NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET

The value of information in portfolio problems with dependent projects

By

Debarun Bhattacharjya, Jo Eidsvik and Tapan Mukerji

PREPRINT
STATISTICS NO. 12/2012

NTNU

This report has URL http://www.math.ntnu.no/preprint/statistics/2012/S11-2012.pdf
Jo Eidsvik has homepage: http://www.math.ntnu.no/~joeid
E-mail: joeid@stat.ntnu.no
Address: Department of Mathematical Sciences, NTNU, N-7491 Trondheim, Norway.

The value of information in portfolio problems with dependent projects

Debarun Bhattacharjya, Business Analytics and Math Sciences, IBM T.J. Watson Research Center, Yorktown Heights, NY, USA

Jo Eidsvik, Department of Mathematical Sciences, NTNU, Trondheim, Norway

Tapan Mukerji, Department of Energy Resources Engineering, Stanford University, CA, USA

Abstract

In the portfolio problem, the decision maker selects a subset out of a set of candidate projects, each yielding an uncertain profit. When the projects in the portfolio are probabilistically dependent, further information regarding any particular project also provides information about other projects, and therefore there is an opportunity to improve value through prudential information gathering. In this paper, we study the value of information in portfolio problems with multivariate Gaussian projects, analyzing the effect of parameters such as the expected values and standard deviations of profits from each project, the accuracy of the information and dependence among projects. We are particularly interested in the role that dependence plays, illustrating the results using examples from the Earth sciences where there is spatial dependence among physical locations. We also present a real-world case study in oil exploration, based on data from the Glitne reservoir in Norway, where we deploy our analytical results to help the decision maker address important acquisition issues pertaining to seismic and electromagnetic information for the reservoir under consideration.

Keywords: value of information, portfolio problem, dependent projects, probabilistic dependence, spatial decision making, portfolio decision analysis

1. Introduction

In the *portfolio problem*, the decision maker is faced with the challenge of selecting a subset from a set of N projects, each yielding an uncertain profit. When the projects in the portfolio are probabilistically dependent, further information regarding any of the projects also provides information about other projects, and therefore there is an opportunity to improve value through prudential and selective information gathering. In this paper, we study portfolio problems where the projects are modeled using a multivariate Gaussian distribution, present several insights from analytical results regarding the effect of project dependence on information value, and discuss potential implications for decision makers.

Ignoring dependence in practical portfolio problems may lead to erroneous results (Killen and Kjaer 2012). Consider the following examples of dependent portfolios: A venture capital fund evaluating start-ups from the same incubator; a salesperson prioritizing deals from a pool of potential deals for the same customer; a bio-technology company choosing among different R&D projects for molecules in the same disease category; etc. Our work was primarily motivated by applications where dependence among projects is of a *spatial* nature – this is a natural way to model portfolio problems in the Earth sciences, such as selecting wildlife bio-conservation sites, choosing locations for mining ores, deciding where to drill oil wells, etc. The last example mentioned is an extension of a classic problem in decision analysis: the oil wildcatter problem (Raiffa 1968). In this paper, we use oil exploration related examples to illustrate the concepts. We also present a detailed case study based on data from the Glitne oil field in Norway to highlight how our analytical results could be used to support information gathering decisions for challenging real-world problems.

In Section 2 we briefly review the relevant literature on the decision-theoretic concept of value of information (henceforth referred to as VOI). The statistical model is formulated in Section 3, where we introduce two motivating examples. In Section 4, we study the VOI for a risk-neutral decision maker's dependent portfolio problem. We make a distinction between situations where information is available for all projects versus where information is available only for a strict subset. When projects are dependent, exploring partial information opportunities can be particularly prudential. A case study on seismic and electro-magnetic testing for reservoir exploration is analyzed in Section 5, and finally, we conclude the paper in Section 6.

2. Literature Review

Although there are many approaches to evaluating the benefits of further information, the decision-theoretic notion of VOI is arguably one of the most powerful approaches because it assigns a monetary value to the effect of refining the probabilities in the decision analysis. The literature is founded upon classic work such as Schlaiffer (1959), Howard (1966) and Raiffa (1968), and continues to gain popularity. However, the hunt for analytical closed-form solutions has often led to more negative results than positive ones (Hilton 1981), and as a result, many researchers have studied canonical problems, i.e. specific classes of decision problems, to understand the general effect of the parameters on information value.

VOI and Gaussian Models

Due to the favorable mathematical properties of the Gaussian distribution, decision problems with Gaussian models have served as some of the most popular canonical problems; see for instance Schlaiffer (1959), Clemen and Winkler (1985), Keisler (2004a), Bickel (2008), Bhattacharjya and Deleris (2012). Among these, the two-action linear loss (TALL) problem, also known as the go/no-go problem, has received particular attention in the literature.

VOI and Portfolio Models

The portfolio problem is a classic operations research problem (Markowitz 1952), and although most of the literature focuses on the case of independent projects, dependence has also been of interest (Weingartner 1966). The existing work related to VOI for portfolio and related problems assumes independence and focuses on separating the effects pertaining to prioritizing projects as opposed to potentially excluding some (Keisler 2004b, Zan and Bickel 2012). Frazier and Powell (2010) consider the Bayesian ranking and selection problem; they also assume independent projects. Our version is a variation on the typical portfolio formulation: we choose to focus on dependent projects where there are no constraints on selection. Although there has been a significant amount of work on evaluating and optimizing portfolios with dependent projects, particularly with deterministic interdependencies (see for instance Santhanam and Kyparisis 1996, Dickinson et al. 2001, Verma and Sinha 2002, Blau et al. 2004, Eilat et al. 2006), the closest work incorporating dependence for information valuation is that of Clemen and Winkler (1985), which studies the value of dependent information sources for a single project. We extend the previous literature on information value for Gaussian models as well as portfolio models.

VOI in the Earth Sciences and Spatial Models

Much of the applied literature on VOI in dependent portfolio problems has occurred in the Earth sciences. The literature has recently generated a significant amount of interest in wildlife conservation (Polasky and Solow 2001, Williams et al. 2011), mining (Phillips et al. 2009), hydrology (Trainor-Guitton et al. 2011), fishing (Hansen and Jones 2008) and forestry (Kangas 2010), to name a few areas. Perhaps the most prolific work using spatial models has been conducted for oil exploration, see e.g. Bickel et al. (2008), Cunningham and Begg (2008), Bratvold et al. (2009) and Martinelli et al. (2011). Eidsvik et al. (2008) and Bhattacharjya et al. (2010) integrate models of dependence from spatial statistics with decision theory; the former considers a logistic model whereas the latter describes a Markov random field approach to valuing information in spatial decision problems. Both these articles solve the computational problems using numerical methods and Monte Carlo simulation. In this paper, we extend previous work by presenting analytical results for the dependent portfolio problem.

3. Model Formulation and Two Motivating Examples

We organize this section by first presenting the notation and model for the basic portfolio problem, motivating dependence among projects with two examples. Then we introduce the notation and likelihood model for information about the projects.

3.1 The Basic Portfolio Problem and Examples

In the portfolio problem, the decision maker is presented with N projects, whose uncertain profits are denoted $\boldsymbol{x} = (x_1, \ldots, x_N)^T$. We assume $\boldsymbol{x}$ is multivariate Gaussian, therefore its probability density function (pdf) is $p(\boldsymbol{x}) = N(\boldsymbol{\mu}, \Sigma)$, where the mean vector is $\boldsymbol{\mu} = [\mu_1, ..., \mu_N]^T$ and the positive definite covariance matrix $\Sigma = \begin{pmatrix} \sigma_1{}^2 & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{N1} & \cdots & \sigma_N{}^2 \end{pmatrix}$. If the projects have a common prior mean, $\boldsymbol{\mu} = \mu_0 \mathbf{1}_N$, where $\mathbf{1}_N$ is a vector of ones. The diagonal entries of the covariance matrix Σ contain the marginal variances of the projects, while the off-diagonal entries are covariance terms. The correlation between project i and j is $\rho_{ij} = \sigma_{ij} / (\sigma_i \sigma_j)$. We assume that the problem is *unconstrained*, i.e. the decision maker can choose as many projects as is profitable. When the projects are independent, the covariance matrix Σ is diagonal. In addition, if the projects have a common variance σ_0^2, the covariance matrix $\Sigma = \sigma_0^2 I_N$. The case of independent projects is an important one in portfolio problems. It is a convenient model to apply in practice because it is not always easy to assess dependence among projects, and most of the literature deals with the independent portfolio problem due to theoretical and computational tractability.

The following two examples are from the domain of oil exploration, illustrating dependence in portfolio problems. Using differing forms of covariance matrices, they are presented in order of coarser to finer spatial granularity (see Figure 1).

A) Equicorrelated projects

Consider an oil company bidding among a set of oil fields in the same petroleum system basin, where all fields are believed to originate from the same geological mechanism; see Martinelli et al. (2011) for a related example. The company chooses to use a simple dependence model – they assume projects have identical pairwise correlation $\rho_{ij} = \rho_0$. This formulation only requires one additional parameter as compared to the independent case, and is therefore popular for modeling projects that share common attributes. The only restriction on ρ_0 is that the resulting covariance matrix should be positive definite, implying that $-1/(N-1) < \rho_0 < 1$. If variances are identical, covariance matrix $\Sigma = (1-\rho_0)\sigma_0^2 I_N + \rho_0 \sigma_0^2 \mathbf{1}_N \mathbf{1}_N^T$. Figure 1 (left) illustrates this pictorially for five projects. The nodes represent potential petroleum prospects, tens of kilometers apart, whose profits are dependent via a common geological mechanism.

B) Spatially dependent projects

Now let us "zoom in" further into an oil company's spatial decision problem at the field level. Suppose the company must decide where to drill oil wells in an oil field. Here, every oil well is a project in the portfolio and it is natural to account for spatial dependence. A common technique for modeling covariance in spatial applications is to let correlations between projects decay as a function of the Euclidean distance between them. There are many functional relations for valid correlation decay, see e.g. Le and Zidek (2006). We use an exponential function $\rho_{ij} = \exp(-\eta d_{ij})$, where d_{ij} is the distance between two projects and $\eta > 0$ determines the *decay of correlation* as a function of distance. Distance $d_{ij} = 3/\eta$ is the *effective range*, resulting in correlation $\exp(-3) \approx 0.05$. Note that this formulation also requires only one additional parameter as compared to the independent case, if project locations are known. However, projects that are spatially further apart are less correlated; the dependence formulation is therefore more nuanced.

If the projects have identical variance σ_0^2, the covariance matrix $\Sigma = \sigma_0^2 \exp(-\eta D)$, where D is the $N \times N$ matrix of Euclidean distances between project locations (the exponent operates element-wise on the matrix). This model is useful for describing the multivariate Gaussian features of spatial projects, such as properties in an oil field that may vary according to the heterogeneous rock composition, e.g. Eidsvik et al. (2008). Figure 1 (right) shows the prior mean values for profits in the case study studied in Section 5, where several reservoir units are defined on a regular lattice covering an area of about 2.5 x 2.5 square kilometers.

3.2 The Portfolio Problem with Information

Our focus in this paper is to analyze how much the decision maker should pay for more information about some or all of the projects, and we are particularly interested in how dependence among projects affects valuation of information sources.

Suppose there is an opportunity to purchase further information about the projects' profits. Examples of information sources include seismic tests for oil exploration, market survey results for a new product launch, etc. Let y_i denote information about the ith project. In many applications, it is natural to assume that the information about a project is *conditionally independent* of attributes of other projects, given the profit from that particular project, i.e. $P\left(y_i \mid \boldsymbol{x}, y_1, y_2 ... y_{i-1}, y_{i+1} ... y_N\right) = P\left(y_i \mid x_i\right) \forall i$. For instance, in Earth sciences applications, it may sometimes be appropriate to assume that the information about a particular physical location is conditionally independent of other locations' properties, given the properties of the current location. We assume information about any particular project has an additive independent Gaussian error term, and along with the conditional independence assumption, this implies that $y_i = x_i + \varepsilon_i$, $\varepsilon_i \sim N\left(0, \tau_i^2\right) \forall i$. This can be written in vector form: $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{E}$, $\boldsymbol{E} = (\varepsilon_1, \ldots, \varepsilon_N)^T$. The likelihood for $\boldsymbol{y} = (y_1, \ldots, y_N)^T$ conditioned on $\boldsymbol{x}$ can be written as $p\left(\boldsymbol{y} \middle| \boldsymbol{x}\right) = N\left(\boldsymbol{x}, \mathrm{T}\right)$, where T is a diagonal matrix with diagonal terms τ_i^2. If the error terms have identical variance t_0^2,

$\mathrm{T} = \tau_0^2 I_N$. Regardless of the Gaussian noise term structure, the marginal distribution of the data is $p(\boldsymbol{y}) = N(\boldsymbol{\mu}, \Sigma + \mathrm{T})$. Please see Appendix A for some properties of the Gaussian distribution.

When information y_i is gathered for all projects, we refer to this case as *total information*. Alternatively, information may be collected for only a strict subset of the projects, in which case we refer to *partial information*. When there is no measurement noise, i.e. $\tau_i = 0 \ \forall i$ or equivalently $y_i = x_i \ \forall i$, there is *perfect information*.

An important generalization of the conditionally independent likelihood model occurs when the information about a project is some linear combination of the profits from all projects, i.e. $\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{E}$, $\boldsymbol{E} \sim N(\boldsymbol{0}, T)$, for general matrix A. We use this model for our case study in Section 5, and all our analytical results extend to this model. Partial information can be handled using this framework by appropriately designing matrix A: the matrix should contain a 1 entry in each row at columns for projects selected for information gathering, and all other entries should be 0. This case is also equivalent to setting some diagonal elements of T to very large values. For ease of exposition, we use $A = I$ to demonstrate the analytical results in the subsequent sections, where we analyze the value of different information gathering schemes. Our results have implications for decision makers who wish to acquire and utilize information in such problems.

4. The Value of Information in Dependent Portfolio Problems

In this section, we study the properties of the VOI, where the decision maker can observe information about K projects. We organize this section based on whether all projects ($K = N$)

or whether only a strict subset ($K < N$) are observed. When there is partial information, dependence between projects is expected to have a strong impact on information value, because information about a subset of the projects would also provide some information about others. We use illustrative examples to highlight issues of interest to managers of dependent portfolios. To compute the VOI, we compare the value of the decision situation without information to the value of the situation if the information were available for free. Without information, a risk neutral decision maker would select those projects that are a-priori profitable. If we denote the set of a-priori profitable projects as P , then prior value $PV = \sum_i \max(0, \mu_i) = \sum_{i \in P} \mu_i$.

4.1 Total Value of Information

Here we study the case where there is total information, i.e. the decision maker has information about all projects. All theorem proofs are summarized in Appendix B.

Theorem 1 (Total Value of Information): For a risk-neutral decision maker, the total value of information (TVOI), i.e. information regarding all projects is:

$$\text{TVOI} = \sum_{i \in \mathrm{P}} \left[-\mu_i \Phi\left(-\mu_i / s_i \right) + s_i \phi\left(\mu_i / s_i \right) \right] + \sum_{i \notin \mathrm{P}} \left[\mu_i \Phi\left(\mu_i / s_i \right) + s_i \phi\left(\mu_i / s_i \right) \right],$$

where s_i^2 is the i$^{\text{th}}$ diagonal element of $S = \Sigma(\Sigma + T)^{-1}\Sigma$, and Φ and ϕ are the cumulative distribution function (cdf) and pdf of the standard Normal distribution respectively. □

When projects are independent, prior covariance matrix Σ is diagonal and $S = \Sigma(\Sigma + T)^{-1}\Sigma$ simplifies to a diagonal matrix with $s_i = \frac{\sigma_i^2}{\sqrt{\sigma_i^2 + \tau_i^2}}$. Moreover, when there is perfect information, there is no noise in the information source $\tau_i = 0$, therefore $s_i = \sigma_i$. This happens regardless of dependence as $T = 0$ results in $S = \Sigma(\Sigma)^{-1}\Sigma = \Sigma$. Since information is perfect and total, dependence no longer matters; only the marginal means and standard deviations matter.

When information is imperfect, dependence among projects has a "coupling effect". We can interpret s_i in Theorem 1 as the *equivalent standard deviation* for projects in this information gathering scheme. The VOI of such a scheme is equivalent to the VOI from a hypothetical equivalent portfolio where the prior uncertainties in the projects of this portfolio are reduced, thereby implying that the value is reduced. The effect of the information source and dependence is to modify the standard deviation term s_i from σ_i to a function of the matrix $S = \Sigma(\Sigma + T)^{-1}\Sigma$ incorporating both prior covariance matrix Σ as well as information noise via T. This may be an intuitive way for managers to envision the interplay between information noise and dependence, particularly as their intuition around independent projects may be more refined. Also, note that the result from Theorem 1 on the total VOI applied to independent projects is a direct extension of a result from Bickel (2008), because the independent unconstrained portfolio problem is a direct extension of the two-action linear loss (TALL) problem for N projects.

4.2 Partial Value of Information (PVOI)

Suppose the decision maker has information about a strict subset of the projects $K \subset N$, with cardinality K. The set of profits for these projects and their information are denoted $\boldsymbol{x_K}$ and $\boldsymbol{y_K}$. The complementary set of projects are denoted $\boldsymbol{x_L}$ and $\boldsymbol{y_L}$. Σ_K is the K by K covariance matrix for projects $\boldsymbol{x_K}$, Σ_{iK} is the 1 by K vector of the covariances between the i^th project and projects in set K, and T_K is the K by K covariance matrix in the likelihood model for observed projects.

Theorem 2 (Partial Value of Information): For a risk-neutral decision maker, the partial value of information, i.e. information regarding projects in set K is:

$$\mathrm{PVOI}(\boldsymbol{y}_K) = \sum_{i \in \mathrm{P}} \left[-\mu_i \Phi\left(-\mu_i / s_i \right) + s_i \phi\left(\mu_i / s_i \right) \right] + \sum_{i \notin \mathrm{P}} \left[\mu_i \Phi\left(\mu_i / s_i \right) + s_i \phi\left(\mu_i / s_i \right) \right]$$

where $s_i^2 = \Sigma_{iK} \left(\Sigma_K + T_K \right)^{-1} \Sigma_{Ki}$. □

We observe similarities between Theorems 1 and 2 – the only difference lies in the equivalent standard deviation computation. There is a coupling effect provided by the information from the subset of surveyed projects, and this effect is less than if total information were available. When there is perfect information, $T_K = 0$. Moreover, if the i^th project has been surveyed, then $s_i = \sqrt{\Sigma_{iK} \left(\Sigma_K \right)^{-1} \Sigma_{Ki}} = \sqrt{\boldsymbol{e_i}^T \Sigma_K \boldsymbol{e_i}} = \sigma_i$, where $\boldsymbol{e_i}$ is a unit vector selecting the i^th project in the survey set. If the i^th project has not been surveyed, then $s_i < \sigma_i$. Using the notion of an equivalent standard deviation as we had done previously for Theorem 1, the situation of partial and imperfect information can be considered equivalent to a hypothetical situation where there is total perfect information but where the prior uncertainties on projects in the portfolio are lower. The effect of the information noise is to effectively decrease the prior uncertainty in the equivalent portfolio and the effect of dependence is to effectively increase it.

4.3 Parameter Sensitivity and Illustrative Examples

The following theorem indicates how the basic model parameters affect the total and partial VOI:

Theorem 3 (Sensitivity of VOI to Basic Parameters):

TVOI (Theorem 1) and PVOI (Theorem 2) vary with the basic model parameters in the following fashion:

(i) μ_i: TVOI and PVOI are maximum when $\mu_i = 0$, and they decrease as μ_i increases or decreases from 0.

(ii) τ_i: TVOI and PVOI increase as τ_i decreases.

(iii) σ_i: TVOI and PVOI increase as σ_i increases. □

Theorem 3(i) specifies that the TVOI is highest when the decision maker is indifferent between selecting and not selecting a project, i.e. when $\mu_i = 0$. This is consistent with previous results such as Fatti et al. (1987) and Delquié (2008). Information is less valuable as the mean increases or decreases from 0 because the decision becomes easier to make. Also, in Theorem 3(ii) and 3(iii), we verify the intuitive result that the TVOI increases with the accuracy of the information as well as with more uncertainty around profits. The theorem proof (please see Appendix B) highlights closed-form expressions for the partial derivatives, i.e. the rate at which the VOI changes as the parameters are varied.

Note that there are no differences in the main effects of the parameters on the PVOI or TVOI. In the examples below and the proof in Appendix B, we discuss the influence of model parameters on TVOI and PVOI in more detail. In particular, the effect of partial information is studied through the equivalent standard deviation s_i, which is always smaller for partial information as compared to total information. The derivative of the equivalent standard deviation is also different for partial versus total information gathering.

It is difficult to analytically gauge the effect of parameters for general models of correlation on the VOI. Based on our experiences through numerical experiments, a “more dependent” structure in S tends to result in higher VOI.

We illustrate the results using the following examples of dependent portfolio problems.

Example A: Equicorrelated projects

Consider a portfolio of $N = 100$ projects, each with an identical marginal distribution and where each pair of projects has the same correlation coefficient. As base case values, we assume each project has mean $m_0 = 0$ and marginal standard deviation $s_0 = 2$. Also, the information accuracy is assumed to be the same for all projects that have been surveyed and equal to 0.8; information accuracy is defined as the correlation between the information and a project’s profits, i.e. $\frac{s_0}{\sqrt{s_0^2 + t_0^2}} = 0.8$, or equivalently $t_0 = 1.5$. The dependence structure is provided by the covariance matrix Σ, which has identical off-diagonal parameters that describe the degree of pairwise correlation. We vary this pairwise correlation coefficient and plot information value as a function of the coefficient.

Suppose that all projects are surveyed. Figure 2 presents sensitivity analyses for the total VOI with respect to the basic parameters. We consider three values of mean, standard deviation and information accuracy in the three parts of the figure, and observe how the TVOI is affected by dependence. We observe that the VOI is maximum at $m_0 = 0$ and that the standard deviation has a strong effect on information value. We also see that as projects become more dependent, the VOI with poorer information accuracy is comparable to that for information with higher accuracy due to the information provided by dependence. This has practical ramifications for decision makers whenever it is possible to purchase lower quality information at a considerably cheaper price that can be as valuable as costlier higher quality information. When information accuracy equals 1, we have perfect information, and the TVOI shown in Figure 2 (right) is the same regardless of the correlation coefficient.

In Figure 3, we compare TVOI with three partial information gathering schemes, in which every 5^{th}, 10^{th} and 20^{th} project is surveyed, respectively. The VOI is computed for the base case values and plotted as a function of correlation coefficient. Information accuracy is assumed to be the same for all projects that have been surveyed, fixed at 0.8. When there is total information, although more dependence makes information more valuable as there are multiple sources of (imperfect) information for every project, the effect is weak. Dependence has a particularly strong effect when there is partial information. Even when only 5 out of 100 projects are sampled for information, the PVOI is comparable to TVOI at a correlation coefficient of around 0.8.

In Figure 4, we study the effect of dependence on the equivalent standard deviation. Here we assume that 20 projects are surveyed (every 5^{th} project out of 100). Due to the identical marginal distribution of projects and identical pairwise correlation between them, there are only two unique values of the equivalent standard deviation – one for projects that are selected (denoted by triangles in the figure) and one for those that are not (denoted by circles). Projects that are selected naturally have a higher equivalent standard deviation, but the effect of dependence is more crucial for projects that are not selected. There is a lot of information value to be gained for these projects, and this grows significantly as there is more dependence. The figure highlights how partial experimentation can be prudential as a practical information acquisition scheme. Note that when projects are independent, a project which is not selected has an equivalent standard deviation of 0 because there is nothing that can be learnt about it from other projects.

When projects have identical standard deviation and correlation coefficient, there is a closed-form analytical solution for the equivalent standard deviation s_i. Using the results in Appendix B, we can write: $s_i^2 = \boldsymbol{e}_i^T \left(\Sigma - T + T\left(\Sigma + T\right)^{-1} T \right) \boldsymbol{e}_i$. Here, $\Sigma + T = \left(\sigma_0^2 \left(1 - \rho_0\right) + \tau_0^2 \right) I_N + \sigma_0^2 \rho_0 1_N 1_N^T$. For a matrix of the form $\alpha I_N + \beta 1_N 1_N^T$, the inverse $\left(\alpha I_N + \beta 1_N 1_N^T \right)^{-1}$ is $c I_N + d 1_N 1_N^T$, where $c = 1/\alpha$ and $d = -\frac{\beta}{\alpha^2 + \alpha\beta N}$. By inserting we get $s_i^2 = \sigma_0^2 - \tau_0^2 + \tau_0^4 \frac{a}{\left(\sigma_0^2 + \tau_0^2 \right) a - \sigma_0^4 \left(N - 1 \right) \rho_0^2}$, where $a = \sigma_0^2 + \tau_0^2 + \sigma_0^2 \left(N - 2 \right) \rho_0$. The last term in the denominator decreases with the absolute pairwise correlation $\left| \rho_0 \right|$. Thus, higher correlation increases s_i and results in higher VOI.

Example B: Spatially dependent projects

Now we study the VOI for a spatial model. Here the $N = 100$ projects are located in a two dimensional domain, identified by randomly selecting the north and east coordinates from a unit square. We assume that projects have positive spatial correlation, modeled using exponential correlation. Again, we assume prior mean 0, standard deviation 2 and information accuracy 0.8. We vary the effective range, which is a measure of spatial dependence in the domain, and observe the effect on the TVOI and PVOI.

Figure 5 depicts the spatial locations of the 100 sites using dots, and also indicates the different partial information schemes, analogous to the previous example. We compare two different types of schemes, shown on the left and right respectively. On the left, the partial tests are chosen as follows: every 5^{th} site that is surveyed is denoted using a circle, every 10^{th} site also includes a plus sign, and when every 20^{th} site is surveyed, a cross sign is also added. The resulting partial tests therefore ensure that surveyed projects are randomly distributed on the grid. On the right side, partial tests are conducted at the grid corners: the first test involves 5 sites on each corner, therefore 20 sites in total (circles), the second involves only the north-east and south-west corners (circles + plus sign), and the third test includes 5 from only the north-east corner (circles + plus sign + cross sign). Note that the two schemes have everything in common except for the location of the surveyed projects. We compare these two schemes to highlight the effect of partial information scheme designs.

Figure 6 plots the TVOI as well as the PVOI for the three partial information tests for the two spatial acquisition designs shown in Figure 5. Both pictures show a trend similar to Figure 3; with the VOI increasing with increasing spatial correlation quantified by the increasing effective range ($3/\eta$). The information schemes on the right, with surveys at the corners, are clearly dominated by the more well spread out designs on the left. When accessibility to sites is an issue, well spread out surveys might be costlier (or even impossible) to perform, while spatially biased surveys (say along roads or around the periphery) might be cheaper and feasible. Designing good spatial experiments is crucial for such problems, and although there is plenty of literature in this area, VOI based techniques are nowhere near as common as entropy or variance reduction techniques, see e.g. Le and Zidek (2006).

The effect of spatial dependence in this model is very similar to the pairwise correlation case, suggesting that the results from these special cases might be applicable for more complex dependence models.

5. An Application in Oil Exploration: Valuing Seismic and Electromagnetic Information

The acquisition and processing of informative data is crucial in oil exploration, due to the significant uncertainty and potential profits/losses in the business. Data can be of various types: advanced geological modeling, electromagnetic measurements, seismic data, observations in wells, and others. We consider seismic data (SD) and electromagnetic data (EM) in this case study, and evaluate them based on the analytical results previously described.

5.1. Case Study Description

The case is based on the Glitne reservoir in the North Sea, also studied in Avseth et al. (2005) and Eidsvik et al. (2008). Production from an offshore field like Glitne is usually done from seabed installations using many deviated wells to drain different reservoir units. We model the reservoir as a two dimensional spatial model of the top reservoir zone and with prior information based on expert understanding of the geology and seismic interpretation. We represent the reservoir units on a lattice of size $N = 25 \times 25 = 625$ covering about 6.25 square kilometers, modeling profits $\boldsymbol{x}$ as a Gaussian prior model with varying mean and variance levels and an exponential spatial covariance function (similar to Section 4, Example B). The means for profits are shown in Figure 1 (right). The current geological and seismic interpretation provides relevant information about the reservoir porosity and thickness, but the profits are also a function of the uncertain oil saturation variable. The decision maker has to decide whether to purchase carefully processed SD or EM, or both. SD and EM can both provide (imperfect) information about saturation (and in turn, the profits) after calibration and interpretation.

We assume conditional independence for SD, EM and between SD and EM. Let $\boldsymbol{y_1} = \left(y_{1,1}, \ldots, y_{1,M_1}\right)$denote the SD, where $M_1 = N = 625$, since SD is processed over the entire domain of interest (total information). Accuracy level τ_{SD}^2 is assumed identical for all reservoir units. Let $\boldsymbol{y_2} = \left(y_{2,1}, \ldots, y_{2,M_2}\right)$ denote the EM. We assume $M_2 = 25$, since EM is usually acquired only along one-dimensional sailing lines (partial information). We have selected the center North-South column for EM collection. Accuracy level τ_{EM}^2 is assumed to be constant. The Gaussian likelihood for data $\boldsymbol{y} = \left(\boldsymbol{y}_1, \boldsymbol{y_2}\right)^T$ is denoted $p\left(\boldsymbol{y} \mid \boldsymbol{x}\right) = N\left(A\boldsymbol{x}, T\right)$, with diagonal A and T matrices. Please see Appendix C for further modeling details.

5.2 Value of Information Analysis

We analyze the VOI of SD and EM to recommend a course of action for information acquisition. We study the VOI as a function of the model parameters, using the analytical results from previous sections. Note that the posterior value at projects only depends on the posterior mean

$\boldsymbol{z} = E(\boldsymbol{x} \mid \boldsymbol{y})$ and its distribution, which is Gaussian with mean vector $\boldsymbol{\mu}$ and covariance matrix $S = \Sigma A^T \left(A \Sigma A^T + T \right)^{-1} A \Sigma$ in this case.

The base case situation for parameters is specified from geological inputs and by likelihood maximization based on prior information regarding the Glitne reservoir. We estimate the accuracy of the EM based on geophysical principles. We assign cost of drilling C = 20 million USD. The sensitivity studies cover parameter values in a specified range (from low to high) around the base case levels. (Please see Appendix C).

In Figure 7, we plot VOI for SD and EM as a function of the cost of drilling, assumed identical for all units in the grid. The VOI is highest for intermediate costs of drilling. All curves show a peak at around C = 15 million USD, around which a combination of SD or EM information is most valuable. For very low drilling costs, there is not much added value in SD or EM since drilling is lucrative enough without any further data. For high drilling costs, the reservoir units are too costly to develop and added information is unlikely to change this decision. Note that the VOI of EM is smaller than the VOI of SD. The total information attained by SD seems more valuable here, but in general the difference depends on the noise levels for EM versus SD.

In Figure 8 (left), we plot the VOI as a function of the prior standard deviation. Recall that the prior standard deviations for profits vary between the reservoir units. Here the first axis represents a constant scaling parameter for the covariate-dependent standard deviation. Figure 8 (right) shows the VOI as a function of the effective spatial range parameter (1000m is about 10 cells). Note the difference in second axes numbering for the two displays. This reflects the VOI changes when varying the first axes within the low-high range around the reference case. Both displays in Figure 8 show three curves: SD+EM, SD and EM. The total test with SD appears very valuable here. The VOI is clearly increasing when the prior standard deviation varies from low to high. The dependence in the reservoir basin also has a strong effect on the VOI.

Figure 9 displays the VOI as a function of the measurement accuracy of SD and EM. We treat the situations separately: First, VOI is evaluated for SD only (left), as a function of the SD accuracy. Second, VOI is evaluated for EM only (right), as a function of the EM accuracy. The first axes represent high to low accuracy of measurements, with different scales on the second axes. Both displays show that VOI decreases as the information is less reliable. The VOI is smaller for less dependent profits (dashed lines).

To determine the optimal acquisition scheme, the decision maker must of course also consider the price of the experiments. We construct optimal acquisition decision regions as a function of the prices of SD and EM by comparing the four static decision options: i) Purchase both EM and SD; ii) Purchase SD only; iii) Purchase EM only; and iv) Purchase neither. Let P_{SD} and P_{EM} be the prices of SD and EM. The decision about information gathering is made according to:

$$\operatorname{argmax}\left(\mathrm{VOI}_{SD+EM}-\mathrm{P}_{SD}-\mathrm{P}_{EM},\mathrm{VOI}_{SD}-\mathrm{P}_{SD},\mathrm{VOI}_{EM}-\mathrm{P}_{EM},0\right).$$

The decision boundaries are computed by equating the values for the different types of data. Figure 10 shows four diagrams with decision regions corresponding to two different prior noise levels and two different effective correlation range values. The high and low levels of parameters are set at (0.9 and 1.1) relative to the base-case specification. Note that the decision regions are very sensitive to the parameter values. For high effective range and high prior uncertainty, the alternative of purchasing neither data type is optimal only for very high price ranges of data.

6. Conclusions

Much of the prior literature on information valuation for portfolio selection assumes independent projects. Here we have derived analytical expressions for the value of information in portfolio problems where the projects are probabilistically dependent, in which case further information regarding any particular project also provides information about others, and therefore there is an opportunity to improve value through prudential information gathering. We modeled project dependencies as a multivariate Gaussian distribution, allowing us to obtain closed-form analytical results. Alternate approaches that do not assume the Gaussian property would in general require Monte Carlo simulations and may therefore be computationally demanding for large problems (see for instance Keisler 2004b, Eidsvik et al. 2008, Bhattacharjya et al. 2010).

Our results can help a risk-neutral decision maker study how much they should pay for more information about some (PVOI) or all (TVOI) of the projects, and how dependence among projects affects valuation of information sources. When information is imperfect, dependence among projects has a coupling effect, and in general the information is more valuable as a result. We studied VOI as a function of the prior uncertainty and noise level in the information, deriving new results for the sensitivity of information value to model parameters.

In this paper, we have made some restrictive assumptions, most notably regarding the lack of constraints for project selection and the risk-neutrality of the decision maker. There are analytical challenges towards incorporating constraints (such as Keisler 2004b) or risk aversion (such as Bickel 2008) for deriving closed-form results for information value in dependent portfolio problems, because the project selection methodology can no longer treat the inclusion/exclusion of each project separately. Bhattacharjya et al. (2010) study the effect of constraints on project selection in spatial decision problems through simulations, noting that the effects appear to be non-monotonic in general; we feel there may be potential for further research along these directions.

In the examples studied here, dependence among projects was primarily of a spatial nature – this is a common aspect of problems in the Earth sciences. We also studied a case study based on data from the Glitne oil field in Norway to highlight how our analytical results could be used to support information gathering decisions for challenging real-world problems. This case study compares the VOI for seismic data versus electromagnetic data, and the spatial dependence of the reservoir property is modeled as an exponential spatial covariance function. Our analytical results can of course also be applied to other industries involving portfolio problems with interdependent projects and uncertain profits.

References

Anderson, T. W., 2003, *An Introduction to Multivariate Statistical Analysis*, Third edition, John Wiley and Sons.

Avseth, P., Mukerji, T., Mavko, G., 2005, *Quantitative Seismic Interpretation*, Cambridge University Press.

Bhattacharjya, D., Deleris, L., 2012, The value of information in some stopping problems, Working paper.

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

Bickel, J. E., 2008, The relationship between perfect and imperfect information in a two-action risk-sensitive problem, *Decision Analysis*, **5** (3), pp. 116 – 128.

Bickel, J. E., Gibson, R. L., McVay, D. A., Pickering, S., Waggoner, J., 2008, Quantifying 3D land seismic reliability and value, *SPE Reservoir Evaluation and Engineering*, **11** (5), pp. 832 – 841.

Blau G. E., Pekny, J. F., Varma, V. A., Bunch, R. R., 2004, Managing a portfolio of interdependent new product candidates in the pharmaceutical industry, *Journal of Product Innovation Management*, **21**, pp. 227 – 245.

Bratvold, R., 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), pp. 630 – 638.

Clemen, R., Winkler, R., 1985, Limits for the precision and value of information from dependent sources, *Operations Research*, **33** (2), pp. 427 – 442.

Cunningham P., Begg S., 2008, Using the value of information to determine optimal well order in a sequential drilling program, *AAPG Bulletin*, **92**, pp. 1393 – 1402.

Delquié, P., 2008, The value of information and intensity of preference, *Decision Analysis*, **5** (3), pp. 129 – 139.

Dickinson, M. W., Thornton, A. C., Graves, S., 2001, Technology portfolio management: Optimizing interdependent projects over multiple time periods, *IEEE Trans. Engineering Management*, **48** (4), pp. 518 – 527.

Eidsvik, J., Bhattacharjya, D., Mukerji, T., 2008, Value of information of seismic amplitude and CSEM resistivity, *Geophysics*, July – August, **73** (4), pp. R59 – R69.

Eilat H., Golany, B., Shtub, A., 2006, Constructing and evaluating balanced portfolios of R&D projects with interactions: A DEA based methodology, *European Journal of Operational Research*, **172**, pp. 1018 – 1039.

Fatti, L. P., Mehrez, A., Pachter, M., 1987, Bounds and properties of the expected value of sample information for a project-selection problem, *Naval Research Logistics*, **34** (1), pp. 141 – 150.

Frazier, P. I., Powell, W. B., 2010, Paradoxes in learning and the marginal value of information, *Decision Analysis*, **7** (4), pp. 378 – 403.

Hansen, G., Jones, M., 2008, The value of information in fishery management, *Fisheries*, **33** (7), pp. 340 – 348.

Hilton, R., 1981, The determinants of information value: Synthesizing some general results, *Management Science*, **27** (1), pp. 57 – 64.

Howard, R., 1966, Information value theory, *IEEE Transactions on Systems Science and Cybernetics*, **2** (1), pp. 22 – 26.

Kangas, A. S., 2010, Value of forest information, *European Journal of Forest Research*, **129**, pp. 863 – 874.

Keisler, J., 2004a, Technical note: Comparative static analysis of information value in a canonical decision problem, *Engineering Economist*, **49**, pp. 339 – 349.

Keisler, J., 2004b, Value of information in portfolio decision analysis, *Decision Analysis*, **1** (3), pp. 177 – 189.

Killen, C. P., Kjaer, C., 2012, Understanding project interdependencies: The role of visual representation, culture and process, *International Journal of Project Management*, **30**, pp. 554 – 566.

Kitanidis, P. K., Lane, R. W., 1985, Maximum likelihood parameter estimation of hydrological spatial processes by the Gauss-Newton method, *Journal of Hydrology*, pp. 53 – 71.

Le, N. D., Zidek J. V., 2006, *Statistical Analysis of Environmental Space-Time Processes*, Springer.

Markowitz, H., 1952, Portfolio selection, *Journal of Finance*, **7**, pp. 77 – 91.

Martinelli, G., Eidsvik, J., Hauge, R., Forland, M.D., 2011, Bayesian networks for prospect analysis in the North Sea, *AAPG Bulletin*, pp. 1423 – 1442.

Phillips, J., Newman, A. M., Walls, M. R., 2009, Utilizing a value of information framework to improve ore collection and classification procedures, *The Engineering Economist*, **54**, pp. 50 – 74.

Polasky, A., Solow, A., 2001, The value of information in reserve site selection, *Biodiversity and Conservation*, **10**, pp. 1051 – 1058.

Pratt, J., 1964, Risk aversion in the small and in the large, *Econometrica*, **32**, pp. 122 – 136.

Raiffa, H., 1968, *Decision Analysis: Introductory Lectures on Choices under Uncertainty*, Addison-Wesley.

Santhanam, R., Kyparisis, G. J., 1996, A decision model for interdependent information system project selection, *European Journal of Operational Research*, **89**, pp. 380 – 399.

Schlaiffer, R., 1959, *Probability and Statistics for Business Decisions*, McGraw-Hill Book Company, Inc., New York.

Trainor-Guitton, W. J., Caers, J. K., Mukerji, T., 2011, A methodology for establishing a data reliability measure for value of spatial information problems, *Mathematical Geosciences*, **8**, pp. 929 – 949.

Verma D., Sinha, K. K., 2002, Toward a theory of project interdependencies in high tech R&D environments, *Journal of Operations Management*, **20**, pp. 451 – 468.

Weingartner, H. M., 1966, Capital budgeting of interrelated projects: Survey and synthesis, *Management Science*, **12** (7), pp. 485 – 516.

Williams, B. K., Eaton, M. J., Breininger, D. R., 2011, Adaptive resource management and the value of information, *Ecological Modelling*, **222**, pp. 3429 – 3436.

Zan, K., Bickel, J. E., 2012, Components of portfolio value of information, Working paper.

Appendix A: Some properties of multivariate Gaussian distribution

The multivariate Gaussian probability density function for random vector $\boldsymbol{x} = (x_1, \ldots, x_N)^T$ is:

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{N/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right).$$

The marginal distribution of x_i is Gaussian with mean μ_i and variance equal to the i^{th} diagonal term of Σ, denoted σ_i^2.

Proofs for the following results can be found in Anderson (2003):

Result 1: If we split the random variable into two sets $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x_L} \\ \boldsymbol{x_K} \end{pmatrix}$, with mean $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_L \\ \boldsymbol{\mu}_K \end{pmatrix}$, and covariance structure $\Sigma = \begin{pmatrix} \Sigma_L & \Sigma_{LK} \\ \Sigma_{KL} & \Sigma_K \end{pmatrix}$, we have a Gaussian conditional distribution for $\boldsymbol{x_L}$ given $\boldsymbol{x_K}$. The mean and covariance matrix are: $\begin{aligned} \boldsymbol{\mu}_{L|K} &= \boldsymbol{\mu}_L + \Sigma_{LK}\Sigma_K^{-1}\left(\boldsymbol{x_K} - \boldsymbol{\mu}_K\right) \\ \Sigma_{L|K} &= \Sigma_L - \Sigma_{LK}\Sigma_K^{-1}\Sigma_{KL}. \end{aligned}$.

Result 2: For a constant matrix A and Gaussian vector $\boldsymbol{x}$, the linear transformation $\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{E}$, $\boldsymbol{E} \sim N\left(\boldsymbol{0}, T\right)$, is also Gaussian and the marginal is given by $p\left(\boldsymbol{y}\right) = N\left(\mathrm{A}\boldsymbol{\mu}, A\Sigma A^T + T\right)$.

Result 3: The conditional distribution of $\boldsymbol{x}$ given $\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{E}$, where $\boldsymbol{E} \sim N\left(\boldsymbol{0}, T\right)$ is Gaussian with mean and variance: $\begin{aligned} \boldsymbol{\mu}_{x|y} &= \boldsymbol{\mu} + \Sigma\mathrm{A}^T\left(A\Sigma\mathrm{A}^T + T\right)^{-1}\left(\boldsymbol{y} - A\boldsymbol{\mu}\right) \\ \Sigma_{x|y} &= \Sigma - \Sigma\mathrm{A}^T\left(A\Sigma\mathrm{A}^T + T\right)^{-1} A\Sigma. \end{aligned}$

Lemma: A closed form expression for expected maximization for the Gaussian distribution

For a Gaussian variable z with mean m and variance s^2, $E\left(\max\left(z, 0\right)\right) = m\Phi(m/s) + s\phi\left(m/s\right)$

Proof: The Gaussian pdf is $p(z) = \frac{1}{\sqrt{2\pi s^2}}\exp\left[-\frac{(z-m)^2}{2s^2}\right]$. The standard normal occurs for $m = 0$ and $s = 1$, and is symmetrical for the pdf and cdf. Since $\frac{d\exp(-\frac{x^2}{2})}{dx} = -x\exp(-\frac{x^2}{2})$, we have $\int x\exp\left(-\frac{x^2}{2}\right)dx = -\exp(-\frac{x^2}{2}) + Const$. Therefore:

$$
\begin{aligned}
&\int_{-\infty}^{\infty} \max(z,0)\,p(z)dz = \int_{0}^{\infty} z\,p(z)dz = \int_{-m/s}^{\infty} (m+xs)\frac{1}{\sqrt{2\pi s^2}}\exp(-\frac{x^2}{2})sdx \\
&= m\int_{-m/s}^{\infty} \frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2})dx + \frac{s}{\sqrt{2\pi}}\int_{-m/s}^{\infty} x\exp\left(-\frac{x^2}{2}\right)dx \\
&= m[1-\Phi(-m/s)] + \frac{s}{\sqrt{2\pi}}\left[-\exp\left(-\frac{x^2}{2}\right)\right]_{-m/s}^{\infty} = m\Phi(m/s) + s\phi(-m/s) \\
&= m\Phi(m/s) + s\phi(m/s)
\end{aligned}
$$

A change of variable $x = \frac{z-m}{s}$, with $dz = sdx$, is used in the integral. □

Appendix B: Theorem Proofs

Proof of Theorem 1:

From the Appendix A results, the conditional mean is:

$$E[\boldsymbol{x} \mid \boldsymbol{y}] = \boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}} = \boldsymbol{\mu} + \Sigma(\Sigma+T)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) = \boldsymbol{z}.$$

The value with this information (for free) is:

$$\sum_{i=1}^{N}\int_{y} \max\left(E\left(x_i \mid \boldsymbol{y}\right),0\right)p\left(\boldsymbol{y}\right)d\boldsymbol{y} = \sum_{i=1}^{N}\int_{y} \max\left(z_i,0\right)p\left(z_i\right)dz_i,$$

since the component z_i is the only relevant part of the data or information $\boldsymbol{y}$ in the integral. We have $p(\boldsymbol{z}) = N(\boldsymbol{\mu}, S)$, where $S = \Sigma(\Sigma+T)^{-1}\Sigma$, with diagonal entries s_i^2. We use the Lemma from Appendix A to compute the value with information. We need to subtract the prior value from the value with free information, therefore:

$$
\begin{aligned}
TVOI &= \sum_{i=1}^{N}\left[\mu_i\Phi\left(\mu_i/s_i\right) + s_i\phi\left(\mu_i/s_i\right)\right] - \sum_{i\in \mathrm{P}}\mu_i \\
&= \sum_{i\in \mathrm{P}}\left[\mu_i\left(\Phi\left(\mu_i/s_i\right)-1\right) + s_i\phi\left(\mu_i/s_i\right)\right] + \sum_{i\notin \mathrm{P}}\left[\mu_i\Phi\left(\mu_i/s_i\right) + s_i\phi\left(\mu_i/s_i\right)\right]
\end{aligned}
$$

Since $1-\Phi\left({\mu_i}/{\sigma_i}\right)=\Phi\left(-{\mu_i}/{\sigma_i}\right)$, we get the required result. □

Proof of Theorem 2:

From the Appendix A results, the conditional mean is:

$$E\left[x_i \mid \boldsymbol{y_K}\right]=\mu_i+\Sigma_{iK}\left(\Sigma_K+T_K\right)^{-1}\left(\boldsymbol{y_K}-\boldsymbol{\mu}_K\right)=z_i .$$

This conditional mean has distribution $p(z_i)=N(\mu_i, s_i^2)$, where $s_i=\sqrt{\Sigma_{iK}\left(\Sigma_K+T_K\right)^{-1}\Sigma_{Ki}}$. The value with information (for free) is:

$$\sum_{i=1}^{N}\int_{\boldsymbol{y_K}}\max\left(\mu_i+\Sigma_{iK}\left(\Sigma_K+T_K\right)^{-1}\left(\boldsymbol{y_K}-\boldsymbol{\mu}_K\right),0\right)p(\boldsymbol{y_K})d\boldsymbol{y_K}=\sum_{i=1}^{N}\int_{z_i}\max\left(z_i,0\right)p(z_i)dz_i .$$

The remainder of the proof is similar to that of Theorem 1, and we get the required result. □

Preliminaries to Proof of Theorem 3:

I) For any parameter θ in the mean vector or covariance matrix, the derivative of the expression for $E\left(\max\left(z,0\right)\right)$ in the Lemma from Appendix A is:

$$\frac{dm\Phi\left({m}/{s}\right)+s\phi\left({m}/{s}\right)}{d\theta}=\frac{dm}{d\theta}\Phi\left({m}/{s}\right)+m\frac{d\Phi\left({m}/{s}\right)}{dm/s}\frac{dm/s}{d\theta}+\frac{ds}{d\theta}\phi\left({m}/{s}\right)+s\frac{d\phi\left({m}/{s}\right)}{dm/s}\frac{dm/s}{d\theta}$$

Since $\dfrac{d\Phi\left(x\right)}{dx}=\phi\left(x\right)$ and $\dfrac{d\phi\left(x\right)}{dx}=\dfrac{d}{dx}\dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{x^2}{2}\right)=-x\dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{x^2}{2}\right)=-x\phi\left(x\right)$,

the derivative simplifies to:

$$\frac{dm\Phi\left({m}/{s}\right)+s\phi\left({m}/{s}\right)}{d\theta}=\frac{dm}{d\theta}\Phi\left({m}/{s}\right)+m\phi\left({m}/{s}\right)\frac{dm/s}{d\theta}+\frac{ds}{d\theta}\phi\left({m}/{s}\right)-s\left(\frac{m}{s}\right)\phi\left({m}/{s}\right)\frac{dm/s}{d\theta}$$
$$=\frac{dm}{d\theta}\Phi\left({m}/{s}\right)+\frac{ds}{d\theta}\phi\left({m}/{s}\right)$$

.

- Sensitivity to the mean: If $\theta = m$, $\dfrac{dE\left(\max\left(z,0\right)\right)}{dm} = \Phi\left(m/s\right)$.

- Sensitivity to the effective standard dev.: If $\theta = s$, $\dfrac{dE\left(\max\left(z,0\right)\right)}{ds} = \phi\left(m/s\right)$.

II) Using the result from Theorem 1 and previous result I), the derivative of the information value with respect to a covariance matrix parameter θ becomes $\sum_{i=1}^{N} \frac{1}{2s_i} \frac{ds_i^{\,2}}{d\theta} \phi\left(\mu_i / s_i\right)$. We will therefore focus on the derivative of the equivalent variance with respect to q, $\frac{ds_i^{\,2}}{d\theta}$.

III) The derivative of a diagonal element of the matrix S is: $\frac{ds_i^2}{d\theta} = \frac{d\boldsymbol{e}_i^T S \boldsymbol{e}_i}{d\theta} = \boldsymbol{e}_i^T \frac{dS}{d\theta} \boldsymbol{e}_i$, where $\boldsymbol{e}_i$ is a unit vector with 1 in the i^th^ entry and 0s elsewhere.

IV) The derivative of a matrix inverse is $\frac{d\Sigma^{-1}}{d\theta} = -\Sigma^{-1} \frac{d\Sigma}{d\theta} \Sigma^{-1}$, where the matrix derivative works for every entry of the matrix.

V) The following matrix identity holds (by Sherman-Woodbury-Morrison):

$\left(\Sigma + T\right)^{-1} = \Sigma^{-1} - \Sigma^{-1}\left(\Sigma^{-1} + T^{-1}\right)^{-1}\Sigma^{-1}$, and $\left(\Sigma^{-1} + T^{-1}\right)^{-1} = T - T\left(\Sigma + T\right)^{-1} T$.

Therefore, $S = \Sigma\left(\Sigma + T\right)^{-1}\Sigma = \Sigma - T + T\left(\Sigma + T\right)^{-1} T$. Or equivalently,

$$\Sigma - \Sigma\left(\Sigma + T\right)^{-1}\Sigma = T - T\left(\Sigma + T\right)^{-1} T.$$

From Appendix A, we see that the matrix of equivalent variances and covariances, $S = \Sigma\left(\Sigma + T\right)^{-1}\Sigma$, can be interpreted as the reduction in variance from prior Σ to the posterior $\Sigma - \Sigma\left(\Sigma + T\right)^{-1}\Sigma$.

Proof of Theorem 3:

(i) Sensitivity to μ_i:

From I), the derivative with respect to the mean of any project is:

$$\frac{dTVOI}{d\mu_i}=\frac{d\sum_{i=1}^{N}\left[\mu_i\Phi\left(\mu_i/s_i\right)+s_i\phi\left(\mu_i/s_i\right)\right]-\sum_{i\in\mathrm{P}}\mu_i}{d\mu_i}=\left\{\begin{matrix}\Phi\left(\mu_i/s_i\right)-1, & i\in\mathrm{P}, & \mu_i>0\\ \Phi\left(\mu_i/s_i\right) & i\notin\mathrm{P} & \mu_i\le 0\end{matrix}\right\}.$$

Considering PVOI, the only difference for this derivative is the equivalent standard deviation s_i. When we have partial information, s_i is smaller, and the absolute derivatives are smaller.

(ii) Sensitivity to τ_i:

Matrices Σ and $\Sigma+T$ are both symmetric. $(\Sigma+T)^{-1}$ is also symmetric since matrix inversion retains the symmetric property. Based on III) and IV) above, for some parameter θ in T,

$$\frac{d\left(\boldsymbol{e}_i^T\Sigma(\Sigma+T)^{-1}\Sigma\boldsymbol{e}_i\right)}{d\theta}=-\boldsymbol{e}_i^T\Sigma(\Sigma+T)^{-1}\frac{dT}{d\theta}(\Sigma+T)^{-1}\Sigma\boldsymbol{e}_i.$$

If $q=t_i^2$, $\frac{dT}{d\theta}$ is a matrix of 0s with i^th diagonal element equal to 1. Note that this is right and left multiplied with a symmetric expression $\boldsymbol{b}=(\Sigma+T)^{-1}T\boldsymbol{e}_i$, therefore:

$$\frac{d\left(\boldsymbol{e}_i^T\Sigma(\Sigma+T)^{-1}\Sigma\boldsymbol{e}_i\right)}{d\tau_i^2}=-\boldsymbol{b}^T\frac{dT}{d\tau_i^2}\boldsymbol{b}=-b_i^2<0.$$

In summary, using II) above, we get negative effects and the derivative is:

$$\frac{dTVOI}{d\tau_i^2}=-\sum_i\left[\frac{1}{2s_i}\phi\left(\mu_i/s_i\right)b_i^2\right].$$

A similar form holds for PVOI, in which case the eqv. standard deviation $s_i = \sqrt{\Sigma_{iK}\left(\Sigma_K + T_K\right)^{-1}\Sigma_{Ki}}$.

We can take derivatives with respect to the measurement noise parameters in the survey set.

(iii) Sensitivity to σ_i :

The derivative of s_i^2 is again the crucial part. Using results I)-V) above:

$$\frac{d\boldsymbol{e}_i^T\left(\Sigma - T + T\left(\Sigma+T\right)^{-1}T\right)\boldsymbol{e}_i}{d\theta} = \boldsymbol{e}_i^T\frac{d\Sigma}{d\theta}\boldsymbol{e}_i - \boldsymbol{e}_i^T T\left(\Sigma+T\right)^{-1}\frac{d\Sigma}{d\theta}\left(\Sigma+T\right)^{-1}T\boldsymbol{e}_i .$$

If $q = s_i^2$, $\frac{d\Sigma}{d\theta}$ is a matrix which is nonzero only for the elements in row i and column i. Let $B = \left(\Sigma+T\right)^{-1}\frac{d\Sigma}{d\sigma_i^2}\left(\Sigma+T\right)^{-1}$, which is non-negative definite. Then, the above expression can also be written as: $\frac{d\boldsymbol{e}_i^T\left(\Sigma - T + T\left(\Sigma+T\right)^{-1}T\right)\boldsymbol{e}_i}{d\sigma_i^2} = \boldsymbol{e}_i^T\left(\Sigma+T\right)B\left(\Sigma+T\right)\boldsymbol{e}_i - \boldsymbol{e}_i^T TBT\boldsymbol{e}_i$.

Since both Σ and T are positive definite, this derivative will be positive, and the TVOI increases as a function of the prior uncertainty.

The case with PVOI can be viewed as a special case of total information where some t_is go to infinity. The same proof therefore holds. Also, using this argument, the inflation caused by $\Sigma + T$ in the derivative above is more comparable with that of T . Thus, for the case of partial information, the VOI tends to increase slower as a function of the prior variance. □

Appendix C: Case Study Details

We describe the modeling assumptions in the case study in more detail.

Prior

Let $\boldsymbol{x}$ be the profits associated with drilling wells and producing oil from N reservoir units. The lateral top-reservoir domain of interest is represented on a lattice of size $N = 25 \times 25 = 625$ units. Each unit is 100m x 100m, and has a reservoir thickness of 20m. The revenues depend on the oil saturation as well as several other physical parameters. The porosity is specified in each unit from geological inputs and preliminary seismic analysis. In addition, the expected price of oil is assumed to be 400 USD per standard cubic meter. The cost of drilling one well is set to a reference value of C=20 million USD, valid for all units. The profits are a priori Gaussian. At reservoir unit i, the expected profit $\mu_i = r_i - C$, where r_i is the revenue. We let $r_i = r_i\left(H_i\right)$ denote the revenue as a function of explanatory variables H_i (porosity, thickness, and prior assumptions about saturation and oil-price). For a-priori level of saturation we use an uncertain depth-dependent oil-water contact.

We let the prior marginal variance terms σ_i^2 vary across different reservoir units, and use a simple parametric form $\sigma_i^2 = \beta H_i$, where explanatory variables H_i incorporate the prior uncertainties in the local saturation levels based on a binomial assumption. The spatial correlation is defined via an exponential correlation function of Euclidean distance.

Likelihood

The individual terms comprising the SD likelihood are $p\left(y_{1,i} \mid x_i\right) = N\left(A_{SD,i,i} x_i, \tau_{SD}^2\right)$, where the expectation term is computed from linearized geophysical relationships between saturation and SD. For EM data at grid unit j, we define likelihood $p\left(y_{2,j} \mid x_{i(j)}\right) = N\left(A_{EM,j,i(j)} x_{i(j)}, \tau_{ED}^2\right)$, where $i(j)$ is the grid cell of EM observation number $j = 1, \ldots, 25$. Similar to SD, this formula is based on linearizing a geophysical model for the EM data. For both SD and EM, the linearization is based on data variables regarded as a function of saturation, and then profits are directly related to saturation, which is the main uncertainty here. For SD, we mainly use Gassmann's fluid substitution formula and Zoeppritz's equation to relate saturation and SD, while the relation between saturation and EM is based on Archie's law, see e.g. Eidsvik et al (2008). Figure C1 shows the SD and EM as a function of profits, where the various saturation levels and the main covariate porosity are indicated. The linearization is carried out at each porosity curve, and at a-priori level of the saturation variable, which depends on the geological considerations. Figure C2 shows the predicted SD and observed SD attribute from the current processing of data at Glitne. The predicted SD is computed from a-priori expected values in the model for the profits. In summary, the likelihood for data $\boldsymbol{y} = (\boldsymbol{y_1}, \boldsymbol{y_2})^T$ is represented by the following joint model:

$$p(\boldsymbol{y} \mid \boldsymbol{x}) = N(A\boldsymbol{x}, T), \qquad A = \begin{bmatrix} A_{SD} \\ A_{EM} \end{bmatrix}, \qquad T = \begin{bmatrix} \tau_{SD}^2 I_{625} & 0 \\ 0 & \tau_{EM}^2 I_{25} \end{bmatrix},$$

where the matrix A_{SD} is diagonal, while A_{EM} picks the reservoir units in the center North-South column in the lattice for the partial EM information.

Posterior and marginal likelihood

Associated with the prior $p(\boldsymbol{x})$ and likelihood model $p(\boldsymbol{y} \mid \boldsymbol{x})$, we have marginal likelihood

$p(\boldsymbol{y}) = N(A\boldsymbol{\mu}, A\Sigma A^T + T)$ and the following posterior model:

$$p(\boldsymbol{x} \mid \boldsymbol{y}) = N\left(\boldsymbol{z}, \Sigma - \Sigma A^T \left(A\Sigma A^T + T\right)^{-1} A\Sigma\right), \qquad \boldsymbol{z} = \boldsymbol{\mu} + \Sigma A^T \left(A\Sigma A^T + T\right)^{-1} (\boldsymbol{y} - A\boldsymbol{\mu}).$$

Based on the currently available SD processing at the Glitne field, the marginal likelihood can be used to estimate most of the statistical model parameters; see e.g. Kitanidis and Lane (1985).

Figures

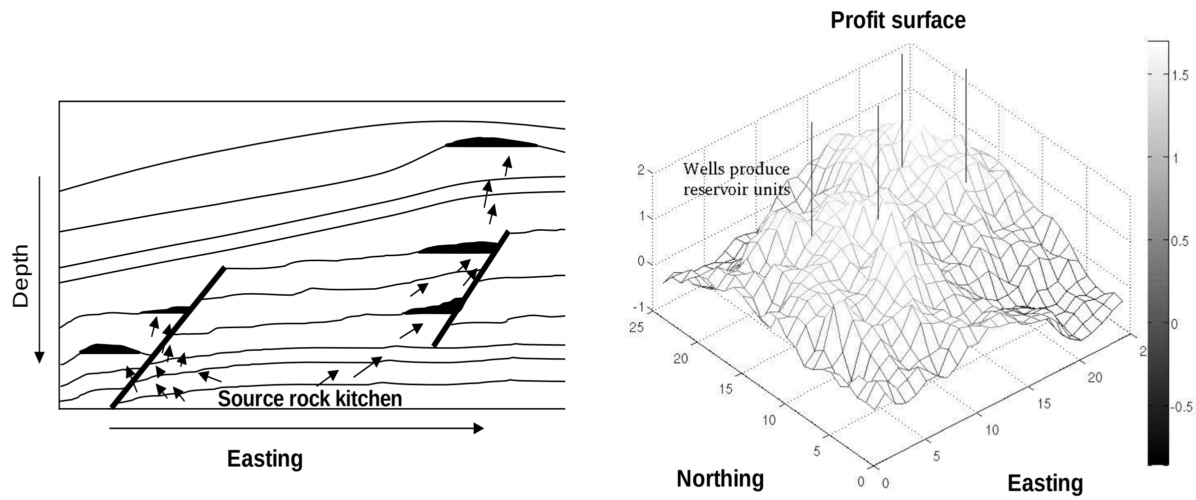

Figure 1: Two motivating examples; Left: Five oil fields in a petroleum system basin are correlated through a common geological mechanism. Right: A reservoir is represented as a 25 by 25 lattice of reservoir units.

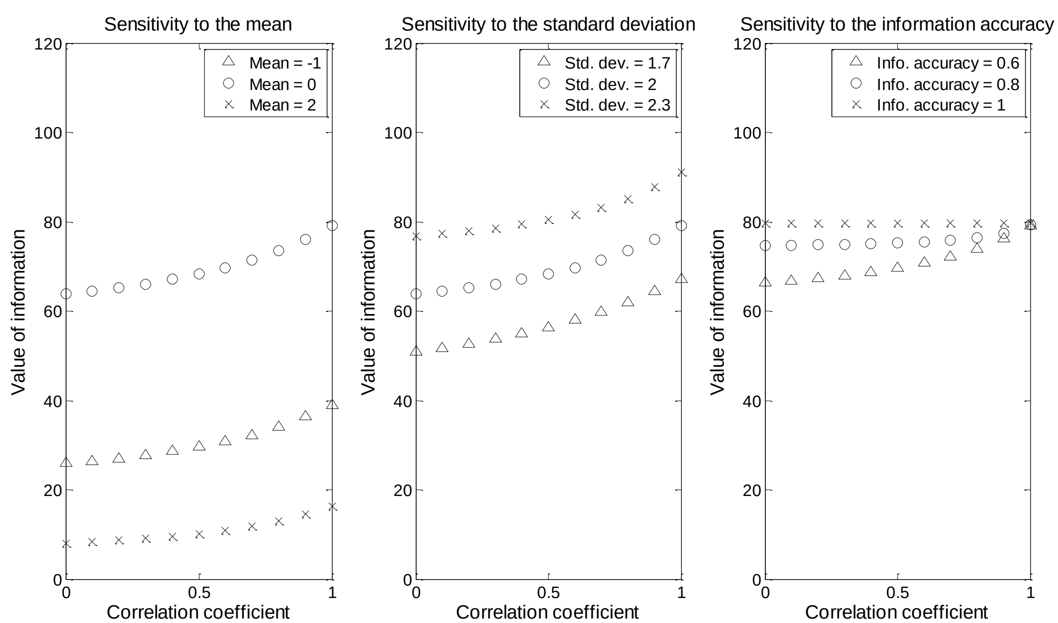

Figure 2: Dependence of TVOI on pairwise project correlation coefficient, and sensitivity to parameters in Example A. Left: Mean. Middle: Standard deviation. Right: Information accuracy.

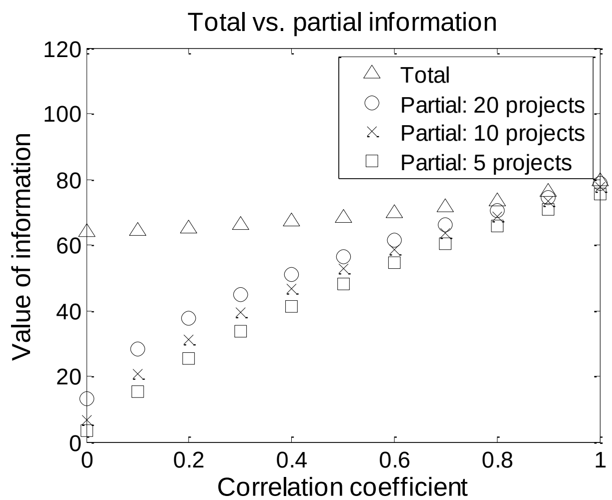

Figure 3: Total vs. partial information in Example A.

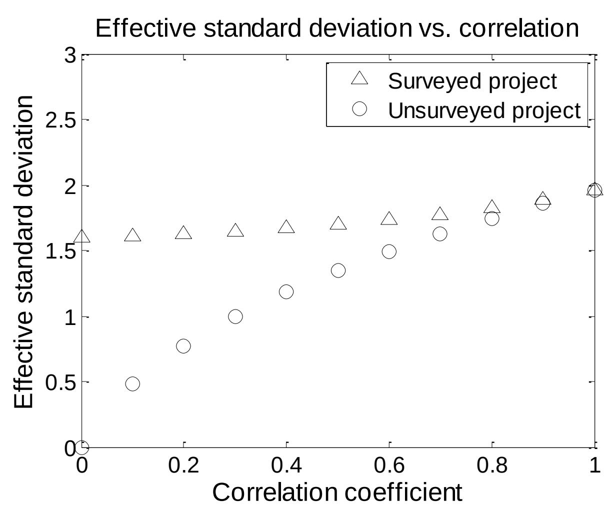

Figure 4: Equivalent standard deviation in Example A.

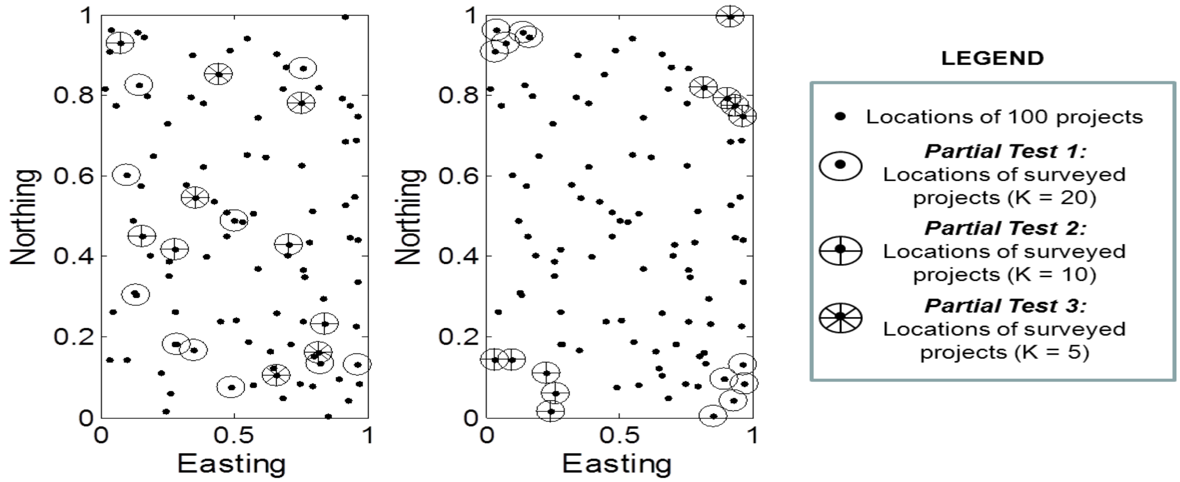

Figure 5: Two types of information gathering schemes for the spatial domain in Example B, with three partial tests of 5, 10 and 20 projects surveyed. Left: Well spread out partial surveys. Right: Partial surveys at corners.

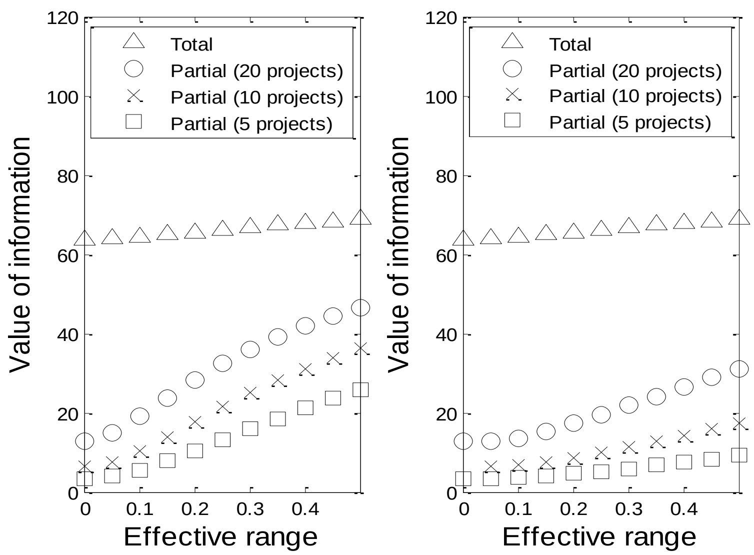

Figure 6: Total vs. partial information in Example B, corresponding to the two types of information schemes in Figure 5. Left: Well spread out partial surveys; Right: Partial surveys at corners.

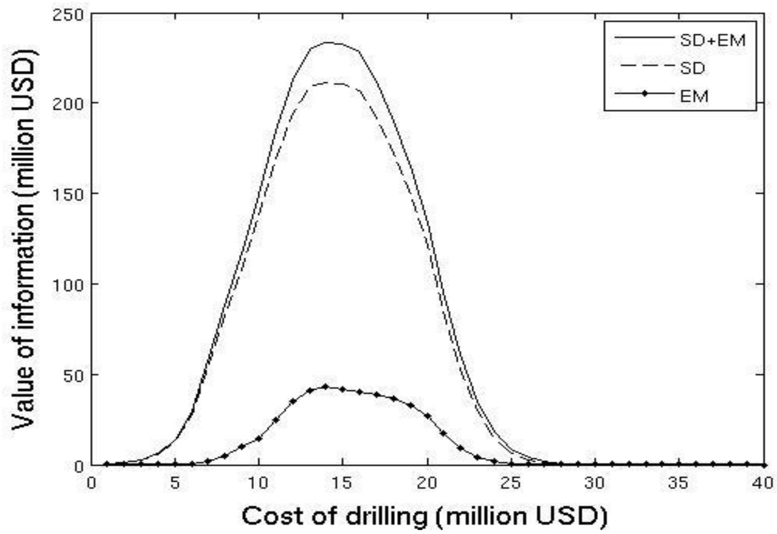

Figure 7: VOI of seismic and electromagnetic data, plotted as a function of drilling cost.

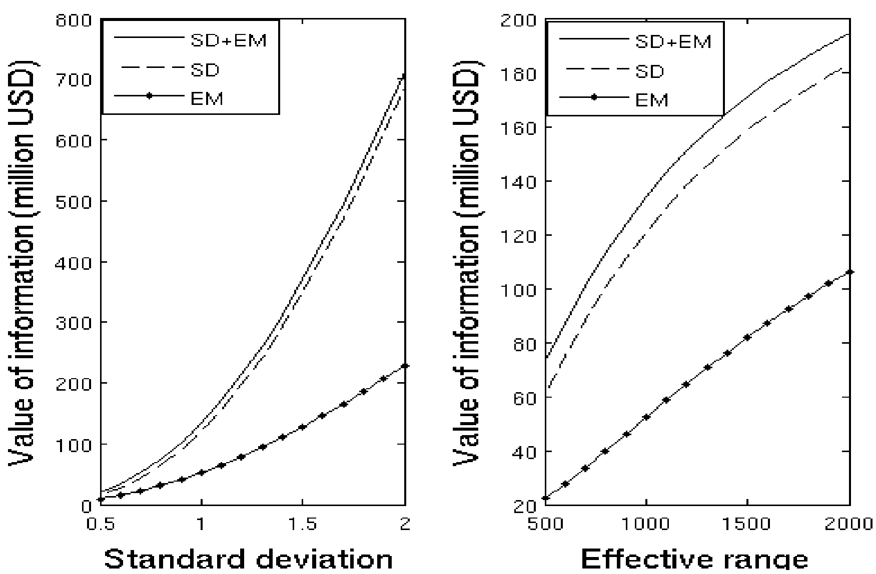

Figure 8: VOI as a function of the standard deviation parameter (left) and as a function of the spatial effective range parameter (right). The first axes are varied from low to high compared with a reference case.

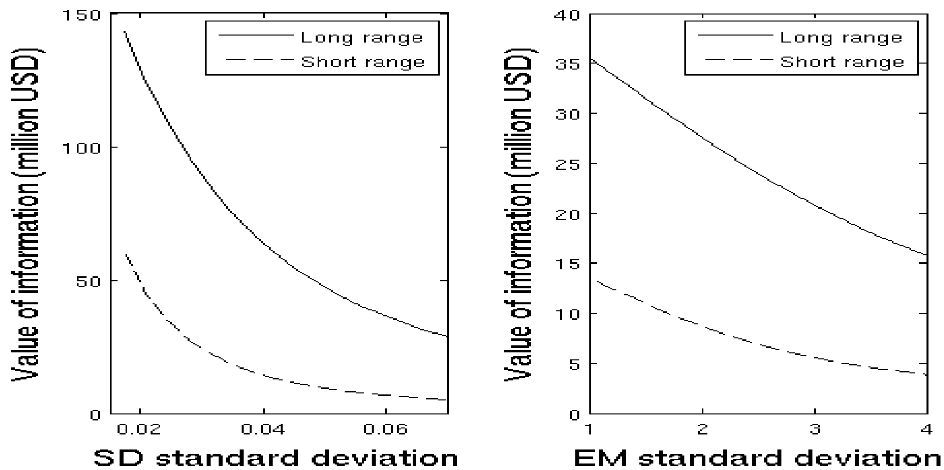

Figure 9: VOI as a function of measurement noise standard deviation parameter for seismic data (left) and as a function of measurement noise standard deviation parameter for electromagnetic data (right). The first axes are varied from low to high compared with a reference case.

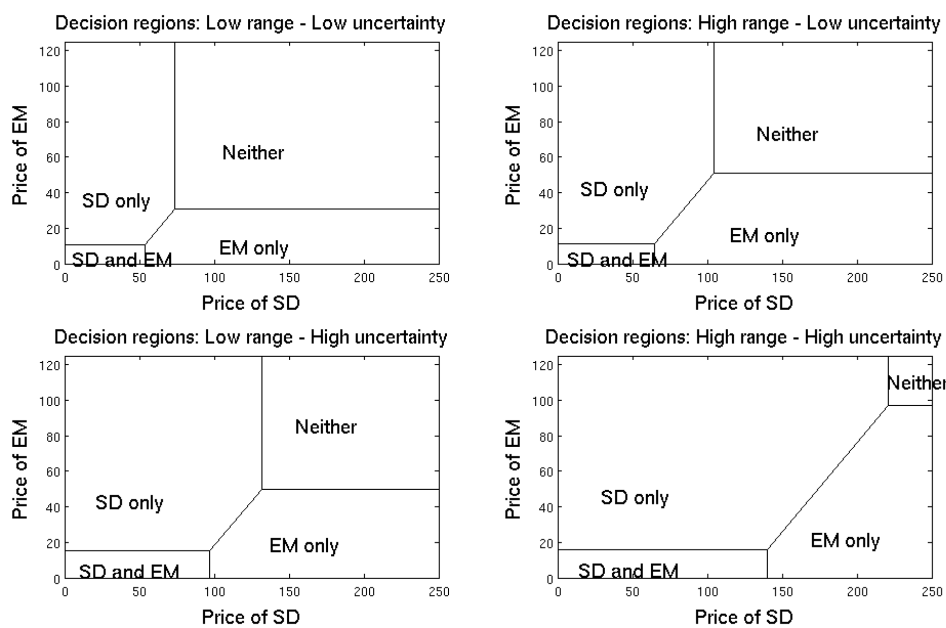

Figure 10: Decision regions for the SD and/or EM information gathering. The first axis represents the price of the SD and the second axis that of EM. The decision maker can purchase either SD and EM, only SD, only EM, or neither.

Figures for Appendix C

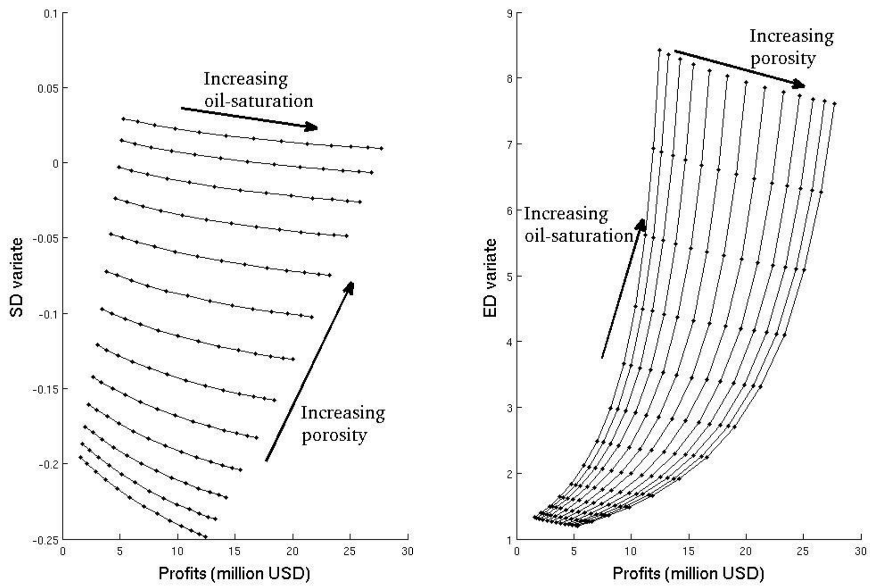

Figure C1: Estimated response values for seismic (left) and electromagnetic data (right) as a function of profits (first axis). The curves represent different levels of porosity.

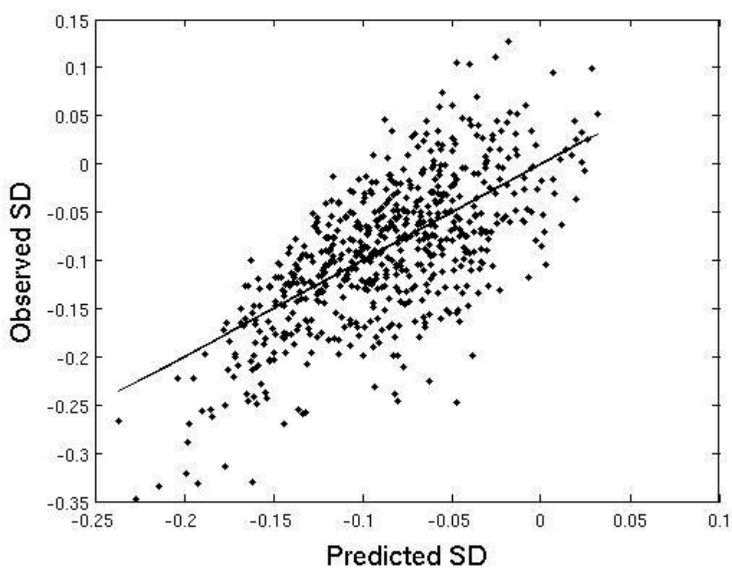

Figure C2: Predicted seismic (first axis) vs. observed seismic data (second axis). The residuals are described by a spatially structured model for the revenues (profits) and independent error terms for the seismic data.