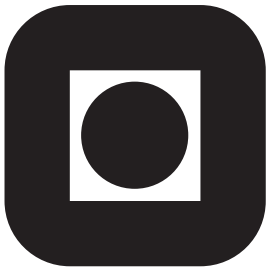
NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

Dynamic Decision Making for Graphical Models Applied to Oil Exploration

by

Gabriele Martinelli, Jo Eidsvik and Ragnar Hauge

PREPRINT
STATISTICS NO. 12/2011



NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL http://www.math.ntnu.no/preprint/statistics/2011/S12-2011.pdf
Gabriele Martinelli has homepage: http://www.math.ntnu.no/∼gabriele
E-mail: gabriele@stat.ntnu.no
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

Dynamic Decision Making for Graphical Models Applied to Oil Exploration

Gabriele Martinelli, Jo Eidsvik and Ragnar Hauge
Department of Mathematical Sciences, NTNU, Norway and Norwegian Computing Center, NR, Norway

We present a framework for sequential decision making in problems described by graphical models. The setting is given by dependent discrete random variables with associated costs or revenues. In our examples, the dependent variables are the potential outcomes (oil, gas or dry) when drilling a petroleum well. The goal is to develop an optimal selection strategy that incorporates a chosen utility function within an approximated dynamic programming scheme. We propose and compare different approximations, from simple heuristics to more complex iterative schemes, and we discuss their computational properties. We apply our strategies to oil exploration over multiple prospects modeled by a directed acyclic graph, and to a reservoir drilling decision problem modeled by a Markov random field. The results show that the suggested strategies clearly improve the simpler intuitive constructions, and this is useful when selecting exploration policies.

Keywords: Bayesian Networks, Dynamic Programming, Graphical models, Heuristics, Petroleum Exploration

1 Introduction

This paper considers the problem of sequential decision making, where the outcome of one decision will influence the others, and the decisions are based on the expected utility. Our motivation and main applications are from oil and gas exploration, where a petroleum company has a set of potential drilling sites, called prospects. For each prospect, we may either drill or not. There is a cost of drilling, but revenues if the well discovers oil or gas. The prospects are statistically dependent, and drilling at one prospect gives information that is used to update the probability of success at other prospects. The goal is to find an optimal drilling sequence, including when to stop drilling and abandon the remaining prospects. Thus, we are interested in designing a strategy or a policy for selecting the sequence of prospects, or at least the first few best prospects in such a sequence.

The optimization of the expected utility function is a trade-off between two factors: the direct gain from the exploitation, and the indirect gain of learning, or exploration, that helps us make informed future decisions. The balance between these is controlled by a discounting factor. With no discounting, the problem becomes a maximization of the value of information (VOI), whereas a high discounting factor leads to a greedy approach where only immediate gain counts.

We have no theoretical restrictions on the underlying statistical model for dependence between outcomes. In practice, there is a requirement that conditional distributions can be computed and updated fast, since many of these will be computed when designing a strategy. For comparing strategies, it is also advantageous if we can easily simulate from the models. In our examples, we use Bayesian networks (BN) and Markov random fields (MRF), which both have these properties.

This sequence selection challenge is a discrete optimization problem and the optimal strategy can be found by Dynamic Programming (DP), see Bellman [1957] and Nemhauser [1966]. However, DP becomes computationally infeasible when the number of possible actions increases. A remedy for this is to apply a heuristic approach. These strategies have been studied in many contexts due to the curse of dimensionality, which affects most DP methods [Powell, 2008]. The simplest heuristic is to run an independent strategy, disregarding the information gain caused by dependent variables. A more sophisticated alternative is to use a myopic strategy. This strategy conditions on past outcomes, but does not account for future scenarios in a proper way.

A possible solution to large DP problems is also offered by approximate DP methods, see Bertsekas and Tsitsiklis [1996] and Powell [2008]. The main idea of approximate DP is to replace the optimization function with a statistical model that captures the impact of decisions now on the future. Approximate DP techniques for solving a multivariate knap-sack problem [Bertsimas and Demir, 2001] resembles the situation of drilling wells, but in our graphical representation of dependent prospects it is not obvious how to find a statistical model that approximates the future value function. Further, our main goal is to find an optimal sequence, and most approximate methods do not give this as a byproduct when approximating the utility function.

When considering a set of independent prospects, the optimal sequential decisions are offered by the Gittins indeces [Gittins, 1979], introduced for solving bandit-problems [Weber, 1992]. These methods were used for a petroleum example by Benkherouf and Bather [1988]. Here, the discovery probabilities in different prospects are apriori independent, and later dependent just through the total number of discoveries. In our context the correlation is much more complex, and the actions influence the model probabilities in a complicated manner.

Branch and bound methods are non-heuristic in the sense that they produce lower and upper bounds of the values [Goel et al., 1979]. In practice the gap between bounds can be wide. Moreover, it is not obvious how to generalize these methods for graphical models with dependence between prospects. In our context we will typically lack monotonicity when computing the best (discounted) sequence. Branch-and-bound methods seem more suited for the actual maximum value of the utility function, instead of an approximate sequential decision strategy.

The challenge of constructing drilling strategies is of course well known in the oil and gas industry, but no one seems to have looked at it from a modern statistical modeling viewpoint applying graphs to couple many dependent prospects. Kokolis et al. [1999] describe a similar problem with a focus towards decision making under uncertainty and the technical risks connected to a project. They do not consider how to design an optimal sequential drilling strategy, but discuss the combinatorial increase of the number of scenarios that has to be considered. Smith and Thompson [2008] analyze the consequences of dependent versus independent prospects, and give drilling guidelines that are optimal in special situations. In Bickel and Smith [2006] and Bickel et al. [2008], DP is used to compute the optimal sequences and profits from six dependent prospects, but they do not

indicate solutions for the large scale challenge.

Our approach is a classical DP procedure with the use of heuristics for approximating the continuation value (CV). The CV is defined as the value of the prospects that have not yet been revealed in the sequential exploration. This value of course depends on the outcome of the current sequence. The simplest form of this is the naive strategy sketched above, where the CV is computed under independence. We use this for benchmarking. In addition, we apply pruning of the decision tree, where we ignore unlikely branches to reduce the combinatorial problem.

We use profit as utility function, which is quite reasonable for a large oil company. Alternatives would be profit given that loss at no time exceeds a given value, or, in the case of entering new exploration areas, minimum loss before concluding that there is no oil present. The profit criterion we use is not dissimilar to the VOI. For instance, Eidsvik et al. [2008], Bhattacharjya et al. [2010] and Martinelli et al. [2011] study the effects of more data acquisition, the ability to make improved decisions, and the associated VOI for spatially dependent variables. However, they do not compute the VOI in a sequential manner [Miller, 1975], neither are they focusing on the best sequential exploration program.

A statistician can of course imagine other non-monetary utility functions within a similar framework. For instance minimum integrated variance, minimum entropy, or other design of experiment criteria, where the goal could be to stabilize the probabilities at nodes in the graph, with the least possible observables. Our approach is in some ways similar to constructing sequential spatial designs.

The paper develops as follows: In Section 2 we introduce the notation, statistical framework, and the assumptions required for applying our methods. In Section 3 we present the DP algorithm for our problem. In Sections 4 and 5 we propose the various heuristic strategies, and the algorithms used to evaluate the properties of the sequential exploration strategies. Finally, in Section 6 we provide results for a small BN model and a BN case study of 25 prospects in the North Sea, and a MRF for a oil reservoir represented on a 5×20 lattice.

2 Assumptions and notation

We consider a set of N prospects with a discrete set of possible outcomes. These N prospect nodes are a subset of the total M nodes in a graph. The remaining $M - N$ auxiliary nodes impose the specified dependency structure in the model, but are not observable. For every node $i = 1, \ldots, M$ we have a discrete random variable $x_i \in \{1, \ldots, k_i\}$. In the examples below we use $k_i = k$, and $k = 3$. The random vector of all variables is $\mathbf{x} = (x_1, ..., x_M)$, where the N first components correspond to the prospect variables.

The directed acyclic graph (DAG) in one of our case studies is built from the causal large scale processes required to make sufficient amounts of oil and gas, see VanWees et al. [2008] and Martinelli et al. [2011]. A DAG defines the joint probability model $p(\mathbf{x})$ from the product of conditional distributions $p(x_i|x_i^{\text{pa}})$, for all nodes $i = 1, \ldots, M$, where x_i^{pa} denotes the set of outcomes at parent nodes of i. In the MRF example for a lattice of cells in a specific reservoir unit, the model is defined over neighborhoods on the lattice, where $p(x_i|\mathbf{x}_{-i}) = p(x_i|x_j; j \in N_i)$, and $\mathbf{x}_{-i}$ is the vector of all variables except x_i, while N_i is the neighborhood of node i. The particular type of model is not critical, but for our purposes fast updating of the conditional probabilties is important. This updating is required when we get sequential evidence. BNs are fast to update using for instance the junction tree algorithm, see e.g. Lauritzen and Spiegelhalter [1988] and Cowell et al. [2007]. Moderate size MRFs can be computed recursively by forward-backward algorithms [Reeves and Pettitt, 2004]. Moreover, we will use Monte Carlo samples to generate realistic future scenarios. It is easy to draw samples $\mathbf{x} = (x_1, \ldots, x_M) \sim p(\mathbf{x})$ from the BNs and MRFs we consider.

Given a probabilistic model with a certain dependence structure, we want to develop a drilling strategy, i.e. a dynamic road map that leads us through the exploration phase of the prospects. Since the prospects are dependent, the outcome of one changes the probability of success in the others. The strategy of continued drilling thus entails a sequential updating of the probability model.

We let ω_i be the observable in node $i = 1, \ldots, N$. If node i is not yet observed, we set $\omega_i = -$. If we choose to observe node i, ω_i is the actual outcome of the random variable x_i at this node. For instance, $\omega_i = 1$ can mean that well i has been drilled and found dry, $\omega_i = 2$ if found gas, and $\omega_i = 3$ if oil. Initially, before acquiring any observables, we have $\boldsymbol{\omega} = (-, \ldots, -)$. If we start to explore nodes, we put the outcomes at the corresponding indices of the vector $\boldsymbol{\omega}$. Say, if node 2 is selected first, and observed in state $\omega_2 = x_2 = 2$, we

set $\boldsymbol{\omega} = (-, 2, -, \ldots, -)$. For the likelihood of this scenario we need the marginal $p(x_2 = 2)$. This is computed by summing out all scenarios that share the second component equal to 2. In order to compute the conditional probabilities of a node i, given evidence, we need $p(x_i = j|\boldsymbol{\omega})$, $j = 1, \ldots, k$, where the empty elements $(-)$ of $\boldsymbol{\omega}$ are unobserved and marginalized out.

The CV associated with the state vector $\boldsymbol{\omega}$ is denoted $v(\boldsymbol{\omega})$. This is the expected value of all currently unobserved states given the observed states, the objective function, and the chosen strategy. One objective is to find the initial value before any sites have been explored, i.e. $v(\boldsymbol{\omega}_0)$ where $\boldsymbol{\omega}_0 = \{-, -, \ldots, -\}$. This initial value is in theory given by DP. As an integral part of the DP algorithm one must evaluate the values $v(\cdot)$ of all possible combinations of evidence. This becomes impossible when we have many nodes in the graph.

The DP algorithm also gives the optimal sequential decisions, but since this is not feasible for large N, we instead construct forward selection strategies, approximating $v(\cdot)$ to different accuracies. When building such strategies we make assumptions about the way decisions are made. First, we assume that the decision maker selects one node at a time. Without this assumption, the problem would grow to allow all orders of two-tuples, three-tuples, etc. Second, we assume that there are fixed revenues and costs associated with each node. If we choose to explore a node, we have to pay a cost. For certain outcomes of the node variable, we receive a revenue. For instance, if the outcome is oil, we get the fixed revenues associated with this outcome. The revenues and costs change from node to node, but introducing random distributions on the costs and revenues for each type of outcome would make our optimization problem harder. Finally, we assume the utility function contains separate parts for every node, without any coupling of the nodes. This utility function expresses the decision makers inclination to collect the revenues or cost at any site. In principle, there could be shared costs or revenues for nodes, say if certain HC prospects have common infrastructure [Martinelli et al., 2011]. We could include this into our framework, but it gives extra computation time, and obscures the presentation of the sequential strategies, that is the focus of our work.

Given these assumptions, we will next show how DP presents a recipe for computing the optimal strategy. We will discuss why this is not possible for a model with many nodes, and we will instead propose strategies to overcome the problem.

3 Dynamic programming

In our context DP recursively explores backwards all the possible sequences that may occur, and it uses these evaluations to select the best dynamic decisions. See e.g. Bickel and Smith [2006] for a similar application of DP.

By the word sequence we mean each of the possible situations that may arise. Sequences are indexed by adding one element $\omega_i \in \{1, \ldots, k\}$ at a time to the evidence vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)$. With $N = 4$ prospects, the state $\boldsymbol{\omega} = \{-, 1, -, 2\}$ means that the node 1 has not yet been explored, node 2 has been observed to be in state 1, node 3 has not yet been explored, and node 4 has been observed to be in state 2. Two different scenarios may correspond to this sequence, one when node 2 is explored before node 4, and another when node 4 is explored before 2. This order is of course relevant when we have only explored node 2, and consider observing node 4, or vice versa, but once both node 2 and 4 have been explored, we no longer distinguish between these two scenarios (except for discounting purposes). Thus, we tend to use the terms sequence and scenario as synonyms.

The decision tree (Figure 1) visualizes the chosen strategy. It works in the following way:

1. First, decide which site, if any, to observe first.
2. Then, depending on the outcome $x_i \in \{1, \ldots, k\}$, which node to observe next, if any, and so on.

DP solves the tree by working backwards:

1. First, decide whether to drill the last prospect, conditional on the first $N - 1$ observables.
2. Then, decide which prospect to drill if there are two nodes left, and so on, to the initial empty set.

In order to pursue this strategy, we have to maximize a certain utility function. We use maximum profit, and $v(\boldsymbol{\omega})$ then represents the expected revenues of future cash flows given that we are in state $\boldsymbol{\omega}$. Initially, the

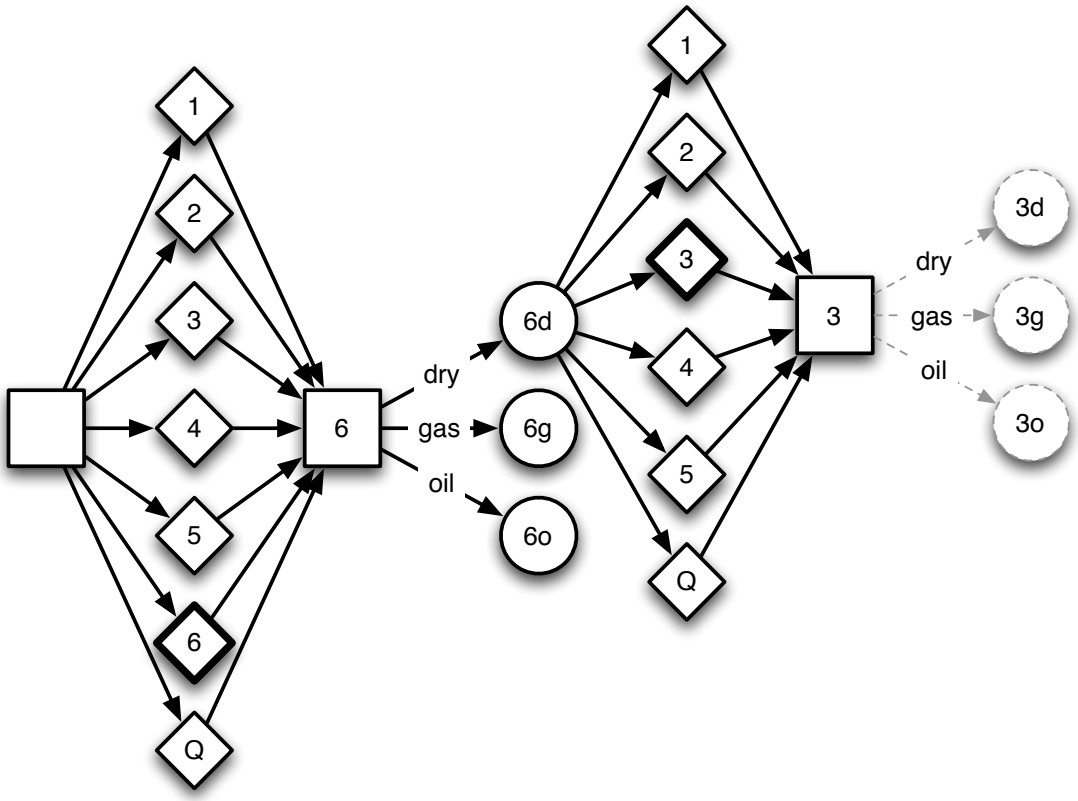


Figure 1: Illustration of a decision tree. At the first branch we can select any of the 6 nodes, or quit (Q). Node 6 is explored first here. If node 6 is dry, we select node 3 at the next branch. The outcome of node 3 can influence which branch to enter next, and so on.

vector of observables is empty: $\boldsymbol{\omega}_0 = \{-, -, \ldots, -\}$. The maximization is among all possible free states:

$$v(\boldsymbol{\omega}) = \max_{i \in N} \left\{ \sum_{j=1}^{k} p(x_i = j) \left[r_i^j + \delta \max_{s \in N-1} \left\{ \sum_{l=1}^{k} p(x_s = l | x_i = j)(r_s^l + \ldots), 0 \right\} \right], 0 \right\}, \tag{1}$$

where the second and the subsequent maximizations are over all nodes not yet considered. Here, δ is a discounting factor that depends on the specific case and on the inclination of the decision maker. The r_i^j are revenues or costs of node i with outcome j. When all the sites have been drilled, the CV is $v(\cdot, \cdot, \ldots, \cdot) = 0$, and we can proceed backwards, one step at a time, to extract the DP solution.

Equation (1) can be rewritten [Bickel and Smith, 2006], and it can be seen as a maximization over all free nodes and 0 (not exploring any further). This means that $v(\boldsymbol{\omega}) = \max_i\{0, v_i(\boldsymbol{\omega})\}$, where:

$$v_i(\boldsymbol{\omega}) = \sum_{j=1}^{k} \left\{ p(x_i = j | \boldsymbol{\omega})(r_i^j + \delta \cdot v(\boldsymbol{\omega}_i^j)) \right\}, \tag{2}$$

where $\boldsymbol{\omega}_i^j = \{\boldsymbol{\omega}_{\text{-i}}, \omega_i = j\}$ and $v(\boldsymbol{\omega}_i^j)$ is the CV of the state $\boldsymbol{\omega}_i^j$, i.e. $v(\boldsymbol{\omega}_i^j) = \max_{l \neq i}\{0, v_l(\boldsymbol{\omega}_i^j)\}$.

The main problem with this optimal DP solution is the exponential growth of the number of scenarios that have to be considered. Bickel and Smith [2006] derives the computational cost for a non-recombining tree, i.e. a tree ignoring the order of the observed nodes. Then,

$$\text{Number of possible scenarios in a non-recombining tree: } : \sum_{i=0}^{N} \left(\begin{array}{c} N \\ i \end{array} \right) k^i (N - i + 1).$$

This entails an order of 10^4 scenarios for six nodes [Bickel and Smith, 2006], and 10^{15} when $N = 25$ nodes. The computational cost (proportional to the number of scenarios) is therefore in the order of $\left(\frac{N}{2}\right)!k^{N/2}$. Bickel and Smith [2006] suggest to save the local results of the computations in order to reduce the number on configurations to consider. Say, for the purposes of the CV, it does not matter whether we first drilled first one well or another, given that we observe their outcomes. Nonetheless, the exact procedure remains unfeasible when the N increases. Furthermore, we need to mention that the introduction of the discounting factor δ makes impossible the use of classical non-recombining algorithms, and gives us few chances other than following the described approach.

4 Heuristic strategies

Because of the rapid growth in scenarios, one must look for approximate solutions. The problem shares some features with that of a chess game. The player has to choose among all the possible moves she can carry out, and at the same time he has to consider all the possible replies of his opponent, and the consequential replies of herself, and so on. What is done in practical chess algorithms is to limit the search to a reasonable amount of moves forward, and to evaluate the best move in that "restricted match", see Shannon [1950] and Feldmann et al. [1994].

Similarly, we push the search through a certain number of steps, figuring out some rules to approximate the remaining value of the scenarios. The idea is to introduce different and simpler rules, in order to approximate the CV in equation (2) without going all the way down through the branches of the decision tree. We will call these rules *heuristics*, following the literature described in Pearl [1984].

4.1 Naive strategy

The naive strategy ignores the dependence among nodes. Therefore, the decision is just based on a priori knowledge. There is no learning. The CV is then estimated as a simple sum of a priori intrinsic values:

$$v_N(\boldsymbol{\omega}) = \sum_{i=1}^{N} \max \left\{ \sum_{j=1}^{k} r_i^j p(x_i = j), 0 \right\}. \tag{3}$$

The best sequence is therefore computed just once, at the beginning of the algorithm, and the nodes are chosen according to:

$$\mathrm{i}_{(1)} = \arg\max_i \left\{ \sum_{j=1}^{k} r_i^k p(x_i = j), 0 \right\}, \quad \mathrm{i}_{(2)} = \arg\max_{i \setminus i_{(1)}} \left\{ \sum_{j=1}^{k} r_i^j p(x_i = j), 0 \right\} \ldots. \tag{4}$$

As we can see, the outcome of the first best prospect is irrelevant when choosing the second best site. This approach, though being very simple (the computational cost is linear in N), still captures a large part of the value if the correlation between nodes is small. The main problem is the individuation of the correct best sequence, since disregarding any correlation effect can lead to focused attention on nodes that might not be appealing given the evidence of the previous steps.

4.2 Myopic approach

A second natural approach is represented by the *myopic* strategy Bollapragada and Morton [1999]. According to this strategy, the best sequence is computed step-by-step in a forward selection scheme. The conditional probabilities in the different nodes are now updated, given the previous outcomes. This represents an intuitive sequential strategy, but it only exploits the dependence in the graph through the past, and does not consider what the future might bring.

The strategy for finding the first best prospect coincides with the naive approach:

$$i_{(1)} = \arg\max_i \left\{ \sum_{j=1}^{k} r_i^j p(x_i = j), 0 \right\}. \tag{5}$$

Given an outcome $x_{(1)}$ at this first selected node $i_{(1)}$, the second myopic best site is then chosen as a function of the observable in the first node:

$$\begin{aligned} \mathrm{i}_{(2j_1)|\mathrm{x}_{i_{(1)}}=j_1} &= \arg\max_{i\setminus i_{(1)}} \left\{ \sum_{j=1}^{k} r_i^j p(x_i = j | x_{i_{(1)}} = j_1), 0 \right\} \\ \mathrm{i}_{(2j_2)|\mathrm{x}_{i_{(1)}}=j_2} &= \arg\max_{i\setminus i_{(1)}} \left\{ \sum_{j=1}^{k} r_i^j p(x_i = j | x_{i_{(1)}} = j_2), 0 \right\}, \dots \end{aligned} \tag{6}$$

Now, the second best choice, therefore, involves k different maximizations, depending on the outcome of $x_{i_{(1)}}$. Thus, using a myopic strategy leads to a decision tree with $\sum_{i=0}^{N} k^i$ scenarios.

The myopic approach approximates the CV in equation (2) by

$$\begin{aligned} v_1 &= \max \left\{ \sum_{j=1}^{k} r_i^j p(x_{i_{(1)}} = j), 0 \right\} \\ v_2 &= \sum_{j=1}^{k} \left(\max \left\{ \sum_{l=1}^{k} r^l_{x_{i_{(2j)}}} p(x_{i_{(2j)}} = l | x_{i_{(1)}} = j), 0 \right\} \right) p(x_{i_{(1)}} = j) \\ v_M(\boldsymbol{\omega}) &= \sum_{i=1}^{N} \delta^{i-1} v_i. \end{aligned}$$

The complexity of designing an entire strategy with this myopic approach is of order k^N. This remains considerably high, keeping in mind that we are just using a small part of the information.

One way of evaluating the myopic strategy is by Monte Carlo sampling $\mathbf{x}^1, \dots, \mathbf{x}^B \sim p(\mathbf{x})$. For each of the B samples the decision is given by the past outcomes, say $x^b_{i_{(1)}} = j$, $x^b_{i_{(2j)}} = l$, ..., and different samples would follow different branches of the decision tree. One could also imagine truncating the myopic evaluation and using the (conditional) naive approach from a certain branch on. We will discuss such approaches in more depth in the next section, when we study more refined forward selection strategies applying the heuristics for the CV at every stage.

5 Look-Ahead and Rolling Horizon strategies

The methods considered in the previous section have the common goal of providing an approximation to the CV. It is therefore natural to use them at different stages of the forward selection procedure. We next propose look-ahead strategies that apply a depth n forward search combining DP with approximations of the CV. The depth n can be chosen by the user. It will depend on the desired accuracy and on the available time and computation power.

In our oil and gas prospect application, there is typically no need to push the forward-backward selection procedure until the very last node. The oil and gas company is most interested in deciding the first few prospects to drill. On the other hand, the approximations we consider apply heuristics for the CV, and in the presence of a large and non-homogeneous number of sites, the associated sequences are not necessarily optimal.

5.1 Look-ahead strategies

Assume that n decisions have been made and that the CV of the field is estimated by a naive or myopic strategy. We propose to assign a large contribution to the first $n < N$ decisions, and a smaller contribution to the remaining $N - n$. We approximate all CVs, and use them to run a restricted n-steps DP. The complexity of the algorithm depends on the size n chosen in the approximation, and it is order of $(\frac{n}{2})! k^{n/2}(N - n)$, when approximating the CV with the naive approach. The strategy is the following:

- Starting point: no nodes have been observed yet: $\boldsymbol{\omega} = \{-, -, \dots -\}$.

- n steps are evaluated with DP, i.e. $v(\boldsymbol{\omega}) = \max\{v_1(\boldsymbol{\omega}), v_2(\boldsymbol{\omega}), \ldots\}$. At each step $v_i(\boldsymbol{\omega})$ is computed according to equation (2).

- After n steps the decision vector has n observed components and $N-n$ still empty (not observed). We define the decision vector at this stage $\boldsymbol{\omega}^*$. For instance, if $N = 6$ and $n = 2$, with observations $x_2 = 2$ and $x_6 = 1$, then $\boldsymbol{\omega}^* = \{-, 2, -, -, -, 1\}$.

- The CV $v(\boldsymbol{\omega}^*)$ is always approximated according to one of the methods introduced in Section 4:

  - *Naive:*
  $$v_N(\boldsymbol{\omega}^*) = \sum_{i=1}^{N-n} \max\left\{\sum_{j=1}^{k} r_i^j p(x_i = j|\boldsymbol{\omega}^*), 0\right\},$$
  We can also fix an order for the $N-n$ prospects, based of their intrinsic values, in order to discount the values in a particular way.

  - *Myopic:*
  Similar to what was has been done in Section 4, we now approximate the CVs with a stepwise procedure, computed in the following way:
  $$\begin{aligned} v_1 &= \max\left\{\sum_{k=1}^{3} r_i^k p(x_{i_{(1)}} = k|\boldsymbol{\omega}^*), 0\right\} \\ v_2 &= \sum_{j=1}^{3}\left(\max\left\{\sum_{k=1}^{3} r^k_{x_{i_{(2j)}}} p(x_{i_{(2j)}} = k|x_{i_{(1)}} = j, \boldsymbol{\omega}^*), 0\right\}\right) p(x_{i_{(1)}} = j|\boldsymbol{\omega}^*), \ldots \\ v_M(\boldsymbol{\omega}^*) &= \sum_{i=1}^{N-n} \delta^{i-1} v_i, \end{aligned}$$

5.2 Rolling horizon look-ahead strategies

We next combine different look-ahead searches and forward selection strategies. We suggest the idea depicted in Figure 2, where one first runs a look-ahead search of depth n. Next, the best node is selected. Given the outcome of this node, a second search of depth n is performed, and so on.

We call these strategies *Depth n* (in the following Dpt n) rolling horizon look-ahead (RHLA) strategies (see Le and Day [1982] and Alden and Smith [1992]). It is interesting to note that a Dpt 0 strategy coincides with a full naive or myopic approach (depending on the approximation chosen for the CV), while a Dpt $N-1$ strategy coincides with a full evaluation of the decision tree, and therefore with the DP presented in equation (1).

This RHLA strategy is a *forward selection*, but it partially accounts for future scenarios in its look-ahead length-n DP procedure. In the RHLA strategy we explore the tree up to a certain fixed depth n, but we draw conclusions just about the first best site. Since at every step we rerun the strategy, we can incorporate at this step the outcome of the sample, instead of exploring all the possible combinations of evidence.

The resulting algorithm has the same computational complexity as the myopic strategy, with an additional factor due to the complexity of the look-ahead strategy in itself. In total we have a complexity of $(\frac{n}{2})! k^{n/2}(N - n) \cdot k^N$. Note that this strategy can always be computed in a forward selection manner. It is however much harder to evaluate the strategy, for instance to compute the associated value, or the variability in the computed sequences over different outcomes.

For a small number of nodes N, one can compute the values probabilistically for different depths n RHLA strategies. For larger dimensions we suggest to use Monte Carlo sampling to evaluate the different strategies.

We then draw samples from the graphical model with joint distribution $p(\mathbf{x})$. We run the RHLA depth n procedure to select nodes, and for each step in the forward selection we plug in the outcomes according to the relevant sample at that node. This approach mimics what would happen in hypothetical scenarios, and we can say that we are *playing the game*.

Given one realization from the graphical model, the pseudo algorithm is presented in Algorithm 1. The algorithm presents two parts: a first one, that constitutes the core of the algorithm from where we call the

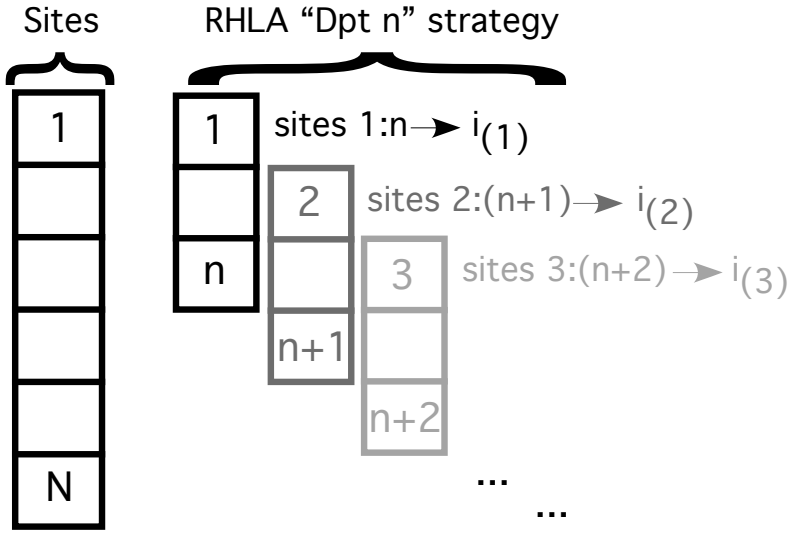


Figure 2: Rolling horizon look-ahead strategy, Dpt n; at every step we run a DP strategy using n sites for finding the best node, and then we update the strategy with the outcome of that node.

recursion, and a second one that presents the recursive function itself. In the core we find a `while` loop that is necessary to terminate the algorithm when all the nodes have been explored and an `if` condition that breaks the process if none of the nodes presents a positive CV. In the recursive function we have an `if` condition that ensures that the correct depth is achieved, and a `for` loop that goes through all the not-yet-explored nodes. When running a RHLA strategy on small examples (cfr. Section 6.1) there is the possibility to run a RHLA for every possible evidence, spanning the whole sample space. By averaging the revenues and costs collected through the strategy, we get a value that coincides (exact and myopic case) or approximates (RHLA case) the estimated final value. In large examples (cfr. Section 6.2) this is not possible and we estimate the final values through a Monte Carlo sampling procedure.

5.3 Pruning strategies

The look-ahead strategies share the idea of choosing a priori the depth n of the search tree. This choice must be done before running an approximation. In practice, we choose n based on the available computation time.

The problem is that we often explore branches of the decision tree that are useless for designing an optimal strategy, and we do not privilege enough branches that can give a stronger contribution to the value. We next design adaptive strategies based on tree-pruning, accounting for the value of the different branches. These idea is inspired by similar ideas applied in contiguous fields, like the chess computer-based algorithms.

We prune the branches of the tree that are very unlikely. In this way we do not have to explore all the combinations, and we reduce the complexity of the algorithm. We define threshold parameter ε such that

$$\text{if } P(\boldsymbol{\omega}_i^j) < \varepsilon \text{ then } v(\boldsymbol{\omega}_i^j) \approx v(\boldsymbol{\omega}_i^*),$$

and we use one of the methods described in Section 4 in order to approximate the CV.

A more refined approach is to decide which branches to explore based on the value of the nodes. This reduces the number of nodes to explore. The method can either be based on the intrinsic value of the individual node under consideration or a look-ahead evaluation of depth 1.

The pseudo-algorithm is the following:

- $\boldsymbol{\omega}_0 = \{-, -, \ldots, -\}$

Algorithm 1 Evaluating a Rolling Horizon Look-Ahead strategy of Depth n

$\boldsymbol{\omega} = [-, -, \ldots, -]$ # Dynamic programming outcome vector
$y = 0$ # Rolling horizon counter
$val = 0$ # Value counter
$seq = [\ ]$ # Best sequence vector
Sample $\mathbf{s} \sim p(\mathbf{x})$ # Current sample
while $\#[\omega_i = \{-\}] > 0$ **do**
  $[v, j] = f(\boldsymbol{\omega}, 1)$
  if $v > 0$ **then** # CV positivity condition
    $\omega_j = s_j$ # Set sampled outcome s_j at selected node j
    $val \mathrel{+}= \delta^y \cdot r_j^{s_j}$ # Discounting of revenues
    $seq_{y+1} = j$ # Selected node is j
  else
    `break`
  end if
  $y{+}{+}$
end while
return val
return seq

function $[v, j] = f(\boldsymbol{\omega}, d)$ # Input: Current state, current depth
if $\#[\omega_i = \{-\}] == 0$ **then** # Last iteration condition, stop
  j=0
  v=0
else if $d \leq n$ **then** # "Depth n" condition, continue DP
  for $i : \omega_i = \{-\}$ **do**
    for $l = 1 : k$ **do**
      $[v, j] = f(\boldsymbol{\omega}_i^l, d + 1)$ # DP iteration at next depth level
      $v_i^l = r_i^l + \delta \cdot v$
    end for
    $v_i = \sum_{l=1}^{k} \{p(x_i = l|\boldsymbol{\omega}) \cdot v_i^l\}$
  end for
  $v = \max_i\{v_i, 0\}$
  $j = \arg\max\{v_i\}$
else # Reached depth n, compute naive CV
  $j = 0$
  $v(\boldsymbol{\omega}) = \sum_{i:\omega_i=\{-\}} \max\left\{\sum_{l=1}^{k} r_i^l \cdot p(x_i = l|\boldsymbol{\omega}), 0\right\}$
end if
end function

- `for i=1:N` we order the segments on the basis on an approximate CV, that can be either of the following:
  - Intrinsic value: $v(\boldsymbol{\omega}_i) = \sum_{j=1}^{k} r_i^k p(x_i = j)$
  - Look-ahead Dpt 1 value:
    $v(\boldsymbol{\omega}_i) = \sum_{j=1}^{k} p(x_i = j) \left[r_i^j + \sum_{s=2}^{N} \max\left\{\sum_{l=1}^{k} \delta^s r_{(s)}^l p(x_{(s)} = l|x_i = j), 0\right\}\right]$
- Keep only the $N - N_{\text{prun}}$ maximum nodes with the highest values and move to the second level of depth in a RHLA framework. For the N_{prun} nodes with minimum values, use the approximated values already computed (Intrinsic or Look-ahead Dpt 1).

In practice, N_{prun} cannot be too small (too many paths to explore), nor too large (we risk to abandon paths that may result being interesting). We will use the pruning strategies to speed up the computations on large graphs.

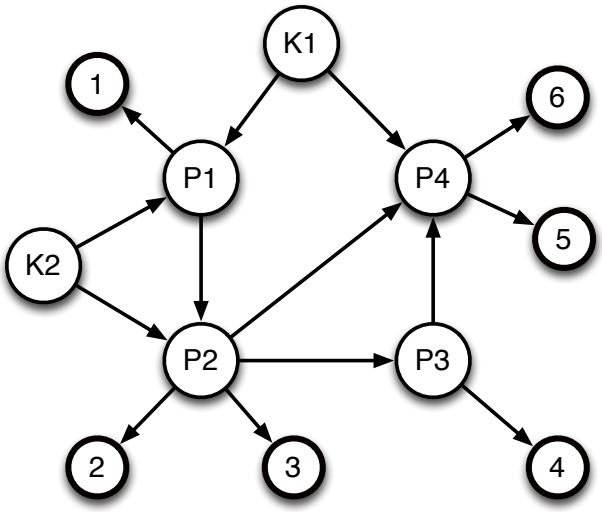


Figure 3: BN used in the 1st case study. We indicate with the letter K the nodes denominated *kitchens*, i.e. zones where HC have been generated, with letter P auxiliary nodes that are functional to establish the desired correlation structure, and with numbers the six nodes where we can drill.

6 Results

We first study a small BN model, where the exact DP solution is available. This allows us to compare the suggested strategies with the exact solution. This synthetic study also anticipates the behavior of the approximations on the BN case study from the North Sea, with 25 prospects. Finally, we analyze a MRF model for an oil reservoir. We construct sequential exploration schemes and interpret the results of different strategies.

6.1 A small Bayesian Network example

We are first interested in exploring the accuracy and the results of our methods on a small BN example (Figure 3). We use a small DAG with $M = 12$ nodes. The nodes denoted $K1$, $K2$, $P1$, $P2$, $P3$ and $P4$ are auxiliary nodes that cannot be drilled. They are motivated by geological mechanisms that are needed to introduce a realistic correlation structure in the network. The two K-nodes represent kitchens, i.e. areas where the hydrocarbon (HC) generation has been or still is in place, and where the migration of HC started. The P-nodes represent geological macro-regions able to store HC. Finally, the bottom numbered nodes, $1, \ldots, 6 = N$ in Figure 3, are prospect nodes where the oil and gas company considers drilling wells. The cost and revenues and marginal probabilities are summarized in Table 1. We designed the DAG to have large variabilities both in the likelihood of finding HC and in the related volumes (revenues). The intrinsic values, i.e. the marginal a priori values of the prospect, are all very close to 0: this makes the case harder to solve. The conditional probabilities defined by the edges are based on geological reasoning and explained in details in [Martinelli et al., 2011]. They impose some learning in the model, once we collect evidence.

In this small case we can compare the result of approximate strategies with the exact DP solution. The discounting parameter δ is fixed, here and in the next simulations, to a realistic value of 0.99, as suggested in [Bickel and Smith, 2006]. The first comparison is presented in Table 2. Here, the result of the strategies up to the third best choice are presented, for the naive and myopic strategies, for exact DP and for *Dpt n* strategies, up to $n = 4$. According to the exact strategy, if oil or gas is found in the first segment chosen (in this case, number 6), the suggestion is to keep drilling in the same area (under P4 node) with segment number 5. If the well reports a negative result, it makes sense to immediately explore another part of the field. The naive approach does not take this dichotomy into account because the sequence is fixed a priori. The myopic approach uses a different strategy for the oil/gas and the dry case, but since the depth of the search is in this case short-sighted, the conclusion is to stop drilling immediately after a dry well.

In addition to comparing strategies, we study the computational time and the final value, $v(\boldsymbol{\omega}_0)$. We notice that, despite slightly different strategies, the final values are quite close to the exact for Dpt 2 or even Dpt 1,

Table 1: Parameters r_i^k for the 1st case study, and relative Intrinsic Values (marginal probability of success/failure times revenues/costs)

k \ i	1	2	3	4	5	6
r_1^k (dry)	-20	-25	- 1	-15	- 22	- 8
r_2^k (gas)	6	3	9	0	4	5
r_3^k (oil)	3	1	6	7	2	1
$p(x_k = 1)$	0.20	0.10	0.80	0.30	0.15	0.34
$p(x_k = 2)$	0.52	0.72	0.01	0.02	0.68	0.53
$p(x_k = 3)$	0.28	0.18	0.19	0.68	0.17	0.13
Intrinsic Value	-0.04	-0.16	0.43	0.15	-0.25	0.05

Table 2: Results of the sequential exploration program for the 1st case study, for naive, myopic, exact and Dpt1 to Dpt 4 strategies. $i_{(1)}$, $i_{(2)}$ and $i_{(3)}$ are respectively the first, the second and the third best site selected. Q means *quit* (the strategy). Final value is $v(\boldsymbol{\omega}_0)$.

	Naive	Myopic	**Exact**	Dpt1	Dpt2	Dpt3	Dpt4
$i_{(1)}$	3	3	6	6	6	6	6
$i_{(2)}\|x_{i_{(1)}} = dry$	4	Q	3	3	3	3	3
$i_{(2)}\|x_{i_{(1)}} = gas$	4	2	5	2	5	5	5
$i_{(2)}\|x_{i_{(1)}} = oil$	4	2	5	2	4	4	5
$i_{(3)}\|x_{i_{(1)}} = dry, x_{i_{(2)}} = dry$	6	Q	Q	Q	Q	Q	Q
$i_{(3)}\|x_{i_{(1)}} = dry, x_{i_{(2)}} = gas$	6	Q	2	2	2	2	2
$i_{(3)}\|x_{i_{(1)}} = dry, x_{i_{(2)}} = oil$	6	Q	2	2	2	2	2
$i_{(3)}\|x_{i_{(1)}} = gas, x_{i_{(2)}} = dry$	6	4	4	5	4	4	4
$i_{(3)}\|x_{i_{(1)}} = gas, x_{i_{(2)}} = gas$	6	4	4	5	4	4	4
$i_{(3)}\|x_{i_{(1)}} = gas, x_{i_{(2)}} = oil$	6	4	4	5	4	4	4
$i_{(3)}\|x_{i_{(1)}} = oil, x_{i_{(2)}} = dry$	6	4	4	5	3	5	4
$i_{(3)}\|x_{i_{(1)}} = oil, x_{i_{(2)}} = gas$	6	4	4	4	2	2	4
$i_{(3)}\|x_{i_{(1)}} = oil, x_{i_{(2)}} = oil$	6	4	4	4	2	2	4
Final Value	0.63	1.67	4.960	3.85	4.84	4.93	4.957
Time	0.24 sec	0.24 sec	85.6 sec	0.43 sec	3.52 sec	16.11 sec	48.22 sec

with a much smaller computational time. The final value reported in the table is only the approximate value found when optimizing the strategy for the Dpt 1-4 algorithms. In practice, their value will be higher, since the approximation is based on using a naive strategy at the end, whereas the algorithm always looks ahead running new Dpt n searches. We therefore believe that the best comparison is not much about comparing values, but more about comparing the proposed strategies on real scenarios.

Since the dimension of the problem is relatively small, we can directly span the whole sample space and compute all RHLA strategies exactly, as anticipated in Section 5.2. This is the approach adopted in Table 3. Here we compare the evaluation of the different strategies (naive, myopic and different depths of look ahead strategies) on the whole sample space generated by the BN of reference. We therefore test $3^6 = 729$ combinations of evidence on the nodes of interest, and we compute the likelihood of these scenarios by summing out the outcome at the top nodes. In this way, we can compute the average performance of the strategies, and the related variance.

The result tells us that, when applied in practice on this simple test case, the two simple strategies perform extremely poorly, while the look ahead strategies perform significantly better. In particular, Dpt 2 and Dpt 3 perform almost as good as Dpt 4 (which in this case corresponds exactly to the Exact Strategy), with a significant reduction in the computational time. An interesting argument in favor of the look-ahead strategies can also be made considering the variance. If we consider the second row of Table 3, we observe an increasing variance

Table 3: Sequential exploration program, methods' comparison following a complete RHLA procedure (Section 5.2)

Revenues Distribution	Naive	Myopic	Dpt1	Dpt2	Dpt3	Dpt4
Average value	0.63	1.68	4.89	4.95	4.959	4.960
Standard deviation	12.664	8.815	15.268	14.878	14.877	14.869

between the simpler strategies and the look-ahead strategies. We first notice that the variance of the revenues distribution under the naive strategy just reflects the variance of the marginal a priori distribution for prospects 3, 4 and 6:

$$\sigma_N^2 = 12.664^2 = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\left\{(r_i^3 + \delta r_j^4 + \delta^2 r_j^6 - \bar{r})^2 \cdot p(x_3 = i, x_4 = j, x_6 = k)\right\}$$

Furthermore, we can relate the low variance of the myopic strategy to a spike on the value '-1', that corresponds (see Table 1) to the loss for a likely ($p = 0.8$) dry observation in segment 3. Since a dry outcome at the first site in the myopic strategy would imply quitting the search, we are ultimately left with a high number of scenarios whose revenues' outcome is simply -1. If we remove these scenarios, the variance shrinks from myopic to Dpt 1 to Dpt 4, providing another argument in favor of these strategies. A lower variance in this case coincides with a more stable estimate and a lower risk when starting an exploration campaign, and this can be almost as important as a high final value.

6.2 A Case Study from the North Sea

We next study a BN model developed for 25 HC prospects in the North Sea. The network (Figure 4) is taken from Martinelli et al. [2011], and represents a model of HC fields in the Norwegian part of the North Sea. The network includes the same characteristics as the small test study, but there are now 25 possible drilling locations (numbered 1 through 25 in Figure 4). We use the same probability model as in Martinelli et al. [2011]. This gives the marginal probabilities in Table 4. The joint model is constructed from the DAG. Many geological assumptions are used when building the model. In particular, gas will tend to replace oil in the HC migration. Thus, with a single edge between two nodes in the graph we have $p(x_k = 1|x_k^{pa} = 2) = 0$, $p(x_k = 2|x_k^{pa} = 1) > 0$, where 1 is gas and 2 is oil. Dry outcomes result from migration failures. Similar to the previous model, the DAG has a three-level structure representing the geological mechanisms. For decision making we are interested in the bottom nodes of the network, that represent identified prospects whose volumes and costs are assumed known. The corresponding revenues and costs (in Million USD) are listed in Table 4. Here, we avoid shared prospect costs that would make the computational task harder, and the interpretation more difficult. In this real case, there are still some nodes where the probability of success (and consequently the intrinsic value) may change substantially given the outcome in other nodes. However, some nodes would be drilled or not drilled in any event, no matter the strategy.

Given the BN model we are interested in identifying a drilling sequence that gives maximum profit under some criterion. Table 5 shows the results of comparing the naive, myopic and three depth (Dpt) level heuristic strategies. Note that final values are now quite close to each other for all the approximations considered. The dynamic decisions depend less on the strategy than in the synthetic case in the previous section. Still, there is a clear increase of about 3000 Million USD when using the Dpt 3 strategy rather than the naive one. We have again run the different strategies on a number of simulated scenarios (Table 6). Since the computational time required by the RHLA strategy is order of hours per step, we have considered a sample size of 200 and followed the algorithm described in Section 5.2. For the same reason we will focus from now on in a comparison between simple strategies, such as naive or myopic, and two RHLA strategies, namely Dpt 1 and Dpt 2.

The difference is not very large, but the Dpt 1 and Dpt 2 strategies perform better than the myopic one. In particular, Dpt 2 strategies give on average around 400 Million USD more than the myopic strategy. It is particularly important to investigate the reason of this improvement. A first hint is given by the last three lines of Table 6. Here we can notice that more complex strategies suggest in general to drill more than simpler

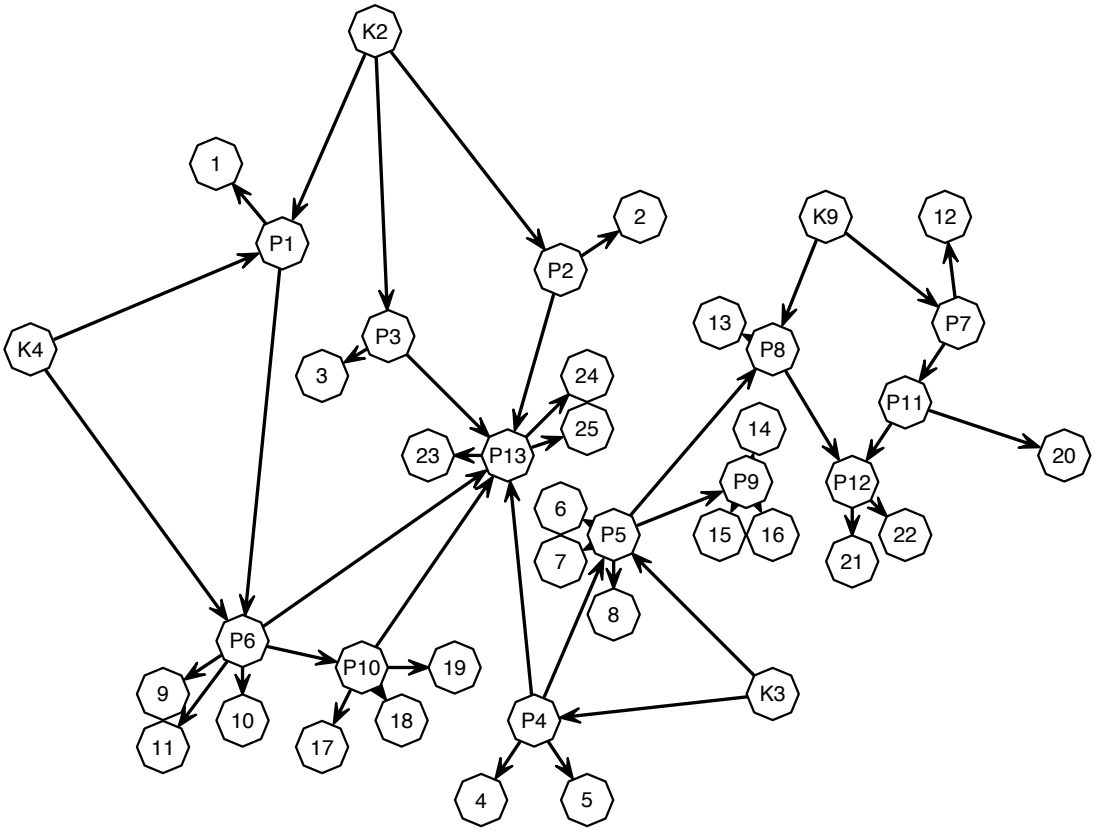


Figure 4: Network used in the 2nd case study. In this case we have 25 drilling prospects, identified with the nodes from 1 to 25, where we can possibly drill. The BN was first presented in [Martinelli et al., 2011].

strategies. The typical case is that whenever an area is found dry, the intrinsic values for all the segments around drop, and just long-sight strategies can look for the potential remaining values. Nonetheless, a higher number of drilled sites translates into an effective improvement of the result just if the newly drilled sites have a positive outcome. This is the case that we are considering, since among the 1.49 sites more drilled with Dpt 2 strategy, just 0.13 are on average dry, while an outstanding 1.36 are found gas or oil.

Figure 5 shows what happens to all the 25 prospects when treated with different strategies. In many cases (segments $2, 6, 7, 8, 9, \ldots$) the marginal probability of a positive discovery is higher for the Dpt 1 approach wrt to the myopic approach. It is interesting to note that, considering for example prospect 8, both the marginal probability of a positive discovery is increased and of a negative one is decreased. This is explained by Table 7, that tells us that we are drilling prospect 8 a smaller number of times with the Dpt 1 strategy, but with higher efficiency. Conversely, in the cases of prospect 14, we have the same marginal accuracy for myopic and Dpt 1 strategy, but we still have a benefit in economical terms, since we are drilling the site a higher number of times: technically, in this case, with Dpt 1 strategy we drill prospect 14 *only and all* the times that this segment is valuable. Finally, for prospect 20, we increase both the accuracy and (substantially!) the percentage of drilled times, resulting in a strong economical return. The results are difficult to interpret in some extreme cases, like prospect 2. Here we note how the accuracy of Dpt 1 strategy is 100%, while the accuracy of myopic strategy is not known (both the bars are 0). This is due to the fact that with myopic strategy we never drill prospect 2, thus we can not say anything about the accuracy of such strategy here; on the other side, with Dpt 1 we drill it just 2% of the times, but in these cases we always find oil or gas, therefore the accuracy boosts at 100%. This is the reason for listing $P(drilled)$ in Table 7 as an important diagnostic factor.

We finally consider (Table 8) what happens in single scenarios, i.e. what are the results when "playing the game" on a few samples with different strategies (myopic, Dpt 1 and Dpt 2). We immediately see that the myopic approach performs either brilliantly (sample 2) or extremely poorly (samples 1 and 3), while the revenues guaranteed by the other two approaches are, in a way, more stable: this is consistent with the type of approach,

Table 4: Costs, revenues, marginal probabilities and intrinsic values for the 25 sites taken into account in the 2nd case study.

Prospect k	r_1^k	r_2^k	r_3^k	$p(x_k = 1)$	$p(x_k = 2)$	$p(x_k = 3)$	Int. Value
1	-3000	3032	2783	0.20	0.52	0.28	1756
2	-900	125	236	0.40	0.21	0.39	-242
3	-2400	1094	1085	0.60	0.26	0.14	-1004
4	-1800	188	377	0.28	0.57	0.15	-337
5	-600	594	1321	0.20	0.29	0.51	729
6	-1500	156	1132	0.21	0.04	0.75	534
7	-3600	406	3255	0.34	0.03	0.63	844
8	-2100	750	6934	0.52	0.02	0.46	2107
9	-2700	2751	1415	0.10	0.72	0.18	1965
10	-1200	2751	1415	0.20	0.64	0.16	1747
11	-2400	500	4576	0.80	0.01	0.19	-1040
12	-2700	125	802	0.19	0.04	0.77	123
13	-4500	0	0	0.36	0.32	0.32	-1620
14	-1800	188	94	0.10	0.45	0.45	-53
15	-2100	563	613	0.10	0.45	0.45	319
16	-3600	31	613	0.10	0.03	0.87	172
17	-3300	250	3161	0.61	0.22	0.17	-1410
18	-1200	688	8963	0.30	0.02	0.68	5697
19	-2100	250	3349	0.37	0.02	0.61	1285
20	-5400	969	660	0.18	0.41	0.41	-312
21	-1800	1375	3444	0.49	0.26	0.25	336
22	-2400	3220	2264	0.41	0.47	0.12	783
23	-3000	156	1274	0.10	0.04	0.86	806
24	-2400	2782	1604	0.10	0.72	0.18	2052
25	-2700	2251	1274	0.30	0.56	0.14	629

Table 5: Results of the sequential exploration program for the 2nd case study, for naive, myopic, and Dpt1-3 strategies. $i_{(1)}$ and $i_{(2)}$ are respectively the first and second best sites selected. Final value is $v(\boldsymbol{\omega}_0)$.

	Naive	Myopic	Dpt1	Dpt2	Dpt3
$i_{(1)}$	18	18	15	22	18
$i_{(2)}\|x_{i_{(1)}} = dry$	8	8	21	18	24
$i_{(2)}\|x_{i_{(1)}} = gas$	8	19	22	18	22
$i_{(2)}\|x_{i_{(1)}} = oil$	8	19	22	18	22
Final Value	20213	21321	21841	22535	23197
Time	$<$ 1 sec	$<$ 1sec	4.72 sec	175 sec	4h

since we understand that being more long-sighted correspond to being more cautious in our decision. The difference in the revenue variances recorded in the two samples confirms this statement, with a strong decrease recorded when comparing myopic strategy with RHLA strategies.

If we look closer, we discover other signs that agree with this statement. The first 5 sites picked by a myopic approach are all on the left part of the network. In simple words, we start our search from the left side (prospect 18), and keep exploring the same side for a long period as long as the results are positive. The Dpt 1 approach suggests to jump 3 times between the left and the right side of the network just in the first five picks (15 and 22, then 18, then 12, then 24), even if the results are very good: this means that while we consolidate the strength of a part of the network, we also explore if other parts of the networks are likewise strong. This way of exploring has the further benefit, in this particular case, to allow a longer series of straight good results (7 versus 5).

Table 6: Sequential exploration program, methods' comparison following RHLA procedure (Section 5.2) with a sample of 200 scenarios.

	Myopic	Dpt1	Dpt2
Average value	24256	24500	24668
Standard deviation	13632	12474	12586
Average # sites drilled	16.62	18.01	18.11
Average # sites found dry	2.89	3.02	3.02
Average # sites found gas or oil	13.73	14.99	15.09

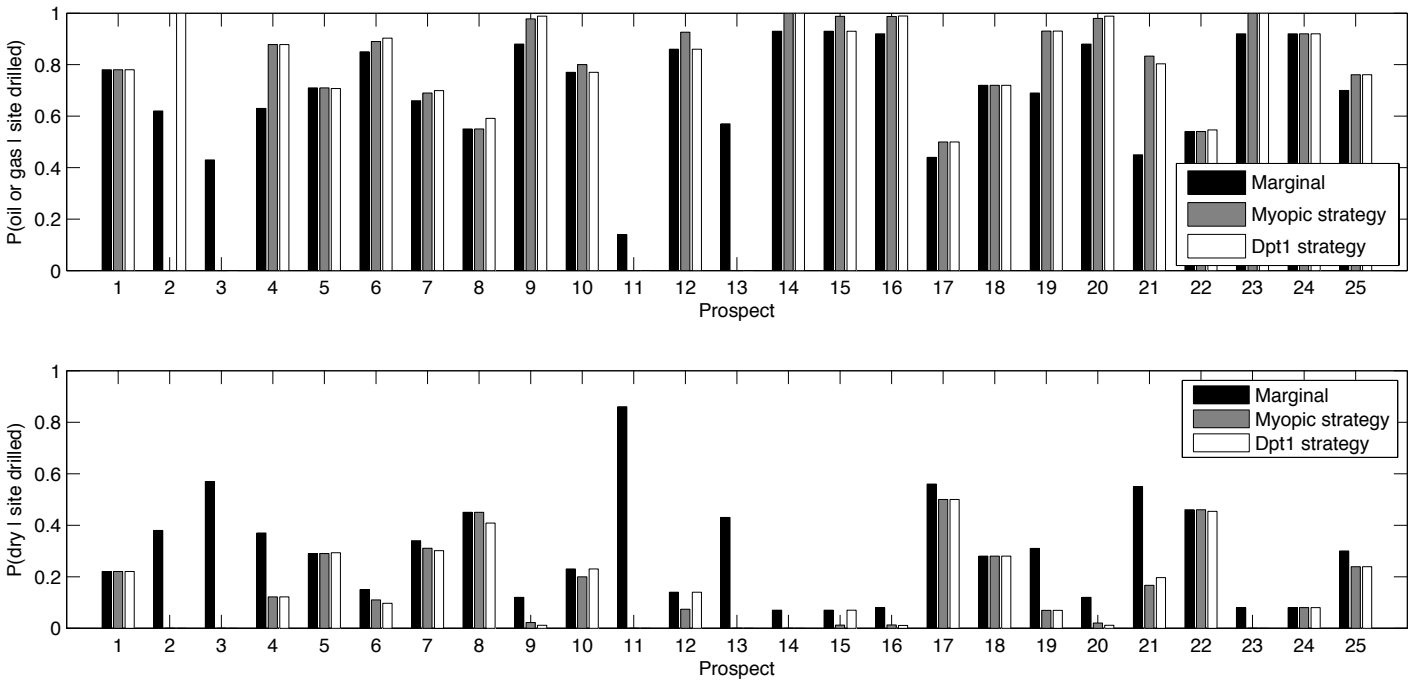


Figure 5: Probabilities of positive and negative discoveries for the 25 sites analyzed in the 2nd case study. We compare the marginal probabilities a priori with the frequency of successes following a myopic or Dpt 1 strategy.

Table 7: Marginal probabilities of positive and negative discoveries and probability of drill for three prospects, namely prospect 8, 14 and 20. $P(drilled)$ reports the frequency of exploration provided by myopic (Myo) or depth 1 RHLA (Dpt 1) strategy.

Prospect	8	14	20
P(oil/gas)	0.55	0.93	0.88
P(dry)	0.45	0.07	0.12
P(oil/gas | Myo)	0.55	1	0.98
P(dry | Myo)	0.45	0	0.02
P(oil/gas | Dpt1)	0.59	1	0.99
P(dry | Dpt1)	0.41	0	0.01
P(drilled, Myo)	1	0.8	0.5
P(drilled, Dpt1)	0.93	0.93	0.86

The myopic strategy looks to perform better in very lucrative scenarios: this is consistent with the theoretical definition of myopic strategy, that *goes for the best first*. In an hypothetical scenario of all prospects containing oil, the myopic strategy would be difficult to beat, and this situation is very similar to the one drawn in the second sample. In such situation an even simpler naive strategy could beat both myopic and RHLA strategies,

provided that there is not enough correlation to confirm the nodes characterized by low probabilities and high volumes.

Table 8: Ordered list of sites chosen with myopic, Dpt 1 and Dpt 2 strategy for 3 different samples taken from the RHLA evaluation.

	1-Myo		1-Dpt1		1-Dpt2		2-Myo		2-Dpt1		2-Dpt2		3-Myo		3-Dpt1		3-Dpt2	
	18	3	15	2	22	3	18	3	15	3	22	3	18	1	15	2	22	1
	19	3	22	3	18	3	19	3	22	3	18	3	8	1	22	1	18	1
	9	2	18	3	15	2	9	2	18	3	15	3	24	3	18	1	24	3
	24	1	12	3	24	1	24	2	12	3	24	2	1	1	24	3	10	3
	10	3	24	1	21	1	10	2	24	2	21	3	23	3	10	3	15	2
	8	1	21	1	19	3	8	1	21	3	19	3	22	1	12	3	8	1
	1	2	19	3	8	1	1	2	19	3	8	1	5	1	8	1	1	1
	22	3	8	1	9	2	23	3	8	1	9	2	25	3	1	1	9	3
	21	1	9	2	10	3	25	2	9	2	10	2	10	3	5	1	5	1
	5	1	10	3	1	2	22	3	10	2	1	2	9	3	9	3	7	3
	12	3	1	2	7	2	21	3	1	2	5	3	0	0	7	3	23	3
	20	3	7	2	5	1	5	3	5	3	7	3	0	0	23	3	6	3
	0	0	5	1	6	3	7	3	7	3	23	3	0	0	6	3	16	3
	0	0	6	3	16	3	6	3	23	3	12	3	0	0	20	1	25	3
	0	0	20	3	12	3	15	3	25	2	25	2	0	0	16	3	12	3
	0	0	16	3	20	3	16	3	6	3	6	3	0	0	25	3	20	1
	0	0	14	3	14	3	12	3	20	3	20	3	0	0	14	2	14	2
	0	0	0	0	0	0	20	3	16	3	16	3	0	0	0	0	0	0
	0	0	0	0	0	0	14	2	14	2	14	2	0	0	0	0	0	0
	0	0	0	0	0	0	4	2	4	2	4	2	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	16081		18126		**18196**		**37293**		36859		37087	0	-2455		-1208		**-1146**	

In summary we learnt that there are clear differences in the suggested drilling strategies for the naive, myopic and Dpt n computations. A myopic strategy gives a large improvement over the naive strategy in our network, and this will always be the case as long as the prospects are dependent and not obviously profitable or unprofitable. The extra gain from running Dpt n strategies is in this 25 prospect case seen as a larger payoff in money for the computing time spent. The Dpt n strategies also suggest other drilling locations. In a practical setting, our recommendation is to run a Dpt n search with as large n as computationally feasible. Note that this can be done stepwise. In many situations we only need to identify the first prospect, and can wait for the result there before computing the next. This is the practical exploration scenario a petroleum company faces.

6.3 MRF case study

In the third application we apply our sequential exploration technique on a larger dataset, where the current knowledge consists of geological knowledge combined with seismic data. The data and the case study are explained in Bhattacharjya et al. [2010]. The MRF model has 3-colors, where the three distinctions of interest represent respectively *oil saturated sand* ($x_i = 1$), *brine saturated sand* ($x_i = 2$) and *shale* ($x_i = 3$). We use a lattice representation of the field, with 20×5 cells, i.e. $M = N = 100$.

The prior model is a categorical first-order MRF [Besag, 1974]:

$$p(\mathbf{x}) \propto \exp\left\{ \beta \cdot \sum_{i \sim j} \mathbb{I}(x_i = x_j) + \sum_{i=1}^{N} \alpha_i(x_i) \right\},$$

where $i \sim j$ denotes the sum over all neighboring lattice nodes (north, east, south, and west). The parameter β imposes spatial interaction. The α_i terms are set from a priori geological knowledge [Bhattacharjya et al., 2010]. We work with a highly correlated MRF (β=0.8).

The seismic data $\mathbf{y}$ are incorporated in the MRF model $\mathbf{x}$ through a Gaussian likelihood model [Eidsvik et al., 2004]. At each cell bivariate seismic data, shown in Figure 6 are modeled by:

$$p(\mathbf{y}_j | x_j) \sim \mathcal{N}\left\{ \begin{pmatrix} \mu_1(x_j) \\ \mu_2(x_j) \end{pmatrix} \begin{pmatrix} 0.06^2 & -0.007 \\ 0.007 & 0.17^2 \end{pmatrix} \right\},$$

where: $\boldsymbol{\mu_1} = (0.03, 0.08, 0.02)$ and $\boldsymbol{\mu_2} = (-0.21, -0.15, 0)$.

The posterior model is defined by:

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{x}) \prod_{j=1}^{100} p(\mathbf{y}_j|x_j).$$

This posterior is a MRF with new α_i terms which now also depend on the data values.

As was done in Bhattacharjya et al. [2010], we assign a fixed cost of 2 Million USD for drilling a dry well (state 2 or 3), while we have a potential revenue of 5 million USD when finding an oil saturated sand (state 1). Before drilling we have the situation represented in Figure 6. In the top row we see the bivariate seismic data, in the bottom row we see the the prior geological knowledge and the posterior oil saturated sand probability.

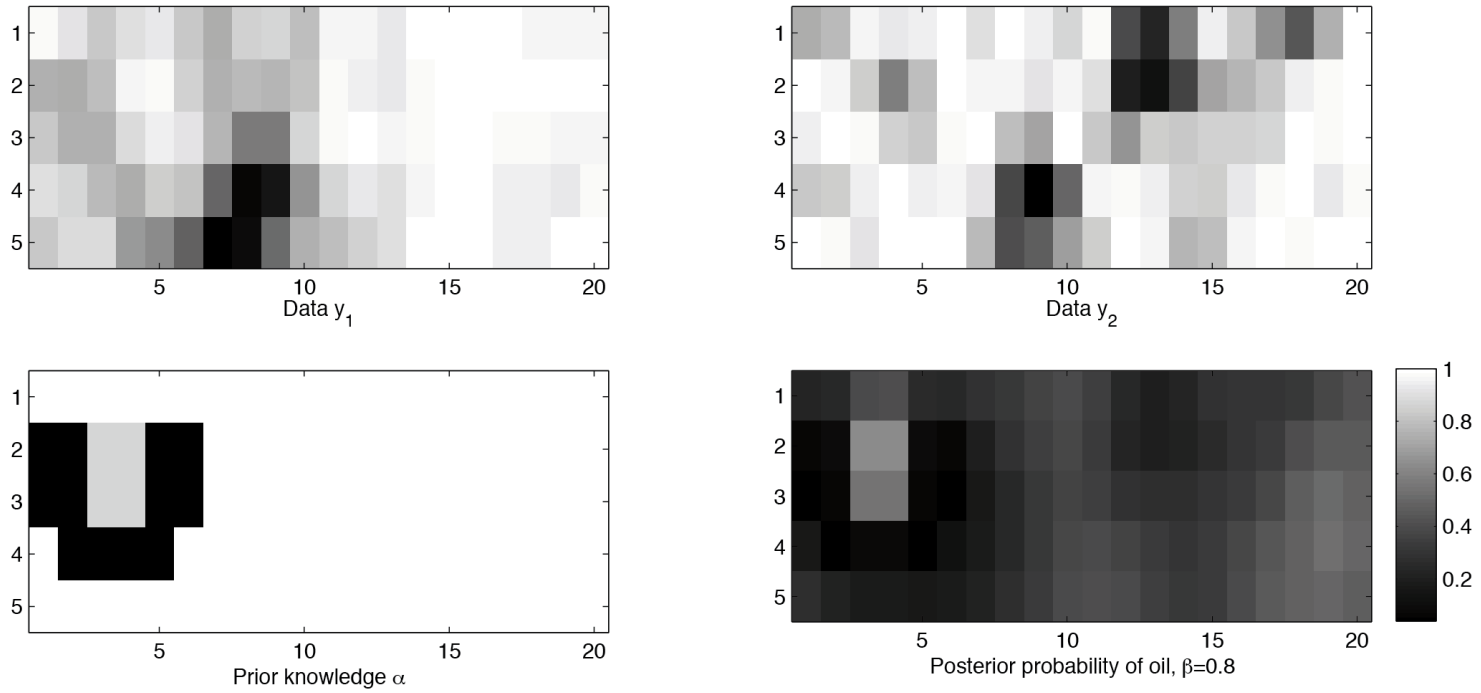


Figure 6: Initial conditions of the MRF described in Section 6.3. Top left: reflectivity seismic data. Top right: amplitude seismic data. Bottom left: prior geological knowledge. Bottom right: Probability of oil saturated sand with interaction parameter $\beta = 0.8$.

The combinatorial complexity prevents us from running a full search, therefore we try different levels of approximations, from the myopic strategy to more complex depth searches. We present in Figure 7 the results of myopic, Dpt 1 and Dpt 2 strategies. While the first myopic strategy reproduces the same pattern that we observe in the posterior probability of oil (bottom right, Figure 6), the second Dpt 1 strategy shows a different pattern. The sites on the eastern part of the basin, those that get the higher expected revenues (due to a strong prior probability of oil sand), are not anymore selected in the first step, because they are surrounded by sites with low profitability. On the other hand, the central sites, whose profitability was not that high, but overall good over a large area, are privileged by a Dpt 1 strategy. The same behaviors appear in the bottom part of Figure 7, that report the best first and second choice for Dpt 2 strategy. We can further note that the expected final values increase with more complex strategies.

For a petroleum company that wants to explore a reservoir zone, we expect the drilling strategy to depend heavily on the amount of data available (seismic data and well data in the neighborhood of the reservoir), and the cost of establishing new infrastructure. In this example we built the first element into the MRF model and the second as part of the case-specific utility function. In our situation, the Dpt n strategies clearly select different drilling locations than the myopic approach. This kind of information is useful in an appraisal stage of a reservoir unit

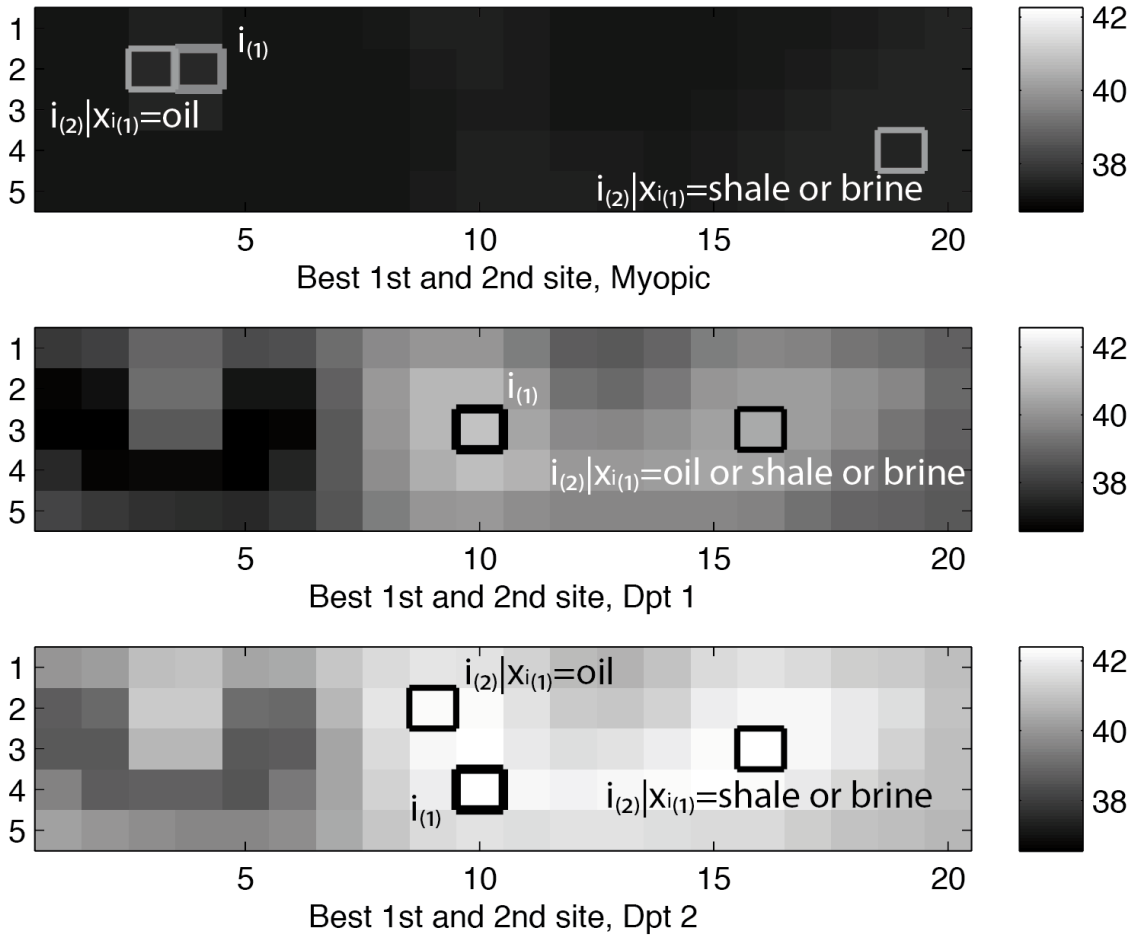


Figure 7: Best 1st and 2nd sites using myopic (top), Dpt1 (center) and Dpt2 (bottom) strategies. The colors correspond to $v_i(\boldsymbol{\omega})$ under the three strategies, where $v_i(\boldsymbol{\omega})$ (cfr. equation 2) represents the CV of the chosen strategy, given that we start drilling at prospect i.

7 Closing remarks

The paper proposes a new approximate solution to sequential decision making. The approximations apply heuristic procedures to estimate the optimization function at different stages of the algorithm. Pruning strategies are also proposed in order to speed up the computation by cutting the less valuable branches of the decision tree.

The methodology is applied to case studies from the petroleum industry. First, a BN model for 25 prospects in the North Sea [Martinelli et al., 2011] is solved. Second, a MRF with 100 lattice cells for a local reservoir is studied. In both cases, we construct approximate drilling sequences. We show how sequential decision making, coupled with a statistical model for the dependence of the field, can yield strategies very different from those based on independent or myopic searches.

We recommend running a strategy of depth n, where n is as large as computationally feasible. In practice a petroleum company would often wait for the outcome of the first well(s) to continue its exploration strategy. It is also possible to run different depth searches and see if results are very dissimilar. In practice the petroleum company can test the depth n strategies over different utility functions, various kinds of risk behavior, and a range of cost and revenue inputs. This means only minor edits to inputs parameters in our implemented algorithms, and provides helpful guidelines when selecting the final exploration policy.

The applications do not limit the scope and the merit of the developed algorithms. One can use the methodology to other selection problems on graphical models. Nodes could for example correspond to clinical tests, in

a problem where the practitioners make sequential decisions. Also, generic variable selection problems or design of experiments for graphs could be envisioned utilizing the same instruments.

We believe that there is large potential for interplay between operational research and recent development for computing multivariate statistical models. The current paper is just one example. Here, the search is built on heuristic strategies, and we have made no attempts to justify the approximation as the optimal solution. It would be interesting to study these problems from a more theoretical perspective, merging knowledge from both operations research, decision theory and statistics.

8 Acknowledgments

We thank the Statistics for Innovation (SFI^2) research center in Oslo, that partially financed GMs scholarship through the "FindOil" project. We acknowledge David Brown and Jim Smith, Duke University, for interesting and useful discussions on this topic.

References

J. M. Alden and R. L. Smith. Rolling horizon procedures in nonhomogeneous markov decision processes. *Operations Research*, 40:S183–S194, 1992.

R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.

L. Benkherouf and J. Bather. Oil exploration: Sequential decisions in the face of uncertainty. *Journal of Applied Probability*, 25(3):529–543, 1988.

D. P. Bertsekas and J. Tsitsiklis. *Dynamic Programming*. Athena Scientific, 1996.

D. Bertsimas and R. Demir. An approximate dynamic programming approach to multidimensional knapsack problems. *Management Science*, 48(4):550–565, 2001.

J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, 36:192–236, 1974.

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

J. Bickel and J. Smith. Optimal sequential exploration: A binary learning model. *Decision Analysis*, 3(1):16–32, 2006.

J. Bickel, J. Smith, and J. Meyer. Modeling dependence among geologic risks in sequential exploration decisions. *SPE Reservoir Evaluation & Engineering*, 11(2):352–361, 2008.

S. Bollapragada and T. E. Morton. Myopic heuristics for the random yield problem. *Operations Research*, 47 (5):713–722, 1999.

R. Cowell, P. Dawid, S. Lauritzen, and D. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Springer series in Information Science and Statistics, 2007.

J. Eidsvik, P. Avseth, H. Omre, T. Mukerji, and G. Mavko. Stochastic reservoir characterization using pre-stack seismic data. *Geophysics*, 69:978–993, 2004.

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

R. Feldmann, P. Mysliwiete, and B. Monien. Studying overheads in massively parallel min/max-tree evaluation. *Proceedings of the sixth annual ACM symposium on Parallel algorithms and architectures*, pages 94–103, 1994.

J. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society, Series B*, 41:148–177, 1979.

V. Goel, I. Grossmann, A. El-Bakry, and E. Mulkay. A novel branch and bound algorithm for optimal development of gas fields under uncertainty of reserves. *Computers and Chemical engineering*, 30:1076–1092, 1979.

G. Kokolis, B. Litvak, W. Rapp, and B. Wang. Scenario selection for valuation of multiple prospect opportunities: a montecarlo simulation approach. *SPE paper 52977, presented at the SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, TX, 20-23 March 1999*, 1999.

S. L. Lauritzen and D. J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B*, 50:157–224, 1988.

K. Le and J. Day. Rolling horizon method: A new optimization technique for generation expansion studies. *IEEE Transactions on Power Apparatus and Systems*, 101(9):3112–3116, 1982.

G. Martinelli, J. Eidsvik, R. Hauge, and M. D. F. rland. Bayesian networks for prospect analysis in the north sea. *AAPG Bulletin*, 95(8):1423–1442, 2011.

A. Miller. The value of sequential information. *Management Science*, 22(1):1–11, 1975.

G. Nemhauser. *Introduction to dynamic programming.* Wiley, 1966.

J. Pearl. *Heuristics: Intelligent Search Strategies for Computer Problem Solving.* Addison-Wesley, 1984.

W. B. Powell. Approximate dynamic programming: Lessons from the field. In S. J. Mason, R. R. Hill, L. Moench, and O. Rose., editors, *Proceedings of the 2008 Winter Simulation Conference*, 2008.

R. Reeves and A. Pettitt. Efficient recursions for general factorisable models. *Biometrika*, 91(3):751–757, 2004.

C. Shannon. Programming a computer for playing chess. *Philosophical Magazine, Series 7*, 41(314):256–275, 1950.

J. . Smith and R. Thompson. Managing a portfolio of real options: Sequential exploration of dependent prospects. *The Energy Journal, International Association for Energy Economics*, 29:43–62, 2008.

J. VanWees, H. Mijnlieff, J. Lutgert, J. Breunese, C. Bos, P. Rosenkranz, and F. Neele. A bayesian belief network approach for assessing the impact of exploration prospect interdependency: An application to predict gas discoveries in the netherlands. *AAPG Bulletin*, 92(10):1315–1336, 2008.

R. Weber. On the gittins index for multiarmed bandits. *The Annals of Applied Probability*, 2(4):1024–1033, 1992.