Modeling Prospect Dependencies with Bayesian Networks*

Gabriele Martinelli[1]

Search and Discovery Article #40553 (2010)
Posted June 30, 2010

*Adapted from oral presentation at AAPG Annual Convention and Exhibition, New Orleans, Louisiana, April 11-14, 2010; reporting joint work with Jo Eidsvik[1], Ragnar Hauge[2], and Maren Drange Førland[2].

[1]Department of Mathematical Sciences, NTNU, Norwegian University of Science and Technology, Trondheim, Norway (gabriele.martinelli@math.ntnu.no)
[2]NR, Norwegian Computing Center, Oslo, Norway.

Abstract

In recent years, great interest has been devoted to the evaluation of prospect dependencies. Here, we look at the use of Bayesian networks based on trap, source, and reservoir probabilities as a tool for analyzing a set of prospects. With such networks in place, we can easily run what-if drilling scenarios to determine drilling sequences. We exemplify this by finding the minimum number of dry wells to drill before abandoning an area.

In oil exploration, the probability of success for a prospect is often split into three independent factors: reservoir, trap, and source. These represent the presence of reservoir rock, a closed trap, and oil migrating into the reservoir, respectively. We model dependencies between prospects by using one Bayesian network for each of these factors.

A Bayesian network is a powerful, yet simple tool for representing dependencies. With one network representing the geological dependencies for reservoir rock, another traps and a third one sources and possible flow paths, it is evident that the networks mirror physical dependencies and that the nodes have physical interpretation. The two first networks use binary success/failure nodes, whereas the source nodes may have four states if both oil and gas are modeled (dry, oil, gas, oil and gas). This discrete state situation fits nicely into the Bayesian network framework.

With the help of a team of experts, we have created such networks. With these networks established, it is easy to evaluate the changes in probability and economic viability for different prospects given new well data. The main strength of this is to feed the model with hypothetical well data, and use it to compute values of information and plan drilling sequences.

Using networks based on a real case, we show how a drilling plan can be created. The setting is that we want to drill as few dry wells as possible before abandoning the area. With a full model for dependencies, it is easy to evaluate the economic viability of the remaining prospects given a set of dry wells, and by running through all small subsets, we find the optimal solution.

We show that Bayesian networks are a powerful tool, both for analyzing and modeling the dependencies between prospects in an area. With such a model in place, it is easy and fast to update probabilities and expected value of prospects as new data become available. This is useful for well planning.

References

Bickel, J.E, and J.E. Smith, 2006, Optimal Sequential Exploration: A Binary Learning Model: Decision Analysis, v. 3/1, p. 16-32, doi:10.1287/deca.1050.0052.

Bratvold, R.B., J.E. Bickel, and H.L. Lohne, 2009, Value of information in the Oil and Gas Industry: Past, Present, and Future: SPE Reservoir Evaluation and Engineering: v. 12/4.

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

Cunningham, P. and S.H. Begg, 2008, Using the value of information to determine optimal well order in a sequential drilling program: AAPG Bulletin, v. 92/10, p. 1393-1402.

Eidsvik, J., D. Bhattacharjya, and T. Mukerji, 2008, Value of information of seismic amplitude and CSEM resistivity: Geophysics, v. 73/4, R59-R69.

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

Van Wees, J.D., H. Mijnlie, J. Lutgert, J. Breunese, C. Bos, P. Rosenkranz, and F. Neele, 2008, A Bayesian belief network approach for assessing the impact of exploration prospect interdependency: An application to predict gas discoveries in the Netherlands: AAPG Bulletin, v. 92/10, p. 1315-1336.

Modeling Prospect Dependencies with Bayesian Networks

Gabriele Martinelli
Norwegian University of Science and Technology, Trondheim, Norway

April 14, 2010,
AAPG 2010 Annual Convention, New Orleans, LA

Joint work with
Jo Eidsvik, Ragnar Hauge, Maren Drange Førland

Index

Motivations and goals

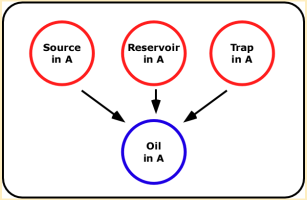
- How do prospect dependencies affect analysis of a geologically appealing area? [Van Wees et al, 2008], [Cunningham and Begg, 2008].
- Find the drilling sequence maximizing criteria of interest [Bickel and Smith, 2006].
- Formulate the problem of VoI [Bratvold et al., 2009], [Eidsvik et al., 2008] in the BN context and propose interpretation of the results

Motivations and goals

- How do prospect dependencies affect analysis of a geologically appealing area? [Van Wees et al, 2008], [Cunningham and Begg, 2008].
- Find the drilling sequence maximizing criteria of interest [Bickel and Smith, 2006].
- Formulate the problem of VoI [Bratvold et al., 2009], [Eidsvik et al., 2008] in the BN context and propose interpretation of the results

Model framework:

Three main components are required for oil and gas to accumulate in sufficient quantities to be worth producing: source, reservoir, and trap.

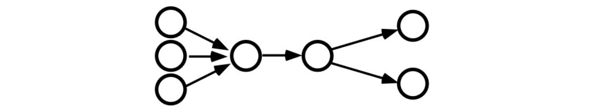


Bayesian Networks

Bayesian Network

A Bayesian network is a graphical model for probabilistic relationships among a set of variables.

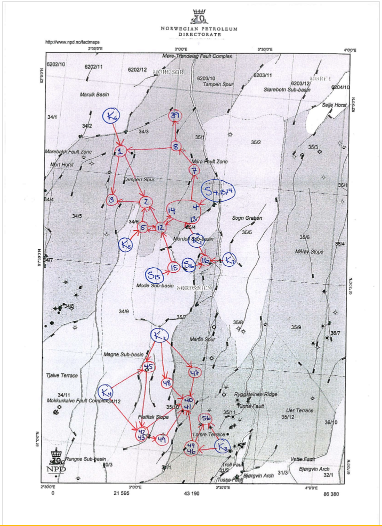
Example:

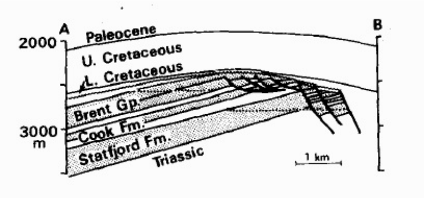


BN are directed acyclic graphs. Nodes represent variables, edges encode conditional dependencies between the variables.

Over the last decade, the Bayesian network has become a popular representation for encoding uncertain knowledge in expert systems.

The North Sea case study





- Several sites (nodes) have been chosen,
- Discrete model ⇒ three possible states for each node (*dry*, *gas* or *oil*).
- Three different models for source, reservoir and trap ⇒ We focus on the source model.

Encoding the geological information in the BN

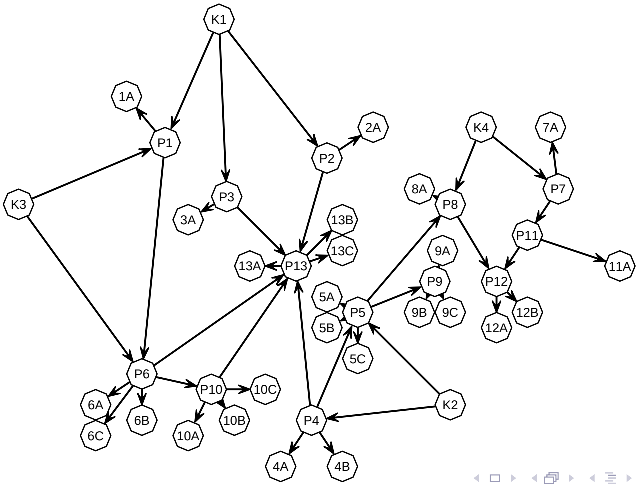
Qualitative part:

- List of nodes: either sites chosen for potential drilling or areas where the HC formation developed
- Possible outcomes at any segment
- Map of possible migration paths (edges)

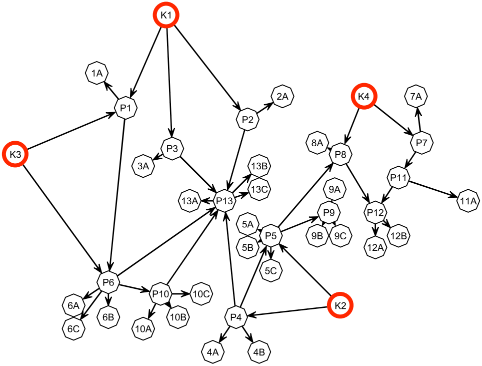
Quantitative part:

- Conditional probabilities: expert belief
- Expected a priori volumes
- Expected values of oil and gas and operational costs

The Bayesian Network of our case study



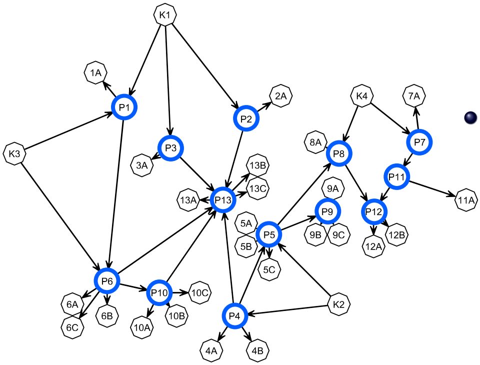
The nodes of the network



The model includes 3 possible kinds of nodes:

- **kitchens:** areas where source rock has reached appropriate conditions of pressure and temperature to generate hydrocarbons.

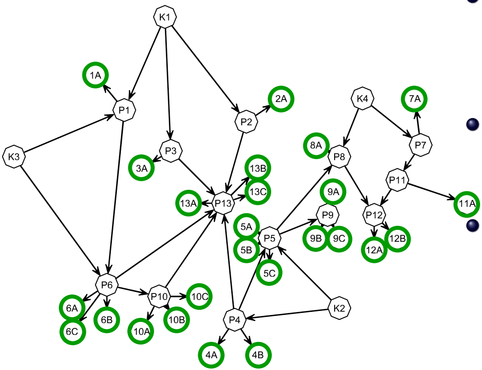
The nodes of the network



The model includes 3 possible kinds of nodes:

- **kitchens:** areas where source rock has reached appropriate conditions of pressure and temperature to generate hydrocarbons.
- **prospects:** key-nodes of the network. Through them we can define the spatial relationships that we need.

The nodes of the network



The model includes 3 possible kinds of nodes:

- **kitchens:** areas where source rock has reached appropriate conditions of pressure and temperature to generate hydrocarbons.
- **prospects:** key-nodes of the network. Through them we can define the spatial relationships that we need.
- **segments:** bottom nodes of our network. They correspond to physical 3D locations. Different segments are distinguished by:
  - Depth
  - Latitude and longitude
  - Relative position wrt a geological element (fault,...)

Model assumptions

1. The rules for the gas and oil flows:

v_i	$P(v_j = \mathrm{Dry}\|v_i)$	$P(v_j = \mathrm{Gas}\|v_i)$	$P(v_j = \mathrm{Oil}\|v_i)$
Dry	x	x	x
Gas	x	x	x
Oil	x	x	x

Model assumptions

1. The rules for the gas and oil flows:

v_i	$P(v_j = \mathrm{Dry}\|v_i)$	$P(v_j = \mathrm{Gas}\|v_i)$	$P(v_j = \mathrm{Oil}\|v_i)$
Dry	x	x	x
Gas	x	x	x
Oil	x	x	x

- Physical constraints
- Few parameters: $v_{ij}^{HC} = P(HC \text{ in } j | HC \text{ in } i)$, and $v_{ij}^{G} = P(v_j = \text{gas} | HC \text{ in } j \wedge \ HC \text{ in } i)$

Model assumptions

1. The rules for the gas and oil flows:

v_i	$P(v_j = \mathrm{Dry} \mid v_i)$	$P(v_j = \mathrm{Gas} \mid v_i)$	$P(v_j = \mathrm{Oil} \mid v_i)$
Dry	1	0	0
Gas	$1 - v_{ij}^{HC}$	$v_{ij}^{HC} * v_{ij}^{G}$	$v_{ij}^{HC} * (1 - v_{ij}^{G})$
Oil	$1 - v_{ij}^{HC}$	0	v_{ij}^{HC}

- Physical constraints
- Few parameters: $v_{ij}^{HC} = P(HC \text{ in } j | HC \text{ in } i)$, and $v_{ij}^{G} = P(v_j = \mathrm{gas} | HC \text{ in } j \wedge \ HC \text{ in } i)$

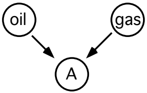
Model assumptions

1. The rules for the gas and oil flows:

v_i	$P(v_j = \mathrm{Dry} \mid v_i)$	$P(v_j = \mathrm{Gas} \mid v_i)$	$P(v_j = \mathrm{Oil} \mid v_i)$
Dry	1	0	0
Gas	$1 - v_{ij}^{HC}$	$v_{ij}^{HC} * v_{ij}^{G}$	$v_{ij}^{HC} * (1 - v_{ij}^{G})$
Oil	$1 - v_{ij}^{HC}$	0	v_{ij}^{HC}

- Physical constraints
- Few parameters: $v_{ij}^{HC} = P(HC \text{ in } j \mid HC \text{ in } i)$, and $v_{ij}^{G} = P(v_j = \mathrm{gas} \mid HC \text{ in } j \wedge \; HC \text{ in } i)$

2. Parents with different states:

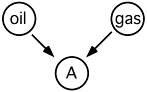


Model assumptions

1. The rules for the gas and oil flows:

v_i	$P(v_j = \mathrm{Dry} \mid v_i)$	$P(v_j = \mathrm{Gas} \mid v_i)$	$P(v_j = \mathrm{Oil} \mid v_i)$
Dry	1	0	0
Gas	$1 - v_{ij}^{HC}$	$v_{ij}^{HC} * v_{ij}^{G}$	$v_{ij}^{HC} * (1 - v_{ij}^{G})$
Oil	$1 - v_{ij}^{HC}$	0	v_{ij}^{HC}

- Physical constraints
- Few parameters: $v_{ij}^{HC} = P(HC \text{ in } j \mid HC \text{ in } i)$, and $v_{ij}^{G} = P(v_j = \text{gas} \mid HC \text{ in } j \wedge \ HC \text{ in } i)$

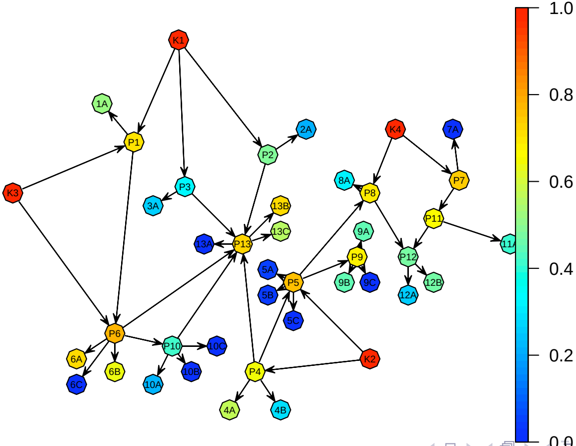


2. Parents with different states:

- Gas is lighter than oil. It forces itself into the pockets and squeezes the oil out of the trap. Assumptions of abundance and time are met.

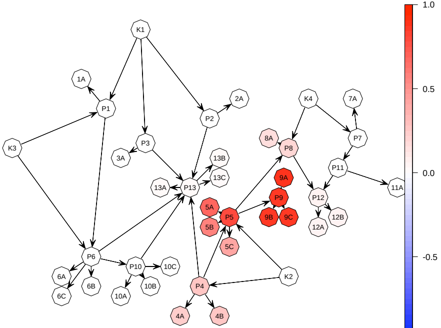
Marginal distribution of state $\{gas\}$

$P(v_j = gas)\ \forall\ j \in \{1, \ldots, 42\}$



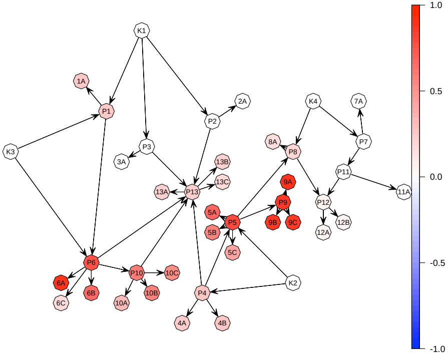
Single-site evidence

- What if we drill at segment i and get **evidence** v_i.
- We evaluate $P(v_j|v_i) - P(v_j)\ \forall\ i \in \{1, \ldots, 42\}$
- In the example, $\{9A\} = dry$



Multiple-site evidence

- What if we drill at segments i and j and get the **evidence** v_i, v_j.
- We evaluate $P(v_k|v_i, v_j) - P(v_k)\ \forall\ k \in \{1, \dots, 42\}$
- In the example, $\{\{9A\} = dry, \{6A\} = dry\}$



Choice of criterion

The optimal decision strategy depends on the choice of the criterion, and by the purpose that we want to achieve, whether it is a sequential strategy [Bickel and Smith, 2006] or a simultaneous one.

Choice of criterion

The optimal decision strategy depends on the choice of the criterion, and by the purpose that we want to achieve, whether it is a sequential strategy [Bickel and Smith, 2006] or a simultaneous one.

Our criteria:

- Which is the best combination of segments in order to maximize the total revenues?
- Which are the best segments to drill in order to give us the maximum confidence that the area is dry?

Choice of criterion

The optimal decision strategy depends on the choice of the criterion, and by the purpose that we want to achieve, whether it is a sequential strategy [Bickel and Smith, 2006] or a simultaneous one.

Our criteria:

- Which is the best combination of segments in order to maximize the total revenues? $\Rightarrow$ Value of Information (VoI) [Bratvold et al., 2009] and [Bhattacharjya et al., 2010]
- Which are the best segments to drill in order to give us the maximum confidence that the area is dry? $\Rightarrow$ Abandoned Revenue (AR)

Technical details

For each segment, we have following data:

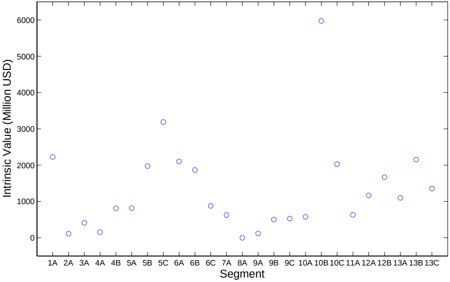
- Expected gas and oil volumes
- Assumed gas and oil values
- Fixed costs (includes the infrastructures and the well costs)
- Variable cost, proportional to the expected volume of gas and oil (includes the operation and the maintenance costs)

Technical details

These quantities allow us to compute:

- PRO_i: Partial Revenues for Oil, associated with segment i
- PRG_i: Partial Revenues for Gas, associated with segment i
- EFC: Exploratory Fixed Costs
- PFC: Prospect Fixed Costs

Intrinsic Value



The Value of Information

Idea:

- Select all the possible combinations of the segments of size p.
- Compute all the possible combinations of evidence (3^p), their prior probability of occurrence, and the conditional for the other segments.
- Include in the expected revenues with evidence, the segments having a balance between costs and revenues times the probability of HC higher than a fixed threshold (*pre-posterior* value).
- Find the VoI by subtracting the expected revenues without evidence

$$\mathrm{ER}(i) = \sum_{\text{evidence } j} \left[\sum_{\text{prospect } k} \max\{\mathrm{Rev}_{kj} - \mathrm{PFC}, 0\} \right] P(v_i = j),$$

where

$$\mathrm{Rev}_{kj} = \sum_{\text{segment } l \in k} \left\{ \mathrm{PRO}_l * P(v_l = oil | v_i = j) + \mathrm{PRG}_l * P(v_l = gas | v_i = j) - \mathrm{EFC} \right\}.$$

.

The Value of Information

Idea:

- Select all the possible combinations of the segments of size p.
- Compute all the possible combinations of evidence (3^p), their prior probability of occurrence, and the conditional for the other segments.
- Include in the expected revenues with evidence, the segments having a balance between costs and revenues times the probability of HC higher than a fixed threshold (*pre-posterior* value).
- Find the VoI by subtracting the expected revenues without evidence

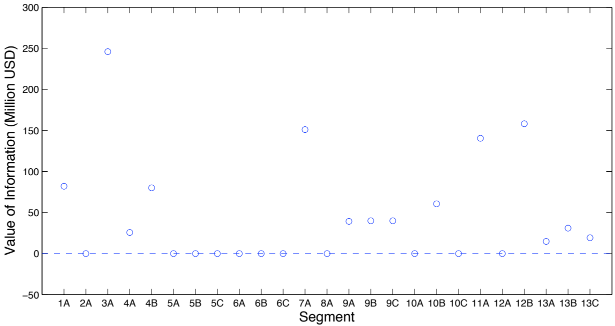
$$VoI(i) = ER(i) - \sum_{\text{prospect } k} \max\{\text{Rev}_k - \text{PFC}, 0\},$$

where

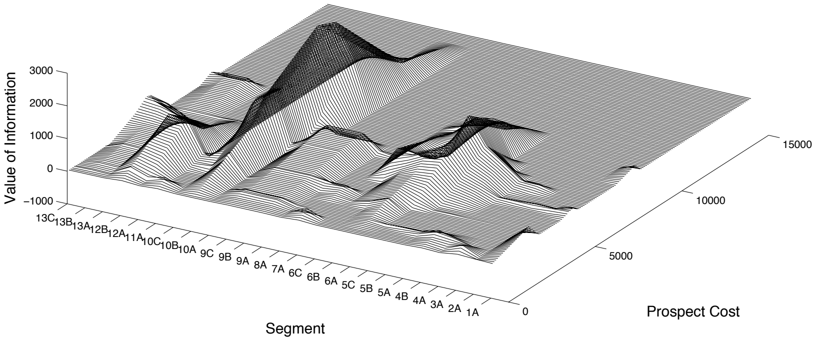
$$\text{Rev}_k = \sum_{\text{segment } l \in k} \Big\{\text{PRO}_l * P(v_l = oil) + \text{PRG}_l * P(v_l = gas) - \text{EFC}\Big\}.$$

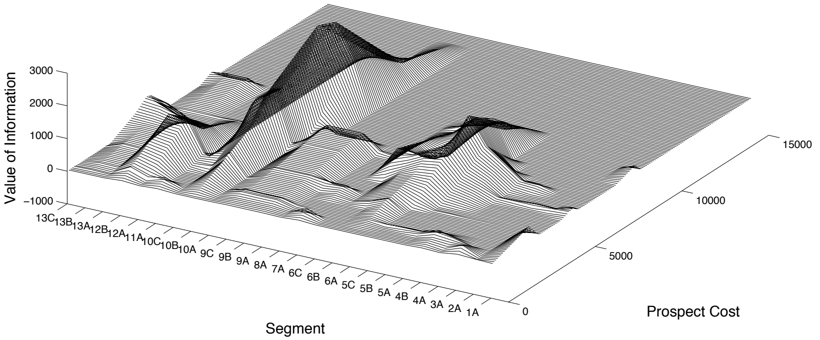
.

The Value of Information, $p = 1$



- Little correlation between the intrinsic values and the VoI
- A particular segment has a high VoI if its impact is able to change the decision that we would make


Value of Information
3000
2000
1000
0
−1000
13C 13B 13A 12B 12A 11A 10C 10B 10A 9C 9B 9A 8A 7A 6C 6B 6A 5C 5B 5A 4B 4A 3A 2A 1A
Segment
0
5000
10000
15000
Prospect Cost



- Similar behavior for all segments: VoI increases, then decreases.
- VoI is not unimodal.
- Continuity of behavior among segments belonging to the same prospect.

The Abandoned Revenue

Idea:

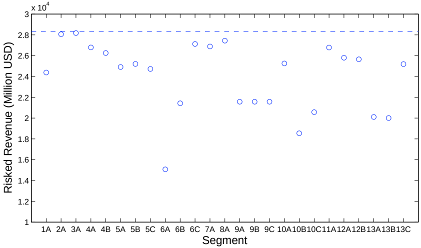
- Select segments that, if dry, give the maximum confidence about the scarce appealing of the area.
- Increase the dimension p, in order to find the joint best combination of segments

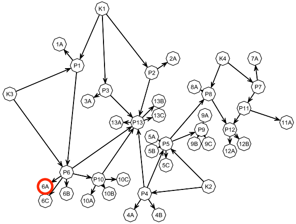
$$AR(i) = \sum_{\text{prospect } k} \max\{\mathrm{Rev}_k - \mathrm{PFC}, 0\},$$

where

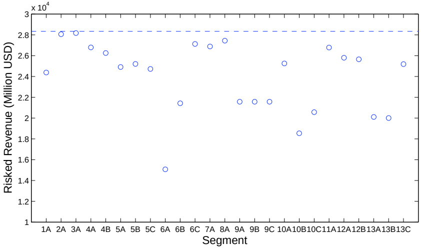
$$\mathrm{Rev}_k = \sum_{\text{segment } l \in k} \left\{ \mathrm{PRO}_l * P(v_l = oil | v_i = dry) + \mathrm{PRG}_l * P(v_l = gas | v_i = dry) - EFC \right\}$$

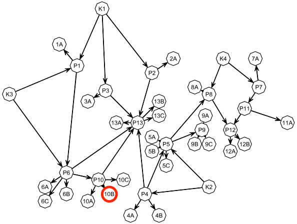
The Abandoned Revenue



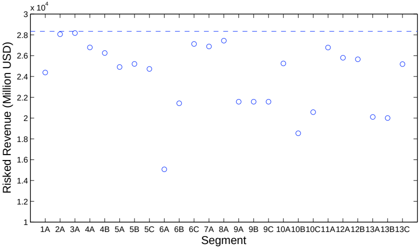


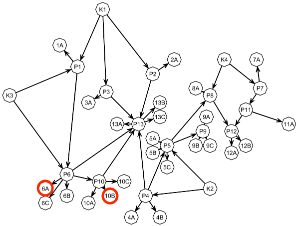
The Abandoned Revenue

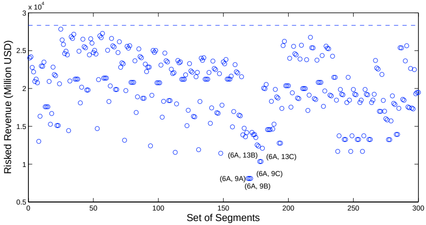


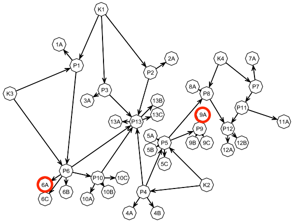


The Abandoned Revenue









Closing remarks

Results achieved:

1. Propose a framework for evaluating and studying prospect dependencies , while maintaining statistical computational benefits
2. Indicate a guideline for interpreting the results coming from two indices, VoI and AR.

Main bottlenecks:

1. The definition of the network and the specification of the parameters
2. The evaluation criterion to choose: based on different utility functions (higher moments of the distribution).
3. Combinatorial problems: these criteria run very fast, but more complex criteria, that involve dynamic programming, have exponential complexity in the number of segments.

Closing remarks

Current work:

1. Understanding and improving in our case the Junction Tree Algorithm, that allows the propagation of the evidence through the network.
2. Proposing different criteria and a strategy for sequential planning rather than simultaneous planning [Bickel and Smith, 2006].
3. Work under the assumption of imperfect information [Eidsvik et al., 2008]
4. Evaluating the sensitivity to the network's structure.
5. Overcoming some model limitations (discreteness, elicitation of parameters,...)

Acknowledgements:

- We thank Statoil for providing the expertise and the data necessary to perform this study.
- We thank the Statistics for Innovation (SFI^2) research center in Oslo, that partially financed this work through the "FindOil" project.

References

- Bickel J.E, J.E. Smith, 2006, Optimal Sequential Exploration: A Binary Learning Model, *Decision Analysis, Vol. 3, No. 1*
- Bratvold R.B., J.E. Bickel, H.L. Lohne, 2009, Value of Information in the Oil and Gas Industry: Past, Present, and Future, *SPE Res Eval & Eng, Vol. 12, No. 4*
- Cunningham P., S. H. Begg, 2008, Using the value of information to determine optimal well order in a sequential drilling program, *AAPG Bulletin, Vol. 92, No. 10*
- Eidsvik J., Bhattacharjya D., and Mukerji, T., 2008, Value of information of seismic amplitude and CSEM resistivity, *Geophysics, Vol. 73, No. 4, R59-R69*
- Lauritzen S. L., D. J. Spiegelhalter, 1988, Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems, *Journal of the Royal Statistical Society, Series B 50, 157-224*
- Van Wees J.D, Mijnlieff H., Lutgert J., Breunese J., Bos C., Rosenkranz P. and Neele F., 2008, A Bayesian belief network approach for assessing the impact of exploration prospect interdependency: An application to predict gas discoveries in the Netherlands *AAPG Bulletin, Vol. 92, No. 10*