Value of Information in the Earth Sciences

Jo Eidsvik,
Department of Mathematical Sciences, NTNU, NORWAY

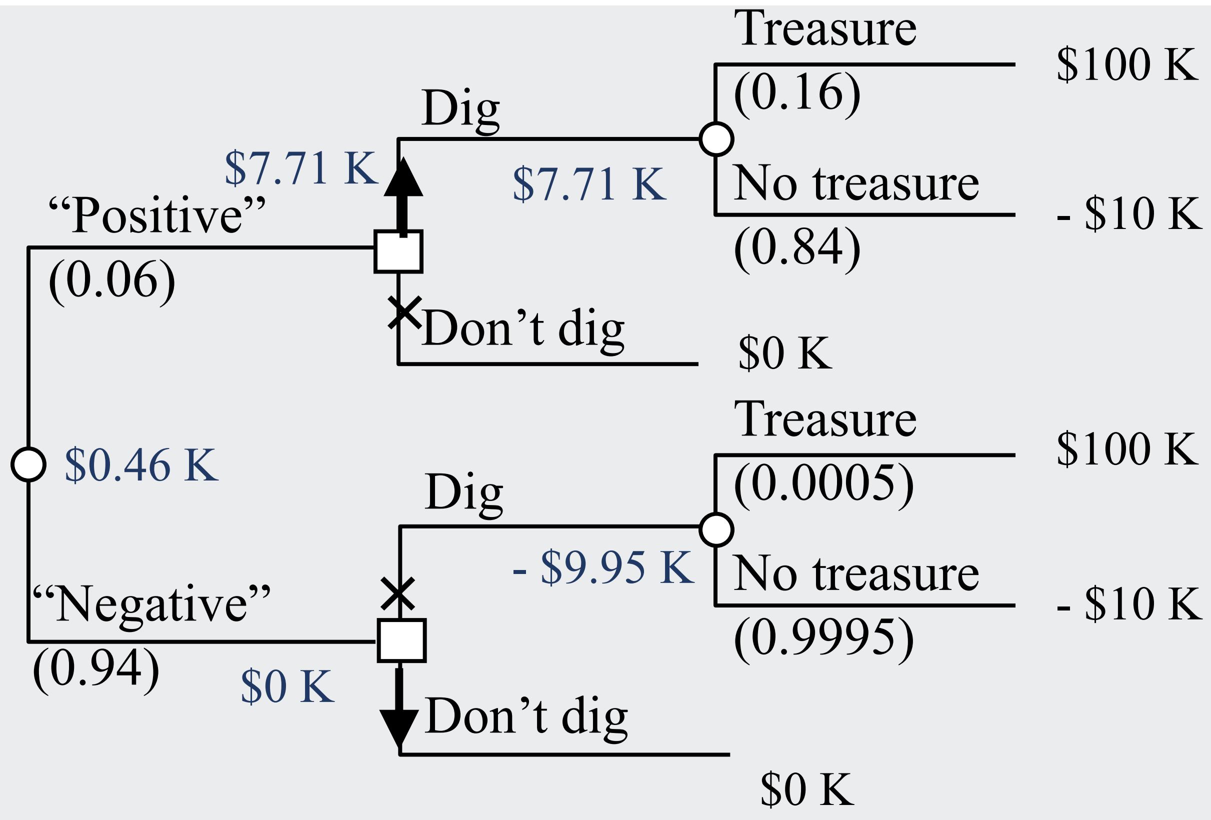
Plan for course

Time	Topic
Day 1	Introduction and motivating examples
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Day 2	**Examples of value of information analysis in Earth sciences**
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.

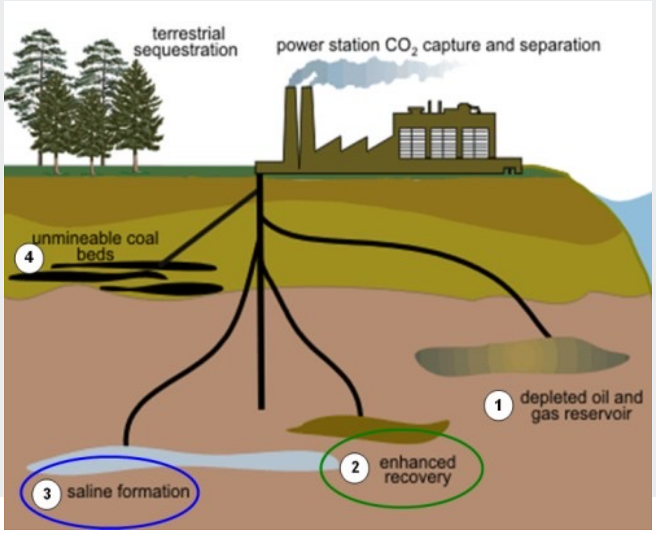
Pirate example
(For motivating decision analysis and VOI)

PoV - imperfect information

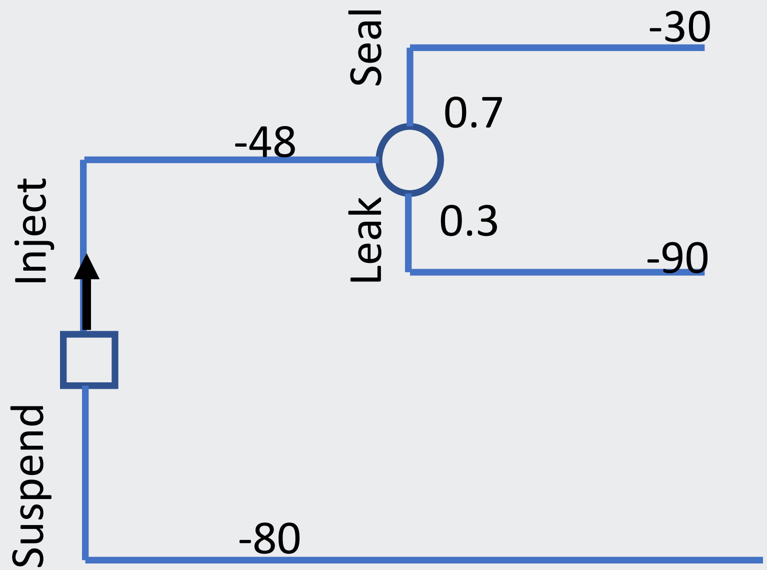


Problem: CO2 sequestration

CO2 is sequestered to reduce carbon emission in the athmosphere and defer global warming.
Geological sequestration involves pumping CO2 in subsurface layers, where it will remain, unless it leaks to the surface.



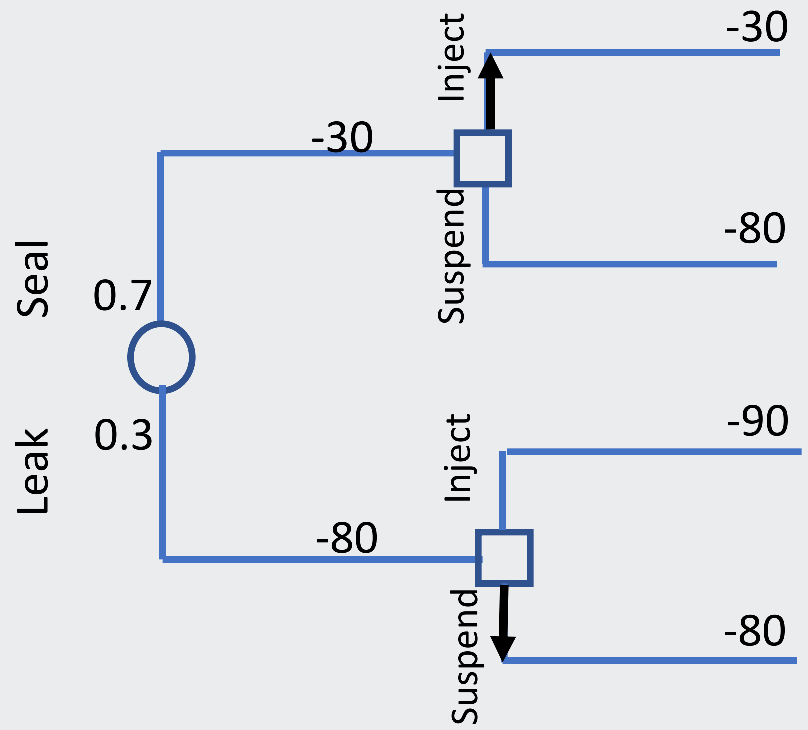
PV



Expected value of Injection alternative : 0.7 (- 30) + 0.3 (- 90) = -21 – 27 = - 48

PV = max (- 80 , - 48) = -48

VOI of Perfect information



max (- 30 , - 80) = - 30

max (- 90 , - 80) = - 80

PoV(x) = 0.7 (- 30) + 0.3 (- 80) = -21 - 24 = - 45

VOI (x) = -45 - (- 48) = 3

Numbers for imperfect information

The decision maker can proceed with CO_2 injection or suspend sequestration. The latter incurs a tax of 80 monetary units. The former only has a cost of injection equal to 30 monetary units, but the injected CO_2 may leak (x=1). If leakage occurs, there will be a fine of 60 monetary units (i.e. a cost of 90 in total). Decision maker is risk neutral.

$$p(x=1)=0.3 \qquad p(x=0)=0.7$$

Data: Geophysical experiment, with binary outcome, indicating whether the formation is leaking or not.

$$p(y=0 \mid x=0)=0.95 \qquad p(y=1 \mid x=1)=0.9$$

Marginalization:

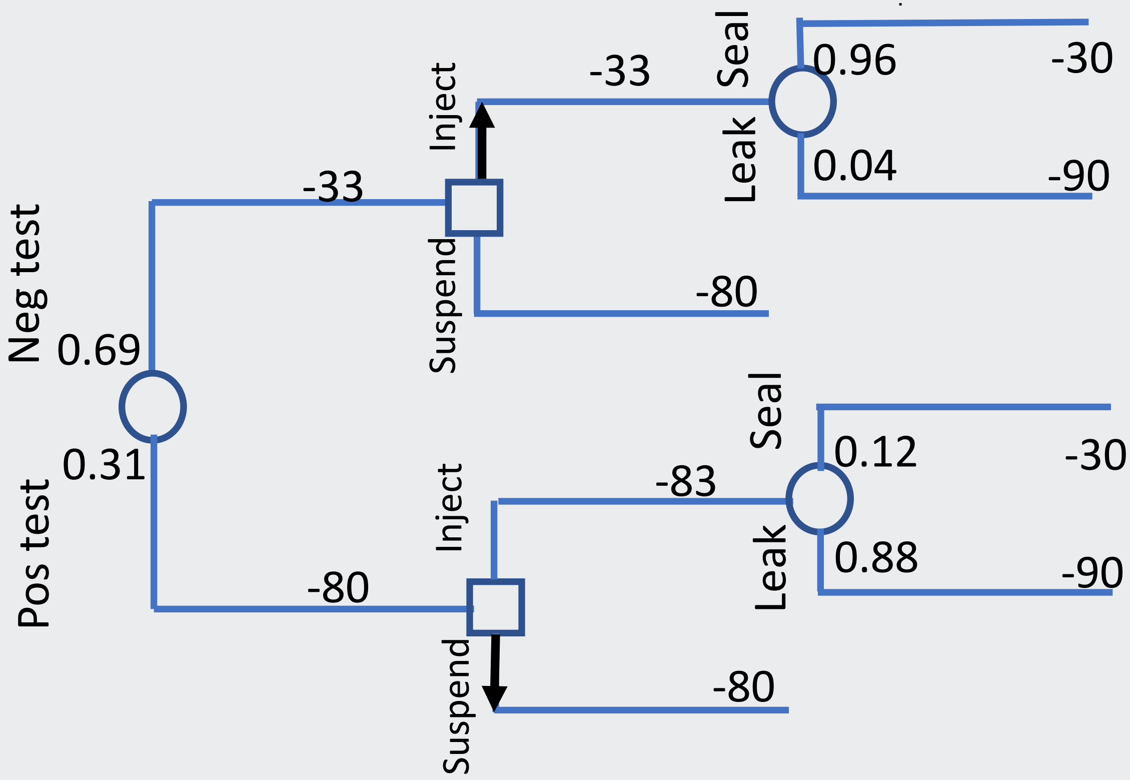
P(pos test) = P (pos test | leak) P (leak) + P (post test | seal) P (seal) = 0.9 (0.3) + 0.05 (0.7) = 0.31

Bayes' rule:

P(leak | pos test) = P (pos test | leak) P (leak) / P (pos test) = 0.9 (0.3) / 0.31 = 0.88

P(leak | neg test) = P (neg test | leak) P (leak) / P (neg test) = 0.1 (0.3) / 0.69 = 0.04

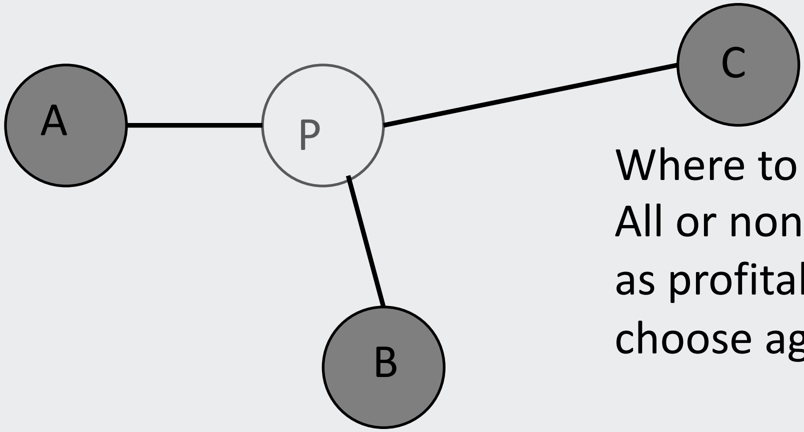
VOI of Imperfect information



PoV(y) = 0.69 (-33) +0. 31 (-80) = - 47.

VOI (y) = -47 - (- 48) = 1

What if several projects / treasures?

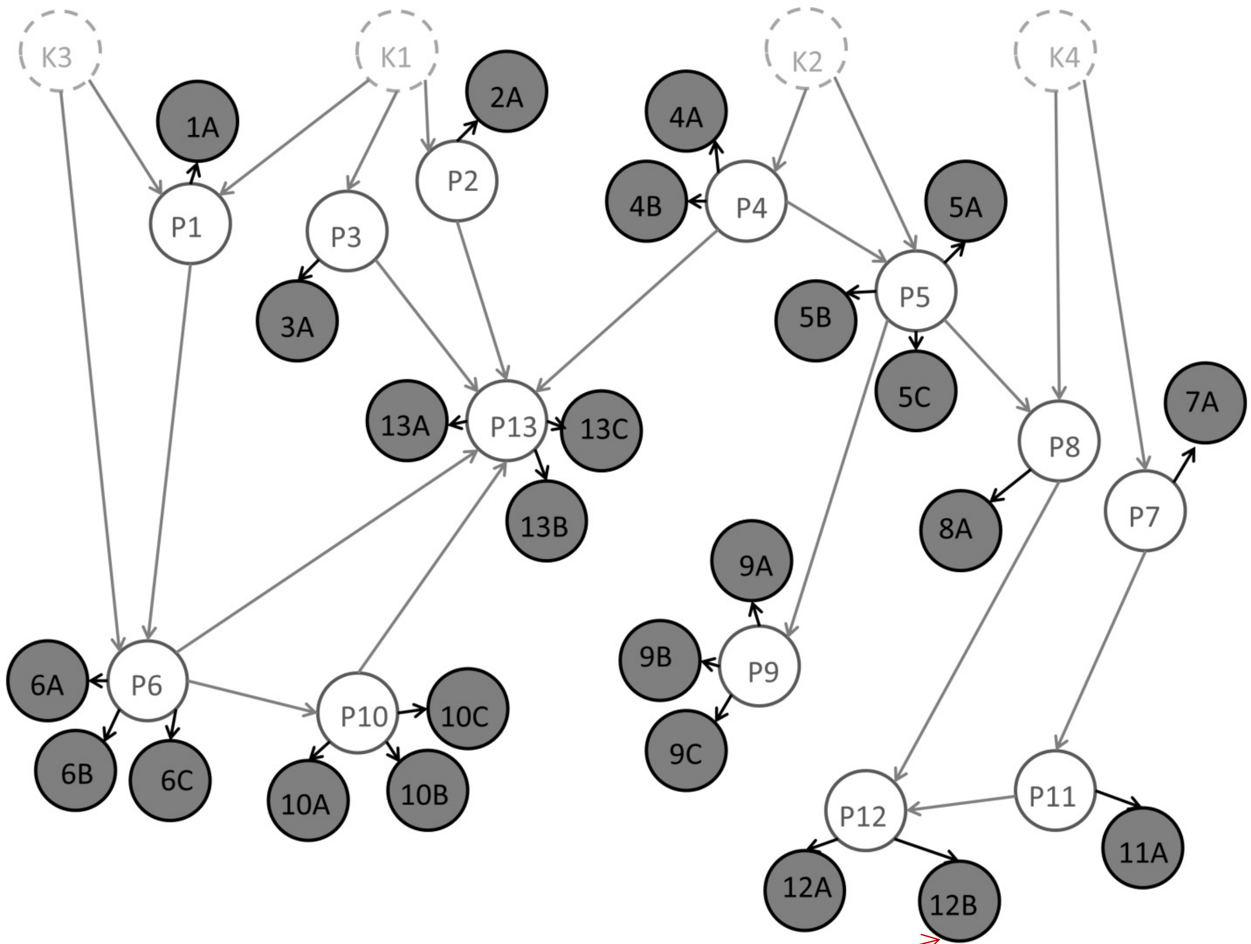


Where to invest?
All or none? Free to choose as many as profitable? One at a time, then choose again?

Where should one collect data? All or none? One only? Or two? One first, then maybe another?

Dependence? Does it matter?

Gray nodes are petroleum reservoir segments where the company aims to develop profitable amounts of oil and gas.



Drill the exploration well at this segment! The value of information is largest.

VOI and Earth sciences

- **Alternatives are spatial**, often with high flexibiliy in selection of sites, control rates, intervention, excavation opportunities, harvesting, etc.

- **Uncertainties are spatial**, with multi-variable interactions . Often both discrete and continuous.

- **Value function is spatial**, typically involving coupled features, say through differential equations. It can be defined by «physics» as well as economic attributes.

- **Data are spatial**. There are plenty opportunities for partial, total testing and a variety of tests (surveys, monitoring sensors, electromagnetic data, , etc.)

Decision situations and values

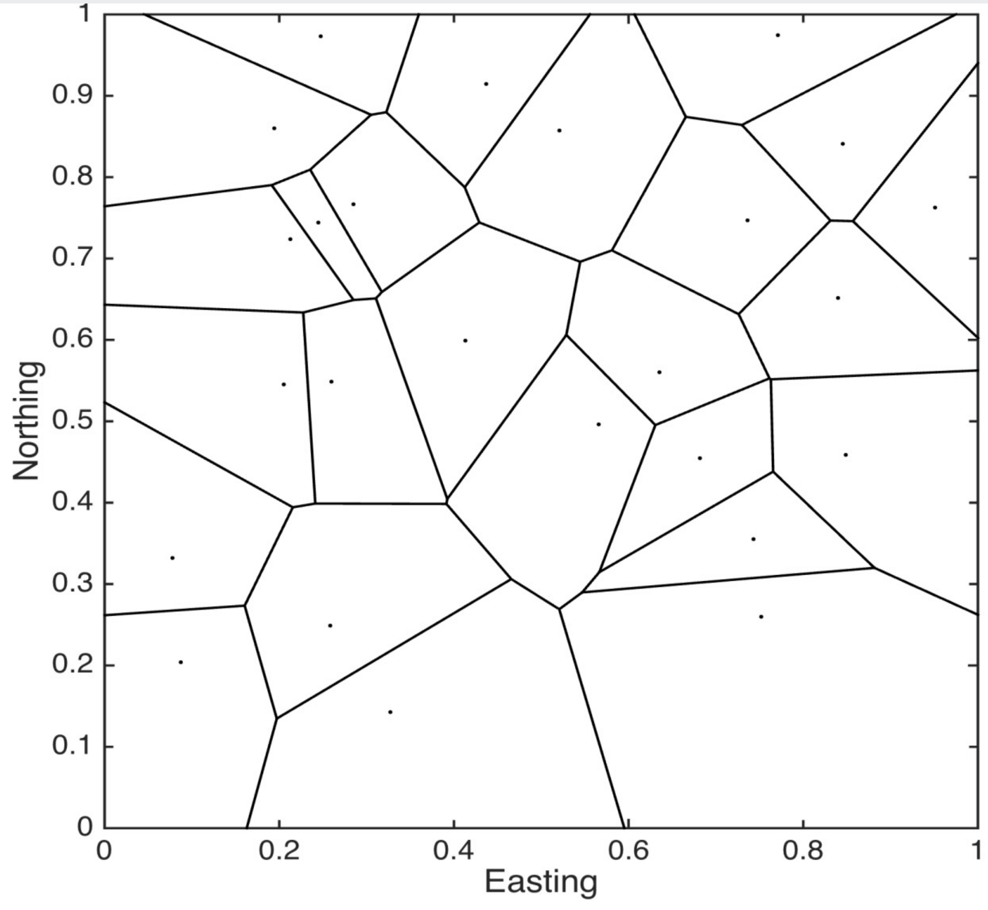
.

	Assumption: Decision Flexibility	Assumption: Value Function
Low decision flexibility; Decoupled value	Alternatives are easily enumerated $$a \in A$$	Total value is a sum of value at every unit $$v(\boldsymbol{x}, a) = \sum_j v(x_j, a)$$
High decision flexibility; Decoupled value	None $$\boldsymbol{a} \in \boldsymbol{A}$$	Total value is a sum of value at every unit $$v(\boldsymbol{x}, \boldsymbol{a}) = \sum_j v(x_j, a_j)$$
Low decision flexibility; Coupled value	Alternatives are easily enumerated $$a \in A$$	None $$v(\boldsymbol{x}, a)$$
High decision flexibility; Coupled value	None $$\boldsymbol{a} \in \boldsymbol{A}$$	None $$v(\boldsymbol{x}, \boldsymbol{a})$$

Low versus high decision flexibility

High flexibility:
Farmer can select individual forest units.

Low flexibility:
Farmer must select all forest units, or none.



Problem: Two projects

High flexibility:
Farmer can select individual forest units.

$$PV = \sum_{i=1}^{2} \max(0, \mu_i)$$

Low flexibility:
Farmer must select all forest units, or none.

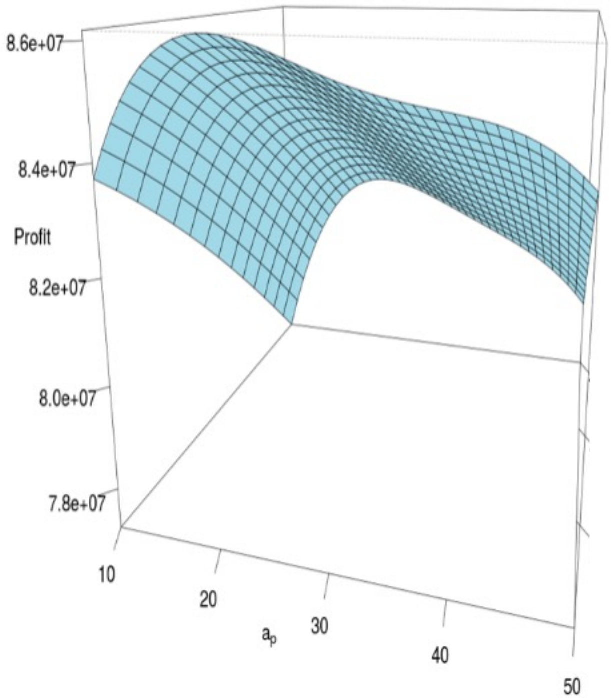
$$PV = \max(0, \sum_{i=1}^{2} \mu_i)$$

Problem:

1. **What's the PV for different mean values?**

Flexibility is related to framing the decision situation

How many options are there for feasible operation?
How can one break this down into something that can be interpreted?



Operations of off-shore windfarm: Number of personell, number of ship type A, shiup type B.

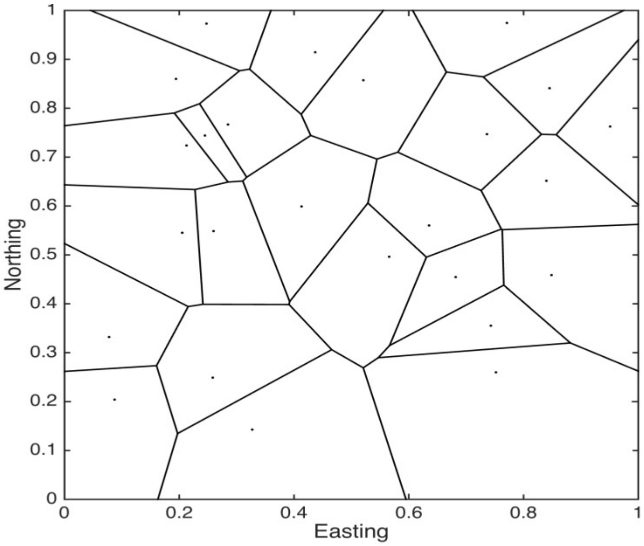
$$a_p \in \mathcal{A}_p = \{10, 11, \ldots, 50\}, \;\; a_{CTV} \in \mathcal{A}_{CTV} = \{0, 1, \ldots, 4\}, \;\; a_{SES} \in \mathcal{A}_{SES} = \{0, 1, \ldots, 4\}.$$

Decoupled versus coupled value

.

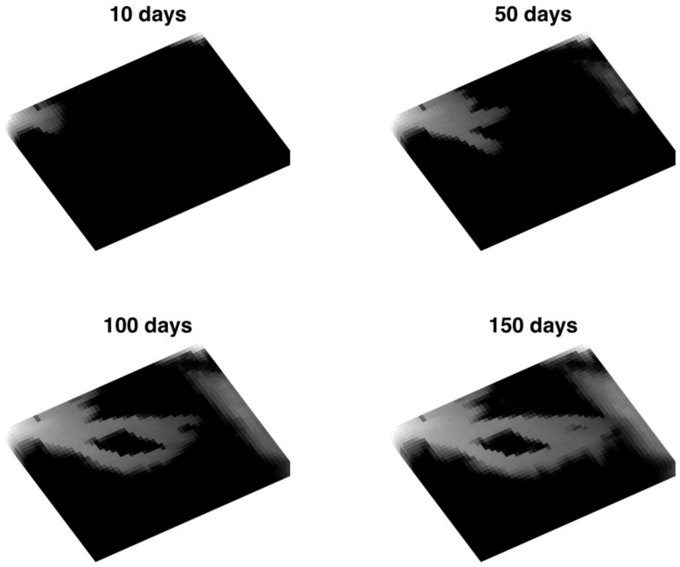
Farmer must decide whether to harvest at forest units, or not.

Value decouples to sum over units.



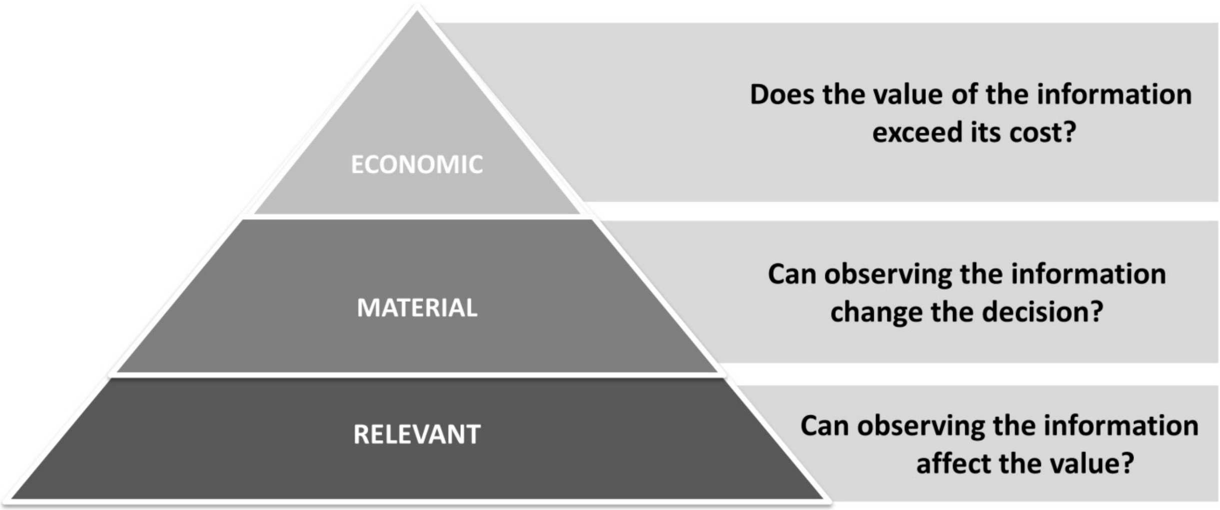
Petroleum company must decide how to produce a reservoir.

Value involves complex coupling of drilling strategies, and reservoir properties.



EAGE EDUCATION

VOI - Pyramid of conditions



Information gathering

.

	Perfect	Imperfect
Total	Exact observations are gathered for all locations. This is rare, occurring when there is extensive coverage and highly accurate data gathering. $$\boldsymbol{y} = \boldsymbol{x}$$	Noisy observations are gathered for all locations. This is common in situations with remote sensors with extensive coverage, e.g. seismic, radar, satellite data. $$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\varepsilon}$$
Partial	Exact observations are gathered at some locations. This might occur, for instance, when there is careful analysis of rock samples along boreholes in a reservoir or a mine. $$\boldsymbol{y}_{\mathrm{K}} = \boldsymbol{x}_{\mathrm{K}}, \qquad \mathrm{K} \ \text{subset}$$	Noisy observations are gathered at some locations. Examples include hand-held (noisy) meters to observe grades in mine boreholes, electromagnetic testing along a line, biological surveys of species, etc. $$\boldsymbol{y}_{\mathrm{K}} = \boldsymbol{x}_{\mathrm{K}} + \boldsymbol{\varepsilon}_{\mathrm{K}}, \qquad \mathrm{K} \ \text{subset}$$

EAGE
EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Problem: Two projects

.

High flexibility:
Farmer can select individual forest units.

$$PV = \sum_{i=1}^{2} \max(0, \mu_i)$$

Low flexibility:
Farmer must select all forest units, or none.

$$PV = \max(0, \sum_{i=1}^{2} \mu_i)$$

Problem:

1. **What's the PV for different mean values?**
2. **How is this influencing the VOI of perfect information on both projects?**

Recall VOI for Gaussian

Gaussian cdf

Gaussian pdf

Result:

$$VOI(x) = m\Phi\left(\frac{m}{r}\right) + r\phi\left(\frac{m}{r}\right) - \max\{0, m\}$$

The analytical form facilitates computing, and it eases the study of VOI properties as a function of the parameters.

$$m = 0,$$

$$VOI(x) = r\phi(0) = \frac{r}{\sqrt{2\pi}}$$

Formula for VOI

$$PV = \max_{a \in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a})\right)\right\} = \max_{a \in A}\left\{\int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a})\, p\left(\boldsymbol{x}\right) d\boldsymbol{x}\right\}$$

$$PoV\left(\boldsymbol{y}\right) = \int \max_{a \in A}\left\{E\left(v\left(\boldsymbol{x},\boldsymbol{a}\right) \mid \boldsymbol{y}\right)\right\} p\left(\boldsymbol{y}\right) d\boldsymbol{y}$$

$$VOI\left(\boldsymbol{y}\right) = PoV\left(\boldsymbol{y}\right) - PV.$$

Formula for VOI

$$PV = \max_{a \in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a})\right)\right\} = \max_{a \in A}\left\{\int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a})\,p(\boldsymbol{x})\,d\boldsymbol{x}\right\}$$

$$PoV(\boldsymbol{y}) = \int \max_{a \in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a}) \mid \boldsymbol{y}\right)\right\} p(\boldsymbol{y})\,d\boldsymbol{y}$$

$$VOI(\boldsymbol{y}) = PoV(\boldsymbol{y}) - PV.$$

Spatial alternatives.

Spatial uncertainties.

Spatial value function.

Spatial data.

Computation - Formula for VOI

VOI=Posterior value – Prior value

$$PV = \max_{\boldsymbol{a}\in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a})\right)\right\} = \max_{\boldsymbol{a}\in A}\left\{\int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a})\,p(\boldsymbol{x})\,d\boldsymbol{x}\right\}$$

$$PoV(\boldsymbol{y}) = \int \max_{\boldsymbol{a}\in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a}) \mid \boldsymbol{y}\right)\right\} p(\boldsymbol{y})\,d\boldsymbol{y}$$

Computations :

- Easier with low decision flexibility (less alternatives).
- Easier if value decouples (sums or integrals split).
- Easier for perfect, total, information (upper bound on VOI).
- Sometimes analytical solutions. Otherwise approximations and Monte Carlo.

Formula for total perfect information

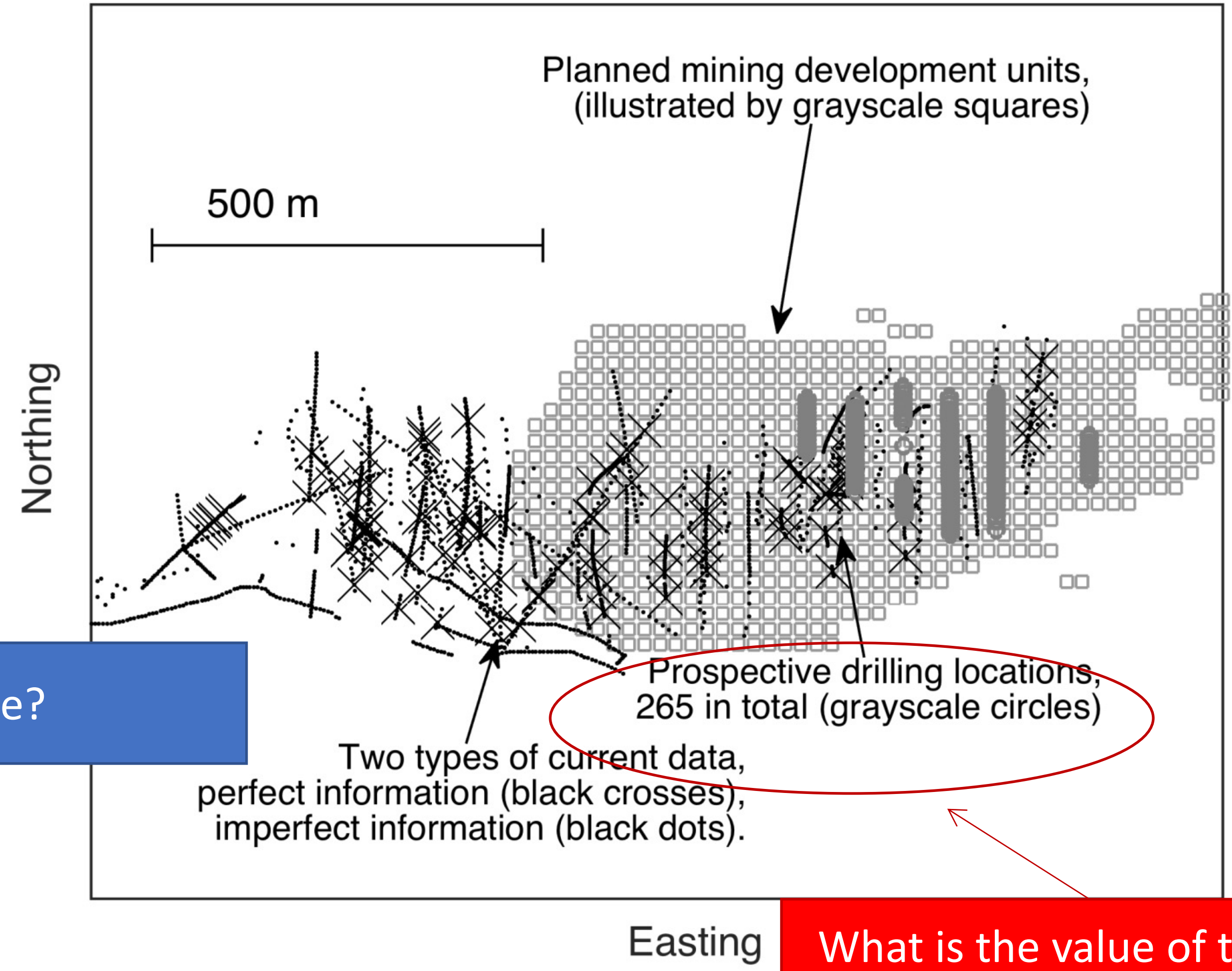
$$PV = \max_{a \in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a})\right)\right\} = \max_{a \in A}\left\{\int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a})\,p(\boldsymbol{x})\,d\boldsymbol{x}\right\}$$

$$PoV(\boldsymbol{x}) = \int \max_{a \in A}\left\{v(\boldsymbol{x},\boldsymbol{a})\right\} p(\boldsymbol{x})\,d\boldsymbol{x}$$

$$VOI(\boldsymbol{x}) = PoV(\boldsymbol{x}) - PV.$$

Upper bound on any information gathering scheme.

Low flexibility. Decoupled value

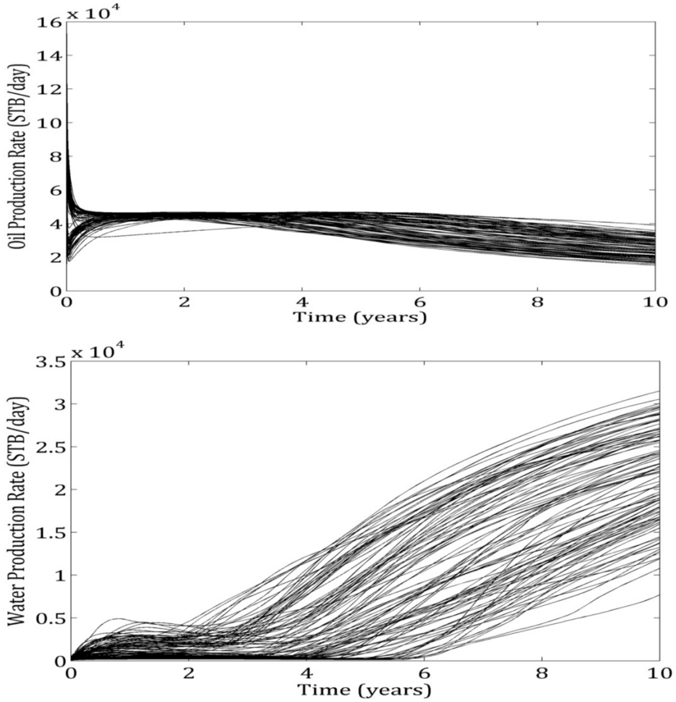


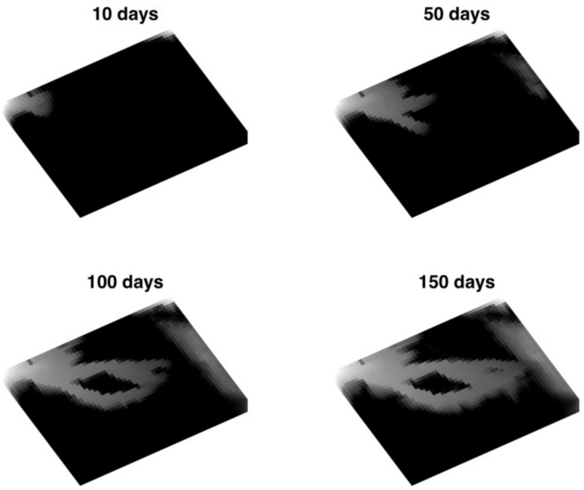
Is mining profitable?

What is the value of this additional information?

Higher flexibility. Coupled value

Infill drilling?





Is time lapse seismic data valuable?

Plan for course

Time	Topic
Day 1	Introduction and motivating examples
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Day 2	Examples of value of information analysis in Earth sciences
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.

Joint pdfs

There are families of joint **pdfs**. Parametrically, or non-parametrically.

Gaussian distribution is very common:

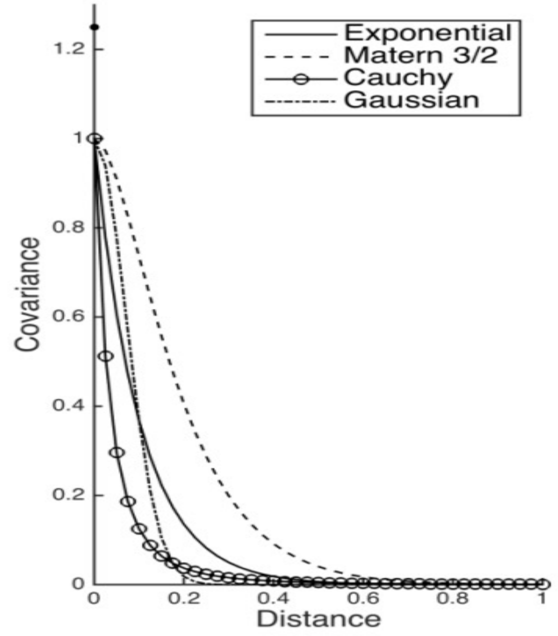
$$p(\boldsymbol{x}) = N(\boldsymbol{0},\boldsymbol{\Sigma}), \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & .. & \Sigma_{21} \\ .. & .. & .. & .. \\ \Sigma_{n1} & \Sigma_{n2} & .. & \Sigma_{nn} \end{pmatrix}$$

For a Gaussian process, in a spatial application, the covariance entries are formed in a particular way.

Spatial covariance functions

$$p(x) = N(\mu, \Sigma), \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & .. & \Sigma_{21} \\ .. & .. & .. & .. \\ \Sigma_{n1} & \Sigma_{n2} & .. & \Sigma_{nn} \end{pmatrix}$$

$$\Sigma_{ij} = \Sigma(|\mathbf{s}_i - \mathbf{s}_j|) = \Sigma(|\mathbf{t}|)$$



Model	Covariance
Exponential	$\Sigma(\lvert\mathbf{t}\rvert) = \sigma^2 \exp(-\eta\lvert\mathbf{t}\rvert)$
Matern 3/2	$\Sigma(\lvert\mathbf{t}\rvert) = \sigma^2 (1+\eta\lvert\mathbf{t}\rvert)\exp(-\eta\lvert\mathbf{t}\rvert)$
Cauchy-type	$\Sigma(\lvert\mathbf{t}\rvert) = \sigma^2 \frac{1}{(1+\eta\lvert\mathbf{t}\rvert)^3}$
Gaussian	$\Sigma(\lvert\mathbf{t}\rvert) = \sigma^2 \exp(-\eta^2\lvert\mathbf{t}\rvert^2)$

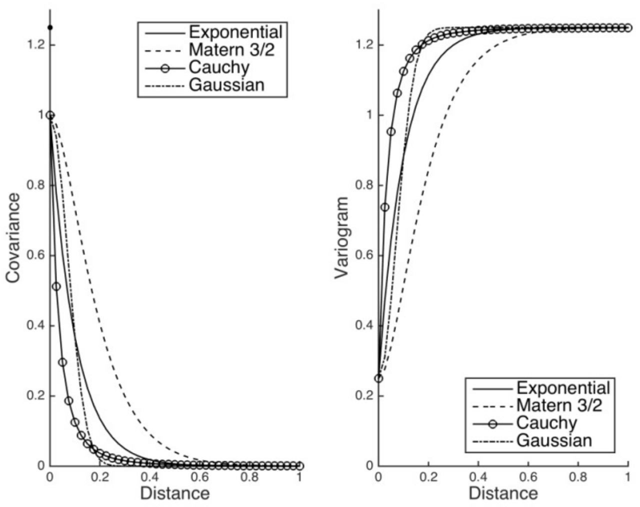
EAGE EDUCATION

Example - Gaussian process

$$p(\boldsymbol{x}) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \qquad \boldsymbol{y} = \boldsymbol{F}\boldsymbol{x} + N(\boldsymbol{0}, \tau^2 \boldsymbol{I})$$

$$p(\boldsymbol{y}) = N(\boldsymbol{F}\boldsymbol{\mu}, \boldsymbol{C}) \qquad \boldsymbol{C} = \boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}^t + \tau^2 \boldsymbol{I}$$

Design matrix: $\boldsymbol{F}$



Model	Covariance
Exponential	$C(\lvert\mathbf{t}\rvert) = \tau^2 I(\lvert\mathbf{t}\rvert = 0) + \sigma^2 \exp(-\eta \lvert\mathbf{t}\rvert)$
Matern 3/2	$C(\lvert\mathbf{t}\rvert) = \tau^2 I(\lvert\mathbf{t}\rvert = 0) + \sigma^2 (1 + \eta \lvert\mathbf{t}\rvert) \exp(-\eta \lvert\mathbf{t}\rvert)$
Cauchy-type	$C(\lvert\mathbf{t}\rvert) = \tau^2 I(\lvert\mathbf{t}\rvert = 0) + \sigma^2 \frac{1}{(1 + \eta \lvert\mathbf{t}\rvert)^3}$
Gaussian	$C(\lvert\mathbf{t}\rvert) = \tau^2 I(\lvert\mathbf{t}\rvert = 0) + \sigma^2 \exp(-\eta^2 \lvert\mathbf{t}\rvert^2)$

Gaussian process - model

$$p(\boldsymbol{x}) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \qquad \boldsymbol{y} = \boldsymbol{F}\boldsymbol{x} + N\left(\boldsymbol{0}, \tau^2 \boldsymbol{I}\right)$$

$$p(\boldsymbol{y}) = N(\boldsymbol{F}\boldsymbol{\mu}, \boldsymbol{C}) \qquad \boldsymbol{C} = \boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}^t + \tau^2 \boldsymbol{I}$$

Bayes' rule

Goal is:

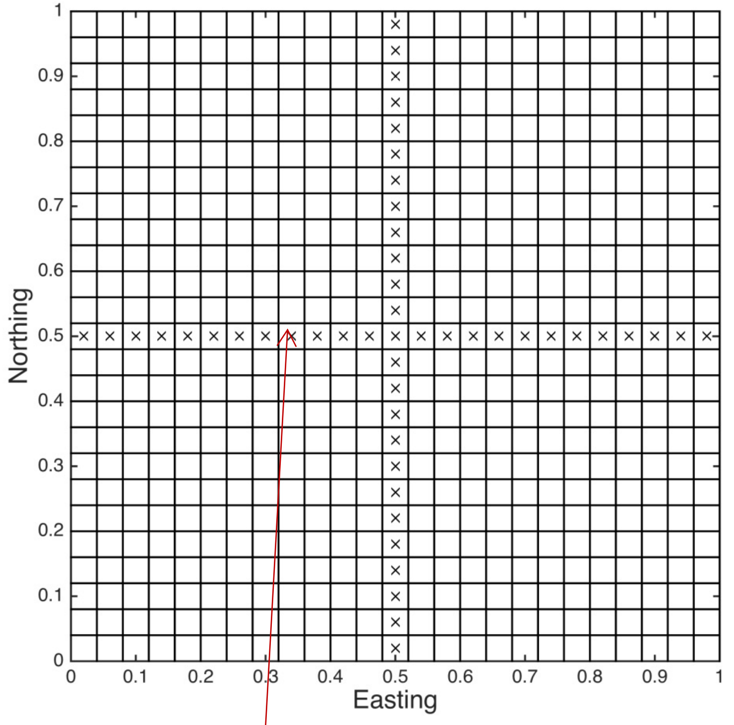
$$p(\boldsymbol{x} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid \boldsymbol{x})\, p(\boldsymbol{x})}{p(\boldsymbol{y})}$$

$$p(\boldsymbol{x} \mid \boldsymbol{y}) = N\left(\boldsymbol{\mu} + \boldsymbol{\Sigma}\boldsymbol{F}^t \left(\boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}^t + \tau^2 \boldsymbol{I}\right)^{-1} (\boldsymbol{y} - \boldsymbol{F}\boldsymbol{\mu}), \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\boldsymbol{F}^t \left(\boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}^t + \tau^2 \boldsymbol{I}\right)^{-1} \boldsymbol{F}\boldsymbol{\Sigma}\right)$$

Norwegian wood - forestry example

Farmer must decide whether to harvest forest, or not. There is uncertainty about timber volumes and profits over the spatial domain.

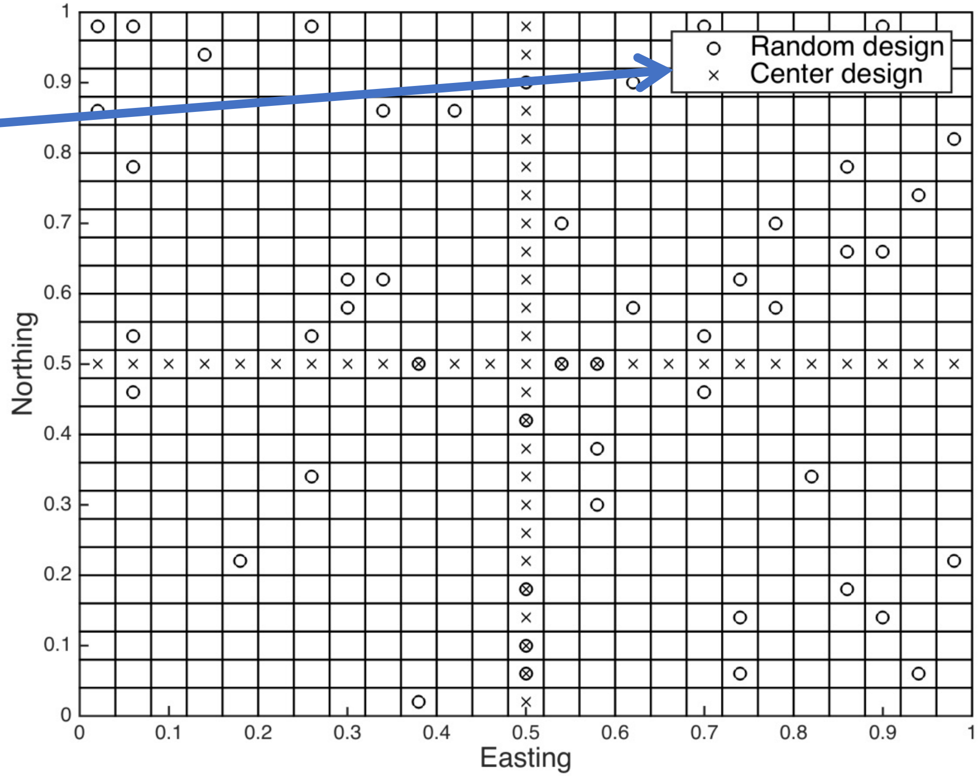
Another decision is whether to collect data before making these decisions. If so, how and where should data be gathered.



Where to put survey lines for timber volumes information?
Typically partial, imperfect information.

Norwegian wood - posterior

Design matrix: $\boldsymbol{F}$



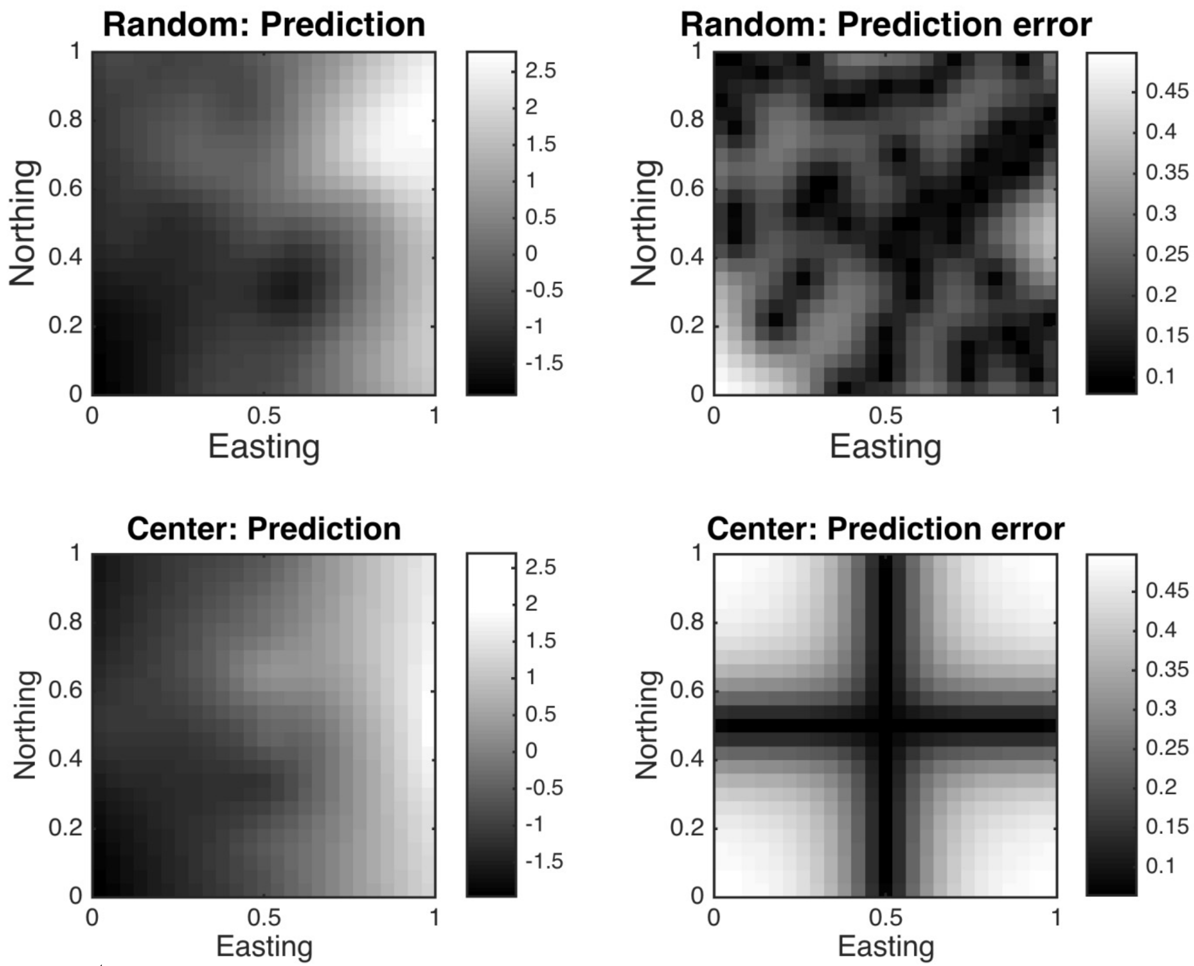
Goal is:

$$p(\boldsymbol{x}\,|\,\boldsymbol{y}) = \frac{p(\boldsymbol{y}\,|\,\boldsymbol{x})\,p(\boldsymbol{x})}{p(\boldsymbol{y})}$$

$$p(\boldsymbol{x}\,|\,\boldsymbol{y}) = N\left(\boldsymbol{\mu} + \Sigma\boldsymbol{F}^t\left(\boldsymbol{F}\Sigma\boldsymbol{F}^t + \tau^2\boldsymbol{I}\right)^{-1}(\boldsymbol{y} - \boldsymbol{F}\boldsymbol{\mu}), \Sigma - \Sigma\boldsymbol{F}^t\left(\boldsymbol{F}\Sigma\boldsymbol{F}^t + \tau^2\boldsymbol{I}\right)^{-1}\boldsymbol{F}\Sigma\right)$$

This is Kriging prediction and associated variance.

Norwegian wood – posterior results

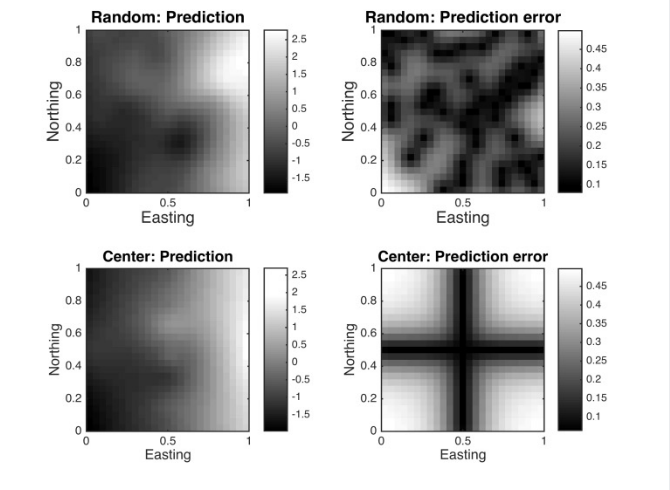


$$p(x\mid y)=N\left(\mu+\Sigma F^{t}\left(F\Sigma F^{t}+\tau^{2}I\right)^{-1}\left(y-F\mu\right),\Sigma-\Sigma F^{t}\left(F\Sigma F^{t}+\tau^{2}I\right)^{-1}F\Sigma\right)$$

Norwegian wood – information

We can base data gathering schemes on different criteria

- Maximum variance reduction
- Maximum entropy
- Value of information (VOI)



VOI is based on decision situation!
Others are not material – not tied to decision situation.

Spatial design

- Geometric criterion (space-filling design).
  - Minimize average distance between data locations.
  - Set a threshold on minimum distance to nearest data location.

  Challenging to compare various data accuracies.
- Variance reduction criterion.
- Kriging-related criteria (slope and weight of mean).
- Entropy reduction criterion.
- Prediction error.

Variance reduction

$$\mathrm{SV} = \sum_{i=1}^{n} Var(x_i) = \sum_{i=1}^{n} \Sigma_{ii} = \mathrm{trace}\left(\boldsymbol{\Sigma}\right)$$

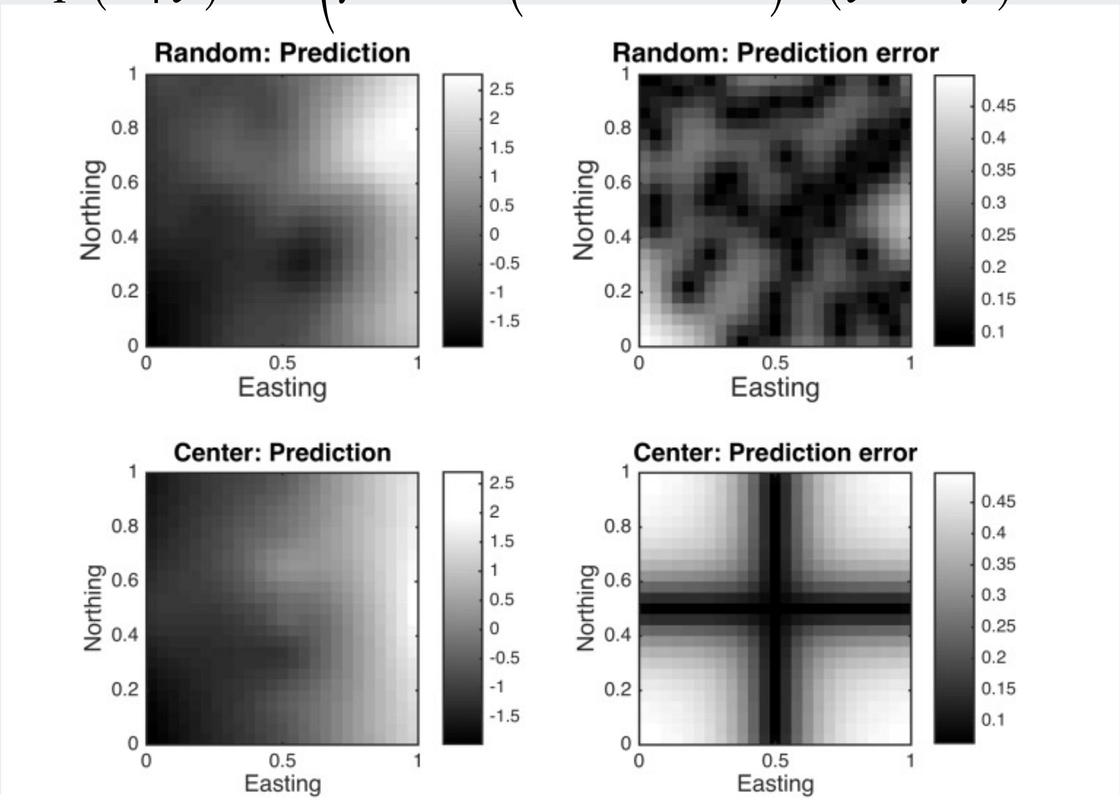
Expected variance reduction:

$$\mathrm{EVR}(\boldsymbol{y}) = \sum_{i=1}^{n} Var(x_i) - E\left(\sum_{i=1}^{n} Var(x_i \mid \boldsymbol{y})\right) = \sum_{i=1}^{n} Var(x_i) - \int \sum_{i=1}^{n} Var(x_i \mid \boldsymbol{y}) p\left(\boldsymbol{y}\right) d\boldsymbol{y}$$

Could use a weighted sum, or choose a subset of variables for prediction.

Variance reduction (Kriging)

$$p(\boldsymbol{x} \mid \boldsymbol{y}) = N\left(\boldsymbol{\mu} + \boldsymbol{\Sigma}\boldsymbol{F}^t\left(\boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}^t + \tau^2\boldsymbol{I}\right)^{-1}\left(\boldsymbol{y} - \boldsymbol{F}\boldsymbol{\mu}\right), \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\boldsymbol{F}^t\left(\boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}^t + \tau^2\boldsymbol{I}\right)^{-1}\boldsymbol{F}\boldsymbol{\Sigma}\right)$$



Overall variance reduction is larger for the random design.

Entropy (Shannon)

$$Ent(\boldsymbol{x}) = -\int p(\boldsymbol{x}) \log p(\boldsymbol{x}) d\boldsymbol{x}$$

$$Ent(\boldsymbol{x} \mid \boldsymbol{y}) = -\int p(\boldsymbol{x} \mid \boldsymbol{y}) \log p(\boldsymbol{x} \mid \boldsymbol{y}) d\boldsymbol{x}$$

Expected mutual information:

$$\mathrm{EMI}\left(\boldsymbol{y}\right) = Ent(\boldsymbol{x}) - \int Ent(\boldsymbol{x} \mid \boldsymbol{y}) p(\boldsymbol{y}) d\boldsymbol{y}$$

EAGE
EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Entropy of a Gaussian

$$Ent(\boldsymbol{x}) = -\int p(\boldsymbol{x}) \log p(\boldsymbol{x}) d\boldsymbol{x}$$

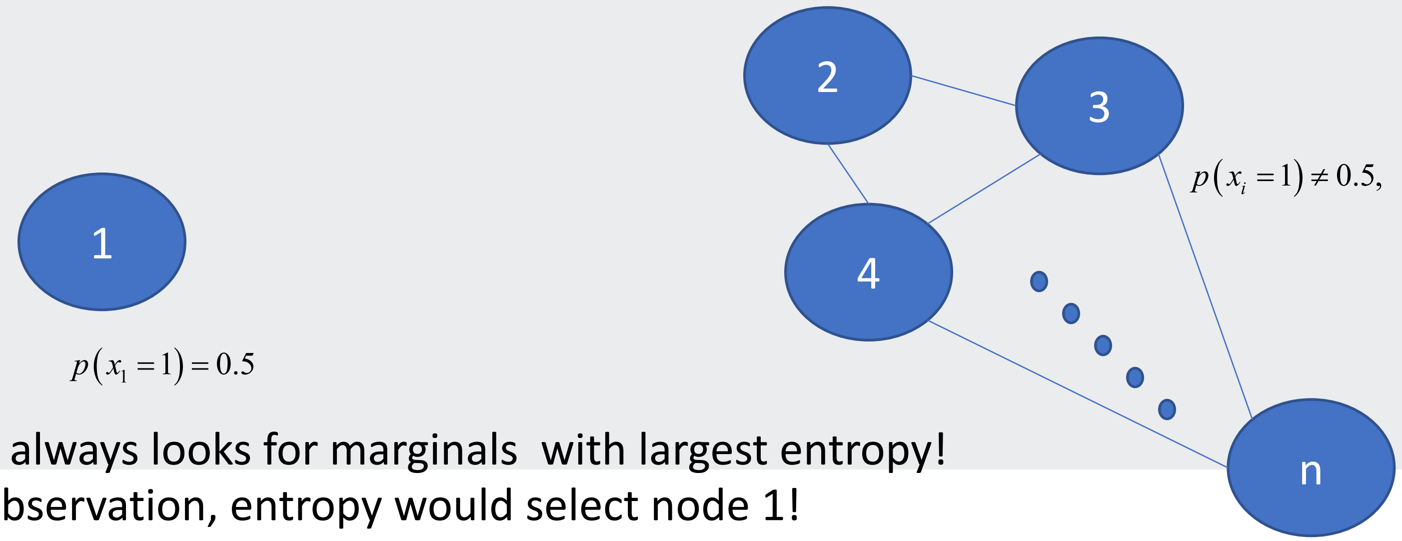
$$Ent(\boldsymbol{x}) = \frac{n}{2}\left(1 + \log(2\pi)\right) + \frac{1}{2}\log|\boldsymbol{\Sigma}|$$

Very commonly used in the design of spatial experiments (air quality monitoring, river monitoring networks, etc.)

Entropy blind-spot

$$EMI(x_K) = Ent(\boldsymbol{x}) - \int Ent(\boldsymbol{x}_L \mid x_K)\, p(x_K)\, dx_K,$$

$$= Ent(x_K) + \int Ent(\boldsymbol{x}_L \mid x_K)\, p(x_K)\, dx_K - \int Ent(\boldsymbol{x}_L \mid x_K)\, p(x_K)\, dx_K = Ent(x_K).$$



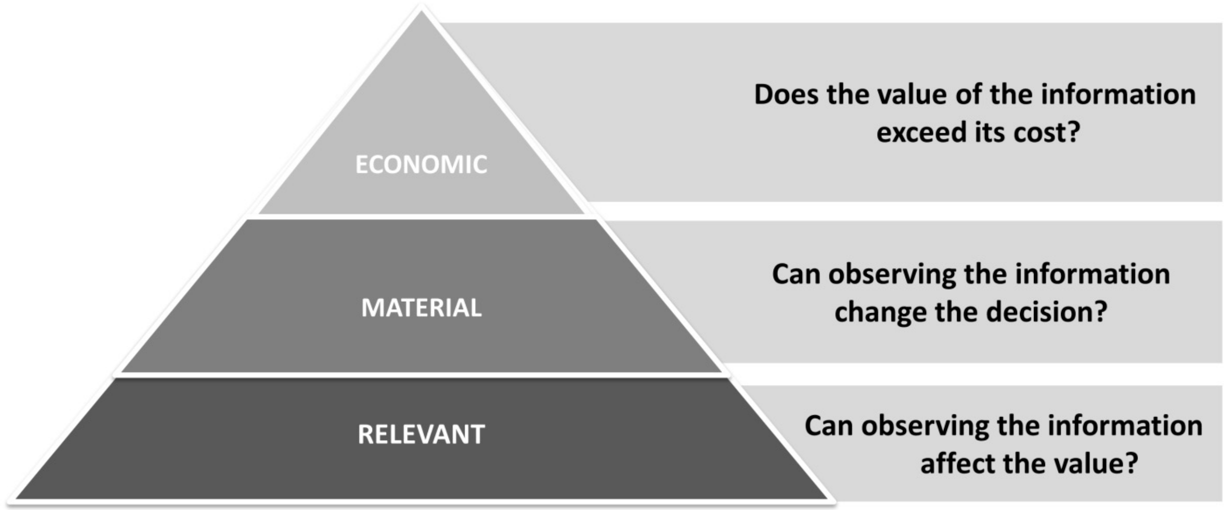
$p(x_1 = 1) = 0.5$

$p(x_i = 1) \neq 0.5, \quad i = 2,..,n$

This criterion always looks for marginals with largest entropy!

For a single observation, entropy would select node 1!

VOI - Pyramid of conditions



Pyramid of conditions - VOI is different from other information criteria (entropy, variance, prediction error, etc.)

$$PV = \max\left\{0, E\left(\sum_{i=1}^{n} x_i\right)\right\} = \max\left\{0, \sum_{i=1}^{n} \mu_i\right\}$$

Low flexibility:
Must select all units, or none.

Value decouples to sum.

$$PoV(\boldsymbol{y}) = \int \max\left\{0, E\left(\sum_{i=1}^{n} x_i \mid \boldsymbol{y}\right)\right\} p(\boldsymbol{y})\, d\boldsymbol{y}$$

VOI – Gaussian models

$$PV = \max\left\{0, E\left(\sum_{i=1}^{n} x_i\right)\right\} = \max\left\{0, \sum_{i=1}^{n} \mu_i\right\}$$

Low flexibility:
Must select all units, or none.

$$\boldsymbol{R} = \boldsymbol{\Sigma F}^t\left(\tau^2 \boldsymbol{I} + \boldsymbol{F\Sigma F}^t\right)^{-1} \boldsymbol{F\Sigma}$$

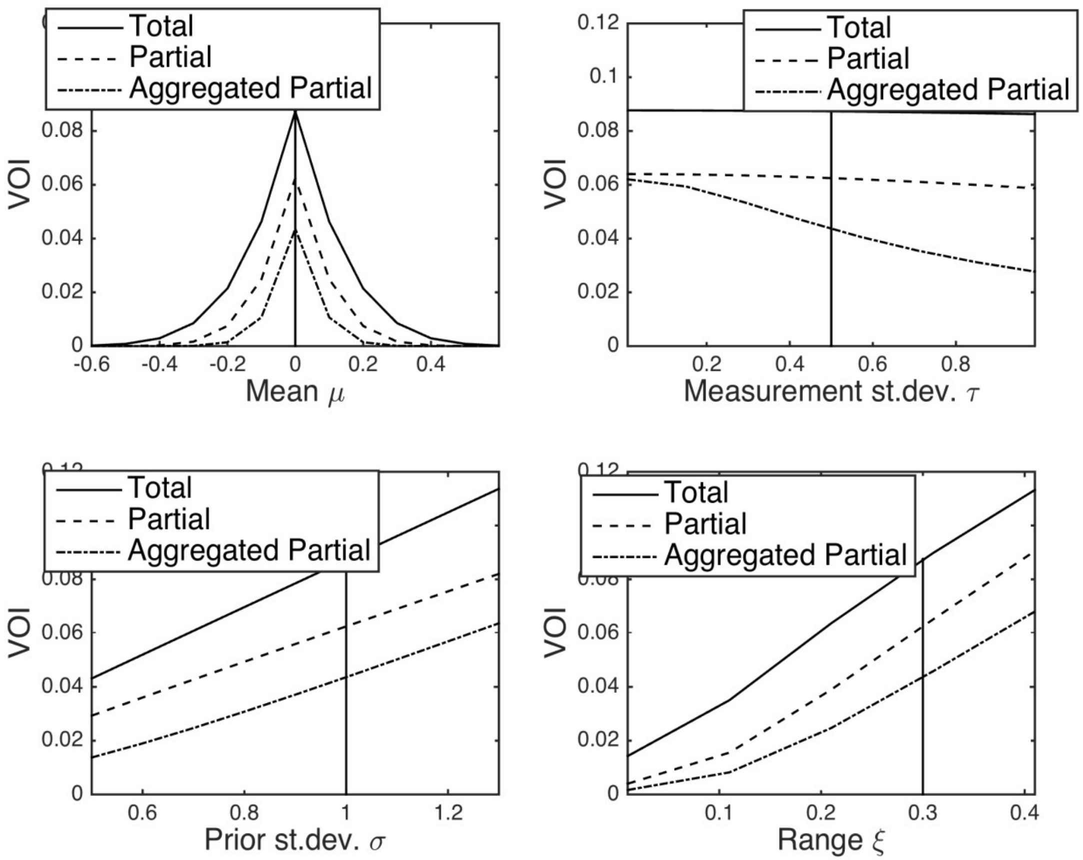
$$\mu_w = \sum_{i=1}^{n} \mu_i \qquad r_w^2 = \sum_{k=1}^{n}\sum_{l=1}^{n} R_{kl}$$

$$PoV(\boldsymbol{y}) = \int \max\left\{0, E\left(\sum_{i=1}^{n} x_i \mid \boldsymbol{y}\right)\right\} p(\boldsymbol{y})\, d\boldsymbol{y} = \mu_w \Phi\left(\mu_w / r_w\right) + r_w \phi\left(\mu_w / r_w\right)$$

$\phi(z), \Phi(z)$ standard Gaussian density and cumulative function

Results - Forestry example

Low flexibility:
Must select all units, or none.



Total: all cells. Partial: Every cell along center lines. Aggregated partial: sums along center lines.

(Results are normalized for area).

Norwegian wood - Insight in VOI

- Total test does not necessarily give much higher VOI than a partial test. It depends on the spatial design of experiment as well as the prior model (mean and dependence).
- VOI increases with larger dependence in spatial uncertainties.
- VOI is largest when we are most indifferent in prior (mean near 0 and large prior uncertainty.
- VOI increases with higher accuracy of measurements.

Implementation : Norwegian wood

Problem:

- Consider profits, modeled as a Gaussian random field represented on a 25 x 25 grid for the 625 units. The mean is m=0 at all cells, the covariance is exponential with st dev r=1 and correlation range r=40. Draw a random realization of this Gaussian process.

- One can gather imperfect data at 100 random design locations, giving unbiased profit measurements, and independent error with st dev 0.5. Draw a random dataset.

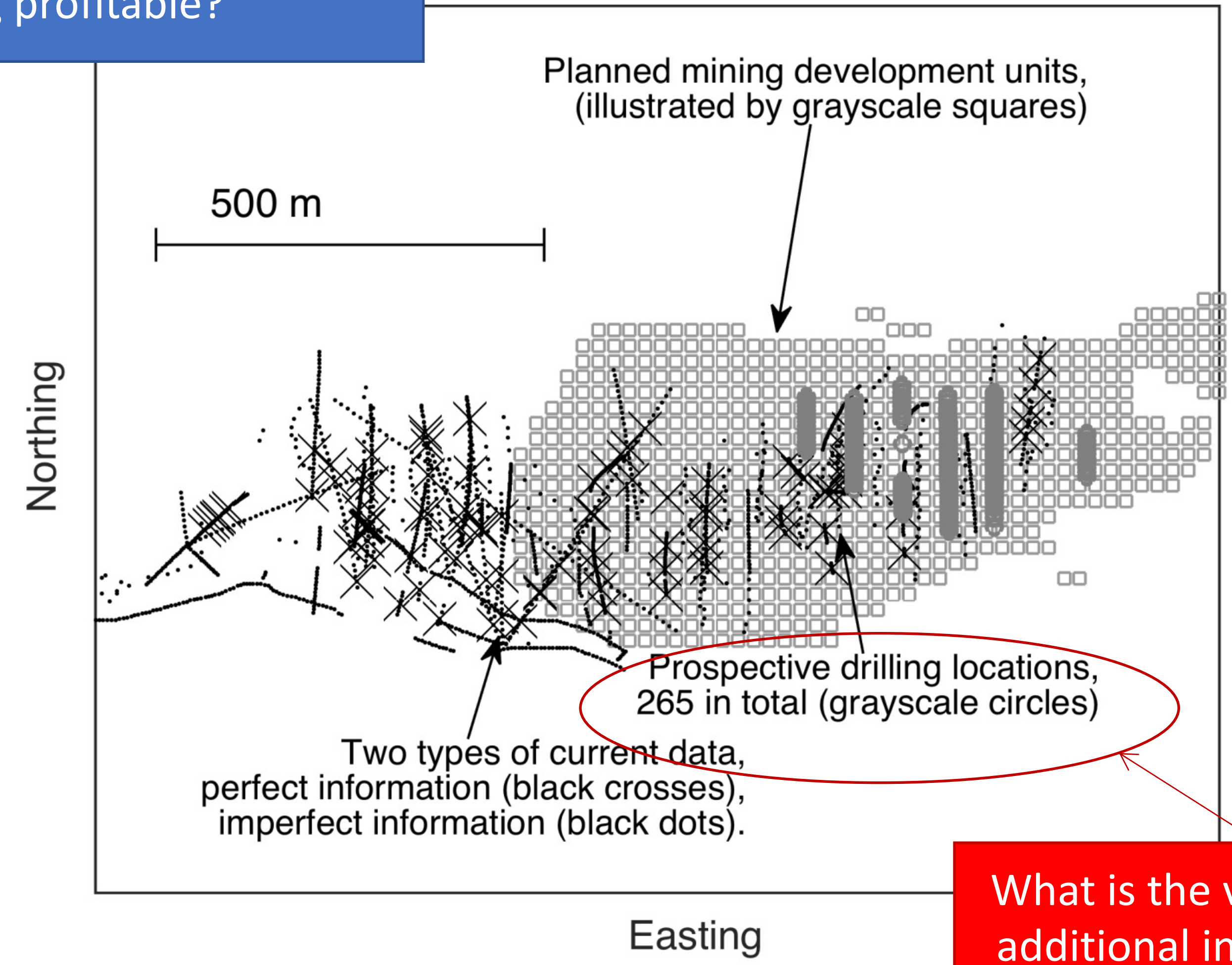
- Compute the Kriging prediction and the associated variances.

- Compute the VOI (using the same data design) of the decision situation where the farmer harvest all units or none.

(MATLAB)

I love rock and ore – mining example

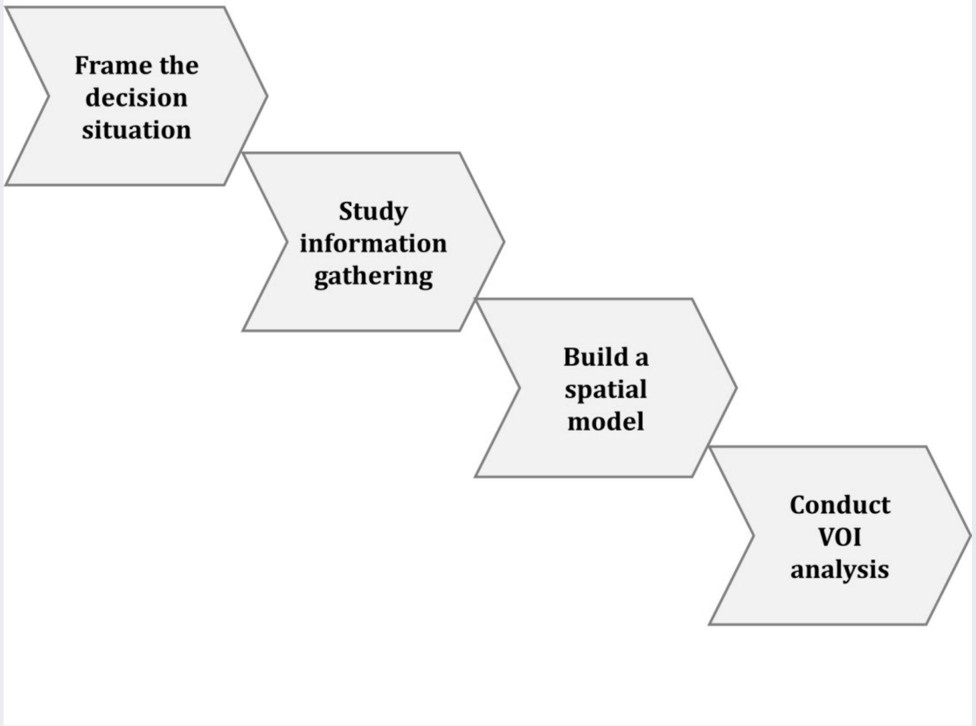
Is mining profitable?



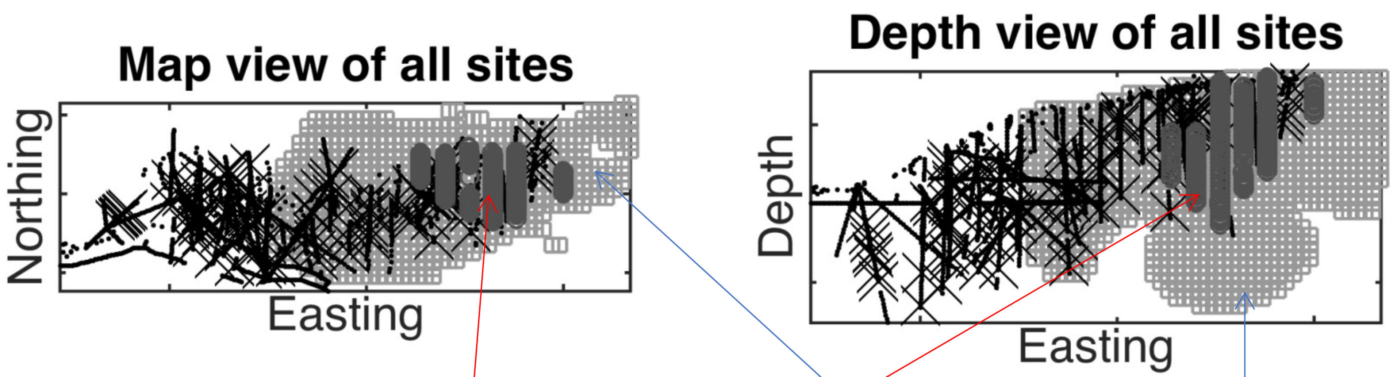
What is the value of this additional information?

VOI workflow

- Low decision flexibility. De-coupled value function.
- Gather information by XRF or XMET in boreholes. No opportunities for adaptive testing.
- Model is a spatial Gaussian process.
- VOI analysis done by exact, Gaussian, computations.



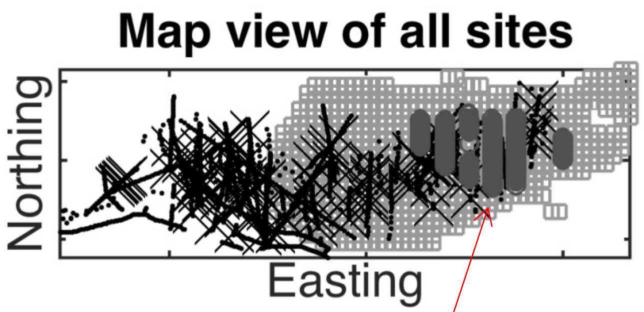
Decision situation and data

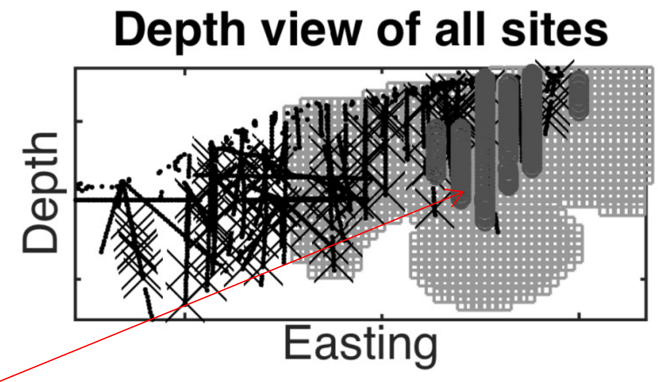


Planned boreholes.

Mining blocks. Some waste rock. Some high-grade.

Information gathering

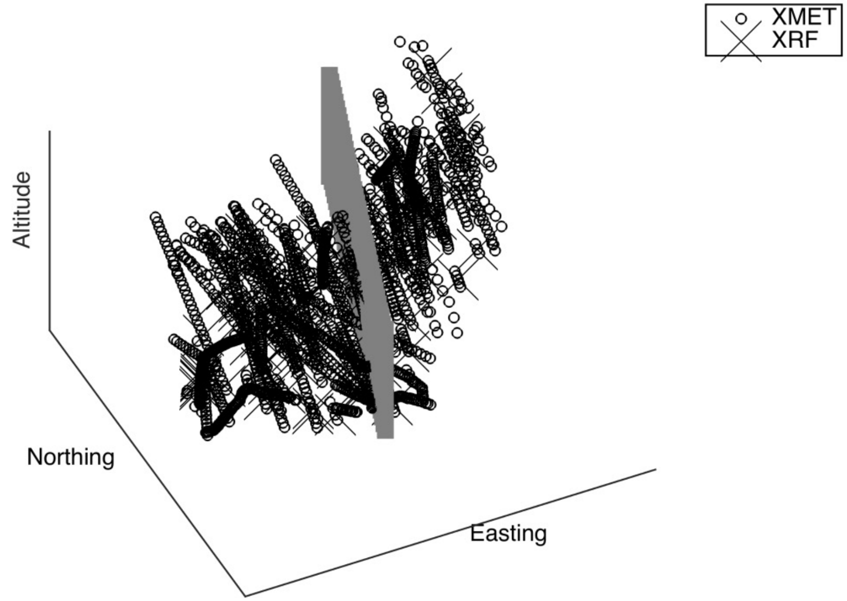


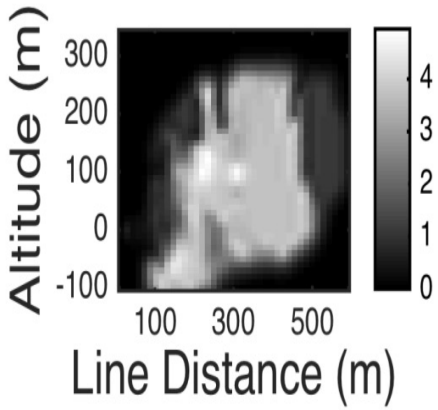


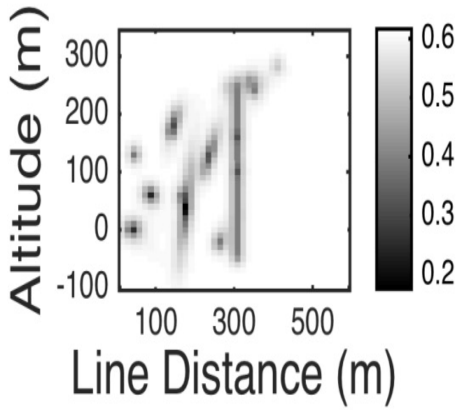
Planned
boreholes.

- **Total test** : 265 measurements in 21 new boreholes.
- **Partial test**: Drilling and sampling data only in a subset of boreholes.
- **Perfect** testing (XRF: done in lab). **Imperfect** testing (XMET: handheld meter).

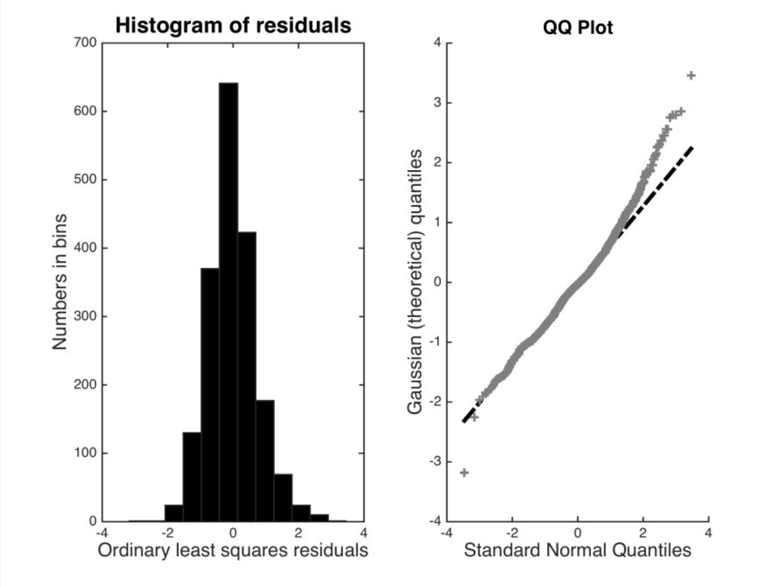
Prior model

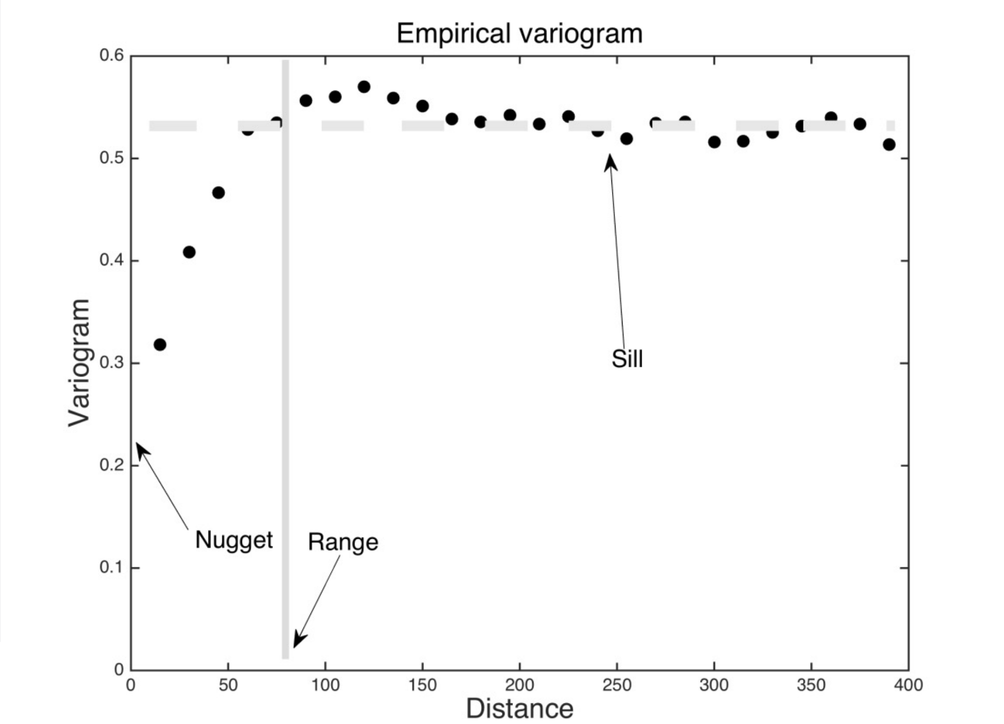

XMET
XRF
Altitude
Northing
Easting


Altitude (m)
300
200
100
0
-100
100
300
500
Line Distance (m)
4
3
2
1
0


Altitude (m)
300
200
100
0
-100
100
300
500
Line Distance (m)
0.6
0.5
0.4
0.3
0.2

Model


Histogram of residuals
Numbers in bins
Ordinary least squares residuals
QQ Plot
Gaussian (theoretical) quantiles
Standard Normal Quantiles


Empirical variogram
Variogram
Distance
Nugget
Range
Sill

Prior and likelihood model

Set from current data.

$$p(\boldsymbol{x}) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$p(\boldsymbol{y} \mid \boldsymbol{x}) = N(\boldsymbol{Fx}, \boldsymbol{T})$$

Defined by test (XRF, XMET).

Defined by design of boreholes.

$$y(\mathbf{s}) = x(\mathbf{s}),$$ -XRF data

$$y(\mathbf{s}) = x(\mathbf{s}) + N(0, \tau^2)$$ -XMET data

VOI

Weights set from block model(waste or ore).

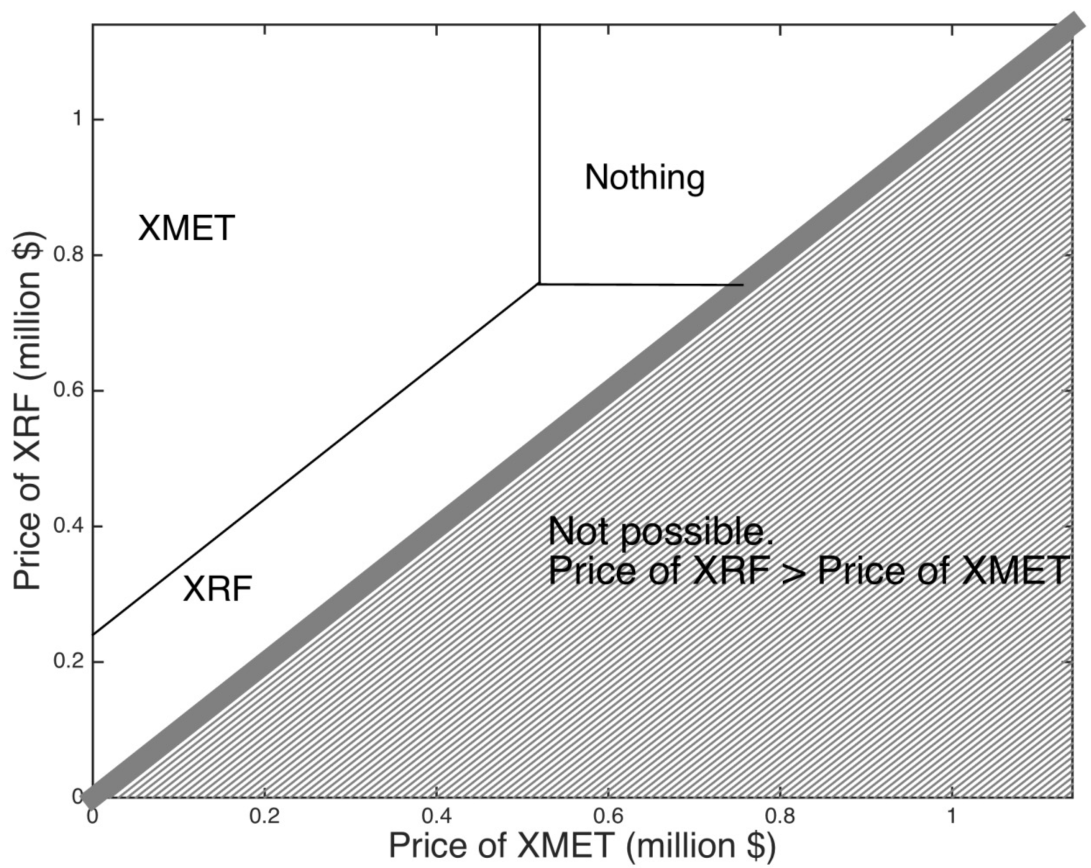
$$PV = \max\left\{E\left(\boldsymbol{w}^t\boldsymbol{x} - \text{Cost}\right), 0\right\}$$

$$PoV\left(\boldsymbol{y}\right) = \int \max\left\{E\left(\boldsymbol{w}^t\boldsymbol{x} - \text{Cost} \mid \boldsymbol{y}\right), 0\right\} p\left(\boldsymbol{y}\right) d\boldsymbol{y}$$

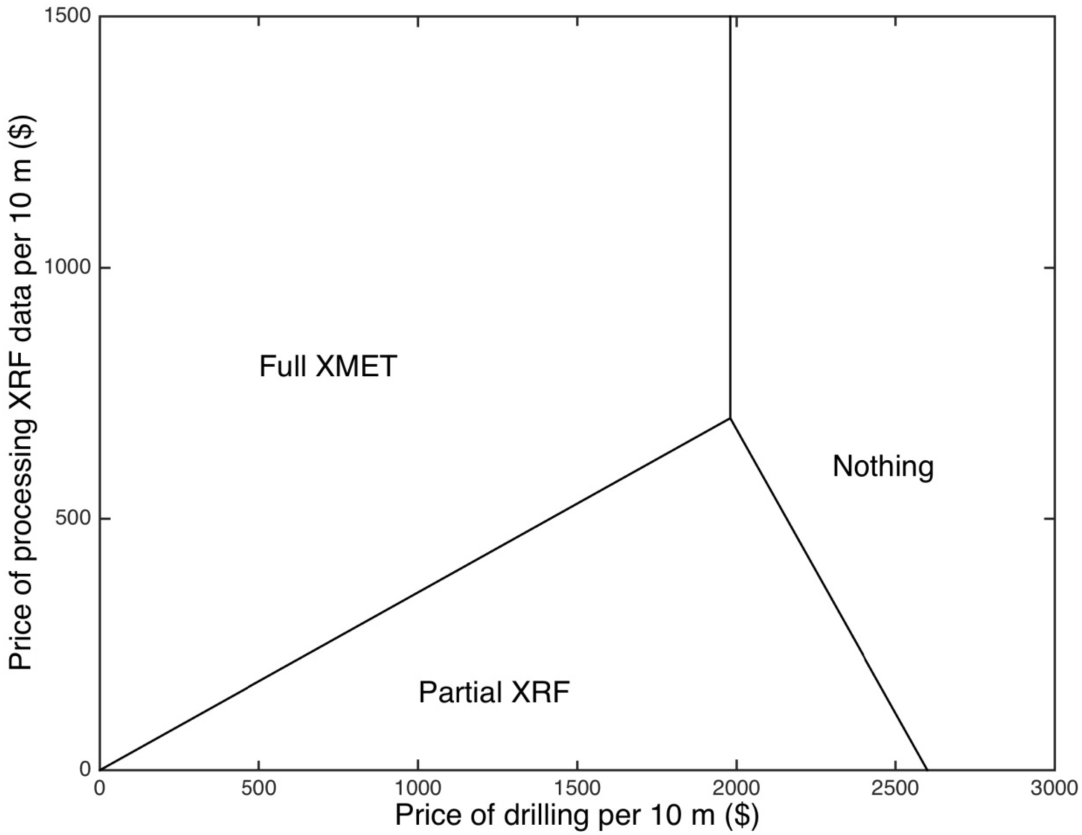
$$VOI\left(\boldsymbol{y}\right) = PoV\left(\boldsymbol{y}\right) - PV.$$

Analytical solution under the Gaussian modeling assumptions.

VOI : Decision regions XRF,XMET.



VOI : Decision regions, partial data.



I love rock and ore - Take home:

.

- Information connected to partial perfect testing can be less/more than total imperfect testing.
- Information criteria depend on design and data accuracy.
- Entropy appears to like perfect information.
- VOI can be connected with decisions and prices (not so easy for other criteria).

Plan for course

Time	Topic
Day 1	Introduction and motivating examples
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Day 2	Examples of value of information analysis in Earth sciences
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.

Computation - Formula for VOI

$$PV = \max_{a \in A} \left\{ E\left(v(\boldsymbol{x}, \boldsymbol{a}) \right) \right\} = \max_{a \in A} \left\{ \int_{\boldsymbol{x}} v(\boldsymbol{x}, \boldsymbol{a}) p(\boldsymbol{x}) d\boldsymbol{x} \right\}$$

$$PoV(\boldsymbol{y}) = \int \max_{a \in A} \left\{ E\left(v(\boldsymbol{x}, \boldsymbol{a}) \mid \boldsymbol{y} \right) \right\} p(\boldsymbol{y}) d\boldsymbol{y}$$

Main challenge.

Computations :

- Easier with low decision flexibility (less alternatives).
- Easier if value decouples (sums or integrals split).
- Easier for perfect, total, information (upper bound on VOI).
- Sometimes analytical solutions, otherwise approximations and Monte Carlo.

Techniques – Computing the VOI

$$PV = \max_{a \in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a})\right)\right\} = \max_{a \in A}\left\{\int_{x} v(\boldsymbol{x},\boldsymbol{a})\, p(\boldsymbol{x})\, d\boldsymbol{x}\right\}$$

$$PoV(\boldsymbol{y}) = \int \max_{a \in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a}) \mid \boldsymbol{y}\right)\right\} p(\boldsymbol{y})\, d\boldsymbol{y}$$

Inner integral.

Outer integral.

Techniques :

- Fully analytically tractable for two-action, linear Gaussian models.
- Analytical or partly analytical for Markovian models, graphs.
  (Monte Carlo sample over data, analytical for inner expectation.)
- Various approximations and Monte Carlo usually applicable.
- Should avoid double Monte Carlo (inner and outer). Too time consuming.

Fully analytical

$$PV = \max\{0, \mu_w\}$$

Inner integral analytical. Linear combination of data.

$$PoV(\boldsymbol{y}) = \int \max\{0, \boldsymbol{\alpha}'\boldsymbol{y}\}\, p(\boldsymbol{y})\, d\boldsymbol{y}$$

Gaussian.

$$= \int \max\{0, w\}\, p(w)\, dv = \mu_w \Phi\left(\frac{\mu_w}{r_w}\right) + r_w \phi\left(\frac{\mu_w}{r_w}\right)$$

Partly analytical, Monte Carlo for rest

$$PV = \max\{0, \mu_w\}$$

Inner integral solved.

$$PoV(\boldsymbol{y}) = \int \max\{0, f(\boldsymbol{y})\}\, p(\boldsymbol{y})\, d\boldsymbol{y}$$

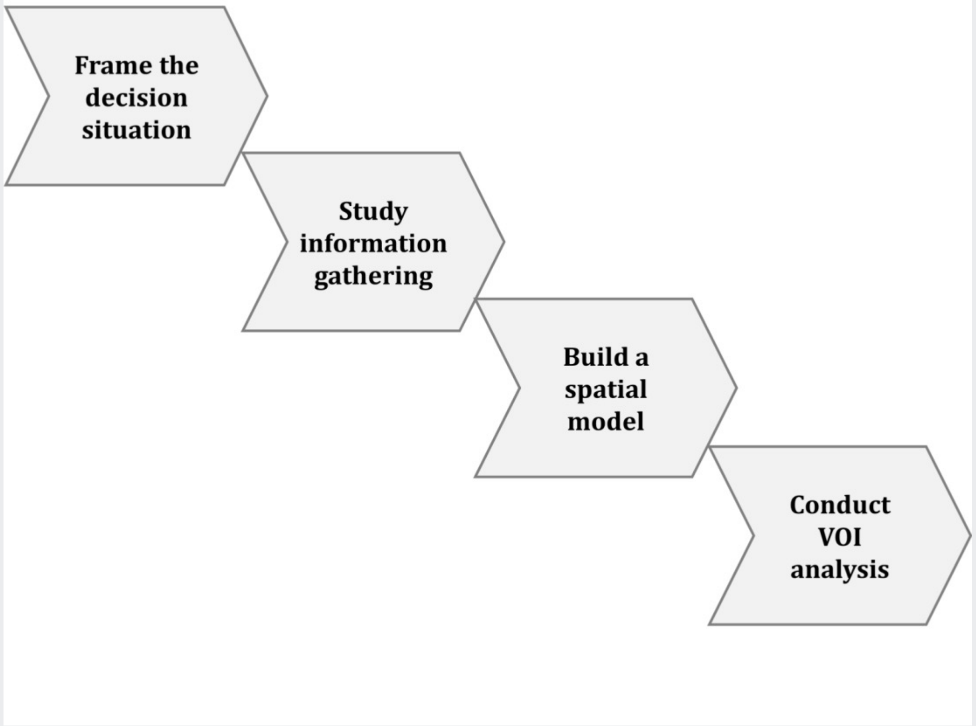
Use sampling.

$$= \frac{1}{B}\sum_{b=1}^{B} \max\{0, f(\boldsymbol{y}^b)\}$$

$$\boldsymbol{y}^b \sim p(\boldsymbol{y}), \qquad b = 1, \ldots, B.$$

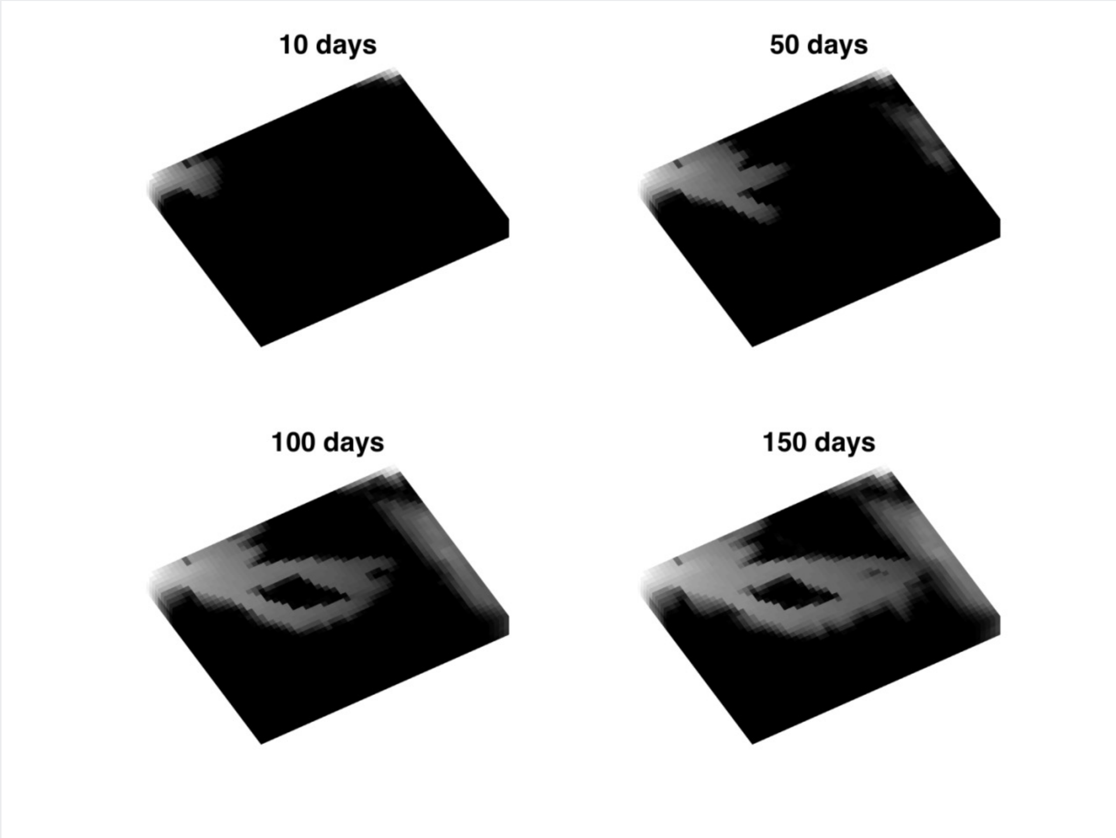
Reservoir dogs - petroleum example

- Decisions about drilling alternatives.
- Seismic information. Which kind?
- Model is a represented by spatial process, obtained by simulations.
- VOI analysis done by a simulation-regression approach.



Reservoir flow in heterogeneous media

Injection of water, pushing oil out.



Flow in the reservoir depends on the composition of rocks, porosity, permeability, faults, etc.

Seismic data can help identify these important reservoir properties.

Very non-linear relations!

Computation - Formula for VOI

$$PV = \max_{a \in A} \left\{ E\left(v(\boldsymbol{x}, \boldsymbol{a})\right) \right\} = \max_{a \in A} \left\{ \int_{\boldsymbol{x}} v(\boldsymbol{x}, \boldsymbol{a}) \, p(\boldsymbol{x}) \, d\boldsymbol{x} \right\}$$

$$PoV(\boldsymbol{y}) = \int \max_{a \in A} \left\{ E\left(v(\boldsymbol{x}, \boldsymbol{a}) \mid \boldsymbol{y}\right) \right\} p(\boldsymbol{y}) \, d\boldsymbol{y}$$

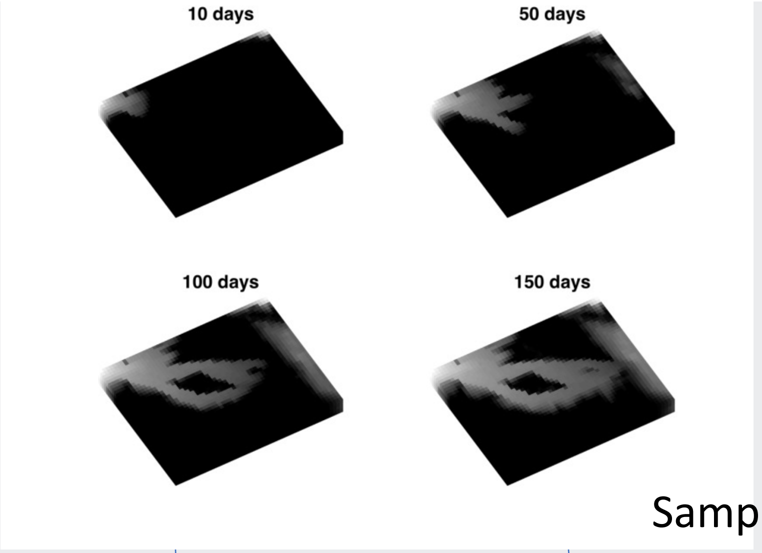
Computationally challenging:

- No decoupling in space. Joint optimization over all alternatives.
- Non-linear value function and seismic data.

Prior - Reservoir uncertainty

Prior is $p(\boldsymbol{x})$.

This distribution of reservoir variables is usually represented by multiple Monte Carlo realizations from the prior distribution.



Sample 1

....

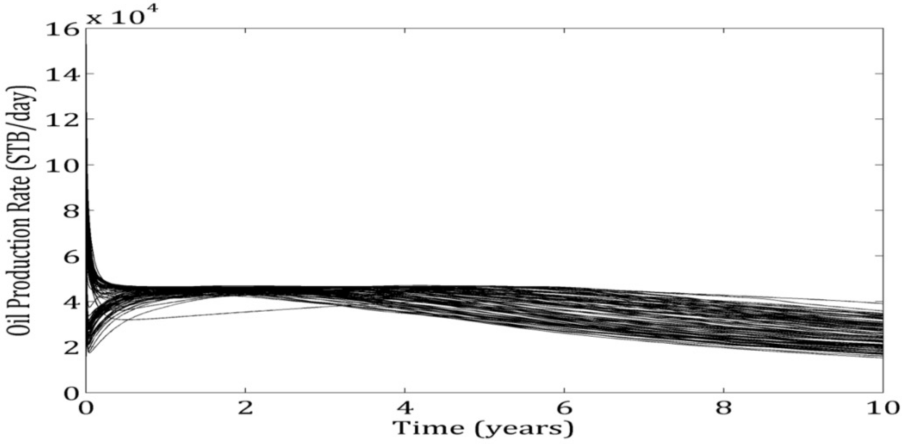
Sample 1000

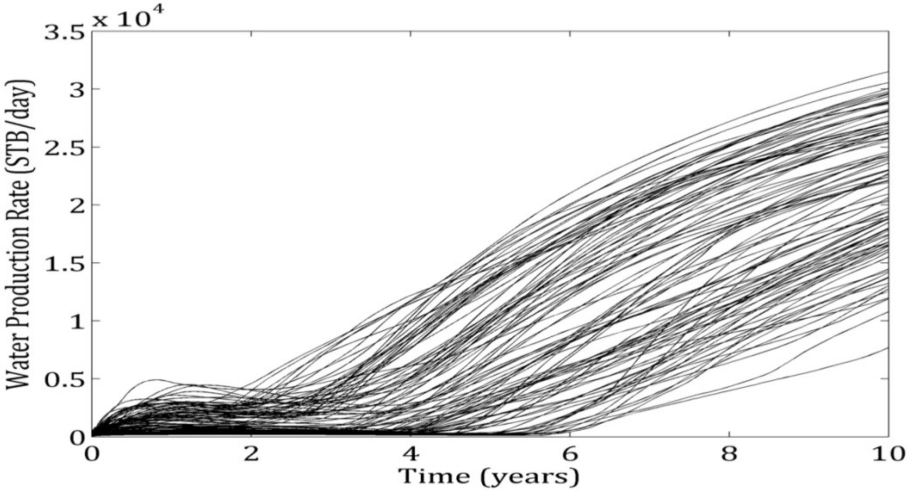
Flow simulation and value

.

$$v\left(\boldsymbol{x}^{b},\boldsymbol{a}\right),\qquad b=1,\ldots,1000$$

Flow simulation gives amount of recoverable oil, for each realization, with the development cost subtracted.



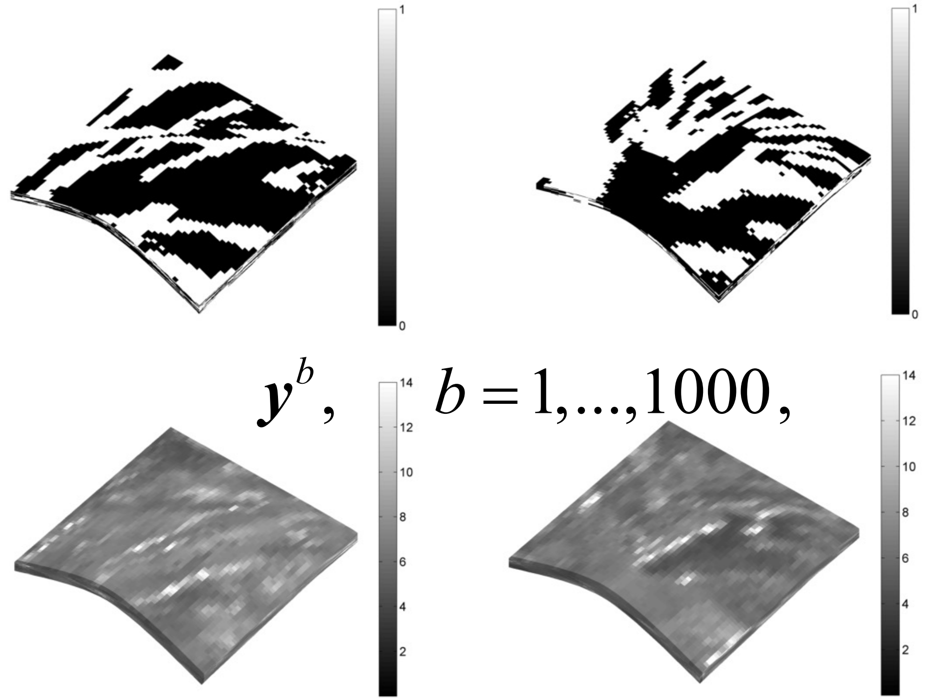


Likelihood - Seismic data

.

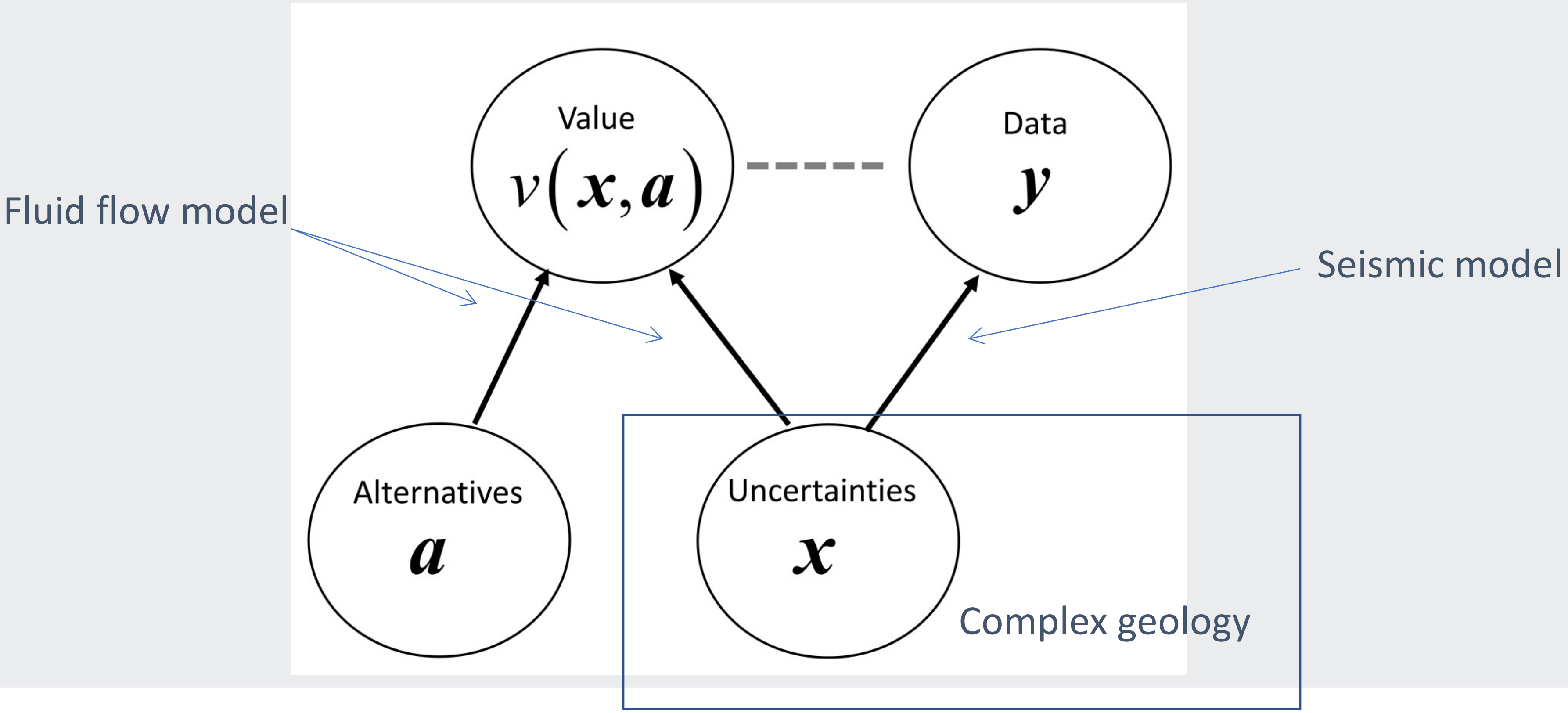
$$p(\boldsymbol{y} \mid \boldsymbol{x}) = N(\boldsymbol{g}(\boldsymbol{x}), \boldsymbol{T})$$

The likelihood is non-linear, but we can generate synthetic seismic data from the likelihood model, given realizations of reservoir properties.



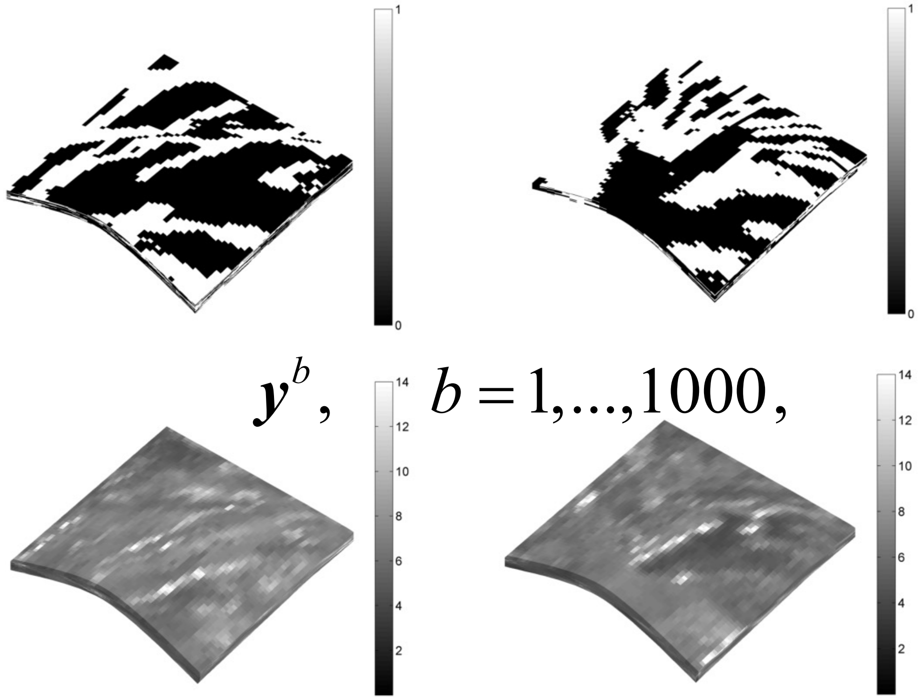
$$\boldsymbol{y}^b, \quad b = 1,\ldots,1000,$$

Linking variables



Value and seismic data

- Random draws of geologic scenario (meandering channels or delta):
- Draw rock-type realizations
- Draw porosity realizations
- Draw permeability realizations
- **Draw value** by fluid flow simulation and economics
- **Draw seismic data** using physics



$$\mathbf{y}^b, \quad b = 1, \ldots, 1000,$$

We next use these samples for VOI approximation.

Computation

Outer expectation: $\boldsymbol{y}$

$$PoV(\boldsymbol{y}) = \sum_{\boldsymbol{y}} \max_{\boldsymbol{a} \in A} \left\{ E\left(v(\boldsymbol{x}, \boldsymbol{a}) \mid \boldsymbol{y} \right) \right\} p(\boldsymbol{y})$$

Inner expectation: $\boldsymbol{x} \mid \boldsymbol{y}$

$$VOI(\boldsymbol{y}) = PoV(\boldsymbol{y}) - PV$$

- Monte Carlo (outer) and simulation-regression for inner expectation!

Simulation-regression algorithm

Outer expectation

$$PoV(\boldsymbol{y}) = \sum_{\boldsymbol{y}} \max_{\boldsymbol{a}\in A}\left\{E\left(v(\boldsymbol{x},\boldsymbol{a}) \mid \boldsymbol{y}\right)\right\} p(\boldsymbol{y})$$

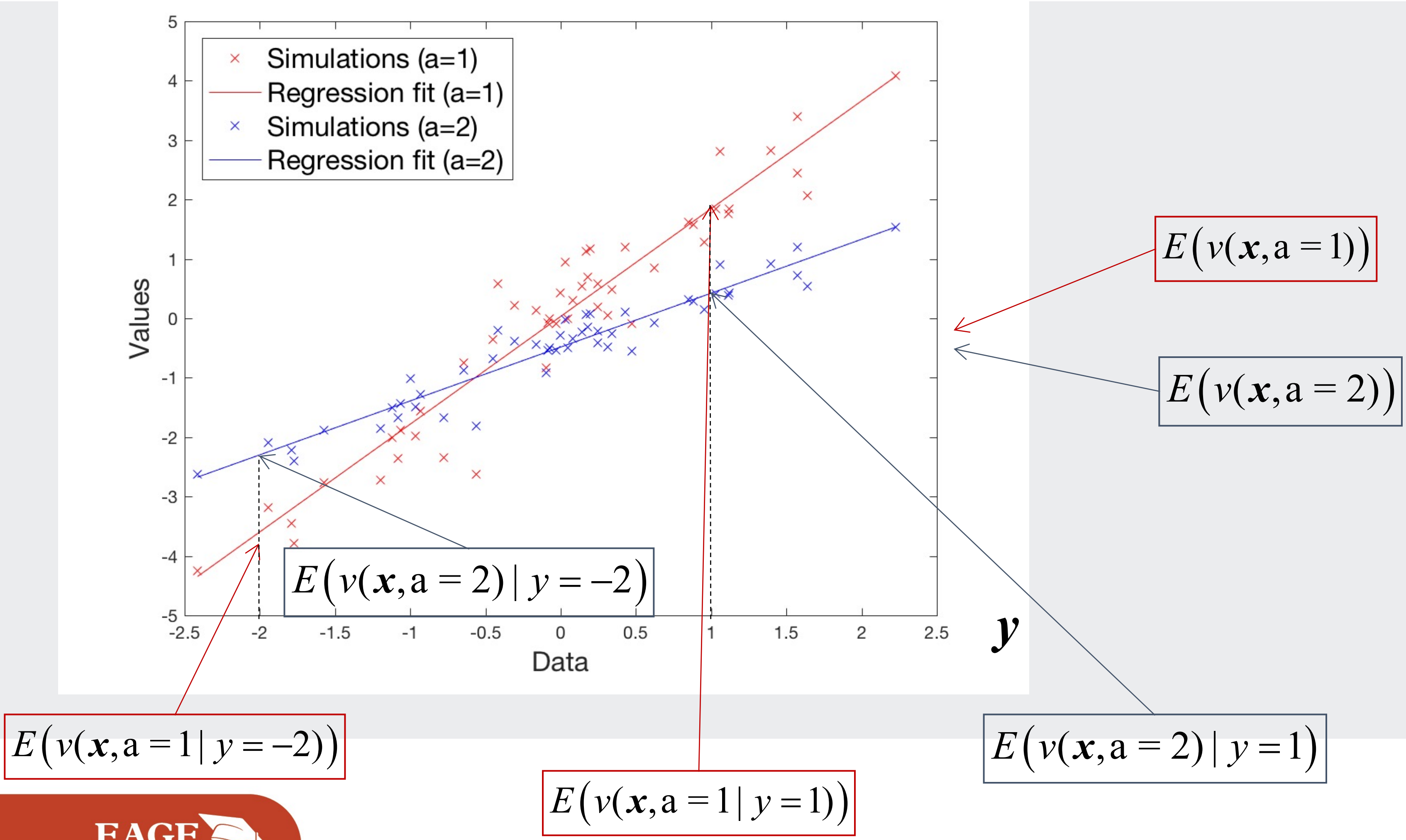
Inner expectation

1. Simulate uncertainties: $\boldsymbol{x}^b, \quad b = 1,\ldots,B$
2. Simulate values, for all alternatives: $v_{\boldsymbol{a}}^b = v\left(\boldsymbol{x}^b,\boldsymbol{a}\right), \quad b = 1,\ldots,B, \quad \boldsymbol{a} \in A$
3. Simulate data: $\boldsymbol{y}^b \sim \left[\boldsymbol{y} \mid \boldsymbol{x}^b\right], \quad b = 1,\ldots,B$
4. Regress samples to fit conditional mean: $\hat{E}\left(v_a \mid \boldsymbol{y}\right)$

$$PoV(\boldsymbol{y}) \approx \frac{1}{B}\sum_{b=1}^{B} \max_{\boldsymbol{a}\in A}\left\{\hat{E}\left(v_{\boldsymbol{a}} \mid \boldsymbol{y}^b\right)\right\}$$

Illustration -regression of samples

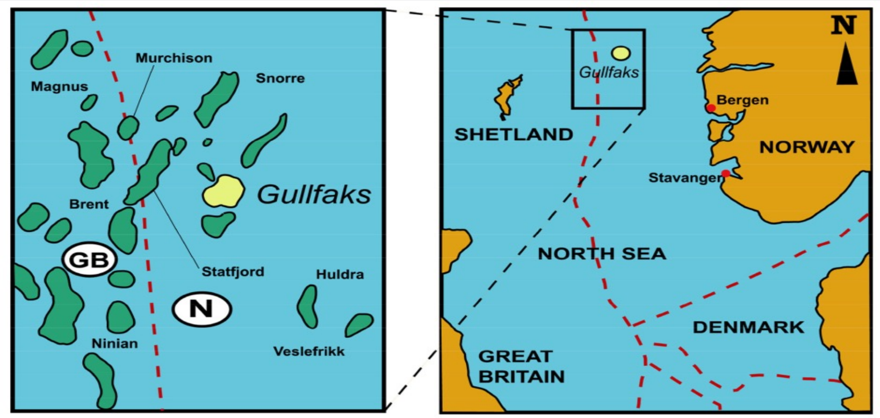
$v(\boldsymbol{x}, \boldsymbol{a})$



Time after time – 4D seismic case

- Decisions about infill drilling.
  - Uncertainty, heterogeneity and dependence make this choice difficult.

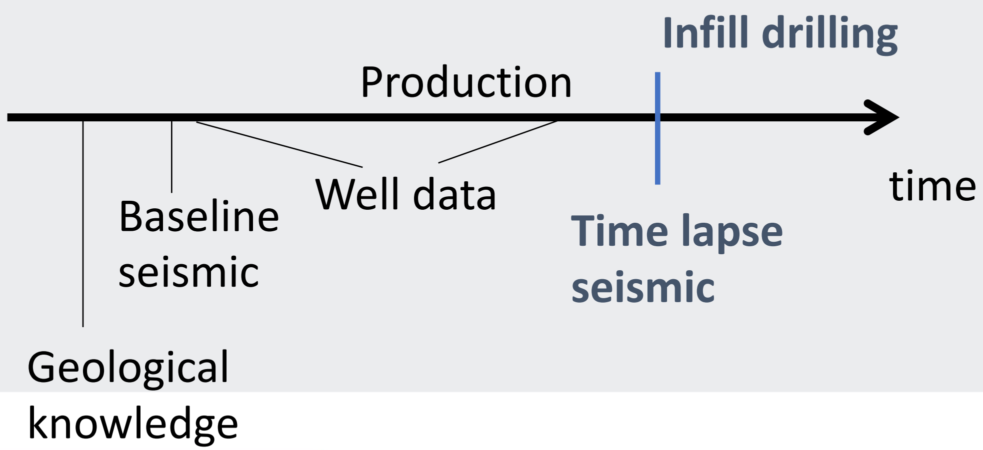
- Data gathering decisions about time-lapse seismic data.
  - Which kind of data are likely to be valuable? How much data is enough?

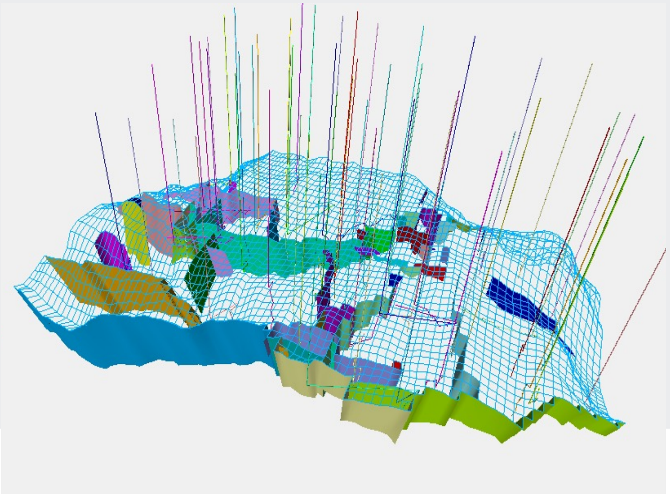


Time after time

- Decisions about infill drilling.
  - Uncertainty, heterogeneity and dependence make this choice difficult.

- Data gathering decisions about time-lapse seismic data.
  - Which kind of data are likely to be valuable? How much data is enough?



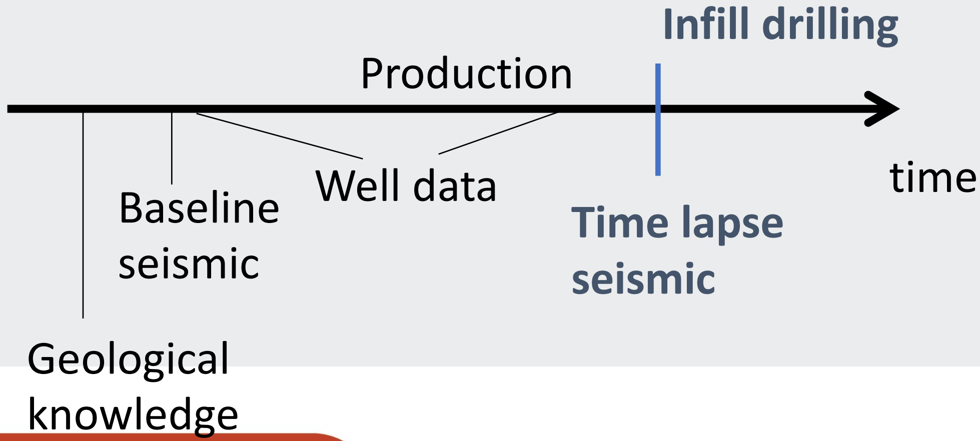


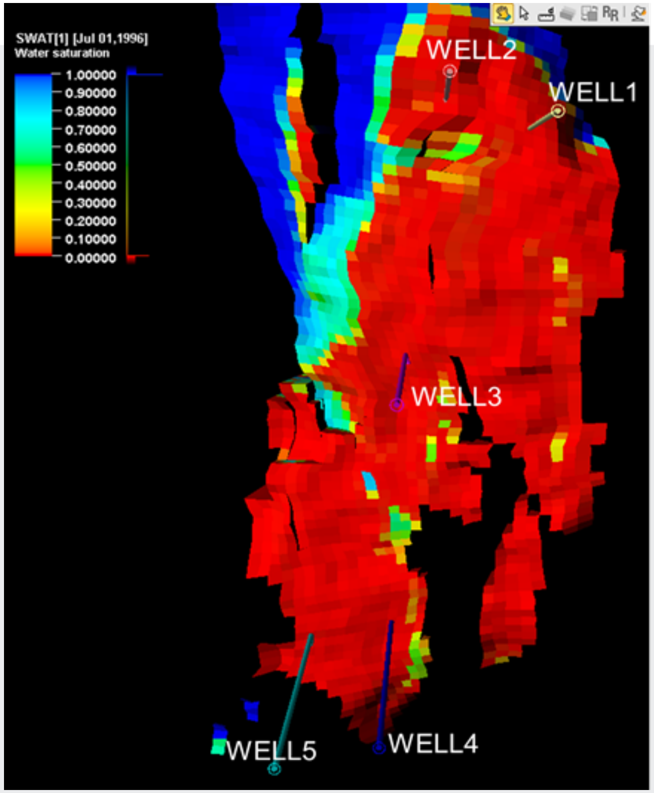
Wells drilled at the Gullfaks field, North Sea.

Gullfaks case (infill drilling and 4D)

- Decisions tied to infill drilling.

- Time-lapse seismic has shown useful here. But no formal VOI analysis was conducted up-front.
- We consider this case in retrospect.



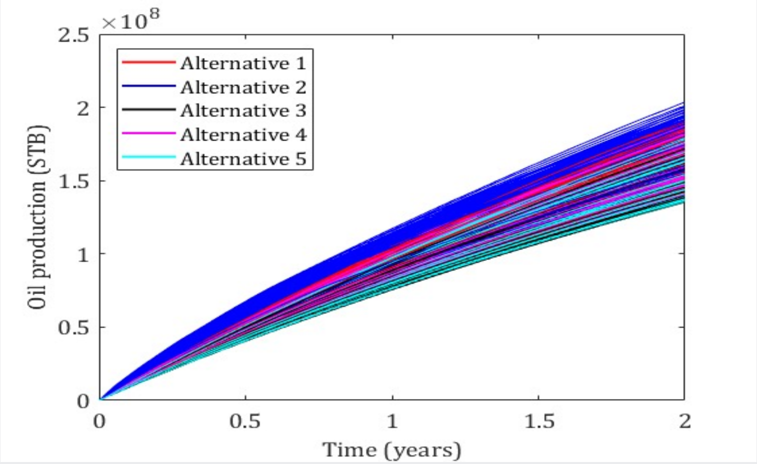


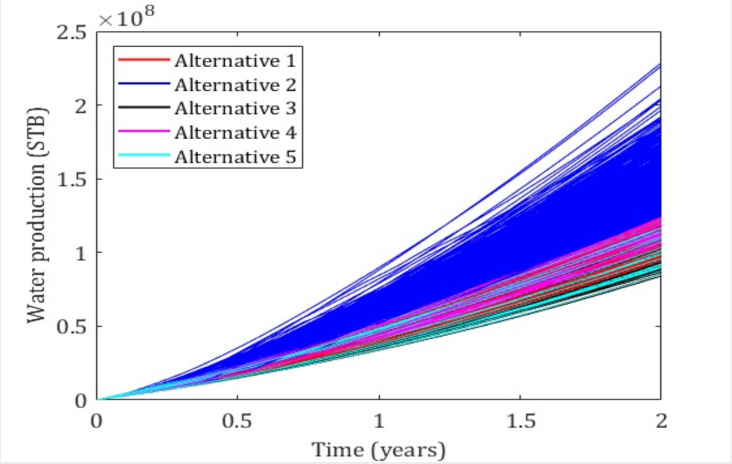
5 decision alternatives.

Gullfaks case (values)

Uncertainties: saturation, pressure, porosity, permeability and fault transmissibilities. (Conditioned on existing data.)

Production for 5 different infill drilling alternatives.

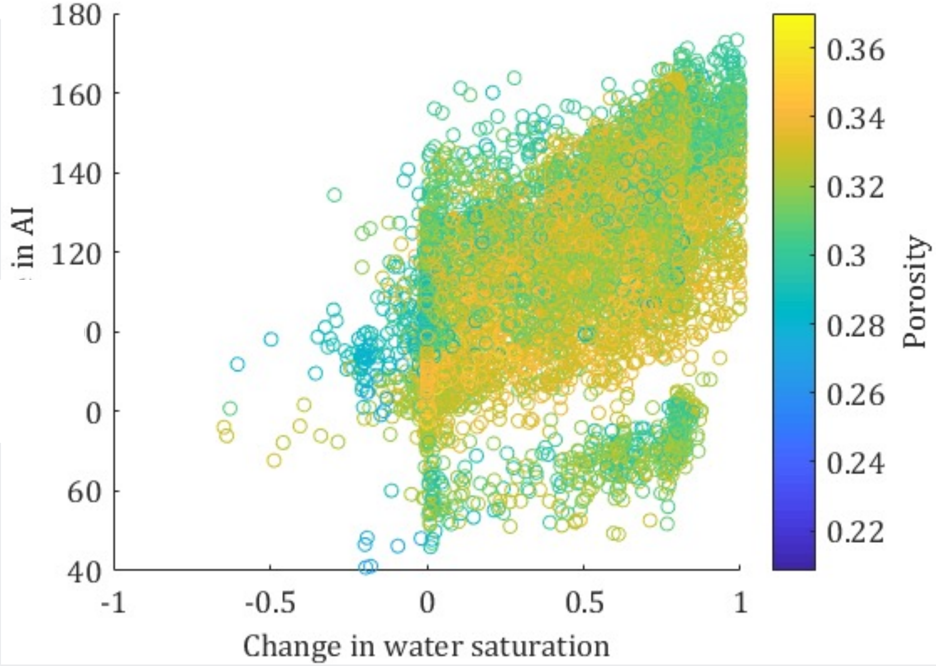




$$v\left(\boldsymbol{x}^b,\boldsymbol{a}\right)=\int\frac{q_o\left(t,\boldsymbol{x}^b,\boldsymbol{a}\right)r_o-q_w\left(t,\boldsymbol{x}^b,\boldsymbol{a}\right)r_w}{\left(1+\alpha\right)^t}dt-C_{drill}\left(\boldsymbol{a}\right),\qquad b=1,\ldots,1000$$

EAGE EDUCATION

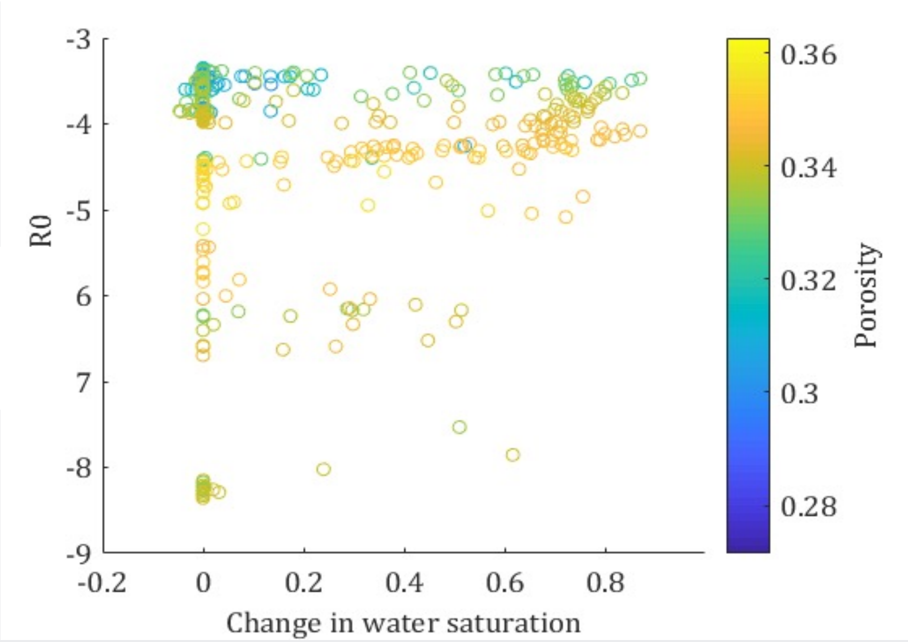
Gullfaks case (likelihood of AI data)

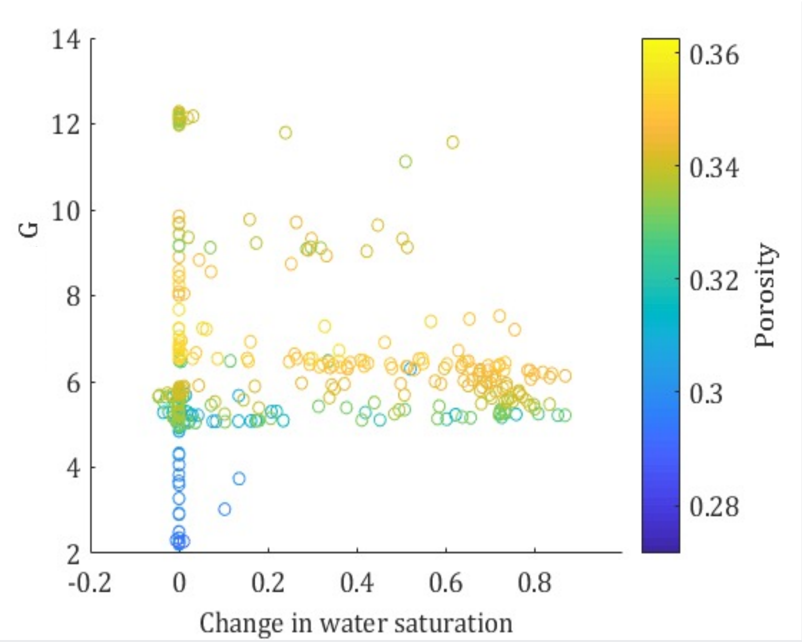


Synthetic time-lapse seismic:
Use rock physics relations connecting reservoir properties to AI, (constant cement model + Gassmann's equation).

Simulations indicate some information about saturation from AI for this case.

Gullfaks case (likelihood of R0,G data)

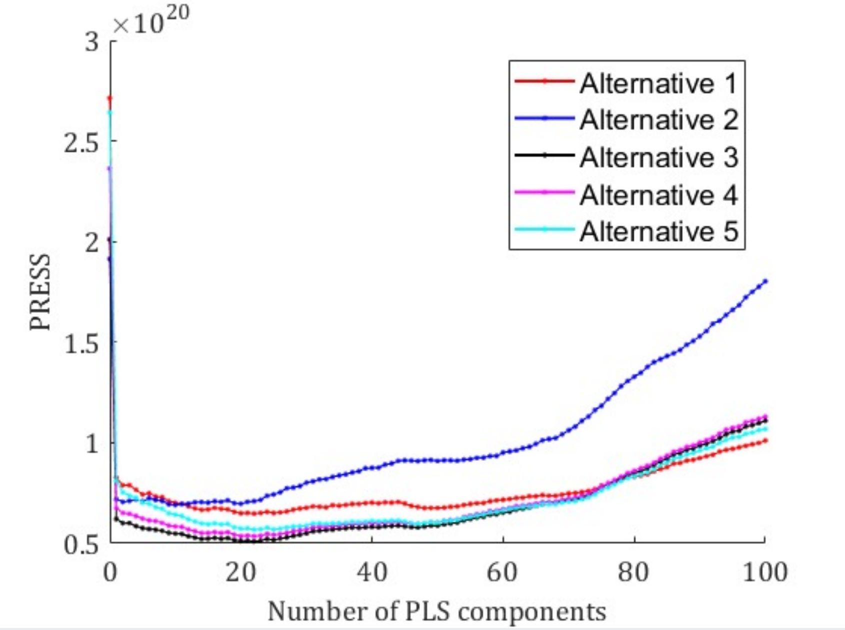




Synthetic time-lapse seismic:
Use rock physics relations connecting reservoir properties to (R0,G).

Simulations indicate limited information about saturation from (R0, G).

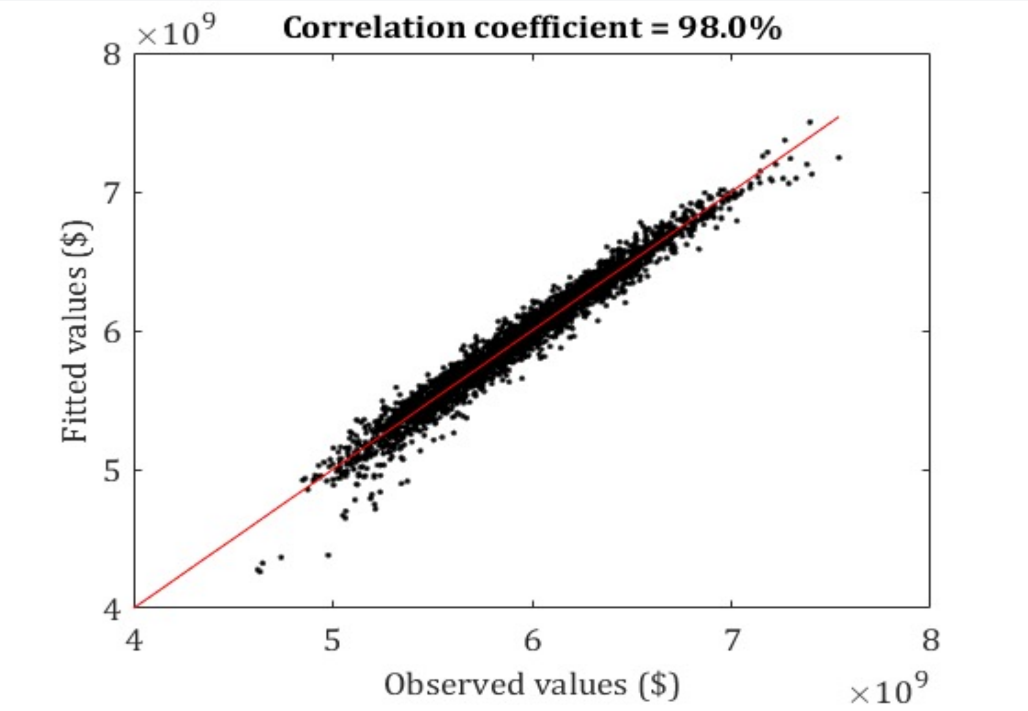
Gullfaks case (PLS for expected values)



$\hat{E}\left(v(\boldsymbol{x},\boldsymbol{a} \mid \boldsymbol{y})\right)$

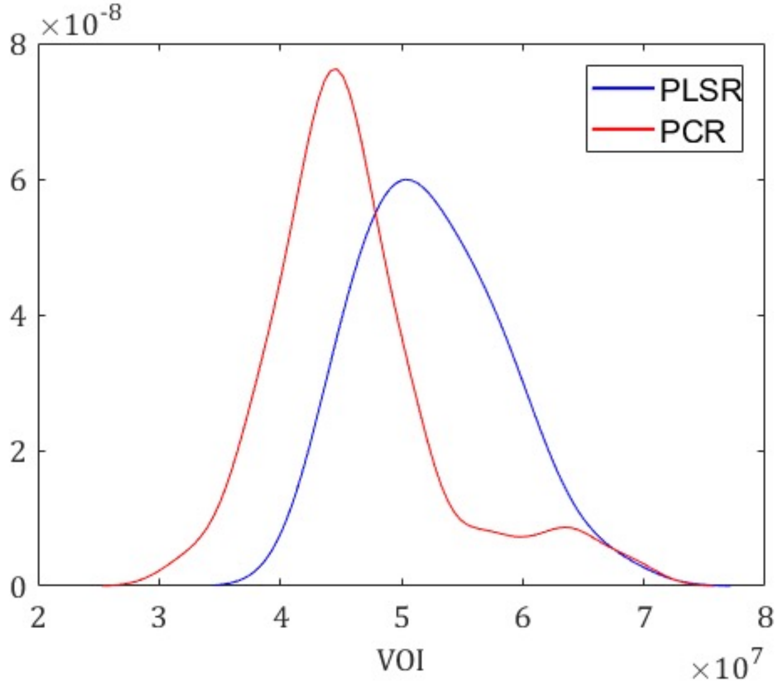
Fit regression model from Monte Carlo samples.
12 regressor components in the PLS regression.

Gullfaks case (predictive performance)

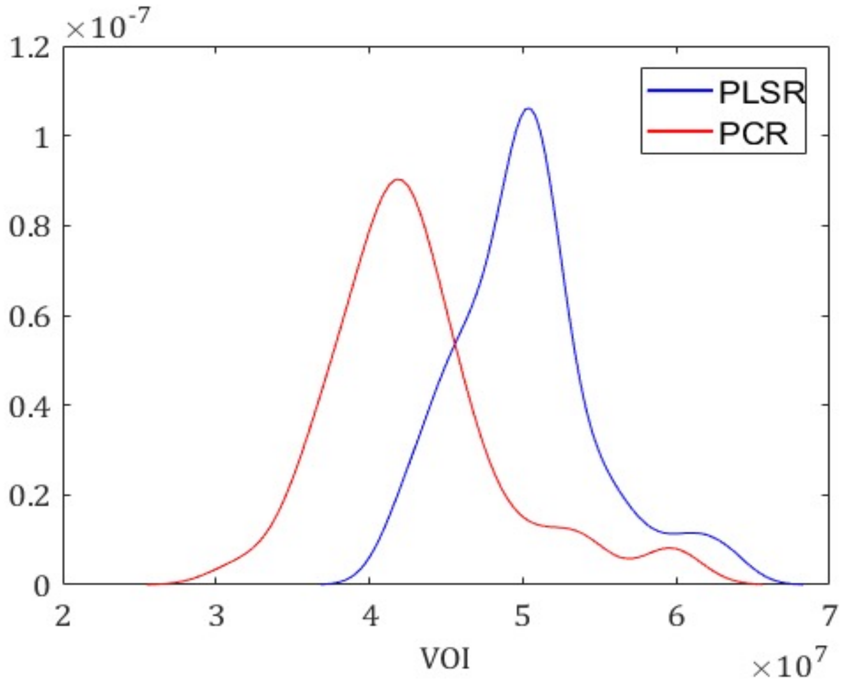


Fit of regression models is reasonable (based on AI data here).

Gullfaks case (VOI results)



Acoustic impedance (AI)



Angle information, (R0,G)

VOI of time-lapse data is about $50 million.
No big differences in VOI of processing methods
(but the price of these likely differ). (Bootstrap used to get distribution.)

EAGE EDUCATION

Take home from Gullfaks case

- VOI for useful time-lapse seismic data gathering plans.
- AVO not necessarily much more informative than AI.

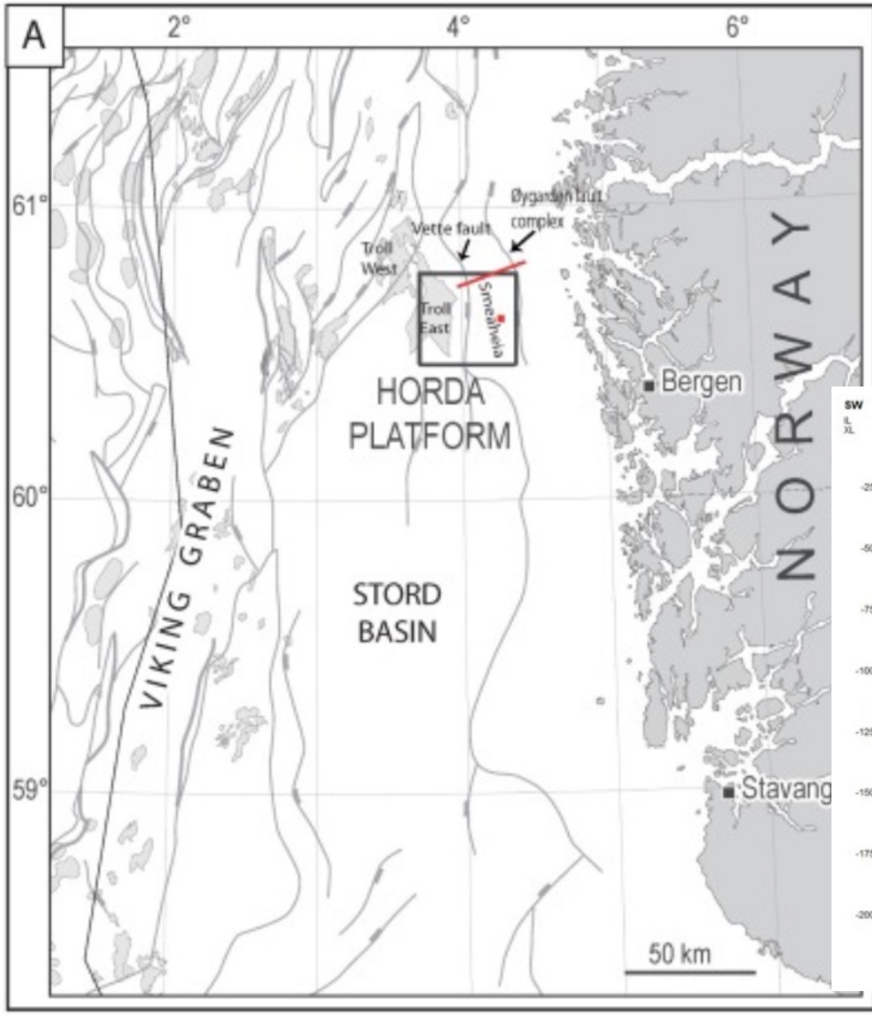
- Frame decision situation - alternatives and uncertainties.
  What is the key question? Here infill drilling plans.

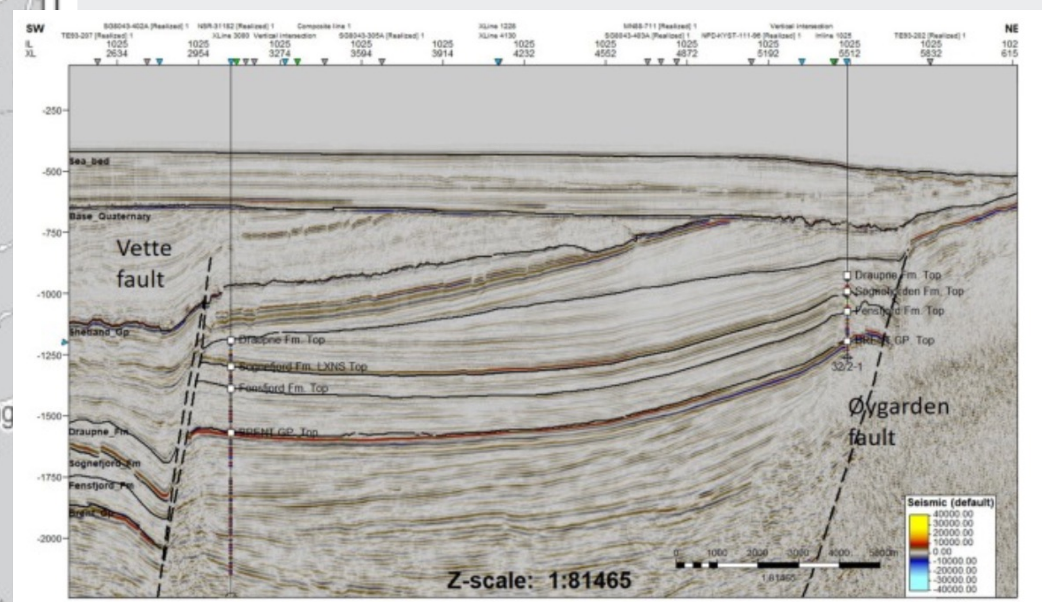
- Computationally difficult - require approximations.
  Simulation-regression : i) generate realizations of values and data,
  ii) fit conditional expectation of values.

Future : Continuous monitoring. (Johan Sverdrup field – digitalization)
But the processing requires calibration , when/where/how is it most valuable.

EAGE
EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Related example on CO2 monitoring:

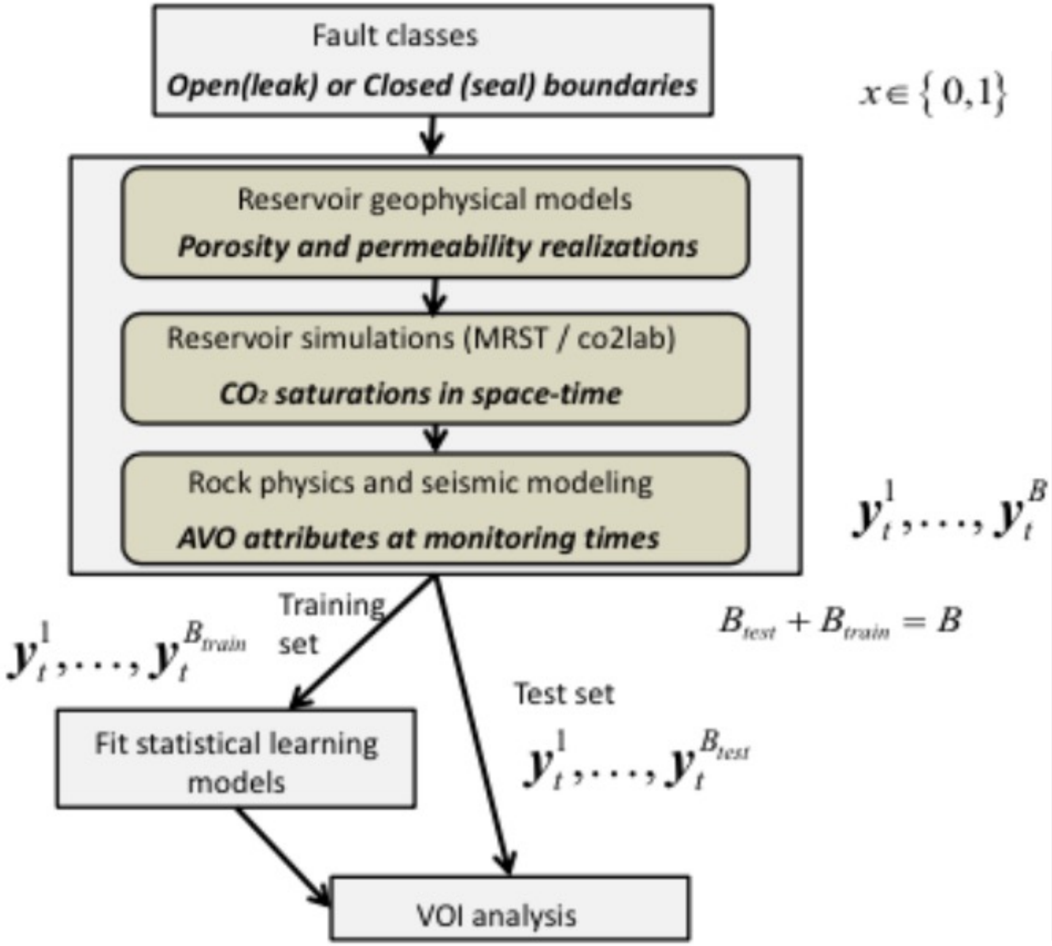




Anyosa et al. (2021), IJGGC.

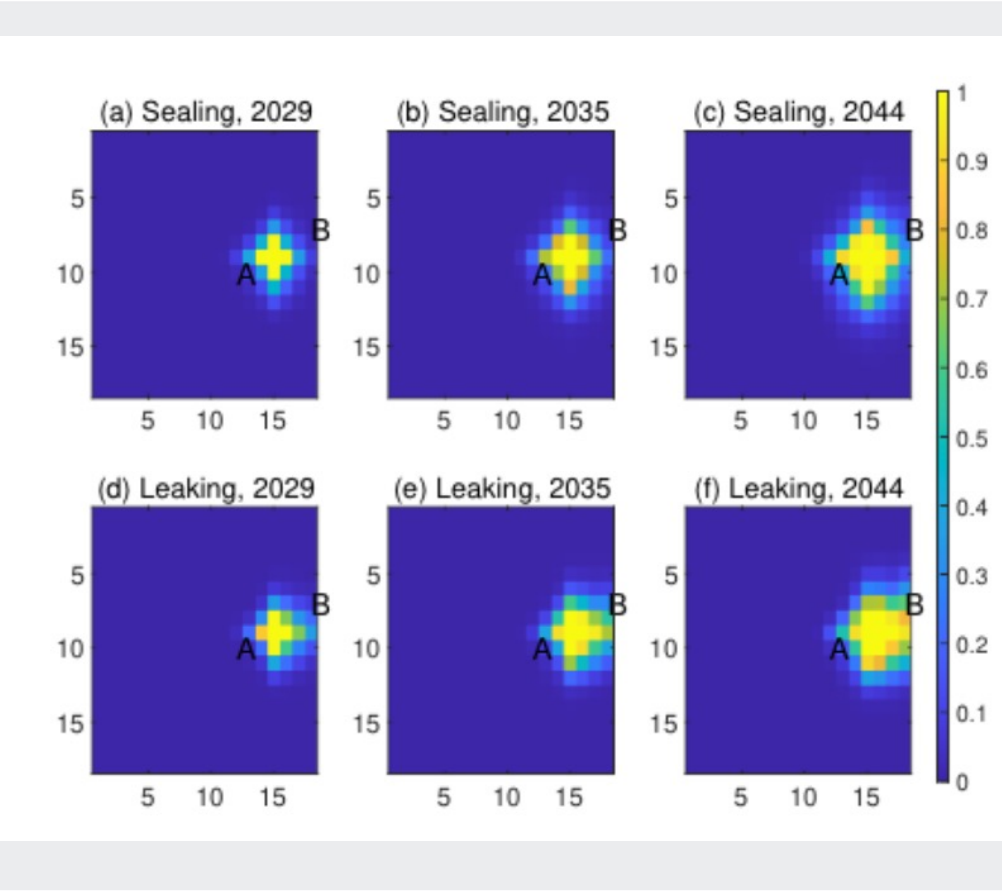
EAGE EDUCATION

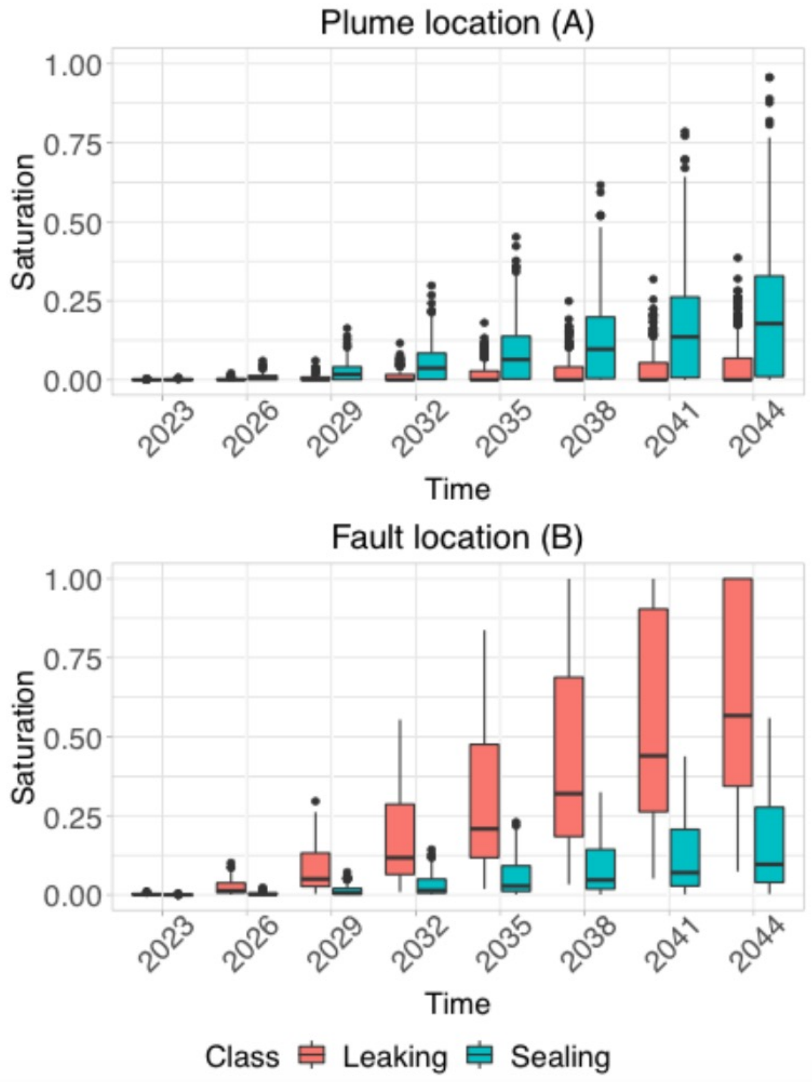
Related example on CO2 monitoring:



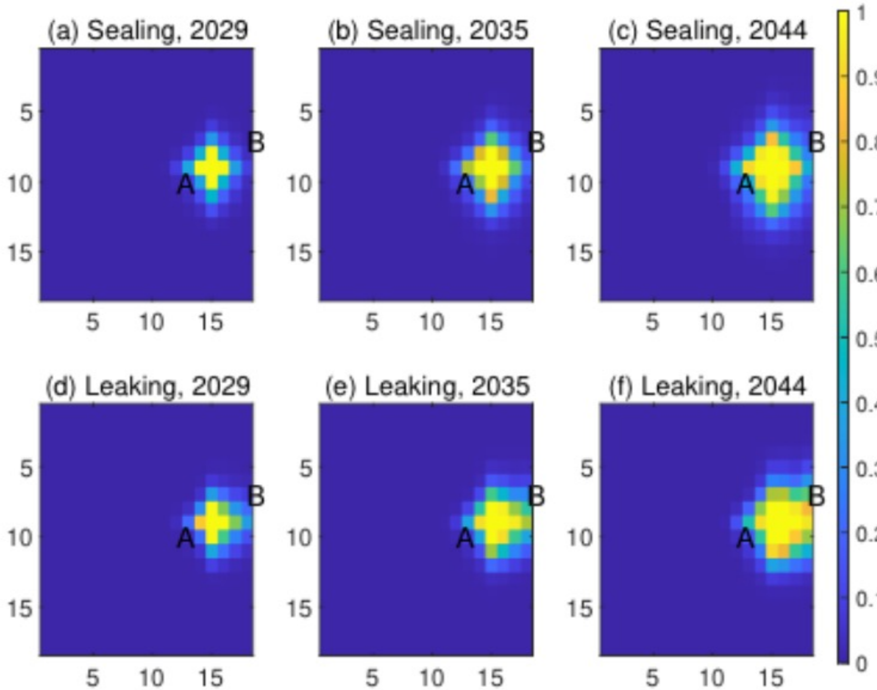
Use simulation and machine learning approaches to approximate the VOI.

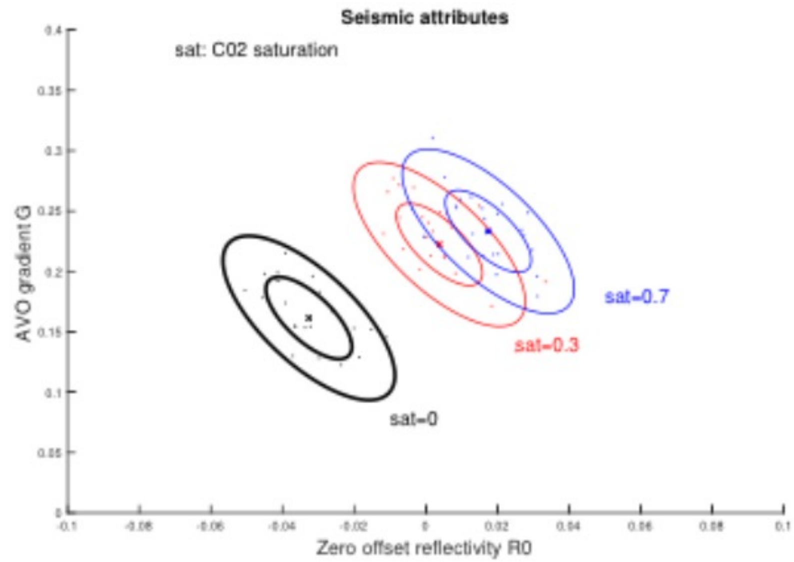
Related example on CO2 monitoring:

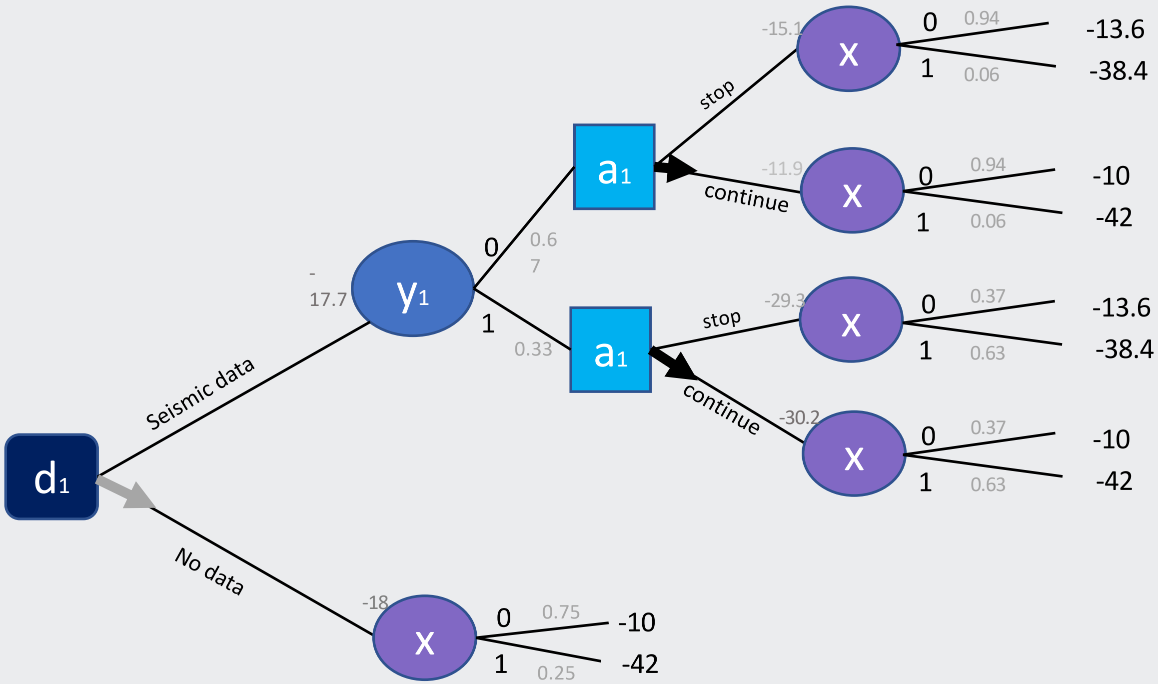




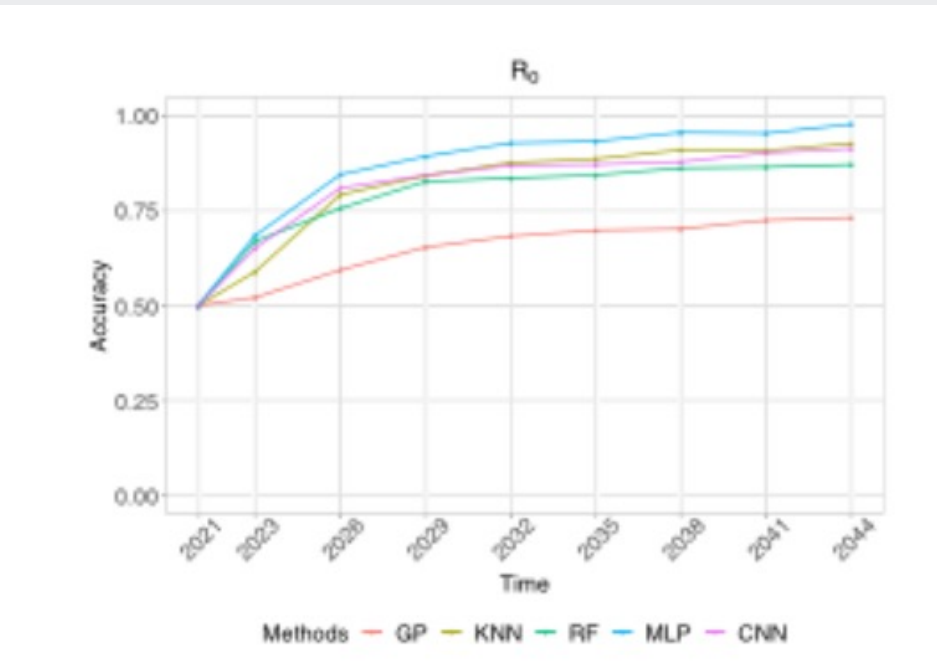
Related example on CO2 monitoring:

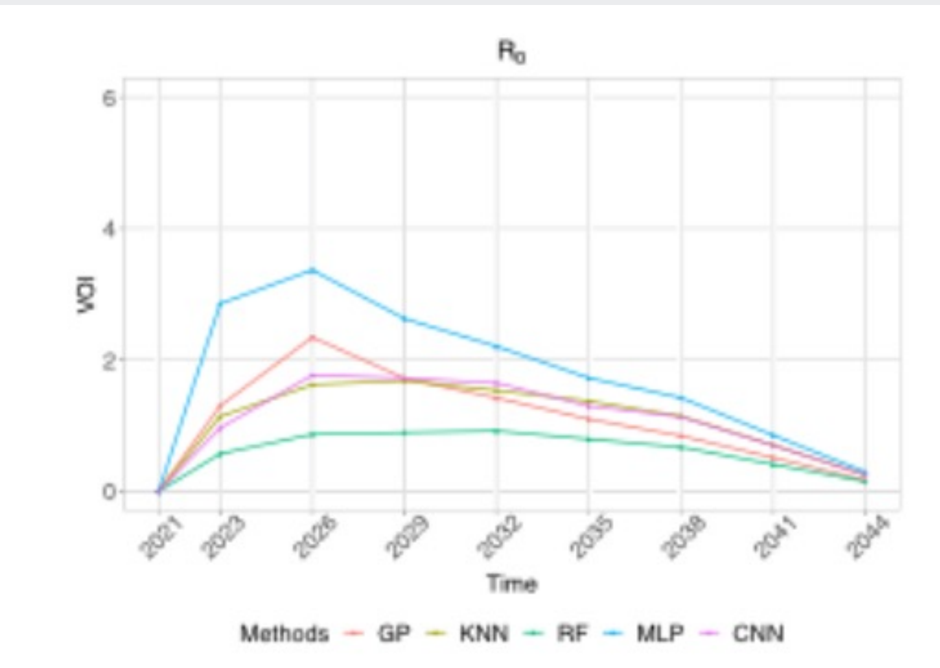





d1
Seismic data
No data
-17.7
y1
0
0.67
1
0.33
a1
a1
stop
continue
stop
continue
-15.1
-11.9
-29.3
-30.2
-18
x
0
0.94
1
0.06
-13.6
-38.4
0
0.94
1
0.06
-10
-42
0
0.37
1
0.63
-13.6
-38.4
0
0.37
1
0.63
-10
-42
0
0.75
1
0.25
-10
-42

Related example on CO_2 monitoring:





Machine learning algorithms to study seismic monitoring times.

Take home from CO_2

- VOI is small early in injection program.
- VOI is small late in program (too late to change decision).

- VOI peaks in-between (10 y)

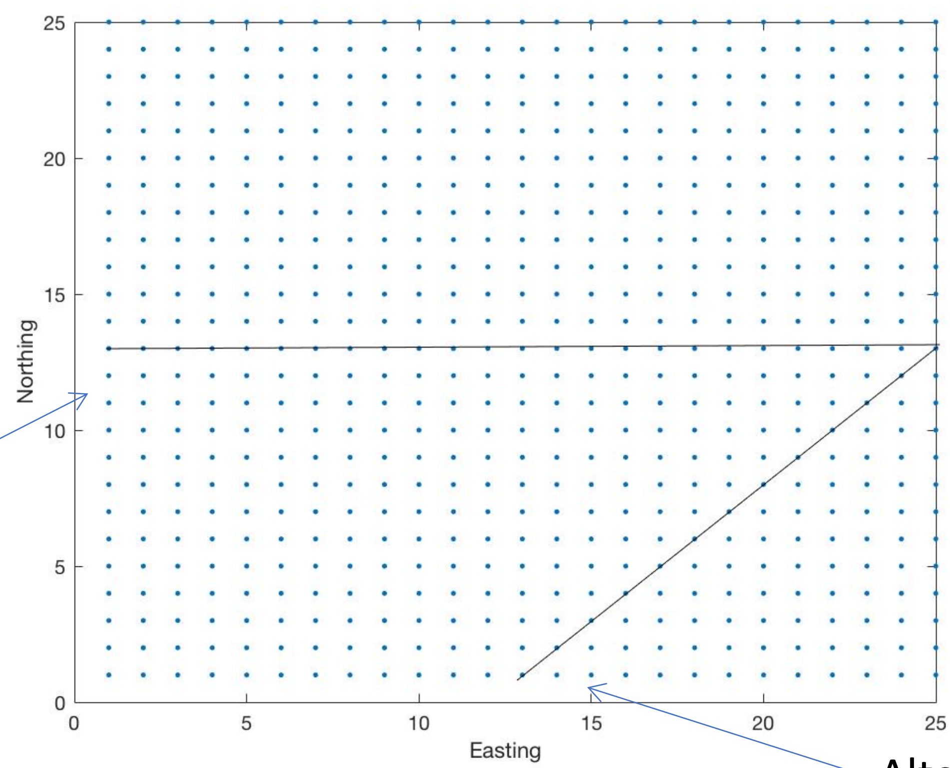
- Trends similar, but actual values depend on regression method

Future : Real time monitoring decision support

Project case : alternatives

Suppose there is a 25 x 25 grid of reservoir variables. We want to flood the reservoir either from the west, or from the south.

Problem:
Where to inject



Project case : models

There is uncertainty in the reservoir properties, possibly a channel with larger permeability in the middle, and some heterogeneity.

We can sample from the model as follows:

$$b = 1,\dots,B$$

- Draw a regression parameter:

$$\beta^b \sim p(\beta) = \mathrm{Gamma}(1,1)$$

- Draw a Gaussian process on the 25 x 25 grid:

$$\boldsymbol{x}_0^b \sim p\left(\boldsymbol{x}_0 \mid \beta^b\right) = N\left(-\beta^b \cdot \frac{(North-13)^2}{144}, \boldsymbol{\Sigma}\right)$$

- Permeability is log-Gaussian:

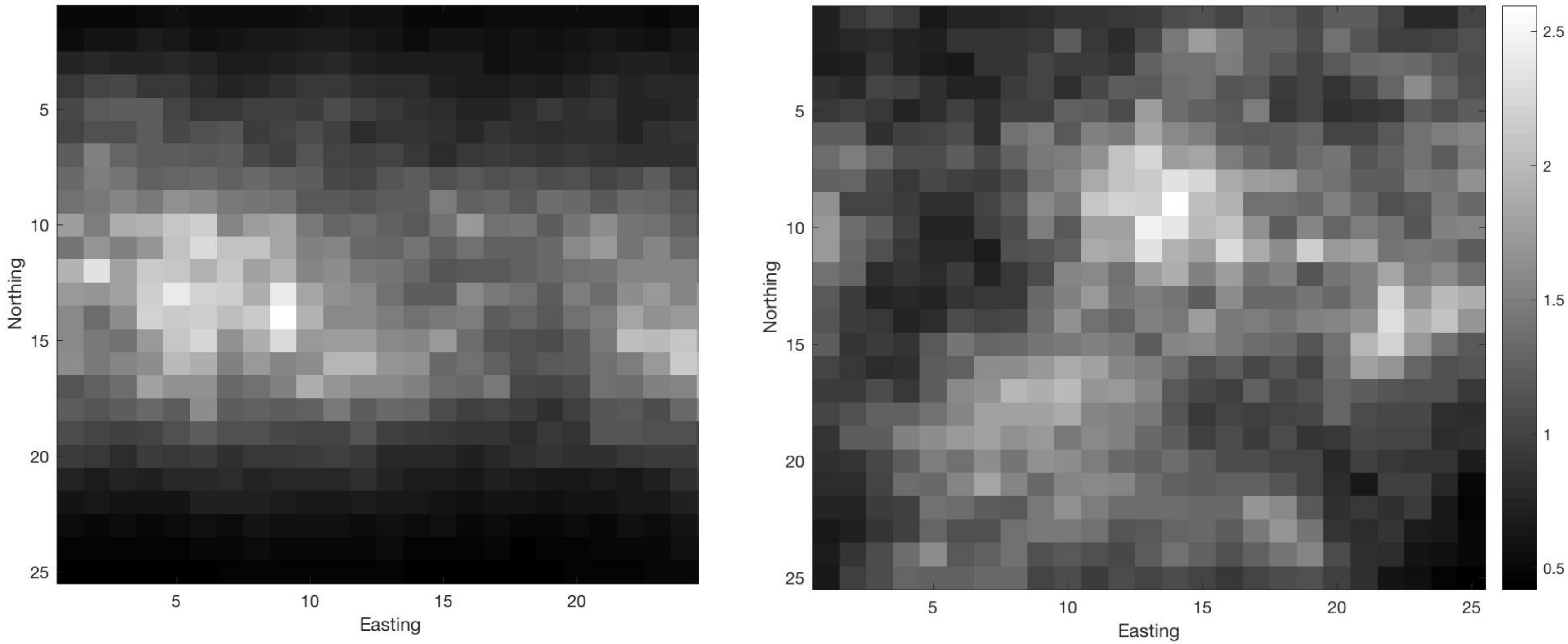
$$\boldsymbol{x}^b = \exp\left(\boldsymbol{x}_0^b\right)$$

Project case : models

$$\Sigma_{ij} = 0.3^2 \exp\left(-\frac{3\left|\mathbf{s}_i - \mathbf{s}_j\right|}{20} \right)$$

Variability due to the regression uncertainty and the spatial heterogeneity.

Two permeability realizations



Project case : values

Value is set as the time-of-flight: time it takes a particle to travel from the injector to the producer. Smaller is better, larger 'value'. (This is used as a proxy for fluid flow.)

For each alternative (west or south), we compute time-of-flight as follows:

$$v_a^b = \sum_{i \in l_a} d_i / x_i$$

$$b = 1, \ldots, B$$

Distance is 1 for 'west' alternative, 1.5 for 'south' alternative.

Sum inverse permeability variables along the line. Large permeability, smaller time of flight.

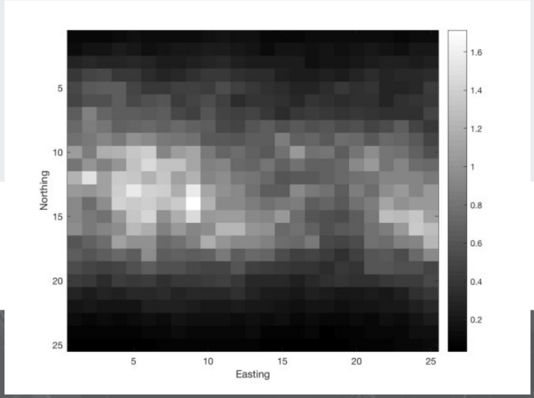
Project case : data

Problem:
Is data helpful?

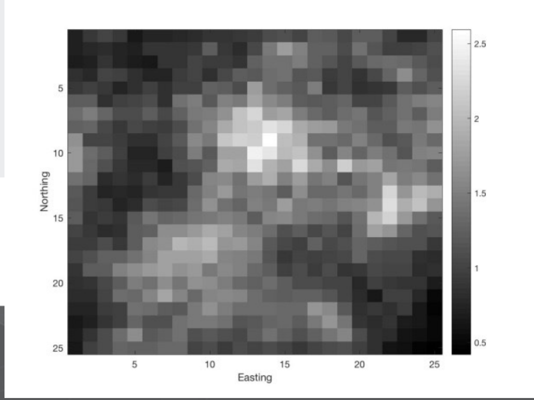
Data is the log-ratio of the variability in the center N-S line compared with the center E-W line. (This might be a result of processing seismic data.)

$$y^b = \log\left[\frac{\frac{1}{24}\sum_{i\in NS}\left(x_i^b - \bar{x}_{NS}^b\right)^2}{\frac{1}{24}\sum_{i\in EW}\left(x_i^b - \bar{x}_{EW}^b\right)^2}\right] \qquad b = 1,\ldots,B$$

Large data.



Small data.



Project case : regression

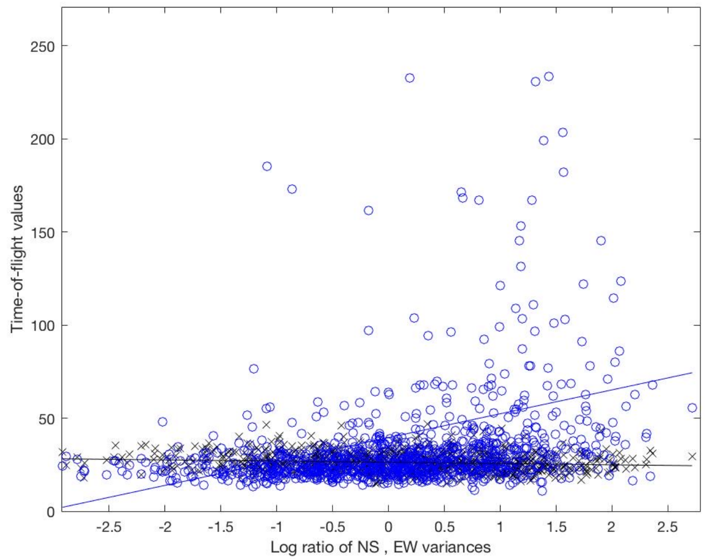
Estimate the conditional expected values by simple linear regression, using the samples of values and data. Do this for both 'west' and 'south' alternative.

$$\hat{E}(v_a \mid y) = \hat{\alpha}_{0,a} + \hat{\alpha}_{1,a} y \qquad \left(y^b, v_a^b\right),$$

$$PoV(\boldsymbol{y}) \approx \frac{1}{B}\sum_{b=1}^{B} \max_{\boldsymbol{a}\in A}\left\{-\hat{E}\left(v_{\boldsymbol{a}} \mid \boldsymbol{y}^b\right)\right\}$$

$$PV \approx \max_{\boldsymbol{a}\in A}\left\{-\frac{1}{B}\sum_{b=1}^{B} v_a^b\right\}$$

$$VOI \approx PoV(\boldsymbol{y}) - \mathrm{PV}$$



Project case : VOI approximation

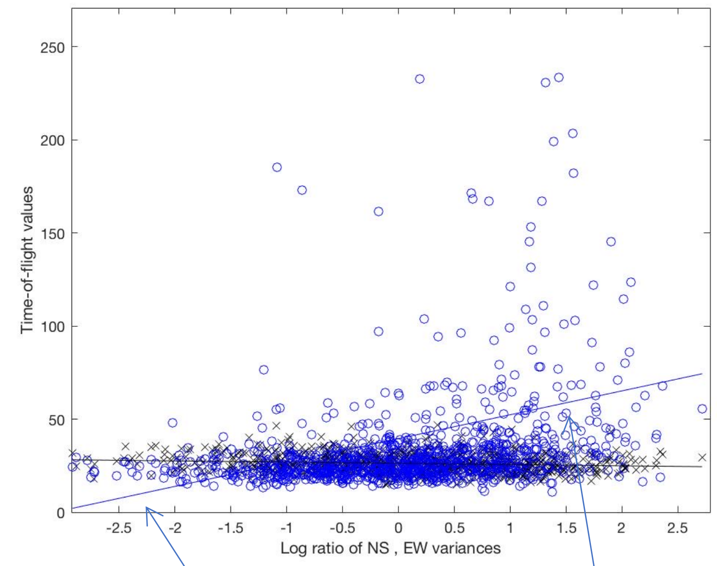
$$\left(y^b, v_a^b\right), \qquad b = 1, \ldots, B$$

$$PV \approx \max_{\boldsymbol{a} \in \boldsymbol{A}} \left\{ -\frac{1}{B} \sum_{b=1}^{B} v_a^b \right\}$$

$$\hat{E}\left(v_a \mid y\right) = \hat{\alpha}_{0,a} + \hat{\alpha}_{1,a} y$$

$$PoV\left(\boldsymbol{y}\right) \approx \frac{1}{B} \sum_{b=1}^{B} \max_{\boldsymbol{a} \in \boldsymbol{A}} \left\{ -\hat{E}\left(v_{\boldsymbol{a}} \mid \boldsymbol{y}^b\right) \right\}$$

$$VOI \approx PoV\left(\boldsymbol{y}\right) - \mathrm{PV}$$



Best alternative seems to depend on high or low data. This should give positive VOI.

Plan for course

Time	Topic
Day 1	Introduction and motivating examples
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Day 2	Examples of value of information analysis in Earth sciences
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.

Formula for VOI

$$PV = \max_{a \in A} \left\{ E\left(v(\boldsymbol{x}, \boldsymbol{a})\right) \right\} = \max_{a \in A} \left\{ \int_{\boldsymbol{x}} v(\boldsymbol{x}, \boldsymbol{a}) p(\boldsymbol{x}) d\boldsymbol{x} \right\}$$

$$PoV(\boldsymbol{y}) = \int \max_{a \in A} \left\{ E\left(v(\boldsymbol{x}, \boldsymbol{a}) \mid \boldsymbol{y}\right) \right\} p(\boldsymbol{y}) d\boldsymbol{y}$$

$$VOI(\boldsymbol{y}) = PoV(\boldsymbol{y}) - PV.$$

The analysis is usually done for **static decisions** and **static tests**:

- We make the one-time decisions here and now.
- We can only collect the data here and now.

Sequential decisions or **sequential tests** can give benefits.

Sequential decision (prior value)

.

$$PV_{seq} = \max_{i_1} \left\{ E\left(v\left(x_{i_1}, a_{i_1} = 1\right)\right) + \int_{x_{i_1}} \max_{i_2 \neq i_1} \left\{0, \mathrm{ContVal}_{i_2}\left(x_{i_1}\right)\right\} p\left(x_{i_1}\right) dx_{i_1}, 0 \right\}$$

$$\mathrm{ContVal}_{i_2}\left(x_{i_1}\right) = E\left(v\left(x_{i_2}, a = 1\right) \mid x_{i_1}\right) + \int_{x_{i_2}} \max_{i_3 \neq i_1, i_2} \left\{0, \mathrm{ContVal}_{i_3}\left(x_{i_1}, x_{i_2}\right)\right\} p\left(x_{i_2} \mid x_{i_1}\right) dx_{i_2}$$

$$\mathrm{ContVal}_{i_n}\left(x_{i_1}, \ldots, x_{i_{n-1}}\right) = \max\left\{0, E\left(v\left(x_{i_n}, a_{i_n} = 1\right) \mid x_{i_1}, \ldots, x_{i_{n-1}}\right)\right\}$$

Solution is a discrete optimization method called dynamic programming:

- Go through all possible decision paths, moving forward.
- Find the optimal values winding backwards in the tree of paths .

This is costly for large systems, and several approximations exists.
Suboptimal strategies, using heuristics, are often used in practice, with success.

Sequential decisions (VOI)

.

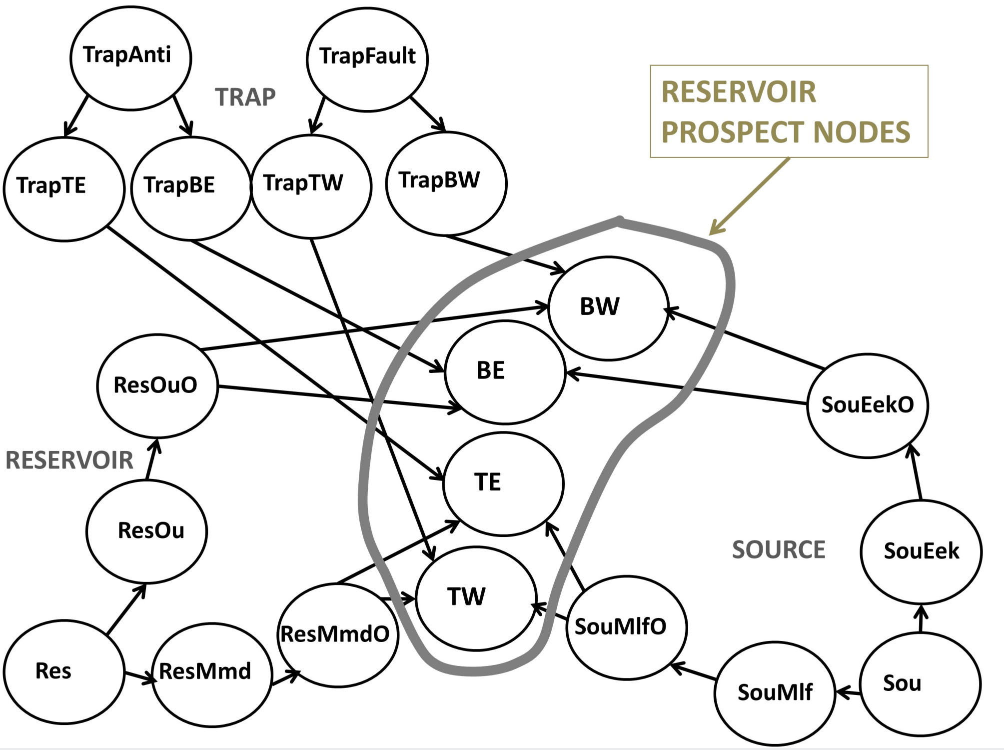
$$PV_{seq} = \max_{i_1}\left\{E\left(v\left(x_{i_1}, a_{i_1}=1\right)\right) + \int_{x_{i_1}} \max_{i_2 \neq i_1}\left\{0, \mathrm{ContVal}_{i_2}\left(x_{i_1}\right)\right\} p\left(x_{i_1}\right) dx_{i_1}, 0\right\}$$

$$PoV_{seq}(\boldsymbol{y}) = \int_{\boldsymbol{y}} \max_{i_1}\left\{\begin{array}{l} 0, E\left(v\left(x_{i_1}, a=1\right) \mid \boldsymbol{y}\right) + \\ \int_{x_{i_1}} \max_{i_2 \neq i_1}\left\{0, \mathrm{ContVal}_{i_2}\left(x_{i_1}, \boldsymbol{y}\right)\right\} p\left(x_{i_1} \mid \boldsymbol{y}\right) dx_{i_1} \end{array}\right\} p(\boldsymbol{y}) d\boldsymbol{y}$$

Must now use dynamic programming for different data outcomes.

VOI is difference between posterior and prior value.

Exploration drilling example

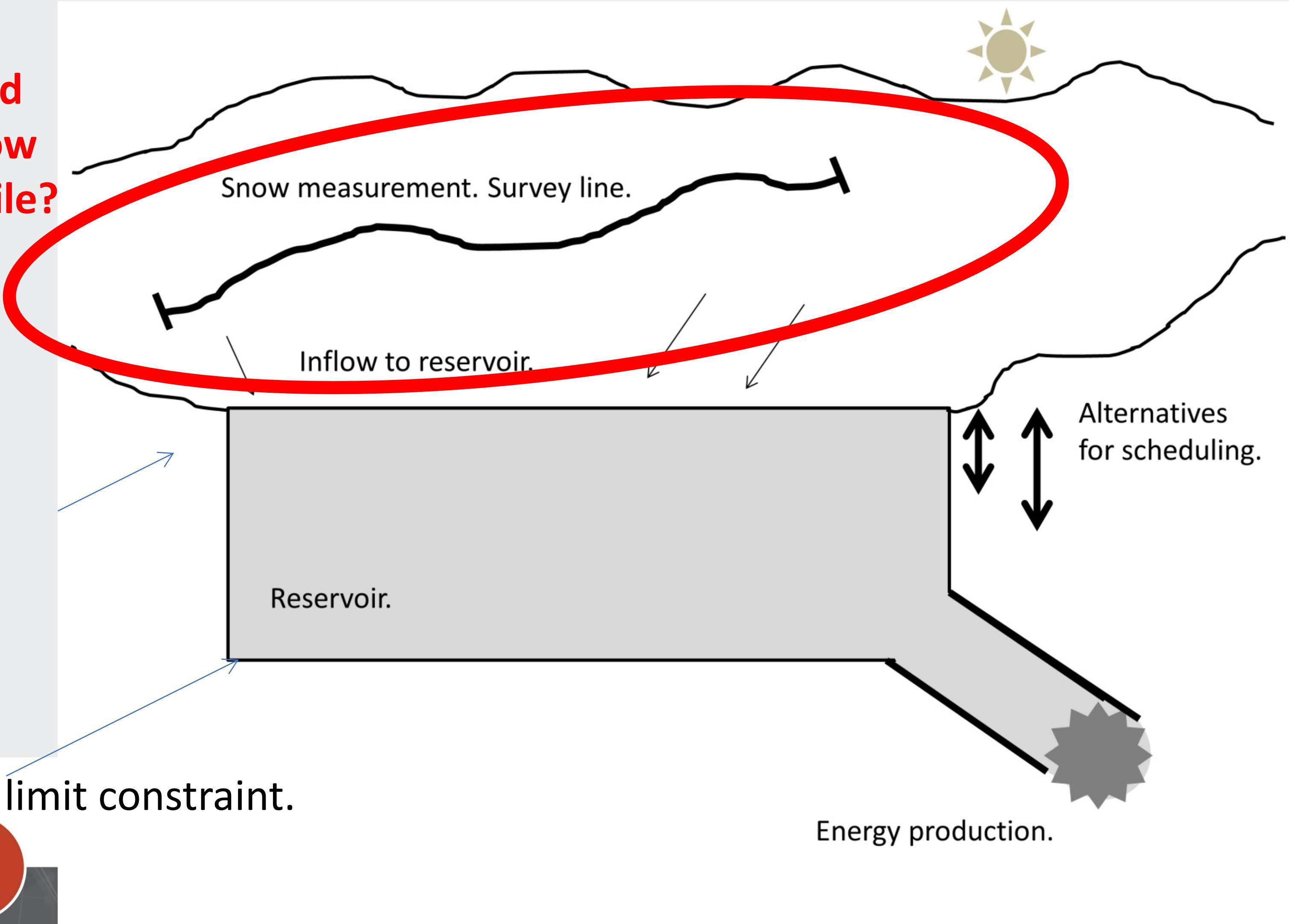


Optimal decision can be solved exactly when there are only a few variables (small size problem).

Otherwise heuristic solutions are sought.

Hydropower

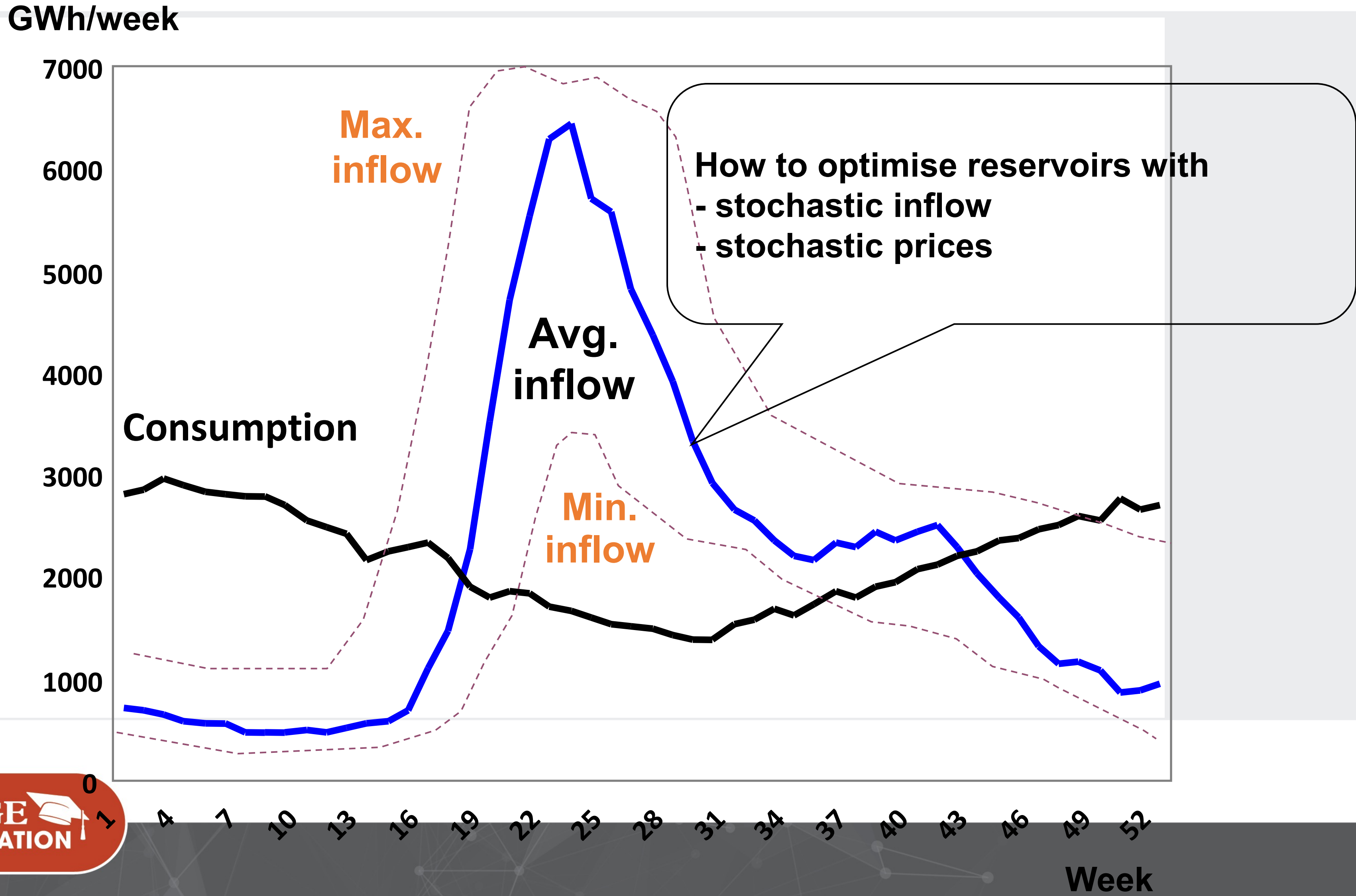
Is acquiring and processing snow data worthwhile?



No money for water going over the dam.

Lower limit constraint.

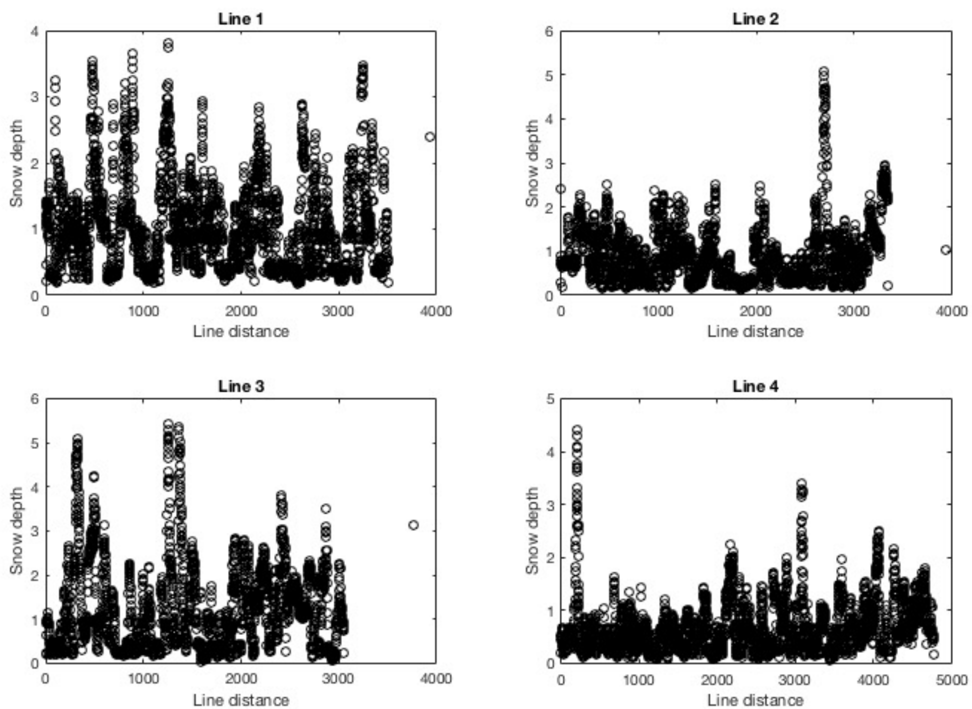
Inflow and hydro scheduling



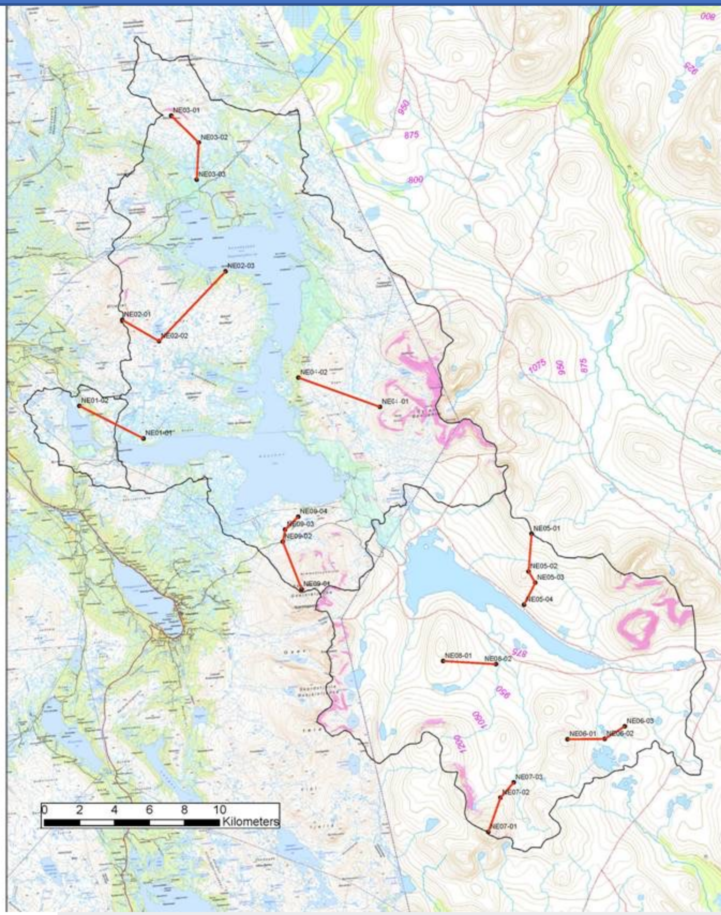
EAGE
EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Snow data

Spatio-temporal datasets.



Map view



- Data is snow level : {low, medium, high} for base case.
  (Dramatic reduction of data dimension.)

EAGE EDUCATION

Scheduling (with information)

$$PoV_{seq}(\boldsymbol{y}) = \int_{\boldsymbol{y}} \max_{a_1} \left\{ \begin{array}{l} E\left(v\left(x_1, a_1\right) \mid \boldsymbol{y}\right) + \\ \int_{x_1} \max_{a_2} \left\{ \text{ContVal}_2\left(x_1, \boldsymbol{y}\right) \right\} p\left(x_1 \mid \boldsymbol{y}\right) dx_1 \end{array} \right\} p(\boldsymbol{y}) d\boldsymbol{y}$$

$$VOI(\boldsymbol{y}) = PoV_{seq}(\boldsymbol{y}) - PV_{seq}$$

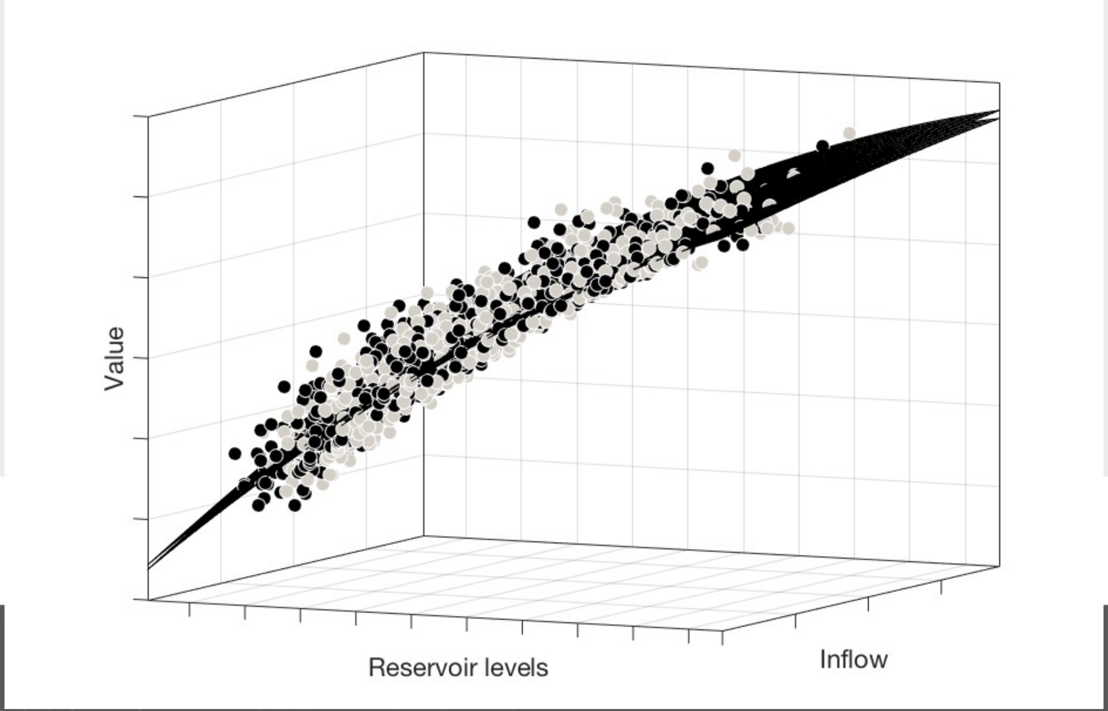
- Alternative is control variable (production).
- Uncertainty is inflow, represented by scenarios.

- Data is snow level : {low, medium, high} for base case.

Least squares Monte Carlo

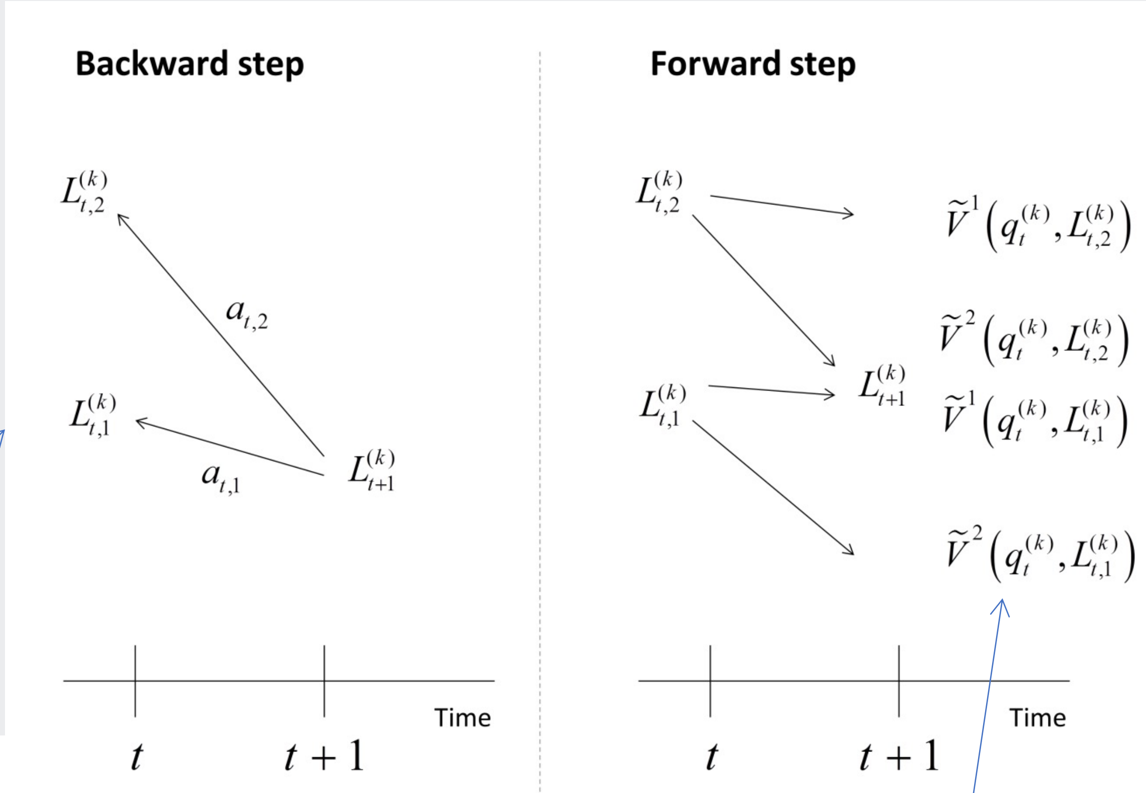
Algorithm:

- Simulate inflow **scenarios** (10 000 models from data and time series fitting)
- Wind-up optimal solution backward in time, by **least-squares fitting** of surfaces from simulated values as a function of inflow and reservoir level.
- Optimal **controls are decided by forward selection**, based on largest regression surface for current reservoir level and inflow.



Simulation regression

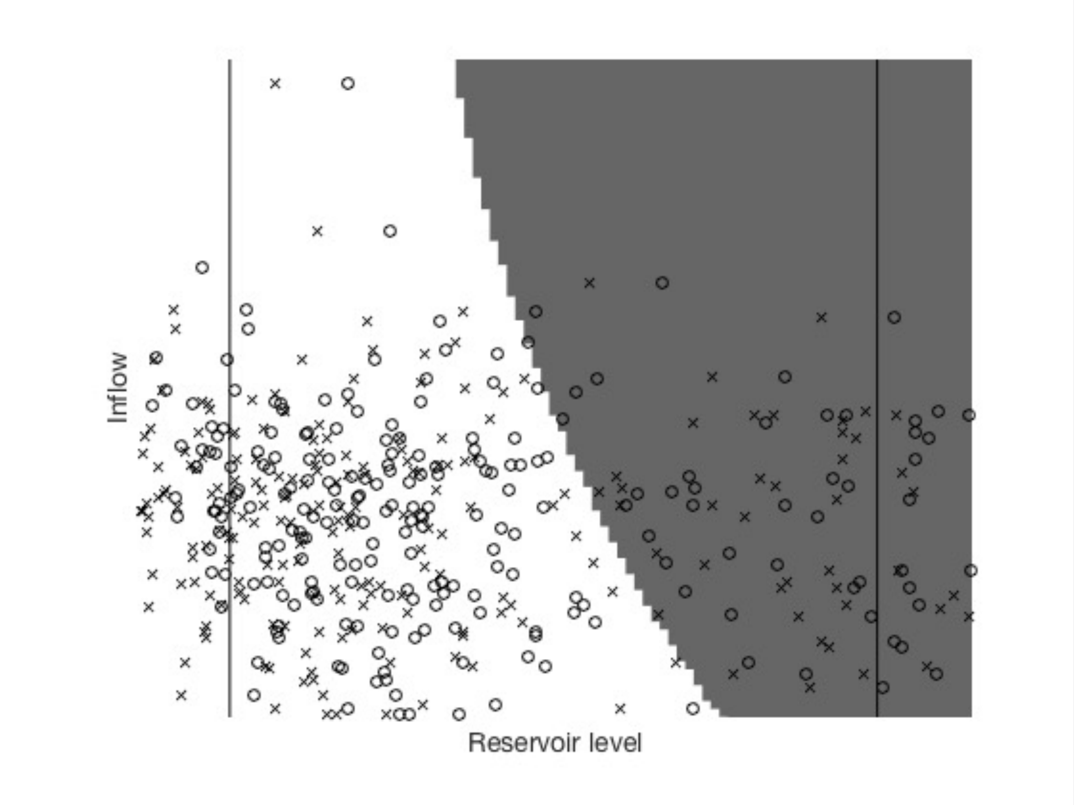
Linking forward runs and backward algorithm using Markov assumption.



Reservoir levels:

Value is function of inflow and reservoir level.

Regression surfaces

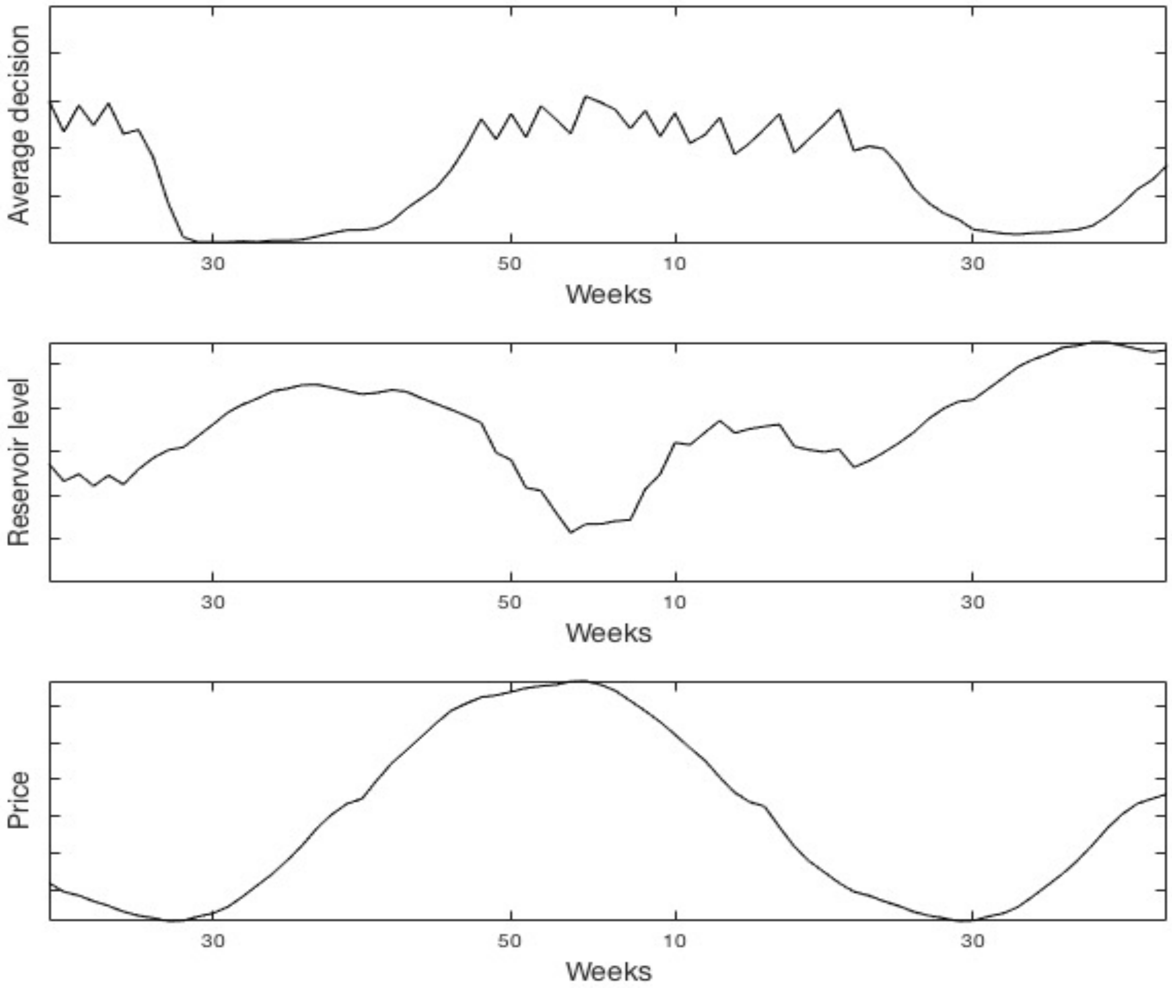


Highest surface sets control.

Quadratic components in regression fitting.

Results – scheduling (prior)

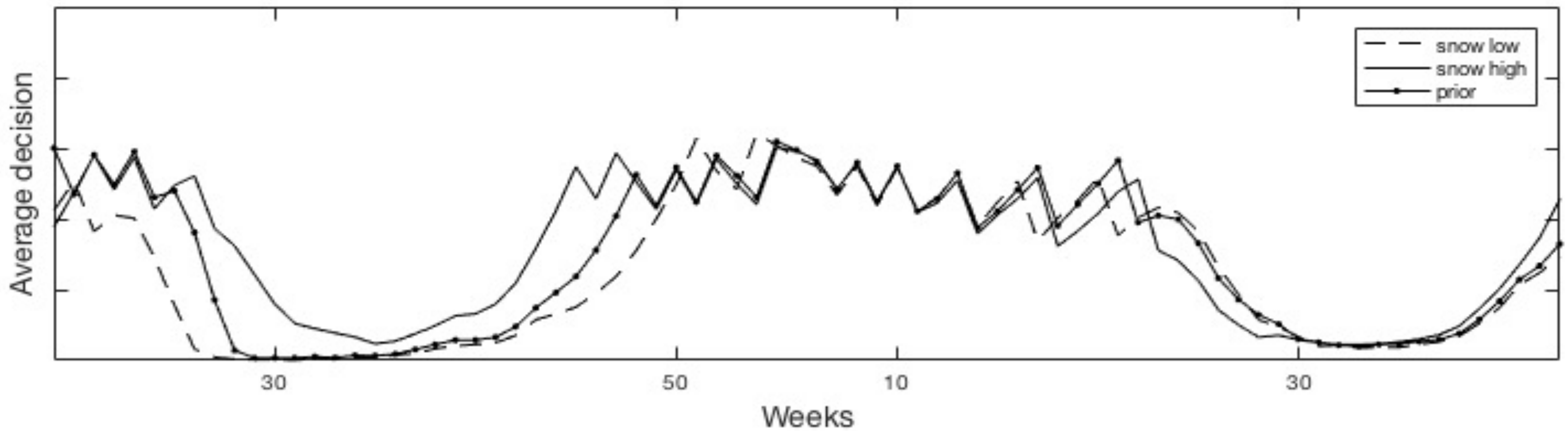
Base case has two controls.

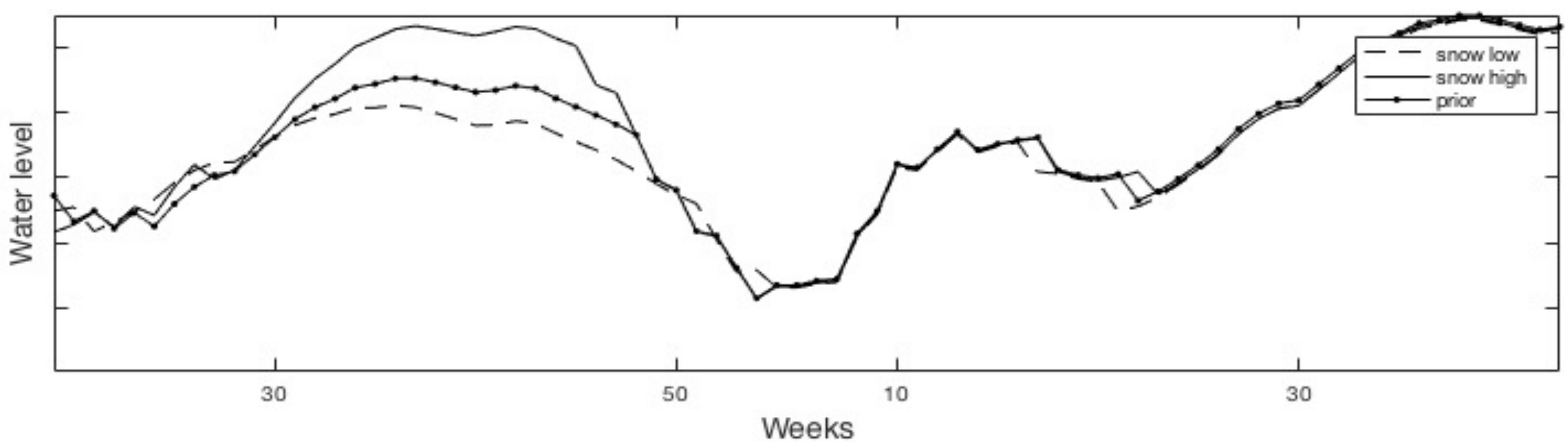


Scheduling with snow data

Data split scenarios in sub-groups.

Base case has three levels for snow measurements → three scenario groups.





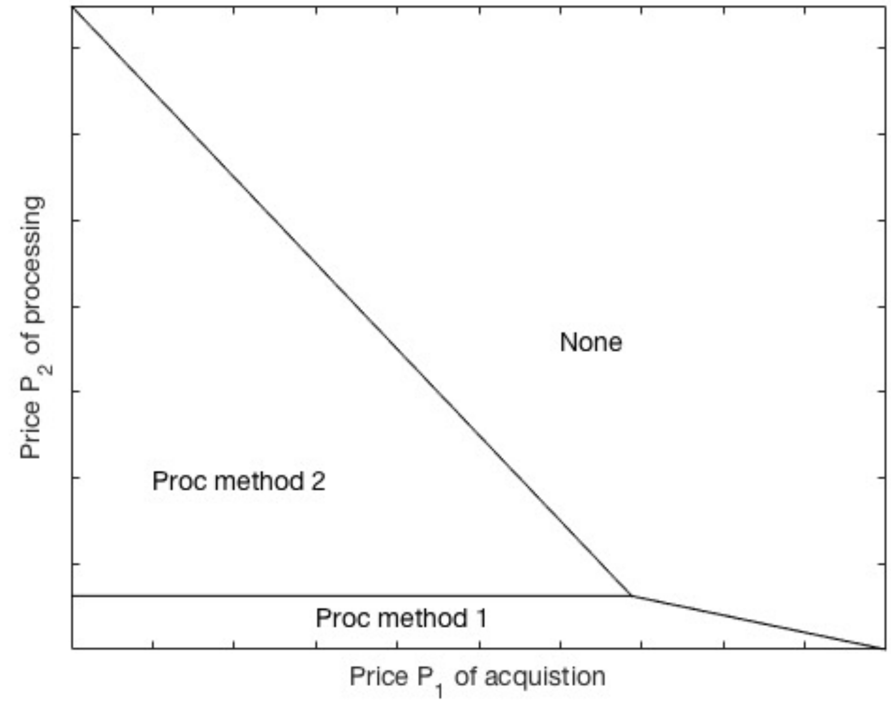
VOI (relative numbers)

	3 snow classes	6 snow classes	9 snow classes	12 snow classes
Two production controls	0	0	9	21
Four production controls	0	23	47	100

More discrimination in snow measurement – larger VOI.

More flexibility in production controls – larger VOI
(of course also depends on time - here week)

VOI and decision regions



Data costs can be divided in acquisition and processing (type 1 or 2 – better or worse).

$$\arg\max\left\{0, VOI_A - P_1 - P_2, VOI_B - P_1 - P_2/5\right\}$$

Take home

- If reservoir is large relative to inflow – not much added value of snow measurements.

- If reservoir is relatively small, and large risk of overflow or small future production (lower limit), there is high value of snow measurements.

- More accurate processing of snow data gives larger VOI.
  Both with more classes, and less mis-classification (imperfect information).

- VOI at 10 % of prior value is significant if there is much production.

Information gathering

	Perfect	Imperfect
Total	Exact observations are gathered for all variables. This is rare, occurring when there is extensive coverage and highly accurate data gathering. $$\boldsymbol{y} = \boldsymbol{x}$$	Noisy observations are gathered for all variables. This is common in situations with remote sensors with extensive coverage, e.g. seismic, radar, satellite data. $$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\varepsilon}$$
Partial	Exact observations are gathered at some variables. This might occur, for instance, when there is careful analysis of some samples, or sensors at some locations. $$\boldsymbol{y}_{\mathrm{K}} = \boldsymbol{x}_{\mathrm{K}}, \quad \mathrm{K} \ \text{subset}$$	Noisy observations are gathered at some variables. Examples include noisy sensors for local monitoring, testing along a survey line, etc. $$\boldsymbol{y}_{\mathrm{K}} = \boldsymbol{x}_{\mathrm{K}} + \boldsymbol{\varepsilon}_{\mathrm{K}}, \quad \mathrm{K} \ \text{subset}$$

Could also have **sequential** (adaptive) information gathering.

Sequential information gathering

Decision maker has the opportunity of dynamic testing, where one can stop testing, or continue testing, depending on the currently available data. The sequential order of tests and the number of tests also depend on the data.

Continue testing.

$$PoV_{seqtest}(\mathbf{y}_1) = \int \max \left\{ \begin{array}{l} \max_{j \neq 1} \left\{ CV(j \mid 1) \right\}, \\ \sum_{i=1}^{n} \max_{a_i} \left\{ E\left(v\left(x_i, a_i\right) \mid \mathbf{y}_1\right) \right\} \end{array} \right\} p(\mathbf{y}_1) d\mathbf{y}_1$$

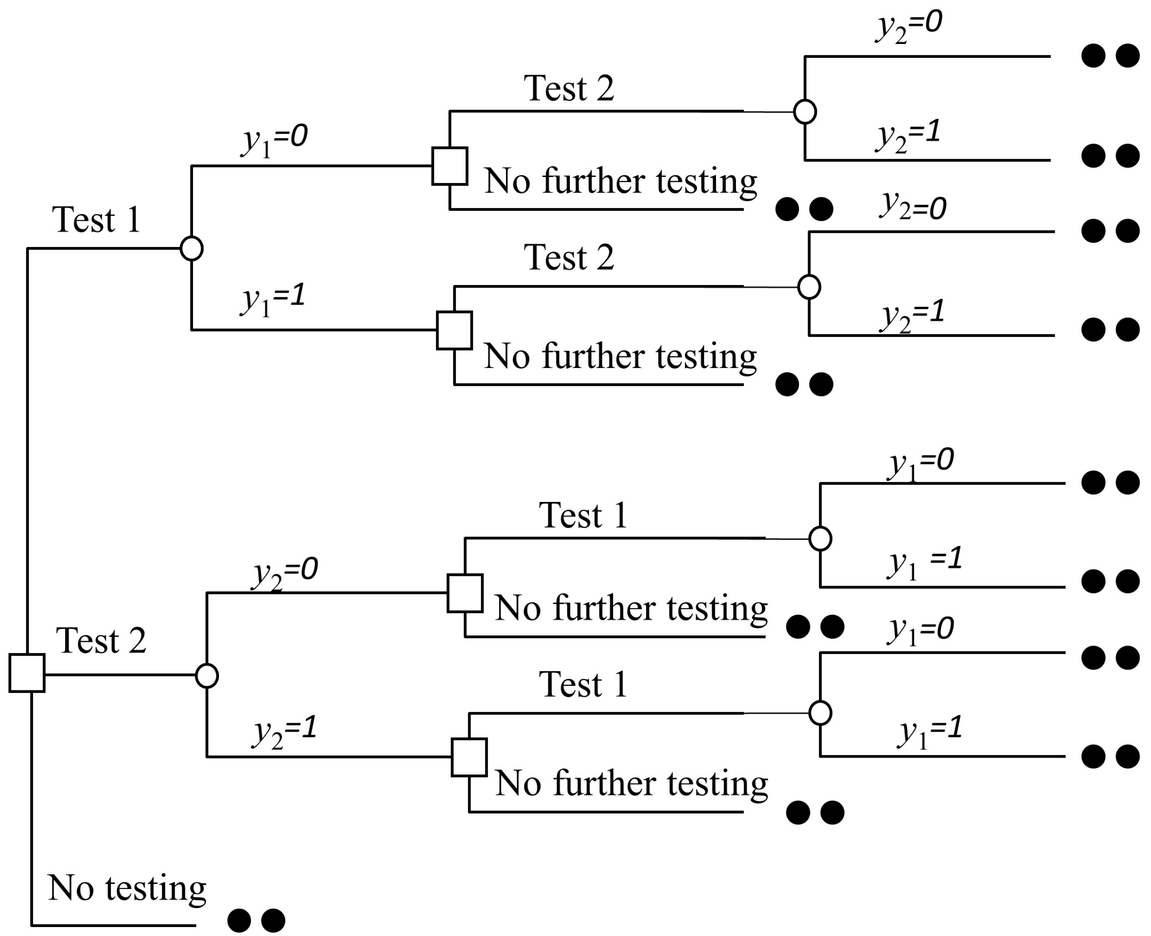
Stop testing.

$$CV(j \mid 1) = Cont(j \mid 1) - P_j$$

Price test.

$$Cont(j \mid 1) = \int \max \left\{ \begin{array}{l} \max_{k \neq 1, j} \left\{ CV(k \mid j, 1) \right\}, \\ \sum_{i=1}^{n} \max_{a_i} \left\{ E\left(v\left(x_i, a_i\right) \mid \mathbf{y}_1, \mathbf{y}_j\right) \right\} \end{array} \right\} p(\mathbf{y}_j \mid \mathbf{y}_1) d\mathbf{y}_j$$

Sequential testing– bivariate illustration



Sequential information (bivariate data)

.

Value with no more testing (after first test):

$$PoV(\mathbf{y}_1) = \int \sum_{i=1}^{n} \max_{a_i \in A_i} \left\{ E\left(v(x_i, a_i) \mid \mathbf{y}_1 \right) \right\} p(\mathbf{y}_1) d\mathbf{y}_1$$

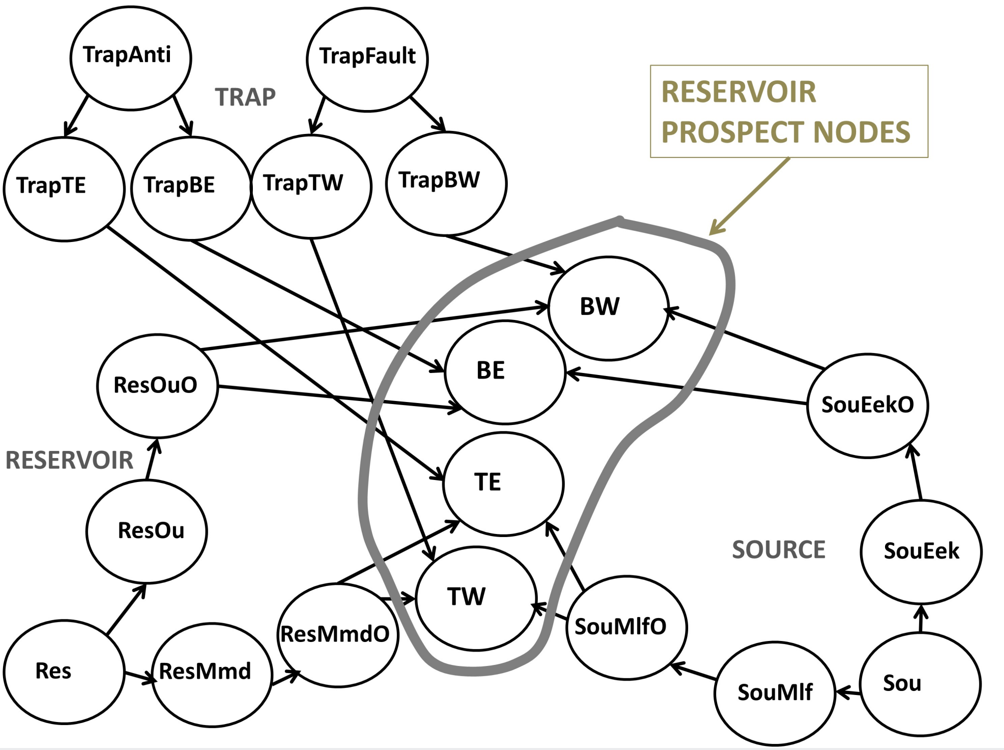
Criterion for continued testing:

$$\int_{\mathbf{y}_2} \sum_{i=1}^{n} \max_{a_i \in A_i} \left\{ E\left(v(x_i, a_i) \mid \mathbf{y}_1, \mathbf{y}_2 \right) \right\} p(\mathbf{y}_2 \mid \mathbf{y}_1) d\mathbf{y}_2 - P_2 > \sum_{i=1}^{n} \max_{a_i \in A_i} \left\{ E\left(v(x_i, a_i) \mid \mathbf{y}_1 \right) \right\}$$

$$PoV_{seqtest}(\mathbf{y}_1) = \int \max \left\{ \begin{array}{l} \int_{\mathbf{y}_2} \sum_{i=1}^{n} \max_{a_i} \left\{ E\left(v(x_i, a_i) \mid \mathbf{y}_1, \mathbf{y}_2 \right) \right\} p(\mathbf{y}_2 \mid \mathbf{y}_1) d\mathbf{y}_2 - P_2, \\ \sum_{i=1}^{n} \max_{a_i} \left\{ E(v(x_i, a_i) \mid \mathbf{y}_1) \right\} \end{array} \right\} p(\mathbf{y}_1) d\mathbf{y}_1$$

Continue testing when the additional expected value of more testing excceds the price.

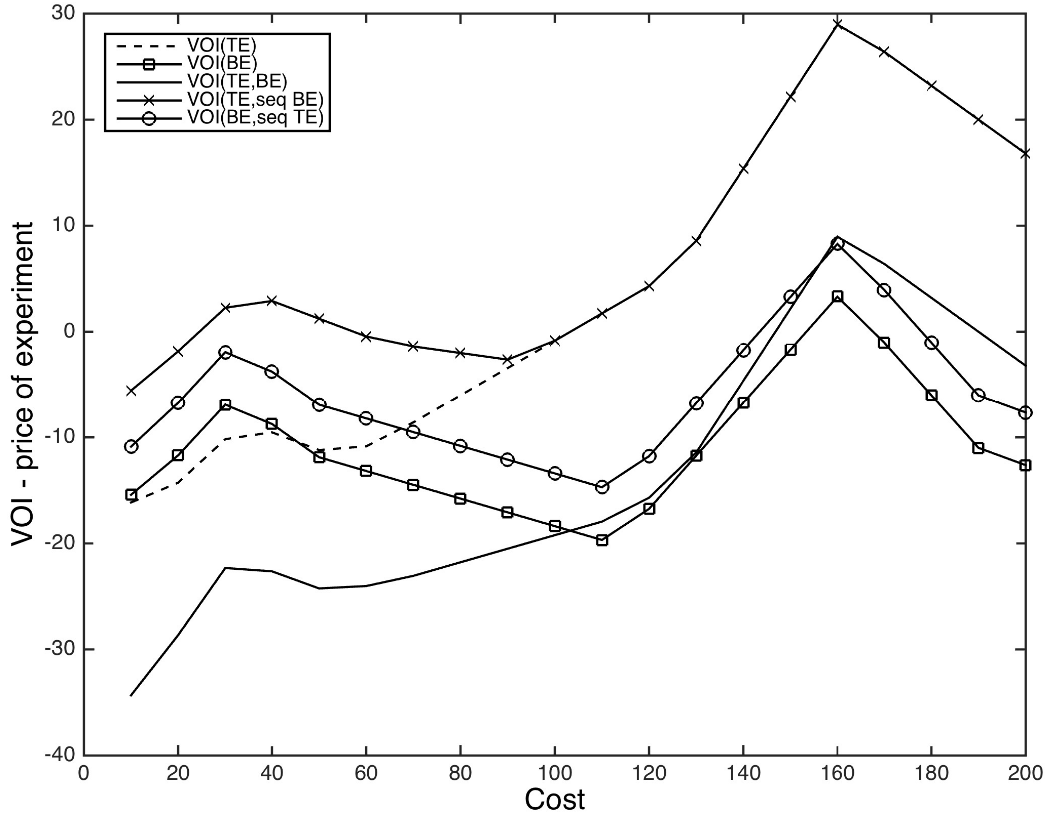
Exploration drilling example



Optimal sequential information gathering can be solved exactly when there are only a few variables.

Exploration drilling example

.



Optimal sequential information gathering can be solved exactly when there are only a few variables.

Myopic strategy for information

Myopic (near-sighted) is a common strategy for sequential problems.
It considers only one-stage at a time, not looking into the 'future'.

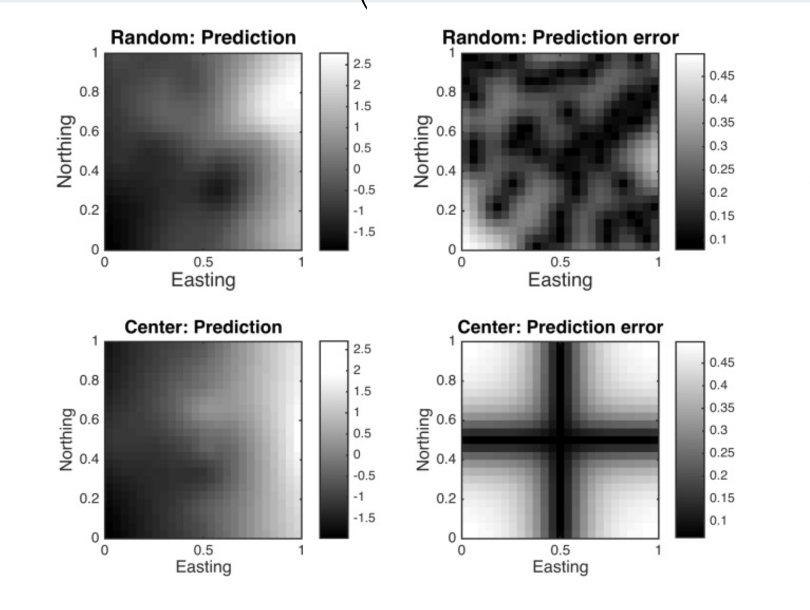
1. Find best first data design, using one-stage, if any give positive VOI. **1 level**
2. Collect data (by simulation) using best design.
3. Update probability distributions, conditional on the data.
4. Find second best data, using one-stage, in new model, if any give positive VOI.
5. Collect second data (by simulation from new model) using best design. **2 level**
6. Update probability distributions, conditional on the data.
7. Find third best data, using one-stage, in new model, if any give positive VOI.

3 level

.....

Sequential line testing(Kriging)

$$p(x \mid y) = N\left(\mu + \Sigma F^t \left(F\Sigma F^t + \tau^2 I\right)^{-1} \left(y - F\mu\right), \Sigma - \Sigma F^t \left(F\Sigma F^t + \tau^2 I\right)^{-1} F\Sigma\right)$$



Which NS line to select

Problem:

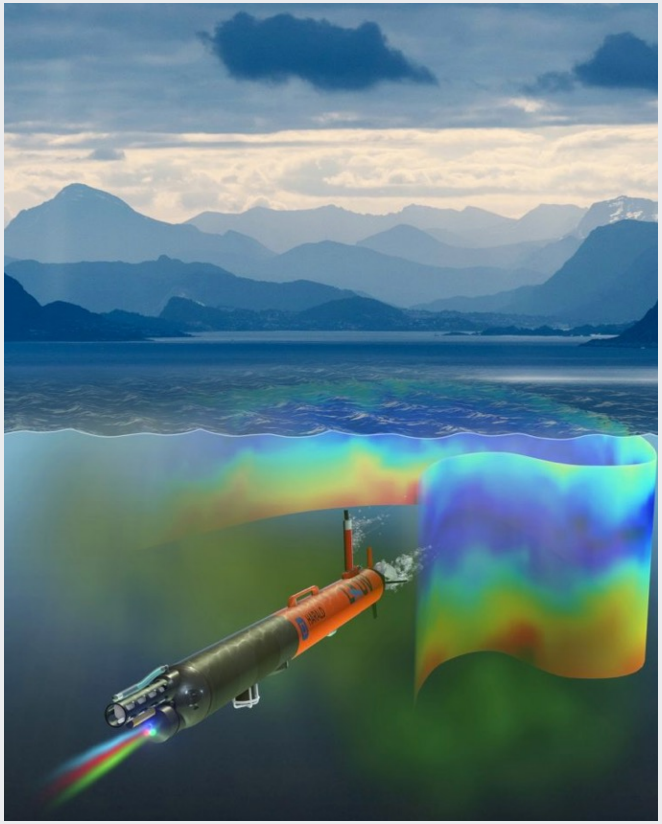
1. **Compute the VOI of NS lines.**
2. **Update in myopic way of thinking.**

(MATLAB)

AUV exploration

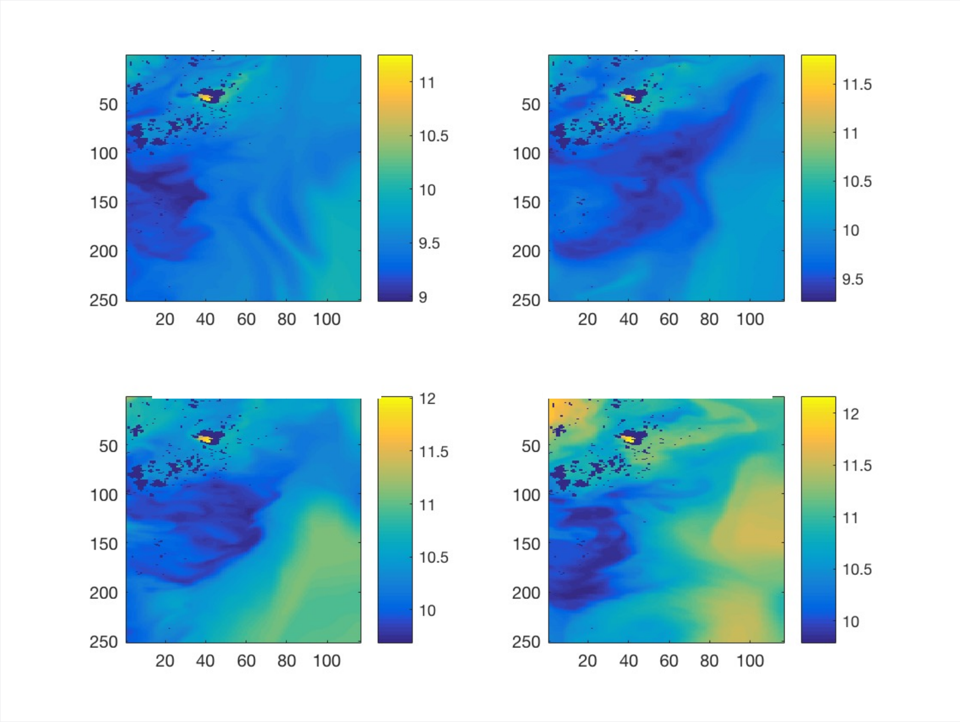
Atonomous sampling to monitor environment. (myopic strategy).

Limited computing resources onboard. Use Gaussian process representation of spatial variables.
This is updated when more data are gathered.



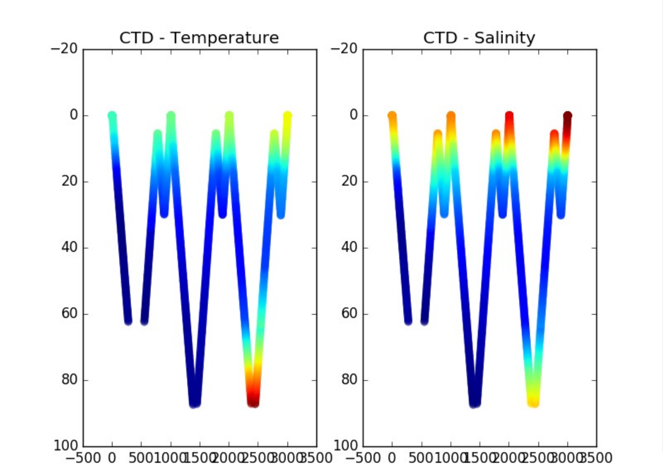
Ocean temperatures and salinity

Prior realizations (ocean models)



Questions:

- Environmental challenges
- Fish farming
- Algae bloom



Typical AUV data

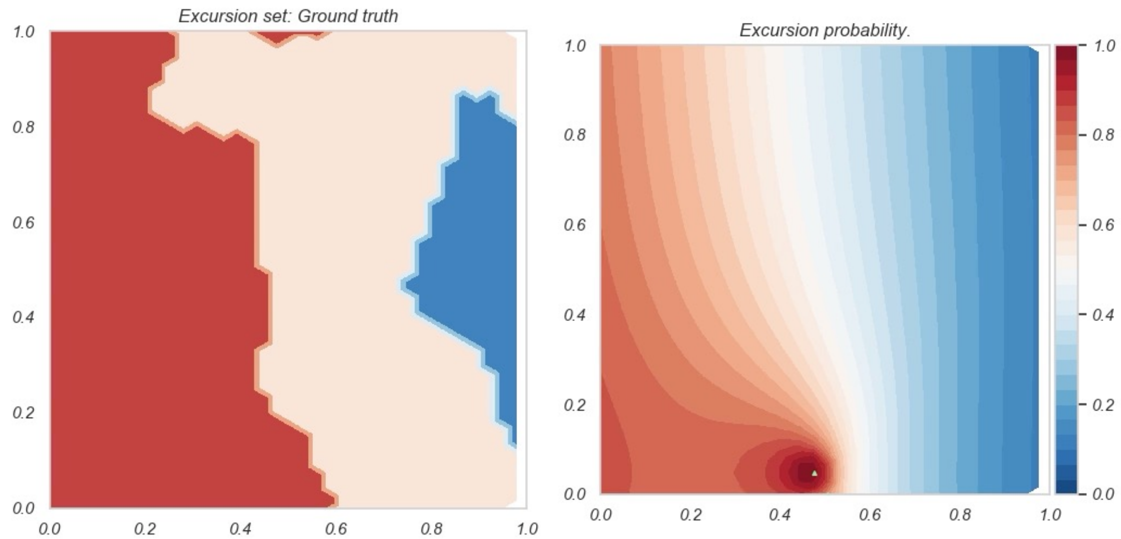
Adaptive sequential algorithm

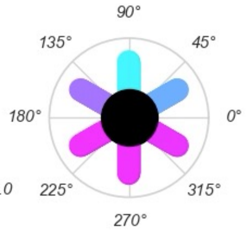
1. Find next best direction from analytic VOI (here connected to excursion sets classification), of all possible survey lines.
2. Collect data along currently best survey line.
3. Update statisical model in entire spatial domain given survey data.
4. Go to 1.

Myopic heuristic for dynamic program.

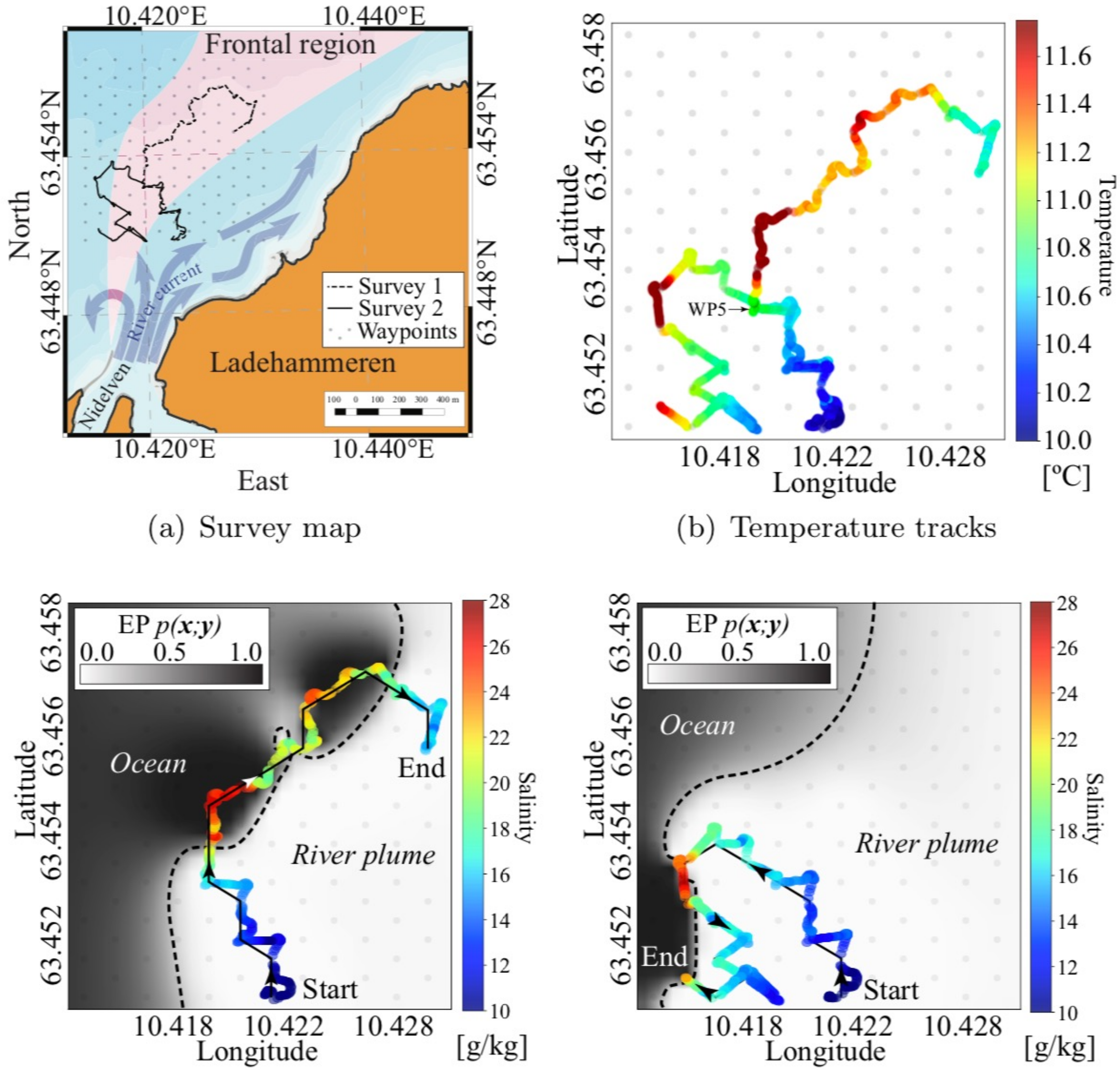
Excursion set algorithm

Excursion sets define regions where variables are above threshold.





Excursion set algorithm



(a) Survey map (b) Temperature tracks

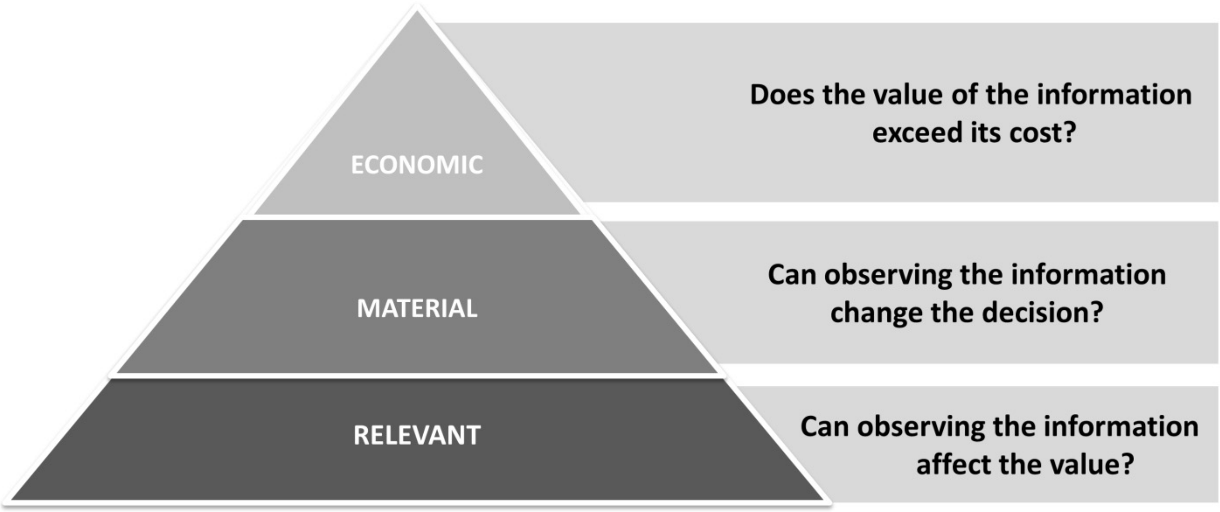
(c) Survey 1 (d) Survey 2

Take home : on sequential elements

- Information criterion is applied to a sequential search for good data designs.
- The design will depend on the data, and the results can be averaged over the data, to approximate value of different strategies.

- Sometimes connected with spatio-temporal applications.

Summary: VOI - Pyramid of conditions



Pyramid of conditions - VOI is different from other information criteria (entropy, variance, prediction error, etc.)

Summary: VOI and Earth sciences

- **Alternatives are spatial**, often with high flexibiliy in selection of sites, etc.
- **Uncertainties are spatial**, with multi-variable interactions .
- **Value function is spatial**, typically involving coupled features, say through differential equations. It can be defined by «physics» as well as economic attributes.
- **Data are spatial**. There are plenty opportunities for partial, total testing and a variety of tests (surveys, monitoring sensors, seismic, electromagnetic data, etc.)

Summary: Information gathering

	Perfect	Imperfect
Total	Exact observations are gathered for all locations. This is rare, occurring when there is extensive coverage and highly accurate data gathering. $$\boldsymbol{y} = \boldsymbol{x}$$	Noisy observations are gathered for all locations. This is common in situations with remote sensors with extensive coverage, e.g. seismic, radar, satellite data. $$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\varepsilon}$$
Partial	Exact observations are gathered at some locations. This might occur, for instance, when there is careful analysis of rock samples along boreholes in a reservoir or a mine. $$\boldsymbol{y}_{\mathrm{K}} = \boldsymbol{x}_{\mathrm{K}}, \quad \mathrm{K} \text{ subset}$$	Noisy observations are gathered at some locations. Examples include hand-held (noisy) meters to observe grades in mine boreholes, electromagnetic testing along a line, biological surveys of species, etc. $$\boldsymbol{y}_{\mathrm{K}} = \boldsymbol{x}_{\mathrm{K}} + \boldsymbol{\varepsilon}_{\mathrm{K}}, \quad \mathrm{K} \text{ subset}$$

Frame decision situations and values

.

	Assumption: Decision Flexibility	Assumption: Value Function
Low decision flexibility; Decoupled value	Alternatives are easily enumerated $$a \in A$$	Total value is a sum of value at every unit $$v(\mathbf{x},a)=\sum_j v(x_j,a)$$
High decision flexibility; Decoupled value	None $$\mathbf{a} \in \mathbf{A}$$	Total value is a sum of value at every unit $$v(\mathbf{x},\mathbf{a})=\sum_j v(x_j,a_j)$$
Low decision flexibility; Coupled value	Alternatives are easily enumerated $$a \in A$$	None $$v(\mathbf{x},a)$$
High decision flexibility; Coupled value	None $$\mathbf{a} \in \mathbf{A}$$	None $$v(\mathbf{x},\mathbf{a})$$

Summary: VOI workflow