# Value of Information

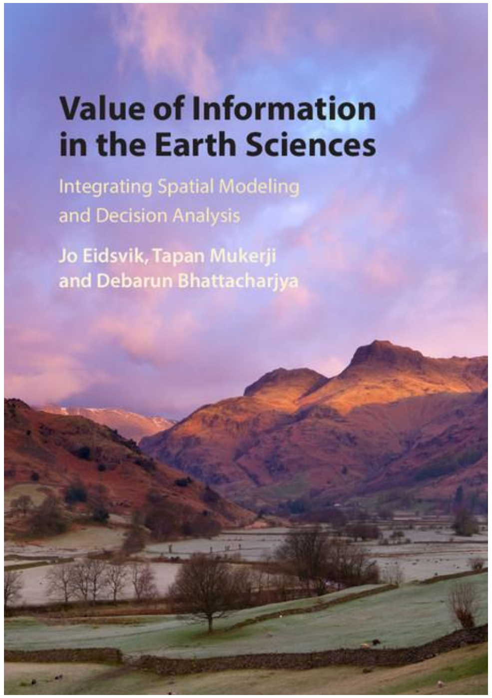
Jo Eidsvik,
Mathematical Sciences, NTNU

Probabilistic AI, Trondheim, June 2019

# Background:

- Computational and spatial statistics.
- Design of experiments / algorithms.
- Integration of statistical modeling and decision analysis.

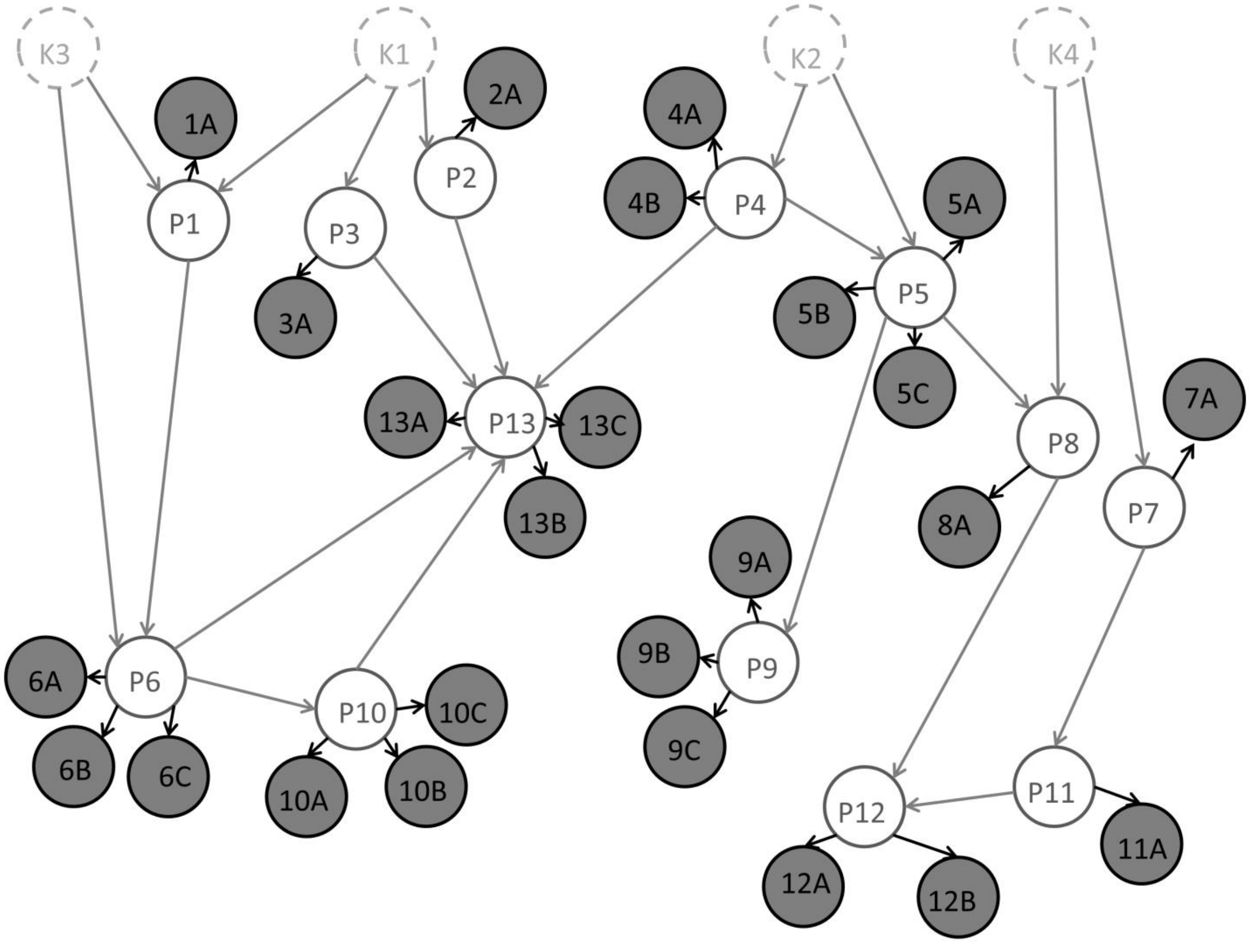
Collect data to resolve uncertainties
and to make informed decisions.




Eidsvik et al., 2015, Cambridge Univ press.

# Motivation
# (a petroleum exploration example)

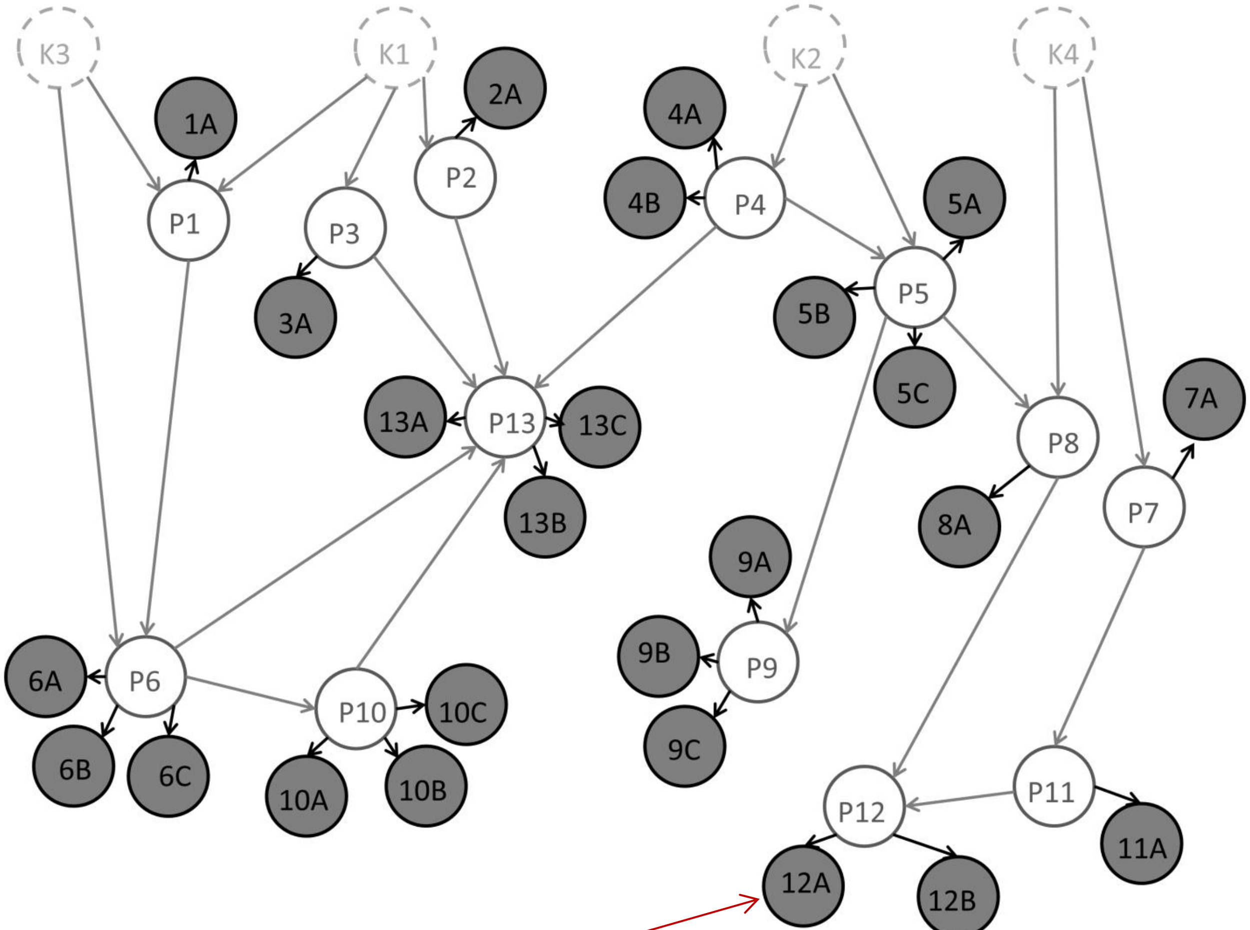
Gray nodes are petroleum reservoir segments where the company aims to develop profitable amounts of oil and gas.




Martinelli et al., 2013, Eur J of Oper Res.

# Motivation
# (a petroleum exploration example)

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

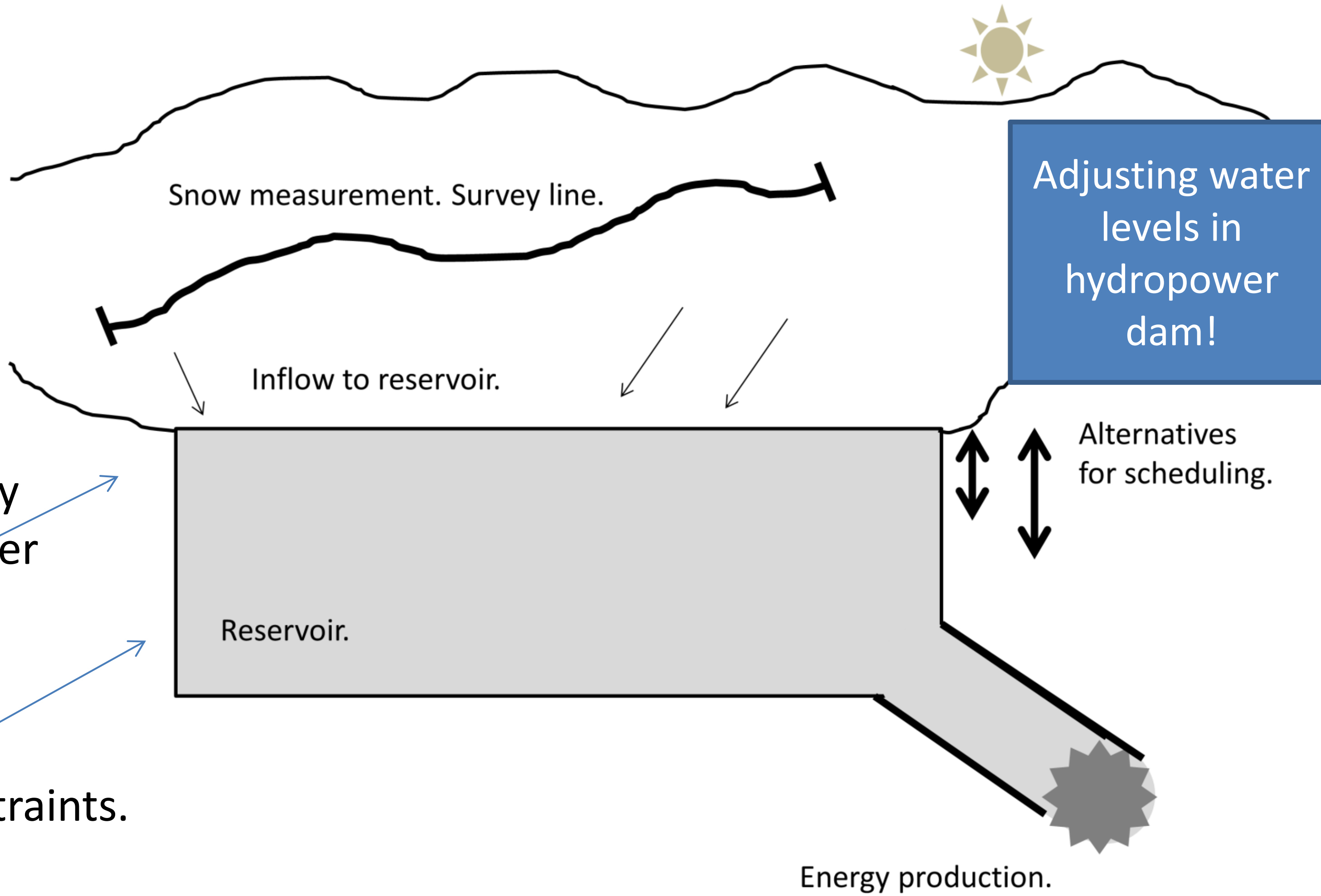



Drill the exploration well at this segment! The value of information is largest.

# Motivation
# (a hydropower example)

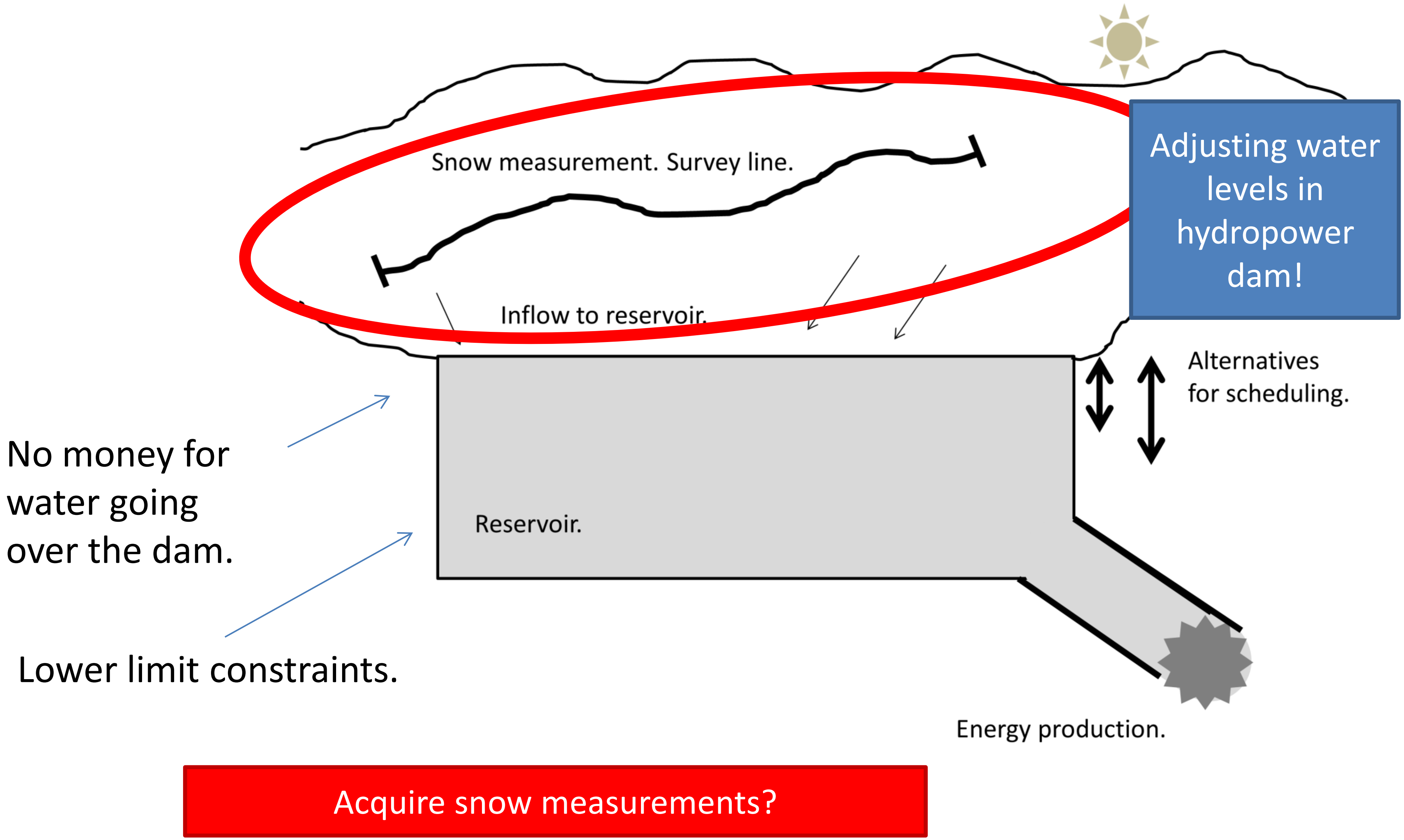
No money / energy for water going over the dam.

Lower limit constraints.




Ødegård et al., 2019, *Energy Systems*.

# Motivation
# (a hydropower example)



# Other applications

- Environmental – how monitor where pollutants are, to minimize risk or damage?

- Robotics - where should drone (UAV) or submarine (AUV) go to collect valuable data?
- Industry reliability – how to allocate sensors to 'best' monitor state of system?
- Internet of things – which sensors should be active now?



# Which data are valuable?

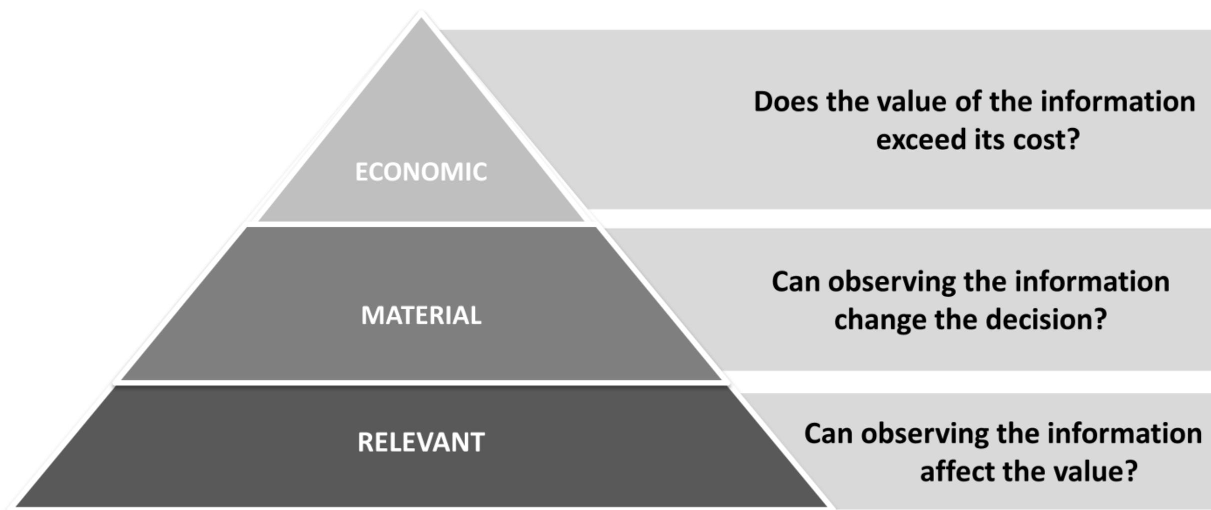
Five Vs of big data:

- Volume
- Variety
- Velocity
- Veracity
- **Value**

We must acquire and process data that has value!
Data should help us answer a key question.

# Value of information (VOI)

We often consider purchasing more data before making difficult decisions under uncertainty.
The value of information (VOI) is useful for quantifying the value of the data, before it is acquired and processed.




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

# Pirate example
(For motivating decision analysis and VOI)

# Pirate example

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).

?

Pirate makes decision based on preferences and maximum **utility** or **value**!

- Digging cost.
- Revenues if he finds the treasure .

# Pirate example

$a \in \{0,1\}$

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).

?

$x \in \{0,1\}$

Pirate makes decision based on preferences and maximum **utility** or **value**!

- Digging cost.
- Revenues if he finds the treasure .

$$\max_{a\in\{0,1\}}\left\{E\left(v\left(x,a\right)\right)\right\}$$

# Mathematics of decision situation:

- **Alternatives**

$$a \in \{0,1\} = A$$

- **Uncertainties (probability distribution)**

$$x \in \{0,1\} = \Omega \qquad p(x=1) = 0.01$$

- **Values**

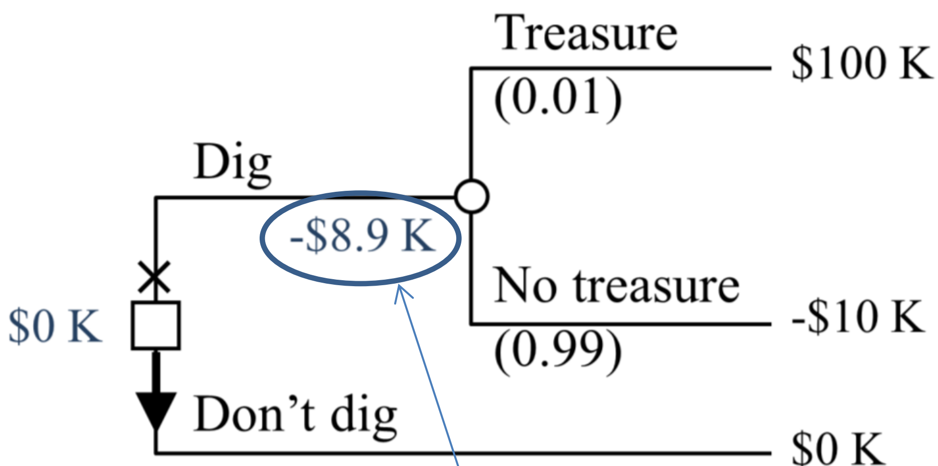
$$v = v(x,a)$$

$$v(x=0, a=1) = -10000 \qquad v(x=1, a=1) = 100000 \qquad v(x, a=0) = 0$$

- **Maximize expected value**

$$a^* = \arg\max_{a \in A} \{E(v(x,a))\}$$

# Pirate's decision situation




Risk neutral!

$$E\left(u\left(v_{dig}\right)\right) = E\left(v_{dig}\right) = 0.01(100000) + 0.99(-10000) = -8900$$

# Pirate example

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).
- Pirate can collect **data** before making the decision, if the experiment is worth its price!

- Perfect information. Clairvoyant!

- Imperfect information. Detector!

# Value of information (VOI)

- VOI analysis is the additional value of making informed decisions.
- If the VOI exceeds the price, the decision maker should purchase the data.

VOI=Posterior value – Prior value

# VOI – Pirate considers clairvoyant

$$PV = 0 = \$0K$$

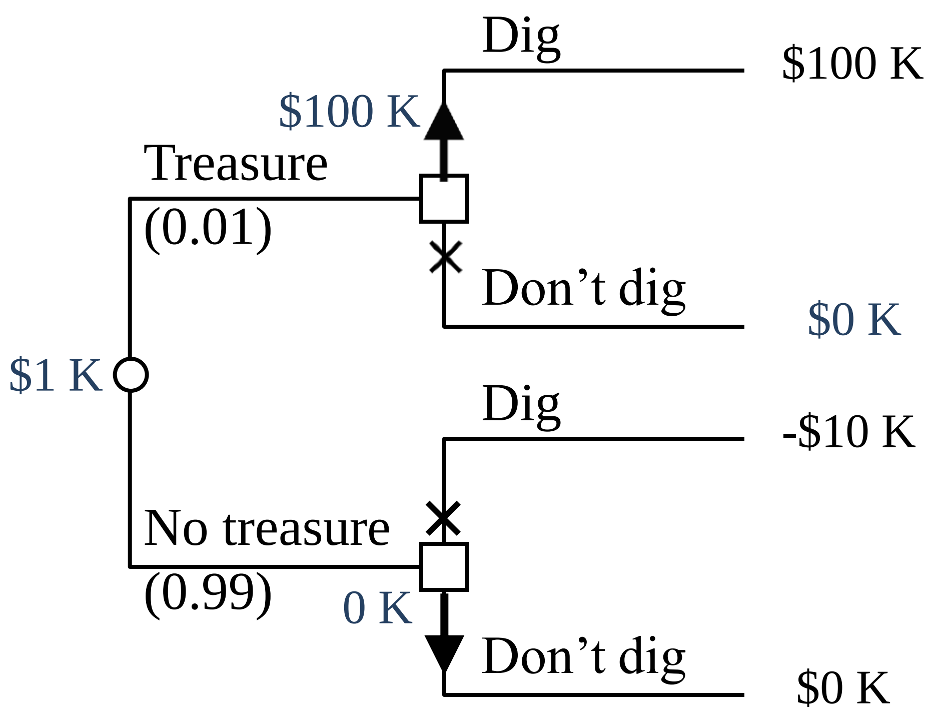
$$PoV(x) = \sum_x \max_{a \in A} \{v(x,a)\} \, p(x)$$

$$= \left(0.01 \cdot \max\{0,100\}\right) + \left(0.99 \cdot \max\{0,-10\}\right) = \$1K$$

$$VoI(x) = PoV(x) - PV = 1 - 0 = \$1K$$

Conclusion: Consult clairvoyant if (s)he charges less than $1000.

# PoV – decision tree, perfect information



# Pirate example - detector

$a \in \{0,1\}$

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).

- Pirate can collect **data** with a detector before making the decision, if this experiment is worth its price!

$x \in \{0,1\}$

$y \in \{0,1\}$

Pirate makes decision based on preferences and maximum expected **value**!

- Digging cost.
- Revenues if he finds the treasure .

$$\max_{a \in \{0,1\}} \{E(v(x,a) \mid y)\}$$

# Detector experiment

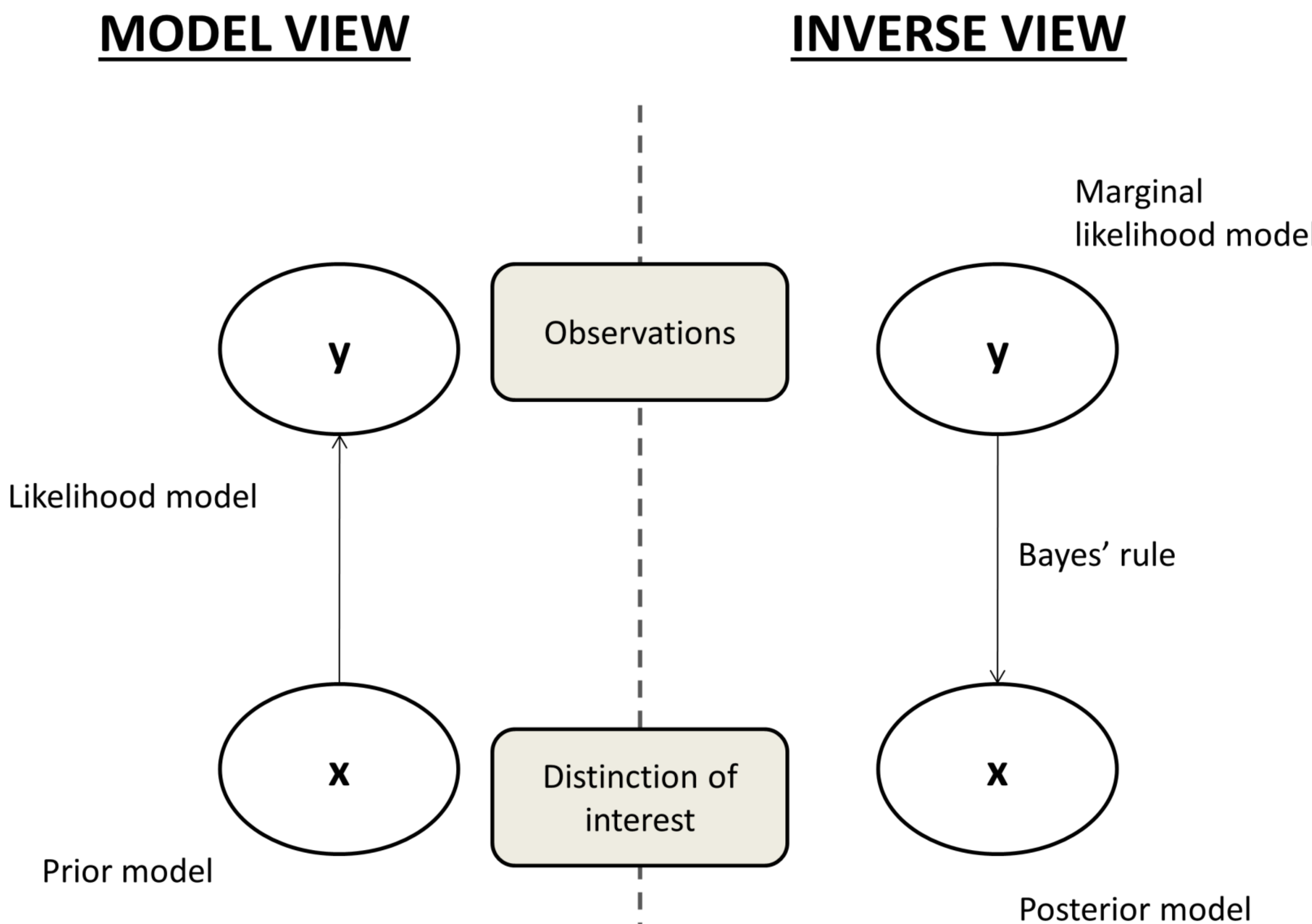
Accuracy of test:

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

Should the pirate pay to do a detector experiment?

Does the VOI of this experiment exceed the price of the test?

# Bayes rule - Detector experiment



# Bayes rule - Detector experiment

Likelihood:

$$p(y=0\,|\,x=0)=p(y=1\,|\,x=1)=0.95$$

Marginal likelihood:

$$p\left(y=1\right)=p\left(y=1\,|\,x=0\right)p\left(x=0\right)+p\left(y=1\,|\,x=1\right)p\left(x=1\right)$$
$$=0.05\cdot 0.99+0.95\cdot 0.01=0.06$$

Posterior:

$$p(x=1\,|\,y=1)=\frac{p(y=1\,|\,x=1)\,p(x=1)}{p(y=1)}=\frac{0.95\cdot 0.01}{0.06}\approx 0.16=16/100.$$

$$p(x=1\,|\,y=0)=\frac{p(y=0\,|\,x=1)\,p(x=1)}{p(y=0)}=\frac{0.05\cdot 0.01}{0.94}\approx 0.0005=5/10000.$$

# VOI – Pirate considers detector test

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

$$= \left(0.06 \cdot \max\left\{0, (100 \cdot 0.16) + (-10 \cdot 0.84)\right\}\right)$$

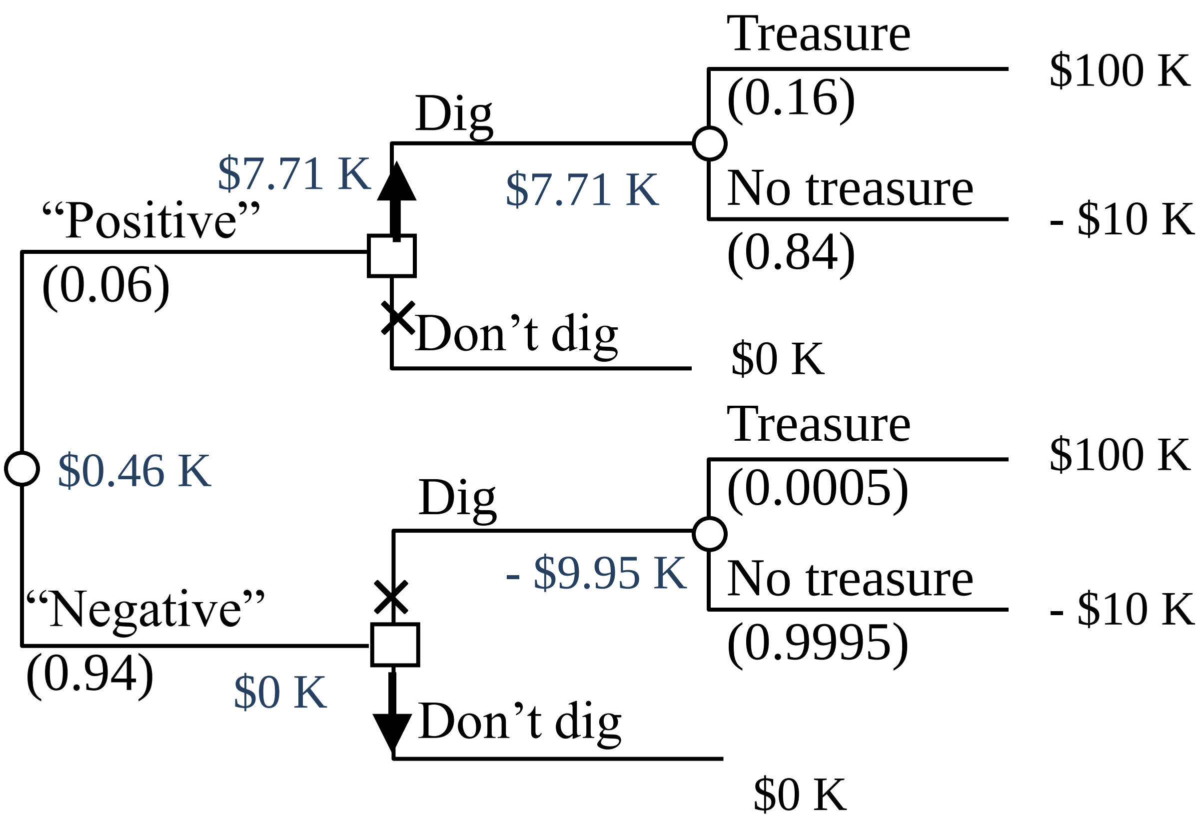
$$+ \left(0.94 \cdot \max\left\{0, (100 \cdot 0.0005) + (-10 \cdot 0.9995)\right\}\right)$$

$$= \left(0.06 \cdot \max\left\{0, 7.71\right\}\right) + \left(0.94 \cdot \max\left\{0, -9.95\right\}\right) = \$0.46K.$$

$$VoI(y) = PoV(y) - PV = 0.46 - 0 = \$0.46K$$

Conclusion: Purchase detector testing if its price is less than $460.

# PoV - imperfect information



# Value of information (VOI)

- VOI analysis is the additional value of making informed decisions.
- If the VOI exceeds the price, the decision maker should purchase the data.
- Rather than values, one can use utility functions to include risk profiles and solve for the certain equivalents to get the VOI.

# VOI - Clairvoyance

Price $P$ of experiment makes the equality.

$$\sum_{x} \max_{a\in A}\left\{v\left(x,a\right)-P\right\} p\left(x\right)=\max_{a\in A}\left\{E\left(v\left(x,a\right)\right)\right\}$$

$$\rightarrow P=VOI=\sum_{x} \max_{a\in A}\left\{v\left(x,a\right)\right\} p\left(x\right)-\max_{a\in A}\left\{E\left(v\left(x,a\right)\right)\right\}$$

VOI=Posterior value – Prior value

Assuming risk neutral decision maker!

# Properties of VOI

a) VOI is always positive

- Data allow better, informed decisions.

$$\max\left\{0, \sum_i v_i\right\} \leq \sum_i \max\{0, v_i\}$$

b) If value is in monetary units, VOI is in monetary units.

c) Data should be purchased if VOI > Price of experiment P.

d) VOI of clairvoyance is an upper bound for any imperfect information gathering scheme.

e) When we compare different experiments, we purchase the one with largest VOI compared with the price:

$$\arg\max\{VOI_1 - P_1, VOI_2 - P_2\}$$

f) Useful even without economic formulation – what is the most informative design / learning strategy : For improved classification, reducing entropy, etc.

# Requirements for VOI analysis

- Statistical model
- Decision situation
- Opportunities for data gathering

# Statistical modeling - Bayes

- All the currently available information about variables:

$$p(\boldsymbol{x})$$

- New data (and the data gathering scheme) is represented by a likelihood model:

$$p(\boldsymbol{y} \mid \boldsymbol{x})$$

- If we collect data, the model is updated to the posterior, conditional on the new observations:

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

# Information gathering

| | **Perfect** | **Imperfect** |
|---|---|---|
| **Total** | Exact observations are gathered for all locations. $$\boldsymbol{y} = \boldsymbol{x}$$ | Noisy observations are gathered for all locations. $$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\varepsilon}$$ |
| **Partial** | Exact observations are gathered at some locations. $$\boldsymbol{y}_{\mathbb{K}} = \boldsymbol{x}_{\mathbb{K}}, \qquad \mathbb{K} \text{ subset}$$ | Noisy observations are gathered at some locations $$\boldsymbol{y}_{\mathbb{K}} = \boldsymbol{x}_{\mathbb{K}} + \boldsymbol{\varepsilon}_{\mathbb{K}}, \qquad \mathbb{K} \text{ subset}$$ |

# Decision analysis and VOI analysis

Prior value:

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

Posterior value:

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

*VOI* = Expected posterior value – Prior value

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

$\boldsymbol{x}$ - Uncertainties

$\boldsymbol{a}$ - Alternatives

$v\left(\boldsymbol{x}, \boldsymbol{a}\right)$ - Value function

$\boldsymbol{y}$ - Data

# Computing the VOI

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

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

Inner integral.

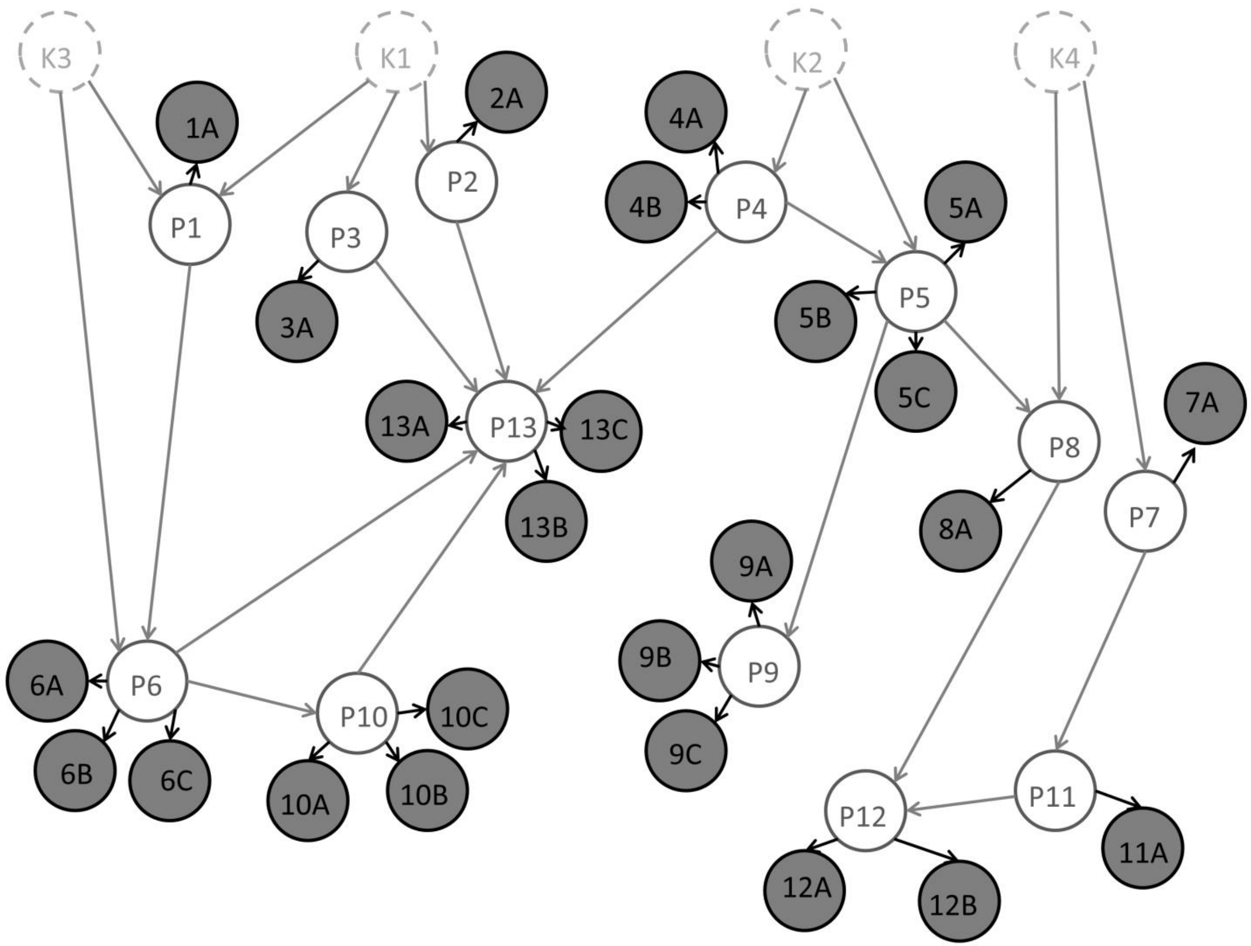
Outer integral.

**Computational techniques** :

- Fully analytically tractable for special cases, **Gaussian-linear models, some discrete models.**
- Various approximations and Monte Carlo approaches usually applicable.

# Petroleum example of VOI analysis

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



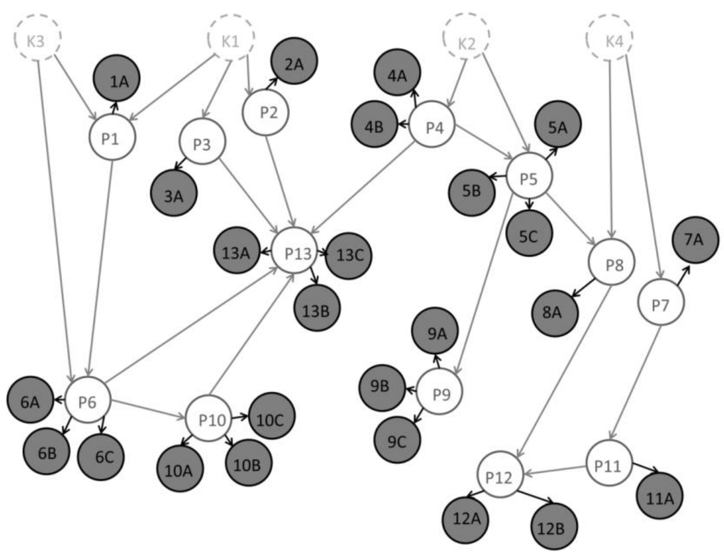

Drill the exploration well at this segment!
The value of information is largest.

# Networks / graphs - computation

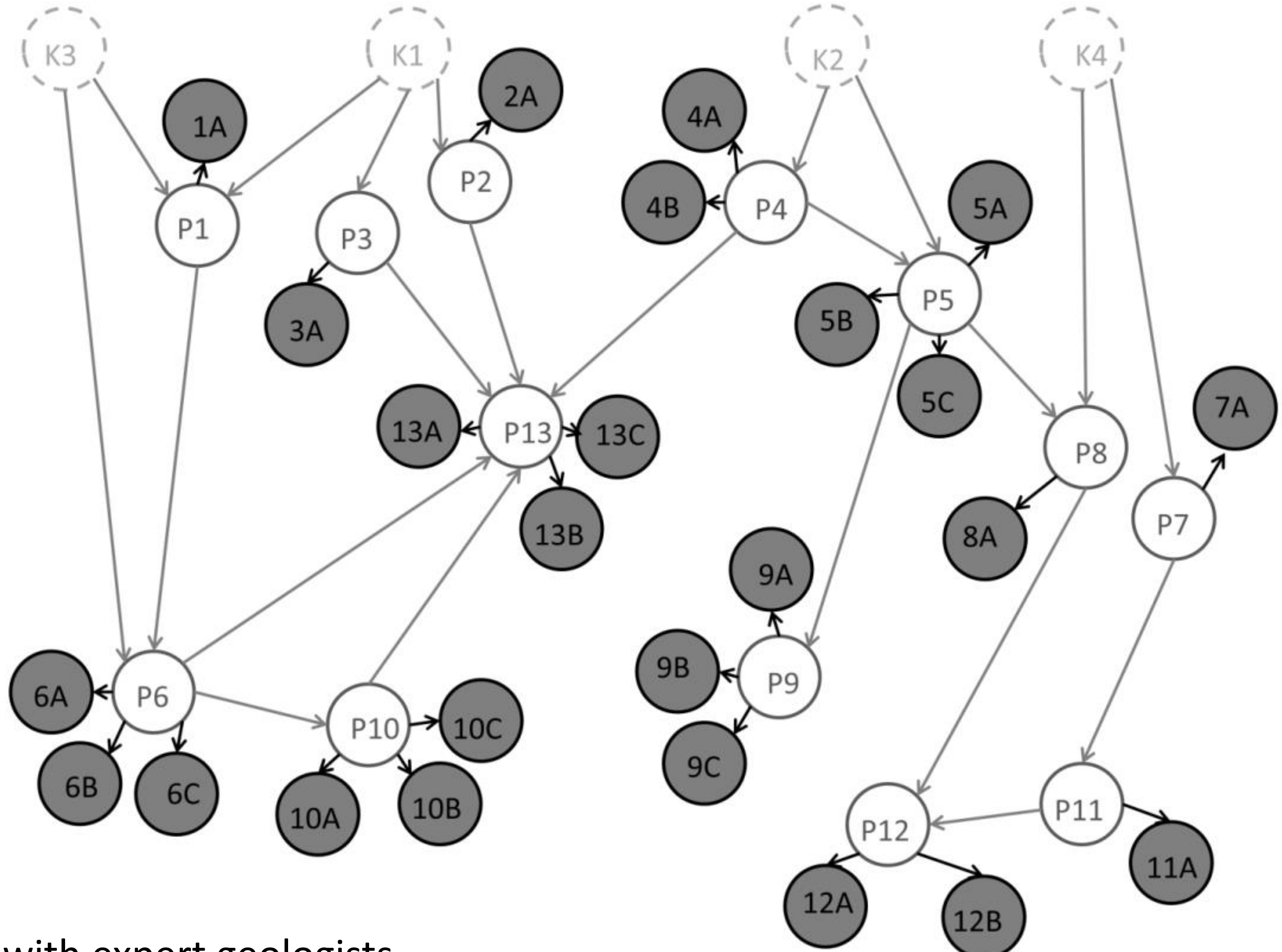
Algorithms have been developed for efficient marginalization and conditioning.

This enables fast evidence propagation and VOI calculations.

- Bayesian network models (Junction tree algorithm).



# Bayesian network , reservoir segments

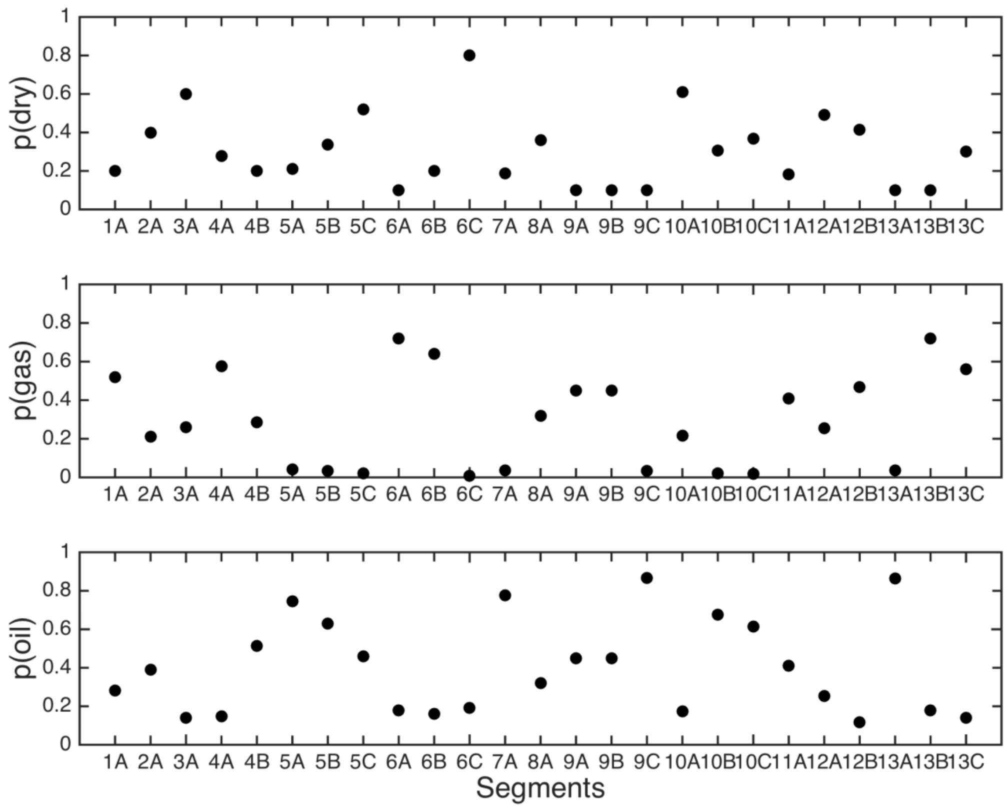



Model elicited with expert geologists.
Source migration of gas (and oil) from kitchens (K).
i) Local failure probability of migration and ii) $p\left(x_i = \text{dry} \mid \boldsymbol{x}_{Pa(i)} = \mathbf{dry}\right) = 1$

# Prior marginal probabilities

Three possible classes at all nodes:

- Dry
- Gas
- Oil



# Prior values

Development fixed cost.
Infrastructure at prospect r.

$$PV = \sum_{r=1}^{13} \max\left\{0, \sum_{i \in \Pr} IV(x_i) - DFC\right\}$$

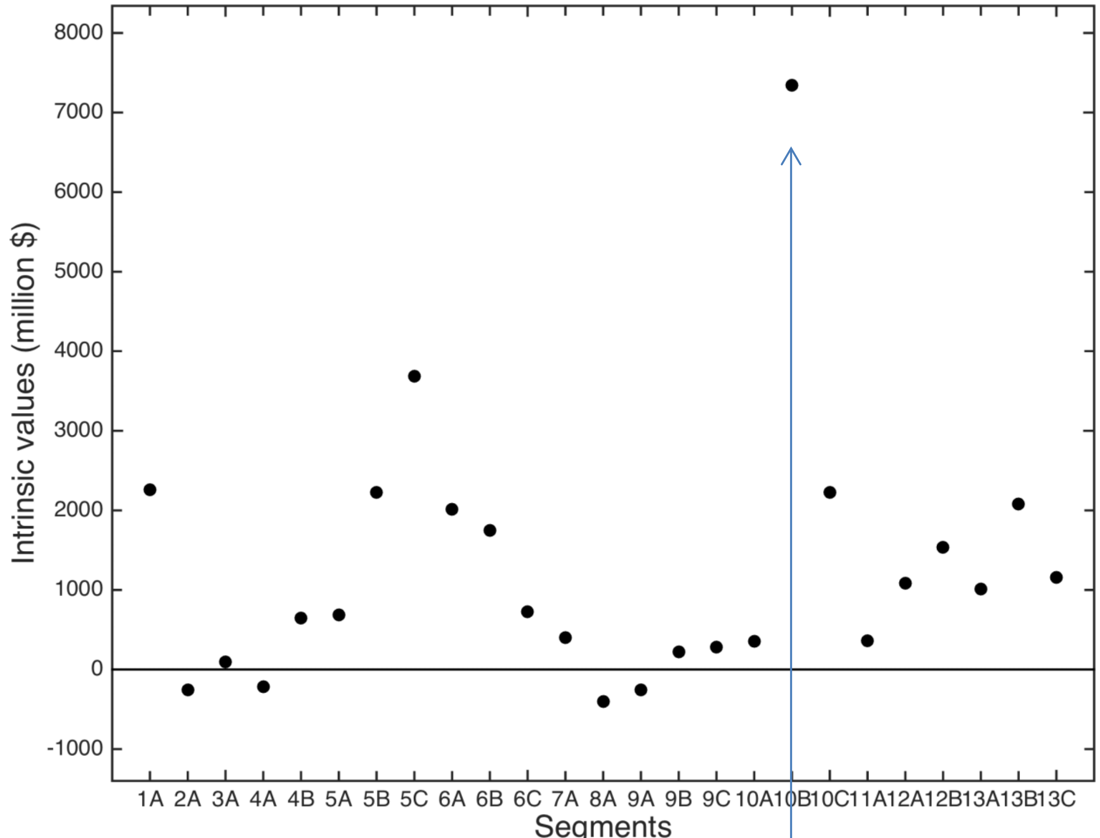
$$IV(x_i) = \sum_{k=1}^{3}\left(\mathrm{Rev}_{i,k}\, p(x_i = k) - \mathrm{Cost}_{i,k}\, p(x_i = k)\right) - \mathrm{Cost}_{i,0}$$

Revenues of oil/gas,
0 otherwise.

Cost if dry,
0 otherwise.

Cost of drilling
segment i.

# Values




Most lucrative. But might not be most informative.

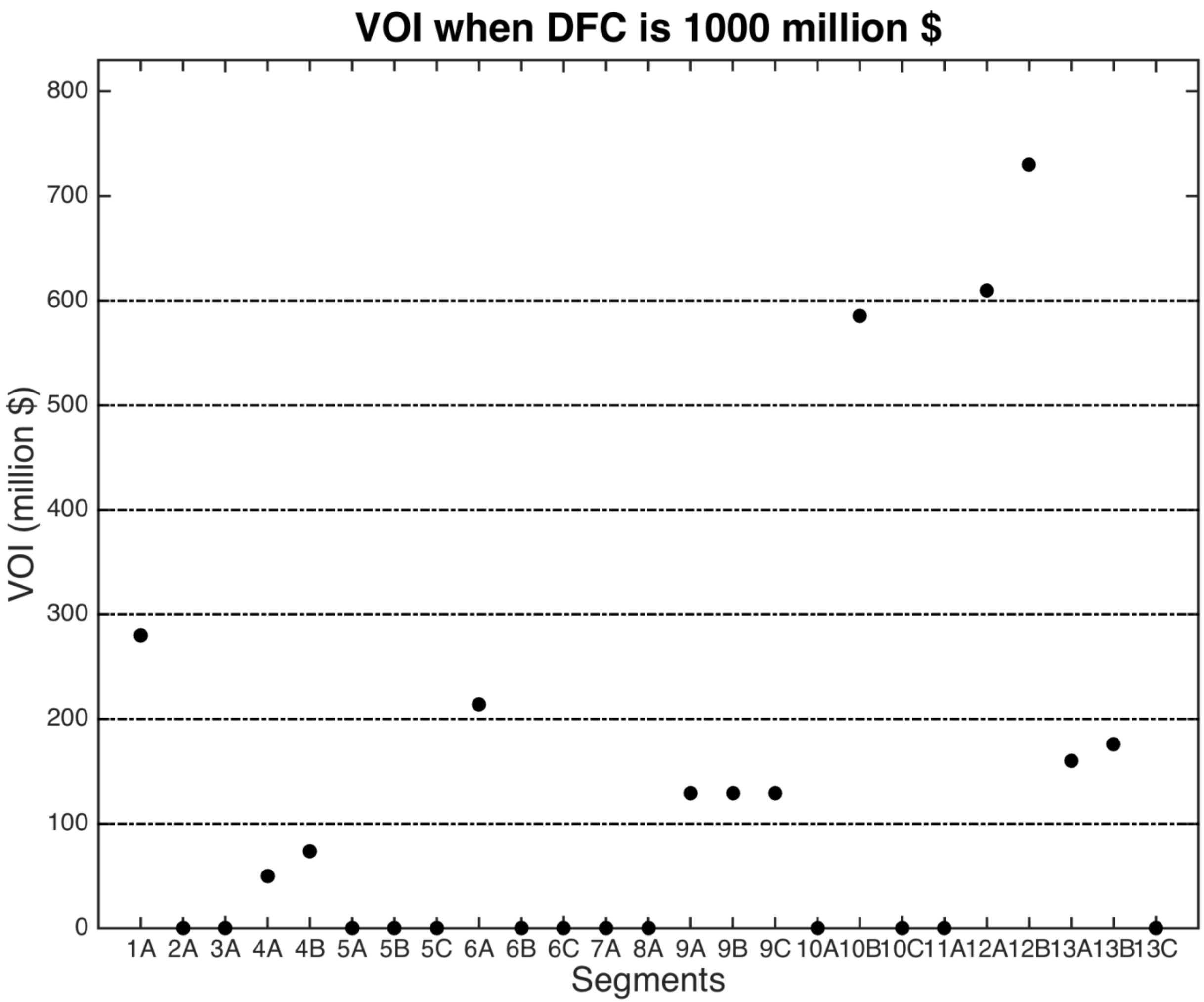
# Posterior values and VOI

$$PoV\left(x_{\mathbb{K}}\right)=\sum_{l=1}^{3}\sum_{r=1}^{13}\max\left\{0,\sum_{i\in \Pr} IV\left(x_i \mid x_{\mathbb{K}}=l\right)-DFC\right\}p\left(x_{\mathbb{K}}=l\right)$$

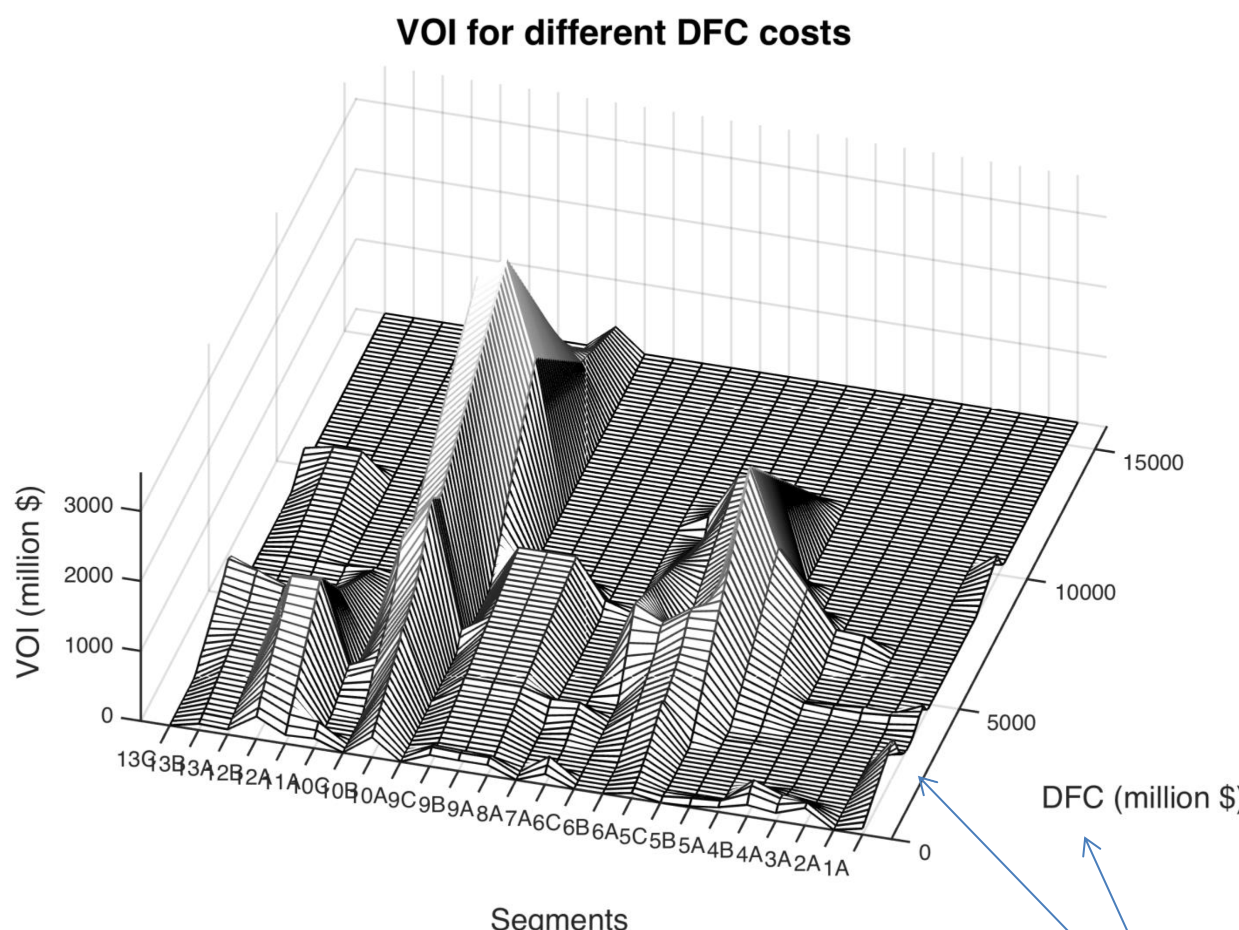
$$VOI\left(x_{\mathbb{K}}\right)=PoV\left(x_{\mathbb{K}}\right)-PV$$

Data acquired at single well.

# VOI single wells



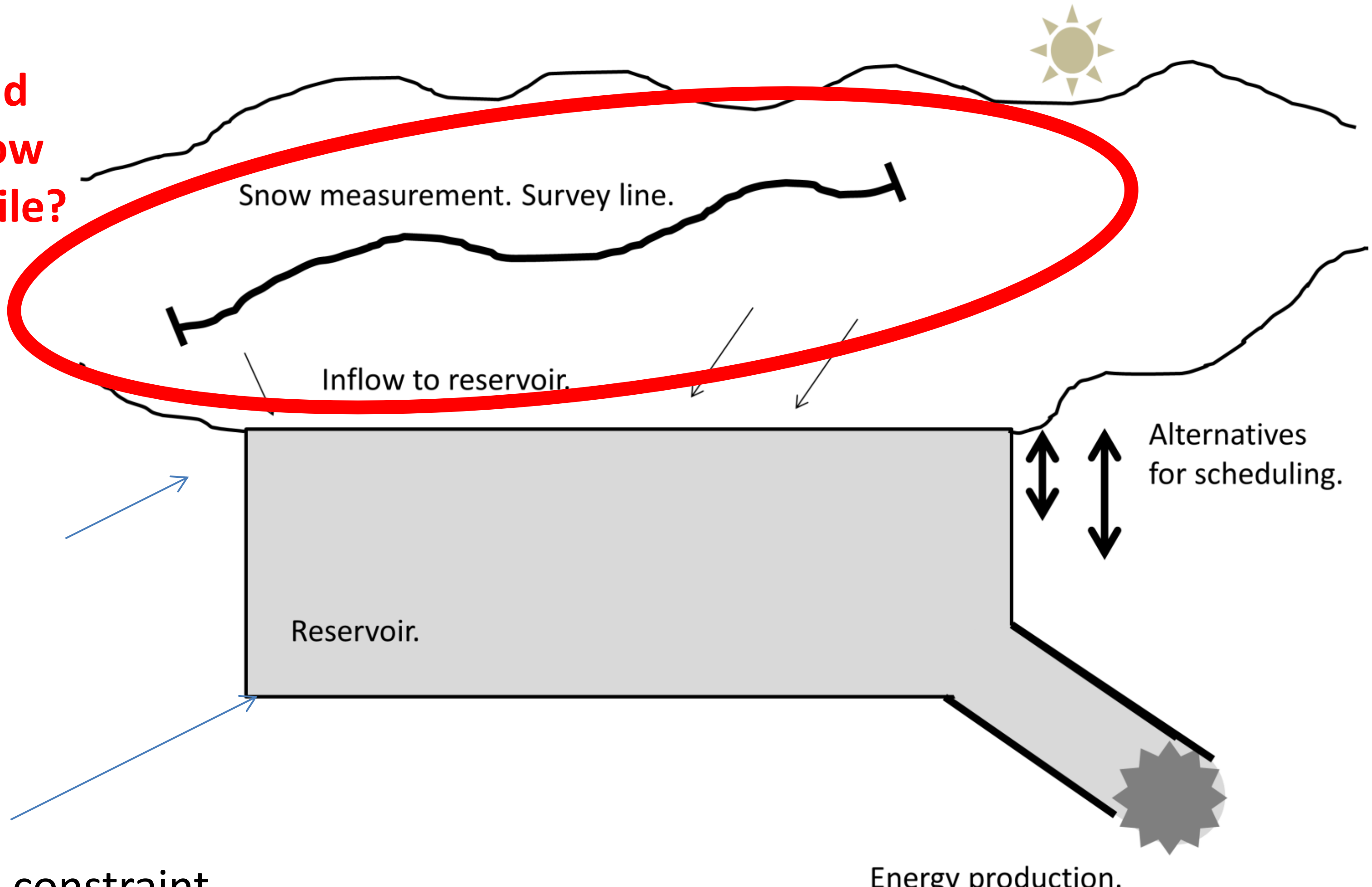
# VOI for different costs




Development fixed cost.

# Hydropower example of VOI analysis

**Is acquiring and processing snow data worthwhile?**




No money for water going over the dam.

Lower limit constraint.

# VOI of snow measurements

$\boldsymbol{x}$ Inflow.

$\boldsymbol{a}$ Scheduling controls.

$\boldsymbol{y}$ Snow measurements.

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

$$PV = \operatorname{seq}\max_{\boldsymbol{a}\in\boldsymbol{A}}\left\{E\left(v\left(\boldsymbol{x},\boldsymbol{a}\right)\right)\right\}$$

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

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

# Approximate VOI computation

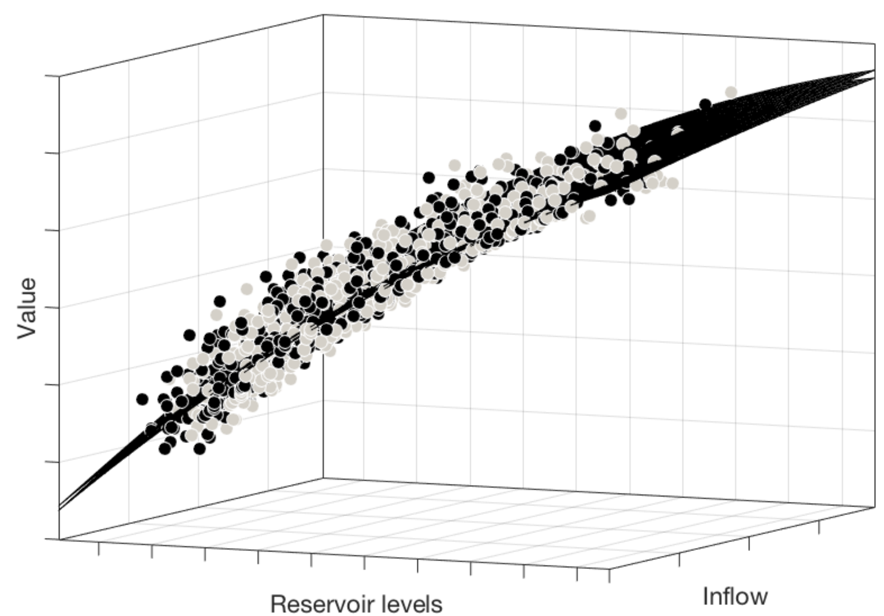
$$PoV\left(\boldsymbol{y}\right)=\overbrace{\sum_{\boldsymbol{y}}\operatorname{seq}\max_{\boldsymbol{a}\in\boldsymbol{A}}\left\{\underbrace{E\left(v\left(\boldsymbol{x},\boldsymbol{a}\right)\mid\boldsymbol{y}\right)}_{\text{Inner expectation: }\ \boldsymbol{x}\mid\boldsymbol{y}}\right\}p\left(\boldsymbol{y}\right)}^{\text{Outer expectation: }\ \boldsymbol{y}}$$

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

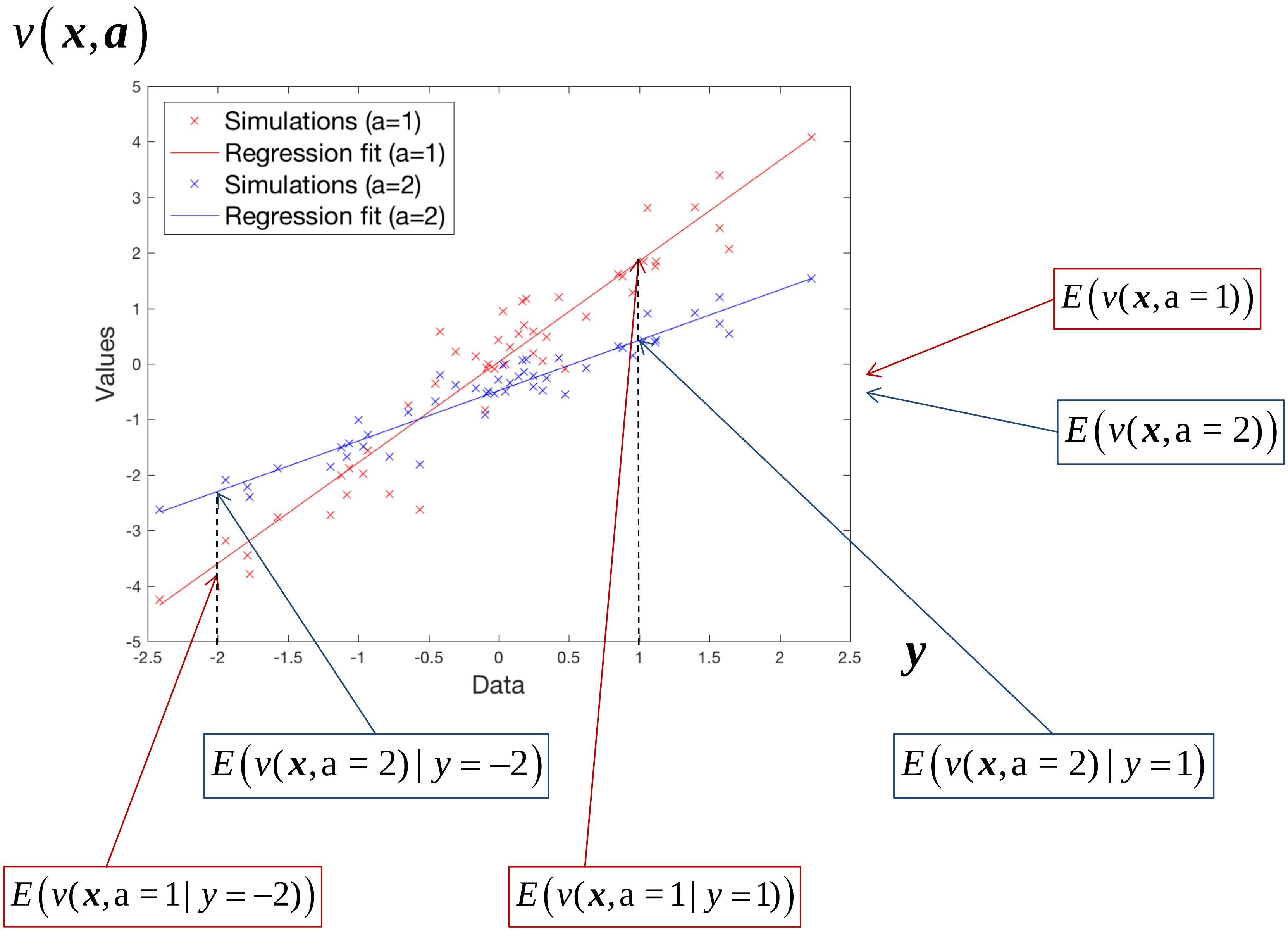
- **Monte Carlo (outer) and regression approximation (inner).**

# Method – Least squares Monte Carlo

- Simulate inflow **scenarios** (10 000 models from data and time series fitting)
- Wind-up **optimal solution for controls over time**.
- Optimal scheduling solution not available. Approximated by **least-squares fitting** of surfaces from simulated values as a function of inflow and reservoir level.



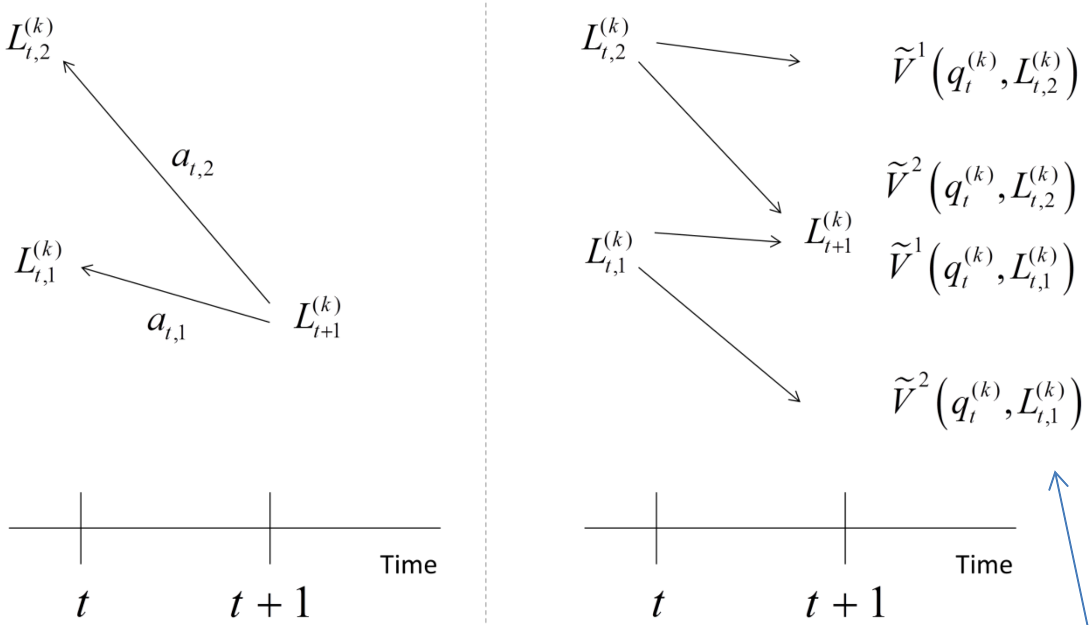
# Illustration - fit regression model to samples



# Method – simulation regression

Linking forward runs and backward algorithm using Markov assumption.

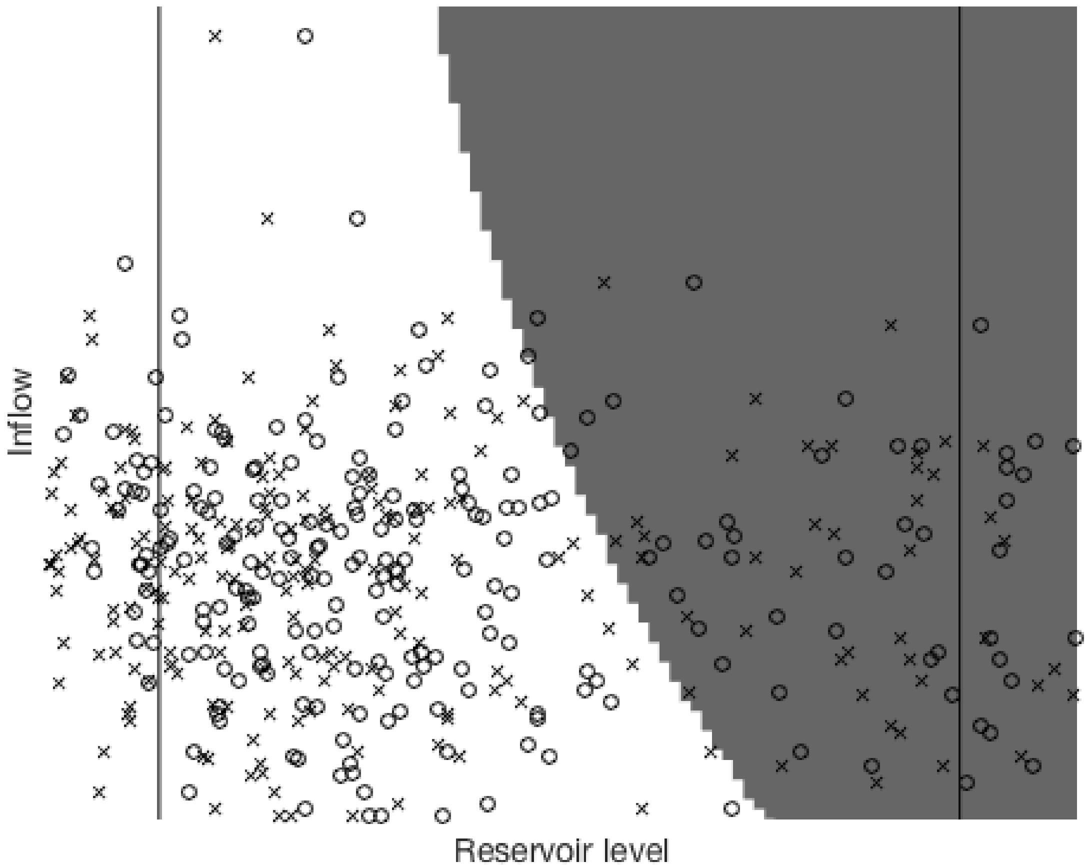
**Backward step** **Forward step**




Reservoir levels:

Value is function of inflow and reservoir level.

# Method – regression surfaces

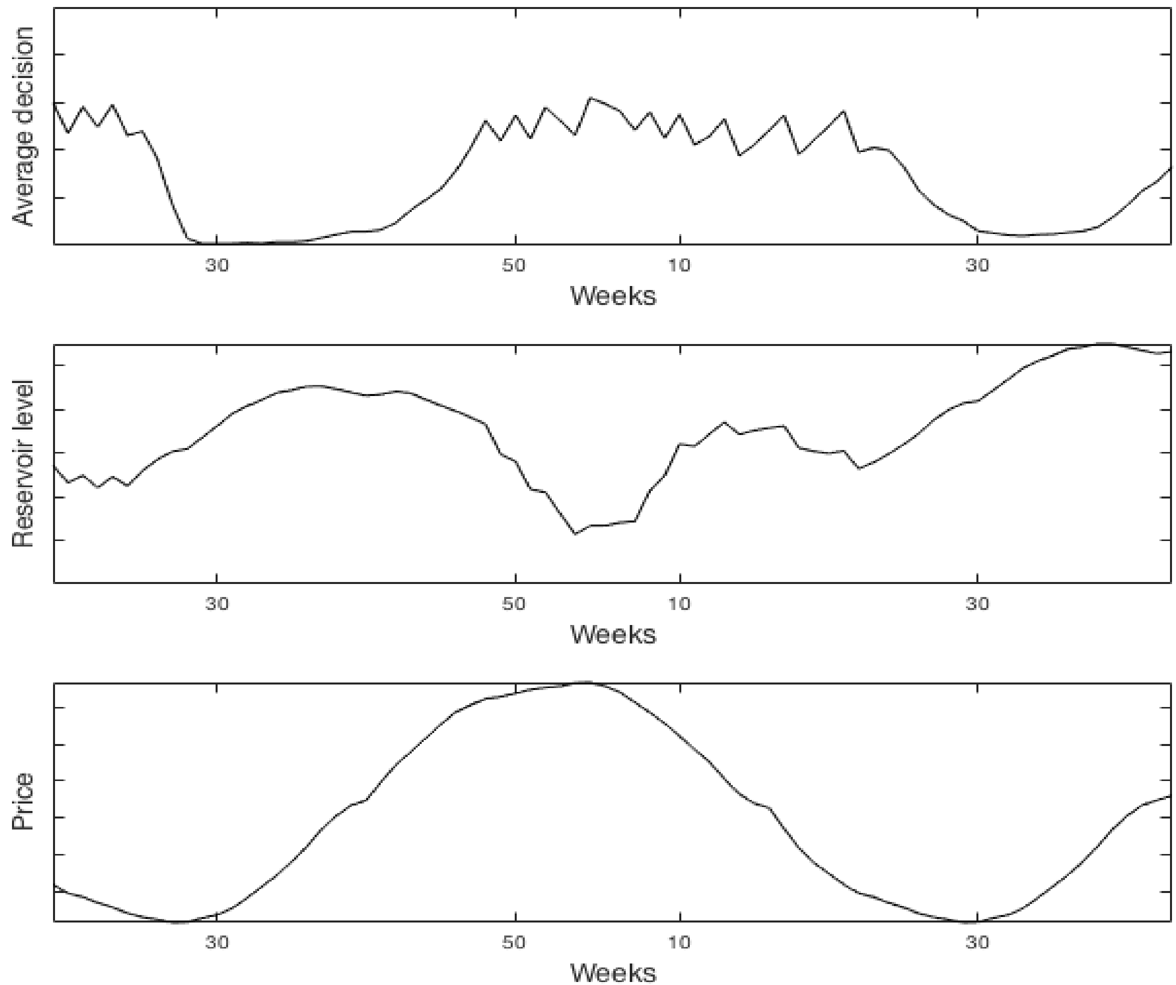



Highest surface sets control.

Quadratic components in regression fitting.

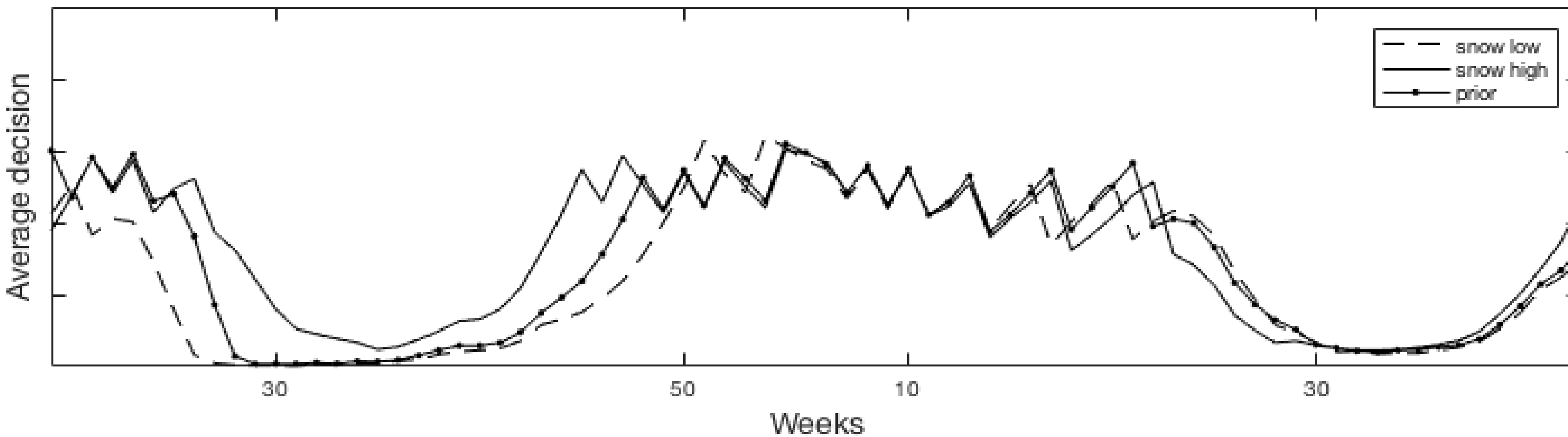
# Results – scheduling (prior)

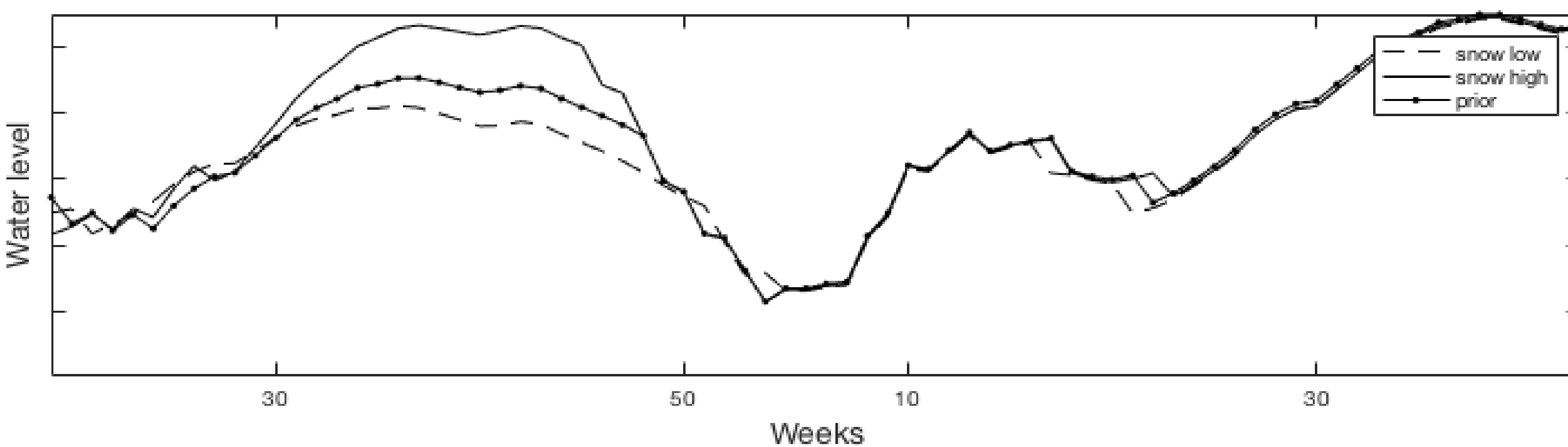
Base case has two controls.



# Results – scheduling with snow data

Data split scenarios in sub-groups.



# Results – 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

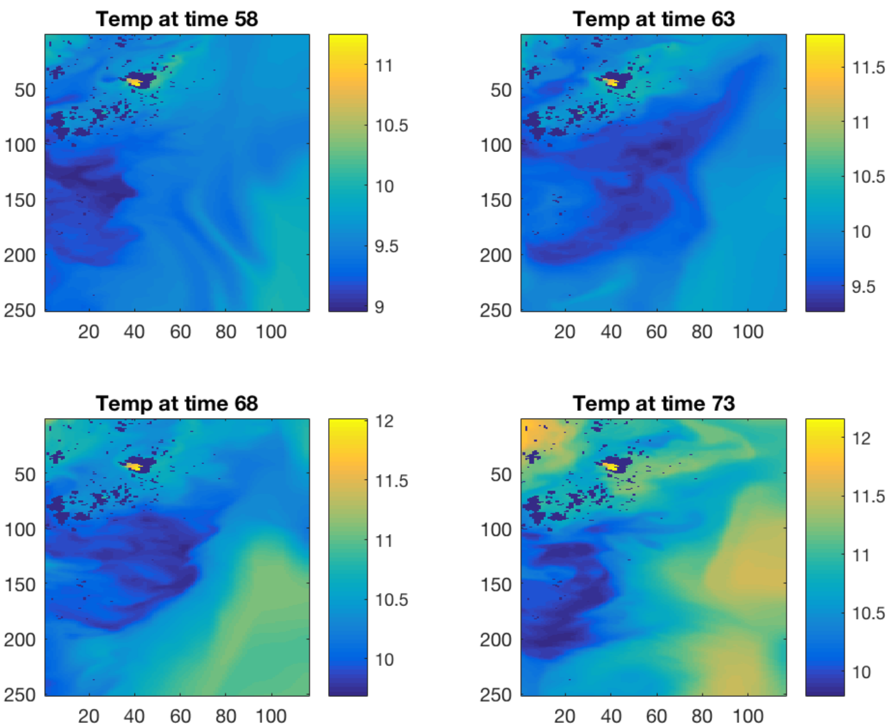
# Oceanographic example of VOI

- Goal (value) is to detect large spatial gradients in ocean temperature.
- Autonomous underwater vehicle (AUV) information. Where? And in what sequence?
- Model for temperature is represented by Gaussian spatial process.
- VOI analysis uses analytical approach and myopic heuristics.

Fossum et al., 2018, Journal of field robotics.

# Mapping ocean temperature variability

Satellite data and ocean models realizations are used to build Gaussian prior mean and covariance.

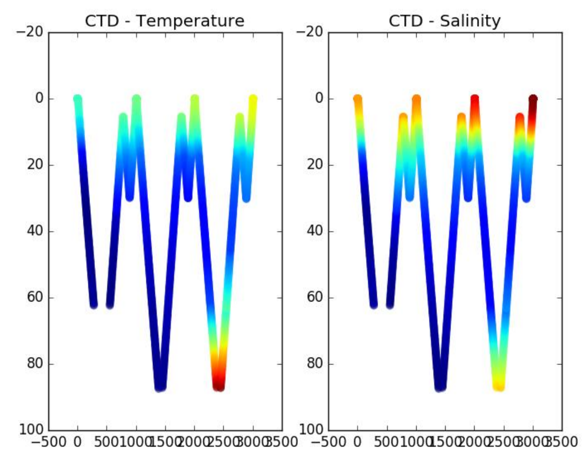



Area outside Trondheim fjord.

Possible questions:

- Environmental challenges
- Fish farming
- Algae bloom

Typical AUV data



# Gaussian prior and likelihood

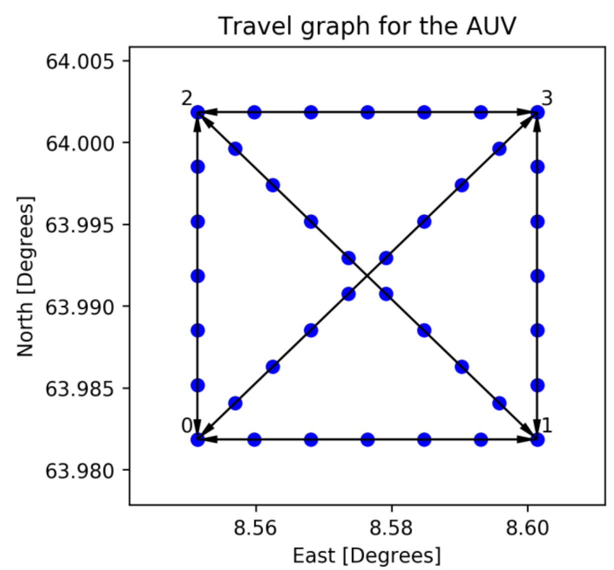
$$p(\boldsymbol{x}) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Gaussian spatial process prior for temperatures (learned from current knowledge).

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

$$p(\boldsymbol{y} \mid \boldsymbol{x}) = N(\boldsymbol{F}\boldsymbol{x}, \tau^2 \boldsymbol{I})$$

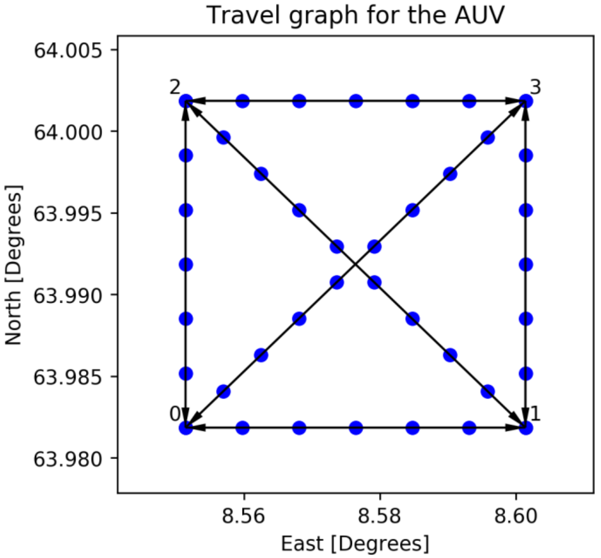
Likelihood, design matrix, picks data locations, for every time step.



# Goal of surveying

The main task for the AUV is to detect large gradients in temperature which are linked to algea bloom.

Waypoints in survey design for AUV.

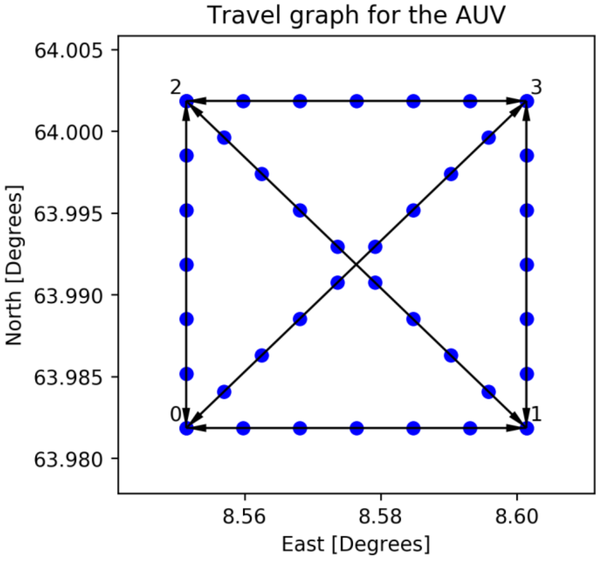


# Adaptive sequential algorithm

1. Find next best survey line (if any) from analytic VOI, of all possible survey lines.
2. Collect temperature data along currently best survey line.
3. Update temperature model in entire spatial domain given survey data.
4. Go to 1.

Myopic heuristic for dynamic program.

Closed form expression for the Gaussian process model.



# Myopic scheme

1. Find best single NS line, if any.

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

$$r_{w,j} = \sqrt{\sum \sum R_{ii',j}}, \qquad \mu_w = \sum \mu_i$$

$$PoV\left(\boldsymbol{y}_j\right) = \left( \mu_w \Phi\left( \mu_w / r_{w,j} \right) + r_{w,j} \phi\left( \mu_w / r_{w,j} \right) \right) - P_j$$

2. Collect data for this line. $\boldsymbol{y}_j$

3. Update the model

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

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma} - \boldsymbol{R}_j,$$

4. Stop testing or continue testing.

$$Stop = \max\{0, \mu_w\}, \qquad \mu_w = \sum_{i=1}^{n} \mu_i$$

Find largest among all k.

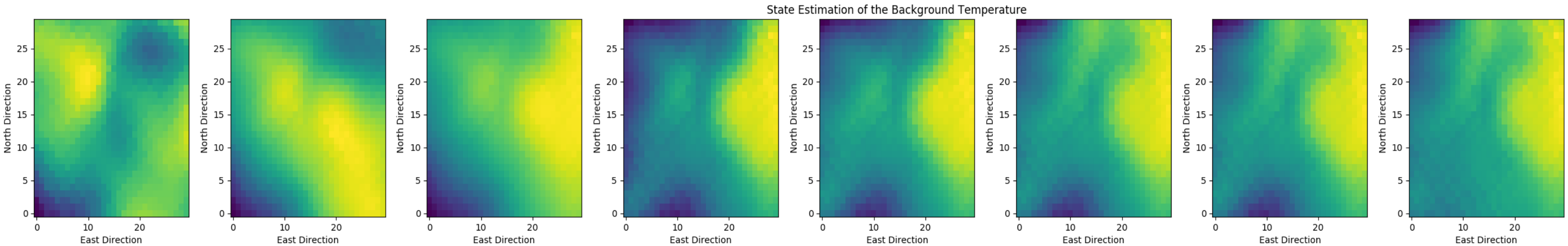
$$Cont\left(\boldsymbol{y}_k\right) = \left( \mu_w \Phi\left( \mu_w / r_{w,k} \right) + r_{w,k} \phi\left( \mu_w / r_{w,k} \right) \right) - P_k$$

Use updated mean and covariances.

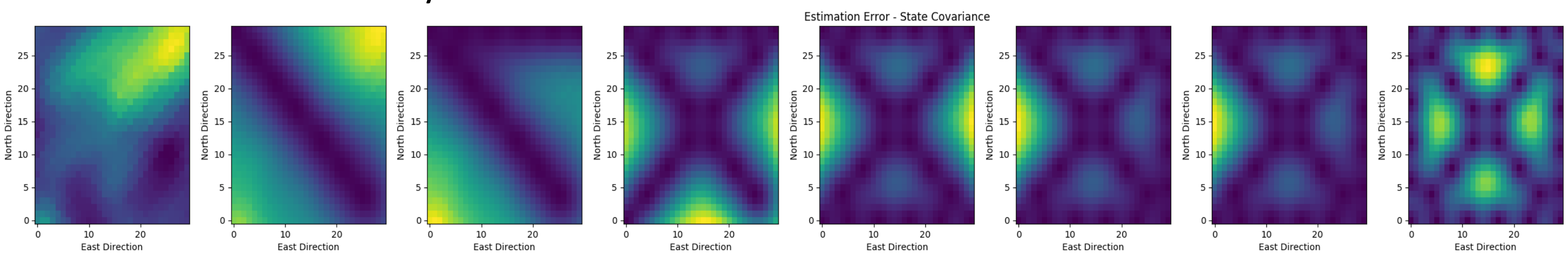
Etc....

# Results of adaptive algorithm

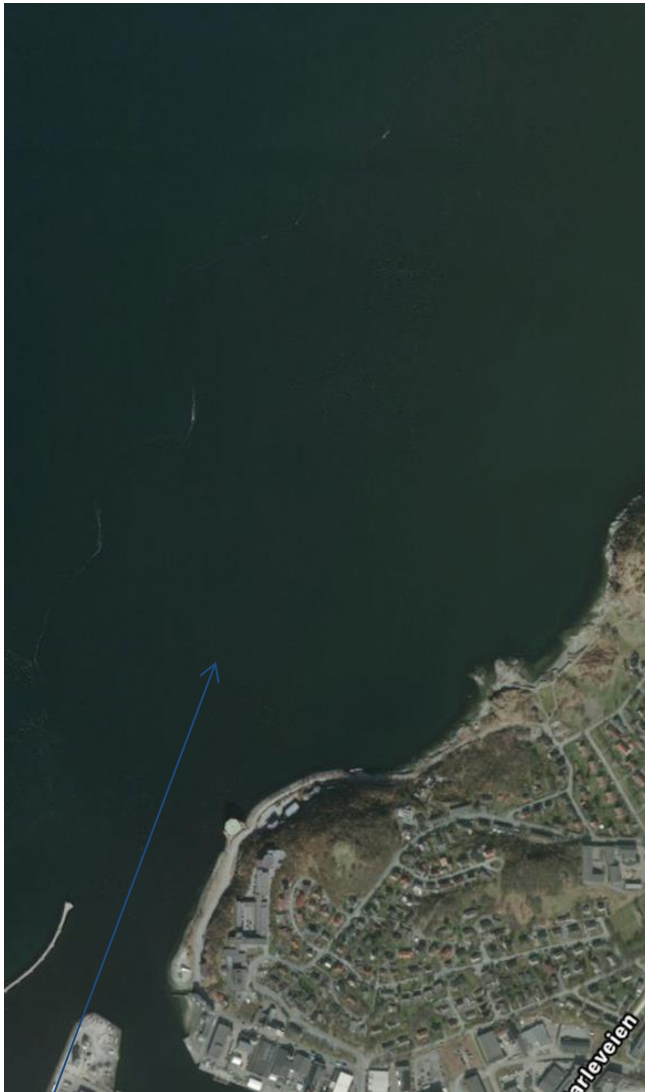
## Mean of one survey




## Variance of one survey



# Fresh cold water & salt warm water




River mouth

River mouth

# Excursion sets and excursion probabilities

$$ES_a = \{\mathbf{s} : x(\mathbf{s}) < a\}$$

Connections to active learning.

$$\boldsymbol{x} \sim GP(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$EP_a(\mathbf{s}) = P(x(\mathbf{s}) < a)$$

$$\boldsymbol{y}_d = \boldsymbol{F}_d \boldsymbol{x} + N(\boldsymbol{0}, \boldsymbol{T}_d)$$

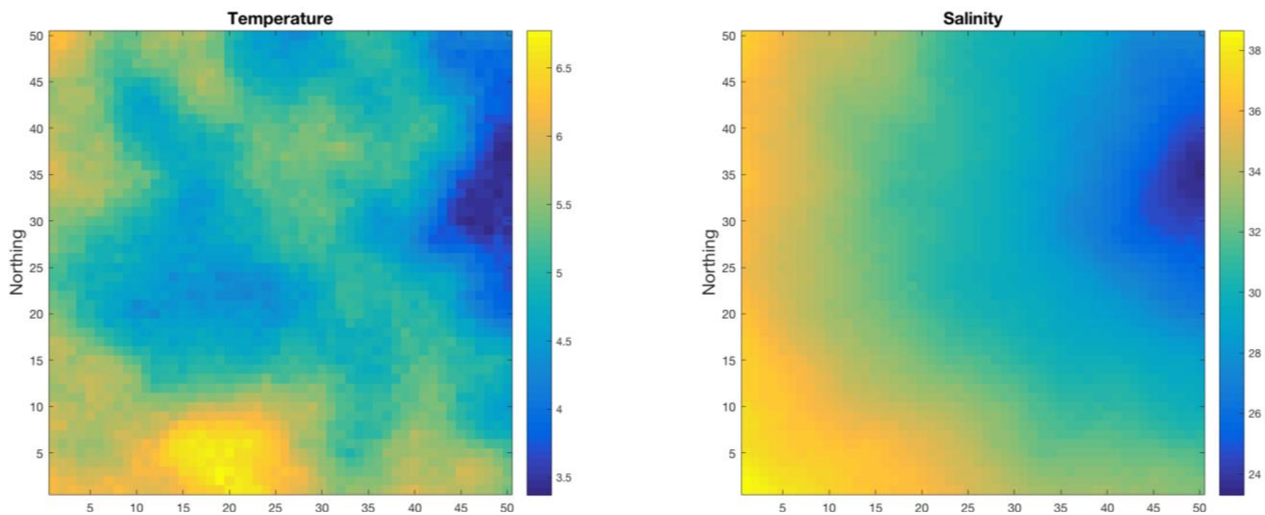
$$EP_a(\mathbf{s} \mid \boldsymbol{y}_d) = P(x(\mathbf{s}) < a \mid \boldsymbol{y}_d)$$

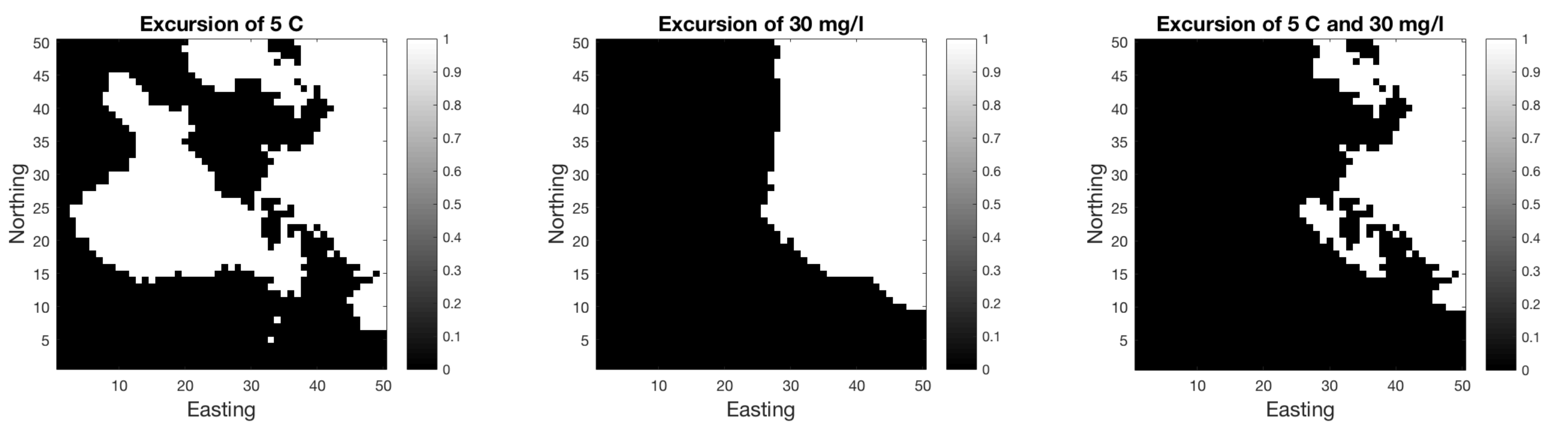
Criterion for path selection:

$$d^* = \arg\min_d \left\{ \int\int \underbrace{EP_a(\mathbf{s} \mid \boldsymbol{y}_d)(1 - EP_a(\mathbf{s} \mid}_{} \boldsymbol{y}_d)) \, p(\boldsymbol{y}_d) \, d\boldsymbol{y}_d \, d\mathbf{s} \right\}$$

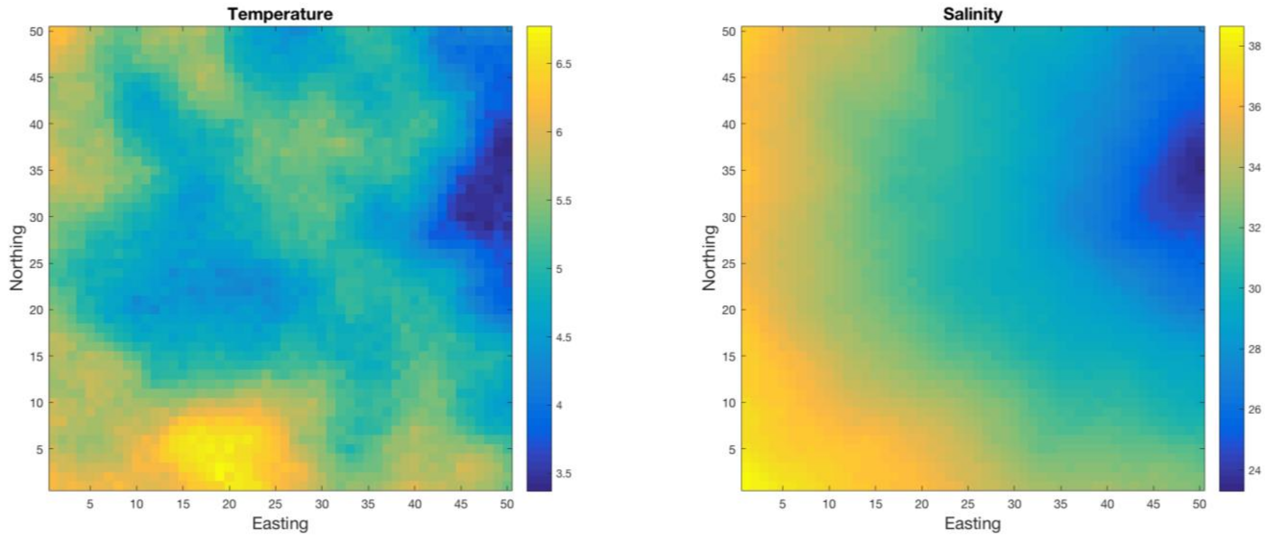
Closed form for Gaussian processes.

# (Bivariate) excursion sets



# Bivariate excursion sets – closed form



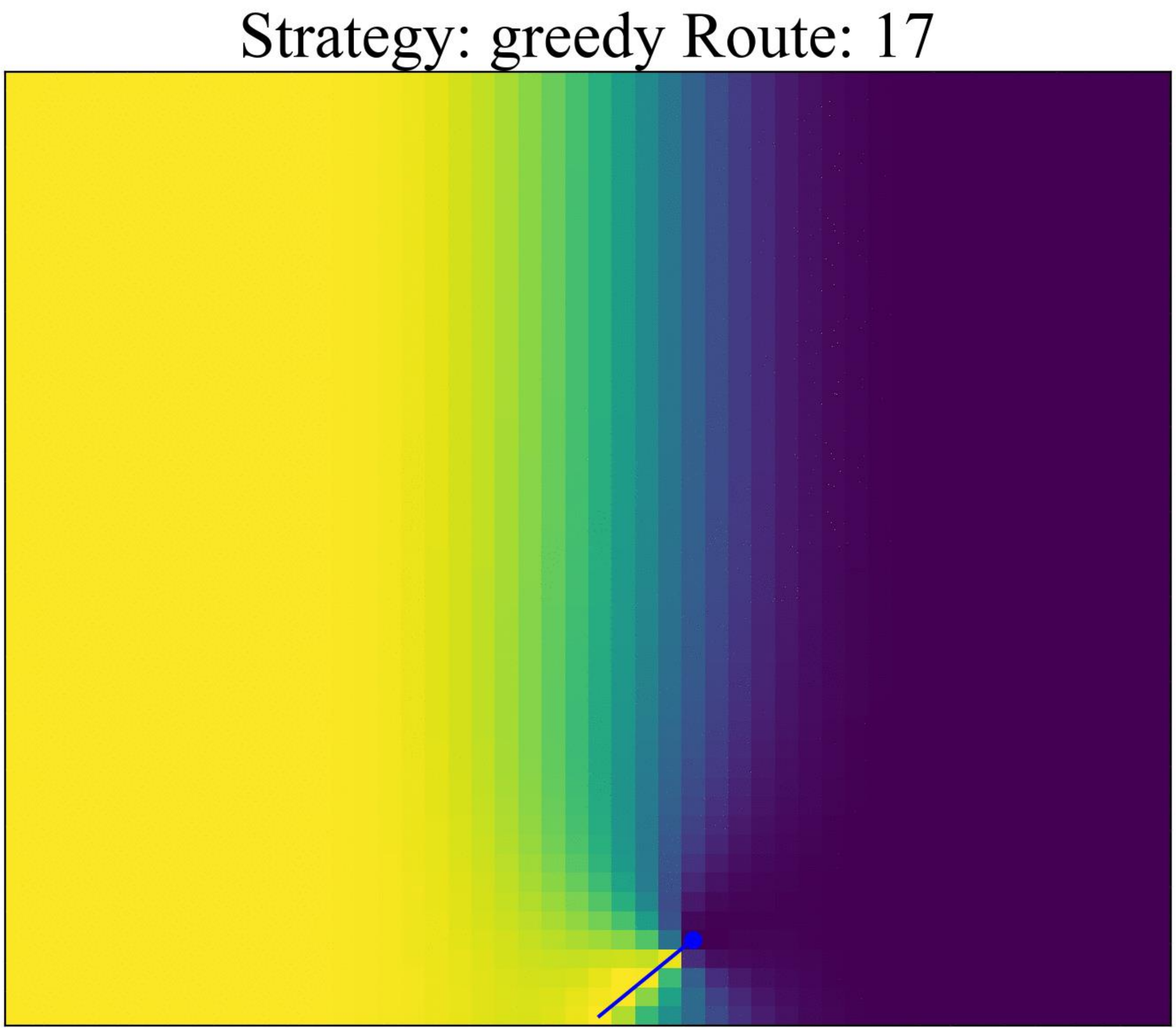

$$\int EP_a(\boldsymbol{s} \mid \boldsymbol{y}_d)\left(1 - EP_a(\boldsymbol{s} \mid \boldsymbol{y}_d)\right) p(\boldsymbol{y}_d)\, d\boldsymbol{y}_d$$

$$= \Phi_2(\boldsymbol{0}_2, \boldsymbol{\beta}_2, \boldsymbol{R}_2) - \Phi_4(\boldsymbol{0}_4, \boldsymbol{\beta}_4, \boldsymbol{R}_4)$$

Multivariate Gaussian cumulative distribution function

Standard matrix –vector computations.

# Myopic path selection for excursions




Real time excursion probability (blue = cold fresh water, yellow = salt warm water.

# Summary:

- VOI is useful for choosing what, where or how to gather data.
- The analysis is done before the actual data gathering, using expected values.
- VOI use value function which can be economic.
  Also related to i) Active learning, ii) Design of experiments.
- VOI is largest when (prior) decision making is difficult, with high data accuracy and lots of decision flexibility.
- With dependence, information will be valuable away from the measurement locations.