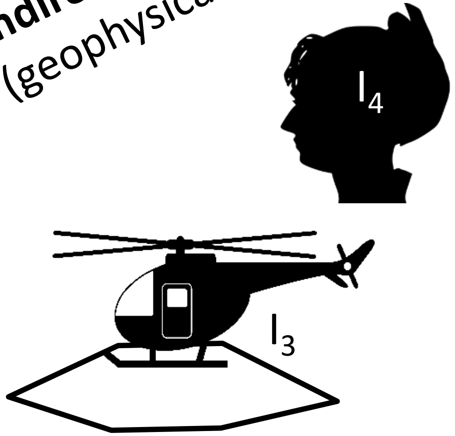
PROBABILISTIC INTEGRATION OF (GEO)-INFORMATION

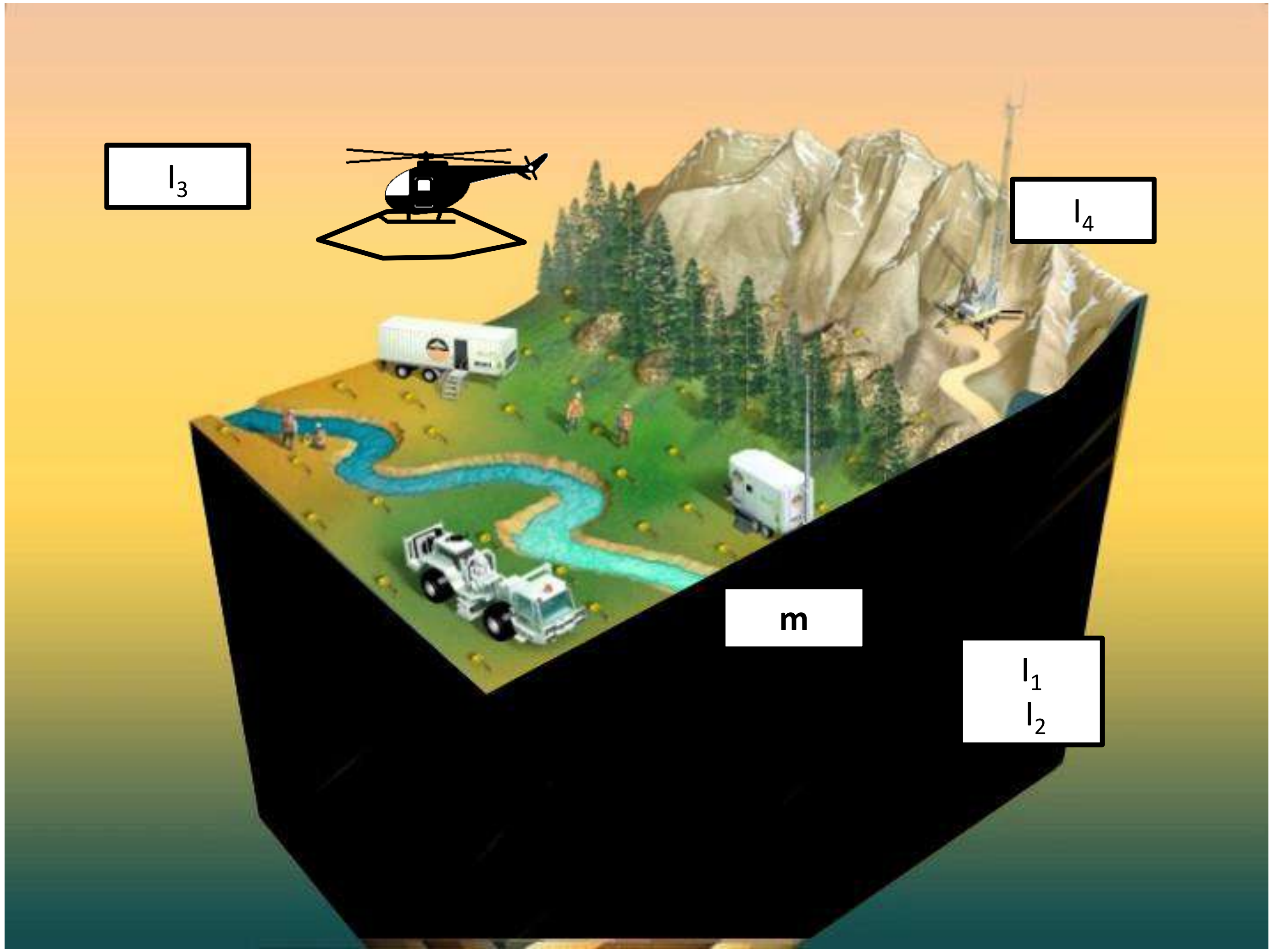
THOMAS MEJER HANSEN

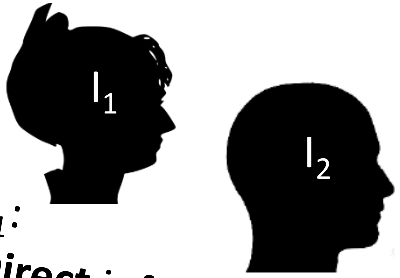
AARHUS UNIVERSITY
DEPARTMENT OF GEOSCIENCE

INTEGRATION OF GEO-INFORMATION

I_3, I_4:
Indirect information
(geophysical data)







I_1:
Direct information
(prior, structural)

(Tarantola and Valette, 1982; Hansen et al., 2016)

AARHUS
UNIVERSITY
DEPARTMENT OF GEOSCIENCE

OUTLINE

Probabilistic integration of GEO-information

- Algorithms
- Prior Information
- Modeling error

Examples (Clay?, UXO, Pollution)

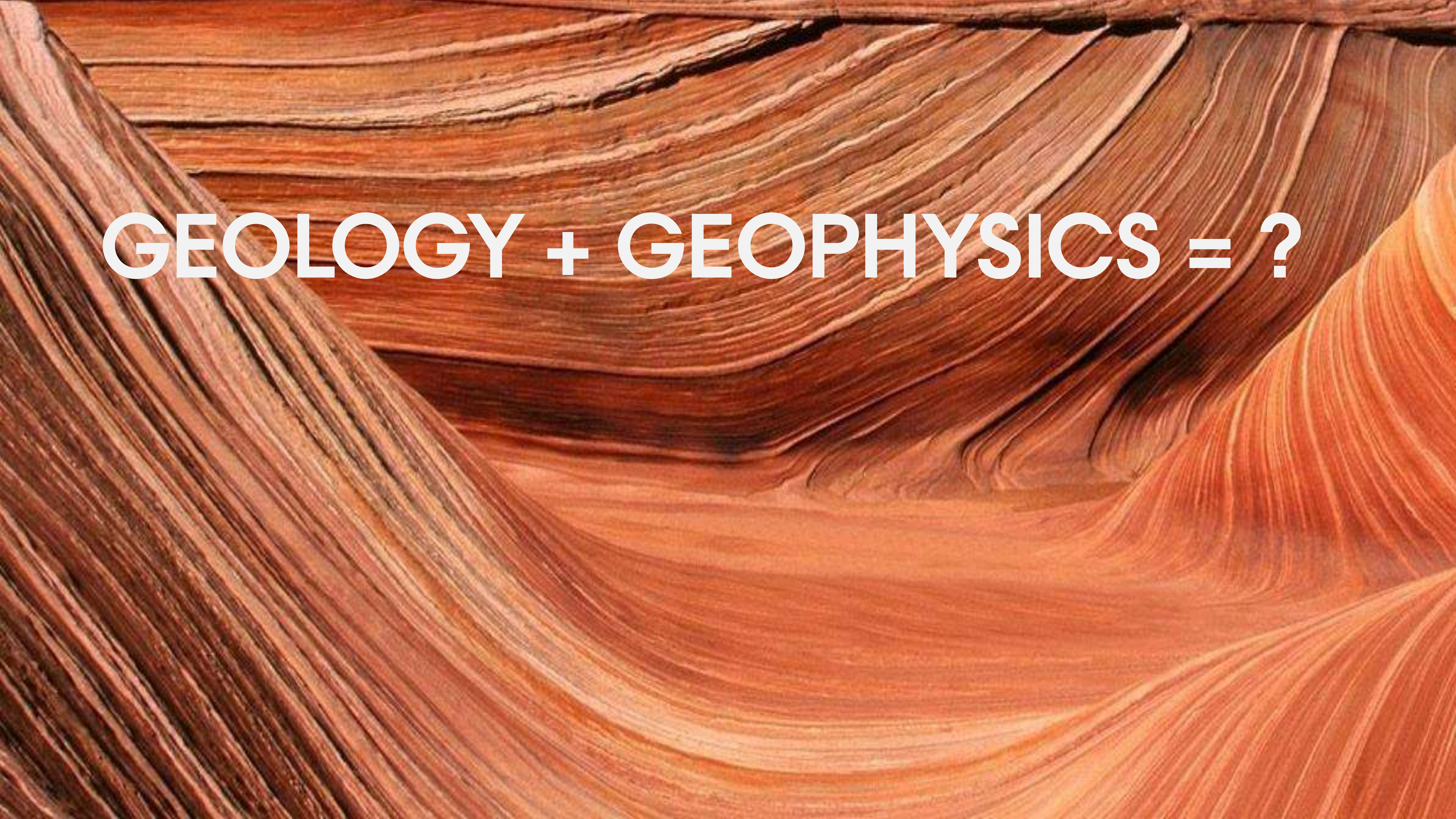
Airborne electromagnetic data

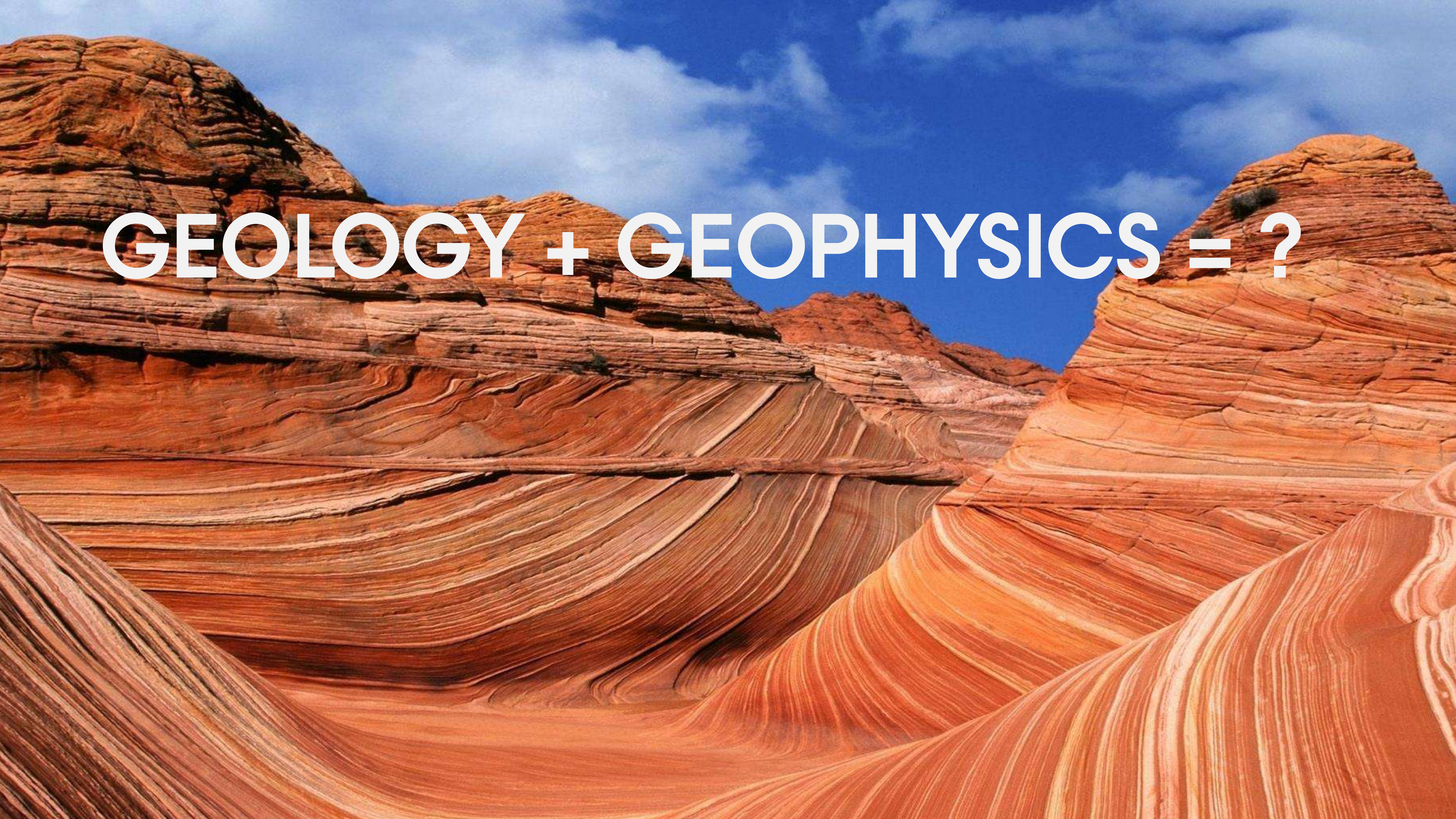
- Extended Metropolis algorithm
- Extended rejection sampling
- Machine learning

Probabilistic PET image analysis

GEOLOGY + GEOPHYSICS = ?

GEOLOGY + GEOPHYSICS = ?

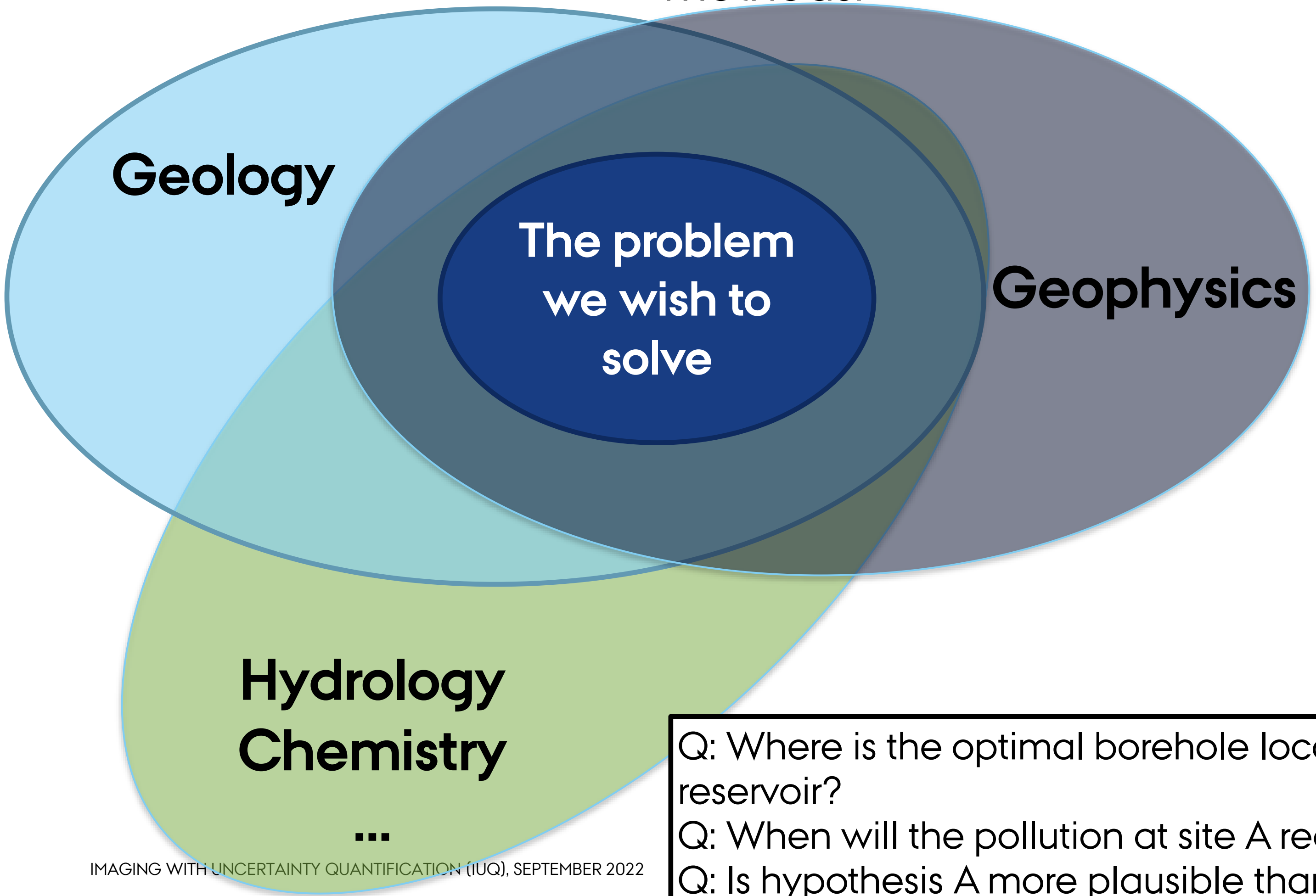



GEOLOGY + GEOPHYSICS = ?

GEOLOGY VS. GEOPHYSICS

"**Geology** is the study of the Earth"

"**Geophysics** is the study of the Earth by quantitative physical methods."

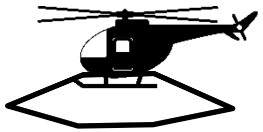


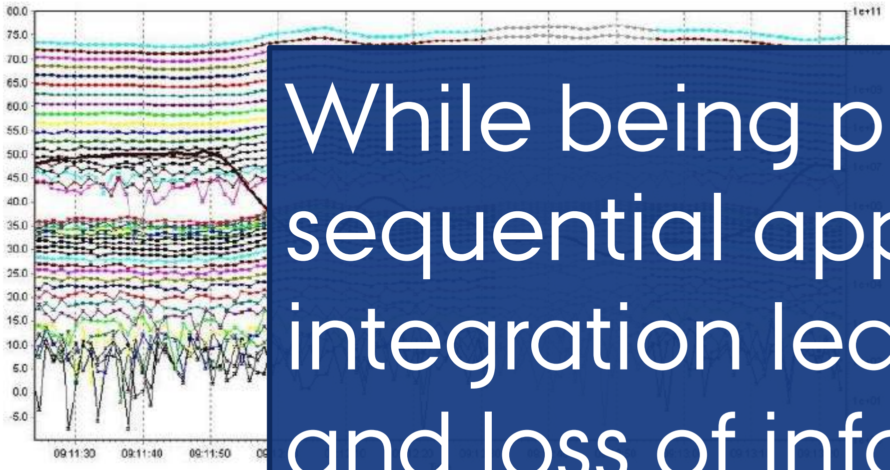
Q: Where is the optimal borehole location to target a water reservoir?
Q: When will the pollution at site A reach point B?
Q: Is hypothesis A more plausible than hypothesis B?

Sequential Information Integration in **GeoScience**

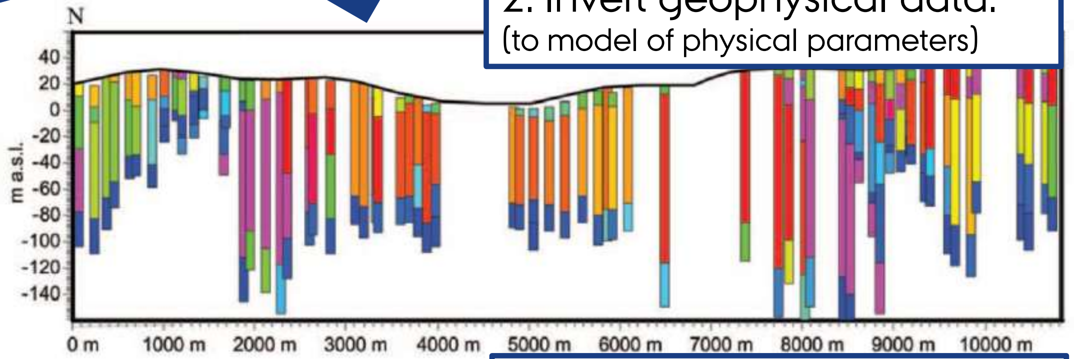
[Jørgensen et al., 2005]

1. Obtain geophysical data

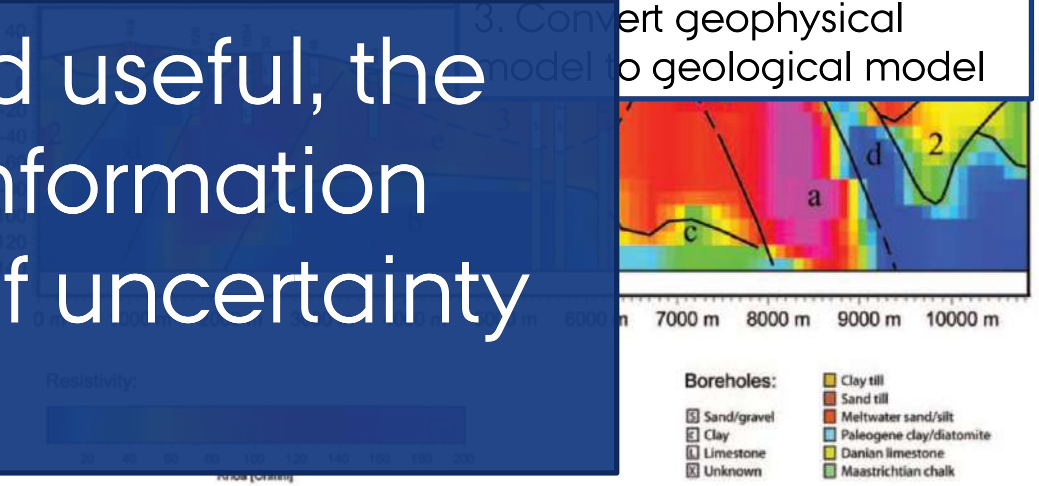




2. Invert geophysical data.
(to model of physical parameters)



3. Convert geophysical model to geological model



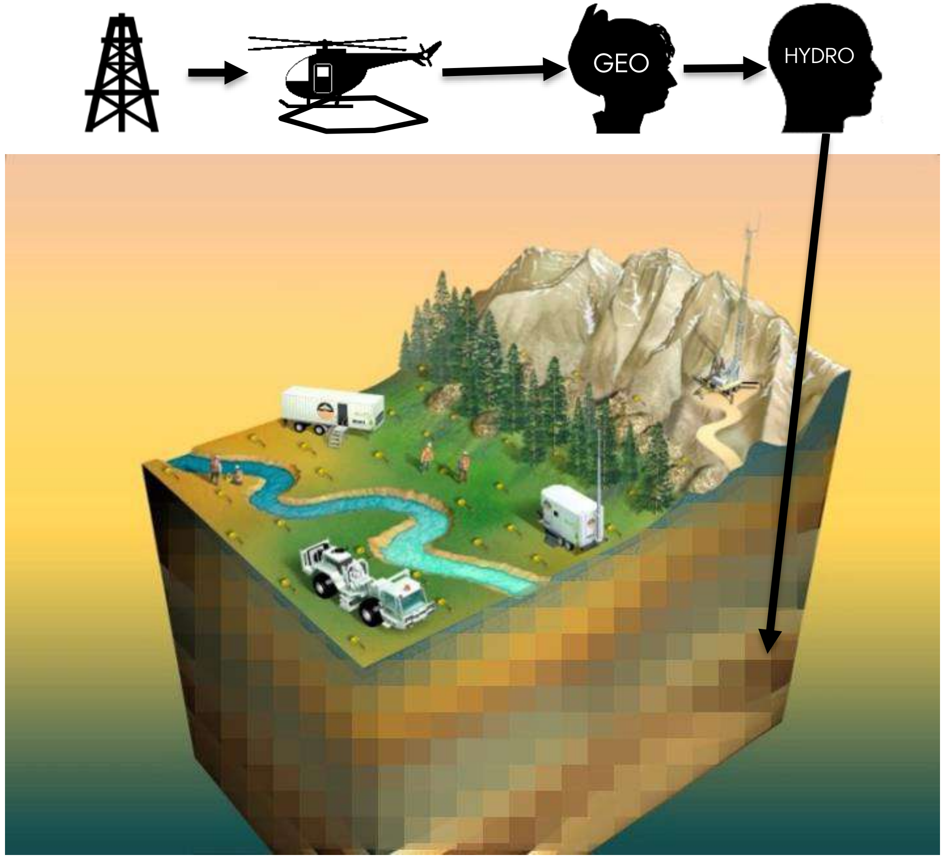
4. Convert geological model to hydrological model

5. Make decisions

While being practical and useful, the sequential approach to information integration leads to loss of uncertainty and loss of information

Sequential workflow

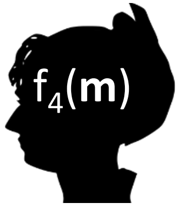
One optimal (smooth) model.
Mathematical regularization.
Consistency ?

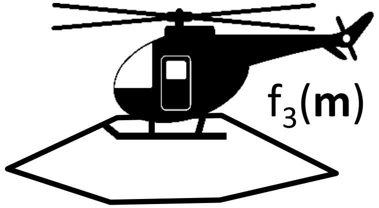


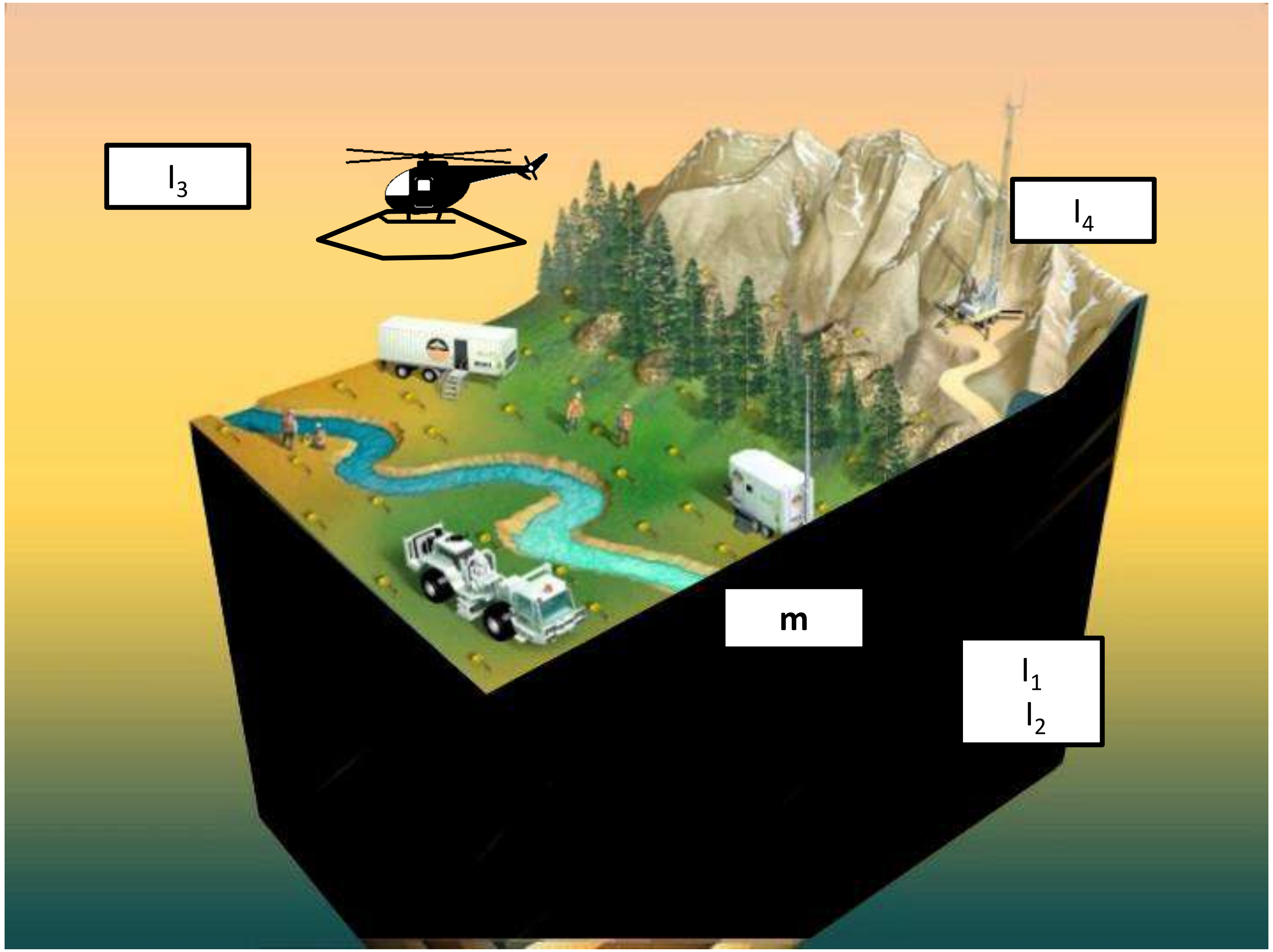
PROBABILISTIC INTEGRATION OF GEO-INFORMATION

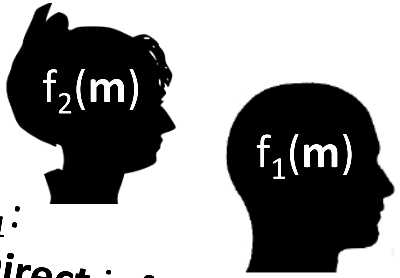
$$f_I(\mathbf{m}) = f(\mathbf{m} \mid I_1, I_2, I_3, I_4 \ldots) = k\, f_1(\mathbf{m})\, f_2(\mathbf{m})\, f_3(\mathbf{m})\, f_4(\mathbf{m}) \ldots$$

I_3, I_4:
Indirect information
(geophysical data)
$L(d-g(\mathbf{m})|I_3)$









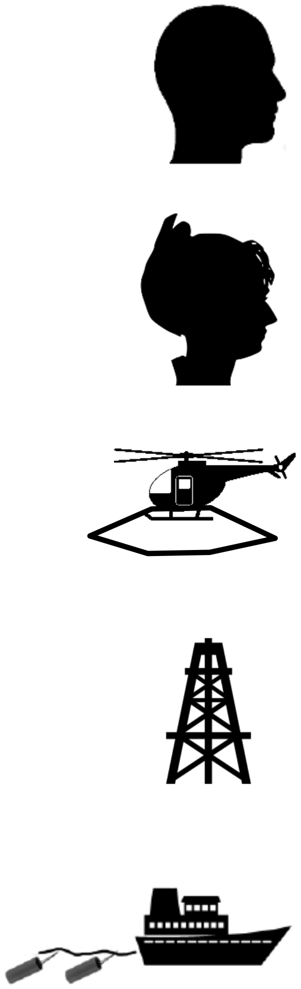
I_1:
Direct information
(prior, structural)

AARHUS UNIVERSITY
DEPARTMENT OF GEOSCIENCE

(Tarantola and Valette, 1982; Hansen et al., 2016)

Data Integration: $f_I(m)$

INFORMATION



$\rightarrow I_1: f_1(m)$

$\rightarrow I_2: f_2(m)$

$\rightarrow I_3: f_3(m)$

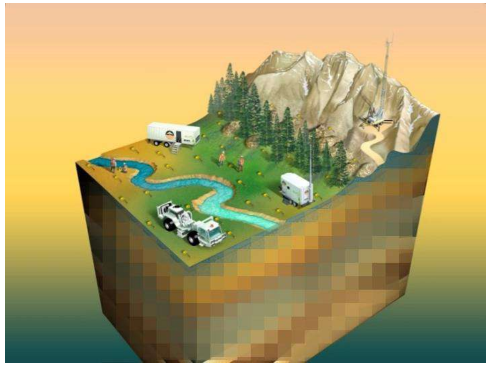
$\rightarrow I_4: f_4(m)$

$\rightarrow I_5: f_5(m)$

INVERSION / INTEGRATION
ALGORITHM

Calibration

Deterministic IP, $m^{optimal}$



INFORMED DECISIONS

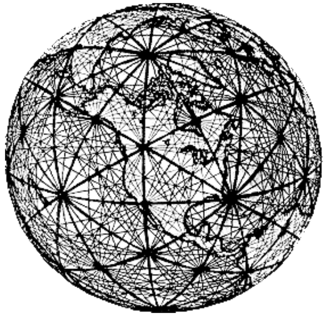
What if … ?

Is the groundwater safe to using pesticides at the surface?

Will the pollution reach XXX?

CONJUNCTION OF INFORMATION

$\mathbf{m} = [m_1, m_2, \ldots]$



[parameterized Earth]

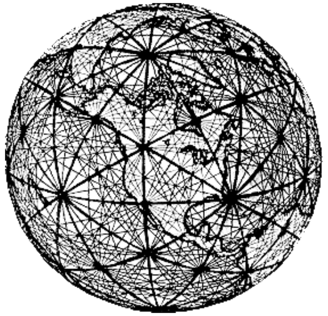
$$f_1(\mathbf{m}) \wedge f_2(\mathbf{m}) \wedge f_3(\mathbf{m}) \wedge f_4(\mathbf{m}) \wedge \ldots \propto \prod_{i=1}^{Ni} f_i(\mathbf{m})$$

$f_i(\mathbf{m})$ must be obtained independently from $f_j(\mathbf{m})$, for all sets of [i,j].

Tarantola and Valette, 1982: Inverse Problems = Quest for information

CONJUNCTION OF INFORMATION

$\mathbf{m} = [m_1, m_2, \ldots]$



[parameterized Earth]

$$\sigma_M(\mathbf{m}) = k\, \rho_M(\mathbf{m})\, L(\mathbf{m})$$

$$L(\mathbf{m}) = \int_D d\mathbf{d}\, \frac{\rho_D(\mathbf{d})\theta(\mathbf{d}|\mathbf{m})}{\mu_D(\mathbf{d})}$$

$$\approx \rho_D(\mathbf{g}(\mathbf{m}))$$

Tarantola and Valette, 1982: Inverse Problems = Quest for information

THE EXTENDED METROPOLIS ALGORITHM

The goal: Sample from $\sigma(\mathbf{m}) = k\, \rho(\mathbf{m})\, L(\mathbf{m})$

0. Propose a starting from $\rho(\mathbf{m}) \rightarrow \mathbf{m}_{cur}$
   - 1. Propose a model from $\rho(\mathbf{m}) \rightarrow \mathbf{m}_{pro}$ *in the vicinity of* $\mathbf{m}_{cur}$
   - 2. Accept the move from $\mathbf{m}_{cur}$ to $\mathbf{m}_{pro}$ with probability
     - $P_{acc} = L(\mathbf{d}_{obs} - g(\mathbf{m}_{pro})) / L(\mathbf{d}_{obs} - g(\mathbf{m}_{cur}))$
   - 3. Store the current model, $\mathbf{m}_{cur}$.

4. Go to 1 (until enough realizations have been generated).

To apply the extended Metropolis algorithm, one must be able to

1) Sample the prior $\rho(\mathbf{m})$, <u>through a random walk.</u>
2) Evaluate the likelihood $L(\mathbf{m})$ for any model, (solve the forward problem, evaluate the noise)

Mosegaard and Tarantola (1995)

THE EXTENDED REJECTION SAMPLER

The goal: Sample from $\sigma(\mathbf{m}) = k\,\rho(\mathbf{m})\,L(\mathbf{m})$

1. Propose a model from $\rho(\mathbf{m}) \rightarrow \mathbf{m}^*$
2. Accept $\mathbf{m}^*$ as a realization from the posterior with probability

   $P_{acc} = L(\mathbf{d}_{obs} - g(\mathbf{m}^*))/\max(L)$

To apply the extended rejection sampler one must be able to

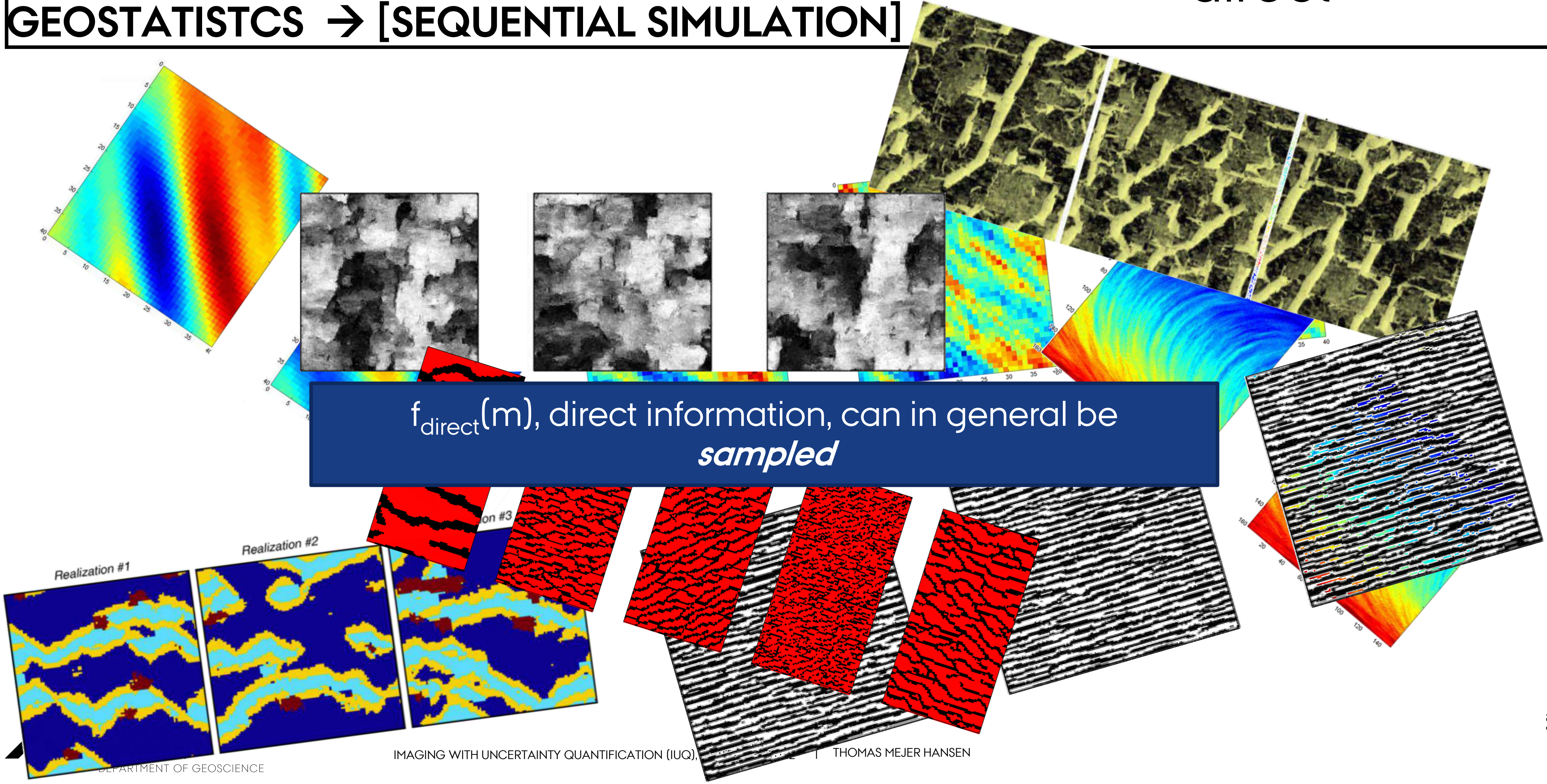
1) Sample the prior $\sigma(\mathbf{m})$
2) Evaluate the likelihood $L(\mathbf{m})$ for any model, (solve the forward problem, evaluate the noise)

Lookup table: [$\mathbf{M}^*$, $\mathbf{D}^*$] → Compute once, apply for any $\mathbf{d}_{obs}$

Hansen (2021)

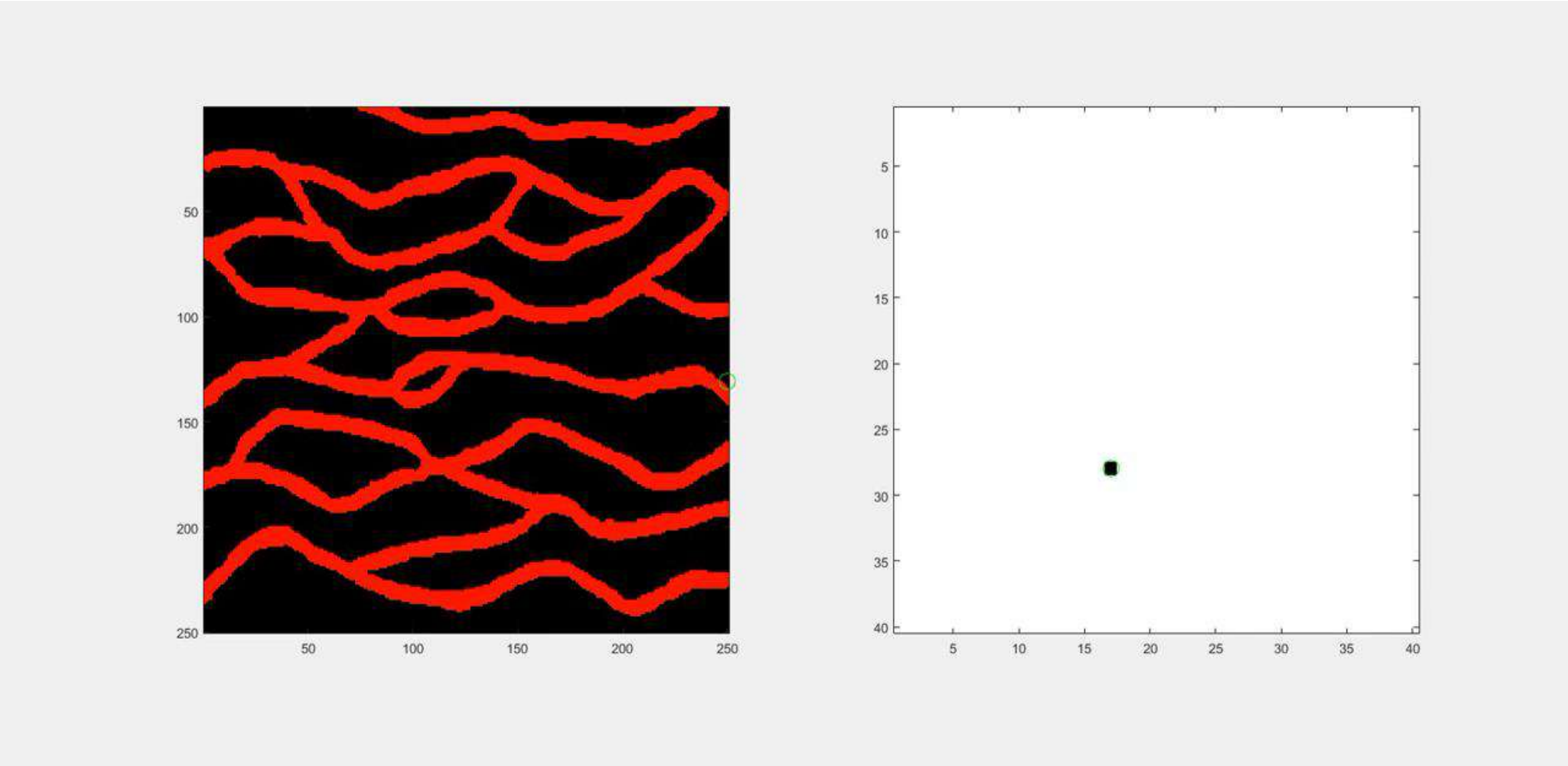
DIRECT (PRIOR) INFORMATION, $f_{direct}(\mathbf{m})$

GEOSTATISTCS → [SEQUENTIAL SIMULATION]



$f_{direct}(m)$, direct information, can in general be ***sampled***

Sequential Simulation
$f(\mathbf{m})=f(m_1)\, f(m_2|m_1^*)\, f(m_3|m_1^*, m_2^*)....$

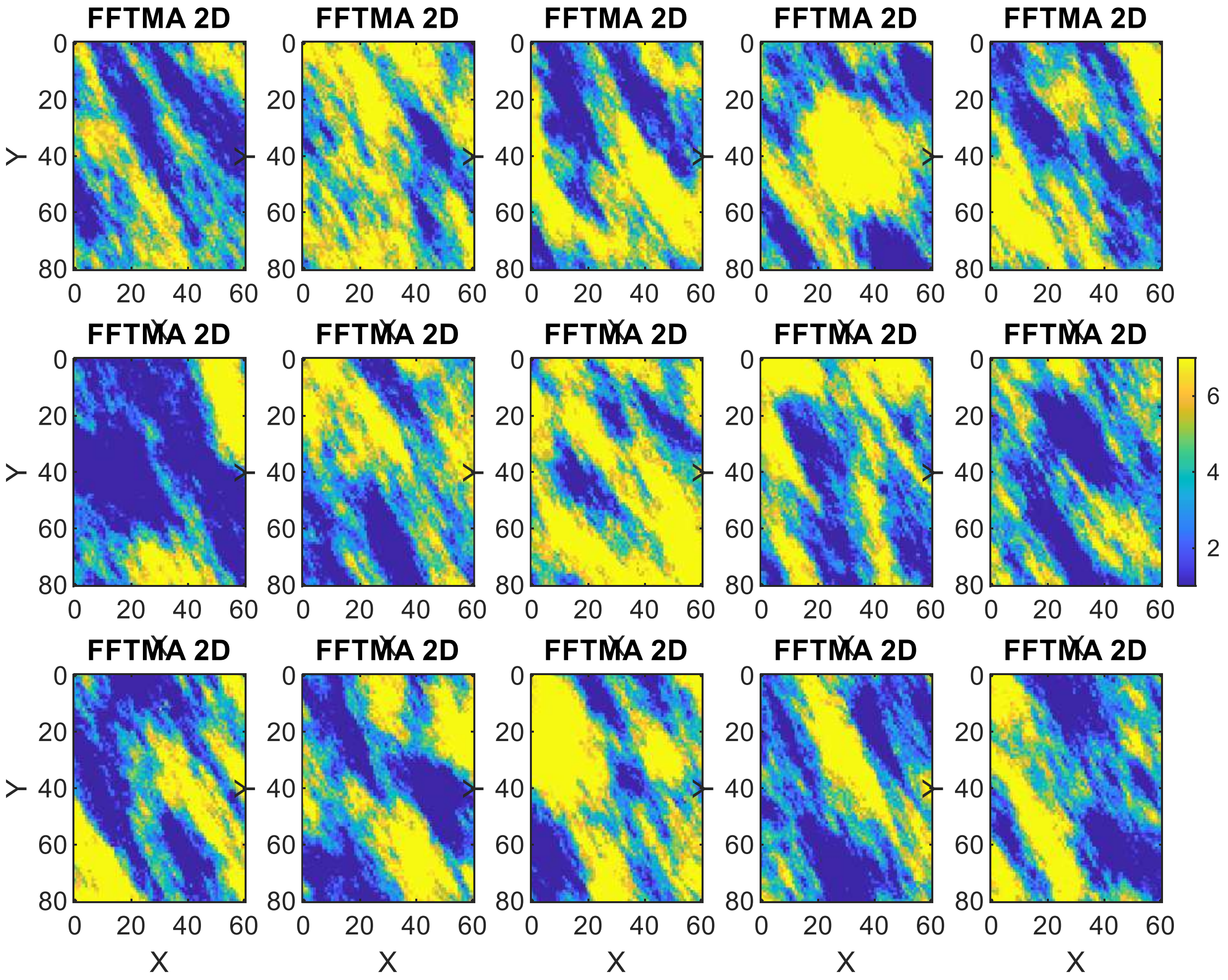


(Guardiano and Srivastava, 1992)

2D Gaussian prior

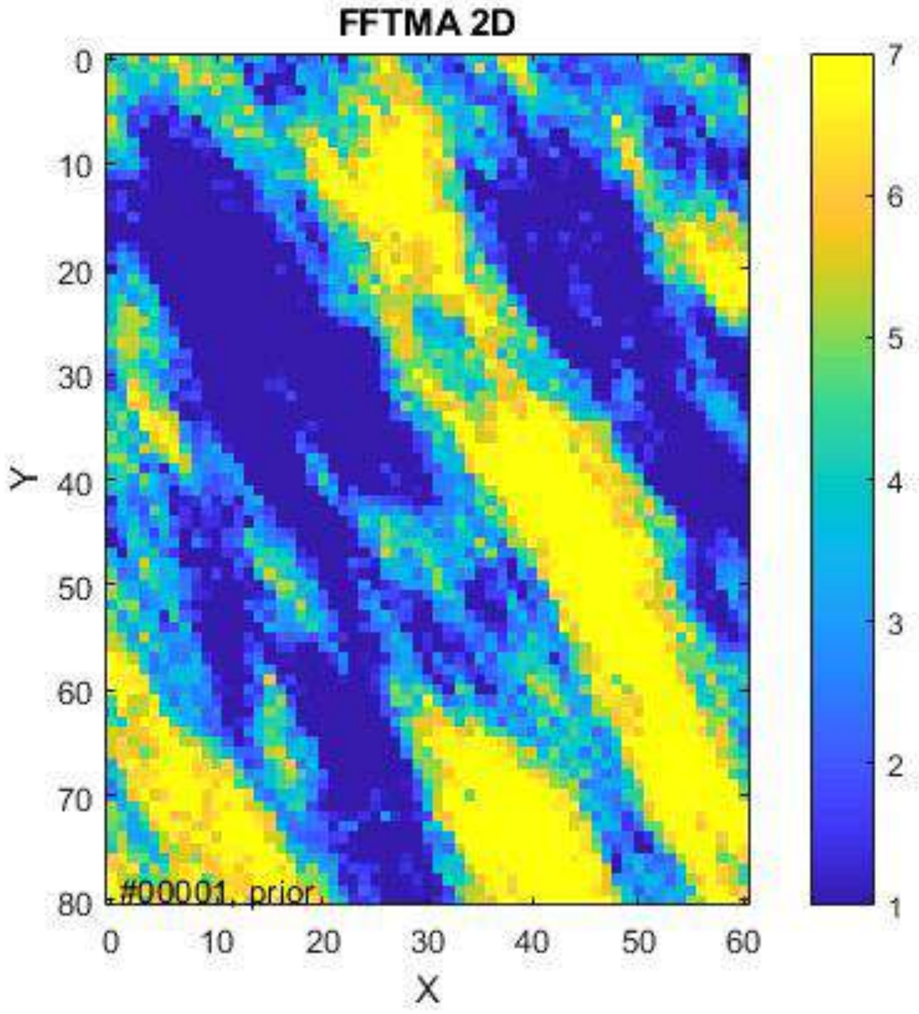
```
ip=1;
prior{ip}.type='FFTMA';
prior{ip}.name='FFTMA 2D';
prior{ip}.x=0:1:60;
prior{ip}.y=0:1:80;
prior{ip}.m0  = 4;
prior{ip}.Va='10 Sph(10,90,0.3)';
sippi_plot_prior_sample(prior);
```

From:
https://github.com/cultpenguin/sippi



2D Gaussian prior – random walk

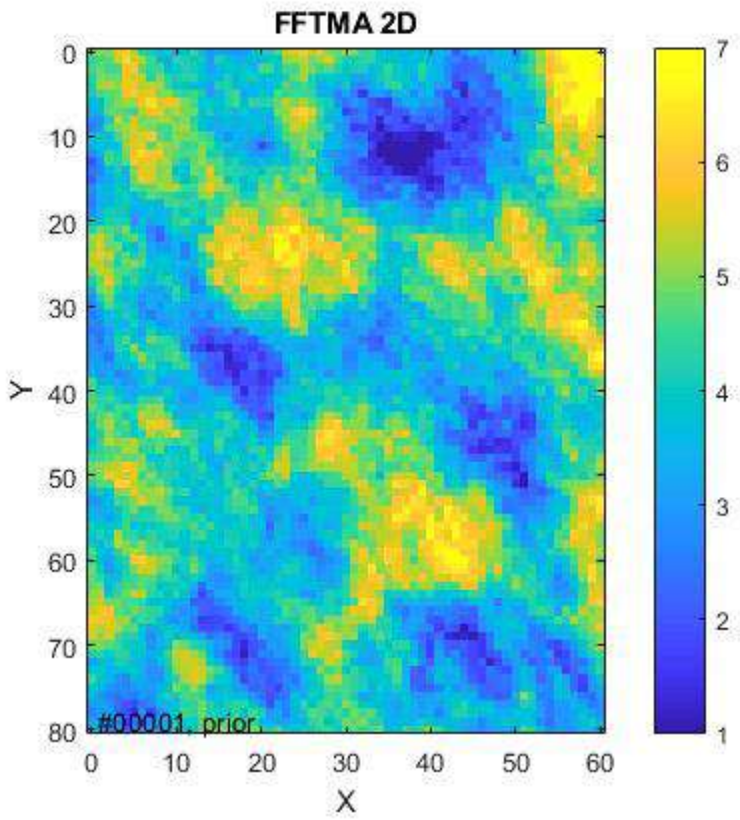
```
ip=1;
prior{ip}.type='FFTMA';
prior{ip}.name='FFTMA 2D';
prior{ip}.x=0:1:80;
prior{ip}.y=0:1:60;
prior{ip}.m0  = 4;
prior{ip}.Va='10 Sph(10,90,0.3)’;

prior{ip}.seq_gibbs.step  = .02;
sippi_plot_prior_movie(prior,200);
```

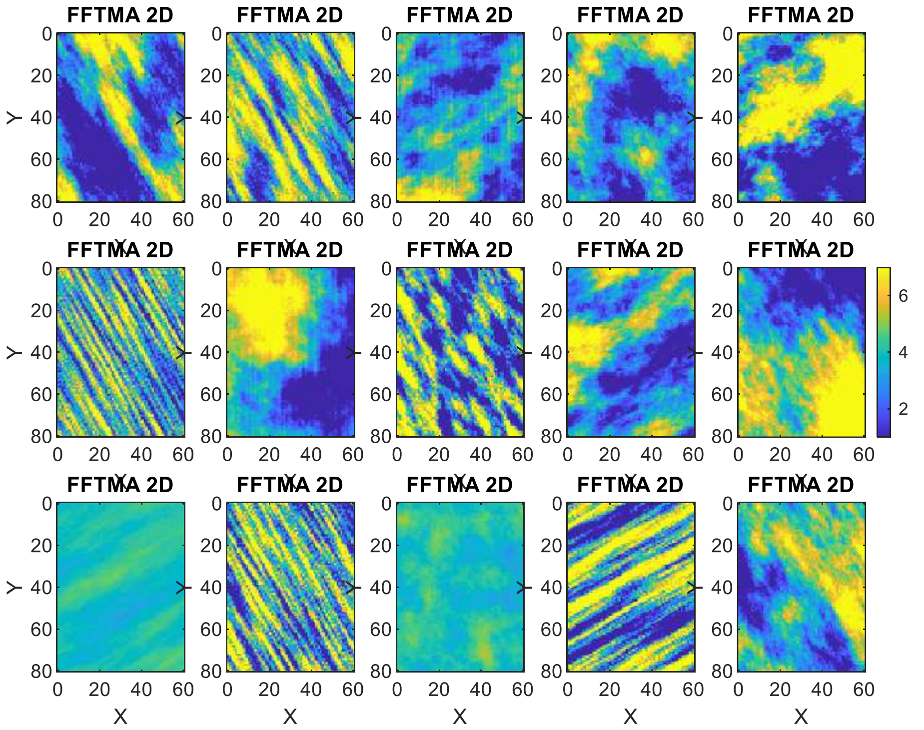


(Hansen et al., 2016)

2D Gaussian prior – random walk – uncertain covariance model parameters

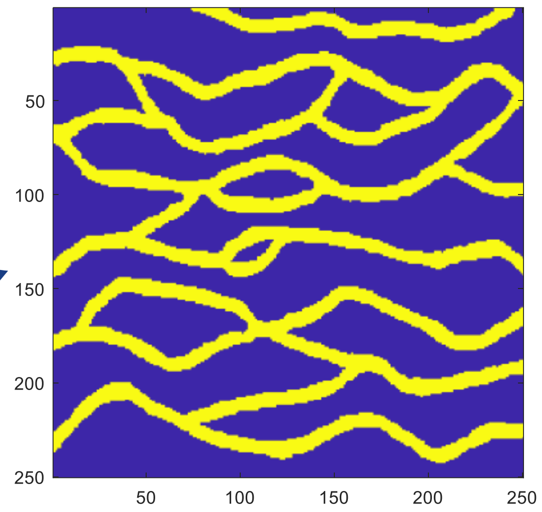
```
clear prior
ip=1;
prior{ip}.type='FFTMA';
prior{ip}.name='FFTMA 2D';
prior{ip}.x=0:1:60;
prior{ip}.y=0:1:80;
prior{ip}.m0  = 4;
prior{ip}.Va='10 Sph(50,30,0.3)';
prior{ip}.cax=[-3 3]+4;

ip=2;
prior{ip}.name='range_1';
prior{ip}.type='uniform';
prior{ip}.min=1;
prior{ip}.max=80;
prior{ip}.prior_master=1;

ip=3;
prior{ip}.name='range_2';
prior{ip}.type='uniform';
prior{ip}.min=1;
prior{ip}.max=80;
prior{ip}.prior_master=1;

ip=4;
prior{ip}.name='angle_2';
prior{ip}.type='uniform';
prior{ip}.min=-30;
prior{ip}.max=30;
prior{ip}.prior_master=1;

ip=5;
prior{ip}.name='sill';
prior{ip}.type='uniform';
prior{ip}.min=0;
prior{ip}.max=10;
prior{ip}.prior_master=1;

sippi_plot_prior_movie(prior,200);
```

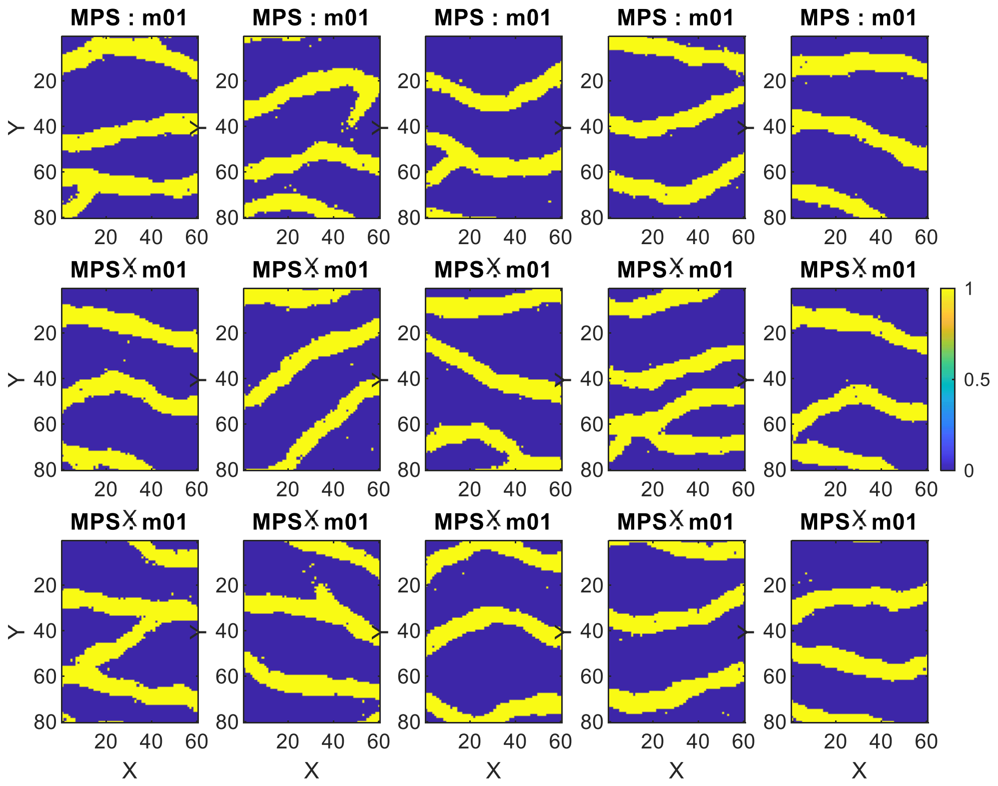




2D Multiple Point Statistical prior

```
clear prior
ip=1;
prior{ip}.type='mps’;
prior{ip}.method='mps_genesim';
prior{ip}.x=1:1:60;
prior{ip}.y=1:1:80;
prior{ip}.ti=channels;

figure(1);
imagesc(prior{ip}.ti);axis image

sippi_plot_prior_sample(prior);
```





PRIOR ASSUMPTION <–> SUBJECTIVE INFORMATION

The use/need of prior information has been debated, Scales and Tenorio (1997)

we try to find a **noninformative**, or conservative, **prior** that **injecting a minimum of artificial information**, that is, information not justified by the physical process.

Can, and should, we avoid prior information?

Benoît Convers

HOW TO AVOID PRIOR/STRUCTURAL INFORMATION?

m=[m_1, m_2 ..., m_{10000}] represents a 2D gray-scale image.

f(**m**) represents our information about **m.**

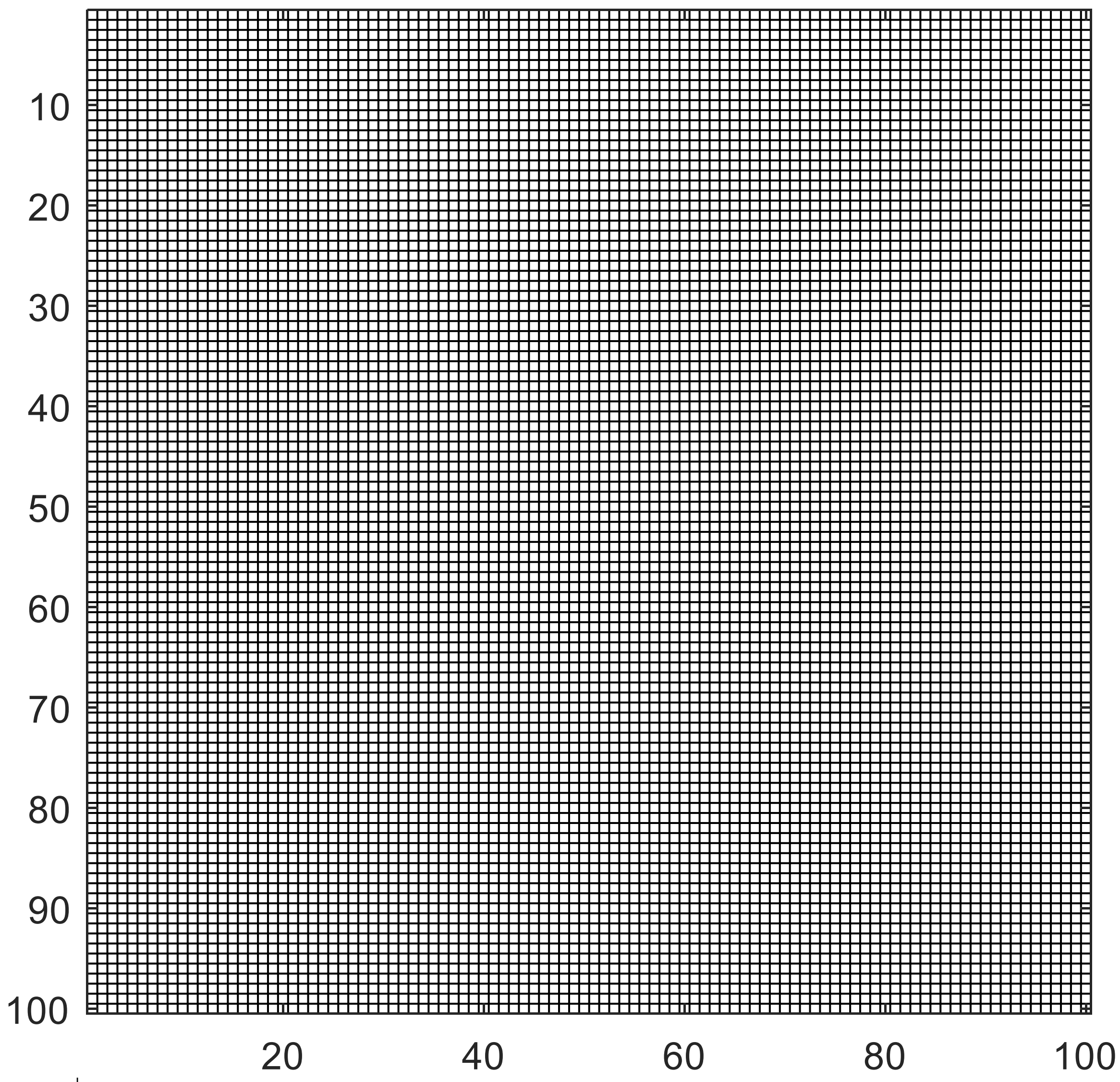
A realization **m*** of f(**m**) represents a specific image.

1. Assume there is no dependence between individual model parameters

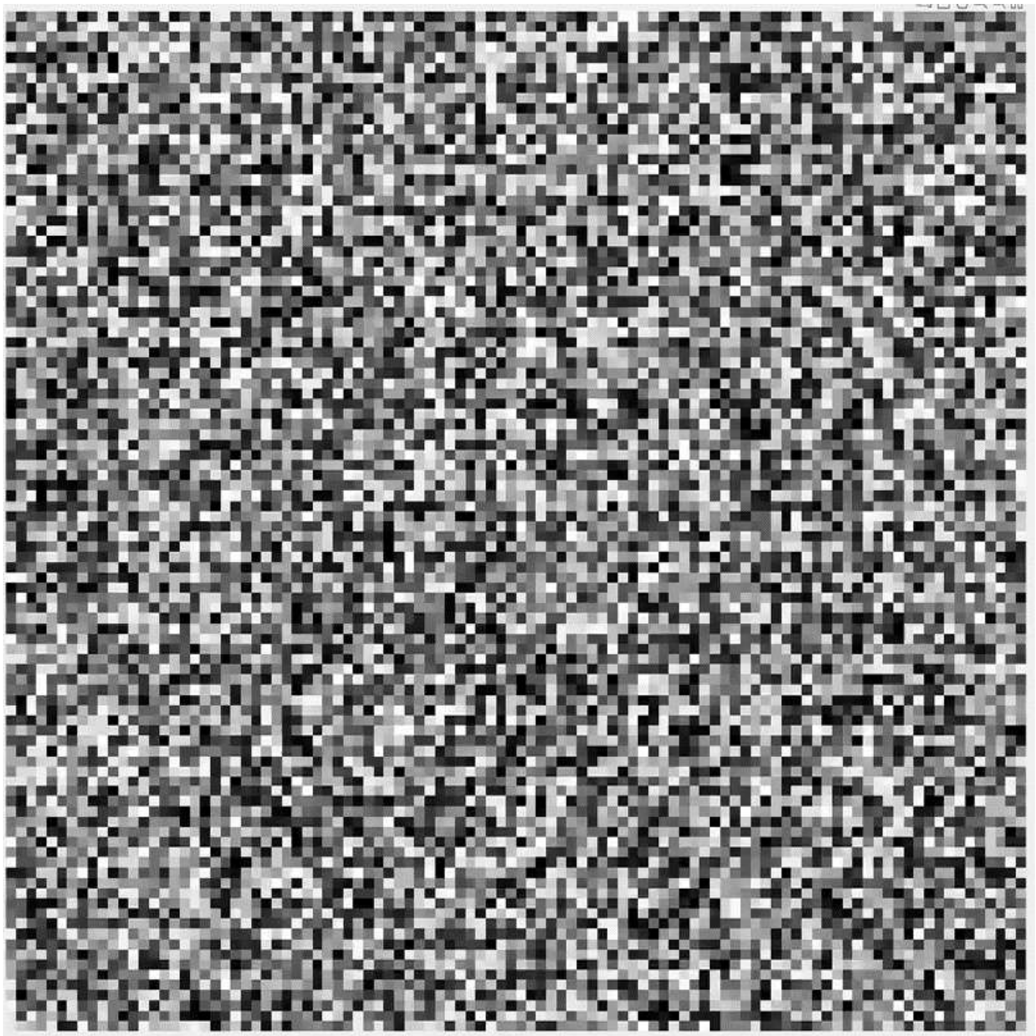
$$f(m_i, m_j)=f(m_i)f(m_j)$$

2. Use a uniform prior distribution

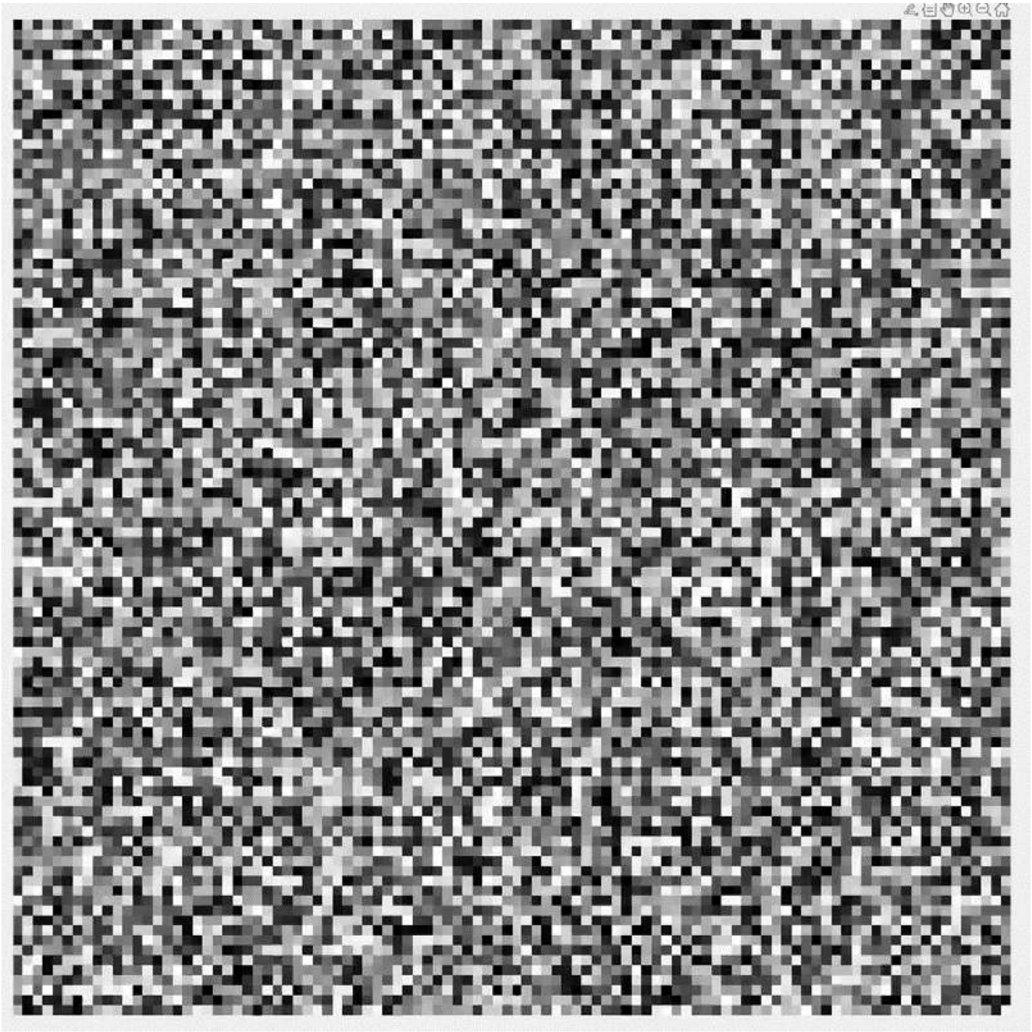
$$f(m_i) = U[0,255]$$

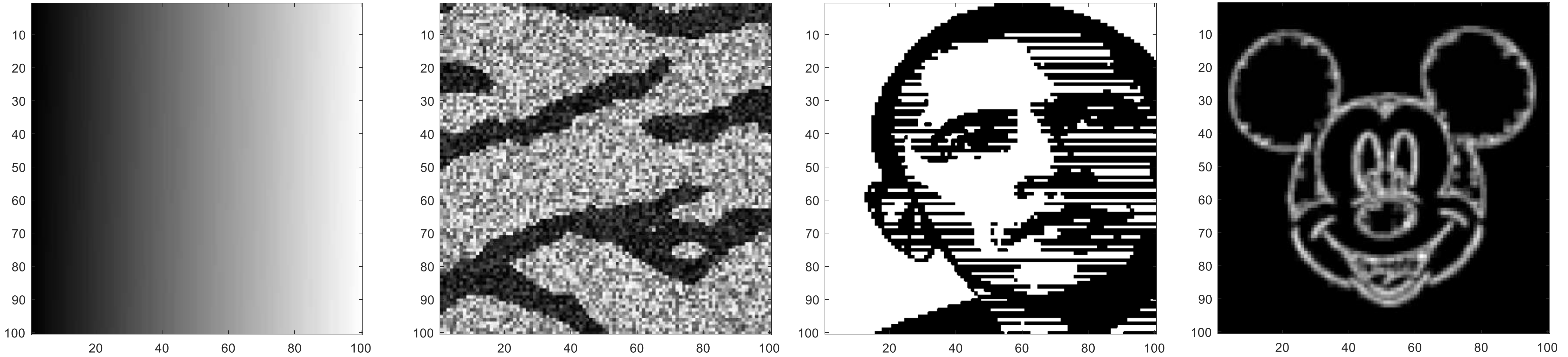


Random walk in f(**m**)



Guided random walk in f(**m**)



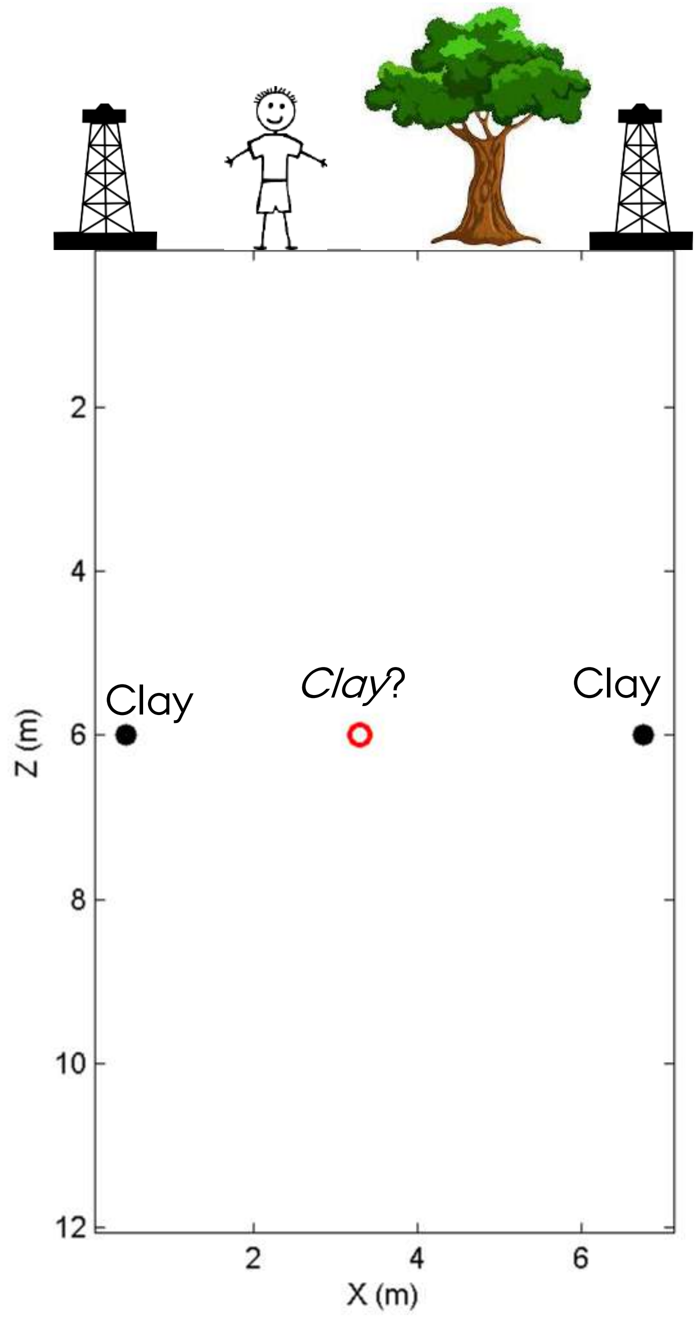


P(Image looks somewhat like Obama) < 1/250000 = 0.000004
P(Image looks somewhat like any structure) < 0.000004
P(Image look like random noise) > 0.999996

Key take home message:
YES, an uniform prior can represent in principle any image/structure.
BUT, the probability of any other feature than random noise exists, tend to zero (as the pixel size becomes smaller).
One cannot assume a uniform prior in order to avoid choosing a prior
A uniform prior is not an “uninformed prior” but a specific choice of maximum disorder!

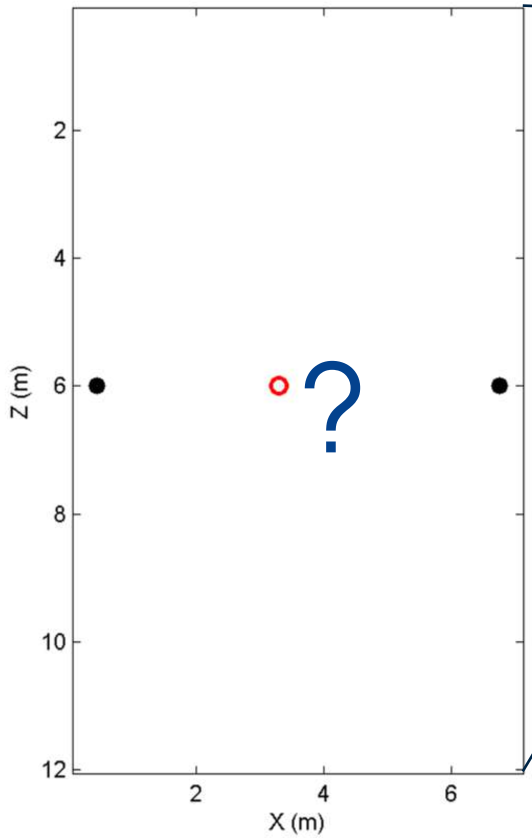
CASE: CLAY OR NOT

EXAMPLE 1: PROBABILITY OF CLAY

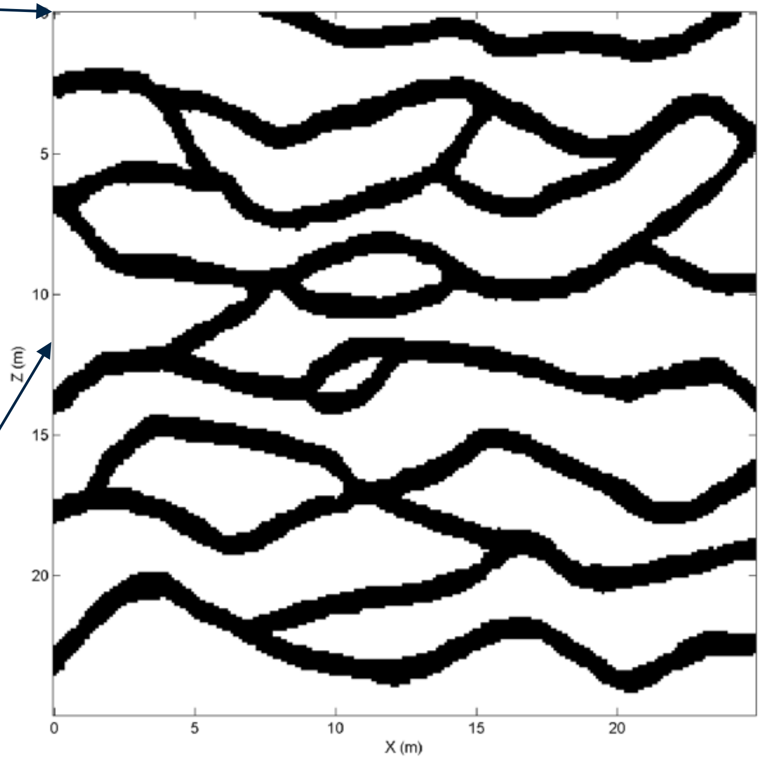


A simple application of geostatistics and inverse problem theory

- We have 2 observations:

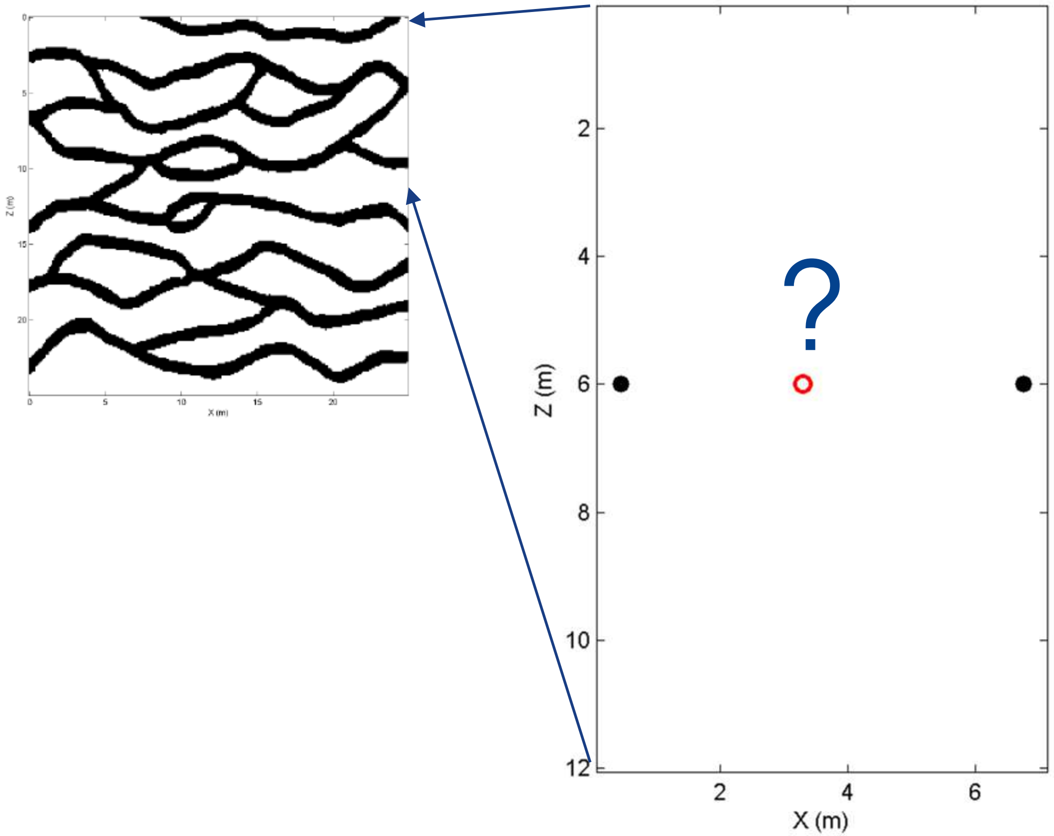


- I_1: Structural information from a training image:



Black: Clay
White: Sand

A simple application of geostatistics and inverse problem theory



Prior probability of channel/clay:

P(o == clay | prior) = 0.30

Posterior probability of channel:

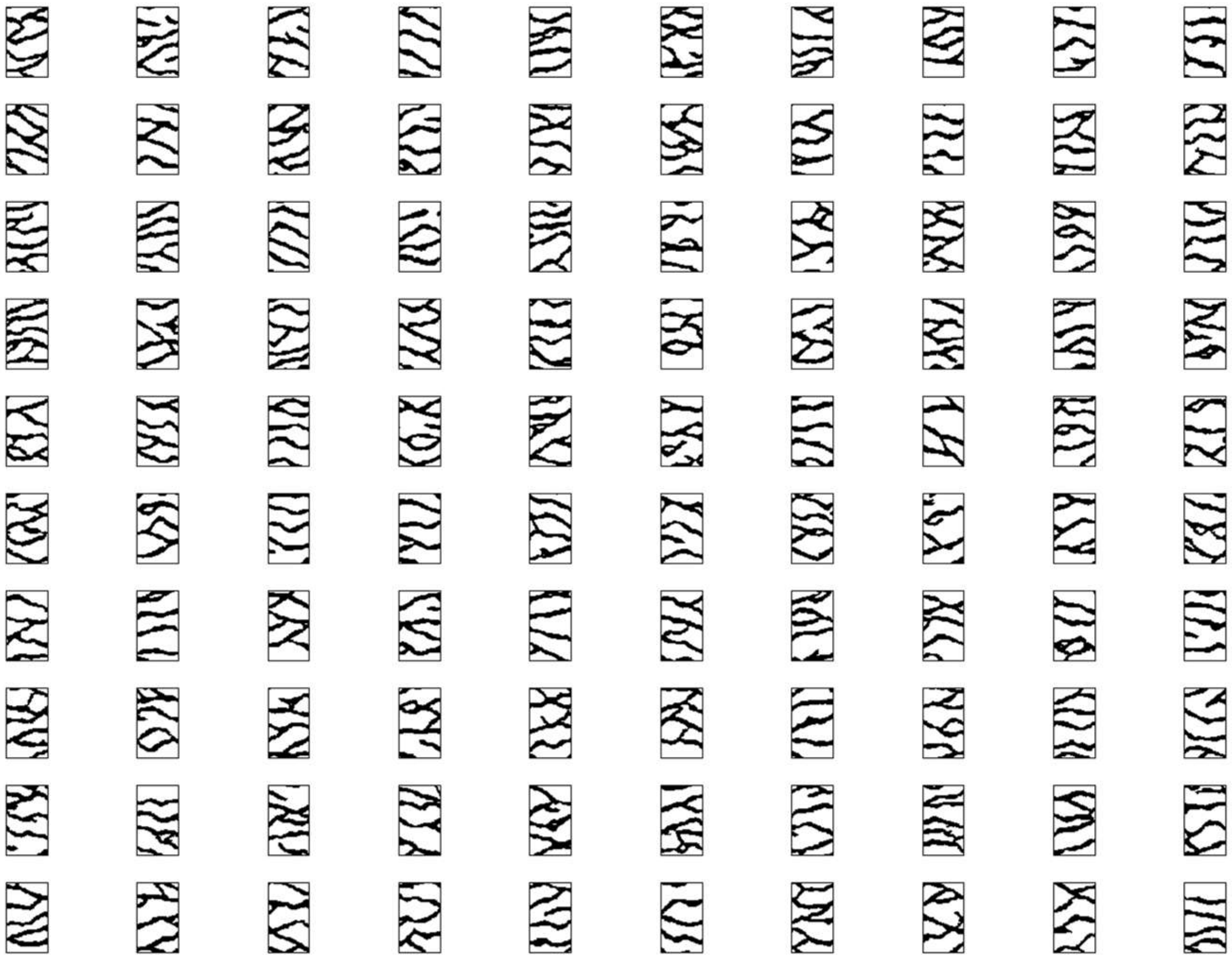
P(o == clay | prior, data) = ?

QUIZ:

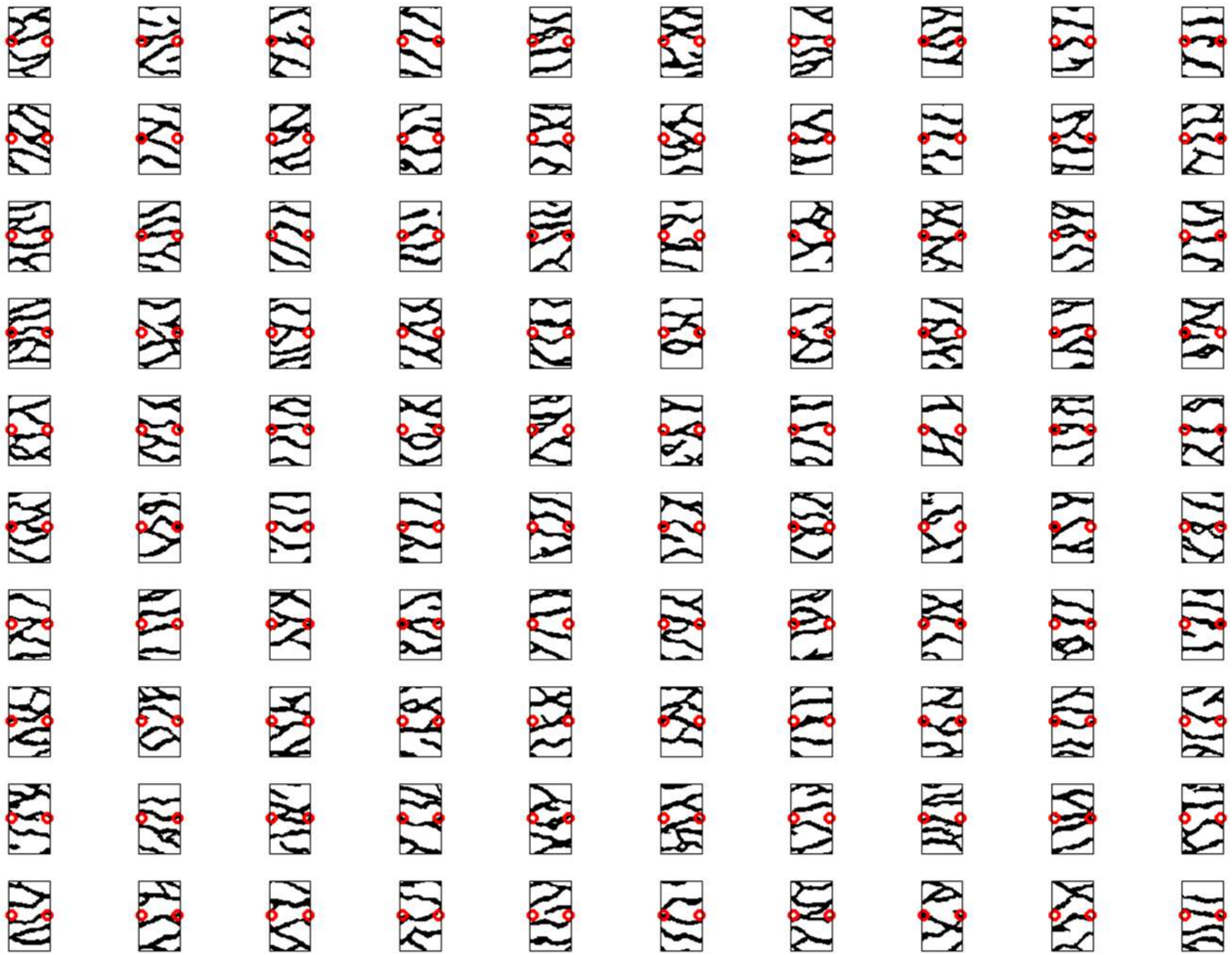
P(o == clay | prior, data) < 0.30

P(o == clay |prior, data) >= 0.30

A simple application of geostatistics and inverse problem theory

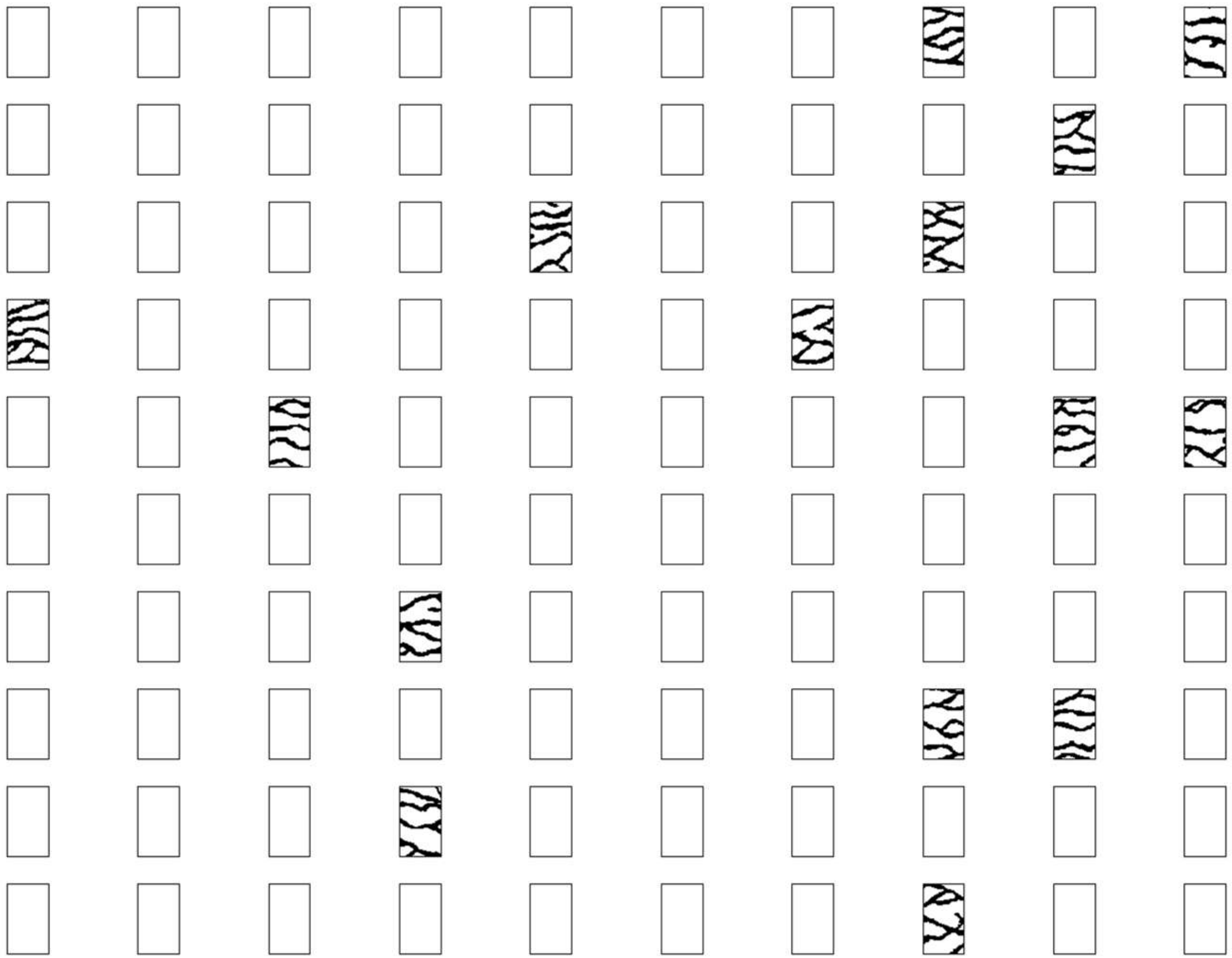


A simple application of geostatistics and inverse problem theory



AARHUS
UNIVERSITY
DEPARTMENT OF GEOSCIEN

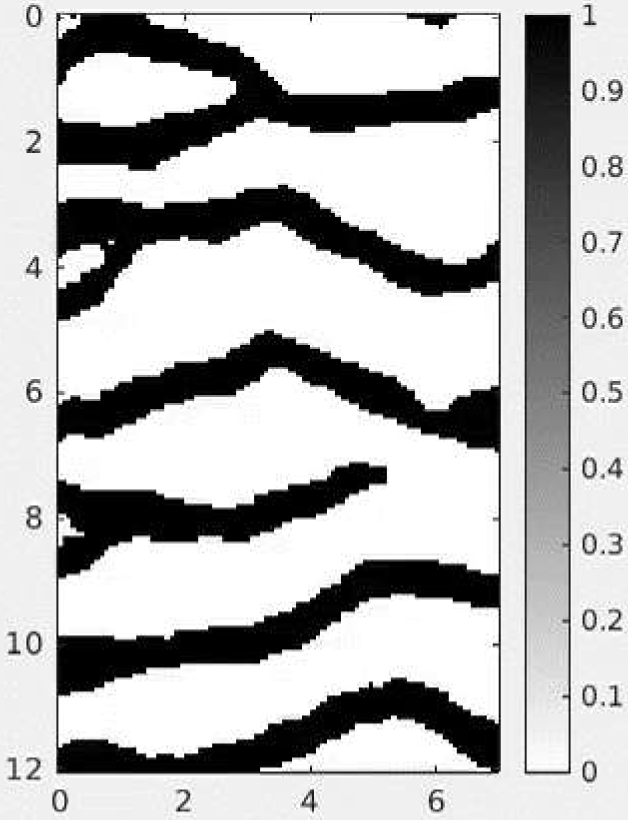
A simple application of geostatistics and inverse problem theory



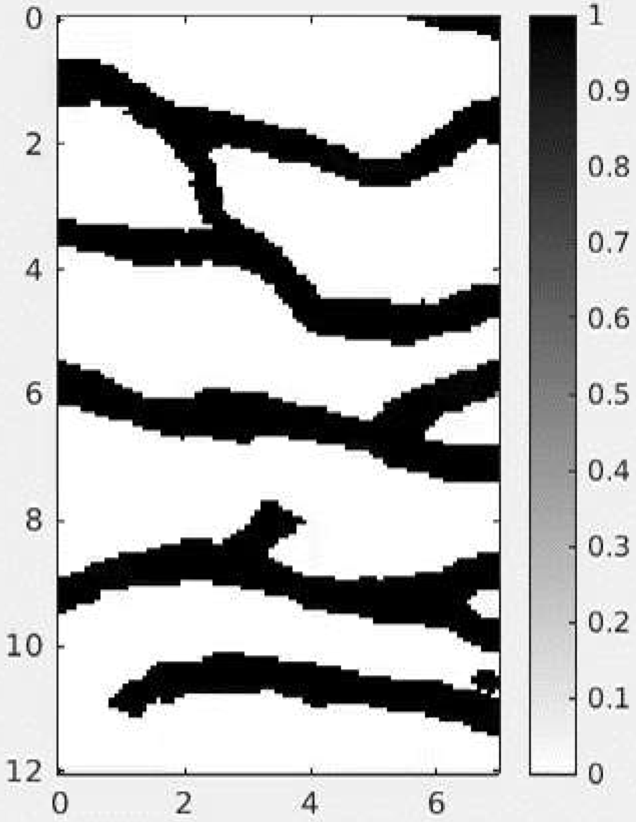
AARHUS UNIVERSITY
DEPARTMENT OF GEOSCIEN

A simple application of geostatistics and inverse problem theory

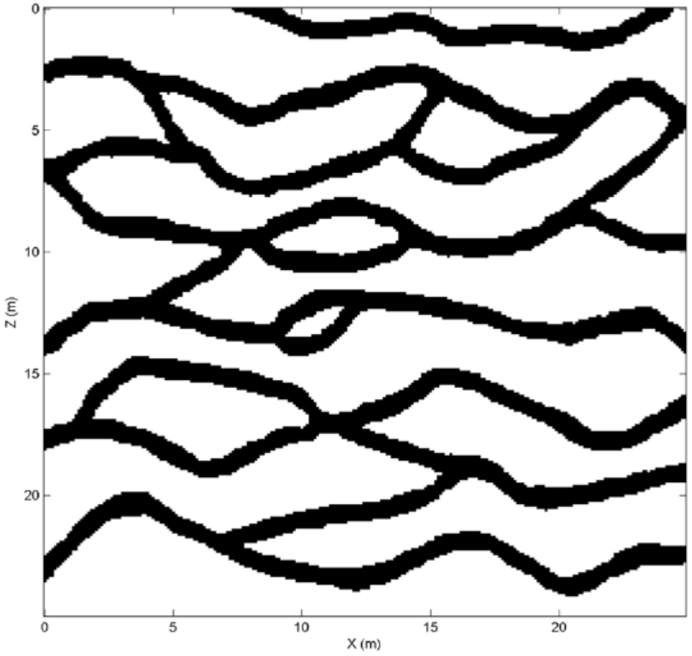
f(m| prior)→m*

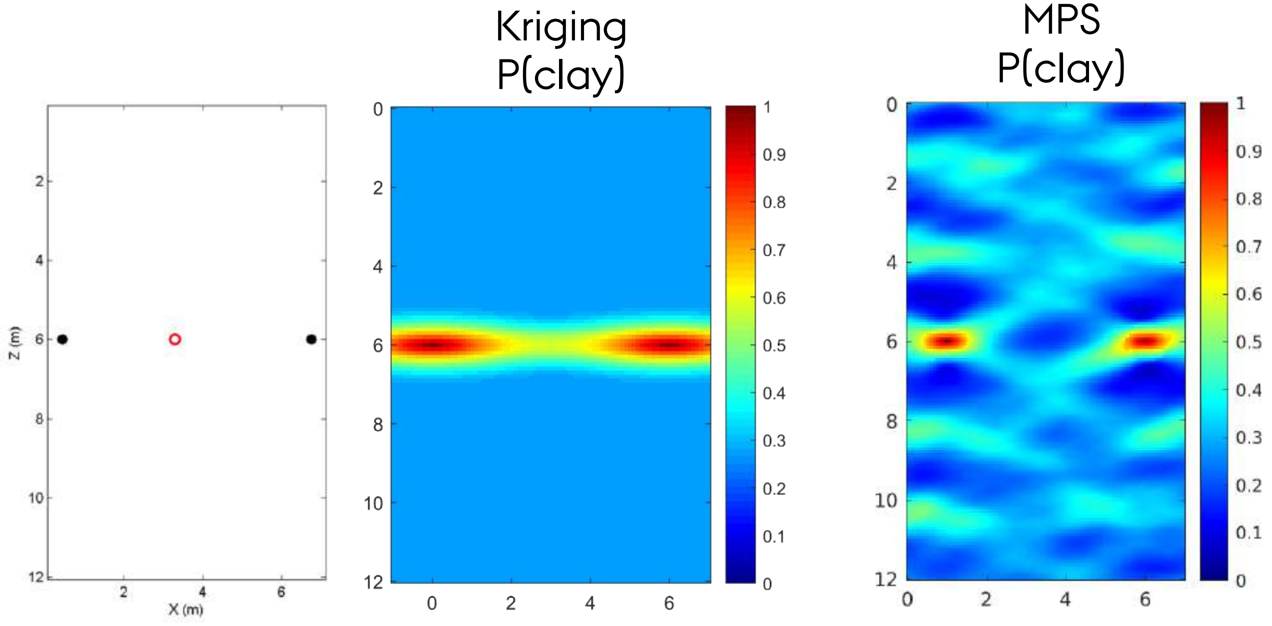


f(m| prior,data)→m*



A simple application of geostatistics and inverse problem theory





P(o == clay | prior) = 0.30

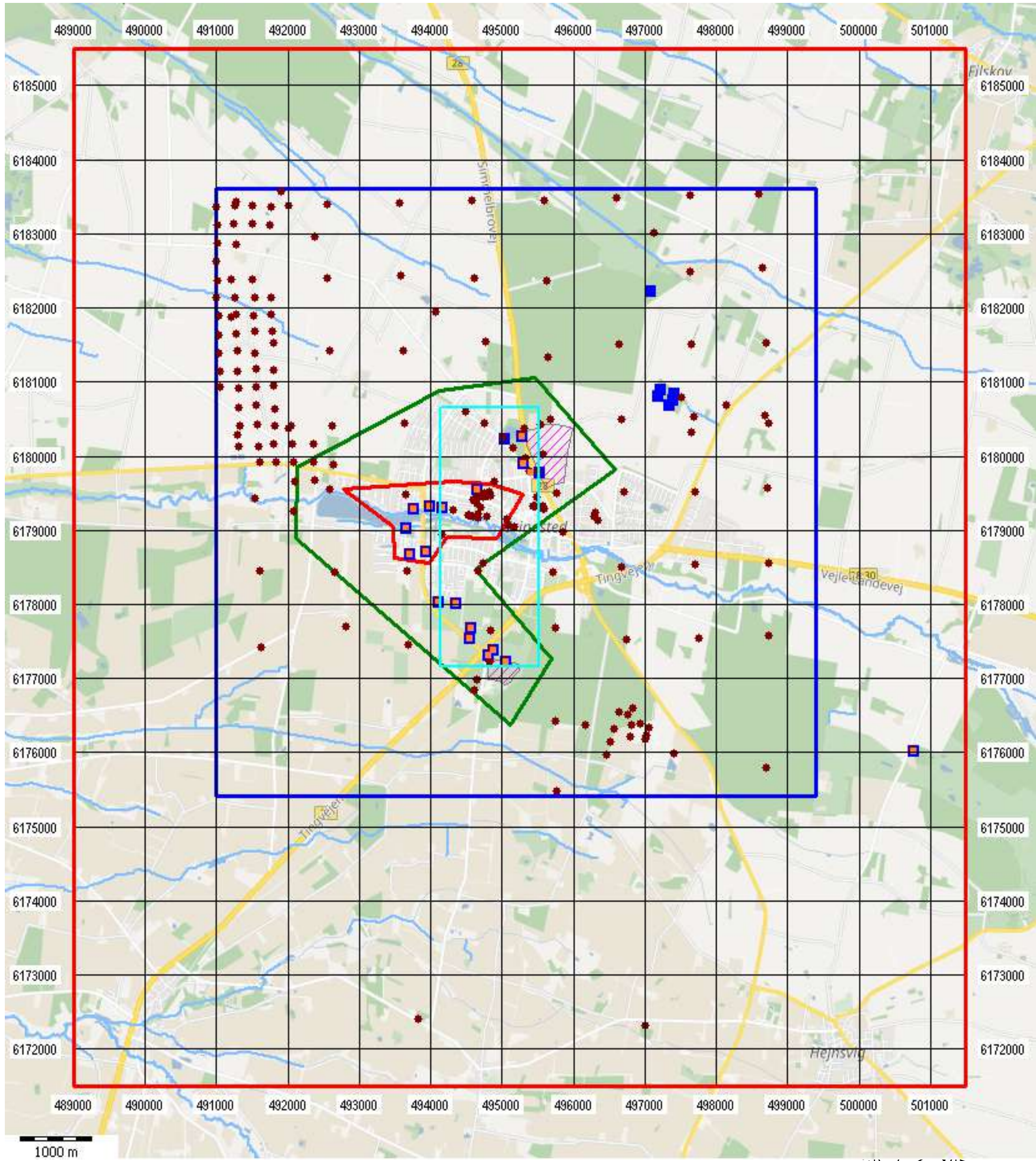
P(o == clay | prior, data) = 0.56 Kriging

P(o == clay | prior, data) = 0.17 MPS

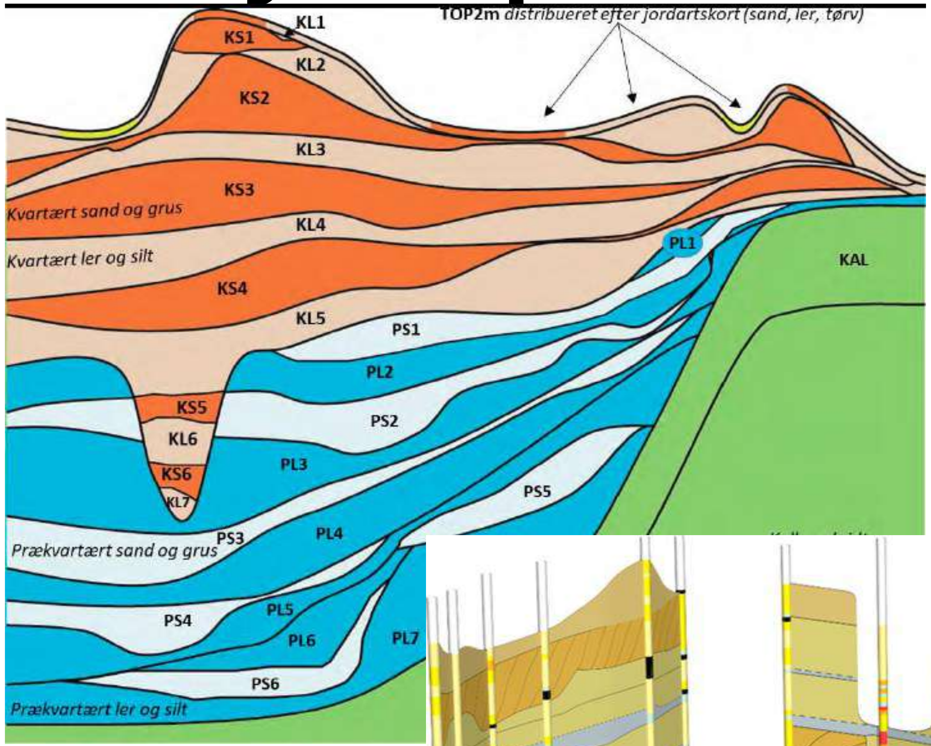
CASE:
MAPING POLLUTION USING GEOLOGICALLY INFORMED

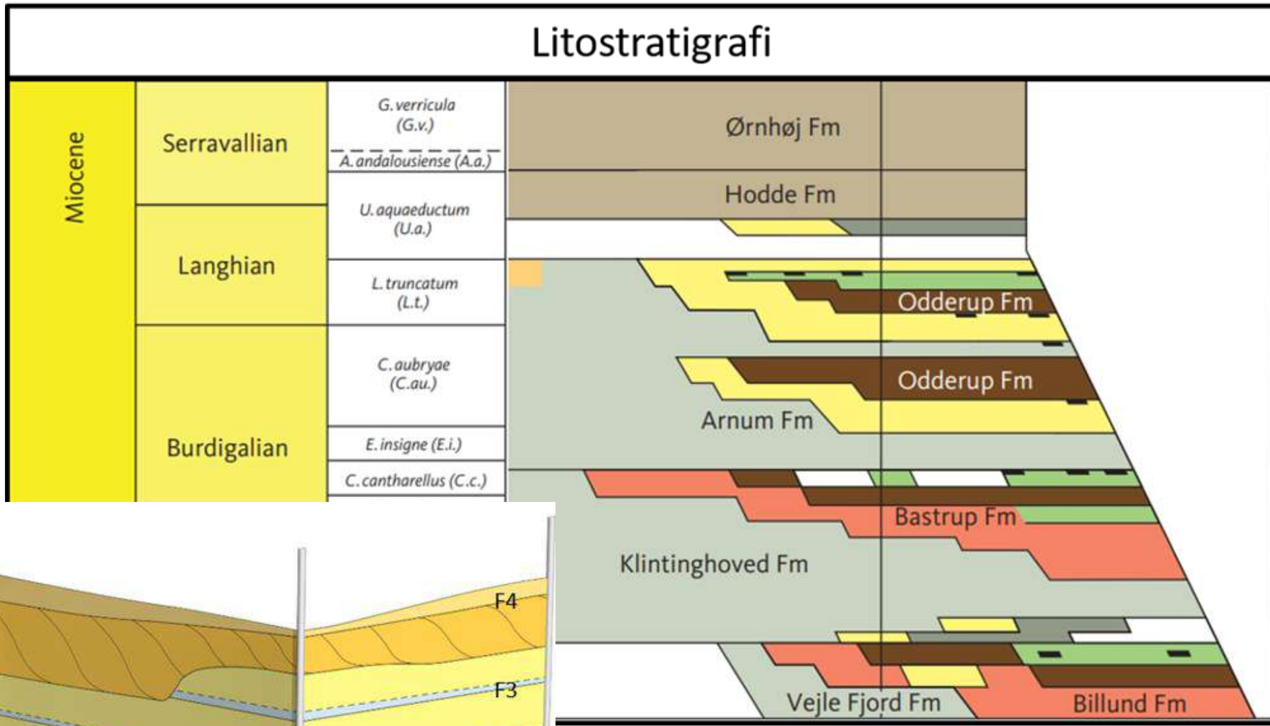
WITH, RASMUS BØDKER MADSEN, INGELISE BALLING, GEUS, REGION MIDT.

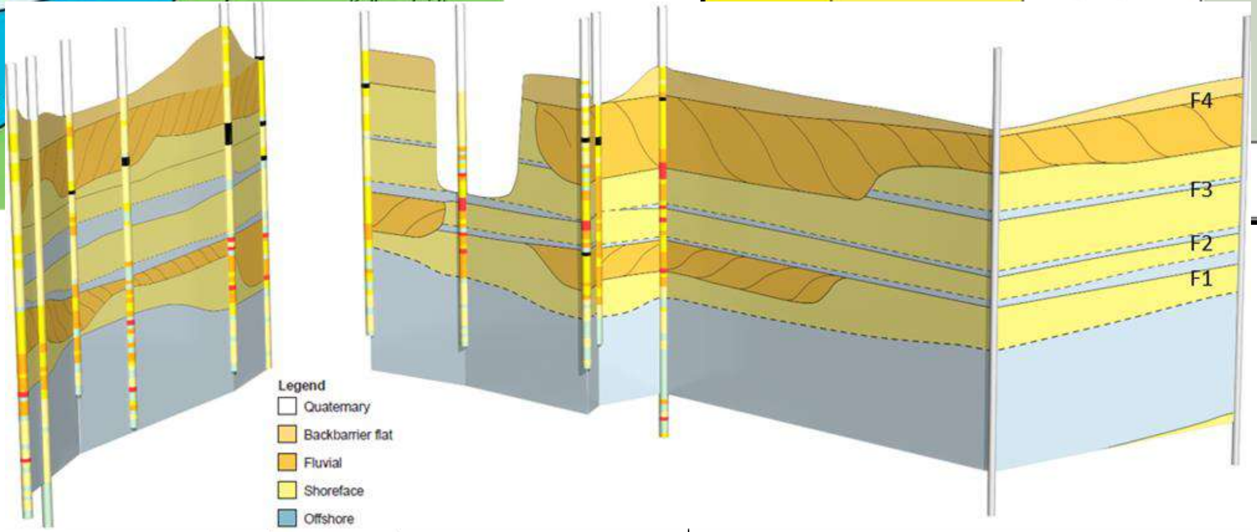
GRINDSTED


489000
490000
491000
492000
493000
494000
495000
496000
497000
498000
499000
500000
501000
6185000
6184000
6183000
6182000
6181000
6180000
6179000
6178000
6177000
6176000
6175000
6174000
6173000
6172000
Hejnsvig
1000 m

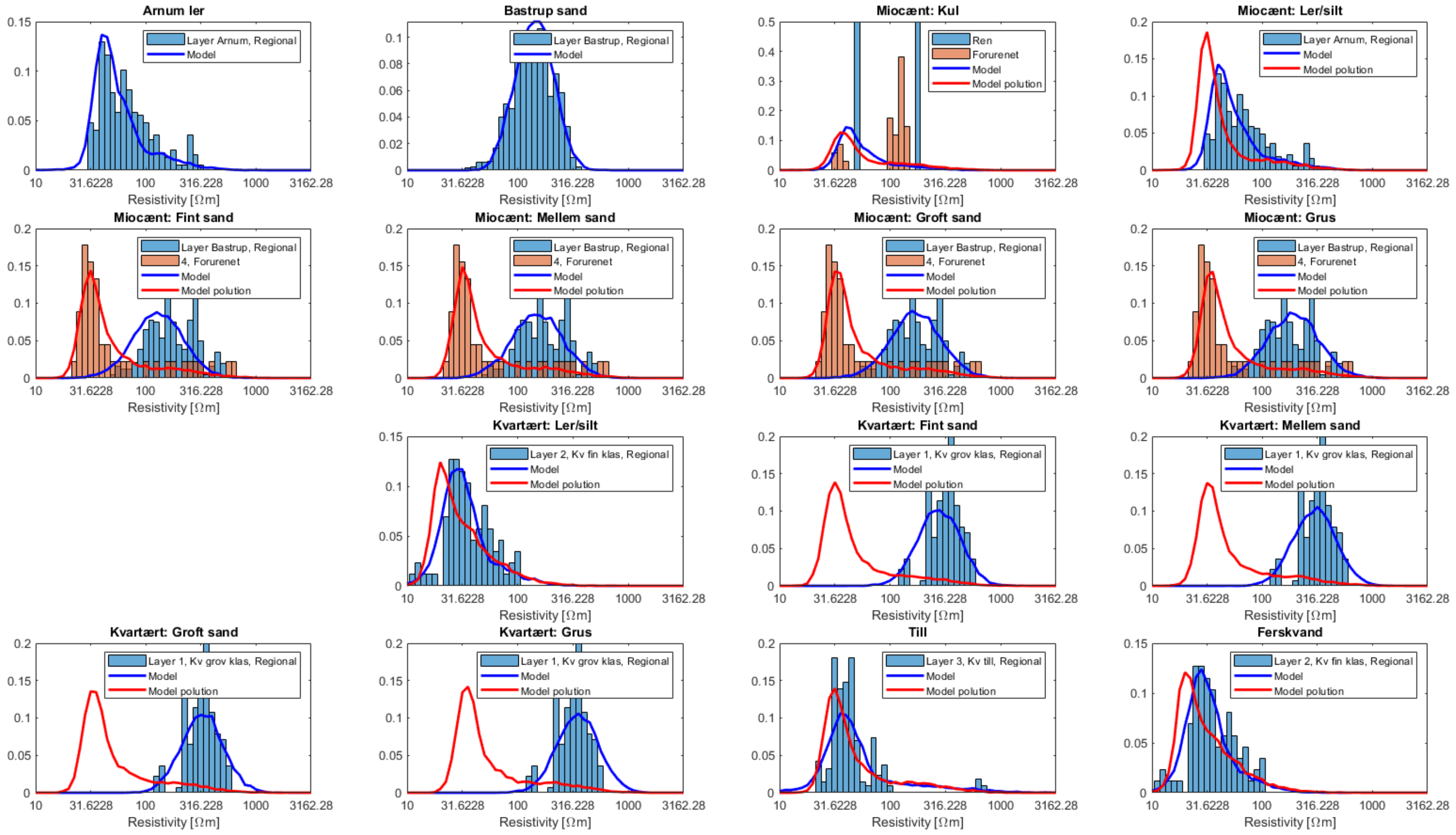
Geological prior information



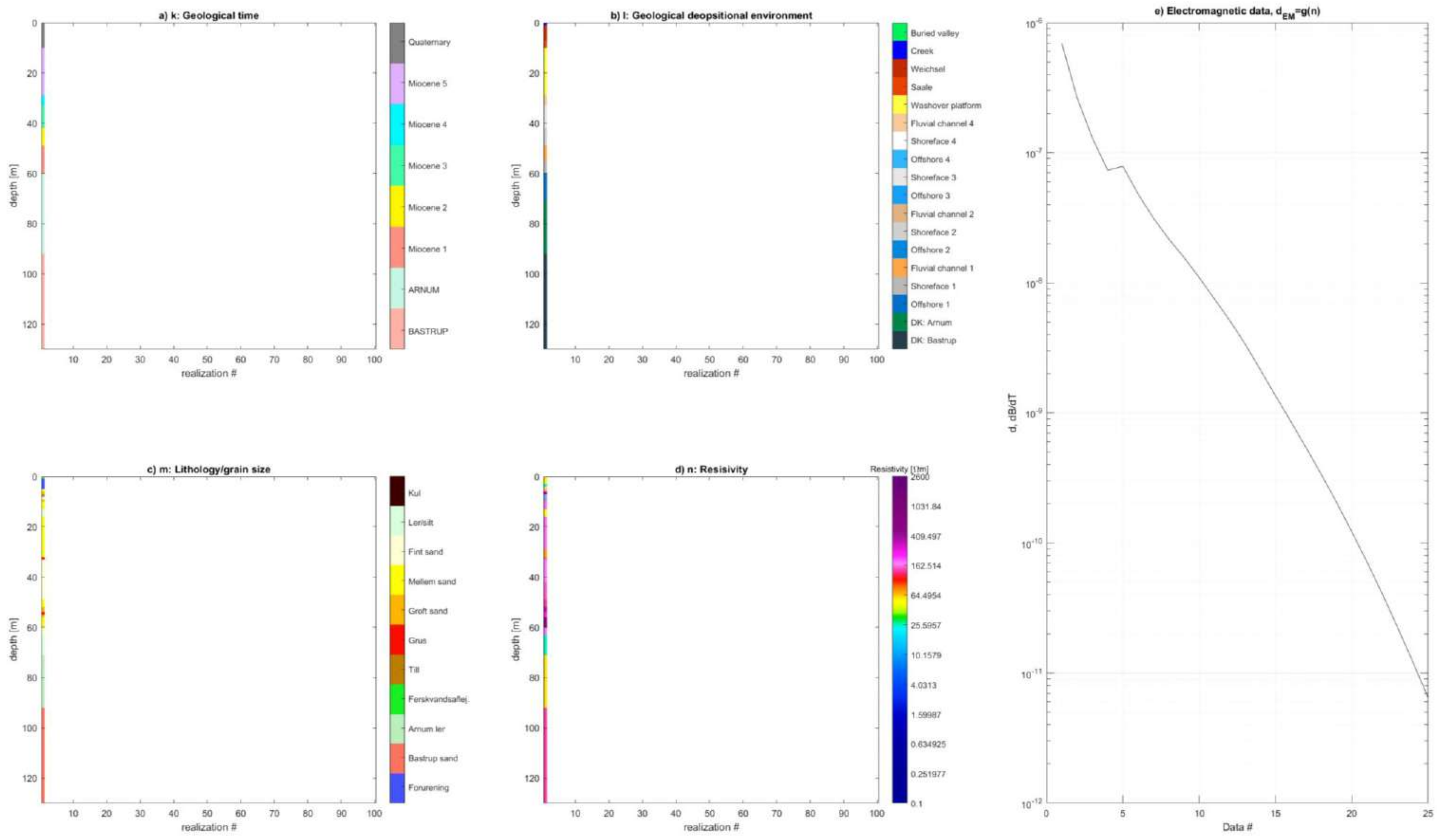




Grainsize -> resistivity modeling
f(resistivity|grainsize)

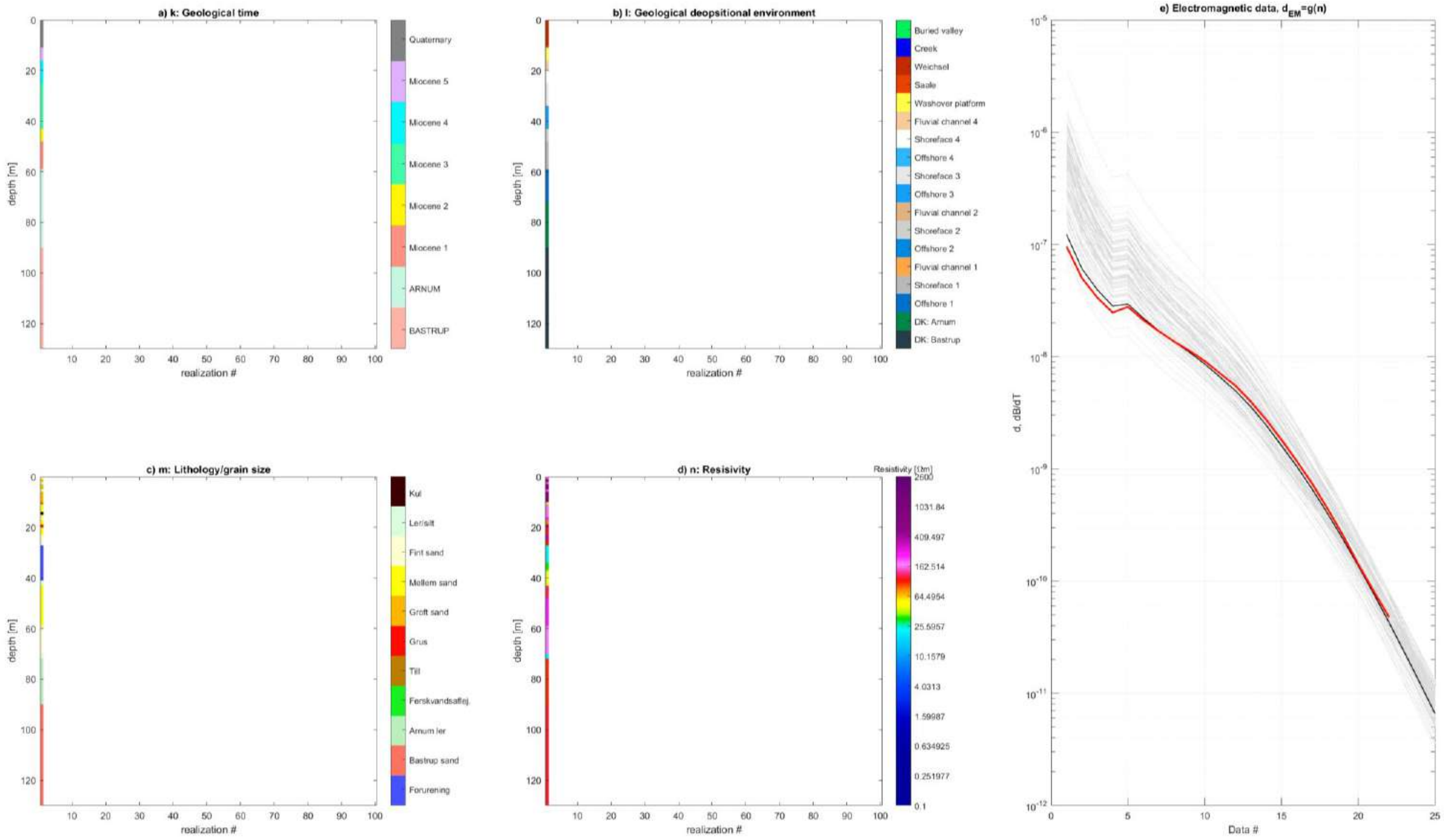


PRIOR

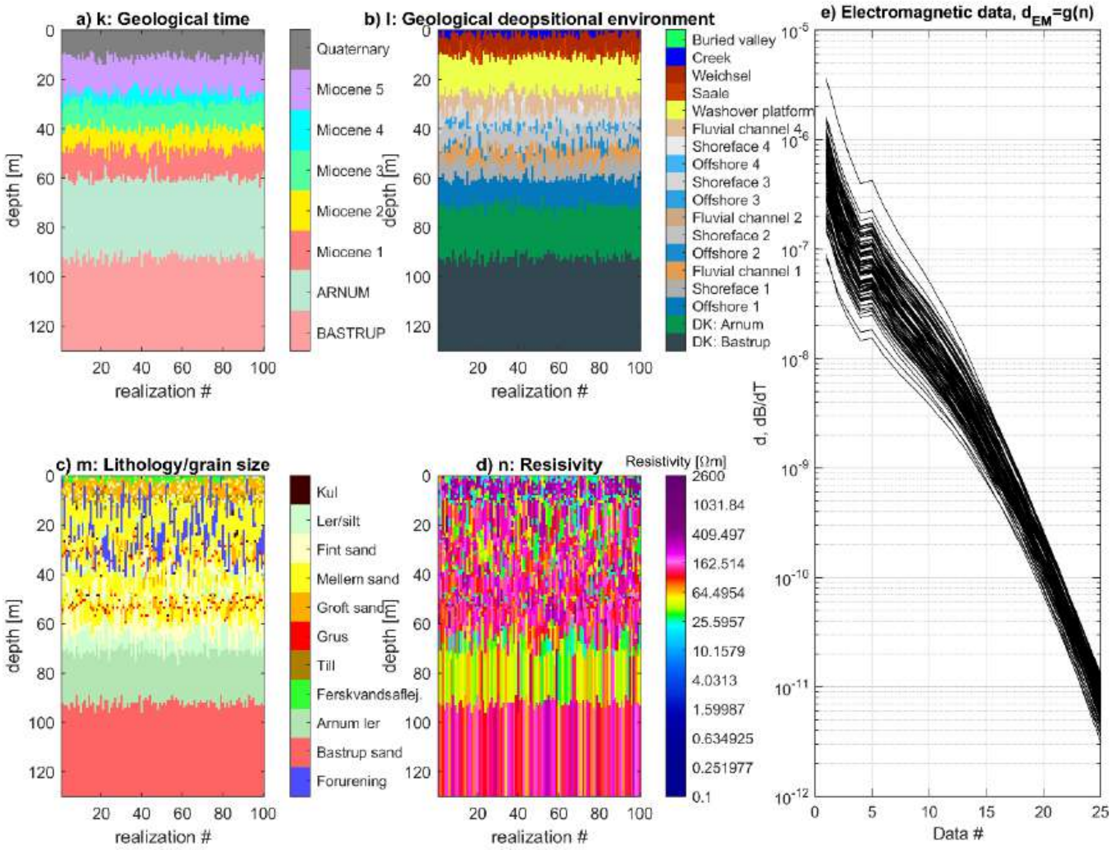


a) k: Geological time
Quaternary
Miocene 5
Miocene 4
Miocene 3
Miocene 2
Miocene 1
ARNUM
BASTRUP
b) l: Geological deopsitional environment
Buried valley
Creek
Weichsel
Saale
Washover platform
Fluvial channel 4
Shoreface 4
Offshore 4
Shoreface 3
Offshore 3
Fluvial channel 2
Shoreface 2
Offshore 2
Fluvial channel 1
Shoreface 1
Offshore 1
DK: Arnum
DK: Bastrup
c) m: Lithology/grain size
Kul
Ler/silt
Fint sand
Mellem sand
Groft sand
Grus
Till
Ferskvandsaflej.
Arnum ler
Bastrup sand
Forurening
d) n: Resisivity
e) Electromagnetic data, $d_{EM}=g(n)$
depth [m]
realization #
d, dB/dT
Data #

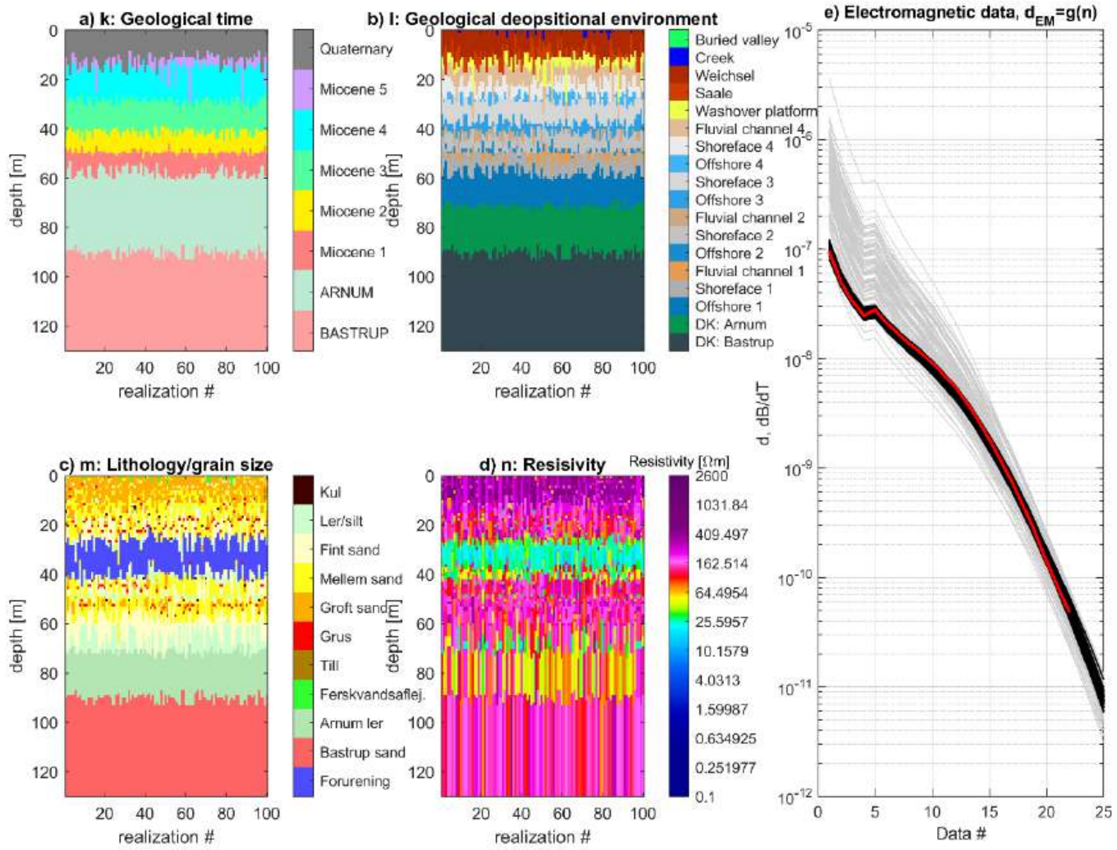
POST



a) k: Geological time
depth [m]
realization #
Quaternary
Miocene 5
Miocene 4
Miocene 3
Miocene 2
Miocene 1
ARNUM
BASTRUP
b) l: Geological deopsitional environment
Buried valley
Creek
Weichsel
Saale
Washover platform
Fluvial channel 4
Shoreface 4
Offshore 4
Shoreface 3
Offshore 3
Fluvial channel 2
Shoreface 2
Offshore 2
Fluvial channel 1
Shoreface 1
Offshore 1
DK: Arnum
DK: Bastrup
c) m: Lithology/grain size
Kul
Ler/silt
Fint sand
Mellem sand
Groft sand
Grus
Till
Ferskvandsaflej.
Arnum ler
Bastrup sand
Forurening
d) n: Resisivity
2600
1031.84
409.497
162.514
64.4954
25.5957
10.1579
4.0313
1.59987
0.634925
0.251977
0.1
e) Electromagnetic data, $d_{EM}=g(n)$
d, dB/dT
Data #

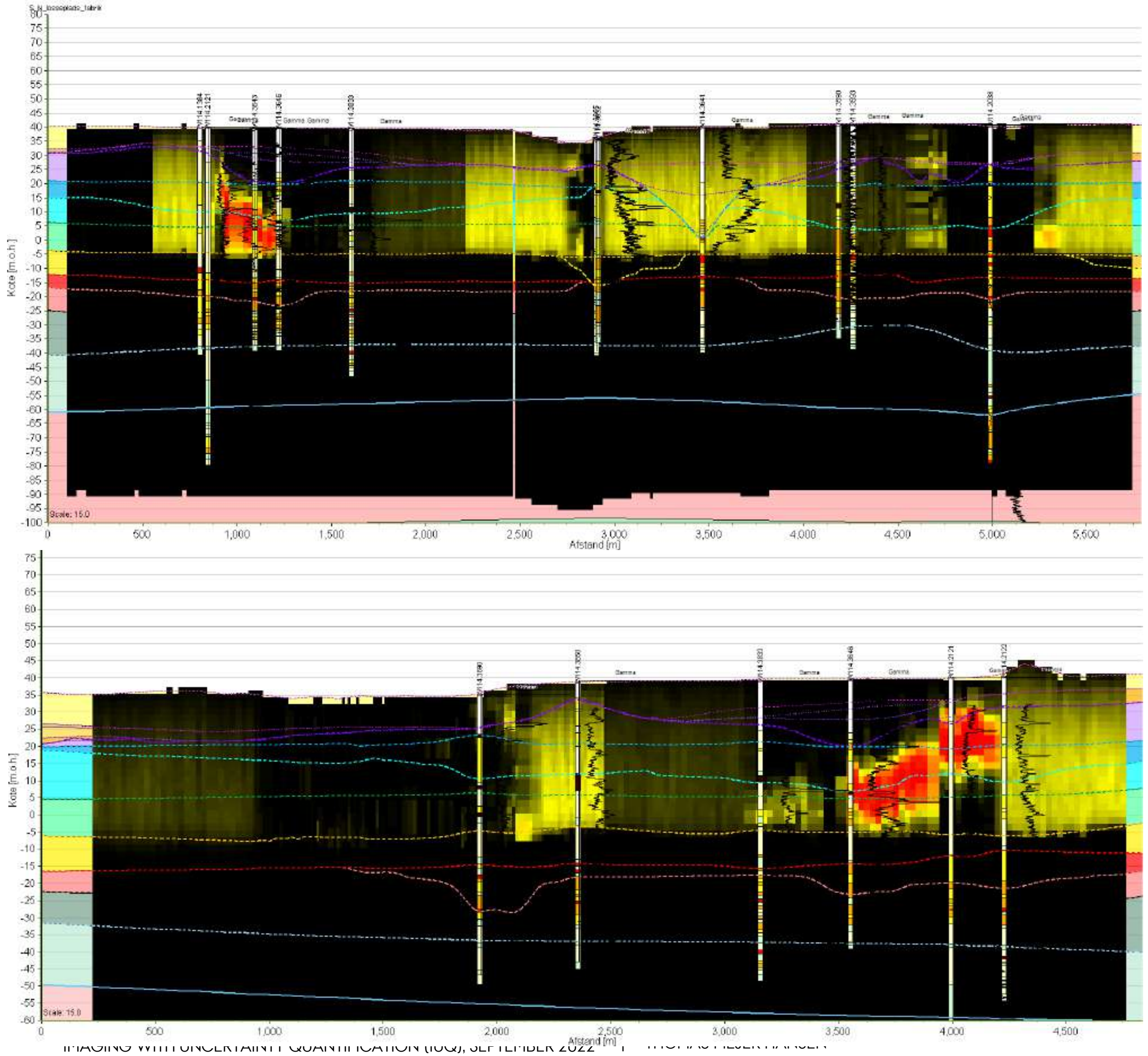


a) k: Geological time
Quaternary
Miocene 5
Miocene 4
Miocene 3
Miocene 2
Miocene 1
ARNUM
BASTRUP
b) l: Geological deopsitional environment
Buried valley
Creek
Weichsel
Saale
Washover platform
Fluvial channel 4
Shoreface 4
Offshore 4
Shoreface 3
Offshore 3
Fluvial channel 2
Shoreface 2
Offshore 2
Fluvial channel 1
Shoreface 1
Offshore 1
DK: Arnum
DK: Bastrup
c) m: Lithology/grain size
Kul
Ler/silt
Fint sand
Mellem sand
Groft sand
Grus
Till
Ferskvandsaflej.
Arnum ler
Bastrup sand
Forurening
d) n: Resisivity
Resistivity [Ωm]
2600
1031.84
409.497
162.514
64.4954
25.5957
10.1579
4.0313
1.59987
0.634925
0.251977
0.1
depth [m]
realization #
e) Electromagnetic data, $d_{EM}=g(n)$
d, dB/dT
Data #



a) k: Geological time
Quaternary
Miocene 5
Miocene 4
Miocene 3
Miocene 2
Miocene 1
ARNUM
BASTRUP
b) l: Geological deopsitional environment
Buried valley
Creek
Weichsel
Saale
Washover platform
Fluvial channel 4
Shoreface 4
Offshore 4
Shoreface 3
Offshore 3
Fluvial channel 2
Shoreface 2
Offshore 2
Fluvial channel 1
Shoreface 1
Offshore 1
DK: Arnum
DK: Bastrup
c) m: Lithology/grain size
Kul
Ler/silt
Fint sand
Mellem sand
Groft sand
Grus
Till
Ferskvandsaflej.
Arnum ler
Bastrup sand
Forurening
d) n: Resisivity
Resistivity [Ωm]
2600
1031.84
409.497
162.514
64.4954
25.5957
10.1579
4.0313
1.59987
0.634925
0.251977
0.1
depth [m]
realization #
e) Electromagnetic data, $d_{EM}=g(n)$
d, dB/dT
Data #

Probability of contaminated plume

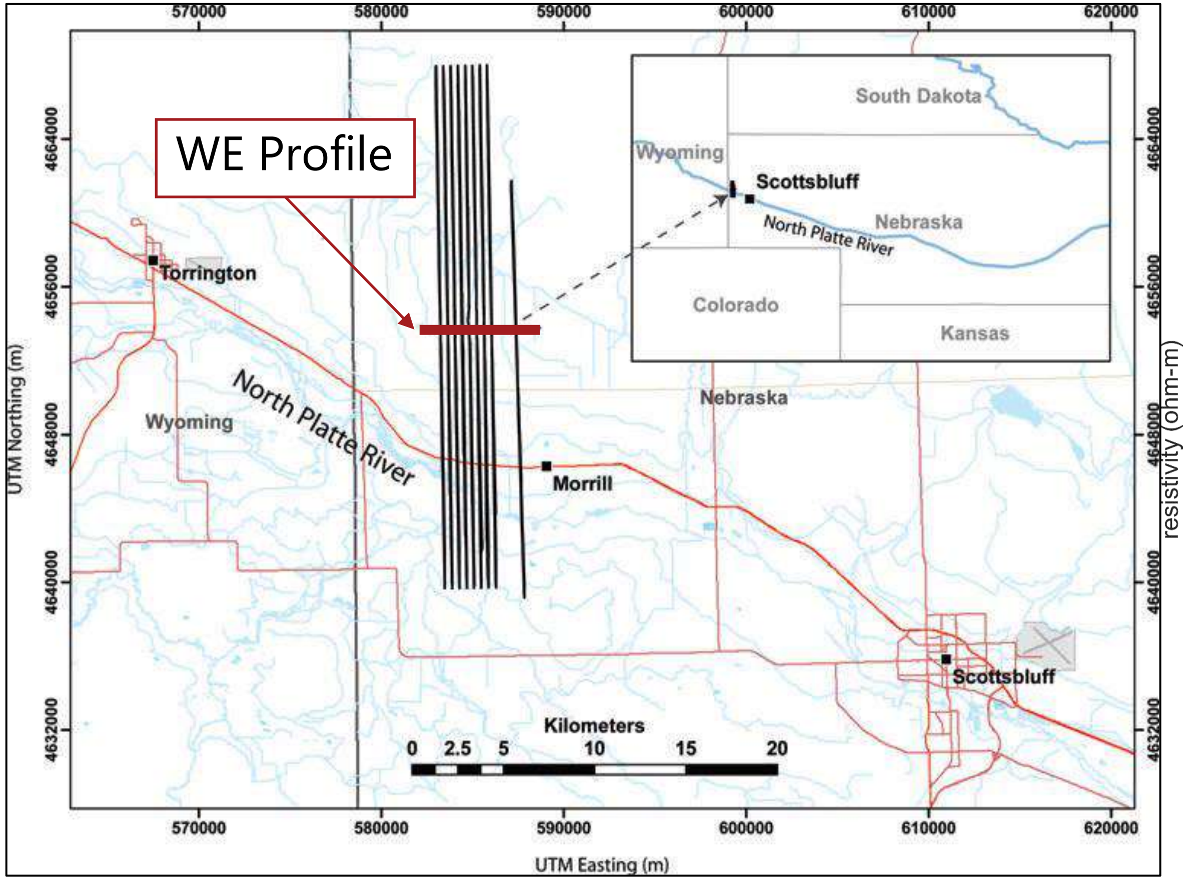


CASE: AIRBORNE EM

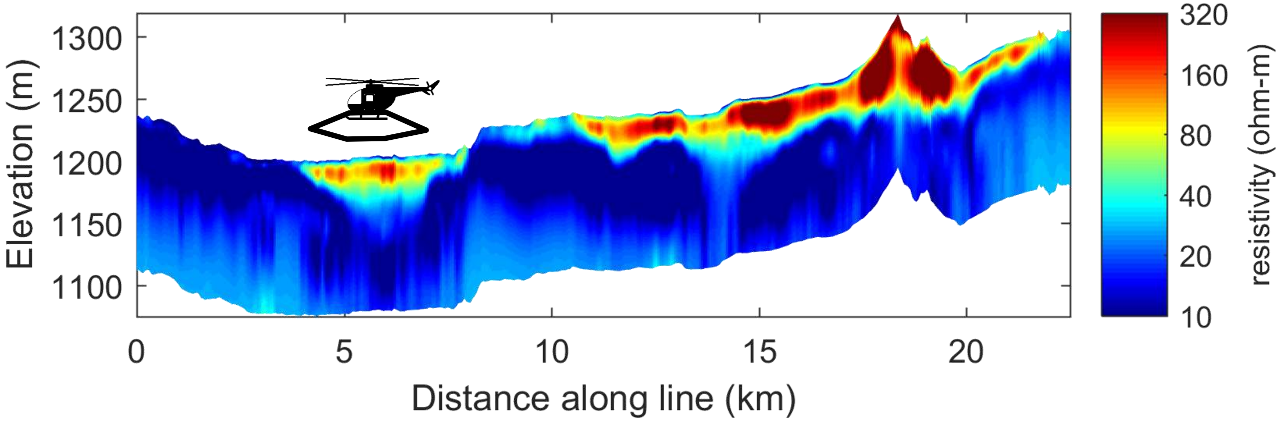
Sampling the posterior distribution using the extended Metropolis algorithm

THOMAS MEJER HANSEN AND BURKE MINSLEY

Airborne FD-EM data from Morill, Nebraska



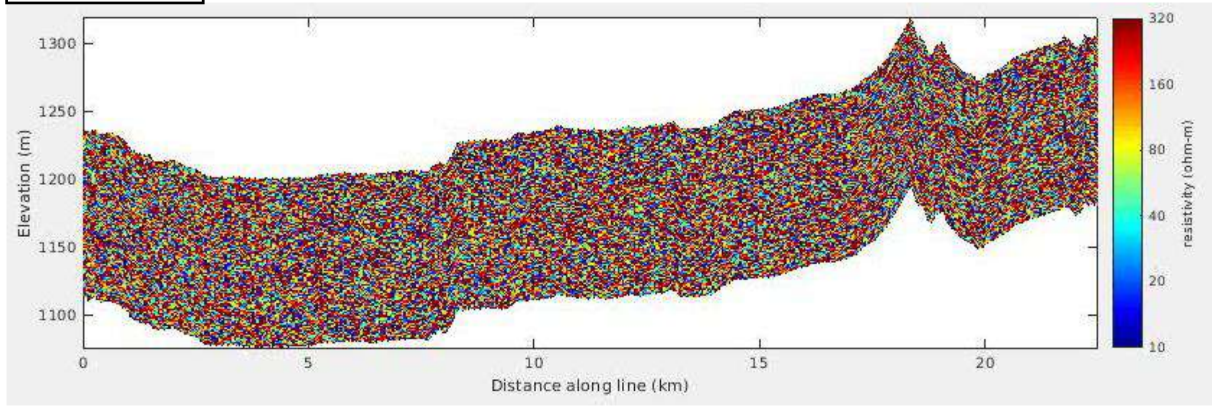
EM1DFM solution



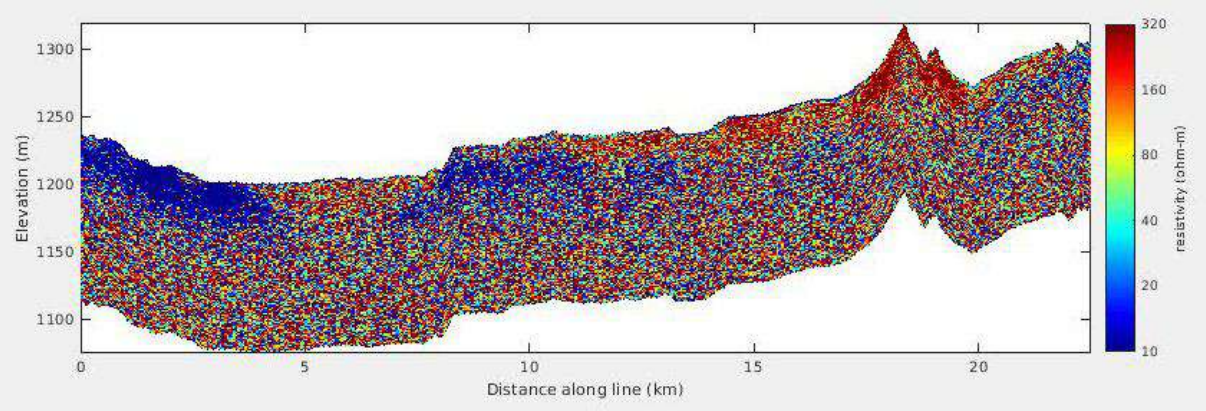
(Minsley, 2011)

1D inversion along profile – 'no prior' $f_{GEO}(\mathbf{m})$=constant
Purely data driven

Prior

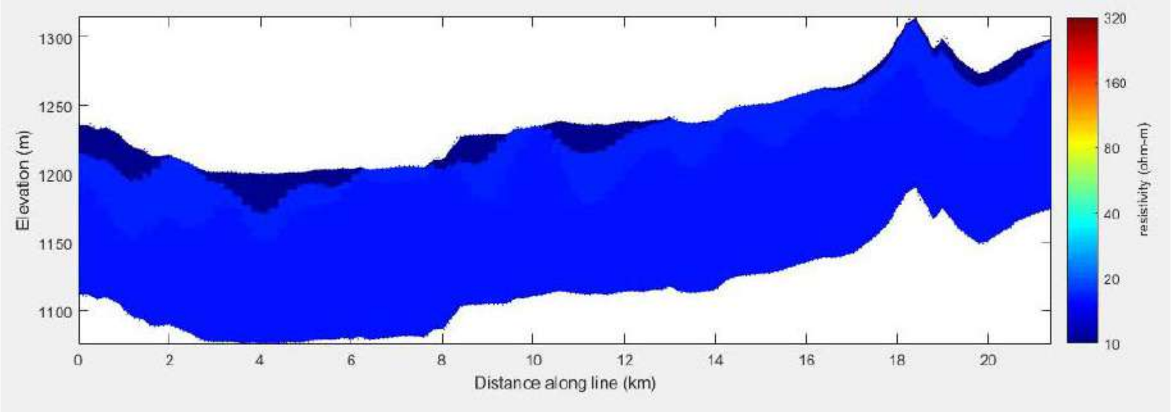


Posterior

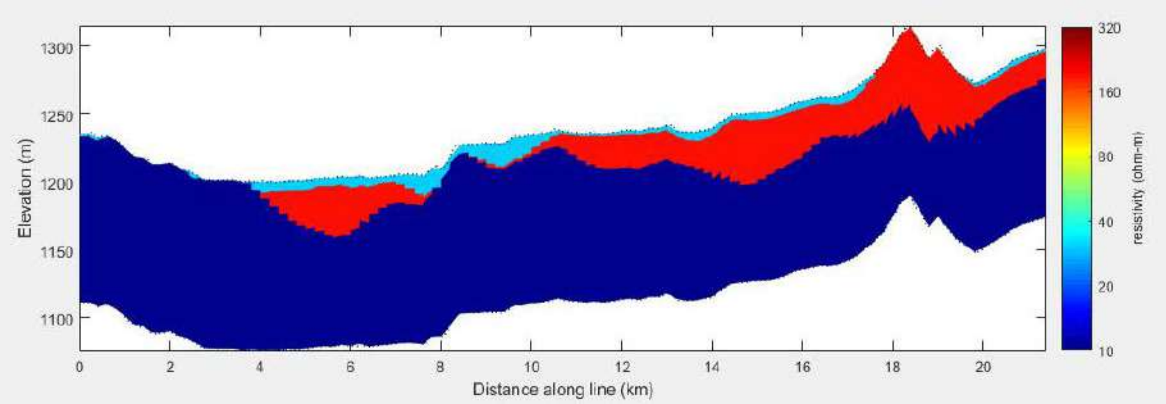


1D inversion along profile – $f_{GEO}(\mathbf{m})$: 3 layer. Too informative prior

Prior

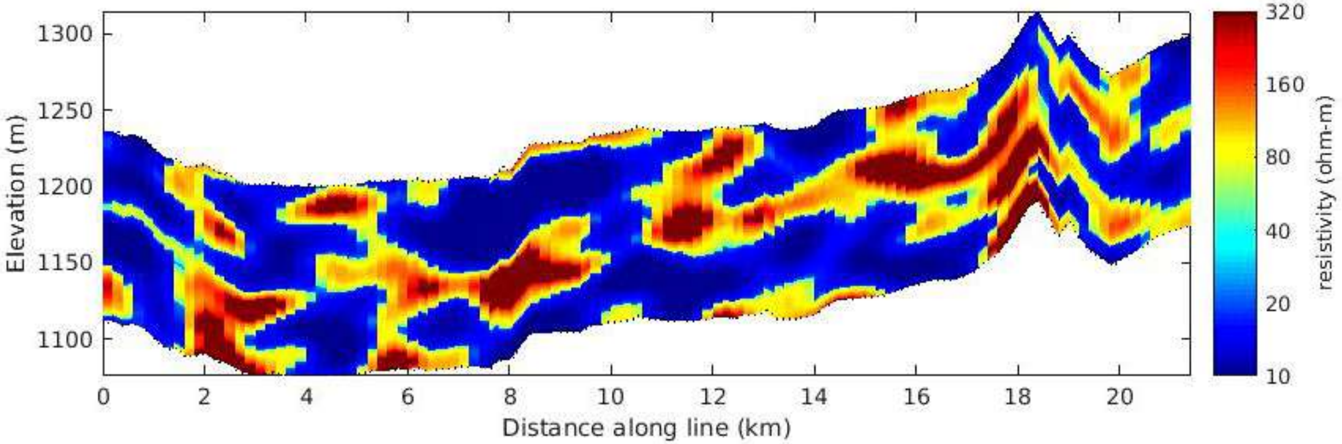


Posterior

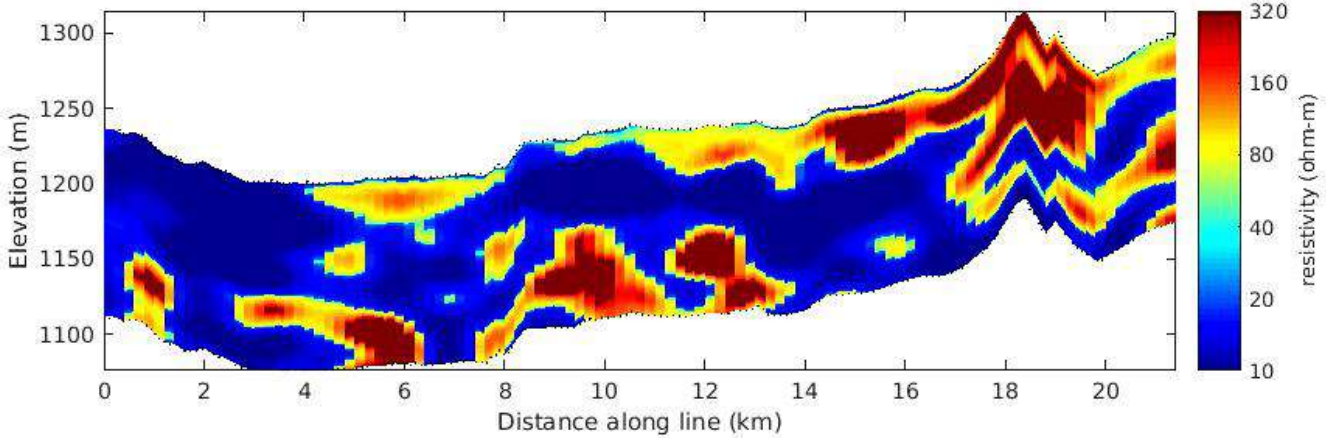


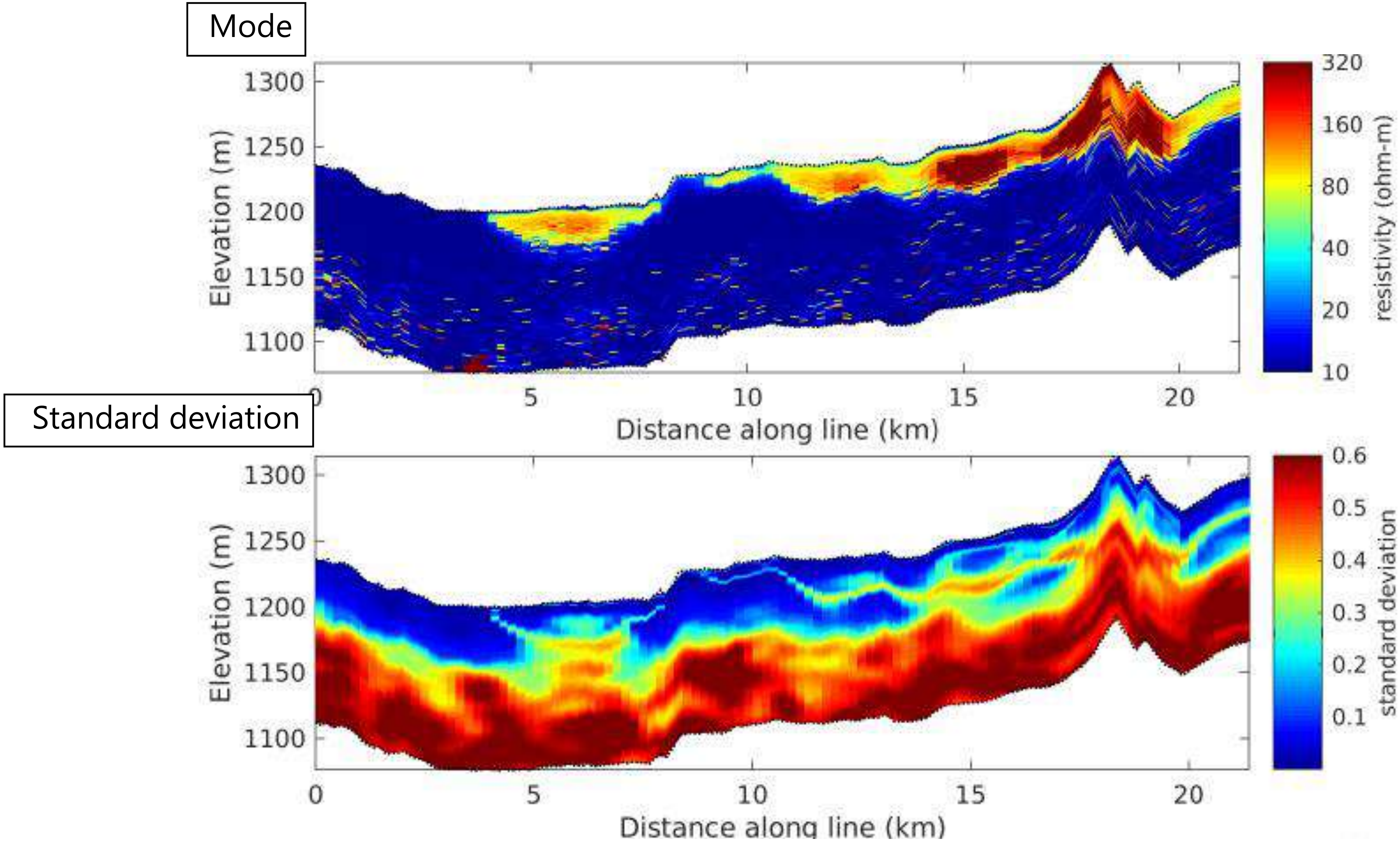
1D inversion along profile –$f_{GEO}(\mathbf{m})$: realistic Trimodal prior inferred from well logs

Prior



Posterior




Mode
1300
1250
1200
1150
1100
Elevation (m)
0
5
10
15
20
Distance along line (km)
320
160
80
40
20
10
resistivity (ohm-m)
Standard deviation
1300
1250
1200
1150
1100
Elevation (m)
0
5
10
15
20
Distance along line (km)
0.6
0.5
0.4
0.3
0.2
0.1
standard deviation

CASE: AIRBORNE EM

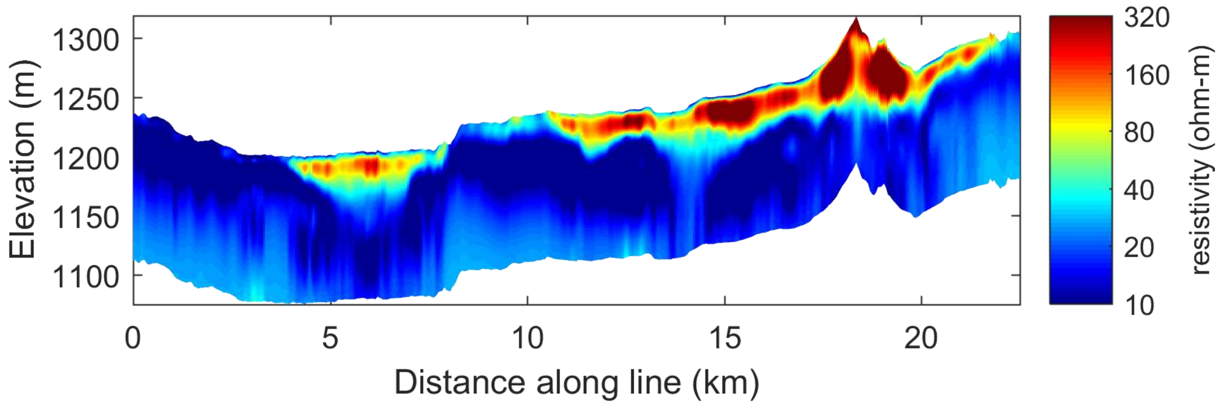
Sampling the posterior distribution using the extended rejection sampler

LOCALIZED PROBABILISTIC INVERSION

$$[N^*, M^*, D^*]$$

$\rho(\mathbf{m})$

$$\left[\begin{array}{c} \mathbf{n}^{*1},\ \mathbf{m}^{*1}, \mathbf{d}^{*1}, \\ \mathbf{n}^{*2},\ \mathbf{m}^{*2}, \mathbf{d}^{*2}, \\ \ldots \\ \mathbf{n}^{*nr},\ \mathbf{m}^{*nr}, \mathbf{d}^{*nr}, \end{array}\right]$$



Localized inversed problem:
When the prior and the forward problem are the same, for different data sets!

THE EXTENDED REJECTION SAMPLER WITH LOOKUP TABLES

The goal: Sample from $\sigma(\mathbf{m}) = k\,\rho(\mathbf{m})\,L(\mathbf{m})$

1. Propose a model from $\rho(\mathbf{m}) \rightarrow \mathbf{m}^*$
2. Accept $\mathbf{m}^*$ as a realization from the posterior with probability

   $P_{acc} = L(\mathbf{d}_{obs} - g(\mathbf{m}^*))/\max(L)$

To apply the extended rejection sampler one must be able to

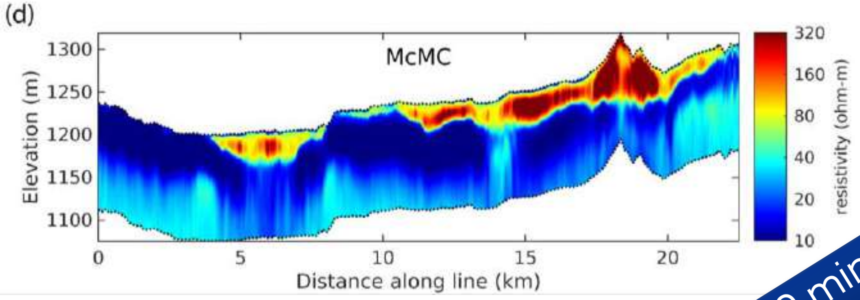
1) Sample the prior $\sigma(\mathbf{m})$
2) Evaluate the likelihood $L(\mathbf{m})$ for any model, (solve the forward problem, evaluate the noise)

Lookup table: [M*, D*] → Compute once, apply for any $\mathbf{d}_{obs}$

Hansen (2021)

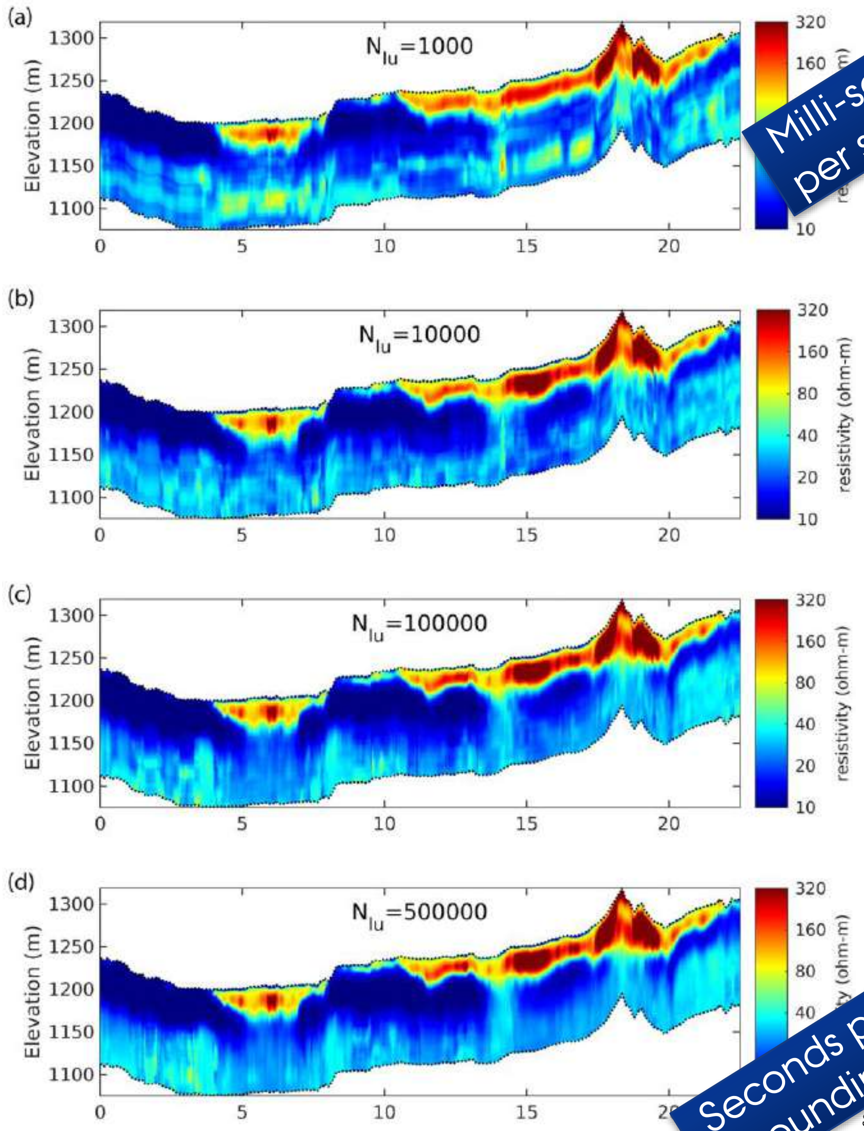
POSTERIOR MEAN

Extended Metropolis algorithm



10 minutes per sounding

Extended rejection sampler with lookup tables



Milli-seconds per sounding

Seconds per sounding

CASE: AIRBORNE EM

Computing statistics of the posterior distribution directly using machine learning,

without sammpling the posterior!

THOMAS MEJER HANSEN AND CHRIS FINLAY

$$\sigma(\mathbf{m}) = k\,\rho(\mathbf{m})\,L(\mathbf{d}_{obs} - g(\mathbf{m}))$$

The forward problem

$\mathbf{d} = g(\mathbf{m})$

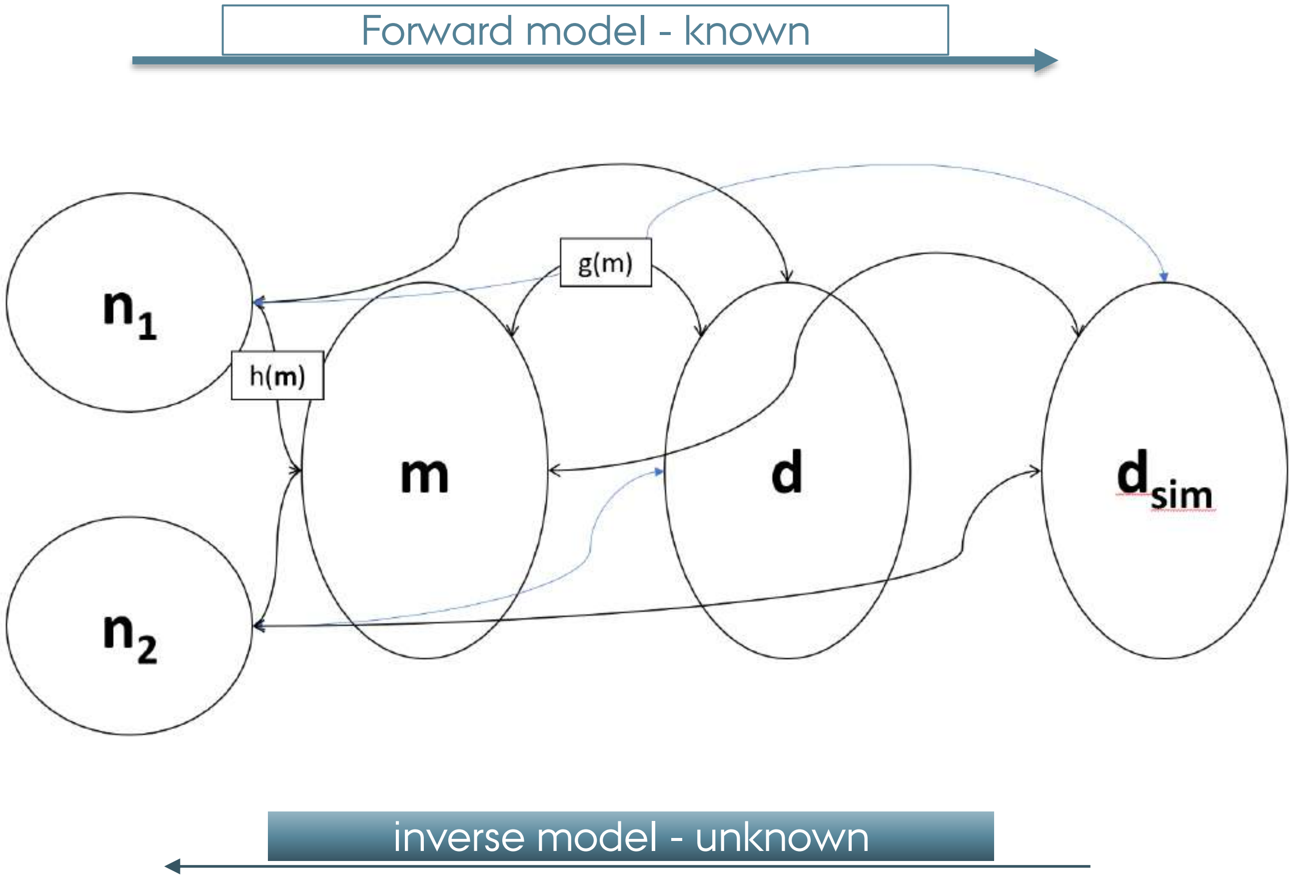
The forward problem -> simulation

$\mathbf{d}_{sim} = \mathbf{d} + e(\mathbf{m}) = g(\mathbf{m}) + e(\mathbf{m})$

Sometimes one is not interested in **m**, but in a property **n** related to **m**

$\mathbf{n} = h(\mathbf{m})$

$\rightarrow \sigma(\mathbf{n})$



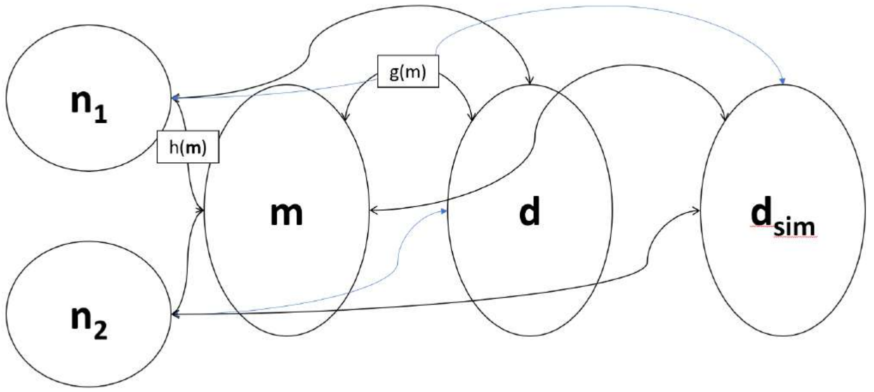
PROBABILISTIC INVERSION USING MACHINE LEARNING

$$[\mathbf{N}^*, \mathbf{M}^*, \mathbf{D}^*, \mathbf{D}\mathrm{sim}^*]$$

$$\mathbf{d}_{sim} = \mathbf{d} + e(\mathbf{m}) = g(\mathbf{m}) + e(\mathbf{m})$$

$$\mathbf{n} = h(\mathbf{m})$$

$$\begin{array}{c} \\ \begin{array}{cccc} & \rho(\mathbf{m}) & & \end{array} \\ \left[\begin{array}{c} \mathbf{n}^{*1},\ \mathbf{m}^{*1}, \mathbf{d}^{*1}, \mathbf{d}_{sim}{}^{*1} \\ \mathbf{n}^{*2},\ \mathbf{m}^{*2}, \mathbf{d}^{*2}, \mathbf{d}_{sim}{}^{*2} \\ \ldots \\ \mathbf{n}^{*nr},\ \mathbf{m}^{*nr}, \mathbf{d}^{*nr}, \mathbf{d}_{sim}{}^{*nr} \end{array}\right] \end{array}$$

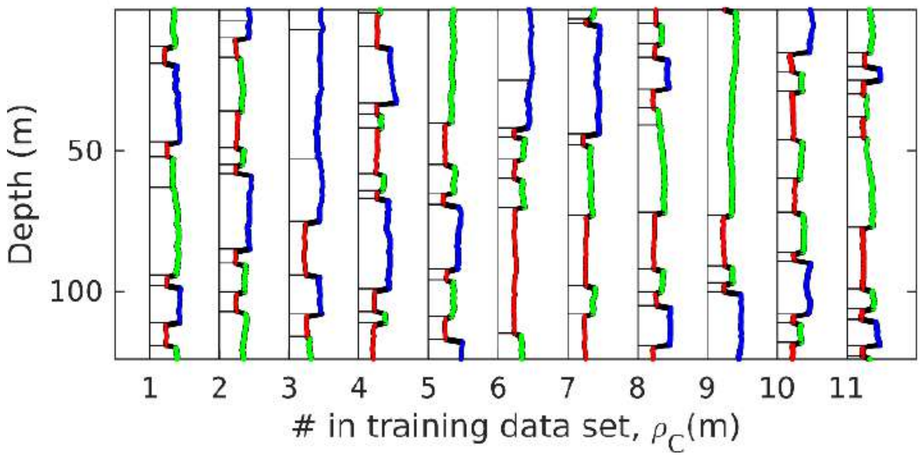


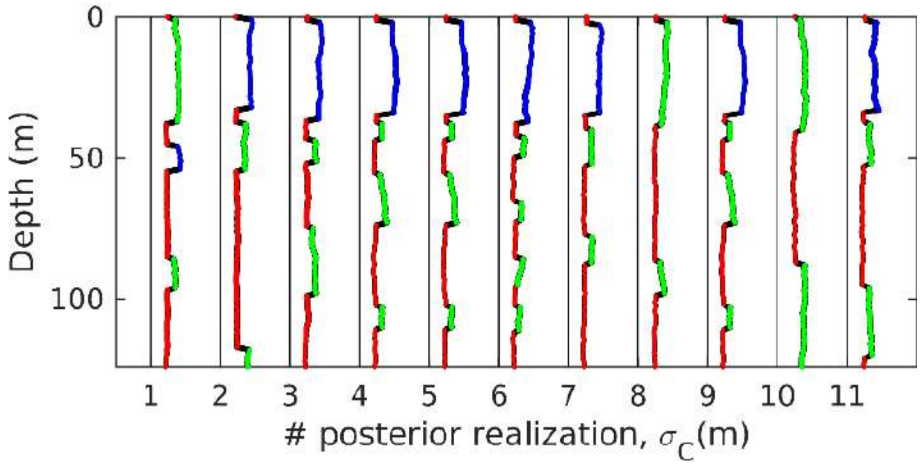
Hansen (in rev)

PRIOR ρ(**m**)

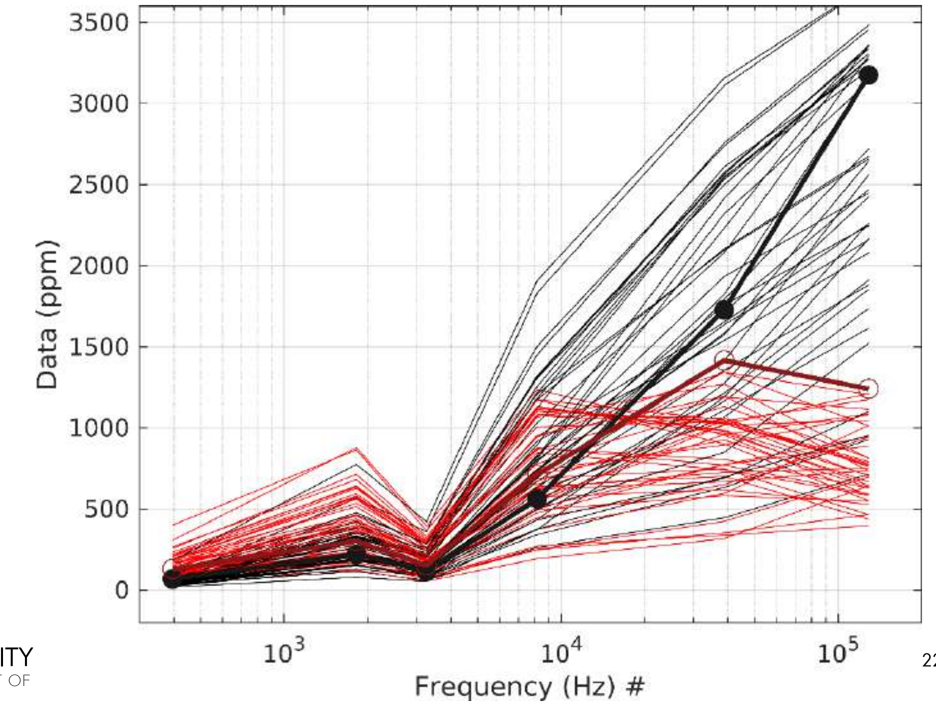
POSTERIOR σ(**m**)

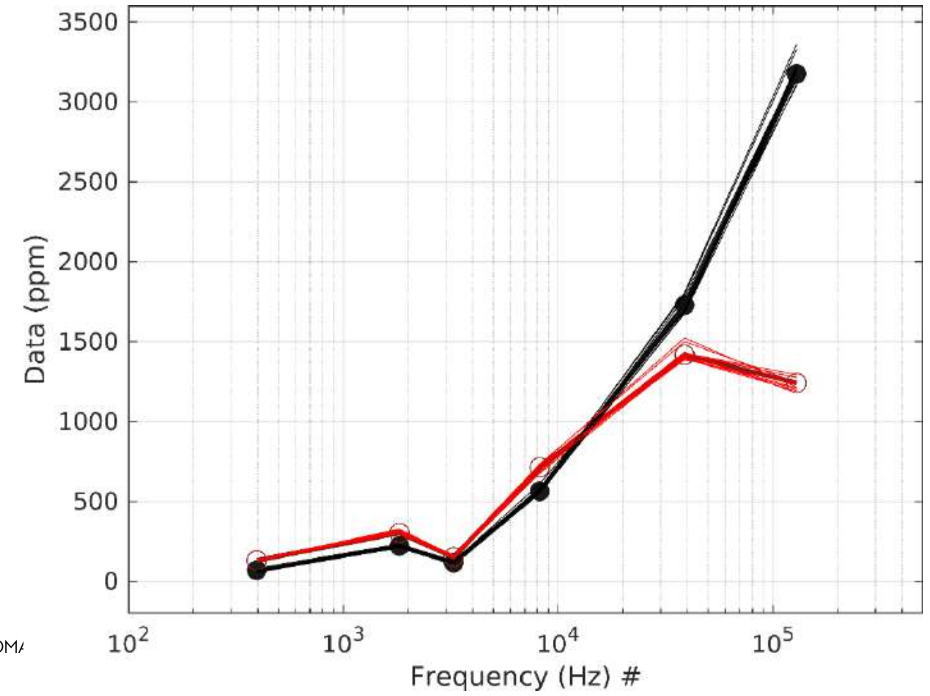
MODEL



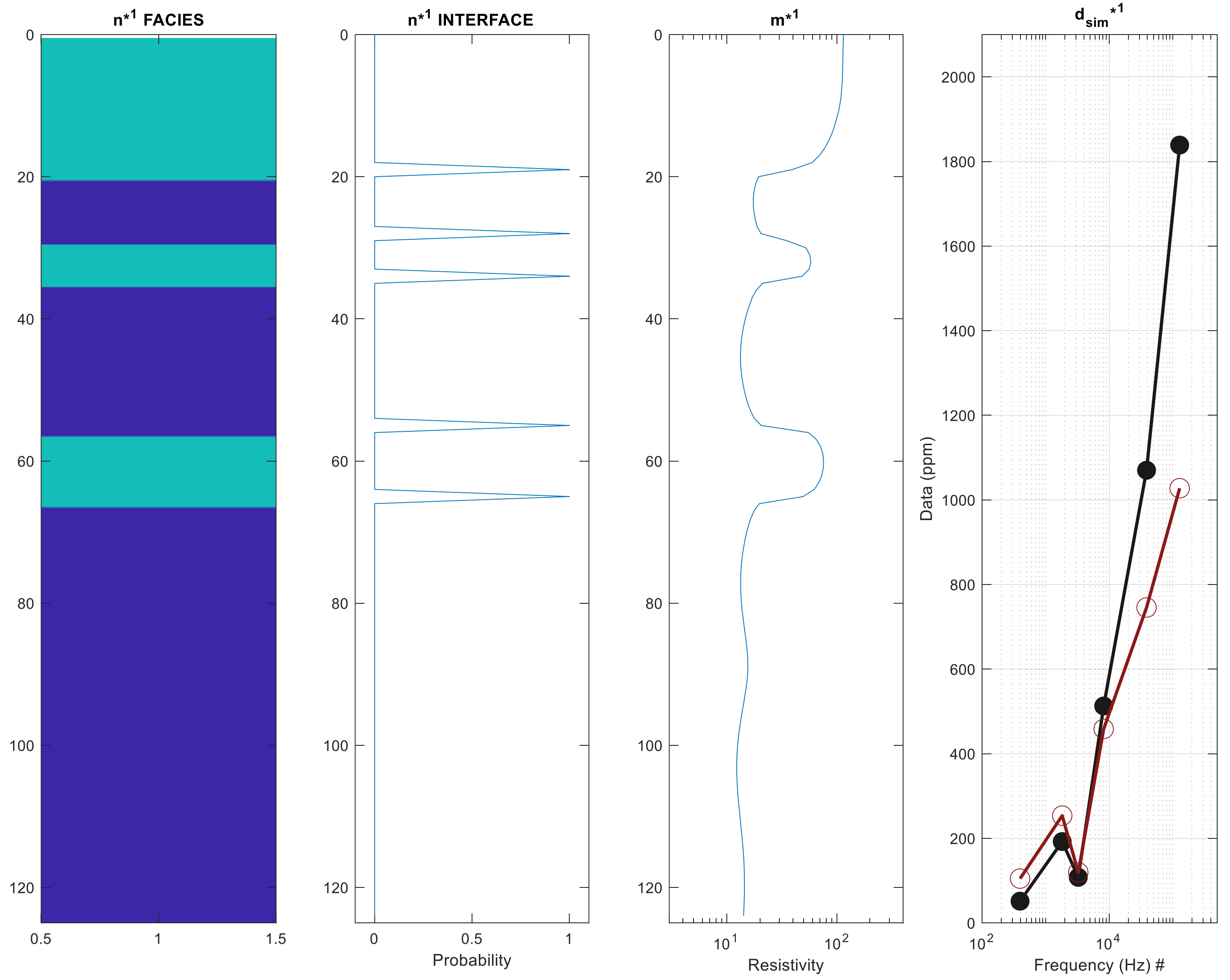


DATA





CONSTRUCTING A TRAINING DATA SET



CONSTRUCTING A TRAINING DA

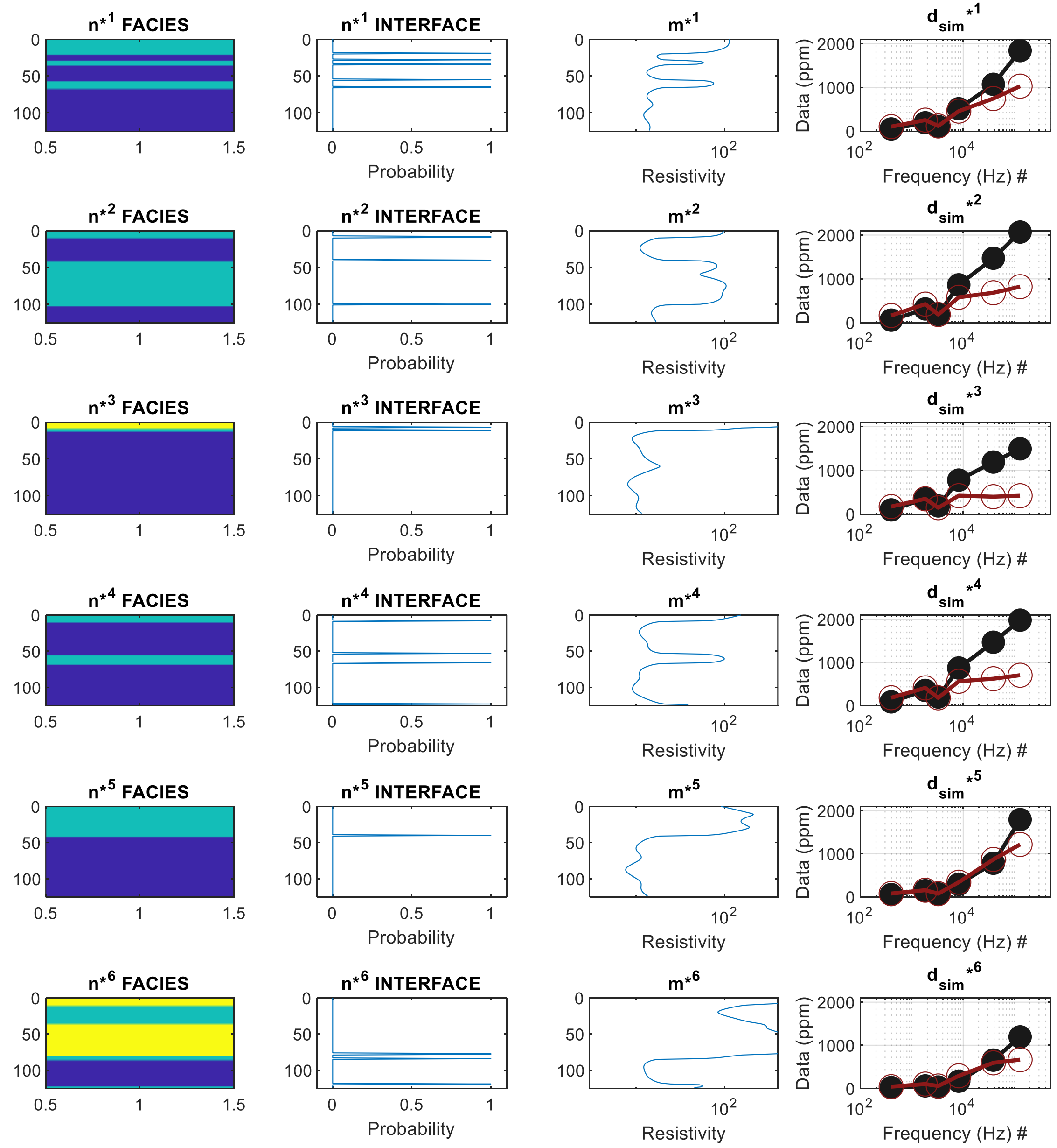
The training data can, in principle, be arbitrarily large

Here the training data set consists of up to 5000000 sets of **n**, **m** and $\mathbf{d}_{sim}$

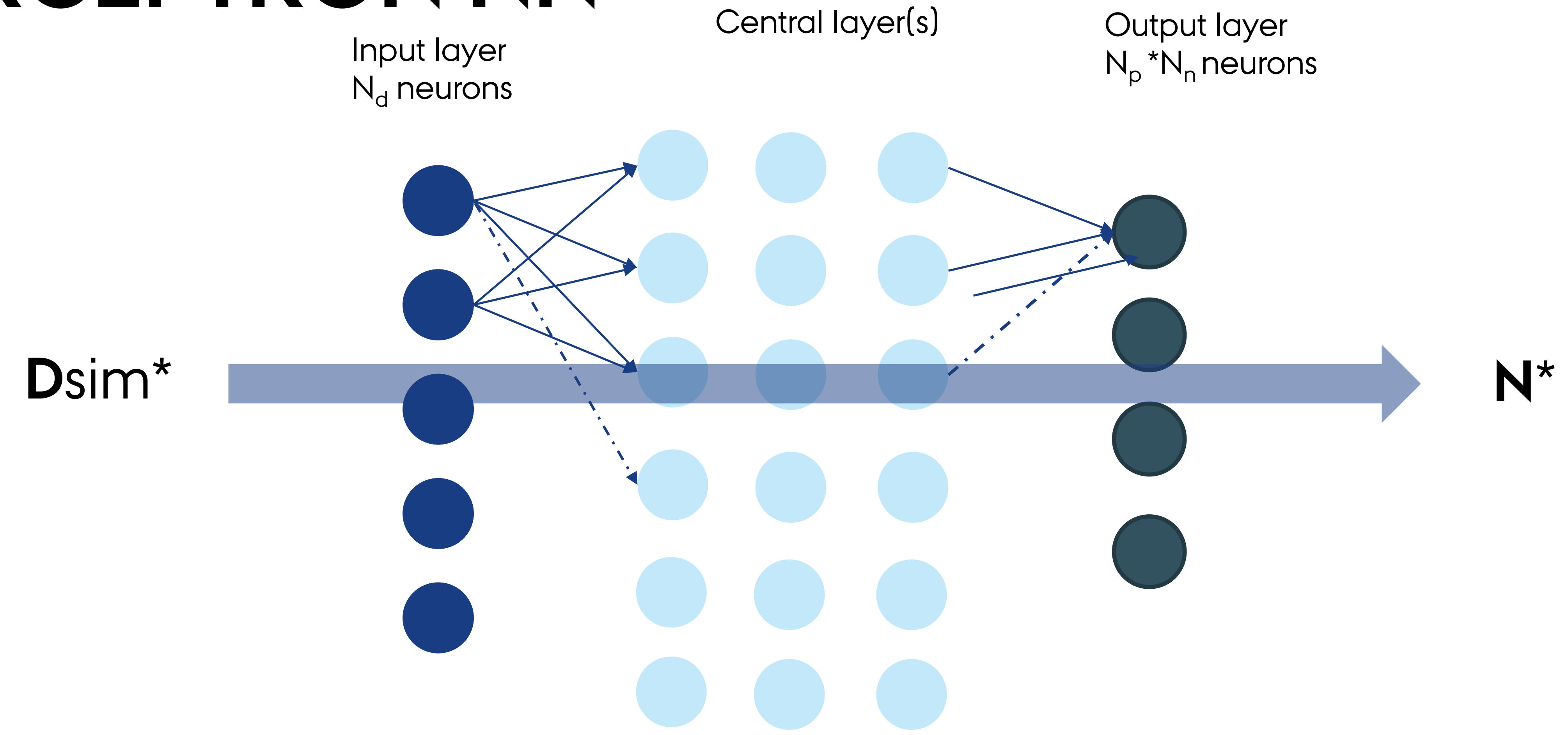
[**N***, **M***, **D**sim*]

This training set express everything we know (as much as can be represented by a finite size sample) about

- The prior
- The forward
- The noise



DESIGNING A SIMPLE MULTILAYER PER~~C~~EPTRON NN



CHOOSING A LOSS FUNCTION TO MINIMIZE WHEN ADJUSTING THE FREE PARAMETERS OF THE NN

Regression, n is a continuous parameter

Linear activation function in output layer

The output of the NN can for example be the parameters of a multivariate normal distribution N(**m**', **C**m'):

$f(\mathbf{m}^*) = k \exp(-0.5\ (\mathbf{m}'-\mathbf{m}^*)\ \mathbf{C}_{m'}^{-1}\ (\mathbf{m}'-\mathbf{m}^*)^T\ )$

$\mathrm{LOSS} = -\log(f(\mathbf{m}^*)) = -0.5\ (\mathbf{m}'-\mathbf{m}^*)\ \mathbf{C}_{m'}^{-1}\ (\mathbf{m}'-\mathbf{m}^*)^T$

A loss function minimizing this, and many more, distributions can easily be implemented using TensorFlow probability.

https://www.tensorflow.org/probability

Classification, n is a discrete parameter

Softmax activation function in output layer

Loss function: Binary cross entropy

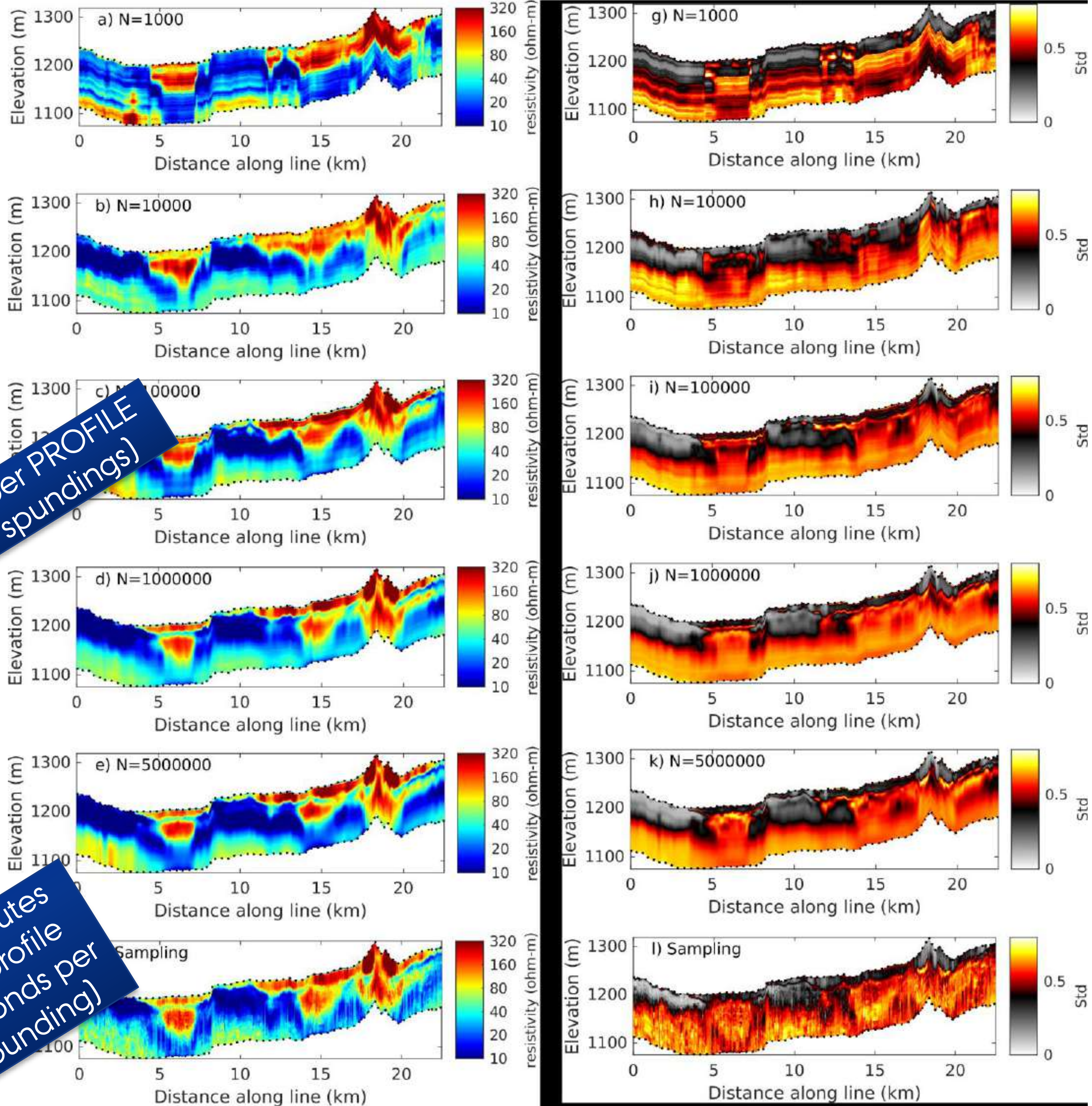
Ensures the output of the NN can be interpreted as P(**n**==n°)

Direct estimation of the pointwise mean and standard deviation for different size of training data set,

compared to using the extended rejection sampler.

Realistic sized EM surveys may contain +100.000 soundings, and they are getting larger.!

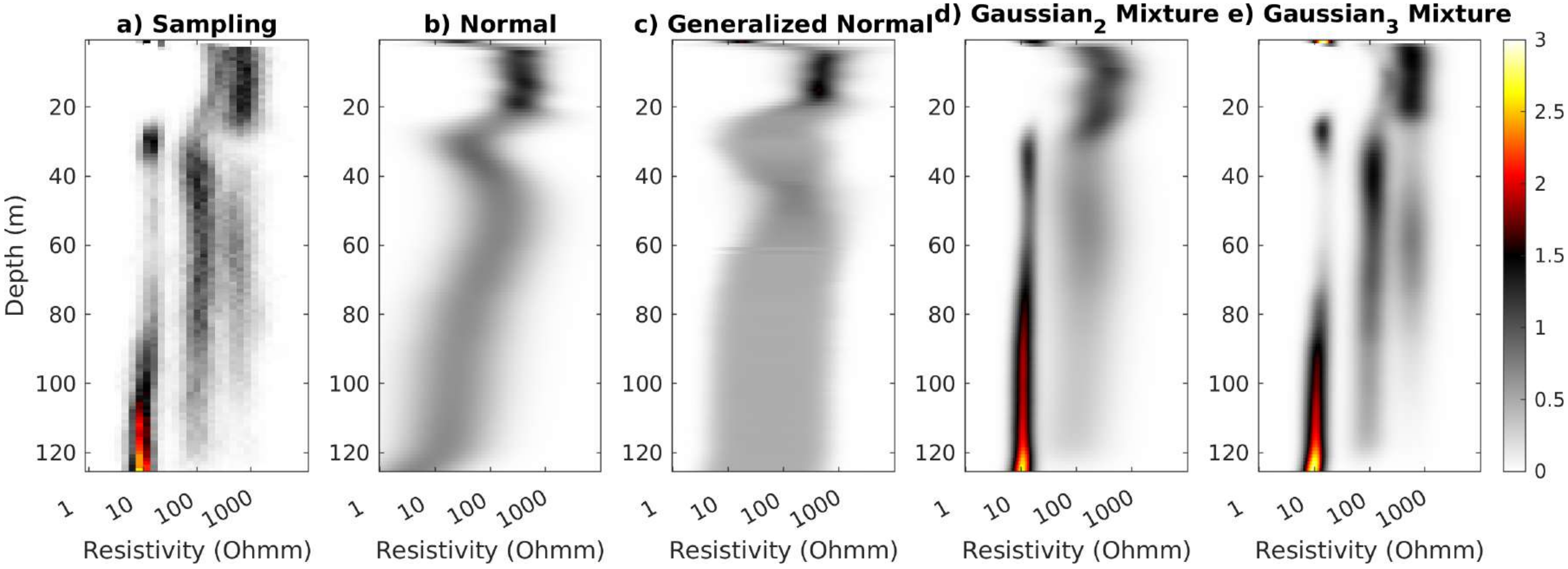
Training time: seconds (N=1000) to 15 minutes (N=5.000.000)



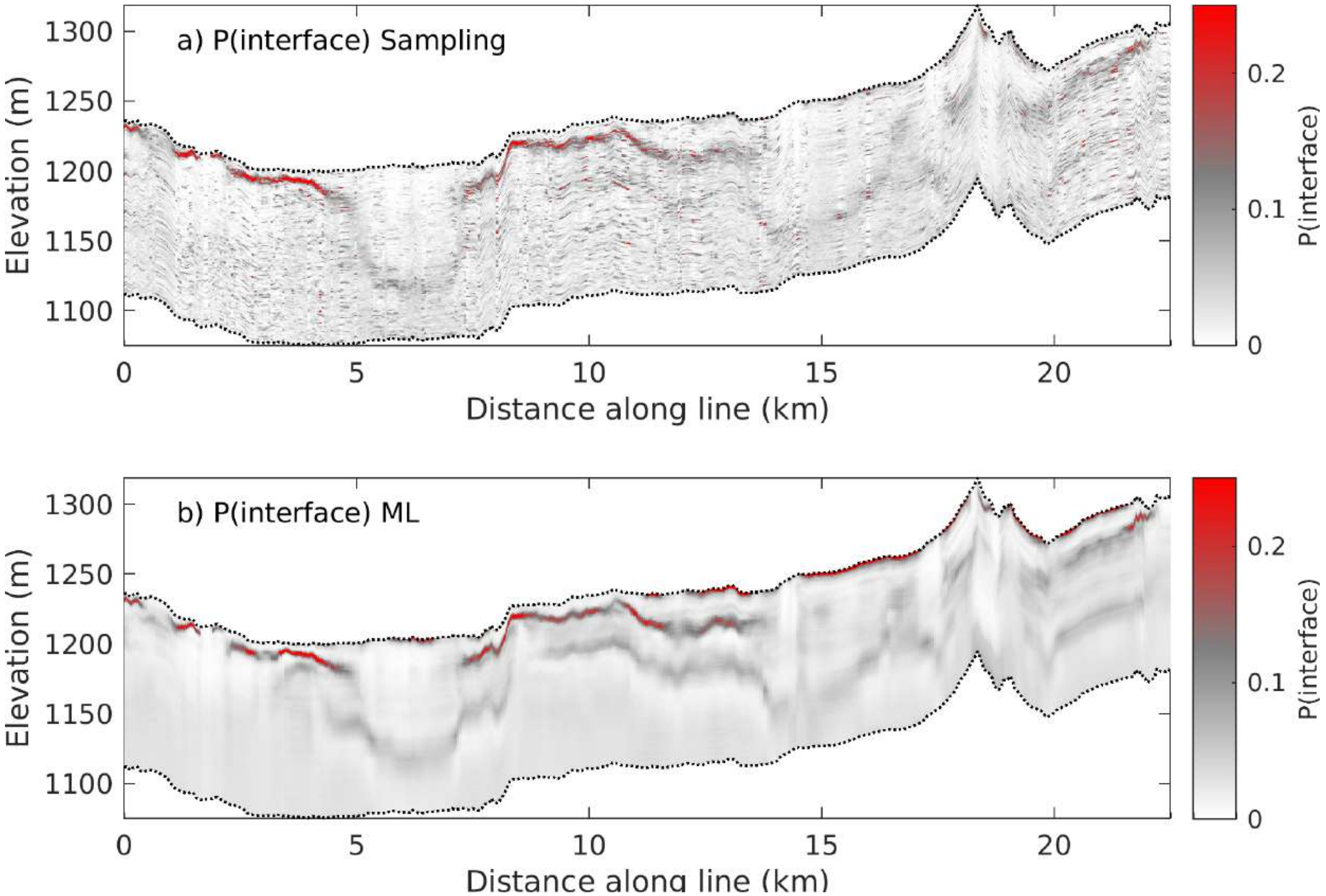
20 ms per PROFILE (451 spundings)

10 minutes per profile (seconds per sounding)

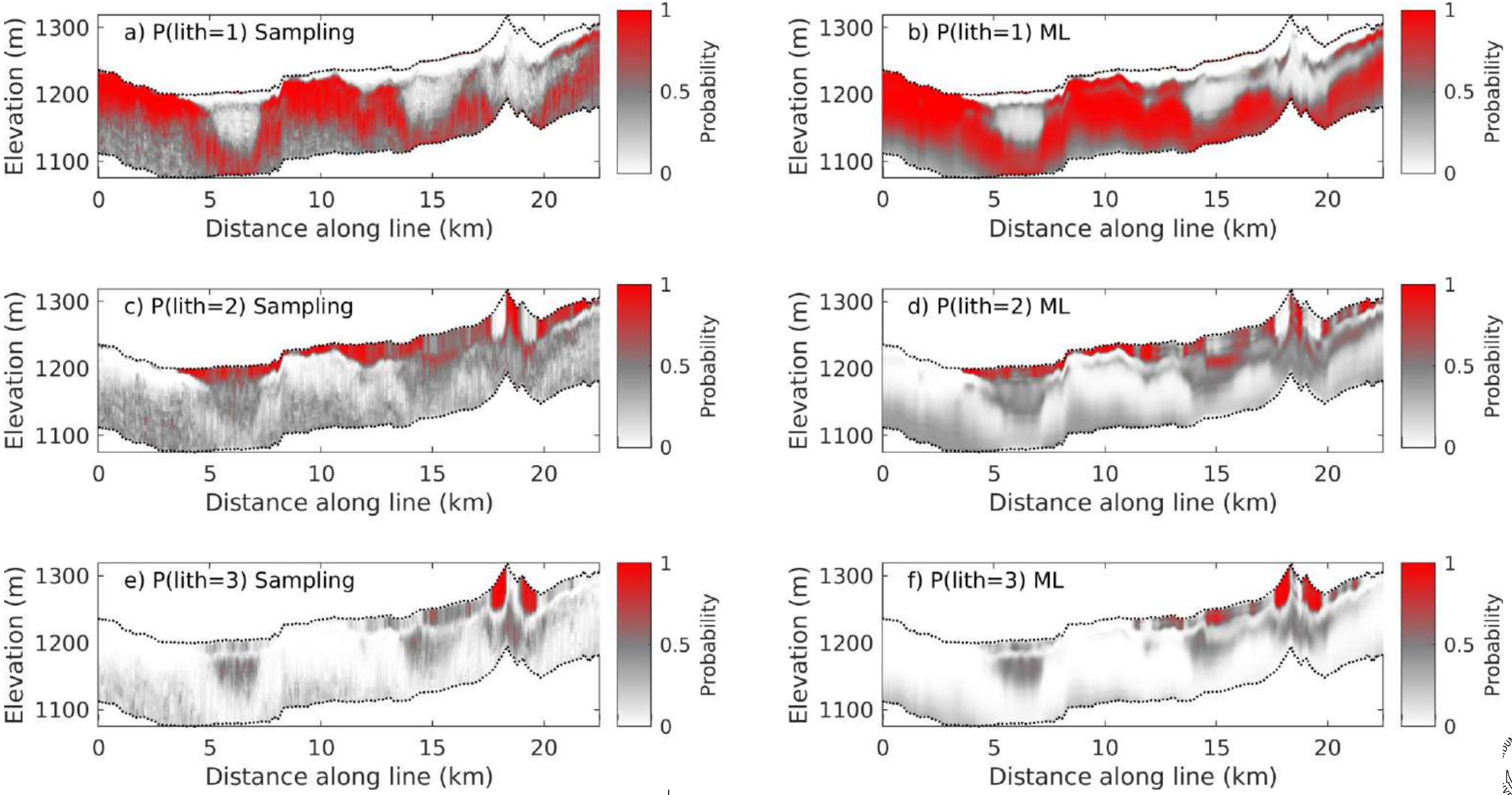
MARGINAL POSTERIOR STATISTICS



Posterior probability of a layer interface



Posterior probability of lithology



ALGORITHMS FOR SAMPLING

Hansen, Thomas M., and Burke J. Minsley. "Inversion of airborne EM data with an explicit choice of prior model." *Geophysical Journal International* 218.2 (2019): 1348-1366.

Minutes per EM sounding

Hansen, Thomas M. "Efficient probabilistic inversion using the rejection sampler—exemplified on airborne EM data." *Geophysical Journal International* 224.1 (2021): 543-557.

Seconds per EM sounding

Hansen, Thomas M., and Christopher C. Finlay. "Use of machine learning to estimate statistics of the posterior distribution in probabilistic inverse problems - an application to airborne EM data." in review/arxiv. https://doi.org/10.31223/X5JS56

Milli-seconds per EM sounding

CASE: PET IMAGE ANALYSIS

THOMAS MEJER HANSEN, KLAUS MOSEGAARD,
SØREN HOLM, FLEMMING ANDERSEN, BARBARA MALENE FISCHER, AND
ADAM ESPE HANSEN

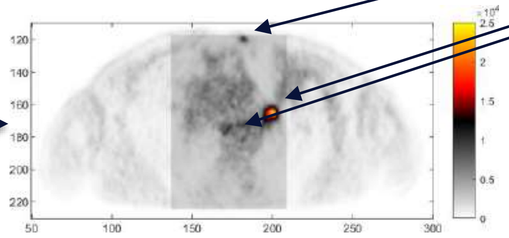
TRADITIONAL PET IMAGE ANALYSIS

Cancer ?

1. PET scanning

2. PET reconstruction

3. PET image analysis and decision making



DECISION

Small cancer cells are difficult to identify due to
- Resolution limitations
- Noise

Identification of large features with high activity is easy.

Early identification of small features with high activity concentration is difficult.

Early identification is key to successful treatment

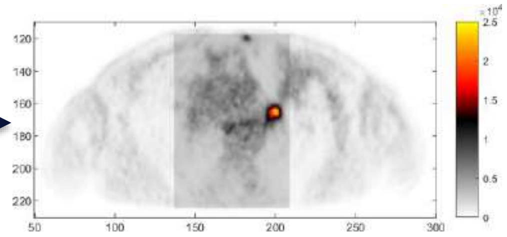
PROBABILISTIC PET IMAGE ANALYSIS

Cancer
99% probability??

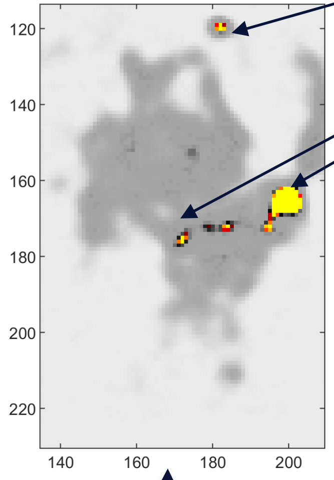
1. PET scanning

2. PET reconstruction

3. PET image analysis and decision making



m: activity concentration, tissue type, cancer



DECISION

g(**m**)

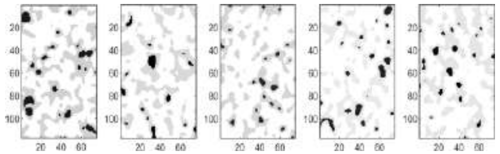
L(**m**)

ρ(**m**)

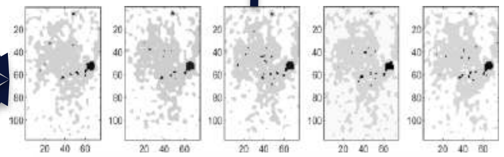
PHYSICS (smoothing)

NOISE model

Medical EXPERTISE



σ(**m**)



In **ProPET** all information available about the PET image is combined into one statistical model

One result is an “Enhanced image that shows the average intensity (high is indicative of cancer) in each cell
This aids the medical expert to take more informed decisions especially regarding the existence of small cancer cells

DEFINITIONS

$\mathbf{\Phi} = [\Phi_1, \Phi_2, \dots]$: The real activity concentration

$\mathbf{\Phi}_{PET}$: The reconstructed PET image

The forward model assumption (convolution)

$\mathbf{\Phi}_{PET} = G\,\mathbf{\Phi} + n(\mathbf{\Phi})$

$\rho(\mathbf{\Phi})$: The prior distribution of activity concentration

$\rho_{\mathbf{\Phi}}(\mathbf{\Phi}_{PET})$: The noise distribution of $\mathbf{\Phi}_{PET}$

$L(\mathbf{\Phi}) = \rho_{\mathbf{\Phi}}(G\mathbf{\Phi})$

PROBABILISTIC PET IMAGE ANALYSIS

A: Quantify information

- A1: $\rho(\Phi)$ Quantify prior information
- A2: $L(\Phi)=\rho_{\Phi}(G\Phi)$ Quantify noise and the forward model

B: Combine information

- Sample from $\sigma(\Phi) \propto k\, \rho(\Phi)\, L(\Phi)$

C: Clinical quantitative analysis

- Statistical analysis of the obtained sample from $\sigma(\Phi)$

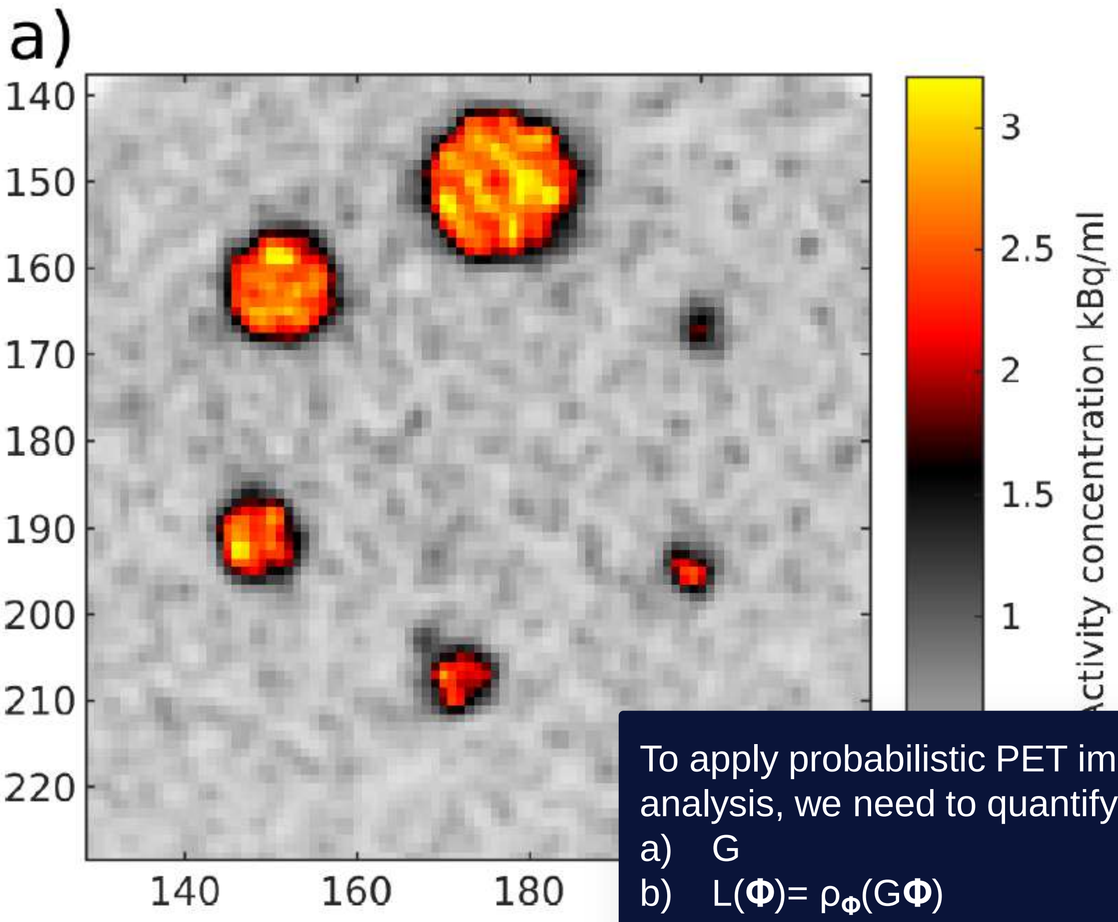
The *extended* Metropolis sampler:
We need to be able to
1) sample $\rho(\Phi)$, and
2) evaluate $L(\Phi)$.

$\rho(\Phi)$ can be represented by an algorithm!

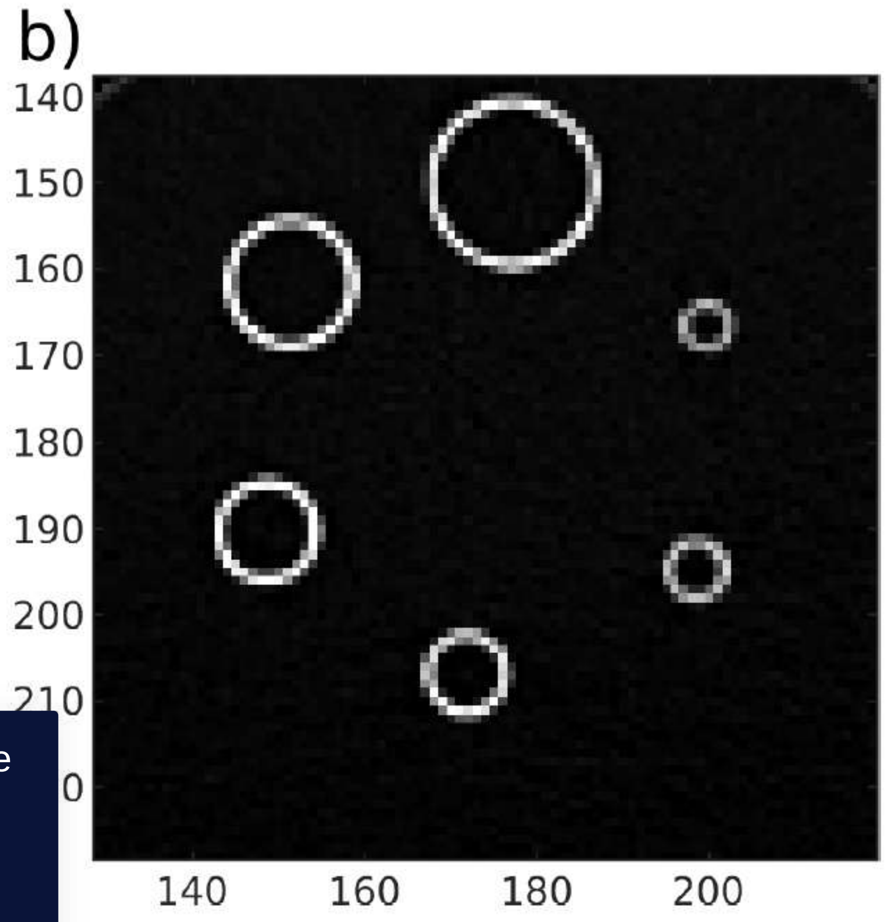
[Mosegaard and Tarantola, 1995]

PHANTOM CASE DATA

PET



CT

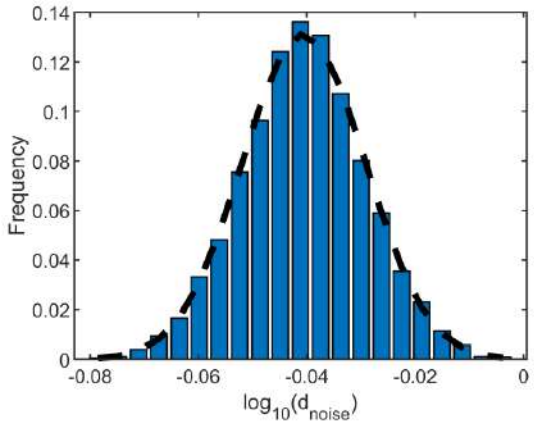


To apply probabilistic PET image analysis, we need to quantify

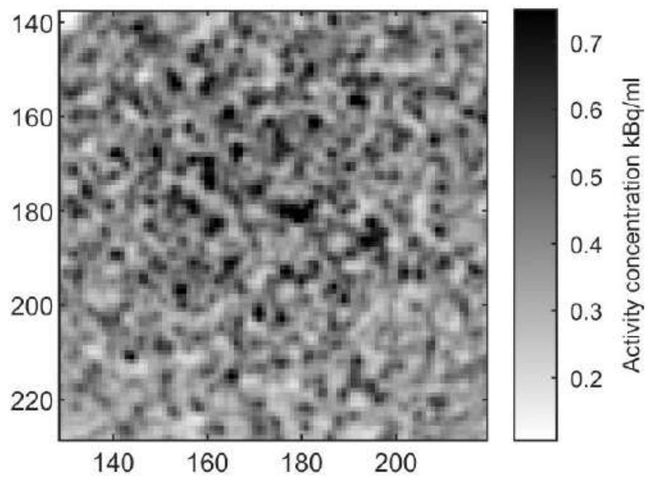
a) G
b) $L(\boldsymbol{\Phi}) = \rho_{\Phi}(G\boldsymbol{\Phi})$
c) $\rho(\boldsymbol{\Phi})$

PHANTOM CASE: THE NOISE, MULTIVARIATE CORRELATED LOG-NORMAL

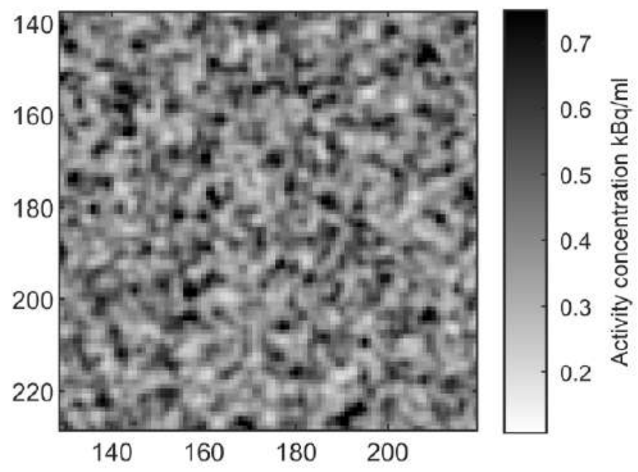
1D distribution



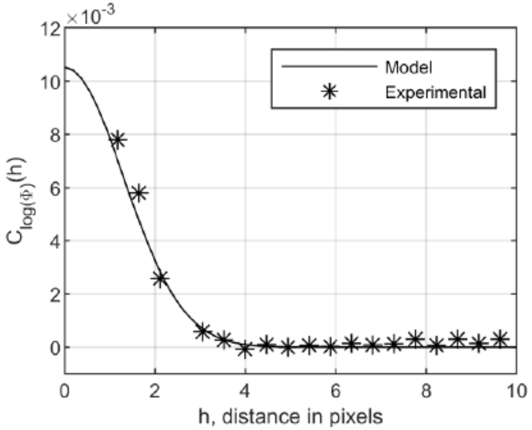
Actual noise



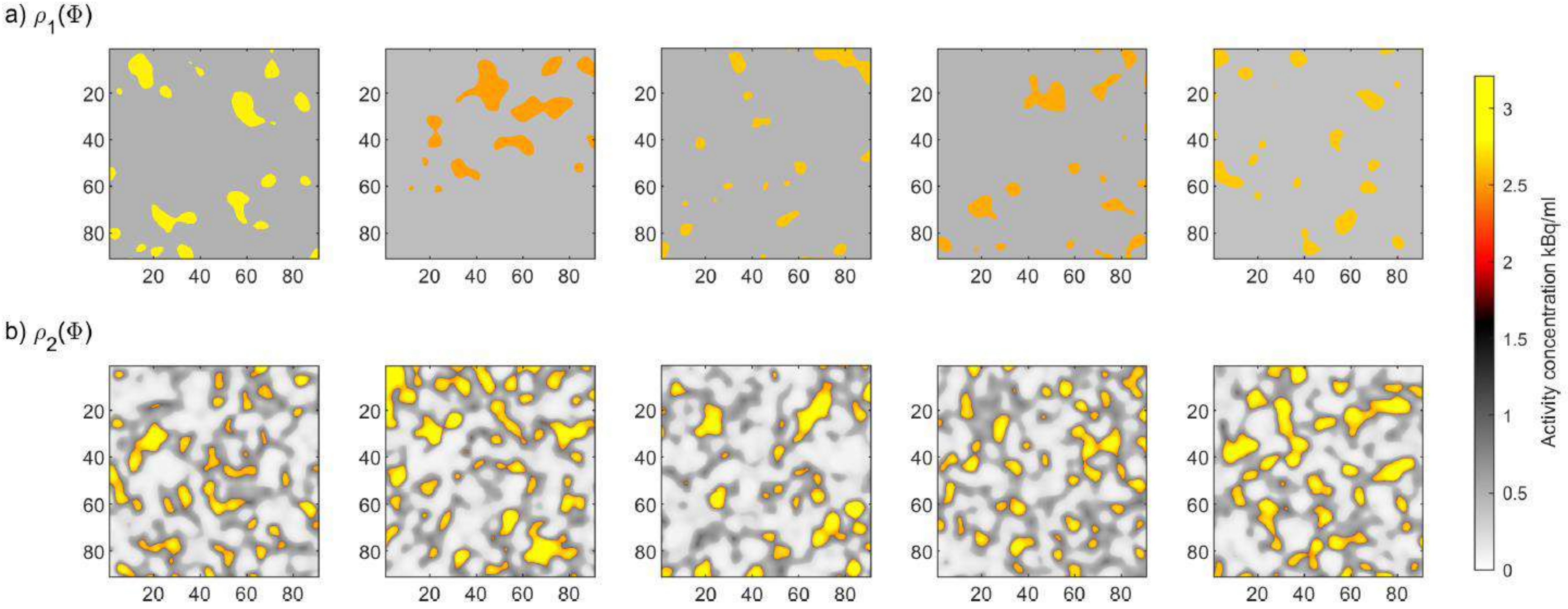
Realization of noise model



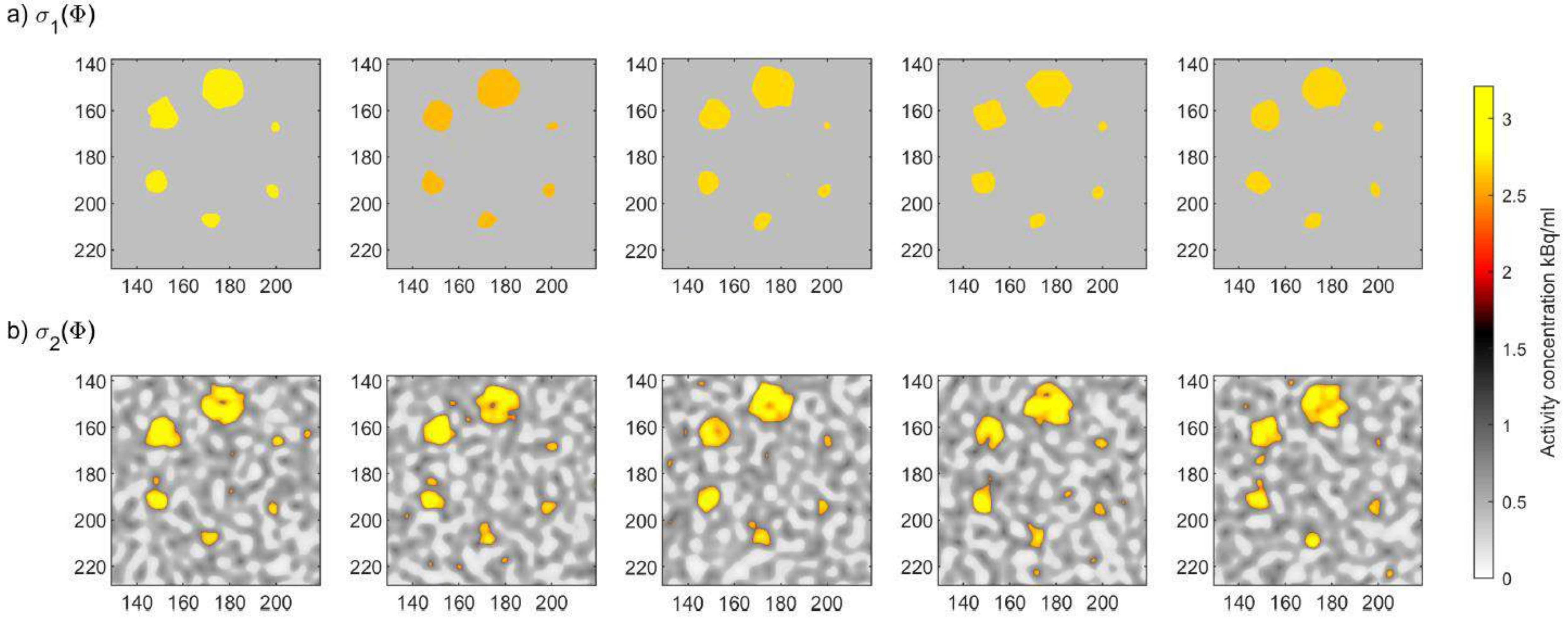
Experimental covariance



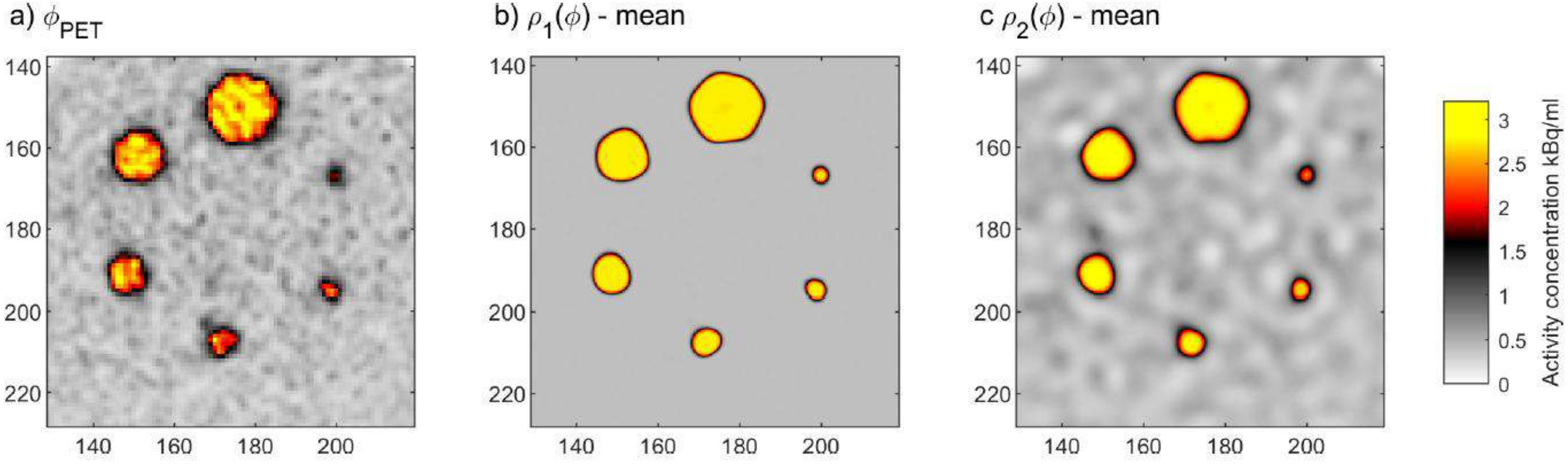
PHANTOM CASE, THE PRIOR



PHANTOM CASE, SAMPLING THE POSTERIOR



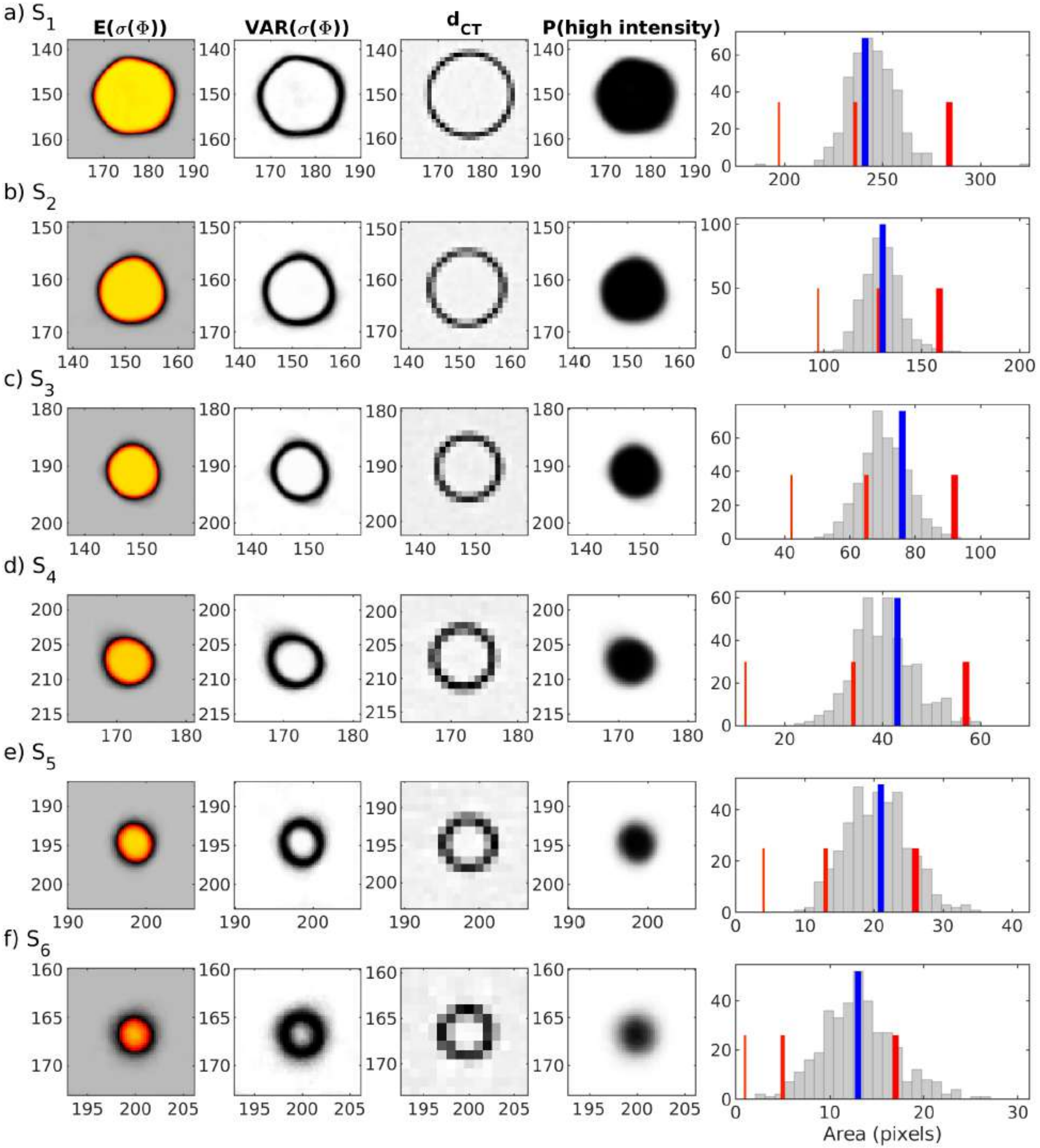
PHANTOM CASE – STATISTICS OF THE ~~POST~~ERIOR



Pixel size 2 mm x 2 mm

Pixel size 0.5 mm x 0.5 mm

SUB PIXEL RE~~SO~~LUTION



CONCLUSIONS

Tarantola and Valette (1982) provide a framework for probabilistic integration of information, that naturally allow accounting for uncertainty information.

In order to implement the method in practice we need to be able to quantify available information:

- (Algorithms that can quantify) prior information
  - The choice of choosing/quantifying prior information cannot be avoided
- Algorithms that compute the forward problem (physics)
- Quantification of modelling errors

- Algorithms for integration of information
  - Linear least squares, rejection sampler, Metropolis algorithm, Machine learning.