An approximate Bayesian inversion framework based on local Gaussian likelihoods

Martin Jullum
University of Oslo, Norway

Odd Kolbjørnsen
Lundin Norway (formerly Norwegian Computing Center)

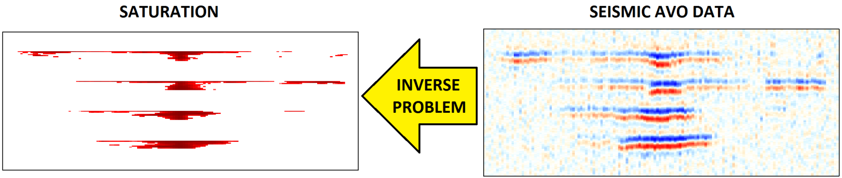
martinju@math.uio.no

September 10, 2015

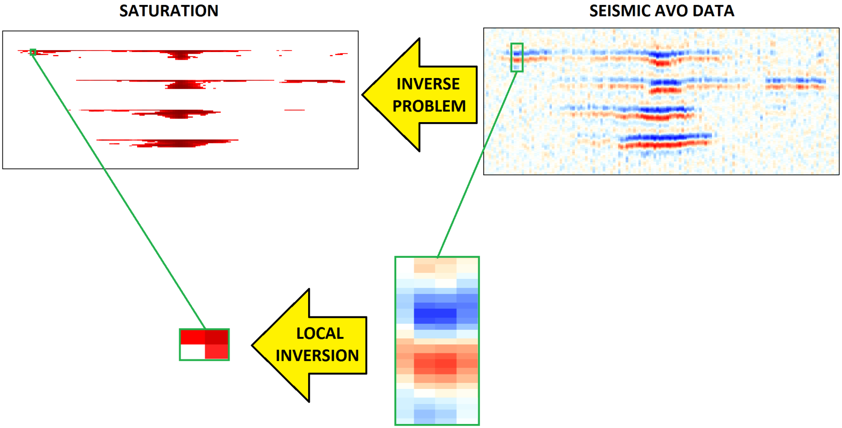
EAGE

PETROLEUM 2015 GEOSTATISTICS

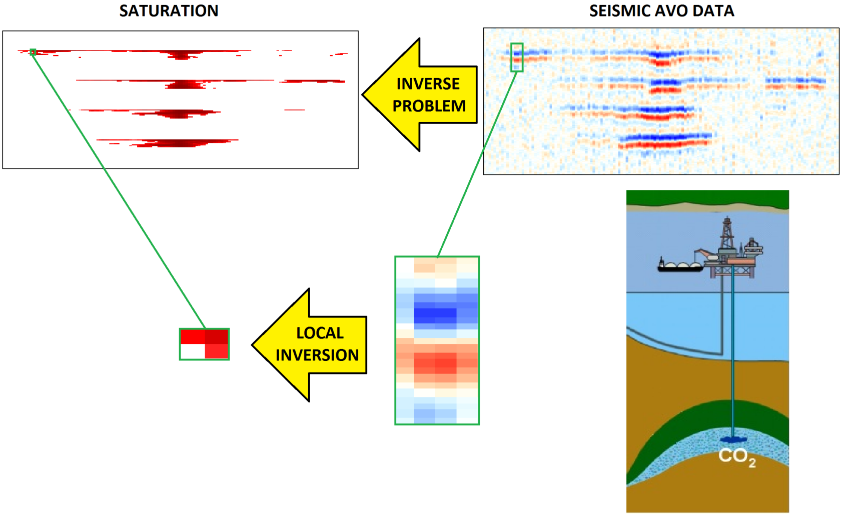
Outline



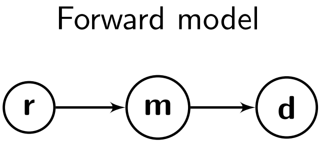
Outline


SATURATION
SEISMIC AVO DATA
INVERSE PROBLEM
LOCAL INVERSION

Outline

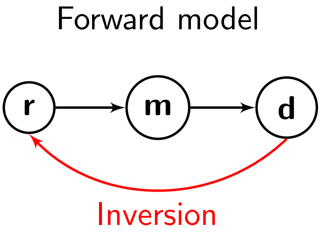

SATURATION
SEISMIC AVO DATA
INVERSE PROBLEM
LOCAL INVERSION
CO_2

Problem setup



- **r**: Rock properties (e.g. saturation, porosity, permeability, lithology, etc.)
- **m**: Geophysical properties (= elastic parameters $\mathbf{V}_p, \mathbf{V}_s$ and $\boldsymbol{\rho}$)
- **d**: Geophysical data (= Seismic AVO data)

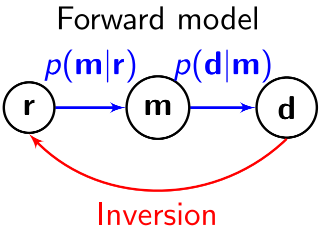
Problem setup



- **r**: Rock properties (e.g. saturation, porosity, permeability, lithology, etc.)
- **m**: Geophysical properties (= elastic parameters $\mathbf{V}_p, \mathbf{V}_s$ and $\boldsymbol{\rho}$)
- **d**: Geophysical data (= Seismic AVO data)

Inverse problem: Which values of **r** generated the observed **d**?

Problem setup

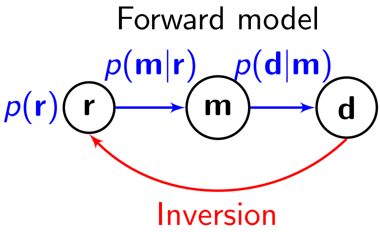


- **r**: Rock properties (e.g. saturation, porosity, permeability, lithology, etc.)
- **m**: Geophysical properties (= elastic parameters $\mathbf{V}_p, \mathbf{V}_s$ and ρ)
- **d**: Geophysical data (= Seismic AVO data)

Inverse problem: Which values of **r** generated the observed **d**?

- Statistical model setup

Problem setup



- $\mathbf{r}$: Rock properties (e.g. saturation, porosity, permeability, lithology, etc.)
- $\mathbf{m}$: Geophysical properties (= elastic parameters $\mathbf{V}_p, \mathbf{V}_s$ and ρ)
- $\mathbf{d}$: Geophysical data (= Seismic AVO data)

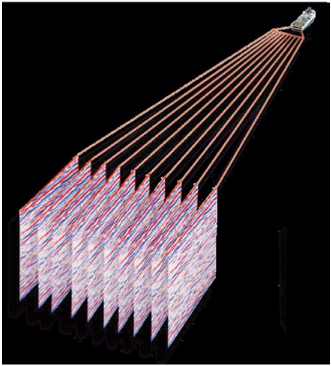
Inverse problem: Which values of $\mathbf{r}$ generated the observed $\mathbf{d}$?

- Statistical model setup
- The Bayesian approach is natural
  - Inversion $\Leftrightarrow$ consult posterior distribution $p(\mathbf{r}|\mathbf{d})$

Computing posterior

Posterior distribution: $p(\mathbf{r}|\mathbf{d}) \propto \int p(\mathbf{d}|\mathbf{m})p(\mathbf{m}|\mathbf{r})p(\mathbf{r})\,\mathrm{d}\mathbf{m}$

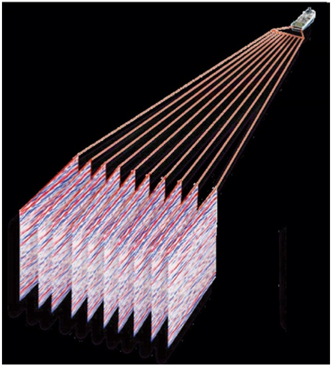
- High dimensional problem
- Enormous amount of highly correlated data **d**
- Complex dependency structures



Computing posterior

Posterior distribution: $p(\mathbf{r}|\mathbf{d}) \propto \int p(\mathbf{d}|\mathbf{m})p(\mathbf{m}|\mathbf{r})p(\mathbf{r})\,d\mathbf{m}$

- High dimensional problem
- Enormous amount of highly correlated data **d**
- Complex dependency structures
- Analytical expression for posterior seldom available
- Full Gaussian linear inversion (Buland and Omre (2003)) and similar approaches may not be valid
- MCMC can be very time consuming



Key elements of our approach

- Divide the global inversion problem into several local inversions
  - Approximate marginal posterior $p(\mathbf{r}_i|\mathbf{d})$ for each cell i in the gridded region, rather than the global $p(\mathbf{r}|\mathbf{d})$ for the full region

Key elements of our approach

- Divide the global inversion problem into several local inversions
  - Approximate marginal posterior $p(\mathbf{r}_i|\mathbf{d})$ for each cell i in the gridded region, rather than the global $p(\mathbf{r}|\mathbf{d})$ for the full region
- Dimension reduction by only using variables spatially close to cell i
- Clever utilization of properties of Gaussian distributions
- Monte Carlo sampling of lower dimensional integral

Key elements of our approach

- Divide the global inversion problem into several local inversions
  - Approximate marginal posterior $p(\mathbf{r}_i|\mathbf{d})$ for each cell i in the gridded region, rather than the global $p(\mathbf{r}|\mathbf{d})$ for the full region
- Dimension reduction by only using variables spatially close to cell i
- Clever utilization of properties of Gaussian distributions
- Monte Carlo sampling of lower dimensional integral
- Can be parallelized under stationarity assumptions

Key elements of our approach

- Divide the global inversion problem into several local inversions
  - Approximate marginal posterior $p(\mathbf{r}_i|\mathbf{d})$ for each cell i in the gridded region, rather than the global $p(\mathbf{r}|\mathbf{d})$ for the full region
- Dimension reduction by only using variables spatially close to cell i
- Clever utilization of properties of Gaussian distributions
- Monte Carlo sampling of lower dimensional integral
- Can be parallelized under stationarity assumptions

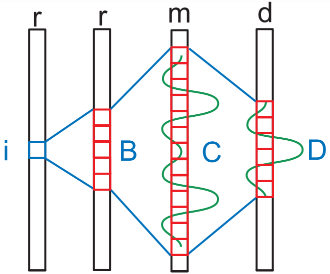
Assumptions:

- $p(\mathbf{d}|\mathbf{m}) \sim N(G\mathbf{m}, \Sigma)$, i.e. $\quad \mathbf{d} = G\mathbf{m} + \varepsilon, \quad \varepsilon \sim N(0, \Sigma)$
- We may sample from $p(\mathbf{m}|\mathbf{r})$ and $p(\mathbf{r})$

Local variables

Inversion for cell i: $p(\mathbf{r}_i|\mathbf{d})$

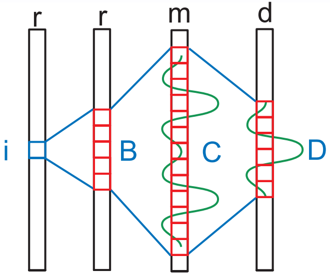
- 3 neighborhoods of cells: $B = B(i), C = C(i)$, $D = D(i)$
- These have corresponding local variables: $\mathbf{r}_B, \mathbf{m}_C, \mathbf{d}_D$



Local variables

Inversion for cell i: $p(\mathbf{r}_i|\mathbf{d})$

- 3 neighborhoods of cells:
  $B = B(i), C = C(i)$,
  $D = D(i)$
- These have corresponding local variables: $\mathbf{r}_B, \mathbf{m}_C, \mathbf{d}_D$



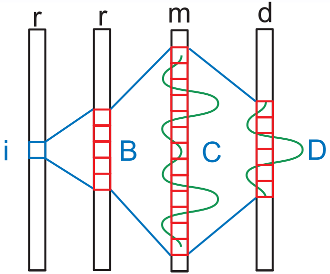
- Rely on the following relations

$$p(\mathbf{r}_i|\mathbf{d}) \approx p(\mathbf{r}_i|\mathbf{d}_D) = \int p(\mathbf{r}_B|\mathbf{d}_D)\,\mathrm{d}\mathbf{r}_{B_{-i}} \propto \int p(\mathbf{d}_D|\mathbf{r}_B)p(\mathbf{r}_B)\,\mathrm{d}\mathbf{r}_{B_{-i}}.$$

Local variables

Inversion for cell i: $p(\mathbf{r}_i|\mathbf{d})$

- 3 neighborhoods of cells: $B = B(i), C = C(i)$, $D = D(i)$
- These have corresponding local variables: $\mathbf{r}_B, \mathbf{m}_C, \mathbf{d}_D$



- Rely on the following relations

$$p(\mathbf{r}_i|\mathbf{d}) \approx p(\mathbf{r}_i|\mathbf{d}_D) = \int p(\mathbf{r}_B|\mathbf{d}_D)\,\mathrm{d}\mathbf{r}_{B_{-i}} \propto \int p(\mathbf{d}_D|\mathbf{r}_B)p(\mathbf{r}_B)\,\mathrm{d}\mathbf{r}_{B_{-i}}.$$

- We shall approximate $p(\mathbf{d}_D|\mathbf{r}_B)$

Local Gaussian likelihood approximation

Need to approximate the local likelihood

$$p(\mathbf{d}_D|\mathbf{r}_B) = \int p(\mathbf{d}_D|\mathbf{m}_C)p(\mathbf{m}_C|\mathbf{r}_B)\,\mathrm{d}\mathbf{m}_C$$

Local Gaussian likelihood approximation

Need to approximate the local likelihood

$$p(\mathbf{d}_D|\mathbf{r}_B) = \int p(\mathbf{d}_D|\mathbf{m}_C)p(\mathbf{m}_C|\mathbf{r}_B)\,\mathrm{d}\mathbf{m}_C$$

- $p(\mathbf{d}|\mathbf{m}) \sim N(G\mathbf{m}, \Sigma) \Rightarrow p(\mathbf{d}_D|\mathbf{m}_C) \approx p^*(\mathbf{d}_D|\mathbf{m}_C) \sim N(G_0\mathbf{m}_C, \Sigma_0)$

Local Gaussian likelihood approximation

Need to approximate the local likelihood

$$p(\mathbf{d}_D|\mathbf{r}_B) = \int p(\mathbf{d}_D|\mathbf{m}_C)p(\mathbf{m}_C|\mathbf{r}_B)\,\mathrm{d}\mathbf{m}_C$$

- $p(\mathbf{d}|\mathbf{m}) \sim N(G\mathbf{m}, \Sigma) \Rightarrow p(\mathbf{d}_D|\mathbf{m}_C) \approx p^*(\mathbf{d}_D|\mathbf{m}_C) \sim N(G_0\mathbf{m}_C, \Sigma_0)$
- $p(\mathbf{m}_C|\mathbf{r}_B) \approx p^*(\mathbf{m}_C|\mathbf{r}_B) \sim N(\boldsymbol{\mu}, \boldsymbol{\Gamma})$, where

$$\boldsymbol{\mu} = \boldsymbol{\mu}(\mathbf{r}_B) = \mathrm{E}^*[\mathbf{m}_C|\mathbf{r}_B] \text{ and } \boldsymbol{\Gamma} = \boldsymbol{\Gamma}(\mathbf{r}_B) = \mathrm{Cov}^*[\mathbf{m}_C|\mathbf{r}_B]$$

Local Gaussian likelihood approximation

Need to approximate the local likelihood

$$p(\mathbf{d}_D|\mathbf{r}_B) = \int p(\mathbf{d}_D|\mathbf{m}_C)p(\mathbf{m}_C|\mathbf{r}_B)\,\mathrm{d}\mathbf{m}_C$$

- $p(\mathbf{d}|\mathbf{m}) \sim N(G\mathbf{m}, \Sigma) \Rightarrow p(\mathbf{d}_D|\mathbf{m}_C) \approx p^*(\mathbf{d}_D|\mathbf{m}_C) \sim N(G_0\mathbf{m}_C, \Sigma_0)$
- $p(\mathbf{m}_C|\mathbf{r}_B) \approx p^*(\mathbf{m}_C|\mathbf{r}_B) \sim N(\boldsymbol{\mu}, \boldsymbol{\Gamma})$, where

$$\boldsymbol{\mu} = \boldsymbol{\mu}(\mathbf{r}_B) = \mathrm{E}^*[\mathbf{m}_C|\mathbf{r}_B] \text{ and } \boldsymbol{\Gamma} = \boldsymbol{\Gamma}(\mathbf{r}_B) = \mathrm{Cov}^*[\mathbf{m}_C|\mathbf{r}_B]$$

  - Sample large number of 'pairs' $(\mathbf{m}_C, \mathbf{r}_B)$ from joint distribution
  - Fit $\boldsymbol{\mu}$-function by nonlinear regression
  - Use residuals to fit different $\boldsymbol{\Gamma}$-functions for different $\mathbf{r}_B$

Local Gaussian likelihood approximation

Need to approximate the local likelihood

$$p(\mathbf{d}_D|\mathbf{r}_B) = \int p(\mathbf{d}_D|\mathbf{m}_C)p(\mathbf{m}_C|\mathbf{r}_B)\,\mathrm{d}\mathbf{m}_C$$

- $p(\mathbf{d}|\mathbf{m}) \sim N(G\mathbf{m}, \Sigma) \Rightarrow p(\mathbf{d}_D|\mathbf{m}_C) \approx p^*(\mathbf{d}_D|\mathbf{m}_C) \sim N(G_0\mathbf{m}_C, \Sigma_0)$
- $p(\mathbf{m}_C|\mathbf{r}_B) \approx p^*(\mathbf{m}_C|\mathbf{r}_B) \sim N(\boldsymbol{\mu}, \boldsymbol{\Gamma})$, where

$$\boldsymbol{\mu} = \boldsymbol{\mu}(\mathbf{r}_B) = \mathrm{E}^*[\mathbf{m}_C|\mathbf{r}_B] \text{ and } \boldsymbol{\Gamma} = \boldsymbol{\Gamma}(\mathbf{r}_B) = \mathrm{Cov}^*[\mathbf{m}_C|\mathbf{r}_B]$$

  - Sample large number of 'pairs' $(\mathbf{m}_C, \mathbf{r}_B)$ from joint distribution
  - Fit $\boldsymbol{\mu}$-function by nonlinear regression
  - Use residuals to fit different $\boldsymbol{\Gamma}$-functions for different $\mathbf{r}_B$

- Final approximation:

$$p(\mathbf{d}_D|\mathbf{r}_B) \approx \int p^*(\mathbf{d}_D|\mathbf{m}_C)p^*(\mathbf{m}_C|\mathbf{r}_B)\,\mathrm{d}\mathbf{m}_C$$
$$= p^*(\mathbf{d}_D|\mathbf{r}_B) \sim N(G_0\mu, \Sigma_0 + G_0\Gamma G_0^t)$$

Approximate inversion quantities

Approximate local posterior:

$$p(\mathbf{r}_i|\mathbf{d}_D) \approx p^*(\mathbf{r}_i|\mathbf{d}_D) \propto \int p^*(\mathbf{d}_D|\mathbf{r}_B)p(\mathbf{r}_B)\,\mathrm{d}\mathbf{r}_{B_{-i}}$$

- Inversion: Compute inversion quantities like the posterior mean, median, sd or credibility intervals

Approximate inversion quantities

Approximate local posterior:

$$p(\mathbf{r}_i|\mathbf{d}_D) \approx p^*(\mathbf{r}_i|\mathbf{d}_D) \propto \int p^*(\mathbf{d}_D|\mathbf{r}_B)p(\mathbf{r}_B)\,\mathrm{d}\mathbf{r}_{B_{-i}}$$

- Inversion: Compute inversion quantities like the posterior mean, median, sd or credibility intervals

Weighted Monte Carlo:

- Sample from $p(\mathbf{r}_B)$
- Weight the sampled $\mathbf{r}_i$ by $p^*(\mathbf{d}_D|\mathbf{r}_B)$ and normalize the weights
- Compute appropriate weighted averages over r_i to approximate inversion quantities

Approximate inversion quantities

Approximate local posterior:

$$p(\mathbf{r}_i|\mathbf{d}_D) \approx p^*(\mathbf{r}_i|\mathbf{d}_D) \propto \int p^*(\mathbf{d}_D|\mathbf{r}_B)p(\mathbf{r}_B)\,\mathrm{d}\mathbf{r}_{B_{-i}}$$

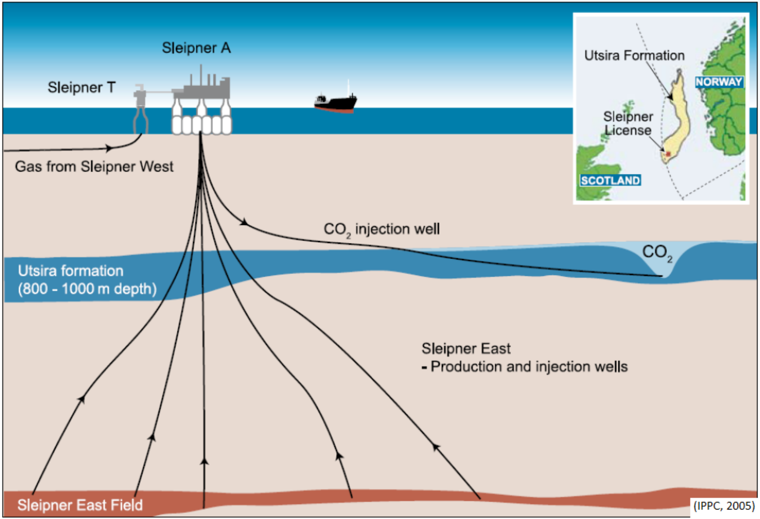
- Inversion: Compute inversion quantities like the posterior mean, median, sd or credibility intervals

Weighted Monte Carlo:

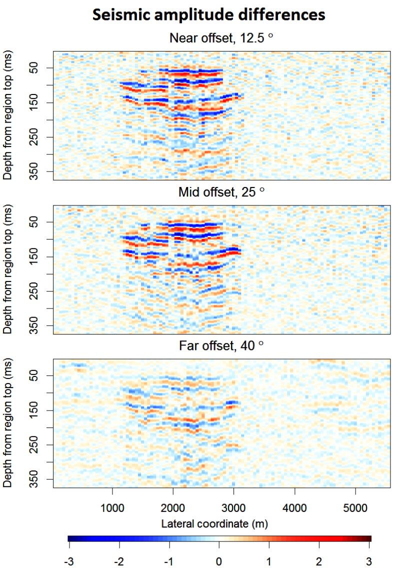
- Sample from $p(\mathbf{r}_B)$
- Weight the sampled $\mathbf{r}_i$ by $p^*(\mathbf{d}_D|\mathbf{r}_B)$ and normalize the weights
- Compute appropriate weighted averages over r_i to approximate inversion quantities

- Under stationarity: Re-use the general form of $p^*(\mathbf{d}_D|\mathbf{r}_B)$ and the samples from $p(\mathbf{r}_B)$ and repeat the Weighted Monte Carlo procedure for all i in the region

Illustration: CO_2 monitoring ⇒ map saturation



Real case

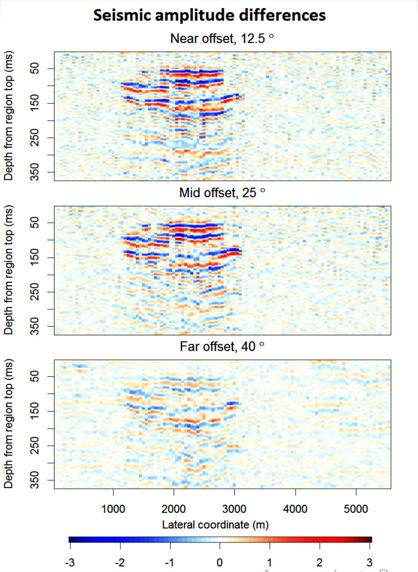


$\mathbf{d}$ = Seismic AVO difference data

$\mathbf{m}$ = change in log of $(\mathbf{V}_p, \mathbf{V}_s, \rho)$

$\mathbf{r}$ = saturation change

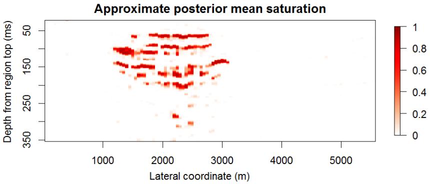
Real case



$\mathbf{d}$ = Seismic AVO difference data

$\mathbf{m}$ = change in log of $(\mathbf{V}_p, \mathbf{V}_s, \rho)$

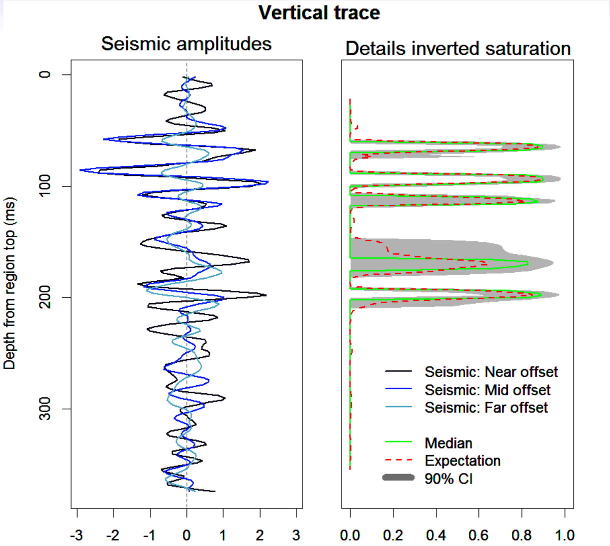
$\mathbf{r}$ = saturation change



18 000 cells inverted $<$ 30 min

on 4-cored Windows laptop

Real case


Vertical trace
Seismic amplitudes
Details inverted saturation
Depth from region top (ms)
0
100
200
300
-3 -2 -1 0 1 2 3
0.0 0.2 0.4 0.6 0.8 1.0
Seismic: Near offset
Seismic: Mid offset
Seismic: Far offset
Median
Expectation
90% CI

Method comparison based on inversion of synthetic data

Regular Gaussian inversion: $p^*(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p^*(\mathbf{m})$, Buland and Omre (2003)

Method comparison based on inversion of synthetic data

Regular Gaussian inversion: $p^*(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p^*(\mathbf{m})$, Buland and Omre (2003)

Define $\mathbf{r}$ = change in log density ($\boldsymbol{\rho}$)

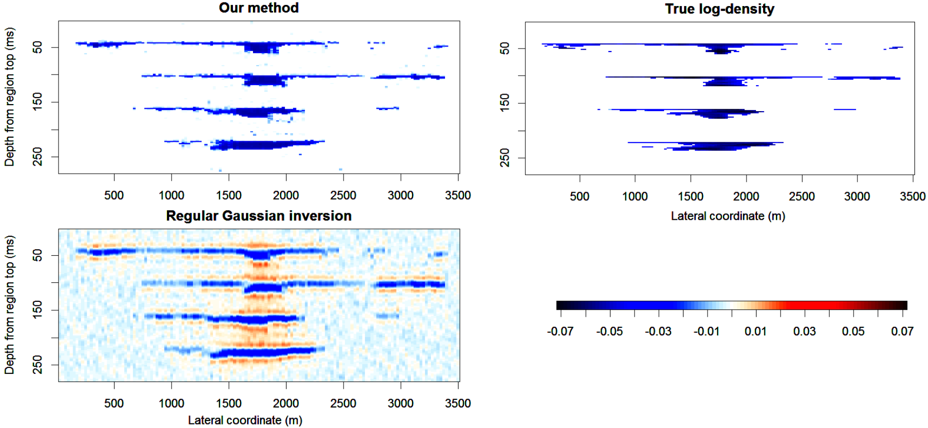
Our method: $p^*(\mathbf{r}_i|\mathbf{d}) \propto \int p^*(\mathbf{d}_D|\mathbf{r}_B)p(\mathbf{r}_B)\, d\mathbf{r}_{B_{-i}}$

Method comparison based on inversion of synthetic data

Regular Gaussian inversion: $p^*(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p^*(\mathbf{m})$, Buland and Omre (2003)

Define $\mathbf{r}$ = change in log density ($\boldsymbol{\rho}$)

Our method: $p^*(\mathbf{r}_i|\mathbf{d}) \propto \int p^*(\mathbf{d}_D|\mathbf{r}_B)p(\mathbf{r}_B)\,\mathrm{d}\mathbf{r}_{B_{-i}}$



Conclusion

Our method summarized:

- Split the global inversion problem into lots of cell-wise inversions
- Use only information in variables spatially close to the cell to invert
- Clever local-Gaussian approximations and weighted Monte Carlo

- Conceptually easy to extend the method
- Well suited for parallelization