Bayesian Inversion and Model Selection in Ocean Acoustics

Stan Dosso & Jan Dettmer

Earth & Ocean Sciences

University of Victoria
Introduction

- Remote sensing of geophysical properties via acoustic or seismic waves represent nonlinear inverse problems of great interest.

- **Bayesian** formulation and **Markov-chain Monte Carlo** methods provide a general fully-nonlinear approach to
  - Parameter estimation and **quantitative uncertainty analysis**
  - **Model selection** (e.g., parameterization)
Outline

- **Goal:** Overview of Bayesian inversion theory and implementation illustrated with examples:
  - Marine seismic amplitude-vs-angle inversion
  - Ocean acoustic reverberation inversion
  - Seabed reflection inversion (controlled source and ambient noise)
  - Seismic hazard site assessment via ambient noise inversion (on land)
Bayesian Formulation

- Let \( m, d \) be random variables → Bayes’ Rule:

\[
P(m \mid d, M) = \frac{P(d \mid m, M) P(m \mid M)}{P(d \mid M)}
\]

where

- \( d \) = Data
- \( M \) = Model: physical theory, parameterization, error distribution
- \( m \) = Model parameters
Prior Information

- Let \( m, d \) be random variables \( \rightarrow \) Bayes’ Rule:

\[
P(m \mid d, M) = \frac{P(d \mid m, M)P(m \mid M)}{P(d \mid M)}
\]

- **Prior distribution**: Existing parameter information independent of the data
  - Represents degree of belief
  - Can be (relatively) non-informative
    - e.g., uniform distribution over wide bounds
Let \( m, d \) be random variables→ Bayes’ Rule:

\[
P(m \mid d, M) = \frac{P(d \mid m, M)P(m \mid M)}{P(d \mid M)}
\]

**Likelihood**: Data information—Conditional data probability interpreted in terms of \( m \) for measured \( d \)

Defining the likelihood requires specifying the error distribution

- Can include variance/covariance parameters in inversion
- Should carry out *a posteriori* statistical validation
Posterior Probability Density

- Let \( m, d \) be random variables → Bayes’ Rule:

\[
P(m \mid d, M) = \frac{P(d \mid m, M) P(m \mid M)}{P(d \mid M)}
\]

- **PPD**: Information content of parameters given data, prior and model

- Interpret multi-dimensional distribution?
**PPD Interpretation**

- PPD interpreted in terms of parameter estimates, uncertainties, inter-relations:

\[
\hat{m} = \text{Arg}_{\max} \left\{ P(m | d, M) \right\} \quad \text{MAP model}
\]

\[
\bar{m} = \int m' P(m' | d, M) \, dm' \quad \text{Mean model}
\]

\[
P(m_i | d, M) = \int \delta (m_i - m_i') P(m' | d, M) \, dm' \quad \text{Marginal distribution}
\]

\[
C_{ij} = \int (m_i' - \bar{m}_i)(m_j' - \bar{m}_j) P(m' | d, M) \, dm' \quad \text{Covariance matrix}
\]

- Analytic solutions for “**Classical**” linear inverse problem with Gaussian errors and prior
PPD Interpretation

- PPD interpreted in terms of parameter estimates, uncertainties, inter-relations:

\[ \hat{m} = \text{Arg}_{\text{max}} \left\{ P(m \mid d, M) \right\} \]

\[ \bar{m} = \int m' \, P(m' \mid d, M) \, dm' \]

\[ P(m_i \mid d, M) = \int \delta(m_i - m_i') \, P(m' \mid d, M) \, dm' \]

\[ C_{ij} = \int (m_i' - \bar{m}_i)(m_j' - \bar{m}_j) \, P(m' \mid d, M) \, dm' \]

- **Nonlinear** inversion requires **numerical optimization** and **integration (sampling)** of PPD
MCMC Sampling

- PPD sampled via **Markov-chain Monte Carlo**
  - Gibbs sampling
  - Metropolis-Hasting sampling (MHS)

- MHS constructs Markov chain based on perturbations \( \mathbf{m} \rightarrow \mathbf{m}' \) generated with proposal distribution \( Q(\mathbf{m}' | \mathbf{m}) \) accepted with probability

\[
p = \text{Min}
\left\{ 1, \frac{P(d | \mathbf{m}', M) P(\mathbf{m}' | M) Q(\mathbf{m} | \mathbf{m}')}{P(d | \mathbf{m}, M) P(\mathbf{m} | M) Q(\mathbf{m}' | \mathbf{m})} \right\}
\]
MCMC Sampling

- PPD sampled via **Markov-chain Monte Carlo**
  - Gibbs sampling
  - Metropolis-Hasting sampling (MHS)

- MHS constructs Markov chain based on perturbations \( \mathbf{m} \rightarrow \mathbf{m}' \) generated with proposal distribution \( Q(\mathbf{m}'|\mathbf{m}) \) accepted with probability

\[
p = \text{Min}\{1, \exp[E(\mathbf{m}) - E(\mathbf{m}')]\}
\]

for fixed dimension, uniform prior, symmetric proposal

\( E(\mathbf{m}) = -\log \text{likelihood or data misfit} \)
Proposal: Principal Components

- **Perturbation direction:** Sample along eigenvectors of covariance matrix (principal components)
- **Perturbation size:** Eigenvalues give length scales (PC variances)

\[ C_m = U W U^T \quad \Rightarrow \quad m' = U^T m \]
Example 1: Amplitude vs Angle

- Inversion of marine seismic AVA data
- Two sites along seismic line in Baltic (mud, till)

Can different sediments be resolved?

Example 1: Amplitude vs Angle

- Inversion of marine seismic AVA data
- Two sites along seismic line in Baltic (mud, till)

Can different sediments be resolved?
Seismic Section

- Mud-Location
- CMP-No.
- Till location

TWT [s]

- Tilt (mud)
- Shallow gas
- Small channel
- Second layer (boulder clay)
- Shallow gas
- Deeper reflectors
- Sea floor (boulder clay)

First water multiple
Internal multiple

20 m

500 m
• Reflection coeffs: ratio of direct and reflected wave amplitudes (time windowed for 1-m penetration)
Marginals quantify the geoacoustic information content and sediment differences resolved by AVA data.
Joint Marginals

- Mud
- Till

The images show the joint marginals of shear wave velocity ($V_s$) and P-wave velocity ($V_p$) for mud and till. The density ($\rho$) is also plotted against $V_p$ for both materials.
Example 2: Solid-Solid Reflection

- Invert single-frequency (simulated) reflection coefficients between 2 solid layers
- Theory: Reflection coeffs depend on parameter contrasts
Principal Components

- PCs represent uncorrelated linear combinations of parameters, ordered by variance

![Graph showing Principal Components](image-url)
Joint Marginals

Parameter sampling:

PC sampling:
Joint Marginals

5000 samples

Parameter sampling:

PC sampling:
Joint Marginals

50,000 samples

Parameter sampling:

PC sampling:
Joint Marginals

50 000 / 500 samples

Parameter sampling:

PC sampling:
Example 3: Reverb Inversion

- Invert mono-static reverberation level vs time (range) for seabed geoacoustic and scattering parameters.

![Graph showing reverberation level vs range](image)
Seabed Parameters

- 2-Layer seabed model includes:
  - geoacoustic parameters (define reflectivity)
  - interface roughness & volume inhomogeneity (define back-scattering)

\[
R_i(k_x, k_y) = \frac{w_i}{(k_x^2 + k_y^2)^{\gamma_i/2}}
\]
Reverb Inversion

\[ \sigma \text{ (dB)} \]
\[ h \text{ (m)} \]
\[ \rho_1 \text{ (g/cm}^3\text{)} \]
\[ \rho_2 \text{ (g/cm}^3\text{)} \]
\[ \alpha_2 \text{ (dB/m/kHz)} \]
\[ \gamma_1 \]
\[ \gamma_2 \]
\[ \log_{10} w_1 \]
\[ \log_{10} w_2 \]
\[ \log_{10} \sigma_v \]
Reverb Inversion
Joint Marginals—MHS
Tempered Sampling

- Wider parameter sampling by tempering at temperature $T > 1$

  i.e., sample $P(m|d, M)^{1/T}$ with acceptance

  $$p = \text{Min}\{1, \exp\left[\frac{(E(m) - E(m'))}{T}\right]\}$$

- Helps transition multi-modal PPDs but:
  - Sampling is **inefficient**
  - Sampling is **biased**
Parallel Tempering

- Run parallel MHS chains at a series of temperatures $T_1 = 1,\ T_2 > T_1,\ T_3 > T_2 \ldots$

- Exchange between chains with probability

$$p = \text{Min} \left\{ 1, \exp \left[ \frac{(E(m') - E(m))}{(1/T' - 1/T)} \right] \right\}$$

- High $T$ chains widely sample parameter space
- $T = 1$ chain gives efficient unbiased sampling
- Interacting chains provide effective ensemble sampler
Parallel Tempering

Monte Carlo Step

Parameter space

Earl & Dean, 2005
Parallel Tempering

- $10^5$ samples
- $2 \times 10^5$ samples
- $5 \times 10^5$ samples
Model Selection: Evidence

- Let \( \mathbf{m}, \mathbf{d} \) be random variables → Bayes’ Rule:

\[
P(\mathbf{m} | \mathbf{d}, M) = \frac{P(\mathbf{d} | \mathbf{m}, M) P(\mathbf{m} | M)}{P(\mathbf{d} | M)}
\]

- **Bayesian Evidence:**
  Likelihood of model \( M \) for measured data \( \mathbf{d} \)

- **Model selection:**
  Maximize evidence to determine most appropriate model (e.g., parameterization)
Bayesian Razor

- **Parsimony:**
  - Evidence discriminates against models with unnecessary predictive power
  - Evidence intrinsically prefers simple models
Evidence & BIC

\[ P(d \mid M) = \int P(d \mid m, M) P(m \mid M) \, dm \]

- Difficult multi-dimensional integral
- **Bayesian information criterion (BIC)** provides asymptotic point estimate:

\[
\text{BIC} \approx -2 \log_e P(d \mid M)
\]

\[
\text{BIC} = -2 \log_e P(d \mid \hat{m}, M) + N_m \log_e N_d
\]

# parameters # data
Model Parameterization

- Determining appropriate parameterization (e.g., number of seabed layers) important

- **Under-parameterization:**
  - Under-fits data leaving seabed structure unresolved
  - Can bias parameter estimates, under-estimate uncertainties

- **Over-parameterization:**
  - Over-fits data
  - Under-constrains parameters leading to spurious structure
  - Over-estimates parameter uncertainties
Model Parameterization

- Determining appropriate parameterization (e.g., number of seabed layers) important

- **Under-parameterization:**
  - Under-fits data leaving seabed structure unresolved
  - Can bias parameter estimates, under-estimate uncertainties

- **Over-parameterization:**
  - Over-fits data
  - Under-constrains parameters leading to spurious structure
  - Over-estimates parameter uncertainties

→ Evidence/BIC identifies simplest parameterization consistent with data resolving power
Model Selection: Trans-D Inversion

“When the number of things you don’t know is one of the things you don’t know”

(Peter Green, 1995)
Model Selection: Trans-D Inversion

- **Trans-Dimensional Inversion:**
  Number of unknown parameters is treated as an unknown which is integrated over (sampled according to its probability)

  - PPD spans multiple parameter spaces of different dimensions
  - Trans-D inversion accounts for uncertainty in parameterization in parameter uncertainty estimates
Consider $m_k$ where $k \in K$ indexes possible models, e.g., $k$ can be number of seabed layers.

Hierarchical trans-D PPD:

$$P(m_k, k | d) = \frac{P(d | m_k, k) P(m_k | k) P(k)}{\sum_{k' \in K} \int P(d | m_{k'}, k') P(m_{k'} | k') P(k') dm_{k'}}$$

sampled with trans-D MCMC.
Trans-D MCMC

- Trans-D MCMC acceptance rule:

\[
p = \text{Min}\left\{1, \frac{P(d | m'_k, k')} {P(d | m_k, k)} \cdot \frac{P(m'_k | k')P(k')}{P(m_k | k)P(k)} \cdot \frac{Q(m_k', k | m_k', k')}{Q(m_k', k' | m_k, k)} \cdot |J| \right\}
\]

where \(J\) is the Jacobian for \((m_k, k) \rightarrow (m'_k, k')\)

- Intrinsic Bayesian parsimony:
  Increase in likelihood associated with increase in parameters \(k \rightarrow k'\) countered by decrease in prior probability due to prior volume increase
Birth-Death Reversible Jump MCMC

- 3 Types of Steps:
  - Parameter perturbations
  - **Birth**—Layer addition
  - **Death**—Layer deletion

- Birth and death proposals formulated so reverse jump always possible and $|J| = 1$

- Dimension jumps can have low acceptance
  - Apply parallel tempering
Example 4: Wide-angle Reflection

Ship fires “Boomer” acoustic source
Reflection Measurements

Acoustic wave interacts with seabed
Reflection Measurements

Direct wave recorded at receiver
Reflection Measurements

Reflected wave recorded at receiver
Reflection Measurements

Moving ship shoots again…
Trans-D Inversion

- Treat number of seabed layers as an unknown sampled in inversion
Spherical Reflection Coeffs

- Geometry requires modelling **spherical-wave reflection coefficients**

\[
R_s(\theta, f) \propto \int_0^{\pi/2-i\infty} J_0(k_0 r \cos \theta') e^{ik_0 z \sin \theta'} R_p(\theta', f) \cos \theta' \, d\theta'
\]

- Computationally intensive:
  - Massively parallel sampling (160 cores)
  - GPU processing (~30 x speed-up)
Reflection Data

Ensemble sampling, spherical-wave reflection model
Dimensional Chain Mixing

![Dimensional Chain Mixing Diagram](image)
Reflection Inversion & Cores
Example 5: Noise Reflection Inversion

- Invert seabed reflectivity (bottom loss) from up- and down-going fields due to surface wave noise

\[ BL = -10 \log |R|^2 \text{ (dB)} \]
Mediterranean Sea Data

Ambient-noise Inversion

Active-source Inversion
Example 6: Seismic Site Assessment

Earthquakes in or near Canada, 1627 - 2007
Crustal EQs (M < 7-8)

Intra-slab EQs (M < 7-8)

Inter-plate (megathrust) EQs (M ~ 9+)

Tectonics & Earthquakes
Site (Hazard) Assessment

- Local **site geology** strongly affects shaking in EQ
  - Soft sediments amplify shaking
  - Layering causes resonance at specific frequencies
  - Saturated sediments liquefy during shaking
  - Weak soils fail in landslides
Site (Hazard) Assessment

- Local **site geology** strongly affects shaking in EQ
  - Soft sediments amplify shaking
  - Layering causes resonance at specific frequencies
  - Saturated sediments liquefy during shaking
  - Weak soils fail in landslides

- **Shear-wave velocity profile** $V_s(z)$ over top 10s m useful indicator of shaking response

Estimated from:
- Geology
- Borehole measurements
- Seismic cone penetration test (SCPT)
- Inversion of dispersion of surface (Rayleigh) waves

Approximate

Invasive, Expensive
Seismic-Noise Inversion

- Invert high-frequency dispersion of ambient seismic noise for shallow $V_s(z)$ for EQ site assessment

- Sources: Urban noise, traffic, wind, waves

- Source and directionality of noise unknown:
  - Geophone array used to extract phase-velocity dispersion
  - Array size varied to resolve different wavelength (penetration depth) ranges
Study Sites

- **Fraser River Delta (Vancouver):**
  - Pop: 2,000,000
  - Up to 500 m of unconsolidated sediments
  - Borehole & 4 SCPT

- **Victoria:**
  - Pop: 300,000
  - 0-30 m soft silts over bedrock
  - 1 SCPT
Fraser Delta Dispersion Data

- Dispersion curve built up over frequency from expanding array configurations
Fraser Delta Dispersion Data

- Three data “segments” from different array sizes
Data Covariance Matrix

- Estimate block-Toplitz covariance matrix from residual analysis
Parameterization: BIC

\[ \text{BIC} = -2 \log_e P(d | \hat{\mathbf{m}}, M) + N_m \log_e N_d \]

**U** … Uniform

**L** … Linear gradient

**P** … Power law
Marginal Probability Profile
Credibility Profile

- **Inversion:**
  MAP estimate and 95% highest probability density credibility interval

- **Direct Measures:**
  $V_s$ average from borehole and 4 SCPTs
Victoria Dispersion Data

![Diagram showing phase velocity vs frequency](image-url)
Parameterization: BIC

- **U**...Uniform
- **L**...Linear gradient
- **P**...Power law
Credibility Profile

- **Inversion:**
  MAP estimate and 95% highest probability density credibility interval

- **Direct Measures:**
  $V_s$ from SCPT (refusal at 17 m)
Probabilistic Site Assessment

- Building code classes based on $V_{S30} = \text{average } V_S(z)$ over upper 30 m
  - A – hard rock
  - B – rock
  - C – soft rock/dense soil
  - D – stiff soil
  - E – soft soil
Peak Ground Acceleration

- Probability distribution for peak ground acceleration relative to bedrock reference
Amplification

- Amplification probability spectra due to reduced impedance of upper soils
Resonance

- Amplification probability spectra due to resonance within near-surface layers
Summary

- **Bayesian** formulation and **Markov-chain Monte Carlo** methods provide quantitative and general nonlinear approach to wave inversion
  - Model selection (e.g., parameterization)
  - Parameter estimation and uncertainty analysis

- **Thanks to:**
  - AVA inversion: Michael Riedel (*PhD*)
  - Scholte-wave noise inversion: Cuilin Li (*PhD*)
  - Reflection inversion: Jan Dettmer (*PhD/Research Assoc.*)
  - Noise reflection inversion: Jorge Quijano (*Postdoc*)
  - Rayleigh-wave noise inversion: Sheri Molnar (*PhD*)