Uncertainty Estimation: Overview

1. Prior information
2. Model selection
3. Data misfit
4. Parameter estimation
5. Uncertainty estimation
6. Uncertainty/variability
7. Joint Inversion
1. Prior Information

- Quantitative information applied to inversion independent of measured data

- **Explicit:**
  - Parameter bounds (bounded uniform distribution)
  - Non-uniform prior distributions
  - Inter-parameter relationships

- **Implicit:**
  - Physics models and parameterizations considered

---

**Hamilton data**

- Plot showing a relationship between material properties and wave velocities.
Prior Information

- Prior information (particularly parameterization, hard bounds) can strongly influence solution
  - Important to specify priors in comparing uncertainty results

- Common goal:
  - Constrain parameters to physically-reasonable values
  - Allow data information to primarily determine solution

- If data and prior disagree:
  - Reassess data and error estimates
  - Reassess prior, including physics model and parameterization
2. Model Selection

- **Physics model**
  - Fluid, elastic or poro-elastic?
  - Range independent/dependent?
  - Plane wave or spherical wave?

- **Model parameterization**
  - Number of layers/segments?

---

n layers—best choice of n?
Model Selection

- Quantitative uncertainty estimation requires appropriate model parameterization
  - **Under-parameterization** can lead to under-fitting data, biased parameter estimates, under-estimated uncertainties
  - **Over-parameterization** can lead to over-fitting data, unconstrained structure, over-estimated uncertainties

- Seek simplest parameterization consistent with resolving power of the data
Model Selection

- Qualitative Model Selection:
  - Based on insight and experience

- Quantitative Model Selection:
  - **Bayesian information criterion** (BIC)—point estimate based on optimization that balances data fit and number of parameters
  - **Evidence**—Integral estimate of parameterization likelihood given the data, based on sampling
  - **Trans-dimensional inversion**
  - **Multiple-model particle filter**

Include number of parameters as unknown in inversion
Example: BIC

- Invert Scholte (interface) wave dispersion curves from ambient noise
  - Invert fundamental mode only
  - Invert first 3 modes
BIC: 1 & 3 Mode Inversions

1 mode: 5 layers resolved

3 modes: 8 layers resolved
MAP Profiles
Marginal Probability Profiles
3. Data Misfit Function

- Misfit quantifies difference between measured and modeled data

- Parameter estimation (optimization):
  - Minimize any reasonable misfit function; result is corresponding best-fit model according
  - Likelihood-based misfit provides \textit{efficient} estimator

- Uncertainty estimation:
  - Generally requires likelihood-based misfit
  - Maxent methods can specify least-informative misfit function for a given constraint
Likelihood Function

- **Likelihood:** Interprets data uncertainty distribution as a function of model parameters
  - Consistent with inversion as mapping data uncertainty distribution (data space) to parameter uncertainty distribution (model space)

- Requires estimating data uncertainties (measurement and theory errors)
  - Form of distribution (Gaussian, Laplace, …)
  - Statistical properties (variance, covariance) estimated from data residuals or included as unknowns in inversion
Examples

- IID Gaussian data errors

\[ P(d,m) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{|d - d(m)|^2}{2\sigma^2}\right] \]

misfit \( E(m) \)

(least squares)

- IID Gaussian errors, unknown source strength

\[ E(m) = \left[ |d|^2 - \frac{|d^T d(m)|^2}{|d(m)|^2} \right] / \sigma^2 \]

(Bartlett processor)
Data Errors

- Specifying likelihood requires quantifying the data error distribution

- Data errors = Inability to model measured data:
  - **Measurement errors**: ambient noise, instrumental uncertainties, etc.
  - **Theory errors**: due to idealized physics and simplified parameterization, etc.

- Ensure modeling is as accurate as possible and data sample over error processes (difficult)
  - Sample over noise, internal waves, variability, etc.
  - Collect multiple data sets (same & different types)
  - Note: beyond a point, denser data lead to correlated errors
4. Parameter Estimation

- Minimize data misfit via optimization
  - Linearized inversion (prone to local minima)
  - Global search
  - Hybrid optimization

- Repeat optimization to ensure stable result

- Mean model via sampling
5. Uncertainty Estimation

- **Linearization**—Analytic result
  (exact solution to approximate problem)
  - Gaussian data uncertainties and unbounded-uniform or Gaussian prior leads to Gaussian parameter uncertainties
  - Efficient, potentially inaccurate

- **Nonlinear**—Numerical sampling
  (approx solution to exact problem)
  - Monte Carlo/Importance sampling
  - Markov-chain Monte Carlo
    (Metropolis Hastings, Gibbs sampling)
  - Parallel-tempering
  - Numerically intensive; sampling/convergence issues
Joint Uncertainties—Reverb
Joint Uncertainties—Reflection

nonlinear

linearized
Uncertainties—Reverb/Scattering
Experiment Planning: Simulation

- Uncertainty estimation for simulations quantifies ideal sensitivity and can help plan experiment factors

Example: Frequencies in MFI
6. Variability & Uncertainty

• Variability
  - Measure of inherent spatial or temporal heterogeneity in an environmental property
  - Ideally quantified statistically/probabilistically
  - Intrinsic property of the environment—cannot be reduced by improved experiment or data analysis, although these can improve variability estimates

• Uncertainty
  - Measure of knowledge of an environmental parameter
  - Ideally quantified statistically/probabilistically
  - Property of environmental knowledge, not of the environment itself—can be reduced by improved experiments or data analysis
Variability & Uncertainty

- Inversion uncertainties quantify accuracy of the model parameter estimates adopted to represent the environment.

- Consider a parameter (e.g., sound speed of upper layer) over an experimental footprint:
  - Uncertainty quantifies accuracy of average sound speed over footprint.
  - Uncertainty does not quantify sound-speed variability over footprint (accurate average could be obtained for a highly variable property).
  - Parameter estimates involve non-uniform averaging so care required in interpretation.
Variability & Uncertainty are distinct but related

- Variability can cause theory/modeling errors which lead to parameter uncertainties
- If theory errors due to variability dominate and are adequately sampled, uncertainty estimates can quantify variability (care required)
Variability & Uncertainty

• Variability study:

  ➢ Localized, high-resolution measurements closely spaced in space or time

  ➢ Significant differences between recovered parameters represent variability

  ➢ Uncertainty estimation essential to determine if observed differences due to environmental variability or uncertain parameter estimates
Sequential Trans-D Inversion

- **AUV-towed source and array:**
  - Reflection data for small seafloor footprint
  - Mobile system for sub-bottom mapping
  - Reduces effects of seabed/ocean variability
Sequential Trans-D Inversion
7. Joint Inversion

- Joint (simultaneous) inversion of different data brings more information to bear

- Different physics for different data can overcome
  - Low sensitivity to some parameters
  - Inter-parameter correlations
Example: Reverb/Prop Inversion

- Invert (separately and jointly):
  - Short-range propagation data
  - Reverb data

\[ z_s, R, z_r, \mu, c_1, \rho_1, \alpha_1, c_2, \rho_2, \alpha_2, h, D \]
Reverb Inversion—Joint Marginals

- Strong inter-parameter correlations from reverb physics
Different parameter correlations arise from different physics.
Reverb + Propagation Inversion

- Geoacoustics & scattering well resolved
Inversion Comparison

Reverb

Prop

Reverb + Prop

Reverb

Prop

Reverb + Prop