Overview of today’s joint inversion methodology and its future into the machine-learning era for geophysical imaging

Monica Maceira

Chengping Chai & Hector Santos-Villalobos (ORNL)
E. Syracuse (LANL), H. Zhang (USTC), C.J. Ammon (PSU)
Seismic tomography – Structure and evolution of the Earth

Imaging patterns of thermo-chemical convection

S-wave velocity of the African superplume (Ritsema et al., 2011)

P-wave velocity beneath Indonesia and northwest Pacific (Simmons et al., 2012)
Seismic tomography – Seismic ground motion predictions

Focusing and defocusing of seismic waves
Strong amplification in sedimentary basins
Seismic tomography – Hydrocarbon exploration

Detection and characterization of reservoirs
Monitoring and time-dependent changes

The Valhall Gas Cloud

(Sirgue et al., 2010)
Looking into a bright future

- numerical wave propagation
- high-performance computing
- dense seismometer networks

global wave propagation LLNL-G3Dv3

ORNLSUMMIT

IberARRAY deployment
Challenges

different Earth models

Schaeffer & Lebedev 2013
Shapiro & Ritzwoller 2002
Debayle & Ricard 2012
Lekic & Romanowicz 2011
Kustowski et al. 2008
Ritsema et al. 2011
Challenges
different Earth models

Necessary Developments
assimilation of all available data

Different data subsets lead to different models
Challenges

different Earth models
small scale structure & the crust

Necessary Developments

assimilation of all available data

✓ Thin layer
✓ Difficult to resolve in global tomography
✓ Usually assumed known and fixed
✓ Hope: details do not affect results too much!
Challenges

different Earth models
small scale structure & the crust

Necessary Developments

assimilation of all available data

CCs from CRUST2.0 (Bassin et al., 2000)
CCs from CRUST07 (Meier et al., 2007)
CCs from 3SMAC (Nataf & Ricard, 1996)

(Ferreira et al., 2010)
Challenges

different Earth models
small scale structure & the crust

Necessary Developments

assimilation of all available data
joint resolution of crust and mantle

CCs from CRUST2.0 (Bassin et al., 2000)

CCs from 3SMAC (Nataf & Ricard, 1996)

CCs from CRUST07 (Meier et al., 2007)

Small-scale details of the crust matter for large-scale mantle structure
Challenges

different Earth models
small scale structure & the crust
computational resources

Necessary Developments

assimilation of all available data
joint resolution of crust and mantle

Compressional waves propagate through the whole Earth at

min period ~ 1s
Challenges

different Earth models
small scale structure & the crust
computational resources

Necessary Developments

assimilation of all available data
joint resolution of crust and mantle

Compressional waves propagate through the whole Earth at

\[ \text{min period } \approx 1 \text{s} \]

Today
Tomography based on fully numerical wave propagation

\[ \text{min period } \approx 10 \text{s} \]

In ~ 50 years
Tomography based on fully numerical wave propagation

\[ \text{min period } \approx 1 \text{s} \]

- Moore’s law continues to hold
- Handling computers 100,000 times bigger
- Code to harness such resources
### Challenges

different Earth models  
small scale structure & the crust  
computational resources

### Compressional waves

Propagate through the whole Earth at

<table>
<thead>
<tr>
<th>Min period</th>
<th>Today</th>
<th>In ~ 50 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>~ 1s</td>
<td>~ 20s</td>
<td>~ 1s</td>
</tr>
</tbody>
</table>

### Necessary Developments

Assimilation of all available data  
Joint resolution of crust and mantle  
Combination of inversion techniques

- Moore’s law continues to hold  
- Handling computers 100,000 times bigger  
- Code to harness such resources

Computer power alone is unlikely to solve the problem
Challenges

different Earth models
small scale structure & the crust
computational resources

Necessary Developments

assimilation of all available data
joint resolution of crust and mantle
combination of inversion techniques

SIMULTANEOUS JOINT INVERSION OF DISPARATE GEOPHYSICAL OBSERVATIONS
Why simultaneous joint inversion of disparate observations?

Multiple benefits

☑ “Standard” geophysical models are developed only to fit one type of data.

(Maceira et al., 2005)
Why simultaneous joint inversion of disparate observations?

Multiple benefits

✓ “Standard” geophysical models are developed only to fit one type of data.
✓ Different data sets have different spatial coverage and resolution.
✓ Different data types have different strengths.

Multiple challenges

○ Deal with different bandwidths.
○ Design responsive misfit norms; relative weighting of data sets.
○ Make assumptions to model the different data; relationships between independent data sets.
Evolution of Analysis Codes:

It takes a village!

SLU/LANL
1989
Ammon & Randall
P Receiver Functions

SLU
1998
Julia, Ammon, Herrmann
+ Surface Wave Dispersion (Saito Functions)

PSU
2002
Julia
+ Local Traveltimes

LANL
Maceira
+ Gravity, “3D”
2006
Ammon
+ Period Weights, Transverse Isotropy (Saito Functions)

LANL/SLU
SLU
P Receiver Functions
+ Surface Wave Dispersion (Saito Functions)

Herrmann
Computer Programs in Seismology

2001-02

LANL/PSU
Chai & Ammon
+ Improvements
2009-15

LANL/MIT
Maceira & Zhang
+ Body-wave traveltimes
2009-14
Challenges

different Earth models
small scale structure & the crust
computational resources

Necessary Developments

assimilation of all available data
joint resolution of crust and mantle
combination of inversion techniques

SIMULTANEOUS JOINT INVERSION OF DISPARATE GEOPHYSICAL OBSERVATIONS

APPLICATIONS
Application I Colombia; several subduction models
Semblance tests of the recovery of 100 randomly perturbed synthetic velocity models

(Syracuse et al., 2016)
Caribbean segment: shallowly dipping, non-volcanogenic subduction of the dry, buoyant Caribbean large igneous province.

Bucaramanga segment: flat-slab, non-volcanogenic subduction of the Nazca plate, with the Bucaramanga nest influenced by interaction with the Caribbean segment.

Cauca segment: steeply dipping, volcanogenic subduction of the Nazca plate, offset from the Bucaramanga segment.

(Syracuse et al., 2016)
Application II Iran and event relocation effects

Can joint inversion methods for tomography improve travel time predictions and hence event locations?
We explore a full range of regularization and data weights.
Application II Iran and event relocation effects

- multiple-event relocation of 2006 M6.1 Silahour earthquake sequence by Ghods et al., 2013
- all earthquakes and events within 100 km of centroid are removed from 3D velocity inversion
- after relocation of target earthquakes, misfits are:
  - 7.12 km for ak135
  - 5.95 km for TT-only model
  - 5.85 km for joint model
Geophysical imaging & the machine learning era

1. Cluster analysis (unsupervised ML) for tectonic regionalization
   ✓ Fast estimates of first-order structure variation

2. Automatic waveform grading (supervised ML)
   ✓ Significant reduction of analyst time

3. 2D image reconstruction with Deep Learning (supervised ML)
   ✓ Significant noise reduction
   ✓ Near-real time processing
   ✓ Higher-resolution images
1. CLUSTER ANALYSIS FOR TECTONIC REGIONALIZATION
Clustering of teleseismic receiver functions

- Blue: recomputed from He et al., 2014
- Red: newly computed

- 1900 seismic stations
- Few techniques to explore receiver functions from dense networks
- Detailed analysis takes much time
- OUR APPROACH: use ML and investigate only representative receiver functions
Receiver Function Smoothing

(after Chai et al, 2015)
Hierarchical

K-medians

K-means

Spectral
The Ps converted phase arrives earlier for eastern clusters, and last for those from Tibet.
Within each cluster the waveforms are similar.
Correlation to Crust Thickness

The spatial distribution of the clustering correlate with crust thickness estimates from previous active-source results.
2. AUTOMATIC WAVEFORM GRADING
Assume 5 sec per waveform, it takes almost 56 work days to label this dataset.

Data courtesy of M. Cleveland (LANL) and J. Kintner (PSU)
Classification Strategy

- Seismograms
- Preprocessing
- Features
- Machine Learning
- Labels
- Examine Results
- Visualization

Using 100 trees, the accuracy is about 92.4% for the testing data.
Comparing Different Algorithms

All algorithms are improving when more data are included.

Random forest and neural network outperform others when large amounts of data are used.

ANN: Artificial Neural Network (7 layers)
RF: Random Forest
KNN: K-Nearest Neighbors
LR: Logistic Regression
SVM: Support Vector Machine (linear)

(with help from Jingyi Luo)
Using probability as output instead of labels, we can leave low probability waveforms to further inspections.
Classification Results
Event 2015-09-18T15:59:42Z

Legend
- Correctly Labeled
- Incorrectly Labeled
- Unclear

Google Earth
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
US Dept. of State Geographer
Image Landsat / Copernicus
© 2018 Google

View from Space (Altitude: 6213 mi)
3. IMAGE RECONSTRUCTION WITH DEEP LEARNING
Above the surface
Image Inversion techniques for Non-Destructive Evaluation

• Image physical characteristics of small and large specimens
  – High resolution (1 cm or better)
  – Fast reconstructions
    • From seconds to minutes
  – Actual physical values
    • Acoustic speed, attenuation coefficients, etc.
Ultrasound Reconstruction (50 kHz)

SAFT: Synthetic Aperture Focusing Technique
SAFT Artifacts

- Assumes narrow band transmitted signal
- Stitched SAFT reconstruction
- Crosstalk wave artifacts
- Low resolution
- Reverberation
- Noise
- Stitching artifacts
- Shadowing
Model-Based Iterative Reconstruction (MBIR)
MBIR + Machine Learning (ML)

- Present ML-based techniques improve the prior model $p(x)$
  - Dictionary learning
  - De-noising

- Other ML-based techniques generate higher quality reconstructions from simpler models
  - For example, predict MBIR reconstruction from a SAFT-generated image

- Our current effort is to define the physics in $f(x)$ from actual data
K-Wave Synthetic Experiments

- Deep Learning (DL) algorithm
  - Trained a modified U-net architecture

\[ x = A^{-1} y \]
K-Wave Synthetic Results
Empirical Experiment

- **Specimen**
  - 7ft x 7ft x 3 ft cement block
  - 20 defects are inserted in the block

- **Sensor**
  - Commercial 10-transducer MIRA system

- **Samples**
  - Over 300 vertical and horizontal scans of specimen
Empirical Results

Defect Diagram  MBIR  Joint DDL

Precision vs Recall Curves

Area Under the Curve

<table>
<thead>
<tr>
<th>Method</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBIR 2.5D</td>
<td>0.2908</td>
</tr>
<tr>
<td>MBIR</td>
<td>0.2836</td>
</tr>
<tr>
<td>SAFT</td>
<td>0.1323</td>
</tr>
<tr>
<td>UTSR</td>
<td>0.1932</td>
</tr>
<tr>
<td>DDL</td>
<td>0.3447</td>
</tr>
</tbody>
</table>
Final Remarks

• Addressing current challenges in seismic imaging with simultaneous joint inversion of disparate geophysical observations

• Incorporating Machine Learning techniques into different aspects of imaging:
  – Tectonic regionalization by clustering the interpolated receiver function wavefield
  – Automatic and fast waveform grading with random forest and neural networks
  – Prediction of 2D object image with Deep Learning
    • Linear operator maps time series to image format
    • No iterative framework -> FAST!
    • Actual physical parameter estimation
Final Remarks

• Currently working on DL-based approach that predicts $\Delta x$ instead of $x$
  – Prediction of optimization gradients
  – Allows for iterative framework

• Currently working on application of DL-based approach to seismic data and imaging.