Open access seismic data with scripts for processing with open software

Karl Schleicher
University of Texas at Austin
There is a big gap between functioning research prototype and a tested program. You learn a lot when you start processing field data. In industry, processing groups can help. They provided data expertise including selecting a suitable dataset, previous results, detailed parameters, partially processed data for input, and an eye to evaluate our new results. This paper starts to build a data library for testing open source seismic software.
Long Term Goals

• Build an open-access seismic library with scripts for open software processing.
• The library can be used by others to recreate my processing.
• The scripts provide detailed processing sequences and parameters that can be used with or without modification.
• Accelerate testing and validation of new seismic algorithms.
Long Term Goals (continued)

- Multiple datasets suitable for testing different research efforts (2d, 3d, land, marine, noise, multiples, sampling, field data, synthetic data, etc).
- Evaluate the relative strengths of open-access seismic software.
- Improve open-access seismic software.
- Make Madagascar the “go to” place for seismic test data.
Previous open seismic processing

- Provide instruction for basic unix, user environments, basic processing, and advanced SU scripting.
  - Demos directory in SU distribution
  - Seismic Data Processing with Seismic Un*x, Forel, Benz, and Pennington
  - Geophysical Image processing with Seismic Unix, Stockwell

- This effort's focus:
  - basic processing scripts for full processing sequence: load, trace header creation, velocity filter, velocity analysis, moveout, and stack.
  - Field data access
<table>
<thead>
<tr>
<th>Dataset</th>
<th>copied</th>
<th>SU processing</th>
<th>Madagascar</th>
<th>In svn</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska 31-81 2d Land</td>
<td>y</td>
<td>y</td>
<td>partial</td>
<td>y</td>
<td>Easy statics Ground roll spikes</td>
</tr>
<tr>
<td>Alaska 635-79 2d Land</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td>Hard statics</td>
</tr>
<tr>
<td>Alaska 51S-75 2d Land</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td>Intermediate statics</td>
</tr>
<tr>
<td>Nankai 2d Marine</td>
<td>y</td>
<td>y</td>
<td>partial</td>
<td></td>
<td>deep water</td>
</tr>
<tr>
<td>Taiwan 2d Marine</td>
<td>y</td>
<td>started</td>
<td></td>
<td></td>
<td>strong random noise</td>
</tr>
<tr>
<td>Usp 2d Land</td>
<td></td>
<td>partial</td>
<td></td>
<td></td>
<td>Partial</td>
</tr>
<tr>
<td>Teapot dome 3d Land</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opunake 3d marine</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Processing NPRA Line 31-81

- Background information about the data
- Data Loading and initial data QC
- Shot record velocity filter
- Shot record edit, mute, cdp sort
- Velocity interpretation
- Stack
- Comparison of the SU and 1981 processing
NPRA Background

Line 31-81 Background

- Line 31-81 selected because it was short line from last year data was collected. Approximate location is marked.
- Data is 96 trace, 12 fold, dynamite. 440 ft shot interval, 110 ft receiver interval.
- Previous processing included spherical divergence correction, velocity filtering, designature, agc, velocity estimation, nmo, residual statics, diversity stack, time variant filter, and agc.
Data loading and initial QC

• Segyread converted the data from SEGY to SU format.
• Surange showed only fldr and tracf in the input trace headers.
• Suximage helped identify the first 10 record to be test data. There is good signal. There is ground roll.
• Xmovie display identifies data channels 1-96, aux channels 97-101.
• First impressions of the data:
  - Good signal. Ground roll, noisy trace segments, repeated ground roll on center traces, “spikes”, small statics.
  - Suitable for testing “land noise attenuation”.


Seismic Unix (SU) processing

- Data load, header dumps, and initial qc plots
- Data observations:
  - Ground roll, noisy traces, spikes.
  + Good signal, small statics.
- Headers contain only flnd and tracf. Custom program to load headers
- Agc, decon, shot record velocity filter, mute, cdpgather
- Velocity analysis (a long script)
- Residual statics
- Stack (compare with previous results)
- Post stack migration
Initial view of data. Do not process first 10 records (test records). Each record has 5 aux channels. Scattered spike traces. Good signal.
Data loading and initial QC:

- Line 31-8.

Groundroll

Good signal

Aux traces
Data loading and initial QC

- Noisy traces
- Repeated Ground roll?
- Noisy trace segment
- Aux traces 97-101
Data loading and initial QC

- I wrote a custom java program to load the headers. The observer log describes the (FFID,EP) relationship. The elevations are in the surveyor's log. I typed this data into two text files. I dumped the ffid, tracf headers (sugethw), assembled the data from these three textfile with a custom java program, and loaded the headers (sugethw).

- Dataload program translated to python.

- Stack section from previous processing
  - loaded with straight forward segyread.
  - Surange showed a couple of headers that required update (I found many segy filed required f1, d1, f2, d2 to be zeroed).
Data loading and initial QC
Velocity filter, CMP gather

- I used sudipfilter to apply a shot record FK filter to remove ground roll.
- I split the positive and negative offsets and used an asymmetrical dip filter (-15,5).
- I developed a looping script to separate the records and apply FK filter.
- I also applied spreading correction, mute, and ags in the same script. I looked at the results using sumovie.
- I removed a bad shot record, muted, and sorted the data to cmp. I used sumovie to view the cmp's.
Data loading and initial QC

Shot record with agc and velocity filter
Data loading and initial QC

Two twelve fold cmp
Velocity Interpretation and stack

- I used a long script that combined several SU commands. This significantly improved Forel et. al. It combined the capabilities of iva.sh and velanQC.sh.
- The velocity is plotted on the semblance and the CVS plot. It can be edited on the semblance plot.
- Gather is plotted with and without NMO.
- After updating velocity the plot is recreated with the new velocity function.
- I found these upgrades were enough to provide a minimal velocity interpretation capability.
Velocity Interpretation and stack
Stack

- I used the velocity field to stack the data. I applied decon in the same script.
- Differences with 1981 processing
  - GSI designature leave a different wavelet on the data than decon.
  - AGC application was different. GSI brackets velocity filter with AGC/inv AGC.
  - GSI used diversity stack
  - GSI applied residual statics
Comments on SU processing

- Considering all the differences, the results are surprisingly similar.
- The 1981 result is better above 400 ms.
- SU software is hard to use:
  - Sudipfilt was not intended for prestack FK filter, so a looping script was developed.
  - The programs do not trap user errors.
  - SU's primary domain has been software prototyping.
Madagascar results

- William has honors for best section. Challenge now is to match his results with a faster script.
Bashkardin – fk analysis
Bashkardin – automatic velocity picking show same lateral variation as manual picking
Bashkardin – stack without scaling
Burnett – fp analysis
Burnett – edit and fp example
Burnett – initial stack without agc or edits
Burnett – stack with edits and fp filter
Burnett – median stack without agc, edits, or fp filter
Scripts are available

- SConstruct are in svn. $RSFSRC/book/data/alaska
- There are three directories:
  - Line31-81 the su scripts
  - bash Bashkardin's Madagascar
  - wburnett-31-81 Burnett's Madagascar
- The Sconstruct downloads the data from usgs and processes and creates displays
- Still pretty dynamic
Strengths and Weakness of Open Software

- Still early days for me, so these comments are very preliminary.
  - Basic programs (nmo, mute, edit) tend to be neglected. No one in a University receives recognition for working on these.
  - In the 1980's the universities were tops in interactive display. There has been little improvement in the packages since.
  - Hale has advocated Java and Python for a years. I think Java and Python will slowly become adopted by industry. I say slowly after observing industry adoption of c and c++. Memory problems (leaks, wild pointers) slow you down in c and c++. Java and Python solve some of these problems. JTK from Mines and Madagascar are already on board.
Strengths and Weakness of Open Software

- Madagascar
  - Steep learning curve. I am still learning basic Madagascar and struggling with tools like scons, svn, latex, and reproducible documents.
  - Seems to be improving.

- SU
  - Based on older software tools (c, make).
  - Does not seem to be improving.
  - Data is just a bunch of traces, so it is slow to extract a subset of a large dataset.

- SEPLIB, Freeusp, DDS, cpseis, javaSeis
  - I have barely scratched the surface. I did not see data viewer or velocity interpretation.
Strengths and Weakness of Open Software

- OpenDdetect
  - Interpretation tool. I briefly worked on the tutorial and it appears this is an interpretation tool. I do not think it is a good tool to review your processing results.
Identified Datasets

- Alaska land line 31-81 (more lines available)
- Avenue data set from University Utah
- Land line from freeUSP “how to process tutorial” previously loaded to Madagascar
- Dragon Land3d
- Teapot Dome Land 3D
- Mobil AVO Viking Graben 2D Line 12
- Nankai 2D deep water line NT62-8 from Seismic Data processing with Seismic Un*x, Forel, et. al.
- Taiwan 2D line from Seismic Data processing with Seismic Un*x, Forel, et. al.
- Teal South from UT
Identified Datasets

- UTIG has several data sets (mostly marine?). Need to contact DeAngelo. Stratton 3D?
- Stockwell's lab project 1-14 already ported to Madagascar.
- CSM summer camp data.
- Blake Ridge data set from USGS.
- William Burnett has one migrated line from 3D Duran Ranch project from GXT. May have restrictions.
- New Zealand Marine 3D available for media cost.
- Synthetics not main objective, but could include:
  - SEG Salt Model, SEAM, Overthrust model, various BP models (tomography salt dome,
Nankai Deep Water 2D

- SU processing sequence
  - Data load, header dumps, and initial qc plots
  - Tx and Fx display of selected records
  - Resample, t**2 gain, cdp gather.
  - Near trace gather display
  - Velocity analysis
  - Stack
  - Post stack migration
- Data observations:
  - good signal.
  - Is velocity variation due to cable feathering?
Near trace gather
Shot 1707 TX and FX
Stack
Stacking velocity
Phase shift migration
Kirchhoff migration
Taiwan 2D Marine

- SU processing sequence
  - Data load, header dumps, and initial qc plots
  - Shot displays before and after lc filter
  - Near trace displays before and after lc filter
Shot 900 before and after lc filter
Near trace gather (before and after lc filter)
Interactive Display Options

- I think we need better seismic visualization programs.
- Possible starting points:
  - VTK (Visualization Tool Kit)
    - Paraview
    - Visit
    - mayavi2
  - Colorado School of Mines Java Tool Kit (JTK)
  - BotoSeis
Conclusions

- Several datasets identified. Scripts started for three.
- Processing sequence is very basic. Each dataset has its own issues (noise, statics, etc).
- The library is in the development version of Madagascar. The directory is $RSFSRC/book/data
- Alaska line 31-81 is the most mature. There are three directories in $RSFSRC/book/data/alaska
- SU is the most mature open geophysical software system. I have found it tricky to use. I think there is a bug in migkt2d.
Future Direction

- Continue to develop basic SU processing scripts
- Compare SU and Madagascar results
- Contribute program improvements to SU
- Reproducible documents describing data and scripts
- Present this paper at The PTTC Workshop - Open Software Tools for Reproducible Computational Geophysics.