Data access in the Seis Lab

Geophysical Institute at UAF

Michael West

All examples can be found on the Seismology Linux Network in:
/home/west/SeisLab/data_access

Examples current as of October 2011

starters

Why bother with a database?

+ Not a lowest common denominator tool
+ The harder the problem, the more a database counts
+ Antelope is a power tool, not always novice-friendly (the Photoshop or ARC GIS of seismology)
+ Strength in numbers. Two dozen Antelope users at GI

Assembling the right database

+ Mix / match: [stations] + [waveforms] + [hypocenters]
+ Making a good “descriptor file” is key

Things to consider in a descriptor file

+ Time span of your waveform database (.wfdisc)?
+ Time span of your hypocenter database?
+ Do you have station info for waveforms and hypocenters (phase arrivals)?

example 1

Task: View the magnitude 5.2 earthquake that occurred on September 28, 2011 at 20:13

Skill: Browsing continuous waveform data

Data: Alaska master_stations +
Alaska waveforms (1 day) +
Automated AEIC earthquake solutions (1 day)

example 1

Create a descriptor file (3 lines)
"points" to relevant database tables elsewhere
{stations} + {waveforms} + {events}

File name: mydb_2011_09_28

```
#
schema css3.0
dbpath /aerun/sum/params/Stations/{master_stations}:/aerun/sum/
db/archive/archive_2011/{archive_2011_09_28}:/aerun/op/run/
dbseg/Quakes_2011/{Quakes_2011_09_28}
```

View the database tables. View the waveforms

```
Wolverine:> dbe mydb_2011_09_28

Wolverine:> smartpick mydb_2011_09_28
```

Scroll to 20:13 (or) select orid 28174 on smartpick panel

example 2

Task: Visualize Denali fault aftershock sequence

Skill: Creating a custom database

Data: AEIC Total earthquake catalog

example 2

Subset the hypocenters by time and location. Join accompanying information location error.

```
Wolverine:>   dbsubset /Seis/catalogs/aeic/Total/Total.origin \
   'time>"11/01/2002" && time<"11/05/2002"' | \
   dbsubset - 'lat>62   && lat<65  && lon>-151 && lon<-141' | \
   dbjoin - event netmag origerr | \
   dbunjoin -o denali –

Wolverine:>  db2kml -oel denali > googleearth_denali.kml
```

<- create new subsetted database

<- export to kml

Open .kml file in Google Earth …

Tip: *long commands can be split across multiple lines with the \ character.*

```
wolverine>  dbsubset denali.origin 'ml>3.0' | dbunjoin –o bigevents –

wolverine>  dbsubset denali.origin 'ml>3.0' | \
                dbunjoin –o bigevents –
```

example 3

Task: Make text output for a tomography code

Skill: Exporting derived data from a database

Data: Mt. Spurr volcano campaign dataset recorded summer 2005. Database already includes relevant stations, events and waveforms.

example 3

Subset for preferred “hypocenters”

+ join phase arrivals and station information
+ subset for P wave arrivals
+ sort them by source-receiver distance
+ write out custom fields

```
wolverine:> dbjoin /home/admin/databases/SP05/origin/SP05.origin event |\
    dbsubset - 'orid==prefor' |\
    dbjoin - assoc arrival site |\
    dbsubset - 'iphase=="P"' |\
    dbsort - 'distance(origin.lat,origin.lon,site.lat,site.lon)' |\
    dbselect - sta site.lat site.lon  origin.lat origin.lon origin.depth 'arrival.time-origin.time'


…
OPT     59.6532 -153.2299   61.3197 -152.2572   0.4100    31.7000000476837
KNK     61.4125 -148.4556   61.3197 -152.2572   0.4100    33.4199998378754
CNP     59.5259 -151.2348   61.3197 -152.2572   0.4100    32.7200000286102
HOM     59.6583 -151.6432   61.5605 -151.9432   3.5600    34.2599999904633
…
```

or pipe results into a file …

```
wolverine:> dbjoin …   … > spurr_traveltimes.txt
```

example 4

Task: Extract SAC files of the 3/27/2010 Maule, Chile earthquake

Skill: Extracting specific waveform data

Data: Alaska master_stations +
Alaska waveforms (1 day) +
PDE global earthquake catalog (1 year)

example 4

Create appropriate descriptor file

File name: pde_2010_03_27

```
#
schema css3.0
dbpath /aerun/sum/params/Stations/{master_stations}:
/aerun/sum/db/archive/archive_2010/{archive_2010_02_27}:
/Seis/catalogs/pde/{pde_2010}
```

Extract segmented waveforms from this database:

```
Wolverine:>  trexcerpt -m event -c 'chan=="BHZ"' -s 'ms>8.0' -o ic  \
   -j 'deg2km(distance(site.lat,site.lon,64.87,-147.86))<500' -de \
   -w DATA_%{orid}/%Y%m%d_%H%M%S_%{net}_%{sta}_%{chan}.sac \
   pde_2010_03_27 dbMaule 'parrival()-300' 3600
```

List the SAC files

```
wolverine:> ls DATA_2363 | more
20100227_064745_AV_PTPK_BHZ.sac
20100227_064746_AV_BMR_BHZ.sac
20100227_064746_AV_DAWY_BHZ.sac
…
```

example 4

trexcerpt deconstruction

```
wolverine:>  trexcerpt -m event -c 'chan=="BHZ"' -s 'ms>8.0' -o ic
    -j 'deg2km(distance(site.lat,site.lon,64.87,-147.86))<500' -de
    -w DATA_%{orid}/%Y%m%d_%H%M%S_%{net}_%{sta}_%{chan}.sac
    pde_2010_03_27 dbMaule 'parrival()-300' 3600
```

Explanation

Extract data from all BHZ channels within 500 km of UAF, for event(s) with ms>8.0, saved as sac files with PC byte order (as opposed to Sun), in directories with the specified file names, from database pde_2010_02_27, beginning 5 minutes before the predicted P wave, lasting 30 minutes. New data will be “wrapped” database called dbMaule.

---

trexcerpt is extremely powerful. See the man page.

example 5

Task: Read travel time information into Matlab

Skill: Reading database fields
(comparable to example 3 but in Matlab)

Data: Mt. Spurr volcano campaign dataset recorded summer 2005. Database already includes relevant stations, events and waveforms.

example 5

Open, subset, join database tables. Close database pointer.

```
db = dbopen('/home/admin/databases/SP05/origin/SP05','r');
db = dblookup(db,'','origin','','');
db1 = dblookup(db,'','event','','');      db = dbjoin(db,db1);
db = dbsubset(db,'orid==prefor');
db1 = dblookup(db,'','assoc','','');      db = dbjoin(db,db1);
db1 = dblookup(db,'','arrival','','');    db = dbjoin(db,db1);
db = dbsubset(db,'iphase=="P"');
db1 = dblookup(db,'','site','','');       db = dbjoin(db,db1);
dbquery(db,'dbRECORD_COUNT')
[siteLat siteLon arrivalTime] = dbgetv(db,'site.lat','site.lon','arrival.time');
[originLat originLon originTime] = dbgetv(db,'origin.lat','origin.lon','origin.time');
dbclose(db);
```

Extract segmented waveforms from this database:

```
>> whos
```

Name	Size	Bytes	Class
arrivalTime	5753x1	46024	double
originLat	5753x1	46024	double
originLon	5753x1	46024	double
originTime	5753x1	46024	double
siteLat	5753x1	46024	double
siteLon	5753x1	46024	double

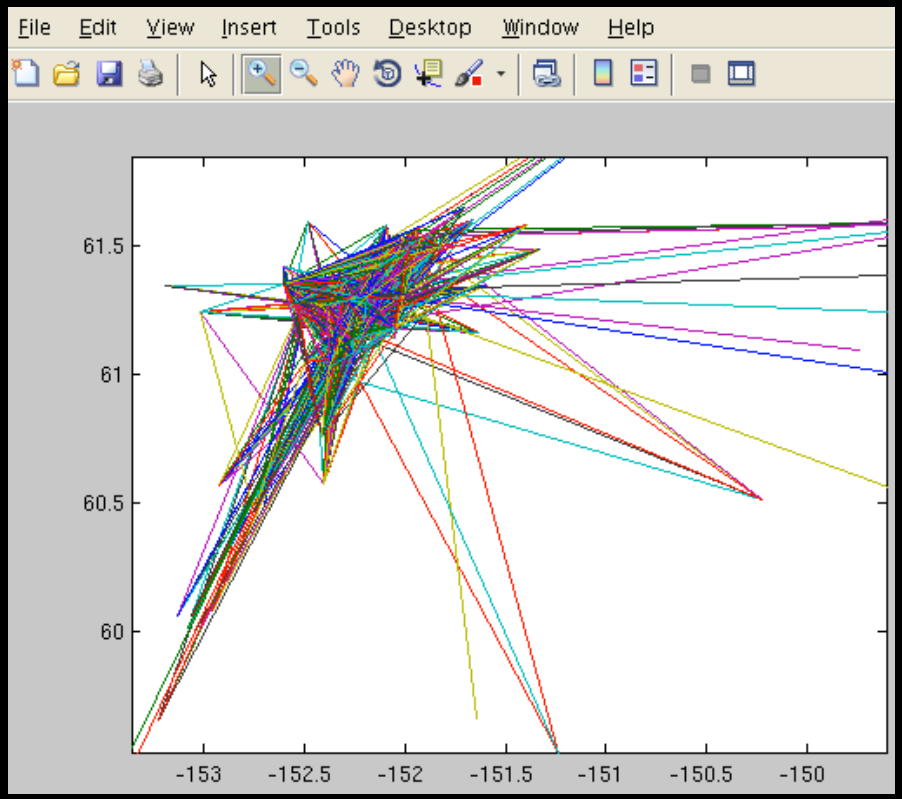
example 5

Change times from epoch to Matlab time format

```
arrivalTime = epoch2datenum(arrivalTime);
originTime = epoch2datenum(originTime);
```

Get on with your research

```
>> plot( [siteLon originLon]' , [siteLat originLat]' );
```



example 6

Task: Extract mag. 4.5 Kodiak earthquake waveforms for Mt. Spurr broadband station SPCP

Skill: Importing waveform data
(Comparable to example 4, but in Matlab)

Data: Alaska waveforms

example 6

Specify the source of waveform data

```
>> datasource('uaf_continuous')
ans =
      type: ANTELOPE
  location: /aerun/sum/db/archive/archive_[YEAR]/archive_[YEAR]_[MONTH]_[DAY]
```

Select sta_chan_net_loc

```
>> scnl=scnlobject( ‘SPCP’ ,{‘BHZ’ ‘BHN’ ’BHE’},'','')
```

Get waveforms for the stated time range

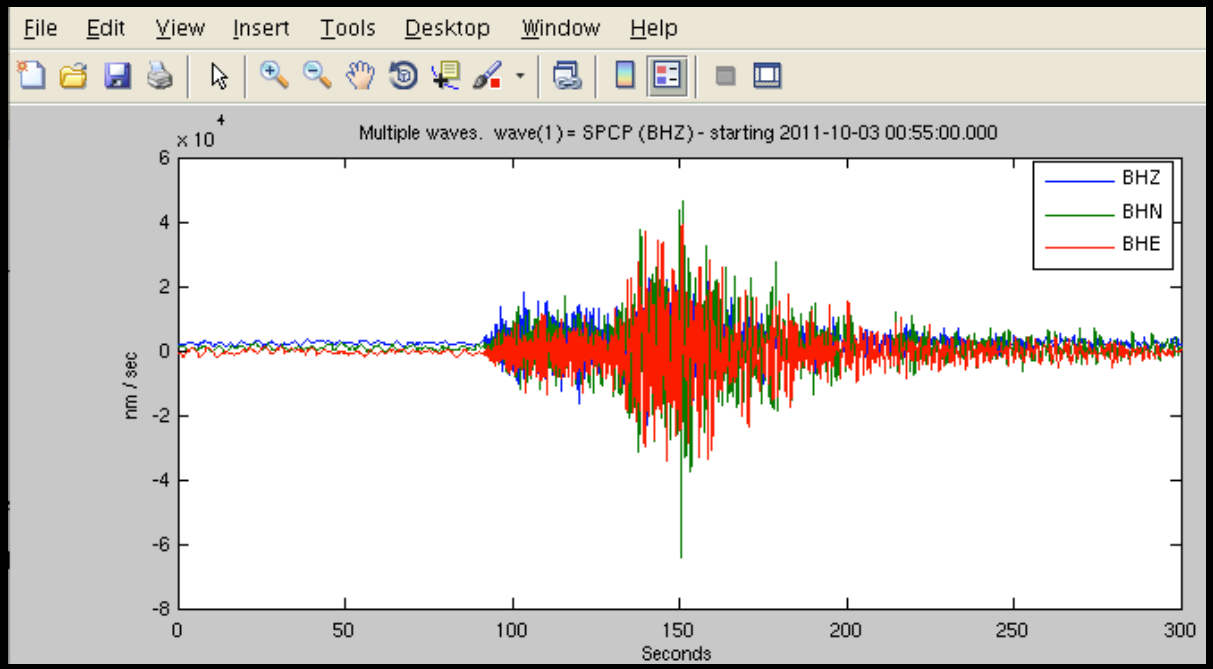
```
>> w = waveform( ds , scnl , ’10/03/2011 00:55’ , ’10/03/2011  00:59’ );
```

datasource, scnlobject and waveform are in-house tools. See help pages

example 6

Get on with your research …

```
>> plot(w)
>> legend(w)
```



See Seismo. Res. Lett. paper for details of waveform object
http://giseis.alaska.edu/input/west/papers/2011_srl_reyes_waveformSuite.pdf

wrap up

Pointers and lessons learned:

A good descriptor file is worth the effort!

Nearly all routines are documented in man pages!

- file:///opt/antelope/5.0-64/antelope.html
- see “Program man pages”

Programming intro in later chapters of Datascope book

- see “Antelope Seismic Information System (Datascope)”

Lab colleagues are your best resource!