PH5: From Field to Archive

Outline

Introduction to PH5
Tools Available In Field
From Raw Data to Archive
The PH5 Web Interface at the DMC
PH5 In The Future
Questions and Answers

What is PH5?

PH5 is a data archiving format based on the Hierarchical Data Format version 5 originally developed by the National Center for Supercomputing Applications.

A single format that can hold both raw data and metadata from a variety sources and output it in SEG-Y, SAC, and other future data formats on the fly.

Why PH5?

Freely Available and Open Source
Extensible and Future Proof
Fast and Reliable
Full Data Set is Archived
Metadata and Data Are Fully Separated
Gathers Can be Quickly Generated (and Regenerated)

What Needs to be Archived

All of Your RAW Data
Experiment Metadata
Station/Receiver Metadata
Source Metadata
Report

Experiment Metadata

PI Names
Institutions
Assembled ID
Network Code
Experiment Name and Nickname
Brief Description of the Experiment
Bounding Coordinates of the Experiment

Station/Receiver Metadata

Station ID
Datalogger Serial Number
Channel Number
Array Number
Deploy Time
Pickup Time
Latitude, Longitude, and Elevation

*Possibly Optional: Datalogger type, sensor type, sensor serial

Source Metadata

Source ID
Latitude, Longitude, Elevation
Source Time

Optional (Depending on Source Type)
Source size
Unit of Measure for Size(kg, lbs, etc)
Depth
Top of Column
Bottom of Column

Field Tools: The Kitchen Suite

The Kitchen Software Suite includes:

PForma: Raw Data to PH5
Experiment_t-gen: Experiment Description Metadata
Noven: Format Station and Event Metadata in KEF format
Geod2kef: Calculate Source-Receiver Offsets
Time-kef-gen: Texan Time Correction
PH5View: Quickly Look at Gathers
PH5toseg and PH5tosac: Output Your Data

Kitchen Suite Requirements

Linux or Mac (CentOS 7 Recommended)
Modern Multi-Core Processor Highly Recommended
Python 2.7
HDF5
Epel-release-noarch (if using CentOS)
PIP 6.0.8 or Higher (if using CentOS)

Complete list of requirements at:
http://jasper.passcal.nmt.edu/wiki/

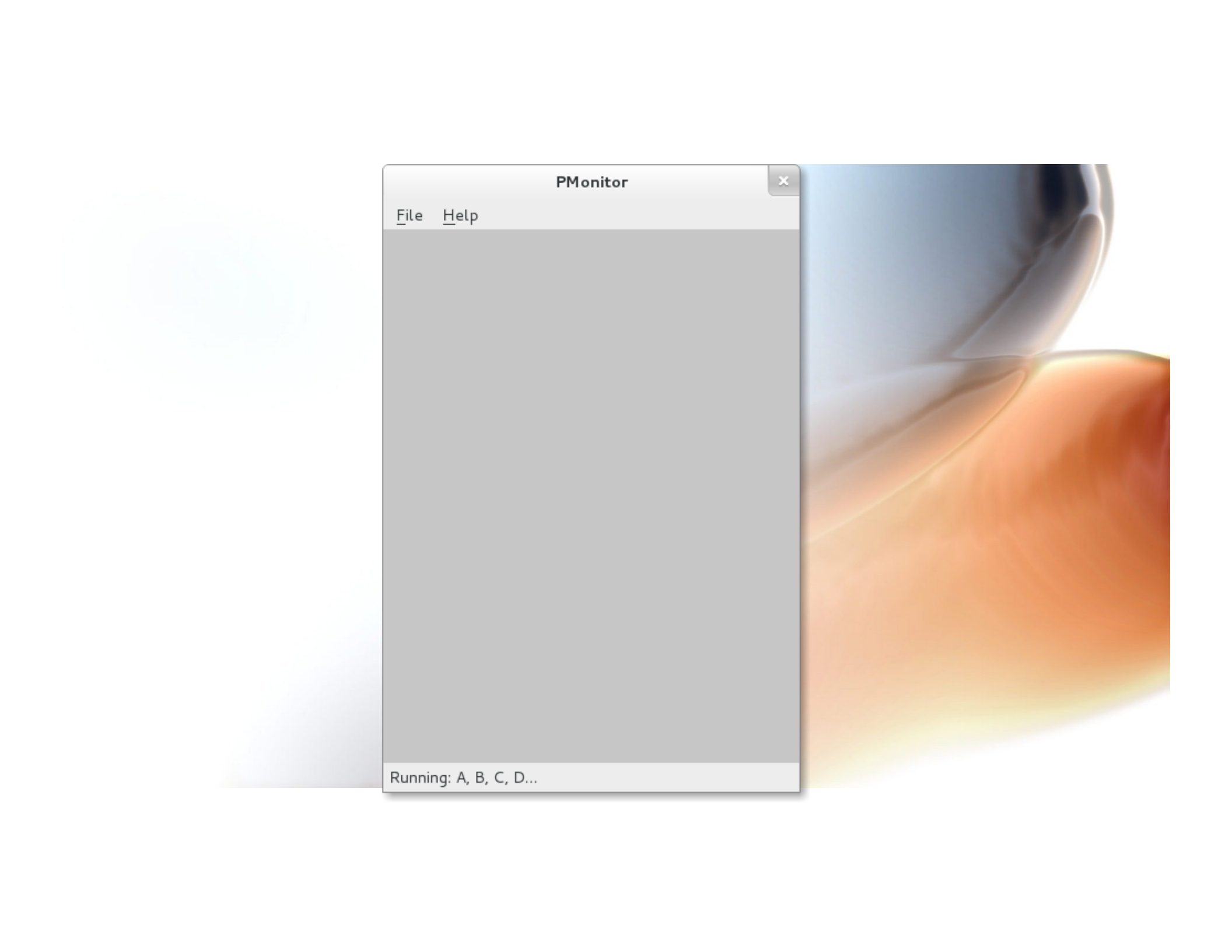
Switching Gears:
Creating a PH5 Archive in the Field

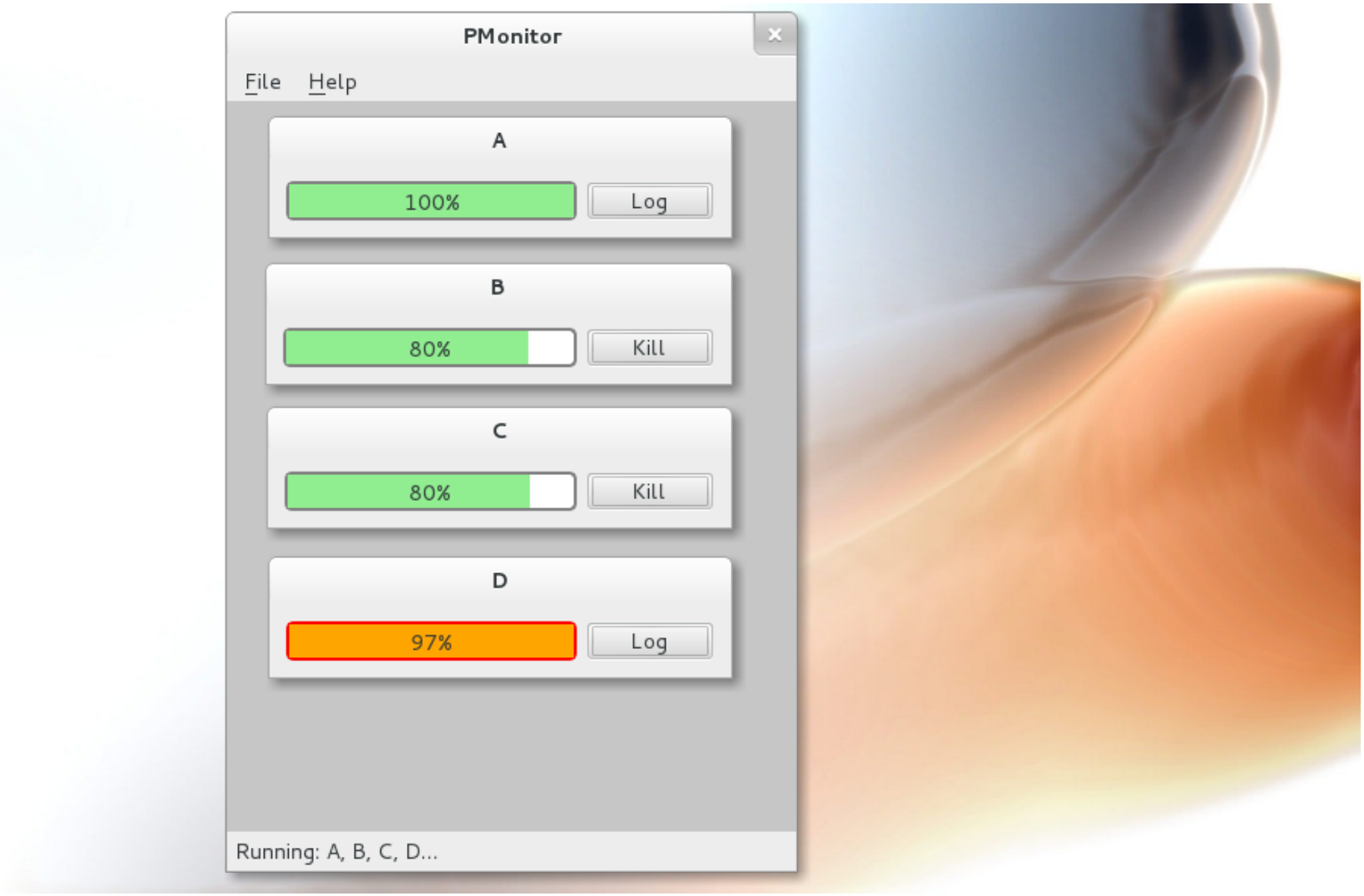
Step 1: PForma

- Create list of Raw files

```
[root@localhost Raw_Data]# ls -d -1 $PWD/*.* > trd_list
```

Note: There are many ways to create a list of raw files. Choose your favorite, but ensure that each line contains the full path to the file.


PMonitor
File
Help
Running: A, B, C, D...


PMonitor
File
Help
A
100%
Log
B
80%
Kill
C
80%
Kill
D
97%
Log
Running: A, B, C, D...


PMonitor
File
Help
A
100%
Log
B
80%
Kill
C
80%
Kill
D
97%
Log
Running: A, B, C, D...
Errors D
File
Processing: /run/media/azevedo/Data/SUGAR_RAW/I3928RAW2014079.TRD 13928 =*** Opening: miniPH5_00001 done
Processing: /run/media/azevedo/Data/SUGAR_RAW/I2983RAW2014079.TRD 12983 =*** Opening: miniPH5_00001 done
Processing: /run/media/azevedo/Data/SUGAR_RAW/I2627RAW2014079.TRD 12627 =*** Opening: miniPH5_00001 done
Processing: /run/media/azevedo/Data/SUGAR_RAW/I2471RAW2014079.TRD 12471 =Updating external references...done,
191 nodes recreated.
Closing remaining open files:./miniPH5_00001.ph5...done
done
Done
Process finished but not all files processed: 40 of 41.

KEF Files: Kitchen Exchange Format

A KEF file is a plain text file that is used to load metadata into PH5.

Tools exist to create the various KEF files you will need to create a complete PH5 archive.

```
/Experiment_g/Sorts_g/Event_t
        id_s = 5001
        description_s =
        time/ascii_s = Mon Nov 24 04:25:00 2014
        time/epoch_l = 1416803100
        time/micro_seconds_i = 0
        time/type_s = BOTH
        location/X/value_d = 21.86500666
        location/X/units_s = degrees
        location/Y/value_d = -18.41244766
        location/Y/units_s = degrees
        location/Z/value_d = 995.702
        location/Z/units_s = meters
        location/coordinate_system_s = geodetic
        location/projection_s = none
        location/ellipsoid_s = WGS84
        location/description_s = 5001
        size/value_d = 1000
        size/units_s = lbs
        depth/value_d =
        depth/units_s = meters
```

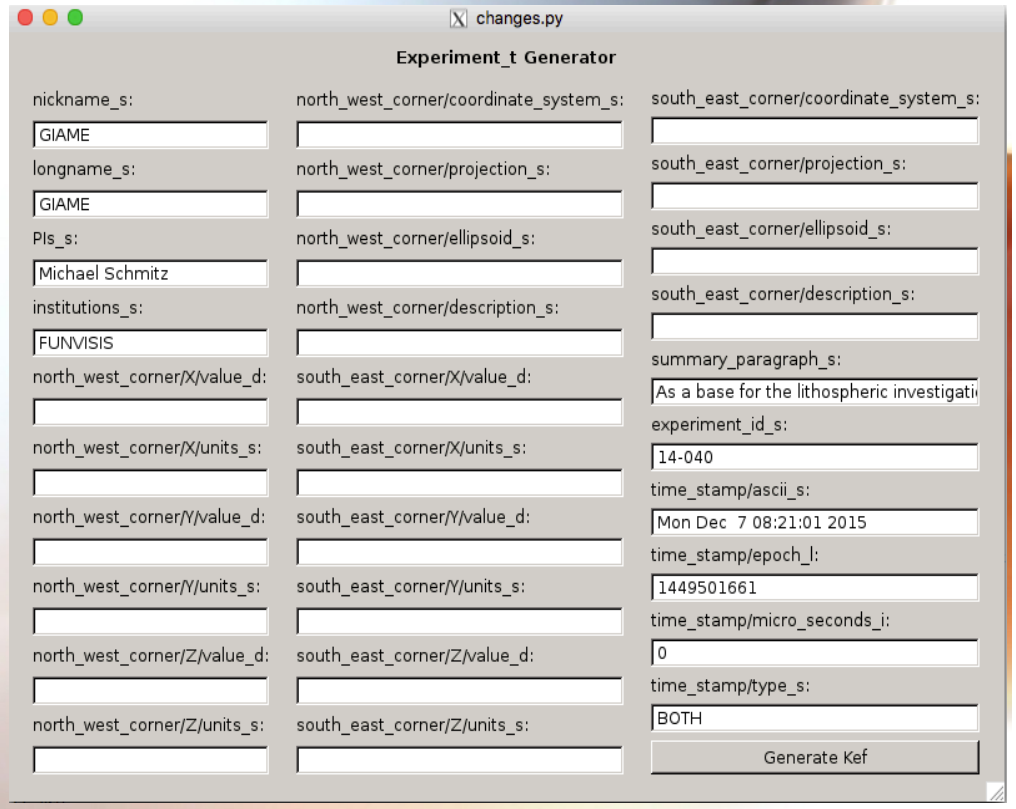
Experiment_t-gen

Generate Experiment Metadata

Save as kef file

Run kef2ph5 to Load Generated kef

```
[field@dhcp-25 ~]$ kef2ph5 -n master -p /path/to/ph5 -k experiment_t.kef
```



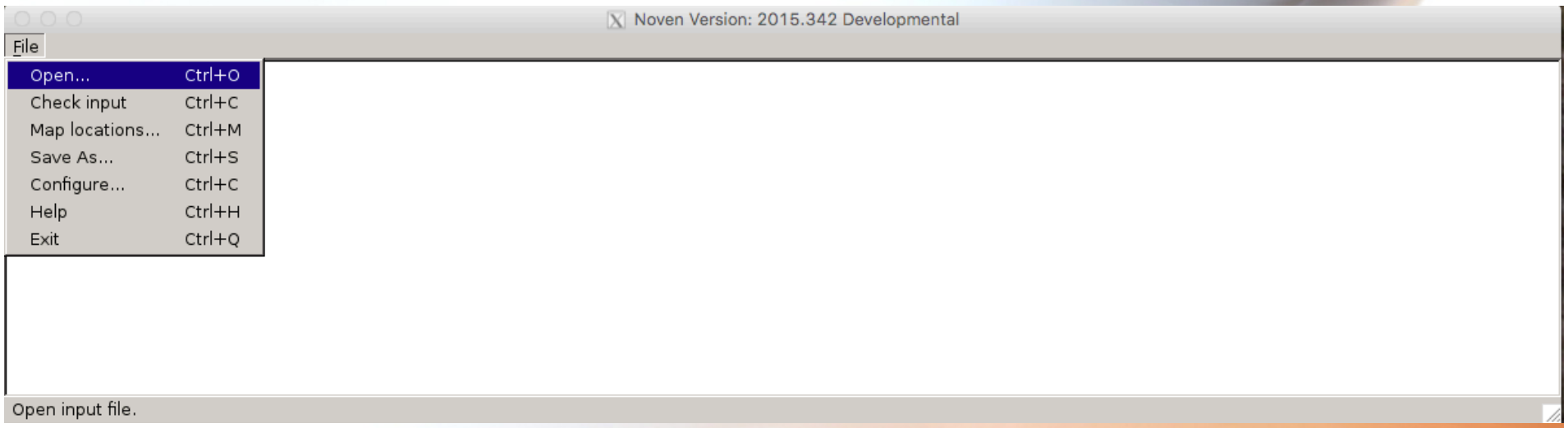
Receiver and Shot Metadata

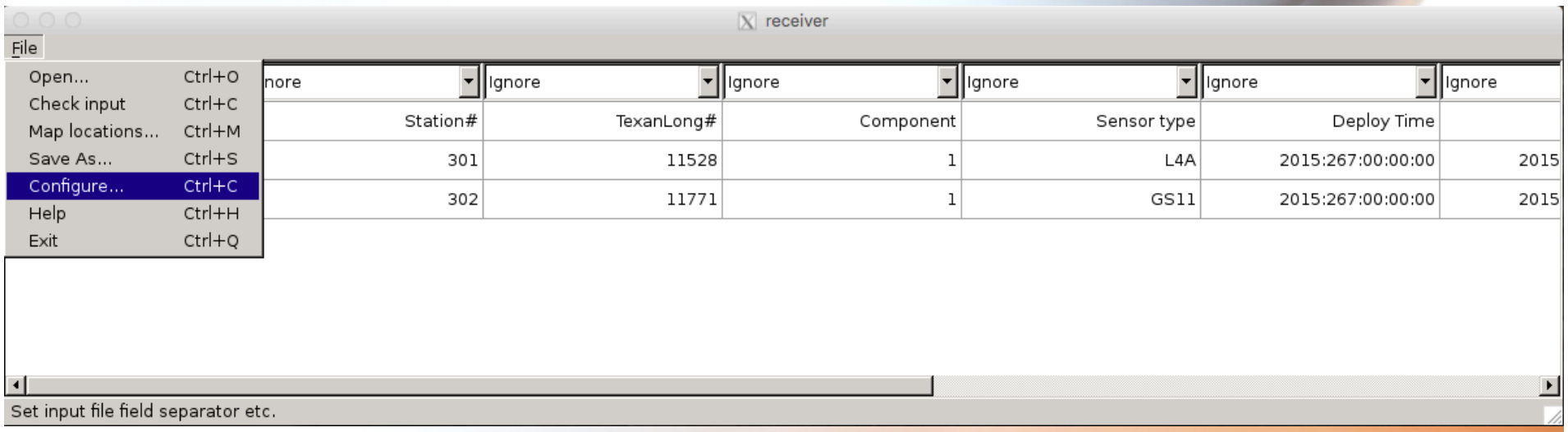
Receiver and Shot Metadata are Saved in CSV Format

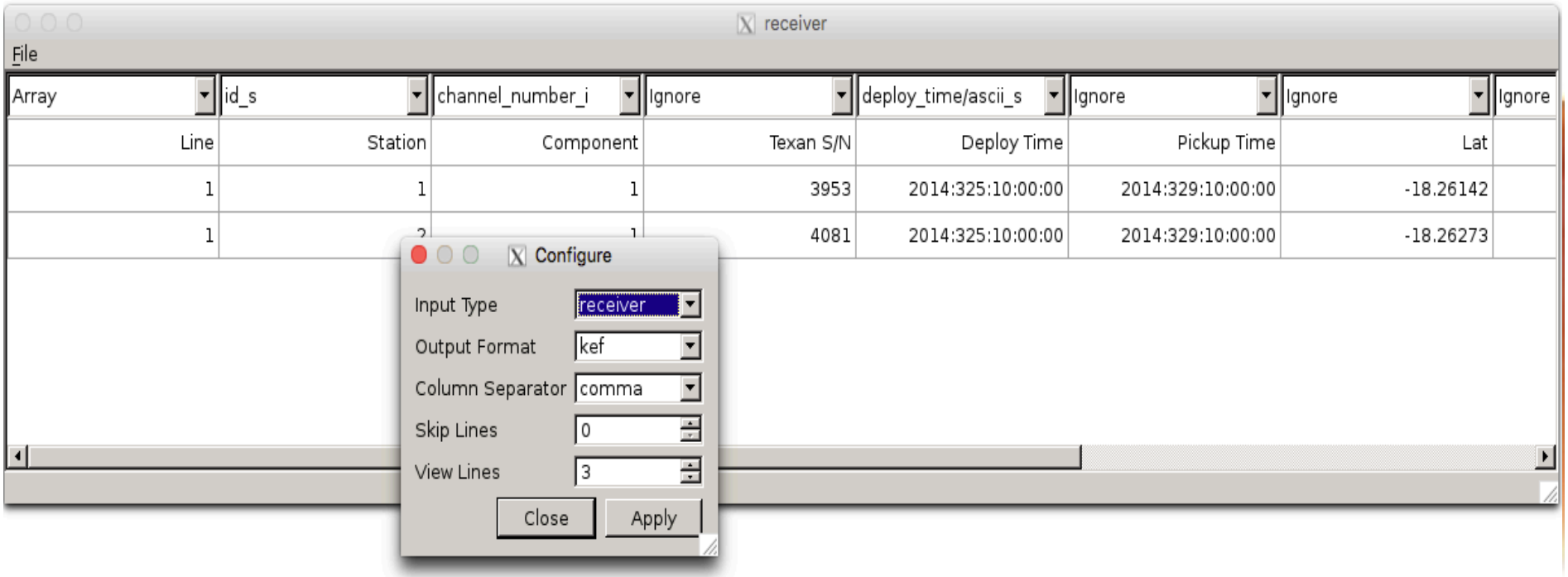
station	serial	lat	lon	elevation	Channel	Array	Deploy	Pickup
1001	12997	8.64808	-71.85546	57	1	1	2015:320:00:00.00	2015:321:00:00.00
1002	13919	8.64698	-71.85546	56.81	1	1	2015:320:00:00.00	2015:321:00:00.00
1003	11891	8.64625	-71.85493	56.61	1	1	2015:320:00:00.00	2015:321:00:00.00
1004	13971	8.64549	-71.85435	56.33	1	1	2015:320:00:00.00	2015:321:00:00.00
1005	14094	8.64478	-71.85378	55.86	1	1	2015:320:00:00.00	2015:321:00:00.00
1006	12945	8.64409	-71.85323	55.43	1	1	2015:320:00:00.00	2015:321:00:00.00
1007	12394	8.64329	-71.85276	55.51	1	1	2015:320:00:00.00	2015:321:00:00.00
1008	11590	8.64212	-71.85288	54.84	1	1	2015:320:00:00.00	2015:321:00:00.00
1009	12545	8.64043	-71.85354	55.2	1	1	2015:320:00:00.00	2015:321:00:00.00
1010	12870	8.63791	-71.85376	55.55	1	1	2015:320:00:00.00	2015:321:00:00.00

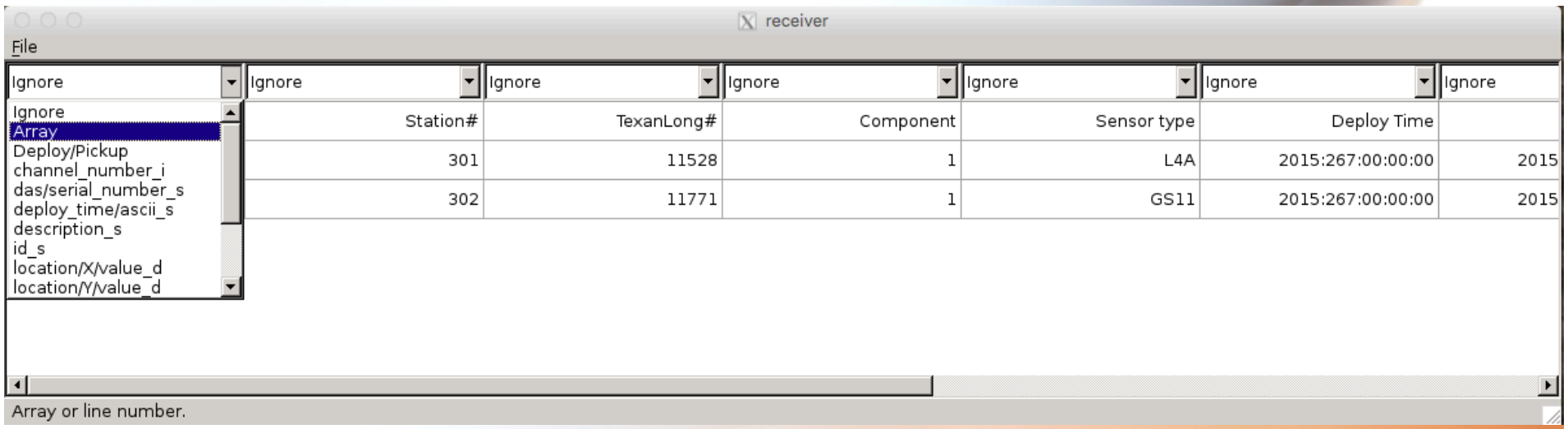
Shot	lat	lon	elev	time	size	unit	depth
5202	33.80847	-83.55728	225	2015:219:03:20:00.020	273	kg	16
5203	33.69766	-83.49372	194	2015:219:07:20:00.020	182	kg	18
5204	33.45043	-83.39242	172	2015:219:05:50:00.020	727	kg	22
5205	33.31253	-83.29161	177	2015:219:08:51:00.020	182	kg	17
5206	33.11601	-83.12384	147	2015:220:03:00:00.020	182	kg	19
5207	32.83597	-82.89922	102	2015:220:04:10:00.020	182	kg	17
5208	32.69918	-82.92459	75	2015:220:05:10:00.020	182	kg	19
5209	32.47656	-82.80255	59	2015:220:06:18:00.020	364	kg	21
5210	32.35993	-82.71048	55	2015:220:05:05:00.020	182	kg	12
5212	31.84319	-82.42969	68	2015:220:03:05:00.020	182	kg	19

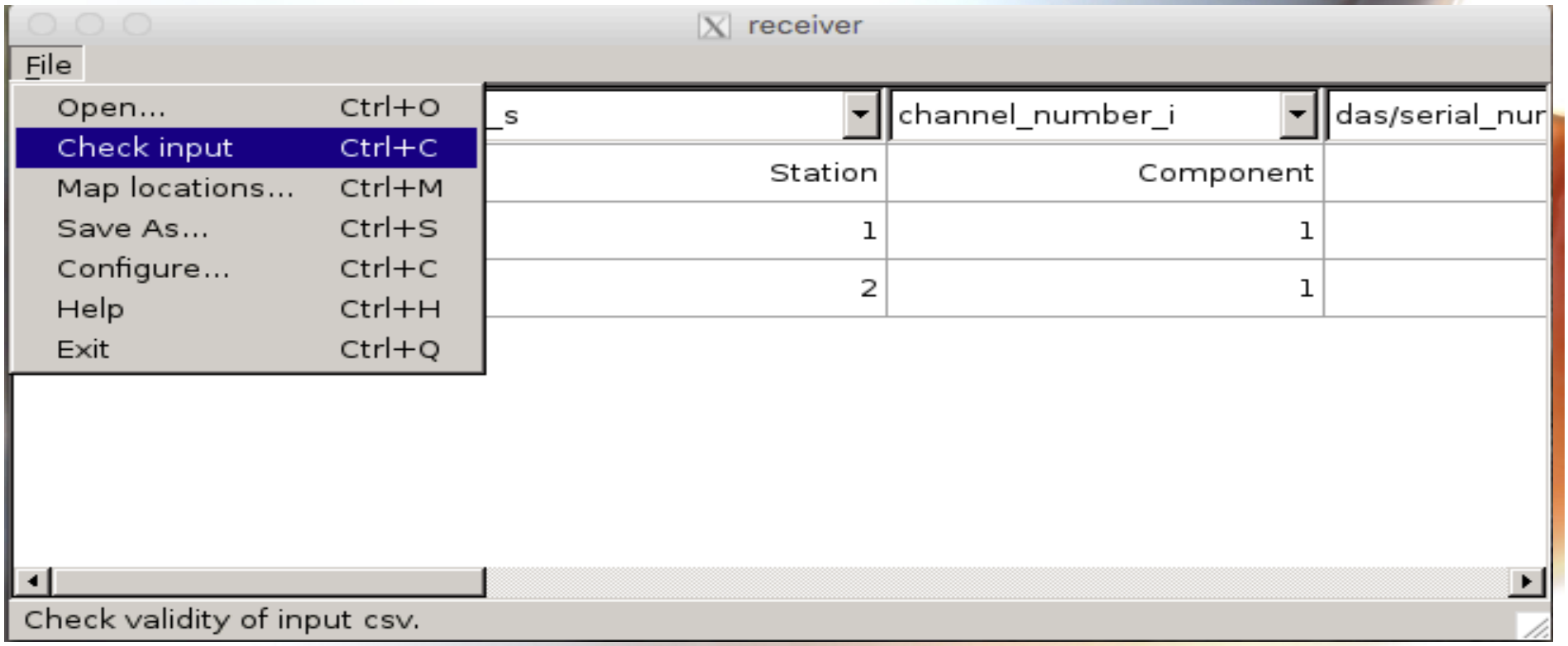
Noven

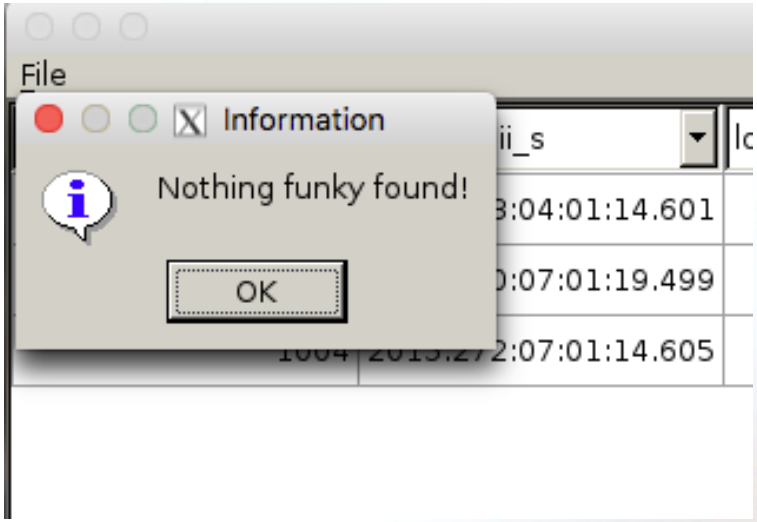

Noven Version: 2015.342 Developmental
File
Open... Ctrl+O
Check input Ctrl+C
Map locations... Ctrl+M
Save As... Ctrl+S
Configure... Ctrl+C
Help Ctrl+H
Exit Ctrl+Q
Open input file.


receiver
File
Open... Ctrl+O
Check input Ctrl+C
Map locations... Ctrl+M
Save As... Ctrl+S
Configure... Ctrl+C
Help Ctrl+H
Exit Ctrl+Q
nore
Ignore
Ignore
Ignore
Ignore
Ignore
Station#
TexanLong#
Component
Sensor type
Deploy Time
301
11528
1
L4A
2015:267:00:00:00
2015
302
11771
1
GS11
2015:267:00:00:00
2015
Set input file field separator etc.


receiver
File
Array
id_s
channel_number_i
Ignore
deploy_time/ascii_s
Ignore
Ignore
Ignore
Line
Station
Component
Texan S/N
Deploy Time
Pickup Time
Lat
1
1
1
3953
2014:325:10:00:00
2014:329:10:00:00
-18.26142
1
4081
2014:325:10:00:00
2014:329:10:00:00
-18.26273
Configure
Input Type
receiver
Output Format
kef
Column Separator
comma
Skip Lines
0
View Lines
3
Close
Apply


receiver
File
Ignore
Ignore
Ignore
Ignore
Ignore
Ignore
Ignore
Ignore
Array
Deploy/Pickup
channel_number_i
das/serial_number_s
deploy_time/ascii_s
description_s
id_s
location/X/value_d
location/Y/value_d
Station#
TexanLong#
Component
Sensor type
Deploy Time
301
11528
1
L4A
2015:267:00:00:00
2015
302
11771
1
GS11
2015:267:00:00:00
2015
Array or line number.


receiver
File
Open... Ctrl+O
Check input Ctrl+C
Map locations... Ctrl+M
Save As... Ctrl+S
Configure... Ctrl+C
Help Ctrl+H
Exit Ctrl+Q
_s
channel_number_i
das/serial_nur
Station
Component
1
1
2
1
Check validity of input csv.



Save the kef and Run:

```
field$ kef2ph5 -n master -p /path/toph5 -k Array_1_t.kef
```

```
field$ kef2ph5 -n master -p /path/toph5 -k Event_t.kef
```

Note: You need to run kef2ph5 for every array in your experiment

Geod2kef

Used to calculate the offset distance between sources and receivers. Geod2kef creates an offset kef file to be loaded into the PH5 Archive

```
field$ geod2kef -n master -p /path/to/ph5 > offset_t.kef
```

```
field$ kef2ph5 -n master -k offset_t.kef
```

Time-kef-gen

Time-kef-gen is used only for Texan(RT125A) RAW data. Time-kef-gen will calculate the timing-drift corrections for the RT125A RAW data and create a kef file to be loaded into the PH5 Archive.

```
field$ time-kef-gen -n master -p /path/to/ph5 > time_t.kef
```

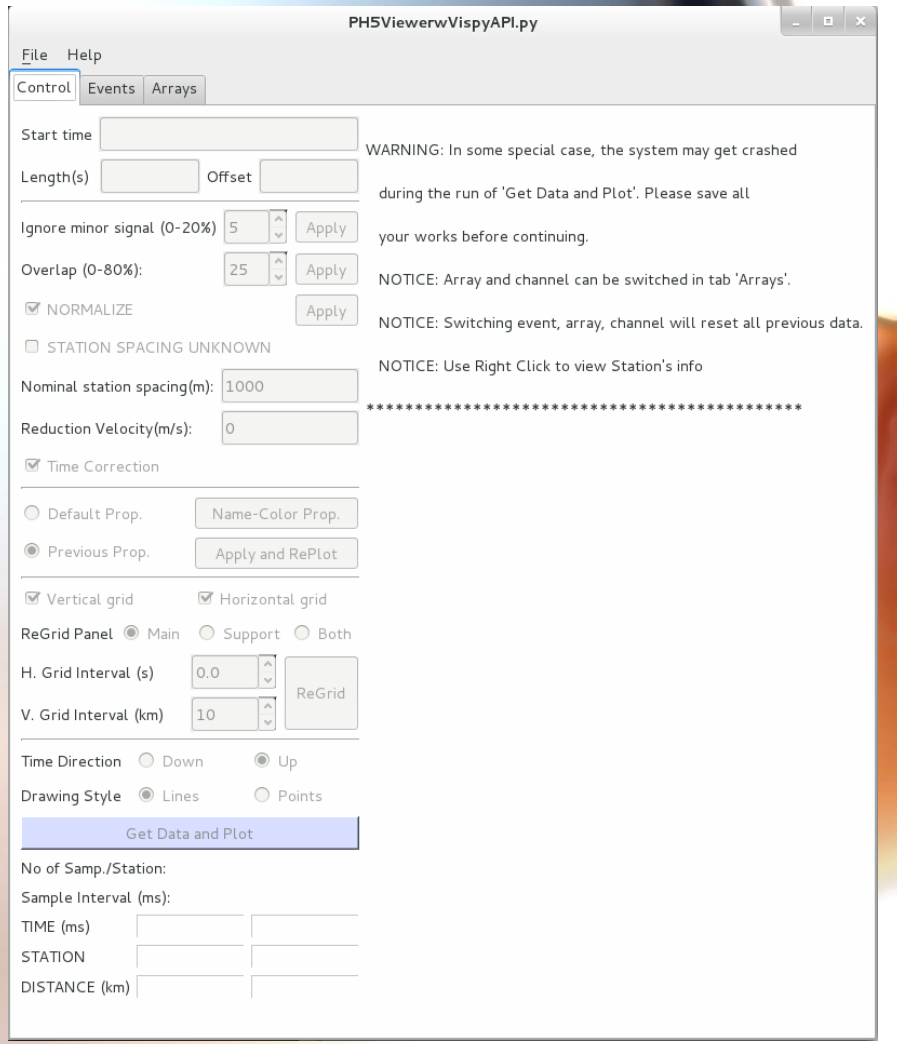
```
field$ kef2ph5 -n master -k time_t.kef
```

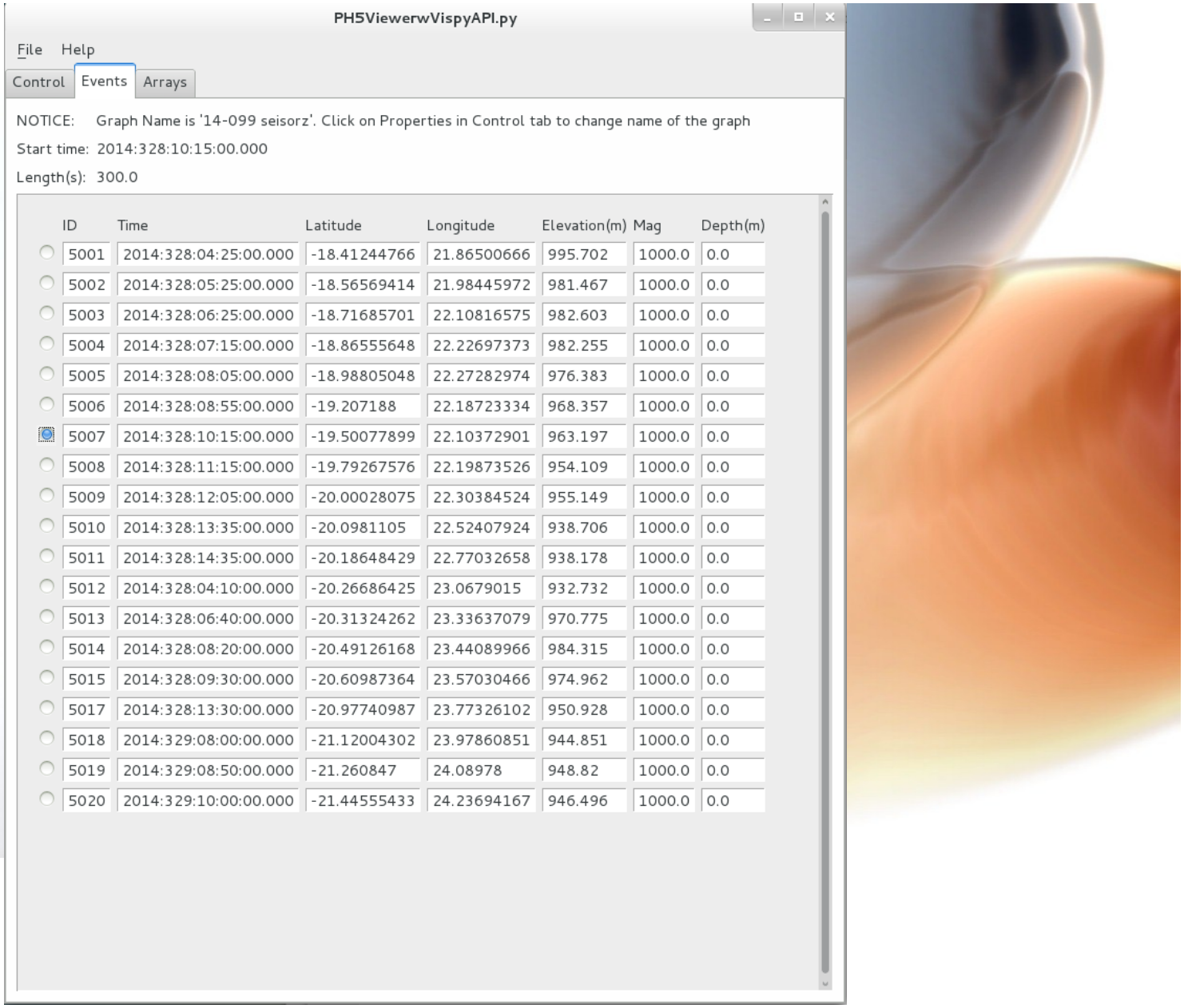
PH5View

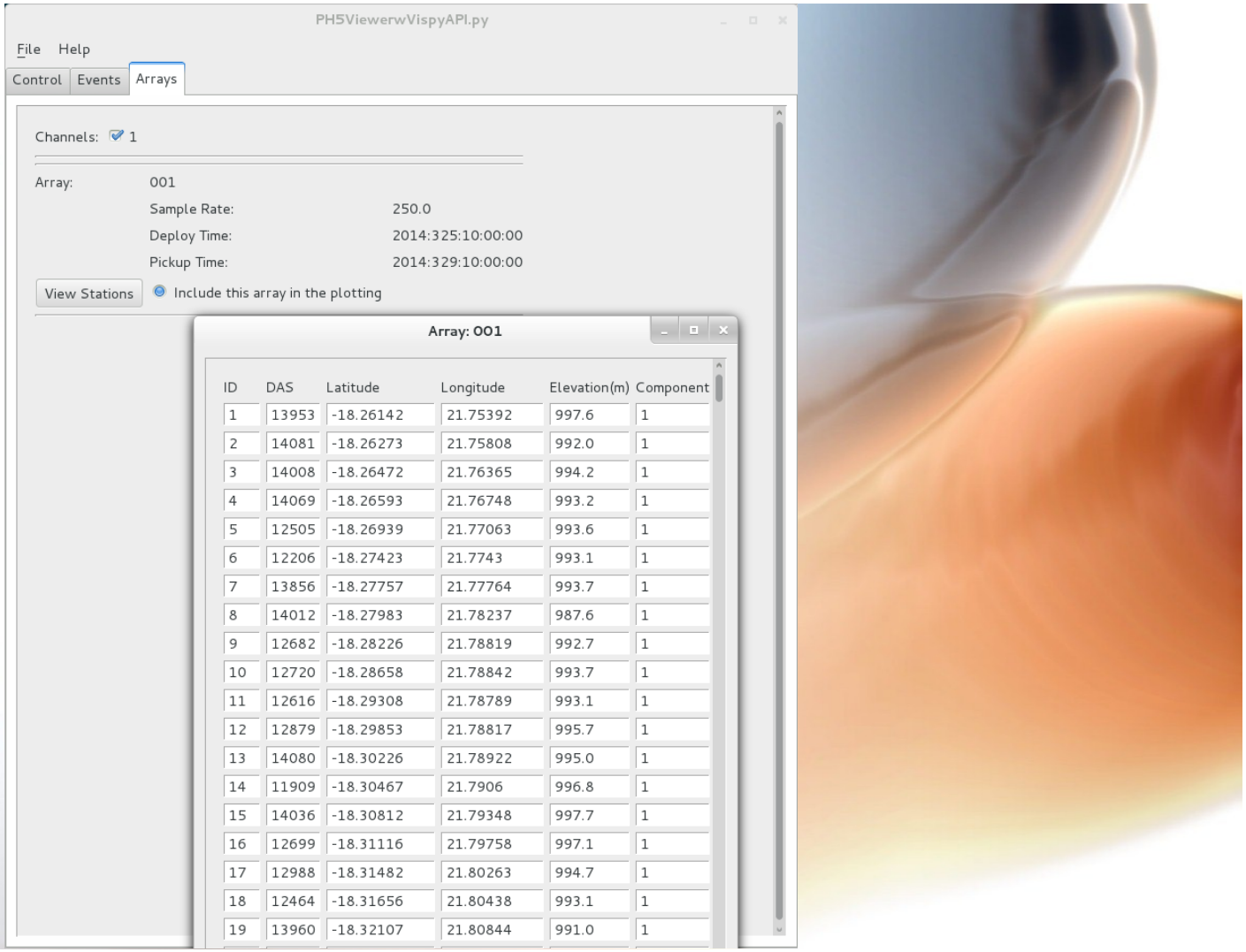
Easily Check Metadata

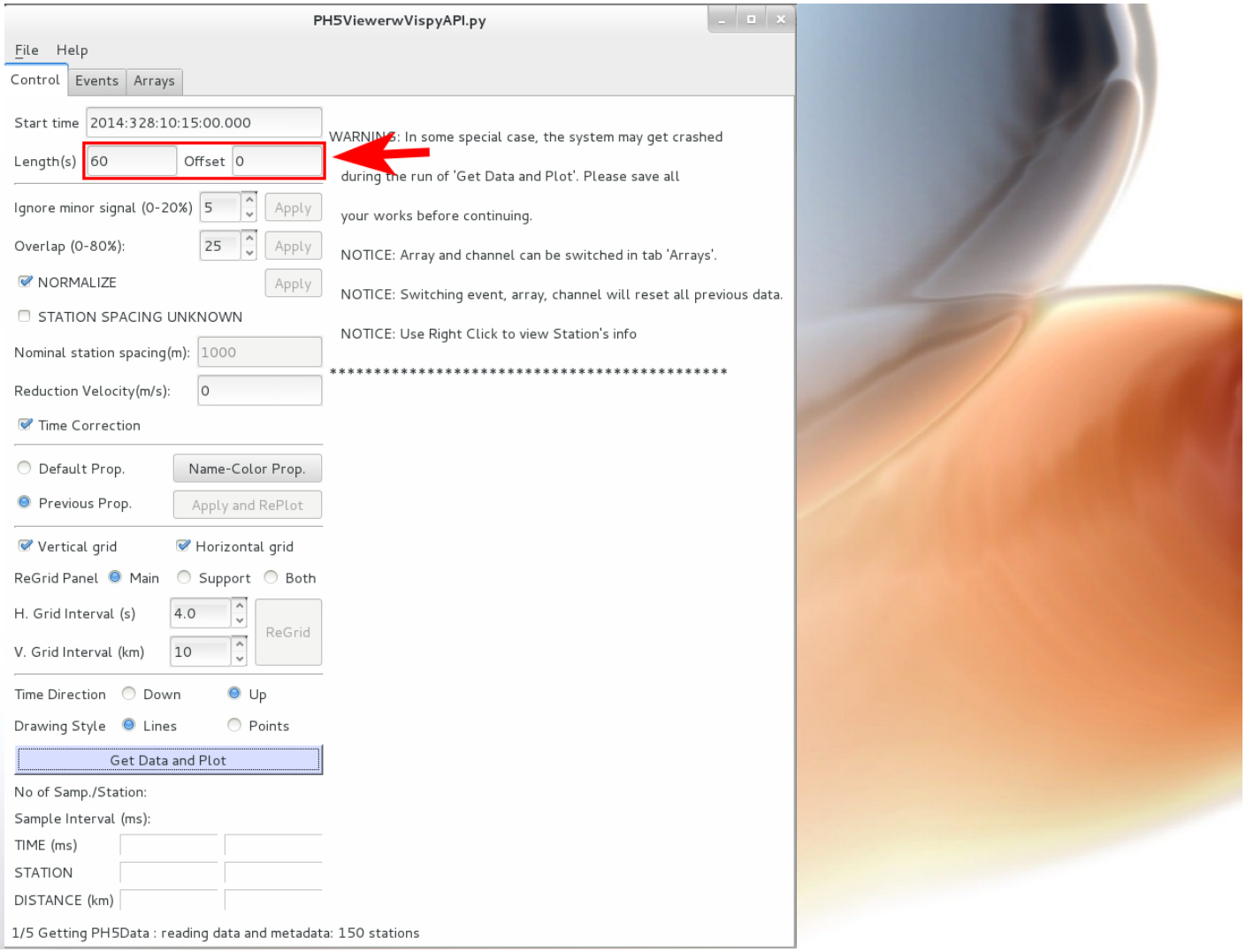
Quickly View Shot Gathers

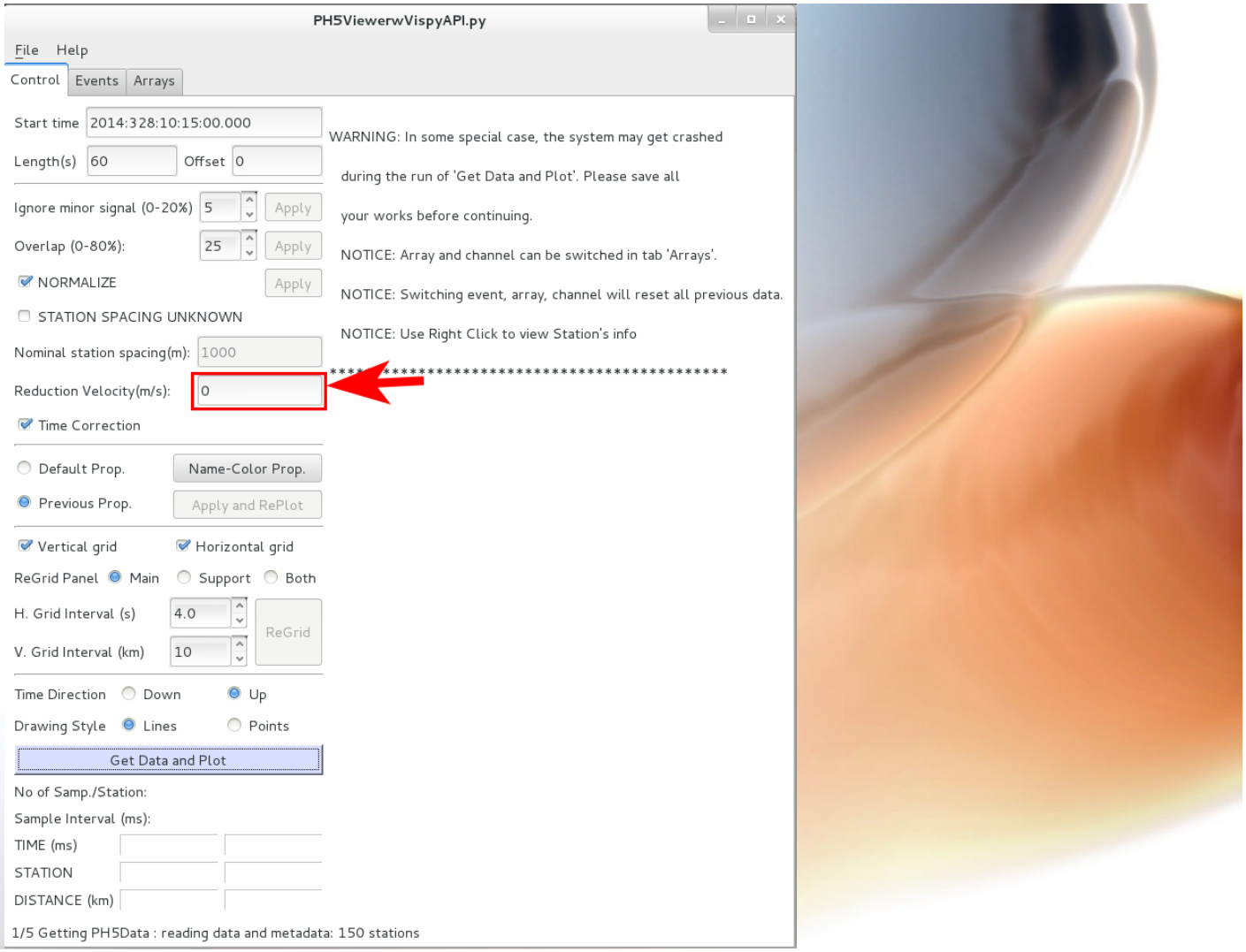
Fast Basic QC Before Creating SEG-Y

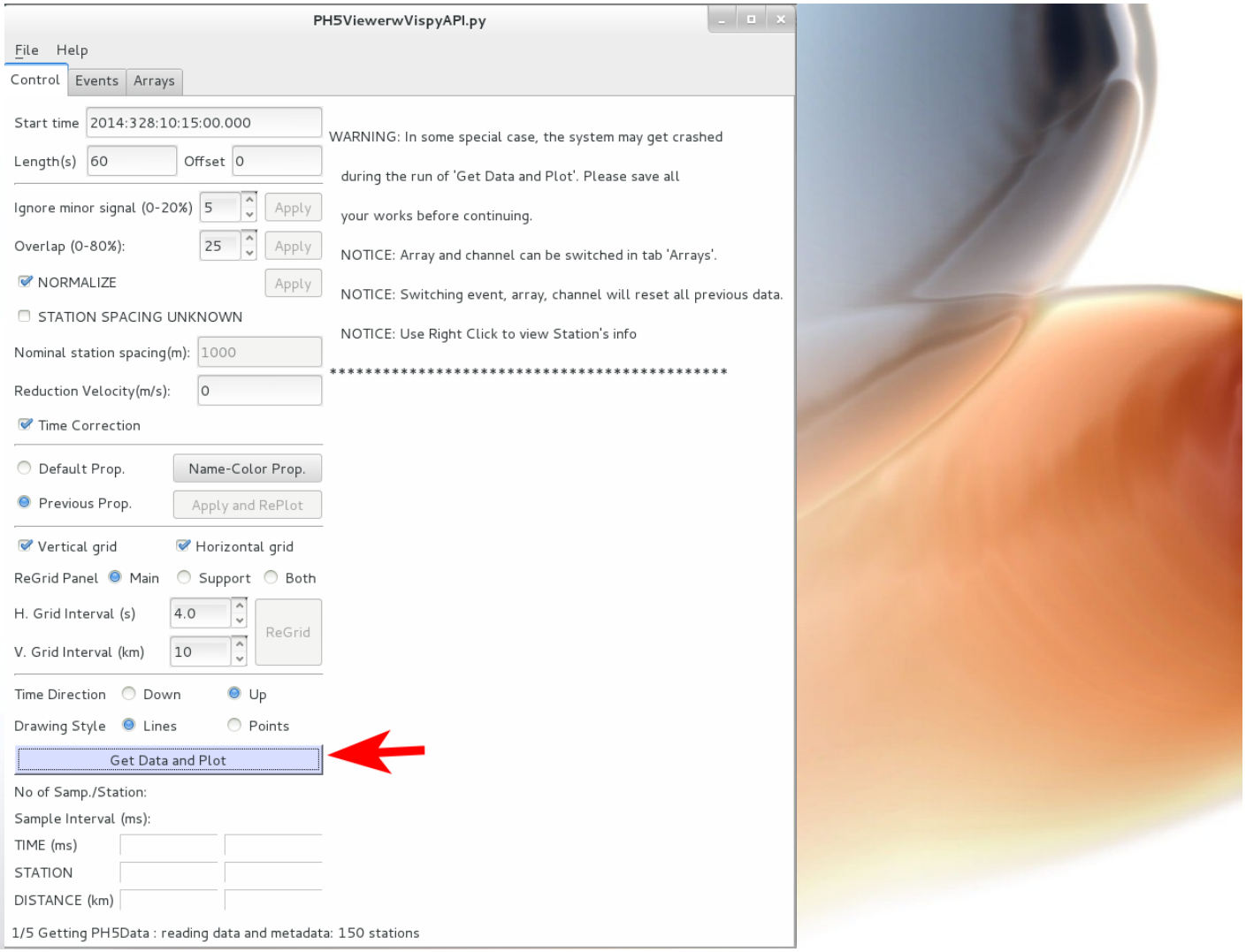


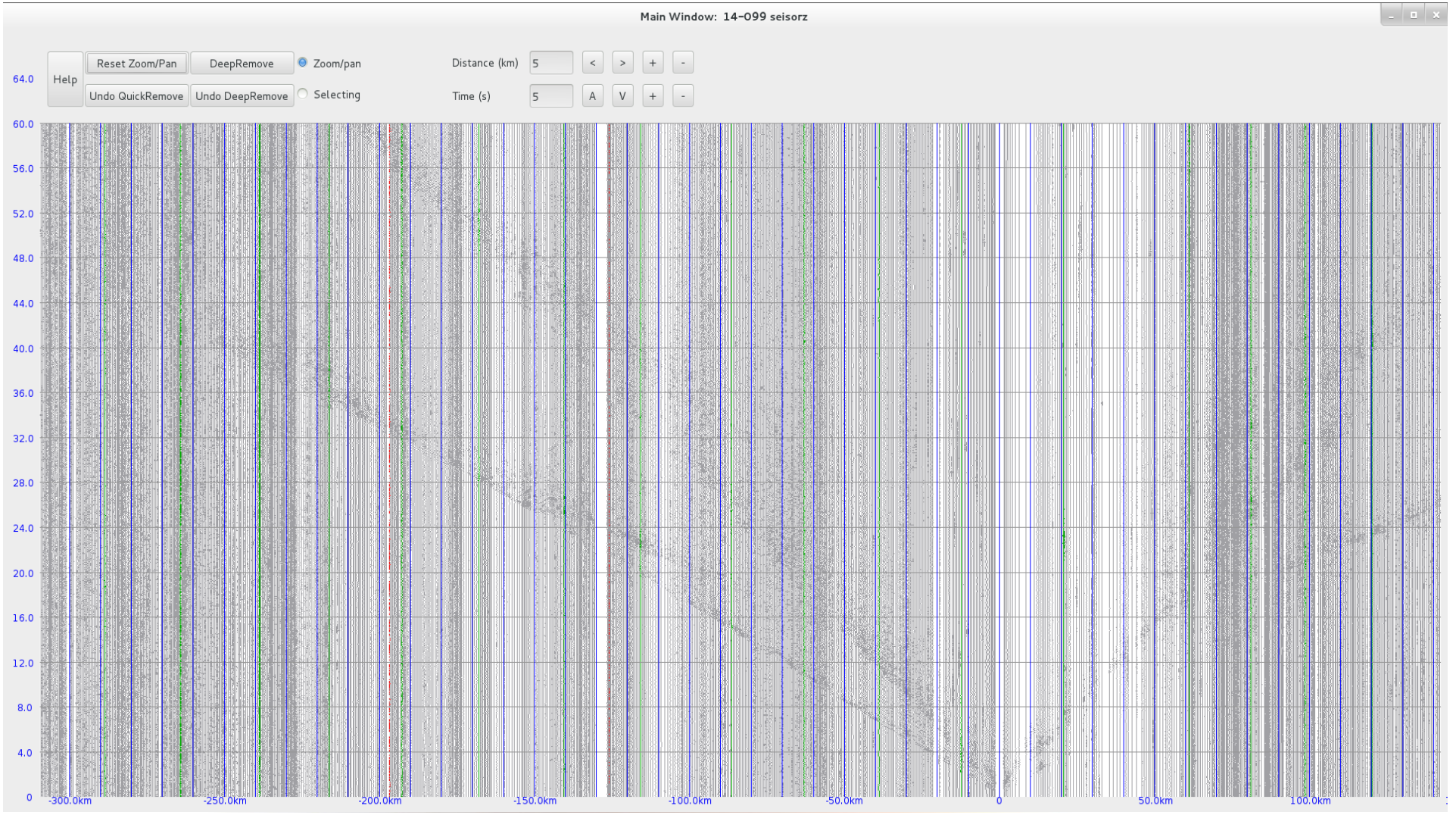

PH5ViewerwVispyAPI.py
File
Help
Control
Events
Arrays
NOTICE: Graph Name is '14-099 seisorz'. Click on Properties in Control tab to change name of the graph
Start time: 2014:328:10:15:00.000
Length(s): 300.0
ID Time Latitude Longitude Elevation(m) Mag Depth(m)
5001 2014:328:04:25:00.000 -18.41244766 21.86500666 995.702 1000.0 0.0
5002 2014:328:05:25:00.000 -18.56569414 21.98445972 981.467 1000.0 0.0
5003 2014:328:06:25:00.000 -18.71685701 22.10816575 982.603 1000.0 0.0
5004 2014:328:07:15:00.000 -18.86555648 22.22697373 982.255 1000.0 0.0
5005 2014:328:08:05:00.000 -18.98805048 22.27282974 976.383 1000.0 0.0
5006 2014:328:08:55:00.000 -19.207188 22.18723334 968.357 1000.0 0.0
5007 2014:328:10:15:00.000 -19.50077899 22.10372901 963.197 1000.0 0.0
5008 2014:328:11:15:00.000 -19.79267576 22.19873526 954.109 1000.0 0.0
5009 2014:328:12:05:00.000 -20.00028075 22.30384524 955.149 1000.0 0.0
5010 2014:328:13:35:00.000 -20.0981105 22.52407924 938.706 1000.0 0.0
5011 2014:328:14:35:00.000 -20.18648429 22.77032658 938.178 1000.0 0.0
5012 2014:328:04:10:00.000 -20.26686425 23.0679015 932.732 1000.0 0.0
5013 2014:328:06:40:00.000 -20.31324262 23.33637079 970.775 1000.0 0.0
5014 2014:328:08:20:00.000 -20.49126168 23.44089966 984.315 1000.0 0.0
5015 2014:328:09:30:00.000 -20.60987364 23.57030466 974.962 1000.0 0.0
5017 2014:328:13:30:00.000 -20.97740987 23.77326102 950.928 1000.0 0.0
5018 2014:329:08:00:00.000 -21.12004302 23.97860851 944.851 1000.0 0.0
5019 2014:329:08:50:00.000 -21.260847 24.08978 948.82 1000.0 0.0
5020 2014:329:10:00:00.000 -21.44555433 24.23694167 946.496 1000.0 0.0


PH5ViewerwVispyAPI.py
File
Help
Control
Events
Arrays
Channels: 1
Array: 001
Sample Rate: 250.0
Deploy Time: 2014:325:10:00:00
Pickup Time: 2014:329:10:00:00
View Stations
Include this array in the plotting
Array: 001
ID DAS Latitude Longitude Elevation(m) Component
1 13953 -18.26142 21.75392 997.6 1
2 14081 -18.26273 21.75808 992.0 1
3 14008 -18.26472 21.76365 994.2 1
4 14069 -18.26593 21.76748 993.2 1
5 12505 -18.26939 21.77063 993.6 1
6 12206 -18.27423 21.7743 993.1 1
7 13856 -18.27757 21.77764 993.7 1
8 14012 -18.27983 21.78237 987.6 1
9 12682 -18.28226 21.78819 992.7 1
10 12720 -18.28658 21.78842 993.7 1
11 12616 -18.29308 21.78789 993.1 1
12 12879 -18.29853 21.78817 995.7 1
13 14080 -18.30226 21.78922 995.0 1
14 11909 -18.30467 21.7906 996.8 1
15 14036 -18.30812 21.79348 997.7 1
16 12699 -18.31116 21.79758 997.1 1
17 12988 -18.31482 21.80263 994.7 1
18 12464 -18.31656 21.80438 993.1 1
19 13960 -18.32107 21.80844 991.0 1


PH5ViewerwVispyAPI.py
File
Help
Control
Events
Arrays
Start time
2014:328:10:15:00.000
Length(s)
60
Offset
0
Ignore minor signal (0-20%)
5
Apply
Overlap (0-80%):
25
Apply
NORMALIZE
Apply
STATION SPACING UNKNOWN
Nominal station spacing(m):
1000
Reduction Velocity(m/s):
0
Time Correction
Default Prop.
Name-Color Prop.
Previous Prop.
Apply and RePlot
Vertical grid
Horizontal grid
ReGrid Panel
Main
Support
Both
H. Grid Interval (s)
4.0
V. Grid Interval (km)
10
ReGrid
Time Direction
Down
Up
Drawing Style
Lines
Points
Get Data and Plot
No of Samp./Station:
Sample Interval (ms):
TIME (ms)
STATION
DISTANCE (km)
1/5 Getting PH5Data : reading data and metadata: 150 stations
WARNING: In some special case, the system may get crashed
during the run of 'Get Data and Plot'. Please save all
your works before continuing.
NOTICE: Array and channel can be switched in tab 'Arrays'.
NOTICE: Switching event, array, channel will reset all previous data.
NOTICE: Use Right Click to view Station's info
****


PH5ViewerwVispyAPI.py
File
Help
Control
Events
Arrays
Start time
2014:328:10:15:00.000
Length(s)
60
Offset
0
Ignore minor signal (0-20%)
5
Apply
Overlap (0-80%):
25
Apply
NORMALIZE
Apply
STATION SPACING UNKNOWN
Nominal station spacing(m):
1000
Reduction Velocity(m/s):
0
Time Correction
Default Prop.
Name-Color Prop.
Previous Prop.
Apply and RePlot
Vertical grid
Horizontal grid
ReGrid Panel
Main
Support
Both
H. Grid Interval (s)
4.0
V. Grid Interval (km)
10
ReGrid
Time Direction
Down
Up
Drawing Style
Lines
Points
Get Data and Plot
No of Samp./Station:
Sample Interval (ms):
TIME (ms)
STATION
DISTANCE (km)
WARNING: In some special case, the system may get crashed
during the run of 'Get Data and Plot'. Please save all
your works before continuing.
NOTICE: Array and channel can be switched in tab 'Arrays'.
NOTICE: Switching event, array, channel will reset all previous data.
NOTICE: Use Right Click to view Station's info
1/5 Getting PH5Data : reading data and metadata: 150 stations


PH5ViewerwVispyAPI.py
File
Help
Control
Events
Arrays
Start time
2014:328:10:15:00.000
Length(s)
60
Offset
0
Ignore minor signal (0-20%)
5
Apply
Overlap (0-80%):
25
Apply
NORMALIZE
Apply
STATION SPACING UNKNOWN
Nominal station spacing(m):
1000
Reduction Velocity(m/s):
0
Time Correction
Default Prop.
Name-Color Prop.
Previous Prop.
Apply and RePlot
Vertical grid
Horizontal grid
ReGrid Panel
Main
Support
Both
H. Grid Interval (s)
4.0
V. Grid Interval (km)
10
ReGrid
Time Direction
Down
Up
Drawing Style
Lines
Points
Get Data and Plot
No of Samp./Station:
Sample Interval (ms):
TIME (ms)
STATION
DISTANCE (km)
1/5 Getting PH5Data : reading data and metadata: 150 stations
WARNING: In some special case, the system may get crashed
during the run of 'Get Data and Plot'. Please save all
your works before continuing.
NOTICE: Array and channel can be switched in tab 'Arrays'.
NOTICE: Switching event, array, channel will reset all previous data.
NOTICE: Use Right Click to view Station's info
****


Main Window: 14-099 seisorz
Help
Reset Zoom/Pan
DeepRemove
Undo QuickRemove
Undo DeepRemove
Zoom/pan
Selecting
Distance (km)
5
Time (s)
5
A
V
64.0
60.0
56.0
52.0
48.0
44.0
40.0
36.0
32.0
28.0
24.0
20.0
16.0
12.0
8.0
4.0
0
-300.0km
-250.0km
-200.0km
-150.0km
-100.0km
-50.0km
0
50.0km
100.0km

PH5toSeg

Allows you to quickly create SEG-Y shot gathers using a variety of parameters.

To get SEG-Y gathers of all events simply run:

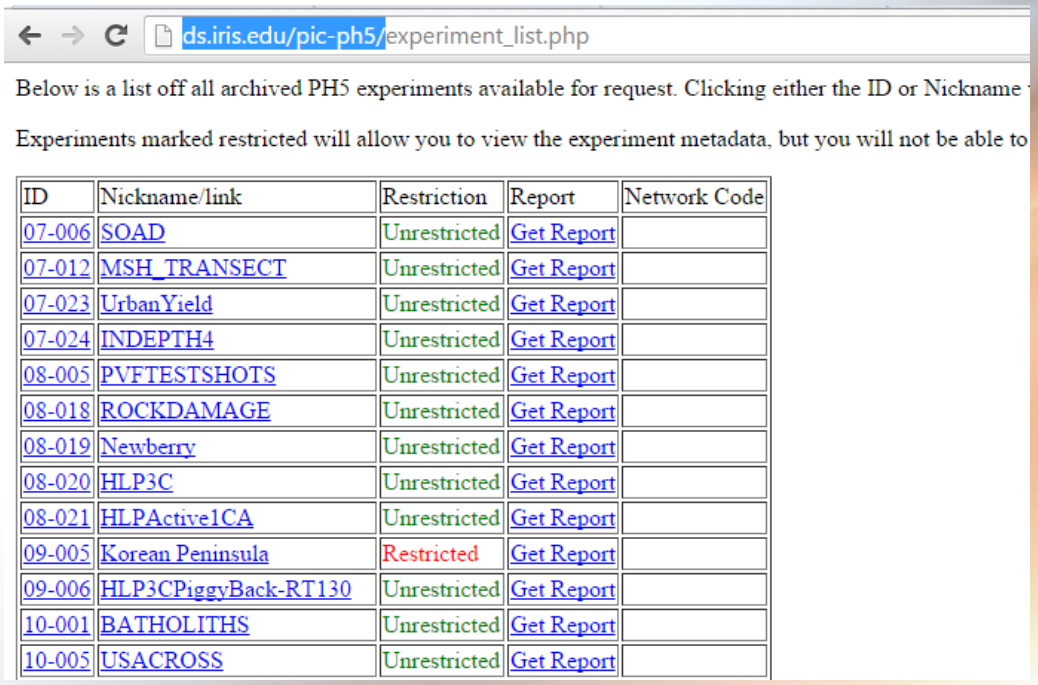
```
[field@dhcp-25 PH5Viewer]$ ph5toseg -n master -p /path/to/ph5 -E -l 60 -o /output_directory
```

```
Convert ph5 file to standard SEG-Y.

Options:
  -h, --help            show this help message and exit
  -e event_number, --eventnumber=event_number
  -E, --allevents
  -s start_time, --starttime=start_time
  -A, --all
  -t stop_time, --stoptime=stop_time
  -a array, --array=array
  -l length, --length=length
  -O offset, --offset=offset
  -n nickname, --nickname=nickname
  -p ph5_path, --ph5path=ph5_path
  -c channel, --channel=channel
  -N, --notimecorrect
  -d decimation, --decimation=decimation
  -f format, --format=format
  -o out_dir, --out_dir=out_dir
  -C, --check_tables
  --use_deploy_pickup   Use deploy and pickup times to determine if data
                        exists for a station.
  -D das_sn, --das=das_sn
  -S station, --station=station
  -Y doy_keep, --doy=doy_keep
                        Comma separated list of julian days to extract.
  -r sample_rate, --sample_rate_keep=sample_rate
  -V red_vel, --reduction_velocity=red_vel
  -U, --UTM             Fill SEG-Y headers with UTM instead of lat/lon.
  -x extended_header_style, --extended_header=extended_header_style
                        Extended trace header style:
                        'P' -> PASSCAL,                         'S' -> SEG,
                        'U' -> Menlo USGS,                        'I' ->
                        Scripts SIOSEIS Not implemented,
                        'N' -> iNova Firefly Not implemented
  --ic
  --break_standard
```

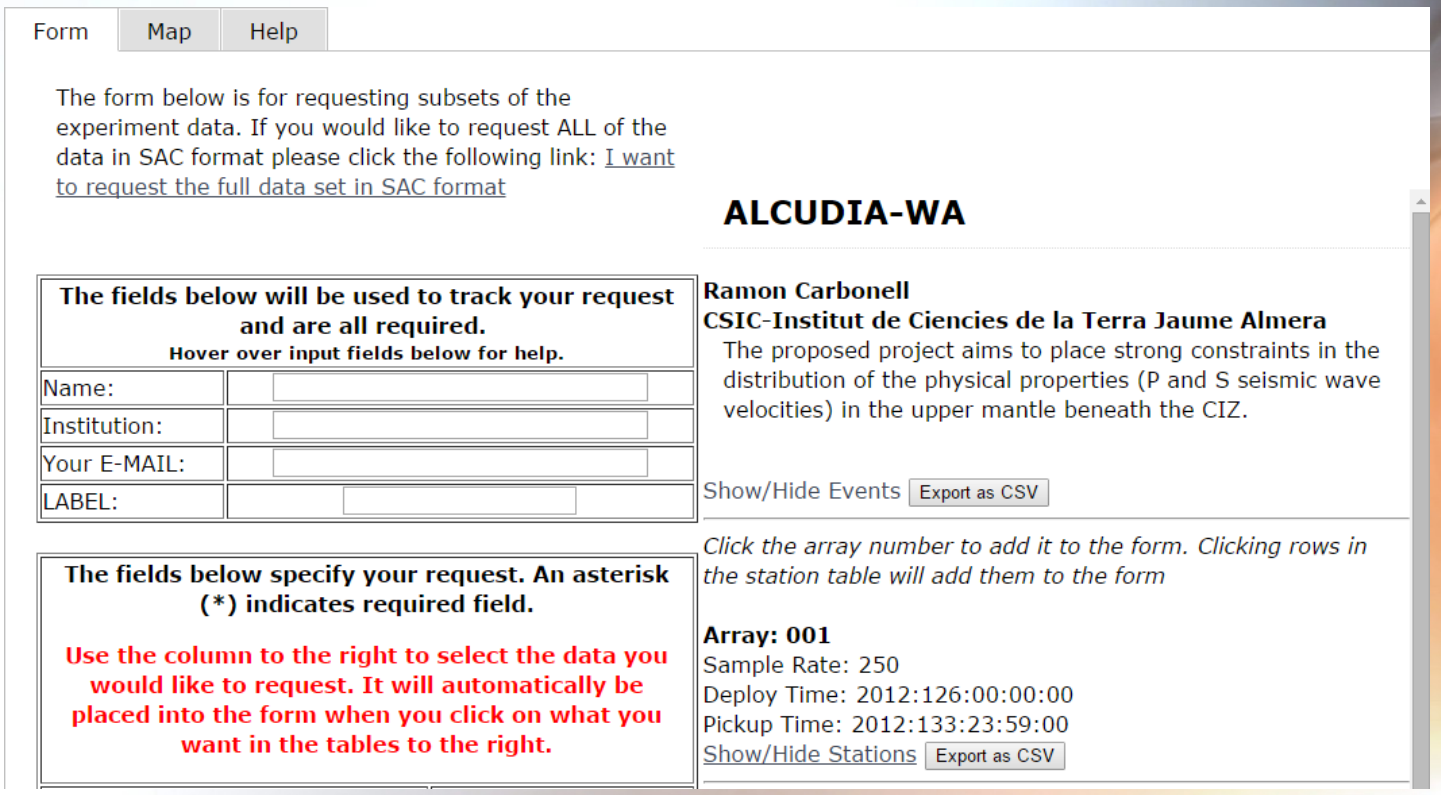
Switching Gears: PH5 Web Interface

The Website: http://ds.iris.edu/pic-ph5/



ID	Nickname/link	Restriction	Report	Network Code
07-006	SOAD	Unrestricted	Get Report	
07-012	MSH_TRANSECT	Unrestricted	Get Report	
07-023	UrbanYield	Unrestricted	Get Report	
07-024	INDEPTH4	Unrestricted	Get Report	
08-005	PVFTESTSHOTS	Unrestricted	Get Report	
08-018	ROCKDAMAGE	Unrestricted	Get Report	
08-019	Newberry	Unrestricted	Get Report	
08-020	HLP3C	Unrestricted	Get Report	
08-021	HLPActive1CA	Unrestricted	Get Report	
09-005	Korean Peninsula	Restricted	Get Report	
09-006	HLP3CPiggyBack-RT130	Unrestricted	Get Report	
10-001	BATHOLITHS	Unrestricted	Get Report	
10-005	USACROSS	Unrestricted	Get Report	

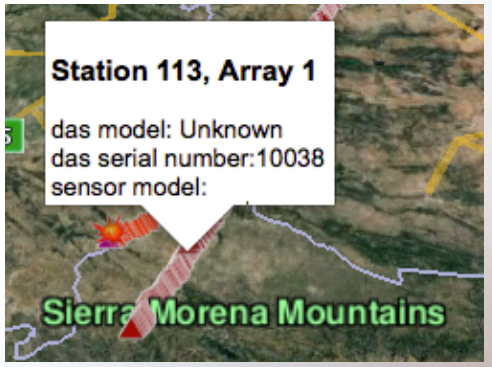
The Main Interface



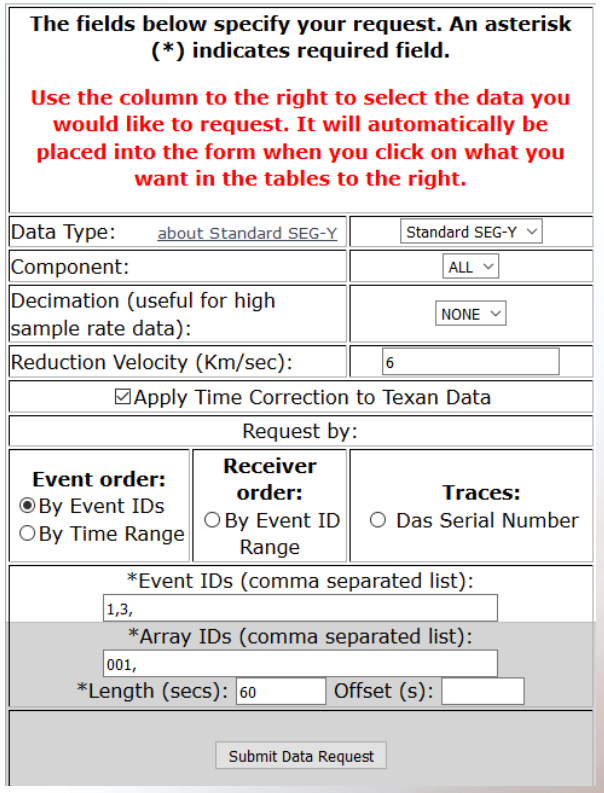
Map View

The map tab can be used to get a quick view of the experiment geometry.

Clicking on any point will bring up metadata about the receiver or shot.



The Main Form Fields Autofilled!



The fields below specify your request. An asterisk (*) indicates required field.

Use the column to the right to select the data you would like to request. It will automatically be placed into the form when you click on what you want in the tables to the right.

Data Type: about Standard SEG-Y	Standard SEG-Y
Component:	ALL
Decimation (useful for high sample rate data):	NONE
Reduction Velocity (Km/sec):	6

☑Apply Time Correction to Texan Data

Request by:

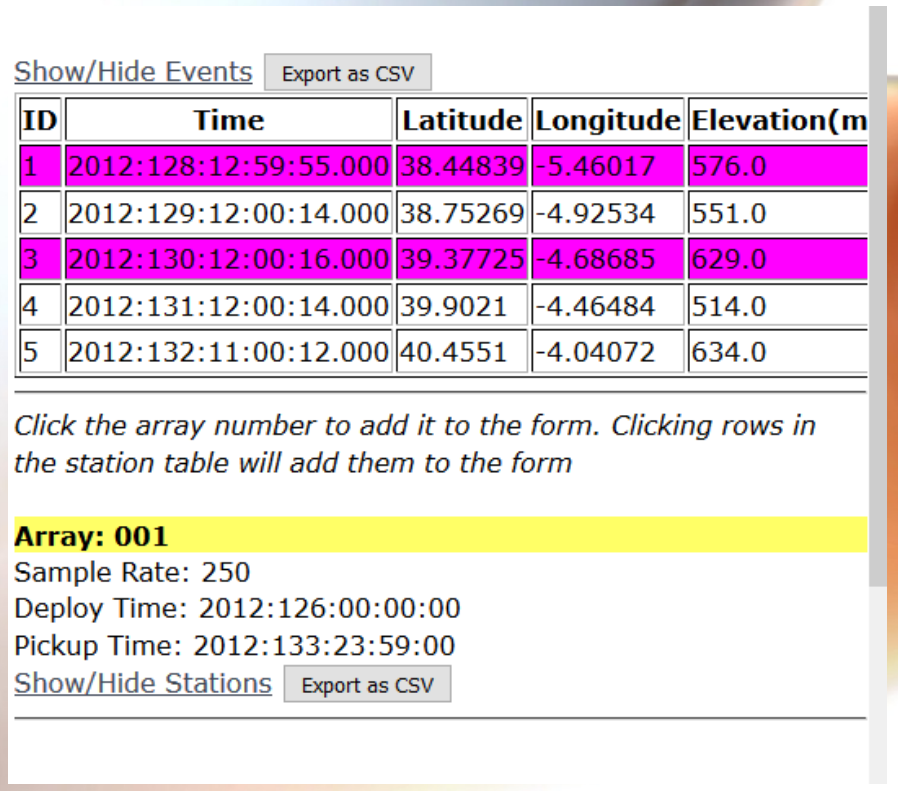
Event order:	Receiver order:	Traces:
◉By Event IDs ○By Time Range	○By Event ID Range	○ Das Serial Number

*Event IDs (comma separated list): 1,3,

*Array IDs (comma separated list): 001,

*Length (secs): 60 Offset (s):

Submit Data Request



Show/Hide Events | Export as CSV

ID	Time	Latitude	Longitude	Elevation(m
1	2012:128:12:59:55.000	38.44839	-5.46017	576.0
2	2012:129:12:00:14.000	38.75269	-4.92534	551.0
3	2012:130:12:00:16.000	39.37725	-4.68685	629.0
4	2012:131:12:00:14.000	39.9021	-4.46484	514.0
5	2012:132:11:00:12.000	40.4551	-4.04072	634.0

Click the array number to add it to the form. Clicking rows in the station table will add them to the form

Array: 001

Sample Rate: 250

Deploy Time: 2012:126:00:00:00

Pickup Time: 2012:133:23:59:00

Show/Hide Stations | Export as CSV

Submit Request-Receive Data

Below is the live output of the processing log.

You can close this page at anytime.

2015-12-07 15:14:14,250 Offset: 130732 m Azimuth: 24.359186 degrees
2015-12-07 15:14:14,324 Gain: 32 Bitweight: 5.96046e-08 volts/count
2015-12-07 15:14:14,325 Component: Z Azimuth: 0.0 degrees Dip: -90.0 degrees
2015-12-07 15:14:14,325 Clock: Start Epoch: 01336123843.000 End Epoch: 01336909319.000
2015-12-07 15:14:14,325 Clock: Offset: -0.0252304 seconds Slope: -3.21212e-08
2015-12-07 15:14:14,326 Sample rate: 250 Number of samples: 15000
2015-12-07 15:14:14,330 Wrote: 15000 samples with 0 sample padding.
2015-12-07 15:14:14,331 733

PH5 Data Request Message 2 of 9969

From PIC@iris.washington.edu
To dhess@passcal.nmt.edu
Date Today 08:20

data size: 87402.16KB
Tar file size: 36654.78KB
Data Set: 12-007

The data you requested can be found at the following link
ftp://dmc.iris.washington.edu/pub/userdata/ph5_data/Derick_Hess/Derick_Hess_test.tar.gz
The data will be available for 7 days before it is removed form the server.If you need more time to download the data please email engine_room@iris.washington.edu and let them know. Please right click the link and select 'save as' to download the file. Clicking the link may just bring up a blank page in some browsers

PH5 In The Future

PH5tomseed (2016)

New PH5 Website (Early 2016)

Integrated Web Services (Early 2016)

Integrated Field Tool (Mid 2016)

New Data Formats: Input and Output (As Community Need Arises)

Questions?

Contact

dhess@passcal.nmt.edu

software@passcal.nmt.edu