LASSO and MUSTANG: How to access web services for station diagnostics

Andy Frassetto, Justin Sweet, Danielle Sumy
• What is MUSTANG?
  – Data quality metrics web service
  – Produces variety of metrics for data at DMC
  – http://service.iris.edu/mustang

• What is LASSO?
  – Tool for accessing/analyzing data quality metrics
  – Runs entirely within web-browser
  – http://lasso.iris.edu
## IRIS DMC Web Services

**Services implementation: MUSTANG**

### Request tools

<table>
<thead>
<tr>
<th>Service interface</th>
<th>Version</th>
<th>Summary</th>
<th>Return options</th>
</tr>
</thead>
<tbody>
<tr>
<td>measurements</td>
<td>v.1</td>
<td>The main MUSTANG web service returning measurements for metrics relating to station data quality.</td>
<td>• XML (default)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• text</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• CSV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• JSON</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• JSONP</td>
</tr>
<tr>
<td>noise-psd</td>
<td>v.1</td>
<td>Returns Power Spectral Density estimates of seismic data and can generate aggregate plots.</td>
<td>• Text – CSV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• XML</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Plot (PNG)</td>
</tr>
<tr>
<td>noise-pdf</td>
<td>v.1</td>
<td>Returns Probability Density Functions in frequency ‘bins’ and can generate aggregate plots.</td>
<td>• Text – CSV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• XML</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Plot (PNG)</td>
</tr>
<tr>
<td>noise-mode-timeseries</td>
<td>v.1</td>
<td>Returns PDF Mode Timelines at select frequencies and can generate plots.</td>
<td>• Text – CSV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• XML</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Plot (PNG)</td>
</tr>
<tr>
<td>metrics</td>
<td>v.1</td>
<td>The metrics web service returns a description of available metrics in a variety of formats</td>
<td>• XML</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• HTML</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• XSD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• JSON</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• JSONP</td>
</tr>
<tr>
<td>targets</td>
<td>v.1</td>
<td>The targets web service returns a list of stations and channels for a given metric.</td>
<td>• Text</td>
</tr>
</tbody>
</table>
IRIS DMC MUSTANG metrics Web Service Documentation

orientation_check Channel orientation check

Summary
This metric takes advantage of Rayleigh waves' polarized retrograde elliptical particle motion to empirically estimate event back azimuths and compare them to calculated values. Radial and vertical components have similar motions but differ in phase by 90 degrees.

For events having $M_s$ or $m_b \geq 7.0$ and depth $< 100$ km, a station's horizontal components are rotated incrementally through 360 degrees to find the bearing from north that maximizes cross-correlation of this "trial radial component" and the Hilbert transform of the vertical component. (The Hilbert transform introduces a 90-degree phase shift.) Two peak correlation coefficients are calculated – they differ slightly in their normalization. One coefficient is maximized to find the estimated back azimuth (radial direction); the other characterizes the quality of the cross-correlation results.

The observed (empirical) channel orientations are reported for channels "X" and "Y" where the sensor is assumed to have two horizontal channels that differ by 90 degrees. Channel "X" is oriented 90 degrees clockwise from channel "Y", just as the X-axis is oriented 90 degrees clockwise from the Y-axis in a Cartesian coordinate system.


Uses
Channel orientation estimates having large correlation coefficients from many events and back azimuths can be averaged to give an empirical orientation estimate. This estimate can be used in the metadata when the channel orientation is unknown (e.g. for ocean-bottom seismometers) or can verify metadata orientations reported from the field. Effects such as multi-pathing make it unreliable to estimate orientation using a statistically small sample of measurements. (See the reference above for details on how to clean and average these measurements. Note that there is an apparent contradiction in the paper about which correlation coefficient to use for data cleaning; max_Czr should be $> 0.4$)

Algorithm
1. For events with $M_s$ or $m_b \geq 7.0$ and depth $< 100$ km...
URL Builder: noise-psd v.1

Use this form to build a URL to the noise-psd web service. Notice that as you edit the form, the link is automatically updated.

Targets

- SNCLQ filter or Target?
  - Filter
  - Target
- Network: IU
- Station: ANMO
- Location: 00
- Channel: BHZ
- Quality: M

Output

- Response
- Correction:
- Format:
  - XML (default)
  - Text
  - Plot

Temporal Constraints

- Parameter: Start/end
- Start time: 2010-01-01T00:00:00
- End Time: 2010-01-02T00:00:00

Click the link:
http://service.iris.edu/mustang/noise-psd/1/query?net=IU&sta=ANMO&loc=00&cha=BHZ&quality=M&starttime=2010-01-01T00:00:00&endtime=2010-01-02T00:00:00&correct=true&format=xml
Noise-Mode-Timeseries

URL Builder: noise-mode-timeseries v.1

Use this form to build a URL to the noise-mode-timeseries web service. Notice that as you edit the form, the link is automatically updated.

**Targets**

- **SNCLQ filter or Target?**
  - Option: Filter
- **Network**: IU
- **Station**: ANMO
- **Location**: 00
- **Channel**: BHZ
- **Quality**: M

**Date Range:**

- **Start time**: 2014-04-01
- **End Time**: 2014-04-07

**Output**

- **Format**: XML
- **Frequency Options**
  - **Auto**
  - **Frequencies**: 0.1, 0.01 Hz
  - **Periods**: 10, 100 sec

**Click the link:**

Latest Assessment of Seismic Station Observations (LASSO)

View groups of MUSTANG metrics designed to examine specific aspects of a seismic station's performance. Groupings focus on seismometer component channels and are available by Virtual Network.

Virtual Network: GSN

View: Mass Positions

Metric(s): m1,m2,m3

Location(s): Any

Ranking: Qualitative

Table Type: Snapshot

Time: 2015-01-01 00:00:00

Display metrics closest to this date and time

Get Measurements
View groups of MUSTANG metrics designed to examine specific aspects of a seismic station's performance. Groupings focus on seismometer component channels and are available by Virtual Network.

Virtual Network: _GSN

View: ✓ Mass Positions  Noise Power  Signal Quality  Time Series Integrity  Metadata Validity

Metric(s):

Location(s): Any

Ranking: ○ Qualitative  ○ Quantitative

Table Type: ○ Snapshot  ○ Mean for Period  ○ Median for Period

Time: 2015-01-01 00:00:00

Show Counts:

Get Measurements
Mass Positions
  - Displayed in volts/day

Noise Power
  - Displays 6 periods commonly of interest per day
  - Mode of PDF shown for 0.101s, 0.964s, 6.484s, 10.905s, 30.844s, and 103.747s
• Signal Quality
  – Grouping provides insight into quality of waveform recordings for a given day
  – Shows daily root-mean-square (RMS) variance of a time series
  – Shows daily ratio of the RMS variance calculated from teleseismic earthquakes
  – Shows % of PDFs in a day exceeding the Peterson(1993) high noise model
Metric Quality Assessment Groups

- Time Series Integrity
  - Grouping provides a sense of robustness of the archived data stream
  - Shows count for data gaps
  - Shows count for data overlaps
  - Shows % of data available during period requested
  - Shows length of largest gap (in seconds)
• Metadata Validity
  – Provides info on the timing quality of the seismic channels and whether calibration pulse is present
  – Shows if timing is questionable because clock has not obtained a lock since system power up
  – Daily avg timing quality based on accuracy of datalogger clock relative to GPS clock
  – Shows whether a timing correction has been applied
Disclaimer: Please note that metrics are still being developed for MUSTANG and back-populated for the entire IRIS archive. Most metrics for many major permanent networks are complete for the last several years, and gaps in metric coverage are being filled. Metric and network coverage reports are periodically updated. Channel labeling conventions are available through IRIS Data Services and Appendix A of the SEED manual.

Time Series Integrity view of _GSN for 2013-07-01T00:00:00 thru 2013-09-29T00:00:00 requested at Thu Apr 23 2015 10:17:02 GMT-0600 (MDT):

<table>
<thead>
<tr>
<th>Target</th>
<th>Rank</th>
<th>num_gaps</th>
<th>num_overlaps</th>
<th>percent_availability</th>
<th>max_gap</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU.MCQ.BHZ.M</td>
<td>100.00</td>
<td>0.06</td>
<td>0.00</td>
<td>99.99</td>
<td>11.25</td>
<td>86/88</td>
</tr>
<tr>
<td>BK.CNB.00.BHZ.M</td>
<td>100.00</td>
<td>0.01</td>
<td>0.00</td>
<td>99.98</td>
<td>13.80</td>
<td>87/87</td>
</tr>
</tbody>
</table>
## Time Series Integrity Results

<table>
<thead>
<tr>
<th>Target</th>
<th>Rank</th>
<th>num_gaps</th>
<th>Count</th>
<th>num_overlaps</th>
<th>Count</th>
<th>percent_availability</th>
<th>Count</th>
<th>max_gap</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU_MCQ_BHZ_M</td>
<td>100.00</td>
<td>0.06</td>
<td>86/88</td>
<td>0.00</td>
<td>88/88</td>
<td>99.99</td>
<td>88/88</td>
<td>11.25</td>
<td>86/88</td>
</tr>
<tr>
<td>BK_CMB_00_BHZ_M</td>
<td>100.00</td>
<td>0.01</td>
<td>86/87</td>
<td>0.00</td>
<td>87/87</td>
<td>99.98</td>
<td>87/87</td>
<td>13.80</td>
<td>86/87</td>
</tr>
<tr>
<td>CI_PASC_00_BHZ_M</td>
<td>100.00</td>
<td>0.11</td>
<td>78/87</td>
<td>0.15</td>
<td>74/87</td>
<td>98.86</td>
<td>87/87</td>
<td>0.86</td>
<td>86/87</td>
</tr>
<tr>
<td>CI_PASC_10_BHZ_M</td>
<td>100.00</td>
<td>0.03</td>
<td>85/88</td>
<td>0.14</td>
<td>76/88</td>
<td>100.00</td>
<td>88/88</td>
<td>0.49</td>
<td>87/88</td>
</tr>
<tr>
<td>CU_ANWB_00_BHZ_M</td>
<td>100.00</td>
<td>0.00</td>
<td>88/88</td>
<td>0.00</td>
<td>88/88</td>
<td>100.00</td>
<td>88/88</td>
<td>0.00</td>
<td>88/88</td>
</tr>
<tr>
<td>CU_BBGH_00_BHZ_M</td>
<td>75.00</td>
<td>0.09</td>
<td>86/88</td>
<td>0.01</td>
<td>87/88</td>
<td>99.87</td>
<td>88/88</td>
<td>85.99</td>
<td>86/88</td>
</tr>
<tr>
<td>CU_BCIP_00_BHZ_M</td>
<td>100.00</td>
<td>0.00</td>
<td>88/88</td>
<td>0.00</td>
<td>88/88</td>
<td>100.00</td>
<td>88/88</td>
<td>0.00</td>
<td>88/88</td>
</tr>
<tr>
<td>CU_GRGR_00_BHZ_M</td>
<td>87.50</td>
<td>0.01</td>
<td>87/88</td>
<td>0.03</td>
<td>85/88</td>
<td>99.96</td>
<td>88/88</td>
<td>35.23</td>
<td>87/88</td>
</tr>
<tr>
<td>CU_GRTK_00_BHZ_M</td>
<td>62.50</td>
<td>0.23</td>
<td>70/80</td>
<td>0.01</td>
<td>79/80</td>
<td>84.15</td>
<td>70/80</td>
<td>4724.87</td>
<td>70/80</td>
</tr>
<tr>
<td>CU_GTBY_00_BHZ_M</td>
<td>75.00</td>
<td>0.07</td>
<td>86/88</td>
<td>0.01</td>
<td>87/88</td>
<td>99.99</td>
<td>88/88</td>
<td>71.04</td>
<td>86/88</td>
</tr>
</tbody>
</table>
Quick Plot of Metric Variance

IC.LSA.10.BHZ.M
2013-07-01T00:00:00 to 2013-07-31T00:00:00

percent_availability

percent_availability for IC.LSA.10.BHZ.M

View Plot As Image Below (may need to scroll it into view)

Link to this plot
### User-defined Performance Criteria

<table>
<thead>
<tr>
<th>Target</th>
<th>num_gaps Rules</th>
<th>num_overlaps Rules</th>
<th>percent_availability Rules</th>
<th>max_gap Rules</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU.MCQ.88.BHZ.M</td>
<td><img src="image1" alt="Rule" /></td>
<td><img src="image2" alt="Rule" /></td>
<td><img src="image3" alt="Rule" /></td>
<td><img src="image4" alt="Rule" /></td>
<td>11.25</td>
</tr>
<tr>
<td>BK.CMB.00.BHZ.M</td>
<td><img src="image1" alt="Rule" /></td>
<td><img src="image2" alt="Rule" /></td>
<td><img src="image3" alt="Rule" /></td>
<td><img src="image4" alt="Rule" /></td>
<td>13.80</td>
</tr>
<tr>
<td>CI.PASC.00.BHZ.M</td>
<td><img src="image1" alt="Rule" /></td>
<td><img src="image2" alt="Rule" /></td>
<td><img src="image3" alt="Rule" /></td>
<td><img src="image4" alt="Rule" /></td>
<td>98.86</td>
</tr>
<tr>
<td>CI.PASC.10.BHZ.M</td>
<td><img src="image1" alt="Rule" /></td>
<td><img src="image2" alt="Rule" /></td>
<td><img src="image3" alt="Rule" /></td>
<td><img src="image4" alt="Rule" /></td>
<td>100.00</td>
</tr>
<tr>
<td>CU.ANWB.00.BHZ.M</td>
<td><img src="image1" alt="Rule" /></td>
<td><img src="image2" alt="Rule" /></td>
<td><img src="image3" alt="Rule" /></td>
<td><img src="image4" alt="Rule" /></td>
<td>99.87</td>
</tr>
<tr>
<td>CU.BBGH.00.BHZ.M</td>
<td><img src="image1" alt="Rule" /></td>
<td><img src="image2" alt="Rule" /></td>
<td><img src="image3" alt="Rule" /></td>
<td><img src="image4" alt="Rule" /></td>
<td>99.96</td>
</tr>
<tr>
<td>CU.BCIP.00.BHZ.M</td>
<td><img src="image1" alt="Rule" /></td>
<td><img src="image2" alt="Rule" /></td>
<td><img src="image3" alt="Rule" /></td>
<td><img src="image4" alt="Rule" /></td>
<td>35.23</td>
</tr>
<tr>
<td>CU.GRGR.00.BHZ.M</td>
<td><img src="image1" alt="Rule" /></td>
<td><img src="image2" alt="Rule" /></td>
<td><img src="image3" alt="Rule" /></td>
<td><img src="image4" alt="Rule" /></td>
<td>87/88</td>
</tr>
</tbody>
</table>
Latest Assessment of Seismic Station Observations (LASSO)

Create a customized view of MUSTANG metrics, with the ability to tune the parameters used to assemble the network-station-channel-metric display.

Metrics:
- data_latency
- dc_offset
- dead_channel_exp
- digital_filter_charging
- digitizer_clipping
- event_begin
- event_end
- event_in_progress
- feed_latency
- glitches
- max_gap
- max_overlap
- max_stalta
- missing_padded_data
- num_gaps
- num_overlaps
- num_spikes
- orientation_check
- pct_above_nhnm
- pct_below_nhnm
- percent_availability
- polarity_check
- pressure_effects
- sample_max
- sample_mean
- sample_median
- sample_min
- sample_rms
- sample_onr
- spikes
- station_completeness
- suspect_time_tag
- telemetry_sync_error
- timing_correction

Click to select metric(s) to display.

[Contract metric list]
Latest Assessment of Seismic Station Observations (LASSO)

Create a customized view of MUSTANG metrics, with the ability to tune the parameters used to assemble the network-station-channel-metric display.

**Metric(s):**
- data_latency
- dc_offset
- dead_channel_exp
- digital_filter_charging
- elevation_angle
-...

Specify:
- By Network and Station
- By Virtual Network

**Network(s):**
- BK

**Station(s):**
- CMB

**Location(s):**
- 00

**Channel(s):**
- BHZ

**Quality:**
- M

**Ranking:**
- Qualitative
- Quantitative

**Table Type:**
- Snapshot
- Mean for Period
- Median for Period

**Start Time:**
- Set to
- 2013-07-01 00:00:00
- Start of metric averaging time window

Click to select metric(s) to display.
66% of stations are between 90-100% timing quality
Estimates are based on time since last GPS lock
Command Line Queries

• MUSTANG products can be pulled directly via correctly-formatted URLs
• Allows users to access MUSTANG metrics from within personal scripts/codes
#!/bin/bash
#
home=`pwd`

START="2013-01-01"; END="2015-01-01"

curl "http://service.iris.edu/fdsnws/station/1/query?net=GSN-BROADBAND&sta=*&loc=00,10,-&cha=BH&starttime=${START}T00:00:00&endtime=${END}T00:00:00&level=channel&format=text&nodata=404" > temp
tail +2 temp > GSN.txt; rm temp

if [ ! -d "PDFPSD" ]; then
  mkdir PDFPSD
fi

while read line; do
  name=$line;
  NET=`echo $name | awk -F' ' '{print $1}`;
  STA=`echo $name | awk -F' ' '{print $2}`;
  LOC=`echo $name | awk -F' ' '{print $3}`;
  CHA=`echo $name | awk -F' ' '{print $4}`;

  case $LOC in
    ""
      LOCS="-";;
  *)
    esac

  echo $NET.$STA.$LOC.$CHA
  curl "http://service.iris.edu/mustang/noise-pdf/1/query?net=${NET}&sta=${STA}&loc=${LOC}&cha=${CHA}&quality=M&starttime=${START}T00:00:00&endtime=${END}T00:00:00&format=text" > temp
tail +7 temp > PDFPSD/$NET.$STA.$LOC.$CHA.bin
  done < GSN.txt
LASSO the MUSTANG

http://lasso.iris.edu
http://service.iris.edu/mustang

What are you waiting for? Go be a cowboy!