Practical examples of LASSOing the MUSTANG for noise analysis and general quality assessment

Andy Frassetto

IRIS

What is MUSTANG?

- Data Services product (Modular Utility for STAtistical kNowledge Gathering)
- Data quality metrics web service
- http://service.iris.edu/mustang/

Service interface	Version	Summary	Return options
measurements	v.1	The main MUSTANG web service returning measurements for metrics relating to station data quality.	• XML (default) • text • CSV • JSON • JSONP
noise-psd	v.1	Returns Power Spectral Density estimates of seismic data and can generate aggregate plots.	• Text – CSV • XML • Plot (PNG)
noise-pdf	v.1	Returns Probability Density Functions in frequency `bins` and can generate aggregate plots.	• Text – CSV • XML • Plot (PNG)
noise-mode-timeseries	v.1	Returns PDF Mode Timelines at select frequencies and can generate plots.	• Text – CSV • Plot (PNG)

46 MUSTANG Metrics Currently

The *measurements* web service returns measurements for metrics relating to station data quality. **This is the primary query interface for MUSTANG users.**

Current list of all metrics

- **amplifier_saturation:**
  The number of times that the 'Amplifier saturation detected' bit in the 'dq_flags' byte is set within a miniSEED file. This data quality flag is set by some dataloggers in the fixed section of the miniSEED header. The flag was intended to indicate that the preamp is being overdriven, but the exact meaning is datalogger-specific.
  Please consult the Detailed Documentation

- **asl_coherence:**
  This metric computes the coherence between two co-located broadband sensors using a gamma-squared coherence. Data windows for comparison are one-day and only like components are compared (e.g. vertical to vertical). The coherence values are averaged into different frequency bands. (contributed by Albuquerque Seismic Laboratory)
  Please consult the Detailed Documentation

- **calibration_signal:**
  The number of times that the 'Calibration signals present' bit in the 'act_flags' byte is set within a miniSEED file. A value of 1 indicates that a calibration signal was being sent to that channel.
  Please consult the Detailed Documentation

- **channel_continuity:**
  This metric reports time durations of continuous data for the specified channel.It represents a combination of continuous days from the channel_up_time metric, so a large value suggests good data continuity.The start and end time will reflect the time extent for this continuous measurement.

IRIS

Example Documentation

IRIS DMC MUSTANG metrics Web Service Documentation

percent_availability Percentage of data available per day

Summary

The portion of data available for each day is represented as a percentage. 100% data available means full coverage of data for the reported start and end time.

Uses

percent_availability values can be averaged to give data completeness for a channel over an integer number of days. The metric also indicates whether or not data for a given channel/day is available for request from the IRIS DMC.

Algorithm

1. Request 24 hours of data for a single channel from web services.
2. Sum the data gaps in seconds for this period of time.
3. Calculate the percentage from the ratio of data found (omitting gaps) to data expected for the full day.

Target Domain

One channel per measurement

Duration

Window size is 24 hours, starting at midnight, UTC.

Formulae

1. percent_available := 100 – (100 * gap_seconds / total_seconds)

Constraints

1. Channels – ALL
2. Restricted data — No – pending

Data Preparation

1. Data is provided by web services
2. For data stamped quality 'M', overlapping segments are merged to eliminate overlaps
3. Merged data segments have the highest SEED quality factor available

IRIS

Example MUSTANG Output

http://service.iris.edu/mustang/measurements/1/query?metric=percent_availability&net=XE&sta=SNP63&loc=01&cha=BHZ&format=text&timewindow=2005-01-01T00:00:00,2008-01-01T00:00:00&orderby=start_asc

```
"100.000","XE.SNP63.01.BHZ.M","2005/10/27 00:00:00","2005/10/28 00:00:00","2015/10/31 16:13:55.456573"
"100.000","XE.SNP63.01.BHZ.M","2005/10/28 00:00:00","2005/10/29 00:00:00","2015/10/31 17:00:02.688316"
"100.000","XE.SNP63.01.BHZ.M","2005/10/29 00:00:00","2005/10/30 00:00:00","2015/10/31 20:00:35.294588"
"100.000","XE.SNP63.01.BHZ.M","2005/10/30 00:00:00","2005/10/31 00:00:00","2015/10/31 20:11:49.896299"
"51.5140","XE.SNP63.01.BHZ.M","2005/10/31 00:00:00","2005/11/01 00:00:00","2015/10/31 16:07:29.388275"
"21.2660","XE.SNP63.01.BHZ.M","2005/11/01 00:00:00","2005/11/02 00:00:00","2015/10/31 20:08:21.378744"
"100.000","XE.SNP63.01.BHZ.M","2005/11/02 00:00:00","2005/11/03 00:00:00","2015/10/31 16:09:15.985581"
"11.7730","XE.SNP63.01.BHZ.M","2005/11/03 00:00:00","2005/11/04 00:00:00","2015/11/01 14:58:07.675667"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/04 00:00:00","2005/11/05 00:00:00","2015/10/31 16:04:13.647447"
"22.6270","XE.SNP63.01.BHZ.M","2005/11/05 00:00:00","2005/11/06 00:00:00","2015/10/31 15:58:29.718914"
"100.000","XE.SNP63.01.BHZ.M","2005/11/06 00:00:00","2005/11/07 00:00:00","2015/10/31 16:07:03.473296"
"41.4430","XE.SNP63.01.BHZ.M","2005/11/07 00:00:00","2005/11/08 00:00:00","2015/10/31 15:58:10.921438"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/08 00:00:00","2005/11/09 00:00:00","2015/10/31 15:59:23.640449"
"22.2220","XE.SNP63.01.BHZ.M","2005/11/09 00:00:00","2005/11/10 00:00:00","2015/10/31 20:00:35.056613"
"60.3220","XE.SNP63.01.BHZ.M","2005/11/10 00:00:00","2005/11/11 00:00:00","2015/10/31 15:54:28.451867"
"21.9660","XE.SNP63.01.BHZ.M","2005/11/11 00:00:00","2005/11/12 00:00:00","2015/10/31 15:51:52.147679"
"79.6030","XE.SNP63.01.BHZ.M","2005/11/12 00:00:00","2005/11/13 00:00:00","2015/10/31 20:10:42.852515"
"54.1060","XE.SNP63.01.BHZ.M","2005/11/13 00:00:00","2005/11/14 00:00:00","2015/10/31 20:07:19.401646"
"54.7430","XE.SNP63.01.BHZ.M","2005/11/14 00:00:00","2005/11/15 00:00:00","2015/10/31 15:53:53.529862"
"54.0680","XE.SNP63.01.BHZ.M","2005/11/15 00:00:00","2005/11/16 00:00:00","2015/10/31 15:53:33.544219"
"27.6840","XE.SNP63.01.BHZ.M","2005/11/16 00:00:00","2005/11/17 00:00:00","2015/10/31 15:51:54.533119"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/17 00:00:00","2005/11/18 00:00:00","2015/10/31 15:52:55.244988"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/18 00:00:00","2005/11/19 00:00:00","2015/10/31 20:06:22.022643"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/19 00:00:00","2005/11/20 00:00:00","2015/10/31 15:53:30.947631"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/20 00:00:00","2005/11/21 00:00:00","2015/10/31 15:54:01.193034"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/21 00:00:00","2005/11/22 00:00:00","2015/10/31 15:54:36.670334"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/22 00:00:00","2005/11/23 00:00:00","2015/10/31 15:53:30.429154"
"0.00000","XE.SNP63.01.BHZ.M","2005/11/23 00:00:00","2005/11/24 00:00:00","2015/10/31 15:53:00.341857"
```

Example MUSTANG Output

http://service.iris.edu/mustang/measurements/1/query?metric=percent_availability&net=XE&sta=SNP63&loc=01&cha=BHZ&format=text&timewindow=2005-01-01T00:00:00,2008-01-01T00:00:00&orderby=start_asc

```
"100.000","XE.SNP63.01.BHZ.M","2005/10/27
"100.000","XE.SNP63.01.BHZ.M","2005/10/28
"100.000","XE.SNP63.01.BHZ.M","2005/10/29
"100.000","XE.SNP63.01.BHZ.M","2005/10/30
"51.5140","XE.SNP63.01.BHZ.M","2005/10/31
"21.2660","XE.SNP63.01.BHZ.M","2005/11/01
"100.000","XE.SNP63.01.BHZ.M","2005/11/02
"11.7730","XE.SNP63.01.BHZ.M","2005/11/03
"0.00000","XE.SNP63.01.BHZ.M","2005/11/04
"22.6270","XE.SNP63.01.BHZ.M","2005/11/05
"100.000","XE.SNP63.01.BHZ.M","2005/11/06
"41.4430","XE.SNP63.01.BHZ.M","2005/11/07
"0.00000","XE.SNP63.01.BHZ.M","2005/11/08
"22.2220","XE.SNP63.01.BHZ.M","2005/11/09
"60.3220","XE.SNP63.01.BHZ.M","2005/11/10
"21.9660","XE.SNP63.01.BHZ.M","2005/11/11
"79.6030","XE.SNP63.01.BHZ.M","2005/11/12
"54.1060","XE.SNP63.01.BHZ.M","2005/11/13
"54.7430","XE.SNP63.01.BHZ.M","2005/11/14
"54.0680","XE.SNP63.01.BHZ.M","2005/11/15
"27.6840","XE.SNP63.01.BHZ.M","2005/11/16
"0.00000","XE.SNP63.01.BHZ.M","2005/11/17
"0.00000","XE.SNP63.01.BHZ.M","2005/11/18
"0.00000","XE.SNP63.01.BHZ.M","2005/11/19
"0.00000","XE.SNP63.01.BHZ.M","2005/11/20
"0.00000","XE.SNP63.01.BHZ.M","2005/11/21
"0.00000","XE.SNP63.01.BHZ.M","2005/11/22
"0.00000","XE.SNP63.01.BHZ.M","2005/11/23
```

Noise Studies - Roll Your Own!

grab_pdf_all_GSN.bash

grab_pdf_all_GSN.bash 〉 No Selection

```
#!/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
   "")
      LOC="--";;
   *)
   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=${START}&endtime=${END}&format=text" > temp
   tail +7 temp > PDFPSD/${NET}.${STA}.${LOC}.${CHA}.bin
done < GSN.txt
```

Download noise probability density function bin file for each network.station.location.channel.

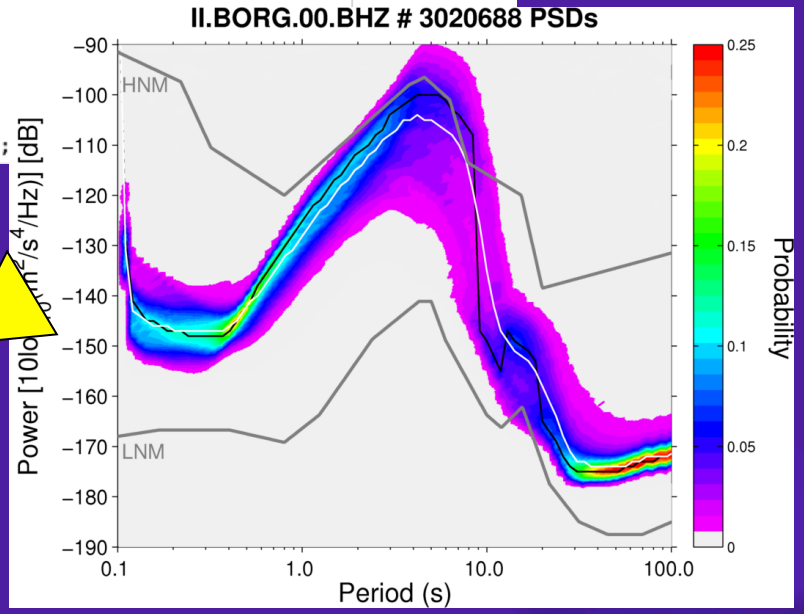
IRIS

Noise Studies - Roll Your Own!

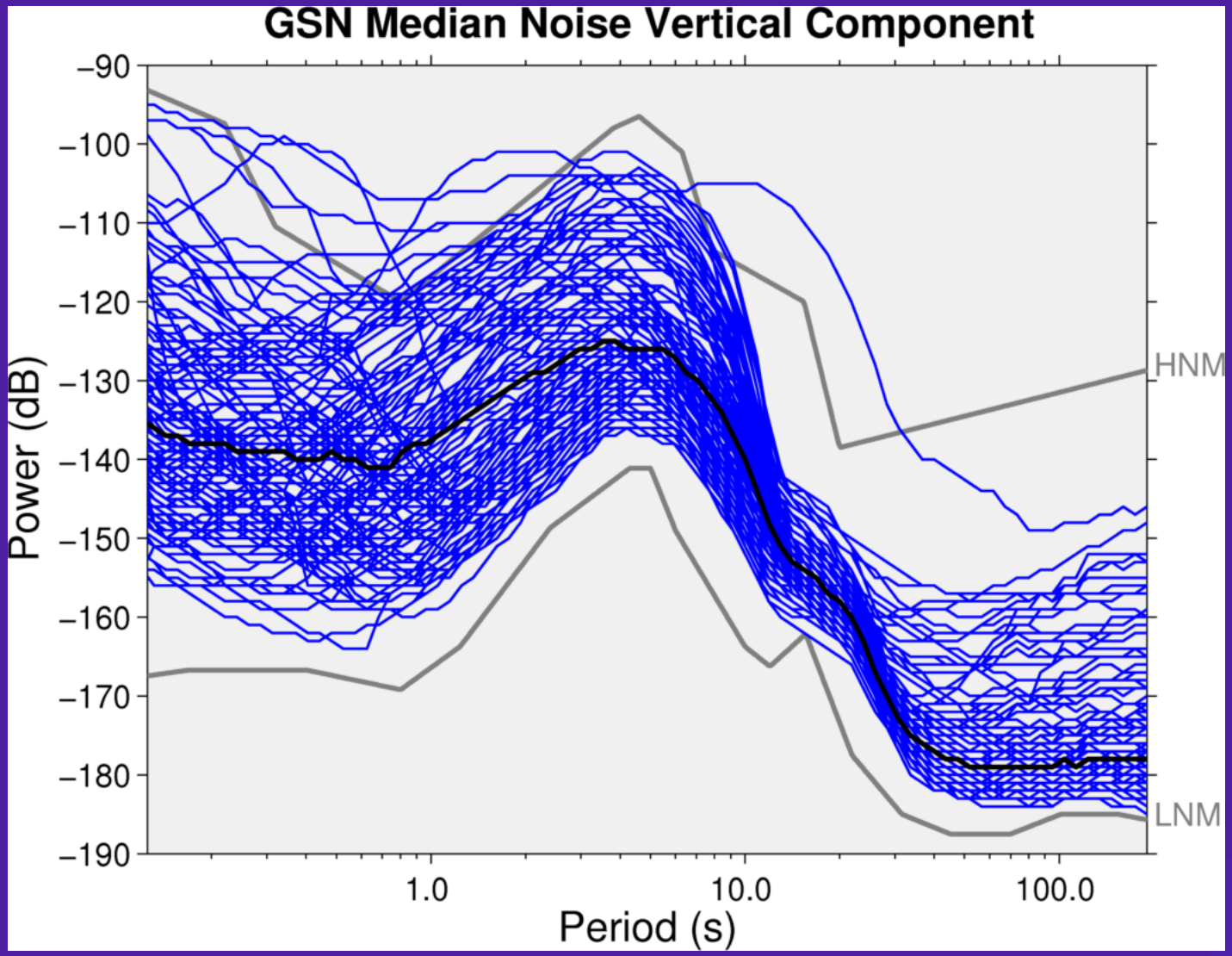
II.BORG.00.BHZ.bin

```
0.0052556, -175, 8
0.0052556, -174, 44
0.0052556, -173, 226
0.0052556, -172, 790
0.0052556, -171, 2428
0.0052556, -170, 4941
0.0052556, -169, 6897
0.0052556, -168, 6685
0.0052556, -167, 4638
0.0052556, -166, 2526
0.0052556, -165, 1160
0.0052556, -164, 748
0.0052556, -163, 474
0.0052556, -162, 409
0.0052556, -161, 343
0.0052556, -160, 280
0.0052556, -159, 187
0.0052556, -158, 221
0.0052556, -157, 141
0.0052556, -156, 118
0.0052556, -155, 116
0.0052556, -154, 86
0.0052556, -153, 64
0.0052556, -152, 73
0.0052556, -151, 55
0.0052556, -150, 54
0.0052556, -149, 49
0.0052556, -148, 44
0.0052556, -147, 55
```

```
89 -      psd = zeros(201,fc);
90        % loop through frequencies
91 -      for icol=skiphigh+1:fc
92 -         freq = frequencylist(icol);
93          % go through frequency list and select the ones match
94 -         for ifreq=1:length(tmp)
95 -            if(tmp(ifreq,1) == freq)
96 -               irow = round (-1*tmp(ifreq,2)) + 1; % round the dB values at this freq
97 -               if irow < 1 || irow > 201 fprintf('unusual irow: %d\n',irow); continue; end;
98                % add psd hits to the appropriate matrix element
99 -               psd(irow,icol) = psd(irow,icol) + tmp(ifreq,3);
100 -              numval = numval + tmp(ifreq,3);
101 -           end; % frequency list
102 -        end;% matrix population
103         % extract mode
104 -        [~,b]=max(psd(:,icol)); if (length(b)>1) disp(b); end
105 -        if (b>0); stn_mode(1,icol) = b; end
106
107         % extract median
108 -        [a,~,~]=find(psd(:,icol) ~= 0); b=nonzeros(psd(:,icol));
109 -        vals = [];
110 -        for i=1:length(a);
111 -           for j=1:b(i)
112 -              vals=[vals a(i)];
113 -           end
114 -        end
115 -        stn_medi(1,icol) = median(vals);
116 -     end;
117 -     stn_mode = -stn_mode; stn_medi = -stn_medi;
```



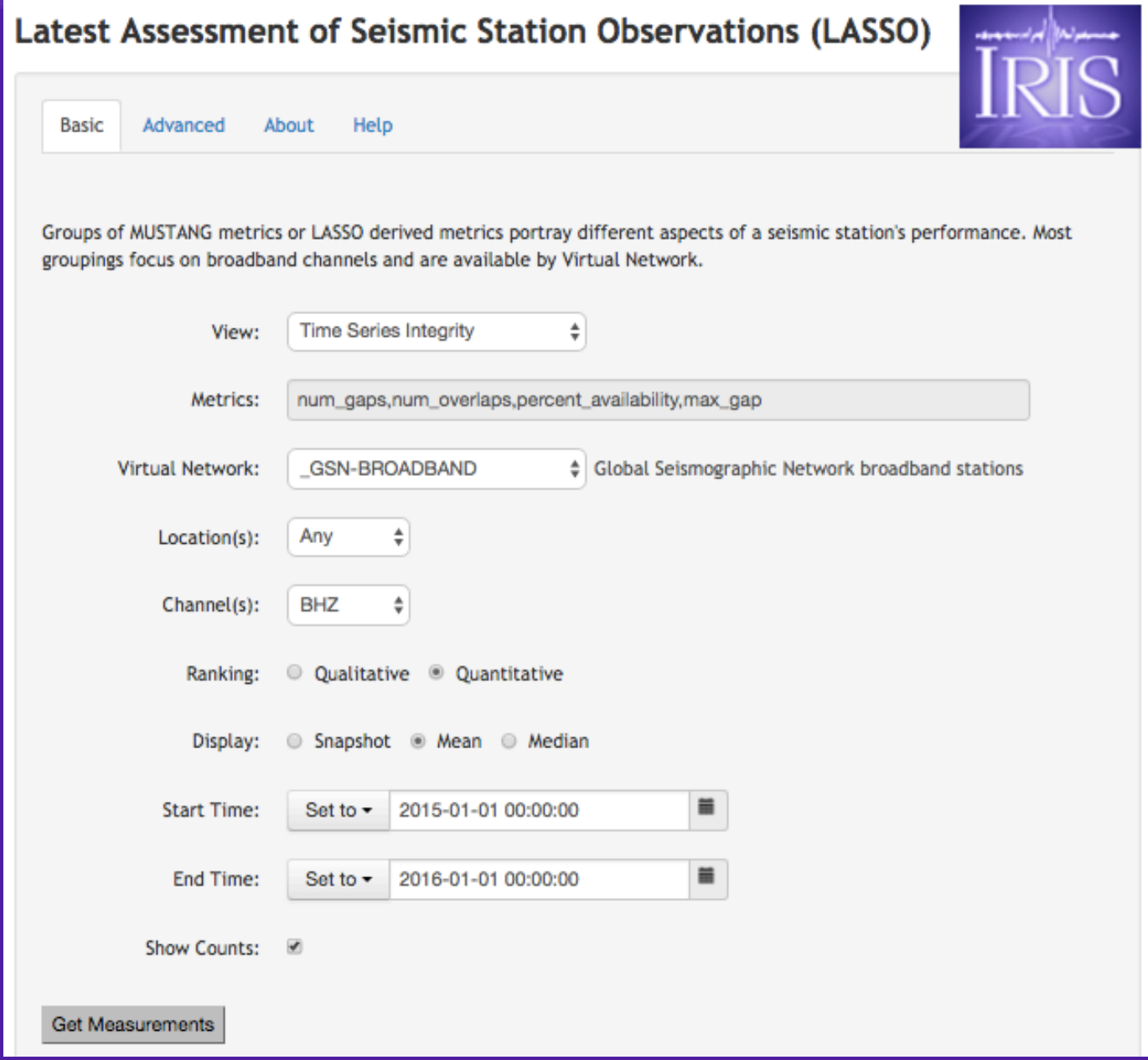
Noise Studies - Roll Your Own!



What is LASSO?

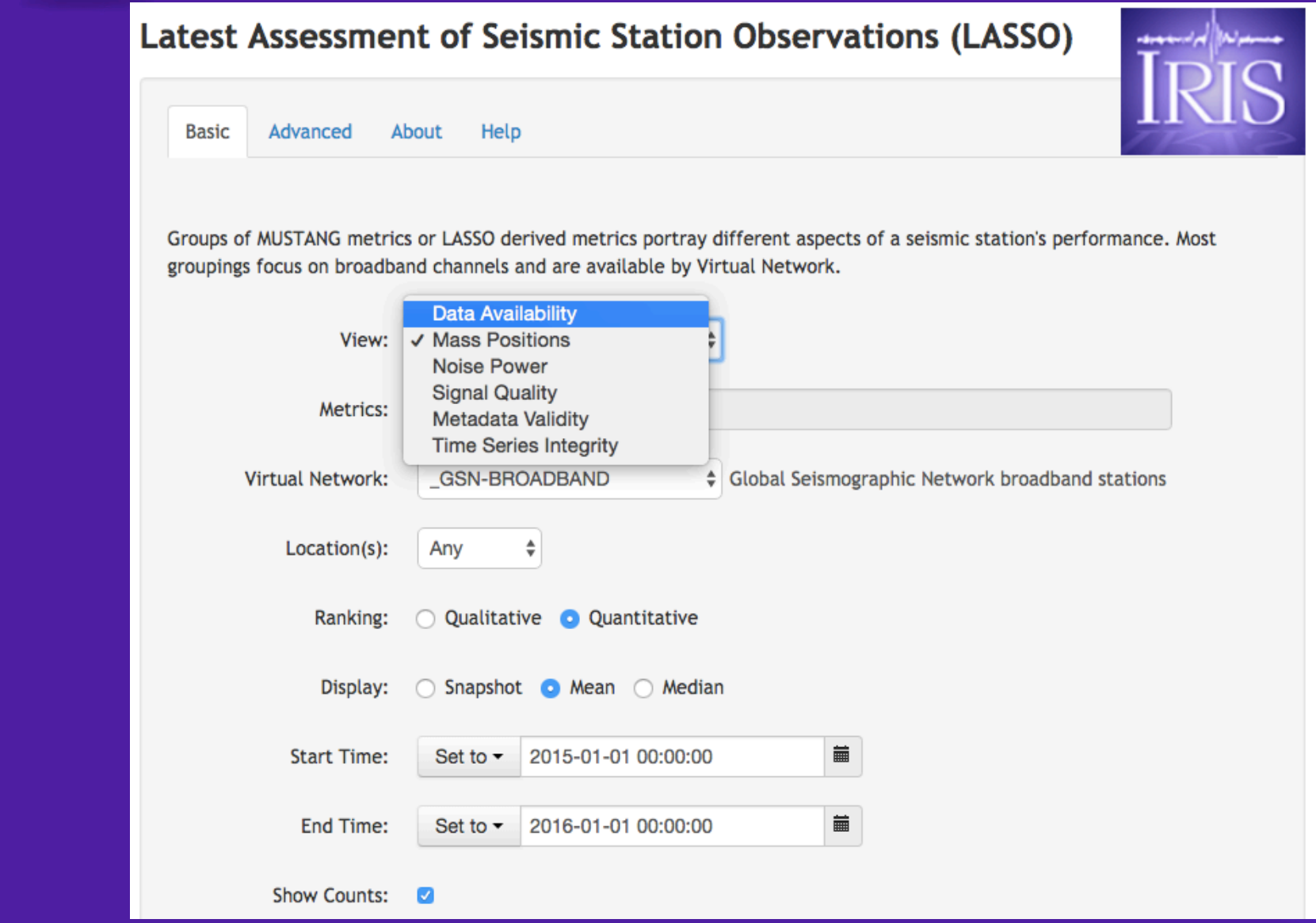
- Instrumentation Services product (Latest Assessment of Seismic Station Observations)
- Tool for accessing/displaying/ranking data quality metrics
- Web-client, runs within browser (chrome, safari, etc.)
- lasso.iris.edu

LASSO Basic View



IRIS

LASSO Metric Groups

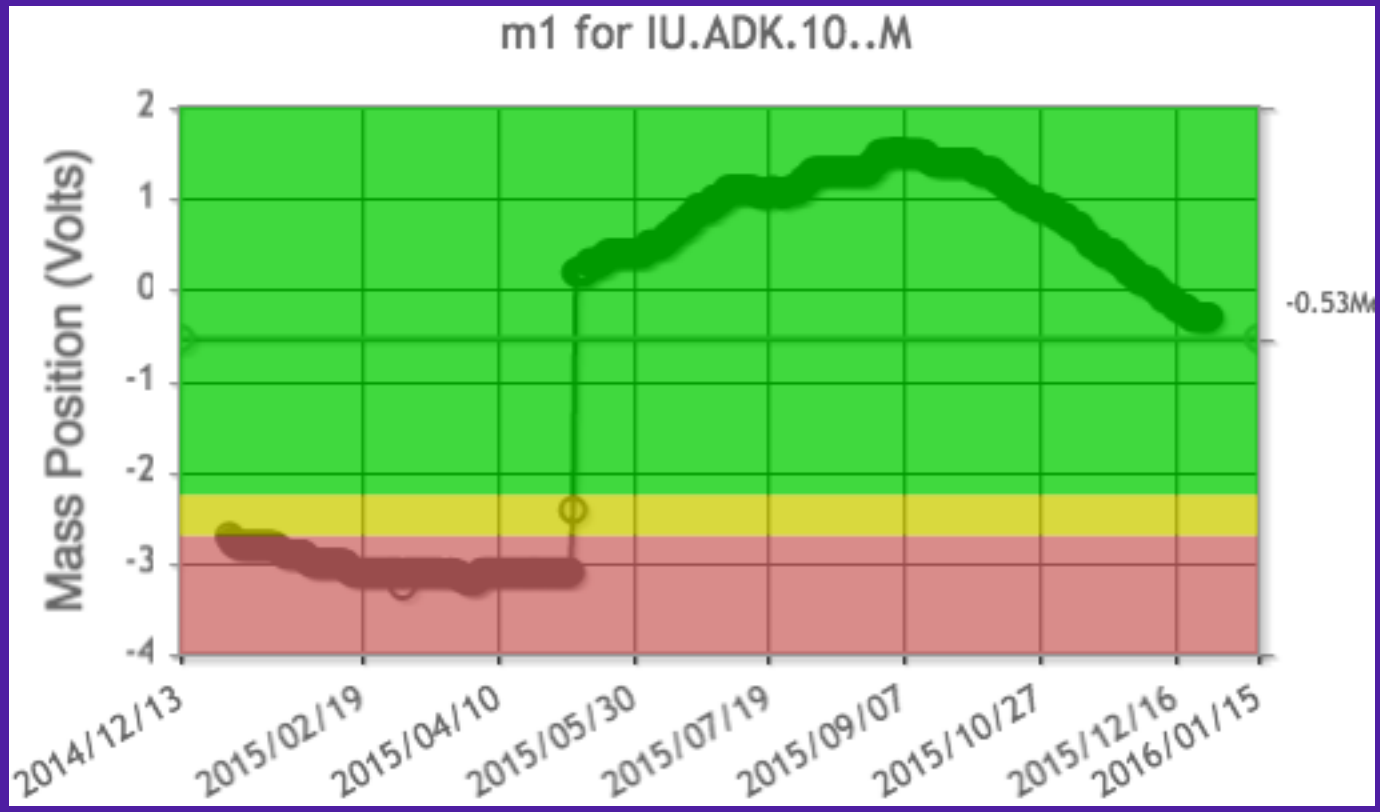


IRIS

LASSO Example - Masses

Target	Rank	m1 ? Rule	m2 ? Rule	m3 ? Rule
IU.QSPA.80..M	00,01,00			-3.88
IU.HKT.00..M	01,00,02	8.63	7.11	-0.50
IU.SNZO.00..M	01,02,00	0.35	4.24	4.10
II.CMLA.00..M	02,00,01	0.10	0.16	4.50
II.EFI.00..M	02,00,01	0.09	0.16	-6.60
II.KAPI.00..M	02,00,01	0.36	0.29	-7.30
II.KDAK.00..M	02,00,01	-0.30	0.10	-8.00
II.MBAR.00..M	02,00,01	0.12	-0.10	-5.40
II.MSEY.00..M	02,00,01	0.22	0.32	-9.80
II.MSEY.10..M	02,00,01	-1.34	-8.20	1.02
II.MSVF.00..M	02,00,01	-0.30	-0.15	-8.70
II.NIL.00..M	02,00,01	0.10	-0.08	-5.00
II.PALK.00..M	02,00,01	0.10	0.10	-5.70
II.SHEL.00..M	02,00,01	0.10	0.10	-6.10
IU.GUMO.00..M	02,00,01	0.10	-2.23	-5.18
IU.MBWA.00..M	02,00,01	0.80	0.30	-4.70
IU.PAYG.00..M	02,00,01	0.46	0.10	-5.40
IU.QSPA.00..M	02,00,01	-0.37	1.82	-8.11
II.BORG.00..M	02,01,00	0.20	1.10	-4.40
IU.RCBR.00..M	02,01,00	-0.45	0.29	-3.93
IU.WCI.00..M	02,01,00	-0.52	0.25	-3.91
BK.CMB.00..M	03,00,00	0.01	-0.00	0.00

LASSO Example - Masses



LASSO Example – Time Series

Time Series Integrity view of _CASCADIA-TA for 2016-01-01T00:00:00 thru 2016-01-31T00:00:00 requested at Tue Apr 12 2016 11:31:47 GMT-0700 (PDT):

QR Weights

0.250 num_gaps
0.250 num_overlaps
0.250 percent_availability
0.250 max_gap
Edit

entries

Search:

percent_availability Rules

bad >= 0
fair >= 75
good >= 90
Edit

	Rank (Weights)	num_gaps (? Performance Rule)	Count	num_overlaps (? Performance Rule)	Count	...ilability (Rule)	Count	max_gap (? Performance Rule)	Count
	100.00	0.00	30/30	0.		100.00	30/30	0.00	30/30
	100.00	0.00	30/30	0.		100.00	30/30	0.00	30/30
	75.00	0.07	28/30	0.		99.29	29/30	611.53	28/30
	62.50	0.77	18/30	0.00	30/30	84.68	18/30	8885.92	18/30
TA.D04E..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30
TA.E04D..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30
TA.F04D..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30
TA.F05D..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30
TA.G03D..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30
TA.G05D..BHZ.M Data	75.00	0.37	26/30	0.00	30/30	98.83	29/30	412.20	26/30
TA.H04D..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30
TA.I02E..BHZ.M Data	25.00	6.70	17/30	0.00	30/30	71.90	17/30	10637.68	17/30
TA.I03D..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30
TA.I04A..BHZ.M Data	62.50	0.32	19/28	0.00	28/28	75.75	19/30	16277.14	19/28
TA.I05D..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30
TA.J01E..BHZ.M Data	50.00	2.23	15/30	0.00	30/30	86.45	15/30	7824.95	15/30
TA.J04D..BHZ.M Data	100.00	0.00	30/30	0.00	30/30	100.00	30/30	0.00	30/30

IRIS

LASSO Example – Noise Power

View: Noise Power

Metrics: p1,p2,p3,p4,p5,p6,p7

Virtual Network: _US-TA USArray Transportable Array (EarthScope) stations

Location(s): Any

Channel(s): BHN

Ranking: Qualitative Quantitative

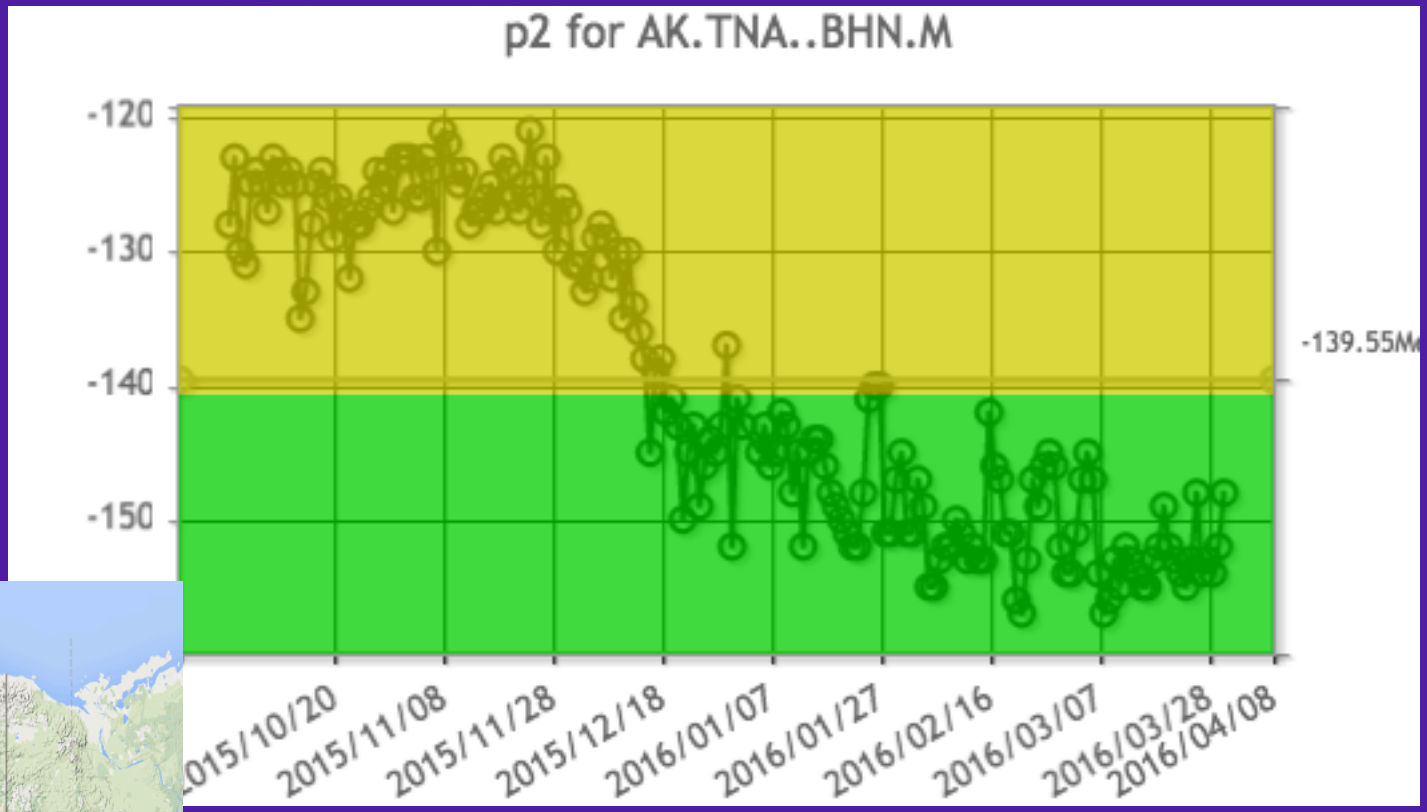
Display: Snapshot Mean Median

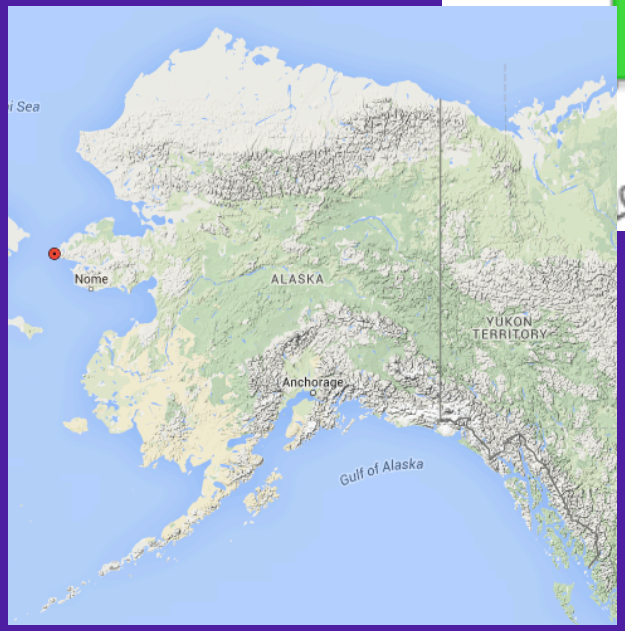
Start Time: Set to 2015-10-01 00:00:00

End Time: Set to 2016-04-01 00:00:00

TA.L61B..BHN.M	Data	01,05,00	-138.84	-137.29	-123.55	-139.27	-156.98	-150.90
TA.M02C..BHN.M	Data	03,03,00	-147.19	-146.32	-120.43	-132.62	-160.64	-150.66
TA.M04C..BHN.M	Data	01,05,00	-131.62	-138.70	-112.86	-126.31	-158.10	-150.04
TA.M19K..BHN.M	Data	03,03,00	-147.32	-145.22	-123.23	-137.13	-164.25	-156.14
TA.M20K..BHN.M	Data	06,00,00	-153.50	-147.75	-127.00	-140.00	-175.00	-176.75
TA.M22K..BHN.M	Data	02,04,00	-117.87	-130.14	-120.05	-135.40	-166.15	-158.00
TA.M24K..BHN.M	Data	01,05,00	-149.72	-134.22	-112.77	-127.74	-154.87	-146.70
TA.M26K..BHN.M	Data	04,02,00	-147.85	-146.57	-122.86	-137.77	-164.91	-161.59
TA.M27K..BHN.M	Data	04,02,00	-154.21	-143.07	-121.81	-135.81	-163.91	-160.91
TA.M30M..BHN.M	Data	04,02,00	-152.59	-150.16	-122.35	-135.49	-169.59	-170.72
TA.M31M..BHN.M	Data	04,02,00	-126.15	-148.66	-121.61	-135.18	-165.18	-160.62
TA.M53A..BHN.M	Data	01,05,00	-128.96	-135.41	-123.68	-138.52	-150.34	-143.31
TA.M54A..BHN.M	Data	01,05,00	-137.15	-137.78	-123.02	-138.32	-153.17	-151.34
TA.M65A..BHN.M	Data	00,03,03	-115.45	-109.11	-122.50	-134.21	-131.08	-127.17
TA.MDND..BHN.M	Data	00,06,00	-123.10	-131.38	-117.72	-134.34	-154.04	-149.83
TA.MSTX..BHN.M	Data	02,04,00	-127.70	-145.34	-123.96	-134.95	-150.79	-146.03
TA.N18K..BHN.M	Data	05,01,00	-151.43	-142.86	-129.00	-142.00	-162.00	-151.86
TA.N19K..BHN.M	Data	03,03,00	-152.16	-140.24	-121.72	-135.81	-164.32	-160.31
TA.N23A..BHN.M	Data	03,03,00	-131.85	-148.90	-125.55	-137.62	-154.11	-148.75
TA.N25K..BHN.M	Data	04,02,00	-153.48	-143.24	-121.98	-135.56	-168.70	-166.81

LASSO Example – Noise Power

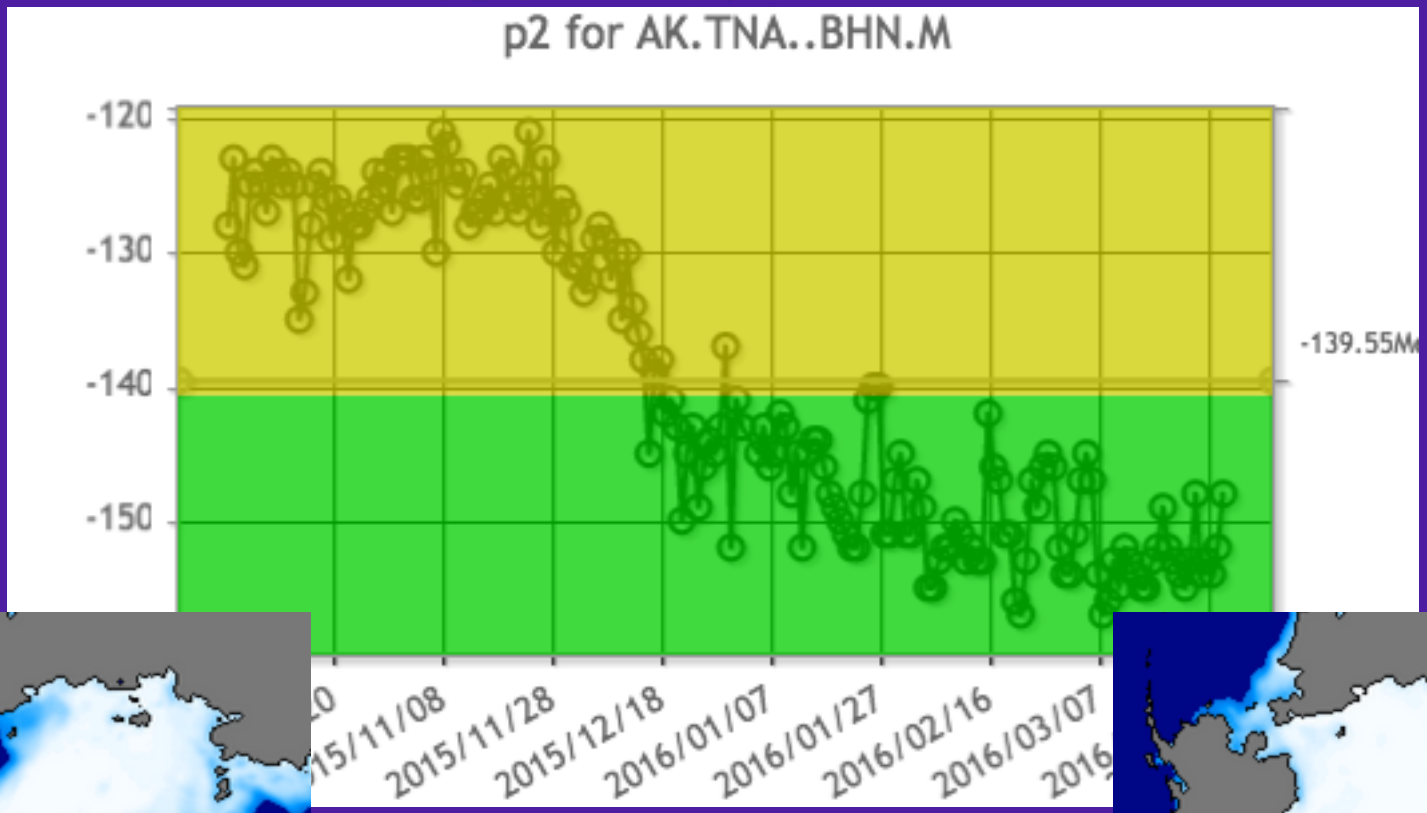


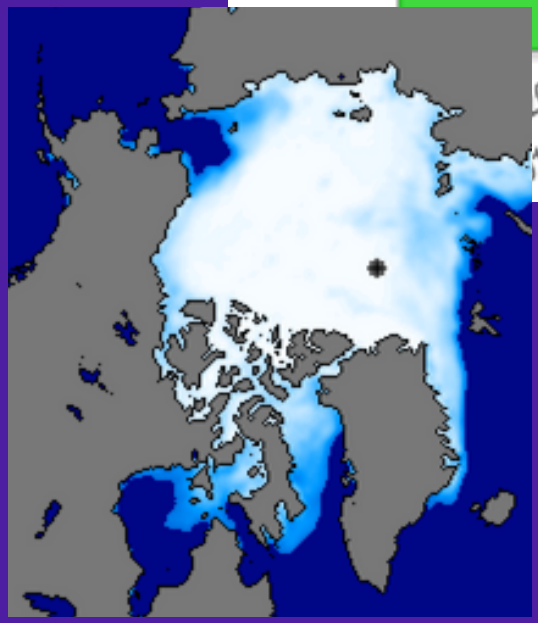


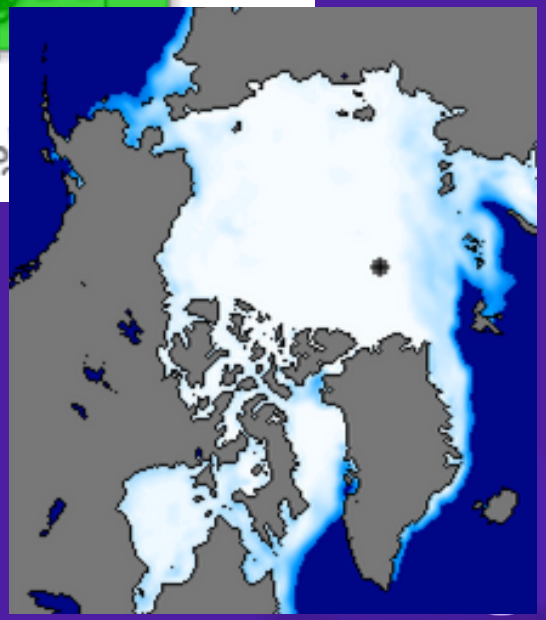
1Hz noise drops by 25 dB at Tin City station.

IRIS

LASSO Example – Noise Power







LASSO Advanced View

Basic | Advanced | About | Help

Develop a customized view of MUSTANG metrics, parameterizing a network- station-location-channel-metric display.

Metric(s):
dead_channel_lin
digital_filter_charging
digitizer_clipping
event_begin
event_end

(•) Metrics ( ) Derived Metrics

Click to select metric(s) to display.

Expand metric list

Hold [control] (PC) or [command] (Apple) to make multiple selections or hold [shift] (any platform) to select metrics en masse.

Search: (•) By Network and Station ( ) By Virtual Network

Network(s): BK Enter "*" for all or use single network code, e.g. "II".
"*" not recommended due to size of potential return

Station(s): CMB Enter "*" for all or station code, e.g. "BORG".

Location(s): 00 Enter "*" or location code, e.g. "00".

Channel(s): BHZ Enter "*" or wildcard with "?", e.g. "VM?".

Quality: M Enter "*" for all; "M" is highly recommend.

Ranking: ( ) Qualitative (•) Quantitative

LASSO Plan

- ca. 2014-2015 version is buggy, never had a formal release.
- Final development w/ISTI has resolved functionality issues, being finalizing currently.
- Will be given a formal release this summer, with documentation/tutorials.