A Semi-Automatic Method for Cepstral Depth Estimates for Shallow Events

Dr. Robert C. Kemerait
Senior Scientist

Dr. Ileana M. Tibuleac
Geophysicist

Marina A. Capuano
MERC-G Systems Engineer

Daniel H. Stayt
Geophysicist

Jose F. Pascual-Amadeo
Securburations Systems Engineer

Dr. Michael Thursby
NDC Consultant

Acknowledgements:

Dr. Mark Woods

Dr. Eileen Fisher

Air Force Technical Applications Center (AFTAC)

Objective

We investigate optimal processing of very shallow events (large SNR explosions and earthquakes with depth less than 3 km) with the goal to:

- **Implement an automatic set of algorithms for depth estimation, based on Cepstral Analysis;**
- **Using analysis metrics, provide a reliable statistical assessment of the measurement confidence and errors.**

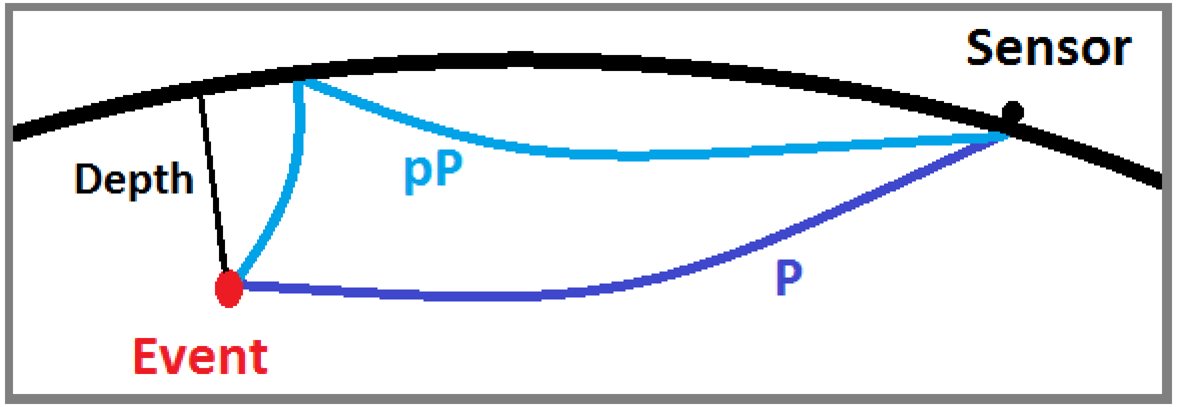
New Developments

- **Use of Complex and Power Cepstrae, versus Power Cepstrum only.**
- **Homomorphic deconvolution allows comparison with the initial signal, deconvolution of the initial signal, phase delay and polarity check;**
- **Metrics are developed, tested and used to stabilize and statistically evaluate the depth estimates:**

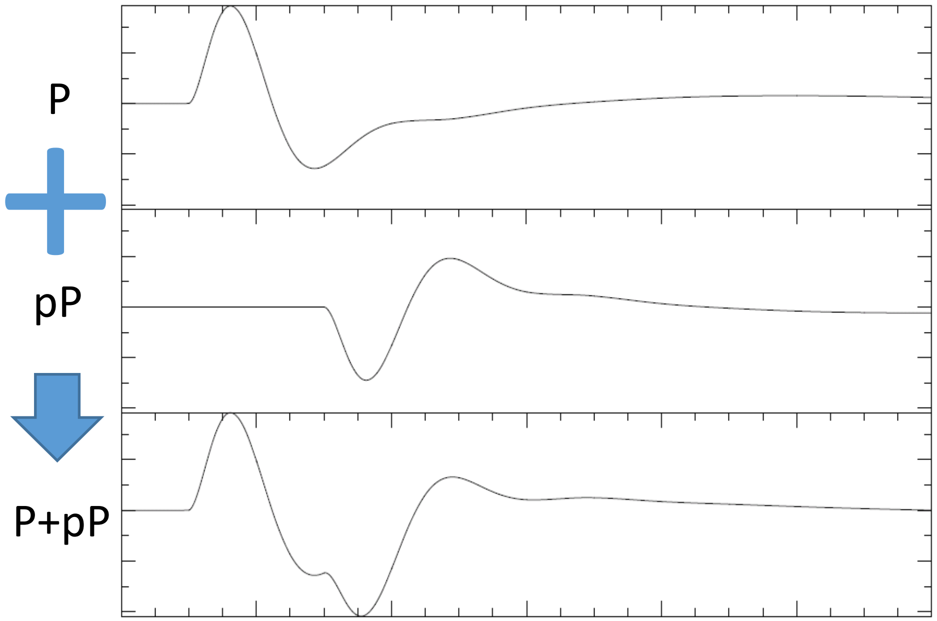
$$\text{TOTAL METRIC} = \Pi_i(\ \text{METRIC}(i)\ \text{WEIGHT}(i)\ )$$

- **New semi-automatic approach: Process a large number of signal windows at a single station, equivalent to array processing. Combine with array processing when available.**
- **Two “flavors” of code:**
  - **WIZARD – Completely interactive**
  - **BATCH – Semi automated with interactive quality control metric**

Assumptions



CEPSTRAL ANALYSIS WITH HOMOMORPHIC DECONVOUTION



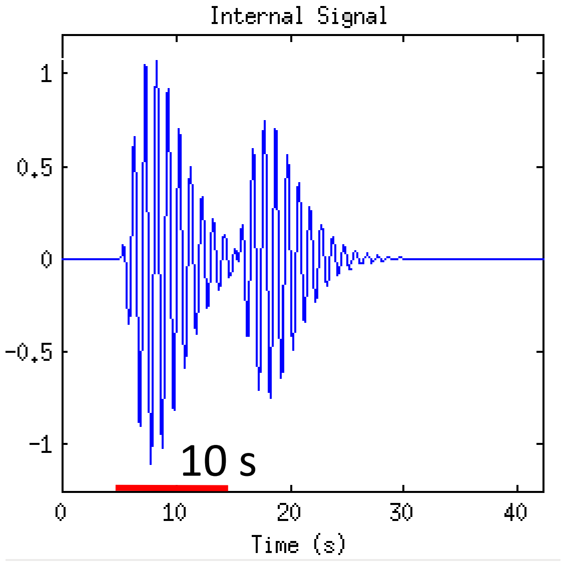
ASSUMPTIONS

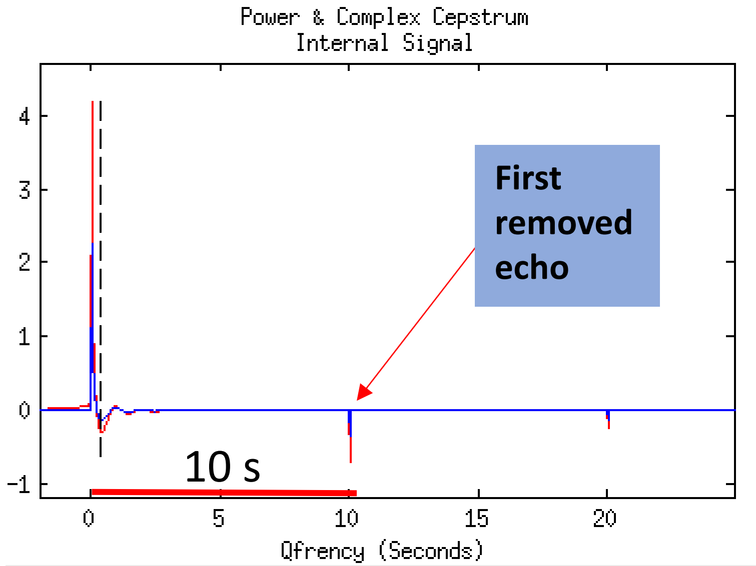
1. For underground explosions, the first arrival (named P) is larger than the reflected phase (pP) ;
2. The signal is a reasonable approximation of a minimum phase signal;
3. The P and pP are much larger than the seismic noise;
4. P and pP have similar shape;
5. A preliminary location (with depth) is available, and seismic phases are identified;
6. A seismic P-velocity model is available at the event location;
7. The event location is shallower than 3 km;
8. The depth is estimated as half the P-pP time delay * velocity at source *1/cos(ray_take-off_angle_deg)
    1. $\text{Depth} = \frac{1}{2} t_{\text{P}-\text{pP time delay}} * v_{\text{Source}} * \frac{1}{\cos[\theta_{\text{ray take off angle}}]}$

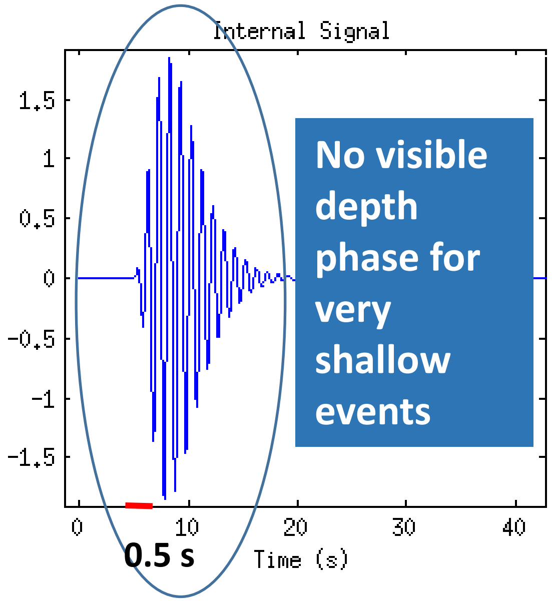
Typical very shallow event:

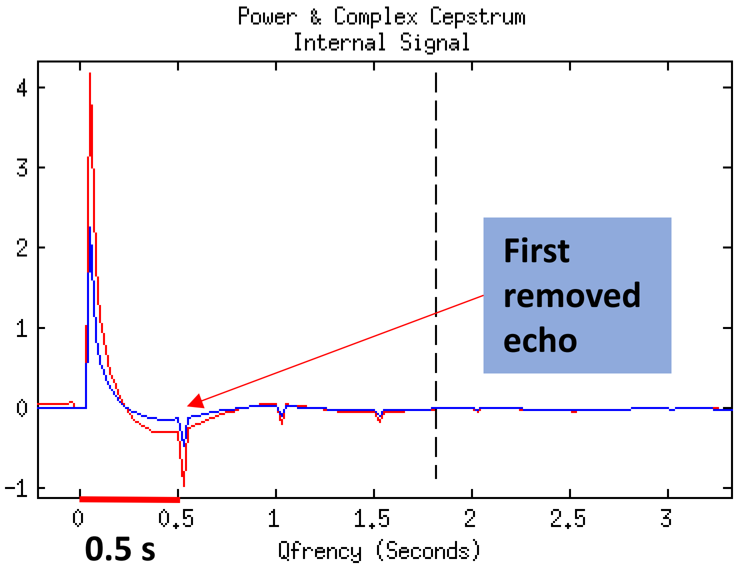
- P and pP are superposed
- P amplitude is affected
- Depending on epicentral distance, these phases are well separated at all stations for earthquakes deeper than 3 km

Pushing the Limits







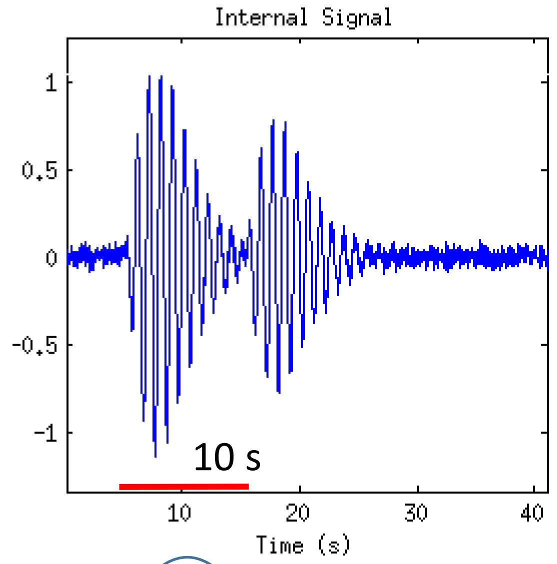


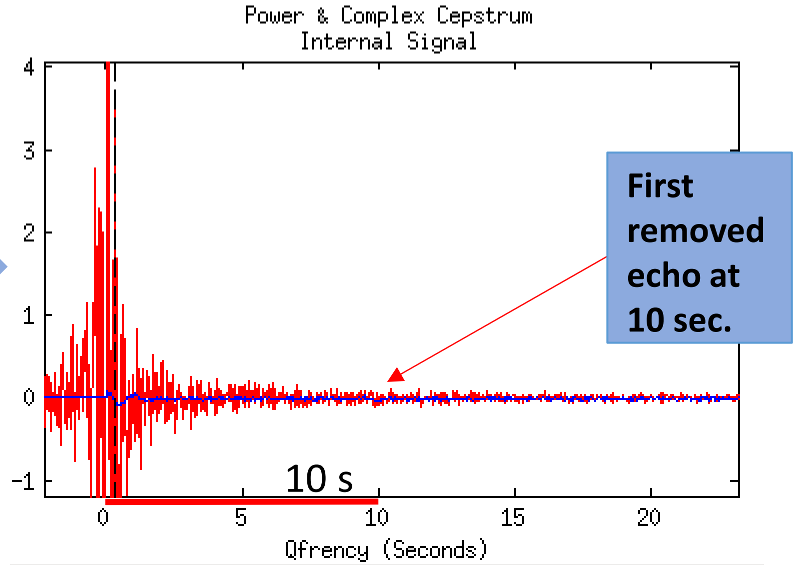
Examples without Noise

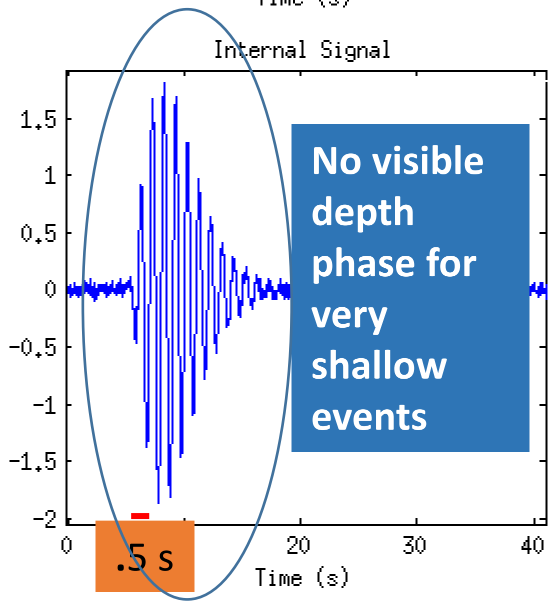
"Perfect" Example of Power and Complex Cepstrum

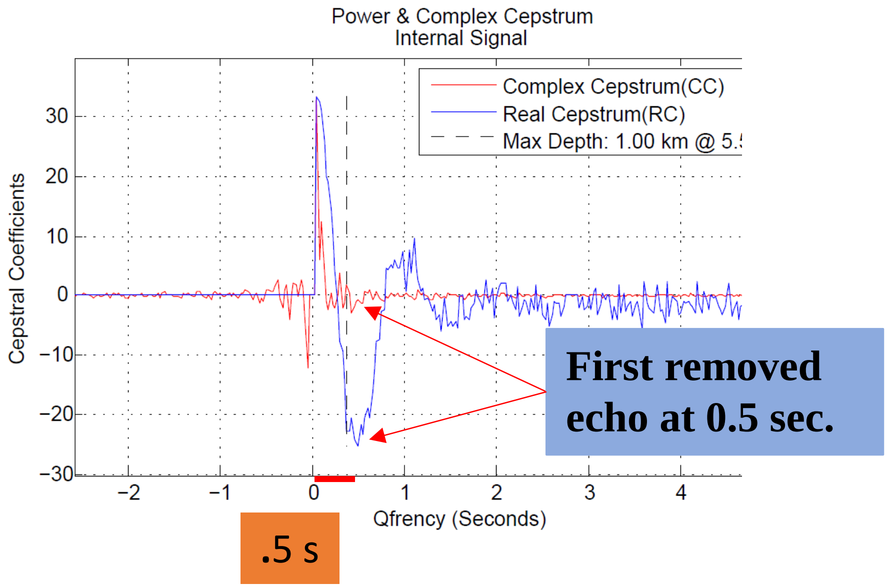
Complex Cepstrum of a Berlage function with an echo similar to the initial wavelet, opposite polarity and 70% reduced amplitude, delayed 15s (above) and 0.5s (below). All the peaks are negative (if the echo has opposite polarity), and the Power and Complex Cepstrums are coincident and of negative sign.

With Noise - Disaster? Not really









Examples with Noise

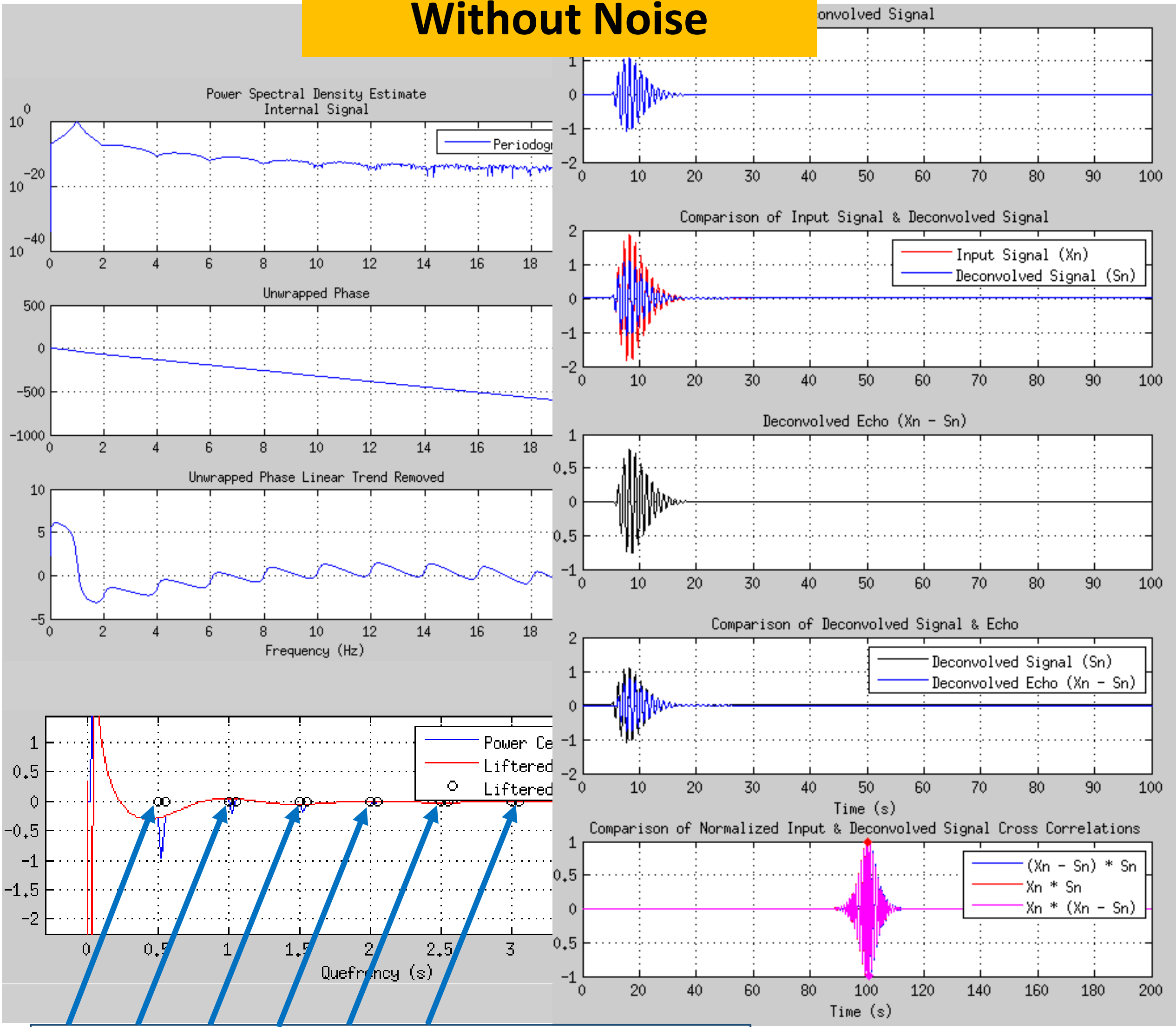
SOLUTION: Although the best peak to lifter is not obvious, the auto-prune feature of the program (WIZARD) lifters every sample and indicates GOOD CANDIDATES for the best deconvolution, based on metric values.

GOOD NEWS:

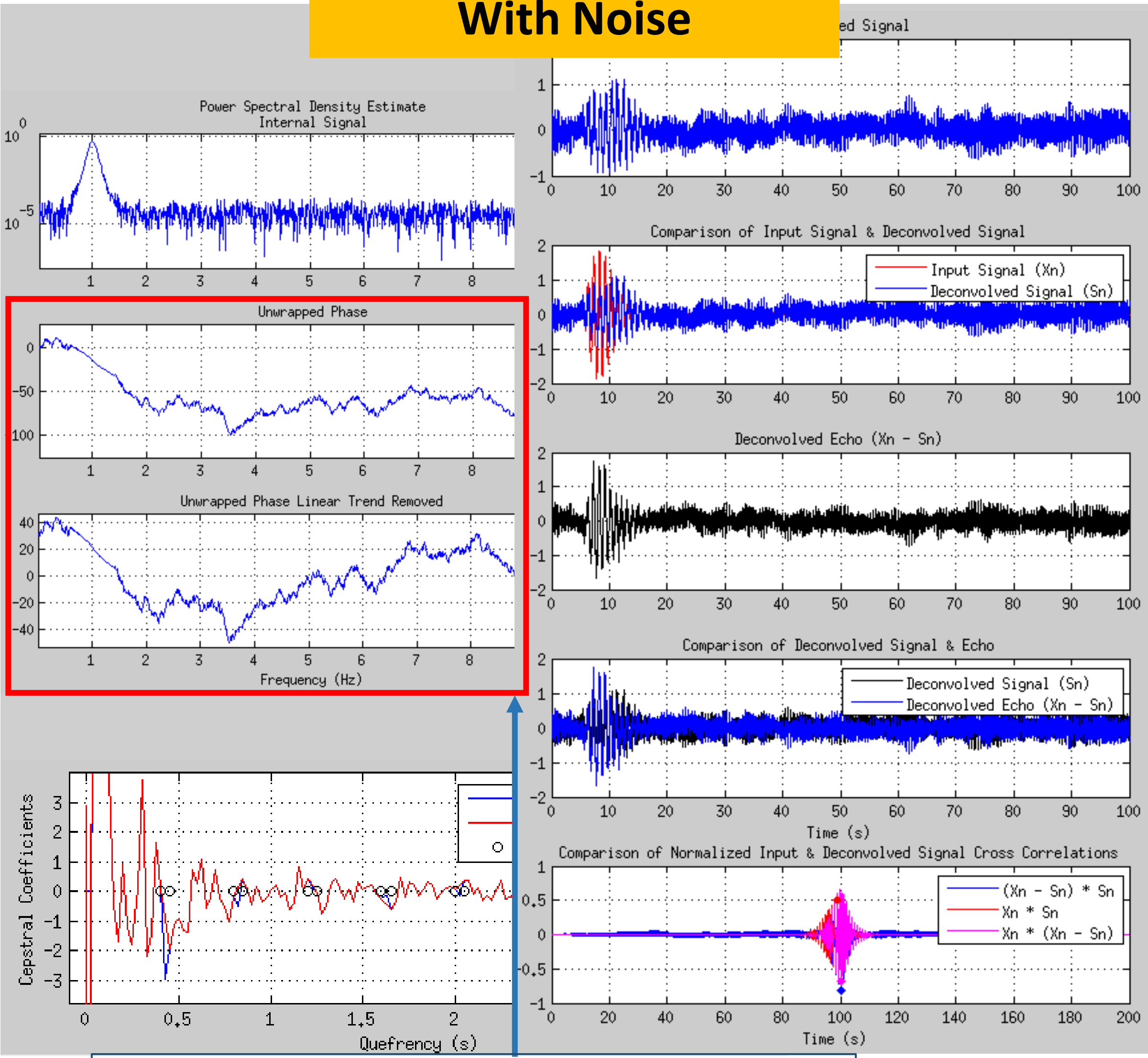
As opposed to signal made of Berlage functions, a real seismic signal has multiple realizations of the depth phase periodicity (when shallow) in the P-coda and multiple crustal /teleseismic phases with similar periodicity, which add in cepstral domain and enhance the cepstral peaks.

Noisy Waveforms Can Still Be Deconvolved

Without Noise



With Noise

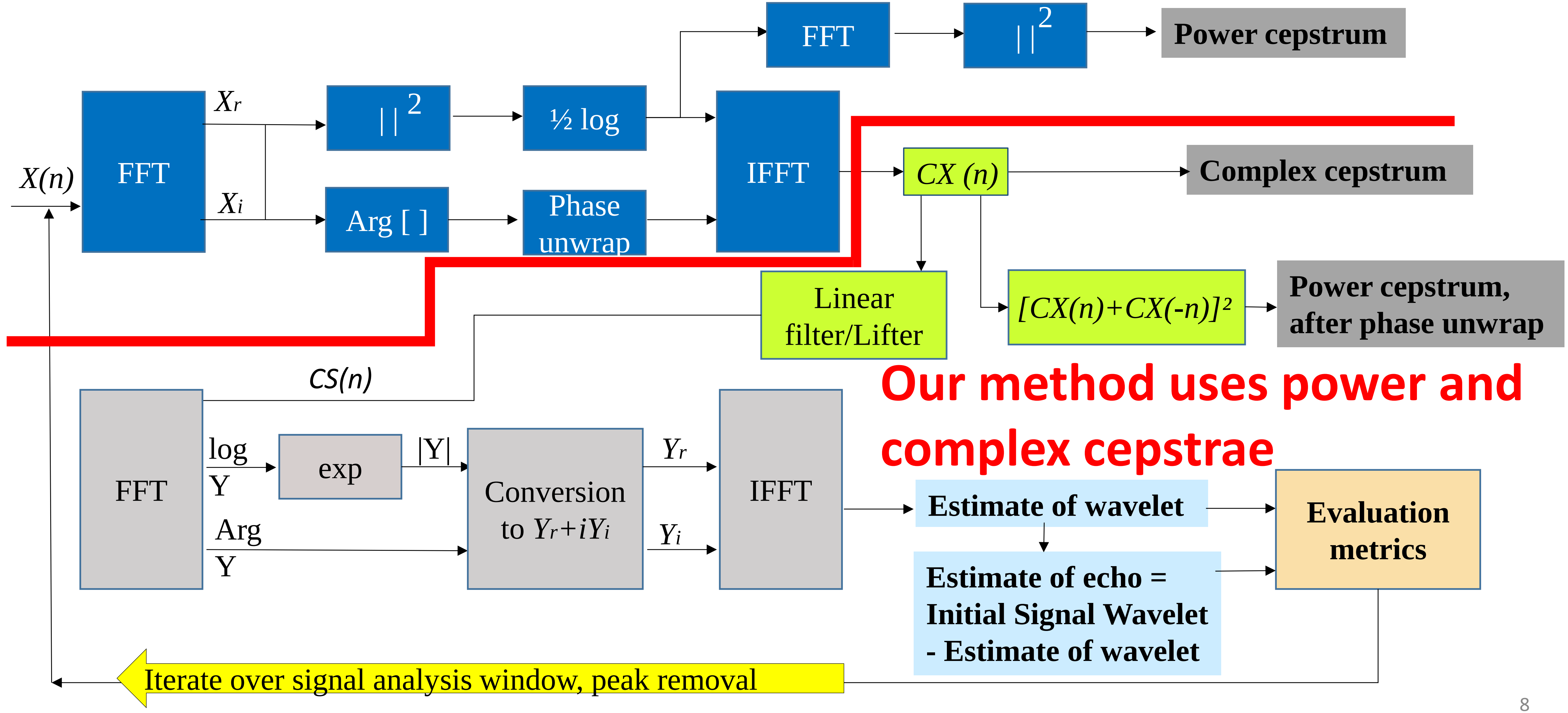


The circles represent the liftered window limits

The lower graph shows the linear trend removed

Cepstral Process Flow Chart

Previous studies used only the power cepstrum



Our method uses power and complex cepstrae

Automatic Cepstral Analysis Tool Modules

Seismic Input Parameter Estimation Module

- Preliminary location, depth range, mechanism
- Seismic phase arrival time prediction
- Source, path, receiver seismic velocity models
- Frequency content, P-phase arrival time and Signal to Noise Ratio (SNR)

Cepstral Analysis Tool

Done for each station

- **Choose a set of analysis windows**
- **Signal – echo time delay**
- **Signal and echo waveforms**
- **Best liftered sample solution score**

Seismic Solution Validation Module

- Array/network pP phase/echo identification
- Analyst validation

Station/Network Depth Estimate

Data for a Textbook Case

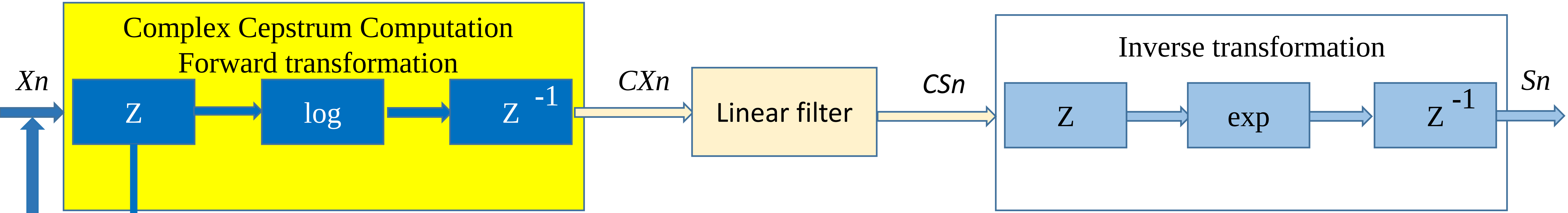
Shock wave data from an explosion in Israel, designed to test the CTBTO-IMS system (1000 sps) recorded at 6.6 km.

This is not a depth phase example but it's a nearly perfect example of an in-phase signal and echo pair from a surface explosion.

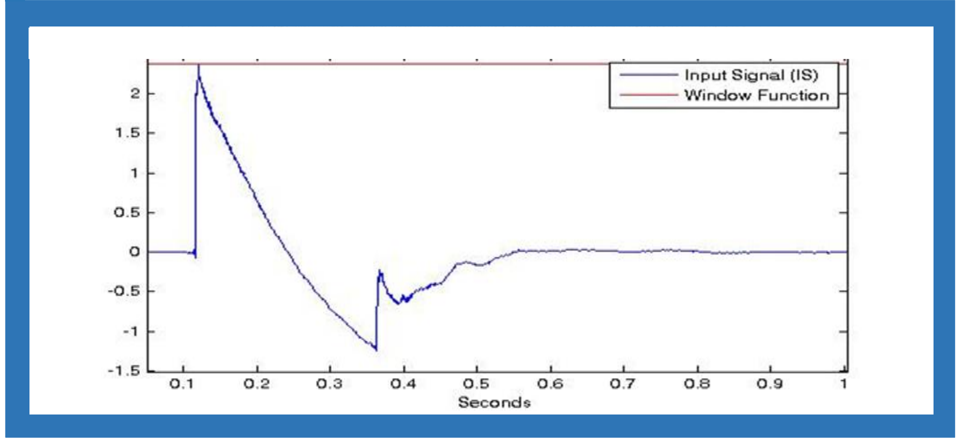
The tests were conducted on February 2011 and recorded at an array of six infrasound sensors located at ~29.9N, 34E.

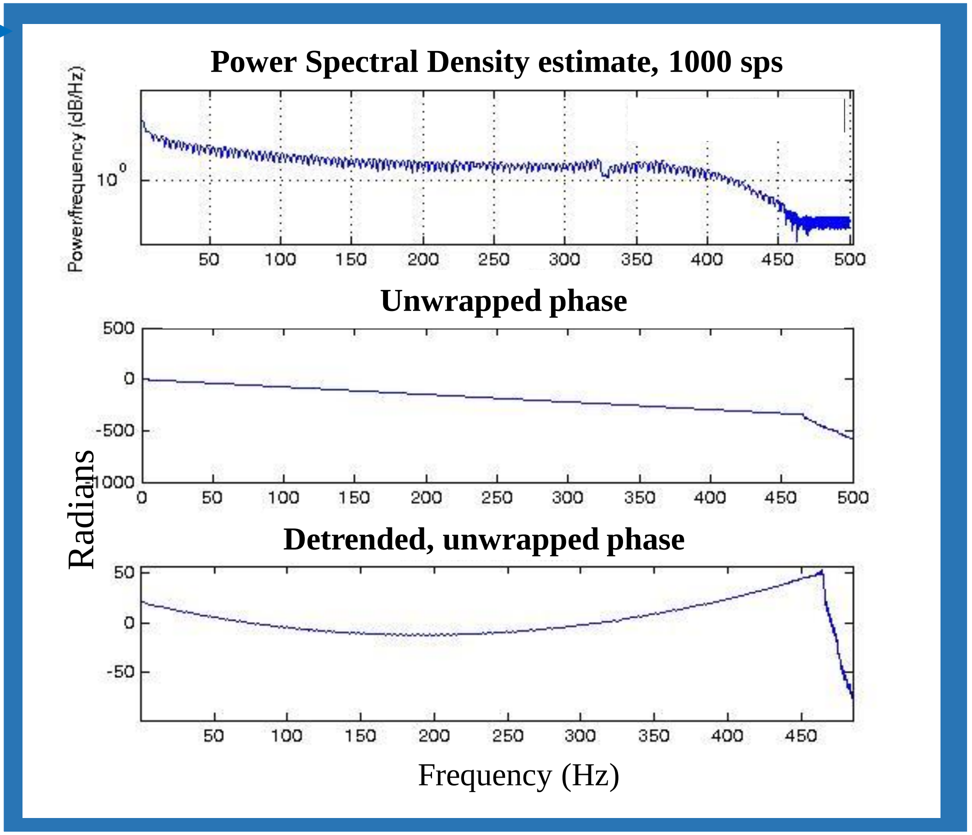
Here we present an analysis of records at station NS2A.

Cepstral Analysis and Metric Application



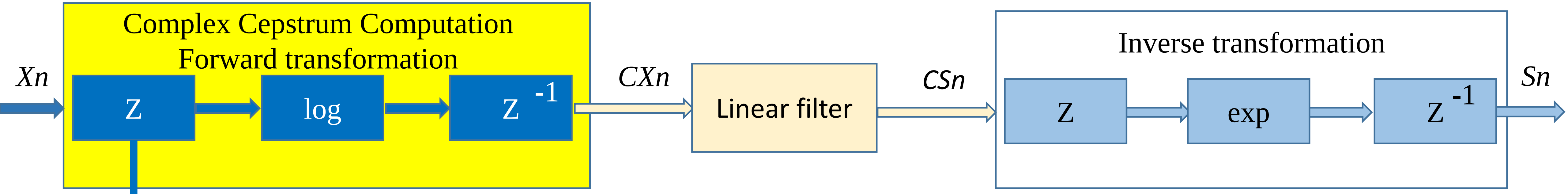
Initial signal





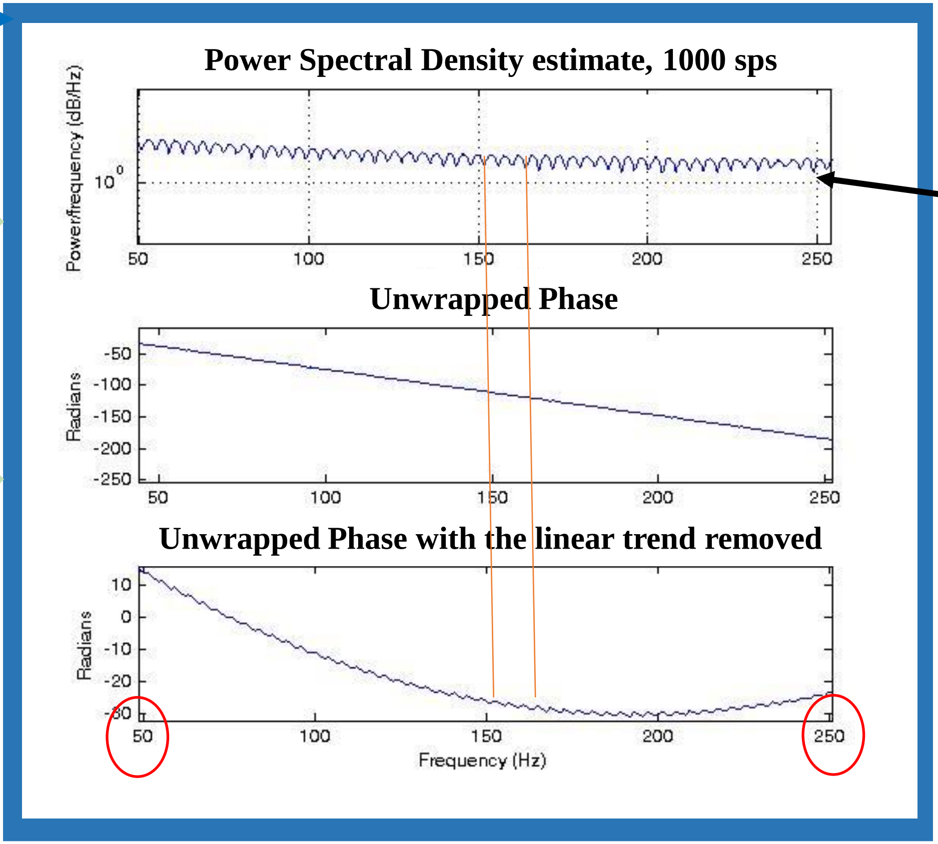
The Power Cepstrum is the power spectrum of the logarithm of the Power spectrum.

Cepstral Analysis Metrics 1 & 2



METRIC1 (BATCH) Quantify the similarity of the initial signal and the signal recovered after phase unwrapping.

METRIC2 (BATCH) Use crosscorrelation (X) to confirm presence of scalloping in the log-spectra and in the unwrapped, detrended phase. Maximum crosscorrelation is the weight and METRIC2 is the weight multiplied by the power cepstrum of the crosscorrelation X for a chosen quefrency interval.

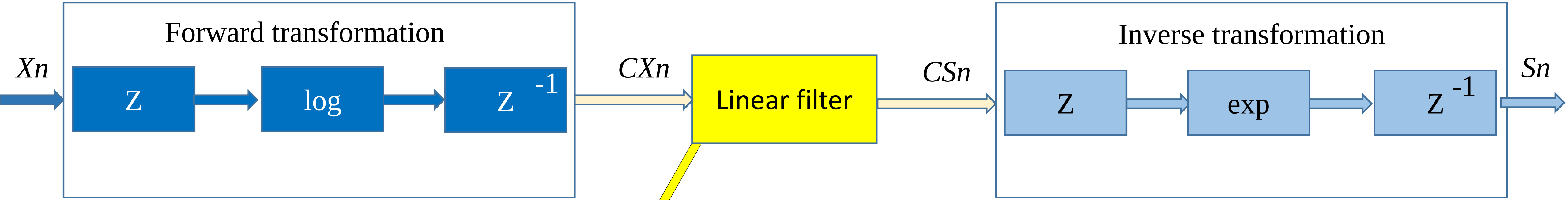


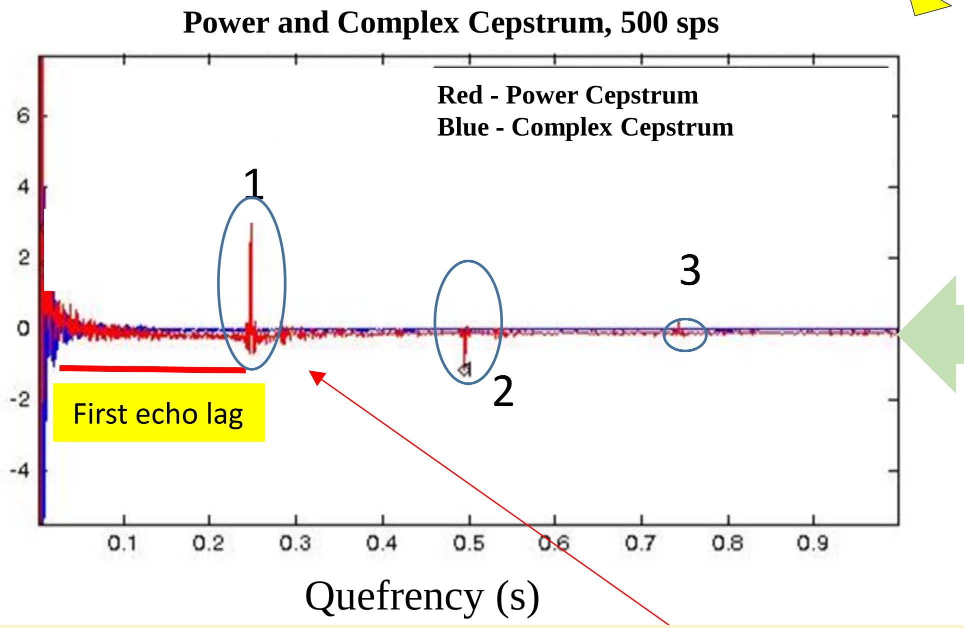
Magnitude and phase scalloping are observed when an echo is present

WS = WINDOW SCORE

CLS = CHOSEN LIFTERED SAMPLE

Cepstral Analysis Metric 3





METRIC3 (BATCH)

Quantifies the minimum phase approximation validity

METRIC3= Metric3_weight_energy *Metric3_weight_PC^2_CC^2

where

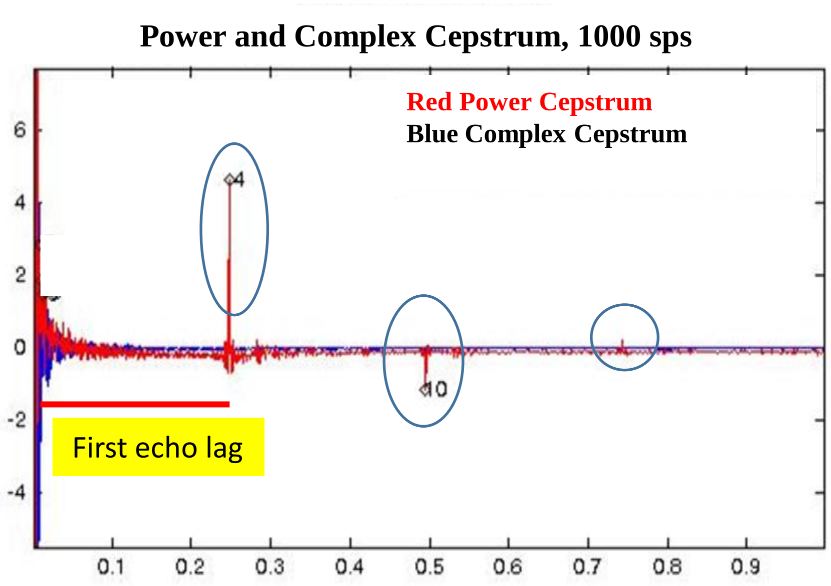
$$\mathbf{WEIGHT}= 1/[1 + \frac{\sum_{i=1}^{N}[(CC_{i,left}[\tau])^2],\ for\ \tau < 0}{\sum_{i=1}^{N}[(CC_{i,right}[\tau])^2],\ for\ \tau > 0}]$$

where $CC_{i,left}[\tau]$ is the CC with negative quefrency and $CC_{i,right}[\tau]$ is the CC with positive quefrency.

METRIC3=WEIGHT* $1/[1 + |\sum_{i=1}^{N}[RR_i - CC_i]|]$, where RR_i is the demeaned PC squared scaled to 1 and CC_i is the sum of the $CC_{i,left}[\tau]$ added to $CC_{i,right}[\tau]$, squared, demeaned and scaled to 1.

The Power and Complex Cepstrum should be equal and entirely positive for a minimum phase signal, and would have peaks at the same lags after ideal phase unwrapping. The location of the highest Complex Cepstrum (*CXn*) (positive in this case) peak due to the echo should also correspond to the largest *CXn* amplitude. The energy in the power cepstrum should be equal to the sum of the energy in the complex cepstrum (at positive and negative quefrency).

Cepstral Analysis Metrics 4 & 5



METRIC4 (BATCH): When the signal is the closest approximation of a minimum phase signal, the power and complex cepstra are very similar, with similar peaks on the positive quefrency axis. This metric uses the PC and CC, only values for positive quefrency, each scaled to 1. The metric WEIGHT is the product of:

- **Max**$[\int_{-\infty}^{\infty} \boldsymbol{CC}[\boldsymbol{\tau}]\ \boldsymbol{PC}[\boldsymbol{\tau}+\boldsymbol{t}]\boldsymbol{d\tau}]$, i.e. the Cross Correlation of CC and PC or *xcorr(CC,PC)*
- a ratio of a sum of the number of times the *diff(CC)* and *diff(PC)* have the same sign and the total number of PC (or CC) samples, where *diff* is a MATLAB function name for the differential of a time series.

METRIC4=WEIGHT*metric4_wf, where **metric4_wf=1** if the differential PC and CC signs at the respective quefrency coincide, and zero otherwise.

METRIC5 (BATCH) The crosscorrelation of the PC and CC autocorrelations should be a good indicator of the position of the first cepstral peak related to the echo, on the positive quefrency axis. The first peak occurs at the echo-signal delay, followed by repeating peaks at the same quefrency intervals. This periodicity should be captured by the PC and CC autocorrelations. The PC and PC autocorrelations are scaled to 1 and detrended prior to crosscorrelation.
WEIGHT=max(xcorr(xcorr(PC),xcorr(CC), 'coeff')); *xcorr* is the MATLAB function for crosscorrelation.
METRIC5_CLSS=WEIGHT* abs(xcor(CC).*xcorr(PC)) for positive quefrency

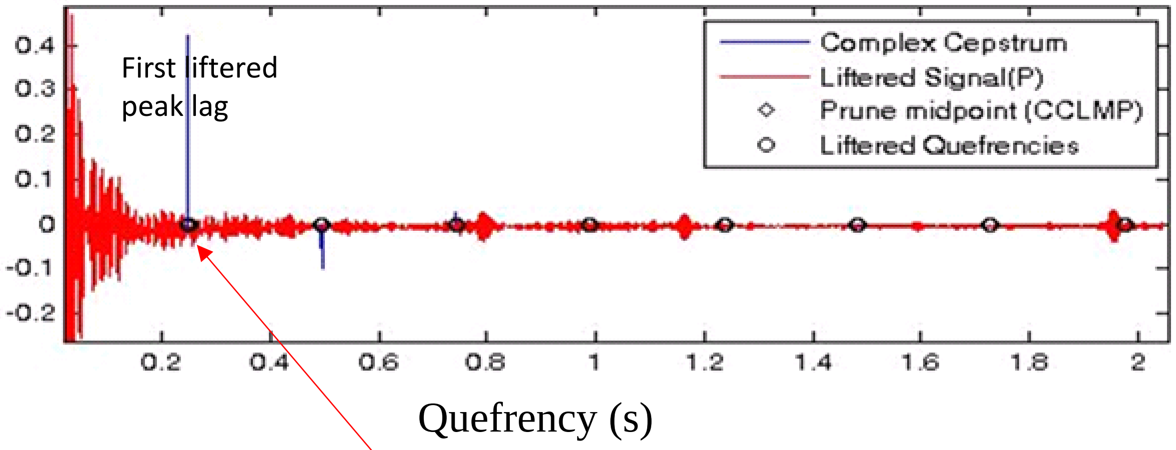
Cepstral Analysis Metrics 6, 7 & 8

METRIC6 (WIZARD, BATCH). Quantifies the polarity in the first CXn liftered peak. Ensures that the polarity of the first peak corresponds to the deconvolved echo-signal polarity. The WIZARD Metric 6 does not have a weight. In BATCH, the weight ensures that the CC liftered peak aligns with a PC peak.

The BATCH WEIGHT is empirically estimated as:

$$\textbf{Weight} = \frac{3}{4} * \frac{\sum_{i=1}^{N}\left[CC_{(diff,liftered,peak)i} * PC_{(diff,liftered,peak)i}\right]}{\sum_{i=1}^{N}\left[\left|CC_{(diff,liftered,peak)i}\right| * \left|PC_{(diff,liftered,peak)i}\right|\right]}$$

- $CC_{(diff,liftered,peak)i}$ is estimated in a window (3 samples here) around the liftered sample as the sign of the differential of the values in the CC window and
- $PC_{(diff,liftered,peak)i}$ is estimated in a similar way in a PC window.



The blue liftered *CXn* peak sign is positive when a similar polarity echo has lower amplitude than the *Sn*.

METRIC7 (WIZARD, BATCH): quantifies liftered (pruned) and deconvolved echo lag-time similarity

- $Lag_{metric} = 0.00001 \; if \; pruned_{midpoint,sps} < 5$
- $Lag_{metric} = \frac{1}{\left[\left|pruned_{midpoint,sps} - (echo\ lag)\right| + 1\right]^3}$; where

$\left|pruned_{midpoint,sps} - (echo\ lag)\right|$ is the difference between the time lag of the liftered/pruned sample (of the first cepstral peak) and the signal-echo time lag recovered from deconvolution.

METRIC8_CLSS (BATCH) is equal to 1/(delta_lag+1). The weight of this metric is 1 if delta_lag is less than an analyst-chosen value, which here is 1, and 0.01 otherwise.

Cepstral Analysis and Metrics 9 & 10

METRIC9 (WIZARD indicator only, BATCH) is equal to the square root of the Xn-Sn and Sn energy ratio in an analyst chosen window. WEIGHT=1 for (empirical) values of the metric between 0.15 and 0.97, otherwise is 0.00001.

METRIC10 (WIZARD, BATCH): If the crosscorrelations of the Sn, Xn and Xn-Sn meet the following expectations:

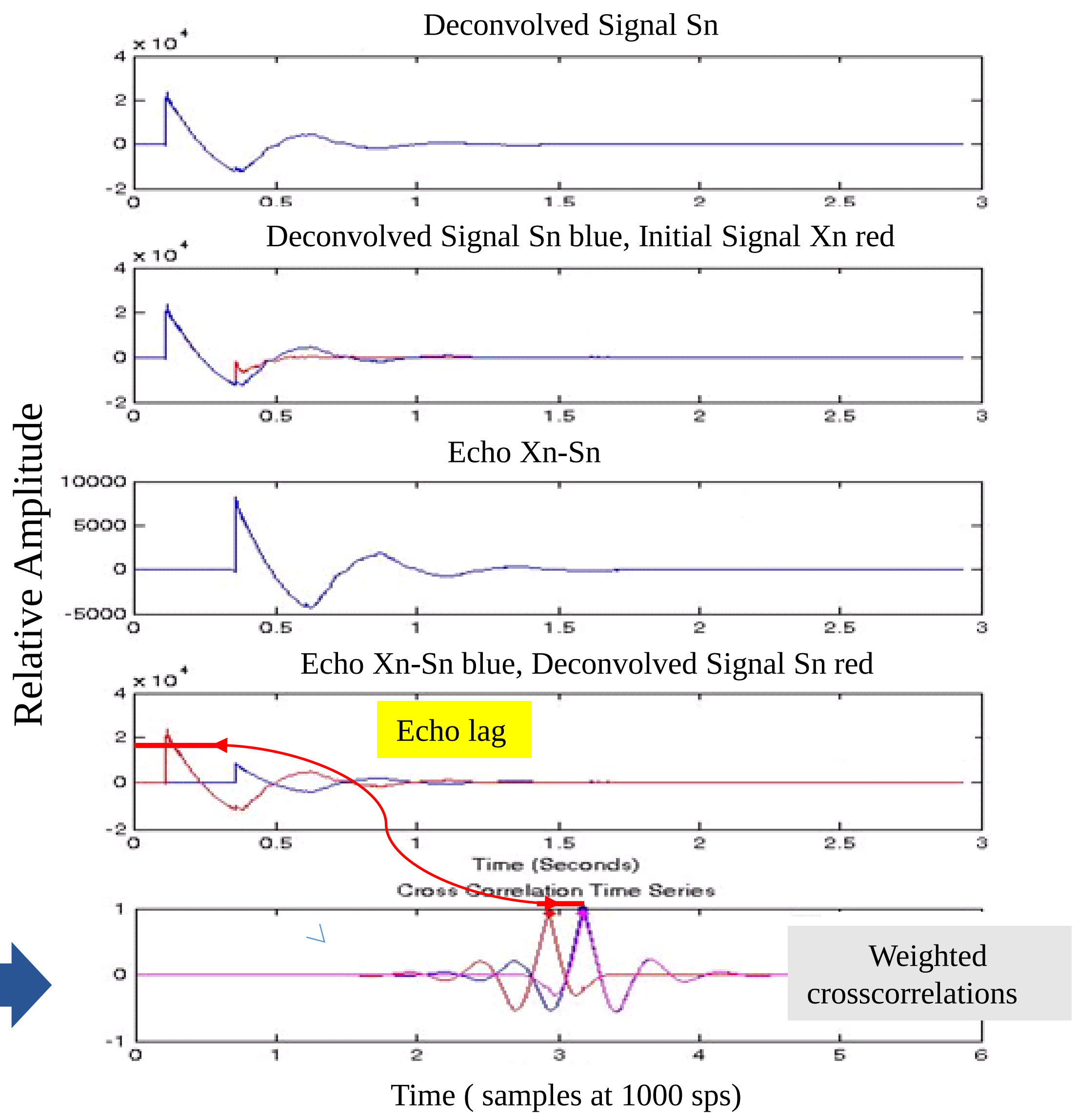
- **0.7< xcorr(Sn,Xn) < 0.95 (0.2< xcorr(Sn,Xn) < 1 in WIZARD)**
- **0.7< xcorr[Sn,(Xn-Sn)]<0.97**

Then METRIC10=1 else METRIC10=0.00001

WEIGHT = the maxim value of the deconvolved echo (Xn-Sn) and signal (Sn) crosscorrelation

Sn: First arrival deconvolved after Complex Cepstrum Lifter;
Xn: Original signal;
Xn-Sn: Echo hypothesis.

Narrow crosscorrelation peaks show high signal to echo similarity.



Cepstral Analysis Metrics 11, 12 & 13

METRIC11_CLSS (BATCH):

$$\frac{\left[\int_{-\infty}^{\infty}(\boldsymbol{x_n}-\boldsymbol{S_n})[\boldsymbol{\tau}]\,\boldsymbol{S_n}[\boldsymbol{\tau}+\boldsymbol{t}]\boldsymbol{d\tau}\right]_{\boldsymbol{maximum}}}{\left[\int_{-\infty}^{\infty}\boldsymbol{x_n}[\boldsymbol{\tau}]\,\boldsymbol{S_n}[\boldsymbol{\tau}+\boldsymbol{t}]\boldsymbol{d\tau}\right]_{\boldsymbol{maximum}}}$$

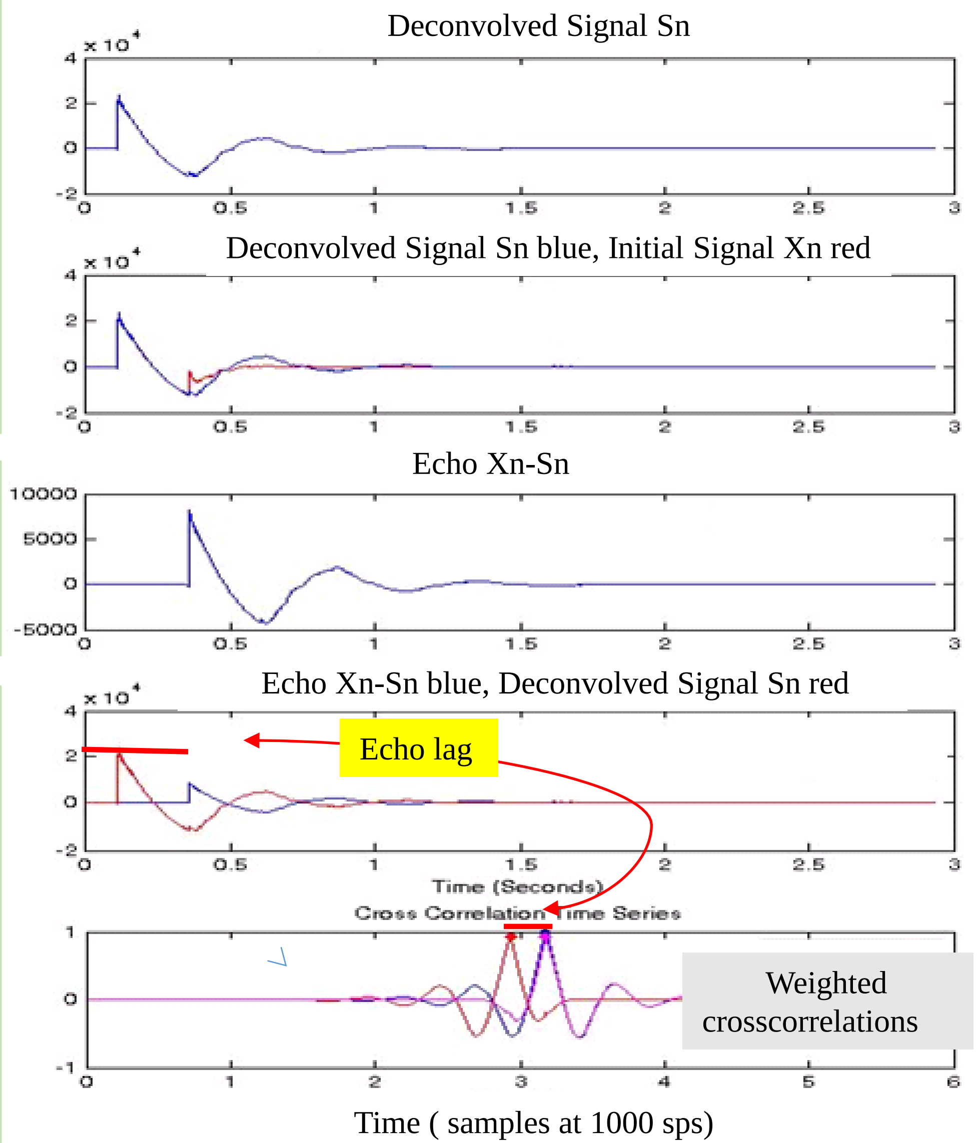
The WEIGHT of this metric is 1, unless the maxim value of the normalized Xn and Sn crosscorrelation is greater than the maxim value of the deconvolved echo (Xn-Sn) and signal (Sn) crosscorrelation, in which case the weight is 0.5.

METRIC12 (BATCH): is the median of the PC and CC of the data autocorrelations, scaled to maximum amplitude 1. **This is the most effective metric so far.** WEIGHT=1.

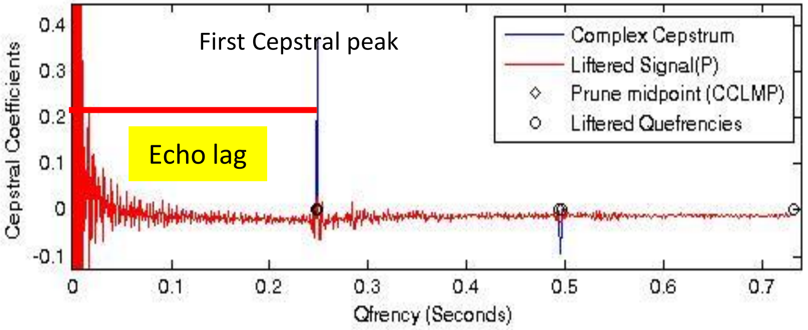
METRIC13 (BATCH): is **WEIGHT*(1-XP)** where XP is the maxim of the crosscorrelation of a unique, minimum phase sequence, named REC_PC, with the same power cepstrum as the analyzed time series, provided as an output of the *rceps* MATLAB@ function and the deconvolved signal (Sn). WEIGHT=1;

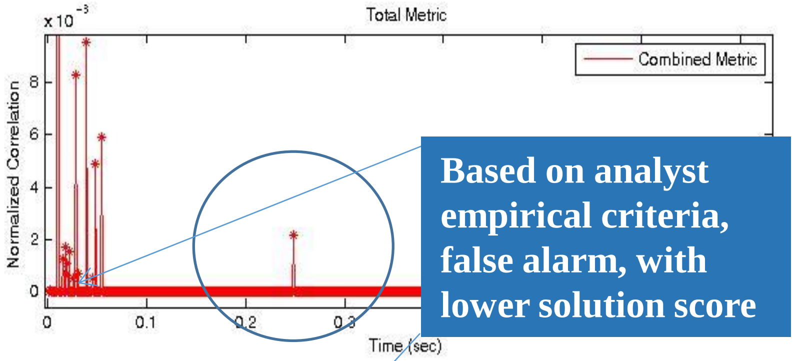
$$\boldsymbol{XP}=\left[\int_{-\infty}^{\infty}\boldsymbol{REC_{PC}}[\boldsymbol{\tau}]\,\boldsymbol{PC_{rceps,deconvolved\,S_n}}[\boldsymbol{\tau}+\boldsymbol{t}]\boldsymbol{d\tau}\right]_{\boldsymbol{maximum}}$$

METRIC13 should decrease once the echo is effectively removed from the initial signal.

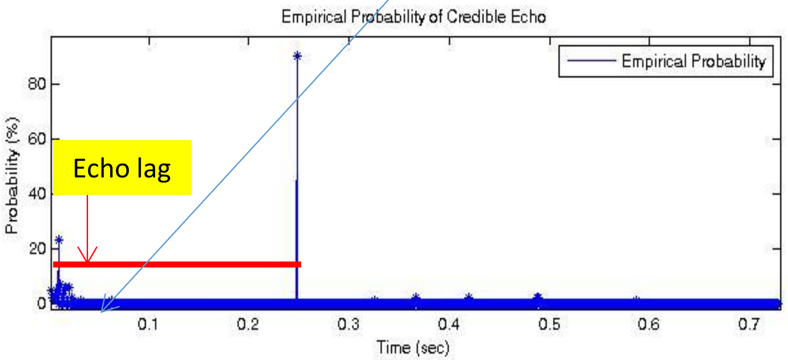


WIZARD RESULTS Semi-automatic cepstral liftering analysis and evaluation of the signal Sn and echo Xn-Sn deconvolution





Based on analyst empirical criteria, false alarm, with lower solution score



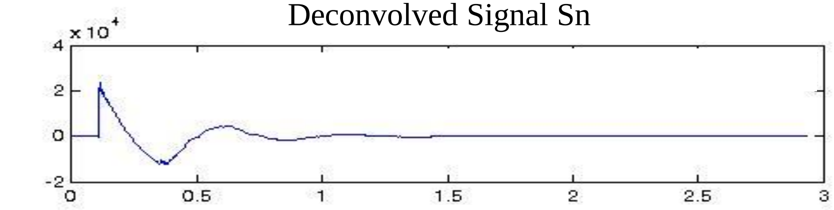
Sn: First arrival deconvolved after CC Lifter
Xn: Original signal   Xn-Sn: Echo hypothesis

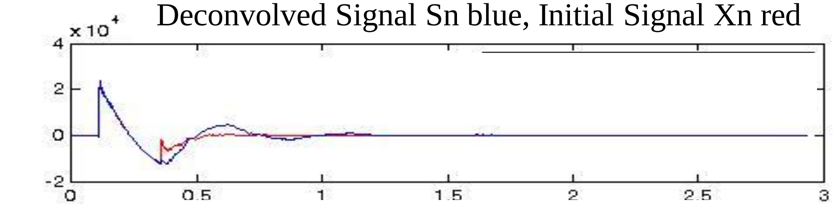
REPORT

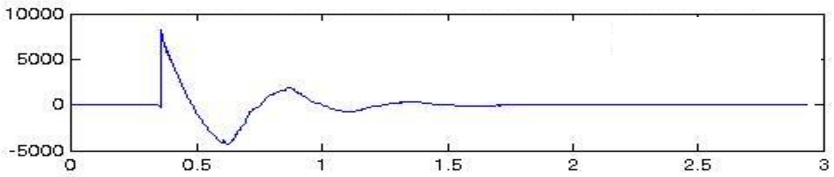
Pruning: Semi-Automatic
Time-range (s): 0.247-0.249
Sample-range (samples): 247-249
Estimated echo time delay: 0.248s

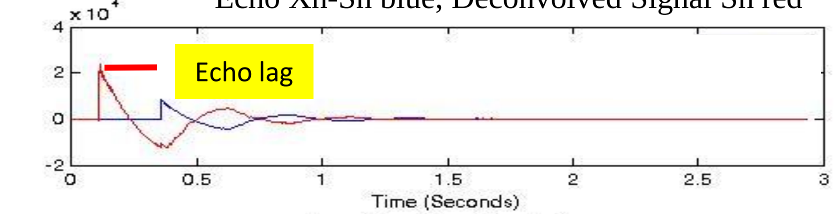
Correlation (Xn*Sn): 0.94
Correlation ((Xn-Sn)*Xn):0.92
Correlation:((Xn-Sn)*Sn): 0.99
Correlation Ratio
Xn*Sn / ((Xn-Sn)*Sn): 1.1
Cross-Correlation Lag (Expected – Estimated) = 1 sample

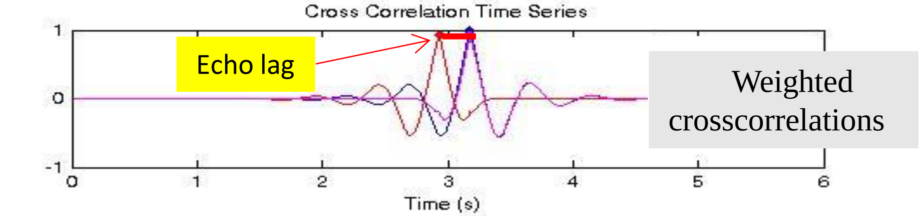
The solution score is obtained from the total metric, using empirical analyst-estimated criteria











Why do we have a WIZARD and a BATCH code?

- **The initial interactive code was named WIZARD.**
  - **This code is properly utilized by an experienced analysist with prior knowledge and experience of processing difficult signals of interest.**
- **The BATCH code was created for fast processing of multiple windows, by stripping the WIZARD and adding new metrics.**

Metrics: WIZARD vs BATCH

WIZARD METRICS or GOOD CANDIDATE Metrics

Less restrictive, allow multiple echo interpretations, enhanced analyst contribution;

Ensure that the echo lag is recovered;

Ensure that the deconvolved echo polarity (+ or -) is consistent with the liftered first peak sign (positive or negative);

Ensure that the pP/P similarity is greater than the P/SI and pP/Si similarity;

Encourage results with large pP/P ratio;

Selects a set of Good Candidates based on the above criteria;

The analyst selects the CHOSEN liftered sample, however, DOES NOT restrict polarity

BATCH METRICS – To be automated

Use the Wizard metrics (GOOD CANDIDATE METRICS), except for the pP/P ratio metric, which is used as a final solution score indicator;

Use additional, new BATCH metrics;

Restrict the polarity of the echo;

Restrict similarity values for pP, P and SI;

Are estimated in a large set of windows of variable length and noise before P, set by an analyst.

Data (Nuclear Explosion)

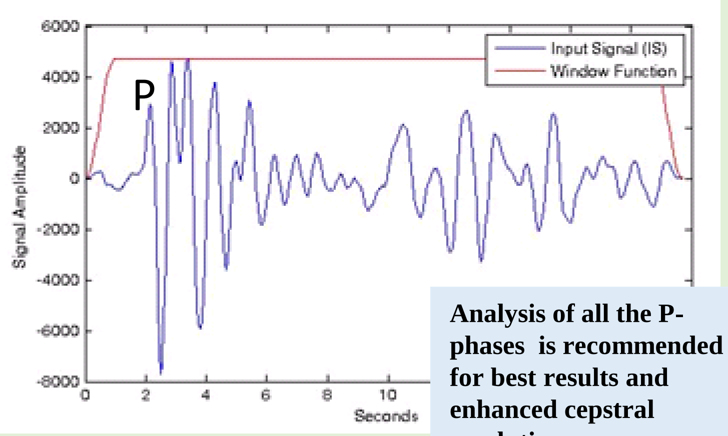
The Bullion Nevada Test Site shaft nuclear explosion, detonated on June 13, 1990, 16:00:00, 37.2615N, 116.4210W, at 2211 ft depth of burial (674 m), 20-150kt yield (United States Nuclear Tests, June 1945-Sept.1992, DOE/NV –209-REV 16, Sept . 2015, pp 124)

The underground nuclear test was detonated at the Silent Canyon Caldera that is within the Pahute Mesa of the Nevada Test Site (NTS).

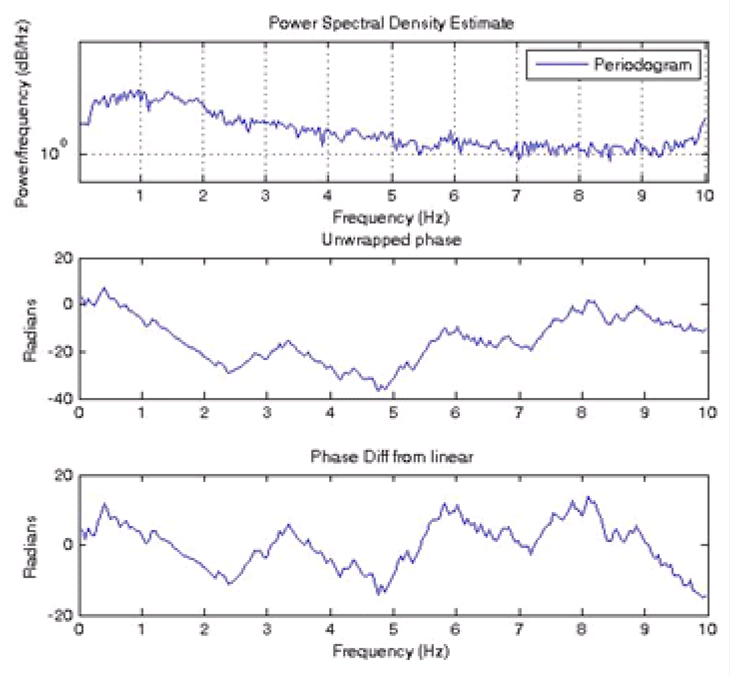
The source geology of this specific nuclear test is hardrock, which has a low porosity feature and a P-wave velocity of 3.3 km/sec.

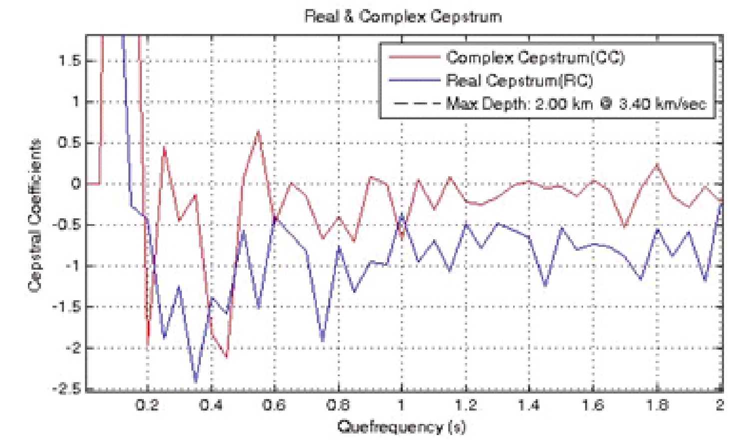
WIZARD CODE, KONO, Bullion, NTS

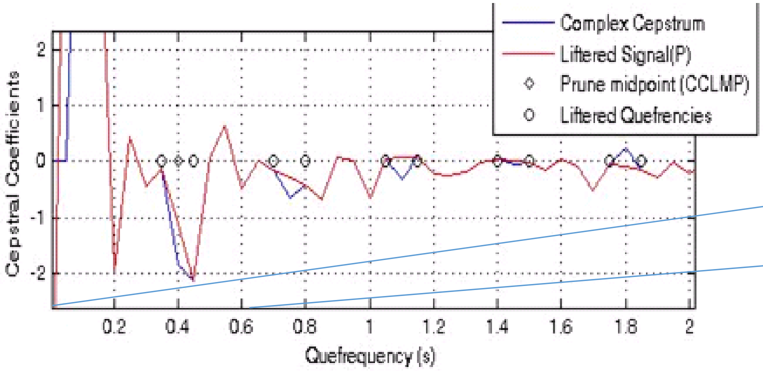
Seismic station KONO, Kongsberg, Norway, 59.64 N, 9.59 E, at 73.6 deg epicentral distance



Analysis of all the P-phases is recommended for best results and enhanced cepstral resolution.

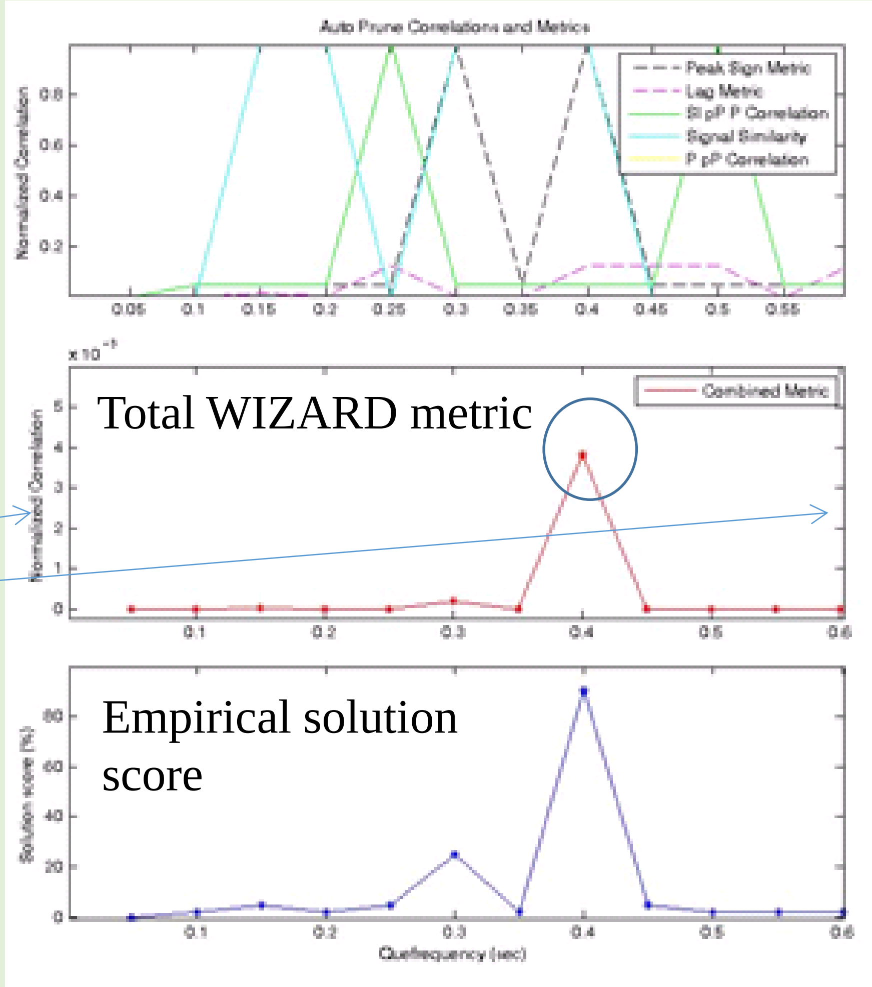






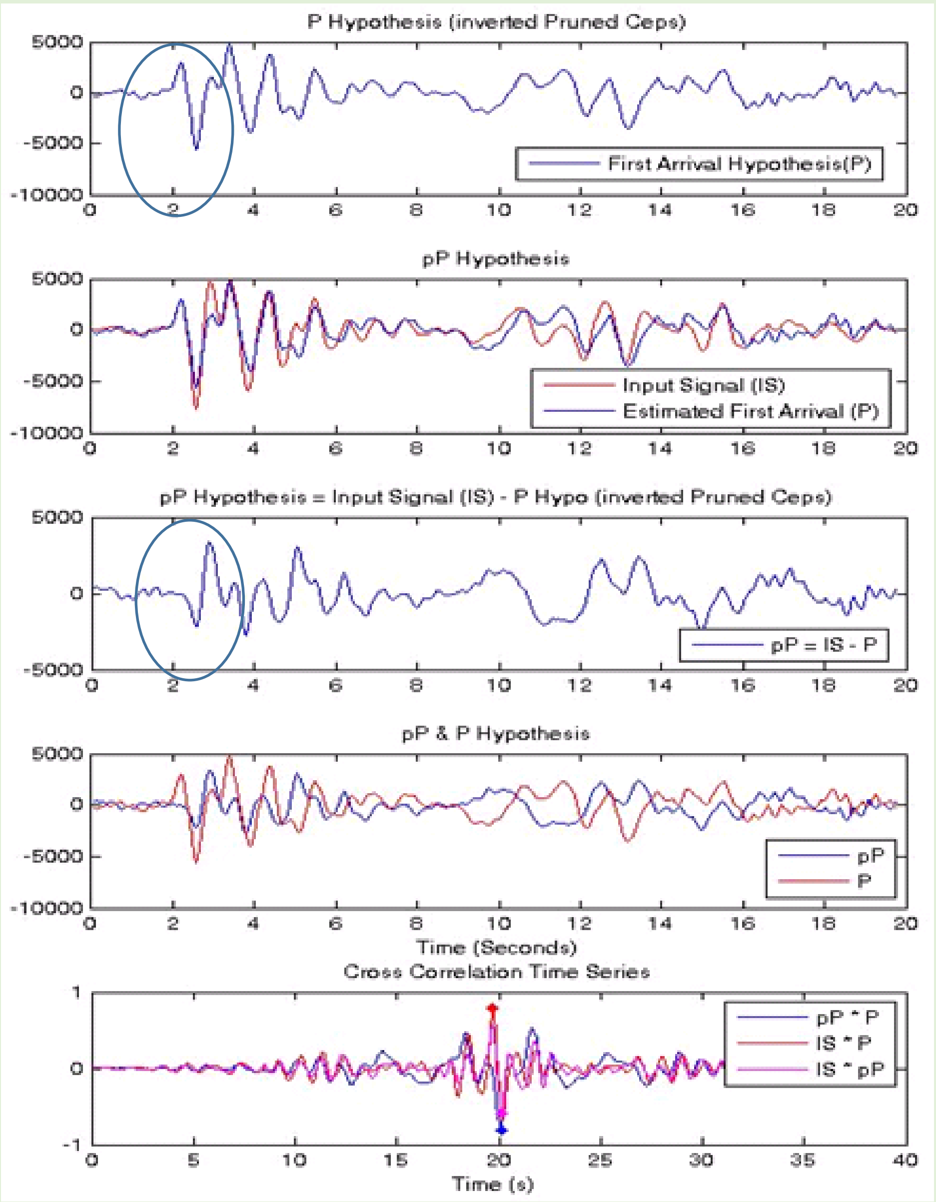
The most likely echo delay has largest combined metric value, and largest solution score

GOOD CANDIDATE METRICS
The solution score is obtained from the total metric, using analyst-estimated criteria



WIZARD CODE, KONO, Bullion, NTS

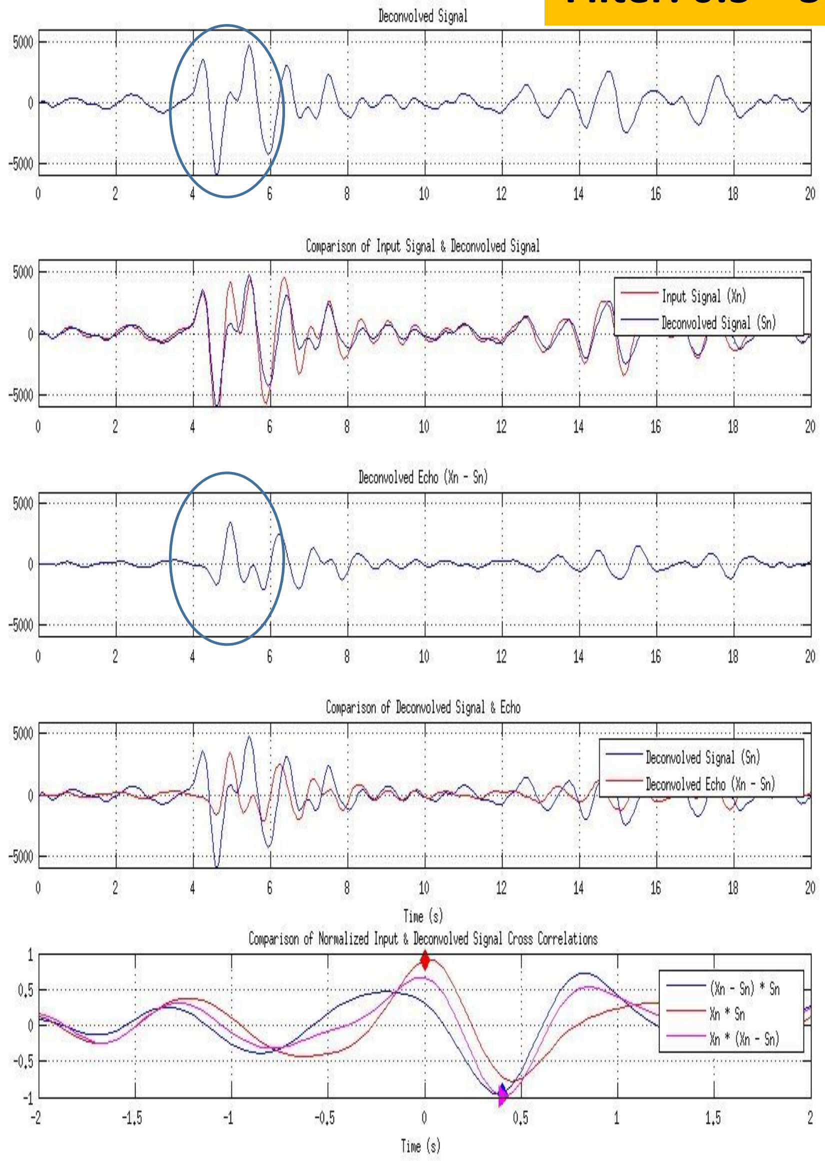
Reported Depth of Burial: 674 m
Our Estimate: 660 m
(at 3.3 km/s source velocity)
The sample interval-related error at 20sps is ~ 82.5m



BATCH

Filter: 0.5 – 3 Hz

WIZARD vs BATCH COMPARISON TABLE	WIZARD FINAL RESULT	BATCH FINAL RESULT
Depth Range (meters)	595 - 765	595 - 765
Estimated Echo Time Delay (seconds)	0.4	0.4
Sample Range (samples)	7.0 - 9.0	7.0 - 9.0
Time Range (seconds)	0.35 - 0.45	0.35 - 0.45
Correlation #1:(Xn*Sn):	0.78	0.91
Correlation #2:((Xn-Sn)*Sn):	-0.81	-0.94
Correlation #3:((Xn-Sn)*Xn):	-0.6	-0.99
Correlation Ratio (#2/#1):	-1.04	-1.03
Power(Xn-Sn)/Power(Sn)	0.74	0.55
Cross-Correlation Lag (Expected - Estimated)	1 sample at 20sps	1 sample at 20sps



Reported Depth of Burial: 674 m **Our Estimate: 660 m**

At 3.3 km/s source P velocity. The error at 20 sps is ~ 82.5 m

Summary

- **Homomorphic deconvolution and solution evaluation metrics have been implemented and initial testing shows that some metrics perform well for a limited number of cases.**
- **Cepstral resolution depends on parameters such as waveform analysis window length, sample rate, signal seismic phase content, signal-to-noise ratios and frequency content. Selecting the best waveform analysis window by hand is a long and tedious process, which we seek to automate.**
- **Thus, we have designed analysis metrics, currently under evaluation.**
- **The BATCH code relies on empirically set parameter bounds and on metrics which will be also tested using synthetic data.**
- **The phase unwrapping algorithms and the liftering choices significantly affect the complex cepstrum shape and resolution, as well as the homomorphic deconvolution, thus are the subject of on-going research.**

SUPPLEMENTARY SLIDES

Sample Rate Discussion

DATA

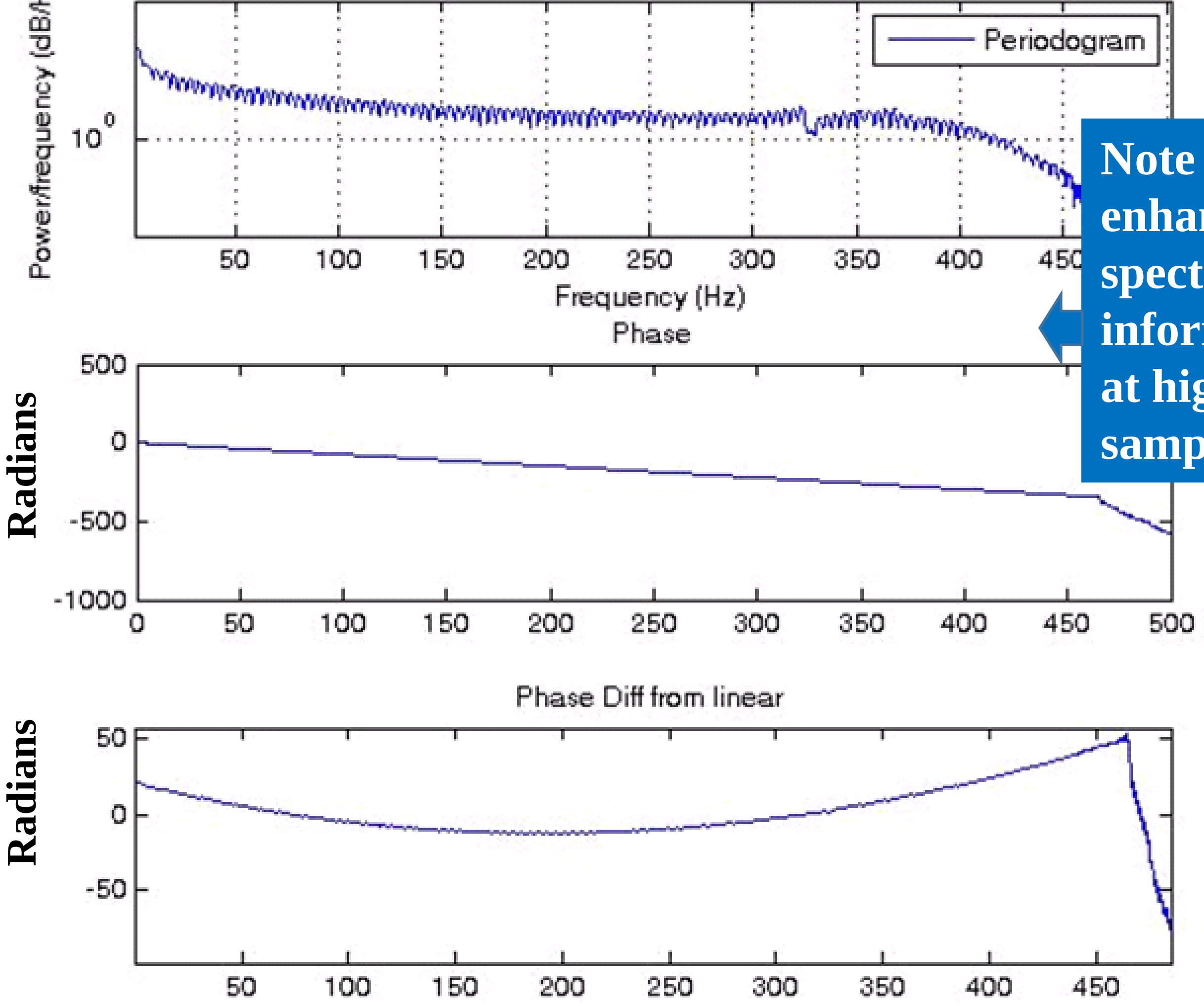
Shock wave data from an explosion designed to test the CTBTO-IMS system (1000 sps vs 100 sps). The tests were conducted at the surface, in February 2011 and recorded at an array of six infrasound sensors located at 29.9N, 34E. Here we present an analysis of records at station NS2A.

A higher sample rate results in lower phase unwrapping errors, however, also results in ambient noise phase distortion at higher sample rates, and lower Signal-to-Noise ratios. Preliminary investigations show that filtering is a possible solution to this problem.

Close-in infrasound data 10-ton surface explosion test of the CTBTO-IMS, 2011

1000 sps

Power Spectral Density
NS2A GDF 1

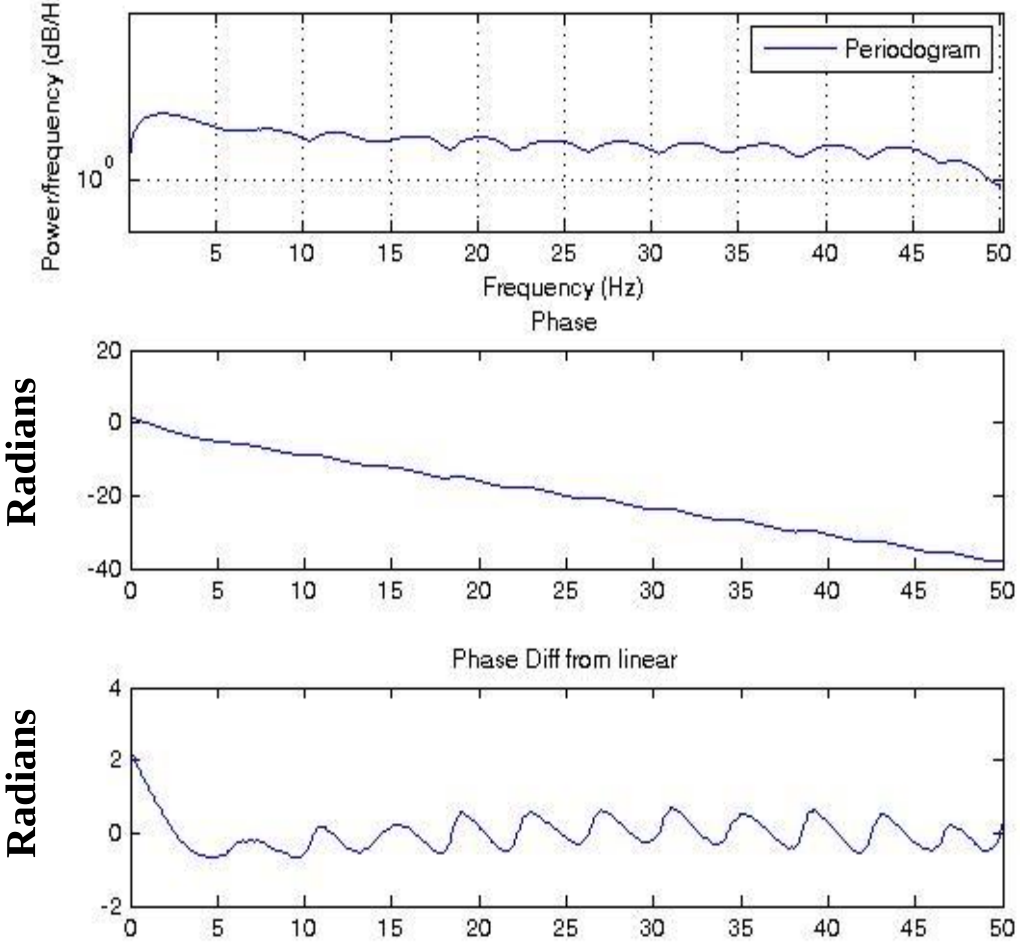


Frequency (Hz)

Note enhanced spectral information at higher sample rate

100 sps

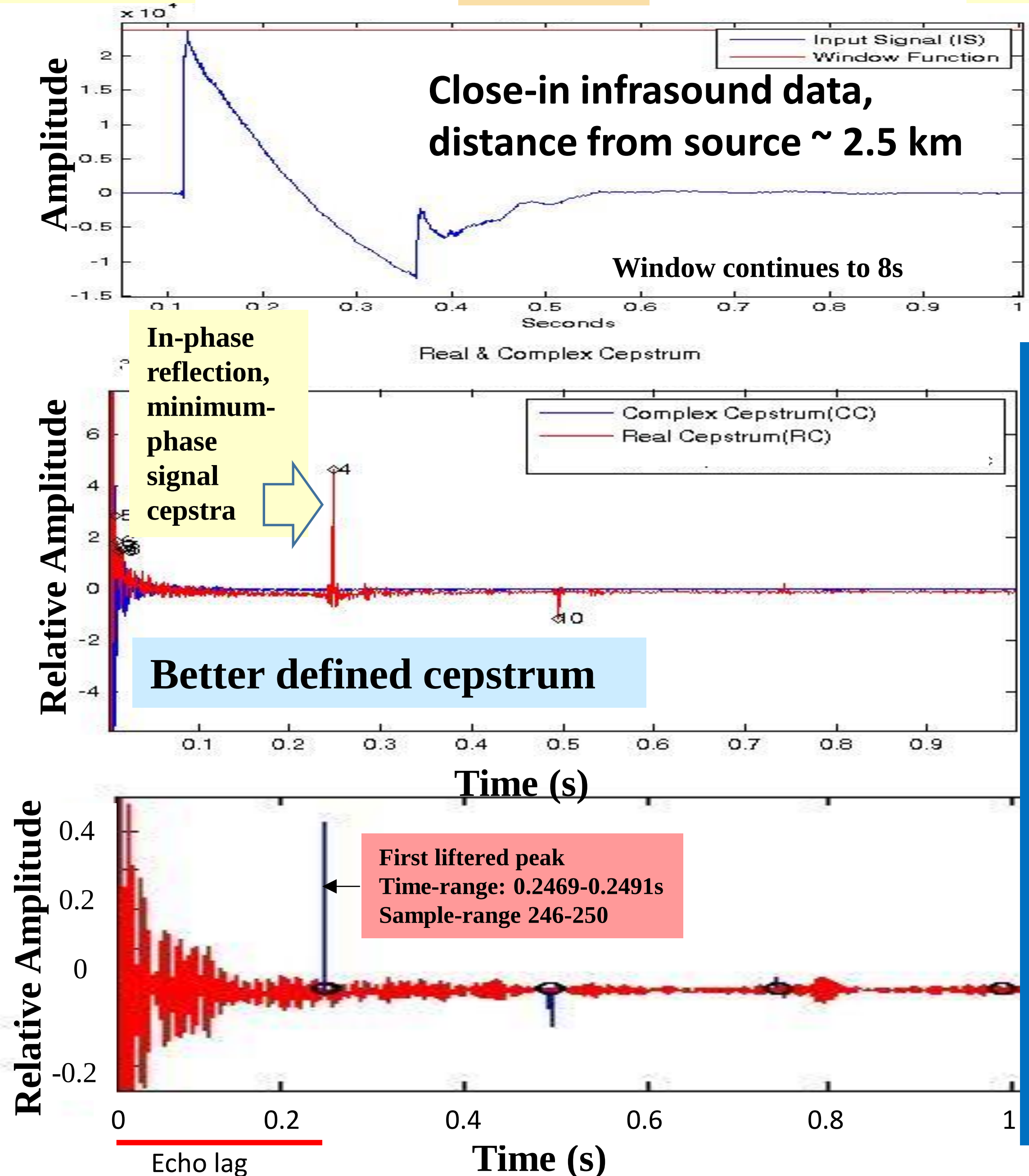
Power Spectral Density
NS2A GDF 1



Frequency (Hz)

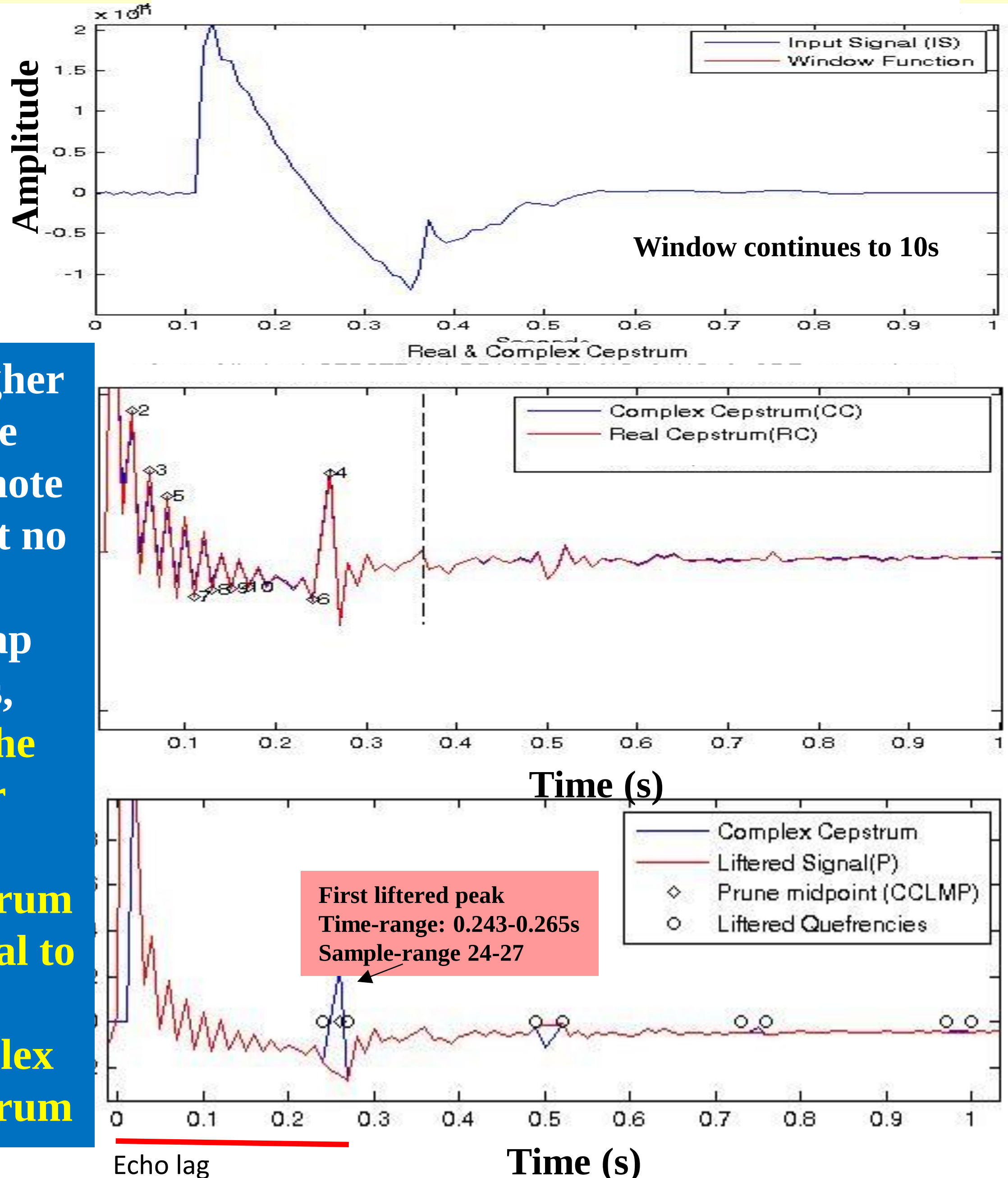
Close-in infrasound data 10-ton surface explosion test of the CTBTO-IMS, 2011

1000 sps



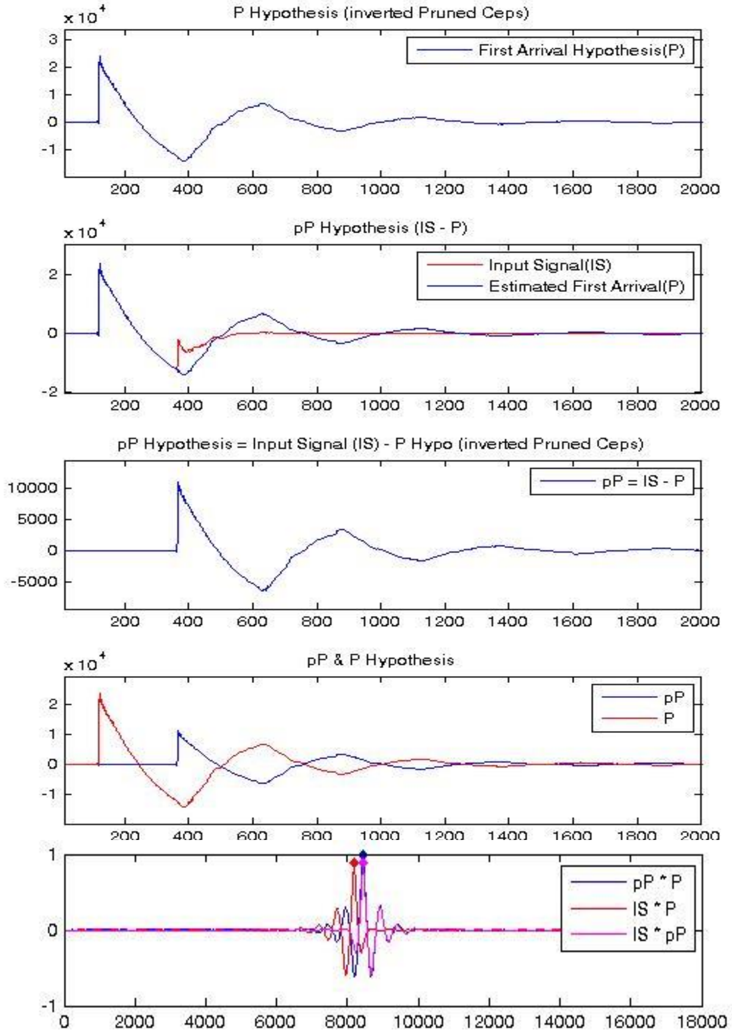
At higher sample rate, note almost no phase unwrap errors, thus **the Power (Real) Cepstrum is equal to the Complex Cepstrum**

100 sps



Close-in infrasound data 10-ton surface explosion test of the CTBTO-IMS, 2011

1000 sps



Relative Amplitude

Time (samples at 1000sps)

100 sps

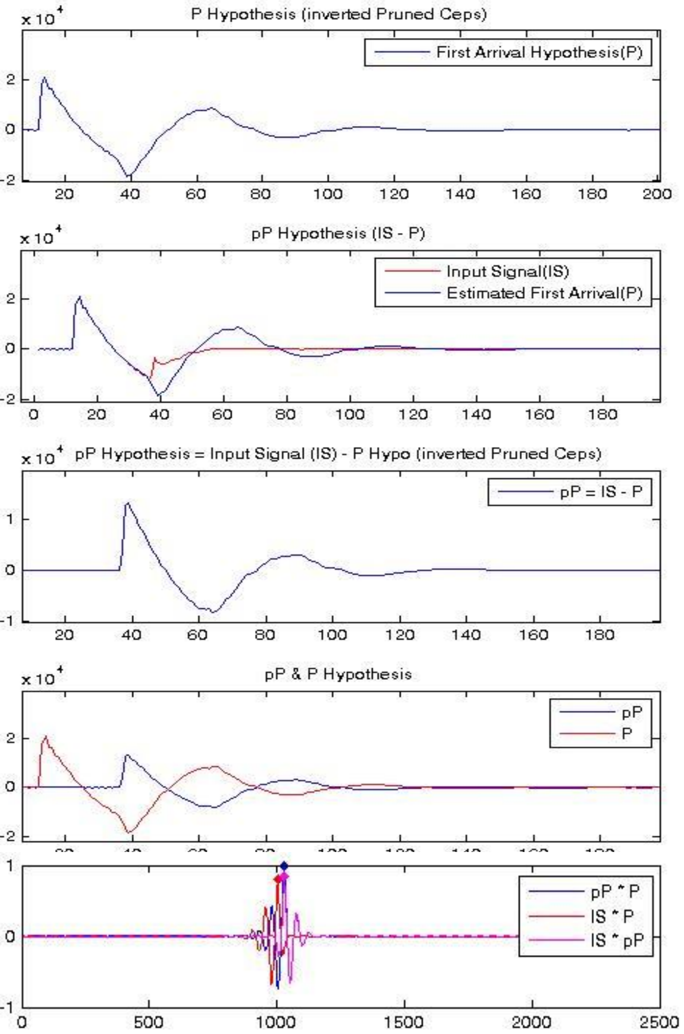
P : First arrival hypothesis based on Complex Cepstrum Lifter;
IS : Original signal;
pP : (IS - P) First echo hypothesis.

Note deconvolved waveform similarity at all sample rates, however, enhanced detail at higher sample rate;

Deconvolved waveform similarity evaluation metrics are slightly better at higher sample rate.

Correlation IS*P: 0.89
Correlation pP*P : 0.9
Correlation pP*IS: 0.8
power(pP)/ power(xcor_P)= 0.46

Correlation IS*P: 0.79
Correlation pP*P: 0.98
Correlation pP*IS: 0.84
power(pP)/ power(xcor_P)= 0.61



Time (samples at 100sps)

The input and estimated echo lag corresponds within two sample points for both sample rates.

References

- Bogert, B. P., Healy, M. J., & Tukey, J. W. (1962). *The Quefrency Alanysis of Time Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum and Saphe Cracking.* New York City: Wiley & Sons.
- Childers, D. G., Skinner, D. P., & Kemerait, R. C. (1976). The Cepstrum: A Guide to Processing. *Proceedings of the IEEE, 65*(10), 1428 - 1443.
- Kemerait, R. C., & Sutton, A. F. (1982). A Multidimensional Approach to Seismic Event Depth Estimation. *Geoexploration, 20*(1), 113-130.
- Otis, R. M., & Smith, R. B. (1977). Homomorphic Deconvolution By Log Spectral Averaging. *Geophysics, 42*(6), 1146-1157.
- Taylor, S. R. (1983, March 10). Three-Dimensional Crust and Upper Mantle Structure at the Nevada Test Site. *Journal of Geophysical Research, 88*(B3), 2220-2232.
- Tribolet, J. M. (1977). A New Phase Unwrapping Algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-25*(2), 170-177.
- Yacoub, N. K., Scott, J. H., & Mckeown, F. A. (1970, August). Computer Ray Tracing Through Complex Geological Models For Ground Motion Studies. *Geophysics, 35*(4), 585-602.