Analysis and classification of microseismic events

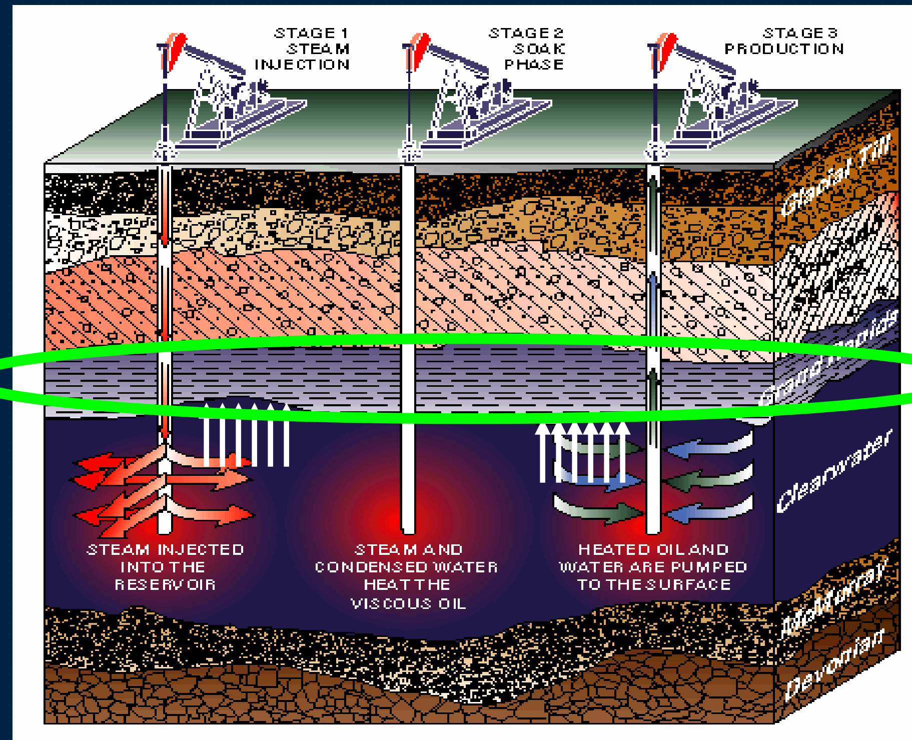
Jeff Tan
Henry Bland
Robert Stewart

19th Annual CREWES Sponsors Meeting
November 29, 2007

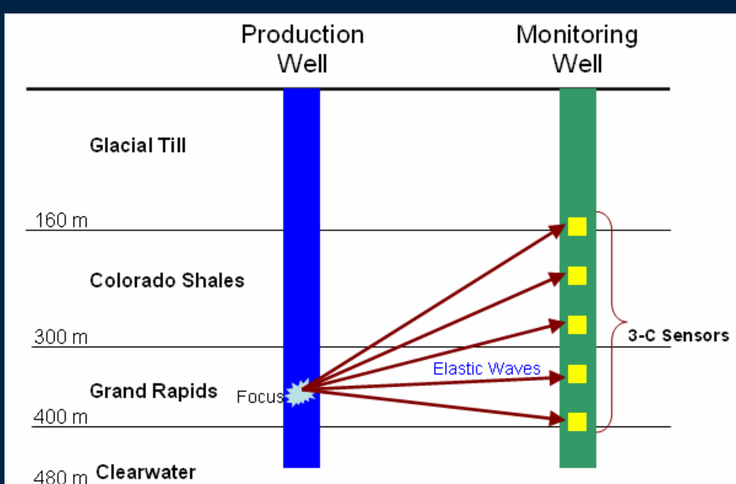
Cold Lake Background

(Imperial Oil Ltd., 2006a)

-Producing formation > 400m deep
-CSS used (Imperial Oil Ltd., 2006c):



~320ºC, 11MPa



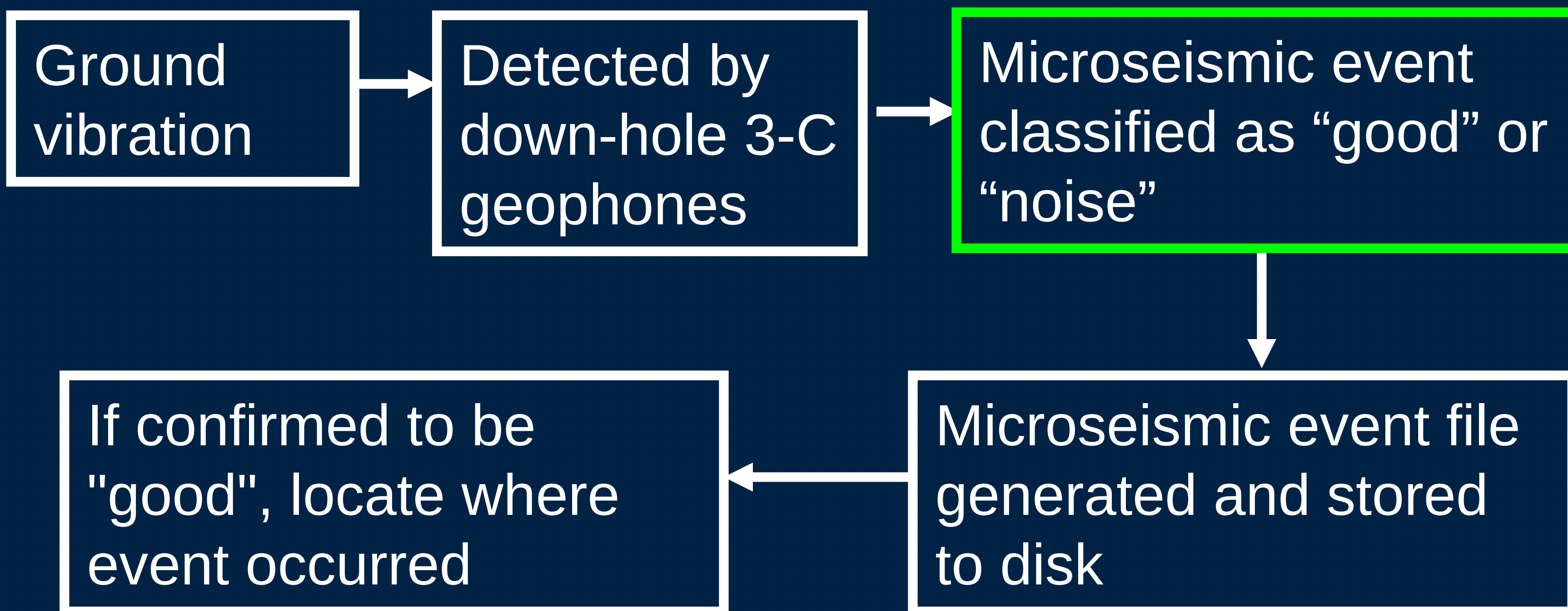
Stress in overburden (Grand Rapids Formation):

Cement cracks, casing failures possible

Passive seismic monitoring required

Cold Lake Background: Passive-Seismic Monitoring

Passive seismic system operation:



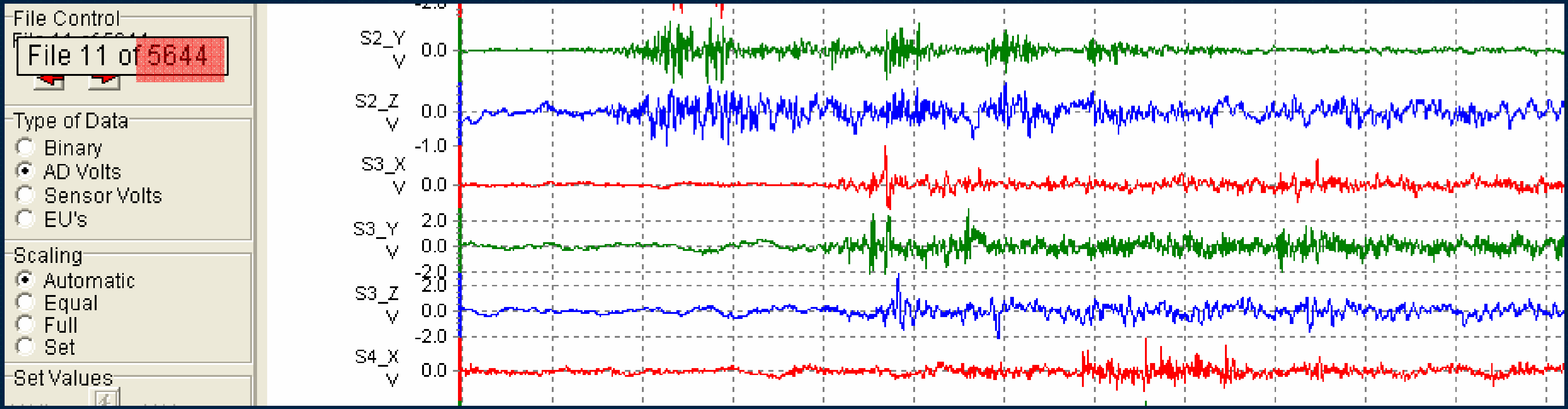
Theoretically investigate all “good” files, discard the rest.

Noise events ~ 99% of all microseismic events detected

Purpose:

Problem: Event-file classification software misclassifies files.

Importance: Manual analysis of thousands of misclassified files time-consuming & inefficient.



Solution: Develop novel and robust algorithms capable of accurately differentiating between “good” and “noise” files. Implement algorithms into user application.

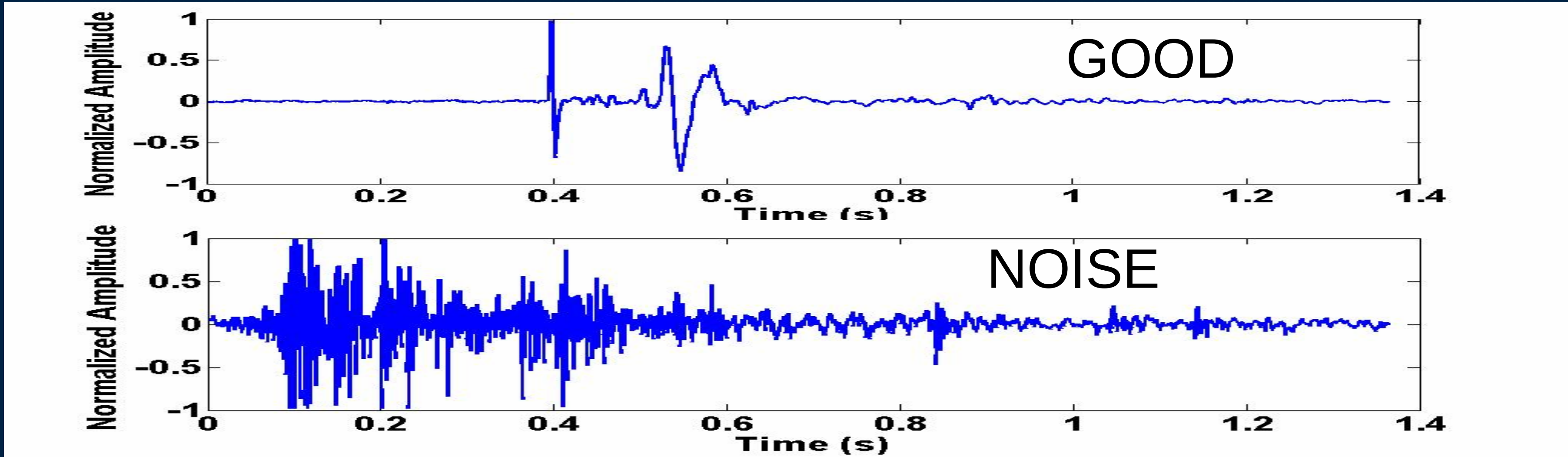
Algorithms Explored:

Classification Techniques:

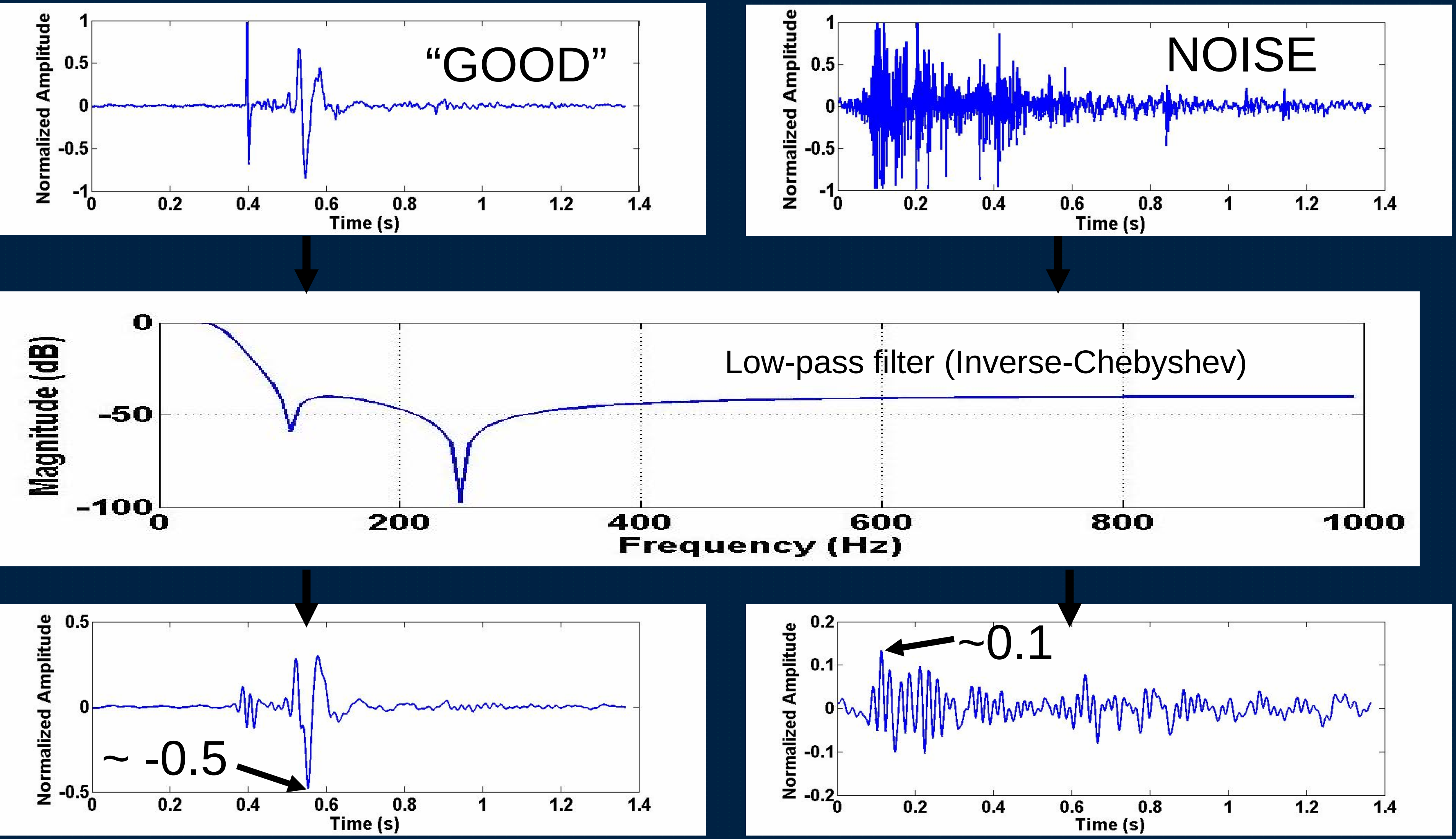
1) Frequency filtering: “Good” signals often contain lower dominant frequencies than noise.

2) Event-length detection: P-wave event-lengths of “good” signals are generally shorter than noise event-lengths.

3) Statistical analysis: "Good" events often have *lower signal variance*, *higher central data distribution* and *less sporadic sequential time-series behaviour* compared to noise.



Frequency Filtering: Low-pass example



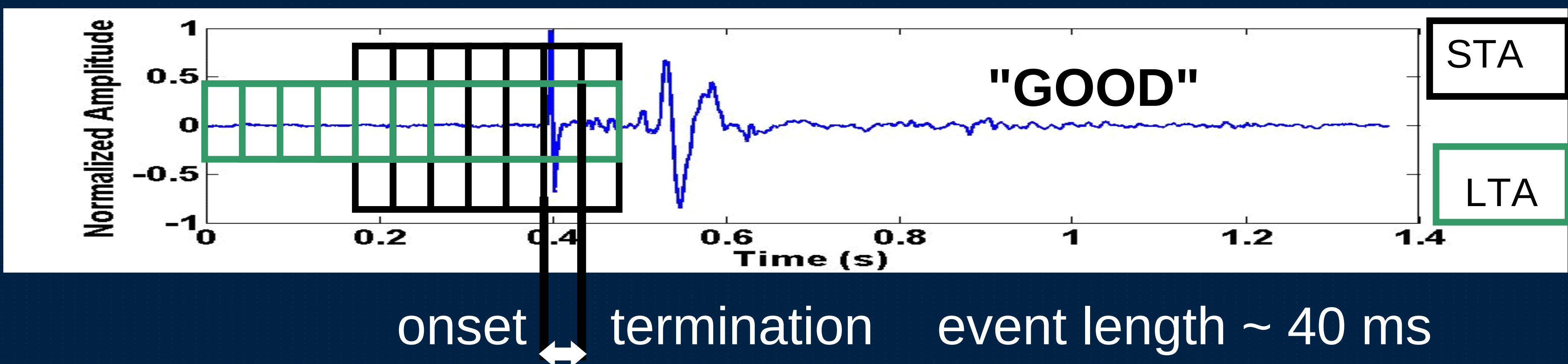
High-pass filtering also used (results in opposite trend shown above)

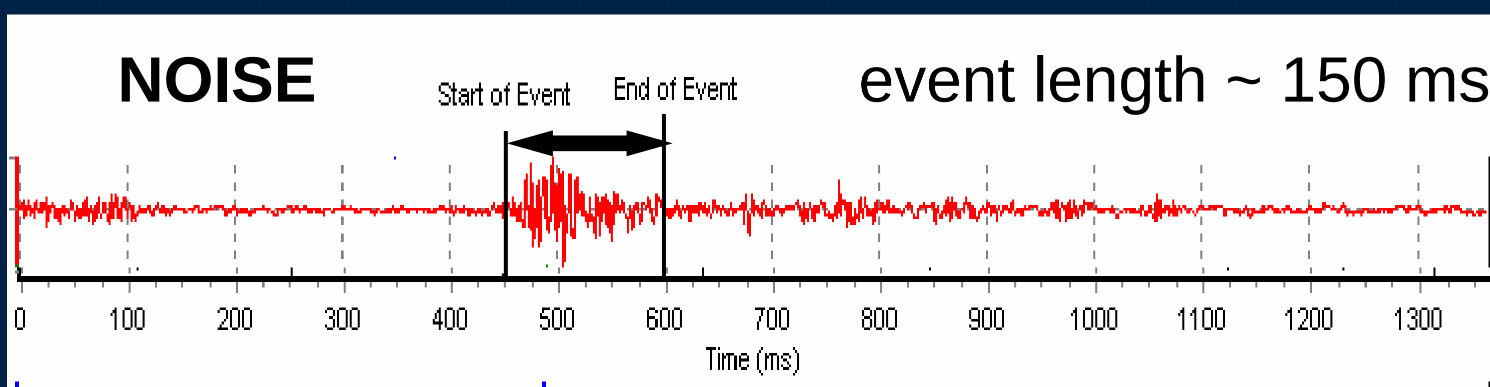
Algorithms Explored:

Event-length detection using a time-domain technique

STA/LTA (Ambuter and Solomon, 1974)

- STA / LTA ratio sharply increases at onset of event
- STA / LTA ratio sharply decreases at termination



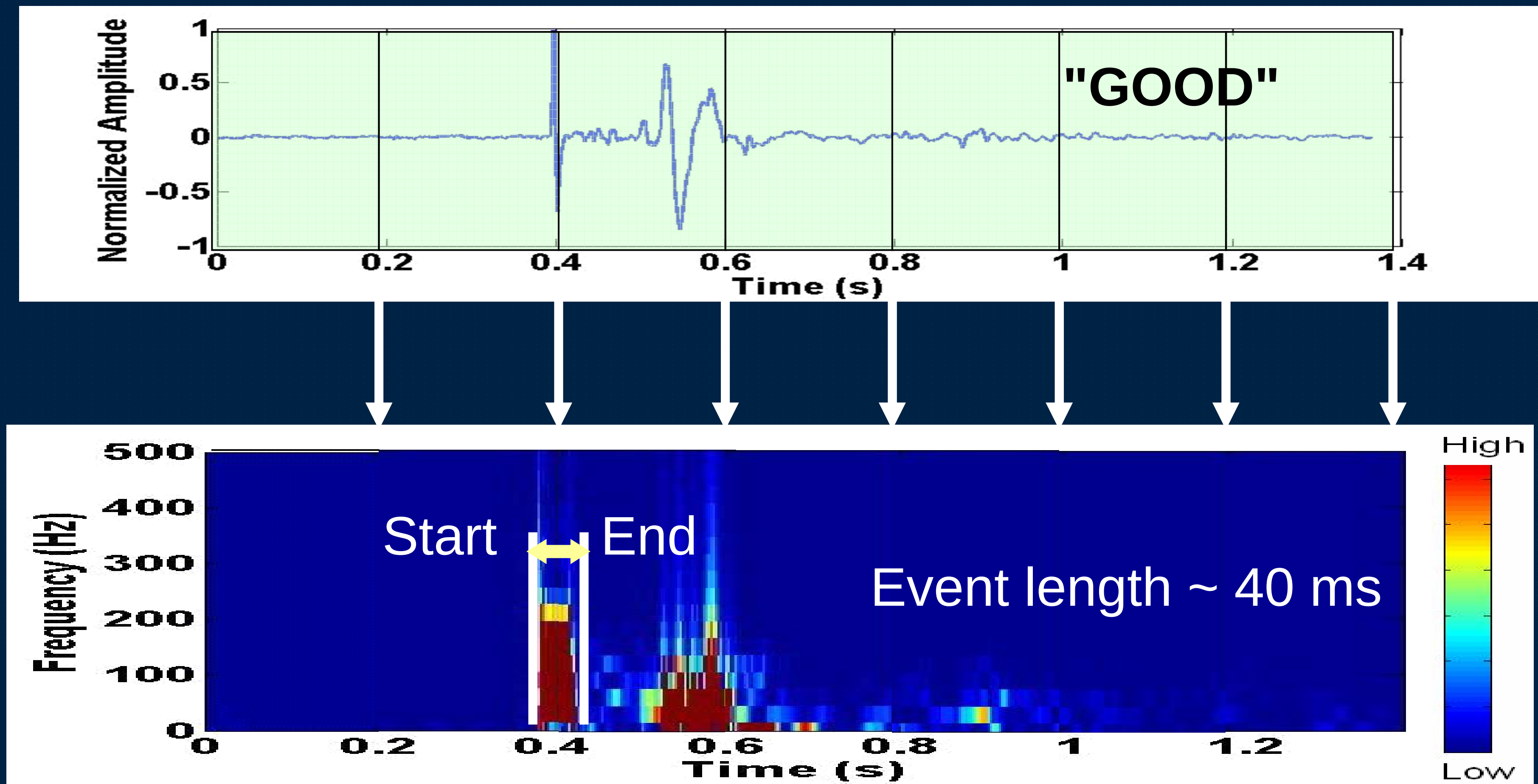


Algorithms Explored:

Event-length detection using a frequency-domain technique

Perform time-localized frequency transforms

- Examine high-frequency content to detect start / end points of event
  - High freq. content sharply increases at onset of event
  - High freq. content sharply decreases at termination



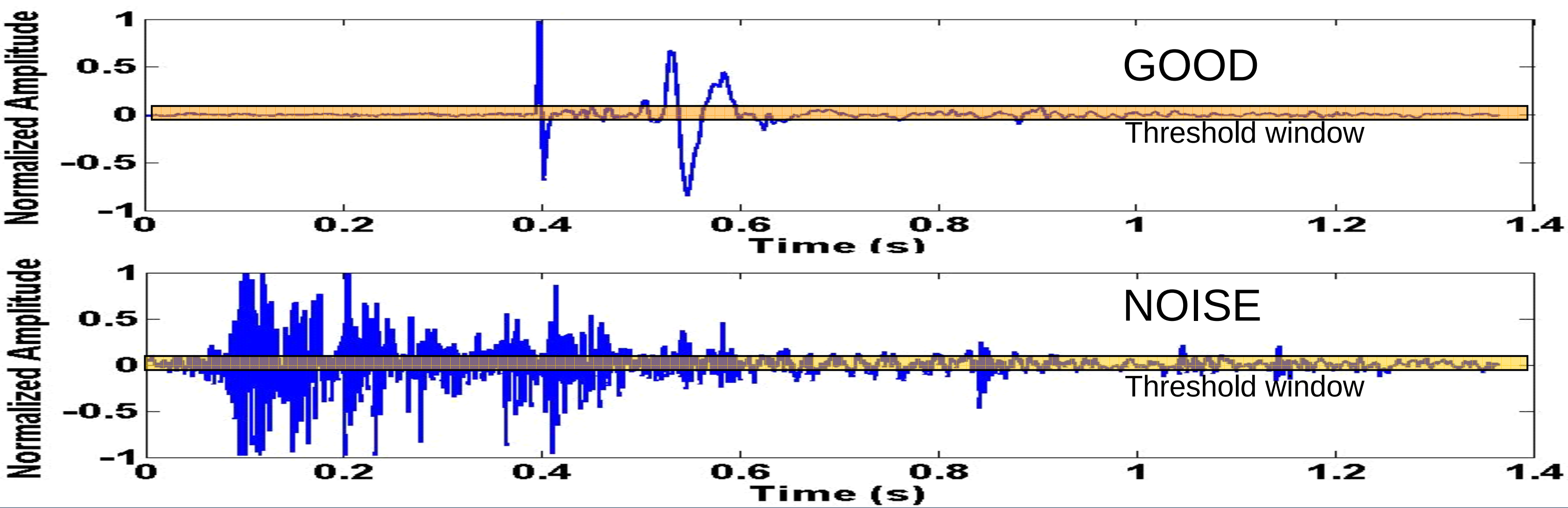
Algorithms Explored:

Chebyshev's Inequality
(e.g. Mitzenmacher and Upfal, 2005)

$$\Pr(|X - \mathbf{E}[X]| \geq a) \leq \frac{\mathbf{VAR}[X]}{a^2}$$

Statistical "Threshold Window" based on signal variance

Example: Set a threshold window between -0.03 and 0.03 (*a* = 0.03) and count all data points in time series that lie *outside* this window.



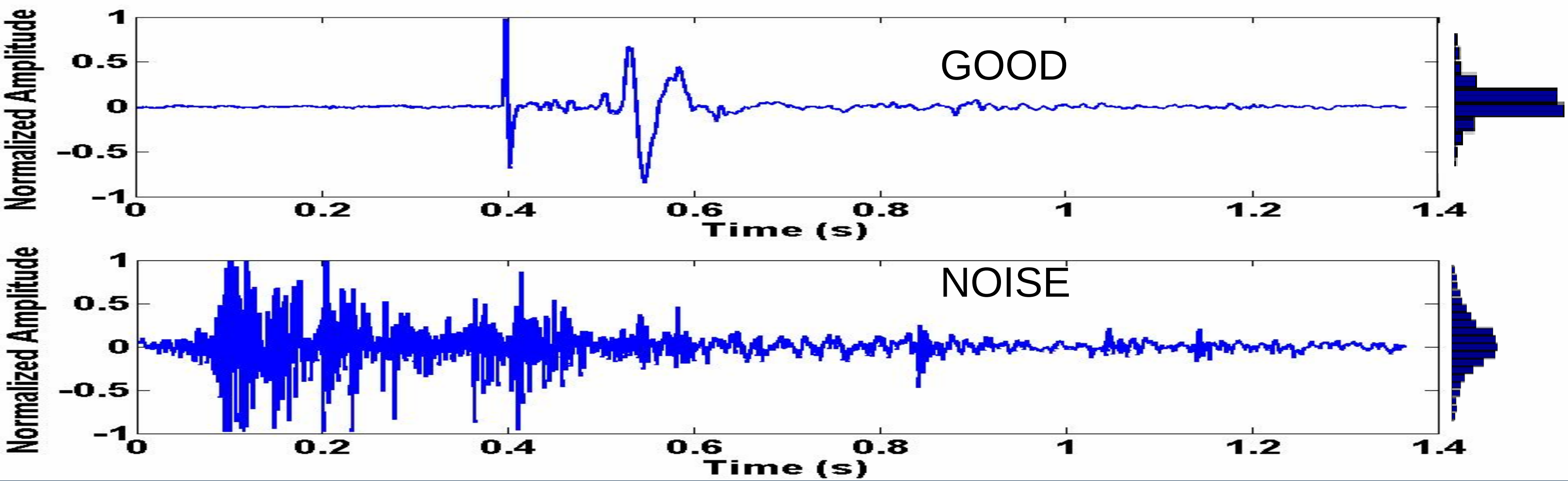
Signal shown	#Pts. Outside	Tot. Pts	% Pts. Outside
Good	850	4096	20.8%
Noise	2795	4096	68.2 %

Algorithms Explored:

Statistical Histogram to determine central data distribution

- "Good" signals generally have higher central data distribution.
- Histogram will be used to determine number of time series data points that fall within disjointed amplitude ranges.
- Look at concentration of points close to time axis.

Example: 99 evenly-spaced bins from -1 to 1, examine # data pts. in 50th bin range.

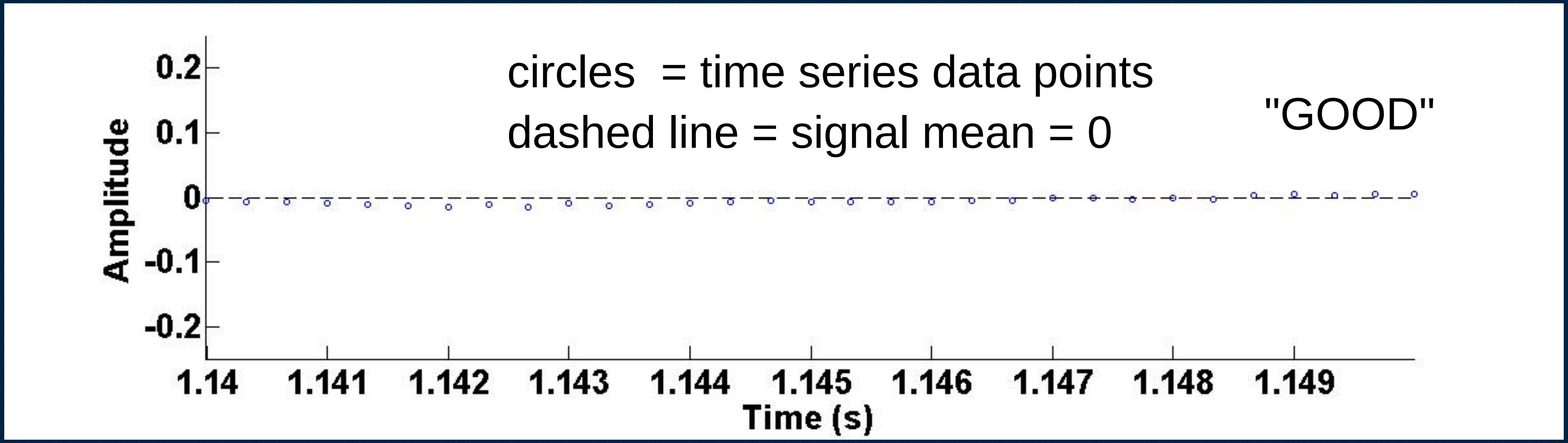


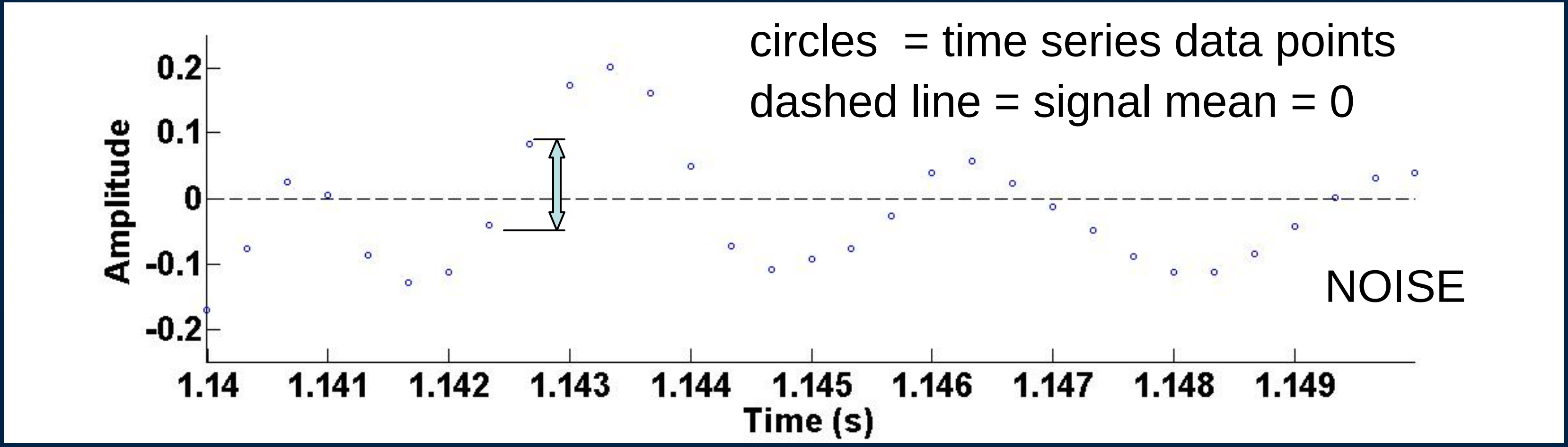
Signal shown	# pts in Bin 50	Total Pts.	% pts in Bin 50
Good	1416	4096	34.6%
Noise	438	4096	10.7%

Algorithms Explored:

Statistical "Specialized Zero-Crossing Count" algorithm

- Generally, "good" signals have less sporadic sequential time series behaviour about its mean.
- Take a look at zoom to very fine time interval to see this.

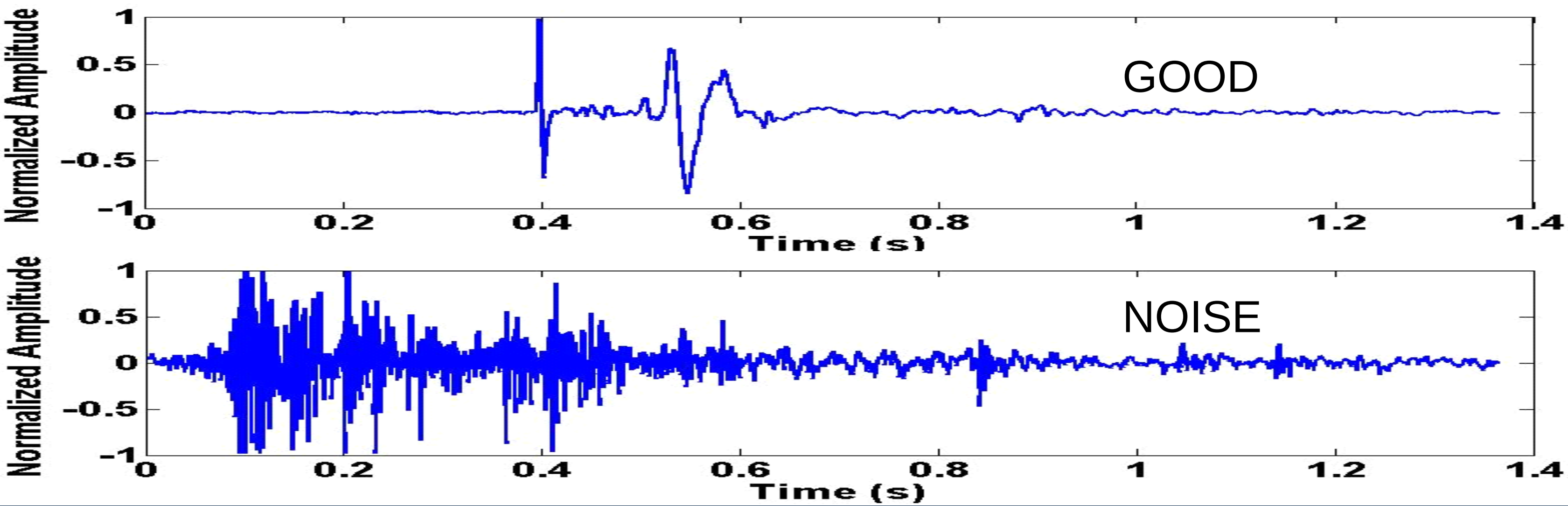




Algorithms Explored:

Statistical "Specialized Zero-Crossing Count" algorithm

Example: Count # times signal goes from strictly +ve to strictly -ve value (or other way) in **adjacent** data samples **after** low amplitude noise (data in range |y| <0.01, for example) is set to y = 0.



Signal shown	# Counted	Tot. Pts	%
Good	1	4096	0.0244%
Noise	298	4096	7.275 %

Algorithms Explored:

Summary:

1) Frequency Filtering (*peak amplitude examined after filtering***)**
- a) Inverse-Chebyshev low-pass filter
- b) Butterworth high-pass filter
- c) Chebyshev band-pass filter

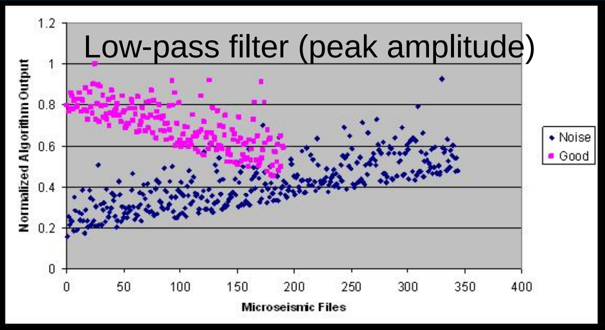
2) Event-Length Detection (*first arrival event-length calculated***)**
- a) Time-Domain (STA / LTA)
- b) Frequency-Domain (time-localized transform)

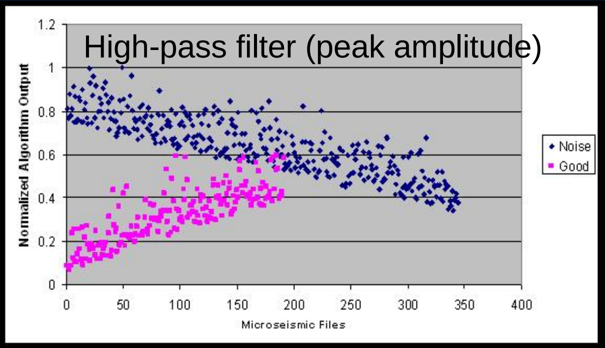
3) Statistical Analysis
- a) "Threshold" technique *(% outlying data points*)
- b) "Histogram" technique *(% pts in center histogram bin*)
- c) "Specialized Zero-Crossing Count" technique *(% adjacent polarity reversals after low-amplitude noise removed*)

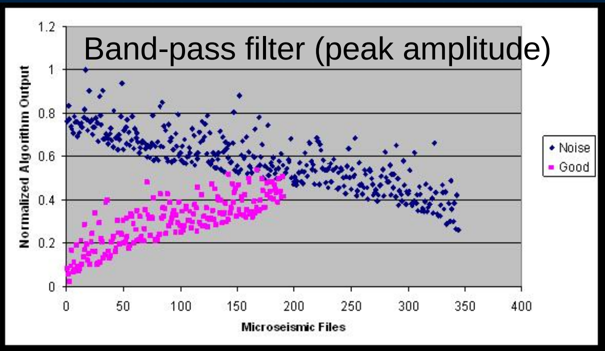
- Eight algorithm outputs (eight dimensional dataset).
- Every microseismic file can be seen as a point in an 8-D data space.
- Apply multivariate data reduction to reduce effective dimensionality of data.
- Use *principal components analysis* (PCA) to resolve data on new set of axes ("principal components") that are linear combinations of algorithm outputs.

Algorithm Outputs (e.g. 540-file test dataset):

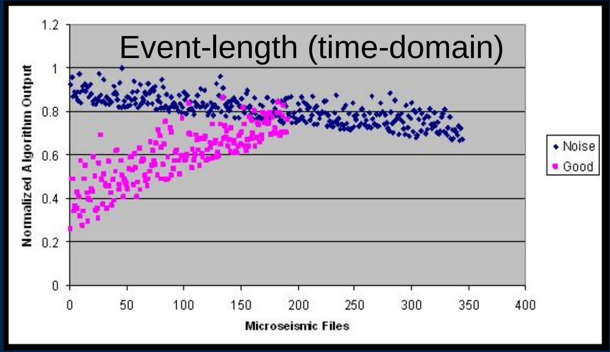
Frequency-Filtering

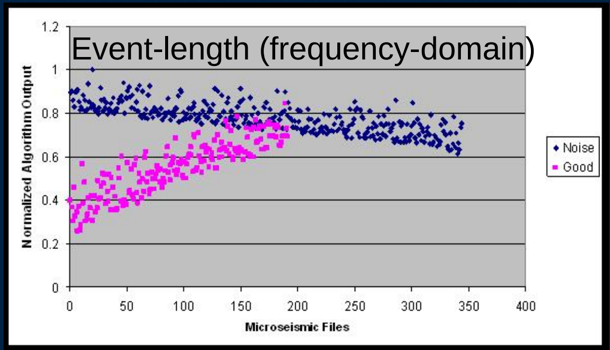






Event-Length Calculation



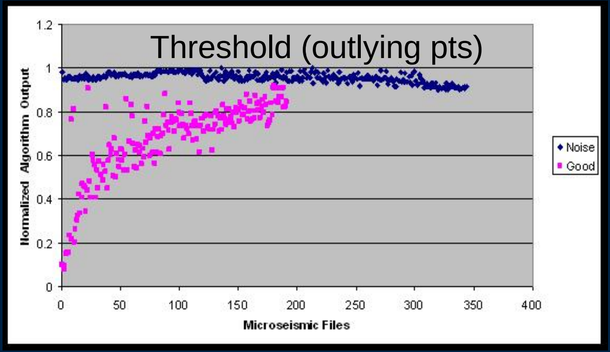


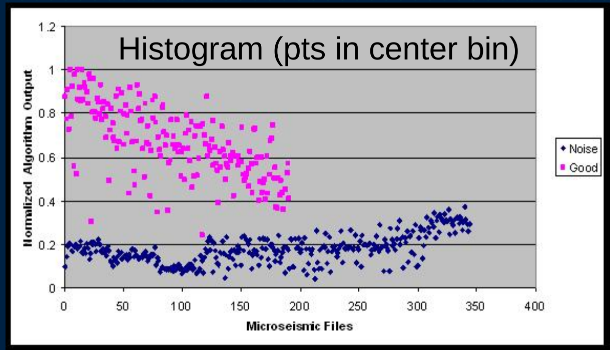
"Good" Files

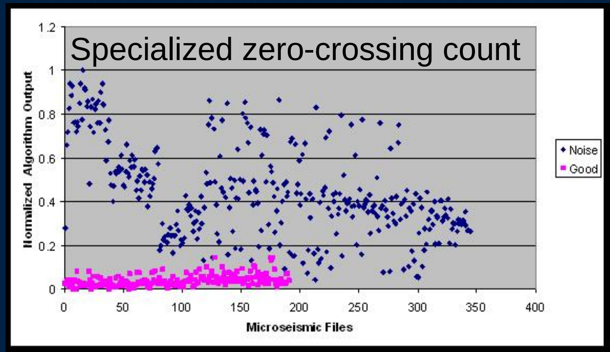
Noise Files

All outputs normalized for PCA application

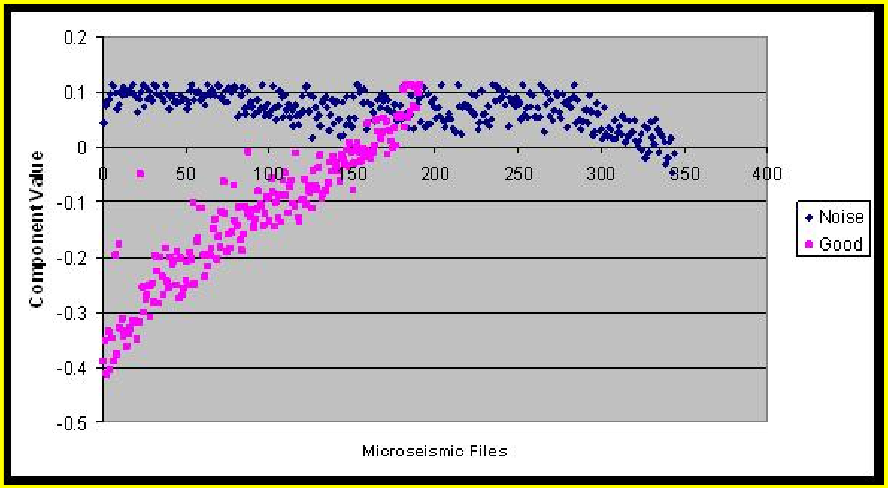
Statistical Analysis



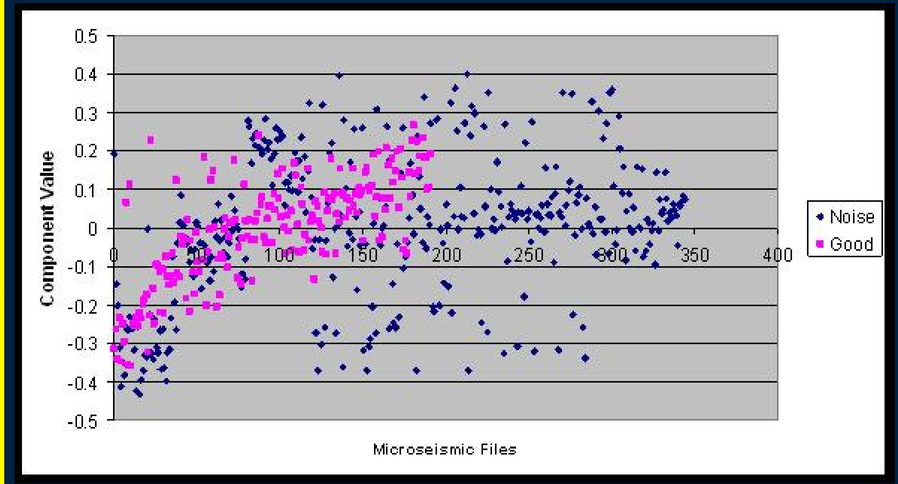




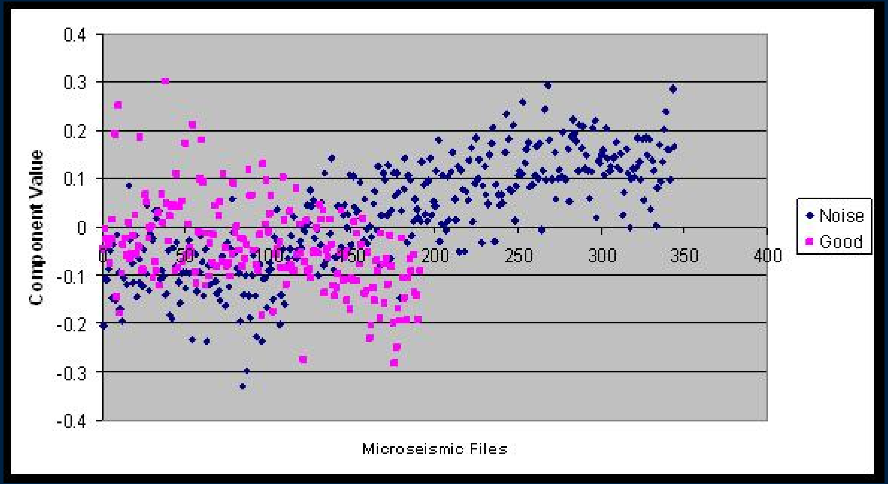
Projection onto Principal Components of 8-D Dataset:



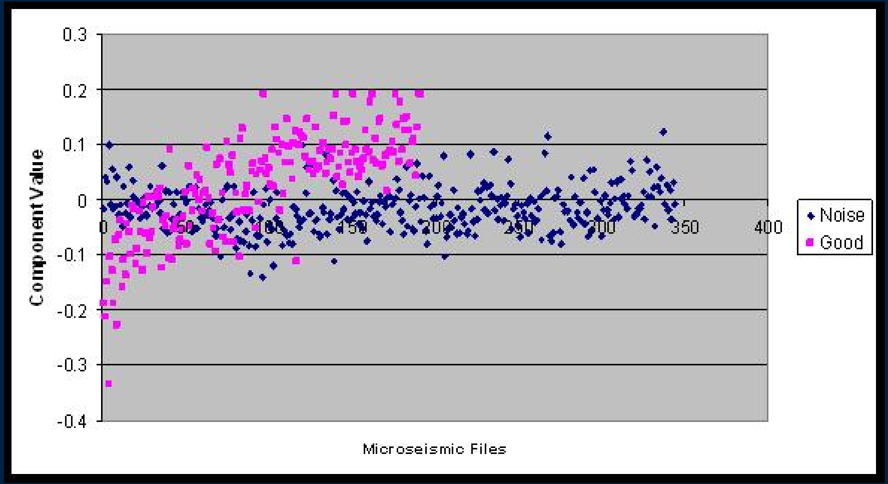
1st Component



2nd Component



3rd Component

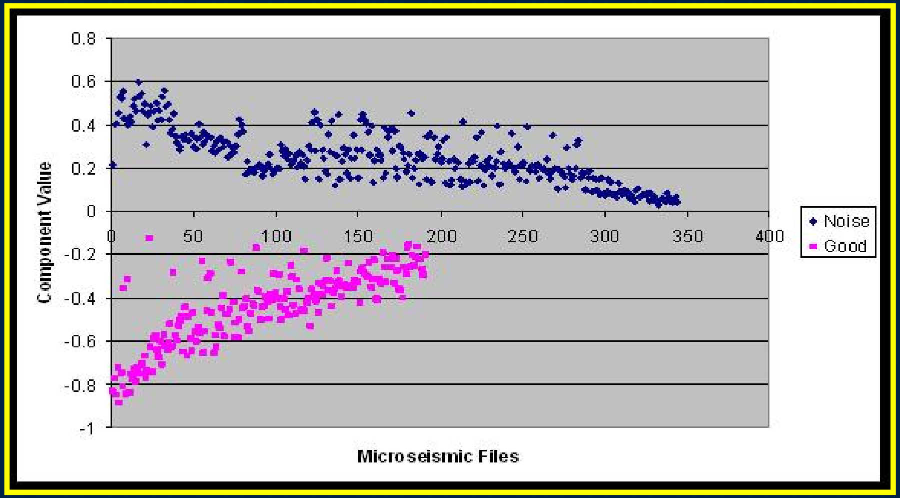


4th Component

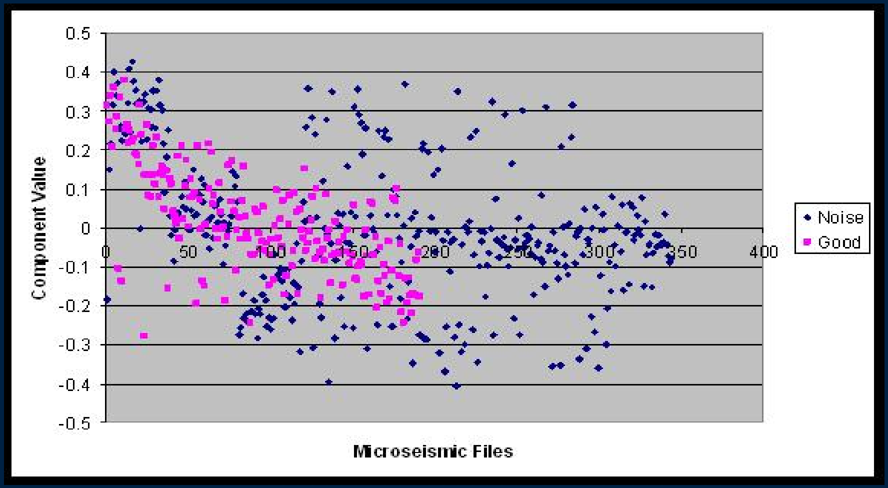
1st component shows improved clustering, but significant overlap still exists

Projection onto Principal Components of 3-D Dataset:

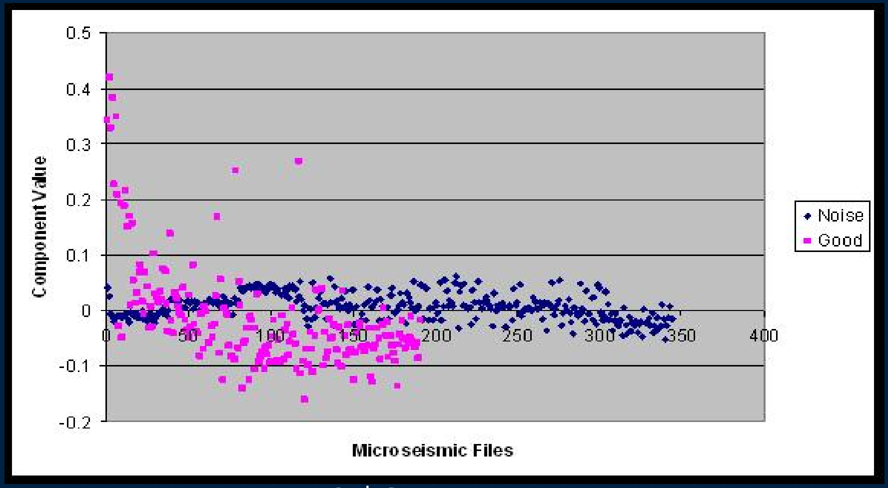
Restrict PCA to 3 statistical analysis algorithms only



1st Component



2nd Component

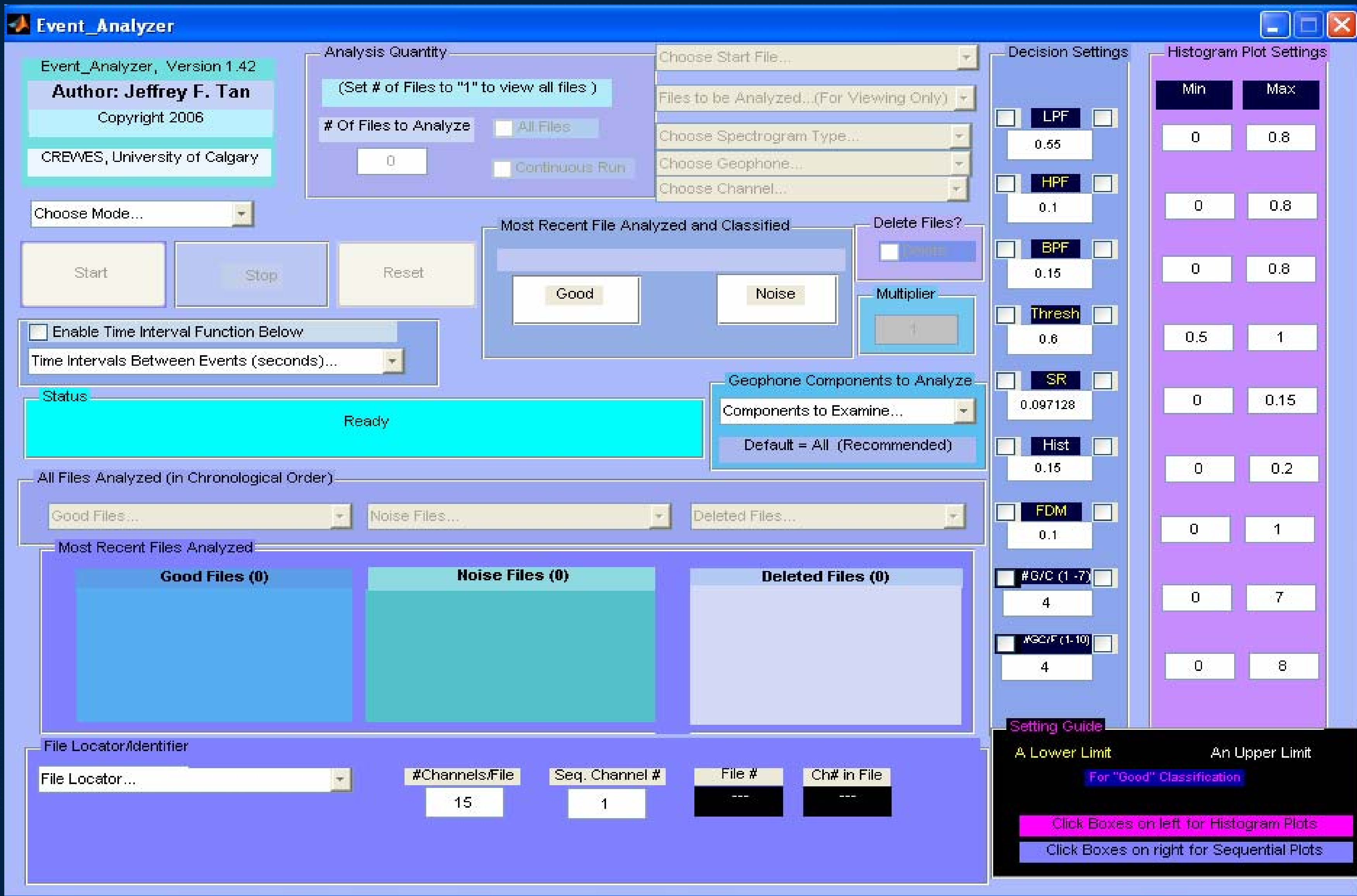


3rd Component

1st component shows clustering with no overlapping data from "good" and noise files (will not always be the case for different datasets, but is a significant improvement).

Implementations:

1) MATLAB Graphical User Interface (GUI) -- **applies most algorithms**.



Implementations:

2) MATLAB function that applies Principal Components Analysis to **statistical algorithm outputs**.

- Get principal components from statistical algorithm measurements on a reference dataset (the more diverse this dataset is, the better).
- Project measurements from an incoming microseismic file onto principal components.
- Analyze projected data for file classification.

Results:

Most consistent results with Implementation 2).

Three datasets tested (results from Implementation 2):

A) Specific dataset (most files from less than 5 pads)

- 99.5% accuracy

B) More diverse dataset (files from 28 pads)

- 98.8% accuracy

C) Most diverse, exhaustive dataset (files from 72 pads)

- 90.0% accuracy

Conclusions:

- Passive-seismic event-classification algorithms developed.
- Principal components analysis performed to reduce dataset dimensionality.
- Potentially significant future impact on Cold Lake operations given magnitude of daily microseismic dataset (sometimes up to 10,000+ events).

Acknowledgements:

- CREWES sponsors
- Robert Stewart, my supervisor
- Henry Bland now with Pinnacle Technologies
- Colum Keith, Richard Smith, and Sophia Follick from Imperial Oil Ltd.
- CREWES staff and students

Presentation References

Ambuter, B.P., and Solomon, S.C., 1974, An event-recording system for monitoring small earthquakes: Bulletin of the Seismological Society of America, **64**, 1181-1188.

Dunteman, G.H., 1989, Principal Components Analysis: Sage Publications.

Feichtinger, H.G., and Strohmer, T., 1998, Gabor Analysis and Algorithms: theory and applications: Birkhäuser.

Imperial Oil Ltd., 2006a, Alberta's Oil Sands: http://www.limperiale.ca/Canada-English/Investors/Operating/Natural_Resources/I_O_NaturalResourcesFig1.asp, internet web page. Accessed November 3, 2006.

Imperial Oil Ltd., 2006b, Proposed Development Areas – Mahihkan North and Nabiye: http://www.limperiale.ca/Canada-English/Investors/Operating/Natural_Resources/I_O_NaturalResourcesFig4.asp, internet web page. Accessed November 5, 2006.

Imperial Oil Ltd, 2006c, Cyclic Steam Stimulation - not to scale: http://www.limperiale.ca/CanadaEnglish/Investors/Operating/Natural_Resources/I_O_NaturalResourcesFig5.asp, internet web page. Accessed November 5, 2006.

Isaac, H.J., 1996, Seismic Methods for Heavy Oil Reservoir Monitoring: PhD. thesis, University of Calgary.

Jackson, J.E., 1991, A User's Guide to Principal Components: John Wiley & Sons.

Jolliffe, I.T., 2002, Principal Component Analysis, 2nd edition: Springer.

Lupton, R., 1993, Statistics In Theory and Practice: Princeton University Press.

Maundy, B., 2005, ENEL 559 Course Notes, University of Calgary.

Mitzenmacher, M., and Upfal, E., 2005, Probability and Computing: Randomized Algorithms and Probabilistic Analysis: Cambridge University Press.

Shlens, J., 2003, A Tutorial On Principal Component Analysis: Derivation, Discussion, and Singular Value Decomposition, University of California, San Diego.

Smith, L.I., 2002, A tutorial on Principal Components Analysis, University of Otago, New Zealand.

Stockwell, R.G., Mansinha, L., and Lowe, R.P., 1996, Localization of the Complex Spectrum: The *S* Transform: IEEE Transactions On Signal Processing, **44**, 998-1001.