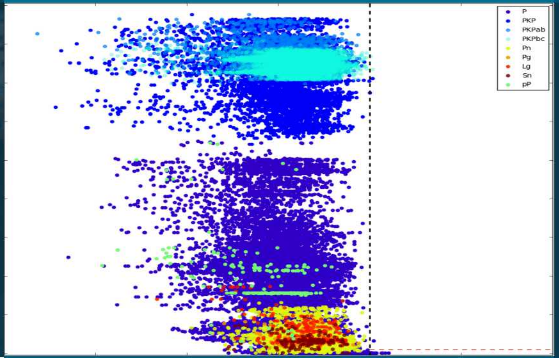
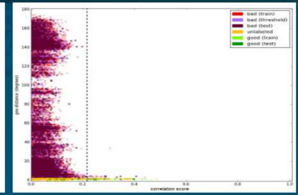
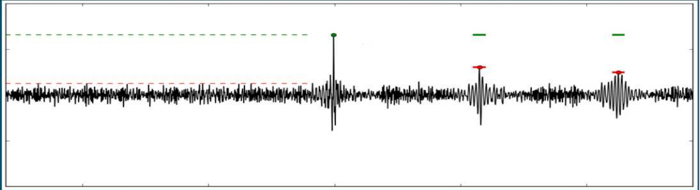
SAND2019-1605PE

Alternate Null Hypothesis Correlation: A New Approach to Seismic Event Detection







PRESENTED BY

Amy Sundermier

COAUTHORS

Tyler Ganter, Sanford Ballard

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

The views expressed here do not necessarily reflect the views of the United States Government, the United States Department of Energy, the National Nuclear Security Administration, or Sandia National Laboratories.

Agenda

Introduction

Waveform correlation uses waveform templates to detect and identify repeating seismic events.

High false alarm rates for global and regional monitoring of sparse networks, where false alarms are caused by geographically non-collocated sources, not noise.

Alternate Null Hypothesis Correlation (ANCorr) applies machine learning algorithms to generate alternative templates that increase the detection/false alarm ratio.

Mathematical Basis of Waveform Correlation Detection Model

Conventional template matching is based on matched filter (MF)

- Optimal linear filter for maximizing the signal-to-noise ratio (SNR)
- In the presence of additive white Gaussian noise (AWGN).

The matched filter detection model consists of binary hypotheses:

$$\mathcal{H}_0 : \mathbf{x} = \sigma \mathbf{n}$$

Null hypothesis $\mathcal{H}_0$:
the data contains only noise†.

$$\mathcal{H}_1 : \mathbf{x} = A\mathbf{w}_{\mathrm{MF}} + \sigma \mathbf{n}$$

Signal present hypothesis $\mathcal{H}_1$:
the data contains an amplitude-scaled copy of the template waveform and noise.

†Noise vector is drawn from a multivariate normal distribution $\mathbf{n}$: $\mathcal{N}(\mathbf{0}, \mathbf{I})$ and scaled by constant σ

Mathematical Basis of Waveform Correlation Detection Model

Conventional template matching is based on matched filter (MF)

- Optimal linear filter for maximizing the signal-to-noise ratio (SNR)
- In the presence of additive white Gaussian noise (AWGN).

The matched filter detection model consists of binary hypotheses:

$$\mathcal{H}_0 \colon \mathbf{x} = \sigma \mathbf{n}$$

Problem: Matched filter detection model does not account for the presence of realistic alternative signals, such as other seismic waveforms.

$$\mathcal{H}_1 \colon \mathbf{x} = A\mathbf{w}_{\mathrm{MF}} + \sigma \mathbf{n}$$

†Noise vector is drawn from a multivariate normal distribution $\mathbf{n}\colon \mathcal{N}(\mathbf{0}, \mathbf{I})$ and scaled by constant σ

Alternate Null-Hypothesis Correlation (ANCorr)

ANCorr detection model consists of binary hypotheses, but with an alternate null hypothesis, $\mathcal{H}_0^*$:

$$\mathcal{H}_0^*: \mathbf{x} = B\mathbf{v} + \sigma\mathbf{n}$$

$\mathcal{H}_0^*$ recognizes that in addition to noise, other seismic arrival waveforms may be present in the data.

$$\mathcal{H}_1: \mathbf{x} = A\mathbf{w}_{\mathrm{MF}} + \sigma\mathbf{n}$$

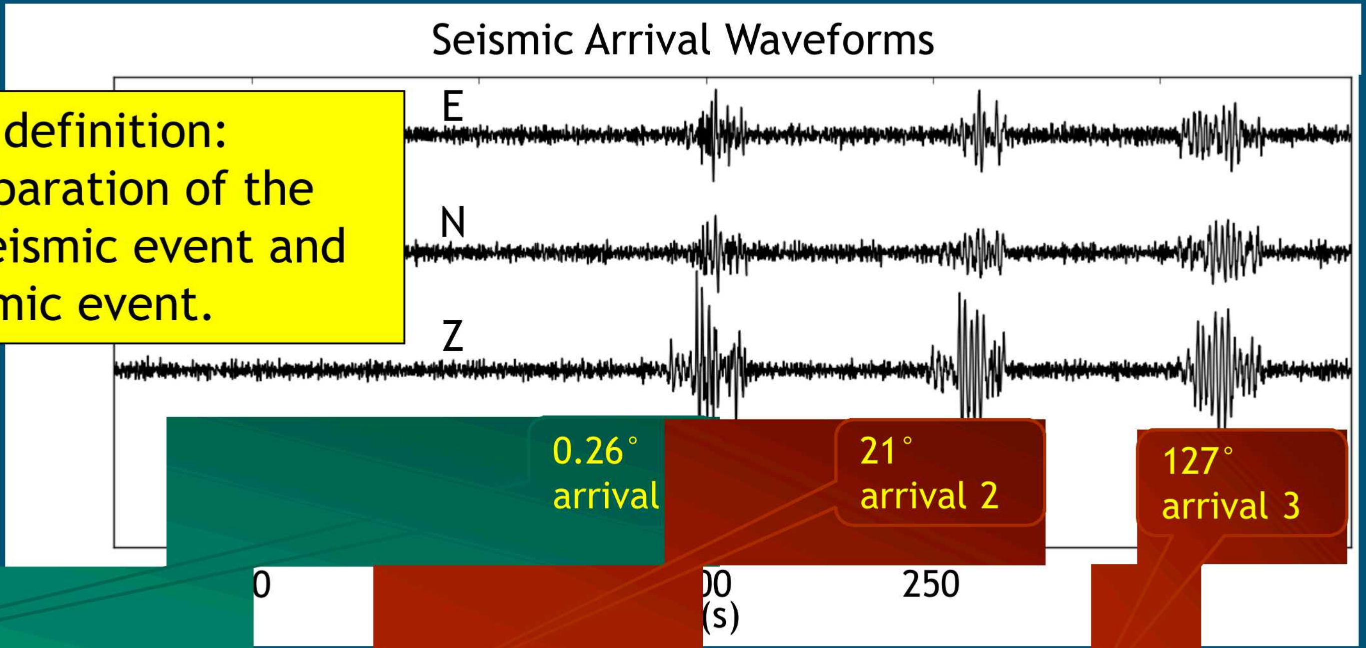
We develop an ANCorr template to maximize the SNR ratio between $\mathcal{H}_0^*$ and $\mathcal{H}_1$.

In other words, an ANCorr template is synthesized to detect collocated events but minimize false alarms from seismic arrivals with a different event location.

Defining Geodistance for 3 Seismic Arrivals at a Three-component Station

Template location
(-7.07, 129.41)

Geodistance definition:
Epicentral separation of the **template's** seismic event and **arrival's** seismic event.



0.26° arrival

21° arrival 2

127° arrival 3

Arrival 1 location
(-6.85, 129.56)

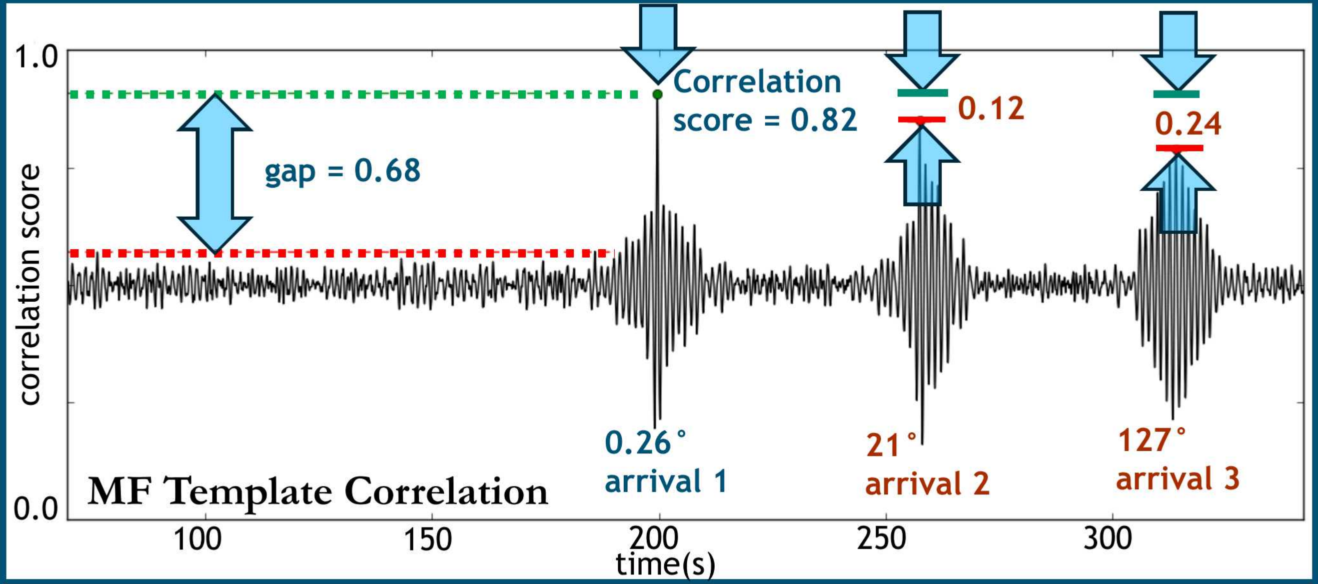
Detection desired

ocation
(11.08, 140.46)

Detection not desired

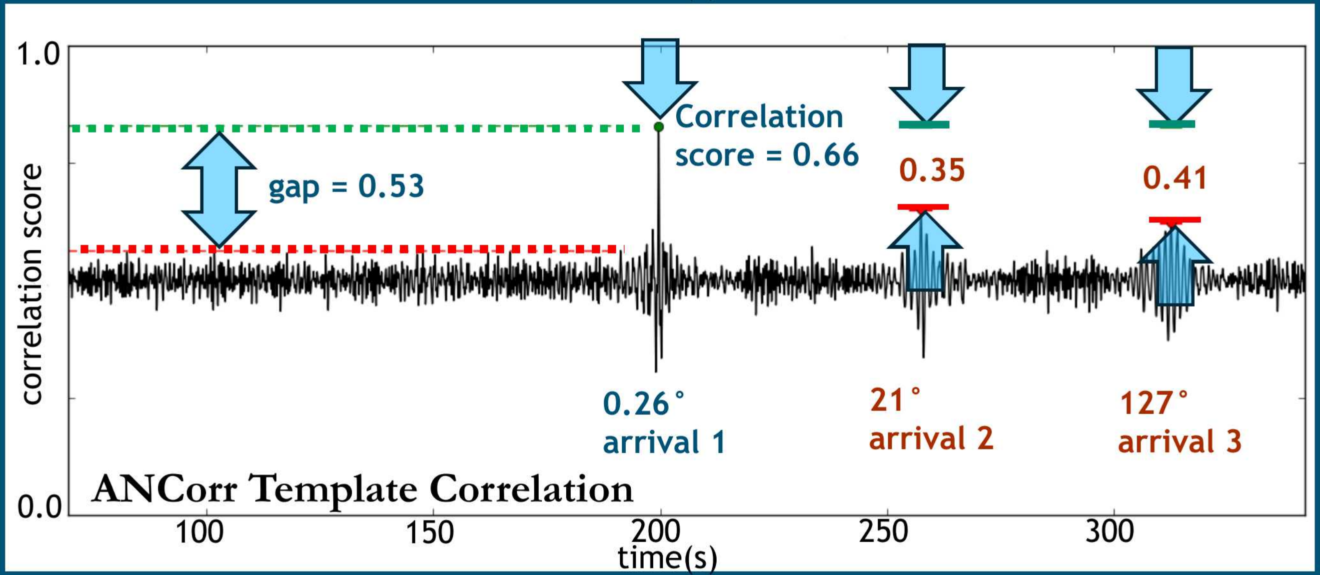
Arrival 3 on
(15.32, -104.57)

Correlation Functions: Matched Filter and ANCorr Templates



Arrival	MF Score
1	0.82
2	0.7
3	0.58
Noise	0.14

False Detections Likely



Arrival	ANCorr Score
1	0.66
2	0.31
3	0.25
Noise	0.13

Reduced Chance of False Detections

Research Challenges of Alternate Null-Hypothesis Correlation Model

How to make an ANCorr template to maximize the SNR ratio between $\mathcal{H}_0^*$ and $\mathcal{H}_1$?

$\mathcal{H}_0^*: \mathbf{x} = B\mathbf{v} + \sigma\mathbf{n}$ $\qquad$ $\mathcal{H}_1: \mathbf{x} = A\mathbf{w}_{\mathrm{MF}} + \sigma\mathbf{n}$

Research Challenges:

1. Unlike white Gaussian noise, seismic arrivals are nonstationary.

Stationary Noise: Swapping windows of noise will not lead to a different answer from the detection algorithm.

Research Challenges of Alternate Null-Hypothesis Correlation Model

How to make an ANCorr template to maximize the SNR ratio between $\mathcal{H}_0^*$ and $\mathcal{H}_1$?

$$\mathcal{H}_0^*: \mathbf{x} = B\mathbf{v} + \sigma\mathbf{n} \qquad \mathcal{H}_1: \mathbf{x} = A\mathbf{w}_{\mathrm{MF}} + \sigma\mathbf{n}$$

Research Challenges:

1. Unlike white Gaussian noise, seismic arrivals are nonstationary.

Swapping windows of nonstationary seismic signals will lead to a different answer from the detection algorithm.

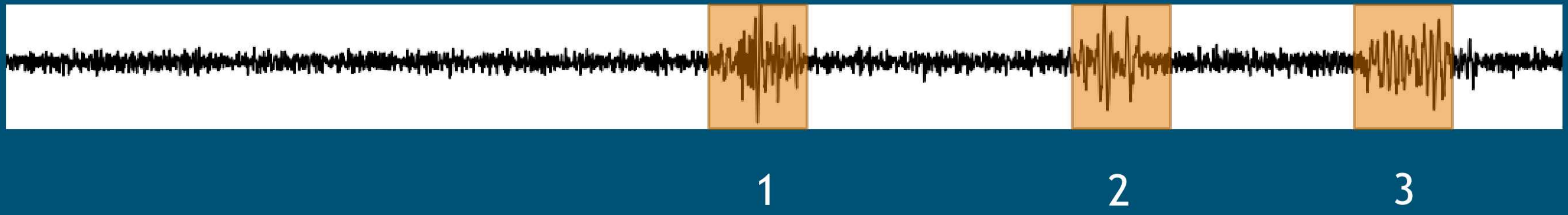
Research Challenges of Alternate Null-Hypothesis Correlation Model

How to make an ANCorr template to maximize the SNR ratio between $\mathcal{H}_0^*$ and $\mathcal{H}_1$?

$$\mathcal{H}_0^*: \mathbf{x} = B\mathbf{v} + \sigma\mathbf{n} \qquad \mathcal{H}_1: \mathbf{x} = A\mathbf{w}_{\mathrm{MF}} + \sigma\mathbf{n}$$

Research Challenges:

1. Unlike white Gaussian noise, seismic arrivals are nonstationary.
2. Nondeterministic: Each possible event and seismic phase that can generate $B\mathbf{v}$ must be considered.



Basis of ANCorr Method

How to make an ANCorr template to maximize the SNR ratio between $\mathcal{H}_0^*$ and $\mathcal{H}_1$?

$$\mathcal{H}_0^*: \mathbf{x} = B\mathbf{v} + \sigma\mathbf{n} \qquad \mathcal{H}_1: \mathbf{x} = A\mathbf{w}_{\mathrm{MF}} + \sigma\mathbf{n}$$

Confluence of three factors enable the ANCorr method:

1. the new model of $\mathcal{H}_0^*$ and $\mathcal{H}_1$;
2. the large sampling of labeled **v**'s (i.e., seismic arrivals) available in seismic bulletins; and
3. the similarity between correlation-based detection and linear discriminant functions (LDFs).

These three factors provide sufficient conditions for supervised machine learning.

We will exploit machine learning to create ANCorr templates.

Agenda

Comparison: Waveform Correlation and LDF equations

Waveform Correlation

The correlation function for data $\mathbf{x}$ and template $\mathbf{w}$ at time n has correlation score:

$$R(\mathbf{x}, n) = \mathbf{w}^T \mathbf{x}_n$$

Incorporate the template threshold to make new function:

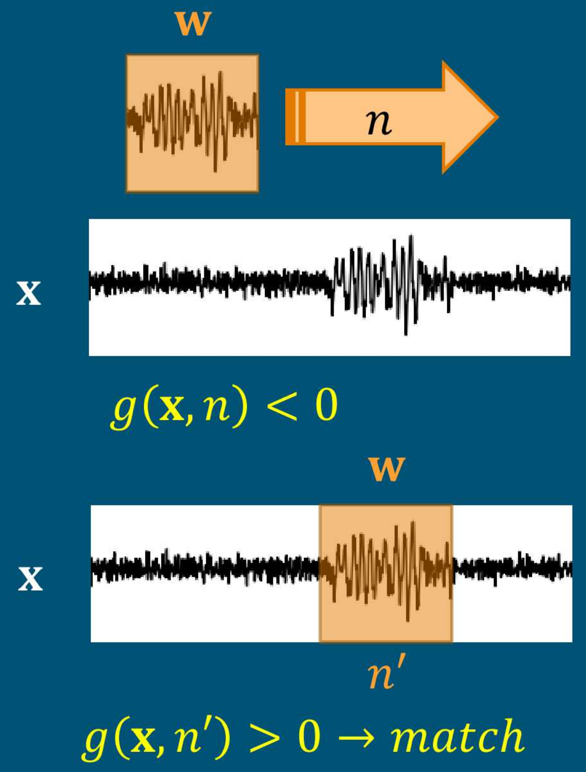
$$g(\mathbf{x}, n) = R(\mathbf{x}, n) - \omega_0 = \mathbf{w}^T \mathbf{x}_n - \omega_0$$

With match criterion at the time when maximum correlation occurs:

$$g(\mathbf{x}, n') > 0 \rightarrow match$$

$$g(\mathbf{x}, n') < 0 \rightarrow no\ match$$

We interpret correlation as a series of dot products by incrementing n and sliding the template $\mathbf{w}$ over the data stream $\mathbf{x}$.



Comparison: Waveform Correlation and LDF equations

Waveform Correlation

The correlation function for data $\mathbf{x}$ and template $\mathbf{w}$ at time n has correlation score:

$$R(\mathbf{x}, n) = \mathbf{w}^T \mathbf{x}_n$$

Incorporate the template threshold to make new function:

$$g(\mathbf{x}, n) = R(\mathbf{x}, n) - \omega_0 = \mathbf{w}^T \mathbf{x}_n - \omega_0$$

With match criterion at the time when maximum correlation occurs:

$$g(\mathbf{x}, n') > 0 \rightarrow match$$

$$g(\mathbf{x}, n') < 0 \rightarrow no\ match$$

Linear Discriminant Functions

A **discriminant function** is a linear combination of the components of the data vector $\mathbf{u}$ where $\mathbf{w}^T$ is the weight vector and ω_0 is the bias (Duda, *et al.*, 2001):

$$g(\mathbf{u}) = \mathbf{w}^T \mathbf{u} - \omega_0$$

In the two-category case, classification of $\mathbf{u}$ goes by the decision rule

$$g(\mathbf{u}) > 0 \rightarrow class\ 1$$

$$g(\mathbf{u}) < 0 \rightarrow class\ -1$$

Comparison: Waveform Correlation and LDF equations

Waveform Correlation

The correlation function for data $\mathbf{x}$ and template $\mathbf{w}$ at time n has correlation score:

$$R(\mathbf{x}, n) = \mathbf{w}^T \mathbf{x}_n$$

Incorporate the template threshold to make new function:

$$g(\mathbf{x}, n) = R(\mathbf{x}, n) - \omega_0 = \boxed{\mathbf{w}^T \mathbf{x}_n - \omega_0}$$

With match criterion at the time when maximum correlation occurs:

$$\boxed{\begin{array}{c} g(\mathbf{x}, n') > 0 \rightarrow match \\ g(\mathbf{x}, n') < 0 \rightarrow no\ match \end{array}}$$

Linear Discriminant Functions

Modified correlation function looks like linear discriminant function:

$$g(\mathbf{u}) = \boxed{\mathbf{w}^T \mathbf{u} - \omega_0}$$

Match criterion looks like LDF classification rules:

$$\boxed{\begin{array}{c} g(\mathbf{u}) > 0 \rightarrow class\ 1 \\ g(\mathbf{u}) < 0 \rightarrow class\ -1 \end{array}}$$

Time Dependence: Waveform Correlation and LDF equations

Linear Discriminant Functions

$$g(\mathbf{u}) = \mathbf{w}^T\mathbf{u} - \omega_0$$

$$g(\mathbf{u}) > 0 \rightarrow class\ 1$$

$$g(\mathbf{u}) < 0 \rightarrow class\ -1$$

Waveform Correlation

$$g(\mathbf{x}, n) = R(\mathbf{x}, n) - \omega_0 = \mathbf{w}^T\mathbf{x}_n - \omega_0$$

Time-dependent linear discriminant function.

$$g(\mathbf{x}, n') > 0 \rightarrow match$$

$$g(\mathbf{x}, n') < 0 \rightarrow no\ match$$

LDF classification is equivalent if

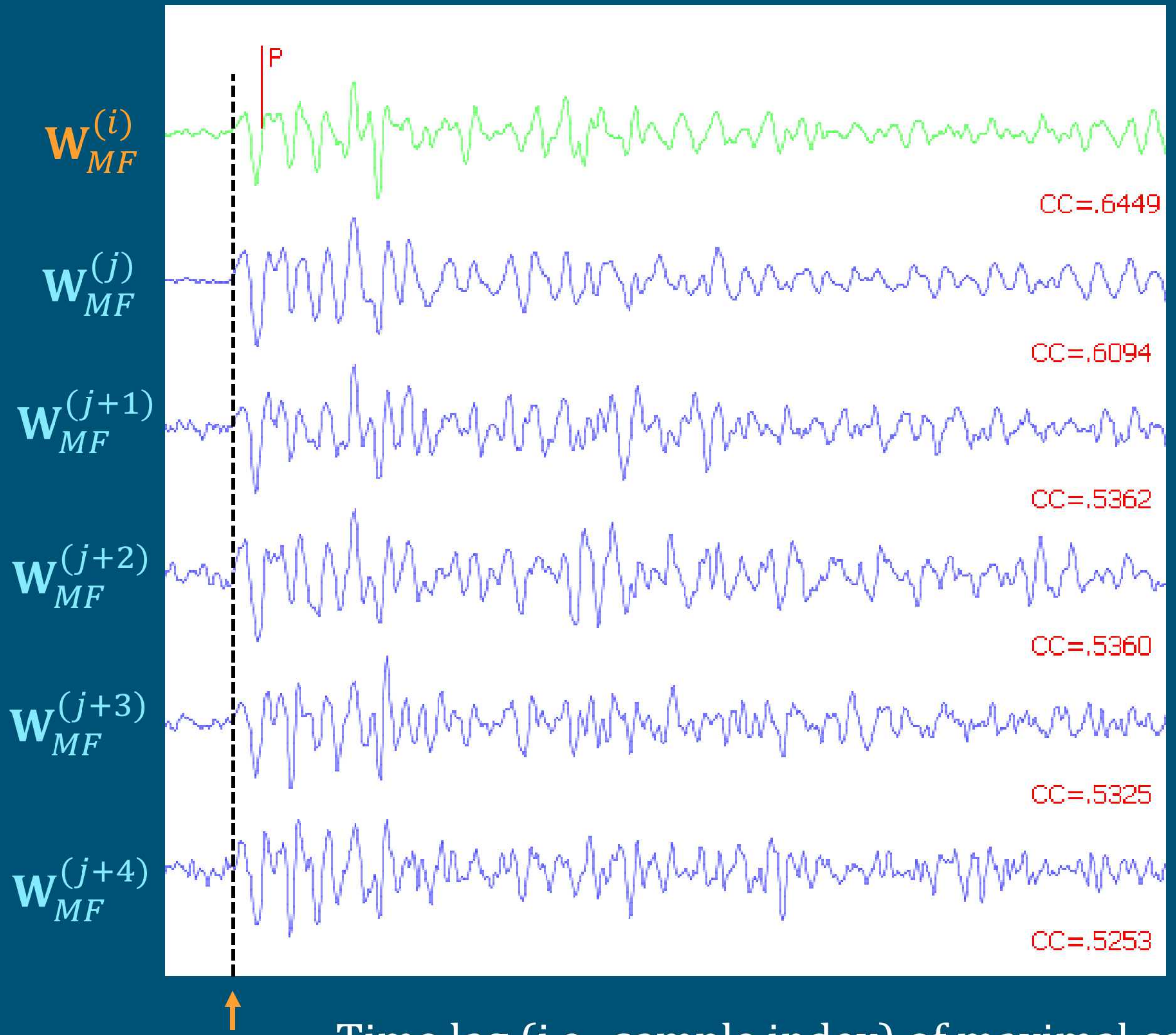
$$\boldsymbol{u} = \mathbf{x}_{n'}$$

Insight: Overcome the time dependence by evaluating the waveform correlation functions for a single time index that is known prior to training the ANCorr templates.

Agenda

Training: Determine time of maximal correlation n'

1. **Start with a matched filter template library of arrival waveforms at a station.**
2. **Choose the *i*th template.**
3. **Circularly shift every other template in the matched filter library to align with the time lag of maximal correlation.**



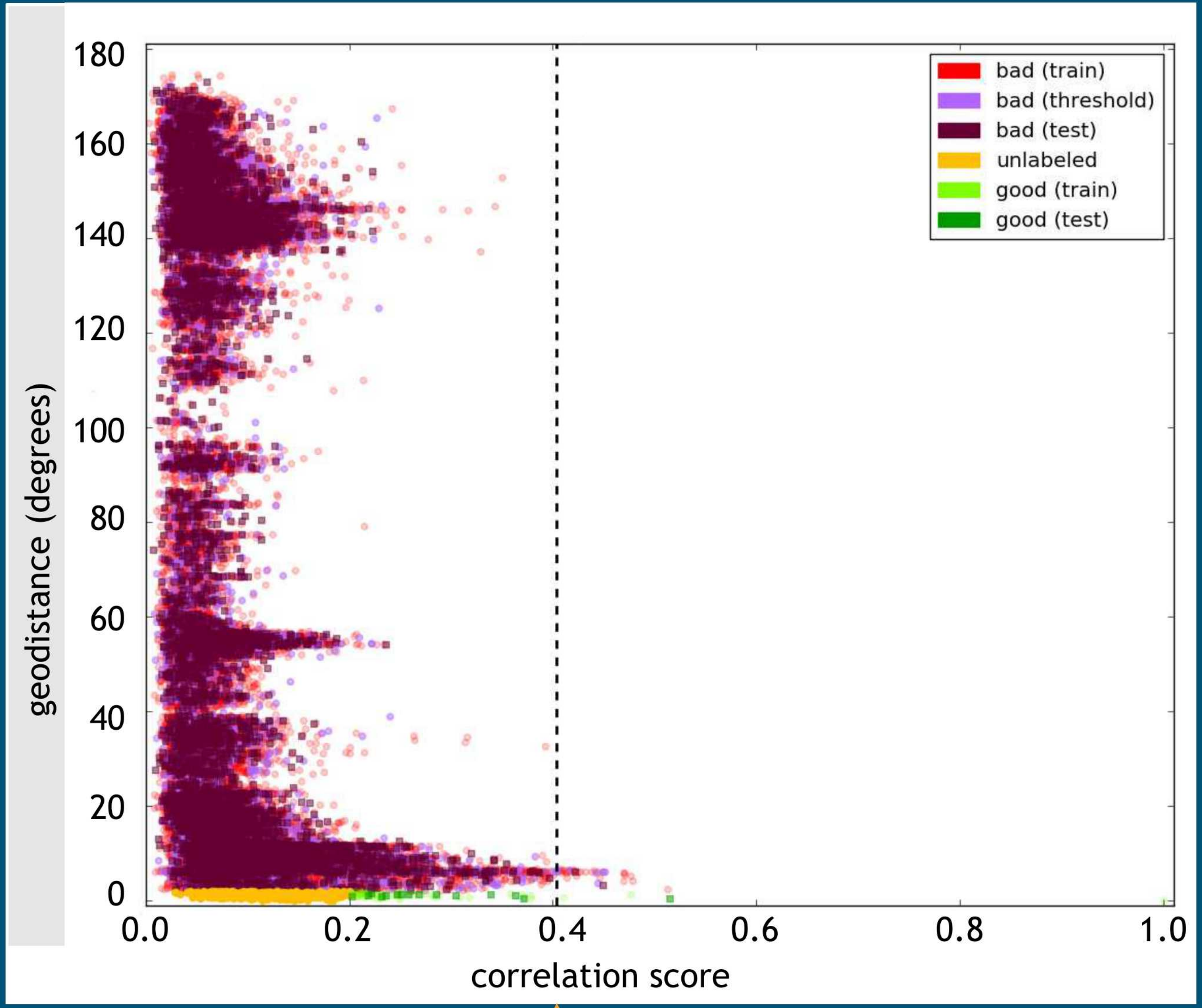
$\boldsymbol{n'}_{\boldsymbol{MF}}^{(\boldsymbol{i},\boldsymbol{j})}$ Time lag (i.e., sample index) of maximal correlation for *i*th template and every *j*th template, i ≠ j

Training: Label Matched Filter Templates

Label templates to make a training set for the *i*th template.

- Only template waveforms from nearby sources are desired as matches
- Most templates in the library are labeled *No match* training class for a geodistance of 2 degrees.

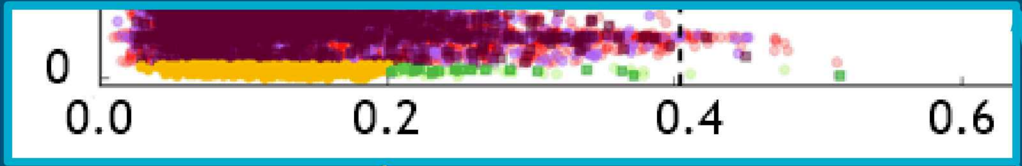
Choose a template correlation threshold that rejects 99.9% of templates in the *No match* training class.



Template Threshold

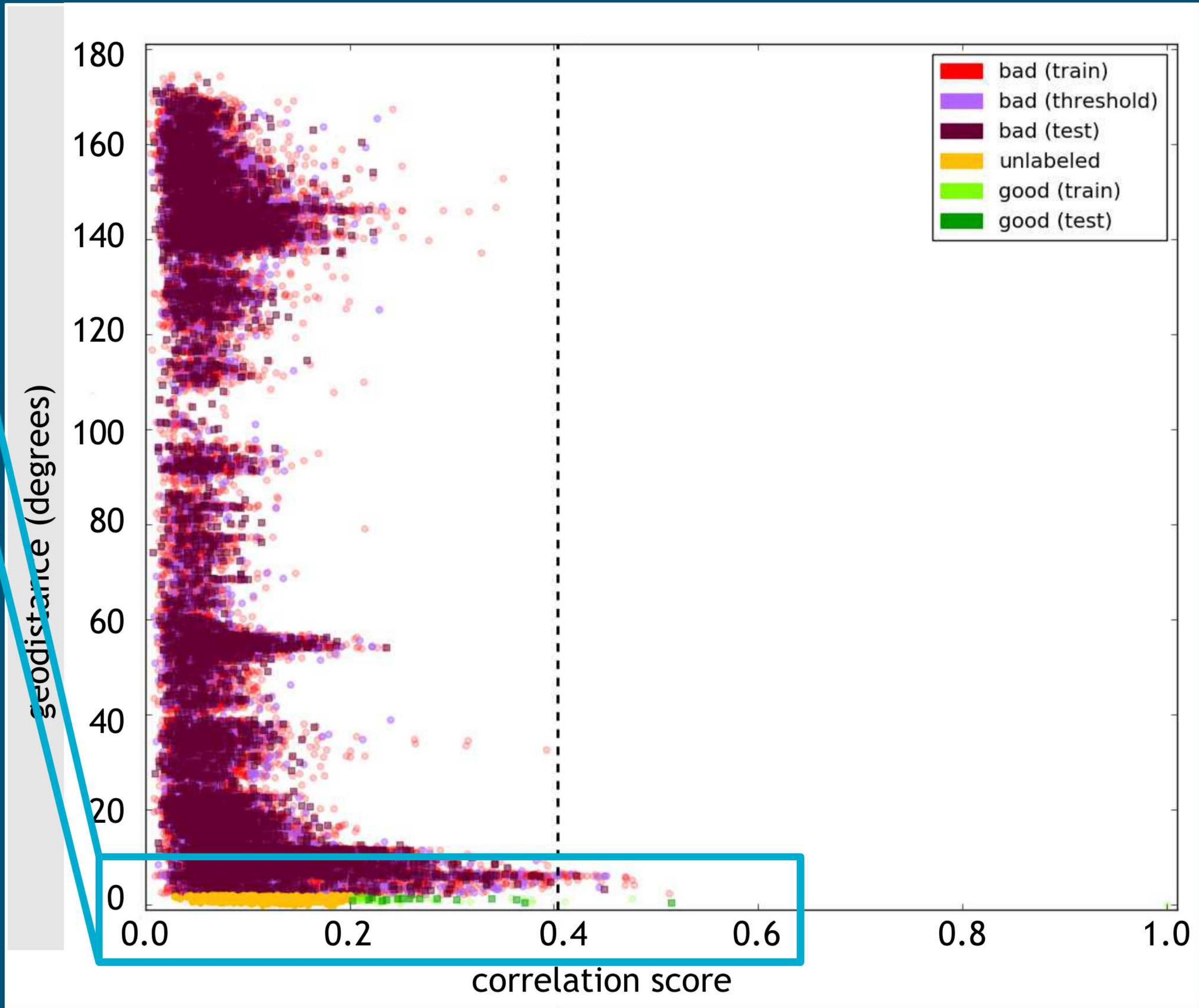
Training: Waveform similarity as a tuning parameter

What about multiple sources or different source types that are geographically collocated?



similarity constraint = 0.5

The **similarity constraint** restricts the waveforms in the *match* class according to correlation with matched filter template. Dissimilar waveforms are removed from the training set.



Template Threshold

Training: Synthesize ANCorr Template from Training Set

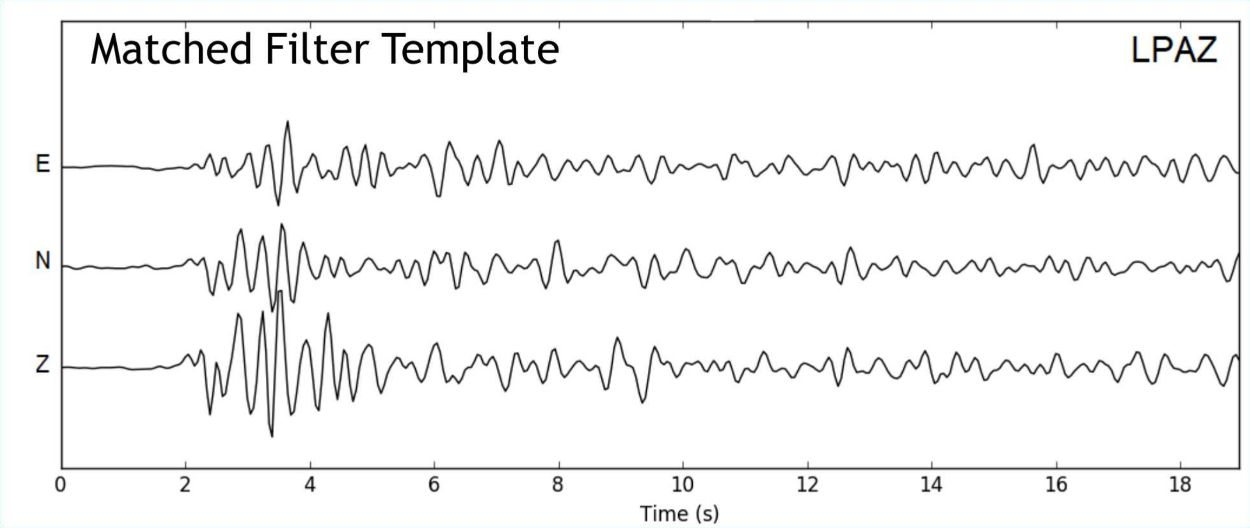
Synthesize the ANCorr template by applying machine learning algorithm to the training set of time-shifted and labeled matched filter templates.

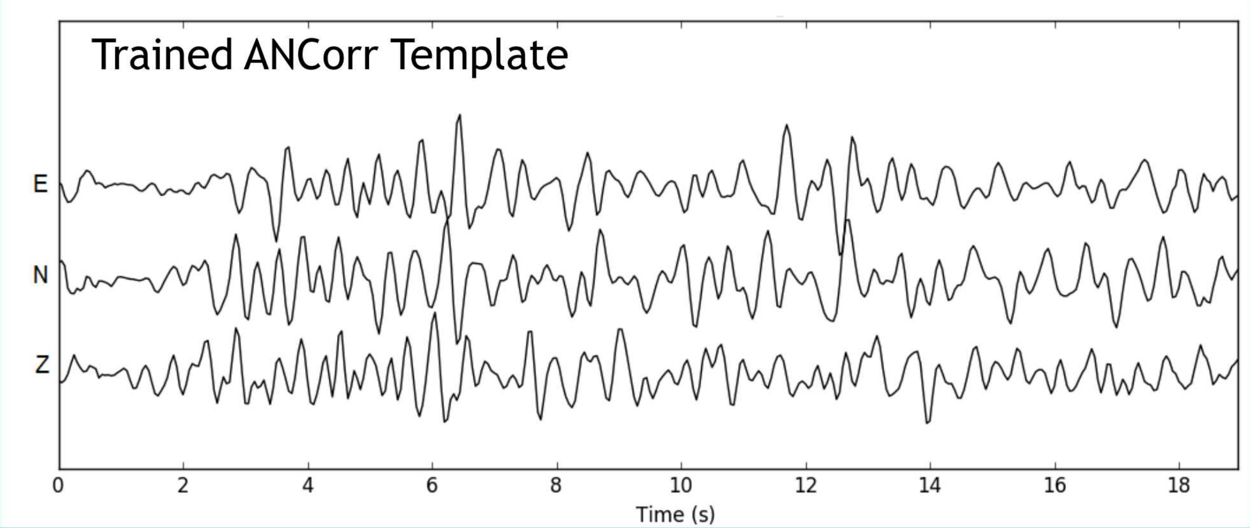
We used a Linear Support Vector Machine (LSVM) implementation in Python.

Set ANCorr Template Threshold

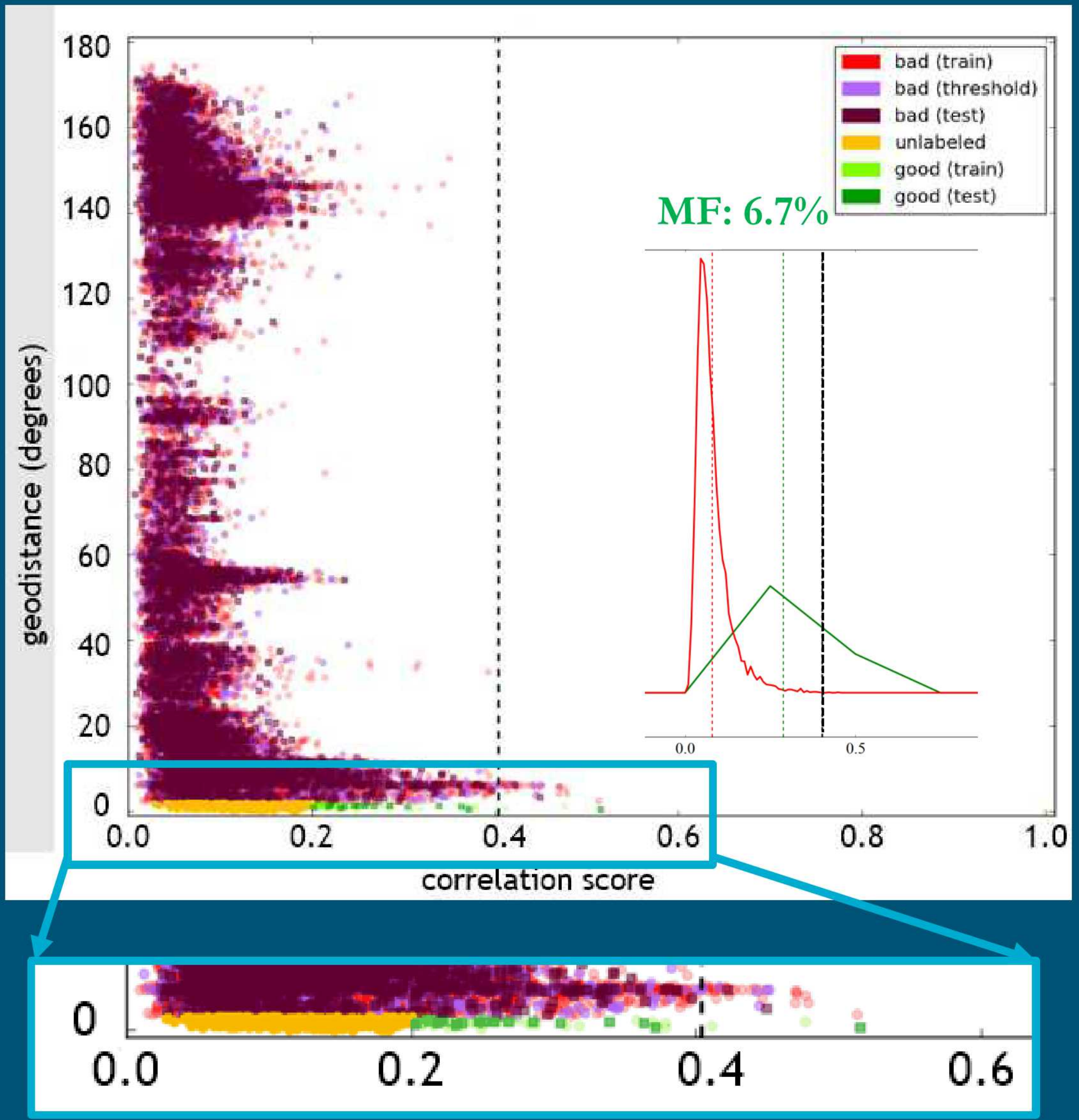
1. Correlate the ANCorr template with the training set.
2. Set the ANCorr template threshold to reject 99.9% of templates in the *No match* training class.

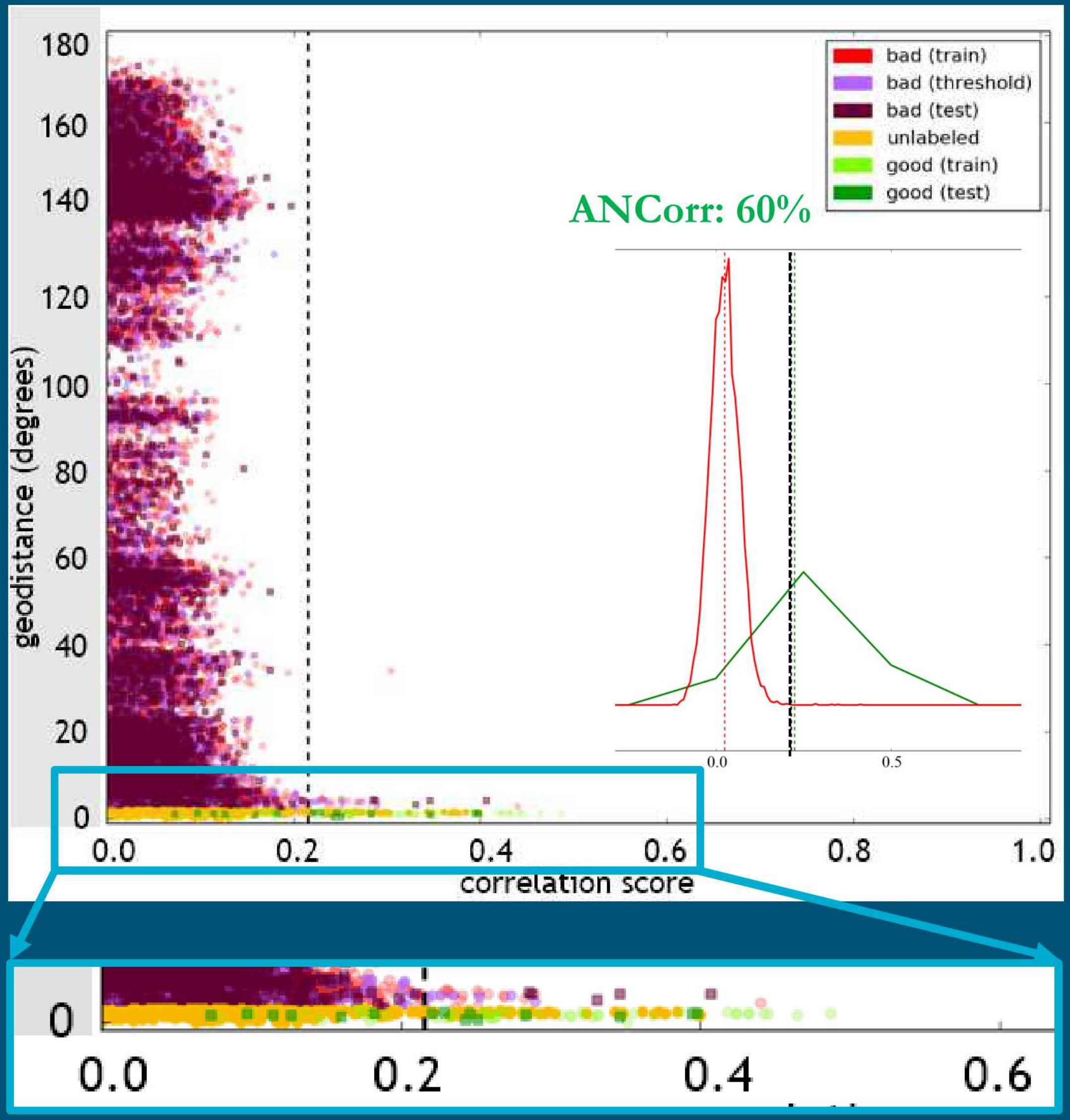
Late Event Bulletin [LEB] P (-15.97, -75.74), 1 January 2006 03:01:39





Matched Filter and ANCorr Template Example

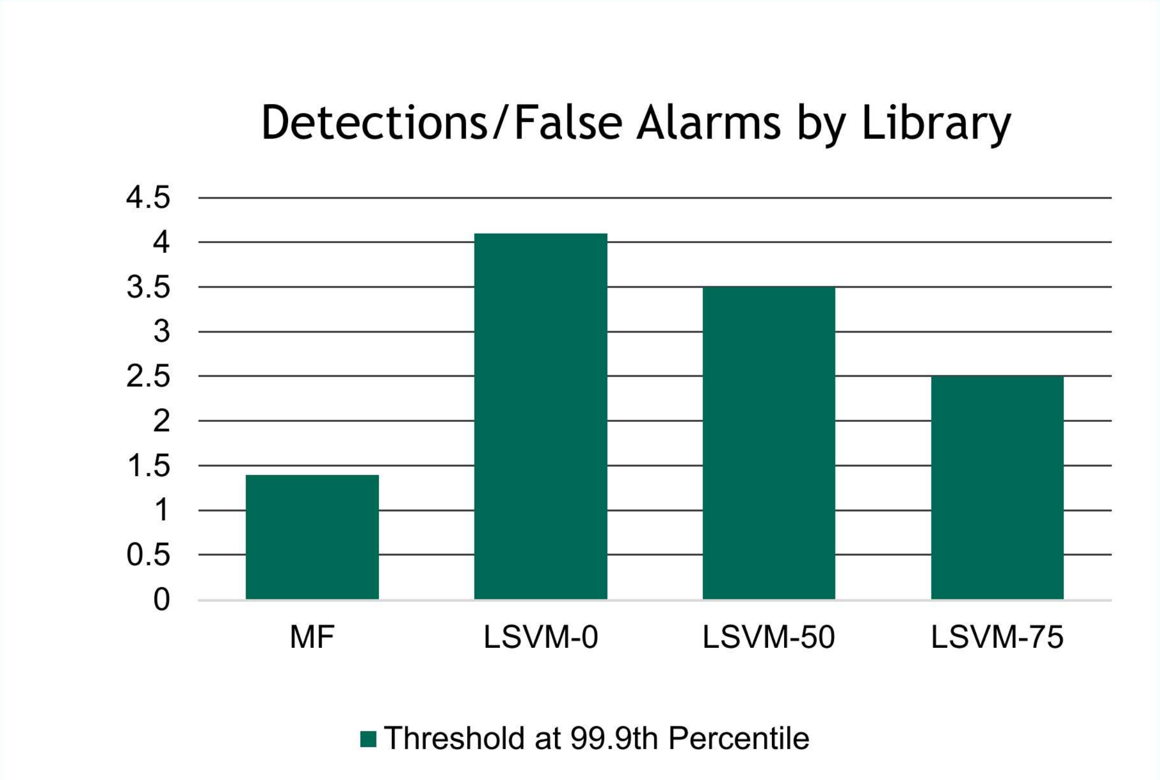




Experimental Set-Up

- We created template libraries from the three-component station LPAZ, located in La Paz, Bolivia.
- We chose a set of 51,690 template arrivals from the International Data Centre (IDC) Late Event Bulletin (LEB) for the time range of 2006 to 2016 as the training set.
- The month of May 2010 was reserved for a labeled test set, not part of training set.
- The template libraries were applied to 663 arrivals from the labeled test set.
- Application Method:
  1. Correlate with MF template to detect matches exceeding threshold.
  2. Correlate with ANCorr template at time of maximal correlation to reject false alarms.
- The detection/false alarm ratio was chosen as the metric for comparison.
- Four ANCorr libraries were trained, with *Match* class parameterized by similarity constraint.

Results



Library	Description
MF	Matched Filter Library from seismic arrivals at station LPAZ.
All ANCorr template libraries were trained by LSVM where the ***Match*** training class membership included MF waveforms for events within 2 degrees geodistance and:	
LSVM-0	no similarity constraint.
LSVM-50	similarity constraint 0.5 (correlation score exceeded 50% of template threshold).
LSVM-75	similarity constraint 0.75 (correlation score exceeded 75% of template threshold).

Overall, the ANCorr templates demonstrated an improved ratio of detections to false alarms over the matched filter library for this dataset.

- The best performing library (LSVM-0) used all waveforms within 2 degree geodistance of template event.
- Removing dissimilar waveforms from the *Match* class decreased the overall detection/false alarm ratio.

Research Contributions

- Defined an alternative detection model for waveform correlation that improves upon the conventionally applied matched filter method by modifying the null hypothesis to account for non-collocated seismicity.
- Recognized the mathematical similarity of linear discriminant functions to waveform correlation equations, justifying the use of machine learning for template synthesis.
- Simplified the time dependence of waveform correlation so that linear discriminant functions can be used for template synthesis.
- Developed a method of template synthesis that uses the new detection model.

References

Ganter, T., A. Sundermier, and S. Ballard (2018). Alternate Null Hypothesis Correlation: A New Approach to Automatic Seismic Event Detection. *Bulletin of the Seismological Society of America,* https://doi.org/10.1785/0120180074.

Scharf, L. L. (1991). *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Chapter 4: Neyman-Pearson Detectors, First Ed., Addison-Wesley Publishing Company, Reading, Mass.

Duda, R. O., P. E. Hart, and D. G. Stork (2001). *Pattern Classification*, Chapter 5: Linear Discriminant Functions, Second Ed., John Wiley & Sons, Inc., New York, New York.