Unsupervised Seismic Anomaly Detection

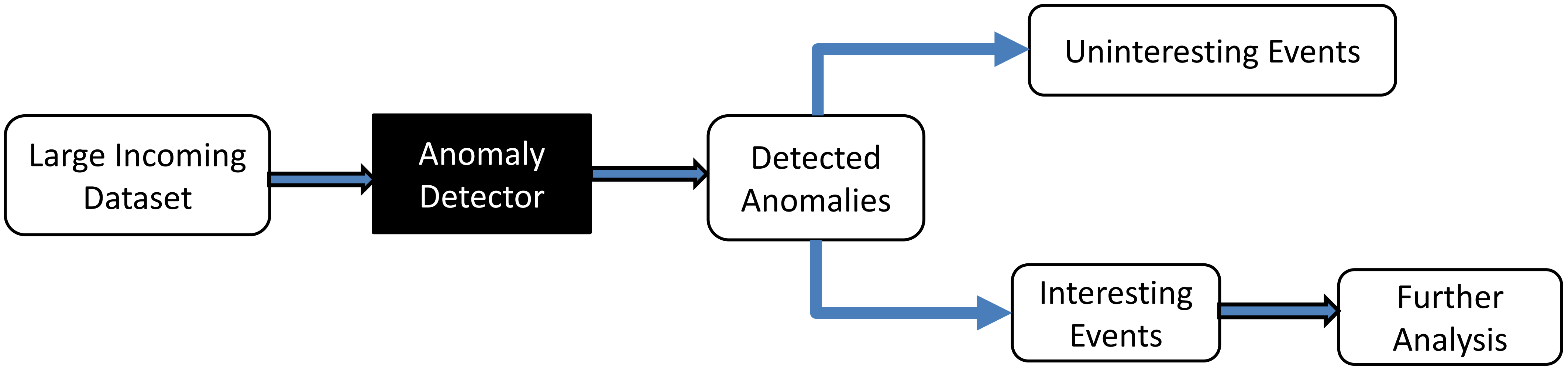
Srija Chakraborty
Maxfield Thompson
Mentors: James Theiler, Diane Oyen
08/10/2017

Outline

- Motivation
- Dataset
- Unsupervised Approach
- Approaches
    - Short Term Average-Long Term Average (STA-LTA)
    - Spectrogram as Feature
    - Mahalanobis Distance
    - Autoencoder
- Results
- Future Work

Motivation

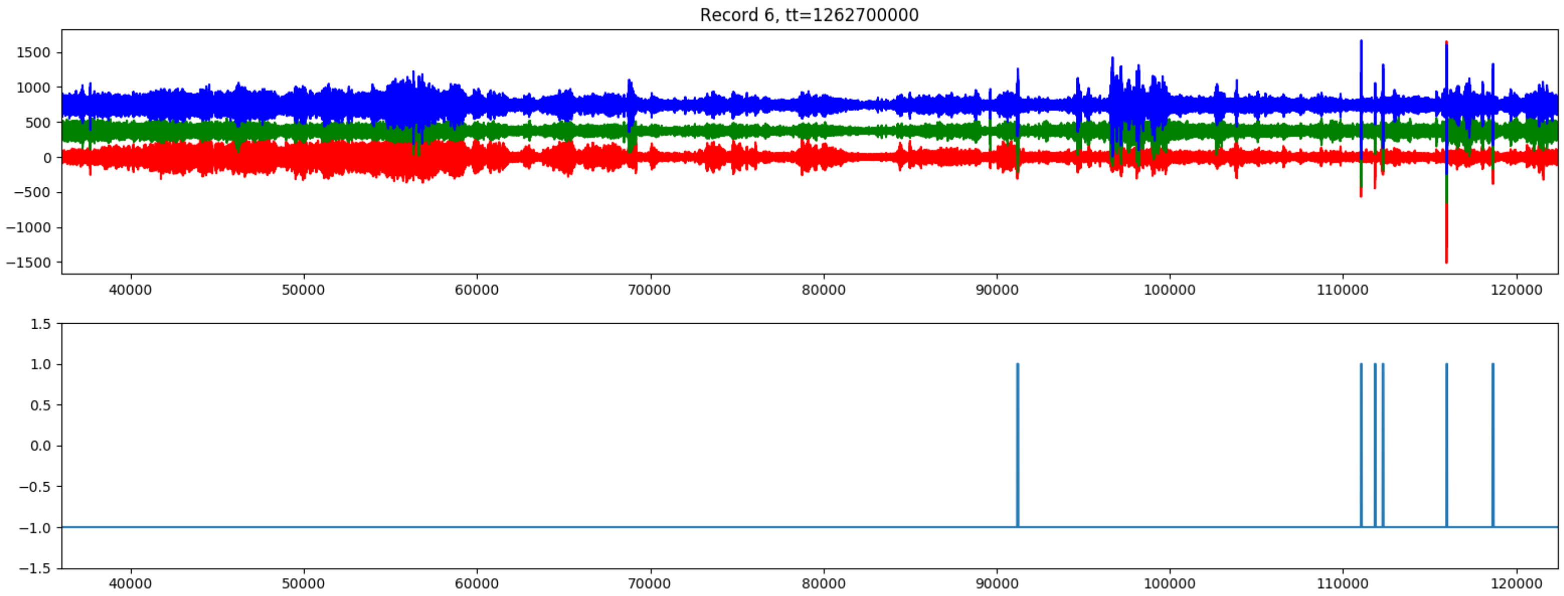
- **Anomalies** : Data samples that are unlike the rest of the data



- **Final Objective –** In a large seismic dataset, identify events that are interesting to seismologists.
- **Intermediate Objective** - Find events that are unusual (includes both interesting and uninteresting events).

Dataset

- Seismic data from a single station
- Three channels
- Labeled events (earthquakes, mine blasts)
- Has events that are not labeled



Top – Signal from all sensors
Bottom – Labeled Events

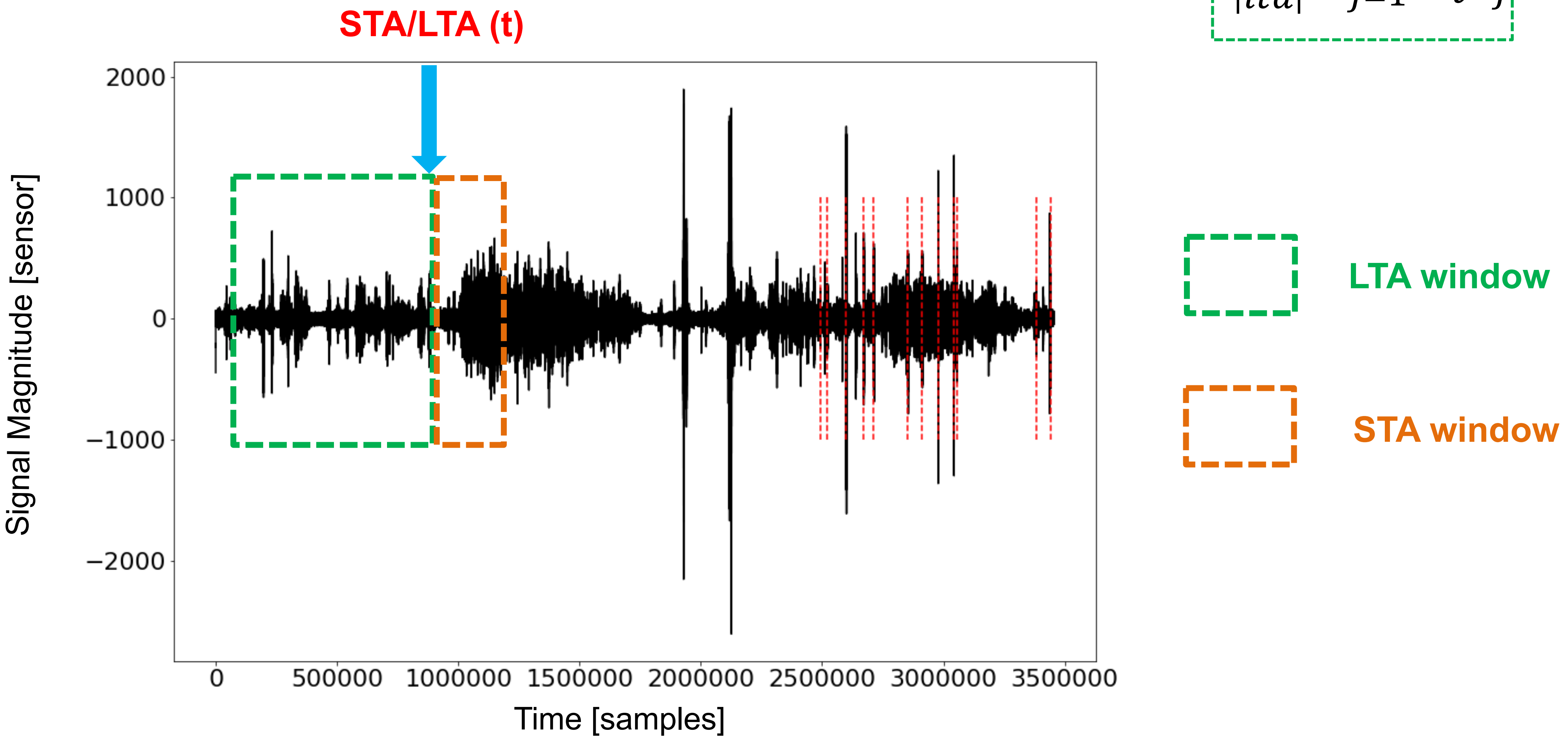
Unsupervised Learning

- Arrival time of primary (P) and secondary (S) waves
  - regional Geology
  - distance of event from sensor
- Incomplete labels

Generic model independent of local properties

Short Term Average/Long Term Average

- Short Term Average/Long Term Average, STA/LTA(t) = $\dfrac{\frac{1}{|sta|}\sum_{i=1}^{|sta|} x_{i+t}^2}{\frac{1}{|lta|}\sum_{j=1}^{|lta|} x_{t-j}^2}$



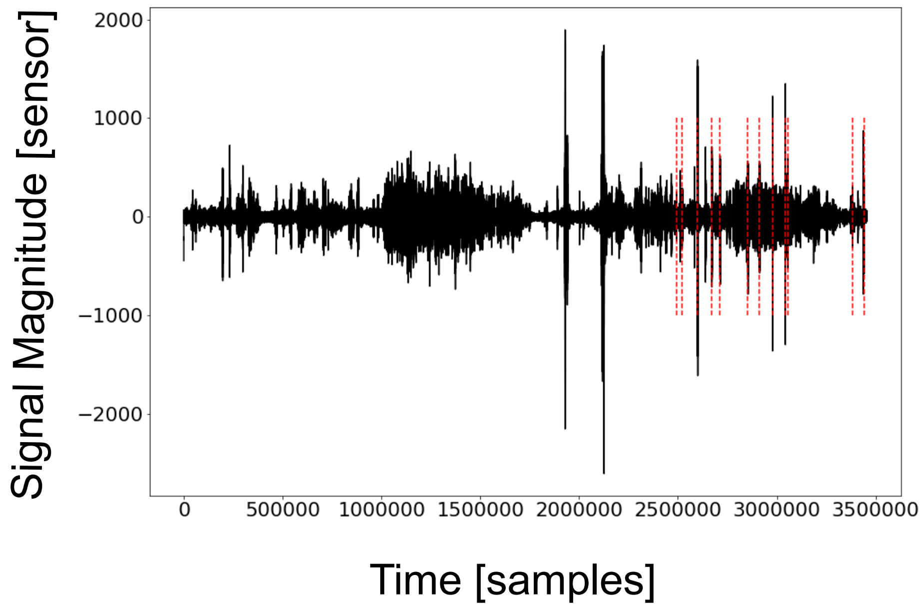
Spectrogram as Feature

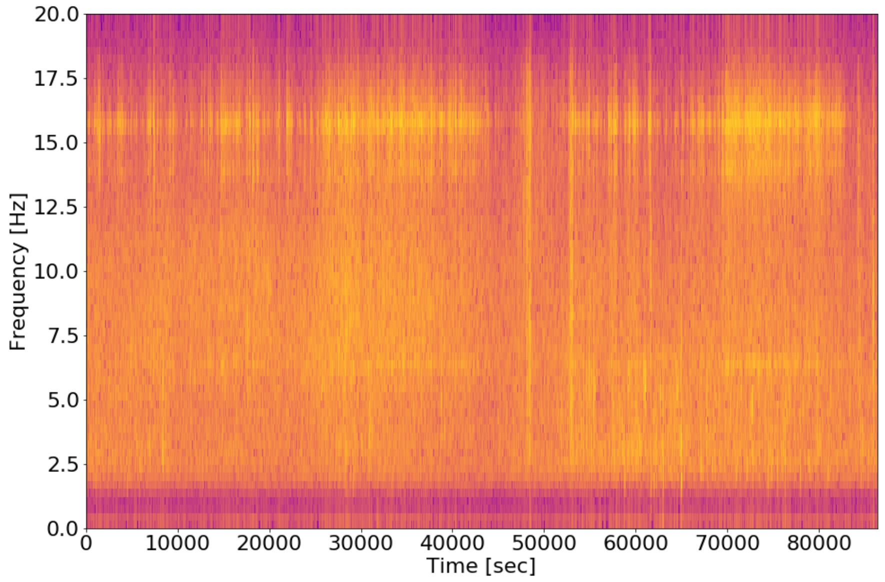
Magnitude Square of Short Term Fourier Transform (STFT)

$$\mathbf{STFT}\{x(t)\}(\tau,\omega) \equiv X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j\omega t}\,dt$$

$$\mathrm{spectrogram}\{x(t)\}(\tau,\omega) \equiv |X(\tau,\omega)|^2$$

Represents raw signal as a time varying power spectrum

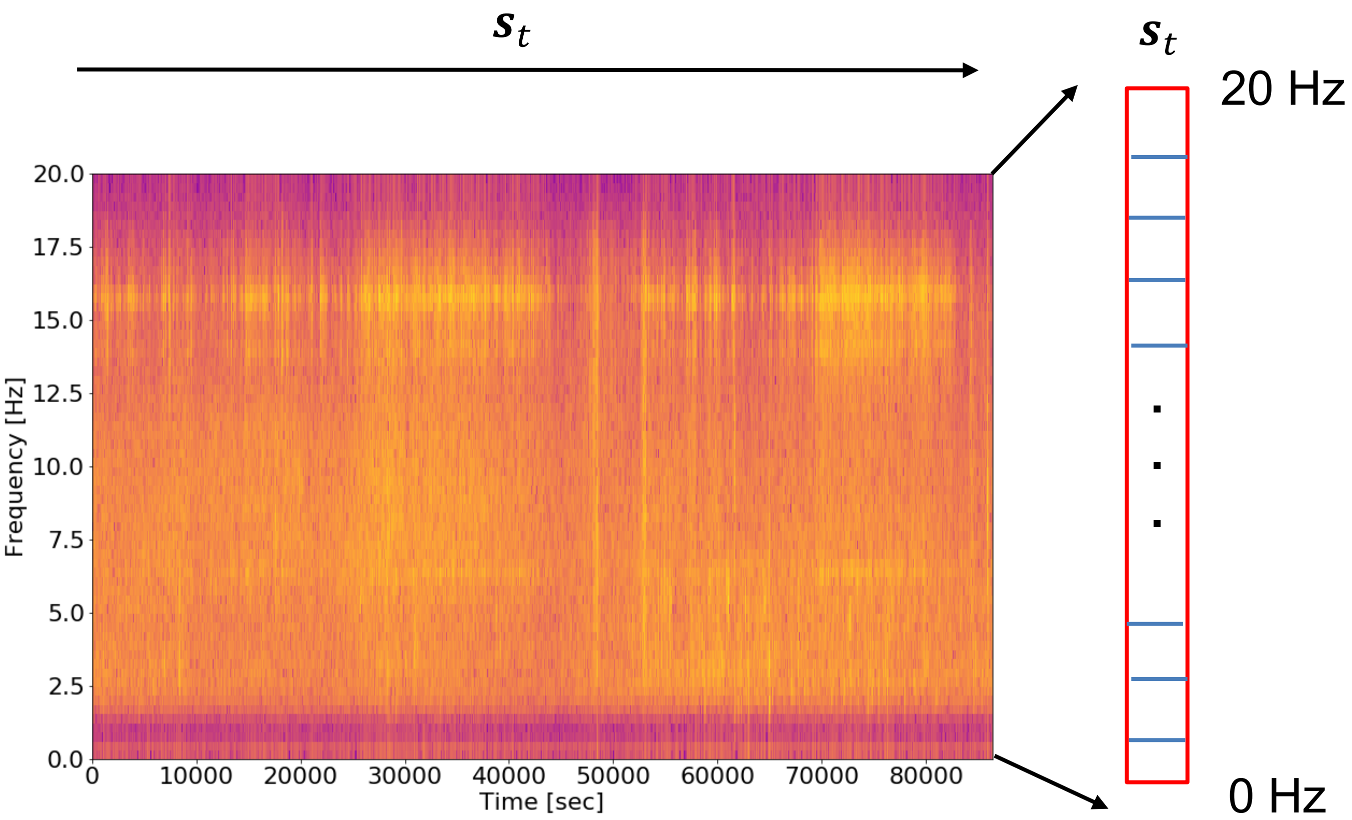




Spectrogram

Mahalanobis Distance

Mahalanobis distance d_t, to spectrogram mean



Spectrogram over a day

$$d_t = \sqrt{(\boldsymbol{s}_t - \bar{\boldsymbol{s}})^T R^{-1} (\boldsymbol{s}_t - \bar{\boldsymbol{s}})},$$

where $\boldsymbol{s}_t$ spectrogram vector at t,

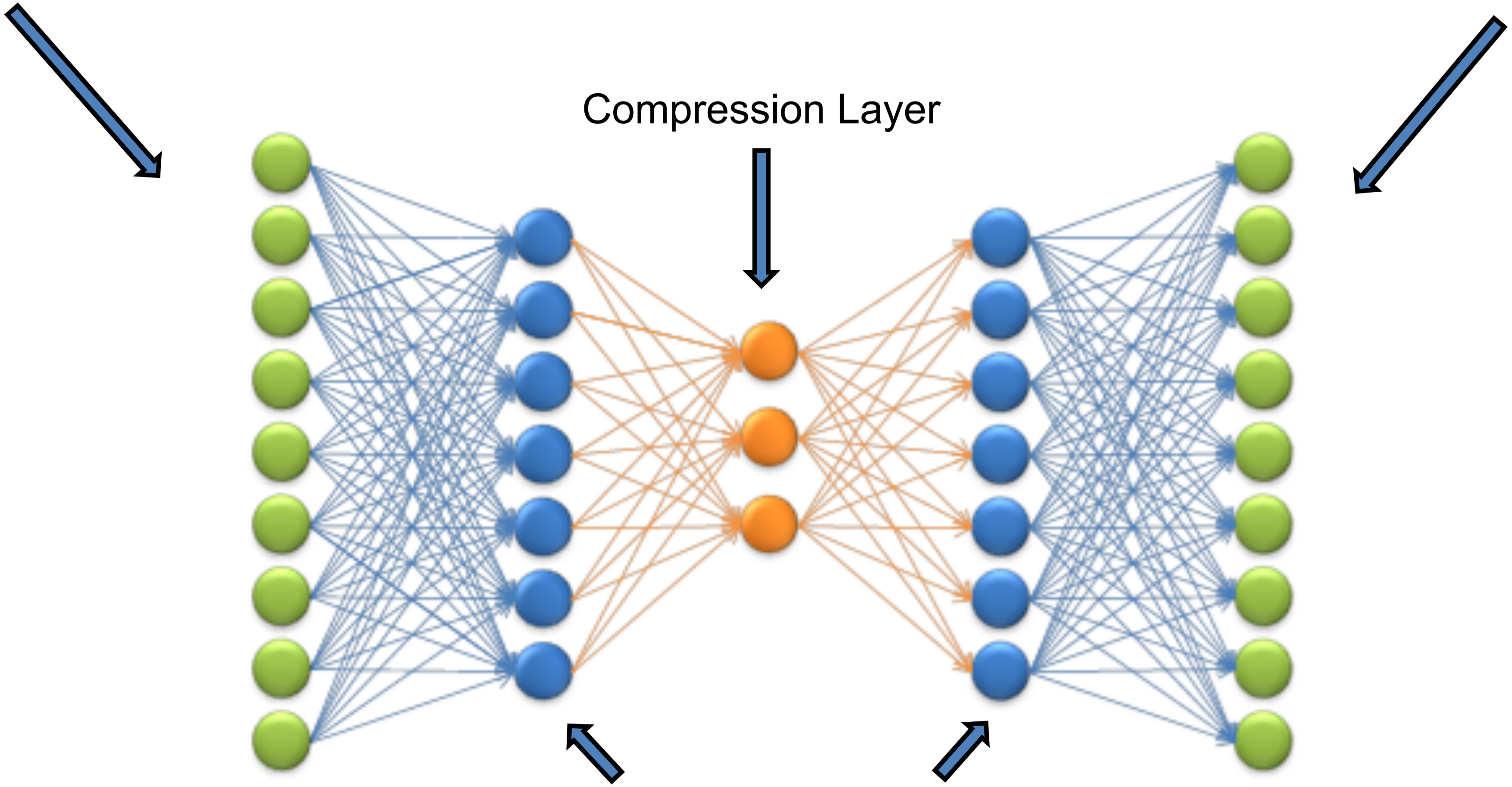
$$\left.\begin{array}{l} R \text{ is the covariance matrix,} \\ \bar{\boldsymbol{s}} \text{ mean of the spectrogram} \end{array}\right\} \text{Over a day}$$

Autoencoder

Input Layer [spectrogram vector]

Output Layer [spectrogram vector]

Compression Layer

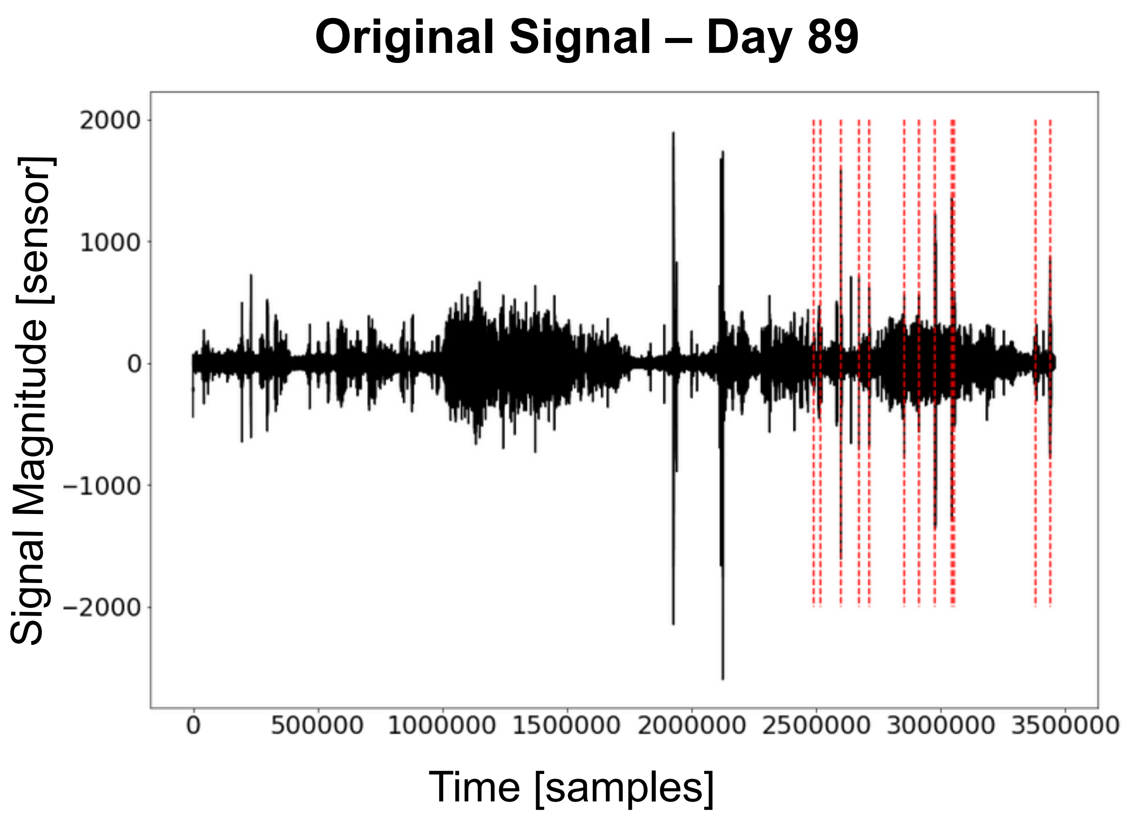


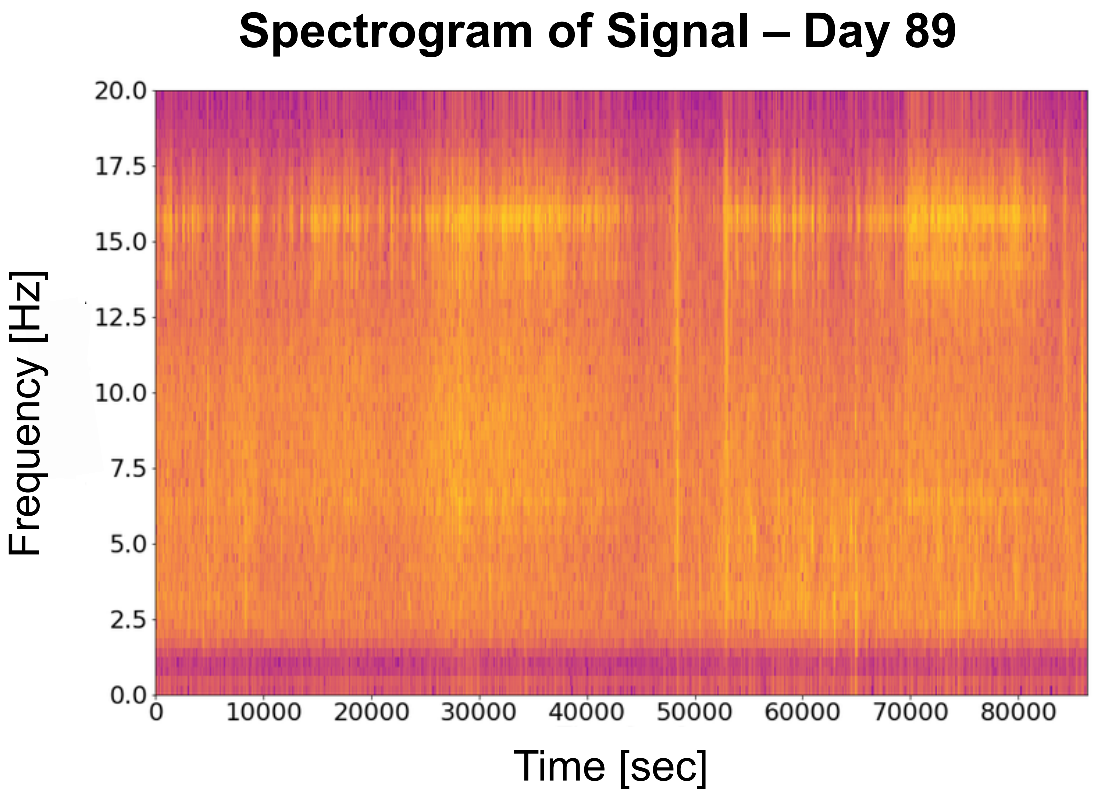
Hidden Layers

https://velesnet.ml/docs/_images/autoencoder.png

Autoencoder Approach

- Unsupervised Learning Approach:
  - **Autoencoder**: Learn a compressed spectrogram representation of the background class
  - **Input Feature**: Three spectrogram vectors
  - **Assumption:** High reconstruction error implies anomalies
  - **Expectation:** Labeled and unlabeled events lie in top k% of reconstruction error



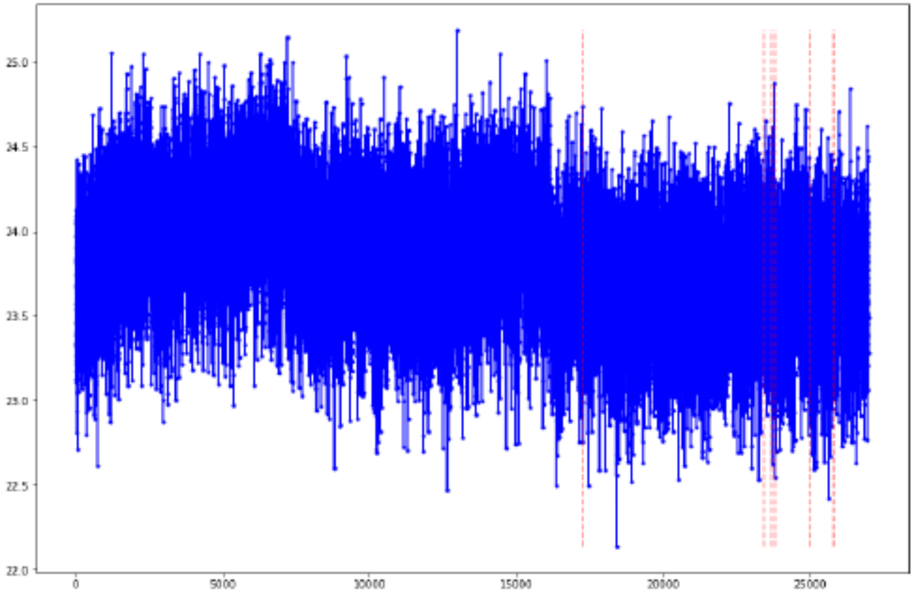


Hyper-parameter Selection

- In order to create the best autoencoder, a series of experiments were created to select the optimal parameters. These experiments systematically ran through hundreds of autoencoders and examined their reconstruction errors on a series of days to see which set of parameters best caught the labeled anomalies.

Reconstruction Error – Day 6

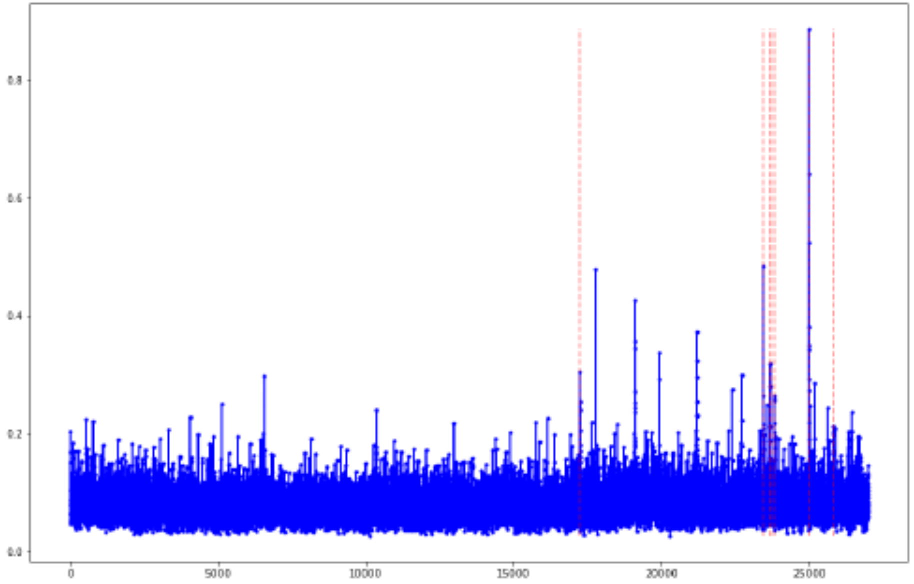
Reconstruction Error



Time [3.2 sec window]

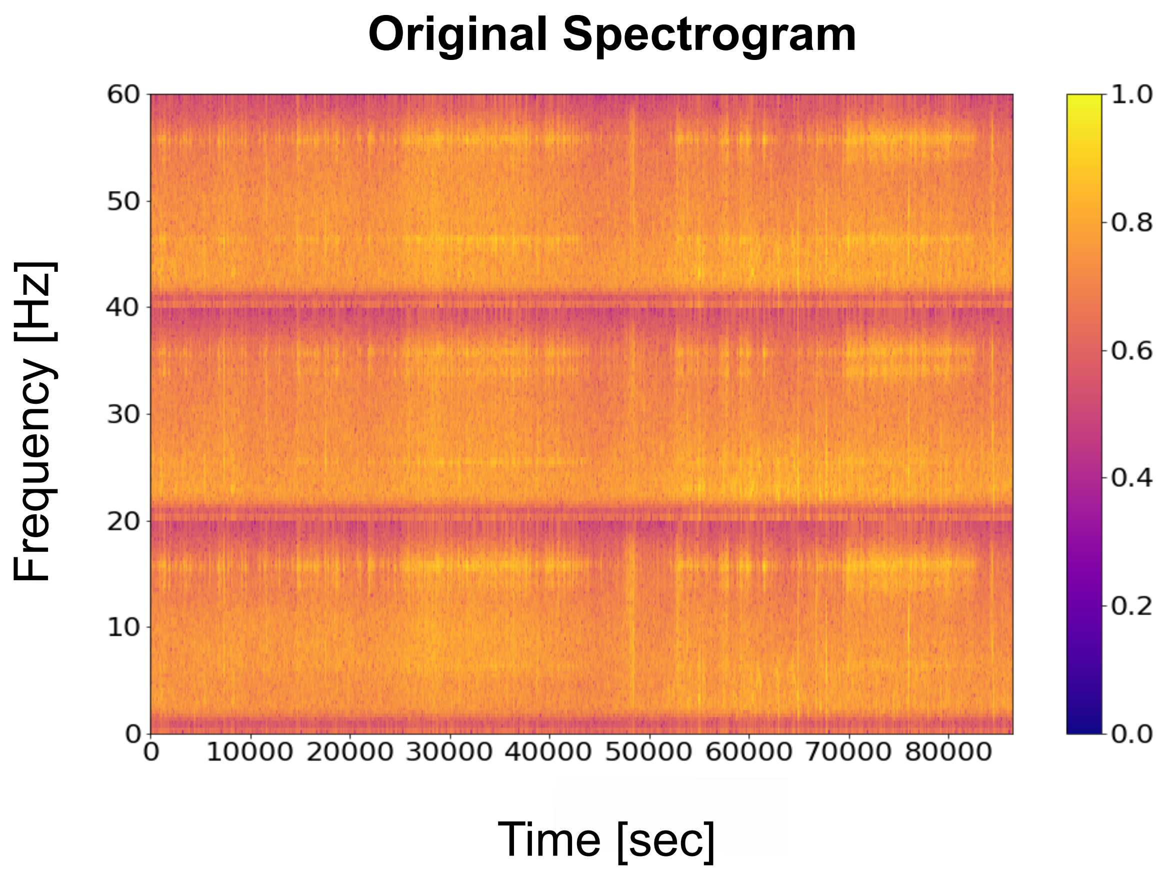
Reconstruction Error – Day 6

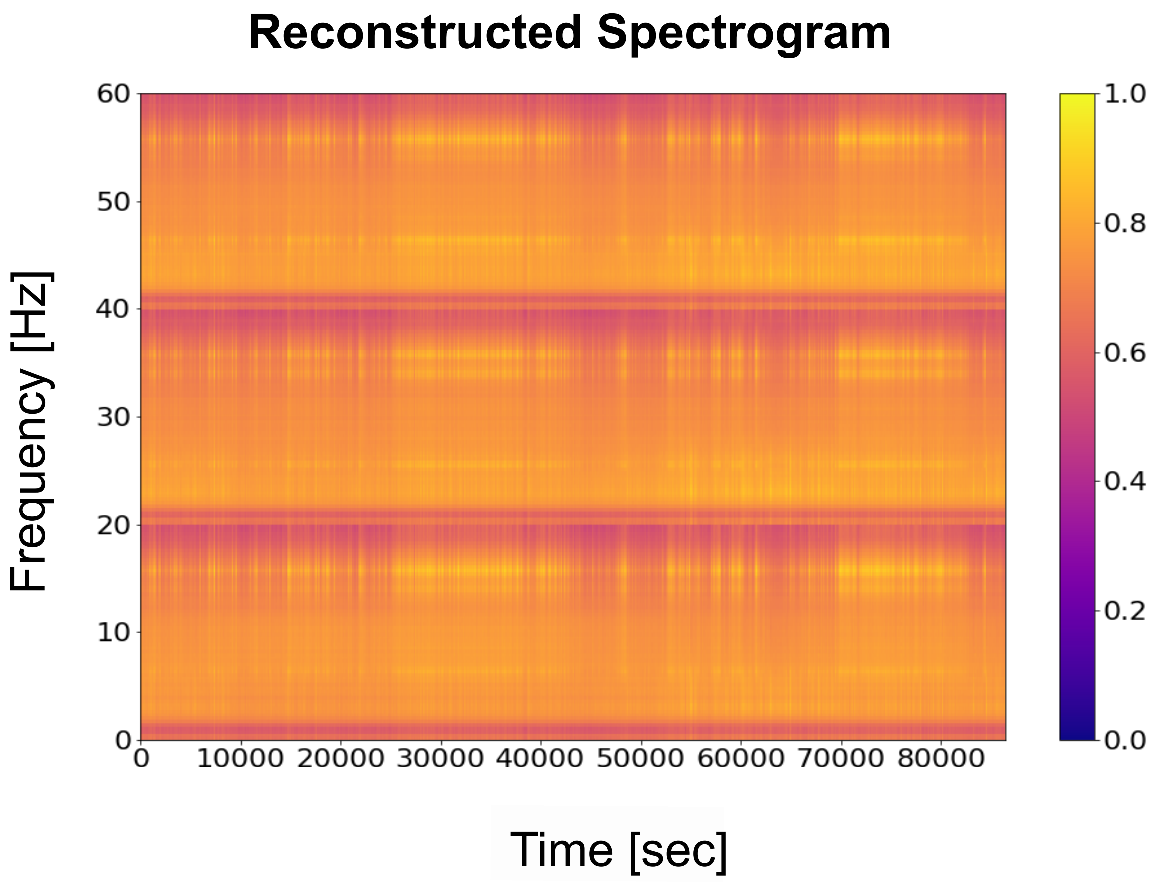
Reconstruction Error



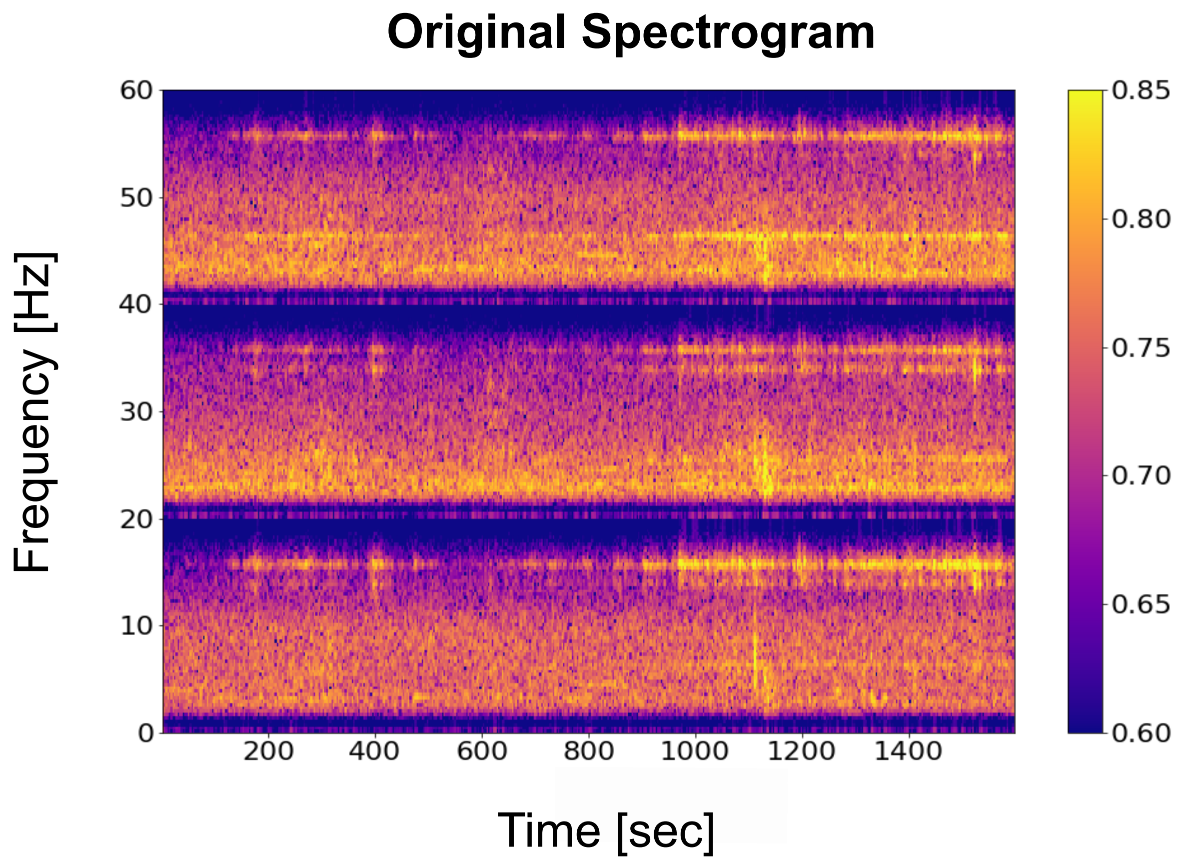
Time [3.2 sec window]

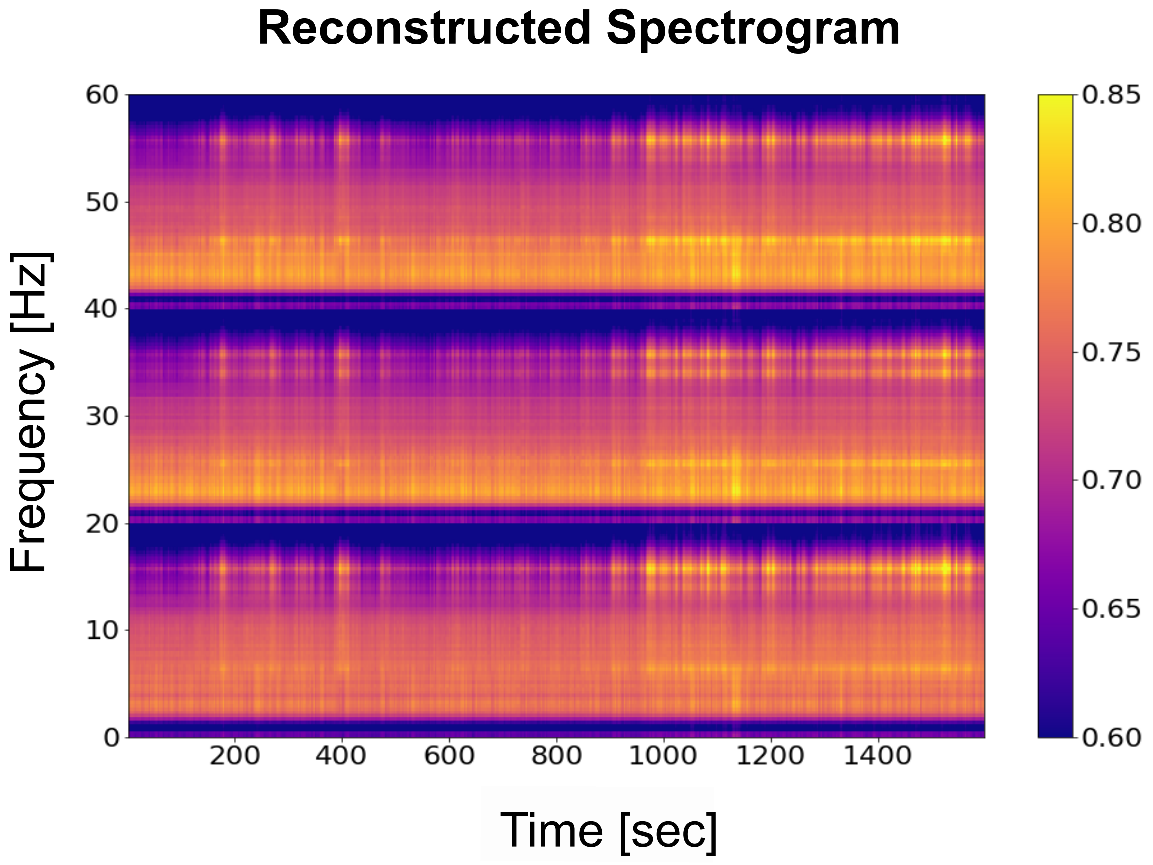
Spectrograms

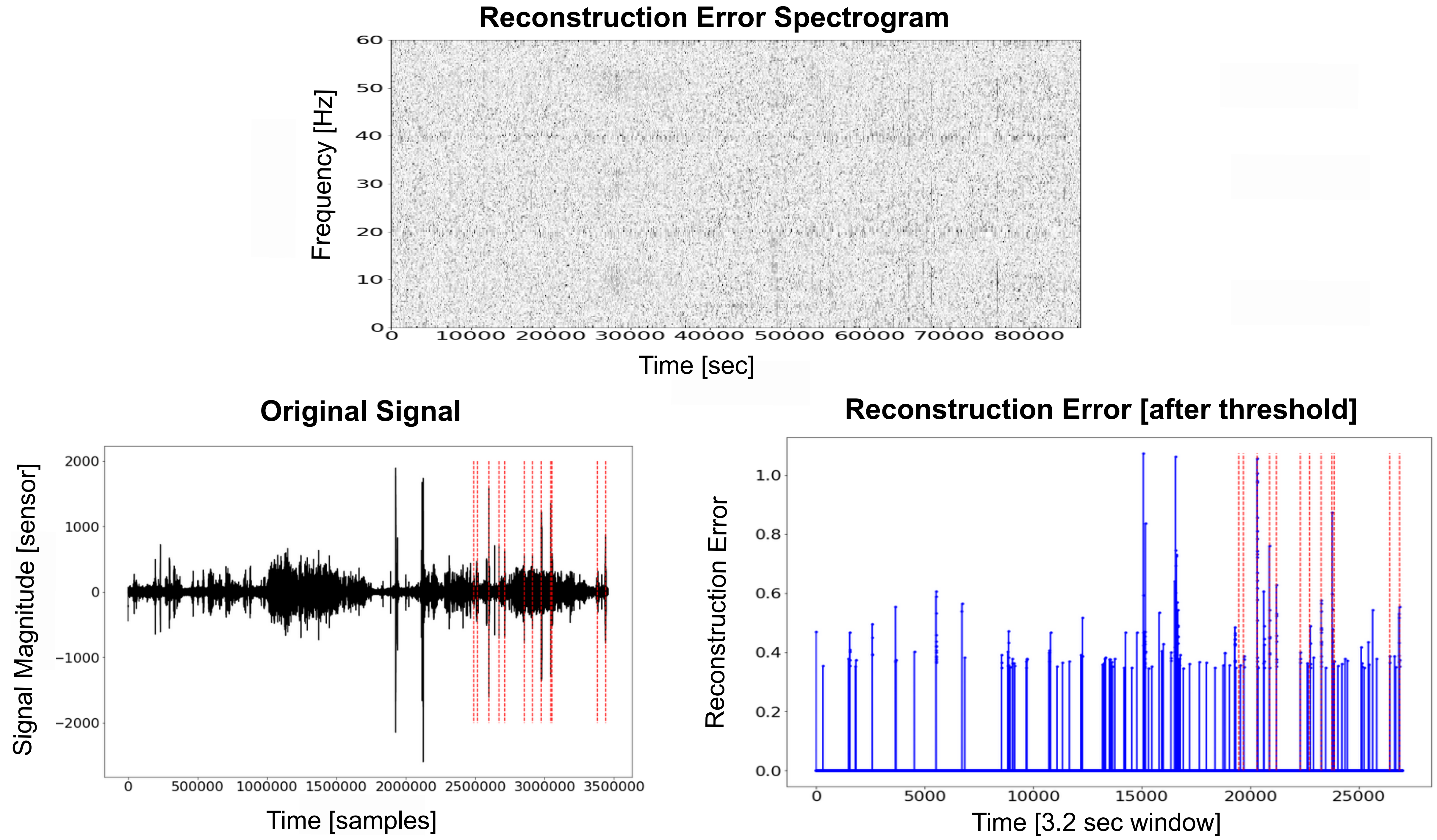




Spectrograms In-depth






Reconstruction Error Spectrogram
Frequency [Hz]
0
10
20
30
40
50
60
0
10000
20000
30000
40000
50000
60000
70000
80000
Time [sec]
Original Signal
Signal Magnitude [sensor]
2000
1000
0
−1000
−2000
0
500000
1000000
1500000
2000000
2500000
3000000
3500000
Time [samples]
Reconstruction Error [after threshold]
Reconstruction Error
0.0
0.2
0.4
0.6
0.8
1.0
0
5000
10000
15000
20000
25000
Time [3.2 sec window]

Channel Selection Experiments

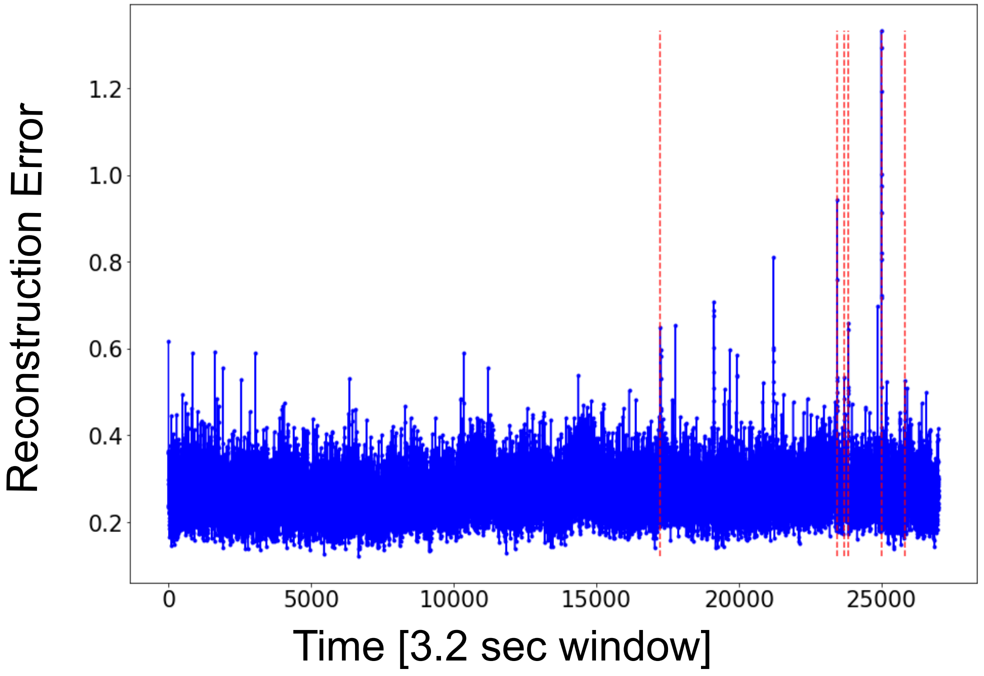
- P wave [4 to 12 Hz] and S wave [1 to 4 Hz]
- Autoencoder on 4-12 Hz
- Significantly lowered the overall accuracy of the autoencoder.

There is important information that characterizes an earthquake in frequency bands not classically associated with the p and s waves.

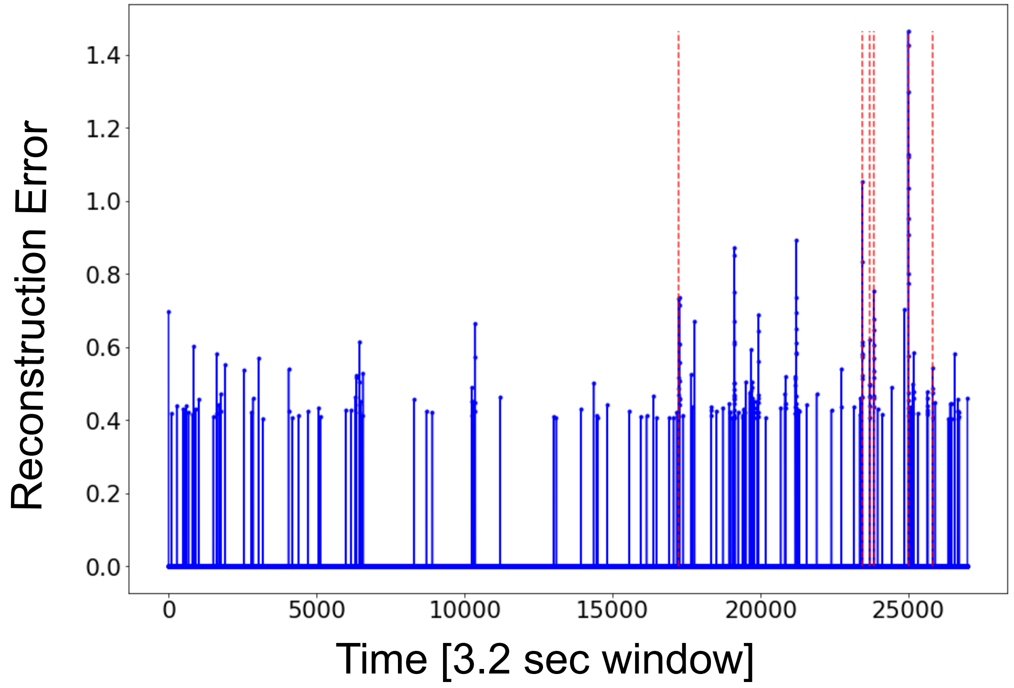
Threshold Selection

- Various thresholds were selected as methods to filter the reconstruction error. The most effective threshold tested was a raw percentile distribution cut. For all anomaly detection methods tested, we chose to examine the top 1 percent of reconstruction error values.

Reconstruction Error – Before Threshold



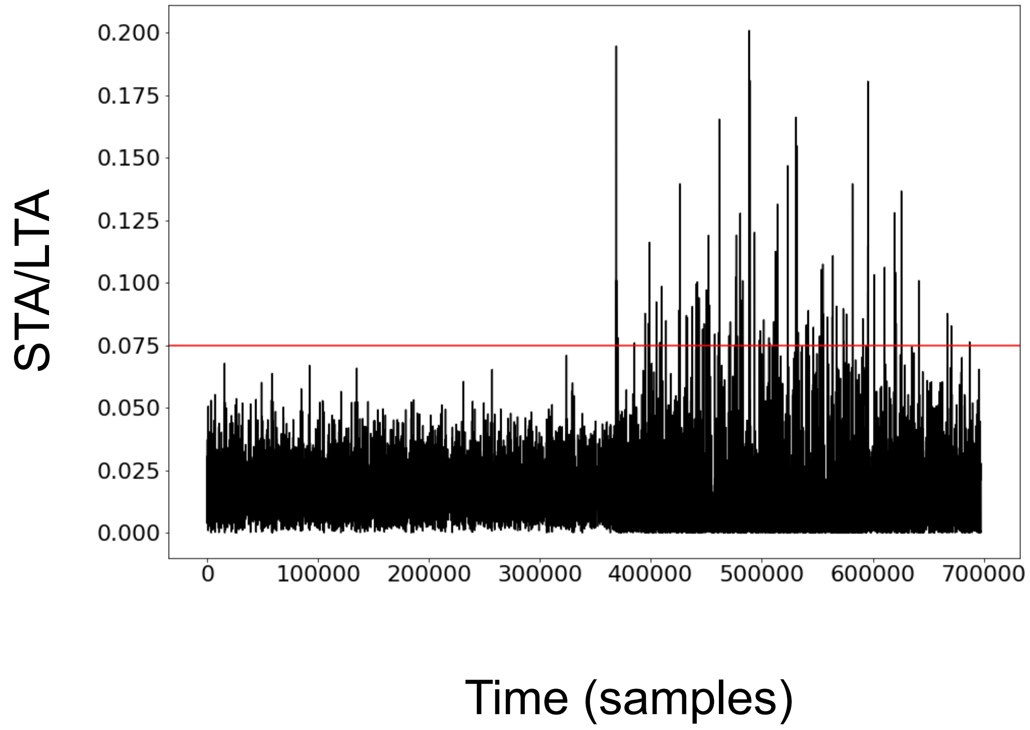
Reconstruction Error – After Threshold



Results

Comparison with other anomaly detectors:

- Mahalanobis distance to mean of the spectrogram
- Metric: percentage of labeled events detected



Approach	True Events Detected [threshold = top 1% of reconstruction error]
Autoencoder	86.45%
Mahalanobis Distance	65.45%

Future Work

- Other anomaly detectors
- Modeling trend on multiple days
- Optimum training data size - diurnal and short term patterns
- Geographic Associations – multiple sensors