HiDA HELMHOLTZ Information & Data Science Academy

Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum

אוניברסיטת בר-אילן Bar-Ilan University

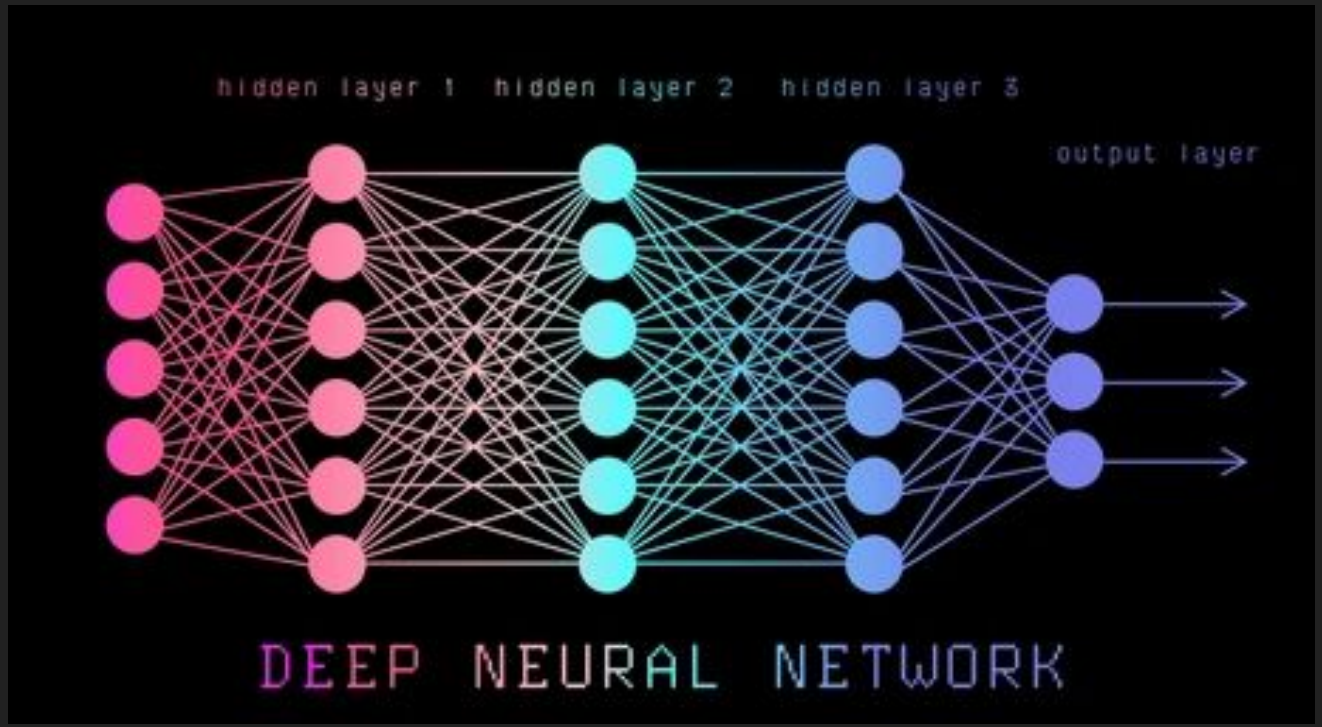
Quantifying Picker Uncertainty Tool

Presented By Moshe Beutel Bar-Ilan University BIU

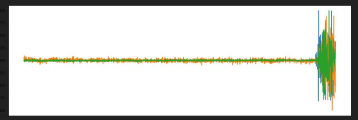
Work done during 6-week internship At GFZ Seismology

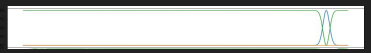
Supervised by: Prof. Frederik Tilmann,
Dr. Joachim Saul

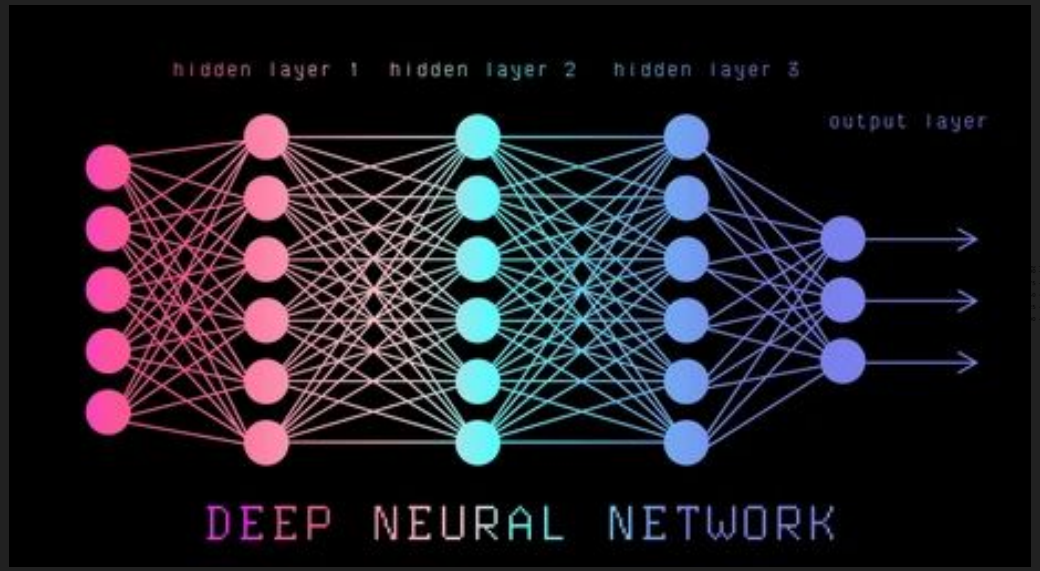
MACHINE LEARNING

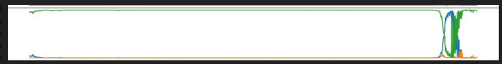


Machine Learning in Seismology









Seismic Datasets

ETHZ

GEOFON

INSTANCE

Iquique

LENDB

NEIC

SCEDC

STEAD

AI Models

PhaseNet

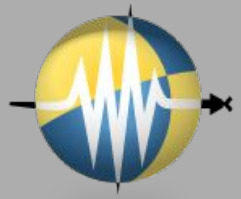
EQTransformer

CRED

GPD

BasicPhaseAE

Software



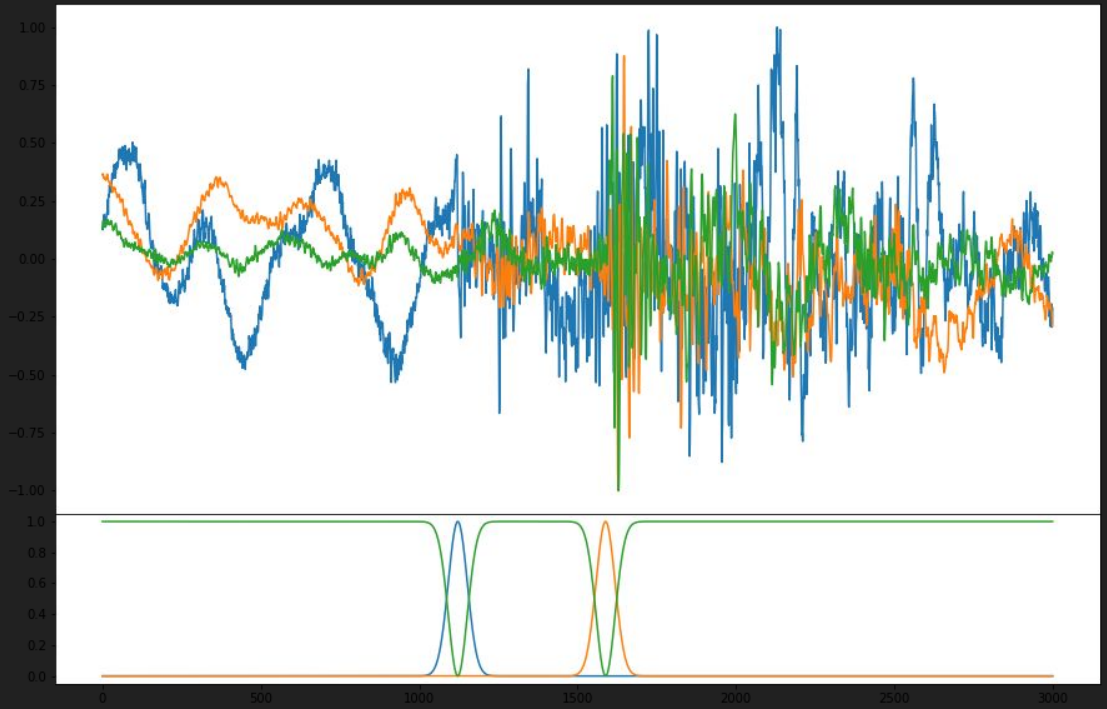
ObsPy

A Python Framework for Seismology

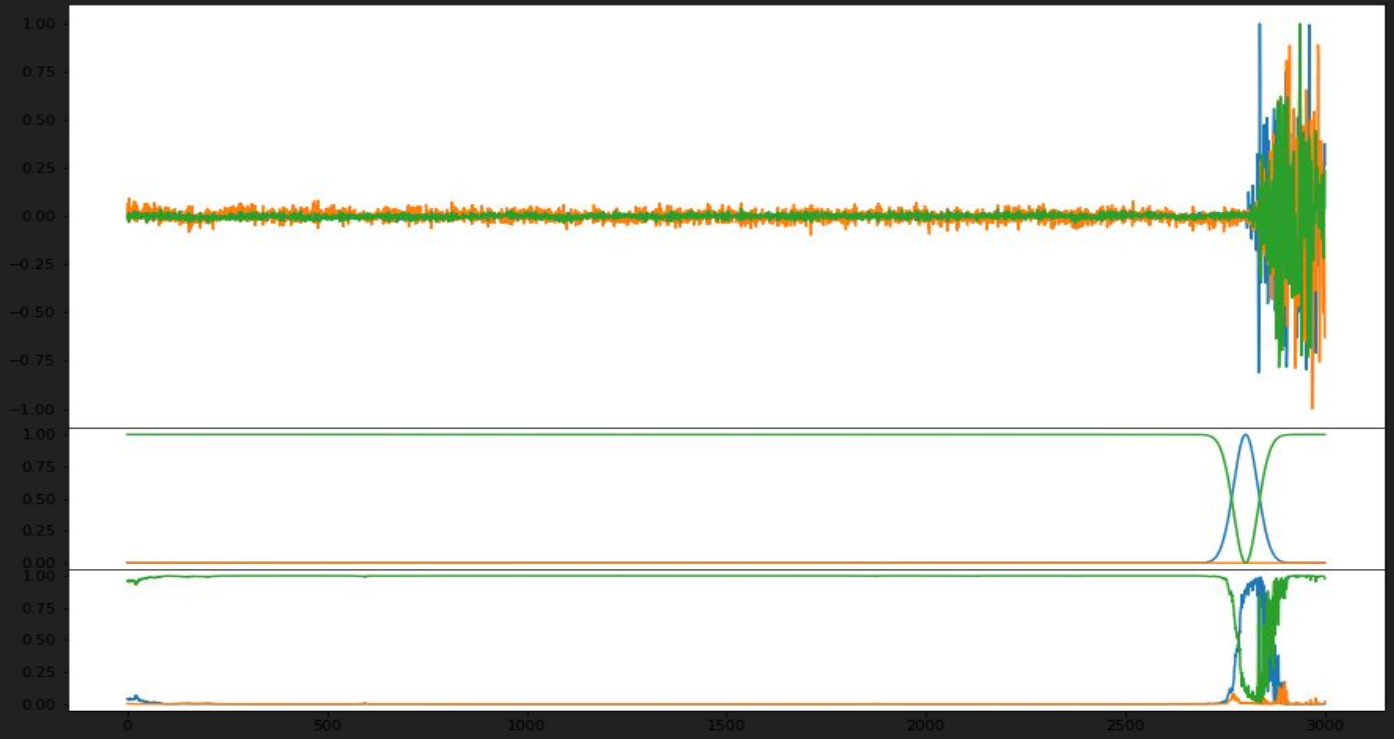
SeisBench

A toolbox for machine learning in seismology

Seismic Trace and Labels



Seismic Trace, Label and Prediction Probability Function



SNR Estimation

Signal/Noise Energy Ratio -

```python
@staticmethod
def _calc_snr_energy(trace, onset):
    En = np.mean(np.square(trace[:onset]))
    Es = np.mean(np.square(trace[onset:]))
    return CalcSNR.to_db(float(Es / En))
```

Max Amplitude/RMS Ratio -

```python
@staticmethod
def _calc_snr_max_amplitude_div_rms(trace, onset, window_length=100):
    max_signal_amp = np.max(np.abs(trace[onset: (onset + window_length)]))
    noise_rms = np.sqrt(np.mean(np.square(trace[:onset])))
    return 2 * CalcSNR.to_db(float(max_signal_amp / noise_rms))
```

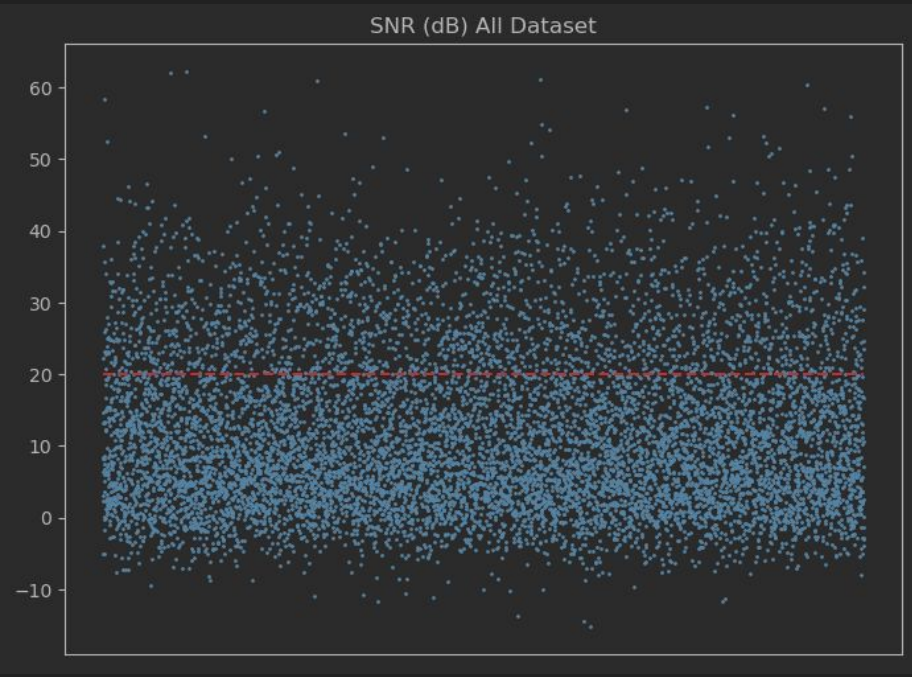
Seisynth Pipeline

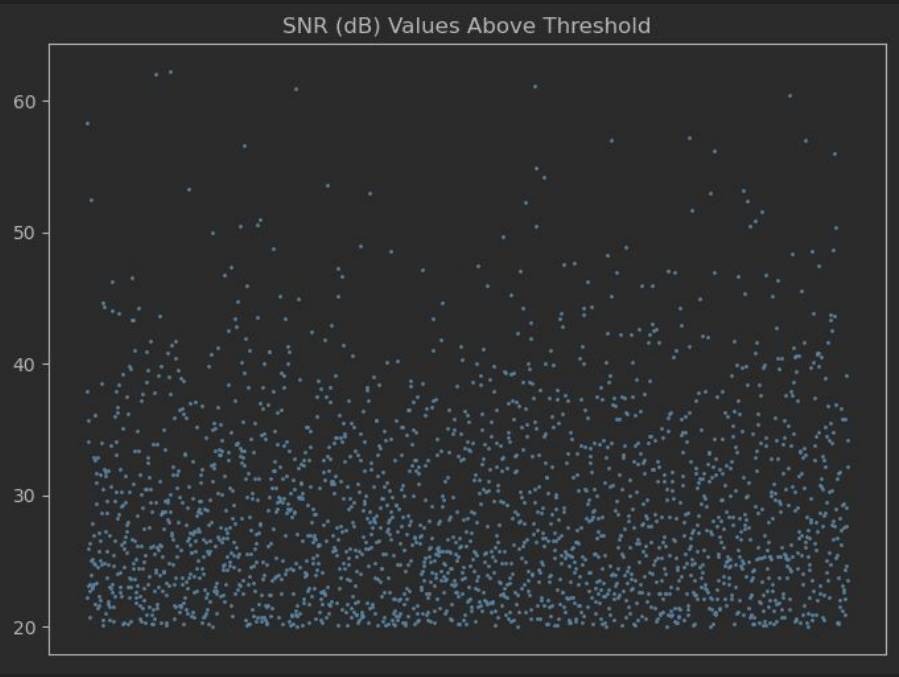
- Extracting high SNR traces from a dataset

- Creating a Noisy Dataset given SNR Definition

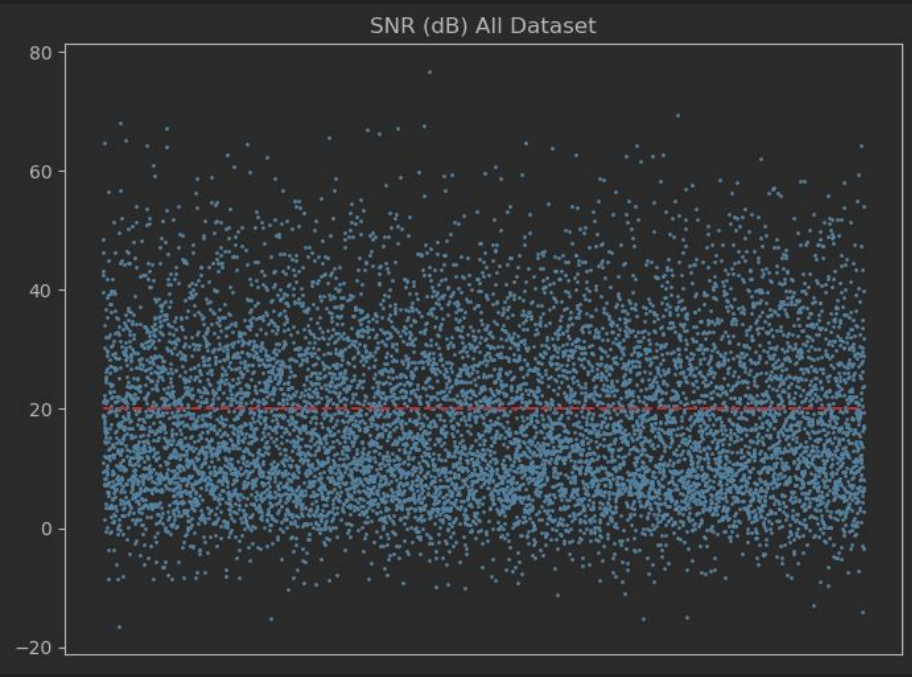
- Evaluating and Benchmarking AI Model on datasets and visualize results

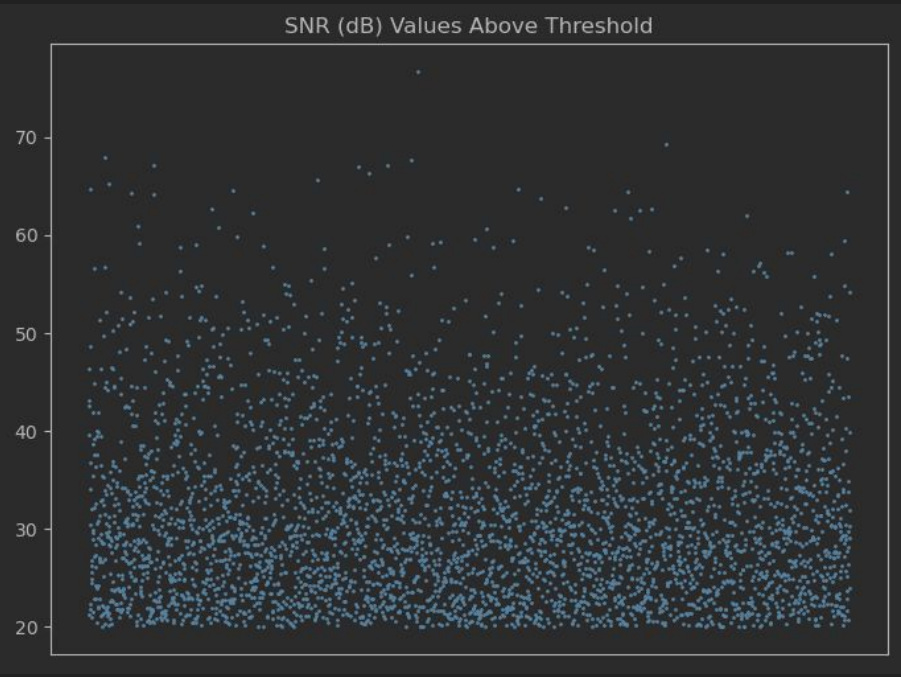
Extract High SNR Traces



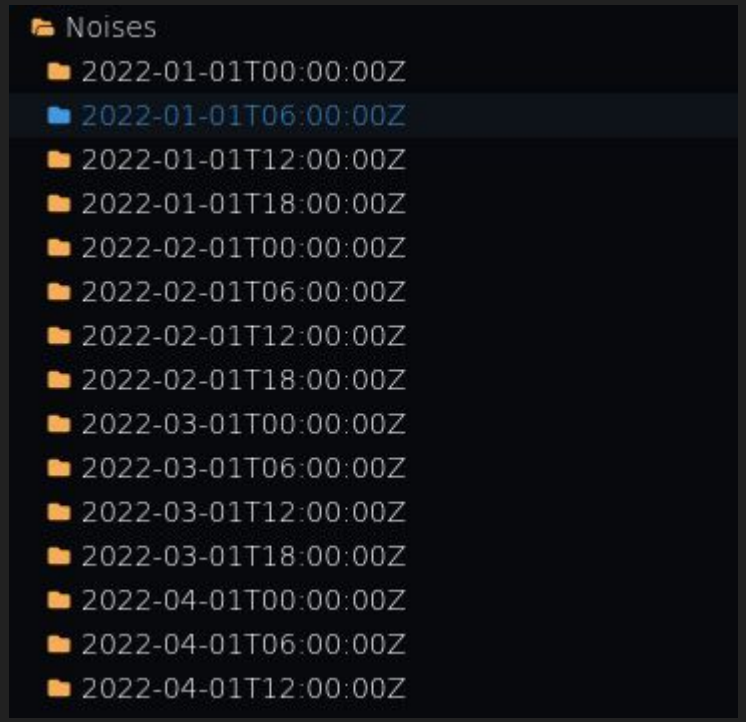


Extract High SNR Traces -



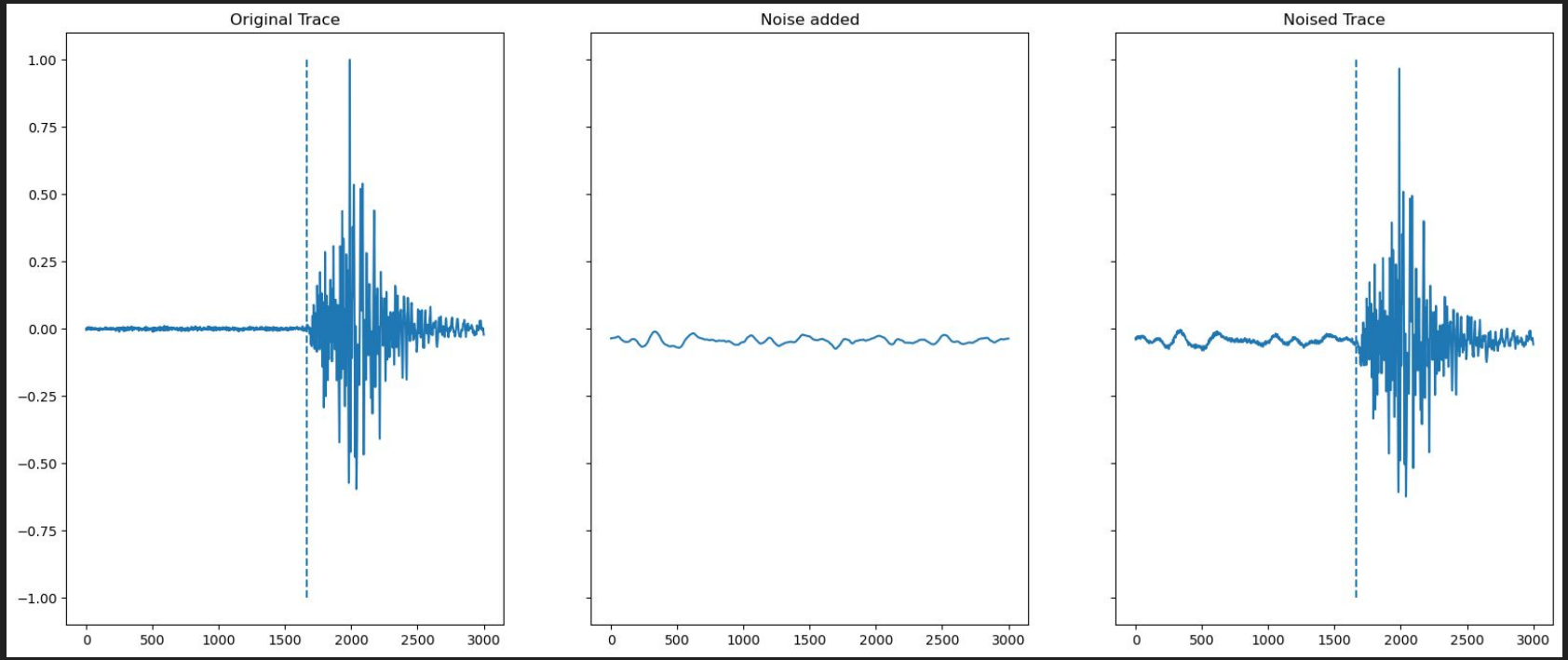


Noises Bank

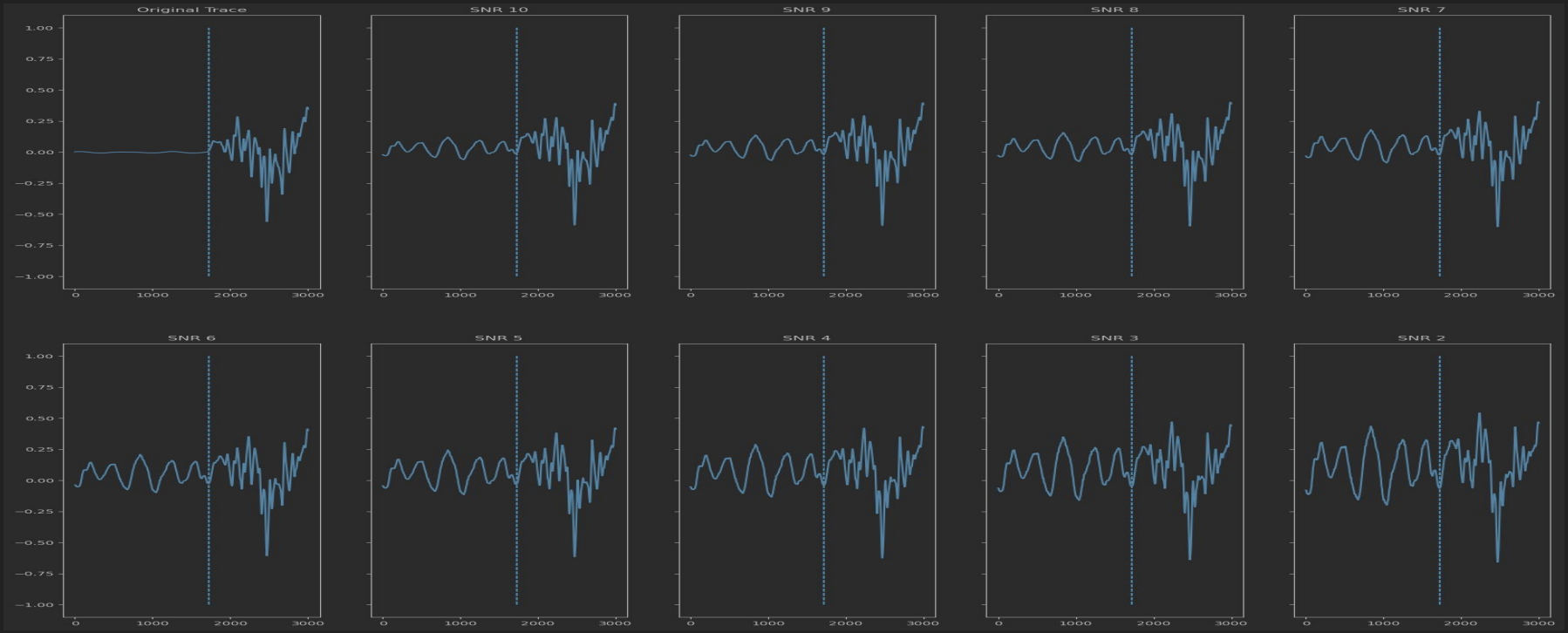

Noises
2022-01-01T00:00:00Z
2022-01-01T06:00:00Z
2022-01-01T12:00:00Z
2022-01-01T18:00:00Z
2022-02-01T00:00:00Z
2022-02-01T06:00:00Z
2022-02-01T12:00:00Z
2022-02-01T18:00:00Z
2022-03-01T00:00:00Z
2022-03-01T06:00:00Z
2022-03-01T12:00:00Z
2022-03-01T18:00:00Z
2022-04-01T00:00:00Z
2022-04-01T06:00:00Z
2022-04-01T12:00:00Z


2022-01-01T06:00:00Z
GE.BNDI..BHE.mseed
GE.BNDI..BHN.mseed
GE.BNDI..BHZ.mseed
GE.BOAB..BHE.mseed
GE.BOAB..BHN.mseed
GE.BOAB..BHZ.mseed
GE.DSB..BHE.mseed
GE.DSB..BHN.mseed
GE.DSB..BHZ.mseed
GE.FALKS..BHE.mseed
GE.FALKS..BHN.mseed
GE.FALKS..BHZ.mseed
GE.FLT1..BHE.mseed
GE.FLT1..BHN.mseed
GE.FLT1..BHZ.mseed
GE.GSI..BHE.mseed
GE.GSI..BHN.mseed
GE.GSI..BHZ.mseed

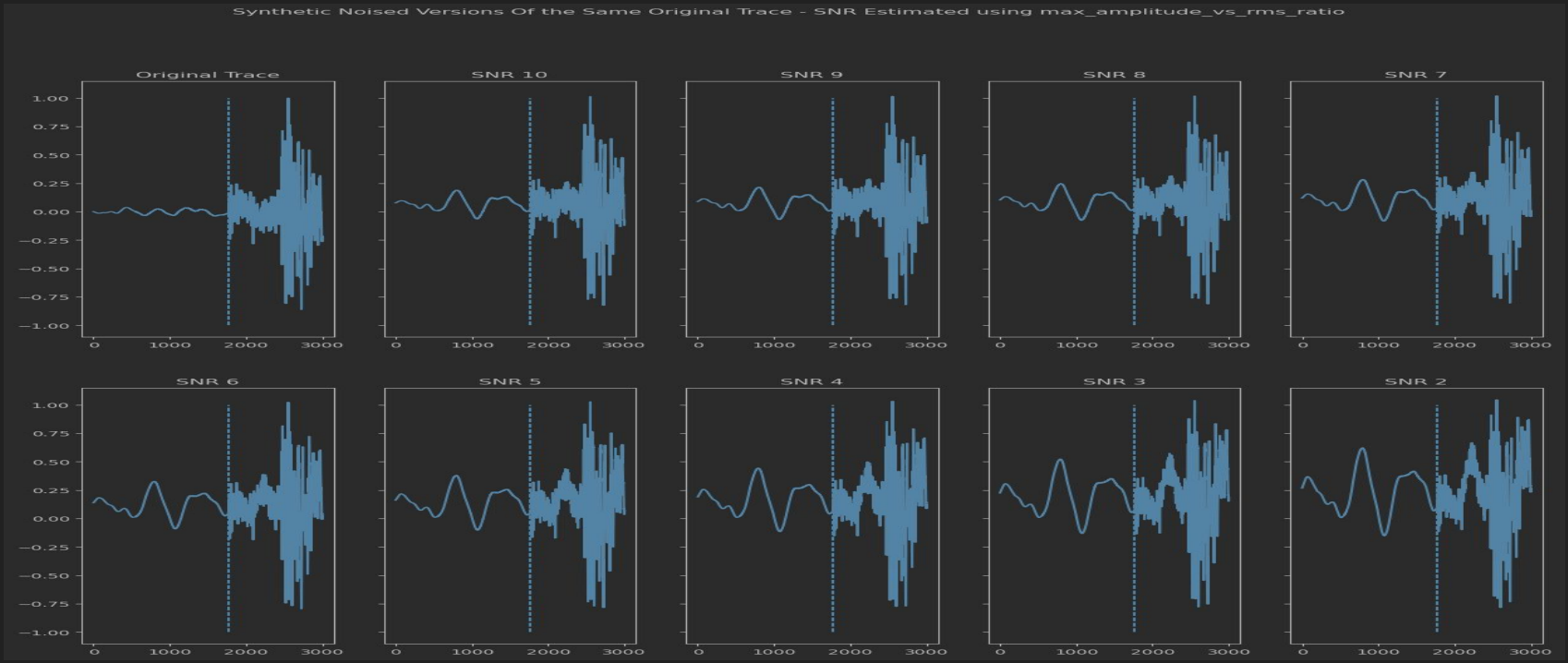
Generate a Synthetic Noisy Dataset



Synthetic Dataset Given SNR Definition- Energy Ratio



Synthetic Dataset Given SNR Definition - Max Amplitude vs. RMS Ratio



Model Evaluation

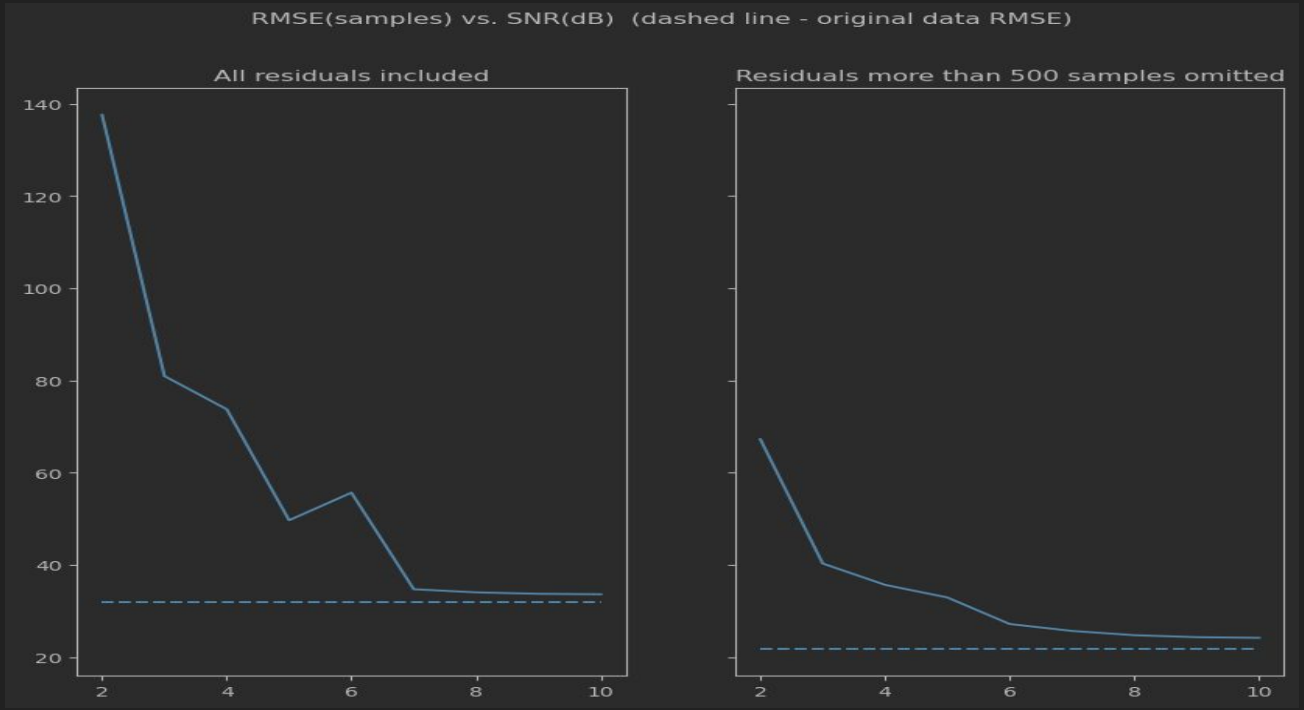
Evaluate PhaseNet\ EQTransformer

on synthetic dataset

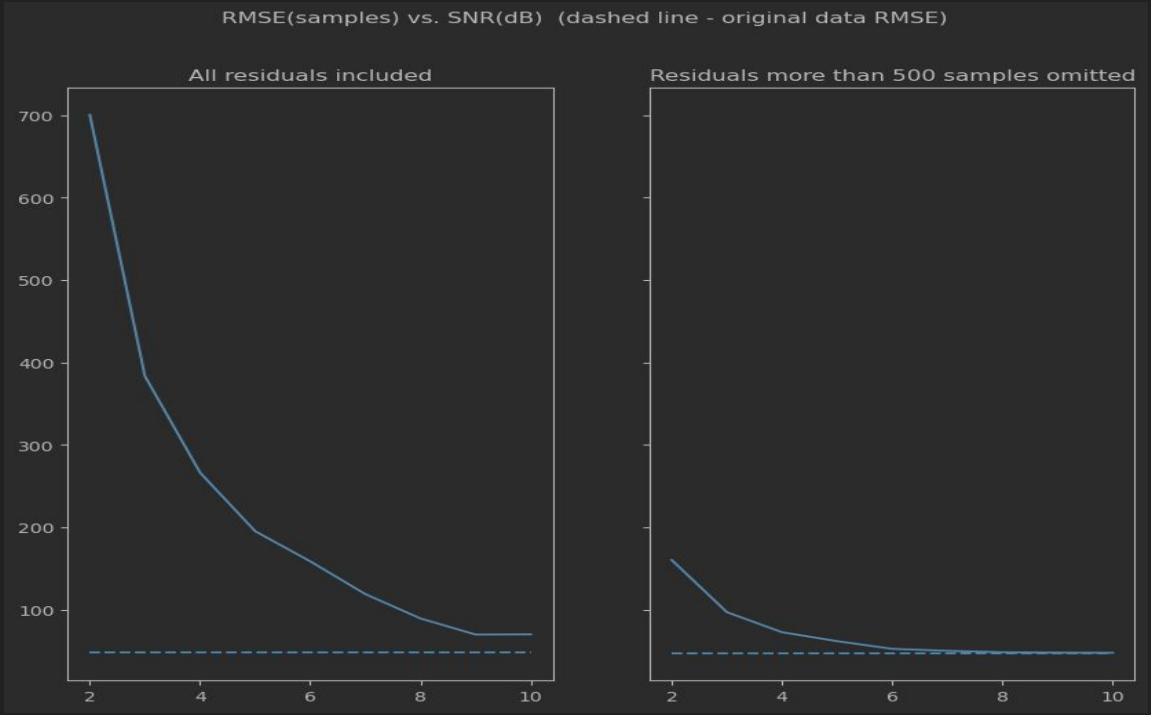
generated using

real traces from ETHZ\GEOFON

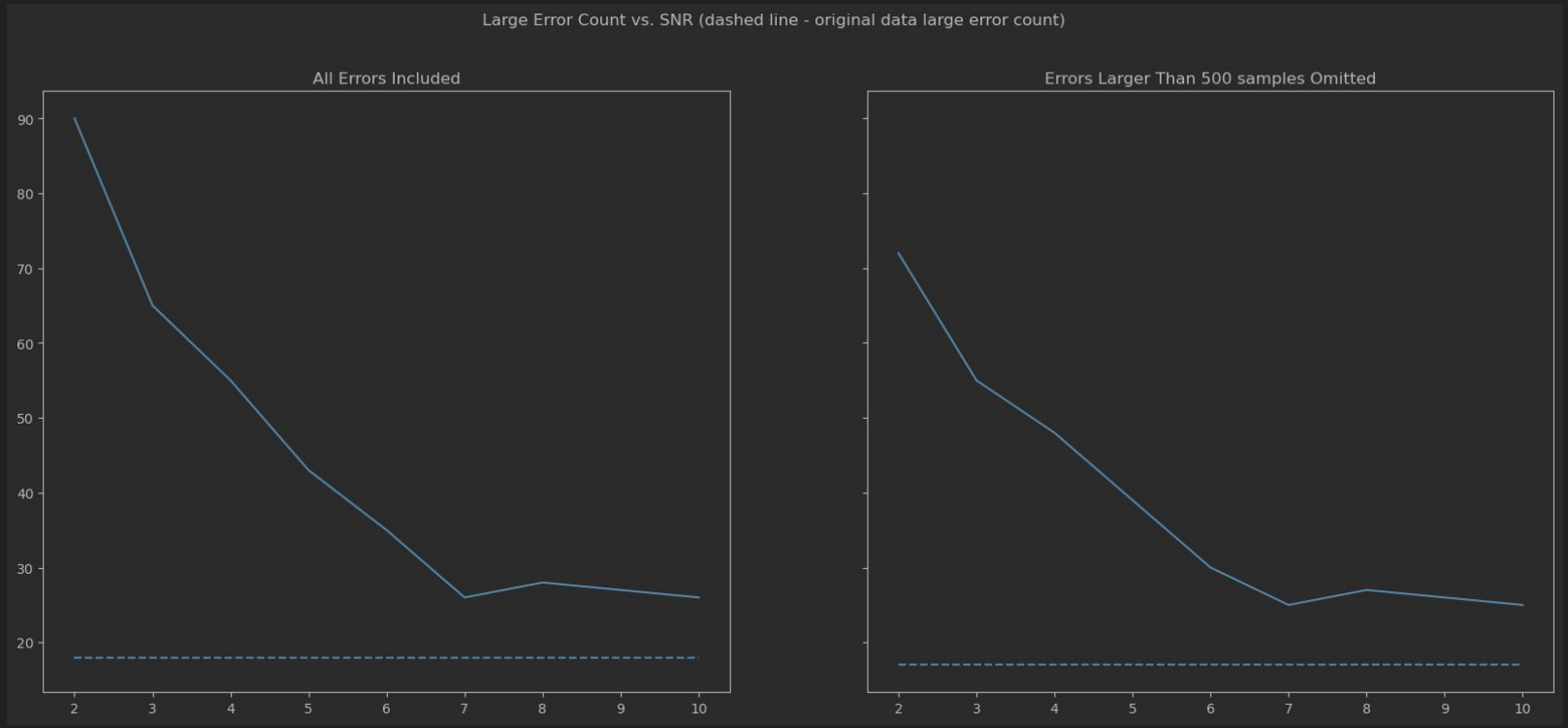
Root-Mean-Squared-Error - PhaseNet ETHZ - Energy Ratio



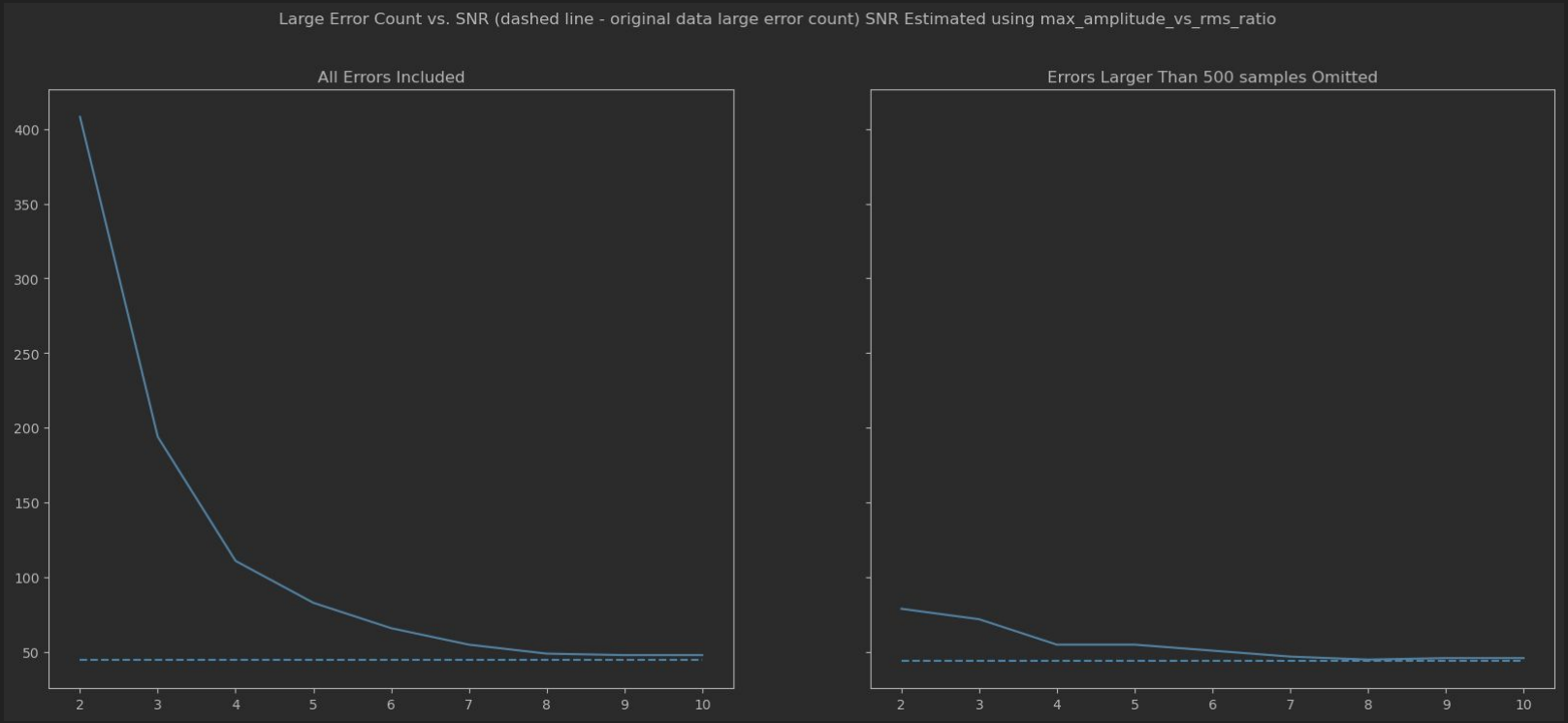
Root-Mean-Squared-Error - PhaseNet ETHZ- Max Amplitude vs. RMS Ratio



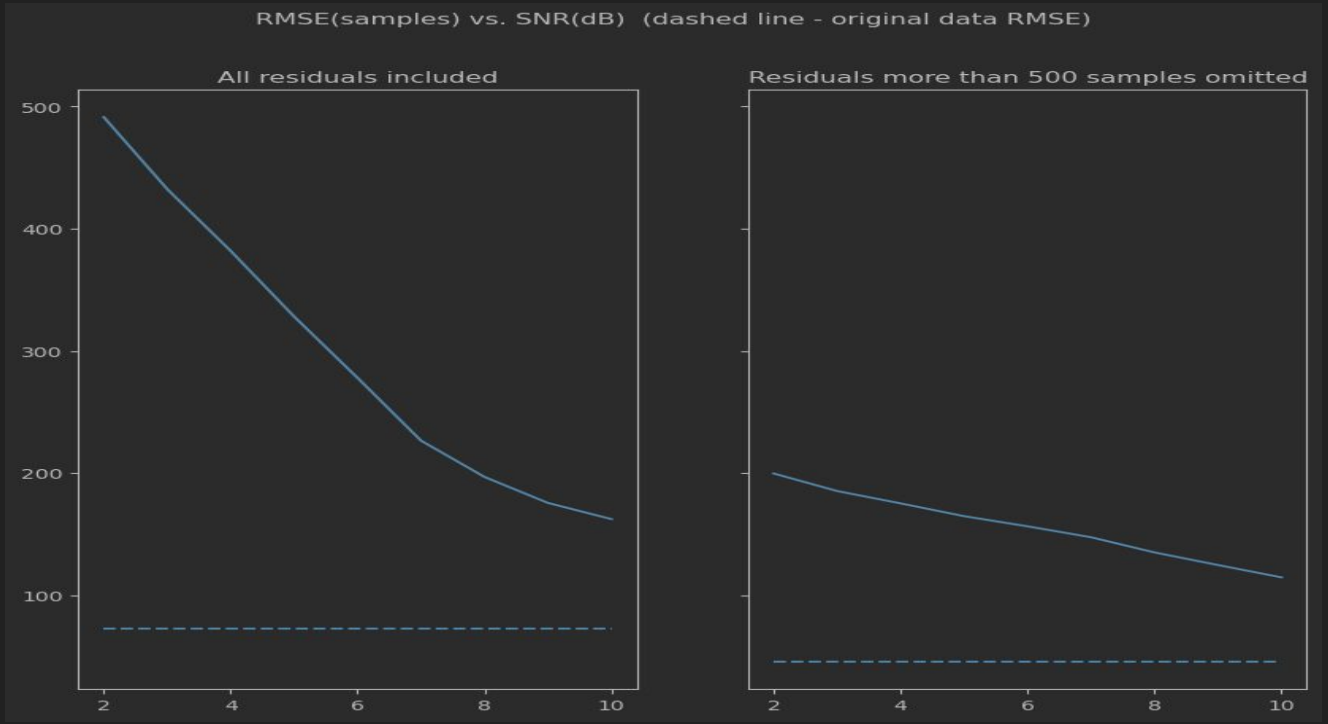
Large Error Count - PhaseNet ETHZ - Energy Ratio



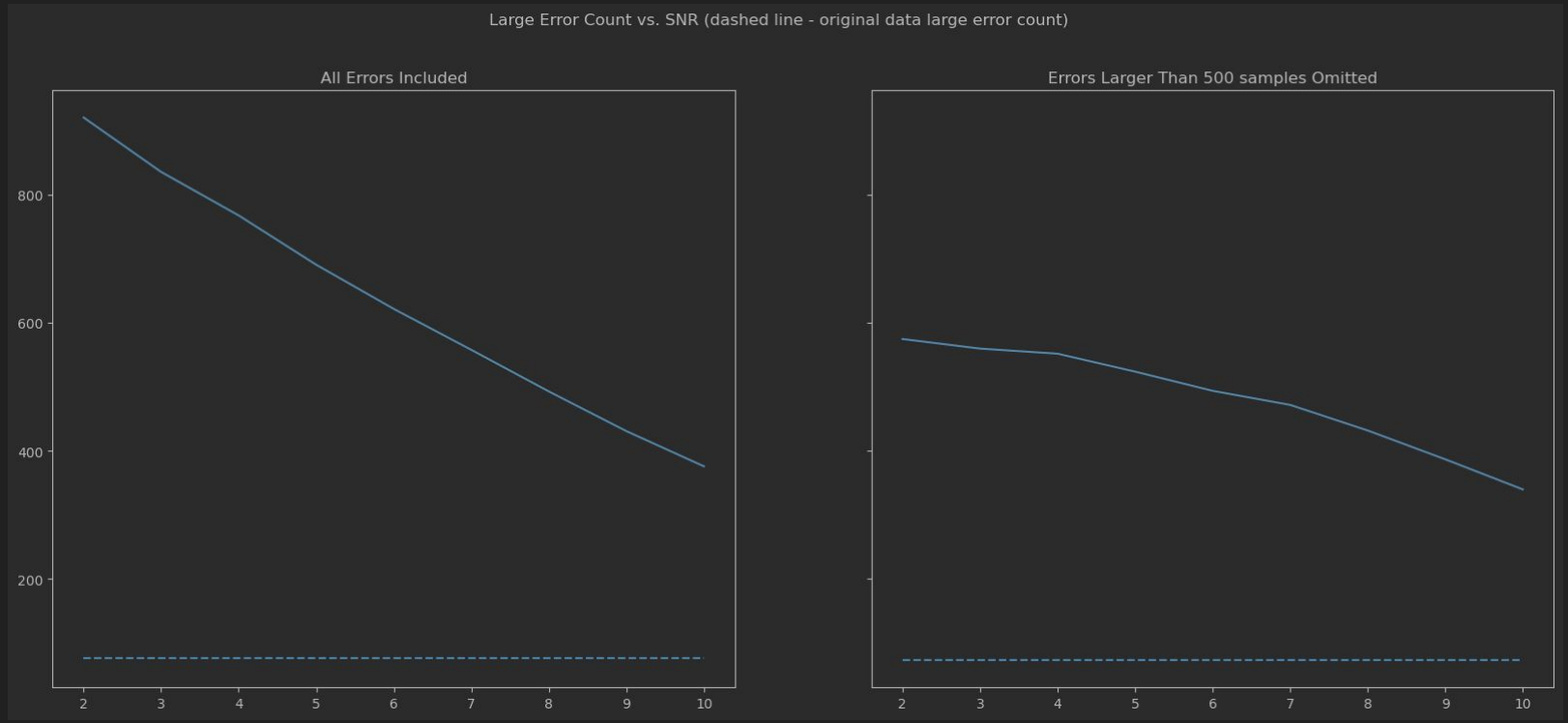
Large Error Count - PhaseNet ETHZ- Max Amplitude vs. RMS Ratio



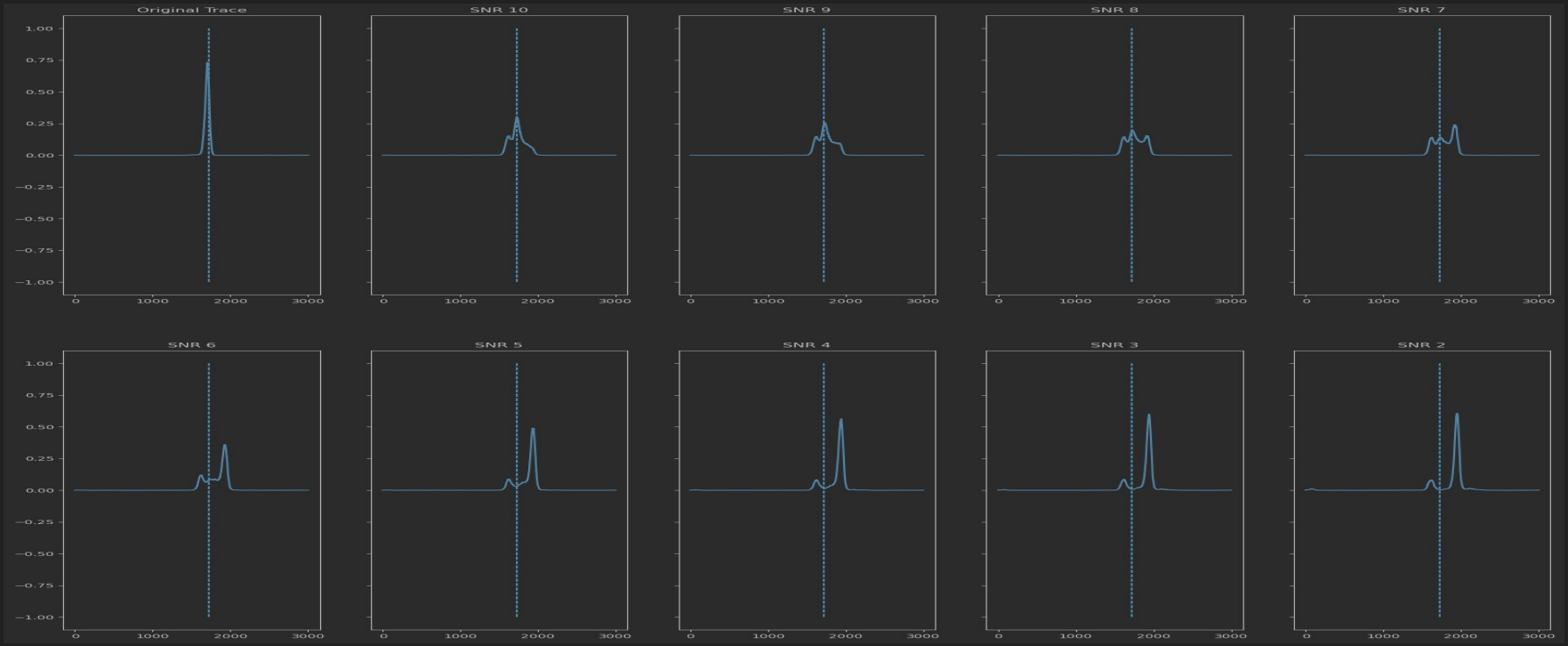
RMSE - PhaseNet GEOFON



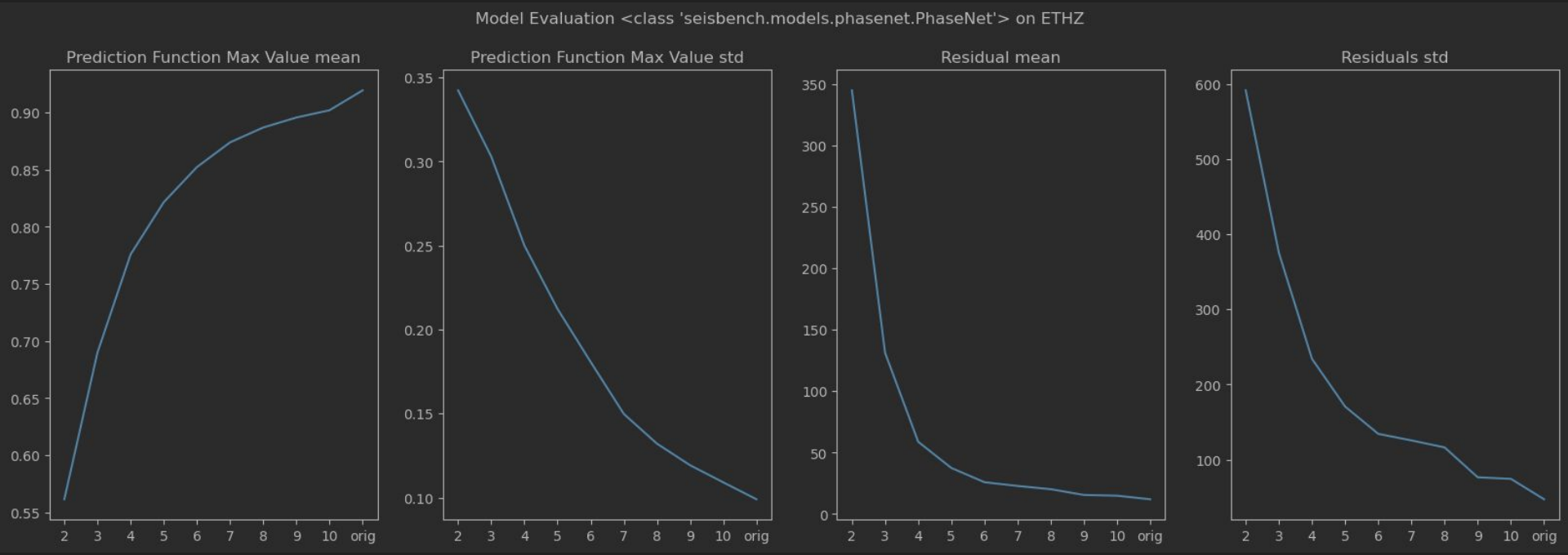
Large Error Count - PhaseNet GEOFON



The Effect Of SNR on Picker Prediction Function



Prediction Function Mean and Standard Deviation



Future Work

- Prediction Function features vs. SNR level - Maximum value, Similar Gaussian bias, variance
- Ablation Study - The effect of training pipelines variants
  - preprocessing
  - Loss function - Cross Entropy \ SoftArgmax\EBD\KL …
  - Label Smoothing Sigma - maybe try scheduling
- Use Bayesian
- Use synthetic dataset as data augmentation to improve model performance

I Really Enjoyed.

THANK YOU FOR LISTENING!

ANY MORE Q?