SnT 2021

CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

CTBTO

PREPARATORY COMMISSION

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi[1], Patrick Hammond[1], Ronald Brogan[2], Christopher Young[1], and Keith Koper[3]

P3.6-124

[1]Sandia National Laboratories

[2]ENSCO, Inc.

[3]Department of Geology and Geophysics, University of Utah

Sandia National Laboratories

NNSA National Nuclear Security Administration

POSTER

Acknowledgments: Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and 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-NA-0003525. The views expressed here do not necessarily reflect the views of the United States Government, the United States Department of Energy. This research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D). The authors acknowledge important interdisciplinary collaboration with scientists and engineers from LANL, LLNL, MSTS, PNNL, and SNL.

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

ABSTRACT

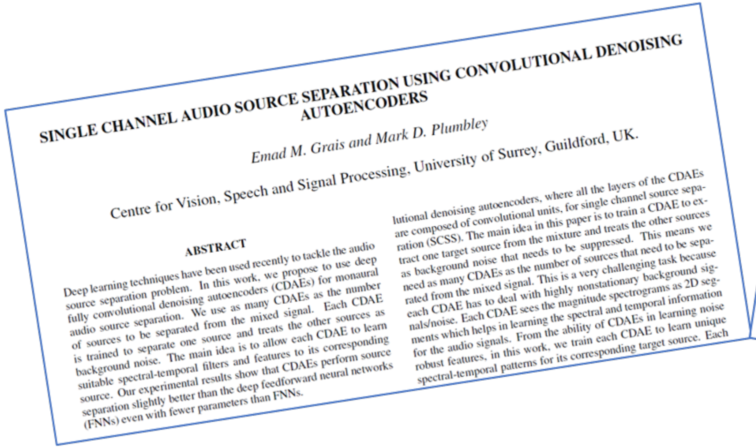
Seismic waveform data are generally contaminated by noise from various sources. To date, the most common noise suppression methods have been based on frequency filtering. These methods, however, are less effective when the signal of interest and noise share similar frequency bands. We implemented a seismic denoising method that uses a trained deep convolutional neural network (CNN) model. In our approach, the CNN provides a signal mask and a noise mask for an input signal. The Short-Time Fourier Transform (STFT) of the estimated signal is obtained by multiplying the signal mask with the STFT of the input signal. To build and test the denoiser, we used carefully compiled signal and noise datasets of seismograms recorded by the University of Utah Seismograph Stations network (United States). Results of test runs involving more than 9,000 constructed waveforms suggest that on average the denoiser improves the SNRs by ~5 dB and that most of the recovered signal waveforms have high similarity with respect to the target waveforms and suffer little distortion. Application to real data suggests that our denoiser achieves on average a factor of up to ~2–5 improvement in SNR over bandpass filtering and can suppress many types of noise that bandpass filtering cannot.

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

INTRODUCTION

- Frequency filtering, commonly used for noise suppression, is ineffective when signal and noise share the same frequency range.
- Frequency filtering is known to distort the signal, in some cases, making phase onsets and polarities difficult to determine.
- Deep learning denoising is widely used in the field of music information retrieval for music source separation (e.g., separation of singing voices from music accompaniment).


SINGLE CHANNEL AUDIO SOURCE SEPARATION USING CONVOLUTIONAL DENOISING AUTOENCODERS

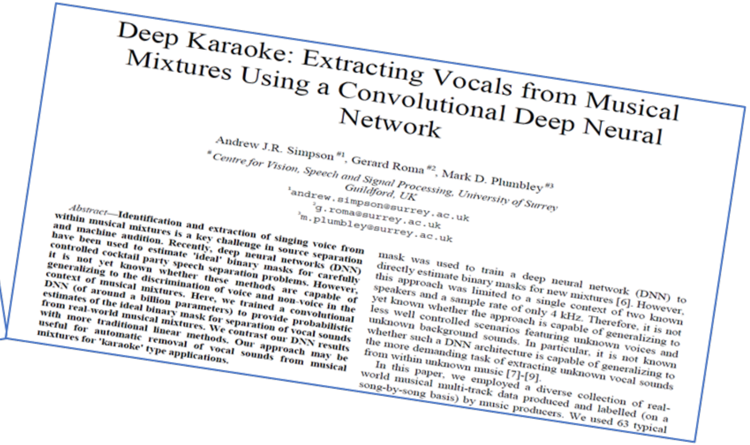
Emad M. Grais and Mark D. Plumbley

Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, UK.

ABSTRACT

Deep learning techniques have been used recently to tackle the audio source separation problem. In this work, we propose to use deep fully convolutional denoising autoencoders (CDAEs) for monaural audio source separation. We use as many CDAEs as the number of sources to be separated from the mixed signal. Each CDAE is trained to separate one source and treats the other sources as background noise. The main idea is to allow each CDAE to learn suitable spectral-temporal filters and features to its corresponding source. Our experimental results show that CDAEs perform source separation slightly better than the deep feedforward neural networks (FNNs) even with fewer parameters than FNNs.

lutional denoising autoencoders, where all the layers of the CDAEs are composed of convolutional units, for single channel source separation (SCSS). The main idea in this paper is to train a CDAE to extract one target source from the mixture and treats the other sources as background noise that needs to be suppressed. This means we need as many CDAEs as the number of sources that need to be separated from the mixed signal. This is a very challenging task because each CDAE has to deal with highly nonstationary background signals/noise. Each CDAE sees the magnitude spectrograms as 2D segments which helps in learning the spectral and temporal information for the audio signals. From the ability of CDAEs in learning noise robust features, in this work, we train each CDAE to learn unique spectral-temporal patterns for its corresponding target source. Each


Deep Karaoke: Extracting Vocals from Musical Mixtures Using a Convolutional Deep Neural Network

Andrew J.R. Simpson [#1], Gerard Roma [#2], Mark D. Plumbley [#3]
[#] *Centre for Vision, Speech and Signal Processing, University of Surrey*
Guildford, UK
[1] andrew.simpson@surrey.ac.uk
[2] g.roma@surrey.ac.uk
[3] m.plumbley@surrey.ac.uk

Abstract—**Identification and extraction of singing voice from within musical mixtures is a key challenge in source separation and machine audition. Recently, deep neural networks (DNN) have been used to estimate 'ideal' binary masks for carefully controlled cocktail party speech separation problems. However, it is not yet known whether these methods are capable of generalizing to the discrimination of voice and non-voice in the context of musical mixtures. Here, we trained a convolutional DNN (of around a billion parameters) to provide probabilistic estimates of the ideal binary mask for separation of vocal sounds from real-world musical mixtures. We contrast our DNN results with more traditional linear methods. Our approach may be useful for automatic removal of vocal sounds from musical mixtures for 'karaoke' type applications.**

mask was used to train a deep neural network (DNN) to directly estimate binary masks for new mixtures [6]. However, this approach was limited to a single context of two known speakers and a sample rate of only 4 kHz. Therefore, it is not yet known whether the approach is capable of generalizing to less well controlled scenarios featuring unknown voices and unknown background sounds. In particular, it is not known whether such a DNN architecture is capable of generalizing to the more demanding task of extracting unknown vocal sounds from within unknown music [7]-[9].

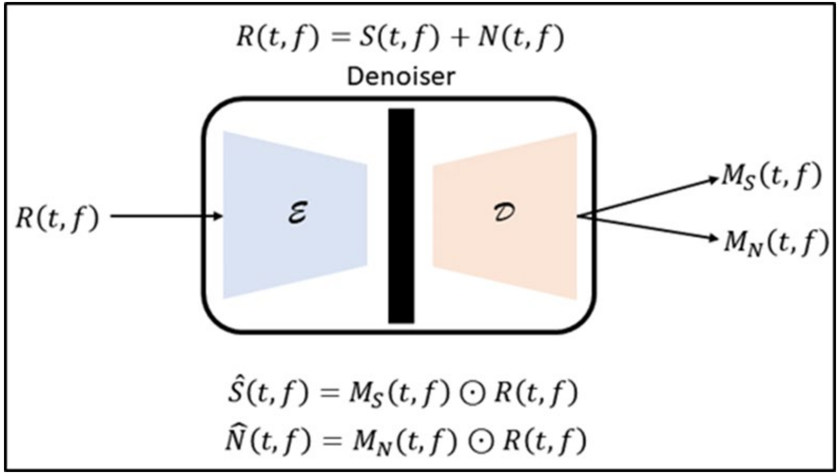
In this paper, we employed a diverse collection of real-world musical multi-track data produced and labelled (on a song-by-song basis) by music producers. We used 63 typical

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

CTBTO PREPARATORY COMMISSION

METHODS



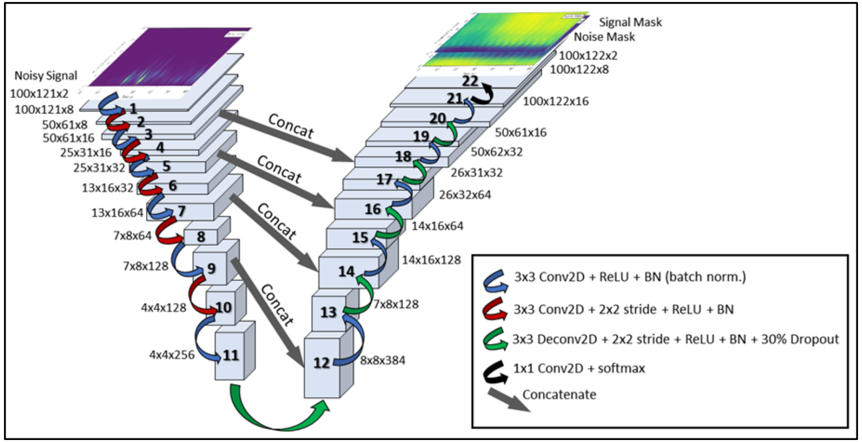
- The network consists of an encoder and a decoder.
- For an input $R(t,f)$, the network provides a signal mask ($M_S(t,f)$) and a noise mask ($M_N(t,f)$).
- The estimated 'clean' signal ($\hat{S}(t,f)$) is obtained by multiplying $M_S(t,f)$ with $R(t,f)$; and the estimated noise ($\hat{N}(t,f)$) is obtained by multiplying $M_N(t,f)$ with $R(t,f)$.

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

CTBTO PREPARATORY COMMISSION

METHODS



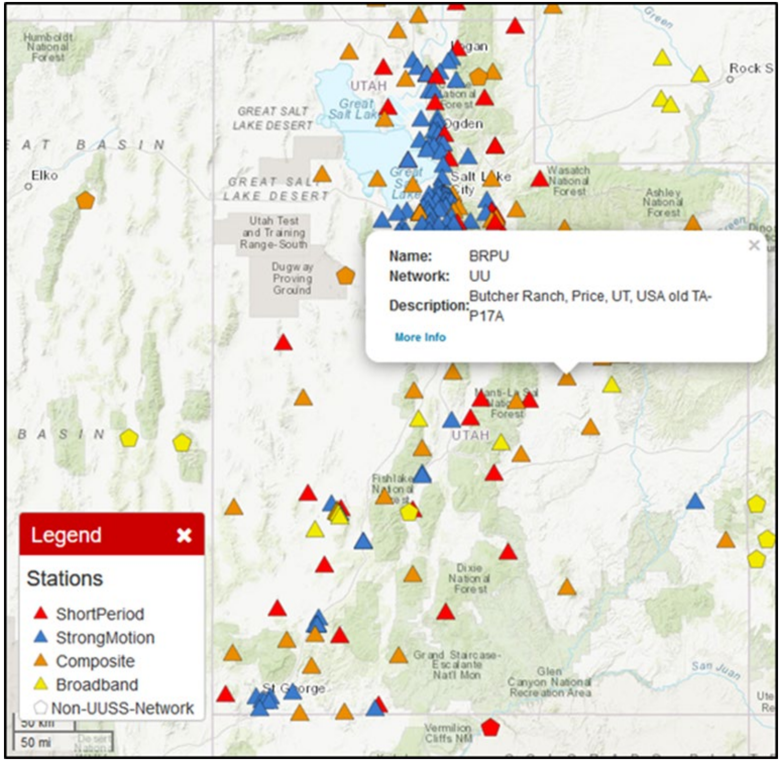
- The network consists of 20 hidden layers
- Half of the layers make up the encoder, and the other half the decoder
- Implemented using Keras on top of TensorFlow
- ~2.4 million trainable parameters
- ~3K non-trainable parameters

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

DATA

- The ‘clean’ signal dataset consists of 3,188 high-SNR Z-component recorded at BRPU from local and near-regional earthquakes.
- The noise dataset contains 15,426 waveforms from various noise sources and various stations.
- The 2 datasets randomly divided into training, validation, and test sets using the ‘70-15-15’ convention.
- Noisy waveforms constructed by summing each ‘clean’ signal waveform and a randomly selected noise waveform. This was repeated 20 times for each set, resulting in:
  - 44,620 waveforms for the training set,
  - 9,580 waveforms for the validation set, and
  - 9,560 waveforms for the test set



SnT 2021
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

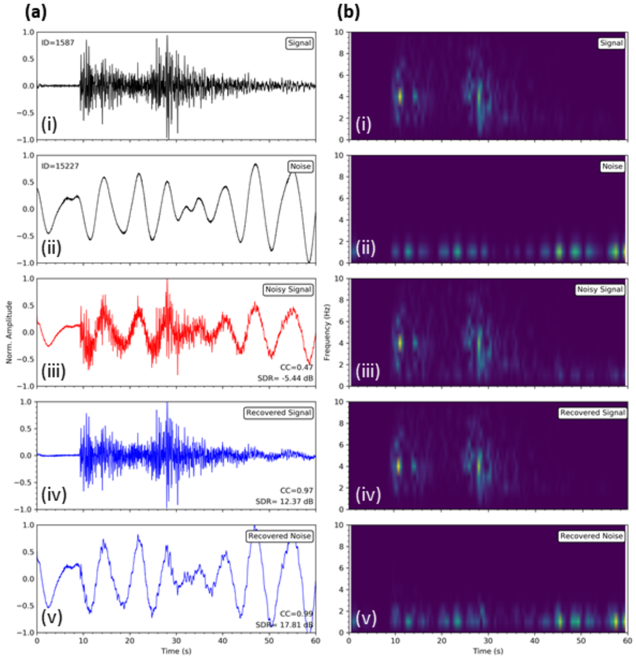
Poster No.: P3.6-124

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

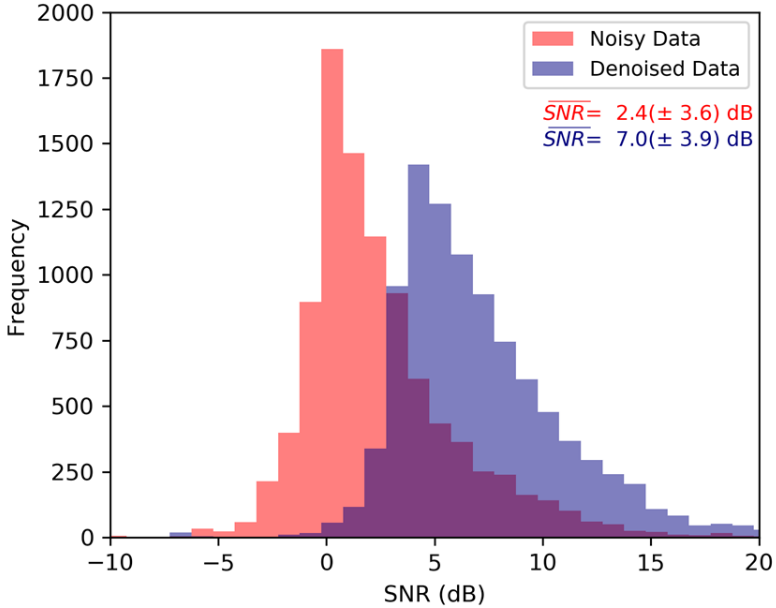
Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

CTBTO PREPARATORY COMMISSION

MODEL EVALUATION



- Recovered waveforms are very similar to the corresponding GTs (CC of 0.97–0.99).
- Recovered seismograms show little distortion with respect to the GTs (SDR of 12.37–17.81 dB).

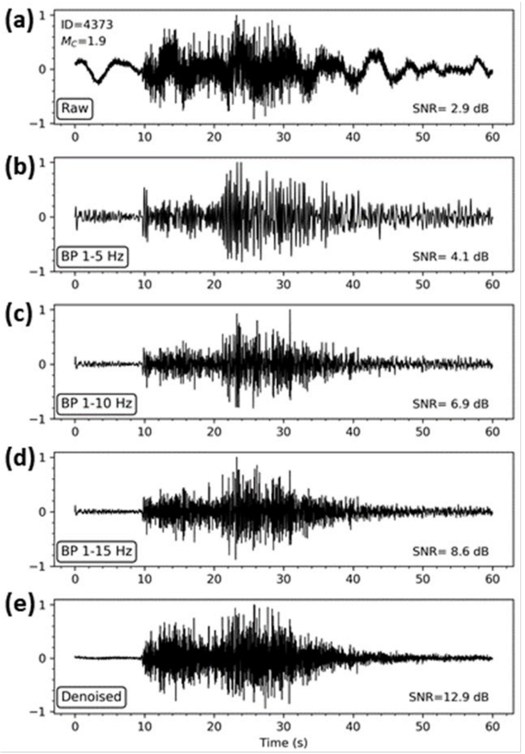


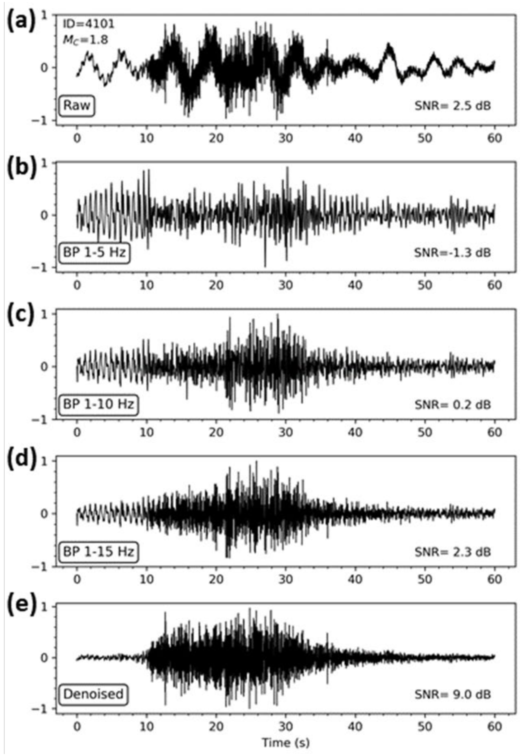
- 9,560 constructed noisy waveforms of the test set
- The denoiser achieves an average improvement in SNR of ~5 dB.

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

DENOISING vs. FREQUENCY FILTERING





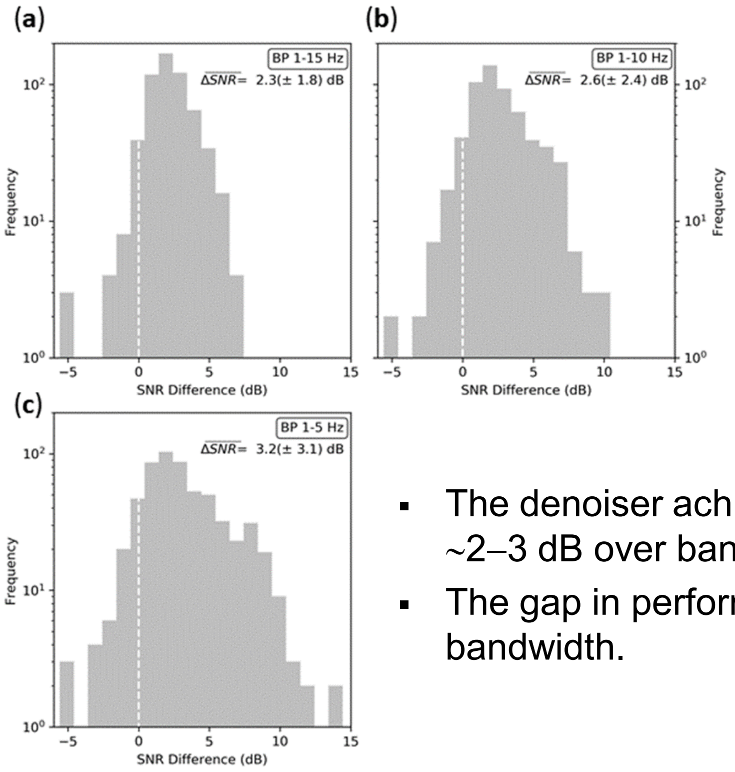
- Compared with both the raw and filtered data, the denoised seismograms show reduced pre-*P* noise and enhanced *P* amplitudes,
- Resulting in improved SNRs.

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

DENOISING vs. FREQUENCY FILTERING

Station: BRPU



- The denoiser achieves an average improvement in SNR of ~2–3 dB over bandpass filter.
- The gap in performance increases with decreasing filter bandwidth.

SnT 2021
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

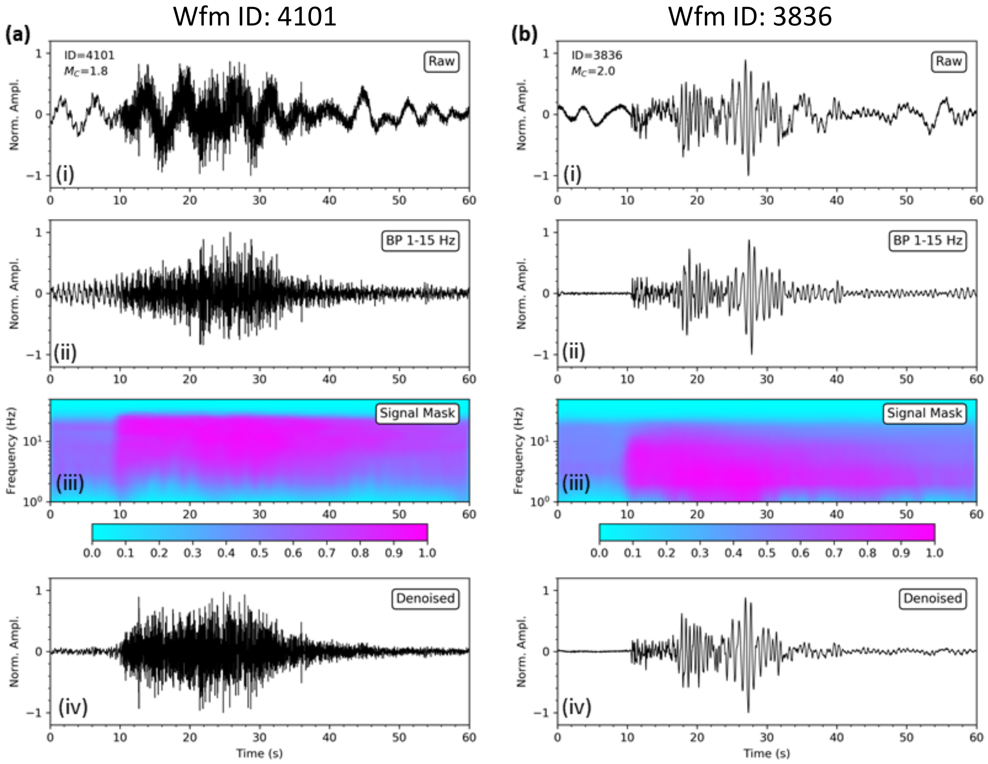
Poster No.: P3.6-124

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

CTBTO PREPARATORY COMMISSION

DENOISING vs. FREQUENCY FILTERING



- The values of the elements of the mask operator vary with both time & frequency in the range of 0–1.
- The operator for a bandpass filter would appear as a streak of 1’s within the passband.
- The mask operator adapts to the changing characteristics of the input signal.
- This adaptability is undoubtedly the reason for the observed edge of the denoiser over bandpass filtering.

SnT 2021

CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

Poster No.: P3.6-124

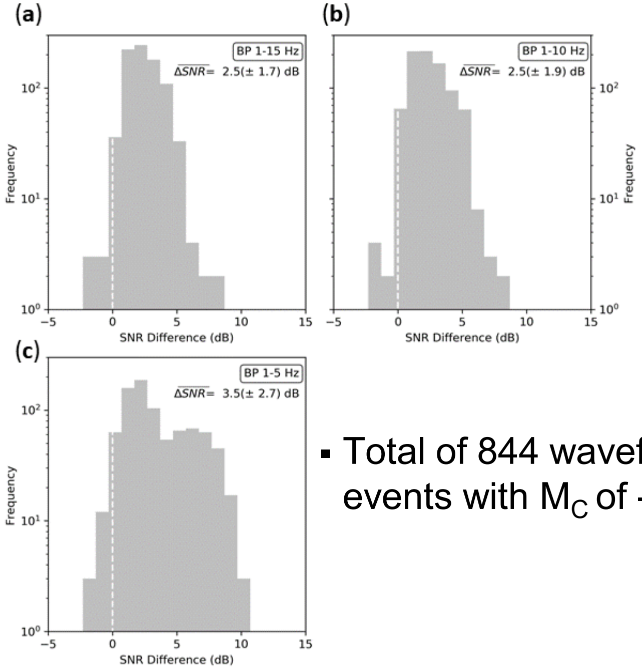
Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

CTBTO PREPARATORY COMMISSION

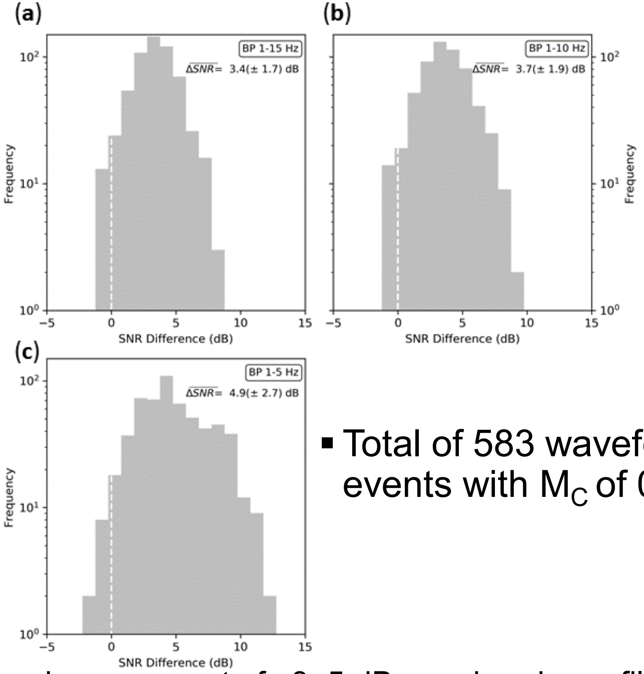
MODEL TRANSPORTABILITY

Station: ZNPU



- Total of 844 waveforms from events with M_C of -0.84–5.3

Station: SPU



- Total of 583 waveforms from events with M_C of 0.1–4.5

- For the stations ZNPU and SPU, denoising achieved an average improvement of ~3–5 dB over bandpass filtering.
- The CNN model is transportable to other stations in Utah, and possibly also in neighboring regions.

Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah

Rigobert Tibi, Patrick Hammond, Ronald Brogan, Christopher Young, and Keith Koper

CONCLUSIONS

- We implemented a seismic denoising method that uses a trained deep CNN model to decompose an input waveform into a signal of interest and noise.
- Test results based on more than 9,000 constructed waveform data suggest that most of the waveforms recovered by the trained deep convolutional network show high degree of fidelity to their respective GTs, in terms of both waveform similarity and amplitudes.
- Processing of real seismograms suggests that the denoiser achieves an average improvement in SNR of ~5 dB and ~2–5 dB over the raw and bandpass filtered data, respectively.
- The CNN model also works well for UUSS stations not involved in model training, suggesting that it is transportable around Utah, and possibly also to neighboring regions with similar wave propagation characteristics and background noise.