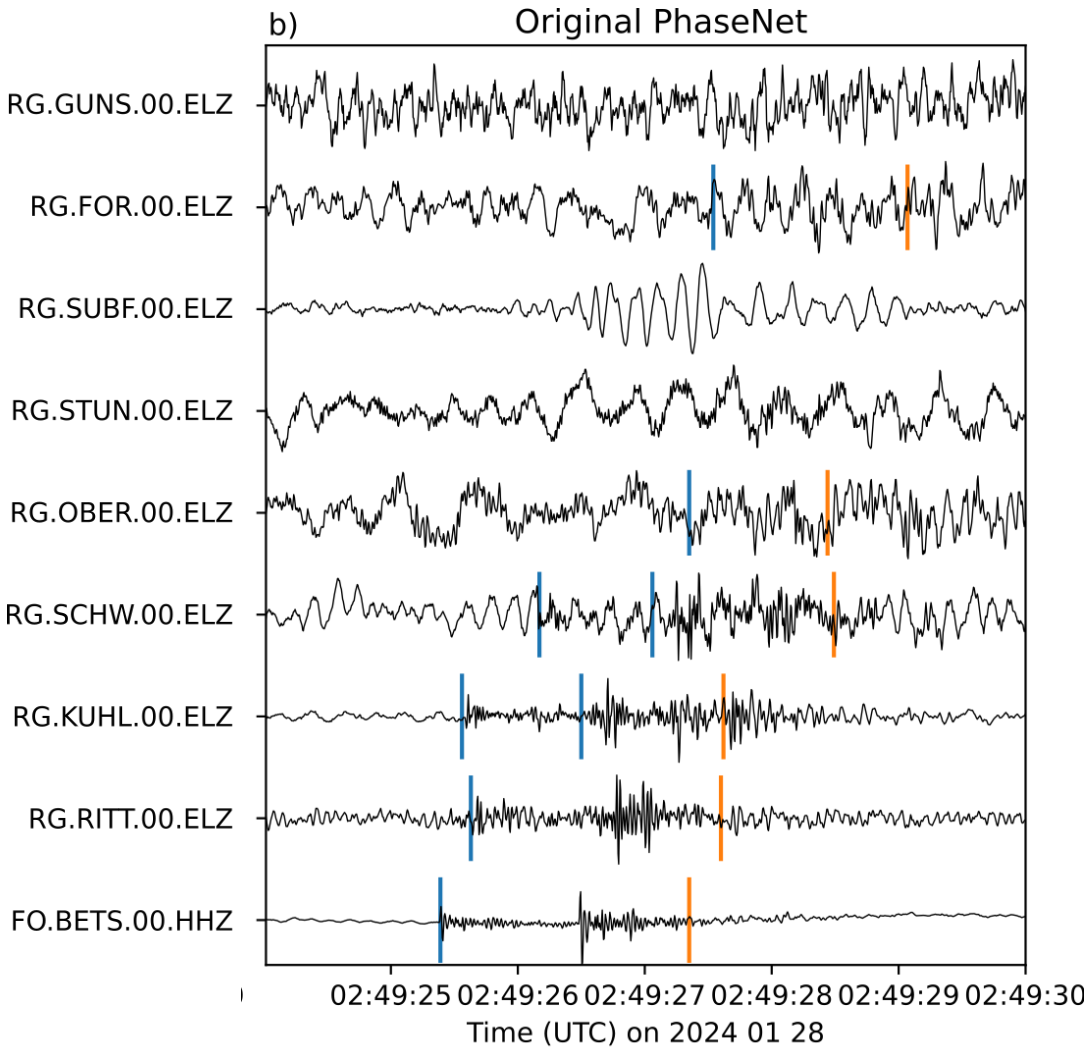
Seismic Phase Picking for Induced Seismicity with Deep Learning

Janis Heuel, V.Maurer, M. Weiel, D. Coquelin, M. Frietsch,
T. Sharia, M. Götz, and A. Rietbrock

4th Induced Seismicity Workshop | Schatzalp | Davos

AIS: AI-based monitoring of induced seismicity

Motivation



Original PhaseNet model: Zhu & Beroza, GJI, 2018

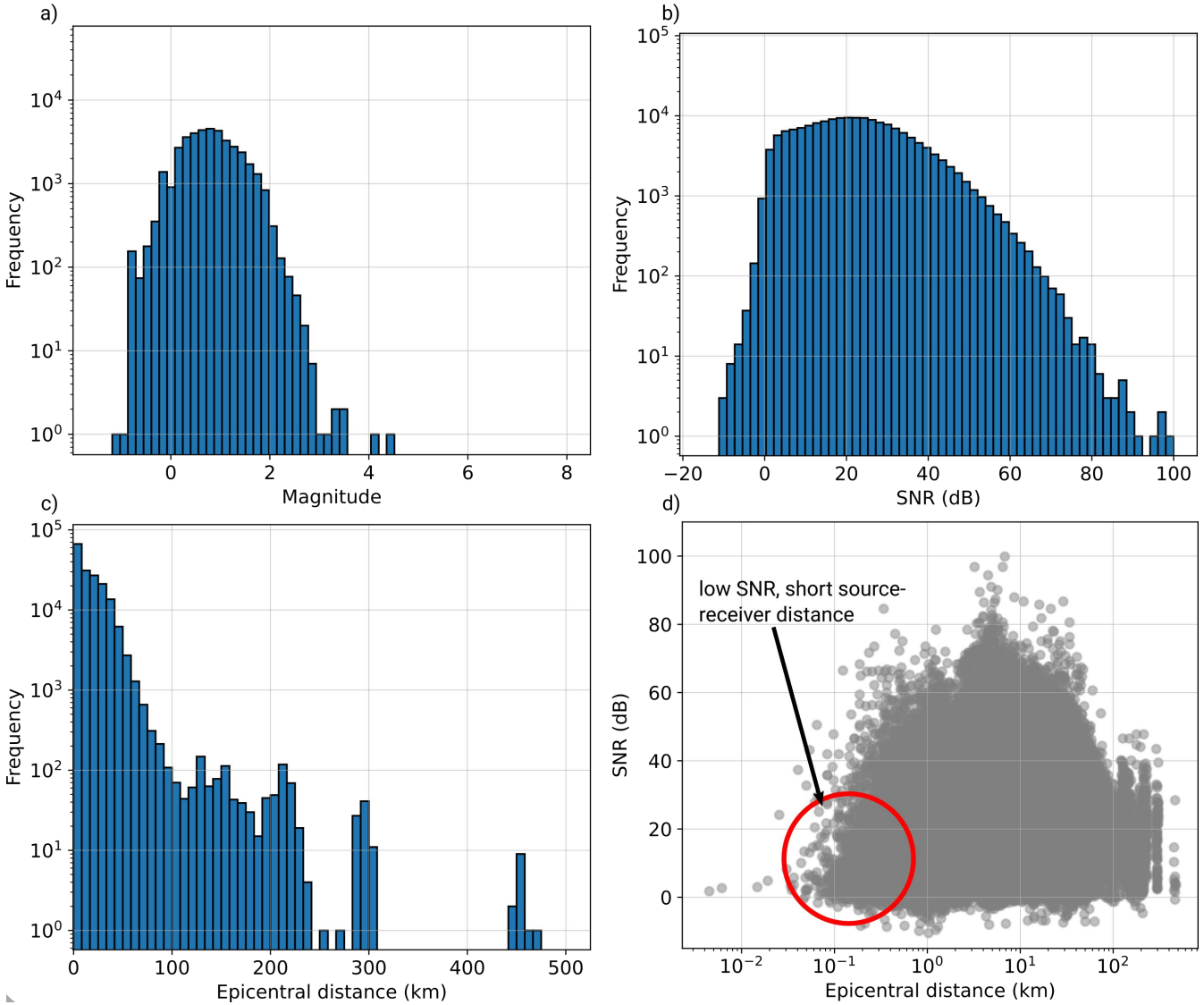
Problems

- Data sets for **AI** in seismology mostly contain **earthquakes** recorded at **local and regional distances**
- When applying the **original published models** for deep-learning pickers to different geographical region, they do not perform well (i.e. **missing events**)
- Published deep-learning picking models **fail** when applied to data sets **with low SNR and/or small epicentral distances**
- Existing PhaseNet models have not been trained with multiple events within a window

Objectives

- Training PhaseNet with induced seismicity and low SNR events
- Apply trained PhaseNet model at a geothermal site in Rittershoffen (France)

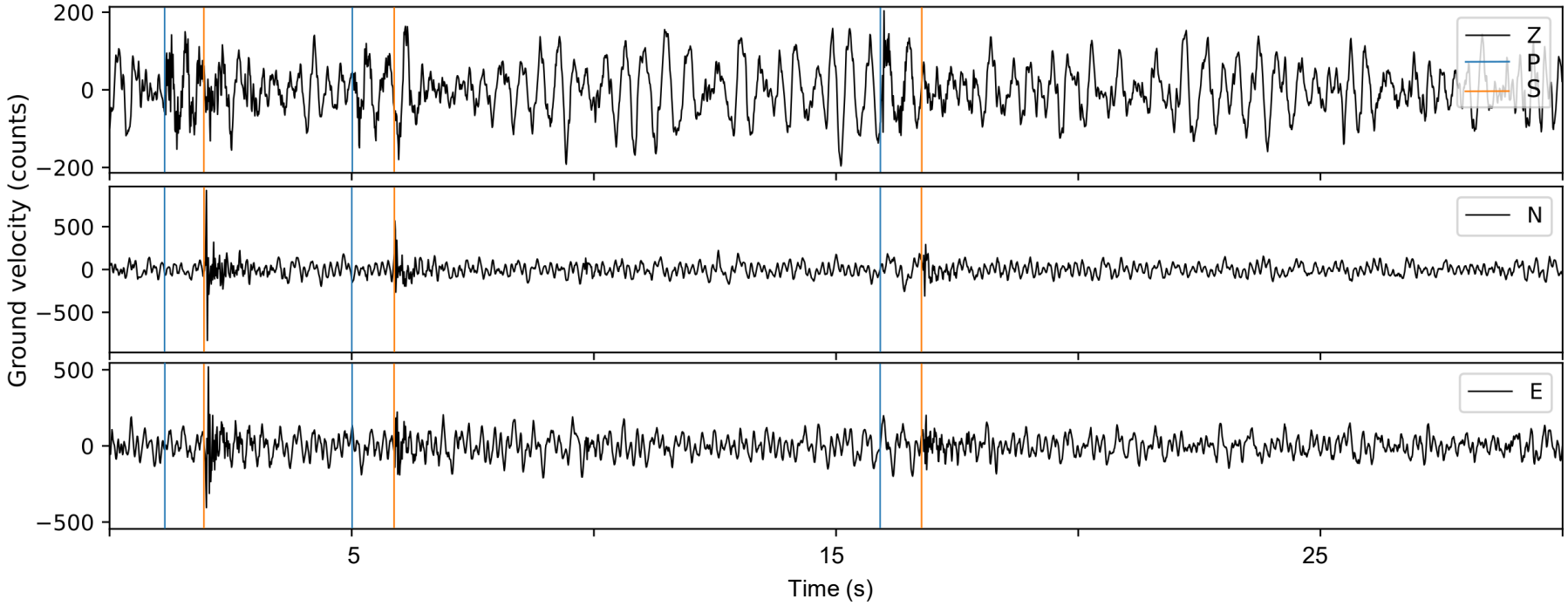
A new dataset for AI in Seismology (induced seismicity)



- low magnitude events
- short distances between source and receiver (e.g. shallow earthquakes)
- low signal-to-noise ratios
- 170,000 three-component waveforms with P- and S-arrival in each window
- 40,228 different seismic events
- 455 seismological stations

- seismic events are from
  - geothermal sites
  - wastewater disposal
  - coal mine flooding
  - low-magnitude events from SED

Training of PhaseNet (with multiple events)

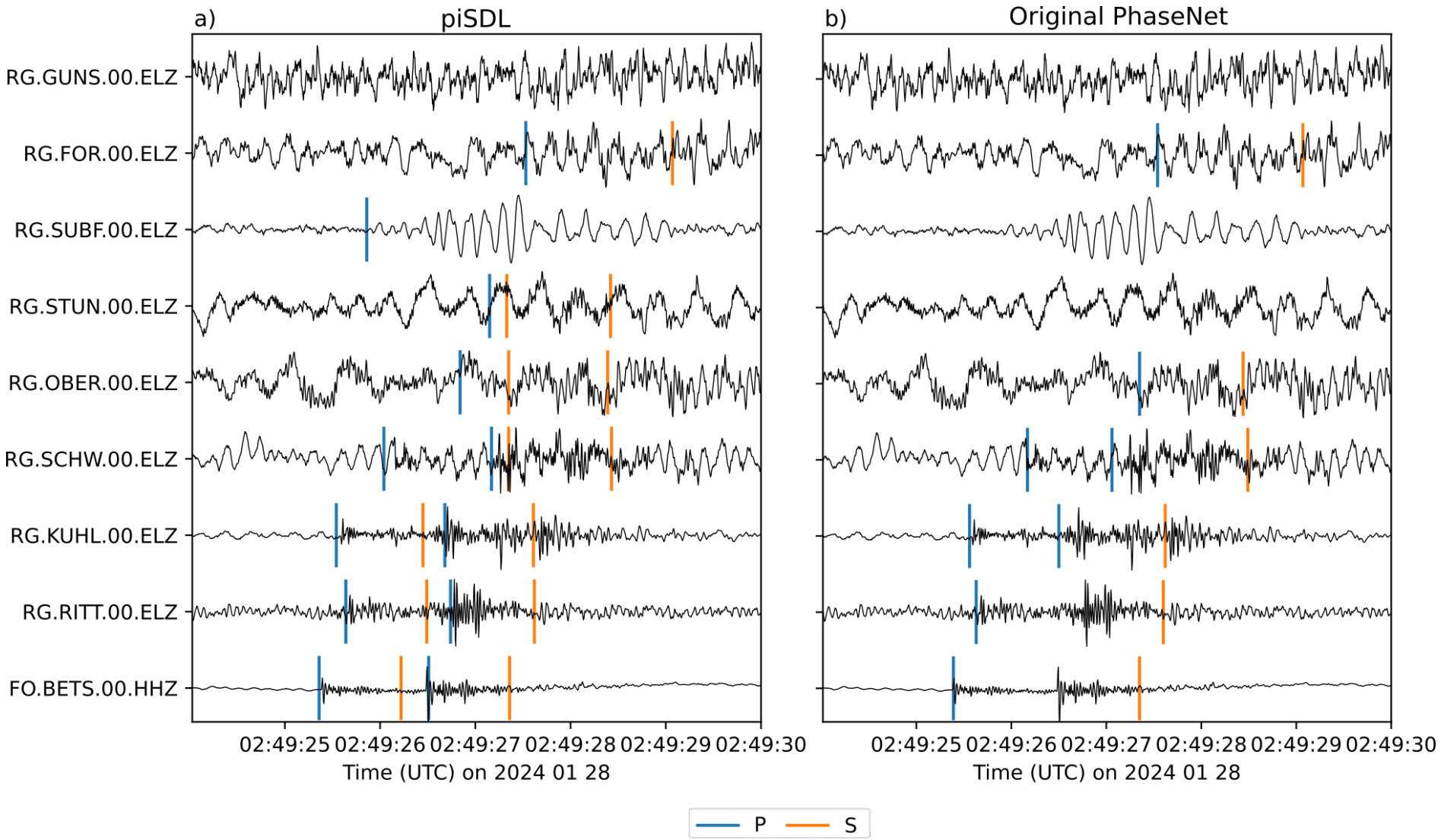


- Split data sets:
  - 70% for training
  - 20% for validation
  - 10% for testing
- Data augmentation
  - Rotating horizontal components
  - Adding noise to earthquake waveforms
  - Adding noise samples to training data set
- PhaseNet and augmentation techniques:

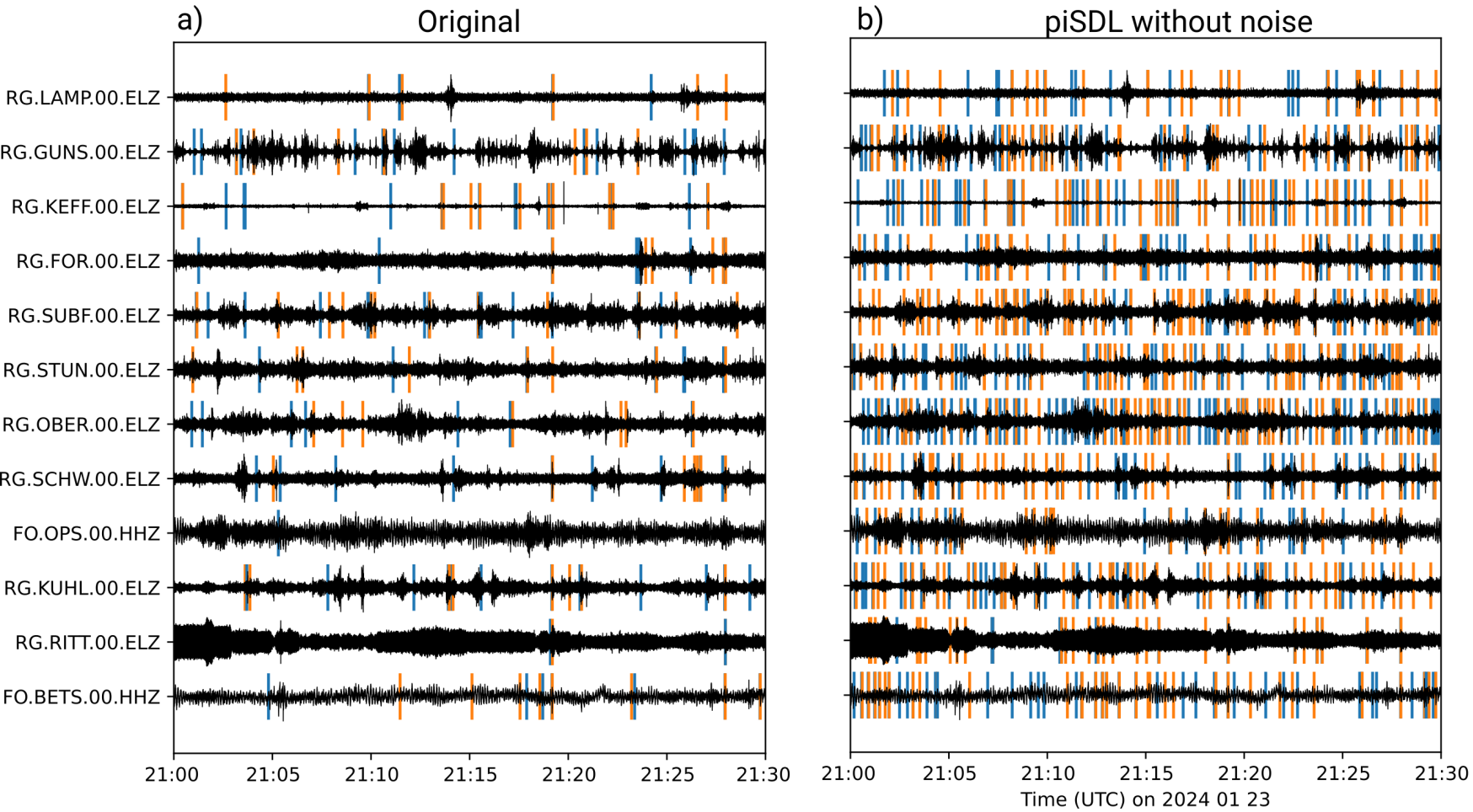
SeisBench
A toolbox for machine learning in seismology

Woollam et al., SSA, 2022

Testing new PhaseNet model (piSDL)



Testing new PhaseNet model (piSDL)

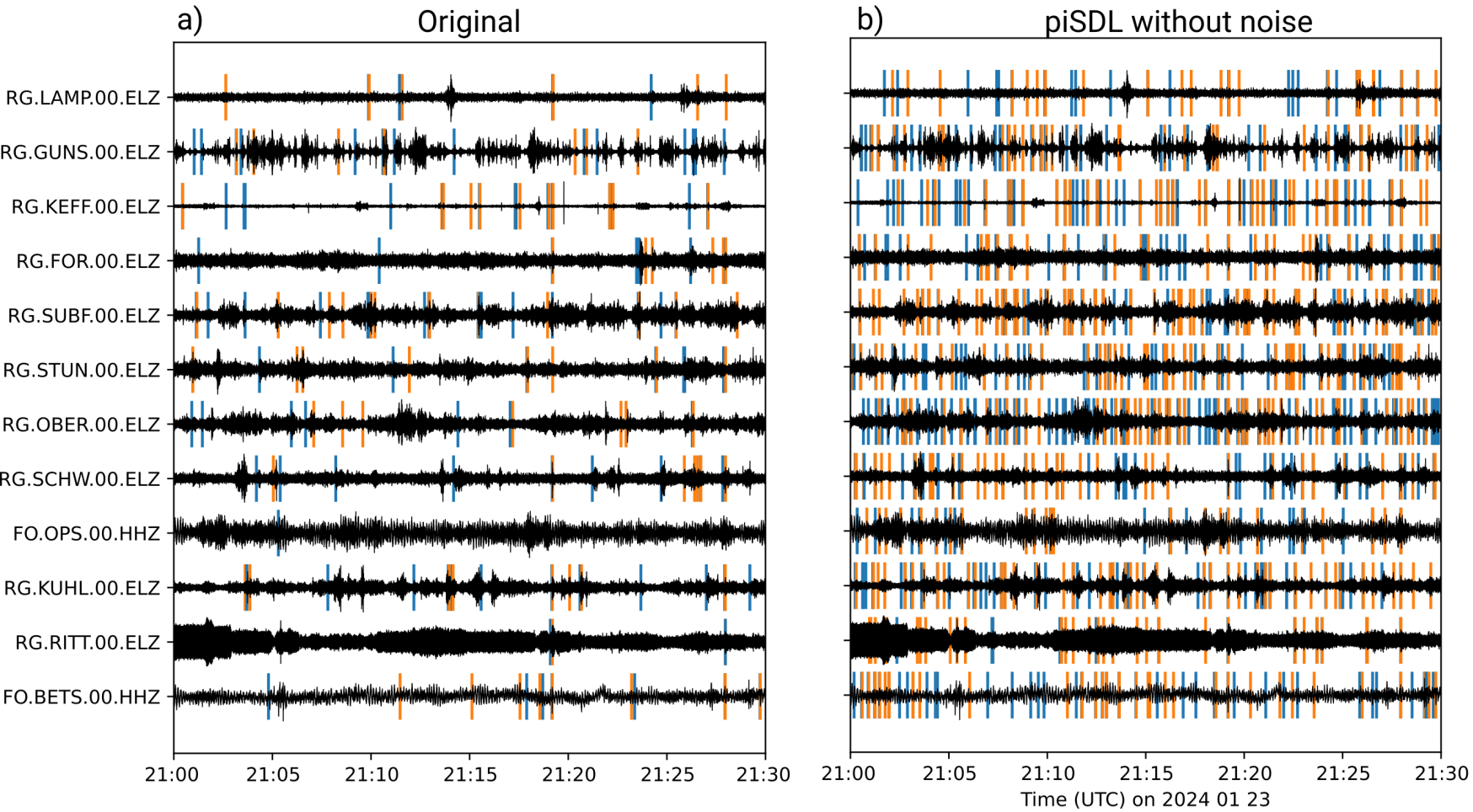


118 P-picks
97 S-picks

524 P-picks
478 S-picks

Too many false picks after training of PhaseNet with data set for induced seismicity

Testing new PhaseNet model (piSDL)



118 P-picks
97 S-picks

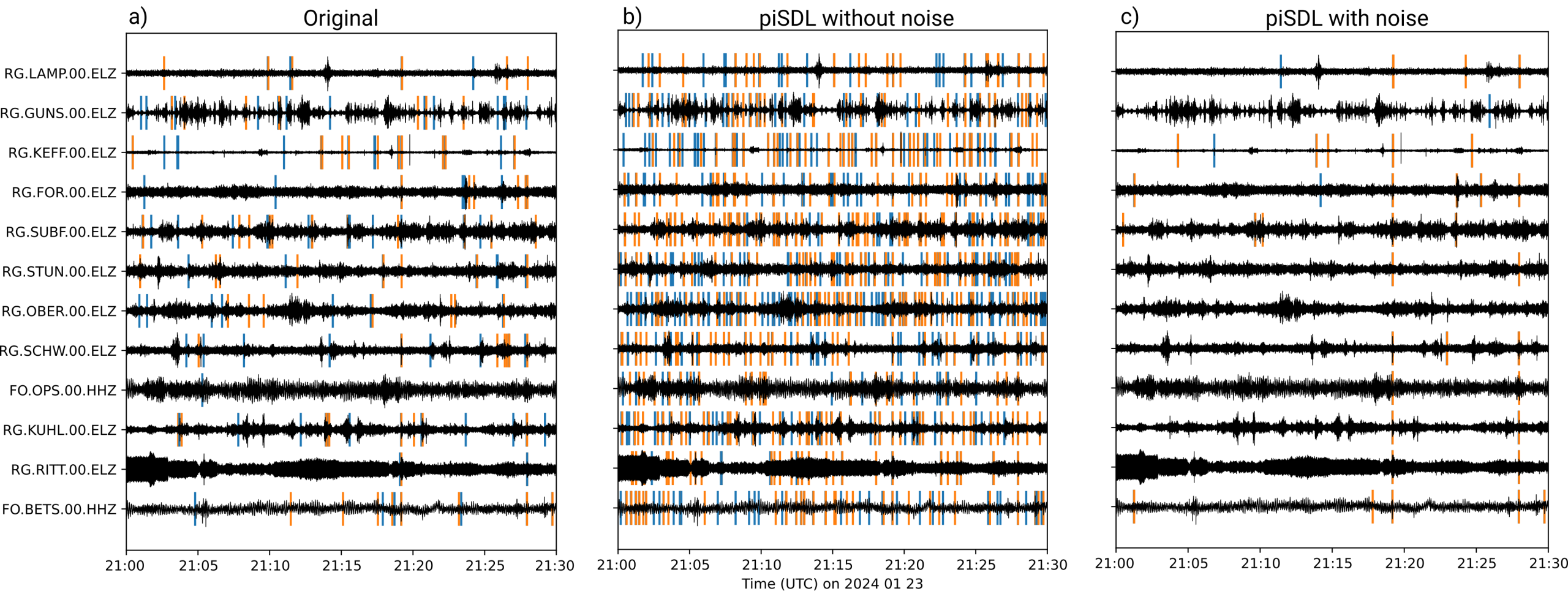
524 P-picks
478 S-picks

Too many false picks after training of PhaseNet with data set for induced seismicity

Solution:
Adding noise samples from STEAD and Rittershoffen to the training

Model learns to distinguish between noise and signal

Testing new PhaseNet model (piSDL)

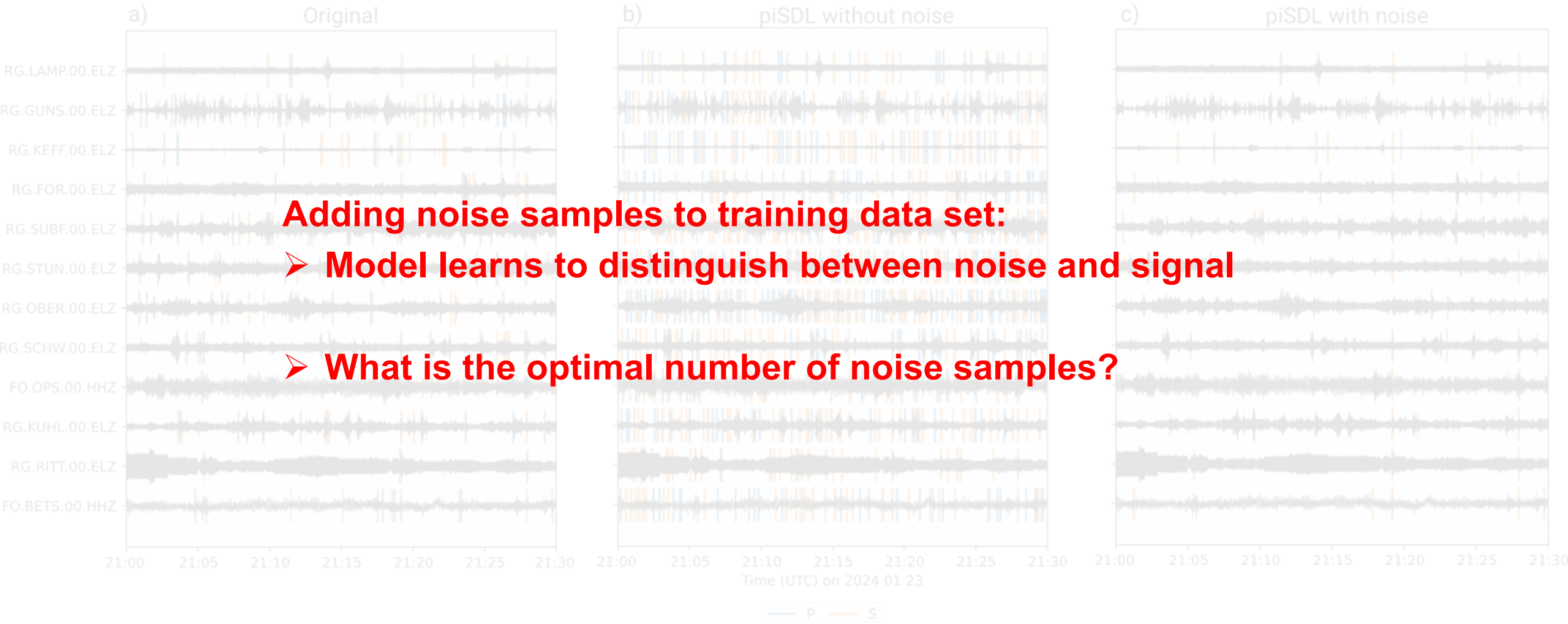


118 P-picks
97 S-picks

524 P-picks
478 S-picks

33 P-picks
36 S-picks

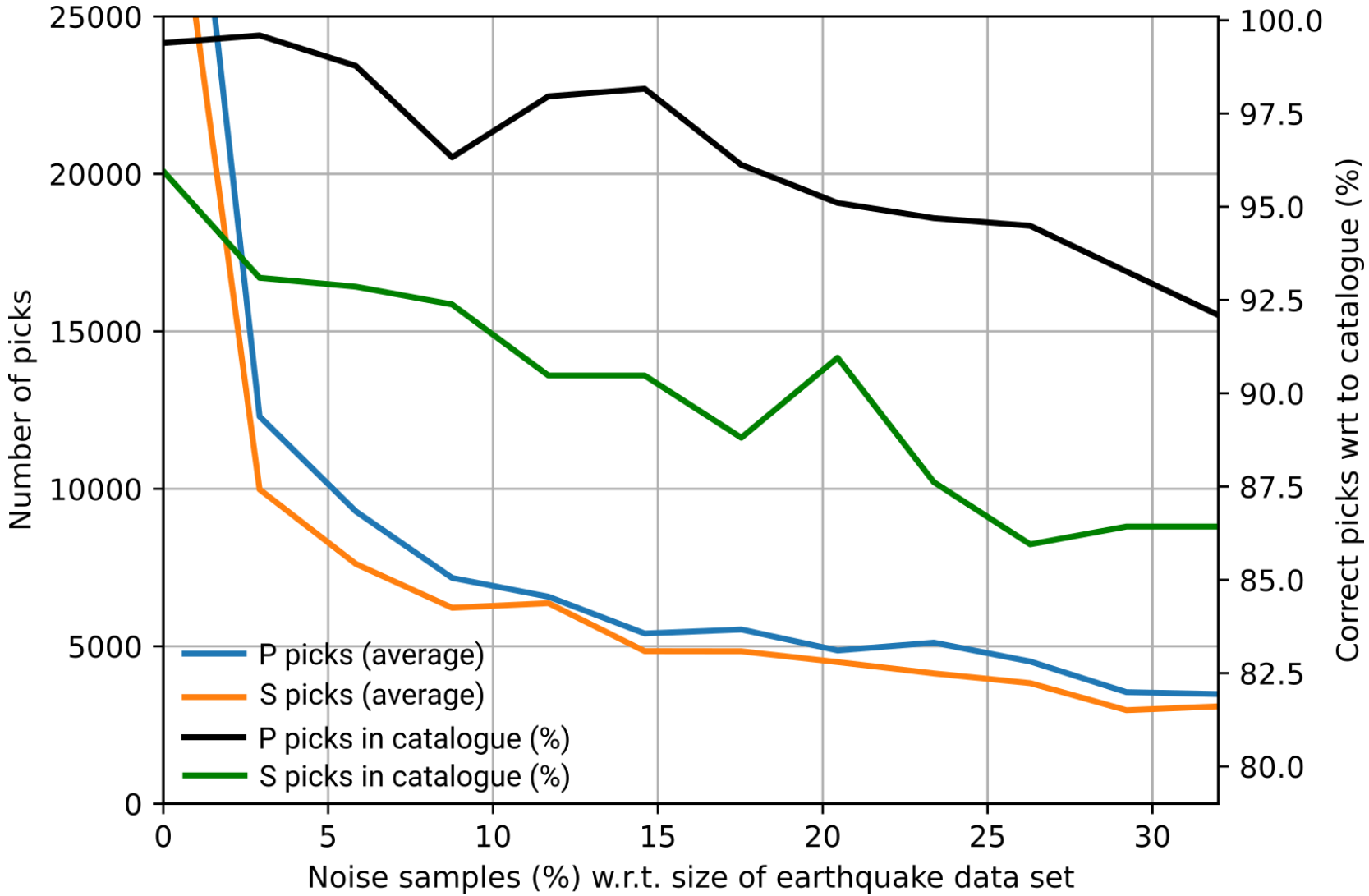
Testing the piSDL model



Adding noise samples to training data set:

- **Model learns to distinguish between noise and signal**
- **What is the optimal number of noise samples?**

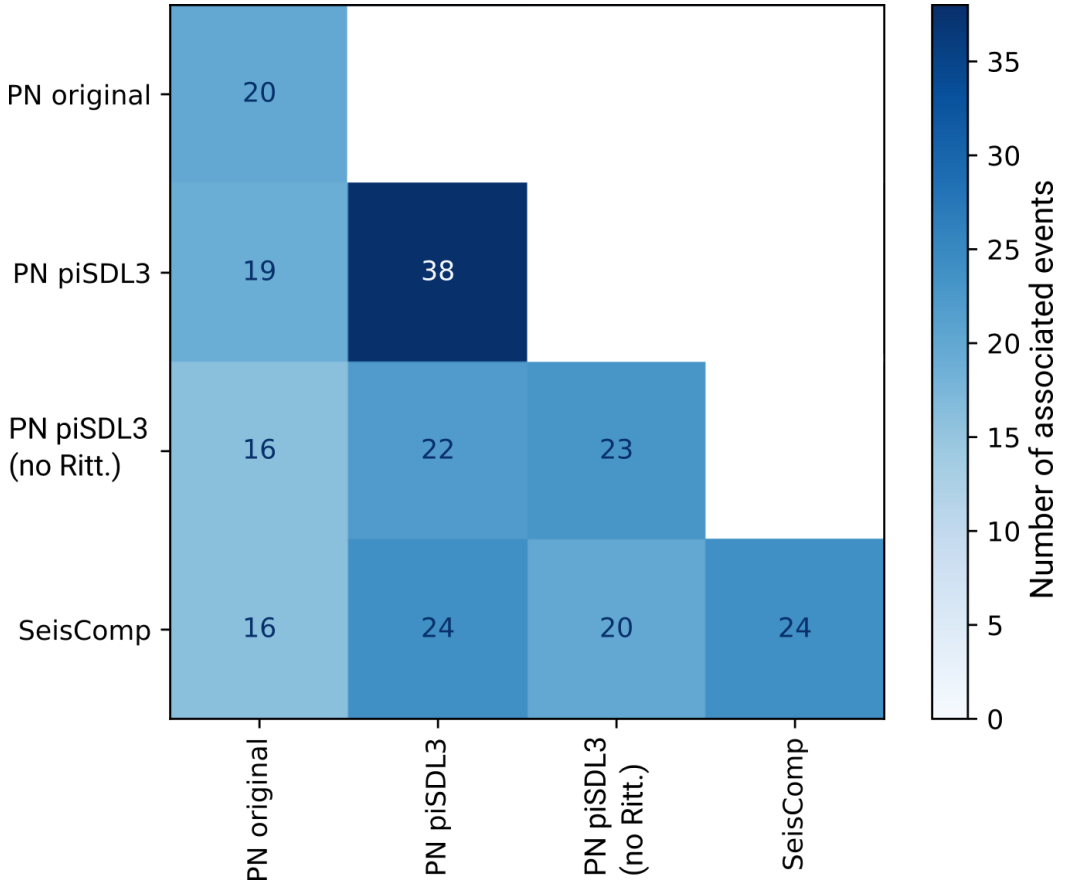
Optimal number of noise samples for piSDL



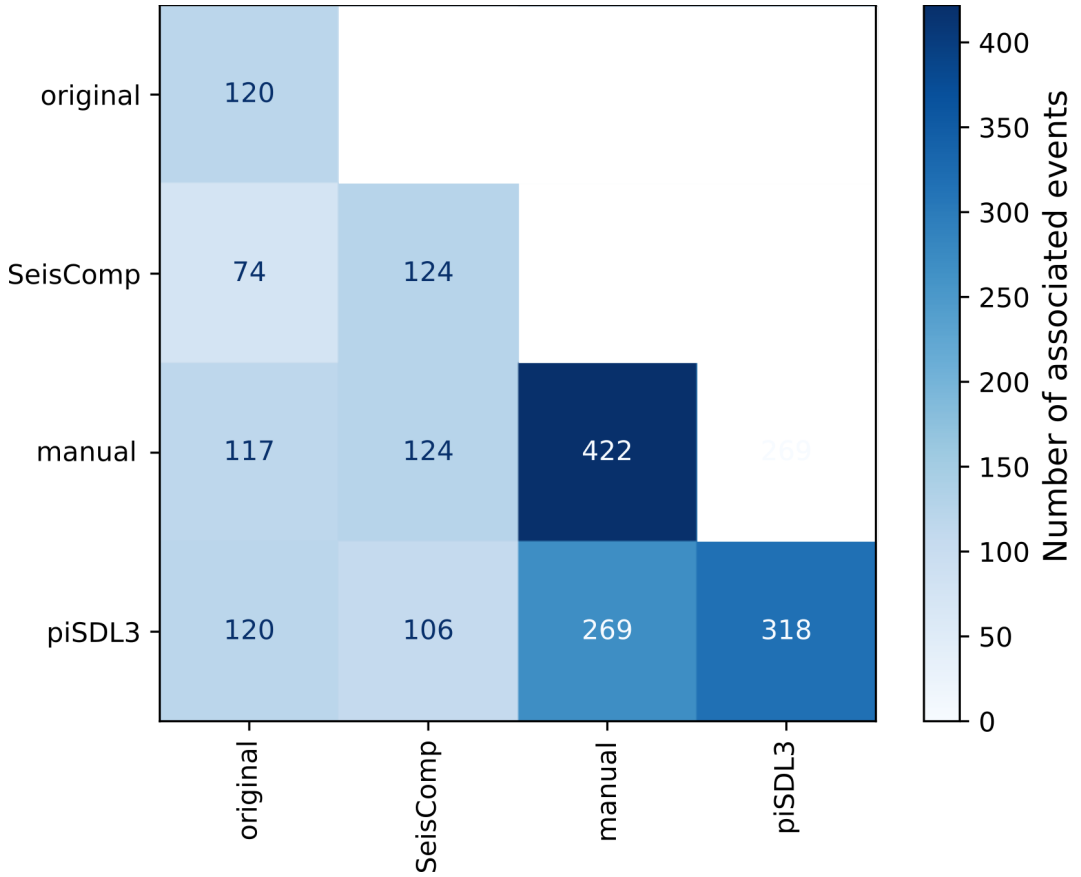
Adding 5 - 10% of noise samples to the training data set is the best way to reduce the number of false picks

Application of the PhaseNet model to continuous data

Catalogue comparison January 2024 (Rittershoffen, France)



Catalogue comparison* 2024-03-02 06:00 - 09:00 (Rittershoffen, France)



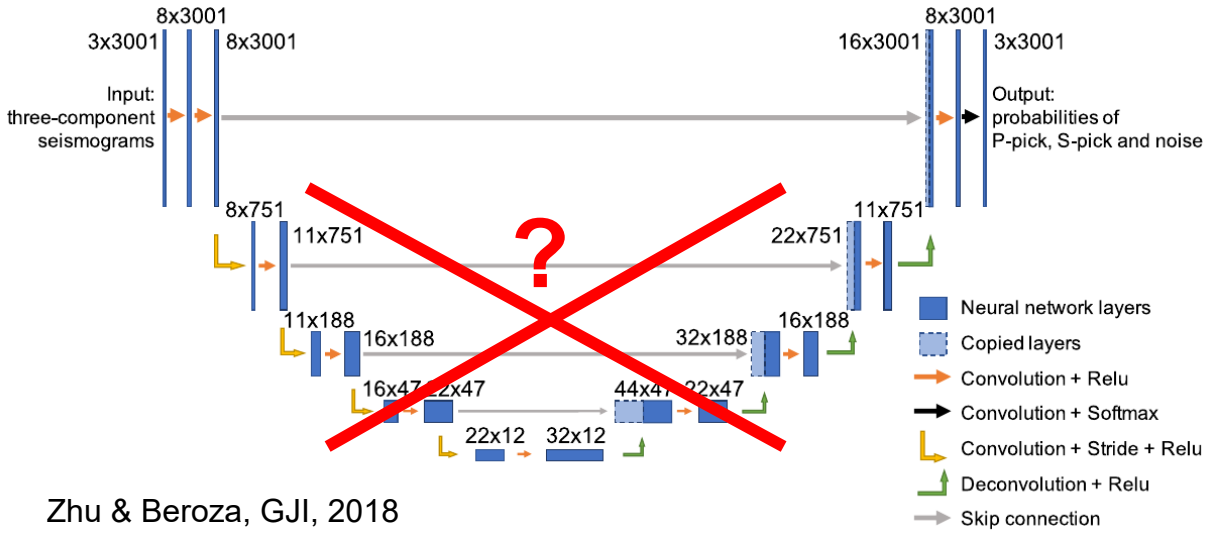
*only compared origin time

Conclusions

- **New data set** for AI in seismology **for induced seismicity**
- Adding **noise samples** to the training data set **reduces the number of false picks**
  - ➢ PhaseNet learns to distinguish between earthquake and noise samples
- **piSDL model performs well** during phases with low- as well as with high-seismicity
- In comparison to manual picked catalogue, a **few events are still missing**
  - ➢ Apply template-matching to detect more events
- Data set and piSDL model **will be available in SeisBench**

Current work

- Finding **best set of hyperparameters** for PhaseNet by evolutionary optimization



Zhu & Beroza, GJI, 2018

References

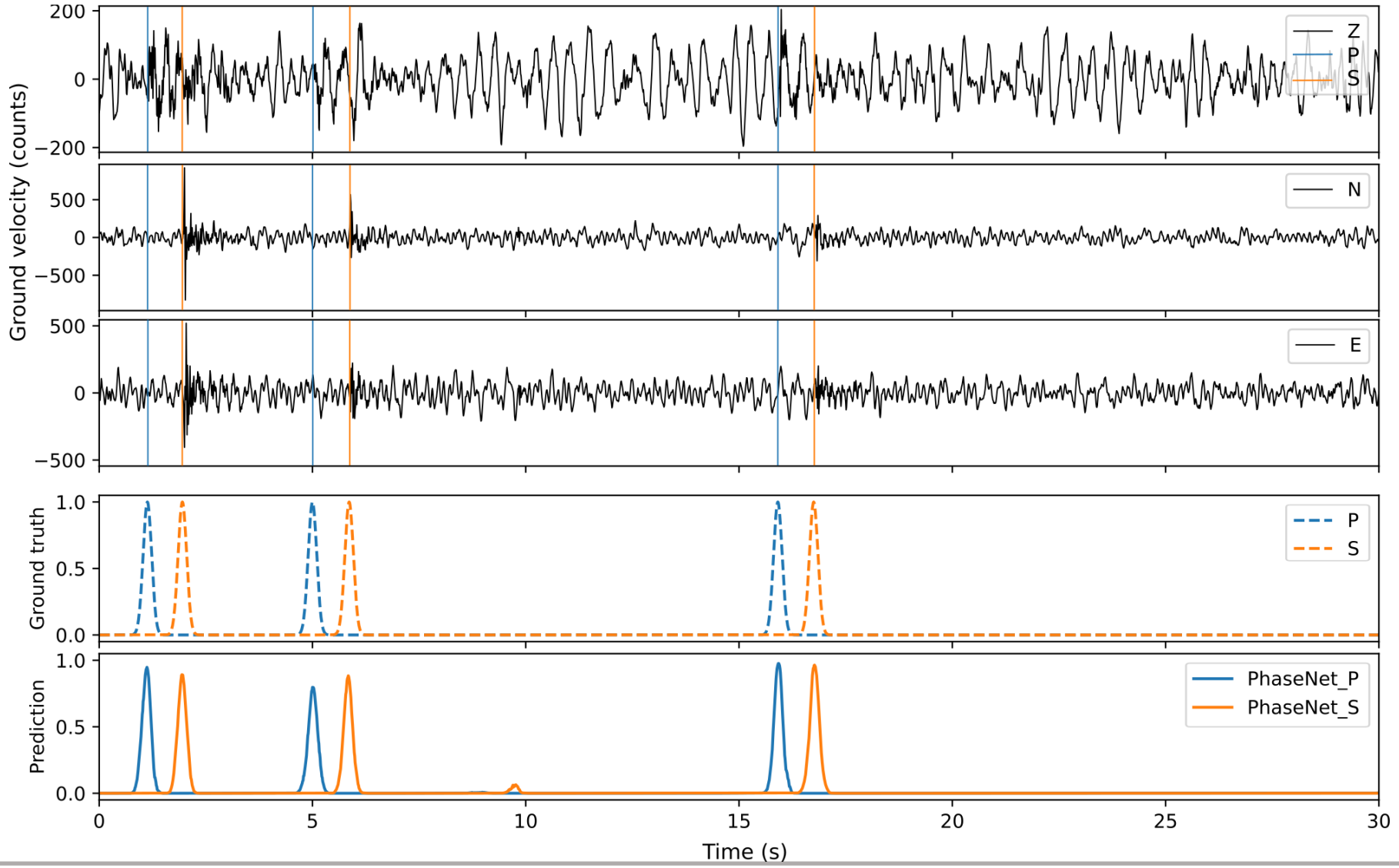
Heuel, J., Maurer, V., Frietsch, M., Rietbrock, A. (2025). Picking Induced Seismicity with Deep Learning (piSDL). *Seismica (submitted)*

Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019). STanford EArthquake Dataset (STEAD): A global data set of seismic signals for AI. *IEEE Access*, *7*, 179464-179476.

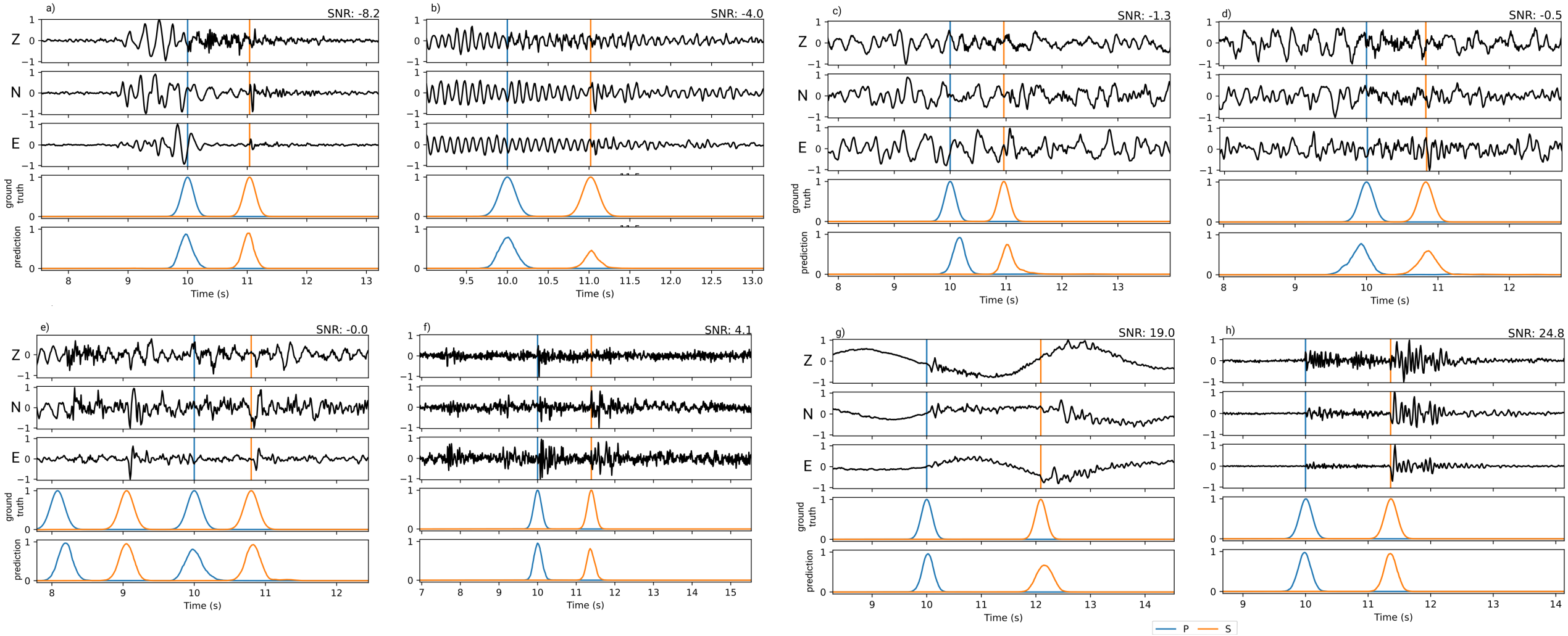
Woollam, J., Münchmeyer, J., Tilmann, F., Rietbrock, A., Lange, D., Bornstein, T., ... & Soto, H. (2022). SeisBench—A toolbox for machine learning in seismology. *Seismological Society of America*, *93*(3), 1695-1709.

Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, *216*(1), 261-273.

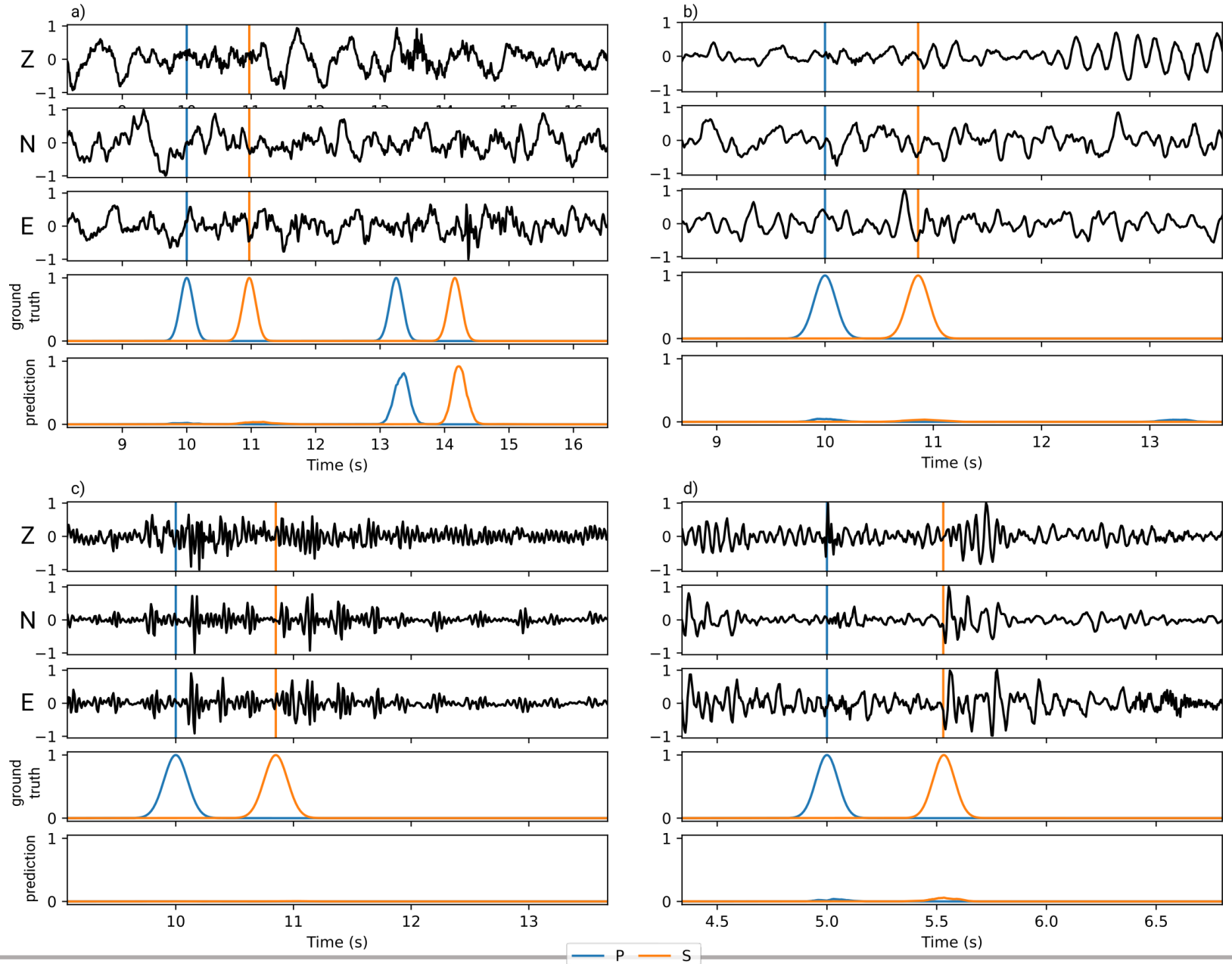
Appendix


Ground velocity (counts)
Z
P
S
N
E
Ground truth
P
S
Prediction
PhaseNet_P
PhaseNet_S
Time (s)

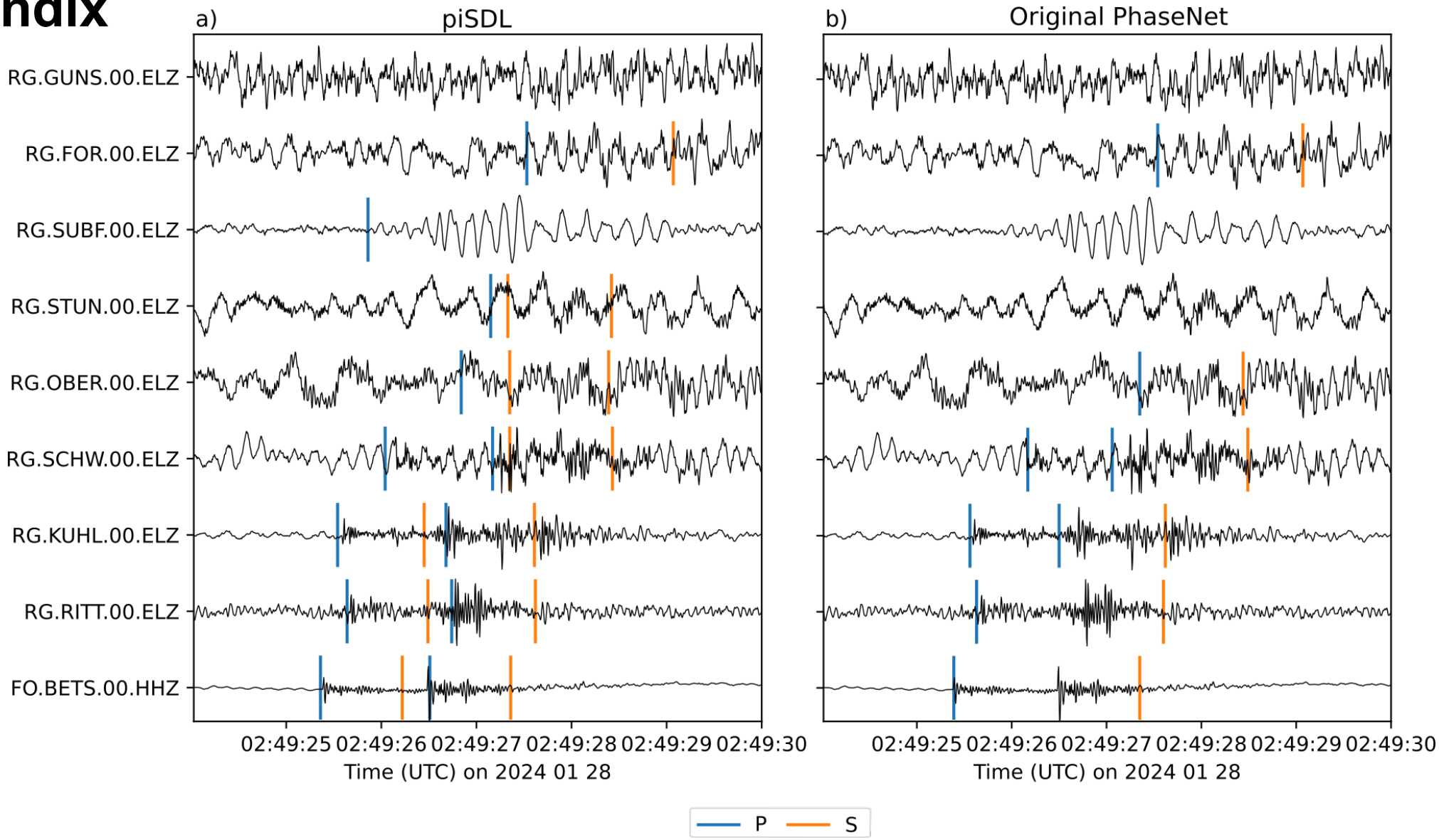
Appendix


a)
SNR: -8.2
b)
SNR: -4.0
c)
SNR: -1.3
d)
SNR: -0.5
e)
SNR: -0.0
f)
SNR: 4.1
g)
SNR: 19.0
h)
SNR: 24.8
Z
N
E
ground truth
prediction
Time (s)
P
S

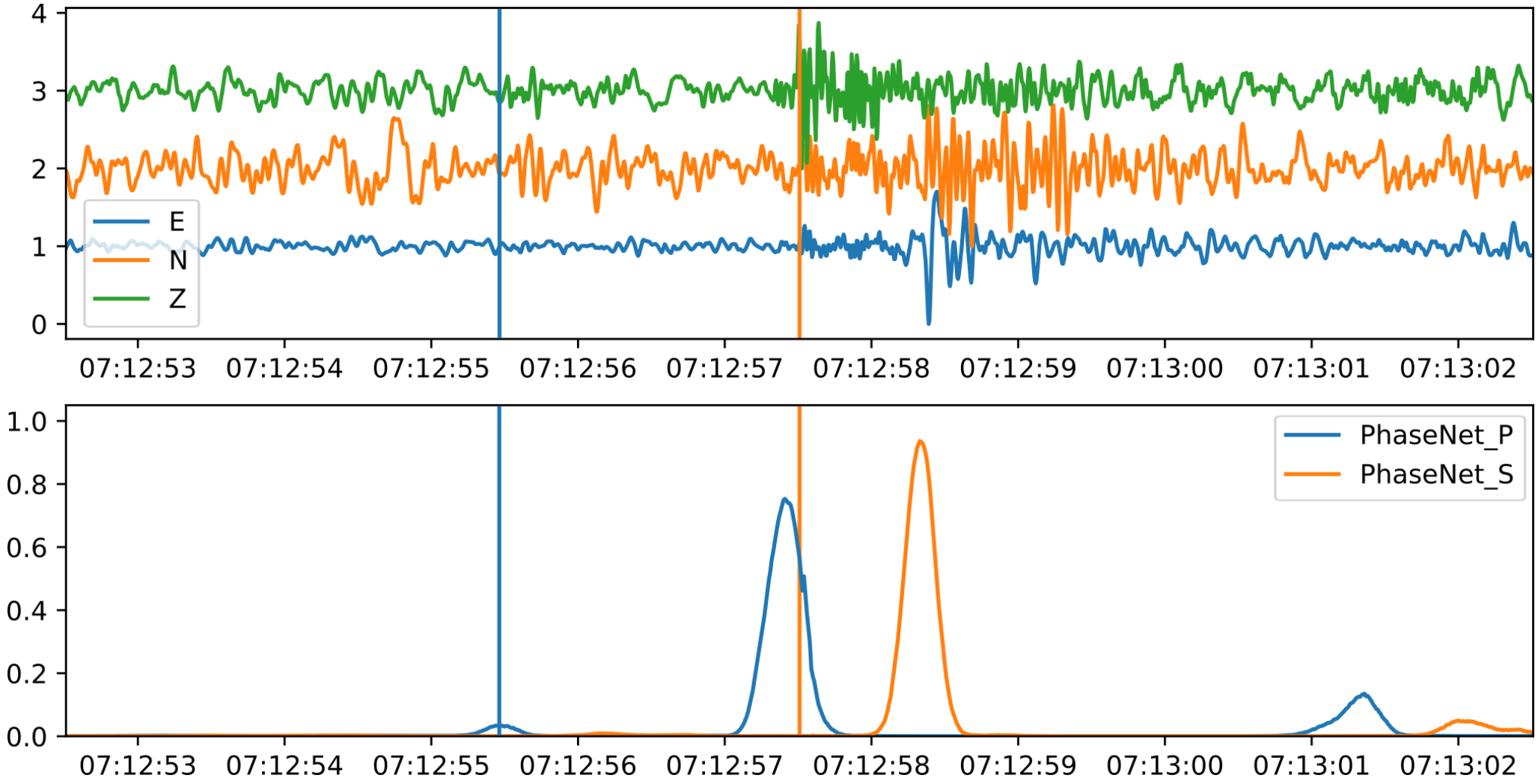
Appendix



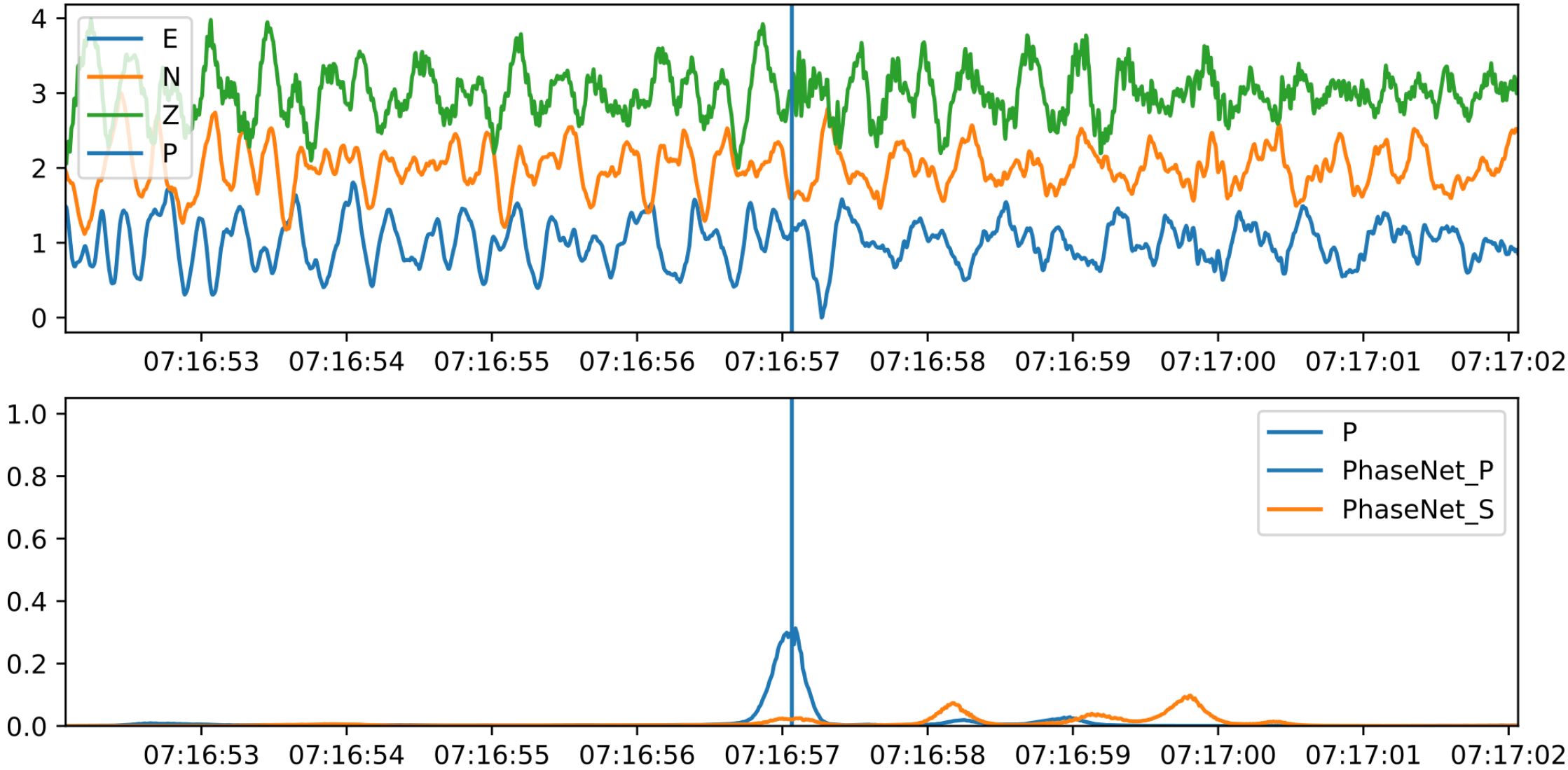
Appendix


a)
piSDL
b)
Original PhaseNet
RG.GUNS.00.ELZ
RG.FOR.00.ELZ
RG.SUBF.00.ELZ
RG.STUN.00.ELZ
RG.OBER.00.ELZ
RG.SCHW.00.ELZ
RG.KUHL.00.ELZ
RG.RITT.00.ELZ
FO.BETS.00.HHZ
02:49:25 02:49:26 02:49:27 02:49:28 02:49:29 02:49:30
Time (UTC) on 2024 01 28
02:49:25 02:49:26 02:49:27 02:49:28 02:49:29 02:49:30
Time (UTC) on 2024 01 28
P
S

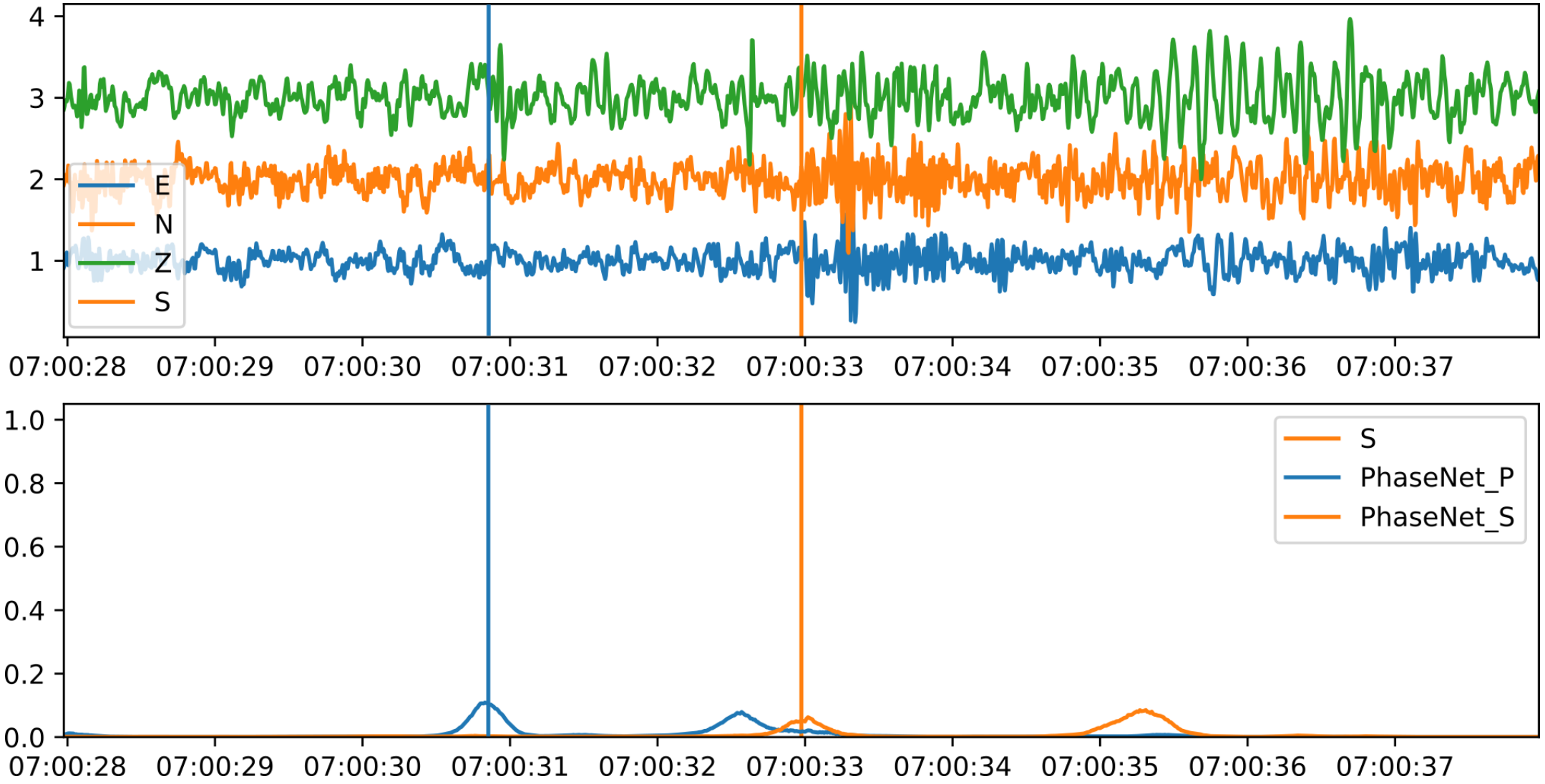
Appendix (Failed example)



Appendix (Failed example)



Appendix (Failed example)



Appendix

An was für Events scheitern wir noch?
Bestes Modell aktuell: noise samples + adding noise on earthquake waveform (SNR >= 10 dB)
Wie viele Beben finden wir mit SeisComp?
Histogramme von P und S mit SeisComp Vergleich
Vergleich welche Beben nicht von PhaseNet Modell, aber von SeisComp gepickt werden
Vergleich welche Beben nicht von PhaseNet Modell, aber von Hand gepickt wurden

Fazit: Man könnte im nächsten Schritt immer noch mit template matching nach weiteren Events suchen

124 Events, die automatisch von SeisComp gepickt worden sind. Nach automatischer und händischer Analyse konnten all Events mit PhaseNet picks gefunden werden
Voller automatischer PhaseNet Katalog hat 318 Events (mehrere Iterationen mit PyOcto)

Summary:
Different training stratetgies for PhaseNet with noise samples