Deep learning prediction of measured earthquake waveforms from synthetic data

A. Bauer, J. Walda and C. Hammer

Institute of Geophysics, University of Hamburg

Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

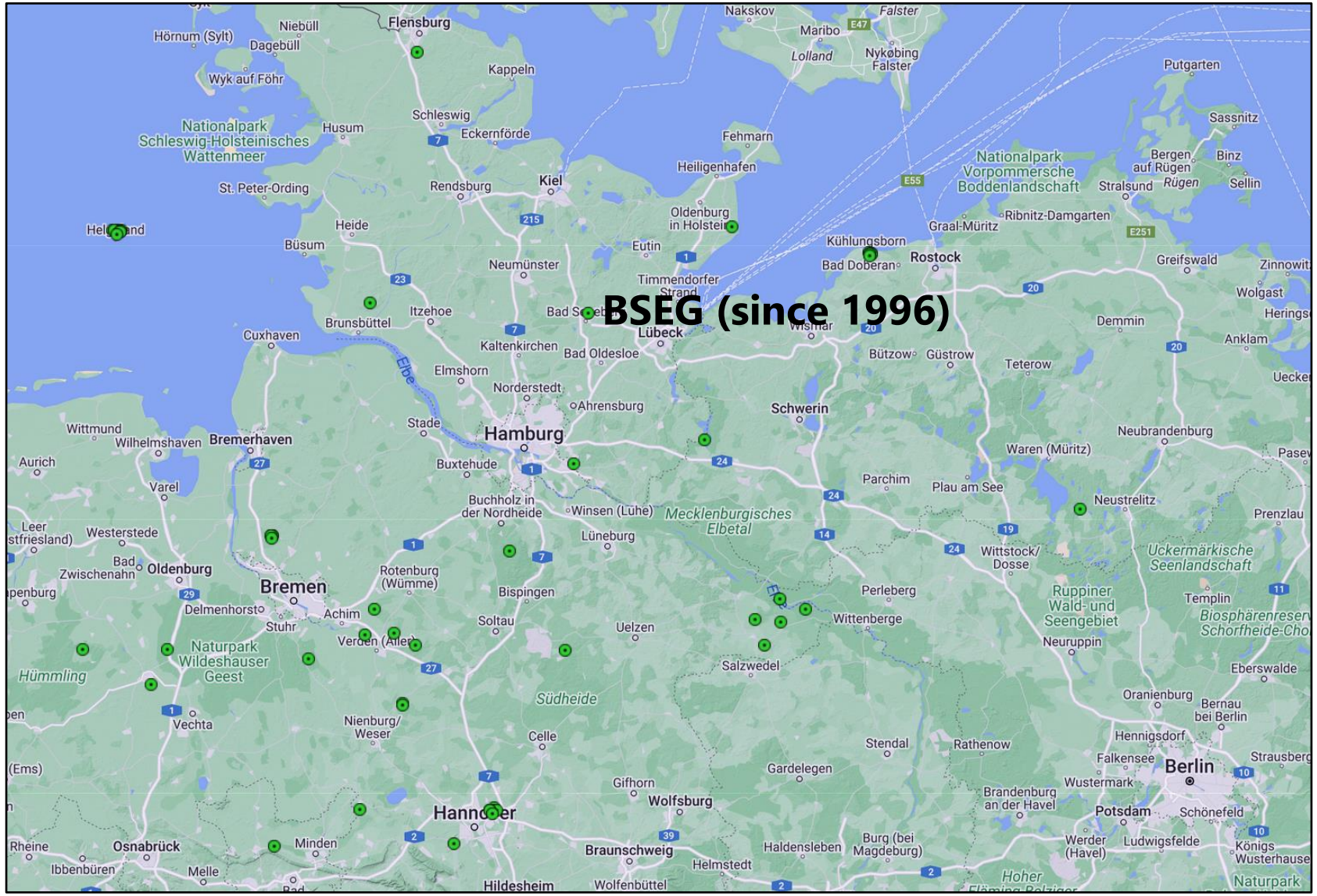
Deep learning prediction of earthquake waveforms

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Detailed knowledge of the seismic wavefield generated by **large teleseismic earthquakes** is crucial for high-precision measurements and experiments
- Long-term goal: **predict measured data** for earthquakes at arbitrary coordinates starting from synthetic data
- **Synthetic earthquake waveforms** can be generated for arbitrary coordinates
- **First step: train a CNN** to predict measured data for existing stations
  - **Input**: synthetic data generated for large earthquakes in the past
  - **Labels**: data measured at an existing seismological station
  - **Test**: application of trained network to earthquakes not part of training data

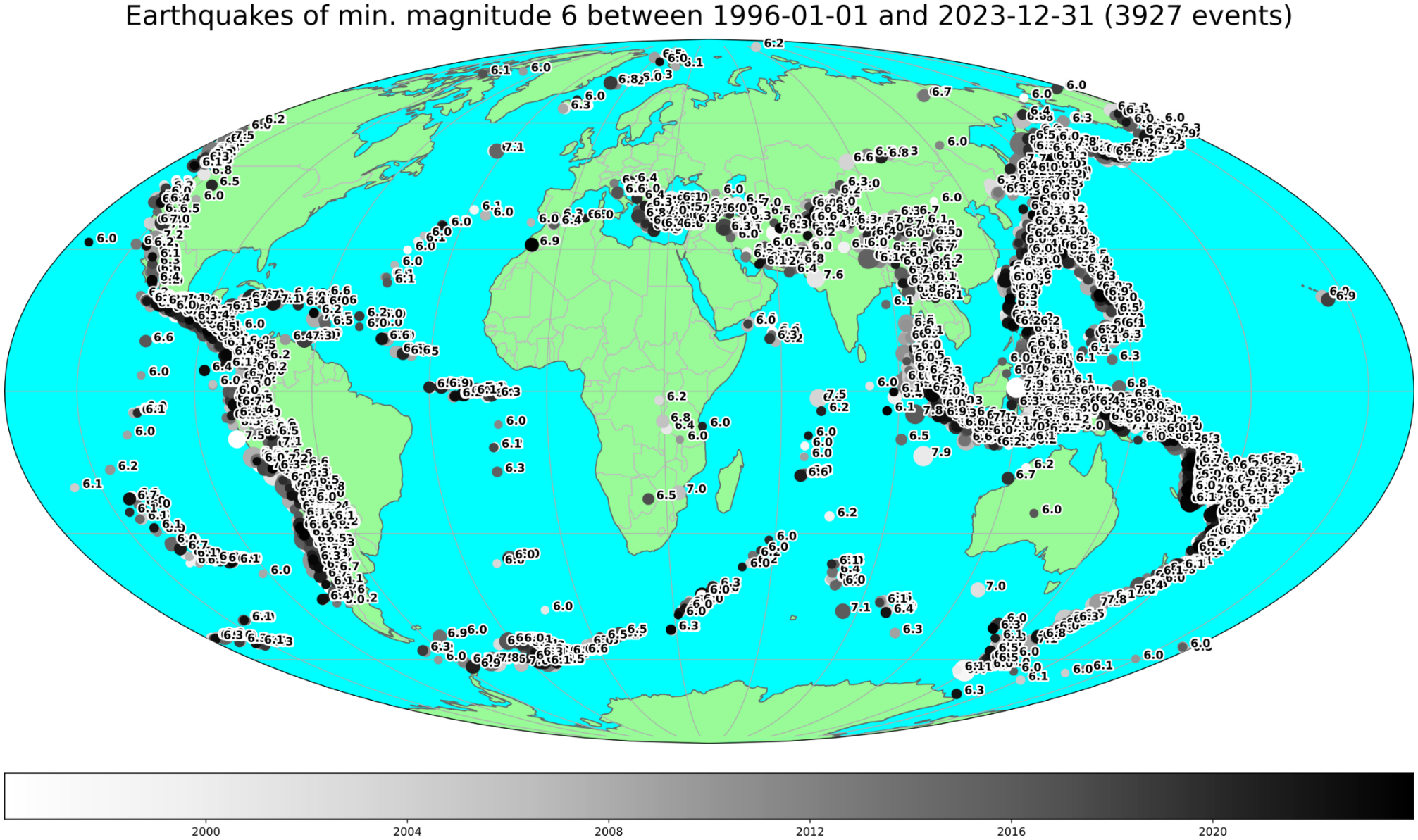
Seismological stations in Northern Germany

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



Source: Seismological Facility for the Advancement of Geoscience (SAGE)

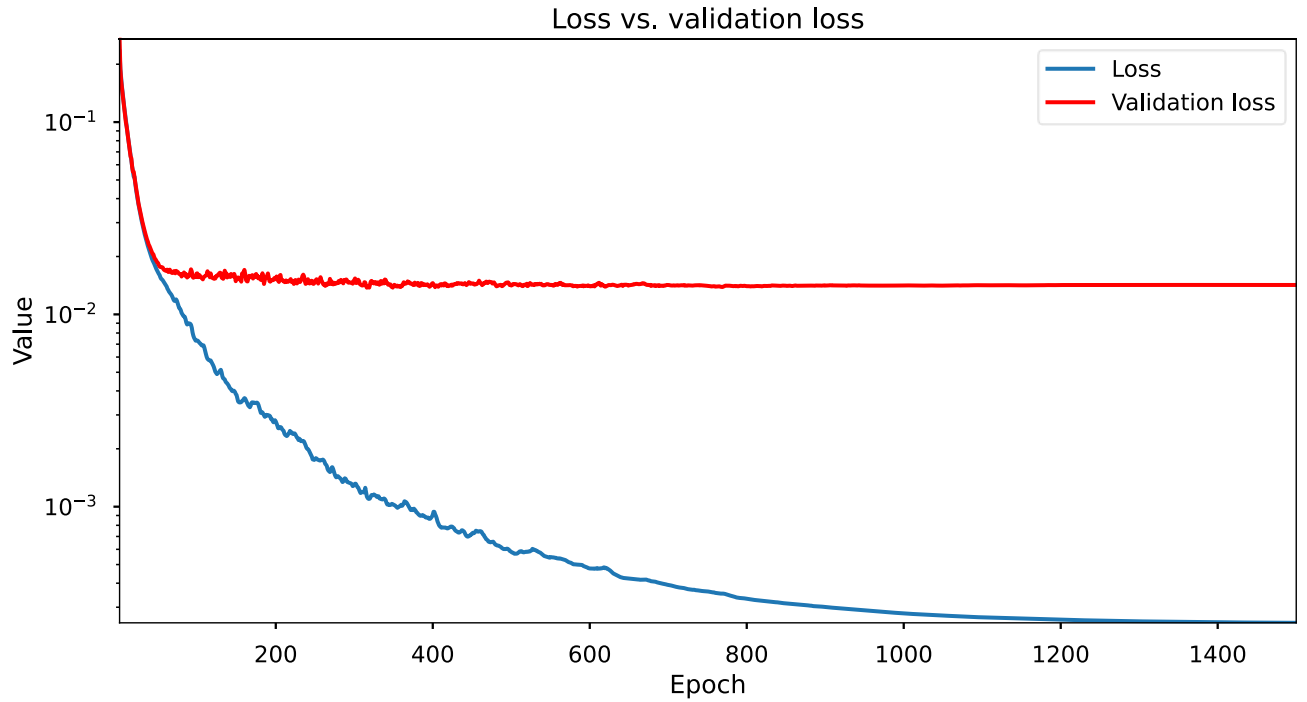
Measured earthquakes ≥ M6.0 (3927 events)



Data preparation and neural network training

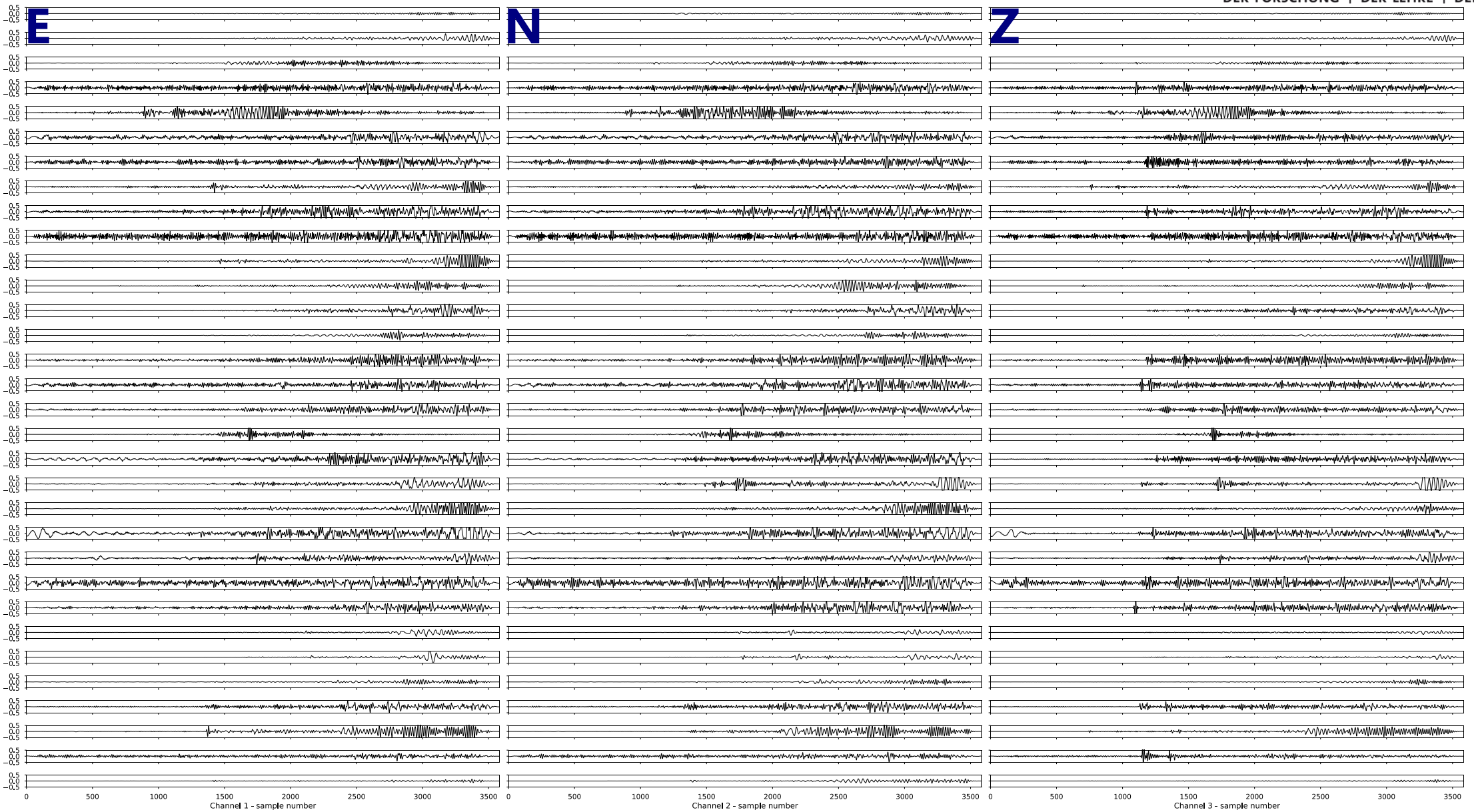
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- **Data size**: 3216 events (+22 events for application)
  - 3584 samples @ 1 Hz per event
- Bandpass-filtered between 0.01 and 0.1 Hz
- Normalized between -1 and 1
- **Neural network**:
  - Convolutional autoencoder
  - Depth 4
  - 2 ResNeXt blocks per depth level
  - Skipping connections
  - Dense bottleneck layer
- Initial learning rate $5 \cdot 10^{-4}$
- 1500 epochs of training



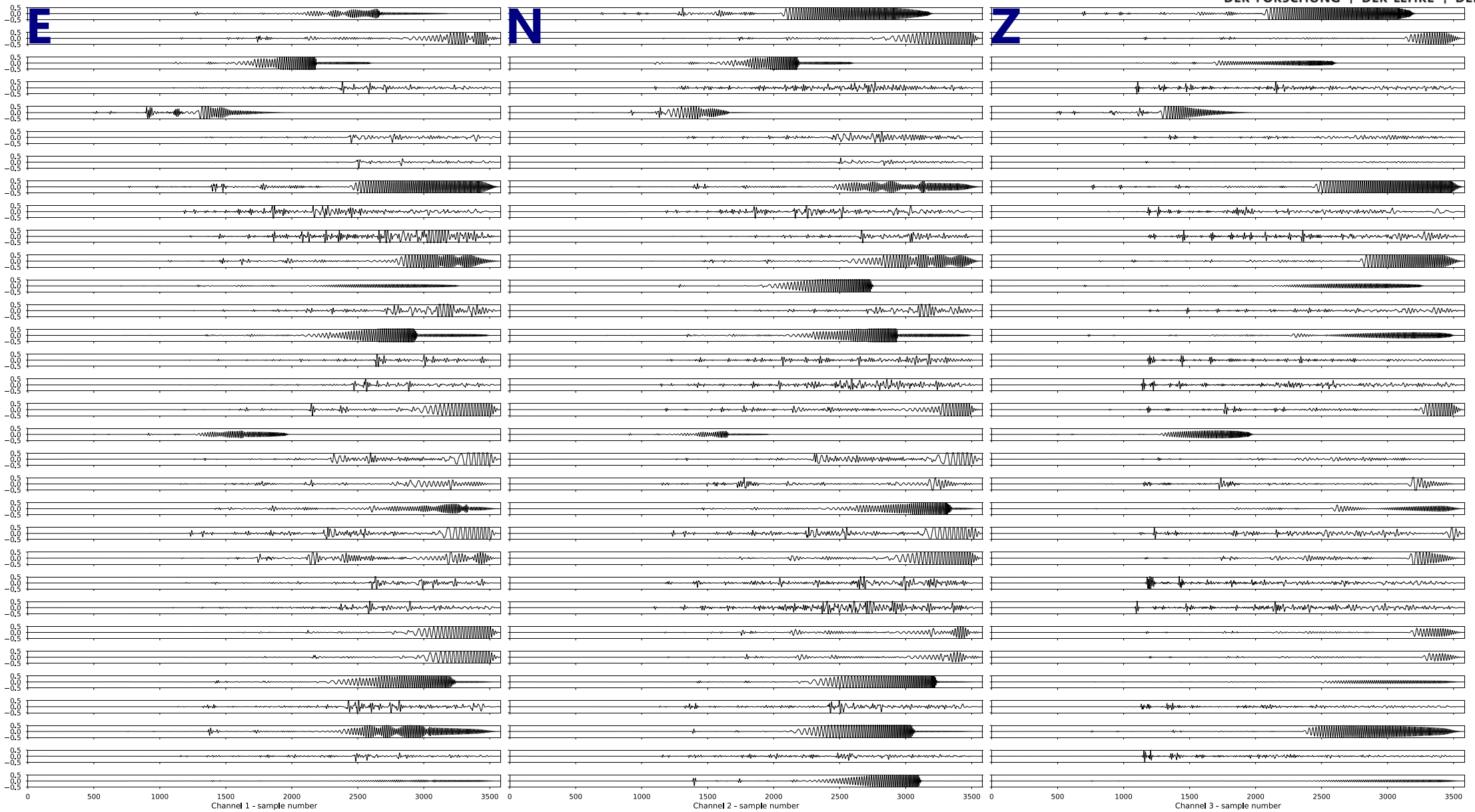
Neural network training: measured data (labels)

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

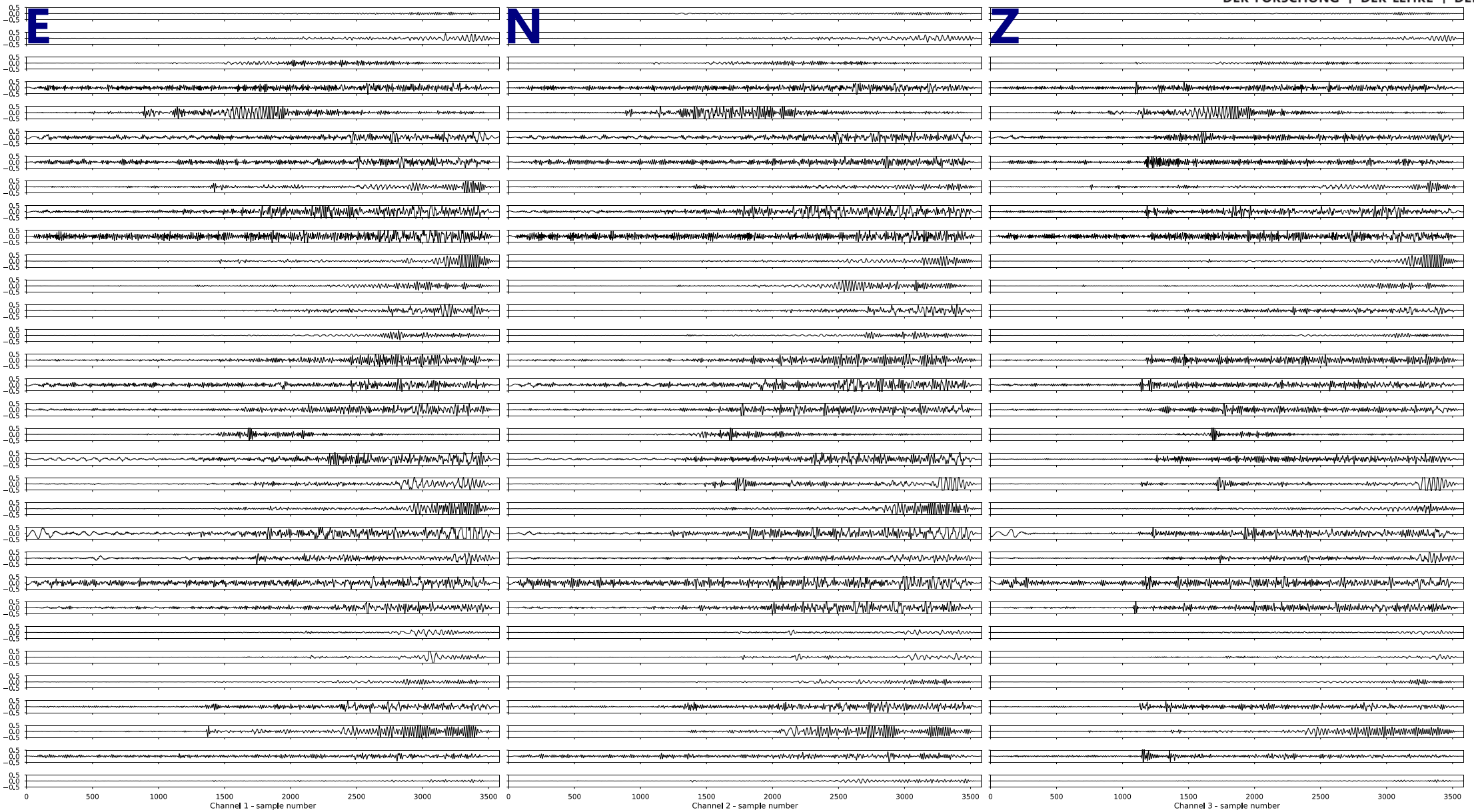


Neural network training: synthetic data (input)

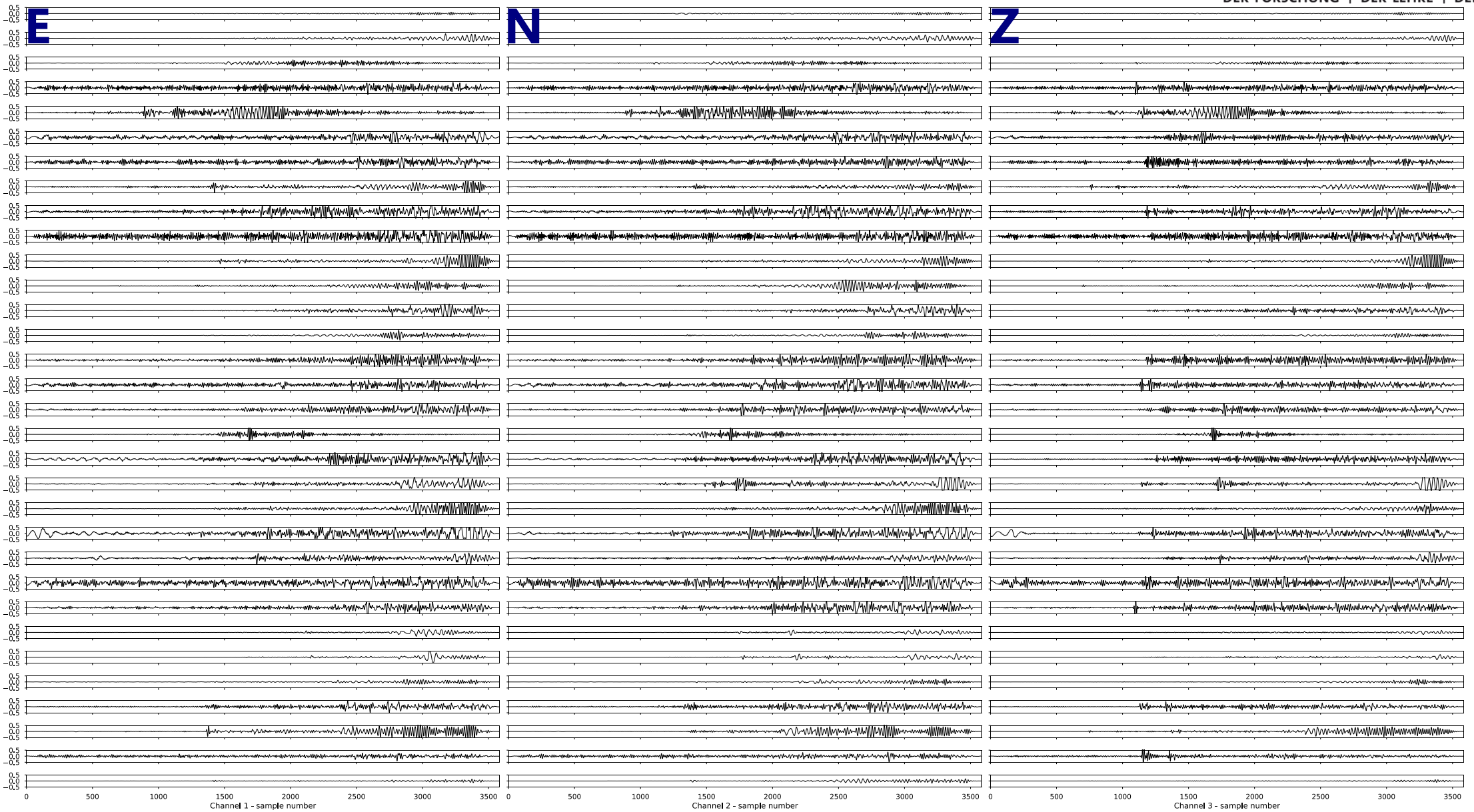
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



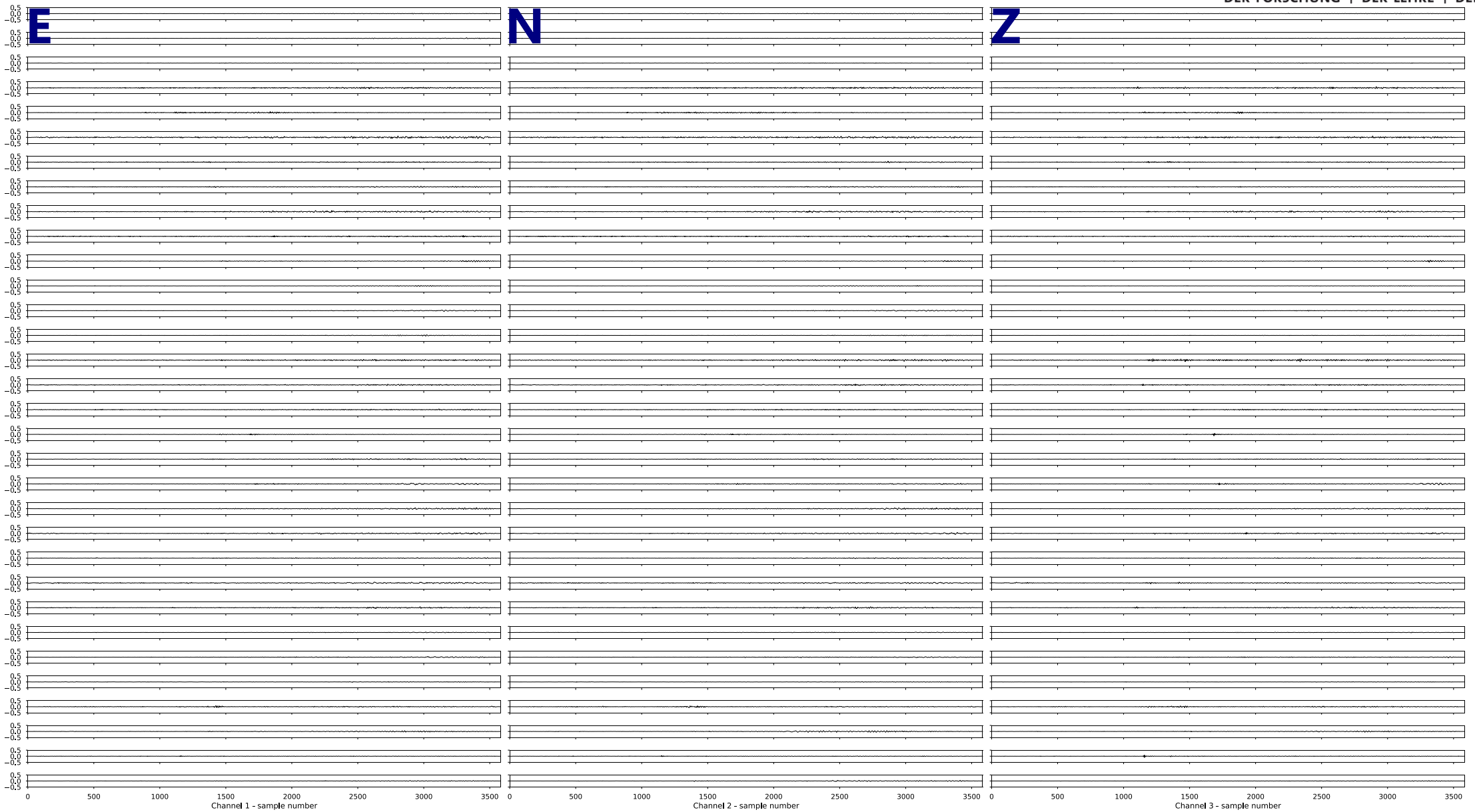
Neural network training: predictions



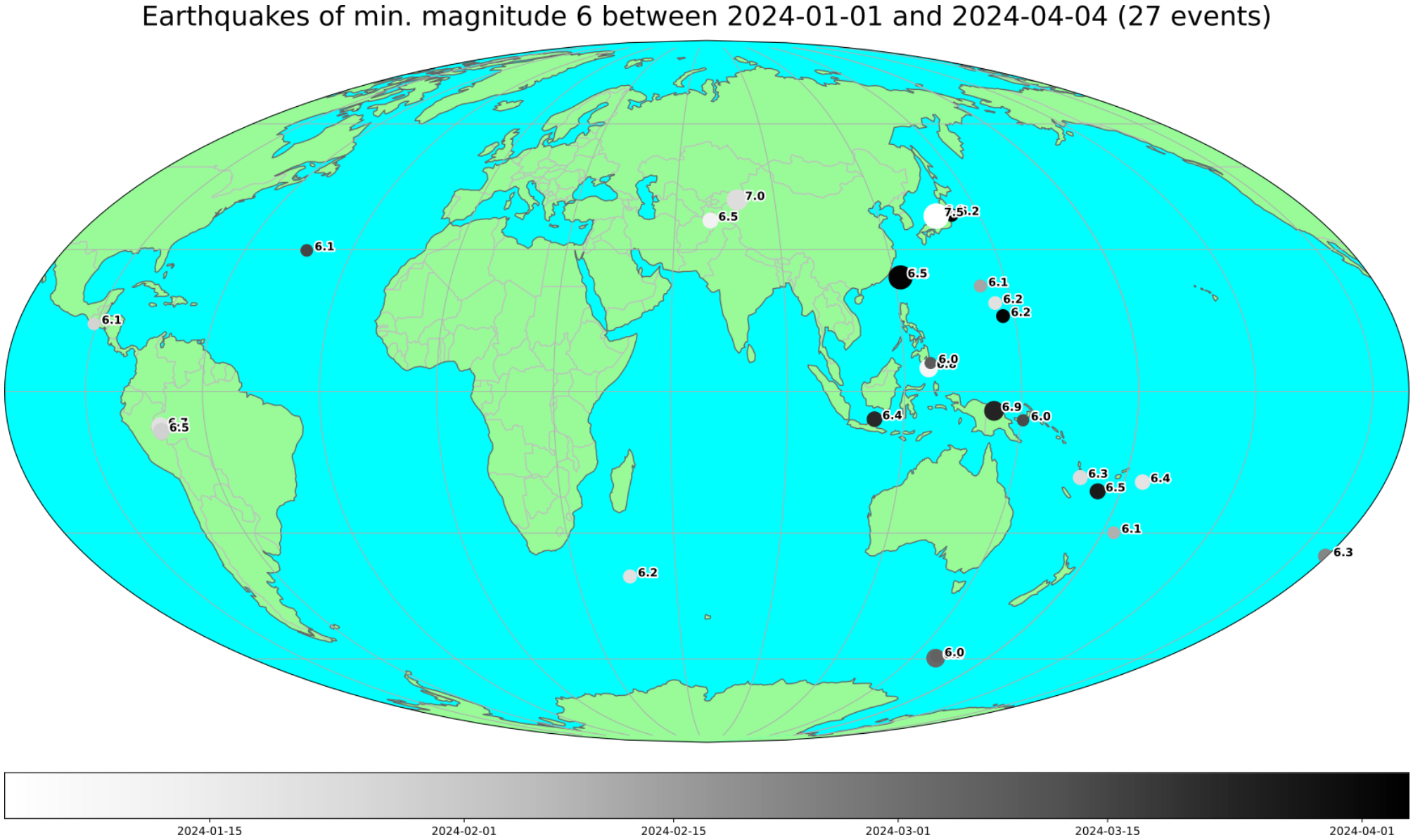
Neural network training: measured data (labels)



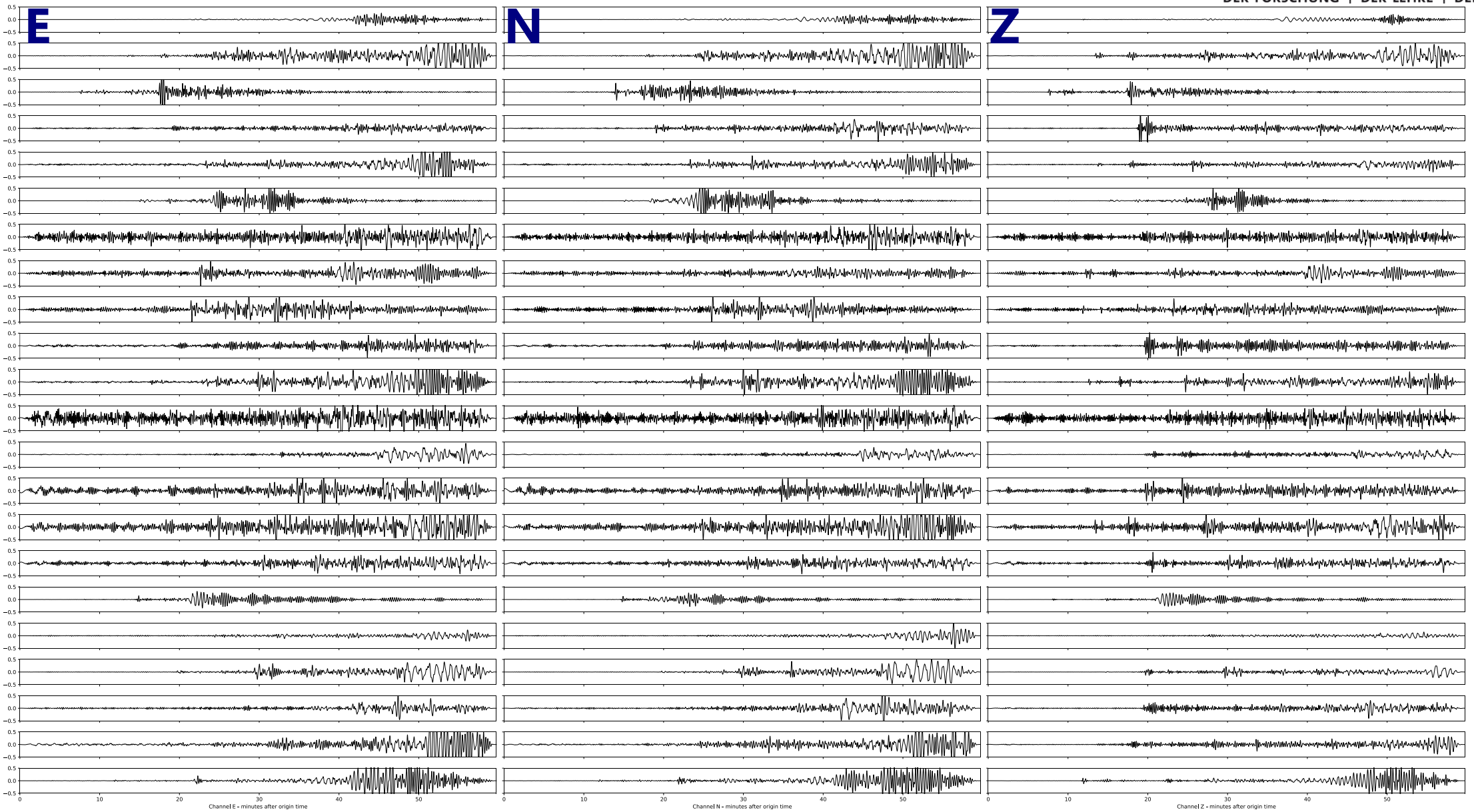
Neural network training: labels – predictions



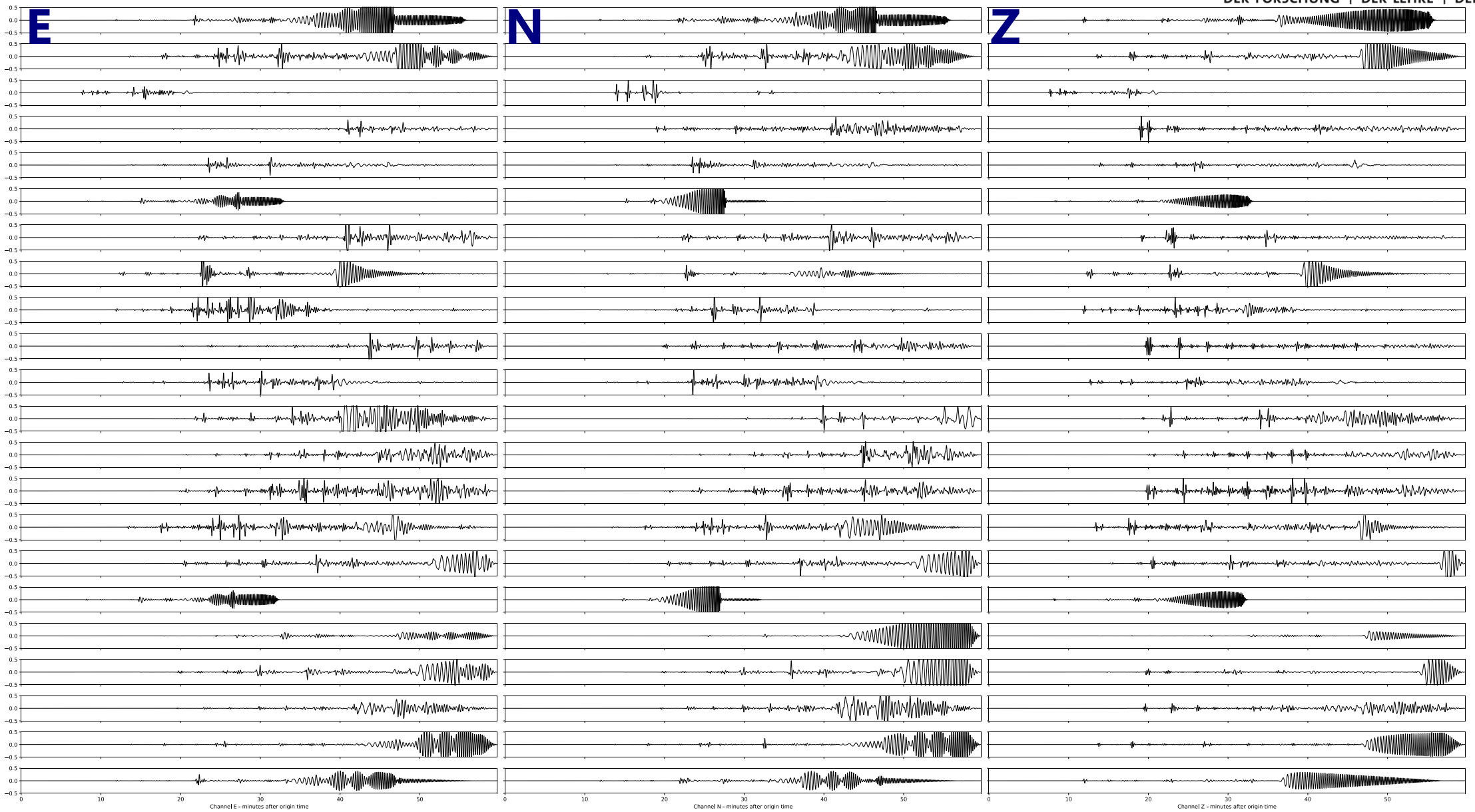
Unseen measured earthquakes ≥ M6.0



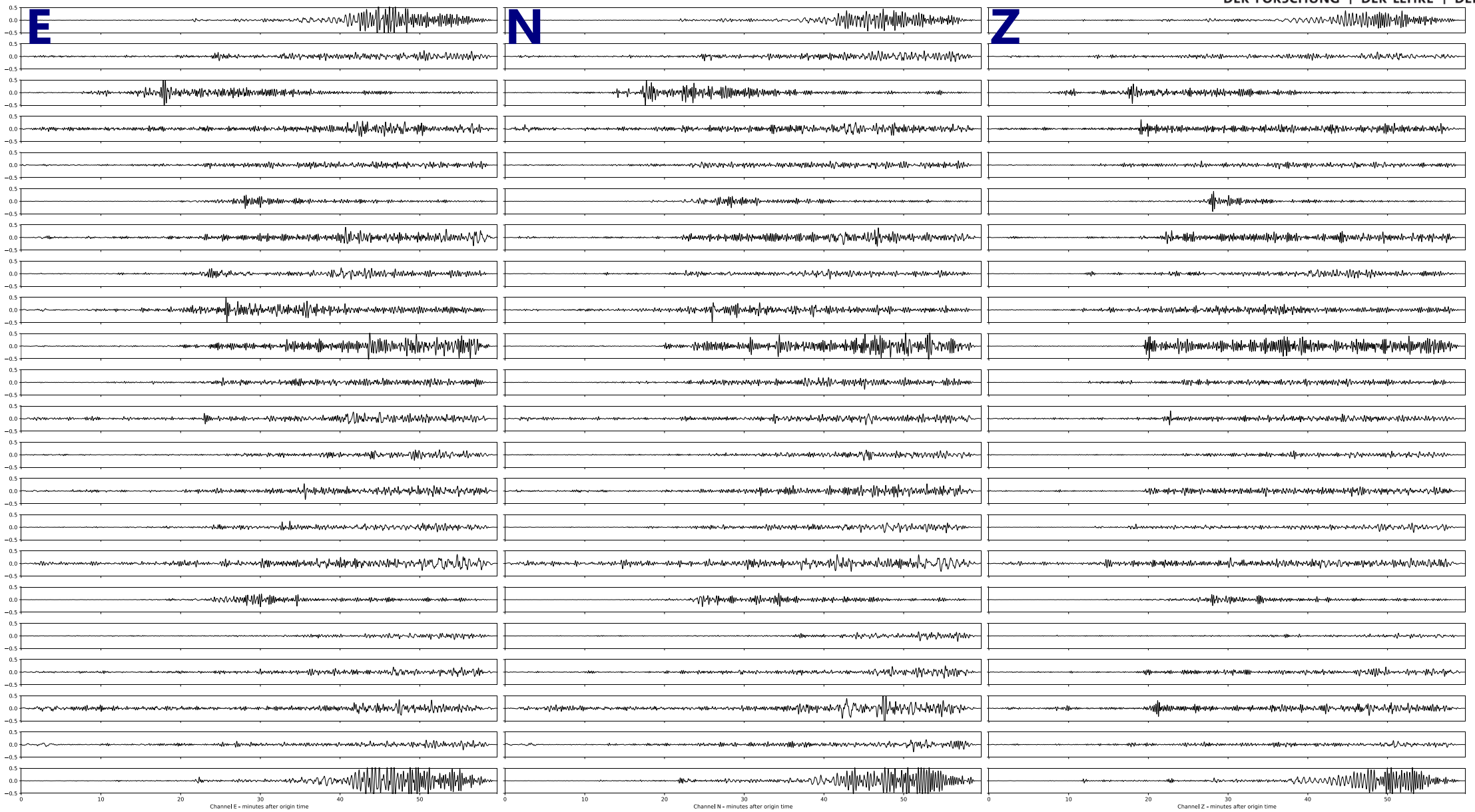
Application: unseen measured data



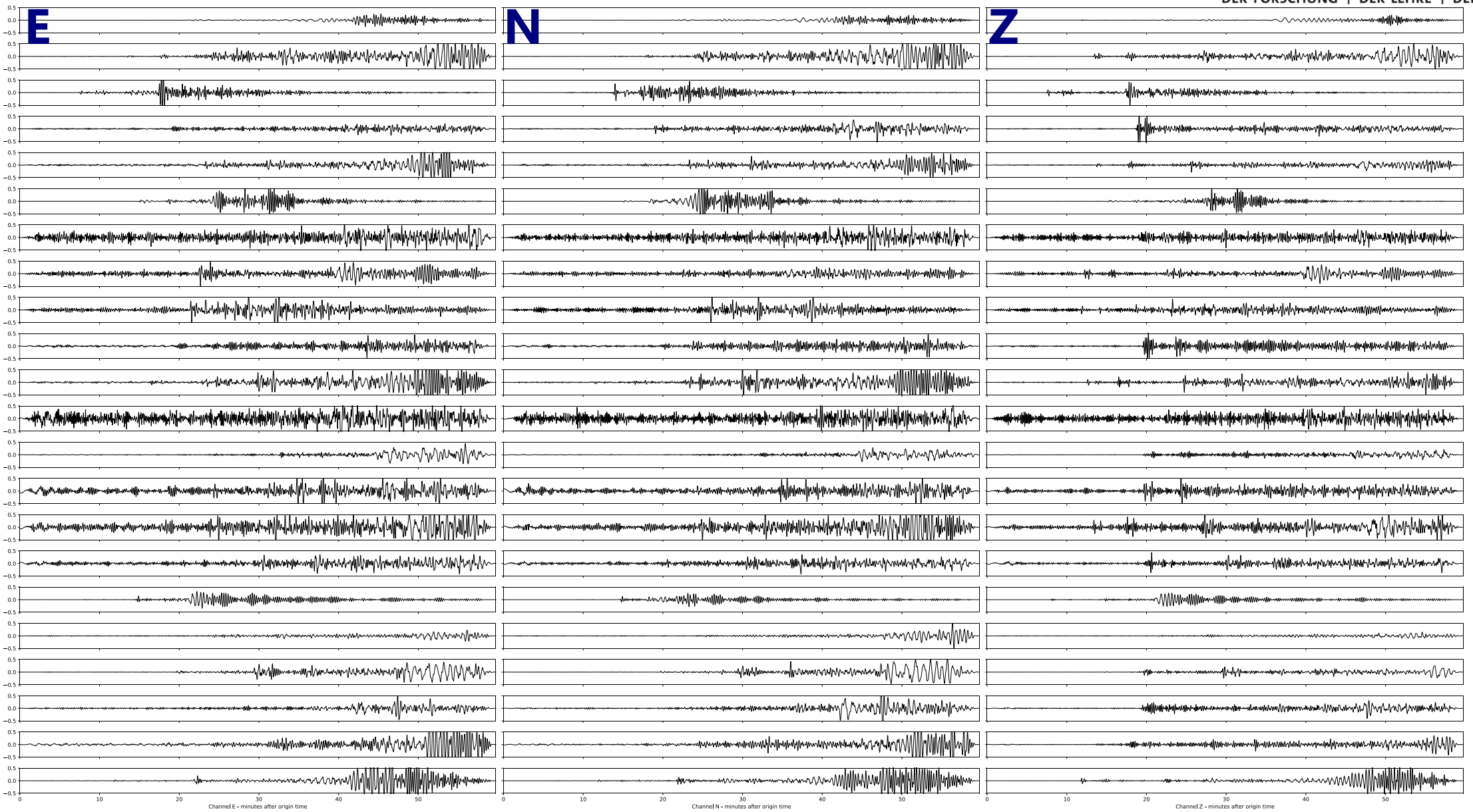
Application: synthetic data (input)



Application: predictions

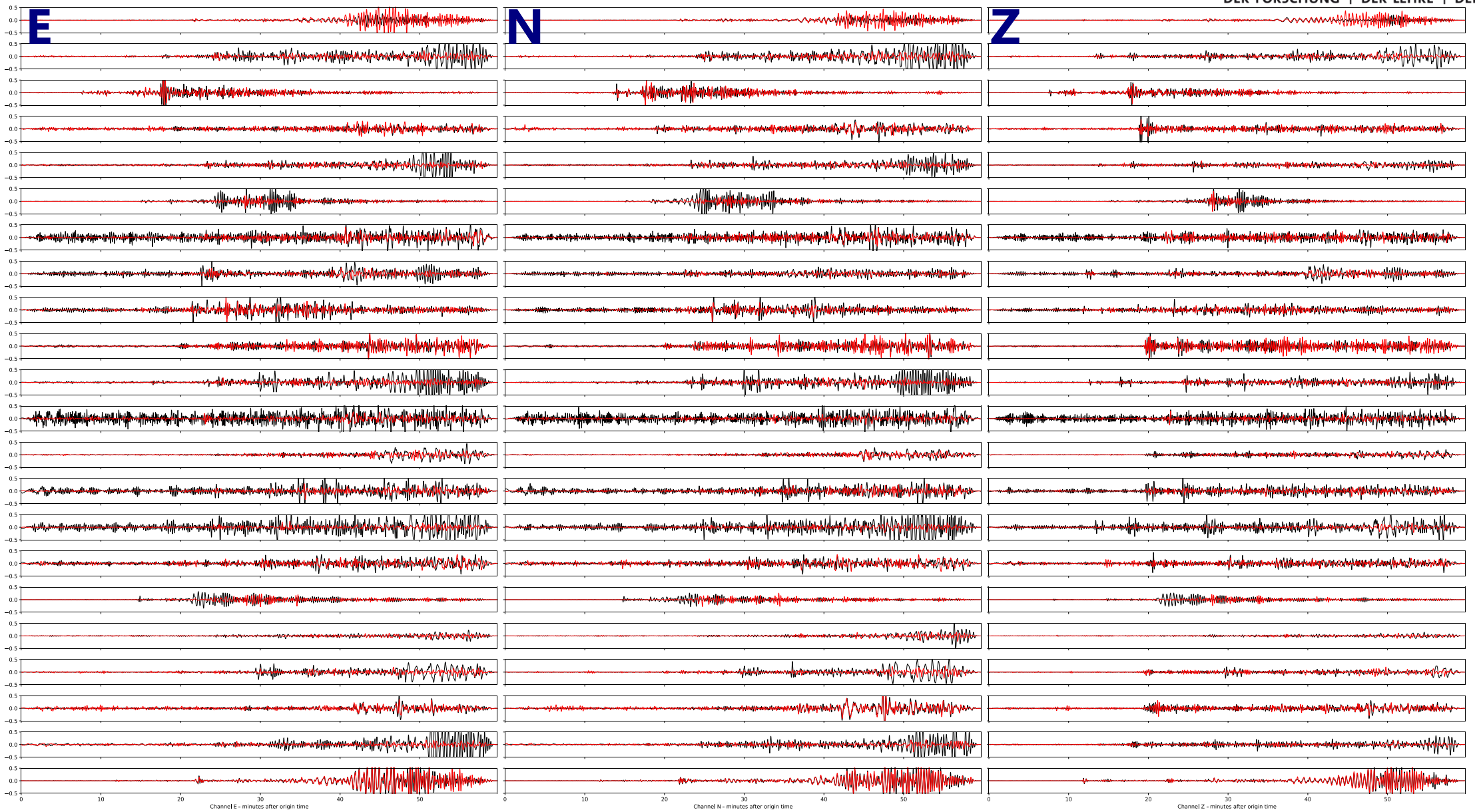


Application: unseen measured data (labels)

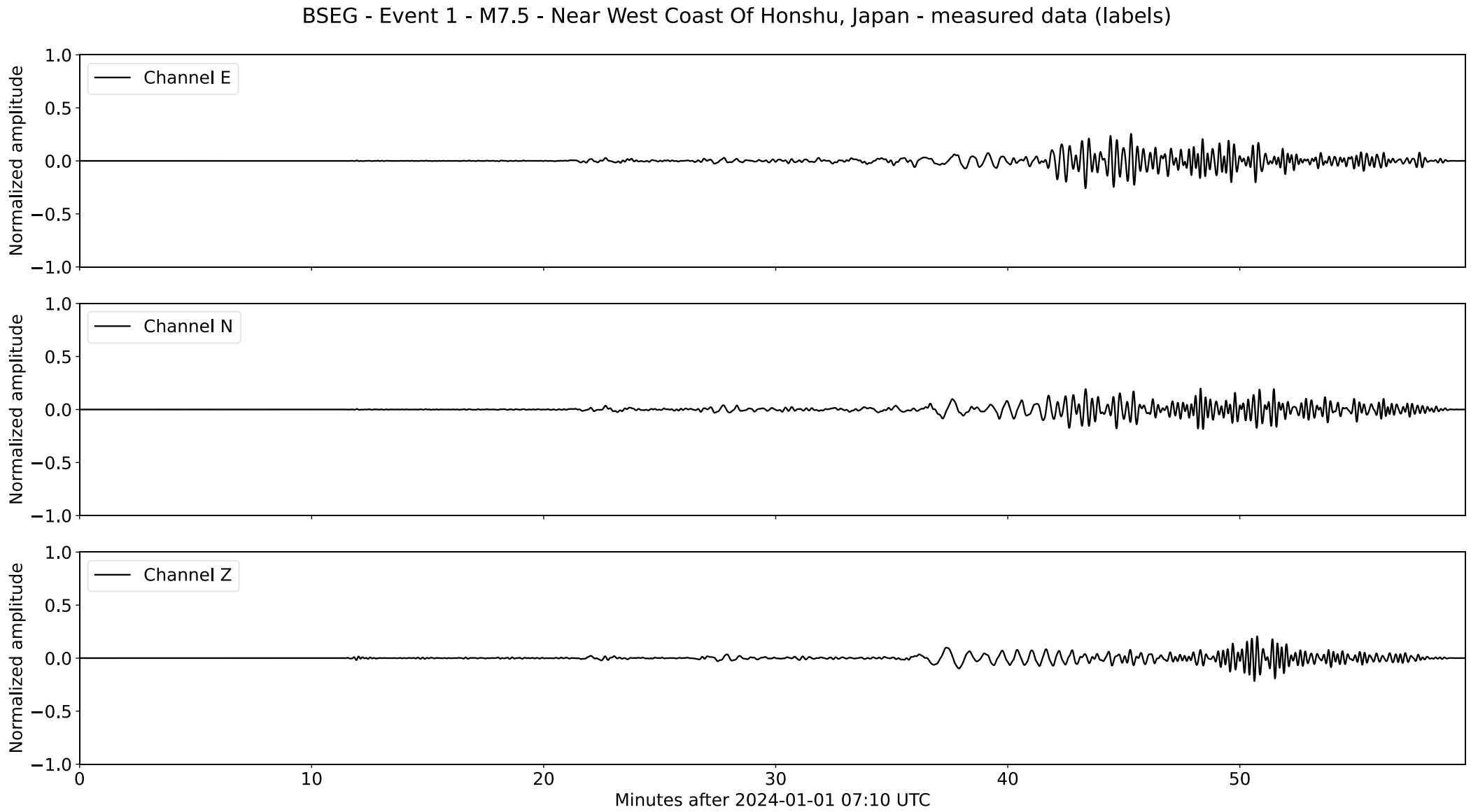


Application: predictions vs. labels

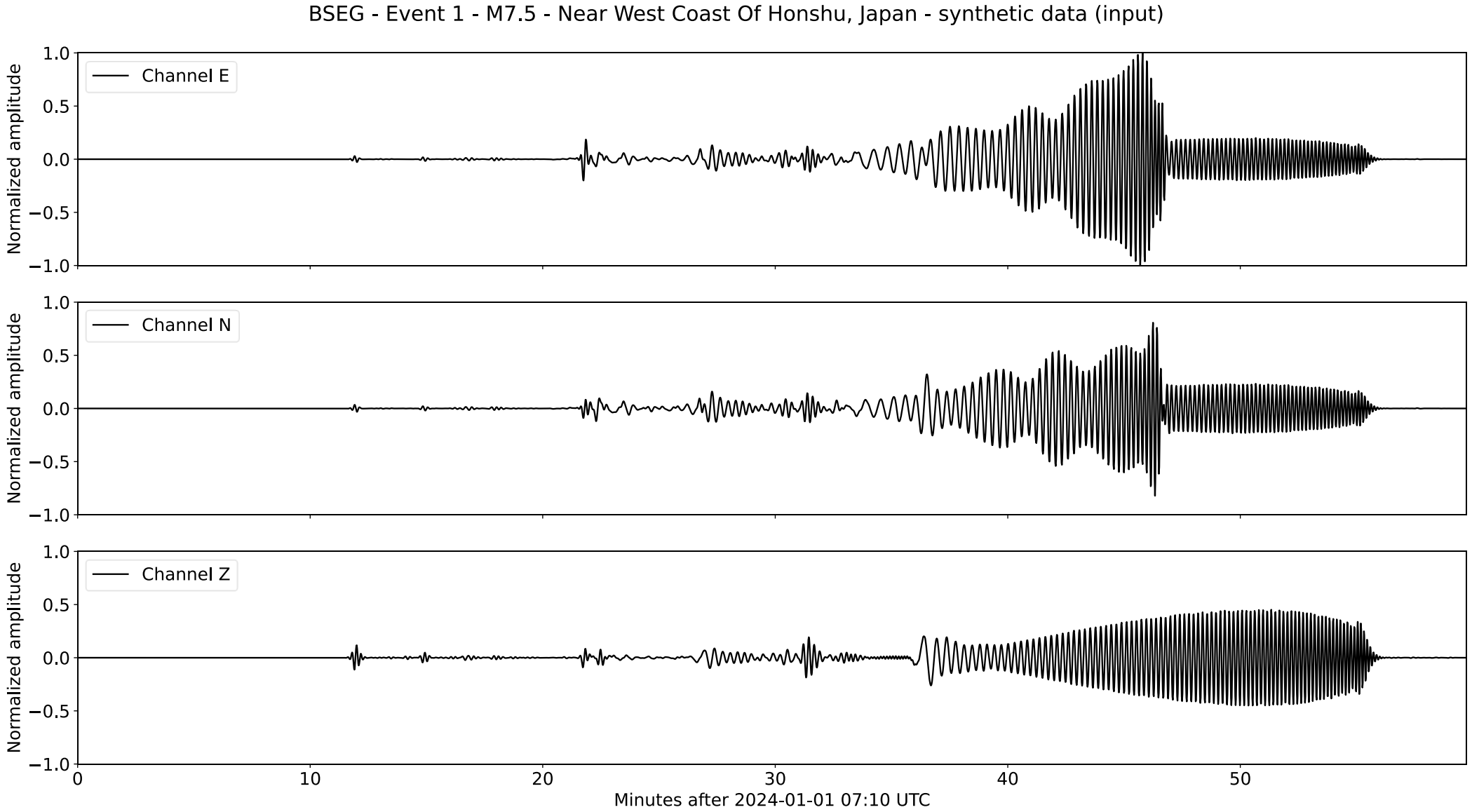
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



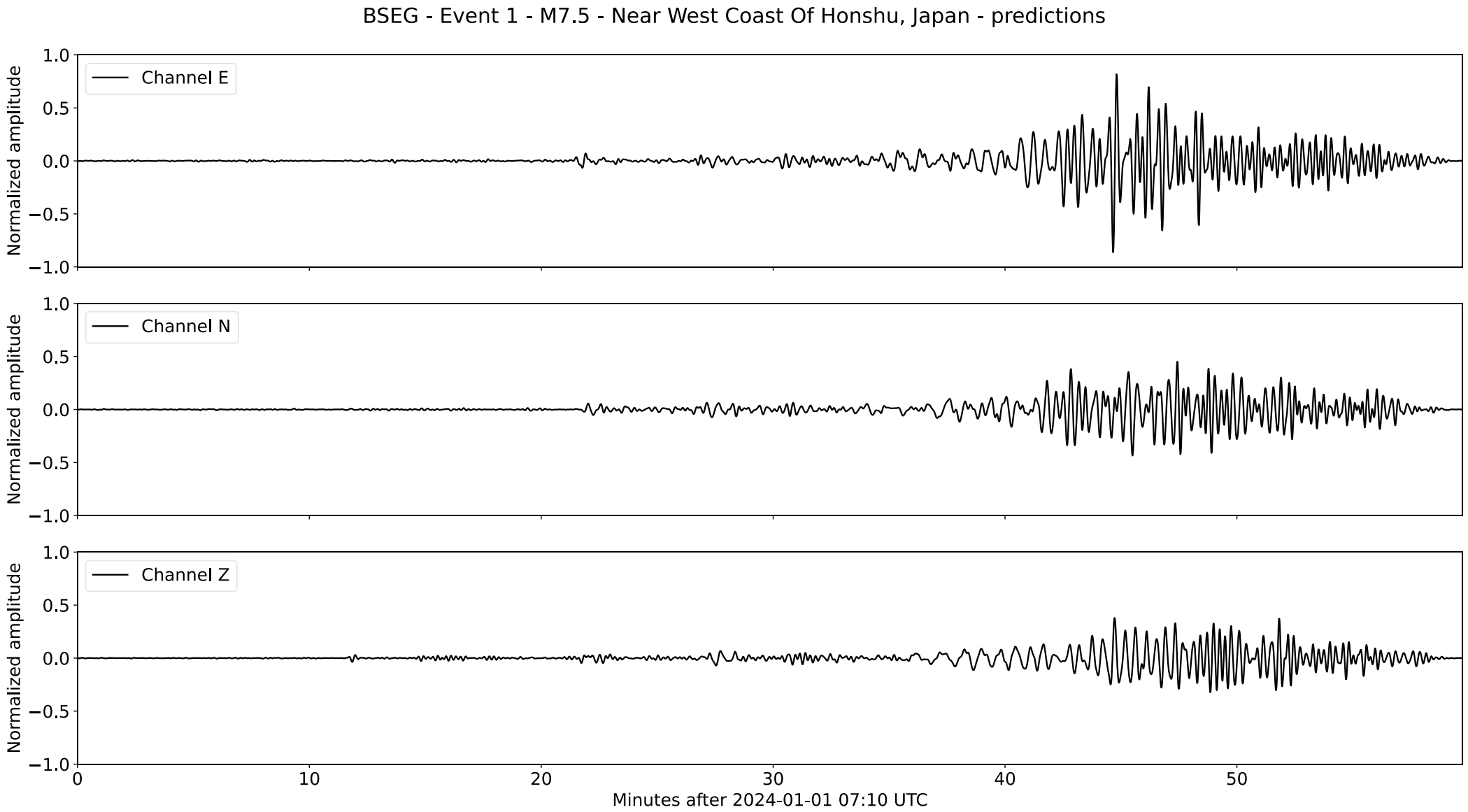
Application: Event 1 – unseen measured data



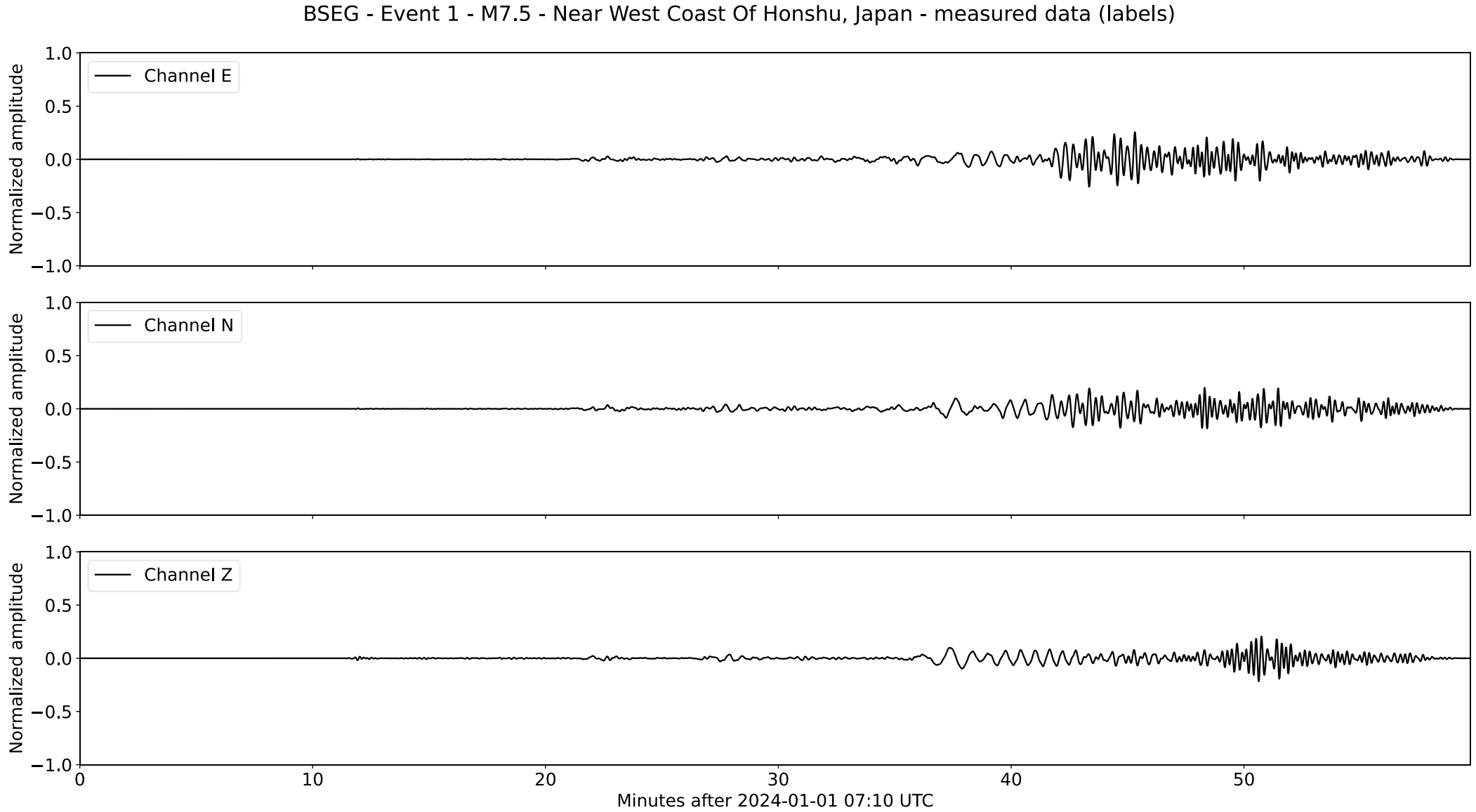
Application: Event 1 – synthetic data (input)



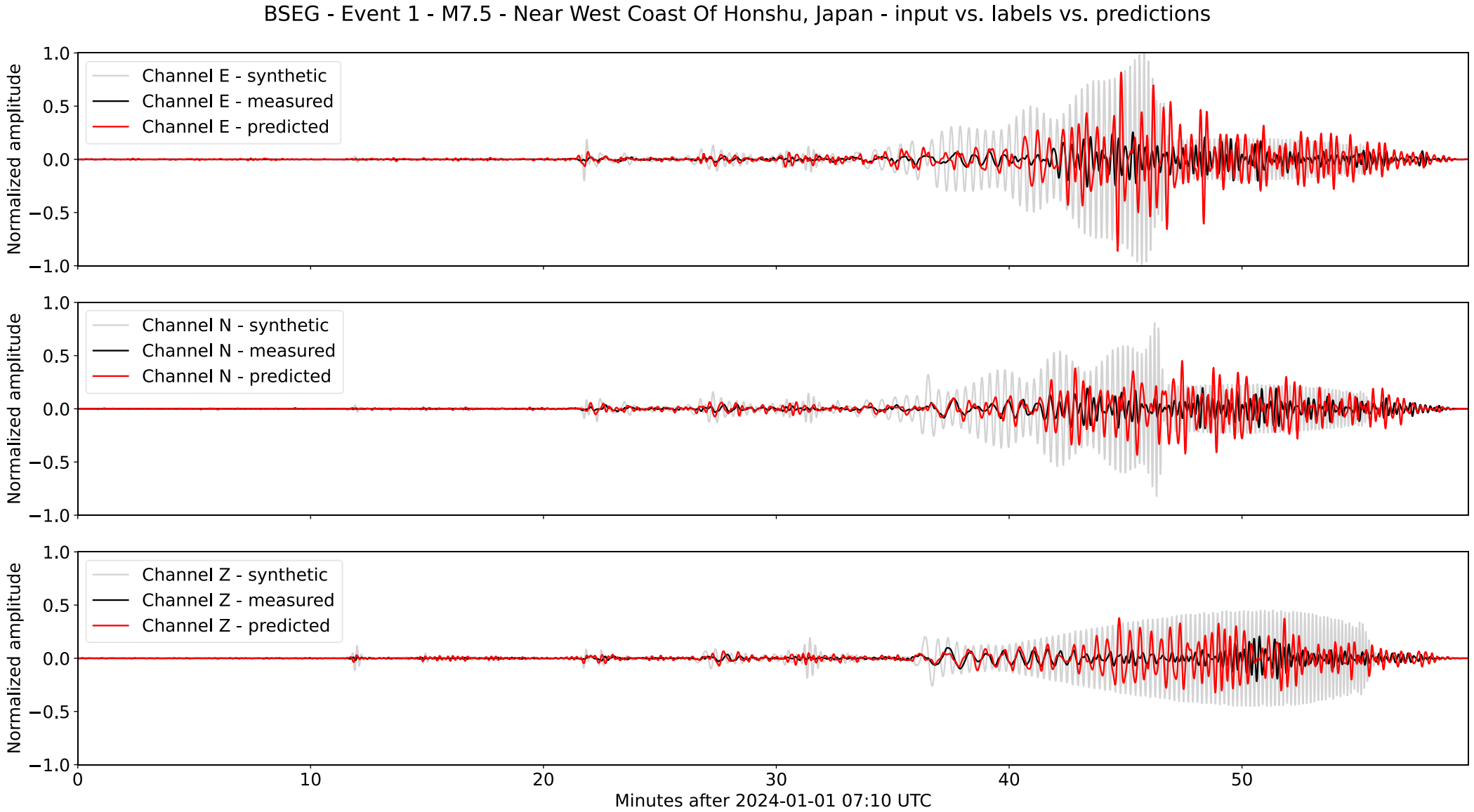
Application: Event 1 – predictions



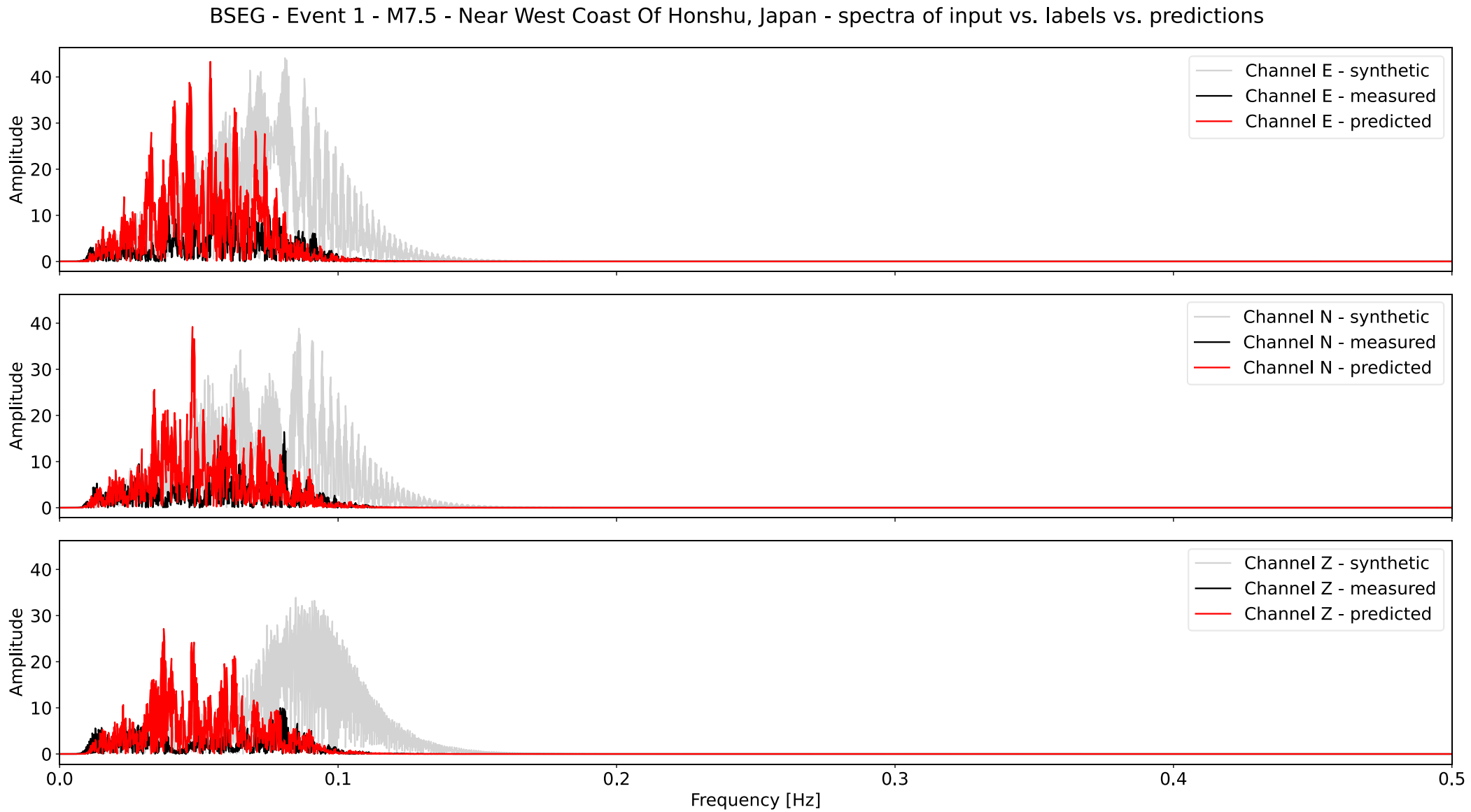
Application: Event 1 – measured data (labels)



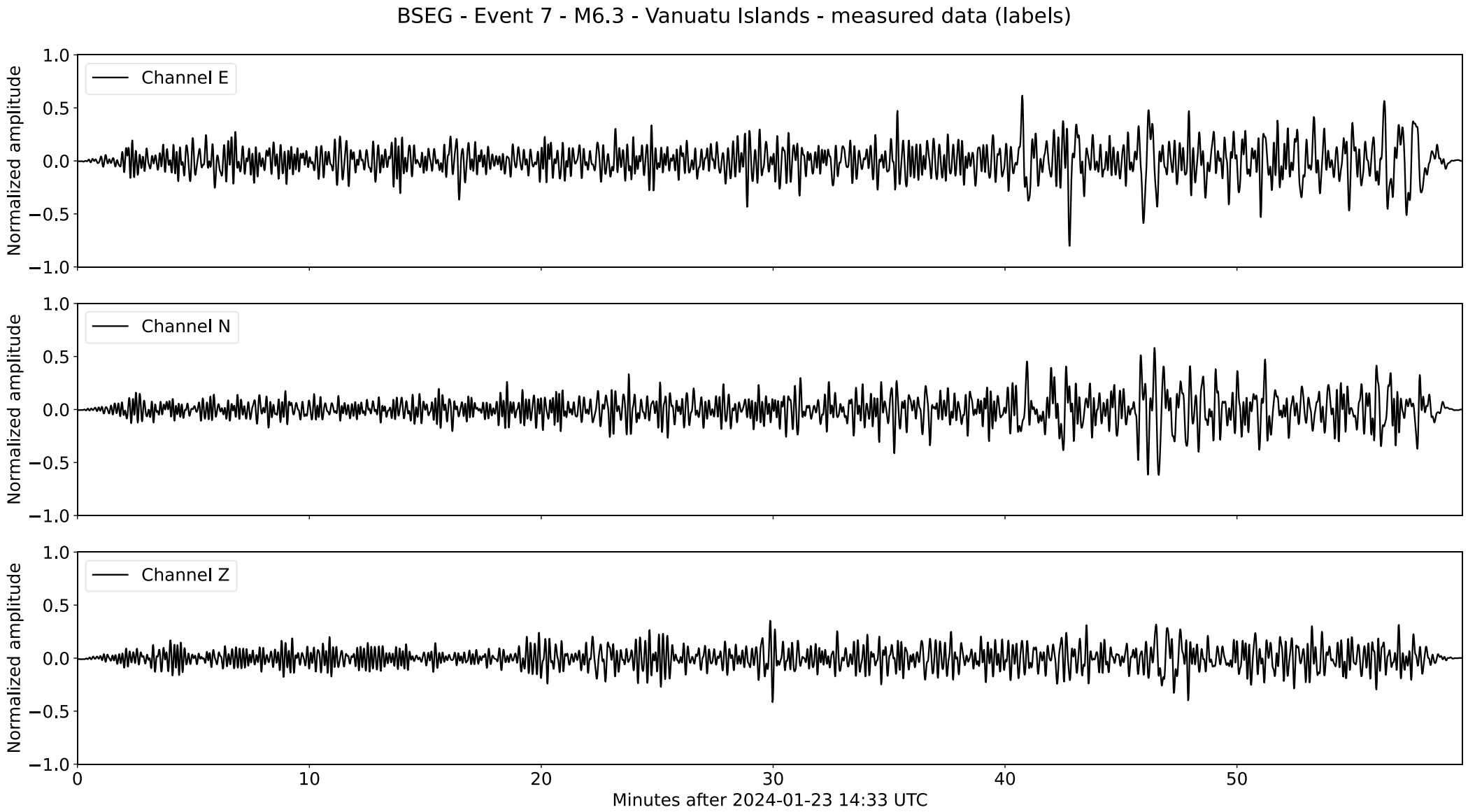
Application: Event 1 – overlay



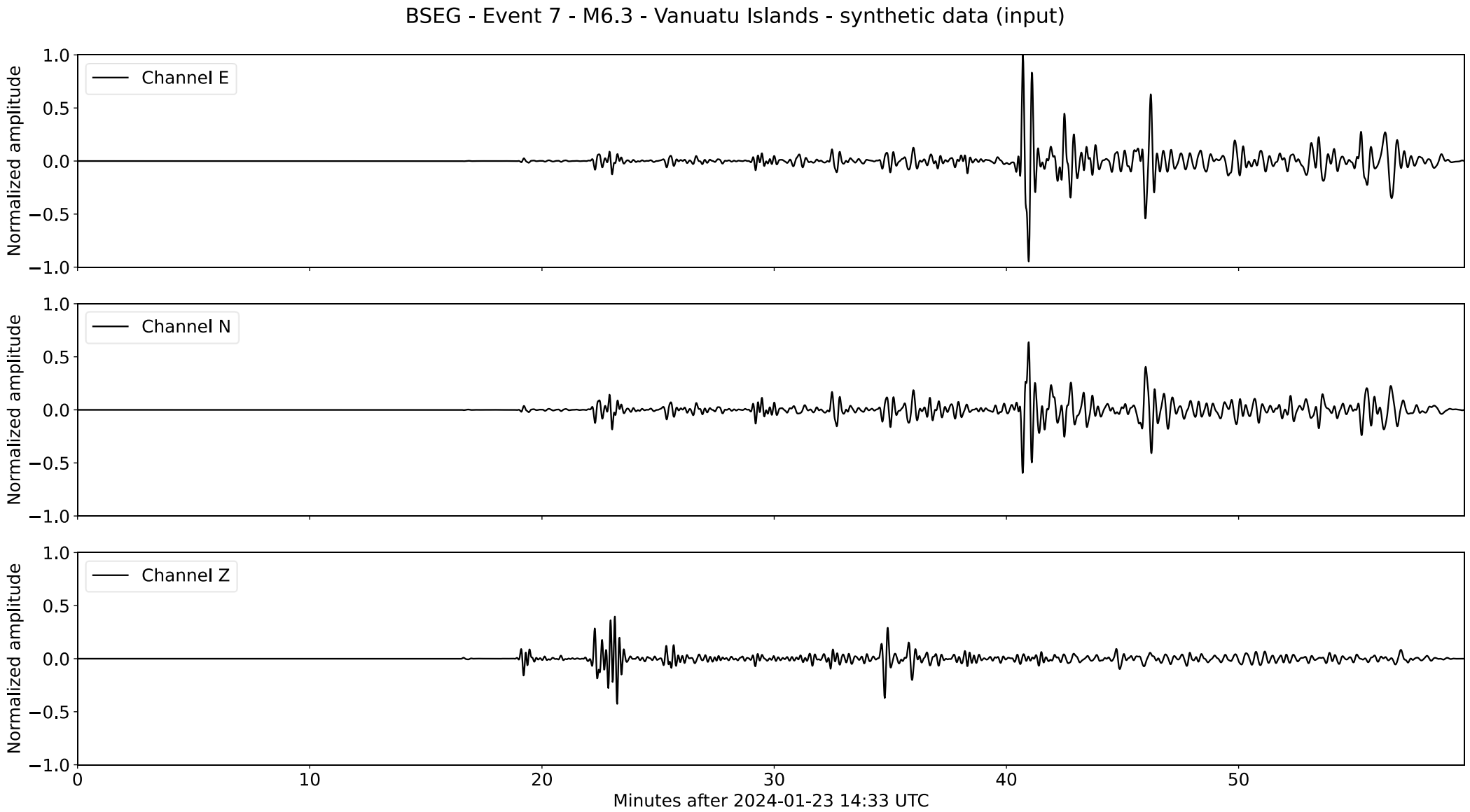
Application: Event 1 – spectra



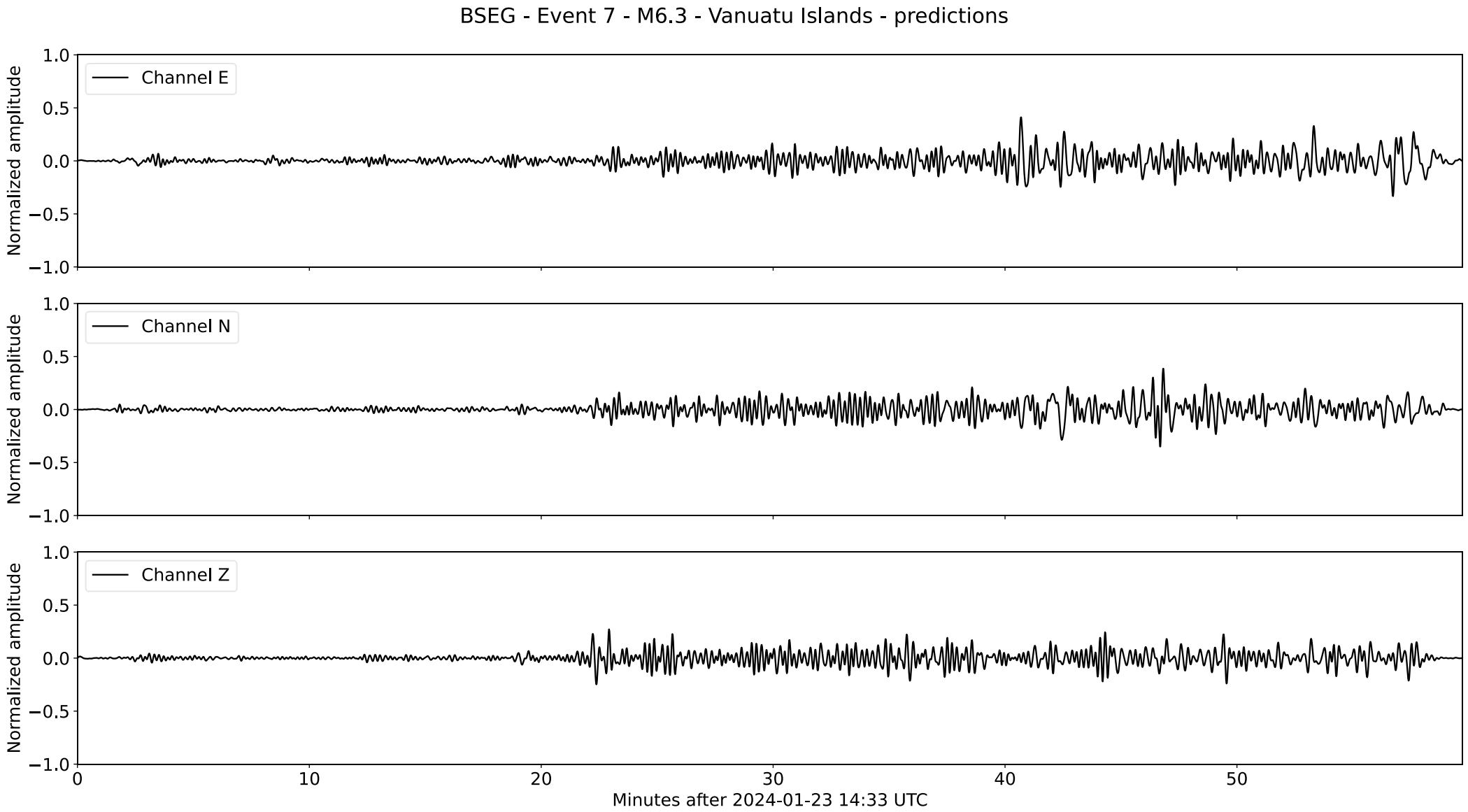
Application: Event 7 – unseen measured data



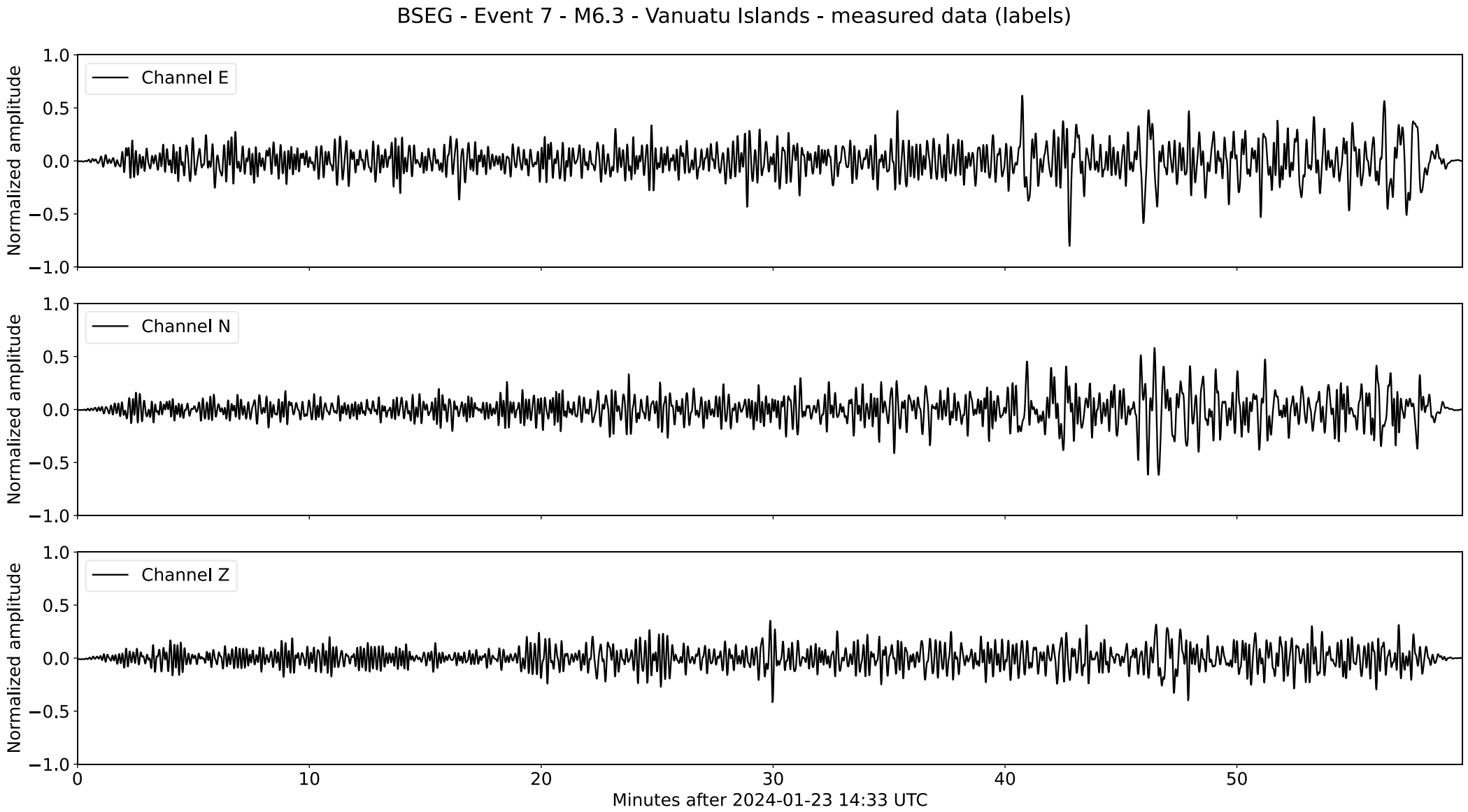
Application: Event 7 – synthetic data (input)



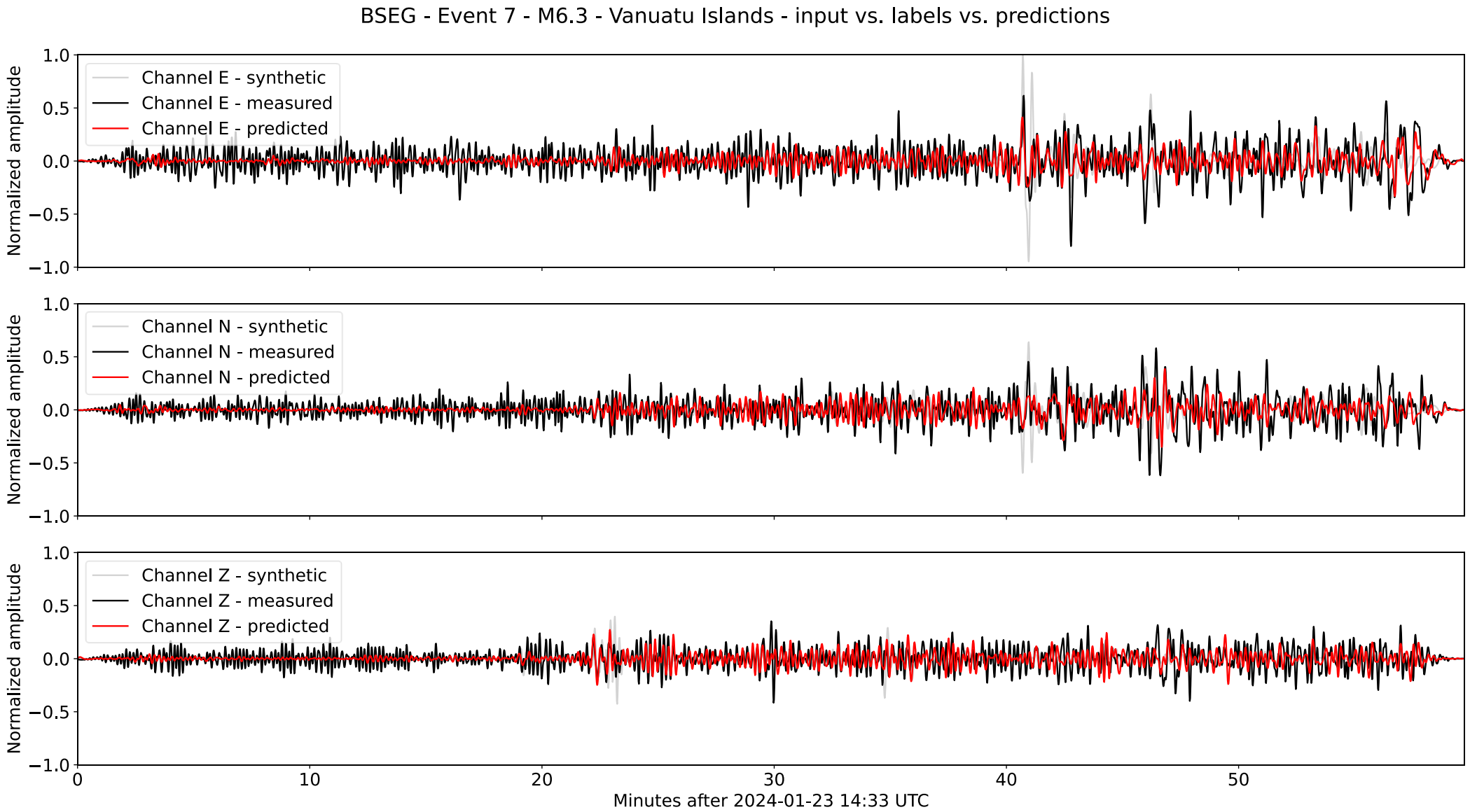
Application: Event 7 – predictions



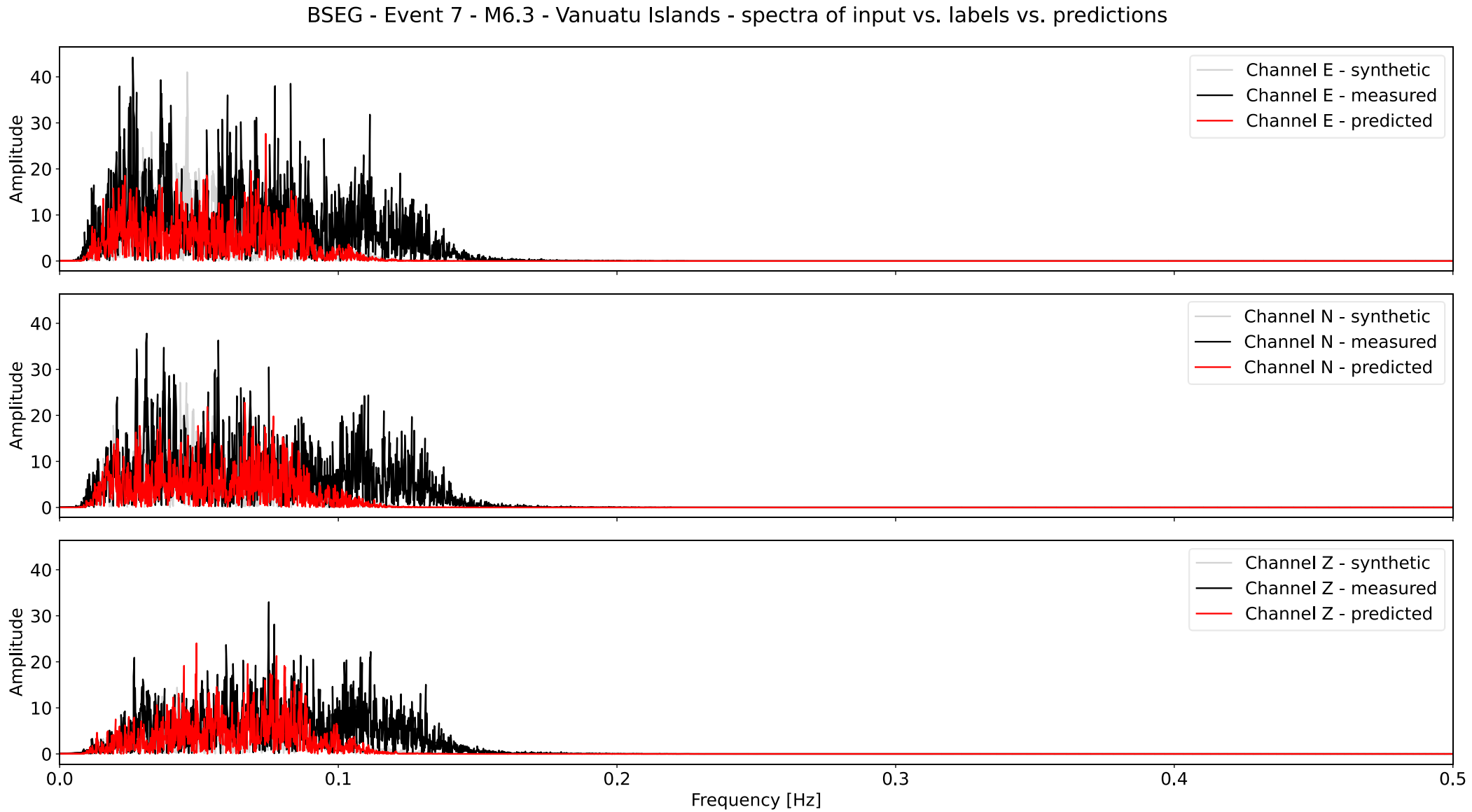
Application: Event 7 – measured data (labels)



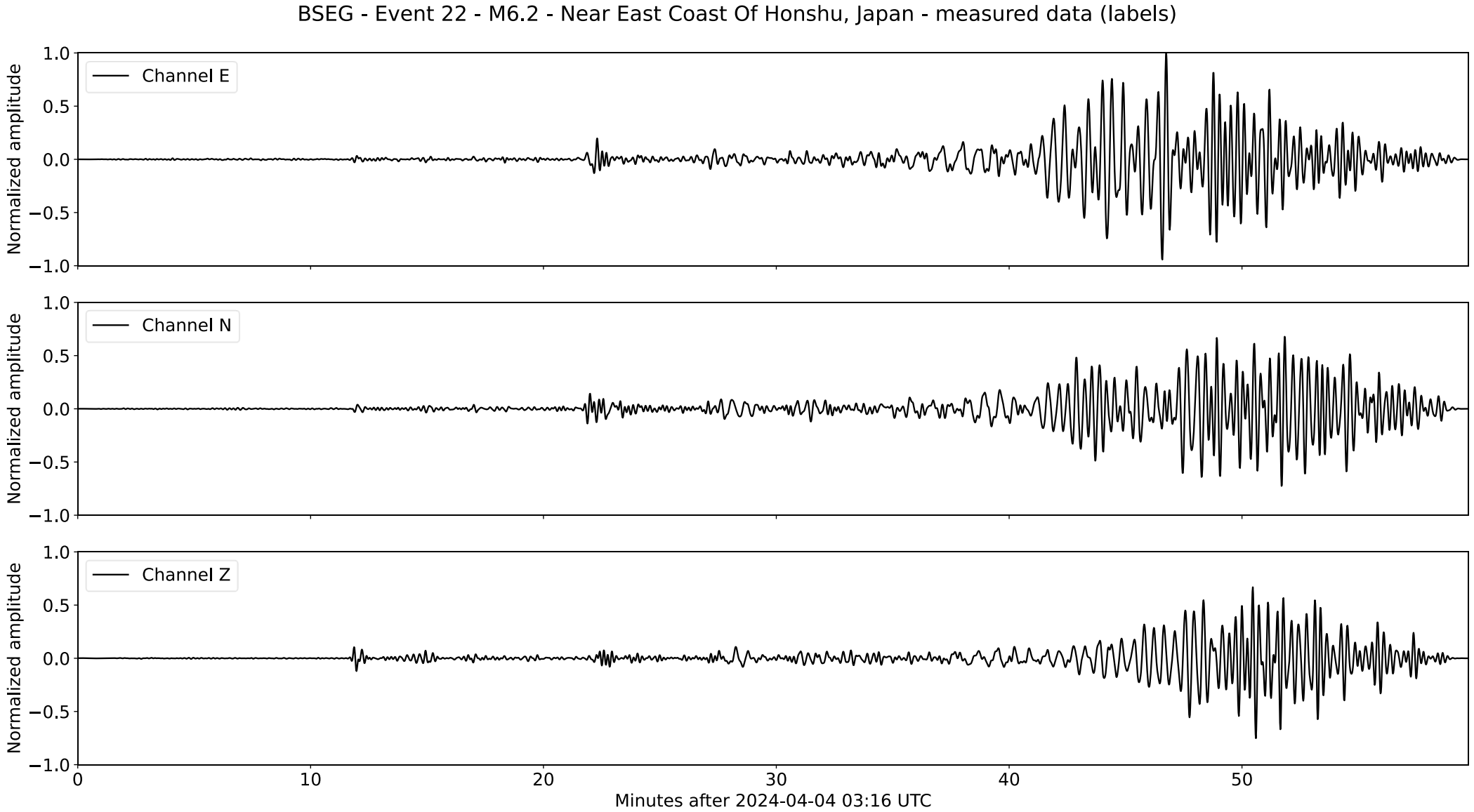
Application: Event 7 – overlay



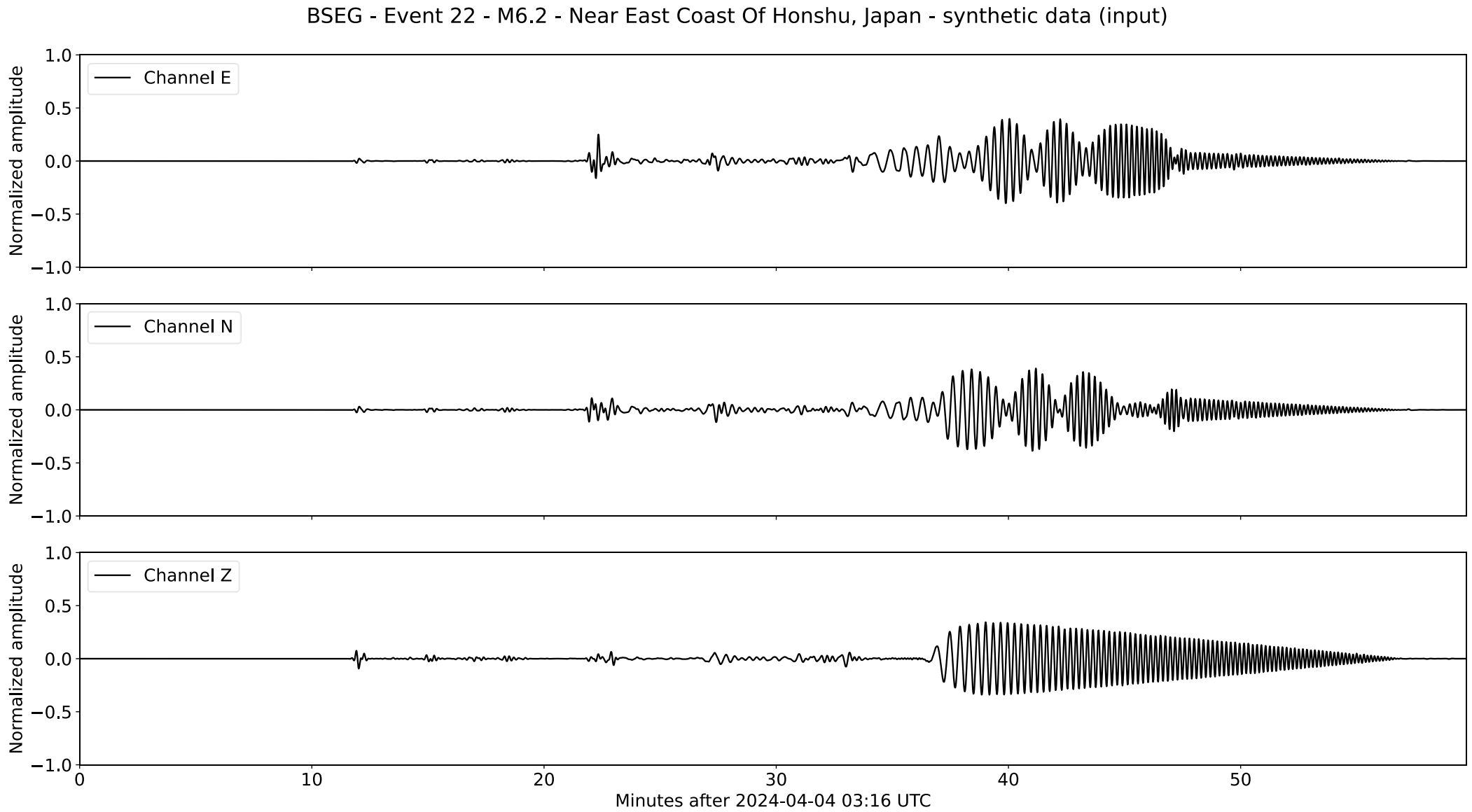
Application: Event 7 – spectra



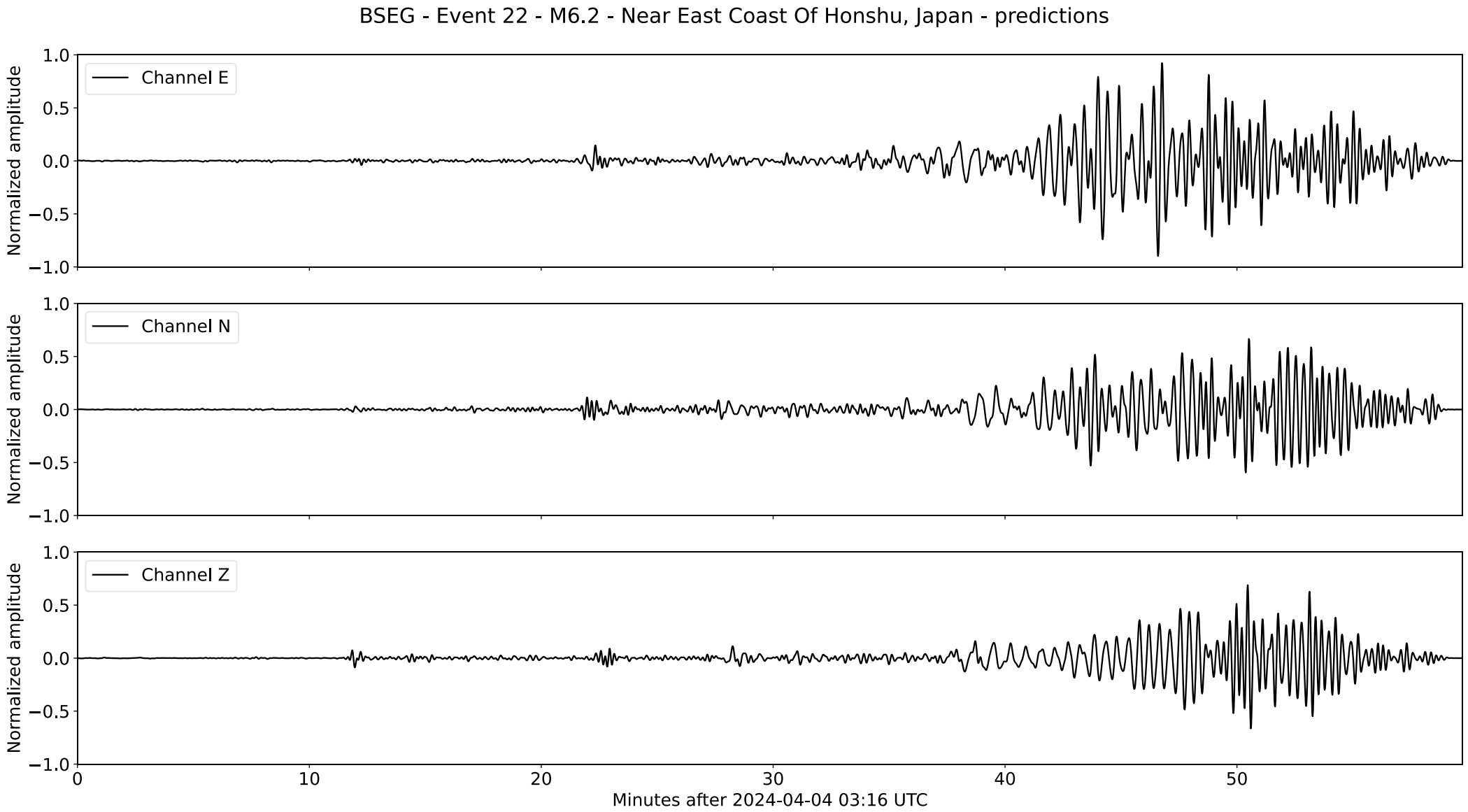
Application: Event 22 – unseen measured data



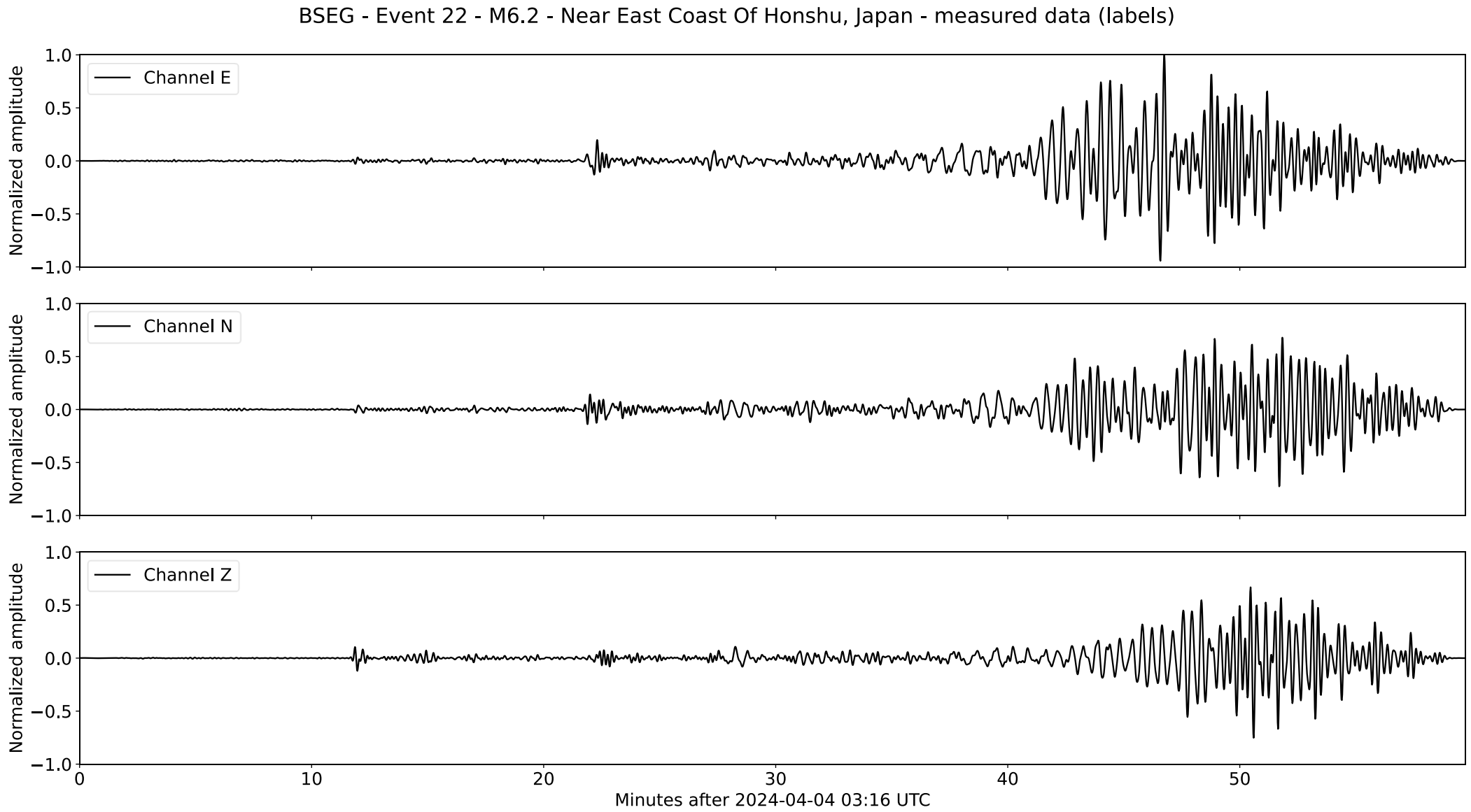
Application: Event 22 – synthetic data (input)



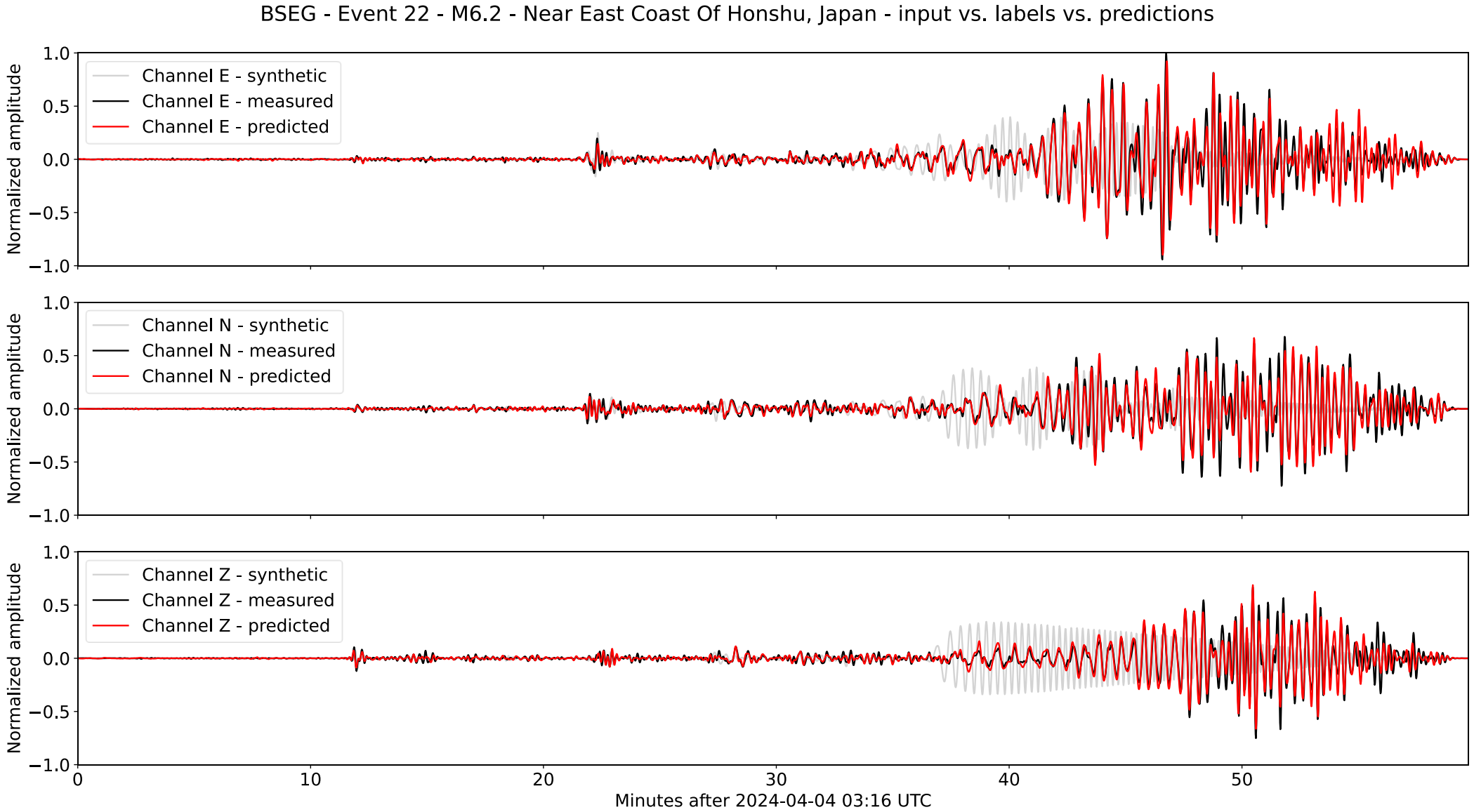
Application: Event 22 – predictions



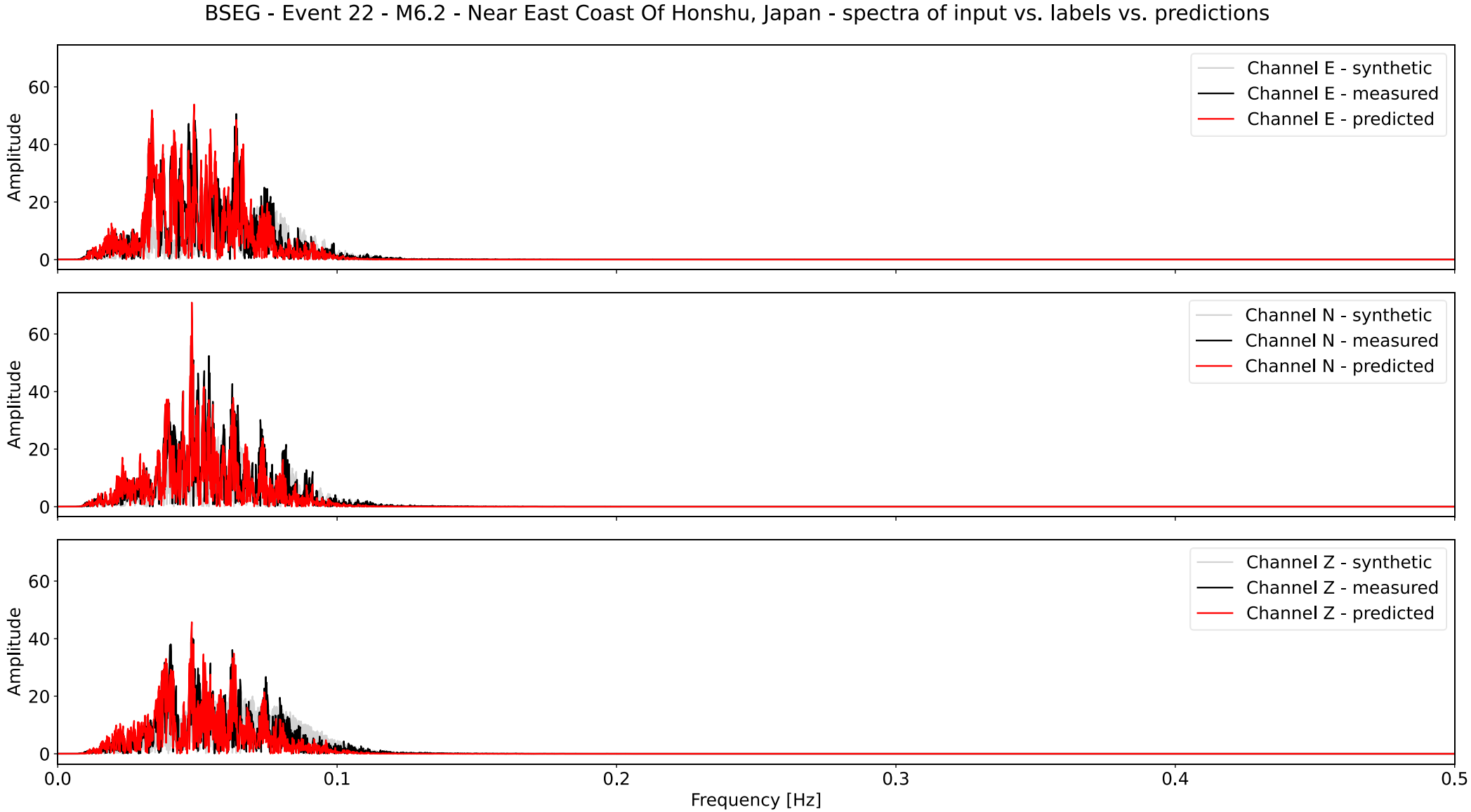
Application: Event 22 – measured data (labels)



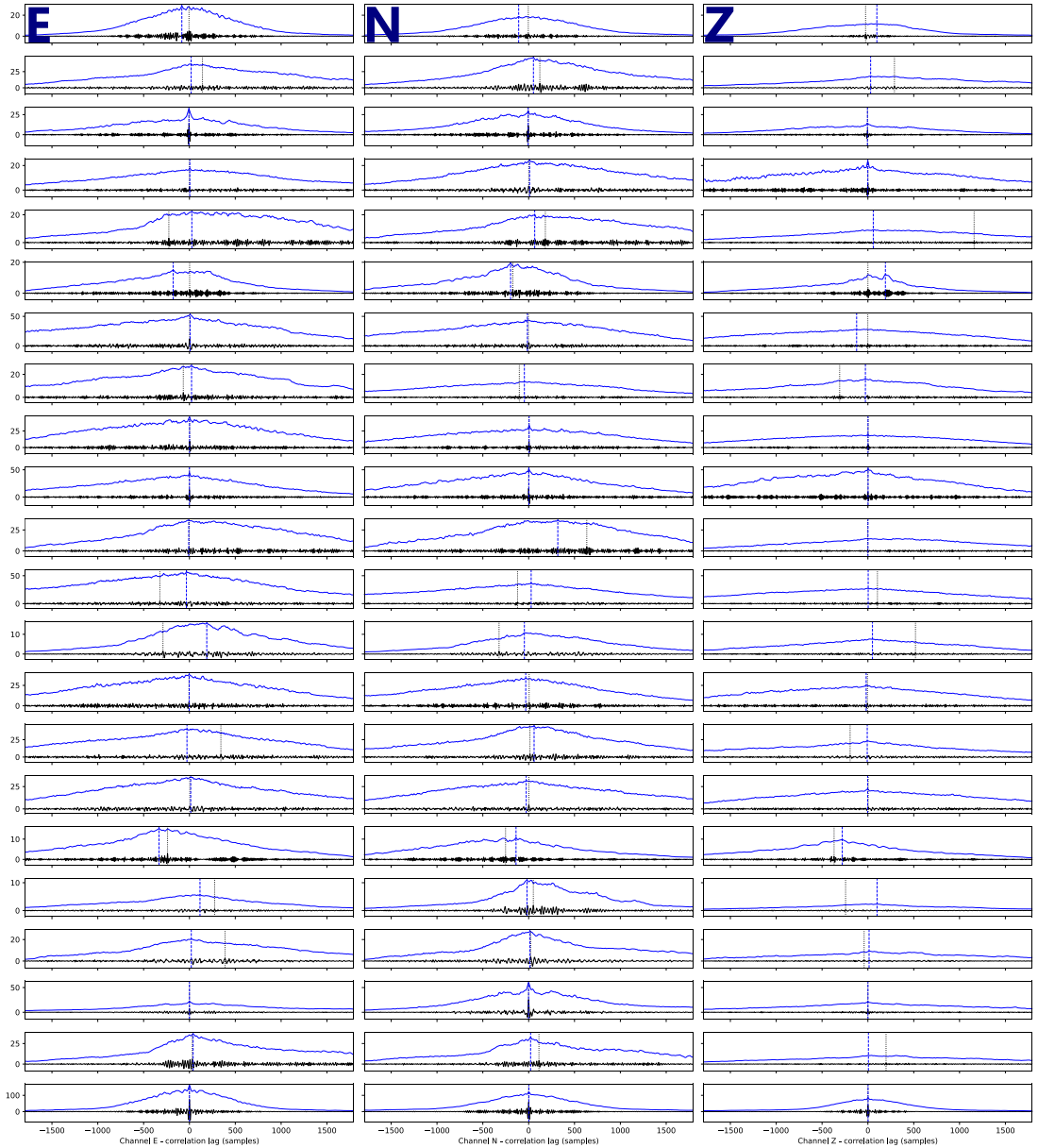
Application: Event 22 – overlay



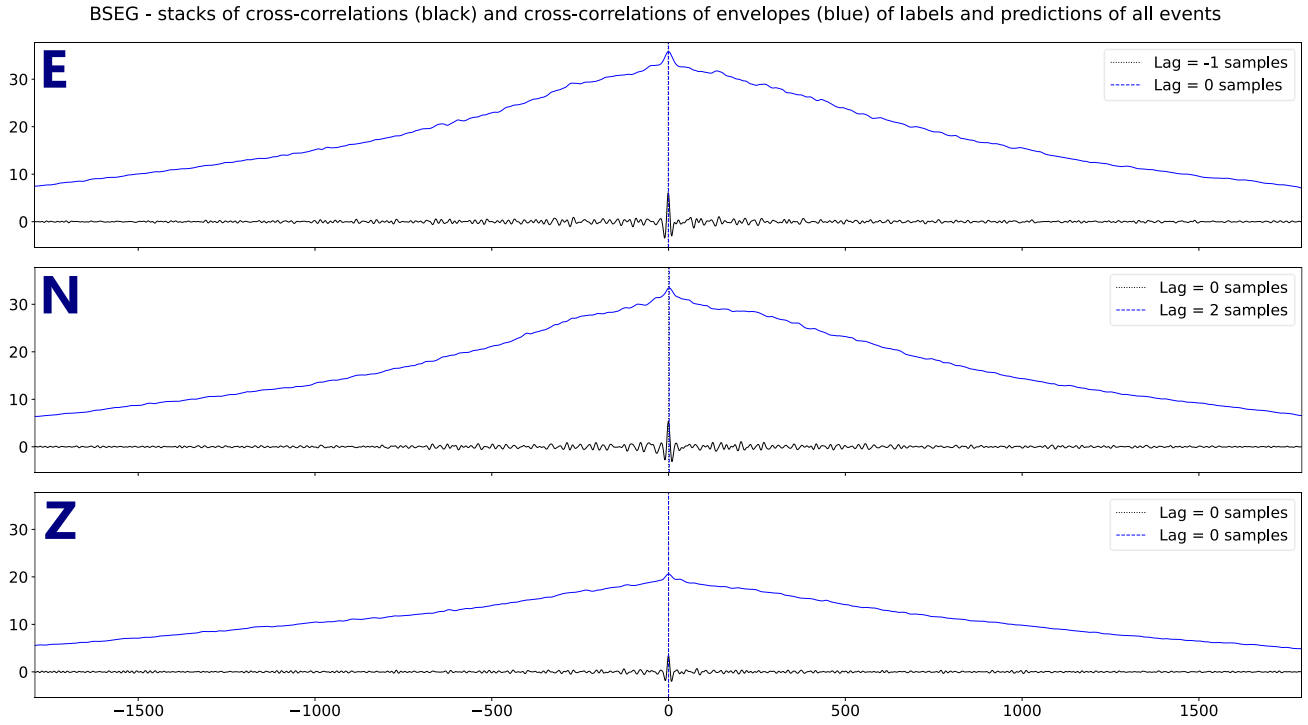
Application: Event 22 – spectra



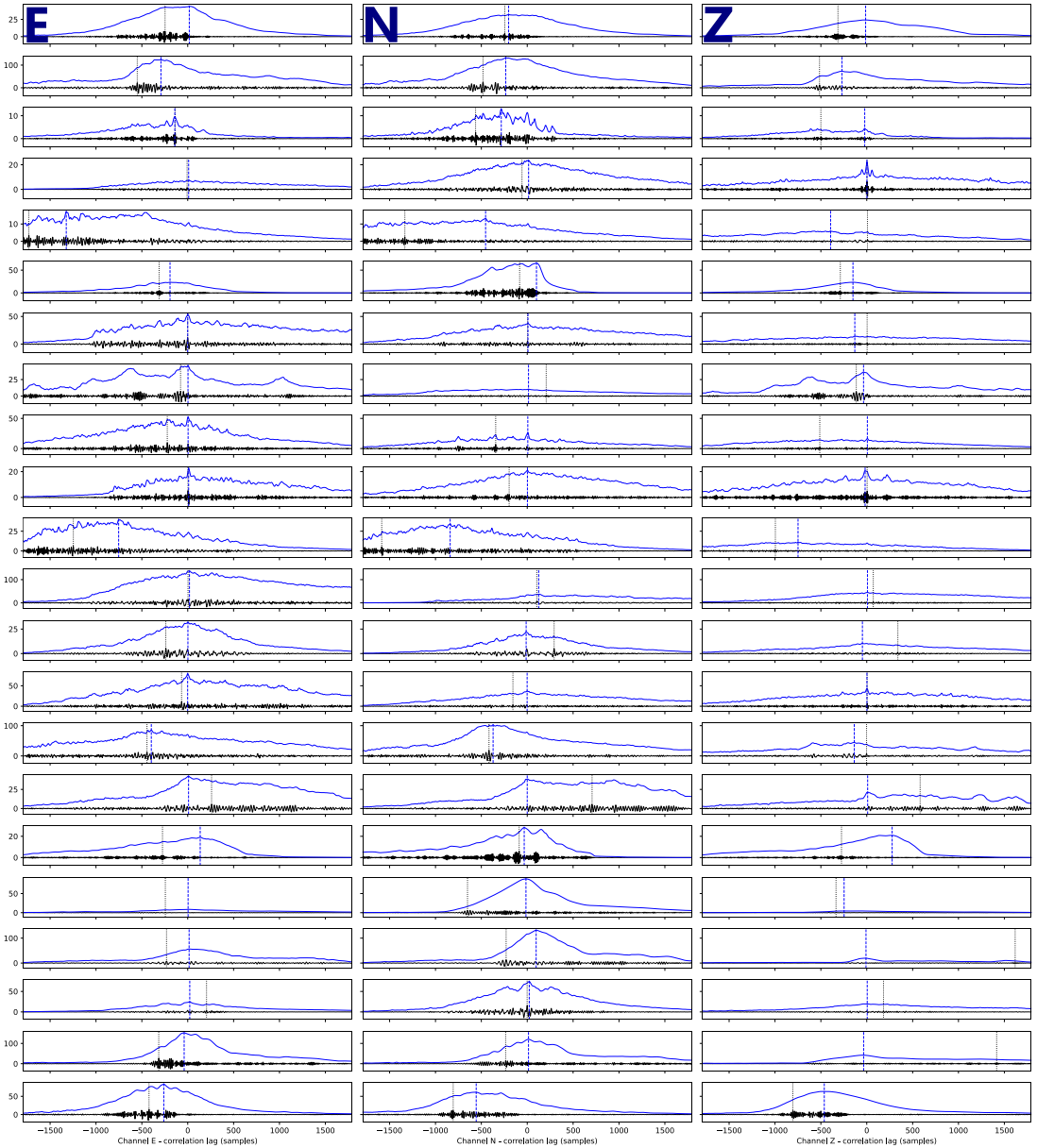
QC: labels vs. predictions



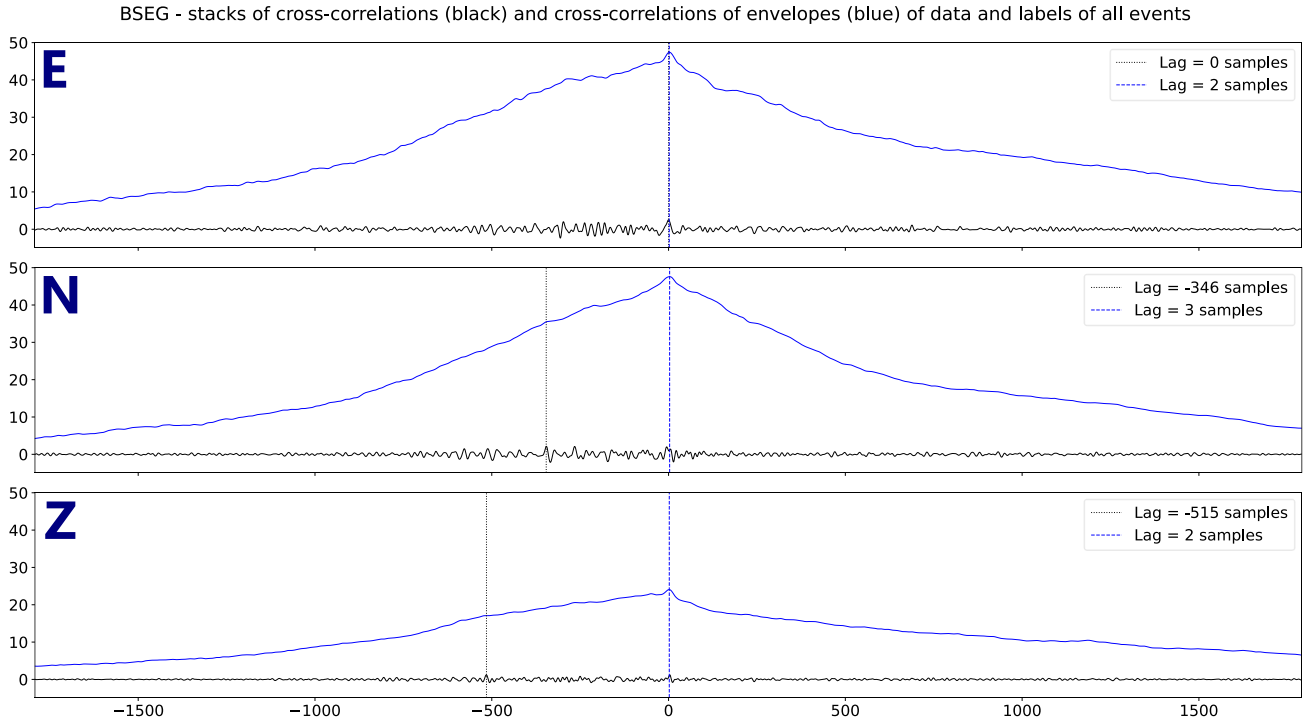
- Cross-correlation of labels and predictions
- Cross-correlation of envelopes
- Identification of maximum lag (ideally 0)



QC: labels vs. input data

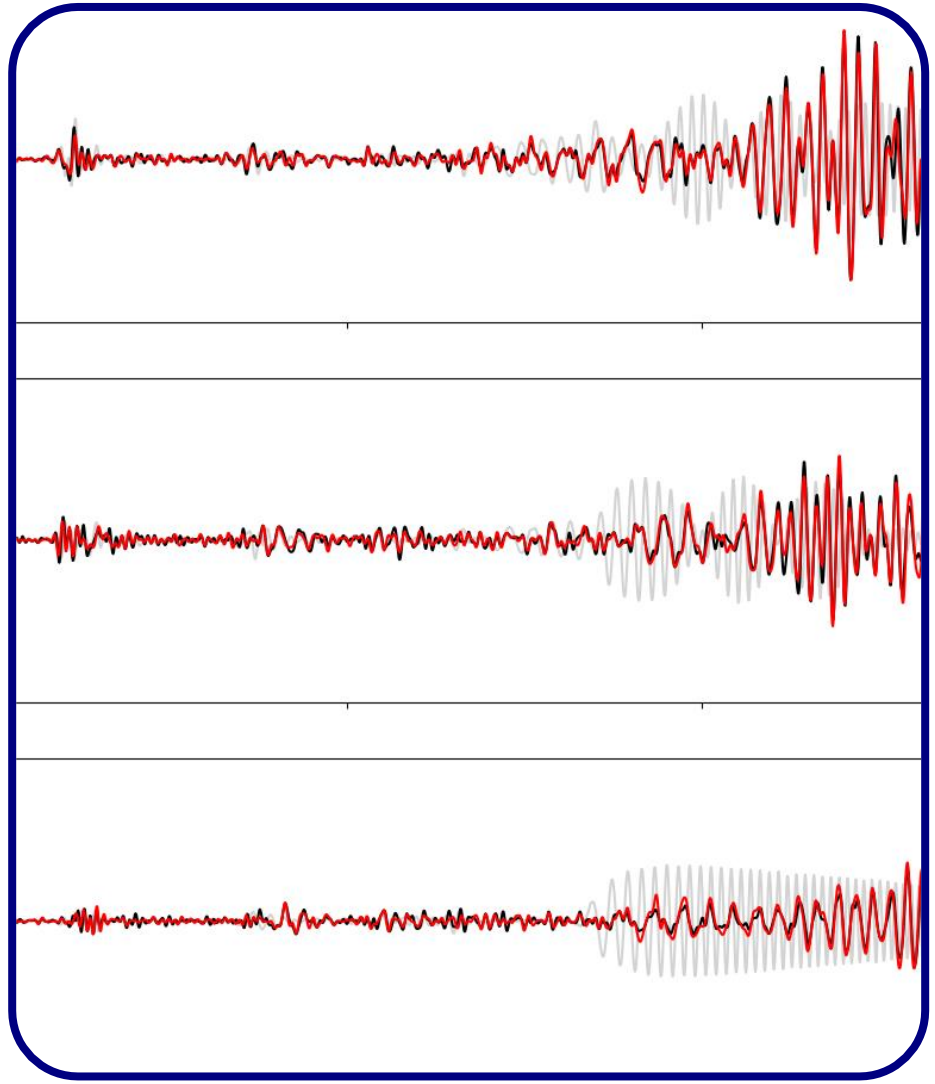


- Cross-correlation of labels and input
- Cross-correlation of envelopes
- Identification of maximum lag (ideally 0)



Conclusions

- Training of a **convolutional autoencoder** to predict earthquake waveforms from synthetic data
- The trained CNN is largely able to **predict** prominent phases of **unseen earthquake waveforms**
- Quality of results depends on **number of measured earthquakes**
- QC by **stacks of cross-correlations** of labels and predictions
- **Next steps**:
  - Combine data of various stations
  - Improve quality of synthetic data



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Thank you for your attention!

✉ alex.bauer@uni-hamburg.de

Acknowledgments

- Federal Ministry of Education and Research of Germany (BMBF, 05D23GU5)
- IRIS, SAGE
- Instaseis / Syngine
- TensorFlow 2
- Stefan Knispel (University of Hamburg), TEEC GmbH