

Smart data selection – Using machine learning for RMT data processing

A. Platz¹, U. Weckmann^{1,2}

¹Helmholtz-Zentrum Potsdam, Deutsches GeoForschungsZentrum, Potsdam, Germany,
aplatz@gfz-potsdam.de

²University of Potsdam, Institute of Geosciences, Potsdam, Germany

SUMMARY

The Radio-Magnetotelluric (RMT) method is a geophysical near-surface imaging technique with a broad range of possible applications. The GFZ Potsdam has recently acquired a newly developed horizontal magnetic dipole transmitter that allows the application of the RMT method even in regions with an insufficient coverage of radio transmitters serving as source signal. First measurements were conducted at three different locations in Chile in 2020. We have full control over the subsequent data processing by storing the raw time series. Processing the entire time series has a couple of drawbacks: The different emitted frequencies are transmitted in a sweep implying that only a smaller fraction of the time series contains the required signal for a particular target frequency and leading to an unfavourable signal-to-noise ratio. However, since it is technically impossible to have the same time base for the data logger and the transmitter with an accuracy of a few nanoseconds, an automated detection scheme is required to find time segments that contain the transmitter signal. Usually, several Gigabytes of raw time series are collected during field measurements, making manual editing and supervision of the time series virtually impossible. However, a careful selection of appropriate time segments is essential for the success of the data processing. Since the underlying physics are comparable to those of low-frequency MT, the same tools for statistical time series analysis and data processing can be used with partly significant modifications. The processing tools at GFZ include the modular processing suite EMERALD, which was originally designed for MT processing, but has recently been quickly adapted to process RMT data. One main difference is that in RMT the transmitter data is considered as signal, while in natural source MT this would be regarded as (near-field) electromagnetic noise that needs to be removed using automated robust statistical approaches. To address the challenge, machine learning algorithms have a high potential to solve both problems. Initial experience was gained with a recurrent neural network approach in order to identify suitable time segments (Patzer & Weckmann, EMTF 2021 – conference contribution and personal communication). However, many questions remained open, e.g. if other machine learning algorithms can result in better performances, which machine learning algorithms are in principle suitable for the characteristics and properties of RMT time series or which parameters should be used as input variables for the algorithms. A large number of machine learning algorithms exist, which can be divided into different groups according to their operating principle and their activity fields. We will test unsupervised methods, especially for clustering the data, as well as supervised methods as e.g. logistic regression, support vector machine or neural networks. We will present first results using the RMT data from Chile as training data set.

Keywords: Radio-Magnetotellurics (RMT), Machine learning, Neural networks