Journal of Physics:
Conference Series

PURPOSE-LED PUBLISHING™

PAPER • OPEN ACCESS

Magnetotelluric Time Series Denoising Using Encoder-Decoder Consisted of LSTM Cells

To cite this article: Sihao Wang *et al* 2023 *J. Phys.: Conf. Ser.* **2651** 012123

View the article online for updates and enhancements.

You may also like

- Resistivity Distribution of Lembang Fault Based on Magnetotelluric Data
Nurhasan, Muhammad Risyad Naufal, Wahyu Srigutomo et al.
- Analysis of the effect of magnetotelluric data quality improvement using rho variance and edit XPR parameters in densely populated areas
A Widodo, W Lestari, D D Warnana et al.
- Determination of Geoelectrical Strike To Investigate The Geological Structures In Surabaya And Gresik Areas
W Lestari, A Widodo, D D Warnana et al.

This content was downloaded from IP address 173.177.195.252 on 26/02/2026 at 21:15

Magnetotelluric Time Series Denoising Using Encoder-Decoder Consisted of LSTM Cells

Sihao Wang[1,2,3], Lanfang He[1,2,3*], Xuben Wang[4], Liang Li[1,2,3]

[1] Key Laboratory of Mineral Resources, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China
[2] College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[3] Innovation Academy for Earth Science, Chinese Academy of Sciences, Beijing, China
[4] State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu 610059, China

*E-mail:mofoo@263.net

Abstract. Electromagnetic signals in geophysics are frequently disturbed by various interference in field data acquisition. Denoising for passive electromagnetic methods such as magnetotelluric (MT) or audio magnetotelluric (AMT) data is significant to improve data quality and finally imaging to the geoelectrical structure. Conventionally, most denoising methods are employed in frequency domain and few of them are applied in time domain. However, a great number of irregular noise in the electromagnetic time series prove difficulty to be removed. We propose a denoising method, using Encoder-Decoder consisted of Long Short-Term Memory cells (ED-LSTM), to reduce the effect of the step noise and the random-impulsive noise. Supervised learning and transductive learning are used for the denoising of the step noise and the random-impulsive noise, respectively. Our results indicate that the step and random-impulsive noise could be successfully removed from the raw time series. The result indicate that ED-LSTM could potentially to be wildly used in the electromagnetic time series denoising and then improve data quality.
Keywords: LSTM, Magnetotelluric, Denoising, Encoder-Decoder, Time Series

1. Introduction

Magnetotelluric sounding (MT) are passive electromagnetic exploration methods to image the subsurface electrical conductivity structure. More and more MT raw data contain various noise, this comes to that MT data denoising is crucial to obtain the reliable estimation of relationship between multiple time series [1]. Approaches for MT data denoising are always applied in frequency and time domain [2]. In frequency domain, robust statistics and remote reference are widely used to obtain the reliable impedance tensor as well as the transfer functions. Differing from least-squares method, robust method uses the adaptive weights which is proportional to reciprocal of variance for each measurement and it can weight source contaminated by outliers down [3]. Remote-reference method requires at least one additional remote correlated noise-free site to overcome the noise which is present in one of the components of electric and magnetic fields [4]. It is hard to find a suitable reference site which is totally free from cultural interference, the remote reference method not always work in many cases [5]. Besides, it does not work while both electric and magnetic fields contains cultural interference.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

Neural network can establish the mappings that is based on the dataset without any priori information, and comes to be one of the major geophysical data processing approaches [6]. Neural network can robustly distinguish the noise from MT data, some studies have demonstrated the feasibility of MT denoising with neural network, the results indicate that the performance of neural network is consistent with expert selection in the case of low signal-to-noise ratio [7]. According to the statistical characteristics of this method, if there is persistent artificial noise existing in time series, using unsupervised learning method could lead unreasonable results. As a modification of recurrent neural network, Long Short-Term Memory (LSTM) neural network can selectively extract the information contained in the previous time series. Moreover, it can effectively solve the problem of vanishing gradient which exist commonly in recurrent network. The LSTM structure allows network to forget useless information in time domain therefore the network can deal with long term time series easier. LSTM neural network has been used to structure an end-to-end denoiser to suppress stable-distributed noise and power frequency interference [8].

We present an encoder-decoder consisted of LSTM cells to process artificial step noise and impulsive interference in time domain. We then introduce the approaches of removing two kinds of noise by different ways of machine learning: supervised learning and transductive learning. Our result indicate that the encoder-decoder LSTM could potentially improve MT data quality.

2. Method

As a particular kind of recurrent neural network, LSTM neural network can efficiently solve the vanishing gradient problem comparing with the traditional RNN for dealing with long-term sequences [9]. The complete LSTM cell is consisted of four units: input unit, forget unit, cell unit, output unit. Differing from traditional recurrent neural network, LSTM depends cell unit to select what information to forget and what information to succeed. Input unit is responsible to select the information which will be transferred to cell unit in the current time step, and then forget unit controls the weight of cell unit from last time step, therefore it can decide the amount of information which will be transferred to the cell unit from current time step. In summary, LSTM introduce the forget unit to screen effective information existed in the data from the current time step. In other words, LSTM is learning how to forget during training phase.

We apply the Encoders and Decoders consisted of LSTM cells to eliminate the noise in AMT time series. For ideal step noise, we know its distribution exactly and supervised learning is suitable. For existed impulsive noise, we have no idea for the distribution of it and transductive learning is suitable. Here we define the magnetotelluric time series denoising problem as f[s] = M(g[n]), where f is the denoising signal we need, g[n] is normalised time series with noise, □ denotes the discrete time series, f[s] is a single point signal the model output where the spike exist, and M is an ideal model we desire. We denoise the g[n] by replacing the noisy signal(spike) with f[s] which is exported from the model. To obtain the dataset in source domain and target domain, we find all the spikes in original time series and intercept 64-points slices which are centered around every spike as unlabeled data in target domain. Then we structure labeled dataset in source domain by randomly selecting the time slices which are undisturbed and add spiked noise on them. The trained model is directly attempt to extend to the target domain. To ensure the robustness of model, we collect statistics of all the outliers as a distribution and extract the degree of outliers from it when structuring labeled dataset. The loss function we use is mean absolute error (MAE) to measure the loss of training dataset and validation dataset. LSTM are powerful to the sequences to sequences learning problems and sensitive to the causality between words [10]. The LSTM cells we used are expressed in [11]. The encoder and decoder is consisted of LSTM cells, for the complete model, the expressions can be written as:

$$\text{code} = \text{E}(\text{input}), \text{and output} = \text{D}(\text{code})$$

where E, D represent encoder and decoder respectively and inputs represent the noisy signal. The input represents the noisy signals for step and impulsive denoising model. The code vector represents the latent information of noisy signal. For step noise, we expect output would be ideal step noise and for impulsive noise we expect output would be denoising signal.

3. Result

3.1. Step noise

AMT data with four components (Ex, Ey, Hx, Hy) observed from northern Myanmar, which has almost no artificial interference in the raw signal according to analysis result of the spectrum, are used for as noise-free data. We slip the time series into 1000 points per sample and apply stateful LSTM cells during the training phase. To ensure continuity of output from the denoising model, the cell state instead of initialization are saved for every batch of the dataset during training phase. MAE loss function and Adam algorithm has been used for training the denoising model. We set 20 LSTM units as encoder and 20 as decoder. The learning rate has been set as 0.001 to ensure the model can converge steadily. We use data from one station to train the denoising model. The synthetic step noise with uniformly distributed signal-to-noise ratio ranges from 1 to 10, and the system only show the case with the minimum signal-to-noise signal (Figure 1a). A section of whole time series which has 246000 sample points is processed (Figure 1b). Figure 1c shows the time series of the raw data Hx(t) from the station. Figure 2. shows comparison of the spectrum of the Hx(t) data containing artificial step noise (a) and denoising spectrum (b) and the original raw spectrum (c). The data with the step noise shows a great number of spikes exist in 1-100Hz and the denseness of them increases with increasing frequencies. The amplitude of step noise is far greater than the raw data and it embodies in wide variation between spikes and raw signal in frequency domain. The LSTM denoising have removed the step noise in time series.

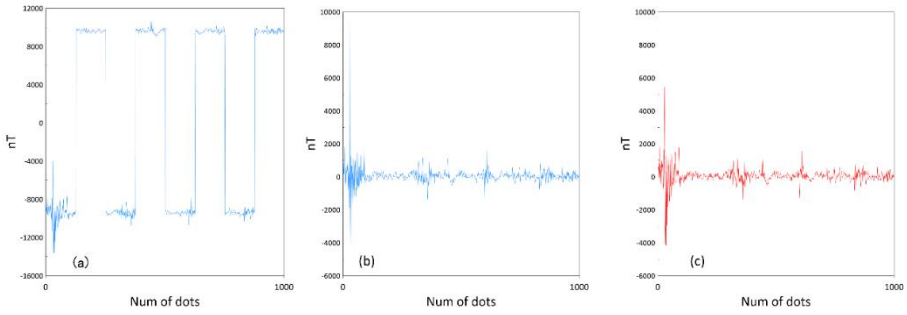


Figure 1. Hx(t) time series from northern Myanmar (a) noisy data, (b) raw data, (c)denoising data. The noisy data is affected by artificial step noise and denoising time series remove step noise accurately from noisy data.

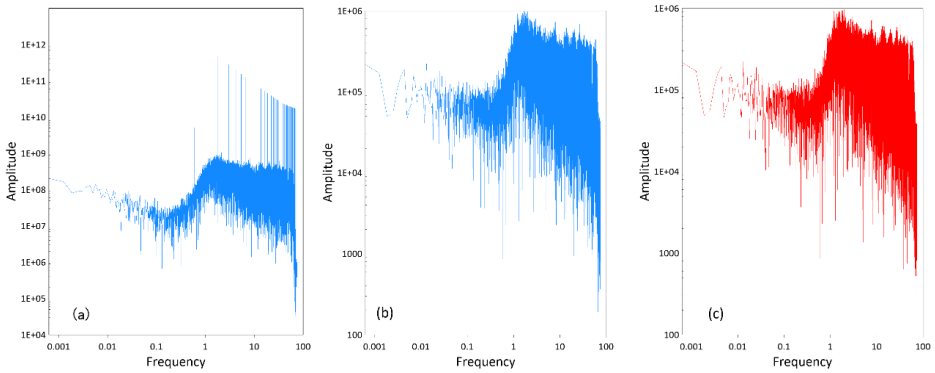


Figure 2. Spectrum for Hx(t) from northern Myanmar (a) noisy data, (b) raw data, (c) denoising data

3.2. Impulsive noise

Another magnetotelluric data with random-impulsive noise was processed for denoising using ED-LSTM. We select data of Ex(t) component to train and vertify the model and Ey(t) component for testing. Figure 3. shows the whole raw of Ey component (a) and denoised (b) time series. A great number of spikes in whole Ey component. The denoising model remove every spikes in raw time series accurately and improve data quality. The encoder-decoder only process signals near the spikes and we can ensure the useful part of raw data maintain unprocessed.

4. Conclusions

We proposes a denoising method to process step noise and impulsive noise using encoder-decoder consisted of LSTM cells for MT data in time domain. For step noise and random-impulsive noise, this method is proved with advantage processing persistent noise in raw data. It makes a new way to denoise MT data with step noise and impulsive noise in a long time duration. The encoder-decoder LSTM could potentially improve MT data quality.

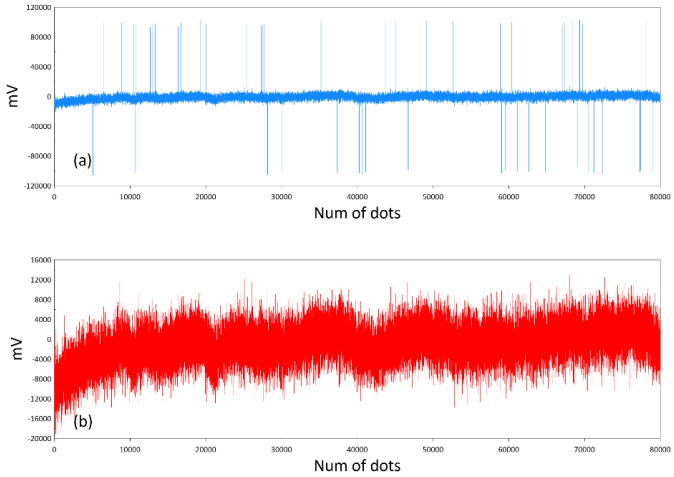


Figure 3. Raw MT time series with random-impulsive noises in terms of spike (a) and the corresponding time series after denoising (b).

References

[1] Chave A, Thomson D, Ander M 1987 On the robust estimation of power spectra coherences and transfer functions *Journal of Geophysical Research* **92** 633-648
[2] Chen H, Guo R, and Liu J 2020 Magnetotelluric data denoising with recurrent neural network: SEG 2019 Workshop *Mathematical Geophysics* 116-118
[3] Clarke J, Gamble T D, and Goubau W M, 1978 SQUIDs and magnetotellurics with a remote reference *AIP Conf Proc* **44**
[4] Escalas M, Queralt P, and Ledo J, Marcuello A 2013 Polarisation analysis of magnetotelluric time series using a wavelet-based scheme: A method for detection and characterisation of cultural noise sources *Physics of The Earth and Planetary Interiors* **218** 31-50
[5] Gers F A, Schmidhuber J, & Cummins F 2000 Learning to forget Continual prediction with LSTM *Neural computation* **12**(10) 2451-2471
[6] Manoj C, Nagarajan N 2003 The application of artificial neural networks to magnetotelluric time-series analysis *Geophysical Journal International* **153**(2) 409-423
[7] Hochreiter S, and Schmidhuber J 1997 Long short-term memory *Neural computation* **9**(8) 1735-1780
[8] Sutarno D, and Vozoff K 1989 Robust M-Estimation of Magnetotelluric Impedance Tensors *Exploration Geophysics* **20**(3) 383-398
[9] Sutskever I, Vinyals O, and Le Q 2014 Sequence to Sequence Learning with Neural Networks *Advances in Neural Information Processing Systems* **27**
[10] Wang H, and Yang J 2020 Current status and application prospect of deep learning in geophysics *Progress in Geophysics* **35**(2) 642-655
[11] Xu T, Wang Z 2020 Magnetotelluric power frequency interference suppression based on LSTM recurrent neural network *Progress in Geophysics* **35**(5) 2016-2022

Acknowledgments

This work was supported by the Key Research Program of the Chinese Academy of Sciences (Grant NO. ZDRW-ZS-2020-4).