Journal of Applied Geophysics 187 (2021) 104290

Contents lists available at ScienceDirect

Journal of Applied Geophysics

journal homepage: www.elsevier.com/locate/jappgeo

Machine learning based fast forward modelling of ground-based time-domain electromagnetic data

Thue Sylvester Bording [a,1], Muhammad Rizwan Asif [a,b,1,*], Adrian S. Barfod [b], Jakob Juul Larsen [b], Bo Zhang [a], Denys James Grombacher [a], Anders Vest Christiansen [a], Kim Wann Engebretsen [a], Jesper Bjergsted Pedersen [a], Pradip Kumar Maurya [a], Esben Auken [a,c]

[a] HydroGeophysics Group, Department of Geoscience, Aarhus University, Aarhus, Denmark
[b] Geophysics, Instrumentation, and Signal Processing, Department of Electrical and Computer Engineering, Aarhus University, Aarhus, Denmark
[c] Geological Survey of Denmark and Greenland, Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 22 January 2020
Received in revised form 23 October 2020
Accepted 21 February 2021
Available online 23 February 2021

Keywords:
Transient electromagnetic
Forward responses
Inversion modelling, neural networks

ABSTRACT

Inversion of large-scale time-domain electromagnetic surveys are computationally expensive and time consuming. Deterministic or probabilistic inversion schemes usually require calculations of forward responses, and often thousands to millions of forward responses are computed. We propose a machine learning based forward modelling approach as a computationally feasible alternative to approximate numerical forward modelling where a neural network is employed to model the relationship between the resistivity models and corresponding forward responses. For training of the neural network, we generated forward responses using conventional numerical algorithm for 93,500 resistivity models derived from different surveys conducted in Denmark representing typical resistivities of sedimentary geological layers. The input resistivity models and the network target outputs, i.e. forward responses, are scaled using a novel normalization strategy to ensure each gate is equally prioritized. The performance of the network is evaluated on two test datasets consisting of 8942 resistivity models by comparing the forward responses generated by the neural network and the conventional algorithm. We also measure the performance for the time derivatives of forward responses, i.e. dB/dt, by incorporating a system response. The results show that the proposed strategy is at least 13 times faster than commonly used accurate modelling methods and achieves an accuracy of 98% within 3% relative error, which is comparable to data uncertainty. Additional experiments on surveys from two other continents show that the results generalize in similar geological settings. Thus, under certain geological constraints, the proposed methodology may be incorporated into the pre-existing inversion structures, allowing for significantly faster inversion of large datasets.

1. Introduction

In exploration geophysics, the goal is to extract information about the subsurface from geophysical data. Translating geophysical data into useful information usually requires geophysical inverse modelling, the end result being a model of the target physical properties in the Earth's subsurface. The transient electromagnetic (TEM) method is a non-invasive geophysical method used to image the spatial variability of the electrical resistivity, or equivalently the conductivity, in the subsurface which reflects the geological structures. Hence, it is an effective non-invasive approach for mapping near surface geology, and it has found applications in groundwater mapping (Fitterman and Stewart, 1986), mineral exploration (Daniels and Dyck, 1984), saltwater intrusion mapping (Pedersen et al., 2017), permafrost mapping (Foley et al., 2019), etc.

The basic principles of TEM are well understood. A strong current is generated in a transmitter coil and is then rapidly turned off. During turn-off, a time varying magnetic field is produced, which in turn induces eddy currents at depth, which generate their own secondary magnetic field. The strength of this secondary magnetic field is measured as a function of time at the surface by a receiver coil. The magnitude and time-dependence of the secondary magnetic field holds information about the subsurface from which they originate. In order to generate images of subsurface properties from TEM data, inverse modelling is employed. Here, the end product is an estimate of the electrical resistivities of layers in the subsurface, consistent with the observed magnitude and time dependence of the secondary magnetic field (Auken et al., 2015).

* Corresponding author at: Room 321, Building 1120, Department of Geoscience, Aarhus University, C.F. Møllers Allé 4, 8000 Aarhus, Denmark.
E-mail address: rizwanasif@geo.au.dk (M.R. Asif).
[1] Equally contributing authors

https://doi.org/10.1016/j.jappgeo.2021.104290

TEM data inversion provides an estimate of the subsurface properties by iteratively updating and refining an initial subsurface model until an optimal model is obtained. Using the initial model, a forward response is numerically calculated using a set of equations that represent the influence of the coil geometries, subsurface properties, and instrument related factors, such as transmitter waveforms and filters. The model is considered to be optimal if the observed data and the calculated forward data agree within some uncertainty bounds. If not, the initial model is adjusted to minimize the error between the calculated and measured response. Model updates are typically performed iteratively using a Gauss-Newton methodology, and continue until either an acceptable data fit criterion is satisfied or the data fit fails to improve (Tarantola, 2005).

TEM inverse modelling based on numerical forward responses, requires several computational steps including inverse Fourier or Laplace transforms and Hankel transforms of kernel functions in the frequency, or Laplace/wavenumber domain (Auken et al., 2015) and can be a computationally expensive task. Especially in the computation of the Jacobian matrix used in the iterative model update, a significant number of forward responses are required. As such, forward response computation speeds are often the bottleneck in TEM inversions. Additionally, TEM surveys often produce massive amounts of data. Hence, inverse modelling based on fast forward responses using machine learning could have significant potential for improving the computational efficiency of geophysical inversion schemes.

Several studies using machine learning for forward modelling have been presented in literature: Artificial Neural Networks (ANNs) for forward modelling of electro-kinetic responses (Ardjmandpour et al., 2011), scatterometer forward responses for weather prediction (Cornford et al., 2001) and forward modelling for seismic waveform inversion (Fu et al., 2019), Modular Neural Network (MNN) for modelling borehole electrical resistivity in a layered formation (Zhang et al., 2002) and deep neural network based forward modelling of borehole resistivity measurements for geo-steering (Shahriaria et al., 2019).

Machine learning based approaches have also been used for inversion: ANNs for seismic amplitude-variation-with-offset (AVO) (Mogensen, 2001) and surface wave inversion (Hou et al., 2019), MNNs for modelling Cole-Cole parameters of TEM data (El-Kaliouby et al., 2001), pruning Bayesian neural networks for the inversion of electrical resistivity imaging (Jiang et al., 2016), deep learning based inversion, coined InversionNet (Wu et al., 2018) for full waveform inversion, 1-D Convolutional Neural Networks (CNNs) (Das et al., 2019) for seismic and for marine controlled-source EM data (Puzyrev et al., 2019).

While it may be possible to use a neural network directly for inversion, we elect to create a network for computing faster forward responses instead. The forward response is a one-to-one mapping, while inversion must deal with the non-uniqueness of the TEM inversion, where noise is also an important factor. Furthermore, such an approach would require training on individual systems as the data are system specific. However, alternate strategies for forward modelling would allow us to keep an inversion structure that is already well established (Auken et al., 2015), and may be incorporated into any inversion algorithm.

In this paper, we present an approach for modelling the forward responses of a ground-based TEM system using a standard 40 m × 40 m central loop configuration by deploying a fully connected feed-forward neural network, coined Fast Forward Modelling (FFM). FFM aims to replace the computationally expensive numerical forward modelling by simple matrix operations in order to provide a computationally efficient alternative. The input of the FFM is the subsurface resistivity model while the output is the magnitude of the system independent secondary magnetic field, i.e. the B-field, produced for a step-response TEM calculation. It should be noted here that the target output is the magnitude of the B-field itself and not the time derivative of the B-field. In order to verify the complete performance of the forward responses, a system impulse response is incorporated after the system independent forward responses are generated. We show that by considering this alternative approach, a significant speed-up factor is realized while achieving satisfactory performance accuracy, ensuring the suitability for use in inversion schemes.

2. Methods and methodology

As previously mentioned, inverse modelling is the key to extract meaningful information from TEM measurements, and any inversion scheme requires a forward modelling procedure. The traditional approach for forward modelling is either an analytical expression or, in the case of TEM, a numerical procedure (Auken et al., 2015; Christensen, 2002; Ward and Hohmann, 1988).

The idea behind FFM is that these hefty computations are replaced by a machine learning algorithm, which maps the inputs (resistivity model parameters) to the outputs (B-field), resulting in a significant speed-up to calculate time-domain forward responses that are sufficiently accurate to be used in geophysical inversion algorithms. Additionally, to make our approach more generalized, we use the neural network to generate forward responses that are independent of the system response, as it may vary depending on the instrument configuration. In the end, to incorporate the system response, the waveform is applied according to the convention used in (Fitterman and Anderson, 1987):

$$V(t) = \sum_{i=1}^{n-1} \frac{A_{i+1} - A_i}{t_{i+1} - t_i} [B]_{t-t_i}^{t-t_{i+1}} \tag{1}$$

where A is the amplitude of waveform, B is the secondary magnetic field produced for a step-response TEM, n is the total number of waveform data points and t represent the gate times of B.

2.1. Artificial neural networks

Inspired from biological neural networks, ANNs "learn" to perform specific tasks by considering a set of examples without being explicitly programmed (Chen et al., 2019). ANNs are composed of connected units generally referred to as artificial neurons, which aim to loosely model the biological brain. Generally, neurons are organized in layers, and all neurons of one layer are connected to all the neurons of neighbouring layers. ANNs use a supervised learning technique called backpropagation for training (Rumelhart et al., 1985). A typical structure of an ANN is shown in Fig. 1.

The underlying principle of an ANN is fairly simple and is briefly described here. Consider an input vector x with m elements and n neurons

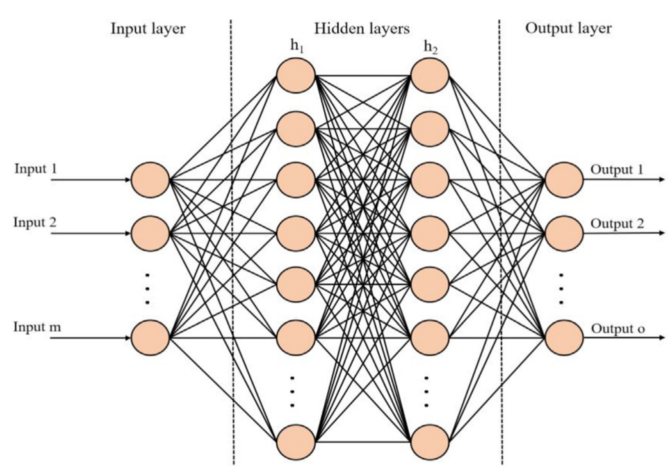

Fig. 1. Network architecture of an ANN with two hidden fully-connected layers.

in the hidden layer. Then, the output of i^{th} neuron in the hidden layer is given by:

$$y_i = f_i\left(\sum_{j=1}^{m} w_{ij}x_j + b_i\right) \tag{2}$$

where w_{ij} are weight factors, b_i is a bias, and f_i is the activation function that adds non-linearity to the system and should be chosen to resemble the characteristics of the underlying problem.

Neural networks are widely used for a range of applications across science and engineering in computer vision (LeCun et al., 2010), healthcare (Alloghani et al., 2019), time-series forecasting (Sezer et al., 2019), etc. They are also applied in many geoscience applications and have great potential in forward and inverse modelling for the approximation of numerical simulations and automated geoscientific processing schemes (Bergen et al., 2019; Reichstein et al., 2019).

ANNs provide the means to model the relationship between inputs and outputs without the need of knowing the physical model of the underlying problem. The relationship, which represents the physical model, is inherent in the input and output pairs that are fed to the ANN for training. Once trained, the ANN can predict the output for any input, and are, depending on the network size, computationally efficient. This characteristic prompts us to consider the use of neural networks in forward modelling of ground-based TEM data, where input and output pairs of the forward mapping are fed to the network for prediction. It should be noted here that proper data normalization is a key criterion for improved network convergence (Yan and Au, 2019). Hence, the input and output pairs are normalized before being given to the neural network for training.

2.2. Data normalization

The network input consists of a 30-layer resistivity model with log-increasing thicknesses with top layer thickness of 2.1 m and a depth to last layer boundary at 250 m. The layer thicknesses are fixed and are therefore not considered as an input parameter. A generalized approach would consider thickness as an input, since it can vary depending on the application, but inversions with predefined layer thicknesses, often called multi-layer or smooth, inversions are common in TEM. Additionally, we consider TEM responses from 10^{-7} s to 10^{-2} s, with 14 gates per decade in time.

It is practical to consider logarithmic variations in resistivity, as the sensitivity of the forward response does not vary linearly with changes in resistivity. For example, a change in resistivity from 5 Ωm to 10 Ωm would lead to much larger variation in data space than a change from 500 Ωm to 1000 Ωm. Therefore, we apply the logarithmic transform on the resistivity model before scaling it between [a, b] using eq. (3) for better optimization. In our case, [a, b] corresponds to [−1,1].

$$R_n = a + \frac{(b-a)(\log_{10}(R) - \log_{10}(R_{min}))}{\log_{10}(R_{max}) - \log_{10}(R_{min})} \tag{3}$$

where R_n is the normalized resistivity model of R, R_{min} is the minimum resistivity value and R_{max} denotes the maximum resistivity value, both obtained from the training dataset.

The neural network target output, i.e. the B-field, has a dynamic range spanning several orders of magnitude, and gate values at late times are relatively close to zero in comparison with the early times (see Fig. 2b). Therefore, the standard scaling normalization on the B-field is not optimal as the late times would have little to no impact when the network is being trained. We also do not consider the typical logarithmic transform of the data, as this effectively stretches the dynamic range of the small values while shrinking the dynamic range of large values (see Fig. 2c).

The logarithmic transform is also not well-suited to data containing sign changes, which may be encountered in offset geometries or in the presence of induced polarization effects. Instead, we propose to normalize each gate of the B-field using eq. (4) resulting in each gate value being weighted equally.

$$B_N(t) = a + \frac{(b-a)(B(t) - B_{min}(t))}{B_{max}(t) - B_{min}(t)} \tag{4}$$

where $B_N(t)$ are the normalized gate values of $B(t)$ while $B_{min}(t)$ and $B_{max}(t)$ is the minimum and maximum dynamic value for each gate at time t respectively.

The output of the neural network is de-normalized using the same parameters by manipulating eq. (4). However, due to the low dynamic range of the target outputs, rounding/truncation errors lead to kinks in the B-curve. These subtle kinks amplify the error when the system response is added to the output signal using eq. (1). Hence, it is necessary to post-process the de-normalized output by a smoothing method.

2.3. Neural network output post-processing

The de-normalized neural network output is processed with a locally estimated scatterplot smoothing (LOESS) algorithm that uses a weighted linear least squares fit and a 2nd degree polynomial model (Cleveland and Devlin, 1988). LOESS is developed based on classical methods, such as linear and nonlinear least squares regression. For each data point, a second order polynomial is fit in a local window of the data. Weighted least squares is used to fit the polynomial, where

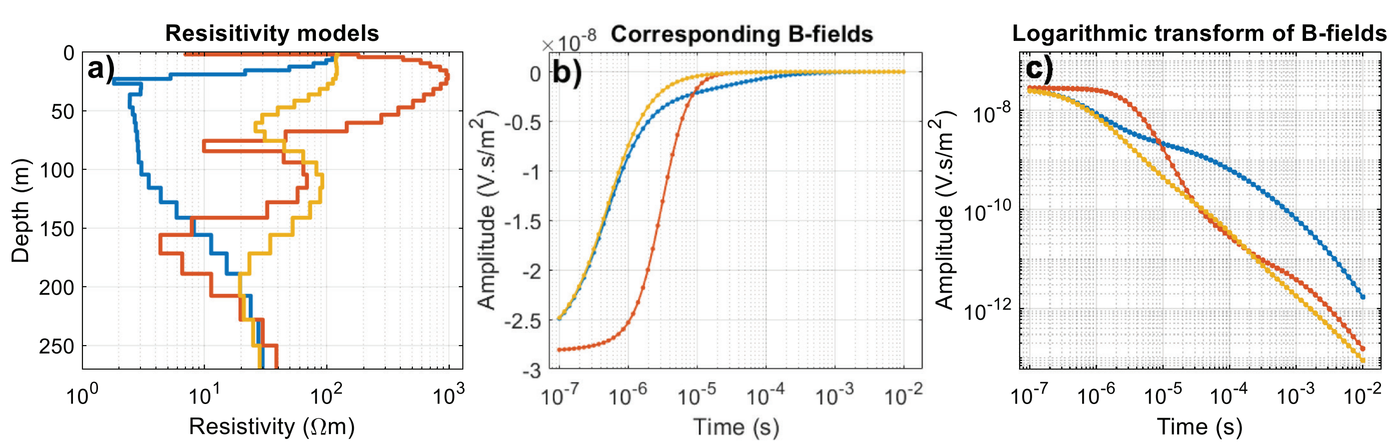

Fig. 2. Three examples of resistivity models and associated B-fields (a) Resistivity models, (b) B-field B (c) $\log_{10}|B|$.

higher weights are assigned to data points in the vicinity of the estimation point, and lower weights are allotted to distant data points. The local polynomial is then evaluated by using the dependent variable for each data point to obtain the values of the regression function. Typically, a tri-cubic function is used to compute the regression weights for each data point within a local window.

$$w_i(x) = \left(1 - \left|\frac{x - x_i}{d(x)}\right|^3\right)^3 \tag{5}$$

where x is the point to be estimated that is associated with the response value, $d(x)$ is the distance between data points x_i and x. Heuristic analysis shows that, in our case, the local window size of 15 gate times results in the best smoothing results.

2.4. Proposed neural network architecture

We use a four-layered fully-connected feed-forward network architecture, where the neurons in the first layer correspond to the number of inputs (the resistivity value in each of the 30 depth layers), yields 30 input neurons. The second and third layers are the hidden processing layers, and have 260 and 180 neurons respectively. The number of neurons in the hidden layers are chosen empirically. The fourth layer, representing the output layer, has 71 neurons which correspond to the number of gates of the B-field. Obviously, one can argue why this specific architecture is chosen. In our experience, not much improvement is gained when deeper network architecture is used. Additionally, heuristic analysis shows that the selected number of neurons in the hidden layers gives the best performance.

A hyperbolic tangent function, i.e. tanh-function, is used as the activation function for eq. (2) in the hidden layers to add non-linearity to the system (Cybenko, 1989). This activation function is well suited for the data ranging between -1 and $+1$ as it smoothly approaches -1 as x goes to $-\infty$ and 1 when x goes to $+\infty$. The tanh-function is also zero centred, which improves the modelling of strongly negative/positive and neutral values. Lastly, at the output layer, a linear activation function is selected for regression.

Once the network architecture is defined, the next phase is to determine the weights and biases that correspond to the features that relate input with the target outputs. The weights and biases of the network are updated iteratively until a good relationship between the target outputs and inputs is found. For this purpose, we use the scaled conjugate gradient (SCG) algorithm to update the network weights by backpropagation. The SCG method generally performs well for the networks with a large number of parameters (Chel et al., 2011). We consider using the full-batch algorithm to avoid the tuning of the additional parameter, i.e. the mini-batch size, as its performance in comparison to the mini-batch algorithm with different batch sizes is similar (Zheng et al., 2016). This trade-off is achieved at the cost of more training time.

Rather than assigning the initial weights randomly, we use the Nguyen-Widrow initialization algorithm (Nguyen and Widrow, 1990) that approximately distributes the active region of each neuron in the layer evenly across the input space (Andayani et al., 2017). Small numbers of random values are assigned in Nguyen-Widrow initialization prior to backpropagation, which helps reduce the time it takes to train a network. If the initial weights are too large, the neurons will fall into a saturation region, where the derivative values of the activation function are small. If the weights are too small, the value of the given neurons will approach to zero, resulting in diminishing gradients, causing little to no learning (Mishra et al., 2014).

Lastly, the goal of the proposed forward modelling approach is to minimize the error between the actual measurements and the predictions made by the network. Hence, we define the loss function as the sum of the squares error (SSE) as in eq. (6).

$$E = \sum_{i=1}^{N} \left(x_i - \widehat{x}_i\right)^2 \tag{6}$$

where x_i is the target output, $\widehat{x}_i$ is the predicted output and N is the number of samples.

A value closer to 0 indicates that the model has a smaller random error, and the fit will be more useful for prediction, similar to the process of geophysical inversion.

3. Results and discussion

We have used MATLAB 2019b on a system with an Intel Xeon Gold 6132 CPU with 2.6GHz, and three NVIDIA GeForce RTX 2080Ti GPUs for training the neural network. As we have used a supervised learning model, we require a database of inputs, i.e. the resistivity models, and corresponding targets, i.e. the B-fields. The database comprises of 112,964 resistivity models acquired from various tTEM (Auken et al., 2018) surveys spread across Denmark, collected by the HydroGeophysics Group (HGG) at Aarhus University, Denmark. The tTEM system is specifically designed for detailed geophysical mapping of the shallow subsurface. It uses a 2 m × 4 m transmitter coil and a z-component receiver coil in an offset configuration from the transmitter, and is towed by an all-terrain vehicle.

The datasets represent different Quaternary sedimentary environments and span a wide range of possible Danish geological environments. Note that the tTEM system can image to depths of approximately 100 m, which is shallower than the 40 m × 40 m coincident TEM system considered in this work. To translate these models to a depth interval relevant for the 40 m × 40 m case, we scale the depths of each model to reach 250 m. This procedure is done to provide a larger training dataset of geologically plausible resistivity models for the 40 m × 40 m case. Although each individual model has been manipulated from those observed in the field, a vertical stretching is still able to produce a geologically realistic scenario.

The database is shuffled prior to dividing it into a training, a validation, and a test set. A total of 93,500 models are used for training, while 16,500 and 2964 models are used for validation and testing purposes, respectively. We also include an entirely different tTEM survey conducted in Gedved, an area in central Denmark, consisting of 5978 resistivity models stretched in the same way. The target outputs, i.e. the forward responses, for the corresponding resistivity models were computed using AarhusInv (Auken et al., 2015). Fig. 3 shows the prevalence of resistivity values for the data used in this study.

The data is normalized by acquiring R_{min}, R_{max}, $B_{min}(t)$ and $B_{max}(t)$ from the training dataset and applying eq. (3) and eq. (4). To ensure that our FFM covers the entire dynamic range of all possible B-fields, the values of $B_{min}(t)$ and $B_{max}(t)$ are extended by 30%. Training the network on the 93,500 samples took ~3 h, with an early stopping criterion to avoid overfitting. The early stopping criteria refers to the point where the validation error starts to increase, while the training error is still decreasing. The training and validation loss of the network with regards to time, while training is shown in Fig. 4, which shows no indication of overfitting.

The output of the neural network is de-normalized using the same parameters by manipulating eq. (4). To evaluate the performance of our FFM, we use the resistivity models from the test set and the Gedved survey. The Gedved survey, which is completely unseen to the network, would give a good indication about the generalization capabilities of the trained network. Fig. 5a shows that the 99.86% of all the output gates lie within a ± 3% relative error, i.e. a typical assumed uncertainty level for many TEM inversions, for the test set comparing to 99.32% for the Gedved survey. In order to evaluate each gate individually, the percentage of gates within ±3% range of relative error is calculated and presented in Fig. 5b. Since the new survey is completely unknown to the network, higher error was expected for the Gedved survey, as compared to the test set. The test set is a subset of the same surveys used for

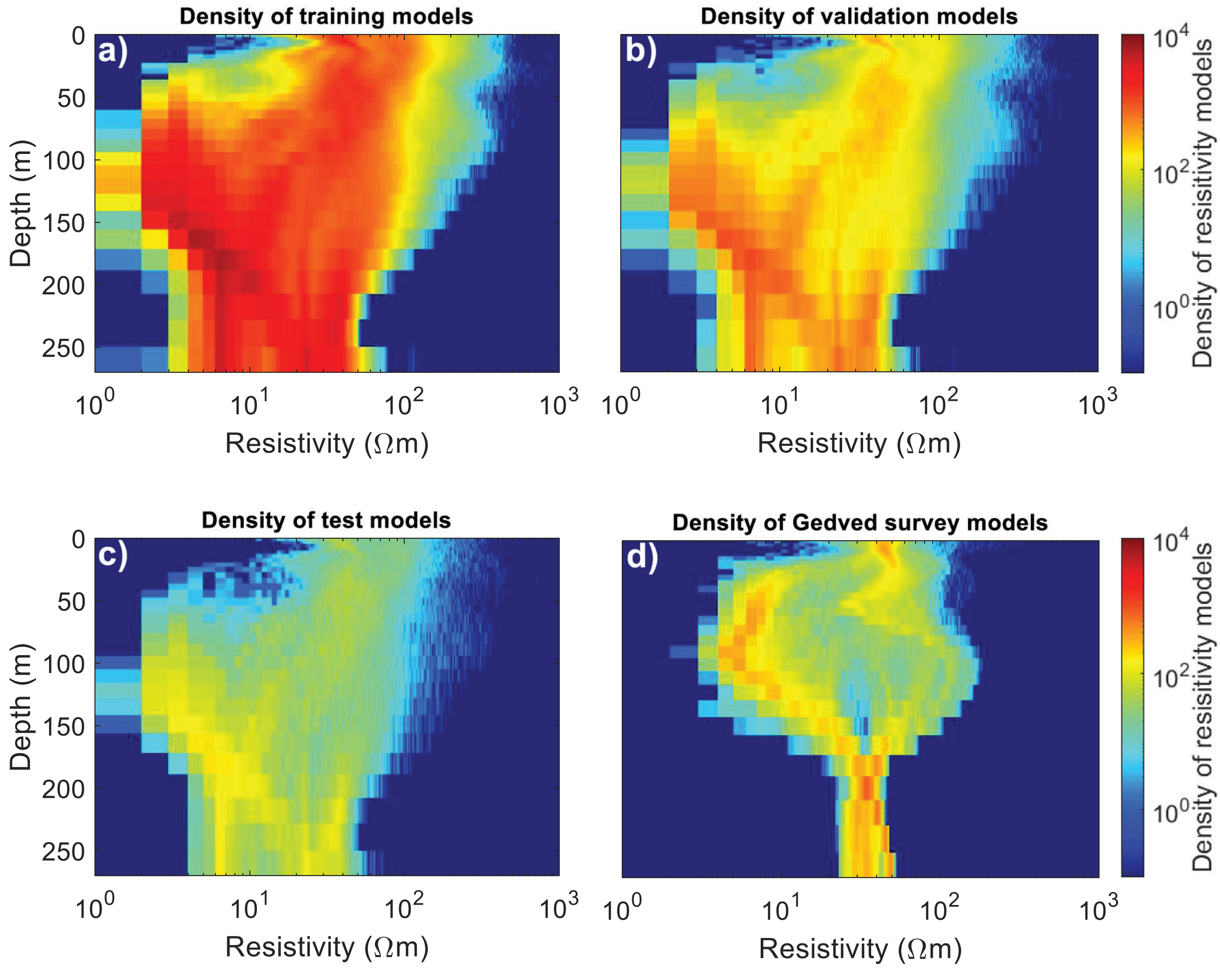

Fig. 3. Density plots of resistivity models for training, validation and test sets (a) Training set (b) Validation set (c) Test set (d) Gedved survey models.

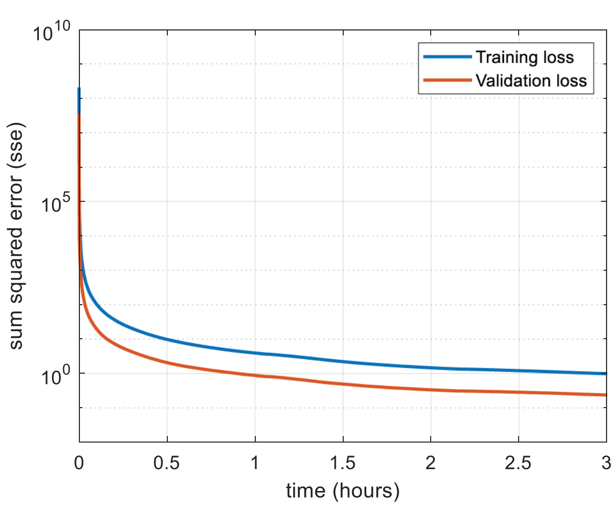

Fig. 4. Performance plot for training and validation data during training.

training, and therefore, was expected to perform better. Nevertheless, for a general ground-based TEM system, the usable gate times typically begin from 5 μs to 100 μs, and end around 1 ms to 3 ms for which the performance of the network is found to be satisfactory.

3.1. Incorporating system response

To include the system response of the instrument, we also compute the time derivatives, $\frac{dB}{dt}$, of the forward responses generated by FFM and AarhusInv, and convolve with the transmitter waveform using eq. (1). A simple input waveform, 3 ms ramp up, 0.2 ms on-time and 3 μs ramp down, is considered. We do not consider gate times before the transmitter turn-off, which means that the first gate is at 3 μs. Also, as the smoothing is performed within a localized portion of the data, late gates do not get a complete window for smoothing and the last gate is at 3 ms that lies within the typical usable data window.

Fig. 6 shows the results when the system response is incorporated for the test set. Fig. 6a illustrates that 97.8% of all the gates lie within the 3% error range for the smoothened output for the test set as compared to the unprocessed output which only has an accuracy of 86.2%. Additionally, Fig. 6b shows the results of individual gates in the ±3% relative error range, where the post-processed output performs better. As mentioned earlier, the error propagates when the system response is incorporated to the un-processed output signal due to kinks in the B-field curve. Therefore, the post-processing of the output using LOESS helps supress the errors and improve the performance of FFM.

Similarly, we show the performance of FFM on the Gedved survey in Fig. 7. Fig. 7a shows that 98.02% of the gate times are within a 3% error range for the post-processed output achieving similar performance as the test set. Moreover, it is evident from Fig. 7b that the performance

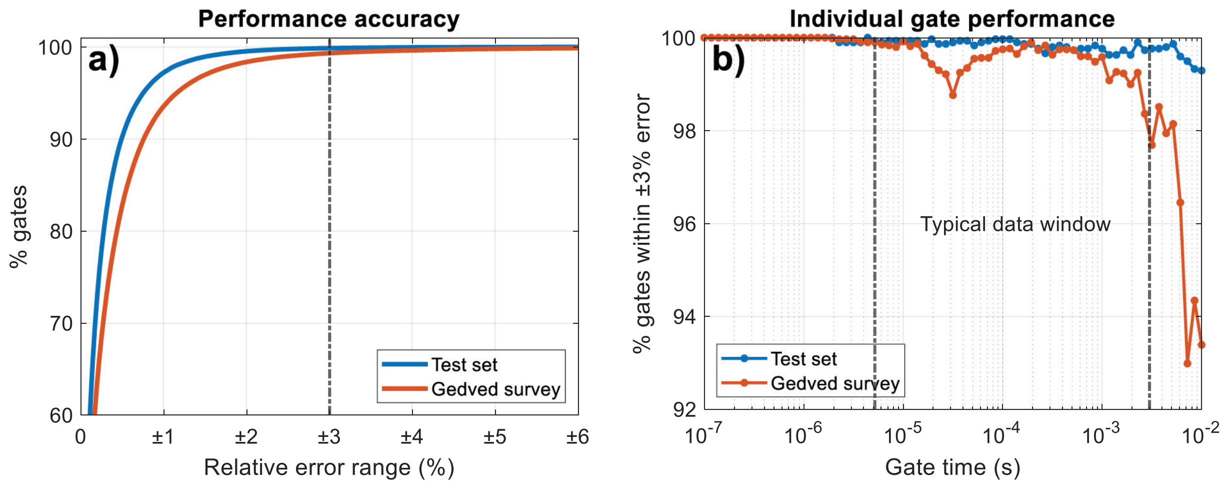

Fig. 5. Performance of FFM for the test set and the unseen Gedved survey (a) Relative error performance (b) Individual gate performance.

of FFM for individual gates for both test sets is almost identical, which means that the network has achieved good generalization capabilities. Furthermore, as expected, the post-processing output dominates the un-processed output.

3.2. Generalizability test

To examine the generalizability of the proposed methodology, we consider two surveys conducted in two different continents, North America and Africa. These surveys are conducted in California, USA and Ladysmith, South Africa, and consist of 3923 and 9360 models, respectively. As the resistivity models from these surveys are also obtained from the tTEM system, the depth of the models are to reach 250 m as explained earlier. The training of the neural network is based on data collected from Denmark, hence an evaluation of the network on the data from significantly different regions would be a good measure to assess the generalization of the network.

Fig. 8a and Fig. 8b illustrate the suite of resistivity models present within the California and Ladysmith surveys, respectively. The California resistivity models are observed to cluster around resistivity values from 5 Ω m to 100 Ω m over the investigated depth interval. Some higher resistivity values are observed at shallow depths, but not in significant abundances. This is consistent with the background geological information of the area suggesting the presence of alluvial sand and clay deposits. Groundwater in the region is not expected to have elevated salinity.

In contrast, the Ladysmith survey exhibits a much wider range of resistivity values, with a significant proportion of the models displaying elevated resistivities exceeding several hundred ohm m. Some low resistivity (<10 Ω m) values are present at shallow depths, but much of the investigated depth interval displays elevated resistivities (> ~ 30 Ω m). The Ladysmith data were collected in the relatively flat terrain surrounding a river. There is an expectation that alluvial deposits are present throughout the area, as well as the presence of Ecca Group shale and sandstone units intermittently intruded by Jurassic dolerite sills and dykes. The dolerites are the likely source of the very high resistivity values in the area.

Fig. 9a shows that the 99.71% of all the output gates lie within a ± 3% relative error for the California survey, while the accuracy performance for the Ladysmith survey is found to be 54.03%. In order to evaluate the gate-wise performance, the percentage of gates within the ±3% range of relative error is calculated and presented in Fig. 9b. The poor performance of the neural network on the Ladysmith survey is not surprising as resistivity models contain an abundance of elevated resistivities at deep layers. Such models where high resistivity is observed at depth are not present in the model range used for training, which is noted by comparing Fig. 3a and Fig. 8b.

We also show the performance of the network for these surveys with the system response on the post-processed output in Fig. 9c and

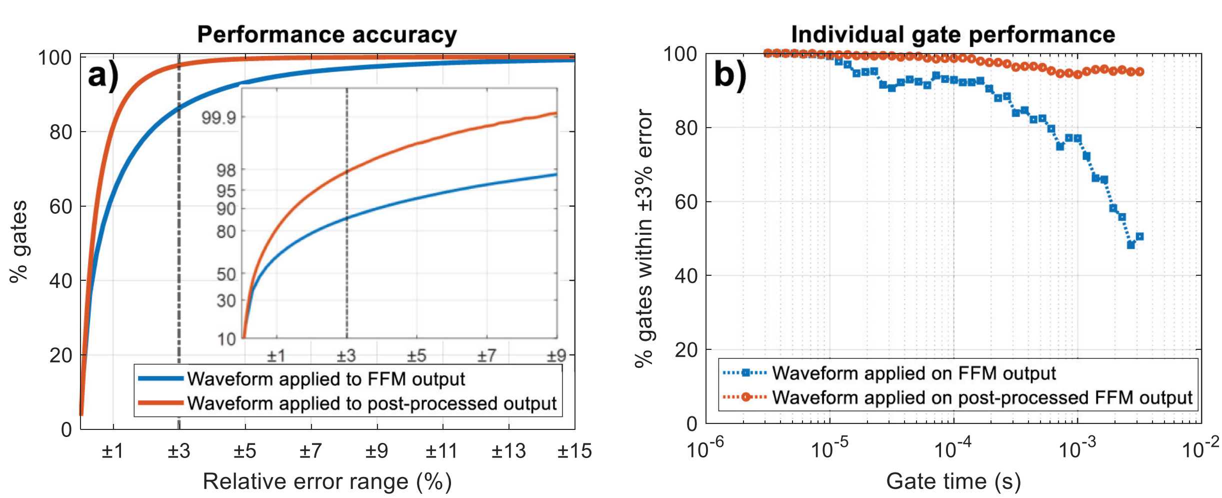

Fig. 6. Performance of FFM for the test set with system response (a) Relative error performance, the inset shows the data in stretched y-axis (b) Gate-wise performance.

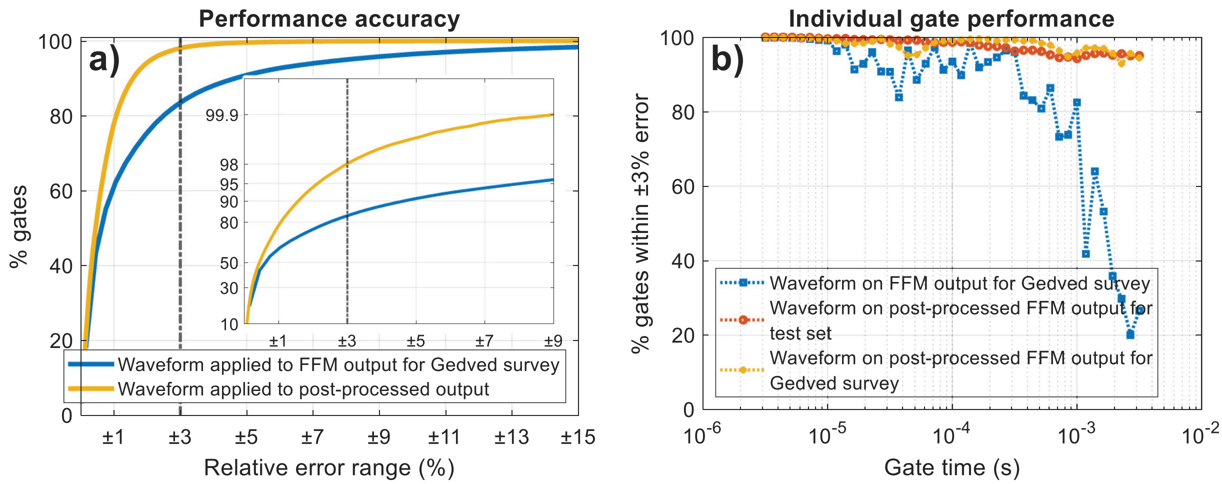

Fig. 7. Performance of FFM for the Gedved survey with system response (a) Relative error performance, the inset shows the data in stretched y-axis (b) Individual gate performance.

Fig. 9d. It can be seen in Fig. 9c that 98.55% of the gates for California survey are within a 3% error range for the post-processed output achieving similar performance as the test set and the Gedved survey. However, the performance achieved on the Ladysmith survey is further reduced to 38.45%.

In order to improve the network's accuracy for such resistive models, additional training data is required to cover the entire model range. Fine-tuning the trained network by considering the whole model space would result in improvement of accuracy for high resistivity models. One way to cover the entire model range is to augment the training models by shifting the existing resistivity models by a factor. Another way is to generate random models to cover the entire model range. However, in our experience, using random models for the training of the network does not result in optimal performance. Therefore, it is desirable to train the neural network on the resistivity models which are similar that are similar to those seen by the TEM system.

Although our FFM is trained on the resistivity models obtained from surveys in Denmark, the performance is found to be satisfactory when the resistivity models are within the same range as that found in the training model space, regardless of the geographical location. However, in order to make the proposed approach viable for a more diverse range of geological settings, it is essential to train the neural network on the resistivity models that cover the entire model range.

3.3. FFM computation time

The main objective of this study is to speed-up the TEM forward modelling. Therefore, we compare the processing time of the proposed FFM approach with two established geophysical 1-D modelling methods, namely AarhusInv (Auken et al., 2015) and AirBeo (Kwan et al., 2015), which are widely used inversion codes for inverse modelling of electric and electromagnetic data. Both of these methods are implemented in FORTRAN, therefore, we extract the weights and biases from the trained network to apply eq. (2) in FORTRAN and get the final output by post-processing with the LOESS algorithm. It should be noted here that the FORTRAN code has not been optimized and the computation time is evaluated on a single CPU core without any parallelization.

Table 1 shows the computation time comparison of FFM with the above-mentioned methods. Evident from Table 1, FFM gives a speed-up of 13 times when compared to AarhusInv and over 17 times when compared with AirBeo. However, optimization of the FORTRAN code may result in improved FFM computational performance. It should also be noted here that the FFM post-processing takes ~60% of the computation time, which may be reduced by considering an alternative post-processing strategy.

The immediate benefit of this speed-up is shorter inversion times, which is a general improvement. The derived opportunities lie in

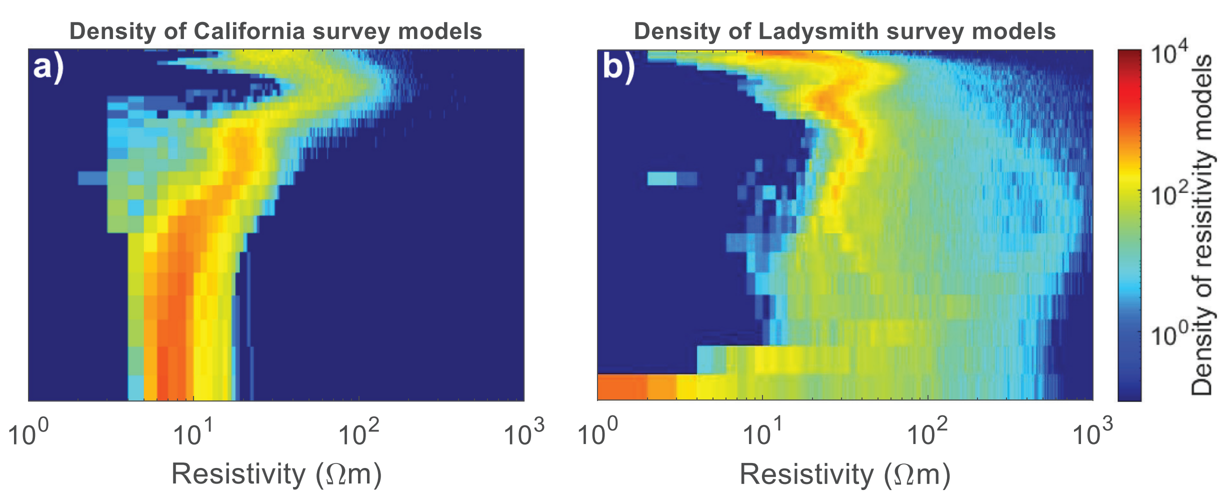

Fig. 8. Density plots of resistivity models of surveys from two different continents (a) California, USA (b) Ladysmith, South Africa.

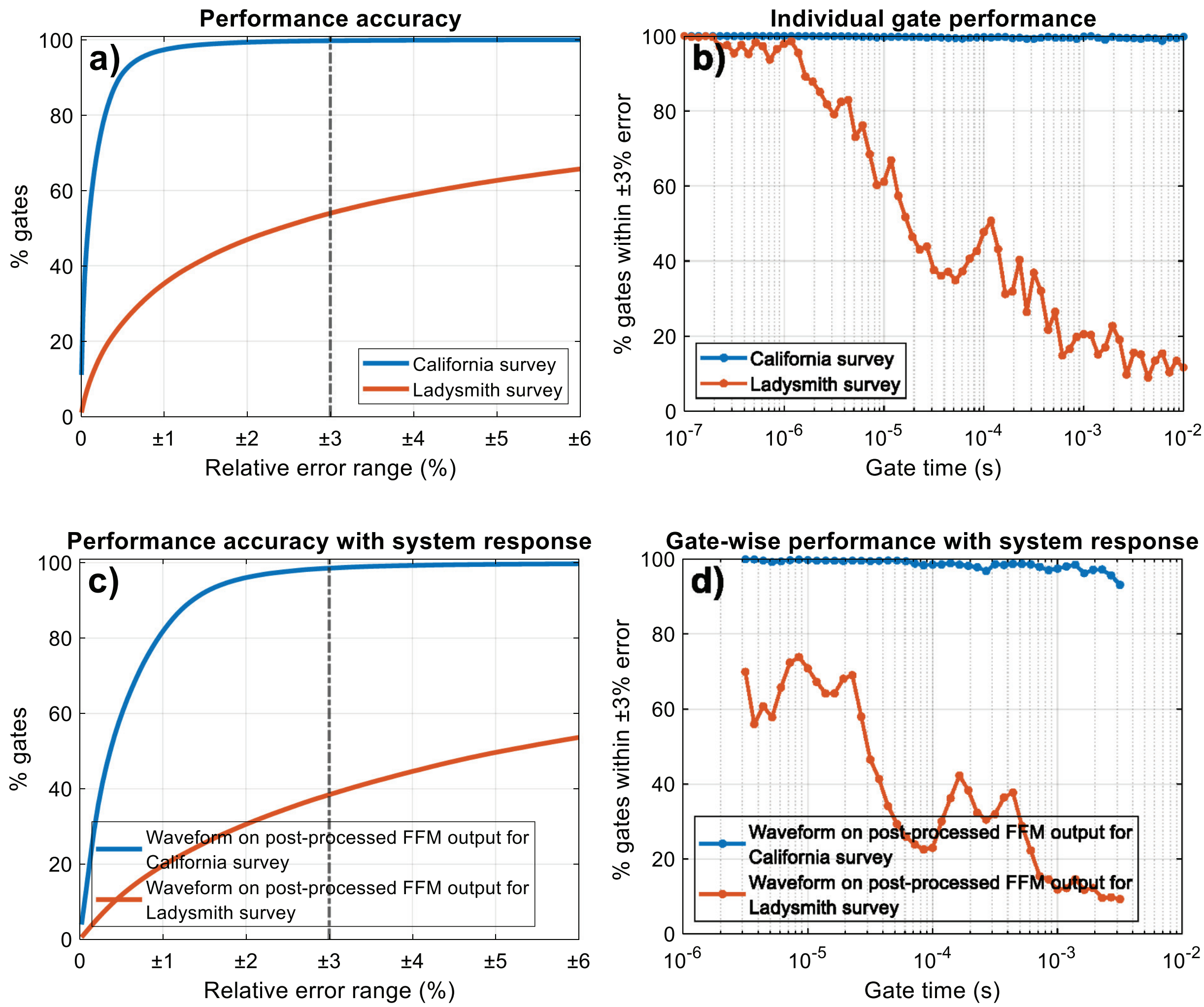

Fig. 9. Performance of FFM for the California and Ladysmith survey (a) Relative error performance (b) Individual gate performance (c) Relative error performance including system response (d) Gate-wise performance with system response.

Table 1
Computation time comparison for forward responses on a single CPU core.

Modelling method	Speed
FFM	208.3 forward responses/s
AarhusInv	16.2 forward responses/s
AirBeo	11.8 forward responses/s

different inversion schemes, such as a two-sided derivatives calculation in the Jacobian matrix, running multiple inversions with different starting models or different inversion settings, or by switching to probabilistic methods.

In this work, we have considered a TEM system with a 40 m × 40 m central loop configuration. However, in principle, any geometry can be applied. Furthermore, based on the proposed scheme, FFM for various TEM instruments, such as tTEM (Auken et al., 2018) and SkyTEM (Sørensen and Auken, 2004) can be modelled by considering extra parameters at the input; such as distance between Tx and Rx for tTEM, and altitude for SkyTEM.

4. Conclusion

The proposed machine learning based methodology for fast forward modelling of time-domain EM data, is shown to be more than an order of magnitude faster than other commonly used numerical forward models. We show that the network output for B-fields after de-normalization is highly accurate without any post-processing required, with 99.86% and 99.32% of all gates from the two test datasets being accurate to within 3% relative error. For dB/dt values to be satisfactorily accurate, we apply a LOESS smoothing function on the B-fields before convolving with a transmitter waveform. For our test datasets, 97.8% and 98% of all gates in the typical time-range of TEM data are accurate to within 3%. This shows that the accuracy of FFM is within typical data uncertainties.

The network is trained and tested on resistivity models derived from Danish surveys, and therefore it performs better in similar geological environments, but other environments may also be incorporated in the training for future applications.

Funding

This work was supported by Innovation Fund Denmark under the 'MapField' project.

Author statement

Conceptualization: Thue Sylvester Bording, Muhammad Rizwan Asif; Data curation: Thue Sylvester Bording, Muhammad Rizwan Asif, Jesper Bjergsted Pedersen; Formal analysis: Muhammad Rizwan Asif, Adrian Sanchez Barfod, Kim Wann Engebretsen; Funding acquisition: Jakob Juul Larsen, Esben Auken; Investigation: Jakob Juul Larsen, Esben Auken, Denys James Grombacher, Anders Vest Christiansen; Methodology: Muhammad Rizwan Asif, Thue Sylvester Bording, Adrian Sanchez Barfod; Project administration: Jakob Juul Larsen, Esben Auken; Resources: Jakob Juul Larsen, Esben Auken; Software: Muhammad Rizwan Asif, Bo Zhang; Supervision: Jakob Juul Larsen, Esben Auken; Validation: Pradip Kumar Maurya; Roles/Writing - original draft: Muhammad Rizwan Asif, Thue Sylvester Bording, Adrian Sanchez Barfod; Writing - review & editing: Muhammad Rizwan Asif, Adrian Sanchez Barfod, Denys James Grombacher, Jakob Juul Larsen, Anders Vest Christiansen.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The data from Ladysmith, South Africa was collected during a project funded by Umgeni Water, JG Africa, and the Poul Due Jensens Foundation.

References

Alloghani, M., Al-Jumeily, D., Aljaaf, A.J., Khalaf, M., Mustafina, J., Tan, S.Y., 2019. The application of artificial intelligence technology in healthcare: a systematic review. International Conference on Applied Computing to support Industry: Innovation and Technology. Springer, pp. 248–261.

Andayani, U., Nababan, E.B., Siregar, B., Muchtar, M.A., Nasution, T.H., Siregar, I., 2017. Optimization backpropagation algorithm based on Nguyen-Widrom adaptive weight and adaptive learning rate. 2017 4th International Conference on Industrial Engineering and applications (ICIEA). IEEE, pp. 363–367.

Ardjmandpour, N., Pain, C., Singer, J., Saunders, J., Aristodemou, E., Carter, J., 2011. Artificial neural network forward modelling and inversion of electrokinetic logging data. J Geophys. Prospect. 59, 721–748.

Auken, E., Christiansen, A.V., Kirkegaard, C., Fiandaca, G., Schamper, C., Behroozmand, A.A., Binley, A., Nielsen, E., Effersø, F., Christensen, N.B., 2015. An overview of a highly versatile forward and stable inverse algorithm for airborne, ground-based and borehole electromagnetic and electric data. J Expl. Geophys. 46, 223–235.

Auken, E., Foged, N., Larsen, J.J., Lassen, K.V.T., Maurya, P.K., Dath, S.M., Eiskjær, T.T., 2018. tTEM—A towed transient electromagnetic system for detailed 3D imaging of the top 70 m of the subsurface. J. Geophys. 84, E13–E22.

Bergen, K.J., Johnson, P.A., Maarten, V., Beroza, G.C., 2019. Machine learning for data-driven discovery in solid Earth geoscience. J. Sci. 363 eaau0323.

Chel, H., Majumder, A., Nandi, D., 2011. Scaled conjugate gradient algorithm in neural network based approach for handwritten text recognition. International Conference on Computational Science, Engineering and Information Technology. Springer, pp. 196–210.

Chen, Y.-Y., Lin, Y.-H., Kung, C.-C., Chung, M.-H., Yen, I., 2019. Design and implementation of cloud analytics-assisted smart power meters considering advanced artificial intelligence as edge analytics in demand-side management for smart homes. J Sensors 19, 2047.

Christensen, N.B., 2002. A generic 1-D imaging method for transient electromagnetic data. J. Geophys. 67, 438–447.

Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. J. Am. Stat. Assoc. 83, 596–610.

Cornford, D., Nabney, I.T., Ramage, G., 2001. Improved neural network scatterometer forward models. J. Geophys. Res.: Oceans 106, 22331–22338.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. J. Math. control, signals 2, 303–314.

Daniels, J.J., Dyck, A.V., 1984. Borehole resistivity and electromagnetic methods applied to mineral exploration. IEEE Transactions on Geoscience Remote Sensing, pp. 80–87.

Das, V., Pollack, A., Wollner, U., Mukerji, T., 2019. Effect of rock physics modeling in impedance inversion from seismic data using convolutional neural network. The 13th SEGJ International Symposium, Tokyo, Japan, pp. 522–525.

El-Kaliouby, H.M., ElDiwany, E.A., Hashish, E., Bayoumi, A., Poulton, M.M., 2001. Forward modeling and inversion of IP effects in TEM response using measured rock samples data. SEG Technical Program Expanded Abstracts 2001. Society of Exploration Geophysicists, pp. 1525–1528.

Fitterman, D.V., Anderson, W.L., 1987. Effect of transmitter turn-off time on transient soundings. J Geoexploration 24, 131–146.

Fitterman, D.V., Stewart, M.T., 1986. Transient electromagnetic sounding for groundwater. J. Geophys. 51, 995–1005.

Foley, N., Tulaczyk, S., Auken, E., Grombacher, D., Mikucki, J., Foged, N., Myers, K., Dugan, H., Doran, P.T., Virginia, R.A., 2019. Mapping geothermal heat flux using permafrost thickness constrained by airborne electromagnetic surveys on the western coast of Ross Island, Antarctica. J Expl. Geophys. 1–10.

Fu, H., Zhang, Y., Ma, M., 2019. Seismic waveform inversion using a neural network-based forward, Physics. Conference Series. IOP Publishing.

Hou, S., Angio, S., Clowes, A., Mikhalev, I., Hoeber, H., Hagedorn, S., 2019. Learn to Invert: Surface Wave Inversion with Deep Neural Network, 81st EAGE Conference and Exhibition 2019 Workshop Programme. Eur. Assoc. Geosci. & Eng. 1–5.

Jiang, F.-B., Dai, Q.-W., Dong, L., 2016. Nonlinear inversion of electrical resistivity imaging using pruning Bayesian neural networks. J. Appl. Geophys. 13, 267–278.

Kwan, K., Prikhodko, A., Legault, J.M., Plastow, G., Xie, J., Fisk, K., 2015. Airborne inductive induced polarization chargeability mapping of VTEM data. J ASEG Extended Abstracts 2015, 1–5.

LeCun, Y., Kavukcuoglu, K., Farabet, C., 2010. Convolutional networks and applications in vision. Proceedings of 2010 IEEE International Symposium on Circuits and Systems. IEEE, pp. 253–256.

Mishra, K., Mittal, N., Mirja, M.H., 2014. Image compression using multilayer feed forward artificial neural network with nguyen widrow weight initialization method. J Int. J. Emerg. Technol. Adv. Eng. 4.

Mogensen, S., 2001. Artificial neural networks solutions to AVO inversion problems. SEG Technical Program Expanded Abstracts 2001. Society of Exploration Geophysicists, pp. 316–319.

Nguyen, D., Widrow, B., 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. 1990 IJCNN International Joint Conference on Neural Networks. IEEE, pp. 21–26.

Pedersen, J.B., Schaars, F.W., Christiansen, A.V., Foged, N., Schamper, C., Rolf, H., Auken, E., 2017. Mapping the fresh-saltwater interface in the coastal zone using high-resolution airborne electromagnetics. J First Break 35, 57–61.

Puzyrev, V., Meka, S., Swidinsky, A., 2019. Deep convolutional neural networks for 1D inversion of electromagnetic data. 81st EAGE Conference and Exhibition European Association of Geoscientists & Engineers, pp. 1–5.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., 2019. Deep learning and process understanding for data-driven Earth system science. J Nature 566, 195–204.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1985. Learning Internal Representations by Error Propagation. California Univ San Diego La Jolla Inst for Cognitive Science.

Sezer, O.B., Gudelek, M.U., Ozbayoglu, A.M., 2019. Financial time series forecasting with deep learning: a systematic literature review: 2005–2019. J arXiv preprint arXiv 13288.

Shahriaria, M., Pardob, D., Mosera, B., 2019. A deep neural network as surrogate model for forward simulation of borehole resistivity measurements. viXra preprint viXra 42 1910.0255.

Sørensen, K.I., Auken, E., 2004. SkyTEM-A new high-resolution helicopter transient electromagnetic system. J Exploration Geophysics 35, 191–199.

Tarantola, A., 2005. Inverse problem theory and methods for model parameter estimation. Society for Industrial and Applied Mathematics (SIAM).

Ward, S.H., Hohmann, G.W., 1988. Electromagnetic theory for geophysical applications. Electromagnetic Methods in Applied Geophysics: Voume 1, Theory. Society of Exploration Geophysicists, pp. 130–311.

Wu, Y., Lin, Y., Zhou, Z., 2018. InversionNet: Accurate and efficient seismic waveform inversion with convolutional neural networks. SEG Technical Program Expanded Abstracts 2018. Society of Exploration Geophysicists, pp. 2096–2100.

Yan, N., Au, O.T.-S., 2019. Online learning behavior analysis based on machine learning. J Asian Association of Open Universities Journal 14, 97–106.

Zhang, L., Poulton, M.M., Wang, T., 2002. Borehole electrical resistivity modeling using neural networks. J. Geophys. 67, 1790–1797.

Zheng, J., Ma, Q., Zhou, W., 2016. Performance comparison of full-batch bp and mini-batch bp algorithm on spark framework. 2016 8th International Conference on Wireless Communications & Signal Processing (WCSP). IEEE, pp. 1–5.