

Artificial neural networks for removal of couplings in airborne transient electromagnetic data

Kristoffer K. Andersen*, Casper Kirkegaard, Nikolaj Foged, Anders V. Christiansen and Esben Auken

Department of Geoscience, Aarhus University, C.F. Møllers Allé 4, DK-8000 Aarhus C, Denmark

Received April 2015, revision accepted July 2015

ABSTRACT

Modern airborne transient electromagnetic surveys typically produce datasets of thousands of line kilometres, requiring careful data processing in order to extract as much and as reliable information as possible. When surveys are flown in populated areas, data processing becomes particularly time consuming since the acquired data are contaminated by couplings to man-made conductors (power lines, fences, pipes, etc.). Coupled soundings must be removed from the dataset prior to inversion, and this is a process that is difficult to automate. The signature of couplings can be both subtle and difficult to describe in mathematical terms, rendering removal of couplings mostly an expensive manual task for an experienced geophysicist.

Here, we try to automate the process of removing couplings by means of an artificial neural network. We train an artificial neural network to recognize coupled soundings in manually processed reference data, and we use this network to identify couplings in other data. The approach provides a significant reduction in the time required for data processing since one can directly apply the network to the raw data. We describe the neural network put to use and present the inputs and normalizations required for maximizing its effectiveness. We further demonstrate and assess the training state and performance of the network before finally comparing inversions based on unprocessed data, manually processed data, and artificial neural network automatically processed data. The results show that a well-trained network can produce high-quality processing of airborne transient electromagnetic data, which is either ready for inversion or in need of minimal manual processing. We conclude that the use of artificial neural network scan significantly reduce the processing time and its costs by as much as 50%.

Key words: Data processing, Electromagnetics.

INTRODUCTION

Currently, it is a very time-consuming task to process large airborne electromagnetic surveys conducted in areas with abundant man-made infrastructure: power lines, fences, railways, etc. Couplings to such infrastructure influence the measured signal to a degree that cannot be corrected for (Kirsch 2006), making it necessary to completely remove coupled data from

the processed dataset. Whereas most of the steps in an airborne transient electromagnetic (ATEM) processing workflow can be automated (Auken *et al.* 2009), one has to manually inspect all data in order to remove data that experience coupling. Leaving coupled data in the dataset will, in most cases, create artificial conductive structures that have no geological origin. We argue that these structures should not be discarded by a geologist during the geological interpretation of the inversion result but rather be removed at the source.

*E-mail: kka@geo.au.dk

Describing the signature of couplings in general programmatic terms is difficult, making it equally difficult to develop an algorithm for detecting them. Currently, removal of couplings is conducted by manual inspection of all soundings of the dataset, which can take up as much as 50%–70% of the total time spent on processing the data. Obviously, it would be desirable to automate this part of the workflow in order to save both time and resources.

Rather than trying to automate the process by describing the behaviour of a coupled sounding in programmatic terms, we instead try to produce a collection of coupled and uncoupled soundings and train an artificial neural network (ANN) to recognize the characteristics of each category. ANNs have previously been used in connection with geophysical surveys but not to remove couplings. A review on the different applications of ANNs in geophysics, especially in connection with petroleum exploration, is given by Poulton (2002). ANNs were suggested as potential alternatives to traditional inversion techniques for electromagnetic data as early as the beginning of the 1990s (Raiche 1991) and have recently been used for simple two-layer inversions of ATEM data (Zhu *et al.* 2012). Inversion of vertical electrical resistivity soundings (Singh, Tiwari, and Singh 2005) and controlled source audio-magnetotellurics data (Spichak *et al.* 2002) using ANNs has also been investigated, with the conclusion remaining that information can only viably be extracted for a very limited number of model parameters.

Previously simple, analytic methods have been used to remove couplings from ATEM data, as described by Auken *et al.* (2009). However, since it is hard to describe the characteristics of the couplings in general terms, these methods only remove a subset of the coupled soundings. Reninger *et al.* (2011) used singular value decomposition for the same purpose and concluded that it could make the processing less time consuming and subjective. We use ANN for the same purpose but with a higher level of ambition. Our goal is to create a network so well-behaved that it is capable of outputting data that can readily be inverted, thus significantly reducing the time requirements and inherent subjectivity associated with manual data processing.

This paper is structured as follows: we first briefly describe ANNs and discuss the specific ANN used for our novel data processing scheme. The key elements are the inputs to the ANN and the structure of the network. Having outlined our methodology, we show how the predictive capabilities of the ANN depend on the size of the reference dataset before using the optimal ANN to perform an automatic processing of a large dataset. The result of automatic processing is then

compared with that of a fully manual processing, and we compare the results of inversions of both datasets. Finally, we discuss the advantages and limitations of our approach before presenting our conclusions.

METHODOLOGY

Neural networks

Artificial neural networks (ANNs) are used in a wide range of applications across both science and engineering, e.g., speech recognition, image compression, detection of credit card fraud, and stock market prediction (Hsieh 2009). In this section, we will briefly introduce the concept of ANNs and present the specific implementation used here.

ANNs are inspired by the network found between neurons in the brain and the idea that one can “teach” a network to recognize patterns. The neurons are connected to each other by different couplings and react to signals from each of the other neurons in different ways. The basic building block is the neuron or processing element (PE). In our network, these are organized in layers and the PEs of one layer are connected to all the PEs of the neighbouring layers. For our purpose, we use a simple three-layer structure where all PEs in the first layer correspond to inputs, the second layer is a hidden processing layer, and the third layer has a single output PE.

The mathematics of an ANN is fairly simple and is briefly described in the following. Consider an input vector \mathbf{x} with m elements and n PEs in the hidden layer. The output of the i th PE in the hidden layer is then given by

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

where w_{ij} are weight factors, b_i is a bias, and f_i is the transfer function. The transfer function should be chosen to resemble the characteristics of the underlying problem as closely as possible. For our purpose, the tan-sigmoid function is a good choice:

$$f(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}}. \quad (2)$$

This function smoothly approaches -1 as t goes to $-\infty$ and 1 when t goes to ∞ , which is very well suited for a network of binary predictions. The outputs of the neurons in the hidden layer are sent to the output neuron in the third layer where a weighting is performed and the transfer function is evaluated. For this PE, we also use the tan-sigmoid function since our target will be either -1 or 1 for coupled and uncoupled soundings, respectively.

Having defined the structure of the network, the next step is to determine weights and biases matching the features of a given “training” dataset containing both inputs and targets. This is called supervised learning since one “tells” the network the target values of different inputs and then optimizes the network for this situation. The weights connecting the neurons are iteratively adjusted until good agreement between the output of the network and the targets is found. For this purpose, we use the iterative Levenberg–Marquardt algorithm to correct the weights by the method of back-propagation.

One central criterion in designing a network of good predictive capability is to ensure proper normalization of the inputs with respect to the initial values of the weights while also taking the nature of the problem into account. If the inputs span several orders of magnitude, which is the case for raw airborne transient electromagnetic dB/dt measurements, one cannot expect the optimization process to work well if the weights are initially chosen randomly and of the same magnitude, which is the standard procedure. In this case, the large inputs simply have a higher natural influence on the network output while not necessarily containing more actual discriminatory power. We thus normalize our inputs as described in the next section to achieve good convergence of the optimization.

While the proper choice of normalization is important, the predictive capabilities of a neural network can obviously never go beyond the patterns actually present in the training set. If the network is to “recognize” specific features in the data, one has to supply inputs that actually enable this type of identification and do not contain too many “confusing” inputs without discriminatory power. This also means that one cannot expect an ANN to work well if it is applied to input data differing significantly from what is actually represented in the training set. Obviously, it is also of vital importance that the training set does not contain any inconsistencies, for example, in the form of subjective inconsistent human labelling of grey zone patterns.

Applications in data processing

Our main goal is to make an ANN that can remove so-called capacitive and galvanic couplings (Danielsen *et al.* 2003), which are abundant in densely populated countries such as Denmark. The two types of couplings give rise to two very different signals (see Fig. 1) due to their different origin. In real situations, one can often not consider a sounding affected by only a galvanic or capacitive coupling. Instead one will often see both types of couplings present to some degree.

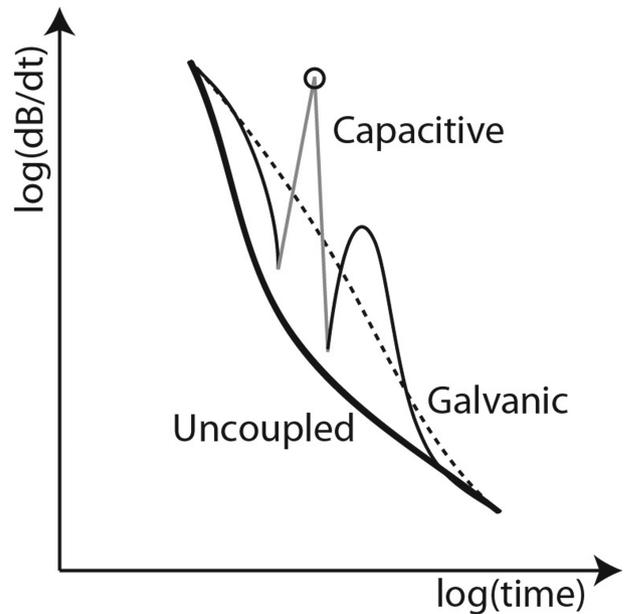


Figure 1 A schematic plot of an uncoupled sounding (fat line), a capacitive coupled sounding (thin line), and galvanic coupled sounding (thin dashed line) on a log-log scale. The capacitive sounding changes sign, indicated by grey connecting line. The negative point itself is shown with a circle.

The capacitive coupling can be described as an LCR circuit in the ground and is fairly easy to detect since it demonstrates large oscillations in the ATEM soundings. A source for such a coupling could be an underground cable, as shown in Fig. 2a). Here the LCR circuit consists of the insulating material around the cable acting as a capacitor, the underground return path of the current acting as the resistor, and the whole circuit as one large inductor. When a current is excited in such a circuit, it will lead to an oscillating and exponential damped current. The oscillations are due to the nature of the LC circuit. A capacitor limits direct currents, and an inductor limits alternating currents. The interplay between these components is therefore an oscillating current. Damping is caused by energy being dissipated in the resistive ground.

Turning to galvanic couplings, these are generally much harder to detect than their capacitive counterparts. The underlying LR circuits do not give rise to an oscillatory signal but only an exponential decay. One common source of these couplings could be, e.g., grounded overhead power lines, as shown in Fig. 2b). The LR circuit consists of the resistive ground and the loop created by the power line mast, the wire connecting two or more masts, and the ground. This type of coupling can only be identified by comparing neighbouring

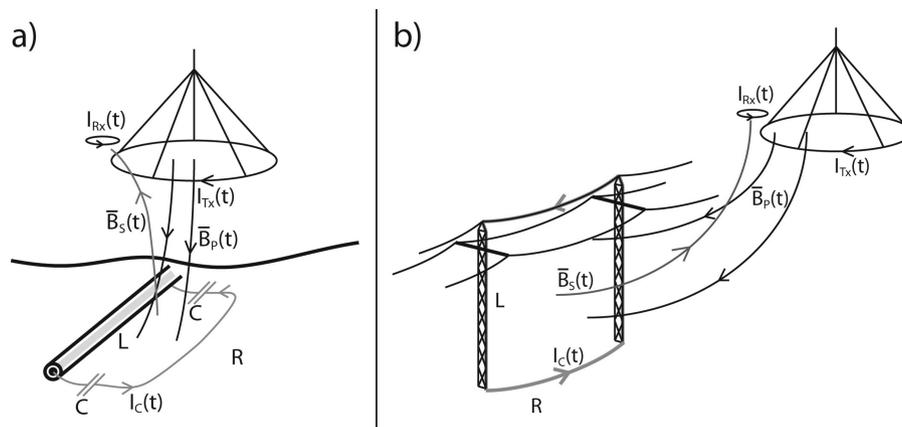


Figure 2 a) A schematic of the source of a capacitive coupling: an LCR circuit generated by the underground cable and the resistive earth. A current in the transmitter $I_{TX}(t)$ generates a time-varying magnetic field $B_P(t)$, which excites a current $I_C(t)$ in the LCR circuit consisting of the underground cable and the earth. The generated current creates a secondary magnetic field $B_S(t)$, which creates a current in the receiver $I_{RX}(t)$. b) A source of a galvanic coupling could be a grounded overhead power line. In this case there is no capacitive effect and the circuit will be an LR circuit only.

soundings and identifying spatial variations deemed too abrupt to be caused by changes in earth resistivity.

In order to find an optimal strategy for automatic detection of both types of coupling patterns, we have conducted numerous experiments in terms of inputs for the network. These included different types of information from both the dB/dt data itself and the instrument navigational sensors. We found that a good choice of ANN input for a given sounding position consists of the raw dB/dt data measured at that very sounding position *and* its two neighbours along the flight line. Thus, an input vector consists of three raw dB/dt measurement vectors stacked on top of each other, with further two elements holding the time derivative of the instrument altitude between the three soundings. Having three neighbouring soundings included in the input, the network will know how the dB/dt signal changes as the measurement position changes. This makes it possible to identify if the variation is well behaved: correlated variations for all gates keep in line with the change in flight altitude. Conversely, if the changes are irregular and chaotic, it will probably be caused by random noise or couplings. Obviously, one can argue that more soundings in the input vector would be an advantage since this would give more information to the ANN. This would also increase the size of ANN parameter space, however, which would result in an unwanted corresponding increase in the required size of training set to get to the same training state. In our experience, three soundings make a good compromise.

The range of input dB/dt data is chosen to only include the range of time gates where couplings can actually be distinguished from the signal and background noise. This im-

plies skipping gates from before the primary field has decayed completely and after the signal becomes dominated by noise. In practice one chooses a time window narrow enough that the signal does not include any significant amount of noise but wide enough that the region where couplings are typically present is fully covered. Since noise levels are affected by factors such as flight height, local geology, and transmitter moment, it is not possible to avoid noise under all circumstances. It is therefore important to ensure that the training set includes data with a similar amount of noise as the data that one intends to use the ANN on. For the SkyTEM system (Sørensen and Auken 2004), which utilizes dual transmitter moments, we found an optimal signal range to be from 20 μs to 180 μs for the low moment and 0.1 ms to 3.5 ms for the high moment. The ground resistivity is mostly between 10 Ωm and 100 Ωm for the training data. We choose to train separate networks for each transmitter moment for generality. This approach also allows for discarding only half the data points of a sounding when only one of the transmitter moments is actually affected by coupling.

Input normalizations

The range of dB/dt measurements within the time interval selected for ANN input spans several decades, which makes it necessary to carefully normalize the data. We normalize by

$$x' = \text{sign}(x) \log|x|, \quad (3)$$

where $\text{sign}(x) = 1$ for $x > 0$ and $= -1$ for $x < 0$. This normalization retains information about both signs of the signal

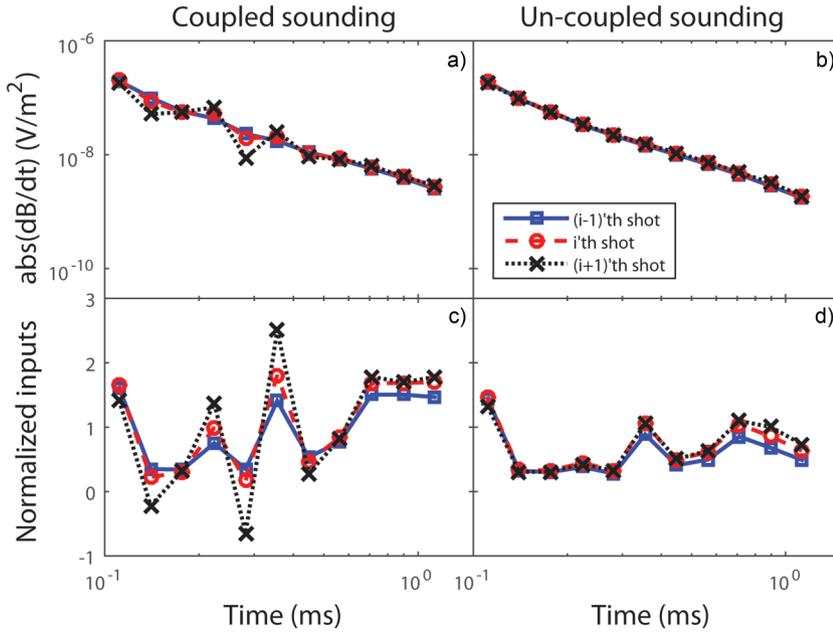


Figure 3 Input data for a coupled (left plots) and uncoupled (right plots) sounding before (upper) and after normalization (lower). The height inputs are not shown. Note that the unnormalized decays span a factor of a least 100 while the normalized decays span the interval of 0–2.

and makes x' of the order of 1 since we further prepare the input vectors by subtracting its mean and dividing by its standard deviation. In other words consider all the inputs as a matrix where each row is an input vector containing the flight height time derivatives, dh/dt , and the x' values for three neighbouring soundings. Let m_j be the mean of the j th column and s_j be the standard deviation of the j th column. We normalize an input x'_{ij} by

$$x''_{ij} = \frac{x'_{ij} - m_j}{s_j}. \quad (4)$$

Unnormalized and normalized inputs are shown in Fig. 3 for both coupled and uncoupled data. From this figure it is clear how the inputs are transformed from having a dynamic range on the order of a factor of 100 into a signal of the order of 1, which facilitates training with random initial weights.

While the target values of the training set are discrete (-1 or 1 for coupled/uncoupled), the output of the trained network can be any number between -1 and 1 . We therefore have to determine the optimal “cut-off” value separating coupled from uncoupled soundings. This is done after the training process by finding the value that maximizes the agreement between the network output and the target. In Fig. 5 we show histograms of targets and output, with a superimposed line indicating the optimal cut-off. These figures will be described in more detail in connection with the field example.

FIELD EXAMPLE

To demonstrate the capabilities of our technique, we now apply it to a SkyTEM dataset acquired by the German/Danish border in 2008–2009. The survey covers an area of 730 km^2 with 3327 km of flight lines. Full survey details are provided by Jørgensen *et al.* (2012). The dataset consists of multiple sub-surveys that were originally processed by different parties, thus containing small subjective differences in the processing patterns. In order not to let the different subjective personal opinions of the original data processors affect the network, we decided to reprocess a small portion of the survey for our training set and do it as consistently as humanly possible.

Figure 4 shows the layout of the survey flight lines north of the Danish border, with the position of the reprocessed data marked in blue. The position of the data for reprocessing was chosen to cover all geologically distinct settings of the survey area, which is essential in order to obtain a well-trained network. In total we reprocessed around 10% of the survey, corresponding to ~ 16000 soundings. These data were randomly divided into three datasets, where one contained 70% of the data and was used for training and the second contained 15% and was used to validate the artificial neural network (ANN) result during training. The remaining 15% was used to test the ANN after the training had terminated. Using a division in three distinct bins of training data ensures that the network is not “over-trained” to a very specific subset of data but performs well on the more general patterns. Primarily, we

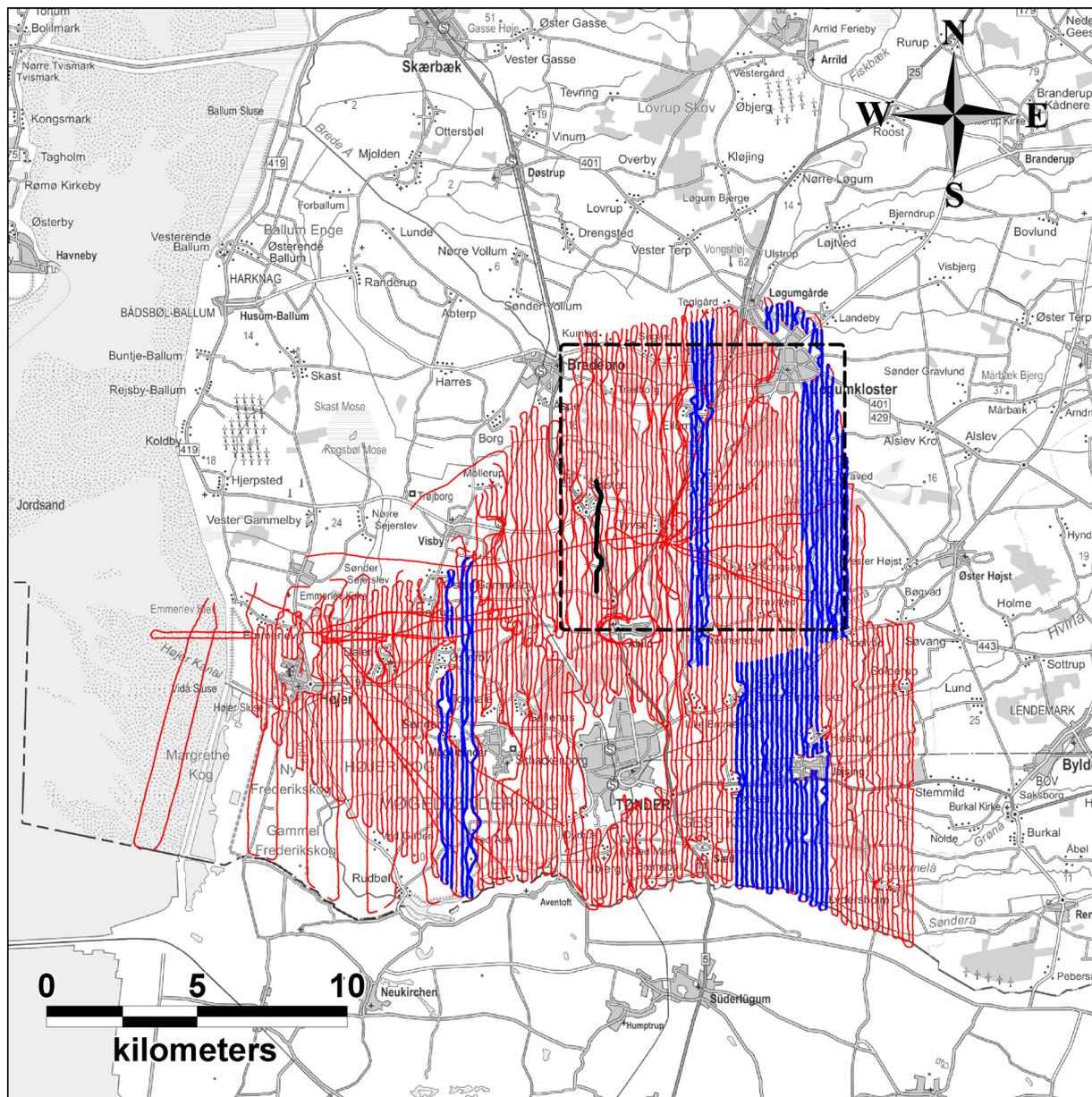


Figure 4 Full flight lines of the Danish part of the survey. The blue lines are reprocessed data, the black line indicates the data shown in Fig. 8, and the black square indicates the data used in Fig. 10.

use ANN configurations ranging from 15 to 35 neurons for the networks and around 35 inputs. With fewer neurons the network will not be able to distinguish all the relevant characteristics of the couplings, and we therefore get bad agreement between the manual processing and the network. When the number of neurons is higher than ~ 35 neurons, we generally do not see any improvement in the agreement. We avoid networks with excessive neurons in the hidden layer since the network training also becomes slower with an increased num-

ber of neurons. The number of inputs depends on the number of gates included per sounding, which can vary and is not necessarily equal for the two moments. If there are many inputs, it will generally be a good idea to use many neurons since there will also be more characteristics in the data.

The output of a 30-neuron network on high-moment data is shown in Fig. 5. The blue bars are the manual processing and the red lines are the output of the network. The black dashed line indicates the optimal cut-off. The output of the

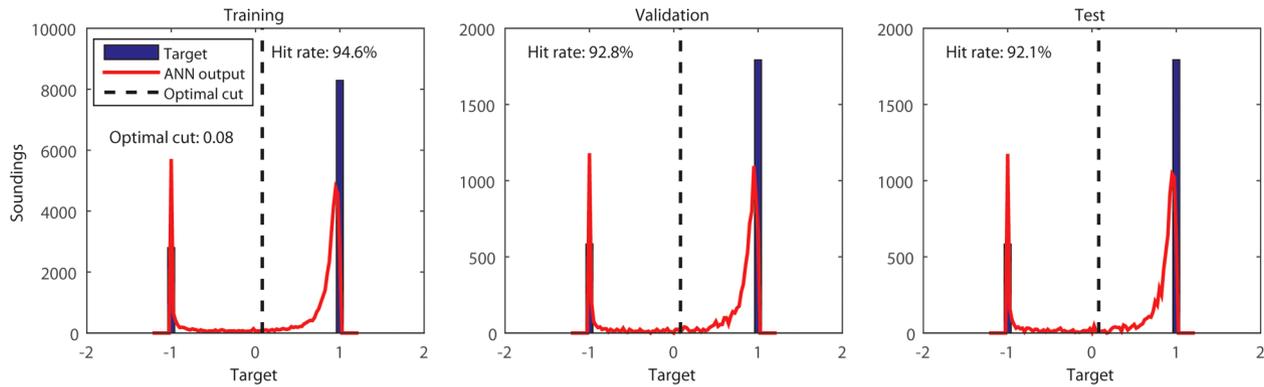


Figure 5 The training state of a 30-neuron network trained on high-moment data. The training data are used to train the network. For each training iteration, the network is evaluated on the validation data and finally tested on the test data after the training has terminated. The hit rate is the percentage of ANN outputs that agree with the target.

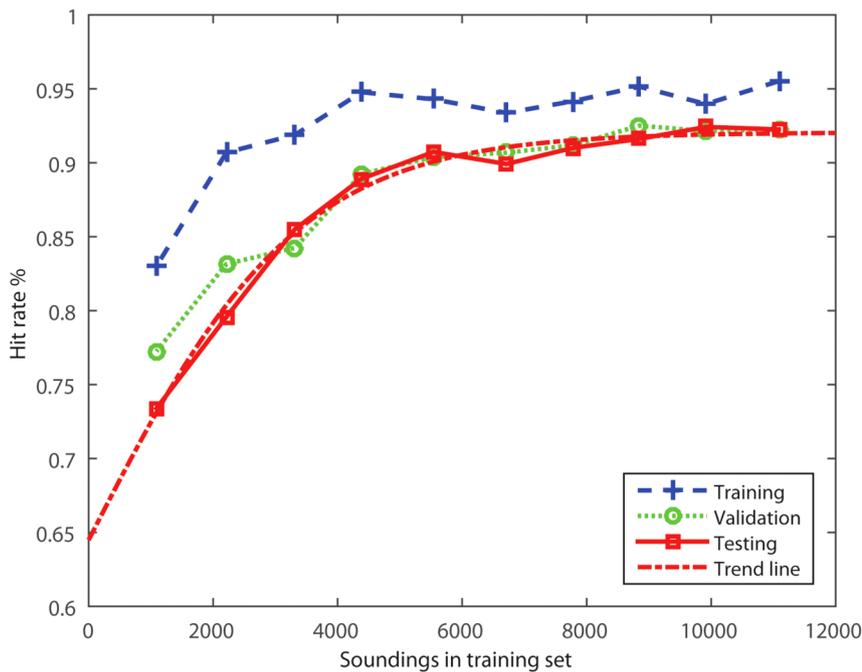


Figure 6 The hit rate of a 30-neuron ANN depending on the size of the training set for high-moment data. A trend line has been fitted to guide the eye.

network has clear and distinct peaks at -1 and 1 as expected, and the distributions of the three training datasets are almost identical. Generally the coupled peak at -1 is narrower than the uncoupled at 1 , suggesting that it is easier for the network to recognize a coupled signal compared to a clean earth signal. The small number of soundings falling in between the peaks can be considered as “grey zone” soundings. If one has a network that is not as well trained, it will show up as more “grass” in between the two peaks where the network is unable to determine if a sounding is either coupled or uncoupled. Furthermore, the number of soundings in the region between the two peaks is also a sign of possible inconsistencies in the

manual processing used to train the network; if the training data contain soundings that are similar but marked differently, it will be difficult for the network to determine the output and it will most likely have an output around 0 in the “grey zone”. Generally we find hit rates (percentage of ANN outputs equal to the target) above 90% for the high moment and around 85% for the low moment.

Before applying the trained network to the actual field data, we first validate whether the training set is of an optimal size or if the network would benefit from a larger training set. In Fig. 6 we show the hit rate of a 30-neuron ANN for high-moment data as a function of the number of soundings used

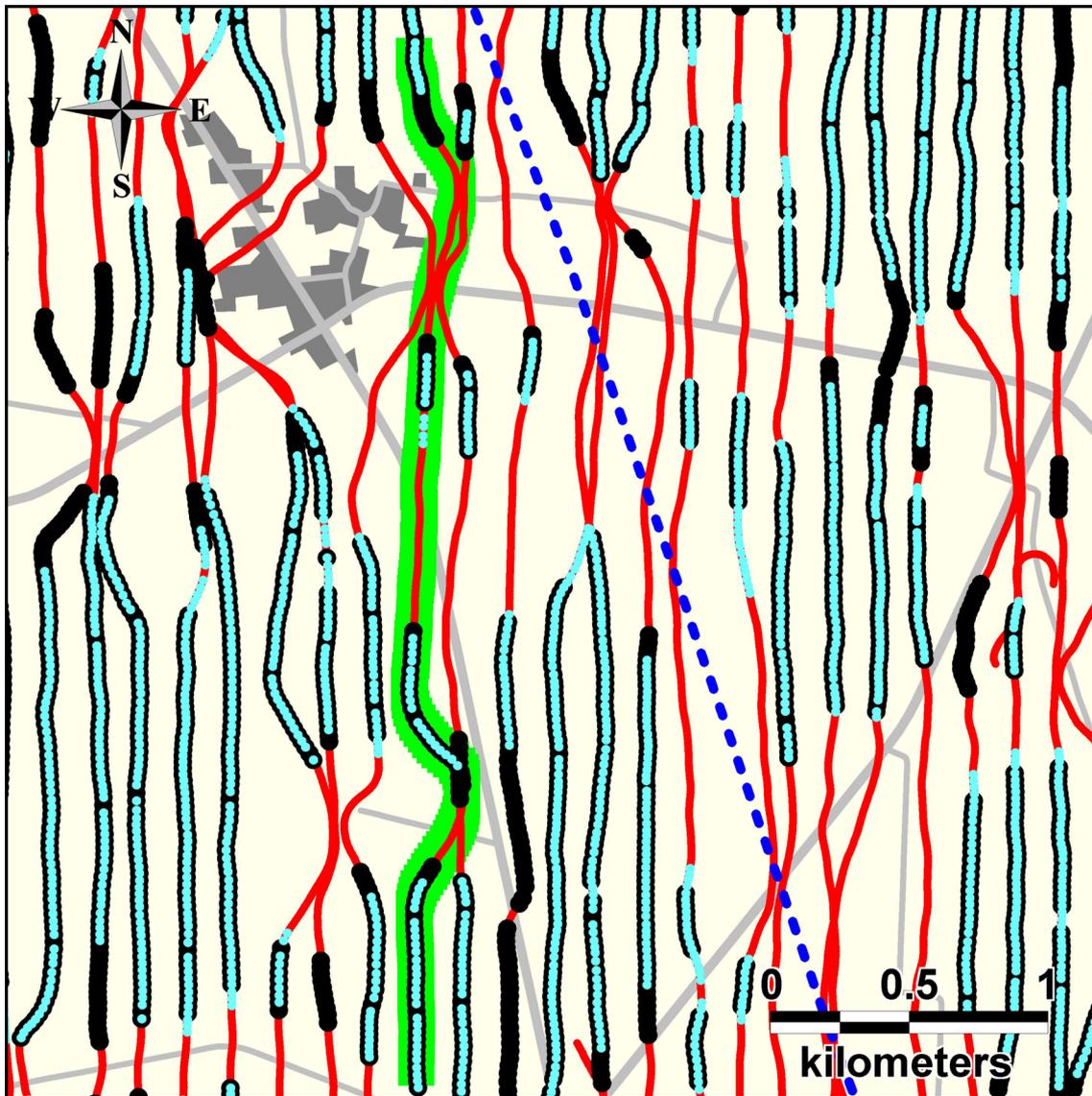


Figure 7 Sounding positions for the high-moment data for a subset of the full survey. Roads (light grey) and towns (dark grey) are marked. The blue dashed line is a power line, the red lines are the flight lines, the black dots show where the ANN finds uncoupled soundings, the blue small dots are the manually processed uncoupled soundings, and the green line is the line shown in Fig. 8. One sees that the ANN generally accepts slightly more data than the manual processing but otherwise removes the same data.

for the training. From this figure it is clear that the hit rate converges around 8000 training soundings, corresponding to approximately 230 km of flight lines. Beyond this point the improvement of the hit rate is very small when the size training dataset is increased. In other words, one can create a custom ANN for a new survey by processing a small representative sub-portion of the total survey area. A similar result is found for the low moment.

Having trained the ANNs for the two transmitter moments and verified their validity, we apply the networks to

the full survey dataset. In Fig. 7 we show a detailed view of a part of the survey, with flight lines (red lines), good soundings based on human opinion (blue dots), and good soundings flagged by ANN (black dots). Apart from the very good overall agreement, we observe several interesting features. The blue dashed line going north–south in the figure indicates a power line, and both processing schemes succeed in removing soundings in close proximity. Furthermore, both schemes also successfully remove data around almost all roads and close to towns. Overall the ANN processing accepts slightly more data

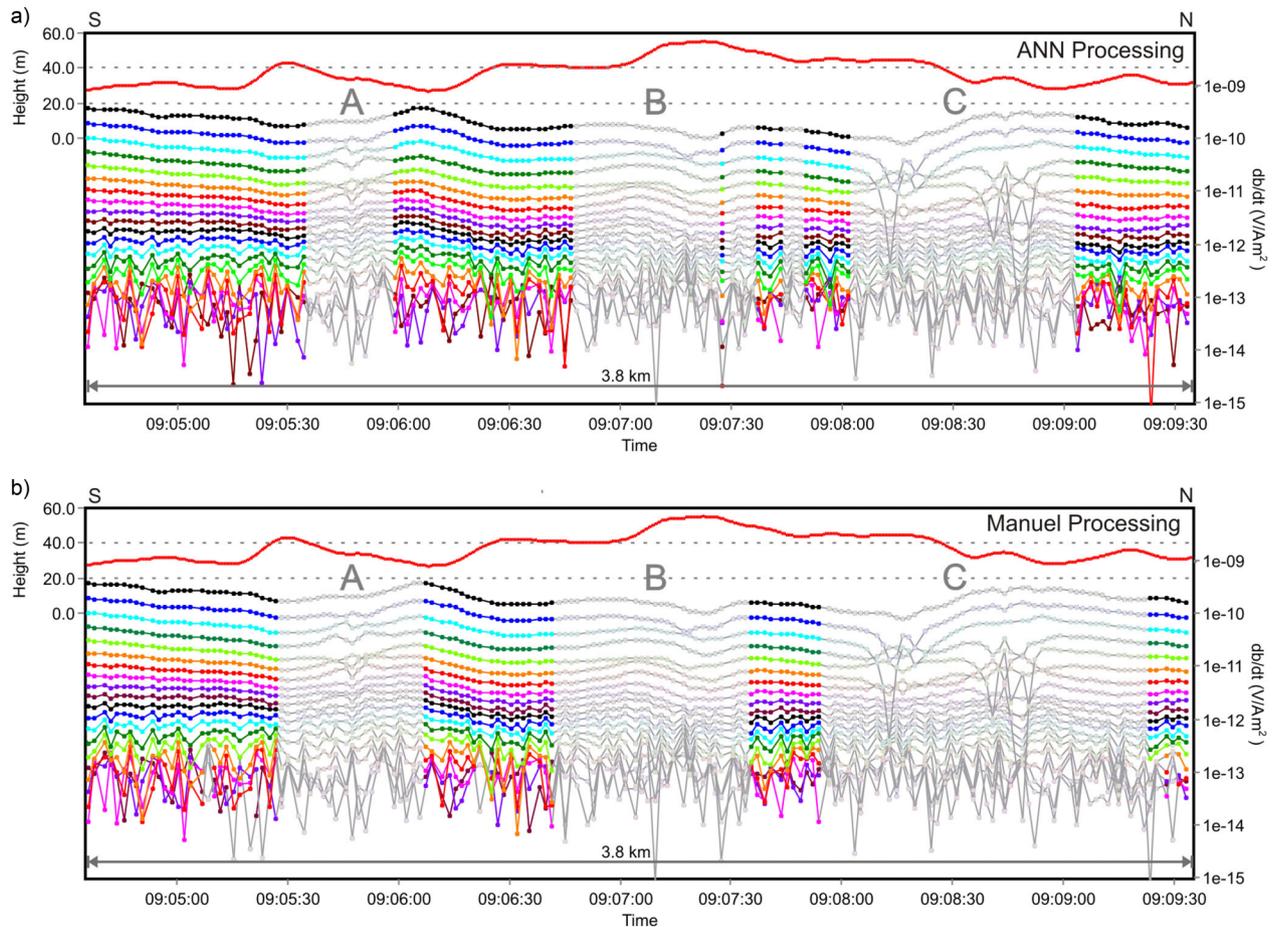


Figure 8 High-moment data sections for the green line in Fig. 7 flying from south to north. a) The output of the ANN and b) the manual processing. In the top of each figure, the flight height is shown in red. There are three large regions (A, B, and C), which have been removed in both cases but there are slight variations in the extent of the exclusion regions.

than the manual processing, which is also seen in several places in the figure. As an example, the manual processing has a clear tendency to exclude more data around roads than the ANN, especially from flight lines that are close and parallel to roads. The reason for this might be that, when the angle of the flight line compared with the road is very small, the signal changes (especially galvanic couplings) very slowly, which is hard for the network to detect, since it only sees three soundings at a time. On the other hand a human inspection of the data allows inspecting many neighbouring soundings at a time, utilizing a GIS map with superimposed maps of roads and other infrastructure. Therefore, the network is not always able to remove the couplings that show up as very small perturbations to the signal, which are also difficult to detect by human inspection.

The high-moment soundings of the flight line marked in green in Fig. 7 are shown in a gate-by-gate data section in Fig. 8. In this figure there are three clearly coupled data re-

gions (marked A, B, and C), where data have been removed by both the manual and ANN processing. From the overview map in Fig. 7, it is evident that these all correspond to roads. The A and C regions contain clear signatures of capacitive couplings, given the large fluctuations and spikes in the data section. The B region features much smoother data from a large galvanic coupling (overlain by signatures of a capacitive coupling), given rise to a smooth signal increase in proximity of the road. In this case, the ANN processing has included a single sounding in the manual exclusion region and excluded three soundings where the manual processing has accepted them. One also clearly sees that the exclusion regions are narrower for the ANN processing, which means that more data are left for the inversion. Obviously, this is only an advantage as long as the data can actually be considered uncoupled.

To investigate the raw data section in Fig. 8 in more detail, we have conducted multiple inversions of this data using

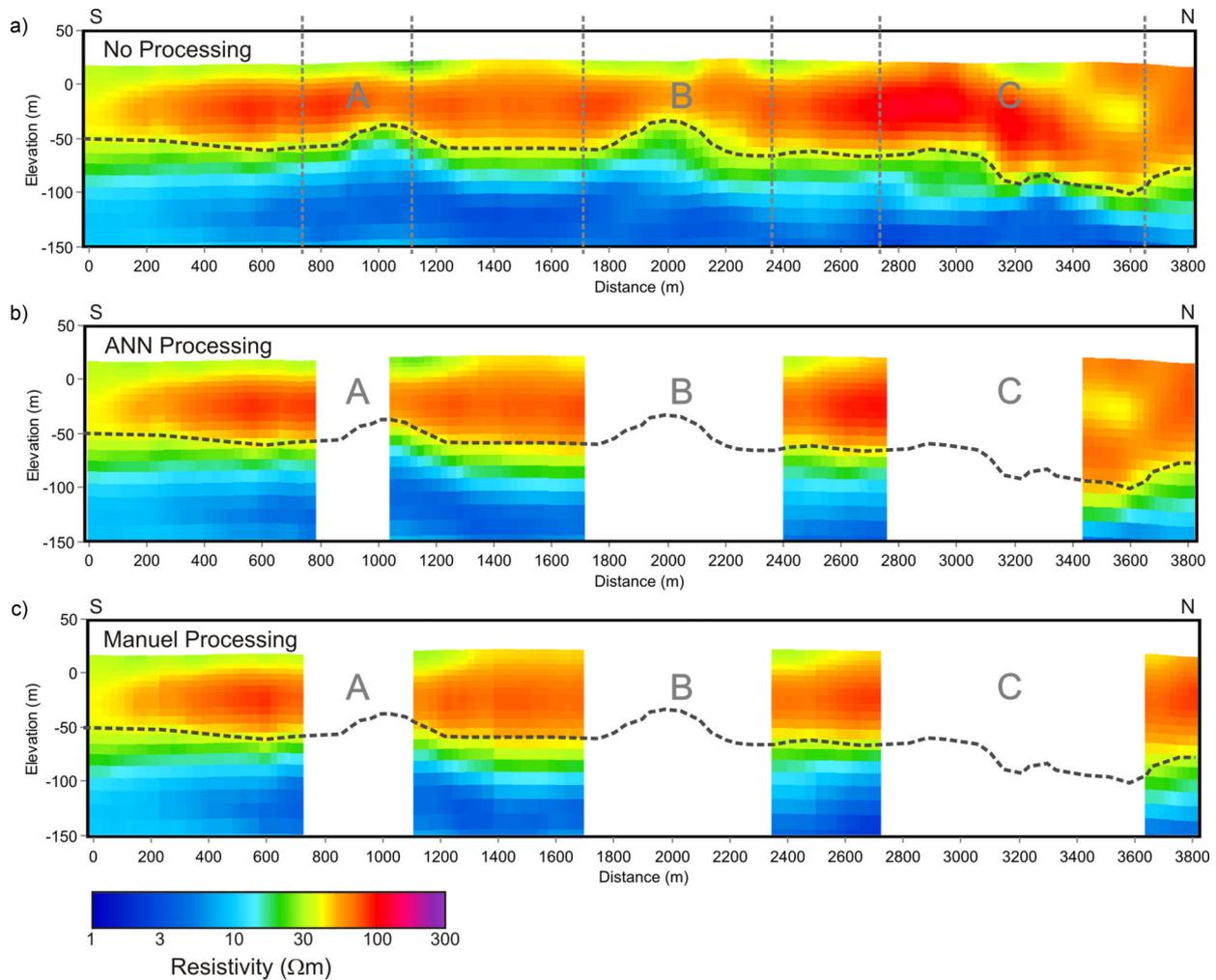


Figure 9 Model sections based on a temporal running mean of the raw high-moment data shown in Fig. 8. a) The inversion result based on the full dataset where couplings have not been removed. b) The inversion based on the ANN processing. c) The inversion of the manually processed data. The dashed line indicates the boundary between the bottom conductive layer and the upper resistive layer for the data without processing. This line is also drawn in b) and c) for better comparison. The vertical dashed lines in a) indicate the boundaries of regions A, B, and C that are excluded by the manual processing.

identical settings. The inversions are performed on temporal running means of the raw data, however, since the signal-to-noise ratio of each individual raw sounding is much too low for a meaningful inversion (Auken *et al.* 2009). The results are presented in Fig. 9 where we show the model sections corresponding to the raw high-moment data in Fig. 8. Note that there is no direct 1:1 correspondence between the sounding positions of two figures since one operates directly in raw data space and the other on temporal averages. Figure 9a) shows the model section obtained from data experiencing no processing at all. For comparative purposes, we have superimposed a dashed line marking the boundary between the conductive bottom layer and the resistive top layer. In regions A and B,

the elevation of this boundary increases by 20 m–30 m, thus showing large deviations from the otherwise flat boundary. In C, the boundary is also found to be somewhat uneven. In Fig. 9b) and c), we show the same results for ANN and manually processed data, respectively. The superimposed dashed line is identical to that in Fig. 9a), allowing for easy comparison of the differences. It is clear that the regions where the boundary line is rapidly changing have mostly been excluded by the ANN processing and almost completely excluded by the manual processing. In A, the coupling shows up from 900 m to 1200 m. The ANN excludes only half of this region, and the manual processing excludes it almost completely but not fully. In region B, both b) and c) remove the coupling

completely, which is quite impressive since this is also the most galvanic coupling, and these are generally harder to detect than the capacitive couplings. In region C a very large amount of data is removed in the manual processing. In the ANN processing, most of these data are removed but there are some models from 3400 m to 3600 m. In this interval one also sees a 50- Ωm region at a depth of 40 m in the 100- Ωm layer which is caused by a small coupling. This small region has not been removed by the ANN. The elevation of the boundary line also changes by around 30 m at 3600 m. It is difficult to determine if this change is present in c) or not.

Finally, in Fig. 10, we compare mean resistivity maps at 30 m–40 m of depth for the area indicated in Fig. 4 based on data with no processing a), an ANN processing b), and a manual processing c) for both high- and low-moment data (for geological interpretations, we refer to Jørgensen *et al.* (2012)). The main feature of the map is a large high-resistivity region in the southern part and a low-resistivity region in the north. The general shape of these regions appears virtually identical for all three inversions; however, looking into details, many small discrepancies can be identified. Starting with unprocessed data in a), the resistivity is very uneven and has many small low-resistivity regions spread over the entire map. These regions often follow power lines, roads, or other infrastructure and are caused by coupled data, e.g., the power line going east–west in the upper part of the map where the resistivity drops to less than 3 Ωm . The inversion based on ANN processed data b) is significantly more even than a) but still experiences some coupling artifacts. For example, one can still see a small low-resistivity region along the power line and the resistivity drops from 100 Ωm to around 50 Ωm along a road in the southwestern corner of the map. Finally, the manual processed data c) result in the smoothest inversion; however, one still finds regions that might be questionable. The small discrepancies between b) and c) do not alter the general good agreement between the two inversion results.

Discussion

In this paper we train and apply our ANN in a very local context where the survey conditions in terms of flight height, signal, and noise level are reasonably constant. We found that such a local network is not readily applicable to surveys conducted under entirely different conditions, which was also not expected. However, one should be able to improve the applicability of an ANN incrementally by adding more data from different locations to the training set. We have investigated this by adding reference data to the training set from

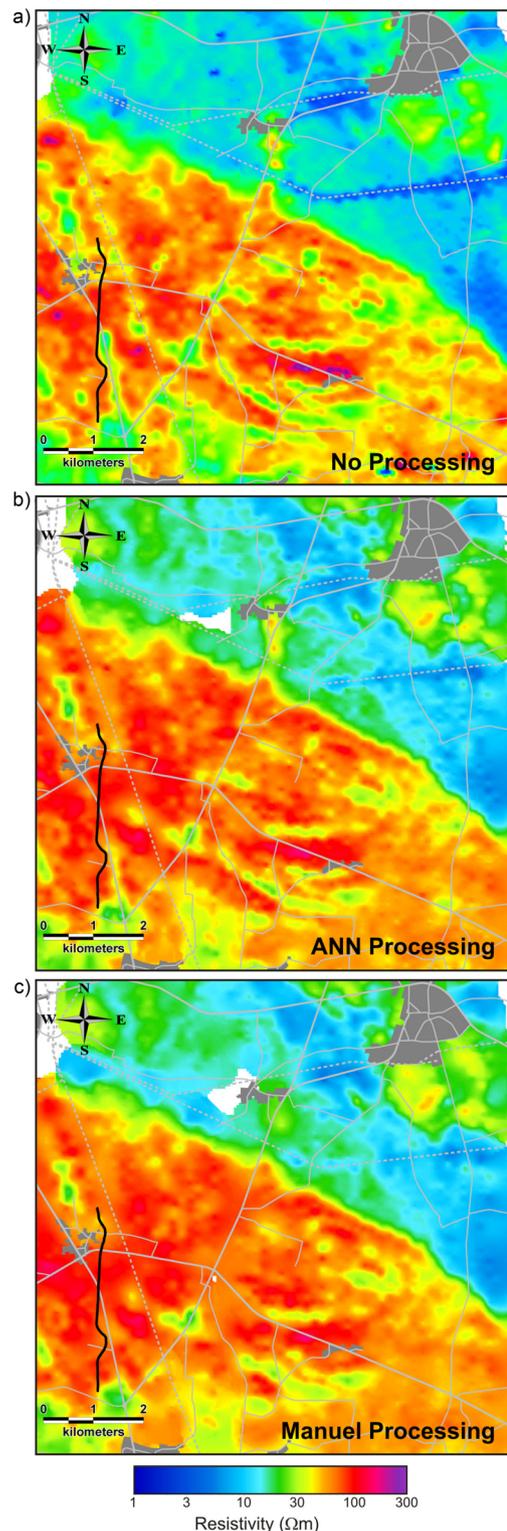


Figure 10 Mean resistivities at a depth of 30 m–40 m based on dual moment data a) without processing, b) ANN processed data, and c) on manually processed data. In the background is a card of the area showing roads, towns, and power lines (grey dashed lines).

an entirely different survey. The resulting combined network worked equally good on both datasets, and we thus feel confident that one can incrementally increase the input parameter space covered by the network by adding more and more data to the training set. For the network presented here, we chose to include three neighbouring soundings for each input vector to the network. If one is only interested in capacitive couplings, one can neglect both the height information and the neighbouring soundings since the distinct signature of capacitive couplings can be identified on a single sounding level. However, for the galvanic couplings, one has to include more soundings since this coupling develops over several soundings and can only be identified by comparing several neighbouring soundings. One could consider using even more soundings in each input vector. The disadvantage of this is that the input parameter space is also drastically increased, which affects the amount of data needed to train a network.

CONCLUSION

We have shown that one can successfully train an artificial neural network (ANN) to recognize couplings to man-made conductors in airborne transient electromagnetic (ATEM) data in order to automatically remove them before conducting an inversion. In order to get the best performance of the network, we had to normalize the data to values of the order of 1 and use 3 dB/dt soundings and the change in the flight altitude as an input. Measuring the performance of the network in terms of how well it reproduces a human processing (the hit rate), we are able to get a hit rate above 90%.

From our experiments, we found that a network needs to be adapted to the local survey and geological conditions. This is not ideal, but we have shown that, in large surveys, one can manually process a small subset (200-km flight lines corresponding to one to two flights) of the full dataset, train a network, and successfully apply it to the remaining data. This can significantly reduce the labour-intensive task of manually processing ATEM data and hence can reduce the data processing costs by up to 50%. Furthermore, it seems possible to continuously improve the applicability of a network by adding more data to the training set, eventually leading to a versatile and more general network applicable under any conditions.

ACKNOWLEDGEMENTS

The authors acknowledge the Villum Center of Excellence HOBE for providing partial funding for Casper Kirkegaard.

REFERENCES

- Auken E., Christiansen A.V., Westergaard J.A., Kirkegaard C., Foged N. and Viezzoli A. 2009. An integrated processing scheme for high-resolution airborne electromagnetic surveys, the SkyTEM system. *Exploration Geophysics* **40**, 184–192.
- Danielsen J.E., Auken E., Jørgensen F., Søndergaard V.H. and Sørensen K.I. 2003. The application of the transient electromagnetic method in hydrogeophysical surveys. *Journal of Applied Geophysics* **53**, 181–198.
- Hsieh W.W. 2009. *Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels*. Cambridge University Press.
- Jørgensen F., Scheer W., Thomsen S., Sonnenborg T.O., Hinsby K., Wiederhold H. et al. 2012. Transboundary geophysical mapping of geological elements and salinity distribution critical for the assessment of future sea water intrusion in response to sea level rise. *Hydrology and Earth System Sciences* **16**, 1845–1962.
- Kirsch R. 2006. *Groundwater Geophysics - A Tool for Hydrogeology*, 1st edn. Springer.
- Poulton M.M. 2002. Neural networks as an intelligence amplification tool: a review of applications. *Geophysics* **67**, 979–993.
- Raiche A. 1991. A pattern recognition approach to geophysical inversion using neural nets. *Geophysical Journal International* **105**, 629–648.
- Reninger P.A., Martelet G., Deparis J., Perrin J. and Chen Y. 2011. Singular value decomposition as a denoising tool for airborne time domain electromagnetic data. *Journal of Applied Geophysics* **75**, 264–276.
- Singh U.K., Tiwari R.K. and Singh S.B. 2005. One-dimensional inversion of geo-electrical resistivity sounding data using artificial neural networks - case study. *Computers & Geosciences* **31**, 99–108.
- Sørensen K.I. and Auken E. 2004. SkyTEM - A new high-resolution helicopter transient electromagnetic system. *Exploration Geophysics* **35**, 191–199.
- Spichak V., Fukuoka K., Kobayashi T., Mogi T., Popova I. and Shima H. 2002. ANN reconstruction of geoelectrical parameters of the Minou fault zone by scalar CSAMT data. *Journal of Applied Geophysics* **49**, 75–90.
- Zhu K.G., Ma M.Y., Che H.W., Yang E.W., Ji Y.J., Yu S.B. et al. 2012. PC-based artificial neural network inversion for airborne time-domain electromagnetic data. *Applied Geophysics* **9**, 1–8.