Estimating noise in AEM data

Aaron Davis
CSIRO
26 Dick Perry Ave, WA, AUSTRALIA, 6151
Aaron.davis@csiro.au

SUMMARY

In this paper, I discuss a method to obtain reliable noise estimates for airborne electromagnetic (AEM) surveys based on the reversible-jump Markov chain Monte Carlo method. In addition to estimating electrical conductivity and thickness using 1D layered-earth models, the method provides estimates of the additive error required to make all measurements of a repeat line agree. The noise estimates can also be obtained from a single line where repeat line information is unavailable. The resulting additive noise estimates then can be used in a general deterministic inversion. Analysis of inversions shows that model regularisation has little effect at depths where the data is informative. This improves the reliability of the inverted models, since it is the noise-adjusted data which is informing the model.

Key words: airborne electromagnetics, noise, RJMCMC.

INTRODUCTION

The airborne electromagnetic (AEM) method was originally developed for use in mineral exploration whereby an explorer searched the survey data for signals that indicated prospective mineralisation targets. More recently, AEM is increasingly applied to near surface investigations for groundwater exploration, hazard identification and geological mapping. Products from AEM surveys for the newer applications usually include conductivity-depth inversions generated from the collected survey data. Inversion of AEM data is a complicated task due to incomplete information, the difficulty of the inverse problem to be addressed, different systems offered for service, and estimation of the utility of the data for the purpose for which the survey was flown. With increased focus on the near surface for resources such as groundwater, the inversion task needs to address resolution of conductive features in the ground.

Inversion of AEM data for electrical conductivity models involves many factors that need to be addressed. Some factors must be addressed prior to the survey being flown and are governed by such things as system availability, cost of acquisition, and the goal of the study. Other factors are part of the inversion problem itself. These include: the description of the AEM system to be modelled, discretisation of the model used to estimate conductivity, the choice of inversion algorithm, and the amount of regularisation needed to ensure reasonable convergence. Perhaps one of the most important factors that needs to be mentioned, and one that is closely tied to model regularisation, is estimates of noise in the AEM survey data. Noise is any unwanted signal that interferes with the electromagnetic signal transmitted and received by the AEM sensor. It can be generated by various sources such as man-made structures, power lines, lightning, and atmospheric disturbances, but can also be caused by the variation in the earth's natural electromagnetic fields and the geology of the survey area. Different types of noise can be present in AEM data, but they are generally classified as random and systematic noise. Systematic noise, caused by a specific source that produces a consistent pattern of interference in the AEM data, can often be corrected by identifying the source and applying appropriate correction methods. Random noise is caused by the statistical variation of the electromagnetic signal received by the AEM sensor. This noise is difficult to remove as it is not correlated with any causes.

Accurate estimation of noise levels in AEM survey data is necessary since they directly influence the accuracy and reliability of the data. This can have a profound impact on our interpretation of the subsurface geology. Despite the importance of obtaining accurate estimates of data noise, there is little in the literature that describes how we obtain them. The purpose of this paper is to address this apparent shortcoming.

DETERMINING NOISE IN AEM SURVEY DATA

It can be difficult to determine the noise characteristics of AEM survey data since the effective use of many of the systems available depends upon them being airborne to operate effectively. That is, there is little opportunity to operate them at full capacity while on the ground. To address this, many contractors now offer measurements of high-altitude data. With the transmitter operating at full capacity, the AEM system is taken to extreme altitude in the assumption that the signals measured by the receiver coils will be uninterrupted by earth responses. These measurements offer an understanding of the noise floor of the receiver assembly: the measurements of the receivers in the absence of any signal other than the system itself. This is a good first step to understanding the bias of the system. An example of high altitude noise recordings is shown with the solid lines in Figure 1.

Green and Lane (2003) suggest a different strategy for estimating noise in survey data through use of repeat lines. The assumption here is that the system should always measure the same responses over the same survey transect. They recommend characterising noise as either additive or multiplicative in nature, meaning that noise levels for a given delay time are composed of some base level of noise plus some factor multiplied by the signal itself at that delay time. This can be written as

$$\sigma_i = \sqrt{\sigma_{iadd}^2 + (\sigma_{im} \cdot d_i)^2},$$

where σ_i is the noise at delay time i, d_i is the measured data for that delay time, and σ_{iadd} and σ_{im} are the additive noise term and the multiplicative factor, respectively. It should be noticed that the noise term σ_i enters the data misfit equations

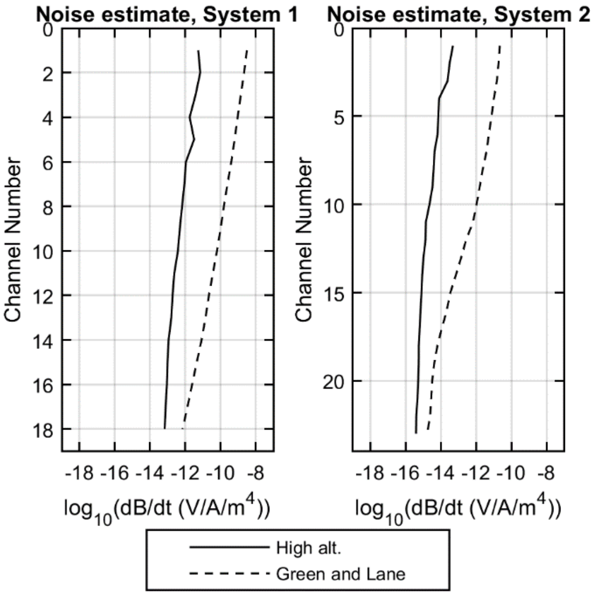

Figure 1. Noise estimates from an AEM system. Solid black lines show the high-altitude measurements, while the dashed lines show the estimates from a Green and Lane (2003) analysis of repeat lines.

as an additive term when used in this manner. An example of an additive noise estimate following this method is show with dashed lines in Figure 1.

One of the drawbacks of the method of Green and Lane (2003) is that repeat lines are impossible to replicate exactly due to the platforms being airborne. Differences in altitude can have a profound effect on the measured response. Another drawback is that many older surveys do not have repeat lines flown (or are not available as part of the delivered data); so, a compromise must be sought. In my approach, I assume that while repeat lines may not have the same measured responses due to variations in acquisition, they should have the same earth *and noise* model provided the repeat lines are flown reasonably close together. At each station of the repeat lines, differences between the recorded data and the forward response for each station can therefore be classified as 'noise'. Noise in this sense incorporates variations in measurement at each station, but also encompasses the choice of model used to determine the earth response.

Reversible-jump Markov chain Monte Carlo

To achieve estimates of electrical conductivity distribution and noise for the repeat lines, I employ the Reversible-jump Markov chain Monte Carlo (RJMCMC) method described by Green, (1995) with a few modifications similar to those employed by Minsley et al., (2021), but with some modifications.

We begin by ensuring that sampling of the repeat line data is consistent across a regular spacing along the survey line. The simplest way of doing this is by taking stations from each repeat line that are close enough that we can assume they are measuring the same volume of the earth (eg, Reid et al., 2006) or by resampling the data to a regular spacing. For every station, we create a 1D layered-earth model of variable electrical conductivity layers (and thickness) and a common noise estimate for every delay time of the system. Using $\boldsymbol{m}_s$ to describe the model at station s, the model is composed of k layers of resistivity $\boldsymbol{\rho}$ with thickness $\boldsymbol{t}$ to describe the earth, and n values of $\boldsymbol{\sigma}$ to describe the additive error applied to the n delay times for the j measurements at location s. Notice that the variables in bold are vectors.

At every iteration in the chain, a new model $\boldsymbol{m'}_s$ is proposed from the previous model $\boldsymbol{m}_s$. The new model is accepted or rejected based on the Metropolis, Hastings, Green (MHG) algorithm according to the following acceptance criterion α

$$\alpha(\mathbf{m}'_s|\mathbf{m}_s) = \min\left[1, \frac{p(\boldsymbol{m}'_s)}{p(\boldsymbol{m}_s)}\frac{p(\mathbf{d}|\boldsymbol{m}'_s)}{p(\mathbf{d}|\boldsymbol{m}_s)}\frac{q(\boldsymbol{m}_s|\boldsymbol{m'}_s)}{q(\boldsymbol{m'}_s|\boldsymbol{m}_s),}|\mathbf{J}|\right],$$

where $p(\boldsymbol{m}'_s)/p(\boldsymbol{m}_s)$ is the prior ratio of the models, $p(\mathbf{d}|\boldsymbol{m}'_s)/p(\mathbf{d}|\boldsymbol{m}_s)$ is the likelihood ratio of the data given the models, $q(\boldsymbol{m}_s|\boldsymbol{m'}_s)\backslash q(\boldsymbol{m'}_s|\boldsymbol{m}_s)$ is the proposal ratio, and $\boldsymbol{J}$ is the Jacobian governing changes between dimensions. Of special interest in this paper is the data likelihood function $p(\mathbf{d}|\boldsymbol{m}_s)$ which will change at every iteration due to choices of perturbations in $k, \boldsymbol{\rho}, \boldsymbol{t},$ or $\boldsymbol{\sigma}$. We write the likelihood function as

$$p(\mathbf{d}|\boldsymbol{m}_s) = \sum_{i=1}^{j} \frac{1}{\sqrt{2\pi|\boldsymbol{C_d}(\boldsymbol{\sigma})|}} \exp\left(-\frac{1}{2}\left((\boldsymbol{f}(\boldsymbol{\rho},\boldsymbol{t},k) - \boldsymbol{d}_{si})^T \boldsymbol{C_d}(\boldsymbol{\sigma})(\boldsymbol{f}(\boldsymbol{\rho},\boldsymbol{t},k) - \boldsymbol{d}_{si})\right)\right)$$

where $\boldsymbol{f}(\boldsymbol{\rho},\boldsymbol{t},k)$ is the predicted data given the model parameters, $\boldsymbol{d}_{si}$ is the measured data at station s for measurement i ,and $\boldsymbol{C}_d(\boldsymbol{\sigma})$ the data covariance matrix that models the error in the system responsible for the measurements. In this paper, $\boldsymbol{C}_d(\boldsymbol{\sigma})$ is assumed to be diagonal.

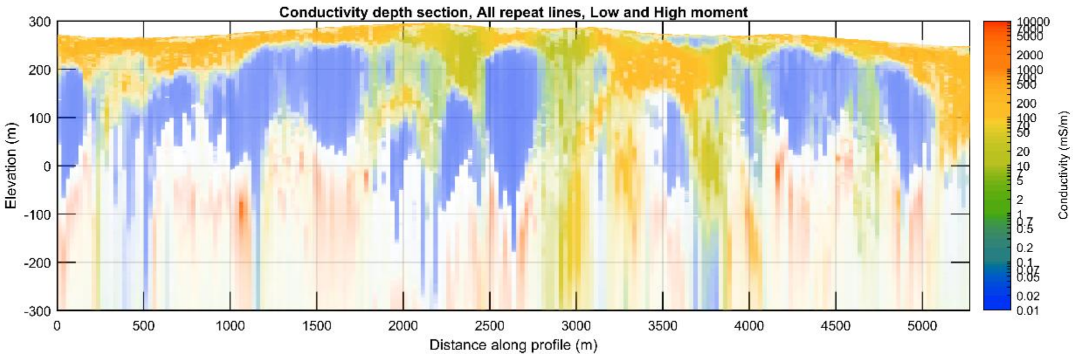

Figure 2. An example of the posterior mean conductivity section resulting from the RJMCMC process. Blanked areas are due to a wide spread of accepted models relative to the prior conductivity.

Model proposals are based on the usual choices for MHG samplers. At every iteration, we choose to: create a conductivity interface, destroy a conductivity interface, change the structure of the existing model (by creating and destroying random interfaces, or vice versa), or changing one of the n noise parameters in $\boldsymbol{\sigma}$. For each station, several chains (8) are run for many iterations (3×10^5). Several thousand models are excluded from the beginning of each of the chains, and the results are accumulated.

RESULTS

Figure 2 shows a posterior mean electrical conductivity distribution transect for the example discussed earlier. Areas that are made transparent reveal that the spread in conductivity of the accepted models is similar to the prior probability of the conductivity proposals (and, therefore, less infomative). The section looks reasonable, and there is clear structure to depths of approximately 300 mAHD. Figure 3 shows the mean distributions of the marginalised noise estimates for every model in the reduced RJMCMC chains, and for every station. The distributions are shown in shaded blue, while the mean additive noise value for each delay time is marked by the solid gold line. The mean values from the gold line are chosen to represent the average additive noise values for the entire survey. Also shown are the high-altitude (black), the Green and Lane (2003) noise estimates (red), and an RJMCMC noise estimate conducted on only one line. Clearly, the RJMCMC noise estimates are consistently higher than the high-altitude noise estimates, but mostly lower than the Green and Lane (2003) estimates.

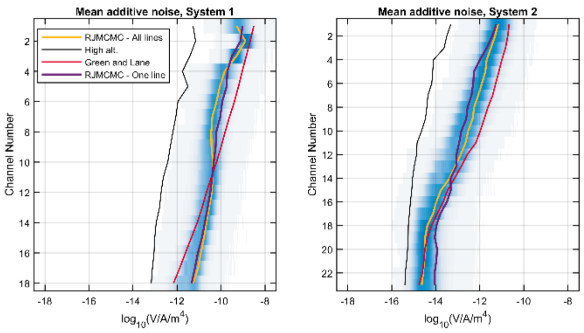

Figure 3. Estimates of additive noise for each delay time of both systems. The blue colour variations show the distributions of the noise estimates. The solid gold line shows the peak of the distributions and is taken as the average measurement noise for the entire survey. The solid black line is from high-altitude tests, the red line is the Green and Lane (2003) estimate, and the purple line is from an analysis conducted with only one repeat line.

What if you don't have repeat lines?

The purple lines in Figure 3 show the result of an RJMCMC noise analysis conducted with only one of the repeat lines. Instead of using all 4 lines to estimate noise, the chains are run using only one measurement per station. Noise is estimated from each model in the chain. We can see that the use of only one line exhibits similar noise levels, indicating this method can be used if there are no repeat lines available.

Effect on regularisation in inversion

Having determined an average noise estimate for the entire survey based on several repeat lines, it is useful to see the effect the noise estimates have on deterministic inversion. Figure 4 shows smooth 1D layered earth inversion models for a wide range of model regularisation values. In this case, an isotropic exponential regularisation with a 25 m correlation length was chosen for the model regularisation; the regularisation matrix was the same and only the weighting was changed. The prior resistivity was chosen to be 10^4 Ωm. A depth of investigation line (DOI) (Christiansen and Auken, 2012) for each inversion is shown in white. The figure shows little variation in inverted conductivity models above the DOI line. This indicates that the model regularisation has little effect in determining model structure where the models are informed by the data, which is precisely what is desired in the inversion.

CONCLUSIONS

In this paper, I have adapted the results of previous authors by using a reversible-jump Markov chain Monte Carlo method that, in addition to yielding conductivity-depth information, provides us with noise estimates for airborne electromagnetic surveys. The noise estimates are shown to be similar to the repeat line method of Green and Lane (2003) but are significantly greater than the high-altitude measurements provided. The additive RJMCMC noise estimates can be applied to the entire survey. The method can be applied when there are no repeat lines available. Analysis of the RJMCMC noise estimates show that model regularisation has little effect at depths where the signal is informative in generating electrical conductivity models.

ACKNOWLEDGMENTS

I would like to acknowledge the CSIRO Deep Earth Imaging Future Science Platform and CSIRO Mineral Resources for support in producing this work.

REFERENCES

Christiansen, A.V. and Auken, E., 2012, A global measure for depth of investigation: *Geophysics*, 77(4), WB171–WB177.

Green, A. and Lane, R., 2003, Estimating Noise Levels in AEM Data: *ASEG Extended Abstracts*, 2003(2), 1.

Green, P.J., 1995, Reversible jump Markov chain Monte Carlo computation and Bayesian model determination: *Biometrika*, 82(4), 711–732.

Minsley, B.J., Foks, N.L. and Bedrosian, P.A., 2021, Quantifying model structural uncertainty using airborne electromagnetic data: *Geophysical Journal International*, 224(1), 590–607.

Reid, J.E., Pfaffling, A. and Vrbancich, J., 2006, Airborne electromagnetic footprints in 1D earths: *GEOPHYSICS*, 71(2), G63–G72.

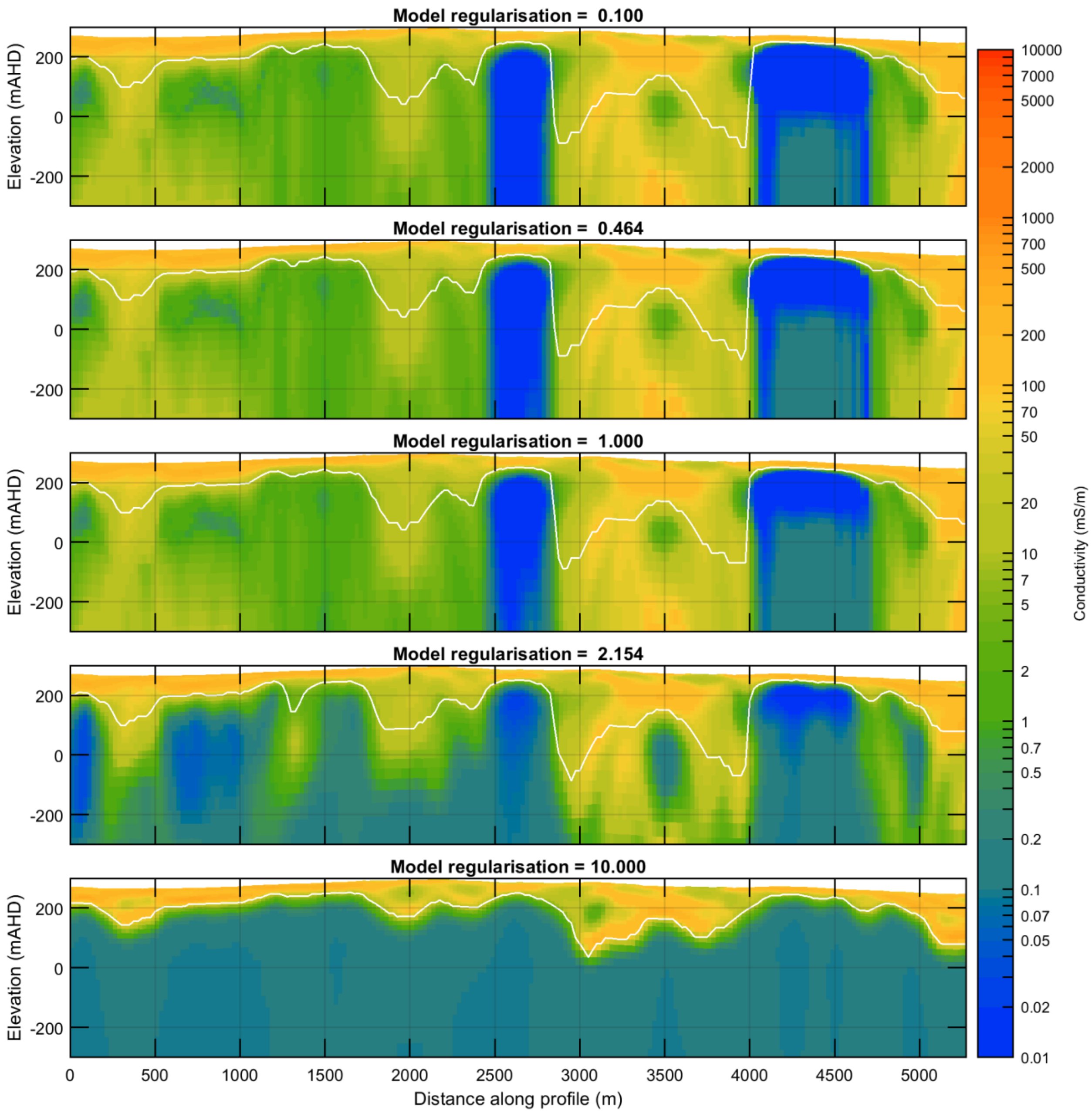

Figure 4. Inversions showing the effect of model regularisation for a wide range of regularisation weighting values. All inversion runs were initialised with the same starting model. The model regularisation structure for each is the same. A depth of investigation (DOI) line is shown in white. There is very little variation in models above the DOI.