Blocking borehole conductivity logs at the resolution of above-ground electromagnetic systems

Aaron Davis¹ and Juerg Hauser²

ABSTRACT

Borehole conductivity logs provide an in situ measurement of the electrical conductivity of the subsurface. Despite the measurements being a proxy for the true earth structure, they are often used as ground truth when inferring subsurface electrical conductivity boundaries between lithologies. Borehole conductivity measurements are therefore commonly used to plan and benchmark electromagnetic (EM) surveys and to establish the credibility of a given inversion technique. A consequence of the diffusion physics of EM prospecting is that not all subsurface features present in a conductivity log can be resolved by an EM system, nor can they be recovered by a subsequent inversion. Quantification of the ability of an EM system to determine layer boundaries in the subsurface is therefore an issue meriting investigation. We have developed a reversible-jump Markov chain Monte Carlo (RJMCMC) method to segment borehole conductivity logs at the scale recoverable by a given EM system as the foundation for an objective comparison between the inversion results and conductivity logs. A common consequence of RJMCMC inversions for EM problems is that few layers are required to fit the data. Similarly, we find that a borehole log blocked at the scale sensed by an EM system consists of a limited number of segments. Segmentation of borehole conductivity logs is determined by the physics of EM prospecting and by factors such as base frequency, number of gates, system geometry, and noise levels. For a survey line intersecting a borehole near Carnarvon, Western Australia, we see that different inversion schemes result in images of the subsurface that are consistent with a borehole conductivity log segmented according to the mechanics of the EM system and accounting for the physics of EM prospecting.

INTRODUCTION

Imaging subsurface geologic features is among the prime motivations for acquiring airborne electromagnetic (AEM) data. An important source of information for a planned AEM survey is available from borehole logs; often, the success of an AEM survey is determined through a comparison of the inverted electrical conductivity model to the recovery of geologic features of interest contained in logs of electrical conductivity and lithologic units. It is therefore imperative to propose adequate measures of comparison and recognize the fundamental differences between AEM systems and downhole induction conductivity logs.

The commonly used induction conductivity logging tools measure a quadrature voltage, which is transformed into an apparent conductivity that represents the equivalent conductivity that a whole space of the earth must have to produce the recorded signal. The voltage recording actually depends on the conductivity distribution of the subsurface and the lines of current that are induced to flow within that distribution (Doll, 1949). For a 1D earth under the low-frequency assumption, apparent conductivity is derived from the voltage through a linear relationship between the logarithm of the signal and the logarithm of the equivalent earth conductivity (Moran and Kunz, 1962). Various efforts have been made to transform the apparent conductivity back to real conductivity (Shen, 1987).
We recognize that it is possible to invert the measured apparent conductivity to an equivalent layered earth model, as is done by Christensen and Lawrie (2018).

When considering the problem of AEM system selection for the determination of the fitness of purpose for a given survey, conductivity and lithology logs are used to build conceptual models. Such models inferred from borehole data then provide the basis for studies seeking to determine if a given AEM system can resolve the salient features of interest (e.g., Brodie and Richardson, 2013; Christensen and Lawrie, 2013). This is often done by subdividing the log into a layered earth model, which is asserted to encompass the true earth model. However, these layered earth models are often computed at the expected layering that most benefits the modeling signals measured by the AEM system.

Borehole logs are also important once the survey data have been collected and preprocessed to be inverted. They contribute to our prior understanding about the geology we seek to recover from the AEM data. In practice, this means that they might influence our choice of model parameterization, inversion algorithm, and starting model. A conductivity log revealing sharp changes in conductivity between lithologic boundaries might motivate an inversion of the AEM data for a blocky model by limiting the number of layers (e.g., Hauser et al., 2015) or by using a sharpening operator (e.g., Vignoli et al., 2015). On the other hand, if the electrical conductivity of the borehole log shows gradual changes, this could indicate that an inference of a vertically smooth model (e.g., Constable et al., 1987) from the AEM data is more appropriate. Accounting for borehole information to guide decisions around model parameterization in such a manner can be considered a qualitative use of the information.

If insufficient lithologic information is collected when drilling or logging a borehole, then formations can be interpreted from the geophysical log. This often-arbitrary process of segmenting the borehole log into distinct homogeneous depth intervals is commonly referred to as well-log blocking. If the goal is to smooth the log data while leaving the boundaries sharp, a median filter with an optimal sliding window may be used. More appropriate blocking of a borehole log is scale dependent. On a tenement scale, a log might be blocked into cover and basement; on the prospect scale, the same log might be subdivided into individual weathering horizons. This scale-dependent nature of borehole blocking has given rise to filtering and denoising techniques based on continuous wavelet transform (CWTs), which allow computation of more optimal blocking that honors natural change points (Cooper and Cowan, 2009). For example, Davis and Christensen (2013) derive a CWT method that was later used to determine the more natural subsurface structure of the alluvial aquifer of the Gascoyne River (Davis et al., 2015a). More quantitative approaches are based on fitting a change-point model to the borehole log (Hawkins and ten Krooden, 1978; Hawkins, 2001).

An electrical conductivity log is by itself only an indirect measurement of the true earth conductivity. Christensen and Lawrie (2018) invert electrical conductivity logs and use the resulting 30-layer models inferred from the data for comparison with AEM inversion results. Transdimensional or reversible-jump Markov chain Monte Carlo (RJMCMC) algorithms applied to the inversion of AEM data (e.g., Minsley, 2011; Brodie and Sambridge, 2012) demonstrated that models with few layers (<7) are sufficient to fit the measurements.

METHODOLOGY

Typically, the first step in a well-log analysis is identifying formation boundaries. We assert our underlying assumption that an abrupt change in a geophysical property, for example, electrical conductivity, corresponds to a boundary between two formations and that within the formations, the variation in the geophysical property can be described using a functional form. One example of the functional form is that there is a linear trend in the log as a function of depth over the formation. Another is that the formation can be adequately described by some statistical measure (mean, or median) of the geophysical measurements across its extent. If we seek to determine an objective segmentation of a well log, then we are tasked with solving an inverse problem in which the number of change points (i.e., boundaries) and the functional form representing the depth dependence of the geophysical parameter of interest within a formation layer are unknown (e.g., Hawkins and ten Krooden, 1978). In our work, we are interested in determining an optimum segmentation based on electrical conductivity as derived by the EM system, although this principle is true for any measurement.

Given an apparent-conductivity trace from a well log, the determination of its change-point model and information about its robustness can be inferred by using transdimensional MCMC algorithms (e.g., Gallagher et al., 2011). We observe that we are not determining the change-point model of the earth from the resolution of the sonde. Rather, we are determining the model from the resolving power of an EM system that is conducting a sounding in the vicinity of the borehole. In this work, we adopt the notion that the apparent-conductivity log of a geophysical borehole trace represents an accurate model of the subsurface electrical conductivity structure.
The optimization we seek is to minimize the misfit between the segmented well log and the original well log through the measurement recorded by an EM system. Because the original well log represents the true earth model, we can generate the geophysical response of the EM system for that model. We call this the true response \( f(m) \), where \( m \) is the borehole earth model, defined later. Our segmented well log produces another, simplified, model that has a separate geophysical response. Using a likelihood function that is based on the difference between the geophysical response of the true model (determined by the original log) and the simplified model (calculated from the segmented well log) allows us to block the well log at the resolution of the EM system. We therefore determine the ability of the EM system to resolve and detect the features of interest, which will assist us in our selection of an appropriate system, if any exists, and to inform us on how well our inversions will be able to solve the problem for which the survey was undertaken.

**RJMC$$\text{MC borehole blocker}$$**

In the following, we will refer to our implementation of the aforementioned approach as the RJMC$$\text{MC}$$ borehole blocker. A conductivity borehole log consists of a finite set of \( N + 1 \) conductivity measurements \( \sigma \) spaced evenly in depth down a vertical well. The depth interval is typically approximately 0.05 m for hydrogeophysical logging in Australia. A measurement of conductivity from the original log at discretized depth \( n \) is given as \( \sigma_{\text{true}} \), and we may express the entire downhole record that represents the true earth model as \( \sigma_{\text{true}} = \{\sigma_{\text{true}}(n) : 1 \leq n \leq N + 1\} \). Here, \( n \) represents the measurement at a given (discrete) depth index, with the first index \( n = 1 \) at the first depth, not including the surface. The true layered earth model is constructed from each of the conductivity-depth pairs, with the last layer whose top is at depth \( n = N + 1 \), representing a half-space with a conductivity of \( \sigma_{\text{true}}(N) \).

The segmented model with a set of \( k \) change points and \( k + 2 \) layers is given as \( m = \{\sigma_{k}, k \leq \sigma_{k+1} \} \), where the electrical conductivity for the \( k + 2 \) half-space is given by \( \sigma_{\text{true}}(n) \) at depth \( N + 1 \). The depths of the \( k \) change points that partition the borehole trace into \( k + 1 \) layers are a discrete subset of \( \{n\} \), that is, depths determined by the borehole measurement. The method of Davis and Christensen (2013) allows change points to occur at any depth determined by the CWT rather than at some discrete depth \( k \); however, we judge that the simplification proposed here is justified because the discretization for depth in the borehole log is on the order of centimeters, whereas for ground-based and AEM systems, we typically have depths on the scale of meters. The conductivity \( \sigma c \) of the simplified model is calculated using a summary statistical measure of the layers of the true model that lie between the \( k \) change points. Let us consider, for example, that the first change point occurs at \( n = 100 \). This means that the segmented model has a layer that extends from the surface of the earth to index 100, and the value of \( \sigma_{c} \) must be derived from \( \{\sigma_{1}, \ldots, \sigma_{100}\} \). We can calculate the total conductance of the layers distributed over these depths and calculate the mean conductivity, we can calculate the median conductivity of the layers in the set, or we can calculate the geometric mean conductivity. We consider each of these for comparison.

RJMC$$\text{MC}$$ algorithms depend upon the use of Bayes’ theorem, and they are often implemented in the Metropolis-Hastings-Green formulation (Green, 1995) in which a model \( m \) is proposed from a previous model \( m' \) with an acceptance criterion \( \alpha \) that we express as

\[
\alpha(m', m) = \min\left\{ 1, \frac{p(m') p(d|m') f_{\text{move}(m')} q(m|m')}{p(m) p(d|m) f_{\text{move}(m)} q(m'|m')} \right\},
\]

where \( p(\cdot), q(\cdot), \) and \( f(\cdot) \) represent the probabilities, \( m' \) denotes a proposal model, and \( d \) is a vector of \( M \) voltages measured by an EM system. The four ratios in equation 1, from left to right, are the prior model ratio, the likelihood (data misfit) ratio, the move-type ratio, and the proposal ratio. The final term is the Jacobian. A proposed model is accepted based on the comparison of \( \alpha \) to 1: If it is equal to 1, we accept the model; if it is less than 1, we accept the proposal with a probability of \( \alpha \). That is, we draw a uniform random number \( u = U(0, 1) \) and accept the change if \( u \geq \alpha \), rejecting it otherwise.

**Move-type ratio**

We propose three different move types for each step in the chain: a birth move, in which a new change-point is proposed; a perturb move, in which we select a change point and move it according to some proposal scheme; and a death move, in which a change point is removed. In each case, the conductivities of the new model are recalculated based on our chosen functional method.

If there are no change points in the model, then we can only propose a birth move because there are no change points to eliminate, nor are there any available for which to move. If we have \( N \) change points, then the set is full, and we cannot propose another or change the position of the ones existing. In the final case, we are free to perturb the position of a given change point (provided that the change point does not fall on an index already occupied), we may propose a new change point, or we may elect to eliminate one from our set. The move-type ratios are therefore

\[
\frac{f_{\text{move}(m')}}{f_{\text{move}(m)}} = \begin{cases} 
1, & 0, 0, \text{ if } k = 0 \\
1/3, & 1/3, 1/3, \text{ if } 1 \leq k \\
0, 0, 1, & \text{if } k = N,
\end{cases}
\]

for the moves birth, perturb, and death, respectively.

**Prior ratio**

We are free to express our knowledge in the formulation of our prior beliefs about the segmented model. This is done by expressing the probability of a given model \( p(m) \) as a hierarchical model such that

\[
p(m) = p(k) p(c|k) p(\sigma|c, k),
\]

where \( p(k) \) is the probability of the number of change points occurring to segment the model, \( p(c|k) \) is the probability of a particular distribution of change points that segment the borehole trace, and \( p(\sigma|c, k) \) is the probability of the conductivity of the layers partitioned by the change points \( c \). We note that the conductivities depend on the distribution of change points and are independent so that \( p(\sigma|c, k) = \prod_{i=1}^{k} p(\sigma_{i}|c, k) \).

Using an uninformative prior, we suppose that any number of change points \( k \) is just as likely as any other \( k' \) a priori. The probability ratio of the prior for \( k \) and \( k' \) is therefore 1. Given several segmentation points \( k \), we make the further assumption that a particular distribution of the change points \( c \) is also as likely as any other. The number of distributions of the change points is equal to the binomial
coefficient for \( k \) points over \( N \) locations \( \binom{N}{k} = N!/(k!(N-k)!).

The probability of a particular distribution is therefore

\[
p(c|k) = \frac{k!(N-k)!}{N!}.
\]  

(4)

It is known that EM systems tend to be most sensitive to the near-surface conductivity structure, but here we keep our prior as uninformative as possible. Finally, although the conductivity values \( \sigma \) are determined through a deterministic function, we assume that they can be distributed uniformly over the range of conductivities recorded by the borehole tool. Thus,

\[
p(\sigma|c, k) = \frac{1}{(\Delta \sigma)^N},
\]  

(5)

where \( \Delta \sigma = \sigma_{\text{max}} - \sigma_{\text{min}} \) and \( \sigma_{\text{max}} \) and \( \sigma_{\text{min}} \) are the maximum and minimum conductivity values measured in the original log, respectively.

When we consider the prior ratio for each of the moves described above, our assertions about indifference to the model become important. For the case in which we perturb a change point, the probability ratio becomes 1, meaning that every model is equally likely.

When a birth move is proposed, and \( k \) increases to \( k + 1 \), we see no change in the prior for the increase in the number of change points. We do, however, allow for a more complicated distribution across the set \( N \), and the prior probability of the \( k + 1 \) change-point distribution becomes

\[
p(c|k+1) = \binom{N+1}{k+1}
\]  

(6)

Furthermore, we have increased the number of conductivities that segment our model, and we pay for this increase in possible parameters with an extra conductivity prior probability term. The prior ratio for a birth move then becomes

\[
\frac{p(m')}{p(m)} \bigg| \text{birth} = \frac{(k+1)!(N-k-1)!}{N!\Delta \sigma} \frac{N!}{k!(N-k)!} = \frac{k+1}{(N-k)\Delta \sigma}.
\]  

(7)

**Likelihood ratio**

We use the usual \( L_2 \)-norm misfit function between the data \( d \) measured by the EM system for the original conductivity log and \( f(m) \), the forward prediction of the data based on the proposed model \( m \), so that the likelihood ratio becomes

\[
\frac{p(d|m')}{p(d|m)} = \frac{\exp((d - f(m))/C_d^{-1}(d - f(m)))}{\exp((d - f(m'))/C_d^{-1}(d - f(m')))}. \tag{8}
\]

Here, \( C_d^{-1} \) is the data covariance matrix that models how noise is distributed in the measurement. In this work, we use a simple multiplicative noise applied to each member \( d_k \) of the measured data so that \( C_d^{-1} \) is a diagonal matrix, and we provide an example of segmentation results that incorporate different values of \( C_d^{-1} \) later. Note that equation 8 has the misfit of our proposed model in the denominator, so that a proposed model \( m' \) that has a lower misfit than the previous model than \( m \) results in a likelihood ratio that is greater than 1 and is therefore more likely to be accepted. When we consider only the prior and the likelihood ratio together, we see that we would be more likely to accept better fitting and more complicated models. We note here that the choice of forward modeler used in the sampler is subject to the choice of the experimenter; we have chosen to use the AMIRA P223 project code LEROI (Raiche et al., 2007).

**Proposal ratio**

The proposal ratio is dependent on the move type, and it expresses the ratio of the joint densities \( q \) and \( q' \) needed for a given proposal. Remembering that the total number of variables and generative random numbers must remain constant across all transitions (Green, 1995), we consider each of the move types.

The simplest move is one that changes the depth of a change point, say \( k_j \), from one location to another. The proposal ratio evaluates to 1 provided that the proposals are symmetric for the forward and return moves. There is a variety of proposal methods that may be used, and care must be taken that proposals outside the boundary of the set are not allowed. We use a Baxtian proposal function, with reflection at the boundaries (Yang and Rodriguez, 2013). Once the position of \( k_j \) is perturbed, the layer conductivity values of the segmented model are recalculated. Note that if the position of \( k_j \) is altered so that it is less than \( k_{j-1} \) (or greater than \( k_{j+1} \)), then the set is reordered and there is no change in the total number of layers.

A birth move proposes a new layer boundary in the segmented model so that the number of change points increases from \( k \) to \( k + 1 \). Because there are \( N - k \) indexes available for the proposal and we have no preference for the distribution of the \( k + 1 \) change points, we choose a single index with probability \( 1/(N-k) \). For the reverse move, there are \( k + 1 \) available change points. If we were to select the one that we just generated, we would do so with a probability of \( 1/(k+1) \). The conductivity of the new layer generated is formed from our statistical function calculated from the original borehole log, so we do not need to propose a new value for it. We note that the prior ratio (equation 7) already includes the probability assigned to the new conductivity value. The proposal ratio for a birth move is thus

\[
\frac{q(m|m')}{q(m'|m)} \bigg| \text{birth} = \frac{N-k}{k+1}.
\]  

(9)

The death move is reciprocal to the birth move above. We select a change point out of the set of \( \{k+1\} \) indexes and remove it with a probability of \( 1/(k+1) \). From the reduced model, we would have \( N - k \) locations to choose from to create the birth move, and we do this with probability \( 1/(N-k) \). The proposal ratio for the death move is thus

\[
\frac{q(m|m')}{q(m'|m)} \bigg| \text{death} = \frac{k+1}{N-k}.
\]  

(10)

**Jacobian**

The Jacobian determinant is necessary to transform the randomly generated numbers required when increasing the number of change points in the segmented model. The mapping from \( k \) to \( k + 1 \) change points (increasing the total number of layers by one) requires only one random number to be drawn, and that number is drawn uniformly from \( U(1,N-k) \). The number drawn is mapped directly to the set
indexes that are available. Once an index is chosen, we update the conductivity of the newly segmented model using a particular statistical function. Because the number of indexes is finite, and the statistical functions that we consider are the mean, median, and geometric mean of the borehole conductivity values that are contained within the true model, there is always a one-to-one correspondence between a particular set of change-point indexes and the corresponding conductivity values allowed. Therefore, correspondence between a change-point set and resulting segmented conductivity values is bijective and, following Denison et al. (2002), the Jacobian evaluates to 1.

We now see that the Metropolis-Hastings-Green acceptance criterion becomes

$$a(m', m) = \min \left\{ 1, \frac{\exp((d-f(m'))/C_d^{-1}(d-f(m))) j_{\text{move}}(m')}{\exp((d-f(m))/C_d^{-1}(d-f(m))) j_{\text{move}}(m)} p_{b-d} \right\},$$

where the move-type ratio values are given in equation 2, and $p_{b-d}$ is the product of the prior and proposal ratios:

$$p_{b-d} = \begin{cases} (\Delta \sigma)^{-1}, & \text{if birth} \\ \Delta \sigma, & \text{if death} \\ 1, & \text{otherwise.} \end{cases}$$

Notice the inherent parsimony of the segmentation strategy. Models that lower the misfit will be more likely to be accepted based on equation 8, so we favor better fitting models. Increasing the model complexity by adding change points will tend to fit the measured data better. Because our prior ratio for model layers is uninformative and we do not prefer one model over another, we could eventually end up with a maximally partitioned model that replicates the structure of the original borehole trace if it were not for the punishment introduced by the prior probability of the conductivity and the data covariance matrix $C_d^{-1}$. These mitigate the tension between the decreased misfit due to the increased model complexity and the desire to have models that fit the data without being overly complex. Eventually, a balance will be struck between the data misfit and the model complexity, and this will be governed by the physics of the diffusive nature of electromagnetics and the measurement capability of the EM system in question.

**EXAMPLES**

In this section, we provide examples of our method to clarify points that provide the foundation for an objective comparison between EM inversion results and borehole conductivity logs. First, we demonstrate the performance and typical statistics of an RJMCMC borehole blocker run. This shows how the Markov chain is summarized into an ensemble of conductivity-depth pairs. We then show how the RJMCMC logs compare to conductivity logs that are blocked using wavelet transforms and that the RJMCMC borehole blocker models honor the scale and resolution ability of the EM system used. We demonstrate the effect of system noise in the EM system and show that increased noise degrades the performance of the blocking algorithm, and we also compare the ability of different EM systems to block a particular conductivity log. We then show how the RJMCMC borehole blocker algorithm compares to other inversion methods such as smooth-model inversion and RJMCMC realizations. Finally, we use the RJMCMC borehole blocker method on a borehole that is adjacent (<20 m) to an AEM survey line in Western Australia.

For every RJMCMC borehole blocker example that we demonstrate in this paper, we use parallel tempering and delayed rejection sampling (Mira, 2001; Sambridge, 2014). Both techniques offer better mixing and reduce burn-in and autocorrelation. In addition to calculating the likelihood function of the measured and forward-modeled data, and hence the data misfit, we follow Christensen and Lawrie (2013) and calculate the model misfit. The model misfit is calculated in the same fashion as the data misfit (i.e., the numerator of equation 8), which offers a measure of how well the segmented conductivity-depth model compares to the measured borehole trace. Every sampling run begins with a randomly chosen three-layer earth. That is, there is one change-point selected in the borehole trace and a half-space at the bottom of the borehole. All data used in these examples are available from the CSIRO Data Access Portal (CSIRO, 2019).

**Sampler performance**

First, we show the performance results of a typical RJMCMC borehole blocking application. Figure 1 shows a summary of results of a run from a single chain of the RJMCMC borehole blocker for the geophysical log from a 120 m induction conductivity borehole log recorded from the Gascoyne River AEM Project outside Carnarvon, Western Australia (Davis et al., 2016). The project, a collaboration between the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the Department of Food and Agriculture, Western Australia, and funded by Royalties for Regions, was for groundwater exploration and aquifer detection. The geophysical log is from a borehole that was drilled after the project was completed, but was used to determine the thickness of the Old and New Alluvium aquifers overlying the Cardabia Calcareinite aquitard (Global Groundwater, 2016). The EM system considered in this example is SkyTEM®304, a central-loop-style helicopter-borne AEM system (HEM).

Figure 1a shows the negative logarithm of the likelihood function for the accepted models in the run. In this case, as in all the examples mentioned here, we ran the sampler for $2 \times 10^5$ proposals. Results were down-sampled by a factor of 10, and a burn-in period of approximately 5000 samples was stripped off the beginning. We found, through trials with different starting models, that 5000 was a sufficient burn-in due to the simplicity of the change-point problem, parallel tempering, and delayed rejection enhancements. Figure 1b and 1d shows the time series and histogram of the number of layers, and Figure 1c shows the average acceptance rate. The original borehole trace is shown in Figure 1e along with a representation of the results of the RJMCMC borehole blocker. Here, we show the 10%, 50%, and 90% ranges for the run. Underlying Figure 1e is a pixelated image of every model in the ensemble. We can draw every model on the subfigure and accumulate the results like an exposure from a photograph. For every depth interval, we have a distribution of conductivity that can be analyzed to determine percentiles and the mode. The mode conductivity for a depth is the most probable. Figure 1f shows a histogram of the change points from the estimation. Clearly, the AEM system is sensitive to the conductive layer in the upper 10 m, but it is less sensitive to the resistive layers that are deeper. It recovers the conductor at 100 m in depth and also the change in conductivity at 30 m, but it is relatively insensitive to the 20 m thick resistive layer at a 60 m depth.
Figure 1. Performance of the RJMCMC borehole blocker algorithm for the example borehole (see Figure 2). (a) Negative log of the likelihood. (b) Number of layers. (c) Averaged acceptance rate of the proposals. (d) Histogram of a number of layers from (b). (e) Original borehole trace (solid black) with 10% (fuchsia, right triangles) and 90% (blue, left triangles) bounds, and the 50% mode (red line with ×). (f) Distribution of change points indicating the segmentation of the models. The median function is used to calculate segmented conductivity values.

Figure 2. Examples of well-log blocking at the scale of a central-loop AEM system such as SkyTEM. (a) Misfit between the AEM responses of the original and blocked well log using different blocking strategies. Included is the data misfit. (b) Borehole logs segmented using the CWT with a cutoff of 5 m minimum layers (red solid), and an eight-layer cutoff (green dashed). (c) Borehole log blocking obtained using the RJMCMC borehole blocker which accounts for the ability of the EM system to resolve the earth model. Results are shown for the median-value (blue), mean-value (magenta dashed), and log-mean blockers (orange dashed).
Application to log data

The same geophysical log is used to demonstrate how the RJMCMC borehole blocker compares to the CWT method as described by Davis and Christensen (2013). Figure 2b and 2c shows the original induction conductivity log (the thick black lines). Figure 2b shows two different CWT realizations: one with a minimum cutoff of 5 m (the solid red line) and one with a cutoff using the eight most important boundaries (the dashed green line). The CWT log block models accurately reproduce the original borehole trace, with model misfits of $\phi_{m}^{\text{CWT}} = 2105$ and $\phi_{m}^{\text{eight-layer CWT}} = 9563$ for the 5 m and eight-layer case, respectively. The RJMCMC borehole blocker results have poorer model recovery, Figure 2c, where we show the most probable models resulting from the median (blue), mean (dashed magenta), and log-mean (dashed orange) estimators for layer conductivity. When we calculate the data misfit for these examples, Figure 2a, we see that the RJMCMC models have lower data misfits than the CWT models. As we would expect, the RJMCMC borehole blocker models more accurately reflect the sensitivity of the AEM systems to the conductivity structure of the earth than the CWT transforms do. That is, the models in Figure 2c show how the AEM system filters the earth’s geoelectric structure measured by the borehole log, reflecting the physics of the EM measurement: An AEM system is sensitive to conductivity changes in the near-surface, and less so with increasing depth. By contrast, the CWT blocker does not reflect the physics of the AEM system measurement and therefore cannot be recommended as a tool for comparing borehole logs to EM models.

Influence of noise

We consider noise in the EM system and how it affects the borehole blocking results. Recall from equation 8 that the misfit between the measured and modeled data is modified in the $L_2$ norm by the data covariance matrix $C_d^{-1}$. In the examples used in this paper, we used a multiplicative error. Each EM receiver channel is ascribed a 3% error, and noise is considered to be independent across channels. This means that $C_d^{-1}$ is diagonal and has entries of $1/((3\% \times d_n)^2)$, where $d_n$ is the measured data of the $n$th channel of the EM system.

Figure 3 shows an artificial borehole conductivity trace composed of a collection of sine waves. The RJMCMC borehole blocker (median function) results are shown for a few different values of multiplicative noise (1%, 3%, 6%, and 10%). The blue-shaded region indicates the 10%–90% range for the 3% noise level. The model misfits increase as the noise is increased, reflecting that the AEM system is less capable of discriminating changes in the electrical structure of the earth. By contrast, the 1% noise case (dashed magenta) recovers the near-surface, but also gives better resolution of the borehole log at depth, as expected. Model misfits are indicated in the figure. The blue shaded region indicates the 10%–90% range for the 3% noise level. The median function is used.

Figure 4. Demonstration of the method with different EM systems to show depths of detection. We use an artificial borehole log for (a) a ground-based system such as NanoTEM, (b) a helicopter-borne central-loop system such as SkyTEM, and (c) a fixed-wing system such as TEMPEST. In each subfigure, we show the original trace (black) and the RJMCMC borehole blocker results for the median (blue) function. The shaded region marks the 10%–90% boundaries.
earth with increasing noise. We see not only reduced depth of detection, but fewer layers in the resulting models. Thus, we see that noise is an important factor when comparing borehole logs to different EM systems.

**Different EM systems**

When planning an EM survey, the most important factor to consider is whether the EM platform chosen is able to resolve the primary subsurface features of interest. The technique we propose is an excellent method to do this. We examine the ability of three different systems to resolve the deep borehole model displayed in Figure 3.

Figure 4 shows the RJMCMC borehole block method applied using a ground-based system such as NanoTEM (Figure 4a), a helicopter-borne AEM system such as SkyTEM304 (Figure 4b), and a fixed-wing AEM system such as TEMPEST (Figure 4c). In each panel, we look at the most probable models using the median function for the segmented conductivity values. We see that the NanoTEM system, which is operating with a 40 × 40 m loop at a base frequency of 32 Hz, can just detect conductivity changes to approximately 500 m depth, but it is extremely sensitive to the near-surface. The SkyTEM304 system, which operates with two moments in a transmitter with an area of approximately 347 m² and base frequencies of 175 and 25 Hz at 30 m altitude, can resolve features to a depth of approximately 800 m. The TEMPEST system is also capable of resolving to this depth. It is modeled with a 154 m² dipole transmitter separated 120 m horizontally and 35 m above a dipole receiver. The transmitter has an altitude of 120 m and operates at 25 Hz. If the feature of interest is the conductive layer at a 300 m depth, then we can surmise that all three EM systems would be able to detect it. If we are interested in the change in conductivity at 800 m depth, then we can discard the ground-based system and opt for either of the AEM platforms. We note here that the depth of detection for each of these systems is greatly enhanced due to the fact that we are heavily constraining our model. Recall from the discussion on move type that we choose change points at random, but the conductivity of the layers contained within the change-point boundaries are calculated from the original borehole log. This means that the models generated from the RJMCMC borehole block method constitute best-case scenarios. If a layer is not detected with this method, then it is extremely unlikely that any other type of inversion will do so either. We elaborate on this next.

**Comparison to other inversion methods**

The use of the borehole log conductivity values as input to the conductivity of the segmented layer that the EM system sees heavily constrains the model and offers a great deal of information that is otherwise not available. This is important when comparing our method to other inversion techniques. In this section, we compare the RJMCMC borehole blocker to a 20-layer smooth inversion that has a Tikhonov type 1 regularization in the conductivity model and logarithmically increasing layer thickness. We also show the mean posterior model from an RJMCMC estimation. The results are shown in Figure 5. Figure 5a shows the original borehole trace (black) as well as the most probable model from the RJMCMC borehole blocker (blue), and an RJMCMC algorithm (dark-blue dashed). Also shown is a 20-layer smooth model. The light-blue-shaded region is the 10%–90% region for the RJMCMC method, and the dark-blue-shaded region is for the RJMCMC borehole blocker method.

Figure 5b shows the percentage difference between the true earth response and the modeled responses for each of the methods used. From the figure, we see that the 20-layer smooth model has the lowest data and model misfit of all methods. However, we must remember that this result makes sense when we consider that the inversion model is constrained to be smoothed within a certain level of tolerance that encourages the inversion to converge. It is arguable that the similarity between the borehole log and the inversion model is a result of serendipity.

The results from the RJMCMC inversion are also shown in Figure 5. The mean posterior model looks similar to the borehole blocker model, but it has a slightly higher model misfit. The data
misfit from the mean posterior model is quite large compared to the other two methods. This is an interesting result because the accepted models from the RJMCMC chain all had data misfits of approximately 20. A collection of many models, each of which individually fits very well, may not necessarily combine to a mean model that fits very well. Unsurprisingly, the model misfit given by the RJMCMC borehole blocker model is similar to the smooth inversion. This is due to the fact that the models allowed during the estimation process must be produced by a deterministic function acting on the original log. The similarity between the results is encouraging and demonstrates that RJMCMC borehole blocking compares well to other methods of inversion.

APPLICATION TO FIELD DATA

As a final example of the use of the borehole blocking algorithm, we present inversion results from AEM survey data with a borehole logged nearby (<20 m). The borehole geophysical log is the one that we have been using throughout this paper. Borehole B 9-16 was drilled after the Gascoyne River AEM project was completed. It is located on the north side of the Gascoyne River outside Carnarvon, Western Australia. The borehole is drilled to a depth of 137 m and logged to 125 m. The drilling log records the Gascoyne New and Old Alluvium exists from surface to a depth of 55 m. This depth range constitutes the main Quaternary aquifer targets of the Gascoyne project that overlay a layer of undifferentiated Cenozoic materials such as calcarenite, calcilutite, clays, and mudstones. There are some minor aquifers consisting of sand and gravel, but they are not very thick and are not laterally extensive. At a 100 m depth, the drilling logs indicate a transition to mudstones and clays that are consistent with the Toolonga Calcilutite. The AEM survey was flown with the SkyTEM304 system, and further details of the system specifications may be found in reports and data sets (Davis et al., 2015a, 2015b, 2015c, 2016). Figure 6a–6c shows layered earth inversions from the GALEISBSTDM, AirBeo, and Aarhus Workbench, respectively (Raiche et al., 2007; Brodie, 2010; Auken et al., 2015). In each case, the RJMCMC borehole blocker segmented model is shown at a distance of 500 m. All sections are color-coded to the color bar at the bottom of the figure. A 50 m diameter circle is shown in Figure 6c to indicate vertical exaggeration.

Each inversion program shows similar results: a resistive cover layer that increases in conductivity to approximately 0 m Australian Height Datum (mAHD), followed by a resistive layer, the bottom of which marks the end of the Old Alluvium. The bottom of the Cainozoic is at approximately −50 mAHD, at the transition to the more-resistive layers of calcarenite and calcilutite. The Toolonga Calcilutite begins at approximately −80 mAHD. The water table is at approximately 12 mAHD. The RJMCMC borehole blocker segmented model shows these main features. GALEI inversions clearly show the top of the relatively conductive Toolonga Calcilutite, but erroneously show a resistive layer near −150 mAHD. Furthermore, the differentiation between the Old Alluvium and the Cainozoic is not very clear.

Figure 6. The RJMCMC borehole blocker results (shown inside the rectangle at 500 m) from the earlier examples overlaid on the inversion results from a 1000 m stretch of SkyTEM survey data near Carnarvon, Western Australia. The colors indicate the conductivity in this figure. A 50 m circle is shown in (c). Panels (d-f) show the borehole blocker model (blue), but also the inverted model from the nearest AEM survey fiducial for examples presented in the left column.
The AirBeo inversions show all four of the zones correctly (unsaturated and saturated New and Old Alluvium, Undifferentiated Cainozoic, and Toolonga Calcilutite Formations), but they are more jagged laterally. The Aarhus Workbench inversion, which uses a spatially constrained inversion algorithm, shows all regions but is perhaps smoother than is necessary. They all broadly reflect the main features that we know are there and agree generally with the borehole blocker model. Our aim in this section is not to compare different inversion techniques; rather to demonstrate that all of the inversion techniques reflect the borehole blocker segmented model with similar levels of agreement. Recalling that the RJMCMC borehole blocker recovers the best model that the AEM system is capable of resolving (because conductivity values of the RJMCMC borehole blocker models are constrained to the borehole logs), we argue that each inversion yields models consistent with a robust method that converts measured borehole induction conductivity logs into the best segmented models that an AEM system can resolve.

CONCLUSION

It is well known that borehole logs contain electrical conductivity data on a much finer vertical scale than is achievable with other EM systems that must make inductive soundings at or above the earth’s surface. Because a disparity of scale exists between surface (or airborne) EM systems and borehole records, there should be some way to reconcile the measurements. In this paper, we have explored the notion of simplifying and segmenting geophysical induction conductivity logs according to the resolving ability and sensitivity of different EM systems.

We have shown that the RJMCMC borehole blocking method is a valid approach to comparing borehole records to inversions of ground-based and airborne EM systems. Models generated using our approach are sensible representations of the subsurface that are only as complicated as the sensitivity of the EM system allows. We simplify the borehole induction conductivity model through the filter of the EM system itself. Our method is based on the machinery of RJMCMC simulations, which methods have been developed extensively in the literature and are now in relatively common practice in geophysics. The method is intrinsically parsimonious toward complex models and hence follows an Occam principle of seeking the simplest models that satisfy the data given. Our field data example on an induction log in an alluvial aquifer demonstrates this nicely and shows that existing inversion techniques for layered earth models are uniformly capable of replicating the blocked borehole log. Differences between the inversion packages are mainly due to vertical and horizontal regularization. The over- and undershoot in the inversion models is typical of $L_2$-type norms coupled with a smoothing constraint.

Models from the RJMCMC borehole blocking technique offer insights into models obtained from commonly available inversion packages of AEM data. Although inversions with vertical regularization restrictions may produce models that are better in terms of data misfit, those same models contain features that are a consequence of the choice of the regularization operator used and are not necessarily a representation of physical reality. Whereas model features obtained using regularized methods might be confirmed by comparison with an unblocked conductivity log, they might not be supported by the data or be detectable by an EM system. We avoid this problem by examining boreholes at the scale and resolution sensed by a chosen EM system.

ACKNOWLEDGMENTS

The authors thank the Commonwealth Scientific and Industrial Research Organisation (CSIRO) for support in publishing this paper. Financial assistance was provided by the Deep Earth Imaging Future Science Platform within CSIRO. We would also like to acknowledge the Department of Primary Industries, Western Australia (formerly the Department of Agriculture and Food), and Global Groundwater for access to borehole record B9-16.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be accessed via the following URL: Note: A digital object identifier (DOI) linking to the data in a general or discipline-specific data repository is strongly preferred.

REFERENCES

Aukén, E., A. V. Christiansen, C. Kirkegaard, G. Fiandaca, C. Schamar, A. A. Behrouzzaman, A. Bingel, E. Nielsen, P. Effersen, N. B. Christensen, K. Sørensen, N. Foged, and G. Vignoli, 2015, An overview of a highly versatile forward and stable inverse algorithm for airborne, ground-based and borehole electromagnetic and electric data: Exploration Geophysics, 46, 233–235, doi: 10.1071/EG1097.

Brodie, R. C., 2010, Holistic inversion of airborne electromagnetic data: Ph.D. thesis, Australian National University.

Brodie, R. C., and M. Richardson, 2013, AEM system target resolvability analysis using a Monte Carlo inversion algorithm: ASEG, Extended Abstracts, 1–4, doi: 10.1071/ASEG2013ab227.

Brodie, R. C., and M. Sambridge, 2012, Transdimensional Monte Carlo inversion of AEM data: ASEG, Extended Abstracts, 1–4, doi: 10.1071/ ASEG2012ab095.

Christensen, N. B., and K. C. Lawrie, 2013, Resolution analyses for selecting an appropriate airborne electromagnetic (AEM) system: Exploration Geophysics, 43, 213–227, doi: 10.1071/EG12005.

Christensen, N. B., and K. C. Lawrie, 2018, A novel approach to comparing AEM inversion results with borehole conductivity logs: Exploration Geophysics, 49, 309–322, doi: 10.1071/EG17029.

Christensen, N. B., and K. C. Lawrie, 2013, Resolution analyses for selecting an appropriate airborne electromagnetic (AEM) system: Exploration Geophysics, 43, 213–227, doi: 10.1071/EG12005.

Cooper, G. R. J., and D. R. Cowan, 2009, Blocking geophysical borehole log data using the continuous wavelet transform: Exploration Geophysics, 40, 233–236, doi: 10.1071/EG08127.

CIRO, 2019, Electromagnetic data for GEO-2019-0095, v1: Data collection: CSIRO, https://doi.org/10.25919/5da45bc9c0edc.

Dadi, S., R. Gibson, and K. Wang, 2016, Velocity log upscaling based on reversible jump Markov chain Monte Carlo Modeling: Geophysics, 81, no. 5, R293–R305, doi: 10.1190/1.495121.

Davis, A., T. Munday, and R. George, 2015a, Gascoyne River airborne electromagnetic (AEM) survey I: System selection report: Water for a Healthy Country, Technical report EP 1311553, CSIRO.

Davis, A., T. Munday, and R. George, 2015b, Gascoyne River airborne electromagnetic (AEM) survey II: Inversion report: Water for a Healthy Country, Technical report EP1311758, CSIRO.

Davis, A., T. Munday, and R. George, 2015c, Gascoyne River airborne electromagnetic (AEM) survey III: Drilling target selection report: Water for a Healthy Country, Technical report EP1312353, CSIRO.

Davis, A., T. Munday, and R. George, 2016, Gascoyne River airborne electromagnetic (AEM) survey IV: Final report: Technical report EP1519929, CSIRO.

Davis, A. C., and N. B. Christensen, 2013, Derivative analysis for layer selection of geophysical borehole logs: Computers and Geosciences, 60, 34–40, doi: 10.1016/j.cageo.2013.06.015.

Dennison, D. G. T., C. C. Homes, B. K. Mallick, and A. F. Smith, 2002, Bayesian methods for nonlinear classification and regression: Wiley, Wiley series in probability and statistics (OCLC: ccm48754149).

Doll, H., 1949, Introduction to induction logging and application to logging of wells drilled with oil base mud: Journal of Petroleum Technology, 1, 148–162, doi: 10.2118/949148-1.

Gallagher, K., T. Bodin, M. Sambridge, D. Weiss, M. Kylander, and D. Large, 2011, Inference of abrupt changes in noisy geochronometric records using transdimensional change point models: Earth and Planetary Science Letters, 311, 182–194, doi: 10.1016/j.epsl.2011.09.015.
Global Groundwater, 2016, Northern Gascoyne river borefield phase III bore completion report; for Department of Food and Agriculture WA (DAFWA): Technical report.

Green, P. J., 1995, Reversible jump Markov chain Monte Carlo computation and Bayesian model determination: Biometrika, 82, 711–732, doi: 10.1093/biomet/82.4.711.

Hauser, J., J. Gunnung, and D. Annetts, 2015, Probabilistic inversion of airborne electromagnetic data under spatial constraints: Geophysics, 80, no. 2, E135–E146, doi: 10.1190/geo2014-0389.1.

Hawkins, D. M., 2001, Fitting multiple change-point models to data: Computational Statistics & Data Analysis, 37, 323–341, doi: 10.1016/S0167-9473(00)00068-2.

Hawkins, D. M., and J. A. ten Krooden, 1978, A review of several methods of segmentation, in D. Gill and D. F. Merriam, eds., Geomathematical and petrophysical studies in sedimentology: Pergamon Press, 117–126.

Lawrie, K. C., R. S. Brodie, K. P. Tan, D. L. Gibson, J. W. Magee, J. D. A. Clarke, L. Halas, L. Gow, P. Somerville, H. Apps, N. B. Christensen, R. C. Brodie, J. Abraham, M. Smith, D. Page, P. J. Dillon, J. Vanderzalm, K. Miotlinski, S. Hostetler, A. Davis, A. Y. Ley-Cooper, G. L. Schoning, K. Barry, and K. Levett, 2012, Securing Broken Hill’s water supply: Assessment of conjunctive water supply options involving managed aquifer recharge and/or groundwater extraction at Menindee Lakes: Data acquisition, processing, analysis and interpretation methods: Technical report, Geoscience Australia (OCLC: 908193406).

Minsley, B. J., 2011, A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data: Geophysical Journal International, 187, 252–272, doi: 10.1111/j.1365-246X.2011.05165.x.

Mira, A., 2001, On Metropolis-Hastings algorithms with delayed rejection: Metron — International Journal of Statistics, 59, 231–241.

Moran, J. H., and K. S. Kunz, 1962, Basic theory of induction logging and application to study of two-coil sondes: Geophysics, 27, 829–858, doi: 10.1190/1.1439108.

Raiche, A., F. Sugeng, and G. Wilson, 2007, Practical 3D EM inversion? The P223F software suite: ASEG, Extended Abstracts, 1–5, doi: 10.1071/ASEG2007ab114.

Reading, A. M., and K. Gallagher, 2013, Transdimensional change-point modeling as a tool to investigate uncertainty in applied geophysical inference: An example using borehole geophysical logs: Geophysics, 78, no. 3, WB89–WB99, doi: 10.1190/geo2012-0384.1.

Sambridge, M., 2014, A parallel tempering algorithm for probabilistic sampling and multimodal optimization: Geophysical Journal International, 196, 357–374, doi: 10.1093/gji/ggt342.

Shen, L. C., 1989, Effects of skin-effect correction and three-point deconvolution on induction logs: The Log Analyst, 30, 217–224.

Vignoli, G., G. Fiandaca, A. V. Christiansen, C. Kirkegaard, and E. Auken, 2015, Sharp spatially constrained inversion with applications to transient electromagnetic data: Geophysical Prospecting, 63, 243–255, doi: 10.1111/1365-2478.12185.

Yang, Z., and C. E. Rodriguez, 2013, Searching for efficient Markov chain Monte Carlo proposal kernels: Proceedings of the National Academy of Sciences of the United States of America, 110, 19307–19312, doi: 10.1073/pnas.1311790110.