

The use of asymmetric time constraints in 4-D ERT inversion

Authors

M.H.Loke¹, P. B. Wilkinson², J.E. Chambers², S. Uhlemann^{2,3}, T. Dijkstra⁴ and T. Dahlin⁵.

¹ Geotomo Software Sdn Bhd, 115 Cangkat Minden Jalan 6, 11700 Gelugor, Penang, Malaysia. email : drmhloke@yahoo.com

² British Geological Survey, Keyworth, Nottingham, NG12 5GG, U.K.
email : pbw@bgs.ac.uk, jecha@bgs.ac.uk

³ Lawrence Berkeley National Laboratory, 1 Cyclotron Road 85B-0104K, Berkeley, CA 94720, USA.
email : suhlemann@lbl.gov

⁴ School of Architecture, Building and Civil Engineering, Loughborough University, Loughborough LE11 3TU, United Kingdom
email : t.a.dijkstra@lboro.ac.uk

⁵ Lund University, Engineering Geology, Lund University, John Ericssons väg 1, S-223 63 Lund, Sweden
email : torleif.dahlin@tg.lth.se

Corresponding author : M.H.Loke, drmhloke@yahoo.com

Publication

Loke, M.H., Wilkinson, P.B., Chambers, J.E., Uhlemann, S., Dijkstra, T. and Dahlin, T., 2022. The use of asymmetric time constraints in 4-D ERT inversion. *Journal of Applied Geophysics*, **197**, 104536.

Abstract

Time-lapse resistivity surveys are commonly used to monitor temporal changes in the subsurface. In certain cases, it is known from other information that the resistivity will only decrease or increase with time. The 4-D resistivity smoothness-constrained inversion method reduces artifacts due to noise by including a temporal roughness filter constraint that ensures the temporal changes vary in a smooth manner. A least-squares optimization method is used to find a solution by attempting to locate the minimum of an objective function that consists of the data misfit and model (spatial and temporal) roughness. In some cases, the 4-D time-lapse inverse models show an increase in the resistivity with time in parts of the subsurface where it is only expected to decrease (or vice versa). We compare two methods, the barrier function and transformation methods, that attempt to minimize or eliminate these artifacts. We incorporate the barrier function constraint into the 4-D inverse method by using a modified difference matrix as a temporal roughness filter. The barrier function constraint includes an additional term that increases the objective function value greatly if the model values cross the allowed thresholds. This greatly minimizes the artifacts but does not completely eliminate them. It has the advantage that there are minimal changes in the objective function in regions of model space that are not close to the imposed thresholds. The method of transformations changes the model parameter such that the additional positivity or negativity constraints are implicitly included in the transformed model parameter. It has the advantage that it can completely eliminate the artifacts. However, it modifies the entire objective function which could be a disadvantage in some cases. We also explore a combination of the two methods, using the barrier function method to generate an initial model that minimizes the artifacts followed by the transformation method. This hybrid technique completely removes the residual artifacts left by the barrier method, and produces an inverse model which is closer to the true model for a synthetic data set. We also describe a post-inversion modification of the L-curve method to determine the optimum model that takes into account the non-linear nature of the inverse problem and the forward modelling method. The technique gave an estimate of the noise level for a field data set and produced a model which is consistent with independent hydrological measurements at the test site.

Keywords : 4-D, resistivity, inversion, asymmetric, time, constraints

1. Introduction

Over the past few decades there has been a rapid expansion in the use of 3-D ERT (Electrical Resistivity Tomography) surveys that are now widely used in engineering, environmental, hydrological, archaeological and mineral exploration surveys (Auken et al., 2006; Loke et al., 2013). In recent years, there has been more interest in time-lapse surveys to measure temporal changes in the subsurface that gives information about hydrogeological variations with time (Cassiani et al., 2006; Singha et al., 2014; Uhlemann et al., 2017; Hojat et al., 2019). These surveys normally use a grid of electrodes where the measurements using the same array configurations are repeated at regular intervals. The 4-D smoothness-constrained least-squares optimization method (Kim et al., 2009; Loke et al., 2014a) has proved to be a robust method for the inversion of time-lapse data sets. The measurements made at different times are inverted simultaneously as a single data set. A temporal smoothness constraint is included to ensure that the resistivity changes in a smooth manner with time which reduces artifacts due to random noise. In some situations, such as a solution injection project (Putman, 2014; Rucker et al., 2014), it is expected that the subsurface resistivity will only decrease (or increase) with time. However, the standard 4-D inversion method does not directly constrain the direction of the changes with time. It might produce artifacts where parts of the model show an increase with time where it is expected to only decrease (or vice versa). In this paper, we examine two different methods of including constraints to explicitly constrain the direction of the changes with time. The first method uses a barrier function (Putman, 2014; Johnson et al., 2015), while the second method uses transformations (Daniels, 1978). We also examine an alternative technique that combines the two methods to take advantage of the strengths of both methods. The following sections describe the inversion techniques followed by a synthetic model test data set and a field survey example.

2. Theory

2.1. General nonlinear optimization

In geophysical interpretation, we seek to find a model for the subsurface physical property (such as the resistivity) so that the calculated model response matches the measured values (such as the apparent resistivity). The solution is invariably non-unique in practical field surveys. Thus, additional regularization terms are added to the objective function based on independent knowledge or reasonable assumptions about the variations of the model

parameters. A nonlinear optimization method is then used to find the minimum of an objective function $\Phi(\mathbf{m})$ such as given below.

$$\Phi(\mathbf{m}) = \Phi_d(\mathbf{m}) + \lambda\Phi_s(\mathbf{m}) + \alpha\Phi_t(\mathbf{m}) \quad (1)$$

where \mathbf{m} is a vector with the model parameters, Φ_d is a measure of the difference between the measured and calculated apparent resistivity values. Φ_s and Φ_t are measures of the model spatial and temporal roughness (Oldenburg and Li, 1994; Farquharson and Oldenburg, 1998; Kim et al., 2010). λ and α are the spatial and temporal damping factor vectors (Loke et al., 2014a) that gives the relative weights of the model roughness constraints compared to reducing the data misfits. 4-D surveys and models frequently have thousands of data points and model parameters (Poje et al., 2018), so a local optimization method such as the smoothness-constrained least-squares method is commonly used (deGroot-Hedlin and Constable, 1990; Oldenburg and Li, 1994). While the solution to equation (1) provides a model that fits the data to the required degree with the prescribed model roughness in space and time, it does not explicitly constrain the direction of the changes with time. In some surveys, it is known from other physical or hydrological information that the resistivity in at least part of the subsurface will only decrease (or increase) with time (Putman, 2014; Rucker et al., 2014; Johnson et al., 2015) within a certain time frame. In this paper, we examine two possible methods for enforcing the selected directional changes with time in the inverse model. To simplify the following discussion, we assume the resistivity is expected to decrease with time for a case with data sets at two time steps. However, the same methods can be generalized for cases where the resistivity is expected to increase with time and for more time steps.

2.2. Barrier (or penalty) function method

This method adds an additional term to the objective function so that the model is directed away from a certain parameter range.

$$\Phi(\mathbf{m}) = \Phi_d(\mathbf{m}) + \lambda\Phi_s(\mathbf{m}) + \alpha\Phi_t(\mathbf{m}) + \beta\Phi_b(\mathbf{m}) \quad (2)$$

Φ_b is a barrier function (Gill and Murray, 1974) term that has very large values as the model values approaches the limits which the model parameters \mathbf{m} can take, and small values elsewhere. β is a vector with the weights given to the barrier function constraint for the model cells. This method is fairly widely used in geophysical inversion. It was used by Li and Oldenburg (2003) to impose a constraint so that the inverse model magnetic susceptibility is always greater than zero by using the logarithm of model value in the barrier function term.

This was also used by Putnam (2014) to impose a constraint so that the resistivity of cells in a later time model are always less than that of the initial time model. An error function (Press et al., 2007) form of the barrier was used by Johnson et al. (2015). Below is the formula for the error function barrier used in this paper so that the resistivity of the second temporal model is less than the first temporal model.

$$\delta_j = 0.5 \operatorname{erf} \left[\left(\frac{m_{2,j} - m_{1,j}}{\gamma} \right) + 1 \right] = 0.5 \left[\operatorname{erf} \left(\frac{\log(\rho_{2,j} / \rho_{1,j})}{\gamma} \right) + 1 \right] \quad (3)$$

$m_{1,j}$ and $m_{2,j}$ are respectively the logarithm of the resistivity for the j th model cell of the first and second temporal models. γ is a parameter (usually 0.01) that controls the steepness of the barrier term. Figure 1 shows a plot of the barrier function. It rises steeply when the ratio of the resistivity of the second and first temporal models (ρ_2/ρ_1) approaches 1.0. At (ρ_2/ρ_1) values less than 0.98 the function has almost zero value so the constraint is not imposed. It has a value of 0.5 when the (ρ_2/ρ_1) ratio is 1.0, and reaches a constant value of 1.0 for (ρ_2/ρ_1) values beyond 1.02. Another form of barrier function (Putnam, 2014) uses the logarithm of the difference between first and second resistivity value, $\log(\rho_1 - \rho_2)$, as the barrier function parameter. However, this function has a discontinuity if ρ_2 is equals to ρ_1 , so special precautions have to be taken so that the threshold is not reached. To avoid the discontinuity, we instead examine the use of the power function below that also increases rapidly near the threshold but still has a continuous form.

$$\delta_j = 0.5 \left(\frac{\rho_{2,j}}{\rho_{1,j}} \right)^q \quad (4)$$

Figure 1 shows a plot of the function when the exponent q is equals to 40. Beyond a resistivity ratio of about 1.01, the function rises steeply above the error function plateau of 1.0 (Figure 1). This function has a longer tail and starts to have significant values for (ρ_2/ρ_1) ratios above 0.9. Using higher values of the exponent (such as 60) will shorten the tail for (ρ_2/ρ_1) below 1.0 and increases the steepness of the function above 1.0, but in practice the effect on the inverse model is not significant. Figure 1 also shows the power function curve when the exponent q is equals to 20. It has a much longer tail extending to resistivity ratio values of 0.8 which is not desirable as this will distort a larger region of the objective function away from the threshold of 1.0.

The barrier function constraint attempts to force the resistivity of a cell in the second temporal model to be close to that of the first temporal model when the resistivity ratio exceeds 1.0, or a resistivity change close to zero. In regions of the model space away from the threshold

value, 1.0 in this case for (ρ_2/ρ_1) , the barrier function term has very small amplitudes. Thus, parts of the inverse model where the resistivity ratio is significantly less than 1.0 should be similar to that obtained without this constraint. However, the barrier function constraint does not prevent the model resistivity values to exceed the threshold although it strongly discourages it.

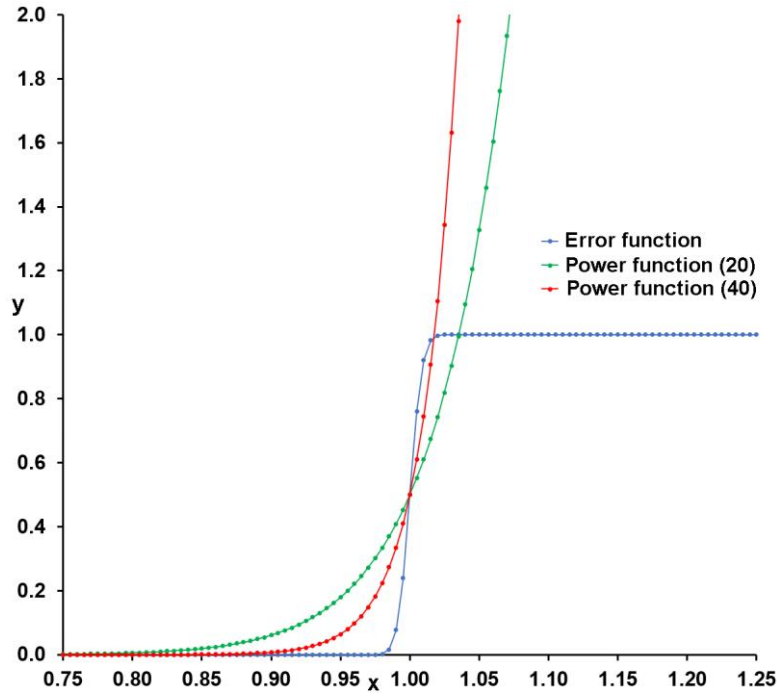


Fig. 1. Shape of the barrier functions used. The x -axis is given as (ρ_2/ρ_1) , or the ratio of the resistivity of the second temporal model to the first temporal model. The y axis is the barrier function value.

In this paper, we incorporate the barrier function constraint into the 4-D smoothness-constrained least-squares method (Kim et al., 2009; Loke et al., 2014a). This method inverts the data from a number of times series simultaneously. It has been found to be less sensitive to noise and gives more accurate results than inverse methods that directly use the differences in the model or apparent resistivity values (Kim et al., 2010; Karaoulis et al. 2011, 2014). This method is more efficient if there are more than two time steps. It is not unusual for field surveys to have data sets with more than 10 time steps measurements (Loke et al., 2014a; Rucker et al., 2014; Poje et al., 2018). We first start with the least-squares equation used by the standard 4-D inversion method which is given by

$$\begin{aligned} & \left[\mathbf{J}_i^T \mathbf{R}_d \mathbf{J}_i + (\lambda_i \mathbf{W}^T \mathbf{R}_m \mathbf{W} + \alpha_i \mathbf{M}^T \mathbf{R}_t \mathbf{M}) \right] \Delta \mathbf{m}_i = \mathbf{J}_i^T \mathbf{R}_d \mathbf{g}_i \\ & - (\lambda_i \mathbf{W}^T \mathbf{R}_m \mathbf{W} + \alpha_i \mathbf{M}^T \mathbf{R}_t \mathbf{M}) (\mathbf{m}_{i-1} - \mathbf{m}_r). \end{aligned} \quad (5)$$

The Jacobian matrix \mathbf{J} contains the partial derivatives of the data with respect to the model parameters. λ and \mathbf{g} are the damping factor and data misfit vectors. \mathbf{m}_{i-1} is the model parameter vector from the previous iteration, while \mathbf{m}_r is a reference model. $\Delta \mathbf{m}_i$ is the change in the model parameters which is to be calculated. \mathbf{W} includes the spatial roughness filters in the x , y and z directions. \mathbf{R}_d , \mathbf{R}_m and \mathbf{R}_t are weighting matrices used by the iteratively reweighted least-squares method if the L1-norm is used (Farquharson and Oldenburg, 1998; Loke et al., 2003; Loke et al., 2014a). \mathbf{M} is the difference matrix applied across the temporal models so that the changes occur in a smooth manner with time (Kim et al., 2009). It has a diagonal and one sub diagonal with values of 1 and -1, respectively. α is the temporal damping factor vector with the weight given to minimizing the temporal changes in the inverse model. The modified least-squares equation that incorporates the barrier function is given below.

$$\begin{aligned} & \left[\mathbf{J}_i^T \mathbf{R}_d \mathbf{J}_i + (\lambda_i \mathbf{W}^T \mathbf{R}_m \mathbf{W} + \alpha_i \mathbf{M}^T \mathbf{R}_t \mathbf{M} + \beta_i \mathbf{B}_i^T \mathbf{B}_i) \right] \Delta \mathbf{m}_i = \mathbf{J}_i^T \mathbf{R}_d \mathbf{g}_i \\ & - (\lambda_i \mathbf{W}^T \mathbf{R}_m \mathbf{W} + \alpha_i \mathbf{M}^T \mathbf{R}_t \mathbf{M} + \beta_i \mathbf{B}_i^T \mathbf{B}_i) (\mathbf{m}_{i-1} - \mathbf{m}_r) \end{aligned} \quad (6)$$

\mathbf{B} is a difference matrix that links the values of the model parameters for the first temporal model with the corresponding model values for the later temporal models, and β_i is the associated damping factor vector. As an example, for a data set with 3 time steps where each temporal model has n model cells, the matrix is given by

$$\mathbf{B} = \begin{pmatrix} D_1 & \dots & -\delta_{1,1} & \dots & -\delta_{2,1} & 0 & 0 & \dots & 0 \\ & D_2 & \dots & -\delta_{1,2} & \dots & -\delta_{2,2} & 0 & \dots & 0 \\ & & D_3 & \dots & -\delta_{1,3} & \dots & -\delta_{2,3} & \dots & 0 \\ & & & \dots & \dots & \dots & \dots & \dots & 0 \\ & & & & D_n & \dots & \dots & \dots & 0 \\ & & & & 0 & 0 & \dots & \dots & 0 \\ 0 & & & & & 0 & 0 & \dots & 0 \end{pmatrix}. \quad (7)$$

The value of the diagonal term D_j for the j th model cell of the first temporal model for a data set with t data time steps is given by

$$D_j = \sum_{i=1}^t \delta_{i,j}. \quad (8)$$

$\delta_{i,j}$ is the barrier function value for the j th model cell of the i th temporal model. Putnam (2014) used a simple diagonal matrix for \mathbf{B} but reported problems in the convergence of the algorithm.

Our tests indicate that using a difference matrix results in a more stable inversion algorithm compared to a diagonal matrix.

In the standard 4-D inversion method, a homogeneous half-space is commonly used as the reference model \mathbf{m}_r as well as the starting model. However, since the barrier function value depends on the ratio of the resistivities of the same model cell for two temporal models, we cannot use same starting model for all the temporal models. To create the starting model for the first time step, we carry out an inversion of the first time step data set (Appendix). The model obtained is used as the starting model \mathbf{m}_1 for the first time step, as well as the reference model for all the time steps. We find that using the model from the inversion of the first temporal data set alone as the starting model for the first temporal model in the 4-D inversion algorithm improves the convergence compared to using a homogeneous starting model. This is probably because the inverse model from a standalone inversion of the first temporal data set is close to the optimum model for the first temporal model in the 4-D inversion algorithm. To set the second starting temporal model, the ratio of the average apparent resistivity value of the second time step to the average value for the first time step is used. We then multiply this ratio with the starting model \mathbf{m}_1 for the first time step to get the starting model \mathbf{m}_2 for the second time series. A maximum value of 0.8 is set for the multiplication factor to ensure that the resistivity of the model cells for the second time step is always less than the first time step (assuming the resistivity is expected to decrease with time). This ensures that the starting model for the second time step is outside the zone where the barrier function is active. The same procedure is used to set the starting models for the other time steps if there are more than two.

2.3. Method of transformations

This method does not add an additional penalty term to the objective function but instead changes the mathematical form of the model parameter so that the constraints are implicitly imposed. The same objective function form, such as in equation (1), is used but the definition of the model parameter \mathbf{m} is changed. This technique has been long used in numerical analysis (Daniels, 1978) to place limits on the range of values a parameter can take, but it is probably less well known in geophysical inversion. One commonly used transformation in geoelectrical inversion is the logarithm of the model and apparent resistivity (Johansen, 1977; Johnson *et al.*, 2015) which is given below.

$$m = \log(\rho), \text{ or } \rho = \exp(m) \quad (9)$$

Resistivity values in a model or data set can have a range of several orders of magnitude. This is reduced to a linear range by using the logarithmic transformation. In automatic inversion of

resistivity sounding data, it has been found that the logarithmic transformation results in a more stable algorithm that converges more rapidly to essentially the same solution (Johansen, 1977). Another benefit of using the logarithmic transformation is that the inverse model resistivity value is always greater than zero. This avoids the use of a barrier function constraint to impose this condition as it is implicitly satisfied by the transformation. However, the effect of the transformation method is dependent on the characteristics of the model parameter as it changes the shape of the objective function. The solution found at the minimum of the transformed objective function is optimal for the transformed parameter which might be different from the optimal value for the original parameter (Box, 1966; Daniels, 1978). One example is the inversion of I.P. data. It is known that intrinsic I.P. has a fixed range of 0.0 to 1000.0 mV/V (although in practice the maximum value is probably less than 300 mV/V for natural minerals). It is possible to use the logarithmic transformation in I.P. inversion to impose a lower limit of 0.0 for the model values. However, Oldenburg and Li (1994) found that using the logarithmic transformation gave a slightly poorer I.P. inverse model. This is probably because the I.P. parameter has a linear rather than a logarithmic range.

As the method of transformations is not well covered in exploration geophysics literature, we briefly described three common transformations (Daniels, 1978). If a parameter y is constrained to have an upper bound y_u , the following transformation can be used.

$$y = y_u - z^2, \text{ or } z = \sqrt{y_u - y} \quad (10)$$

This was used by Loke et al. (2018a) where the positions of the electrodes along a survey line (which had moved down a slope due to soil movements) for the second time step were estimated from the apparent resistivity data, the initial electrodes positions and apparent resistivity data from a prior survey. It is known that the electrodes movements (if any) will only be in the downwards direction, so the previous position of an electrode (y_u) forms an upper limit. Carrying out an inversion with the transformed variable z ensures that the new position of an electrode estimated from the inverse routine is never higher than the original position. If a parameter y has a lower bound y_l , the corresponding transformation is given by

$$y = y_l + z^2, \text{ or } z = \sqrt{y - y_l} \quad (11)$$

The following transformation is used if the parameter has both a lower and upper bound.

$$y = y_l + (y_u - y_l)(\sin z)^2, \text{ or } z = \sin^{-1} \left[\frac{y - y_l}{y_u - y_l} \right]^{0.5} \quad (12)$$

One use of this transformation is in I.P. inversion where the model values are limited to between 0.0 and 1000 mV/V (Loke, 2021). Other more complex transformations are discussed in Box (1966).

2.4. Algorithm to use the method of transformations in 4-D time-lapse inversion

To enforce a constraint so that the resistivity of the second temporal model is always less than that of the first temporal model, both the lower and upper bound transformations are used. Equations (10) and (11) are meant for a case with a single model parameter. However, in time-lapse inversion, the constraints are applied across two sets of model parameters corresponding to the resistivity values for the first and second temporal models. The following algorithm is used to adapt the method of transformations for this type of cross-model constraint.

1). Initially the standard 4-D least-squares smoothness-constrained inversion technique is used as this generally produces a stable inverse model. This model is then used as the base model for the modified inversion method. The temporal models from this initial inversion, where the direction of changes in time is not constrained, might have regions where the resistivity increases with time instead of the expected decrease. The next two steps are then used to remove these artifacts.

2). The resistivity of each model cell of the first temporal model is compared with the corresponding cells in the second temporal model to identify the model cells where the resistivity had increased with time. The mean value (p_j) of the j th model cell is calculated from the resistivity first (r_{1j}) and second (r_{2j}) temporal models.

$$p_j = (r_{1j} + r_{2j}) / 2 \quad (13)$$

If the j th model cell (r_{1j}) for the first temporal model is lower than p_j , it is reset to this mean value. If the same cell for the later temporal model (r_{2j}) is higher than p_j , it is also reset to p_j . This procedure will reset the resistivity temporal change for cells with an increase in the resistivity to zero, but does not affect the other model cells. Resetting the resistivity values of some model cells sometimes cause an increase in the data misfit. The following refinement step is then used to decrease the data misfit. If there are more than two time steps, the maximum resistivity value of the same model cell for all the later temporal models is used in place of r_{2j} .

3). The optimization process is then continued using the transformed parameters so that the resistivity of any cell in the second temporal model is always lower than the corresponding cell in the first temporal model. The cell resistivity (r_{1j}) for the first temporal model is replaced by

a new variable, s_{lj} , using the following equation that ensures the cell resistivity is always higher than the mean value p_j .

$$r_{lj} = p_j + s_{lj}^2, \text{ or } s_{lj} = \sqrt{r_{lj} - p_j} \quad (14)$$

A similar transformation is used for the cell resistivity (r_{2j}) for the second temporal model to ensure that it is always lower than the mean value p_j .

$$r_{2j} = p_j - s_{2j}^2, \text{ or } s_{2j} = \sqrt{p_j - r_{2j}} \quad (15)$$

The two transformations ensure that the resistivity of the second temporal model is always less than the first temporal model. We then carry out a one or two more iterations using the least-squares optimization method with the transformed variables. The mean values p_j are kept fixed during this model refinement step. The data misfit is usually reduced to a level similar to that achieved without the additional constraint while producing a model without the temporal artifacts.

The least-squares optimization method uses the sensitivity values such as $\frac{\partial \phi_i}{\partial r_j}$ that gives the change in the measured potential at the i th data point due to a change in the resistivity for the j th model cell. When the transformed parameters are used, the required sensitivity values are replaced as follows for the transformation in equation (14).

$$\frac{\partial \phi_i}{\partial s_{lj}} = \frac{\partial \phi_i}{\partial r_{lj}} \frac{\partial r_{lj}}{\partial s_{lj}} = 2s_{lj} \frac{\partial \phi_i}{\partial r_{lj}} \quad (16)$$

A similar calculation is made for the transformation in equation (15).

3. Results

This section shows results obtained with the barrier function and the transformation methods for one synthetic data set (where the true structures are known) and a field data set. The 4-D smoothness-constraint least-squares optimization method (Kim et al., 2009; Kim et al., 2010) was used to invert the data sets. We use the L1-norm method for the data misfit and the model roughness (both in the spatial and temporal domains). Further details on the practical use of this method, such as the use of different norms and selection of the damping factor, can be found in the references (Farquharson and Oldenburg, 1998; Farquharson and Oldenburg, 2004; Loke et al., 2014a).

3.1. Synthetic data set

In the studies by Putman (2014) and Johnson et al. (2015), the temporal artifacts occurred in regions of the model that were relatively far from the electrodes and had low sensitivity values. For the synthetic model test, we examine a case where the artifacts occur near the surface close to the electrodes that should be well resolved. The synthetic model used to generate the test data sets is shown in Figure 2a. The first temporal model consists of a 500 Ωm structure (between depths of 0.0 to 1.08 m) embedded in a 100 Ωm homogeneous background medium (Figure 2a). The second model has a low resistivity 20 Ωm region (between depths of 0.0 to 1.74 m) next to the high resistivity structure to simulate an infiltration experiment. The low resistivity region extends below the high resistivity structure.

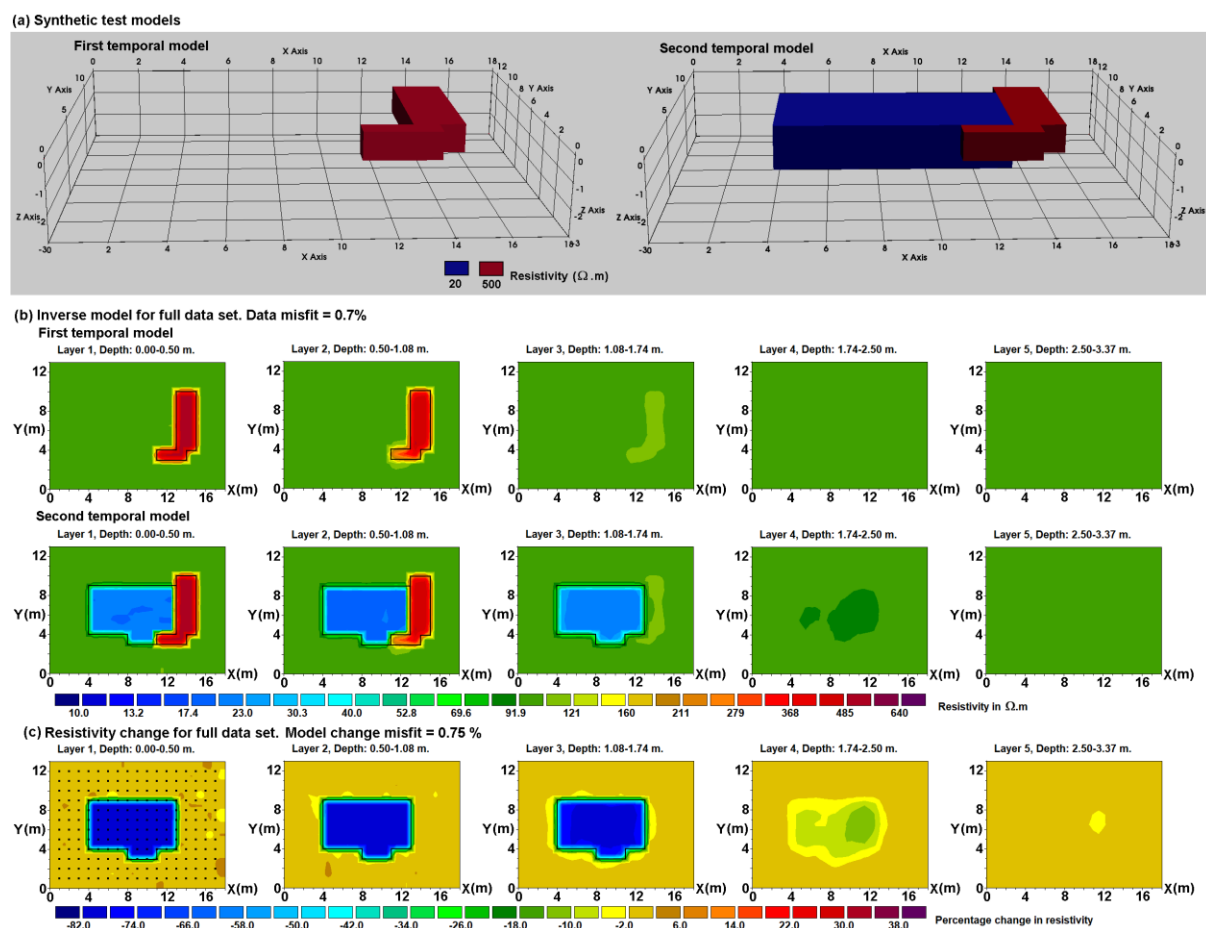


Fig. 2. (a) Synthetic test models used to generate the first and second data sets. The homogeneous background medium has a resistivity of 100 Ωm . (b) Inverse models for synthetic model full test data set. (c) Difference sections from the inversion of the full data set. The small black dots in the leftmost section in (c) show the positions of the electrodes used to generate the data set. The model blocks shown in (a) are marked by black outlines in (b) and (c).

The electrode array used to generate the test data set is the dipole-dipole array with the ‘ a ’ dipole length fixed at 1 m, and the dipole separation factor ‘ n ’ varying from 1 to 8. The electrode spacing used is 1 m with the electrodes located along lines 1 m apart in the y -direction. Only measurements along the x -direction are used. The electrodes used are located between 1 and 17 m in the x direction, and 1 and 12 m in the y direction giving a total of 204 electrodes positions (Figure 2c). The resulting data set has 1008 data points for each temporal model. Voltage dependent Gaussian random noise (Zhou and Dahlin, 2003) of 1 m Ω was added to the calculated resistance values before they were converted to apparent resistivity values. This resulted in an average noise level of about 0.7% in the final time-lapse data set.

Figure 2b shows the first and second temporal resistivity models obtained with the standard 4-D inversion method. The models agree well with the true structures. We follow the approach by Maurer and Friedel (2006) in extending the region covered by the model cells slightly beyond the grid of electrodes (Figure 2c). The data misfit is about 0.7% which is similar (to within 0.05%) to the added noise level. We next calculate the percentage change in the resistivity for each model cell using the following equation.

$$\varepsilon_{21}(j) = 100 \cdot [r_2(j) - r_1(j)] / r_1(j) \quad (17)$$

Figure 2c shows a plot of the percentage difference sections. It shows an area with changes of up to about -80% that corresponds to the low resistivity structure in the second model. Most of the regions outside of this structure in the top three layers (as well as layers 5 and 6) have changes of less than 2%. There are small areas with changes of up to about +5% in the top layer, particularly at the edges of the model beyond the grid of electrodes (Figure 2c). This is probably partly due to the noise added and poor data coverage. Layer 4 shows a region with changes of up to about -14% below the low resistivity structure that is probably due to the decrease in resolution with depth where the bottom boundary is not well resolved. To assess the accuracy of the time-lapse model, we calculate the average absolute difference (δ_A) between the true percentage change in the resistivity (δ_T) and that obtained by the inverse model (δ_M).

$$\varepsilon_A = \frac{1}{n} \sum_{j=1}^n |\varepsilon_M(j) - \varepsilon_T(j)| \quad (18)$$

n is the number of model cells. The average difference in the true and calculated changes for the time-lapse model is 0.75%. The largest contribution comes from the region in Layer 4 (Figure 2c) with the decrease in resistivity values below the bottom boundary of the low resistivity region.

We next carry out another test with a reduced data set by only using the data points where the x -positions of all the electrodes are located between the 3 and 14 m marks giving a data set with 528 data points. This was done to explain the source of artifacts observed in a field data set (described in the following section). The difference sections from the inverse models for the reduced data set show significant artifacts with increases in the resistivity of up to about +29% at the position of the high resistivity structure (Figure 3b). This is probably caused by the large contrast (of about 25 times) between the low and high resistivity structures, and that there is insufficient data to fully resolve the high resistivity anomaly as the survey lines do not completely cross the structure (Figure 3b). The average difference between the true and calculated resistivity changes is 0.87% (Table 1).

Since it is known that the resistivity should only decrease with time in the synthetic data set, we next try different methods to reduce the artifacts in areas where the resistivity increases with time. The result with the barrier method using the error function is shown in Figure 4a. It does reduce the artifact with the increase in the resistivity at the edges of the low resistivity structure. The maximum increase in the resistivity is about +5% in the top layer at the x - y position of (4.5 m, 9.5 m) near the edge of the low resistivity region (Figure 4a). The barrier function constraint does significantly reduce the artifacts with an increase in the resistivity but does not completely eliminate them. This is similar to the results obtained by Johnson et al. (2015) where it was expected the conductivity of the region in the saturated zone should only decrease due to influx of colder water from the Columbia River at the Hanford survey site. The addition of the barrier function constraint did greatly reduce zones with increases of about +1.5 S/m but still left residual artifacts of about +0.15 S/m. The overall difference in the calculated and true changes in the resistivity (equation 18) was reduced to 0.73%.

Figure 4b shows the results obtained when the power function (with the exponent q equals to 40) was used to set the barrier threshold. The areas with an increase in the resistivity were reduced (maximum increase of +4%) and the average difference between the true and calculated changes was slightly lower at 0.68%. The method of transformations (Figure 4c) eliminated the artifacts with positive resistivity changes but had a slightly higher average misfit in the changes of 0.76% (Table 1). Comparing Figures 3b and 4c, the transformation method did remove the regions with positive resistivity changes but had no effect on the artifacts with negative resistivity changes.

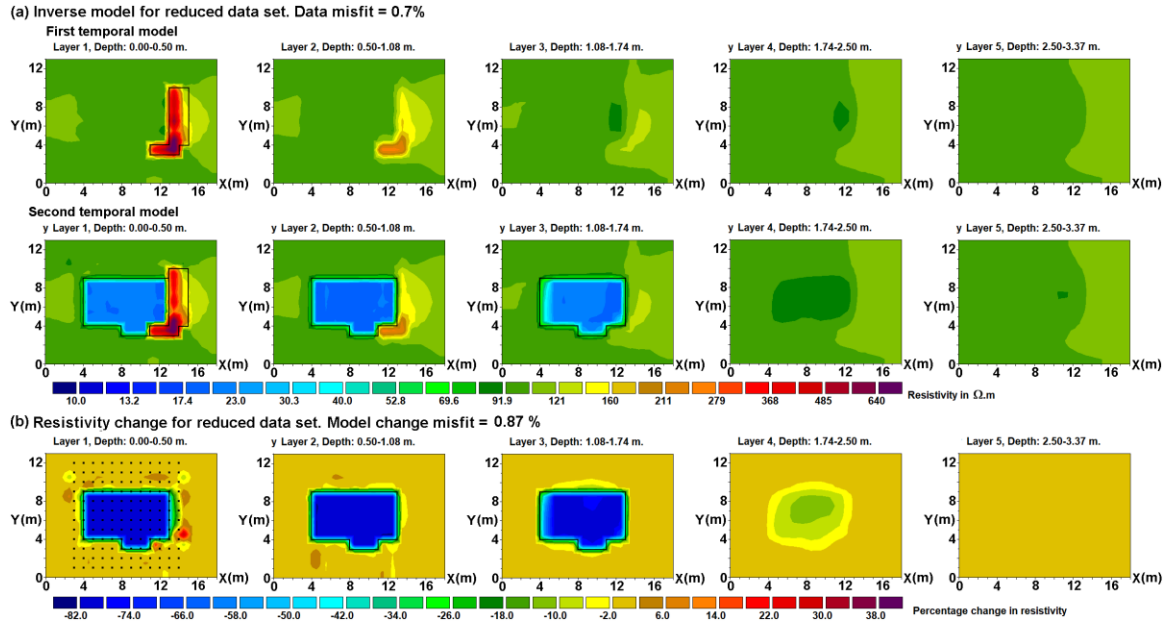


Fig 3. (a) Inverse models for synthetic model reduced test data set. (b) Difference sections from the inversion of the reduced data set using standard 4-D inversion method. The positions of the anomalous blocks in the synthetic model are marked by black outlines in (a) and (b).

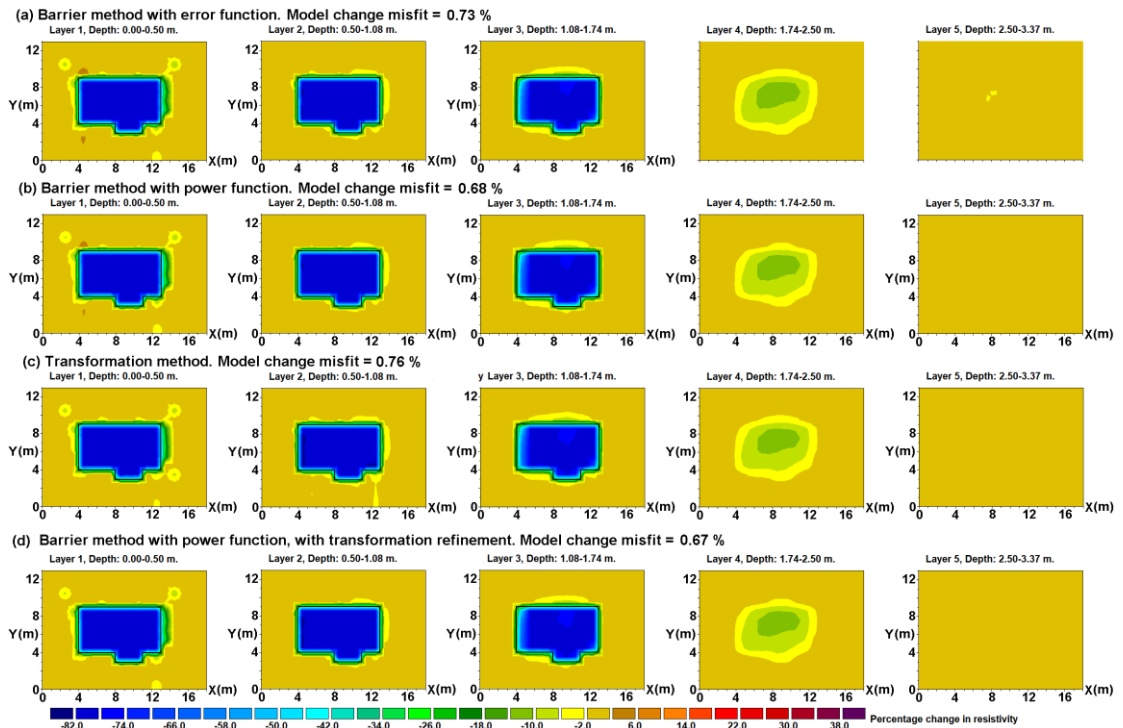


Fig. 4. Difference sections from the inversion of the reduced data set using (a) barrier method with error function, (b) barrier method with power function, (c) transformation method and (d) barrier method (with power function) followed by transformation method. The data misfits for all the inverse models are about 0.7%. The position of the low resistivity region where the temporal change occurs in the synthetic model is marked by a black outline in the sections.

The method of transformations starts with the model produced by the conventional 4-D inversion method (Figure 3b) with significant artifacts. The barrier function methods significantly reduced the artifacts, including those with negative resistivity changes. We next use the model produced by the barrier method using the power function as the starting model for the transformation method as this had the lowest model change misfit. The transformation method is now used as a refinement step (after an initial inversion with the barrier function method) to eliminate any residual artifacts with positive resistivity changes. The results are shown in Figures 4d. The transformation method refinement removed the positive resistivity changes while leaving the rest of the model unchanged which resulted in a small reduction in the model change misfit value to 0.67% (Table 1). This method combines the strengths of the barrier function and transformation methods.

Table 1. Error in the recovered percentage resistivity changes by the different inversion methods for the synthetic model reduced data set.

Inversion method	Average error in percentage resistivity change
Standard 4-D inversion method	0.87
Error function barrier method	0.73
Power function barrier method	0.68
Transformation method	0.76
Power function barrier plus transformation method	0.67

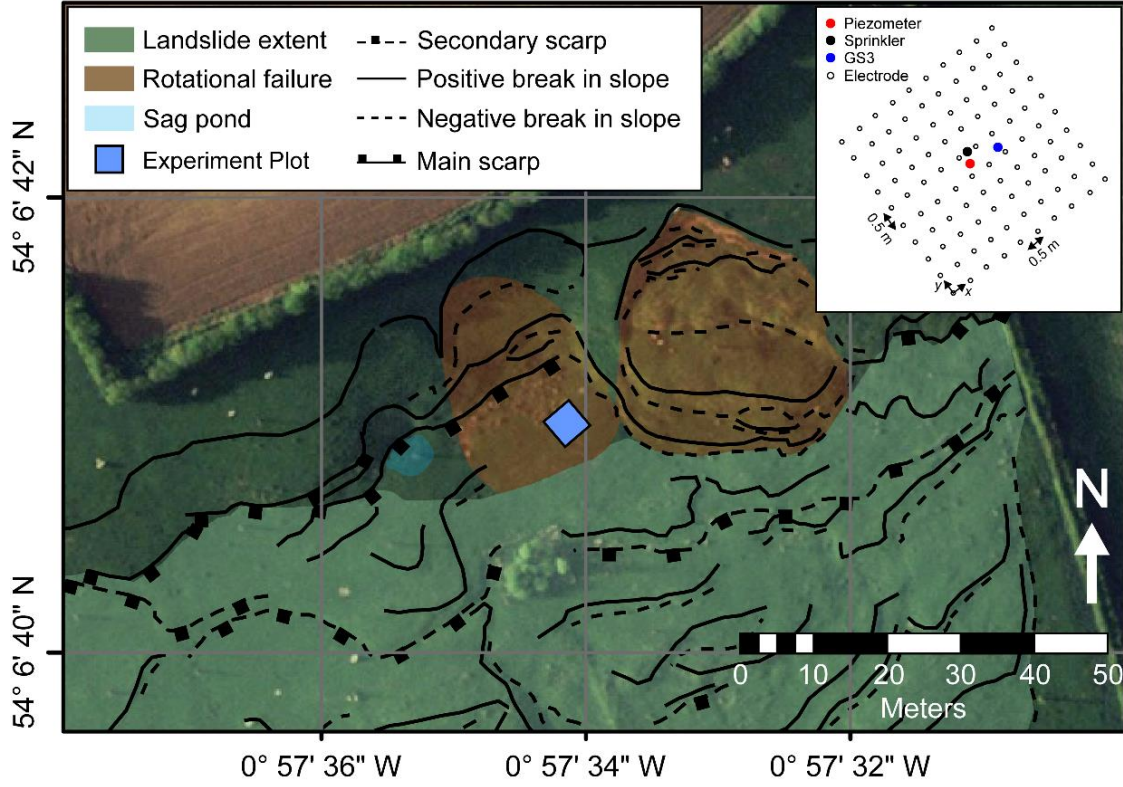
3.2. Hollin Hill field data set

An infiltration experiment was conducted at the BGS Hollin Hill landslide research site (Uhlemann et al., 2017) where a saline solution was sprinkled on the surface to study the effects of fissuring on hydrological processes and landslide initiation. Figure 5a shows the location of the test plot, which had an area of $\sim 20 \text{ m}^2$. The inset shows the grid of 100 electrodes, and the relative locations of a GS3 soil moisture sensor and piezometer probe. The piezometer was installed from 2.15 m depth up, with sand filling up to 0.35 m below ground surface and the top filled with bentonite. The depths of the soil moisture probes were 0.05, 0.1, 0.2 and 0.4 m. The saline solution had a conductivity of 0.26 S/m (resistivity = $3.8 \text{ } \Omega\text{m}$), and $\sim 1 \text{ m}^3$ was sprinkled over the plot for about 6.4 hours at a rate of 7.5 mm/hour. Figure 5b shows the

piezometer values before, during and after the infiltration experiment. There was a significant increase in the piezometer values after the start of the saline solution sprinkling, and they remain well above the background values before the start of the experiment even after the sprinkling had ceased. Eight time-lapse resistivity data sets were collected that spanned the infiltration experiment. Each took about 1 hour to measure (horizontal bars in Figure 5b) and the first set (starting at $t = 0$) was collected before infiltration began to determine the background resistivity. Sprinkling started at $t = 1.3$ hours and lasted until $t = 7.7$ hours. The resistivity data were measured with an AGI Supersting R8 multichannel resistivity meter using a combination of linear and equatorial dipole-dipole configurations. Each set contained 1832 measurements comprising: linear dipole-dipoles along the x , y and diagonal directions with dipole lengths (a) of 1, 2, and 3 electrode spacings (x or y : $a \approx 0.5$ m, diagonal: $a \approx 0.71$ m) and all possible dipole separations (na) with a maximum of $n = 7$; and equatorial dipole-dipoles in the x and y directions with $a \approx 0.5$ m and all possible dipole separations (na) with a maximum of $n = 9$. Reciprocal measurements were made for the last set, several hours after infiltration had finished (indicated by 7 and 7R in Figure 5b). For each reciprocal measurement pair, the standard error in the mean of the data was used as an indicator of the level of noise in the data. The distribution of reciprocal errors indicated very good data quality with 98.2% of the data having errors of less than 5%.

Figure 6a shows the resistivity inverse model from the initial survey data set before the start of sprinkling. The cells in the inverse model have widths of 0.25 m (half the electrode spacing) in the x and y directions. We use a post-inversion L-curve method to determine the optimum damping factor and the associated target data misfit (Appendix) for the initial survey data set. The top drier layer has a higher resistivity of over 15 Ωm , while the deeper zone below 0.5 m has lower values of less than 15 Ωm . Figure 6b shows the model at $t = 7.3$ hours (using the standard 4-D inversion method) when the infiltration has almost reached its maximum extent. There are visible changes in the resistivity in the top 3 layers. Figure 6c shows the percentage resistivity model change between $t = 0$ and $t = 7.3$ hours. It shows a large region with negative changes in the top two layers due to the addition of saline solution. However, there is a band to the right of the infiltration zone in the top layer with increases of up to +30%, which is an artifact. There had been no rainfall for several days prior to the infiltration experiment, so there was unlikely to have been any significant drying over the 8 hours of the sprinkling. This is supported by Figure 5b, which shows the data from moisture content sensors at four depths during the course of the experiment.

(a) Hollin Hill experimental site



(b) Plot of saturation and piezometric level measurements

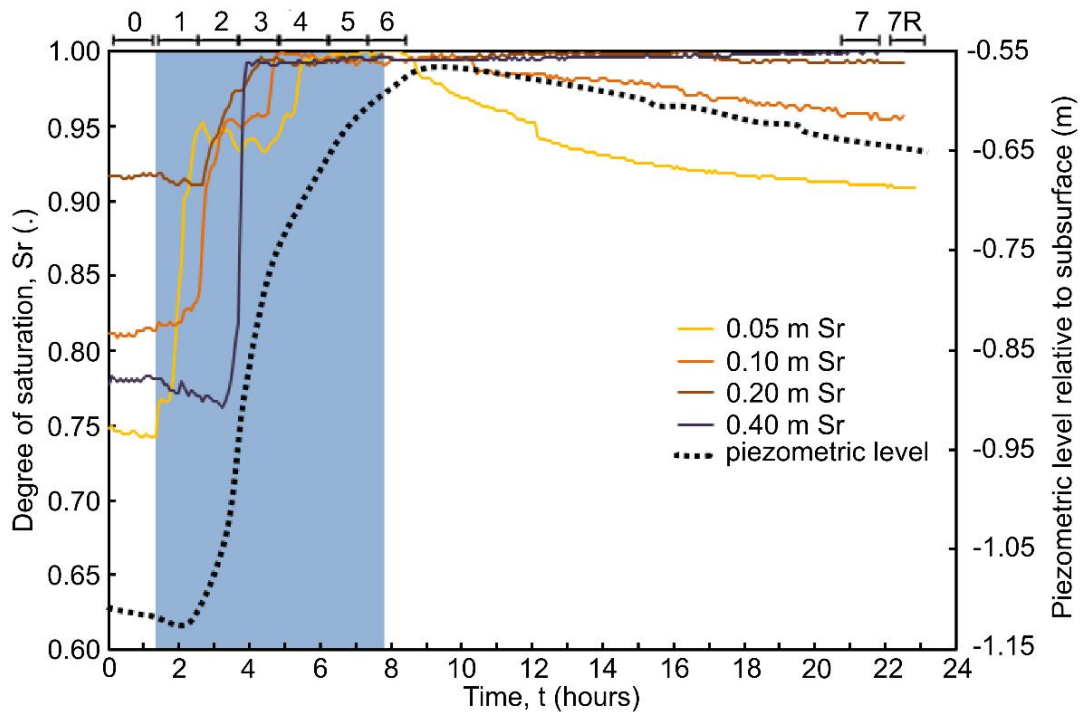


Fig. 5. (a) Location of Hollin Hill infiltration experimental site. GS3 marks the position of the soil moisture sensor. (b) Change of the saturation (S_r) and piezometric measurements over the course of the infiltration experiment. Sprinkling period indicated by light blue shaded area, and ERT data collection periods shown by labelled horizontal bars.

To enable better comparison of the state of the soils we converted the volumetric moisture content (VMC) to the degree of saturation, S_r . S_r is a ratio that indicates how much water is filling the available pore space (at 0 the pore space is completely dry, at 1 the soil is completely saturated). It is calculated by dividing the VMC by the porosity of the relevant layer (taken as the maximum VMC encountered in the time series of observations for that layer). The degree of saturation shows a clear pattern (Figure 5b). The top three layers (0.05 m, 0.10 m and 0.20 m) show no significant changes in the degree of saturation before the test commences. The material at 0.40 m has a lower ambient moisture content than the layers above at 0.10 m and 0.20 m. During the experiment, the degree of saturation rapidly increases as water infiltrates the soil. At 0.05 m the response is immediate and the delay of the responses at 0.10 and 0.20 m depth show that infiltration in the top three layers are well connected and an infiltration front is progressing through the soil. Particularly the upper two layers display this infiltration process very well. These hover around 95% saturation ($S_r=0.95$), an indication that this is the degree of saturation where vertical flow of water is most effective. It is interesting to note that the degree of saturation at 0.40 m depth does not respond initially, but after about 2 hours 45 minutes into the test the infiltration front has reached this depth and it rapidly fills up the pores (it is likely to be impeded in its progress by deeper, less permeable strata where infiltration is assumed to occur mainly through local fissure flow). Once the pores at 0.40 m depth are filled with water, it can be observed that the strata above are gradually filling up their pore space as well. There is an interval sequence of these layers reaching saturation ($S_r=1$) with 0.40-0.20 m of 20 minutes, 0.20-0.10 m of 35 minutes, and 0.10-0.05 m of 45 minutes. In turn, this generates a further head that drives water into the deeper strata. The result is that the piezometric level gradually increases until it reaches a maximum level (about 0.6m above pre-test level) after about 8 hours since the test started. After the completion of the test, the soil gradually drains as indicated by the sequential drop in degree of saturation of the four layers and the piezometric level. Taking the 0.10 m and piezometric level as a guide, it can be estimated that the bulk permeability of this soil is somewhere around 4×10^{-7} m/s.

Clément et al. (2009) have shown that small scale structure can cause changes in apparent resistivity that have the opposite sign to changes in ground resistivity caused by shallow infiltration. Here, the regions with an increase in resistivity correspond approximately with the high resistivity areas of above 25 Ωm at the right side in the top layer. During dry conditions, shrinkage and soil cracking at the site cause near-surface fissuring and routes for preferential infiltration (Gunn et al., 2013) and higher resistivities (although any fissures did not reach the surface in this area). In addition, we note that the electrodes in the survey grid do

not completely cross the high resistivity structure towards the right end of the survey area (Figure 6c). This situation is similar to the reduced data set in the synthetic model data set, and probably accounts for the artifacts with an increase in resistivity in the time-lapse model.

Figure 6d shows the model using the barrier method (with the power function) followed by the transformation method that removes the artifacts with an increase in the resistivity. The artifacts where the resistivity increased with time have been removed, and it also shows changes in the fourth layer which were not detected by the standard inversion method. To show the small changes in the third and fourth layers more clearly, a logarithmic contour scale is used for the percentage change in the resistivity (Figures 6c, 6d and 6e). Figure 7a shows the percentage change in the inverse model resistivity with time at four different depths at the position of the GS3 probe. The average of four model cells in the x and y directions at the location of the GS3 probe was used (giving a combined cell width of 0.5 m in the x and y directions) to reduce effects of noise on an individual model cell. The changes in the resistivity at depths of 0.063 and 0.197 m depths (Figure 7a) are mirror images of the changes in the S_r values at 0.05 and 0.20 m depths. The plot for the model cell at 0.063 m shows a slight increase in the resistivity after the sprinkling has ceased (Figure 7a) which mirrors the changes in the S_r values at 0.05 and 0.10 m depths (Figure 5b). The inverse model shows relatively small resistivity changes at depth of 0.351 m. Note the drop in resistivity (Figure 7a) first occurs at 0.063 m depth, followed by 0.197 m and 0.351 m depths in succession due to the downwards migration of the saline solution.

Figure 7b shows the variation of the model resistivity values with time. The resistivity values at 0 hours (before the start of the infiltration) shows that the background resistivity decreases with depth, from about 23.6 Ωm at 0.063 m to 12.5 Ωm at 0.529 m depth. For this reason, the maximum percentage change in the resistivity values decreases from -35.8% in layer 1, to -16.2% in layer 2 to -4.0% in layer 3 (Figure 7a). The resistivity in layers 1 and 2 decreases to a minimum value of about 15.2 Ωm after 7.3 hours. The starting resistivity value at layer 3 of 14.0 Ωm is lower than this value, so the addition of the saline solution (which probably is also diluted as it migrates downwards) causes a relatively small change in the resistivity. The model also shows effects of water ingress into deeper strata (0.43 - 0.84 m), which were not resolved using the standard 4-D inversion. Deeper ingress was expected due to preferential flow through the piezometer borehole, although its location at (2.17 m, 2.23 m) does not exactly coincide with the center of the changes in the deeper layers.

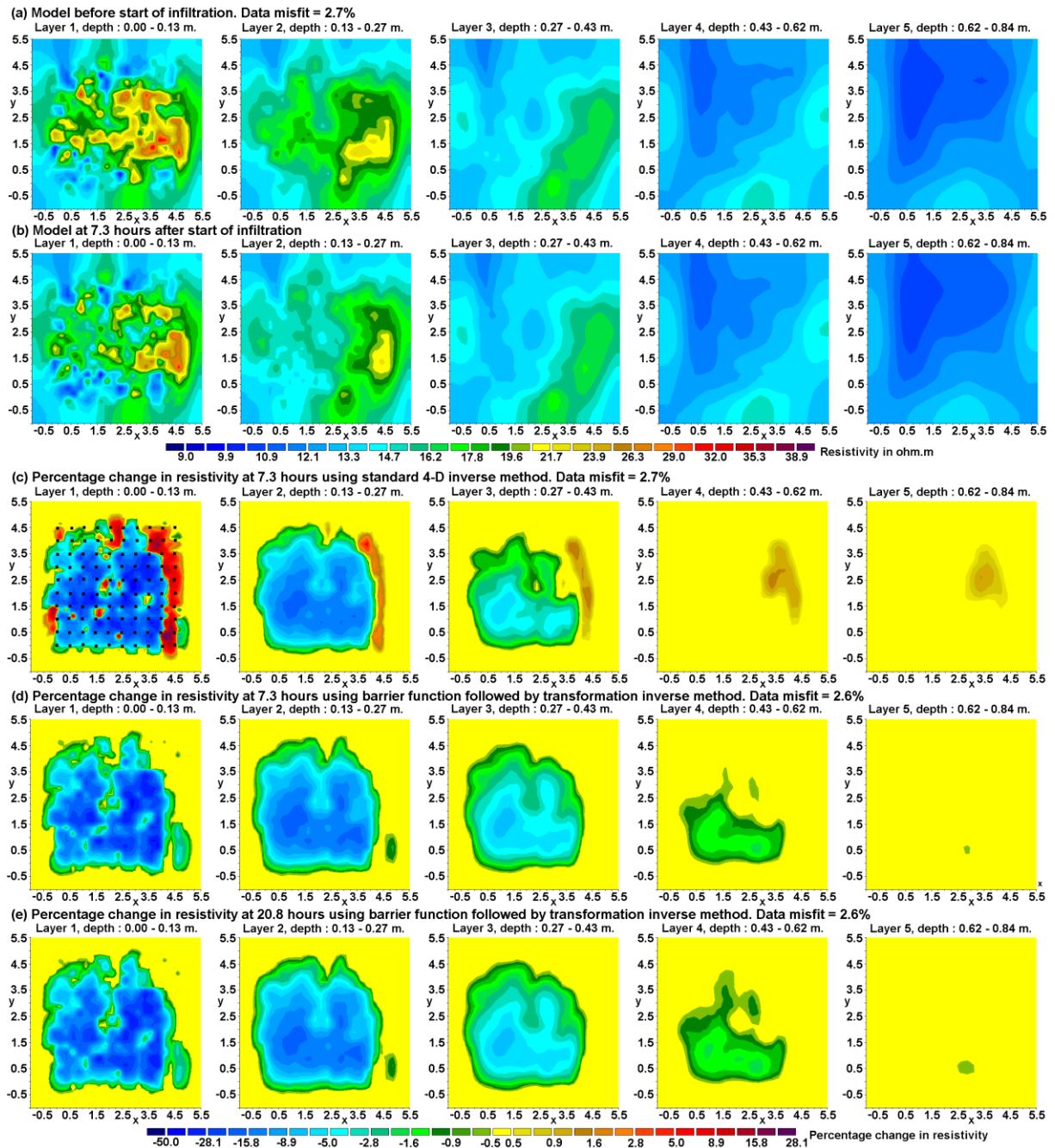


Fig. 6. Inverse models for Hollin Hill field data set using standard 4-D inversion method at (a) 0.0 hours before the start of the infiltration, and (b) 7.3 hours after start of infiltration. (c) Change in resistivity between 0.0 and 7.3 hours using standard 4-D inversion. Positions of the electrodes marked by black dots in the topmost section. Change in resistivity (d) between 0.0 and 7.3 hours, and (e) between 0.0 and 20.8 hours using barrier method (with power function) followed by transformation method. The yellow regions have resistivity change of -0.5 to +0.5 % in (c), (d) and (e).

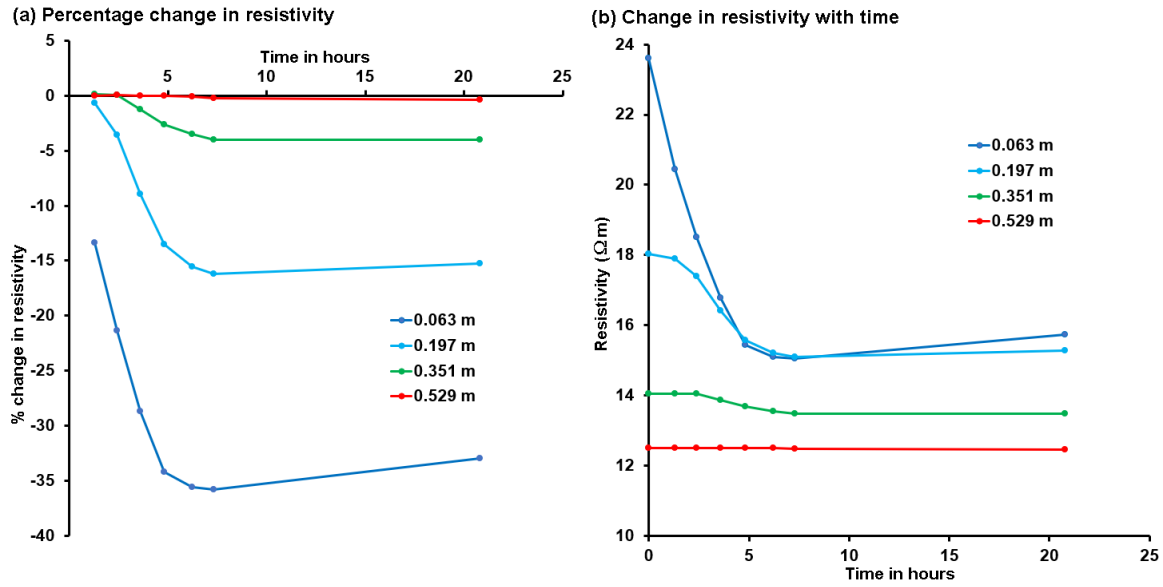


Fig. 7. (a) Variation of percentage change of resistivity in inverse model (average of 4 model cells) with time at different depths at location of GS3 probe. (b) Variation of inverse model resistivity values with time.

Figure 8 shows plots of the model resolution (using the diagonal elements of the resolution matrix) for the different layers for the data sets before the infiltration and at 7.3 hours. The model resolution (Day-Lewis et al., 2005) has a fixed range of between 0.0 and 1.0. The resolution values in the top layer (which is nearest to the electrodes) reaches a maximum value of about 0.8, and decreases exponentially with depth to less than 0.03 in the fourth layer. Thus, it is expected that resistivity changes in the deeper layers are less well resolved. The addition of the saline solution also affects the resolution that can be achieved by the data set. Note the resolution values in Layers 3 and 4 for the 7.3 hours data set (Figure 8b) are visibly lower compared to the initial data set (Figure 8a). This is because the lower resistivity values in the top layer reduces the amount of current that penetrates into the deeper layers. This makes the detection of changes in the deeper layers even more difficult.

Figure 9 shows a 3-D plot of the change in the resistivity with time from the model produced by the hybrid inversion algorithm using the barrier method (with the power function) followed by the transformation method. There are some lateral variations in the areas with significant decreases in the resistivity probably due to variations in the soil permeability, porosity and fissuring. The largest decrease in the resistivity occurs next to the high resistivity zone. The lateral spread of the areas with a decrease in the resistivity are shown by the sequence of models in Figures 9a to 9c (upper row). Note also the slight decrease in the intensity of areas with the highest negative resistivity changes in 9d compared to 9c (upper row) about 12 hours

after the end of the infiltration. The lower row in Figure 9a with a side view of the 3-D plots shows the downwards migration of the saline solution with time. The central part of the model (between the 1 and 4 m marks in the x -direction), which has better data coverage, shows a consistent downward movement of the plume from 2.4 to 7.3 hours. Between 7.3 to 20.8 hours (about 12 hours after the sprinkling had ceased), there is a smaller increase in the region with lowered resistivity values caused by the continuous downwards migration of the saline solution. This is also shown in Figures 6d and 6e by a slight increase in the extent of the region with lower resistivity values in the fourth layer.

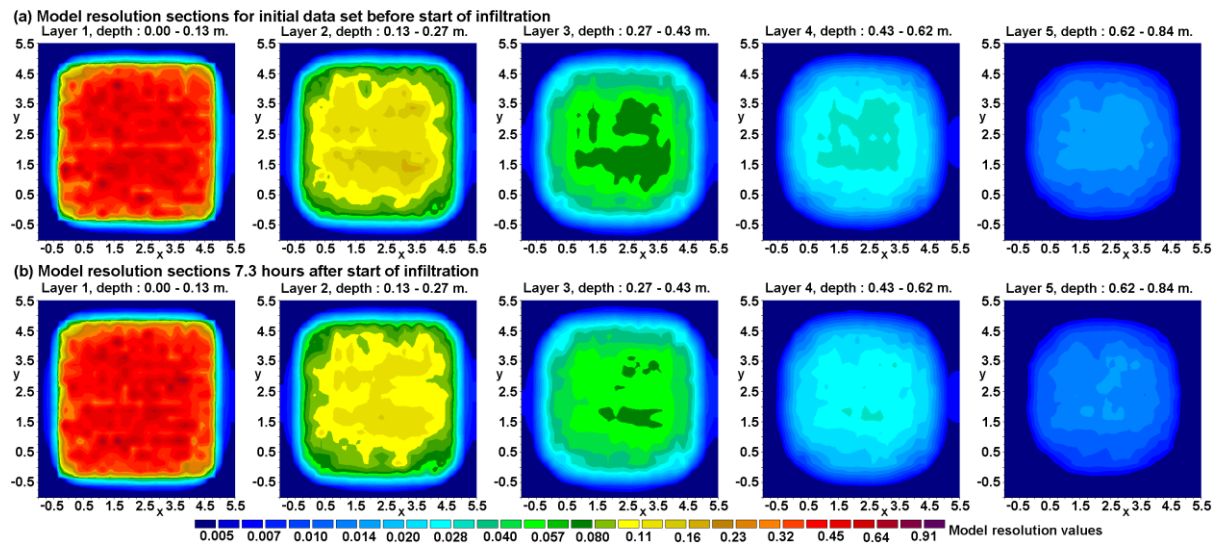


Fig. 8. Model resolution sections for data set measured (a) before start of infiltration, and (b) 7.3 hours after start of infiltration.

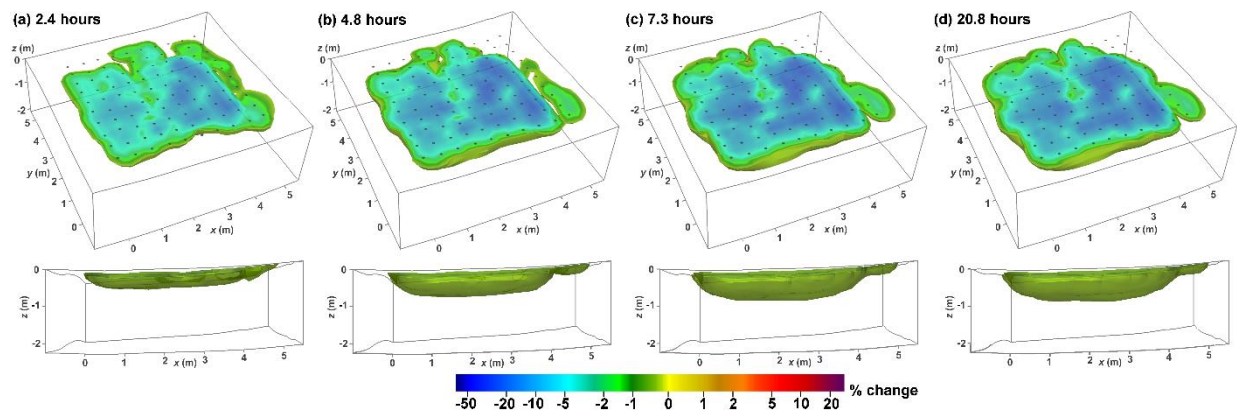


Fig. 9. 3-D plots of the change in the model resistivity at different times. The isosurface of -0.5% is shown. The upper row shows a high angle view while the lower row shows a side view. The positions of the electrodes are shown by small black dots on the surface in the upper row.

4. Discussion

The standard 4-D inverse method can produce artifacts in images of the temporal changes in the resistivity where the direction of the change with time is different from the expected change based on other information. This could be caused by areas with large resistivity contrasts, or areas far from the electrodes with poor resolution. The barrier function method, using either the error or power function formulations, are quite successful in greatly reducing the artifacts. In the synthetic data set, they reduced an artifact with an increase of 29 percent to about 4 to 5 percent. They leave a small residual artifact which can be completely removed by a model refinement step using the method of transformations. We also describe a modified implementation of the L-curve method to determine the optimum damping factor for the model smoothness constraint by using a slow cooling sequence for the damping factor in the iterative inversion algorithm. The optimum damping factor is determined after the inversion by examining the change in the data misfit and model roughness from the sequence of inverse models generated. The advantage of this method is that it takes into account the nonlinear nature of the inversion process and the forward modelling method.

The Hollin Hill field data set illustrates the effect of the background resistivity on the changes detected using a low resistivity tracer solution. The percentage change measured decreases with depth partly because the resistivity of the layers decreases with depth. Together with the decrease in resolution with depth for a survey using electrodes on the surface, it makes the detection of the downwards migration of the tracer more difficult below the top 0.5 m. Possible solutions that require further research include the use of subsurface electrodes (Poje et al., 2018) and optimized arrays (Loke et al., 2014b; Uhlemann et al., 2018).

5. Conclusions

As the standard 4-D inversion method does not include a constraint on the direction of the changes with time, it can result in models with changes opposite to that expected based on other information. In this paper, we incorporate the barrier function method into the 4-D inversion method using a modified temporal difference filter which resulted in a more stable inversion algorithm (compared to a simple diagonal damping matrix). The 4-D barrier function method greatly reduces the temporal artifacts but does not completely remove them. The residual artifacts can be removed using a refinement step with the method of transformations. These additional temporal constraints can be modified so that they are only applied to selected regions in the model (Johnson et al., 2015; Loke et al., 2018b).

Acknowledgements

Wilkinson, Chambers and Uhlemann publish with the permission of the Executive Director, British Geological Survey (UKRI-NERC).

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Auken, E., Pellerin, L., Christensen, N. B., Sørensen, K. I., 2006. A survey of current trends in near-surface electrical and electromagnetic methods. *Geophysics* 71, G249-G260. <https://doi.org/10.1190/1.2335575>
- Box, M.J., 1966. A comparison of several current optimization methods and the use of transformations in constrained problems: *Computer Journal* 9(1), 67-77. <https://doi.org/10.1093/comjnl/9.1.67>
- Cassiani, G., Bruno, V., Villa, A., Fusi, N., Binley, A.M., 2006. A saline trace test monitored via time-lapse surface electrical resistivity tomography: *Journal of Applied Geophysics* 59, 244-259. <https://doi.org/10.1016/j.jappgeo.2005.10.007>
- Clément, R., Descloitres, M., Günther, T., Ribolzi, O., Legchenko, A., 2009. Influence of shallow infiltration on time-lapse ERT: Experience of advanced interpretation. *Comptes Rendus Geoscience* 341, 886–898. <https://doi.org/10.1016/j.crte.2009.07.005>
- Daniels, R.W., 1978. *An introduction to numerical methods and optimization techniques*. Elsevier North-Holland.
- Day-Lewis, F.D., Singha, K., Binley A., 2005. Applying petrophysical models to radar travel time and electrical resistivity tomograms: Resolution-dependent limitations. *Journal of Geophysical Research* 110, B08206. <https://doi.org/10.1029/2004jb003569>
- deGroot-Hedlin, C., Constable, S., 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics* 55, 1613-1624. <https://doi.org/10.1190/1.1442813>
- Farquharson, C.G., Oldenburg, D.W., 1998. Nonlinear inversion using general measures of data misfit and model structure. *Geophysical Journal International* 134, 213-227. <https://doi.org/10.1046/j.1365-246x.1998.00555.x>
- Farquharson, C.G., Oldenburg, D.W., 2004. A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems. *Geophysical Journal International* 156, 411-425. <https://doi.org/10.1111/j.1365-246x.2004.02190.x>
- Gill, P.E., Murray, W., 1974. *Numerical methods for constrained optimization*. Academic Press.
- Gunn, D. A., Chambers, J. E., Hobbs, P. R. N., Ford, J. R., Wilkinson, P. B., Jenkins, G. O., Merritt, A., 2013. Rapid observations to guide the design of systems for long-term monitoring of a complex landslide in the Upper Lias clays of North Yorkshire, U.K. *Q. J. Eng. Geol. Hydrogeol.* 46, 323–336. <https://doi.org/10.1144/qjegh2011-028>

- Gunther, T., Rucker C., Spitzer K., 2006. Three-dimensional modelling and inversion of dc resistivity data incorporating topography – II. Inversion. *Geophysical Journal International* 166, 506-517. <https://doi.org/10.1111/j.1365-246x.2006.03011.x>
- Haber, E., Oldenburg, D.W., Shekhtman, R., 2007. Inversion of time domain three-dimensional electromagnetic data. *Geophysical Journal International* 171, 550–564. <https://doi.org/10.1111/j.1365-246x.2007.03365.x>
- Hojat, A., Arosio, D., Ivanov, V.I., Longoni, L., Papini, M., Scaioni, M., Tresoldi, G., Zanzi, L., 2019. Geoelectrical characterization and monitoring of slopes on a rainfall triggered landslide simulator. *Journal of Applied Geophysics* 170, 103844. <https://doi.org/10.1016/j.jappgeo.2019.103844>
- Johansen, H.K., 1977. A man/computer interpretation system for resistivity soundings over horizontally stratified earth. *Geophysical Prospecting* 25, 667–691. <https://doi.org/10.1111/j.1365-2478.1977.tb01196.x>
- Johnson, T., Versteeg, R., Thomle, J., Hammond, G., Chen, X., Zachara, J., 2015. Four-dimensional electrical conductivity monitoring of stage-driven river water intrusion: Accounting for water table effects using a transient mesh boundary and conditional inversion constraints. *Water Resources Research* 51(8), 6177-6196. <https://doi.org/10.1002/2014wr016129>
- Karaoulis, M., Kim J.H. and Tsourlos P.I., 2011. 4D Active Time Constrained Inversion. *Journal of Applied Geophysics*, 73, 25–34.
- Karaoulis, M, Tsourlos, P., Kim, J.H. and Revil, A., 2014. 4D time-lapse ERT inversion: introducing combined time and space constraints. *Near Surface Geophysics*, 12, 25-34.
- Kim, J. H., Yi, M J., Park, S G., Kim, J.G., 2009. 4-D inversion of DC resistivity monitoring data acquired over a dynamically changing earth model. *Journal of Applied Geophysics* 68, 522-532. <https://doi.org/10.1016/j.jappgeo.2009.03.002>
- Kim, J.H., Yi, M.J., Ahn, H.Y., Kim, K.S., 2010. 4-D inversion of resistivity monitoring data using L1 norm minimization. *Near Surface 2010 - 16th EAGE European Meeting of Environmental and Engineering Geophysics, Zurich, Switzerland, Expanded Abstracts*, A15. <https://doi.org/10.3997/2214-4609.20144780>
- LaBrecque, D.J., Miletto, M., Daily, W., Ramirex, A., Owen, E., 1996. The effects of noise on Occam's inversion of resistivity tomography data. *Geophysics* 61, 538–548. <https://doi.org/10.1190/1.1443980>

- Li Y., Oldenburg, D.W., 2003. Fast Inversion of large-scale magnetic data using wavelet transforms and logarithmic barrier method. *Geophysical Journal International* 152, 251 - 265. <https://doi.org/10.1046/j.1365-246x.2003.01766.x>
- Loke, M.H., Barker, R.D., 1996. Practical techniques for 3D resistivity surveys and data inversion. *Geophysical Prospecting* 44, 499-523. <https://doi.org/10.1111/j.1365-2478.1996.tb00162.x>
- Loke M.H., Acworth, I., Dahlin, T., 2003. A comparison of smooth and blocky inversion methods in 2D electrical imaging surveys. *Exploration Geophysics* 34, 182-187. <https://doi.org/10.1071/eg03182>
- Loke, M.H., Chambers, J.E., Rucker, D.F., Kuras, O., Wilkinson, P.B., 2013. Recent developments in the direct-current geoelectrical imaging method. *Journal of Applied Geophysics* 95, 135-156. <https://doi.org/10.1016/j.jappgeo.2013.02.017>
- Loke, M.H., Dahlin, T., Rucker, D.F., 2014a. Smoothness-constrained time-lapse inversion of data from 3-D resistivity surveys. *Near Surface Geophysics* 12, 5-24. <https://doi.org/10.3997/1873-0604.2013025>
- Loke, M.H., Wilkinson, P.B., Uhlemann, S.S., Chambers, J.E., Oxby, L.S., 2014b. Computation of optimized arrays for 3-D electrical imaging surveys. *Geophysical Journal International* 199, 1751-1764. <https://doi.org/10.1093/gji/ggu357>
- Loke, M.H., Wilkinson, P.B., Chambers, J. E., Meldrum, P.I., 2018a. Rapid inversion of data from 2-D resistivity surveys with electrodes displacements. *Geophysical Prospecting* 66, 579-594. <https://doi.org/10.1111/1365-2478.12522>
- Loke, M.H., Wilkinson, P.B., Dahlin, T., Chambers, J.E., Uhlemann, S., Dijkstra, T., 2018b. Time-Lapse 4-D Resistivity Imaging Inversion with Positivity Constraints. 24th European Meeting of Environmental and Engineering Geophysics, 9-12 September 2018, Porto, Portugal, We 24A 02. <https://doi.org/10.3997/2214-4609.201802625>
- Loke, M.H., Papadopoulos, N., Wilkinson, P.B., Oikonomou, D., Simyrdanis, K., Rucker, D.F., 2020. The inversion of data from very large 3-D ERT mobile surveys. *Geophysical Prospecting*, 68, 2579-2597.
- Loke, M.H., 2021. Tutorial : 2-D and 3-D electrical imaging surveys. Geotomo Software Sdn Bhd, Malaysia. www.geotomosoft.com.
- Maurer, H., Friedel, S., 2006. Outer-space sensitivities in geoelectrical tomography. *Geophysics* 71, 1942-2156. <https://doi.org/10.1190/1.2194891>
- Oldenburg, D.W, Li, Y. 1994. Inversion of induced polarization data. *Geophysics* 59, 1327-1341. <https://doi.org/10.1190/1.1443692>

- Poje, M., Salgado, L., Rucker, D.F., Levitt, M., Strohmeyer, J., McNeill, M., Rucker, K., Berumen, R., Galbraith, H., 2018. Optimizing Gold Leaching with the Aid of Electrical Resistivity Geophysics. *Fast Times* 23(2), 78-84.
- Press, W.H, Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. *Numerical Recipes: The Art of Scientific Computing* (3rd Edition). Cambridge University Press.
- Putman, B.D., 2014. Aspects of time-lapse electrical resistivity monitoring in geotechnical and reservoir problems. Unpubl. PhD thesis, Colorado School of Mines, U.S.A.
- Rucker, D.F, Crook, N., Winterton, J., McNeill, M., Baldyga, C.A., Noonan, G., Fink, J.B., 2014. Real-time electrical monitoring of reagent delivery during a subsurface amendment experiment. *Near Surface Geophysics* 12, 151-163. <https://doi.org/10.3997/1873-0604.2013017>
- Singha, K., Day-Lewis, F.D., Johnson, T., Slater, L.D., 2014. Advances in interpretation of subsurface processes with time-lapse electrical imaging. *Hydrological Processes* 29(6), 1549-1576. <https://doi.org/10.1002/hyp.10280>
- Uhlemann, S., Chambers, J.C., Wilkinson, P.B., Maurer, H., Merritt, A., Meldrum, P.I., Kuras, O., Gunn, D., Smith, A., Dijkstra, T., 2017. 4D imaging of moisture dynamics during landslide reactivation. *Journal of Geophysical Research: Earth Surface* 122(1), 398-418. <https://doi.org/10.1002/2016jf003983>
- Uhlemann, S., Wilkinson, P.B., Maurer, H., Wagner, F.M., Johnson, T.C., Chambers, J.E., 2018. Optimized survey design for electrical resistivity tomography: combined optimization of measurement configuration and electrode placement. *Geophysical Journal International* 214, 108–121. <https://doi.org/10.1093/gji/ggy128>
- Vasco, D.W., 1998. Regularization and trade-off associated with nonlinear geophysical inverse problems: penalty homotopies. *Inverse Problems* 4, 1033-1052.
- Zhou, B., Dahlin, T., 2003. Properties and effects of measurement errors on 2D resistivity imaging surveying. *Near Surface Geophysics* 1, 105-117. <https://doi.org/10.3997/1873-0604.2003001>

Appendix. A post-inversion L-curve method to determine the optimum damping factor

The barrier function method requires the inversion of the first time step data set that is later used as a reference model in the 4-D time-lapse inversion. For the synthetic data set, the noise level is known. Thus, we choose the inverse model at the iteration where the data misfit is close to the noise level. In some field surveys, reciprocal measurements (LaBrecque et al., 1996) are made to estimate the noise level. However, in many field surveys such reciprocal measurements are not made. Furthermore, the reciprocal measurements do not take into account errors from other sources such as errors in the electrode positions (Zhou and Dahlin, 2003). A typical least-squares equation used for the inversion of a single data set is as follows.

$$\left[\mathbf{J}_i^T \mathbf{R}_d \mathbf{J}_i + \lambda_i \mathbf{W}^T \mathbf{R}_m \mathbf{W} \right] \Delta \mathbf{m}_i = \mathbf{J}_i^T \mathbf{R}_d \mathbf{g}_i - \lambda_i \mathbf{W}^T \mathbf{R}_m \mathbf{W} (\mathbf{m}_{i-1} - \mathbf{m}_r). \quad (\text{A.1})$$

The least-squares equation is a linear approximation to a non-linear optimization problem that leaves out the second and higher order terms of the partial derivatives of the calculated model response with respect to the model parameters (Gill and Murray, 1974; Daniels, 1978). It contains the damping factor (λ) that provides the relative weight given to reducing the model roughness compared to the data misfit. Farquharson and Oldenburg (2004) describes two methods to estimate the optimum damping factor value, the Generalized Cross Validation (GCV) and L-curve methods. The GCV method requires a matrix inversion where the time taken is proportional to the cube of the number of model parameters. It is impractical to use in cases where the number of model parameters might be several hundred thousand (Loke et al., 2020). The standard L-curve method (Farquharson and Oldenburg, 2004; Loke et al., 2014a) attempts to find the optimum damping factor for each iteration by comparing the data misfit versus the model roughness for a range of damping factor values that includes the optimum value. For example, 25 damping factor values were used by Loke et al. (2014a) ranging from 0.01 to 100 times the trial value. Equation (A.1) is solved for each damping factor value to determine the change in the model parameter $\Delta \mathbf{m}_{iL}$. The calculated apparent resistivity values (\mathbf{d}_{iL}) that would be obtained from the new model, $\mathbf{m}_i + \Delta \mathbf{m}_{iL}$, is estimated using the following linear approximation.

$$\mathbf{d}_{iL} \approx \mathbf{d}_i + \mathbf{J}_i^T \Delta \mathbf{m}_{iL} \quad (\text{A.2})$$

The corresponding approximate data misfit and model roughness are then calculated for each trial model. The optimum damping factor is then estimated from the curve with the data misfit and model roughness values. However, as noted by Gunther et al. (2006), the L-curve algorithm does not always provide a satisfactory solution. Very often, a simple model such as a homogeneous half-space, is used as a starting model (Loke and Barker, 1996). The data misfit

vector \mathbf{g}_i in the first few iterations usually have high amplitudes. If the damping factor estimated by the L-curve algorithm is too low, this could produce a model with resistivity variations that are too large as equation (A.1) is a linear approximation to a non-linear inverse problem. Furthermore, equation (A.2) is a linear approximation to a non-linear forward modelling routine. Haber et al. (2009) used a cooling sequence that starts with a large damping factor that is slowly reduced after each iteration. This requires more iterations, and thus a longer calculation time, but leads to a more stable solution. The output from the slow cooling sequence inversion is a series of models each with its own data misfit and model roughness. We use the L-curve method as a post-inversion technique to select the optimum model. This avoids the linear approximations used in the normal L-curve method. We use the following equations to calculate the data misfit and model roughness if the L1-norm method (Farquharson and Oldenburg, 2004) is used.

$$D(i) = \frac{1}{m} \sum_{j=1}^m |d_{oi}(j) - d_{ci}(j)| \quad (\text{A.3})$$

d_{oi} and d_{ci} are the logarithms of measured and calculated apparent resistivity values, and m is the number of data points. The model roughness is calculated using the following equation.

$$M(i) = \frac{1}{n} |\mathbf{W}\mathbf{m}_i| \quad (\text{A.4})$$

Figure A.1a shows a plot of the model roughness versus the data misfit (together with the damping factors) for the Hollin Hill data set. The data misfit decreases from about 4.6% to 1.3% in the sequence of models generated. The model roughness gradually increases as the data misfit (and damping factor) decreases. Figure A.1b shows a plot of the curvature of the L-curve which has a maximum for a damping factor of about 0.067. This provides an estimate of the optimum damping factor to use for the data set.

The proposed post-inversion L-curve analysis can be used for any data set where the inversion can be carried out. The advantage is that it takes into account the non-linear nature of the inverse problem and the forward modelling routine, and gives a stable solution by using a slow cooling sequence. The main disadvantage is that it might require more iterations (and thus more computer time) in order to generate sufficient points for the L-curve analysis. Examples of L-curves for different data sets and some guidelines in setting the range of damping factor values are given in Loke (2021). Research is still being conducted with this variation of the L-curve method with different data sets, and refinements (such as a more robust method to determine the maximum of the curvature curve) will probably be made in future. It should be considered an additional tool to current L-curve analysis methods (Vasco 1988;

Farquharson and Oldenburg 2004), particularly for large 3-D data sets that have already been processed and an output file with the inverse models and data misfits information is available.

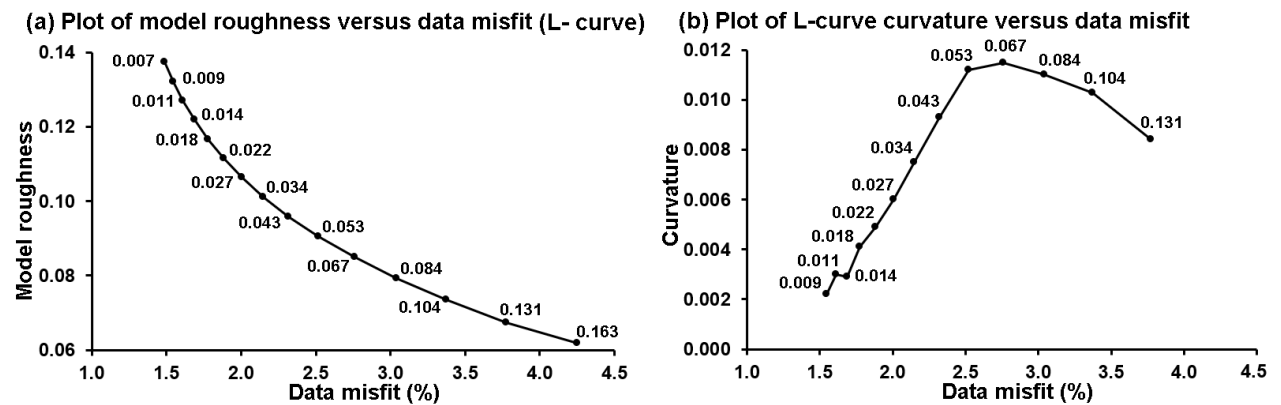


Figure A.1. (a) Plot of model roughness versus data misfit (L-curve) for the Hollin Hill data set. (b) Plot of L-curve curvature versus data misfit. The damping factor values are given next to the points on the curves.