Title : Optimized arrays for 2-D cross-borehole electrical tomography surveys

Authors : M. H. Loke[1*] (Geotomo Software), P. B. Wilkinson[2], J.E. Chambers[2] and M. Strutt[2] (British Geological Survey)

[1] Geotomo Software, 115 Cangkat Minden Jalan 5, Minden Heights, Gelugor 11700, Penang, Malaysia

[2] British Geological Survey, Kingsley Dunham Centre, Keyworth, Nottingham, United Kingdom NG12 5GG

[*] Email : drmhloke@yahoo.com

Abstract

The use of optimized arrays generated using the ‘Compare R’ method for cross-borehole resistivity measurements is examined in this paper. We compare the performances of two array optimization algorithms, one that maximizes the model resolution and another that minimizes the point spread value. Although both algorithms give similar results, the model resolution maximization algorithm is several times faster. A study of the point spread function plots for a cross-borehole survey shows that the model resolution within the central zone surrounded by the borehole electrodes is much higher than near the bottom end of the boreholes. Tests with synthetic and experimental data show that the optimized arrays generated by the ‘Compare R’ method have significantly better resolution that a ‘standard’ measurement sequence used in previous surveys. The resolution of the optimized arrays is less if arrays with both current (or both potential) electrodes in the same borehole are excluded. However, they are still better than the ‘standard’ arrays.

Keywords : Acquisition, Borehole geophysics, Inversion, Resistivity, Tomography

INTRODUCTION

Within the past two decades there have been many significant developments in the resistivity exploration method and it is now one of the standard techniques used in engineering, environmental and mining surveys. The development of multi-electrode resistivity meter systems (Griffiths *et al.* 1990) has led to the widespread use of two-dimensional (2-D) and even three-dimensional (3-D) resistivity surveys in areas of complex geology (Dahlin 1996; Auken *et al.* 2006; Chambers *et al.* 2006). The field applications range from agriculture, groundwater exploration, engineering site investigation, environmental assessment, and mineral exploration to hydrocarbon mapping (Loke *et al.* 2011). The resolution of surveys carried out with electrodes on the ground surface decreases rapidly with depth. Cross-borehole surveys have been carried out to improve the resolution across the entire depth range of interest (Slater *et al.* 2000; Wilkinson *et al.* 2006a; Chambers *et al.* 2010).

Automatic multi-electrode survey instruments have made it possible for the user to select the optimum arrays for a survey. There have been many significant developments in algorithms to automatically select arrays to maximize the resolution of the inversion model for linear surface arrays and cross-borehole surveys (Stummer *et al.* 2004; Wilkinson *et al.* 2006a, 2006b; Loke *et al.* 2010a; Nenna *et al.* 2011; Hagrey 2012). A non-linear method that calculates the model resolution (the 'Compare R' method) by Wilkinson *et al.* (2006b) proved to be the best method (Loke *et al.* 2010a). The 'Compare R' method uses only the diagonal elements of the model resolution matrix but does not take into account the off-diagonal elements. A possibly better measure of the resolution capability of the data is the point spread function (Friedel 2003; Miller and

Routh 2007; Oldenborger and Routh 2009) that uses the entire resolution matrix. The use of the point-spread-function for selecting field survey configurations was studied by Routh *et al.* (2005).

In this paper, we compare the 'Compare R' (CR) method with a modified algorithm that minimizes the spread value. We also compare optimized arrays generated using the CR method with a standard measurement sequence previously used by Wilkinson *et al.* (2006a). Point spread function plots for selected model cells for a cross-borehole survey setup are first examined to help understand the resolving power of the measurements. Inversions of data sets from two synthetic models are then made to illustrate the resolving capability of the different data sets. Finally tests from a laboratory tank experiment using the standard measurement sequence and optimized array data sets are presented.

THEORY

Data inversion and model resolution

The smoothness-constrained least-squares optimization method is frequently used for 2-D inversion of resistivity data (Loke *et al.* 2003). The subsurface model usually consists of a large number of rectangular cells. The linearized least-squares equation that gives the relationship between the model parameters and the measured data is given below.

$$\left(\mathbf{J}^{\mathrm{T}}\mathbf{J}+\lambda\mathbf{C}\right)\Delta\mathbf{r}_{\mathbf{i}}=\mathbf{J}^{\mathrm{T}}\mathbf{d}-\lambda\mathbf{C}\mathbf{r}_{\mathbf{i-1}}, \qquad (1)$$

where $\Delta\mathbf{r}_{\mathbf{i}}=\mathbf{r}_{\mathbf{i}}-\mathbf{r}_{\mathbf{i-1}}$.

The Jacobian matrix J contains the sensitivities of the (logarithm of the) apparent resistivity values with respect to the (logarithm of the) model resistivity values (Loke and Barker 1995). The **C** matrix contains the roughness filter constraint, λ is the damping factor and **d** is the data misfit vector. $\mathbf{r_{i-1}}$ is the model parameter vector (the logarithm of the model resistivity values) for the previous iteration, while $\mathbf{\Delta r_i}$ is the change in the model parameters. Various modifications have been made to the above equation to incorporate desired characteristics in the data misfit or model roughness, such as using a L1-norm criterion for the data misfit and model roughness filter (Loke *et al.* 2003) and to include known data errors using a data weighting matrix (Ellis and Oldenburg 1994).

It can be shown that the model resolution matrix **R** (Menke 1989; Loke *et al.* 2010a) is given by

$$\mathbf{R} = \mathbf{B}\ \mathbf{A} \quad \text{, where} \quad \mathbf{A} = \mathbf{J}^{\mathbf{T}}\mathbf{J} \text{ and } \mathbf{B} = \left(\mathbf{J}^{\mathbf{T}}\mathbf{J} + \lambda\mathbf{C}\right)^{-1} . \qquad (2)$$

The main diagonal elements of **R** give an estimate of the model cells resolution.

The 'Compare R' method

The ‘Compare R’ method (Wilkinson *et al.* 2006b) attempts to determine the set of array configurations that will maximize the average resolution value for a homogeneous earth model. For a system with *N* electrodes, there are $N(N-1)(N-2)(N-3)/8$ independent four-electrode configurations. To reduce the number of possible arrays, arrays where the current and potential electrodes are interleaved (Carpenter and Habberjam 1956) of the Wenner-γ type configuration as well as those large geometric factors that exceed a set limit are also excluded (Stummer *et al.* 2004). The remaining configurations form the ‘comprehensive’ data set. A small base data set consisting of the dipole-dipole

configurations with the '*a*' dipole length of 1 unit electrode spacing is used as the starting base data set. The change in the model resolution matrix **R** for each new array when added to the base set is then calculated. A selected number of the configurations that result in the largest increase in the model resolution, and have a suitable degree of orthogonality to the existing configurations, are then added to the base data set (Wilkinson *et al.* 2012). This is repeated until the desired number of optimized array configurations is selected.

The Sherman-Morrison Rank-1 update (Golub and van Loan 1989) is used to calculate the change in the model resolution matrix when a new test configuration is added to the base set. The following set of updating formulae (Loke *et al.* 2010b) is used to calculate the new resolution matrix $\mathbf{R_{b+1}}$ when a new array is added to the base set

$$\mathbf{R_{b+1}} = \mathbf{R_b} + \Delta\mathbf{R_b}, \tag{3}$$

where $\Delta\mathbf{R_b} = \frac{\mathbf{z}}{1+\mu}\left(\mathbf{g}^{\mathbf{T}} - \mathbf{y}^{\mathbf{T}}\right)$, $\mathbf{z} = \mathbf{B_b g}$, $\mathbf{y} = \mathbf{A_b z}$ and $\mu = \mathbf{g.z}$

The vector **g** (that has *m* elements, where *m* is the number of model cells) contains the sensitivity values of the model cells for the new test configuration. The following function F_{CR} (Wilkinson *et al.* 2012) that uses the ratio of the change in the model resolution to the comprehensive data set resolution is used to the rank the improvement in the model resolution ~~with *m* model cells~~ due to an add-on array.

$$F_{CR} = \frac{1}{m}\sum_{j=1}^{j=m}\frac{\Delta R_b(j,j)}{R_c(j,j)}. \tag{4}$$

The 'Compare R' method selects the arrays that have the largest F_{CR} values. The average relative model resolution that is given by

$$S_r = \frac{1}{m}\sum_{j=1}^{j=m}\frac{R_b(j,j)}{R_c(j,j)}. \quad (5)$$

This is used to assess the performance of the array optimization methods. R_b and R_c are the base and comprehensive data set model resolutions.

A method using the spread value

Routh *et al.* (2005) proposed a method that uses the point spread function for selecting optimal arrangements of sources and receivers in EM surveys. The point spread function (PSF) for a model cell consists of the elements of corresponding column of the resolution matrix. The distribution of the resolution values in the point spread function show the degree of resolution. A PSF that is wide and/or has large side lobes show that the model cell is poorly resolved. A spread criterion value is frequently used to reduce the information in the PSF to a single number. In this paper, we use a discretized form of the spread criterion proposed by Miller and Routh (2007). The spread criterion value for the *i*th model cell, *S(i)*, is given by the following equation.

$$S(i) = \frac{\sum_{j=1}^{j=m} W(i,j)\left[R(i,j) - \Delta_{ij}\right]^2 \delta_j}{\alpha + \sum_{j=1}^{j=m} R(i,j)^2 \delta_j} \quad (6)$$

where $W(i,j) = 1 + d_{ij}$,

$\Delta_{ij} = 1$ for $i=j$, and $\Delta_{ij} = 0$ for $i \neq j$.

d_{ij} is the normalized distance (distance divided by the unit electrode spacing) between the centres of *i*th and *j*th model cells, α is a small value (0.0001) and δ_j is the normalized area of the *j*th cell.

In this paper, we use the methodology previously described for the ‘Compare R’ method to select electrode arrays that minimize the spread value. We use the following ranking function for the spread criterion to evaluate the effect of adding an array to the base set on the spread value.

$$F_{PSF} = \frac{1}{m}\sum_{j=1}^{j=m}\left[S_b(j) - S_{b+1}(j)\right] \tag{7}$$

This ranking function uses the average reduction in the spread criterion after adding the test configuration to the base data set. The optimization algorithm seeks the new array configurations that cause a maximum reduction in the average spread value. We denote this method as the 'Compare S’ (CS) method in the following discussion. We also calculate the average spread criterion value

$$S_v = \frac{1}{m}\sum_{j=1}^{j=m} S_b(j) \tag{8}$$

to display the change in the spread value with the number of arrays in the base data set. Tests were also conducted with other ranking functions but the results were not significantly different from that obtained using equation (7) and will not be listed in this paper.

RESULTS AND DISCUSSION

Region of model coverage

In the model used to generate the optimized data sets for surface arrays, only the region between the first and last electrode of the survey line was considered by Wilkinson *et al.* (2006b) and Loke *et al.* (2010a). In this section we examine of influence of the regions outside the grid of electrodes used for surface and cross-borehole surveys. This will be used to determine the model space used for calculating the optimized arrays.

The model resolution for the comprehensive data set with 61820 data points is calculated for a linear surface survey line with 31 electrodes with a spacing of 1 meter. The maximum geometric factor is set at 2262 m (that corresponds to a dipole-dipole array with a=1 m and n=8). The model resolution values are calculated for regions up to 15 meters beyond the first and last electrodes. The width and thickness of all the model blocks were set at 1 meter so as to avoid effects due to non-uniform block sizes. A value of 0.001 (Wilkinson *et al.* 2012) was used for the damping factor λ in equation (1). A resolution value of 0.05 is typically used as the cut-off value to mark regions that have significant resolution. The model resolution section for the surface survey line shows resolution values of about 0.05 up to about 3.5 meters beyond the ends of the line (Figure 1a).

A similar calculation is made for a cross-borehole survey with 11 electrodes along the surface and 20 electrodes along two boreholes that are 10 meters apart. The cross-borehole comprehensive data set has 466934 data points. The larger size of the cross-borehole comprehensive data set is due to the larger number of electrodes (51 compared

to 31 for the surface survey line). For the cross-borehole data set the region with resolution values above 0.05 near the surface extends to about 7 meters beyond the first and last surface electrodes. Thus cross-borehole measurements are more greatly influenced by structures outside the line compared to surface only measurements. The model resolution values between the two boreholes are lower towards the bottom of the boreholes. The region with significant resolution values extends to about 3 meters below the deepest borehole electrode. In the array optimization calculations, we use models with cells up to 12 meters beyond the left and right electrodes on the surface and 6 meters below the bottommost borehole electrode. This ensures that all the regions that have a significant influence on the results are accounted for.

Array optimization using the model resolution and spread function

The results from the array optimization calculations using the 'Compare R' (CR) method and the 'Compare S' (CS) methods are presented. For the 'Compare R' method, we show the results where the number of data points in the base set is increased by 5% or by 10% after each iteration. The results using the 'single step' method (Loke *et al.* 2010b) are also shown for the CR method as it gives the maximum resolution possible with the array optimization algorithm used. The calculations were carried out on a computer with a 2.8 GHz Intel quad-core i7 CPU and an Nvidia 460GTX GPU (Farber 2011). The single-precision version of the GPU based CR method (Loke *et al.* 2010b) was used to reduce the calculation time.

Figure 2a shows the change of the average relative model resolution (equation (5)) for the different methods for up to 20000 data points. The CR method with a 5% step

size (red line) achieves higher resolution values than the CS method (black line). This is not surprising as the mandate of the CR method is to select arrays with the highest resolution values. Figure 2b shows the change of the average spread criterion value S_v (equation (8)) with the number of arrays for the different methods. While the CS method achieves lower spread values than the CR method over most of the range tested (particularly for more than 2000 data points), the difference is less than 2% and probably not significant. The spread value is a weighted average of the resolution values over a large number of cells. Thus it is not as sensitive to the addition of a new array compared to the model resolution value of a single cell. Different objective functions using the spread function were tested (such as using the ratio of the base set to the comprehensive set) but the results were similar and are not presented here. The CR method only requires the calculation of the diagonal of the model resolution matrix while the CS method requires the entire resolution matrix (actually about half because of symmetry). Thus as shown in Table 1 the CS method takes a significantly longer time compared to the CR method. Therefore the CR method alone is used for the tests in the following sections.

The CR method with a 5% step size and using the new selection method proposed by Wilkinson *et al.* (2012) achieves resolution values that are very close to the 'single step' method while the calculation time is much less. The CR method with 10% step size achieves arrays that are almost as good for over 12000 data points (Figure 2), and thus is an attractive option for large data sets due to the lower computer time required.

Plots of the point spread function for cross-borehole data sets

It is generally expected that the model cells near the central region between the boreholes are better resolved compared to areas near the bottom. To illustrate this, plots of the point spread function for two selected cells are shown. Figure 3a shows the point spread function plot for a model cell near the center of the cross-borehole region using the comprehensive set with 466934 data points. Since this data set contains all the viable measurements, it displays the maximum resolution possible for the cross-borehole setup. For this cell, the maximum resolution value (0.337) occurs at the cell location. Figure 3b shows a similar plot for a cell near the bottom edge of the boreholes. In this case, the maximum resolution value (0.156) is much lower. Furthermore the maximum value occurs about 1 meter to the right of the actual cell location. This shows that even if all the possible measurements are made, it is still not possible to accurately determine the horizontal position of an anomaly in this region. Figure 4 shows similar plots using the optimized data set with 1875 data points generated using the Compare R method with a 5% step size. The maximum resolution values for the two cells are 0.199 and 0.128 respectively. Although the number of data points is only about 0.4% of the comprehensive set, the maximum resolution values achieved at the two cells are about 59% and 82% respectively of the comprehensive set. Note for the optimized data set, the relative maximum resolution value for the cell near the bottom (82%) is closer to that achieved by the comprehensive data set compared to the cell near the centre (59%). This is because the array optimization program attempts to achieve a similar relative model resolution throughout the entire model region.

Figure 5 shows similar point spread function plots for a ‘standard’ cross-borehole measurement sequence with 1875 data points that was used by Wilkinson *et. al.* (2006a). The resolution value for the cell near the centre of the cross-borehole region (Figure 5a) is much lower compared to the optimized data set (Figure 4a). It reaches a maximum value of 0.106 compared to 0.199 for the optimized data set. Note that unlike the optimized data set, the maximum value is located 1 meter to the right of the cell location. The performance of the ‘standard’ measurement sequence for the model cell near the bottom is even poorer. The maximum resolution value achieved is 0.084 (Figure 5b) compared to 0.128 for the optimized data set (Figure 4b).

It was shown in Figure 1 that the regions with significant resolution extend well outside the area between the boreholes. To assess the possible resolution of structures located outside the borehole region, Figure 6 shows plots of the point spread function for a cell to the right of the borehole region for the optimized and 'standard' data sets. The cell has a width of 1 meter with the center at 3.5 meters from the right column of borehole electrodes. The optimized data set has a maximum resolution value of 0.125 while the 'standard' data set has a much lower value of 0.089. Furthermore, the maximum resolution value for the 'standard' data set is located well to the left of the cell location (Figure 6b). Thus it is expected that the 'standard' data set will be significantly poorer at resolving structures located outside the borehole region.

Figure 7 shows plots of the ratios of the model resolution values of the optimized and 'standard' data sets, $R_b(j,j)/R_c(j,j)$, to those of the comprehensive data set. Both measurement sequences have high relative resolution values near the electrodes, and lower values with increasing distance from the boreholes. The optimized data set has a

much higher average relative resolution value (0.603) compared to the 'standard' data set (0.297). There is a slight asymmetry in the section for the 'standard' data set as the array configurations do not have a symmetrical distribution, while the algorithm used for generating the optimized data set ensures a symmetrical distribution (Loke *et al.* 2010a). The plots give an overall picture of the model resolution achieved by the two measurement sequences compared to the comprehensive data set. However they do not show the finer details, such as the poorer horizontal resolution compared to the vertical resolution near the bottom of the boreholes (Figures 4b and 5b), that are revealed by the point spread function plots. Similar plots were made for the difference in the spread values of the two measurement sequences and the comprehensive data sets, $S_b(j)$-$S_c(j)$. As the plots are essentially the inverse of the relative model resolution plots (with small values near the electrodes and larger values with increasing distance from the borehole region) they are not included in this paper.

Tests with synthetic models

In this section, we present results of tests with two synthetic models. The first test model consist of four high resistivity blocks of different sizes with resistivity of 1500 ohm-m embedded within a homogeneous medium of 15 ohm-m. The outlines of the blocks are shown in Figure 8. The three topmost blocks are located well within the region covered by the electrodes on the surface and in the boreholes. The deepest block is located near the bottom end of the boreholes. The first test is with the optimized data set (with a 5% step size) with 1875 data points. The calculated resistance values using the finite element method for the arrays configurations in this data set range from 1.32 to 2396.62 ohm.

Gaussian random noise (Press *et al.* 2007) with a maximum amplitude of 0.25 ohm was added to the resistance values before they were converted to apparent resistivity values. The average noise level in the final apparent resistivity values was 1.6%. An inversion of the apparent resistivity data set with the added noise was then carried out using the smoothness-constrained least-squares method. The L1-norm method was used for both the data misfit and model roughness filters (Loke *et al.* 2003). The roughness filter with diagonal components was used as the high resistivity blocks also have sloping sides (Farquharson 2008). A cooling sequence method was used to set the damping factor in equation (1). A large value of about 0.10 is initially used which is then reduced by half after each iteration until it reached a preset minimum value (about 0.001). Since the average noise level of the data set is known, we used the 'discrepancy principle' method (Farquharson and Oldenburg 2004) where the inversion model with a data misfit similar to the known noise level is selected. The selected inversion model is shown in Figure 8a. The top three high resistivity blocks are well resolved although the maximum resistivity values achieved (Table 2) are well below the true value of 1500 ohm-m. This is expected as the smoothness constraint used to stabilize the inversion method in the presence of noise reduces the model resolution. The highest resistivity value of 335.2 ohm-m is achieved at the topmost block that is located near the electrodes at the surface and the right borehole. Not surprisingly, a much lower value of 114.2 ohm-m is obtained in the second deepest block that has a smaller size. A slightly higher value of 152.7 ohm-m is obtained at the third deepest block which is the largest. It is located near the central region between the borehole electrodes but well away from the surface electrodes. The deepest block located near the bottom end of the boreholes is barely resolved where the

maximum value obtained is only 21.9 ohm-m (Table 2). Note the vertical extent of the block is much better resolved compared to the horizontal extent. This agrees with the point spread function plot in Figure 4b that shows the vertical location of an anomaly near the ends of the borehole will be more accurately determined compared to the horizontal location.

Figure 8b shows the inversion model (with a data misfit of 0.1%) obtained using the 'standard' array configurations. After adding the random noise of 0.25 ohm to the resistance values, the average noise level in the apparent resistivity values is only 0.1% that is much lower than the optimized data set. This is because the average geometric factor for the arrays used is much lower than that used in the optimized data set. Although the data misfit in the inversion model is lower, the deepest block is not resolved. The highest resistivity values at the three upper blocks are also significantly lower than the corresponding model for the optimized data set (Table 2). This agrees with the lower resolution values displayed by the point-spread-function plot for this data set (Figure 5).

In cases where the resistivity of the fluid filling the boreholes is much lower than the surrounding subsurface materials, current channelling that occurs when two current electrodes are located within the same borehole can be a significant problem (Wilkinson et al. 2006a). Another set of optimized arrays was generated that excludes configurations with both current (or both potential) electrodes in the same borehole. This is referred to as the 'reduced' optimized data set in this paper. Figure 6c shows the inversion model with a data misfit of 0.6% (that is the same as the added noise level). It fails to detect the deepest block but the maximum resistivity values achieved at the top three blocks are

significantly higher than for the ‘standard’ data set model (Table 2). The maximum values at the first and third blocks are even higher than the model for the full optimized data set. This is probably because the data set is less noisy.

The second test model has a more complex structure with both conductive and resistive blocks in a two-layer medium (Figure 9a). A block is also placed beyond the right column of borehole electrodes to assess the resolution of the different arrays for structures outside the borehole region. Gaussian random noise with a maximum amplitude of 1.0 ohm was added to the resistance values before they were converted to apparent resistivity values. This resulted in average noise levels of 2.6%, 0.1% and 1.0% in the final apparent resistivity values for the full optimized, 'standard' and reduced optimized data sets. The inversion models for the three data sets are shown in Figure 9. The boundary between the two layers is well resolved by all the three inversion models. The background resistivity values, away from the embedded rectangular blocks, in the models are also generally within a few percent of the true values of 100 and 30 ohm-m in the two layers. The topmost high resistivity block is well resolved by all three data sets. The models with the full and reduced optimized data sets give maximum values of 888 and 980 ohm-m respectively (compared to the true value of 1000 ohm-m), while the 'standard' data set achieved a maximum value of only 653 ohm-m. The second (low resistivity) block within the borehole region is also detected by the three data sets. The full optimized data set model gives a minimum resistivity value of 50 ohm-m (true value 30 ohm-m), while the reduced optimized and 'standard' data sets give values of 56 and 65 ohm-m respectively. The third deepest block between the boreholes is well resolve by all the three data sets, with lowest model values of 15, 16 and 22 ohm-m (true value 10 ohm-

m) respectively for the full optimized, reduced optimized and 'standard' data sets. There is a slight distortion in the shape in the full optimized data set model, probably due to the higher noise level. The deepest block is resolved only in the full optimized data set (Figure 9b). In this test model, the deepest block is placed 1 meter further from the bottom end of the borehole (Figure 9a) compared to the first test model (Figure 8a). It is comparatively much better resolved in the second test model as there is better data coverage by the borehole electrodes. The high resistivity block to the right of the borehole region is resolved by the full and reduced optimized data sets giving values of about 138 and 135 ohm-m (true value 1000 ohm-m) that are well above the background value of 100 ohm-m, while the 'standard' data set model fails to detect it.

Experimental data sets

The measurements were made in a tank with dimensions of 1.0 by 1.4 by 1.5 meters filled with water with resistivity value of 17.02±0.03 ohm-m that was measured with a conductivity meter. The arrangement of the cross-borehole setup is shown in Figure 10. Plastic tubes were placed in between two vertical columns of subsurface 'borehole' electrodes as high resistivity objects to test the resolution of the different arrays. Laboratory measurements made in a tank are more convenient than a field survey but presents special challenges in processing the data. The medium is usually assumed to be an infinite half-space in a field survey but the finite distance to the sides of the tank can have significant effect on the measurements. In 2.5-D modelling, it is assumed that the structures do not change in the *y*-direction. The sides of the tank in the *x* and *z*- directions can be easily incorporated into a 2-D resistivity model, but the sides in the *y*-direction

cannot be as easily handled. The 2.5-D finite-difference and finite-element modelling algorithms use a Fourier transform to convert a 3-D potential function into a series of 2-D potential functions in the wave-number domain (Dey and Morrison 1979). An inverse Fourier transform of the potential values in the wave-number domain is then carried out to calculate the potential in the space domain. A truncation of the resistivity structure in the *y*-direction with a box function in the space domain is equivalent to a convolution with the sinc function (James 2011) in the wave-number domain. This makes the potential function in the wave-number domain more complicated and more difficult to accurately evaluate due to the oscillatory nature of the sinc function. To incorporate the sides of the tank accurately, the inversion of the data sets was carried out using a 3-D inversion program (Loke 2011) and the resistivity model in the *x*-*z* plane at the center of the tank was used. Note this approach also makes it possible to handle measurements made in a non-rectangular tank.

After measuring each of the three data sets, they were remeasured using their reciprocal configurations to estimate the levels of random noise. For each configuration, the data was taken to be the mean of the forward and reciprocal measurements, and the reciprocal error as the standard error in this mean. The distributions of reciprocal errors for each data set are shown in Table 3. The greater levels of random noise in the optimized data sets are consistent with their higher proportions of measurements with larger geometric factors. However, reciprocal error estimates cannot account for certain systematic effects, such as those caused by errors in the electrodes positions and geometry. In these experiments, the distance between adjacent electrodes in the tank setup was only 0.05 meter (50 mm). Small errors in the positions of the electrodes and

their finite sizes can have significant effects on the results for certain cross-borehole configurations (Wilkinson *et al.* 2008). To minimize these geometric errors, the electrodes were mounted on rigid rods and anchored to fixed points on the base of the tank. This ensured that the positions of the electrodes were known to within 1 mm. The finite sizes of the electrodes (about 5 mm) were accounted for by representing each electrode in the inversion model as points located at their geometric centers (Rücker and Günther 2011). In the inversion of the data set, the L-curve method (Farquharson and Oldenburg 2004) was used to estimate the optimum damping factor for each data set. This method attempts to determine the damping factor that balances the opposing requirements to minimize both the model roughness and data misfit.

Figure 11 shows the inversion models obtained for the three data sets. The two larger cylinders are well resolved in the model with the full optimized data set (Figure 11a). The highest resistivity values are located near the centres of the cylinders. The highest resistivity value at the anomaly corresponding to the second deepest cylinder is slightly offset to the left of the cylinder centre. This is partly because the cylinder is smaller and thus produces a smaller change in the apparent resistivity values. The anomaly in the inversion model is thus more sensitive to noise. There is a region of slightly higher resistivity values of over 23 ohm-m near the location of the deepest cylinder but it is poorly resolved. The centre of the region is well outside the actual location of the cylinder. It was shown previously for the synthetic data set the resolution, even for the optimized data set, near the bottom of the borehole zone is expected to be poor. Note the artefacts located along the positions of the two boreholes that are probably caused by small errors in the positions of the borehole electrodes. The L-curve method

selected a damping factor of 0.029 for this data set that resulted in a data misfit of 1.3%. In comparison, the L-curve method selected a lower damping factor of 0.012 for the 'standard' array configurations giving a lower data misfit of 0.6% (Figure 11b). Despite the lower data misfit, the cylinders are much less well resolved in the model for the 'standard' borehole array data set. As an example, the maximum resistivity at the topmost cylinder is less than 38 ohm-m (Figure 11b) compared to over 53 ohm-m (Figure 11a) for the optimized data set model. It also fails to detect the deepest cylinder. The maximum resistivity values at the positions of the three upper cylinders achieved by the model for the reduced optimized data set (Figure 11c) are significantly higher that for the 'standard' arrays (Figure 11b) but it also fails to resolve the deepest cylinder. The L-curve method selected a damping factor of 0.019 for this data set giving a data misfit of 1.0%. These values are in between those obtained for the full optimized and 'standard' arrays data sets. This is in agreement with the average geometric factors of 32.6, 2.4 and 15.5 m. for the full optimized, standard and reduced optimized data sets. A data set with larger geometric factors is not only more sensitive to random noise, but generally also more sensitive to errors in the electrode positions. A larger geometric factor results in a higher overall noise level that consequently causes the L-curve method to select a higher damping factor.

CONCLUSIONS

For a cross-borehole survey, the region with significant model resolution values extends well outside the area between the boreholes. It is important to extend the model region horizontally at least 7 times (and vertically 3 times) the unit electrode spacing so that all

regions that have significant resolution values are included in the model used to calculate the optimized arrays. The ‘Compare R’ method that seeks to optimize the model resolution values provides the best results in terms of the quality of the optimized arrays as well as the lowest calculation time required. A modification of the array optimization algorithm that attempts to minimize the spread function did not produce arrays that have significantly lower spread values but was several times slower than the ‘Compare R’ method. However, a study of the point spread function plots for a cross-borehole survey is useful in understanding the areas that are likely to be well resolved. The regions close to the bottom end of the boreholes will have poor resolution even if all the possible measurements are made. The poor resolution towards the bottom region is due to its remoteness from the electrodes leading to poor data coverage.

Tests with synthetic and experimental data show that the optimized arrays generated by the ‘Compare R’ method have significantly better resolution than a ‘standard’ measurement sequence used in previous surveys. The resolution is less if arrays with both current (or both potential) electrodes in the same borehole are excluded but it is generally still better than the ‘standard’ arrays.

ACKNOWLEDGEMENTS

This paper is published with permission of the Executive Director of the British Geological Survey (NERC). We would also like to thank two anonymous reviewers and Prof. A. T. Basokur for their constructive comments that have helped to improve the paper.

REFERENCES

Auken E., Pellerin L., Christensen N.B. and Sørensen, K. 2006. A survey of current trends in near-surface electrical and electromagnetic methods. *Geophysics* **71**, G249-G260.

Carpenter E.W. and Habberjam G.M. 1956. A tri-potential method of resistivity prospecting. *Geophysics* **11**, 455-469.

Chambers J.E., Kuras O., Meldrum P.I., Ogilvy R.O. and Hollands J. 2006. Electrical resistivity tomography applied to geologic, hydrogeologic, and engineering investigations at a former waste-disposal site. *Geophysics* **71**, B231-B239.

Chambers J.E., Wilkinson P.B., Wealthall G.P., Loke M.H., Dearden R., Wilson R. and Ogilvy R.D. 2010. Hydrogeophysical Imaging of Deposit Heterogeneity and Groundwater Chemistry Changes during DNAPL Source Zone Bioremediation. *Journal of Contaminant Hydrology* **118**, 43–61.

Dahlin T. 1996. 2D resistivity surveying for environmental and engineering applications. *First Break* **14**, 275-284.

Dey A. and Morrison H.F. 1979. Resistivity modelling for arbitrary shaped two-dimensional structures. *Geophysical Prospecting* **27**, 106-136.

Ellis R.G. and Oldenburg D.W. 1994. Applied geophysical inversion, *Geophysical Journal International* **116**, 5-11.

Farber R. 2011. *CUDA application design and development.* Elsevier. IBSN 9780123884268.

Farquharson C.G., and 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.

Farquharson C.G. 2008. Constructing piecewise-constant models in multidimensional minimum-structure inversions. *Geophysics* **73**, K1-K9.

Friedel S. 2003. Resolution, stability and efficiency of resistivity tomography estimated from a generalized inverse approach, *Geophysical Journal International* **153**, 305–316.

Golub G.H. and van Loan C.F. 1989. *Matrix Computations (2nd edn)*. The John Hopkins University Press. ISBN 0801854148.

Griffiths D.H., Turnbull J. and Olayinka A.I. 1990. Two-dimensional resistivity mapping with a computer- controlled array. *First Break* **8**, 121-129.

James J.F. 2011. *A Student's Guide to Fourier Transforms: With Applications in Physics and Engineering (3rd Edition)*. Cambridge University Press. ISBN 9780521176835.

Hagrey S. A. al, 2012. 2D optimized electrode arrays for borehole resistivity tomography and CO2 sequestration modelling. *Pure and Applied Geophysics* **169**, 1283-1292.

Loke M.H. and Barker, R.D., 1995. Least-squares deconvolution of apparent resistivity pseudosections. *Geophysics* **60**, 1682-1690.

Loke M.H., Acworth I. and Dahlin T., 2003. A comparison of smooth and blocky inversion methods in 2D electrical imaging surveys. *Exploration Geophysics* **34**, 182-187.

Loke M.H., Wilkinson P. and Chambers J. 2010a. Fast computation of optimized electrode arrays for 2D resistivity surveys. *Computers & Geosciences* **36**, 1414-1426.

Loke M.H., Wilkinson P. and Chambers J. 2010b. Parallel computation of optimized arrays for 2-D electrical imaging. *Geophysical Journal International* **183**, 1202-1315.

Loke M.H. 2011. Electrical resistivity surveys and data interpretation. In: *Encyclopaedia of Solid Earth Geophysics,* Vol. 1 (ed. H. K. Gupta), pp. 276-283, Springer-Verlag. IBSN 9789048187010.

Loke M.H., Chambers J.E. and Kuras O., 2011. Instrumentation, electrical resistivity. In: *Encyclopaedia of Solid Earth Geophysics,* Vol. 1 (ed. H. K. Gupta), pp. 599-604, Springer-Verlag. IBSN 9789048187010.

Menke W. 1989. *Geophysical data analysis: Discrete inverse theory (Revised edition).* Academic Press Inc. IBSN 0124909213.

Miller C.R. and Routh P.S. 2007. Resolution analysis of geophysical images: Comparison between point spread function and region of influence measures. *Geophysical Prospecting* **55**, 835-852.

Nenna V., Pidlisecky A. and Knight R., 2011, Informed experimental design for electrical resistivity imaging. *Near Surface Geophysics* **9**, 469-482.

Oldenborger G.A. and Routh P.S. 2009. The point-spread function measure of resolution for the 3D electrical resistivity experiment. *Geophysical Journal International* **176**, 405-414.

Press W.H. , Teukolsky S.A., Vetterling W.T. and Flannery B.P. 2007. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press. IBSN 9780521880688.

Routh P. S., Oldenborger G.A. and Oldenburg D.W. 2005. Optimal survey design using the point spread function measure of resolution. 75th Annual International Meeting, SEG, Expanded Abstracts, 1033–1036.

Rücker C. and Günther T. 2011. The simulation of finite ERT electrodes using the complete electrode model. *Geophysics* **76**, F227-F238.

Slater L., Binley A.M., Daily W. and Johnson R. 2000. Cross-hole electrical imaging of a controlled saline tracer injection. *Journal of Applied Geophysics* **44**, 85-102.

Stummer P., Maurer H. and Green A. 2004. Experimental design: Electrical resistivity data sets that provide optimum subsurface information. *Geophysics* **69**, 120-129.

Wilkinson P.B., Chambers J.E., Meldrum P.I., Ogilvy R.D. and Caunt S., 2006a. Optimization of array configurations and panel combinations for the detection and imaging of abandoned mineshafts using 3D cross-hole electrical resistivity tomography. *Journal of Environmental and Engineering Geophysics* **11**, 213-221.

Wilkinson P.B., Meldrum P.I., Chambers J.E., Kuras O. and Ogilvy R.D. 2006b. Improved strategies for the automatic selection of optimized sets of electrical resistivity tomography measurement configurations. *Geophysical Journal International* **167**, 1119-1126.

Wilkinson P B, Chambers J E, Lelliott M., Wealthall P. and Ogilvy R. D. 2008. Extreme sensitivity of crosshole electrical resistivity tomography measurements to geometric errors. *Geophysical Journal International* **173**, 49-62.

Wilkinson P.B., Loke M.H., Meldrum P.I., Chambers J.E, Kuras O., Gunn D.A. and Ogilvy R.D. 2012. Practical aspects of applied optimised survey design for Electrical Resistivity Tomography. *Geophysical Journal International* **189**, 428-440.

List of captions

Table 1 Results for the different optimization methods for an optimized data set with the cross-borehole setup using 11 electrodes on the surface and 20 electrodes in two boreholes using the model resolution (CR) and spread function (CS) optimization methods. S_r = average relative model resolution, S_v = average spread value.

Table 2 Maximum resistivity value at location of high resistivity blocks. Block 1 is located near the surface while block 4 is the deepest. The background resistivity is 15 ohm-m and the true resistivity of all the blocks is 1500 ohm-m.

Table 3 Percentage of each data set with reciprocal error less than the stated error level for the standard, reduced optimized, and full optimized experiments.

Figure 1 Comprehensive data set model resolution sections for (a) survey line with 31 electrodes on the ground surface, (b) cross-borehole configuration with 20 electrodes along 2 boreholes and 11 electrodes on the surface.

Figure 2 (a) Change of the average relative model resolution with number of data points in the optimized data sets generated with using different array optimization methods. (b) Similar plots showing change of the average spread criterion value with the number of data points.

Figure 3 Point spread function plots for the cross-borehole configuration using the comprehensive data set. The spread values for a (a) model cell near the central region between the two boreholes, (b) a model cell near the bottom edge of the boreholes.

Figure 4 Point spread function plots for the cross-borehole configuration using the optimized data set with 1875 points. The spread values for a (a) model cell near the central region between the two boreholes, (b) a model cell near the bottom edge of the boreholes.

Figure 5 Point spread function plots for the cross-borehole configuration using the ‘standard’ data set with 1875 points. The spread values for a (a) model cell near the central region between the two boreholes, (b) a model cell near the bottom edge of the boreholes.

Figure 6 Point spread function plots for a model cell to the right of the borehole region. The spread values for (a) the optimized data set and (b) the 'standard' data set with 1875 points.

Figure 7 Relative model resolution sections for (a) the optimized data set and (b) the 'standard' data set with 1875 points.

Figure 8 Inversion models for (a) optimized data set with all arrays, (b) 'standard' data set and (c) the reduced optimized data set that excludes arrays with both current (or

potential) electrodes in the same borehole. All the data sets have 1875 data points. The true model has resistive anomalies of 1500 ohm-m (outlines shown) within a homogenous medium of 15 ohm-m.

Figure 9 (a) Two-layer test model with conductive and resistive anomalies. Inversion models for (b) optimized data set with all arrays, (c) 'standard' data set and (d) the reduced optimized data set that excludes arrays with both current (or potential) electrodes in the same borehole. All the data sets have 1875 data points. The outlines of the rectangular blocks showing their true positions are also shown.

Figure 10 Cross section of the setup for the tank experiment with 11 electrodes on the surface and 20 electrodes in two boreholes. The spacing between the electrodes is 0.05 m. The distance of the electrodes from the sides of the tank in the y-direction is 0.7 m. The outlines of the cylindrical plastic tubes within a water medium are shown.

Figure 11 Inversion models for the cross-borehole tank experiment using (a) the full optimized data set, (b) the ‘standard’ measurement sequence and (c) the reduced optimized data set. All the data sets have 1875 data points. The actual locations of the plastic cylinders are marked by black circles.

Table 1 Results for the different optimization methods for an optimized data set with the cross-borehole setup using 11 electrodes on the surface and 20 electrodes in two boreholes using the model resolution (CR) and spread function (CS) optimization methods. S_r = average relative model resolution, S_v = average spread value.

	1875 data points			4000 data points		
Method	Time (s.)	S_r	S_v	Time (s.)	S_r	S_v
CR Single step	4633	0.6079	10.4524	10909	0.6867	9.323
CR 5%	248	0.5891	10.786	338	0.6763	9.489
CS 5%	5508	0.5661	10.790	7568	0.6616	9.354
CR 10%	141	0.5567	11.428	189	0.6622	9.680

Table 2 Maximum resistivity value at location of high resistivity blocks. Block 1 is located near the surface while block 4 is the deepest. The background resistivity is 15 ohm-m and the true resistivity of all the blocks is 1500 ohm-m.

Method	Block 1	Block 2	Block 3	Block 4
Optimized data set	335.2	114.1	152.7	21.9
Standard arrays	262.4	71.3	97.9	17.0
Reduced optimized set	543.0	106.0	204.1	18.3

Table 3 Percentage of each data set with reciprocal error less than the stated error level for the standard, reduced optimized, and full optimized experiments.

Error level (%)	Standard	Reduced optimized	Full optimized
0.1	39.2 %	10.7 %	9.8 %
0.2	66.1 %	20.8 %	18.1 %
0.5	93.5 %	43.4 %	38.5 %
1.0	99.6 %	70.1 %	57.4 %
2.0	99.8 %	94.1 %	78.8 %
5.0	100 %	98.7 %	99.0 %

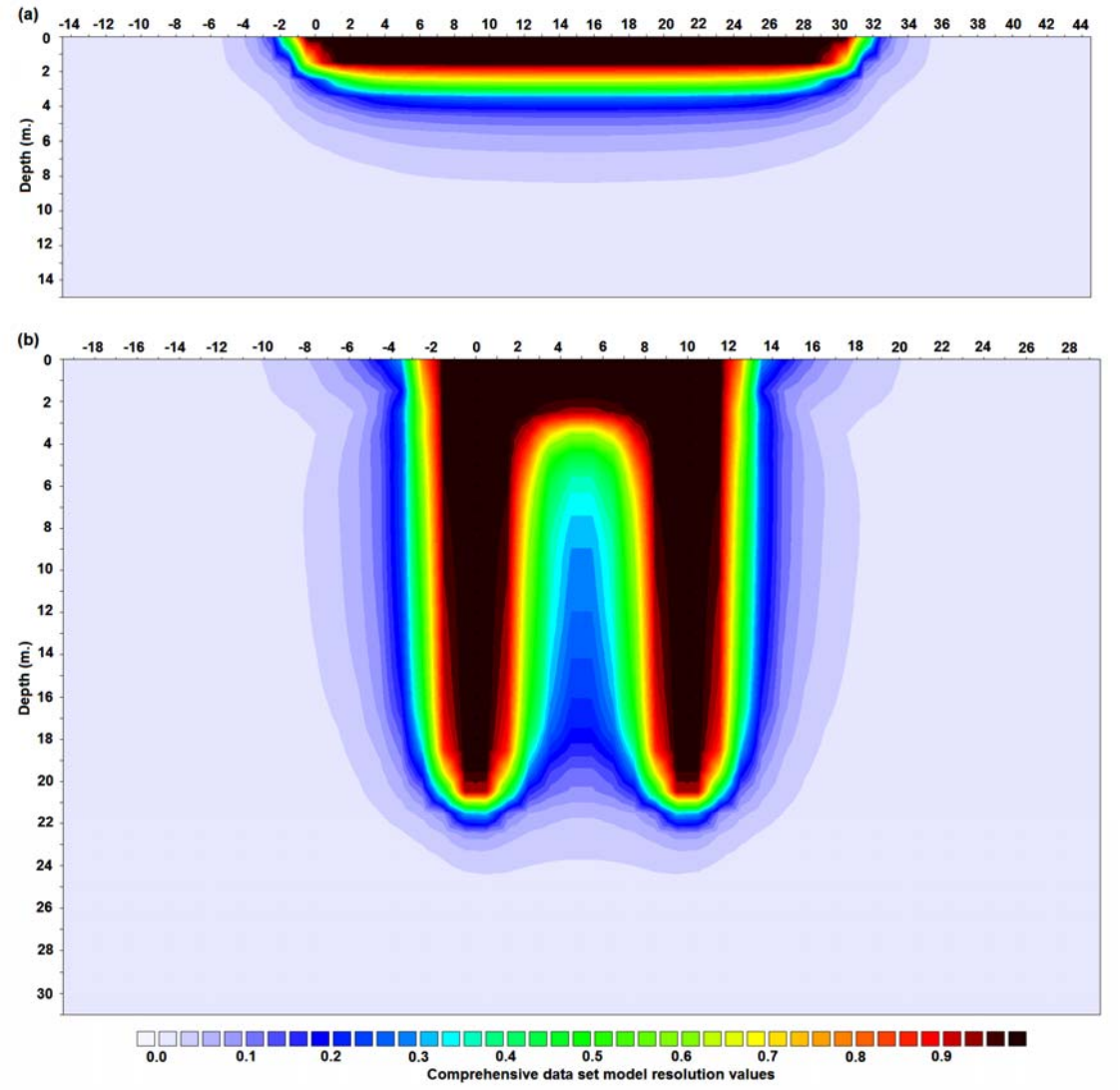

Figure 1 Comprehensive data set model resolution sections for (a) survey line with 31 electrodes on the ground surface, (b) cross-borehole configuration with 20 electrodes along 2 boreholes and 11 electrodes on the surface.

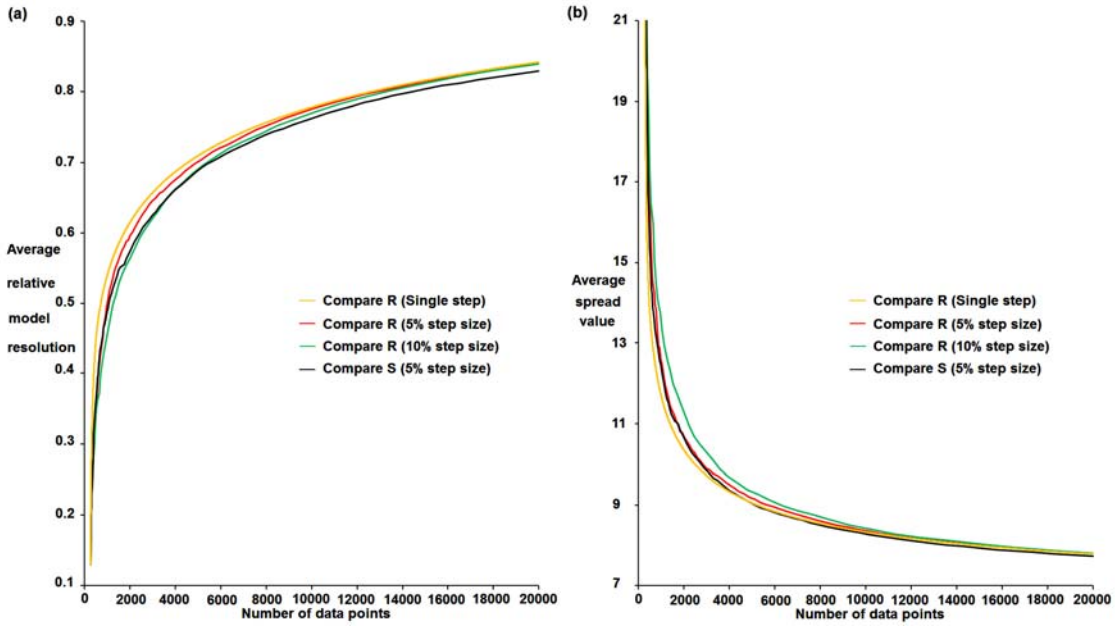

Figure 2 (a) Change of the average relative model resolution with number of data points in the optimized data sets generated with using different array optimization methods. (b) Similar plots showing change of the average spread criterion value with the number of data points.

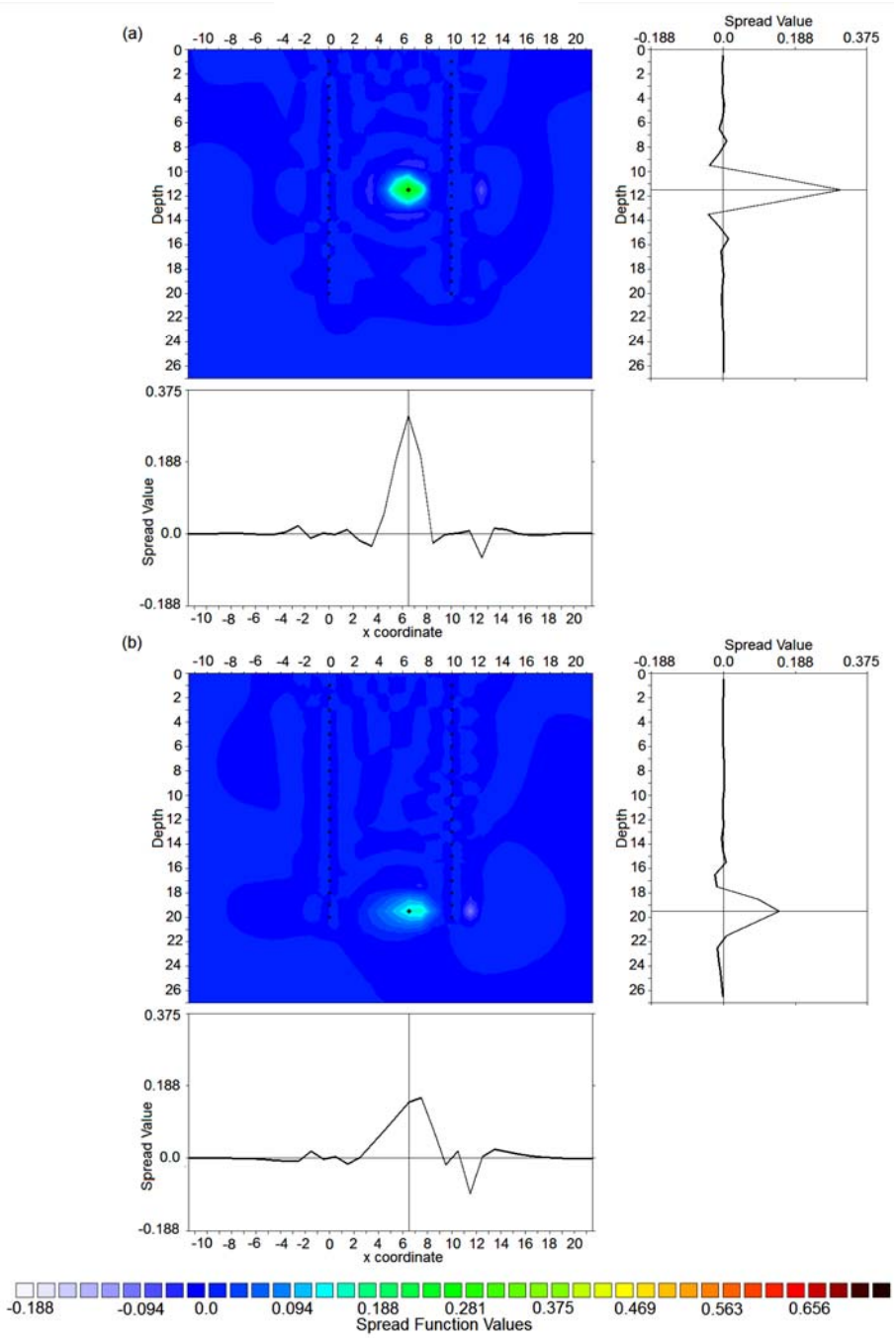

Figure 3 Point spread function plots for the cross-borehole configuration using the comprehensive data set. The spread values for a (a) model cell near the central region between the two boreholes, (b) a model cell near the bottom edge of the boreholes.

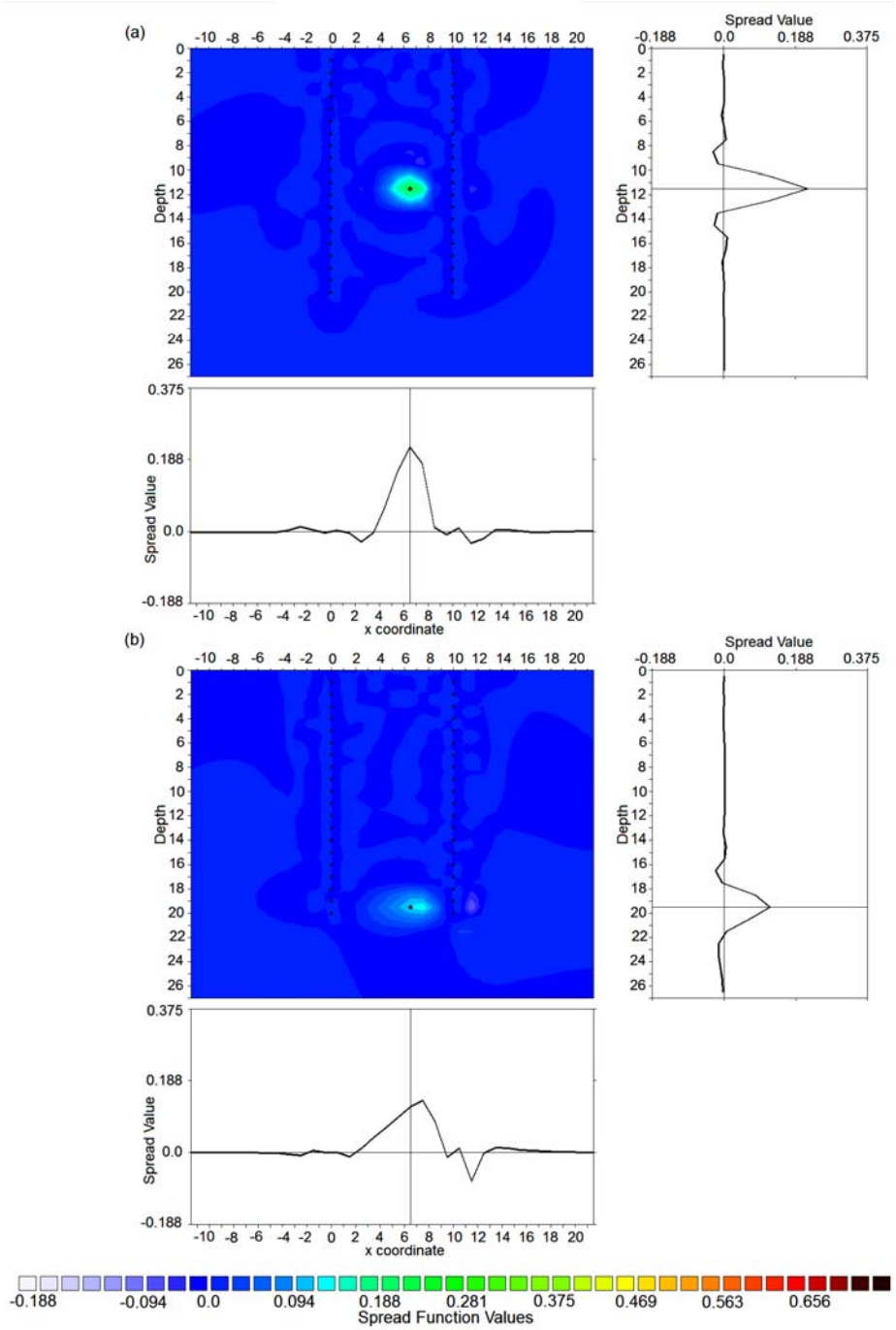

Figure 4 Point spread function plots for the cross-borehole configuration using the optimized data set with 1875 points. The spread values for a (a) model cell near the central region between the two boreholes, (b) a model cell near the bottom edge of the boreholes.

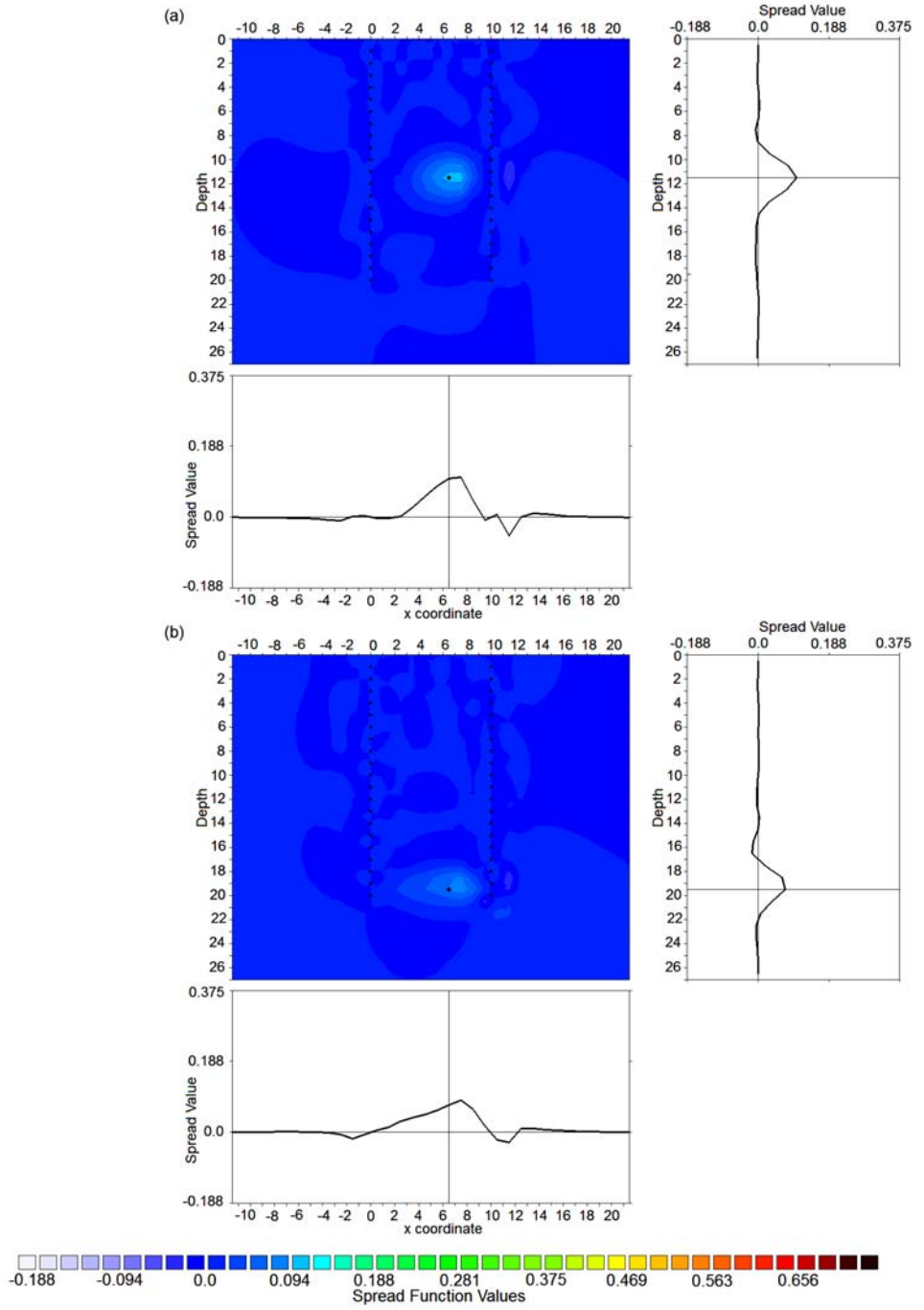

Figure 5 Point spread function plots for the cross-borehole configuration using the 'standard' data set with 1875 points. The spread values for a (a) model cell near the central region between the two boreholes, (b) a model cell near the bottom edge of the boreholes.

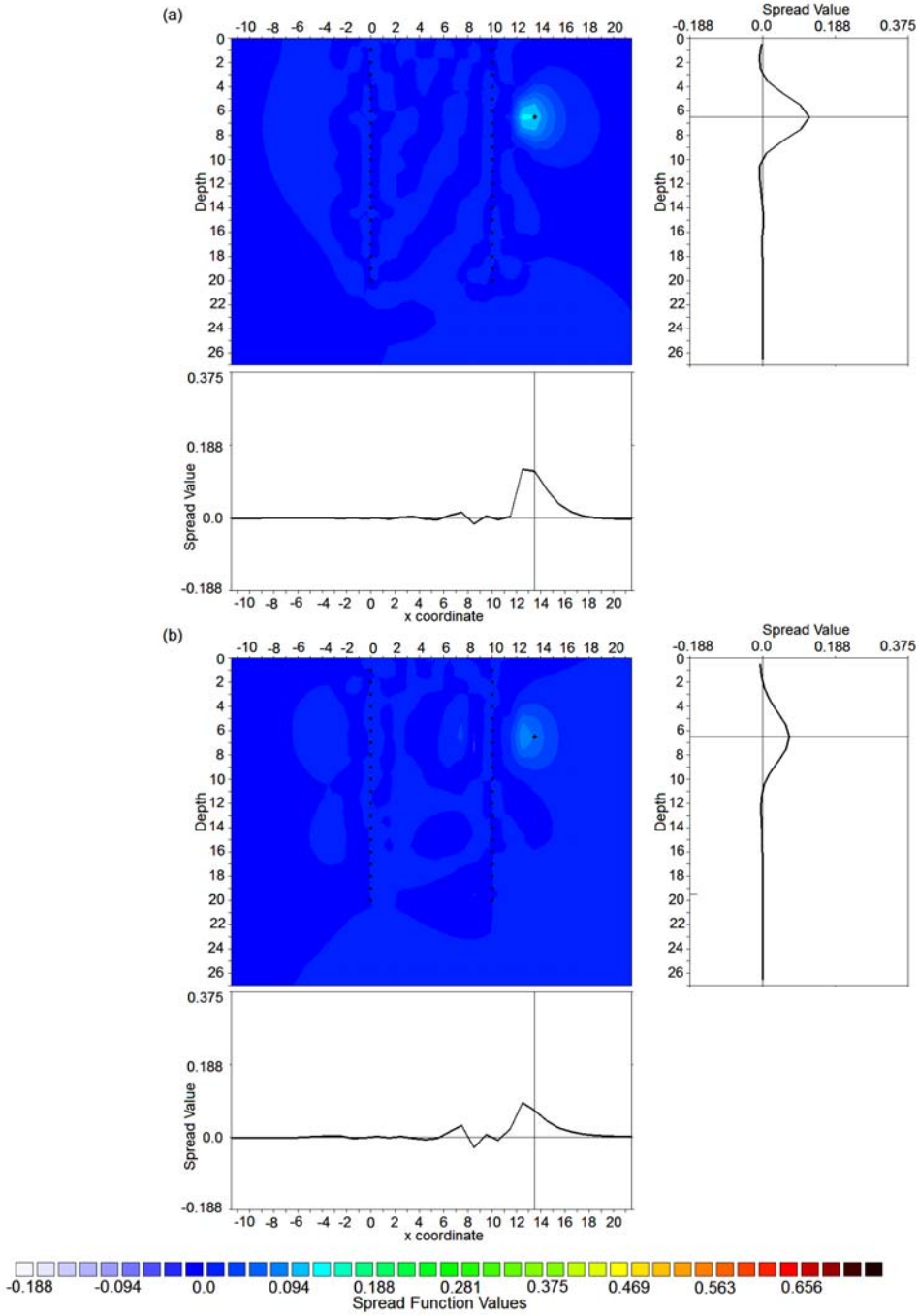

Figure 6 Point spread function plots for a model cell to the right of the borehole region. The spread values for (a) the optimized data set and (b) the 'standard' data set with 1875 points.

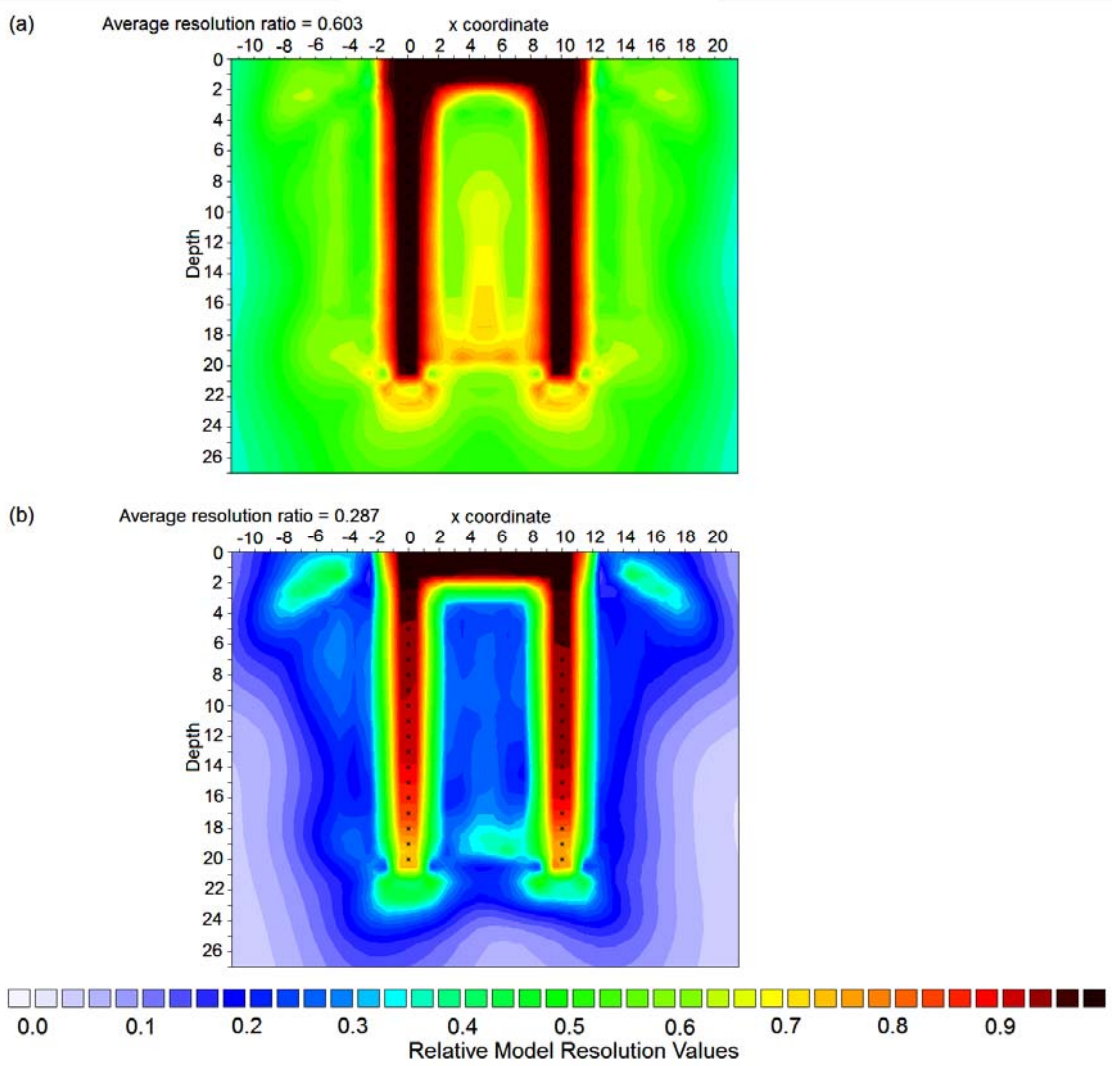

Figure 7 Relative model resolution sections for (a) the optimized data set and (b) the 'standard' data set with 1875 points.

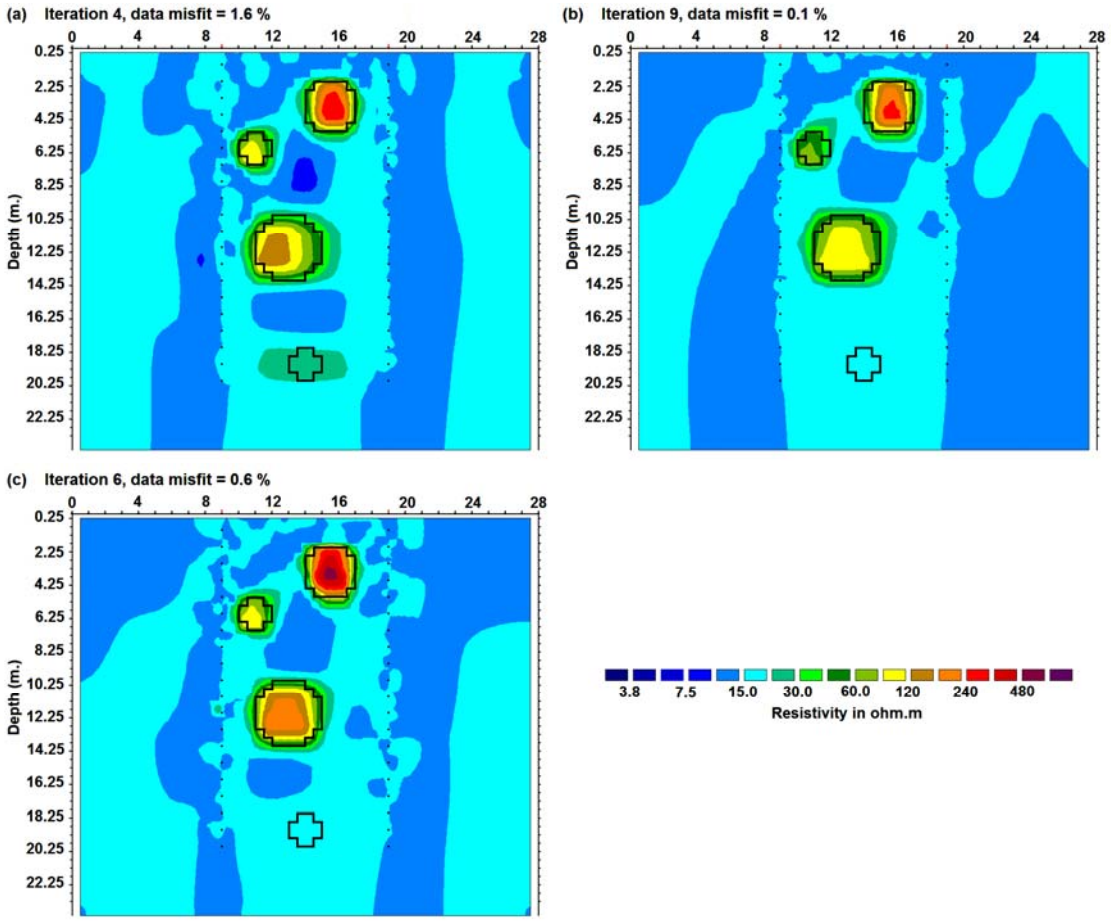

Figure 8 Inversion models for (a) optimized data set with all arrays, (b) 'standard' data set and (c) the reduced optimized data set that excludes arrays with both current (or potential) electrodes in the same borehole. All the data sets have 1875 data points. The true model has resistive anomalies of 1500 ohm-m (outlines shown) within a homogenous medium of 15 ohm-m.

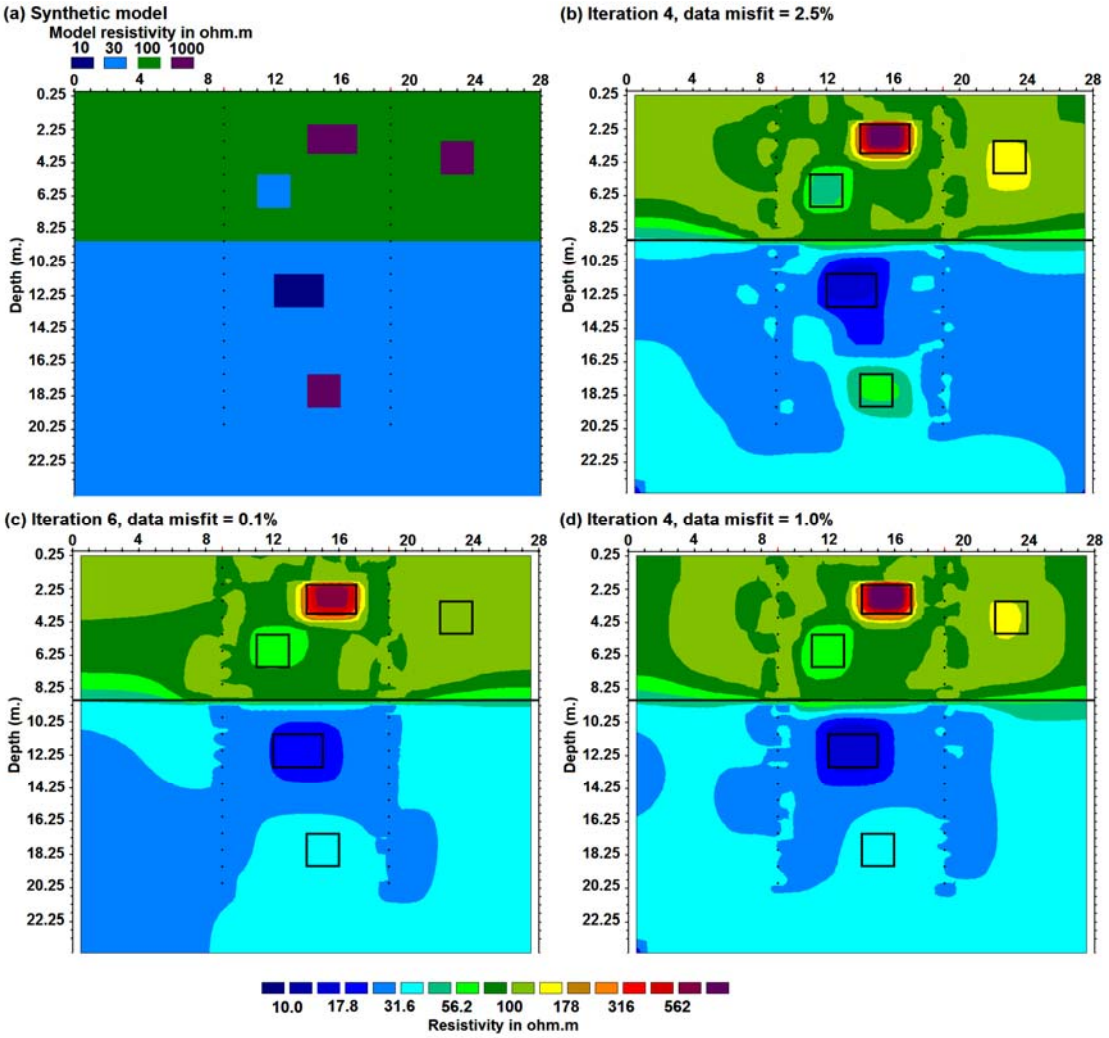

Figure 9 (a) Two-layer test model with conductive and resistive anomalies. Inversion models for (b) optimized data set with all arrays, (c) 'standard' data set and (d) the reduced optimized data set that excludes arrays with both current (or potential) electrodes in the same borehole. All the data sets have 1875 data points. The outlines of the rectangular blocks showing their true positions are also shown.

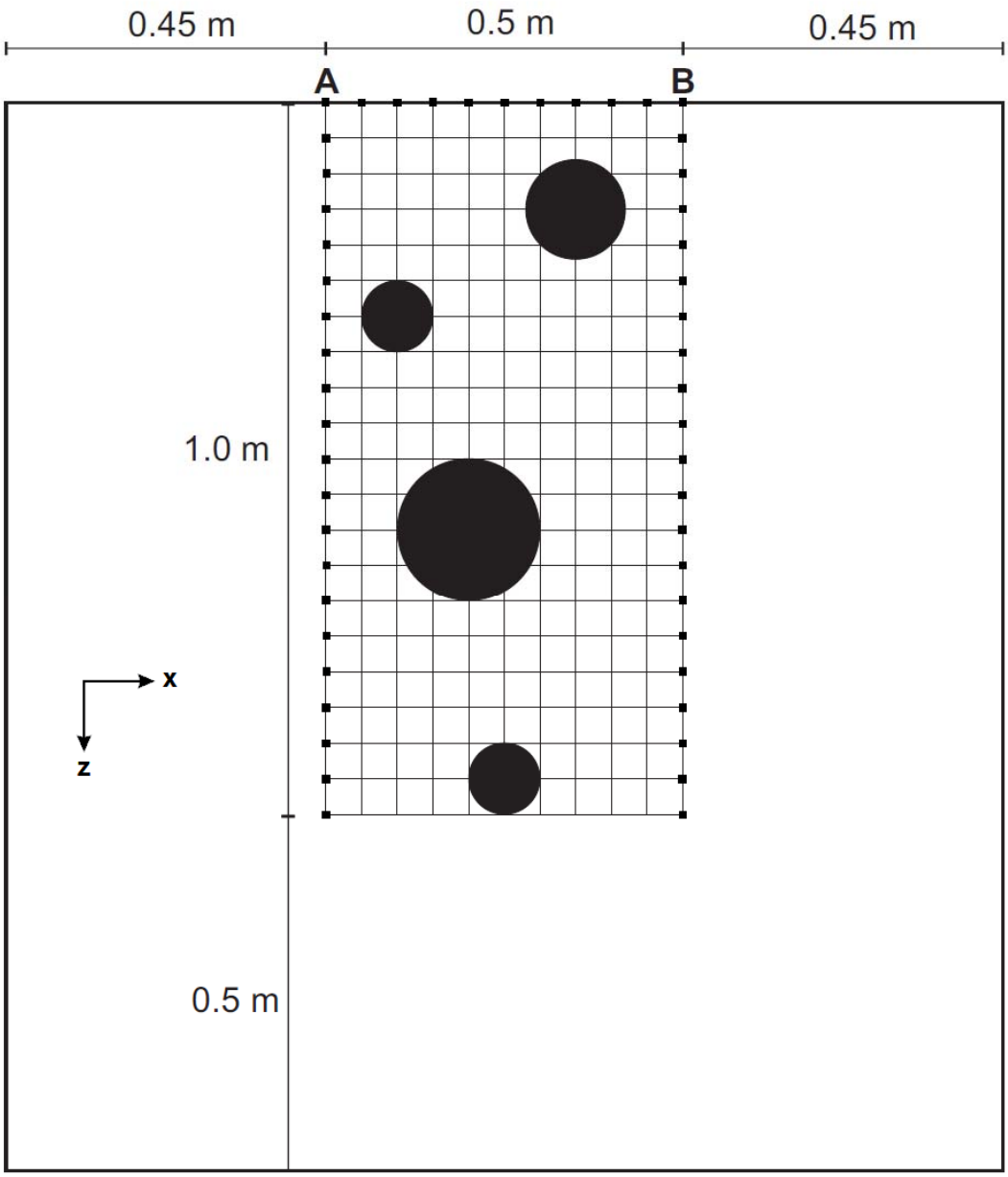

Figure 10 Cross section of the setup for the tank experiment with 11 electrodes on the surface and 20 electrodes in two boreholes. The spacing between the electrodes is 0.05 m. The distance of the electrodes from the sides of the tank in the y-direction is 0.7 m. The outlines of the cylindrical plastic tubes within a water medium are shown.

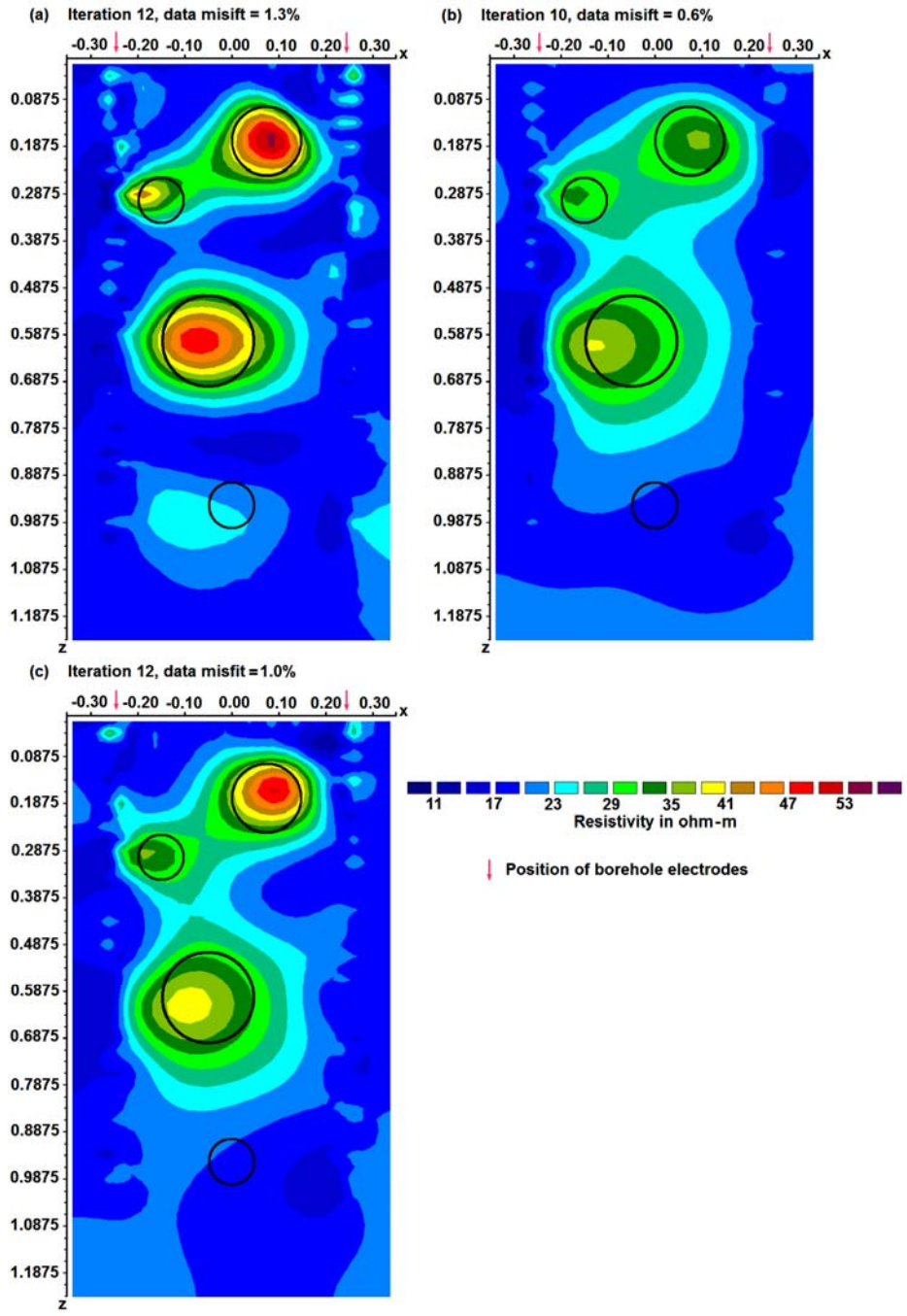

Figure 11 Inversion models for the cross-borehole tank experiment using (a) the full optimized data set, (b) the 'standard' measurement sequence and (c) the reduced optimized data set. All the data sets have 1875 data points. The actual locations of the plastic cylinders are marked by black circles.