

Field test of imaging methodology, resolution and uncertainty determination using optimized data acquisition strategies

Work package 2
Milestone 2.2

Roland Martin, Andreas Kemna
University of Bonn (UBO)

December 20, 2011

Introduction

In the last two decades, electrical imaging has become a widely used exploration tool in the field of geophysics (see, e.g., Binley & Kemna [2005]). However, to date there is only a limited number of studies addressing image appraisal based on parameter uncertainty computation in electrical imaging, although this is considered essential especially for quantitative imaging applications. The understanding of the propagation of data errors and of uncertainties associated with the regularization constraint into the geophysical image represents an essential prerequisite for using the imaging results in an integrated data fusion approach aiming at parameter uncertainty prediction and risk assessment (input to WP3 and WP8).

In order to assess the characterization capabilities of the method for the present purpose, resolution and uncertainty in terms of parameters describing the geophysical signatures have been analyzed for a given field example on the basis of enhanced data acquisition methodology using already existing EIT codes.

We compare qualitative and quantitative means for image appraisal by comparing coverage and resolution for a given field example and incorporate qualitative and quantitative means for image appraisal. To address resolution measures, we compare Sensitivities and Resolution quality for a given field example. We indicate the resolution borders by enhanced image processing incorporating resolution as transparency map within the conductivity and phase images. Also, we supply means of a quantitative error estimate: Based on linearized computation schemes, parameter variance related on one hand to data errors, and on the other hand, to uncertainty in the prior model, is combined to assess the standard deviation of our final estimate.

Optimal model parameter estimation

The non-linear inverse problem in electrical impedance tomography, i.e., the optimal estimation of M complex model parameters $\mathbf{m}^* \in \mathbb{C}^M$ from N given measured (complex) data $\mathbf{d} \in \mathbb{C}^N$, contaminated by noise, is usually solved in a deterministic way by solving the normal equations of the linearized and regularized [Tikhonov & Arsenin, 1977] problem for the n 'th model update $\delta\mathbf{m}_n$, starting at the prior guess \mathbf{m}_{m0} ,

$$(\mathbf{A}_n^H \mathbf{C}_d^{-1} \mathbf{A}_n + \lambda \mathbf{R}^T \mathbf{R}) \delta\mathbf{m}_n = (\mathbf{A}_n^H \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{f}(\mathbf{m}_n)) - \lambda \mathbf{R}^T \mathbf{R} (\mathbf{m}_n - \mathbf{m}_{m0})) , \quad (1)$$

and forming a converging sequence to determine the optimal estimate (OE) after n^* iterations:

$$\mathbf{m}^* = \mathbf{m}_{m0} + \sum_n^{n^*} \alpha_n \delta\mathbf{m}_n . \quad (2)$$

Here, \mathbf{f} describes the non linear model response which maps parameter space into data space: $(\mathbf{f}(\mathbf{m}) : \mathbb{C}^M \rightarrow \mathbb{C}^N)$. \mathbf{A} is the Jacobian matrix of partial derivatives (i.e. $A_{ij} := \partial f_i(\mathbf{m}) / \partial m_j$) displaying the sensitivity of the i 'th forward model prediction with respect to

changes of the j 'th model parameter m_j , \mathbf{C}_d the data covariance matrix, which is a diagonal matrix for uncorrelated data errors (i.e. $\mathbf{C}_d = \{\epsilon_{ii}^2\}$), \mathbf{R} the roughness matrix imposing a smoothness constraint on the model (regularization), λ the regularization parameter, and $0 < \alpha \leq 1$ is the step length of each model update¹. H denotes the conjugate transpose operator. The smoothness constraint, also known as OCCAM inversion [Constable et al., 1987], maintains active regularization between neighboring model cells via a discretized first-order finite-difference operator matrix (see for example [Kemna, 2000]).

Complex data and model parameter representation

Within the optimal estimation, log transformed parameters and data are used to account for the wide dynamic range of conductivity magnitudes for earth materials. The model vector \mathbf{m} and data vector \mathbf{d} are defined as

$$m_j = \ln \rho_j \quad (3)$$

$$= \rho'_j + \mathbf{i}\rho''_j \quad (4)$$

$$= \ln |\rho_j| + \mathbf{i}\phi_j, \quad (j = 1, \dots, M), \quad (5)$$

where ρ' denotes the real and ρ'' denotes the imaginary part $\mathbf{i}^2 = -1$ is the imaginary unit and

$$d_i = \ln Z_i \quad (6)$$

$$= \ln |Z_i| + \mathbf{i}\phi_i, \quad (i = 1, \dots, N), \quad (7)$$

are the transfer impedances. The inverse conductivity (resistivity)

$$\sigma_j^{-1} = \rho_j \quad (8)$$

$$= |\rho_j| e^{\mathbf{i}\phi_j} \quad (9)$$

is complex valued and discretized on a finite element mesh.

$$\phi_j = \arctan \left(\frac{\rho''_j}{\rho'_j} \right) \quad (10)$$

$$= \arctan \left(-\frac{\sigma''_j}{\sigma'_j} \right) \quad (11)$$

are the corresponding phase values for resistivity and conductivity, respectively. Please note, that the resistivity phase equals the negative conductivity phase.

$$Z_i = \frac{V_i}{I_i} \quad (12)$$

$$= |Z_i| e^{\mathbf{i}\phi_i} \quad (13)$$

¹usually $\alpha \approx 1$, but it can be optimized in a last parabola fitting routine [Kemna, 2000] as well.

is the measured complex transfer impedance. Note, that the complex log function in equations (5) and (7) separates the log magnitude and phase of its argument into real and imaginary part, for example $\Re(d) = \ln |Z|$ and $\Im(d) = \phi$.

From the equations (5) and (7) follows, that the elements of the Jacobian matrix are altered as well. By applying the complex version of the chain rule, the complex valued sensitivity is computed like

$$A_{ij} = \frac{\partial f_i(\mathbf{m})}{\partial m_j} \quad (14)$$

$$\begin{aligned} (I_i = 1) \quad & \frac{\rho_j}{V_i} \frac{\partial V_i}{\partial \rho_j} \\ & = \end{aligned} \quad (15)$$

The right-hand side of equation (15) is computed via the superposition of adjoint fields (e.g. Kemna [2000]).

Although for real valued problems, the data can vary only in one dimension, i.e., any data noise is one-dimensional, a complex number may be subject to errors in two dimensions if the real and imaginary parts yield stochastic independent random variables. To describe the misfit of the complex quantities, we express the denominators of the chi-squared data misfit functional according to [Kemna, 2000]:

$$\psi_d(\mathbf{m}) = [\mathbf{d} - \mathbf{f}(\mathbf{m})]^H \mathbf{C}_d^{-1} [\mathbf{d} - \mathbf{f}(\mathbf{m})] \quad (16)$$

$$= \sum_i \frac{|d_i - f_i(\mathbf{m})|^2}{|\epsilon_i|^2}, \quad (17)$$

with

$$\epsilon_i = s(\ln |Z_i|) + i s(\phi_i), \quad (18)$$

where $s()$ denotes the standard deviation of the quantity. These values may be determined from reciprocal analysis (i.e. Flores Orozco et al. [2011]; Slater et al. [2000]), by statistical means or estimated according to numerical considerations (electrode spacing, mesh quality, etc.).

Measurements at the Trecate test site

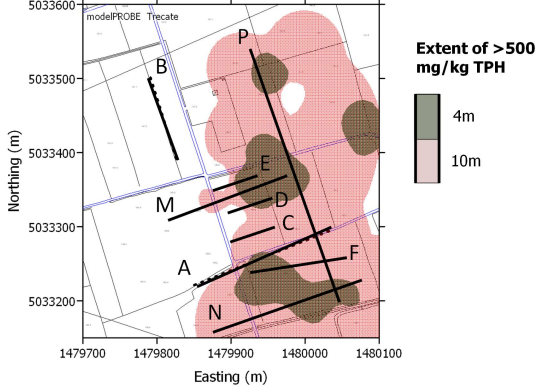


Figure 1: Schematic representation of the complex resistivity measurements performed at the Trecate test site. Solid lines represent the different electrical lines, and the contour plots the contaminant distribution at different depths.

order to estimate the data error.

Based on the imaging results, different locations were selected to collect information by means of direct-push technologies. In particular 10 locations were selected to run measurements with the EC-probe. Here data were collected for 10 different sampling points as depicted in Figure 2, for a maximum depth of 12 m.

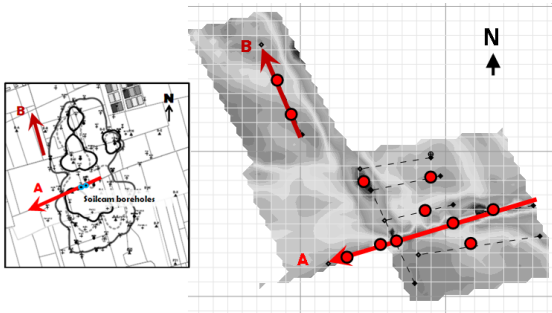


Figure 2: Location of the measurements performed with the EC-probe (solid circles) in reference with the position of the lines for acquisition of complex resistivity data (solid lines).

Images of the electrical conductivity for data collected along line A are presented in

To test our resolution and uncertainty determination we used a field example with optimized data acquisition strategies. In order to characterize the site with a high spatial resolution, complex resistivity images have been collected at nine different arrays, as depicted in Figure 1. The extension and location of the arrays have been selected taking into the consideration the expected distribution of the contaminants in the subsurface, as measured in previous studies (e.g., Cassiani et al. [2009]).

The measuring protocol has been carefully designed in order to increase the signal-to-noise ratio for deep measurements, to minimize unwanted electromagnetic coupling effects and to avoid potential measurements with electrodes previously used for current injection, i.e., with polarized electrodes. All measurements in each profile consist of normal and reciprocal pairs in

For a better comparison between the complex resistivity images and the data collected by means of the EC-probe, imaging results will be displayed in terms of the complex conductivity - real (σ') and imaginary (σ'') components and phase (ϕ). Meanwhile, data collected by means of the EC-probe is presented as overlapping pixels in the complex conductivity images. The position of the groundwater level is indicated with the black line at ≈ 10 m below ground surface (bgs). The position of the electrodes at the surface is indicated by means of the solid circles.

Figure 3. The Plots show a high consistency between the collected σ' values by means of the EC-probe and those computed within the inversion of tomographic data. Agreement is also observed between an increase in the conductivity values (σ') and the presence of the groundwater table. Furthermore, Figure 3 shows a significant contrast between the left and right side of the electrical images: whereas clean locations (in the first 50 m along profile direction) reveal relatively low σ' values, the areas associated with high contaminants concentrations (between 50 and 200 m along profile direction) exhibit high σ' values, likely associated with natural degradation of the contaminants increasing the conductivity in pore water through the degradation of soil materials. The same contrast is observed for polarization values (σ''); however, here the occurrence of contaminants are associated with a decrease in the polarization effects, most probably due to the presence of contaminants coating the soil material and impeding the formation of the electrical double layer (EDL) required to enhance polarization. This pattern is also noticeable for phase values (ϕ) and likely controlled by the same mechanisms.

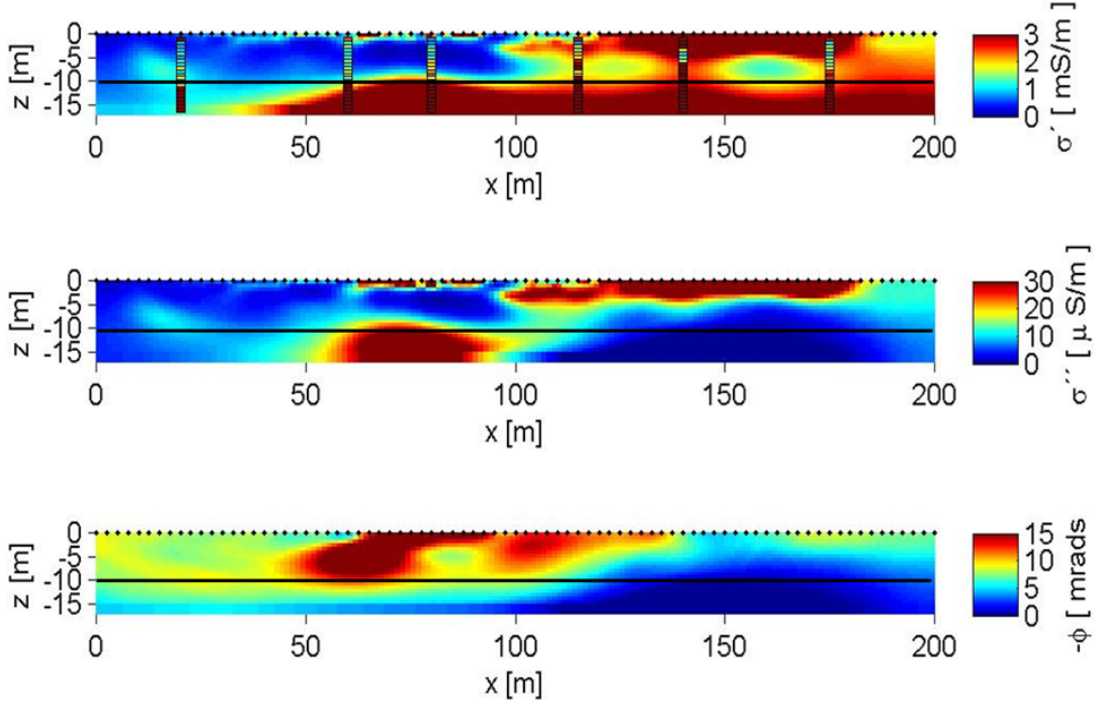


Figure 3: Complex conductivity images computed for measurements collected along line A. This line crosses an uncontaminated area (the first 50 m along profile direction) and an area with high concentrations of contaminants (at ≈ 120 m along profile direction). For comparison, the data collected by means of direct-push technologies with the EC-probe is presented with the overlapping pixels in the image of the real component of the complex conductivity (σ'). The position of the groundwater table is indicated with the solid line.

Resolution determination for the field example

We compare Sensitivity and Resolution measures for the field example using optimized data acquisition strategy and provide means to incorporate resolution measures within the EIT.

Sensitivity

The sensitivity is defined as mapping small changes of the forward model with respect to small changes in model space. The idea of the accumulated sensitivity is, instead of examining every entry of the Jacobian matrix, which will be too complex to investigate for large data sets, to compute the absolute influence of one parameter change to the complete data set, i.e. to map how specific regions of the model space are "covered" by the data.

Analogous to seismic tomography, and taking data covariance into account, we compute the sensitivity as

$$c_j^d = \sum_i^N \frac{|A_{ij}|^2}{|\epsilon_i|^2} = (\mathbf{A}^H \mathbf{C}_d^{-1} \mathbf{A})_{jj} , \quad (19)$$

which represents exactly the diagonal entries of our Hessian approximation within OE, except that the sensitivity is computed for the optimal model after the iterative process (Equation 2) terminates.

Obviously, a poorly covered region is unlikely to be well resolved and thus, the sensitivity may give some crude indication for how well the model parameter is represented by the data set. However, it must be emphasized, that high sensitivity does not necessarily imply high resolution but rather represents a favoring factor.

Resolution measures

The Resolution matrix, as used in classic inversion theory (e.g. [Menke, 1984]), addresses the question if all model parameters are independently resolved or not (with emphasis on their independence). For non linear inversion, we consider the case, where the data vector is ideally mapped via the forward operator of the true model, i.e. $\mathbf{d} = \mathbf{f}(\mathbf{m}^{**})$. We now want to find the expression how good this assumption is represented by the actual model estimate.

We start our investigations with a Taylor series around the n^{*} 'th model estimate:

$$\mathbf{d} = \mathbf{f}(\mathbf{m}^{**}) = \mathbf{f}(\mathbf{m}_{n^*}) + \mathbf{A}_{n^*} (\mathbf{m}^{**} - \mathbf{m}_{n^*}) + \mathcal{O}(\mathbf{m}^{**} - \mathbf{m}_{n^*})^2 . \quad (20)$$

Let us now assume, that the model estimate is already close to the true model. That means that higher order terms ($\mathcal{O}((\mathbf{m}^{**} - \mathbf{m}_{n^*})^2)$) can be neglected. For simplification, we define the generalized inverse for the regularized case:

$$\mathbf{A}_{n^*}^\dagger := (\mathbf{A}_{n^*}^H \mathbf{C}_d^{-1} \mathbf{A}_{n^*} + \lambda \mathbf{R}^T \mathbf{R})^{-1} . \quad (21)$$

Without loss of generality we take the next model update as our final model estimate, that is

$$\mathbf{m}^* = \mathbf{m}_{n^*} + \delta \mathbf{m}_{n^*} \quad (22)$$

$$= \mathbf{m}_{n^*} + \mathbf{A}_{n^*}^\dagger [\mathbf{A}_{n^*}^H \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{f}(\mathbf{m}_{n^*})) - \lambda \mathbf{R}^T \mathbf{R} (\mathbf{m}_{n^*} - \mathbf{m}_{m0})] , \quad (23)$$

and replace $\mathbf{d} - \mathbf{f}(\mathbf{m}_{n^*}) = \mathbf{A}_{n^*} (\mathbf{m}^{**} - \mathbf{m}_{n^*})$ as the remainder term of equation (20) in equation (23). Thus, the last estimate (i.e. the OE) reads

$$\mathbf{m}^* = \mathbf{m}_{n^*} + \mathbf{A}_{n^*}^\dagger \mathbf{A}_{n^*}^H \mathbf{C}_d^{-1} \mathbf{A}_{n^*} (\mathbf{m}^{**} - \mathbf{m}_{n^*}) - \lambda \mathbf{A}_{n^*}^\dagger \mathbf{R}^T \mathbf{R} (\mathbf{m}_{n^*} - \mathbf{m}_{m0}) \quad (24)$$

$$= \mathbf{m}_{n^*} - \mathbf{A}_{n^*}^\dagger \mathbf{A}_{n^*}^H \mathbf{C}_d^{-1} \mathbf{A}_{n^*} \mathbf{m}_{n^*} - \lambda \mathbf{A}_{n^*}^\dagger \mathbf{R}^T \mathbf{R} \mathbf{m}_{n^*} \quad (25)$$

$$+ \mathbf{A}_{n^*}^\dagger \mathbf{A}_{n^*}^H \mathbf{C}_d^{-1} \mathbf{A}_{n^*} \mathbf{m}^{**} + \lambda \mathbf{A}_{n^*}^\dagger \mathbf{R}^T \mathbf{R} \mathbf{m}_{m0} . \quad (26)$$

with

$$\mathbf{A}^\dagger \mathbf{A}^H \mathbf{C}_d^{-1} \mathbf{A} + \lambda \mathbf{A}^\dagger \mathbf{R}^T \mathbf{R} = \mathbf{A}^\dagger (\mathbf{A}^H \mathbf{C}_d^{-1} \mathbf{A} + \lambda \mathbf{R}^T \mathbf{R}) = \mathbf{I} , \quad (27)$$

we define the non linear model Resolution matrix of the n 'th model iterate as

$$\mathbf{R}_{n^*}^M := \mathbf{A}_{n^*}^\dagger \mathbf{A}_{n^*}^H \mathbf{C}_d^{-1} \mathbf{A}_{n^*} . \quad (28)$$

With the identities of (27) and the definition (28) we can express the final estimate in terms of resolution matrices:

$$\mathbf{m}^* = \mathbf{R}_{n^*}^M \mathbf{m}^{**} + (\mathbf{I} - \mathbf{R}_{n^*}^M) \mathbf{m}_{m0} . \quad (29)$$

In practice, we compute the Resolution matrix for the OE, and not for the last but one model estimate ($\mathbf{R}_{n^*}^M \approx \mathbf{R}_{m^*}^M$).

Figure 4 shows the exemplary Sensitivity and Resolution measures. The Sensitivity distribution just reveals the influence of each model parameter on the data points, mostly dominated by the radial distance from the electrodes, except for very low conductive parts of the model, whereas the Resolution maps the true dependency of the non linear optimal estimation taking both, regularization and data quality into account. However, the Sensitivity, which is easier to compute can be taken as a first qualitative impression on the resolution capabilities of the EIT method, at least for the data acquisition scheme and model parameters show in this example. For a qualitative interpretation, we use the model parameters with normalized Resolution entries $> 10^{-4}$.

Incorporating resolution information within EIT visualization

In order to use the model resolution as indicator for good and poor resolved structures, we decided use it as weighting coefficient for the OEs image intensity. The images transparency (alphamap) maps color intensity as factors within the interval $f \in [0, 1]$. Thus, we transform the resolution matrix dynamic range, above a certain threshold $a \in \mathbb{R}_{>0}$, into this interval by using the weighting coefficients

$$w(a) = \max \left\{ \frac{\log_{10} \left(\frac{\text{diag}\{\mathbf{R}\}}{\max\{\text{diag}\{\mathbf{R}\}\}} \right) + a}{a}, 0 \right\} . \quad (30)$$

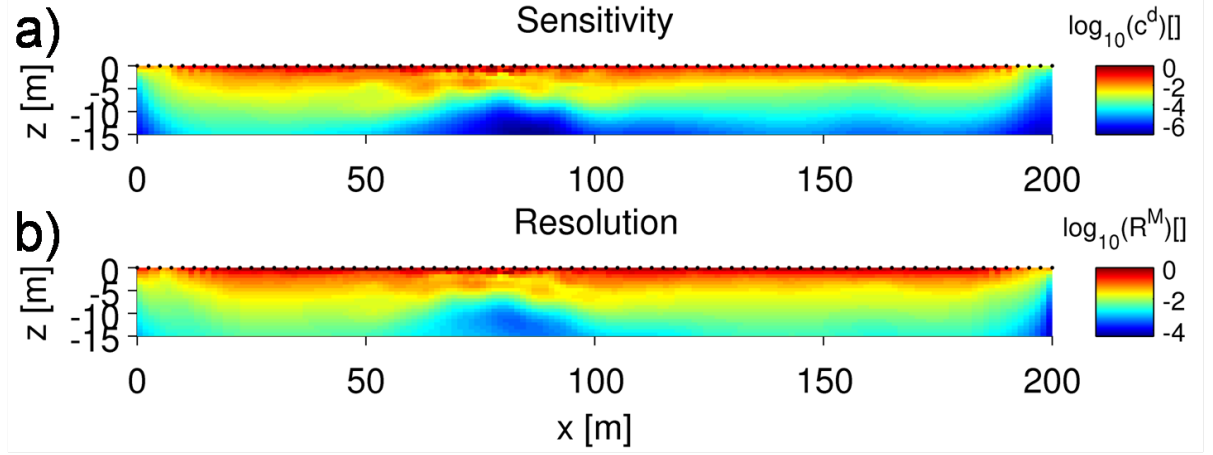


Figure 4: Normalized Sensitivity (a) and Resolution (b) measures computed for the OE given in Figure 3 using equations (19) and (28). The diagonal entries of the Resolution matrix (b) are obtained using a regularization parameter of $\lambda = 28.85$ and scaled through the maximum value of $\max \{diag \{ \mathbf{R}_M \} \} = 0.76$. The red and yellow parts of the Sensitivity and Resolution distribution show good resolved model parameters whereas turquoise and blue parts reveal poor resolved areas.

The heuristic threshold a should reflect a realistic means of resolution measure.

Figure 5 shows the transparency weighting factor for a resolution threshold of 4 decades mapping all values above $> 10^{-4}$ logarithmically onto the transparency table ($a = 4 \approx .1\%$). Red orange and yellow values are corresponding to good resolved regions whereas green and blue values reflect regions of the model with less to poor resolution. All model cells with resolution values below $.1\%$ of the maximal resolution are treated to have no influence on the data and mapped as dark blue.

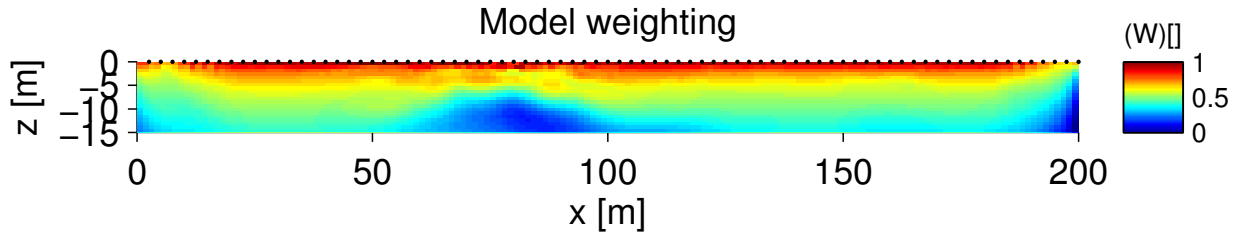


Figure 5: Resolution weighting coefficient according to Equation 30 using the resolution tradeoff of 4 decades ($a = 4$).

Figure 6 shows basically the same optimal estimates as Figure 3, but we included the resolution information as transparency weighting coefficient transformation within the estimated model parameters.

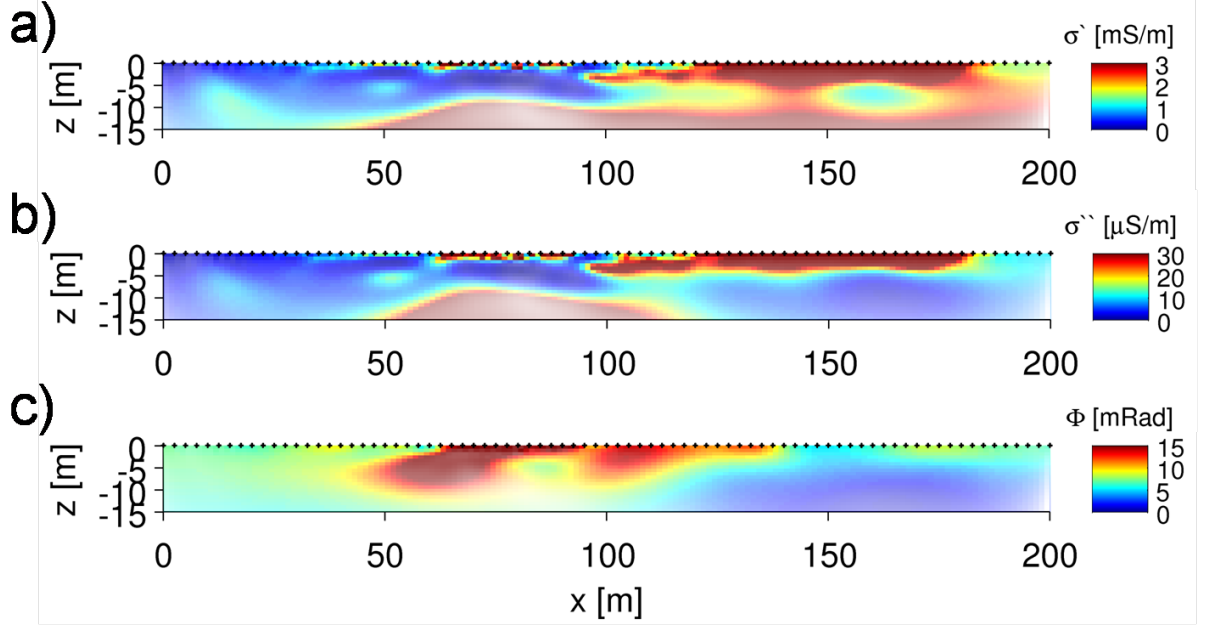


Figure 6: Weighted optimal estimation of complex conductivity (a & b) and phase (c) images using resolution tradeoff of 4 decades ($a = 4 \approx .1\%$) with transparency mapping according to Equation 30.

Uncertainty determination for the field example

Following Tarantola [1987], the a posteriori model covariance matrix \mathbf{C}_{m^*} , describing the covariance of the OE given the prior model covariance matrix \mathbf{C}_{m0} , is given by

$$\mathbf{C}_{m^*} = (\mathbf{A}_{m^*}^H \mathbf{C}_d^{-1} \mathbf{A}_{m^*} + \mathbf{C}_{m0}^{-1})^{-1}. \quad (31)$$

For the current investigations we identify the regularization term in equation (1) with the inverse prior model covariance matrix (i.e., considering λ as an inverse measure of a priori model variance magnitude, see for instance Alumbaugh [2000]) and compute the covariance of the final estimate like

$$\mathbf{C}(\mathbf{m}^*, \lambda) = (\mathbf{A}_{m^*}^H \mathbf{C}_d^{-1} \mathbf{A}_{m^*} + \lambda \mathbf{R}^T \mathbf{R})^{-1}. \quad (32)$$

Interestingly, since the matrix product (\bar{A} denotes the complex conjugate)

$$\mathbf{A}^H \mathbf{A} = \bar{A}_{ji} A_{ij}, \quad (33)$$

is real valued, $\mathbf{C}(\mathbf{m}^*, \lambda)$ is a real valued and symmetric matrix. This has the far-reaching consequence, that not only the variance of the complex valued model estimate, which is computed from the main diagonal, but also the covariance of the complex model estimate is real valued and is likewise for log magnitude and phase.

Figure 7 displays the uncertainty for the complex valued \mathbf{m}^* with respect to the inverse measure of the a priori variance magnitude, i.e. our regularization amplitude. Interestingly, it also reflects the Sensitivity pattern, which is in perfect agreement to the mathematical concept of error propagation, i.e. low uncertainty in good resolved regions and higher uncertainty in poor resolved areas of the imaging result. The values of $\approx 5\%$ are relatively close to the surface electrodes, and the standard deviation of the conductivity for both, real and imaginary part of the model parameters is below 10% for depths of ≈ 5 m.

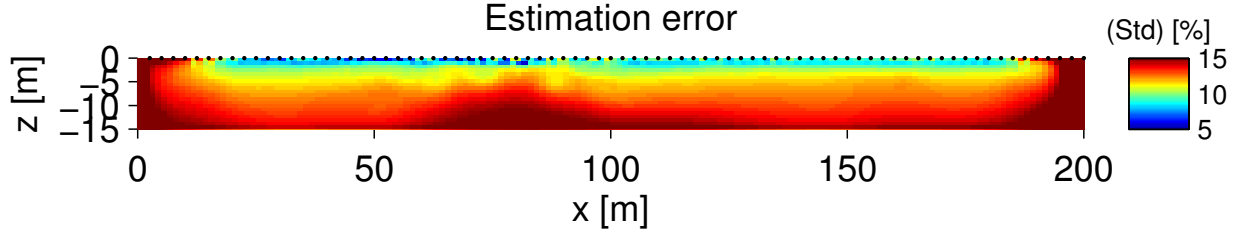


Figure 7: Standard deviation of the non linear a posteriori model based on linear uncertainty computation using equation (32) with a regularization parameter of $\lambda = 28.85$. Blue and green parts reflect low uncertainty with a standard deviation of less than 5%, whereas orange and red parts are due to higher uncertainty within the estimated parameter distribution.

Conclusion and discussion

The consequences of a real valued, relative and likewise treatment of the uncertainty for log magnitude and phase in terms of quantitative imaging of soil contaminants suggest, that the contaminated area, associated with increase in conductivity and phase values between 50m and 200m profile length, can be quantified within reasonable error bars (Figures 3 and 7). The standard error is $< 10\%$ for this particular depth range (Figure 7). For the very surface near region, we can recognize a standard deviation of $\approx 5\%$, which is reasonable for the data quality and error treatment within the imaging process. From the resolution measures (Figure 4) we can see, that the subsurface admittance is good resolved by the EIT method for at least 5 – 10m below surface. The small variances and high resolution measures are also reflected by the good agreement to the EC logs, which can be observed in this depth range (Figure 3). The resolution measures, which were incorporated within the subsurface images of complex conductivity and phase values via transparency weighting (Figure 6), provide a new and quantitative measure of resolution for EIT results. Last not least, we combine the resolution measures and the uncertainty computation within Figure 8 to combine both concepts for enhanced imaging methodology using optimized data acquisition strategies.

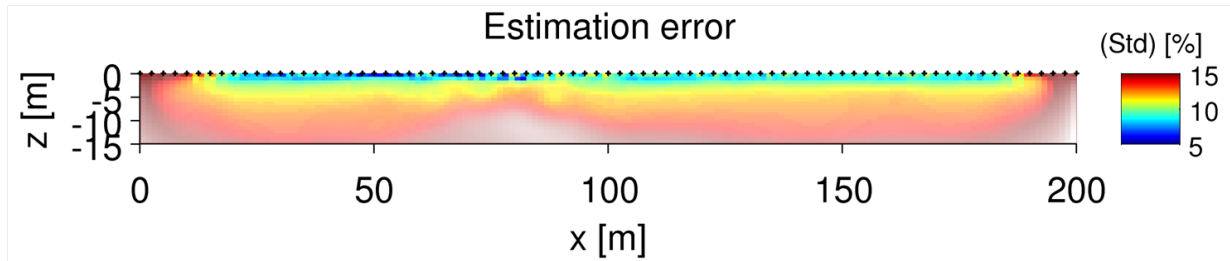


Figure 8: Weighted standard error of the OE taking resolution means into account.

References

- Alumbaugh, D. L., 2000. Linearized and nonlinear parameter variance estimation for two-dimensional electromagnetic induction inversion, *Inverse Problems*, **16**(5), 1323.
- Binley, A. & Kemna, A., 2005. *Hydrogeophysics*, chap. DC resistivity and induced polarization methods, pp. 129–156, Springer.
- Cassiani, G., Kemna, A., Villa, A., & Zimmermann, E., 2009. Spectral induced polarization for the characterization of free-phase hydrocarbon contamination of sediments with low clay content, *Near Surface Geophysics*, **7**, 547–562.
- Constable, S. C., Parker, R. L., & Constable, C. G., 1987. Occam’s inversion: a practical algorithm for generating smooth models from em sounding data, *Geophysics*, **52**, 289–300.
- Flores Orozco, A., Williams, K. H., Long, P. E., Hubbard, S. S., & Kemna, A., 2011. Using complex resistivity imaging to infer biogeochemical processes associated with bioremediation of an uranium-contaminated aquifer, *J. Geophys. Res.*, **116**(G3), G03001–.
- Kemna, A., 2000. *Tomographic inversion of complex resistivity - theory and application*, Ph.D. thesis, Ruhr-Universität Bochum.
- Menke, W., 1984. *Geophysical data analysis: discrete inverse theory*, Academic Press inc.
- Slater, L., Binley, A., Daily, W., & Johnson, R., 2000. Cross-hole electrical imaging of a controlled saline tracer injection, *Journal of Applied Geophysics*, **44**(2-3), 85 – 102.
- Tarantola, A., 1987. *Inverse problem theory, methods for data fitting and model parameter estimation*, Elsevier, Amsterdam.
- Tikhonov, A. N. & Arsenin, V. A., 1977. *Solution of Ill-posed Problems*, Winston & Sons, Washington.