Estimation of information content and efficiency for different data sets and inversion schemes using the generalized singular value decomposition

Th. Günther†, S. Friedel* & K. Spitzer† (†TU Bergakademie Freiberg, *ETH Zürich)

Abstract

For both experimental design and interpretation of inversion results it is useful to analyze resolution power and model uncertainties. The model resolution matrix can significantly help to understand the quality of the inversion process. On this basis, it is possible to find measures for information content and efficiency of data sets depending on the noise level. On the other hand, it is possible to investigate how different inversion and regularization schemes affect resolution power.
The singular value decomposition (svd) can be used to calculate generalized inverse matrices, which lead to resolution matrices for truncated svd (tsvd) inversion and Tikhonov regularization. With some additional effort, a generalized svd (gsvd) can also be applied to arbitrary model restrictions like smoothness constraints.
For reasons of comparability, a one step linearized inversion scheme is applied. Whereas tsvd and Tikhonov regularization rather produce noise artifacts, smoothness constraints and weighting by coverage tend to produce artifacts near the boundaries.
As far as the comparison of data sets is concerned, the highest information contents are achieved using pole-dipole and dipole-dipole configurations, the lowest ones using pole-pole and Wenner type arrays. In all cases enlarged dipole lengths for large separations and augmented data sets using so called circulated data significantly increase the information content.

1 Theory

Inversion and resolution

Most inverse problems arising in geophysics are non-unique and ill-posed. The methodology is to find a model $\mathbf{m}$ by minimizing some model constraint functional

$$\Phi_m = \|\mathbf{C}(\mathbf{m} - \mathbf{m}^0)\|_2^2 \quad ,$$

while fitting the model's forward response $\mathbf{f}(\mathbf{m})$ to the data $\mathbf{d}$, weighted by their standard deviations σ, in a least squares sense

$$\Phi_d = \sum_i^{n_d} \left(\frac{d_i - f_i(\mathbf{m})}{\sigma_i} \right)^2 = \|\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}))\|_2^2 = \Phi_d^* \quad . \tag{1}$$

$\mathbf{C}$ and $\mathbf{D}$ are model and data weighting matrices, respectively and $\mathbf{m}^0$ is a starting model. A Langrangian parameter λ is used to combine model and data functional

$\Phi = \Phi_d + \lambda\Phi_m$. Generally, the target value Φ_d^* equals the number of data n_d, so that the data are fitted within standard deviation.
The application of an Gauss-Newton method minimizing Φ leads to iteratively updating $\mathbf{m}^{k+1} = \mathbf{m}^k + \Delta\mathbf{m}^k$. In every iteration k, the well-known regularized normal equations (Tarantola & Valette, 1982) are solved as follows:

$$((\mathbf{DS})^T\mathbf{DS} + \lambda\mathbf{C}^T\mathbf{C}) \cdot \Delta\mathbf{m}^k = (\mathbf{DS})^T\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}^k)) - \lambda\mathbf{C}^T\mathbf{C}(\mathbf{m}^k - \mathbf{m}^0) \quad , \quad (2)$$

where $\mathbf{S}$ is the Jacobian or sensitivity matrix with the elements

$$S_{ij} = \frac{\partial f_i(\mathbf{m})}{\partial m_j} \quad .$$

This equation can be solved efficiently using a variant of the CGLS algorithm (Günther & Donner, 2001), which avoids assembling the left-hand side and its inverse. With $\hat{\mathbf{S}} = \mathbf{DS}$ equation (2) can be written as

$$\Delta\mathbf{m}^k = \hat{\mathbf{S}}^\dagger\mathbf{D}(\mathbf{d} - \mathbf{f}(\mathbf{m}^k)) - \mathbf{C}^\dagger\mathbf{C}(\mathbf{m}^k - \mathbf{m}^0) \quad \text{with} \quad (3)$$

$$\hat{\mathbf{S}}^\dagger = (\hat{\mathbf{S}}^T\hat{\mathbf{S}} + \lambda\mathbf{C}^T\mathbf{C})^{-1}\hat{\mathbf{S}}^T \quad \text{and}$$

$$\mathbf{C}^\dagger = \lambda(\hat{\mathbf{S}}^T\hat{\mathbf{S}} + \lambda\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T \quad \Rightarrow \quad \hat{\mathbf{S}}^\dagger\hat{\mathbf{S}} + \mathbf{C}^\dagger\mathbf{C} = \mathbf{I}.$$

The regularization parameter λ has to be chosen to satisfy $\Phi_d = n_d$ by appropriate methods. Assuming that $\mathbf{m}^k$ is already close to the true model $\mathbf{m}^{true}$, linearization yields

$$\mathbf{d} = \mathbf{f}(\mathbf{m}^{true}) + \mathbf{n} = \mathbf{f}(\mathbf{m}^k) + \mathbf{S}(\mathbf{m}^{true} - \mathbf{m}^k) + \mathbf{n} \quad , \quad (4)$$

where $\mathbf{n}$ denotes noise. Inserting (4) into (3) leads to the resolution equation

$$\begin{aligned} \mathbf{m}^{k+1} &= \mathbf{m}^k + \hat{\mathbf{S}}^\dagger\hat{\mathbf{S}}(\mathbf{m}^{true} - \mathbf{m}^k) + \hat{\mathbf{S}}^\dagger\mathbf{Dn} - \mathbf{C}^\dagger\mathbf{C}(\mathbf{m}^k - \mathbf{m}^0) \\ &= \mathbf{R}^M\mathbf{m}^{true} + (\mathbf{I} - \mathbf{R}^M)\mathbf{m}^0 + \hat{\mathbf{S}}^\dagger\mathbf{Dn} = \mathbf{m}^{est} \end{aligned} \quad (5)$$

The model estimate $\mathbf{m}^{est}$ is thus constructed by the true and the starting model and contaminated by noise effects. The resolution matrix $\mathbf{R}^M = \hat{\mathbf{S}}^\dagger\hat{\mathbf{S}}$ describes how information is mapped from reality into $\mathbf{m}^{est}$. Hence, regularization can be interpreted as trade-off between good resolution and small noise artifacts.
The sum of the main diagonal elements $\mathbf{R}^M_{ii}$ is called information content (IC), dividing it by the number of data an information efficiency (IE) can be defined:

$$IC = \sum_i \mathbf{R}^M_{ii} \quad \text{and} \quad IE = \frac{IC}{n_d} \quad .$$

Resolution in terms of svd and gsvd

The singular value decomposition (svd) can be used to find two orthonormal matrices $\mathbf{U}$ and $\mathbf{V}$ transforming $\hat{\mathbf{S}}$ into diagonal form

$$\mathbf{U}^T\hat{\mathbf{S}}\mathbf{V} = \mathbf{\Sigma} = \mathrm{diag}(s_i) \Rightarrow \hat{\mathbf{S}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad . \quad (6)$$

Omitting the zero singular values $s_i \; \forall i > r = \mathrm{rank}(\hat{\mathbf{S}})$ and introducing filter factors f_i leads to a generalized inverse

$$\hat{\mathbf{S}}^{\dagger} = \mathbf{V}_r \mathrm{diag}\left(\frac{f_i}{s_i}\right) \mathbf{U}_r^T \quad , \tag{7}$$

valid for the truncated svd (Friedel, 2003), where $f_i = 1 \, \forall \, i \leq p$ (cutting value) else 0, and Tikhonov inversion with $\mathbf{C} = \mathbf{I}$ and $f_i = \frac{s_i^2}{s_i^2+\lambda}$.
When $\mathbf{C} \neq \mathbf{I}$, the generalized svd can be applied to diagonalize both $\hat{\mathbf{S}}$ and $\mathbf{C}$ with two orthonormal matrices $\mathbf{U}$ and $\mathbf{W}$ and an invertible matrix $\mathbf{V}$

$$\mathbf{U}^T \hat{\mathbf{S}} \mathbf{V} = \mathbf{\Sigma} = \mathrm{diag}(s_i) \quad \text{and} \quad \mathbf{W}^T \mathbf{C} \mathbf{V} = \mathbf{M} = \mathrm{diag}(\mu_i) \tag{8}$$

It can be shown (Golub & van Loan, 1996), that for a quadratic stabilizer $\mathbf{C}$ the generalized inverse can be written as

$$\hat{\mathbf{S}}^{\dagger} = \mathbf{V}_r \mathrm{diag}\left(\frac{f_i}{s_i}\right) \mathbf{U}_r^T \quad , \tag{9}$$

computing the filter factors by the generalized singular values $\gamma_i = s_i/\mu_i$ in the same way as in svd ($f_i = \frac{\gamma_i^2}{\gamma_i^2+\lambda}$). Then the model resolution reads

$$\mathbf{R}^M = \mathbf{V} \mathrm{diag}(\frac{f_i}{s_i}) \mathbf{U}_r^T \mathbf{U}_r \mathrm{diag}(s_i) \mathbf{V}^{-1} = \mathbf{V} \mathrm{diag}(f_i) \mathbf{V}^{-1} \quad . \tag{10}$$

The information content is computed by $IC = \sum_i R_{ii}^M = \sum_i f_i$ without inverting $\mathbf{V}$.

2 Numerical setup

Model Parameterization

Figure 1 shows the synthetic model investigated in this study. It is equally discretized in x from -1 to 42m and in z from 0 to 6m in block sizes of 1m×1m and consists of bodies with resistivities of 50 and 200 Ωm within a homogeneous background of 100 Ωm .

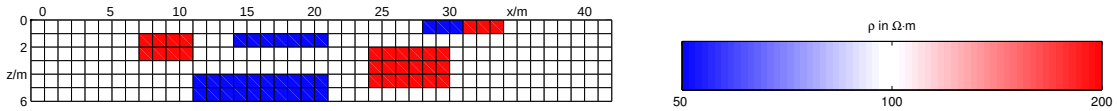


Figure 1: Model parameterization and synthetic model consisting of 6 bodies with 50 Ωm and 100 Ωm , respectively, in a homogeneous half-space of 100 Ωm , the colorscale holds for subsequent figures

Data and model parameters are the logarithms of the apparent resistivities and the cell resistivities, respectively.

Data and noise

The following numerical investigations employ 42 equally spaced electrodes with the positions x=0,1,...,41m. Results of the simulations can easily be transformed to larger problems by a scale factor, which changes the configuration factor k and therefore the error level $\delta\rho^a$. Due to the use of logarithms the standard deviation reads $\sigma_i = \log(1 + \delta\rho_i^a/\rho_i^a)$.
Several widely used electrode configurations are investigated: pole-pole, pole-dipole, dipole-dipole, Schlumberger and Wenner (in form of α, β and γ arrays). Since the small dipoles of pole-dipole, dipole-dipole and Schlumberger data are associated with a large k for large separations n, they are sensitive to noise. Therefore, alternative sets are created with enlarged dipole lengths $\overline{MN} = \mathrm{mod}(n + 3, 4) \cdot a$. While information is lost by reducing the data set in such a way, additional data can be obtained by measuring so-called circulated data as presented by Friedel (2000).

Dataset	tag	n	data	max(k)	min/max ρ_a
Pole-pole	c–p	1-8	300	50.3	79.3/143.4
Pole-dipole	c–p.p	1-8	292	452.4	50.0/211.2
(enlarged dipoles)	c–p:p	1-15	280	447.7	50.0/210.8
(+circulated data)	c–p:p+c	1-10	289	449.2	50.0/210.8
Dipole-dipole	c.c–p.p	1-8	284	2261.9	41.6/238.9
(enlarged dipoles)	c:c–p:p	1-15	265	1287.1	41.6/238.9
(+circulated data)	c:c–p:p+c	1-10	271	1319.5	41.6/238.9
Schlumberger	c–p.p–c	1-10	300	345.6	61.1/181.8
(enlarged dipoles)	c–p:p–c	1-15	285	1287.1	73.1/167.6
Wenner-α	c–p–p–c	1-13	273	81.7	63.1/165.1
Wenner-β	c–p–c–p	1-13	273	245.0	41.6/238.9
Wenner-γ	c–c–p–p	1-13	273	122.5	41.9/281.2

Table 1: Definition of data sets. c stands for current, p for potential electrode; "–" denotes the separation $n \cdot a$, "." denotes dipole length a and ":" denotes enlarged dipole length, +c are additional circulated data

Table 1 shows the electrode arrangements investigated. To guarantee a comparability, the maximum separation is chosen such that the number of data are approximately equal (269-300). Note, that the synthetic data of pole-dipole and dipole-dipole type generally show the largest data anomalies.
The noise $\delta\rho^a$ consists of two parts. The first part is a relative deviation ε in the order of several percent and the second part results from the minimum voltage that conventional devices are able to distinguish. This voltage usually lies at several hundred microvolts. Thus, the following relation holds

$$\delta\rho_i^a = \varepsilon \cdot \rho_i^a + \frac{U_{min}}{I} k_i \quad .$$

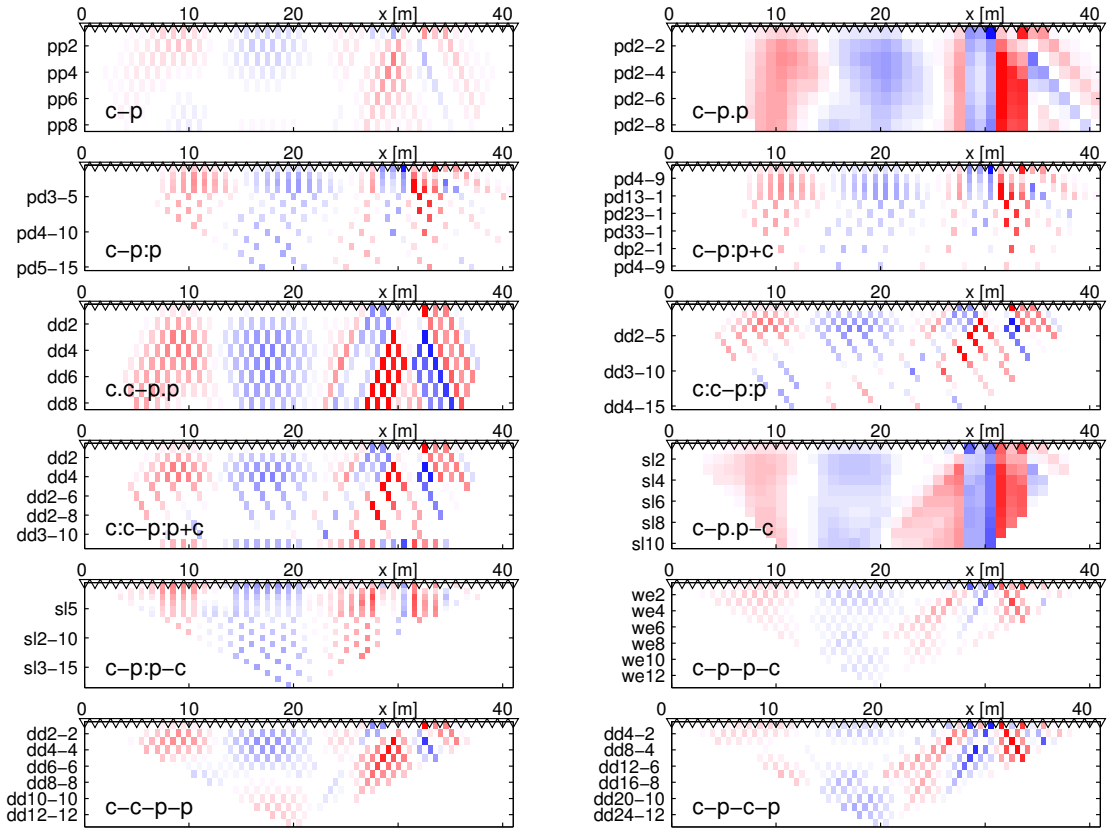


Figure 2: Synthetic data sets for the investigated data

Consequently, configurations with large k (e.g. dipole-dipole) obtain large errors, which significantly affect resolution properties.

Simulation

Following the scheme of Friedel (2003) we assume the sensitivity to be a first order approximation based on a homogeneous half-space. It is calculated by numerical integration and used to produce the synthetic data, so that the nonlinear problem is transformed into a linear one and can thus be solved in one iteration step. The linearization is valid for small contrasts and thus the resolution becomes independent of the individual model.
The synthetic data, displayed in Figure 2, are contaminated with Gaussian noise of standard deviation σ_i. The parameter λ is determined in such a way that $\Phi_d = n_d$. This procedure is repeated 10 times to get independent of the noise realization and the mean λ is used to compute model and resolution matrix from the gsvd.

3 Numerical results

Comparison of inversion schemes

For a fixed data set (c:c–p:p) different inversion or regularization schemes are compared with respect to information content. Besides tsvd and Tikhonov regularization ($\mathbf{C} = \mathbf{I}$), a 2^{nd} order smoothness constraint matrix is investigated representing a discrete differential operator with Dirichlet boundary conditions. Additionally, a diagonal weighting matrix is applied using the coverage (or cumulative sensitivity) as the sum of the absolute sensitivity values for each data:

$$\mathbf{C} = \mathrm{diag}(\sqrt{\mathbf{c}}) \quad \text{with} \quad c_j = \sum_{i=1}^{n} \|S_{ij}\| \quad .$$

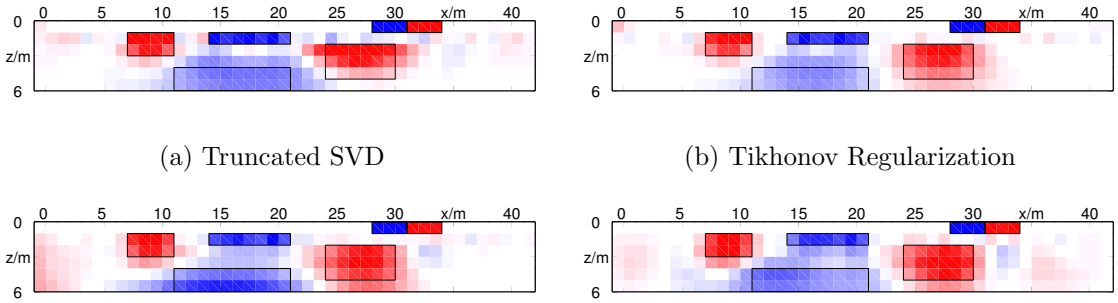
Inversion scheme	regul.	IC	deviation
tsvd	$p = 69$	69.0	17.9%
Tikhonov	$\lambda = 39.3$	71.5	16.5%
weighting by coverage	$\lambda = 27.7$	70.0	16.0%
smoothness constraints	$\lambda = 58.1$	67.3	18.5%

Table 2: Comparison of different inversion methods, data set c:c–p:p, ε=1%, $U_{min} = 100\,\mu V$, I=100 mA, a=1m, deviation is rms of estimated and synthetic model

Table 2 shows the results of the numerical simulation for low data noise. All schemes possess comparable information contents of around 70. Higher values correspond to lower deviations between synthetic and estimated model.
In Figure 3 the inversion results are displayed. It is to be seen, that tsvd and Tikhonov produce noise artifacts resulting from the third term of the right hand side of eq. (5). On the other hand, smoothness constraints and weighting by coverage produce constraint artifacts near the boundaries.

Comparison of data sets

In the following, the data denoted above are compared regarding information content and efficiency. For all simulations the second order smoothness constraints are applied.
Figure 4 shows the individual inversion results. It can be seen, that all data sets are able to locate the anomalous bodies, but with different quality.
Table 3 shows, that the resolution properties correspond to the model quality. The highest information contents are achieved using pole-dipole and dipole-dipole type sets, the lowest using pole-pole and Wenner. In all three cases the dipole enlargement significantly increases the IC values. Also, the added circulated data increase the IC for dipole-dipole type.



(a) Truncated SVD

(b) Tikhonov Regularization

(c) Weighting by coverage

(d) Smoothness constraints

Figure 3: Inversion results for different inversion schemes using data set c:c–p:p

data set	IC	IE	deviation	λ	maxerr
c–p	61.7	20.6	21.9	33.64	1.1
c–p.p	70.8	24.2	18.9	39.63	1.7
c–p:p	73.4	26.2	16.4	27.35	1.6
c–p:p+c	72.4	25.1	18.7	37.23	1.6
c.c–p.p	71.7	25.2	18.8	47.81	5.0
c:c–p:p	75.3	28.4	15.2	35.11	2.6
c:c–p:p+c	78.8	29.1	16.0	36.23	2.6
c–p.p–c	65.6	21.9	20.0	34.70	1.6
c–p:p–c	81.5	28.6	15.4	29.92	1.5
c–p–p–c	63.6	23.3	18.7	25.39	1.1
c–c–p–p	69.2	25.3	15.5	31.97	1.2
c–p–c–p	64.7	23.7	20.7	52.51	1.1

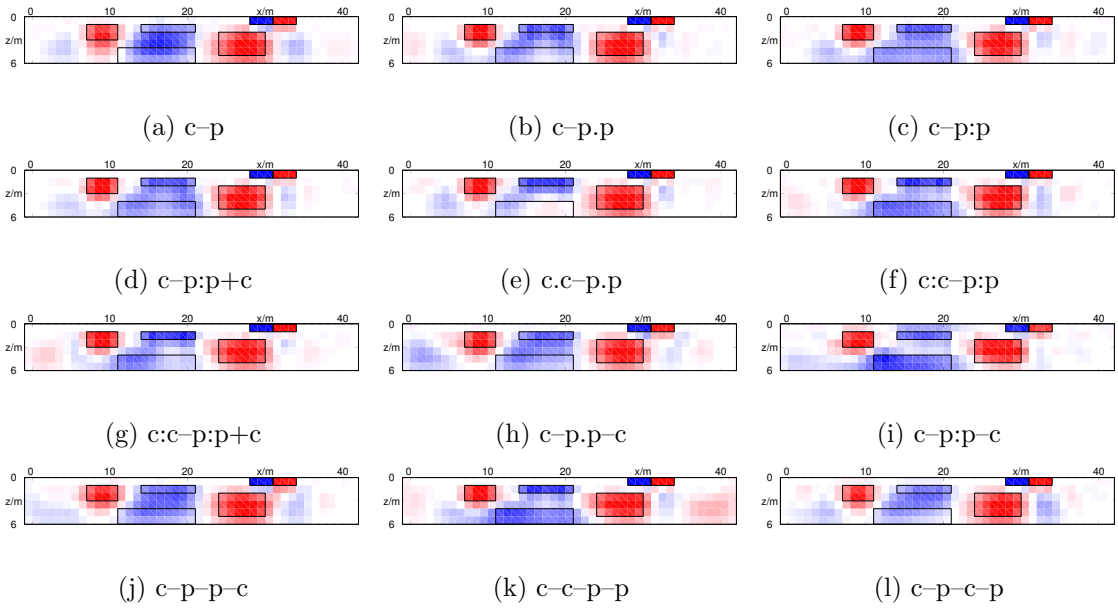
Table 3: Comparison of different data sets corresponding to Figure 4

Most information is obtained using the c–p:p–c set, whereas the c:c–p:p+c set yields the largest efficiency. The IE values, ranging between 20 and 30%, show a quite large variability.
It can clearly be seen, that good model estimates denoted by small deviations correspond to large information contents and vice versa. Thus, the defined information content proves to be a reliable measure of inversion quality.

Resolution radius

The individual model resolutions for each cell i can be used to estimate a model resolution radius r_i. Assuming the model resolution to be locally constant, it is defined by the equivalent circle of model resolution 1, yielding for the cell dimensions Δx_i and Δz_i

$$r_i = \sqrt{\frac{\Delta x_i \Delta z_i}{\pi R_{ii}^M}} \quad .$$



(a) c–p (b) c–p.p (c) c–p:p

(d) c–p:p+c (e) c.c–p.p (f) c:c–p:p

(g) c:c–p:p+c (h) c–p.p–c (i) c–p:p–c

(j) c–p–p–c (k) c–c–p–p (l) c–p–c–p

Figure 4: Inversion results for different data sets for ε=1%, U_{min}=100 μV and I=100 mA

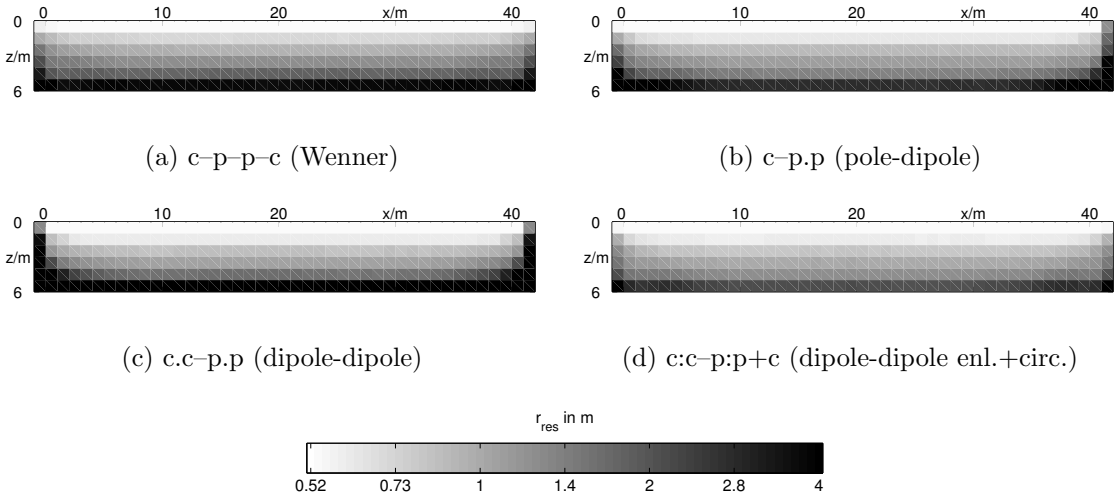
Figure 5 shows that for all selected configurations the resolution radius near the surface is about equal (0.5 m) and rapidly increases with depth to 3-4 m at z=6 m. However, the decrease of resolution is different for the individual data sets. The Wenner array (a) has moderate resolution at depth while the dipole-dipole (c) array resolves best at medium depth. This is strongly improved by increasing dipole lengths and adding circulated data (d).

Effect of increased noise

Table 4 shows the resolution properties for an assumed noise of ε=2% and voltage resolution U_{min}=1 mV and I=100 mA current, or, equivalently U_{min}=100 μV and I=10 mA driving current.
As expected, higher noise leads to worse inversion results and lower information content (see columns 2 and 4) for all data sets. Again, enlarged dipole-dipole and Schlumberger configurations yield the best results. As the noise level rises, pole-dipole arrays and the Wenner-β configuration perform better.

Further investigations

The question arises, how a variable model parameterization can affect model resolution. According to the increasing resolution radii, the horizontal grid lines are defined by the values (z=0 0.4 1 1.8 2.8 4.2 6). Furthermore, weakly covered cells, which usually occur at the boundaries of the modeling domain, can be either neglected in the inversion process or combined to larger cells.



(a) c–p–p–c (Wenner)

(b) c–p.p (pole-dipole)

(c) c.c–p.p (dipole-dipole)

(d) c:c–p:p+c (dipole-dipole enl.+circ.)

Figure 5: Resolution radii for data sets c–p–p–c, c–p.p, c.c–p.p and c:c-p:p-c

Although the quality of the inversion results is slightly improved, the information content stays surprisingly constant during all steps of parameter change. This supports the idea, that for a given (noisy) data set a maximum information content exists, which cannot be enhanced by a finer discretization.
If the resistivity contrasts of the anomalous bodies are increased, the resolution properties improve due to the bigger data anomalies. The essential property is the signal-to-noise ratio, which determines regularization strength and therefore the resolving power.

4 Conclusions

It has been shown, how the resolution analysis can be applied to arbitrary inversion and regularization schemes. On the basis of a linearized Gauss-Newton inversion scheme several regularization techniques were compared. Additionally, different data sets were investigated concerning information content and efficiency. Compared with classical configurations, data sets with increased dipole lengths yield superior results.
Amongst the investigated data sets, dipole-dipole and Schlumberger configuration sets with increased dipole lengths obtained the largest resolution qualities. Noise and resistivity contrasts are important factors with respect to resolution.
It can be seen, that the presented quantity information content is an appropriate measure for the reliability of the model. Furthermore, it can be applied to improve the experimental design. However, it has to be proved if this applies to the full nonlinear non-linear case as well. The crucial point in nonlinear inversion is the choice of the regularization parameter, which strongly influences resolution properties and thus has to be chosen carefully.

data set	IC	IE	deviation	λ	maxerr
c–p	49.3	16.4	23.0	39.63	2.5
c–p.p	56.9	19.5	22.7	43.70	8.6
c–p:p	58.6	20.9	20.9	35.25	7.7
c–p:p+c	60.1	20.8	22.2	34.03	8.0
c.c–p.p	59.0	20.8	22.8	35.09	42.0
c:c–p:p	62.8	23.7	20.4	20.56	17.9
c:c–p:p+c	61.3	22.6	21.8	39.79	18.4
c–p.p–c	54.0	18.0	22.8	32.35	7.7
c–p:p–c	64.9	22.8	18.3	22.32	7.1
c–p–p–c	50.2	18.4	23.7	41.37	2.9
c–c–p–p	60.7	22.3	18.2	21.89	4.5
c–p–c–p	56.5	20.7	22.3	41.86	3.4

Table 4: Comparison of different data sets for ε=2%, $U_{min} = 1000\,\mu$V at $I = 100\,$mA

Acknowledgements

This work was partly funded by the Deutsche Forschungsgemeinschaft DFG (Ja 590/18-1 & 2).

References

Friedel, S. (2000). *Über die Abbildungseigenschaften der geoelektrischen Impedanztomographie unter Berücksichtigung von endlicher Anzahl und endlicher Genauigkeit der Messdaten.* Dissertation. Shaker, Aachen. (Universität Leipzig)

Friedel, S. (2003). Resolution, stability and effiency of resistivity tomography estimated from a generalized inverse approach. *Geophys. J. Int.*, *153*, 305-316.

Golub, G. H., & van Loan, C. (1996). *Matrix computations.* The Johns Hopkins University Press.

Günther, T., & Donner, F. (2001). Der Multi-Lösungs-Algorithmus CGLSPAR und seine Anwendung auf die 3D-Inversion geoelektrischer Daten. In A. Junge & A. Hördt (Eds.), *Tagungsband Elektromagnetische Tiefenforschung.* Burg Ludwigstein.

Tarantola, A., & Valette, B. (1982). Generalized nonliear inverse problems solved using the least squares criterion. *Reviews of Geophysics and Space Physics*, *20*(2), 219-232.