

# Das inverse Problem der Geoelektrik: Ein nichtlinearer multikriterieller Lösungsansatz mit genetischen Algorithmen

Christoph Schwarzbach und Ralph-Uwe Börner

Institut für Geophysik, TU Bergakademie Freiberg

## Einführung

Beim inversen Problem der Geoelektrik werden räumliche Verteilungen der elektrischen Leitfähigkeit  $\sigma(\mathbf{r})$  derart gesucht, dass die Diskrepanz zwischen gemessenen Daten  $\rho_a^{obs}$  und synthetisch erzeugten Daten  $\rho_a^{syn}(\sigma)$  unter ein messfehlerabhängiges Niveau gesenkt wird,

$$f_1(\sigma) = \|\rho_a^{obs} - \rho_a^{syn}(\sigma)\| < const. \quad (1)$$

Durch Einschränkungen des Modellraums wird die Vielzahl möglicher Lösungen reduziert. Die Verwendung eines FD-Vorwärtsoperators für die Berechnung der synthetischen Daten legt die Betrachtung von stückweise konstanten Leitfähigkeitsmodellen nahe. Mittels geeigneter Codierung der Modellparameter im genetischen Algorithmus kann der Wertevorrat von  $\sigma$  auf eine diskrete Menge begrenzt werden. Die Einführung einer zweiten Zielfunktion  $f_2(\sigma) = \|R(\sigma)\|$  ermöglicht weitere Einschränkungen der Modellklassen, z. B. auf „glatte“ Modelle. Wird außerdem auch  $f_1(\sigma) \rightarrow \min!$  gefordert, so kann das inverse Problem als nichtlineare multikriterielle Minimierungsaufgabe mit dem vektorwertigen Zielfunktional

$$\mathbf{f} : \sigma \mapsto \mathbf{f}(\sigma) = \begin{pmatrix} f_1(\sigma) \\ f_2(\sigma) \end{pmatrix} = \begin{pmatrix} \|\rho_a^{obs} - \rho_a^{syn}(\sigma)\| \\ \|R(\sigma)\| \end{pmatrix} \rightarrow \min! \quad (2)$$

formuliert werden. Im Allgemeinen sind neben dem Vorwärtsoperator, d. h. der Abbildung  $\sigma \mapsto \rho_a^{syn}(\sigma)$ , auch die Modellrestriktion  $R : \sigma \mapsto R(\sigma)$  und die verwendeten Normen nichtlineare Abbildungen.

Die Lösungsmenge dieser Optimierungsaufgabe wird als die Menge der Pareto-optimalen Lösungen (Pareto-Menge, Menge der vektoroptimalen Kompromisslösungen) bezeichnet. Für zwei Punkte  $\sigma$  und  $\sigma^*$  gelte

$$\left. \begin{array}{l} f_i(\sigma) \leq f_i(\sigma^*) \quad \text{für } i = 1, 2 \\ f_i(\sigma) < f_i(\sigma^*) \quad \text{für wenigstens ein } i \end{array} \right\} \text{ kurz: } \mathbf{f}(\sigma) < \mathbf{f}(\sigma^*). \quad (3)$$

Dann sagt man, der Punkt  $\sigma$  dominiert den Punkt  $\sigma^*$  (Abb. 1). Existiert kein  $\sigma \neq \sigma^*$  mit  $\mathbf{f}(\sigma) < \mathbf{f}(\sigma^*)$ , so heißt  $\sigma^*$  nicht-dominiertes Punkt. Alle nicht-dominierten Punkte bilden die Menge der Pareto-optimalen Lösungen.

Für die Lösung der multikriteriellen Minimierungsaufgabe werden die Komponenten der vektorwertigen Zielfunktion (2) häufig durch Linearkombination zu einer skalaren Zielfunktion

$$f_\lambda(\sigma) = f_1(\sigma) + \lambda f_2(\sigma) \rightarrow \min! \quad (4)$$

zusammengefasst. Für ein festes  $\lambda$  kann mit einer beliebigen Optimierungsmethode nur ein Punkt der Pareto-Menge gefunden werden. Die populationsbasierte Optimierungsstrategie der genetischen Algorithmen erlaubt es hingegen, vektorwertige Zielfunktionen direkt zu behandeln und die gesamte Pareto-Menge in einem Evolutionszyklus zu approximieren.

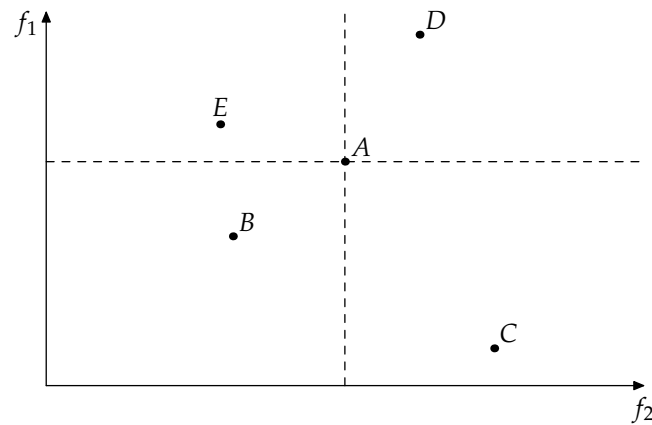


Abbildung 1: Sortierung nach Nicht-Dominanz: Punkt A dominiert Punkt D und wird von Punkt B dominiert. Die Punkte B, C, E gehören zur ersten, Punkt A zur zweiten und Punkt D zur dritten nicht-dominierten Menge (Pareto-Front).

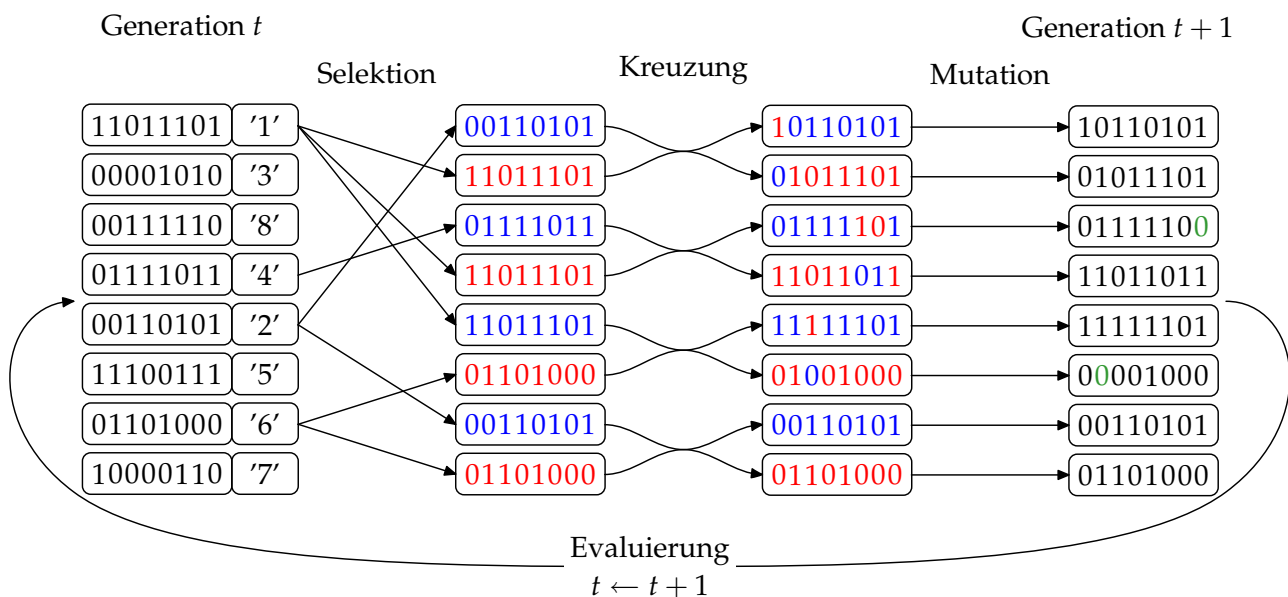


Abbildung 2: Schematischer Ablauf eines Evolutionsschritts mit binär codierten Strings. '1', '2', ... bezeichnet den Rang (die Fitness) eines Strings, der sich aus dem Vergleich aller Zielfunktionswerte ergibt.

## Genetischer Algorithmus

Genetische Algorithmen gehören zur Gruppe von stochastischen Optimierungsverfahren, die unter dem Begriff Evolutionsalgorithmen zusammengefasst werden. Sie übertragen Prozesse der biologischen Evolution, die darwinsche Idee vom Überleben der Bestangepassten, Mechanismen der Genetik und die Anpassung von Lebewesen an ihre Umwelt, auf künstliche Systeme. Genetische Algorithmen operieren auf einer Anzahl von codierten Repräsentanten der Punkte im Modellraum, die eine Population von Strings bilden. Den Strings wird nach Decodierung, Zielfunktionsauswertung und Vergleich aller Werte innerhalb der Population ein Fitnesswert zugeordnet. Durch die Anwendung der genetischen Operatoren Selektion, Kreuzung und Mutation wird eine Zahl neuer Strings erzeugt (Abb. 2). Diese werden wiederum decodiert, evaluiert und für die Zuordnung von Fitnesswerten verglichen. Aus den neuen und alten Strings wird eine neue Population zusammengestellt und der gesamte Prozess wiederholt. Die Evolution endet, wenn hinreichend gute Lösungen gefunden worden sind, eine definierte Anzahl an Zielfunktionsauswertungen oder Generationen erreicht ist oder die Suche stagniert.

Die Zuordnung von Fitnesswerten entsprechend der vektorwertigen Zielfunktion erfolgt in der ver-

wendeten Implementation eines genetischen Algorithmus NSGA-II (Deb et al., 2000) durch die Sortierung nach Nicht-Dominiertheit (Abb. 1). Zunächst werden alle nicht-dominierten Strings in der Population bestimmt. Ihnen wird der Rang '1' zugewiesen (erste nicht-dominierte Menge). Dieser Prozess wird für die Ränge '2' (zweite nicht-dominierte Menge), '3', etc. wiederholt, wobei dafür nur die verbleibenden Strings betrachtet werden, denen noch kein Rang zugeordnet wurde. Weiterhin wird den Strings mit gleichem Rang der Abstand untereinander, gemessen im Zielfunktions- oder Modellraum, als Hilfsfitness zugewiesen, um eine gleichmäßige Pareto-Front bzw. möglichst diverse Lösungen zu erzielen.

Auf der Grundlage von NSGA-II wurde ein genetischer Inversionsalgorithmus für die Geoelektrik entwickelt. Neben der Implementierung der problemspezifischen Zielfunktionen wurden einige Modifizierungen vorgenommen, um die Effektivität zu verbessern. Dies waren u. a. die Integration der gesteuerten Elitetechnik (Deb und Goel, 2000), der Einsatz von binärer Gray-Codierung und die Parallelisierung nach dem Inselmodell. Beim Inselmodell wird die globale Population in eine Anzahl von Teilpopulationen (Inselpopulationen) unterteilt. Jede Inselpopulation ist ihrem eigenen Evolutionsprozess unterworfen, tauscht jedoch in regelmäßigen Abständen Individuen mit anderen Inseln aus. Dadurch kann die langsame Konvergenz großer Populationen auf die schnellere Konvergenz kleiner Populationen zurückgeführt werden, wobei die große Gesamtpopulation eine hinreichende Abtastung des Modellraums gewährleistet.

## Zielfunktionen

Genetische Algorithmen erlauben eine große Flexibilität in der Wahl der Zielfunktionen, da sie z. B. keine Anforderungen bezüglich der Differenzierbarkeit stellen. Für die Datenanpassung wird die Familie  $f_1^p(\sigma)$ ,  $1 \leq p < \infty$ ,

$$f_1^p(\sigma) := \left\| \frac{\rho_a^{obs} - \rho_a^{syn}(\sigma)}{\sqrt{n-1} s^{obs}} \right\|_p = \left( \sum_{i=1}^n \left| \frac{\rho_{a,i}^{obs} - \rho_{a,i}^{syn}(\sigma)}{\sqrt{n-1} s_i^{obs}} \right|^p \right)^{1/p} \quad (5)$$

verwendet. Dabei bezeichnen  $n$  die Anzahl der Datenpunkte und  $s_i^{obs}$  die Standardabweichung von  $\rho_{a,i}^{obs}$ . Als Modellrestriktion  $f_2(\sigma)$  wird im folgenden Beispiel ein Maß für die Rauheit des Modells, die  $L_1$ -Norm der ersten Ableitung der Leitfähigkeit

$$f_2(\sigma) := \left\| \frac{\partial \sigma}{\partial x} \right\|_1 + \left\| \frac{\partial \sigma}{\partial z} \right\|_1 = \int_{\Omega} \left| \frac{\partial \sigma}{\partial x} \right| dx dz + \int_{\Omega} \left| \frac{\partial \sigma}{\partial z} \right| dx dz \quad (6)$$

eingesetzt, wobei  $\Omega \subset \mathbb{R}^2$  das für die Inversion betrachtete Modellgebiet bezeichnet.

## Vorwärtsoperator

Genetische Algorithmen erfordern eine Vielzahl an Zielfunktionsauswertungen. Deshalb wurde ein schneller 2D-Vorwärtsoperator entwickelt, der auf der Methode der Finiten Differenzen beruht (Dey und Morrison, 1979). Durch die Verwendung der Technik der Singularitätsbeseitigung (Lowry et al., 1989; Zhao und Yedlin, 1996), die Rücktransformation aus dem Wellenzahlbereich mit Hilfe von Gauss-Quadraturformeln und einen direkten Gleichungslöser, der die Matrixstruktur der rechten Seite des linearen Gleichungssystems aufgrund mehrerer Quellen ausnutzt, konnten sowohl Geschwindigkeit als auch Genauigkeit verbessert werden. Mit 1.6 Millionen Vorwärtsrechnungen für das folgende Beispiel und einer Gesamtlauzeit von etwa 6 Tagen auf einem 2.4 GHz-PC mit 1 GB RAM stellt die Inversion mit einem 2D-Vorwärtsoperator derzeit die Grenze des Machbaren dar.

## Anwendungsbeispiel

Als geeignetes Untersuchungsobjekt für die Anwendung des genetischen Inversionsalgorithmus wurde der Rothschnberger Stolln gewählt, welcher der Entwässerung des Freiburger Grubenreviers dient. Eine profilhafte Dipol-Dipol-Messung wurde quer zum Streichen des Stollens angelegt,

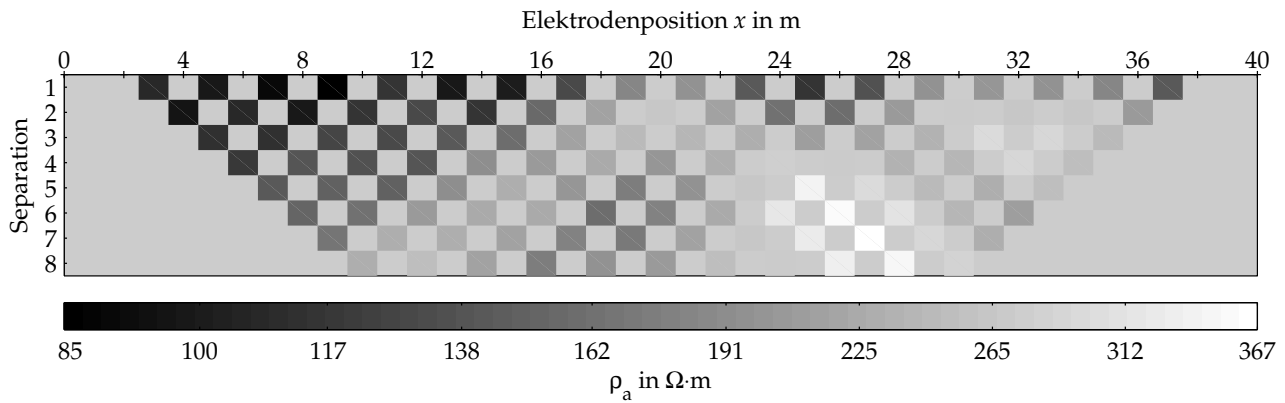


Abbildung 3: Dipol-Dipol-Pseudosektion über dem Rothschönberger Stolln. Dipollänge 2 m.

sodass eine zweidimensionale Betrachtung gerechtfertigt scheint. Eine Pseudosektion der  $n = 114$  Messwerte ist in Abb. 3 dargestellt.

Die Messung wurde mit dem beschriebenen Algorithmus und den Zielfunktionen (5) ( $p = 2$ ) und (6) invertiert. Die Evolution von 16 Teil-(Insel-)populationen à 200 Individuen wurde nach 500 Generationen abgebrochen. Der Austausch zwischen den Inseln erfolgte alle 25 Generationen.

Die Menge der nicht-dominierten Lösungen im Zielfunktionsraum ist in Abb. 4 dargestellt. Die mit Pfeilen gekennzeichneten Lösungen sind in Abb. 5 wiedergegeben. Alle Modelle zeigen eine befriedigende Datenanpassung  $f_1$ , der Zielwert von  $f_1^2 \cong 1$  (d.h.  $\chi^2 \cong n - 1$ ) wird jedoch nur der Größenordnung nach erreicht. Zunehmende Modellrestriktion (abnehmendes  $f_2$ ) bewirkt verringerte Kontraste im spezifischen elektrischen Widerstand sowie ein Verschmieren der Strukturen. Der Stolln kann in allen Modellen als hochohmiges Objekt bei  $x \approx 18 \dots 20(22)$  m,  $z \approx 2 \dots 3.5$  m identifiziert werden.

## Zusammenfassung

Das inverse Problem der Geoelektrik wurde als Minimierungsaufgabe eines nichtlinearen, vektorwertigen Zielfunktional formuliert und ihre Lösung als die Menge der Pareto-optimalen Lösungen identifiziert. Zur Konstruktion der Lösungsmenge wurde ein genetischer Inversionsalgorithmus entwickelt. Dieser Ansatz konnte erfolgreich bei der zweidimensionalen Inversion eines realen Felddatensatzes eingesetzt werden.

## Literatur

- Deb, K., Agrawal, S., Pratap, A. und Meyerivan, T. (2000). *A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II* (Technischer Bericht Nr. 2000001). Kanpur, Indien: Kanpur Genetic Algorithms Laboratory (KanGAL).
- Deb, K. und Goel, T. (2000). *Controlled elitist non-dominated sorting genetic algorithms for better convergence* (Technischer Bericht Nr. 2000004). Kanpur, Indien: Kanpur Genetic Algorithms Laboratory (KanGAL).
- Dey, A. und Morrison, H. F. (1979). Resistivity modelling for arbitrarily shaped two-dimensional structures. *Geophysical Prospecting*, 27, 106–136.
- Lowry, T., Allen, M. B. und Shive, P. N. (1989). Singularity removal: A refinement of resistivity modeling techniques. *Geophysics*, 54(6), 766–774.
- Zhao, S. und Yedlin, M. J. (1996). Some refinements on the finite-difference method for 3-d dc resistivity modeling. *Geophysics*, 61(5), 1301–1307.

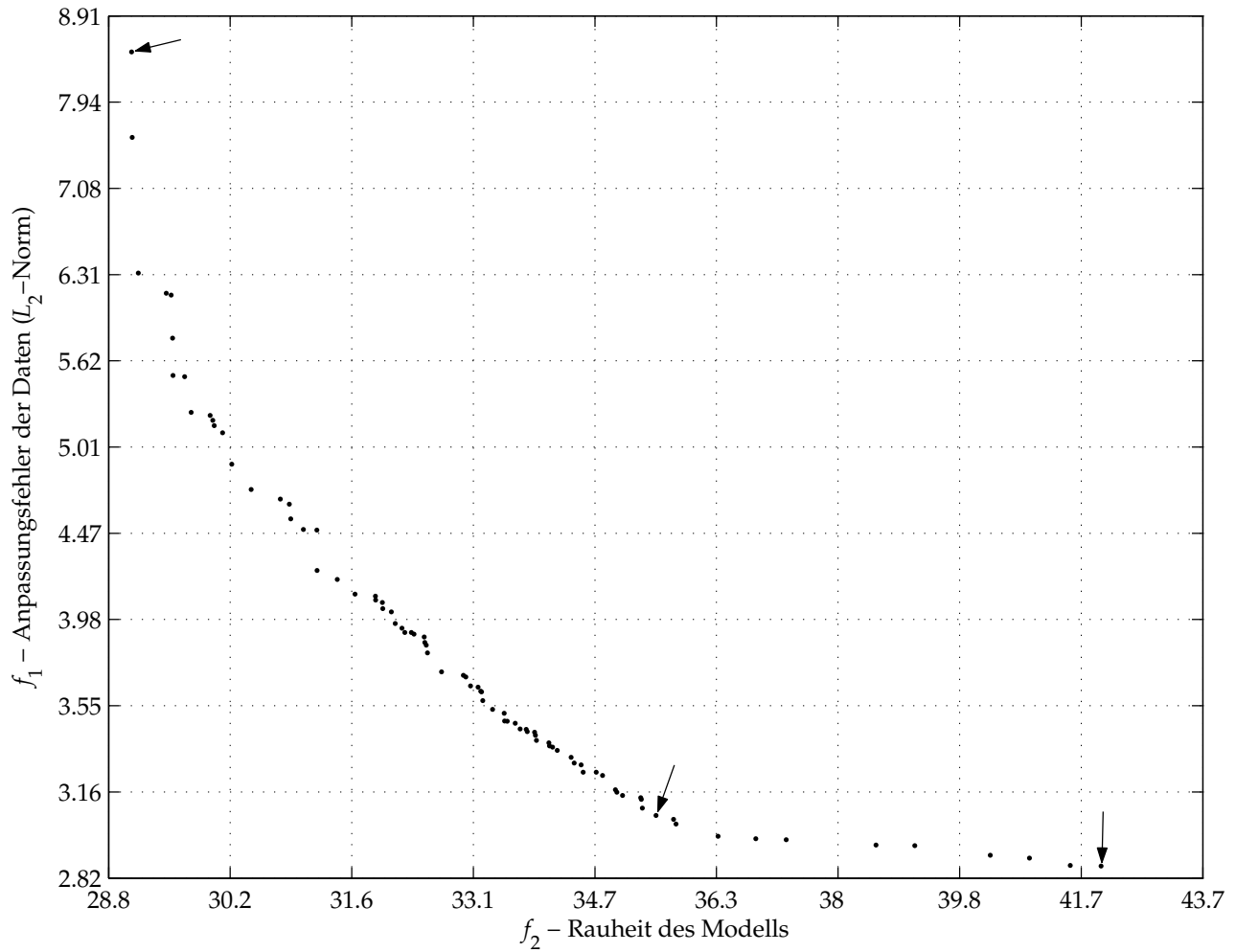


Abbildung 4: Menge der nicht-dominierten Lösungen im Zielfunktionsraum. Die mit Pfeilen gekennzeichneten Modelle sind in Abb. 5 dargestellt.

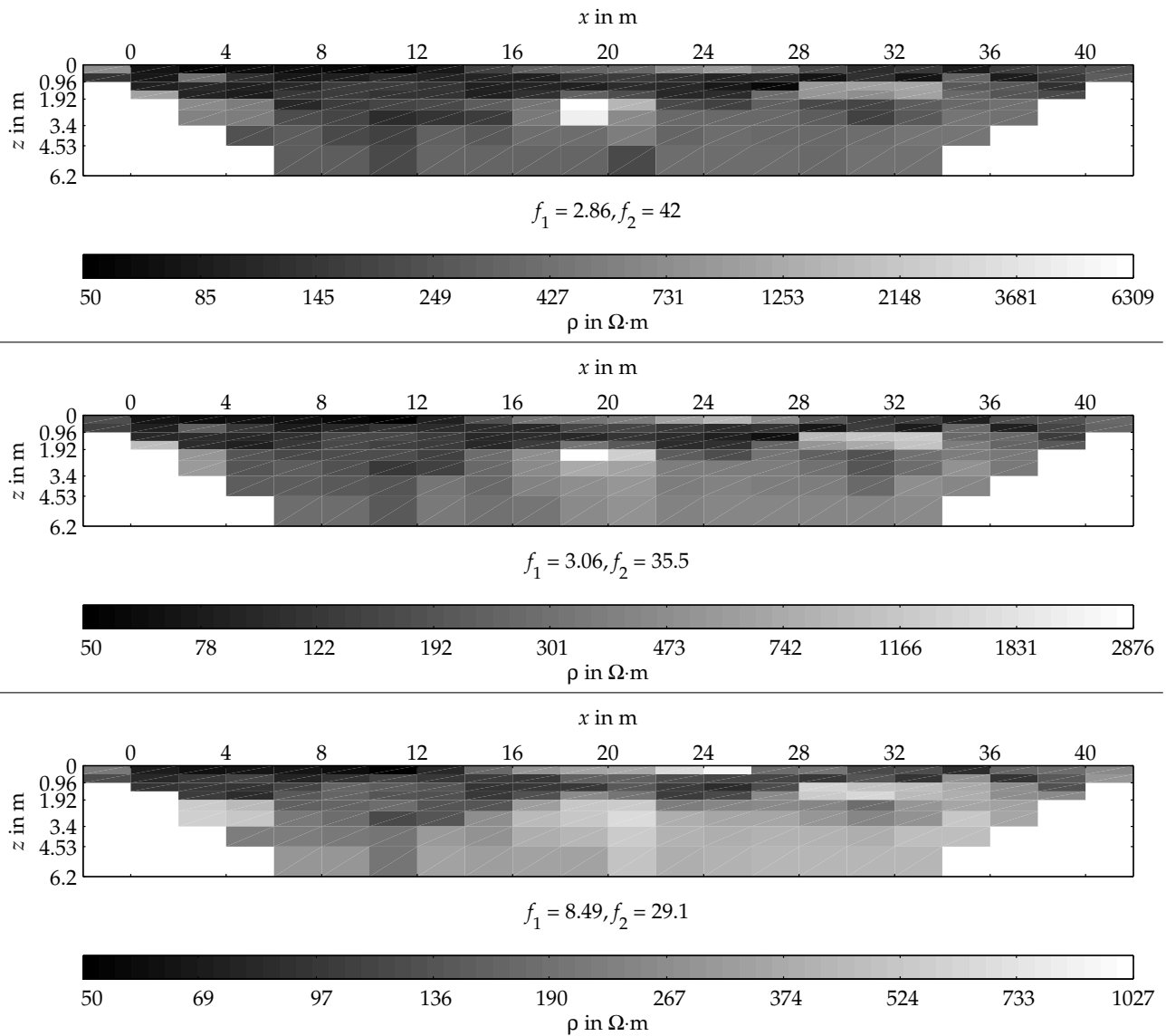


Abbildung 5: Drei ausgewählte Inversionsergebnisse.  $f_1$  bezeichnet den Anpassungsfehler der Daten und  $f_2$  das Maß für die Rauheit des Modells. Der Stolln stellt sich in allen Modellen als hochohmiges Objekt bei  $x \approx 18 \dots 20(22)$  m,  $z \approx 2 \dots 3.5$  m dar.