Fachhochschule Aachen

Abteilung Jülich

Fachbereich: Angewandte Naturwissenschaft und Technik

Studiengang: Technomathematik

Numerische Untersuchungen zur Bestimmung von Bodenstrukturen durch elektrische Widerstandstomographie (ERT)

Diplomarbeit von Andreas Kleefeld

Jülich, den 12. Dezember 2005

Die vorliegende Diplomarbeit wurde in Zusammenarbeit mit der Forschungszentrum Jülich GmbH, Institut für Festkörperforschung (IFF), angefertigt.

Diese Diplomarbeit wurde betreut von:

Referent: Prof. Dr. M. Reißel
Koreferent: Privatdozent Dr. H. Lustfeld

Diese Arbeit ist von mir selbständig angefertigt und verfasst. Es sind keine anderen als die angegebenen Quellen und Hilfsmittel benutzt worden.

Jülich, den 12. Dezember 2005

Inhaltsverzeichnis

Kapitel 1

Einführung

1.1 Einleitung

Die Electric Resistance Tomography (ERT) ist ein tomographisches Verfahren aus der Geopyhsik, das zur Ermittlung von geologischen Strukturen in einem Boden dient. Dieses Verfahren wird zum Beispiel auf dem Testfeld von Krauthausen angewendet. Es befindet sich etwa 10 km nordwestlich von Düren und ca. 7 km südöstlich von Jülich auf knapp 100 m über N.N. Höhe und damit in der Nähe des Forschungszentrums Jülich. Das Gelände ist 200 m x 70 m groß und wurde 1993 im Rahmen eines EU-Projektes errichtet.

Abbildung 1.1: Versuchsgelände bei Niederzier-Krauthausen von Nordwesten gesehen[1]

Beim ERT werden an der Oberfläche und in Bohrlöchern eines Bodens mehrere Elektroden angelegt, die entweder Ströme injizieren oder Spannungen messen können. Damit ist es möglich, durch eine der Elektroden Strom in den Boden zu injizieren, welcher an einer anderen wieder abfließt, und anschließend die sich einstellenden Potentiale an den restlichen Elektroden zu messen. Anhand der Messungen erhofft man sich, die Anordnung und Beschaffenheit des Bodens beschreiben zu können. Dies ist ein

[1] Quelle: http://www.fz-juelich.de/icg/icg-iv/index.php?index=62

typisches tomographisches Problem, bei dem aus gemessenen elektrischen Potentialen, in diesem Zusammenhang Beobachtungen genannt, auf die Bodenstruktur (Leitfähigkeit), auch Ursache genannt, geschlossen werden soll. Wie immer bei tomographischen Problemen ist das inverse Problem schlecht gestellt. Das bedeutet, daß kleine Fehler in den Beobachtungen, hervorgerufen durch die endliche Genauigkeit der Meßapparatur, große Fehler in den gesuchten Ursachen bewirken können. Der Grund ist, daß man das Potential nur an wenigen Stellen kennt (schließlich kann man ja nicht überall Elektroden verlegen), so daß gegebenenfalls eine Vielzahl von verschiedenen Parameterverteilungen gefunden werden kann, die zu den Beobachtungen zu passen scheinen. Um dieses Problem der Lösungsinstabilität zu mildern, wird versucht, möglichst viele Informationen über die Leitfähigkeitsverteilung des Bodens zu erhalten, indem man sich die Frequenzabhängigkeit der Leitfähigkeit verschiedener Bodentypen zu Nutze macht. Untersuchungen hinsichtlich dieser Informationsgewinnung sind bereits ausführlich in [8] durchgeführt worden.
Diese Arbeit beschäftigt sich mit dem Lösen des inversen Problems, das heißt, mit welchem Verfahren und welchem Erfolg die auftretenden Instabilitäten zu beherrschen sind, um bei gegebenen Meßdaten die Anordnung und Beschaffenheit des Bodens zu beschreiben. Hierbei sind sogenannte Regularisierungen unabdinglich, um zu verhindern, daß Rauschen in den Meßdaten die Approximationsqualität einer Lösung extrem verfälscht.
Der im folgenden dargestellte Lösungsalgorithmus benötigt weiterhin ein Modul, das bei einer gegebenen Verteilung das Potential in einem Volumen berechnet, also das direkte Problem löst. Dieses Modul ist in [9] beschrieben und in dieser Arbeit verwendet worden.
In Kapitel 2 werden die Grundlagen dargelegt. Hier werden unter anderem die verschiedenen Verfahren zur Lösung von linearen und nichtlinearen Ausgleichsproblemen mit integrierter Regularisierung erläutert.
Im nachfolgenden Kapitel 3 wird das entwickelte Programm erklärt. Dies beinhaltet das Erzeugen von gestörten Meßdaten aus eingelesenen Cole-Cole Parametern, die im Abschnitt 2.6 eingeführt werden, sowie das Lösen des inversen Problems.
Das Kapitel 4 beschäftigt sich mit den erzielten Ergebnissen anhand verschiedener Testbeispiele.
Abschließend folgt in Kapitel 5 eine Zusammenfassung und es wird ein Ausblick auf mögliche Verbesserungen zur Performance-Steigerung gegeben.

1.2 Aufgabenstellung

Es wird reihum an einer Elektrode Strom in den Boden injiziert, der an einer festen Stelle abgeführt wird. An den restlichen Elektroden wird das Potential gemessen. Dies erfolgt nicht nur mit Gleichstrom, sondern auch mit Wechselstrom. Anhand dieser Informationen sollen die Parameter, die die Leitfähigkeit charakterisieren, bestimmt werden. Diese ist zusätzlich frequenzabhängig.
Das mathematische Modell, das die Vorgänge im Boden beschreibt, beruht auf den Maxwellschen Gleichungen und ist ausführlich in [8] beschrieben.
Sei σ die Leitfähigkeit, u das Potential, I der eingeleitete Strom, c die Cole-Cole Parameter, die die Leitfähigkeit bestimmen, und ω die Frequenz, dann gilt:

$$\begin{aligned} -\mathrm{div}(\sigma(c,\omega)\nabla u) &= I \text{ in } \Omega \\ u &= 0 \text{ auf } \Gamma_0 \\ \partial_n u &= 0 \text{ auf } \Gamma_1 \end{aligned}$$

mit dem Gebiet $\Omega = [0, x_{\max}] \times [0, z_{\max}]$, dem Rand $\partial\Omega = \Gamma_0 \cup \Gamma_1$ mit $\Gamma_0 = ([0, x_{\max}] \times \{0\}) \cup ([0, z_{\max}] \times \{0\}) \cup ([0, z_{\max}] \times \{x_{max}\})$ und $\Gamma_1 = [0, x_{\max}] \times \{z_{max}\}$.
Es gilt:

$$\sigma(c, \omega) = \frac{1}{R_0 \left[1 - m\left(1 - \frac{1}{1+(i\omega\tau)^\gamma}\right)\right]} \text{ mit } c = (R_0, m, \tau, \gamma)^T$$

Der Nenner ist hierbei das sogenannte Cole-Cole Gesetz. Es beschreibt den Frequenzgang der Leitfähigkeit. Eine genauere Erklärung findet sich in Kapitel 2.6.

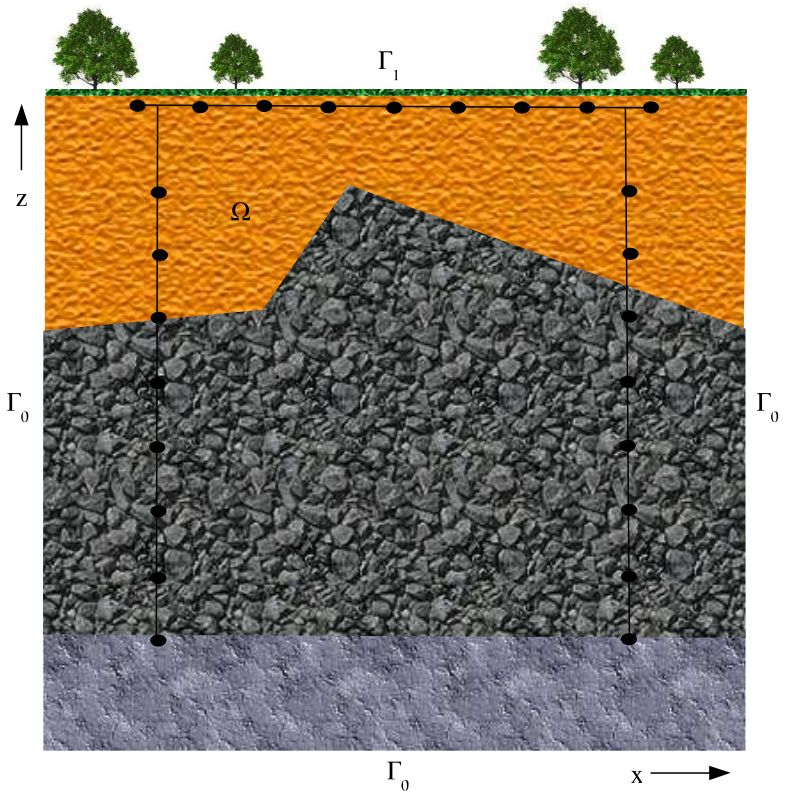


Abbildung 1.2: Gebiet Ω, Ränder Γ_0 und Γ_1 und Elektroden •

Beim direkten (Vorwärts-) Problem sind das Gebiet Ω, die Parameterverteilung c, die Frequenz ω und der Strom I gegeben. Gesucht wird das Potential u.
Beim inversen Problem hingegen sind das Gebiet Ω, der Strom I, die Frequenz ω und das komplexe Potential u_i, $i = 1, \ldots, n$ an n-Stellen gegeben. Gesucht sind die Cole-Cole Parameter c auf dem ganzen Gebiet Ω.
Wir nutzen in unserem Fall mehrere Ströme. Dort, wo diese eingeleitet werden und abfließen, werden diese Potentiale nicht gemessen. Außerdem geschieht dieses Vorgehen mit mehreren Frequenzen ω. Die Lösung $u = u(\sigma(c, \omega), I)$ des direkten Problems hängt nichtlinear von $\sigma(c, \omega)$ ab. Daher wird zur Lösung des inversen Problems folgende Methode verwendet:
Zu bestimmen sind die diskreten Cole-Cole Parameter $c = (R_0, m, \tau, \gamma)$ auf dem ganzen Gebiet Ω, so daß

$$H(c) = \frac{1}{2} \frac{1}{n \cdot m \cdot p} \sum_{k=1}^{p} \sum_{j=1}^{m} \sum_{i=1}^{n} (u(\sigma(c, \omega_k), I_j)_i - u_i(\omega_k, I_j))^2$$

minimal wird, wobei $u_i(\omega_k, I_j)$ die im Ort (x_i, z_i) gemessenen Potentialwerte mit dem Strom I_j bei

der Frequenz ω_k sind. Wir erhalten somit das folgende nichtlineare Ausgleichsproblem:

$$H(c) = \frac{1}{2}||\hat{u}(\hat{\omega}, \hat{I}) - \hat{u}(\sigma(c, \hat{\omega}), \hat{I})||_2^2 \tag{1.1}$$

mit

$$\hat{u} = \begin{pmatrix} \hat{u}_1 \\ \vdots \\ \hat{u}_{\hat{n}} \end{pmatrix}, \ \hat{I} = \begin{pmatrix} I_1 \\ \vdots \\ I_m \end{pmatrix} \text{ und } \hat{\omega} = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_p \end{pmatrix}$$

Mit $h(c) = \hat{u}(\hat{\omega}, \hat{I}) - \hat{u}(\sigma(c, \hat{\omega}), \hat{I}), \quad h : \mathbb{R}^n \longrightarrow \mathbb{C}^{\hat{n}}$ suchen wir ein Minimum von $||h(c)||_2$. Dies zu minimieren, ist äquivalent mit

$$||\mathbf{Re}(h(c))||_2 + ||\mathbf{Im}(h(c))||_2.$$

Diese Minimierung kann auch als

$$||f(c)||_2 \equiv ||\begin{pmatrix} \mathbf{Re}(h(c)) \\ \mathbf{Im}(h(c)) \end{pmatrix}||_2, \quad f : \mathbb{R}^n \longrightarrow \mathbb{R}^m \tag{1.2}$$

mit $m = 2 \cdot \hat{n}$ umformuliert werden.
Dies ist äquivalent mit

$$\hat{c} = \arg\min_c F(c),$$

wobei

$$F(c) = \frac{1}{2}||f(c)||_2^2 = \frac{1}{2} f(c)^T f(c) \tag{1.3}$$

das quadratische Fehlerfunktional ist.

Es gibt mehrere Optimierungsverfahren, die dieses Problem lösen. Hier werden jedoch nur Verfahren präsentiert, deren Konvergenzgeschwindigkeit mindestens linear ist. Die vorgestellten Verfahren konvergieren manchmal quadratisch und benötigen keine zweiten Ableitungen.
Wir nehmen an, daß $f(c)$ zweimal stetig ableitbar ist und erhalten mit Hilfe der Taylorentwicklung für das lineare Modell $l(h)$ von f in der Umgebung c

$$f(c + h) \simeq l(h) \equiv f(c) + J(c)h,$$

wobei $||h||_2$ klein ist. $J(c) \in \mathbb{R}^{m \times n}$ ist hierbei die Jacobimatrix mit den partiellen Ableitungen $(J(c))_{ij} = \frac{\partial f_i}{\partial c_j}(c)$.
Setzt man dies in (1.3) ein, so erhält man die Näherung $L(h)$ von F

$$\begin{aligned} F(c + h) \simeq L(h) &\equiv \frac{1}{2} l(h)^T l(h) \\ &= \frac{1}{2} f(c)^T f(c) + h^T J(c)^T f(c) + \frac{1}{2} h^T J(c)^T J(c) h \\ &= F(c) + h^T J(c)^T f(c) + \frac{1}{2} h^T J(c)^T J(c) h. \end{aligned}$$

Der Gradient von F ist:

$$g \equiv F'(c) = J(c)^T f(c).$$

Alle Methoden für die nichtlineare Optimierung sind iterativ, das heißt, ausgehend von einem Startpunkt c_0 wird eine Folge $c_1, c_2, \ldots$ generiert, die gegen ein lokales Minimum $\hat{c}$ konvergiert.

Verschiedene Strategien sind hierbei möglich, um diese Folge $c_{k+1} = c_k + \lambda_k h_k$ zu erzeugen. In dieser Arebit werden untersucht:

- Gauß-Newton Verfahren
- Levenberg-Marquardt Verfahren
- Powell's Dog Leg Verfahren

Beim Gauß-Newton Verfahren wird $L(h)$ in jedem Schritt minimiert. Beim Dog Leg Verfahren wird hingegen $L(h)$ unter der Nebenbedingung $h^T h \leq \Delta$ minimiert, wobei Δ der Radius des Vertrauensbereichs ist. Diese Verfahren werden in Kapitel 2.5 ausführlich erklärt.
In jedem Iterationsschritt muß sowohl die Jacobimatrix $J(c_k)$ mit den Ableitungen als auch der Vektor $f(c_k)$ neu bestimmt werden.
In jedem Verfahren muß ein lineares Ausgleichsproblem geeignet gelöst werden, wobei der Löser eine Regularisierung enthalten sollte, damit Rauschen in den Meßdaten die Approximationsqualität nicht extrem verschlechtert (siehe Kapitel 2.4).
Weiterhin muß die Schrittlänge λ_k in jedem Iterationsschritt optimal bestimmt werden. Ferner soll die Minimierung einfache lineare Nebenbedingungen beachten, da sich die gesuchten Parameter in einem bestimmten Bereich befinden (siehe Kapitel 2.6).
Des weiteren werden die Parameter nicht im vollen 3-D Volumen bestimmt, sondern im 2-dimensionalen Raum. Die physikalische Bedeutung dieser Einschränkung ist hierbei, daß sich in y-Richtung die Leitfähigkeit nicht ändert. Damit die Potentiale sich nicht in y-Richtung ändern, wird angenommen, daß die Stromeinleitungen und -ausleitungen in y-Richtung konstant sind. Grundsätzlich ist es möglich, das volle 3-D Problem zu lösen, sofern die Funktionen von [9] dies bewerkstelligen. Dies war zum Zeitpunkt dieser Diplomarbeit aber nicht der Fall, da sich diese Funktionen noch in der Weiterentwicklung befanden.

1.3 Ziel

Das Ziel dieser Arbeit ist es, ein speziell an dieses Problem angepaßtes Programm in MATLAB zu entwickeln, welches das inverse Problem löst. Dabei werden verschiedene Verfahren zur Lösung des nichtlinearen Ausgleichsproblems getestet. Zum einen wird ein Gauß-Newton Verfahren mit Liniensuche verwendet. In diesem Algorithmus wird in jedem Iterationsschritt eine optimale Suchrichtung ermittelt, was eine geschickte Handhabung zur Lösung von Randwertproblemen erfordert. Zum Aufstellen einer Spalte der Ableitungsmatrix muß jeweils ein Randwertproblem gelöst werden, was hohe Rechenzeiten nach sich zieht. In unserem Fall ist es aber möglich, direkt Ax und $A^H z$ zu berechnen, was ebenfalls vom Aufwand her, dem Lösen eines Randwertproblems entspricht (siehe [9]).
Anschließend muß ein lineares Ausgleichsproblem gelöst werden, das Regularisierungstechniken verwendet, um zu vermeiden, daß Störungen in den Meßdaten die Approximationsqualität verschlechtern und somit die Lösung unbrauchbar machen. Ferner müssen optimale Regularisierungsparameter gefunden werden.
Die Ermittlung der optimalen Suchlänge ist ein entscheidender Faktor, denn sie bestimmt das Konvergenzverhalten stark.
Algorithmen, die dies automatisieren, sind zum Beispiel das Levenberg-Marquardt und das Dog Leg Verfahren.
Des weiteren soll die Minimierung einfache lineare Nebenbedingungen (siehe Tabelle 3.4) beachten, da die Cole-Cole Parameter in einem bestimmten Bereich liegen.

Natürlich soll aus der gefundenen Lösung des inversen Problems eine Aussage über die Parameterverteilung im Boden gemacht werden können, welche der wahren Verteilung möglichst nahe kommen sollte.

Kapitel 2

Grundlagen

2.1 Schlechtgestelltheit

Das inverse Problem ist schlecht gestellt, da bei Standardverfahren kleine Störungen in den gemessenen Potentialen große Störungen in den gesuchten Parametern bewirken. Die Potentiale liegen nicht exakt vor, sondern werden aus Messungen gewonnen. Die Meßapparaturen erfassen das Potential ungenau. Dies führt dazu, daß Rauschen in den Messungen zu dramatischen Rekonstruktionsfehlern führt (siehe Tabelle 2.1).
Bei einer festen Meßanordnung gibt es eine Vielzahl an möglichen Lösungen, die das nichtlineare Ausgleichsproblem minimieren. An den Stellen, an denen kaum Strom fließt, können die Parameter nicht identifiziert werden. Das heißt, daß Variationen in diesen Bereichen kaum die Potentiale an den gemessenen Stellen verändern. Das bedeutet wiederum umgekehrt, daß es mehrere Lösungen geben kann.
Man erhält also im Rahmen der Meßungenauigkeit auch noch andere Lösungen. Aus diesen muß die „Richtige“ ausgewählt werden. Bei dieser Wahl hängt es davon ab, was als „richtig“ angesehen wird. Eine Möglichkeit ist diejenige zu nehmen, die möglichst glatt ist.

In unserem Fall liegt eine Abbildung h von $\mathbb{R}^n$ nach $\mathbb{C}^m$ vor. Wir betrachten im folgenden nur reelle Größen, da es möglich ist, die Minimierung von $||h(c)||_2$ umzuformulieren. Der Vektor $h(c)$ und die Jacobimatrix werden nach Real- und Imaginärteil aufgespalten und untereinander zusammengefügt (siehe (1.2)).

2.2 Pseudoinverse

Satz 1

Sei $A \in \mathbb{R}^{m\times n}$, $m \geq n$, mit Rang, der kleiner als n ist, dann ist das Minimalproblem

$$\min_{x\in\mathbb{R}^n} ||Ax - b||_2$$

nicht mehr eindeutig lösbar. Das modifizierte Minimalproblem

$$x = \min_{z\in X} ||z||_2, \quad X = \{z|\ ||Az - b||_2 \text{ minimal}\} \tag{2.1}$$

besitzt genau eine Lösung.

Definition 1

Sei $A \in \mathbb{R}^{m\times n}$. Dann gibt es genau eine Lösung $x \in \mathbb{R}^n$ von (2.1). Diese hängt linear von b ab, das heißt es existiert eine eindeutige, von b unabhängige Matrix $A^+ \in \mathbb{R}^{n\times m}$ mit

$$x = A^+ b, \quad \forall b \in \mathbb{R}^m.$$

A^+ heißt Pseudoinverse oder auch Moore-Penrose-Inverse zu A.

Bemerkung 1

Die Pseudoinverse A^+ zu A ist durch die folgenden vier Penrose-Axiome eindeutig festgelegt

$$(A^+A)^T = A^+A, \quad (AA^+)^T = AA^+, \quad A^+AA^+ = A^+, \quad AA^+A = A$$

2.3 Singulärwertzerlegung

Die Singulärwertzerlegung einer Matrix A steht in engem Zusammenhang mit der Pseudoinversen A^+.

Definition 2

Sei $A \in \mathbb{R}^{m\times n}$ beliebig, dann gibt es orthonormale Matrizen $U \in \mathbb{R}^{m\times m}$ und $V \in \mathbb{R}^{n\times n}$, so daß gilt:

$$U^T AV = \Sigma$$

mit

$$\Sigma \in \mathbb{R}^{m\times n}, \quad \Sigma = diag(\sigma_1, \ldots, \sigma_p), \quad p = \min(m, n)$$

und

$$\sigma_1 \geq \sigma_2 \geq, \ldots, \sigma_p \geq 0.$$

$U^T AV = \Sigma$ bzw. $A = U\Sigma V^T$ heißt Singulärwertzerlegung von A mit Singulärwerten σ_i.

Definition 3

Sei r=rang(A) und $A = U\Sigma V^T$ mit

$$\Sigma = diag(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0) \in \mathbb{R}^{m\times n}$$

dann ist die Pseudoinverse berechenbar über

$$A^+ = V\Sigma^+ U^T \in \mathbb{R}^{n\times m} \quad \text{mit} \quad \Sigma^+ = diag(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}, 0, \ldots, 0) \in \mathbb{R}^{n\times m}$$

2.4 Lineares Ausgleichsproblem

Bei dem zu behandelnden Problem ist ein lineares Ausgleichsproblem $||Ax - b||_2$ zu lösen. Die Matrix A ist hierbei beliebig und nicht quadratisch. Ferner ist ihr Rang nicht maximal. Damit geht die Eindeutigkeit eines Minimierers verloren, das heißt, kennt man eine Lösung, so ist es möglich weitere Minimierer dadurch zu erzeugen, daß man zur Lösung ein beliebiges Element aus dem Kern von A hinzuaddiert. Ein eindeutiges Minimum erhält man nur, falls der Defekt Null ist.
Wird allerdings der kleinste Minimierer im Sinne der $||.||_2$ gesucht, so ist dies über die eindeutige lineare Abbildung A^+ möglich.
Aufgrund der Schlechtgestelltheit müssen Regularisierungsverfahren verwendet werden.

Es gibt sowohl direkte Verfahren, wie die Tikhonov Regularisierung oder abgeschnittene Singulärwertzerlegung, als auch iterative Verfahren, wie das Landweber Verfahren und das CGLS-Verfahren (Conjugate Gradient Least Square). Hierbei übernimmt der Iterationsindex die Rolle des Regularisierungsparameters.
Im folgenden werden diese Verfahren erklärt.

2.4.1 Abgeschnittene Singulärwertzerlegung

Das lineare Ausgleichsproblem $||Ax - b||_2$ kann mit Hilfe der Pseudoinversen eindeutig gelöst werden, indem man $x = A^+b$ berechnet. Das Problem hierbei ist, daß Störungen der Messungen die Approximationsqualität extrem verschlechtern. Eine Möglichkeit, dies zu beseitigen, ist der Abschnitt kleiner Singulärwerte. Schneidet man zuviel ab, ist dies allerdings schlecht für die Approximationsqualität. Die Vorgehensweise kann folgendermaßen beschrieben werden:
Man erzeugt sich von der Matrix A eine Singulärwertzerlegung und erhält damit die Matrizen U, Σ und V. In der Matrix Σ setzt man alle Singulärwerte, die kleiner als die angegebene Grenze sind, auf Null und erhält die neue Matrix $\Sigma_\star$. Anschliessend multipliziert man die Matrizen wieder zusammen. Dies liefert die neue abgeschnittene Pseudoinverse $A^+_\star = V\Sigma^+_\star U^T$.
Die Wahl der Grenze ist hierbei abhängig von der Größe der absoluten Störung in den Daten. Ist diese bekannt, dann sollte die Grenze auf den Wert der absoluten Störung gesetzt werden.

Beispiel 1

Gegeben seien die Matrix $A \in \mathbb{R}^{10\times 9}$, der Vektor $b \in \mathbb{R}^{10}$ und $b^\delta \in \mathbb{R}^{10}$.

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} & \frac{1}{12} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} & \frac{1}{12} & \frac{1}{13} \\ \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} & \frac{1}{12} & \frac{1}{13} & \frac{1}{14} \\ \frac{1}{7} & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} & \frac{1}{12} & \frac{1}{13} & \frac{1}{14} & \frac{1}{15} \\ \frac{1}{8} & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} & \frac{1}{12} & \frac{1}{13} & \frac{1}{14} & \frac{1}{15} & \frac{1}{16} \\ \frac{1}{9} & \frac{1}{10} & \frac{1}{11} & \frac{1}{12} & \frac{1}{13} & \frac{1}{14} & \frac{1}{15} & \frac{1}{16} & \frac{1}{17} \\ \frac{1}{10} & \frac{1}{11} & \frac{1}{12} & \frac{1}{13} & \frac{1}{14} & \frac{1}{15} & \frac{1}{16} & \frac{1}{17} & \frac{1}{18} \end{pmatrix}, \quad b = \begin{pmatrix} 2.8290 \\ 1.9290 \\ 1.5199 \\ 1.2699 \\ 1.0968 \\ 0.9682 \\ 0.8682 \\ 0.7879 \\ 0.7217 \\ 0.6661 \end{pmatrix}, \quad b^\delta = \begin{pmatrix} 2.8167 \\ 1.8818 \\ 1.5234 \\ 1.2780 \\ 1.0644 \\ 1.0019 \\ 0.9019 \\ 0.7868 \\ 0.7310 \\ 0.6711 \end{pmatrix}$$

Wenn man auf den Vektor b eine normalverteilte Störung mit Mittelwert 0 und Standardabweichung $\max\{|b|\}\cdot 1\%$ *addiert, erhält man den Vektor b^δ. Löst man anschließend das Gleichungssystem $Ax = b^\delta$ mit der Pseudoinversen A^+ und mit der abgeschnittenen Singulärwertzerlegung $A^+_\star$, wobei die Grenze zum Abschneiden* $\max\{|b|\}\cdot 1\%$ *war, dann bekommt man die in Tabelle* (2.1) *präsentierten Ergebnisse. Die wahre Lösung ist $x = (1,1,1,1,1,1,1,1,1)^T$.*

Pseudoinverse	abges. SVD
-37566.362206528	1.308551300
2452820.531510646	0.222794986
-39757941.774206281	0.602917385
274108640.471090555	0.933350020
-976982617.619923115	1.146236697
1947394688.194134712	1.273295091
-2191122422.490019798	1.343956282
1300288281.068540573	1.378136738
-316360421.951824188	1.388678554

Tabelle 2.1: Links: $x = A^+ b^\delta$, rechts: $x = A_\star^+ b^\delta$

2.4.2 Tikhonov Regularisierung

Sei $A \in \mathbb{R}^{m\times n}$ beliebig vorgegeben und soll $J(x) = ||Ax - b||_2^2$ minimiert werden, so kann hierfür die Tikhonov Regularisierung verwenden werden. Hierbei betrachtet man

$$J_\alpha(x) = ||Ax - b||_2^2 + \alpha^2 ||x||_2^2,$$

wobei α ein Regularisierungsparameter ist.

Satz 1

Den Ausdruck $J_\alpha(x) = ||Ax - b||_2^2 + \alpha^2||x||_2^2$ zu minimieren, ist äquivalent mit

$$(A^T A + \alpha^2 I)x_\alpha = A^T b.$$

Beweis 1

Der Ausdruck $J_\alpha(x) = ||Ax - b||_2^2 + \alpha^2||x||_2^2$ kann auch so formuliert werden:

$$J_\alpha(x) = || \underbrace{\begin{pmatrix} A \\ \alpha I \end{pmatrix}}_{\hat{A}} x - \underbrace{\begin{pmatrix} b \\ 0 \end{pmatrix}}_{\hat{b}} ||_2^2 \quad I \in \mathbb{R}^{n\times n} \quad \alpha > 0.$$

Für $\alpha > 0$ hat $\hat{A} \in \mathbb{R}^{(m+n)\times n}$ vollen Rang. Daraus folgt, daß ein Minimierer x_α existiert, der durch folgenden Ausdruck gegeben ist:

$$\begin{aligned} & \hat{A}^T \hat{A} x_\alpha = \hat{A}^T \hat{b} \\ \iff \quad & (A^T, \alpha I) \begin{pmatrix} A \\ \alpha I \end{pmatrix} x_\alpha = (A^T, \alpha I) \begin{pmatrix} b \\ 0 \end{pmatrix} \\ \iff \quad & (A^T A + \alpha^2 I) x_\alpha = A^T b \end{aligned} \tag{2.2}$$

□

Satz 2

Die eindeutige Lösung des Tikhonov Problems konvergiert für $\alpha \longrightarrow 0$ gegen $A^+ b$.

Beweis 2

Es ist $A = U\Sigma V$ mit U und V orthonormal. Durch eine geeignete Basistransformation geht A also in Σ über. Deshalb nehmen wir Œ

$$A = \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$$

an. Für $\hat{A} = \binom{A}{\alpha I}$ erhalten wir

$$\hat{A}^T \hat{A} \equiv \begin{pmatrix} \sigma_1^2 + \alpha^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_r^2 + \alpha^2 & & & \\ & & & \alpha^2 & & \\ & & & & \ddots & \\ & & & & & \alpha^2 \end{pmatrix} x_\alpha = \begin{pmatrix} \sigma_1 b_1 \\ \vdots \\ \sigma_r b_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \equiv \hat{A}^T b$$

also:

$$x_\alpha = \begin{pmatrix} \frac{\sigma_1}{\sigma_1^2+\alpha^2} b_1 \\ \vdots \\ \frac{\sigma_r}{\sigma_r^2+\alpha^2} b_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_1+\frac{\alpha^2}{\sigma_1}} b_1 \\ \vdots \\ \frac{1}{\sigma_r+\frac{\alpha^2}{\sigma_r}} b_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \overset{\alpha \to 0}{\longrightarrow} \begin{pmatrix} \frac{b_1}{\sigma_1} \\ \vdots \\ \frac{b_r}{\sigma_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A^+ b$$

□

Zu sehen ist, daß für $\frac{\alpha}{\sigma_i} \ll 1$ der Anteil b_i behandelt wird wie bei $A^+ b$. Ist hingegen $\frac{\alpha}{\sigma_i} \gg 1$, dann werden Singulärwerte $\sigma_i \ll \alpha$ abgeschnitten.
Der Regularisierungsparameter α steuert hierbei gegenläufige Trends. Wird α eher groß gewählt, so werden die Fehler in b gedämpft. Wird α hingegen eher klein gewählt, so ist J_α eine gute Approximation von J.
A-priori weiß man nicht, wie groß ein günstiges α zu wählen ist, da man dafür die Singulärwerte benötigt.

Erwähnt sei noch, daß eine modifizierte Variante der Tikhonov Regularisierung existiert:

$$J_\alpha(x) = ||Ax - b||_2^2 + \alpha^2 ||B(x - d)||_2^2$$

Wird für die Matrix B der diskrete Gradient gewählt, so werden hohe Ableitungen von x bestraft. Das heißt, es werden Lösungen bevorzugt, die glatt sind. Der Vektor d kann benutzt werden, falls eine Näherungslösung von x bekannt ist. So werden Lösungen, die weit entfernt von x liegen, bestraft.

2.4.3 Landweber Verfahren

Das Landweber Verfahren ist ein einfaches Gradienten Verfahren, wobei μ ein Relaxationsparameter ist. Es lautet:

$$x_{k+1} = x_k + \mu A^T(b - Ax_k)$$

Satz 3
Das Landweber Verfahren konvergiert gegen die Pseudoinverse, falls $\forall\, i$ gilt: $|1 - \mu\sigma_i^2| < 1$

Beweis 3
Das Landweber-Verfahren kann folgendermaßen geschrieben werden:

$$x_{k+1} = \underbrace{(I - \mu A^T A)}_{C_\mu} x_k + \underbrace{\mu A^T b}_{d_\mu}$$

Sei jetzt $x_0 = 0$, dann erhalten wir:

$$\begin{aligned} x_1 &= d_\mu \\ x_2 &= C_\mu d_\mu + d_\mu \\ &\vdots \\ x_k &= \sum_{j=0}^{k-1} C_\mu^j d_\mu \end{aligned} \tag{2.3}$$

Es sei Œ

$$A = \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

Dann folgt daraus:

$$C_\mu = I - \mu A^T A = \begin{pmatrix} 1-\mu\sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & 1-\mu\sigma_r^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = \begin{pmatrix} \tau_1 & & & & & \\ & \ddots & & & & \\ & & \tau_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

mit $\tau_i = 1 - \mu\sigma_i^2$ und

$$d_\mu = \mu A^T b = \begin{pmatrix} \mu\sigma_1 b_1 \\ \vdots \\ \mu\sigma_r b_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Damit folgt:

$$x_k = \sum_{j=0}^{k-1} C_\mu^j d_\mu = \sum_{j=0}^{k-1} \begin{pmatrix} \tau_1^j & & & & & \\ & \ddots & & & & \\ & & \tau_r^j & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} d_\mu = \begin{pmatrix} \sum_{j=0}^{k-1} \tau_1^j & & & & & \\ & \ddots & & & & \\ & & \sum_{j=0}^{k-1} \tau_r^j & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} d_\mu$$

Es gilt:

$$\sum_{j=0}^{k-1} \tau_i^j = \frac{1-\tau_i^k}{1-\tau_i} = \frac{1-\tau_i^k}{\mu\sigma_i^2}.$$

Wir erhalten also insgesamt:

$$x_{k,i} = \frac{1-\tau_i^k}{\mu\sigma_i^2}\mu\sigma_i b_i = (1-\tau_i^k)\frac{b_i}{\sigma_i}, \text{ für } i = 1, \ldots, r.$$

Für $i > r$ gilt $x_{k,i} = 0$. Da $|\tau_i| < 1$ ist, folgt

$$x_k \overset{k\to\infty}{\longrightarrow} \begin{pmatrix} \frac{b_1}{\sigma_1} \\ \vdots \\ \frac{b_r}{\sigma_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A^+ b$$

□

Dabei ist zu beachten, daß das Verfahren umso schneller konvergiert, je kleiner der Ausdruck τ_i im Betrag ist.

Satz 4

Die höchste Konvergenzgeschwindigkeit erreicht man mit

$$\mu_{opt} = \frac{2}{\sigma_1^2 + \sigma_r^2}.$$

Beweis 4

Es wird der Parameter μ so bestimmt, daß $\max\limits_{i=1,\ldots,r} \tau_i$ minimal wird. Da $|\tau_i| < 1$ gelten muß, ergibt sich: $\tau\sigma_i^2 \in (0,2)$. Daraus folgt, daß $\mu > 0$ sein muß und $\mu < \frac{2}{\sigma_i^2}$, das heißt: $\mu < \frac{2}{\sigma_1^2}$.
Mit $\sigma_1 \geq \ldots \geq \sigma_r$ folgt $\tau_1 \leq \ldots \leq \tau_r$ und $\max\limits_{i=1,\ldots,r} |\tau_i|$ wird minimal, falls $\tau_1 = -\tau_r$. Durch Umrechnen erhält man:

$$\mu = \frac{2}{\sigma_1^2 + \sigma_r^2}$$

□

Bemerkung 2

Bei gestörten Daten $\hat{b} = b + \Delta b$ sollte der Ausdruck $1 - \mu\sigma_i^2 \approx 1$ sein. Dies wirkt sich stabilisierend aus. Der Nachteil ist hierbei allerdings, daß das Verfahren sehr langsam konvergiert.

2.4.4 Konjugiertes Gradientenverfahren

Das klassische CG-Verfahren zur Lösung von linearen Gleichungssystemen $Ax = b$ mit $A \in \mathbb{R}^{n \times n}$ und $b \in \mathbb{R}^n$ kann als 3-Term Rekursion geschrieben werden:

begin
  x_0 gegeben
  $p_0 := r_0 := b - Ax_0$
  $k := 0$
  while $||r_k - r_{k-1}||_2 > \epsilon_1$
    $\alpha_k := \frac{<r_k, r_k>}{<p_k, Ap_k>}$
    $x_{k+1} := x_k + \alpha_k p_k$
    $r_{k+1} := r_k - \alpha_k A p_k$
    $\beta_k := \frac{<r_{k+1}, r_{k+1}>}{<r_k, r_k>}$
    $p_{k+1} := r_{k+1} + \beta_k p_k$
  end
end

Abbildung 2.1: Konjugiertes Gradienten Verfahren

Zur Approximation der Normalengleichung $A^T Ax = A^T b$ mit $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$ mit $m > n$ gibt es eine angepaßte Variante, nämlich das CGLS-Verfahren (Conjugate Gradient Least Square):

begin
  x_0 gegeben
  $s_0 := b - Ax_0$
  $p_0 := r_0 := A^T s_0$
  $k := 0$
  while $||s_k - s_{k-1}||_2 > \epsilon_1$
    $\alpha_k := \frac{<r_k, r_k>}{<Ap_k, Ap_k>}$
    $x_{k+1} := x_k + \alpha_k p_k$
    $s_{k+1} := s_k - \alpha A p_k$
    $r_{k+1} := A^T s_{k+1}$
    $\beta_k := \frac{<r_{k+1}, r_{k+1}>}{<r_k, r_k>}$
    $p_{k+1} := r_{k+1} + \beta_k p_k$
  end
end

Abbildung 2.2: Konjugierte Gradienten für Normalengleichung

Die Vorteile des CGLS sind hierbei: Die Matrix $A^T A$ muß nicht explizit aufgestellt werden. Außerdem ist die Speicheranforderung des Verfahrens von fünf Vektoren sehr gering und es konvergiert in der Regel sehr schnell.

In dieser Arbeit ist die Matrix A und der Vektor b aber komplex. Des weiteren wird ein Strafterm $\mu^2 ||B(x - d)||_2^2$ benötigt. Um das Minimierungsproblem

$$||Ax - b||_2^2 + \mu^2 ||B(x - d)||_2^2,$$

das äquivalent zu

$$(A^T A + \mu^2 B^T B)x = A^T b + \mu B^T B d$$

ist, mit Hilfe des CGLS zu lösen, muß das Verfahren noch modifiziert werden. Da aber ein reelles x gesucht wird, wird das Minimum von

$$(A_r^T A_r + A_i^T A_i + \mu^2 B^T B)x = A_r^T b_r + A_i^T b_i + \mu B^T B d$$

gesucht, wobei A_r, b_r der Realteil und A_i, b_i der Imaginärteil von A und b ist.
Eine Möglichkeit ist, an das oben beschriebene CGLS Verfahren die Matrix

$$\hat{A} = \begin{pmatrix} A_r \\ A_i \\ \mu B \end{pmatrix}$$

und den Vektor

$$\hat{b} = \begin{pmatrix} b_r \\ b_i \\ \mu B d \end{pmatrix}$$

zu übergeben.

Da es aber sehr aufwendig ist, die Matrix A und A^H zu berechnen, habe ich Funktionen, wie in [9] beschrieben, verwendet, die in der Lage sind, direkt Matrix-Vektor-Produkte Ax und $A^H z$ zu berechnen. Im folgenden werden diese Funktionsauswertungen $DF(c, x)$ und $DFH(c, z)$ genannt.

Das angepaßte CGLS Verfahren lautet somit:

begin
- x_0 gegeben
- $s_0 := \binom{b - DF(c,x_0)}{\mu B d - \mu B x}$
- $p_0 := r_0 := \mathbf{Re}(DFH(c, \overline{s}_0)) + \mu B^T \underline{s}_0$
- $k := 0$
- **while** $||s_k - s_{k-1}||_2 > \epsilon_1$
  - $q_k := \binom{DF(c,p)}{\mu B p}$
  - $\alpha_k := \frac{<r_k, r_k>}{<q_k, q_k>}$
  - $x_{k+1} := x_k + \alpha_k p_k$
  - $s_{k+1} := s_k - \alpha q_k$
  - $r_{k+1} := \mathbf{Re}(DFH(c, \overline{s}_k)) + \mu B^T \underline{s}_k$
  - $\beta_k := \frac{<r_{k+1}, r_{k+1}>}{<r_k, r_k>}$
  - $p_{k+1} := r_{k+1} + \beta_k p_k$
- **end**

end

Abbildung 2.3: Angepaßtes konjugiertes Gradientenverfahren

Die Formate der einzelnen Komponenten sind in der Tabelle 2.2 aufgelistet.

$c, x, d, p, r, B^T\underline{s}, \mathbf{Re}(DFH(c,\overline{s}))$	$\in \mathbb{R}^n$
$Bd, \underline{s}$	$\in \mathbb{R}^k$
$b, DF(c,x), \overline{s}$	$\in \mathbb{C}^m$
s	$\in \mathbb{C}^{m+k}$

Tabelle 2.2: Formate der Komponenten

$\overline{s}$ sind hierbei die ersten m Elemente des Vektors s. $\underline{s}$ sind die letzten k Elemente des Vektors s.

Satz 5

Das CGLS-Verfahren minimiert

$$\min_x ||Ax - b||_2$$

über dem Krylovraum

$$\mathcal{K}_k(A^TA, A^Tb) = \text{span}\{A^Tb, (A^TA)A^Tb, \ldots, (A^TA)^{k-1}A^Tb\}.$$

Am Anfang der Iterationen werden die Komponenten mit den größten Singulärwerten einbezogen. Das bedeutet, daß ab einem bestimmten Iterationsindex gestoppt werden sollte, um zu verhindern, daß Richtungen mit kleinen Singulärwerten mit einbezogen werden, da sonst das Rauschen aus den Meßdaten die Qualität der Lösung negativ beeinflußt.

Satz 6

Falls die rechte Seite b zu $\mathcal{R}(A)$ gehört, dann konvergiert die Folge x_k des CGLS-Verfahren für $k \to \infty$ gegen A^+b.

Satz 7

Nun sei die rechte Seite b nicht in $\mathcal{R}(A)$. Dann gilt entweder $||x_k||_2 \to \infty$ mit $k \to \infty$ oder das Verfahren stoppt nach $\kappa < \infty$ Schritten. Im zweiten Fall ist $Ax_\kappa = Pb$, wobei Pb die Projektion von b auf $\mathcal{R}(A)$ ist.

Die Beweise zu Satz 5 und Satz 6 sind in [12] zu finden.

Man kann zeigen, daß die Folge $||s_k||_2$ monoton fallend ist. Des weiteren gilt, daß die Folge $||s_{k,1}||_2 = ||Ax - b||_2$ monoton fällt. Die Folge $||x_k||_2$ ist eine monoton wachsende Folge. Stoppt man frühzeitig, entspricht dies einer Regularisierung.

Das in dieser Arbeit verwendete angepaßte CGLS-Verfahren wird gestoppt, wenn zwei aufeinanderfolgende Folgenglieder $||s_k||_2$ sich nicht mehr als um 10^{-2} unterscheiden. Das Verfahren stoppt außerdem, falls sich zwei aufeinanderfolgende $||s_{k,1}||_2$ nicht mehr als 10^{-2} unterscheiden.

2.4.5 Beispiel Hilbertmatrix

Eine typisch schlecht konditionierte Matrix ist die Hilbertmatrix. Sie ist folgendermaßen definiert: $H(i,j) = \frac{1}{i+j-1}$. Die Konditionszahl von $H \in \mathbb{R}^{n\times n}$ in der 2-Norm läßt sich mit der Hilfe der Formel $K_2(H) = ||H||_2||H^{-1}||_2 = \frac{\sigma_1(H)}{\sigma_n(H)}$, wobei $\sigma_1(H)$ und $\sigma_n(H)$ der größte und kleinste Singulärwert von H ist, berechnen. Nachzulesen in [7].
Im Falle n=15 erhält man durch Singulärwertzerlegung, wie in Kapitel 2.3 beschrieben, für $\sigma_1(H) =$

1.8459 und $\sigma_{15}(H) = 2.1748 \cdot 10^{-18}$, so daß sich für die Konditionszahl $K_2(H) = 8.4880 \cdot 10^{17}$ ergibt.

Erzeugt man für $x = (1, 1, \ldots, 1)^T \in \mathbb{R}^{15}$ den Vektor $b \in \mathbb{R}^{15}$, so daß $Hx = b$ gilt und löst anschließend die Gleichung für gegebenes H und b, so erhält man die in der Tabelle 2.3 präsentierten Ergebnisse.

Inverse	Pseudoinverse	abges. SVD	Tikhonov	Landweber	CGLS
1.000000003	1.000000025	1.000000001	0.999995084	0.996907778	1.000000720
0.999997688	0.999997318	0.999999949	1.000109307	1.024981632	0.999976117
1.000159067	1.000057220	1.000000766	0.999512309	0.976729691	1.000175183
0.996177589	0.999191284	0.999995500	1.000473348	0.975174983	0.999588021
1.045325929	1.003051758	1.000011422	1.000295280	0.989341002	1.000154885
0.690664172	0.987304688	0.999989947	1.000288870	1.004467037	1.000332617
2.315187199	1.020019531	0.999994360	0.999360122	1.015485895	1.000095142
-2.616423449	0.990966797	1.000007663	0.999289883	1.021257319	0.999805755
7.450261017	0.990234375	1.000008635	0.999920050	1.022086454	0.999693341
-6.117508823	1.011474609	0.999998614	0.999846345	1.018729269	0.999795618
4.955226430	1.008300781	0.999990325	1.000587390	1.012006007	1.000028377
1.420256882	0.992919922	0.999993086	1.000638438	1.002661629	1.000256261
-1.155387887	0.987792969	1.000005148	1.000692888	0.991326160	1.000338636
2.279486575	1.015625000	1.000012573	0.999931672	0.978514167	1.000153217
0.736577538	0.995483398	0.999991976	0.999050381	0.964638016	0.999605227

Tabelle 2.3: Berechnung von x ohne gestörtes b

Für die abgeschnittene Singulärwertzerlegung wurden $\sigma_{10}, \ldots, \sigma_{15}$ auf Null gesetzt. Bei der Tikhonov Regularisierung wurde $\alpha = 10^{-6}$ gewählt. Bei dem Landweber Verfahren ist ein μ von 0.3 gewählt worden. Die Anzahl der Iterationen betrug 1.000.000 Schritte. Beim CGLS-Verfahren wurden nur 17 Iterationsschritte benötigt.

Wird auf den Vektor $b \in \mathbb{R}^{10}$, der durch Lösen von $Hx = b$ mit $x = (1, 1, \ldots, 1)^T \in \mathbb{R}^{10}$ erzeugt wurde, eine normalverteilte Störung mit Mittelwert 0 und Standardabweichung $\max\{|b|\} \cdot 10^{-6}$ addiert, so erhält man die in Tabelle 2.4 dargestellten Ergebnisse.

Inverse	Pseudoinverse	abges. SVD	Tikhonov	Landweber	CGLS
8.23644048	8.23583958	1.00031860	1.00103880	1.00145350	0.99912327
-712.70336798	-712.65172489	0.99689175	0.98307039	0.99731575	1.01005811
16600.364011244	16599.26850319	1.00641659	1.06056220	0.98715029	0.98263899
-160705.89732182	-160695.97012330	0.99891867	0.95380560	1.00110775	0.99200364
803698.31746908	803651.09147644	0.99634881	0.95697840	1.01343283	1.00570647
-2292471.16536095	-2292341.62695313	0.99806408	1.00375955	1.01762416	1.01334951
3874349.97150233	3874137.83703613	1.00076430	1.03824206	1.01377909	1.01338040
-3836215.59211013	-3836010.92019653	1.00227630	1.04045180	1.00359286	1.00685011
2055280.57528633	2055173.27627564	1.00160134	1.00954710	0.98883857	0.99531298
-459831.98771971	-459808.42039871	0.99846954	0.95183157	0.97098056	0.98019147

Tabelle 2.4: Berechnung von x mit gestörtem b^δ

Die Singulärwerte $\sigma_6, \ldots, \sigma_{10}$ sind bei der abgeschnittenen Singulärwertzerlegung auf Null gesetzt worden. Der Parameter α ist bei der Tikhonov Regularisierung auf 10^{-5} gesetzt. Beim Landweber Verfahren ist als $\mu = 0.3$ gewählt worden. Die Iterationsanzahl betrug 1.000.000 Schritte, wohingegen das CGLS nur fünf Schritte benötigte.

2.5 Nichtlineares Ausgleichsproblem

Es gibt mehrere Verfahren, die ein nichtlineares Ausgleichsproblem lösen können (siehe auch [10]). Es werden nun das Gauß-Newton Verfahren, das Levenberg-Marquardt Verfahren und Powells Dog Leg Verfahren vorgestellt.
J(c) erhält man, indem man die Matrix mit den komplexen Fréchet Ableitungen der Potentiale nach den Cole-Cole Parametern an der Stelle c nach Real- und Imaginärteil aufspaltet und untereinandersetzt. Die Differenz zwischen dem ausgewerteten Potential an der Stelle c und den gemessenen Potentialwerten ist ebenfalls komplex. Diese wird auch nach Real- und Imaginärteil aufgespalten und zu einem neuen Vektor $f(c)$ zusammengefügt, indem man die beiden Teile untereinandersetzt (siehe (1.2)).
Ferner ist

$$F(c) = \frac{1}{2}||f(c)||_2^2 = \frac{1}{2}f(c)^T f(c), \quad f : \mathbb{R}^n \to \mathbb{R}^m \tag{2.4}$$

das quadratische Fehlerfunktional.
Mit der Taylorentwicklung erhält man für das lineare Modell $l(h)$ von f in der Umgebung c

$$f(c+h) \simeq l(h) \equiv f(c) + J(c)h,$$

wobei $||h||_2$ klein ist. Setzt man dies in (2.4) ein, so erhält man für das quadratische Modell $L(h)$ von F

$$F(c+h) \simeq L(h) = F(c) + h^T J(c)^T f(c) + \frac{1}{2}h^T J(c)^T J(c)h.$$

Der Gradient von F ist:

$$g \equiv F'(c) = J(c)^T f(c).$$

Die einfachste Möglichkeit wäre es, das Abstiegsverfahren (Steepest Descent) zu benutzen, um die Iterationsfolge $c_{k+1} = c_k + \lambda_k h_k$ zu generieren. Dabei ist $h_k = -g$ und λ_k wird über eine exakte Liniensuche bestimmt. Diese Methode konvergiert nur linear und damit sehr langsam. Es ist weiterhin möglich, einfache Beispiele zu konstruieren, bei denen das Verfahren scheitert, das Minimum eines Polynoms zweiter Ordnung zu finden (siehe [10]).
Das Gauß-Newton Verfahren berechnet die Richtung als Lösung des linearen Ausgleichsproblems von $L'(h) = 0$. Damit hat man zwar schnellere Konvergenz (i.a. superlinear), aber diese ist nur lokal.
Deshalb werden Verfahren verwendet, die die Vorteile von Steepest Descent und Gauß-Newton kombinieren.
Typische Vertreter sind das Levenberg-Marquardt und das Dog Leg Verfahren. Diese berechnen sich die Abstiegsrichtungen als gemischte Kombinationen der Abstiegsrichtung von Steepest Descent und Gauß-Newton.
Beim Levenberg-Marquardt Verfahren wird dies über den Dämpfungsparameter μ erreicht, wohingegen das Dog Leg Verfahren einen Vertrauensbereichsradius Δ verwendet.

2.5.1 Gauß-Newton Verfahren

Das Gauß-Newton Verfahren basiert auf dem Modell L von F in einer Umgebung c. Gesucht ist der Gauß-Newton Schritt h_{gn}, der $L(h)$ minimiert. Da

$$L'(h) = J(c)^T f(c) + J(c)^T J(c)h$$

gilt, so erhält man h_{gn}, indem man

$$J(c)^T J(c) h_{gn} = -J(c)^T f(c)$$

löst. Dies ist äquivalent mit dem Lösen eines linearen Ausgleichsproblems. Abgesehen davon, daß es in jedem Iterationschritt gelöst werden muß, muß zusätzlich eine optimale Schrittlänge λ_k bestimmt werden, so daß $F(c_{k+1})$ möglichst klein wird. Dann kann man

$$c_{k+1} = c_k + \lambda_k h_{gn}$$

berechnen.

```
begin
    k := 0
    while ||c_neu − c||_2 > ϵ and k < k_max
        k := k + 1
        löse ||J(c)h_gn + f(c)||_2
        bestimme λ mit Hilfe einer Liniensuche
        c_neu := c + λh_gn
    end
end
```

Abbildung 2.4: Gauß-Newton Verfahren mit Liniensuche

Eine Möglichkeit ist es, $\Theta(\lambda) = F(c + \lambda h_{gn})$, $\lambda \geq 0$ zu minimieren, um F so weit wie nur möglich zu reduzieren. Das Problem hierbei ist allerdings, daß die Funktion $\Theta(\lambda)$ nichtlinear ist und das Finden eines globalen Minimierers von Θ sehr schwierig ist.
Eine andere Möglichkeit ist der folgende Backtracking Algorithmus, wobei ϱ und γ vom Benutzer vorgegeben werden. Dies ist eine sogenannte inexakte Liniensuche.

```
begin
    λ := 1
    while F(c + λh_gn) ≤ F(c) + γλ∇F(c)h_gn
        λ := ϱλ
    end
end
```

Abbildung 2.5: Backtracking Algorithmus zur Liniensuche

Hierbei wird

$$F(c + \lambda h_{gn}) \leq F(c) + \gamma \lambda \nabla F(c) h_{gn}$$

mit $\gamma \in (0, 1)$ verwendet. Diese Bedingung ist auch als die Armijo Bedingung bekannt. Hierbei wird ein Abstieg mit einem bestimmten Wert garantiert. Man beginnt mit $\lambda = 1$ und verringert es solange, bis diese Bedingung erfüllt ist.

Da es in unserem Fall aber sehr aufwendig ist, den Gradienten von F zu berechnen, wird die Funktion $\Theta(\lambda)$ minimiert. Diese wird über dem Intervall $]0, 1]$ mit Schrittweite h gesampelt. An der Stelle, an der das Minimum angenommen wird, erhält man die gesuchte Schrittweite λ.

Das Abbruchkriterium für das Gauß-Newton Verfahren ist eine kleine Änderung von c:

$$||c_{k+1} - c_k||_2 \leq \epsilon.$$

Außerdem wird ein Abbruchkriterium benötigt, um eine Endlosschleife zu verhindern, also:

$$k \geq k_{\max}.$$

2.5.2 Levenberg-Marquardt Verfahren

Das Levenberg-Marquardt Verfahren ist ein gedämpftes Newton Verfahren. Das bedeutet, daß dies eine Hybridmethode ist, die zwischen der Abstiegsmethode und der Gauß-Newton Methode wechselt. Hierbei wird die Suchrichtung h_{lm} von $c_{k+1} = c_k + h_{lm}$ nach der Formel

$$(J(c)^T J(c) + \mu I)h_{lm} = -g = -J(c)^T f(c) = -F'(c) \tag{2.5}$$

berechnet, wobei $\mu \geq 0$ ein Dämpfungsparameter ist. Er hat mehrere Effekte:

a) Für alle $\mu > 0$ ist die Matrix $J(c)^T J(c) + \mu I$ positiv definit und dies garantiert, daß h_{lm} eine Abstiegsrichtung ist. Dies folgt, wenn man Gleichung (2.5) von links mit h_{lm} multipliziert. Dann gilt: $h_{lm}^T F'(c) = h_{lm}^T g = -h_{lm}^T (J(c)^T J(c) + \mu I)h_{lm} < 0$.

b) Für große Werte μ erhalten wir $h_{lm} \simeq -\frac{g}{\mu} = -\frac{1}{\mu}F'(c)$. Dies bedeutet, daß ein kleiner Schritt in Richtung des negativen Gradientens stattfindet. Das heißt, der Parameter μ beeinflußt nun die Richtung. Dies ist gut, wenn man sich noch weit von der Lösung entfernt befindet, da man dann nicht lokal an einer Stelle viele kleine Abstiegsschritte macht.

c) Für sehr kleine Werte von μ ist $h_{lm} \simeq h_{gn}$. Diese Schritte erhält man in der Nähe des Optimums. Befindet man sich bereits in der Nähe der Lösung, dann konvergiert das Verfahren in einigen Fällen sogar quadratisch.

Der Dämpfungsparameter μ beeinflußt in diesem Verfahren sowohl die Richtung als auch die Länge der Schritte, so daß eine Liniensuche entfällt.
Die Wahl des ersten μ-Wertes ist abhängig von der Größe der Diagonalelemente in $A(c_0) = J(c_0)^T J(c_0)$. Mit folgender Formel wird man dieser Forderung gerecht:

$$\mu_0 = \tau \cdot \max_i \{a_{ii}(c_0)\}$$

Die Wahl von τ ist abhängig vom Startwert c_0. Befindet man sich bereits in der Nähe der wahren Lösung, sollte ein kleiner Wert $\approx 10^{-6}$ gewählt werden, ansonsten sollte τ im Bereich von 10^{-3} bis 1 benutzt werden.

Das Gewinnverhältnis ϱ

$$\varrho = \frac{F(c) - F(c + h_{lm})}{L(0) - L(h_{lm})} \tag{2.6}$$

beschreibt die Güte des quadratischen Modells L gegenüber dem Funktional F bei einem berechneten Schritt h_{lm}. Der Nenner ist aufgrund seiner Konstruktion immer positiv. Der Zähler ist negativ, falls der Funktionswert größer als der vorherige ist. Ein großer Wert für ϱ bedeutet, daß $L(h_{lm})$ eine gute Approximation von F ist und der Dämpfungsparameter μ reduziert werden kann. Das bedeutet, daß

der nächste Levenberg-Marquardt Schritt eher einem Gauß-Newton Schritt entspricht.
Falls ϱ klein oder sogar negativ ist, ist $L(h_{lm})$ eine schlechte Approximation von F. Das heißt, daß μ erhöht werden muß, um zu erreichen, daß der nächste Schritt eine Abstiegsrichtung ist und die Schrittlänge verkürzt wird.
Der Nenner von (2.6) kann noch umgeformt werden:

$$\begin{aligned}
L(0) - L(h_{lm}) &= ||f(c)||_2^2 - ||f(c) + J(c)h_{lm}||_2^2 \\
&= \frac{1}{2}f(c)^T f(c) - (f(c) + J(c)h_{lm})^T (f(c) + J(c)h_{lm}) \\
&= \frac{1}{2}f(c)^T f(c) - \frac{1}{2}(f(c)^T f(c) + 2h_{lm}^T J(c)^T f(c) + h_{lm}^T J(c)^T J(c)h_{lm}) \\
&= -h_{lm}^T J(c)^T f(c) - \frac{1}{2}h_{lm}^T J(c)^T J(c)h_{lm} \\
&= -\frac{1}{2}h_{lm}^T (2g + J(c)^T J(c)h_{lm}) \\
&= -\frac{1}{2}h_{lm}^T (2g + (J(c)^T J(c) + \mu I - \mu I)h_{lm}) \\
&= -\frac{1}{2}h_{lm}^T (2g + (J(c)^T J(c) + \mu I)h_{lm} - \mu h_{lm}) \\
&= -\frac{1}{2}h_{lm}^T (2g - g - \mu h_{lm}) \\
&= \frac{1}{2}h_{lm}^T (\mu h_{lm} - g)
\end{aligned}$$

Somit ergibt sich für ϱ:

$$\varrho = \frac{F(c) - F(c + h_{lm})}{\frac{1}{2}h_{lm}^T(\mu h_{lm} - g)}.$$

Eine Strategie, den Dämpfungsparameter μ zu kontrollieren, stammt von Marquardt:

$$\mu = \begin{cases} \mu \cdot 2 & \varrho < 0.25 \\ \frac{\mu}{3} & \varrho > 0.75 \end{cases}$$

Die Grenzen 0.25 und 0.75 sind hierbei so gewählt, um eine Oszillation der μ-Werte zu verhindern. Das Problem ist aber, daß in vielen Fällen die Konvergenzgeschwindigkeit sehr langsam ist.
In [10] wird eine Strategie vorgeschlagen, die dieses Problem beseitigt:

$$\mu = \begin{cases} \mu \cdot \max\{\frac{1}{3}, 1 - (2\varrho - 1)^3\}, \quad \nu = 2 & \varrho \geq 0 \\ \mu \cdot \nu, \quad \nu = \nu \cdot 2 & \text{sonst} \end{cases}$$

Der Faktor ν wird hierbei mit $\nu = 2$ initialisiert.

Ein Stoppkriterium ist, falls $||F'(c)||_\infty$ klein ist. Das heißt

$$||g||_\infty \leq \epsilon_1.$$

Ein anderes Stoppkriterium ist, wenn die Änderung in c_k sehr klein ist, also gilt:

$$||c_{k+1} - c_k||_2 \leq \epsilon_2(||c_k||_2 + \epsilon_2).$$

Außerdem wird ein Abbruchkriterium benötigt, um eine Endlosschleife zu verhindern, also:

$$k \geq k_{\max}.$$

Die Wahl von ϵ_1, ϵ_2 und $k_{\max}$ ist vom Benutzer abhängig.

begin
  $k := 0 \quad \nu := 2 \quad c = c_0$
  $A := J(c)^T J(c) \quad g := J(c)^T f(c)$
  found $:= (||g||_\infty \leq \epsilon_1) \quad \mu := \tau \cdot \max_i\{a_{ii}\}$
  while (**not** found) **and** $(k < k_{\max})$
    $k := k + 1$ löse $(A + \mu I)h_{lm} = -g$
    if $||h_{lm}||_2 \leq \epsilon_2(||c||_2 + \epsilon_2)$
      found := **true**
    else
      $c_{neu} := c + c_{lm}$
      $\varrho := \frac{F(c)-F(c+h_{lm})}{\frac{1}{2}h_{lm}^T(\mu h_{lm}-g)}$
      if $\varrho > 0$
        $c := c_{neu}$
        $A := J(c)^T J(c) \quad g := J(c)^T f(c)$
        found $:= (||g||_\infty \leq \epsilon_1)$
        $\mu := \mu \cdot \max\{\frac{1}{3}, 1 - (2\varrho - 1)^3\}$
        $\nu := 2$
      else
        $\mu := \mu \cdot \nu \quad \nu := 2 \cdot \nu$
      end
    end
  end
end

Abbildung 2.6: Levenberg-Marquardt Methode

2.5.3 Powells Dog Leg Verfahren

Die Dog Leg Methode von Powell arbeitet mit einer Kombination von Gauß-Newton Richtungen und Abstiegsrichtungen. Diese werden durch den Radius Δ eines Vertrauensbereichs kontrolliert, das heißt, daß $L(h)$ unter der Nebenbedingung $h^T h \leq \Delta^2$ minimiert wird. Die Idee, dieses Minimierungsproblem in eine Minimierung ohne Nebenbedingungen umzuformulieren, stammt von Powell.
Die Gauß-Newton Richtung h_{gn} im aktuellen Punkt c ist die kleinste Quadrate Lösung der Gleichung $J(c)h_{gn} \simeq -f(c)$. Diese kann durch Lösen der Normalengleichung

$$J(c)^T J(c)h_{gn} = -J(c)^T f(c) \tag{2.7}$$

berechnet werden.
Die Abstiegsrichtung des Steepest Descent Verfahrens h_{sd} ist durch $h_{sd} = -g = -J(c)^T f(c)$ gegeben. Da dies nur die Richtung ist und wir die Länge eines Schrittes nicht kennen, muß zusätzlich die

Funktion $F(c+\alpha h_{sd})$ nach α minimiert werden. Wegen

$$\begin{aligned} f(c+\alpha h_{sd}) &\simeq f(c)+\alpha J(c)h_{sd} \\ \Rightarrow F(c+\alpha h_{sd}) &\simeq \frac{1}{2}||f(c)+\alpha J(c)h_{sd}||_2^2 = F(c)+\alpha h_{sd}^T J(c)^T f(c)+\frac{1}{2}\alpha^2||J(c)h_{sd}||_2^2, \end{aligned}$$

ist diese Funktion für festes c und h_{sd} minimal, wenn

$$\alpha = -\frac{h_{sd}^T J(c)^T f(c)}{||J(c)h_{sd}||_2^2} = \frac{||g||_2^2}{||J(c)g||_2^2}. \tag{2.8}$$

ist.
Nun haben wir ausgehend vom aktuellen Punkt c zwei Richtungen zur Verfügung:

$$\begin{aligned} a &= \alpha h_{sd} \\ b &= h_{gn}. \end{aligned}$$

Für die Berechnung von

$$h_{dl} \tag{2.9}$$

für einen gegebenen Radius Δ hat Powell folgendes vorgeschlagen:

$$h_{dl} = \begin{cases} h_{gn} & ||h_{gn}||_2 \leq \Delta \\ \frac{\Delta}{||h_{sd}||_2} h_{sd} & ||\alpha h_{sd}||_2 \geq \Delta \\ \alpha h_{sd} + \beta(h_{gn} - \alpha h_{sd}) & \text{sonst} \end{cases}$$

β wird dabei so gewählt, daß $||h_{dl}||_2 = \Delta$ ist.
Der letzte Fall ist in Abbildung 2.7 dargestellt.

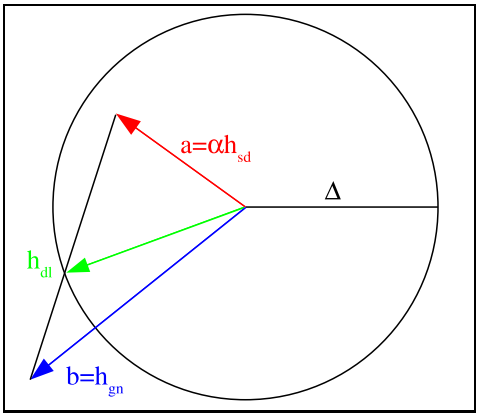


Abbildung 2.7: Vertrauensbereich und Dog Leg Schritt

Der Radius Δ schaltet also zwischen dem Abstiegsverfahren und dem Gauß-Newton Verfahren.

Die Berechnung von β mit der Vereinfachung $\Lambda = a^T(b-a)$ führt auf eine einfache Nullstellensuche.

$$\beta = \begin{cases} \frac{(-\Lambda+\sqrt{\Lambda^2+||b-a||_2^2(\Delta^2-||a||_2^2)})}{||b-a||_2^2} & \Lambda \leq 0 \\ \frac{(\Delta^2-||a||_2^2)}{(\Lambda+\sqrt{\Lambda^2+||b-a||_2^2(\Delta^2-||a||_2^2)})} & \text{sonst} \end{cases}$$

Auch hier wird wie bei der Levenberg-Marquardt Methode das Gewinnverhältnis ϱ über die Formel

$$\varrho = \frac{F(c) - F(c+h_{dl})}{L(0) - L(h_{dl})} = \frac{F(c) - F(c+h_{dl})}{-\frac{1}{2}h_{dl}^T(2g + J(c)^T J(c)h_{dl})}$$

berechnet.
Ein großer Wert zeigt an, daß das quadratische Modell $L(h_{dl})$ gut ist und der Radius Δ vergrößert werden kann, um größere Schritte zu machen, die der Gauß-Newton Richtung entsprechen.
Falls ϱ klein oder sogar negativ ist, ist $L(h_{dl})$ eine schlechte Approximation und der Radius Δ wird verkleinert, was bedeutet, daß kleine Schritte gemacht werden. Diese entsprechen dann den Abstiegsrichtungen.

Um den Radius des Vertrauensbereichs Δ zu modifizieren, benutzen wir eine Strategie, die verhindert, daß die Δ-Werte oszillieren (siehe [10]).

$$\Delta = \begin{cases} \frac{\Delta}{2} & \varrho < 0.25 \\ \max\{3 \cdot ||h_{dl}||_2, \Delta\} & \varrho > 0.75 \end{cases}$$

Im Levenberg-Marquardt Verfahren wird ϱ zum Kontrollieren des Dämpfungsparameters benutzt, beim Verfahren von Powell wird jedoch der Radius Δ des Vertrauensbereichs kontrolliert.

Die Stoppkriterien für diesen Algorithmus sind die folgenden:
Falls

$$||g||_\infty \leq \epsilon_1$$

oder

$$||f(c)||_\infty \leq \epsilon_3$$

gilt, dann ist ein globales Minimum gefunden worden.
Wenn die Änderung in c_k sehr klein ist, dann greift das Kriterium

$$||c_{k+1} - c_k||_2 \leq \epsilon_4(||c_k||_2 + \epsilon_4).$$

Ein zusätzliches Abbruchkriterium ist

$$\Delta \leq \epsilon_2(||c_k||_2 + \epsilon_2).$$

Das bedeutet, daß im nächsten Iterationsschritt das vorherige Kriterium erfüllt ist.
Das Abbruchkriterium, um eine Endlosschleife zu verhindern, ist $k \geq k_{\max}$.

Die Wahl von Δ_0, ϵ_1, ϵ_2, ϵ_3 und $k_{\max}$ ist vom Benutzer abhängig.

begin
  $k := 0 \quad c = c_0 \quad \Delta := \Delta_0 \quad g := J(c)^T f(c)$
  found := $(||f(c)||_\infty \leq \epsilon_3)$ **or** $(||g||_\infty \leq \epsilon_1)$
  while (**not** found) **and** $(k < k_{\max})$
    $k := k + 1$ berechne α nach (2.8)
    $h_{sd} := -\alpha g$ löse $J(c)h_{gn} \simeq -f(c)$ nach (2.7)
    berechne h_{dl} nach (2.9)
    if $||h_{dl}||_2 \leq \epsilon_2(||c||_2 + \epsilon_2)$
      found := **true**
    else
      $c_{neu} := c + h_{dl}$
      $\varrho := \frac{F(c)-F(c+h_{dl})}{-\frac{1}{2}h_{dl}^T(2g+J(c)^T J(c)h_{dl})}$
      if $\varrho > 0$
        $c := c_{neu} \quad g := J(c)^T f(c)$
        found := $(||b(c)||_\infty \leq \epsilon_3)$ **or** $(||g||_\infty \leq \epsilon_1)$
      end
      if $\varrho > 0.75$
        $\Delta := \max\{\Delta, 3 \cdot ||h_{dl}||_2\}$
      elseif $\varrho < 0.25$
        $\Delta := \frac{\Delta}{2} \quad found := (\Delta \leq \epsilon_2(||c||_2 + \epsilon_2))$
      end
    end
  end
end

Abbildung 2.8: Dog Leg Methode

2.5.4 Beispiel Parameterbestimmung

Es sollen die Datenpunkte $(t_1, y_1), \ldots, (t_n, y_n)$, wie in Abbildung 2.9 dargestellt, an das Modell

$$f(x,t) = x_1 \cdot e^{x_2 \cdot t} + x_3 \cdot e^{x_4 \cdot t}$$

angepaßt werden. Die Datenpunkte sind $t = [0 : 0.5 : 1]$ und $y = f([-4, -3, 4, -4], t)$. Zu minimieren ist

$$F(x) = \frac{1}{2}\sum_{i=1}^{n}(y_i - f(x, t_i))^2$$

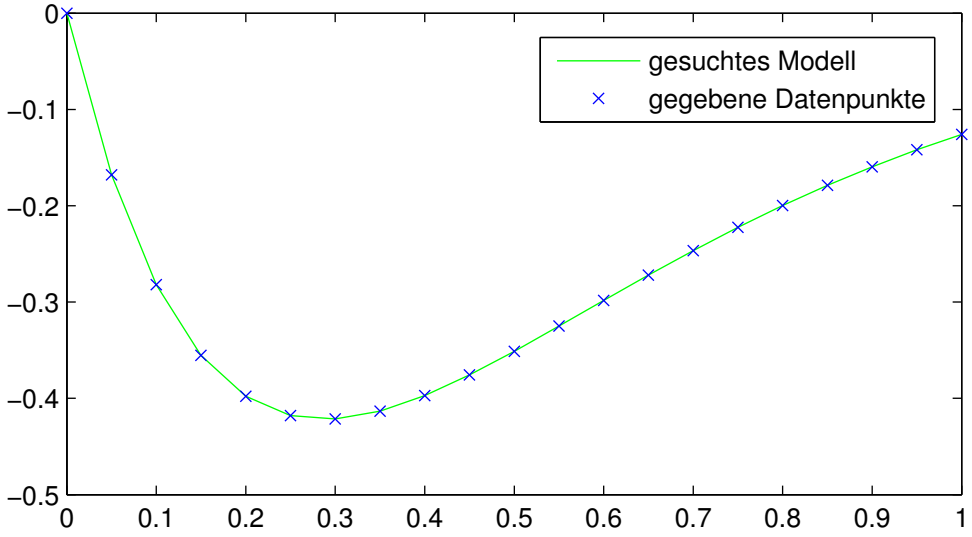


Abbildung 2.9: Datenpunkte (t_i, y_i) [×] und gesuchtes Modell $f([-4, -3.4. - 4], t)$ [—]

Bei diesen gegebenen Punkten erhalten wir nach 51 Iterationen mit dem Levenberg-Marquardt Verfahren die Parameter $x = [-4, -3, 4, -4]$, wobei die Startwerte $x = [-1, -2, 1, -1]$, $\tau = 10^{-3}$ und $\epsilon_1 = \epsilon_2 = 10^{-8}$ war. Mit dem Dog Leg Verfahren erhält man nach 49 Iterationen dieselben Parameter, wenn man mit dem gleichen Startvektor x und $\Delta = 1$ sowie $\epsilon_1 = \epsilon_2 = \epsilon_3 = 10^{-8}$ beginnt. Das Gauß-Newton Verfahren benötigt allerdings 224 Iterationen, um die Parameter zu berechnen.

2.6 Cole-Cole Modell

Die Leitfähigkeit $\sigma(\omega)$ ist umgekehrt proportional zum Widerstand $\rho(\omega)$. Es gilt also:

$$\sigma(\omega) = \frac{1}{\rho(\omega)}.$$

Es hat sich gezeigt, daß Widerstände von Böden durch die sogenannten Cole-Cole Parameter charakterisiert werden können. Damit erhält man für den Widerstand

$$\rho(\omega) = R_0 \left[1 - m\left(1 - \frac{1}{1 + (i\omega\tau)^\gamma}\right)\right].$$

Dieses klassische Cole-Cole Modell beschreibt den Frequenzgang der Leitfähigkeit eines Bodentypes. R_0 ist dabei der Gleichstromwiderstand, m die Aufladefähigkeit, γ die Frequenzabhängigkeit und τ die Zeitkonstante des Relaxationsprozesses. Die Herleitung der Cole-Cole Formel findet man in [8].

Für drei oft vorkommende Böden (siehe [2]) findet man die Parameter in Tabelle 2.5.

	R_0	m	τ	γ
polarisierbarer Boden	100	0.25	10^{-2}	0.1
steiniger Boden	1000	0.1	$10^{-1.5}$	0.4
Bruchzone	500	0.4	10	0.3

Tabelle 2.5: Cole-Cole Parameter

Für diese drei Bodentypen ist in Abbildung 2.10 der Real- und Imaginärteil des Widerstandes ρ über den Frequenzbereich 10^{-8} bis 10^{6} Hertz dargestellt. Zu sehen ist, daß der Imaginärteil stark frequenzabhängig ist, was zur Informationsgewinnung für die Inversion ausgenutzt werden kann.

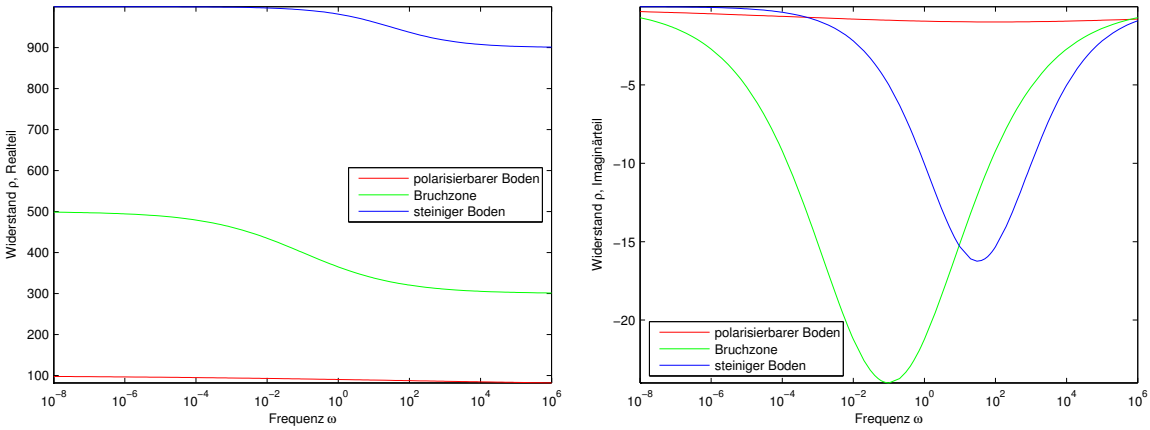


Abbildung 2.10: Widerstand der drei oft vorkommenden Bodentypen

Um möglichst viel Informationen zu gewinnen, sollten Frequenzen so gewählt werden, daß sich die Verhältnisse der Widerstände stark unterscheiden (siehe Abbildung 2.11).

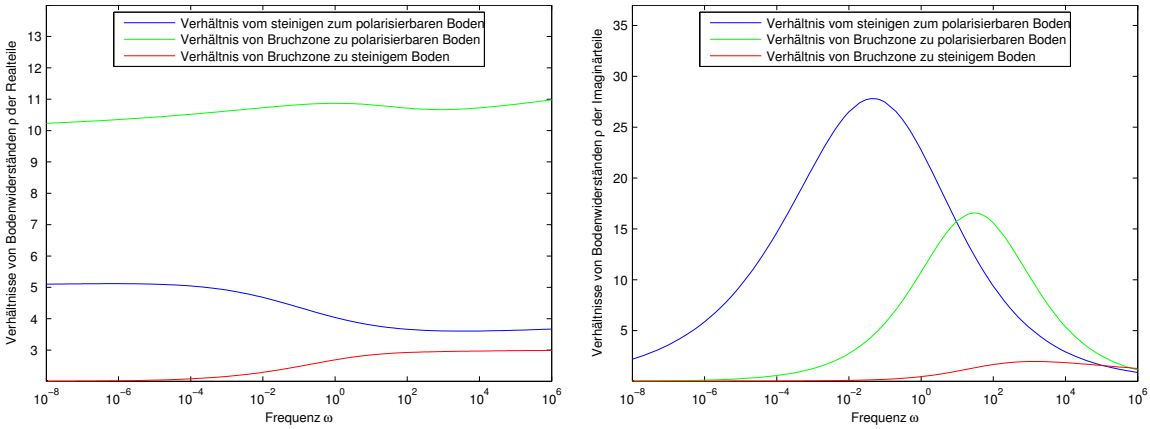


Abbildung 2.11: Verhältnisse der Widerstände der drei oft vorkommenden Bodentypen

In den Testbeispielen werden deshalb die Frequenzen $\omega = 0, 10^{-1}, 1, 10, 100$ und 1000 Hertz verwendet.

2.7 Ableitungen der Cole-Cole Funktion

Für die Berechnung der Fréchet Ableitungen des Potentials nach den Cole-Cole Parametern werden die Ableitungen der Cole-Cole Formel benötigt.

$$\partial_{R_0}\sigma(c_k) = -\frac{1}{R_0^2(1 - m(1 - \frac{1}{1+(i\omega\tau)^\gamma}))}$$

$$\partial_{m}\sigma(c_k) = \frac{1 - \frac{1}{1+(i\omega\tau)^\gamma}}{R_0(1 - m(1 - \frac{1}{1+(i\omega\tau)^\gamma}))^2}$$

$$\partial_\tau \sigma(c_k) = -\frac{1}{R_0(1-m(1-\frac{1}{1+(i\omega\tau)^\gamma}))^2}\frac{mc(i\omega\tau)^\gamma}{\tau(1+(i\omega\tau)^\gamma)^2}$$

$$\partial_\gamma \sigma(c_k) = \frac{1}{R_0(1-m(1-\frac{1}{1+(i\omega\tau)^\gamma}))^2}\frac{-m(i\omega\tau)^\gamma \ln(i\omega\tau)}{(1+(i\omega\tau)^\gamma)^2}$$

2.8 Fréchet Ableitungen

In den Iterationsschritten wird die Linearisierung der nichtlinearen Funktion f, das heißt die Approximation durch eine lineare Abbildung, benötigt. Da im kontinuierlichen Fall die Abbildung f die Cole-Cole Parameterfunktion auf Potentiale abbildet, wird der Differenzierbarkeitsbegriff auf Banachräumen gebraucht.

Satz 8 Fréchet-Differenzierbarkeit

Seien X und Y Banachräume. Eine Abbildung $f : X \to Y$ heißt Fréchet-differenzierbar an der Stelle $x_0 \in X$, falls es eine beschränkte, lineare Abbildung $f'_{x_0} : X \to Y$ gibt, so daß

$$\lim_{\|h\|_X \to 0} \frac{1}{\|h\|_X} \|f(x_0+h) - f(x_0) - f'_{x_0}h\|_Y = 0$$

Die Abbildung $f'(x_0)$ heißt Fréchet-Ableitung oder Linearisierung von f in x_0.

2.9 Implizite Funktionen

Das vorliegende Problem ist ein elliptisches Randwertproblem, was bedeutet, daß die Funktion $f(c)$ nur implizit gegeben ist. $f(c)$ hängt hierbei von $u(c)$ ab, wobei gilt:

$$-\mathrm{div}(\sigma(c,\omega)\nabla u(c)) = I \ + \ \text{Randbedingungen.}$$

Satz 9 Implizite Funktionen

Seien X, Y und Z Banachräume und $F : X \times Y \to Z$. Für $x_0 \in X$ und $y_0 \in Y$ gelte $F(x_0, y_0) = 0$. Ist die Fréchet-Ableitung $\partial_y F_{x_0,y_0}$ umkehrbar, dann ist $F(x,y) = 0$ im Punkt (x_0, y_0) eindeutig nach y auflösbar.
Es existiert also lokal um x_0 die stetige Abbildung $y : X \to Y$ mit $F(x, y(x)) = 0$.

2.10 Berechnung der Matrix mit den Fréchet Ableitungen

Die Diskretisierung des Randwertproblems $-\mathrm{div}(\sigma(c,\omega)\nabla u) = I$ + Randbedingungen sieht folgendermaßen aus (siehe [9]):

$$(-DA(c,\omega)G + B)u(c,\omega) = I, \tag{2.10}$$

wobei D die diskrete Divergenz, A die Diagonalmatrix mit den Cole-Cole Parametern, G der diskrete Gradient und B die diskreten Randbedingungen sind.
Gesucht ist die partielle Ableitung von u an der Stelle c_0 und ω in Richtung von c. Leitet man (2.10) nach der Produktregel ab, so ergibt sich:

$$-DA'(c_0,\omega)(c)Gu(c_0,\omega) - DA(c_0,\omega)Gu'(c_0,\omega)(c) + Bu'(c_0,\omega)(c) = 0.$$

Substituiert man $v = u'(c_0, \omega)(c)$, so erhält man:

$$(-DA(c_0, \omega)G + B)v = DA'(c_0, \omega)(c)Gu(c_0, \omega).$$

Benennt man nun

$$R_0 = -DA(c_0, \omega)G + B,$$

so ergibt sich für eine Spalte der Ableitungsmatrix:

$$v = R_0^{-1} DA'(c_0, \omega)(c)Gu(c_0, \omega).$$

Die direkte Berechnung von $A(c_0, \omega) \cdot x$ ist möglich durch

$$A(c_0, \omega) \cdot x = R_0^{-1} DA'(c_0, \omega)(x)Gu(c_0, \omega).$$

Die direkte Berechnung von $A(c_0, \omega)^H \cdot z$ läßt sich durch

$$A(c_0, \omega)^H \cdot z = \begin{pmatrix} (\overline{Gu(c_0, \omega)})_1 \cdot (D^H R_0^{-H} z)_1 \\ \vdots \\ (\overline{Gu(c_0, \omega)})_n \cdot (D^H R_0^{-H} z)_n \end{pmatrix}$$

realisieren. Eine ausführliche Beschreibung dieser Berechnungen sind in [9] zu finden.

Kapitel 3

Programmbeschreibung

Das Programm ist mit MATLAB Version 7 auf einem Rechner AMD Athlon XP 2600+ und 1024 MByte Arbeitsspeicher mit Betriebssystem Fedora Core release 2 (Tettnang) entwickelt worden.
Es ist folgendermaßen aufgebaut:
Zunächst werden aus einem Satz Cole-Cole Parametern gestörte Meßdaten erzeugt. Anschließend werden aus diesen Meßdaten durch Lösen des inversen Problems alle vier Cole-Cole Parameter berechnet. Eine graphische Ausgabe zeigt sowohl die Originalgleichstromwiderstände R_0 als auch die rückwärtsberechneten R_0. Außerdem werden die Singulärwerte und die aus den Iterationen gewonnenen Funktionswerte $||f(c)||_2$ graphisch dargestellt. Alle Cole-Cole Parameter werden zusätzlich auf dem Bildschirm ausgegeben.

Der Aufruf des Programmes in MATLAB lautet:
invers2D('Parameter', 'Meßprogramm', Verfahren, $\mathbf{x_r}$***,*** $\mathbf{z_r}$***,*** $\mathbf{x_g}$***,*** $\mathbf{z_g}$***, Gewichtung)***

Die Übergabeparameter des Programms haben hierbei folgende Bedeutung:

Argument	**Datentyp**	**Bedeutung**
Parameter	String	Datei, die die Cole-Cole Parameter enthält
Meßprogramm	String	Datei enthält komplettes Meßprogramm
Verfahren	Integer	1 = gedämpftes Gauß-Newton 2 = Levenberg-Marquardt 3 = Powells Dog Leg Methode
x_r	Integer	rechnerische Auflösung in x-Richtung
z_r	Integer	rechnerische Auflösung in z-Richtung
x_g	Integer	gewünschte Auflösung des Gleichstromwiderstands in x-Richtung
z_g	Integer	gewünschte Auflösung des Gleichstromwiderstands in z-Richtung
Gewichtung	Integer	Gewichtung der Matrix B mit den diskreten Gradienten

Tabelle 3.1: Argumente für den Programmaufruf

3.1 Erzeugen von gestörten Meßdaten

Das Erzeugen der gestörten Meßdaten ist unterteilt in das Einlesen der Cole-Cole Parameter, der Potentialmeßstellen, der Stromstellen und der Frequenzen, in das Lösen des Vorwärtsproblems, um das Potential zu berechnen, sowie die Störung der gemessenen Potentiale.

3.1.1 Einlesen der Cole-Cole Parameter

Die Cole-Cole Parameter und ihre Anzahl werden aus einer angegebenen Datei gelesen. Die Datei hat dabei folgenden Aufbau:

#			
dimX	dimZ		
#			
$R_{0_{1,1}}$	$m_{1,1}$	$\tau_{1,1}$	$\gamma_{1,1}$
$R_{0_{2,1}}$	$m_{2,1}$	$\tau_{2,1}$	$\gamma_{2,1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$R_{0_{x,1}}$	$m_{x,1}$	$\tau_{x,1}$	$\gamma_{x,1}$
$R_{0_{1,2}}$	$m_{1,2}$	$\tau_{1,2}$	$\gamma_{1,2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$R_{0_{x,z}}$	$m_{x,z}$	$\tau_{x,z}$	$\gamma_{x,z}$

Tabelle 3.2: Aufbau der Parameterdatei

In der ersten Zeile steht ein # als Begrenzer. Hinter diesem kann beliebiger Kommentar stehen. Es folgen die Raumdimensionen in x- und z-Richtung. Anschließend folgt ein weiteres # als Trenner. Danach werden die Cole-Cole Parameter R_0, m, τ, γ für jede Zelle definiert.

```
# Anzahl Zellen c-Gitter
1 16
# Cole-Cole Parameter
100   0.25 0.01   0.1
100   0.25 0.01   0.1
100   0.25 0.01   0.1
100   0.25 0.01   0.1
100   0.25 0.01   0.1
500   0.4  10     0.3
500   0.4  10     0.3
500   0.4  10     0.3
500   0.4  10     0.3
500   0.4  10     0.3
500   0.4  10     0.3
1000  0.1  0.0316 0.4
1000  0.1  0.0316 0.4
1000  0.1  0.0316 0.4
1000  0.1  0.0316 0.4
1000  0.1  0.0316 0.4
```

Abbildung 3.1: Beispiel für eine Parameterdatei

3.1.2 Einlesen des Meßprogramms

Das Meßprogramm enthält die Koordinaten der Potentialmeßpunkte, das Arbeitsprogramm für jede einzelne Sonde, die Koordinaten der Stromeinleitungen und die Zuordnung für die Stromeingänge und Stromausgänge und deren Höhe in Ampere, sowie die Frequenzen ω in Hertz.
Diese Daten werden aus einer Datei gelesen. Diese hat dabei folgenden Aufbau:

#				
x_1	x_2	$\cdots$	x_n	
z_1	z_2	$\cdots$	z_n	
#				
$m_{1,1}$	$m_{1,2}$	$\cdots$	$m_{1,n}$	
$\vdots$	$\vdots$		$\vdots$	
$m_{p,1}$	$m_{p,2}$	$\cdots$	$m_{p,n}$	
#				
x_1	x_2	x_3	$\cdots$	x_m
z_1	z_2	z_3	$\cdots$	z_m
#				
$s_{1,1}$	$s_{1,2}$	$s_{1,3}$	$\cdots$	$s_{1,m}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
$s_{p,1}$	$s_{p,2}$	$s_{p,3}$	$\cdots$	$s_{p,m}$
#				
ω_1	$\cdots$	ω_k		

Tabelle 3.3: Aufbau der Meßprogrammdatei

In der ersten Zeile steht das Zeichen # als Begrenzungszeichen. Danach folgt eine Zeile, in der alle x-Koordinaten der Potentialmeßstellen stehen. Diese Werte liegen im Intervall $[0, 1]$. In der nächsten Zeile stehen die z-Koordinaten. Nach einem weiteren Begrenzer folgt in jeder Zeile das Meßprogramm, das heißt, hier wird angegeben, ob eine Meßstelle mißt oder nicht.
Nach einem weiteren Begrenzer folgen die x- und z-Koordinaten der Stromquellen, gefolgt von einem # als Trenner. Anschließend wird angegeben, welche Menge Strom in Ampere injiziert bzw. abgeleitet wird.
Nach einem #, folgen die Frequenzen ω in Hertz.

Hinweise:

Es sollte für die Potentialmeßstellen und Stromquellen ein Abstand zum Dirichlet Rand von 0.2-0.3 Einheiten eingehalten werden. Zum Rand hin wird das Potential auf Null gesetzt, was nicht der Realität entspricht (künstliche Randbedingung).

Der eingeleitete Strom sollte genauso groß sein, wie der abfließende Strom, um die Realität möglichst gut beschreiben zu können.

An den Stellen, an denen Strom ein- und ausgeleitet werden, sollte nicht das Potential gemessen werden, da dort das Potential singulär und somit unbeschränkt ist. Die Elektroden können entweder das Potential messen oder Ströme ein- und ausleiten.

```
# Position der Potential-Messpunkte
0.3 0.7  0.3  0.7 0.3 0.5 0.7
0.3 0.3 0.65 0.65 1.0 1.0 1.0
# Messprogramm P
0  0  1  1  1  1  1
0  1  0  1  1  1  1
0  1  1  0  1  1  1
0  1  1  1  0  1  1
0  1  1  1  1  0  1
0  1  1  1  1  1  0
# Position der Deltas
0.3 0.7  0.3  0.7 0.3 0.5 0.7
0.3 0.3 0.65 0.65 1.0 1.0 1.0
# Stroeme J in Ampere
-1  1  0  0  0  0  0
-1  0  1  0  0  0  0
-1  0  0  1  0  0  0
-1  0  0  0  1  0  0
-1  0  0  0  0  1  0
-1  0  0  0  0  0  1
# Frequenzen Omega in Hertz
0 10 100
```

Abbildung 3.2: Beispiel für eine Meßprogrammdatei

3.1.3 Generierung der Potentiale

Zur Generierung der Potentiale wird die Funktion F(c) aufgerufen (siehe [9]). Diese liefert das Potential über alle Stromkonfigurationen und Frequenzen ω, ausgewertet an den angegebenen Meßstellen, das heißt:

$$(u_{11}(f_1,\omega_1),\cdots,u_{n1}(f_1,\omega_1),u_{12}(f_2,\omega_1),\cdots,u_{n2}(f_2,\omega_1),\cdots,u_{1m}(f_m,\omega_1),u_{nm}(f_m,\omega_1),u_{11}(f_1,\omega_2),\cdots$$
$$\cdots,u_{nm}(f_m,\omega_2),\cdots,u_{nm}(f_m,\omega_p))^T$$

3.1.4 Störung der Messung

Nachdem man nun das Potential an den gemessenen Punkten erhalten hat, wird dieses mit einer Störung versehen, um die Ungenauigkeit der Messung zu simulieren. Dazu werden normalverteilte Zufallsvariablen mit Mittelwert $\mu = 0$ und Standardabweichung $\sigma = \max\{|u_j|\}\cdot 1‰$ erzeugt und auf die gemessenen Potentiale addiert.
Normalverteilte Zufallszahlen erhält man durch den Box-Muller Algorithmus (siehe [11]). Für zwei unabhängige auf dem Einheitsintervall gleichverteilte Zufallsvariablen U, V kann man zwei unabhängige standardnormalverteilte Zufallsvariablen X, Y erzeugen, indem man $X = \sqrt{-2\log U}\cos(2\pi V)$ und $Y = \sqrt{-2\log(U)}\sin(2\pi V)$ berechnet.
Benötigt man normalverteilte Zufallszahlen mit Mittelwert μ und Varianz σ^2, so transformiert man die standardnormalverteilte Zufallsvariable mit der Formel

$$Z = \sigma X + \mu.$$

um. Z ist dann normalverteilt mit Mittelwert μ und Varianz σ^2.

3.2 Lösen des inversen Problems

Das inverse Problem kann optional mit Hilfe des gedämpften Gauß-Newton, mit dem Levenberg-Marquardt Verfahren oder mit Powells Dog Leg Methode gelöst werden.

3.2.1 Startvektor

Bevor ein Iterationsverfahren gestartet werden kann, benötigt man einen geeigneten Startvektor. In den Untersuchungen wurde als Startvektor der homogene Boden des Typs „Bruchzone“ angenommen.

$$c_{Start} = (500, 0.4, 10, 0.3, 500, 0.4, 10, 0.3, \ldots, 500, 0.4, 10, 0.3)^T$$

3.2.2 Iterationsverfahren

Alle drei Verfahren benötigen zum Lösen des nichtlinearen Ausgleichsproblems ein Verfahren, das ein lineares Ausgleichsproblem löst. Hierbei wird ausschließlich das konjugierte Gradientenverfahren verwendet. Eine Erläuterung für diese Entscheidung findet sich im nachfolgenden Kapitel.

Außerdem wird ein einfaches Verfahren verwendet, damit die unterschiedliche Skalierung der zu bestimmenden Parameter keinen Einfluß auf die Qualität hat. Zunächst wird eine neue Suchrichtung für R_0 bestimmt. Dann wird diese für m, τ und γ berechnet. Danach fängt man wieder bei R_0 an. Wendet man dieses Verfahren nicht an, dann werden die Parameter m, τ und γ stärker gewichtet und die Änderungen von R_0 sind sehr gering. Das bedeutet, daß das Iterationsverfahren selbst nach einer großen Anzahl von Iterationen noch nicht in die Nähe einer brauchbaren Approximation gekommen ist.

Da die drei Iterationsverfahren einfache lineare Nebenbedingungen, die in Tabelle 3.4 aufgelistet sind, beachten sollen, ist es notwendig die neue Iterierte auf den Rand zu projizieren, falls ein Verlassen des zulässigen Bereichs droht. Nach der Projektion befindet sich die neue Iterierte auf dem Rand, der von den Nebenbedingungen aufgespannt wird.

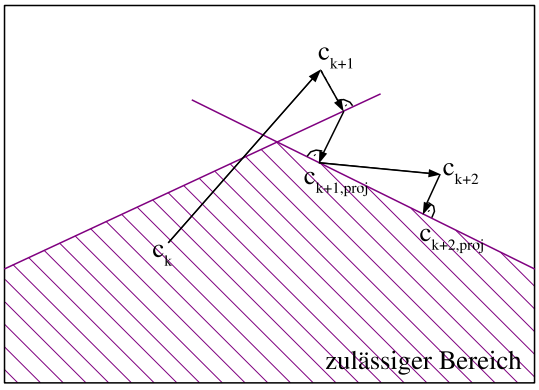


Abbildung 3.3: Projektion der neuen Iterierte

Die Grenzen dieses Bereichs ergeben sich als Minima und Maxima der drei häufig vorkommenen Bodentypparameter.

100	$\leq$	R_0	$\leq$	1000
0.1	$\leq$	m	$\leq$	0.4
0.01	$\leq$	τ	$\leq$	10
0.1	$\leq$	c	$\leq$	0.4

Tabelle 3.4: Nebenbedingungen

Für alle drei Verfahren ist die maximale Anzahl der Iterationen beschränkt. Der Grund liegt darin, daß nach einer hinreichend großen Anzahl an Schritten der Informationsgewinn vernachlässigbar klein wird.
In den verwendeten Verfahren wurden folgende Parameter benutzt:

- Gauß-Newton: $k_{\max} = 100, \epsilon = 10^{-4}$
- Levenberg-Marquardt: $k_{\max} = 100, \epsilon_1 = \epsilon_2 = 10^{-4}$, $\mu = \frac{1}{2}$
- Powells Dog Leg: $k_{\max} = 100, \epsilon_1 = \epsilon_2 = \epsilon_3 = 10^{-4}, \Delta_0 = 100$

3.2.3 Strafterm

Es wird in jedem Iterationsschritt ein lineares Ausgleichsproblem $||Ac - b||_2^2$ gelöst. Dieses ist schlecht gestellt und es muß regularisiert werden. Um hohe Variationen in der Approximation c zu bestrafen, ist es möglich, eine Matrix B zu verwenden, die den diskreten Gradienten auf dem Gitter, auf dem die Cole-Cole Parameter berechnet werden, enthält. Das zu lösende Ausgleichsproblem lautet dann $||Ac - b||_2^2 + \alpha^2 ||B(c - d)||_2^2$.
Der Vektor d kann hierbei als Strafterm benutzt werden, um Lösungen, die weit von d entfernt liegen, zu bestrafen.

Das Aufstellen der Matrix B wird zur Vereinfachung an der Gittereinteilung $x = 3$ und $z = 2$ erklärt. Bei dieser ist somit die Schrittweite $h_x = \frac{1}{x-1} = 0.5$ und $h_z = \frac{1}{z-1} = 1$. Es gilt somit für X:

$$X = \begin{pmatrix} -\frac{1}{h_x} & \frac{1}{h_x} & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{h_x} & \frac{1}{h_x} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{h_x} & \frac{1}{h_x} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{h_x} & \frac{1}{h_x} \\ -\frac{1}{h_z} & 0 & 0 & \frac{1}{h_z} & 0 & 0 \\ 0 & -\frac{1}{h_z} & 0 & 0 & \frac{1}{h_z} & 0 \\ 0 & 0 & -\frac{1}{h_z} & 0 & 0 & \frac{1}{h_z} \end{pmatrix} = \begin{pmatrix} -2 & 2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 & -2 & 2 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix}$$

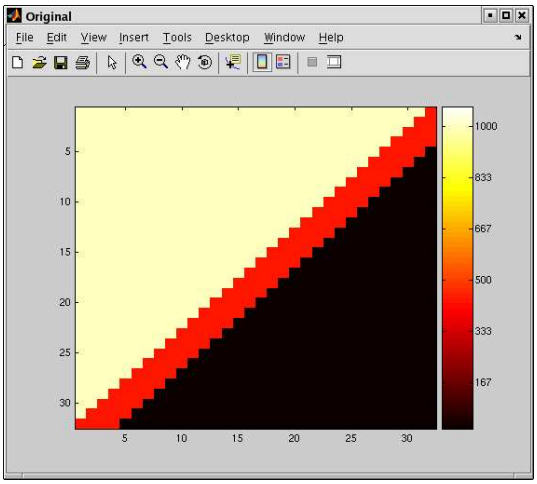
Da diese diskrete Ableitung auf alle 4 Cole-Cole Parameter angewandt wird, erhalten wir schließlich:

$$B = \begin{pmatrix} X & 0 & 0 & 0 \\ 0 & X & 0 & 0 \\ 0 & 0 & X & 0 \\ 0 & 0 & 0 & X \end{pmatrix}$$

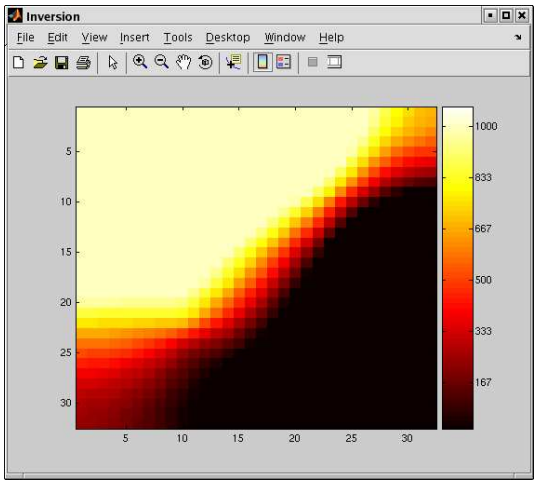
3.3 Ausgabe des Programms

Die rückwärts berechneten Gleichstromwiderstände R_0 werden graphisch angezeigt. Zusätzlich werden die eingelesenen Gleichstromwiderstände R_0 ausgegeben. Ferner werden die Singulärwerte der Ableitungsmatrix an der Stelle der identifizierten Cole-Cole Parameter dargestellt. Außerdem werden die

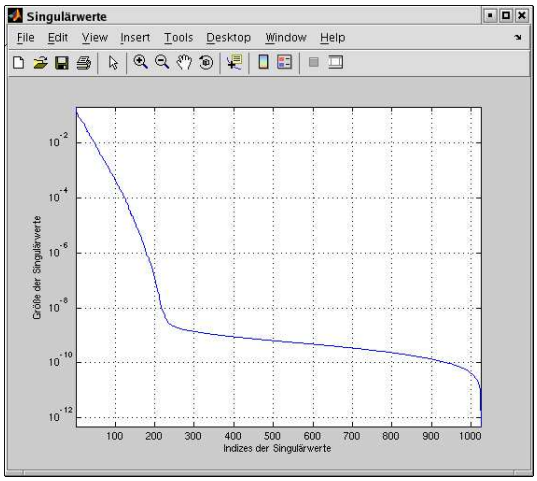
in den Iterationen gewonnenen Funktionswerte $||f(c)||_2$ gezeichnet, um nachvollziehen zu können, was das Iterationsverfahren geleistet hat.
Diese Ausgaben werden nicht nur graphisch dargestellt, sondern es folgt zusätzlich eine Ausgabe auf dem Bildschirm, bei der zusätzlich noch die Parameter m, τ und γ ausgegeben werden.



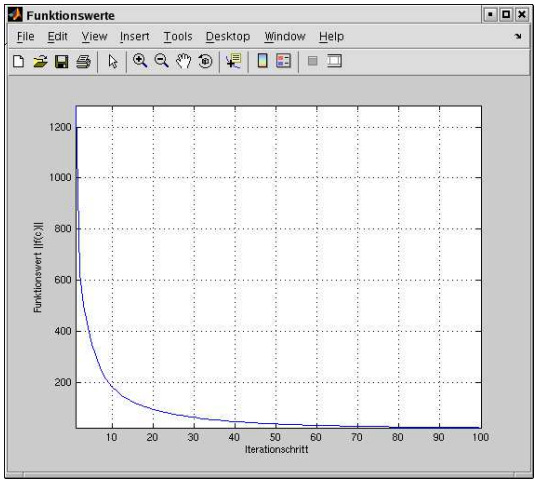
(a) Originalparameter R_0



(b) Berechnete R_0



(c) Singulärwerte



(d) Funktionswerte

Abbildung 3.4: Beispiel einer Ausgabe

Kapitel 4

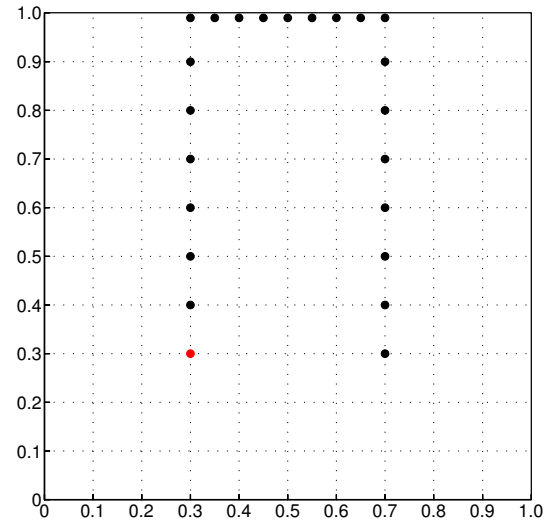
Testbeispiele und Ergebnisse

Zunächst stellte sich die Frage, ob direkte oder iterative Verfahren zur Lösung der linearen Ausgleichsprobleme zu bevorzugen sind. In diesem Fall sind die iterativen den direkten Verfahren überlegen, und zwar aus folgendem Grund:
Bei direkten Verfahren wie der Tikhonov Regularisierung oder der abgeschnittenen Singulärwertzerlegung benötigt man die vollständige Ableitungsmatrix A. Diese zu berechnen, bedeutet bei n gesuchten Parametern, n Vorwärtsprobleme zu lösen. Die iterativen Verfahren benutzen hingegen Matrix-Vektor Produkte Ax und $A^H z$, so daß eine komplette Ableitungsmatrixberechnung entfällt. Die Auswertung eines dieser Produkte entspricht der Lösung eines Vorwärtsproblems. Das CGLS-Verfahren braucht im Durchschnitt 40 Schritte. Abgesehen davon, daß die iterativen Verfahren bevorzugt gewählt werden sollten, sei hier noch erwähnt, daß die Erzeugung einer singulären Wertezerlegung einer sehr großen Matrix, wie es hier der Fall ist, sehr aufwendig ist und hohen Speicherbedarf benötigt.
Nun stellt sich die Frage, welches der iterativen Verfahren die beste Wahl darstellt. Einfache Untersuchungen haben ergeben, daß das Landweber Verfahren für bestimmte Relaxationsparameter über 100 Iterationsschritte benötigt, um eine schlechte Approximationsqualität zu erhalten. Wünscht man hingegen eine gute Approximationsqualität, so werden in einigen Fällen 1000 Schritte benötigt. Deshalb ist das beste Verfahren für dieses Problem das speziell angepaßte konjugierte Gradientenverfahren (siehe 2.3). Daher wird in allen Verfahren, die das nichtlineare Ausgleichsproblem lösen, das konjugierte Gradientenverfahren verwendet.

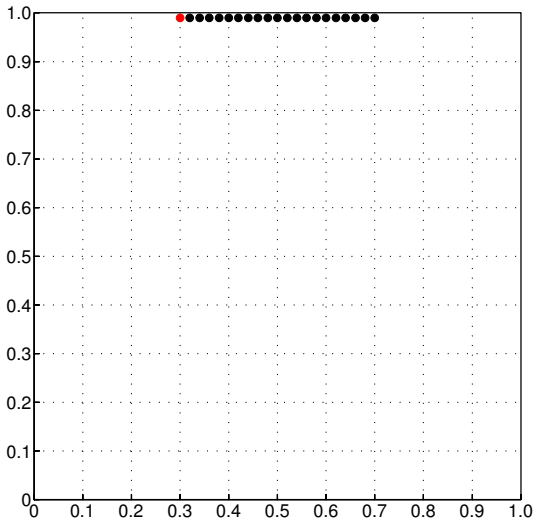
4.1 Verwendete Meßprogramme

In den nachfolgenden Tests sind die in Abbildung 4.1 dargestellten Meßprogramme verwendet worden. Im ersten Meßprogramm sind die Elektroden • sowohl im Boden als auch auf der Oberfläche verteilt. Hier fließt der Strom aus der linken unteren Elektrode • ab. An den restlichen Elektroden wird reihum einzeln Strom injiziert. An den übrigen werden die Potentialwerte gemessen. Bei diesem Programm sind 23 Elektroden verwendet worden.
Im zweiten Meßprogramm sind die Elektroden nur auf der Oberfläche verteilt. Aus der linken Elektrode fließt der Strom raus und in die anderen wird nacheinander Strom injiziert. Dort, wo kein Strom rein- oder rausfließt, wird das Potential gemessen. Die Anzahl der Elektroden beträgt 21.
Das dritte Meßprogramm arbeitet folgendermaßen: An der linken unteren Elektrode • fließt der Strom ab. An den restlichen Elektroden wird reihum der Strom injiziert und an den übrigen Elektroden das Potential gemessen. Nach diesem Durchgang ist die nächste Elektrode • die Stromauslaßstelle und es

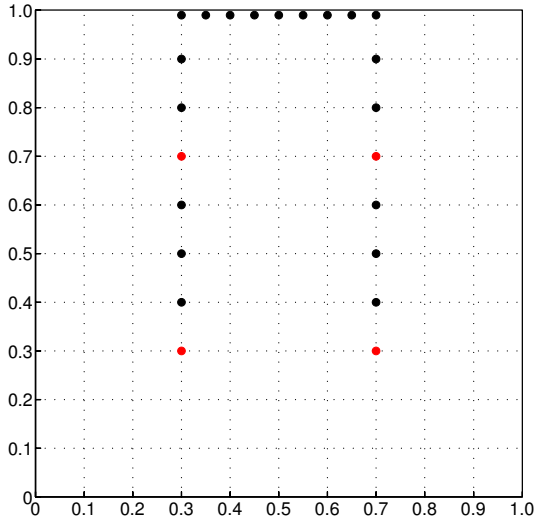
geschieht wieder das Injizieren und das Messen wie zuvor beschrieben. Insgesamt wird dies solange fortgesetzt bis alle roten Elektroden als Stromablaßstelle benutzt worden sind.
Hierbei wurde ein Abstand von 0.3 Einheiten vom Dirichlet Rand eingehalten.



(a) Meßprogramm 1 mit 21 Elektroden •



(b) Meßprogramm 2 mit 23 Elektroden •



(c) Meßprogramm 3 mit 21 Elektroden •

Abbildung 4.1: Verwendete Meßprogramme

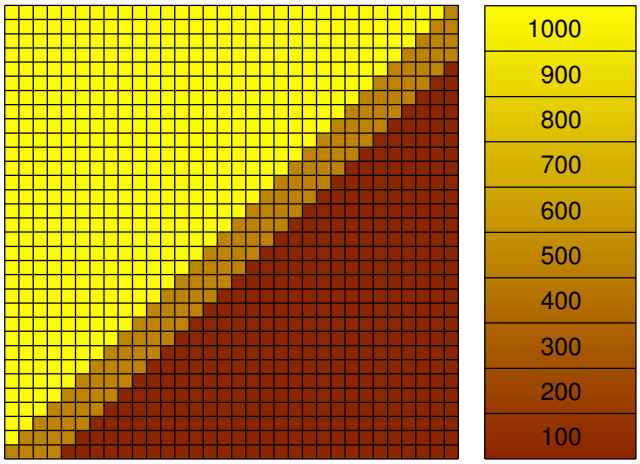
Das Vorwärtsproblem wird mit einer Auflösung des Potentials u von 100×100 gelöst. Die gemessenen Potentialwerte werden anschließend, wie in Kapitel 3.1.4 beschrieben, gestört.

In den nachfolgenden Tests stehen unter den Abbildungen das verwendete Verfahren, die rechnerische Auflösung, die gewünschte Auflösung der Cole-Cole Parameter für die Inversion, die Gewichtung der diskreten Ableitungsmatrix B und das verwendete Meßprogramm. Die Bilder zeigen nur den rückwärtsberechneten Gleichstromwiderstand R_0, da er die größte Aussagekraft besitzt. Die anderen Cole-Cole Parameter werden ebenfalls bestimmt, aber in den nachfolgenden Tests nicht dargestellt.

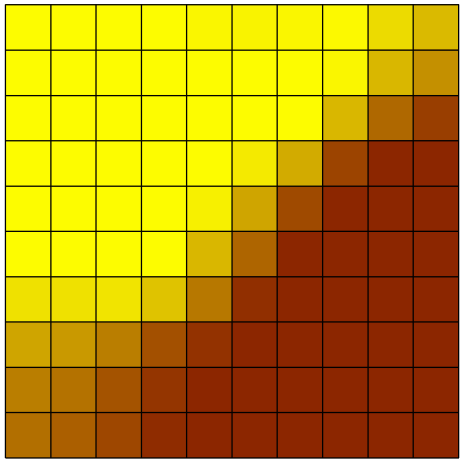
4.2 Vergleich der unterschiedlichen Verfahren

In diesem Test werden die unterschiedliche Verfahren miteinander verglichen. Getestet werden das Gauß-Newton, das Levenberg-Marquardt und das Dog Leg Verfahren. Dieser Vergleich geschieht zum einen mit Gleichstrom und zum anderen mit Wechselstrom.
Hierbei wird das Meßprogramm 1 zur Berechnung benutzt, wobei die rechnerische Auflösung mit 50×50 gegeben ist. Alle Verfahren sind durch $k_{\max} = 100$ gestoppt worden. Als Faktor für die Matrix B wird 1 benutzt. Es sollen $10 \times 10 = 100$ Parameter bestimmt werden.

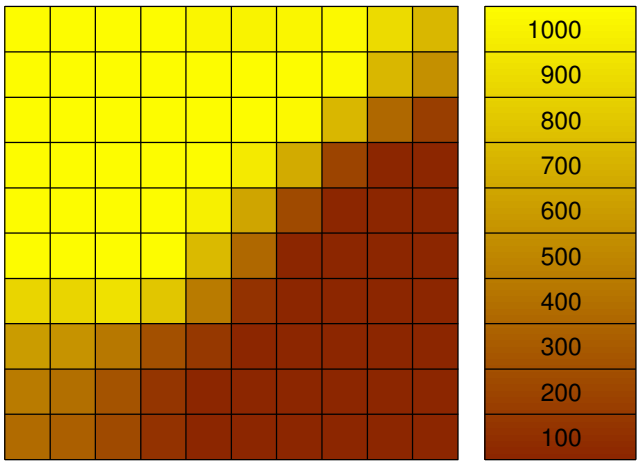
Vergleich mit Gleichstrom:
In diesem Test wird die Frequenz ω=0 verwendet.



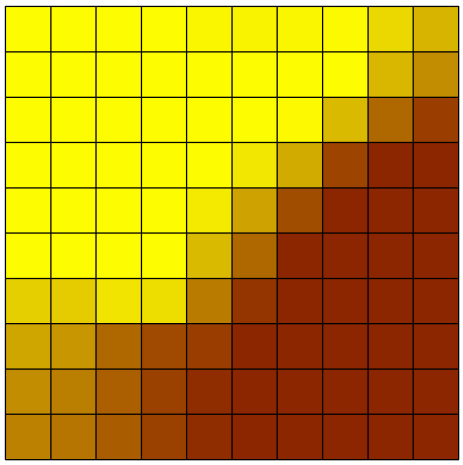
(a) Original 32×32



(b) Gauß-N. 50×50, 10×10, B=1, MP=1



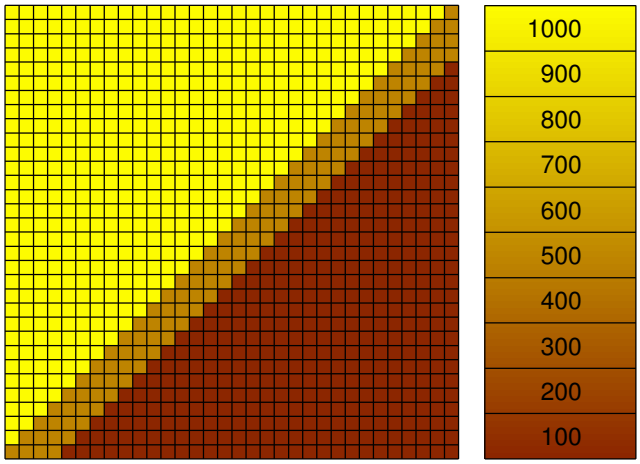
(c) Levenberg-M. 50×50, 10×10, B=1, MP=1



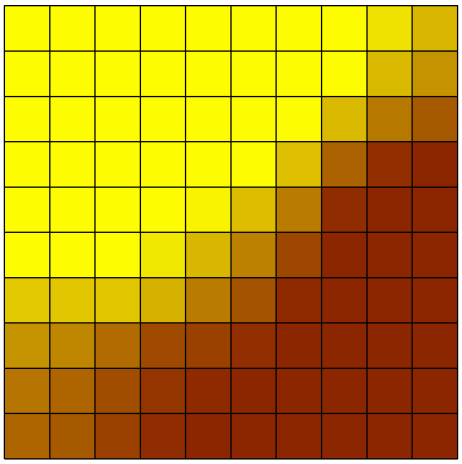
(d) Dog Leg 50×50, 10×10, B=1, MP=1

Ergebnisse:
Alle Verfahren erzeugen bei Gleichstrom vergleichbare Ergebnisse. Das Gauß-Newton Verfahren ist allerdings aufwendiger als das Levenberg-Marquardt und das Dog Leg Verfahren, da noch eine zusätzliche Liniensuche durchgeführt wird. Deshalb wird bei der Verwendung von Gleichstrom empfohlen, das Levenberg-Marquardt oder das Dog Leg Verfahren zu verwenden.

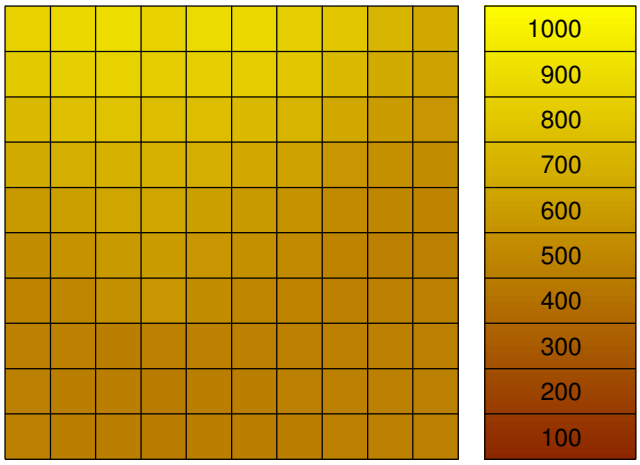
Vergleich mit Wechselstrom:
In diesem Test wird die Frequenz ω=1 benutzt.



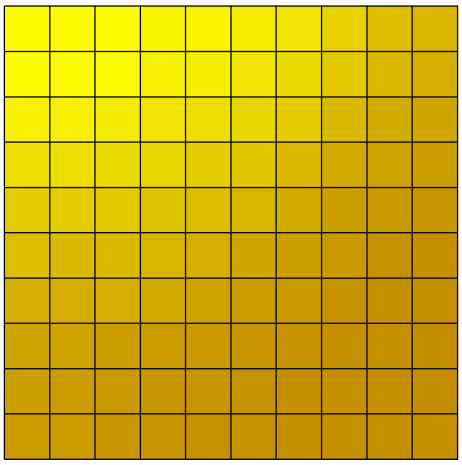
(a) Original 32×32



(b) Gauß-N. 50×50, 10×10, B=1, MP=1



(c) Levenberg-M. 50×50, 10×10, B=1, MP=1



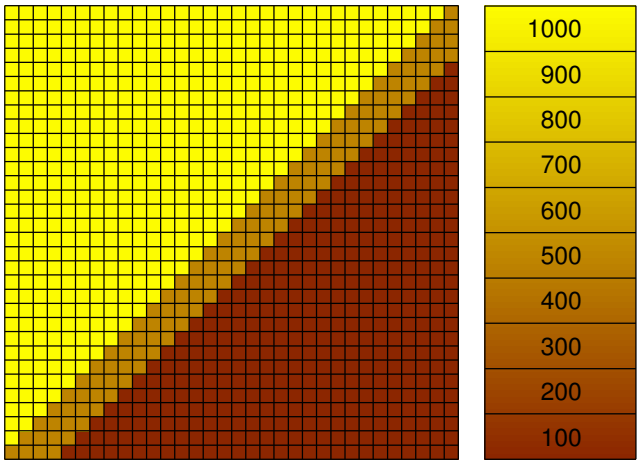
(d) Dog Leg 50×50, 10×10, B=1, MP=1

Ergebnisse:
In diesem Vergleich zeigt sich, daß das Gauß-Newton Verfahren dem Levenberg-Marquardt und dem Dog Leg Verfahren überlegen ist. Die beiden letzten Verfahren konvergieren extrem langsam, da die Cole-Cole Parameter m, τ und γ einen negativen Einfluß auf das Gewinnverhältnis ϱ haben. Deshalb sollte bei der Verwendung von Wechselstrom das Gauß-Newton Verfahren benutzt werden.

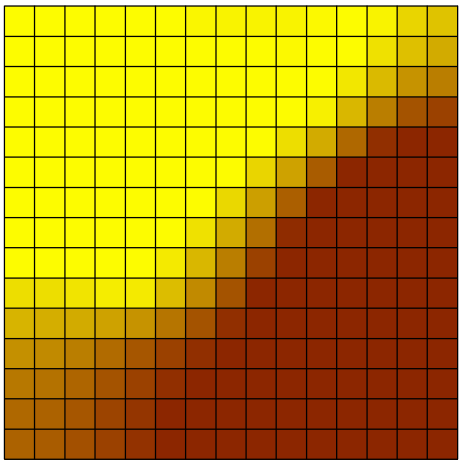
Es stellt sich nun die Frage, ob die Anzahl der zu identifizierenden Parameter auf 25×25 gesteigert werden kann, um eine kontrastreichere Auflösung des Bodens zu erhalten.

4.3 Erhöhung der Auflösung

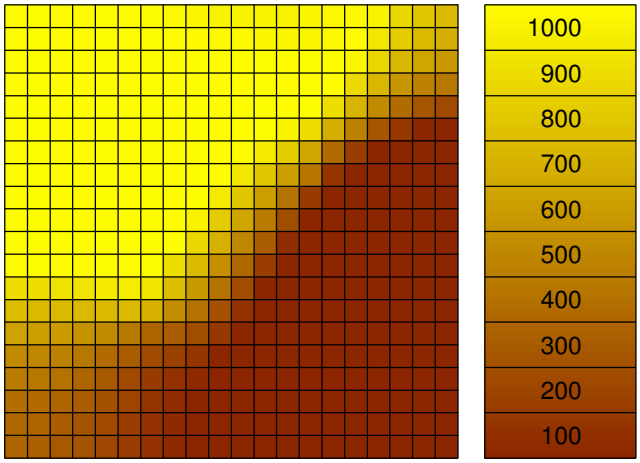
In diesem Test wird demonstriert, daß eine große Anzahl an Parametern identifiziert werden kann, obwohl nur wenige Potentialmeßwerte zur Verfügung stehen. Beim ersten Test werden 15×15, beim zweiten Test 20×20 und beim dritten Test 25×25 Parameter bestimmt.
Es wird mit einer Diskretisierung von 50×50 unter Verwendung des Meßprogramms 1 gerechnet. Für diesen Test ist das Levenberg-Marquardt Verfahren verwendet worden, das nach jeweils 100 Iterationen durch $k_{\max}$ abgebrochen wurde. Für die Gewichtung der diskreten Gradientenmatrix B ist der Faktor 1 verwendet worden. Es ist mit Gleichstrom gerechnet worden.



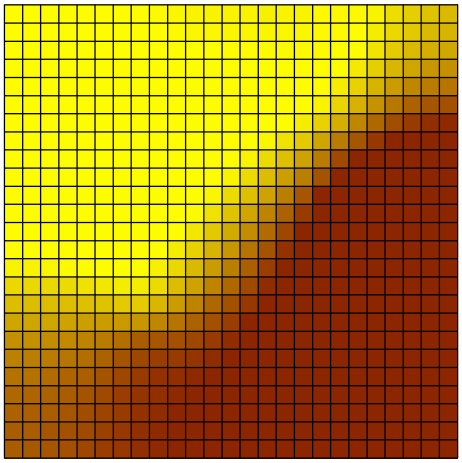
(a) Original 32×32



(b) Levenberg-M. 50×50, 15×15, B=1, MP=1



(c) Levenberg-M. 50×50, 20×20, B=1, MP=1



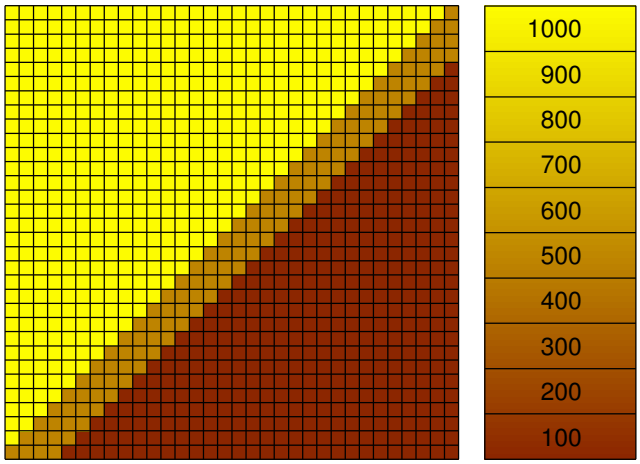
(d) Levenberg-M. 50×50, 25×25, B=1, MP=1

Ergebnisse:
Es kann festgehalten werden, daß auch eine große Anzahl an Parametern bestimmt werden kann. Das heißt, daß auch bei starker Unterbestimmtheit das Verfahren stabil bleibt. Dies wird durch die Verwendung der Matrix B erreicht, weil man mit dieser die Möglichkeit hat, hohe Variationen der Cole-Cole Parameter zu unterbinden.
Die linke untere Ecke kann nicht so gut identifiziert werden, was durch die Anordnung der Elektroden erklärt werden kann, da dort kein Strom durchfließt.

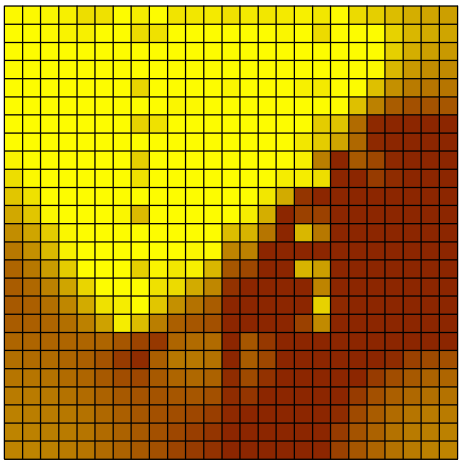
Deshalb stellt sich die Frage, wie groß der Einfluß der Matrix B ist.

4.4 Einfluß des diskreten Gradienten

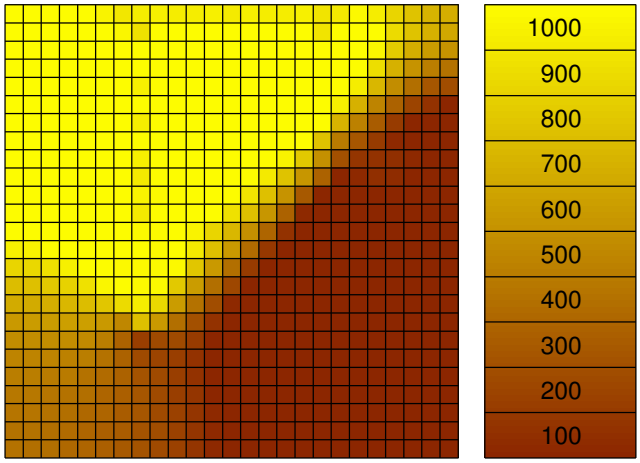
Hier wird nach einem optimalen Faktor für die Gewichtung der Matrix B gesucht. Die Lösungen sollen in den jeweiligen homogenen Teilbereichen möglichst glatt sein. Dies hat zur Folge, daß mehr Iterationen benötigt werden, da starke Variationen in Strukturen nur langsam identifiziert werden.
Es wird mit einer Diskretisierung von 50×50 mit dem Meßprogramm 1 gerechnet, wobei das Dog Leg Verfahren, das nach 100 Iterationen durch $k_{\max}$ abbricht, verwendet wird. Als Faktor für die Matrix B werden 0, 0.1, 0.3, 0.5, 0.9, 2 und 5 benutzt. Es werden $25 \times 25 = 625$ Parameter gesucht. Es wird hierbei nur mit Gleichstrom gerechnet.



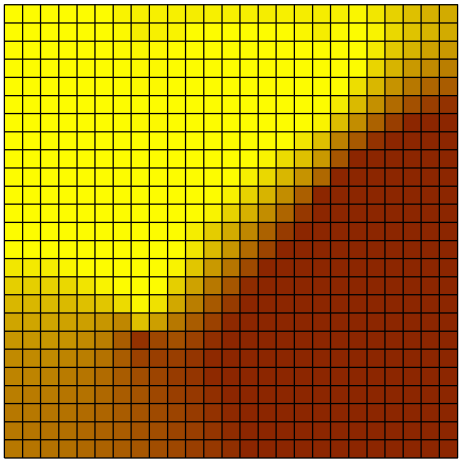
(a) Original 32×32



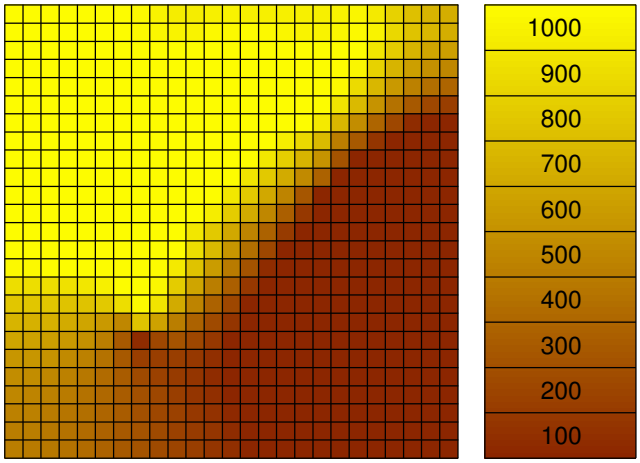
(b) Dog Leg 50×50, 25×25, B=0, MP=1



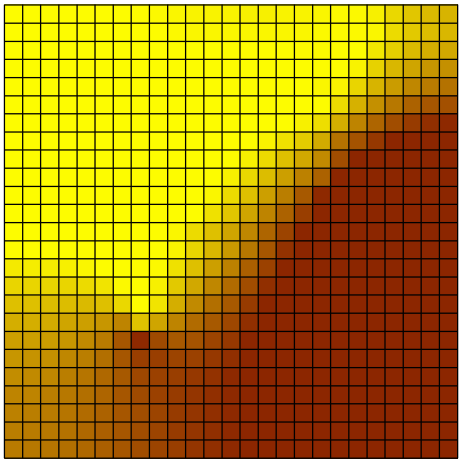
(c) Dog Leg 50×50, 25×25, B=0.1, MP=1



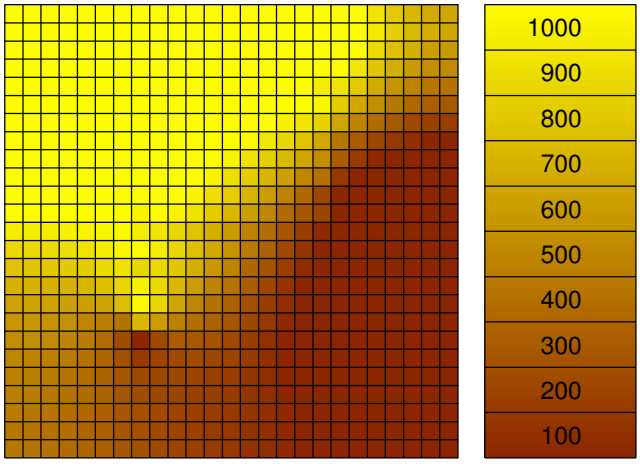
(d) Dog Leg 50×50, 25×25, B=0.3, MP=1



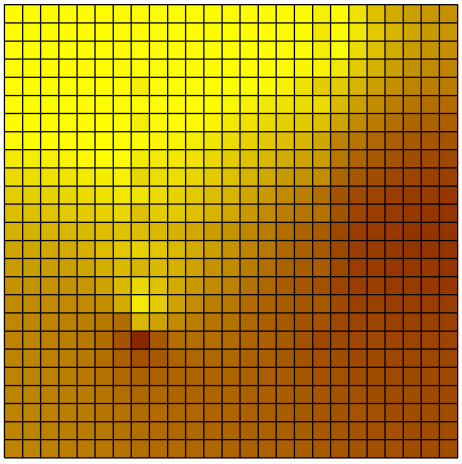
(e) Dog Leg 50 × 50, 25 × 25, B=0.5, MP=1



(f) Dog Leg 50 × 50, 25 × 25, B=0.9, MP=1



(g) Dog Leg 50 × 50, 25 × 25, B=2, MP=1



(h) Dog Leg 50 × 50, 25 × 25, B=5, MP=1

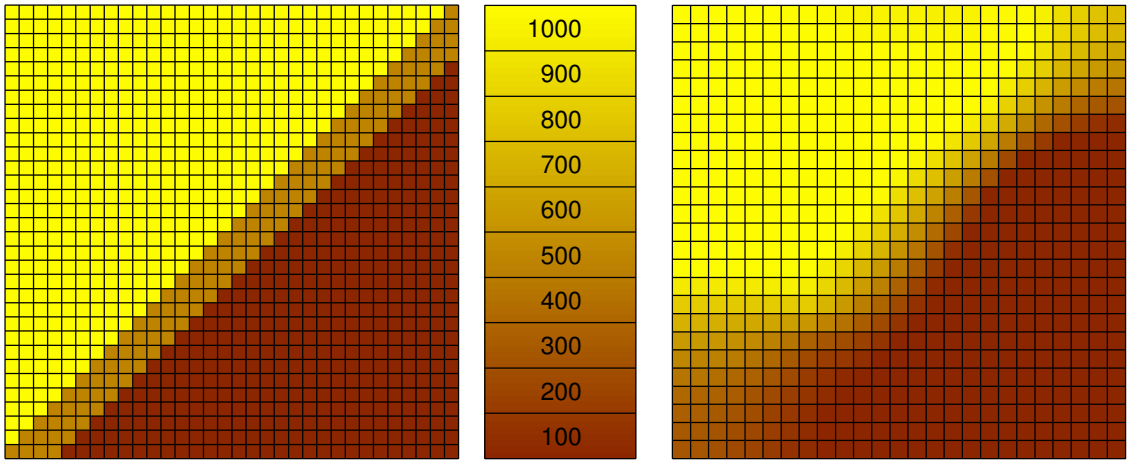
Ergebnisse:

Die Matrix B sollte benutzt werden, da sonst Ergebnisse wie in (b) auftreten. Wird der Faktor zu groß gewählt, so konvergiert das Verfahren sehr langsam, was beispielsweise bei (h) der Fall ist. Als gute Faktoren ergeben sich in diesem Fall 0.1, 0.3, 0.5, 0.9 und 1.0 (siehe vorheriger Test). Bei Verwendung von kleinen Faktoren, wie 0.1 und 0.3 sind die Übergänge stärker, während bei 0.9 und 1 die Übergänge sehr breit sind. Das bedeutet, daß man für die Identifizierung von kleinen lokalen Strukturen 0.1 verwenden sollte, wohingegen man bei globalen Strukturen den Faktor 1 wählen sollte.

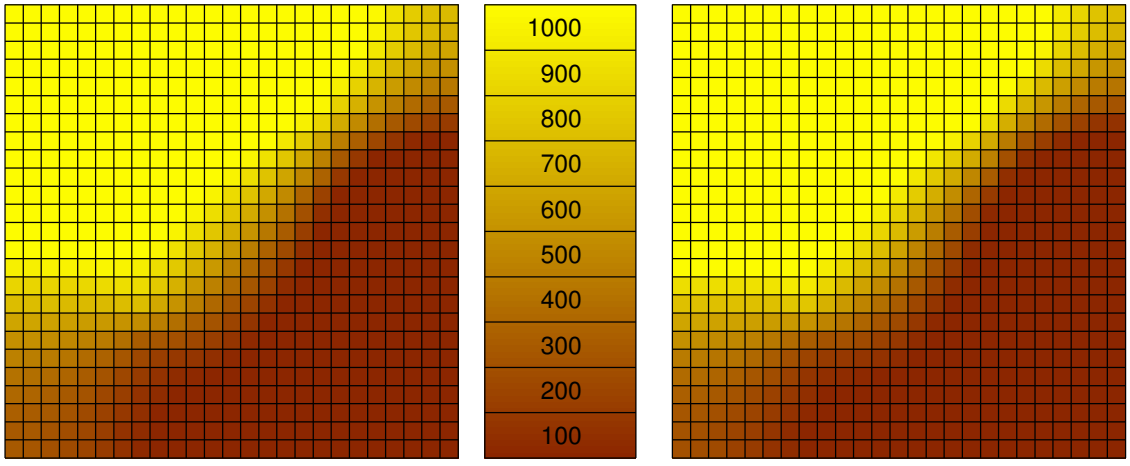
Im nächsten Test wird untersucht, ob eine Erhöhung der Anzahl der Frequenzen für eine bessere Identifikation sorgt.

4.5 Erhöhung der Anzahl der Frequenzen

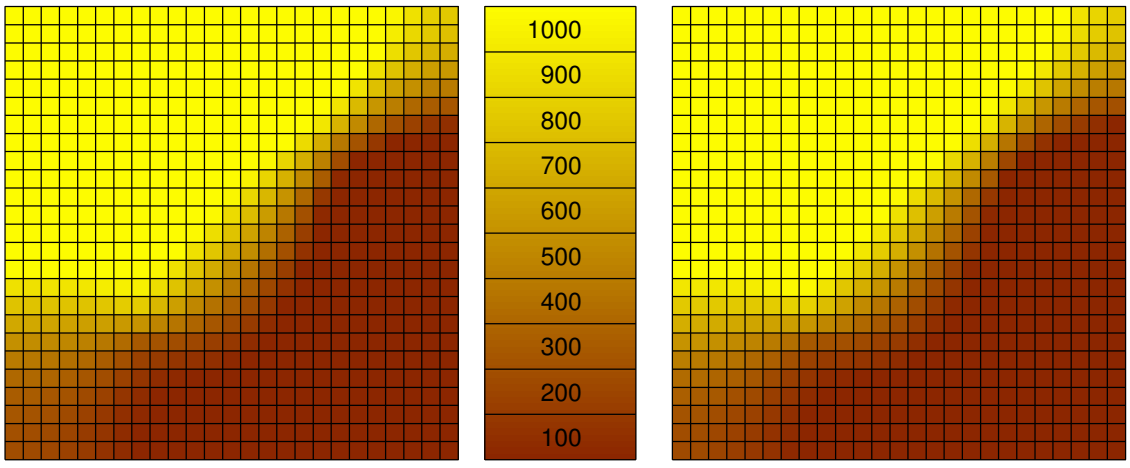
In diesem Test wird das Meßprogramm 1 verwendet. In jedem Test wird eine Frequenz von $\omega = 0, 10^{-1}, 1, 10, 100, 1000$ Hertz hinzugenommen. Bei dieser Wahl unterscheidet sich der Frequenzgang des Imaginärteils der verschiedenen Bodentypen am deutlichsten (siehe Abbildung 2.10). Die verwendete rechnerische Auflösung ist 50×50. Als Faktor vor der Matrix B wird 1 benutzt. Es sollen $25 \times 25 \times 4 = 2500$ Parameter bestimmt werden. Es wird dabei das Gauß-Newton Verfahren gebraucht, das nach 100 Iterationen durch $k_{\max}$ gestoppt wird.



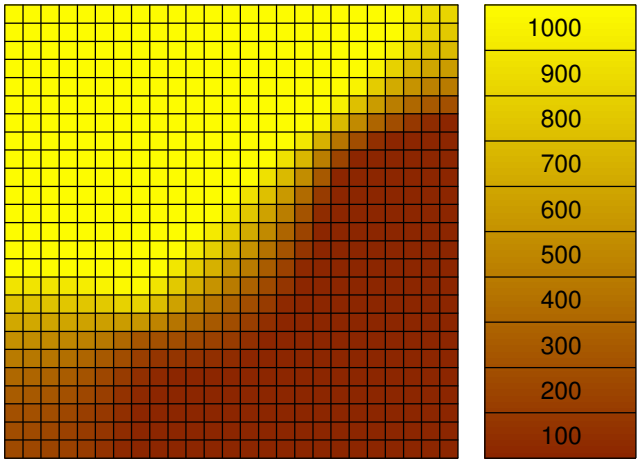
(a) Original 32×32 (b) Gauß-N. 50×50, 25×25, B=1, MP=1



(c) Gauß-N. 50×50, 25×25, B=1, MP=1 (d) Gauß-N. 50×50, 25×25, B=1, MP=1



(e) Gauß-N. 50×50, 25×25, B=1, MP=1 (f) Gauß-N. 50×50, 25×25, B=1, MP=1



(g) Gauß-N. 50×50, 25×25, B=1, MP=1

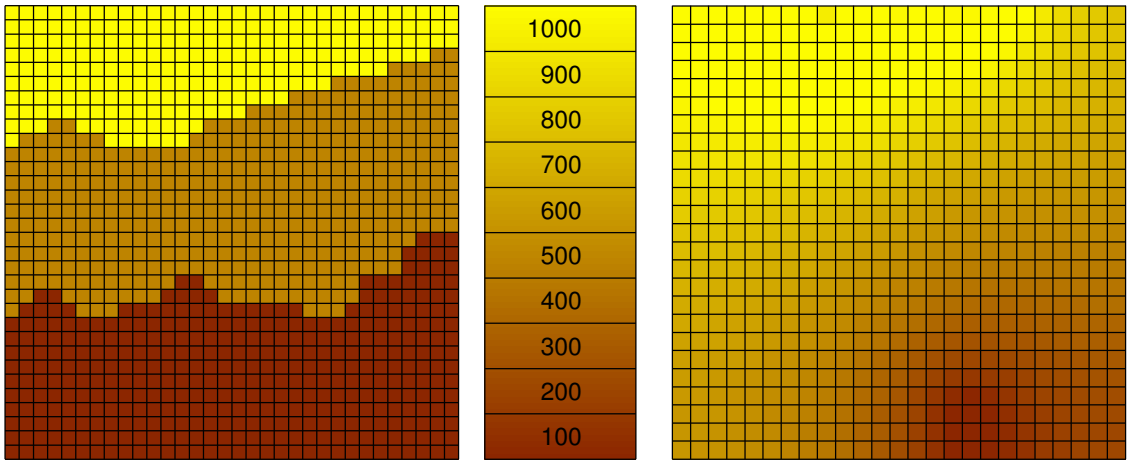
Ergebnisse:

An diesen Bildern ist leider keine Identifikationsverbesserung erkennbar. Benutzt man allerdings eine gröbere Auflösung als 25×25, dann ist diese deutlich zu erkennen. Dann sieht man, daß eine Erhöhung ab der fünften Frequenz keine Identifikationsverbesserung bringt(siehe auch [8]). Es reicht, wenn man zum Beispiel die Frequenzen $\omega = 0, 10^{-1}, 1, 10$ und 100 verwendet (siehe Abbildungen (b)-(e)).

Es stellt sich nun die Frage, wie tief in den Boden geschaut werden kann. Aus diesem Grund wird im nächsten Test das Meßprogramm 2 verwendet, bei dem sich Elektroden nur an der Oberfläche befinden.

4.6 Vertikaler Identifikationsbereich

Bei diesem Test wird das Meßprogramm 2 benutzt, wobei die Elektroden unmittelbar unter der Oberfläche angeordnet sind. Es soll demonstriert werden, wie weit in die Tiefe identifiziert werden kann. Hierbei wird das Gauß-Newton Verfahren mit rechnerischer Auflösung von 50×50 verwendet. Diese Methode ist durch $k_{\max} = 100$ abgebrochen worden. Es sollen $25 \times 25 \times 4 = 2500$ Parameter identifiziert werden. Die Gewichtung der Matrix B ist auf 1 gesetzt. Es wird mit den Frequenzen ω=0, $10^{-1}, 1, 10$ und 100 Hertz identifiziert.



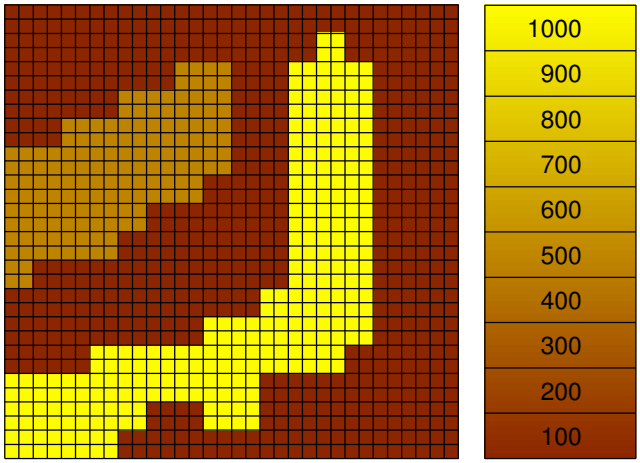
(a) Original 32×32 (b) Gauß-N. 50×50, 25×25, B=1, MP=2

Ergebnisse:
Es ist ein Trend zu erkennen, jedoch können selbst einfache horizontale Strukturen nicht identifiziert werden, da der Strom nur an der Oberfläche fließt. Daher macht es keinen Sinn, das Meßprogramm 2 zu verwenden.

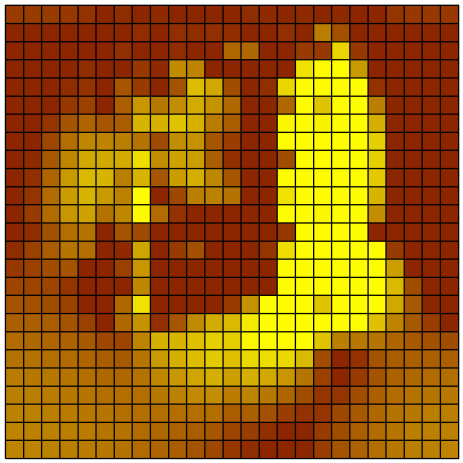
Es bleibt offen, ob mit dem ersten Meßprogramm komplexe Strukturen identifiziert werden können. Im nachfolgenden Test wird dieser Fragestellung nachgegangen.

4.7 Identifikation von komplexen Strukturen

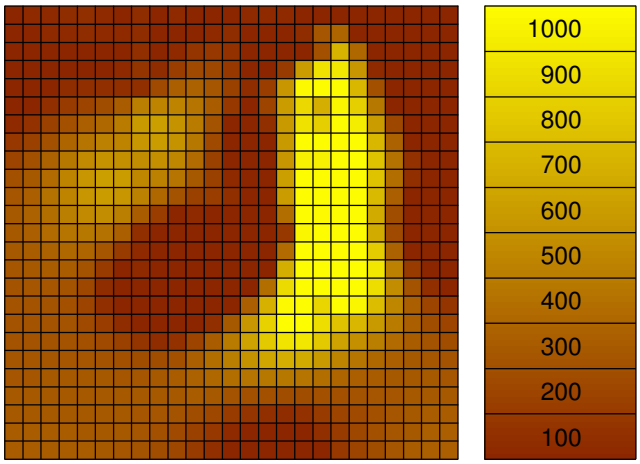
In diesem Test wird eine komplexe Struktur identifiziert. Dabei wird das Meßprogramm 1 verwendet. Die Diskretisierung ist 50×50. Es sollen mit der Gewichtung 0, 0.1, 0.2, 0.3, 0.4, 0.5 und 1 für die Matrix B $25 \times 25 = 625$ Parameter identifiziert werden. Es wird das Dog Leg Verfahren verwendet, das nach 100 Iterationen durch $k_{\max}$ beendet wird. Es wird mit der Frequenz $\omega = 0$ gerechnet.



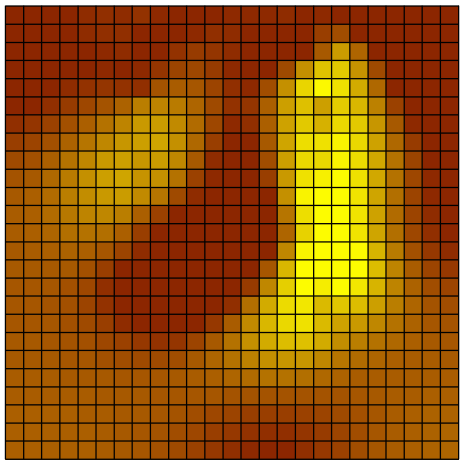
(a) Original 32×32



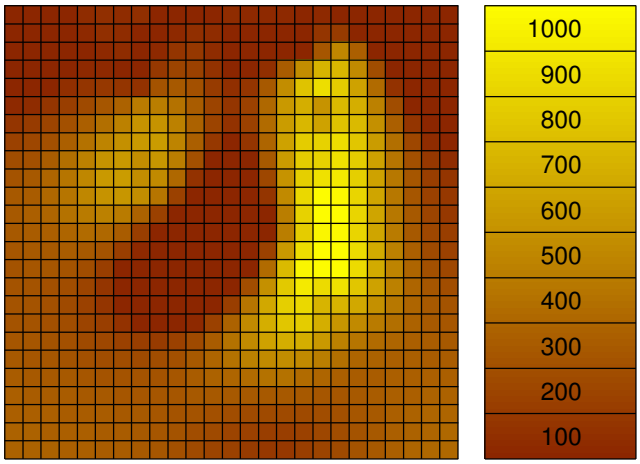
(b) Dog Leg 50×50, 25×25, B=0, MP=1



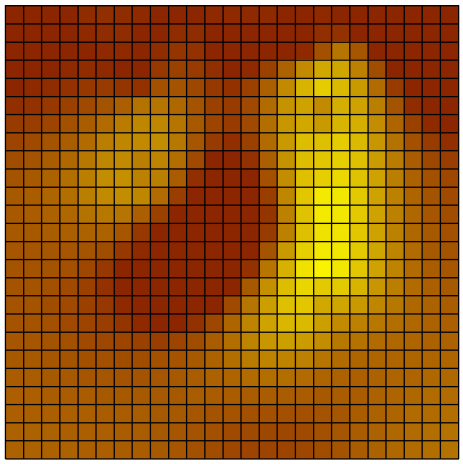
(c) Dog Leg 50×50, 25×25, B=0.1, MP=1



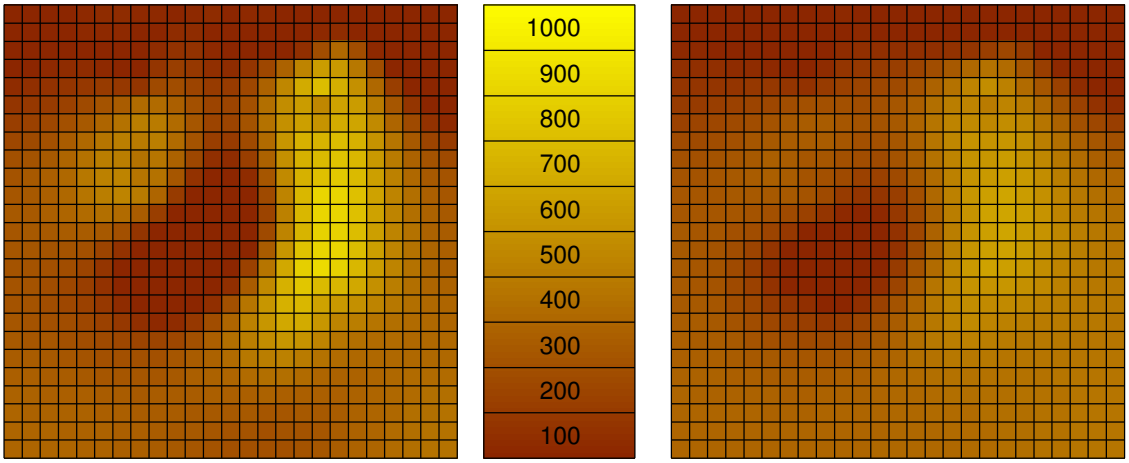
(d) Dog Leg 50×50, 25×25, B=0.2, MP=1



(e) Dog Leg 50×50, 25×25, B=0.3, MP=1



(f) Dog Leg 50×50, 25×25, B=0.4, MP=1



(g) Dog Leg 50 × 50, 25 × 25, B=0.5, MP=1

(h) Dog Leg 50 × 50, 25 × 25, B=1, MP=1

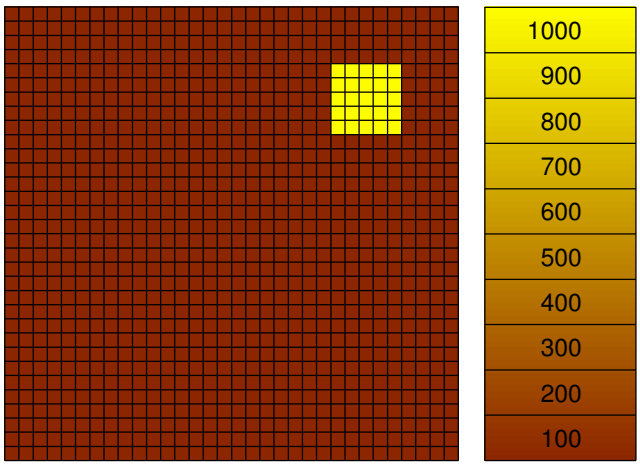
Ergebnisse:

Das Identifizieren der komplexen Struktur ist nicht ganz möglich. Zu sehen ist, daß das untere Drittel nicht identifiziert wird, was durch die Anordnung der Elektroden erklärt werden kann. Im unteren Drittel fließt kein Strom und es wird dort auch kein Potential gemessen. Die Matrix B bewirkt, daß die Struktur geglättet wird. In diesem Fall ist der Faktor 0.1 am besten.
Im Fall (e) ist nicht das absolute Minimum gefunden worden (siehe Test 4.13).

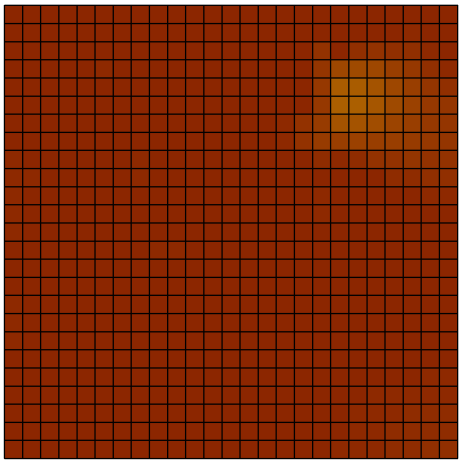
Hat dies zur Folge, daß eventuell einzelne Blöcke nicht identifiziert werden können?

4.8 Identifikation einzelner Blöcke

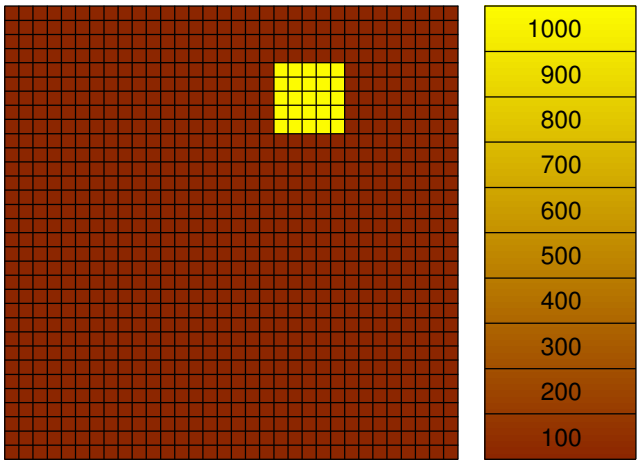
In diesem Test wird untersucht, ob auch einzelne Blöcke an verschiedenen Positionen identifiziert werden können. Es wird das Meßprogramm 1 verwendet. Die rechnerische Auflösung ist 50×50. Der Faktor der Matrix B ist 0.1. Es sollen $25 \times 25 \times 4 = 2500$ Parameter identifiziert werden. Das Gauß-Newton Verfahren stoppt nach 100 Iterationen durch $k_{\max}$. Es werden dabei die Frequenzen $\omega = 0, 10^{-1}, 1, 10$ und 100 Hertz verwendet.



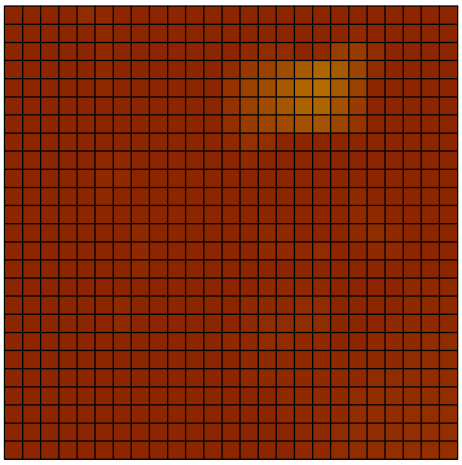
(a) Original 32×32



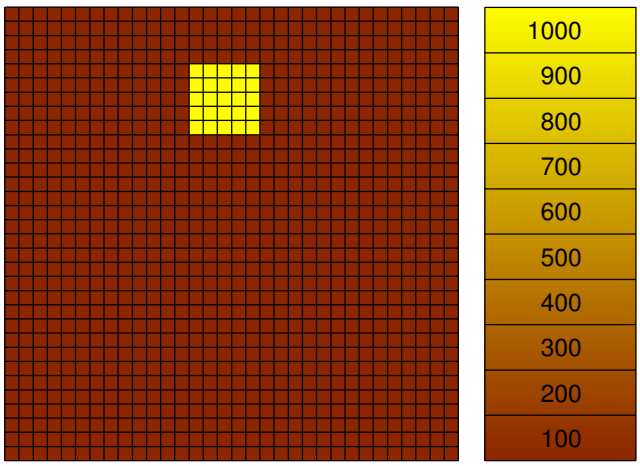
(b) Gauß-N. 50×50, 25×25, B=0.1, MP=1



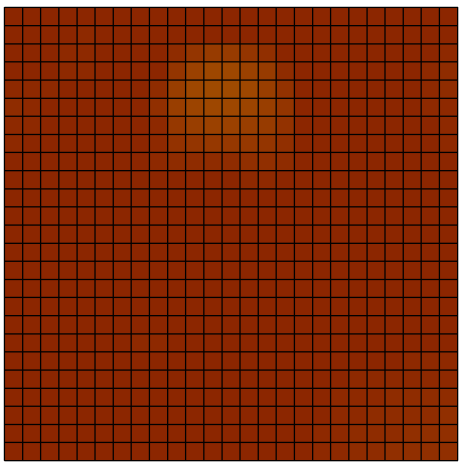
(c) Original 32×32



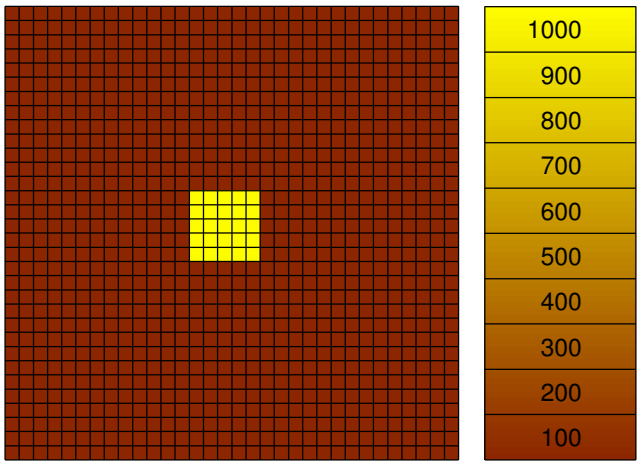
(d) Gauß-N. 50×50, 25×25, B=0.1, MP=1



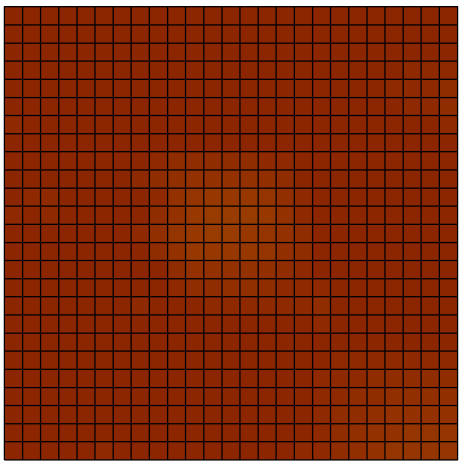
(e) Original 32×32



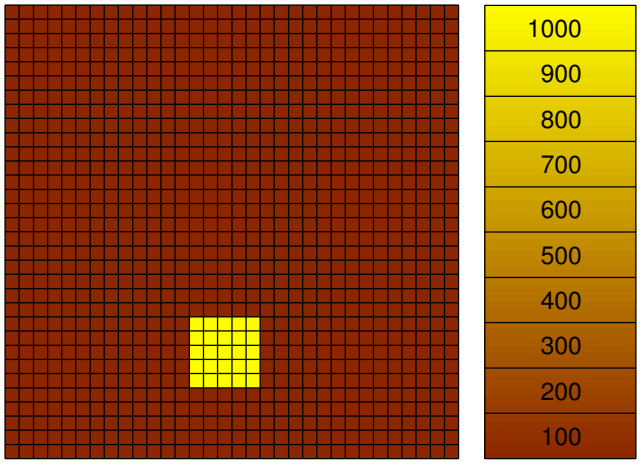
(f) Gauß-N. 50×50, 25×25, B=0.1, MP=1



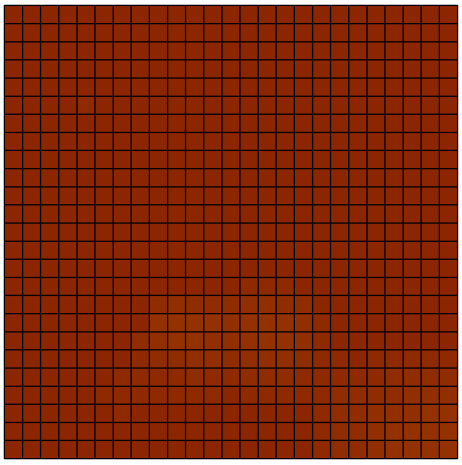
(g) Original 32×32



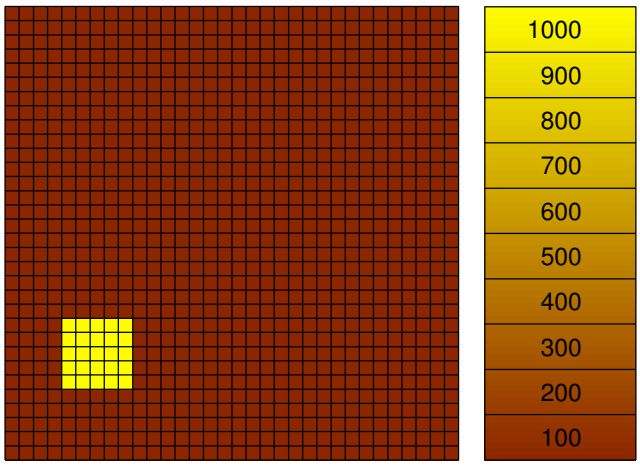
(h) Gauß-N. 50×50, 25×25, B=0.1, MP=1



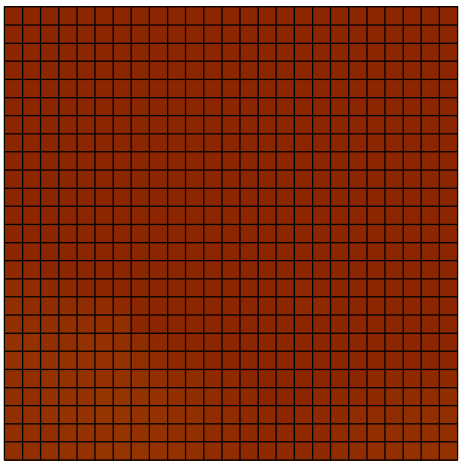
(i) Original 32×32



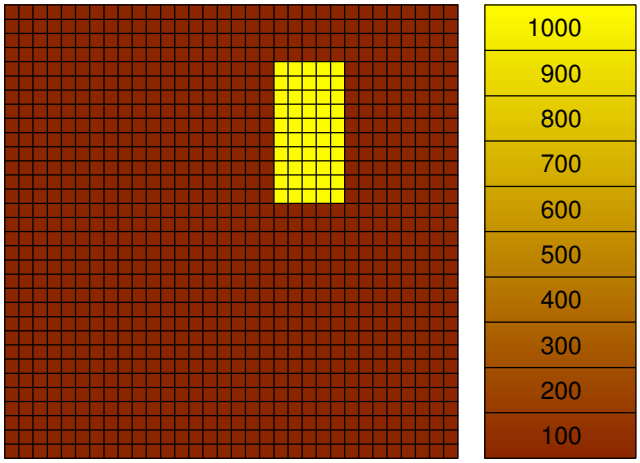
(j) Gauß-N. 50×50, 25×25, B=0.1, MP=1



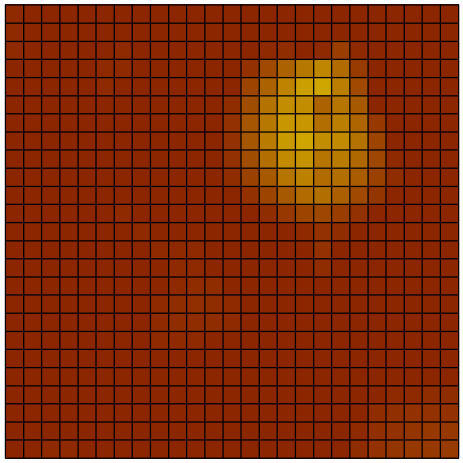
(k) Original 32×32



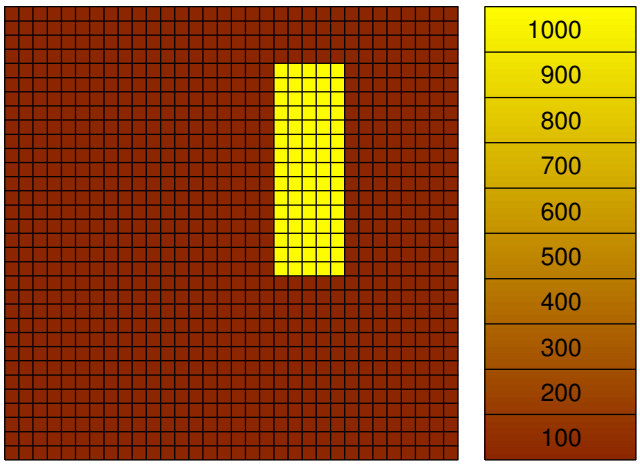
(j) Gauß-N. 50×50, 25×25, B=0.1, MP=1



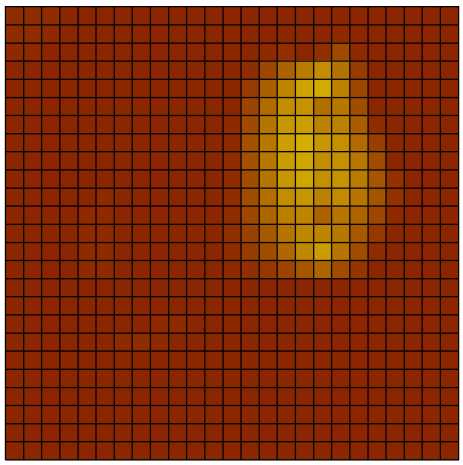
(k) Original 32×32



(l) Gauß-N. 50×50, 25×25, B=0.1, MP=1



(m) Original 32×32



(n) Gauß-N. 50×50, 25×25, B=0.1, MP=1

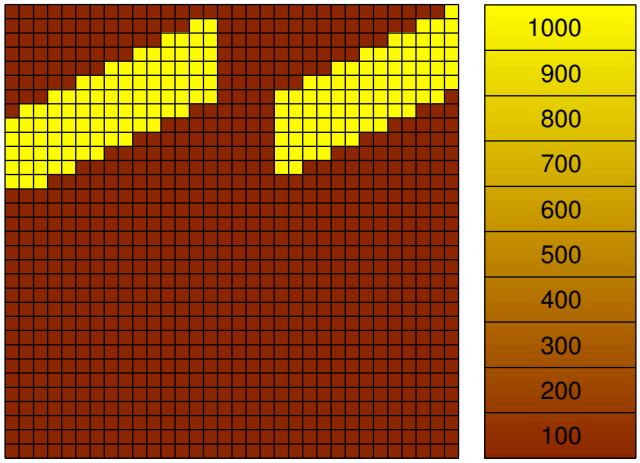
Ergebnisse:

Einzelne Blöcke können im unteren Drittel nicht identifiziert werden (siehe Abbildung (h) und (j)). Ansonsten ist eine Identifikation möglich. Ein leichter Schatten läßt vermuten, daß sich dort eine Struktur befindet. Die Sichtbarkeit ist hierbei zusätzlich abhängig von der Größe der Struktur (siehe Abbildung (l) und (n)). Ferner hängt es davon ab, ob die Elektroden an dem zu identifizierenden Klotz liegen (siehe Abbildung (d)).

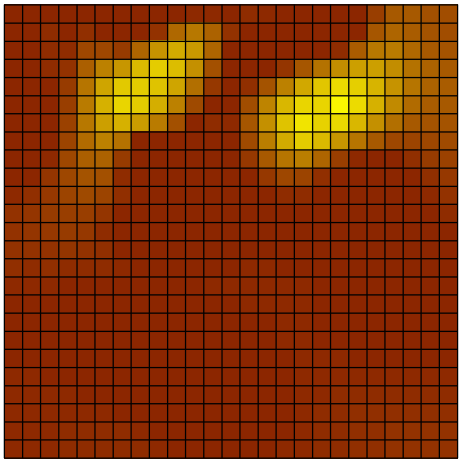
Es stellt sich die Frage, ob die Größe der zu identifizierenden Struktur hierbei eine wesentliche Rolle spielt.

4.9 Identifikation eines Durchbruchs

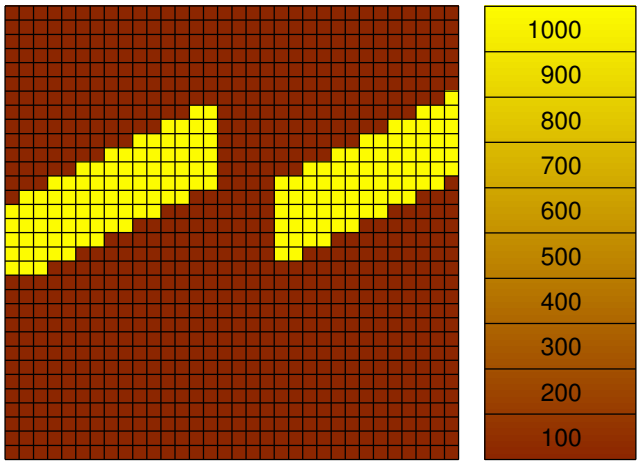
In diesem Test wird versucht, ein Durchbruch im Boden zu identifizieren. Dieser ist wesentlich größer als der im vorherigen Test verwendete Block. Beim ersten Test befindet sich der Durchbruch fast an der Oberfläche. Dieser wird bis hin zum letzten Test immer weiter abgesenkt. Es wird das Gauß-Newton Verfahren mit der rechnerischen Auflösung von 50×50 verwendet, das nach 100 Schritten durch $k_{\max}$ abgebrochen worden ist. Es sollen $25 \times 25 \times 4 = 2500$ Parameter mit dem Meßprogramm 1 bestimmt werden. Die Gewichtung der Matrix B ist 0.3. Die verwendeten Frequenzen sind $\omega = 0, 10^{-1}, 1, 10$ und 100 Hertz.



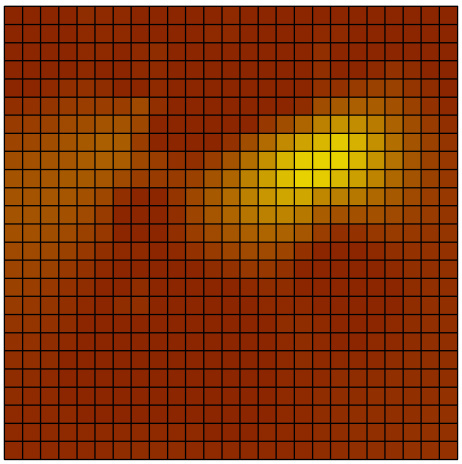
(a) Original 32×32



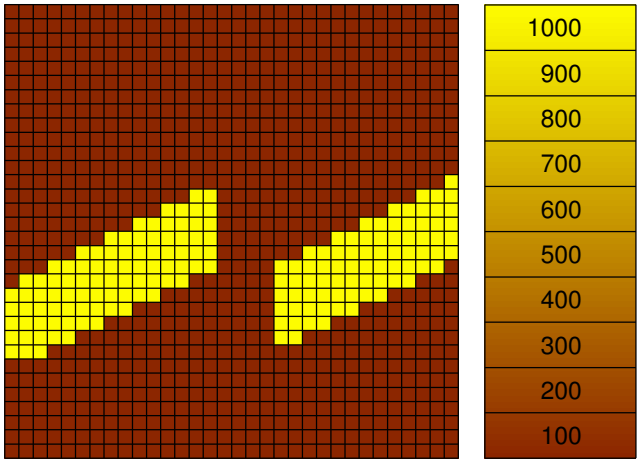
(b) Gauß-N. 50×50, 25×25, B=0.1, MP=1



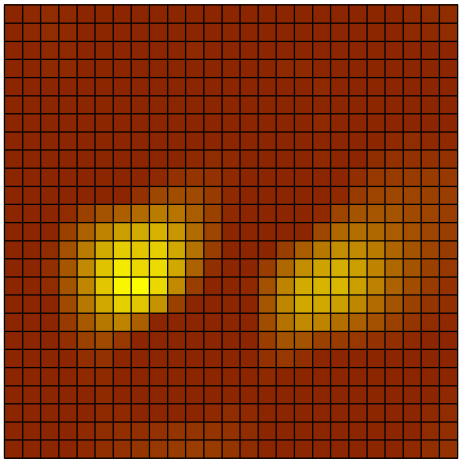
(c) Original 32×32



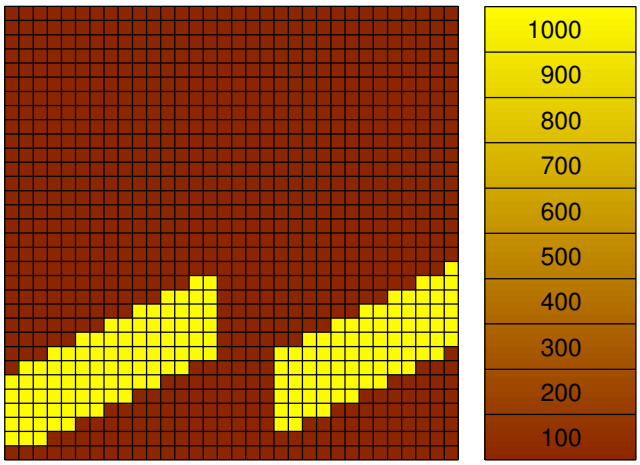
(d) Gauß-N. 50×50, 25×25, B=0.1, MP=1



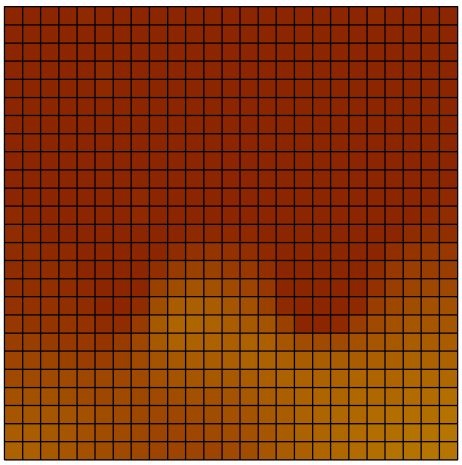
(e) Original 32×32



(f) Gauß-N. 50×50, 25×25, B=0.1, MP=1



(g) Original 32×32



(h) Gauß-N. 50×50, 25×25, B=0.1, MP=1

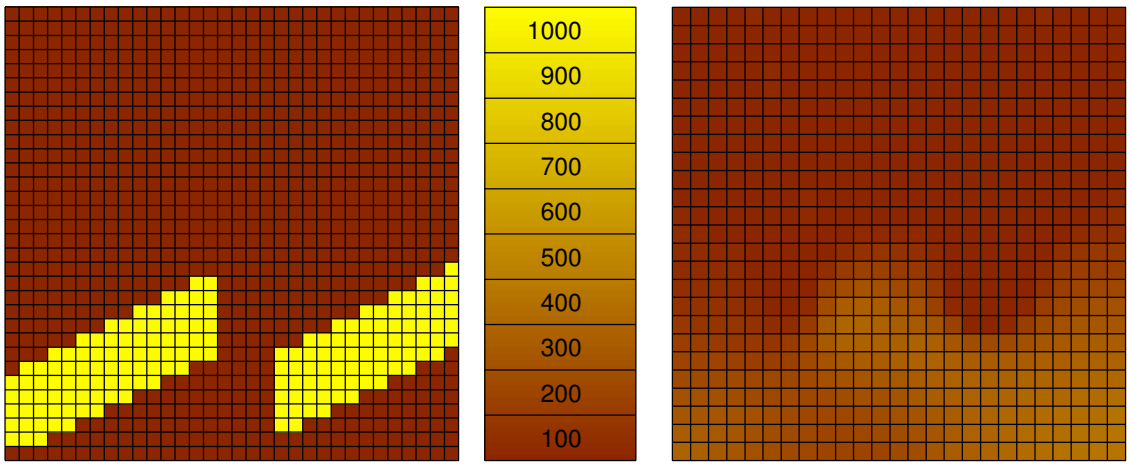
Ergebnisse:

Der Durchbruch an der Oberfläche kann identifiziert werden. Auch in der Mitte kann er einigermaßen erkannt werden. Liegt er hingegen im unteren Drittel, so ist keine Identifikation möglich, was wiederum durch die Anordnung der Elektroden erklärt werden kann.

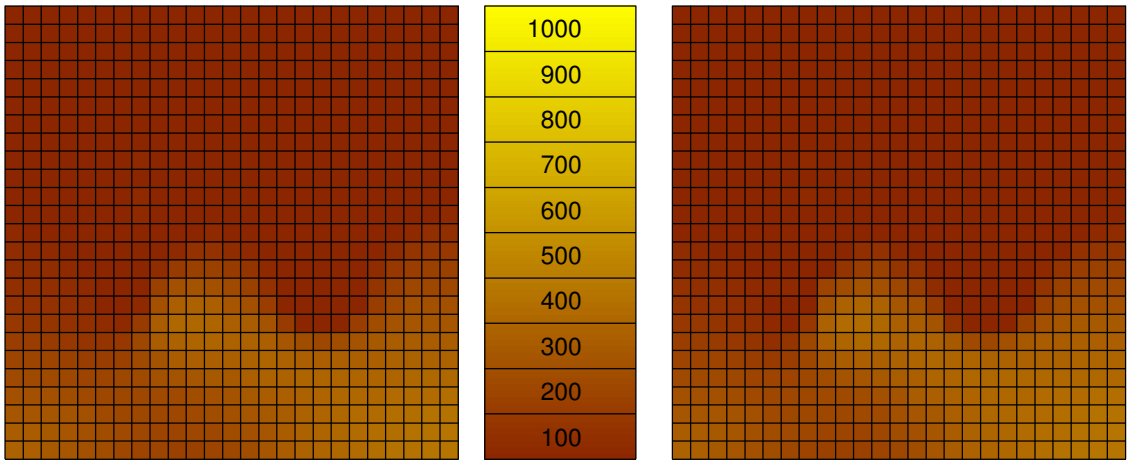
Im folgenden soll betrachtet werden, ob ein Erhöhen der Iterationsanzahl für eine bessere Identifikation sorgt.

4.10 Erhöhung der Iterationen

In diesem Test wird die Iterationsanzahl erhöht, um den Durchbruch im untersten Drittel besser identifizieren zu können. Es wird einmal mit 100, 200 und 500 Iterationen bei einer rechnerischen Auflösung von 50×50 gerechnet. Es wird das Gauß-Newton Verfahren verwendet, um $25 \times 25 \times 4 = 2500$ Parameter zu identifizieren. Der Faktor der Matrix B ist 0.3. Die verwendeten Frequenzen sind $\omega = 0, 10^{-1}, 1, 10$ und 100 Hertz.



(a) Original 32×32 (b) Gauß-N. 50×50, 25×25, B=0.1, MP=1



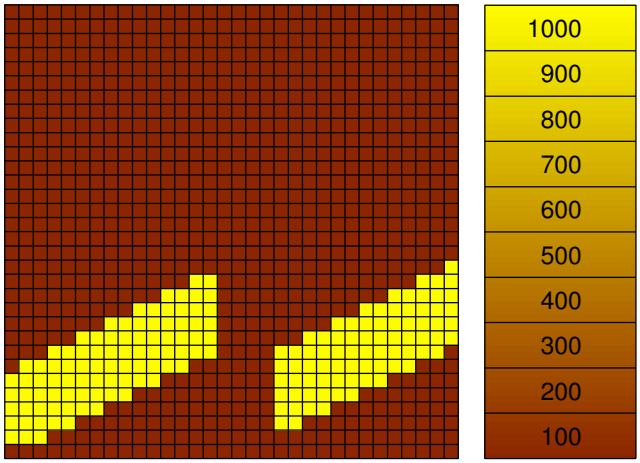
(c) Original 32×32 (d) Gauß-N. 50×50, 25×25, B=0.1, MP=1

Ergebnisse:
Eine Erhöhung der Iterationsanzahl bringt keine Verbesserung der Identifikation.

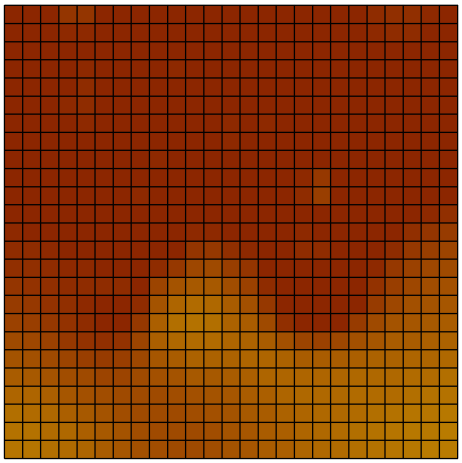
Kann eventuell durch Verwenden des Meßprogramms 3, das vier Stromauslaßstellen besitzt, eine Verbesserung erreicht werden?

4.11 Verwendung mehrerer Stromauslaßstellen

In diesem Test wird das Meßprogramm 3 verwendet, da hier vier Stromauslaßstellen existieren. Es wird mit dem Gauß-Newton Verfahren mit der Auflösung 50×50 gerechnet, welches nach 100 Iterationen durch $k_{\max}$ abgebrochen worden ist. Mit dem Faktor 0.3 für die Matrix B und den Frequenzen $\omega = 0, 10^{-1}, 1, 10$ und 100 Hertz sollen $25 \times 25 \times 4 = 2500$ Parameter bestimmt werden.



(a) Original 32×32



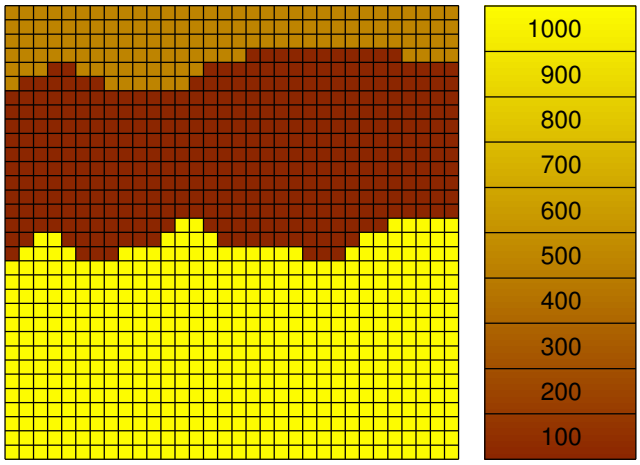
(b) Gauß-N. 50×50, 25×25, B=0.1, MP=3

Ergebnisse:

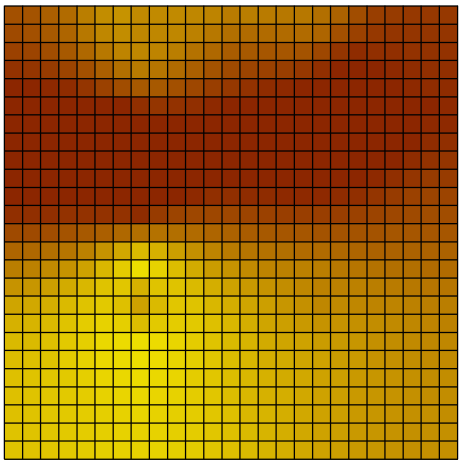
Auch bei der Verwendung des Meßprogramms 3 ist keine Verbesserung der Identifikation möglich. Dies kann dadurch erklärt werden, daß keine neuen Informationen hinzugefügt wurden. Die Verwendung von mehreren Stromquellen liefert Gleichungen, die linear abhängig von den bisher aufgestellten Gleichungen sind. Ein Durchbruch in dieser Tiefe kann also nicht identifiziert werden.

4.12 Identifikation von horizontalen Schichten

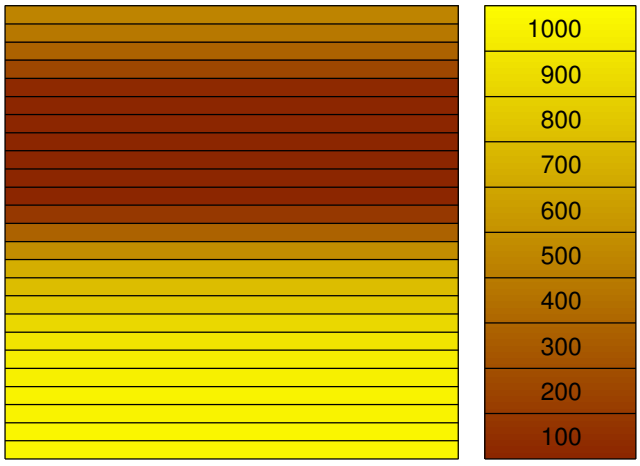
In diesem Test wird demonstriert, daß das Programm auch in der Lage ist, nur horizontale Schichten zu identifizieren. Dies erreicht man, indem die Auflösung in x-Richtung auf 1 gesetzt wird. Neben der Bestimmung von $25 \times 25 = 625$ Parametern werden einmal 25 und 50 horizontale Schichten bestimmt. Dazu wird das Levenberg-Marquardt Verfahren mit der rechnerischen Auflösung 50×50 und dem Meßprogramm 1 verwendet. Der Faktor der Matrix B ist 0.3. Es wird hier nur mit Gleichstrom gerechnet. Das Verfahren wird durch $k_{\max}$ nach 100 Iterationen gestoppt.



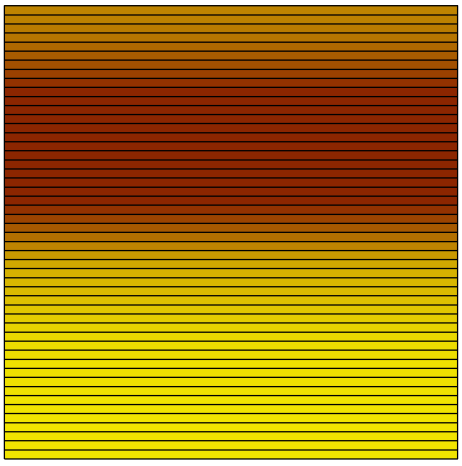
(a) Original 32×32



(b) Levenberg-M. 50×50, 25×25, B=1, MP=1



(c) Levenberg-M. 50×50, 1×25, B=1, MP=1



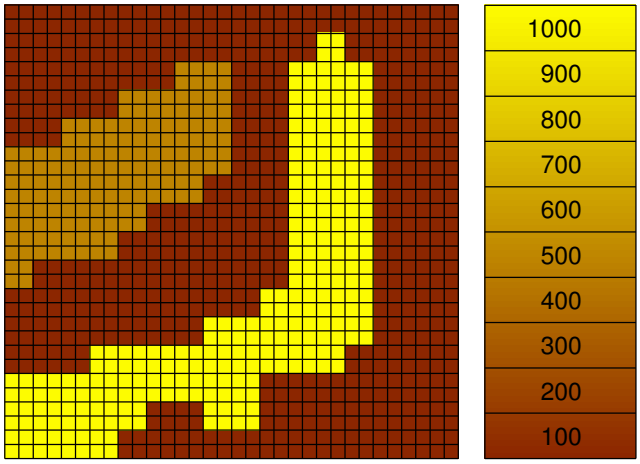
(d) Levenberg-M. 50×50, 1×50, B=1, MP=1

Ergebnisse:

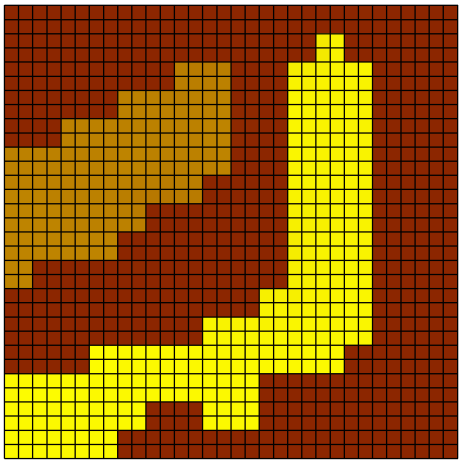
Eine Identifikation mit horizontalen Schichten ist sehr gut, da die Anzahl der Parameter hierbei signifikant verringert wird. Im letzten Fall werden nur 50 Parameter identifiziert. Diese Annahme ist oft gerechtfertigt, da Bodenschichten häufig horizontal verlaufen.

4.13 Einfluß der Fehler

In diesem Test wird mit der wahren Lösung gestartet. Es wird getestet, wie sich das Verfahren verhält. Dazu wird die Gauß-Newton Methode mit dem Meßprogramm 1 verwendet. Die rechnerische Auflösung ist 64×64. Es sollen $32 \times 32 \times 4 = 4096$ Parameter bestimmt werden. Der Faktor der Matrix B ist 0.3. Die verwendeten Frequenzen ω sind 0, $10^{-1}, 1, 10$ und 100 Hertz. Das Verfahren wird nach 100 Iterationen durch $k_{\max}$ gestoppt.



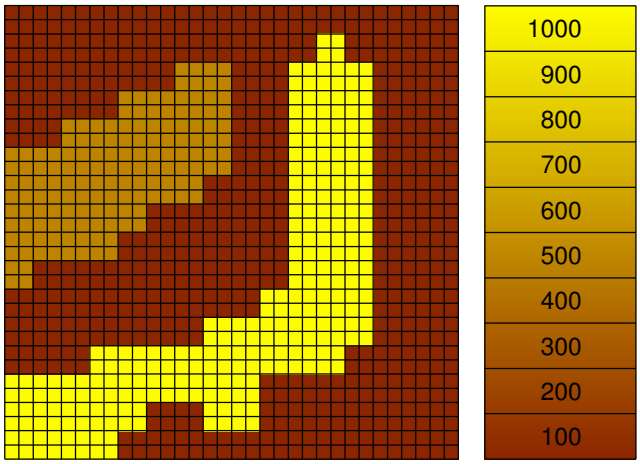
(a) Original 32×32



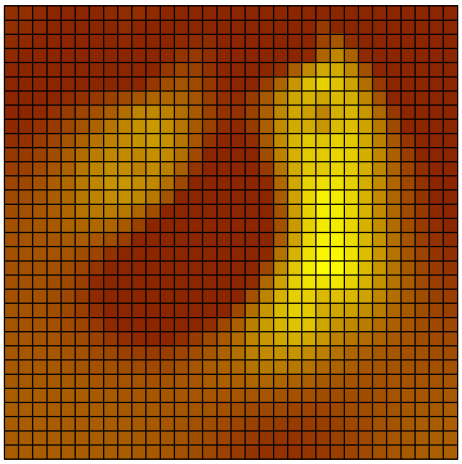
(b) Gauß-N. 64×64, 32×32, B=0.3, MP=1

Ergebnisse:

Das Gauß-Newton Verfahren iteriert 100 Schritte und sucht einen anderen Minimierer, was durch den Einfluß der Fehler erklärt werden kann. Die Abweichungen sind aber sehr gering und nicht sichtbar (siehe Abbildung (b)). Der Funktionswert $||f(c_{100})||_2$ hat nach 100 Schritten den Wert 19,037515.
Dem wird die Lösung gegenübergestellt, die man erhält, falls mit dem Bodentyp „Bruchzone“ gestartet wird (siehe Abbildung (d)). Der Funktionswert $||f(c_{500})||_2$ beträgt nach 500 Schritten 40,520878 und liegt damit deutlich über dem vorherigen Wert. Das heißt, daß das absolute Minimum nicht erreicht worden ist.



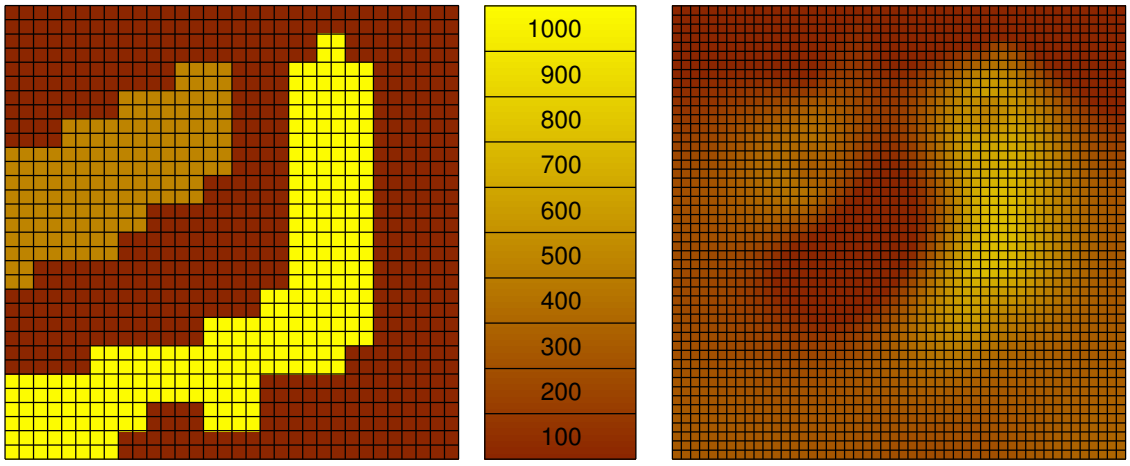
(c) Original 32×32



(d) Gauß-N. 64×64, 32×32, B=0.3, MP=1

4.14 Maximalbeispiel

In diesem Test wird das Vorwärtsproblem mit einer rechnerischen Auflösung von 200×200 gelöst. Es sollen mit dem Gauß-Newton Verfahren $50 \times 50 \times 4 = 10000$ Parameter bestimmt werden. Es wird das Meßprogramm 1 verwendet. Die Diskretisierung ist hierbei 100×100. Der Faktor vor der Matrix B ist 0.3 und die verwendeten Frequenzen ω sind 0, 10^{-1}, 1, 10 und 100 Hertz. Dieser Test fand auf einem Rechner mit Intel XEON mit 2 GHz und 2048 MByte Arbeitsspeicher statt, da 1024 MByte nicht ausreichten.



(a) Original 32×32 (b) Gauß-N. 100×100, 50×50, B=0.3, MP=1

Ergebnisse:

Die Verwendung einer hohen Auflösung ist möglich, falls genug Arbeitsspeicher zur Verfügung steht. Zu sehen ist außerdem, daß das Verfahren auch bei einer solch hohen Auflösung stabil bleibt.

Kapitel 5

Zusammenfassung und Ausblick

In dieser Arbeit wird ein inverses Problem zur Bestimmung von Bodenstrukturen im zweidimensionalen Raum behandelt. Diese Einschränkung bedeutet physikalisch, daß sich in y-Richtung die Leitfähigkeit nicht ändert. Damit sich die Potentiale in y-Richtung nicht ändern, wird angenommen, daß die Stromeinleitungen und -ausleitungen in y-Richtung konstant sind. Diese Approximationen erscheinen berechtigt, da sie am hier behandelten Problem die Schlechtgestelltheit des inversen Problems nichts wesentliches ändern.
Das inverse Problem wird durch verschiedene Verfahren, wie das Gauß-Newton, das Levenberg-Marquardt und das Dog Leg Verfahren gelöst. In diesen Methoden müssen lineare Ausgleichsprobleme gelöst werden. Zusätzlich muß das Lösen eine Regularisierung enthalten, damit Rauschen in den gemessenen Potentialen nicht die Approximationsqualität der zu bestimmenden Cole-Cole Parameter verfälscht. Dafür ist ausschließlich das konjugierte Gradientenverfahren für Normalengleichungen verwendet worden, da hierbei nicht die vollständige Jacobimatrix aufgestellt werden muß (siehe Einleitung von Kapitel 4).
Bei der Verwendung von Gleichstrom erzeugen das Gauß-Newton, das Levenberg-Marquardt und das Dog Leg Verfahren vergleichbare Ergebnisse. Es sollte aber nicht das Gauß-Newton Verfahren verwendet werden, da es mehr Aufwand für die zusätzliche Liniensuche benötigt. Wird allerdings mit Wechselstrom gerechnet, ist das Gauß-Newton Verfahren dem Levenberg-Marquardt und dem Dog Leg Verfahren überlegen, da die beiden letzten genannten Verfahren extrem langsam konvergieren. Dies wird durch die Cole-Cole Parameter m, τ und γ verursacht, da diese einen negativen Einfluß auf das Gewinnverhältnis ϱ haben (Test 4.2).
Durch die Verwendung eines Strafterms wird erreicht, das starke Variationen in der Approximation der Cole-Cole Parameter bestraft werden. Dabei wird eine Matrix B verwendet, die den diskreten Gradienten der Cole-Cole Parametern enthält. Außerdem beachten die verwendeten Verfahren einfache Nebenbedingungen, die sich durch die Minima und Maxima der Cole-Cole Parameter der drei typischen Bodentypen ergeben. Droht ein Verlassen des zulässigen Bereichs, werden die neu iterierten Cole-Cole Parameter auf den Rand dieses Bereichs projiziert (Test 4.4).
Ferner kann festgehalten werden, daß alle drei Verfahren auch bei starker Unterbestimmtheit stabil bleiben. Es kann somit eine sehr große Anzahl an Parametern bestimmt werden (Test 4.3 und Test 4.14).
Des weiteren lohnt es sich nicht, mehr als fünf Frequenzen zu verwenden. Man erhält keine neuen Informationen, da eine Sättigung eintritt (Test 4.5).
Eine Anordnung der Elektroden nur an der Oberfläche sollte nicht erfolgen, da dann höchstens ein Trend zu beobachten ist. Selbst einfache Strukturen können nicht erkannt werden (Test 4.6).

Benutzt man hingegen das Meßprogramm 1, bei dem sich neben Elektroden an der Oberfläche auch Elektroden im Boden befinden, dann können sogar komplexe Strukturen identifiziert werden (Test 4.7).
Auch das Identifizieren von einzelnen kleinen Blöcken ist begrenzt möglich. In den zwei oberen Dritteln sind sie zu erkennen, wohingegen sie im unteren Drittel nicht erkennbar sind, was aufgrund der Anordnung der Elektroden erklärt werden kann. Wird der Block größer gewählt oder befinden sich Elektroden in diesem Block, dann ist die Identifikation besser (Test 4.8).
Auch Durchbrüche können in den zwei oberen Dritteln sichtbar gemacht werden (Test 4.9).
Bei allen Testbeispielen ist das jeweils verwendete Verfahren durch $k_{\max}$=100 abgebrochen worden. Ab diesem Iterationsindex waren die Veränderungen so gering, daß sich der Aufwand nicht lohnte, das Verfahren weiter iterieren zu lassen. Die Verfahren sind nicht durch Finden eines globalen Minimums oder durch zu kleine Schritte gestoppt worden. Es hat sich aber auch gezeigt, daß ab der 100. Iteration kaum sichtbare Änderungen geschehen (Test 4.10).
Wechselnde Ein- und Auslaßquellen für den Strom bringen keine neuen Informationen. Es reicht, wenn an einer Elektrode der Strom abfließt (Test 4.11).
Es können globale Strukturen mit einer sehr großen Auflösung identifiziert werden. Auch horizontale Schichten können sehr gut und bei deutlich reduziertem Aufwand bestimmt werden (Test 4.12).
Die Verwendung des Faktors vor der Matrix B sollte folgendermaßen gewählt werden: Falls globale Strukturen identifiziert werden sollen, ist 0.5-1.0 eine gute Wahl. Für lokale Strukturen sollte hingegen 0.1-0.3 verwendet werden. Der Faktor Null sollte nicht benutzt werden.
Startet man mit der exakten Lösung, so werden 100 Iterationsschritte durchgeführt. Dies kann durch die Fehler, die auf den gemessenen Potentialen liegen, erklärt werden. Startet man mit dem Bodentyp „Bruchzone“, dann liegt der Funktionswert $||f(c)||_2$ deutlich über dem Wert, den man bei Benutzung der exakten Lösung erhält. Das heißt, daß das absolute Minimum noch nicht gefunden worden ist (Test 4.13).

Es sollten Überlegungen gemacht werden, andere Matrizen als die Matrix B, die den diskreten Gradienten enthält, zu verwenden.
Zur Erhöhung der Konvergenzgeschwindigkeit sollte ein Hybrid-Verfahren [10] getestet werden. Hierbei wird das Levenberg-Marquardt Verfahren mit einer Quasi-Newton Methode verknüpft. Der Vorteil dieses Verfahrens ist die superlineare Konvergenz, falls sich der Algorithmus für einen Quasi-Newton Schritt entscheidet. Außerdem wird dabei die Ableitungsmatrix mit Broydens Rang-Update erneuert. Versuche, insgesamt die Konvergenzgeschwindigkeit mit dieser vorgeschlagenen Methode zu erhöhen, wären interessant.
Eine Erweiterung auf das Dreidimensionale sollte ebenfalls vorgenommen werden, da dies die Wirklichkeit besser beschreibt.
Eventuell sollten auch Überlegungen angestellt werden, das Programm nach C++ umzuschreiben, um es anschließend zu parallelisieren.
Ein völlig anderer Ansatz wäre es, Heuristiken zur Optimierung zu verwenden. Es seien hier das Simulated Annealing Verfahren und der genetische Algorithmus erwähnt.

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

[1] T. Hohage: Inverse Problems, Universität Göttingen, 2002

[2] H. Lustfeld: Vorlesung Moderne Probleme der Geophysik mit Methoden der theoretischen Physik II, Forschungszentrum Jülich, 2004

[3] A. Rieder: Keine Probleme mit Inversen Problemen, Vieweg Verlag, Wiesbaden, 2003

[4] H. W. Engl, M. Hanke, A. Neubauer: Regularization of Inverse Problems, Kluwer Academic Publishers, Dordrecht Boston London, 1996

[5] J. E. Dennis, Jr., R. B. Schnabel: Numerical Methods for unconstrained Optimization and nonlinear Equations, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1983

[6] P. Deuflhard: Newton Methods for Nonlinear Problems, Springer Verlag, Berlin Heidelberg, 2004

[7] A. Quarteroni, R. Sacco, F. Saleri: Numerical Mathematics, Springer Verlag, New York, Inc., 2000

[8] A. Erven: Untersuchungen zur Informationsoptimierung in der geoelektrischen Widerstandstomografie, Diplomarbeit, FH Aachen – Abteilung Jülich – in Zusammenarbeit mit Forschungszentrum Jülich GmbH – Institut für Festkörperforschung, 2005

[9] S. Jongen: Entwicklung schneller Löser für Randwertprobleme in der geoelektrischen Widerstandstomographie, Diplomarbeit, FH Aachen – Abteilung Jülich – in Zusammenarbeit mit Forschungszentrum Jülich GmbH – Institut für Festkörperforschung, 2005

[10] K. Madson, H. B. Nielsen, O. Tingleff: Methods for Non-Linear Least Squares Problems, Informatics and Mathematical Modelling – Technical University of Denmark, 2nd Edition, April 2004

[11] G. Dikta: Stochastik, Vorlesungsskript, Wintersemester 2004

[12] M. Hanke: Conjugate gradient type methods for ill posed problems, Longman, 1995

[13] M. Jürgens: LaTeX– Eine Einführung und ein bißchen mehr, Benutzerhandbuch, KFA-ZAM-BHB-0134, 1. Auflage, April 1995

[14] M. Jürgens: LaTeX– Fortgeschrittene Anwendungen oder: Neues von den Hobbits ..., Benutzerhandbuch, FZJ-ZAM-BHB-0135, 1. Auflage, Oktober 1995

[15] J. Heinen: GLI Graphics Language Interpreter Reference Manual, Benutzerhandbuch, KFA-ZAM-BHB-0120, 1. Auflage, November 1995

Danksagungen

Mein besonderer Dank gilt

Herrn Prof. Dr. Reißel für die Übernahme des Hauptreferats und

Herrn Dr. Lustfeld als Koreferent sowie

beiden vorher genannten für ihre fachliche Unterstützung.