

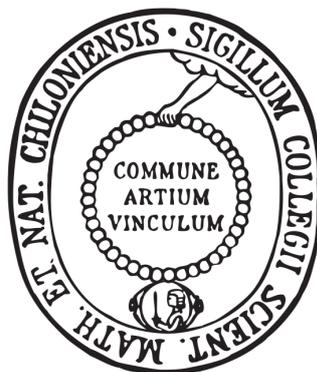
Christian-Albrechts-Universität zu Kiel
Institut für Geowissenschaften

DISSERTATION
ZUR ERLANGUNG DES AKADEMISCHEN GRADES
DOCTOR RERUM NATURALIUM
DER MATHEMATISCH-NATURWISSENSCHAFTLICHEN FAKULTÄT
DER CHRISTIAN-ALBRECHTS-UNIVERSITÄT ZU KIEL

Zur Anwendung von Schwarmintelligenz-Optimierung auf die Dispersionsanpassung von Scholtewellen

VORGELEGT VON
DENNIS WILKEN
GEB. AM 18.12.1979 IN OLDENBURG(OLDB)

KIEL, DEZEMBER 2009



Referent: Prof. Dr. Wolfgang Rabbel

Korreferent: Prof. Dr. Hans-Jürgen Götze

Tag der mündlichen Prüfung: 11.12.2009

Zum Druck genehmigt: Kiel,

Der Dekan

„Das Wort heißt NUKULAR...“
Homer J. Simpson

Zusammenfassung

Die Bestimmung der Scherwellengeschwindigkeit in flachmarinen Sedimenten ist für die Abschätzung der Stabilität des Meeresbodens als Baugrund von großem Interesse. In dieser Arbeit werden die dispersiven Eigenschaften von Scholtewellen genutzt, um Scherwellengeschwindigkeitsprofile als Teil einer Offshore-Baugrunduntersuchung und Überwachung zu invertieren. Dieser Anwendungsbereich stellt hohe Anforderungen an die Auflösung eines Inversionsverfahrens dar, da die erwarteten Änderungen in der Scherwellengeschwindigkeit durch mechanische Belastungen im Baugrund unterhalb des bisherigen Auflösungsvermögens der Methodik vermutet werden. Zur optimierten Anpassung von vorwärts modellierten Dispersionskurven an gemessene Phasenlangsamsamkeitsfrequenzspektren werden daher unterschiedliche Schwarmintelligenz-Optimierungsalgorithmen vorgestellt und in ihrer Leistung untereinander und mit anderen Verfahren wie Simulated-Annealing, der Evolutionsstrategie und der Kovarianz-Matrix-Adaptions Evolutionsstrategie verglichen.

Die Motivation der Nutzung globaler, stochastischer Verfahren in der Inversion von Oberflächenwellen ist die hohe Multimodalität des Problems, welche eine starke Abhängigkeit vom Startmodell bei der Nutzung lokaler Optimierungsverfahren induziert und die Auflösung der Inversion bereits technisch begrenzt. Die eingesetzten Schwarmintelligenzmethoden Partikel-Schwarm-Optimierung und künstlicher Bienenschwarmalgorithmus zeigen sich als effektivste Methoden zur Inversion mit hoher Exploration des Parameterraumes und entsprechend geringer Anforderung an a priori Informationen. Es konnten durch den Einsatz dieser Verfahren Auflösungen von weniger als 8% in den Scherwellengeschwindigkeiten erreicht werden. Im Hinblick auf zu erwartende Korrelation zwischen Schichtmächtigkeiten und Scherwellengeschwindigkeiten sowie den Einfluß der Dichte als Inversionsparameter wird eine Auflösungsanalyse einer gemeinsamen Inversion von Reflektivitäten und Dispersionkurven anhand eines synthetischen Modellbeispiels durchgeführt. Diese Inversion beruht auf dem Multi-Objective-Particle-Swarm Ansatz und nutzt das Prinzip der Pareto-Dominanz, um die Menge gleichwertiger Modelllösungen zu erhalten, ohne eine Gewichtung der gemeinsam invertierten Zielfunktionen vornehmen zu müssen.

Im Rahmen der Baugrundvoruntersuchung des FINO3 Projektes zum Bau einer Offshore-Forschungsplattform werden mehrere profilhafte Messungen zur Anregung von Scholtewellen an der geplanten Lokation der FINO3-Plattform in der Nordsee durchgeführt. Um die Menge der Anwendungsgebiete zu vergrößern wird ein weiterer Testdatensatz in der Ostsee hinzugefügt. Die erhaltenen Scholtewellenfelder wurden dann mit dem letztlich identifizierten effizientesten Optimierungsverfahren, einer hybriden Variante aus Partikel-Schwarm-Optimierung und Downhill-Simplex, invertiert. Die invertierten Scherwellengeschwindigkeitsprofile werden mit hochauflösenden Reflexionsseismiksektionen verglichen und zeigen sehr gute Korrelation mit entsprechenden strukturellen Ergebnissen.

Abstract

The estimation of shear-wave velocities in shallow marine sediments is of major interest when dealing with the estimation of soil stability as foundation soil for offshore buildings. This work will derive shear-wave velocity profiles as part of an offshore building foundation soil analysis by exploiting the dispersive properties of solid-liquid interface waves, so called Scholtewaves. The offshore investigation area makes great demands on the resolution of the inversion when trying to provide the possibility of monitoring changes in shear-wave velocity structure due to mechanical loads by a building. These changes are considered to be smaller than the resolution of so far inversion schemes.

To enhance resolution and performance of optimized matching of surface-wave dispersion curves, different computational swarm-intelligence optimization approaches are tested and compared for performance. To classify these methods, some state-of-the-art stochastic optimization methods like Evolutionary Algorithms, Simulated Annealing and Covariance Matrix Adaption Evolutionary Strategy are also compared with the swarm approaches.

The motivation of using global, stochastic optimization methods for the inversion of surface-waves is the high multimodality of the problem. This induces a strong dependence of the solution from starting models when using local convergent optimization methods and thus reduces the resolution of the derived model. The used swarm-intelligence approaches, Particle Swarm Optimization P.S.O. and Artificial Bee Colony algorithm A.B.C., appear to be the most effective methods with high exploration of the multiple minima of the problem and thus fewer requirements of a priori information. Using these methods, it was possible to enhance the resolution of shear-wave velocity models from the inversion of artificial test datasets to less than 8%.

As the inversion suffers from an expected trade-off between layer thicknesses, densities and shear-wave velocities, a resolution analysis of a joint inversion approach of reflectivity data and Scholtewave dispersion data was tested for an artificial dataset. The optimization was done using a pareto-dominance-based multi-objective particle swarm approach, providing a quantity of equal solutions.

In the course of the FINO3 project aiming at the construction of an offshore research platform in the North Sea, a foundation soil analysis was performed including multiple profiles where Scholtewaves were excited. These profiles were then inverted for shear-wave structure using the identified most effective optimization method, a hybrid interplay of downhill-simplex and particle swarm optimization. A second testing site was chosen in the baltic sea, comprising a single Scholtewave profile. Resulting shear-wave velocity profiles were compared with high-resolution reflection seismic sections and revealed very good correlation with the identified structural events in the reflection data.

Inhaltsverzeichnis

I	1
1 Einführung	3
2 Oberflächenwellen	9
2.1 Ausbreitung von Grenzflächenwellen	10
2.2 Scholtewellen	18
3 Technische Implementierung Scholtewellenmessung	21
3.1 Meßkonzept und Dimensionierung	21
3.2 Entwicklung Seebodengeophon	27
4 Allgemeine Inversionsmethodik	31
4.1 Wellenfeldtransformation	32
4.2 Anpassung der Wellenfeldspektren	34
II	43
5 Motivation eines globalen Suchmusters	45
6 Anforderungen an ein globales Suchmuster	51
6.1 Zufallszahlen	51
6.2 Anpassung der Spektren	52
6.3 Verteilung der Startpopulation	57
7 Angewendete Optimierungsmethoden	59
7.1 Partikel-Schwarm-Optimierung (P.S.O.)	60
7.2 Varianten der PSO: Ein Überblick	64
7.3 ABC	75
7.4 SA/Metropolis	78
7.5 CMA-ES	80
7.5.1 PSO-CMAES-Hybrid	83
8 Methodenvergleich und Auflösung	85

8.1	Das Testproblem	86
8.2	Durchführung der Tests	94
8.3	Auflösung und Effizienz	96
8.4	Statistische Beurteilung	114
III		123
9	Gemeinsame Inversion	125
9.1	Optimierung multipler Zielfunktionen	126
9.2	Multi-Objective-PSO	128
9.3	Gemeinsame Inversion eines synthetischen Datensatzes	132
IV		137
10	Fallstudien zur Inversion	139
10.1	Die FINO3 Baugrundvoruntersuchung	141
10.1.1	Das FINO3 Hintergrundmodell	149
10.2	Lokation Adler-Grund	153
11	Zusammenfassung und Diskussion der Ergebnisse	161
12	Abkürzungsverzeichnis	167
	Literaturverzeichnis	169

Teil I

1 Einführung

Die Baugrunduntersuchung für Offshorebauwerke im flachmarinen Umfeld wie beispielsweise in Schelfgebieten der Nord- und Ostsee wird in den nächsten Jahren aufgrund verstärkter Bauvorhaben weitaus höheren Anforderungen genügen müssen, als dies bislang der Fall ist. Offshore-Windenergieanlagen, Brückenkonstruktionen oder auch Pipeline- und Plattformkonstruktionen sind besondere statische Belastungen für den Meeresboden. Gründungspfähle von Offshorebauwerken werden nach heutigem Stand der Bauverfahrenstechnik vorzugsweise durch Rammen auf Tiefe gebracht. Mit diesem Rammvorgang ist stets ein Verfahrensrisiko durch Auftreffen auf Hindernisse wie Findlinge oder Schichten sehr hoher Lagerungsdichte verbunden (Richwien & Lesny, 2004). Die Entwicklung von geotechnischen Methoden zur hochauflösenden In-Situ Baugrundvorerkundung und Bestimmung der Festigkeit flachmariner Sedimente, also ihres Widerstandes gegenüber statischen und dynamischen Kräften ist daher seit einiger Zeit im Fokus mariner geophysikalischer Forschung. Hierbei ist die Inversion von Scherwellengeschwindigkeiten von zentraler Bedeutung, da die Festigkeit des Meeresbodens durch empirische Beziehungen zwischen Scherfestigkeit, also dem Widerstand, den ein Festkörper Scherkräften entgegensetzt, und Scherwellengeschwindigkeit quantifiziert werden kann (Ayres und Theilen, 1999). Da die Scherwellengeschwindigkeit sowohl In-Situ als auch im Labor gemessen werden kann, können Laboruntersuchungen genutzt werden, um die Felddaten zu kalibrieren. Modelle der Scherwellengeschwindigkeit können somit fundamentale Informationen für die Baugrundvoruntersuchung und Überwachung von Offshorebauwerken liefern.

Bohlen et al. (2004), Klein et al. (2005), Kugler et al. (2005 & 2007) präsentieren eine Methode zur Bestimmung der 2D- und 3D-Verteilung der Scherwellengeschwindigkeit in flachmarinen Sedimenten aus der dispersiven Propagation von aktiv angeregten Oberflächenwellen am Meeresboden, sogenannten Scholtewellen, basierend auf den Arbeiten von McMechan & Yedlin (1981) zur Spektralanalyse von Oberflächenwellen (SASW, Spectral Analysis of Surface Waves) und Luke & Stokoe (1998) zur Anwendung im marinen Bereich. Diese Methode profitiert von der hohen Sensitivität der Scholtewellen gegenüber der Scherwellengeschwindigkeit im Vergleich zu reflektivitätsorientierten Verfahren und ist weniger aufwendig als beispielsweise das direkte Anregen von Scherwellen am Meeresboden (Gimpel, 1987) und ermöglicht eine effektive Inversion von Scherwellengeschwindigkeitsmodellen auf großen Gebieten.

Unter dynamischer Beanspruchung, verursacht durch Strömung, Wellenschlag und Winddruck, kann es in der unmittelbaren Umgebung von Offshorebauwerken

1 EINFÜHRUNG

zu Veränderungen der Sedimentstruktur und deren mechanischen Eigenschaften kommen (Abbildung 1.1). Im Boden werden hierbei mit zyklischer Belastung in unterschiedlichen Richtungen plastische Formveränderungen erzeugt, wodurch eine Pfahlbettung dramatisch herabgesetzt werden kann (Richwien & Lesny, 2004). Zu den Gefährdungen zählen neben zyklischer Entfestigung, die bis zur Boden-

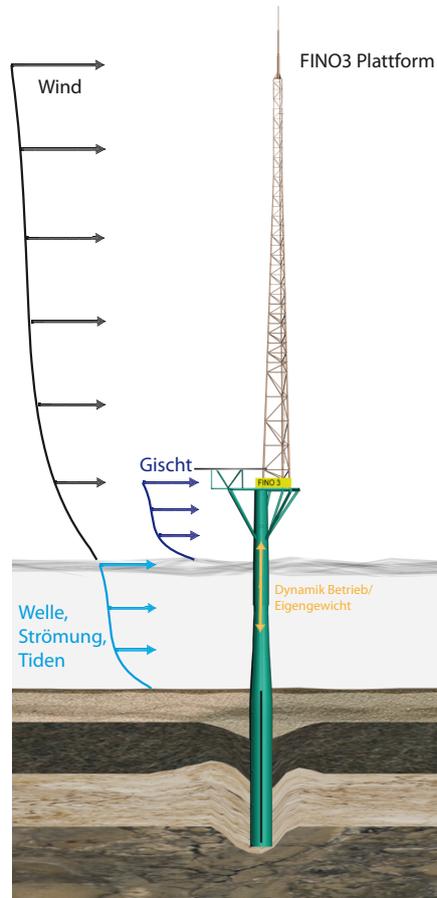


Abbildung 1.1: Skizze zu Belastungspotentialen am Offshorebauwerk FINO3.

verflüssigung führen kann, zum Beispiel auch Kolkbildung, also erosionsbedingte Vertiefungen um das Bauwerk. Solche Veränderungen stellen ein beträchtliches Gefährdungspotential dar. Nicht nur eine auf die Festigkeit des Sediments abzielende Baugrundvoruntersuchung und hochauflösende bildgebende Methoden zur Quantifizierung von Hindernisgefährdungen, sondern in Ergänzung dazu eine In-Situ Überwachung der zeitlichen Entwicklung und räumlichen Ausdehnung der genannten Effekte in der näheren Umgebung des Bauwerks mit seismischen Methoden zu quantifizieren, kann somit von großer Bedeutung sein.

Im Rahmen des FINO3 (**F**orschungsplattformen **I**n Nord- und **O**stsee) - Projektes wird als Teil der Installation einer Offshoreforschungsplattform in der Nordsee ein solches In-Situ Meßprogramm entwickelt und durchgeführt. Es umfaßt hochauflösende reflexionsseismische Verfahren sowie Messungen mit Scholtewellen zur Baugrundvoruntersuchung und Überwachung. Die Reflexionsseismik liefert den strukturellen geologischen Aufbau des Baugrundes, während die Scholtewellen-seismik Informationen über azimutale und laterale Veränderungen der Scherwellengeschwindigkeit respektive Scherfestigkeit ergibt. Tomographische Wiederholungsmessungen für die Überwachung der Bodenveränderungen sollen dann mit Hilfe von Hydrophonen durchgeführt werden, die an der Außenwand der Plattform unterhalb des Meeresbodens installiert sind.

Mit bodenmechanische Laboruntersuchungen wurde des weiteren versucht, die Zusammenhänge zwischen Scherwellengeschwindigkeit und Bodenverformungen bei zyklischer Belastung, wie sie durch die Gründung der Plattform hervorgerufen werden zu quantifizieren. Bei einer Gefährdung mit zunehmendem Verflüssigungspotential ist hier eine Abnahme der Scherwellengeschwindigkeiten zu erwarten. Rabbel & Feeser (2009) führten hierzu Versuche mit isotroper Belastung an einer Kernprobe aus 16 m bis 17 m Tiefe vom Standort der FINO3 Plattform durch. Bei einer zyklischen Belastung von ± 200 kPa bei einer Frequenz von 1 Hz über einen Zeitraum von 4000 Sekunden konnte eine Abnahme der Scherwellengeschwindigkeit von 300 m/s auf 150 m/s festgestellt werden. Mit zunehmender Zahl von Belastungszyklen stieg der Porenwasserdruck sukzessive an. Dieser Anstieg führte zu einer Verringerung der effektiven Spannungen. Die Verflüssigung des Sediments tritt mit Beginn des überproportionalen Anstiegs des Porenwasserdrucks ein, der mit überproportionaler Abnahme der Scherwellengeschwindigkeit verknüpft ist (Rabbel & Feeser, 2009).

Der Fokus dieser Arbeit liegt auf der Bereitstellung eines verbesserten Inversionsverfahrens zur Bestimmung der Scherwellengeschwindigkeiten durch die Anpassung von Dispersionsspektren von Scholtewellen. Des weiteren wird die Entwicklung eines hochauflösenden Seebodengeophons zur Registrierung der Scholtewellen vorgestellt.

Die bei der Überwachung des FINO3 -Baugrundes zu erwartenden Veränderungen in der Scherwellengeschwindigkeit, bedingt durch Veränderungen der Porosität bei Krafteinwirkung durch das Bauwerk, bewegen sich nach Abschätzungen im Bereich von 30 m/s bis 50 m/s. Solche Änderungen sind mit bisherigen Inversionsmethoden nicht aufzulösen. Die Überwachung des Baugrundes von Offshorebauwerken stellt somit hohe Anforderungen an die Auflösung der Inversion als auch an die technische Durchführung der geplanten Überwachungsmessungen. Es muß mitunter die Möglichkeit zur genauen Positionierung der Quellen und Sensorik unter Wasser auf wenige Meter gegeben sein, was durch den Einsatz eines eigens entwickelten hochauflösenden Seeboden-Geophon/Hydrophon-Systems und den

1 EINFÜHRUNG

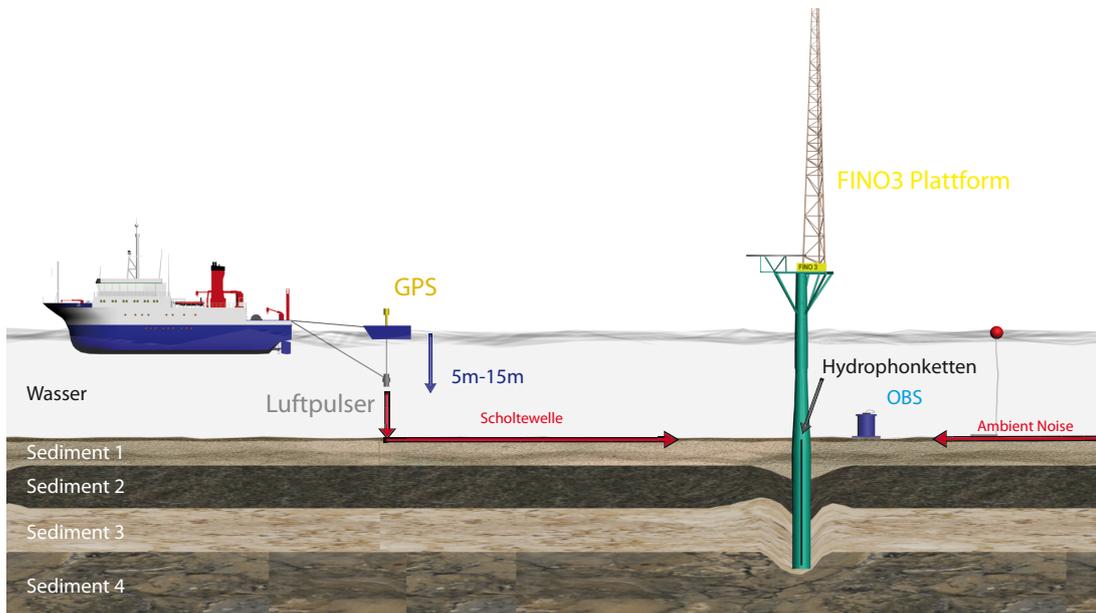


Abbildung 1.2: Meßkonzept an der FINO3 Forschungsplattform. Scholtwellen werden aktiv mit einem Luftpulsler (Airgun) angeregt oder als passiv im Hintergrundrauschen angenommen. Registriert wird mit Seebodenseismometern um die Plattform und Hydrophonketten an der Außenhaut des Monopile.

Einsatz hochfrequenter seismischer Quellen möglich wird (siehe Abschnitt 3.2).

Die oft verwandten Gradientenverfahren zur optimierten Anpassung von modellierten Datensätzen (hier Dispersionskurven von Oberflächenwellen) an gemessene konvergieren effizient, doch sind stets abhängig von der Wahl eines Startmodells. Die Inversion ist somit nicht nur abhängig von a priori Informationen um das Startmodell, sondern konvergiert ausschließlich in das benachbarte Minimum, ohne die Möglichkeit einer weiteren Exploration des umliegenden Parameterraumes. Schon für einen zweidimensionalen Fall läßt sich jedoch zeigen, daß eine Anpassung des Scherwellengeschwindigkeitsmodells durch den Vergleich von Dispersionskinematiken ein hochgradig multimodales Problem darstellt (siehe Kapitel 5). Auch im Nahbereich eines globalen Minimums zeigen sich im Testfall eine hohe Zahl lokaler Minima oder mitunter gleichwertige Lösungen resultierend aus der Unterbestimmtheit des Problems. Es ist daher von großer Bedeutung, nicht nur die Wahrscheinlichkeit zu erhöhen, einem globalen Minimum näher zu kommen, sondern auch die Menge an möglichen Lösungen zu bestimmen.

Es stellt sich somit zuallererst die Aufgabe, zunächst unabhängig der Einflüsse von Meßfehlern und Rauschen, durch die Anwendung stochastischer, globaler Verfahren die algorithmische Auflösung der Inversion in die geforderten Bereiche

zu erhöhen. Dieses Ziel soll durch die Anwendung von Schwarmintelligenzverfahren erreicht werden, welche durch ihre Eigenschaft der Exploration des Parameteraumes und simultanen Konvergenz eine effektive Optimierungsmethode für multimodale Probleme darstellen (Mishra, 2006). Es werden hierbei verschiedene Verfahren der **Partikel-Schwarm-Optimierung (PSO)** (Eberhart & Kennedy, 1995) und das Verfahren der künstlichen Bienenschwärme (Karaboga & Basturk, 2007) auf das Problem angewendet. Die Methode der Partikel-Schwarm-Optimierung nutzt eine Population zufällig erzeugter Parametervektoren, der Partikel, und ordnet jenen Vektoren Bewegungs- oder Geschwindigkeitsvektoren zu. Diese Partikel bewegen sich somit im Parameterraum und teilen untereinander ihre Erfahrung über bereits explorierte Punkte der Zielfunktion und somit erstrebenswerte Positionen im Raum. Zu jedem Iterationsschritt werden die Partikel bewegt und ihr Geschwindigkeitsvektor wird angepaßt durch Addition gewichteter Vektoren in Richtung der global besten gefundenen Position des gesamten Schwarms, ihrer Ursprungsgeschwindigkeit und der besten gefundenen Position jedes individuellen Partikels. Des weiteren werden unterschiedliche Varianten der PSO-Methode untersucht, welche das Verhalten von Konvergenz oder Exploration modulieren.

Zur Einordnung der Methoden in die Menge der stochastischen Optimierungsmethoden wird zusätzlich ein Kovarianzmatrix-Adaptions Evolutionsalgorithmus nach Hansen & Ostermeier (2001), eine $(\mu + \lambda)$ -Evolutionsstrategie (Rechenberg, 1994) sowie eine hybride Optimierungsvariante aus Metropolis-Algorithmus und Downhill-Simplex implementiert und auf das Problem angewendet (siehe Kapitel 7).

Im Rahmen der Baugrundvoruntersuchung des FINO3-Projektes wurden mehrere profilhafte Messungen zur Anregung von Scholtewellen an der geplanten Lokation der FINO3-Plattform in der Nordsee durchgeführt. Diese Messungen sowie eine weitere profilhafte Messung in der südlichen Ostsee werden als Datengrundlage zur ersten Anwendung (siehe Kapitel 10) eines letztlich erhaltenen effizientesten Optimierungsverfahren (siehe Abschnitt 8.3) genutzt. Die invertierten Scherwellengeschwindigkeitsprofile werden mit hochauflösenden reflexionsseismischen Sektionen verglichen und auf ihre Kohärenz geprüft, auch um die Güte der Inversion zu bewerten. Zusätzlich wird die große Anzahl ausgewerteter Vergleichsmodelle während des Inversionsprozesses genutzt, um statistische Aussagen über Auflösung und Parameterkorrelationen der Lösungen zu machen.

Zunächst sei aber eine Einführung in die Grundlagen der Eigenschaften von seismischen Oberflächenwellen sowie in die genutzte Offshoremeßmethodik gegeben.

1 EINFÜHRUNG

2 Oberflächenwellen

Oberflächenwellen sind elastische Wellen, die im Bereich von Grenzflächen signifikante Teilchenbewegung aufweisen. Sie sind hierbei als Eigenschwingungen des Grenzflächenbereichs zu sehen. Es wird in verschiedene Arten von seismischen Grenzflächenwellen unterschieden. Die physikalischen Hauptgruppen definieren sich über die Polarisierung in jene Grenzflächenwellen, welche einen kompressiven Anteil und einen vertikal polarisierten Scheranteil haben (P-SV-Wellen) und jene, welche nur einen horizontal polarisierten Scheranteil aufweisen (SH-Wellen). Eine Klassifizierung der Wellen des P-SV Typs wird des weiteren nach Art der

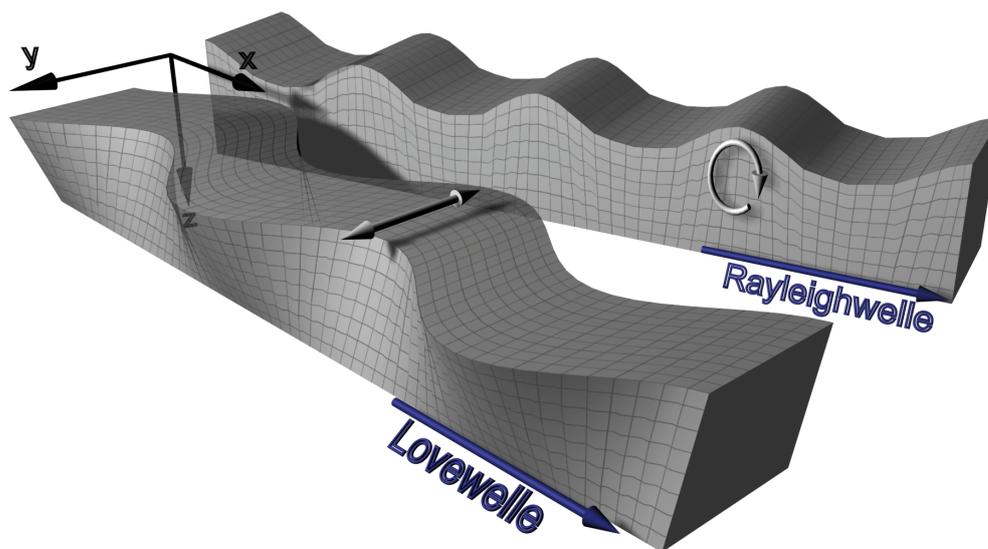


Abbildung 2.1: Illustration der Propagation von Rayleigh(P-SV)- und Love(SH)-Wellen. Die grauen Pfeile zeigen die Teilchenbewegung an.

Grenzfläche vorgenommen. Bei einer Grenzschicht zwischen Festkörper und Vakuum wird von einer Rayleighwelle gesprochen. Die Welle an der Grenzschicht Festkörper zu Flüssigkeit heie Scholtewelle und Grenzflchenwellen an Fest-Fest Grenzflchen werden als Stoneleywellen bezeichnet. Die Wellen des SH-Typs seien als Lovewellen bezeichnet (siehe Abbildung 2.1).

Ist das Propagationsmedium vertikal heterogen, treten also Schichtungen mit

2 OBERFLÄCHENWELLEN

verschiedenen seismischen Parametern auf, so ist die Propagation der Welle dispersiv. Es zeigen sich frequenzabhängige Ausbreitungsgeschwindigkeiten, da die unterschiedlichen Frequenzanteile der Welle verschiedene Amplitudentiefenverteilungen aufweisen und somit in unterschiedlichen Tiefen sensitiv sind (siehe Abbildung 3.3).

2.1 Ausbreitung von Grenzflächenwellen

Die dispersiven Eigenschaften von Oberflächenwellen

In diesem Abschnitt werden nun die grundlegenden Eigenschaften, insbesondere die Dispersion, von Oberflächenwellen anhand der Propagation in einem vertikal heterogenen Medium diskutiert. Die Betrachtungen geschehen hierbei in einem kartesischen Koordinatensystem (x,y,z) mit Tiefe anzeigender z -Achse (siehe Abbildung 2.1). Im Falle einer P-SV-Welle wie Rayleigh- oder Scholtewellen mit einer Polarisierung in der x - y -Ebene beschreiben die homogenen Wellengleichungen für das Vektorpotential $\vec{\Psi} = (\Psi_1, \Psi_2, \Psi_3)^T$ und das Skalarpotential Φ des Vektorfeldes der Verschiebungen $\vec{u}(x, y, z, t)$ die Ausbreitung elastischer Wellen (siehe dazu auch Müller, 1986). Die homogene elastische Wellengleichung für die Verschiebung \vec{u} lautet

$$\rho \frac{\partial^2 \vec{u}}{\partial t^2} = (\lambda + \mu) \vec{\nabla} (\vec{\nabla} \cdot \vec{u}) + \mu \Delta \vec{u} \quad (2.1)$$

oder

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial p_{ij}}{\partial x_j},$$

wobei λ und μ (Schermodul) die Lamé-Konstanten und ρ die Dichte sind. Die zweite Gleichung bezeichnet die Wellengleichung der Komponenten der Verschiebung u_i mit p_{ij} als Komponenten des Spannungstensors (Siehe dazu Kasten 1).

Kasten 1

Ein Körper sei durch Einwirkung äußerer Kräfte deformiert. Gesucht ist eine Größe, welche die Verzerrung eines Körperelementes bei Kraftereinwirkung auf den Körper beschreibt. Der Tensor

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$$

mit $i, j = 1, 2, 3$ liefert diese Beschreibung in Darstellung der Kompressions- (Hauptdiagonalelemente) und Scherdehnungen (Nebenelemente) bei Verschiebung $\vec{u}(x, y, z, t) = (u_1, u_2, u_3)^T$.

Das Hook'sche Gesetz bietet des weiteren eine lineare Beziehung zwischen der Dehnung eines Körpers und den auftretenden Spannungen, p , im Körper im Falle infinitesimaler Deformation.

$$p_{ij} = \lambda \underbrace{\epsilon_{kk}}_{\sum_i \epsilon_{ii}} \delta_{ij} + 2\mu \epsilon_{ij}$$

Hierbei ist δ_{ij} das Kroneckersymbol.

Aus dem zweiten Newton'schen Axiom folgt die Proportionalität von Beschleunigung und der Kraft der inneren Spannungen, $\frac{\partial p_{ik}}{\partial x_k}$.

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial p_{ij}}{\partial x_j}$$

(siehe auch Aki & Richards (1980) und Müller (1986))

Durch Substitution des Vektorfeldes \vec{u} durch die Potentiale $\vec{\Psi}$ und Φ

$$\vec{u} = \vec{\nabla} \phi + \vec{\nabla} \times \vec{\Psi}$$

nach dem Helmholtz-Theorem ist dann mit $\alpha^2 = \frac{\lambda+2\mu}{\rho}$ und $\beta^2 = \frac{\mu}{\rho}$

$$\vec{\nabla}^2 \Phi - \frac{1}{\alpha^2} \ddot{\Phi} = 0 \tag{2.2}$$

$$\vec{\nabla}^2 \Psi_2 - \frac{1}{\beta^2} \ddot{\Psi}_2 = 0 . \tag{2.3}$$

Im Fall von Oberflächenwellen des Typs SH, welche in y-Richtung polarisiert sind, gilt ohne Volumenkraftanteil in Gleichung 2.1

2 OBERFLÄCHENWELLEN

$$\vec{\nabla}^2 v - \frac{1}{\beta^2} \ddot{v} = 0 . \quad (2.4)$$

Betrachtet sei in unserem kartesischen Koordinatensystem (x,y,z) ein lateral homogenes Medium mit vertikaler z -Achse, welche die Tiefe anzeigt. Angenommen sei nun eine Oberflächenwelle der Frequenz ω , welche in x -Richtung propagiert. Für den Verschiebungsvektor \vec{u} sei für Oberflächenwellen horizontaler Ausbreitung eine mit der Tiefe abnehmende Amplitude zu erwarten. Für jede Schicht j sei also folgender Ansatz einer ebenen Welle gemacht

$$\Phi_j = A_j(z) e^{i(kx - \omega t)} \quad (2.5)$$

$$\Psi_j = B_j(z) e^{i(kx - \omega t)} \quad (2.6)$$

$$v_j = C_j(z) e^{i(kx - \omega t)} , \quad (2.7)$$

wobei k die Wellenzahl ist. Ausgehend von den Bedingungen an existierende Wellen, daß

- obige Wellengleichungen in jeder Schicht j gelten,
- eine freie Oberfläche an $z = 0$ vorliegt, also die Spannungen verschwinden,
- die Stetigkeit der Lösung an Schichtgrenzen gegeben ist
- und die Amplitude der Oberflächenwelle für $z \rightarrow \infty$ verschwinden muß,

kann auf Bedingungen für Lösungen oder Wellenrealisierungen geschlossen werden. Dies soll im folgenden nun geschehen. Es ist zu bedenken, daß obige Lösung bei diesen Bedingungen für unabhängige ω und k nicht möglich ist (Aki & Richards, 1980). Es kann jedoch eine Lösung gefunden werden, wenn ω unabhängig und $k = k_n(\omega)$. Hierbei wird mehr als nur ein k mögliche Lösungen für Oberflächenwellen repräsentieren.

Es sei nun das einfache Beispiel eines Mediums bestehend aus einer einzelnen Schicht (Dichte ρ_1 , Scherwellengeschwindigkeit β_1 und Kompressionswellengeschwindigkeit α_1) über einem Halbraum (Dichte ρ_2 , Scherwellengeschwindigkeit β_2 und Kompressionswellengeschwindigkeit α_2) in der Tiefe $z = H$ betrachtet. Es folgt dann aus den obigen Ansätzen eingesetzt in die Wellengleichungen mit

$$\frac{\partial \vec{W}_j}{\partial t} = \frac{\partial}{\partial t} (\Phi_j, \Psi_j, v_j)^T = -i\omega \vec{W}_j$$

$$\frac{\partial \vec{W}_j}{\partial x} = ik \vec{W}_j$$

für jede Schicht j

2.1 Ausbreitung von Grenzflächenwellen

$$\frac{\partial^2 C_j}{\partial z^2} + k^2 \left(1 - \frac{c^2}{\beta_j^2}\right) C_j = 0 \quad (2.8)$$

$$\frac{\partial^2 A_j}{\partial z^2} + k^2 \left(1 - \frac{c^2}{\alpha_j^2}\right) A_j = 0 \quad (2.9)$$

$$\frac{\partial^2 B_j}{\partial z^2} + k^2 \left(1 - \frac{c^2}{\beta_j^2}\right) B_j = 0, \quad (2.10)$$

wobei $c^2 = \frac{\omega^2}{k^2}$. Erhalten bleiben also Differentialgleichungen für die z-Abhängigkeit. Diese Art von Differentialgleichung besitzt nach Aki & Richards (1980) folgende allgemeine Lösung.

$$\begin{aligned} C_j(z) &= C_{1j} e^{-g_j z} + C_{2j} e^{g_j z} \\ A_j(z) &= A_{1j} e^{-g'_j z} + A_{2j} e^{g'_j z} \\ B_j(z) &= B_{1j} e^{-g_j z} + B_{2j} e^{g_j z} \end{aligned} \quad (2.11)$$

mit

$$g_j = \sqrt{k^2 - \frac{\omega^2}{\beta_j^2}} = \sqrt{\frac{\omega^2}{c^2} - \frac{\omega^2}{\beta_j^2}}; \quad g'_j = \sqrt{k^2 - \frac{\omega^2}{\alpha_j^2}}$$

Aus der Bedingung, daß die Amplitude der Welle für $z \rightarrow \infty$ im Halbraum verschwinden muß, folgt daß g_j reell sein muß (also $c < \beta_j$) und mit (am Beispiel der Lösung für die $A_j(z)$)

$$\lim_{z \rightarrow \infty} A_{1j} e^{-g'_j z} = 0$$

$$\lim_{z \rightarrow \infty} A_{2j} e^{g'_j z} \rightarrow \infty$$

folgt, daß $A_{2j} = 0$. Also lauten die Lösungen der Differentialgleichungen für die Tiefenabhängigkeit im Halbraum

$$A_j(z) = A_{1j} e^{-g'_j z}$$

$$B_j(z) = B_{1j} e^{-g_j z}$$

$$C_j(z) = C_{1j} e^{-g_j z}.$$

Es seien nun im folgenden exemplarisch die vertikalen Verschiebungen v in y-Richtung betrachtet, also SH-Wellen. Für die obere Schicht $j = 1$ des Beispielm mediums war eine freie Oberfläche vorausgesetzt. Diese Nebenbedingung bringt

$$\frac{\partial v}{\partial z} \Big|_{z=0} = 0.$$

2 OBERFLÄCHENWELLEN

woraus mit der allgemeinen Lösung für $C_j(z)$ folgt, daß

$$C_{11} = C_{21} \quad [0 < z < H].$$

Es bleiben also für die Verschiebungen v folgende Gleichungen für die beiden Bereiche des Mediums, obere Schicht und Halbraum:

$$v = [C_{11}e^{-g_1z} + C_{11}e^{g_1z}] e^{i(kx-\omega t)} \quad [0 < z < H]$$

$$v = [C_{12}e^{-g_1z}] e^{i(kx-\omega t)} \quad [z > H]$$

Für die beiden Bereiche gelte des weiteren Stetigkeit an der Schichtgrenze $z = H$ für die Verschiebungen sowie die Spannungskomponenten. Es ist also für die Verschiebungen

$$\begin{aligned} C_{11} \underbrace{(e^{-g_1H} + e^{g_1H})}_{2\cos(ig_1H)} &= C_{12}e^{-g_2H} \\ \Rightarrow \frac{C_{12}}{C_{11}} &= \frac{2\cos(ig_1H)}{e^{-g_2H}}. \end{aligned}$$

Für die Stetigkeit der Spannung gelte $\mu \frac{\partial v}{\partial z} = p_{zy} = \text{konstant}$, also

$$\frac{C_{12}}{C_{11}} = \frac{2i\mu_1g_1\sin(ig_1H)}{\mu_2g_2e^{-g_2H}}.$$

Durch Gleichsetzen der beiden Gleichungen folgt

$$\tan(ig_1H) = \frac{\mu_2g_2}{i\mu_1g_1}.$$

Aus dieser Bedingung folgen Wellenrealisierungen mit einer Wellenzahl k_n . Mit eingesetztem g_j ist somit

$$\tan(i\omega H \sqrt{\frac{1}{c^2} - \frac{1}{\beta_1^2}}) = \frac{\mu_2}{i\mu_1} \frac{\sqrt{\frac{1}{c^2} - \frac{1}{\beta_2^2}}}{\sqrt{\frac{1}{c^2} - \frac{1}{\beta_1^2}}}$$

oder

$$\tan(\omega H \sqrt{\frac{1}{\beta_1^2} - \frac{1}{c^2}}) = \frac{\mu_2}{\mu_1} \frac{\sqrt{\frac{1}{c^2} - \frac{1}{\beta_2^2}}}{\sqrt{\frac{1}{\beta_1^2} - \frac{1}{c^2}}} \quad (2.12)$$

2.1 Ausbreitung von Grenzflächenwellen

Die damit zu erhaltenden Lösungen der Wellengleichung für die Verschiebungen v beschreiben damit eine Sinus-Oszillation in Abhängigkeit von der Tiefe in der oberen Schicht und einen exponentiellen Abfall der Amplitude mit der Tiefe im Halbraum (Aki & Richards, 1980).

Die möglichen Lösungen der Gleichung 2.12 und damit realisierte Wellen mit Phasengeschwindigkeit $c(\omega)$ lassen sich am besten graphisch darstellen. Substituiere dazu

$$X(c) = \omega H \sqrt{\beta_1^{-2} - c^{-2}},$$

so wird die rechte Seite von Gleichung 2.12 zu

$$f(x, \omega)_1 = \frac{\mu_2}{\mu_1 X} \sqrt{\omega^2 H^2 (\beta_1^{-2} - \beta_2^{-2}) - X^2}$$

und die linke Seite zu

$$f(x, \omega)_2 = \tan(X).$$

Trägt man nun diese beiden Funktionen über X auf, so ergibt sich Abbildung 2.2 oben. Lösungen von Gleichung 2.12 entsprechen hier Schnittpunkten der Funktionen, also $f_1(\text{hellblau}) = f_2(\text{dunkelblau})$. Die Funktion f_1 ist zwischen $X = 0$ und ihrer Nullstelle

$$X_0 = \omega H \sqrt{\beta_1^{-2} - \beta_2^{-2}}$$

reell. Diese Nullstelle wandert in Abhängigkeit von ω nach rechts. Mit steigendem ω existieren also immer mehr Lösungen mit Phasengeschwindigkeiten $c_n(\omega)$. Der Schnittpunkt $X_i = X_i(\omega)$ ergibt mit $c(X)$ die Dispersionsgleichung einer i -ten Mode der Welle.

$$c_i(\omega) = \frac{\omega H}{\sqrt{\frac{\omega^2 H^2}{\beta_1^2} - X_i^2(\omega)}}$$

Es entstehen Dispersionskurven verschiedener Oberschwingungen der Welle, wie im unteren Teil der Abbildung 2.2 dargestellt ist. Die 1. Mode sei als Fundamentalmode bezeichnet. Der hier betrachtete Zwei-Schichten-Fall erläutert so die grundlegendsten Charakteristika, insbesondere die Dispersion von Oberflächenwellen. Für den n -Schichten-Fall ist das Vorgehen jedoch komplexer und soll im folgenden Abschnitt kurz behandelt werden.

Der Thomson-Haskell-Formalismus

Für ein geschichtetes Medium können Verschiebungsvektoren des Wellenfeldes als Lösungen bestimmter Dispersionsrelationen mit verschiedensten Lösungsverfahren für die Gleichungen zur z -Abhängigkeit von Verschiebungen und Spannungskomponenten bestimmt werden. Einen Überblick dazu bieten bspw. Buchen & Ben-Hador (1996). Zur Vorwärtsmodellierung von Phasenlangsamer-Frequenzspektren wurde in dieser Arbeit auf das Verfahren von Wang (1999) zurückgegriffen. Die Methodik von Wang (1999) gehört zu einer Klasse von Methoden

2 OBERFLÄCHENWELLEN

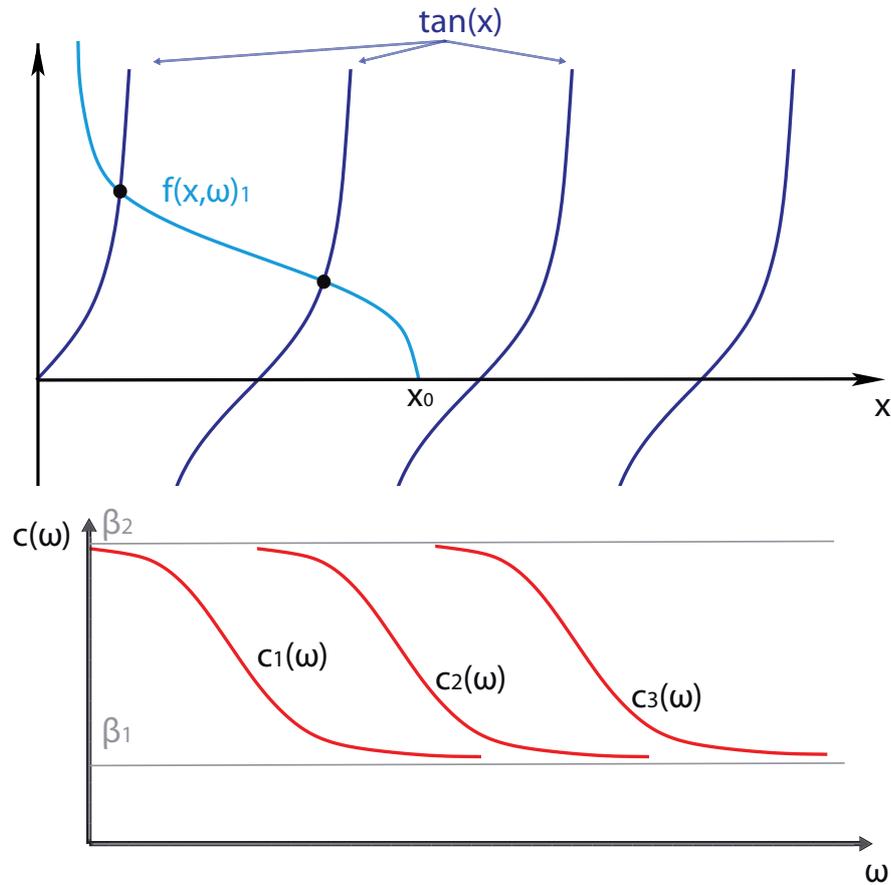


Abbildung 2.2: Darstellung von möglichen Wellenlösungen für Oberflächenwellen (oben) und ihrer Dispersionskurven (unten).

basierend auf dem Verfahren von Thomson (1950) und Haskell (1951). Im folgenden soll der Formalismus dieses Verfahrens kurz erläutert werden.

Am Beispiel der SH-Wellen sei für die Lösung der Wellengleichung

$$\rho \frac{\partial^2 \vec{u}}{\partial t^2} = (\lambda + \mu) \vec{\nabla} (\vec{\nabla} \cdot \vec{u}) + \mu \Delta \vec{u}$$

analog zum Ansatz 2.7

$$\vec{u} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ Z_1(k, z, \omega) e^{i(kx - \omega t)} \\ 0 \end{pmatrix}.$$

Die Spannungskomponenten sind in diesem Fall

$$p_{xx} = p_{yy} = p_{zz} = p_{zx} = 0$$

2.1 Ausbreitung von Grenzflächenwellen

$$p_{yz} = \mu \frac{dZ_1}{dz} e^{i(kx - \omega t)}$$

$$p_{xy} = ik\mu Z_1 e^{i(kx - \omega t)}$$

Das Einfügen von obigen Verschiebungen und Spannungskomponenten in die Wellengleichung bringt eine Differentialgleichung für Z_1

$$-\omega^2 \rho(z) Z_1 = \frac{d}{dz} \left[\mu(z) \frac{dZ_1}{dz} \right] - k^2 \mu(z) Z_1, \quad (2.13)$$

analog dem Vorgehen zur Bestimmung der A_j , B_j , C_j im vorigen Abschnitt. Definiere nun die z -Abhängigkeit der Spannungskomponente p_{yz} mit

$$p_{yz} = Z_2(k, z, \omega) e^{i(kx - \omega t)}$$

wobei also

$$\frac{dZ_1}{dz} = \frac{Z_2}{\mu}$$

und nach der Differentialgleichung 2.13

$$\frac{dZ_2}{dz} = (k^2 \mu(z) - \omega^2 \rho(z)) Z_1$$

In Matrixform lauten diese Gleichungen dann

$$\frac{d}{dz} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 0 & \mu(z)^{-1} \\ k^2 \mu(z) - \omega^2 \rho(z) & 0 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \quad (2.14)$$

Der Vektor $(Z_1, Z_2)^T$ ist somit der Bewegungs-Spannungs-Vektor für SH-Wellen. Z_1 beschreibt die z -Abhängigkeit der Verschiebung v und $Z_2 = \mu \frac{dZ_1}{dz}$ die Spannungskomponente.

Der Bewegungs-Spannungs-Vektor für P-SV-Wellen kann in analoger Weise erhalten werden. Mit den Ansätze der Verschiebungen u und w ($v = 0$) und den Gleichungen für nicht verschwindende Spannungskomponenten sowie eine erneute Substitution der z -Abhängigkeiten folgt ein Gleichungssystem vierter Ordnung (siehe Aki & Richards, 1980).

Die Propagatormatrix-Methode nach Thomson (1950) und korrigiert von Haskell (1951) löst nun die beschriebenen Gleichungen der allgemeinen Form

$$\frac{d}{dz} \vec{f}(z) = A(z) \vec{f}(z)$$

2 OBERFLÄCHENWELLEN

mit \vec{f} als n-dimensionaler Verschiebungs-Spannungs-Vektor und A eine $n \times n$ -Matrix mit $n=2$ für SH-Wellen und $n=4$ für Rayleighwellen.

Ein zugrunde liegendes, vertikal heterogenes Medium ($\mu(z)$, $\rho(z)$ für SH) mit einer freien Oberfläche in der Tiefe z_0 sei repräsentiert durch eine Reihe homogener Schichten (μ_k , ρ_k) $k = 1, n_S$ über einem Halbraum. Das Lösungsverfahren basiert nun auf der Formulierung einer Schichtmatrix und dem Propagieren der Lösung durch die Schichten unter der Bedingung der Stetigkeit. Eine sogenannte Propagatormatrix sei also definiert durch

$$P(z, z_0) = I + \int_{z_0}^z A(\xi_1) d\xi_1 + \int_{z_0}^z A(\xi_1) \int_{z_0}^{\xi_1} A(\xi_2) d\xi_2 d\xi_1 + \dots$$

mit I als Einheitsmatrix. Es läßt sich zeigen, daß P hierbei folgende Bedingung erfüllt (Aki & Richards, 1980).

$$\vec{f}(z) = P(z, z_0) \vec{f}(z_0)$$

P erzeugt also den Bewegungs-Spannungs-Vektor an z durch Anwendung auf den Vektor an der Oberfläche $z = z_0$. Nun sei, wie vorausgesetzt, $A(z)$ konstant in einer Schicht. Es ist dann für die Propagatormatrix

$$P(z, z_0) = I + (z - z_0)A + \frac{1}{2}(z - z_0)^2 AA + \dots = e^{(z-z_0)A}$$

Für ein geschichtetes Medium aneinander gereihter homogener Schichten ist so für $z_{k-1} < z < z_k$

$$\vec{f}(z) = P(z, z_{k-1})P(z_{k-2}, z_{k-1}) \dots P(z_1, z_0) \vec{f}(z_0)$$

Es kann also mit Hilfe der Randbedingungen der freien Oberfläche der Bewegungs-Spannungs-Vektor für jede Schicht mit einer solchen Kettenregel gewonnen werden.

Wang (1999) erweiterte nun dieses Verfahren zur Korrektur verschiedener numerischer Instabilitäten bei Wellen sehr kleiner Amplitude. Darüber hinaus bietet der Code von Wang (1999) nicht nur Lösungen für Love- und Rayleighwellen sondern auch für Scholtewellen.

2.2 Scholtewellen

Wie bereits für dieses Kapitel einleitend erwähnt, können neben Rayleighwellen, Stoneleywellen und Lovewellen auch Oberflächenwellen an einer fest-flüssig Grenzschicht angeregt werden. Diese Wellen wurden als Scholtewellen eingeführt. Die Propagation dieser Wellen ist ähnlich jener der Rayleighwellen.

Betrachtet sei hierzu ein einfaches Modellbeispiel eines Halbraumes festen Materials (Kompressions- und Scherwellengeschwindigkeit α_s , β_s und Dichte ρ_s) und

einem an der Tiefe $z=0$ anschließenden Halbraum flüssigen Materials (Kompressionswellengeschwindigkeit α_f und Dichte ρ_f).

Für Oberflächenwellen an der Grenzschicht der beiden Halbräume gilt, daß alle Scherspannungen verschwinden müssen, für alle anderen Spannungskomponenten Stetigkeit an der Grenzschicht

$$p_{xz}|_{z=0} = 0$$

$$(p_{zz})_f|_{z=0} = (p_{zz})_s|_{z=0}$$

und Stetigkeit für die vertikalen Verschiebungen vorliegen muß

$$(u_z)_f|_{z=0} = (u_z)_s|_{z=0} .$$

Eine Gleichung zur Realisierung und Beschreibung der Propagationsgeschwindigkeit solcher Wellen, analog der Bedingung 2.12, kann aus den Ansätzen für die z -Abhängigkeit ebener Wellen, 2.5, 2.6 und 2.7 und den obigen Randbedingungen an der Oberfläche analog Abschnitt 2.1 erhalten werden:

Aus den Ansätzen ebener Wellen für die Wellengleichung folgen Differentialgleichungen für die z -Abhängigkeiten der Amplituden. Für diese wird ein Ansatz in der Form von Gleichung 2.11 gemacht. Aus den Randbedingungen und der Bedingung, daß die Amplitude mit zunehmendem Abstand von der Grenzfläche verschwinden muß, können dann Bedingungen für Kombinationen der konstanten Vorfaktoren dieser Ansätze bestimmt werden. Dies liefert die Scholtegleichung (Scholte, 1942) welche die Realisierungslösungen und die Propagationsgeschwindigkeit von Scholtewellen an einer fest-flüssig Grenzschicht beschreibt.

$$4\sqrt{1 - \frac{c^2}{\alpha_s^2}}\sqrt{1 - \frac{c^2}{\beta_s^2}} - (2 - \frac{c^2}{\beta_s^2})^2 = m(2 - \frac{c^2}{\beta_s^2})^2 \frac{\sqrt{1 - \frac{c^2}{\alpha_s^2}}}{\sqrt{1 - \frac{c^2}{\alpha_f^2}}}$$

Hierbei ist $m = \rho_f/\rho_s$ und c die Propagationsgeschwindigkeit der Scholtewelle. Für Rayleighwellen gelten hingegen im Falle eines Halbraumes mit freier Oberfläche, daß alle Spannungskomponenten an der Oberfläche verschwinden müssen,

$$p_{xz}|_{z=0} = p_{zz}|_{z=0} = 0 .$$

Die Scholtewellengleichung gilt so für Rayleighwellen mit $m=0$ und liefert die Rayleighgleichung

$$4\sqrt{1 - \frac{c^2}{\alpha_s^2}}\sqrt{1 - \frac{c^2}{\beta_s^2}} - (2 - \frac{c^2}{\beta_s^2})^2 = 0 .$$

Die Scholtewelle kann somit als eine modifizierte Rayleighwelle gesehen werden. Diese Scholtegleichung hat reelle Lösungen für mögliche Oberflächenwellen, wenn

2 OBERFLÄCHENWELLEN

$0 \leq c \leq \beta_s$. Mit der Berechnung der Verschiebungen aus den erhaltenen Lösungen läßt sich zeigen, daß die Teilchenbewegung der Scholte- und Rayleighwellen elliptisch in Richtung der Ausbreitung und senkrecht zur freien Oberfläche orientiert verläuft (siehe Abbildung 2.1 und Abbildung 2.3) und damit als P-SV Welle charakterisiert wird (z.B. d'Arnaud Gerkens (1989), Meegan et al. (1999)). Die Amplitude der Partikelbewegung fällt exponentiell mit dem Abstand zur Oberfläche in den Halbräumen. Die Horizontalkomponente der Verschiebungen weist eine Sprungstelle an der Oberfläche auf (siehe Abbildung 2.3). Im hier be-

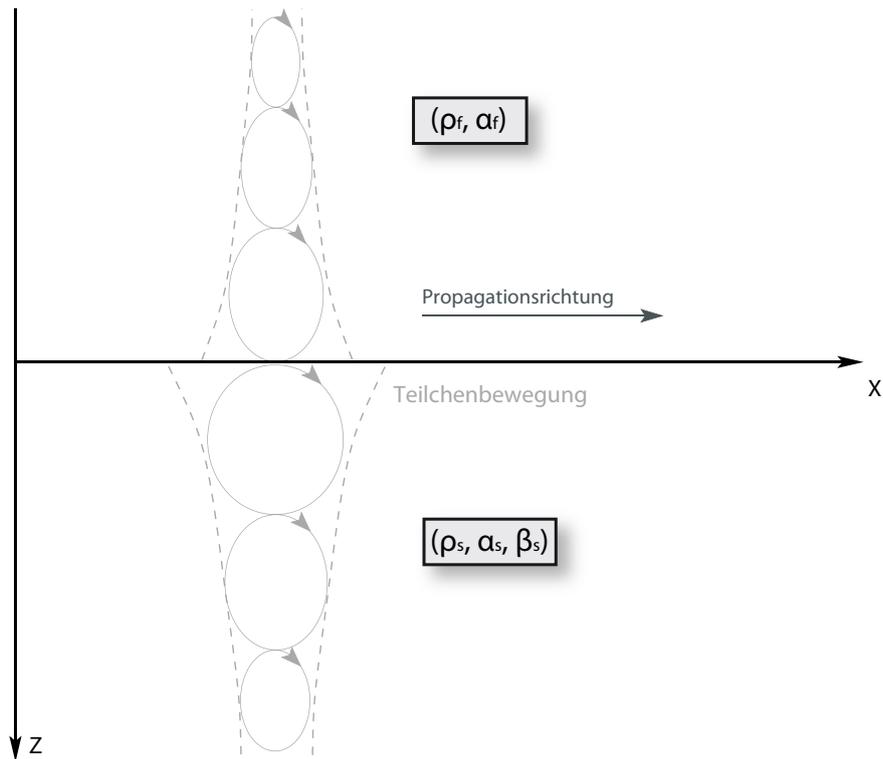


Abbildung 2.3: Schematische Darstellung der Partikelbewegung einer Scholtewelle an einer Grenzfläche zwischen einem Halbraum festen und einem Halbraum flüssigen Materials.

trachteten Fall von zwei Halbräumen festen und flüssigen Materials weist sowohl die Rayleigh- als auch die Scholtewelle keine Dispersion auf. Die Scholtgleichung ist unabhängig von der Frequenz. Im Falle weiterer Schichten werden die Wellenrealisierungen jedoch frequenzabhängig analog der Betrachtung zur SH-Welle in Abschnitt 2.1.

3 Technische Implementierung einer Scholtewellenmessung

Die Anregung und Registrierung von Scherwellen und Oberflächenwellen am Meeresboden wurde bereits mit verschiedensten Meßkonstellationen durchgeführt. Gehrman et al. (1984), Gimpel (1987) und Essen et al. (1998) zum Beispiel konzentrieren sich auf die Anregung und mehrkomponentige Aufzeichnung direkt am Meeresboden und setzen hierbei Torsions-Quellen zur Anregung von Scherwellen oder Implosionsquellen und stationäre Unterwassergeophone ein. Diese Verfahrenstechnik erlaubt jedoch nur langsamen Meßfortschritt bei hohem technischen Aufwand (Bussat, 2006). Klein et al. (2005) schlugen die Anregung von Scholtewellen in der Wassersäule durch einen geschleppten Luftpulser (Airgun) vor. Die Aufzeichnung sollte durch einen tief geschleppten Streamer erfolgen. Hierbei ist jedoch keine mehrkomponentige Auswertung des Wellenfeldes möglich, da nur der Kompressionsanteil der Oberflächenwellen in der Wassersäule registriert werden kann (Bussat, 2006). Kugler et al. (2005), Bohlen et al. (2004) und Bussat (2006) modifizierten diese Methode daher um den Einsatz von Ozeanbodenseismometern (OBS), wobei die vom Schiff geschleppte Luftpulserquelle an unterschiedlichen Entfernungen zum stationären OBS gezündet wird. Man erhält über einer Profifahrt somit eine entfernungsabhängige Sektion mit gemeinsamem Empfänger (Common-Receiver-Gather). Je nach Dämpfung der Oberflächenwellen müssen gegebenenfalls mehrere OBS auf einem Profil eingesetzt werden. Dieses Kapitel befaßt sich nun mit der genaueren Beschreibung der letztgenannten Methode, welche im Rahmen der FINO3-Baugrundvoruntersuchung eingesetzt wurde (Wilken et al., 2009). Des weiteren folgt ein Abschnitt über die Entwicklung eines hochauflösenden Seebodengeophons, welches speziell für die Anwendungen an FINO3 im Rahmen dieser Arbeit entwickelt wurde. Das System zeichnet sich durch hohe mögliche Abtastraten (>11 kHz) aus, was eine Relokalisierung des Systems am Meeresboden unter Benutzung einer hochfrequenten seismischen Quelle ermöglicht.

3.1 Meßkonzept und Dimensionierung

Abbildung 3.1 zeigt das Prinzip der Anregung und Akquise von Scholtewellen im flachmarinen Umfeld. Es werden Scholtewellen durch den Einsatz eines Luftpulser in der Wassersäule am Meeresboden angeregt und mit einem am Meeresbo-

3 TECHNISCHE IMPLEMENTIERUNG SCHOLTEWELLENMESSUNG

den abgesetzten Geophon, Ozeanbodenseismometer oder Hydrophon registriert (Klein et al., 2005; Bohlen et al., 2004; Kugler et al., 2005; Bussat, 2006). Der Luftpulser mit einem eingesetzten Kammervolumen von 0.6 Litern bis 2.5 Litern wird hierbei hinter einem Schiff geschleppt und sein Schwimmkörper mit einer GPS-Antenne zur Positionierung versehen. Während der Überfahrt des Schiffes über den Sensor wird die Quelle in definierten Zeitintervallen von einigen Sekunden ausgelöst. Die Schüsse erfolgen damit an zu- oder abnehmender Entfernung („Offset“) vom Sensor. Nach dem Schneiden der registrierten Daten nach den Schußzeiten (siehe Abschnitt 3.2) entsteht eine entfernungsabhängige Sektion *Common Receiver Gather* (CRG) als Profelsektion verankert am abgesetzten Sensor. Abbildung 3.2 zeigt ein Beispiel einer solchen Seismogrammsektion.

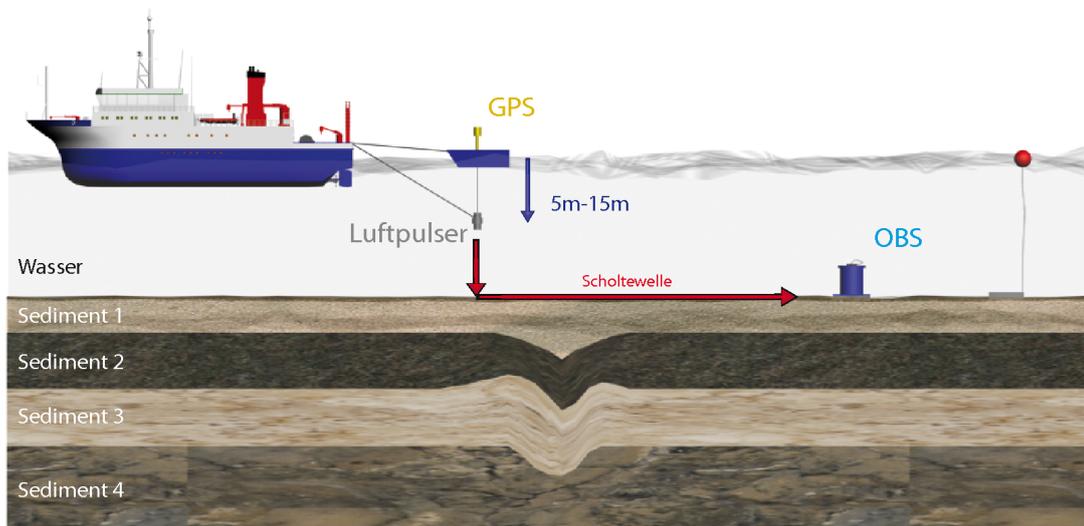


Abbildung 3.1: *Prinzipskizze der eingesetzten marinen Scholtewellen-Akquisitionsmethodik.*

Das Frequenzband der aktiv angeregten Scholtewelle ist definiert durch den Frequenzgehalt der eingesetzten Quelle. Es wurde daher der Einfluß der Quelltiefe (Abb. 3.1) auf das Spektrum der angeregten Scholtewelle von Kugler et al. (2005) und Bussat (2006) numerisch sowie experimentell untersucht. Kugler et al. (2005) zeigten mit Hilfe einer Modellierungsstudie, daß die Amplitude der Scholtewelle mit zunehmendem Abstand vom Meeresboden exponentiell abnimmt und gleichzeitig wird der Frequenzgehalt des Spektrums stark durch die Quelltiefe beeinflusst, da der exponentielle Abfall der Scholtewellenenergie in der Wassersäule mit höheren Frequenzen steiler wird (siehe Abb. 3.3). So wird nicht nur die unterschiedliche Dämpfung der Frequenzen bereits in der Wassersäule durch die

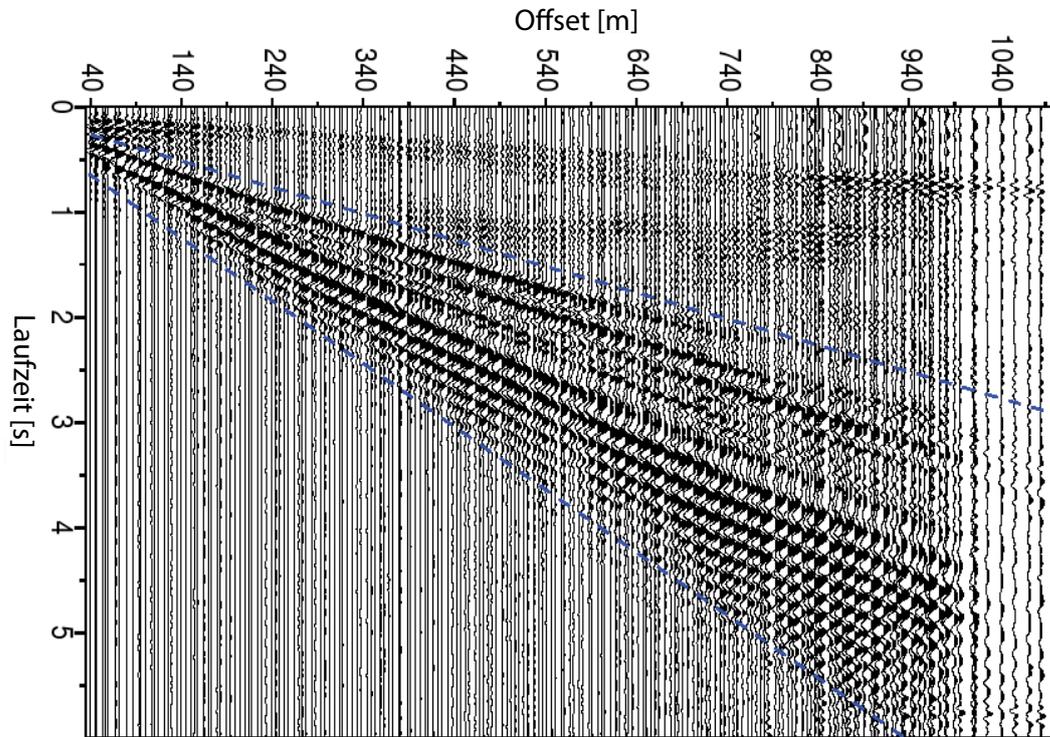


Abbildung 3.2: Beispiel für ein Common Receiver Gather, aufgenommen nach dem Prinzip in 3.1. Der Sensor befindet sich bei Offset 0.0. Es läßt sich klar der Einsatz der Scholtewelle erkennen (eingegrenzt durch gestrichelte, blaue Linie). Die Daten dieses Beispiels sind tiefpaßgefiltert (<20 Hz) und spurnormiert.

Entfernung von Quelle zum Meeresboden moduliert, auch die Resonanzfrequenz der Luftpulserblase ist definiert durch die Tiefe (Safar 1979). So gilt für die Zentralfrequenz

$$f_c = 51.5V^{-1/3}P^{-1/3.24}(1 + z/10)^{0.81}$$

wobei V das Volumen der Airgunkammer und P der Kammerdruck und z die Tiefe sind. Es ist also zu beachten, daß sich mit zunehmender Quelltiefe die Amplitude sowie der Frequenzgehalt zu höheren Frequenzen erhöht. Jedoch wird damit die Quellposition schwerer zu bestimmen, da GPS-Antenne und Schwimmkörper zur Quelle stärker verschert werden können.

Zur Wahl der weiteren Meßparameter wie Schußintervall, Aufzeichnungsdauer, Offsetintervall und Geschwindigkeit des Schiffes gibt es zwei definierende Bedingungen. Zum einen muß das volle Wellenfeld der Scholtewelle registriert sein, bevor der nächste Schuß abgegeben wird, um Interferenzen zu vermeiden, und zum anderen muß der Schußabstand so gewählt sein, daß kein räumlicher Alias-Effekt („Aliasing“) im interessanten Frequenz-Langsamkeitsbereich auftritt. So gilt für

3 TECHNISCHE IMPLEMENTIERUNG SCHOLTEWELLENMESSUNG

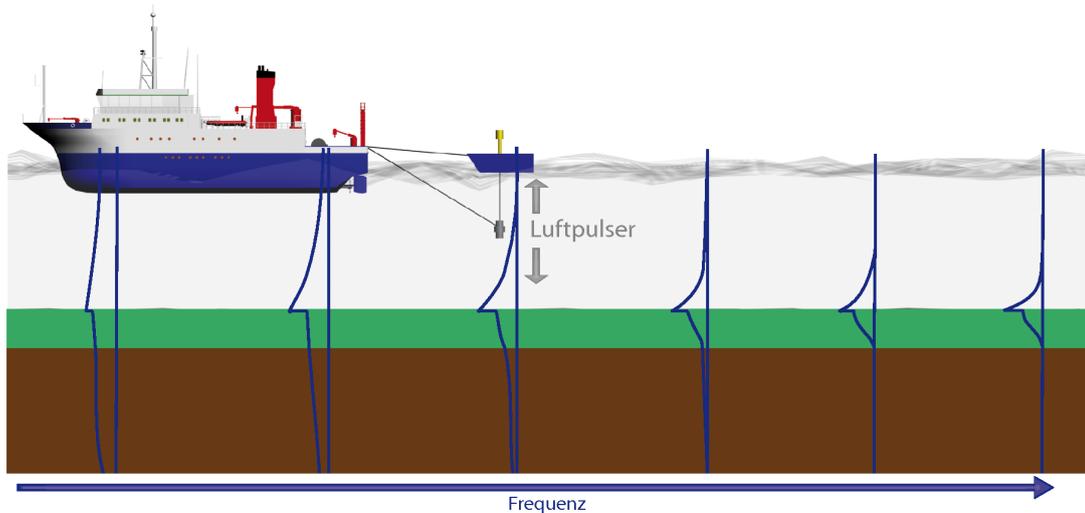


Abbildung 3.3: Schema zur Amplitudentiefenverteilung der Scholtewelle bei verschiedenen Frequenzen.

die Distanz zwischen zwei Schüssen Δx , um räumliches Aliasing zu vermeiden,

$$\Delta x \leq \frac{v_{min}}{2f_{max}} \quad (3.1)$$

mit V_{min} als langsamste Scholtewellengeschwindigkeit der Fundamentalmode und f_{max} als entsprechende maximaler Frequenz der Scholtewelle. Um Interferenz zwischen den von verschiedenen Schüssen angeregten Scholtewellenfeldern zu vermeiden, muß das Wellenfeld den maximalen Offset L_{max} durchlaufen haben. Es gilt also mit $p_{max} = 1/v_{max}$ für das minimale Schußintervall Δt

$$\Delta t = L_{max} \cdot p_{max} \quad (3.2)$$

Mit den Gleichungen 3.1 und 3.2 folgt dann mit $\Delta x = \Delta t \cdot v$ (v Schiffsgeschwindigkeit)

$$p_{max} \leq \frac{1}{\sqrt{L_{max} f_{max} v}},$$

und definiert mitunter die benötigte Schiffsgeschwindigkeit. Anhand der zu erwartenden Eigenschaften der Scholtewelle und dem damit verbundenen größten Δx sind also die Meßparameter L_{max} bzw. Δt und v zu wählen. v ist hierbei zu kleinen Geschwindigkeiten begrenzt durch die zu erhaltende Manövrierbarkeit

3.1 Meßkonzept und Dimensionierung

des Meßschiffes. Es sind also Abschätzungen oder Testmessungen durchzuführen, wobei die Maximalfrequenz in flachmarinen Bereichen der Nord- und Ostsee zumeist unter 25 Hz zu erwarten ist (Bohlen et al., 2004) und verfestigte Sedimente zumeist Scholtewellengeschwindigkeiten über 200 m/s aufweisen. Bei einer maximalen Scholtewellenfrequenz von 25 Hz wäre somit ein Schußabstand von $\Delta x = 4 \text{ m}$ ausreichend. Eine vorausgehende Spektralanalyse mit dem verfahrensdefinierten minimalsten Schußabstand ist somit zu empfehlen, um weitere, größere Kampagnen zu dimensionieren.

Es bleibt, gegeben durch technische Umstände wie mangelnde Verfügbarkeit von Druckluft zum Betrieb des Luftpulsers, so daß ein ausreichendes Schußintervall nicht zu erfüllen ist, oder eine zu hohe Manövriergeschwindigkeit des Schiffes aufgrund von schlechten Wetterbedingungen, die Möglichkeit, daß eine optimale Dimensionierung der Messung nicht möglich ist. Für diesen Fall wurde im Rahmen dieser Arbeit ein Hilfsmittel implementiert, welches eine optimale Anpassung des Schußabstandes in jedem Fall möglich macht. Das Konzept basiert auf einer Änderung der Luftpulseransteuerung zu Gunsten der Position. So werden vorab ein optimaler Schußabstand festgelegt und entsprechende Schußpunktpositionen auf dem zu fahrenden Profil definiert. Es werden dann die mögliche Geschwindigkeit des Schiffes und das mögliche Schußintervall ermittelt. Basierend auf diesen Voraussetzungen werden multiple Überfahrten des Profils berechnet, deren Schußpunkte ineinander fassen und zusammen die vorab definierte Schußverteilung ergeben. Während der Überfahrten nutzt das verwendete Ansteuerungsprogramm die Positionsinformationen der Quelle zur Zündung derselbigen. Überfährt also die Quelle eine vorab definierte Position, so wird gefeuert. Der Schuß wird hierbei ausgelöst, überfährt die Quelle eine zum Profil senkrechte Gerade, die durch den definierten Schußpunkt verläuft. Auf diese Weise wird gewährleistet, daß auch bei leichten Abweichungen des Kurses vom Profil gefeuert wird (siehe Abbildung 3.4).

3 TECHNISCHE IMPLEMENTIERUNG SCHOLTEWELLENMESSUNG

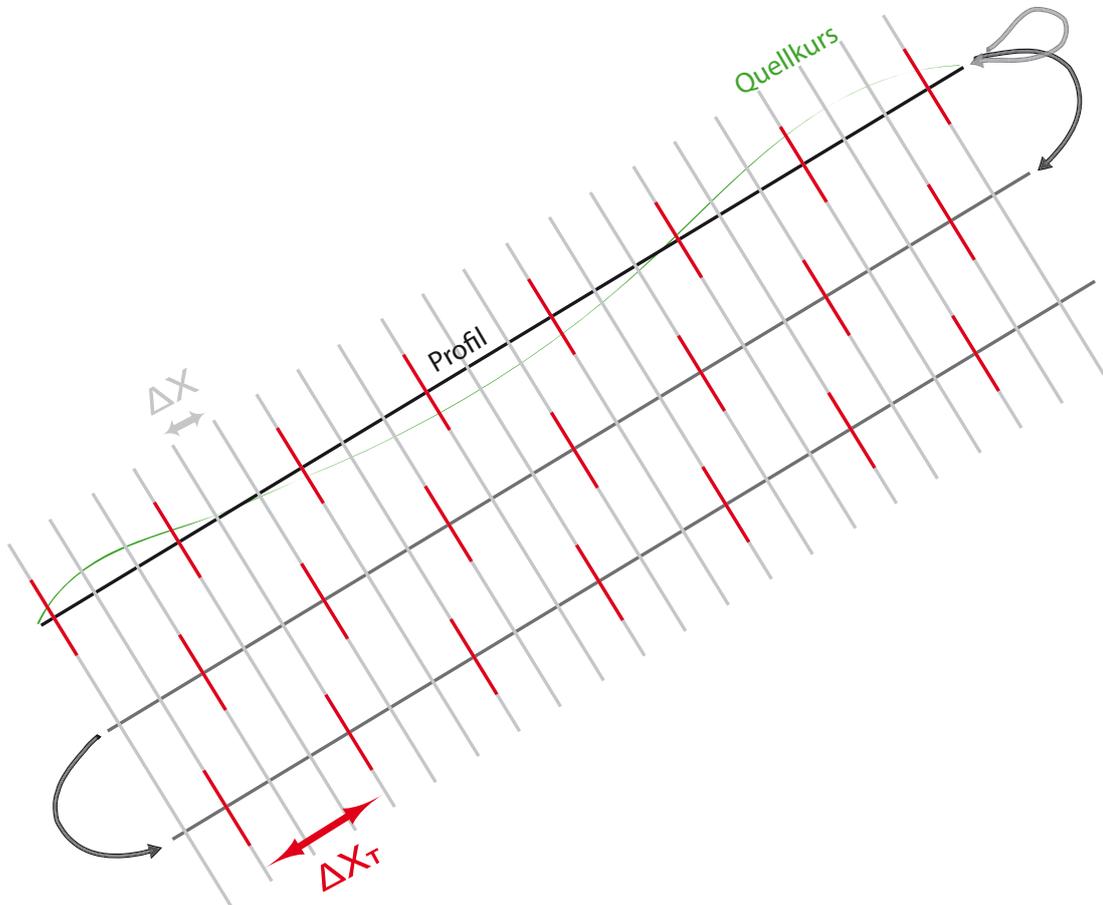


Abbildung 3.4: *Prinzipskizze des eingesetzten Triggerkonzeptes zum Auslösen der Quelle in Abhängigkeit von der Position. Die schwarze Linie zeigt das zu fahrende Profil mit den Soll-Schußpositionen (mit Abstand ΔX) (hellgraue Linien). Die roten Linien zeigen die Auslösepunkte auf den unterschiedlichen Überfahrten (dunkelgraue Linien). ΔX_T definiert den technisch erreichbaren Schußabstand.*

3.2 Entwicklung eines hochauflösenden Seebodengeophons

Im Rahmen der Ansätze zur Verbesserung der Auflösung der Scholtwellenanalyse wurde innerhalb des FINO3-Projektes und dieser Arbeit ein Hochfrequenz-Seebodengeophon (SBG) entwickelt. Motiviert ist diese Entwicklung durch die Anforderungen der Überwachungsmessungen an der Forschungsplattform FINO3. Diese Messungen erfordern eine auf wenige Meter genaue Relokalisierung der Sensoren zur örtlichen Korrelation der Wiederholungsmessungen. Das genutzte Konzept der Lokalisierung basiert auf der Verwendung einer hochfrequenten Boomerquelle, deren Signale von dem mit bis zu 11 kHz auflösenden SBG aufgezeichnet werden. Die Quelle wird hierbei von einem Schiff auf Kreisbahnen verschiedener Radien um die SBG geschleppt. Die Laufzeiten der Signale im Wasser können dann zur Relokalisierung des Gerätes am Meeresboden verwendet werden. Die Schallgeschwindigkeit der Kompressionswellen sei hierbei als konstant angenommen. Das SBG-System beinhaltet einen Embedded PC, verbunden mit einem Zweikanal-24Bit-A/D-Wandler und Vorverstärkern zur Akquisition von Signalen eines 4.5 Hz-Vertikalgeophons und eines omnidirektionalen Breitbandhydrophons. Das vorgestellte SBG-Konzept wurde im Rahmen von mehreren Meßfahrten an verschiedenen Lokationen in Nord- und Ostsee zu refraktionsseismischen Messungen und Scholtwellenmessungen erfolgreich eingesetzt.

Abbildung 3.5 zeigt eine Skizze des eingesetzten SBG. Dargestellt ist das Druckgehäuse aus PVC-U, welches aus einem mit einer quadratischen Grundplatte verklebten Druckrohr besteht. Das Rohr ist am oberen Ende mit einem Flansch versehen. Ein Deckel, versehen mit einem O-Ring, ist mit dem Flansch verschraubt und bietet so Zugang zum System. Im Deckel befindet sich eine druckdichte Durchführung zum Anschluß der Stromversorgung und Datenleitung des außen liegenden Hydrophons. Der Boden des Gehäuses wird mit zwei Betonplatten verschraubt, welche als Grundgewicht mit einem Gesamtgewicht von bis zu 60 kg dienen. Der zur Akquise verwendete Embedded PC vom Typ LE-363 der Firma Spectra trägt einen 400 MHz-AMD-Geode Prozessor, welcher die geringe Stromlast von ca. 1 A für das Gesamtsystem ermöglicht. Zur Speicherung der Daten ist eine 2.5 Zoll-Festplatte mit einer Speicherkapazität von 80 Gigabyte integriert. Die Festplatte hat sich auch unter extremen Belastungen als ausreichend lauffähig erwiesen. Die Signale des 4.5 Hz-Vertikalgeophons werden des weiteren analog vorverstärkt. Das außen am Gehäuse befestigte Hydrophon vom Modell AQ-17 der Firma Benthos enthält einen integrierten Vorverstärker. Die von den beiden Sensoren gelieferten Signale werden zunächst von einem 24Bit Zweikanal-A/D-Wandler aus dem Audio-Bereich (Spezifikationen siehe Tabelle 3.1) gewandelt und dann über eine optische Schnittstelle digital an die Soundkarte des PCs übertragen. Die Soundkarte puffert die eintreffenden Daten und eine

3 TECHNISCHE IMPLEMENTIERUNG SCHOLTEWELLENMESSUNG

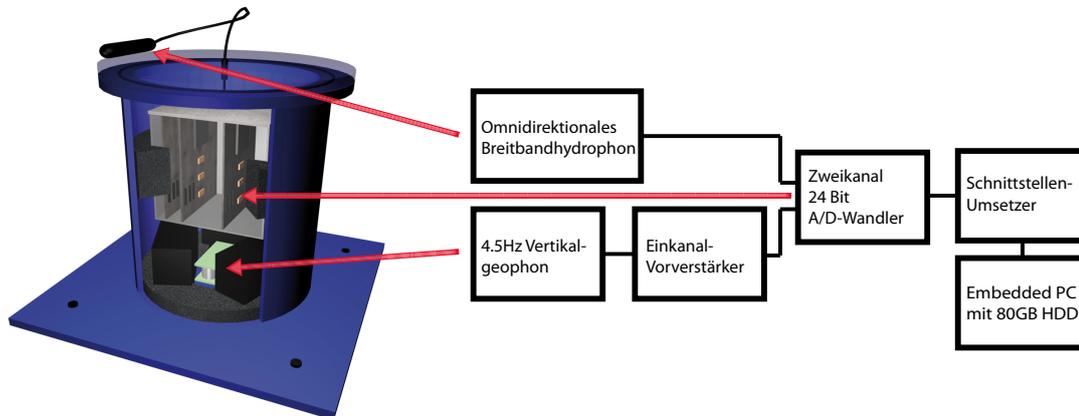


Abbildung 3.5: Skizze der eingesetzten Seebodensensorik.

Akquisitionsoftware sorgt lediglich für ein konstantes additives Speichern der Pufferinhalte, jeweils versehen mit einem Zeitstempel der Systemzeit des PCs.

Es wurde eine Akquisitionsoftware für das SBG geschrieben, welche neben der Speicherung der eintreffenden Datenpakete auch das simultane Speichern von Zeitsignalen ermöglicht. Hierzu wird das System vor der Messung über die serielle Schnittstelle des PCs mit einer GPS-Uhr verbunden. Diese sendet auf ein Steuersignal ein auf $\pm 2 \mu\text{s}$ genaues Zeitsignal, welches mit Hilfe der Akquisitionsoftware gemeinsam mit der entsprechenden PC-Zeit gespeichert wird. Diese Information wird später genutzt, um Datenbereiche mit GPS-Zeit verknüpften Schußzeiten zu korrelieren.

Das System liefert somit eine zweikanalige Datenspur mit einer der Akquisitionsdauer entsprechenden Länge. Das Format der Daten ist hierbei eine binäre Aneinanderreihung der Pufferblöcke der Soundkarte, jeweils versehen mit der Systemzeit des PCs. Die Spur wird mit Hilfe eines externen Programmes anhand der GPS-Schußzeiten der seismischen Quelle geschnitten. Hierbei werden automatisch GPS-Koordinaten der Schüsse eingesetzt. Das resultierende Datenformat eines so erhaltenen Common Receiver Gathers ist Seismic-Unix-Standard.

Zur Relokalisierung der Position des SBG am Meeresboden wird das System in Kreisbahnen von unterschiedlichem Radius (kleinster Radius ist hierbei abhängig von der Wassertiefe wählbar mit $R \gg z$) mit einer hochfrequenten seismischen Quelle ($> 500 \text{ Hz}$) umfahren. Besonders geeignet ist hierbei eine Boomerquelle mit einem Frequenzbereich von 500 Hz bis 6000 Hz. Die Wellenlänge der Signale beträgt so für eine Wassergeschwindigkeit nach Standardwerten von $v_{\text{Wasser}} = 1500 \text{ m/s}$ maximal 3 m. Jeder Quellschuß wird mit einem GPS-Zeitstempel versehen. Es können so die Laufzeiten der Signale im Wasser durch Anreißen des Ersteinsatzes in der Seismogrammsektion bestimmt werden. Laufzeit-Positions-

3.2 Entwicklung Seebodengeophon

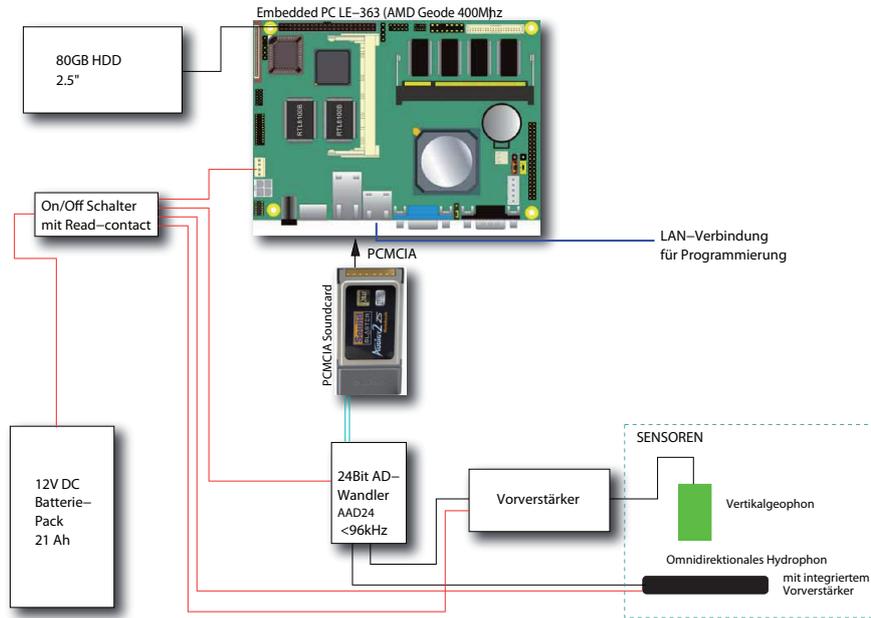


Abbildung 3.6: Konzeptskizze des technischen Interieurs des SBG.

Paare werden dann nach der Position des Empfängers invertiert. Hierzu wird die Laufzeitgleichung 3.2 mit Hilfe eines Referenzpaares linearisiert.

$$d_i = v_{Wasser} \cdot t_i = \|\vec{P} - \vec{P}_{Si}\| = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} \quad (3.3)$$

wobei $\vec{P} = [x, y, z]^T$ die Position des SBG und $\vec{P}_{Si} = [x_i, y_i, z_i]^T$ die Schußposition. Betrachtet man nun einen Referenzpunkt $i = R$,

$$d_R = v_{Wasser} \cdot t_R = \|\vec{P} - \vec{P}_{SR}\| = \sqrt{(x - x_R)^2 + (y - y_R)^2 + (z - z_R)^2} \quad (3.4)$$

so ergibt sich für $d_i^2 - d_R^2$ unter Kürzen der quadratischen Terme in x , y und z

$$\begin{pmatrix} T_1 \\ \dots \\ T_n \end{pmatrix} = \begin{pmatrix} a_{x1} & a_{y1} \\ \vdots & \vdots \\ a_{xn} & a_{yn} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_{x1} + b_{y1} \\ \dots \\ b_{xn} + b_{yn} \end{pmatrix} \quad (3.5)$$

3 TECHNISCHE IMPLEMENTIERUNG SCHOLTEWELLENMESSUNG

Gehäuse	
Material	PVC-U
Bodenplatte	500 mm x 500 mm x 15 mm
Druckrohr	∅ 320 mm, Höhe 370 mm
Flansch	Öffnung 260 mm, ∅ 380 mm
Deckel	∅ 380 mm, Nut 4 mm mit ∅ 324 mm
PC	
CPU	Embedded AMD Geode GX533 400 MHz
Speicher	Memory 1 x 200-pin DDRSO-DIMM 512 MB
Spannungsversorgung	DC 10.5 V 13 V
Stromversorgung	12 V/0.7 A 8.4 W
Maße	146 mm x 101 mm (L x B)
A/D-Wandler	
Typ	ELV AAD-24
Auflösung:	24 Bit pro Kanal
Abtastrate:	<48 kHz, 96 kHz, 192 kHz
Frequenzgang:	
48 kHz:	2 Hz – 23,5 kHz
96 kHz:	2 Hz – 47 kHz
192 kHz:	2 Hz – 67 kHz
Sensoren	
Hydrophon	Omnidirektionales Benthos AQ17
Geophon	Vertikal OYO 4.5 Hz
Zeitgeber	
GPS	Meinberg GPS167TGP

Tabelle 3.1: *Technische Daten SBG*

mit n Anzahl der Schüsse und $i = 1, \dots, n$ ($i \neq R$), $a_{xi} = 2x_R - 2x_i$, $a_{yi} = 2y_R - 2y_i$, $b_{xi} = x_i^2 - x_R^2$, $b_{yi} = y_i^2 - y_R^2$ und $T_i = v_{Wasser}^2(t_i - t_R)^2$. Die Wassertiefe z sei hierbei als konstant angenommen.

Man erhält somit ein überbestimmtes lineares Gleichungssystem, welches nach den Positionen des SBG über Singulärwertzerlegung gelöst wird. Die Linearisierung wird hierbei mit verschiedenen Referenzpaaren durchgeführt, so daß ein statistischer Fehler bestimmt werden kann. In Kapitel 10 ist für eine Scholtewellenmessung im Mai 2007 dieses Verfahren erfolgreich eingesetzt worden. Die Positionen der SBG ließen sich mit einer Standardabweichung von 4 m für verschiedene Referenzpositionen relokalisieren.

4 Allgemeine Inversionsmethodik

Die Analyse von Scholtewellen findet ihren Ursprung in der Adaption von Verfahren für Rayleighwellen in der Landseismik. Diese Verfahren basieren auf der Ausnutzung von dispersiven Eigenschaften der Oberflächenwellen bei vertikal heterogen verteilten seismischen Parametern. Die spektral trennbaren Propagationsgeschwindigkeiten lassen so Rückschlüsse auf vertikale Verteilungen der elastischen Eigenschaften des Mediums zu. Das in dieser Arbeit genutzte Konzept zur Inversion von Scherwellengeschwindigkeiten mit Hilfe der Spektralanalyse von Scholtewellen basiert auf dem von Muzyert (2000), Park et al. (2000), Ritzwoller & Levshin (2002), Bohlen et al. (2004) und Klein et al. (2005) für marine Anwendungen angepassten Konzept der Mehrkanalanalyse von Oberflächenwellen (Multichannel Analysis of Surface Waves, MASW). Diese Methode arbeitet mit einer Wellenfeldtransformation der mehrkanaligen Schußsektionen oder Empfängersektionen vom Weg-Zeit-Bereich in den Frequenz-Langsamkeitsbereich mittels der Local-Slant-Stack-Methode (McMechan and Yedlin, 1981). Diese Transformation resultiert in Phasenlangsamkeits-Frequenz-Spektren, aus denen die Dispersionskurven der Scholtewellen abgeleitet werden können. Dieser Spektralanalyse folgt eine iterative Anpassung vorwärts modellierter Spektren. Die Modellierung erfolgt hierbei nach der modifizierten Thomson-Haskell-Methode nach Wang (1999). Die Zielfunktion des Problems sei zunächst als Summe der Langsamkeitsresiduen von beobachteten und modellierten Dispersionskurven definiert (siehe Abbildung 4.1). Es sind gegebenenfalls mehrere Moden der Scholtewelle zu identifizieren. Die Kurven der einzelnen Moden der Scholtewellendispersion werden angerissen und den Amplitudenmaxima entsprechender Frequenz in den vorwärts gerechneten Spektren angepaßt.

Die laterale Variation der Scherwellengeschwindigkeiten kann ebenfalls nach dieser Methode untersucht werden. Es werden hierbei 1D-Scherwellengeschwindigkeitsprofile von lokalen Dispersionsspektren an verschiedenen Offsetpunkten invertiert. Die Empfängersektion (CRG) wird an unterschiedlichen Offsetpunkten mit einem Gaußschen Amplitudenfilter definierter Halbwertsbreite multipliziert und dann in den Frequenz-Langsamkeitsbereich transformiert. Die so transformierten lokalen Wellenfelder zeigen die lokalen Dispersionseigenschaften der Scholtewelle auf (Bohlen et al., 2004; Kugler et al., 2005).

4 ALLGEMEINE INVERSIONSMETHODIK

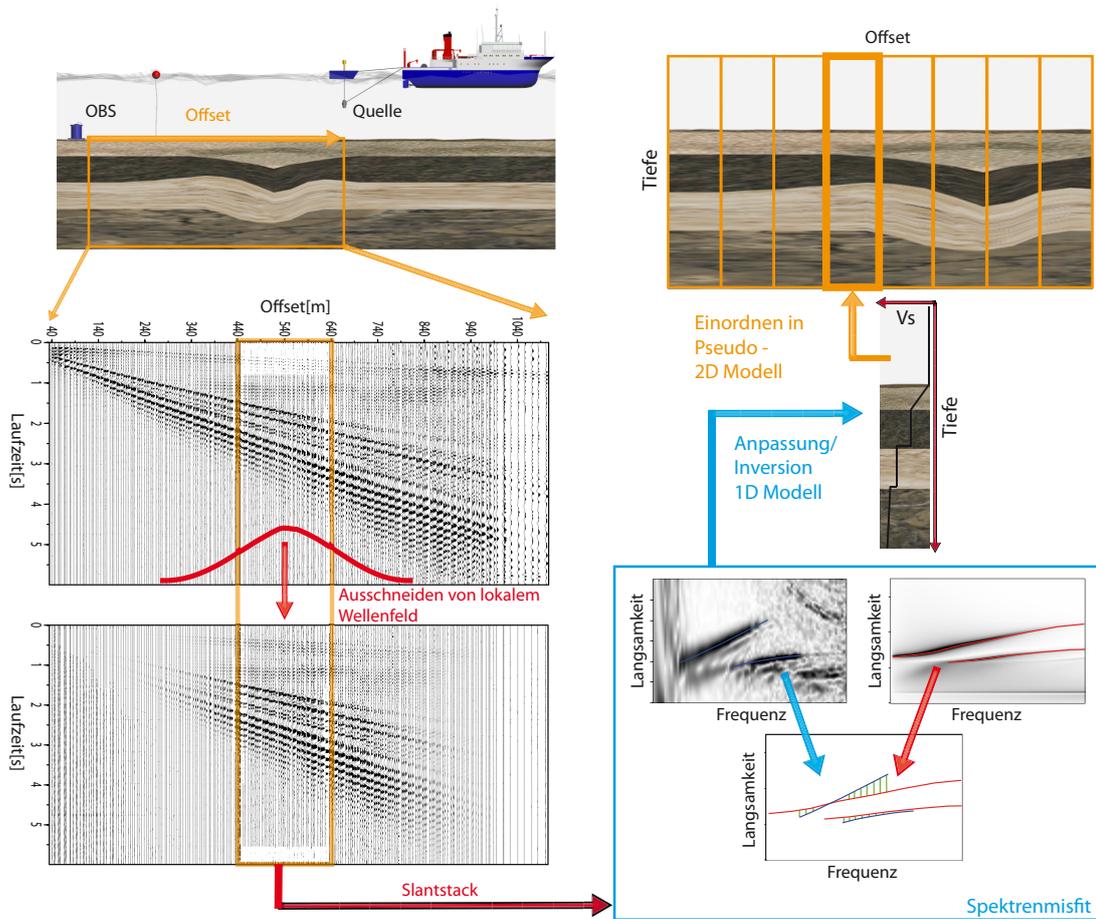


Abbildung 4.1: *Prinzip der Inversion für Pseudo-2D-Geschwindigkeitsmodelle.*

4.1 Wellenfeldtransformation

Transformationen des Wellenfeldes vom Weg-Zeit-Bereich in den Phasenlangsamkeits-Frequenzbereich sind beschrieben in McMechan and Yedlin (1981), Park et al. (1998) und Forbriger (2003a). Der Arbeit von McMechan and Yedlin (1981) folgend, wird das Wellenfeld (Beispiel in Abbildung 3.2) durch zwei lineare Transformationen in den Phasenlangsamkeitsfrequenzbereich (p , f) transformiert, um die Dispersionseigenschaften der Scholtewelle, also die Abhängigkeit der Phasengeschwindigkeit von der Frequenz erkennbar zu machen. Die erste Transformation ist ein „Slantstack“ des Wellenfeldes. Hierbei wird ein mit dem Offset linear ansteigender Zeitversatz mit definierter Steigung bzw. Geschwindigkeit auf die seismischen Spuren angewandt. Anschließend werden die Spuren gestapelt. Sind die Seismogramme hierbei in Phase und entsprechen damit der angesetzten

4.1 Wellenfeldtransformation

Geschwindigkeit, so werden sich die Amplituden konstruktiv addieren. Diese Stapelung wird für verschiedene Geschwindigkeiten durchgeführt. Diese Transformation ist analog einer *diskreten* Radon- bzw. τ - p -Transformation (siehe Gleichung 4.1).

$$\bar{U}(p, \tau) = \int_{-\infty}^{\infty} U(x, \tau - px) dx \quad (4.1)$$

Die zweite Transformation ist eine Fouriertransformation der erhaltenen Spuren. Das seismische Wellenfeld $U(x, t)$ wird so in den Frequenz-Wellenzahlbereich überführt.

Bohlen et al. (2004) beschrieben eine Frequenzbereich-Slantstackmethode zur Transformation lokaler Wellenfelder. Hierbei betrachte man ein Wellenfeld $U(x_n, t)$ für eine Sektion von $n = 1, \dots, N$ Schüssen an den Offsets x_n . Das lokale Wellenfeld $u(x_n, t)$ um den Offsetpunkt x_c wird erhalten, indem ein Gaußscher Amplitudendämpfungsoperator mit dem Wellenfeld multipliziert wird (siehe Abbildung 4.1 und Gleichung 4.1),

$$u(x_n, t) = U(x_n, t) e^{-\left(\frac{x_n - x_c}{H/2}\right)^2} \quad (4.2)$$

wobei $H/2$ den Abstand der Halbwertsbreite definiert. Das lokale Wellenfeld wird dann in den Frequenzbereich überführt,

$$\bar{u}(x_n, f) = \int_{-\infty}^{\infty} u(x_n, t) e^{i2\pi ft} dt \quad (4.3)$$

worauf der Slantstack im Frequenzbereich mit einer offsetabhängigen Phasenverschiebung $-2\pi fp \cdot x_n$ folgt.

$$\tilde{u}(p, f) = \sum_{n=1}^N \bar{u}(x_n, f) e^{-i2\pi fp \cdot x_n} \quad (4.4)$$

4 ALLGEMEINE INVERSIONSMETHODIK

Kugler (2001) hat ein modelliertes Wellenfeld mit nur einem Einsatz bei einer bestimmten Langsamkeit \tilde{p} mittels der Slantstackmethode spektral analysiert. Hierbei war zu beachten, daß das zu erwartende diskrete spektrale Maximum bei \tilde{p} nicht auftrat. Vielmehr wurde eine frequenzabhängige Verbreiterung des Maximums und das Auftreten von Nebenmaxima beobachtet. Das Ausbleiben eines diskreten Einsatzes bei \tilde{p} läßt sich jedoch durch die Transformation einer endlichen Anzahl Spuren erklären. Nur bei infiniten Spurenzahl verschwindet das Integral der Radontransformation für Langsamkeiten ungleich \tilde{p} . Die Verbreiterung des Hauptmaximums ist somit festgelegt durch,

$$\Delta p \approx \frac{1}{\Delta L \cdot f} \quad (4.5)$$

wobei ΔL der betrachtete Offsetbereich ist.

Abbildung 4.2 zeigt das Beispiel eines Phasenlangsamkeits-Frequenz-Spektrums, erhalten mit der Methode des lokalen Slantstacks. Das Spektrum gehört zu einem lokalen Wellenfeld mit einer Breite von 400 m, aufgezeichnet in der Nordsee an der Lokation FINO3 (siehe Abbildung 10.1 und Abschnitt 10.1). Zu höheren Frequenzen außerhalb des Frequenzbereiches der Scholtewelle (blaue, gestrichelte Linie) tritt Aliasing auf (grüne, gestrichelte Linie). Hier wurde das räumliche Abtastkriterium (Abschnitt 3.1) verletzt. Auch eine Verbreiterung bzw. Unschärfe der Moden der Scholtewelle zu niedrigen Frequenzen ist zu erkennen. Das Spektrum zeigt einige Komplexität mit mehreren enthaltenen Moden der Scholtewelle, welche nicht in jedem Bereich klar trennbar oder zuzuordnen sind.

4.2 Anpassung der Wellenfeldspektren

Die Anpassung der Spektren der Scholtewellen an modellierte Spektren zur Inversion des vorliegenden Scherwellengeschwindigkeitsmodells geschieht analog Bohlen et al. (2004) durch eine Anpassung von extrahierten (gepickten) Dispersionskurven für die verschiedenen Moden der Scholtewelle. Die Zielfunktion des hier verwendeten Vergleichs berechnet sich also durch Summation der Langsamkeitsresiduen zwischen gemessenen und modellierten Dispersionskurven jeder Frequenz zu jeder Mode der Scholtewelle. Abbildung 4.3 zeigt das Prinzip dieser Anpassung. Die aus der Slantstack-Transformation erhaltenen Spektren werden zunächst manuell angerissen, das heißt die Dispersionskurven der Moden werden gepickt (grüne Kurven). Die Modellparameter des Problems lassen sich in einem Vektor \vec{X} zusammenfassen, welcher die Scherwellengeschwindigkeit in verschiedenen Tiefen $\vec{v}_s(z)$, die Kompressionswellengeschwindigkeit $\vec{v}_p(z)$ sowie die Dichte $\vec{\rho}(z)$ und die Tiefe der Schichtpakete \vec{z} enthält.

4.2 Anpassung der Wellenfeldspektren

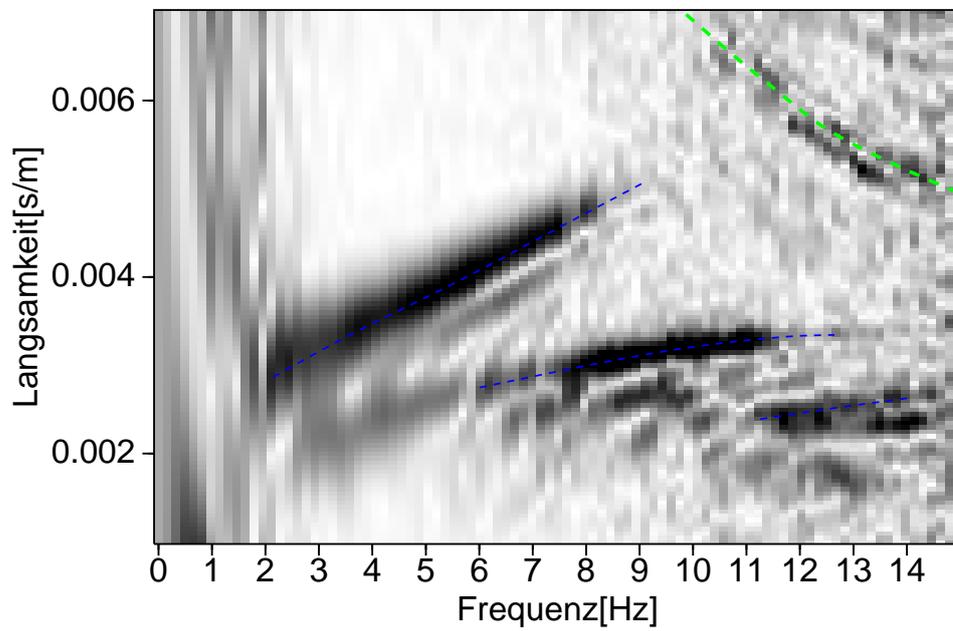


Abbildung 4.2: *Beispiel eines Phasenlangsamkeits-Frequenz-Spektrums. Die blauen Linien reißen enthaltene Scholtewellenmoden an. Die grüne, gestrichelte Linie zeigt Aliasing-Energie auf, welche jedoch außerhalb des interessanten Frequenzbereiches liegt.*

4 ALLGEMEINE INVERSIONSMETHODIK

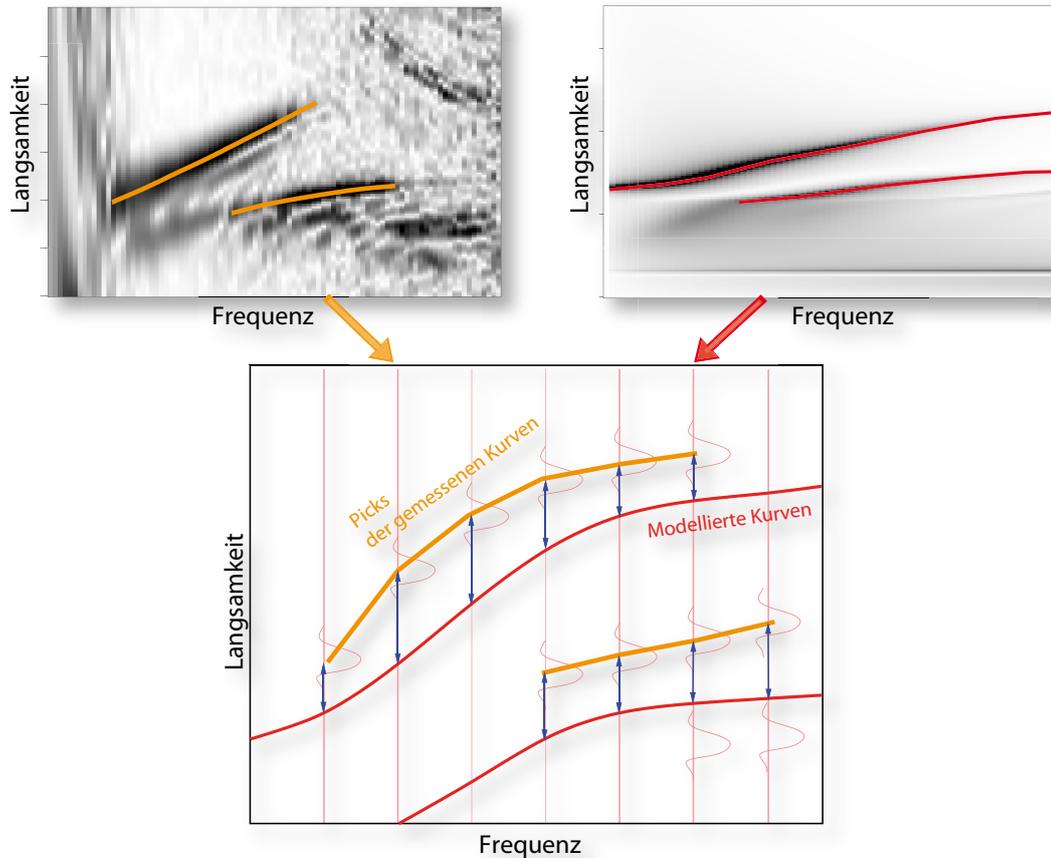


Abbildung 4.3: Prinzip zur Bestimmung des Anpassungsmaßes von modellierten (rot) und gemessenen (grün) Dispersionskurven.

$$\vec{X} = (\vec{v}_s(z), \vec{v}_p(z), \vec{\rho}(z), \vec{z})^T \quad (4.6)$$

Mit diesen Parametern wird ein Spektrum modelliert, in welchem die Amplitudenmaxima der Moden der Scholtewelle automatisch detektiert werden (rote Kurven). Die Residuen der Langsamkeiten (blaue Pfeile) für jede enthaltene Frequenz werden dann summiert und als Maß der Anpassung betrachtet.

Unter der Voraussetzung, daß die Thomson-Haskell-Variante von Wang (1999) eine Matrix als Amplitudenspektrum im Phasenlangsamkeits-Frequenzbereich liefert (siehe Abschnitt 2.1), und unter der Annahme, daß dieses Spektrum nur die Fundamentalmode der Scholtewelle enthält, ließe sich die exakte Berechnung der Zielfunktion wie folgt angeben,

4.2 Anpassung der Wellenfeldspektren

$$E = \sum_{f=f_{min}}^{f_{max}} \|p(f)^{obs} - \max(PF(p, f))\| \quad (4.7)$$

wobei $\max(PF(p, f))$ das Amplitudenmaximum der Fundamentalmode innerhalb einer Spalte mit Frequenz f des Spektrums $PF(p, f)$ anzeigt und $p(f)^{obs}$ die gemessene Dispersionskurve ist. Die Definition $\max(PF(p, f))$ ist hierbei kein funktionaler Zusammenhang und läßt sich nicht differenzieren, so daß eine Auswertung des Gradienten $\nabla E(\vec{v}_s(z), \vec{v}_p(z), \vec{\rho}(z), \vec{z})$ nach den Modellgrößen $\vec{v}_s(z), \vec{v}_p(z), \vec{\rho}(z), \vec{z}$ der Zielfunktion möglich wäre. Unter der Voraussetzung, daß ein Thomson-Haskell-ähnlicher Algorithmus zur Vorwärtsmodellierung verwendet wird, welcher nicht direkt die Dispersionskurven in Abhängigkeit der Modellparameter liefert, ist also ein numerisches Berechnen des Gradienten der Zielfunktion notwendig.

Da die Abhängigkeit der modellierten Phasenlangsamkeiten und damit der Dispersionskurven von \vec{X} ein nichtlineares Problem darstellt, wurde in den Arbeiten von Bohlen (2004), Kugler (2005) und (2006) das Optimierungsverfahren der Sequentiellen Quadratischen Programmierung (SQP) zur Lösung des Problems angewendet. Das SQP-Verfahren ist ein Verfahren zur Lösung restrukturierter, nichtlinearer Optimierungsprobleme (siehe bspw. Boggs & Tolle, 1995). Es basiert auf dem Broyden-Fletcher-Goldfarb-Shanno (BFGS) Quasi-Newton-Sekantenverfahren, welches simultan zur iterativen Anpassung eines aktuellen Lösungsvektors x_k (Iterationsschritt k) des Problems $\min f(x)$ ($x \in D \subset \mathbb{R}^n$) ein Analogon zur Hessematrix, H_k , der Funktion f nach einer definierten Aufdatierungsvorschrift iterativ anpaßt (Kosmol, 1993). Nebenbedingungen $g_i(x)$ werden unter Gewichtung durch Lagrange Faktoren (λ) in eine Lagrangefunktion L integriert, die es nun zu optimieren gilt.

$$L(x, \lambda) = f(x) + \sum \lambda_i g_i(x)$$

Die Aufdatierung des Lösungsvektors x erfolgt nach dem BFGS-Verfahren über

$$x_{k+1} = x_k + \alpha_k d_k = x_k + \alpha_k \cdot \underbrace{H_k \nabla f_k}_{\text{gradientenorientierte Folge}}$$

wobei d_k die Abstiegsrichtung definiert. α_k ist die Schrittweite, welche i. A. durch eine Schrittweitenregel (Kosmol, 1993) angepaßt werden muß.

Der SQP-Ansatz definiert nun ein an f angepaßtes quadratisches Problem zur Bestimmung der optimalen Abstiegsrichtung d_k . Hierzu wird das Unterproblem

$$P(d) = \frac{1}{2} d^T H_k d + \nabla f(x_k)^T d$$

unter der Bedingung $\nabla g_i(x) + g_i(x) \geq 0$ gelöst.

Auflösung

Das genutzte lokale SQP-Optimierungsverfahren bedarf eines Startvektors, mit welchem über Parametervariationen eine Auflösungsanalyse durchgeführt werden kann. Es gilt prinzipiell zu prüfen, wie gut die Modellparameter durch die Daten definiert sind und wie sich interparameträre Zusammenhänge (Trade-Off) auf die Lösung auswirken. Hierzu schlugen Bohlen et al. (2004) ein Verfahren nach Forbriger (2003b) vor, bei dem zunächst ein einzelner Modellparameter des Startvektors \vec{X}_0 einer Inversion gewählt und modifiziert wird. Die Modifikation des Parameters entspricht hierbei einer Erhöhung der Zielfunktion um 10%. Es wird eine Inversion mit dem variierten Startvektor durchgeführt, in welcher der modifizierte Parameter konstant gehalten wird. Diese Vorgehensweise wird dann auf alle anderen Parameter angewandt. Es kann so die Variation der einzelnen Parameter zum Referenzergebnismodell der ursprünglichen Inversion untersucht werden. Sind die Zusammenhänge unter den Parametern (Trade-Off) groß, so verhindert bspw. ein festgehaltener Parameter die Anpassung eines anderen und es kommt zu großen Abweichungen.

Misfitdefinition und Fehler

Die einfachste Methode, um inverse Probleme zu bemessen, basiert auf der Messung einer Größe p bzw. Länge abhängig von den angepassten Modellparametern \vec{m}^{est} und den damit vorhergesagten Daten $\vec{p}^{pred} = g(\vec{m}^{est})$, wobei die Funktion g der Vorwärtsmodellierung entspricht. Für jede Beobachtung i betrachte man nun den Vorhersagefehler oder „Misfit“ $e_i = p_i^{obs} - p_i^{pred}$ von vorhergesagten Daten p_i^{pred} und beobachteten Daten p_i^{obs} . Der beste Satz von m_i^{est} bietet somit den kleinsten Gesamtfehler (Methode der kleinsten Quadrate) $E = \sum_{i=1}^n e_i^2$, welcher der quadrierten, euklidischen Länge des Vektors \vec{e} entspricht. Diese Länge, also der euklidische Abstand von den Beobachtungen wird den vorhergesagten Daten zugeordnet.

Die euklidische Norm ist hierbei nur eine Möglichkeit des Abstandsmaßes. Norm bezeichnet nur das Maß einer Länge. Die gängigsten Normen gehören zur Familie der L_n -Normen.

$$L_1 : \|\vec{e}\|_1 = \left[\sum_i |e_i| \right]$$

$$L_2 : \|\vec{e}\|_2 = \left[\sum_i |e_i|^2 \right]^{1/2}$$

4.2 Anpassung der Wellenfeldspektren

...

$$L_n : \|\vec{e}\|_n = \left[\sum_i |e_i|^n \right]^{1/n}$$

Normen höherer Ordnung n verleihen hierbei großen Elementen e_i größeres relatives Gewicht. Die gängige Methode der kleinsten Quadrate nutzt die L_2 -Norm als Längenmaß. Diese Wahl ist bestimmt durch die Frage, wie Daten gewichtet werden, die als Ausreißer gelten.

Allgemein gilt, wenn die Abweichungen der Daten zum generellen Trend klein sind, ist ein Ausreißer j mit großem Vorhersagefehler e_j stark zu gewichten. Normen höherer Ordnung sind also geboten. Sind die Daten jedoch um den Trend stark gestreut, so kann keine Priorität mehr auf wenige einzelne große Vorhersagefehler fallen. Eine Norm geringerer Ordnung wäre also angemessen. Betrachtet man die Fehler der Daten als Gaußverteilt, so ist im allgemeinen die L_2 -Norm zu wählen, da die Gaußverteilung eher steile Flanken aufweist. Verteilungen mit flacheren Flanken induzieren eine breitere Streuung, denn die Wahrscheinlichkeit für einen Meßwert, abseits des Trends zu liegen, hängt von der Form der Verteilung ab. Die Verteilungen für einen einzelnen Meßwert bei wiederholter Messung definiert also die Streuung der Abweichungen um den Gesamttrend und diktiert die Wahl der Norm. (Menke, 1989)

Wie bereits in Abbildung 4.3 dargestellt, ist die Zielfunktion E des Problems der Anpassung der Wellenfeldspektren von Scholtewellen durch eine Summation der Langsamkeitsresiduen zwischen gemessenen und modellierten Dispersionskurven jeder Frequenz der Scholtewelle definiert.

$$E = \sum_{f=f_{min}}^{f_{max}} \left\| p_f^{obs} - p_f^{pred} \right\|$$

Mit der Existenz von Rauschen in den Daten sollte die Zielfunktion einen bestimmten Residualwert nicht unterschreiten, um nicht Rauscheffekte anzupassen. Es sei nun angenommen, daß Rauschfehler in den Daten Gaußverteilt mit der Standardabweichung σ_i und einem Mittelwert von 0 sind. Diese Annahme ist zu treffen, da jeder Meßpunkt der Langsamkeit p bei einer Frequenz f (p_f^{obs}) im Phasenlangsamkeits-Frequenz-Spektrum nur einmal registriert wurde, da mehrfache Messungen zu aufwendig wären. Die Annahme ermöglicht nun eine Gewichtung der eingehenden Datenpunkte mit der Abweichung,

$$E = \sum_{f=f_{min}}^{f_{max}} \frac{\left\| p_f^{obs} - p_f^{pred} \right\|}{\sigma_f}$$

4 ALLGEMEINE INVERSIONSMETHODIK

Dies ergibt die resultierende Definition der Zielfunktion mit $\sigma_f = \frac{1}{\Delta L \cdot f}$ nach Gleichung 4.5 und 4.2:

$$E = \sum_{f=f_{min}}^{f_{max}} \frac{\|p(f)^{obs} - p(\vec{X}, f)\|}{\sigma_f} \quad (4.8)$$

mit $\vec{X} = (\vec{v}_s(z), \vec{v}_p(z), \vec{\rho}(z), \vec{z})^T$. Eine automatische Erkennung von Amplitudenmaxima in den Phasenlangsamkeits-Frequenz-Spektren, auch beim Auftreten mehrerer Schwingungsmoden der Scholtewelle, wird zur Wahl der Norm eine entscheidende Rolle spielen. Es sei daher an diesem Punkt auf den Abschnitt 6.2 hingewiesen, welcher die Wahl der Norm im Hinblick auf erwartete Fehler des Qualitätsmaßes diskutiert.

Parametrisierung

Da Scholtewellen insbesondere in Analogie zur Rayleighwelle hohe Sensitivität gegenüber dem Scherwellengeschwindigkeitstiefenmodell haben (Ewing et al., 1957), untersuchten Bohlen et al. (2004) den Einfluß unterschiedlicher linearer Relationen für die Abhängigkeit der Dichte und der Kompressionswellengeschwindigkeit von der Scherwellengeschwindigkeit auf ihre Inversionsergebnisse. Da die Sensitivität der Scholtewelle gegenüber ρ und v_p als gering anzunehmen ist, ist eine Auflösung dieser Größen durch die Inversion nicht zu erwarten. Des weiteren vergrößert die Inversion von Dichte und Kompressionswellengeschwindigkeit den zu untersuchenden Parameterraum immens, so daß ρ und v_p entweder mittels a priori Informationen abgeschätzt und konstant gehalten werden oder durch empirische Relationen in Abhängigkeit von den aktuell angepaßten Scherwellengeschwindigkeiten gewählt werden. Bohlen et al. (2004) prüften die Inversion auf Sensitivität für verschiedene Relationen und fanden, daß das Ergebnismodell nahezu unabhängig von der gewählten Relation ist. Auch bei fest gewählter Dichte und Kompressionswellengeschwindigkeit variierte das Ergebnismodell nur leicht, was die geringe Sensitivität der Scholtewelle gegenüber diesen Größen im betrachteten flachmarinen Fall bestätigte. Kugler et al. (2007) zeigten allerdings, daß bei hoher Datendichte, also beim Auftreten vieler Moden der Scholtewelle, eine Inversion grober Auflösung in der Dichte erhalten werden kann. Die Sensitivität für das Schermodul ist somit dominierend und Änderungen der Dichte und v_p sind im allgemeinen als klein dem gegenüber zu sehen. Bei Dispersionsanalysen im tieffrequenteren Fall und größeren Tiefen der Modelle ist jedoch zu erwarten, daß Dichte und v_p einen signifikanteren Einfluß haben. Die zu invertierenden Größen beschränken sich so für die bislang untersuchten flachmarinen Lokationen auf die Scherwellengeschwindigkeiten und Schichtmächtigkeiten.

$$\vec{X} = (\vec{v}_s, \vec{h})^T$$

4.2 Anpassung der Wellenfeldspektren

Zur Wahl der Schichtanzahl läßt sich in einfachster Form ein Vergleich von Inversionsergebnissen in Abhängigkeit von der Schichtzahl durchführen, um eine Anpassung der optimalen Schichtzahl zu erreichen. Hierbei wird zunächst mit geringer Schichtanzahl invertiert und dann die Anzahl gesteigert. So läßt sich sehen, ob mit erhöhter oder verringerter Schichtzahl die Anpassung verbessert werden kann oder ob sich eine Sättigung der Anpassung mit steigender Schichtzahl einstellt. Zu beachten sei hierbei jedoch, daß mit steigender Schichtzahl und damit steigender Dimensionalität des Problems auch die Anzahl der Mehrdeutigkeiten bzw. Nebenminima der Qualitätsfunktion steigen kann, was ebenfalls die Auflösung verschlechtern kann. Ein Auftragen des Inversionsergebnisses in ein Achsensystem aus Parameterzahl und Schichtzahl läßt so gegebenenfalls eine Analyse zu, die einen Kompromiß aus Parameterzahl und Schichtzahl zum Ergebnis haben kann. Das Auffinden der minimalen Anzahl von Parametern sowie Grenzen der Parameter, die zur Erklärung des Problems ausreichend sind, ist somit eine Hauptaufgabe bei der Dimensionierung der Inversion, um die Unterbestimmtheit des Problems zu verringern. Zur weiteren Regularisierung können noch zusätzliche Kriterien bzw. a priori Informationen bezüglich der Lösung eingeführt werden, um Mehrdeutigkeiten zu umgehen. Im vorliegenden flachmarinen Fall sind oftmals reflexionsseismische Daten vorhanden, deren Horizonte zur Dimensionierung des Parameterraums genutzt werden können und die Anzahl der Schichten des Sedimentmodells bereits festlegen. Des weiteren kann in den meisten Fällen einfacher Sedimentation von einem positiven Geschwindigkeitsgradienten in den Modellen ausgegangen werden.

Die gewählte Parametrisierung der Modelle ist zunächst durch die Vorgaben der Vorwärtsmodellierung definiert. Der Algorithmus von Wang (1999) bedient sich einer einfachen linearen Tiefenparametrisierung, bei der jede Schicht unterhalb des Meeresbodens durch einen linearen Gradienten zwischen einem Oberkanten- und Unterkanten-Wert innerhalb einer Schicht bestimmter Mächtigkeit parametrisiert wird (siehe Abbildung 4.4).

Eine Unterteilung in diskrete Schichtpakete konstanter Mächtigkeit ist aufgrund der dadurch erhöhten Parameterzahl nur unter Einbindung weiterer Nebenbedingungen wie Modellglattheit zu empfehlen. Beaty et al. (2002) untersuchten ergänzend das Verhalten einer auf einem Metropolis-Algorithmus basierenden Rayleighwellen-Inversion in Abhängigkeit von der Wahl der Parametrisierung. Die Schichten wurden zum einen in viele gleichmächtige, diskrete, homogene Schichten unterteilt und zum anderen wurde das Scherwellengeschwindigkeitstiefenmodell durch Chebyshev-Polynome fünfter und siebter Ordnung parametrisiert (skizziert als graue und grüne Parametrisierung in Abb. 4.4). Die Inversion basierend auf polynomialer Parametrisierung zeigte sich stabiler und schneller, jedoch war im Vergleich zur Variante mit freier Parameterwahl in diskreten Schichten die Anpassung schlechter. Die durch die Polynome induzierten Abhängigkeiten zeigten nicht die nötige Variabilität für bspw. Bereiche starken Gradienten. Die Ap-

4 ALLGEMEINE INVERSIONSMETHODIK

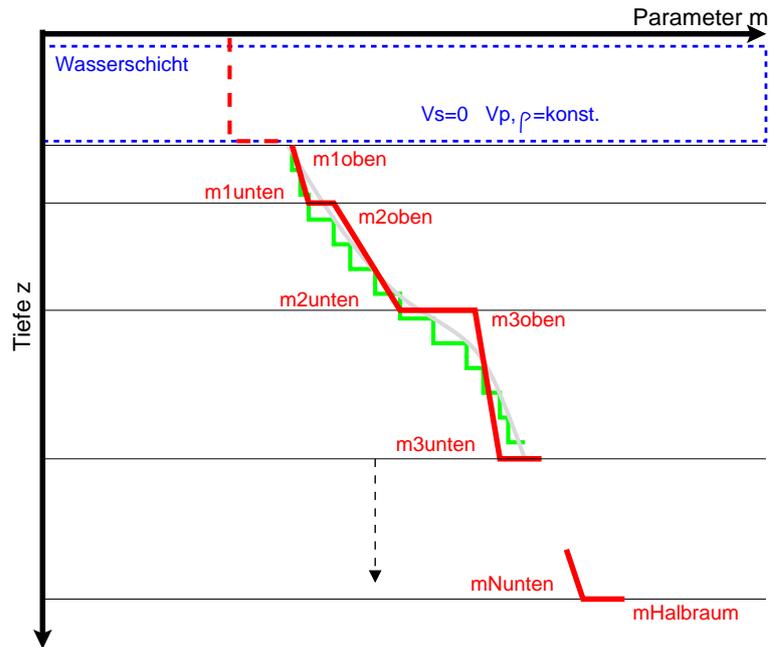


Abbildung 4.4: Skizze zur Parametrisierung der Modellparameter m . Die in dieser Arbeit verwendete lineare Gradientenparametrisierung ist rot dargestellt. Das graue und grüne Modell skizzieren alternative Parametrisierungen mittels Polynomapproximation (grau) oder diskreter Schichtung (grün).

proximation durch lineare Gradienten mit frei wählbaren Parametern erscheint dahingegen in den Arbeiten von Bohlen et al. (2004), Kugler et al. (2005) & (2006) als hinreichend.

Teil II

5 Motivation: Baugrundüberwachung und die Anforderungen an die Inversion

Eine grundlegende Frage bei der Überwachung des Baugrundes eines Offshorebauwerkes wie der FINO3-Plattform ist die nach der Amplitude der Schermoduländerung bei langsamer Entfestigung des Sedimentes. Bei vollständiger Verflüssigung würde der Schermodul verschwinden. Das heißt, sollten die Überwachungen zeigen, daß mit der Standzeit des Bauwerks der Schermodul sinkt, so ist die Wahrscheinlichkeit einer Gefährdung als zunehmend zu erwarten. Die Zeitskala, mit welcher solche Veränderungen fortschreiten, und die Amplitude der Änderungen sind jedoch unbekannt. Eine hohe Auflösung des Gradienten der Schergeschwindigkeitsresultate einer Überwachungskampagne ist daher von großem Interesse. Die von Bohlen et al. (2004) und Kugler et al. (2005) angegebene Auflösung einer Scholtewelleninversion liegt bei 10% bis 14% der Scherwellengeschwindigkeit in Schichten bis 40 m.

Als Studie zur Baugrundvoruntersuchung und Überwachung an der FINO3 Plattform wurde hierzu ein einfaches numerisches Experiment in Form von elastischen und viskoelastischen 2D-Finite Differenzen Modellierungen (Bohlen et al., 2001) des Wellenfeldes in der Umgebung der Plattform durchgeführt (Wilken et al., 2007), um eine Abschätzung der Laufzeitunterschiede von Scholtewellen und Kompressionswellen durch eine künstliche Kompaktierungszone an FINO3 zu erhalten.

Die genutzten Modelle beinhalten vereinfachte Geometrien einer künstlichen Verdichtungszone um den Monopile der FINO3 in den oberen 20 m der Sedimente. Die physikalischen Rahmenbedingungen für Scherwellengeschwindigkeiten, Dichte und Kompressionswellengeschwindigkeiten wurden aus den ersten Ergebnissen einer Baugrundvoruntersuchung (Wilken et al., 2009) und den Befunden aus einer Bohrung im Rahmen dieser Voruntersuchung abgeschätzt. Die vereinfachten Modelle enthalten keinen Gründungspfahl und das Wellenfeld wird am Rand der Kompaktierungszone, der sich an den Pfahl anschließt, absorbiert. Die laterale Ausdehnung der Kompaktierungszone wurde für mehrere Modellierungen mit 15 m, 20 m, 25 m, und 50 m definiert und nimmt invers mit der Tiefe ab. Die Wahl der Parameter des Geschwindigkeits- und Dichtegradienten in der Kompaktierungszone beruht auf Modellierungen von Wiemann & Lesny (2004), welche die

5 MOTIVATION EINES GLOBALEN SUCHMUSTERS

Veränderung des Korngerüsts des Sedimentes um einen Monopile, resultierend aus einer äußeren Horizontalbelastung des Pfahls, untersucht haben. Die Beispiele dieser Studie ergaben eine Abschätzung der Veränderung der Porosität, also des Anteils Porenraum im Gesteinsvolumen, $\Delta\Theta$, innerhalb der Kompaktierungszone. Für das Beispiel FINO3 kann so eine Änderung der Porosität von $\Delta\Theta = 0.08$ angenommen werden. Aus dieser Änderung wird eine theoretische Dichteänderung berechnet sowie die Gesamtänderung der Scher- und Kompressionswellengeschwindigkeit mittels empirischer Zusammenhänge für gesättigte Sande (Prasad & Meissner, 1992; Schön, 1983) abgeschätzt. Ausgehend von der so erhaltenen Größenordnung von $\Delta v_s^{erw} = 40 \text{ m/s}$ als Änderung in der Scherwellengeschwindigkeit läßt sich die benötigte Auflösung der zu invertierenden Modelle bemessen, welche benötigt wird, um Änderungen im Rahmen von Überwachungsmessungen zu quantifizieren.

Als Motivation eines globalen Suchmusters sollte die Multimodalität des Inversionsproblems, also das Maß für die Komplexität der Zielfunktion durch multiple Minima abgeschätzt werden. Existiert eine eindeutige Lösung, so sei ein Problem als unimodal bezeichnet. Bei mehr als nur einem Optimum ist das Problem multimodal (siehe Abbildung 5.1). Dazu sei ein einfaches Modellbeispiel betrachtet.

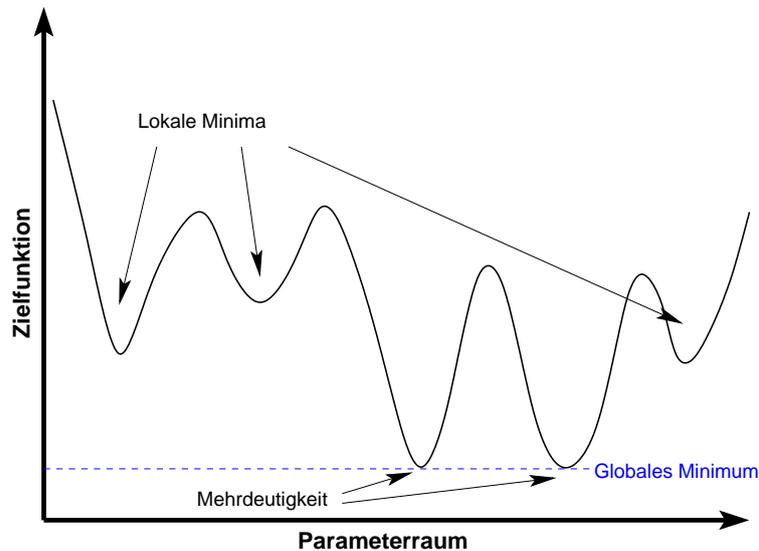


Abbildung 5.1: Skizze zur Erläuterung der Multimodalität einer Zielfunktion (schwarze Kurve). Lokale Nebenminima sind durch Pfeile markiert. Die Zielfunktion hat ein uneindeutiges, globales Minimum.

Ausgehend von einem definierten Scherwellengeschwindigkeitstiefenmodell wird ein Phasenlangsamspektrum modelliert. Es werden des weiteren Modelle mit variierender Scherwellengeschwindigkeit der am Meeresboden anschließen-

den Schicht und des Halbraumes (v_{s1} und v_{sn}) gerechnet. Das ungestörte Modell sei nun als Referenzmodell angenommen, so daß das Anpassungsmaß für dieses Modell und die variierten Modelle bestimmt werden kann. Dies wird in einer Beispielstudie für 30000 gleichverteilt, zufällig gewählte Variationen der Paare v_{s1}, v_{sn} und zwei Referenzmodelle durchgeführt. Tabelle 5.1 zeigt das verwendete Referenzmodell (v_s mit der Tiefe), die genutzten Intervalle für v_{s1}, v_{sn} sowie das zweite Referenzmodell mit verringerter Geschwindigkeit der obersten Schicht, \tilde{v}_s .

Tiefe[m]	v_s [m/s]	v_s -Intervall[m/s]	\tilde{v}_s [m/s]	\tilde{v}_s -Intervall[m/s]
0.0	0.0	0.0	0.0	0.0
23.0	0.0	0.0	0.0	0.0
23.0	135.0	120.0-150	115.0	100.0-130
27.0	174.0	0.0	174.0	0.0
27.0	225.0	0.0	225.0	0.0
33.0	279.0	0.0	279.0	0.0
33.0	320.0	0.0	320.0	0.0
56.0	373.0	0.0	373.0	0.0
56.0	422.0	370.0-500.0	422.0	370.0-500.0

Tabelle 5.1: *Referenzmodell und Variationsintervalle zur Untersuchung der Multimodalität des Inversionsproblems.*

Bei einer Veränderung der Geschwindigkeiten der obersten Schicht durch Kompaktion ließe sich vermuten, daß die Minima nah eines globalen Optimums nahezu äquidistant translatiert werden, die Landschaft der Zielfunktion also nur verschoben wird. Durch die Wahl der gleichen Startpopulation oder die Wahl des zuletzt invertierten Modells als Startmodell könnten dann immer die gleichen Minima erreicht werden und damit Variationen der Scherwellengeschwindigkeit im Bereich Δv_s^{est} als relative Änderungen auflösbar werden. Um dies zu untersuchen, ist die Zielfunktion in Abhängigkeit von v_{s1} und v_{sn} für zwei verschiedene Referenzmodelle $v_s(z)$ und $\tilde{v}_s(z)$ in Abbildung 5.2 dargestellt. Es gilt hierbei nach Tabelle 5.1 $v_{s1} = 135 \text{ m/s}$ und $\tilde{v}_{s1} = 115 \text{ m/s}$.

Das berechnete Anpassungsmaß zu beiden Referenzmodellen ist nun in Abbildung 5.2 aufgetragen. Die Funktion zeigt einen hohen Grad an Multimodalität, insbesondere im Nahbereich des globalen Minimums. Betrachtet man die unteren beiden Graphen in Abbildung 5.2, so zeigt die relative Lage der Minima keine vergleichbare Struktur für die Referenzmodelle $v_s(z)$ und $\tilde{v}_s(z)$. Die obige Vermutung einer relativen Verschiebung kann also schon in diesem vereinfachten, zweidimensionalen Beispiel nicht bestätigt werden. Es sei nun zu beachten, daß sich für höher dimensionale Fälle das Problem weitaus multimodaler verhalten wird. Die Lösungen werden mitunter nicht eindeutig sein (siehe auch Abbildung 5.1,

5 MOTIVATION EINES GLOBALEN SUCHMUSTERS

verschiedene Parameterkonstellationen führen zu gleichen Werten der Zielfunktion). Um die oben beschriebene benötigte Auflösung also überhaupt im Sinne der Optimierung erreichen zu können, ist es zu empfehlen, Optimierungsverfahren zu verwenden, die ein hohes Maß an Parameterraumexploration garantieren und global suchen, d.h. den gesamten Parameterraum explorieren, aber gleichzeitig in Minima konvergieren. Gradientenverfahren, also Verfahren die ausgehend von einem Startpunkt die Lösung in Richtung des stärksten Abstiegs suchen, bieten diese Eigenschaften nicht. Die in Frage kommenden Verfahren gehören zur Klasse der stochastischen Optimierungsverfahren und werden im Kapitel 7 beschrieben. Diese Verfahren gehen von einer zufällig gewählten Startpopulation oder einem zufällig gewählten Startpunkt aus und wählen dann neue Punkte durch Selektion, basierend auf den Werten der Zielfunktion. Durch den Erhalt einer Zufallskomponente bei der Wahl eines neuen Punktes wird die Exploration des Parameterraumes erhalten.

Zunächst jedoch sollen neben der Multimodalität weitere Gründe für ein globales, stochastisches Suchmuster benannt werden:

- *Black box Optimierung.* Die meisten Inversionsprobleme, so auch das hier behandelte, beinhalten eine Modellierung anzupassender Datensätze, die meist die numerische Berechnung eines nichtlinearen Problems zu Erzeugung künstlicher Datensätze beinhaltet. Die Verfahren bieten daher oft keine Möglichkeit, die Ableitung der Zielfunktion analytisch zu bestimmen. Ein gradientenorientiertes Verfahren wäre somit auf die Approximation des Gradienten durch Differenzenquotienten angewiesen.
- Differenzenquotienten können numerisch instabil sein. Die Güte der Approximation des Gradienten definiert sich über das Infinitesimal des Differenzschrittes. Wird dieses jedoch zu klein gewählt, so ist oft die Zielfunktion nicht sensitiv für derart kleine Parameteränderungen und die Approximation des Gradienten wird instabil.
- Die Zielfunktion birgt oft Mehrdeutigkeiten durch Rauschen. Hierbei können Rauschartefakte bei der Inversion angepaßt werden.
- Die Zielfunktion kann Diskontinuitäten enthalten. Diese können zum Beispiel durch das Einführen von Straftermen für Nebenbedingungen entstehen. Die Konvergenz der Gradientenverfahren läßt sich jedoch nur für wenigstens einfach differenzierbare Funktionen zeigen (Kosmol, 1993).

Für den betrachteten Fall liegen alle genannten Probleme vor. Die Anwendung eines globalen, gradientenfreien Suchmusters ist also für die Untersuchung von Uneindeutigkeiten oder von Nebenminima im Nahbereich der Lösung und zum Umgehen gradienteninduzierter Instabilitäten von fundamentaler Bedeutung.

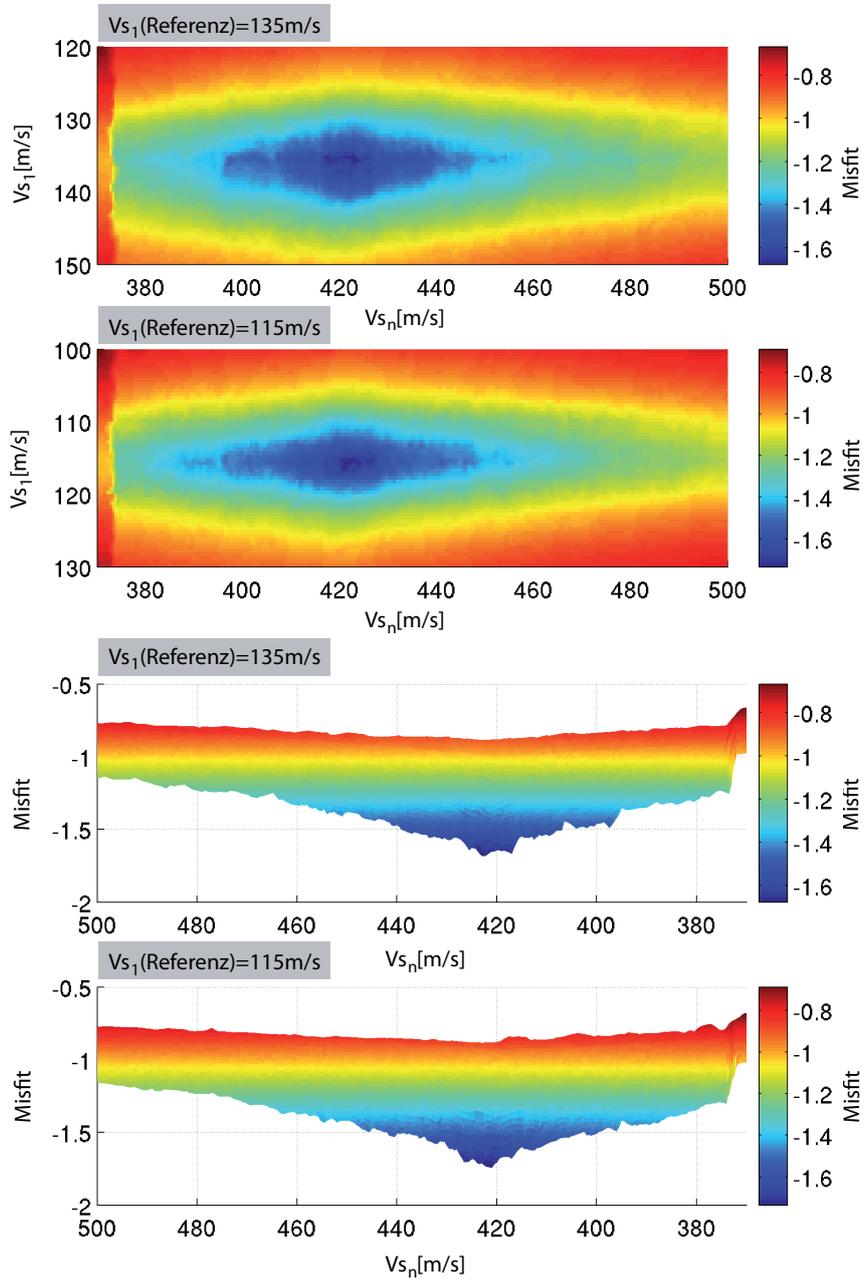


Abbildung 5.2: Zielfunktion oder Anpassungsmaß für Modelle mit variierten v_{s1} und v_{sn} zu zwei Referenzmodellen $v_s(z)$ und $\tilde{v}_s(z)$.

5 MOTIVATION EINES GLOBALEN SUCHMUSTERS

6 Anforderungen an ein globales Suchmuster

Globale, stochastische Optimierungsverfahren erfordern Zufallszahlen zur Erzeugung von Startvektoren oder Populationen von Startvektoren. Auch werden oftmals während der Iterationen Parameter durch Zufallsvariationen verändert. Die Auswahl eines getesteten und adäquaten Pseudo-Zufallszahlengenerators ist somit von essentieller Bedeutung. Auch die Verteilung der Parameter bedarf bestimmter Voraussetzungen. So sind Startpopulation und Variationen der Parameter als gleichverteilt zu wählen. Nebenbedingungen an die Startvektoren und Zusammenhänge zwischen den Parametern verändern aber die Struktur ihrer Verteilung, so daß die Methode zur Anwendung solcher Nebenbedingungen ebenfalls mit Bedacht zu wählen ist.

Des weiteren können zufällig generierte Geschwindigkeitsmodelle Spektren sehr hoher Komplexität erzeugen, in denen eine Zuordnung der Dispersionsmoden zum Vergleich der Phasenlangsamkeiten nicht mehr automatisiert werden kann. Es wird deshalb ein neues, zuordnungsfreies Vergleichsverfahren zur Berechnung des Anpassungsmaßes vorgeschlagen.

6.1 Zufallszahlen

Stochastische Optimierungsverfahren benötigen Zufallszahlen, um die Startpopulation, den Startpunkt oder iterative Parametervariationen behandeln zu können. Die Kriterien an die Zufallszahlen wären die Unabhängigkeit der Zahlen und eine möglichst homogene Verfüllung des gegebenen Parametervolumens. Des weiteren ist es zur Fehleranalyse von großem Nutzen, wenn die Folge der Zufallszahlen reproduzierbar ist. Eine Rechenmaschine kann jedoch keine echten Zufallszahlen erzeugen. Es wurde daher in dieser Arbeit auf einen Algorithmus von Marsaglia et al. (1990) zurückgegriffen. Zur Initialisierung eines Zufallszahlengenerators verwendet man Startwerte x_0, x_1, \dots, x_{k-1} , die als Saat bezeichnet werden. Sie können fest vorgegeben werden, wenn die erzeugte Zahlenfolge wiederholbar sein soll, oder sie werden zufällig initialisiert. Das Verfahren von Marsaglia et al. (1990) baut auf dem Verzögerten-Fibonacci-Generator auf. Beim Fibonacci-Zufallszahlengenerator werden aus der Saat mit der rekursiven Folge der Fibonacci-Zahlen der Vorschrift $x_{i+2} = x_{i+1} + x_i$ weitere Zahlen erzeugt: $x_{i+1} = x_i + x_{i-1} \bmod(m)$, dabei bezeichnet \bmod den Divisionsrest mit dem Modul

m. Der Verzögerte-Fibonacci-Generator ist eine Verallgemeinerung, bei dem nicht die beiden letzten, sondern weiter zurückliegende Zustandswerte zur Erzeugung der neuen Zufallszahl verwendet werden.

6.2 Anpassung der Spektren

Automatisches Extrahieren der gemessenen Dispersionskurven

Die Lokalisierung der Amplitudenmaxima in gemessenen Spektren entlang der Dispersionskurven wurde im Rahmen dieser Arbeit automatisiert. Ziel des Verfahrens ist eine Phasenlangsamkeits-Frequenz-Matrix, welche nur an den Amplitudenmaxima der Dispersionkurven von Null verschieden ist. Hierzu wurde ein Verfahren entwickelt, welches kohärente Amplitudensektionen in Spektren oder auch Seismogrammsektionen extrahiert. Das Konzept des Verfahrens soll der menschlichen Verfahrensweise, kohärente Einsätze in Bildern zu erkennen, möglichst nahe kommen. Hierbei wird zunächst angenommen, daß der menschliche Betrachter zum einen Bereiche ähnlicher Amplitude und zum anderen Bereiche starker, spaltenübergreifender Signalkohärenz und Glattheit als zusammengehörig identifiziert. Hierbei sollte möglichst keine Grenzbedingung, wie eine Amplitudenschwelle verwendet werden, da diese meist nicht adaptiv wählbar ist.

Als Beispiel sei nun ein Phasenlangsamkeits-Frequenz-Spektrum betrachtet (siehe Abbildung 6.1 a)). Die Spalten konstanter Frequenz des Spektrums entsprechen hierbei einer Spur und die Phasenlangsamkeiten seien Samples der Spur. Das Verfahren geht nun wie folgt vor:

1. Alle Punkte des Spektrums werden durchlaufen und es werden zunächst nur die Amplitudenwerte erhalten, die innerhalb eines Intervalls von n_s Samples ein lokales Amplitudenmaximum sind (Abbildung 6.1 a)) (n_s hängt hierbei von der Gesamtzahl der Samples ab und kann mit der Signalbreite abgeschätzt werden). Alle anderen Punkte werden auf 0 gesetzt (Abbildung 6.1 b)).
2. Alle in 1. gewählten lokalen N Amplitudenmaxima werden in einer Schleife ausgewählt. Um jeden Punkt N_i wird nun ein Fenster gelegt, dessen Breite w (in Spuren) und Höhe h (in Samples) vordefiniert sei. In diesem Fenster soll nun der Pfad höchster Kohärenz entlang der enthaltenen Amplitudenmaxima (P_j in Abbildung 6.1 b2)) gewählt werden. Die Wahl der Höhe und Breite des Fensters hängt somit von der maximalen Steigung der zu detektierenden kohärenten Einsätze ab. Je stärker ein Segment geneigt ist, desto höher muß ein Fenster bei konstanter Breite gewählt werden. Die Größe der zusammenhängenden Segmente im untersuchten Bild definiert ebenfalls die Breite des Fenster.

6.2 Anpassung der Spektren

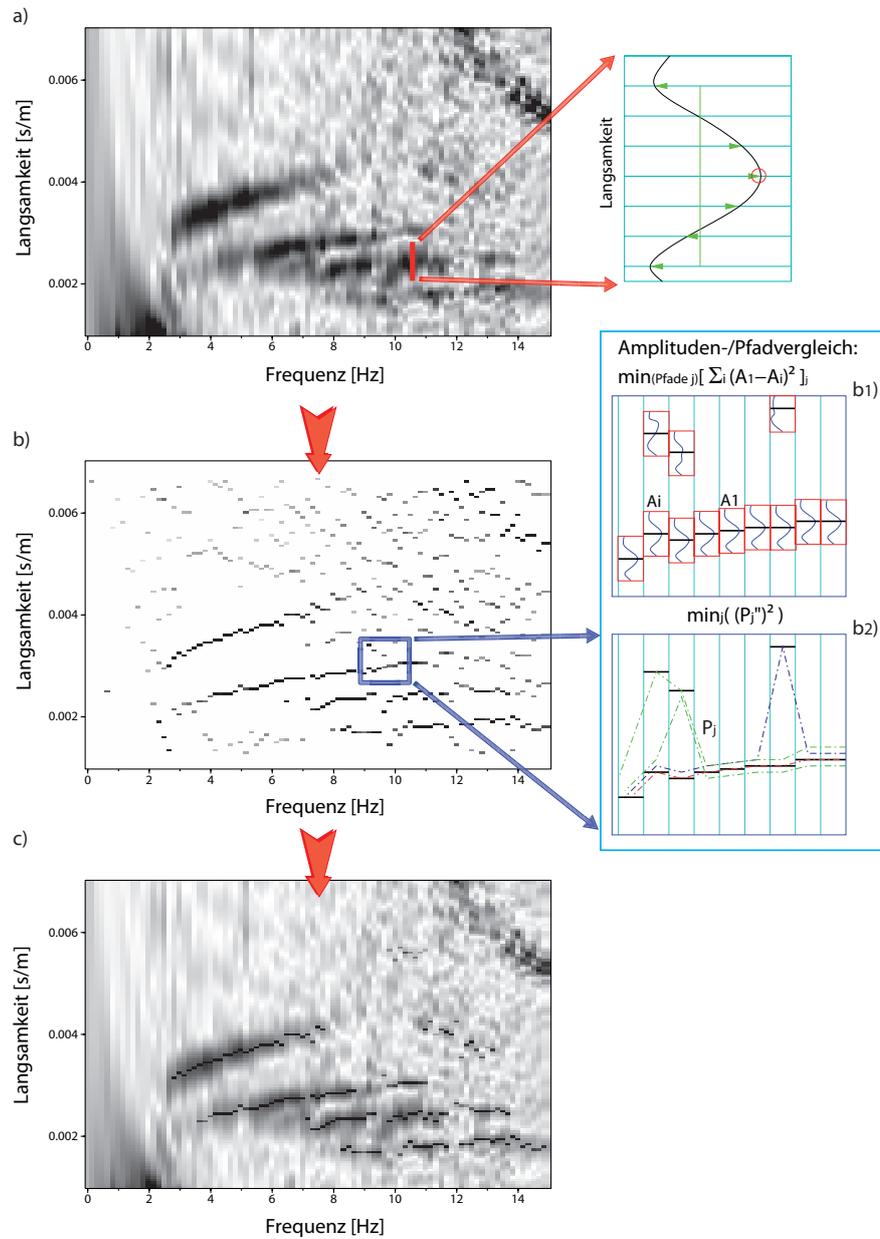


Abbildung 6.1: Konzept des automatischen Extrahierens der Amplitudenmaxima von Dispersionskurven, dargestellt anhand eines Beispielspektrums aus der Nordseelokation FINO3 (siehe Karte 10.1).

Je kleiner zusammenhängende Bereiche im Bild sind, desto kleiner muß auch die Breite gewählt sein. Die Wahl dieser Werte ist daher mit Hilfe einiger Testdurchläufe zu optimieren.

Es sind nun in jedem Fenster, das auf einem gesuchten kohärenten Segment

6 ANFORDERUNGEN AN EIN GLOBALES SUCHMUSTER

liegt, mehrere Maximalamplitudenpunkte aus 1. enthalten. Nun werden alle möglichen **vollständigen** Pfade, mit denen diese Punkte über die Spuren hinweg verbunden werden können, bestimmt (Abbildung 6.1 b2)). Die Anzahl der erhaltenen Pfade D berechnet sich mit $D = \prod_{i=1}^w M_i$, wobei M_i die Anzahl Amplitudenmaxima in Spur i ist.

3. In jedem Fenster werden nun für jeden Pfad zwei Werte berechnet, die als Maß der Kohärenz eines Pfades dienen. Es werde zunächst entlang jedes Pfades die zweite Ableitung approximiert und ihr Quadrat als Maß für die Glattheit des Pfades (Whittaker & Robinson, 1946) aufsummiert. Die Approximation der zweiten Ableitung geschieht hierbei analog einem Teil des Verfahrens zur Cubic Spline Interpolation nach Press et al. (2007). Angenommen sei eine Splinefunktion als Approximation des Pfades (x_i, y_i) . Aus der Voraussetzung für Splinefunktionen, daß die erste Ableitung stetig sei, auch zwischen den Intervallen (x_{j-1}, x_j) (x_j, x_{j+1}) , folgt ein Gleichungssystem für die zweite Ableitung des Splines. Dieses wird hier durch das Verfahren der konjugierten Gradienten (Kosmol, 1993) numerisch gelöst. Des weiteren wird die mittlere Amplitudenabweichung entlang eines Pfades berechnet. Hierzu wird nicht nur das Amplitudenmaximum, sondern ein definiertes Intervall der Ausgangsspur, um die Maxima genutzt. Die Abweichungen der Amplitudenbereiche zum Intervall um den Ausgangspunkt werden summiert (Abbildung 6.1 b1)).
4. Ist nun der Pfad der minimalen Ableitungssumme und der Pfad der minimalsten Amplitudenabweichung der selbe, so wird das entsprechende Pfadsegment als vermeintlich kohärentes Segment gewählt und gespeichert.

Mit diesem Verfahren werden kohärente Segmente extrahiert und hinterher zusammengefügt (siehe auch Datenbeispiel in Abbildung 6.1 c)). Es bleiben in der Ergebnismatrix nur die Punkte verschieden von Null, die in einem Segment enthalten sind. Ausreißer und fehlerhafte Segmente können dann nachträglich manuell entfernt werden. Dieses Verfahren ermöglicht einen einfachen Zugang zu Eingangsdaten für die Inversion, ohne den zusätzlichen Einfluß menschlicher Fehler durch manuelles Anreißen der Dispersionskurven.

Anpassung modellierter Dispersionskurven ohne Modenzuordnung

Im vorliegenden Fall der Scholtewelleninversion, werden Parameterkonstellationen für $v_s(z)$ -Modelle zufällig generiert, um Startpunkt oder Startpopulation für ein stochastisches Optimierungsverfahren zu bestimmen. Die Modelle werden in definierten Bereichen des Parameterraumes, abhängig von den Randbedingungen des Problems, erzeugt. Durch die zufällige Wahl der Parameter eines Modells

6.2 Anpassung der Spektren

kann es zu Konstellationen von Parametern kommen, welche modellierte Spektren mit extremen Eigenschaften erzeugen. Diese Spektren enthalten oftmals Oskulationen (Berührungspunkte der Moden), oder ein Ausbleiben von Energie in bestimmten Frequenzbereichen (siehe Beispiel 6.2). Ein einfaches Vorgehen zur Erkennung der Dispersionskurven und Zuordnung der Moden ist in den Spektren so nicht mehr möglich. Es muß daher eine neue Automatisierung der Erkennung der Dispersionskurven und gegebenenfalls eine neue Definition des Anpassungsmaßes überlegt werden. Zur Lösung dieses Problems wird ein Ansatz verfolgt, der

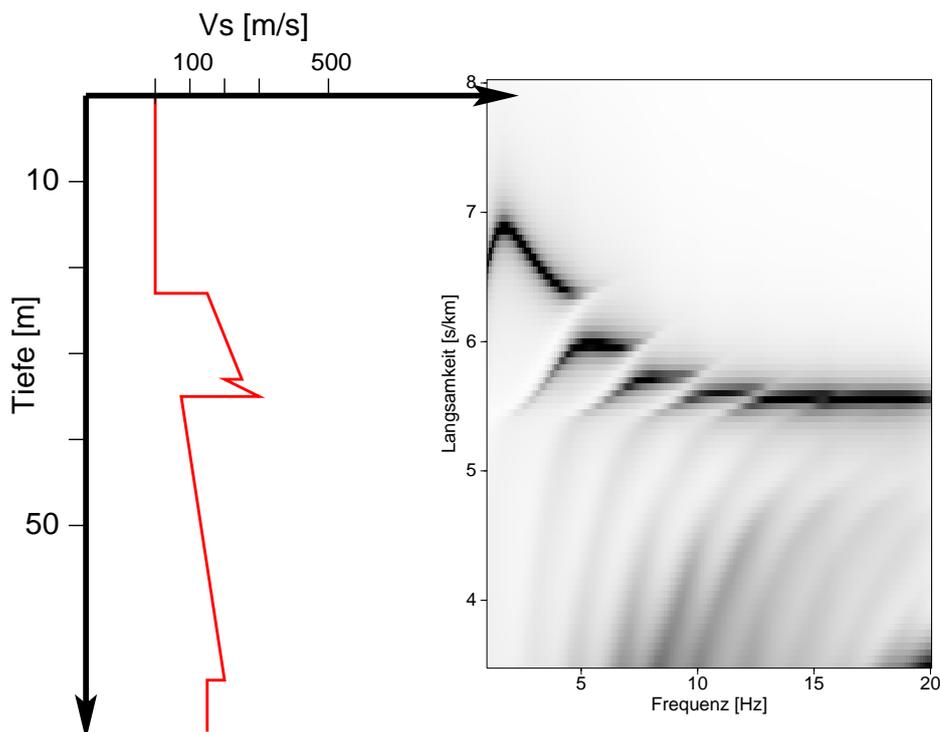


Abbildung 6.2: Beispiel für ein modelliertes Phasenlangsamkeits-Frequenz-Spektrum hoher Komplexität. Die linke Seite zeigt das zugehörige Schergeschwindigkeitsmodell.

die Zuordnung zu den Moden außer Acht läßt. So wird ein modelliertes Spektrum zunächst nach lokalen Maxima durchsucht. Es bleiben nur die Amplitudenwerte erhalten, die innerhalb eines Intervalls von n Samples das lokale Amplitudenmaximum sind (siehe auch voriger Abschnitt). Dies erzeugt ein skelettiertes Spektrum, welches nur die Amplitudenmaxima der Moden enthält. Zur Berechnung der Zielfunktion/Anpassungsmaßes wird nun zu jedem Punkt der gemessenen Dispersionskurve das nächstgelegene Amplitudenmaximum bei selbiger Frequenz im modellierten Vergleichsspektrum gesucht (siehe Abbildung 6.3). Die Abstände werden für jeden gemessenen Phasenlangsamkeits-Frequenz-Punkt bestimmt und summiert. Die Zuordnung des nächst gelegenen Maximums kann zu Beginn

6 ANFORDERUNGEN AN EIN GLOBALES SUCHMUSTER

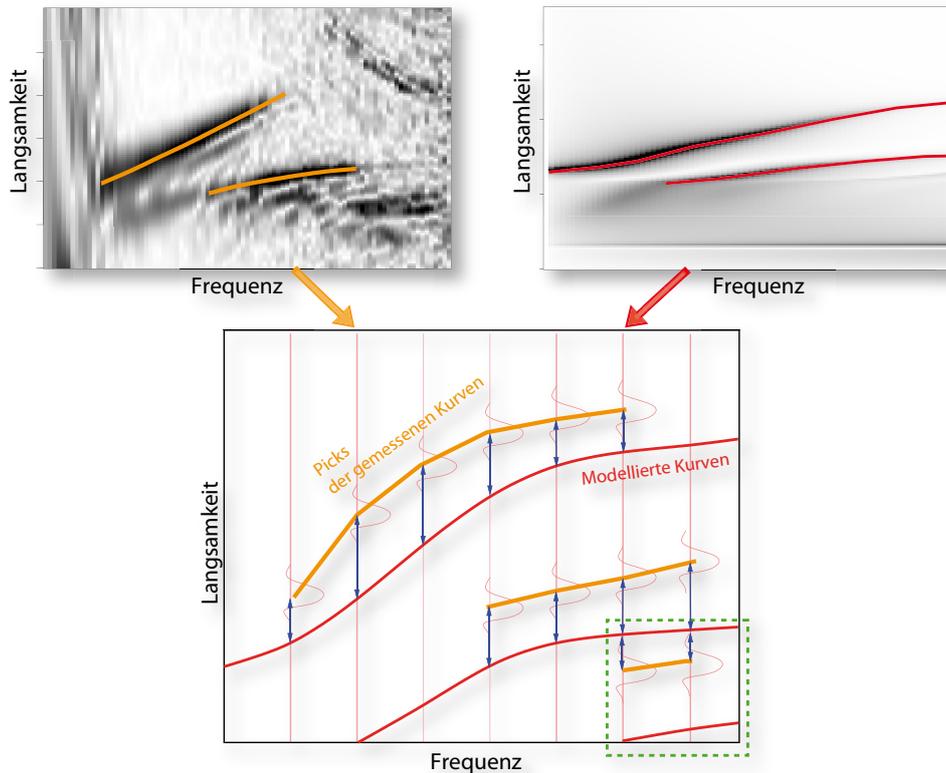


Abbildung 6.3: Konzept des Anpassungsmaßes. Jedem Punkt der gemessenen Dispersionskurven (grüne Kurven) wird bei konstanter Frequenz das nächstgelegene Amplitudenmaximum des modellierten Spektrums (rote Kurven) zugeordnet. Die ermittelten Abstände der Phasenlangsamkeiten werden summiert. Hierbei kann es in anfänglichen Iteration zu Zuordnungsfehlern kommen (grüner Kasten); in diesem Beispiel sind Punkte der zweiten höheren Mode der ersten zugeordnet.

der Inversion, also bei großen Abständen der Moden, durch falsche Zuordnung zu Fehlern führen; so können Punkte zu falschen, unterschiedliche Moden zugeordnet werden (siehe grüner Kasten in Abbildung 6.3). Nähert sich jedoch das angepaßte Spektrum dem gemessenen, so minimiert sich auch der Zuordnungsfehler. Es gilt daher zu prüfen, ob diese Methode der modenzuordnungsfreien Zielfunktionsbestimmung zu adäquaten Auflösungen gereicht und ob anfängliche Zuordnungsfehler die Tendenz der Inversion beeinflussen. Des weiteren gilt zu prüfen, welche Norm für das Anpassungsmaß zu wählen ist.

Die Verteilungen für einen einzelnen Meßwert bei wiederholter Messung definieren die Streuung der Abweichungen um den Gesamttrend (siehe Abschnitt 4.2). Abweichungen in den Vorhersagefehlern sind daher im allgemeinen definiert durch die Statistik der einzelnen Meßpunkte, da die Modellierung keine Fehler beinhaltet.

tet.

Die hier beschriebene Methode der Punktzuordnung kann jedoch zu Ausreißern in den Vorhersagefehlern für die einzelnen Meßpunkte führen. Da ein Meßpunkt nur allgemein dem im vorwärts modellierten Spektrum nächsten Amplitudenmaximum zugeordnet wird, kann es zu falschen Zuordnungen zu anderen Moden der modellierten Scholtewelle kommen, deren Maxima zufällig näher am Pick liegen. Die Wahrscheinlichkeit solcher Zuordnungsfehler sollte mit fortschreitender Anpassung der Dispersionskurven geringer werden. Da dieser Fehler jedoch keinen physikalischen Hintergrund hat, dürfen trotz schwacher Variation in den Datenpunkten aufgrund von gaußverteilten Fehlern, jene Ausreißer nicht stark gewertet werden. In diesem Fall ist die L_1 -Norm als bessere und robustere Wahl vorgeschlagen. Es bleibt aber offen, wie echte Ausreißer in den Punkten entlang der gemessenen Dispersion durch die Wahl der L_1 -Norm bewertet werden. Es kann jedoch davon ausgegangen werden, daß gravierende Ausreißer durch die Auswahl des messenden Wissenschaftlers im Vorgang des Anreißens der Dispersionskurven nicht gewertet werden. Eine Auflösungsanalyse mit definiert verrauschten Daten im Vergleich von L_1 - und L_2 -Norm findet sich in Abschnitt 8.3.

6.3 Verteilung der Startpopulation

Die Startpopulation der Modellparameter muß im definierten Parameterraum gleichverteilt zufällig gewählt sein. Dies gilt auch für Parametervektoren, die im Laufe der Optimierung den durch Randbedingungen begrenzten Raum verlassen. An die Modelle Nebenbedingungen zu stellen, die neben den Definitionsintervallen gültig seien, beinhaltet im Hinblick auf diese Voraussetzung Probleme. Solche Nebenbedingungen könnten beispielsweise nicht-negative Geschwindigkeitsgradienten mit der Tiefe sein, um den Parameterraum weiter einzuzugrenzen. Sobald eine solche Bedingung appliziert wird, bedingt diese auch die Wahrscheinlichkeit des Würfelns neuer Modelle und die Gleichverteilung kann nicht mehr garantiert sein. Es werden nicht mehr alle Bereiche des Parameterraumes gleichgewichtet durchsucht. Abbildung 6.4 zeigt hierzu Histogramme der gewürfelten Scherwellengeschwindigkeit einer exemplarischen Startpopulation von 3000 Modellen. Die roten Histogramme repräsentieren eine nicht bedingte Verteilung und zeigen klare Gleichverteilung. Die blauen Histogramme zeigen die Verteilung der Geschwindigkeiten bei der Vorgabe eines positiven Geschwindigkeitsgradienten. Eine gewählte Geschwindigkeit bedingt dabei die nächste. Die Verteilung zeigt nun offenbar keine Uniformität mehr. Da also im vorliegenden Problem alle Intervallgeschwindigkeiten gleichverteilt gewählt werden müssen, lassen sich negative Gradienten in den Geschwindigkeiten nicht ausschließen. Eine mögliche Korrektur der Geschwindigkeitstiefenmodelle verschiebt und diskretisiert die Verteilung der Intervallgeschwindigkeiten. Zur Lösung des Problems werden aus den genannten Gründen Nebenbedingungen wie nicht-negative Geschwindigkeitsgradienten mit

6 ANFORDERUNGEN AN EIN GLOBALES SUCHMUSTER

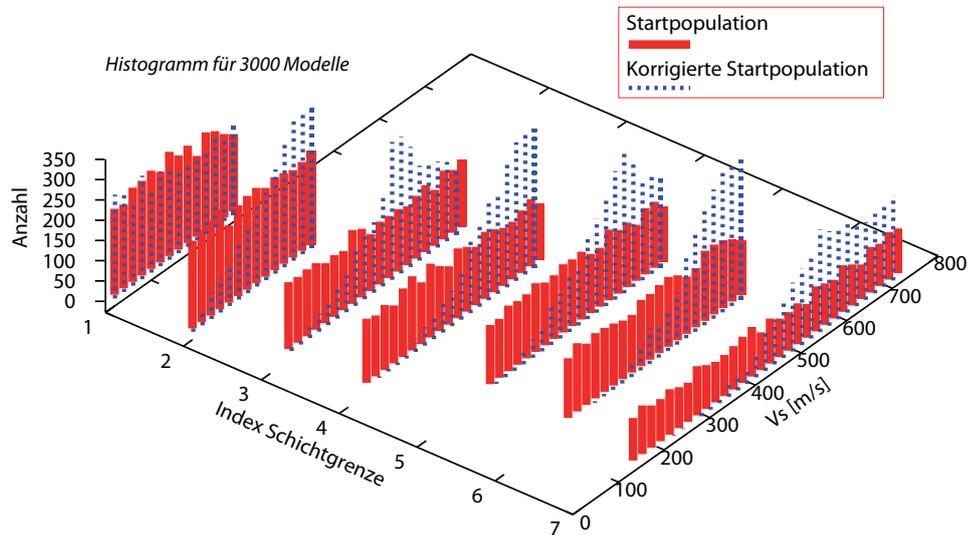


Abbildung 6.4: Geschwindigkeits-Histogramme für eine Startpopulation von 3000 Modellen. Die rote Verteilung zeigt rein zufällig gewählte Modelle, die blaue Verteilung ist bedingt durch nicht-negative Gradienten.

Straftermen belegt, welche den Zielfunktionswert jener Modelle erhöhen. So werden die $i = 1, \dots, n_v$ Intervallgeschwindigkeiten v_{s_i} eines Modells mit zunehmender Tiefe durchlaufen und für jede tiefer als v_{s_i} liegende Intervallgeschwindigkeit v_{s_j} mit geringerer Geschwindigkeit wird ein konstanter Strafwert δ_p zur Zielfunktion addiert. Die Parametrisierung der Vorwärtsmodellierung (siehe Abschnitt 4.2) verwendet die Tiefe als definierenden Parameter, welcher mit angepaßt wird. Die Schichttiefe als Inversionsparameter zu wählen, ist jedoch aus oben genannten Gründen ebenfalls nicht sinnvoll. Schichttiefen bedingen sich automatisch als stets tiefer liegend. Es wurden somit Schichtmächtigkeiten als Inversionsparameter gewählt, welche zur Vorwärtsmodellierung in Tiefen umgerechnet werden können.

7 Angewendete Optimierungsmethoden

Zur Lösung nichtlinearer Optimierungsprobleme existieren viele sehr unterschiedliche aufgestellte Verfahren von einfachen Gradientenverfahren wie der Levenberg-Marquardt Ansatz über Sekantenverfahren mit iterativer Anpassung der Hesse-Matrix, wie das BFGS-Verfahren (Kosmol, 1993) und auch die Sequentielle Quadratische Programmierung (siehe Abschnitt 4.2), bis hin zu stochastischen Verfahren wie den Evolutionsstrategien, Genetischen Algorithmen und Schwarmintelligenzoptimierungsmethoden. Der Fokus dieser Arbeit liegt auf der Anwendung von den in den letzten zwei Jahrzehnten entwickelten Schwarmintelligenzoptimierungsmethoden auf das in den vorigen Kapiteln beschriebene Problem der Inversion von Scherwellengeschwindigkeitsmodellen aus der dispersiven Propagation von Scholtewellen. Der Einsatz dieser Methoden soll hierbei die Abhängigkeit von a priori Informationen durch zu wählende Startmodelle einschränken, numerische Instabilitäten durch ableitungsfreies Vorgehen umgehen, aber vor allem die Auflösung durch globale Exploration des Parameterraumes verbessern, um im Anwendungsbereich der Offshorebaugrundüberwachung Einsatz zu finden.

In den folgenden Abschnitten werden zwei Verfahren der Schwarmintelligenzoptimierung vorgestellt. Dies ist zum einen die Partikel-Schwarm-Optimierung (Eberhardt & Kennedy, 1995) und zum anderen die Optimierung durch Simulation künstlicher Bienenschwärme (Karaboga & Basturk, 2007). Auf der Basis dieser Verfahren werden des weiteren Varianten vorgestellt, die in verschiedener Weise die Waage zwischen Exploration und Konvergenz der einzelnen Verfahren beeinflussen sollen. Als Ergebnis der Untersuchungen wird sich eine hybride Variante aus Partikel-Schwarm-Optimierung und Downhill-Simplex als effektivste Variante aufzeigen lassen. Des weiteren werden die Schwarmintelligenzoptimierungsverfahren mit anderen gängigen stochastischen Methoden verglichen. So wurde ein Simulated Annealing Algorithmus (Metropolis et al., 1953; Kirkpatrick et al., 1983) implementiert sowie eine ableitungsfreie Variante des sogenannten Basin-Hopping (Wales & Doye, 1997) entwickelt und getestet. Ein Evolutionsansatz mit adaptiver Schrittweite (Hansen & Ostermeier, 2001) sowie die einfache Evolutionsstrategie (Rechenberg, 1994) werden ebenfalls zum Vergleich untersucht. Die Verfahren werden hierbei anhand dreier modellierter Spektren als synthetische Eingangsdaten getestet und verglichen.

7.1 Partikel-Schwarm-Optimierung (P.S.O.)

Kooperatives und kollektives Verhalten innerhalb von Gruppierungen kann in verschiedenen Bereichen der Fauna beobachtet werden. So ist es offensichtlich effektiver, wenn beispielsweise bei der Suche einer Gruppe nach einem ersehnten Gegenstand, welcher sich in einem bestimmten Umfeld befindet, alle Gruppenmitglieder an der Suche teilnehmen, aber vor allem ihre gewonnenen Informationen während der Suche untereinander austauschen. Eine bestimmte Art von Gruppierungen sei nun als Schwarm bezeichnet. Die Gruppe sei selbstorganisiert und habe keine Hierarchie oder Anführer. Das einzelne, im Schwarm gleichberechtigte Mitglied hat des weiteren keine Information über das allgemeine oder globale Verhalten der gesamten Gruppe oder die eigene Umgebung. Schwärme haben aber trotz allem die Möglichkeit, sich gemeinsam zu bewegen, nur basierend auf lokalen Interaktionen der Mitglieder untereinander. Aus der einfachen lokalen Interaktion (bspw. dem Informationsaustausch) können Schwärme sehr komplexe Bewegungen und Exploration ihres Umfeldes ausführen (Engelbrecht, 2005). Beobachtungen zum Verhalten von Vogelschwärmen oder Fischeschwärmen können Ideen vermitteln, wie eine effektive und kollektive Suche durchzuführen ist. Man denke sich als einfaches Beispiel einen Fischeschwarm, welcher die ergiebigste unter vorhandenen Futterquellen sucht. Dies ließe sich als Analogon für die Suche nach dem globalen Optimum einer Funktion in einem definierten Bereich eines Parameterraumes sehen.

Ein vielversprechender Ansatz zur Optimierung nichtlinearer, multimodaler Probleme ist in den letzten zwanzig Jahren aus dem Verhalten von Schwarmgruppierungen erwachsen. Die sogenannte Partikel-Schwarm-Optimierung (**P.S.O.**) basiert auf den Prinzipien eines theoretischen Schwarmintelligenzmodells nach Millonas (1994), welches als Hilfsmittel zum Verständnis der Schwarmdynamik dienen soll (Engelbrecht, 2005). Dieses Modell beinhaltet, daß eine Gruppe von Individuen fähig sein sollte, auf Qualitätsfaktoren in ihrer Umgebung zu reagieren (Qualitätsprinzip), einfache Transformationen in Raum und Zeit durchzuführen (Lageprinzip), ihre Ausdehnung in einem gewissen Maße zu erhalten, um die Informationsbreite aus dem Parameterraum zu erhalten (Prinzip der verschiedenartigen Informationen), ihr Verhalten auch bei dynamischem Umfeld zu erhalten (Stabilitätsprinzip) und ihr Verhalten zu ändern, wenn es den Aufwand (Qualitätsmaß) lohnt (Adaptionsprinzip). Um zu verstehen, wie ein Optimierungsverfahren diesen Modellprinzipien entspricht, soll nun zuerst das Grundprinzip des PSO-Verfahrens vorgestellt werden. Grundlage des Verfahren ist ein Schwarm oder eine Menge sogenannter Partikel. Diese Partikel sind Vektoren im Parameterraum. Die Partikel entsprechen den Individuen eines Schwarms und können sich im Parameterraum bewegen. Diese Bewegung durch den multidimensionalen Parameterraum, welche bedingt ist durch drei Anteile, birgt die Grundidee der PSO. Zum einen ist dies die eigene Erfahrung des einzelnen Individuums, welche dem ausgewerteten Qualitätsmaß oder der Zielfunktion an schon bereisten Orten

7.1 Partikel-Schwarm-Optimierung (P.S.O.)

entspricht, und zum anderen die Erfahrung der Nachbarn des Individuums und dem kollektiven, lokalen Informationsaustausch untereinander sowie die bisherige Bewegung der Partikel.

Betrachte nun das Optimierungsproblem mit dem Qualitätsmaß f ,

$$\min f(\vec{x}) \quad f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$$

mit $\vec{x} \in D$. Sei nun $\vec{x}_i(t)$ die Position des Partikels i bei Iteration t . Die Bewegung eines Partikels kann durch Addition eines Verschiebungsvektors geschehen, der als Geschwindigkeit $\vec{v}_i(t)$ bezeichnet sei:

$$\vec{x}_i(t+1) = \vec{x}_i(t) + \vec{v}_i(t) \quad (7.1)$$

Dieser Verschiebungsvektor beinhaltet Informationen über die Erfahrungen des individuellen Partikels und die ausgetauschte Information der Schwarmpartikel in der Nachbarschaft des Individuums. Die individuelle Erfahrung eines Partikels wird als *kognitive Komponente* bezeichnet und sei proportional zur Entfernung des Partikels von seiner eigenen, bisher erreichten besten Position im Raum. Die kollektive Erfahrung des Schwarms, also die sozial ausgetauschte Erfahrung aller Partikel wird als *soziale Komponente* bezeichnet und ist proportional zur Entfernung des Partikels von der besten Position einer nachbarschaftlichen Gruppe (Engelbrecht, 2005). Mit einfachen Worten bedeutet dies, ein Partikel merkt sich die Position mit dem besten Qualitätswert, die es bislang auf seinem Weg erreicht hat. Des Weiteren wird den Partikeln einer bestimmten Nachbarschaftsumgebung die Information über die beste Position, welche die Nachbarschaftsgruppe bislang gefunden hat, bereitgestellt. Nun versucht jedes Partikel, sich anteilig in Richtung dieser beiden Positionen zu bewegen.

Global Best PSO

Die in dieser Arbeit verwendete Grundvariante der PSO ist die *Global Best PSO* oder global beste PSO. Diese Form der PSO definiert den gesamten Schwarm als beeinflussende Nachbarschaft für ein Individuum. Der Geschwindigkeits- oder Verschiebungsvektor der individuellen Partikelbewegung berechnet sich so mit

$$\vec{v}_i(t+1) = \omega \vec{v}_i(t) + c_1 r_{1j} (\vec{y}_{K_i}(t) - \vec{x}_i(t)) + c_2 r_{2j} (\vec{y}_G(t) - \vec{x}_i(t)) \quad (7.2)$$

oder

7 ANGEWENDETE OPTIMIERUNGSMETHODEN

$$v_{ij}(t+1) = \underbrace{\omega v(t)_{ij}}_I + \underbrace{c_1 r_{1j}(y_{Kij}(t) - x_{ij}(t))}_{II} + \underbrace{c_2 r_{2j}(y_{Gj}(t) - x_{ij}(t))}_{III} \quad (7.3)$$

mit $v_{ij}(t)$ der Verschiebung von Partikel i in den Dimensionen $j = 1, \dots, n_x$ (n_x : Dimension des Parameterraumes) zum Zeitschritt t . c_1 , c_2 und ω sind positive Gewichtungskonstanten für die Verschiebungsanteile *I-III*. r_{1j} und r_{2j} sind n_x gleichverteilte Zufallszahlen aus dem Intervall $[0, 1]$. Verschiebungsanteil *I* entspricht der Initialbewegung der Schwarmpartikel aus der letzten Iteration, die dem Schwarm eine Art Trägheit verleiht. Anteil *II* ist der kognitive Anteil der Partikelbewegung. Die von einem individuellen Partikel i bislang nach dem Qualitätsmaß (bspw. eine zu minimierende Qualitätsfunktion f) beste erreichte Position ist in dem Vektor $\vec{y}_{K_i}(t)$ gespeichert (kognitive Komponente). Diese Position wird nach folgender Vorschrift neu belegt (Engelbrecht, 2005).

$$\vec{y}_{K_i}(t+1) = \begin{cases} \vec{y}_{K_i}(t) & \text{wenn } f(\vec{x}_i(t+1)) \geq f(\vec{y}_{K_i}(t)) \\ \vec{x}_i(t+1) & \text{wenn } f(\vec{x}_i(t+1)) < f(\vec{y}_{K_i}(t)) \end{cases}$$

Die global beste Position des gesamten Schwarmes ist gespeichert im Vektor $\vec{y}_G(t)$ und ist die beste der individuell besten Positionen, also

$$\vec{y}_G(t) \in \{\vec{y}_{K_0}(t), \dots, \vec{y}_{K_n}(t)\} \mid f(\vec{y}_G(t)) = \min \{f(\vec{y}_{K_0}(t)), \dots, f(\vec{y}_{K_n}(t))\}$$

($n = \text{Anzahl Partikel im Schwarm}$). Dieser wird neu gesetzt, wenn der Schwarm eine bessere Position gefunden hat. Ein Schwarmpartikel bewegt sich somit anteilig entlang der ursprünglichen Bewegung, in Richtung der besten Position in seiner Erinnerung und in Richtung der kollektiv besten Position. Die letzten beiden Anteile werden hierbei noch mit einer zufälligen Streuung durch Multiplikation mit den Zufallszahlen r_{1j} und r_{2j} versehen.

Der Algorithmus der PSO geht nun in folgenden Schritten vor.

1. Initialisiere eine Menge von $i = 1, \dots, n$ Partikeln \vec{x}_i und evaluiere Qualitätsfunktion $f(\vec{x}_i)$. Iterationsschritt $t = 0$
2. Berechne die individuell beste erreichte Position jedes Partikels $\vec{y}_{K_i}(t)$ ($\vec{y}_{K_i}(t) = \vec{x}_i$ für $t = 0$).
3. Berechne global beste Position als

$$\vec{y}_G(t) \in \{\vec{y}_{K_0}(t), \dots, \vec{y}_{K_n}(t)\} \mid f(\vec{y}_G(t)) = \min \{f(\vec{y}_{K_0}(t)), \dots, f(\vec{y}_{K_n}(t))\}$$

4. Verschiebe Partikel nach Gleichung 7.2
5. Prüfe Abbruchbedingung und gehe zu Schritt 2.

7.1 Partikel-Schwarm-Optimierung (P.S.O.)

Es läßt sich zeigen, daß PSO die oben genannten Schwarmmodelleigenschaften besitzt. Durch die stete Anpassung der Partikelpositionen nach Auswertung der Qualitätsfunktion in allen Iterationsschritten wird das Qualitätsprinzip erfüllt. Außerdem reagiert die PSO auf Qualitätsänderung durch Anpassung der individuell besten und der global besten Position. Das Lage-Prinzip ist somit ebenfalls gegeben. Der Schwarm ändert seinen Zustand nur, wenn sich $\vec{y}_G(t)$ und $\vec{y}_{K_i}(t)$ ändern, was eine Stabilität bestimmt. Der Schwarm ändert seinen Zustand dementsprechend adaptiv und paßt sich an, wenn $\vec{y}_G(t)$ und $\vec{y}_{K_i}(t)$ sich ändern. Die Wechselwirkung von *kognitiven* und *sozialen* Anteilen sorgt für eine Anpassung der Ausdehnung des Schwarms.

Abbildung 7.1 ist eine geometrische Illustration der Scharmbewegung und der Anpassung des Verschiebungsvektors. Die Abbildung zeigt eine Qualitätsfunktion F im Parameterraum $(x_1, x_2) \subset \mathbb{R}^2$. Der weiße Pfad zeigt die letzten vier Positionen $\vec{x}(t-3)$ bis $\vec{x}(t)$ des dargestellten Partikels \vec{x} . Ebenso sind die global beste Position \vec{y}_G (rot) und die individuell beste Position des Partikels $\vec{y}_K = \vec{x}(t-2)$ (blau) eingetragen. Die Pfeile zeigen die Komponenten des Verschiebungsvektors an, resultierend in der neuen Position $\vec{x}(t+1)$ (gelb).

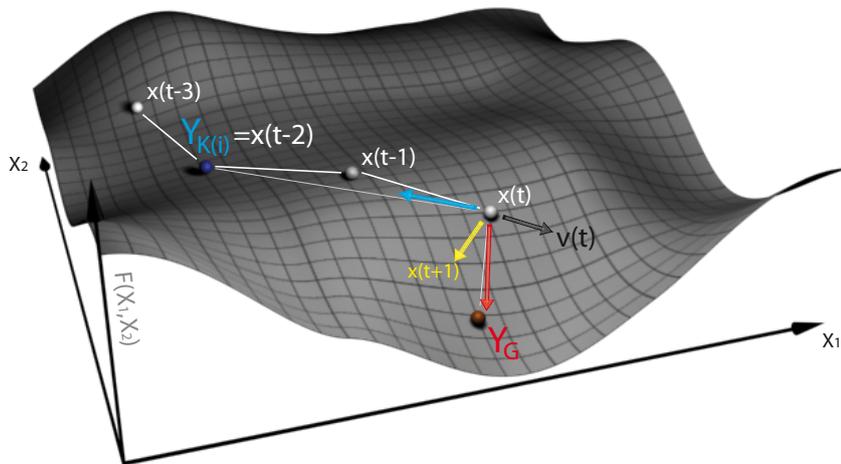


Abbildung 7.1: Geometrische Illustration der Bewegung eines Schwarmpartikels.

7.2 Varianten der PSO: Ein Überblick

Die oben aufgezeigte Grundform der PSO wurde in verschiedensten Optimierungsproblemen erfolgreich angewandt (bspw. normale Funktionsoptimierung (Mishra, 2006), aber auch bereits in geophysikalischen Inversionsproblemen (Shaw & Srivastava, 2007)). Jedoch zeigen sich oft Probleme in der effektiven Konvergenz des Verfahrens. So zeigt Engelbrecht (2005), daß die PSO nicht zwingend gegen ein globales oder sogar lokales Optimum konvergiert. Die PSO ist kein lokalsuchender Algorithmus, da sie keine garantierte Konvergenz zu lokalen Optima besitzt. Während verschiedenste empirische Arbeiten die Fähigkeit der PSO aufzeigten, Optimierungsprobleme zu lösen, zeigen sie aber auch die Probleme der konsistenten lokalen Konvergenz. So kann die Konvergenz des Schwarms stagnieren, aber auch seine Ausdehnung zu früh zu klein werden, so daß kein Optimierungsfortschritt mehr möglich ist, obwohl das Optimum noch nicht erreicht ist. Dieses grundlegende Problem ließ im Laufe der Jahre unterschiedlichste Varianten der PSO entstehen, welche Konvergenz und Exploration des Schwarms modulieren. Da aufgrund stets begrenzter Zeit- und Rechenkapazitäten nicht alle entwickelten Varianten am Problem dieser Arbeit getestet werden können, soll hier eine Auswahl der vorab aussichtsreichsten Varianten vorgestellt werden.

PSO mit garantierter lokaler Konvergenz: Guaranteed Convergence PSO - GCPSO

Die GCPSO-Variante umgeht das Problem der nicht garantierten lokalen Konvergenz. So führt das global beste Partikel \vec{y}_G parallel zur globalen Schwarmsuche eine lokale, stochastische Suche durch. Dies wird durch Variation der Verschiebungsberechnung erreicht. Das global beste Partikel j ändert nun seine Position nach

$$\vec{x}_j(t+1) = \vec{y}_G + \omega \vec{v}_j(t) + \underbrace{\tau(t)(1 - 2 \cdot r_2(t))}_{IV} \quad (7.4)$$

mit $r_2(t) \in [0, 1]$. $\tau(t)$ ist hierbei ein Skalierungsfaktor, welcher die Schrittweite der Zufallsuche IV steuert. Dieser wird nach folgender Bedingung gewählt.

$$\tau(t+1) = \begin{cases} 2\tau(t) & \text{wenn } N_{Erfolge}(t) > \epsilon_E \\ 0.5\tau(t) & \text{wenn } N_{Misserfolge}(t) > \epsilon_M \\ \tau(t) & \text{sonst} \end{cases} \quad (7.5)$$

Diese Bedingung läßt den Bereich der lokalen Suche um \vec{y}_G anwachsen, wenn die Anzahl an erfolgreichen Schritten der Suche einen Grenzwert ϵ_E überschreitet. Es wird somit probiert, ob in diesem vermeintlich erfolgreichen Suchbereich auch größere Schritte möglich sind. Führt die Folge der Suchposition mehr als ϵ_M -mal zu keinem Optimierungserfolg, so wird der Suchbereich eingeschränkt und die Wahrscheinlichkeit erhöht, eine Abstiegsrichtung zu entdecken. GCPSO verfügt also über eine zusätzliche stochastische, lokale Suche mit adaptivem Suchbereich. Die Effizienz der Variante hängt dabei von der Wahl der Parameter $\tau(0)$, ϵ_E und ϵ_M ab, welche problemabhängig zu wählen sind. Van den Bergh et al. (2002) empfahlen jedoch $\epsilon_E=15$ und $\epsilon_M=5$ für hochdimensionale Probleme. Das Verfahren tadelt so eine schlechte Wahl von τ schneller als daß es erfolgreiche Werte von τ begünstigt (Engelbrecht, 2005). Die GCPSO ermöglicht garantierte lokale Konvergenz (Engelbrecht, 2005), da die Wahrscheinlichkeit, einen neuen optimaleren Punkt zu erreichen, durch stetes Generieren absolut neuer Suchpunkte immer größer Null ist.

Reset-PSO

Nach Engelbrecht (2005) ist ebenfalls nicht zwingend gegeben, daß die PSO global konvergiert. So erfüllt die PSO die Bedingungen für eine globale Suche nicht und kann somit auch nicht als globale Optimierungsmethode benannt werden. Diese Bedingung beinhaltet zusammengefaßt, daß für eine Teilmenge M des Suchraumes die Wahrscheinlichkeit, M wiederholt durch Zufallssuche nicht zu erreichen, Null sein muß. Mit anderen Worten, die Wahrscheinlichkeit, einen Punkt in einer Umgebung um das globale Optimum zu erreichen, muß verschieden von Null sein. Um das Problem der nichtglobalen Suche umgehen zu können, seien nun zwei Varianten der PSO benannt, die das Kriterium einer globalen Optimierungsmethode nachweislich erfüllen.

1. **Zufalls-Partikel PSO:** Diese Variante erfüllt die Bedingung der globalen Suche durch zufällige Neuverteilung einiger (wenigstens eines) zufällig gewählter Partikel zu jeder Iteration. Die Verschiebung der Partikel folgt nach wie vor Gleichung 7.2.
2. **Neustart PSO:** Diese Variante nutzt ebenfalls die Aufdatierungsgleichung 7.2 für die Teilchenverschiebung. Jedoch wird stets ein Konvergenzmaß berechnet, welches beim mehrfachen Unterschreiten eines bestimmten Grenzwertes die Konvergenz des Schwarms zu einem stabilen Punkt anzeigt und dem Algorithmus so das Signal gibt, einen Neustart durchzuführen. Die bisherige beste Position wird gespeichert und die Suche beginnt mit neuer Startpopulation von vorn. Die globale Konvergenzbedingung ist erfüllt, wenn sich zeigen läßt, daß der Schwarm so eine unendliche Anzahl von Neustarts durchführen kann (Engelbrecht, 2005). So ist die Wahl des Konvergenzkriteriums von essentieller Bedeutung für diese Variante.

7 ANGEWENDETE OPTIMIERUNGSMETHODEN

- Zum einen läßt sich ein Neustart nach einer festen Anzahl Iterationen festlegen. Jedoch birgt dies die Gefahr, daß das System zu früh reinitialisiert wird, bevor ein Optimum erreicht wurde. Zum anderen können Konvergenzbedingungen definiert werden, die als Maß für einen Neustart gelten.
- Es läßt sich zeigen, daß die PSO stets den Zustand $\vec{x}_i = \vec{y}_{K_i} = \vec{y}_G$ anstrebt. So läßt sich zum Beispiel der maximale Schwärmdurchmesser nutzen, um Konvergenz anzuzeigen. Unterschreitet der Schwärmdurchmesser einen bestimmten Wert für eine gewisse Anzahl von Iterationen, so läßt sich Konvergenz zu einem stabilen Optimalpunkt annehmen.
- Auch kann man die Frage nach dem Erfolg der Iterationsschritte stellen. So kann ein Schwarm reinitialisiert werden, wenn über eine definierte Anzahl Iterationen die Verbesserung des global besten Qualitätsmaßes stagniert, bzw. die Kurve der Verbesserung, \tilde{f} , abflacht. So wird jede Iteration als Mißerfolg gewertet für die gilt

$$\tilde{f} = \frac{f(\vec{y}_G(T)) - f(\vec{y}_G(T + 1))}{f(\vec{y}_G(T))} < \epsilon$$

wobei $f(\vec{y}_G(T))$ und $f(\vec{y}_G(T + 1))$ das alte und das neu gefundene global beste Qualitätsmaß sind.

- Auch läßt sich nach Xie et al. (2002) diese Bedingung für die Reinitialisierung einzelner Partikel definieren. So wird ein Partikel $\vec{x}_i \neq \vec{y}_G$ reinitialisiert, wenn für eine bestimmte Anzahl Iterationen $f(\vec{x}_i - \vec{y}_G) < \epsilon$.

Geladene PSO - Charged PSO

Ein weiteres Problem ist, daß die soziale Komponente des Schwarms zu stark wirken kann, und die Ausdehnung des Schwarms zu früh verringert wird. Dies kann zu nicht ausreichender Exploration des Parameterraumes führen, was wiederum die Wahrscheinlichkeit eines Schrittes aus einem lokalen Minimum oder des Findens besserer Suchbereiche verkleinert. Eine Möglichkeit, die Diversität eines Schwarms zu erhalten, ist das Einführen von Abstoßungsbedingungen, die verhindern, daß Partikel sich zu nahe kommen. Damit dies nicht die Effizienz des Algorithmus in der Konvergenz vermindert, sollte nur ein Teil des Schwarms mit dieser Eigenschaft belegt sein. Blackwell & Bentley (2002) entwickelten zu diesem Problem die sogenannte „geladene PSO“. Dieses Verfahren nutzt ein pseudo-elektrostatisches Potential. Die Partikel werden hierbei als geladene Teilchen betrachtet und mit einem vereinfachten Coulomb'schen Potential umgeben. Es

7.2 Varianten der PSO: Ein Überblick

werden zwei gegensätzliche Kräfte innerhalb der Schwarmdynamik eingeführt. Zum einen ist dies die PSO-eigene, sozial induzierte anziehende Wirkung, welche die Konvergenz zu einer Lösung antreibt. Zum anderen sei dies die abstoßende pseudo-elektrostatische Kraft, die dem Schwarm die Diversität oder Explorationsfähigkeit erhält.

Der Ansatz von Blackwell & Bentley (2002) führt einen einfachen „Beschleunigungs“-term, $\vec{a}_i(t)$, in der Verschiebungsgleichung 7.2 ein,

$$\vec{v}_i(t+1) = \omega \vec{v}_i(t) + c_1 r_{1j} (\vec{y}_{K_i}(t) - \vec{x}_i(t)) + c_2 r_{2j} (\vec{y}_G(t) - \vec{x}_i(t)) + \vec{a}_i(t)$$

$$\vec{a}_i(t) = \sum_{l=1, l \neq i}^{n_s} A_{il}(t)$$

$$A_{il}(t) = \begin{cases} \left(\frac{Q_i Q_l}{\|\vec{x}_i(t) - \vec{x}_l(t)\|^3} \right) (\vec{x}_i(t) - \vec{x}_l(t)) & \text{wenn } R_k \leq \|\vec{x}_i(t) - \vec{x}_l(t)\| \leq R_w \\ \left(\frac{Q_i Q_l}{R_k^2 \|\vec{x}_i(t) - \vec{x}_l(t)\|} \right) (\vec{x}_i(t) - \vec{x}_l(t)) & \text{wenn } \|\vec{x}_i(t) - \vec{x}_l(t)\| < R_k \\ 0 & \text{wenn } \|\vec{x}_i(t) - \vec{x}_l(t)\| > R_w \end{cases}$$

(Engelbrecht, 2005). Die $A_{il}(t)$ beschreiben die Wechselwirkung zwischen Teilchen i und l , Q_i und Q_l sind die jeweiligen „Ladungen“, also die Magnitude der abstoßenden Wechselwirkung. R_k ist der Kernradius des Teilchens und R_w der Wirkungsradius jedes Teilchens. Die Abstoßung zweier Partikel passiert also nur dann, wenn die Entfernung der beiden innerhalb des Intervalls $[R_k, R_w]$ ist. In diesem Fall nimmt die abstoßende Wirkung mit sinkendem Abstand invers zu. Die Partikel werden also niemals konvergieren, aufgrund der sehr starken Wechselwirkung bei kleinen Abständen. Um dies zu umgehen, wird die Abstoßungswirkung am Kernradius R_k festgesetzt für Abstände kleiner R_k . Außerhalb von R_w ist $\vec{a}_i = 0$; Partikel können sich hier also unbeeinflusst in Richtung der Schwarmverschiebung bewegen.

Eine Variation der geladenen PSO ist die sogenannte „atomare“-PSO. Diese Variante definiert die Ladung für den halben Schwarm auf $Q_i > 0$ und $Q_i = 0$ sonst. Der halbe Schwarm ist somit neutral. Die atomare PSO hat sich in empirischen Studien als effektivste geladene Variante gezeigt, denn wenn die Entfernungen der Partikel niemals unterhalb von R_k sinken sollten, wird der Schwarm nicht konvergieren. Das Einführen neutraler Partikel erlaubt eine Konvergenz dieses Schwarmanteils, während der geladene Anteil die Diversität des Systems erhält. Abbildung 7.2 zeigt nochmal das Konzept der atomaren PSO. Die Wahl der Parameter R_k , R_w und Q ist problemabhängig und es sei hierzu auf Abschnitt 8.1 verwiesen.

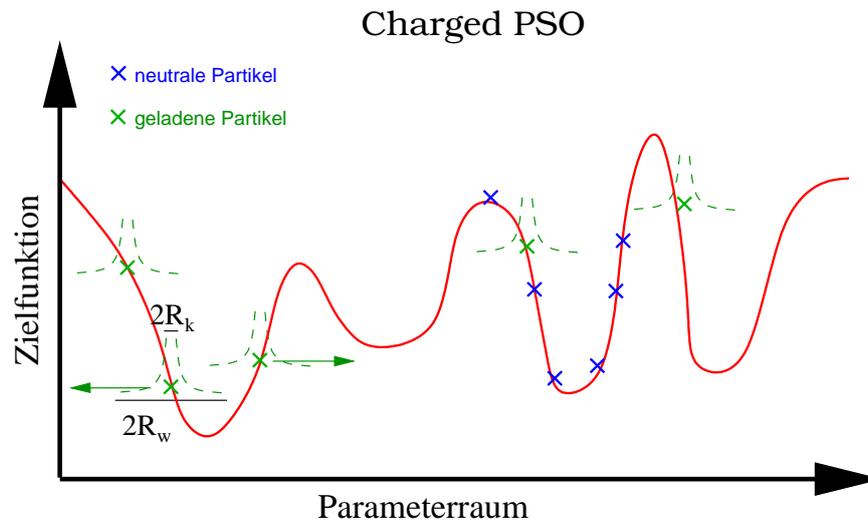


Abbildung 7.2: Illustration zum Prinzip der atomaren PSO. Die grünen Kreuze zeigen die geladenen Partikel, die im Falle einer Annäherung jenseits R_w abstoßende Impulse erhalten. Die neutralen (blauen) Partikel garantieren mögliche Konvergenz.

Rasternde PSO - Niching PSO

Die Neustart-PSO ermöglicht eine theoretische globale Konvergenz. Jedoch ist hierbei nicht der Rechenaufwand erwähnt, den vielmaliges Neustarten des Verfahrens beinhaltet. So steht oftmals aus Zeitgründen eine nur begrenzte Anzahl an Neustarts zur Verfügung. Es liegt daher auf der Hand, daß Möglichkeiten bedacht werden müssen, die verhindern, daß ein erneut gestarteter Schwarm mehrfach in die selben Optima konvergiert. Die bereits gewonnenen Informationen über entdeckte Optima sollten effektiv genutzt werden, um Rechenzeit zu sparen und die Wahrscheinlichkeit der globalen Konvergenz innerhalb begrenzter Anzahl von Neustarts zu erhöhen. Des weiteren ist es von großem Vorteil, wenn bei nicht eindeutigen Problemen die Menge aller vorhandenen Optima abgerastert wird, um gegebenenfalls nachträglich aus dieser Menge wählen zu können. Parsopoulos & Vrahatis (2004) veröffentlichten eine Arbeit über die Möglichkeiten zur Berechnung aller Minima eines Problems mittels PSO. Grundidee war zum einen die Veränderung der Zielfunktion nach den bereits erhaltenen Informationen. So kann die Zielfunktion in den Bereichen gefundener Minima verzerrt werden, so daß sie schlechtere Werte liefert. Der suchende Algorithmus würde so von den bisherigen Optima abgelenkt. Zum anderen lassen sich im Falle der PSO auch die Partikel von bereits gefundenen Optima entfernungsabhängig abstoßen.

1. **Abstoßende PSO - Repulsion PSO:** Nach erfolgreicher Detektion eines Optimums wird dessen Position gespeichert. Wenn ein Partikel sich in Richtung dieser gespeicherten Positionen bewegt, so wird es analog der ge-

$$H(\vec{x}) = G(\vec{x}) + \gamma_2 \frac{\text{sign}(f(\vec{x}) - f(\vec{y}^*)) + 1}{\tanh(\mu(G(\vec{x}) - G(\vec{y}^*)))}$$

wobei γ_1 , γ_2 und μ problemabhängige Konstanten sind. γ_1 kontrolliert hierbei das entfernungsabhängige Verzerren, γ_2 und μ bestimmen den Wirkungsradius und die Amplitude der zweiten Verzerrungsphase. Ein Problem, welches die zweite Phase mit sich bringt, ist der sogenannte „Mexican-Hat“ Effekt. Das Anheben des gefundenen Minimums und seiner Umgebung kann bei zu kleiner Ausdehnung der Verzerrung ein neues Minimum in dieser Umgebung schaffen. Wird nur der innere Bereich eines Minimums umgebenen „Tals“ angehoben, so entsteht die Form des Mexican-Hat und neue lokale Minima entstehen. Eine Kombination mit der Variante des abstoßenden Schwarms aus Punkt 1. kann bei diesem Problem Abhilfe schaffen.

Hybride Ansätze

Viele Ansätze zur Verbesserung der Eigenschaften der PSO nutzen die Verknüpfung des Verfahrens mit anderen Optimierungsalgorithmen. Die Hoffnung bei diesem Vorgehen ist oftmals die erfolgreiche Verbindung der positiven Eigenschaften unterschiedlicher Verfahren. Die Anwendung der PSO auf das Scholtewelleninversionsproblem zeigte von Anfang an Probleme in der effektiven Konvergenz im Nahbereich eines Minimums. Dies wurde getestet, indem eines der Partikel der Startpopulation in die Nähe eines bekannten Minimums der Qualitätsfunktion platziert wurde. Nachdem der Schwarm sich dem Nahbereich des Minimums genähert hatte, konnte keine effiziente Konvergenz mehr festgestellt werden. Ein weiteres, stets präsent Problem ist mangelnde Diversität des Schwarms, was zu mangelnder Exploration des Parameterraumes führt. Auch dieses Problem kann durch Assimilation von Eigenschaften anderer Optimierungsverfahren angegangen werden. Die folgenden Abschnitte zeigen ein vielversprechendes Konzept der Verbindung der PSO mit dem lokalen Downhill-Simplex-Verfahren auf, das im Rahmen dieser Arbeit getestet wurde.

„Memetische“ PSO - Memetic PSO

Viele globale Optimierungsverfahren zeigen zu späteren Stadien des Optimierungsprozesses schlechte Konvergenz- und Schrittweiteigenschaften. Die Algorithmen beweisen oft schnelle Konvergenz zu Bereichen guter Qualitätsfunktion, verlieren jedoch im Nahbereich des Optimums ihre Effektivität. Obwohl die Konvergenz- und Explorationseigenschaften durch die Schwarmparameter gesteuert werden können, gilt dieses Problem auch bei der PSO (Engelbrecht, 2005). Eine vielversprechende Anpassung der PSO beinhaltet das Integrieren lokaler Suchalgorithmen, welche zwischen den Iterationen der PSO für das global beste Partikel eingeschaltet werden. Dieses können stochastische lokale Suchverfahren

sein oder Gradientenverfahren.

Da die numerische Auswertung von Gradienten über Differenzenquotienten im vorliegenden Inversionsproblem numerische Instabilitäten mit sich bringen kann, bleiben nur gradientenfreie, lokale Optimierungsmethoden. Dies wären stochastische Suchverfahren wie Metropolis-(siehe Abschnitt 7.4) und Hill-Climbing Verfahren oder die Downhill-Simplex Methode.

Downhill-Simplex Algorithmus DHS

Nelder & Mead (1965) stellten einen ableitungsfreien, lokalen Optimierungsalgorithmus vor. Das Verfahren basiert auf der geometrischen Figur des „Simplex“. Betrachte erneut das Problem

$$\min f(\vec{x}) \quad f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$$

Ein Simplex sei nun ein $(n+1)$ -dimensionales Polygon. Für $n=2$ wäre ein Simplex ein Dreieck und für $n=3$ ein Tetraeder. Ausgehend von einem Startvektor $\vec{x}_0 \in D$ erzeuge man nun ein Simplex, indem n weitere Punkte ausgehend von \vec{x}_0 in Richtung der Basisvektoren im Abstand λ erzeugt werden (siehe Abbildung 7.4 a)). Die n weiteren Eckpunkte des Simplex definieren sich also mit $\vec{x}_{S_i} = \vec{x}_0 + \lambda \vec{e}_i$ ($i=1, \dots, n$) mit \vec{e}_i den Einheitsbasisvektoren. An allen Simplexpunkten wird nun die Qualitätsfunktion ausgewertet. Der Punkt mit dem schlechtesten (größten) Funktionswert wird durch den Mittelpunkt der restlichen Simplexpunkte gespiegelt (siehe Abbildung 7.4 c₁)) und an seiner neuen Position ausgewertet. Diese Vorgehensweise wird wiederholt. Der Simplex transformiert sich so in jeder Iteration und bewegt sich „downhill“, in Bereiche besserer Qualitätswerte. Diese Vorgehensweise ist in Abbildung 7.4 b) mit den Schritten 1-3 skizziert. Der neue Simplexpunkt, der aus dieser *Reflexion* entsteht, berechnet sich durch

$$\vec{x}_R = \vec{x}_M + \alpha(\vec{x}_M - \vec{x}_S) \quad (7.6)$$

wobei \vec{x}_S der Punkt mit dem schlechtesten Qualitätswert, α der Reflexions-Skalierungsfaktor und \vec{x}_M der Mittelpunkt der restlichen Punkte ist.

$$\vec{x}_M = \frac{1}{n} \sum_{i \neq S} \vec{x}_i$$

Dieses Vorgehen ist nur sinnvoll, wenn die Reflexionspunkte stets bessere Qualitätswerte besitzen. Da dies nicht gegeben sein muß, kann das Verfahren drei verschiedene Transformationen durchführen, die den Simplex entweder ausdehnen oder kontrahieren lassen, um seine Schrittweite anzupassen. Der Algorithmus sei nun im folgenden schrittweise beschrieben:

7 ANGEWENDETE OPTIMIERUNGSMETHODEN

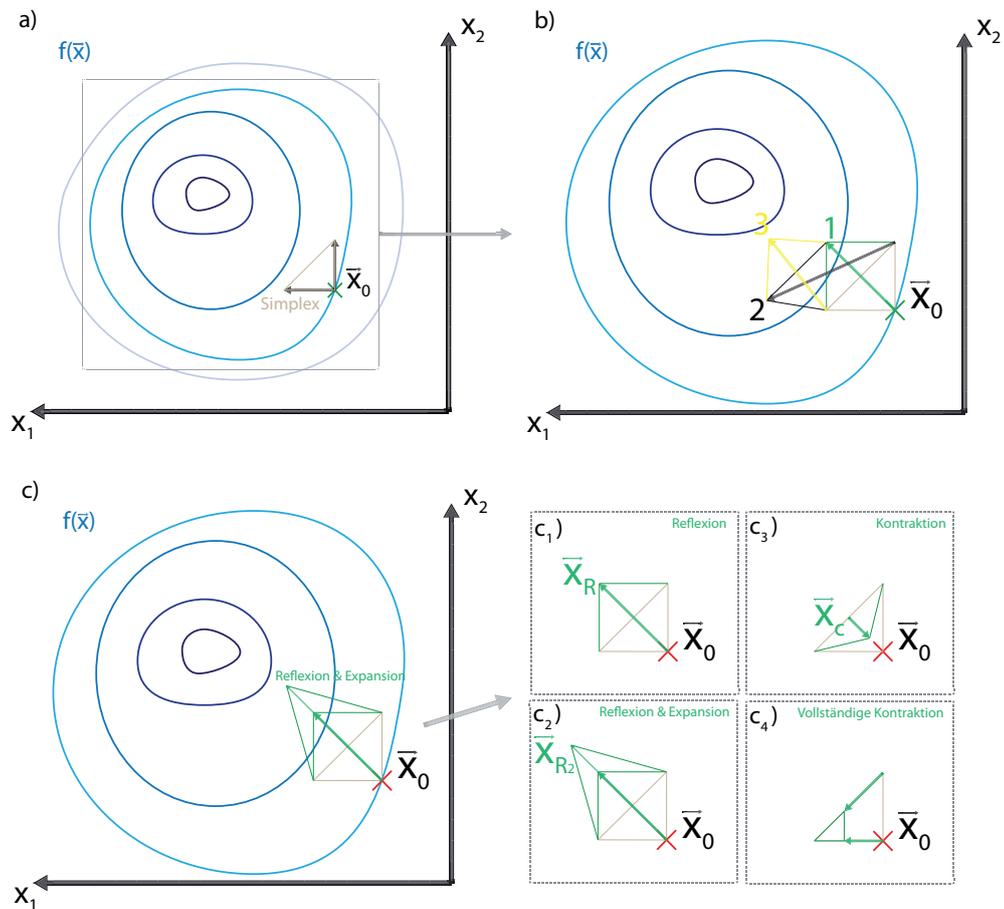


Abbildung 7.4: Verfahren des Downhill-Simplex anhand eines zweidimensionalen Beispiels mit einer skizzierten Qualitätsfunktion $f(x_1, x_2)$ (dunkeres blau entspricht kleineren Werten). Downhill Schritt b) sowie Simplextransformationen c1)-c4) sind hier skizziert.

1. Berechne die Reflexion nach Gleichung 7.2. (siehe Abbildung 7.4 c₁)
2. Ist der Reflexionspunkt R besser als Punkt S ($f(\vec{x}_B) \leq f(\vec{x}_R) < f(\vec{x}_S)$, $\vec{x}_B = \min\{\vec{x}_i\}$), so ersetze Punkt S durch den Reflexionspunkt und beschließe Iterationsschritt
3. Wenn der neue Punkt R der beste bisherige Punkt ist ($f(\vec{x}_R) < f(\vec{x}_i) \forall i$), dann expandiere den Simplex über den Reflexionspunkt hinaus.

$$\vec{x}_{\hat{R}} = \vec{x}_M + \gamma(\vec{x}_M - \vec{x}_S)$$

Ist der neue Reflexionspunkt besser als der alte, dann ersetze Punkt S durch den neuen Reflexionspunkt und beschließe den Iterationsschritt. Andernfalls ersetze Punkt S durch den alten Reflexionspunkt. (siehe Abbildung 7.4 c₂)

4. Wenn der Reflexionspunkt schlechter oder gleich dem zweitschlechtesten Simplexpunkt ($f(\vec{x}_R) \geq f(\vec{x}_{S2})$), dann berechne einen kontrahierten Simplex mit

$$\vec{x}_C = \vec{x}_S + \eta(\vec{x}_M - \vec{x}_S)$$

(siehe Abbildung 7.4 c_3). Wenn $f(\vec{x}_C) \leq f(\vec{x}_S)$, dann ersetze Punkt S durch den Kontraktionspunkt und beschließe den Iterationschritt.

5. Wenn durch Reflexion, Expansion oder Kontraktion keine Verbesserung eintritt, dann kontrahiere den gesamten Simplex um den besten Punkt (siehe Abbildung 7.4 c_4).

$$\vec{x}_i = \vec{x}_B + \sigma(\vec{x}_i - \vec{x}_B)$$

Die Parameter λ , α , η , σ und γ modulieren die Transformationen und müssen problemabhängig angepaßt werden. Standardwerte sind jedoch $\alpha = 1$, $\gamma = 2$, $\eta = 0.5$ und $\sigma = 0.5$.

Hybrid aus DHS und PSO

Ausgehend von der Beobachtung, daß die PSO im Nahbereich eines Optimums stagnierende Konvergenz aufweisen kann und den schwierigen Voraussetzungen für die numerische Bestimmung des Gradienten des Inversionsproblems, liegt es auf der Hand, die PSO mit einem gradientenfreien, lokalen Optimierungsverfahren zu verknüpfen. Dies wurde mit dem Downhill-Simplex (analog Koduru et al. (2006)) als auch mit dem stochastischen Metropolis-Algorithmus (siehe Abschnitt 7.4) durchgeführt. Abbildung 7.5 zeigt ein Flußdiagramm des resultierenden Algorithmus, bei dem nach der Schwarmbewegung und der Evaluierung des global besten Partikels eine lokale Suche angeschlossen wird. Startvektor der Suche ist das global beste Partikel. Die Suche wird für eine feste Anzahl Iterationen durchgeführt.

7 ANGEWENDETE OPTIMIERUNGSMETHODEN

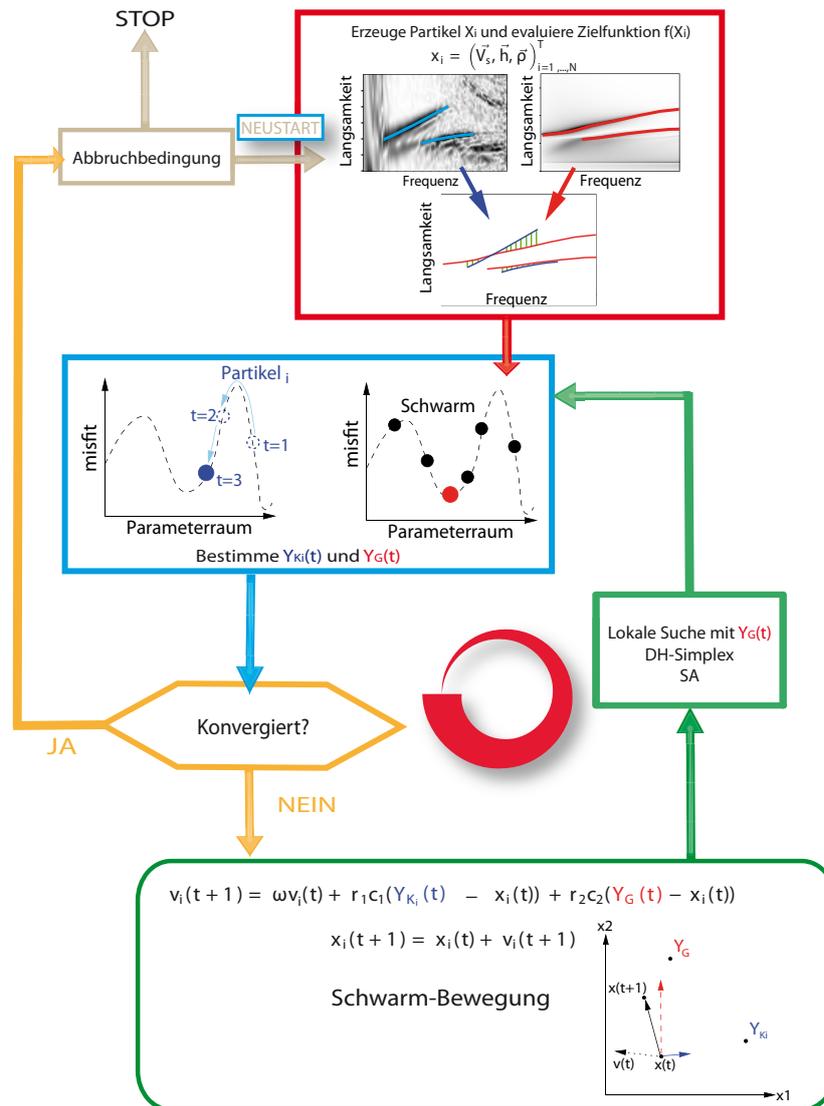


Abbildung 7.5: Flußdiagramm des Hybriden aus PSO und lokaler, gradientenfreier Suche.

Weitere Varianten des PSO

Ausgehend von der Formulierung des Problems der Anpassung von Scholtwellendispersion läßt sich an dieser Stelle vermuten, daß der dominierende Faktor des Problems die Multimodalität der Qualitätsfunktion ist. So wird erwartet, daß Varianten mit erhöhtem Fokus auf Exploration und Multistart-Varianten die nennenswertesten Verbesserungen in der Auflösung der Inversion mit sich bringen. Ungeachtet dieser Vermutung seien hier noch weitere Varianten der PSO

vorgestellt, die sich mit verschiedenen Verfahrensoptimierungen beschäftigen. Alle erwähnten Varianten wurden im Rahmen dieser Studien implementiert, fielen jedoch schon in der Beobachtung des Konvergenzverhaltens und Anwendung an einer Testfunktion aus der engeren Auswahl und werden daher im Vergleich am geologischen Testmodell in Kapitel 8, auch aufgrund begrenzter Rechenzeiten, nicht mehr berücksichtigt.

Zu diesen Varianten gehören Ansätze zur adaptiven Geschwindigkeitsregulierung, um Schrittweiten ggf. abzdämpfen, Ansätze zur adaptiven Anpassung des Inertialbewegungsanteiles, denn der Gewichtungsfaktor ω in der Aufdatierungsgleichung 7.2 moduliert den Einfluß der bisherigen Bewegung des Schwarms auf die neue Bewegung. ω definiert somit das Zusammenspiel von lokaler und globaler Exploration des Schwarms (Parsopoulos & Vrahatis, 2004), denn hohes ω bedingt hohe globale Exploration durch den Erhalt der Bewegung. Die Wahl von ω ist somit von großer Bedeutung. Da in frühen Iterationen eine hohe globale Exploration sinnvoll erscheint, um qualitativ gute Bereiche im Suchraum zu finden und in späten Iterationen die lokale Exploration forciert werden sollte, bietet sich ein mit Verlauf der Optimierung kleiner werdendes ω an.

Auch ein Verfahren, welches zur Erhöhung der globalen Exploration sogenannte Jägerteilchen einführt, wurde getestet. Diese Variante definiert ein Partikel als für die anderen Partikel abstoßendes Teilchen, welches stets in Richtung der besten erreichten Position strebt (Silva et al., 2002). Das Jägerpartikel \vec{x}_p bewegt sich nach der Verschiebungsformel $\vec{v}_p(t+1) = r_j(\vec{y}_G - \vec{x}_p(t))$, wobei r_j für jede Dimension j eine Zahl zwischen 0 und 1, die Maximalgeschwindigkeit des Jägers definierenden Konstanten. Die Partikel werden ihre Verschiebung nun durch einen weiteren Summanden $c_3 r_{3j} D(d)$ erweitern, wobei d der euklidische Abstand zum Jäger und $D(d) = \alpha e^{-\beta d}$ ist (c_3 ist eine Gewichtungskonstante, r_{3j} Zufallszahlen für jede Dimension und α, β den Jagdeinfluß kontrollierende Konstanten). Auch Varianten, bei denen ein dritter Summand einen einfachen Zufallsvektor enthält (Repulsive PSO (Mishra, 2006)), wurden getestet.

Ebenfalls getestet wurden Varianten, die evolutionstrategische Rekombinationsoperatoren in die PSO integrieren. So können beispielsweise die schlechtesten Partikel durch Partikel ersetzt werden, die Eigenschaften des global besten Partikels und eines zufällig gewählten über Rekombination erhalten (Pant et al., 2007).

7.3 „Artificial Bee Colony“ (ABC) Algorithmus

Analog zur PSO, welche das Verhalten von Fisch- oder Vogelschwärmen adaptiert, entwickelte Karaboga & Basturk (2007) ein Verfahren, welches auf dem Verhalten von Bienenschwärmen basiert. Die treibende Kraft für die Bewegungen eines Bienenschwarms ist die Suche und das effektive Ausbeuten von Futterquellen in der Umgebung des Bienenstocks. Ein fundamentales Modell für diese Dynamik, wel-

7 ANGEWENDETE OPTIMIERUNGSMETHODEN

ches die kollektive Intelligenz eines Bienenschwarmes simuliert, besteht aus drei Komponenten (Karaboga & Basturk 2007). Dies sind die Futterquellen selbst, die an einer Quelle „angestellte“ Arbeiterdronen, welche die Quelle ausbeuten sowie Dronen, welche keiner Quelle fest zugeordnet sind. Es werden des weiteren zwei grundlegende Handlungen benötigt: Die Zuordnung von Arbeiterdronen zu einer Futterquelle und das Verlassen einer erschöpften Quelle. Der Wert einer Futterquelle kann für einen Bienenschwarm durch mehrere Faktoren bestimmt sein. Dies sei zum einen der Gehalt der Quelle, zum anderen aber auch die Nähe zum Bienenstock. In Analogie zum Problem einer Optimierung entspräche eine mögliche Futterquelle einem Bereich des Parameterraumes. Der Bereich wird von den Dronen abgesucht und der Ertrag dieser Quelle bestimmt sich nach dem Wert der dortigen Qualitätsfunktion und nach dem Erfolg der Optimierung. Die angestellten Dronen sind somit Partikel, die in einem solchen Bereich eine stochastische Suche vornehmen, den Bereich sozusagen abernten. Die Information über die bearbeitete Futterquelle wird dem Schwarm mitgeteilt, um mit einer bestimmten Wahrscheinlichkeit Dronenverstärkung an besagter Quelle anzuordnen. Nicht angestellte Dronen haben die Aufgabe, mögliche Futterquellen zu suchen. Es gibt zwei Arten nicht angestellter Dronen. Dronen, die per Zufallssuche neue Bereiche des Parameterraumes explorieren und Dronen, die sich anhand der Informationen der Arbeiterdronen eine bekannte Quelle aussuchen und dort mitarbeiten.

Für den ABC-Algorithmus bedeutet dies nun, daß die Bienenkolonie aus drei Arten von Partikeln besteht. Angestellte Dronen, Springerdronen und Explorationsdronen. Die angestellten Dronen werden sich eine Futterquelle suchen und diese ausbeuten, also in einem Bereich des Parameterraumes eine lokale Suche durchführen. Sie teilen die Information über diese Quelle im Schwarm und die Springerdronen wählen eine Quelle als Arbeitsbereich. Eine angestellte Biene, deren Futterquelle leer ist (kein Konvergenzfortschritt) wird zur Explorationsdrone und sucht sich eine zufällige neue Quelle. Der Pseudocode des ABC sieht damit wie folgt aus.

1. Definiere die Anzahl n von angestellten und springenden Bienen. Der Schwarm hat somit eine Größe von $2n$.
2. Initialisiere die angestellten Bienen bzw. den ihnen zugeordneten Futterquellen zufällig gleichverteilt als Vektoren im Parameterraum.
3. Berechne den Wert einer Futterquelle, also das Qualitätsmaß an jeder Position einer angestellten Biene.
4. Berechne Zuordnungswahrscheinlichkeit P für jede Futterquelle als normierter Qualitätsfunktionswert

$$P_i = \frac{f_{max} - f_i}{\sum_{j=1}^n (f_{max} - f_j)}$$

5. Wähle eine Futterquelle (Position einer angestellten Biene) für jede springende Biene nach den Wahrscheinlichkeiten P .
6. Erzeuge die springende Biene \vec{x}_c aus der entsprechenden angestellten Biene \vec{x}_d der Futterquelle durch Variation in der zufällig gewählten Dimension j und in Richtung einer zufällig gewählten anderen Futterquelle \vec{x}_k .

$$x_{cj}(t) = x_{dj}(t) + (2.0r - 1.0)(x_{dj}(t) - x_{kj}(t))$$

($j=1,\dots,n$ n:Dimension des Parameterraumes).

7. Bestimme das Minimum der Qualitätsfunktionswerte für alle angestellten und springenden Bienen, M_a, M_h , und speichere das beste Bienenmodell \vec{x}_G .
8. Beginne Schleife.
9. Mutiere die angestellten Bienen in der zufällig gewählten Dimension j in Richtung einer zufällig gewählten anderen Futterquelle/Biene k :

$$x_{dj}(t) = x_{dj}(t) + (2.0r - 1.0)(x_{dj}(t) - x_{kj}(t))$$

(Die angestellten Dronen suchen in der Umgebung der Futterquelle.)

10. Wenn der Qualitätsfunktionswert der neuen Position besser ist, ersetze die Futterquelle und ihre angestellte Biene durch die neue.
11. Berechne die Zuordnungswahrscheinlichkeiten der Futterquellen neu.
12. Ordne die springenden Bienen nach den neuen Wahrscheinlichkeiten zu und bewege sie entsprechend (siehe Punkt 4).
13. Prüfe die angestellten Bienen auf eventuelle Konvergenz durch Zählen der Iterationen, in denen die Verbesserung der Qualitätsfunktion

$$df_i = \frac{f_{i_{neu}} - f_{i_{alt}}}{f_{i_{neu}}}$$

kleiner bleibt als eine definierte Konstante. Ist die Anzahl der Iterationen mit geringem Erfolg größer einer Konstante, so wird eine lokale Konvergenz vermutet. Ersetze dann die Futterquelle respektive die angestellte Biene durch eine neue, zufällig gewählte (Die Drone wird zur explorierenden Drone).

14. Evaluiere M_a, M_h und ersetze ggf. das beste Modell. Prüfe globale Abbruchbedingung und gehe zu Punkt 7.

Exploration reguliert sich somit selbst, parallel zu der Konvergenz des Verfahrens an verschiedenen Stellen des Parameterraumes. Der ABC ist somit ein sich selbst neustartender Algorithmus, der damit die globale als auch lokale Konvergenzbedingung erfüllt. Parameter des Verfahrens sind nur die Anzahl der Dronen sowie die Anzahl Iterationen, nach denen eine sich nicht verbessernde Quelle als leer definiert wird.

7.4 Simulated Annealing oder Metropolis Algorithmus

Standard Metropolis Algorithmus

Das Grundkonzept des Simulated Annealing (simulierte Abkühlung)-Algorithmus (SA) (Metropolis et al., 1953; Kirkpatrick et al., 1983) ist es, in den ersten Iterationen global in allen Bereichen gleichwertig zu suchen und dann langsam die lokale Suche zu intensivieren, analog einer „freien“ Bewegung in einem geschmolzenen Material und einer dann gebundenen Bewegung in einem sich durch langsame Abkühlung bildenden Festkörper. Die Basis des SA ist ein einfacher Monte-Carlo bzw. *Hill-Climbing* Suchalgorithmus.

1. Eine Startposition $\vec{x}_t |_{t=0}$ wird zufällig gewählt und die Qualitätsfunktion ausgewertet. Die Position wird gespeichert.
2. Variiere die aktuelle Position zufällig in einer Umgebung

$$\vec{x}_{t+1} = \vec{x}_t + \delta(1 - 2\vec{z})$$

(\vec{z} Zufallsvektor aus $[0,1]^n$; δ Magnitude des Variationsschrittes) und werte die Qualitätsfunktion aus.

3. Ist der Funktionswert der neuen Position besser, so wird die neue Position gespeichert. Gehe zu Punkt 2.

Das Verfahren ist in dieser Form ein lokales Suchmuster, denn es ist nicht zwingend gegeben, daß ein Tal um ein Minimum wieder verlassen werden kann. Größere Funktionswerte werden nicht zugelassen und die Schrittweite des Verfahrens ist durch die Wahl von δ ggf. zu klein, um aus dem Tal in einen noch optimaleren Bereich zu springen. Der SA gibt nun dem Verfahren die Möglichkeit größere Funktionswerte mit einer bestimmten Wahrscheinlichkeit anzunehmen. Diese Wahrscheinlichkeit sinkt mit fortlaufender Iteration. So ist die Bewegung des Suchvektors anfänglich freier und konzentriert sich auf einen Bereich zu späteren Iterationen. Die Wahrscheinlichkeit für die Annahme eines schlechteren Qualitätswertes ist definiert durch

$$P = e^{-\frac{\Delta f}{T}}$$

wobei Δf der Unterschied der Qualitätswerte und T das Analogon zur Temperatur des Systems ist. Die Temperatur regelt die Abkühlung, indem sie iterationsabhängig verkleinert wird. Die ideale Abkühlungsrate zu bestimmen ist problemabhängig und sollte empirisch angegangen werden. Eine in verschiedenen Arbeiten als hinreichend angegebene Abkühlvorschrift ist $T(t) = \frac{T_0}{\ln(t)}$. Im SA wird nun der Punkt drei des obigen Algorithmus ersetzt durch die Vorschrift: Ist der neue qualitätswert optimaler, so akzeptiere die neue Position. Sonst wähle zufällig eine Zahl aus $[0,1]$. Ist diese Zahl größer als $P(t) = e^{-\frac{\Delta f}{T(t)}}$, so akzeptiere die neue Position.

Basin Hoppin Simulated Annealing

Um die Effizienz der Suche des SA-Algorithmus auf topographisch komplexen bzw. multimodalen Qualitätsfunktionen zu verbessern, beschreiben Wales & Doyle (1997) die Anwendung einer hybriden Variante aus Metropolis/SA-Algorithmus und Gradientenverfahren, das sogenannte *Basin-Hoppin*. Idee des Verfahrens ist die Reduktion der Qualitätsfunktion auf eine Stufenfunktion, auf welcher der SA seine Suche durchführt. So wird an jedem Punkt, den der SA anspricht eine lokale Suche mit einem Gradientenverfahren durchgeführt, so daß sich der ausgewählte Punkte zum nächstgelegenen lokalen Minimum verschieben läßt. Die neue Qualitätsfunktion ist also

$$\tilde{f}(\vec{x}) = \min(f(\vec{x}))$$

wobei sich der Operator *min* auf die Durchführung eines Gradientenverfahrens bezieht. Abbildung 7.6 illustriert dieses Vorgehen. Ausgehend vom vorliegenden

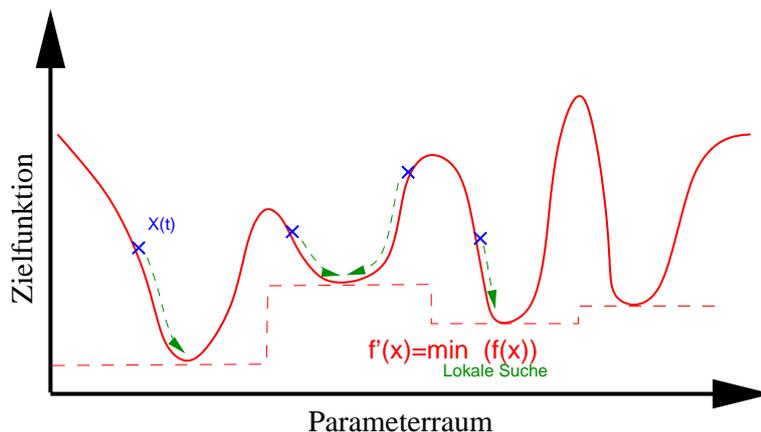


Abbildung 7.6: Den Punkten der globalen SA-Suche (blaue Kreuze) wird durch eine lokale Suche das nächstgelegene Minimum zugeordnet (grüne Pfeile). Die Qualitätsfunktion transformiert sich so zu einer Treppenfunktion.

Inversionsproblem ist der Einsatz von Gradientenverfahren aufgrund numerischer Instabilitäten nicht ratsam. Es wurde daher in dieser Arbeit ein Hybrid aus SA und Downhill-Simplex getestet und in den Effizienzvergleich der Methoden mit einbezogen. Es kann zwar nicht davon ausgegangen werden, daß das Simplexverfahren zu jedem lokalen Minimum konvergiert, jedoch wird sich für jeden Metropolispunkt eine Verbesserung erreichen lassen.

7.5 Kovarianzmatrix-Adaptions-Evolutionsstrategie (CMA-ES)

Evolutionsstrategische Optimierungsverfahren werden bereits in unterschiedlicher Form (Genetische Algorithmen (Goldberg, 1989), $(\lambda+\mu)$ Evolutionsstrategie, Evolutionäres Programmieren, Derandomisierte Evolutionsstrategie (CMA-ES)) in sehr vielen Bereichen der Geophysik eingesetzt (z.B. Dal Moro et al., 2007; Dal Moro & Pipan, 2007; Alvers, 1998; Baggett & Skahill, 2008). Eine sehr effektive Variante der Evolutionsstrategie, welche Informationen über vormals erfolgreiche Iterationsschritte in der Bestimmung der Optimalschrittweite verwendet, ist die Kovarianzmatrix-Adaptions-Evolutionsstrategie (CMA-ES) (Hansen & Ostermeier, 2001). Dieses Verfahren soll in diesem Abschnitt als weiteres Vergleichsverfahren vorgestellt werden. Das Verfahren wurde implementiert und erfolgreich auf das Scholtewelleninversionsproblem angewendet.

Evolutionsstrategie

Die Evolutionsstrategie (EA: Evolutionäre Algorithmen) (Rechenberg, 1994) verfolgt die einfache Strategie durch Rekombination und Variation von Parametern und letztendlich durch Selektion von Modellen eine Modelloptimierung zu erreichen, ähnlich der biologisch evolutionären Entwicklung, welche effektive Adaption und Anpassung von Lebewesen an ihr Umfeld verspricht. Die EA imitieren die evolutionäre Entwicklung in verschiedener Weise, jedoch immer mit dem folgenden Grundgerüst.

1. Die **Initialisierung** einer Population von Parametervektoren bzw. Modellen
2. Die **Selektion** wählt aus der Population eine definierte Zahl, μ , von Elternmodellen aus.
3. Durch Auswahl oder **Rekombination** werden λ Nachkommen aus den μ Elternmodellen gewonnen. Es werden entweder λ Kopien von zufällig ausgewählten Elternmodellen erzeugt, oder durch Rekombination der Eigenschaften mehrerer Elternmodelle neue Modelle erschaffen.

4. Eine darauf folgende **Mutation** verändert die Nachkommen und entspricht dem Anteil der Zufallssuche. Die Modelle \vec{x} werden hier in einfachster Form verändert mit

$$\vec{x} = \vec{x} + (2\vec{z} - 1)\delta$$

wobei \vec{z} ein Vektor normalverteilter Zufallszahlen aus dem Intervall $[0,1]$ und δ die Streubreite oder Schrittweite der Suche, welche der Hauptparameter der Evolutionsstrategie ist.

5. Mit der Qualitätsfunktion wird nun der Wert jedes einzelnen Modells bewertet und die **Auswahl** einer neuen Population vorgenommen. Die μ besten („fittesten“) Modelle werden als neue Generation Elternmodelle gewählt. Hierbei kann bspw. aus den Nachkommen gewählt werden (λ , μ) oder aus Nachkommen und voriger Elternpopulation ($\lambda + \mu$).

Es gibt verschiedene Varianten der EA, welche sich in erster Linie in Art der Selektion und der Auswahlmenge für die Elternpopulation unterscheiden. Hierzu zählen

- $(\mu + \lambda)$: Ein Population umfaßt μ Individuen. Aus μ Elternmodellen werden λ Nachkommen generiert. Die neue Population wird aus Eltern- und Nachkommenmodellen nach der Fitneß/Qualitätsmaß gewählt ($\mu + \lambda$).
- (μ, λ) : Ein Population umfaßt μ Individuen. Aus den μ Elternmodellen werden λ ($\lambda > \mu$) Nachkommen generiert. Die neue Population wird aus den Nachkommenmodellen nach der Fitneß/Qualitätsmaß gewählt.
- $(1+1)$: Jede Population umfaßt nur ein Modell, welches mutiert wird. Dieses Verfahren entspricht dem einfachen Monte-Carlo-Algorithmus.
- $(\mu + 1)$: Eine Population umfaßt μ Modelle. Eines wird als Elternmodell gewählt und mutiert, um ein Nachkommenmodell zu generieren. Dieses ersetzt dann das im Qualitätsmaß schlechteste Modell.
- $(\mu/\eta, \lambda)$: Eine Population umfaßt μ Modelle. η definiert die Anzahl Elternmodelle, die zur Rekombination genutzt werden, um λ Nachkommenmodelle zu erzeugen.

Kovarianzmatrix-Adaptions-Evolutionsstrategie

Der Hauptoperator der Evolutionsstrategie ist die Mutation. Hat man eine gute Mutationsstrategie und eine gute Schrittweitenanpassung, so ist der Erfolg am wahrscheinlichsten. Im Hinblick auf dieses Problem stellten Hansen & Ostermeier (2001) die sogenannte Kovarianzmatrix-Adaptions-Evolutionsstrategie (Covariance Matrix Adaption-Evolution Strategy oder CMA-ES) vor. Das Grundkonzept dieses Verfahrens ist eine adaptive Anpassung der Schrittweite unter Benut-

7 ANGEWENDETE OPTIMIERUNGSMETHODEN

zung der Kovarianzmatrix der Mutationsverteilung bewertet mit der Qualitätsfunktion. Die Kovarianzmatrix beschreibt die paarweisen Abhängigkeiten zwischen den Variablen (siehe auch Kapitel 8). Vergleichbar mit Quasi-Newton Verfahren (siehe Abschnitt 4.2) approximiert die CMA-ES iterativ ein Analogon zur inversen Hesse-Matrix. Hierzu wird eine Kovarianzanalyse der zuvor gewählten Mutation im Hinblick auf die Qualitätswerte durchgeführt, um die neue Mutationsstreuung zu definieren. Die aktuelle Mutationsverteilung wird derart geändert, daß die Wahrscheinlichkeit, einen bereits getätigten guten Mutationsschritt zu wiederholen, gesteigert wird. Die Auswahl eines guten Mutationsschrittes erfolgt über eine gewichtete Summation von Generationsschritten, den Differenzen zweier jeweils aufeinanderfolgender Populationsschwerpunkte.

Die CMA-ES nutzt keine Gradienten und auch keine Differenzenquotienten, was sie zu einem guten Verfahren zur Anwendung auf multimodale und auch nichtstetige Probleme macht. Das Verfahren zeigt darüber hinaus nicht nur gutes Konvergenzverhalten bei lokaler Suche, sondern auch, daß es zur globalen Suche anwendbar ist (Hansen & Ostermeier, 2001). Es läßt sich daher vermuten, daß das Verfahren ebenfalls gut für das Inversionsproblem dieser Arbeit geeignet ist. Für eine detaillierte Beschreibung des Verfahrens sei auf Hansen & Ostermeier (2001) verwiesen.

Es existieren darüber hinaus verschiedene Varianten der CMA-EA, welche sich in der Wahl der Parameter λ und μ , aber auch in der Rekombination der Modellvektoren unterscheiden. In dieser Arbeit wurde die von Hansen & Ostermeier (2001) vorgestellte (μ_w, λ) -Variante gewählt und implementiert. Die Aufdatierungsgleichung dieses Verfahrens für die Parametervektoren \vec{x}_i ist in Kürze

$$\vec{x}_i(t+1) = \langle \vec{x} \rangle_w + \sigma(t) \cdot B(t) \cdot D(t) \cdot \vec{z}_i(t+1)$$

wobei

-

$$\langle \vec{x} \rangle_w(t) = \frac{1}{\sum_{i=1}^{\mu} \Omega_i} \sum_{i=1}^{\mu} \Omega_i \vec{x}_{iB}(t)$$

dem mit vordefinierten Gewichten Ω_i gewichteten, mittleren Modellvektor aus den μ besten Individuen der Generation t .

- $\sigma(t)$ die Schrittweite in Generation t .
- $\vec{z}_i(t+1)$ ein Vektor aus normalverteilten Zufallszahlen aus dem Intervall $(0,1)$.
- Die Matrizen $B(t)$ und $D(t)$ sind Bestandteile einer Singulärwertzerlegung der Kovarianzmatrix $C(t)$,

$$C(t) = B(t)D(t)(B(t)D(t))^T = B(t)(D(t))^2(B(t))^T.$$

Die Kovarianzmatrix und $\sigma(t)$ sind hierbei abhängig von der gewichteten Summation der Generationsschritte (gewichtete Differenzen der Punkte $\langle \vec{x} \rangle_w$).

Es bezieht sich der index w in der Bezeichnung (μ_w, λ) -CMA-ES auf das gewichtete Mittel (gewichtete Rekombination) der μ besten Modelle aus der Population von λ Modellen, welches Ausgangspunkt der neuen Generation ist. Hansen & Ostermeier (2001) schlagen für die Wahl der Parameter λ und μ folgenden Bedingungen vor,

$$\lambda = 4 + 3\log(n) \quad \mu = \lambda/2,$$

wobei n die Dimension des Problems ist. Allgemein wird jedoch empfohlen, daß $\mu \leq \lambda/2$. Um für das Verfahren stärkere Exploration zu erreichen, wird eine Erhöhung von λ bei multimodalen Problemen empfohlen. Es wurden daher für das vorliegende Inversionsproblem verschiedene Werte $\lambda \geq 4 + 3\log(n)$ probiert.

7.5.1 PSO-CMAES-Hybrid

Analog zu Abschnitt 7.2 wurde ebenfalls ein Hybrid aus oben vorgestellte CMA-ES und einem Reset-PSO implementiert und am Problem getestet. Nach Hansen & Ostermeier (2001) ist der CMA-ES ein effektives lokal suchendes Verfahren, welches im Falle multimodaler oder nichtstetiger Zielfunktionen erfolgreich eingesetzt wurde. Die Implementierung als lokale Suche nach jedem Iterationsschritt der PSO für das global beste Partikel ist somit sachgemäß.

7 ANGEWENDETE OPTIMIERUNGSMETHODEN

8 Methodenvergleich und Auflösungsanalyse

Um ein Effizienzmaß für einen Inversionsalgorithmus zu quantifizieren, seien hier drei Fragen angegeben, die es zu beantworten gilt.

1. Wie konvergiert das verwendete Optimierungsverfahren? Hierbei sind zum einen Konvergenzgeschwindigkeit und Laufzeit, aber auch die Wahrscheinlichkeit das globale Optimum zu finden, gemeint. Zusammengefaßt läßt sich fragen, wie gut ist die Auflösung des globalen Optimums bei idealen Datenbedingungen und bei bestimmter Laufzeit. Dies sei die „algorithmische“ Auflösung.
2. Wie ist der Einfluß von Rauschen auf das Inversionsergebnis?
3. Wie groß ist der Einfluß von Korrelationen zwischen Parametern oder „Trade-Off“?

So muß die Verlässlichkeit eines Inversionsergebnisses also im allgemeinen anhand zweier Größen untersucht werden. Zum einen gilt es, die Unsicherheiten der Lösung zu quantifizieren und zum anderen ist es notwendig, die interparameträren Zusammenhänge, den „Trade-Off“, verifizieren zu können. „Trade-Off“ Zusammenhänge bestehen im Fall von unterbestimmten Problemen, bei denen Beziehungen zwischen verschiedenen Parametern nicht aufgelöst werden. Bohlen et al. (2004) untersuchten diese Zusammenhänge, wie in Abschnitt 4.2 beschrieben, mit Hilfe multipler Inversionsdurchgänge und variierten Startmodellen. Für stochastische Verfahren ist ein solches Vorgehen nicht möglich, da kein Startmodell zu Verfügung steht, sondern eine Startpopulation mehrerer Modelle. Eine Evaluation der Verlässlichkeit der Lösung einer Inversion kann unter anderem auch durch wiederholtes Durchführen der Inversion mit unterschiedlichen Startpopulationen erreicht werden. Hierbei wird das Verfahren auch auf eventuelle Abhängigkeiten der Lösung von der Startpopulation untersucht. So können ebenfalls ein mittleres Modell sowie Variationen berechnet werden. Dieses Vorgehen bedarf eines großen Rechenaufwandes und untersucht in erster Linie nicht die Einflüsse eventueller Parameterkorrelationen.

Es existiert jedoch eine große Menge vorwärts modellierter Modelle mit ausgewertetem Qualitätswert, die zur Bestimmung von Kovarianzen und einem mittleren Modell genutzt werden können. So können, wie in Abschnitt 8.4 vorgestellt

wird, Zustandswahrscheinlichkeiten der Modellensembles für die Zustandsgröße der Qualitätsfunktion approximativ für den Nahbereich einer gefundenen Lösung bestimmt werden.

Die Wahrscheinlichkeiten können dann zur Berechnung qualitätsmaßgewichteter Kovarianzen der Parameter sowie eines Erwartungsmodells genutzt werden.

Dieses Kapitel beschäftigt sich mit der Effizienzanalyse der vorgestellten Verfahren und einem Vergleich der algorithmischen Auflösung der Verfahren anhand mehrfacher Inversion synthetischer Daten und anschließender Bestimmung eines mittleren Modells und Varianzen. Ebenso wird eine statistische Auflösungsanalyse vorgestellt zur Beurteilung der Inversion gemessener Datensätze.

8.1 Das Testproblem

Um die Frage zu beantworten, wie effizient ein Optimierungsverfahren arbeitet und wie nah es einem globalen Optimum kommt, sei das Verfahren auf ein Problem angewendet, bei dem das globale Optimum bekannt ist. Es bietet sich daher an, ein Referenz-Phasenlangsamkeits-Frequenz-Spektrum nach bekanntem Modell zu modellieren und es als Datengrundlage für eine Inversion zu nutzen. Hierzu wird ein Testmodell gewählt, welches dem mittleren Ergebnis der Baugrundvoruntersuchung am FINO3 Standort (Wilken et al., 2009) entspricht. Die Parametrisierung des Modells entspricht dem in Abschnitt 4.2 vorgestellten Konzept. Es werden Schichtmächtigkeiten und eine obere sowie untere Scherwellengeschwindigkeit je Schicht als Inversionsparameter definiert. Ein Partikel der PSO ist also definiert durch

$$\vec{X} = (\vec{h}, \vec{v}_s)^T$$

mit $\vec{h} \in D_h \subset \mathbb{R}^n$ und $\vec{v}_s \in D_{v_s} \subset \mathbb{R}^n$ (D : Definitionsintervalle, n : Anzahl Schichten).

Zur Untersuchung der Effizienz bei Problemen höherer Dimensionalität wird nicht nur anhand des FINO3 Modells mit fünf Schichten, sondern auch an einem Sechs- und Siebenschichtfall getestet. In Abschnitt 4.2 wurde angeführt, daß die Sensitivität der Inversion für Dichte und Kompressionswellengeschwindigkeit klein ist gegenüber der Sensitivität für die Scherwellengeschwindigkeit. Dichte, Kompressionswellengeschwindigkeit aber auch Mächtigkeit der Wassersäule sind daher für das Testproblem und während der Inversion als konstant gewählt.

Eine Inversion der Testspektren wird dann für jedes zu untersuchende Verfahren und jede zu untersuchende Parameterkonstellation jeweils mit 56 unterschiedlichen Startpopulationen durchgeführt. Aus den 56 Inversionen konnte so ein mittleres Modell und Abweichungen berechnet und mit der bekannten Lösung verglichen werden.

Referenzmodelle

Es werden im folgenden drei Modelle vorgestellt, mit welchen Testdatensätze modelliert werden sollen. Das erste Modell sei als FINO3 Referenzmodell bezeichnet. Das Modell wurde als mittleres Modell einer bereits durchgeführten Inversion am Standort FINO3 erhalten, und wird als angenommen geologisch sinnvolles, rauschfreies Testdatenbeispiel gewählt. Das Modell enthält fünf Schichten, wobei die erste als Wassersäule mit $v_s = 0.0 \text{ m/s}$, $v_p = 1500 \text{ m/s}$, $\rho = 1.0 \text{ t/m}^3$ und Mächtigkeit $h = 23 \text{ m}$ konstant ist. Es folgen drei Schichten und der Halbraum. Zwei weitere Modelle mit höherer Dimensionalität des Parameterraumes durch sechs und sieben Schichten wurden ausgehend vom FINO3 Modell definiert. Die drei Testmodelle sind in Abbildung 8.1 dargestellt (rote Linien). Darüber hinaus zeigt die Abbildung bereits den Parameterraum der Inversion, also die Intervallbeschränkungen der Parameter. Hierbei ist jeweils das Intervall der Geschwindigkeiten mit der Tiefe aufgetragen. Die Tiefe eines Blockes bewegt sich zwischen der Summe der über dem Block liegenden minimalen Schichtmächtigkeiten bis zur Summe der maximalen Schichtmächtigkeiten. Eine Schichtmächtigkeit von 1 m bis 10 m in der ersten und zweiten Sedimentschicht bedingt so eine mögliche Tiefe von 2 m bis 20 m unter dem Meeresboden für Geschwindigkeiten in der zweiten Schicht. Die so entstehende Verschachtelung der Intervalle ist in Abbildung 8.1 als verschieden grau eingefärbte Rechtecke dargestellt. Bei einer Rastersuche mit einer Auflösung von 5 m/s in der Scherwellengeschwindigkeit und 1 m in den Schichtmächtigkeiten würden so nur für den fünf Schichten Fall $4.0 \cdot 10^{16}$ mögliche Modellen ausgewertet werden. Die Tabelle 8.1 gibt ebenfalls die Parameterbereiche für Scherwellengeschwindigkeit und Schichtmächtigkeiten an, die der Anpassung der Modelle als Suchbereich zu Grunde liegen. Die für die Voruntersuchungen mit modellierten Spektren genutzte Dichteverteilung ist konstant während der Inversion und entspricht der Verteilung des Referenzmodells. Diese ist nach der gemessenen Tiefenverteilung der Dichte eines am Baugrund FINO3 gezogenen Bohrkerns gewählt worden. Ergebnis der Bohrung waren dicht bis sehr dicht gelagerte Sande ab einer Tiefe von 6 m unter dem Meeresboden mit einer Dichte von 1.92 t/m^3 in 6 m Tiefe bis 2.16 t/m^3 in einer Tiefe von 29 m. Die Dichtewerte für die erweiterten Sechs- und Siebenschichtfälle wurden linear extrapoliert. Die Kompressionswellengeschwindigkeit wurde ebenfalls als konstant definiert und wird von 1500 m/s in der Wassersäule und 1550 m/s am Meeresboden bis 1850 m/s im Halbraum linear interpoliert und für die erweiterten Modelle extrapoliert.

Abbildung 8.2 zeigt die erhaltenen Referenzdatensätze, die zur Inversion genutzt werden. Die Abbildung zeigt die identifizierbaren Picks der Dispersionskurven der drei Testmodelle. Es werden also Punkte der Fundamentalmode und der ersten höheren Mode zur Inversion genutzt. Die Frequenz- und Langsamkeitsauflösung der Spektren beträgt $df=1.0 \text{ Hz}$ und $dp=0.05 \text{ s/km}$, um die Rechenzeiten der Vorwärtsmodellierung möglichst gering zu halten, aber auch genügend Datenpunkte

8 METHODENVERGLEICH UND AUFLÖSUNG

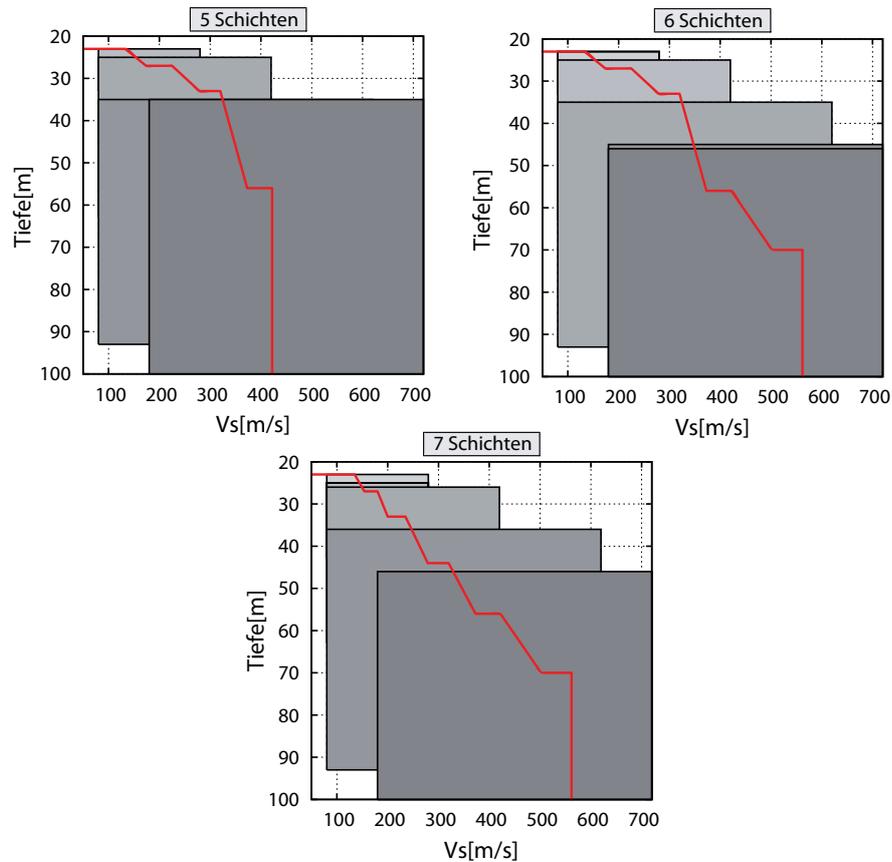


Abbildung 8.1: Darstellung der drei verwendeten Testmodelle in Scherwellengeschwindigkeit mit der Tiefe (rote Linien). Darstellung der Intervallbeschränkungen als Parameterraum der Inversion. Es ist jeweils das Intervall der Geschwindigkeiten von der Summe der darüber liegenden minimalen Schichtmächtigkeiten bis zur Summe der maximalen Schichtmächtigkeiten aufgetragen (graue Rechtecke).

zur Verfügung zu stellen. Die Frequenzauflösung ist hierbei der dominierende Parameter im Bezug auf gebrauchte Rechenzeit.

Grundsätzliche Voreinstellungen

Im folgenden sollen zum einen die zu vergleichenden Optimierungsverfahren, aber zum anderen auch ihre globalen Einstellungen und Parameter benannt und diskutiert werden. Zu diesen Parametern zählen die Gewichtungsfaktoren der Verschiebungskomponenten der PSO, Wirkungsradien für Abstoßungsprozesse, Suchbereiche für lokale Verfahren oder auch die Anzahl der auszuführenden Iterationsschritte. Grundsätzlich gilt für die Suche nach optimaler Parametereinstellung, daß eine Rastersuche für alle Zahlenkombination mehrerer Parameter in den mei-

FINO3 Modell		7 Schicht Modell	
$v_s \pm \Delta v_s$ [m/s]	h [m]	$v_s \pm \Delta v_s$ [m/s]	h [m]
0.0 ± 0.0	23.0	0.0 ± 0.0	23.0
180.0 ± 100.0	1.0→10.0	180.0 ± 100.0	1.0→10.0
250.0 ± 170.0	1.0→20.0	180.0 ± 100.0	1.0→10.0
350.0 ± 270.0	10.0→30.0	250.0 ± 170.0	1.0→20.0
450.0 ± 270.0	Halbraum	350.0 ± 270.0	10.0→30.0
		450.0 ± 270.0	10.0→30.0
		450.0 ± 270.0	Halbraum
6 Schicht Modell			
$v_s \pm \Delta v_s$ [m/s]	h [m]		
0.0 ± 0.0	23.0		
180.0 ± 100.0	1.0→10.0		
250.0 ± 170.0	1.0→20.0		
350.0 ± 270.0	10.0→30.0		
450.0 ± 270.0	10.0→30.0		
450.0 ± 270.0	Halbraum		

Tabelle 8.1: *Parameterraumintervalle der Testmodelle für Scherwellengeschwindigkeit, v_s , und Schichtmächtigkeit, h .*

sten Fällen zu rechenintensiv ist. Sollen Auflösung und Inversionsergebnis als Maß für die Güte der Parameter gelten, so wäre die Durchführung mehrerer Inversionen des Testmodells für jeden Satz von Parametern nötig. Eine Betrachtung des Konvergenzverhaltens für eine einzelne Inversion eines Testmodells mit jeweils gleichen Startbedingungen oder das Zurückgreifen auf publizierte numerisch empirische Studien an multimodalen Testfunktionen sind demnach das plausible und auch hinreichende Vorgehen.

- Es wurden die in Tabelle 8.2 aufgeführten **Varianten der PSO und Verfahren** zur Inversion der drei Testmodelle genutzt und untereinander verglichen.
- **Neustarts und Iterationen:** Es ist nicht zu erwarten, und die Ergebnisse der Voruntersuchungen stützen diese Annahme, daß ein einfaches Optimierungsverfahren ohne die Möglichkeit des Neustarts bei lokaler Konvergenz das globale Optimum der Anpassung erreicht. Die in Kapitel 5 untersuchte Multimodalität läßt annehmen, daß bei der hohen Zahl von Nebenminima eine große Streuung der Lösungen bei nur einmaligem Einsatz der Optimierungstrategien eintritt. Es wurde daher jedes verwendete Verfahren mit der Möglichkeit des Neustarts bei Konvergenz implementiert. Der Neustart

8 METHODENVERGLEICH UND AUFLÖSUNG

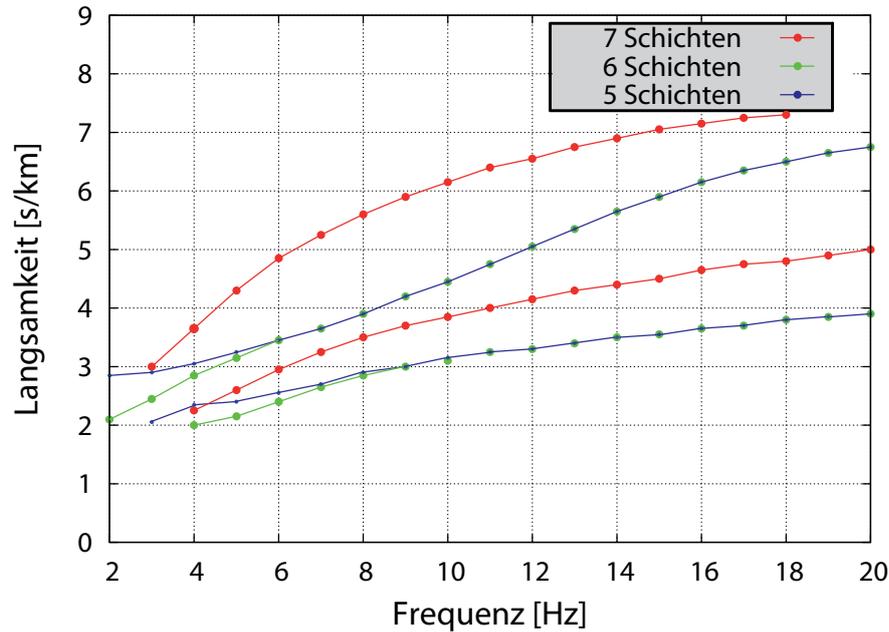


Abbildung 8.2: Angerissene Dispersionskurven für Fundamentalmode und erste höhere Mode der drei Testmodelle.

eines Verfahrens erfolgt im Falle zu kleinen Konvergenzfortschrittes von

$$df = \frac{f(t) - f(t-1)}{f(t)} \leq 6.0e^{-3}$$

über 20 Iterationen.

Das Ende des globalen Inversionsvorganges wird nicht durch das Konvergenzmaß ermittelt, sondern mit einer maximalen Anzahl von Iterationsschritten festgelegt. Die Zahl maximaler Schritte sei hierbei so groß, daß eine „genügend“ große Anzahl von Neustarts erfolgt. Wichtig für die Wahl der Iterationszahl ist der Punkt, an dem der bislang beste global gefundene Wert der Qualitätsfunktion eine Sättigung erreicht, d.h. daß kein besseres als das bislang gefundene vermeintlich globale Minimum durch weitere Neustarts erreicht wurde. Die Anzahl der Schritte wurde so zunächst auf 1000 Iterationen gesetzt. Aus den Lösungen der einzelnen Neustarts wird dann Anhand der Qualitätsfunktion das beste Modell als Inversionslösung gewählt. Die Voruntersuchungen werden des weiteren zeigen, daß sich in fast allen Fällen bereits vor 1000 Iterationsschritten eine Sättigung des besten Anpassungswertes einstellt.

- **Lokale Konvergenz:** In Abschnitt 7.2 wurde bereits die mangelnde Fähigkeit der PSO zur lokalen Konvergenz erläutert und ein Verfahren vorgestellt,

PSO	„Global Beste“ P artikel- S chwarm- O ptimierung
DHS-PSO	Hybridverfahren aus D ownhill- S implex(DHS) und PSO
cDHS-PSO	Hybridverfahren aus D ownhill- S implex(DHS) und geladener PSO (ch arged PSO)
rDHS-PSO	Hybridverfahren aus D ownhill- S implex(DHS) und rasternder PSO (rep elling PSO)
(c+r)DHS-PSO	Hybridverfahren aus D ownhill- S implex(DHS) und geladener (ch arged) und rasternder PSO
CMA-ES	K ovarianz m atrix- A daptions- E volutionsstrategie
CMAES-PSO	Hybridverfahren aus CMA-ES und PSO
BH-SA	Hybrid aus Downhill-Simplex und S imulierter A bkühlung
ABC	Künstlicher Bienenschwarm/ (A rtificial B ee C olony)-Verfahren

Tabelle 8.2: *Verwendete Verfahren und Verfahrensvarianten.*

das bei diesem Problem durch Einführen einer stochastischen lokalen Suche des besten Partikels Abhilfe schafft (GCP SO: Guaranteed Convergence PSO). Diese Ergänzung der PSO wird in den folgenden Untersuchungen nur in der einfachen PSO implementiert sein. Die Variante des DHS-PSO enthält mit dem Downhill-Simplex bereits ein lokal suchendes Verfahren.

- **Schwarmgewichtungsparameter:** Die Einstellung der Schwarmgewichtungsparameter, c_1 , c_2 und ω definiert den Kompromiß aus Exploration und Konvergenz des Schwarms und ist daher von essentieller Bedeutung. Mishra (2006) empfiehlt eine Wahl der Parameter c_1 und c_2 nahe 1.0. Grundsätzlich ist im Falle des hier betrachteten, multimodalen Inversionsproblems $c_1 > c_2$ zu wählen, um hohe Exploration zu erwirken. Die Partikel sind dann stärker von ihrem kognitiven Anteil beeinflusst. Angenommen die Parameter werden nach Mishra (2006) aus dem Intervall $(0,2) \subset \mathbb{R}$ gewählt. Eine Rastersuche in Schritten zu 0.1 ergäbe mit drei Parametern 2000 Rechnungen. Dies ist zu rechenaufwändig im Falle der Inversion. Daher wird diese Frage mit Hilfe eines Testproblems mit geringerem Rechenaufwand untersucht, welches das Potential zwischen ungeladenen, nicht chemisch gebundenen Atomen in einem nicht-idealen Gas annähert. In großer Entfernung überwiegen hierbei anziehenden Kräfte (Van-der-Waals Kräfte und Dipol-Dipol Kräfte) und ab einer bestimmten Entfernung überwiegt der abstoßende Anteil. Dieses „Lennard-Jones“ Potential bestimmt sich mit

$$P_{LJ} = 4\epsilon \sum_{i \neq j} \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right]$$

wobei r_{ij} der Abstand zweier Partikel ist, ϵ die Tiefe des Potentialtop-

8 METHODENVERGLEICH UND AUFLÖSUNG

fes und σ der Teilchendurchmesser (Wales & Doye, 1997). Gesucht sind minimale Zustandsenergien als Gleichgewichtslösungen für Teilchencluster. Das Problem ist vergleichbar hochgradig multimodal. Für die Analyse der Gewichtungparameterkonstellationen wurden 6 Teilchen, also eine Dimensionalität von 18 mit der einfachen global besten PSO implementiert. Die Gewichtungparameter wurden in Schritten zu 0.1 zwischen 0.1 und 2.0 durchgezählt. Die Optimierung wurde jeweils 10 Mal mit unterschiedlicher Startpopulation von 20 Partikeln zu 2000 Iterationen durchgeführt. Es wurde die beste Lösung der 10 Durchläufe als einem Tupel (c_1, c_2, ω) zugehörig gewählt. Die Auswertung dieser Rasteranalyse ließ auf eine klare Verdichtung guter Lösungen um einen Bereich im (c_1, c_2, ω) -Raum hoffen. Das Ergebnis läßt jedoch leider keine solch eindeutige Aussage zu. Es sind aber einige Tendenzen zu erkennen. Wie zu erwarten war, wird die Menge guter Lösungen größer, je größer c_1 gewählt war. Hier muß also der explorative Anteil forciert werden. Für konstantes c_1 zeigen c_2 und ω bessere Werte im mittleren Bereich des Intervalls (0,2). Eine gute relative Konstellation wäre demnach $c_2 \simeq \omega = 2 \cdot c_1$. Für die Inversion wurden so letztendlich die Werte $c_1 = 0.9$, $c_2 = 0.5$ und $\omega = 0.5$. Diese Wahl legt hohes Gewicht auf Exploration durch $c_1 > c_2$ und hohe Anpassungsfähigkeit des Schwarms durch kleines ω .

- **Parametereinschränkungen:** Die Intervallbeschränkungen der Parameter, welche den Suchraum definieren, werden zu jeder Iteration geprüft. Bewegt sich ein Partikel außerhalb dieser Bereiche, so wird der fehlerhafte Parameter neu gewählt.
- **Modellinterpolation:** Die Methode von Wang (siehe Abschnitt 2.1) interpoliert Gradienten in den Modellen in kleineren Subschichten, so daß keine Entartung nach mittleren Geschwindigkeiten in einer Gradientenschicht entsteht. Schichten mit einem konstanten Gradienten werden so in homogenen Schichten diskretisiert. Die Mächtigkeiten der Subschichten werden nach vorab definierten Auflösungen in den Parametern gewählt. Ist diese Zahl kleiner als 1% der charakteristischen Wellenlänge, so wird diese als Mächtigkeit verwendet. Dies ist im flachmarinen Fall, wo die Modelle im Mittel eine Tiefe von 40 m haben, zu groß gewählt. Die Modelle zur Vorwärtsrechnung werden daher bereits vorher in 2 m Schritten interpoliert. Die Inversion kontrolliert so selbst die Interpolation. Dies führt jedoch aufgrund der höheren Schichtzahl in der Vorwärtsrechnung zu höheren Rechenzeiten.
- **Strafterme:** Wie schon in Abschnitt 6.3 beschrieben, werden negative Geschwindigkeitsgradienten in einem Modell durch Addition eines Strafwertes auf die Zielfunktion verhindert, denn eine Selektion nur positiver Gradienten würde die Gleichverteilung der Modelle zerstören. Es wird hierbei

jede einzelne Schichtgeschwindigkeit betrachtet. Für jede Schichtgeschwindigkeit, tiefer als die betrachtete, die einen kleineren Wert aufweist, wird eine Konstante $\delta_p=1.0$ auf das Qualitätsmaß addiert. Die folgenden Untersuchungen beinhalten ebenfalls einen Vergleich von Inversionsergebnissen mit und ohne diese Modifikation. Es konnte gezeigt werden, daß der einfache konstante Strafwert die Auflösung durch die Eingrenzung auf positive Gradienten signifikant verbessert.

- **Zusätzliche Parameter der rPSO und cPSO:** Für die „geladene, atomare“ cPSO bedarf es der Definition von Wirkungsradius R_w , Kernradius R_k und Ladung Q der Partikel. Blackwell & Bentley (2002) schlugen $R_k = 1.0$, $R_w = \sqrt{3}x_{max}$ (x_{max} =größtmöglicher Wert eines Parameters) und $Q=16.0$ für verschiedene Testprobleme vor. Das Studium des Konvergenzverhaltens der Inversion des FINO3 Testmodells hat in Voruntersuchungen diese Zahlen weitestgehend bestätigt, so daß $R_K = 1.0$, $R_w=3.8$ und $Q=16.0$ eine effiziente Wahl darstellen. R_w muß hierbei den Dimensionen des Parameterraumes angepaßt werden.

Für die „abstoßende“ rPSO bedarf es der Definition von Wirkungsradius d_{ij} und Abstoßungsamplitude p_{ij} . Diese wurden ebenfalls in Voruntersuchungen anhand von mehreren Inversionsdurchläufen und dem Vergleich des Konvergenzverhaltens zu $d_{ij}=40.0$ und $p_{ij}=0.1$ gewählt.

- **Initialer Suchbereich für lokale Verfahren:** Der Suchbereich, also die Verteilungsbreite δ des zufälligen Suchschrittes $\vec{x} = \vec{x} + (2\vec{z} - 1)\delta$ der SA-Optimierung und die Ausdehnung des Downhill-Simplex sei zu $\delta_{v_s}=5$ m/s und $\delta_h=2$ m gewählt.

Für das Downhill-Simplex Verfahren als Teil des DHS-PSO sind des weiteren 100 Iterationen für die lokale Suche des global besten Partikels vorgesehen. Im BH-SA Verfahren sind es 20.

- **Scouts des ABC:** Eine angestellte Biene wird im Falle eines Konvergenzfortschritts von

$$df = \frac{f(t) - f(t-1)}{f(t)} \leq 6.0e^{-3}$$

über 100 Iterationen neu initialisiert und wird damit zur suchenden Biene. Die von Karaboga & Basturk (2007) empfohlene Zahl von Produkt aus Dimensionalität des Problems und Anzahl Bienen entspräche 2000 Schritten. Dies hat jedoch zu geringe globale Exploration zur Folge und wurde daher reduziert.

8.2 Durchführung der Tests

Zur Dimensionierung der verwendeten Verfahren im Hinblick auf die Populationsgröße und zur Analyse der algorithmischen Auflösung bei zunächst rauschfreien Daten sowie dem Vergleich der Verfahren, werden die implementierten Inversionsalgorithmen jeweils mit unterschiedlichen Populationsgrößen und Konstellationen der Varianten auf den oben vorgestellten Testdatensatz angewendet. Hierzu werden für jeden Test 56 Rechnungen mit verschiedenen Startpopulationen durchgeführt. Nicht nur der Vergleich des resultierenden mittleren Modells mit dem bekannten Testmodell, sondern auch die Standardabweichungen der Parameter der 56 Lösungen sollen eine Evaluation der Verlässlichkeit der Lösung der jeweiligen Inversion ermöglichen.

Eine einzelne Inversion mit 1000 Iterationen und im Schnitt 400 Vorwärtsrechnungen pro Iterationsschritt benötigt ca. 13 Stunden auf einem Prozessorkern eines *Intel* Quadcore 3.16 GHz. Die 56 Rechnungen zum Test eines Verfahrens werden auf einem Clustersystem auf sieben Knoten mit jeweils 8 Kernen dieses Typs verteilt gerechnet.

In Tabelle 8.3 ist eine Auflistung der verschiedenen Vergleichsanalysen aufgestellt, nach denen die Verfahren zunächst im Hinblick auf erste und zweite Frage im einleitenden Abschnitt dieses Kapitels untersucht werden. Untersucht werden Standardabweichung der mittleren Inversionsergebnisse und die Anpassung zum Referenzmodell, um die Verlässlichkeit und algorithmische Auflösung des Verfahrens zu prüfen. Des weiteren wird durch Testläufe mit drei Modellen verschiedener Schichtanzahl die Abhängigkeit der Auflösung von der Dimensionalität des Problems abgeschätzt, da ein mathematischer Zugang zu dieser Fragestellung verwehrt bleibt. Zur Dimensionierung gilt es zwei weitere Fragen zu beantworten. Zum einen werden Inversionen mit verschiedener Populationsgröße an Testmodellen mit unterschiedlicher Schichtzahl durchgeführt, um möglichst eine optimale Populationsgröße für die unterschiedlichen Probleme zu finden. Zum anderen wird durch die Wahl von 1000 Iterationsschritten auf eine hinreichende Zahl Neustarts einer Optimierung geschlossen, die dann erreicht ist, wenn der Wert der besten Anpassung aller Neustarts mit der Iteration eine Sättigung erreicht. Des weiteren wird in Einzeluntersuchungen mit dem besten Optimierungsverfahren der Einfluß von Fehlern in den Langsamkeitswerten der Testdaten auf die Auflösung untersucht und in diesem Zusammenhang auch der Einfluß der Wahl von L_1 - oder L_2 -Norm. Die Symbole in Tabelle 8.3 finden sich nun in Tabelle 8.4 wieder, in welcher sie, zum Zwecke der Übersicht markieren, welche Verfahren und Parameterkonstellationen und welche Analysen ausgeführt wurden.

Aus Tabelle 8.4 geht damit hervor, welche Untersuchungen an welchen Verfahren durchgeführt wurden. Um Rechenzeit zu sparen, wurde diese Auswahl nach einem Ausschlußverfahren getroffen. So wurden die Analysen in einer bestimmten Reihenfolge durchgeführt, um weniger effiziente Verfahren frühzeitig ausschließen zu können. Die Abhängigkeit der Auflösung und des Konvergenzverhaltens von der

8.2 Durchführung der Tests

- | | |
|---|--|
| ① | Vergleich der Varianten der PSO |
| ② | Abhängigkeit der Auflösung für ein 5-Schicht-Modell von der Schwarmgröße |
| ③ | Abhängigkeit der Auflösung von 5,6,7-Schicht-Modellen von der Schwarmgröße des DHS-PSO |
| ④ | Wahl der Populationsgröße des CMA-ES |
| ⑤ | Methodenvergleich an der Inversion eines 5-Schicht-Modells |
| ⑥ | Methodenvergleich an der Inversion eines 7-Schicht-Modells |

Tabelle 8.3: *Durchgeführte Analysen zur algorithmischen Auflösung.*

Verfahren	PSO		cDHS PSO		rDHS PSO		(c+r) DHS PSO		CMA ES PSO		DHS PSO			CMA -ES		BH -SA		ABC		
	5	7	5	7	5	7	5	7	5	7	5	6	7	5	7	5	7	5	7	
50p	②												③							
100p	②												③							
300p	①	①	①	①	①	①	①	①	①	①	①	③	①				⑤		⑤	⑥
	②											③	③	③			*1		*2	*2
500p	②											③	③	③						
1000p												③	③							
$\lambda = 10$															④					
$\lambda = 50$															④					
$\lambda = 100$															④					
$\lambda = 300$															④	⑥				
															⑤					

Tabelle 8.4: *Durchgeführte Auflösungsanalysen.* *₁: Der BH-SA wurde nach 10000 Iterationen gestoppt; pro Iterationsschritt wurden 20 DHS Iterationen durchgeführt. *₂: Die Bienenpopulation betrug 200+200 und entspricht damit im Aufwand an Vorwärtsrechnungen dem DHS-PSO. Der Index p in der linken Spalte bezieht sich auf die Schwarmgröße.

Teilchenzahl der global besten PSO sowie für die DHS-PSO für den fünf, sechs und den sieben Schichten Testdatensatz, ② und ③, wurde zuerst betrachtet. Dies ermöglichte den Ausschluß der global besten PSO gegenüber der DHS-PSO sowie eine Dimensionierung des Verfahrens im Hinblick auf die Schwarmgröße und das Studieren des Verhaltens der Auflösung mit steigender Parameterzahl. Die Untersuchung der Abhängigkeit der Auflösung von der Populationsgröße für die CMA-ES für den fünf Schichten Testdatensatz ermöglichte eine optimale Wahl der Populationsgröße, ④. Es folgte die Applikation und Vergleich der Varianten der geladenen und der rasternden DHS-PSO und der PSO-CMAES, ①. Letztendlich wurde die Auflösung der verschiedenen Verfahren DHS-PSO, CMA-ES, ABC, BH-SA, cDHS-PSO, rDHS-PSO und (c+r)DHS-PSO bei vergleichbarem Aufwand an Vorwärtsmodellierungen verglichen. Dies geschah sowohl für den fünf Schichten Datensatz, ⑤, als auch anschließend für die erfolgreicherer Verfahren

ren DHS-PSO, CMA-ES und ABC für den sieben Schichten Datensatz, ©.

8.3 Auflösung und Effizienz der verschiedenen Verfahren

Um die eingangs gestellte erste und zweite Frage zu beantworten, wurden wie im vorigen Abschnitt beschrieben für verschiedene Verfahren und Verfahrensvariationen mehrere Rechnungen mit unterschiedlichen Startpopulationen durchgeführt. Hier sollen nun im einzelnen die Analysen und ihre Ergebnisse vorgestellt werden.

Die Auflösung und der Einfluß der Startpopulation

Die Antwort auf die Frage, wie gut ein Verfahren gegen das globale Minimum konvergiert, läßt sich in einfachster Form durch Abweichung der Lösungen von einem vorher bekannten Inversionsziel bemessen. Eine Betrachtung mehrerer Inversionsdurchgänge mit jeweils unterschiedlicher Startpopulation erlaubt darüber hinaus eine Beurteilung des Ergebnisses in Abhängigkeit von den Initialvoraussetzungen und ermöglicht das Erheben eines mittleren Modells und seiner Standardabweichungen, was als Maß der algorithmischen Auflösung interpretiert werden kann. Zur Bemessung der n_R ($n_R=1, \dots, 56$) Lösungen \vec{x}^* aller Inversionsdurchgänge gilt es zu entscheiden, ob der Median oder Zentralwert der Modellparameter

$$\overline{\vec{x}^*} = \frac{1}{2}(\vec{x}_{n_R/2}^* + \vec{x}_{n_R/2+1}^*)$$

mit dem Median der Abweichungen, oder das arithmetische Mittel mit Standardabweichung zu wählen ist. In Abbildung 8.3 sind für 56 Inversionsdurchgänge des DHS-PSO mit einer Populationsgröße von 300 Partikeln das mittlere Modell und der Median der Modellparameter mit ihren jeweiligen Abweichungen dargestellt. Zusätzlich zeigt die Abbildung die Verteilung von zwei Inversionsparametern, der Scherwellengeschwindigkeit an den Modellschichtpunkten vier und sechs. Diese Verteilungen zeigen nur geringe Asymmetrie, so daß das mittlere Modell nach beiden Berechnungen nahezu gleich erscheint. Die Abweichungen werden bei Betrachtung der Einzelmodelle (unten rechts) mit der Standardabweichung besser repräsentiert. Auch sind in Abbildung 8.3 keine Ausreißer in der Gesamtheit aller Lösungsmodelle zu erkennen, so daß arithmetisches Mittel und Standardabweichung als Maß der Auflösung und Effizienz gewählt wurden.

PSO und DHS-PSO: Teilchenzahldimensionierung

Zuerst sollen nun die Ergebnisse der Suche nach der optimalen Populationsgröße bei verschiedener Schichtanzahl sowie der Vergleich von einfacher global bester

8.3 Auflösung und Effizienz

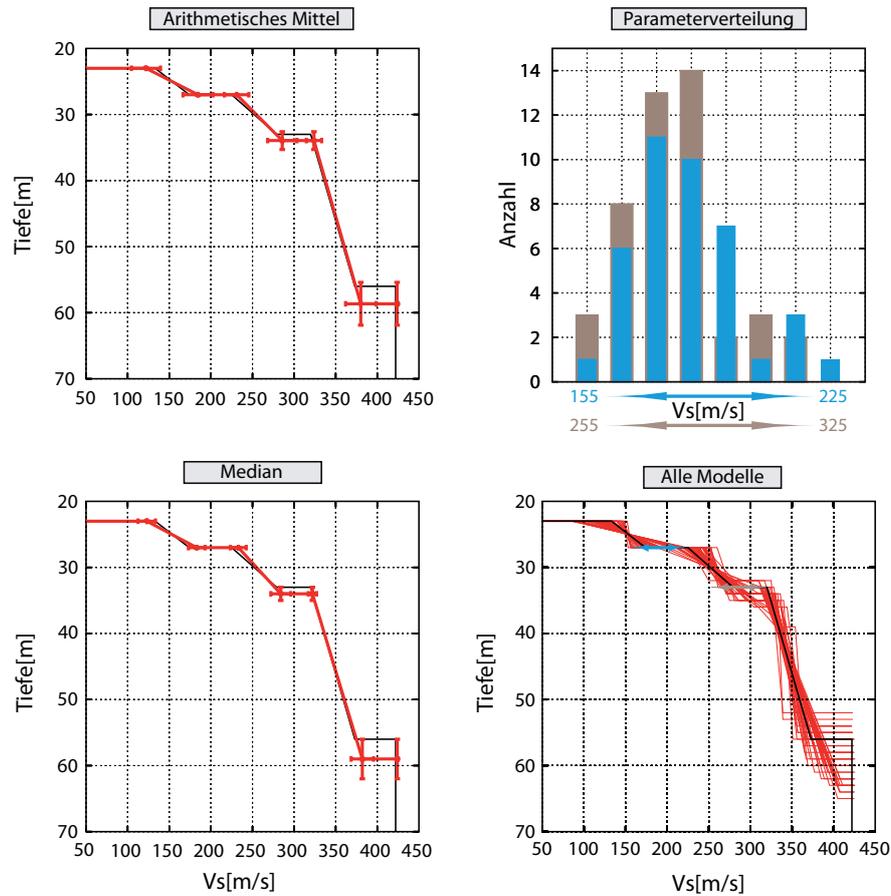


Abbildung 8.3: Beispielmenge von 56 Inversionslösungen des DHS-PSO am fünf Schichten Modell. Mittlere Modelle für die Lösungsmenge (unten rechts) wurden mit Median und arithmetischem Mittel bestimmt (linke Seite). Eine Verteilung von Scherwellengeschwindigkeiten für Modellpunkt vier und sechs (siehe farbige Pfeile unten rechts) ist oben rechts aufgetragen (Bin von 10 m/s).

PSO und DHS-PSO vorgestellt werden. Abbildung 8.4 zeigt hierzu die mittleren Modelle und Standardabweichungen der Parameter Scherwellengeschwindigkeit, v_s , und aus den Schichtmächtigkeiten berechnete Schichtgrenzen für 56 Lösungen von Inversionen mit 1000 Iterationen und Neustartmöglichkeit. Die Inversionen wurden jeweils am FINO3 Referenzmodell mit Populationsgrößen von 50, 100, 300 und 500 Partikeln der PSO und DHS-PSO durchgeführt. In der oberen Zeile der Abbildung sind die mittleren Modelle der PSO in rot aufgetragen und stehen im Vergleich zum Referenzmodell (schwarz). Es lässt sich eine Verbesserung der Auflösung von im Mittel bereits sehr guten 8% auf 6% mit steigender Partikelzahl feststellen. Schlechteste Auflösung zeigt sich grundsätzlich im tieferen Bereich der Modelle, welcher offenbar nur durch geringe Datendichte im tieffrequenten Be-

8 METHODENVERGLEICH UND AUFLÖSUNG

reich des Spektrums abgedeckt wird. Eine Änderung mit der Teilchenzahl ist ebenfalls in der Anpassung des mittleren Modells an das Zielmodell zu erkennen. Hier zeigt sich bei 300 Partikeln eine sehr gute Anpassung, die sich mit 500 Partikeln nicht nennenswert verbessert. Es sei daher eine Populationsgröße von 300 Partikeln vermerkt. Die untere Zeile der Abbildung zeigt mittlere Lösungsmodelle bei einer Anpassung durch die DHS-PSO (hellblau). Hier ist zum einen eine klare Verbesserung der Anpassung des mittleren Modells im direkten Vergleich mit der PSO zu erkennen, zum anderen aber auch eine klare Erhöhung der Auflösung mit der Populationsgröße. Die Erhöhung der Populationsgröße auf 500 Partikel zeigt auch hier keine wesentlich bessere Anpassung des mittleren Modells an das Referenzmodell.

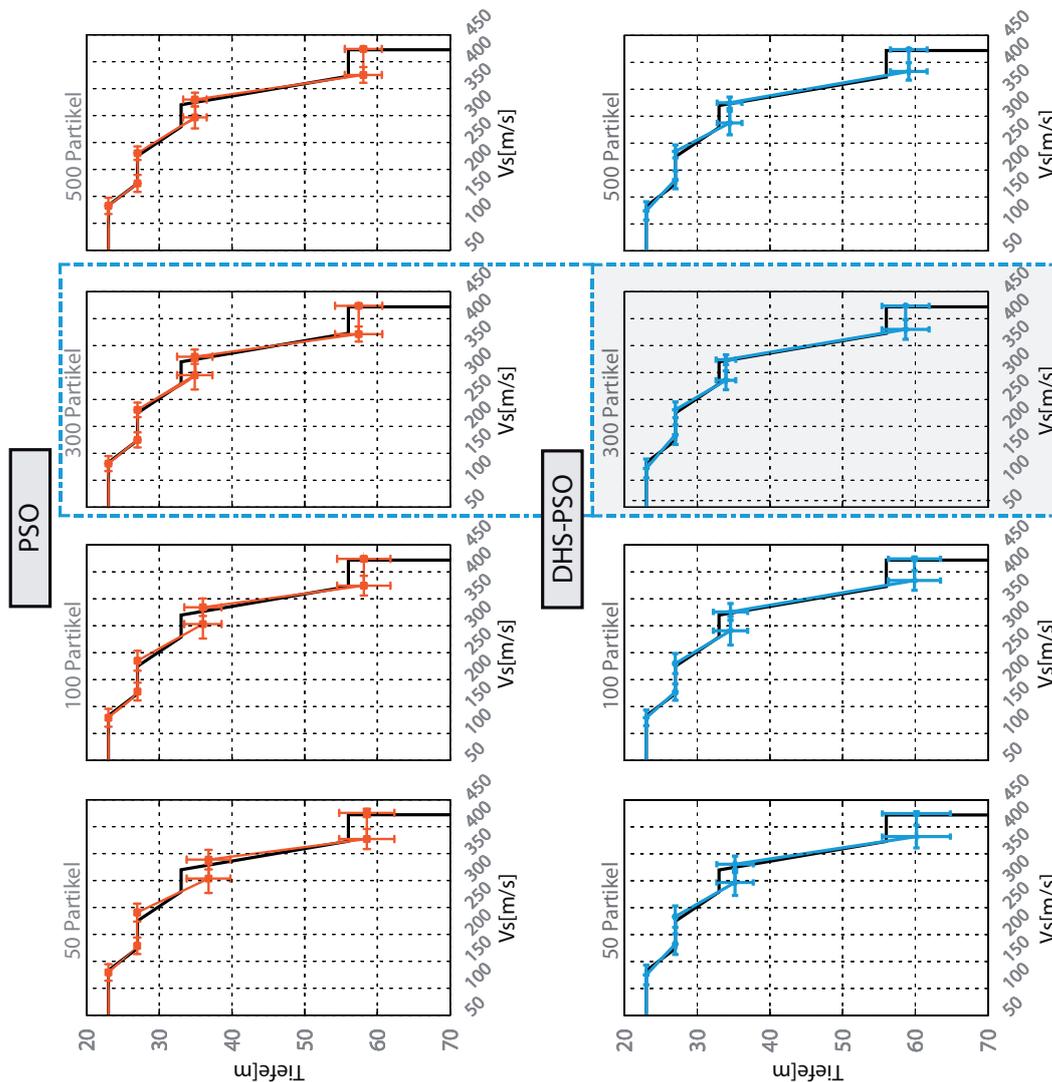


Abbildung 8.4: Darstellung aller mittleren Modelle und Standardabweichungen von jeweils 56 Inversionen bei variabler Populationsgröße und den Verfahren PSO und DHS-PSO. Obere Spalte zeigt die Lösungen der PSO in rot im Vergleich mit dem Referenzmodell (schwarz). Die untere Spalte zeigt die Lösungen der DHS-PSO in hellblau.

Zusätzlich sei in Abbildung 8.5 der Verlauf des Qualitätsmaßes aufgetragen. Da die Verfahren im Neustartmodus verwendet wurden, sei hier zu erwähnen, daß die Verfahren unterscheiden, zwischen dem absolut besten Modell, daß alle Neustarts bislang gefunden haben und dem beste Modell eines einzelnen gestarteten Optimierungsvorganges. Die Abbildung zeigt daher nur das absolut beste Qualitätsmaß einer Inversion, welches jeder Neustart eines Verfahren stets zu überbieten hat, im Mittel über die 56 Inversionen. Hier zeigt sich die klare Überlegenheit der

8 METHODENVERGLEICH UND AUFLÖSUNG

DHS-PSO gegenüber der PSO erneut. Die 50 Partikel DHS-PSO zeigt sogar effizientere Konvergenz als die 100 Partikel PSO. Die drei besten Konstellationen sind (gekennzeichnet durch das graue Feld) die 300 Partikel DHS-PSO sowie beide 500 Partikel Beispiele. Aber auch hier offenbart sich keine nennenswerte Verbesserung im Vergleich von beiden 500 Partikel Beispielen zur 300 Partikel DHS-PSO, so daß letztere als bislang effizientestes Verfahren identifiziert wird.

Nichtsdestotrotz ist zu erwarten, daß die benötigte Partikelzahl für vergleichba-

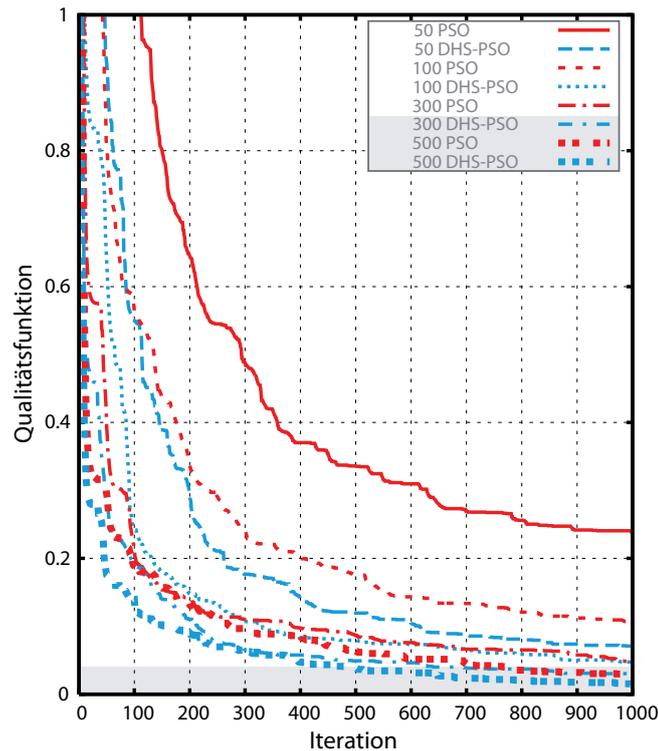


Abbildung 8.5: Fortschritt des mittleren absolut besten Qualitätsmaßes mit der Iteration für die PSO und DHS-PSO bei 50,100,300 und 500 Partikeln.

re Auflösungen und Anpassungen bei höherer Parameterzahl ansteigt. Zu diesem Zweck sind in Abbildung 8.6 die mittleren Lösungsmodelle und Standardabweichungen der Parameter für den fünf, sechs und sieben Schichten Fall mit 300 und 500 Partikeln dargestellt. Der sechs Schichten Fall zeigt ähnlich gute Anpassung wie im Falle des FINO3 Referenzmodells. Im Falle von sieben Schichten zeigt sich jedoch eine starke Verschlechterung von Anpassung und Auflösung, zumindest im tieferen Bereich mit geringerer Datendichte. Die Erhöhung der Partikelzahl auf 500 zeigt hier eine Verbesserung der Auflösung. Es ist jedoch nicht zu erwarten, daß eine Analyse dieser Größenordnung quantitative Aussagen über die Wahl der Populationsgröße in Abhängigkeit von der Dimensionalität des Problems zuläßt. Dies würde eine viel größere Menge an Rechnungen erfordern. Es ist aber klar,

daß die Erhöhung der Parameterzahl die Multimodalität des Problem gravierend erhöht, so daß wiederum eine erhöhte Exploration benötigt wird. Die Anpassung des sieben Schichten Falls in Abbildung 8.6 stützt dies, denn das erhaltene mittlere Modell zeigt einen dem Referenzmodell sehr ähnlichen Gradienten, jedoch sind die Fehler in den Tiefen der unteren Schichten sehr groß. Es läßt sich hier vermuten, daß ein Modell entsprechend eines Nebenminimums angepaßt wurde. Der im Fokus stehende Fall der Baugrundüberwachung in der Nordsee bewegt sich jedoch im Bereich von fünf, maximal sechs Schichten, so daß die Wahl der 300 Partikel DHS-PSO zunächst als gut abgeschätzt sei. Allerdings muß der Inversion von komplexeren Modellen eine Analyse der minimal benötigten Partikelzahl vorausgehen.

Varianten der PSO

In diesem Abschnitt seien die aussichtsreichsten Varianten der einfachen PSO auf das Testproblem des FINO3 Referenzmodells angewendet und verglichen. Es werden die Varianten des geladenen Schwarms, welcher die Exploration des Parameterraumes erhöhen soll und des rasternden Schwarms, welcher Information zuvor gefundener Minima nutzt, um durch Abstoßung der Partikel eine erneute Konvergenz zu diesen Minima zu vermeiden sowie Kombinationen aus beiden getestet. Die Variationen wurden jeweils an der schon als sehr effektiv erkannten Variante des DHS-PSO durchgeführt. Abbildung 8.7 zeigt mittlere Modelle und Standardabweichung für diese Variationen sowie erneut das mittlere Lösungsmodell der global besten PSO zu Referenzierung der Lösungen.

8 METHODENVERGLEICH UND AUFLÖSUNG

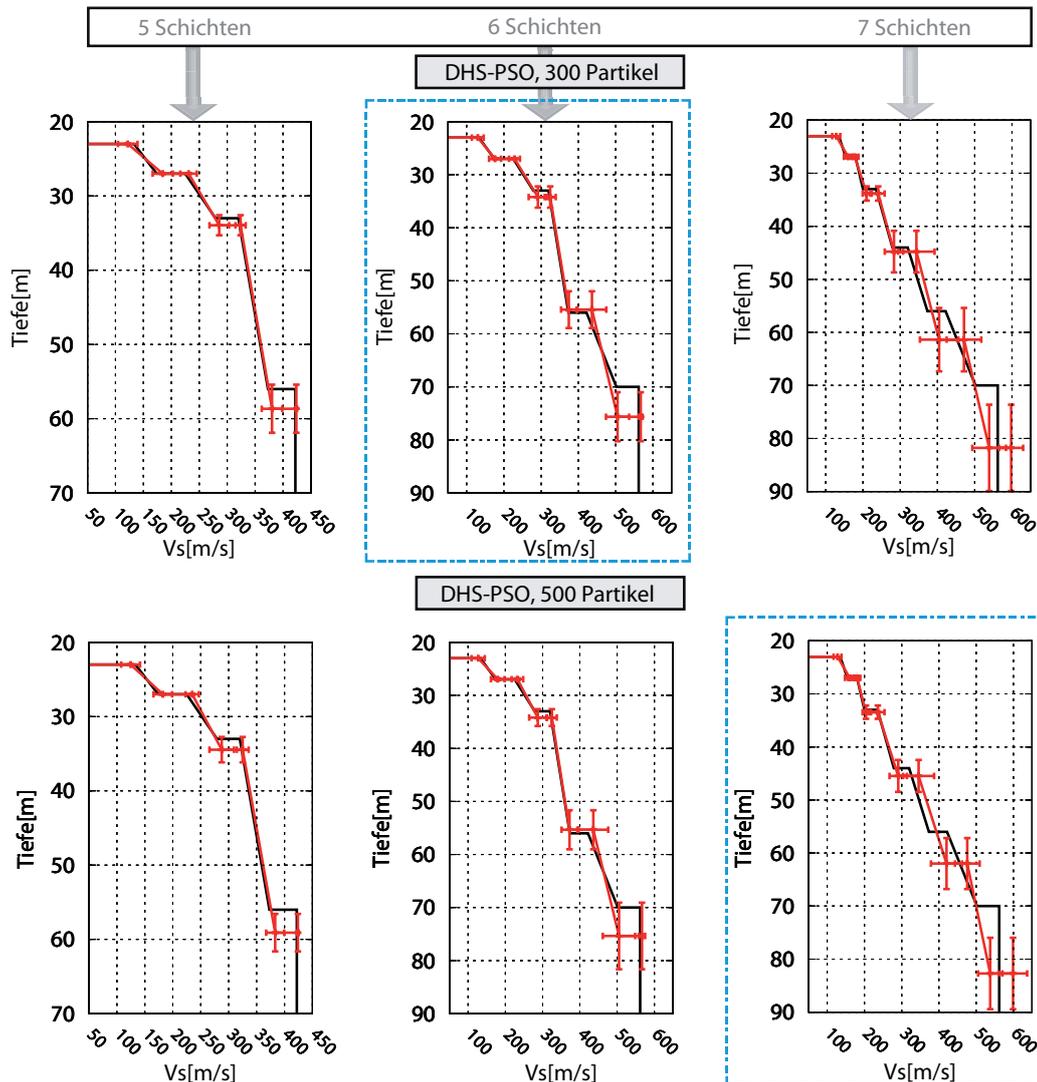


Abbildung 8.6: Darstellung der mittleren Lösungsmodelle und Standardabweichungen von jeweils 56 Inversionen unterschiedlicher Startpopulation für die DHS-PSO mit Populationsgröße 300 und 500 bei fünf, sechs und sieben Modellschichten.

Auf den ersten Blick zeigen die Ergebnisse aller DHS-PSO Varianten keine großen Unterschiede in Anpassung und Auflösung. Der Vergleich des Fortschritts des besten Qualitätswertes zeichnet jedoch die cDHS-PSO als effektivste Variante mit maßgeblicher Verbesserung der Konvergenz aus. Bei genauer Betrachtung der Ergebnismodelle aller gezeigten Modifikationen der PSO zeigen sich geringe Unterschiede in der Auflösung der Scherwellengeschwindigkeit und der Schichttiefe (siehe auch Hilfslinien in Abbildung 8.7). Die Anpassung des mittleren Modells (rot) an das Referenzmodell (hellblau) ist bei allen gezeigten Varianten vergleichbar

8.3 Auflösung und Effizienz

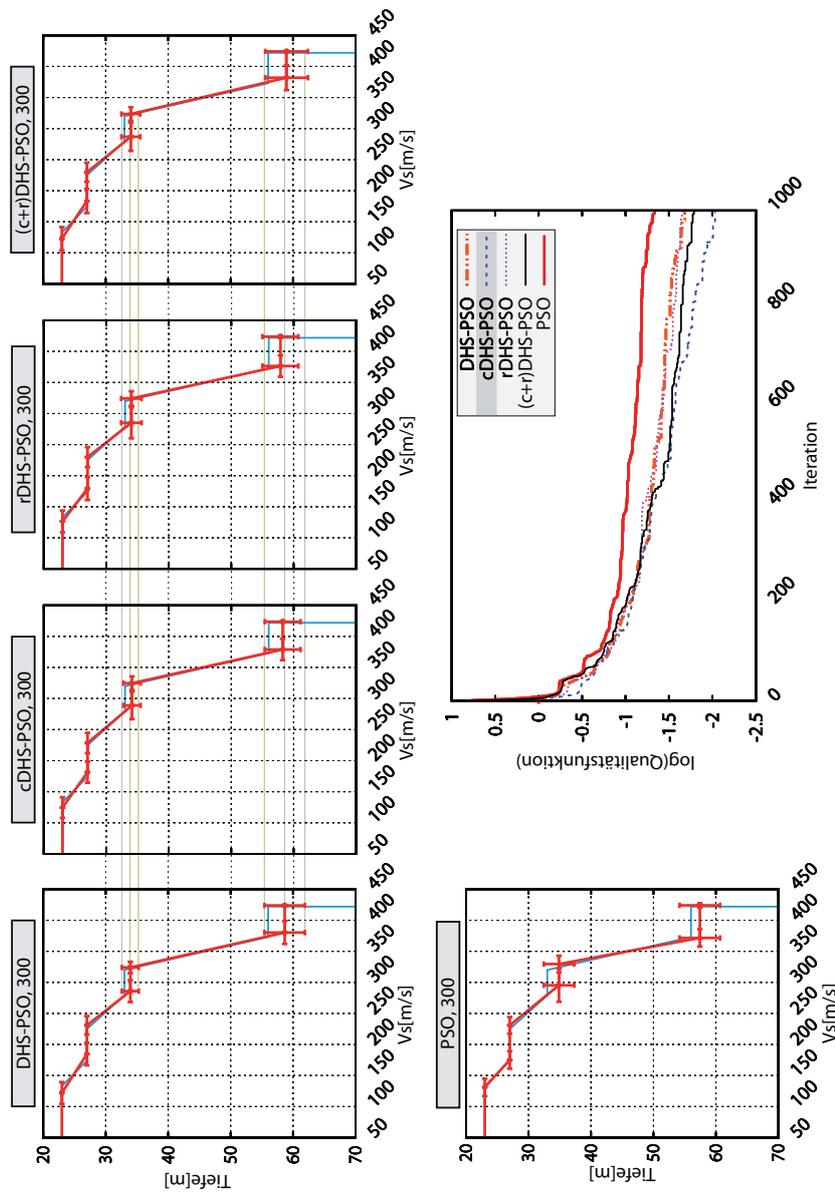


Abbildung 8.7: Darstellung der mittleren Lösungsmodelle und Standardabweichungen von jeweils 56 Inversionen unterschiedlicher Startpopulation für verschiedene Modifikationen der PSO. Der Graph unten rechts zeigt den Fortschritt des mittleren absolut besten Qualitätswertes mit der Iteration für eben diese Varianten. Der graue Balken in der Legende kennzeichnet die effizienteste Variante.

gut, jedoch läßt sich bei der cDHS-PSO eine Verbesserung im Tiefenbereich von 26 m bis 34 m erkennen. Die kleinen Unterschiede in Auflösung und Anpassung lassen vermuten, daß die erhöhte Exploration hauptsächlich zu einer Verbesse-

8 METHODENVERGLEICH UND AUFLÖSUNG

zung der Ausbeute von Nebenminima im Nahbereich der Optimallösung (siehe auch Abbildung 5.2) führt.

Vergleich verschiedener Verfahren

Da nun die für das Scholtewellen Inversionsproblem effektivste Variante der PSO in der Menge der getesteten Varianten als cDHS-PSO identifiziert wurde, gilt es diese mit anderen gängigen stochastischen Optimierungsverfahren zu vergleichen. In Kapitel 7 wurden hierzu die Evolutionsstrategie und die Evolutionsstrategie mit Kovarianzmatrixadaption zur Anpassung des Mutations- bzw. Schrittintervalls sowie ein Hybrid aus Downhill-Simplex und Metropolis Algorithmus vorgestellt. Des weiteren wurde eine zweite Schwarmintelligenzmethode mit dem Verfahren des künstlichen Bienenschwarms eingeleitet. Diese Verfahren wurden nun in der Anwendung auf das FINO3-Referenzmodell und das sieben Schichten Referenzmodell mit der cDHS-PSO verglichen. Hierbei galt es zu beachten, daß jedem Verfahren eine vergleichbare Menge an Vorwärtsmodellierungen zur Verfügung steht. So ist die Populationsgröße des ABC zu 200 angestellten und 200 springenden Bienen definiert. Dies entspricht der Populationsgröße von 300 Schwarmpartikeln mit anschließender DHS-Suche mit 100 Iterationen. Die implementierte $(\mu + \lambda)$ -Evolutionsstrategie erhielt eine Populationsgröße von $\lambda=400$ Nachkommen und $\mu=300$ Eltern. Der BH-SA wurde auf 10000 Iterationen festgelegt mit jeweils 20 Iterationen für den DHS. Die CMA-ES wurde mit $\lambda=300$ initialisiert. Hierzu sei zunächst eine Voruntersuchung betrachtet, welche das Verhalten der CMA-ES mit der Populationsgröße λ zeigt. Abbildung 8.8 zeigt das mittlere Modell von 56 Inversionsdurchläufen mit Standardabweichung für CMA-ES mit $\lambda = 300$, $\lambda = 100$ und $\lambda = 50$, ebenso wie das Konvergenzverhalten der drei Populationsgrößen. Die Modellanpassung ist in allen drei Fällen vergleichbar. Es zeigt sich aber eine Verkleinerung der Standardabweichungen der Parameter mit steigender Populationsgröße, ohne daß das mittlere Modell sich aber dem Referenzmodell nähert und das Referenzmodell sogar außerhalb der Parameterabweichungen liegt. Dies läßt vermuten, daß das Verfahren in allen drei Fällen in ein Nebenminimum konvergiert und die Erhöhung der Populationsgröße nur die lokale Suche verstärkt, was verbesserte Auflösung des Nebenminimums zur Folge hat. Die CMA-ES scheint unabhängig ihrer Populationsgröße eine zu geringe Exploration trotz des mehrfachen Neustarts des Verfahrens aufzuweisen. Das Verfahren scheint also schon bei geringer Populationsgröße ein sehr effektives lokal suchendes Verfahren zu sein, so daß es auf der Hand liegt, einen Hybriden aus CMA-ES und PSO zu testen.

Die Abbildung 8.9 zeigt nun das arithmetische Mittel der 56 Inversionen für die verschiedenen Verfahren, zusätzlich dazu einem Hybriden aus global PSO und CMA-ES („CMAES-PSO“), welcher dem global besten Teilchen der PSO eine lokale CMA-ES-Suche erlaubt. In erster Linie zeigt die Abbildung, daß alle Verfahren eine tendenziell vernünftige Anpassung erreichen, bis auf die $(\mu + \lambda)$ -

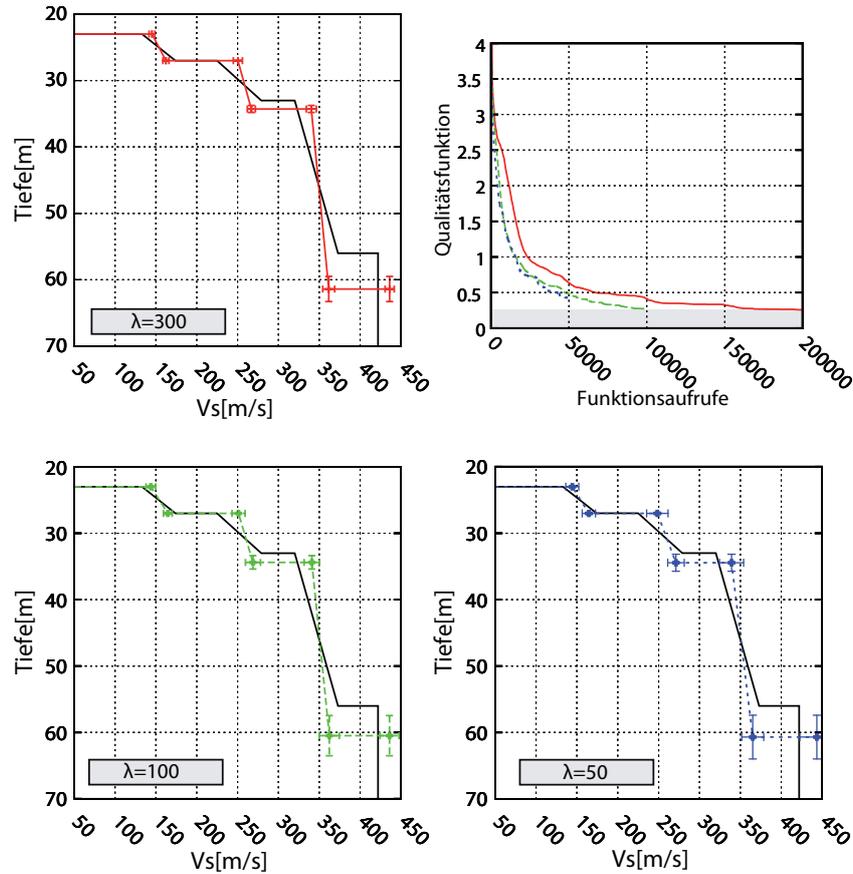


Abbildung 8.8: Darstellung der mittleren Lösungsmodelle und Standardabweichungen von jeweils 56 Inversionen unterschiedlicher Startpopulation für verschiedene Populationsgrößen der CMA-ES.

Evolutionstrategie. Das mittlere Modell dieses Verfahrens zeigt negative Geschwindigkeitssprünge und sehr große Abweichungen. Grund hierfür ist vermutlich die nicht adaptive Wahl des Mutationsbereiches, welche zu geringer lokaler Konvergenz führt. Dies läßt dem Verfahren eine hohe Exploration, aber geringe Effizienz bei lokaler Suche, was zu diesem schlecht aufgelösten mittleren Modell mit aber grob richtiger Tendenz führen würde. Die CMA-ES, welche dieses Problem durch adaptive Schrittweite umgeht, zeigt ein ganz anderes Verhalten. Das Ergebnismodell zeigt vernünftige aber im Vergleich mit den anderen Verfahren die schlechteste Anpassung, wohingegen das Ergebnismodell die kleinste Standardabweichung aufweist. Dies deutet, wie schon in der Voruntersuchung vermutet, erneut darauf hin, daß das Verfahren in ein Nebenminimum konvergiert, dort die lokale Suche verstärkt und so eine sehr gute Auflösung des Nebenminimums aufweist. Die CMA-ES zeigt damit nach wie vor zu geringe Exploration,

8 METHODENVERGLEICH UND AUFLÖSUNG

bzw. die Exploration wird mit fortschreitenden Iterationen nicht erhalten. Der in Abbildung 8.9 auch untersuchte Hybrid aus PSO und CMA-ES zeigt hingegen eine sehr viel bessere Anpassung des mittleren Modells an das Referenzmodell. Das Ergebnis ist hier vergleichbar mit dem BH-SA. Beide zeigen gute Anpassung aber geringe Auflösung der Parameter. Als effizienteste Verfahren in Anpassung und Auflösung zeigen sich die beiden vorgestellten Schwarmintelligenzmethoden cDHS-PSO und ABC, welche offenbar einen hinreichenden Kompromiß aus Exploration und lokaler Konvergenz ermöglichen.

8.3 Auflösung und Effizienz

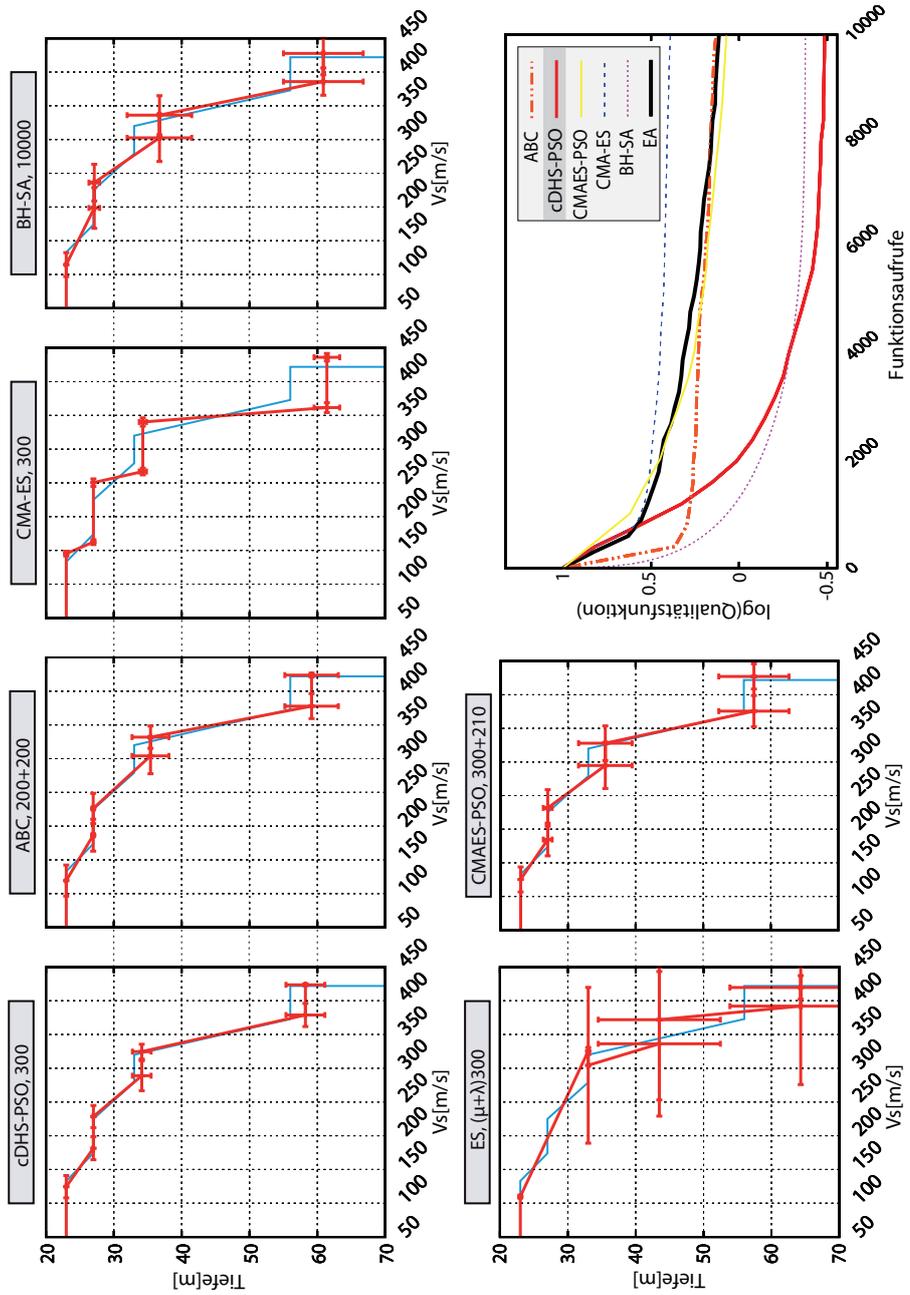


Abbildung 8.9: Darstellung der mittleren Lösungsmodelle und Standardabweichungen von jeweils 56 Inversionen unterschiedlicher Startpopulation für verschiedene Optimierungsverfahren am FINO3 Referenzmodell. Unten rechts: Fortschritt des mittleren absolut besten Qualitätswertes der Verfahren mit der Iteration. Der graue Balken in der Legende kennzeichnet das effizienteste Verfahren.

8 METHODENVERGLEICH UND AUFLÖSUNG

Abbildung 8.10 kann dies noch einmal für das sieben Schichten Referenzmodell bestätigen. Die Abbildung zeigt den Vergleich der nun drei effizientesten Verfahren CMA-ES, ABC und cDHS-PSO. Die CMA-ES zeigt nach wie vor die beste Auflösung, aber auch die schlechteste Anpassung.

Die Auflösung des besten Verfahrens, cDHS-PSO, am FINO3 Referenzmodell ist

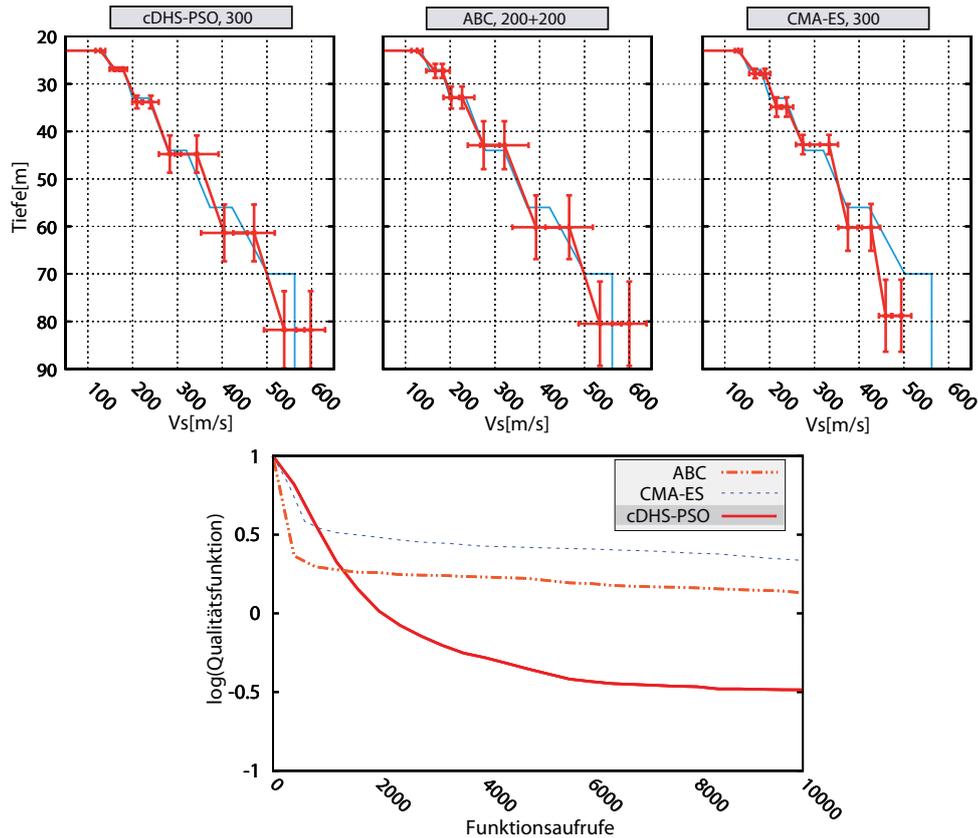


Abbildung 8.10: Darstellung der mittleren Lösungsmodelle und Standardabweichungen von jeweils 56 Inversionen unterschiedlicher Startpopulation für verschiedene Optimierungsverfahren am sieben Schichten Referenzmodell. Unten rechts: Fortschritt des mittleren absolut besten Qualitätswertes der Verfahren mit der Iteration. Der graue Balken in der Legende kennzeichnet das effizienteste Verfahren.

in Zahlen noch einmal in Tabelle 8.5 angegeben. Es wird hier in der Scherwellengeschwindigkeit eine algorithmische Auflösung von im Mittel 6.5% bzw. 17 m/s erreicht. Die Unterschiede von mittlerem Modell und Referenzmodell betragen im Mittel nur 6 m/s in den Scherwellengeschwindigkeiten. Dies führt nun zur zweiten eingangs gestellten Frage, wie sich Fehler in den Phasenlangsamen auf die Auflösung des Ergebnisses auswirken.

Z[m]	ΔZ [m]	v_s [m/s]	Δv_s [m/s]	v_s [m/s](Referenz)
0.00	0.00	0.0	0.0	0.0
23.0	0.00	0.0	0.0	0.0
23.0	0.00	124.5	16.6 (13%)	133.0
27.0	0.00	181.4	17.0 (9.4%)	174.0
27.0	0.00	228.4	16.4 (7.2%)	225.0
34.1	1.37	288.9	22.1 (7.6%)	279.0
34.1	1.37	324.6	10.8 (3.3%)	320.0
58.2	2.86	378.9	17.1 (4.5%)	373.0
58.2	2.86	423.7	1.30 (0.3%)	422.0

Tabelle 8.5: Mittleres Modell der cDHS-PSO mit 300 Partikeln für den fünf Schichten Fall (Referenz) sowie Standardabweichungen.

Fehlereinfluß und Vergleich der Normen L_1 und L_2

In Abbildung 8.11 sind mittlere Modelle für die cDHS-PSO mit 300 Partikeln bei unterschiedlich gewählter Norm und integrierten künstlichen Fehlern in den Phasenlangsamkeiten dargestellt. Hierzu wurden normalverteilte Fehler auf die Phasenlangsamkeiten der Eingangsdaten des FINO3 Referenzmodells (siehe Abbildung 8.2) addiert, wobei für die 56 Inversionsdurchläufe das künstliche Spektrum mit jeweils unterschiedlichen Langsamkeitsfehlern versehen wurde, so daß unterschiedliche Repräsentationen eines möglichen realen Datensatzes zu Grunde liegen. Dieses Vorgehen erzeugt eine Fehlersynergie aus der Streuung des Ergebnisses durch den Einfluß der Startpopulation und der algorithmischen Auflösung sowie einer zusätzlichen Streuung der Ergebnisse durch verschiedene Fehler der Phasenlangsamkeiten. Die Standardabweichung der Fehlerverteilung entsprach hierbei der Phasenlangsamkeitsauflösung bei der jeweiligen Frequenz f (Gleichung 4.5), bei einem Offsetbereich von 1km (Abbildung 8.12 zeigt die so angenommenen Standardabweichungen für die Fehler der Phasenlangsamkeit). Die Ergebnisse in Abbildung 8.11 a) zum Vergleich der Normen L_1 und L_2 bestätigen die Wahl der L_1 -Norm aus Abschnitt 6.2 durch die erreichte, leicht bessere Auflösung. Jedoch stellte sich hier die Frage, ob dies sich bei verrauschten Daten ebenso verhält. Hierzu wurde eine Gewichtung der Eingangsdaten nach der Phasenlangsamkeitsauflösung eingeführt. Die Qualitätsfunktion berechnet sich so nach Gleichung 4.8. In Abbildung 8.11 b) wurden hierzu das Ergebnismodell mit gewichteter L_1 -Norm-Qualitätsfunktion in den selben Vergleich mit der L_2 -Norm angeführt. Die Anpassung der gewichteten Abweichungssumme erzielt hier, wie zu erwarten, eine bessere Anpassung und Auflösung als die ungewichtete Version. Die Veränderung der Auflösung durch fehlerbehaftete Daten im Vergleich zu Daten ohne Fehler zeigt sich in Abbildung 8.11 c). Die ursprüngliche algorithmische Auflösung von 6.5% verschlechtert sich hierbei auf nur 8.4%. Des weiteren wird im

8 METHODENVERGLEICH UND AUFLÖSUNG

für FINO3 interessanten oberen Bereich des Modells noch immer eine Auflösung von 18 m/s erreicht, da hier mit höheren Moden auch die Datengrundlage größer ist und der Auflösungsfehler klein. Die vorwärts modellierte Dispersionskurve bildet stets eine Art physikalische Ausgleichskurve, auch bei verrauschten Daten. So werden Fehler von großer Bedeutung sein, die zum einen den tieffrequenten Bereich der Dispersionskurve betreffen, aber auch solche, die der Dispersionskurve in Bereichen eine falsche Tendenz vorgeben. In diesem Fall würde das Rauschen angepaßt. Eine Gewichtung der eingehenden Daten ist daher sehr wichtig, auch wenn der Auflösungsgewinn im betrachteten Fall merkbar, aber doch gering war.

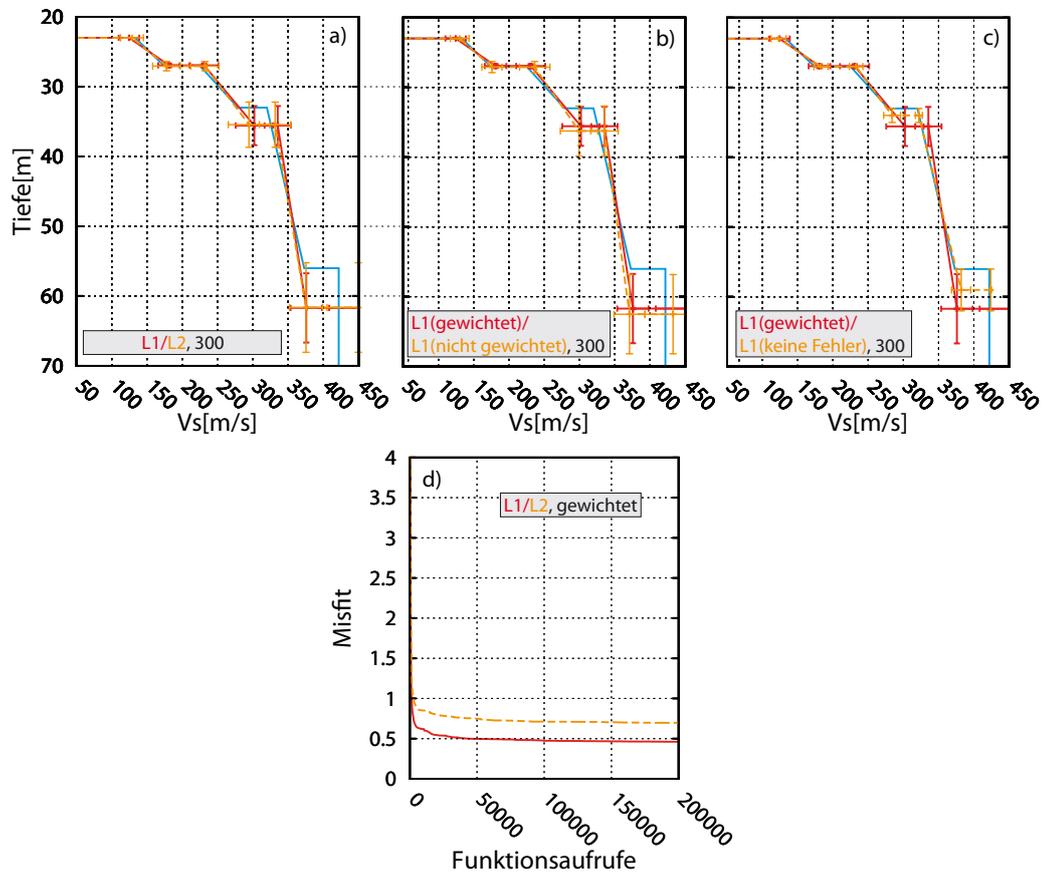


Abbildung 8.11: Darstellung der mittleren Lösungsmodelle und Standardabweichungen von jeweils 56 Inversionen unterschiedlicher Startpopulation für die cDHS-PSO mit unterschiedlichen Normen und künstlich verrauschten Eingangsdaten.

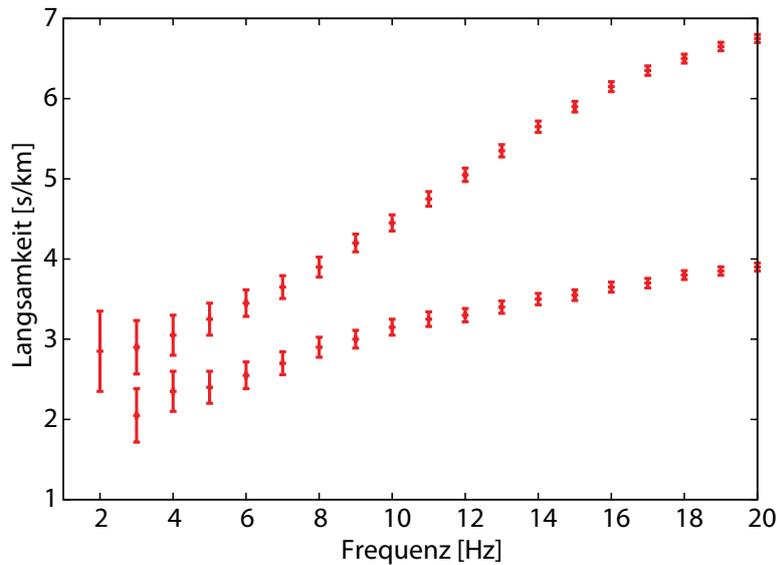


Abbildung 8.12: Standardabweichungen für künstliche, Gaußverteilte Fehler der Phasenlangsamkeiten zum FINO3 Referenzmodell.

Einfluß weiterer Parameter (Dichte, Tide, Strafterm)

Bei der Frage nach einer optimierten Auflösung der Scholtewelleninversion zur Anwendung in der Baugrundüberwachung müssen weitere zusätzliche Fragen angegangen werden. Bisher wurde die Auflösung bei einer Inversion von Scherwellengeschwindigkeiten und Schichtmächtigkeiten betrachtet. Jedoch bleibt die Frage, in wie weit der Einfluß der Wahl der Dichte nun bei höheren Auflösungen steigt. Auch der Einfluß der Tiden und der damit verbundenen Änderungen der Mächtigkeit der Wassersäule sei im Bereich von Auflösungen von 16 m/s zu untersuchen. Zu diesem Zweck werden zunächst die Änderungen der Dispersionskurven durch Änderung des Dichteprofiles von $\Delta\rho = -400, -200, +200, +400 \text{ t/m}^3$ und durch Änderungen der Wassersäule um $\Delta h_w = -3, +3 \text{ m}$ betrachtet. Die Änderungen der Phasenlangsamkeitspics durch Variation der Wassersäule um $\pm 4 \text{ m}$ zeigen sich als kleiner als ein Langsamkeitssample von 0.05 s/km und sind damit vernachlässigbar klein zu erkennen. Die Änderung der Dichte (siehe Abbildung 8.13 „Dichte“) hat jedoch einen signifikanten Einfluß, der sogar Änderungen außerhalb der Phasenlangsamkeitsauflösung bei höheren Frequenzen erzeugt. Eine Untersuchung des Dichteinflusses auf ein eventuelles Inversionsergebnis ist also zu untersuchen. Daher sei hier erneut das mittlere Modell und die Auflösung von mehrfachen Inversionsergebnissen für das FINO3 Referenzmodell mit dem ursprünglichen Dichtemodell und einem um 200 t/m^3 erhöhtes Dichtemodell in Abbildung 8.13 und Tabelle 8.6 aufgezeigt („ ρ_1, ρ_2 “). Diese Untersuchung prüfe den Einfluß eines „falschen“ Dichteprofiles auf das Ergebnis in Scherwellen-

8 METHODENVERGLEICH UND AUFLÖSUNG

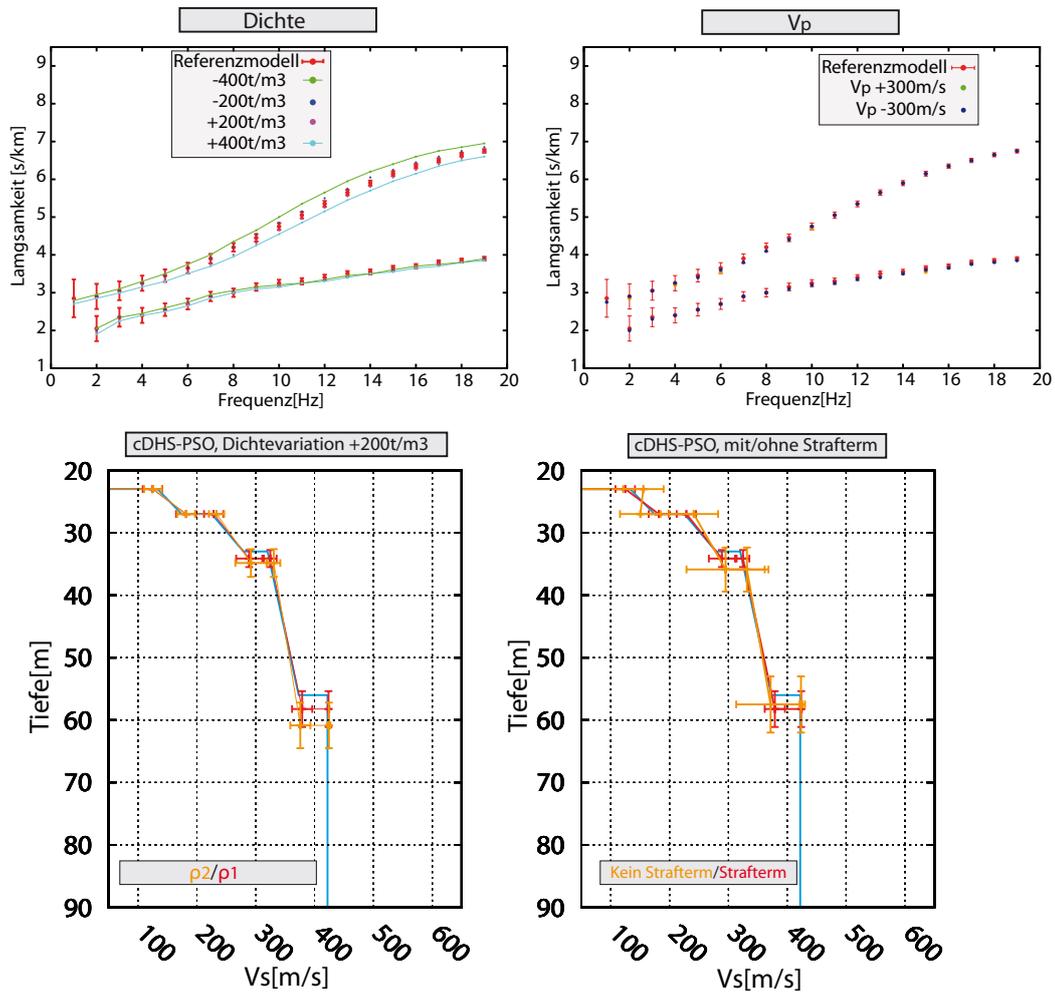


Abbildung 8.13: Änderungen der Phasenlangsamkeitspics durch Variation der Compressionswellengeschwindigkeit V_p und Dichte (oben). Einfluß einer Dichtevariation von $\pm 200 \text{ t/m}^3$ auf das Inversionsergebnis (unten links) und Einfluß des Strafterms für negative Gradienten auf das Inversionsergebnis (unten rechts).

geschwindigkeit und Schichtmächtigkeiten, denn die FINO3-Referenzdaten sind mit dem ursprünglichen Dichteprofil erzeugt worden. Bei geringen Abweichungen der Ergebnisse überwiege damit die Sensitivität der Scherwellengeschwindigkeiten gegenüber der nicht anpaßbaren Dichten. Sollte sich eine starke Änderung der Anpassung einstellen, so läge die Vermutung nahe, daß der Zusammenhang zwischen Dichte und Scherwellengeschwindigkeit dafür sorgt, daß letztere das nicht anpaßbare Dichteprofil ausgleicht. Betrachtet man nun die Ergebnisse in Abbildung 8.13, so zeigt die Auflösung der Scherwellengeschwindigkeiten kaum Unterschiede, wohl jedoch die Auflösung der Schichtmächtigkeiten/Schichttiefen, welche sich

leicht verschlechtert. Die Anpassung verschlechtert sich jedoch hauptsächlich in tieferen Bereichen des Modells. Für eine hohe Auflösung auch tieferer Bereiche ist somit eine eingehende Betrachtung der Dichte zu empfehlen und diese durch a priori Informationen sinnvoll zu setzen und mittels empirischer Bedingungen zu variieren, oder sogar mit Hilfe weiterer Datensätze in einer gemeinsamen Inversion aktiv mit zu invertieren (siehe folgendes Kapitel). In flacheren Bereichen des Mo-

Z[m]	v_s [m/s]	Δv_s [m/s]	v_s [m/s](Referenz)	v_s [m/s]	Δv_s [m/s]	Z[m]
$\rho(z)_1$	$\rho(z)_1$	$\rho(z)_1$		$\rho(z)_2$	$\rho(z)_2$	$\rho(z)_2$
0.00	0.0	0.0	0.0	0.000	0.00	0.00
23.0	0.0	0.0	0.0	0.000	0.00	23.0
23.0	124.5	16.6	133.0	125.4	15.1	23.0
27.0	181.4	17.0	174.0	181.6	15.1	27.0
27.0	228.4	16.4	225.0	233.2	12.9	27.0
34.1	288.9	22.1	279.0	291.7	26.5	34.8
34.1	324.6	10.8	320.0	330.5	11.0	34.8
58.2	378.9	17.1	373.0	375.7	17.0	60.8
58.2	423.7	1.30	422.0	424.4	2.40	60.8

Tabelle 8.6: Mittleres Modell der cDHS-PSO mit 300 Partikeln für den fünf Schichten Fall (Referenz, $\rho(z)_1$) und ein um $+200t/m^3$ verschiedenes Dichteprofil ($\rho(z)_2$), also ein zweites Referenzmodell.

dells bis 50 m, kann jedoch die Inversion aufgrund der v_s -Anpassung und bei stets falscher Dichteinformation eine vernünftige Anpassung vornehmen. Die Aussage von Bohlen et al. (2004) zur Entartung der Inversionsergebnisse bei unterschiedlicher Wahl einer Dichterelation, die im Abschnitt 4.2 bereits aufgegriffen wurde, konnte also auch mit neuer Qualitätsfunktion und Anpassungsoptimierung verifiziert werden.

Die letzte Prüfung dieses Abschnitts bezieht sich auf die Effizienz des eingeführten, einfachen Summations-Straftermes zur Unterbindung negativer Geschwindigkeitsgradienten. Durch dieses Unterbinden von einer großen Zahl von Parameterkonstellationen wird eine massive Verbesserung der Auflösung erwartet. Abbildung 8.13 zeigt nun ebenfalls die mittleren Inversionsergebnisse für Inversionen des FINO3-Referenzmodells mit der cDHS-PSO zum einen mit und zum anderen ohne Nutzung des Strafterms. Die Ergebnisse bestätigen die Vermutung, zeigt sich doch eine stark verminderte Auflösung der Inversion ohne Strafterm. Die Anpassung scheint, bis auf den negativen Geschwindigkeitsgradienten in der ersten Schicht aber vernünftig.

8.4 Statistische Beurteilung der Inversionslösungen

Der folgende Abschnitt soll in ein weiteres statistisches Verfahren zur Parameterauflösungs- und insbesondere Korrelationsanalyse einführen. Das Verfahren beruht auf einer Gewichtung von Parametervarianzen zu einem Erwartungsmodell mit einer von der Qualitätsfunktion abhängigen Wahrscheinlichkeit. Die vorgestellte Methode bietet neben der Bestimmung eines mittleren Modells aus verschiedenen Inversionsdurchgängen eine weitere Möglichkeit zur Analyse der Auflösung unter Berechnung von Varianzen und Kovarianzen der einzelnen Inversionsparameter. Das Verfahren nutzt hierbei die Informationsgrundlage vieler ausgewerteter Modelle in der Umgebung der Lösung. Es kann so eine erwartete Lösung angeben und zu erwartende Varianzen durch eventuelle Nebenminima in der Umgebung.

Hierzu sei zunächst ein Beispiel betrachtet: Vogel (1995) erhält die Boltzmannverteilung als Maß der Wahrscheinlichkeit eines diskreten Zustandes eines Vielteilchensystems aus der Analogie zur Wahrscheinlichkeit der Auswahl einer bestimmten Buchstabenkomposition aus einer zufälligen Menge von Buchstaben. Es besteht hierbei die Frage nach der Wahrscheinlichkeit für das zufällige Entstehen einer bestimmten Folge. Die Komposition der Buchstaben erfülle folgende Bedingung:

- Die Zeilenbreite B sei konstant (analog zum Zustand eines abgeschlossenen Systems/Parameterraum). Gesucht werden also nur diejenige Anzahlen n_i der Breite b_i (ein Analogon zum Energieinkrement), so daß

$$\sum_i n_i b_i = B.$$

Die Wahrscheinlichkeit einer Komposition sei darüber hinaus gegeben durch,

$$P_{komp}(n_1, n_2, \dots) = \frac{N!}{n_1! n_2! \dots} p_1^{n_1} p_2^{n_2} \dots = \frac{N!}{\prod n_i!} \prod p_i^{n_i} \quad (8.1)$$

(siehe auch Abbildung 8.14) wobei die n_i die Anzahl des jeweiligen Symbols i sind. Es gilt weiterhin die Nebenbedingung

$$n_1 + n_2 + \dots = \sum_i n_i = N$$

Gesucht sei nun eine Komposition/Zustand, bestimmt durch die n_i , bei dem sich die Wahrscheinlichkeit P oder $\ln(P)$ (eines solchen Zustandes) unter einer kleinen Änderung ν_i der n_i nicht verändert. Gesucht ist also ein Gleichgewichtszustand. Es sei wie folgt vorgegangen:

8 METHODENVERGLEICH UND AUFLÖSUNG

- Wir identifizieren α durch die Normierungsbedingung $\sum n_i = N$, dann ist

$$\sum n_i = e^\alpha \sum e^{\beta b_i} = N$$

Es folgt

$$e^\alpha = \frac{N}{\sum p_i e^{\beta b_i}} =: \frac{N}{Z}$$

Z ist hierbei die sogenannte Zustandssumme, die den Gesamtzustand des Systems beschreibt. Es folgt für die Verteilung:

$$n_i = \frac{N}{Z} e^{\beta b_i}$$

die Boltzmann- oder kanonische (regelmäßige) Verteilung.

- β wird über die definierte Gesamtbreite B definiert:

$$B = \sum b_i n_i = \frac{N}{Z} \sum b_i p_i e^{\beta b_i}$$

und wird Verteilungsmodul genannt.

Da nun die p_i die a priori Häufigkeiten oder Wahrscheinlichkeiten für das Auftreten für den einzelnen Parameter (Buchstaben) sind, gilt bei Gleichverteilung:

$$p_i = 1/N$$

und somit

$$n_i = \frac{e^{\beta b_i}}{Z}$$

Die Breite b_i sei nun analog zur Energie/Ausdehnung eines Einzelzustandes, oder eben zu einer Potentialfunktion auf einem Parameterraum zu betrachten.

Man betrachte nun das Optimierungsproblem einer Inversion bezüglich der Analogie zu Partikeln in einem physikalischen System. Angenommen ein Festkörpersystem werde erhitzt, so seien alle Partikel zufällig verteilt in einer flüssigen Phase. Wird das Material langsam zu einem Festkörper abgekühlt, so ist die Wahrscheinlichkeit P , daß sich das System bei der Temperatur T im Zustand i mit Energie E befindet analog dem vorigen Abschnitt definiert mit Hilfe der Zustandssumme Z

$$P(E_i) = \frac{\exp(-E_i/k_B T)}{Z} = \frac{\exp(-E_i/k_B T)}{\sum_j \exp(-E_j/k_B T)}$$

Die Frage ist also, wie groß ist die Wahrscheinlichkeit P , daß sich ein Teilchen i im Zustand E_i befindet. Für jedes Ensemble (und damit auch System) gibt es

eine wohldefinierte Zustandssumme, und diese Zustandssumme enthält alle Informationen über das betrachtete System. Summiert wird über alle verfügbaren Energieniveaus des Systems, womit eine „Energiesumme“ entsteht. Der von der Temperatur abhängige Zahlenwert der Zustandssumme für ein bestimmtes System gibt also an, wie viele Zustände dem System thermisch „zugänglich“ sind, also bei der gegebenen Temperatur besetzt sein werden. Sei nun die Energie der Wert des Anpassungsresiduums und T bringe die Information des Fehlers der Daten. Bei einer Inversion wäre es somit wünschenswert, daß nur **ein** Zustand sich als besetzbar erweist, welcher der eindeutigen Lösung entspricht; die Zustandssumme wäre somit klein. Sollte eine Vieldeutigkeit auftreten, so vergrößert sich die Wahrscheinlichkeit für weitere Zustände und die Zustandssumme wächst. Es sei nun angenommen, daß die Elemente des Parameteraumes (Ein Modell wäre ein Ensemble jener Elemente) wie obig beschrieben Boltzmanverteilt sind mit der sogenannten Gibbs'schen Verteilung (Martinez, 2000)

$$P(m = m_i) = \frac{\exp(-E(m_i))}{\sum_j \exp(-E(m_j))}$$

wobei $P(m = m_j)$ die Wahrscheinlichkeit für ein Modell $m_j \in D (D \subset \mathbb{R}^n : \text{Parameterraum})$ ist, so übernimmt die Zielfunktion die Rolle der Zustandsenergie. Der wahrscheinlichste Zustand wäre somit die Lösung des Inversionsproblems und entspräche dem globalen Minimum der Zielfunktion. Unter Berücksichtigung, daß sehr viele Modelle den Parameterraum in seiner Gesamtheit aller möglichen Zustände beschreiben, wäre eine Berechnung aller Zustandsenergien bzw. Zielfunktionswerte nicht empfehlenswert. Es läßt sich also die Wahrscheinlichkeit für die optimale Lösung nicht angeben. Es wird aber oft ein (stochastisches) Optimierungsverfahren angewendet, um die optimale Lösung bzw. den wahrscheinlichsten Zustand zu finden. So kann im Anschluß an eine Inversion die große Zahl der bereits generierten Modelle im Nahbereich der Lösung (Modelle fortgeschrittener Iterationen) genutzt werden, um diesen Parameterbereich und seine Zustandssumme approximativ zu repräsentieren. Mit diesen Modellen können dann über die Gibbs'sche Wahrscheinlichkeit ein Erwartungsmodell und Abweichungen und Korrelation über die Kovarianzen bestimmt werden. Die Varianzen der Parameter sind hierbei abhängig davon, wie viele Zustände ähnlich gewichtet, entartet sind, also wieviele Nebenminima es in einer Umgebung der Lösung gibt. Oft wird im Anschluß an eine Inversion eine Zufallssuche im Nahbereich der Lösung durchgeführt, um die Zustandssumme zu approximieren (Martinez, 2000). Durch die hohe Exploration und die Anzahl an Schwarmpartikeln kann dies analog durch die einfache Nutzung aller Schwarmpositionen aus fortgeschrittenen Iterationsschritten passieren. Dies bietet die Möglichkeit einer Bewertung der Inversionslösung, ohne die Durchführung vieler Inversionen, um ein tatsächliches mittleres Modell und Standardabweichung zu bestimmen.

Das Erwartungsmodell berechnet sich mit

8 METHODENVERGLEICH UND AUFLÖSUNG

$$\langle m \rangle = \frac{1}{N} \int_D m \xi(m) dD, \quad (8.2)$$

die Kovarianzmatrix ist,

$$C_m = \frac{1}{N} \int_D (m - \langle m \rangle)(m - \langle m \rangle)^T \xi(m) dD, \quad (8.3)$$

wobei $\xi(m)$ die Wahrscheinlichkeitsdichtefunktion ist,

$$\xi(m_j) = \exp[-E(m_j)] \quad (8.4)$$

und N die Zustandssumme.

$$N = \sum_i \exp[-E(m_i)]. \quad (8.5)$$

Die Diagonalelemente der Kovarianzmatrix entsprechen den Varianzen der einzelnen Parameter und ihre Quadratwurzel damit der Standardabweichung. Die Nichtdiagonalelemente bieten die Kovarianzen verschiedener Parameter zueinander. Sie sind entweder positiv, wenn zwei Parameter korreliert variieren oder negativ, wenn sie invers korreliert variieren. Mit diesen Informationen lassen sich somit Korrelationen zwischen Parametern erkennen, welche für weitere Unterbestimmtheit des Problems sorgen. Da die invertierten Parameter in ihrer Dimension variieren sei zunächst die normierte Korrelationsmatrix aus der Kovarianzmatrix mit

$$\tilde{C}_{i,j} = \frac{C_{i,j}}{\sqrt{C_{i,i}C_{j,j}}}$$

berechnet.

Berechnet man nun als Beispiel das Erwartungsmodell und die Kovarianzen für die Inversion des FINO3-Referenzmodells mittels der Modelle des zehnten bis zwanzigsten Iterationsschrittes bei allen Neustartdurchläufen eines Inversionsdurchganges, so ergeben sich das in Abbildung 8.15 dargestellte Erwartungsmodell, Standardabweichungen und Korrelationsmatrix. Es sei zunächst die Korre-

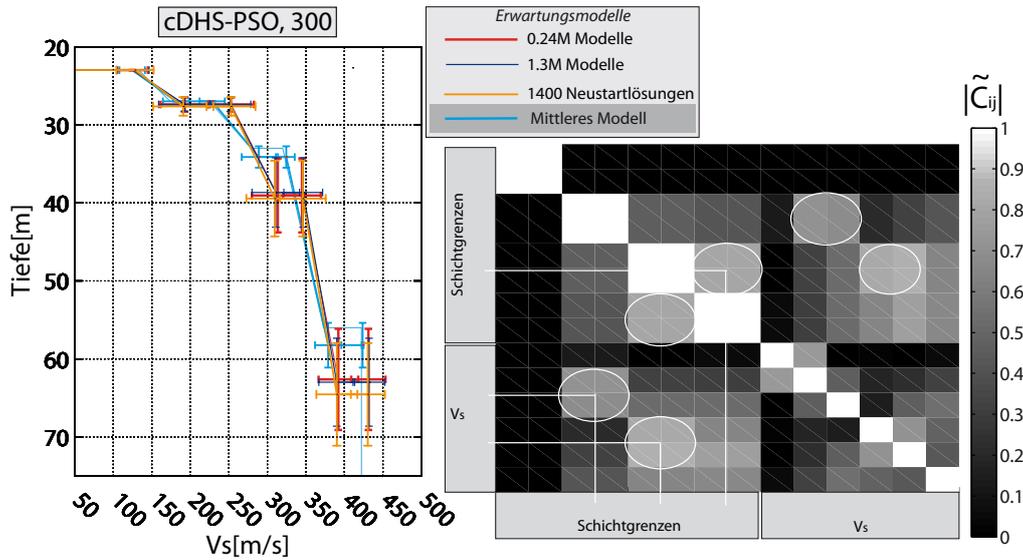


Abbildung 8.15: Darstellung des Erwartungsmodells und der Standardabweichung für Modelle der jeweils zehnten bis zwanzigsten Iteration von 16 Inversionsdurchgängen (blau), 3 Inversionsdurchgängen (rot) und ausschließlich aller Neustartlösungen von 16 Inversionsdurchgängen (grün). Zum Vergleich ist das tatsächliche mittlere Modell von 56 Inversionen eingetragen (hellblau). Unten rechts ist die Korrelationsmatrix farbkodiert angezeigt. Die weißen Ellipsen indizieren Bereiche hoher Korrelation.

lationsmatrix betrachtet. Die Hauptdiagonalelemente sind hier nicht auszuwerten. Die eindeutige Korrelation der Schichtgrenzen in Abbildung 8.15 neben der Hauptdiagonalen ist darauf zurückzuführen, daß die eigentlichen Inversionsparameter die Schichtmächtigkeiten sind, so daß obere und untere Schichtgrenze über die Mächtigkeit korrelieren. Neben diesem Artefakt sind jedoch zwei Bereiche hoher Korrelation zu erkennen (unteres linkes Quartier). Es zeigt sich hier eine relativ starke Korrelation der Geschwindigkeiten und Schichtmächtigkeiten im mittleren Bereich des Modells. Dieses bedeutet einen Trade-Off zwischen Mächtigkeit der Schicht und dem Verlauf des Geschwindigkeitsgradienten, den auch schon Bohlen et al. (2004) beobachteten.

Zu den Erwartungsmodellen in Abbildung 8.15 links sei zu erwähnen, daß zunächst die Erwartungsmodelle und Standardabweichungen für den Fall von zugrunde liegenden 1.3 Millionen Modellen der zehnten bis zwanzigsten Iteration aller Neustartinversionen von insgesamt 16 kompletten Inversionsdurchgängen und von 0.24 Millionen Modellen von drei Inversionsdurchgängen vergleichbar sind. Werden nur einzelne Inversionsdurchgänge als Basis gewählt, so zeigen sich stärkere Abweichungen. Eine Datengrundlage von Modellen von drei Inversionsdurchgängen erscheint hier also als vernünftige Repräsentation der untersuchten

8 METHODENVERGLEICH UND AUFLÖSUNG

Umgebung. Des weiteren liegt zum Vergleich das Erwartungsmodell berechnet aus den 1400 Lösungen aller Neustarts der PSO von 16 Inversionsdurchgängen vor. Dieses Modell ist den anderen ebenfalls nahezu gleich und läßt damit die Wahl der Modelle aus der zehnten bis zwanzigsten Iteration vorerst als vernünftig aussehen. Im Vergleich zum ebenfalls dargestellten mittleren Modell der 56 besten Lösungen von 56 Inversionen zeigt sich jedoch ein starker Unterschied in der Auflösung und Anpassung. Die Auflösung des Erwartungsmodells repräsentiert hier den Bereich **aller** konvergierten Lösungen aller Neustarts. Das mittlere Modell berechnet sich aber aus den nach Qualitätsmaß ausgewählten besten Lösungen der jeweiligen kompletten Inversionsdurchgänge. Die dargestellten Erwartungsmodelle berücksichtigen diese Auswahl nicht. Da aber das ausgewählte beste Modell eines Inversionsdurchganges gleich der Lösung ist, gilt es, seinen Nahbereich zu untersuchen.

Wenn also die Auswertung der besagten 10 Iterationen nur für den Bereich des ausgewählten besten Modells eine ähnliche Anpassung wie die vom mittleren Modell der 56 Inversionen zeigt, so ist die Auswahl dieses Parameterbereichs sinnvoller. So ließe sich mit nur einem Inversionsdurchgang ein vernünftiges Maß für die Abweichungen eines Lösungsmodells gewinnen. Abbildung 8.16 zeigt diesen Vergleich von mittlerem Modell der 56 Inversionen (rote Kurve) und Erwartungsmodell aus den Modellen der zehnten bis zwanzigsten Iteration der Anpassung des besten Modells eines Inversionsdurchganges (schwarze Kurve). Das erhaltene Erwartungsmodell liegt im durch die tatsächliche Auflösung vorgegebenen Bereich und zeigt vernünftige Anpassung. Die Standardabweichungen sind jedoch zu klein um repräsentativ zu sein. Hier wurde nun offenbar eine zu kleine Umgebung eines Nebenminimums gewählt. Es liegt daher nahe, die zu Grunde liegenden Modelle auf weitere Inversionsdurchgänge zu erweitern. Die blaue Kurve zeigt daher das Erwartungsmodell bestimmt aus den Iterationen zu drei ausgewählten besten Modellen dreier Inversionsdurchgänge. Dieses Modell zeigt eine mit dem mittleren Modell vergleichbare Anpassung, so daß bereits mit der Auswertung von drei Inversionsdurchgängen von einem repräsentativen Erwartungsmodell gesprochen werden kann.

Die Standardabweichungen des Erwartungsmodells für alle konvergierten Lösungen (siehe Abbildung 8.15, grüne Kurve) ist allerdings erneut ein Maß für die hohe Multimodalität des Problems.

Diskussion zu den drei Fragen an die Inversion

Es konnte in diesem Kapitel durch Analyse verschiedenster mittlerer Modelle von Inversionsergebnissen unterschiedlicher Verfahren gezeigt werden, daß die PSO in Kombination mit einer neuen Berechnung der Zielfunktion eine sehr gute algorithmische Auflösung bietet und dies bei sehr großem Parametersuchraum. Mit der modenfreien Zuordnung der Dispersionsdatenpunkte durch eine Methode der nächsten Nachbarn konnte erfolgreich eine neue Möglichkeit zur Quali-

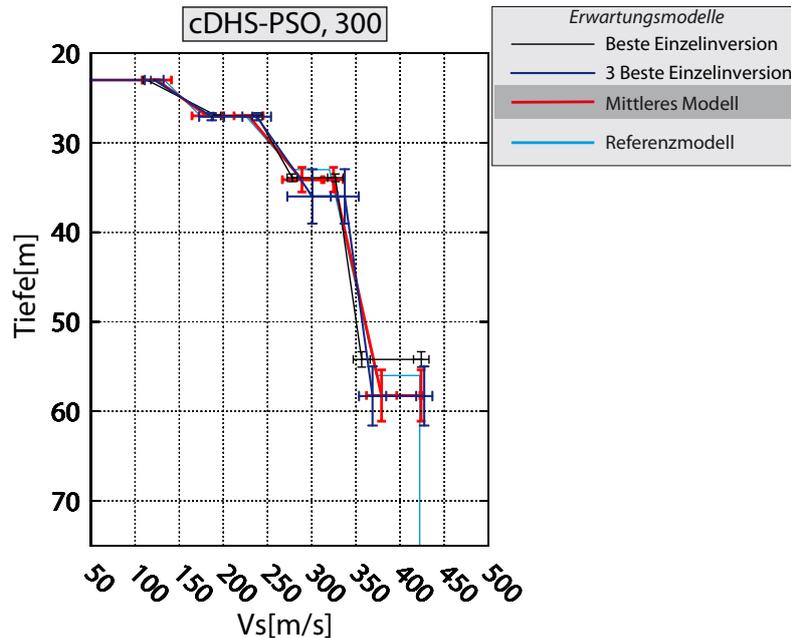


Abbildung 8.16: Darstellung des Erwartungsmodells und der Standardabweichung für Modelle der jeweils zehnten bis zwanzigsten Iteration vor Erreichen des besten Lösungsmodells von einem Inversionsdurchgang (schwarz), drei Inversionsdurchgängen (dunkelblau). Zum Vergleich ist das tatsächliche mittlere Modell von 56 Inversionen eingetragen (rot).

tätsfunktionsberechnung und zur optimierten Anpassung implementiert werden. Der Vergleich verschiedener Schwarmintelligenzvarianten offenbarte den gradientenfreien Hybrid aus Downhill-Simplex für lokale Konvergenz und geladener, atomarer PSO zur globalen Exploration und Konvergenz als effizientestes Verfahren mit einer algorithmischen Auflösung in der Scherwellengeschwindigkeit von 6.5%. Im Vergleich mit anderen gängigen Optimierungsverfahren zeigt sich eine klare Überlegenheit der cDHS-PSO. Die eingangs gestellte erste Frage nach der algorithmischen Auflösung konnte so mit 6.5% Standardabweichung bei nicht verrauschten Daten beantwortet werden. Die zweite Frage nach dem Einfluß von Fehlern auf die Auflösung wurde ebenfalls in einer Studie an synthetischen Daten untersucht. Es konnte eine Verschlechterung der Auflösung auf im Mittel 8% bei gewichtet bewerteten Eingangsdaten gezeigt werden. Die cDHS-PSO ist somit eine vielversprechende Optimierungsmethode für das vorliegende Problem und besitzt die algorithmisch-, technische Potenz, die abgeschätzte erforderliche Auflösung zur Baugrundüberwachung an der Plattform FINO3 zu gewährleisten. Eine Analyse der Kovarianzen der Inversionsparameter auf Grundlage sehr vieler ausgewerteter Modelle mehrerer Inversionsdurchgänge konnte ebenfalls erfolgreich durchgeführt werden. Es zeigte sich ein Zusammenhang zwischen Mächtigkeit

8 METHODENVERGLEICH UND AUFLÖSUNG

keit der Schichten und dem Verlauf des Geschwindigkeitsgradienten in mittleren Tiefen des untersuchten Modells, welcher erwartungsgemäß zu weiterer Mehrdeutigkeit von Lösungen führt. Hier gilt es nun, die Mehrdeutigkeit durch A priori Informationen aus reflexionsseismischen Daten, über das Verhältnis von Schichtmächtigkeiten oder aber die Kohärenz zu Nachbarmodellen in einer 2D Modellsektion zu verringern.

Teil III

9 Auflösung einer gemeinsamen Inversion von 1D-Reflektivitäten und Dispersionskurven

Die Aufgabe der Lösung von Mehrzieloptimierungsproblemen stellt sich bei vielen geophysikalischen Inversionsproblemen. So sind oftmals unterschiedliche Meßmethoden für gleiche Untergrundparameter sensitiv. Die Anpassung verschiedener Datensätze liefert mehr Daten für die Inversion und kann so Auflösungen verbessern und Mehrdeutigkeiten verringern. Auch können die selben Parameter in unterschiedlicher Konstellation oder durch Parameter Trade-Off in verschiedenen Verfahren Sensitivitäten aufweisen. Sogar die Zusammenlagerung (Clustering) von Parametern anhand von Strukturinformationen oder in einer Art Expertensystem kann Grundlage zur gemeinsamen Inversion von Datensätzen sein. So setzt sich auch das seismische Wellenfeld aus verschiedenen Wellentypen mit zum Teil unterschiedlichem Frequenzgehalt zusammen. Diese unterschiedlichen Bereiche des Wellenfeldes zeigen Sensitivitäten für unterschiedliche Parameter, können aber über Struktur oder Dichte miteinander gekoppelt sein. Eine gemeinsame Inversion verschiedener Wellentypen kann somit eine bessere Auflösung des Untergrundmodells liefern, denn es wird eine größere Datengrundlage bereitgestellt bei eventuell gemeinsamen Inversionsparametern. So ist die Bestimmung von Reflektivitäten aus Amplituden/Entfernungs-Informationen (Castagna und Backus, 1993) sensitiv für die Struktur des Untergrundes, und die seismische Impedanz, also Produkt aus Kompressionswellengeschwindigkeit und Dichte. Die Inversion von Scholtewellen zeigt hohe Sensitivität ebenfalls für die Struktur sowie für die Scherwellengeschwindigkeiten. Geringe Sensitivität zeigt sich für die Dichte (Kugler et al., 2006) und nur geringe für die Kompressionswellengeschwindigkeit. Eine gemeinsame Anpassung von Datensätzen aus beiden Methoden auf Grundlage eines Parameterraumes D mit $\vec{X} = (\vec{v}_p, \vec{v}_s, \vec{\rho}, \vec{M}) \in D$ ($\vec{v}_{p/s}$: Vektor der Schichtgeschwindigkeiten; $\vec{\rho}$: Vektor der Schichtdichten; \vec{M} : Vektor der Schichtmächtigkeiten) kann Ergebnisse für alle Parameter liefern sowie Mehrdeutigkeiten in den einzelnen Methoden verringern.

Ziel dieses Abschnittes ist die gemeinsame Anpassung von synthetischen Reflektivitätsdaten einer Zero-Offset Sektion und den Dispersionskurven von Scholtewel-

9 GEMEINSAME INVERSION

len, zur Erzeugung von eindimensionalen Modellen $\vec{X} = (\vec{v}_p, \vec{v}_s, \vec{\rho}, \vec{M})$ (Schema siehe dazu in Abbildung 9.1).

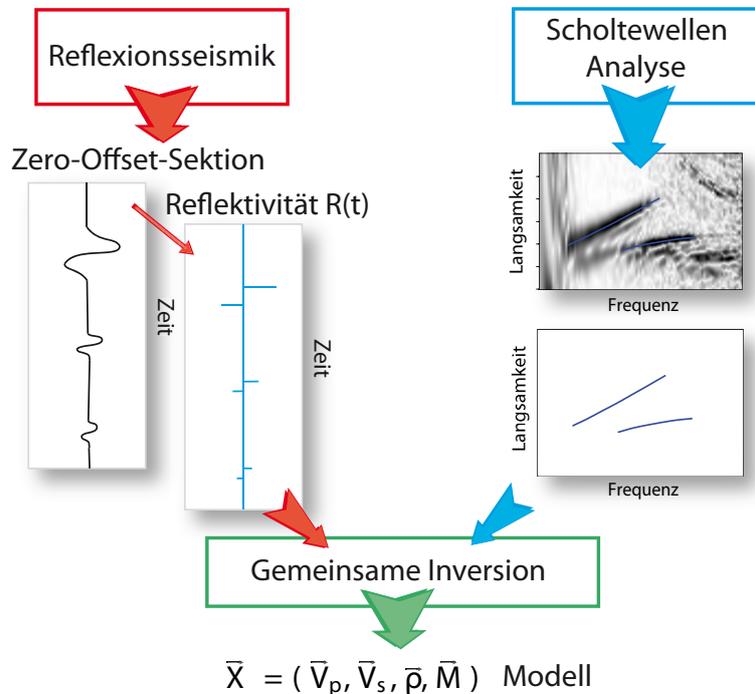


Abbildung 9.1: Schema zur gemeinsamen Inversion von Reflektivitätsspektren und Scholwellendispersionskurven.

9.1 Optimierung multipler Zielfunktionen

Die gemeinsame Optimierung mehrerer Funktionen stellt kein einfaches Problem dar. Ein simultanes Optimum aller Funktionen muß neu definiert werden, da nicht jede Funktion den gleichen Lösungsvektor haben muß. Ein großes Problem hierbei sind mögliche Konflikte zwischen den Zielfunktionen, wenn Verbesserung des einen Ziels eine Verschlechterung des anderen mit sich bringt. Es können Zusammenhänge zwischen den Zielfunktionen bestehen, die eine eindeutige Lösung nicht zulassen. Das Ausbalancieren dieser Zusammenhänge mit dem Suchen einer Lösung, welche die effektivste Balance oder den besten Kompromiß im Hinblick auf beide Zielfunktionen darstellt, sollte damit Ziel einer Optimierung sein. Einfache Ansätze wie das Überführen beider Zielfunktionen in eine einzelne durch Addition und Einführen eines Gewichtungsfaktors lassen dabei die Frage nach der Wahl dieses Gewichtungsparameters offen. Es müßten mehrere Optimierungen durchgeführt werden, um den optimalen Gewichtungsfaktor über eine L-Kurven- oder

9.1 Optimierung multipler Zielfunktionen

Trade-Off-Analyse anzupassen (siehe Abbildung 9.2).

Das Problem multipler Zielfunktionen sei nun wie folgt definiert.

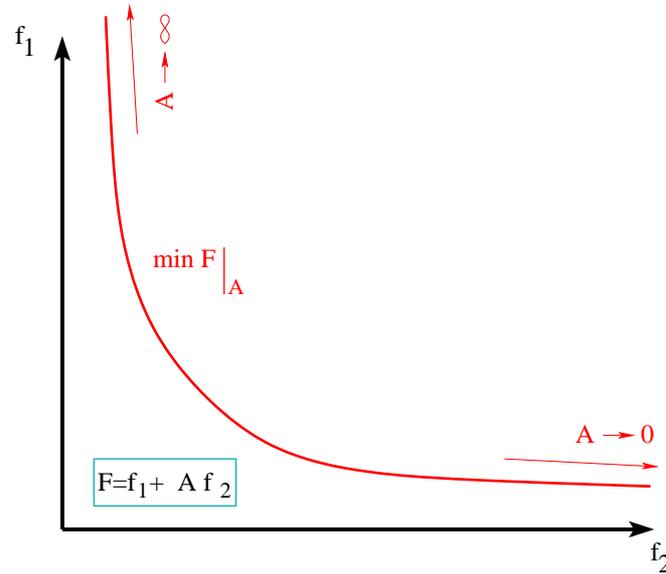


Abbildung 9.2: Konzept einer Trade-Off Analyse zur Bestimmung des Gewichtungsparameters A der gemeinsamen Optimierung der Funktionen f_1, f_2 mit $F = f_1 + A \cdot f_2$. Das Auftragen der Optimierungsergebnisse für verschiedene A ermöglicht die Wahl des Optimierungsergebnisses. Eine optimale Wahl kann das Paar f_1, f_2 am Punkt maximaler Krümmung der L -Kurve sein.

Minimiere $\vec{f}(\vec{x})$

mit Nebenbedingungen $g_i(\vec{x}) \leq 0, i = 1, \dots, n_g$ und $h_i(\vec{x}) = 0, i = n_g+1, \dots, n_g+n_h$, wobei $\vec{f}(\vec{x}) = (f_1(\vec{x}), f_2(\vec{x}), \dots, f_{n_k}(\vec{x}))$ und $\vec{x} = x_1, x_2, \dots, x_{n_x}$. Die Bedeutung des Optimums muß hierbei neu definiert werden, denn Zielfunktionen können in umgekehrt proportionaler Form in Konflikt stehen und voneinander abhängen. Man betrachte hierzu einmal ein Beispiel

$$\begin{aligned}\vec{f}(\vec{x}) &= (f_1(\vec{x}), f_2(\vec{x})) \\ f_1(\vec{x}) &= \frac{1}{n_x} \sum_{j=1}^{n_x} x_j^2 \\ f_2(\vec{x}) &= \frac{1}{n_x} \sum_{j=1}^{n_x} (x_j - 2)^2\end{aligned}$$

welches keine eindeutige Lösung besitzt. Das Ziel einer Optimierung multipler Zielfunktionen ist also das Finden eines Lösungssatzes, welcher einen guten Kompromiß repräsentiert. Dieser Lösungssatz sei als Satz nicht-dominierter Lösungen oder *Pareto-Optimalsatz* definiert (Engelbrecht, 2005).

Pareto-Dominanz

Die Dominanz einer Lösung gegenüber einer anderen ist definiert als: Ein Lösungsvektor \vec{x}_1 dominiert einen anderen Lösungsvektor \vec{x}_2 , genau dann wenn gilt

- \vec{x}_1 ist nicht schlechter als \vec{x}_2 in allen Zielfunktionen; oder

$$f_k(\vec{x}_1) \leq f_k(\vec{x}_2), \forall k = 1, \dots, n_k$$

und

- \vec{x}_1 ist ungleich kleiner als \vec{x}_2 in wenigstens einer Zielfunktion; oder

$$\exists k = 1, \dots, n_k : f_k(\vec{x}_1) < f_k(\vec{x}_2)$$

Damit wäre ein Vektor \vec{x}^* Pareto-Optimal, wenn kein Vektor existiert, der \vec{x}^* dominiert. Der Pareto-Optimalsatz P^* ist dann jener Satz von Lösungen \vec{x}_i^* , die nicht dominiert sind und somit den Trade-Off zwischen den Zielfunktionen ausbalanciert haben.

$$P^* = \{\vec{x}^* \in F : \nexists \vec{x} \text{ dom. } \vec{x}^*\}$$

mit F Parameterraum. Die Pareto-Optimalfront, welche im Zielfunktionsraum definiert ist, sei dann die Menge aller Zielfunktionsvektoren, die nicht dominiert sind.

$$PF^* = \left\{ \vec{f} = (f_1(\vec{x}^*), f_2(\vec{x}^*), \dots, f_k(\vec{x}^*) : \vec{x}^* \in P^* \right\}$$

(Engelbrecht, 2005)

9.2 Multi-Objective-PSO

Es gibt verschiedene Ansätze für die Implementierung einer PSO zur Optimierung mehrerer Zielfunktionen (siehe dazu Reyes-Sierra & Coello, 2006). Die Ziele aller Ansätze seien hierbei

1. Die Minimierung des Abstandes zwischen den Lösungen und der Pareto-Optimalfront.
2. Die Maximierung der Diversität der Lösungen auf der Optimalfront, um so viel der Front zu repräsentieren wie möglich.
3. Der Erhalt bereits gefundener nicht-dominiertes Lösungen.

(Engelbrecht, 2005). Die verwendete Variante der PSO sollte somit auf dem Prinzip der Pareto-Dominanz basieren. Zu diesem Zweck werden Archive im Algorithmus der PSO eingeführt, um auf alle erreichten nicht-dominierten Lösungen im Optimierungsprozeß zurückgreifen zu können. Die Archive werden dann genutzt,

um eine neue Form des individuell „besten“ und des global „besten“ Partikels zu wählen. Die ausgewählten nicht-dominierten Partikel dienen somit als Anführerpartikel, welche den Schwarm in Richtung der Pareto-Optimalfront führen sollen. Die verbleibende Aufgabe für einen einfachen Multi-Objective-PSO Ansatz ist die Wahl des global besten Partikels. Hierzu wurde in dieser Studie auf das Auswahlkriterium nach Kalyanmoy et al. (2002) zurückgegriffen. Das Grundprinzip dieser Auswahlmethode, der nächster-Nachbar Abschätzung, ist in Abbildung 9.3 dargestellt. Die Abschätzung gibt Auskunft darüber, wie dicht liegen die nächsten Nachbarn eines gegebenen Partikels im Raum der Qualitätsmaße. Bezogen auf obigen Punkt 2 würde die Wahl des global besten Partikels aus dem Archiv aller nicht-dominierten Lösungen nach der geringsten Punktdichte im Raum der Qualitätsmaße ausfallen. Hierzu wird der Flächeninhalt eines Rechtecks im Zielfunktionsraum mit den nächsten Nachbarn eines Partikels als Eckpunkte als Maß verwendet. Die Wahl des Partikels mit maximalem Flächeninhalt sorgt so für die Diversität der Pareto-Front. Der verwendete Algorithmus basiert nun

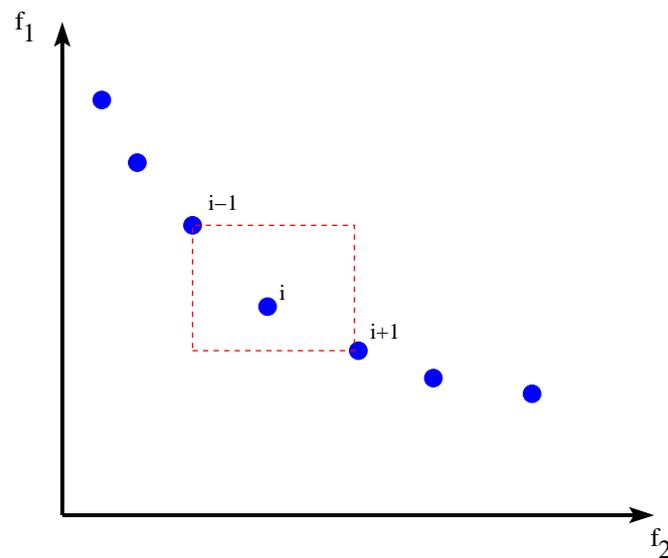


Abbildung 9.3: Darstellung des nächster Nachbar Konzeptes zur Evaluation des besten Anführerpartikels.

auf dem von Coello & Lechuga (2002) vorgestellten Grundkonzept, jedoch mit nächster-Nachbar Auswahlmethode, und ist wie folgt aufgebaut.

- Definiere Zielfunktionen f_1 (Abstand Dispersionskurven), f_2 (Unterschied Reflektivität (siehe folgender Abschnitt)).
- Initialisiere Partikel und evaluiere f_1 und f_2 .
- Initialisiere die individuell beste Position der Partikel $\vec{y}_K i = \vec{x}_i$.

9 GEMEINSAME INVERSION

- Definiere und speichere ein Anführer-Depot (Repository oder REP) und seine Größe (Das Depot beinhalte die Gruppe aller nicht-dominierten Lösungen).
- Wähle einen Anführer aus REP mittels der nächster-Nachbar Abschätzung und setze ihn als \vec{y}_G .
- Vollführe Schwarmbewegung unter Benutzung von \vec{y}_G und $\vec{y}_{K i}$.
- Aktualisiere das REP (alle neuen Vektoren werden mit der Anführerliste verglichen, ob sie nicht-dominierte Lösungen sind; Prüfe anschließend alle Einträge in REP auf Dominanz untereinander, um Dominanzen durch neue Mitglieder auszuschließen.)
- Aktualisiere $\vec{y}_{K i}$, wenn das alte $\vec{y}_{K i}$ durch die neue Position der Partikels dominiert wird. Ist das nicht der Fall behalte $\vec{y}_{K i}$. Wenn keines der beiden dominiert, wähle zufällig.
- Ergebnis des Verfahrens: Das REP enthält alle nicht-dominierten Lösungen.

Die zweite Zielfunktion: Reflektivität

Ausgehend vom Modell $\vec{X} = (\vec{v}_p, \vec{v}_s, \vec{\rho}, \vec{M})$ wird der Reflexionskoeffizient $R(\vec{X})$ für den senkrechten Strahlweg einer Zero-Offset Sektion durch die Serie fiktiver Schichten in Abhängigkeit von der Tiefe berechnet. Das Modell wird hierzu in Tiefenschritten Δz von einem Meter interpoliert.

$$R(z) = \frac{\rho(z+1) \cdot v_p(z+1) - \rho(z) \cdot v_p(z)}{\rho(z+1) \cdot v_p(z+1) + \rho(z) \cdot v_p(z)}$$

Die so erhaltene Reihe von $(z, R(z))$ -Punkten wird dann in den Zeitbereich $(t(z), R(t(z)))$ transformiert.

$$t(z) = \int_0^z \frac{dz}{v_p(z)} \approx \sum_{i=0}^{i=\frac{z}{\Delta z}} \frac{\Delta z}{v_p(i \cdot \Delta z)}$$

Die erhaltene Zeitreihe wird dann entsprechend des Samplings Δt der Vergleichsdaten, die in diesem Fall eine synthetische Reflektivitätenreihe eines Referenzmodells sind, interpoliert.

Als Maß der Anpassung und damit zweite Zielfunktion wurden zwei Varianten gewählt und verglichen.

Zum einen wird der Korrelationskoeffizient der beiden Zeitreihen für modellierte Reflektivitäten, R_i^{pred} , und gemessene Reflektivitäten, R_i^{obs} , als Maß für den Grad des linearen Zusammenhangs, berechnet. Es ist hierbei $i = 1, \dots, t_{max}$,

$$t_{max} = \max(t^{pred}(z), t^{obs}(z)) / \Delta t$$

und $\bar{R} = \frac{1}{t_{max}} \sum_{i=1}^{t_{max}} R_i$.

$$C(R^{pred}(\vec{x}), R^{obs}) = \frac{\frac{1}{t_{max}} \sum_{i=1}^n (R_i^{pred} - \bar{R}^{pred})(R_i^{obs} - \bar{R}^{obs})}{\sqrt{\frac{1}{t_{max}} \sum_{i=1}^n (R_i^{pred} - \bar{R}^{pred})^2} \sqrt{\frac{1}{t_{max}} \sum_{i=1}^n (R_i^{obs} - \bar{R}^{obs})^2}}$$

$$f_2(\vec{x}) = |1 - C(R^{pred}(\vec{x}), R^{obs})|$$

Zum anderen wird ein einfacher Amplitudenvergleich verwendet.

$$f_2(\vec{x}) = \sqrt{\frac{1}{t_{max}} \sum_i (R_i^{pred}(\vec{x}) - R_i^{obs})^2}$$

Das Verhalten des Korrelationskoeffizienten und des Amplitudenvergleichs in Abhängigkeit von einer definierten Änderung addiert auf ρ oder v_p aller Schichten eines bekannten Vergleichsmodells ist in Abbildung 9.4 aufgezeigt. Für jeden Punkt der Abbildung wurde das Referenzmodell (siehe Tabelle 5.1 in Kapitel 5) in einer jeweils zufällig gewählten Schicht um einen zufälligen Wert aus dem Intervall $[-500, 500]$ m/s bzw. $[-0.5, 0.5]$ t/m^3 variiert und die Qualitätsfunktion f_2 in beiden Formen berechnet. Der Korrelationskoeffizient zeigt ein komplexeres Verhalten mit teilweise hoher Streuung im Vergleich zum wenig gestreuten, eine klarere Tendenz aufzeigenden Amplitudenvergleich (siehe Abbildung 9.4). Da der Zusammenhang zwischen den beiden Reflektivitätsreihen nicht linear ist (es handelt sich um keine einfache Translation und Amplitudenänderung der Reihen), erzielt die Korrelation als Maß offenbar ein schlechteres Ergebnis. In Abbildung 9.5

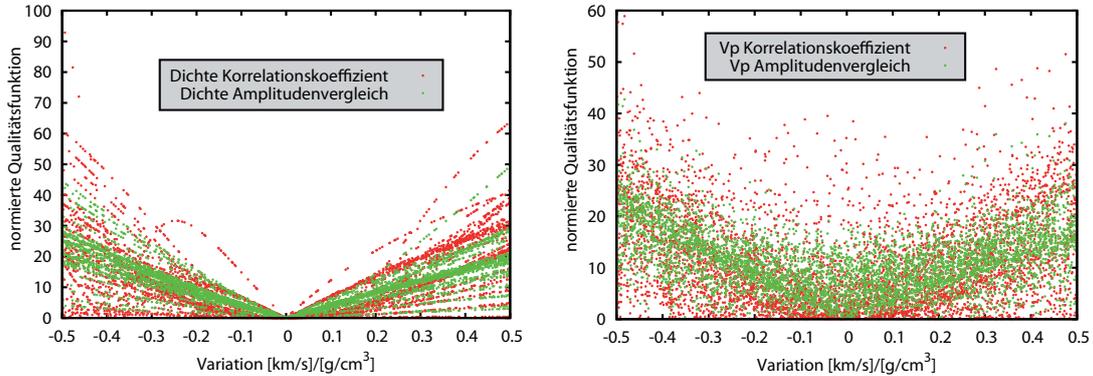


Abbildung 9.4: Vergleich von Korrelationskoeffizient und Amplitudenabstand für zufällige Variationen von v_p und ρ in zufällig gewählten Schichten des Referenzmodells.

ist zusätzlich eine Auflösungsanalyse für beide Vergleichsformen gezeigt, um zu klären welches Vergleichsmaß gewählt werden kann. Es wurden hierbei 56 Inversionsdurchgänge mit der cDHS-PSO mit 300 Partikeln für beide Anpassungsmaße

9 GEMEINSAME INVERSION

durchgeführt. Inversionsparameter waren die Schichtmächtigkeiten sowie jeweils einmal nur v_p , nur ρ oder v_p und ρ . Die Anpassung mittels Amplitudenvergleich führt in jedem Beispiel zu besserer Anpassung des mittleren Modells.

Der gravierende Verlust von Auflösung bei gemeinsamer Inversion von v_p und ρ zeigt die Sensitivität des Verfahrens gegenüber der seismischen Impedanz, also dem Produkt aus v_p und ρ . Im Fall der Zero-Offset Reflektivitäten sind die Faktoren dieses Produktes einzeln nicht auflösbar. Dies führt zur Uneindeutigkeit der Paare v_p und ρ und zur beobachteten schlechten Auflösung. In der Praxis der marinen reflexionsseismischen Exploration liegt aber im Allgemeinen ein Empfängersystem vor, welches eine offsetabhängige Analyse der Laufzeiten ermöglicht. Es werden Empfängersektionen in Streamern oder Ozeanbodenseismometer verwendet, um verschiedene Offsets zwischen Quelle und Empfänger zu realisieren. Dies ermöglicht eine einfache Bestimmung der Kompressionswellengeschwindigkeit (bspw. Korrektur der offsetabhängigen Laufzeitversätze von Einsätzen eines gemeinsamen Reflexionspunktes durch Anpassung eines Geschwindigkeitsmodells (Normal-Moveout-Korrektur) oder Anpassung der Einsätze von Mintropwellen (siehe bspw. Kearey et al., 2006)), so daß praktisch von einem bekannten v_p -Modell ausgegangen werden kann. Des Weiteren sind Dichte und Struktur die Hauptkopplungsparameter mit nennenswerter Sensitivität in beiden Verfahren. Es wird daher im folgenden Abschnitt die gemeinsame Inversion beider Datensätze mit den freien Inversionsparametern Dichte, Scherwellengeschwindigkeit und Schichtmächtigkeit versucht.

9.3 Gemeinsame Inversion eines synthetischen Datensatzes

Dieser Abschnitt beschreibt nun erste Ergebnisse einer gemeinsamen Inversion synthetischer Datensätze modellierter Reflektivitätsreihen und Dispersionskurven für ein bekanntes Untergrundmodell mit der MOPSO. Hierzu wird erneut das FINO3-Referenzmodell verwendet, welches bereits in Abschnitt 8.1 vorgestellt wurde. Darüber hinaus sind Parameterbereiche für die Dichte definiert, welche in Abbildung 9.6 und in Tabelle 9.1 gezeigt sind.

Es sind nun auf Grundlage dieser Testbedingungen Testinversionen mit der MOPSO durchgeführt worden. Es wurden 56 Inversionsdurchgänge gerechnet. Die MOPSO wurde mit einer Neustartmöglichkeit implementiert, so daß jeder Inversionsdurchgang wieder mehrere Einzeloptimierungen beinhaltet. Es wurden Rechnungen mit Populationsgrößen von 300 und 500 Partikeln durchgeführt. Abbildung 9.6 zeigt nun als Ergebnis das mittlere Modell und die Standardabweichungen für die Inversionsparameter M , v_s und ρ . Zur Berechnung wurden zunächst alle nicht dominierten Lösungen zusammengestellt und dann diejenigen Lösungen genutzt, deren Dispersionsanpassung unter dem Wert 0.3 und Reflekti-

9.3 Gemeinsame Inversion eines synthetischen Datensatzes

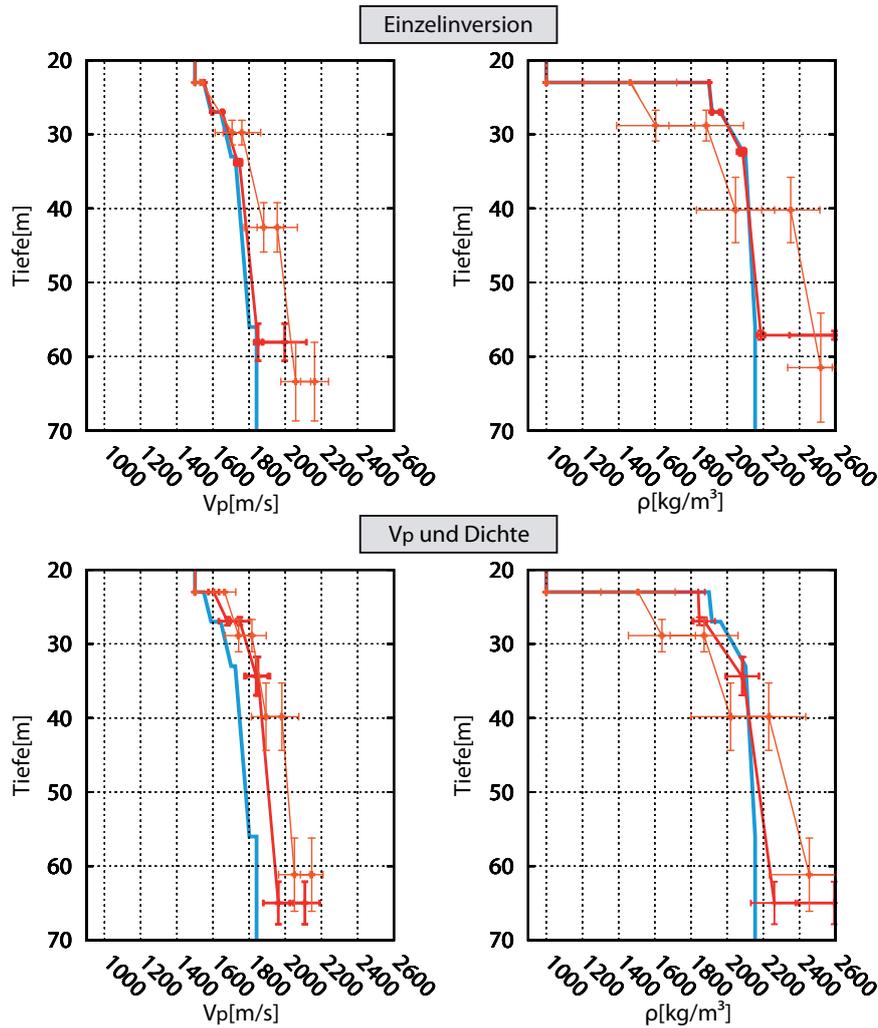


Abbildung 9.5: Darstellung der mittleren Lösungsmodelle und Standardabweichungen von jeweils 56 Inversionen unterschiedlicher Startpopulation für die cDHS-PSO mit Populationsgröße 300. Es sind jeweils beide Vergleichsmaße (rot: Amplitudenresiduen, orange: Korrelationskoeffizient, blau: Referenzmodell) dargestellt. Die oberen Bilder zeigen Ergebnisse von Einzelinversion von v_p oder ρ als Inversionsparameter. Die unteren Bilder zeigen Ergebnisse mit v_p und ρ als Inversionsparameter

vitätsanpassung unter 0.1 liegt. Anpassungswerte über diesen Werten können im Hinblick auf die Inversionen der einzelnen Datensätze als nicht genügend ausgeleitet werden.

Das MOPSO Verfahren zeigt zunächst vernünftige Anpassung für alle Parameter. Abbildung 9.6 zeigt ebenfalls den direkten Vergleich der mittleren Modelle und der Standardabweichung für Dichte und v_s von einfacher und gemeinsamer Inversion. Es zeigt sich im Vergleich zu den mit der cDHS-PSO erhaltenen Ergebnissen

9 GEMEINSAME INVERSION

$v_s \pm \Delta v_s$ [m/s]	M [m]	$\rho \pm \Delta\rho$ [kg/m ³]
0.0 ± 0.0	23.0	1000.0 ± 0.0
180.0 ± 100.0	1.0→10.0	1900.0 ± 400.0
250.0 ± 170.0	1.0→20.0	1950.0 ± 400.0
350.0 ± 270.0	10.0→30.0	2100.0 ± 400.0
450.0 ± 270.0	Halbraum	2150.0 ± 400.0

Tabelle 9.1: Parameterraumintervalle des Testmodells für Scherwellengeschwindigkeit, v_s , und Schichtmächtigkeit, M und Dichte, ρ .

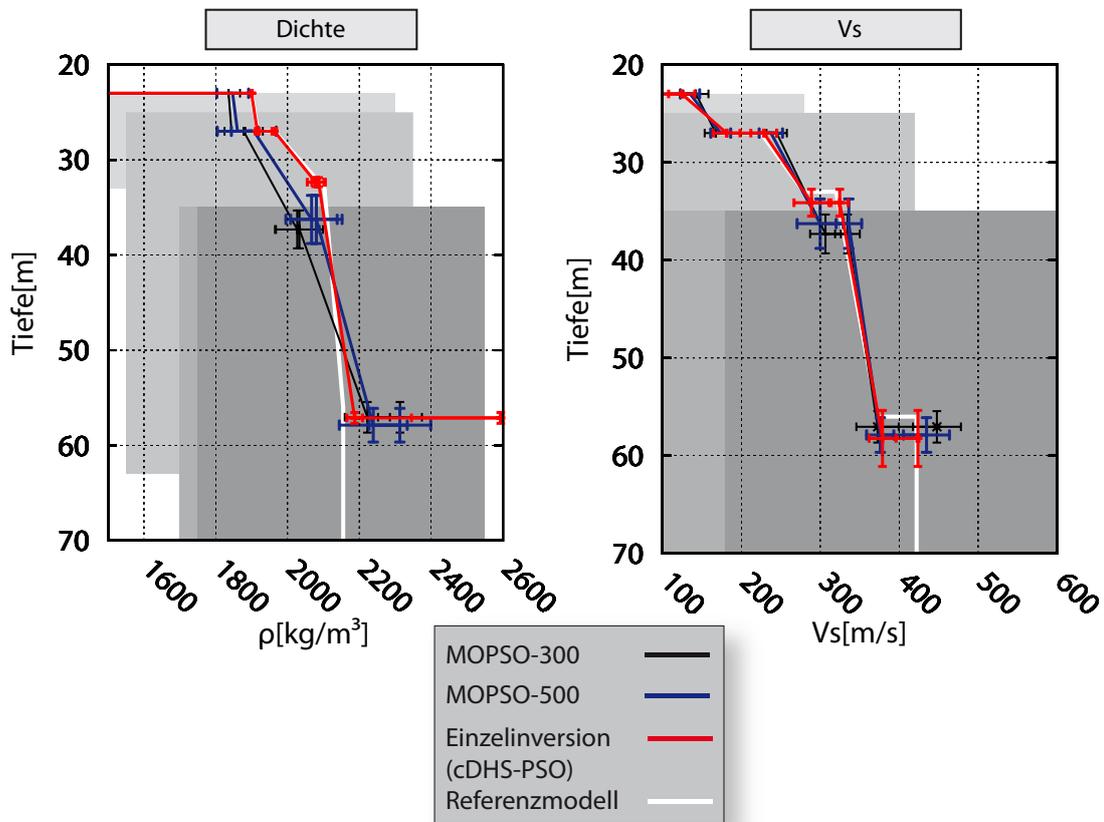


Abbildung 9.6: Darstellung der mittleren Lösungsmodelle und Standardabweichungen für ρ und v_s von jeweils 56 Inversionen unterschiedlicher Startpopulation für die MOPSO mit Populationsgröße 300 (schwarze Kurven) sowie 500 Partikeln (blaue Kurven) und für die Einzelinversionen mit der cDHS-PSO (300 Partikel, rote Kurven). In weiß dargestellt ist das Referenzmodell und in unterschiedlichen Graustufen die Parameterbereiche unterschiedlicher Tiefe.

9.3 Gemeinsame Inversion eines synthetischen Datensatzes

für alle Parameter eine schlechtere Anpassung durch die MOPSO. Dies läßt die Vermutung zu, daß trotz gemeinsamer Inversion und der Kopplung von Dichte und Schichtmächtigkeiten noch immer eine starke Multimodalität das Problem bestimmt. Die erhoffte Verminderung von Mehrdeutigkeiten und damit verbundene Verbesserung der Auflösung scheint somit nicht ausreichend, um mit dem einfachen MOPSO eine Anpassungsverbesserung zu erreichen. Die MOPSO basiert auf dem Konzept der einfachen Schwarmoptimierung und kann daher nicht die gleiche Effektivität im Ausbeuten möglichst vieler Minima vorweisen wie die optimierte cDHS-PSO. Diese Vermutung wird gestützt, betrachtet man im Vergleich die Anpassung durch die MOPSO mit 300 und 500 Partikeln in Abbildung 9.6. Eine Erhöhung der Partikelzahl bedingt stärkere Exploration des Parameterraumes und zeigt eine bessere Anpassung zum Referenzmodell. Es stellt sich daher an dieser Stelle die Aufgabe, das einfache MOPSO-Verfahren in seiner Exploration und lokalen Konvergenz zu optimieren. Dieses Unterfangen gestaltet sich jedoch als sehr schwierig, da Verfahren wie der Downhill-Simplex nicht nach dem Konzept der Pareto-Dominanz funktionieren. Abbildung 9.7 zeigt zusätzlich

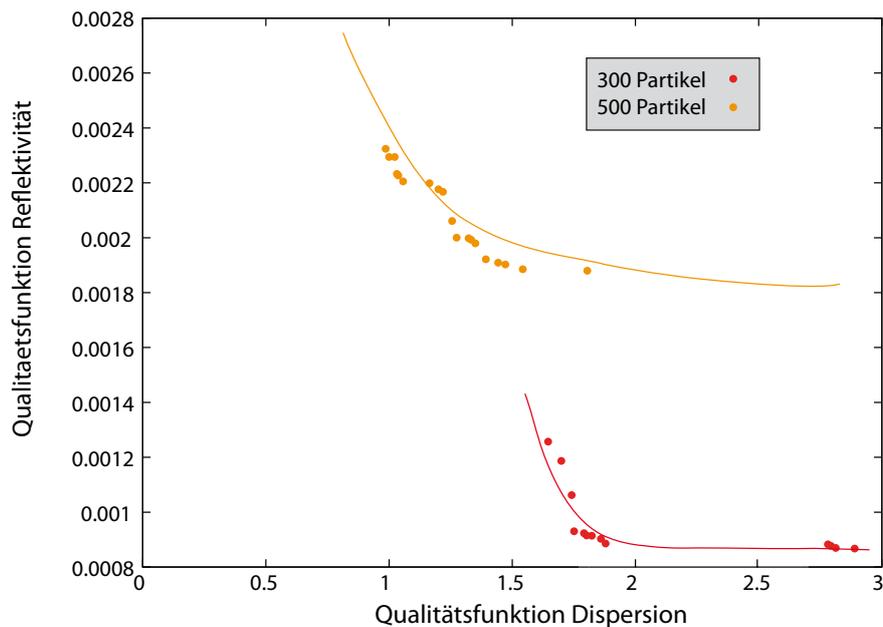


Abbildung 9.7: Darstellung der Paretofronten nicht dominierter Lösungen einer Beispielinversion mit 300 und 500 Partikeln.

die Paretofronten der nichtdominierten Lösungen eines beispielhaften Inversionsdurchganges für 300 und 500 Partikel. Hier kann jedoch keine eindeutige Verbesserung erkannt werden. Während der 500 Partikel Schwarm bessere Ergebnisse in der Dispersionskurvenanpassung zeigt, liefert der 300 Partikel Schwarm bessere Anpassung der Reflektivitätskurve.

9 GEMEINSAME INVERSION

Es sei daher als ein erster Versuch ein Pareto-Dominanz orientierter Downhill-Simplex vorgeschlagen, welcher die lokale Suche der MOPSO verstärken soll. Das Verfahren erstellt analog zum Standard-DHS einen Simplex um das aus dem Anführer-Depot gewählte global beste Partikel. Die erhaltenen $n+1$ -Simplexpunkte werden dann dahingehend geprüft, ob einer der Punkte, den Ausgangspunkt dominiert. Ist dies der Fall, so ist der dominierende Punkt das neue Zentrum des Simplex. Wird kein Simplexpunkt den Ausgangspunkt dominieren, so wird der Simplex zufällig in seiner Größe variiert. Dieses Verfahren wird nicht die Effizienz des Nelder-Mead-Algorithmus aufweisen, ist jedoch fähig, eine Dominanzbasierte lokale Suche durchzuführen. Es wurden daher 56 weitere Testinversionen des synthetischen Datensatzes durchgeführt mit einer Populationsgröße von 300 Partikeln und 15 Iterationen Dominanz-Downhill-Simplex pro Iterationsschritt. Die Ergebnisse des Tests zeigen jedoch keine nennenswerten Verbesserungen in der Anpassung der Modelle.

Abschließend betrachtet kann die gemeinsame Inversion von Reflektivitäten und Dispersionskurven von Scholtewellen mit der MOPSO keine besseren Ergebnisse liefern als die Einzelinversionen. Die noch immer erhaltene Multimodalität scheint nach wie vor das Problem zu dominieren. Es gilt also, weitere Verfahren und Verfahrensvarianten zu testen, die eine noch effektivere Exploration der vielen Minima dieses Problems ermöglichen.

Teil IV

10 Fallstudien: Inversion von Scherwellengeschwindigkeits-Modellen

In diesem Kapitel werden Inversionsergebnisse aus Daten von drei Meßkampagnen in Nord- und Ostsee sowie Auflösungsanalysen der Ergebnisse vorgestellt. Die Inversionen wurden jeweils mit dem für das Problem als effektivstes erkannten Schwarmoptimierungsverfahren (hybrides cDHS-PSO Verfahren) durchgeführt. Im Mai 2006 wurde eine Voruntersuchung des zukünftigen Baugrundes der FINO3-Plattform in der Nordsee durchgeführt. Die Messungen hatten das Ziel, einen 2 km mal 2 km großen Bereich (siehe Karte in Abbildung 10.1, Ausschnitt a)) im Hinblick auf geeignete Bauplätze für die Plattform zu untersuchen. Die Kampagne umfaßte eine hochauflösende 3D reflexionsseismische Studie des Gebietes sowie drei 2 km lange Profile auf denen Scholtewellen mit einer Airgun angeregt und mit Ozeanbodenseismometern registriert wurden (Wilken et al., 2009). Eine Kartierung von Diffraktionen in den Reflexionsdaten ergab eine Verteilung möglicher Gefährdungen bei der Gründung der Plattform durch Steine oder Findlinge. Die Spektralanalyse der erhaltenen Daten der drei Scholtewellenprofile ermöglichte eine Inversion von Pseudo-2D-Scherwellengeschwindigkeitsmodellen nach dem in den Abschnitten 3.1 und 4.1 beschriebenen Vorgehen. Die Optimierung wurde mit einer einfachen „global best“ PSO (Abschnitt 7.1) durchgeführt und die Berechnung der Qualitätsfunktion war noch abhängig von einer Modenidentifikation. Die Ergebnisse dieser Voruntersuchung enthielten neben der hochauflösenden Verteilung von Gefahrenbereichen mit hoher Anzahl an Diffraktionen auch Scherwellengeschwindigkeiten, die in Korrelation mit den dreidimensionalen Reflexionshorizonten eine Abschätzung über das Anstehen stärker kompakterer Sande zuließen. Die FINO3 wurde somit in einem Bereich gegründet, der wenig Diffraktionen enthielt und einen hohen Anteil vermeintlich stärker kompakterer Sande aufwies.

Im Mai 2007 wurde eine weitere Studie am gewählten Standort der FINO3 Plattform durchgeführt. Diese hatte das Ziel, die Scherwellengeschwindigkeitsverteilung im Nahbereich der Plattform (400 m) in drei Dimensionen zu bestimmen. Zu diesem Zweck wurden auf fächerförmig angeordneten Profilen Scholtewellen mit einer Airgun angeregt und am Scheitelpunkt der Profile mit einem der neu

10 Fallstudien zur Inversion

entwickelten SBG aufgezeichnet (siehe Kartenausschnitt b) in Abbildung 10.1). Beide Untersuchungen von 2006 und 2007 am FINO3 Standort wurden in dieser Arbeit erneut mit der cDHS-PSO invertiert. Eine weitere Studie zur Anwendung der cDHS-PSO auf einen Datensatz beschränkt sich auf ein Einzelprofil, welches im Gebiet „Adlergrund“ in der südlichen Ostsee im Juni 2008 aufgezeichnet wurde. Die invertierten Scherwellengeschwindigkeitsmodelle zeigen in allen Fällen sehr gute Kohärenz entlang der Profile und gute Korrelation zu reflexionsseismischen Sektionen.

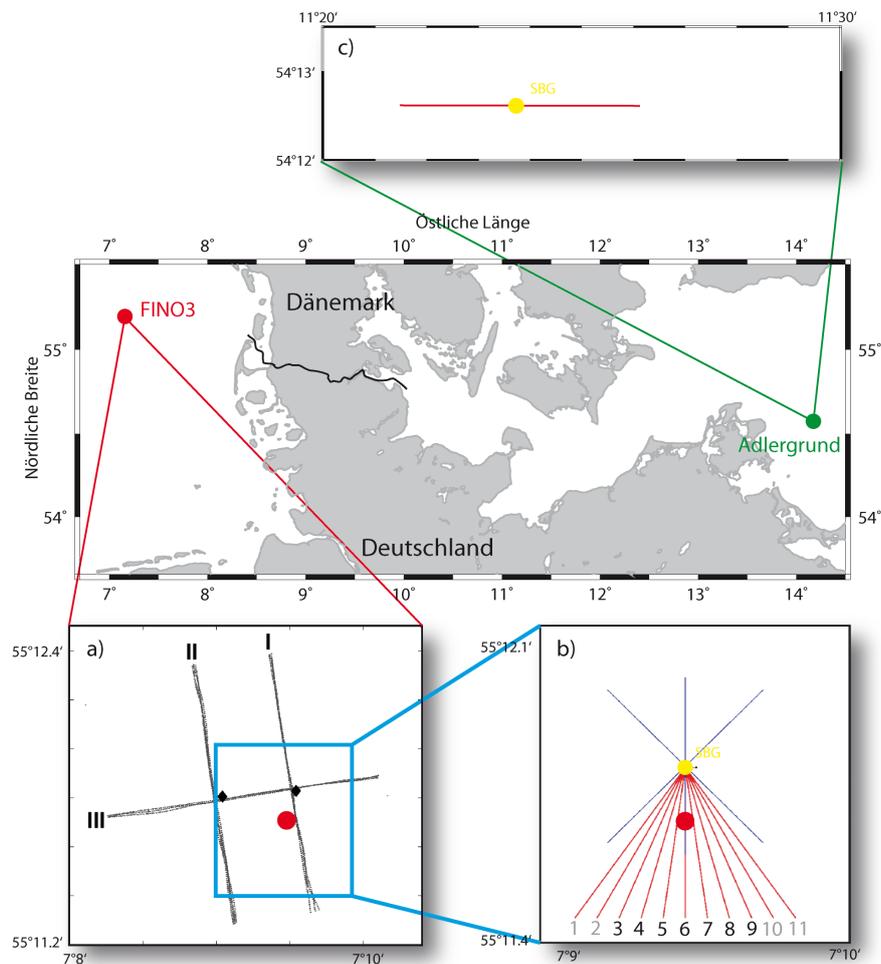


Abbildung 10.1: Karte der Untersuchungsgebiete.

Zur Akquisition wurde nach der im Kapitel 3 vorgestellten Methode nach Bohlen et al. (2004), Klein et al. (2005) vorgegangen. Die Anregung der Scholtewellen erfolgte mittels einer vom Schiff geschleppten Airgun in der Wassersäule. Die Scholtewellen werden dann mit einem Ozeanbodenseismometer registriert. Die

Schußabstände wurden so gewählt, daß räumliches Aliasing möglichst vermieden wird (siehe Abschnitt 3.1). Ausgewertet werden somit Common Receiver Gather der OBS, also Profilkompilationen verschiedener Schüsse an verschiedenen Entfernungen zum OBS.

10.1 Die FINO3 Baugrundvoruntersuchung

Inversion mit „global best“ PSO

Zur Baugrundvoruntersuchung und für ein erstes Maß der Scherwellengeschwindigkeitsverteilung an der Lokation von FINO3 wurden auf drei 2 km langen Profilen (siehe Karte 10.1, a)) Scholtewellen angeregt. Hierzu wurde eine Airgun mit einem Kammervolumen von 2.0 Litern hinter dem Forschungsschiff *Alkor* geschleppt. Die Scholtewellen wurden mit einem stationären Ozeanbodenseismometer (Sensoren: 3 Komponenten-Geophon (4 Hz Zentralfrequenz) und omnidirektionales Hydrophon) registriert. Es wurde auf den Profilen mit ungefähr 15 m Abstand geschossen, wobei die Schüsse durch die GPS Position der Quelle gesteuert wurden (siehe Abschnitt 3.1). So konnte durch das dreifache Überfahren des Profils und das Ineinanderschachteln der Schüsse ein effektiver Schußabstand von 5 m erreicht werden. Abbildung 10.2 zeigt exemplarisch an Profil I die erreichte Datenqualität der Common-Receiver-Gather. Bei einer minimalen Scherwellengeschwindigkeit von 200 m/s und einem maximalen Offset von 1000 m ergäbe sich ein minimales Schußintervall von 5 s. Bei einer minimalen Geschwindigkeit von 4 kn von FS *Alkor* ergibt das einen Schußabstand von 10 m. Der benötigte Schußabstand ist aber nach Abschnitt 3.1 bei einer zu erwartenden maximalen Frequenz von 25 Hz 4 m. Die Spektren werden jedoch zeigen, daß Scholtewellenenergie nur bis maximal 20 Hz zu beobachten ist, so daß ein Schußabstand von 5 m ausreichend ist (siehe Beispielspektrum in Abbildung 10.3).

Zu den Ergebnissen der Inversion mittels „global best“ PSO sei nun an dieser Stelle auf die Arbeit von Wilken et al. (2009) verwiesen.

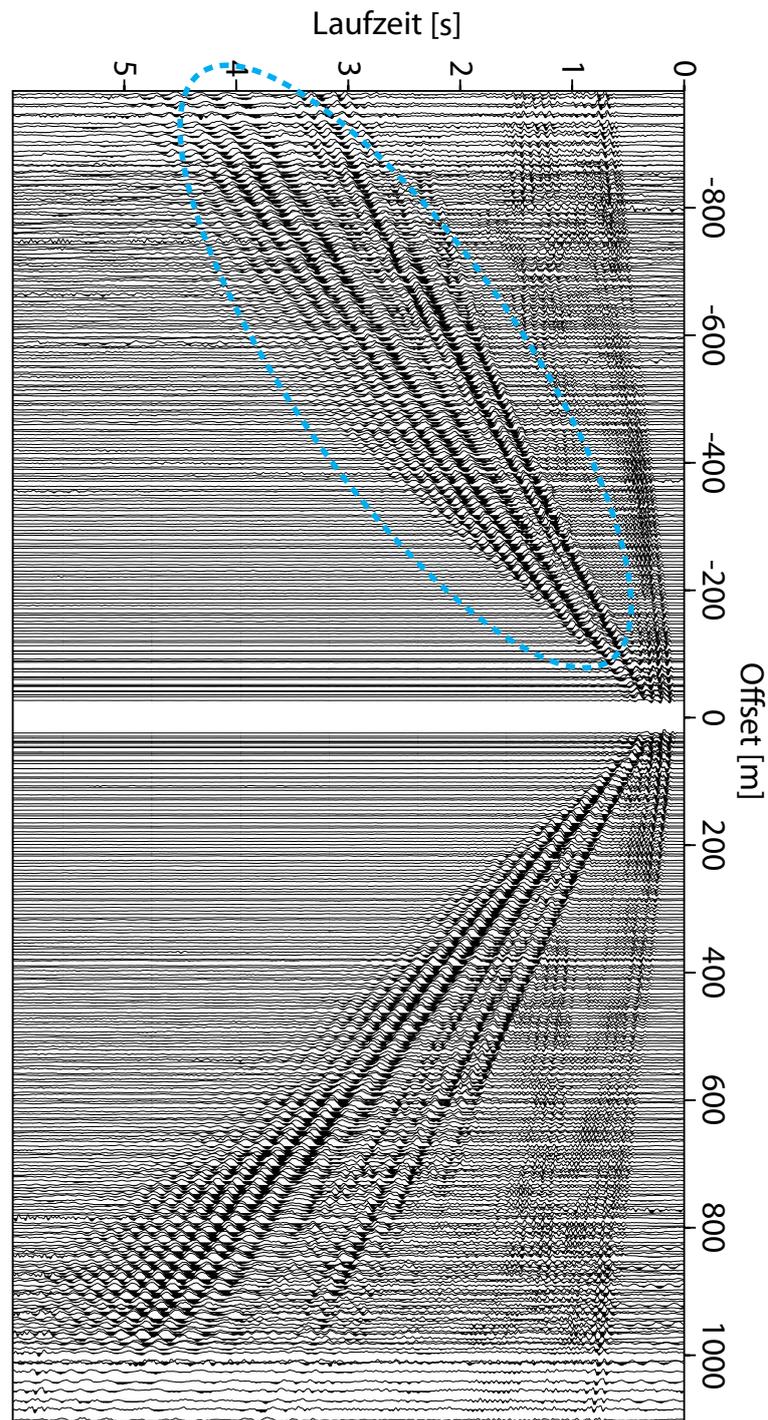


Abbildung 10.2: *Common Receiver Gather* des Profils I vom Mai 2006. Die Daten sind tiefpaßgefiltert mit einem Durchlaß unterhalb 20 Hz sowie spurnormiert. Der grau markierte Bereich zeigt die Einsätze der Scholtewelle für den negativen Offsetbereich.

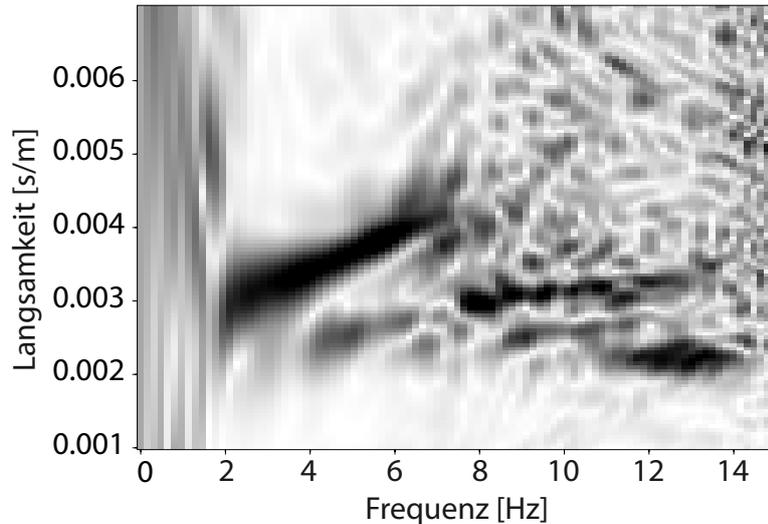


Abbildung 10.3: *Beispiel eines lokalen Phasenlangsamerkeits-Frequenz-Spektrums des Profils II am Offsetpunkt 550m.*

Inversion mit hybridem cDHS-PSO

Inversionsparameter:

Von den drei FINO3-Profilen wurden zwei erneut mit der cDHS-PSO invertiert. Der Parameterbereich ist derselbe wie der in Abschnitt 8.1 zum FINO3-Referenzmodell vorgestellte. Es wurden alle 100 m lokale Wellenfelder mit einer Fensterbreite von 200 m in Phasenlangsamerkeits-Frequenz-Spektren transformiert und die Dispersionskurven dann angepaßt. Als Datengrundlage dienen Dispersionskurvenpicks der identifizierbaren Fundamentalmode und der ersten höheren Mode mit einem Frequenzinkrement von 0.16 Hz. Abbildung 10.4 zeigt exemplarisch die Picks aller lokalen Dispersionskurven des Profils II. Die Picks wurden mit dem in Abschnitt 6.2 vorgestellten Verfahren automatisch extrahiert. Die Abbildung zeigt neben der Superposition aller Picks in einem Spektrum als Datengrundlage der Inversion (rechts) auch die nach Moden getrennten gepickten Phasenlangsamerkeiten über Offset und Frequenz. Hier zeigt sich in beiden Moden eine starke laterale Variation der Phasenlangsamerkeiten, die es im Vergleich mit dem 2D-Inversionsergebnis zu betrachten gilt. Das Qualitätsmaß der Inversion wird nun nach der „nächster Nachbar“ Bedingung, wie in Abschnitt 6.2 beschrieben, berechnet.

Sowohl Kompressionswellengeschwindigkeit als auch Dichte wurden in Konsistenz mit den Ausführungen in den Abschnitten 4.2 und 8.3 konstant gewählt. Die Kompressionswellengeschwindigkeit beträgt 1500 m/s in der Wassersäule und wird linear von 1550 m/s auf 1850 m/s in den Sedimenten erhöht. Die Dichte ist nach

10 Fallstudien zur Inversion

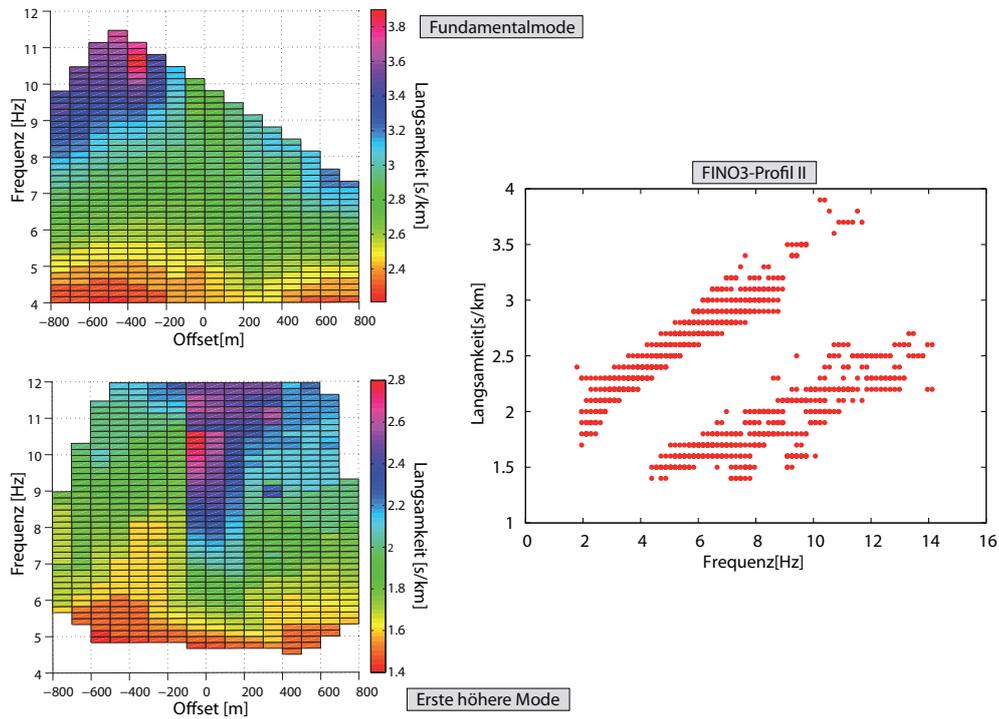


Abbildung 10.4: Darstellung der Dispersionskurvenpicks als Datengrundlage für das Profil II. Rechts: alle Picks der lokalen Dispersionskurven. Links: Picks aller Dispersionskurven entfernungstreu und nach Moden getrennt aufgetragen.

der Tiefenverteilung der Dichte eines am Baugrund der FINO3 gezogenen Bohrkerns gewählt worden und entspricht ebenfalls dem FINO3-Referenzmodell. Die Anzahl der Schichten wurde anhand der drei aufgelösten Horizonte der Reflexionsseismik der Baugrundvoruntersuchung analog dem FINO3 Referenzmodell auf vier Schichten plus Wasserschicht festgelegt.

Die eDHS-PSO wurde mit einer Populationsgröße von 300 Partikeln und einer Laufzeit von maximal 1000 Iterationen durchgeführt und entspricht damit dem in Kapitel 8.3 identifizierten effektivsten Verfahren. Das Verfahren ist so dimensioniert, daß ausreichend Neustarts der Optimierung möglich sind. Die Auswahl der invertierten 1D Modelle aus den Ergebnissen dieser Neustarts erfolgt sowohl nach dem Wert der Qualitätsfunktion als auch nach Kohärenzbetrachtungen entlang eines Profils.

Inversionsergebnisse:

Die Abbildungen 10.5 und 10.6 zeigen nun die erhaltenen Pseudo-2D-Scherwellengeschwindigkeiten der Profile II und I. Zusätzlich zu den Geschwindigkeitsmodellen sind hochauflösende reflexionsseismische Sektionen der Profile gezeigt. Die für das 2D-Modell genutzten 1D-Teilmodelle werden aus den Ergebnissen der Inversionen primär nach dem besten Qualitätswert aber auch nach Kohärenz

10.1 Die FINO3 Baugrundvoruntersuchung

entlang des Profils gewählt. So können gegebenenfalls Teilmodelle enthalten sein, die nicht der nach dem Qualitätsmaß besten Lösung sondern einer anderen Neustartlösung entsprechen, welche ein kohärenteres 2D-Modell ermöglicht. Mit den schwarzen Linien in der Modelldarstellung wird versucht, einige in der Reflexionsseismik identifizierte kohärente Einsätze (grüne Linien) wiederzufinden. Als Hilfslinie ist die 30 m Linie eingezeichnet, deren Tiefe in der Reflexionssektion mit einer mittleren Geschwindigkeit von 1600 m/s berechnet wurde. Es ergibt sich in beiden Ergebnissen eine sehr gute Korrelation der Scherwellengeschwindigkeitsmodelle und der Reflexionseismik. Ebenso bieten die Inversionen eines Durchganges Lösungsmodelle, die zu einem kohärenten zweidimensionalen Bild führen.

10 Fallstudien zur Inversion

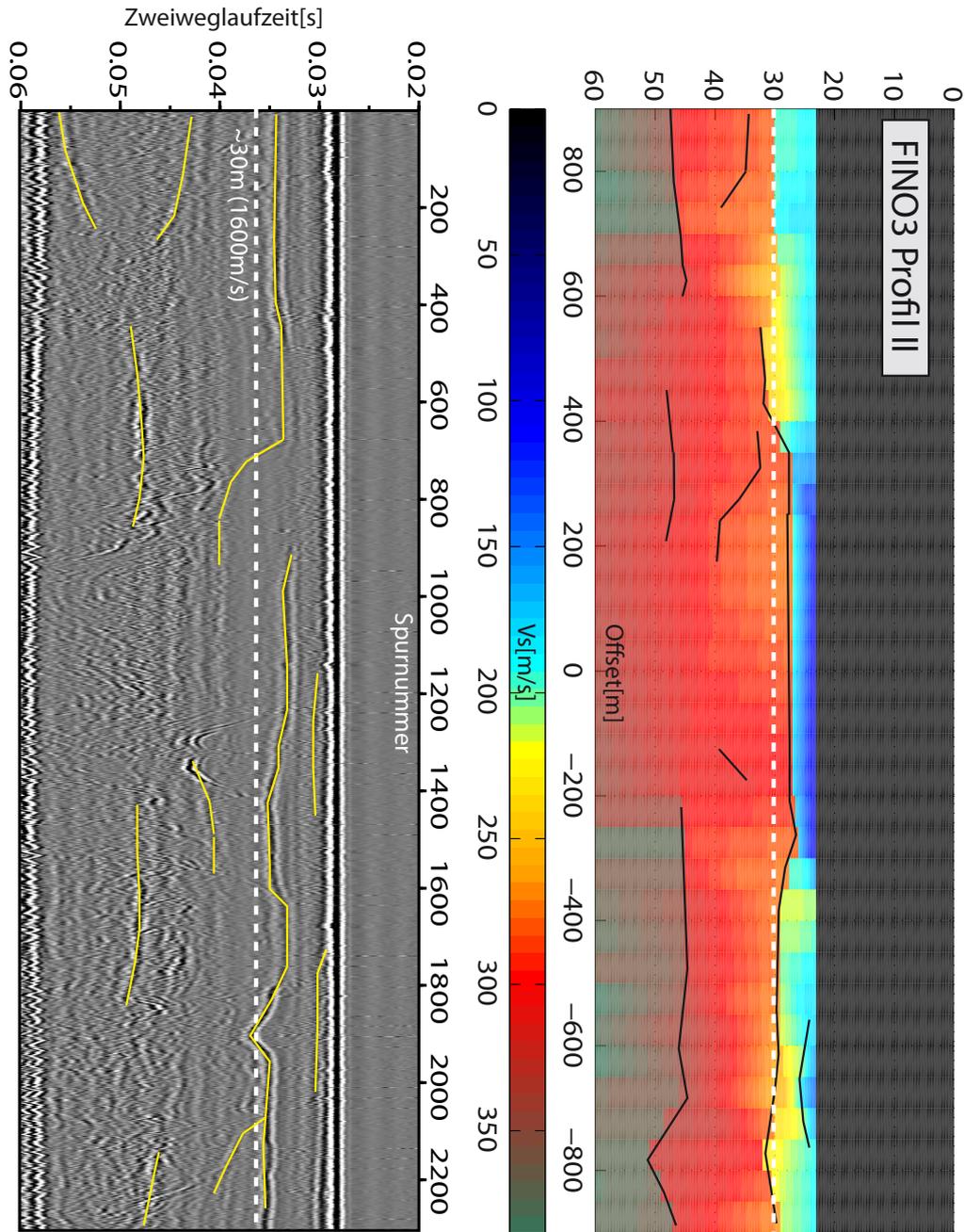


Abbildung 10.5: Pseudo-2D-Scherwellengeschwindigkeitsmodell (rechts) des Profils II und entsprechende reflexionsseismische Boomersektion. Die gestrichelte weiße Linie entspricht einer Tiefe von 30 m bei einer mittleren Geschwindigkeit von 1600 m/s. (Sektion bandpaßgefiltert mit Trapezfilter mit einer Öffnungsflanke von 100 Hz bis 200 Hz und einer Schlußflanke von 2000 Hz bis 3000 Hz)

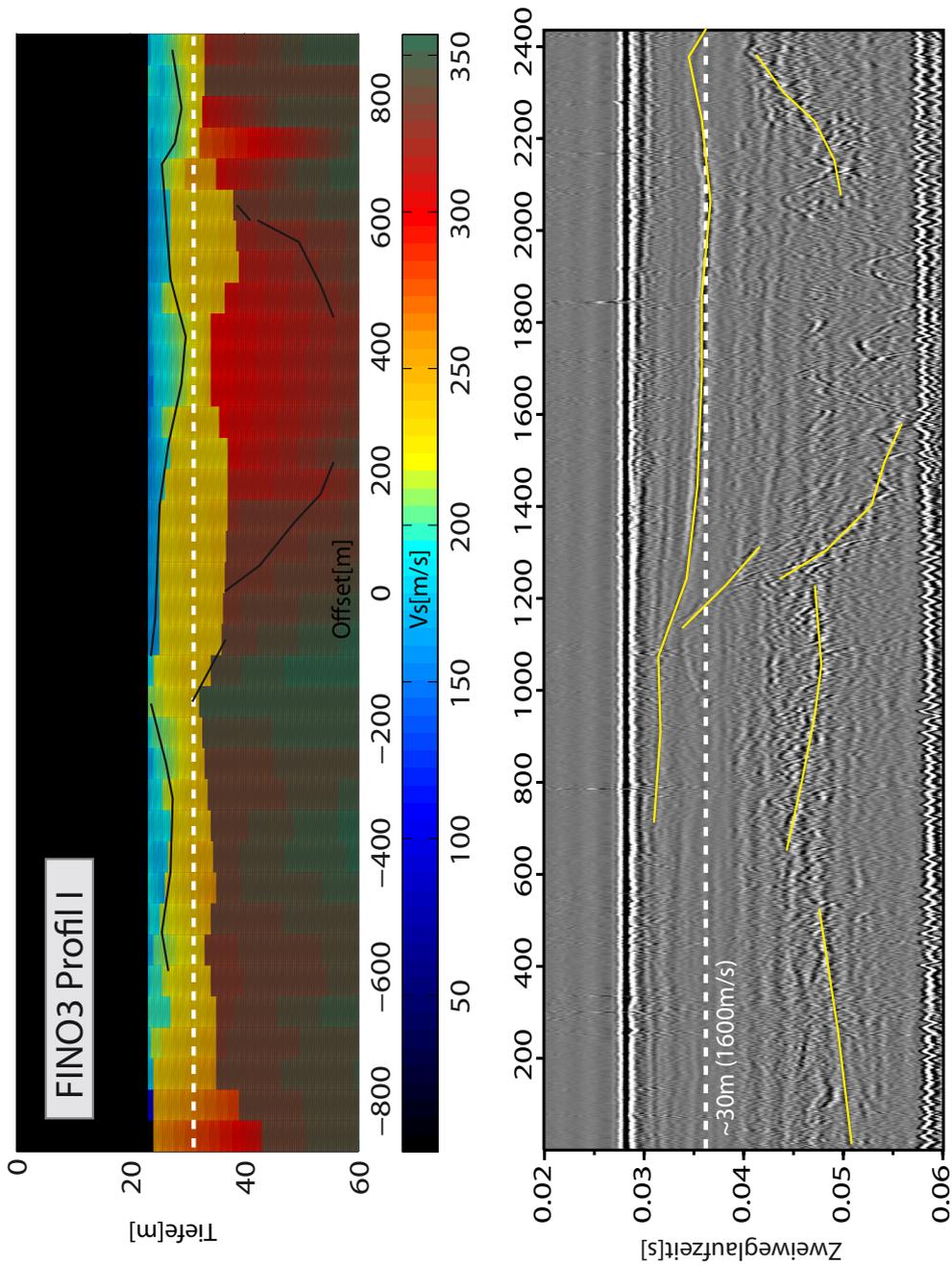


Abbildung 10.6: Pseudo-2D-Scherwellengeschwindigkeitsmodell (links) des Profils I und entsprechende reflexionsseismische Boomersektion. Die gestrichelte weiße Linie entspricht einer Tiefe von 30 m bei einer mittleren Geschwindigkeit von 1600 m/s. (Sektion bandpaßgefiltert mit Trapezfilter mit einer Öffnungsflanke von 100 Hz bis 200 Hz und einer Schlußflanke von 2000 Hz bis 3000 Hz)

10 Fallstudien zur Inversion

Die geringere Geschwindigkeit im senkenartigen Bereich des Modells von Profil II zwischen -200 m und 400 m Offset und unterhalb 30 m Tiefe korreliert mit einer Erhöhung der Langsamkeiten der lokalen Spektren in Abbildung 10.4 bei niedrigen Frequenzen in Fundamentalmode und erster höherer Mode. Dies gilt für beide Profile. Die drei sich abzeichnenden Geschwindigkeitsbereiche von im Mittel 180 m/s, 280 m/s und 360 m/s lassen sich durch die Referenzbohrung Sandschichten verschiedener Kompaktierung zuordnen (Wilken et al., 2009). Der sehr flache Bereich bis 26 m Tiefe mit Geschwindigkeiten unter 200 m/s entspricht vermutlich einer am Meeresboden aufliegenden Schlickschicht.

Auflösungsanalyse:

Abbildung 10.7 zeigt nun die Auswertung von drei Inversionsdurchgängen an einem exemplarischen lokalen Modell des Profils II nach dem in Abschnitt 8.4 vorgestellten und dimensionierten Verfahren. Die Abbildung enthält das Modell

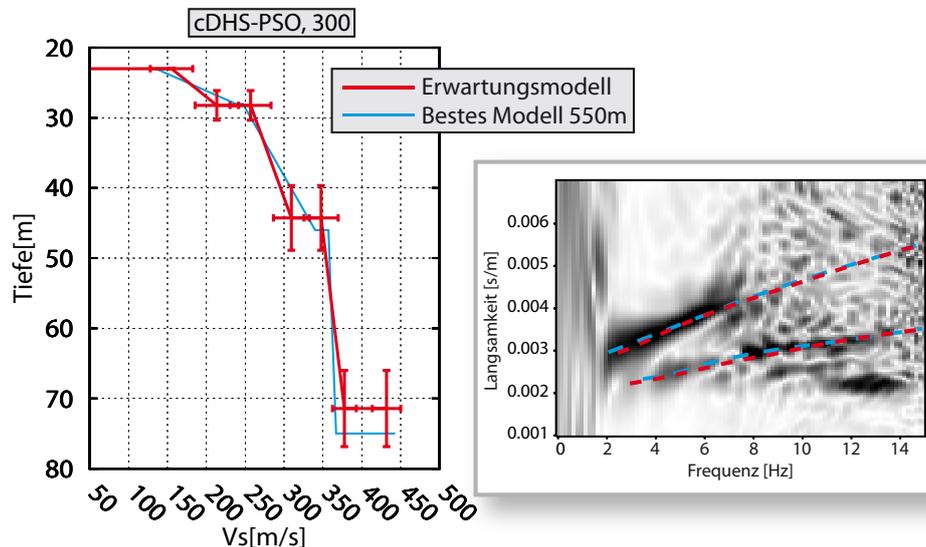


Abbildung 10.7: Erwartungsmodell (rot) und bestes Modell (blau) aus drei Inversionsdurchgängen des lokalen Wellenfeldes am Offsetpunkt 550 m des Profils II. Rechts ist das Phasenlangsamkeits-Frequenz-Spektrum des lokalen Wellenfeldes sowie die Anpassung durch das beste Modell (blau) und das Erwartungsmodell (rot) dargestellt.

bester Anpassung (blaue Kurve) und das aus 5000 Modellen der zehnten bis zwanzigsten Iteration dreier Inversionen ermittelte Erwartungsmodell. Die Standardabweichungen zeigen eine Verschlechterung der Auflösung zu flacheren Schichten, was mit der geringeren Datenmenge zu höheren Frequenzen (die Fundamentalmode zeigt einen Frequenzbereich von 2 Hz bis 8 Hz) und fehlerhaften Picks durch Interferenzen bei höheren Frequenzen verknüpft werden kann. Die Auflösung beträgt jedoch im Mittel noch immer 9%. Zu geringen Tiefen entspricht dies aber bis zu 27 m/s.

10.1.1 Das FINO3 Hintergrundmodell

Meßlokation:

Als Nahbereichsuntersuchung zur Scherwellengeschwindigkeitsverteilung in drei Dimensionen wurden im Mai 2007 auf elf fächerförmig angeordneten Profilen Scholtewellen mit einer Airgun angeregt und am Scheitelpunkt der Profile mit einem der neu entwickelten SBG (siehe Abschnitt 3.2) aufgezeichnet. Die Profile sind im vergrößerten Bereich b) in der Karte (Abbildung 10.1) eingetragen.

Meßparameter:

In Tabelle 10.1 sind die Parameter der Messung zusammengestellt. Zur Positionierung des SBG wurden neben den elf Airgunprofilen der Scholtewellenmessung noch drei weitere Profile innerhalb des Profilfächers mit einer Boomerquelle gefahren. In den Hydrophondaten des SBG konnte so der Ersteinsatz des Boomers ermittelt werden und die erhaltenen Laufzeiten im Wasser zur Positionierung des Systems herangezogen werden. Die Position des SBG konnte auf $54^{\circ} 11.851'$ Nord und $7^{\circ} 9.512'$ Ost mit einer Abweichung von 4 m ermittelt werden. Die Quelltiefe wurde zu dieser Messung auf 12 m erhöht sowie eine Airgun mit kleinerem Kammervolumen verwendet, um höhere Frequenzen anregen zu können und die Auflösung im flachen Modellbereich damit eventuell zu verbessern. Die erhaltenen

Quelle(Volumen/Betriebsdruck)	Airgun (1.2 l, ~ 90 bar)
Quelltiefe [m]	12
Wassertiefe [m]	23
Sensorik SBG	4.5 Hz vertikal Geophon
	omnidirektionales Hydrophon
Abtastrate [Hz]	10000 Hz pro Kanal
Schußintervall [s]	6 s

Tabelle 10.1: *Meßparameter der Scholtewellenmessung vom Mai 2007.*

Common Receiver Gather der Profile zeigen stark unterschiedliche Datenqualität. So sind Offsetbereiche in einigen Profilen stark verrauscht und entsprechende lokale Wellenfelder lassen sich nicht auswerten (siehe Abbildung 10.8, Profil 2). Es wurden daher nur die Profile drei bis neun mit relativ guter Datenqualität zur Inversion der Scherwellengeschwindigkeiten genutzt.

Inversionsparameter:

Die Inversionsparameter wurden im Hinblick auf die Inversion der Profile I-II vom Mai 2006 nicht verändert. Es werden alle 50 m lokale Wellenfelder mit einer Fensterbreite von 150 m in Phasenlangsamkeits-Frequenz-Spektren transformiert. Exemplarisch sind in Abbildung 10.9 die Picks aller lokalen Dispersionskurven des Profils 3 dargestellt. Die Picks wurden ebenfalls mit dem in Abschnitt 6.2 vorgestellten Verfahren automatisch extrahiert. Die Fundamentalmode zeigt hier Frequenzen bis 10 Hz die erste höhere Mode bis 14 Hz. Des weiteren sind einige Picks

10 Fallstudien zur Inversion

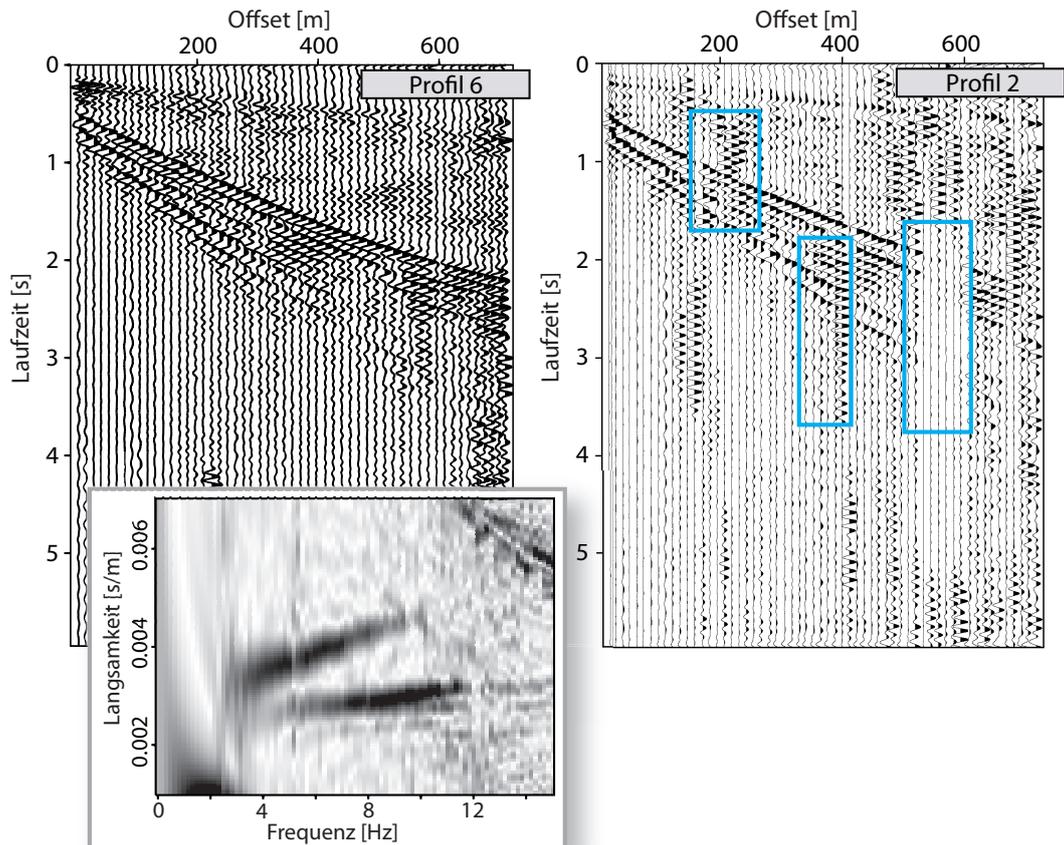


Abbildung 10.8: Common Receiver Gather der Profile 2 und 6 Profildächer vom Mai 2007. Die Daten sind bandpaßgefiltert mit einem Trapezfilter des Verlaufs 2 Hz, 4 Hz, 10 Hz, 15 Hz. Die blauen Kästen in Profil 2 markieren Bereiche besonders schlechter Datenqualität. Ebenfalls dargestellt ist das Phasenlangsamkeits-Frequenz-Spektrum des gesamten Profils 6. Es lassen sich Amplitudenmaxima einer Fundamental- und einer höheren Mode der Schottewelle erkennen. Oberhalb von 11 Hz und 0.005 s/m erscheint Energie zurückgefalteter Signale außerhalb des wichtigen Signalbereichs.

10.1 Die FINO3 Baugrundvoruntersuchung

einer zweiten höheren Mode zuzuordnen. Die Abbildung zeigt ebenfalls die nach Moden getrennten gepickten Phasenlangsamkeiten über Offset und Frequenz. Es zeigt sich in der Fundamentalmode kaum laterale Variation der Phasenlangsamkeiten. Die erste höhere Mode zeigt jedoch im tiefen Frequenzbereich von 2 Hz bis 6 Hz eine Erhöhung der Langsamkeit zur Mitte des Profils.

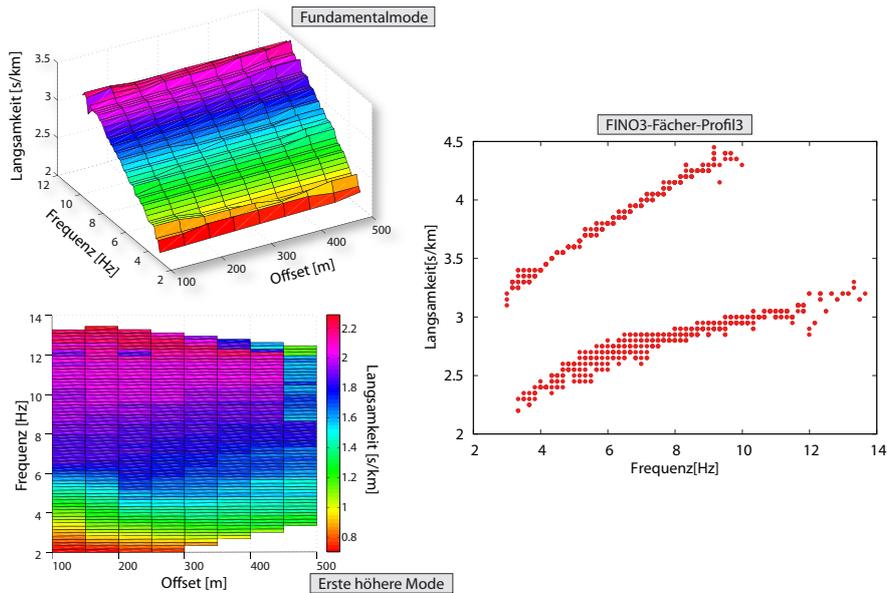


Abbildung 10.9: Darstellung der Dispersionskurvenpicks als Datengrundlage für das Profil 6. Rechts: alle Picks der lokalen Dispersionskurven. Links: Picks aller Dispersionskurven entfernungsgetreu und nach Moden getrennt aufgetragen.

Inversionsergebnisse:

Abbildung 10.10 zeigt die invertierten Scherwellengeschwindigkeitsmodelle der Profile 3 bis 9. Die Modelle zeigen gute Kohärenz entlang jedes Profils (exemplarisch ist Profil 6 abgebildet) und ebenfalls zwischen den Profilen. Die erhaltenen Geschwindigkeitsbereiche entsprechen den Geschwindigkeiten von Profil I der Messung von 2006, welches durch das Meßgebiet verläuft. Abbildung 10.10 zeigt insbesondere zwischen den Profilen interpolierte Scherwellengeschwindigkeitstiefenkarten für Tiefen von 30 m, 40 m, 50 m und 55 m unterhalb der Wasseroberfläche. Die lateralen Variationen der Scherwellengeschwindigkeit innerhalb eines Tiefenschnittes liegen im Bereich von 50 m/s bis 65 m/s und sind damit eher klein. Die gute Kohärenz innerhalb der Profile sowie die Glattheit der Phasenlangsamkeitskarten implizieren jedoch eine Auflösbarkeit dieser Variationen.

10 Fallstudien zur Inversion

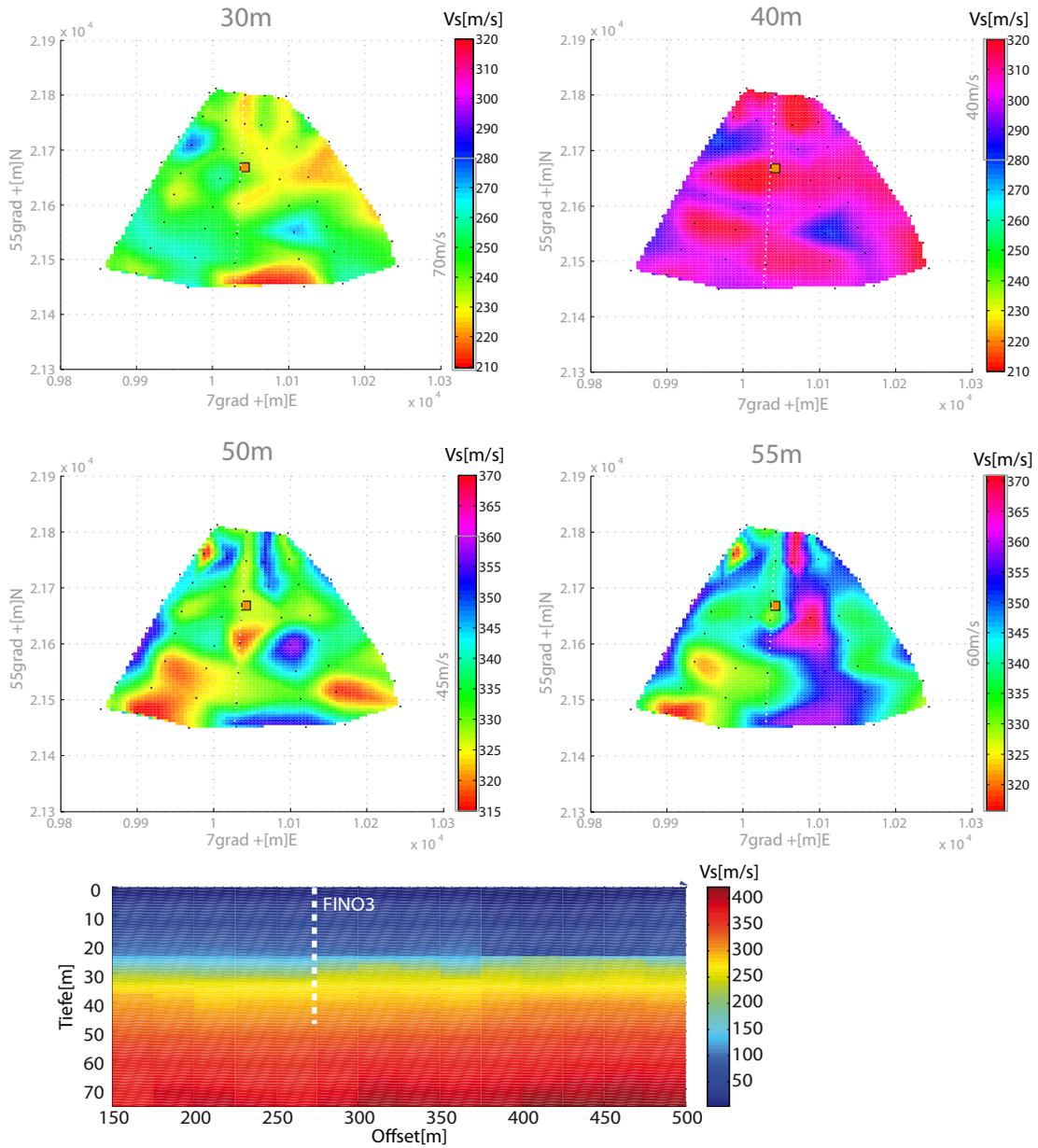


Abbildung 10.10: Invertierte pseudo-3D Scherwellengeschwindigkeitstiefenkarten für 30 m, 40 m, 50 m und 55 m Tiefe unterhalb der Wasseroberfläche. Die schwarzen Punkte zeigen die Stützstellen der invertierten 1D Modelle auf den Profilen. Das untere Bild zeigt exemplarisch das Profil 6 angezeigt durch eine weiße gestrichelte Linie in den Geschwindigkeitskarten. Die Position der FINO3-Plattform ist im Profil (weiße Linie) und in den Karten (grünes Rechteck) markiert.

Auflösungsanalyse:

Abbildung 10.11 zeigt dazu die Auswertung von drei Inversionsdurchgängen an einem exemplarischen lokalen Modell des Profils 3 nach dem in Abschnitt 8.4 vorgestellten und dimensionierten Verfahren. Die Abbildung enthält das Modell

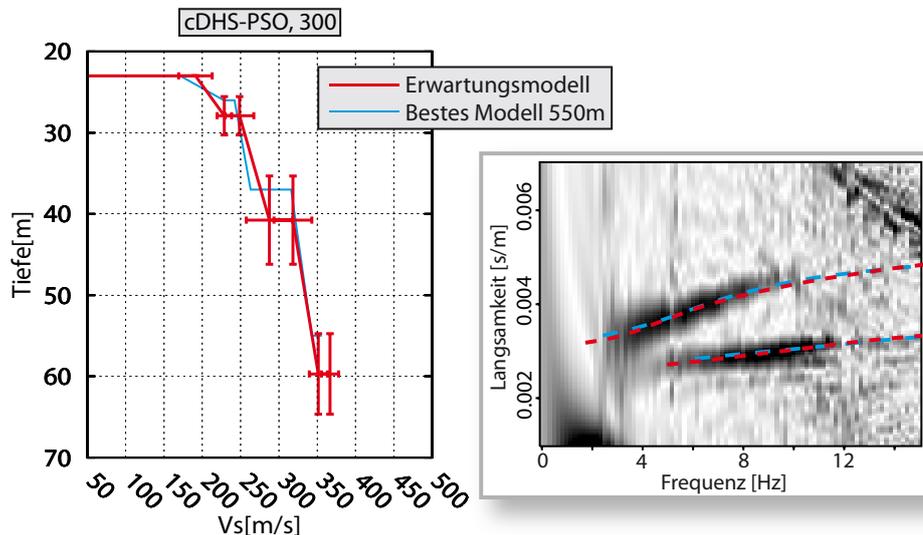


Abbildung 10.11: Erwartungsmodell (rot) und bestes Modell (blau) aus drei Inversionsdurchgängen des lokalen Wellenfeldes am Offsetpunkt 150 m des Profils 3. Rechts ist das Phasenlangsamerkeits-Frequenz-Spektrum des lokalen Wellenfeldes sowie die Anpassung durch das beste Modell (blau) und das Erwartungsmodell (rot) dargestellt.

bester Anpassung (blaue Kurve) und das aus 5000 Modellen aus dem Nahbereich dieser Lösung ermittelte Erwartungsmodell. Die Standardabweichungen zeigen eine Verschlechterung der Auflösung im mittleren Modellbereich. Die Auflösung beträgt im Mittel 8% und übersteigt 11% nicht. Im mittleren Modellbereich bedeutet dies aber bis zu 30 m/s, sonst 11 m/s bis 20 m/s.

10.2 Lokation Adler-Grund

Meßlokation:

Im Juni 2008 wurde mit dem Forschungsschiff *Poseidon* ein einzelnes Airgunprofil zur Anregung von Scholtewellen an der Lokation Adlergrund in der südlichen Ostsee (siehe Abbildung 10.1 c)) gefahren, um ein Fallbeispiel mit zur Nordsee verschiedener Scherwellengeschwindigkeitsstruktur zu erhalten. Auch in diesem Gebiet sind Vorhaben zur Windparkerrichtung geplant (BSH, 2002). Innerhalb des Gebietes Adlergrund ist in weiten Teilen Restsediment mit Grobsand, Feinkies und Steinen auf anstehendem Geschiebemergel zu erwarten. Im Südosten bildet der Geschiebemergel bisweilen Aufragungen (BSH, 2002).

10 Fallstudien zur Inversion

Meßparameter:

Die Parameter der Messung sind in Tabelle 10.2 zusammengestellt. Der Bereich

Quelle(Volumen/Betriebsdruck)	Airgun (1.2 l, ~80 bar)
Quelltiefe [m]	5
Wassertiefe [m]	22
Sensorik SBG	4.5 Hz vertikal Geophon omnidirektionales Hydrophon
Abtastrate [Hz]	10000 Hz pro Kanal
Schußintervall [s]	6 s

Tabelle 10.2: *Parameter der Scholtewellenmessung vom Juni 2008 am Adlergrund.*

des erhaltenen Common Receiver Gathers des Profils, in welchem Scholtewellenenergie beobachtet werden konnte (-750 m bis 550 m Offset) ist in Abbildung 10.13 dargestellt. Eine ebenfalls aufgezeichnete hochauflösende einkanal Boomer-reflexionssektion des Profils zeigte wenig Charakteristika (siehe Abbildung 10.15). Unter dem Meeresboden ist nur ein Einsatz zu erkennen, welcher jedoch zur Mitte des Profils ansteigt und eventuell einer Aufragung des Geschiebemergels entspricht. In Abbildung 10.12 ist exemplarisch das lokale Phasenlangsamkeits-Frequenz-Spektrum vom Offsetpunkt 450m dargestellt. Scholtewellenenergie ist hier von 4 Hz bis 18 Hz zu beobachten.

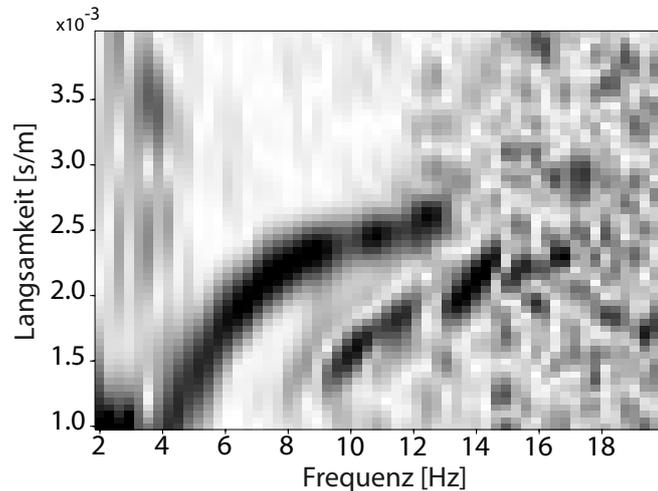


Abbildung 10.12: *Beispiel eines lokalen Phasenlangsamkeits-Frequenz-Spektrums des Profils am Adlergrund am Offsetpunkt 450m.*

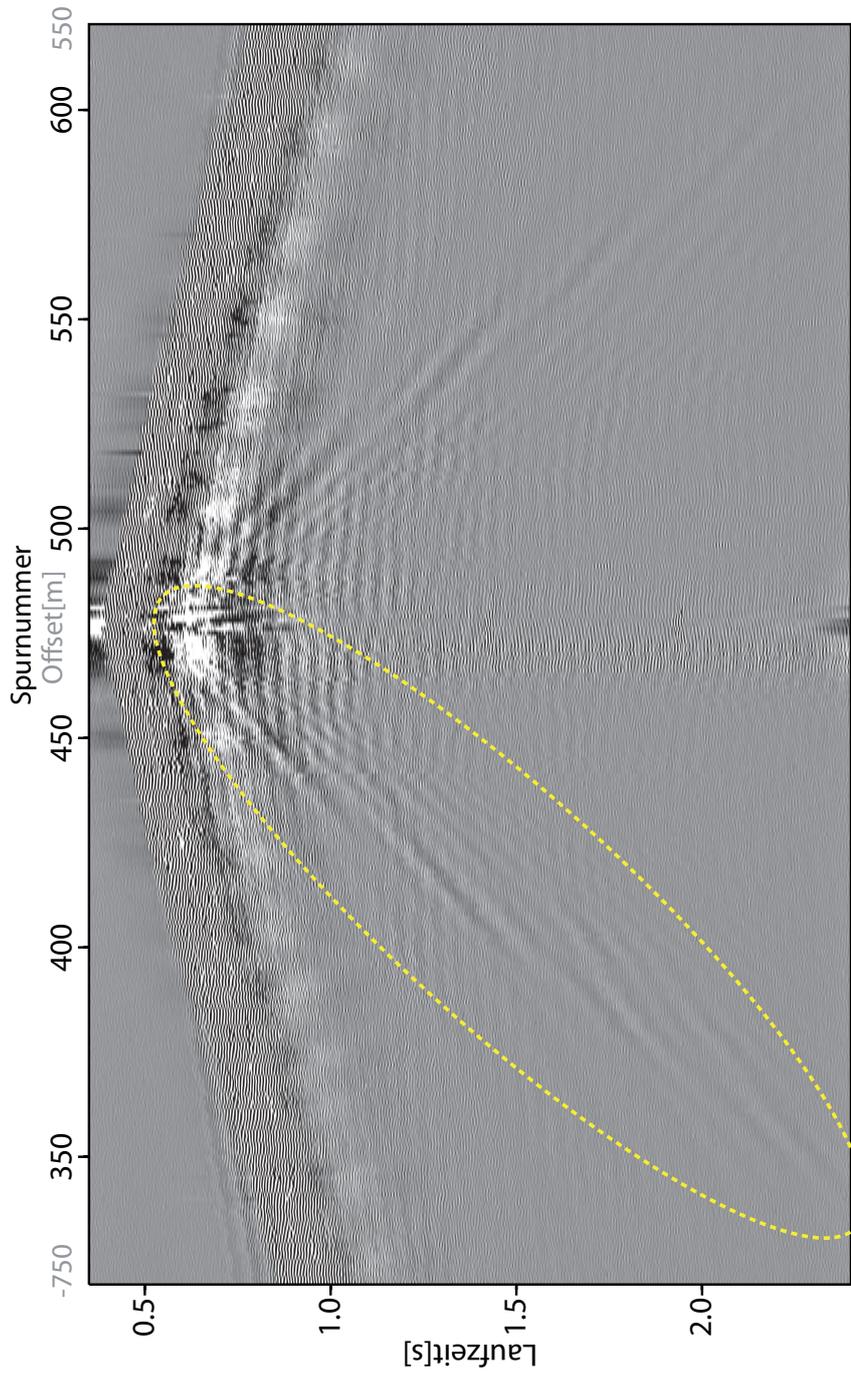


Abbildung 10.13: *Common Receiver Gather* des Profils Adlergrund. Die Daten sind bandpaßgefiltert mit einem Trapezfilter des Verlaufs 2 Hz, 4 Hz, 100 Hz, 200 Hz. Der gelb markierte Bereich zeigt die Einsätze der Scholtewelle für den negativen Offsetbereich.

Inversionsparameter:

Lokale Spektren des Profils wurden nun mit der cDHS-PSO invertiert. Es wurden alle 50 m lokale Wellenfelder mit einer Fensterbreite von 150 m in Phasenlangsamkeits-Frequenz-Spektren transformiert und die gepickten Dispersionskurven dann angepaßt. Als Datengrundlage dienten Dispersionskurvenpicks der in den Spektren identifizierbaren Fundamentalmode und der ersten höheren Mode mit einem Frequenzinkrement von 0.32 Hz. Abbildung 10.14 zeigt die Picks aller lokalen Dispersionskurven des Profils. Die Picks wurden automatisch extrahiert. Die Abbildung zeigt ebenfalls die nach Moden getrennten Phasenlangsamkeiten über Offset und Frequenz. Hier zeigt sich in beiden Moden eine leichte laterale Variation der Phasenlangsamkeiten. Auffällig ist hier die Erhöhung der Phasenlangsamkeiten zu tiefen Frequenzen zur Mitte des Profils, welche voraussichtlich dem ansteigenden Einsatz der Reflexionsseismik entsprechen. Ausgehend von den

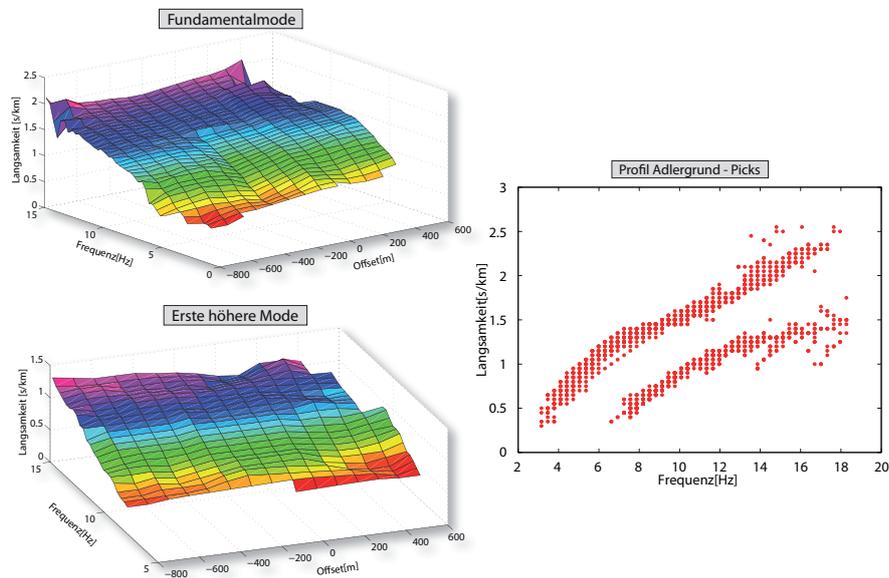


Abbildung 10.14: Darstellung der Dispersionskurvenpicks als Datengrundlage für das Profil Adlergrund. Rechts: alle Picks der lokalen Dispersionskurven. Links: Picks aller Dispersionskurven entfernungsgetreu und nach Moden getrennt aufgetragen.

Reflexionsdaten wurde zunächst versucht, die spektralen Charakteristika durch verschiedene Modellierungen von Dispersionskurven mit einem 4-Schichten Modell (Wassersäule, zwei Schichten, Halbraum) zu reproduzieren. Dies gelang, so daß mit diesem Modell invertiert werden kann. Es zeigen sich allerdings bereits an den Dispersionskurven höhere Geschwindigkeitsbereiche als in der Nordsee. So ist ein Scherwellengeschwindigkeitsbereich von 200 m/s bis 1200 m/s zu erwarten. Abbildung 10.15 stellt den gewählten Parameterbereich für die Inversionsparameter Scherwellengeschwindigkeit und Schichtmächtigkeit der drei variablen Schich-

ten dar. Sowohl Kompressionswellengeschwindigkeit v_p als auch Dichte wurden

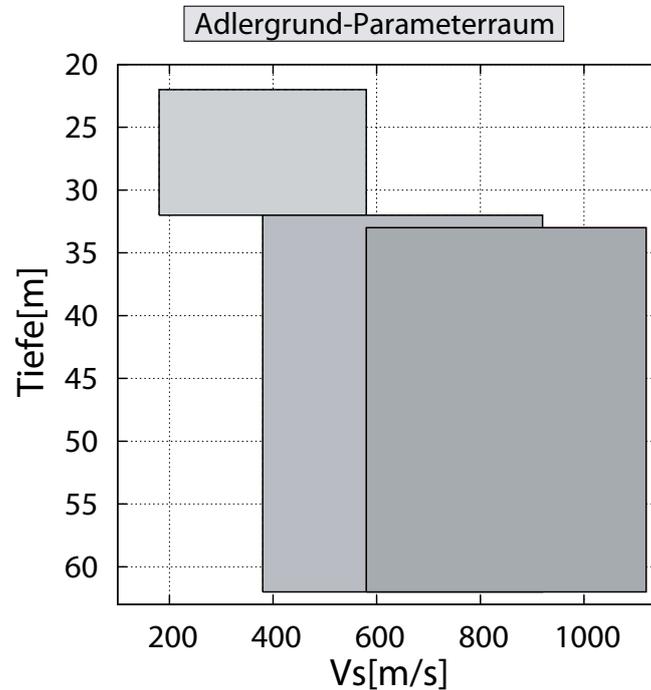


Abbildung 10.15: Darstellung des gewählten Parameterraumes. Die verschiedenen Graustufen entsprechen den Bereichen für Scherwellengeschwindigkeit und Schichtlage der drei variablen Schichten.

erneut konstant gewählt. v_p beträgt 1500 m/s in der Wassersäule und wird in den folgenden Schichten auf im Mittel 1550 m/s, 1750 m/s und 1850 m/s abgeschätzt. Die Dichte wurde auf 1.0 t/m^3 in der Wassersäule gesetzt und auf 1.9 t/m^3 , 2.1 t/m^3 und 2.15 t/m^3 in den Sedimentschichten abgeschätzt. Die cDHS-PSO wurde mit einer Populationsgröße von 300 Partikeln und einer Laufzeit von maximal 1000 Iterationen durchgeführt. Die Auswahl der invertierten 1D-Modelle aus den Ergebnissen dieser Neustarts erfolgt sowohl nach dem Wert der Qualitätsfunktion als auch nach Kohärenzbetrachtungen entlang des gesamten Profils.

Inversionsergebnisse:

Die Abbildung 10.16 zeigt die erhaltenen Pseudo-2D-Scherwellengeschwindigkeiten des Profils. Zusätzlich zum Geschwindigkeitsmodell ist wieder eine hochauflösende reflexionsseismische Sektionen des Profils gezeigt. Mit den weißen Linien in der Modelldarstellungen wurde versucht, den in der Reflexionsseismik ansteigenden Einsatz wiederzufinden. Auch im Scherwellengeschwindigkeitsmodell ist eine ansteigende Schicht höherer Geschwindigkeit sowohl im flachen Bereich (23 m bis 30 m) korreliert mit der Reflexionsseismik, als auch im tieferen Bereich (30 m bis 40 m) identifizierbar. Das Verfahren konnte somit erneut mit einem

10 Fallstudien zur Inversion

Inversionsdurchgang ein kohärentes und mit der Reflexionsseismik korrelierendes Modell invertieren.

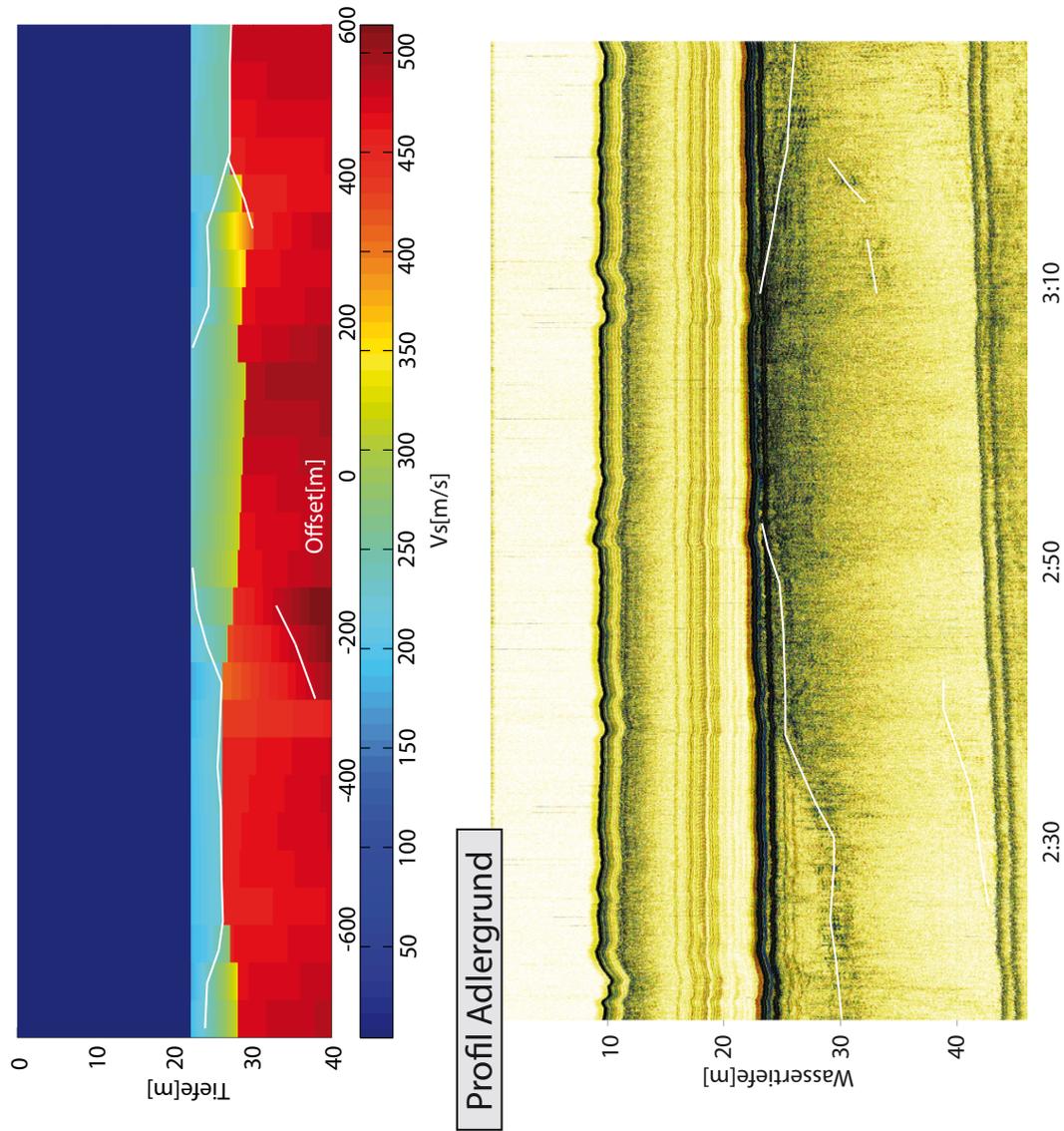


Abbildung 10.16: *Pseudo 2D* Scherwellengeschwindigkeitsmodell (links) des Profils Adlergrund und eine entsprechende reflexionsseismische Boomersektion.

Auflösungsanalyse:

Abbildung 10.17 zeigt die Auswertung von drei Inversionsdurchgängen an einem exemplarischen lokalen Modell des Profils nach dem in Abschnitt 8.4 vorgestellten und dimensionierten Verfahren. Die Abbildung enthält das Modell bester

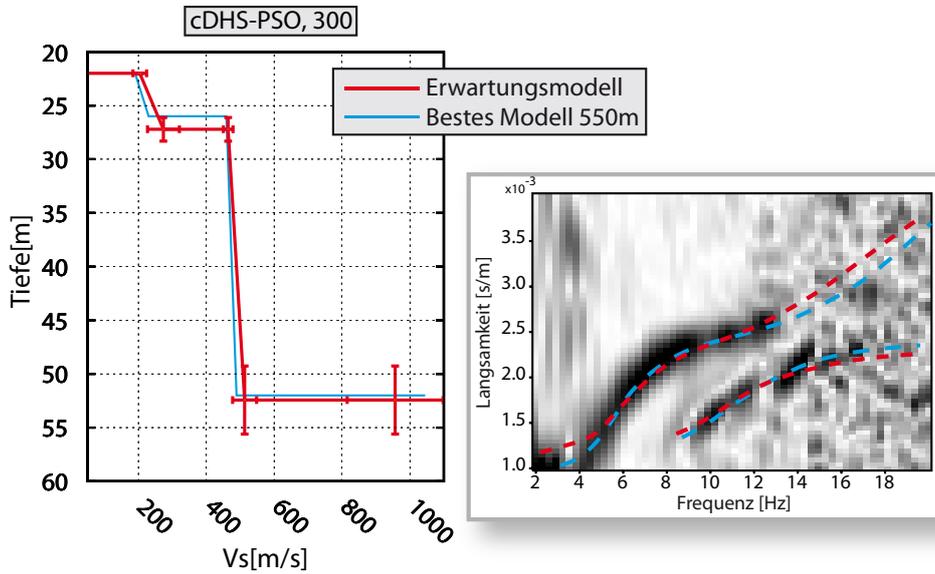


Abbildung 10.17: Erwartungsmodell (rot) und bestes Modell (blau) aus drei Inversionsdurchgängen des lokalen Wellenfeldes am Offsetpunkt 450 m. Rechts ist das Phasenlangsamkeits-Frequenz-Spektrum des lokalen Wellenfeldes sowie die Anpassung durch das beste Modell (blau) und das Erwartungsmodell (rot) dargestellt.

Anpassung (blaue Kurve) und das aus 5000 Modellen aus dem Nahbereich dieser Lösung ermittelte Erwartungsmodell. Die Standardabweichung beträgt im Mittel 10% und übersteigt nicht die 17 m/s. Schichttiefen sind mit im Mittel 5% sehr gut aufgelöst.

11 Zusammenfassung und Diskussion der Ergebnisse

Wie einleitend erläutert, kann eine dynamische Beanspruchung von Offshorebauwerken durch Strömung, Wellenschlag und Winddruck in der Umgebung solcher Bauwerke zu Veränderungen der Sedimentstruktur und deren mechanischen Eigenschaften führen. Es kann hierbei durch zyklische Belastungen in unterschiedlichen Richtungen die Stabilität einer Gründung dramatisch herabgesetzt werden. Informationen über die Verteilung der Scherwellengeschwindigkeiten des Baugrundes können aufgrund der Korrelation der Geschwindigkeiten mit der Scherfestigkeit bei einer Überwachung des Baugrundes von zentraler Bedeutung sein. Die Motivation dieser Arbeit lag daher in der Bestimmung von Scherwellengeschwindigkeitsmodellen, welche aus der Dispersionsanpassung von Scholtewellen erhalten werden. Hierbei sollte insbesondere eine erhöhte Auflösung der Modelle im Vergleich zu bisherigen Ansätzen erreicht werden. Hierzu wurden in erster Linie verschiedene stochastische Optimierungsverfahren mit einem Schwerpunkt auf Schwarmintelligenzverfahren auf das Anpassungsproblem angewandt sowie eine neue Methode zur Bestimmung des Qualitätsmaßes entwickelt. Mit dem neuen Verfahren sollten Bauwerk gefährdende Veränderungen mariner Sedimente, wie beispielsweise sukzessive Verflüssigung, in ihrem Einfluß auf die Scherwellengeschwindigkeit auflösbar werden.

Die Auflösung eines durch Dispersionsanpassung invertierten Scherwellengeschwindigkeitsmodells ist meßtechnisch abhängig von der Genauigkeit der Messung der Phasenlangsamkeiten der Scholtewelle, also durch Bestimmungsfehler der Positionen von Quelle und Empfänger. Des weiteren beschränkten Phasenlangsamkeitsauflösung durch eine endliche Zahl von Meßpunkten sowie die algorithmische Auflösung des Inversionsverfahrens die Auflösung der erhaltenen Modelle. Das Anpassungsproblem leidet grundsätzlich unter Mehrdeutigkeiten und damit verknüpfter Multimodalität, so daß eine adäquate Auflösung bei der Nutzung lokaler Optimierungsverfahren nur durch a priori Informationen zur Definition eines sehr gut gewählten Startmodells erreicht werden kann.

Mit dem speziellen Hintergrund einer zukünftigen Scherwellengeschwindigkeitsüberwachungskampagne am Baugrund der Offshoreforschungsplattform FINO3 wurden in dieser Arbeit erstmals verschiedene Schwarmintelligenzverfahren auf das Problem der Scholtewellendispersionsanpassung angewendet. Die Applikati-

11 ZUSAMMENFASSUNG UND DISKUSSION DER ERGEBNISSE

on dieser Verfahren induzierte ebenfalls eine neue Methode zur Berechnung des Anpassungsmaßes von vorwärts modellierten und gemessenen Dispersionskurven. Es wurde des weiteren eine Methode zum automatischen Extrahieren von Dispersionskurvenpunkten aus den gemessenen Phasenlangsamsamkeits-Frequenz-Spektren auf der Basis kohärenter Segmente (Abschnitt 6.2) entwickelt und erfolgreich eingesetzt. Ein Anreißen der Dispersionskurven geordnet nach verschiedenen Moden in den gemessenen Spektren ist nun nicht mehr nötig, genauso wie die modenabhängige Zuordnung von Dispersionskurvenpunkten der modellierten Spektren. Um eine weitere Fehlerquelle der Methodik zu minimieren, wurde darüber hinaus ein hochauflösendes Seebodengeophon entwickelt, welches durch eine maximale Abtastfrequenz von 11 kHz die Signale hochfrequenter seismischer Quellen auflösen kann. Damit ist die Möglichkeit gegeben, das System unter Wasser durch eine Laufzeitinversion sehr genau zu relokalisieren. Der Einsatz während einer Baugrundvoruntersuchung am FINO3 Standort ergab eine statistische Positionsauflösung von 4 m. Mit dem Einsatz dieses Systems kann damit eine Fehlerquelle der Offsetberechnung und damit der Phasenlangsamsamkeiten umgangen werden.

Die Schwarmintelligenzoptimierungsmethoden Partikel Schwarm Optimierung, PSO, Künstliche Bienenkolonie Algorithmus, ABC, und verschiedene Varianten der PSO wurden untereinander und mit gängigen stochastischen Optimierungsverfahren wie der Evolutionsstrategie mit adaptiver Schrittweite, CMA-ES, und dem Metropolis- oder SA-Algorithmus an verschiedenen synthetischen Modellen verglichen. Eine hybride Variante aus geladener PSO und Downhill-Simplex zur Unterstützung der lokalen Suche zeigte sich hierbei als besonders effizient und hochauflösend. Durch den Einsatz dieser Variante konnte eine algorithmische Auflösung, also die Auflösung einer Anpassung von synthetischen Daten ohne Rauschen, von im Mittel 15 m/s (6.5%) in den Scherwellengeschwindigkeiten erreicht werden und damit auch das neue Qualitätsmaß erfolgreich getestet werden. Die Auflösungsangabe beschreibt hierbei die Standardabweichungen der Parameter von Lösungsmodellen mehrfacher Inversionsdurchgänge mit verschiedenen Startpopulationen. Das mittlere Modell der verschiedenen Inversionsdurchgänge zeigte eine sehr gute Anpassung an das bekannte Lösungsmodell mit welchem die Eingangsdaten erzeugt wurden. Die Unterschiede der beiden Modelle liegen hier im Mittel bei nur 6 m/s (siehe Tabelle 8.5). Das Verfahren liegt damit im Auflösungsbereich von den am Anfang von Kapitel 5 abgeschätzten erwarteten Änderungen in der Scherwellengeschwindigkeit bei einer Belastung durch Offshorebauwerke. Eine Untersuchung, bei der künstliche Fehler auf die Phasenlangsamsamkeitspunkte der Testdaten addiert wurden, zeigte wie zu erwarten eine Verschlechterung der Auflösung, jedoch lag diese in flachen Bereichen der Modelle immer noch bei 18 m/s (im Mittel 8.5%). Die angegebenen Auflösungen beschreiben die Glaubwürdigkeit eines einzelnen Inversionsdurchganges. Für eine hochauflösende Überwachung ist jedoch die Durchführung vieler Inversionsdurchgänge verschiedener

Startpopulationen zu empfehlen, so daß ein statistisch relevantes mittleres Modell als beste Anpassung bestimmt werden kann.

Der Vergleich verschiedener Modifikationen der Partikel Schwarm Optimierung in Kapitel 8 zeigte, daß große Fortschritte im Bereich der Auflösung und Anpassung hauptsächlich durch die Implementierung einer Neustartmöglichkeit und einer zusätzlichen lokalen Suche des global besten Partikels erreicht werden. Die Neustartfunktion ermöglicht es dem Algorithmus, bei stagnierter Konvergenz eine Reinitialisierung des Verfahrens durchzuführen. Die Vorteile solcher selbstinduzierter Neustarts liegen nicht nur in der erhöhten Abtastung des Parameterraumes durch eine Vielzahl von Inversionsdurchgängen und der damit verbundenen Zugänglichkeit einer Vielzahl von Lösungen, sondern auch in der Möglichkeit, bereits gefundene Minima in nachfolgenden Suchen auszuschließen. Maßnahmen zur Erhöhung der globalen Exploration des Parameterraumes, wie die Neustartvariante, aber auch der getestete geladene Schwarm führten so zu signifikanten Verbesserungen der Anpassung. Dies erinnert noch einmal an die hohe Multimodalität des Problems, welche erhöhte Auflösung nur durch effizientes Absuchen möglichst vieler Minima erlaubt. Eine gesteigerte lokale Konvergenz ist für ein effizientes Abtasten dieser einzelnen Minima ebenso wichtig, so daß sich auch durch Einführen einer lokalen Suche für das global beste Teilchen nicht nur die Konvergenz des Verfahrens an sich garantieren, sondern auch die wohl signifikanteste Verbesserung der Auflösung erreichen ließ. Der Schwarm stellt somit die Konvergenz in Bereiche guter Anpassung bereit und der Downhill-Simplex führt dort eine lokale Konvergenz in die Minima durch (Abbildung 11.1 soll dieses Verhalten illustrieren.).

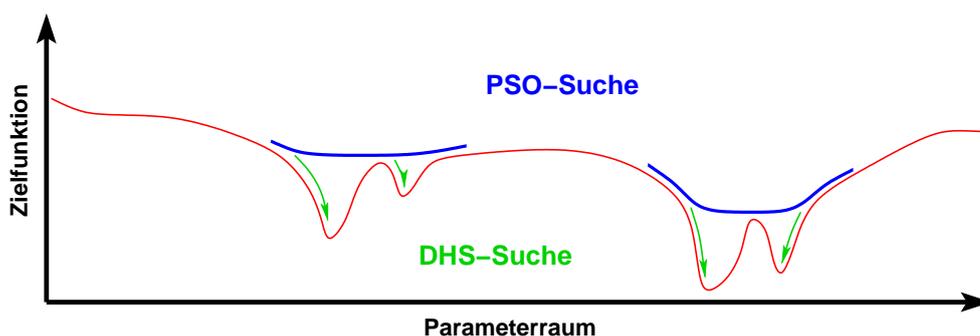


Abbildung 11.1: *Illustration des Zusammenspiels von PSO und DHS im hybriden Ansatz DHS-PSO. Die blauen Bereiche stellen lokale Bereiche guter Anpassung dar.*

Der Nachteil erhöhter Exploration durch multiple Neustarts und forciertes lokaler Ausbeutung liegt auf der Hand und spiegelt sich in einer Zahl von 400000

11 ZUSAMMENFASSUNG UND DISKUSSION DER ERGEBNISSE

Vorwärtsmodellierungen wieder. Diese Zahl steht aber, bei der gewählten Größe des Parameterraumes und bei einer Auflösung von 5 m/s in der Scherwellengeschwindigkeit und 1 m in den Schichtmächtigkeiten für den fünf Schichten Fall, $4.0 \cdot 10^{16}$ möglichen Modellen einer Rastersuche gegenüber. Umgekehrt entspräche eine Rastersuche mit den 400000 Rechnungen der Schwarmoptimierung einem Scherwellengeschwindigkeitsbereich von ca. 10.0 m/s in jeder Schicht. Der hohe Rechenaufwand könnte sicher noch verringert werden, könnte statt des Downhill-Simplex-Verfahrens ein Gradientenverfahren mit adaptiver Schrittweite gewählt werden. Hier zeigten sich jedoch Instabilitäten bei der numerischen Approximation des Gradienten der Zielfunktion, so daß eine solche Anwendung nicht ratsam ist.

Der Aufwand an Vorwärtsmodellierungen wird durch gut aufgelöste Modelllösungen gerechtfertigt, aber auch durch gut angepasste Datenbeispiele, die trotz geringer a priori Informationen und großem Parameterraum erreicht werden konnten. Ausschließlich die Information über die Anzahl der zugrunde liegenden Schichten ist zunächst zu erbringen, ließe sich aber auch durch eine L-Kurven Analyse untersuchen. Eine Abhängigkeit von Startmodellen, welche durch den Wissenschaftler gewählt werden müssen, liegt nicht mehr vor, jedoch muß der Parameterraum in seiner Größenordnung definiert werden.

Die erreichten Anpassungen konnten mit der in Abschnitt 6.2 vorgestellten neuen, modenzuordnungsfreien Vergleichsmethode erbracht werden. Dies spricht dafür, daß die aufgezeigten möglichen Zuordnungsfehler der angewendeten nächster Nachbar Methode durch die Menge der Partikel und die Wahl der L_1 Norm statistisch geringes Gewicht erhalten. Die Methodik macht es ebenfalls möglich, ohne ein manuelles Anreißer der Dispersionskurven, eine Spektralmatrix von automatisch extrahierten Amplitudenmaxima direkt in die Inversion als Datengrundlage einzugeben. Auch das Zuordnen der Amplitudenmaxima in den modellierten Spektren zu den Moden ist somit nicht nötig. Es bleibt jedoch in der Verantwortung des Wissenschaftlers, eventuelle Fehlpicks oder Ausreißer der Dispersionskurvenpunkte manuell aus den Matrizen der Eingangsdaten zu entfernen.

Der Einsatz stochastischer Methoden und die damit zur Verfügung stehende große Zahl von vorwärts modellierten Datensätzen und ihres Anpassungsmaßes ermöglicht ferner eine automatische Analyse von Parameterkorrelationen und eine repräsentative Auflösungsabschätzung. Das Verfahren benötigt hierbei nicht eine statistisch hinreichende Zahl von Inversionslösungen zur Bestimmung eines mittleren Modells und Standardabweichungen, sondern basiert auf einer Abschätzung der Zustandsumme des Nahbereichs der Lösungen aufgrund der Auswertung von Modellen nur weniger Inversionsdurchgänge. Die Analyse hängt jedoch sehr stark von der Wahl der Anzahl der genutzten Modelle und Populationen ab. Es konnte

aber eine Auswahl getroffen werden, welche eine Auflösungsabschätzung und ein Erwartungsmodell vergleichbar mit dem mittleren Modell und der Standardabweichung von 56 Inversionsdurchgängen ermöglicht.

Die Anpassung unterschiedlicher Datensätze in einer gemeinsamen Inversion kann, aufgrund der verbesserten Datenlage und ggf. gemeinsamer Parameter, Mehrdeutigkeiten verringern und so ebenfalls Auflösungen verbessern. Aus diesem Grund wurde in Teil III versucht, eine gemeinsame Anpassung von synthetischen Reflektivitätsdaten einer Zero-Offset Sektion und den Dispersionskurven von Scholtewellen zu realisieren. Hierzu wurde ein Pareto-Dominanz basierter Mehrzieloptimierungsansatz auf Basis der Partikel-Schwarm-Optimierung implementiert und getestet. Der Algorithmus wurde dann erfolgreich an einem synthetischen Datensatz bestehend aus einer modellierten Reflektivitätsreihe und Dispersionskurven getestet. Ausgehend von einem bekannten Modell der Kompressionswellengeschwindigkeit lieferte die gemeinsame Inversion der Datensätze vernünftige Anpassung für Dichte und Scherwellengeschwindigkeit. Es zeigte sich jedoch eine große Streuung der Lösungen und im Vergleich zu den einfachen Inversionen mit der cDHS-PSO eine schlechtere Auflösung der Modellstruktur.

Ein Vergleich der Ergebnisse mit Resultaten einer Inversion mit erhöhter Partikelzahl konnte Anpassungsverbesserungen aufzeigen. Dies legt die Vermutung nahe, daß noch immer eine hohe Multimodalität das Problem dominiert und ebenso hohe Anforderungen an das genutzte Optimierungsverfahren stellt, wie es bereits bei der Einzelinversion der Scholtewellendispersion der Fall war. Es gilt also auch für die betrachtete Mehrzieloptimierung, weitere Verfahren und Verfahrensvarianten zu testen, die eine noch effektivere Exploration der vielen Minima dieses hochdimensionalen Problems ermöglichen. Ein erster Ansatz mit einer Abwandlung des DHS als Pareto-Dominanz orientiertes Verfahren brachte jedoch keine nennenswerten Fortschritte.

Die Anwendung des cDHS-PSO-Verfahrens auf verschiedene Datenbeispiele an der FINO3-Lokation in der Nordsee sowie in der Ostsee lieferte 2D-Modelle guter Auflösung und lateraler Kohärenz (Kapitel 10). Ein Vergleich mit hochauflösenden reflexionsseismischen Daten der untersuchten Profile zeigte hohe Korrelation. Die meisten grundlegenden Charakteristika bzw. Horizonte der Reflexionssektionen ließen sich in den invertierten Pseudo-2D-Modellen wiederfinden. Die Auswahl der Lösungsmodelle der einzelnen lokalen Dispersionsanpassungen aus den multiplen Lösungen der Neustartdurchgänge geschah sowohl nach dem Qualitätswert aber auch nach der Kohärenz entlang des gesamten Profils. So liefert die Exploration multipler Optima durch die cDHS-PSO die Möglichkeit einer unabhängigen Auswahl. Die Tatsache, daß auch die Korrelation zur Wahl genutzt werden mußte, zeigt jedoch, daß die lokale Exploration eines Minimums durch das Optimierungsverfahren mitunter nicht ausreichend ist und womöglich durch nicht gradientenorientierte Verfahren nicht weiter verbessert werden kann. Ein

11 ZUSAMMENFASSUNG UND DISKUSSION DER ERGEBNISSE

zum Nachbarmodell kohärentes Modell mit aber etwas schlechterem Qualitätswert zeigt, daß die lokale Konvergenz in das Minimum nicht beendet ist.

Neben den profilhaften Modellen der Beispiele aus Kapitel 10 konnte mit der cDHS-PSO ein gut aufgelöstes FINO3-Scherwellengeschwindigkeitshintergrundmodell invertiert werden. Dieses Modell basiert auf den Inversionen lokaler Spektren entlang von sieben Profilen, welche fächerförmig angeordnet den Bereich des zukünftigen Baugrundes der FINO3 Plattform abdecken. Das Modell wurde dann mit Cubic-Splines in drei Dimensionen interpoliert. Das resultierende pseudo 3D-Modell zeigt nur geringe laterale Variationen der Scherwellengeschwindigkeit innerhalb der dargestellten Tiefenschnitte von 50 m/s bis 65 m/s, welche aber aufgrund der Kohärenz innerhalb der einzelnen Profile und einem abgeschätztem Auflösungsvermögen von 8% als glaubwürdig erscheinen.

12 Abkürzungsverzeichnis

Die folgende Liste enthält diejenigen Symbole und Abkürzungen, welche in dieser Arbeit wiederholt verwendet werden. Die Zahlen am Ende jeder Zeile entsprechen der Seitenzahl auf welcher das Symbol oder die Abkürzung eingeführt wird.

- ABC** Artificial Bee Colony Algorithm; künstlicher Bienenschwarm Algorithmus, 75
- BFGS** Broyden–Fletcher–Goldfarb–Shanno Quasi-Newton-Sekantenverfahren, 37
- BH-SA** Basin Hopping - Simulated Annealing, 79
- CMA-ES** Covariance Matrix Adaptive Evolution Strategy; Kovarianzmatrix-Adaptions-Evolutionsstrategie, 80
- CMAES-PSO** Hybridverfahren aus Kovarianzmatrix-Adaptions-Evolutionsstrategie und Partikel-Schwarm-Optimierung, 83
- cPSO** Geladene Partikel-Schwarm-Optimierung, 66
- cDHS-PSO** Hybridverfahren aus Downhill-Simplex und geladener Partikel-Schwarm-Optimierung, 66 & 73
- CRG** Common Receiver Gather; Entfernungsabhängige Sektion mit gemeinsamem Empfänger, 21
- (c+r)DHS-PSO** Hybridverfahren aus Downhill-Simplex und geladener und rasternder Partikel-Schwarm-Optimierung, 68 & 73
- DHS** Downhill-Simplex-Verfahren, 71
- DHS-PSO** Hybridverfahren aus Downhill-Simplex und Partikel-Schwarm-Optimierung, 73
- EA** Evolutionäre Algorithmen, 80
- ES** Evolutionsstrategie, 80

12 Abkürzungsverzeichnis

- FINO** Forschungsplattformen in Nord- und Ostsee, 4
- GCPSO** Guaranteed Convergence Particle-Swarm-Optimization;
Partikel-Schwarm-Optimierung mit garantierter Konvergenz, 64
- GPS** Global Positioning System, 23
- MASW** Multichannel Analysis of Surface Waves, 31
- MOPSO** Multi-Objective-Particle-Swarm-Optimization;
Mehrziel-Partikel-Schwarm-Optimierung, 128
- OBS** Ozeanbodenseismometer, 21
- PSO** Particle-Swarm-Optimization; Partikel-Schwarm-Optimierung, 60
- rPSO** Rasternde Partikel-Schwarm-Optimierung, 68
- SA** Simulated Annealing Algorithmus, 78
- SASW** Spectral Analysis of Surface Waves, 3
- SBG** Seebodengeophon, 27
- SQP** Sequentielle Quadratische Programmierung; 37
- v_p Kompressionswellengeschwindigkeit in [m/s], 34
- v_s Scherwellengeschwindigkeit in [m/s], 34

Literaturverzeichnis

- [1] Aki, K. & Richards, P.G., 1980, Quantitative Seismology. W.H. Freeman and Company
- [2] Alvers, M.R., 1998. Zur Anwendung von Optimierungsstrategien auf Potentialfeldmodelle. Dissertation, Fachbereich Geowissenschaften der Freien Universität Berlin
- [3] d'Arnaud Gerkens, J.C., 1989. Foundation of Exploration Geophysics. Amsterdam: Elsevier.
- [4] Ayres, A., Theilen, F., 1999. Relationship between P- and S-wave velocities and geological properties of near-surface sediments of the continental slope of the Barent Sea. *Geophysical Prospecting*, 47, 431-441
- [5] Baggett, J. S., Skahill, B. E., 2008. More efficient evolutionary strategies for model calibration with watershed model for demonstration. American Geophysical Union, Fall Meeting 2008, Abstract H43D-1043
- [6] Beaty, K. S., Schmitt, D. R., and Sacchi, M., 2002. Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure: *Geophysical Journal International*, 151, 622–631
- [7] Blackwell, T.M. & Bentley, P.J., 2002. Dynamic Search with Charged Swarms. *Proceedings of the Genetic and Evolutionary Computation Conference*, p. 19-26
- [8] Boggs, P. T., and Tolle, J. W., 1995. Sequential quadratic programming: *Acta Numerica*, 1–51
- [9] Bohlen, T., 2001. Parallel 3-D viskoelastic finite difference seismic modelling. *Computers & Geosciences* 28(8), 887-899
- [10] Bohlen, T., Kugler, S., Klein, G., Theilen, F., 2004, 1.5-D Inversion of lateral variation of Scholte wave dispersion, *Geophysics*, vol. 69, no. 2, 330-344
- [11] BSH: German Federal Department for Maritime Shipping and Hydrography, 2003. Standard foundation soil analysis for offshore wind turbines. Report 2003

Literaturverzeichnis

- [12] BSH: German Federal Department for Maritime Shipping and Hydrography, 2002. Festlegung des Eignungsgebietes Westlich Adlergrund
- [13] Buchen, P.W. & Ben-Hador, R., 1996. Free-mode surface-wave computations. *Geophys. J. Int.*(1996), 124, 869-887
- [14] Bussat, S., 2006. Anregung von Grenzflächenwellen am Meeresboden durch Quellen in der Wassersäule. Dissertation, Institut für Geowissenschaften, Abteilung Geophysik, Universität Kiel
- [15] Castagna, J.P. & Backus, M.M., 1993. Offset dependent reflectivity - theory and practice of AVO analysis. SEG, Investigations in Geophysics Nr. 8
- [16] Dal Moro, G., Pipan, M., Gabrielli, P., 2007. Rayleigh wave dispersion curve inversion via genetic algorithms and Marginal Posterior Probability Density estimation. *Journal of Applied Geophysics* 61, 39-55
- [17] Dal Moro, G., Pipan, M., 2007. Joint Inversion of surface wave dispersion curves and reflection travel times via multi-objective evolutionary algorithms. *Journal of Applied Geophysics* 61, 56-81
- [18] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan, 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolution-ary Computation*, 6(2):182-197
- [19] Eberhart R.C. and Kennedy J. 1995. A New Optimizer using Particle Swarm Theory, Proceedings Sixth Symposium on Micro Machine and Human Science, pp. 39-43
- [20] Engelbrecht, A.P., 2005. Fundamentals of Computational Swarm Intelligence. Wiley Publishing Company
- [21] Essen, H.-H., Grevemeyer, I., Herber, R. & Weigel, W., 1998. Shear-wave velocity in marine sediments on young oceanic crust: constraints from dispersion analysis of Scholte waves. *Geophysical Journal International*, 132, 227-234
- [22] Ewing, W., Jardetzky, W., and Press, F., 1957. Elastic waves in layered media. McGraw-Hill Book Co.
- [23] Rabbel, W., Feeser, V., 2009. Geotechnische Grundlagenforschung zur Berücksichtigung des dynamischen Verhaltens des Untergrundes im Hinblick auf die standsichere und wirtschaftliche Gründung. Abschlussbericht BMU/PTJ-FINO3 0327533
- [24] Forbriger, T., 2003a, Inversion of shallow-seismic wavefields: I. Wavefield transformation. *Geophysical Journal International*, 153, 719-734

- [25] Gehrman, T., Gimpel, P. & Theilen, F., 1984. Marine shear wave profiling. 54th Ann. Intern. Meeting Soc. of Expl. Geophys., S13.5
- [26] Gimpel, P., 1987. Marine flachseismische Untersuchungen in der Kieler Bucht unter besonderer Berücksichtigung von Scherwellenmessungen. Dissertation, Institut für Geowissenschaften, Abteilung Geophysik, Universität Kiel
- [27] Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Publishing Company
- [28] Hansen, N. & Ostermeier, A., 2001. Completely Derandomized Self-Adaptation in Evolution Strategies. *Evolutionary Computation* 9(2): 159-195
- [29] Haskell, N.A., 1951. The dispersion of surface waves on multilayered media. *Bulletin of the Seismological Society of America*
- [30] Karaboga, D. & Basturk, B., 2007. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of Global Optimization*, 39, 459-471
- [31] Kearey, P., Brooks, M. & Hill, I., 2006. *An Introduction to Geophysical Exploration* (third edition). Blackwell Publishing
- [32] Kirkpatrick, S. & Vecchi, M., 1983. Optimization by simulated annealing. *Science*, 220
- [33] Klein, G., Bohlen, T., Theilen, F., Kugler, S., Forbriger, T., 2005, Acquisition and Inversion of dispersive seismic waves in shallow marine environments, *Marine Geophysical Researches*. 26:287–315
- [34] Koduru, P., Das, S., Welch, S. M., 2006. A Particle Swarm Optimization-Nelder Mead Hybrid Algorithm for Balanced Exploration and Exploitation in Multidimensional Search Space. *Proceedings, International Conference on Artificial Intelligence, Las Vegas, Nevada*, 457–464
- [35] Kosmol, P., 1993. *Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben*. Teubner Verlag
- [36] Kugler, S., 2001. Bestimmung der Scherwellengeschwindigkeit flacher mariner Sedimente aus der Analyse von Scholtewellen. Diplomarbeit, Institut für Geowissenschaften, Abteilung Geophysik, Universität Kiel
- [37] Kugler, S., Bohlen, T., Bussat, S., Klein, G., 2005. Variability of Scholte-wave Dispersion in Shallow-water Marine Sediments. *JEEG*, June 2005, Volume 10, Issue 2, pp 203-218

Literaturverzeichnis

- [38] Kugler, S., Bohlen, T., Forbriger, T., Bussat, S. and Klein, G., 2007. Scholte-wave tomography for shallow-water marine sediments, *Geophysical Journal International*, 168, 2, 551-570
- [39] Luke, B., and Stokoe, K., 1998. Application of SASW method underwater. *Journal of Geotechnical and Geoenvironmental Engineering*, 124, 523-531
- [40] Marsaglia, G., Narasimhan, B. & Zaman, A., 1990. A random number generator for PC's. *Computer Physics Communications* 60, 345—349
- [41] Martinez, M.D., Lana, X., Olarte, J., Badal, J. & Canas, J.A., 2000. Inversion of Rayleigh wave phase and group velocities by simulated annealing. *Physics of the Earth and Planetary Interiors* 122, 3–17
- [42] McMechan, G. and Yedlin, M., 1981. Analysis of dispersive waves by wave field transformation. *Geophysics*, 46: 869-874
- [43] Meegan, G. D., Hamilton, M. F., Il'inskii, Yu. A. & Zabolotskaya, E. A., 1999. Nonlinear Stoneley and Scholte waves. *Journal of Acoustical Society of America* 106 (4), Pt. 1
- [44] Menke, W., 1989. *Geophysical Data Analysis: Revised Edition: Discrete Inverse Theory (International Geophysics)*. Academic Press Inc.
- [45] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller A., Teller, E., 1953. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* 21, 1087-1092
- [46] Millonas, M., 1994. *Swarms, Phase Transitions and Collective Intelligence. Artificial Life III*, Addison-Wesley, 417-445
- [47] Mishra, S. K., 2006. Repulsive Particle Swarm Method on Some Difficult Test Problems of Global Optimization. MPRA Paper No. 1742
- [48] Müller, G., 1986. *Theorie elastischer Wellen. Vorlesungsskript Theoretische Geophysik*, Universität Frankfurt
- [49] Muyzert, E., 2000. Scholte wave velocity inversion for a near surface S-velocity model and P-S-statics: 71st Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1197–1200
- [50] Nath, K.S., Chakraborty, S., Singh, S.K. und Ganguly, N., 1999. Velocity Inversion in cross-hole seismic tomography by counter-propagation neural network, genetic algorithm and evolutionary programming techniques, *Geophys. J. Int.*, 138, 108-124

- [51] Nelder, J.A. & Mead, R.A., 1965. A simplex method for function minimization, *Computer Journal*, 7(4), 308-313
- [52] Pant, M., Thangaraj, R. & Abraham, A., 2007. A New PSO Algorithm with Crossover Operator for Global Optimization Problems. *Innovations in Hybrid Intelligent Systems*, ASC 44, pp. 215–222
- [53] Park, C. B., Miller, R. D., and Xia, J., 1998, Imaging dispersion curves of surface waves on multi-channel record: 68th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1377–1380
- [54] Park, C., Miller, R., Xia, J., Ivanov, J., Hunter, J., Good, R., and Burns, R., 2000. Multichannel analysis of underwater surface waves: 70th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1303–1306
- [55] Parsopoulos, K.E. & Vrahatis, M.N., 2004. On the Computation of All Global Minimizers Through Particle Swarm Optimization. *IEEE Transactions on evolutionary computation*, Vol. 8, No. 3
- [56] Prasad, M. & Meissner, R., 1992. Attenuation mechanisms in sands: Laboratory versus theoretical (Biot) data. *Geophysics* (57), 710-719
- [57] Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 2007. *Numerical Recipes: The Art of Scientific Computing*. Third Edition, Cambridge University Press
- [58] Rechenberg, I., 1994. *Evolutionsstrategie '94*. Frommann-Holzboog
- [59] Richwien, W., Lesny, K., 2004. Windfarmen in der Nordsee, Fundamente für Windmühlen auf hoher See; in: *Essener Unikate - Berichte aus Forschung und Lehre, Ingenieurwissenschaften*, Heft 23, Universität Duisburg-Essen
- [60] Ritzwoller, M. H., and Levshin, A. L., 2002. Estimating shallow shear velocities with marine multicomponent seismic data: *Geophysics*, 67, 1991–2004
- [61] Safar M.H. 1976, The radiation of acoustic waves from an Airgun, *Geophysical prospecting* 24, 756-772
- [62] Shaw, R., Srivastava, S., 2007. Particle swarm optimization: A new tool to invert geophysical data. *Geophysics*, VOL. 72, Nr.2, F75-F83
- [63] Silva, A., Neves, A., Costa, E., 2002. An Empirical Comparison of Particle Swarm and Predator Prey Optimisation. *Proceedings of the Thirteenth Irish Conference on Artificial Intelligence and Cognitive Science*, v. 2464, 103-110, Springer-Verlag

Literaturverzeichnis

- [64] Scholte, J.G., 1942. On the Stoneley wave equation. Proc. K. Ned. Akad. Wet. 45, Pt. 1: 20–25, Pt. 2: 159–164.
- [65] Schön, J., 1983. *Petrophysik*. Akademie Verlag, Berlin
- [66] Thomson, W.T., 1950. Transmission of elastic waves through a stratified solid medium. J. Appl. Phys., 21: 89-93
- [67] van den Bergh, F., 2002. *An Analysis of Particle Swarm Optimizers*. PhD thesis, Department of Computer Science, University of Pretoria, Pretoria, South Africa
- [68] Vogel, H., 1995. *Gerthsen Physik*. Springer-Verlag
- [69] Wales, D.J., Doye, P.K.D., 1997, Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. J. Phys. Chem. A, 101, 5111-5116
- [70] Wang, R., 1999. A simple orthonormalization method for stable and efficient computations of Green's functions. Bulletin of the Seismic Society of America, 89(3) 733-741
- [71] Wiemann J., Lesny K., Richwien W. 2004, Evaluation of Pile Diameter Effects on Soil-Pile Stiffness, Dokumentation of the 7th German Wind Energy Conference DEWEK, Wilhelmshaven
- [72] Wilken, D., Wölz, S., Müller, C., Rabbel, W., 2007. Numerische Studie zur seismischen Analyse von Kompaktierungszonen am FINO3 Monopile. Tagung der Deutschen Geophysikalischen Gesellschaft 2007. Abstract UI 273
- [73] Wilken, D., Wölz, S., Müller, C. & Rabbel, W., 2009. FINOSEIS: A new approach to offshore-building foundation soil analysis using high resolution reflection seismic and Scholte-wave dispersion analysis. Special Issue: Techniques of near-surface seismic and georadar imaging, Journal of Applied Geophysics, 68, 117-123
- [74] Whittaker, E. & Robinson, G., 1946. *The Calculus of Observations*. pp. 303–316
- [75] Xie, X., Zhang, W., Yang, Z., 2002. Adaptive Particle Swarm Optimization in Individual Level. Proceedings of the Sixth International Conference on Signal Processing, v. 2, 1215-1218

Danksagung

Ich möchte mich an dieser Stelle bei all denen bedanken, die zum Gelingen dieser Arbeit maßgeblich beigetragen haben. Dies sind

- **Herrn Prof. Dr. Wolfgang Rabbel**, der die Anfertigung der vorliegenden Arbeit überhaupt erst ermöglicht hat und mir dabei die letzten Jahre stets mit Rat und Tat und Engagement zu Seite stand.
- **Dr. Christof Müller** für seine geduldige Hilfestellung während meiner ersten Zeit als Doktorand; für seine Begeisterung, die er mir für viele Projekte vermitteln konnte und für die Motivation und Bestätigung, meine Ideen zu verfolgen.
- **Herrn Prof. Dr. Hans-Jürgen Götze** für die Übernahme des Korreferats, seine vielen Hinweise und für engagierte Diskussionen
- **den Mitgliedern der AG** für das sehr angenehme und kreative Arbeitsumfeld
- **Dr. Susanne Wölz** für die tolle Zusammenarbeit und für ihre Hinweise und immer aufbauenden Worte, die mir den Einstieg in die Geophysik sehr erleichtert haben.
- **Detlef Schulte-Kortnack** für seine beständige Unterstützung in technischen Fragen, seine stete Hilfsbereitschaft und insbesondere für die tolle gemeinsame Arbeit am Seismometer-Projekt.
- **Dr. Sven Willert** für gemütliche Diskussionen mit vielen Zukunftsvisionen und seine immer hilfreiche Kritik, ohne die ich mir einige wichtige Fragen sicherlich nicht gestellt hätte. Es geht weiter...
- **Jan-Hendrik Palic** nicht nur für das schöne Arbeitsklima im Büro, sondern auch für Rat und Hilfe in Dingen der Informatik und Administration, insbesondere aber für den Satz „Ich hab da mal irgendwas von Schwärmen gehört“.
- **Claudia Haase** ebenfalls für das tolle Arbeitsklima in unserem interdisziplinären Büro und für die Diskussionen und Anmerkungen in Dingen der Inversionsproblematik.

Literaturverzeichnis

- **Dr. Christian Henning, Dr. Martin Ziegler und Michael Becker** für ihre Hilfe und ihr offenes Ohr für Probleme während der Physiker-Kaffeepausen am Mittwoch oder an diversen Abenden.
- **Tobias Jokisch** für seine Hilfe und Freundschaft
- **Dr. Sascha Bussat & Dr. Simone Kugler** für ihre Einweisung in die Kunst der Scholtewellenmessung und für viele ideenreiche Diskussionen
- **Claudia Podolski** für die gute Zusammenarbeit und den tollen Start in das FINO-Projekt
- **Tomislav Hrkac** für seine Freundschaft und ein gemeinsames Hobby, welches stets den Blick auf das Wesentliche schärft.
- **Prof. Dr. Sebastian Krastel** für die gute Zusammenarbeit auf See und an Land
- **Ercan Erkul** für seine Hilfsbereitschaft und seinen Einsatz für unsere Meßfahrten
- **Prof. Dr. Detlev Köster**, ohne den ich nie den Ablauf und die Struktur wissenschaftlicher Arbeit gelernt hätte.

- meinen lieben **Eltern Sybille und Reiner Wilken** und meiner **Großmutter Ilse Wilken**
- ganz besonders **Meike Quitzau**, die mich so sehr unterstützt hat. Du warst und bist mein Halt.

Erklärung

Diese Dissertation ist von mir in Inhalt und Form unter Anleitung meiner akademischen Lehrer und nur mit den angegebenen Hilfsmitteln angefertigt worden. Diese Arbeit hat weder in gleicher, noch in ähnlicher Form hier oder an anderer Stelle im Rahmen eines Prüfungsverfahrens vorgelegen.

Kiel, den _____

(Dennis Wilken)

Lebenslauf Dennis Wilken

Geboren am: 18.12.1979 in Oldenburg(Oldb)

Staatsangehörigkeit: deutsch

Familienstand: ledig

Schulbildung

1992-1999 Besuch des Gymnasiums Cäcilienchule in Oldb

07/1999 Allgemeine Hochschulreife

Ersatzdienst

1999-2000 Zivildienst beim Verein für Natur- und Vogelschutz im Hochsauerlandkreis
Einsatz im Bereich der Landschaftspflege in Marsberg

Studium

ab 10/2000 *Christian Albrechts Universität zu Kiel:*
Studium der Physik(Dipl.)
Nebenfach: Anorganische Chemie(Grundstudium),
Numerische Mathematik(Hauptstudium)

10/2002 Physik-Vordiplom

12/2004 Diplomarbeit:
*Diffusionszeitskalen und Akkretionsraten
für Weiße Zwergsterne des Spektraltyps DAZ*

11/2004 Physikdiplom

ab 04/2006 *Christian Albrechts Universität zu Kiel:*
Promotionsstudium
am Institut für Geowissenschaften
Abteilung Geophysik