DGK Deutsche Geodätische Kommission
bei der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 628

Michael Heinert

Systemanalyse der seismisch bedingten Kinematik Islands

München 2009

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

ISSN 0065-5325 ISBN 978-3-7696-5040-2

Diese Arbeit ist gleichzeitig veröffentlicht in:
Geodätische Schriftenreihe der Technischen Universität Braunschweig,
Nr. 22, Braunschweig 2008, ISBN 3-926146-17-6

Deutsche Geodätische Kommission
bei der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 628

Systemanalyse der seismisch bedingten Kinematik Islands

Von der Fakultät Architektur, Bauingenieurwesen und Umweltwissenschaften
der Technischen Universität Carolo-Wilhelmina zu Braunschweig
zur Erlangung des Grades Doktor-Ingenieur (Dr.-Ing.)
genehmigte Dissertationon

von

Dipl.-Ing. Michael Heinert

aus Bremervörde

München 2009

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München

ISSN 0065-5325 ISBN 978-3-7696-5040-2

Diese Arbeit ist gleichzeitig veröffentlicht in:
Geodätische Schriftenreihe der Technischen Universität Braunschweig,
Nr. 22, Braunschweig 2008, ISBN 3-926146-17-6

Adresse der Deutschen Geodätischen Kommission:

Deutsche Geodätische Kommission

Alfons-Goppel-Straße 11 ● D – 80 539 München

Telefon +49 – 89 – 23 031 1113 ● Telefax +49 – 89 – 23 031 - 1283 / - 1100
e-mail hornik@dgfi.badw.de ● http://www.dgk.badw.de

Prüfungskommission

Berichterstatter: Univ.-Prof. Dr.-Ing. habil. Wolfgang Niemeier
Univ.-Prof. Dr. rer. nat. Harro Schmeling

Tag der Einreichung: 24.01.2008

Tag der mündlichen Prüfung: 04.06.2008

:

ISSN 0065-5325 ISBN 978-3-7696-5040-2

Für Bastiaan und Hannah.

Summary

The trajectories of the permanent stations of the Global Positioning System (GPS) are effected by several seasonal climatic processes. The question is whether the trajectories of the stations in and around Iceland include a seasonal spreading signal as well.
Iceland is situated, namely, on top of the Mid Atlantic Ridge that marks the plate boundary between North America and Eurasia. Nevertheless, the station trajectories on both sides of the plate boundary are almost identical for several months during the winter period.
To exclude the atmosphere as having the main impact it had to be tested whether the seismicity in and around Iceland is highly correlated with the geometrical motion processes. Therefore, a *semi-parametric* model type had to be created to compute motions across the plate boundary using the seismicity data. A non-parametric model like an artificial neural network becomes semi-parametric, if the input data is transformed into a linear relationship to the output data by using the well-known equations of the physical laws.
Within this investigation this pre-step to the modelling is put in action by transforming earthquake magnitudes into equivalents of motion via energy. These equivalents had to be weighted according to their spatial distribution to describe the impact of a tectonical feature. Volcanically induced events, for example, are assumed not to contribute to the tectonic processes.
Further, the weighted earthquake data had to be brought to formal temporary equidistance to enable recursive modelling, because of the non-equidistance of the occurrence of earthquakes.
To develop such a semi-parametric model, based on suitable network architecture, it is necessary to combine the rules from the *statistical learning theory* with the knowledge of *non-linear optimisation methods* from the operations research. Additionally, the development of a new *robust estimator* based on the limited and re-descending psi-function, is required. Several networks with a different architecture can be tested using such tools. All these methods derive from the *system theory.* The techniques of *time series analysis*, taken from the system theory as well, deliver significant correlations between climate and seismic processes in the same regions.
Using synthetic data a thorough model inference was established to get rid of the typical *black box* character of a neural network. Using the real data it was now possible to define the temporary and spatial energy contributions of each tectonic feature. The existence of seasonality within the spreading process could only be verified sufficiently during the periods of a few years before and after the major earthquakes.
Furthermore, it emerged that, in retrospect, the two major earthquakes that took place in Iceland in 2000 occurred following a cascade of observable events and reliable model patterns within the motion process.
The model inference made clear that the use of the global parameters for the computation of seismic energy did not yield the optimal solution. In any future research it may be necessary to adapt these parameters to the seismicity in and around Iceland.

Inhaltsverzeichnis

1 Einführung

„Aus der Schule wissen wir,
dass die kürzeste Verbindung zwischen zwei Punkten die Gerade ist.
Und aus dem Leben wissen wir,
dass man auf Umwegen schneller ans Ziel kommt.
Warum sollte man also den geraden Weg wählen,
wo es doch so viele bequeme Umwege gibt?“
CURT GOETZ

Island, die Insel aus Feuer und Eis im Nordatlantik, ist bereits seit vielen Jahrzehnten eines der herausragendsten natürlichen Laboratorien, das einen Blick auf die Prozesse unserer Erde ermöglicht. Nur hier kann ein Ort betreten werden, an dem sich in jedem Moment neue Kruste bildet. Überall sonst findet dieser Prozess in der Tiefsee an den mittelozeanischen Rücken statt. Aber Island nimmt zugleich auch eine Sonderstellung ein. Die Insel ist eine Anomalie des Mittelatlantischen Rückens. Ein einzigartiges Zusammentreffen von Prozessen hat die Insel aus der Tiefe des Meeres gehoben. Unter Island befindet sich ein sogenannter *hot spot*, eine Stelle in der Kruste der Erde, wo aus großer Tiefe relativ heißes Material an die Oberfläche gefördert wird.
Mit all diesen Prozessen bildet Island die Grenze zwischen den Platten Nordamerikas und Eurasiens. Die Kontinente, das wissen wir seit ALFRED WEGENER, sind in Bewegung. Auf Island kann man die unterschiedliche Bewegung zwischen zwei Platten unmittelbar messen. Lange hat es gedauert, bis die Plattengrenze in der Örtlichkeit festgelegt werden konnte, denn es ist eine Abfolge aktiver und stabiler Zonen. Es ist keineswegs nur der imposante Grabenbruch Þingvellir, der interessierten Islandtouristen gern als Plattengrenze präsentiert wird.
Jetzt nachdem weitgehend bekannt ist, *wo* überall etwa *welche* Bewegungen auftreten, stellt sich die Frage, welche Bewegungen *wann* auftreten und *wie* sie mit den sicht- und spürbaren Prozessen dieser Insel, nämlich den Erdbeben oder den Spalteneruptionen, verknüpft sind.

1.1 Die Fragestellung

Im Rahmen dieser Arbeit soll die Frage gelöst werden, ob das Auseinanderdriften der Platten Nordamerika und Eurasien am Mittelatlantischen Rücken zwischen den großen Riftepisoden zyklisch oder kontinuierlich verläuft. In den Datenreihen der permanenten Satellitenempfangsstationen des *Global Positioning System* (GPS) finden sich jährliche Zyklizitäten der Lagekoordinaten. Es ist daher zu klären, inwieweit diese Zyklizitäten tatsächlich die Folge einer Veränderung der großräumigen Geometrie sind. Es ist bereits bekannt, dass saisonale Artefakte aus der Signalausbreitung durch die Atmosphäre existieren (BAUER 2003, S. 112). Ebenso gibt es lokale Veränderungen um die Empfangsstation selbst, die wetterbedingt sind, wie Änderungen des Grundwasserspiegels oder Verformungen des Gründungsbauwerkes aufgrund von einseitiger Sonneneinstrahlung und viele andere mehr.
Die übliche Strategie zur Untersuchung von Prozessen und deren Zeitreihen, bei der ein Effekt nach dem nächsten herausgerechnet wird, sei es entweder durch Filterung, wie auch durch Trendbereinigungen oder durch Zerlegung der erregenden Periodizitäten im Spektrum, wird

zur Lösung dieser Fragestellung allein nicht zielführend sein. Ein wetterbedingter saisonaler Prozess auf das Messsystem ist geeignet, die saisonalen Prozesse am Objekt zu verschleiern.
Das Ziel ist die Modellierung der Systembeziehung zwischen den Erdbeben auf Island und der gemessenen Bewegung zwischen den Platten Eurasien und Nordamerika. Dazu sollen Magnituden der Erdbeben durch ein zeit- und raumabhängiges, nichtparametrisches Modell über den Weg der Energie in Bewegung überführt werden. Weist das Resultat des überführenden Modells die gleichen Zyklizitäten auf, wie die gemessenen Daten des GPS, so ist davon auszugehen, dass auch eine geometrische Veränderung als ursächlich angesehen werden muss.
Im Konkreten muss daher der Nachweis erbracht werden, dass die Erdbeben in und um Island Verursacher oder Vermittler von zyklischen Bewegungen sind. Der Nachweis enthält zwei Aspekte: Ist die Seismizität in und um Island repräsentativ für das tektonische Auseinanderdriften im Bereich des Nordatlantik und wenn ja, ist ihr Energieeintrag tatsächlich geeignet, diese Zyklizitäten auszulösen?
Es ist zunächst unumgänglich, die Repräsentativität der tektonischen Prozesse in und um Island zu verstehen und aufzuzeigen. Sowohl die Entstehung und auch die Entwicklung der ozeanischen und kontinentalen Kruste des Nordatlantik als auch die gegenwärtigen tektonischen Gegebenheiten können wichtige Hinweise liefern.
Hierzu ist zu klären, inwieweit die Seismizität eines Gebietes einen Beitrag zur Bewegung liefert und wie sich dieser Beitrag quantifizieren lässt. Zur Lösung dieser Frage soll ein neuer Modelltyp auf der Grundlage einer nichtparametrischen Modellbildung favorisiert werden, da ein vollständiges parametrisches Modell zu viele unerkannte Prozesse mit noch mehr unbekannten Parametern abzubilden hätte.
Dieses Vorgehen erfordert aber auch den Nachweis, dass die gewählten nichtparametrischen Modelle grundsätzlich zur Abbildung der realen Systeme geeignet sind.

1.2 Das Ausgangsproblem

Im Sommer 1999 fand in Südwestisland eine Netzmessung mit dem satellitengestützten Messverfahren des *Global Positioning System* (GPS) statt (Perlt & Heinert 2006; Perlt 2006). Während der Auswertung der hierbei gewonnenen Daten stellte sich die Frage der absoluten Lagerung des gemessenen Netzes in einem globalen Bezugsrahmen. Hierzu boten sich die beiden permanenten GPS-Stationen in Reykjavík, Südwestisland (REYK) und in Höfn, Südostisland (HOFN) des *International GNSS Service* an (Dow et al. 2005). Die beiden Stationen befinden sich nahe des Mittelatlantischen Rückens. An ihm driften die Platten Nordamerika und Eurasien mit einer mittleren Geschwindigkeit von etwa zwei Zentimetern im Jahr auseinander.
Zum Zweck der absoluten Lagerung wurden die GPS-Daten dieser beiden Permanentstationen mit den eigenen Messungen zusammen ausgewertet. Als schwierig erwies sich aber die Auswahl der „richtigen“ Stationskoordinaten dieser beiden permanenten Stationen. Die verfügbaren Tageslösungen des *Jet Propulsion Laboratory* besitzen ein erhebliches Rauschen in allen drei Koordinatenkomponenten (Anh. B.1), so dass eine einzelne Tageskoordinate nicht als geeignete Trägerin der gesuchten Information in Frage kommt.
Die Prozessanalyse der durchgängig verrauschten Koordinatenzeitreihen offenbarte, dass die tektonische Bewegung und damit die Spreizung am Mittelatlantischen Rücken keineswegs konstant ablaufen konnte. Dieses Phänomen verdeutlicht sich in der Betrachtung der Differenz zwischen den beiden Stationen. Nachdem sich im Sommer des Jahres 2000 in der Südisländischen Seismischen Zone zwei Erdbeben mit Momentmagnituden $\mathrm{M}_w = 6,4$ und $\mathrm{M}_w = 6,5$ ereignet hatten (Árnadóttir et al. 2001), zeigte sich in der Differenz der beiden Stationsbewegungen, also der Entfernung beider Stationen zueinander, ein starker Anstieg von zwei Zentimetern als unmittelbare Folge auf diese Beben (Abb. 1.1). Nach diesen beiden Ereignissen öffnete sich der Mittelatlantische Rücken auf der geographischen Breite Islands also in

wenigen Tagen genauso weit wie sonst innerhalb eines ganzen Jahres.
Aus dieser Beobachtung lässt sich folgern, dass Vorgänge in der Südisländischen Seismischen Zone einen Einfluss auf die Stationsbewegungen haben müssen. Damit ergab sich die Frage, ob nicht die gesamte Bewegung jeweils an Erdbeben in diesem speziellen Gebiet und einer weiteren seismisch aktiven Zone im Norden Islands gekoppelt sein können. Im Rauschen der Differenzzeitreihe verborgen fanden sich ferner zyklische Veränderungen. Würde damit auch das Riftverhalten des Mittelatlantischen Rückens in Jahreszyklen verlaufen? Tatsächlich sind die zyklischen Veränderungen im anfänglich verfügbaren Datenfenster von 1998 bis 2000 abgesehen von Phasenverschiebungen auffallend ähnlich: so fallen die Graphen der Erdbebenenergie[1] aus der Südisländischen Seismischen Zone, die Temperaturdaten der lokalen Wetterstation und die Divergenzgeschwindigkeit der beiden permanenten GPS-Stationen nazu zusammen (Abb. 1.3). Damit drängte sich die Schlussfolgerung auf, dass es eine Zyklizität in der Driftbewegung geben könne (Heinert & Perlt 2002; Heinert 2003). Am eindrücklichsten hingegen ist die temporale Identität der Stationstrajektorien diesseits und jenseits der Plattengrenze jeweils im Winter und Frühjahr (Abb. 1.2). Für den Vergleich der beiden Stationen REYK und HOFN fällt dieser Effekt sehr deutlich aus. Doch selbst die etwa 2000 km lange Basislinie zwischen REYK und ONSA, der inzwischen mit in die Betrachtung gezogenen Station Onsala in Südschweden, zeigt diese systematischen Identitäten. Mit einem bloßen klimatischen Einfluss auf das Messsystem ist das nicht einfach zu erklären.

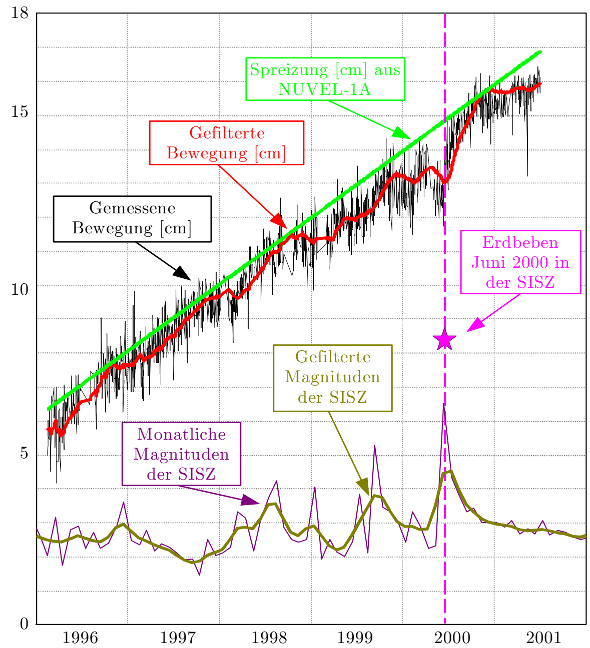


Abb. 1.1: Die theoretische (grün), gemessene und Kalman-gefilterte (rot) Spreizungsbewegung zwischen den GPS-Permanentstationen REYK und ONSA sowie die Seismizität der Südisländischen Seismischen Zone in monatlichen (lila) und gemittelten monatlichen Magnituden (oliv).

Zu dieser Ausgangsthese wurden folgende Kritikpunkte formuliert:

- Eine hohe positive Kreuzkorrelation weise lediglich auf mögliche Kausalitäten hin. Es gäbe jedoch genügend Beispiele für unsinnige Kausalitäten, die aus hohen Kreuzkorrelationen gefolgert werden könnten.
- Die Jahreszyklizität sei ausschließlich ein direkter Einfluss des Wetters auf die GPS-Empfänger. Man könne auch an als stabil anzunehmenden Stationen Zyklizitäten erkennen.
- Eine jahreszeitliche Temperaturverteilung könne nicht kausal für eine etwaige Zyklizität von seismischen Ereignissen sein. Die Wärmeenergie könne unmöglich in die Tiefen propagieren, in welcher sie einen Einfluss auf Erdbeben besitzen könnten.

Diese Kritikpunkte sind berechtigt, daher ist die weitergehende Untersuchung hier erforderlich. Eine zyklische divergente Bewegung an den Plattenrändern für sich genommen, wäre ein

[1] Die Magnituden sind hierzu zunächst in Energie umgerechnet und monatlich aufaddiert worden. Das Ergebnis wurde wieder in eine Magnitude zurückgerechnet.

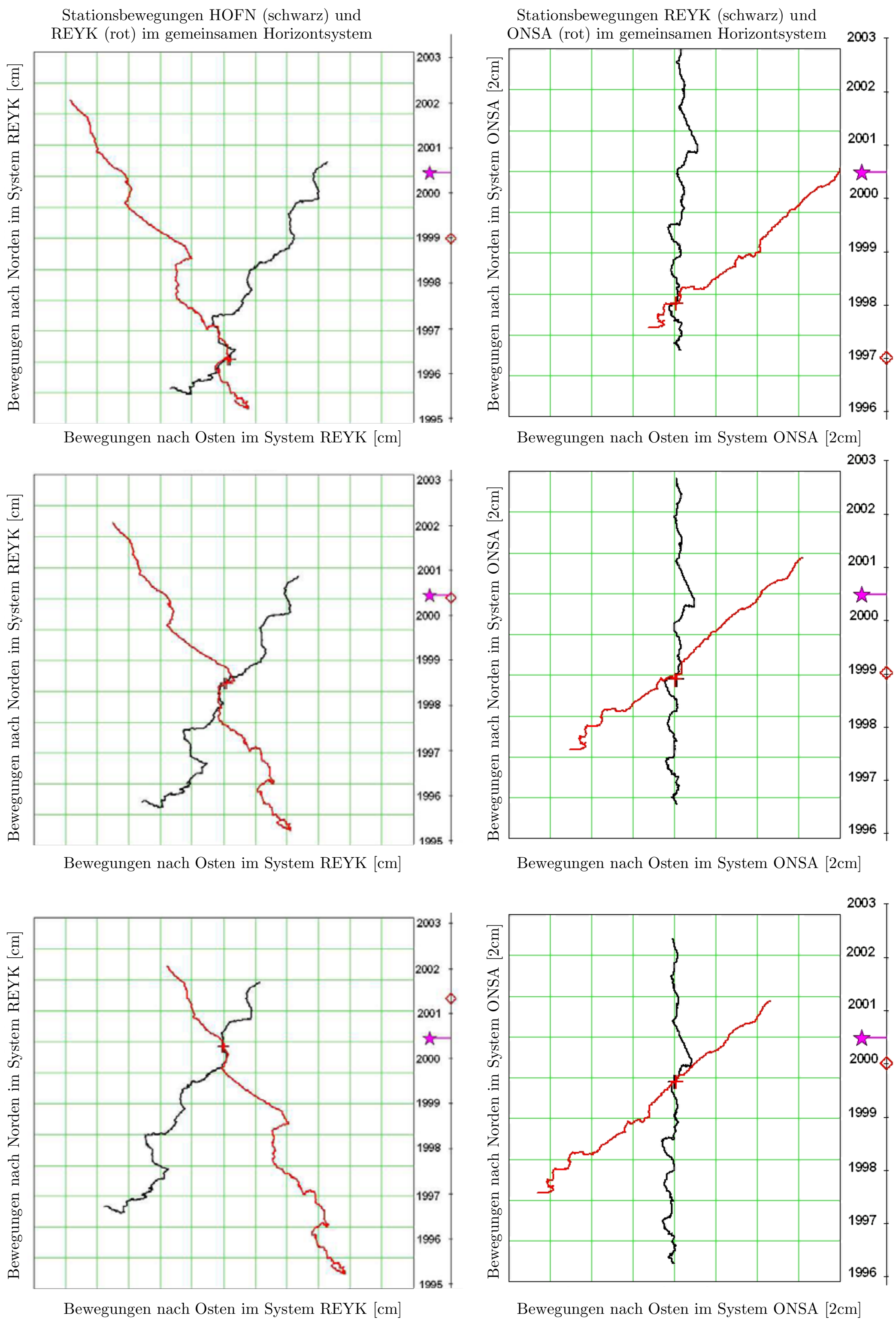


Abb. 1.2: Permanentstationsbewegungen als Trajektorie: Die Stationen REYK auf der Nordamerikanischen Platte und HOFN auf der Eurasischen Platte (links) zeigen jeweils im Winter und Frühling (roter Ring auf der Zeitskala) ein nahezu identisches Bewegungsverhalten. Der Effekt tritt auch zwischen den Stationen REYK und ONSA auf der Eurasischen Platte auf (rechts).

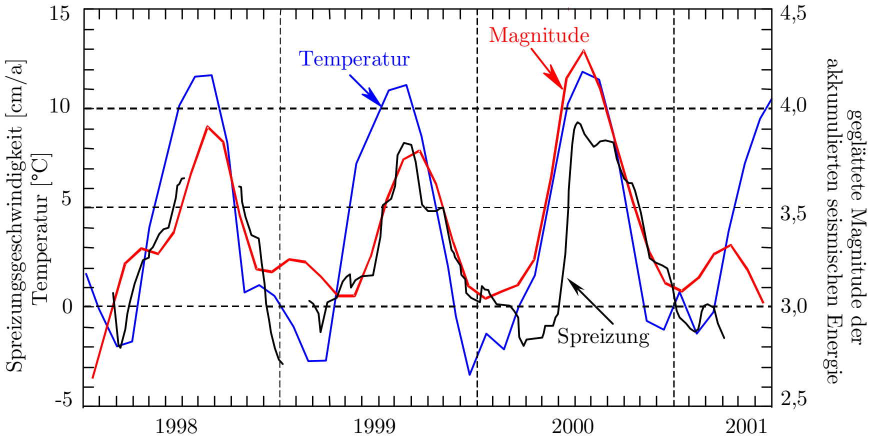


Abb. 1.3: Die Spreizungsgeschwindigkeit (schwarz) aus der Differenz der permanenten GPS-Stationen Reykjavík und Höfn ist hoch korreliert zur freigegebenen seismischen Energie der Südisländischen Seismischen Zone (rot). Beide phasenfrei geglätteten Zeitreihen folgen der lokalen Temperatur der Wetterstation in Selfoss (blau).

Ergebnis, das die Sicht auf die Prozesse im Erdinneren maßgeblich verändern würde. Es hat sich aber mit längeren Zeitreihen herausstellt, dass die Zyklizitäten sowohl der Seismizität der Südisländischen Seismischen Zone als auch der Differenzen zwischen den Stationen vor den Beben des Frühsommers 2000 zunächst kaum existierte, dann anstieg, um nach den Beben wieder nahezu vollständig zu verschwinden.

Eine zweite erweiterte These zu diesen Vorgängen lautete wie folgt: Zunächst fand ein nicht näher bekanntes Ereignis vor 1997 statt, dass die tektonischen Platten in der Südisländischen Seismischen Zone derart verkoppelte, dass die Seismizität signifikant zurückging, also Energie in dieser Zone gespeichert wurde. Im Winter konnte diese Energie schlechter abgebaut werden als im Sommer. Ursächlich erscheint hierfür die Verfügbarkeit von Schmelzwasser in dieser Zone. Etwa ein Drittel der gesamten Gletscherkappen werden über dieses Gebiet entwässert. Mit steigender Temperatur ist mehr Schmelzwasser vorhanden, dessen zusätzlicher Druck in die Tiefe der Spalten propagiert und dort die Wahrscheinlichkeit für Erdbeben erhöht. Mit der steigenden Zyklizität der Prozesse verzögerte sich auch die Plattenbewegung, bis schließlich im Frühsommer 2000 zwei Erdbeben die Platten wieder entkoppelten. Da sich dieses Phänomen der zunehmenden Zyklizität nicht gleich nach den schweren Beben wiederholte, kann davon ausgegangen werden, dass die ursprüngliche Ursache aufgehört hat zu wirken.

1.3 Vorüberlegungen zum Lösungsansatz

Es wird darzustellen sein, weswegen Island ein geeigneter Ort ist, um die Fragen nach den zeitlichen Verläufen des Bewegungsablauf an der Plattengrenze zwischen Nordamerika und Eurasien zu klären. Hierzu soll in Kapitel 2 auch ein weiterführender Überblick über die Prozesse der Plattentektonik gegeben werden.

Um eine Modellierung der Bewegungen an der Plattengrenze in Island vornehmen zu können, müssen zuvor die Systembeziehungen überdacht werden. Welche Systeme wirken hier zusammen und welche Qualitäten besitzen sie? Diese vorbereitenden theoretischen Fragen werden in Kapitel 3 geklärt.

Die Kruste wirkt durch ihre Erdbeben und ihre Deformation auf die Seismometer und die GPS-Stationen. Durch den klimatischen Einfluss wirkt die Atmosphäre auf die Kruste und – natürlich erwünscht – auf alle Wetterstationen, aber eben auch signifikant auf das GPS-System (Abb. 1.4). Von Interesse im Rahmen dieser Untersuchung ist nur der Übertragungspfad von der Kruste auf GPS. Aber dieser Pfad ist durch den direkten Atmosphäreneinfluss gestört.

Daher ist es wichtig zu wissen, ob und wie die Atmosphäre auf die Kruste und ferner die Kruste auf die Seismometer wirkt. An ihnen sollte schließlich ein komplementäres Signal angekommen sein, das mit dem des GPS korrespondiert.
Um einen Datensatz zu analysieren, bedient man sich für gewöhnlich der Methoden der Zeitreihenanalyse (Schlittgen & Streitberg 1997).

Ebenso sind eine Reihe von adaptiven Filtern etabliert, die bereits Modelle darstellen (Haykin 2002). Diese Ansätze geben allerdings keine Systembeschreibung. Hierfür ist es notwendig, ein eigenes physikalisch-mathematisches Modell aufzustellen und mit den Methoden der *Operation research* bestmöglich an den Datensatz anzupassen (Rardin 1998).
Wenn aber, wie in der vorliegenden Fragestellung, die physikalisch-mathematischen Zusammenhänge nur unzureichend bekannt sind, kommen nur noch die nichtparametrischen Methoden, wie beispielsweise Neuronale Netze in Frage (Haykin 1999). Diese standen und stehen aufgrund ihres *black box*-Charakters in der Kritik. Diese Eigenschaft verhindert die Interpretation des Modells und sie verhindert auch eine geeignete Kontrolle des Modells. Die *Statistische Lerntheorie* stellt daher Beurteilungsmaßstäbe für die richtige Modellkapazität zur Verfügung, um den *black box*-Charakter zu beseitigen (Vapnik 1998). Diese Aspekte sollen in Kapitel 4 erörtert werden.

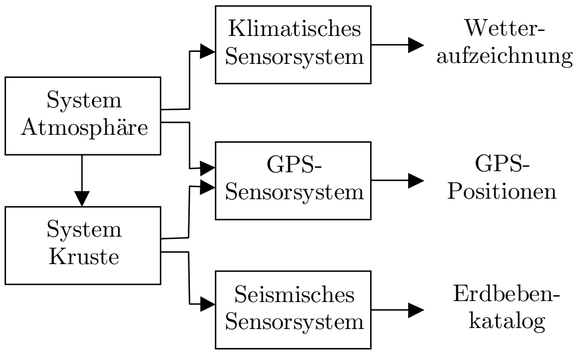


Abb. 1.4: Die Wirkungsweise der Subsysteme *Kruste* und *Atmosphäre* mit- und untereinander sowie auf die jeweiligen Sensoren.

Alle diese Ansätze setzen aber die Äquidistanz der Messwerte voraus. Es ist im Falle von Datenlücken häufig möglich, diese schadlos zu füllen (Lecolazet 1956; Melchior 1983; Heinert & Riedel 2007). Aber bei *per se* inäquidistanten Daten wie beispielsweise Erdbebendaten greifen diese Methoden nicht. Hier muss demnach eine wesentliche Weiterentwicklung vorgenommen werden: Ein nichtparametrisches Modell muss also eine *Zeitfunktion* enthalten, welche die Ankunft der remanenten Verformung nach einem Erdbeben an einer GPS-Station regelt. Im Weiteren muss das Modell eine *Raumfunktion* enthalten, die jene seismisch aktiven Regionen hervorhebt, die tektonisch wirken, gegenüber solchen, deren Seismizität auf vulkanische Aktivitäten zurückzuführen ist. Alle notwendigen algorithmischen Adaptionen und Modelle, die zur Lösung dieser Fragen dienen, werden in Kapitel 5 bereitgestellt.
Die praktischen Ergebnisse der im Rahmen dieser Arbeit entstandenen Adaptionen und Modelle finden sich in Kapitel 6. Der Bezug der Adaptionen, Modelle und Ergebnisse zum Stand der Forschung, wie er ausführlich in den vorangehenden Kapiteln erarbeitet worden ist, wird dort dargelegt. Mit einem Resümee und einem Ausblick auf die möglichen Weiterentwicklungen auf der Grundlage der vorgestellten Untersuchungen schließt diese Arbeit mit Kapitel 7 ab.

2 Geowissenschaftliche Grundlagen

„Um von der willkürlichen Wahl des Bezugsystems freizukommen, könnte man vielleicht ausgeglichene Kontinentverschiebungen definieren, die relativ zur Gesamtheit der Erdoberfläche an Stelle nur eines Teiles derselben zu bestimmen wären. Deren Bestimmung wäre aber praktisch mit großen Schwierigkeiten verknüpft und kommt einstweilen nicht in Betracht.“
Alfred Lothar Wegener

2.1 Theorie der allgemeinen Plattentektonik

2.1.1 Historische Entwicklung

Der flämische Kartograph Ortelius machte im Jahre 1596 die weitreichende Aussage, Amerika sei von Europa und Asien „durch Erdbeben und Flut fortgerissen“ (Anh. C.1). Die Konformität der Küstenlinien insbesondere von Afrika und *Peru* fiel auch Bacon (1620) auf. Er erwähnte sie als eine unter vielen anderen konformen Erscheinungen der Natur, wobei er aber gerade diese unter den vielen nicht als Zufall betrachten wollte[1]. Im Jahre 1858 ging der Katastrophist[2] Snider-Pellegrini einen Schritt weiter, als er die erste Karte veröffentlichte, auf der die Alte und die Neue Welt ohne trennenden Ozean zu sehen war (Abb. 2.1). Er mutmaßte, dass es die biblische Sintflut gewesen sei, welche die Kontinente voneinander getrennt habe.
Der österreichische Geologe Suess vertrat in seiner Buchreihe über das „Antlitz der Erde“ zwar weiterhin die 'Landbrücken-Theorie'[3], aber er konnte die Existenz eines früheren nördlichen und südlichen Großkontinentes belegen. Insbesondere die Verbreitung der sogenannten *Glossopteris-Flora* (Abb. 2.2) spielt hierbei eine herausragende Rolle. Der erstmals als *Gondwána* (Suess 1909, Bd. 2, S. 318) bezeichnete südliche Kontinent umfasst in seinem Werk zwar wechselnde Landgebiete, bei seiner letzten Beschreibung sind in diesem Großkontinent alle Kontinente der südlichen Hemisphäre einschließlich Indien zusammengefasst (Suess 1909, Bd. 3, 2. Hälfte, S. 574 u. S. 766ff). Auch für die nördliche Hemisphäre nimmt Suess eine Gruppenbildung vor, die sich aber hauptsächlich durch das Fehlen der Südflora auszeichnet. Wegweisend für die Existenz eines nördlichen Großkontinentes ist allerdings seine Beschreibung der Geologie der Küsten um den Nordatlantik[4]. Der so beschriebene nördliche Großkontinent wird heute als *Laurasia* bezeichnet.
Unabhängig von den Beobachtungen aller seiner Vorgänger fiel auch Wegener im Jahre 1910 die offensichtliche Kongruenz der Küstenlinien dies- und jenseits des Atlantiks auf[5].

[1] Nach der Deutung von Keary & Vine (1990) bezieht sich Bacon hier vielmehr auf die Ähnlichkeit der beiden Westküsten. Dem Autor ist allerdings keine kartographische Darstellung Südamerikas bis zur Entstehung des *Novi Organi* bekannt, welche Bacon diesen Schluss hätte nahelegen können.

[2] Die *Katastrophisten* vertraten im Gegensatz zu den *Aktualisten* die Ansicht, dass sich geologische und biologische Entwicklungen auf der Erde sprunghaft im Zusammenhang mit Naturkatastrophen ereignet haben.

[3] Die sogenannte 'Landbrücken-Theorie' besagt, dass Kontinente, wie z.B. *Atlantis* oder *Archhelenis*, im Atlantischen Ozean versunken sind. Diese versunkenen Kontinente sollten zu unterschiedlichen Zeiten Landbrücken insbesondere über den Atlantik gebildet haben. Dies erschien als einzige mögliche Erklärung für die Übereinstimmungen in der fossilen Flora und Fauna dies- und jenseits des Atlantiks.

[4] „Die vielen angeführten Einzelheiten lassen erkennen, dass zu beiden Seiten des Oceans eine Art von unvollständiger Symmetrie besteht. In einzelnen Fällen versagt der Vergleich; in anderen ist die Uebereinstimmung auffallend, wenn auch schwer zu erklären, [...]“, Zitat aus (Suess 1909, Bd. 2, S. 164ff)

[5] Dieses beteuert er glaubhaft in der vierten Auflage seines Buches (Wegener 1929, S. 1).

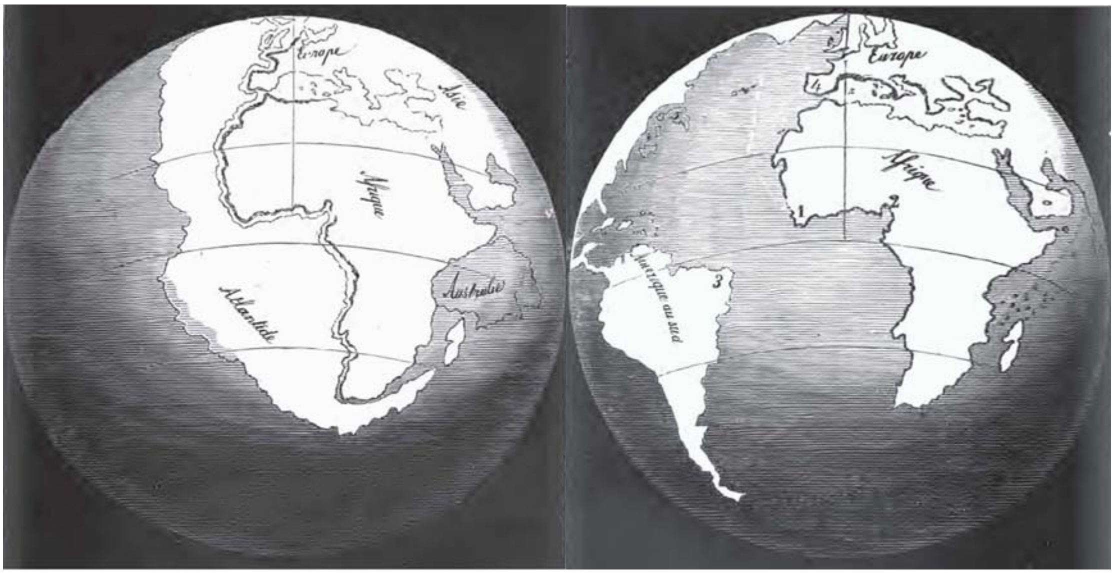


Abb. 2.1: Die erste kartographische Darstellung von Pangaea unter der Annahme einer Kontinentalverschiebung (Snider-Pellegrini 1858, Tafel 9 u. 10, zw. S. 314 u. 315).

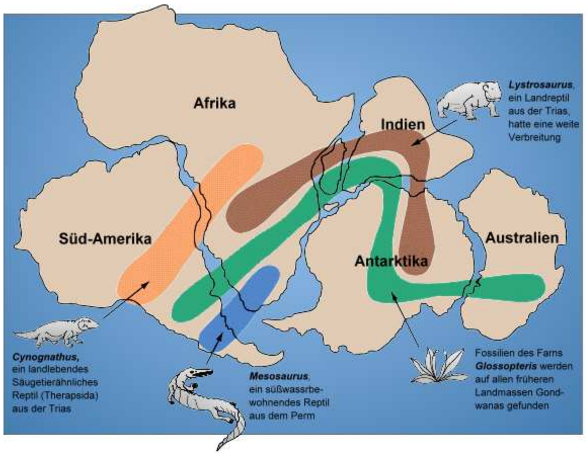


Abb. 2.2: Verteilung von Flora und Fauna auf Gondwána. Quellen: USGS 1999.

Die Möglichkeit einer Kontinentalverschiebung kam ihm aber erst 1911 in den Sinn. Seine erste Publikation im Jahre 1912 rief heftige Kritik hervor. Eine erste geschlossene Darstellung gelingt ihm mit seinem Buch im Jahre 1915[6].

Es waren die Diskrepanzen aus den Erkenntnissen der Paläoklimatologie, welche in der vierten Auflage seines Buches die bestechendsten Hinweise auf die Richtigkeit seiner Annahmen lieferten. Es erschien auffällig, dass sich sowohl nördlich (Indien) als auch südlich (Südafrika, Südamerika, Australien) des Äquators eiszeitliche Spuren in Form von glazialen Blocklehmen in geologischen Schichten fanden, welche während des Übergangs vom Perm[7] in das Karbon[8] vor etwa 285 Mill. Jahren gebildet sein mussten. Viele dieser Spuren des permokarbonischen Eiszeitalters wurden zuvor fälschlicherweise als pseudoglazial eingestuft, also der Gebirgsvergletscherung zugeordnet (Wegener 1929, Kap. 7).

Andererseits liegen die Kohlevorkommen des Karbon in hohen Breiten (Nordamerika, Westeuropa, Kleinasien, China). Derartige Lagerstätten bilden sich entweder in den Tropen, Subtropen oder den gemäßigten Breiten aus, weil hier die klimatischen Bedingungen die nötige

---

[6] „Betrachten wir z. B. die Küstenlinien Südamerikas und Afrikas, wo jeder Vorsprung auf der einen Seite in eine entsprechende Ausbuchtung der anderen Seite hineinpaßt; durch den Einbruch des Zwischenlandes (von 5000 km Breite) kann keine derartige Kongruenz entstehen. Schon bei flüchtiger Betrachtung der Karte erkennt man auch, wie sich hüben und drüben Gebirge (Grönland – Skandinavien), Bruchzonen (Mittelamerika – Mittelmeer) und Tafelländer (Südamerika – Afrika) entsprechen." Zitat aus (Wegener 1915, S. 73 (- 59 -))

[7] Geologisches System der Ära des Paläozoikums, entstanden vor 295-251 Mill. Jahren, Zeitalter der Verbreitung von Reptilien, Höhepunkt der paläozoischen Geokratie.

[8] Geologisches System der Ära des Paläozoikums, entstanden vor 358-296 Mill. Jahren, Zeitalter der Verbreitung der ersten Amphibien, Ausgehende paläozoische Thalattokratie.

Verfügbarkeit von Wasser und das Wachstum der zunächst vertorften und später kohlebildenden Flora zulassen. Jedoch belegen die fossilen Einschlüsse insbesondere der Baumfarne, dass die genannten Vorkommen den Tropen und Subtropen zuzurechnen sind. Unter den klimatischen Bedingungen der heutigen Fundorte hätte diese Flora nicht existieren können. Damit nicht genug, werden die glazialen und kohlebildenden Formationen durch Salzlagerstätten getrennt, welche einem Wüstenklimat zuzurechnen sind.
Die Möglichkeit einer Polwanderung galt zu diesem Zeitpunkt als weithin anerkannte These. Die Erkenntnis über die Kombination dieser drei Lagerstätten war aber nur durch eine Polwanderung allein nicht mehr zu erklären. Es ließ sich keine Pol- bzw. Äquatorlage finden, welche jeder Lagerstätte die richtige Klimazone hätte zuweisen können. Wegener untermauerte seine „Verschiebungstheorie“ mit den gesammelten Erkenntnissen anderer Fachgebiete, wie in der Biologie, Paläontologie, Geologie, Geophysik und Geodäsie. Allerdings konnten die Argumente der Geodäsie aufgrund ihrer damals unzureichenden Genauigkeit nicht überzeugen (Wegener 1929, Kap. 3).

2.1.2 Plattengrenzen

Es stellte sich die Frage, wo die Trennung der Lithosphärenplatten lokalisiert werden könne. Aus vereinzelten Echoprofilen war die Existenz einer „mittelatlantische Bodenschwelle“ durchaus bekannt. Allerdings wurde ihre Existenz zunächst falsch interpretiert: es handele sich hierbei um Basaltdecken der unteren kontinentalen Kruste, welche bei der Trennung der Kontinente zwischen dem Tiefseeboden und der oberen granitischen kontinentalen Kruste hervorgequollen seien (Wegener 1929, S. 214f. u. Abb. 59). Schon 1938 lokalisierten Geodäten einen Abschnitt der Plattengrenze in Nordostisland (Niemczyk 1943).
Unklar blieb lange der „Motor“ der Kontinentaldrift. Eine Summe von Kräften scheint die Platten in Bewegung zu halten: aufsteigendes Mantelmaterial hebt die über ihm liegende junge ozeanische Kruste an. Diese wiederum gleitet von dem aufsteigenden Rücken ab. Die Konvektionsströmung begünstigt die Bewegung der auf ihr befindlichen Platte vom Rücken weg (Wilson 1966). Schließlich taucht die alte, erkaltete und damit spezifisch dichte ozeanische Kruste wieder in den oberen Mantel ein und zieht dabei jüngere akkretierte Sedimente mit sich in die Subduktion[9] (Stüwe 2000, Kap. 5.3.2).

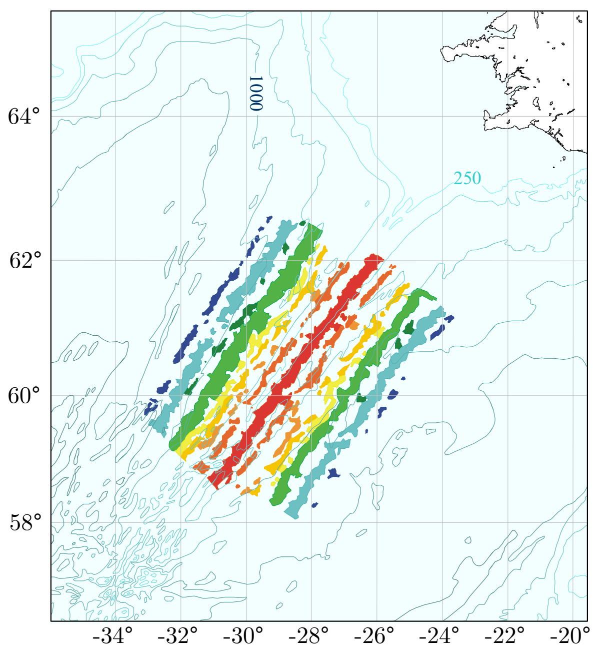


Abb. 2.3: Die Magnetisierung der Gesteine des Reykjanes-Rückens (Vine 1968, S. 82): Die Farben der Streifen mit einer dem heutigen Erdmagnetfeld entsprechenden Magnetisierung geben die symmetrischen Gesteinsformationen gleichen Alters an (Press & Siever 2003, Abb. 20.9).

[9] Trotz der Vielzahl der geophysikalischen Erkenntnisse über die Prozesse an den Plattenrändern und auch der redundanten Daten aus der Geodäsie wird noch heutzutage alternativ die Theorie der Expansion der Erde für die Form und Verteilung der kontinentalen Kruste angeführt (Maxlow 2001) – allerdings ohne den immensen Massenzuwachs ansatzweise zu erklären.

Divergente Plattengrenzen

Durch die Entdeckung des symmetrischen magnetischen Streifenmusters (Abb. 2.3) entlang des Mittelatlantischen Rückens bekam Wegeners Theorie in den 1960'er Jahren neuen Auftrieb, nachdem die mittelozeanischen Rücken als *divergente*[10] Plattengrenzen erkannt worden sind, an denen ständig neue ozeanische Kruste gebildet und an die bereits bestehende angelagert wird. Jene sind das fehlende Glied zu den konvergenten Plattengrenzen.

Konvergente Plattengrenzen

Der sogenannte „Feuergürtel“ um den Pazifik war als aktives Beben- und Vulkangebiet (Gutenberg & Richter 1949, S. 30ff) und auch das lokale Abgleiten ozeanischer Kruste unter eine andere Platte war bekannt (Benioff 1954). Aber erst in der Zusammenschau mit den Prozessen an den mittelozeanischen Rücken konnten die *konvergenten*[11] Plattengrenzen als Teil der Konvektionsbewegung der Erdoberfläche erkannt werden (Hess 1962). Die hier stattfindende Subduktion der jeweils spezifisch dichteren unter die jeweils spezifisch dünnere Platte hat die schwersten Erdbeben und, in der Wadati-Benioff-Zone oberhalb der bereits subduzierten Platte, äußerst explosiven Vulkanismus zur Folge. Typisch für eine Subduktion ist die Ausbildung zweier Spannungszonen (Stüwe 2000, Abb. 6.24). Bis heute ist allerdings noch der Beginn einer Subduktion nicht völlig verstanden. In erster Linie wird die Sedimentauflast auf die ozeanische Kruste an der Grenze zur kontinentalen Kruste verantwortlich gemacht (Branlund et al. 2000). Ist der Prozess erst in Gang gekommen, taucht die kalte und damit spezifisch dichtere ozeanische Kruste wieder in die Asthenosphäre ein[12].

Transformzonen

Gleiten zwei Platten an ihren Rändern ohne Subduktion oder Krustenneubildung aneinander vorbei, so spricht man von einer *Transformzone*. Diese Zonen können sowohl divergenten als auch konvergenten Plattengrenzen unmittelbar benachbart sein und wirken als Verbindungsstücke. Sie sind entweder durch eine zentrale Spalte wie beispielsweise die St.-Andreas-Spalte oder durch rotierende Blöcke[13] wie in der Südisländischen Seismischen Zone gebildet. Kommt diese Transformbewegung aufgrund einer temporär erhöhten Reibung vollständig zum Stillstand, wird das die Ursache eines baldigen schweren Erdbebens in dieser Zone sein. Der Mechanismus eines einzelnen seismischen Ereignisses wird als *Blattverschiebung* bezeichnet. Im Nahbereich zu einer divergenten Plattengrenze, deren Divergenzbewegung ansonsten weitgehend gleichmäßig erfolgen würde, bestimmen die benachbarten Transformzonen wesentlich den Ablauf der Bewegung. Die Transformbewegungen werden aufgrund ihres gedachten Drehsinns in *sinistral* und *dextral*, also linksdrehend und rechtsdrehend unterschieden.

Nichttransformzonen

Ein besonderes tektonisches Element ist eine Zone zwischen einem Versatz von Plattengrenzen an denen keine ausgleichende geradlinige Transformbewegung stattfindet, sogenannte *non-transform offsets*. Sie entstehen unter der Ausbildung von Trögen an kurzen Versätzen von bis zu 30 km. Besonders unter zwei Bedingungen bilden sich diese Zonen aus: extrem langsame Divergenz (Dick et al. 2003) oder auch als Diskontinuitäten der schiefen Divergenz[14].

[10] Aufgrund des vergleichsweise ruhigen Vulkanismus und der weniger starken Seismizität spricht man auch von *passiven* Plattengrenzen.

[11] Im Gegensatz zu den passiven Plattengrenzen werden diese Zonen *aktive* Plattengrenzen genannt.

[12] Dieser Prozess wird als *slab pull* bezeichnet. Neben dem Prozess des Abgleitens der jungen ozeanischen Kruste von der hochgewölbten Asthenosphäre unter einem mittelozeanischen Rücken (*rigde-push*) gilt jener Prozess als eine treibende Kraft für die Plattentektonik.

[13] Bei dieser speziellen Form spricht man aufgrund des Vergleichs mit kippenden Büchern im Regal von *bookshelf faulting*.

[14] Engl.: *oblique spreading*.

2.1.3 *Hot spots* und Manteldiapire

Abseits der Plattengrenzen existieren vulkanische Zentren, deren Ursprung nicht in tektonischen Prozessen zu suchen ist. WILSON beschrieb 1963 erstmals wie ein *hot spot*[15] aufgrund der Bewegung der ozeanischen Platte die Inselketten im Pazifik und im Speziellen die von Hawaii den bildete. Solche als *hot spots* bezeichnete Regionen befinden sich häufig oberhalb von Manteldiapiren[16] (PRESS & SIEVER 2003, S. 89f).
Über die Entstehungsursache und insbesondere über die Entstehungstiefe des Island-*hot spot* bestehen noch immer verschiedene Deutungen: einerseits wird diese Anomalie schon seit Jahren erfolgreich als Manteldiapir, also als eine möglicherweise zyklisch aufsteigende Blase partieller Schmelze von der Kern-Mantelgrenze modelliert (RUEDAS ET AL. 2004; KREUTZMANN ET AL. 2004; MARQUART & SCHMELING 2004).
Andererseits wird diese Anomalie als lokaler Aufschmelzprozess infolge von Druckschwankungen im oberen Mantel der Erde beschrieben und damit ihr Charakter als Manteldiapir massiv in Frage gestellt (FOULGER ET AL. 2005; FOULGER & ANDERSON 2005). Die Hauptbegründung liegt in dem fehlenden oder überaus schwachen Signal im Geschwindigkeitsfeld der seismischen Wellen (FOULGER ET AL. 2000; FOULGER & PEARSON 2001; FOULGER ET AL. 2001). Ein solches Signal wird aber an der Grenze des Diapirs auch unterhalb der Mohorovičić-Diskontinuität[17] erwartet.
Jedoch ist das seismische Geschwindigkeitsfeld durchaus nicht einseitig für eine asthenosphärische Anomalie interpretierbar (WOLFE ET AL. 1997; WOLFE ET AL. 2002). Eine der neuesten Untersuchungen zur Auflösung eines *plume tails*, also des Stammes, weist nach, dass das Fehlen des Signals in der seismischen Tomografie kaum geeignet ist, die Entstehungstiefe an der Kern-Mantelgrenze zu bezweifeln (ISMAIL-ZADEH ET AL. 2006).
Damit ist diese Diskussion um den *hot spot* noch nicht erschöpfend abgeschlossen (NOLET ET AL. 2007).

2.1.4 Der Superkontinent-Zyklus

Der Begriff des *Superkontinents* ist in der vorhandenen Literatur durchaus vage. So existieren Definitionen, ein Superkontinent umfasse mehrere *Kratone*[18] und *Terrane*[19]. In der Superkontinentdiskussion tauchen daher auch verhältnismäßig kleine Agglomerationen von Landmassen auf (SANKARAN 2003). Die heutige Landmasse aus Afrika, Indien und Eurasien würde dabei vermutlich „Eufrasindia“ geheißen und als Superkontinent kategorisiert. Es müssten daher die sperrigen Definitionen von einem *Paläokontinent*, *Paläogroßkontinent*, *Paläosuperkontinent* und *Protosuperkontinent* getroffen werden. In dieser Arbeit soll der Begriff des Superkontinents ausschließlich für den Zusammenschluss nahezu aller heutigen Kratone und Terrane verwendet werden.
WILSON beschrieb 1966 den nach ihm benannten Zyklus mit Dauer von etwa 550 Mill. Jahren, in der sich alle kontinentalen Landmassen zu einem Superkontinent vereinigen. Existiert ein Superkontinent, so wirkt diese Landmasse als eine große Isolierschicht unter der sich die Wärme der Erde staut. Die Ausdehnung des oberen Mantels führt zu Bruchzonen in der Landmasse in der sich vermehrt Vulkane ausbilden, welche die Wärme und das geschmolzene

[15] Engl.: *hot spot* 'heißer Fleck', die deutsche Übersetzung gibt nur unzureichend den so im Englischen definierten Prozess wider. Daher wird im Folgenden der englische Terminus bevorzugt.

[16] Altgr.: $\delta\iota\alpha\pi\varepsilon\acute{\iota}\rho\varepsilon\iota\nu$ 'durchbohren', häufiger gebraucht: Engl.: *mantle plume*.

[17] Grenze zwischen der Kruste und dem oberen Mantel.

[18] Auch *Kontinentalschild* von temporär unabhängigen Kontinenten bis ein Superkontinent sei prinzipiell der Zusammenschluss aller Kratone: Teil eines Kontinents, der seit dem Präkambrium oder älterem Paläozoikum keiner stärkeren Deformation unterlag (PRESS & SIEVER 2003, S. 691).

[19] Auch *Mikroplatte*: Bestandteil eines Grundgebirges, der sich scharf von seiner Umgebung abgrenzen lässt und als ursprünglicher unabhängiger Kontinentalabbruch, Tiefseeberg oder Inselbogen an einer konvergenten Plattengrenze an einen größeren Kontinent angeschweißt wurde (PRESS & SIEVER 2003, S. 693).

Material abführen. Am äußeren Rand des aufbrechenden Superkontinentes ist die alte ozeanische Kruste bereits tief eingesunken. Zum einen liegt ihre spezifische Dichte sowohl oberhalb der des Kontinents als auch über der des äußeren Mantels, zum anderen ist sie beladen von der Fracht der aufliegenden Sedimente. Die Übergangszone an der Grenze zur kontinentalen Kruste hat sich in bereits erheblichem Maße plastisch verformt (Branlund et al. 2000, Abb. 2 u. 3). Der äußere Rand des Superkontinents beginnt sich auf die ozeanische Kruste hinzu zu bewegen. Die Kruste bricht und die plastisch unverformten Bereiche der ozeanischen Kruste haben die Form eines Keils, der sich nun zwischen die kontinentale Kruste und die Asthenosphäre schieben kann. Der alte schwere Rand der ozeanischen Kruste taucht schnell in die spezifisch dünnere Asthenosphäre ein und beginnt die ganze Platte mit sich zu ziehen. Infolge dessen kommt eine Konvektionszelle in Gang, die den Superkontinent an den Dehnungsbrüchen auseinanderreißt. Mitten im Superkontinent entsteht aus der Schwächezone des Grabenbruchs ein neuer Ozean (Press & Siever 2003, S. 561ff).

Pangäa

Zu Beginn der Trias vor 251 Mill. Jahren, als die ersten Dinosaurier das Land erobern, bilden alle bekannten Kontinente den Superkontinent *Pangäa*[20] (Abb. 2.4). Er ist aus der Vereinigung der Großkontinente Gondwána, Laurentia, Baltica und Sibiria hervorgegangen. Dieser Prozess ist bereits zu Beginn des Perm vor etwa 300-290 Mill. Jahren vollendet. Pangäa wird etwa bis in den Jura vor 200-180 Mill. Jahren existieren, wenn sich der Nordkontinent *Laurasia* und der Südkontinent *Gondwána* wieder teilen. Der vereinten Landmasse steht die geschlossene Wasserfläche des *Panthalassa*[21], des Vorläufers des heutigen Stillen Ozeans, gegenüber. Während der Existenz von Pangäa schließen Laurasia und Gondwána eine große Meeresbucht, die *Thetys*[22] ein. Zum Ende des Perm lösen sich große Kontinentabbrüche von Gondwána und driften nordwärts bis sie mit Laurasia kollidieren. So heißt die Bucht nördlich der Abbrüche *Paläothetys*, südlich der Abbrüche *Neothetys*.

Rodinia

Zum Ende des Mesoproterozoikums[23] vereinigen sich die Bruchstücke der heutigen Kontinente im Superkontinent *Rodinia*[24] (Dalziel 1991). In dieser Zeit bestehen bereits die Großkontinente Westgondwána mit Ostantarktika, Indien und Australien sowie Ostgondwána mit Kalahari, Kongo, Westafrika, Rio Plata und Amazonia (Abb. 2.5). Während des mittleren Neoproterozoikum[25] sind sie noch durch das mit Baltica vereinigte Laurentia voneinander getrennt (Dalziel 1995; Dalziel 1997). Bis zum Ende des Neoproterozoikum wird Ostgondwána eine Rotation um Laurentia vollführen und mit Westgondwána verschmelzen (Press & Siever 2003, Abb. 20.23a u. b). Beide werden den Südkontinent Gondwána bilden.
Die Diskussion um die genaue Form Rodinias ist noch nicht beendet (Torsvik 2003). Sogar ob die Landmassen zu einem Zeitpunkt zusammengehangen haben oder vielmehr auf engem Raum nur verschiedene kurzzeitige Verbindungen eingegangen sind, wird wieder diskutiert.

[20] Altgr.: *παν-γαῖα*, 'All-Erde'

[21] Altgr.: *παν-ϑᾰ́λασσα*, 'All-Ozean'

[22] Altgr.: Θέτῐς: Titanin und Meeresgöttin der griechischen Mythologie.

[23] Ära des Präkambrium mit den Geologischen Systemen Stenium, Ectasium und Calymmium, vor 1600-1000 Mill. Jahren, Zeitalter der Entstehung der Eukaryoten (Zellen mit Zellkern).

[24] Russ.: *родина*: 'Mutterland'

[25] Ära des Präkambrium mit den Geologischen Systemen Ediacarium, Cryogenium und Tonium, vor 1000-542 Mill. Jahren, Zeitalter der Entstehung des vielzelligen Lebens, lange Phase der globalen Vereisung.

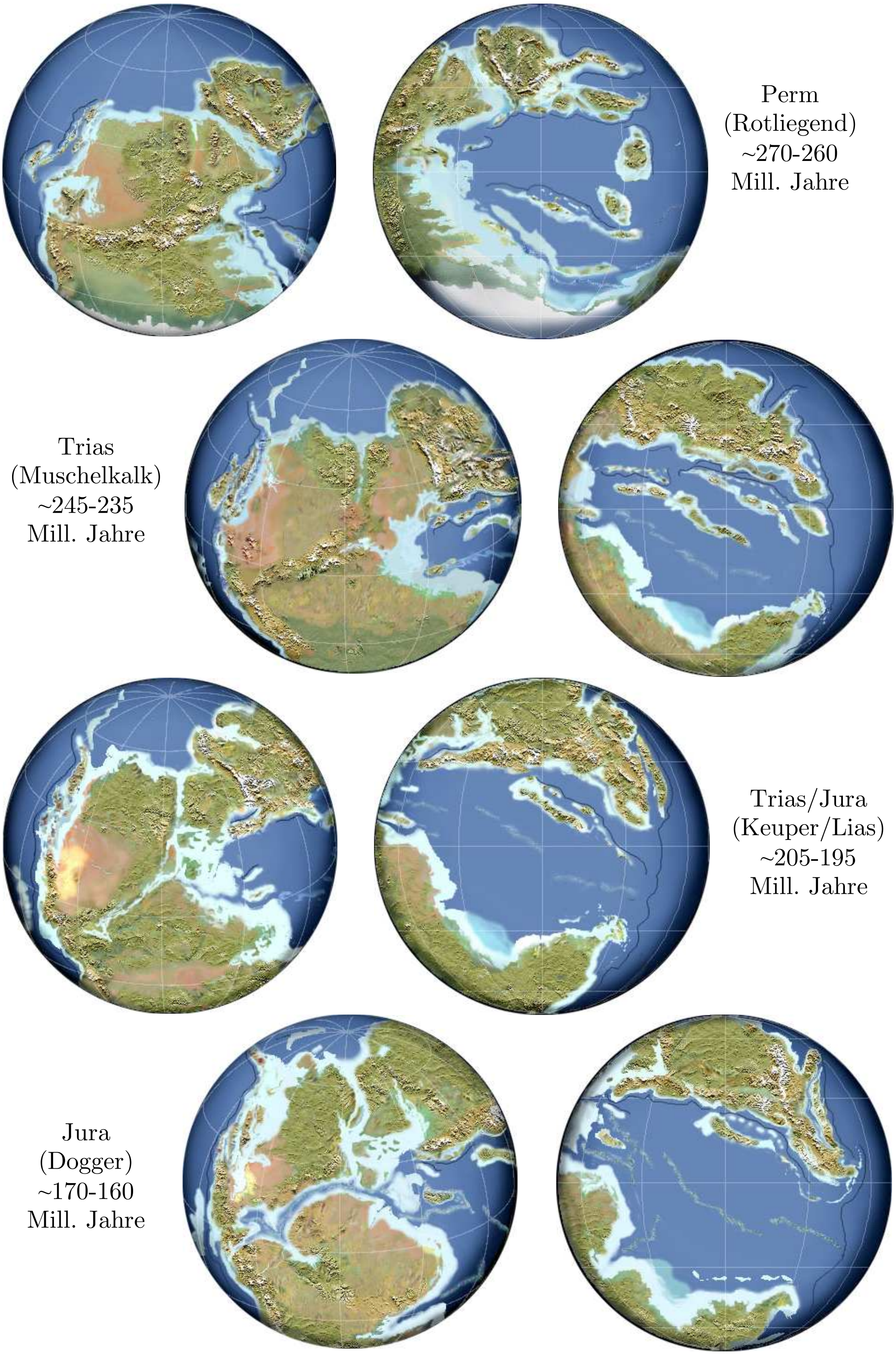


Abb. 2.4: Pangäa zur Zeit von der permokarbonischen Vereisung im unteren Perm bis zur Teilung von Laurasia und Gondwána im mittleren Jura (Blakey 2007).

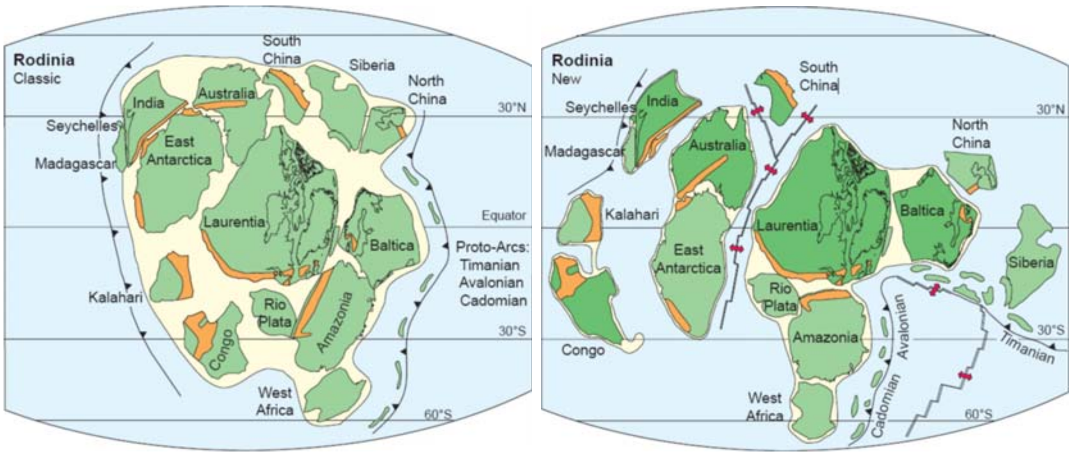


Abb. 2.5: Der Superkontinent Rodinia (die Orogene in orange): links findet sich die klassische Anordnung der Landmassen (Dalziel 1997) und rechts eine Anordnung der Landmassen nach neuerer Interpretation der Daten (Torsvik 2003).

Weitere Groß- und Superkontinente

Im Mesoproterozoikum bildeten die drei Großkontinente *Ur*, *Atlantika* und *Nena* den hypothetischen Superkontinent *Columbia*[26] (Rogers & Santosh 2004). Von diesen Teilen scheint Ur der überhaupt älteste größere Zusammenschluss von kontinentalen Massen zu sein (Sankaran 2003). In ihm finden sich vor etwa 3 Mrd. Jahren die Kratone *Kaapvaal* (Südafrika) und *Madagaskar* im Zusammenschluss mit *Dhawar*, *Bhandara* und *Singhbhum* (Indien) sowie *Pilbara* (Westaustralien). Vor etwa 2,5 Mrd. Jahren sollen die Kratone *Zimbabwe* (Südafrika) und *Yilgarn* (Westaustralien) akkretiert worden sein (Sankaran 2003, S. 1121). Zur hypothetischen Existenz von Ur gibt es noch Vorläufer: So wurde zunächst der Kontinent *Vaalbara*, ein Zusammenschluss von den Kratonen *Kaapvaal* und *Pilbara* vermutet und erweitert wurde die Diskussion, indem auch der Kraton Zimbabwe dieser Akkretion zugerechnet wurde. Damit ergab sich der hypothetische Kontinent *Zimvaalbara*. Erst später wurden auch die Kratone Indiens einbezogen (Bandopadhyay & Sengupta 2005), unter anderem auch deshalb, weil eine hypothetische Verbindung nur über größere Bindeglieder wahrscheinlich erscheint (Evans et al. 2000). Andere Untersuchungen mit umfangreichen Daten zur Paläomagnetik, Isotopen-Altersbestimmungen und Formationsvergleichen scheinen alle frühen Rekonstruktionen wie Vaalbara, Zimvaalbara oder Ur zunächst wieder in Frage zu stellen (Pesonen et al. 2003, Abb. 1a). Aber selbst diese sehr umfangreiche Untersuchung muss aber aufgrund der Schwächen der Rekonstruktionstechniken den ultimativen

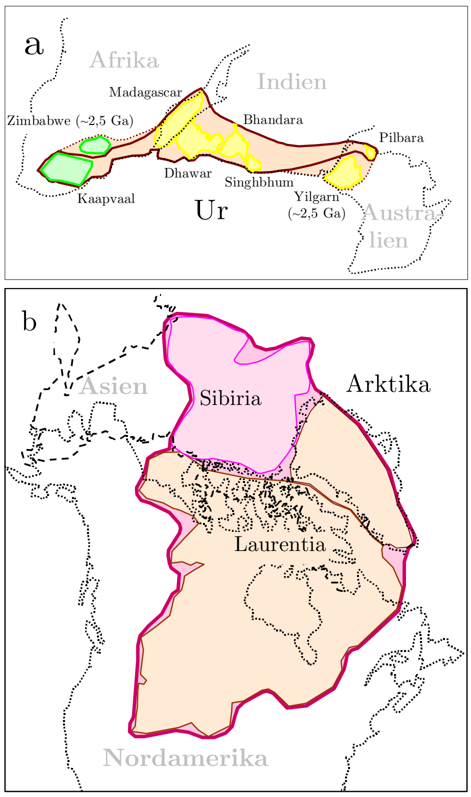


Abb. 2.6: Die mutmaßlichen Paläogroßkontinente *Ur* (a) und *Arktika* (b) (Rogers & Santosh 2004, vergl. Abb. 6.5 u. 6.9).

[26] Dieser Superkontinent ist auch als *Hudsonland* bekannt (Pesonen et al. 2003).

Gegenbeweis zur Existenz Urs schuldig bleiben (Pesonen et al. 2003, Kap. 5, Abb. 3 u. 6). Atlantika umfasst die Kratone West- und Zentralafrikas sowie die des östlichen Südamerika. Nena besteht einerseits aus Baltika und wird anderseits aus den Kratonen Sibiriens und Nordamerikas gebildet, die sich bereits zuvor zu Arktika zusammengeschlossen hatten.

2.1.5 Die Entstehung des Nordatlantik

Die Kontinente Baltica und Laurentia, beides Bruchstücke des Superkontinents Rodinia, vereinigen sich wieder im späten Devon. Damit schließt sich der östliche Teil des *Iapetus*[27] und es entsteht durch die Kollision ein Gebirge dessen Spuren von Spitzbergen bis Florida noch immer präsent sind (Wilson 1966). Diese Verbindung währt etwa 150 Mill. Jahre. Während sich allerdings alle Kontinente zum Superkontinent Pangäa vereinigen, entstehen zu beiden Seiten Grönlands wieder Grabenbrüche. Nach der Teilung Pangäas in den Nordkontinent Laurasia und den Südkontinent Gondwána im Jura nimmt die Öffnung des Atlantischen Ozeans im Norden einen anderen Verlauf als im Süden. Das Öffnen des mittleren und südlichen Atlantiks erfolgt bis auf den heutigen Tag am selben mittelozeanischen Rücken wie zu Beginn (Nürnberg & Müller 1991)[28]. Im Norden hingegen existiert nicht ein einziger dauerhafter mittelozeanischer Rücken, sondern die Hauptachse der ozeanischen Krustenbildung befindet sich in der Kreidezeit und zu Beginn des Tertiär noch westlich und anschließend östlich von Grönland (Braun & Marquart 2001). Verantwortlich hierfür ist die relative Position einer thermischen Anomalie, die als der Island-*plume*[29] bezeichnet wird, unter dem grönländischen Kontinentalschild (Torsvik et al. 2001).
Vom Beginn des Jura vor etwa 190 Mill. Jahren bis zum Beginn des Tertiär vor etwa 57 Mill. Jahren entsteht zunächst zwischen Grönland und Skandinavien ein einige hundert Kilometer breiter Grabenbruch. Seine Schollen bilden in großen Bereichen ein Flachmeer, das den Vorläufer des heutigen Nordatlantik bilden soll (Braun & Marquart 2001). Verschiedene untermeerische Plateaus erscheinen wie ein verstreutes Puzzle des Meeresbodens, der sich am Grunde eines Flachmeeres aus den Schollen des früheren ausgedehnten Grabenbruchs gebildet hat.
In der frühen Kreidezeit hat sich auch ein Grabenbruch zwischen der Baffininsel und dem grönländischen Kontinentalschild ausgebildet. Zum Ende der Kreidezeit erreicht der Manteldiapir diesen Grabenbruch und eröffnet ihn zu einer vulkanischen Riftzone. Im frühen Tertiär vor etwa 60 Mill. Jahren, als sich der Manteldiapir noch unter Grönland befindet, beginnt er bereits Material zum Grabenbruch östlich von Grönland zu fördern (Torsvik et al. 2001, Abb. 5). Die Flutbasalte aus dieser Zeit finden sich im Grönland-Islandrücken und im äußeren Vøringplateau wieder (Vink 1984; Vink et al. 1995; Sigmundsson 2006). Etwa drei Mill. Jahre später bilden sich der Reykjanes- und der Ægirrücken aus. Etwa gleichzeitig kommt die Spreizung in der nun entstandenen Baffinbay zum Erliegen.
Seit sich der Mittelatlantische Rücken östlich von Grönland befindet, ist seine relative Lage aber noch immer instabil. Am deutlichsten wird dieser Umstand durch die Existenz sowohl des östlichen und verödeten Ægirrückens (Abb. 2.7, ÆR) als auch des westlichen und aktiven Kolbeinseyrückens (Abb. 2.7, KR).

---

[27] Vorläufer des Atlantik benannt nach Altgr.: Ἰαπετός: Titan aus der griechischen Mythologie, Vater des Atlas und des Prometheus.

[28] Dieser Umstand ist auch für die frühe Erkenntnis der Plattentektonik von entscheidender Bedeutung, denn andernfalls hätte die Kongruenz der Küsten Afrikas und Südamerikas nicht immer wieder zu Überlegungen in diese Richtung angeregt (Kap. 2.1.1).

[29] In dieser Arbeit soll die Theorie bevorzugt werden, dass es sich bei der thermischen Anomalie nicht um einen *hot spot* im oberen Mantel aufgrund von lokaler Druckentlastung, sondern um einen Manteldiapir handelt (Kap. 2.1.3).

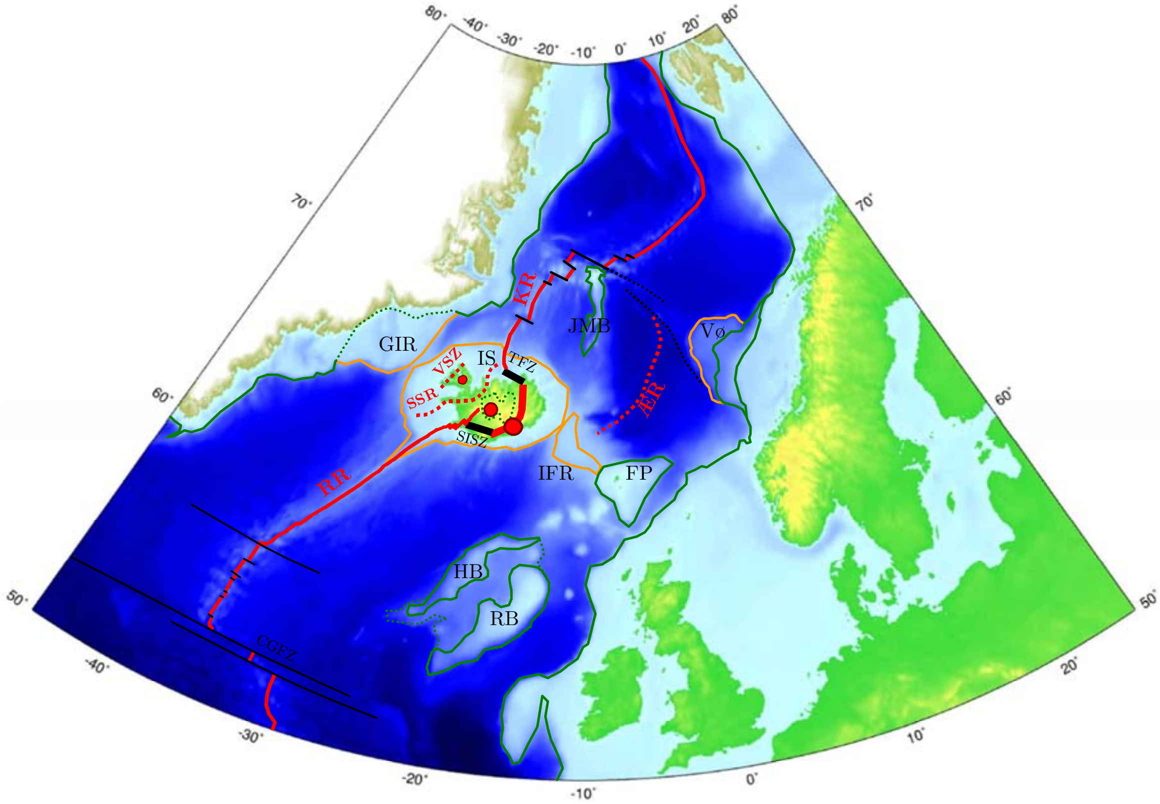


Abb. 2.7: Der Mittelatlantische Rücken in seinen Abschnitten: dem Reykjanesrücken (RR) von der Charly-Gibbs-Bruchzone (CGFZ) im Süden bis zur Südisländischen Seismischen Zone (SISZ), dem Kolbeinseyrücken (KR), und den verödeten Teilen des Ægirrückens (ÆR), der Vestfirðir-Störzone (VSZ) und des Snæfellsnes-Skagi-Rückens (SSR). In der Umgebung des mittelozeanischen Rückensystems befinden sich verschiedene Plateaus und Bänke, deren Entstehung mit der des Rückensystems eng verbunden ist: in der zentralen Achse der transverse Grönland-Island-Rücken (GIR), das Islandschelf (IS), der transverse Island-Färöer-Rücken (IFR) und das Färöerplateau (FP); im Norden der Jan-Mayen-Block (JMB) und das innere und äußere Vøringplateau Vø); im Süden die Hatton- und die Rockallbank (HB u. RB.).

Die kontinentalen Plateaus

Es scheint gesichert, dass es sich bei den Kernbereichen der Hatton- und der Rockallbank[30] (Abb. 2.7, HB und RB) um Teile des Eurasischen Kontinentalschelfes handelt (Bott 1974, S. 39). Die flacheren Randbereiche der beiden Bänke gehen vermutlich auf geneigte Schollen aus dem Grabenbruch zurück, die mit Schichten tertiärer Flutbasalte überlagert worden sind (Sigmundsson 2006, Abb. 2.4). Auch das Färöerplateau (Abb. 2.7, FP) scheint ein kontinentaler Abbruch zu sein (Braun & Marquart 2001; Heller & Marquart 2002). Mächtige Lagen von Flutbasalten aus der Anfangsphase der Spreizung haben die kontinentale Kruste dieser Plateaus teilweise überlagert.
Als ein weiterer kontinentaler Abbruch zeigt sich der Jan-Mayen-Block (Müller et al. 2002). Anders als die übrigen kontinentalen Abbrüche ist er vollständig von ozeanischer Kruste umgeben[31]. Zunächst verblieb der Jan-Mayen-Block bei Grönland. Vermutlich als das Aktivitätszentrum des Manteldiapirs den östlichen Rand der grönländischen Kruste erreichte,

[30] Die Rockallbank enthält Gesteine aus dem Proterozoikum. Damit gehört dieses Plateau den ältesten Teilen kontinentaler Kruste an (Pesonen et al. 2003, S. 290f).

[31] Diese Formation wird daher auch als *Jan-Mayen-Mikrokontinent* bezeichnet (Foulger et al. 2005; Foulger 2006).

entstanden die Flutbasalte des Grönland-Island-Rückens und es brach hier eine neue Spreizungszone westlich des alten Ægirrückens auf. Etwa zwischen 26 und 44 Mill. Jahren war der neu entstandene Kolbeinseyrücken südwestlich des Jan-Mayen-Blocks aktiv und der Ægirrücken nordöstlich (Foulger et al. 2005, Abb. 5). In dieser Phase verödete der Ægirrücken von Süden nach Norden zunehmend. In Folge davon rotierte der Jan-Mayen-Block in die heutige Lage.

Die vulkanischen Plateaus

Das Vøringplateau ist ein Konglomerat kontinentaler und ozeanischer Kruste. Im Aufbruch des frühen Grabens zwischen Grönland und Skandinavien haben sich hier große Bruchstücke kontinentaler Kruste weit verteilt (Heller & Marquart 2002). Später wurde es durch die Flutbasalte überdeckt, die den ganzen Nordatlantik durchziehen (Sigmundsson 2006, Abb. 2.4). Auch eine frühe und weit entfernte singuläre Materialzufuhr durch den Manteldiapir Islands ist diskutiert worden (Vink 1984).
Der Island-Färöerrücken, das Islandschelf und der Grönland-Islandrücken (Abb. 2.7, IFP, IS und GIR) scheinen überwiegend vulkanischen Ursprungs zu sein. Der Island-Färöerrücken am Ende des verödeten Ægirrückens weist die Form eines flachen 'V' auf. Diese entsteht, wenn sich während der Spreizung eine Materialzufuhr eines *hot spots* langsam entlang der Rückenachse bewegt (Vink et al. 1995, Abb. 4)[32].
Die westlichste Formation des Grönland-Färöer-Rückens, nämlich der Grönland-Island-Rücken, scheint aus Flutbasalten zu bestehen, die gefördert wurden, als das Zentrum des Manteldiapirs den östlichen Rand des grönländischen Kontinentalschildes erreicht hat.
Das Islandschelf erscheint in seiner klar abgegrenzten topografischen Form wie eine Durchprägung des Kopfes eines Manteldiapirs. Diese Form findet sich auch in den gravimetrischen Beobachtungen wieder (Abb. 2.8). Die scharfe Abgrenzung des östlichen Schelfs gegenüber dem Meeresboden gibt Anlass zur Annahme, dass hier Teile kontinentaler Kruste eingebettet sind (Heller & Marquart 2002).

Mögliche kontinentale Kruste unter Island

Im Zentrum Islands befindet sich ein Bereich mit einer verdickten Kruste, deren Entstehung noch diskutiert wird: so wird einerseits Krustenakkretion durch die Aktivität des Manteldiapirs vermutet, andererseits kann es sich auch um ältere Kruste handeln, die unter jüngeren Laven abgetaucht ist (Foulger 2006). Die gleichen Argumente, die für die Existenz älterer ozeanischer Kruste unter Island sprechen, sind gleichzeitig die schwerwiegendsten Gegenargumente: woher stammt die mutmaßlich auf 210 km Breite abgetauchte ozeanische Kruste, die existieren soll, damit bei der derzeitigen Bewegungsrate doch genug Kruste zustande kommt. Es gibt allerdings Hinweise, dass die Bewegungsraten in der geologischen Vergangenheit tatsächlich variierten (Mosar et al. 2002).
Existiert dennoch ein Mangel an Kruste, so erscheint alternativ es eher sinnvoll, alte kontinentale Kruste unter Island zu vermuten, denn in Grabenbruchstrukturen kann sich kontinentale Kruste sehr weit ausdehnen. Diese müsste etwa zeitgleich mit dem Jan-Mayen-Block von Grönland abgelöst worden sein (Abb. 2.9g). Ihren Ursprung würde dieses Stück Kruste südlich des Färöerplateaus gehabt haben. Diese überlagerte kontinentale Kruste kann die bisher ungeklärte Frage nach der Herkunft der explosiven silikatischen Laven auf Island beantworten. Aufgrund neuerer Erhebungen über die Verteilung von Tephren[33] in Mittel- und Westeuropa

[32] Ursprünglich wurde diese Annahme für den Island-Färöer-Rücken einschließlich des Färöerplateaus und Teile des Islandschelfes getroffen.

[33] Altgr.: *τέφρᾱ* 'Asche', pyroklastisches , zumeist ryolithisches Material, das bei explosionsartigen Vulkanausbrüchen weiträumig verteilt wird. Die Einteilung erfolgt über die Korngröße in Asche, Lapilli und Bomben (Press & Siever 2003, S. 78 u. 110).

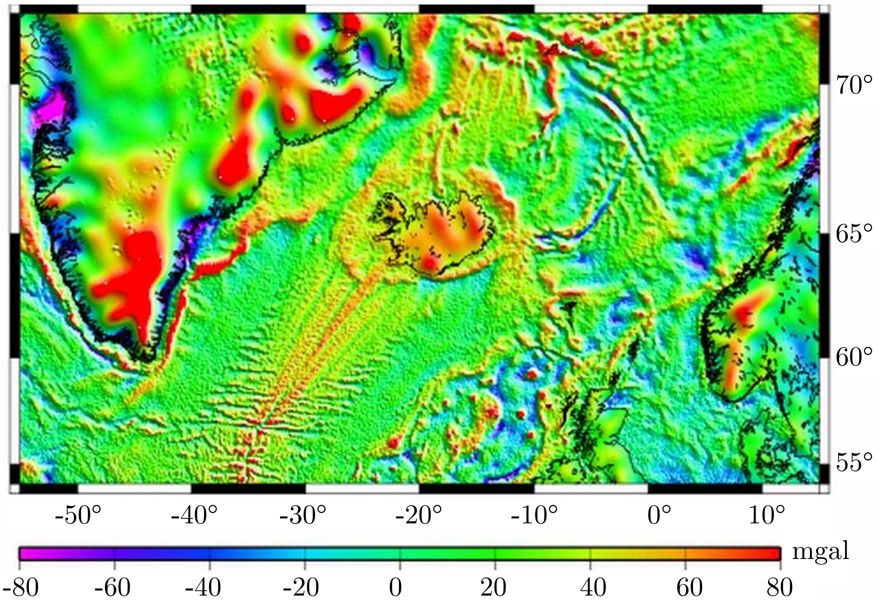


Abb. 2.8: Gravimetrische Freiluftanomalien im Nordatlantik (Sandwell & Smith 1997).

muss die Fördermenge silikatischen Materials mit mehr als 20% am isländischen Vulkanismus veranschlagt werden (Van der Bogaard 2002). Dass bisher aufgrund der Befunde zum Chemismus (Van der Bogaard 2002), zur Herkunft (Foulger 2006), zur Dichte und zur Mächtigkeit (Hartmann et al. 2007, S. 5) keine klare Entscheidung für eine reine ozeanische Kruste gefällt werden kann, lässt kaum einen anderen Schluss zu, dass sich alte kontinentale Bruchstücke eingebettet in junger ozeanische Kruste befinden.
Um die Bruchstücke der kontinentalen Kruste an einen Platz unter Island oder an Islands Ostküste zu befördern, ist eine bestimmte Reihenfolge des Erscheinens und Verödens von Rift- und Transformzonen im Nordatlantik notwendig. Das führt auf eine wiederkehrende Nördliche Vulkanische Riftzone und zwei stationäre Transformzonen (Abb. 2.9).

2.2 Seismologische Grundlagen

Dieser Abschnitt soll einen Überblick über die notwendigen geophysikalischen Termini und Verfahren in der Erfassung und Beschreibung von Erdbeben bereitstellen, auf die später im adaptiven Teil der Arbeit zugegriffen werden kann.

2.2.1 Ursachen von seismischen Ereignissen

Über die Ursachen von Erdbeben ist schon lange phantasievoll spekuliert worden. So vermutete Aristoteles noch ein Phänomen, dass er als *unterirdische Winde* beschrieb[34]. Es war bereits im 18. Jh. bekannt, dass sich Erdbeben häufig in sicht- und fühlbaren Wellen ausbreiten (Michell 1760, S. 8ff) und auch die Beziehung zwischen Vulkanen und Erdbeben schien ersichtlich (Michell 1760, S. 6 §5). Der physikalische Prozess als solcher war unerkannt und so amüsiert sich ein unbekannter Verfasser 1757 über die Vielzahl der widersprüchlichen Erklärungen seiner Zeitgenossen (Siehe Anh. C.3).

[34] Siehe hierzu das übersetzte Zitat aus der *Meteorologica* (Dombois 1998, S. 35)

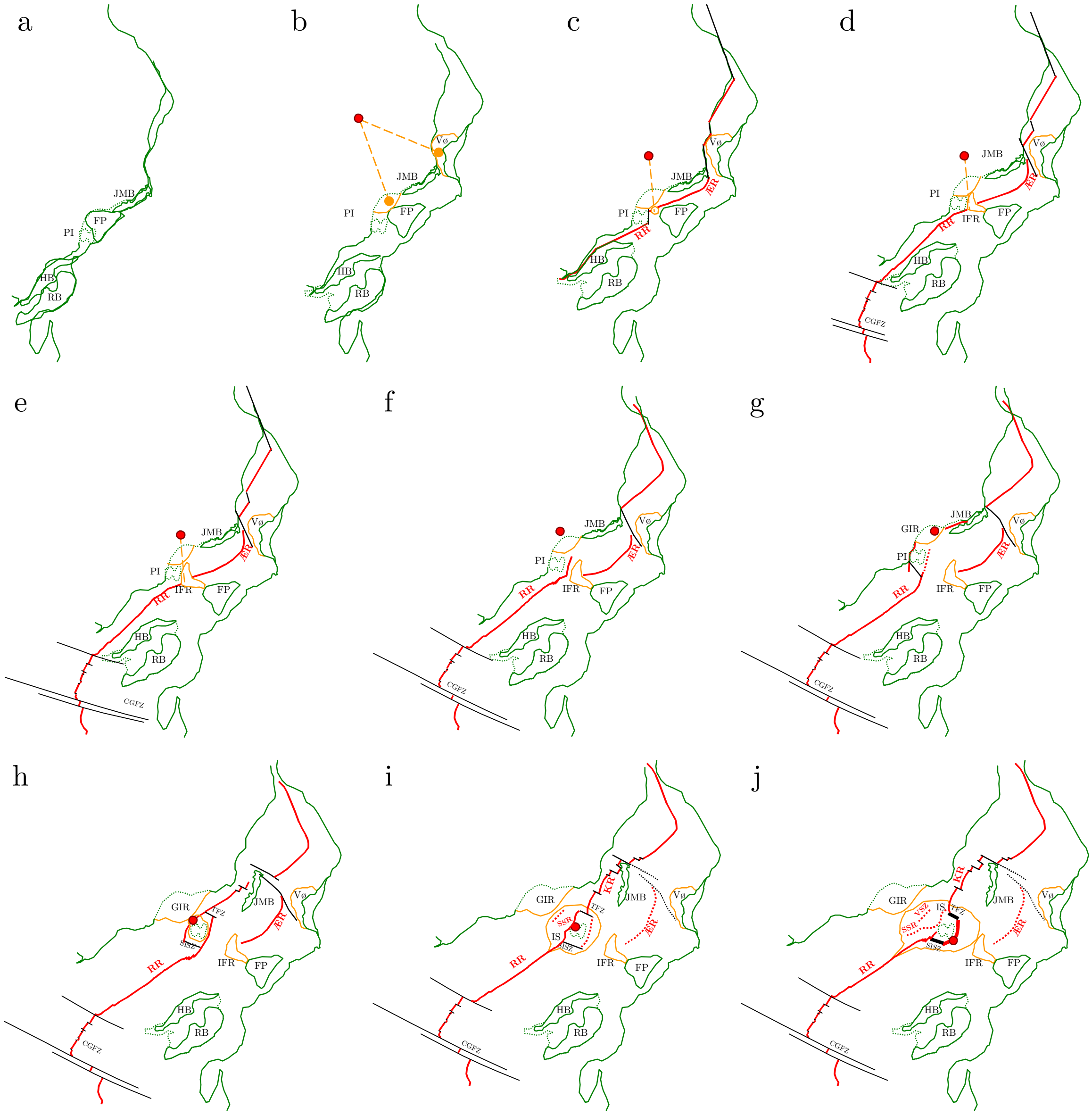


Abb. 2.9: a-b) Ab etwa 160 Ma entwickelt sich ein Grabenbruch. Durch den Manteldiapir entstehen ab etwa 60 Ma das Vøringplateau (Vø) und Teile des Grönland-Island-Rückens (GIR). c-e) Etwa 57-42 Ma entsteht der Island-Färöer-Rücken (IFR) und der Ægirrücken (ÆR) ist aktiv. Die Hatton- und die Rockallbank (HB, RB) und das Färöerplateau (FP) lösen sich. f-h) Etwa 38-25 Ma verödet der Ægirrücken. Der Kolbeinseyrücken (KR) wird aktiv. Der Jan Mayen-Block (JMB) und Protoisland (PI) lösen sich. i,j) Die Transformzonen (SISZ, TFZ) werden aktiv. Das Islandplateau entsteht.

Tektonische Beben

Tektonische Erdbeben sorgen für den Spannungsabbau am Rande elastischer Lithosphärenplatten. Diese Beben besitzen zumeist geringe Herdtiefen. An den Plattengrenzen kommt es zu verschiedenen Typen von tektonischen Beben. An den divergenten Plattengrenzen bilden sich unter dem Einfluss von vertikalen Bewegungen Grabenstrukturen heraus (PRESS & SIEVER 2003, Abb. 18.15). Die Erdbeben hier sind nicht sehr energiereich, da die betroffenen Strukturen eher klein sind und auch nicht viel Energie gespeichert werden kann.
An den Transformzonen leisten tektonische Beben den Abbau der Energie, die als elastische Verformung in den Platten gespeichert ist. Hier gleiten die Platten in einer ‚stotternden' Bewegung aneinander vorbei. In Abhängigkeit von dem orthogonalen Anpressdruck auf diese Zone können die Beben hier mittlere Stärken erreichen.
Die verheerendsten Beben haben sich aber bisher an konvergenten Plattengrenzen ereignet. Hier sind die Subduktionszonen noch mehr betroffen als die Kollisionszonen. Die überschiebende Platte lastet mit ihrem Gewicht auf der unterschiebenden Platte und somit auch auf der Scherfläche (PRESS & SIEVER 2003, Abb.18.16). Können zwei Platten in einer Transformzone oder in einer Grabenstruktur auch aseismisch aneinander vorbeigleiten, so ist das durch die Reibung an der Scherfläche einer Subduktionszone nahezu unmöglich.

Vulkanische Beben

Ein zweiter Prozess, der direkt zu Erdbeben führt, ist der Vulkanismus. Einerseits muss im Untergrund der Materialtransport ermöglicht werden. Dies geschieht durch Eröffnung von Spalten und Gängen einerseits durch eine partielle Aufschmelzung, aber eben andererseits auch durch Sprödbrüche. Diese lassen sich als seismischer Tremor erfassen.
An der Oberfläche hingegen müssen Spannungen des Deckengebirges abgebaut werden, wenn sich eine Magmenkammer in der Tiefe aufbläht (PRESS & SIEVER 2003, S. 127ff).

Tiefherdbeben

Nach den lange Zeit gängigen Vorstellungen konnten Erdbeben nur als Folge von Sprödbrüchen erfolgen. Sprödbrüche sind aber an rigides Material, also Lithosphärenplatten gebunden. Diese erreichen eine Tiefe von bis zu 200 km (PRESS & SIEVER 2003, S. 19). Innerhalb der darunter liegenden Asthenosphäre verhält sich das Material des oberen Mantels viskos. Der Grad der Viskosität ist abhängig vom Anteil an partieller Schmelze im Material. In der Asthenosphäre können demnach Erdbeben nicht stattfinden. Eine Ausnahme bilden hier subduzierte Lithosphärenplatten. Sie behalten auch noch in größerer Tiefe ihre rigiden Eigenschaften bei. Der hohe Druck auf eine abtauchende Lithosphärenplatte würde aber dazu geführt haben, dass Sprödbrüche und damit Erdbeben unterhalb von 400 km endgültig nicht mehr möglich sind. Denn ein Sprödbruch setzt neben

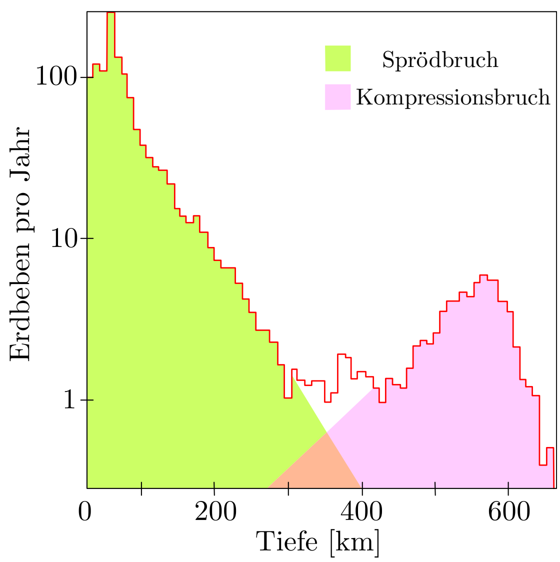


Abb. 2.10: Logarithmische Verteilung von Erdbeben in Abhängigkeit von der Tiefe (GREEN II 2001; GREEN II 2005).

der Existenz eines rigiden Materials auch die Möglichkeit zur Ausdehnung voraus. Bemerkenswerterweise steigt die Häufigkeit von Tiefherdbeben unterhalb von 400 km wieder an. Daher muss ein anderer Prozess als ein Sprödbruch diese Erdbeben verursachen. Aufgrund des hohen Druckes in der subduzierten Lithosphärenplatte tritt eine Umkristallisation von Olivin zu Spinell ein. Die sich komprimierenden Linsen aus Spinell wirken wie 'Antirisse' im Kristallgefüge (Green II 2001; Green II 2005). Mit zunehmender Anzahl dieser Linsen erhöht sich der Druck auf das verbliebene Olivin, so dass die Umkristallisation in einer Kettenreaktion ähnlich des Sprödbruches erfolgt. Solche Beben ereignen sich bis zu einer Tiefe von etwa 700 km.

Intraplattenbeben

Katastrophale Erdbeben können sich auch weit abseits von den aktuellen Plattengrenzen ereignen[35]. Diese Ereignisse sind höchst selten und geschehen damit kaum vorhersehbar. Eine Erklärung für diese Ereignisse besagt, dass der Stress von den Plattengrenzen in die Platte hinein propagiert und an Schwachstellen innerhalb der Platte abgebaut wird. Zwei Typen von Schwachstellen werden hierzu angefügt: einerseits in Gebieten mit einem großen Vorkommen radioaktiver Elemente (Pysklywec & Beaumont 2004), andererseits in Gebieten, in denen es schon ältere Verwerfungen sowie inaktive oder schwach aktive Grabenbrüche gibt, womit die Lithosphäre bereits teilweise vorgeschädigt ist (Iio & Kobayashi 2002). Während es bei tektonischen Erdbeben möglich ist, aufgrund der Plattengeschwindigkeiten statistische Aussagen über die Häufigkeit von Beben verschiedener Stärken zu tätigen, hängen Intraplattenbeben maßgeblich mit von der Viskosität der unteren Kruste ab (Iio et al. 2004), was eine statistische Aussage teilweise unmöglich macht.

2.2.2 Wirkung von seismischen Ereignissen

Zerstörungen

Wie sich aus den Niederschriften aller Jahrhunderte verfolgen lässt, besitzen die Zerstörungen, die durch ein Erdbeben verursacht werden, die eindrücklichste Wirkung auf die Menschen. Dabei sind die primären Schäden eines Bebens oft weniger katastrophal als die sekundären, die durch Flutwellen[36], Erdrutsche, Brände und Explosionen[37] ausgelöst werden. Ein bleibendes Relikt der meisten oberflächennahen Beben sind sichtbare Spalten und Verwerfungen. Diese Verschiebungen können leicht mehrere Meter betragen.

Seismische Wellen

Die Wissenschaftler des 18. Jh. erkannten als erstes die Erdbewegung durch die vertikalen Oberflächenwellen (Michell 1760, S. 8ff). Diese sogenannten Rayleigh-Wellen (Abb. 2.11a) werden zumeist begleitet von den horizontalen Transversalwellen, den Love-Wellen (Abb. 2.11b) (Press & Siever 2003, Abb. 18.8a u. 18.8b). Weit interessanter für

[35] Beispielhaft für diesen Bebentyp ist das Beben in New Madrid im US-Bundesstaat Missouri im Jahre 1812 (Press & Siever 2003, S. 493 u. Abb. 18.11), sowie das Beben bei Basel 1356 (Lambert et al. 2005).

[36] Der Tsunami des 25/26.12.2004 im Indischen Ozean war weit verheerender als primären die Auswirkungen des Sumatrabebens (Vigny et al. 2005), das diese Welle auslöste, obgleich dieses Beben das stärkste war, das bis heute aufgezeichnet worden ist.

[37] Auch das San Francisco-Beben von 1906 hatte nur mäßige Schäden hinterlassen. Erst die verschiedenen Brände, teilweise auch von den Bewohnern selbst gelegt, um später den Versicherungsschaden erstattet zu bekommen, führte zu der nahezu vollständigen Zerstörung der Stadt.

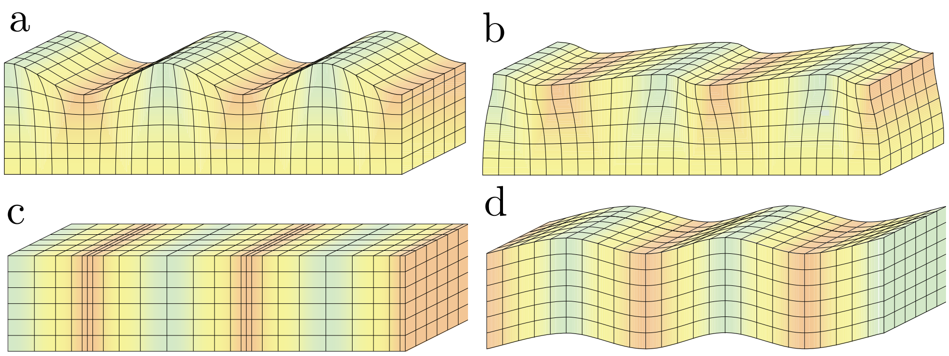


Abb. 2.11: Seismische Wellen: a) die vertikale, retrograde Rayleigh-Welle und die b) horizontale Love-Welle finden sich an der Oberfläche; in der Tiefe hingegen c) die kompressive primäre und d) die transversale sekundäre Raumwelle.

seismologische Untersuchungen sind aber die Tiefenwellen im Untergrund. Die Kompressionswelle, die im 45°-Winkel zur Herdfläche ausgesandt wird, ist die erste Welle, die ein entferntes Seismometer erreicht, weswegen sie auch als Primärwelle oder kurz als P-Welle bezeichnet wird (Abb. 2.11c). Eine kompressive Welle passiert – wenngleich langsamer – Flüssigkeiten, demnach auch den flüssigen äußeren Erdkern. In normaler Richtung zur Herdfläche entsteht die transversale Sekundärwelle oder kurz S-Welle (Abb. 2.11d). Transversale Wellen laufen in Flüssigkeiten abhängig von der Viskosität auf kürzesten Distanzen tot. Diese Wellen können den Erdkern nicht passieren, so dass die gegenüberliegende Seite der Erde im Wellenschatten liegt.
An jeder Grenzfläche werden P- und S-Wellen reflektiert. Die ursprüngliche Welle und ihre Reflexionen benötigen in Abhängigkeit des Untergrundes verschiedene Laufzeiten. Hierzu gibt es allgemeine Laufzeitmodelle (Press & Siever 2003, Abb. 19.7). Es werden aber aufgrund einer besonderen Rheologie auch immer lokale Modelle benötigt, die umgekehrt Aussagen über den Untergrund treffen (Wolfe et al. 1997; Foulger et al. 2000; Wolfe et al. 2002).

2.2.3 Quantifizierung von seismischen Ereignissen

Quantifizierung nach Schäden

Um zu einer ersten nachvollziehbaren Quantifizierung von Erdbeben zu gelangen, stellte Mercalli zu Beginn des 20. Jahrhunderts eine sogenannte *Skala*[38] auf, in der anhand der Wirkung des Bebens eine numerische Zuordnung möglich ist (Cara 1994, S. 45). Diese Aufstellung wurde bereits mehrfach angepasst, so dass heute auch die Modifizierte Mercalli-Skala (MM), die Mercalli-Cancani-Sieberg-Skala (MCS) von 1930 und die für Mitteleuropa gültige Medvedev-Sponheuer-Karnik-Skala (MSK) von 1963 verwendet werden. Die Einteilung von Beben nach ihren Schäden hat ihre Bedeutung noch nicht verloren, wie die Arbeiten an der Europäischen Makroseismischen Skala (EMS) belegen (Grünthal 1998).

Gutenberg-Richter-Magnitude

Die auch als „lokale“ Magnitude bezeichnete Maßeinheit für Erdbeben entstand eigentlich als Kalibrierfunktion

$$M_L = \log A(\Delta) - \log A_0(\Delta) = \log \left(\frac{A(\Delta)}{A_0(\Delta)} \right) \tag{2-1}$$

[38] Ital.: *la scala*, 'die Leiter'.

der ANDERSEN-WOOD-Torsionspendel-Seismometer (KERTZ 1969, S. 32). Darin ist $A(\Delta)$ die Amplitude der hochfrequenten Anteile aller seismischen Wellen. Sie wird als dekadischer Logarithmus angegeben. Um zu berücksichtigen, dass entferntere Beben gleicher Stärke einen geringeren Ausschlag verursachen, wird der ursprüngliche Wert um den entfernungsabhängigen *offset* $\log A_0(\Delta)$ korrigiert. Dieser liegt nicht als vollständig geschlossene Funktion der Entfernung Δ zum Erdbebenherd vor, sondern ist vertafelt (RICHTER 1935, Tab. 1). In einem Entfernungsbereich von 200 bis 600 km ist die Näherungsformel

$$\log A_0(\Delta) = 3,37 - 3\log\Delta \Longleftrightarrow A_0(\Delta) = 2,35\Delta^{-3} \tag{2-2}$$

aber hinreichend genau (RICHTER 1935, S. 13). Um mit dieser Magnitudeninformation zu einer Näherungsgröße der Energie zu gelangen, wurde die Herdtiefe z in der Form

$$\log E(z) = \begin{cases} 19,4 + 0,9M_L & z \leq 30\,\text{km} \\ 21,2 + 0,6M_L & 30\,km < z \leq 300\,\text{km} \\ 20,5 + 0,6M_L & z > 300\,\text{km} \end{cases} \tag{2-3}$$

berücksichtigt (GUTENBERG & RICHTER 1949, S. 19). Das Ergebnis E ist in $[10^{19}$ J] angegeben[39]. Grundsätzlich sind die Parameter in Formel (2-3) nur für Kalifornien gültig. Weltweit kann vereinfacht mit

$$\log E = 11,4 + 1,5M_L \tag{2-4}$$

gerechnet werden (KERTZ 1969, S. 33).

Magnitude der Oberflächenwellen

Die Definition der lokalen Magnitude anhand eines spezifischen Messgerätes war zwar pragmatisch, hatte aber eine Vielzahl von Nachteilen. Der entscheidende ist, das die verschiedenen Wellentypen (Abb. 2.11) nicht voneinander zu unterscheiden waren. Die langlebigste seismische Welle mit der höchsten Amplitude ist die Oberflächenwelle. Ihr vertikaler Anteil, gebildet durch eine retrograde RALEIGH-Welle (Abb. 2.11a) kann bei starken Erdbeben nahezu auf der ganzen Welt erfasst werden. Damit ergibt sich die Magnitude als

$$M_s = \log\frac{A_s}{T} + 1,66\log D + 3,3, \tag{2-5}$$

wobei A_s die Amplitude der vertikalen Komponente der Oberflächenwelle innerhalb eines festgelegten Frequenzbereiches ist. Weiterhin werden die Umlaufzeit T[s] und die Entfernung D[°] benötigt (PERLT 2006, Anh. D).

Magnitude der P-Wellen

Die erste Welle, die ein Seismometer erreicht, ist eine kompressive primäre Raumwelle oder kurz P-Welle. Aus dem Quotienten ihrer Bewegungsamplitude im Untergrund A und der Umlaufzeit T ergibt sich unter Einschränkungen eine annähernd konstante Geschwindigkeit für alle Erdbeben (GUTENBERG & RICHTER 1956). Aus ihr lässt sich eine Magnitude

$$M_b = \log\frac{A}{T} + Q(D, z) \tag{2-6}$$

berechnen, wobei $Q(\cdot)$ eine Funktion der Entfernung D und der Herdtiefe z ist.

[39] GUTENBERG und RICHTER rechnen hier noch in [erg], das entspricht $[10^{-7}$J].

Moment-Magnitude

Um eine quantitative Aussage zu einem Beben machen zu können, der auch eine physikalische Modellvorstellung entspricht, ist die Moment-Magnitude

$$M_w = \frac{2}{3} \log M_0 - 10,7 \tag{2-7}$$

entwickelt worden (HANKS & KANAMORI 1979). Das hierzu benötigte skalare seismische Moment

$$M_0 = \mu F_0 d_0 \tag{2-8}$$

errechnet sich aus dem Schermodul μ, für das zumeist $3 \cdot 10^{10}\mathrm{Nm}^{-2}$ angesetzt wird, der Herdfläche F_0 und der mittleren Scherbewegung d_0 auf dieser Fläche (CARA 1994, S. 53f).

Herdflächenlösungen

Eine zweite wesentliche numerische Festlegung eines seismischen Ereignisses ist die Ausrichtung der Herdfläche im Raum. Zur Berechnung bedient man sich der Phasenlagen sowohl der kompressiven primären als auch der transversalen sekundären Raumwellen, so wie sie an den weltweit verteilten Seismometer eintreffen.
Betrachtet man das Beispiel einer Blattverschiebung mit zwei gekreuzten Kräftepaaren[40], so wird das Maximum der kompressiven Phase der P-Welle jeweils in einem rechtsdrehenden 45°-Winkel von der Hauptbewegungsrichtung ausgesandt (Abb. 2.12). Die Maxima der dilatativen Phase stehen orthogonal auf den kompressiven Maxima. Die Ebene zwischen den unterschiedlichen Phasenmaxima wird als Nodalebene[41] bezeichnet. Die Maxima der S-Wellen befinden sich in den Nodalebenen der P-Wellen und umgekehrt.
Ist die Position des Bebens über das HERGLOTZ-WIECHERT-Verfahren (CARA 1994, S. 65ff) bestimmt, kann jede Phasenlage an einem Seismometer in seine ursprüngliche Aussenderichtung zurückgerechnet werden. Die Vielzahl der Ergebnisse von den verschiedenen Stationen beschreibt dann die Lage der Herdfläche im Raum und damit auch

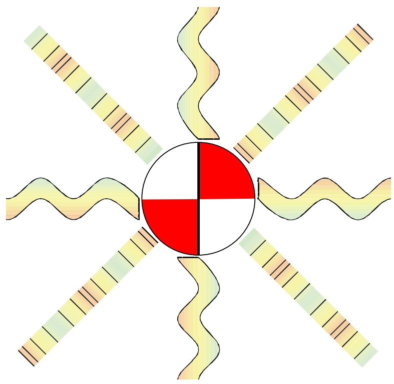


Abb. 2.12: Herdflächenlösung eines Erdbebens einer Blattverschiebung mit zwei gekreuzten Kräftepaaren: Um die Herdfläche (schwarz) ergeben sich Bereiche der Kompression (rot) und der Dilatation (weiß). Die P- und S-Wellen mit verschiedenen Phasenlagen in die unterschiedlichen Richtungen ausgesandt.

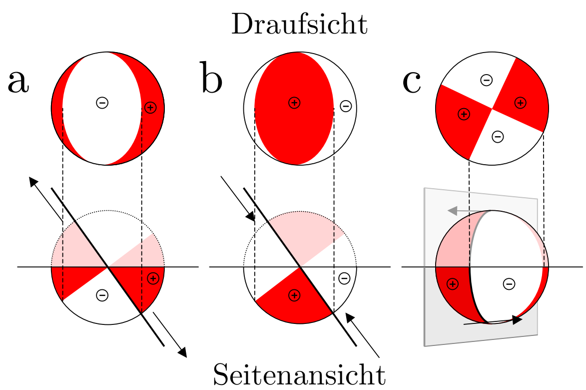


Abb. 2.13: Darstellungen von Herdflächenlösungen: Die sogenannten *beachballs* sind die Draufsicht in die untere Halbschale der kugelförmig aufgetragenen Dilatationen und Kompressionen für a) eine Abschiebung, b) eine Aufschiebung und c) eine Blattverschiebung.

[40] Engl.: *strike slip with double couple*.
[41] Lat.: *nodus* 'der Knoten'.

den Herdmechanismus. Liegt die Schnittgerade aller Nodalebenen parallel zur Erdoberfläche, so kommen zwei Mechanismen in Frage. Zum einen besteht die Möglichkeit einer Abschiebung (Abb. 2.13a) oder einer Aufschiebung (Abb. 2.13b) (Stüwe 2000, S. 196ff). Im anderen Extremfall, wenn die Schnittgerade der Nodalebenen lotrecht steht, handelt es sich um eine Blattverschiebung, also ein Beben in einer lotrechten Transformstörung (Abb. 2.13c). Zwischen diesen beiden Extrema existieren alle möglichen schiefen Herdflächenlösungen.

2.2.4 Vorzeichen von seismischen Ereignissen

Es ist immer versucht worden, Erdbeben zuverlässig vorherzusagen. Denn Erdbeben geschehen nicht, ohne sich zuvor anzukündigen. Verschiedene Vorzeichen sind bereits beschrieben oder vermutet worden. Manchen Erdbeben gehen Vorbeben voraus und es ist auch schon gelungen, aus deren Charakteristik das folgende Hauptbeben vorherzusagen (Press & Siever 2003, S. 499).
Ein weiteres Vorzeichen für einige Beben ist der vermehrte oder verminderte Austritt von Radon aus den sich zuvor im Untergrund öffnenden Spalten (Einarsson et al. 2008). Die Konzentrationen scheinen dabei gelegentlich so hoch zu sein, dass in den Nächten vor den Beben auch von einem diffusen Leuchten berichtet worden ist.
Auch die Veränderung des elektrischen Potentials ist ein Indikator für Deformationen im Untergrund (Trique et al. 1999; Freund 2007a; Freund 2007b). In der Diskussion befindet sich weiterhin ein seismischer Tremor im Infraschallbereich. Er sorgt dafür, dass viele Tiere für Beobachter aus unerklärlichem Anlass unruhig werden (Press & Siever 2003, S. 497).
Die Schwierigkeit liegt aber darin, dass diese Phänomene nicht immer auftreten und vor allem, wenn sie auftreten, in aller Regel nicht gezielt überwacht werden.
Umso mehr Gewicht erhalten die bereits regelmäßig erfassten geodätischen und geophysikalischen Größen in einem gefährdeten Gebiet. Insbesondere deren unerwartete Veränderungen sollten Aufmerksamkeit erregen.

2.3 Geodätische Grundlagen

Dieser Abschnitt soll lediglich die notwendigen Grundlagen aus der Geodäsie zusammenstellen und deren Bezug zur Thematik herstellen. Ansonsten wird auf die einschlägige weiterführende Literatur verwiesen.

2.3.1 Geodätische Bezugssysteme

Jede absolute oder relative Ortsbestimmung ist maßgeblich an die Festlegung eines Koordinatensystems geknüpft. Die Beschreibung des Ortes innerhalb dieses Systems ist der Ortsvektor $\mathfrak{x}$. Für die Nutzung von Satellitenbeobachtungssystemen sind zunächst nur die globalen Systeme von Bedeutung. Erst die kleinräumige Nutzung benötigt eine Überführung in ein lokales, zumeist an das Schwerefeld gekoppeltes System (Seeber 2003, S. 21ff).

2.3.2 Globale terrestrische geodätische Bezugssysteme

Die *terrestrischen* oder erdfesten Bezugssysteme sind im Gegensatz zu den *zälestischen* Systemen an die Rotation der Erde gebunden (Torge 2003, S. 23ff). Astronomische Messungen der Zeit, der Erdrotationsgeschwindigkeit und der Lage der Rotationsachse verbinden terrestrische und zälestische Systeme. Die terrestrischen Bezugssysteme orientieren sich mit ihrer XY-Ebene an der Äquatorebene. Die Z-Achse des Systems beschreibt eine Näherung der Rotationsachse der Erde (Torge 2003, S. 29f).

Ein geodätisches *Datum* beschreibt die Festlegung des Bezugssystems auf einem Rotationsellipsoid sowie dessen Form und exakte Ausrichtung, sowie den Maßstab des Systems (Hooijberg 1997). Die Verknüpfung eines Ellipsoids mit dem kartesischen Koordinatensystem erfolgte zumeist mittels eines Ursprungspunktes nahe des Masseschwerpunktes der Erde. Ältere lokale Datumsfestlegungen beziehen sich häufig auf einen einzelnen Hauptpunkt an der Erdoberfläche, wie zum Beispiel der Hauptpunkt Rauenberg oder der Normalhöhenpunkt in der Sternwarte Berlin. Neuere Festlegungen haben zwar auch eine solche punktweise Definition, diese ist aus Realisierungsgründen aber unter der Annahme der linearen Punktbewegung auf viele Punkte verteilt und kann demnach ausgeglichen werden (Perlt 2006, Abb. 2.1).
In der Praxis hat sich der Internationale Terrestrische Bezugsrahmen[42] ITRF durchgesetzt. Über 500 permanente Satellitenempfangsstationen für GPS, SLR, VLBI und DORIS[43] werden hier mit ihrer Position und Geschwindigkeit jährlich neu festgelegt. Das ITRF ist so gelegt, dass sich keine signifikante Rotation gegen das globale Krustenbewegungsmodell NUVEL-1A-NNR ergibt.

2.3.3 Globale Krustenbewegungsmodelle

Die relativen Bewegungen von zwei Lithosphärenplatten an einer Plattengrenze lassen sich in der lokalen und auch regionalen Betrachtung leicht durch verebnete Vektoren beschreiben. Um aber zu einem globalen Modell zu gelangen, muss man sich von der ebenen Anschauung lösen und mit möglichst einfachen Parametern die Bewegung einer Lithosphärenplatte beschreiben. Das erscheint zunächst nicht einfach, wenn man sich das verebnete Vektorfeld auf einer Weltkarte ansieht. Um die Bewegung einer einzelnen Lithosphärenplatte zu verstehen, stelle man sich eine halbe Kugelschale vor, die auf einer entsprechend kleineren Kugel verschoben werden kann. Indem sich ein Großkreis der Halbschale geradlinig bewegt, existieren zwei Randpunkte, welche stillstehen und um welche sich die Halbschale dreht (Abb. 2.14). Ein solcher Punkt heißt *Euler*-Pol. Gibt man die Position eines der beiden *Euler*-Pole und die Rotationsgeschwindigkeit um die Achse durch die beiden Pole an, so ist die Bewegung der Halbschale relativ zur Kugel vollständig beschrieben.

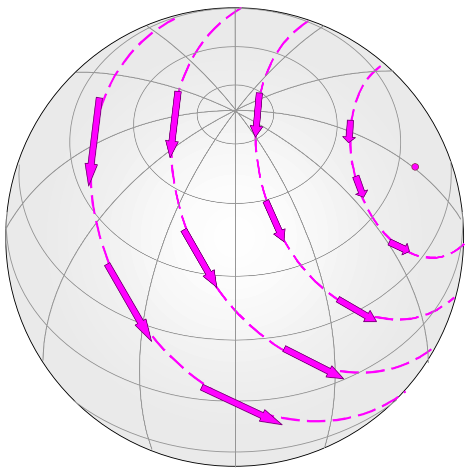


Abb. 2.14: Eulerpol-Bewegung

Wie für diese eine Halbschale können nun für jede Lithosphärenplatte, welche idealisiert als Kugelschale angesehen wird, die drei Parameter der geographischen Länge und Breite des Eulerpoles, sowie die Rotationsgeschwindigkeit am zugehörigen Großkreis angegeben werden.
Grundsätzlich existieren drei verschiedene Gruppen von Bewegungsmodellen. Die erste ist repräsentiert durch das NUVEL-1 (DeMets et al. 1990), mit der Bedingung, dass ein Kontinent als Referenzkontinent festgehalten wird. Zumeist wird die Geschwindigkeit von Afrika zu Null gesetzt, aber es sind hier auch andere Lösungen etabliert, wie beispielsweise die Festlegung auf die Pazifische Platte (PCFC) (Tab. 2.1).
In der zweiten Gruppe wurde im NUVEL-1A-NNR (siehe auch Tab. 2.1) mit der sogenannten *no net rotation*-Bedingung oder kurz NNR eingeführt[44] (DeMets et al. 1994). Diese Be-

[42] Engl.: *International Terrestrial Reference Frame*. Es sei hier unterschieden zwischen dem *Bezugssystem* als geophysikalisch-mathematische Definition dem *Bezugsrahmen* als dessen messtechnische Realisierung.
[43] SLR: *satellite laser ranging*, VLBI: *very long baseline interferometry*, DORIS: *Doppler orbitography and radio-positioning integrated by satellite*.
[44] Vergleiche das Zitat am Kapitelanfang aus (Wegener 1929, S. 152).

Tab. 2.1: Divergente tektonische Bewegung der Plattengrenze in Island aus verschiedenen Globalbewegungsmodellen (Perlt 2006, Tab. 5.2). Modelle unter verschiedenen Minimierungsbedingungen: NNR – Summe der Plattenbewegung ist Null (*no net rotation*), PCFC – Pazifische Platte stabil, HS3 – *Hot spot*-Positionen stabil. Δv steht für die relative Geschwindigkeit zwischen zwei Platten im Azimuth ΔAz.

Modell	$\rightarrow min$	Δv $^{mm}/_{yr}$	ΔAz $^\circ$	Quelle
ITRF 2000	NNR	17,28	284,10	(Drewes & Angermann 2001)
GRSM v1.2	NNR	17,93	283,58	(Kreemer et al. 2003)
APKIM 2000.0	NNR	18,61	283,01	(Drewes & Angermann 2001)
NUVEL 1A	HS3	18,84	284,70	(Gripp & Gordon 2002)
NUVEL 1A	NNR	18,84	284,70	(DeMets et al. 1994)
NUVEL 1	PCFC	19,70	284,71	(DeMets et al. 1990)
NUVEL 1	NNR	19,86	286,68	(DeMets et al. 1990)
REVEL 2000	NNR	19,85	282,54	(Sella et al. 2002)

dingung besagt, dass die Summe aller Kontinentbewegungen den Wert Null annehmen muss. Aufgrund der Relativbewegungen eines Kontinents zu allen seinen Nachbarn ist jetzt das Bewegungsfeld eindeutig definiert. Die Bewegung der Kontinente muss aber nicht in einer geophysikalischen Beziehung zum Erdkörper stehen.
Eine an den Erdkörper gebundene Definition benutzt die Theorie, dass ein Manteldiapir ortsfest ist, beziehungsweise gegenüber den Platten vernachlässigbar kleine Bewegungen ausführt und sich dessen Spur, ein sogenannter aseismische Rücken (Press & Siever 2003, S. 126), gut verfolgen lässt[45]. Damit ist jede numerische Plattenbewegung optimal an die Plattenbewegung bezüglich der Manteldiapire anzupassen. Damit ergibt sich die dritte Gruppe von Bewegungsmodellen, woraus das NUVEL-HS3 das bekannteste Modell ist (Gripp & Gordon 2002).

2.3.4 Permanentstationen des *Global Positioning System*

Vom *International GNSS Service* werden weltweit verteilte GPS-Permanentstationen gemeinsam betrieben (Abb. 2.15). Das bedeutet, dass es eine gemeinsame Datensammlung, Datenvorhaltung und Auswertung bei den verschiedenen Partnern gibt. Derzeit sind etwa 350 Stationen in dieses Netz eingebunden (IGS 2008). Als Daten werden die Satellitenbahnen in verschiedenen Prozessierungsstadien aber auch die empfangenen Satellitendaten gehalten. Ebenfalls von großer Bedeutung sind die Stationsbeschreibungen, die Auskunft über den derzeitigen Status und alle zurückliegenden Veränderungen der Station geben.

2.3.5 Saisonale und Pseudosaisonale Einflüsse auf das *Global Positioning System*

Die Funktionsweise des *Global Positioning System* (GPS) und die Qualität der jeweiligen Messergebnisse sind bereits breit diskutiert worden (Zumberge et al. 1997; Riedel & Heinert 1998; Dach 2000; Wieser 2001; Seeber 2003; Heinert & Riedel 2007, u.v.a.). Im Rahmen geodätischer und geophysikalischer Projekte auf Island sind auch die lokalen Gegebenheiten für GPS-Messungen wiederholt umfangreich beschrieben worden (Hackman 1991; Jahn 1992; Foulger et al. 1993; Sturkell et al. 1994; Völksen & Seeber 1998; Alex et al. 1999; Hreinsdóttir 1999; Völksen 2000; Árnadóttir et al. 2001; Perlt 2006; Perlt

[45] Inzwischen gibt es eine Reihe von Arbeiten, welche die absolute als auch die relative Ortstabilität von Manteldiapiren anzweifeln (Anretter 2001; Stock 2003; Tarduno et al. 2003).

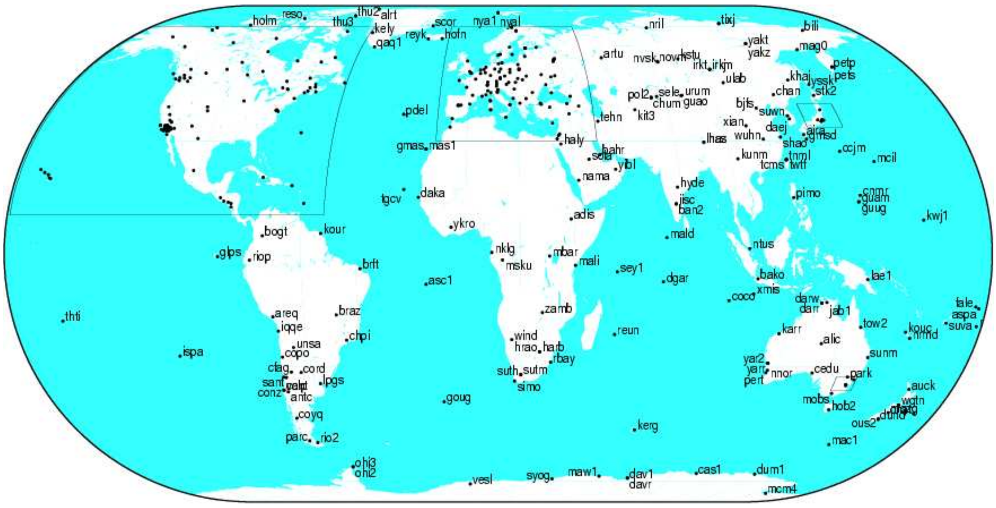


Abb. 2.15: Das globale IGS-Stationsnetz (IGS 2008).

ET AL. 2008, u.v.a.). Diese Fragestellungen sollen in der vorliegenden Arbeit nicht nochmals eröffnet werden. Vielmehr soll davon ausgegangen werden, dass GPS-Resultate nach dem besten Stand des Wissens erzeugt worden sind, dass diese also im Rahmen der Messgenauigkeit und des Fehlerbudgets als *richtig* zu betrachten sind. Hierauf wird im Kapitel 5.1.1 detaillierter eingegangen.

Offen bleibt allerdings die Frage, inwieweit Troposphäreneffekte die Zeitreihen der prozessierten Koordinaten von permanenten GPS-Messungen beeinflussen kann. Es kann hierbei konstatiert werden, dass alle Zeitreihen von permanenten GPS-Stationen (SAPOS, IGS u.a.) in allen Koordinatenkomponenten ein regelmäßiges und jährlich wiederkehrendes Muster aufweisen (STRERATH 2006). Das *Budget der systematischen Einflüsse*, das hier als deutlich geschieden vom *Budget der zufälligen Fehler*[46] zu betrachten sein soll, weist für jede Station sehr individuelle Einflüsse auf:

- wiederkehrende Satellitenkonstellation,
- Abhängigkeit elektromagnetischer Wellen von der Wasserdampfsättigung der unteren Atmosphäre,
- Schneebedeckung der Antenne,
- wechselnde Hoch- und Tiefdruckgebiete,
- wechselnder Grundwasserspiegel,
- einseitige Sonneneinstrahlung auf die Antenne und ihre Gründung.

Ein besonderes Phänomen tritt auf, wenn einer der genannten atmosphärischen Einflüsse direkt mit der Satellitenkonstellation zusammenwirkt. Alle atmosphärischen Einflüsse folgen zumeist einem ausgeprägten Tagesgang. Dieser ist aber vom Sonnentag abhängig. Die Satellitenkonstellation ihrerseits ist an den Sternentag gebunden. Daraus ergibt sich eine theoretische zeitliche Phasenverschiebung von 3'56"71 pro Tag. Über das Jahr wirkt diese tägliche Phasenverschiebung als *Alias* (Kap. 3.3.2) mit der Umlaufzeit eines Jahres. Die Tatsache, dass häufig GPS-Tageskoordinaten vorliegen, sorgt so für das Auftreten von Pseudosaisonalitäten.

---

[46] Die Beschreibung der zufälligen Einflüsse auf das GPS-System wurden bereits ausführlich diskutiert (SEEBER 2003, S. 299).

2.4 Die Tektonik Islands

Die passive Plattengrenze auf Island gliedert sich in mehrere Abschnitte mit ganz eigenen Charkteristika. Diese Gliederung ist eine Folge der Wechselwirkung zwischen der Konvektionswalze des Mittelatlantischen Rückens und dem Manteldiapir unter Island.

2.4.1 Die vulkanischen Riftzonen

Auf Island finden sich drei oberflächliche Aktivitätszentren des Manteldiapirs, ähnlich der Inselkette Hawaii (Wilson 1966), aber mit dem entscheidenden Unterschied, dass sie sich hier als topografisch herausragende Formationen in einer Reihe auf der Insel wiederfinden lassen (Foulger et al. 2005, Abb. 2): die älteste Formation wird durch die Westfjorde (Vestfirðir) im äußersten Nordwesten der Insel gebildet. Eine andere Formation scheint sich durch die Zentralvulkane unter den Gletschern Hofsjökull und Langjökull zu zeigen (Lawver & Müller 1994)[47]. Schließlich findet sich das vulkanische Orogen unter Europas größtem Gletscher, dem Vatnajökull, mit den Calderen der Zentralvulkane Bárðarbunga, Kverkfjöll, Grímsvötn, Esjufjöll und Þórðarhyrna. Bei dieser Formation besteht heute kein Zweifel mehr, dass sie das aktuelle Zentrum über dem Manteldiapir bildet (Einarsson et al. 1997). Kongruent zu diesen drei Zentren finden sich auch seitliche *rift zones* parallel zur Hauptachse des Mittelatlantischen Rückens (Sigmundsson 2006, S. 25), die sich jeweils zeitgleich ausgebildet haben müssen (Trønnes 2002, Abb. 2). Die jüngste aktive vulkanische Zone ist bis zu 150 Kilometer aus der Achse des Mittelatlantischen Rückens verschoben. Diese Exzentrizität lässt sich vereinfacht dahingehend deuten, dass der Manteldiapir einen Teil des Materialtransports der Konvektionswalze in seiner Umgebung mitgezogen hat (Vink et al. 1995).
Im Norden der Zentralvulkane unter dem Vatnajökull ist für die divergente Bewegung nur die *Nördliche Vulkanische Riftzone* (NVRZ)[48] verantwortlich.
Ein herausragendes Ereignis in diesem Gebiet war die Riftepisode des Krafla Zentralvulkansystems von 1975 bis 1984 (Sigmundsson 2006, Kap. 5.4). Das etwa 100 km lange, N-S gerichtete Spaltensystem vom Myvatn bis Axarfjörður durchschneidet die zentrale Caldera der Krafla (Björnsson et al. 1979, Abb. 4.). Im Zeitraum dieser neun Jahre hat sich das Spaltensystem in Ost-West-Richtung um bis über 7 m geöffnet (Möller 1989)[49].
Die *Südliche Vulkanzone* erschien lange Zeit als diffuse Plattengrenze (Schutzbach 1985; Czubik 1989, Abb. 19). Sie unterteilt sich in die ältere, tektonisch weniger aktive *Westliche Vulkanische Riftzone* (WVRZ) und die jüngere und tektonisch, wie vulkanisch aktivere *Östliche Vulkanische Riftzone* (EVRZ). Zwischen diesen beiden Zonen befindet sich die weitgehend inaktive Hreppar Mikroplatte. Diese räumliche Aufteilung dieser beiden Spreizungszonen ist deutlichste Erscheinung des Einflusses des Manteldiapirs auf die Plattengrenze an der Oberfläche. Während die Aktivität des westlichen Teils langfristig abklingen wird, nimmt die Aktivität im östlichen Teil noch zu.
An der Schnittstelle zwischen Westlicher Vulkanischer Riftzone, Reykjanesrücken und Südisländischer Seismischer Zone befindet sich die Hengill *triple junction*, benannt nach dem Zentralvulkan Hengill (Abb. 2.16). In der Zeit von 1993 bis 1998 ereignete sich am Hengillmassiv eine seismisch hochaktive Phase (Sigmundsson 2006, S. 100). Sie ging mit einer Spreizung der Flanken von 25 cm in WNW-OSO-Richtung (Perlt & Heinert 2006) und einer Aufwölbung von 6 cm[50] einher.

[47] Die Annahme, dass diese Formation eine Spur des *hot spot* sein kann, legen auch die Ergebnisse neuerer Untersuchungen nahe (Eysteinsson & Gunnarsson 1995; Allen et al. 2002a; Allen et al. 2002b).

[48] Bisher wurde diese Zone in der Literatur durchgehend als *Nördliche Vulkanzone (NVZ)* bezeichnet. Neuerdings werden die vulkanischen Riftzonen von den vulkanischen Flankenzonen unterschieden (Sigmundsson 2006, Abb. 3.9).

[49] Siehe auch Perlt et al. 2008, Abb. 2.

[50] Ergebnis einer bisher unveröffentlichten Deformationsanalyse des Autors.

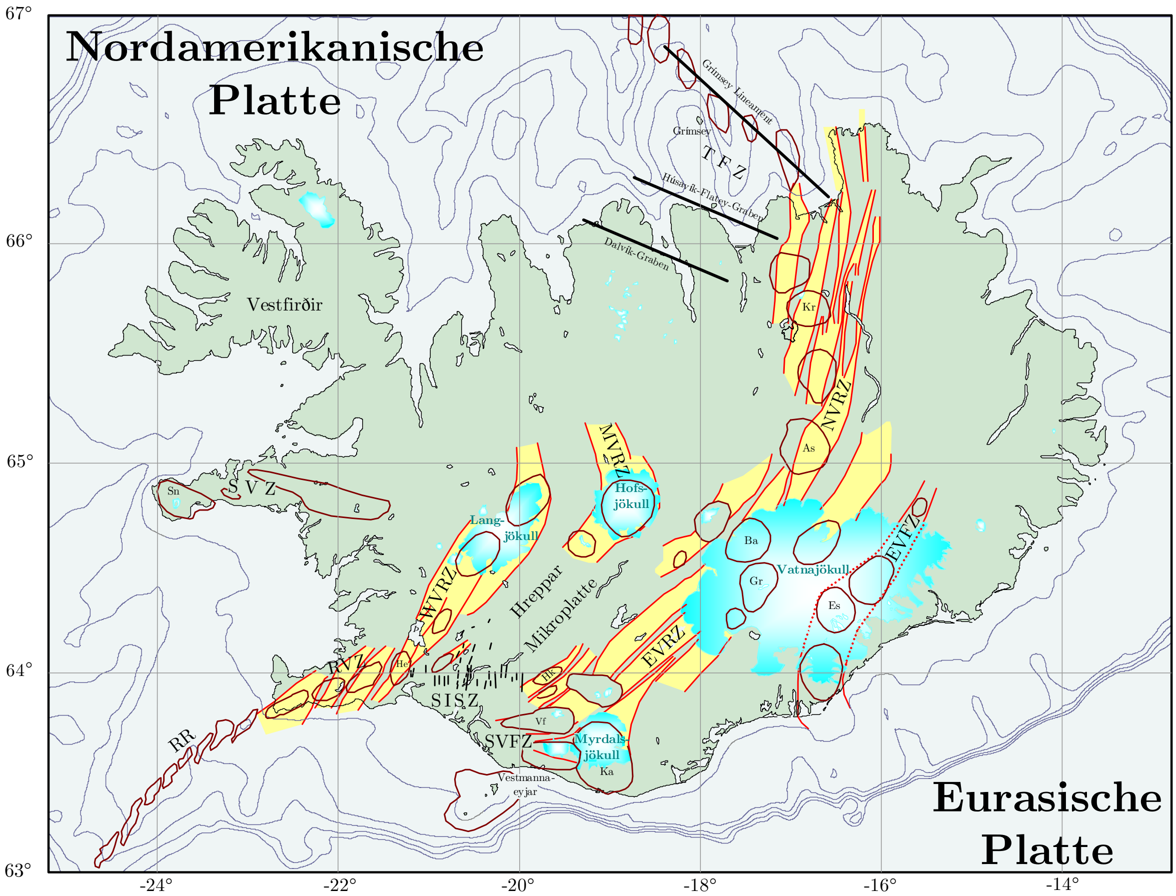


Abb. 2.16: Tektonik Islands: die Nördliche (NVRZ), die Westliche (WVRZ) und die Östliche Vulkanische Riftzone (EVRZ) mit ihrer südlichen Verlängerung der Südisländischen Vulkanischen Frontzone (SVFZ) sind hellgelb mit roten Rändern dargestellt. Diese Zonen beinhalten die Calderen der Zentralvulkane (braune Umrisslinien) Hengill (He), Hekla (Hk), Vatnafjöll (Vf), Katla (Ka), Grimsvötn (Gr), Barðarbunga (Ba), Askja (As) und Krafla (Kr) (Einarsson & Sæmundsson 1987, modifiziert). Die Reykjanes Vulkanzone (RVZ) zeichnet sich durch ein schiefes Öffnen aus. Die Snæfellsnes Vulkanzone (SVZ) ist wie Vestmannaeyjar auf der gegenüberliegenden Seite der WVRZ ein alter Abschnitt des Rückens. Die Südisländische Seismische Zone (SISZ) und die Tjörnes Bruchzone (TFZ) sind die seismisch aktivsten Zonen der Insel. Die Gletscher sind hellblau geschummert.

Die *Mittelisländische Vulkanische Riftzone*[51] mit dem Gletscher Hofsjökull im Zentrum koppelt das Nordende der WVRZ an die NVRZ und lockert die Verbindung der Hreppar Mikroplatte zur Nordamerikanischen Platte.
Verschiedene Versuche einer FE-Simulation der tektonischen Bewegungen in Island lieferten allesamt Ergebnisse, die sich nicht in Einklang mit den messbaren Oberflächenbewegungen bringen ließen (Witt 2002). Die neue Erkenntnis über diese vergleichsweise kleine Riftzone hatte noch nicht berücksichtigt werden können und verhinderte damit eine aussagekräftige Modellierung.

[51] Diese Zone ist bereits als *Mittelisländische Vulkanische Zone* und als *Zentralisländisches Transform* bezeichnet worden (Sigmundsson 2006, S. 37). Ihr wird aber bisher keine Aktivität im Spreizungsprozess zugeschrieben.

2.4.2 Die Transformzonen

Die Tjörnes Bruchzone

Im Norden der Insel, größtenteils unter Wasser, befindet sich die dreigliedrige Tjörnes Bruchzone (TFZ), die als Ganzheit eine dextrale Transformbewegung durchführt. Das zentrale tektonische Element bildet der Husavík-Flatey-Graben (HFF) mit einer Länge von etwa 100 km. Jeweils nördlich und südlich befindet sich parallel zum HFF eine geradlinige seismisch aktive Zone. Beiden sogenannten *lineaments* ist zu eigen, dass sie sich nicht durch einfache Spaltenstrukturen an der Oberfläche lokalisieren lassen (Riedel 2001, S. 13). Vielmehr zeichnen sich am besser untersuchten Grímsey-Lineament (Abb. 2.16) *non-transform offsets* ab (Riedel 2001, Abb. 83). Guðmundsson (2007, Kap. 4.3) bringt ihre Entstehung mit der Verlängerung einerseits des Kolbeinseyrückens nach Süden und andererseits der NVRZ nach Norden in Verbindung.

Die Südisländische Seismische Zone

Die WNW–ESE gerichtete Südisländische Transformzone (SITZ) von der Reykjanes Vulkanzone (RVZ) bis zur Südisländischen Vulkanischen Frontzone (SVRZ) liegt an Land. Die zentrale Transformzone, nämlich die Südisländische Seismische Zone (SISZ) weist keine signifikante vulkanische Aktivität auf. Statt aus einer einzigen etwa von Ost nach West gerichteten sinistralen Spalte wird die SISZ aus vielen von Nord nach Süd gerichteten, parallelen, dextralen Spalten gebildet (Einarsson 1991). Damit erreicht diese Zone auf 70 km Länge eine Breite von etwa 10-15 km (Guðmundsson 2007).
Geodätische Messungen gaben den Anlass, den Transformprozess als *bookshelf faulting* (Kap. 2.1.2) zu beschreiben (Sigmundsson et al. 1995; Sigmundsson 2006). Darüberhinaus wird die Lage der SISZ nicht als stabil betrachtet (Guðmundsson 2007). Vielmehr existiert die Annahme, dass diese Zone sich fortwährend mit der Östlichen Vulkanischen Riftzone nach Süden verschoben habe (Einarsson 1991). Weiterhin besteht die Annahme, dass diese Zone sich in eine schiefe Riftzone ähnlich des nördlichen Reykjanesrückens entwickeln wird und dass dieser Prozess an den Vulkanspalten des Hekla bereits begonnen habe (Guðmundsson 2007).
Einige Fakten sprechen jedoch gegen eine Mobilität der Südisländischen Seismischen Zone. So ist der Hreppar Mikroplatte mit Spalten durchzogen deren Hauptrichtung etwa NNE–SSW entspricht. Die Spalten der SISZ sind aber ziemlich genau N–S gerichtet. Unter der Annahme einer konstanten Riftgeschwindigkeit wären 800.000 Jahre notwendig, um zu dieser Rotation zu kommen. Realistischer ist eine Riftgeschwindigkeit, die mit dem Südwärtsstreben der Östliche Vulkanische Riftzone ständig zugenommen hat. Dann wären eher bis zu zwei Mill. Jahren für die Spalten- und damit Blockrotation erforderlich. Vor dieser langen Zeitspanne existierte noch keine dominante Östliche Vulkanische Riftzone (Trønnes 2002, Abb. 2), womit auch eine frühere Existenz einer mobilen SISZ obsolet würde. Es deutet also einiges daraufhin, dass sich die SISZ gleich an der heutigen Stelle manifestiert hat.

Aus der Entwicklung Islands kann vermutet werden, dass Vorläufer der beiden Transformzonen älter sind als die vulkanische Kruste Islands. Sie sind immer wieder aktiviert worden, auch mit einer gegenläufigen Transformbewegung. Ursächlich wäre die anzunehmende kontinentale Kruste unter Island. Dieser These zufolge waren die Gesteine im Untergrund insbesondere am südlichen Rand der kontinentalen Fragmente bereits geschwächt.

Dafür spricht, dass die rezenten Spalten der jüngsten Beben sich überwiegend in den Ergussgesteinen des Holozän befinden (GUÐMUNDSSON 2007). Die relativ schmale SISZ hätte demnach immer am Südrand der kontinentalen Kruste gelegen, die weit aufgefächerte TFZ hingegen in einiger Entfernung nördlich dieser Fragmente.
Für diese Arbeit bedeuten die genannten Umstände, dass die Annahme einer tektonischen Reaktion auf Erdbeben in diesen Zonen zwingend ist.

2.5 Die Kinematik Islands

Das Bewegungsverhalten der Plattengrenze in Island ist bereits seit 1930er Jahren ein Gegenstand der Forschung. Frühere Messkampagnen der Dänischen Landesvermessung erwiesen sich als zu ungenau, um von ihnen Bewegungsraten ableiten zu können. Auf Island existieren verschiedene und historisch gewachsene geodätische Netze, vorzugsweise Lagenetze, die entweder für wissenschaftliche oder für Katasterzwecke von verschiedenen Instituten angelegt, vermarkt und wiederholt vermessen worden sind. Waren diese Netze ursprünglich terrestrische Netze, so sind bereits viele Wiederholungsmessungen mit GPS ausgeführt worden. Hierbei sind in Abhängigkeit des Untersuchungszieles verschiedene Punktgruppen besetzt worden. Aufgrund der seismischen Aktivität auf Island ist die Frage der Gründung der Vermarkung von besonderer Wichtigkeit: damit vermarkte Punkte stabil oder zumindest repräsentativ für die Bewegung des Umlands sind, befinden sie sich auf anstehendem vulkanischen Felsgestein. Eine Gründung in den quartären vulkanischen Sedimenten hätte sehr individuelle und wenig repräsentative Bewegungen zur Folge.
Dieser Umstand führt auf Island zu der unglücklichen Situation, dass sich auf einem Felsen manchmal bis zu drei Vermarkungen unterschiedlicher Institute im Umkreis weniger Meter finden lassen. Trotz dieser unmittelbaren Nachbarschaft ist es in der Vergangenheit oft versäumt worden, dass Exzentrizitätsmessungen durchgeführt worden sind. Erst in der GPS-Messkampagne Südwestisland 1999 ist dieser Mangel größtenteils behoben worden. Die damit möglichen kombinierten Netzlösungen liefern neue Einblicke in das kinematische Verhalten Islands (PERLT & HEINERT 2006; PERLT ET AL. 2008).

2.5.1 Feldarbeiten

Im Folgenden können nicht alle geodätischen Feldarbeiten erwähnt werden. Gerade in Island gibt es derer zu viele. Dennoch soll ein kurzer Überblick über die Arbeiten gegeben werden, die durch die braunschweigischen und hannoveraner Geodäten durchgeführt wurden und solche Arbeiten, die hiermit im Zusammenhang stehen.
Die ersten geodätischen Feldarbeiten mit dem ausgewiesenen geophysikalischen Ziel, präzise Lagekoordinaten und damit die Riftbewegung in Nordisland erfassen zu können, wurden 1938 durchgeführt (NIEMCZYK 1943). Aufgrund des Zweiten Weltkrieges konnten diese Messungen nicht mehr zeitnah wiederholt werden. Erst 1965 gelang es, eine Wiederholungsmessung dieses Lagenetzes durchzuführen (GERKE 1966; HEUMANN 1972). Jedoch auch dieser erste Epochenvergleich war noch nicht zufriedenstellend. In den Folgejahren wurden ausgedehnte terrestrische Deformationsmessungen sowohl im Nordosten (MÖLLER & RITTER 1980) als auch im Südwesten der Insel rund um den Þingvallavatn, Islands größten See, durchgeführt (RITTER 1982; RITTER 1986; MÖLLER 1989). Parallel wurden Nivellements teils als Schleifen um den Þingvallavatn (CZUBIK 1989) oder auch als Querschnitte durch das Þingvellir in der Westlichen Vulkanischen Riftzone durchgeführt (TRYGGVASON 1974).
Im Jahr 1986 entstand das erste über ganz Island ausgeweitete GPS-Netz, das 1989 allerdings ohne die Stationen in der Nördlichen Vulkanischen Riftzone wiederholt wurde (HACKMAN 1991; FOULGER ET AL. 1993). Diese Lücke sollte aber durch die Messungen 1987 und 1990

geschlossen werden (Jahn 1992). Die Isländische Landesvermessung führte in Kooperation 1993 die erste verdichtete GPS-Kampagne über ganz Island aus (Magnusson et al. 1997). Aus ihr entstand das ISNET93.
Sowohl im Bereich der Südisländischen Seismischen Zone (Sigmundsson et al. 1992; Sigmundsson et al. 1995; Sigmundsson et al. 1997; Alex et al. 1999; Perlt & Heinert 2006), als über die Nördliche Vulkanische Riftzone wurden nun immer wieder GPS-Kampagnen durchgeführt (Völksen & Seeber 1998; Völksen 2000). Eine der letzten größeren GPS-Kampagnen in Island war die Neumessung des Grundlagennetzes ISNET2004, das jüngst vollständig ausgewertet und mit der ursprünglichen Grundlagennetzmessung ISNET93 verglichen worden ist (Valsson et al. 2007, Abb. 15).
Einen ausgezeichneten grafischen Überblick über alle diese Feldarbeiten gibt Perlt 2006 (Abb. 3.4–3.9).

2.5.2 Auswertung

Die Grundlage einer jeden Bewegungslösung bilden die Verfahren der Deformationsanalyse (Niemeier 1981). Diese besteht im einfachsten Falle aus einer Transformation der Koordinaten einer wiederholten Messepoche auf die Koordinaten einer sogenannte Nullepoche (Perlt 2006, Kap. 2.4.3). Ein stochastisches Modell erlaubt es, schlecht bestimmte Punkte aus der Analyse herauszuhalten (Pelzer 1985). Während die Deformationen mit dieser Technik im Nordosten der Insel verhältnismäßig günstig zu berechnen waren, erlaubten die besonderen tektonischen Bedingungen im Zusammenspiel mit der damals gewählten Netzkonfiguration im Südwesten erst jüngst eine plausible Herleitung der Bewegungen aus den terrestrischen Messungen (Heinert et al. 2004).

Deformationsanalysen und kinematische Modelle

Ein reiner Zwei-Epochen-Vergleich ist trotz vieler Adaptionen wie die der Hypothesentests (Niemeier 1985a), der Datumspunkt-Objektpunkt-Trennung und einer Vorwärts-Rückwärts-Strategie (Niemeier & Tengen 1988) gegenüber vielerlei Einflüssen anfällig. Insbesondere wenn es sich eben nicht um gut angelegte Ingenieursnetze, sondern vielmehr um Deformationsnetze in tektonisch aktiven Regionen handelt, neigen alle Zwei-Epochen-Analysen zum Versagen[52] (Heinert et al. 2004, Kap. 4).

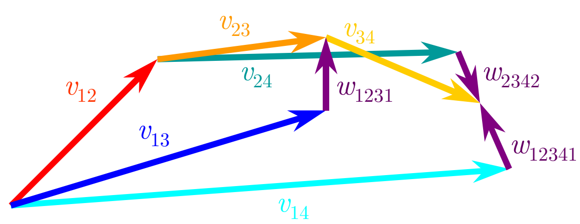


Abb. 2.17: Ansatz einer Mehrepochen-Deformationsanalyse zur Bestimmung nichtlinearer Bewegungen: Minimierung aller v und w (Heinert et al. 2004).

Es gibt grundsätzlich zwei verschiedene Ansätze der Mehrepochendeformationsanalyse. Zum einen solche, welche die Hypothese einer *linearen Bewegung* für Einzelpunkte und Punktgruppen aufstellen (Altamimi et al. 2002). Im Falle für großräumige und über lange Zeit beobachtete Netze ist der Ansatz der konstanten Geschwindigkeit für Einzelpunkte und Punktgruppen berechtigt und auch in tektonisch aktiven Gebieten erfolgreich (Perlt &

[52] Ein Beispiel für eine unglückliche Umsetzung der Zwei-Epochen-Deformationsanalyse in tektonisch aktiven Gebieten ist die großräumig angenommene Radialdeformation um das Vulkansystem der Krafla, das ein erstes nachgewiesenes Wirken eines Manteldiapirs hätte sein sollen. Die Wahl der Datumspunkte im Zentrum des Messgebietes führte zu der vielfach extrapolierten Radialkomponente an den um mehrere hundert Kilometer entfernten Punkten am Rand des Netzes aufgrund eines nur minimalen Netzmaßstabsfehlers (Völksen & Seeber 1998; Perlt et al. 2008).

Heinert 2006; Perlt et al. 2008). Die wesentliche Variation ist die Annahme über das Geschwindigkeitsfeld: Handelt es sich als separat zu betrachtende Blöcke mit erkennbaren Grenzen oder ist wegen einer elastischen Reaktion der Erdoberfläche ein fließender Geschwindigkeitsübergang von Punkt zu Punkt zu definieren.
Hingegen müssen *nichtlineare Bewegungen* bestenfalls ohne Hypothesen über die Geschwindigkeit analysiert werden (Heinert et al. 2004). Erfolgreich kann hier die Oberbedingung sein, dass summierte Bewegungen aus einzelnen Zwei-Epochen-Vergleichen auf das exakt gleiche Ergebnis eines Vergleichs zwischen Anfangs- und Endepoche führen müssen (Abb. 2.17).

2.5.3 Zusammenstellung der Ergebnisse

Die Zonen größter Deformation korrespondieren gut mit den bereits vorgestellten vulkanischen Riftzonen (Kap. 2.4.1). Die Mehrepochendeformationsanalyse über die GPS-Kampagnen von 1986 bis 2002 erbringt in der Nördlichen Vulkanischen Riftzone Deformationen in der Größenordnung, wie sie für die Plattengrenze zwischen NOAM und EURA von den verschiedenen globalen Bewegungsmodellen (Tab. 2.1 und Abb. 2.18, oben, schwarze Vektoren) vorhergesagt werden. Im langjährigen Mittel sind diese Bewegungen mit 23 $^{\mathrm{mm}}/_{\mathrm{a}}$ (Abb. 2.18, unten) hier allerdings etwas größer als es die globalen Bewegungsmodelle vorhersagen (Perlt 2006; Perlt et al. 2008). Diese hohe Bewegung kann als Abklingen der Riftepisode im Krafla Vulkansystem der Jahre 1975 bis 1984 angesehen werden (Sigmundsson 2006, Kap. 6.3).
Das alternative Bewegungsfeld aus der Helmert-Transformation der ISNET2004- auf die ISNET93-Kampagne zeigt insbesondere im Osten der Insel Differenzen zur Mehrepochenlösung (Valsson et al. 2007). Gerade östlich des Gletschers Vatnajökull ist die Diskrepanz in Richtung und Geschwindigkeit am größten (Abb. 2.18, oben, graue Vektoren). Ein Artefakt aus einer Nullpunktsverschiebung der relativen Vektoren ist ausgeschlossen, dafür ist die Übereinstimmung im Westen der Insel zu groß. Die Differenzen spiegeln eher die tatsächliche Veränderung der Bewegung wider. Die Zweiepochenlösung über das ISNET beinhaltet nur die Bewegung ab 1993, als die Krafla-Riftphase als vollständig abgeschlossen gelten kann. Die Mehrepochenlösung schließt aber noch Daten von 1986 mit ein. In dieser Zeit wird noch eine Postrift-Bewegung in die Lösung eingehen. Die Richtungsänderung der Zweiepochenlösung östlich des Vatnajökull deutet an, dass sich die größten Bewegungsraten und der stärkste Richtungsgradient auf eine ansteigende Aktivität der NVRZ zwischen den Zentralvulkanen Askja und Barðarbunga zurückführen lassen.
Für den südlichen Abschnitt der Westlichen Vulkanischen Riftzone lassen sich aber noch Bewegungen von $4 - 8\,^{\mathrm{mm}}/_{\mathrm{a}}$ nachweisen (Heinert et al. 2004; Perlt & Heinert 2006; Perlt et al. 2008). Diese werden nach Norden hin immer kleiner, bis sie nördlich des Langjökull vollständig verschwinden. Zwischen 1983 und 1995 haben sich unerwartet die beiden Randspalten des Zentralgrabens Þingvellir der Westlichen Vulkanischen Riftzone geschlossen (Heinert et al. 2004).

Die relativen Bewegungen zwischen der Nordamerikanischen Platte und dem Hreppar Mikroplatte sind sehr großen Schwankungen unterworfen. Das wird aus den Epochenvergleichen der frühen terrestrischen Messungen um Islands größten See, den Þingvallavatn, deutlich (Abb. 2.19). Demnach ist die Hreppar Mikroplatte möglicherweise eine Ausgleichsfläche für alle tektonischen Ereignisse um sie herum. Die Anbindung an die anderen Platten ist gering: Im Süden durch die rotierenden Blöcke der Südisländischen Seismischen Zone von der Eurasischen Platte weitgehend entkoppelt und im Norden ist die Anbindung an die Nordamerikanische Platte über die Mittelisländische Vulkanische Riftzone hinweg auch eher gering. Diese zeigt nämlich unerwartet eine Divergenzbewegung von bis zu 6 $^{\mathrm{mm}}/_{\mathrm{a}}$. Beide Lösungen sind hier weitgehend identisch. Damit erreicht die Spreizung hier immerhin eine Größenordnung, wie sie in der als aktiv geltenden Westlichen Vulkanischen Riftzone zu finden ist (Abb. 2.18, unten). Lange Zeit ist jedoch davon ausgegangen worden, dass die Zone um den Hofsjökull tektonisch nahezu inaktiv ist.

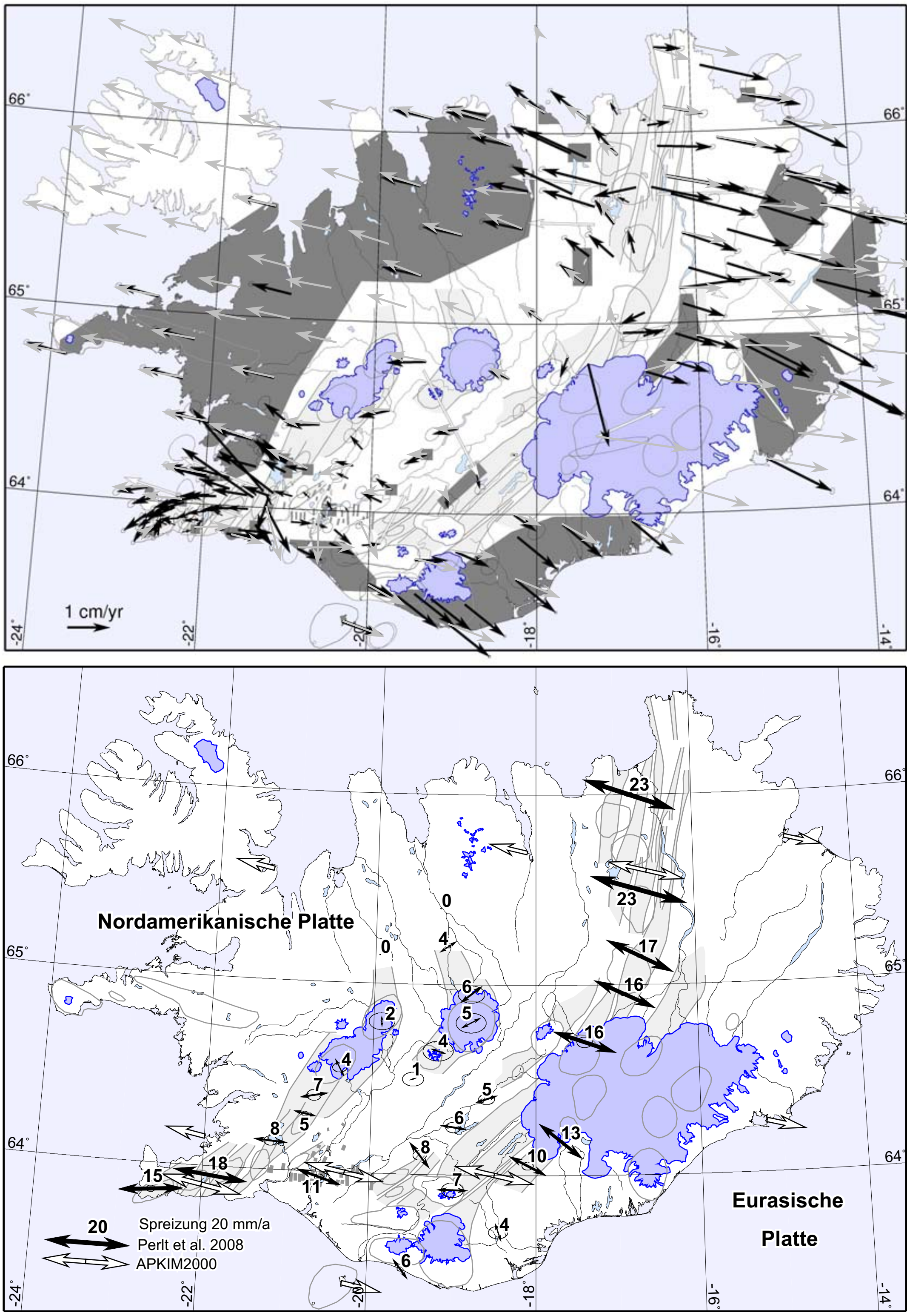


Abb. 2.18: Oben: Bewegungsfeld Islands aus GPS: zum einen aus kombinierten GPS-Lösungen von 1986-2002 (Perlt 2006, schwarze Pfeile) und zum anderen aus der Helmert-Transformation des ISNET2004 auf das ISNET93 (Valsson et al. 2007, hellgraue Pfeile). Unten: Spreizungsraten aus der Differenz zwischen benachbarten Blockbewegungen (Perlt et al. 2008, schwarze Pfeile) und die Lösung aus dem globalen Bewegungsmodell (Drewes & Angermann 2001).

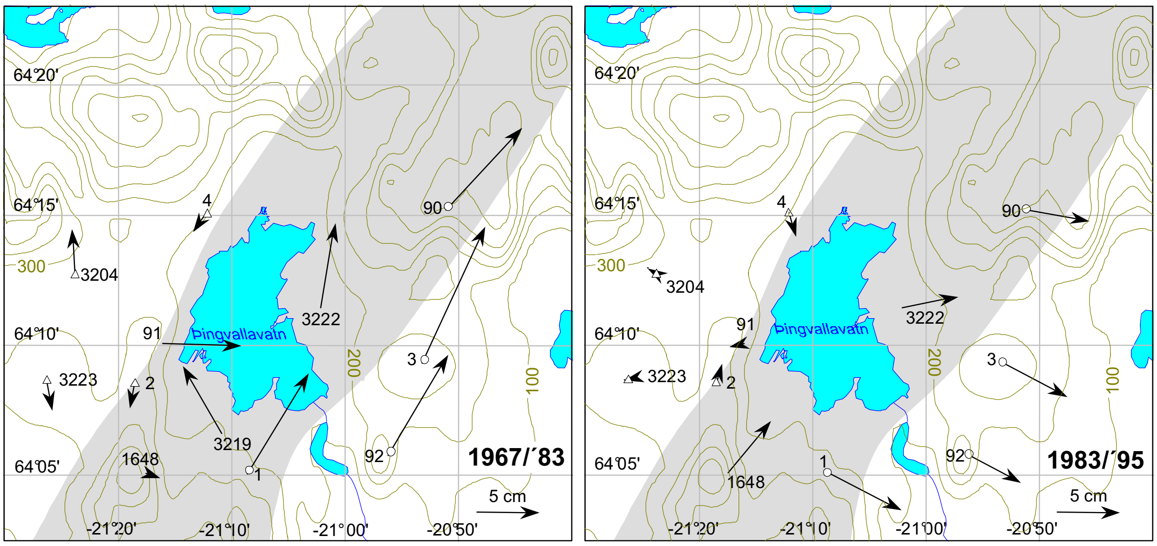


Abb. 2.19: Deformationen in der Westlichen Vulkanischen Riftzone: Wechsel der Bewegung der Hreppar Mikroplatte von Nordnordwest vor 1983 nach Ost nach 1983.

Diese Untersuchung erbringt aber ein weiteres wichtiges Detail: Die Südisländische Transformzone ist *nicht* identisch mit der Südisländischen Seismischen Zone, diese ist nur ein Element jener Transformzone. Erst der westlich des Mýrdalsjökull (Abb. 2.16) gelegene stabile Block führt gegenüber der Nordamerikanischen Platte die vollständige Bewegung der Eurasischen Platte aus (Abb. 2.18), damit bildet die Südisländische Vulkanische Frontzone tektonisch betrachtet ein Element der Transformzone. Gleiches gilt für die Reykjanes Vulkanzone im Westen. Auch sie bildet nach dieser Definition einen Teil der gesamten Transformzone.

2.6 Die Seismizität Islands

In und um Island finden sich entlang der Plattengrenzen sowohl tektonische (Kap. 2.2.1), als auch vulkanische Beben (Kap. 2.2.1). Die Herdflächen der auftretenden Beben reichen daher in Tiefen bis etwa 25 km, nur sehr vereinzelt treten Beben bis an die Grenze der isländischen Kruste auf. Diese befindet sich in einer Tiefe von 15-46 km (Allen et al. 2002a). Die Erdbeben in Island sind daher zumeist von mittlerer Stärke. Das schwerste Erdbeben in historischer Zeit, das tatsächlich gemessen worden ist, war das Ereignis von 1912 in der SISZ mit einer Magnitude M_s 7,0 (Stefánsson et al. 1993, Tab. 1).
Ein immer wieder berichtetes Phänomen sind saisonale Schwankungen in der seismischen Aktivität in genau begrenzten Gebieten Islands. So sind jährliche Periodizitäten im westlichen, also dem der Südisländischen Seismischen Zone (SISZ) zugewandten Teil des Zentralvulkansystems Katla (Goðabunga) im Zeitraum 1978-85 beobachtet worden (Einarsson & Brandsdóttir 2000). Eine vergleichbare temporäre Periodizität der Aktivität ist auch im Kernbereich der SISZ nachgewiesen (Heinert & Perlt 2002; Heinert 2003). Die Flüsse Hvitá, Þjórsá und Markarflót entwässern durch die SISZ etwa ein Drittel der Gletscherflächen von Langjökull, Hofsjökull und Vatnajökull. Je nach Verfügbarkeit von Schmelzwasser von den Gletschern muss vermutet werden, dass der schwankende Porenwasserdruck in die Tiefe der Spalten propagiert und hier als hydraulischer Auslöser für Erdbeben wirkt[53]. Es ist daher zu vermuten, dass sich gegebenfalls zumindest temporär eine schwache Periodizität in der Seismik in dieser Zone ergeben wird (Zschau 2002).

[53] Auch wenn diesbezüglich für Island noch keine Untersuchungen publiziert sind, so zeigen Ergebnisse aus anderen Gebieten eindrucksvoll, dass diese Korrelationen existieren müssen (Hainzl et al. 2006; Kraft et al. 2006; Miller 2008).

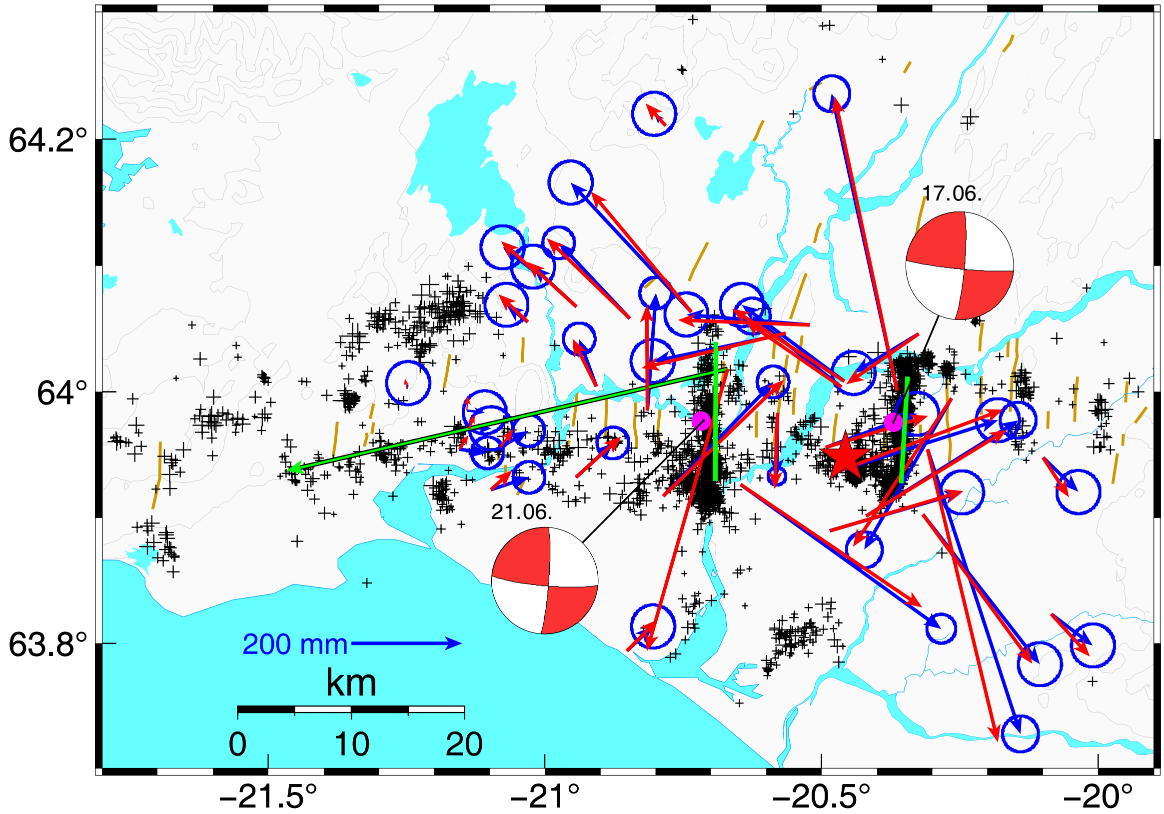


Abb. 2.20: Die Erdbeben des Jahres 2000 in der Südisländischen Seismischen Zone mit ihren Hauptverwerfungen (grün), den modellierten (rot) und den gemessenen koseismischen Bewegungen (blau), Nachbeben (schwarze Kreuze) bis zum 22.11.00 einschließlich des schwersten einzelnen Nachbebens (roter Stern). Eine GPS-Station befand sich unmittelbar in der Nähe der oberflächennahen Verschiebungen während des zweiten Bebens (grüner Pfeil) (Árnadóttir et al. 2001).

Von besonderer Bedeutung sind die Erdbeben des 17. und 21. Juni 2000 in der Südisländischen Seismischen Zone (Stefánsson et al. 2000). Es ist ein besonderer Umstand, dass diese Erdbeben erwartet worden sind (Stefánsson et al. 1993) und dass daher schon im Vorfeld diese Region mehrfach vermessen worden ist (Alex et al. 1999), was dann eine umfangreiche Analyse der Bewegungen der Beben ermöglichte (Árnadóttir et al. 2001).
Im Rahmen dieser Arbeit spielen diese Beben eine andere Rolle: Zum einen tragen sie ein massives postseismisches Signal in die Bewegungen der umliegenden GPS-Permanentstationen. Zum anderen sind diese Beben der Abschluss eines physikalischen Prozesses, der schon lange unerkannt beobachtet worden ist und erst jetzt entschlüsselt werden kann.

3 Systemtheoretische Grundlagen

„Causa aequat effectum,
nil fit ad nihilum,
nil fit ex nihilo.“
Julius Robert von Mayer

3.1 Systeme

Das Wort System hat seinen Ursprung im griechischen Wort $\sigma\acute{\upsilon}\sigma\tau\eta\mu\alpha$ und bedeutet *ein aus vielen Teilen vereinigtes Ganzes*. Dieser Bedeutung folgend fassen Zadeh und Polak 1969 und später Ogata 1978 die fundamentalen Aussagen über Systeme zusammen: die allgemeine Systemtheorie befasse sich mit der Untersuchung von Mengen aus *Elementen*. Jedoch stellt ein System mehr dar als nur eine Menge im mathematischen Sinne. Vielmehr beschreibt der Systembegriff, inwieweit Elemente in einer Beziehung stehen und sich durch die Art ihrer Beziehung zusammenfassen lassen.

Ein System lässt sich weiterhin durch eine *Systemoberfläche* (Abb. 3.1) gegen seine Umgebung abgrenzen[1]. Häufig stellt diese Oberfläche je nach Art des Systems auch eine Beschränkung der Untersuchungsmöglichkeiten dar. Aus diesem Grund ist es oft nur möglich, einen Teil des Systemeinganges bzw. Systemausganges, also die Wechselwirkungen eines Systems mit seiner Umwelt zu erfassen, während die systeminternen Relationen zwischen den Systemelementen nicht selten verborgen bleiben (Zadeh & Polak 1969, S. 4). Diese internen Beziehungen können fernerhin dergestalt sein, dass sich auch innerhalb eines Systems Elementgruppen ergeben, welche sich ihrerseits gegen die übrigen Elemente abgrenzen lassen. Hierbei spricht man von einem *Subsystem*. Eine Systemoberfläche grenzt das System jedoch nicht nur räumlich und inhaltlich, sondern auch zeitlich ein.

Unterhalb der Systemoberfläche findet eine zeitliche Veränderung des Systems von seiner Entstehung bis zu seinem Zusammenbruch statt. Eine grundlegende Forderung der Systemtheorie besteht deshalb darin, ein System nicht nur in einem Zustand, sondern auch in seiner zeitlichen Veränderung zu untersuchen.

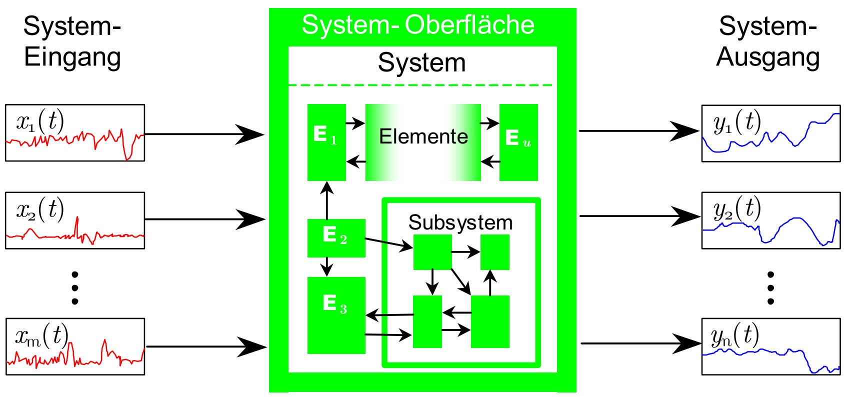


Abb. 3.1: Darstellung eines Systems mit Systemein- und -ausgängen, vergl. (Unbehauen 2002, S. 5).

[1] Diese Eigenschaft wird häufig durch den Begriff *Systemgrenze* ausgedrückt. Um aber die gelegentliche Unschärfe dieser Grenze zu verdeutlichen, soll im folgenden von einer Oberfläche gesprochen werden.

3.1.1 Systemeigenschaften

Ein physisches System lässt sich durch drei Komponenten *Systemeingang*, *Systemausgang* und *Systemübertragungseigenschaften* mathematisch vollständig beschreiben (NATKE 1992, Kap. 1.1). In der Praxis müssen die *Systemprozesse* hierzu zunächst *realisiert* werden (Kap. 3.2). Eine Realisierung im Sinne einer *Erfassung, Messung, Beobachtung* oder *Stichprobe* kann das System und seine Prozesse nur unter den vorhandenen technischen Voraussetzungen oder entsprechend der Qualitäten des Beobachters abbilden. Demzufolge geht die Art dieser Abbildung in das Realisierungsresultat mit ein.
Die erste und in der Geodäsie zumeist bekannte Komponente ist der realisierte zeitvariante Systemausgang $y(t)$. Eine Zustandsgeometrie und ihre Veränderung wird an der Systemoberfläche erfasst und ist damit Bestandteil des Systemzustands und seiner Veränderung. Die zweite realisierbare Komponente ist der zeitvariante Systemeingang $x(\tau)$. Dabei müssen τ und t nicht zwingend einem gleichen Intervall $\mathbb{T}$ entstammen. Die dritte Komponente besteht aus den Übertragungseigenschaften des Systems, die durch den Transformationsoperator $\mathcal{T}[\,\cdot\,](t)$ ausgedrückt werden. Damit lässt sich ein physisches System in der Form

$$\{y_i(t) : i = 1, \ldots, n; t \in \mathbb{T}_2\} = \mathcal{T}\left[\{x_i(\tau) : i = 1, \ldots, m; \tau \in \mathbb{T}_1\}\right](t), t \in \mathbb{T}_2 \tag{3-1}$$

vollständig beschreiben (MINKLER & MINKLER 1993, S. 71). Unter der Beachtung der beiden Intervalle $\mathbb{T}_1$ und $\mathbb{T}_2$ soll (3-1) verkürzt als

$$\mathbf{y}(t) = \mathcal{T}\left[\mathbf{x}(\tau)\right](t) \tag{3-2}$$

ausgedrückt werden (UNBEHAUEN 2002, S. 7).

3.1.2 Lineare Systeme

Ist die Summe über die m Systemantworten von einzelnen Systemeingängen gleich dem Systemausgang aus der Summe der m Systemeingänge

$$\sum_{i=0}^{m} \mathcal{T}\left[x_i(\tau)\right](t) = \mathcal{T}\left[\sum_{i=0}^{m} x_i(\tau)\right](t), \tag{3-3}$$

dann erfüllt das System die Voraussetzung der *Additivität*[2]. Das System ist weiterhin *homogen*, wenn die Multiplikation des Systemeingangs mit α zum gleichen Ergebnis führt wie die Multiplikation des Systemausgangs mit α:

$$\mathcal{T}\left[\alpha x(\tau)\right](t) = \alpha\mathcal{T}\left[x(\tau)\right](t). \tag{3-4}$$

Sind (3-3) und (3-4) zugleich erfüllt und lässt sich zeigen, dass das System auf die Nullfolge $\mathbf{x}(t) \rightarrow \mathbf{0}$ mit einer Nullfolge des Systemausgangs reagiert (UNBEHAUEN 2002, S. 8), so gilt die erweiterte Linearitätsbedingung mit $m \rightarrow \infty$:

$$\sum_{i=0}^{\infty} \alpha_i\mathcal{T}\left[x_i(\tau)\right](t) = \mathcal{T}\left[\sum_{i=0}^{\infty} \alpha_i x_i(\tau)\right](t). \tag{3-5}$$

[2] Die Additivität wird auch als *Superposition* bezeichnet (KOCH & SCHMIDT 1994; HEINE 1999; VAN BRUSSEL 2002).

3.1.3 Inkremental lineare Systeme

Auch wenn der Vektor des Systemausgangs $\mathbf{y}(t)$ durch eine lineare Gleichung

$$\mathbf{y}(t) = \mathcal{T}\left[\mathbf{x}\right](t) = \alpha\mathbf{x}(t) + \mathbf{b} \;\forall\; b \neq 0 \tag{3-6}$$

gebildet wird, ist das resultierende System nichtlinear. Es ist einfach ersichtlich, dass die Homogenitätsbedingung für $\mathcal{T}[\mathbf{0}] = \mathbf{0}$ aus (3-4) nicht erfüllt sein kann (Minkler & Minkler 1993, S. 75f). Dennoch reagiert dieses System auf Änderungen des Systemeingangs $\mathbf{\Delta x}(\tau)$ mit der linearen Änderung der Systemantwort

$$\mathbf{\Delta y}(t) = \mathcal{T}\left[\mathbf{x}(\tau) + \mathbf{\Delta x}(\tau)\right](t) - \mathcal{T}\left[\mathbf{x}(\tau)\right](t). \tag{3-7}$$

3.1.4 Stationäre und statische Systeme

Wenn der gleiche Systemeingang um ein beliebiges reelles t_0 verschoben[3] werden kann und sich gemäß des Ausdrucks

$$\mathcal{T}\left[\mathbf{x}(\tau - t_0)\right](t) = \mathbf{y}(t - t_0) \;\forall\; t_0 \in \mathbb{R} \tag{3-8}$$

ein gleicher Systemausgang an der korrespondierenden, um t_0 verschobenen Stelle ergibt, so ist das System *zeitinvariant* oder *stationär* (Minkler & Minkler 1993, S. 76f). Für eine Analyse des Systems ist es zunächst nur erforderlich, dass das System diese Bedingung auf den Intervallen $\tau - t_0 \in \mathbb{T}_1$ und $t - t_0 \in \mathbb{T}_2$ erfüllt. Ein System wird als *statisch*[4] bezeichnet, wenn es gedächtnislos nur auf den unmittelbar zeitgleichen Systemeingang $\mathbf{x}(t - t_0)$ reagiert (Unbehauen 2002, S. 11f):

$$\mathcal{T}\left[\mathbf{x}(t - t_0)\right](t) = \mathbf{y}(t - t_0) \;\forall\; t_0 \in \mathbb{R}. \tag{3-9}$$

Diese Eigenschaft lässt sich dahingehend umformulieren, dass auf einen gleichen Systemeingang auch der gleiche Systemausgang folgen muss (Van Brussel 2002, S. 1.12):

$$\mathbf{x}(t - t_0) = \mathbf{x}(t) \;\Rightarrow\; \mathbf{y}(t - t_0) = \mathbf{y}(t) \;\forall\; (t - t_0) < t \in \mathbb{T}. \tag{3-10}$$

3.1.5 Dynamische und kausale Systeme

Bei einem *dynamischen* System ist (3-10) nicht für alle $(t - t_0) < t$ erfüllt. Hat das System die endliche Gedächtnislänge Δt, so gilt die engere Stationaritätsbedingung (3-10) nur, wenn $t_0 > \Delta t$ (Zadeh & Desoer 1963, S. 45f). Bei einem *unendlichen* Gedächtnis des Systems ist (3-10) nie erfüllt[5]. Ein dynamisches System ist *kausal*, wenn der Systemausgang $y(t)$ zwar aus einem $\mathcal{T}[x(t - t_0)]$ folgen kann, aber sich kein Systemausgang $y(t - t_0)$ auf ein beliebiges $\mathcal{T}[x(t)]$ zurückführen lässt (Minkler & Minkler 1993, S. 78f).

3.1.6 Systemantworten

Ein System werde mit vereinheitlichten Systemeingängen angeregt. Die Systemantwort zeigt sich im Systemausgang und gibt damit wesentliche Informationen über ein System preis.

---

[3] Daher wird auch der Begriff *verschiebungsinvariant* verwendet (Koch & Schmidt 1994, S. 98).

[4] Die Bezeichnung *statisches* System wird meistens als Synonym für ein *stationäres* System verwendet.

[5] Systeme werden aufgrund ihrer Gedächtnislänge auch in FIR- (*finite impulse response*) und in IIR- (*infinite impulse response*) unterteilt (Unbehauen 2002, S. 41f).

Die Sprungantwort

Gegeben sei hierzu die reelle Funktion

$$r(t)_\infty = \begin{cases} 0 & \forall\, t < 0 \\ \frac{t}{\varepsilon} & \forall\, 0 \leq t < \varepsilon \\ 1 & \forall\, t \geq \varepsilon \end{cases} . \tag{3-11}$$

mit unendlicher Ausdehnung (Unbehauen 2002, S. 37). Mit Hilfe dieser Funktion ergibt sich die *Sprungfunktion*

$$s(t) = \lim_{\varepsilon \to 0} r(t)_\infty \text{mit} \int_{-\infty}^{\infty} s(t)\partial t = \infty, \tag{3-12}$$

deren Anstieg von 0 auf 1 entlang der Richtung der Abszissenachse infinitisimal klein ist. Nach dem Grenzübergang handelt es sich nicht mehr um eine reelle Funktion, sondern um eine *reguläre Distribution*[6]. Die *Sprungantwort* $h(t)$ eines Systems ergibt sich nach (3-1) als

$$h(t) = \mathcal{T}\left[s(t)\right]. \tag{3-13}$$

Die Impulsantwort

Die Rechteckfunktion mit der infinitisimalen Breite ε kann als die Differenz zweier unendlicher Sprungfunktionen

$$\begin{aligned} \Delta s(t)_\varepsilon &= \frac{1}{\varepsilon}\left(s(t)_\infty - s(t+\varepsilon)_\infty\right) \qquad (3\text{-}14) \\ &= \begin{cases} 0 & \forall\, t < 0 \\ \frac{1}{\varepsilon} & \forall\, 0 \leq t < \varepsilon \\ 0 & \forall\, t \geq \varepsilon \end{cases} \end{aligned}$$

angesehen werden. Mit Hilfe dieser Rechteckfunktion ergibt sich die *Deltafunktion* (Unbehauen 2002, S. 15), die streng genommen als *Deltadistribution*

$$\delta(t) = \lim_{\varepsilon \to 0} \Delta s(t)_\varepsilon \text{ mit} \int_{-\infty}^{\infty} \delta(t)\partial t = 1, \tag{3-15}$$

zu bezeichnen ist und deren Ausdehnung entlang der Richtung der Abszissenachse infinitisimal klein und parallel zur Ordinatenachse unendlich groß ist. Nach dem Grenzübergang handelt es sich nicht mehr um eine reelle Funktion, sondern um eine *irreguläre Distribution* und wird auch als Dirac-*Impuls* bezeichnet. Für praktische Anwendungen wird der Impuls häufig auf den Wert eins normiert. Die *Impulsantwort* $h(t)$ eines Systems ergibt sich nach (3-1) als

$$h(t) = \mathcal{T}\left[\delta(t)\right] \tag{3-16}$$

(Koch & Schmidt 1994, S. 98f). Ein bisher unbekannter Systemoperator $\mathcal{T}[\,\cdot\,]$ wird durch die Impulsantwort $h(t)$ vollständig charakterisiert (Marko 1995, S. 60f).

[6] Der Begriff der *Distribution* stelle in dieser Arbeit im Unterschied zum Begriff der statistischen *Verteilung* eine Verallgemeinerung der reellwertigen Funktion dar. Eine reguläre Distribution lässt sich durch Funktionen geschlossen darstellen. Die Sprungfunktion ist auch durch die Heaviside-Funktion definiert (vergl. Tab. 4.1, Schwellwertfunktion). Eine irreguläre Distribution ist nicht mehr durch Funktionen darstellbar.

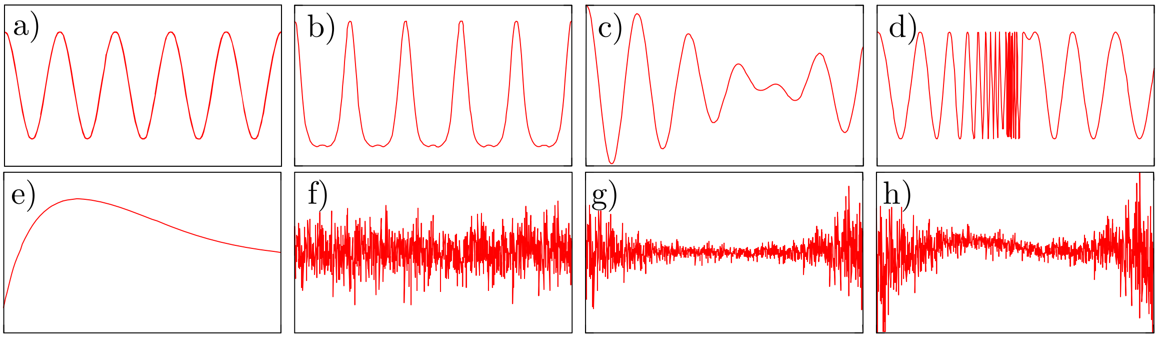


Abb. 3.2: Vergleich verschiedener Prozesse: a) harmonisch, b) periodisch, c) fastperiodisch, d) quasiperiodisch, e) transient, f) schwach stationär, g) mittelwertstationär und h) instationär.

3.2 Prozesse

Die Gesamtheit der Umformung bzw. der Transport von Materie, Energie, sowie Information, welche in einem System stattfindet, wird als Prozess bezeichnet (DIN 19226-1:02/1994). Bei der physikalischen *Realisierung* eines Prozesses wird häufig der Begriff des *Signals* verwendet (Koch & Schmidt 1994; Marko 1995; Unbehauen 2002). Ist ein Prozess zeitabhängig und wird er diskret erfasst, so spricht man klassisch von einer *Zeitreihe* (Natke 1992; Schlittgen & Streitberg 1997). Zur Beobachtung eines Prozesses erfolgt in der Praxis häufig die epochale, also in diskreten Zeitabständen vorgenommene Erfassung verschiedener Systemein- und -ausgänge (Unbehauen 2002). Der nächste Schritt einer Betrachtungsweise des Systems stellt der Übergang von epochalen zu quasikontinuierlichen Erfassungen dar. Erfassungen können als kontinuierlich angesehen werden, wenn sie hinreichend dicht diskretisiert worden sind. Bei dem Übergang eines allgemeinen Systems zu einem physischen System wird aus der Erfassung eine Messung[7]. Damit ist neben der Detektion von Systemrelationen die Analyse von Prozessen in ihrer zeitlichen Abhängigkeit entscheidend für die Untersuchung eines Systems. Physikalische Prozesse lassen sich in verschiedene deterministische und stochastische *Subprozesse* unterteilen, deren Detektion jeweils eine eigene Vorgehensweise verlangt. Die Analyse eines Subprozesses wird im Allgemeinen nur durch Beseitigung der jeweils anderen Subprozesse ermöglicht (Schlittgen & Streitberg 1997).

3.2.1 Deterministische Prozesse

Gehorchen Systembeziehungen kausalen Gesetzmäßigkeiten, sind die resultierenden Prozesse deterministisch. Es existieren demnach funktionale Zusammenhänge, die sich eindeutig beschreiben lassen. Ein deterministischer Prozess ist vollständig vorhersagbar und reproduzierbar (Natke 1992, S. 11).

Periodische Prozesse

Eine entscheidende Rolle kommt in der Praxis den periodischen Prozessen zu. Ist der Prozess $X(t)$ abhängig von der Zeit t, so finden sich in seinen Realisierungen häufig sogenannte Tages- oder Jahresgänge[8]. Solche und andere periodische Prozesse lassen sich in *harmonische*

[7] Ein allgemeines System kann beispielsweise aus der Gesamtheit der Kommunikation innerhalb eines menschlichen Sozialgefüges bestehen. Mit *Mess*methoden der Ingenieurswissenschaften können Systemein- und -ausgänge hier nicht abgebildet werden (Luhmann & Baecker 2002).

[8] Man spricht hier auch von *diurnalen* bzw. *saisonalen* Prozessen.

und *nichtharmonische* Prozesse[9] weiter untergliedern. Ein *harmonischer Prozess* lässt sich vollständig durch *eine* Dreiecksfunktion beschreiben (Abb. 3.2a). Der harmonische Prozess $x(t)$ habe eine Amplitude c und schwinge mit der Frequenz ν und der Phasenverschiebung t_0 (Natke 1992, S. 12). Der Prozess lässt sich weiterhin als eine komplexe Exponentialfunktion oder auch als eine komplexe Summe darstellen (Merziger & Wirth 1993, S. 79ff).

$$\begin{aligned} x(t) = &\ c \cdot \cos\left(2\pi\nu\left(t - t_0\right)\right) & \\ = &\ c \cdot \sin\left(\omega t - \phi + \tfrac{\pi}{2}\right) & \text{mit } \omega = 2\pi\nu \text{ und } \phi = \omega t_0 \\ = &\ c \cdot e^{i(\omega t - \phi)} & \\ = &\ c \cdot \left(\cos(\omega\left(t - t_0\right)) + i \cdot \sin(\omega\left(t - t_0\right))\right) & \text{mit } e^{i\alpha} = \cos\alpha + i \cdot \sin\alpha \\ = &\ a \cdot \cos(\omega t) + b \cdot i \cdot \sin(\omega t) & \text{mit } a \perp b \text{ und } c^2 = a^2 + b^2 \end{aligned} \tag{3-17}$$

Mit Hilfe des Additionstheorems

$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) \tag{3-18}$$

ergibt sich für die Parameter

$$a = c \cdot cos(\phi) \text{ und } b = c \cdot sin(\phi). \tag{3-19}$$

Für die Phasenverschiebung lässt sich damit nachweisen, dass

$$2\pi\nu t_0 = \phi = \arctan\left(\frac{b}{a}\right) = \arctan\left(\frac{\Im(c)}{\Re(c)}\right) \text{ mit } c \in \mathbb{C}. \tag{3-20}$$

Ein periodischer, aber *nichtharmonischer Prozess* kann jeder beliebige, mit der Differenz τ aus t_{k+1} und t_k streng wiederkehrende Prozess sein (Abb. 3.2b). Ein solcher Prozess lässt sich als die unendliche Summe

$$x(t) = \sum_{k=0}^{\infty} \left(a_k \cos(\omega_k t) + b_k \sin(\omega_k t)\right) \text{ mit } b_0 = 0 \text{ und } \omega_k = 2\pi\nu_0 k,\ k \in \mathbb{Z} \tag{3-21}$$

darstellen (Butz 2000, S. 17).

Transiente Prozesse

Ein Prozess, der sich funktional beschreiben lässt, aber *aperiodisch* ist, bezeichnet man als *transient*. Ein transienter Prozess kann beispielsweise durch ein Taylor-Polynom n-ter Ordnung, eine Exponentialfunktion oder logarithmische Funktion gebildet werden (Abb. 3.2e). Eine spezielle Untergruppe der Exponentialfunktionen bildet die der entarteten Schwingungen, der sogenannte *Kriechfall* oder auch der *aperiodische Grenzfall* (Koch & Schmidt 1994, S. 3). Besondere Bedeutung kommt den transienten Prozessen bei der Schätzung von Trendfunktionen zu.

Quasiperiodische Prozesse

Ist die Frequenz eines Prozesses nicht konstant und bildet ihre Folge für sich genommen einen transienten Prozess, so handelt es sich bei jenem um einen *quasi-periodischen Prozess*

$$x(t) = \sum_{k=0}^{\infty} \left(a_k(t)\cos(\omega_k(t) \cdot t) + b_k(t)\sin(\omega_k(t) \cdot t)\right). \tag{3-22}$$

[9] Man spricht hier auch von *sinusoiden* und *nichtsinusoiden* Prozessen.

Eine häufig auftretende Gruppe quasiperiodischer Prozesse ist an der sogenannten Doppler-Verschiebung zu erkennen (Abb. 3.2d)[10]. Ein Spezialfall des quasiperiodischen Prozesses ist der *fast-periodische Prozess*[11]. Seine beteiligten Einzelperioden sind konstant. Durch die Überlagerung von Perioden, die nicht das ganzzahlige Vielfache k einer Grundfrequenz sind (vergl. 3-21), stellen sich Amplitudenvariationen ein, die als *Schwebung* bekannt sind (Abb. 3.2c).

$$x(t) = \sum_{k=0}^{\infty} \left(a_k \cos(\omega_k t) + b_k \sin(\omega_k t)\right) \text{ mit } b_0 = 0 \text{ und } \omega_k = 2\pi\nu_0 k,\ k \in \mathbb{R}. \tag{3-23}$$

Folgt der Betrag der komplexen Amplitude c einer Exponentialfunktion, so handelt es sich um einen *gedämpften* oder *erregten periodischen Prozess* (NATKE 1992, S. 17ff).

$$c(t) = c_0 \cdot e^{-a \cdot t} = \sqrt{a_0^2 + b_0^2} \cdot e^{-a \cdot t} \text{ und } \omega(t) = \text{const.} \tag{3-24}$$

Entscheidend für die meisten quasiperiodischen Prozesse ist, dass sie sich einer Analyse periodischer Prozesse weitgehend entziehen.

3.2.2 Stochastische Prozesse

Da sich unter praktischen Messbedingungen deterministische Prozesse als solche nicht realisieren lassen oder schon von vornherein nicht als rein deterministische Prozesse existieren, muss noch eine weiterer grundlegender Prozesstyp definiert werden. Dieser Prozess muss anstelle einer reinen deterministischen Systemantwort eine Zufallskomponente zulassen. Ein solcher *stochastischer Prozess* lässt sich folglich nicht mehr durch funktionale Zusammenhänge ausdrücken. Er ist vielmehr eine Reihe von Zufallsereignissen, die als Menge $X(t)$ von Zufallsvariablen in Abhängigkeit eines Parameters t dargestellt werden können (DIN 18709-5:Entwurf-2005). Der so bezeichnete Parameter ist in den meisten Anwendungsfällen entlang des Ortes oder der Zeit definiert. Da ein stochastischer Prozess keinen funktionalen Zusammenhängen folgt, ist er an keiner Stelle t vollständig vorhersagbar und lässt sich nicht reproduzieren. Ein stochastischer Prozess kann aber mittels der Zentralmomente seiner Verteilungsfunktion beschrieben werden. Ein einzelnes Ereignis kann damit zwar nicht exakt vorhergesagt werden, aber es lässt sich jedoch eine Wahrscheinlichkeit für sein Eintreten quantitativ schätzen.

Stationäre Prozesse

Ein statisches (3-10) oder stationäres System (3-8) erzeugt stationäre Prozesse, sofern der Systemeingang seinerseits stationär ist. Der Umkehrschluss, dass ein stationärer Prozess die Folge eines stationären Systems ist, gelingt schon aus der Stationaritätsforderung des Systemeinganges nicht. Abweichend vom Stationaritätsbegriff bei einem System, wird ein Prozess als *stationär* bezeichnet, wenn er sich als unveränderlich gegenüber der Zeit oder anderen Grundparametern, etwa den Dimensionen des Raumes, erweist.
Bleiben wir aber bei der eingeschränkten Betrachtung in der Zeit, denn jene entspricht den meisten praktischen Aufgaben. Eine Realisierung eines stationären Prozesses kann zu einem beliebigen Zeitpunkt t_0 beginnen und auch nach einer hinreichenden Beobachtungszeit wieder enden, ohne dass wesentliche Informationen über den Prozess unerfasst bleiben. Aus diesem Grund ist die Feststellung von verschiedenen stationären Eigenschaften eines diskret realisierten Prozesses $x(t)$ von entscheidender Bedeutung für seine spätere Analyse. Die erste

[10] Ein vielen Computer-Nutzern bekannte Anwendung eines Such-Algorithmus nach quasiperiodischen Prozessen ist das als Bildschirmschoner weit verbreitete Programm SETI@home.

[11] Die Definitionen von fastperiodischen und quasiperiodischen Prozessen werden in der Literatur sowohl widersprüchlich als auch synonym gebraucht (NATKE 1992, S. 17ff).

Grundeigenschaft ist die *Mittelwertstationarität* (SCHLITTGEN & STREITBERG 1997, S. 100). Sie besagt, dass ein Prozess (KOCH & SCHMIDT 1994, S. 167) einen Erwartungswert

$$\mathrm{E}\left\{x\left(t\right)\right\}=\mu_x=\text{const.} \tag{3-25}$$

besitzt, der über die Zeit konstant ist und damit auch über den Zeitraum der gesamten Realisierung geschätzt werden kann (Abb. 3.2g). Die zweite Grundeigenschaft ist die *Varianzstationarität* eines Prozesses. Diese besagt, dass sein Streuverhalten des Prozesses um seinen Erwartungswert als erstes Moment[12] über die Zeit konstant bleibt.

$$\mathrm{E}\left\{\left(x\left(t\right)-\mu_x\right)^2\right\}=\sigma_x^2=\text{const.} \tag{3-26}$$

Die dritte Grundeigenschaft ist die *Kovarianzstationarität*. Sie besagt, dass ein Prozess als zweites Moment[13] einen Erwartungswert

$$\mathrm{E}\left\{x^2\left(t\right)\right\}=\mathrm{E}\left\{\left(x\left(t_1\right)-\mu_x\right)\cdot\left(x\left(t_2\right)-\mu_x\right)\right\}=\sigma_{x_1x_2}^2=r_{x_1x_2}\sigma_{x_1}\sigma_{x_2}=\text{const.} \tag{3-27}$$

besitzt, der über die Zeit konstant ist und damit auch über den Zeitraum der gesamten Realisierung geschätzt werden kann. Ein kovarianzstationärer Prozess ist immer auch varianzstationär, denn es gilt

$$\sigma_x^2=\mathrm{E}\left\{\left(x\left(t\right)-\mu_x\right)^2\right\}=\mathrm{E}\left\{\left(x\left(t_1\right)-\mu_x\right)\cdot\left(x\left(t_2\right)-\mu_x\right)\right\}\ \forall\ t_1=t_2. \tag{3-28}$$

Ein Prozess, der sowohl die Kriterien der Mittelwert- als auch der Kovarianzstationarität erfüllt, wird als *stationär im weiteren Sinne*[14] bezeichnet (Abb. 3.2f).
Ein stochastischer Prozess ist *stationär im engeren Sinne*, wenn *alle* seine stochastischen Eigenschaften invariant gegenüber der Zeit sind. Das bedeutet, dass alle Verteilungs- und Verbundverteilungsdichten nicht vom Zeitnullpunkt der Erfassung, sondern ausschließlich von der Zeitdifferenz τ abhängig sein dürfen (UNBEHAUEN 2002, Kap. I § 4.2). Das heißt, dass alle Momente höherer Ordnung

$$\mathrm{E}\left\{x^3\left(t\right)\right\}=const.\ \ldots\ \mathrm{E}\left\{x^n\left(t\right)\right\}=\text{const.} \tag{3-29}$$

ebenfalls konstant sein müssen (NATKE 1992, vergl. Kap. 3.2.1). Ein stationärer Prozess kann auf einem periodischen Prozess (3-21) aufgebaut werden, denn er erfüllt in einem unendlichen Intervall die Bedingung der Stationarität. Unter der Annahme, dass die Parameter a und b den Erwartungswert null besitzen und unkorreliert sind, kann der stationäre stochastische Prozess

$$x(t)=\sum_{k=0}^{n}\left(a_k\cos(\omega_k t)+b_k\sin(\omega_k t)\right) \tag{3-30}$$

gebildet werden (SCHLITTGEN & STREITBERG 1997, S. 155).

$$\text{mit } \mathrm{E}(a)=\mathrm{E}(b)=0,\ \sigma_a=\sigma_b \text{ und } \mathrm{E}(a\cdot b)=0.$$

---

[12] Die Begriffe *Zentralmoment erster Ordnung* (TAUBENHEIM 1969, S. 42) und *Moment erster Ordnung* (SCHLITTGEN & STREITBERG 1997, S. 497) werden synonym gebraucht. Abweichend davon, sowie vom gelegentlichen umgangssprachlichen Gebrauch, bezeichnet der Begriff des Prozesses *n-ter Ordnung* einen Prozess, der aus n skalaren Prozessen besteht (PELZER 1985, S. 269).

[13] Sowohl die Varianz, als auch die Kovarianz werden in der Literatur voneinander abweichend als *Zentralmoment zweiter Ordnung* definiert. Vergl. hierzu TAUBENHEIM (1969, S. 42) und SCHLITTGEN & STREITBERG (1997, S. 497) mit KOCH & SCHMIDT (1994, S. 167).

[14] Der Begriff *stationär im weiteren Sinne* ist gleichbedeutend mit *schwach stationär* (KOCH & SCHMIDT 1994, S. 166).

Für die Wahrscheinlichkeit P jeder Frequenz ν dieses Prozesses gilt

$$\mathrm{P}\{\nu\} = \text{const.} \tag{3-31}$$

In Anlehnung an die Optik wird ein solcher Prozess als *weißes Rauschen*[15] bezeichnet, denn wie beim weißen Licht ist jede Frequenz des optischen Spektrums an diesem Prozess beteiligt. Für die Praxis ergeben sich einige wichtige Spezialfälle stationärer Prozesse. Folgt ein Prozess zu jedem Zeitpunkt t der GAUSS*schen Normalverteilung*

$$x(t) \sim N(\mu, \sigma^2) \tag{3-32}$$

und gilt für das zweite Zentralmoment nach (3-27)

$$\mathrm{E}\left\{(x(t_1) - \mu_x) \cdot (x(t_2) - \mu_x)\right\} = 0 \;\forall\; t_1 \neq t_2 \tag{3-33}$$

so spricht man von einem stationären GAUSS-Prozess. Die Bedingung (3-33) besagt, dass alle Einzelereignisse dieses Prozesses voneinander stochastisch unabhängig sind.
Ist hingegen (3-33) nicht erfüllt und es existiert eine stochastische Abhängigkeit eines Einzelereignisses zum jeweils vorangegangenen, ausgedrückt durch die Wahrscheinlichkeit

$$\mathrm{P}\left\{X(t_{n+1}) \geq x_{n+1} | X(t_n) = x(t_n)\right\} \;\; \forall\; x_{n+1} \in \mathbb{R}, \tag{3-34}$$

und damit eine Erhaltensneigung des Prozesses, so handelt es sich um einen stationären GAUSS-MARKOV-Prozess (MINKLER & MINKLER 1993, Kap. 2.11). Anders formuliert beinhaltet die gegenwärtige Realisierung[16] $x(t_n)$ aus dem Prozess $X(t_n)$ alle Informationen, um eine Aussage über die Zukunft des Prozesses $X(t_{n+1})$ zu treffen (SCHLITTGEN & STREITBERG 1997, S. 144).

Instationäre Prozesse

In praktischen Anwendungen in natürlicher Umgebung treten stationäre Prozesse, bzw. eine stationäre Realisierung höchst selten auf. Die Ursache liegt in der Systemdynamik begründet. Ein natürliches System unterliegt der Neigung sich innerhalb langer Zeiträume zu verändern. Täte es das nicht, kämen alle seine Prozesse zum Stillstand. Die Prozesse werden in Rückkoppelung von den Änderungen des Systems betroffen, was sich als eine Überlagerung durch einen transienten Prozess zeigt. Bleibt ein Prozess von diesen Veränderungen unberührt, so dass sich kein natürlicher Trend ausbildet, unterliegt die resultierende Erfassung aber ihrerseits einem transienten Prozess[17]. Für die Realisierung eines natürlichen Prozesses bedeutet das den Verlust der Mittelwertstationarität (3-25), während das Zentralmoment zweiter Ordnung (3-27) oder gar noch jene höherer Ordnung die Stationaritätsbedingung erfüllen. Es ist daher möglich, einen *mittelwertinstationären* GAUSS-Prozess

$$x(t) \sim N(\mu_t, \sigma^2) \tag{3-35}$$

zu definieren, der als Erwartungswerte eine Mittelwertfolge μ_t besitzt (SCHLITTGEN & STREITBERG 1997, S. 99). Dieser Prozess lässt sich zu einem mittelwertinstationären GAUSS-MARKOV-Prozess ergänzen, indem eine Erhaltungsneigung des Prozesses definiert wird (RASMUSSEN 2003, S. 64).

---

[15] Der *white noise process* kann nur auf einem endlichen Frequenzintervall definiert werden, da seine Gesamtenergie sonst unendlich groß werden müsste (TAUBENHEIM 1969, S. 228). Für die Praxis gilt daher, dass eine maximale Frequenz $\nu_{\max}$ für diesen Prozess *implicit* vorausgesetzt wird.

[16] Der Begriff der Realisierung entspricht hier dem der *Stichprobe*.

[17] Dieses Phänomen ist dem Geodäten als Instrumentendrift geläufig.

Autoregressive Prozesse

Unter der Annahme, dass ein System dynamisch reagiert (Kap. 3.1.5) oder anders ausgedrückt, der Systemausgang Y entsprechend der Gedächtnislänge des Systems eine Erhaltensneigung besitzt, lässt sich eine einzelne Realisierung y_i als gewichtete Summe seiner benachbarten Realisierungen darstellen. In Abhängigkeit der p benachbarten Realisierungen und deren Autokorrelation lässt sich ein autoregressiver Prozess p-ter Ordnung AR[p]

$$y_i = \sum_{k=-p}^{p} \alpha_k y_{i-k} + \varepsilon_i, i \in \mathbb{Z} \tag{3-36}$$

mit dem Zufallsstoß ε_i definieren. Sei der Prozess abhängig von dem äquidistanten Zeitschritt i und die gewichtete Summe erfolge ausschließlich in negativer Richtung entlang der Zeitachse, so spricht man von einem *kausalen* AR[p]-Prozess der Ordnung p (Koch & Schmidt 1994, S. 191).

$$y_i = \sum_{k=1}^{p} \alpha_k y_{i-k} + \varepsilon_i \tag{3-37}$$

Das bedeutet mit anderen Worten, dass sich ein aktueller Wert des Prozesses y vollständig aus einer Summierung vergangener Systemausgänge rekonstruieren lässt (Schlittgen & Streitberg 1997, S. 122). Die Parameter α_k können mit Hilfe der Yule-Walker-Gleichung (Anhang A.2) berechnet werden.

Gleitende Mittelwertprozesse

Sei der Systemeingang ein unabhängiger Zufallsstoß zum äquidistanten Zeitschritt i mit dem Erwartungswert μ und der Varianz σ^2, so erzeugt ein lineares System (3-5) als Ausgang den *allgemeinen linearen* Prozess

$$y_i = \sum_{k=-\infty}^{\infty} \beta_k \varepsilon_{i-k},\ i \in \mathbb{Z}. \tag{3-38}$$

Erfolgt die gewichtete Summe ausschließlich in negativer Richtung entlang der Zeitachse, so handelt es sich um einen Spezialfall eines allgemeinen linearen Prozesses

$$y_i = \sum_{k=0}^{q} -\beta_k \varepsilon_{i-k} \text{ mit } \beta_0 = -1 \tag{3-39}$$

und man spricht von einem *kausalen* gleitenden Mittelwertprozess[18] der Ordnung q oder kurz von einem MA[q]-Prozess (Schlittgen & Streitberg 1997, S. 116). Die Parameter β_k lassen sich in der Levinson-Durbin-Rekursion schätzen (Schlittgen & Streitberg 1997, S. 196f).

Autoregressive gleitende Mittelwertprozesse

Die Summe von einem autoregressiven Prozess der Ordnung p (3-37) und einem gleitenden Mittelwertprozess der Ordnung q (3-39) ergibt sich zum *autoregressiven gleitenden Mittelwertprozess* der Ordnung p und q oder kurz ARMA[p, q] (Schlittgen & Streitberg 1997, S. 132).

$$y_i = \sum_{k=1}^{p} \alpha_k y_{i-k} + \sum_{k=0}^{q} -\beta_k \varepsilon_{i-k} \text{ mit } \beta_0 = -1 \tag{3-40}$$

[18] Engl. *moving-average process.*

Berücksichtigt man die Tatsache, dass der Systemeingang $x(t)$ zum diskreten Zeitpunkt i nur aus ε_i besteht, so muss ein System, dass einen ARMA-Prozess erzeugt, die Eigenschaft der inkrementalen Linearität (Kap. 3.1.3) erfüllen. Das erzeugende System ist weiterhin dynamisch, denn außer für den Spezialfall ARMA[0,0] kann (3-10) nicht erfüllt sein. Schließlich, da es keine negativen Ordnungen gibt, ist das erzeugende System kausal (Kap. 3.1.5).

3.3 Prozessanalyse

Der erste Schritt, um sich einem System $\mathcal{T}$ und seinen Prozessen X, Y zu nähern, ist die Analyse der Realisierungen x, y. Auch wenn diese Analyse sich an einer objektiv-mathematischen Vorgehensweise orientieren sollte, wird die grafische Darstellung der Realisierungen und ihre visuelle Erfassung maßgeblich den Fortgang der Analyse beeinflussen (Jäger et al. 2005, S. 49ff). Zweifelsfrei wird diese sinnliche Wahrnehmung subjektive und hypothesenbehaftete Entscheidungen zur Folge haben.
Dennoch soll die Prozessanalyse gleich am Anfang alle möglichen Erkenntnisse aus der Realisierung der Prozesse x, y extrahieren, bevor mit einer Modellierung, sei sie nun imitierend oder identifizierend (Kap. 4.1), begonnen wird. Einen eindeutigen Weg der Analyse gibt es hierbei nicht (Heinert & Niemeier 2004). Die Art der zu analysierenden Prozesse bedingt das weitere Vorgehen zur Analyse und Beschreibung. Dabei sind nicht selten Kompromisse aufgrund der Qualität der Realisierung einzugehen. Die Standardverfahren zur Analyse stochastischer Prozesse, so wie sie häufig in der Geodäsie zur Verwendung kommen (DIN 18709-5:Entwurf-2005; Koch & Schmidt 1994; Niemeier 1980; Schlittgen & Streitberg 1997; Taubenheim 1969, u.a.), liefern nur eine Beschreibung der schwach stationären Anteile eines Prozesses. Sobald eine Änderung eines der beiden Erwartungswerte $E\{x(t)\}$ und $E\{x^2(t)\}$ eintritt, beispielsweise durch einen überlagerten transienten oder quasiperiodischen Subprozess, versagen die Standardverfahren. Sie können nur periodische oder stochastische Prozesse (Kap. 3.2.1 u. 3.2.2) analysieren. Damit ist ihre Nutzung an die Bedingung der schwachen Stationarität gekoppelt. Ist die Bedingung nicht erfüllt, muss sie vor einer Nutzung erzwungen werden.

3.3.1 Empirische Schätzung der Momente eines Prozesses

Wie in Kap. 3.2.2 bereits dargelegt, beschreiben die Zentralmomente eines Prozesses dessen stochastisches Verhalten. Zumeist wird nur eine Schätzung der ersten beiden Zentralmomente vorgenommen, da sie für viele Zwecke bereits hinreichend ist.

Schätzung des Momentes erster Ordnung

Unter der Annahme eines mittelwertstationären Prozesses (3-25) kann der Erwartungswert empirisch mit verschiedenen Verfahren geschätzt werden. In den meisten Fällen erfolgt die Schätzung als arithmetisches Mittel $\bar{y}$ (Niemeier 2002, S. 3). Diese hat ihre Gültigkeit für alle Prozesse mit einer symmetrischen Verteilungsfunktion. Da eine rein symmetrische Verteilungsfunktion jedoch nicht immer vorliegt, kommen weitere Schätzer zum Einsatz, wie der Median oder ein robuster Schätzer auf der Basis beschränkter oder zurückfallender Einflussfunktionen $\Psi(v)$ (Kap. 4.3.2). Der Mittelwertschätzer wird daher mit der Gewichtsfunktion g^* in der diskreten Form

$$\bar{y} = \frac{\sum\limits_{i=-N}^{N} g_i^* \cdot y_i}{\sum\limits_{i=-N}^{N} g_i^*} \text{ mit } g_i^* = \frac{\Psi(v_i)}{v_i} \text{ und } v_i = y_i - \bar{y}^* \tag{3-41}$$

verallgemeinert. Der bekannte Mittelwertschätzer des arithmetischen Mittels auf der Basis der kleinsten Fehlerquadrate ergibt sich damit für $g_i^* = 1$.
Eine robust gewichtete Mittelbildung startet mit einer Näherung des Mittels $\bar{y}^*$. Je nach deren Qualität ist nur eine Berechnung erforderlich. Andernfalls wird eine Iteration notwendig. Erfahrungsgemäß ist der Median $\mathrm{med}(\mathbf{y})$ als Approximation hinreichend. Ein mittelwertstationärer Prozess mit einem konstanten Erwartungswert ist aber erfahrungsgemäß die Ausnahme. Alle weiteren Untersuchungen eines Prozesses, welche die Stationarität voraussetzen, würden bereits jetzt ausscheiden müssen, wie beispielsweise alle auf der diskreten FOURIER-Transformation (3-53) basierenden Methoden zur Frequenzanalyse. Aus diesem Grund wird häufig versucht, die Mittelwertstationarität zu erzwingen (SCHLITTGEN 2001, S. 5f), indem man lineare, transiente oder langwellige Subprozesse als Trend abspaltet. Im Falle des robusten Schätzers gelänge eine empirische Schätzung der Mittelwertfolge μ_t (3-35) durch die Erweiterung mit einer weiteren lokalen Gewichtsfunktion

$$\bar{y}_k = \frac{\sum\limits_{i=-N}^{N} (g_i \cdot g_i^*) y_{k-i}}{\sum\limits_{i=-N}^{N} (g_i \cdot g_i^*)}. \tag{3-42}$$

Das Resultat ist ein gleitendes, robust gewichtetes Mittel. Die Frage nach einer geeigneten Gewichtsfunktion g_i (3-70) wird im Kap. 3.3.3 erneut aufgegriffen, denn formal entspricht dieses Vorgehen einer amplitudentreuen, nichtrekursiven Filterung.

Schätzung des Momentes zweiter Ordnung

Um eine erste praktische Aussage über die Erhaltensneigung oder die Periodizität eines Prozesses treffen zu können, wird die *empirische Autokovarianzfunktion*

$$C_{yy}(\tau) = \frac{1}{N} \sum_{\tau} (y_t - \bar{y})(y_{t+\tau} - \bar{y}) \tag{3-43}$$

berechnet (SCHLITTGEN & STREITBERG 1997, S. 7f). In der dargestellten Form entspricht sie der Schätzung des zweiten Zentralmomentes eines schwach stationären Prozesses (Kap. 3.2.2). In der Gebrauchsform für diskrete Realisierungen

$$C_{yy}(i) = \frac{1}{N} \sum_{k=0}^{N-i-1} (y_k - \bar{y}_k)(y_{k+i} - \bar{y}_{k+i}) \tag{3-44}$$

kann der robuste Schätzer für die Erwartungswertfolge $\bar{y}_k$ aus Formel (3-42) verwendet werden, um auch für mittelwertinstationäre Prozesse zu einer Autokovarianzfunktion zu gelangen. Der verbleibende mittelwertstationäre Prozess weist aber noch Ausreißer auf, die eine Schätzung der Autokovarianzfunktion empfindlich verfälschen können. Hier lässt sich SUTORS Vorschlag (1997, S. 31) mittels der Psi-Funktionen über die verkürzten Daten $\Psi(\Delta y)$ die robuste Autokovarianzfunktion

$$C_{yy}(i) = \frac{1}{(s_k^*)^2} \sum_{k=0}^{N-i-1} \Psi(y_k - \bar{y}_k)\, \Psi(y_{k+i} - \bar{y}_{k+i}) \tag{3-45}$$

zu berechnen, abermals erweitern, indem durch die Nutzung der Folge

$$s_k^* = \sqrt{\frac{\sum\limits_{i=-N}^{N} (g_i \cdot g_i^*) v_{k-i}}{\sum\limits_{i=-N}^{N} (g_i \cdot g_i^*)}} \tag{3-46}$$

eine lokale Definition erreicht wird. Der Skalenfaktor s_k^* ergibt sich analog zu (3-42) unter der Verwendung der beiden Gewichtsfunktionen g und g^*. Der hieraus resultierende Schätzer schneidet unter den verschiedenen robusten Schätzern recht gut ab, wobei auch der rechentechnische Aufwand gering bleibt (SUTOR 1997, Kap. 3.3.6). Der letzte Schritt ist allerdings nur bedingt sinnvoll, da die Robustheit des Schätzers überwiegend von der Mittelwertschätzung abhängig ist. Liegen fastperiodische Subprozesse (Kap. 3.2.1) vor, wirkt sich der Ansatz der Ψ-Funktion der robusten Schätzer (Kap. 4.3.2) auf die Differenzen Δy in (3-45) eher kontraproduktiv aus.
Aus der Normierung der Autokovarianzfunktion als Division durch dessen erstes Element erhält man die *empirische Autokorrelationsfunktion*

$$R_{yy}(\tau) = \frac{C_{yy}(\tau)}{C_{yy}(0)} = \frac{C_{yy}(\tau)}{\sigma_y^2}. \tag{3-47}$$

Diese nimmt gemäß der umgeformten CAUCHY-SCHWARZ-Ungleichung

$$1 \leq \frac{\left(\sum_i a_i^2\right) \cdot \left(\sum_i b_i^2\right)}{\left(\sum_i a_i b_i\right)^2} \tag{3-48}$$

nur Werte für die Korrelationen von -1 bis 1 an (SCHLITTGEN & STREITBERG 1997, S. 97 u. S. 498). Die Verallgemeinerung der Autokovarianzfunktion stellt die *empirische Kreuzkovarianzfunktion*

$$C_{xy}(k) = \frac{1}{N} \sum_{k=0}^{N-i} (x_i - \bar{x})\,(y_{i+k} - \bar{y}) \tag{3-49}$$

dar (SCHLITTGEN & STREITBERG 1997, S. 56). Die *empirische Kreuzkorrelationsfunktion*

$$R_{xy}(\tau) = \frac{C_{xy}(\tau)}{\sqrt{\sigma_x^2 \cdot \sigma_y^2}} \geq \frac{C_{xy}(\tau)}{\sqrt{C_{xx}(0) C_{yy}(0)}} \tag{3-50}$$

wird analog zu (3-47) gebildet (NATKE 1992, S. 101ff) und kann auch entsprechend (3-45) robustifiziert werden.

3.3.2 Spektralanalyse

In der Prozessanalyse sind die periodischen Subprozesse von besonderem Interesse. Die Tatsache, ob sich ein Phänomen eines Ausgangsprozesses täglich, wöchentlich, jährlich oder andersartig zyklisch wiederholt, liefert schon eine Reihe von Hinweisen auf die möglicherweise kausalen Systemeingänge und schließt gleichermaßen andere als Ursache aus. Die gebräuchlichste Übertragung eines realisierten Prozesses in den Frequenzraum erfolgt auf der Grundlage einer FOURIER-Transformation (TAUBENHEIM 1969, Kap. 7.3).

Die FOURIER-Transformation

Nach dem WIENER-KHINTCHINE-Theorem[19] existiert für die Autokovarianzfunktion C_{yy} ein Leistungsspektrum über die Frequenzen ν des Prozesses Y (TAUBENHEIM 1969, S. 226). Dieses Spektrum lässt sich als FOURIER-Kosinus-Transformation der Autokovarianzfunktion

$$F(\nu) = 4 \int_0^\infty C_{yy}(\tau) \cos(2\pi\nu\tau) \partial\tau \tag{3-51}$$

[19] Dieses Theorem ist auch als EINSTEIN-WIENER-KHINTCHINE-Relation bekannt.

berechnen (Kuhlmann 1996, S. 34). Die Einschränkung auf den Realteil der Fourier-Transformation ist zulässig, da die empirische Autokovarianzfunktion positiv definit ist. Eine Ausnahme stellen hier einige robuste Schätzer für C_{yy} dar (Sutor 1997). Das Parseval-Theorem erweitert den Einsatz der Fourier-Transformation erheblich (Schlittgen & Streitberg 1997, S. 80). Demnach lässt sich eine beliebige Funktion mittels der kontinuierlichen Fourier-Transformation

$$F(\nu) = \int_{-\infty}^{\infty} f(x) e^{-i\nu x} \partial x \tag{3-52}$$

in einen komplexen Frequenzraum übertragen. Diese Eigenschaft ist nicht nur für reine Analysezwecke wichtig, sondern ist auch für das Design von Filteralgorithmen von erheblicher Bedeutung (Abb. 3.5). Für die Transformation von diskret realisierten Prozessen kommt die Summenform

$$F_j = \sum_{k=0}^{n-1} y_k e^{-i2\pi \frac{j}{n} k} \tag{3-53}$$

zum Einsatz. Die diskreten Frequenzen ν_j im Frequenzraum errechnen sich zu

$$\nu_j = \frac{j}{n\Delta t}, \tag{3-54}$$

wobei die Abtastrate Δt als Kehrwert zur Abtastfrequenz ν_A gebildet werden kann. Die größte im Frequenzraum darstellbare Frequenz, die Nyquist-Frequenz ν_N, muss mit genau zwei Realisierungen pro Zyklus $\frac{1}{2\Delta t}$ erfasst werden. Für diese Frequenz kann keine Amplitude oder Phasenlage mehr berechnet werden. Größere Frequenzen führen zu dem Phänomen des *aliasing* (Schlittgen & Streitberg 1997, Kap. 1.6.3). Daher ist es zweckmäßig, bei der Realisierung von unbekannten Prozessen auch kurzzeitig eine höhere Diskretisierung vorzunehmen.

Gegen die Frequenzen ν_j lassen sich die Amplituden

$$A_j = 2|F_j| \tag{3-55}$$

als normierte Beträge des komplexen Ergebnisses der Fourier-Transformation auftragen.

Trotz der nützlichen Eigenschaften der Fourier-Transformationen stellen mit ihnen erzeugte Spektren keine geeigneten Schätzer für die theoretische Spektraldichte eines realisierten Prozesses dar (Schlittgen & Streitberg 1997, S. 353)[20]. Ein wesentliches Merkmal der diskreten Fourier-Transformationen ist die Anfälligkeit der Amplituden in Abhängigkeit von der Länge des transformierten Intervalls (Heinert 1998, Abb. 4.14). Entspricht das Transformationsintervall nicht einem ganzzahligen Vielfachen der Wellenlänge, so kann die Amplitude im schlechtesten Fall auf das $2/\pi$-fache absinken (Abb. 3.3a). Dieser Umstand führt zu der Merkwürdigkeit, dass im Einzelfall ein kürzeres Intervall eine bessere Schätzung zum Ergebnis hat. Zusätzlich sind gerade die langwelligen Periodizitäten von der geringen Anzahl der Summationsstellen im Frequenzbereich beeinträchtigt.

Die Intervallverlängerung (*zero padding*)

Die Verlängerung des Intervalls um angehängte Nullen (*zero padding*) kann diese Beeinträchtigung zwar beheben, aber die Unstetigkeitsstelle zwischen dem ursprünglichen Intervall und

[20] Obwohl an dieser Stelle Bezug auf das Periodogramm als das Ergebnis einer Fourier-Transformation auf der Zeitreihe (3-53) selbst genommen wird, hat die Aussage gemäß des Parseval-Theorems in gleicher Weise auch Gültigkeit für die Fourier-Kosinus-Transformation.

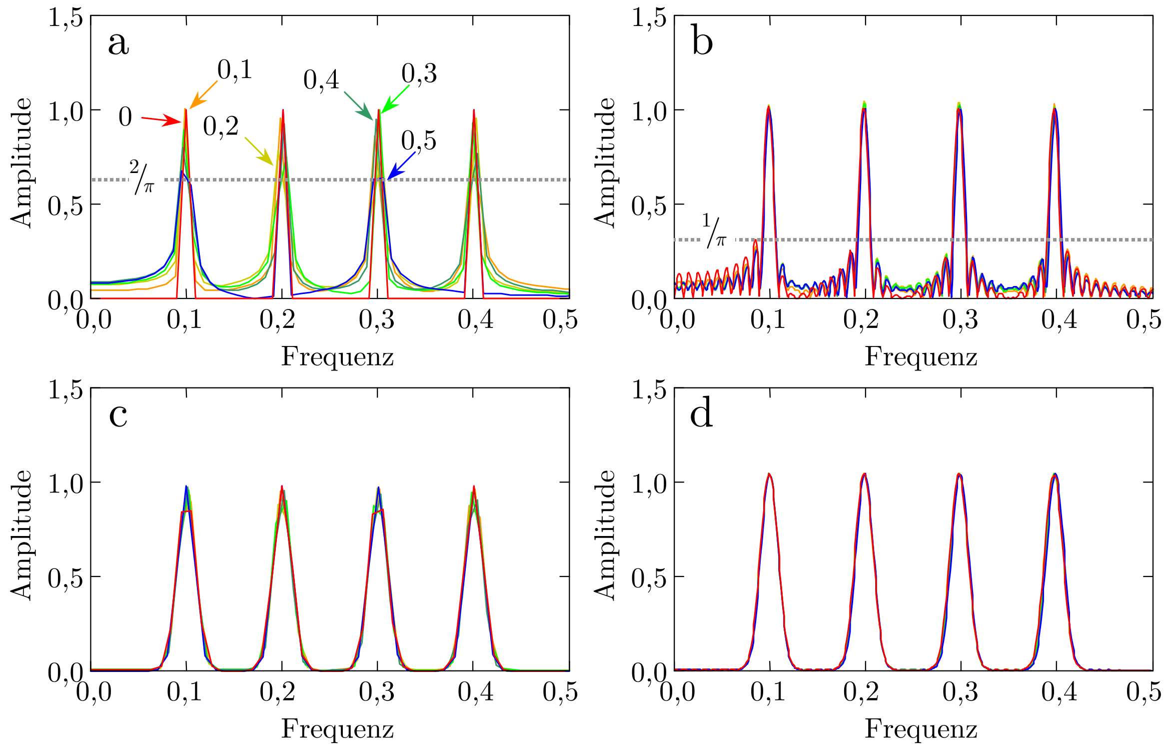


Abb. 3.3: Spektren im Vergleich, wenn das transformierte Interval um das 0,1- bis 0,5-fache der Wellenlänge der Grundfrequenz verlängert wird: a) bei einer einfachen FOURIER-Transformation, b) mit *zero padding*, c) gefenstert und d) gefenstert mit *zero padding*.

den angehängten Nullen wirkt wie ein Rechteckfenster. Dessen FOURIER-Transformierte

$$H(\nu) = \frac{\sin(\pi \Delta t f)}{\pi \Delta t f} \tag{3-56}$$

ergibt eine abklingende Welle (MARKO 1995, S. 248). Nach deren Betragsbildung im Frequenzbereich bilden sich typischen Nebenspitzen (*sidelobes*) aus, die das $1/\pi$-fache der größten Amplitude erreichen können (Abb. 3.3b). Diese sind geeignet, benachbarte Frequenzen zu maskieren. Jede Form der Fensterung vermindert die Wirkung der Nebenspitzen, die Hauptspitze verbreitert sich dafür erheblich. Dicht benachbarte Frequenzen werden so ebenfalls in Amplitude und Frequenz falsch geschätzt.

Die Fensterung

Bei einer FOURIER-Transformation über ein endliches Intervall eines realisierten Prozesses existieren aufgrund des sogenannten Abschneideeffektes Unstetigkeiten an den Intervallgrenzen (NATKE 1992, S. 135ff). Während einerseits die Amplituden über den enthaltenen Periodizitäten absinken, kommt es zu einem Einsickern von Energie in die benachbarten Frequenzen (*leakage*) (SCHLITTGEN & STREITBERG 1997, S. 389ff). Das kann bei einem unverlängerten Invervall zu einer monoton fallenden Funktion ausgehend vom Amplitudenzentrum führen (Abb. 3.3a) oder im Falle eines mit Nullen verlängerten Intervalls die typischen Nebenspitzen ausbilden (Abb. 3.3b). In beiden Fällen führt die Transformation

$$F_j^* = \sum_{k=0}^{n-1} g_{k-\frac{n}{2}} y_k e^{-i2\pi \frac{j}{n} k} \tag{3-57}$$

nach der Multiplikation mit einer zu den Intervallgrenzen hin monoton fallenden Fensterfunktion $g_{k-\frac{n}{2}}$ zu einer stabileren Schätzung der Amplituden. Bei einer gefensterten FOURIER-Transformation über ein unverlängertes Intervall bleiben die Amplituden stabil, aber die Schätzung der zugehörigen niederen Frequenzen hängt nach wie vor stark von den Summationsstellen ab. Zusätzlich verbreitern sich die *peaks* (BUTZ 2000, Abb. 3.10 u. vergl. i.d.Arb. Abb. 3.3c). Bei einer gefensterten FOURIER-Transformation mit *zero padding* sind die Frequenzen und Amplituden relativ stabil, die verbreiterten *peaks* bleiben aber erhalten (Abb. 3.3d).

Die Zeit-Frequenz-Analyse

Die Zeit-Frequenz-Analyse bietet die Möglichkeit, die Instationaritäten eines quasiperiodischen Prozesses (Kap. 3.2.1) über die Zeit zu verfolgen (UNBEHAUEN 2002, S. 259ff). Die Veränderungen des Prozesses können sich sowohl als eine Schwankung der Amplitude oder als eine Frequenzdrift eines Subprozesses im Zeit-Frequenz-Spektrum zeigen (Abb. 3.4). Eine Schwankung der Amplitude wäre ein Indiz für erregte und abklingende Reaktionen. Eine Frequenzdrift kann auf eine Veränderung oder gar Schädigung eines Beobachtungsobjektes hinweisen.

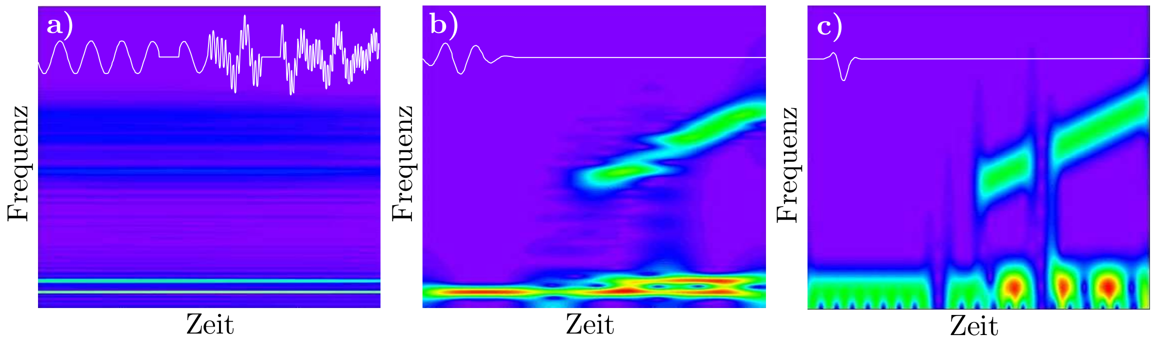


Abb. 3.4: Die Zeit-Frequenz-Analyse mit GAUSS-Fenstern verschiedener Länge: a) ein breites Fenster entspricht nahezu der ungefensterten FOURIER-Transformation des Datensatzes (weiße Kurve), so dass quasiperiodische Prozesse in schwache statische Periodizitäten zerlegt werden. Ein Fenster mittlerer Breite wie in b) ist in Frequenz- und Zeitauflösung gleichermaßen unscharf. Ein kurzes Fenster wie in c) führt zu einer geringen Schärfe in der Frequenz, zeigt aber deutlich die zeitlichen Veränderungen. Am Datensatz (weiß) lässt sich die Wirkung der Fensterung ablesen.

Diese Analyse basiert auf der FOURIER-Transformation

$$^{*}\mathbf{I}^{j}_{t_c} = \mathrm{F}\left(\mathrm{diag}\left[\mathbf{g}^{b_j}_{t_c} \times \mathbf{y}\right]\right) \quad \text{mit} \quad j = 0 \ldots n \tag{3-58}$$

des Produktes des Prozesses $\mathbf{y}$ mit einer gleitenden Fensterfunktion, wie zum Beispiel der GAUSS-Funktion

$$g^{b_j}_{t_c}(t) = e^{-\frac{(t-t_c)^2}{b_j}} \tag{3-59}$$

mit der Bezugszeit t_c der Filterbreite b_j. Die Elemente der resultierenden Amplitudenmatrix errechnen sich aus den normierten Beträgen der komplexwertigen FOURIER-Transformation:

$$^{*}A^{j}_{t_c,\nu} = \frac{2\left|^{*}I^{j}_{t_c,\nu}\right|}{n_{\mathrm{eff},j}} \quad \text{mit} \quad n_{\mathrm{eff},j} = \sum_{t=t_0}^{t_n} g^{b_j}_{t_c}(t). \tag{3-60}$$

Bei jeder einzelnen Analyse kann eine hohe Auflösung nur in der Frequenz oder der Zeit erreicht werden (UNBEHAUEN 2002, S. 265)[21]. Eine Kompromisslösung besteht in der Wahl einer gleichermaßen eingeschränkten Auflösung in Frequenz und Zeit (Abb. 3.4).

[21] In der Diskussion fällt hier häufig der Verweis auf die HEISENBERGsche Unschärfe-Relation, dass keine hohe Auflösung in Zeit und Frequenz gleichermaßen möglich sei. Allgemeingültige Literatur findet sich hierzu nicht.

3.3.3 Die Filterung

In der Prozessanalyse ist es häufig nötig, gesuchte Eigenschaften eines Prozesses gegenüber unerwünschten hervorzuheben. Um die Prozesse besser untersuchen und verstehen zu können, bedient man sich linearer Filteroperatoren.

Nichtrekursive Filter

In der geodätischen Praxis werden zumeist jene linearen Operatoren als Filter bezeichnet, welche im systemtheoretischen Kontext als nichtrekursive akausale Filter bekannt sind. Aus den zeitlich oder räumlich geordneten Werten eines Prozesses $x(t)$ führt ein linearer Filteroperator eine Integration[22]

$$y(t) = \int_{-\infty}^{\infty} g(\tau) \cdot x(t-\tau)\partial\tau \tag{3-61}$$

über dem Produkt mit einer geeigneten Gewichtsfunktion $g(\tau)$ durch (Taubenheim 1969, Kap. 9). Handelt es sich bei dem zu untersuchenden Prozess nicht um einen deterministischen Zusammenhang im Sinne einer bekannten Funktion, sondern um eine diskrete Realisierung des Prozesses, so wird die lineare Operation

$$y(t) = \sum_{k=-\infty}^{\infty} g_k \cdot x_{t-k} \tag{3-62}$$

zu einer gewichteten Summe über die zeitlich oder räumlich äquidistanten Werte der Prozessrealisierung. Das Ergebnis, die gewichtete Summe, verstärkt einzelne Frequenzen oder Frequenzbänder dieses Prozesses, und dämpft die übrigen ab (Niemeier 1980). Abweichend kann diese Summierung mit Hilfe des *back shift*-Operators B^k in der Form

$$y(t) = \sum_{k=-\infty}^{\infty} g_k \mathrm{B}^k x_t \tag{3-63}$$

notiert werden. Zur Darstellung dieser Eigenschaften des linearen Operators wird seine Gewichtsfunktion $g(\tau)$ mittels der komplexen Fourier-Transformation

$$H(\nu) = \mathrm{F}(g(\tau)) = \int_{-\infty}^{\infty} g(\tau) e^{-i2\pi\nu\tau}\partial\tau \tag{3-64}$$

in den Frequenzbereich übertragen[23]. Das Ergebnis dieses Integrals ist die komplexwertige Durchlasscharakteristik $H(\nu)$ des Filteroperators[24]. Multipliziert man die Durchlasscharakteristik $H(\nu)$ mit der Fourier-Transformierten des Ausgangsprozesses $\mathrm{F}(y(t))$ und führt man anschließend die inverse Fourier-Transformation durch, so ergibt sich ebenfalls der gefilterte Folgeprozess (Abb. 3.5).

$$y(t) = \mathrm{F}^{-1}(\mathrm{F}(y(t)) \cdot H(\nu)) \tag{3-65}$$

Eine weitere Eigenschaft eines Filters ist die Phasenverschiebung. Sie wird durch die Phasen-

[22] Dieses Integral wird auch als Duhamel- oder Faltungsintegral bezeichnet.

[23] Bildlich gesprochen wird durch eine Gewichtsfunktion eine komplexe harmonische Schwingung der Amplitude 1 in Abhängigkeit von der Frequenz ν geschickt (Leiner 1998, S. 75).

[24] Fasst man die diskreten Frequenzen eines Prozesses in der Weise auf, dass sie Elemente einer Menge sind, so ordnet sie ein linearer Operator einer unscharfen Menge mit der Zugehörigkeit von 0 bis 1 zu.

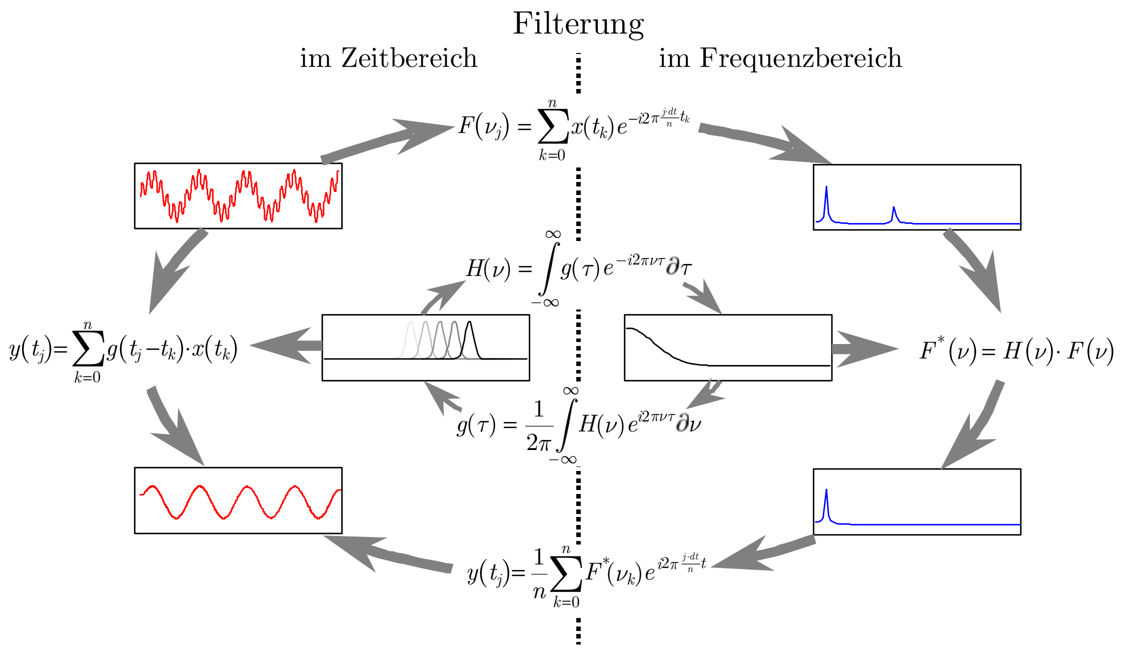


Abb. 3.5: Die praktischen Beziehungen zwischen der Tiefpassfilterung im Zeit- und im Frequenzbereich (NATKE 1992, Abb. 2.32).

charakteristik $\phi(\nu)$, also Quotienten aus dem imaginären $\Im(\cdot)$ und dem realen Anteil $\Re(\cdot)$ der Durchlasscharakteristik ausgedrückt (vergl. (3-20)).

$$\phi(\nu) = \arctan\left(\frac{\Im(H(\nu))}{\Re(H(\nu))}\right) \tag{3-66}$$

Die empirische Bestimmung der Durchlasscharakteristik eines Filters kann durch eine FOURIER-Transformation über das Produkt der Gewichte mit dem DIRAC-Impuls δ_t (3-15) erfolgen:

$$H(\nu_j) = \left|\sum_{j=0}^{n}\left(\sum_{k=0}^{n}\delta(k)\cdot g_{j+k}\right)e^{-i2\pi\frac{j\cdot k}{n}}\right|. \tag{3-67}$$

Dieses Vorgehen ist besonders dann hilfreich, wenn sich die Gewichte nicht als geschlossene Funktion ausdrücken lassen, sondern nur als Distribution vorliegen. Liegen keine Informationen über den Filteroperator vor, so ist es möglich, den Quotienten

$$H(\nu_j) = \frac{|\mathrm{F}(y(t))|}{|\mathrm{F}(x(t))|} \tag{3-68}$$

aus dem Ausgangs- und Eingangsprozess zu bilden. Dieser gibt je nach Qualität der Spektren einen guten Eindruck von der Durchlasscharakteristik.

Amlitudentreue nichtrekursive Filter

Soll der Folgeprozess $y(t)$ amplitudentreu sein, so muss das unendliche Integral über der Gewichtsfunktion

$$\int_{-\infty}^{\infty} g(\tau)\partial\tau = 1 \tag{3-69}$$

den Wert eins annehmen. Weicht die Gewichtsfunktion von dieser Forderung ab, so ist es möglich, einen diskreten Filteroperator

$$y(t) = \frac{\sum\limits_{\tau=-\infty}^{\infty} g_\tau \cdot x_{t-\tau}}{\sum\limits_{\tau=-\infty}^{\infty} g_\tau} \;\forall\; \tau \in \mathbb{T} \subset \mathbb{R} \tag{3-70}$$

mit einer entsprechenden Normierung zu versehen.

Phasentreue nichtrekursive Filter

Der Folgeprozess eines Filteroperators ist phasentreu, wenn der Funktionswert der Phasencharakteristik $\phi(\nu)$ für alle Frequenzen ν den Wert null annimmt. Das ist gerade dann erfüllt, wenn der Imaginärteil der Durchlasscharakteristik

$$\phi(\nu) = \arctan\left(\frac{\Im(H(\nu))}{\Re(H(\nu))}\right) \stackrel{!}{=} 0 \;\forall\; \nu \quad \Rightarrow \quad \Im(H(\nu)) = 0 \;\forall\; \nu \tag{3-71}$$

für alle Frequenzen den Wert null hat. Um die Eigenschaften der zugehörigen Gewichtsfunktion des linearen Filteroperators $g(\tau)$ bestimmen zu können, wird der Imaginärteil der Durchlasscharakteristik

$$\begin{aligned}\Im(H(\nu)) &= \Im\left(\int\limits_{-\infty}^{\infty} g(\tau) e^{-i2\pi\nu\tau} \partial\tau\right)\\ &= \int\limits_{-\infty}^{0} g(\tau)\sin(2\pi\nu\tau)\partial\tau + \int\limits_{0}^{\infty} g(\tau)\sin(2\pi\nu\tau)\partial\tau\end{aligned} \tag{3-72}$$

aus dem Integral der FOURIER-Transformation errechnet. Durch die Variation der Integrationsgrenzen des ersten Terms kann das Integral aufgelöst werden. Aufgrund der Ungeradheit der Sinusfunktion kann das negative Vorzeichen des Inkrementes auf den gesamten ersten Integralterm übertragen werden:

$$-\int\limits_{0}^{\infty} g(-\tau)\sin(2\pi\nu\tau)\partial\tau + \int\limits_{0}^{\infty} g(\tau)\sin(2\pi\nu\tau)\partial\tau \stackrel{!}{=} 0. \tag{3-73}$$

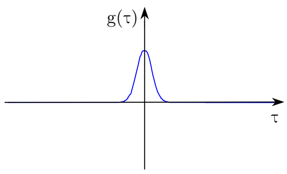
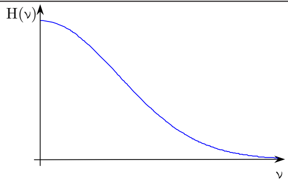
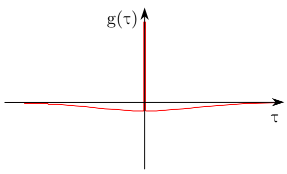
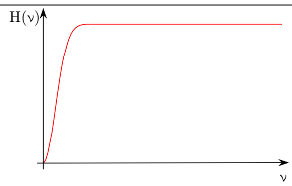
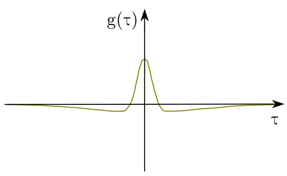
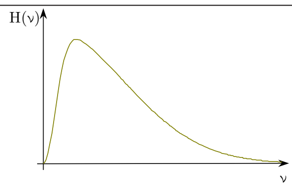
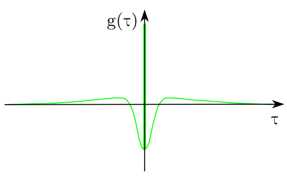
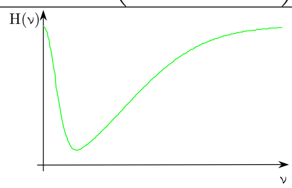
Damit ist die Gleichung (3-71) genau dann erfüllt, wenn die Gewichtsfunktion $g(\tau)$ gerade ist:

$$\Im(H(\nu)) \stackrel{!}{=} 0 \;\forall\; g(\tau) = g(-\tau). \tag{3-74}$$

Nichtrekursive Filter im Frequenzbereich

Im Frequenzbereich werden nichtrekursive Filter durch ihre Durchlasscharakteristik beschrieben. Die Durchlasscharakteristik eines *Tiefpasses* $H_T(\nu)$ (Tab. 3.1), eines Filters, das hochfrequente Periodizitäten wie beispielsweise das Rauschen dämpft, zeichnet sich in den meisten Fällen durch eine monoton fallende Funktion aus (DIN 18709-5:Entwurf-2005). Jedoch bildet gerade der einfachste Tiefpass, das *gleitende Mittel*, eine Ausnahme. Dessen Durchlasscharakteristik ist zunächst monoton fallend und beginnt dann auszuschwingen (WELSCH ET AL. 2000, Abb. 8.2-3). Die Durchlasscharakteristik entspricht der FOURIER-Transformation einer

Tab. 3.1: Filteroperatoren und ihre Durchlasscharakteristika am Beispiel für Exponentialfunktionen (DIN 18709-5:Entwurf-2005)

	$g(\tau)$	$H(\nu)$
Tiefpass	 $\frac{1}{\theta\sqrt{\pi}}e^{-\frac{\tau^2}{\theta^2}}$	 $e^{-\pi^2\tau^2\theta^2}$
Hochpass	 $\delta(\tau) - \frac{1}{\theta\sqrt{\pi}}e^{-\frac{\tau^2}{\theta^2}}$	 $1 - e^{-\pi^2\tau^2\theta^2}$
Bandpass	 $\frac{1}{\theta\sqrt{\pi}}e^{-\frac{\tau^2}{\theta^2}} - \frac{1}{\sqrt{\pi\theta^2+\pi\vartheta^2}}e^{-\frac{\tau^2}{\theta^2+\vartheta^2}}$	 $e^{-\pi^2\tau^2\vartheta^2}\left(1 - e^{-\pi^2\tau^2\theta^2}\right)$
Bandsperre	 $\delta(\tau) - \frac{1}{\theta\sqrt{\pi}}e^{-\frac{\tau^2}{\theta^2}} + \frac{1}{\sqrt{\pi\theta^2+\pi\vartheta^2}}e^{-\frac{\tau^2}{\theta^2+\vartheta^2}}$	 $1 - e^{-\pi^2\tau^2\vartheta^2} + e^{-\pi^2\tau^2\left(\theta^2+\vartheta^2\right)}$

Rechteckfunktion (3-56). Die Durchlasscharakteristik des *Hochpasses* $H_H(\nu)$ (Tab. 3.1), eines Filters, das niederfrequente Periodizitäten wie beispielsweise Trends dämpft, ist eine monoton steigende Funktion. Sie stellt meistens das Komplement des entsprechenden Tiefpasses dar, so dass gilt

$$H_H(\nu) = 1 - H_T(\nu)\ \forall \nu \in \mathbb{R}^+. \tag{3-75}$$

Die Durchlasscharakteristik eines *Bandpasses* $H_P(\nu)$ ist die additive oder multiplikative Kombination zweier oder mehrerer Durchlasscharakteristiken von Hoch- und Tiefpässen. Jene ist eine Funktion, die auf einem endlichen Frequenzintervall signifikant von null verschieden ist. Im einfachsten Falle handelt es sich um die Hintereinanderschaltung eines Hoch- und eines Tiefpasses (Tab. 3.1). Auch hier existiert das Komplement, die sogenannte *Bandsperre*[25]

$$H_S(\nu) = 1 - H_P(\nu)\ \forall \nu \in \mathbb{R}^+. \tag{3-76}$$

Deren Durchlasscharakteristik ist eine Funktion, die auf einem endlichen Frequenzintervall signifikant von eins verschieden ist.

[25] Dieser Operator wird auch als *Ausblendfilter* bezeichnet (Welsch et al. 2000, S. 8.3-2).

Rekursive Filter

Um einen trendbehafteten Ausgangsprozess eines dynamischen Systems Y (Kap. 3.1.5) mittels eines Filters nachbilden zu können, sind nichtrekursive Filter nur dann geeignet, wenn zuvor eine Trendschätzung durchführbar gewesen ist. Soll das unendliche Gedächtnis eines dynamischen Systems in der Filterung mit berücksichtigt werden, so muss das rekursive Filter nicht nur die Realisierung des Systemeingangs x, sondern auch die des Systemausgangs y der diskreten Vorepochen mit berücksichtigen (KOCH & SCHMIDT 1994, S. 58ff):

$$y(t) = \int\limits_{0}^{\infty} g_\tau x_{t-\tau} \partial\tau - \int\limits_{\tau}^{\infty} g_\tau^* y_{t-\tau} \partial\tau. \tag{3-77}$$

Die autoregressive Rekursion über die zurückliegenden $y_{t-\tau}$ bildet auch die Eigenschaft eines IIR-Systems[26] ab. Dieses Filter ist *kausal*, wenn die Integrationsgrenzen in (3-77) positiv, beziehungsweise *antikausal*, wenn die Integrationsgrenzen negativ sind.

Verschiedene numerische Anwendungen in der Geodäsie lassen sich auf (3-77) zurückführen. Eine direkte Verwendung der kausalen rekursiven Filter findet sich in den ARX-Modellen (4-29). Von diesen lassen sich unmittelbar die Neuronen (4-30) der Künstlichen Neuronalen Netze herleiten.
Aber auch die Systemgleichung (4-49), die zur Prädiktion im KALMAN-Filter in Gebrauch ist, weist große Ähnlichkeiten zum allgemeinen rekursiven Filter auf. Die große Ähnlichkeit wird deutlich, indem die Elemente der Standard-Transitionsmatrix eines KALMAN-Filters über eine endliche Reihenentwicklung unter der Nutzung des PASCALschen Dreiecks in die Parameter eines AR[p]-Modells überführt werden können.

[26] Engl.: *infinite impulse response system* 'System mit unendlicher Impulsantwort' (UNBEHAUEN 2002, S. 42).

4 Abbildung von Systemen durch Modelle

„Imagination
is more important than
knowledge.“
Albert Einstein

Um ein komplexes reales System verständlich werden zu lassen, wird eine Vereinfachung des Systems entworfen (Abb. 4.1). Das bedeutet, ein System wird durch den Operator $A[\cdot]$ in ein anderes abgebildet. Dieses durch Annahmen oder Bedingungen vereinfachte System nennt man ein *Modell* (DIN 19226-1:02/1994). Das Modell kann das Verhalten eines Systems einerseits in seinem gegenwärtigen Zustand aus seiner Vergangenheit heraus beschreiben, es andererseits aber auch in seiner zukünftigen Entwicklung vorhersagen.

4.1 Modelltypen

Modelle lassen sich in vielerlei Hinsicht einteilen. Üblicherweise geschieht die Einteilung in *allgemeine*, *physische*, *mathematische* oder *logische* Modelle.
Ein *physisches* Modell[1] ist ein körperliches Abbild eines Systems, welches in Bezug auf die Fragestellung die gleichen oder zumindest ähnliche Eigenschaften besitzt wie das abzubildende System. Das Modell unterscheidet sich zumeist in der räumlichen Ausdehnung oder in der Anzahl der räumlichen Dimensionen von dem realen physischen System.
Die *mathematischen* Modelle lassen sich in zwei Modellgruppen unterteilen (Vapnik 1998, S. 21). Ein *imitierendes* Modell versucht die Beziehungen zwischen dem Systemeingang und dem Systemausgang bestmöglich abzubilden. Aufgrund dieser Eigenschaften werden imitierende Modelle auch als Ein-Ausgangsmodelle, Verhaltensmodelle[2] oder nichtparametrische Modelle[3] bezeichnet, die einer *experimentellen* Analyse[4] bedürfen (Natke 1992, Kap. 1.1 u. 1.2).

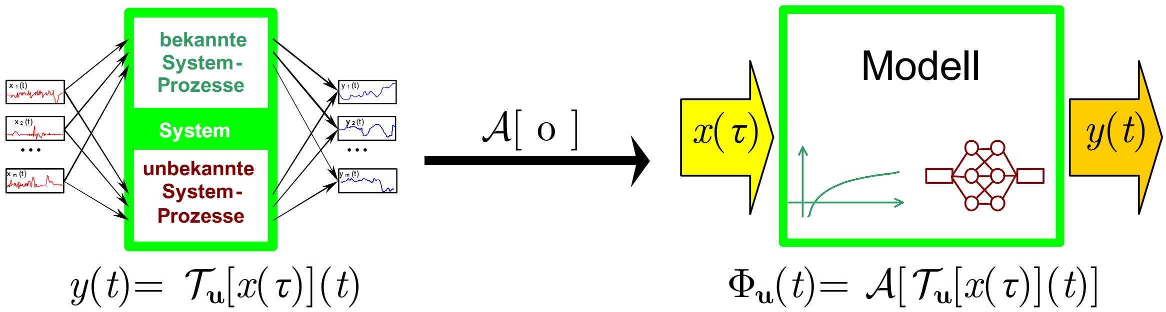


Abb. 4.1: Abbildung eines Systems mit seinen realisierten Systemeingängen und Systemausgängen in ein Modell.

[1] Abweichend wird hier auch von einem *physikalischen* Modell gesprochen (DIN 19226-1:02/1994).
[2] Vergleiche hierzu auch (Heine 1999, S. 15).
[3] Man spricht hier auch von *nichtparametrischer Identifikation* (Welsch et al. 2000, S. 355), was jedoch nicht einem *identifizierenden Modell* nur unter bestimmten Bedingungen entspricht.
[4] Hier spricht (Natke 1992, Abb. 1.5) abweichend von einer *Identifikation*.

Imitierende Modelle kommen immer dann zur Anwendung, wenn ein System und seine Prozesse nicht von vornherein die kausal-funktionalen Zusammenhänge in den realisierten Zeitreihen offenbaren. Oftmals existiert *a priori* auch kein vollständiges funktionales Modell, das sich auf physikalische Vorüberlegungen basieren lässt. Dennoch besteht nicht selten die Aufgabe, das Prozessverhalten adäquat zu prädizieren. Ein solches Modell trifft daher zunächst keine Aussagen über das System selbst. Darüber hinaus neigt es dazu, seine Prädiktionsfähigkeit einzubüßen, wenn ein bisher unbekannter Systemeingang auftritt. Aufgrund ihrer Abhängigkeit von den Systemein- und -ausgängen werden diese Modelle zumeist *diskret* oder *diskretisiert* realisiert sein.
Ein *identifizierendes* Modell hingegen versucht den Zustand des Systems selbst bestmöglich abzubilden. Aufgrund dieser Eigenschaften werden identifizierende Modelle auch als Zustandsmodelle, Strukturmodelle[5] oder parametrische Modelle bezeichnet, die einer *theoretischen* Analyse bedürfen (Natke 1992, Kap. 1.1). Ein identifizierendes Modell bedient sich der theoretischen Gesetzmäßigkeiten, die entweder ausformuliert oder als Differentialgleichungen zur Anwendung kommen (Ogata 1978, Kap. 2-5). Damit ist die Modellbildung nicht mehr durch eine Diskretisierung beschränkt, sondern kann auch *kontinuierlich* erfolgen.
Handelt es sich bei der identifizierenden Beschreibung eines Systems um physikalische Gesetzmäßigkeiten, so spricht man von einem *physikalisch-mathematischen* Modell. Für Beschreibung von Prozessen der Erdoberfläche wird die weitere Unterscheidung in *kinematische* und *dynamische* Modelle vorgenommen. Während sich ein kinematisches Modell allein der Nutzung der Geometrie und ihrer Ableitungen nach der Zeit bedient, fließen in ein dynamisches Modell die physikalischen Gesetzmäßigkeiten mit ein.
Die Statistik unterscheidet ihre mathematischen Modelle in Anlehnung an die Prozesse (Kap. 3.2) auch in *funktionale* und *stochastische* Modelle (Niemeier 2002). Das funktionale Modell beschreibt hierbei häufig in linearisierter Form die Zusammenhänge der unverfälschten Systemein- und -ausgänge. Die Ergänzung um ein stochastisches Modell soll hierbei der zufälligen Verfälschungen von Systemein- und -ausgängen gerecht werden.
Ein *logisches* Modell wird auf der Basis von Regeln errichtet (Van Brussel 2002, S. 1.4f) und stellt damit zunächst nur ein *qualitatives Modell* dar. Die Regeln sollen hierbei möglichst allgemeingültig sein.
In der Praxis lassen sich die Regeln nur für einen speziellen Fall erstellen (Bothe 1995, S. 120) und bilden so ein *Expertensystem* (Zimmermann 1996, S. 173ff). Die häufig linguistischen Regeln dieses Expertensystems werden entweder durch einen Satz von Hypothesen vorgegeben oder aus der Dateninferenz gewonnen (Bothe 1995, S. 126ff).
Die vorgenommene Einteilung in Modelltypen erlaubt es, Modelle von Modellen[6] zu entwerfen, was einer zweimaligen Abbildung eines Systems entspricht.

4.2 Modelldimensionen

Ein wesentliches Kriterium eines mathematischen Modells ist die Dimension des abzubildenden Systems. Während ein System und sein Prozess auf vielen Systemeingängen basiert und mehrere Systemausgänge erzeugt, kann ein Modell von vornherein auf die Abbildung der Wechselwirkung genau eines Systemeingangs mit wiederum genau einem Systemausgang ausgerichtet sein. Man spricht hierbei von einem *single-input, single-output* Modell oder kurz SISO (Miima 2002, S. 9). Eine derartige Modellbildung ist an eine bedingte Realisierung der Prozesse geknüpft. Außerdem kann ein SISO-Modell nur lineare Systeme (Kap. 3.1.2) oder inkremental lineare Systeme (Kap. 3.1.3) erschöpfend beschreiben. In der geodätischen Praxis entscheidet der Messaufbau maßgeblich, ob eine derartige vereinfachte Modellbildung hinreichend ist (Heine 1999, S. 17).

[5] Elastomechanische Systeme heißen im Englischen dementsprechend *structures* (Natke 1992, S. 1)

[6] Ein Wellenkanal stellt zunächst ein physikalisches Modell einer Küste dar, von dem sich seinerseits mathematische Strömungsmodelle ableiten lassen.

In den meisten Fällen werden eine Vielzahl von Systemeingängen einen Systemausgang bilden. Damit entfallen die Linearitätsbedingung (3-3) und (3-4). Man spricht hierbei von einem *multiple-input, single-output* Modell oder kurz MISO.
Die Systembeziehungen lassen sich allerdings stabiler und schneller in einem Modell erfassen, wenn dieses mehrere Systemausgänge berücksichtigt. Die Anzahl der mathematischen Bedingungen steigt und richtige Lösungen für die Systembeziehungen lassen sich schneller eingrenzen. Dabei spricht man von einem *multiple-input, multiple-output* Modell oder kurz MIMO.
Eine neuere Entwicklung stellen sogenannte *demixer* dar. Sie sind am einfachsten als *zero-input, multiple-output* Modelle zu charakterisieren. Die Aufgabe besteht hierbei nicht nur in der Beschreibung der Systembeziehungen, sondern auch in der Entmischung der unbekannten Systemeingänge (Haykin 2002, S. 766ff).

4.3 Statische Modelloptimierungen

Die Abbildung $\mathcal{A}[\,\cdot\,]$ eines Systems $\mathcal{T}[\,\cdot\,](t)$ erfolge durch die Modellfunktion $\Phi_{[\,\cdot\,]}(t)$. Dabei sei es hier zunächst unerheblich, ob es sich um einen imitierenden oder identifizierenden beziehungsweise einen parametrischen oder nichtparametrischen Operator (Kap. 4.1) handelt. Ein imitierendes Modell kann numerisch unabhängig von einem Systemeingang $x(t)$ sein, da es diesen auch *implicit* beschreiben kann[7]. Das Modell soll nun in Abhängigkeit seiner Modellparameter u einen Systemausgang $y(t)$ bestmöglich beschreiben, was durch die Forderung

$$\Phi_{\mathbf{u}}(t) - y(t) \Rightarrow \min \text{ mit } \Phi_{\mathbf{u}}(t) = \mathcal{A}[\mathcal{T}[\mathbf{u}](\mathbf{t})] \tag{4-1}$$

zum Ausdruck gebracht wird. Die Realisierung $y(t)$ des Systemausgangs $Y(t)$ enthält aber nicht nur die deterministische Antwort auf stochastische Systemeingänge. Da das Messsystem seinerseits ein System ist, kommt es zur Mischung des Systemausganges des beobachteten Objektes mit der deterministischen Antwort der auf das Messsystem wirkenden stochastischen Systemeingänge. Damit ist ausgeschlossen, dass das Minimum in (4-1) den Wert null erreicht. Gleichermaßen wird klar, dass an die Stelle einer eindeutigen Lösung eines Gleichungssystems, eine Schätzung treten muss, da alle $y(t)$ zur Bestimmung der Parameter u herangezogen werden müssen.

4.3.1 Die Zielfunktion

Die *Zielfunktion*[8] $\Omega(\,\cdot\,)$ beschreibt die verschiedenen Kriterien, die zu einer Optimierung herangezogen werden müssen. Die Bestimmung einer optimalen Lösung, sofern sie existiert, kann nur mit ihrer Hilfe erfolgen, womit (4-1) nun zu

$$\Omega(\Phi_{\mathbf{u}}(t) - y(t)) \Rightarrow \min \tag{4-2}$$

erweitert werden muss. Den Kern der Zielfunktion

$$\Omega(\Phi_{\mathbf{u}}(t) - y(t)) = \mathcal{K}\,[\,p(t) \cdot \varrho(v(t))\,] \text{ mit } v(t) = \Phi_{\mathbf{u}}(t) - y(t) \tag{4-3}$$

bildet die mit p gewichtete *Verlustfunktion*[9] über den Verbesserungen $\varrho(v)$ (Niemeier 2002, S. 186f) in Verbindung mit einem Kriteriumsoperator[10] $\mathcal{K}[\,\cdot\,]$. Die in der Geodäsie am häufig-

[7] Ein Gezeitenmodell kann aus Messdaten parametrisiert werden, ohne dass zuvor die wirkenden Kräfte in das Modell eingeführt werden (Heinert & Riedel 2007).

[8] Die Zielfunktion wird alternativ auch als *Fehlerkriterium*, *Verlustfunktion* oder *Zielfunktional* bezeichnet (Natke 1992, S. 80).

[9] Die Verlustfunktion ist in der Geodäsie eher als Norm bekannt. So besagt die L2-Norm, dass es sich um die quadrierten Differenzen handelt (Niemeier 2002, S. 325).

[10] Der Kriteriumsoperator sei hier definiert, da er sozusagen die Ausführungsbestimmung der Verlustfunktion darstellt. Dadurch, dass diese Definition bisher weitgehend unterblieben ist, ergeben sich unterschiedliche Darstellungen der Ziel- und der Verlustfunktion. Vergleiche Niemeier (2002, S. 186f), Natke (1992, S. 80) und Grafarend (1979, S. 41).

sten verwendete Zielfunktion ist das Minimum der mit p gewichtete Fehlerquadratsumme[11]

$$\Omega(v)_{\mathrm{L}_2} = \sum p \cdot v^2 \text{ mit } \varrho(v) = v^2 \text{ und } \mathcal{K}[\,\cdot\,] = \sum p \cdot \varrho(v). \tag{4-4}$$

Hierbei repräsentiert das Gewicht p eine stochastische Vorinformation, die als stochastisches Modell in die Optimierung mit eingeführt wird.

Das Inkrement der Verlustfunktion des Typs $\varrho(v) = |v|^s$ mit dem Exponenten $s \in \mathbb{Q}$ lässt sich mit verschiedenen Kriteriumsfunktionen ($\sum(\,\cdot\,)$, med$(\,\cdot\,)$, max$(\,\cdot\,)$, ...) verknüpfen. Damit ergeben sich bereits bekannte Zielfunktionen

$$\begin{aligned}
\Omega_{\mathrm{L}_1}(v) &= \sum p \cdot |v|, \\
\Omega_{\mathrm{L}_{\max}}(v) &= \max\left(p \cdot\ |v|\right), \\
\Omega_{\mathrm{L}_{\mathrm{ms}}}(v) &= \operatorname{med}\left(p \cdot\ |v|\right), \\
\Omega_{\mathrm{L}_{\mathrm{ms2}}}(v) &= \operatorname{med}\left(p \cdot\ v^2\right).
\end{aligned}$$

Da einige Verlustfunktionen auf Intervallen gebildet werden (Anhang A.3), bietet sich die Normierung der Verbesserung mithilfe des Skalenparameters[12] σ_y an (Borutta 1988, S. 65). Es gilt für die vollständige Zielfunktion

$$\Omega(t) = \mathcal{K}\left[p \cdot \varrho\left(\frac{\Phi_{\mathbf{u}}(t) - y(t)}{\sigma_y}\right)\right]. \tag{4-5}$$

4.3.2 Die Einflussfunktion

Die Abhängigkeit der Verlustfunktionen von der einzelnen Verbesserung v werden mithilfe ihrer ersten Ableitung

$$\Psi(v) = \frac{\partial \varrho(v)}{\partial v} \tag{4-6}$$

diskutiert. Wenn für die resultierende *Einflussfunktion* oder $\Psi(v)$-Funktion die Beschränktheit

$$|\Psi(v)| \ll \infty \ \forall\ v$$

erfüllt ist, gilt die erzeugende Verlustfunktion als robust. Die Einflussfunktion von $\varrho(v) = v^2$ ist dagegen nicht beschränkt (Abb. 4.2). Verschiedene alternative Ψ-Funktionen[13] (Anhang A.3) sollen diesen Mangel beheben (Schlittgen & Streitberg 1997, S. 20ff).

Der Huber-Schätzer $\varrho_{\mathrm{Hu}}(v)$ erweist sich als nicht hinreichend robust gegen extreme Ausreißer. Die deshalb entwickelten *redescending Ψ-functions* hingegen sind unzureichend für Optimierungen, denen geeignete Startwerte fehlen, denn die Zielfunktion besitzt für diese Bereiche gemäß der Einflussfunktionen Ψ_{Tu} und Ψ_{Ha} eine erste Ableitung mit dem Wert null. Die Näherung an ein Optimum ist damit ausgeschlossen.

Die Nutzung einer robusten Verlustfunktion hat ihre Berechtigung, wenn die theoretische Verteilung der Differenzen v zwischen Modell und realisiertem Systemausgang entweder mehr als ein Maximum besitzt oder nicht achsensymmetrisch ist. Letzteres ließe sich beispielsweise durch die Nutzung einer asymmetrischen Ψ-Funktion berücksichtigen.

Der Wechsel der Verlustfunktion $\varrho(v)$ einschließlich des Kriteriumsoperators $\mathcal{K}[\,\cdot\,]$ während einer Optimierung kann die globale Optimierung begünstigen, falls bereits ein unplausibles lokales Optimum erreicht worden ist (Kap. 4.3.5).

---

[11] „*Systema itaque maxime probabile valorum incognitarum* p, q, r, s etc. *id erit, in quo quadrata differentiarum inter functionum V, V', V"* etc. *valores obseruatos et computatos summam minimam effierunt*, siquidem in omnibus obseruationibus idem praecisionis gradus praesumendus est." Zitat in Liber secundus, sectio tertia Determinatio orbitae obseruationibus quotcunque quam proxime satisfacientis (Gauss 1809, S. 213 §179).

[12] Im Falle der Normalverteilung ist hier die Standardabweichung einzusetzen.

[13] Die Gruppe der zugehörigen Verlustfunktionen wird auch als *M-Schätzer* bezeichnet (Jäger et al. 2005, Kap. 4.5).

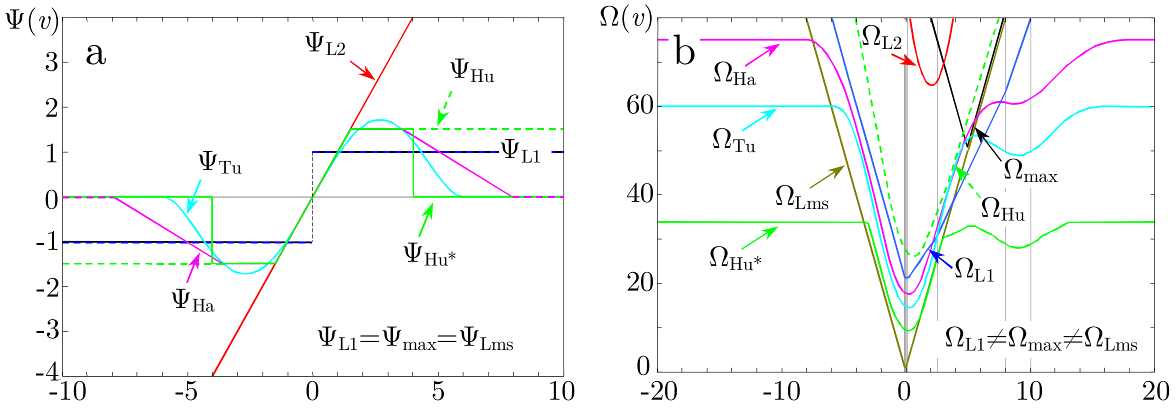


Abb. 4.2: a) Ψ-Funktionen der Standardzielfunktion L2, der robusten Zielfunktionen L1, Lms und Max, sowie die Funktionen von TUKEY und HAMPEL sowohl die zurückfallende Ψ_{Hu^*} als auch die beschränkte Funktion von HUBER Ψ_{Hu} und schließlich die Ψ-Funktion des Ln-Schätzers. Die Reaktion der zugehörigen Zielfunktionen $\Omega(v)$ mit $\varrho(v)$ in b) kann durch die Schätzung eines Mittelwertes aus einer kontaminierten Stichprobe von zehn Werten (graue Striche) verdeutlicht werden.

4.3.3 Die Risikofunktion

In der *statistischen Lerntheorie* (VAPNIK 1998) ist die Betrachtung der Risikofunktion[14]

$$R(\mathbf{u}) = \int \varrho(v(t))\partial P\{\mathbf{x}|y\} = \int \varrho\left(\frac{\Phi_{\mathbf{u}}(t) - y(t)}{\sigma_y}\right)\partial P\{\mathbf{x}|y\} \tag{4-7}$$

als Integral der Verlustfunktion (Kap. 4-3) über die bedingte Verteilung $P\{\mathbf{x}|y\}$ des realisierten Systemeingangsvektors $\mathbf{x}$ zu einem realisierten Systemausgang y das Kernanliegen. In der Praxis liegt die Schwierigkeit aber in der Unkenntnis dieser Verbundverteilung. Daher muss auf der Grundlage der verfügbaren Paare $\{\mathbf{x}^{\mathcal{P}}, \mathbf{y}^{\mathcal{P}}\}$ von realisierten Eingangs- und Ausgangsvektoren die Minimierung des empirischen Risikos gelingen.
Diese Betrachtung führt zu der zunächst etwas erstaunlichen Erkenntnis, dass die Bedeutung der Verlustfunktion $\varrho(v)$ und damit auch die ihrer Ableitung, der Einflussfunktion $\Psi(v)$, gegenüber der Wahl der Schar der Modellfunktionen $\mathcal{F}_{\mathbf{u}}(t) \ni \Phi_{\mathbf{u}}(t)$ deutlich zurückfällt (HAYKIN 1999, S. 91). *Expressis verbis*: Eine robuste Schätzung kann kein fälschlich gewähltes funktionales Modell korrigieren.

4.3.4 Das Identifikationsproblem

Der Nachweis für die Bedeutung der richtig gewählten Modellfunktion $\Phi_{\mathbf{u}}(t)$ lässt sich anhand des empirischen Risikos führen. Es gehöre der Systemoperator $\mathcal{T}[\,\cdot\,](t)$ einer Funktionsschar $\mathcal{F}_{\mathbf{u}^*}(t)$ an (Abb. 4.3). Dann lässt es sich beispielhaft für die quadratische Verlustfunktion ϱ_{L_2} zeigen, dass die Risikofunktion

$$R(\mathbf{u}^*) = \sum (y(t) - \Phi_{\mathbf{u}}(t))^2 + \sum (\mathcal{F}_{\mathbf{u}^*}(t) - \Phi_{\mathbf{u}}(t))^2 \tag{4-8}$$

sowohl von der stochastischen Verfälschung des realisierten Systemausgangs $y(t)$ direkt abhängig ist, als auch im gleichen Maße von der Zugehörigkeit der Modellfunktion $\Phi_{\mathbf{u}}(t)$ zur Funktionsschar $\mathcal{F}_{\mathbf{u}^*}(t)$ (VAPNIK 1998, S. 26ff). Gehört die Modellfunktion $\Phi_{\mathbf{u}}(t)$ ebenfalls der Funktionsschar $\mathcal{F}_{\mathbf{u}^*}(t)$ an, so gilt

$$\sum (\mathcal{F}_{\mathbf{u}^*}(t) - \Phi_{\mathbf{u}}(t))^2 = 0 \;\forall\; \mathbf{u} = \mathbf{u}^*. \tag{4-9}$$

[14] Im Deutschen ist zunächst der Terminus *Kostenfunktion* gebräuchlich gewesen (WAPNIK & TSCHERWONENKIS 1979, Kap. V §1.).

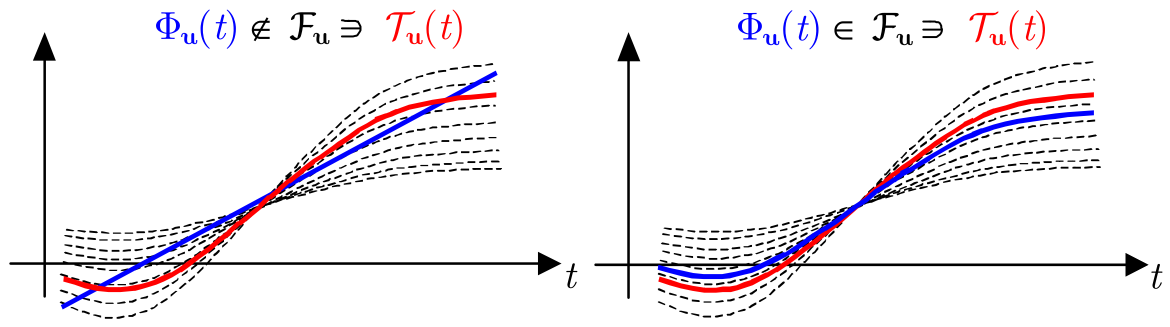


Abb. 4.3: Ein imitierendes Modell $\mathcal{F}_{\mathbf{u}}$ entspringt nicht der gleichen Funktionsschar $\Phi_{\mathbf{u}}(t)$ wie der Systemoperator $\mathcal{T}$ (links). Die Funktionsschar eines identifizierendes Modell ist gleich der des Systemoperators (rechts).

Dieser Idealfall ist für ein imitierendes Modell mit $\Phi_{\mathbf{u}}(t) \notin \mathcal{F}_{\mathbf{u}^*}(t)$ (Abb. 4.3) *a priori* ausgeschlossen. Wenn auch eine Anpassung innerhalb der Wertebereiche X und Y des realisierten Systemein- und -ausgangs für ein imitierendes Modell gelingt, so ist die Gültigkeit außerhalb dieses Wertebereiches nicht mehr gewährleistet[15].

4.3.5 Methoden der numerischen Optimierung

Unter dem Begriff *Operations Research* sind eine ganze Reihe von Methoden der numerischen Optimierung zusammengefasst, wie sie heute überwiegend in den Wirtschaftswissenschaften ihre Anwendung finden. Einen wesentlichen und auch für die Geodäsie nicht uninteressanten Anteil stellen die Methoden der *linearen, nichtlinearen, dynamischen und ganzzahligen Optimierung*[16] dar. Diese Auswahl aus den Optimierungsverfahren lösen Minimierungs- oder Maximierungsaufgaben. In der geodätischen Praxis sind wir daran gewöhnt, die optimierte Lösung einer Fragestellung auf eine Minimierungsaufgabe zurückzuführen. Im Standardfall handelt es sich dabei um eine vermittelnde Ausgleichungsrechnung nach der Methode der kleinsten Fehlerquadratsumme (Niemeier 2002).
Die Methoden der numerischen Optimierung sind eher verwandt mit der bedingten Ausgleichung mit Unbekannten[17] und bieten Alternativen zu den üblichen in der Geodäsie verwendeten Lösungswegen an: insbesondere für die Lösung nichtlinearer oder ganzzahliger Probleme.

Im Zweiten Weltkrieg standen alle Logistiker des Militärs vor der Aufgabe, ihre Material- und Truppentransporte so effizient wie möglich zu gestalten. Die hierzu von den Briten und US-Amerikanern entwickelten Berechnungsmethoden wurden nach dem Krieg durch die Wirtschaft übernommen und erweitert (Hillier & Liebermann 2002, S. 1ff). Zu dem Umfang der *Operations Research* gehören neben den Methoden der Optimierung die Netzplantechnik, die Warteschlangentheorie, die Lagerhaltungstheorie, die Spieletheorie und eine Reihe von Prognoseverfahren (Domschke & Drexl 2002; Grundmann 2002; Hillier & Liebermann 2002; Rardin 1998; Winston 1994). Die Netzplantechnik hat bereits zur wirtschaftlichen Optimierung von Netzmessungen Einzug in die Geodäsie gehalten. Die Nutzungsmöglichkeiten insbesondere der numerischen Methoden für geodätische Fragestellungen sind, obgleich sie den bekannten Ausgleichungsmethoden vielfach stark ähneln, hingegen doch weitgehend unbekannt (Mautz 2001, S. 23).

[15] Das Modell der Newtonschen Mechanik hat nur Geltung unter den Verhältnissen, wie sie Menschen innerhalb ihrer „Wertebereiche“ wahrnehmen können, dennoch hat dieses Modell keine Allgemeingültigkeit.

[16] In der Fachliteratur wird der Begriff der *linearen, nichtlinearen, dynamischen* und *ganzzahligen Programmierung* gleichwertig verwendet.

[17] In diesem Kontext: Parameter.

Die Parameterschätzung einer geodätischen Prozess- oder Deformationsanalyse (HEINERT ET AL. 2004) lässt sich immer auf eine geschlossene Minimierungsaufgabe zurückführen. Über die Lösbarkeit dieser Aufgabe entscheidet die Menge und Qualität des Beobachtungsmaterials.

Die Lösung der Minimierungsaufgabe ist ein m-dimensionaler Parametervektor $\mathbf{u}_{\min}$. Sein Funktionswert ist das globale Minimum der Lösungsfunktion $f(\mathbf{u})$. Im Lösungsraum stellt die Funktion eine $m-1$-dimensionalen Lösungshyperfläche dar.

$$\mathbf{u}_{\min} \in \mathbb{R}^m \quad \text{für} \quad f(\mathbf{u}) \to \min \tag{4-10}$$

Handelt es sich ausschließlich um lineare Funktionen, so kommen in der *Operations Research* die Methoden der linearen Optimierung zum Einsatz. Das bekannteste Verfahren dieser Gruppe ist der sogenannte *Simplex* (HILLIER & LIEBERMANN 2002, Kap. 4).

Die meisten geodätischen Fragestellungen enthalten jedoch naturgemäß nichtlineare *Zielfunktionen*[18]. Zu diesem Zweck wird die Zielfunktion Ω_{Ls} definiert als die Funktion der vermittelten Differenzen v zwischen Messwerten l und Modellwerten Φ unter der Verlustfunktion

$$\Omega_{\text{Ls}}(\mathbf{v}) \to \min \quad \Leftrightarrow \quad f(\mathbf{u}) \to \min \quad \forall \quad \mathbf{v} = \mathbf{l} - \Phi_{\mathbf{u}}. \tag{4-11}$$

Die Modellwerte $\Phi_{\mathbf{u}}$ sind abhängig von einem Modell mit den Parametern u (NIEMEIER 2002, S. 112). Die Lösungsverfahren auf der Basis der Methode der kleinsten Fehlerquadratsumme sind auf die Linearisierung der Funktionen mit allen ihren Problemen angewiesen. Das entscheidende Kriterium stellt hierbei die Qualität der Näherung der Lösung dar. Diese muss zuvor unabhängig berechnet worden sein.

Die Methoden der nichtlinearen Optimierung[19] sind vom theoretischen Ansatz her gleich, unterscheiden sich aber doch durch ihre praktische Umsetzung. Das nachstehend erläuterte *Gradienten-Verfahren* bedient sich ebenfalls durch die Nutzung der ersten Ableitungen der Lösungsfunktionen nach den Unbekannten einer Linearisierung, wie sie prinzipiell einer vermittelnden Ausgleichung entspricht (NIEMEIER 2002, S. 135). Existieren Bedingungen, so lässt sich das *Reduzierte Gradienten-Verfahren* aus der Gruppe der *bedingten nichtlinearen Programmierung* (RARDIN 1998, S. 837 ff.) ebenso wie der Allgemeinfall der Ausgleichungsrechnung auf die LAGRANGEsche Multiplikatormethode zurückführen (BARTSCH 1999, S. 326). Es gibt Ansätze der nichtlinearen Ausgleichung, wo an die Stelle der Linearisierung verschiedene Stufen der Polygonisierung treten. Diese basieren auf der Nutzung von TAYLOR-Polynomen höherer Ordnung und der Abschätzung des Restgliedeinflusses. Sie entsprechen den hier noch zu erläuternden NEWTON- und Quasi-NEWTON-Verfahren, welche zusätzlich noch die lokale HESSE-Matrix der Lösungsfunktion berücksichtigen. Die Minimierungsaufgabe kann mithilfe der nichtlinearen Optimierung weniger abhängig von Näherungen iterativ gelöst werden. Ein stochastisches Modell kann ganz entsprechend einer Ausgleichung an die Lösungsfunktion adaptiert werden. Dabei kann die Minimierungsbedingung, sei es beispielsweise die Minimierung der Fehlerquadratsumme, des Medians der Fehlerquadrate oder einer beliebigen anderen Verlustfunktion in Kombination mit einem Kriteriumsoperator (4-3), frei gewählt werden.

Das Gradienten-Verfahren

Eines der einfachsten Verfahren der nichtlinearen Optimierung ist das sogenannte Gradienten-Verfahren[20] (HILLIER & LIEBERMANN 2002, S. 431f). Als Gradient wird das vollständige

---

[18] Die *Zielfunktion* oder *Zielfunktional* ist ein Synonym für das *Fehlerkriterium* (NATKE 1992, S. 80) bzw. die *Verlustfunktion* (NIEMEIER 2002, S. 186). Da aus der Definition einer Verlustfunktion $\sigma(v) = v^2$ zum Beispiel nicht hervorgeht, ob die Summe über die Verbesserungsquadrate fester Bestandteil der Funktion ist, soll in dieser Arbeit der Begriff *Zielfunktion* verwendet werden.

[19] Alternativ findet sich auch der Begriff *Konvexe Optimierung* (COLLATZ & WETTERLING 1966, Kap. II).

[20] Das Verfahren wird auch als NEWTON-RAPHSON-Verfahren bezeichnet (BORUTTA 1988, S. 81f).

partielle Differential einer Funktion nach allen ihren m Eingangswerten bezeichnet (MERZIGER & WIRTH 1993, S. 523). Unter der Voraussetzung, dass die Ableitungen der Zielfunktion existieren, gilt daher

$$\mathrm{grad}(f(\mathbf{u}_k)) = \nabla f(\mathbf{u}_k) = \frac{\partial f(\mathbf{u}_k)}{\partial \mathbf{u}_k} \quad \text{mit} \quad k = 1 \dots m. \tag{4-12}$$

Das Gradienten-Verfahren lässt sich nun aus dem unendlichen TAYLOR-Polynom

$$f(\mathbf{u}_{k+1}) = \sum_{i=0}^{\infty} \frac{1}{i!} \frac{\partial^i f(\mathbf{u}_k)}{\partial^i \mathbf{u}_k} (\mathbf{u}_{k+1} - \mathbf{u}_k)^i \tag{4-13}$$

entwickeln (BARTSCH 1999, S. 531). Für den Gradienten liegen nur erste Ableitungen vor. Damit lässt sich auch nur das TAYLOR-Polynom erster Ordnung

$$f(\mathbf{u}_{k+1}) = f(\mathbf{u}_k) + \frac{\partial f(\mathbf{u}_k)}{\partial \mathbf{u}_k} (\mathbf{u}_{k+1} - \mathbf{u}_k) \tag{4-14}$$

berechnen. Da für die Minimierungsaufgabe der stationäre Vektor $\mathbf{u}_{\mathrm{opt}}$ gesucht wird, an welcher der Funktionswert der Zielfunktion $f(\mathbf{u}_{\mathrm{opt}})$ ein lokales Minimum, beziehungsweise ein lokales Maximum annimmt, wird das Polynom erster Ordnung nach

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \Delta f(\mathbf{u}_k) \cdot \nabla f(\mathbf{u}_k) \tag{4-15}$$

aufgelöst. Interpretiert man die Differenz der Funktionswerte als eine Schrittweite η, so erhält man mit

$$\mathbf{u}_{k+1} = \mathbf{u}_k \pm \eta_k \nabla f(\mathbf{u}_k) \tag{4-16}$$

die allgemeine Rekursionsvorschrift des Gradienten-Verfahrens (RARDIN 1998, S. 757ff). Eine Variation dieses Algorithmus stellt das *Gradientensuchverfahren* dar. Um die aufwendige Neuberechnung der Ableitungen einzusparen, wird die Richtung des zuerst berechneten Gradient solange beibehalten, bis sich eine seiner Komponenten der null nähert. Dann wird durch die Neuberechnung des Gradienten eine Korrektur der Richtung vorgenommen (HILLIER & LIEBERMANN 2002, S. 432ff).

Das NEWTON-Verfahren

Das Gradienten-Verfahren nähert sich in kleinen linearen Schritten dem Optimum der Zielfunktion an (Abb. 4.4a). Um eine schnelle Konvergenz zu erzwingen, bietet es sich an, von einem TAYLOR-Polynom zweiter Ordnung (vergl. 4-13) auszugehen (PUSKORIUS & FELDKAMP 2001, S. 23).

$$f(\mathbf{u}_{k+1}) = f(\mathbf{u}_k) + \frac{\partial f(\mathbf{u}_k)}{\partial \mathbf{u}_k} (\mathbf{u}_{k+1} - \mathbf{u}_k) + \frac{\partial^2 f(\mathbf{u}_k)}{\partial^2 \mathbf{u}_k} (\mathbf{u}_{k+1} - \mathbf{u}_k)^2 \tag{4-17}$$

Zunächst definiere man die Schrittlänge entlang der Zielfunktion und stelle die Gleichung danach um[21]

$$\Delta f \mathbf{u}_k = \nabla f(\mathbf{u}_k) \Delta \mathbf{u}_k + \frac{1}{2} \nabla^2 f(\mathbf{u}_k) \Delta \mathbf{u}_k^2 \quad \text{mit} \quad \Delta \mathbf{u}_k = \mathbf{u}_{k+1} - \mathbf{u}_k. \tag{4-18}$$

[21] Die sich resultierende Ersatzfunktion stellt nur eine Näherung dar. Problematisch sind die Abschätzungen der Nebeneffekte, die sich aus dieser Vereinfachung ergeben (MAUTZ 2001, S. 21).

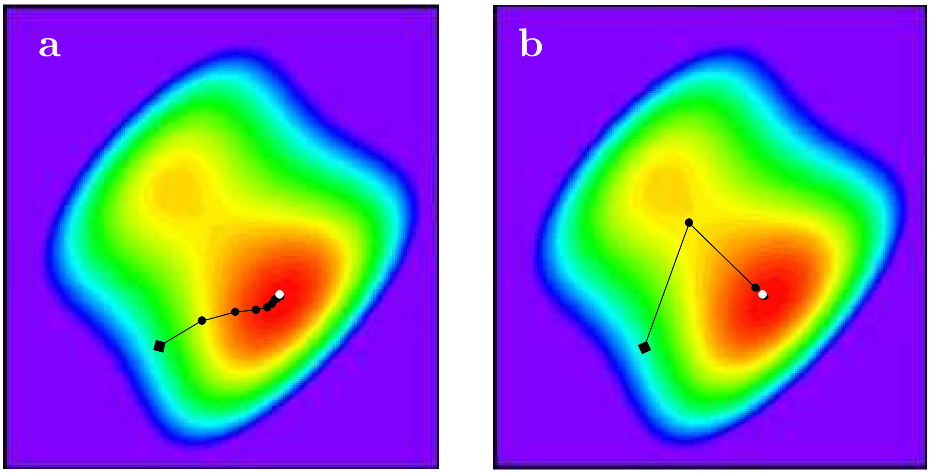


Abb. 4.4: Das Konvergenzverhalten eines Gradientenverfahrens (a) zeichnet sich durch viele Schritte auf direktem Wege zum Optimum aus. Das NEWTON-Verfahren (b) konvergiert trotz des höheren Rechenaufwandes schneller und nimmt gleichzeitig einen anderen Weg zum Optimum.

Die Ableitung nach $\Delta\mathbf{u}_k$ ermöglicht, den Einfluss der Schrittlänge des Parametervektors $\mathbf{u}_k$ auf die Zielfunktion

$$\frac{\partial \Delta f\left(\mathbf{u}_k\right)}{\partial \Delta \mathbf{u}_k} = \nabla f\left(\mathbf{u}_k\right) + \nabla^2 f\left(\mathbf{u}_k\right) \Delta \mathbf{u}_k \tag{4-19}$$

zu bestimmen. Entscheidend für den Parametervektor ist genau die Stelle, wo eine Ableitung der Ersatzfunktion den Wert null annimmt:

$$\frac{\partial \Delta f\left(\mathbf{u}_k\right)}{\partial \Delta \mathbf{u}_k} \stackrel{!}{=} 0 \tag{4-20}$$

Hier erreicht die Zielfunktion einen stationären Punkt (RARDIN 1998, S. 762). Ein weiteres Fortschreiten entlang des Gradienten führte dazu, dass man sich wieder vom Optimum entfernt. Die Ersatzfunktion (4-19) wird umgestellt und man erinnere sich, dass $\Delta\mathbf{u}_k$ in Formel (4-18) als Differenz der beiden Stellen im Parameterraum $k+1$ und k definiert worden ist:

$$-\nabla f\left(\mathbf{u}_k\right) = \nabla^2 f\left(\mathbf{u}_k\right)\left(\mathbf{u}_{k+1} - \mathbf{u}_k\right). \tag{4-21}$$

Mit der Auflösung nach $\mathbf{u}_{k+1}$ ergibt sich die Rekursionsvorschrift des NEWTON-Verfahrens mit seinem sogenannten *Newton-Schritt*

$$\mathbf{u}_{k+1} = \mathbf{u}_k \pm \left[\nabla^2 f(\mathbf{u}_k)\right]^{-1} \nabla f(\mathbf{u}_k). \tag{4-22}$$

Ein Vergleich mit der Rekursionsvorschrift des Gradienten-Verfahrens (4-16) macht deutlich, dass an die Stelle einer selbst definierbaren Schrittlänge η_k die inverse HESSE-Matrix $\mathbf{H}^{-1}(\mathbf{u}_k)$ getreten ist. Die HESSE-Matrix

$$\mathbf{H}(\mathbf{u}_k) = \nabla^2 f(\mathbf{u}_k) = \frac{\partial^2 f(\mathbf{u}_k)}{\partial u_{k,i} \partial u_{k,j}} \tag{4-23}$$

enthält alle numerisch bestimmten zweiten partiellen Ableitungen der Zielfunktion nach ihren Parametern. Das bedeutet für den Rechengang: Wie beim Gradientensuchverfahren wird die Richtung beibehalten, bis ein stationärer Punkt der Zielfunktion erreicht worden ist. Das NEWTON-Verfahren ermöglicht allerdings diesen Weg in einem Schritt zurückzulegen, was die Konvergenz erheblich beschleunigt (Abb. 4.4b).

Das Quasi-Newton-Verfahren

Ein entscheidender Nachteil des Newton-Verfahrens ist das, was zunächst als sein Vorteil gedacht war, nämlich die Nutzung der zweiten Ableitungen der Zielfunktion. Die numerisch berechneten zweiten Ableitungen können in zweierlei Weise eine schlecht konditionierte Hesse-Matrix nach sich ziehen. Einerseits können sie singulär werden und damit nicht invertierbar (Mautz 2001, S. 21f), andererseits kann die quadratische Näherung eines linearen oder genähert linearen Problems zu nur schwacher Information in der Matrix führen (Rardin 1998, S. 766). Ersteres führt zum unweigerlichen Abbruch der Iteration, letzteres kann nahezu chaotische Entwicklungen annehmen, insbesondere fernab eines lokalen Optimums. Daher ist es erforderlich, die Wirkung der Inversion der Hesse-Matrix in den genannten Problemfällen auszuschalten. Aus diesem Grund wird eine Approximation der inversen Hesse-Matrix $\mathbf{D} = \mathbf{H}^{-}$ verwendet (Anh. A.4):

$$\Delta\mathbf{u}_{k+1} = \eta_k \left[\nabla^2 f(\mathbf{u}_k)\right]^{-} \nabla f(\mathbf{u}_k) = \eta_k \mathbf{H}^{-}\left(\mathbf{u}_k\right) \nabla f(\mathbf{u}_k) = \eta_k \mathbf{D}\left(\mathbf{u}_k\right) \nabla f(\mathbf{u}_k). \tag{4-24}$$

Da zu Beginn der Iteration noch keine Informationen zur Approximation von $\mathbf{D}$ existieren und die zweiten partiellen Ableitungen nicht verwendet werden sollen, startet die Iteration mit einer Einheitsmatrix. Wird die Lösungsfunktion $f(\mathbf{u})$ im Verlauf der Iteration wieder zunehmend linear, so nähert sich $\mathbf{D}$ wieder einer Einheitsmatrix an. Im anderen Falle approximiert $\mathbf{D}$ die inverse Hesse-Matrix bestmöglich.

Das Levenberg-Marquardt-Verfahren

Für die Anpassung des Parametervektors $\mathbf{u}$ ist alternativ zu den Newton-Verfahren die vereinfachte Korrektur

$$\Delta\mathbf{u}_{k+1} = \left[\mathbf{J}^{\mathrm{T}}\left[f(\mathbf{u}_k)\right]\mathbf{J}\left[f(\mathbf{u}_k)\right] + \lambda\mathbf{I}\right]^{-1} \mathbf{J}^{\mathrm{T}}\left[f(\mathbf{u}_k)\right] f(\mathbf{u}_k) \tag{4-25}$$

mittels der schneller zu berechnenden Jacobi-Matrix $\mathbf{J}$ entwickelt worden. Das sogenannte Levenberg-Marquardt-Verfahren basiert auf der Variation zwischen einem Gradienten-Verfahren für $\lambda \gg 0$ und einem Newton-Verfahren für $\lambda \to 0$. Das Verfahren startet daher mit den Werten für λ zwischen 0 und 0,1 und wird mit fortschreitender Optimierung vergrößert (Bothe 1998, S. 179ff).

Attraktoren und Attraktionsbereiche

Der Begriff des *Attraktors* kommt ursprünglich aus der *Chaostheorie* und beschreibt das Konvergenzziel einer *fraktalen* Reihenentwicklung wie beispielsweise der bekannten komplexen Mandelbrot-Menge

$$c \in \mathbb{M} \subset \mathbb{C}\ \forall\ z_{n+1} = z_n^2 + c \ll \infty, z_0 = 0. \tag{4-26}$$

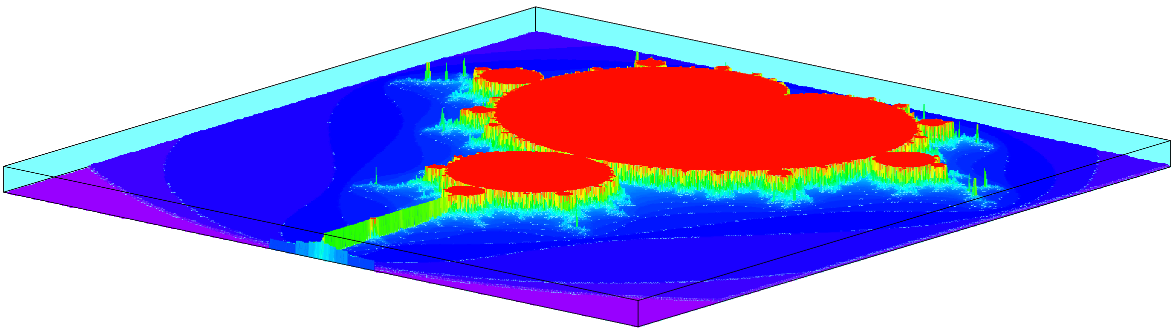


Abb. 4.5: 3D-Darstellung der Mandelbrotmenge nach 250 Iterationen.

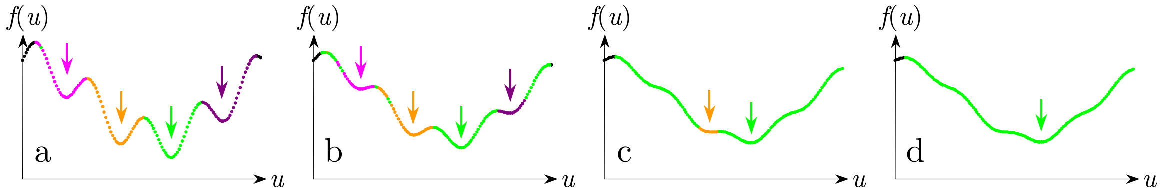


Abb. 4.6: Die Attraktoren einer Lösungsfunktion (bezeichnet durch Pfeile) und ihre zugehörigen Attraktionsgebiete.

Während für die Reihenentwicklung des Fraktals auf der Grundlage einer MANDELBROT-Menge $\mathbb{M}$ nur die Beschränktheit als Konvergenzziel zu erreichen ist, muss eine Optimierung auch zu einem lokalen Optimum führen. Die Menge, für die ein Attraktor erreicht wird, ist der zugehörige *Attraktionsbereich*. Eine solche Reihenentwicklung resultiert aufgrund ihrer Startwerte in einer *Fraktalgeometrie* im Raum der komplexen Startwerte $\mathbb{C}$ (Abb. 4.5).
Zunächst ist die Koinzidenz zwischen einer Optimierung und einer Fraktalgeometrie nicht ersichtlich. Eine Optimierung nach den verschiedenen NEWTON-Verfahren ist formal ebenfalls eine Reihenentwicklung. Eine durch starke Attraktoren gegliederte Lösungsfunktion (Abb. 4.6a) lässt nur Lösungen im nächstgelegenen lokalen Minimum zu (HAYKIN 1999, S. 674f). Im Startwertraum $\mathbb{R}^n$ lässt sich in Abhängigkeit von der Gliederung (Abb. 4.6b) einer Lösungsfunktion Ω eine ganz ähnliche Fraktalgeometrie erzielen, wenn analog auch die Anzahl der notwendigen Iterationen zum Erreichen eines der Optima mit berücksichtigt wird. Die günstige Wahl des Startpunktes kann nämlich bei einer weniger starken Gliederung das Überspringen der lokalen Minima bewirken, so dass das globale Minimum (Abb. 4.6, grün) erreicht wird. Plateaus (Abb. 4.6c) können noch als Attraktoren wirken (Abb. 4.6, orange). Eine hinreichend konvexe Lösungsfunktion (Abb. 4.6d) lässt die globale Minimumsdetektion zu, so dass sich keine Mengenunterscheidung mehr vornehmen lässt.

4.4 Autoregressive Modelle

Um das Verhalten eines dynamischen Systems (Kap. 3.1.5) nichtparametrisch modellieren zu können, ist eine Gruppe von rekursiven Algorithmen nötig, die eine große Fähigkeit zur Identifikation bieten. Die autoregressiven Modelle weisen eine formale Ähnlichkeit zu den autoregressiven Prozessen (Kap. 3.2.2) und zu den rekursiven Filtern (3-77) auf. Jene sind aber in ihrem Aufbau keineswegs so eingeschränkt, wie die Beschreibung der zugehörigen Prozesse zunächst vermuten lässt. Andererseits erfordert die starre Rekursivität dieser Algorithmen unbedingt eine äquidistante Diskretisierung der Ein- und Ausgangsgrößen.

4.4.1 Lineare AR-Modelle

Das Grundelement des Modells entspricht formal der Beschreibung eines autoregressiven Prozesses (3-37). Dabei muss berücksichtigt werden, dass nur der deterministische Subprozess durch die unbekannten Parameter u_k beschrieben wird. Das bedeutet, dass $\bar{y}_i$ hier nur von den vorangegangenen Systemausgängen abhängt (Abb. 4.7a) und dass auf diesem Wege das Gedächtnis eines dynamischen Systems (Kap. 3.1.5) modelliert wird. Daher ist es zweckmäßig, die Gleichung des diskreten autoregressiven Prozesses (3-37), wie folgt umzustellen:

$$\Phi_i(\mathbf{u}) - \varepsilon_i = y_i + v_i(\mathbf{u}) - \varepsilon_i = \sum_{k=-p}^{p} u_k y_{i-k}, \; i \in \mathbb{Z}. \tag{4-27}$$

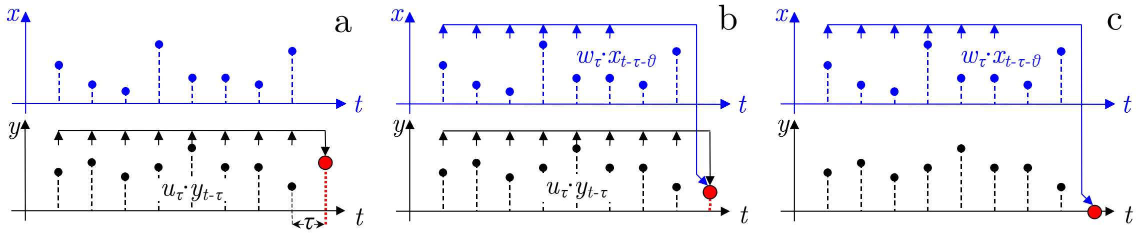


Abb. 4.7: Beziehung zwischen a) einem linearen AR-Modell, b) einem linearen ARX-Modell, das der Kernfunktion eines rekursiven Neurons entspricht und c) der Kernfunktion eines nichtrekursiven, linearen Neurons.

Die zentrale Aufgabe ist es, die unbekannten Parameter u_k so zu wählen, dass entsprechend (4-3) gilt:

$$\mathcal{K}\left[\, p_i \cdot \varrho(v_i)\,\right] \;\Rightarrow\; \text{min.} \tag{4-28}$$

Die Verbesserungen v_i enthalten zunächst noch die Modellfehler der nicht optimierten Parameter u_k. Weiterhin muss die Trennung zwischen den Fehlern des Modells und dem Verhalten der Zufallsstöße ε_i gelingen. Die Gewichte des stochastischen Modells p_i geben diese Trennung vor. Ist der Systemeingang bekannt, so kann ein AR[p]-Modell zu einem ARX[p, e]-Modell der Form

$$\Phi_i(\mathbf{u}, \mathbf{w}) = \bar{y}_i = \mathcal{N} + \sum_{k=1}^{p} u_k y_{i-k} + \sum_{\kappa=d}^{e+d} w_\kappa x_{i-\kappa-d} \tag{4-29}$$

erweitert werden (Abb. 4.7b). Der autoregressive Prozess mit externem Eingang reagiert mit der Verzögerung d und dem Bias $\mathcal{N}$ auf die Systemeingänge x_t (Bunke 1997). Es besteht eine formale Identität des Modells zu einem allgemeinen kausalen Filter (3-77) und zur Systemgleichung eines Kalman-Filters (Heunecke 1995, S. 14).

4.4.2 Nichtlineare AR-Modelle

Der einfachste nichtlineare und zugleich nichtparametrische Algorithmus ist das *nichtrekursive Neuron*, das den Grundbaustein eines *künstlichen neuronalen Netzes* bildet (Patterson 1996; Schlittgen & Streitberg 1997; Vapnik 1998; Haykin 1999). Wie bei einem ARX[0,e]-Prozess wird zunächst die Summe

$$\Phi_i(\mathbf{w}) = \mathcal{N} + \sum_{\kappa=d}^{e+d} w_\kappa x_{i-\kappa-d} \tag{4-30}$$

gebildet (Abb. 4.7c). Anschließend wird der Systemausgang

$$\bar{y}_i = \varphi(\Phi_i(\mathbf{w})) \tag{4-31}$$

durch die Anwendung der Aktivierungsfunktion $\varphi(\cdot)$ nichtlinear. Als Aktivierungsfunktionen werden vornehmlich vollständig differenzierbare Funktionen verwendet (Miima 2002, S. 40f).

Die Anpassung der als Gewichte w bezeichneten Parameter kann auch hier mithilfe des Newton-Verfahrens oder des Quasi-Newton-Verfahrens erfolgen (Kap. 4.3.5). Ein lineares ARX[p,e]-Modell (4-29) kann analog mit (4-31) zu dem nichtlinearen autoregressiven Modell mit externem Eingang NARX[p,e]

$$\bar{y}_i = \varphi\left(\mathcal{N} + \sum_{k=1}^{p} u_k y_{i-k} + \sum_{\kappa=d}^{e+d} w_\kappa x_{i-\kappa-d}\right) \tag{4-32}$$

erweitert werden (Haykin 1999, S. 734f). In dieser Form entspricht es einem einzelnen *rekursiven Neuron*.

4.5 Künstlich Neuronale Netze

Es ist eine Aufgabe, die wir täglich ausführen: wir gehen über eine Straße und je nach Verkehrsdichte müssen wir entscheiden, ob der Verkehr das Passieren der ganzen Straße oder möglicherweise nur einer Fahrbahn zulässt. Wir müssen dabei nicht selten gleich mehrere Fahrzeuge, seien es nun Autos, Motorräder oder Fahrräder, im Auge behalten. Wie aber lösen wir diese Aufgabe? Wir messen weder die Geschwindigkeit aller Fahrzeuge, noch die Breite der zu überquerenden Straße. Wir errechnen auch keine Korrekturen für Windeinfluss, Lichteinfall oder Steigungsänderungen auf den Fahrbahnen. Alle diese Parameter werden uns sogar nach der gelungenen Überquerung unbekannt bleiben. Auf welche Weise hat unser Gehirn nun die Leistung vollbracht? Erinnern wir uns: entweder wir sind mit dem zunehmenden Verkehr gewachsen oder wir haben als Angehöriger einer jüngeren Generation Verkehrsunterricht genossen. Dabei wurden wir wiederholt mit der Situation konfrontiert, eine Straße zu überqueren. Wir haben es üben und stetig wiederholen müssen. Wir haben eine Vielzahl von Einzelsituationen erlebt und in generalisierter Form gespeichert. Die Übung hat dazu geführt, dass wir die gestellte Aufgabe in kürzester Zeit lösen können. Was liegt also näher, als ein mathematisches Pendant unseres Gehirns zu erschaffen, welches in der Lage ist, zu lernen, zu speichern, zu generalisieren und – schließlich die höchste Leistung eines Modells – zu prädizieren. Das Arbeitsprinzip eines Künstlich Neuronalen Netzes entstammt den Erkenntnissen über den Aufbau des Gehirnes und ist das mathematische Abbild der Mustererkennung in der Großhirnrinde (cerebraler Kortex) von Säugetieren. In kürzester Zeit können hier Schätz- und Entscheidungsaufgaben auf chemo-elektrischem Wege gelöst werden. Man schätzt den Gesamtumfang aller ein- und ausgehenden Signale auf 150 MB/s (Spitzer 2002, S. 51ff). Künstliche Neuronale Netze (KNN) simulieren die Verarbeitungsprinzipien bei der Mustererkennung eines natürlichen Gehirns.
In den letzten zehn Jahren sind KNN verstärkt für die Datenanalyse eingesetzt worden. Ohne *a priori*-Kenntnisse über die Systemstruktur und die Zusammenhänge des zu analysierenden Prozesses haben zu müssen, können nichtlineare und komplexe Prozesse abgebildet werden. Durch einen Trainingsprozess mit Beispieldaten erfolgt eine iterative Gewichtsanpassung, so dass am Ende des Trainings der Gewichtungssatz optimiert ist. Es können sowohl Gesamtprozesse abgebildet, als auch einzelne Parameter geschätzt bzw. adaptiert werden. Theoretisch sind KNN in der Lage, bei entsprechender Netzwerkarchitektur jeden tatsächlich existierenden funktionalen Zusammenhang mit beliebiger Genauigkeit abzubilden (Patterson 1996).

4.5.1 Der Lernprozess

Grundsätzlich existieren drei verschiedene Formen des Lernens: das *überwachte*, das *bestärkende* und das *nichtüberwachte* Lernen (Patterson 1996, S. 47). Überwachtes Lernen erfordert einen *Lehrer* (Vapnik 1998, S. 20), also einen Algorithmus der jede Veränderung des neuronalen Netzes direkt beantwortet, während bei den anderen beiden Formen keine oder nur eine zeitlich und logisch eingeschränkte Reaktion erfolgt. Im Rahmen dieser Arbeit soll nur das überwachte Lernen weiter betrachtet werden.
Desweiteren existieren verschiedene Lernziele: die *Assoziation*, die *Mustererkennung* und die *Approximation* (Haykin 1999, S. 66ff). Das *autoassoziative* Lernen speichert die Qualitäten sämtlicher Systemausgänge. Damit werden sich für einige Qualitäten im Objektraum Haufenbildungen ergeben, die Aufschluss über die Eigenschaften der Systemausgänge ermöglicht. Das *heteroassoziative* Lernen speichert auch die Qualitäten sämtlicher Systemeingänge, so dass sich aus der Haufenbildung im Objektraum generalisierte Beziehungen zwischen Systemeingang und Systemausgang angeben lassen. Diese selbst organisierenden Netze oder Kohonen-Netze können zu Fuzzy-Kohonen-Clustering Netzwerken erweitert werden (Bothe 1998, S. 215ff). Diese Aufgabe ist analog zur Extraktion unscharfer Relationen (Bothe 1995, vergl. Kap. 6).

Die Mustererkennung ist eine erweiterte Assoziation, denn jene soll die assoziierten Muster im Objektraum durch Hyperflächen bestmöglich voneinander trennen, um für jeden Systemeingang eine diskrete Entscheidung zu fällen, die einem diskreten[22] Systemausgang entspricht. Die Approximation ist gleichsam die inverse Aufgabenstellung zur Mustererkennung: sollte bei einer Mustererkennung eine Hyperfläche die Muster im Objektraum bestmöglich voneinander trennen, so ist es bei einer Approximation die Aufgabe, eine Hyperfläche bestmöglich durch alle Muster im Objektraum hindurch zu legen.

4.5.2 Methoden des überwachten Lernens

Lernen nach Irrtum

Die Gewichte w, wie sie bereits bei den autoregressiven Modellen (Kap. 4.4) zu Einsatz gekommen sind, werden im Lernprozess für jeweils ein einzelnes Muster[23] $\mathcal{P}$ bestehend aus Eingangs- und Ausgangsvektoren $\{\mathbf{x}^{\mathcal{P}}, \mathbf{y}^{\mathcal{P}}\}$ angepasst. Diese Anpassung

$$\Delta w_{ij,k}^{\mathcal{P}} = -\eta \nabla \Omega^{\mathcal{P}}(w_{ij,k}) \tag{4-33}$$

mit der *Lernrate* η wird häufig nach den Methoden der numerischen Optimierung (Kap. 4.3.5) vorgenommen und heißt *gradientielles Lernen*. Das Ziel dieses Lernens besteht in der Minimierung des Irrtums $\Phi_{\mathbf{w},k}^{\mathcal{P}} - y_j^{\mathcal{P}}$, weswegen diese Methode ein *Lernen nach Irrtum* darstellt. Der Ausdruck (4-33) kann mithilfe der Definition der Zielfunktion (4-5) und Berücksichtigung der nichtlinearen Aktivierungsfunktion (4-31) zu

$$\Delta w_{ij,k}^{\mathcal{P}} = -\eta \frac{\partial \varphi\left(\mathcal{K}\left[p \cdot \varrho\left(\frac{\Phi_{\mathbf{w},k}^{\mathcal{P}} - y_j^{\mathcal{P}}}{s_y}\right)\right]\right)}{\partial w_{ij,k}} \tag{4-34}$$

erweitert werden.
Das Lernen erfolgt ganz analog gemäß des Quasi-Newton-Verfahrens oder Levenberg-Marquardt-Verfahrens (Kap. 4.3.5). Auch die Methode nach Widrow-Hoff gehört zum Lernen nach Irrtum (Haykin 1999, S. 52). Diese sogenannte *Delta-Regel*

$$\Delta w_{ij,k}^{\mathcal{P}} = -\eta \mathcal{K}\left[p \cdot \varrho\left(\frac{\Phi_{\mathbf{w},k}^{\mathcal{P}} - y_j^{\mathcal{P}}}{s_y}\right)\right] x_i^{\mathcal{P}} = \eta \delta_k^{\mathcal{P}} x_i^{\mathcal{P}} \tag{4-35}$$

weist die Differenz zwischen modelliertem und realisiertem Systemausgang $\Phi_{j,k}^{\mathcal{P}} - y_j^{\mathcal{P}}$ entsprechend der Stärke des Systemeinganges x_i als Korrektur dem Gewicht w_{ij} zu.

Korrelatives Lernen

Das Hebb*sche Lernen* ist die als erstes mathematisch umgesetzte und einfachste Form des *korrelativen Lernens*. Das Produkt $x_i^{\mathcal{P}} y_j^{\mathcal{P}}$ ist vergleichbar mit dem ersten Element der Kreuzkovarianzfunktion (3-49) und damit ein Indikator für ein proportional oder antiproportional korreliertes Verhalten. Korrelierte Ein- und Ausgangspaare sollen entsprechend der Lernanweisung

$$\Delta w_{ij,k}^{\mathcal{P}} = \eta x_i^{\mathcal{P}} y_j^{\mathcal{P}} \tag{4-36}$$

[22] *Diskret* ist hier nicht im Sinne der zeitlich oder räumlich diskreten Realisierung zu verstehen, sondern der Systemausgang selbst ist diskret. Beispiel: anhand eines Digitalbildes soll entschieden werden, aus welchen Ziffern die Postleitzahl auf einem handschriftlich ausgefüllten Briefumschlag besteht.

[23] Engl.: *pattern*.

gestärkt werden (BOTHE 1998, S. 128). Im Vergleich zur Definition der Kreuzkovarianzfunktion (3-49) fällt aber die fehlende Normierung auf jeweils einen Mittelwert und eine Standardabweichung auf, so dass das verallgemeinerte *korrelative Lernen*

$$\Delta w_{ij}^{\mathcal{P}}(k) = \eta \frac{y_j^{\mathcal{P}} - \bar{y}}{s_y} \cdot \frac{x_i^{\mathcal{P}} - \bar{x}}{s_x} \tag{4-37}$$

erfolgreicher ist, denn Neuronen, die gemäß (4-36) trainiert werden, sind schnell gesättigt. Damit reagieren sie nicht mehr auf weitere Eingänge und sind für das weitere Lernen verloren (HAYKIN 1999, S. 57). Diese Problematik beschränkt sich aber keinesfalls nur auf das korrelative Lernen. Vielmehr sollten die realisierten Systemein- und -ausgänge generell auf Werte zwischen 1 und -1 normiert werden, um die vollständige Sättigung von Neuronen im Netzwerk zu vermeiden.

Konkurrierendes Lernen

Eine modifizierte Form des korrelativen Lernens ist das *konkurrierende Lernen*, bei dem nur jeweils ein einzelnes Paar von Ein- und Ausgängen $\left\{x_i^{\mathcal{P}}, y_j^{\mathcal{P}}\right\}$ im Lernschritt berücksichtigt wird (PATTERSON 1996, S. 44). In jeder Schicht wird daher nur das Gewicht des aktivsten Neurons gemäß

$$\Delta w_{ij}^{p}(k) = \begin{cases} \eta(x_i - w_{ij}) & \Phi^p(k) = 1 \\ 0 & \Phi^p(k) = 0 \end{cases} \tag{4-38}$$

angepasst[24] (HAYKIN 1999, S. 58f). Diese Art des Lernens ist besonders für diskrete Systemausgänge in der Mustererkennung gebräuchlich.

4.5.3 Aufbau eines Neurons

Die kortikale Zelle empfängt an den Dendriten chemo-elektrische Signale ihrer Nachbarzellen. Erreicht die Summe dieser Signale einen zellspezifischen Schwellwert, sendet die erregte Zelle ihrerseits ein chemo-elekrisches Signal über das Axon zu ihren Nachbarzellen. Es ist durchaus möglich, dass zwei Zellen auf diese Weise eine Rückkopplung bilden können (BOTHE 1998, S. 77). Aus der Neurophysiologie ist es bekannt, dass häufig benutzte Zellverbindungen über die Synapsen verstärkt werden. Diese Flexibilität ermöglicht das Lernen (SPITZER 2002, 44ff u. 64f).

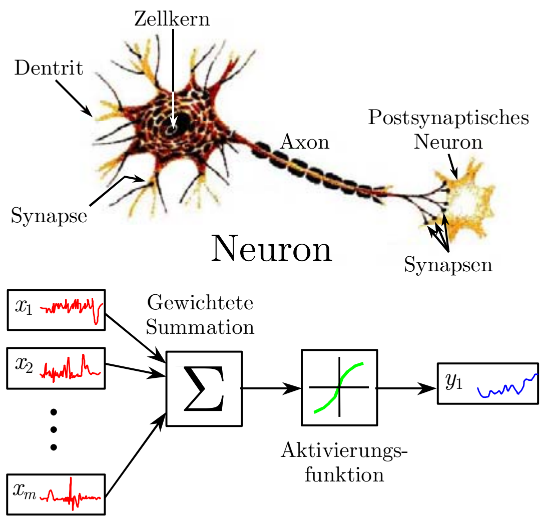


Abb. 4.8: Das natürliche und das künstliche Neuron (SITTER 2001).

Die Grundbausteine eines künstlichen neuronalen Netzes sind ebendiese Neuronen. Daher sind sie nach den gleichnamigen Gehirnzellen benannt, deren Funktion sie in idealisierter Form kopieren. Bei der mathematischen Umsetzung entspricht die Aktivierungsfunktion $\varphi(\,\cdot\,)$ dem Reaktionsverhalten des Zellkerns (Abb. 4.8) und die Stärke der synaptischen Kontakte ist durch die numerischen Gewichte w realisiert (BOTHE 1998, Kap. 3.1).

[24] In der englischsprachigen Literatur wird daher von *the-winner-takes-all neurons* gesprochen.

Die Kernfunktionen des Neurons

Die klassische Kernfunktion eines nichtlinearen, nichtrekursiven Neurons entspricht dem ARX$[0, e]$-Modell (4-30). Diese Kernfunktion hat eine Analogie zur Funktionsweise der kortikalen Zellen. Eine Rekursion, wie sie dem ARX[p,e]-Modell zu eigen ist, kann entweder innerhalb einer einzelnen Zelle oder auch über mehrerer Zellen generiert werden. Daneben funktioniert das RBF[25]-Neuron

$$\Phi_i(w) = w\varphi\left(\|\mathbf{x}_i - \boldsymbol{\chi}\|\right) \tag{4-39}$$

auf der Grundlage der euklidischen Norm $\|\cdot\|$ (Haykin 1999, S. 264). Über die Gewichtung zwischen den Systemeingängen $\mathbf{x}$ entscheidet deren Abstand vom Stützvektor des Neurons $\boldsymbol{\chi}$. Es existiert hier eine Analogie zu den sogenannten *Ortsneuronen* im Hippocampus (Spitzer 2002, S 24). Diese dienen der Orientierung und werden gerade dann aktiv, wenn sich ein Lebewesen dem Ort nähert, den das RBF-Neuron repräsentiert.

Die Aktivierungsfunktionen des Neurons

Die natürliche Reaktion eines kortikalen Neurons folgt der Schwellwertfunktion[26] (Tab. 4.1). Übersteigt die Summe aller eingehenden Signale von den Nachbarzellen einen zelltypischen Schwellwert, so wird diese Zelle ihrerseits aktiv[27]. Die Schwellwertfunktion ist besonders geeignet für das konkurrierende Lernen (Kap. 4.5.2) bei der Mustererkennung. Eine Nutzung für das gradientielle Lernen ist aufgrund der Undifferenzierbarkeit nicht sinnvoll.
Auch die beschränkte lineare Aktivierung eignet sich aus dem gleichen Grunde nicht für das gradientielle Lernen. Allerdings kann sie auch für das konkurrierende Lernen in der letzen Schicht eines Netzwerkes nicht uneingeschränkt verwendet werden, da diese Funktion auch Werte zwischen 0 und 1 zulässt, die für eine diskrete Entscheidung nicht vorgesehen sind.
Für die meisten Anwendungen ergeben sich typische Kombinationen aus Kern- und Aktivierungsfunktionen (Miima 2002, S. 41ff). Zum einen werden Neuronen eines mehrschichtigen gradientiell, überwacht lernenden Netzes zur Musterkennung häufig mit der Sigmoidfunktion

$$\Phi_i(\mathbf{w}) = \left(1 + e^{a\left(\mathcal{N} + \sum_{\kappa=d}^{e+d} w_\kappa x_{i-\kappa-d}\right)}\right)^{-1} \tag{4-40}$$

oder mittels des Tangens hyperbolicus

$$\Phi_i(\mathbf{w}) = \tanh\left(\mathcal{N} + \sum_{\kappa=d}^{e+d} w_\kappa x_{i-\kappa-d}\right) \tag{4-41}$$

aktiviert. Zum anderen werden RBF-Neuronen eines dreischichtigen gradientiell, überwacht lernenden Netzes

$$\Phi_i(\mathbf{w}) = e^{-a(\|\mathbf{x}_i - \boldsymbol{\chi}_k\|)^2} + w_{i0} \tag{4-42}$$

zur Approximation häufig mit der Gauss-Funktion (Tab.4.1) aktiviert.

Die Lösung des XOR-Problems

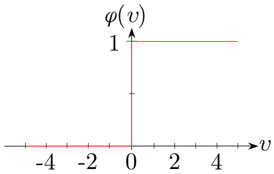
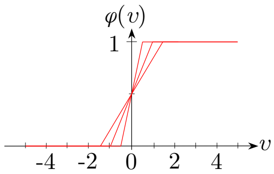
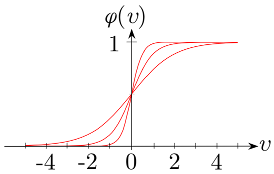
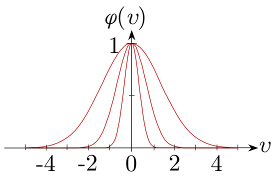
Der Boole-Operator des exklusiven Oders (XOR) führt zu Funktionswerten, die sich im zweidimensionalen Zustandsraum (Abb. 4.9a) nicht linear separieren lassen (Haykin 1999, S. 175ff

[25] Engl.: *radial basis-function* 'Radiale Basisfunktion'.

[26] Diese Funktion wird auch Heaviside-Funktion oder Θ-Funktion genannt.

[27] Die Schwellwertfunktion für die sogenannte Boltzmann-Maschine wird zwischen 1 und -1 definiert (Haykin 1999, S. 60f).

Tab. 4.1: Aktivierungsfunktionen

Schwellwert-Aktivierung	$\varphi_{\mathcal{H}}(\upsilon) = \frac{1}{2}\left(\frac{\upsilon}{\lvert\upsilon\rvert} + 1\right)$	
beschränkt lineare Aktivierung	$\varphi_{[/]}(\upsilon) = \begin{cases} 0 & \forall \quad \upsilon \leq \frac{1}{2m} \\ m\upsilon + \frac{1}{2} & \forall \quad -\frac{1}{2m} < \upsilon \leq \frac{1}{2m} \\ 1 & \forall \quad \upsilon > \frac{1}{2m} \end{cases}$	
sigmoidale Aktivierung	$\varphi_{\Sigma}(\upsilon) = \frac{1}{1+e^{-a\upsilon}}$	
GAUSS Aktivierung	$\varphi_{N}(\upsilon) = e^{-\left(\frac{\upsilon}{\sigma}\right)^2}$	

u. S. 259ff). Die Transformation in den Objektraum lässt den Zustandsraum als ungeeignete Abbildung des Objektraums erscheinen (Abb. 4.9b, Pfeil von oben). Im Objektraum sind die beiden Punktgruppen durch eine lineare Hyperfläche, in diesem Beispiel eine Ebene, eindeutig trennbar. Die alternative Abbildung (Abb. 4.9c) zeigt einen orthogonalen Zustandsraum, in dem eine Klassifizierung mühelos gelänge. Der Kern der Aufgabe besteht darin, einen geeigneten Operator im Zustandsraum zu finden (Abb. 4.9d), der einer Abbildung der Hyperebene im Objektraum entspricht. Die ideale Klassifizierung kann zum einen durch drei Neuronen mit einer Sigmoid-Aktivierung erfolgen (Abb. 4.9e, graue gestrichelte Linien) oder durch zwei RBF-Neuronen mit einer GAUSS-Aktivierung (Abb. 4.9e, schwarze Linien). Die in der Praxis erreichbare Lösung wird hingegen ein weniger ideales Erscheinungsbild haben (Abb. 4.9f). Hiermit lässt sich verdeutlichen, dass die Mustererkennung durch ein neuronales Netz einer Separation der Musterpaare durch eine Hyperebene im Objektraum entspricht.

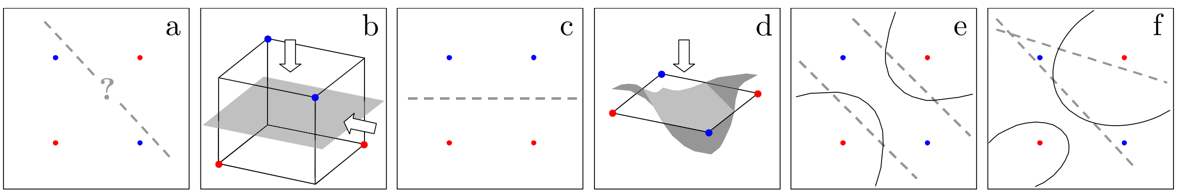


Abb. 4.9: Die XOR-Dichotomie a) im zweidimensionalen Zustandsraum, b) im dreidimensionalen Objektraum mit den Normalenvektoren der Abbildungen, c) in einen orthogonalen Zustandsraum, d) die Abbildung der klassifizierenden Ebene in den Zustandsraum, e) die theoretische Klassifizierung durch drei Neuronen (grau) und zwei RBF-Neuronen (schwarz) im Vergleich und f) praktische Lösung.

4.5.4 Netzwerkarchitekturen

Die Neuronen eines sogenannten *Multilayer perceptron* (MLP) sind in Schichten (Layern) angeordnet (MIIMA 2002, S. 42f). Diese bestehen aus einer Eingabeschicht, einer oder mehreren verdeckten Schichten und einer Ausgabeschicht (Abb. 4.10). Die Strategie, dass die Berechnungen beginnend mit der Eingabe schichtweise in Richtung der Ausgabe erfolgt, wird als *multilayer feed forward* bezeichnet (PATTERSON 1996, S. 47).

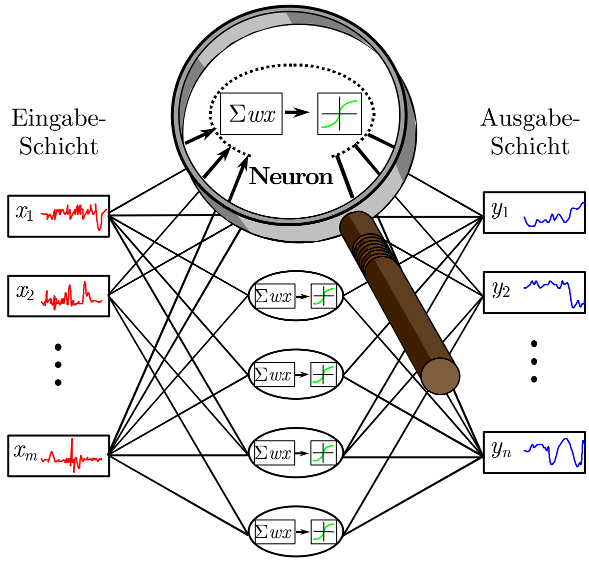


Abb. 4.10: Schema eines *Multilayer perceptron* (MLP) mit drei Schichten (SCHLITTGEN & STREITBERG 1997, vergl. Abb. 8.1.3.1).

Aufbau eines mehrschichtigen Perzeptrons

Die meisten mehrschichtigen Perzeptronen sind aus Neuronen entsprechend einem ARX$[0, e]$-Modells mit sigmoidaler Aktivierung aufgebaut. Mithilfe der sogenannten *backpropagation* auf der Grundlage des gradientiellen Lernens können auch den Gewichten in den verborgenen Schichten Korrekturen zugewiesen werden. Die Nichtlinearität der hintereinander geschalteten Neuronen hat aber zur Folge, dass sich bei der Optimierung lokale Minima ergeben können. Ein neuronales Netz mit RBF-Neuronen kommt bereits mit einer verborgenen Schicht aus und lässt sich für die Anzahl von $\mathcal{H}$ RBF-Neuronen mit

$$\Phi_i(\mathbf{w}) = \sum_{k=1}^{\mathcal{H}} w_{ik} \left(e^{-a(\|\mathbf{x}_i - \boldsymbol{\chi}_k\|)^2} + w_{i0} \right) \tag{4-43}$$

beschreiben (HAYKIN 1999, S. 263). Es stellt sich aber die Frage, ob das Netz mit vielen statischen Stützvektoren $\boldsymbol{\chi} = \text{const}$ oder mit einer deutlich geringeren Anzahl beweglicher Stützvektoren mit variablen Formfaktoren zu bestücken ist.

Die VAPNIK-CHERVONENKIS-Dimension

Ein theoretisches Gütekriterium für ein Modell muss die Aussage treffen können, ob die verwendeten Operatoren jeweils Funktionsscharen angehören, deren Anpassungsfähigkeit nicht zu groß gegenüber der Anzahl von Paaren von Eingangs- und Ausgangsvektoren $\{\mathbf{x}, \mathbf{y}\}$ ist. Ist dieses Gütekriterium eingehalten und ist eine Anpassung des Modells an die Trainingsdaten gelungen, so ist es wahrscheinlich, dass das Modell identifizierend ist (Abb. 4.3), also die Struktur zwischen den Eingängen und Ausgängen eines Systems abbildet. Ist dieses Gütekriterium nicht erfüllt, so besteht eine große Wahrscheinlichkeit, dass die Operatoren ihre Anpassungsfähigkeit zur ungeneralisierten Speicherung der Trainingsdaten verwendet haben. Die VAPNIK-CHERVONENKIS-Dimension eines neuronalen Netzes $\mathcal{N}$ oder kurz $\mathcal{VC}(\mathcal{N})$ sagt aus, wieviele binäre Muster in einem n-dimensionalen Ursprungsraum $\mathbb{R}^n$ irrtumsfrei durch eine Funktionsschar separiert werden können (HAYKIN 1999, S. 94f). Damit wird jeder Funktionsschar und weiterhin ihren verschiedenen Kombinationen eine *Kapazität* zur Trennung

Tab. 4.2: Obere und untere Schranke der Vapnik-Chervonenkis-Dimension für verschiedene Aktivierungsfunktionen in Abhängigkeit der Anzahl der Gewichte n und der Neuronen $\mathcal{H}$ bei einer einzigen verborgenen Schicht.

Aktivierung	$\varphi(\upsilon)$	υ	h_{min}	h_{max}
linear	$\varphi_{/}(\upsilon)$	$\forall\upsilon$	$\Theta(n)$	$\mathcal{O}(n^2)$
beschränkt linear	$\varphi_{[/]}(\upsilon)$	$\lvert\upsilon\rvert > \frac{1}{2m}$ $\lvert\upsilon\rvert < \frac{1}{2m}$	$\Theta(n)$	$\mathcal{O}(n\mathrm{lb}(n))$ $\mathcal{O}(n^2)$
Schwellwert	$\varphi_H(\upsilon)$	$\forall\upsilon$	$\Theta(n)$	$\mathcal{O}(n\mathrm{lb}(n))$
Kombination	$\varphi_H(\upsilon) \oplus \varphi_{/}(\upsilon)$	$\lvert\upsilon\rvert > \frac{1}{2m}$ $\lvert\upsilon\rvert < \frac{1}{2m}$	$\Theta(n)$	$\mathcal{O}(n\mathrm{lb}(n))$ $\mathcal{O}(n^2)$
Sigmoid	$\varphi_\Sigma(\upsilon)$	$\lvert\upsilon\rvert < \frac{1}{a}$ $\lvert\upsilon\rvert > \frac{1}{a}$	$\Theta(n)$	$\mathcal{O}(n^2)$ $\mathcal{O}(n\mathrm{lb}(n))$
RBF	$\varphi_N(\lVert\upsilon\rVert)$	$\forall\upsilon$	$\Theta(n)$	$\mathcal{O}(n\mathrm{lb}(H))$

zugeschrieben. Da allerdings auch unscharfe Mengenzugehörigkeiten von Mustern oder unscharfe Zuweisungen durch die Funktionsschar möglich sind, erweitert sich $h = \mathcal{VC}(\mathcal{N})$ zur *Kardinalität*[28] der irrtumsfrei trennbaren Muster. Die Berechnung dieser Dimension kann bisher noch nicht verallgemeinert werden (Koiran & Sontag 1996; Elisseeff & Paugam-Moisy 1997; Sontag 1998; Schmitt 2001; Schmitt 2005). Vielmehr ist es oft nur möglich, die oberen und unteren Grenzen in Form von Landau-Bachmann-Symbolen[29] in Abhängigkeit von einer Netzarchitektur und den verwendeten Aktivierungsfunktionen $\varphi(\upsilon)$ anzugeben (Tab. 4.5.4).
Im Regressionsfall ist die VC-Dimension die Kardinalität der Muster, die durch eine Funktionsschar irrtumsfrei approximiert werden.

Verteilungsunabhängige Grenzen der Risikofunktion

Die Grenze der Risikofunktion oder Risikoschranke $R(h)$ (Abb. 4.11) berücksichtigt durch die Summe

$$R(h) = R_{\mathrm{emp}} + \epsilon_1(N, h, \alpha, R_{\mathrm{emp}}) \tag{4-44}$$

sowohl das empirische Risiko[30] R_{emp} als auch das Konfidenzintervall

$$\epsilon_1(N, h, \alpha, R_{\mathrm{emp}}) = 2\epsilon_0^2(N, h, \alpha)\left(1 + \sqrt{1 + \frac{R_{\mathrm{emp}}}{\epsilon_0^2(N, h, \alpha)}}\right) \tag{4-45}$$

in Abhängigkeit des empirischen Risikos und des Konfidenzintervalls der VC-Dimension (Haykin 1999, S. 99ff). Dieses Konfidenzintervall der VC-Dimension (Vapnik 1998, S. 219ff)

$$\epsilon_0(N, h, \alpha) = \sqrt{\frac{h}{N}\left(\log\left(\frac{2N}{h} + 1\right) - \frac{1}{N}\log\alpha\right)} \tag{4-46}$$

ist eine Komplexitätsschranke abhängig von der Anzahl der Muster N, der VC-Dimension h und der Wahrscheinlichkeit

$$\alpha = \left(\frac{2eN}{h}\right)^h e^{-\eta^2 N}. \tag{4-47}$$

[28] Die *Mächtigkeit* oder *Kardinalität* einer endlichen Menge ist die Summe aller charakteristischen Funktionswerte ihre Elemente (Bothe 1995, S. 35). Dabei beschreibt die charakteristische Funktion eines Elementes dessen unscharfe Zugehörigkeit zu einer Menge mit Werten zwischen 0 und 1.

[29] Aus zweckmäßigen Gründen ist hier von den Landau-Bachmann-Symbolen leicht abgewichen worden.

[30] Im Falle eines neuronalen Netzes repräsentiert das empirische Risiko den Trainingsfehler.

Darin ist η die Genauigkeit, mit der die Approximation erfolgen soll.
Entscheidend für die Architektur eines neuronalen Netzes ist der Bereich um das Minimum der Risikoschranke (Abb. 4.11). In diesem Bereich[31] ist ein Modell mit den entsprechenden VC-Dimensionen richtig bestimmt. Ist die VC-Dimension kleiner, besteht Grund zur Annahme, dass das Modell *überbestimmt* ist, also zuviele Muster für ein simples Modell aufweist (Haykin 1999, S. 100f). Es ist damit aber keineswegs ausgeschlossen, dass das Modell die Muster so „intelligent" generalisiert hat, dass das Modell dennoch ein kleines empirisches Risiko hat und damit durchaus identifizierend ist. Das hierzu notwendige Phänomen wird als *weight sharing* bezeichnet (Haykin 1999, S. 28f). Es tritt dann auf, wenn verschiedene Kombinationen von realisierten Systemein- und -ausgängen jeweils Beziehungen gleichartiger Natur besitzen und sich damit dieselben Neuronen teilen können.
Umgekehrt ist ein Modell mit hoher VC-Dimension *unterbestimmt*[32]. Es gibt also möglicherweise zuwenig Muster um das Modell zu stützen. Bei einer nur schwachen Unterdeterminierung kommt es zunächst zu dem Phänomen, dass sich die Abbildung von Subprozessen additiv auf mehr Neuronen verteilt, als zur Abbildung notwendig wären. Im schlimmsten Falle kann das Modell seine Kapazität darauf verwendet haben, statt eine Generalisierung zu erreichen, jedes einzelne Muster zu memorisieren, was man als *Übertrainieren*[33] bezeichnet. Es wird sich also im Extremfall an jedes einzelne Muster „erinnern". Das Minimum der Risikoschranke liegt also im Optimum zwischen „Wissen" und „Intelligenz" des Modells.

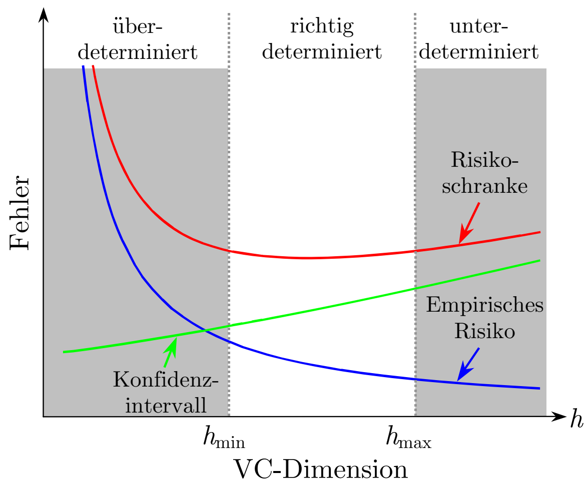


Abb. 4.11: Die Risikoschranke als obere Hüllende über dem empirischen Risiko und dem Konfidenzintervall als Maß der Modellkapazität hat ihr Minimum im Bereich der optimalen VC-Dimension. Überbestimmte Modelle generalisieren möglicherweise zu stark, während unterbestimmte Modelle zu stark memorisieren.

Knotenanzahl eines neuronalen Netzes aus der Risikoschätzung

Für ein mehrschichtiges *feedforward* Netzwerk ist es bereits gelungen, eine geeignete Anzahl von *Knoten* $\mathcal{H}$ anzugeben (Elisseeff & Paugam-Moisy 1997). Knoten sei hier der Sammelbegriff sowohl für Neuronen (4-30) als auch für RBF-Neuronen (4-39). Demzufolge soll ein neuronales Netz über

$$\left\lceil \frac{\mathcal{P}\dim(\mathbf{y})}{\dim(\mathbf{x})+\dim(\mathbf{y})} \right\rceil < \mathcal{H} < 2\left\lceil \frac{\mathcal{P}\dim(\mathbf{y})}{\dim(\mathbf{x})+\dim(\mathbf{y})} \right\rceil \tag{4-48}$$

Knoten in den verborgenen Schichten verfügen, um eine VC-Dimension im Bereich der minimalen Risikoschranke zu besitzen. Dabei sei $\mathcal{P}$ die Anzahl von Paaren von Eingangs- und Ausgangsvektoren, $\dim(\mathbf{x})$ die Dimension des Eingangsvektors und $\dim(\mathbf{y})$ die Dimension des Ausgangsvektors.

[31] Die immer schärfere Beschränkung dieses Intervalls ist die Aufgabe der *advanced learning theory*.

[32] Der Begriff der *Unterbestimmtheit* in Bezug auf die Kapazität eines neuronalen Netzes ist nicht exakt gleichbedeutend mit der Unterbestimmtheit eines Ausgleichungsproblems. Ein unterbestimmtes Ausgleichungsproblem ist nicht lösbar, wohingegen ein unterbestimmtes neuronales Netz durchaus noch konvergiert.

[33] Geläufiger ist hier der englische Terminus *overfitting*.

4.5.5 Praktische Eingriffe in den Lernprozess

Häufig wird in der Architektur von neuronalen Netzen die Kapazität der Neuronen unterschätzt, so dass zu viele Neuronen in ein Netz eingebaut werden. Als Folge ergeben sich unterbestimmte Netze, die man auf verschiedenen Wegen versucht, vom Übertrainieren abzuhalten. Hierzu seien zwei Lösungen vorgestellt.

Die Kreuzvalidierung

Ein Indikator für den Verlust der Identifikationsfähigkeit eines Netzes und damit den Beginn des Memorisierens ist die Verschlechterung der Prädiktion des Modells. Aus diesem Grund werden 10% der Musterpaare vom Training als *Validierungsdatensatz* ausgeschlossen. Die Eingänge dieser Paare werden zur Prädiktion verwendet. Solange sich das empirische Risiko aus diesen Prädiktionen vermindert, wird unterstellt, dass das Netz noch zu einem identifizierenden Modell konvergiert. Beginnt sich während des Lernens das empirische Risiko über dem Validierungsdatensatz wieder zu verschlechtern, wird der Lernprozess gestoppt (MIIMA 2002, S. 60).

Die Klassifizierung

Anstelle in den Algorithmus einzugreifen, können auch die realisierten Systemein- und -ausgänge vorklassifiziert werden. Hierzu werden die Systemein- und -ausgänge entsprechend ihrer Verteilung zu Klassen zusammmengefasst. Dabei können diskrete Untermengen anhand von scharfen Klassengrenzen gebildet werden. Es können ferner semidiskrete Untermengen gebildet werden, wobei die charakteristische Funktion der Klassen auf begrenzten Intervallen über der Verteilung die Werte null oder eins annimmt. Nur im Bereich der unscharfen Klassengrenzen nimmt die charakteristische Funktion Werte zwischen null und eins an, wobei zwei benachbarte Klassen im Übergangsbereich immer komplementäre Mengen bilden. Schließlich ist auch eine vollständig fuzzifizierte Klassenbildung möglich wobei $N \neq \sum_{i=1}^{N} \mu_i$.

Bei jeder Form der Klassifizierung der Daten wird nicht mehr das einzelne möglicherweise stochastisch verfälschte Musterpaar zum Lernen verwendet, sondern es findet bereits vor dem Lernprozess eine Art Mittelwertbildung statt. Damit ist es auch bei Netzen mit zu hoher Kapazität weitgehend ausgeschlossen, dass einzelne Muster memorisiert werden können.

4.6 Sequentielle Modelloptimierungen im KALMAN-Filtermodell

Das KALMAN-Filter läuft in seiner Rekursion in drei Schritten ab (MINKLER & MINKLER 1993, vergl. Kap. 6.5). Es entspricht in seiner Gänze der Anwendung eines allgemeinen rekursiven Filters (3-77).

4.6.1 Das Prädiktionsmodell

In der ersten Phase, der *Prädiktion*, wird ausgehend von der derzeitig ausgeglichenen Realisierung des Systemausgangs $\hat{\mathbf{y}}_t$ der zukünftige Systemausgang[34]

$$\bar{\mathbf{y}}_{t+\tau} = \mathbf{T}_{t,t+\tau} \cdot \hat{\mathbf{y}}_t + \mathbf{B}_{t,t+\tau} \cdot \hat{\mathbf{u}}_t + \mathbf{C}_{t,t+\tau} \cdot \hat{\mathbf{w}}_t \tag{4-49}$$

[34] Der aus der Regeltechnik bekannte und in die Geodäsie importierte Begriff des *Zustandes* (DIN 18709-5:Entwurf-2005) ist hier bewusst zur Verdeutlichung des Zusammenhangs durch den Begriff *Systemausgang* ersetzt.

geschätzt. Die *Transitionsmatrix* $\mathbf{T}_{t,t+\tau}$ enthält ein in aller Regel parametrisches Modell in Form von Übertragungsfunktionen, wie vom derzeitigen ausgeglichenen Systemausgang zum Systemausgang der Folgeepoche $t+\tau$ zu kommen ist. Diese Übertragungsfunktionen sind formal analog zum autoregressiven Modell eines AR[p]-Modells (Kap. 3.2.2).
Die *Stellgrößenmatrix* $\mathbf{B}_{t,t+\tau}$ enthält die Übertragungsfunktionen der realisierten Systemeingänge x bezüglich der Stellgrößen u. Jene entsprechen dem externen Eingang eines ARX[p,e]-Modells.
Die *Störgrößenmatrix* $\mathbf{C}_{t,t+\tau}$ schließlich enthält die Übertragungsfunktionen der realisierten nichtdeterministischen Prozesse des Systems sowie der Realisierung und ist somit vergleichbar mit dem gleitenden Mittelwert des MA[q]-Modells.
Somit gleicht die Systemgleichung eines KALMAN-Filters der eines ARMAX[p,q,e]-Modells und beide lassen sich auf die Grundform eines allgemeinen rekursiven, kausalen Filters (3-77) zurückführen (HEUNECKE 1995, vergl. S. 15f).
Ein wesentlicher Unterschied besteht aber darin, dass die Übertragungsfunktionen anders als beim nichtparametrischen ARMAX-Modell in mathematisch-physikalisch deutbarer Form polynomisiert werden. Ein zweiter Unterschied liegt in der sequentiellen Bearbeitung des KALMAN-Filters, während die nichtparametrischen Modelle auch semisequentiell oder statisch optimiert werden.
Diese sequentielle Bearbeitung erlaubt die Berücksichtigung eines stochastischen Modells

$$\begin{aligned}\mathbf{Q}_{\bar{y}\bar{y},t+\tau} \quad = \quad & \mathbf{T}_{t,t+\tau} \cdot \mathbf{Q}_{\hat{y}\hat{y},t} \cdot \mathbf{T}^{\mathrm{T}}_{t,t+\tau} \\ & + \mathbf{B}_{t,t+\tau} \cdot \mathbf{Q}_{xx,t} \cdot \mathbf{B}^{\mathrm{T}}_{t,t+\tau} \\ & + \mathbf{C}_{t,t+\tau} \cdot \mathbf{Q}_{ww,t} \cdot \mathbf{C}^{\mathrm{T}}_{t,t+\tau},\end{aligned} \tag{4-50}$$

das die Qualität der verschiedenen Realisierungen mithilfe der Kofaktormatrizen $\mathbf{Q} = \frac{1}{\sigma_0^2}\mathbf{\Sigma}$ beschreibt. Die Addition des stochastischen Modells eines Störprozesses $\mathbf{Q}_{ww,t}$ ist aufgrund des *Divergenzproblems* des KALMAN-Filters schon numerisch zwingend erforderlich (HAYKIN 2002, S. 494), denn die Kofaktormatrix der Prädiktion $\mathbf{Q}_{\bar{y}\bar{y},t+\tau}$ muss positiv definit gehalten werden. Das Divergenzproblem entsteht durch die beschränkten digitalen Wortlängen von Fließkommazahlen. Bei der Nutzung der elektronischen Datenverarbeitung müssen unendlich lange Fließkommazahlen zwangsweise gerundet werden. Dieser Effekt wirkt sich auf die fortwährende Iteration in zeitdiskreten Filtern dahingehend aus, dass die Kovarianzmatrizen – bedingt durch wiederholten Rundungsfehler – negativ definit werden.

4.6.2 Die Innovation

Im zweiten Schritt, der sogenannten *Innovation*

$$\mathbf{d}_{t+\tau} = \mathbf{l}_{t+\tau} - \mathbf{A}\bar{\mathbf{y}}_{t+\tau}, \tag{4-51}$$

findet der Vergleich der prädizierten Systemausgänge mit den tatsächlich realisierten statt. Dazu muss der Systemausgang über das funktionale Modell $\mathbf{A}$ in die prädizierten Systemausgänge $\bar{\mathbf{y}}_{t+\tau}$ vermittelt werden (NIEMEIER 2002, Kap. 4.2.3 u. 12.4.3). Entsprechend der *Varianz-Kovarianzfortpflanzung*[35] einer Differenz zweier Realisierungen aus jeweils einem stochastischen Prozess ergeben sich die Kofaktoren der Innovation zu

$$\mathbf{D}_{t+\tau} = \mathbf{Q}_{ll,t+\tau} + \mathbf{A}\mathbf{Q}_{\bar{y}\bar{y},t+\tau}\mathbf{A}^{\mathrm{T}}. \tag{4-52}$$

[35] Die Begriffe *Fehlerfortpflanzung*, *Varianzfortpflanzung* und *Kovarianzfortpflanzung* werden synonym verwendet.

4.6.3 Die Filterung

Der dritte Schritt der Rekursion ist die *Ausgleichung* zum Zwecke der Filterung. Hierzu ist die *Durchlassmatrix*[36]

$$\mathbf{K}_{t+\tau} = \mathbf{Q}_{\bar{y}\bar{y},t+\tau}\mathbf{A}^{\mathrm{T}}\mathbf{D}_{t+\tau}^{-1} \tag{4-53}$$

zu berechnen, mit deren Hilfe die Innovation anteilig an die Prädiktion angebracht werden kann:

$$\hat{\mathbf{y}}_{t+\tau} = \bar{\mathbf{y}}_{t+\tau} + \mathbf{K}_{t+\tau}\mathbf{d}_{t+\tau}. \tag{4-54}$$

Die Kofaktormatrix des ausgeglichenen Systemausgangs der Folgeepoche

$$\mathbf{Q}_{\hat{y}\hat{y},t+\tau} = \mathbf{Q}_{\bar{y}\bar{y},t+\tau} - \mathbf{K}_{t+\tau}\mathbf{D}_{t+\tau}\mathbf{K}_{t+\tau}^{\mathrm{T}} \tag{4-55}$$

die Kofaktormatrix der Prädiktion abzüglich des sogenannten Ausgleichungsvorteils. An dieser Stelle stehen alle Informationen für die nächste Iteration zur Verfügung.

4.6.4 Funktionsweise der Durchlassmatrix

Um die Arbeitsweise des KALMAN-Filters verstehen zu können, wird die Annahme getroffen, es handele sich bei **A** und **T** um zwei [1,1]-Matrixen mit jeweils einem konstanten Wert eins. Damit reduziert sich die ausformulierte Durchlassmatrix

$$\begin{aligned}\mathbf{K}_{t+\tau} \quad = \quad & \left(\mathbf{T}_{t,t+\tau}\cdot\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{\hat{y}\hat{y},t}\cdot\mathbf{T}_{t,t+\tau}^{\mathrm{T}}+\ldots\right)\mathbf{A}^{\mathrm{T}} \\ & \cdot\left(\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{ll,t+\tau}+\mathbf{A}\left(\mathbf{T}_{t,t+\tau}\cdot\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{\hat{y}\hat{y},t}\cdot\mathbf{T}_{t,t+\tau}^{\mathrm{T}}+\ldots\right)\mathbf{A}^{\mathrm{T}}\right)^{-1}\end{aligned} \tag{4-56}$$

auf den Ausdruck

$$\mathbf{K}_{t+\tau} = \frac{\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{\hat{y}\hat{y},t}+\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{xx,t+\tau}+\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{ww,t+\tau}+\ldots}{\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{ll,t+\tau}+\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{\hat{y}\hat{y},t}+\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{xx,t+\tau}+\frac{1}{\sigma_0^2}\mathbf{\Sigma}_{ww,t+\tau}+\ldots}. \tag{4-57}$$

Aus der Bedingung, **A** und **T** seien zwei [1,1]-Matrixen, folgt auch, dass die Kovarianzmatrizen sich auf jeweils eine einzelne Standardabweichung reduzieren lassen. Damit vereinfacht sich die Durchlassmatrix abermals auf:

$$\mathbf{K}_{t+\tau} = \frac{\sigma_{\hat{y}}^2+\sigma_x^2+\sigma_w^2+\ldots}{\sigma_l^2+\sigma_{\hat{y}}^2+\sigma_x^2+\sigma_w^2+\ldots}. \tag{4-58}$$

Damit entsteht der Spezialfall eines KALMAN-Filters, der nur noch eine Zustandsgröße direkt mit einer Realisierung vergleicht, so dass die ausgeglichene Zustandsgröße

$$\hat{y}_{t+\tau} = \bar{y}_{t+\tau} + \frac{\sigma_{\hat{y}}^2+\sigma_x^2+\sigma_w^2+\ldots}{\sigma_l^2+\sigma_{\hat{y}}^2+\sigma_x^2+\sigma_w^2+\ldots}\cdot\left(l_{t+\tau}-\bar{y}_{t+\tau}\right). \tag{4-59}$$

eine Summe des prädizierten Zustands mit der gewichteten Innovation bildet. Die Arbeitsweise der Gewichtung lässt sich mit Grenzwertbildung

$$\lim_{\sigma_l\to\infty}\left(\frac{\sigma_{\hat{y}}^2+\sigma_x^2+\sigma_w^2+\ldots}{\sigma_l^2+\sigma_{\hat{y}}^2+\sigma_x^2+\sigma_w^2+\ldots}\right) = 0 \Rightarrow \hat{y}_{t+\tau} = \bar{y}_{t+\tau} + 0\cdot\left(l_{t+\tau}-\bar{y}_{t+\tau}\right) = \bar{y}_{t+\tau} \tag{4-60}$$

[36] Die in der deutschsprachigen Literatur anzutreffende Übersetzung *Verstärkungsmatrix* für *gain matrix* wird dem englischen Terminus in diesem Kontext weder linguistisch oder semantisch noch mathematisch gerecht: es wird vielmehr *gedämpft* zum Zwecke einer *Regelung*.

verdeutlichen: eine unendlich schlechte Messung hätte zur Folge, dass der ausgeglichene Zustand $\hat{y}_{t+\tau}$ ausschließlich durch die Prädiktion $\bar{y}_{t+\tau}$ gebildet wird (HEINERT & REISER 2002). Existierte hingegen eine unendlich gute Messung, so würde der ausgeglichene Zustand

$$\lim_{\sigma_l \to 0} \left(\frac{\sigma_{\hat{y}}^2 + \sigma_x^2 + \sigma_w^2 + \dots}{\sigma_l^2 + \sigma_{\hat{y}}^2 + \sigma_x^2 + \sigma_w^2 + \dots} \right) = 1 \Rightarrow \hat{y}_{t+\tau} = \bar{y}_{t+\tau} + 1 \cdot (l_{t+\tau} - \bar{y}_{t+\tau}) = l_{t+\tau} \quad (4\text{-}61)$$

vollständig durch diese Messung $l_{t+\tau}$ bestimmt werden und die Prädiktion wäre damit verworfen.

4.6.5 Probleme des KALMAN-Filtermodells

Das KALMAN-Filter ist allerdings in der beschriebenen Form nicht phasentreu. Der Vergleich mit einem kausalen ARMAX[p,q,e]-Modell zeigt, dass sämtliche Gewichte auf frühere Epochen bezogen sind, womit die Gewichtsfunktionen oder -distributionen nicht der Forderung der Symmetrie (3.3.3) entsprechen. Für alle Anwendungen, welche nicht den Zwängen der Echtzeitanwendung unterliegen, bietet sich daher an, das Filter in zwei Richtung zu durchlaufen und die Resultate zu mitteln.

Einerseits ist es gerade die Stärke dieses Filters, schnell, wenn auch – wie soeben dargelegt – nicht gleich phasentreu zu reagieren, andererseits kann genau diese Eigenschaft die größte Schwäche darstellen. Die als *recency phenomenon* bezeichnete Schwierigkeit besteht darin, dass das Filter die Neigung besitzt, schnell auf den gerade aktuellen Zustand zu reagieren (PUSKORIUS & FELDKAMP 2001, S. 25f). Die Ursache des Problems findet sich in der unzureichenden Gedächtnislänge dieses sequentiell arbeitenden Algorithmus'. So werden die stationären Zustandsparameter ebenso geändert, wie jene, welche auf die kurzfristigen Veränderungen reagieren sollen. Insbesondere wenn die Werte extrem sind, werden alle Parameter unmittelbar angepasst.

Schließlich läßt sich das KALMAN-Filtermodell nicht ohne weiteres auf andere Zielfunktionen als die der Summe der kleinsten Fehlerquadrate anwenden (PUSKORIUS & FELDKAMP 2001, S. 64). Falls doch ein anderer Schätzer zu Einsatz kommen soll, muss dieser über gewichtete Beobachtungen auf einen L_2-Schätzer zurückgeführt werden (JÄGER ET AL. 2005, Kap. 4.5). Hier gibt es den Vorschlag, die Aufdatierung des Zustandsvektors über ein Quasi-NEWTON-Verfahren (Kap. 4.3.5) durchzuführen (SEYFFERT 2002, S. 39).

5 Adaptierte Analysen und Modelle

„Einfaches Handeln, folgerecht durchgeführt,
wird am sichersten das Ziel erreichen.“
Helmuth Karl Bernhardt von Moltke

Ein Modell soll ein reales System abbilden. Das System, um das es sich in diesem Untersuchungsfall handelt, soll als erstes entsprechend der Vorüberlegungen zu Systemen (Kap. 3.1) und hinsichtlich des Ziels der Abbildung in diesem Kapitel kategorisiert werden. Als zweites werden in diesem Kapitel die verfügbaren Daten, ihre Quellen und ihre Qualität beschrieben. Schließlich werden in diesem Kapitel die eigens zur Lösung dieser Aufgabe neu entwickelten Adaptionen vorgestellt. Diese Neuentwicklungen basieren auf den bereits dargestellten theoretischen Ansätzen zur Analyse (Kap. 3) und Modellbildung (Kap. 4).

Die geophysikalischen Rahmenbedingungen des Modells

Island mit den beiden Transformzonen und seinen vulkanischen Riftzonen nimmt eine tektonische Sonderstellung im Nordatlantik ein. Es lässt sich aus der Rekonstruktion der Entstehung Islands zeigen, dass insbesondere die Nördliche Vulkanische Riftzone und deren südliche Verlängerung, die Östlichen Vulkanischen Riftzone, immer wieder vom Manteldiapir aktiviert wird, nachdem sie bereits verödet war. In Folge dessen sind die beiden Transformzonen im Norden und Süden Islands im Wechsel immer wieder aktiv gewesen (Kap. 2.1.5).

Die resultierende Form der Plattengrenze weist daher die Form des Querschnittes durch eine Steckverbindung von *Loch und Zapfen* auf (Abb. 5.1), die geeignet ist, die beiden Platten zeitweise miteinander zu verbinden. Die langgestreckten mittelozeanischen Rücken nördlich und südlich Islands sind hingegen weitgehend ungestört. Und tatsächlich haben vulkanische Ereignisse in jüngster geologischer Vergangenheit eine feste Verkantung, wenn nicht ausgelöst, so doch äußerst begünstigt: So soll die Krafla-Rift-Episode von 1975 bis 1984, bei der Spreizungen von bis zu über 7 m auftraten (Möller 1989), die anschließende relative seismische Ruhe der Tjörnes-Bruchzone ausgelöst haben (Guðmundsson 2004). Ebenfalls ging den schweren Beben des Sommers 2000 in der Südisländischen Seismischen Zone eine Aufwölbung des Vulkanes Hengill zwischen 1993 und 1998 voraus (Feigl et al. 2000).

Die systemtheoretischen Rahmenbedingungen des Modells

Die tektonischen Erdbeben führen bezüglich der Kontinentalbewegungen zu irreversiblen Veränderungen des Systems. Damit hat das System eine unendliche Gedächtnislänge und es kann sich gemäß Kap. 3.1.5 nur um ein dynamisches System handeln. Eine gerichtete Kausalität zwischen Erdbeben und Bewegung herrscht zunächst nicht, da die Ursache sowohl der Erdbeben als auch der lokalen Bewegung die großräumige Plattenbewegung ist.
Dennoch ist zu vermuten, dass es temporär zu einer gerichteten Kausalität von Erdbeben zu Bewegung kommt, wenn sich die Platten in den Transformzonen verhaken. Diese Kausalität ist für den beobachteten Zeitraum gegeben, weswegen die Erdbeben im Folgenden als Systemeingang betrachtet werden. Die Bewegungen und Deformationen sind vernachlässigbar klein sowohl gegenüber den Ausdehnungen der Platten, als auch den Veränderungen in geologischen Zeiträumen. Daher kann das System in guter Näherung als inkremental linear angesehen werden (Kap. 3.1.3). Das bedeutet für das Modell, dass ein gleicher Systemeingang zwar eine

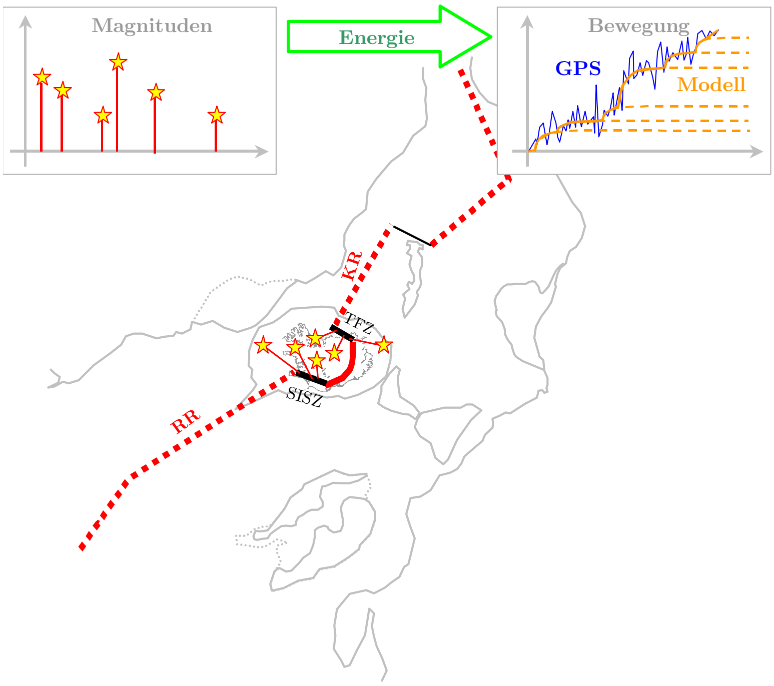


Abb. 5.1: Die vereinfachte Plattengrenze (rot) mit dem Reykjanes- (RR) und dem Kolbeinsey-Rücken (KR) zwischen Nordamerika und Eurasien auf Island hat die Gestalt von *Loch und Zapfen*. Die seismische Aktivität der Transformzonen (schwarz) ist entscheidend für die divergenten Plattenbewegungen im gesamten Nordatlantik.

gleiche relative Systemantwort zur Folge hat, dass sich der Zustand des Systems mit jedem neuen Systemeingang fortwährend in eine Richtung verändert. Die inkrementale Linearität hat aber auch weiterhin zur Folge, dass der Systemausgang auch in einen transienten Prozess (Kap. 3.2.1) und die verbleibende Residualbewegung bei einem periodischen Systemeingang näherungsweise in einen periodischen Prozess (Kap. 3.2.1) zerlegt werden kann.
Unter der Annahme, ein einzelnes Erdbeben löse eine exponentielle Systemantwort der Bewegung aus (Abb. 5.2a, schwarz), kann die Bewegung (Abb. 5.2a, rot) als Summe aller Systemantworten verstanden werden. Die Steigung der Trendfunktion entspräche somit der Geschwindigkeit aus den Globalbewegungsmodellen. Die trendfreie Bewegung (Abb. 5.2b, rot) erscheint als periodisch oder fast periodisch und kann durch eine Fourier-Reihe (Abb. 5.2b, grün) genähert werden. Dabei ergeben sich numerische Artefakte, die der Trendbefreiung geschuldet sind. Diese Artefakte zeigen sich als Pseudo-Periodizitäten in einem Spektrum. Die zugehörigen Nebenspitzen des Spektrums weisen demnach nicht auf Frequenzen einer physikalisch begründbaren Ursache hin, sondern stellen die Differenz zwischen einer harmonischen Modellfunktion und einem nichtharmonischen Prozess dar (Drewes 2006). Wird der Trend wieder zur Fourier-Reihe addiert, so kann die periodische, aber auch unendliche Systemantwort des dynamischen Systems approximiert werden (Abb. 5.2c).
Ein Kernproblem der vorliegenden Modellbildung ist es, wie ein inäquidistanter, distributiver Systemeingang in einen äquidistanten, funktionalen Systemausgang mit unendlicher Gedächtnislänge zu transformieren ist. Die unendliche Gedächtnislänge erfordert zwingend einen rekursiven Algorithmus. Jede Rekursion wiederum erfordert eine Äquidistanz oder eine Übertragung des vollständigen Systemzustandes von einer Epoche in die jeweils folgende. Da eine äquidistante Rekursion einem inkremental linearen System am ehesten gerecht wird, ist es erforderlich, die Information über Erdbeben in geeignete abgeleitete Daten zu überführen. Daher werden die Erdbeben tageweise zusammengefasst. Da aber die Bewegungen weder von

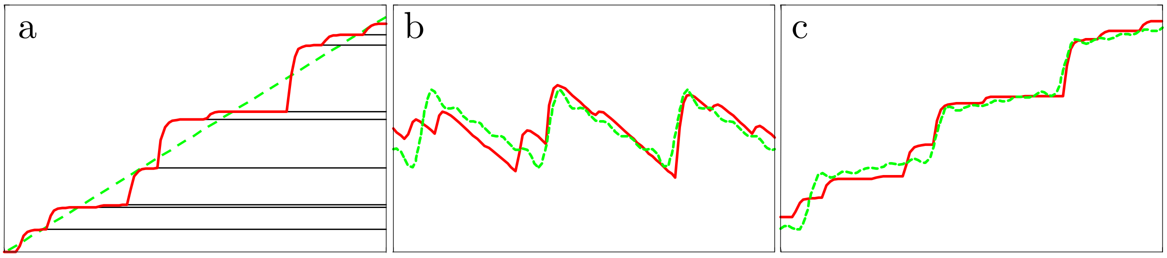


Abb. 5.2: a) Die Summe (rot) der einzelnen exponentiellen Systemantworten ergibt die modellhafte Bewegung. Die Trendfunktion entspricht einer linearen Modellbewegung. b) Die trendfreie Bewegung (rot) wird durch eine FOURIER-Reihe (grün) genähert werden. c) Systemantwort wird durch die Summe von Trend und FOURIER-Reihe approximiert.

der Anzahl, noch von der Stärke einzelner dominanter Ereignisse abhängig ist, werden für die nachfolgenden Modelle repräsentative Tagesenergien errechnet. Diese können mit den Tageslösungen der IGS-Stationen verglichen werden.

5.1 Datensätze

Verschiedene Datensätze anderer Einrichtungen sind zur Untersuchung herangezogen worden. Nicht alle haben die notwendige Datendichte oder Relevanz behalten, weshalb nur die tatsächlich in der Modellbildung verwendeten Datensätze vorgestellt werden. Es ist nicht das Ziel dieser Arbeit, diese Datensätze an sich zu verbessern oder in Frage zu stellen. Grundsätzlich wird davon ausgegangen, dass die bereitgestellten Daten nach dem Stand der Technik erzeugt worden sind, weswegen ihnen im Rahmen der technischen Möglichkeiten und der erzielbaren Messgenauigkeiten *Richtigkeit* unterstellt wird.

5.1.1 GPS-Zeitreihen

Eine Vielzahl von Institutionen stellt zurzeit Koordinatenlösungen der Permanentstationen im IGS zur Verfügung. Zum Beginn dieser Untersuchung waren jedoch nur am *Jet Propulsion Laboratory* des *Carlifornia Institute of Technology* tägliche Koordinatenlösungen frei erhältlich. Diese täglichen Koordinatenlösungen werden durch eine undifferenzierte Prozessierung der Daten auf der Basis von Gipsy-Oasis erstellt (ZUMBERGE ET AL. 1997, S. 5014). Die Lösungen werden in einem unterschiedlichen Turnus, mal wochen-, mal monatsweise veröffentlicht (JPL 2006). Für jede Station wird eine Residuengrafik für jede Koordinatenkomponente in einem individuellen lokalen verebneten Koordinatensystem zur Verfügung gestellt (Anh. B.1). Weiterhin sind diese Residuen auch digital abrufbar. Die Lösungen des JPL gehörten lange Zeit international zu den besten (BEUTLER ET AL. 1999). Hierzu darf als Indikator die Genauigkeit der finalen Orbits betrachtet werden (Abb. 5.3). Allerdings beginnt die Lösung des JPL seit etwa 2004 international zunehmend den Anschluss zu verlieren (GENDT 2008).

5.1.2 Erdbebenkatalog

Der Isländische Wetterdienst[1] stellt die vorläufigen und unüberprüften automatischen Aufzeichnungen seines Südisländischen Flachland-Seismometernetzwerkes[2] (SIL) zur Verfügung. Das Sensornetz ist zunächst nur im Süden in und um die Südisländische Seismische Zone

[1] Isl.: *Veðurstofa Íslands.*
[2] *South Icelandic lowland seismic system.*

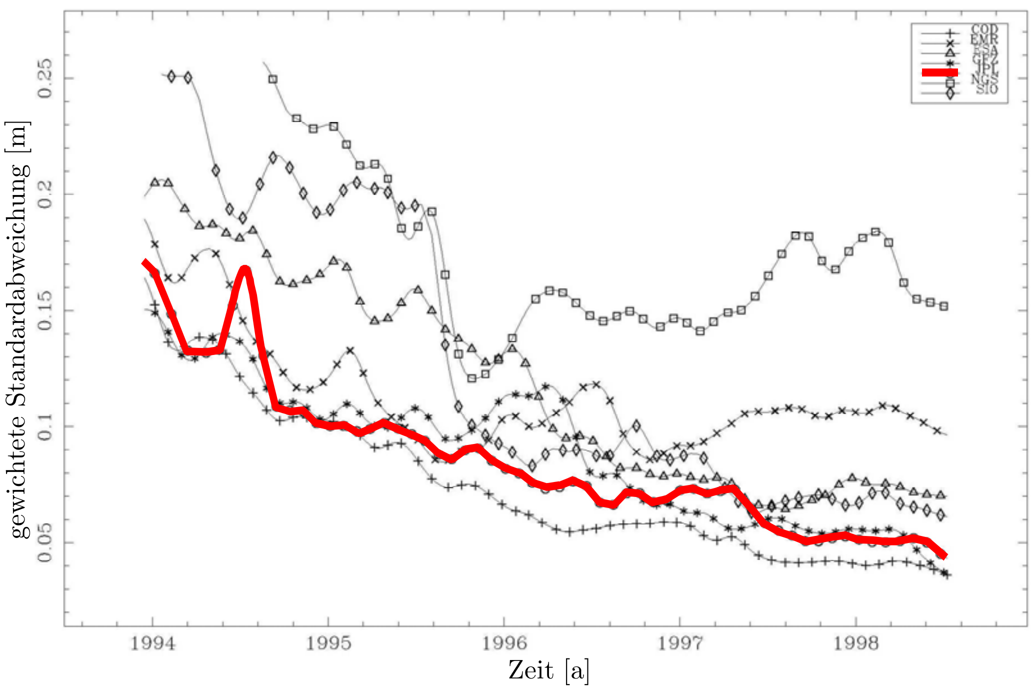


Abb. 5.3: Qualität der IGS-Analysezentren im Vergleich: Finale Orbits.

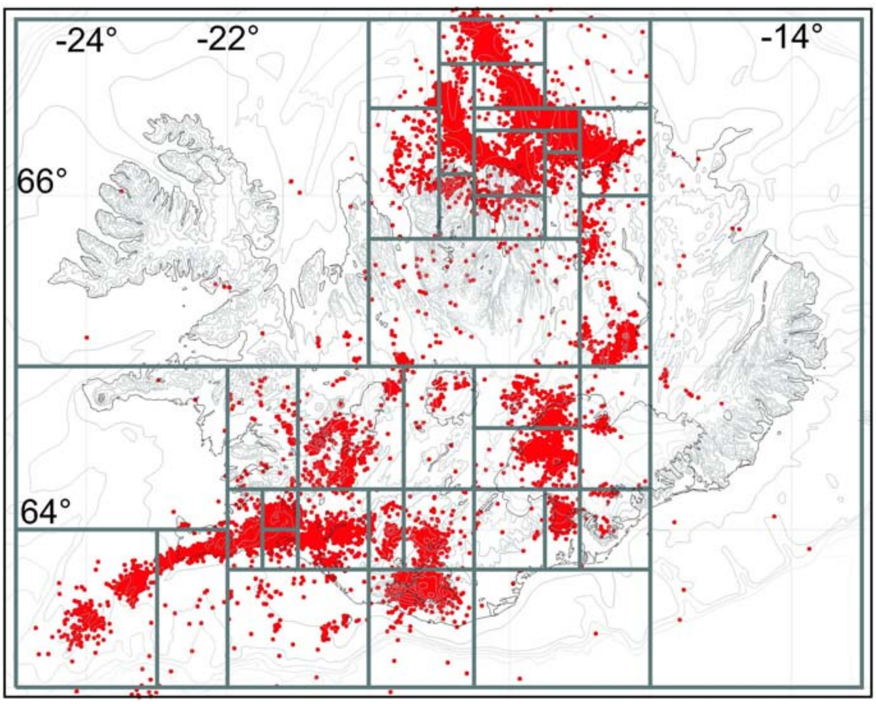


Abb. 5.4: Erdbebenepizentren auf Island von 1995-2005 (rot) ermittelt mittels des SIL-Seismometernetzwerk. Die Zoneneinteilung (grau) dient dem noch zu erläuternden 2D-Modell (Kap. 5.4).

herum eingerichtet worden. Inzwischen erstrecken sich die Sensoren über ganz Island (Jakobsdóttir et al. 2000). In diesem Netz ist eine Vielzahl verschiedener Sensoren[3] zum Einsatz gekommen, so dass jeweils unterschiedliche Kalibrierungen nötig waren. Das System ist in der Lage, auch besonders kleine Ereignisse ab einer lokalen Magnitude von M_L -0,5 zu erfassen und auszuwerten (Stefánsson et al. 1993).
Der resultierende Katalog enthält die Uhrzeit, den Ort, die etwaige Herdtiefe und die lokale Magnitude sowie die Momentmagnitude (Kap. 2.2.3). Eine Vergleichbarkeit mit anderen internationalen Katalogen ist nicht ohne weiteres herzustellen. Der Vorteil besteht darin, dass alle Ereignisse im gleichen System erfasst und weiterverarbeitet worden sind und damit untereinander gut vergleichbar sind. Aus der Überlagerung mehrfach automatisch detektierter Phasenlagen wird für einzelne Ereignisse auch die Herdfläche ermittelt (Stefánsson et al. 1993). Diese ist jedoch im Katalog nicht frei verfügbar.

5.1.3 Wetteraufzeichnungen

Nahe den IGS-Permanentstationen REYK und HOFN auf Island befindet sich jeweils eine Wetterstation direkt vor Ort. Obwohl sich auch in direkter Umgebung der Permanentstation Onsala das Weltraumobservatorium der Technischen Hochschule Göteborg[4] befindet und dort eine Wetterstation betreibt, waren diese Daten nicht zu erhalten. Daher musste auf die Wetterstation des 30 km landeinwärts entfernten Flugfeldes Landvettern ausgewichen werden.
Die Datenhaltung aller Wetterstationen ist jedoch enttäuschend (Abb. 5.5). Daher können die Informationen über Tageshöchst-, Tagestiefsttemperaturen, relative Luftfeuchte und Luftdruck nicht als Eingang für eine Modellierung verwendet werden. Die Datenlücken sind zu groß, als dass sie geeignet füllen ließen (Heinert 1998, Kap. 3.6). Weitere Informationen, wie Windgeschwindigkeiten, Schneelasten und viele mehr sind noch weit unzuverlässiger gespeichert. Dennoch sollen die Wetterdaten von Reykjavík exemplarisch in Verbindung mit den Erdbebendaten aus der Südisländischen Seismischen Zone einer adaptierten Kreuzkorrelationsanalyse unterzogen werden (Kap. 5.2.2 und 6.1.1).

5.2 Datenaufbereitung

Für die vorliegenden Daten hat sich in Hinblick auf die zu lösende Fragestellung die Notwendigkeit ergeben, verschiedene bereits erläuterte Standardansätze der Zeitreihenaufbereitung und Prozessanalyse in geeigneter Weise zu erweitern. Einerseits wird im Folgenden die

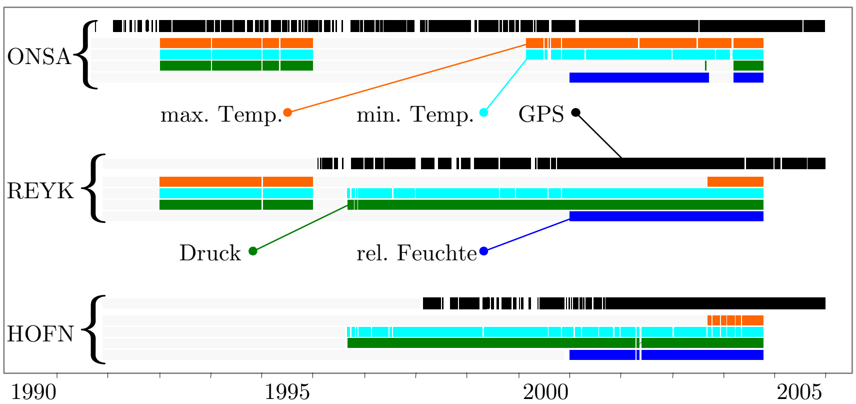


Abb. 5.5: Verfügbarkeit von GPS- und Wetterdaten.

[3] Eingebaut wurden die Seismometer Guralp CMG-3ESP, Guralp CMG-3T, Streckheisen STS-2 sowie die Geophone 1 Hz Lennartz LE-3D und 0.2 Hz Lennartz LE-3D/5s (Rögnvaldsson et al. 1997).

[4] Schwed.: *Onsala Rymdobservatorium – Chalmers Tekniska Högskola.*

KALMAN-Filterung (Kap. 4.6) mit ihren verschiedenen Rekursionsausgängen zur Datenbereinigung verwendet (Kap. 5.2.1), andererseits wird das Potenzial von Korrelationsbetrachtungen (Kap. 3.3.1) und Spektren (Kap. 3.3.2) erweitert, um sie für die in dieser Arbeit notwendigen Analysen zum Einsatz bringen zu können.

5.2.1 Innovationsanalyse

Wie jeder andere Datensatz haben GPS-Zeitreihen grobe Ausreißer als Folge der Prozessierung. Die Häufigkeit von Ausreißern kann durch eine durchgreifende Qualitätsanalyse der *pseudo-ranges* teilweise oder ganz unterdrückt werden (WIESER 2001). Wird jedoch auf die Ergebnisse einer bereits erfolgten Prozessierung zugegriffen, so existieren diese in jedem Falle besseren Möglichkeiten nicht. Als erschwerender Effekt kommen Datensprünge in den Zeitreihen vor, die oft nicht eine sprunghafte Veränderung des beobachteten Systems darstellen[5], sondern vielmehr eine gewollte (Sensortausch) oder ungewollte Veränderung des Messsystems (Manipulation durch Dritte). Beide Störungen, sowohl die groben Ausreißer als auch die Datensprünge können auf der Basis der Analyse des Innovationsprozesses in einem KALMAN-Filter korrigiert werden (HEINERT & REISER 2002; HEINERT & NIEMEIER 2004). Das KALMAN-Filter verfügt über drei separate Filterausgänge: den Prädiktions- oder Modellprozess $\bar{\mathbf{y}}$, den Innovationsprozess $\mathbf{d}$ und den ausgeglichenen Filterprozess $\hat{\mathbf{y}}$ (MINKLER & MINKLER 1993; HAYKIN 2002). Letzterer wird im Allgemeinen als das eigentliche Resultat der Filterung betrachtet.
Das KALMAN-Filter prädiziert mithilfe der Transitionsmatrix $\mathbf{T}_{t,t+\tau}$ (4-50) auf der Grundlage eines angepassten Modells den jeweils nächsten zu messenden Systemzustand $y_{t+\tau}$. Unter der vereinfachenden Annahme, der Systemzustand könne direkt gemessen werden, entspricht der Vektor der aktuellen Messung direkt dem prädizierten Zustandsvektor. Hieraus kann die Innovation durch die einfache Differenz

$$\mathbf{d}_{t+\tau} = \mathbf{y}_{t+\tau} - \bar{\mathbf{y}}_{t+\tau} = \mathbf{y}_{t+\tau} - \mathbf{T}_{t,t+\tau}\hat{\mathbf{y}}_i \tag{5-1}$$

bestimmt werden. Hierzu wird ein KALMAN-Filter auf die GPS-Zeitreihe eines beobachteten Prozesses angepasst und sowohl vorwärts als auch rückwärts durchlaufen. Der resultierende gemittelte Innovationsprozess darf als mittelwertstationär angesehen werden (Abb. 5.6).

Eine mittelwertstationäre Zeitreihe kann durch die gängigen Hypothesentests von Ausreißern befreit werden. Dazu folge man dem Hypothesentest eines Einzelwertes gegen die Studentverteilung (NIEMEIER 2002, S. 88)

$$t_f = \frac{d}{s_d} = \frac{\|\mathbf{d}\|}{\sqrt{\mathbf{e}^{\mathrm{T}}\mathbf{D}\mathbf{e}}}, \tag{5-2}$$

womit sich die Definition des Ausreißertests ergibt

$$\mathbf{l}_{t+\tau} := \begin{cases} \mathbf{l}_{t+\tau} & \forall \|\mathbf{d}_{t+\tau}\| \leq t_f \sqrt{[\mathbf{e}^{\mathrm{T}}\mathbf{D}_{t+\tau}\mathbf{e}]} \\ \mathbf{A} \cdot \bar{\mathbf{y}}_{t+\tau} & \forall \|\mathbf{d}_{t+\tau}\| > t_f \sqrt{[\mathbf{e}^{\mathrm{T}}\mathbf{D}_{t+\tau}\mathbf{e}]} \end{cases}, \tag{5-3}$$

welche die Eliminierung hochsignifikanter Ausreißer erlaubt (HEINERT & NIEMEIER 2007). Folgen mehrere Ausreißer mit gleichem Vorzeichen aufeinander, so ist dies ein sicheres Indiz für einen Datensprung, dessen Größenordung aus der Zeitreihe heraus ermittelt und anschließend aus der Zeitreihe beseitigt werden kann (HEINERT & REISER 2002).

[5] Für den Fall, dass sprunghafte Veränderungen in den Zeitreihen tatsächlich das Systemverhalten repräsentieren, bietet sich eine Analyse auf der Basis von Wavelets an (NEUNER 2008, Kap. 5).

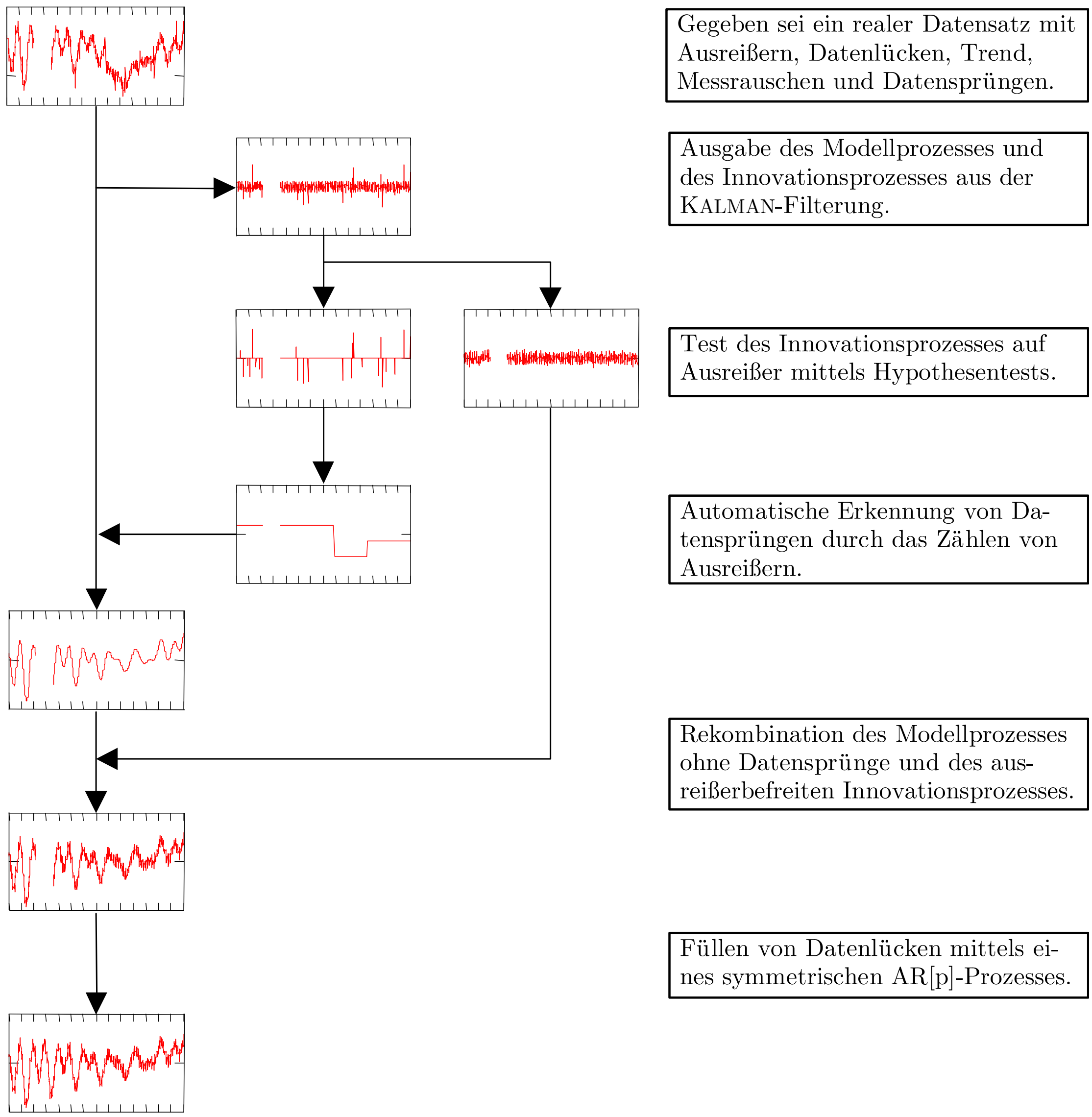


Abb. 5.6: Schema der Datenaufbereitung mit Hilfe der Innovationsanalyse im KALMAN-Filter zur Beseitigung von Ausreißern und Datensprüngen (HEINERT & NIEMEIER 2004, Abb. 5).

5.2.2 Prozessanalyse der realen Daten

Die in Kapitel 3.3 vorgestellten Verfahren zur Prozessanalyse mussten für die vorliegenden Realisierungen in geeigneter Weise erweitert werden. Zur Lösung der konkreten Probleme wurden die nachstehenden Lösungen entwickelt.

Empirische Kreuzkovarianzfunktion verschieden diskretisierter Prozesse

Ein häufiges Problem ergibt sich aus der unterschiedlichen Abtastrate des Eingangsprozesses X und des Ausgangsprozesses Y. Dieses lässt sich lösen, indem eine der beiden Abtastraten entweder künstlich erhöht oder über eine geeignete Mittelwertbildung verringert wird. Dazu sei die Annahme getroffen, dass die Abbildung $\mathcal{A}$ die mögliche vollständige Realisierung des Ausgangsprozesses Y als

$$y_t^* = \mathcal{A}_t\,(Y) \;\text{ mit } t = 0 \ldots N-1 \in \mathbb{N}^0 \qquad (5\text{-}4)$$

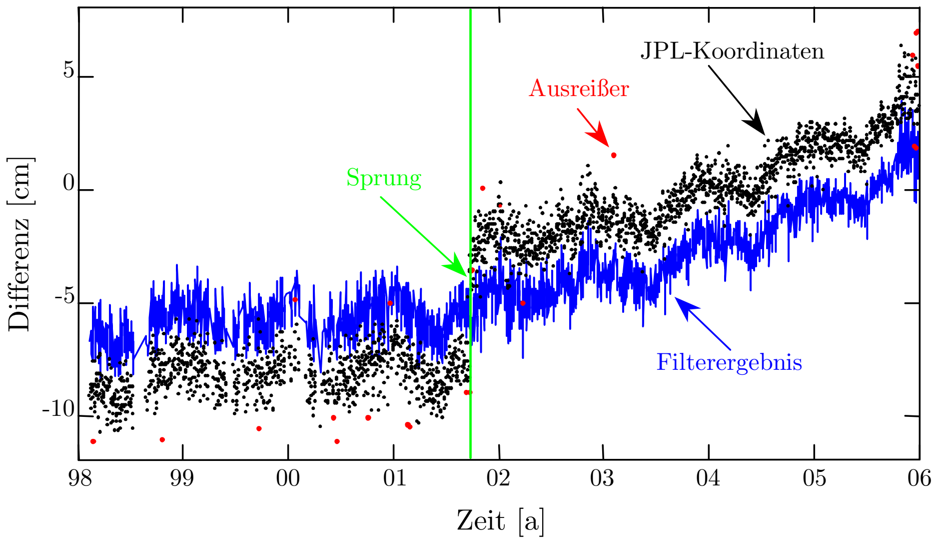


Abb. 5.7: Ergebnis der Sprung- und Ausreißerbeseitigung mithilfe der Innovationsanalyse für die Höhenkomponente der Tageskoordinatenlösungen des JPL für die IGS-Station HOFN.

darstelle, y_j hingegen die tatsächlich realisierte:

$$y_j = \mathcal{A}_j\,(Y) \text{ mit } b \cdot n = N \text{ und } j = 0 \ldots n-1 \in \mathbb{N}^0. \tag{5-5}$$

Weiterhin gelte, dass $\bar{\bar{y}}_j^*$ ein geeignetes Mittel über alle y_i^* mit $i = 0 \ldots b-1$ sei. Das Intervallmittel errechne sich zu

$$\bar{\bar{y}}_j^* = \frac{1}{b}\sum_{i=0}^{b-1} y_i^* = \frac{1}{b}\sum_{t=j\cdot b}^{(j+1)b-1} y_t^*. \tag{5-6}$$

Damit definiere man die Schrittfunktion

$$\bar{\bar{y}}_j = \bar{\bar{y}}^*_{\lfloor \frac{N\cdot j}{n} \rfloor} \tag{5-7}$$

mithilfe der positiven *Integerfunktion* (Bartsch 1999, S. 343)

$$\lfloor x \rfloor = n \;\forall\; n \leq x < n+1, n \in \mathbb{N}^0.$$

Setzt man die Schrittfunktion in die erweiterte Kreuzkovarianzfunktion ein, so ergibt sich

$$C_{xy}(\tau) = \sum_\tau \left(\left(x_{t+\tau} - \frac{1}{N}\sum_{t=0}^{N-1} x_t \right) \left(\frac{1}{b} \sum_{j=\lfloor \frac{t}{n} \rfloor}^{\lfloor \frac{t}{n} \rfloor + b-1} y_j^* - \frac{1}{N}\sum_{t=0}^{N-1} y_t^* \right) \right), \tag{5-8}$$

was sich verallgemeinert als

$$C_{xy}(\tau) = \sum_\tau \left(\left((x_{t+\tau} - \sum_{t=0}^{N-1} g_t \cdot x_t \right) \left(\sum_{j=\lfloor \frac{t}{n} \rfloor}^{\lfloor \frac{t}{n} \rfloor + b-1} g^*_{j-\lfloor \frac{t}{n} \rfloor} \cdot y_j^* - \sum_{t=0}^{N-1} g_t \cdot y_t^* \right) \right) \tag{5-9}$$

ausdrücken lässt. Der erste Term innerhalb der Summe über τ bildet für sich genommen einen Hochpassfilter, der zweite Term hingegen weist formale Ähnlichkeit mit einem Bandpassfilter auf (Tab. 3.1). Diese Ähnlichkeit wirkt sich auch auf die praktische Anwendung aus: So werden zwar die höheren Frequenzen des dichter abgetasteten Datensatzes im Bandpassfilter gedämpft, aber die hohe Auflösung bleibt erhalten.

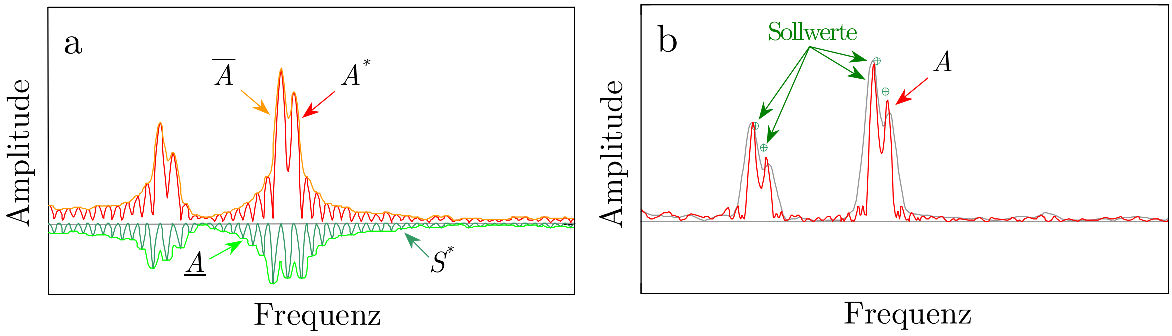


Abb. 5.8: a) Ein hochauflösendes Spektrum entsteht, wenn die maximierten Nebenspitzen vom Spektrum abgezogen werden. b) Damit erhöht sich die Trennschärfe im Frequenzbereich. Dies wird im Vergleich zu einer gefensterten FOURIER-Transformation mit *zero padding* (grau) deutlich.

Hochauflösendes Spektrum

Das zentrale Problem der Schätzung eines Spektrums ist, dass es nur schwer gelingt, eine stabile relative Amplitudenschätzung mit einer guten Auflösung im Frequenzraum zu verbinden (MAUTZ 2001, vergl. Kap. 7). Die meisten Versuche, die Energie einer Periodizität mit einer geeigneten Fensterung und gleichzeitigem *zero padding* (Kap. 3.3.2) in einem *peak* zu bündeln, haben die Verbreiterung des *peak* zur Folge (Abb. 3.3). Für dicht benachbarte Frequenzen ist das untauglich (RIEDEL & HEINERT 1998; RIEDEL ET AL. 1999; HEINERT & RIEDEL 2003; HEINERT & RIEDEL 2007).
Zur Bildung eines *hochauflösenden Spektrums*[6] wird hier eine gänzlich neue Strategie verfolgt. Zunächst wird ein Standardspektrum berechnet. Anschließend werden die störenden Fehler in Form der *side lobes* in einem zweiten Spektrum soweit maximiert bis es sogar zum Zusammenbruch des eigentlichen spektralen Signals kommt. Dieses „Störspektrum“ wird dann von einer Standardlösung abgezogen. Das resultierende Spektrum ist damit von den störenden Seitenspitzen nahezu vollständig befreit. Dazu ist es notwendig, zunächst mittels *zero padding* den Datensatz auf

$$N = 2^b \;\forall\; \left\{ b = \lceil \mathrm{lb}(10 \cdot n) \rceil = \left\lceil \frac{1 + \log n}{\log 2} \right\rceil \right\} \in \mathbb{N} \tag{5-10}$$

Werte mithilfe der *ceiling*-Rundung

$$\lceil x \rceil = n \;\forall\; n - 1 < x \leq n, n \in \mathbb{N}^0.$$

gegenüber der bisherigen Intervalllänge von n Werten mit Nullen aufzufüllen. Die dabei entstehenden Nebenspitzen des Amplitudenspektrums

$$A_i^* = \frac{2N|I_i|}{n} \text{ mit } \mathbf{I} = \mathrm{F}(\mathbf{y}) \tag{5-11}$$

werden zunächst bewusst zugelassen. Es gilt jetzt, ein zweites Spektrum zu erhalten, welches über maximierte phasentreue *sidelobes* verfügt. Hierzu werden die beiden halbseitig gefensterten Realisierungen transformiert:

$$\hat{I}^+ = \mathrm{F}(\mathrm{diag}\left[\mathbf{g}^+ \times \mathbf{y}\right]) \text{ und } \hat{I}^- = \mathrm{F}(\mathrm{diag}\left[\mathbf{g}^- \times \mathbf{y}\right]), \tag{5-12}$$

[6] Im Gegensatz zu einem *robusten Spektrum* (SUTOR 1997) wird der Algorithmus verändert und nicht die Schätzung.

wobei die halbseitigen Dreiecksfunktionen

$$g_i^+ = \frac{i}{n} \text{ und } g_i^- = \frac{n-i}{n} \tag{5-13}$$

hinreichend sind. Diese Art der Fensterung bewirkt, dass der Abschneideeffekt an jeweils einer Intervallgrenze maximiert wird. Das addierte Spektrum

$$\overline{A}_i = \frac{2N(|\hat{I}_i^+| + |\hat{I}_i^-|)}{n} \tag{5-14}$$

bildet dementsprechend eine obere Hüllende über die *sidelobes* (Abb. 5.8a). Die Extraktion der maximierten Nebenspitzen gelingt durch die Differenz

$$\mathbf{S}^* = \mathbf{A}^* - \overline{\mathbf{A}} \tag{5-15}$$

Diese haben jedoch zunächst alle den Wert null, so dass ihre untere Hüllende

$$\underline{A}_i = \frac{\sum\limits_{k=i-n}^{n-i} g_{i-k} \min[S_{i-1} \dots S_{i+1}]}{\sum\limits_{k=i-n}^{n-i} g_{i-k}} \tag{5-16}$$

mit einer Tiefpassfilterung mittels des Gauss-Kerns (Tab. 3.1)

$$g(\vartheta) = \frac{1}{\theta\sqrt{2}} e^{-\frac{\vartheta^2}{\theta^2}} \text{ mit } \theta = \frac{\sqrt{2}}{e} \text{ und } \vartheta = \frac{2(i-k)}{b}; b = \frac{N}{n} \tag{5-17}$$

zu bilden ist. Die abermalige Differenz

$$\mathbf{S} = \mathbf{S}^* - \underline{\mathbf{A}} = \mathbf{A}^* - \overline{\mathbf{A}} - \underline{\mathbf{A}} \tag{5-18}$$

liefert die maximierten Nebenspitzen, worum nun das ursprüngliche Spektrum korrigiert werden kann:

$$\mathbf{A} = \frac{e}{2}|\mathbf{A}^* - \mathbf{S}| = \frac{e}{2}|\underline{\mathbf{A}} + \overline{\mathbf{A}}|. \tag{5-19}$$

Als beste Anpassung ergibt sich demnach die Summe aus der oberen und unteren Hüllenden.

Multiple Zeit-Frequenz-Analyse

Um den Effekt des sogenannten *frequency leakage* teilweise zu unterdrücken, können kleine Amplituden aus der Zeit-Frequenz-Analyse (3-58) unterhalb des Schwellwertes ε beschnitten werden:

$${}^*A^j_{t_c,\nu} = 0 \;\; \forall \;\; {}^*A^j_{t_c,\nu} < \varepsilon. \tag{5-20}$$

Diese Amplitudenbeschneidung wird im besonderen Maße effektiv, wenn mehrere Zeit-Frequenz-Spektren mit verschiedenen Gewichtsfunktionen b_j kombiniert werden. Da ein einzelnes Amplitudenelement ${}^*A^j_{t_c,\nu}$ unterhalb des Schwellwertes fallen kann, erreicht man durch die geometrische Mittelung

$$A_{t_c,\nu} = \left(\prod_{j=1}^{m} {}^*A^j_{t_c,\nu}\right)^{\frac{1}{m}}, \tag{5-21}$$

dass das kombinierte Spektrum ebenfalls den Wert null besitzt. Das führt zu einer schärferen Zeit-Frequenz-Auflösung als ein Einzelspektrum.

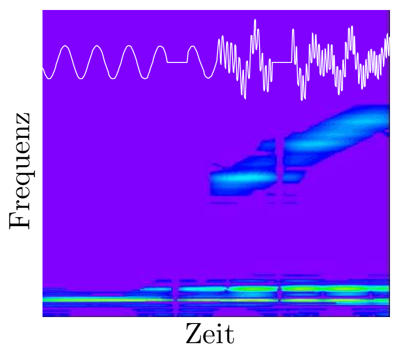


Abb. 5.9: Multiples Zeit-Frequenz-Spektrum (vergl. Abb. 3.4): Besonders die Lücken der transformierten Zeitreihe (weiß) treten in der Zeit-Frequenz-Darstellung deutlich hervor.

5.3 Modellinferenz

In der Literatur wird immer wieder angeführt, dass künstlich neuronale Netze gegenüber Fuzzy-Regelsystemen den Nachteil des *black box*-Charakters besäßen und dass die linguistischen Regeln aus der Inferenz eines Fuzzy-Reglers immerhin einen Einblick in die Modellierung gewährten (Patterson 1996; Heine 1999; Miima 2002). Gerade die Gewichte eines neuronalen Netzes seien im physikalisch-mathematischen Sinne nicht interpretierbar.
Dabei ist anzumerken, dass hier besonders die leicht unterbestimmten Netzarchitekturen (Kap. 4.5.4) die Probleme bereiten, denn die Gewichte der konvergierten Lösung können hier je nach den zufälligen Startwerten völlig unterschiedlich ausfallen, selbst wenn jeweils das globale Minimum erreicht worden ist. Hier spielt die additive Speicherung ein und derselben Beziehung zwischen den realisierten Systemein- und -ausgängen in mehreren Neuronen die entscheidende Rolle: Ausgehend von den zufällig gewählten Startwerten kann diese Additivität, verstärkt durch die nichtlineare Aktivierung (Tab. 4.1), dabei völlig verschleiert werden. Hinzu kommt, dass dennoch ein *weight sharing* bei einer Untermenge der an der Addition beteiligten Neuronen möglich ist (Kap. 4.5.4).
Eine Untersuchung der Gewichte könnte demnach nur bei ideal bestimmten Netzen, also solchen mit etwa der richtigen VC-Dimension, zu einer Interpretierbarkeit führen. Solche Netze konvergieren global immer zu den gleichen Gewichten (Heinert 2008).
Ein anderer Weg kann durch die Untersuchung einer Systemantwort beschritten werden (Kap. 3.1.6), indem einem trainierten neuronalen Netz standardisierte Eingänge auf jeweils einem Eingangsknoten zugewiesen werden. Die übrigen Knoten sollten am besten konstant auf dem Mittelwert der hier sonst üblichen Eingänge gehalten werden. Die Netzausgabe wird zweierlei Informationen beinhalten: erstens, ob die Modellantwort identifizierend ist (Kap. 4.3.4), zweitens, wenn ja, welcher Funktionsschar sowohl Modell- als auch Systemantwort angehören. Ob die Modellantwort des global konvergierten neuronalen Netzes identifizierend ist, erschließt sich aus ihrer Gedächtnislänge und Beschränktheit. Imitierende Modelle rekursiver MLP-Netze (Abb. 4.10) neigen beispielsweise dazu, exponentiell divergierende Einzelantworten des Typs $-a^k e^{bk}$ zu produzieren, die sich in der Summe auslöschen. Das Modell erscheint dabei als vollständig konvergiert. Ein solches Verhalten deutet an, dass ein Netz möglicherweise unterbestimmt ist. Die Reduktion von Neuronen führt fast immer zur Stabilisierung.
Ist ein identifizierendes Modell erreicht, kann der bis jetzt noch nichtparametrische Modellausgang durch eine parametrische Funktion mithilfe einer numerischen Optimierung (Kap. 4.3.5) approximiert werden, deren Parametern jetzt möglicherweise eine konkrete mathematisch-physikalische Bedeutung zugewiesen werden kann (Kap. 6.1.3).

5.4 Semiparametrisches 2D-Modell

Ein entscheidender Vorteil nichtparametrischer Modelle ist die Möglichkeit, Ein- und Ausgangsgrößen miteinander in Beziehung zu setzen, ohne dass eine funktionale Gesetzmäßigkeit bekannt sein muss. Dennoch wird ein nichtparametrisches Modell das Verhalten des zugrundeliegenden Systems bestmöglich imitieren oder gar identifizieren (Kap. 4.1). Der dazu notwendige Lernprozess lässt sich erheblich beschleunigen, wenn die Ein- und Ausgangsgrößen in einer linearen funktionalen Beziehung stehen oder sich auf der Grundlage von Vorinformationen zueinander in einen linearisierten Grundzustand gebracht werden können (Abb. 5.10 unten). In diesem Fall soll in dieser Arbeit von einem *semiparametrischen* Modell gesprochen werden.

5.4.1 Mechanischer Ansatz

Um die divergente Bewegung mit der Freisetzung seismischer Energie vergleichen zu können, ist weder die Anzahl der Erdbeben noch die Betrachtung einzelner herausragender Ereignisse

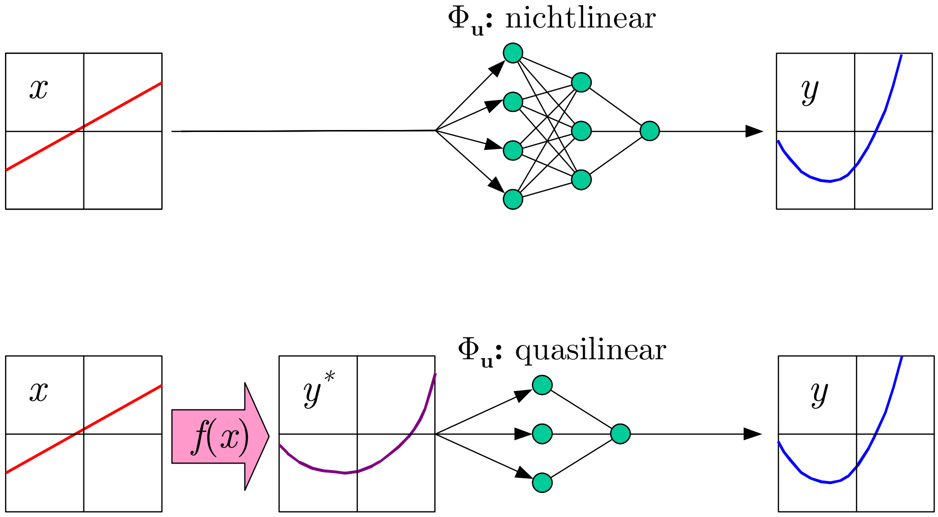


Abb. 5.10: Nicht- und semiparametrisches Modell: Das nichtparametrische Modell muss eine nichtlineare Modellfunktion bilden (oben), während eine Übertragung mit bereits bekannten physikalischen Zusammenhängen das nichtparametrische Modell entlastet (unten). Die Modellfunktion muss das System mit weniger Neuronen nunmehr näherungsweise linear abbilden können.

für sich allein genommen geeignet. Denn es ist ebenso erforderlich, auch die Wirkung vieler kleiner Erdbeben mit berücksichtigen zu können. Der hier gewählte neuartige Ansatz sieht vor, dass die seismische Energie in Abhängigkeit von der lokalen Magnitude M_L und dem logarithmischen *offset* β zunächst gleich der kinetischen Energie einer bewegten Äquivalenzmasse m gesetzt wird und diese in eine Geschwindigkeit

$$\dot{\Phi}_{y_s} = \frac{\partial \Phi_{y_s}}{\partial t} = \frac{a_0}{m}\sqrt{10^{\frac{3}{2}M_L+\beta}} \tag{5-22}$$

übertragen werden kann. Der Vorfaktor a_0 drücke den Anteil der überführten Energie aus. Das Zeitintegral über die Geschwindigkeit

$$\Phi_{y_s}(t) = \int_t \frac{\partial \Phi_{y_s}(t)}{\partial t} dt. \tag{5-23}$$

lässt sich aufgrund der äquidistant realisierten Systemein- und -ausgänge in der diskreten Summenform

$$\Phi_{y_s}(t) = \sum_t \frac{\Delta \Phi_{y_s}(t)}{\Delta t} \tag{5-24}$$

ausdrücken. Der äquidistante Abstand Δt entspreche jeweils genau einem Tag. Damit kann der zurückgelegte Weg an einem Tag

$$\Phi_{y_s}(\Delta t) = \frac{a_0}{m} \sum_{\tau \geq t}^{t+\Delta t} \sqrt{10^{\frac{3}{2}(M_L)_\tau+\beta}} \tag{5-25}$$

in Abhängigkeit der Energie aus allen Erdbeben eines Tages beschrieben werden. Die Größe $\beta = 11,4$ ist in (2-4) global definiert. Die Äquivalenzmasse m der bewegten Kruste hingegen kann nur als Formparameter vorab grob geschätzt werden. Summiert man die Wege eines Tages $\Phi_{y_s}(\Delta t)$ über die Zeit t, so ergibt sich der gesamte Weg

$$\Phi_{y_s}(t) = \frac{a_0}{m} \sum_t \left(\sum_{\tau \geq t}^{t+\Delta t} \sqrt{10^{\frac{3}{2}(M_L)_\tau+\beta}} \right) \tag{5-26}$$

als Summe aller diskreten Tagesenergien. Nun sei Vorfaktor a_0 nicht konstant für alle Ereignisse, sondern variiere in Abhängigkeit von der Zone z, in welcher das Erdbeben stattfindet (Abb. 5.4). Der Weg wird dann vollständig durch die Dreifachsumme

$$\Phi_{y_s}(t) = \frac{1}{m} \sum_t \left(\sum_{\tau \geq t}^{t+\Delta t} \left(\sum_z a_0^z \sqrt{10^{\frac{3}{2}(M_L)_\tau + \beta}} \right) \right) \tag{5-27}$$

beschrieben.

5.4.2 Periodischer Ansatz

Wie durch die spektrale Analyse bereits ersichtlich geworden ist, enthalten die trendfreien Prozesse der Längenänderungen zwischen den IGS-Stationen eine Reihe von periodischen wie auch quasi-periodischen Subprozessen.

Die täglichen und jährlichen Variationen der Atmosphäre und des Klimas erscheinen als ein maßgeblicher Einfluss auf die Messungen. Diese Variationen ergeben sowohl physikalisch als auch numerisch eine Reihe sich überlagernder Jahreswellen. Einerseits interferieren tägliche Änderungen in Verbindung mit der sich jährlich vollständig wiederholenden Satellitengeometrie in der Art, dass die tägliche Auswertung einen numerischen Jahresalias produziert. Andererseits existieren echte jährliche klimatische Variationen mit einem Einfluss auf die Messungen. Diese Einflüsse wirken entweder nur auf die Messung selbst, ohne dass sich der Sensor tatsächlich bewegt, oder es gibt eine reale Wechselwirkung zwischen den klimatischen Variationen mit den Eigenschaften des Untergrundes direkt auf die Position des Sensors.
So wurde für die Messungen auf Island bereits erfolgreich ein Modell verwendet, das die jährliche Schwankung der Schneelast als mögliche Ursache für jährliche Positionsänderungen erscheinen lässt (Grapenthin et al. 2006). Da die Schneelast allerdings nicht direkt aus Messdaten, sondern ihrerseits als singuläre Jahreswelle modelliert wurde, ist die Aussagekraft dieses Modells hinsichtlich der Ursachen nur eingeschränkt tauglich. Auf diese Weise ergibt sich hier allenfalls ein korrelativer Zusammenhang.
Die Summe verschiedener jährlicher und täglicher, aber als jährlich wirkender, Einflüsse sind geeignet, um über den gesamten Zeitraum in Frequenz ν, Phasenlage ϕ und Amplitude A einen stationären Systemeingang zu erzeugen. Daher wird auch der entsprechende Systemausgang in Form der prozessierten Koordinaten hier die gleiche Stationarität besitzen. Um diesen stationären Subprozess mit der Grundfrequenz $\nu_c = 1\mathrm{a}^{-1}$ darzustellen, werden fünf weitere Frequenzen mit dem ganzzahligen Vielfachen $n_i = 2, 3, 4, 6, 8$ der Grundfrequenz[7] mit aufaddiert:

$$\Phi_{y_c}(t) = \sum_i A_i \cos\left(2\pi \left(t - t_0\right) \nu_c \cdot n_i + \phi_i\right) \; \forall n_i \in \mathbb{N}. \tag{5-28}$$

Diese Darstellung kann gemäß (Kap. 3.2.1) überführt werden in die linearisierte Form

$$\Phi_{y_c}(t) = \sum_i a_i \cos\left(2\pi \left(t - t_0\right) \nu_c \cdot n_i\right) + b_i \cdot i \sin\left(2\pi \left(t - t_0\right) \nu_c \cdot n_i\right). \tag{5-29}$$

mit den komplexen Fourier-Parametern a_i und b_i.

Dieser periodische Ansatz dient demnach lediglich der Reduktion der Summe verschiedener jährlicher oder jährlich wirkender Einflüsse, soll und wird aber keine Aussage in Bezug auf deren Ursache erlauben.

[7] Diese Vielfachen der Grundfrequenz sind heuristisch bestimmt und haben sich in dieser Kombination bereits mehrfach zum Modellieren von klimatischen Einflüssen bewährt.

5.4.3 Transienter Ansatz

Als transienter Prozess wird die lineare zeitabhängige Funktion

$$\Phi_{y_{as}}(t) = a_0^* \, (t - t_0) \tag{5-30}$$

gewählt. Sie hat die Aufgabe, ein mögliches aseismisches Kriechen mit abzubilden. Dazu ist der Übertragungsparameter a_0^* zu bestimmen. Die Festlegung von a_0^* wird aber nur schwache Aussage über aseismische Bewegungen treffen können. Bedingt durch die kurzen Zeitreihen wirken die Zonenparameter a_0^z durchaus ähnlich. Darüber hinaus sind die Parameter a_0^* und a_0^z gemeinsam auch noch maßgeblich von der Wahl der als konstant in (5-27) eingeführten Größen β und m abhängig. Aus diesem Grund ist es annähernd unmöglich, den Anteil des aseismischen Kriechens a_0^* an der Gesamtbewegung mithilfe dieses Ansatzes präzise zu bestimmen.

5.4.4 Semiparametrische Modellbildung

Das Modell ergibt sich aus der Summe der drei einzeln vorgestellten Ansätze

$$\Phi_{y_t} = \Phi_{y_0}(t) + \Phi_{y_s}(t) + \Phi_{y_c}(t) + \Phi_{y_{as}}(t). \tag{5-31}$$

mit einer zusätzlichen Additionskonstante Φ_{y_0}. In der so beschriebenen Form enthält es 16 Unbekannte, die in Abhängigkeit zueinander bestimmt werden müssen. Hierzu wird ständig der Abstand $v = y_t - \Phi_{y_t}$ zwischen Modell und realisiertem Prozess bestimmt, um die Anpassung der Parameter zu steuern.

5.4.5 Optimierungsansatz

Zur Optimierung dieses Modellansatzes bedarf es eines neuartigen robusten Schätzers. Dieser darf nicht durch die in Kap. 4.3.2 diskutierten Nachteile der robusten Schätzung in Hinblick auf die Anwendung von nichtlinearen Optimierungsverfahren eingeschränkt sein. Eine Verlustfunktion sollte sich in der Nähe des Optimums gemäß der Fehlerquadratmethode verhalten, da hier nach vor von der Normalverteilung ausgegangen werden kann. Mit zunehmender Entfernung vom Optimum darf einer Stichprobe nur noch ein stetig abnehmender Einfluss zugebilligt werden. Einen konstanten Wert darf sie allerdings auch in größter Entfernung in Hinblick auf die Optimierung nicht annehmen. Die Einflussfunktion

$$\Psi(v)_{\ln} = \frac{kv}{k + v^2} \tag{5-32}$$

erfüllt zunächst die Gleichheit zur Einflussfunktion des L2-Schätzers in dem durch k bestimmten Nahbereich des Optimums (Abb. 5.11a). Sie ist weiterhin beschränkt und fällt schließlich ähnlich den *redescending* Ψ*-functions* zurück. Dennoch nimmt diese Einflussfunktion niemals den Wert null an, so dass eine polynomiale Optimierung (Kap. 4.3.5) auch mit schlechten Startwerten möglich bleibt. Anders als die übrigen *redescending* Ψ*-functions* ist (5-32) vollständig differenzierbar. Die zugehörige Verlustfunktion

$$\varrho(v)_{\ln} = \int \frac{kv}{k + v^2} dv = \frac{k}{2} \ln(k + v^2) + C, \tag{5-33}$$

soll daher in dieser Arbeit favorisiert werden. Sie genügt allen oben genannten Forderungen. Um andere stochastische Subprozesse als einen Gauss-Prozess in der Schätzung berücksichtigen zu können, werde eine Unschärfe um das Optimum

$$v^* = \begin{cases} (|v| - \xi) \cdot \operatorname{sign}(v) & \forall \quad |v| \geq \xi \\ 0 & \forall \quad |v| \leq \xi \end{cases} \tag{5-34}$$

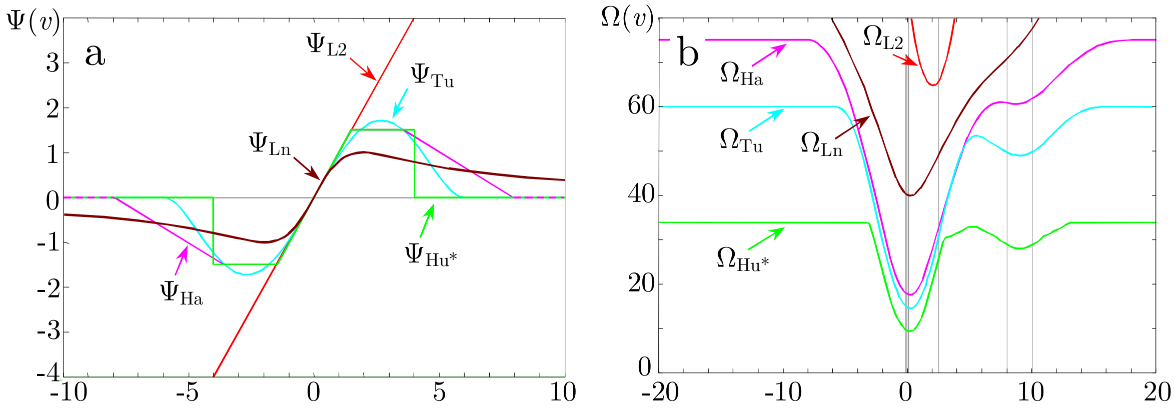


Abb. 5.11: Vergleich der a) Einflussfunkion des Ln-Schätzers und b) der Verlustfunktion über eine kontaminierte Stichprobe (grau) mit den Charakteristika des L2-Schätzers und den *redescendent* Ψ-*functions*.

mit

$$v(t) = y(t) - \Phi_y(t)$$

mittels einer sogenannten *slack variable* ξ definiert. Die Parameter des Modells (5-31) können optimiert werden, indem das Minimum der Risikofunktion

$$\Omega(v) = \frac{k}{2} \sum_t \ln \left(k + \left(\frac{v^*(t)}{s(t)} \right)^2 \right) \tag{5-35}$$

bestimmt wird.

5.5 Semiparametrisches 3D-Modell

Ein entscheidender Unterschied dieses Modells zum zuvor vorgestellten semiparametrischen 2D-Modell ist die Trennung von Zeit- und Ortsfunktionen. Es trägt damit der Tatsache Rechnung, dass Erdbeben zu nichtäquidistanten Zeiten in unterschiedlichen Regionen erfolgen können. Aus diesem Grund wird die infinite Impulsantwort auf jedes Erdbeben als ein allen Ereignissen gemeinsamer autoregressiver Prozess approximiert. In Abhängigkeit des Ortes erhält jedes Ereignis eine Gewichtung. Diese ist weder individuell noch in Klassen sortiert. Vielmehr wird eine Nachbarschaftsähnlichkeit erzwungen, indem die Gewichte durch eine glättende Oberfläche verbunden werden. Als besonders gutmütig haben sich hierbei die radialen Basisfunktionen (RBF) (4-39) aus der Nutzung der heteroassoziativen neuronalen Netze erwiesen.

5.5.1 Ansatz des Zeitmodells

Das Zeitmodell hat die Aufgabe, die Systemantwort auf einen beliebigen Systemeingang zu erlernen und später zu prädizieren. Während im vorhergehenden Modell die Systemantwort fest vorgegeben worden ist, wird sie in diesem Modell zum Gegenstand des Lernens und damit auch des Modells.

Dabei sind zunächst drei Fragen durch eine Modellbildung mit synthetischen Datensätzen zu klären: Erstens, hat eine gewählte Netzarchitektur eine hinreichende Kapazität, die Funktionsschar der zu erwartenden Systemantworten abzubilden und damit Teil eines identifizierenden Modells zu werden, zweitens, kann ein Zeitmodell zwar rekursiv, aber dennoch inäquidistant sein und drittens, ist es sinnvoller mit der gewählten Architektur eine Impuls- oder eine Sprungantwort zu modellieren?

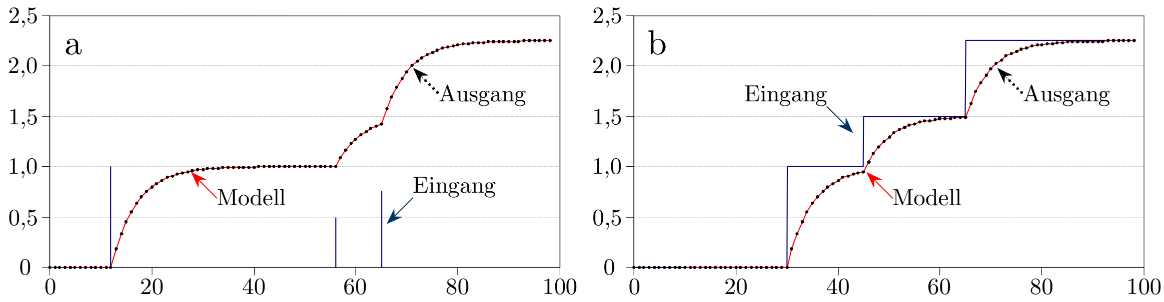


Abb. 5.12: Reaktion eines ARX-Modells bzw. eines linearen einschichtigen MLFF-Netzes als a) Impulsantwort und b) Sprungantwort auf eine äquidistante Distribution.

Synthetische Modellierung

Es existiere eine äquidistante Realisierung x_k mit der die Folge von Ereignissen x_i erfasst werden kann. Diese habe eine negativ-exponentielle Systemantwort

$$y_k = \sum_{i=1}^{n} \begin{cases} a_i \left(1 - e^{-b_i(x_k - x_i)}\right) & \forall\,(x_k - x_i) > 0 \\ 0 & \forall\,(x_k - x_i) \leq 0 \end{cases} \tag{5-36}$$

eines dynamischen Systems mit einer unendlichen Gedächtnislänge zu Folge. In der praktischen Modellbildung lässt sich die Antwort, sei sie nun eine Impuls- oder eine Sprungantwort (Abb. 5.12), mit *einem* rekursiven Neuron mit unbeschränkt linearer Aktivierung

$$y_k = \mathcal{N} + \sum_{i=0}^{2} w_i x_{k-i} + \sum_{i=1}^{3} u_i y_{k-i} \tag{5-37}$$

vollständig abbilden. Ein zusätzliches Neuron trägt zur Modellbildung nicht mehr bei, sondern die Abbildung wird lediglich willkürlich auf zwei Neuronen verteilt. Damit lässt sich zwar einerseits die Anzahl der zur Konvergenz notwendigen Iterationen etwas vermindern, aufgrund der größeren Anzahl von freien Gewichten erhöht sich damit aber auch wieder die Rechenzeit einer einzelnen Iteration. Befindet sich dieses Doppelneuron in einer größeren Netzwerkarchitektur, kann diese Überkapazität aufgrund des *weight sharings* zur Unterbestimmung führen (Kap. 4.5.4), so dass auf ein zweites Neuron verzichtet werden sollte. Weiterhin erschwert die überflüssige zusätzliche Architektur die spätere Inferenz des Modells.

Die Modellierung äquidistanter Daten ist erwartungsgemäß lösbar (Abb. 5.12). Es stellt sich die Frage, ob auch inäquidistante Daten hinreichend genau modellierbar sind. Daher soll auch ein solches Beispiel getestet werden.

Es existiere eine inäquidistante Realisierung $x(t_k)$ mit der die Folge von n Ereignissen zum jeweiligen Zeitpunkt (t_0, i) erfasst werden kann. Diese habe eine negativ-exponentielle Systemantwort

$$y(t_k) = \sum_{i=1}^{n} \begin{cases} a_i \left(1 - e^{-b_i(x(t_k) - (x(t_0,i))}\right) & \forall\,(x(t_k) - (x(t_0,i)) > 0 \\ 0 & \forall\,(x(t_k) - (x(t_0,i)) \leq 0 \end{cases} \tag{5-38}$$

eines dynamischen Systems mit einer unendlichen Gedächtnislänge zu Folge. Der erste Lösungsansatz, der nach verschiedenen Fehlversuchen erstmals vertretbare Resultate lieferte, ist ein zweischichtiges Netz

$$\begin{aligned} y(t_k) \quad = \quad & w_{0,1}\varphi_{/}\left(\mathcal{N}_1 + w_{1,2}x(t_{k-1}) + u_{1,1}y(t_{k-1})\right) \\ & + w_{0,2}\varphi_{/}\left(\mathcal{N}_2 + \sum_{i=1}^{2} w_{2,i}x(t_{k-i}) + u_{2,i}y(t_{k-i})\right) \\ & + w_{0,3}\varphi_{/}\left(\mathcal{N}_3 + \sum_{i=2}^{3} w_{3,i}x(t_{k-i}) + u_{3,i}y(t_{k-i})\right) \end{aligned} \tag{5-39}$$

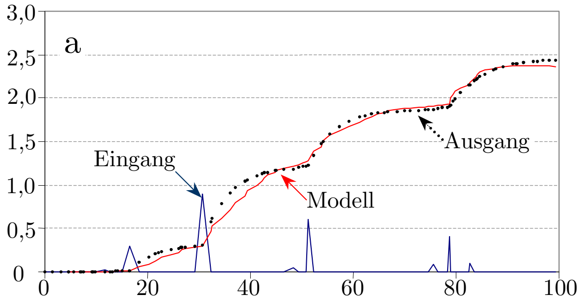


Abb. 5.13: Reaktion eines a) linearen, rekursiven, zweischichtigen Differenzen-Netzes als Impulsantwort und b) eines linearen, rekursiven, zweischichtigen Transitions-Netzes als Sprungantwort auf eine inäquidistante Distribution.

mit drei Neuronen in der verdeckten Schicht und einem Ausgabeneuron. Den drei Neuronen wird nur jeweils Zugriff auf den Systemeingang, bestehend aus Impulsen, und dem Systemausgang aus zwei Beobachtungsepochen gewährt, so dass sich in den Gewichten ablesbar von selbst Differenzen ergeben. Dieses „Differenzen"-Netz approximiert den Systemausgang zwar recht gut (Abb. 5.13a), das trainierte Netz ist mit neuen Systemeingängen aber nicht prädiktionsfähig. Das heißt, es imitiert lediglich die Prozesse.
Der folgende Netztyp hat sich als erfolgreicher herausgestellt. Zunächst bestand der Systemeingang bereits aus einer Sprungfolge (Abb. 5.13b). In Analogie zur Standardtransition eines Kalman-Filters (Niemeier 2002, S. 384) werden die drei Neuronen der verdeckten Schicht auch mit den ersten und zweiten Ableitungen der Systemein- und -ausgänge aus drei Beobachtungsepochen trainiert. Dieses „Transitions-"Netz

$$\begin{aligned} y_k \;=\; & \mathcal{N}_0 + \sum_{\iota=1}^{2} w_\iota \varphi_/ \Bigg(\mathcal{N}_\iota + \sum_{i=1}^{4} w_{\iota i} \varphi_/ \Bigg(\mathcal{N}_i + w_{\iota i0} x_k + \sum_{j=1}^{3} \left(w_{\iota ij} x_{k-j} + u_{\iota ij} y_{k-j} \right) \\ & + \sum_{j=1}^{3} \left(w_{\iota ij} \frac{\partial x_{k-j}}{\partial t} + u_{\iota ij} \frac{\partial y_{k-j}}{\partial t} \right) + \sum_{j=1}^{3} \left(w_{\iota ij} \frac{\partial^2 x_{k-j}}{\partial t^2} + u_{\iota ij} \frac{\partial^2 y_{k-j}}{\partial t^2} \right) \Bigg) \Bigg) \end{aligned} \tag{5-40}$$

approximiert den Systemausgang recht gut und prädiziert auch den Systemausgang von bisher unbekannten Systemeingängen. Dieses Ergebnis hat sich allerdings erst eingestellt, nachdem jedes der drei Neuronen der verdeckten Schicht nur Zugriff auf je eine Beobachtungsepoche erhielt. Die Beschränkung der Anzahl der Gewichte war hier unbedingt notwendig.

Reale Modellierung

Die Problematik der Rekursion über inäquidistante Systemeingänge ist aus den synthetischen Modellierungen deutlich geworden. Eine Rekursion ist zwar möglich, aber eine vollständige Abbildung misslingt auch mit der besten der verfügbaren Lösungen (Abb. 5.13b). Aus diesem Grund soll die Rekursion über äquidistante Eingangsdaten durchgeführt werden. Da der Systemeingang aus Erdbeben aber *per se* nichtäquidistant ist, muss die Äquidistanz ohne Verlust oder Verfälschung der Eingangsinformation erzwungen werden. Das erfolgt nun über einen neuartigen Ansatz, der eine stabilisierende Wirkung auf die Rekursion über inäquidistante Daten besitzt: Über die Uhrzeit eines Bebens kann die unscharfe Zugehörigkeit eines Ereignisses zum jeweiligen Tag

$$\mu_{0,i} = \left(1 - \frac{t_{h,i}}{24} \right) \tag{5-41}$$

bestimmt werden. Die anteilige Energie des Ereignisses wird mit der aller anderen Ereignisse des Tages summiert. Die unscharfe Zugehörigkeit zum Folgetag

$$\mu_{1,i} = 1 - \mu_{0,i} \tag{5-42}$$

bestimmt das Energiekomplement, das erst dann wirksam wird. Die Summation aller Energieäquivalente muss aber noch mit der ortsabhängigen dreidimensionalen Gewichtung $a_\tau(\mathfrak{x})$ angepasst werden. Damit ergibt sich die modellierte Distanzänderung aufgrund der Seismizität zwischen zwei IGS-Stationen zu

$$\Phi_{y_s}(t) = \frac{1}{m} \sum_t \left(\sum_{\tau > t}^{t+\Delta t} a_\tau(\mathfrak{x}) \mu_{0,\tau} \sqrt{10^{\frac{3}{2}(M_L)_\tau + \beta}} + \sum_{\vartheta \geq t-\Delta t}^{t} a_\vartheta(\mathfrak{x}) \mu_{1,\vartheta} \sqrt{10^{\frac{3}{2}(M_L)_\vartheta + \beta}} \right). \tag{5-43}$$

Mithilfe der ortsabhängigen dreidimensionalen Gewichtungen $a_\tau(\mathfrak{x})$ bekommen die sonst an dieser Stelle üblichen nichtparametrischen Gewichte w einen deterministischen Charakter. Auch die Systemeingänge stehen in einem deterministisch linearisierten Verhältnis zu den Systemausgängen. Damit ist dieses Netz nicht mehr rein nichtparametrisch, sondern kann als semiparametrisch angesehen werden. Damit ist es nicht mehr notwendig, die Linearisierung von einem neuronalen Netz mit einer höheren Kapazität vornehmen zu lassen. Damit sinkt die Gefahr eines *overfitting* (Kap. 4.5.4) und es ist eine zuverlässigere Konvergenz zu erwarten.

5.5.2 Ansatz des Ortsmodells

Das Ortsmodell hat die Aufgabe, den Systemeingang, hier sind es Erdbeben, gemäß seines Ursprungsortes zu gewichten. Eine derartige Gewichtung repräsentiert den Einfluss der Erdbeben in Abhängigkeit der Position des Epizentrums und der Tiefe des Hypozentrums. Aber die Stärke eines Bebens kann auch auf Grund mittelbarer Informationen gewichtet werden: Zwar werden die Ausrichtung der Herdflächen zur Bewegungsrichtung nicht direkt berücksichtigt, aber die Annahme, dass eine bestimmte Region eine bevorzugte Ausrichtung der Herdflächen besitzt und Beben entsprechend mehr oder weniger Einfluss auf die Bewegung besitzen, ist zwar vereinfachend aber dafür praktisch und effizient.

Synthetische Modellierung

Es stellt sich die Frage, welche Netzarchitektur geeignet ist, um eine kontinuierliche dreidimensionale Ortsfunktion richtig abzubilden und dabei auch ein zuverlässiges Konvergenzverhalten besitzt. Die Versuche mit MLP-Netzen für zunächst nur eine zweidimensionale Aufgabe waren sehr unterschiedlich erfolgreich. Gerade die gebräuchlichsten Netze mit einer sigmoidalen Aktivierung (Tab. 4.1) kommen zu verfälschten Abbildungen, die aber für diese Aktivierungsfunktion charakteristisch sind (Abb. 5.14a, b). So neigen überbestimmte MLP-Netze mit sigmoidaler Aktivierung dazu, dreieckige (vergl. Abb. 4.9f) oder lanzettförmige Gewichtstopografien auszubilden, auch wenn der reale Zusammenhang beispielsweise eine elliptische Topografie besitzt (Abb. 5.14f). Wenngleich das 8-2-1-Netz (Abb. 5.14d) eine doch gute Approximation liefert, kann man ersehen, dass das gleiche Netz unter anderen Startbedingungen völlig anders konvergiert (Abb. 5.14c).
Dieses Verhalten von Netzen mit überwiegend sigmoidaler Aktivierung zeigt, dass die Sigmoidfunktion besser für Aufgaben der Mustererkennung, als für regressive Approximationsaufgaben verwendet werden sollte.
Überraschend erfolgreich und zuverlässig erwiesen sich Versuche mit einer Sinus-Aktivierung (Abb. 5.14e). Auch wenn die Approximation deutliche Schwächen aufweist, so konvergierte diese Lösung durchaus zuverlässig. Außerdem war auch nur dieses Netz in der Lage, die kleine

Topografie im oberen Bildfeld zu berücksichtigen. Zwar ist sie in diesem Netz überrepräsentiert, aber das ist auch der Randlage und der Stichprobenverteilung zuzuschreiben.
Die Zuverlässigkeit der Dreiecksfunktionen für Approximationen zeigt sich auch bei der FOURIER-Transformation und der Kugelfunktionsentwicklung.

Reale Modellierung

Die Vorinformationen aus der Literatur werden durch die eigenen Tests mit den MLP-Netzen bestätigt. Demnach eignen sich die RBF-Netze besser für regressive Approximationen von kontinuierlichen Ortsfunktionen. Deshalb wird zur Berechnung jeder ortsabhängigen dreidimensionalen Gewichtung

$$a_\tau(\mathfrak{x}) = \sum_{k=1}^{\mathcal{H}} w_{k,0} \cdot e^{\left(\sum\limits_{\kappa=1}^{3} (\mathfrak{x}_\tau - \chi_{k,\kappa})^2 w_{k,\kappa}\right)} \tag{5-44}$$

angesetzt. Dabei bezeichnet $w_{k,0}$ die Amplitude der RBF-Funktion, $w_{k,\kappa}$ die inverse Varianz der GAUSS-Aktivierung und $\chi_{k,\kappa}$ die dreidimensionalen Koordinaten des Stützvektors.

5.5.3 Kombinierter Ansatz

Im letzten Schritt der Modellbildung werden das Orts- und das Zeitmodell zusammengefasst. Die besondere Schwierigkeit dieses Schrittes liegt in der Kombination der Gewichte des Ortes a, der Erhaltensneigung α, β und der Tageszugehörigkeit μ eines einzelnen Bebens. Während a und μ vor der Erzwingung der Äquidistanz des Systemeingangs anzubringen sind, wirken α, β (auto-)regressiv über bereits äquidistanten Systemeingänge.

Synthetische Modellierung

Der modellierte Systemausgang $\Phi_{y_s}(t)$ ist entsprechend (5-43) eine mit $a_\tau(\mathfrak{x})$ gewichtete Summe über alle seismischen Ereignisse innerhalb eines äquidistanten Zeitintervalls Δt. Dieses zu äquidistanten Zeitabständen errechnete Zwischenergebnis wird über alle Systemeingänge des Beobachtungszeitraumes summiert. Um die Erhaltensneigung modellieren zu können, werden die mit β gewichteten Systemeingänge summiert. Sie entsprechen dem in (4-29) eingeführten externen Eingang eines ARX[p, e]-Prozesses. Die autoregressive Antwort des Modells ergibt sich aus der mit α_δ gewichteten Summe der zurückliegenden äquidistanten Antworten $\Phi_{y_s}(t - \delta)$. Der modellierte Gesamtsystemausgang

$$\begin{aligned} \Phi_{y_s}(t) \quad = \quad & \sum_{\delta=0}^{q} \left(\beta_\delta \sum_{t} \left(\sum_{\tau > t-\delta}^{t-\delta+\Delta t} a_\tau(\mathfrak{x}) \mu_{0,\tau} \cdot y_\tau^* + \sum_{\vartheta \geq t-\delta-\Delta t}^{t-\delta} a_\vartheta(\mathfrak{x}) \mu_{1,\vartheta} \cdot y_\vartheta^* \right) \right) \\ & + \sum_{\delta=1}^{p} \left(\alpha_\delta \cdot \Phi_{y_s}(t-\delta) \right) \end{aligned} \tag{5-45}$$

wird zum Zwecke einer Vereinfachung der synthetischen Modellierung aber nicht über Magnituden m_t von Erdbeben durchgeführt, sondern über imaginäre maximale Streckenäquivalente $y_t^* = f(m_t)$.

Dabei sollen abschließend verschiedene Aktivierungsfunktionen der RBF-Funktion getestet werden.

$$a_\tau(\mathfrak{x}) = \sum_{k=1}^{\mathcal{H}} w_{k,0} \cdot \varphi\left(\sum_{\kappa=1}^{3} (\mathfrak{x}_\tau - \chi_{k,\kappa})^2 w_{k,\kappa} \right) \tag{5-46}$$

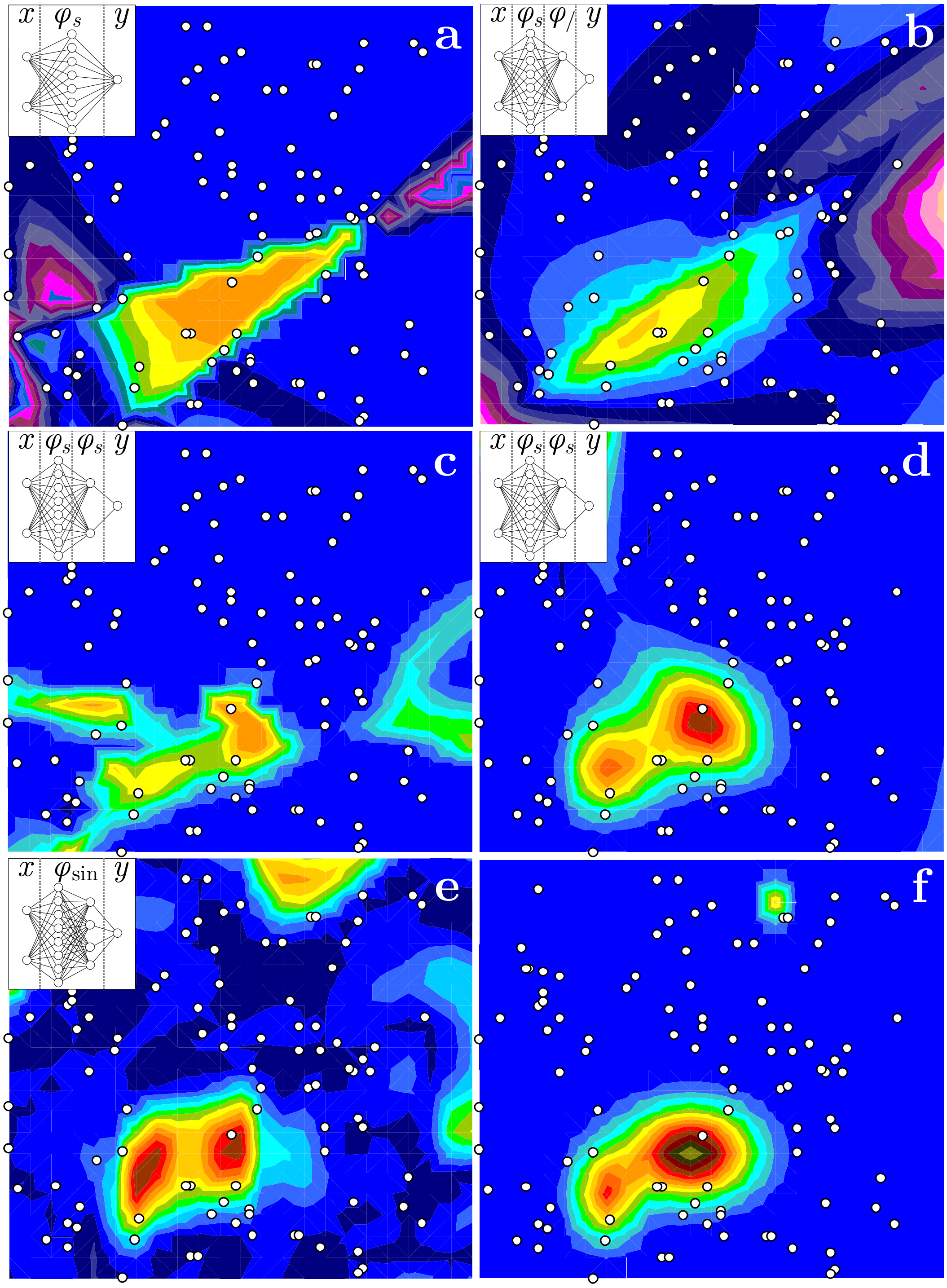


Abb. 5.14: Verschiedene Approximationen einer kontinuierlichen zweidimensionalen Ortsfunktion (f) in Abhängigkeit der MLP-Netzarchitektur und den Aktivierungsfunktionen: a) ein 2-8-1 Netz mit sigmoidaler Aktivierung, b) ein 2-8-2-1 Netz mit sigmoidaler Aktivierung in der ersten und unbegrenzt linearer Aktivierung in der zweiten Schicht, c) und d) jeweils ein 2-8-2-1 Netz mit sigmoidaler Aktivierung aber unterschiedlichen Ergebnissen in Abhängigkeit der Startwerte und e) ein 2-8-4-1 Netz mit sinusoider Aktivierung. Die Netzarchitektur ist im Obereck abgebildet. Zur Approximation stand ein synthetischer Testdatensatz von 100 Stichproben (weiße Punkte) zur Verfügung.

Reale Modellierung

Die Modellierung mithilfe der realen Daten muss auf die Vereinfachung des Systemeingangs über die Streckenäquivalente entsprechend (5-45) wieder verzichten. Statt dessen wird hier auf die bereits getestete Linearisierung (5-26) zurückgegriffen. Anders als im ersten Modell existieren weder diskrete Zonen gleichen Einflusses noch Modellvorgaben hinsichtlich der Systemantwort. Damit errechnet sich der seismische Anteil der modellierten Bewegung zwischen zwei IGS-Stationen mit

$$\begin{aligned}\Phi_{y_s}(t) \quad = \quad & \sum_{\delta=0}^{q} \left(\beta_\delta \sum_{t} \left(\sum_{\tau > t-\delta}^{t-\delta+\Delta t} \left(\sum_{k=1}^{\mathcal{H}} w_{k,0} \cdot e^{\left(\sum\limits_{\kappa=1}^{3} \left(\mathfrak{x}_\tau - \chi_{k,\kappa} \right)^2 w_{k,\kappa} \right)} \right) \mu_{0,\tau} \cdot 10^{\frac{3}{4}(M_L)_\tau + \frac{1}{2}\beta} \right.\right. \\ & \left.\left. \qquad + \sum_{\vartheta \geq t-\delta-\Delta t}^{t-\delta} \left(\sum_{k=1}^{\mathcal{H}} w_{k,0} \cdot e^{\left(\sum\limits_{\kappa=1}^{3} \left(\mathfrak{x}_\tau - \chi_{k,\kappa} \right)^2 w_{k,\kappa} \right)} \right) \mu_{1,\vartheta} \cdot 10^{\frac{3}{4}(M_L)_\vartheta + \frac{1}{2}\beta} \right) \right) \\ & + \sum_{\delta=1}^{p} \left(\alpha_\delta \cdot \Phi_{y_s}(t-\delta) \right). \end{aligned} \tag{5-47}$$

Die vollständige modellierte Systemantwort besteht aus einer Additionskonstante Φ_{y_0}, den periodischen Subprozessen $\Phi_{y_c}(t)$ auf eine individuelle Station (5-28) und einem aseismischen Kriechen $\Phi_{y_{as}}(t)$ (5-30). Analog zum ersten Modell (5-31) ergibt sich die modellierte Streckenänderung zu

$$\Phi_{y_t} = \Phi_{y_0}(t) + \Phi_{y_s}(t) + \Phi_{y_c}(t) + \Phi_{y_{as}}(t). \tag{5-48}$$

Kapazität des 3D-Modells

Es liegen 2880 Datenpaare vor, mit einem vierdimensionalen Systemeingang, nämlich dem dreidimensionalen Ort und der Magnitude. Folgend der Anweisung zur Berechnung der geeigneten Knotenanzahl (4-48) wären demnach $\mathcal{H} = 576$ RBF-Knoten notwendig und $\mathcal{H} = 1142$ hinreichend. Tatsächlich verwendet werden jedoch 16 Knoten. Demnach müsste das Modell extrem überdeterminiert sein (Abb. 4.11).
Die Anweisung berücksichtigt nicht die hier gewählte rekursive Architektur. Zum einen geht jedes Ereignis unscharf an zwei auf einander folgenden Tagen ein, zum anderen in ein ARX[8,8]-Modell. Das bedeutet, dass ein einzelnes Ereignis achtmal zeitlich versetzt gewichtet und abermals achtmal gewichtet autoregressiv als Eingang eingeht. Der Einfluss jedes der 16 Knoten ist also nicht jeweils einmal, sondern 128-mal. Damit liegt die effektive Knotenzahl bei $\mathcal{H} = 1024$, also bereits am Rande zur Unterdeterminiertheit. Die Unterdeterminiertheit würde zur Folge haben, dass das Modell zwar sehr gut approximiert, dass aber die sowohl die Raum- als auch die Zeitfunktion einer Kreuzvalidierung bedürften. Dennoch wäre nicht gewährleistet, dass nun noch eine brauchbare Prädiktion möglich wäre.

6 Beschreibung der seismischen Kinematik

„Erst leugnen die anderen, dass es stimmt,
dann leugnen sie, dass es wichtig ist;
und schließlich schreiben sie das Verdienst dem Falschen zu.“
Alexander von Humboldt

6.1 Ergebnisse

Nach der Bearbeitung der vorhandenen Daten mit den eigens angepassten Algorithmen zur Analyse und Modellbildung (Kap. 5) ergeben sich im Rahmen dieser Arbeit grundsätzlich drei Ergebnisarten: Analysen basierend auf den Daten, Approximation der Daten durch Modellanpassungen und weiterhin die Inferenz der Modelle.
Hierzu müssen zwei Fragen beantwortet werden: zum einen, inwieweit die unterschiedlichen Ergebnisse konsistent und in Bezug auf das Untersuchungsobjekt plausibel sind und zum anderen, ob die getroffenen Adaptionen hinsichtlich der Fragestellung geeignet und zuverlässig sind.

6.1.1 Analyseergebnisse

Die Wetterdaten ermöglichen zwar keine Nutzung als Modelleingang, eröffnen aber die Möglichkeit korrelativer Betrachtung in Hinblick auf die Erdbebendaten der Südisländischen Seismischen Zone (Kap. 2.6). Ein Zusammenhang zwischen der Verfügbarkeit von Schmelzwasser und dem Auftreten von Erdbeben müsste sich mittelbar in der Kreuzkorrelation zwischen den Erdbebendaten und den Wetterdaten zeigen. Eine kurze Vorüberlegung ist notwendig, um die erwartungsgemäß kleinen Korrelationen hinsichtlich ihrer Signifikanz bewerten zu können. Dazu ist die empirische Varianzschwelle als maximaler empirischer Kreuzkorrelationskoeffizent zwischen einer harmonischen Welle und einem Gauss-Prozess (Kap. 3.2.2) in Abhängigkeit der Anzahl der Datenpaare errechnet worden (Abb. 6.1, links).
Es zeigt sich, dass die Kreuzkorrelationsfunktion bis zu einer Zeitverschiebung von vier Jahren zyklisch signifikante Korrelationen bis zu 0,25 erbringt (Abb. 6.1, rechts). Diese sind eine

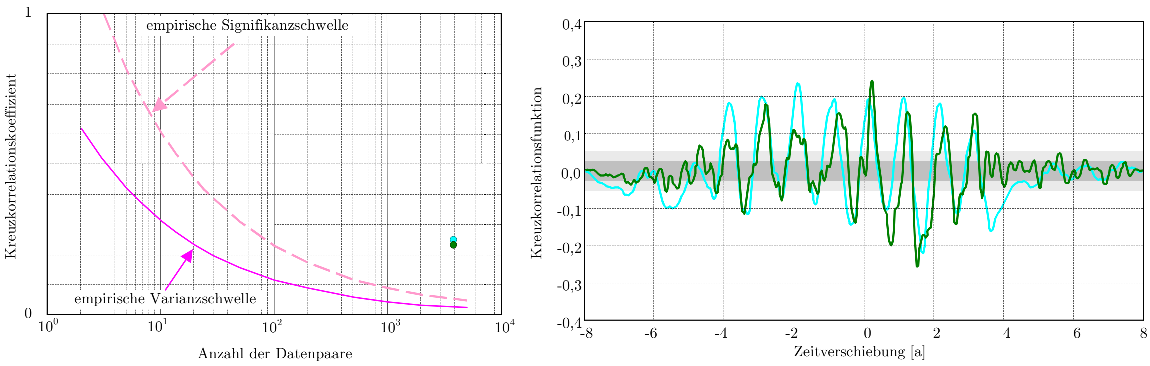


Abb. 6.1: Kreuzkorrelationen zwischen Wetter- und Erdbebendaten sowie deren Signifikanz: Die Anzahl der Datenpaare für die Kreuzkorrelation der Temperatur (links hellblau) und des Luftdruckes (dunkelgrün) mit den Erdbebendaten der SISZ liegt der empirischen Signifikanzschwelle. Die Kreuzkorrelationsfunktion ist innerhalb von vier Jahren signifikant (rechts, hellgrau) und übersteigt bei größeren Zeitverschiebungen nur gelegentlich über die Varianzschwelle (dunkelgrau).

Folge und abermals Nachweis der temporär auftretenden Zyklizität der seismischen Ereignisse in der Südisländischen Seismischen Zone (Kap. 2.6). Die multiple Zeitfrequenzanalyse (Kap. 5.2.2) über die Längenänderungen der Basislinie REYK-HOFN (Abb. 6.2a) zeigt für sich zunächst keine Zeit-Frequenzanteile, die geeignet wären, signifikante Veränderungen vor Beginn der Beben des Jahres 2000 zu belegen, denn der Datensatz beginnt erst 1998. Wesentliche Veränderung in den Periodizitäten sind vermutlich bereits zuvor erfolgt. Dementsprechend anders zeigt sich die zweite Analyse der Basislinie REYK-ONSA (Abb. 6.2b): In der Mitte des Jahres 1997 beginnt die Jahresperiode zu verschwimmen. Die Frequenz erhält höhere Anteile bis zu $1{,}7\,\mathrm{a}^{-1}$ mit einer maximalen Amplitude bei $1{,}2\,\mathrm{a}^{-1}$. Weiterhin kommt zu Beginn des Jahres 1998 eine Halbjahresperiode hinzu. Sie erreicht zu Beginn 1999 ihre maximale Amplitude. Gleichzeitig bricht die Amplitude der Jahresfrequenz ein. Es zeigt sich hiermit, dass der Bewegungsprozess jeweils Jahre vor und nach den Beben deutliche Veränderungen durchläuft.

6.1.2 Vergleich der Modellanpassungen

Das *zweidimensionale semiparametrische Modell* (5-31) ist an die Zeitreihe der Entfernungsänderung zwischen den permanenten GPS-Stationen Reykjavík (REYK, Südwestisland) und Onsala (ONSA, Südschweden) durch die Änderung der Zonengewichtsfunktion angepasst worden (Abb. 6.3a). Die hierzu verwendete Datenreihe ist zuvor um Ausreißer und Sprünge durch die Innovationsanalyse im adaptierten KALMAN-Filter (Kap. 5.2.1) bereinigt worden. Die Gewichtsfunktion besitzt keine zonenübergreifenden Korrelationsbedingungen, wodurch eine freie Anpassung zu recht extremen Gewichten führt. Die Gewichte in einigen Zonen neigen dazu, hohe negative Werte anzunehmen. Sie kompensieren häufig die übersteuerten Signale einer Nachbarzone mit ähnlichen Energieeinträgen, um so ein bestimmtes singuläres Signal in der Gesamtlösung hervorzuheben. Das ist nicht akzeptabel, daher muss dieses Modell dahingehend gedämpft werden, dass negative Gewichte *a priori* ausgeschlossen sind. Mit zunehmender Länge der Eingangs- und Ausgangszeitreihen wird dieser Effekt verschwinden und die Einschränkung obsolet.
Die Anpassung ist bereits weitgehend gelungen. Sie kann aufgrund der systemtheoretischen Vorüberlegungen allerdings nicht vollständig sein:

① Nicht alle Eingänge auf das System Island und auf das Sensorsystem GPS sind erfasst,
② die Aufgliederung in Zonen entspricht nicht der Systemortsfunktion,
③ die Festlegung der Impulsantwort in diesem Modell kann für alle Erdbeben falsch sein.

Als Systemeingang sind lediglich seismische Ereignisse automatisch realisiert (Kap. 5.1.2). Die Wetteraufzeichnungen konnten nur eingeschränkt und mittelbar berücksichtigt werden.

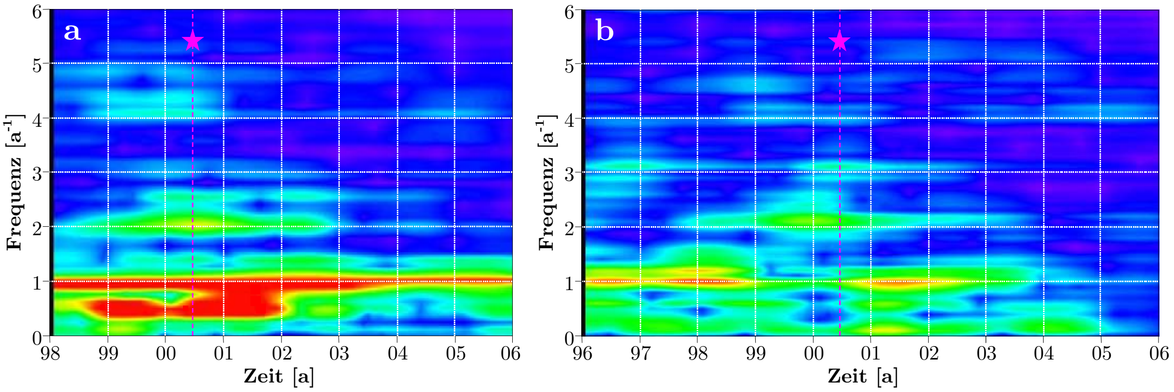


Abb. 6.2: Multiple Zeit-Frequenzanalyse über die Längenänderungen der beiden Basislinien a) REYK-HOFN und b) ONSA-REYK. Der Stern markiert jeweils den Zeitpunkt der Erdbeben des Jahres 2000 in der SISZ.

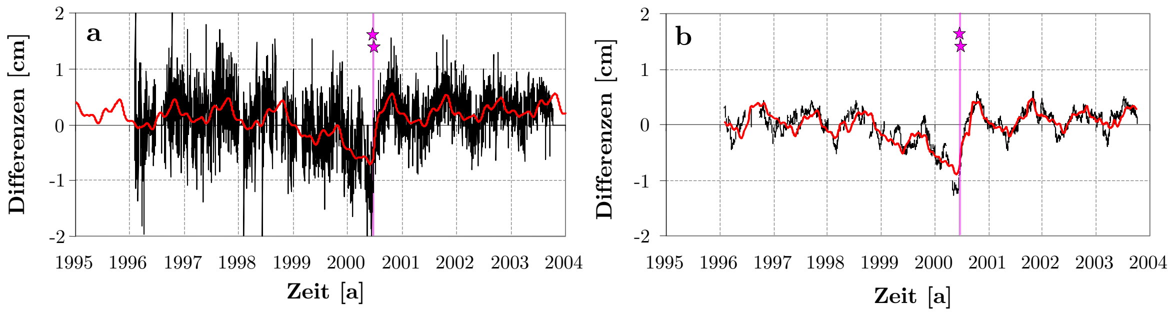


Abb. 6.3: a) Vergleich der Änderungen der Basislinie ONSA-REYK (schwarz) mit der trendbefreiten Lösung des 2D-Modells (rot). Die Basislinie ist um die Trendbewegung aus NUVEL-1A, um ihre Ausreißer und ihre Sprünge bereinigt. b) Vergleich der um Trendbewegung aus NUVEL-1A bereinigten KALMAN-gefilterten Zeitreihe (schwarz) mit der trendbefreiten Lösung des 3D-Modells (rot). Die Sterne markieren die Zeitpunkte der zwei M 6,5 Erdbeben im Juni des Jahres 2000 in der Südisländischen Seismischen Zone.

Zusätzliche Informationen wie die einseitige Sonneneinstrahlung auf die Stationen oder Hebungen und Senkungen mit dem Grundwasserspiegel oder andere unbekannte Einflüsse sind weder im Umfeld der Station realisiert noch in der Modellbildung berücksichtigt (Kap. 5.1.3). Die Zonenaufteilung erzwingt ihrerseits willkürlich die Gleichgewichtung von seismischen Ereignissen, die auf völlig unterschiedliche Prozesse zurückgehen können und damit naturgemäß einen anderen Einfluss auf das Bewegungsverhalten haben müssen. Genaugenommen handelt es sich um eine Mittelbildung über die Einflüsse innerhalb einer Zone.
Die größte Schwäche dieses Modells liegt aber in seiner erzwungenen Sprungantwort. Das *dreidimensionale semiparametrische Modell* (5-47) ist an eine KALMAN-gefilterte Zeitreihe der Entfernungsänderung zwischen den permanenten GPS-Stationen Reykjavík (REYK, Südwestisland) und Onsala (ONSA, Südschweden) durch die Änderung der Zonengewichtsfunktion angepasst worden (Abb. 6.3b). Die hierzu verwendete Datenreihe ist zuvor auch um das hochfrequente Rauschen im adaptierten KALMAN-Filter (Kap. 5.2.1) bereinigt worden.
Die Gewichtsfunktion zeichnet sich durch eine hohe räumliche Korrelation aus und lässt nur langwellige Änderungen zu. Eine Dämpfung ist daher nicht notwendig. Aber auch dieses Modell kann zwar eine gute, jedoch keine vollständige Anpassung erreichen. Die Gründe hierfür sind mit denen des ersten Modells vergleichbar:

① Nicht alle Eingänge auf das System Island und auf das Sensorsystem GPS sind erfasst,
② die langwelligen Radialbasisfunktionen entsprechen nicht der Systemortsfunktion,
③ die Festlegung der Impulsantwort in diesem Modell ist nur eine Mittelung über alle Impulsantworten.

Dennoch fällt bei beiden Modellen die Anpassung erstaunlich gut aus, was sich auch in der Vorwärts- und Rückwärtsprädiktion des 2D-Modells zeigt (Abb. 6.3a).

6.1.3 Inferenz der Modelle

Räumliche Inferenz

Eine erste Plausibilitätskontrolle über die Modellbildung erhält man durch die Georeferenzierung der räumlichen Gewichtsfunktion ((5-27) u. (5-46)). Diese Gewichtsfunktion

$$w_0 = (\mathfrak{x}, \vec{\chi}) \tag{6-1}$$

erhält hiermit folgende Bedeutung: ‚Wieviel Energie überträgt ein Erdbeben in welcher Zone $\vec{\chi}$ auf die Spreizungsbewegung Islands?‘

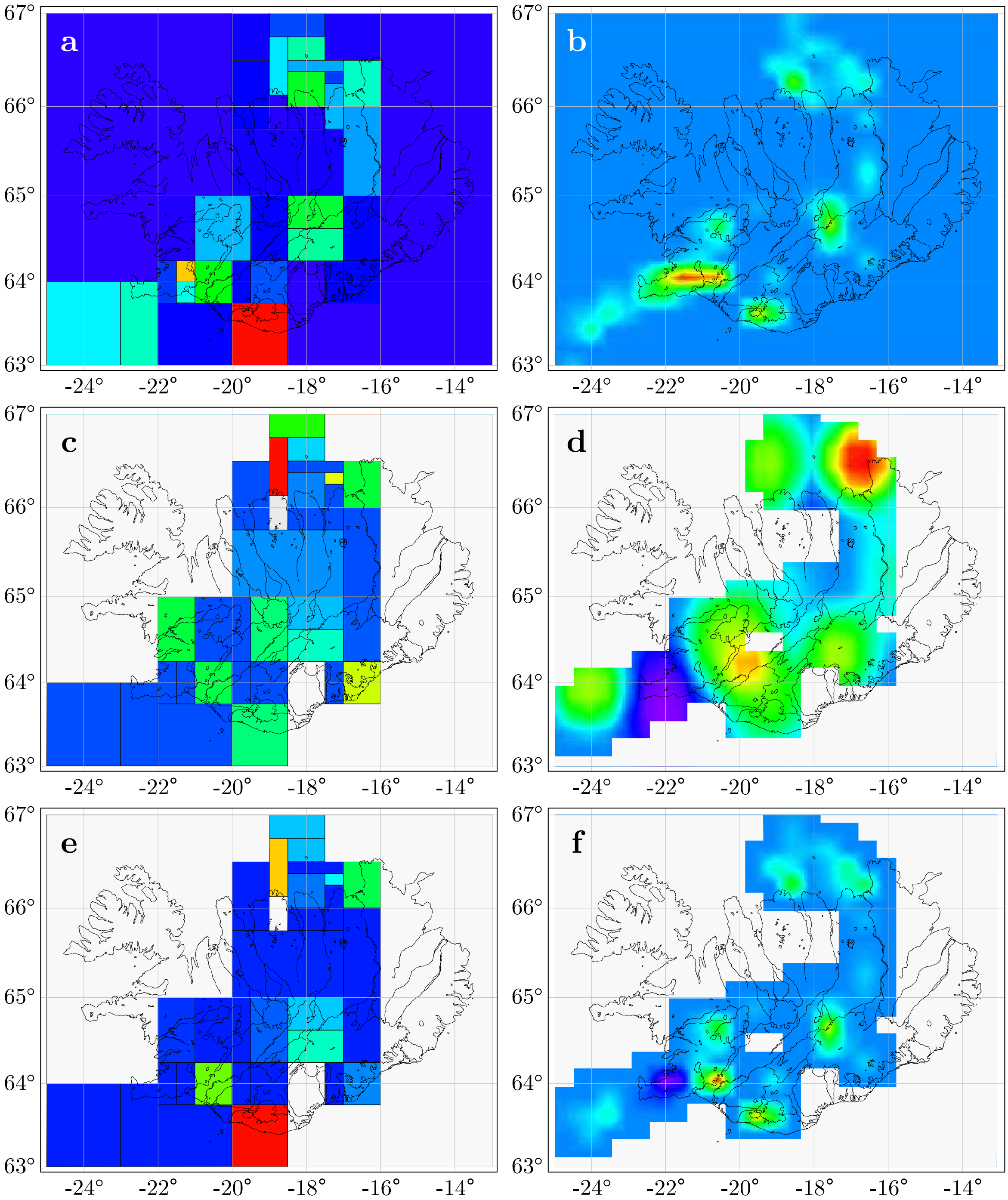


Abb. 6.4: Inferenz der semiparametrischen Modelle der Basislinie ONSA-REYK: normierte seismische Energie als Systemeingang (blau=0, rot=1) im 2D-Säulen-Modell (a) und im 3D-Knoten-Modell (b), Gewichte (violett=-0,3, rot=1,0) zur Übertragung von Energie in Bewegung im Gültigkeitsbereich des 2D-Modells (c) und des 3D-Modells (d), e-f) normierter Anteil der Energie an der Bewegung (violett=-0,3, rot=1,0).

Für die Modellbildung ist als Realisierung des Systemeingangs zunächst der Energieeintrag geeignet zu summieren gewesen. So ergeben sich einerseits Summen über die Zonen Islands im 2D-Modell (Abb. 6.4a) und andererseits Summen an den Knoten im 3D-Modell (Abb. 6.4b). Schon hier zeigt sich in der Darstellung, dass sich für den gleichen Systemeingang unterschiedliche Realisierung für das jeweilige Modell ergeben. Beiden Modellen ist jedoch zu eigen, dass sich die tektonisch aktiven Zonen wie der Reykjanes Rücken, die Südisländische Seismische Zone, die Westliche, Östliche und Nördliche Vulkanische Riftzone wie auch die Tjörnes Bruchzone deutlich abbilden. Allerdings erscheinen im Energiebild nicht die gesamten tektonischen Zonen, wie es sich zunächst aus der Darstellung der Epizentren erwarten ließe (Abb. 5.4). Finden sich dort noch alle Transformzonen deutlich vertreten, so stechen in der Abbildung der eingehenden Energie besonders die östlichen und westlichen Enden dieser Zonen hervor.

Zeitliche Inferenz

Es ist bei beiden Modellen möglich, jeweils die Modelldurchgangskanäle der Seismik und des Klimas voneinander zu trennen. So zeigt der Vergleich des klimatischen Signals mit den Änderungen der Basislinie ONSA-REYK erwartungsgemäß eine klare Saisonalität (Abb. 6.5a, blau). Aber es handelt sich keineswegs um ein Signal, das harmonisch oder in Phase der Klimabeobachtungen wäre. Dennoch spiegelt es sehr gut die quasisaisonalen[1] Schwankungen der trendbefreiten Änderungen der Basislinie ONSA-REYK wider. Der Vergleich des trend-

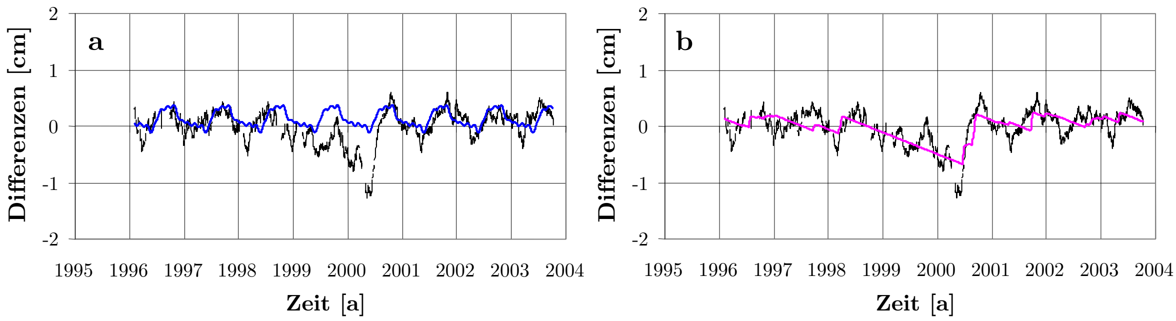


Abb. 6.5: Vergleich des a) klimatischen Signals und des b) seismischen Signals mit den trendbefreiten Änderungen der Basislinie ONSA-REYK.

befreiten Signals mit den trendbefreiten Änderungen der Basislinie ONSA-REYK zeigt ein quasisaisonales Signal mit schwankender Phasenlage und variierender Amplitude (Abb. 6.5b). Das wird insbesondere nach den Beben des Jahres 2000 deutlich. Einzig das Jahr 1999 ist völlig signalfrei. Die einzelnen Zeitsignale eines jeden aktiven Knotens können zunächst aus dem Modell separiert werden. Fasst man die Signale gemäß ihrer tektonischen Zonen zusammen (Abb. 6.6, links), so ergeben sich die Zeitfunktionen einer jeden Zone über den jeweiligen Beitrag an der Gesamtbewegungsfunktion. Damit erhält man eine gute Einsicht in die Arbeitsweise des Modells und die Qualität der Lösung. So fällt besonders auf, dass die Bewegungsbeiträge der Südisländischen Seismischen Zone (SISZ), dominiert von den Beben des Juni 2000, nahezu gänzlich von den Negativsignalen der Hengill *triple junction* (HENG) und des Reykjanesrückens (REYK) kompensiert werden (Abb. 6.6). Im Süden liefern darüber hinaus noch die Region um den Zentralvulkan Katla (KATL) und die Westliche Vulkanische Riftzone (WVRZ) sowie im Norden das Grimsey *lineament* und der Húsavík-Flatey-Graben einen signifikanten Beitrag. Die übrigen Zonen tragen nicht oder nur kaum zur Bewegung bei.

[1] Die Überlagerung mehrerer quasiperiodischer Prozesse (Kap. 3.2.1), die im Wesentlichen ein instabiles saisonales Signal bilden, seien hier als *quasisaisonal* bezeichnet.

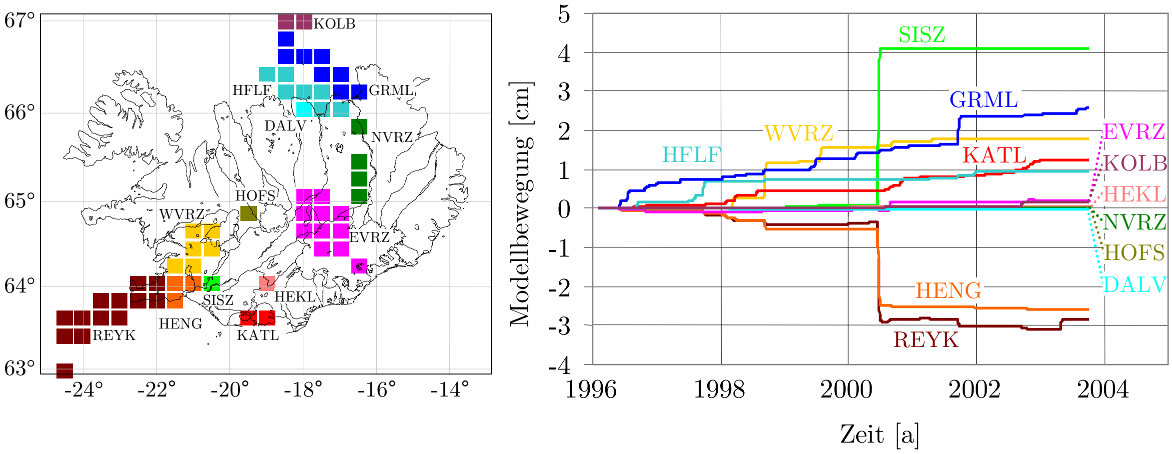


Abb. 6.6: Zeitliche Modellinferenz der tektonischen Zonen: Aktive Knoten des Modells sind zu Zonen zusammengefasst (links). Jede liefert einen eigenen Bewegungsbeitrag in das Modell (rechts).

6.2 Interpretation

6.2.1 Einordnung der Ergebnisse in die Fragestellung

Die Korrelationsanalyse zeigt eine signifikante, wenngleich schwache Korrelation zwischen den Wetterbeobachtungen und der seismischen Aktivität der SISZ. Daher liegt nahe, dass die eingangs getroffene Annahme, das Subsystem *Atmosphäre* wirke direkt und saisonal auf das Subsystem *Kruste*, korrekt ist (Abb. 1.4). Eine nichtkausale Korrelation erscheint hier wenig wahrscheinlich, denn ein weiteres bisher nicht direkt benanntes Subsystem *Wasser* vermittelt den Systemausgang des Subsystems *Atmosphäre* zu einem Systemeingang des Subsystems *Kruste*.

Wie bereits aus der Differenzbewegung der isländischen GPS-Permanentstationen, kündigen sich die Beben des Jahres 2000 auch im Frequenzraum an. Besonders deutlich wird dies in der multiplen Zeit-Frequenzanalyse der Basislinie REYK-HOFN. Ab 1998, dem Jahr nach der bisher noch ungeklärten seismischen Ruhe innerhalb der SISZ, prägt sich eine zweite dominante Periodizität mit der Umlaufzeit von zwei Jahren heraus (Abb. 6.2a). Diese bleibt das ganze Jahr 2001 erhalten und klingt erst danach ab. Der anomale Bewegungsprozess vor und nach den Beben ist damit abgeschlossen. In der langen Basislinie ONSA-REYK ist dieser Effekt zwar vorhanden, aber weitaus schwächer.

Die ursprüngliche Annahme, dass die Seismizität ein klares saisonales Signal in der Spreizungsbewegung in Konkurrenz zu den klimatischen Einflüssen verursachen werde, kann in erster Hinsicht nicht eindeutig verifiziert werden.

Die Wechselwirkungen zwischen Bewegungen und Seismizität stellen sich als komplexer heraus. Besonders überraschend ist der Ausfall des seismischen Signals in der Inferenz des 3D-Modells im Jahr 1999 (Abb. 6.5b). Diese Erscheinung will auf den ersten Blick nicht mit der stetig steigenden Periodizität in der SISZ ab 1998 korrespondieren (Abb. 1.1). Damit drängt sich die Frage auf, wohin die Energie der Erdbeben aus diesem Zeitraum fließt, wenn in der Energieübertragung gleichzeitig eine „Totzeit“ herrscht (Abb. 6.5b).

Zwei Aspekte sind geeignet, diese Diskrepanz zu erklären: Zum einen herrscht diese ansteigende Periodizität vor allem im Inneren der SISZ, ausgerechnet also in dem Gebiet, welches durch das 3D-Modell abgewichtet worden ist. Hier kann ein fehlender Systemeingang zum Tragen kommen, denn es liegen dem Modell keine Informationen zur Ausrichtung der Herdflächen vor. Diese sind aber hier orthogonal zur Richtung der eigentlichen Transformbewegung (Kap. 2.6). Der Großteil der Energie geht daher möglicherweise für die Riftbewegung verloren.

Zum anderen kann nicht ausgeschlossen werden, dass die seismischen Ruhe des Jahres 1997 der Beginn des Verhakens der SISZ gewesen ist. Anschließend haben sich jeweils im Sommer unter dem Einfluss der Gletscherschmelze die Spannungen innerhalb der SISZ abgebaut, ohne dass aber die eigentliche Verhakung gelöst worden ist.

6.2.2 Einordnung in die Tektonik Islands

Ereignisse vor den Beben des Juni 2000

Die Zeitreihen der seismischen Aktivität in der SISZ, die Differenzbewegungen der Permanentstationen (Abb. 1.1), aber eben auch die Kreuzkorrelationen zwischen dem Wetter und der Seismizität (Kap. 6.1.1) sowie schließlich die Zeit-Frequenzanalyse (Kap. 6.1.1) deuten auf

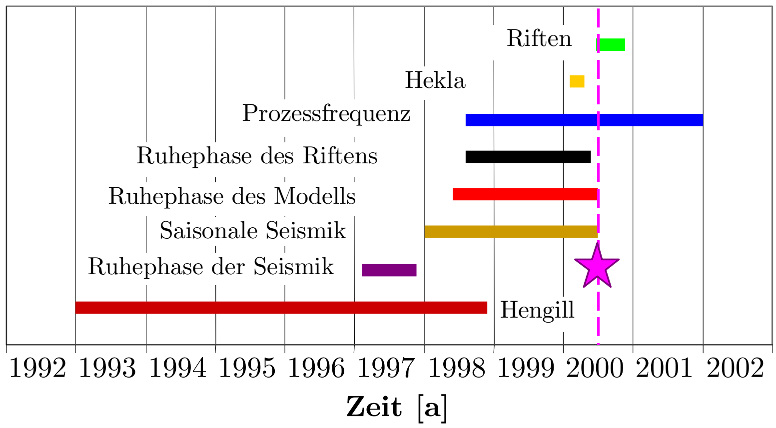


Abb. 6.7: Tektonische Ereignisse vor den Erdbeben des Jahres 2000: Schon Jahre vor den Erdbeben (Stern) verändert sich das tektonische System. Es kommt zu Veränderung der Seismizität der Südisländischen Seismischen Zone und zu veränderten Bewegungsprozessen.

überraschend kurzfristige Veränderungen vor den Beben des Jahres 2000 hin. Überraschend deshalb, weil für die freizusetzende Energie zunächst über lange Zeiträume eine gewisse Spannung aufgebaut werden muss. Diese scheint zwar die Ursache, nicht jedoch der kurzfristige Auslöser der Beben gewesen zu sein. Es ereignen sich nacheinander verschiedene Veränderungen im tektonischen System: als erstes tritt 1997 die relative seismische Ruhe in der SISZ ein, der eine Phase der zunehmenden Saisonalität der seismischen Aktivität folgt. Im Frühling des Jahres 1998 setzt über ganz Island die Übertragung der seismischen Energie in die kinetische Energie aus (Abb. 6.5b), worauf in der Mitte des Jahres sich die Spreizungsbewegung zu verlangsamen beginnt. Gleichzeitig beginnt sich eine neue Prozessfrequenz der Bewegung von $^1/_2 a^{-1}$ im Zeit-Frequenzspektrum zu zeigen (Abb. 6.2). Im Frühjahr des Jahres 2000 bricht die Hekla aus.
Es ist nicht sicher, ob ein einzelnes dieser Ereignisse nicht doch nur zufällig auftritt, aber die Summe und die Reihenfolge der Ereignisse ist doch auffällig.
Nur ist unklar, worin der Auslöser der Ereignisse zu suchen ist. Ein Ereignis, das ab 1993 die Seismologen beschäftigte, war die seismische Aktivität um den Zentralvulkan Hengill. Sie ging einher mit einer signifikanten Aufwölbung (Kap. 2.4.1). Es ist nicht auszuschließen, dass der Stress dieser Aufwölbung in die SISZ propagiert ist und hier die Spalte geschlossen hat, woraufhin hier die relative seismische Ruhe einsetzte. Immerhin haben sich in diesem Zeitraum auch die zwei Randspalten des Þingvellir in zwölf Jahren um 2 cm wieder geschlossen: anstelle sich um etwa 5 cm zu öffnen (Heinert et al. 2004, Abb. 6d).

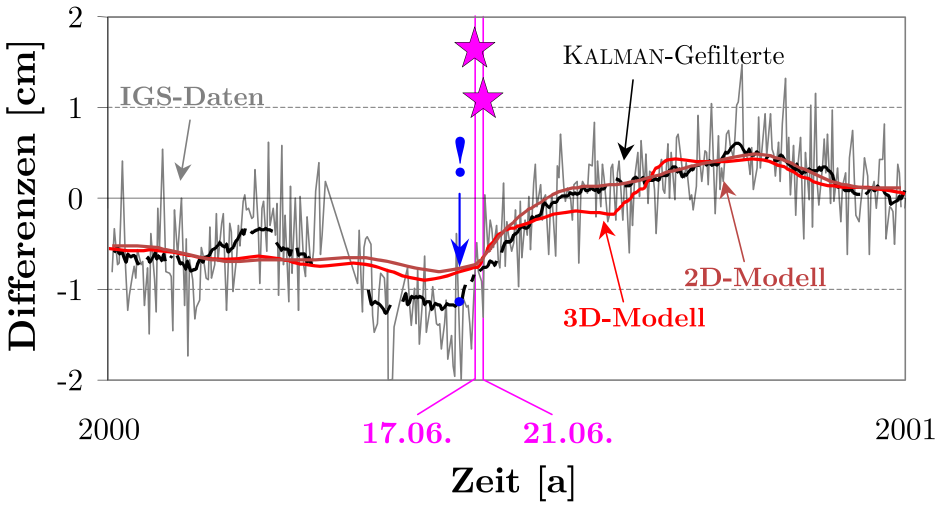


Abb. 6.8: Vergleich verschiedener trendbefreiter Bewegungszeitreihen: Beide Modelle (rot, braun) setzen mit ihrer Bewegung nach den Beben ein, in der KALMAN-gefiltere (schwarz) und der ungefilterte Bewegungszeitreihe (grau) bereits sechs Tage zuvor.

Kurzfristiges Vorzeichen zu den Beben des Juni 2000

Bei dem Vergleich der verschiedenen trendbefreiten Bewegungszeitreihen fällt im Zeitraum der Erdbeben im Juni des Jahres 2000 ein bemerkenswertes Phänomen auf: Gemäß des Modellaufbaus reagieren beide Modelle kausal nach den beiden Beben mit einer verstärkten Divergenzbewegung (Abb. 6.8). In der lediglich ausreißer- und sprungbereinigten Differenzzeitreihe aus GPS hingegen, setzt die Bewegung etwa sechs Tage früher ein. Gleichzeitig verändert sich das Rauschen dieser Zeitreihe. Deutlicher wird die vorzeitig einsetzende Bewegung anhand der KALMAN-gefilterten Zeitreihe, denn auch sie reagiert mit einer verstärkten Bewegung *vor* den beiden Beben. Die Qualität der GPS-Daten und deren Auflösung ist jedoch noch nicht geeignet, um die Existenz dieses Phänomen endgültig zu beweisen. Eine solche Qualität wird aber bald durch eine stetig verbesserte Empfangs- und Prozessiertechnik erreicht werden, so dass damit diese singuläre Beobachtung verifiziert werden könnte.

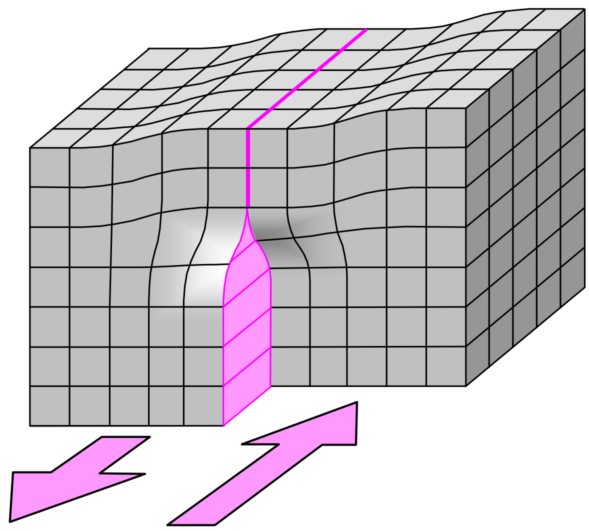


Abb. 6.9: Ein inverses *screw dislocation model*: Plastische Verformung im Untergrund ermöglichen eine Entlastung im weiteren Umfeld der zukünftigen Herdfläche. Die oberen Lithosphärenschichten sind noch rigide verkantet.

Verschiedene Vorzeichen von Erdbeben sind bereits diskutiert worden (Kap. 2.2.4). Für den Fall, dass sich diese Beobachtung verifizieren lässt, muss der vorzeitige Einsatz von lange Verzögerter oder gar ausgebliebener Bewegung ein weiteres Vorzeichen sein. Erklären lässt es sich mit einem inversen *screw dislocation model*. Während die oberen Lithosphärenschichten noch rigide verkantet sind, setzt im Untergrund bereits ein Bruch oder eine plastische Verformung ein. Deren *strain* propagiert nach außen und führt bereits zur Entlastung im weiteren Umfeld der zukünftigen Herdfläche. Die Bewegung wird messbar.

Dafür spricht die Feststellung, dass die Beben in den weitgehend noch ungestörten Schichten mit quartären Basalten stattgefunden haben (GUÐMUNDSSON 2007).
Für das Modell muss daher die in beiden Modellen als Grundlage angenommene temporäre Kausalität kritisch betrachtet werden (Kap. 5).

Einfluss der tektonischen Zonen auf die Bewegung

Eine andere ursprüngliche Annahme hat sich als falsch erwiesen. Die Gewichtung der Modelle hat nicht etwa tektonische Zonen wie die Südisländische Transformzone (SITZ) oder die Tjörnes Bruchzone (TFZ) als Ganzes gegenüber den vulkanischen Zentren hervorgehoben (Kap. 1.3), vielmehr ist die übertragene Energie im Osten und Westen der SITZ und TFZ besonders hoch. Gerade an den beiden Enden der SITZ finden sich mit dem Hengill und der Katla aktive Zentralvulkane. Beide haben in der jüngsten Vergangenheit eine starke seismische Aktivität gezeigt (Kap. 2.6).
Diese zunächst unerwartete Beobachtung in der Modellinferenz lässt sich jedoch wie folgt erklären: Die meiste Energie wird demnach dort in Bewegung übertragen, wo die Lithosphäre noch nicht durch die lange Zeit seismischer und tektonischer Aktivität so vorgeschwächt ist, so wie es in den zentralen Bereichen der Transformzonen bereits der Fall ist. Der Stress ist aus den Zentren bereits an die Enden propagiert, wo vielfach noch rigide Lithosphäre gekoppelt ist und elastisch reagiert. Dieses Ergebnis bestätigt bereits die Beobachtung aus den Bewegungen, dass die Südisländische Transformzone über die Südisländische Seismische Zone hinaus gefasst werden muss.

6.2.3 Einordnung in die Systemtheorie

Die Innovationsanalyse auf der Basis eines adaptiven KALMAN-Filters hat die Nutzung der GPS-Permanentstationsdaten ohne „manuellen“ Aufwand erst möglich gemacht (Kap. 5.2.1). Die implementierte Ausreißer- und Sprungkorrektur lief für diesen Datentyp zuverlässig und ohne notwendige Korrekturen.
Die recht einfach konzipierte Multiple Zeit-Frequenzanalyse ermöglicht je nach Berechnungsaufwand eine weit höhere Auflösung als jede Einzelanalyse. Zeitreihen sind hiermit deutlich einfacher hinsichtlich ihrer instationären Periodizitäten zu interpretieren[2] (Kap. 5.2.2).
Die in dieser Arbeit definierte und favorisierte semiparametrische Modellbildung hat die Überführung der inäquidistanten und logarithmischen Magnituden in äquidistante und lineare Bewegungsgrößen erst ermöglicht, und dies gleich in verschiedener Hinsicht:

- Die inäquidistanten Daten wurden durch eine Fuzzifizierung der Energien für einen rekursiven Modellansatz verwendbar,
- die üblicherweise verwendete Anzahl von Neuronen konnte auf ein einziges rekursives Neuron für die Zeitfunktion und 16 RBF-Neuronen für die Ortsfunktion reduziert werden,
- die jetzt sichere Einhaltung der Schranken der VC-Dimension eröffnet die Möglichkeit der Modellinferenz,
- die Modellinferenz ihrerseits gibt Auskunft über die Plausibilität des Modells und damit die Systemidentifikation, was mit der bisher üblichen Kreuzvalidierung nicht möglich ist.

[2] Die Deutung der HEISENBERGschen Unschärfe ist auf diese Analyse nicht tragfähig, da es sich im Vorliegenden Fall um keine unwiederholbare Echtzeitanwendung handelt. Auch bei Echtzeitanwendungen in der Nachrichtentechnik sollte eine Signalspaltung und damit eine multiple Analyse erwogen werden.

Diese Vorteile sind allerdings auch mit dem entscheidenden Nachteil erkauft, dass eine im Sinne des Modells unvollständige oder falsche funktionale Überführung der Daten eine Modellbildung erschwert. Allerdings kann dies durch die Inferenz sogleich aufgedeckt werden, so dass dieser aus einem ursprünglich *black box*-Algorithmus entwickelte Ansatz kontrollierbar wird.

Im vorliegenden Fall weist die Überführung der Magnituden in Energien die Schwierigkeit auf, dass die funktionale Definition der Energie eines Erdbebens nach (2-4) nicht zwingend für Island gültig ist, was sich auch in der zeitlichen Inferenz widerspiegelt (Abb. 6.6): Zwei benachbarte Zonen zur SISZ, die ab Juni 2000 auch eine erhöhte seismische Aktivität besitzen, kompensieren das zu hohe Signal der SISZ. Dass aber das Beben von 2000 einen viel zu hohen Einfluss gegenüber allen übrigen seismischen Ereignissen in der SISZ besitzt, deutet eine unzureichende Übertragung in Energie an. Die Lösung besteht darin, auch die Konstanten der Energiedefinition (2-4) in erster Näherung für ganz Island mitzuschätzen.
Ein hierzu vollständig alternativer Weg ist eine Semiparametrisierung über das seismische Moment.
In einem weiteren Aspekt sind beide Modelle noch nicht hinreichend: Zu dem Ansatz der propagierenden elastischen Entlastung nach einem Erdbeben gehört auch deren Laufzeit zwischen dem Zentrum des Bebens bis zu der GPS-Permanentstation. Damit müsste ein Modell auch eine raumabhängige zeitliche Gewichtsfunktion $a_\delta(\mathfrak{x})$ mit einschließen. Ganz analog dazu bedürfte es für längere Datensätze auch einer zeitabhängigen räumlichen Gewichtsfunktion $w_0(\mathfrak{x}, \chi, t)$, die der Tatsache gerecht wird, dass sich die Rheologie jedes tektonischen Elementes nach einem Erdbeben remanent verändert. Erst ein solches Modell wird einem dynamischen System (Kap. 3.1.5) vollständig gerecht werden.

7 Zusammenfassung und Ausblick

„'Oui' et 'Non' sont les mots
les plus courts et les plus faciles à prononcer
et ceux qui demandent le plus d'examen.“
CHARLES MAURICE DE TALLEYRAND-PÉRIGORD

Die vorliegende Arbeit beschäftigt sich mit der Frage, wie das Auseinanderdriften der Platten Nordamerikas und Eurasiens zeitlich verläuft und ob sich ein jährlicher Zyklus darin nachweisen lässt. Eine Vielzahl von Indizien deuten auf die Existenz dieses Phänomens hin. So war schon vor Beginn dieser Arbeit bekannt, dass die Trajektorien von permanenten Satellitenempfangsstationen des *Global Positioning System* dies- und jenseits des Mittelatlantischen Rückens jeweils im Winter und Frühling für einige Monate innerhalb der Messgenauigkeit vollständig identisch verlaufen. Diese Kongruenz der Trajektorien kann aber nicht uneingeschränkt als ein rein geometrisches und damit auch reales geodynamisches Phänomen gedeutet werden. Vielmehr kann sie auch als ein atmosphärischer Einfluss auf die Signalausbreitung gedeutet werden: Die Signale müssen die Troposphäre passieren, die den jährlichen klimatischen Schwankungen unterliegt.
Es kann in dieser Arbeit gezeigt werden, dass es signifikante räumliche und zeitliche Zusammenhänge zwischen den jährlichen klimatischen Schwankungen und dem Auftreten von seismischen Ereignissen gibt. Die beiden entwickelten Modelle liefern vergleichbar gute Vorhersagen des Auseinanderdriftens der Platten. Es ergibt sich ferner aus den Modellen die zeitliche und räumliche Verteilung der Energiebeiträge einer jeden tektonischen Zone, die tatsächlich in den Bewegungsprozess einfließen. Die gemeinsame Betrachtung der Modellinferenz mit den Datenanalysen ergibt, dass die beiden schweren Erdbeben des Jahres 2000 in der Südisländischen Seismischen Zone sich durch eine regelrechte Kaskade von beobachtbaren Ereignissen angekündigt haben.
Um aus Erdbebendaten zu einer Herleitung der Bewegung zu gelangen, ist ein neuer Modelltyp entwickelt worden, der hier als *semiparametrisch* bezeichnet werden soll. Dieser Modelltyp baut auf einem nichtparametrischen Modell in der Form eines künstlichen neuronalen Netzes auf. Semiparametrisch wird das Modell durch die deterministische Überführung der Erdbebendaten zunächst in Energieäquivalente und die anschließende Überführung dieser in ein Sollbewegungen. Erst diese Sollbewegungen werden dem neuronalen Netz als Eingang zur Verfügung gestellt.
Doch zuvor müssen die Erdbebendaten ortsabhängig gewichtet werden und formal in eine zeitliche Äquidistanz überführt werden, um der rekursiven Modellierung zu genügen. Die Inferenz eines konvergierten Modells entspricht dem Öffnen der *black box* eines neuronalen Netzes.
Als Grundlage zur Entwicklung dieser semiparametrischen Modelle, dienten sowohl Vorüberlegungen aus der *Statistischen Lerntheorie* für die zu wählende Netzwerkarchitektur, als auch die Betrachtungen zur Modelloptimierung aus der *Operations Research*. Verschiedene Netzwerkarchitekturen für die Gewichtung der Erdbeben in Abhängigkeit ihres Ortes mussten zuvor getestet werden.
Es war ferner nötig, einen geeigneten *robusten Schätzer* für die Optimierung abzuleiten. Die Einflussfunktion dieses Schätzers ist beschränkt und zurückfallend und basiert auf der Logarithmisierung der Fehlerquadrate.
Die gesamte Modellbildung basiert auf Ansätzen aus der *Systemtheorie*, deren ausführliche

Darstellung und Herleitung die mathematischen Grundlagen der Modellbildung lieferte. Durch die Nutzung der hier einzugliedernden *Prozessanalyse* ergaben sich im Zuge der Untersuchungen grundlegende Ergebnisse über die ablaufenden Prozesse. Mit diesen Erkenntnissen konnten die Modelle auf die wesentlich notwendigen Eigenschaften reduziert werden.

Für die Zeiten eines ungestörten Bewegungsprozesses der Platten konnte mithilfe der dargestellten Modelle eine Zyklizität der Bewegungen aus den Beobachtungen der Erdbeben hergeleitet werden. Diese ist aber instationär in Amplitude und Phasenlage. Diese zyklische Bewegung erlosch, als sich die Platten vor den Beben des Jahres 2000 verhakten.
Die Semiparametrisierung der Modelle hat bei ihrem ersten Einsatz zu Artefakten geführt, die einen Nachweis erschweren. Verantwortlich hierfür sind die als Konstanten angesetzten globalen Parameter aus der physikalisch-parametrischen Überführung der Gutenberg-Richter-Magnituden in Energie. Deshalb müssen diese physikalischen Effektivparameter bei einem neuerlichen Modellansatz speziell für Island mit angepasst werden.
Die Frage nach dem endgültigen Nachweis der Zyklizität der Plattenbewegungen bleibt daher zunächst noch offen. Diese wird sich aber lösen lassen, wenn einerseits weitere Daten zur Verfügung stehen werden und andererseits die Zuverlässigkeit der Modelle durch weitere Ergänzungen zunehmen wird.
Grundsätzlich darf man zuversichtlich sein, was die Weiterentwicklung von semiparametrischen Modellen im Zusammenspiel mit den Erkenntnissen der Statischen Lerntheorie betrifft. Letztere stellt bereits den numerischen Nachfolger für neuronale Netze in der Form der *Support Vector Machines* bereit. Aber auch durch die strikte Kontrolle der Netzarchitektur, wie sie von der *Erweiterten Lerntheorie* weiter vorangetrieben wird, ist eine Modellinferenz erst denkbar. Im Zusammenspiel mit der Semiparametrisierung wird sie interpretierbar. Hier stehen die praktischen Modellbildungen noch weitgehend am Anfang und es darf erwartet werden, dass diese Entwicklungen noch vielseitige Möglichkeiten bereithalten.
Diese Arbeit soll einen Beginn darstellen, diese Form der Modellbildung für den Geodäten nutzbar und nützlich zu machen. Sowohl von der Seite der Modellbildung als von der Seite des Untersuchungsobjektes sind die aufgezeigten Möglichkeiten bei weitem nicht ausgeschöpft.

Literaturverzeichnis

ALEX, N., EINARSSON, P., HEINERT, M., NIEMEIER, W., RITTER, B., SIGMUNDSSON, F. & WILLGALIS, ST. (1999): GPS-Kampagne 1995 zur Bestimmung von Deformationen der Erdkruste in Südwestisland. *Z. f. Verm.wesen* **124**: 377–388. ISSN 0340-4560.

ALLEN, R. M., NOLET, G., MORGAN, W. J., VOGFJÖRD, K., BERGSSON, B. H. ERLENDSSON, P., FOULGER, G. R., JAKOBSDÓTTIR, S., JULIAN, B., PRITCHARD, M., RAGNARSSON, ST. & STEFÁNSSON, R. (2002): Plume driven plumbing and crustal formation in Iceland. *J. Geophys. Res.* **107**, 2163, DOI 10.1029/2001JB000584.

ALLEN, R. M., NOLET, G., MORGAN, W. J., VOGFJÖRD, K., BERGSSON, B. H. ERLENDSSON, P., FOULGER, G. R., JAKOBSDÓTTIR, S., JULIAN, B., PRITCHARD, M., RAGNARSSON, ST. & STEFÁNSSON, R. (2002): Imaging the mantle beneath Iceland using integrated seismological techniques. *J. Geophys. Res.* **107**, 2325, DOI 10.1029/2001JB000595.

ALTAMIMI, Z., SILLARD, P. & BOUCHER, C. (2002): ITRF 2000. A new release of the international terrestrial reference frame for Earth science applications. *J. Geophys. Res.* **107**: 2214. DOI 10.1029/2001JB000561.

ANRETTER, M. (2001): *Moving hotspots – Evidence from paleomagnetism and modeling.* Dissertation. Ludwig-Maximilians-Universität München.
`http://deposit.ddb.de/cgi-bin/dokserv?idn=963185136` ⋆.

ÁRNADÓTTIR, Þ., HREINSDÓTTIR, S., GUÐMUNDSSON, G., EINARSSON, P., HEINERT, M. & VÖLKSEN, CHR. (2001): Crustal deformation measured by GPS in the South Icelandic Seismic Zone due to two large earthquakes in June 2000. *Geophys. Res. Let.* **28** (21): 4031–4033. DOI 10.1029/2001GL013332.

BACON, F. (1620): *Novum Organum, sive Indicia vera de Interpretatione Naturae.* KROHN, W. (Hrsg., 1990): *Neues Organon : lateinisch - deutsch / Francis Bacon.* Bd.1 u. Bd.2., Meiner-Verlag, Hamburg. ISBN 3-7873-0757-9.

BANDOPADHYAY, P. C. & SENGUPTA, S. (2005): The Paleoproterozoic Supracrustal Kolhan Group in Singhbhum Craton, India and the Indo-African Supercontinent. *Gondwana Research* **7** (4): 1228–1235. DOI 10.1016/S1342-937X(05)71097-2.

BARTSCH, H.-J. (1999): *Taschenbuch Mathematischer Formeln.* 18., verb. Aufl. Fachbuchverlag Leipzig im Carl-Hanser-Verlag, München-Wien. ISBN 3-446-21048-2.

BAUER, M. (2003): *Vermessung und Ortung mit Satelliten GPS und andere satellitegestützte Navigationssysteme.* 5., neu bearb. u. erw. Aufl., Wichmann Heidelberg. ISBN 3-87907-360-0.

BAZARAA, M. S. & SHETTY, C. M. (1979): *Nonlinear Programming Theory and Algorithms.* John Wiley & Sons, New York-Chichester-Brisbane-Toronto-Singapore. ISBN 0-471-78610-1.

BENIOFF, H. (1954): Orogenesis and deep crustal structure – additional evidence from seismology. *Bull. Geol. Soc. Am.* **65**: 385–400.

BEUTLER, G., ROTHACHER, M., SCHAER, S., SPRINGER, T. A., KOUBA J. & NEILAN R. E. (1999): The international GPS service (IGS): an interdisciplinary service in support of earth sciences. In: *Advances in Space Research* **23** (4): 631–653. DOI 10.1016/S0273-1177(99)00160-X.

BJÖRNSSON, A., JOHNSEN, G., SIGURDSSON, S., ÞORBERGSSON, G., TRYGGVASON, E. (1979): Rifting of the plate boundary in North Iceland 1975-1978. *J. Geophys. Res.* **84** (B6): 3029–3038. DOI 10.1029/JB084iB06p03029.

⋆ Zuletzt aufgesucht zur Drucklegung

Blakey, R. C. (2007): *Paleogeographic globes.* `http://jan.ucc.nau.edu/~rcb7/globaltext.html`★

Borutta, H. (1988): *Robuste Schätzverfahren für geodätische Anwendungen.* Dissertation. Schriftenreihe Studiengang Vermessungswesen **33**. Univ. d. Bundeswehr Neubiberg. ISSN 0173-1009.

Bothe, H.-H. (1995): *Fuzzy Logic – Einführung in Theorie und Anwendung.* 2. erw. Aufl., Springer-Verlag, Berlin-Heidelberg-New York-London-Paris-Tokyo-Hong Kong-Barcelona-Budapest. ISBN 3-540-56967-7.

Bothe, H.-H. (1998): *Neuro-Fuzzy-Methoden – Einführung in Theorie und Anwendung.* Springer-Verlag, Berlin-Heidelberg-New York-Barcelona-Budapest-Hong Kong-London-Mailand-Paris-Santa Clara-Singapur-Tokyo. ISBN 3-540-57966-4.

Bott, M. H. P. (1974): *Deep structure and origin of the Icelandic transverse ridge.* In: Kristjansson, L. (Ed.): *Geodynamics of Iceland and the North Atlantic area.* Proceedings of the NATO Advanced Study Institute: 33–47. Reidel Publishing Company, Dordrecht-Boston/USA. ISBN 90-2077-5050-4.

Branlund, J., Regenauer-Lieb, K. & Yuen, D. A. (2000): Fast ductile failure of passive margins from sediment loading. *Geophys. Res. Lett.* **25** (13): 1989–1992. DOI 10.1029/1999GL008396.

Braun, A. & Marquart, G. (2001): Die bewegte Geschichte des Nordatlantiks. *Spektrum der Wissenschaft* 06/2001: 50–59. ISSN 0170-2971.

Bunke, J. (1997): *Künstliche Neuronale Netze zur Systemidentifikation aus gestörten Meßwerten.* Fortschrittsber. VDI Reihe 8, **667**, VDI-Verlag Düsseldorf. ISBN 3-18-366708-8.

Butz, T. (2000): *Fouriertransformation für Fußgänger.* 2. durchgesehene Aufl., Teubner, Stuttgart-Leipzig. ISBN 3-519-10202-1.

Cara, M. (1994): *Geophysik.* Springer Verlag, Berlin-Heidelberg-New York. ISBN 3-540-55859-4.

Collatz, L. & Wetterling, W. (1966): *Optimierungsaufgaben.* Springer Verlag, Berlin-Heidelberg-New York. LCCN✠ 66022473.

Czubik, E. (1989): Beobachtung rezenter Krustenbewegungen: 650 km Feinnivellements in Island. *Z. f. Verm.wesen* **114**: 25–34. ISSN 0340-4560.

Dach, R. (2000): *Einfluß von Auflasteffekten auf präzise GPS-Messungen.* Dissertation. DGK Reihe C **519**, München. ISBN 3-7696-9558-8.

Dalziel, I. W. D. (1991): Pacific margins of Laurentia and East Antarctica-Australia as a conjugate rift pair: Evidence and implications for an Eocambrian supercontinent. *Geology* **19** (6): 598–601. DOI 10.1130/0091-7613(1991)019<0598:PMOLAE>2.3.CO;2.

Dalziel, I. W. D. (1995): *Die Erde vor Pangäa.* In Giese, P. (Hrsg., 1995): *Geodynamik und Plattentektonik.* Spektrum Akad. Verlag, Heidelberg, ISBN 3-86025-373-5.

Dalziel, I. W. D. (1997): Neoproterozoic-Paleozoic Geography and Tectonics: Review, Hypothesis, Environmental Speculation. *Geol. Soc. Am. Bull.* **109** (1): 16–42. DOI 10.1130/0016-7606(1997)109<0016:ONPGAT>2.3.CO;2.

DeMets, C., Gordon, R. G., Argus, D. F. & Stein, S. (1990): Current plate motions. *Geophys. J. Int.* **101**: 425–478. DOI 10.1111/j.1365-246X.1990.tb06579.x.

DeMets, C., Gordon, R. G., Argus, D. F. & Stein, S. (1994): Effect of Recent Revisions to the Geomagnetic Reversal Time Scale on Estimates of Current Plate Motions. *Geophys. Res. Let.* **21**: 2191–2194. DOI 10.1029/94GL02118.

Deuflhard, P. & Hohmann, A. (1991): *Numerische Mathematik Eine algorithmisch orientierte Einführung.* Walter de Gruyter, Berlin-New York. ISBN 3-11-012918-3.

Dick, H. J. B., Lin, J. & Schouten, H. (2003): An ultraslow-spreading class of ocean ridge. *Nature* **426**: 405–412. DOI 10.1038/nature02128.

✠ *Library of Congress Control Number*: Der gesuchte Titel kann unter `http://lccn.loc.gov/nnnnnnnn` aufgesucht werden. Diese Zuordnung ist geeignet für Bücher ohne *International Standard Book Number.*

DIETRICH, R. (1988): *Untersuchung zur Nutzung künstlicher Erdsatelliten für die geodätische Koordinatenbestimmung.* Dissertation B. Veröffentl. d. Zentralinst. f. Physik d. Erde. Potsdam.

DEUTSCHES INSTITUT FÜR NORMUNG E. V. (02/1994): *DIN 19226-1: Leittechnik, Regelungstechnik und Steuerungstechnik, Allgemeine Grundbegriffe.* Beuth Verlag, Berlin-Wien-Zürich.

DEUTSCHES INSTITUT FÜR NORMUNG E. V. (Entwurf-2005)[ħ]: *DIN 18709-5: Begriffe, Kurzzeichen und Formelzeichen in der Geodäsie, Auswertung kontinuierlicher Messreihen.* Beuth Verlag, Berlin-Wien-Zürich.

DOMBOIS, F. (1998): *Über Erdbeben. Ein Versuch zur Erweiterung seismologischer Darstellungsweisen.* Dissertation, Humboldt-Universität zu Berlin.
`http://edoc.hu-berlin.de/dissertationen/geologie/dombois-florian/PDF/...`
`...Dombois.pdf`★

DOMSCHKE, W. & DREXL, A. (2002): *Einführung in Operations Research.* 5. überarb. u. erw. Aufl., Springer Berlin-Heidelberg. ISBN 3-540-42950-6.

DÖRNER, D. (1989): *Die Logik des Mißlingens - Strategisches Denken in komplexen Situationen.* Rowohlt Taschenbuch Verlag, Hamburg. ISBN 3-499-19314-0.

DOW, J. M., NEILAN, R. E. & GENDT, G. (2005): The International GPS Service (IGS): Celebrating the 10th Anniversary and Looking to the Next Decade. *Adv. Space Res.* **36**(3): 320–326. DOI 10.1016/j.asr.2005.05.125.

DREWES, H., & ANGERMANN, D. (2001): *The Actual Plate Kinematic and Crustal Deformation Model 2000 (APKIM2000) as a geodetic reference system.* IAG 2001 Scientific Assembly, Budapest, 2-8 Sept. 2001.
`http://www.dgfi.badw.de/fileadmin/DOC/2001/DS_APKIM.pdf`★

DREWES, H. (2006): Zum Wandel in der Zielsetzung geodätischer Forschung. *Z. f. Verm.wesen* **131**: 292–305. ISSN 1618-8950.

EINARSSON, P. & SÆMUNDSSON, K. (1987): Earthquake epicenters 1982 – 1985 and volcanic systems in Iceland (map). In Í Hlutarsins Eðli: Festschrift for Þorbjörn Sigurgeirsson, edited by T. Sigfússon, Menningarsjóður, Reykjavik.

EINARSSON, P. (1991): Earthquakes and present-day tectonism in Iceland. *Tectonophysics* **189**: 261–279. DOI 10.1016/0040-1951(91)90501-I.

EINARSSON, P., BRANDSDÓTTIR, B., GUÐMUNDSSON, M. T., BJÖRNSSON, H., GRÖNVOLD, K. & SIGMUNDSSON, F. (1997): Center of the Iceland Hotspot – Experiences Volcanic Unrest. *Eos* **78** (35): 369, 374–375. DOI 10.1029/97EO00237.

EINARSSON, P. & BRANDSDÓTTIR, B. (2000): Earthquakes in the Myrdalsjökull area, Iceland, 1978-1985: Seasonal correlation and connection with volcanoes. *Jökull The Icelandic Journal of Earth Science* **49**: 59–73. ISSN 0449-0576.

EINARSSON, P., THEODÓRSSON, P., HJARTARDÓTTIR, Á. R. & GUÐJÓNSSON, G. I. (2008): Radon Changes Associated with the Earthquake Sequence in June 2000 in the South Iceland Seismic Zone. *Pure appl. geophys.* **165**: 63–74. DOI 10.1007/s00024-007-0292-6.

EISENBACHER, G. H. (1991): *Einführung in die Tektonik.* Stuttgart. ISBN 3-432-99251-3.

ELISSEEFF, A. & PAUGAM-MOISY, H. (1997): Size of multilayer networks for exact learning: analytic approach. *NeuroCOLT Techn. Rep. Series* NC-TR-97-002.

ERLINGSSON, S. & EINARSSON, P. (1985): *Fjarlægðarmælingar á jarðskálftasvæði Suðurlands árið 1984.* Raunvísindastofnun Háskólans, RH-04-85.

EVANS, D. A. D., MARTIN, D. MCB., NELSON, D. R., POWELL, C. MCA. & WINGATE, M. T. D. (2000): *The Vaalbara hypotheses reviewed.* International Geological Congress 2000, Brazil.
`http://physics.curtin.edu.au/staff/docs/nelsond/092. Evans et al 2000.pdf`★

EYSTEINSSON, H. & GUNNARSSON, K. (1995): *Maps of Gravity, Bathymetry and Magnetics for Iceland and Surroundings.* Orkustofnun report, OS-95055/JHD-07. ISBN 9979-827-63-7.

[ħ] Entstanden unter der Mitwirkung des Autors.

Feigl, K. L., Gaspari, J., Sigmundsson, F. & Rigo, A. (2000): Crustal deformation near Hengill volcano, Iceland 1993-1998: Coupling between magmatic activity and faulting inferred from elastic modelling of satellite radar interferograms. *J. Geophys. Res.* **105**: 25655–25670. DOI 10.1029/2000JB900209.

Foulger, G. R., Beutler, G., Bilham, R., Einarsson, P., Fankhauser, S., Gurtner, W., Hugentobler, U., Morgan, W. J., Rothacher, M., Þorbergsson, G. & Wild, U. (1993): The Iceland 1986 GPS geodetic survey: Tectonic goals and data processing results. *Bull. Géod.* **67**: 148–172. DOI 10.1007/BF00806254.

Foulger, G. R., Pritchard, M. J., Julian, B. R., Evans, J. R., Allen, R. M., Nolet, G., Morgan, W. J., Bergsson, B. H., Erlendsson, P., Jakobsdóttir, S., Ragnarsson, S., Stefánsson, R. & Vogfjörd, K. (2000): The seismic anomaly beneath Iceland extends down to the mantle transition zone and no deeper. *Geophys. J. Int.* **142**: F1–F5. DOI 10.1046/j.1365-246x.2000.00245.x.

Foulger, G. R., Pearson, D. G. (2001): Is Iceland underlain by a plume in the lower mantle? Seismology and helium isotopes. *Geophys. J. Int.* **145**: F1–F5. DOI 10.1046/j.0956-540x.2001.01457.x.

Foulger, G. R., Pritchard, M. J., Julian, B. R., Evans, J. R., Allen, R. M., Nolet, G., Morgan, W. J., Bergsson, B. H., Erlendsson, P., Jakobsdóttir, S., Ragnarsson, S., Stefánsson, R. & Vogfjörd, K. (2001): Seismic tomography shows that upwelling beneath Iceland is confined to the upper mantle. *Geophys. J. Int.* **146**: 504–530. DOI 10.1046/j.0956-540x.2001.01470.x.

Foulger, G. R., Natland, J. H. & Anderson, D. L. (2005): *Genesis of the Iceland melt anomaly by plate tectonic processes.* In Foulger, G. R., Natland, J. H., Presnell, D. C. & Anderson, D. L.: *Plates, Plumes and Paradigms.* Boulder, CO, Geol. Soc. Ame. Special Paper **388**: 347–626. DOI 10.1130/0-8137-2388-4.595.

Foulger, G. R. & Anderson, D. L. (2005): A cool model for the Iceland hot spot. *J. Volc. Geoth. Res.* **141**: 1–22. DOI 10.1016/j.jvolgeores.2004.10.007.

Foulger, G. R. (2006): Older Crust underlies Iceland. *Geophys. J. Int.* **165**: 672–676. DOI 10.1111/j.1365-246X.2006.02941.x.

Freund, F. T. (2007): Pre-earthquake signals - Part I: Deviatoric stresses turn rocks into a source of electric currents. *Nat. Hazards Earth Syst. Sci.* **7** (5): 535–541.

Freund, F. T. (2007): Pre-earthquake signals - Part II: Flow of battery currents in the crust. *Nat. Hazards Earth Syst. Sci.* **7** (5): 543–548.

Gauss, C. F. (1809): *Theoria motvs corporvm coelestivm in Sectionibvs conicis solem ambientivm.* Hamburgi svmtibvs Frid. Perthes et I. H. Besser. Univ. Bibl. Techn. Univ. Braunschweig, Sig. 3000-5282.

Gendt, G. & Nischan, T. (2008): *2003/2004 Analysis Coordinator Report.* In: IGS Central Bureau *2003-2004 Technical Reports.* Pasadena, CA: Jet Propulsion Laboratory, 2008.

Gerke, K. (1966): Deutsche Geo-Forschungsarbeiten auf Island 1964/67. *Polarforschung* **36** (1/2): 99–102.

Grafarend, E., Heister, H., Kelm, R., Kropff, H. & Schaffrin, B. (1979): *Optimierung geodätischer Meßoperationen.* Sammlung Wichmann Neue Folge, Bd. 11, Karlsruhe. ISBN 3-87907-052-0.

Grapenthin, R., Sigmundsson, F., Geirsson, H., Árnadóttir, Þ. & Pinel, V. (2006): Iceland rhythmics: Annual modulation of land elevation and plate spreading by snow load. *Geophys. Res. Lett.* **33**: L24305. DOI 10.1029/2006GL028081.

Green II, H. W. (2001): *Der Mechanismus von Tiefenbeben.* In : *Die unruhige Erde.* Dossier 2/2002: 14–21. Spektrum Akad. Verlag, Heidelberg.

Green II, H. W. (2005): *New Light on Deep Earthquakes.* In : *Our Everchanging Earth.* Special edition: 97–105. Scientific American.

GRIPP, A. E. & GORDON, R. G. (1990): Current Plate Velocities Relative to the Hotspots Incorporating the NUVEL-1 Global Plate Motion Model. *Geophys. Res. Let.* **17**: 1109–1112. DOI 10.1029/GL017i008p01109.

GRIPP, A. E. & GORDON, R. G. (2002): Young tracks of hotspots and current plate velocities. *Geophys. J. Int.* **150**: 321–361. DOI 10.1046/j.1365-246X.2002.01627.x.

GRUNDMANN, W. (2002): *Operations Research Formeln und Methoden.* Teubner, Stuttgart-Leipzig-Wiesbaden. ISBN 3-519-00421-6.

GRÜNTHAL, G. (Hrsg. 1998): *European Macroseismic Scale 1998.* CONSEIL DE L'EUROPE Cahiers du Centre Européen de Géodynamique et de Séismologie **15**, Luxembourg. ISBN 2-87977-008-4.

GUÐMUNDSSON, A. (2004): persönliche Mitteilung.

GUÐMUNDSSON, A. (2007): *Infrastructure and evolution of ocean-ridge discontinuities in Iceland.* In: JACOBY, W. R. & GUÐMUNDSSON, A. (eds.): *Hotspot Iceland.* J. Geodyn. **43** (1): 6–29. DOI 10.1016/j.jog.2006.09.002.

GUTENBERG, B. & RICHTER, C. F. (1949): *Seismicity of the Earth and Associated Phenomena.* Princeton University Press. LCCN 49-11473.

GUTENBERG, B. & RICHTER, C. F. (1956): Earthquake Magnitude, Intensity, Energy, and Acceleration (Second Paper). *Bull. Seismol. Soc. Am.* **46** (2): 105–145. ISSN 0037-1106.

HACKMAN, M. C. (1991): *A study of crustal deformation in Iceland using boundary element modeling and the Global Positioning System.* Ph. D. thesis. University of Colorado, Boulder.

HAINZL, S., ZÖLLER, G. & MAIN, I. (2006): Introduction to special issue: Dynamics of seismicity patterns and earthquake triggering. *Tectonophysics* **424**: 135–138. DOI 10.1016/j.tecto.2006.03.034.

HANKS, T. C. & KANAMORI H. (1979): A moment magnitude scale. *J. Geophys. Res.* **84**: 2348–2350. DOI 10.1029/JB084iB05p02348.

HARTMANN, O., JACOBY, W. R., WOLF, D., KLEMANN, V. & SASGEN, I. (2007): *Interpretation glazial-isostatischer Ausgleichsvorgänge im Südosten Islands unter Berücksichtigung des Island-Plumes.* GeoForschungsZentrum Potsdam, Scient. Techn. Rep. **07** (07). ISSN 1610-0956.

HAYKIN, S. (1996): *Adaptive Filter Theory.* 3rd ed., Prentice Hall, Upper Saddle River NJ, ISBN 0-13-322760-X.

HAYKIN, S. (1999): *Neural Networks – A Comprehensive Foundation.* 2nd ed., Prentice Hall, Upper Saddle River NJ, ISBN 0-13-908385-5.

HAYKIN, S. (2002): *Adaptive Filter Theory.* 4th ed., Prentice Hall, Upper Saddle River NJ, ISBN 0-13-048434-2.

HEINE, K. (1999): *Beschreibung von Deformationsprozessen durch Volterra- und Fuzzy-Modelle sowie Neuronale Netze.* Dissertation. Deutsche Geodätische Kommission, Reihe C **516**, München. ISBN 3-7696-9554-2.

HEINERT, M. (1998): *Untersuchung des Gezeiteneinflusses auf die Oberflächenbewegung des Ekströmisens, abgeleitet aus kontinuierlichen GPS-Messungen.* Diplomarbeit. Techn. Univ. Braunschweig, unveröffentlicht.

HEINERT, M. & PERLT, J. (2002): *Relationship of Seismic Events and Divergent Plate Motion in Iceland.* Abstract. XXVII. General Assembly of the European Geophysical Society (EGS), Nice (France), 21.-26. April 2002. Abstr. EGS02-A-03771.

HEINERT, M. & REISER, ST. (2002): *Continuous building monitoring using adaptive Kalman-Filtering for real-time data screening and pre-processing.* In KAHMEN, H., NIEMEIER, W. & RETSCHER, G. (Hrsg.): *Second Symposium on Geodesy for Geotechnical and Structural Engineering II.* Proceedings, Berlin 21.-24. Mai 2002: 82–91. ISBN 3-9501492-1-X.

HEINERT, M. (2003): *Periodical Behaviour of seismic events and divergent plate motion in Iceland.* Abstract. EGS-AGU-EUG Joint Assembly, Nice (France), 06.-11. April 2003. Abstr. EAE03-A-10623.

Heinert, M. & Riedel, B. (2003): *A parametric model of ice-ocean interaction in the grounding zone derived from extremely short time series.* Abstract. 7th International Symposium on Antarctic Glaciology, University of Milano-Bicocca, Italy.

Heinert, M., Ritter, B. & Niemeier, W. (2004): Angepasste Verfahren zur Deformationsanalayse für die geodätischen Messungen in Südwestisland. *Z. f. Verm.wesen* **129**: 399–406. ISSN 1618-8950.

Heinert, M. & Niemeier, W. (2004): *Zeitreihenanalyse bei der Überwachung von Bauwerken.* In Schwarz, W. (Hrsg., 2004): *DVW-Fortbildungsseminar Interdisziplinäre Messaufgaben im Bauwesen - Weimar 2004.* DVW-Schriftenreihe **46**: 157–174. ISBN 3-89639-451-4.

Heinert, M. & Riedel, B. (2007): A parametric model of the geometrical ice-ocean interaction in the grounding zone derived from short time series. *Geophys. J. Int.* **169**: 407–420. DOI 10.1111/j.1365-246X.2007.03364.x.

Heinert, M. & Niemeier, W. (2007): From fully automated observations to a neural network model inference: The Bridge "Fallersleben Gate" in Brunswick, Germany 1999 - 2006. *J. Appl. Geodesy* **1**: 71–80. DOI 10.1515/JAG.2007.010.

Heinert, M. (2008): *Artificial neural networks – how to open the black boxes?* In Reiterer, A. & Egly, U. (Hrsg.): *Application of Artificial Intelligence in Engineering Geodesy.* IAG-Proceedings, Wien 01. Dez. 2008: 42–62. ISBN 3-9501492-4-1.

Heller, D.-A. & Marquart, G. (2002): An admittance study of the Reykjanes Ridge and elevated plateaux between the Charlie-Gibbs and the Senja fracture zones. *Geophys. J. Int.* **148**: 65–76. DOI 10.1046/j.1365-246x.2002.01565.x.

Hess, H. H. (1962): *History of Ocean Basins.* In Engel, A. E. J., James H. L. & Leonard, B. F. (Hrsg): *Petrologic studies: a volume in honor of A. F. Buddington.* Geol. Soc. Am.: 599–820.

Heumann, F. W. (1972): *Untersuchungen im geodätischen Sondernetz in Nordost-Island zu Messungen von 1938 und 1965.* Dissertation. Fakultät Bauwesen, Techn. Univ. Braunschweig.

Heunecke, O. (1995): *Zur Identifikation und Verifikation von Deformationsprozessen mittels adaptiver KALMAN-Filterung (Hannoversches Filter).* Dissertation. Wiss. Arb. d. Fachr. Verm.wesen d. Universtät Hannover **208**. ISSN 0174-1454.

Hill, D. P., Pollitz, F. & Newhall, Chr. (2002): Earthquake-Volcano Interactions. *Physics Today.* November 2002: 41–47. DOI 10.1063/1.1535006.

Hillier, F. S. & Liebermann, G. J. (2002): *Operations Research Einführung.* 5. Aufl., unveränd. Nachdr. d. 4. Aufl., Oldenbourg Verlag, München-Wien. ISBN 3-486-23987-2.

Hooijberg, M. (1997): *Practical Geodesy.* Springer-Verlag, Berlin-Heidelberg-New York. ISBN 3-540-61826-0.

Hreinsdóttir, S. (1999): *GPS Geodetic Measurements on the Reykjanes Peninsula, SW Iceland: Crustal Deformations from 1993 to 1998.* M. Sc. Thesis. University of Iceland, Reykjavík.

Iio, Y. & Kobayashi, Y. (2002): A physical understanding of large intraplate earthquakes. *Earth Planets Space* **54**: 1001–1004. ISSN 1343-8832.

Iio, Y. Sagiya, T. & Kobayashi, Y. (2004): What controls the occurrence of shallow intraplate earthquakes? *Earth Planets Space* **56**: 1077–1086. ISSN 1343-8832.

International GNSS Service (2008): The IGS Tracking Network.
`http://igscb.jpl.nasa.gov/network/complete.html`★

Ismail-Zadeh, A., Schubert, G., Tsepelev, I. & Korotkii, A. (2006): Three-dimensional forward and backward numerical modeling of mantle plume evolution: Effects of thermal diffusion. *J. Geophys. Res.* **111**: 1–15, DOI 10.1029/2005JB003782.

Jakobsdóttir, S. S., Guðmundsson, G. B. & Stefánsson, R. (2000): Seismicity in Iceland 1991-2000 monitored by the SIL seismic system. *Jökull The Icelandic Journal of Earth Science* **51**: 87–94. ISSN 0449-0576.

Jäger, R., Müller, T., Saler, H. & Schwäble, R. (2005): *Klassische und robuste Ausgleichsverfahren – Ein Leitfaden für Ausbildung und Praxis von Geodäten und Geoinformatikern.* Wichmann-Verlag, Heidelberg. ISBN 3-87907-370-8.

Jahn, C.-H. (1992): *Untersuchungen über den Einsatz des Global Positioning Systems (GPS) zum Nachweis rezenter Erdkrustenbewegungen im Spaltengebiet Nordost-Islands.* Wiss. Arb. d. Fachr. Verm.wesen d. Universtät Hannover **182**. ISSN 0174-1454.

Jet Propulsion Laboratory (2006): IGS-Solution.
`http://sideshow.jpl.nasa.gov`★

Kahmen, H. (1993): *Vermessungskunde.* Walter de Gruyter, Berlin-New York. ISBN 3-11-013733-X.

Kertz, W. (1969): *Einführung in die Geophysik Band 1 Erdkörper.* Bibliographisches Institut, Mannheim-Wien-Zürich. ISBN 3-8602-5695-5.

Keary, P. & Vine, F. J. (1990): *Global Tectonics.* Blackwell Science Pub., Oxford. ISBN 0-632-02425-9.

Kobarg, W. (1988): *Die gezeitenbedingte Dynamik des Ekström-Schelfeises, Antarktis.* Dissertation. Berichte zur Polarforschung Nr. 50, Bremerhaven.

Koch, K. R. & Schmidt, M. (1994): *Deterministische und stochastische Signale.* Dümmler Verlag, Bonn. ISBN 3-427-78911-X.

Koiran, P. & Sontag, E. D. (1996): Neural Networks with Quadratic VC Dimension. *NeuroCOLT Techn. Rep. Series* NC-TR-95-044.

Kraft, T., Wassermann, J., Schmedes, E. & Igel, H. (2006): Meteorological triggering of earthquake swarms at Mt. Hochstaufen, SE-Germany. *Tectonophysics* **424**: 245–258. DOI 10.1016/j.tecto.2006.03.044.

Kreemer, C., Holt, W. E., & Haines, A. J. (2003): An integrated global model of present-day plate motions and plate boundary deformation. *Geophys. J. Int.* **154**: 8–34. DOI 10.1046/j.1365-246X.2003.01917.x.

Kreutzmann, A., Schmeling, H., Junge, A., Ruedas, T., Marquart, G. & Bjarnason, I. Þ. (2004): Temperature and melting of a ridge-centered plume with application to Iceland. Part II: Predictions for the electromagnetic and seismic observables. *Geophys. J. Int.* **159**: 1097–1111. DOI 10.1111/j.1365-246X.2004.02397.x.

Kuhlmann, H. (1996): *Ein Beitrag zur Überwachung von Brückenbauwerken mit kontinuierlichen Messungen.* Dissertation. Wiss. Arb. d. Fachr. Verm.wesen d. Universtät Hannover **218**. ISSN 0174-1454.

Lambert, J., Winter, Th., Dewez, T. J. B. & Sabourault, Ph. (2005): New hypotheses on the maximum damage area of the 1356 Basel earthquake (Switzerland). *Quaternary Science Reviews* **24** (3-4): 381–399. DOI 10.1016/j.quascirev.2004.02.019.

Lawver, L. A. & Müller, R. D. (1994): Iceland hotspot track. *Geology* **22**: 311–314. ISSN 0091-7613.

Lecolazet, R. (1956): Application a L'analyse des Observations de la Marée Gravimétrique, de la Méthode de H. & Y. Labrouste dite par Combinaisons Linéaires D'ordonnées. *Annales de Géophysique* **12** (1): 59–71. ISSN 0003-4029.

Leiner, B. (1998): *Grundlagen der Zeitreihenanalyse.* 4. Aufl., München-Wien. ISBN 3-486-24756-5.

Luhmann, N. & Baecker, D. (Hrsg., 2002): *Einführung in die Systemtheorie.* 1. Aufl., Carl-Auer-Systeme-Verlag, Heidelberg. ISBN 3-89670-292-0.

Magnusson, I. Þ., Þorbergsson, G. & Björnsson, J. Þ. (1997): *GPS-mælingar í grunnstöðvaneti 1993 og ný viðmiðun ISN93 við landmælingar á Íslandi.* Landmælingar Íslands, Akranes.

Marko, H. (1995): *Systemtheorie: Methoden und Anwendungen für ein- und mehrdimensionale Systeme.* 3. neu bearb. und erw. Aufl. Springer-Verlag, Berlin-Heidelberg-New York-London-Paris-Tokyo-Hong Kong-Barcelona-Budapest. ISBN 3-540-58232-0.

Marquart, G. & Schmeling, H. (2004): A dynamic model for the Iceland Plume and the North Atlantic based on tomography and gravity data. *Geophys. J. Int.* **159**: 40–52. DOI: 10.1111/j.1365-246X.2004.02398.x.

Mathsoft (1995): *Benutzerhandbuch Mathcad.* Cambridge MA 02142, USA.

Mautz, R. (2001): *Zur Lösung nichtlinearer Ausgleichungsprobleme bei der Bestimmung von Frequenzen in Zeitreihen.* Dissertation. Deutsche Geodätische Kommission, Reihe C **532**, München. ISBN 3-7696-9571-7.

Maxlow, J. (2001): *Quantification of an Archaean to Recent Earth Expansion Process Using Global Geological and Geophysical Data Sets.* Dissertation, Curtin University of Technology. `http://adt.curtin.edu.au/theses/available/adt-WCU20020117.145715/unrestricted/`⋆

Melchior, P. (1983); *Tides of the Planet Earth.* Pergamon Press Oxford-New York-Toronto. ISBN 0-08-026248-1.

Merziger, G. & Wirth, Th. (1993): *Repetitorium der Höheren Mathematik.* 2. Aufl., Hannover. ISBN 3-923-923-33-3.

Michell, J. (1760): *Conjectures Concerning the Cause, and Observations Upon the Phenomena of Earthquakes: Particularly of That Great Earthquake of the First of November 1755, Which Proved So Fatal to the City of Lisbon and Whose Effects Were Felt as Far as Africa, and More or Less Throughout Almost All Europe.* London. British Library, Sig. HMNTS T.1559.(8.).

Miima, J. B. (2002): *Artificial Neural Networks and Fuzzy Logic Techniques for the Reconstruction of Structural Deformations.* Dissertation. Geod. Schriftenreihe d. Techn. Univ. Braunschweig Nr. 18. ISBN 3-926146-13-3.

Miller, H. (1992): *Abriß der Plattentektonik.* Stuttgart. ISBN 3-432-99731-0.

Miller, S. A. (2008): Note on rain-triggered earthquakes and their dependence on karst geology. *Geophys. J. Int.* **173**: 334–338. DOI 10.1111/j.1365-246X.2008.03735.x.

Minkler, G. & Minkler, J. (1993): *Theory and Application of Kalman Filtering.* Magellan Book Comp, Palm Bay, U.S.A. ISBN 0-9621618-2-9.

Möller, D. & Ritter, B. 1980: Geodetic Measurements and Horizontal Crustal Movements in the Rift Zone of NE-Iceland. *J. Geophys.* **47**: 110–119.

Möller, D. (1989): Terrestrische geodätische Arbeiten zur Erfassung horizontaler rezenter Oberflächenbewegungen. *Z. f. Verm.wesen* **114**: 10–25. ISSN 0340-4560.

Mosar, J., Lewis, G. & Torsvik, T. H. (2002): North Atlantic sea floor spreading rates: implications for Tertiary development of inversion structures of the Norwegian–Greenland Sea. *J. Geol. Soc.*, London **159**: 503–515. DOI 110.1144/0016-764901-135.

Müller, R. D., Gaina, C., Roest, W. R. & Hansen, D. L. (2002): A recipe for microcontinent formation. *Geology* **29** (3): 203–206. DOI 10.1130/0091-7613(2001)029<0203:ARFMF>2.0.CO;2.

Murphy, J. B. & Nance, R. D. (1995): *Gebirgsgürtel und der Superkontinentzyklus.* In Giese, P. (Hrsg.): *Geodynamik und Plattentektonik.* Spektrum Akad. Verlag, Heidelberg, ISBN 3-86025-373-5.

Natke, H. G. (1992): *Einführung in Theorie und Praxis der Zeitreihen- und Modalanalyse.* überarb. Aufl. Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig-Wiesbaden. ISBN 3-528-28145-6.

Neuner, H. B. (2008): *Zur Modellierung und Analyse instationärer Deformationsprozesse.* Deutsche Geodätische Kommission, Reihe C **616**, München. ISBN 3-7696-5055-7.

Niemeier, W. (1980): Zur Auswertung geodätischer Meßreihen. *Allgem. Verm.nachr.*: 41–59, Wichmann Verlag. ISSN 0002-5968.

Niemeier, W. (1981): Statistical Tests for Detecting Movements in Repeatedly Measured Geodetic Networks. *Tectonophysics* **71**: 335–351. DOI 10.1016/0040-1951(81)90076-7.

Niemeier, W. (1985a): *Deformationsanalyse.* In Pelzer H. (Hrsg.): *Geodätische Netze in Landes– und Ingenieurvermessung II.* Konrad Wittwer, Stuttgart: 559–619. ISBN 3-87919-140-9.

NIEMEIER, W. (1985b): *Netzqualität und Optimierung.* In PELZER H. (Hrsg.): *Geodätische Netze in Landes– und Ingenieurvermessung II.* Konrad Wittwer, Stuttgart: 153–224. ISBN 3-87919-140-9.

NIEMEIER, W. & TENGEN, D. (1988): *PANDA - A Menu Driven Software Package on a PC for Optimization, Adjustment and Deformation Analysis of Engineering Networks.* Proceedings. 5. Int. FIG-Symposium „Deformationsmessungen“, Fredericton, Canada: 374–376.

NIEMEIER, W., RENNEN, M., SALBACH, H. (2002): *Bestimmung regionaler und globaler Deformationen im Bereich der Antarktischen Halbinsel.* Deutsche Geodätische Kommission, Reihe B **310**: 109–126, München. ISBN 3-7696-8590-3.

NIEMEIER, W. (2002): *Ausgleichsrechnung. Eine Einführung für Studierende und Praktiker des Vermessungs- und Geoinformationswesens.* Walter de Gruyter, Berlin-New York. ISBN 3-11-014080-2.

NIEMCZYK, O. (1943): *Spalten auf Island.* Konrad Wittwer Verlag, Stuttgart. LCCN af47-06647.

NOLET, G., ALLEN, R. & ZHAO, D. (2007): Mantle plume tomography. *Chem. Geol.* **241**: 248–263. DOI 10.1016/j.chemgeo.2007.01.022.

NÜRNBERG, D. & MÜLLER, R. D. (1991): The tectonic evolution of the South Atlantic from Late Jurassic to present. *Tectonophysics* **191**: 27–53. DOI 10.1016/0040-1951(91)90231-G.

NÜTZLICHE SAMLUNGEN (1758): *Neue Erklärungen der Erdbeben.* Nützliche Samlungen, vom Jahre 1757. Dritter Theil. Hannover, gedruckt bey H: E. E. Schlüter, Landschaftl. Buchdrucker, Univ.-Bibl. TU Braunschweig, Sig. 2001-4610.

OGATA, K. (1978): *System Dynamics.* Prentice Hall. New Jersey. ISBN 0-13-880385-4.

ORTELIUS, A. (1587): *Abrahami Ortelij Thesaurus Geographicus: In quo omnium totius regionum, montium nomine & appellationes veteres, additis magna ex parte etiam recentioribus, concilientur.* Plantinus, Antverpiæ. Herzog August Bibliothek, Wolfenbüttel. Sig. S: Alv.: Ma 82.

ORTELIUS, A. (1611): *Abrahami Ortelii Antverpiani Thesavrvs Geographicvs: recognitvs et avctvs; in qvo Omnivm Totivs Terræ Regionvm, Montivm, Promontoriorvm nomina & appellationes veteres; additis magna ex parte etiam recentioribus* [In fine:] Hanoviæ, apud Hæredes Guilielmi Antonii. Herzog August Bibliothek, Wolfenbüttel. Sig. A: 4 Geogr.

PATTERSON, D. W. (1996): *Künstliche neuronale Netze: das Lehrbuch.* Prentice Hall. München-London-Mexiko-New York-Singapur-Sydney-Toronto. ISBN 3-8272-9531-9.

PELZER, H. (1985): *Theoretische Grundlagen.* In PELZER, H. (Hrsg.): *Geodätische Netze in Landes– und Ingenieurvermessung II.* Konrad Wittwer Verlag, Stuttgart. ISBN 3-87919-140-9.

PERLT, J. & HEINERT, M. (2002): *Kinematic modelling of surface velocities in SW-Iceland.* Abstract. XXVII. General Assembly of the European Geophysical Society (EGS), Nice (France), 21.–26. April 2002.

PERLT, J. & HEINERT, M. (2006): Kinematic modelling of surface velocities in SW-Iceland. *Geophys. J. Int.* **164**: 168–175. DOI 10.1111/j.1365-246X.2005.02795.x.

PERLT, J. (2006): *Ein Geokinematisches Modell für Island.* Dissertation. Geod. Schriftenreihe d. Techn. Univ. Braunschweig **20**. ISBN 3-926146-15-X.

PERLT, J., HEINERT, M. & NIEMEIER, W. (2008): The continental margin in Iceland – A snapshot derived from combined GPS networks. *Tectonophysics* **447**: 155–166. DOI 10.1016/j.tecto.2006.09.020.

PESONEN, L. J., SALMINEN, J., DONADINI, F. & MERTANEN, S. (2003): *Paleomagnetic Configuration of Continents During the Proterozoic with a Special Focus on Baltica.* XXI Geofysiikan Päivät Oulu 2003, Proceedings.
`http://spaceweb.oulu.fi/geofys03/pdf/...`
`...No30_Paleomagnetic_Configuration_of_Continents_.pdf`★

PESONEN, L. J., ELMING, S.-Å., MERTANEN, S., PISAREVSKY, S., D'AGRELLA-FILHO, M. S., MEERT, J. G., SCHMIDT, P. W., ABRAHAMSEN, N. & BYLUND, G. (2003): Paleomagnetic Configuration of Continents During the Proterozoic. *Tectonophysics* **375**: 289–324. DOI 10.1016/S0040-1951(03)00343-3.

PRESS, F. & SIEVER, R. (2003): *Allgemeine Geologie Einführung in das System Erde.* 3. Aufl. Spektrum Akademischer Verlag, München, ISBN 3-8274-0307-3.

PUSKORIUS, G. V. & FELDKAMP, L. A. (2001): *Parameter-based* KALMAN *Filter training: Theory and Implementation.* DOI 10.1002/0471221546.ch2, In HAYKIN, S. (Hrsg.): KALMAN *Filtering and Neural Networks.* John Wiley & Sons, New York-Chichester-Weinheim-Brisbane-Toronto-Singapore: 23–68. ISBN 0-471-36998-5

PYSKLYWEC R. N. & BEAUMONT, CHR. (2004): Intraplate tectonics: feedback between radioactive thermal weakening and crustal deformation driven by mantle lithosphere instabilities. *Earth and Planetary Sci. Lett.* **221**: 275–292. DOI 10.1016/S0012-821X(04)00098-6.

RARDIN, R.-L. (1998): *Optimization in Operation Research.* Prentice Hall, Upper Saddle River, USA. ISBN 0-02-398415-5.

RASMUSSEN, C. E. (2003): *Gaussian Processes in Machine Learning.* In BOUSQUET, O., VON LUXBURG, U., RÄTSCH, G. (Hrsg.): *Advanced Lectures on Machine Learning.* Tutorial in CARBONELL, J. G. & SIEKMANN, J. (Hrsg.): *Lecture Notes in Artificial Intelligence.* **3176**: 63–71. Springer, Berlin-Heidelberg-New York. ISBN 3-540-23122-6.

RICHTER, C. F. (1935): An Instrumental Earthquake Magnitude Scale. *Bull. Seismol. Soc. Am.* **25** (1): 1–32. ISSN 0037-1106.

RIEDEL, B. & HEINERT, M. (1998): *First results of GPS array observations in the grounding zone of Ekstroem Ice Shelf.* In OERTER, H. (Hrsg.): *Filchner-Ronne Ice Shelf Programme (FRISP)* Report **12**: 74–76, Alfred-Wegener-Institut für Polar- und Meeresforschung, Bremerhaven.

RIEDEL, B., NIXDORF, U., HEINERT, M., ECKSTALLER, A. & MAYER, C. (1999): The response of the Ekströmisen (Antarctica) grounding zone to tidal forcing. *Ann. Glac.* **29**: 239–242. DOI 10.3189/172756499781821247.

RIEDEL, B. (2002): *Modelle zur Beschreibung des gezeitenbedingten Bewegungsverhaltens von Schelfeisen in der Übergangszone.* Dissertation. Geod. Schriftenreihe d. Techn. Univ. Braunschweig **17**. ISBN 3-926146-12-5.

RIEDEL, C. (2001): *Self-Organizing Seafloor Mapping of the Tjoernes Fracture Zone / Iceland.* Dissertation. Christian-Albrecht. Univ. Kiel. Elektronische Veröffentlichung:
`http://eldiss.uni-kiel.de/macau/servlets/MCRFileNodeServlet/...`
`.../dissertation_derivate_00000500/d500.pdf`★

RITTER, B. (1982): *Untersuchungen geodätischer Netze in Island zur Analyse von Deformationen von 1965 bis 1977.* Deutsche Geodätische Kommission, Reihe C **271**, München. ISBN 3-7696-9323-X.

RITTER, B. (1986): Geodätische Messungen für geowissenschaftliche Zielsetzungen, insbesondere in Island. *Z. f. Verm.wesen* **111**: 548–552. ISSN 0340-4560.

ROGERS, J. J. W. & SANTOSH, M. (2004): *Continents and Supercontinents.* Oxford Univ. Press, Oxford-New York. ISBN 0-19-516589-6.

RÖGNVALDSSON, S. Þ., BÖÐVARSSON, R., SLUNGA, R. & STEINUNN, J. (1997): *The transfer function of the SIL seismic data acquisition system.* Icelandic Meteorological Service, Reykjavík:
`http://hraun.vedur.is/ja/skyrslur/silresp/resp-html.html`★

ROMM, J. (1994): A new forerunner for continental drift. *Nature* **367**: 407–408. DOI 10.1038/367407a0.

ROTH, V. (2001): *Kernel Methods for Regression and Classification.* Fortschr.-Ber. VDI Reihe 10 **671**, VDI Düsseldorf. ISBN 3-18-367110-7.

RUEDAS, T., SCHMELING, H., MARQUART, G., KREUTZMANN, A. & JUNGE, A. (2004): Temperature and melting of a ridge-centered plume with application to Iceland. Part I: Dynamics and crust production. *Geophys. J. Int.* **158**: 729–743. DOI 10.1111/j.1365-246X.2004.02311.x.

SANDWELL, D. T. & SMITH, W. H. F. (1997): Marine gravity anomaly from Geosat and ERS 1 satellite altimetry. *J. Geophys. Res.* **102**: 10039–10054. DOI 10.1029/96JB03223.

SANKARAN (2003): The supercontinent medley: Recent views. *Current Science* **85**: 1121–1124. ISSN 0011-3891.

Schlittgen, R. & Streitberg, B. H. J. (1997): *Zeitreihenanalyse.* 7. Aufl., R. Oldenbourg Verlag, München-Wien. ISBN 3-486-24175-3.

Schlittgen, R. (2001): *Angewandte Zeitreihenanalyse.* R. Oldenbourg Verlag, München-Wien. ISBN 3-486-25805-2.

Schmitt, M. (2001): *Radial basis function neural networks have superlinear VC dimension.* In Helmbold, D. & Williamson, B. (Eds.): *Proceedings of the 14^{th} Annual Conference on Computational Learning Theory COLT 2001 and 5^{th} European Conference on Computational Learning Theory EuroCOLT 2001, Lecture Notes in Artificial Intelligence.* **2111**:14–30. Springer-Verlag, Berlin. ISBN 978-3-540-42343-0.

Schmitt, M. (2005): On the capabilities of higher-order neurons: A radial basis function approach. *Neural Computation* **17** (3): 715–729. DOI 10.1162/0899766053019953.

Schölkopf, B. (1997): *Support Vector Learning.* GMD - Forschungszentrum Informationstechnik GmbH - **287**. Oldenbourg, München-Wien. ISBN 3-486-24632-1.

Schutzbach, W. 1985: *Island Feuerinsel am Polarkreis.* Dritte, völlig neue und erweiterte Auflage. Dümmler, Bonn. ISBN 3-427-88613-1.

Seeber, G. (1989): *Satellitengeodäsie.* Walter de Gruyter, Berlin-New York. ISBN 3-11-010082-7.

Seeber, G. (2003): *Satellite Geodesy.* 2^{nd} completely revised and extended edition, Walter de Gruyter, Berlin-New York. ISBN 3-11-017549-5.

Sella, G. F., Dixon, T. H. & Mao, A. (2002) REVEL: A model for recent plate velocities from space geodesy. *J. Geophys. Res.* **107**: 2081–2121. DOI 10.1029/2000JB000033.

Sens-Schönfelder, Chr. (2003): *Seismicity Patterns in the Hengill area in South West Iceland.* Diplomarbeit, Univ. Leipzig.
`http://hraun.vedur.is/ja/prepared/reports/christoph_thesis.pdf`★

Seyffert, J. (2002): *Ein stochastisches Volatilitätsmodell.* Diplomarbeit. Techn. Univ. Braunschweig. Unveröffentlicht.

Shen, Y., Solomon, S. C., Bjarnason, I. Þ& Wolfe C. J. (1998): Seismic evidence for a lower-mantle origin of the Iceland plume. *Nature* **395**: 62–65. DOI 10.1038/25714.

Sigmundsson, F., Einarsson, P. & Bilham, R. (1992): Magma chamber deflation recorded by the Global Positioning System: The Hekla 1991 eruption. *Geophys. Res. Lett.* **19**: 1483–1486. DOI 10.1029/92GL19.1483S.

Sigmundsson, F., Einarsson, P., Bilham, R., Sturkell, E. (1995): Rift-Transform kinematics in South Iceland: Deformation from Global Positioning System measurements, 1986 to 1992. *J. Geophys. Res.* **100**: 6235–6248. DOI 10.1029/95JB00155.

Sigmundsson, F., Einarsson, P., Rognvaldsson, S. Þ., Foulger, G. R., Hodgkinson, K. M. & Þorbergsson, G. (1997): The 1994-1995 seismicity and deformation at the Hengill triple junction, Iceland: Triggering of earthquakes by minor magma injection in a zone of horizontal shear stress. *J. Geophys. Res.* **102**: 15151–15161. DOI 10.1029/97JB00892.

Sigmundsson, F. (2006): *Iceland Geodynamics – Crustal Deformation and Divergent Plate Tectonics.* Springer publ. in assoc. with Praxis, Chichester, UK. ISBN 3-540-24165-5.

Sitter, R. (2001): *Neuronen.*
`http://home.arcor.de/ralf.sitter/kyb/neuro/neur.htm`★

Sleep, N. & Fujita, K. (1997): *Priciples of Geophysics.* Blackwell Science, Malden, USA. ISBN 0-86542-076-9.

Snider-Pellegrini, A. (1858): *La création et ses mystères dévoilés.* Librairie A. Franck & Librairie E. Dentu, Paris.

Sontag, E. D. (1998): *VC Dimension of Neural Networks.* In Bishop, C. (Ed.): *Neural networks and machine learning*: 69–95. Springer-Verlag, Berlin. ISBN 3-54-064928-X.

Spitzer, M. (2002): *Lernen - Gehirnforschung und die Schule des Lebens.* Spektrum Akademischer Verlag, Heidelberg. ISBN 3-82741-396-6.

STEFÁNSSON, R., BÖÐVARSSON, R., SLUNGA, R., EINARSSON, P., JAKOBSDÓTTIR, S. S., BUNGUM, H., GREGERSEN, S., HAVSKOV, J., HJELME, J., & KORHONEN, H. (1993): Earthquake prediction research in the South Iceland seismic zone and the SIL project. *Bull. Seismol. Soc. Am.* **83** (3): 696–716. ISSN 0037-1106.

STEFÁNSSON, R., GUÐMUNDSSON, G. B., HALLDÓRSSON P. (2000): *The two large earthquakes in the South Icelandic seismic zone on June 17 and 21, 2000.* Icelandic Meteorological Service, Reykjavík:
`http://hraun.vedur.is/ja/skyrslur/June17and21\_2000/index.html`★

STOCK, J. (2003): Hotspots Come Unstuck. *Sience* **301**: 1059–1060. DOI 10.1126/science.1089049.

STRERATH, G. (2006): Landesamt für Geoinformation Niedersachsen, persönliche Mitteilung.

STURKELL, E., SIGMUNDSSON, F., EINARSSON, P. & BILHAM, R. (1994): Strain accumulation 1986-1992 across the Reykjanes Penisula plate boundary, Iceland, determined from GPS measurements *Geophys. Res. Lett.* **21** (2): 125–128. DOI 10.1016/0148-9062(94)91084-7.

STÜWE, K. (2000): *Einführung in die Geodynamik der Lithosphäre Quantitative Behandlung geowissenschaftlicher Probleme.* Springer-Verlag, Berlin-Heidelberg-New York. ISBN 3-540-67516-7.

SUESS, E. (1883–1909): *Das Antlitz der Erde.* 3 Bde., unterteilt, Tempsky, Wien; Freytag Leipzig. Univ. Bibl. Techn. Univ. Braunschweig, Sig. 2224-3913/..26/..39/..42/..55/..68.

SUNDQUIST, U. & TRYGGVASON, E. (1982): *Deformation Measurement in the Hengill Region - Initial Measurement in 1979.* Nordic Volcanological Institute 8204, University of Iceland, Reykjavík.

SUTOR, TH. (1997): *Robuste Verfahren zur Analyse stochastischer Prozesse im Spektralbereich.* Dissertation. Schriftenreihe Studiengang Vermessungswesen **56**. Univ. d. Bundeswehr Neubiberg. ISSN 0173-1009.

TAUBENHEIM, J. (1969): *Statistische Auswertung geophysikalischer und meteorologischer Daten.* In BUCHHEIM, W., FANSELAU, G., HILLER, W. & MENZEL, H. (Hrsg.): *Geophysikalische Monographien.* Band 5, Akademische Verlagsgesellschaft, Leipzig. LCCN 71423056.

TARDUNO, J. A., DUNCAN, R. A., SCHOLL, D. W., COTTRELL, R. D., STEINBERGER, B., ÞORDARSON, Þ.,KERR, B. C., NEAL, C. R., FREY, F. A., TORII, M. & CARVALLO, C. (2003): The Emperor Seamounts: Southward Motion of the Hawaiian Hotspot Plume in Earth's Mantle. *Sience* **301**: 1064–1068. DOI 10.1126/science.1086442.

TORGE, W. (1989): *Gravimetry.* Walter de Gruyter, Berlin-New York. ISBN 3-11-010702-3.

TORGE, W. (2001): *Geodesy.* Walter de Gruyter, Berlin-New York. ISBN 3-11-017072-8.

TORGE, W. (2003): *Geodäsie.* 2. Aufl., Walter de Gruyter, Berlin. ISBN 3-11-017545-2.

TORSVIK, T. H., MOSAR, J. & EIDE, E. A. (2001): Cretaceous-Tertiary geodynamics: a North Atlantic exercise. *Geophys. J. Int.* **146**: 850–866. DOI 10.1046/j.0956-540x.2001.01511.x.

TORSVIK, T. H. (2003): The Rodinia Jigsaw Puzzle. *Sience* **300**: 1379–1381. DOI 10.1126/science.1083469.

TRIQUE, M., RICHON, P., PERRIER, F., AVOUAC, J. P. & SABROUX, J. C. (1999): Radon emanation and electric potential variations associated with transient deformation near reservoir lakes. *Nature* **399**: 137–140. DOI 10.1038/20161.

TRØNNES, R. G. (2002): *Geology and Geodynamics of Iceland.* Nordic Volcanological Institute, Reykjavík:
`http://norvol.hi.is/html/geol/intro/introduction.pdf`★

TRYGGVASON, E. (1974): *Vertical crustal movement in Iceland.* In: KRISTJANSSON, L. (Ed.): *Geodynamics of Iceland and the North Atlantic area.* Proceedings of the NATO Advanced Study Institute: 241–262. Reidel Publishing Company, Dordrecht-Boston/USA. ISBN 90-2077-5050-4.

UNBEHAUEN, R. (2002): *Systemtheorie 1. Allgemeine Grundlagen, Signale und lineare Systeme im Zeit- und Frequenzbereich.* 8. korr. Aufl., München-Wien. ISBN 3-486-25999-7.

USGS (1999): *Rejoined continents.* United States Geological Service:
`http://pubs.usgs.gov/gip/dynamic/continents.html`★

VALSSON, G. Þ., SIGURDSSON, Þ., VÖLKSEN, CHR., RENNEN, M. (2007): *ISNET 2004 Niðurstöður úr endurmælingum grunnstöðvanets Íslands.* Landmælingar Íslands, Akranes.

VAN BRUSSEL, H. (2002): *Systeemtheorie, deel 1.* H770 Systeemtheorie. Scriptum. Katholieke Universiteit Leuven.

VAN DER BOGAARD, C. (2002): persönliche Mitteilung.

WAPNIK, W. N.‡ & TSCHERWONENKIS, A. J. (1979): *Theorie der Zeichenerkennung*, in deutscher Spr. v. UNGER, S. & FRITZSCH, K., in FRÜHAUF, H., KÄMMERER, W., THIELE, H. & VÖLZ, H. (Hrsg.): *Elektronisches Rechnen und Regeln.* Sonder-Bd. 28, Akademie-Verlag, Berlin.

VAPNIK, V. N. (1998): *Statistical Learning Theory.* In HAYKIN, S. (Ed.): *Adaptive and Learning Systems for Signal Processing, Communications and Control* John Wiley & Sons, New York-Chichester-Weinheim-Brisbane-Singapore-Toronto. ISBN 0-471-03003-1.

VEÐURSTOFA ÍSLANDS (ICELANDIC METEROLOGICAL OFFICE) (2006): *Jarðskálftar.* Icelandic Meteorological Service, Reykjavík:
`http://hraun.vedur.is/ja/...`★

VIGNY, C., SIMONS, W. J. F., ABU, S., BAMPHENYU, R., SATIRAPOD, CH., CHOOSAKUL, N., SUBARYA, C., SOCQUET, A., OMAR K., ABIDIN, H. Z. & AMBROSIUS, B. A. C.(2005): Insight into the 2004 Sumatra-Andaman earthquake from GPS measurements in southeast Asia. *Nature* **436**: 201–206. DOI 10.1038/nature03937.

VINE, F. J. (1968): *Magnetic anomalies associated with mid-ocean ridges.* In: PHINNEY, R. A. (Ed.): *The History of the Earth's Crust.* Contributions to a conference held at the Goddard Institute for Space Studies, November 10 - 11, 1966, Princeton Univ. Press. ISBN 0-691-02379-4.

VINK, G. E. (1984): A hotspot model for Iceland and the Vøring Plateau. *J. Geophys. Res.* **89**: 9949–9959. DOI 10.1029/JB089iB12p09949.

VINK, G. E., MORGAN, W. J. & VOGT, P. R. (1995): *Hot Spots: heiße Flecken auf der Erde.* In GIESE, P. (Hrsg., 1995): *Geodynamik und Plattentektonik.* Spektrum Akad. Verlag, Heidelberg, ISBN 3-86025-373-5.

VÖLKSEN, C. & SEEBER, G. (1998): Nachweis von rezenten Krustenbewegungen in Nordisland mit GPS. *Z. f. Verm.wesen* **123**: 68–75. ISSN 0340-4560.

VÖLKSEN, C. (2000): *Die Nutzung von GPS für die Deformationsanalyse in regionalen Netzen am Beispiel Islands.* Dissertation. Wiss. Arb. d. Fachr. Verm.wesen d. Universtät Hannover **237**. ISSN 0174-1454.

WEGENER, A. L. (1915): *Die Entstehung der Kontinente und Ozeane.* In VOGEL, A. (Hrsg., 1980): *Nachdr. d. 1. u. 4. Aufl. u. mit e. Einl. u. e. Nachw. von Andreas Vogel.* Vieweg Verlag, Braunschweig-Wiesbaden. ISBN 3-528-07066-8.

WEGENER, A. L. (1929): *Die Entstehung der Kontinente und Ozeane.* 4., umgearb. Aufl. In WESTPHAL W. (Hrsg., 1962): *DIE WISSENSCHAFT Sammlung von Einzeldarstellungen aus allen Gebieten der Naturwissenschaft.* Bd. 66, Vieweg Verlag, Braunschweig.

WELSCH, W., HEUNECKE, O. & KUHLMANN, H. (2000): *Auswertung geodätischer Überwachungsmessungen.* In MÖSER, M., MÜLLER, H. & SCHLEMMER, H. (Hrsg.) *Handbuch Ingenieurgeodäsie.* Herbert Wichmann Verlag, Heidelberg. ISBN 3-87907-295-7.

WICKI, F. (2001): Robustes Schätzverfahren für die Ausgleichung geodätischer Netze. *Vermessung, Photogrammetrie, Kulturtechnik* **3**: 151–155. ISSN 0252-9424.

WIESER, A. (2001): *Robust and fuzzy techniques for parameter estimation and quality assessment in GPS.* Dissertation. Techn. Univ. Graz. ISBN 3-8265-9807-5.

WILSON, J. T. (1963): A possible origin of the Hawaiian Islands. *Canad. J. Phys.* **41**: 863–870. ISSN 1208-6045.

‡ Владимир Н. Вапник: Deut. Transkription: WLADIMIR N. WAPNIK und Engl. Transkription: VLADIMIR N. VAPNIK .

WILSON, J. T. (1966): Did the Atlantic close and then re-open? *Nature* **211**: 676–681. DOI 10.1038/211676a0.

WINSTON, W. L. (1994): *Operations Research: Applications and Algorithms.* 3rd ed., Duxbury Press, Belmont, California. ISBN 0-534-20971-8.

WITT, H. (2002): *Ansätze der finiten Elemente auf tektonische Transform-Bewegungen der Südisländischen Seismischen Zone.* Studienarbeit. FB Bauingenieurwesen, Techn. Univ. Braunschweig. Unveröffentlicht.

WOLFE, C. J., BJARNASON, I. Þ., VANDECAR, J. C. & SOLOMON, S. C. (1997): Seismic structure of the Iceland mantle plume. *Nature* **385**: 245–247. DOI 10.1038/385245a0.

WOLFE, C. J., BJARNASON, I. Þ., VANDECAR, J. C. & SOLOMON, S. C. (2002): Assessing the depth resolution of tomographic models of upper mantle structure beneath Iceland. *Geophys. Res. Lett.* **29** (2): 1015. DOI 10.1029/2001GL013657.

ZADEH, L. A. & DESOER, CH. A. (1963): *Linear System Theory The State Space Approach.* McGraw-Hill Book Comp., Inc., New York-San Francisco-Toronto-London. LCCN 63-14581.

ZADEH, L. A. & POLAK, E. (1969): *System Theory* McGraw-Hill Book Comp., Inc., New York-St. Louis-San Francisco-London-Sydney-Toronto-Mexico-Panama. LCCN 68-9051.

ZEGERS, T. E., DE WIT, M. J., DANN, J. & WHITE, S. H. (1998): Vaalbara, Earth's oldest assembled continent? A combined structural, geochronological, and palaeomagnetic test. *Terra Nova* **10**: 250–259. DOI 10.1046/j.1365-3121.1998.00199.x.

ZIMMERMANN, H. J. (1996): *Fuzzy Set Theory - and Its Aplications.* 3rd ed., Kluwer Academic Publishers, Boston-Dordrecht-London. ISBN-0-7923-9624-3.

ZSCHAU, J. (2002): Geoforschungszentrum Potsdam, persönliche Mitteilung.

ZUMBERGE, J. F., HEFLIN, M. B., JEFFERSON, D. C., WATKINS, M. M. & WEBB, F. H. (1997): Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J. Geophys. Res.* **102** B3: 5005–5017. DOI 10.1029/96JB03860.

Abbildungsverzeichnis

Tabellenverzeichnis

A Notation und Algorithmen

A.1 Symbolverzeichnis

A.1.1 Verwendete Mengen

$\mathbb{N}$	Menge der natürlichen Zahlen
$\mathbb{Z}$	Menge der ganzen Zahlen
$\mathbb{Q}$	Menge der rationalen Zahlen
$\mathbb{R}$	Menge der realen Zahlen
$\mathbb{C}$	Menge der komplexen Zahlen
$\mathbb{M}$	MANDELBROT-Menge
$\mathbb{T}$	Menge von Zeitpunkten

Dabei gelte weiterhin:
$\mathbb{N} = \mathbb{Z}^+ \subset \mathbb{N}^0 = \mathbb{Z}_0^+ \subset \mathbb{Z}$

A.1.2 Allgemeine Notation

a	Variable, Skalar
$\mathbf{a}$	Vektor
$\mathbf{A}$	Matrix
A	Operator

A.1.3 Formelzeichen

A	Amplitude
C_{xx}, C_{xy}	Autokovarianzfunktion, Kreuzkovarianzfunktion
$E\{\}$	Erwartungswert
F, F	FOURIER-Transformation, FOURIER-Transformierte
$\mathbf{H}$	HESSE-Matrix
$\mathbf{J}$	JACOBI-Matrix
$P\{\}$	Wahrscheinlichkeit
$R()$	Risikofunktion
R_{xx}, R_{xy}	Autokorrelationsfunktion, Kreuzkorrelationsfunktion
X, Y	Systemeingang, Systemausgang
g, g^*	Gewichtsfunktionen
h	Vapnik-Chervonenkis-Dimension
p	stochastisches Gewicht
s	empirische Standardabweichung - Skalenfaktor
$s()$	Sprungfunktion
u	Unbekannte - Parameter
w	nichtparametrisches Gewicht
x, y	realisierter Systemeingang, realisierter Systemausgang
Δ	Differenzoperator
$\Theta()$	asymptotisch untere Schranke
$\Phi()$	Modellfunktion
$\Psi()$	Einflussfunktion
$\Omega()$	Zielfunktion
$\delta()$	Deltadistribution - DIRAC-Impuls
η	Schrittweite, Lernrate
ν	Frequenz
$\varrho()$	Verlustfunktion
σ	theoretische Standardabweichung
χ	Zentrum eines RBF-Neuron
$\vec{\chi}$	Stützvektor einer beliebigen Ortsfunktion
ϕ	Phasenlage - Phasenverschiebung
$\varphi()$	Aktivierungsfunktion
ω	Umlaufzeit
$\mathcal{A}[\,]$	Abbildungsoperator
$\mathcal{F}$	Funktionsschar
$\mathcal{H}$	Anzahl von Neuronen
$\mathcal{K}[\,]$	Kriteriumsoperator der Verlustfunktion
$\mathcal{N}$	Bias, konstanter *offset* der Aktivierungsfunktion
$\mathcal{O}()$	asymptotisch obere Schranke
$\mathcal{T}$	Systemoperator
$\mathfrak{x}$	Ortsvektor
$\Re()$	Realteil einer komplexen Zahl
$\Im()$	Imaginärteil einer komplexen Zahl
∂	Differenzialoperator
∇	Gradientoperator

A.2 Die YULE-WALKER-Gleichung

Um für einen mittelwertstationären[1] Prozess y die Gewichte des zugehörigen AR[p]-Prozesses zu erhalten, benötigt man zunächst einen Vektor $\boldsymbol{\varrho}$ der mit p Korrelationswerten aus der Autokorrelationsfunktion R_{yy} besetzt wird. Dabei wird allerdings der 0-te Funktionswert übersprungen (KOCH & SCHMIDT 1994, S. 194, vergl. Formel 325.17).

$$\boldsymbol{\varrho}_k = R_{yy}(k+1) \quad \text{mit} \quad k = 0 \dots p-1 \tag{A-1}$$

Im zweiten Schritt wird die untere Korrelationsmatrix $\mathbf{R}^u$ aufgestellt, welche die ersten p Funktionswerte der Autokorrelationsfunktion einschließlich des 0-ten Funktionswertes enthält.

$$\mathbf{R}^u_{k,l} = R_{yy}(k-l) \quad \text{mit} \quad k = 0 \dots p-1 \quad \text{und} \quad l = 0 \dots k \tag{A-2}$$

Die vollbesetzte Korrelationsmatrix ergibt sich aus der Addition der unteren Korrelationsmatrix mit ihrer Transponierten. Da jetzt die Hauptdiagonale nur mit Elementen des Wertes 2 besetzt ist, muss eine Einheitsmatrix der Dimension p subtrahiert werden.

$$\mathbf{R} = \mathbf{R}^u + (\mathbf{R}^u)^{\mathrm{T}} - \mathbf{E}(\mathbf{p}) \tag{A-3}$$

Aus der umgestellten YULE-WALKER-Gleichung (SCHLITTGEN & STREITBERG 1997, S. 126, vergl. Formel 2.3.4.4) ergibt sich der unbekannte Gewichtsvektor $\mathbf{u}$ mit den p Elementen, welche den AR[p]-Prozess beschreiben.

$$\alpha = \mathbf{R}^{-1} \cdot \boldsymbol{\varrho} \tag{A-4}$$

A.3 Die Ψ-Funktionen und ihre Verlustfunktionen

Die Formel (4-6) beschreibt den Zusammenhang zwischen der Verlustfunktion $\varrho(v)$ und der Einflussfunktion $\Psi(v)$. Die Ableitungsvorschrift für Beträge lautet (NIEMEIER 2002, S. 193)

$$\frac{\partial |v|^s}{\partial v} = s \cdot v \cdot |v|^{s-2}. \tag{A-5}$$

Die inverse Integrationsvorschrift

$$\varrho(v) = \int \Psi(v) \partial v \tag{A-6}$$

ergibt sich für Beträge demnach zu

$$\int |v|^s \partial v = \frac{|v|^{s+2}}{(s+2) \cdot v} + C \text{ oder } \int v|v|^s \partial v = \frac{|v|^{s+2}}{(s+2)} + C. \tag{A-7}$$

Dabei ist gerade für intervallweise definierte Ψ-Funktionen auf die korrekte Bestimmung der Absolutglieder C zu achten.

[1] Diese Einschränkung ist zur Berechnung der Gewichte α_k an dieser Stelle erforderlich. Sollte der Prozess nicht mittelwertstationär sein, ist es hinreichend, vor der Berechnung eine Trendbereinigung vorzunehmen. Alternativ hierzu ist die Berechnung des empirischen Variogramms nach dem KRIGE- bzw. Kriging-Verfahren, bei dem sich die Mittelwertstationarität nicht voraussetzen lässt.

A.3.1 Die Ψ-Funktion der quadratischen Verlustfunktion

Die Methode der kleinsten Fehlerquadrate nutzt die quadratische Verlustfunktion

$$\varrho(v)_{\mathrm{L2}} = v^2. \tag{A-8}$$

Die zugehörige Einflussfunktion ist die unbeschränkte Gerade

$$\Psi(v)_{\mathrm{L2}} = v. \tag{A-9}$$

Obwohl die Beziehung zwischen Einflussfunktion und Verlustfunktion als Ableitung und Stammfunktion definiert ist, wird der Vorfaktor 2 nicht berücksichtigt. Zu diesem Zweck wird die Verlustfunktion des L2-Schätzers beispielsweise mit $\frac{1}{2}v^2$ angegeben (Jäger et al. 2005, S. 103).

A.3.2 Die Ψ-Funktion einer robusten Verlustfunktion

Die robusten Verlustfunktionen $\varrho(v)_{\mathrm{L1}} = |v|$, $\varrho(v)_{\mathrm{Lms}} = \mathrm{med}(|v|)$ und $\varrho(v)_{\max} = \max(|v|)$ besitzen trotz unterschiedlicher Minimierungseigenschaften die gemeinsame Einflussfunktion

$$\Psi(v)_{\mathrm{L1,\ Lms,\ max}} = \mathrm{sign}(v) = \frac{v}{|v|}. \tag{A-10}$$

Die Einlussfunktion der ebenfalls robusten Huber-Verlustfunktion

$$\Psi(v)_{\mathrm{Hu}} = \begin{cases} v & \forall\ |v| < b \\ b \cdot \mathrm{sign}(v) & \forall\ |v| \geq b \end{cases} \tag{A-11}$$

definiert eine intervallweise Nutzung der quadratischen und der linearen Verlustfunktion[2].

A.3.3 Die zurückfallende robuste Ψ-Funktion

Um die Wirkung von Ausreißern völlig auszuschließen, sind die *redescending Ψ-functions* entwickelt worden, wie beispielsweise Tukey*'s biweight* (Schlittgen & Streitberg 1997, S. 21)

$$\Psi(v)_{\mathrm{Tu}} = \begin{cases} v\left(1 - \frac{v^2}{b^2}\right)^2 & \forall\ |v| \leq b \\ 0 & \forall\ |v| > b \end{cases} \quad \text{mit } b = 6 \tag{A-12}$$

mit der resultierenden Verlustfunktion

$$\varrho(v)_{\mathrm{Tu}} = \begin{cases} \frac{v^6}{6b^4} - \frac{v^4}{2b^2} + \frac{v^2}{2} & \forall\ |v| \leq b \\ \\ \frac{b^2}{6} & \forall\ |v| > b \end{cases} \tag{A-13}$$

und die Ψ-Funktion von Hampel

$$\Psi(v)_{\mathrm{Ha}} = \begin{cases} v & \forall\ |v| \leq a \\ a \cdot \mathrm{sign}(v) & \forall\ a < |v| \leq b \\ \frac{a}{c-b} \cdot (c - |v|) \cdot \mathrm{sign}(v) & \forall\ b < |v| \leq c \\ 0 & \forall\ |v| > c \end{cases} \quad \text{mit } \begin{array}{rcl} a & = & 1,5 \\ b & = & 3,6 \\ c & = & 8,0 \end{array} \tag{A-14}$$

mit ihrer Verlustfunktion (Borutta 1988, vergl. S. 59f)

$$\varrho(v)_{\mathrm{Ha}} = \begin{cases} \frac{1}{2}v^2 & \forall\ |v| \leq a \\ a|v| - \frac{1}{2}a^2 & \forall\ a < |v| \leq b \\ \frac{a}{c-b}(c|v| - \frac{1}{2}v^2) - \frac{a}{2}\left(\frac{b^2}{c-b} + a\right) & \forall\ b < |v| \leq c \\ \frac{a}{2}(c + b - a) & \forall\ |v| > c \end{cases}. \tag{A-15}$$

---

[2] In der gegebenen Form wird diese Einflussfunktion bereits im BIBER-Schätzer standardmäßig zur Netzausgleichung durch das Schweizer Bundesamt für Landestopografie verwendet (Wicki 2001, S. 151ff).

A.3.4 Herleitung der Ln-Schätzer für die Momente eines Prozesses

Sei $\varrho(v) = \frac{k}{2}\ln(v^2 + k)$ robuste Verlustfunktion und sei unter dem Kriterium

$$\frac{k}{2}\sum_{i=1}^{N}\ln\left(\left(\frac{v_i}{s_i}\right)^2 + k\right) \rightarrow \min \tag{A-16}$$

der stationäre Punkt des Optimums zu bilden, so nehme die zugehörige Einflussfunktion als erste Ableitung der Verlustfunktion unter dem Summenkriterium

$$\sum_{i=1}^{N}\frac{\frac{k\cdot v_i}{s_i}}{(\frac{v_i}{s_i})^2 + k} = 0 \tag{A-17}$$

den Wert Null an. Mit der Erweiterung $v_i = \bar{y} - y_i$ ergibt sich

$$\sum_{i=1}^{N}\frac{\frac{k\cdot(\bar{y}-y_i)}{s_i}}{(\frac{\bar{y}-y_i}{s_i})^2 + k} = 0.$$

Die Trennung der Summe führt zu

$$\sum_{i=1}^{N}\frac{\frac{k\cdot\bar{y}}{s_i}}{(\frac{\bar{y}-y_i}{s_i})^2 + k} = \sum_{i=1}^{N}\frac{\frac{k\cdot y_i}{s_i}}{(\frac{\bar{y}-y_i}{s_i})^2 + k}.$$

Da $\bar{y}$ =const ist, kann es vor die Summe gezogen werden und in der Form

$$\bar{y} = \frac{\sum\limits_{i=1}^{N} k\cdot y_i\left[\left(\frac{\bar{y}-y_i}{s_i}\right)^2 + k\right]^{-1}}{\sum\limits_{i=1}^{N} k\cdot\left[\left(\frac{\bar{y}-y_i}{s_i}\right)^2 + k\right]^{-1}} \tag{A-18}$$

aufgelöst werden. Entsprechend der Darstellung von JÄGER ET AL. (2005, Kap. 4.5) ergibt sich ein M-Schätzer

$$\bar{y} = \frac{\sum\limits_{i=1}^{N} g_i\cdot y_i}{\sum\limits_{i=1}^{N} g_i} \tag{A-19}$$

mit der Gewichtsfunktion

$$g_i = \frac{1}{\left(\frac{\bar{y}-y_i}{s_i}\right)^2 + k} = \frac{1}{(v^*)^2 + k}. \tag{A-20}$$

Da $\bar{y}$ von sich selbst abhängig ist, kann (A-18) iterativ gelöst werden, wobei sich die Näherung $\bar{y} \approx \mathrm{med}(v)$ als günstig erwiesen hat. Oft ist eine weitere Iteration nicht erforderlich.

A.4 Iterative Berechnung der pseudo-inversen Hesse-Matrix

Der als erstes entwickelte Ansatz von Davidson, Fletcher, Powell oder kurz DFP-Algorithmus ist zugleich auch der einfachste (Bazaraa & Shetty 1979, S. 300ff). Auf eine ausführliche Herleitung sei hier verzichtet:

$$\mathbf{H}^{-}(\mathbf{u}_{k+1}) = \mathbf{H}^{-}(\mathbf{u}_k) - \frac{\mathbf{H}^{-}(\mathbf{u}_k) \cdot \Delta\nabla f(\mathbf{u}_{k+1}) \cdot \left[\mathbf{H}^{-}(\mathbf{u}_k) \cdot \Delta\nabla f(\mathbf{u}_{k+1})\right]^{\mathrm{T}}}{\left[\Delta\nabla f(\mathbf{u}_{k+1})\right]^{\mathrm{T}} \cdot \mathbf{H}^{-}(\mathbf{u}_k) \cdot \Delta\nabla f(\mathbf{u}_{k+1})}. \tag{A-21}$$

Der neueste Nachfolger ist zugleich auch der beste derzeit bekannte Algorithmus zur Schätzung der pseudo-inversen Hesse-Matrix (Haykin 1999, S. 243f). Dieser Ansatz nach Broyden, Fletcher, Goldfarb, Shanno oder kurz BFGS-Algorithmus ergibt sich wie folgt[3] (Grundmann 2002, S. 119):

$$\begin{aligned}
\mathbf{H}^{-}(\mathbf{u}_{k+1}) \quad &= \quad \mathbf{H}^{-}(\mathbf{u}_k) - \frac{\mathbf{H}^{-}(\mathbf{u}_k)\,\Delta\nabla f(\mathbf{u}_{k+1})\left[\mathbf{H}^{-}(\mathbf{u}_k)\,\Delta\nabla f(\mathbf{u}_{k+1})\right]^{\mathrm{T}}}{\left[\Delta\nabla f(\mathbf{u}_{k+1})\right]^{\mathrm{T}}\mathbf{H}^{-}(\mathbf{u}_k)\,\Delta\nabla f(\mathbf{u}_{k+1})} \\
&\quad + \frac{\Delta\mathbf{u}_{k+1}\left[\Delta\mathbf{u}_{k+1}\right]^{\mathrm{T}}}{\left[\Delta\mathbf{u}_{k+1}\right]^{\mathrm{T}}\Delta\nabla f(\mathbf{u}_{k+1})} + \left[\Delta\mathbf{u}_{k+1}\right]^{\mathrm{T}}\Delta\nabla f(\mathbf{u}_{k+1}) \cdot \mathbf{h} \cdot \mathbf{h}^{\mathrm{T}} \\
&\text{mit} \\
\mathbf{h} \quad &= \quad \frac{\Delta\mathbf{u}_{k+1}}{\left[\Delta\mathbf{u}_{k+1}\right]^{\mathrm{T}}\Delta\nabla f(\mathbf{u}_{k+1})} \\
&\quad - \frac{\mathbf{H}^{-}(\mathbf{u}_k)\,\Delta\nabla f(\mathbf{u}_{k+1})}{\left[\Delta\nabla f(\mathbf{u}_{k+1})\right]^{\mathrm{T}}\mathbf{H}^{-}(\mathbf{u}_k)\,\Delta\nabla f(\mathbf{u}_{k+1})}.
\end{aligned} \tag{A-22}$$

[3] Es existiert eine weitere formal unterschiedliche Darstellung dieses Algorithmus (Rardin 1998, S. 766ff).

B Daten

B.1 GPS-Koordinatenlösungen

Das *Jet Propulsion Laboratory* am *California Institute of Technology* stellt auf seinen Internetseiten `http://sideshow.jpl.nasa.gov/mbh/series.html` (JPL 2006) die täglichen globalen Koordinatenlösungen für die Stationen des *International GNSS Service* (IGS) zur Verfügung.

Hier können für jede Station die Residuen in den drei Komponenten geografische Breite (SSSS.lat) und Länge (SSSS.lon) sowie lokale ellipsoidische Höhe (SSSS.rad) abgerufen werden. Die Dateien sind spaltenweise aufgebaut. Die Spalten enthalten: die dezimale Zeit in Jahren, das Residuum in Metern, die Standardabweichung in Metern, die alphanumerische Namenskürzel SSSS der Station, die Koordinatenkomponente und schließlich das Datum:

```
1996.0848  -0.222824464845475E+02   0.271219750863664E+00  REYK  LAT  96FEB01
1996.0876  -0.222162155723943E+02   0.289081471365867E+00  REYK  LAT  96FEB02
1996.0903  -0.221973432781450E+02   0.262594028476085E+00  REYK  LAT  96FEB03
1996.0958  -0.226464063210154E+02   0.264456141606555E+00  REYK  LAT  96FEB05
1996.0985  -0.223749879672828E+02   0.266561203076293E+00  REYK  LAT  96FEB06
1996.1013  -0.223520995020818E+02   0.281548274111164E+00  REYK  LAT  96FEB07
1996.1040  -0.219748820085881E+02   0.260822681271048E+00  REYK  LAT  96FEB08
1996.1068  -0.224467672740406E+02   0.273119790449057E+00  REYK  LAT  96FEB09
1996.1095  -0.219596741849649E+02   0.292822530162114E+00  REYK  LAT  96FEB10
1996.1123  -0.221960611392846E+02   0.307812126610519E+00  REYK  LAT  96FEB11
1996.1150  -0.215688424745915E+02   0.280406905197091E+00  REYK  LAT  96FEB12
1996.1177  -0.217370134032509E+02   0.271749542452775E+00  REYK  LAT  96FEB13
1996.1232  -0.212414362389046E+02   0.293940699384698E+00  REYK  LAT  96FEB15
1996.1260  -0.218995268707308E+02   0.285194087891956E+00  REYK  LAT  96FEB16
```

...

B.1.1 Station Höfn in Südostisland (HOFN)

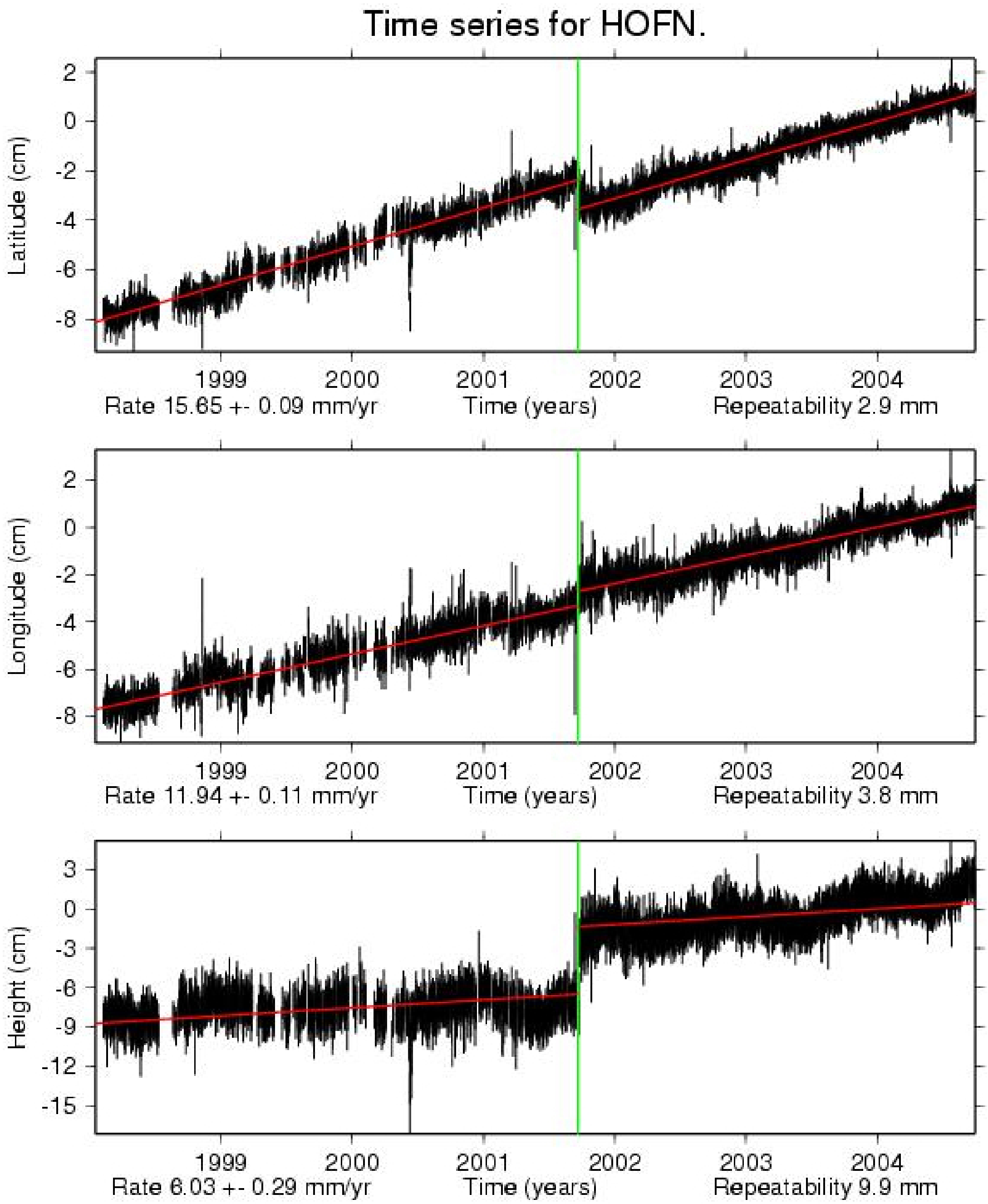


Abb. B.1: Tageslösungen der Station Höfn (Südostisland).

B.1.2 Station Onsala in Südschweden (ONSA)

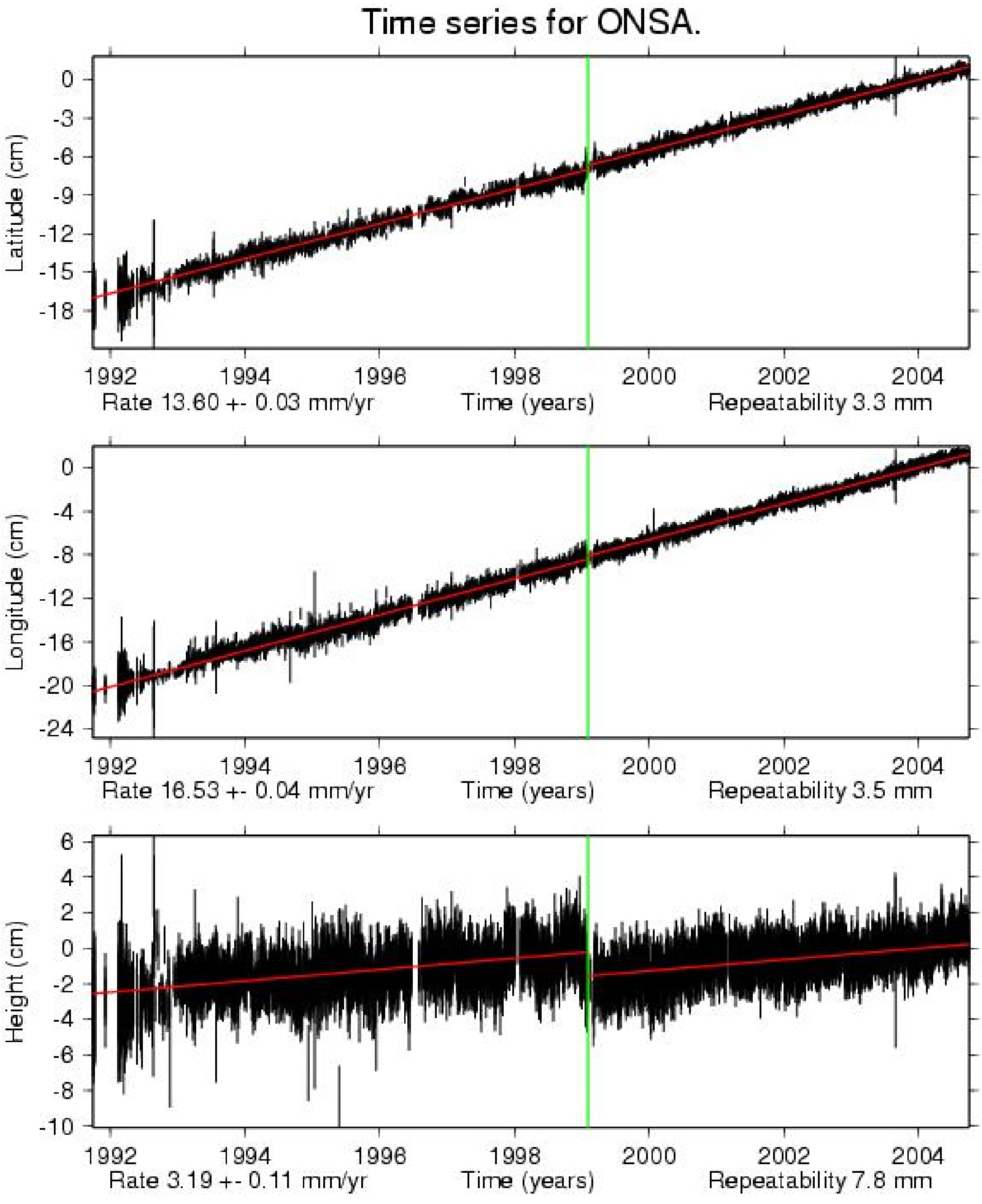


Abb. B.2: Tageslösungen der Station Onsala (Südschweden).

B.1.3 Station Reykjvík in Südwestisland (REYK)

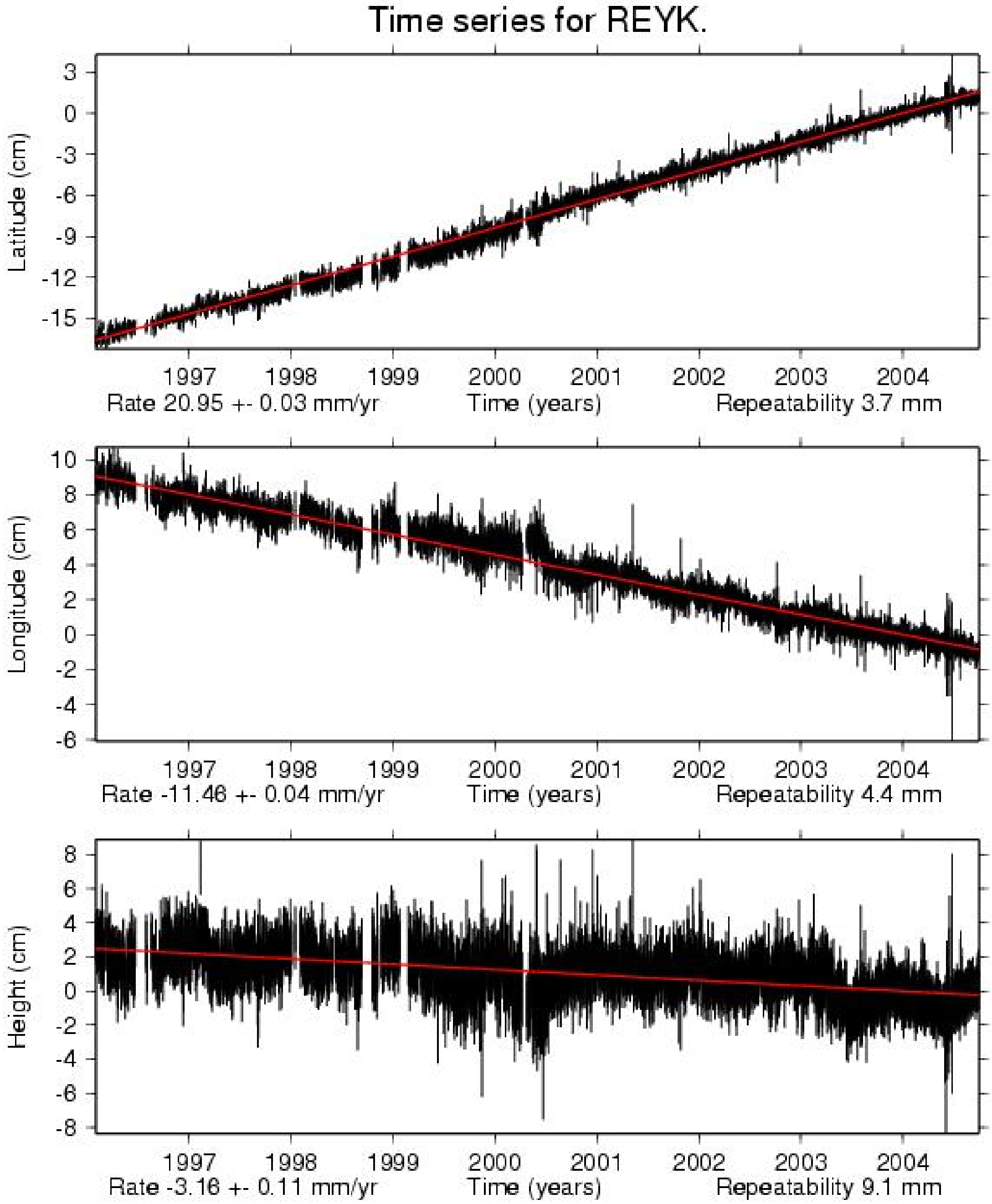


Abb. B.3: Tageslösungen der Station Reykjavík (Südwestisland).

B.2 Erbebendaten

Auf den Seiten des Isländischen Wetterdienstes (*Veðurstofa Íslands*) befinden sich die vorläufigen Daten des automatischen Seismometernetzes. Die jeweiligen Wochenlösungen können als Karte (`http://hraun.vedur.is/ja/viku/2000/vika_24/index.html#mark`) und als Erdbebenliste (`http://hraun.vedur.is/ja/viku/2000/vika_24/listi`) eingesehen werden. Die

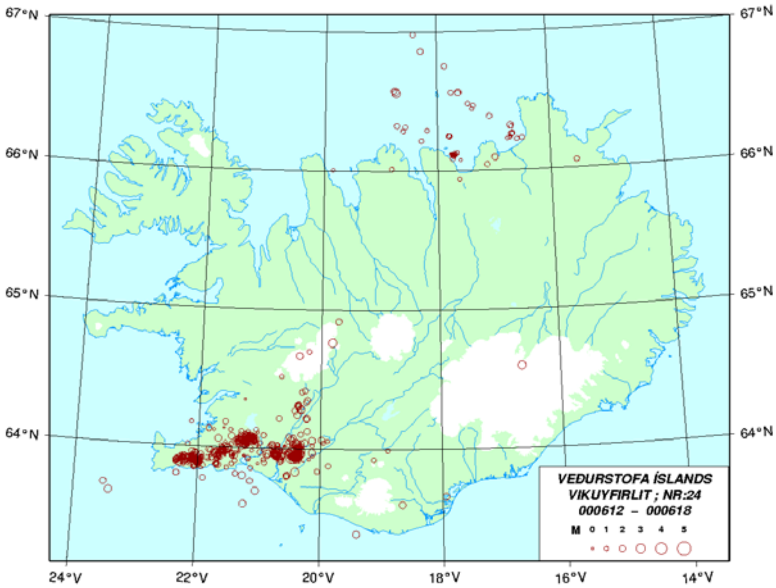


Abb. B.4: Erdbebenkarte einer Wochenlösung des automatischen Seismometernetzes: Lösung der Woche 24 des Jahres 2000, Quelle: Isländischer Wetterdienst.

Listendateien sind spaltenweise aufgebaut. Die Spalten enthalten die laufende Nummer (Nr) des Ereignisses, die Datumsangabe (Dags.) in der Form aaaammdd, die Tageszeit (Timi) in der Form hhmmss.sss, die geografische Breite (Breidd), die geografische Länge (Lengd), die Herdtiefe (Dypi), die Momentmagnitude (M) und die lokale Gutenberg-Richter-Magnitude (ML):

Nr	Dags.	Timi	Breidd	Lengd	Dypi	M	ML
1	20000612	000249.411	63.92150	-22.02515	7.287	0.79	0.61
2	20000612	004231.740	63.94273	-21.24223	7.519	0.01	0.33
3	20000612	013356.466	63.90697	-21.82173	8.421	0.65	0.50
4	20000612	014528.612	66.05756	-15.54608	12.192	1.39	1.48
5	20000612	020540.879	63.98978	-18.89152	48.751	0.70	0.85
6	20000612	041348.608	66.31555	-18.57417	4.998	0.96	1.12
7	20000612	041532.107	64.69887	-20.21022	10.866	1.52	1.06
8	20000612	052550.664	66.86067	-18.29583	14.481	2.07	1.71
9	20000612	060513.423	63.99319	-20.94340	6.190	-0.09	-0.07
10	20000612	080831.314	63.94141	-21.25605	7.611	0.13	-0.07
11	20000612	081924.073	66.56184	-17.75323	0.135	1.56	1.33
12	20000612	095818.011	66.48480	-17.45254	2.990	1.37	1.13
13	20000612	111657.520	66.11916	-17.76249	8.446	0.10	0.70
14	20000612	131832.975	66.12560	-17.70076	8.698	0.66	1.03

...

C Historische Quellen

C.1 Abraham Ortelius

„Die Spuren des Bruches treten zutage, wenn sich jemand die Küsten dieser drei Erdteile, wo sie sich gegenüber liegen, auf einer Weltkarte ansieht und die herausragenden Formen Europas ebenso wie auch die Afrikas mit den Ausbuchtungen Amerikas genauer betrachten würde.“ Die Brisanz dieser Entdeckung war Ortelius sehr wohl bewusst. Er erlebte die religiös-politische Destabiliserung vom Religionsfrieden (1555) bis zum Beginn der Gegenreformation als Auslöser des 80-jährigen Krieges in den Niederlanden. Und obgleich ihm die Worte zugeschrieben werden: „In diesen Zeiten ist es weise zu schweigen“, mag er seine Erkenntnis dennoch nicht völlig verschweigen. Er platziert seine Aussage höchst unauffällig unter einem unwichtigen Stichwort, wo kaum ein Rezensor nachlesen würde:

> Graeci nomen EUMELI, $\epsilon\upsilon\mu\eta\lambda o\upsilon$, olim habuit, vernaculae linguae Gadiri appellationem referens. Ut refert Plato, in Critia, sive Atlantide. Nisi fabula sit, Gadir sive Gades pars erit reliqua Atlantidis sive Americae insulae, atque haec non tam submersa (ut idem refert in Timaeo) quam ab Europa atque Africa terrae motu et illuvione abrupta: et rectà occidentem versus elongata videbitur. Quod si quis hoc Fabulam fabula compensare vocet, per me quidem licebit. **Ostendunt se rupturae vestigia, si quis harum trium dictarum terrae (adhibita geographica universali tabula) partium littora, quo se mutuo aspiciunt, eminentiasque Europae nempe atque Africae cum concavitatibus Americae penitius consideraverit**. Adeo ut quis posset cum Strabone 2 dicere, non esse figmentum quod Plato ex Solonis sententia, de Atlantis insula prodiderit.

Zitat aus Ortelius (1611), siehe hierzu auch Ortelius (1587) und Romm (1994).

C.2 Sir Francis Bacon

> Inter praerogativas instantiarum ponemus sexto loco *Instantias Conformes*, sive *Proportionatas*; quas etiam *Parallelas*, sive *Similitudines Physicas*, appellare consuevimus. [...] etiam in ipsa configuratione mundi in majoribus non sunt negligendae instantiae conformes; veluti *Africa*, et regio *Peruviana* cum continente se porrigente usque ad *Fretum Magellanicum*. Utraque enim regio habet similes isthmos et similia promontoria, quod non temere accidit. Item Novus et Vetus Orbis; in eo quod utrique orbes versus septentriones lati sunt et exporrecti, versus austrum autem angusti et acuminati. [...]

Zitat aus Bacon (1620), *Aphorismorum de Interpretatione Naturae sive de Regno Hominis* [Liber Secundus], XXVII.

C.3 Unbekannter Verfasser

Neue Erklärungen der Erdbeben
Es ist besonders und zeuget von unse≈rer Schwachheit und unvollkomme≈nen Erkäntniß, daß wir eben da die mehresten Zweifel finden, und die entferntesten Ursachen suchen, wo die natür≈lichen Begebenheiten am klärsten, und die würkenden Ursachen davon am nächsten sind. Die Erdbeben haben ein solches Schicksal gehabt. So natürlich bekannt solche sind, so sehr haben sich die Naturkündiger den Kopf zerbrochen, um die Ursache davon aus≈fündig zu machen. Der Eine stellete mit der Elektrisir≈Maschine allerley Spiele an, und es schien ihm gewiß, daß die Elektricität den ganzen Erdboden erschüttern könne; Ein Anderer fand unter der Erde sulphurische Theile; Ein Dritter hörete gar Gewitter unter der Erde; Ein Vierter roch unterirdische Dünste; Ein Fünfter wolte schon, um solche auszuführen, aus dem Innersten der Erde Canäle heraus bauen, Einem Sechsten brauseten die unterirdischen Winde in seinen Ohren: Ein Siebenter fand zwischen den feuerspeyenden Bergen und dem Erdbe≈ben einen Zusammenhang; Ein Achter berief sich auf die unterirdischen Wasser; Ein Neunter verband die unterirdischen Feuer und Quellen zusammen; Ein Zehnter bemerkte in der magnetischen Materie eine ganze Zerrüttung; Ein Elfter sahe weiter über sich, und fand in dem veränderten Druk≈ke der Schwere eine neue Ursache, ohne zu bedenken, daß eine anderweitige würkende Ursache nötig sey, von welcher dieser veränderte Drucke in der Schwere herrühre; Ein Zwölfter stieg wol gar mit seinen Gedanken in den Mond, und mich wundert,
daß noch niemand seine Zuflucht zu des *Pontoppidans* grossen See≈Ungeheuer, dem Kra≈ken oder zu dessen Seewurm genommen hat. Glaubte der eine seine Meynung gegründet, so fanden die übrigen Eilfe desto mehr Widerspruch darin. [...]

Zitat aus *Nůtzliche Samlungen* (1758) [Dritter Theil], S. 545ff.

Dank

Mein erster Dank gilt Herrn Univ.-Prof. Dr.-Ing. habil. Wolfgang Niemeier, der mir einen weiten inhaltlichen Freiraum für diese Arbeit eingeräumt und mit Langmut die Fertigstellung dieser Arbeit erwartet hat. Herrn Univ.-Prof. Dr. rer. nat. Harro Schmeling danke ich für viele Diskussionen auch im Rahmen des Forschungsverbundes „Deutsche Islandforschung“, für die stete und weitreichende Unterstützung sowie für die bereitwillige Übernahme des Zweitreferats. Ferner möchte ich dem Vorsitzenden Herrn Univ.-Prof. Dr.-Ing. Dietrich Möller herzlich für sein stetes Interesse an allen Ergebnissen im Arbeitsschwerpunkt Island und seine vielen Anregungen danken.

Für die vielen fruchtbaren und ermutigenden Diskussionen beim Kaffee bin ich Herrn Dr.-Ing. Björn Riedel sehr dankbar. Seine herausfordernden Fragen bargen nicht selten bereits den Kern der Lösung manchen Problems in sich. In der gleichen Kaffeerunde war mir das breite und zugleich detailreiche geodätische Wissen von Herrn Dr.-Ing. James Perlt eine große Hilfe. Von großem Wert war auch seine stete Bereitschaft, mich bei Problemen mit LaTeX zu unterstützen. Herr Stefan Reiser und Herr Wolfgang Schellin haben mich in der Lehre und in verschiedenen Projekten nach Kräften entlastet.

Herr ir. Jeroen Meeus hat mir im richtigen Moment die Feinheiten der Optimierung nahegebracht und mir zusammen mit Herrn ir. Willem Waegeman fundamentale Hinweise zur Nutzung der nichtparametrischen Modelle gegeben.

Ebenfalls im Rahmen des Forschungsverbundes „Deutsche Islandforschung“ habe ich viel Unterstützung von Frau Doz. Dr. phil. nat. Gabriele Marquart erfahren. In den Diskussionen lernte ich ihre bewundernswerte Gabe schätzen, um komplexe geophysikalische Zusammenhänge allgemeinverständlich erklären zu können.

Für konstruktive Kritik, hilfreiche Diskussionen, angenehme Zusammenarbeit und/oder Erlaubnis zur Nutzung von eigenen Daten danke ich Ph.D. Ingi Þ. Bjarnason, O. Univ.-Prof. Dipl.-Ing. Dr. techn. Fritz K. Brunner, Univ.-Prof. Dr.-Ing. habil. Reinhard Dietrich, Prof. Ph.D. Páll Einarsson, Ph.D. Michael Heflin, Univ.-Prof. Dr. rer. nat. Wolfgang R. Jacoby, Univ.-Prof. Dr.-Ing. habil. Hansjörg Kutterer, Dr.-Ing. John-Bosco Miima, Dipl.-Ing. Hilke Perlt, geb. Witt, Dr. rer. nat. Carsten Riedel, Univ.-Prof. Dr.-Ing. Bernhard Ritter, Ph.D. Freysteinn Sigmundsson, Ph.D. Ragnar Stefánsson und Dipl.-Ing. Dieter Tengen.

Für tiefe Einblicke in die geisteswissenschaftliche Forschungsmethoden und die freundschaftliche Ermutigung danke ich Christian Werner.

Meiner Ehefrau Handelsing. Elke Meeus danke ich, dass sie für diese Arbeit eigens nach Deutschland gezogen ist, mich mit liebender Geduld gewähren ließ und mir mit großem Organisationsgeschick viele Freiräume geschaffen hat.

Schließlich danke ich der Deutschen Forschungsgemeinschaft (DFG). Diese Arbeit hat sich aus der Bearbeitung der geförderten Projekte NI220/5-2, 7-1, 8-1 u. 2, 9-1 entwickelt.

Lebenslauf

Michael Heinert

	geboren am 22.03.1972 in Bremervörde verheiratet seit 2002 mit der Handelsingenieurin Elke Meeus zwei Kinder (Bastiaan, Hannah)
1978 - 1980	Grundschule Bremervörde-Engeo
1980 - 1982	Grundschule WittingenII (Landkreis Gifhorn)
1982 - 1984	Orientierungsstufe Wittingen (Landkreis Gifhorn)
1984 - 1986	Gymnasium Hankensbüttel (Landkreis Gifhorn)
1986 - 1991	Julianum Gymnasium Helmstedt
Abschluss	Abitur
1991 - 1992	Grundwehrdienst
1992	Studienpraktikum am Katasteramt Helmstedt
1992 - 1994	Vorstudium des Vermessungswesens an der Technischen Universität Carolo-Wilhelmina zu Braunschweig
1994 - 1998	studentische Hilfskraft am Institut für Vermessungskunde (seit 1995 Institut für Geodäsie und Photogrammetrie) der Technischen Universität Carolo-Wilhelmina zu Braunschweig
1994 - 1998	Hauptstudium des Vermessungswesens an der Universität Hannover
Abschluss	Diplomingenieur
1998	wissenschaftliche Hilfskraft am Institut für Geodäsie und Photogrammetrie der Technischen Universität Carolo-Wilhelmina zu Braunschweig
1999 - 2006	wissenschaftlicher Mitarbeiter in den DFG-Projekten „Kinematische und dynamische Modellierung der südwest-isländischen Plattengrenzen“, “Analyse von Oberflächendeformationen in Folge aktueller Erdbeben in der Südisländischen Seismischen Zone“, “Analyse kontinuierlicher Beobachtungen mit Hilfe von systemtheoretischen Lösungsverfahren“
seit 2006	planmäßiger wissenschaftlicher Mitarbeiter