

Institut für Geodäsie und Geoinformation

Professur für Theoretische Geodäsie

Strömungsmodellierung im Nordatlantik durch
Auswertung von Oberflächendrifterdaten

Masterarbeit

im Masterstudiengang Geodäsie und Geoinformation

an der Hohen Landwirtschaftlichen Fakultät

der Rheinischen Friedrich–Wilhelms–Universität

zu Bonn

vorgelegt am 29. September 2014 von

B. Sc. Jessica Franken

aus Geilenkirchen

Bonn 2014



Erster Prüfer: Prof. Dr. techn. W.-D. Schuh

Zweiter Prüfer: Dr.-Ing. Silvia Müller, geb. Becker

1. April 2014

Masterarbeit Jessica Franken

Thema: Strömungsmodellierung im Nordatlantik durch Auswertung von Oberflächendrifterdaten

Die Oberflächendrifter des Global-Drifter-Programms liefern seit 1979 Informationen über die Ozeanströmungen in den Weltmeeren. Derzeit werden circa 1100 über die Weltmeere verteilte Oberflächenbojen durch das satellitengestützte ARGOS-System ständig überwacht. Aus der Doppler-Frequenzverschiebung der aktiven Sender der Bojen können Positionen mit einer Genauigkeit von besser als 150 m bestimmt werden. Durch die Satellitenkonstellation kann somit eine Ortung spätestens alle vier Stunden gewährleistet werden. Da sich diese Oberflächenbojen mit der horizontalen Strömung der Ozeanoberfläche bewegen, ist es möglich mit diesen Lagrange-Messungen eine Strömungsmodellierung der Ozeane durchzuführen. Informationen über die Strömungen werden aus den aus Positionsdaten abgeleiteten Geschwindigkeiten gewonnen. Letzte Literaturhinweise zur Verarbeitung der Daten sind allerdings aus dem Jahr 2005 und verweisen auf eine Auswertung der Daten des Jahres 1996.

- In einem ersten Schritt soll geprüft werden, ob die Modellierung von Strömungen im Nordatlantik basierend auf den bereits vorverarbeiteten Datensätzen sinnvoll ist. Auf einem regelmäßigen Zeitraster vorprozessierte Geschwindigkeitsfelder dienen dabei als Ausgangspunkt. Eine Analyse dieser Daten soll Rückschlüsse auf die geeignete Auflösung und Modellierung der Strömungen zulassen und zeigen, inwieweit die Strömung überhaupt ausreichend von den Bojen erfasst werden kann. Als Vergleichsdaten sollen unabhängige Modelle (z.B. Rifugio-Modelle, RIO) herangezogen werden, wo aus Altimetermessungen, Gradiometermessungen und den Tiefenprofilen aus ARGO-Floatern geostrophische Oberflächenströmungen im Nordatlantik abgeleitet werden.
- In einem zweiten Schritt sollen die von dem Surface-Velocity-Programm genutzten Standardverfahren zur Positionsinterpolation und Geschwindigkeitsbestimmung für den nördlichen Atlantik auf Eignung geprüft und gegebenenfalls die Anwendung alternativer Verfahren untersucht werden. Durch Korrelationsanalysen soll damit der geostrophische Anteil im Bewegungsverhalten der Oberflächenbojen untersucht werden.
- In einem letzten Schritt soll eine zeit- und ortsabhängige Modellierung der Strömungen mittels finiter Elemente durchgeführt werden. Eine Trennung in den geostrophischen und ageostrophischen Anteil ist zu untersuchen.

Ziel der Arbeit ist die Analyse des Potentials von Oberflächendrifterdaten zur Strömungsmodellierung. Das Standardverfahren zur Datenanalyse ist nachzuvollziehen und es sind Verfahren zu entwickeln, welche die Datenanalyse für die spezielle Anwendung im Nordatlantik optimieren. Es soll ein funktionsfähiges und gut dokumentiertes Programm zur Verfügung gestellt werden, welches eine Auswertung der Daten für den Nordatlantik ermöglicht.

Aufgabensteller: Prof. Dr.techn. Wolf-Dieter Schuh
Betreuer: Dr.-Ing. Silvia Becker

Tag der Ausgabe: 1. April 2014
Dauer: 6 Monate
Tag der Abgabe:

Bonn, den 17. April 2014



Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht.

Bonn, 29. September 2014

Jessica Franken

Summary

The flow modeling is an important component in understanding the circulation of oceans. By transporting heat, salt and other substances, they control the local and global climate conditions. In recent decades the understanding of surface currents has benefited particularly from the installation of satellites.

Since 1979 standardized surface buoys deliver position data via the Advanced Research and Global Observation Satellite (ARGOS) system using Doppler frequency measurements within the Surface Velocity Program. Currently over 1300 of these buoys can be found in the oceans. Velocities can be derived from their position data, over which a flow modeling is possible.

First, the interpolated position data and the derived velocities by the SVP standard procedures are evaluated mathematically and qualitatively. Furthermore, the influence of high frequency, partly wind driven flow on the SVP pre-processed buoy velocities is analyzed. By this the proportion of unwanted flow in the modeling is visualized and provide information about whether a model is meaningful at all. The final model should no longer contain these effects and represent the geostrophic currents that result from the balance between pressure gradient and Coriolis force. The so-called absolute dynamic topography (ADT) maps the height of the water column above an equipotential surface, called the geoid and is measured by satellites. The velocities of this surface correspond to the geostrophic flow. The analysis of the high-frequency flows and other supplementary effects is done by differentiation with the velocities derived from the ADT of the CNES-CLS09 (RIO et al. 2011), an estimated model made up by measurements of the satellite altimetry and other observation types.

This paper shows a possible way of modeling flow from the raw data of the ARGOS system to the dimensional modeling of velocities on a $2^\circ \times 2^\circ$ -grid in the North Atlantic. For smoothing the raw data and for determining velocities, a method is introduced, which is based on Multivariate Adaptive Regression Splines (MARS) by FRIEDMAN (1991) and enables a parametrisation via cubic B-splines, which are best limited in their oscillation in accordance to the statistical test procedure. This alternative method is robust against buoys noisy traces occurring from the ARGOS system.

To represent the geostrophic flow, the wind-driven slip, which distorts the detection of the flow velocity, must be largely eliminated from the velocities of the buoys. Additionally, a model is developed that allows to eliminate the proportion of the Ekman drift from the buoys velocities. It results from the wind stress acting tangential on the ocean surface and a therefrom resulting shear force. Afterwards a spatial model of the flow with respect to the geostrophic flow is possible. This will be described spatially on the basis of finite elements. For that purpose rectangular elements are spanned by bi-cubic polynomials. The estimated parameters of the rectangular elements define the functional values of the bounding nodes. The resulting velocity fields can be used as an independent data set in combination with other observables, such as the sea surface heights from satellite altimetry, to improve the spatial resolution of the velocity fields.

Zusammenfassung

Die Strömungsmodellierung ist ein wichtiger Bestandteil für das Verständnis der Zirkulationsmechanismen der Ozeane. Diese steuern durch den Transport von Wärme, Salzen und anderen Stoffen, die lokalen und globalen Klimabedingungen. Das Verständnis bezüglich der Oberflächenströme hat in den letzten Jahrzehnten besonders durch die Installation von Satelliten profitiert.

Die innerhalb des Surface Velocity Programms standardisierten Oberflächenbojen liefern seit 1979 Positionsdaten über das satellitengestützte ARGOS-System via Doppler-Frequenzmessungen. Über 1300 dieser Bojen sind zur Zeit in den Weltmeeren zu finden. Aus ihren Positionsdaten können Geschwindigkeiten abgeleitet werden, über die eine Strömungsmodellierung möglich wird.

Zunächst werden diesbezüglich die durch das SVP-Standardverfahren interpolierten Positionsdaten und die daraus abgeleiteten Geschwindigkeiten rechnerisch und qualitativ bewertet. Zudem wird der Einfluss hochfrequenter, zum Teil windgetriebener Strömung auf die durch das SVP vorverarbeiteten Bojen-Geschwindigkeiten analysiert. So wird der Anteil der in der Strömungsmodellierung unerwünschten Strömung visualisiert und Auskunft darüber geben, ob eine Modellierung überhaupt sinnvoll ist. Die endgültige Modellierung soll diese Effekte nicht mehr enthalten und somit die geostrophischen Strömungen abbilden. Sie ergibt sich aus dem Gleichgewicht zwischen Druckgradient und Corioliskraft. Die sogenannte absolute dynamische Topographie (ADT) bildet die Höhe der Wassersäule über einer Äquipotentialfläche, dem Geoid, ab und wird durch Satellitenmessungen bestimmt. Die Analyse der hochfrequenten Strömung und sonstiger zusätzlich zur geostrophischen Strömung in den Bojengeschwindigkeiten enthaltener Effekte erfolgt über die Differenzbildung mit den aus der ADT abgeleiteten Geschwindigkeiten des CNES-CLS09 (RIO et al. 2011), einem durch verschiedene Beobachtungstypen geschätztes Modell.

Diese Arbeit zeigt einen möglichen Weg der Strömungsmodellierung von den Rohdaten des ARGOS-Systems bis hin zur flächenhaften Modellierung der abgeleiteten Geschwindigkeiten auf ein $2^\circ \times 2^\circ$ -Gitter im nördlichen Atlantik. Für die Glättung der Bojenspur und zur Geschwindigkeitsbestimmung wird ein Verfahren eingeführt, welche auf dem multivariaten adaptiven Regressionsspline nach FRIEDMAN (1991) basiert und eine Parametrisierung mittels kubischer B-Splines ermöglicht, die in ihrer Oszillation, den statistischen Prüfverfahren gemäß, bestmöglich eingeschränkt sind.

Um nun die geostrophischen Strömungen abbilden zu können, muss der windgetriebene Schlupf, der die Erfassung der Strömungsgeschwindigkeit verfälscht, weitestgehend aus den Geschwindigkeiten der Bojen herausgerechnet werden. Zudem wird ein Modell entwickelt, dass es ermöglicht, den Anteil des Ekman-Transports aus den Bojengeschwindigkeiten zu eliminieren. Sie resultiert aus der tangential auf die Ozeanoberfläche wirkenden Windschubspannung. Anschließend ist eine räumliche Modellierung von Geschwindigkeitsfeldern der geostrophischen Strömungen möglich. Diese sollen auf Grundlage von finiten Elementen räumlich beschrieben werden. Dazu werden Rechteckelemente durch bikubische Polynome aufgespannt. Die sich ergebenden Geschwindigkeitsfelder können genutzt werden, um als unabhängiger Datensatz in Kombination mit anderen Beobachtungsgrößen, wie den Meereshöhen aus Altimetermessungen, eine verbesserte räumliche Auflösung der Geschwindigkeitsfelder zu ermöglichen.

Inhaltsverzeichnis

1	Einleitung	1
2	Strömung	3
2.1	Geostrophische Strömung	4
2.2	Ageostrophische Strömung	6
2.3	Skalen der Ozeanströmungen	9
3	Oberflächenbojen des Surface Velocity Programms (SVPs)	12
3.1	Datenerfassung	12
3.2	Datenverteilung	14
3.3	Datenformate	15
3.4	Übergang von Roh- zu Kriging-Daten	16
3.5	Berechnung der Geschwindigkeit	24
3.6	Kritische Betrachtung der Kriging-Daten des Surface Velocity Programms	25
4	Vergleich zwischen den aus dem SVP und den aus der absoluten dynamischen Topographie abgeleiteten Geschwindigkeiten	29
4.1	Modell der dynamischen Topographie	29
4.2	Untersuchungsgebiet	30
4.3	Datenhomogenisierung	32
4.4	Varianzanalyse	32
4.5	Empirische Autokorrelationsfunktion	38
5	Alternative Positionsinterpolation und Geschwindigkeitsbestimmung	43
5.1	Stückweise kubische Polynome über trunkeierte Potenzen	43
5.2	Wahl der Knoten	44
5.3	Bestimmung der Zielgrößen und deren Kovarianzinformation	52
5.4	Implementierung	54
5.5	Ergebnisse am synthetischen und praktischen Beispiel	55
5.6	Diskussion	61

6	Reduktion der Bojendaten um windgetriebene Geschwindigkeiten	63
6.1	Der windgetriebene Schlupf	63
6.2	Modellierung des Ekman-Transports	65
7	Strömungsmodellierung über stückweise bikubische Polynome	74
7.1	Stückweise bikubische Polynome	74
7.2	Ergebnisse der Strömungsmodellierung über Oberflächenbojengeschwindigkeiten . . .	77
8	Schlussbemerkungen	80
8.1	Fazit	80
8.2	Ausblick	81
A	Anhang	i
A.1	Redundanz in einer Spline-Approximation	i
B	Bezeichnungen	iii
C	Abkürzungen	v
D	Verzeichnisse	vii
	Abbildungsverzeichnis	ix
	Tabellenverzeichnis	xi
	Algorithmenverzeichnis	xii
	Literatur	xiii

1. Einleitung

Vorrangiges Ziel der Meeresforschung ist es, genaue Kenntnisse über die Zirkulationsmechanismen in den Ozeanen zu erhalten. Diese Mechanismen sind vor allem aus der Strömung und der durch sie geformten Meerestopographie abzuleiten. Oberflächenströme in ozeanischer Umgebung sind von entscheidender Bedeutung, da sie den Antrieb für den Transport von Wärme, Salz und anderen Stoffen über große Entfernungen geben, die sowohl die lokalen als auch globalen Klimabedingungen regeln. Zudem haben sie Einfluss auf die Verteilung lebenserhaltender Ressourcen. Die Abbildung 1.1 zeigt die thermohalinen Wasserkreisläufe, welche gleichbedeutend mit den Zirkulationsmechanismen der Ozeane sind. In den letzten Jahrzehnten hat das Verständnis der Oberflächenströme von durch Satelliten beobachtete Meereshöhen profitiert. Die absoluten dynamischen Topographie (ADT) spiegelt die geostrophischen Strömungen wieder, welche sich durch das sich einstellende Gleichgewicht zwischen Druckgradient und Corioliskraft ergeben. Sie wird in vielen Modellen durch die Satellitenaltimetrie, welche die Höhe der Meeresoberfläche über dem Ellipsoid bestimmt, möglich (ALBERTELLA et al. 2012; ANDERSEN und KNUDSEN 2009). Um die ADT aus diesen Messwerten abzuleiten, muss ein Geoidmodell, also eine Äquipotentialfläche abgezogen werden. Diese liegt meist mit einer geringeren räumlichen Auflösung vor. Ein Geoid beschreibt die Meeresoberfläche in völliger Ruhe, also unter der Annahme, dass keine Strömung existiert. Die Strömungsmodellierung über Bojengeschwindigkeiten benötigt daher kein Geoid-Modell denn aus ihnen kann die geostrophische Strömung direkt abgeleitet werden. Die Geschwindigkeiten der Bojen werden meist in Verbindung mit anderen in-situ Beobachtungen verwendet, um die schlechte räumliche Auflösung der durch Satellitenaltimetrie bestimmten ADT zu verbessern (RIO et al. 2011; MAXIMENKO et al. 2009; RIO und HERNANDEZ 2004). In dieser Arbeit ist die Bojengeschwindigkeit die dominierende Beobachtungsgröße, welche zur Bestimmung der geostrophischen Strömung verwendet wird.

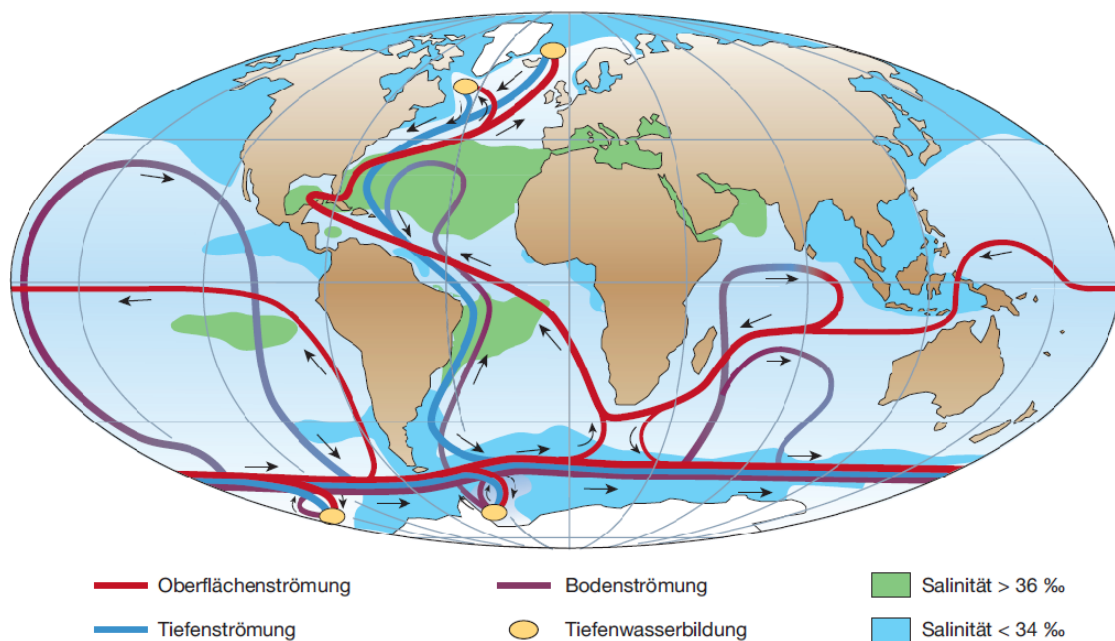


Abbildung 1.1: Das System der globalen Meeresströmungen nach RAHMSTORF (2002)

Vor der Installation von Satelliten wurden Strömungen auf Schiffen durch die Differenz zwischen absoluter Geschwindigkeit und derjenigen durch das Wasser bestimmt. Auch das Tracking von Objekten visuell oder über Funk von der Küste oder Schiffen aus, diente der Strömungsmessung. Seit den 1970er Jahren werden Strömungsmessungen durch Satelliten-Tracking um eine weitere in-situ Erfassungsmöglichkeit ergänzt. Die Strömungen werden durch den Objekt-Transport von „Near Surface Driftern“ (NSD) erfasst. Es erfolgt so eine Lagrange Analyse über diese Oberflächen-Bojen, welche innerhalb des Surface Velocity Programms (SVPs) standardisiert sind. Das SVP, mittlerweile unter dem Namen Global Drifter Programm bekannt, wurde in der Tropical Ocean Global Atmosphere Studie und im World Ocean Circulation Experiment, welche beide Teil des World Climate Research Programms (WCRP) sind, beschlossen. Dort wird seit 1982 die Entwicklung der kostengünstigen, leichten und einfach entwickelten Bojen beziehungsweise deren Standardisierung vorangetrieben. Die Bojen bestehen hauptsächlich aus drei Komponenten, einem Schwimmer an der Oberfläche, einem Stoffzylinder, welcher die Aufgabe eines Treibankers übernimmt und ein diese beiden Komponenten verbindendes Halteseil. Zur Zeit sind etwas über 1300 dieser Bojen in den Ozeanen verteilt. Besonders dicht ist die Verteilung um den Golfstrom. Neben einer satellitengestützten Positions-Erfassung mit Hilfe der Dopplerfrequenz via ARGOS (Advanced Research and Global Observation Satellite), erfolgen in-situ Beobachtungen der Oberflächen-Temperatur, teilweise des Salzgehaltes, sowie des Drucks, der Windgeschwindigkeit und in Ausnahmefällen auch der Ozeanfarbe. Diese Daten werden unter anderem durch das Global Drifter Programm vertrieben und sind frei verfügbar.

In der hier vorgestellten Analyse werden die erfassten Positionsdaten thematisiert, aus welchen die für die Strömungsmodellierung wichtigen Geschwindigkeitsvektoren abgeleitet werden. Besonderer Fokus wird in dieser Arbeit zudem auf den Nord-Atlantik gelegt, auf den die durchgeführten Berechnungen und angesetzten Modelle spezifiziert sind.

Die Positionsdaten von dem ARGOS-System kommend, können mit einer Genauigkeit je Koordinatenwert besser als 150 m gemessen werden, wobei eine Positionsbestimmung spätestens alle vier Stunden stattfindet (BUSHNELL 1996). Diese hohe Genauigkeit wird aber nur bei einer Qualitätsangabe der Klasse 3 gemäß dem ARGOS-System erreicht. Empirische bestimmte Werte weisen allerdings schlechtere Genauigkeiten aus, sodass die Positionen der Roh-Daten bei bester Genauigkeitsklasse eine Standardabweichungen von 240 m beschreiben. Das Verfahren zur Ableitung der Geschwindigkeiten aus diesen Positionsverfahren muss auf diese Genauigkeiten, die im ungünstigsten Fall über 6 km betragen kann, zugeschnitten sein.

Nachfolgend wird der Aufbau der Arbeit erläutert. Nach einem Überblick über die Strömungsmechanismen der Ozeane (Kapitel 2) gliedert sie sich in vier Teilgebiete. Zum Einen beinhaltet sie die Analyse des Ist-Zustandes, die neben der Beschreibung der Datenakquise auch die mathematische Grundlage der von dem SVP bereitgestellten Daten aufzeigt und eine kritische Hinterfragung dieser enthält (Kapitel 3). Zum Zweiten die Entwicklung eines alternativen Verfahrens zur Ableitung der Geschwindigkeiten aus den Bojenspuren, welches vorteilhafte Eigenschaften zur Weiterverarbeitung haben soll, wie die konsistente Ableitung der vollen Kovarianzinformation (Kapitel 5). Zum Dritten die Analyse von Effekten, die aus den Bojengeschwindigkeiten eliminiert werden müssen, damit eine Modellierung der geostrophischen Strömung möglich ist (Kapitel 6). Und zuletzt die räumliche Modellierung der Strömung selbst (Kapitel 7).

2. Strömung

Die Ozeane bedecken rund 70 % der Erdoberfläche (STEWART 2008, S. 24). Allein der Atlantik bedeckt nach MENARD und SMITH (1966) eine Fläche von $106.57 \times 10^6 \text{ km}^2$ mit einer minimalen Ost-West-Ausbreitung von 1500 km und einer maximalen Nord-Süd-Ausbreitung von 13 000 km, sowie einer durchschnittlichen Tiefe von 3 bis 4 km. Aufgrund dieser Geometrie lässt sich auf deutlich geringere vertikale als horizontale Geschwindigkeiten schließen. Nach STEWART 2008 entspricht die vertikale Geschwindigkeit über einige hundert Kilometer maximal 1 % der horizontalen Geschwindigkeit.

Die horizontale Strömung ist jedoch im besonderen mit der großräumigen vertikalen Umwälzung der Ozeane verbunden. In höheren Breiten kühlt das Wasser ab, zudem steigt der Salzgehalt. Das dadurch schwerere Wasser sinkt ab und breitet sich von der Antarktis und dem subpolaren Atlantik in die Weltmeere aus. Durch die Oberflächenströmungen wird das absinkende Wasservolumen ersetzt (JELIC 2012, S. 3). Da Strömungen thermo- und hydrostatischen Gesetzmäßigkeiten folgen, ist die Ozeandynamik über die Erhaltung von Masse, Impuls, Drehimpuls und Energie sowie der Newtonsche Mechanik zu beschreiben. Es lassen sich die Impulsgleichungen in kartesischen Koordinaten nach STEWART (2008, S. 109) mit

$$\frac{du}{dt} = -\frac{1}{\rho} \frac{\partial p}{\partial x} + fv + F_x \quad (2.1)$$

$$\frac{dv}{dt} = -\frac{1}{\rho} \frac{\partial p}{\partial y} - fu + F_y \quad (2.2)$$

$$\frac{d\varpi}{dt} = -\frac{1}{\rho} \frac{\partial p}{\partial z} + 2\Omega u \cos \phi - g + F_z \quad (2.3)$$

formulieren. Das totale Differential zur Linken der Gleichungen ergibt sich über

$$\frac{d}{dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + \varpi \frac{\partial}{\partial z}, \quad (2.4)$$

wobei u die Geschwindigkeit in östlicher, v in nördlicher und ϖ in vertikaler Richtung ist. ρ steht für die Dichte, p für den Druck, ϕ für die geozentrische Breite, g für die Schwerebeschleunigung und f für den Coriolisparameter. Der Coriolisparameter bestimmt sich über

$$f = 2\Omega \sin \phi, \quad (2.5)$$

wobei Ω für die Winkelgeschwindigkeit der Erde mit $7.2921 \times 10^{-5} \frac{\text{rad}}{\text{s}}$ steht. Diese Gleichung wird aufgrund der Ergänzung um die Reibungskräfte F_x, F_y, F_z als Navier-Stokes-Gleichung bezeichnet. Bezüglich der vertikalen Änderung der Geschwindigkeit mit der Zeit (Gleichung (2.3)) ist $2\Omega u \cos \phi \ll g$ und somit zu vernachlässigen.

Ohne die Reibungskräfte wird die Impulsgleichung als Eulersche Gleichung bezeichnet. Wird der Ozean als stationär angesehen, sodass sich die Strömungsgeschwindigkeiten nicht mit der Zeit ändern und damit das totale Differential auf der linken Seite Null wird, ergibt sich das geostrophische

Gleichgewicht. Dieses beschreibt die geostrophische Strömung.

Neben der geostrophischen Strömung wird noch die ageostrophische Strömung benannt, wobei erstgenannte die deutlich größeren Wassertransporte verursachen. Die kleineren Terme der Navier-Stokes-Gleichung, deren Vernachlässigung zu dem geostrophischen Gleichgewicht führt, beschreiben die ageostrophischen Strömungen (CUSHMAN-ROISIN und BECKERS 2011, S. 280). Beide Strömungsarten werden in den nachfolgenden Abschnitten näher erläutert.

2.1 Geostrophische Strömung

Die geostrophische Strömung ergibt sich aus dem geostrophischen Gleichgewicht. Dieses setzt den horizontalen Druckgradient in Balance mit der Corioliskraft, welche aus horizontalen Strömen folgt.

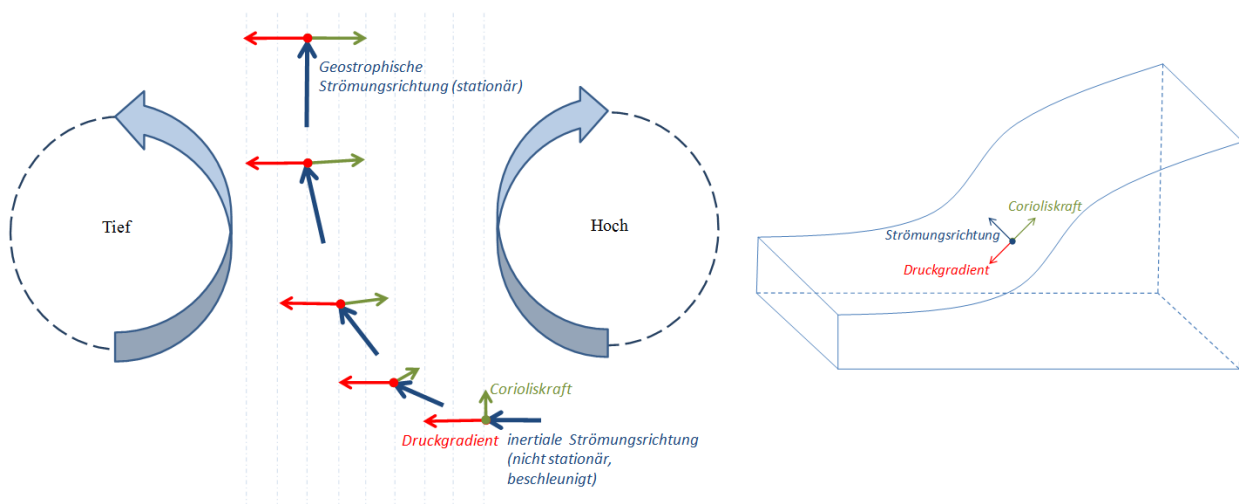


Abbildung 2.1: Schematische Darstellung des sich einstellenden Gleichgewichts zwischen Corioliskraft und Druckgradienten auf der Nordhalbkugel. Die geostrophische Strömungsrichtung stellt sich im stationären Zustand senkrecht zu diesen Kräften.

Würde keine Corioliskraft existieren, so würde das Wasser dem Druckgradienten entsprechend von einem höheren Punkt zu einem tieferen fließen. Die Corioliskraft wirkt auf der Nordhalbkugel nach rechts, sodass die Strömungsrichtung abgelenkt wird (s. Abbildung 2.1). Dadurch ergibt sich um Tiefdruckgebiete eine gegen den Uhrzeigersinn, um Hochdruckgebieten eine mit dem Uhrzeigersinn geostrophische Strömung. Über das geostrophische Gleichgewicht, abgeleitet aus den Gleichungen (2.1) und (2.2) und unter den im letzten Abschnitt erläuterten Vereinfachungen lässt sich die horizontale Strömungsgeschwindigkeit an der Oberfläche wie folgt bestimmen

$$u_g = -\frac{1}{\rho f} \frac{\partial p}{\partial y} = -\frac{g}{f} \frac{\partial \zeta}{\partial y} \quad (2.6)$$

$$v_g = \frac{1}{\rho f} \frac{\partial p}{\partial x} = \frac{g}{f} \frac{\partial \zeta}{\partial x} \quad (2.7)$$

mit den Einträgen des Druckgradienten $\frac{dp}{dy}$ und $\frac{dp}{dx}$ (CUSHMAN-ROISIN und BECKERS 2011, S. 206). Der Druck ergibt sich zu $p(\zeta) = g\rho\zeta$ mit der Höhe ζ der Wassersäule über der Äquipotentialfläche. Die Äquipotentialfläche ist eine Niveaufäche gleichen Schwerepotentials, welche die Meeresoberfläche ohne Strömungen, also in völliger Ruhe abbilden würde. Sie wird Geoid genannt. Da geostrophische Strömungen senkrecht zum Druckgradienten verlaufen, entspricht ∂y der Ableitung in nördlicher Richtung und ∂x der Ableitung in östlicher Richtung. g und ρ gelten bezüglich der Oberflächenströmung in der Horizontalen über die Wassersäule als konstant und sind daher aus dem Differenzial herauszunehmen. So lässt sich die geostrophische Strömung über die Schwerebeschleunigung, den Coriolisparameter und den horizontalen Gradienten der dynamischen Topographie bestimmen. Die absolute dynamische Topographie (ADT) beschreibt die Abweichung der Meeresoberfläche zum Geoid. Sie bildet die stationären Ozeanströmungen ab, sodass aus ihr direkt die geostrophische Strömung abgeleitet werden kann.

Um die Meeresoberfläche beziehungsweise deren Höhe beispielsweise über dem Referenzellipsoid, festhalten zu können, werden unterschiedliche Berechnungsgrößen definiert. Diese Berechnungsgrößen werden in Abbildung 2.2 dargestellt.

Die Höhe der Ozeanoberfläche über der Referenzfläche, einem mathematisch festgelegtem Ellipsoid, wird als Meeresoberflächenhöhe (SSH) bezeichnet. Wird die SSH zeitlich gemittelt so ergibt sich die mittlere Meeresoberflächenhöhe (MSSH). Die Meeresspiegelanomalien (SLA) bezeichnen die Abweichungen der SSH zu der MSSH. Mehr ozeanischer Wärmeinhalt (wärmere Gewässer, eine tiefere Thermokline) zeigt sich durch positive, während sich weniger Wärmeinhalt (kühlere Gewässer, eine flachere Thermokline) durch negative Anomalien zeigt.

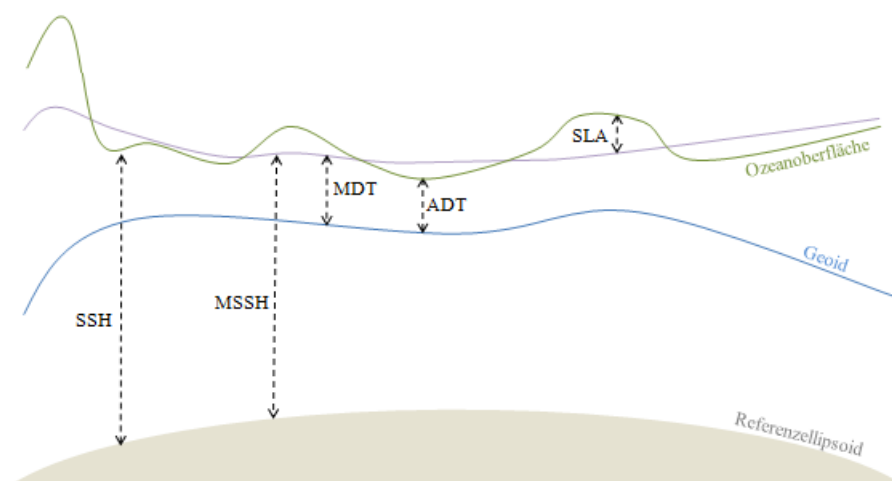


Abbildung 2.2: Berechnungsgrößen zur Höhenkartierung der Ozeanoberfläche

Wie bereits beschrieben, bildet das Geoid die Meeresoberfläche in völliger Ruhe ab. Dort, wo die Anziehungskraft höher ist, wird das Wasser stärker angezogen und es entsteht eine Beule, wo sie schwächer ist, entsteht eine Delle. Das Geoid lässt sich demnach als eine sich ohne Strömung rein hypothetisch ergebende, unbewegte, durch die Landoberflächen durchgehende Meeresoberfläche beschreiben. Da das Geoid eine Fläche gleichen Schwerepotentials abbildet, ergibt sich aufgrund der Variation des Schwerepotentials mit dem Ort eine um ± 100 m ändernde Meereshöhe über dem Ellipsoid.

In der SSH sind demnach vor allem die Höhenunterschiede aufgrund der Geoidhöhen über dem Ellipsoid enthalten. Die Höhenunterschiede, welche sich durch die Strömungen ergeben, sind deutlich klein-skaliger. Um nun Analysen bezüglich der Strömungen durchführen zu können, muss das Geoid aus der SSH herausgerechnet werden. Diese Differenz zwischen SSH und Geoid wird durch die ADT ausgedrückt, welche die geostrophischen Strömungen abbildet. Vielmehr wird die Oberfläche von Variationen im Schwerfeld der Erde geformt, wodurch das Schwerepotential von Ort zu Ort entlang des Ellipsoids variiert. Die mittlere dynamische Topographie (MDT) gibt die zeitgemittelte Topographie mit einer Größenordnung von 1 – 2 m wieder. Mathematisch ausgedrückt bedeutet dies

$$\text{ADT} = \text{SSH} - \text{MSSH} + \text{MDT} = \text{SLA} + \text{MDT}. \quad (2.8)$$

Die ADT kann aus Altimetermessungen und einem Geoidmodell bestimmt werden und zur Berechnung der geostrophischen Strömung nach Gleichung (2.6) und (2.7) für die Höhe der Wassersäule über der Äquipotentialfläche ζ genutzt werden. Vorteil der im nachfolgenden Kapitel beschriebenen Oberflächendrifter beziehungsweise deren in situ Beobachtungen ist, dass daraus direkt die Strömungsgeschwindigkeit bestimmt werden kann. So muss nicht zunächst das Geoid, welches zur Zeit nur mit einer sehr beschränkten Auflösung erfasst ist, von der MSSH subtrahiert werden. Allerdings erfassen die Oberflächenbojen die durch die Summe der geostrophischen und ageostrophischen Strömung beeinflussten Geschwindigkeiten.

2.2 Ageostrophische Strömung

Zu der ageostrophischen Strömung zählen überwiegend windgetriebene Strömungen in Form der Ekman- und Trägheitsströmung. Außerdem zählen neben der Trägheitsströmung andere hauptsächlich hochfrequente Strömungseffekte wie die Gezeitenströmung, Schwerewellen, küstennaher Auftrieb, cyclostrophische Wellen und viele weitere (RIO und HERNANDEZ 2004, S. 7) zu der ageostrophischen Strömung. Die ersten vier genannten Strömungsarten sollen nachfolgend erläutert werden.

2.2.1 Ekman-Strömung

Die Dynamik, welche sich durch Reibung zwischen Ozeanoberfläche und Wind ergibt, bewirkt eine vertikale klein-skalig isotrope turbulente Vermischung. Dieses Vermischen reicht in erster Linie von der Oberfläche bis in eine Tiefe von 30 bis 50 m. Die Wirkspanne wird als Ekman-Schicht (nach Vagn Walfrid Ekman) an der Oberfläche bezeichnet. Die Vermischung bewirkt eine direkte windgetriebene Oberflächenströmung, dessen vertikales Integral orthogonal zur Windrichtung und proportional zur Windstärke ist. Die induzierte Bewegung des Wasservolumens wird Ekman-Transport genannt (OLBERS et al. 2012, S. 441).

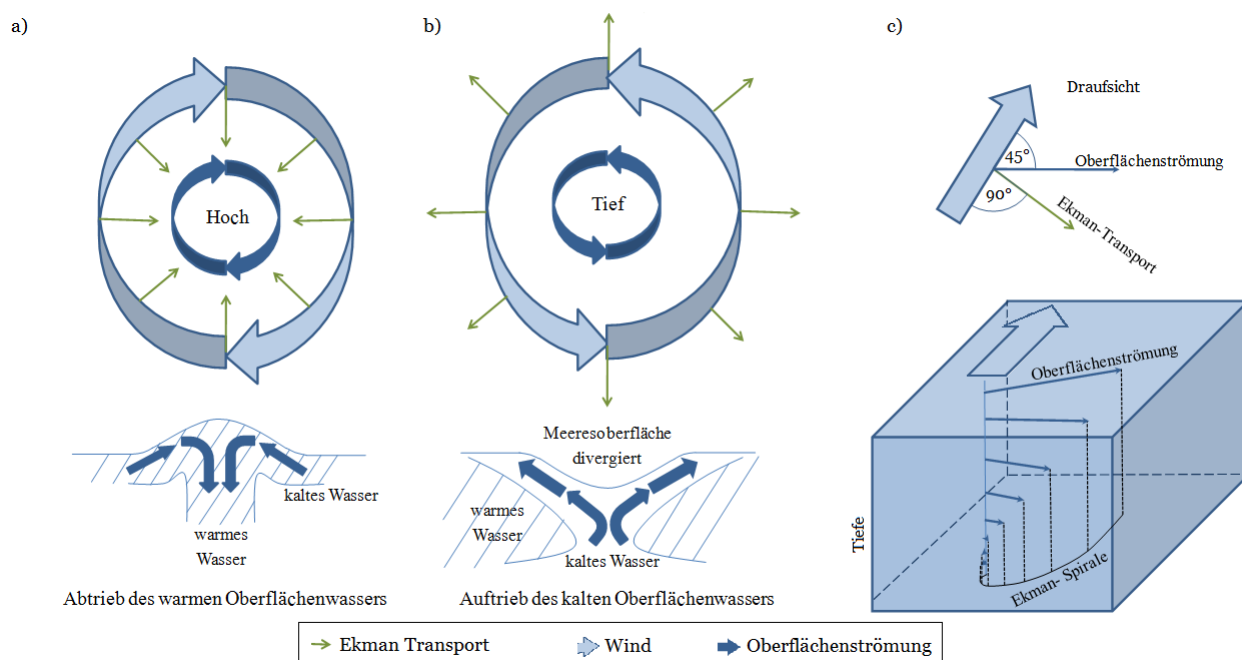


Abbildung 2.3: Schematische Darstellung der windgetriebenen Ozeanströmung in Verbindung mit dem Ekman-Transport für je ein Hoch- (a) und ein Tiefdruckgebiet (b) in der nördlichen Hemisphäre, sowie der Ekman-Spirale (c) (in Anlehnung an CUSHMAN-ROISIN und BECKERS 2011, S.253-254)

Der durch den Westwindgürtel in den mittleren Breiten erzeugte Ekman Transport ist überwiegend äquatorwärts gerichtet, während die subtropischen östlichen oder auch Passatwinde einen polwärts gerichteten Ekman-Transport herbeiführen. In der Mitte des subtropischen Wirbelsystems erzeugt der Ekman-Transport eine Konvergenz, was einen Auftrieb der Meeresoberfläche und eine Abwärtsbewegung der konvergierenden Wassermassen an der unteren Grenze der Ekman-Schicht zur Folge hat. Dieser Mechanismus wird Ekman-Pumpen genannt (CUSHMAN-ROISIN und BECKERS 2011, S. 249). Die Abbildung 2.3 bildet diesen Effekt schematisch ab. Die horizontal übereinander liegenden Dichte-Schichten schließen sich dem Abwärtstrend an und bilden eine Schüssel in dem subtropischen Wirbel. Die daraus resultierenden horizontalen Dichtegradienten führen zu Druckkräften und damit zu einer geostrophischen Zirkulation, in welcher Druck und Corioliskräfte im Gleichgewicht zueinander stehen. Dieser geostrophische Transport unter der Schicht muss das windgetriebene Ekman-Pumpen im stationären Zustand kompensieren. Das Einstellen dieses Gleichgewichts findet mit der Trägheitsperiode

$$T = 2\pi/f \quad (2.9)$$

statt. Die für den Ekman-Transport wichtigen mathematischen Zusammenhänge ergeben sich innerhalb der Impulsgleichung (2.1) und (2.2) durch das Herausnehmen des geostrophischen und des zeitabhängigen Anteils, sodass sich der Anteil des Ekman-Transports wie folgt ausdrücken lässt

$$-fv = F_x \quad (2.10)$$

$$fu = F_y \quad (2.11)$$

Die Abbildung 2.3 zeigt zudem die Ekman-Spirale, welche das vertikale Geschwindigkeitsprofil, welches den Ekman-Transport tiefenabhängig charakterisiert.

2.2.2 Trägheitsströmungen

Der Trägheitsstrom wird durch Trägheitswellen, welche sich durch das Gleichgewicht von Beschleunigungs- und Corioliskraft ergeben, hervorgerufen. Die Wellen werden von plötzlichen Windereignissen hervorgerufen. Im Idealfall bewegen sich die Wasserteilchen in horizontaler Richtung mit der Periode T auf dem Trägheitskreis (s. Gleichung (2.9)). Das spontane Kreisen freier Teilchen mit einer Anfangsgeschwindigkeit in einer drehenden Umgebung (Erde) trägt den Namen Trägheitsschwingung (CUSHMAN-ROISIN und BECKERS 2011, S. 47ff). Aus den Impulsgleichungen (2.1) und (2.2) lässt sich die Trägheitsströmung isolieren, indem die Reibungskräfte und die geostrophische Strömung zu Null gesetzt werden

$$\frac{du}{dt} = fv \quad (2.12)$$

$$\frac{dv}{dt} = -fu. \quad (2.13)$$

2.2.3 Gezeiteneffekte

Sturmfluten können dramatisch werden, wenn sie die Flut überlagern. Fluten sind durch die Anziehungskraft von Mond und Sonne erzwungene Schwerewellen. Gezeiteneffekte sind in der Regel in der Atmosphäre zu vernachlässigen, während Gezeitenströmungen im Ozean in ihrer Größenordnung andere Strömungen übertreffen können (CUSHMAN-ROISIN und BECKERS 2011, S.303).

Im Gegensatz zu windgetriebenen Strömungen welche mit der Tiefe über den Faktor $1/z$ abnehmen, wirken Gezeiteneffekte über die gesamte Wassersäule. Dies führt zu einer zunehmenden Bedeutung der Gezeitenströmung in tieferen Schichten des Ozeans. Die größten Gezeitenströme werden allerdings küstennah beobachtet, wo die Tiefe gering ist. Innerhalb des Ozeans werden die Gezeitenströme in tieferen Regionen erzeugt, wo die Gezeitenkraft auf eine dicke Wasserschicht wirkt und lokal die Oberfläche modifiziert. Die Oberflächenerhöhung wird dann in Küstenregionen und Schelfmeere getragen, wo durch die reduzierte Tiefe die Amplitude verstärkt wird.

Anschaulich entspricht in einem Ozeanbecken mit einer Tiefe von 10 km die Gezeitenkraft einer mit Windgeschwindigkeiten von 75 m/s erzeugten Oberflächenreibung (CUSHMAN-ROISIN und BECKERS 2011, S.308). Zu den ageostrophischen Strömungen gehören nur diejenigen Gezeiteneffekte, welche einen hochfrequenten, also auf wenige Tage beschränkten Einfluss auf die Meereshöhe haben. Die Ozeangezeiten lassen sich in drei Gruppen einteilen. Gelistet werden dazu die täglichen und halbtäglichen Tiden, sowie die langperiodischen Tiden mit halbmonatlichen, monatlichen und halbjährlichen Perioden (DAUBRAWA 2007, S. 15).

2.2.4 Schwerewellen

Schwerewellen, auch interne Wellen genannt, sind an inneren Dichtestufen in den Ozeanen zu finden. Die Dichtestufen ergeben sich über die Temperatur und Salinitätsvariation, was zum Beispiel durch Sonneneinstrahlung oder durch Süßwassereinfluss herbeigeführt wird. Niederfrequente interne

Wellen werden auch als interne Gezeiten bezeichnet. Die Grenzflächen der Dichtestufen liegen in Tiefen von weniger als 100 Metern.

Wenn die Oberfläche des Wassers durch Erwärmung aufgrund der erhöhten Dichte nach oben verschoben wird, zieht sie die Schwerkraft wieder nach unten und es entsteht eine Vertikalgeschwindigkeit, da potentielle in kinetische Energie verwandelt wird. Aufgrund der Trägheit dringt die Oberfläche unter das Niveau des Gleichgewichts und es resultiert eine Schwingung (CUSHMAN-ROISIN und BECKERS 2011, S. 397 ff.).

2.3 Skalen der Ozeanströmungen

Durch die vielen Variationen in den die Strömung beeinflussenden Faktoren, ist es schwierig Größenordnungen bezüglich der Strömungsgeschwindigkeit, die räumliche oder zeitliche Ausdehnung einzelner Strömungsarten aufzulisten. Zu den die Variation beeinflussenden Faktoren zählen beispielsweise die Temperatur, der Salzgehalt, die Meeresbodentopografie, die vertikale Durchmischung, das Verhältnis zu anderen Strömungen, Verdunstung und Niederschlag, sowie Wind- und Druckverhältnisse. CUSHMAN-ROISIN und BECKERS (2011, S. 8 ff.) beschreiben eine vereinfachte Handhabung der Strömungen und ermöglichen so eine ungefähre Angabe bezüglich der Skalen. Sie setzen Relationen zwischen der Rotationsrate der Erde Ω und der Zeit t , die ein Partikel braucht, um mit einer Geschwindigkeit w eine Strecke l zu passieren. Indem sie noch die Dichteschwankung $\Delta\rho$ über die Höhe h , welche sich im Meer vor allem durch Salz- und Temperaturschwankungen ergibt, in dieser Verhältnisrechnung ergänzen, können sie so für den Ozean und die Atmosphäre folgenden Zusammenhang herstellen

$$w \sim \sqrt{\frac{\Delta\rho}{\rho_0}gh}, \quad l \sim \frac{w}{\Omega}, \quad t = \frac{l}{w}, \quad \text{mit } g = 9.81 \frac{\text{m}}{\text{s}^2}. \quad (2.14)$$

Es ergeben sich gesamtheitlich folgende Werte

$$\rho_0 = 1.2 \frac{\text{kg}}{\text{m}^3}, \Delta\rho = 0.03 \frac{\text{kg}}{\text{m}^3} \quad \text{mit } h = 5000 \text{ m für die Atmosphäre} \rightarrow l_A \sim 500 \text{ km}, w_A \sim 30 \frac{\text{m}}{\text{s}},$$

$$\rho_0 = 1028 \frac{\text{kg}}{\text{m}^3}, \Delta\rho = 2 \frac{\text{kg}}{\text{m}^3} \quad \text{mit } h = 1000 \text{ m für den Ozean} \rightarrow l_O \sim 60 \text{ km}, w_O \sim 4 \frac{\text{m}}{\text{s}}.$$

Auffallend ist der große Unterschied zwischen den charakteristischen Geschwindigkeiten, denn die der Atmosphäre w_A ist um mehr als das siebenfache größer als die des Ozeans w_O . Dieser Unterschied hat im besonderen Auswirkungen auf die Schubspannung an der Grenzfläche der beiden Medien. CUSHMAN-ROISIN und BECKERS (2011, S. 15) gibt auch für die einzelnen Strömungsarten charakteristische Längen-, Geschwindigkeits- und Zeitskalen, welche in Tabelle 2.1 aufgelistet sind. Abbildung 2.4 zeigt graphisch die charakteristischen Längen- und Zeitskalen einer alternativen Unterteilung der Strömungen. Zudem ändert sich nach CUSHMAN-ROISIN (1987, S.185) die Geschwindigkeit des geostrophischen Flusses in einer Tiefe von 50 m über eine Strecke von 500 km ungefähr $1 \frac{\text{cm}}{\text{s}}$ pro Jahr.

Phänomen	Skalen		
	Länge l	Geschwindigkeit w	Zeit t
Mikroturbulenz	1-100 m	1-10 $\frac{\text{cm}}{\text{s}}$	10-100 s
interne Wellen	1-20 km	5-50 $\frac{\text{cm}}{\text{s}}$	Minuten bis Stunden
Gezeiten	Ozeangröße	1-100 $\frac{\text{m}}{\text{s}}$	Stunden
Küstennaher Auftrieb	1-10 km	0.1-1 $\frac{\text{m}}{\text{s}}$	mehrere Tage
Fronten	1-20 km	0.5-5 $\frac{\text{m}}{\text{s}}$	wenige Tage
Wirbel	5-100 km	0.1-1 $\frac{\text{m}}{\text{s}}$	Tage bis Wochen
Hauptströme	50-500 km	0.5-2 $\frac{\text{m}}{\text{s}}$	Wochen bis Jahreszeiten
Großskalige Wirbel	Ozeangröße	0.01-0.1 $\frac{\text{m}}{\text{s}}$	Jahrzehnte und darüber

Tabelle 2.1: Längen-, Geschwindigkeiten- und Zeitskalen für Strömungsphänomene im Ozean, nach CUSHMAN-ROISIN und BECKERS (2011, S. 15)

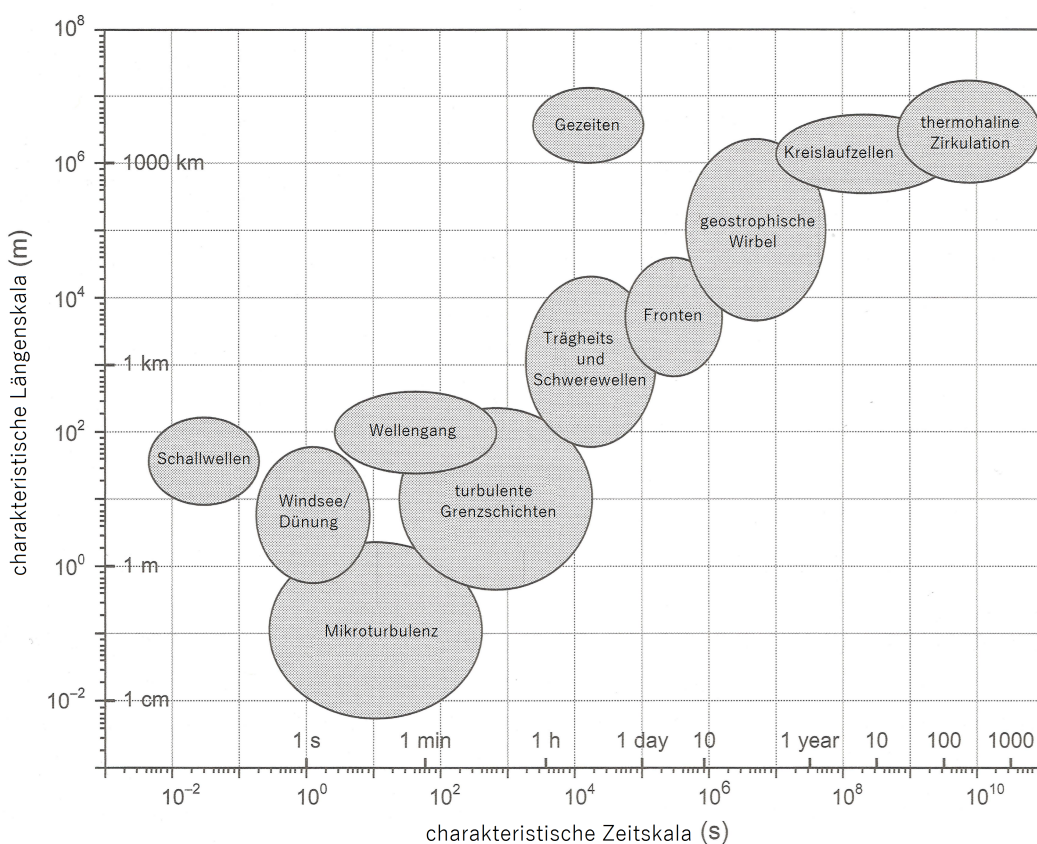


Abbildung 2.4: Längen- und Zeitskala verschiedener Ozeanprozesse nach CUSHMAN-ROISIN und BECKERS (2011)

Bezüglich der windgetriebenen Effekte ist es sehr viel schwieriger eine Zeit- oder Längenskala zu beziffern. Sie sind abhängig von den Bedingungen der Atmosphäre, deren Strömungsmechanismen genauso vielfältig sind wie die der Ozeane. Zudem weist beispielsweise die Ekman-Drift Geschwindigkeiten auf, die mit der Tiefe über wenige Meter stark abnehmen. Einen Eindruck über den zeitgemittelten Einfluss der Ekman-Drift auf die Bojen ermöglicht die Abbildung 2.5. Allgemein wird sie mit einer Stärke von 3% der Windgeschwindigkeit angegeben. Die charakteristische Längenskala der Ekman-Drift für mittlere Breiten ergibt sich nach KANTHA und CLAYSON (2000, S. 69) zu 40 m für den Ozean und hat so nach den Verhältnissen (2.14) eine charakteristische Geschwindigkeit von $3 \frac{\text{cm}}{\text{s}}$.

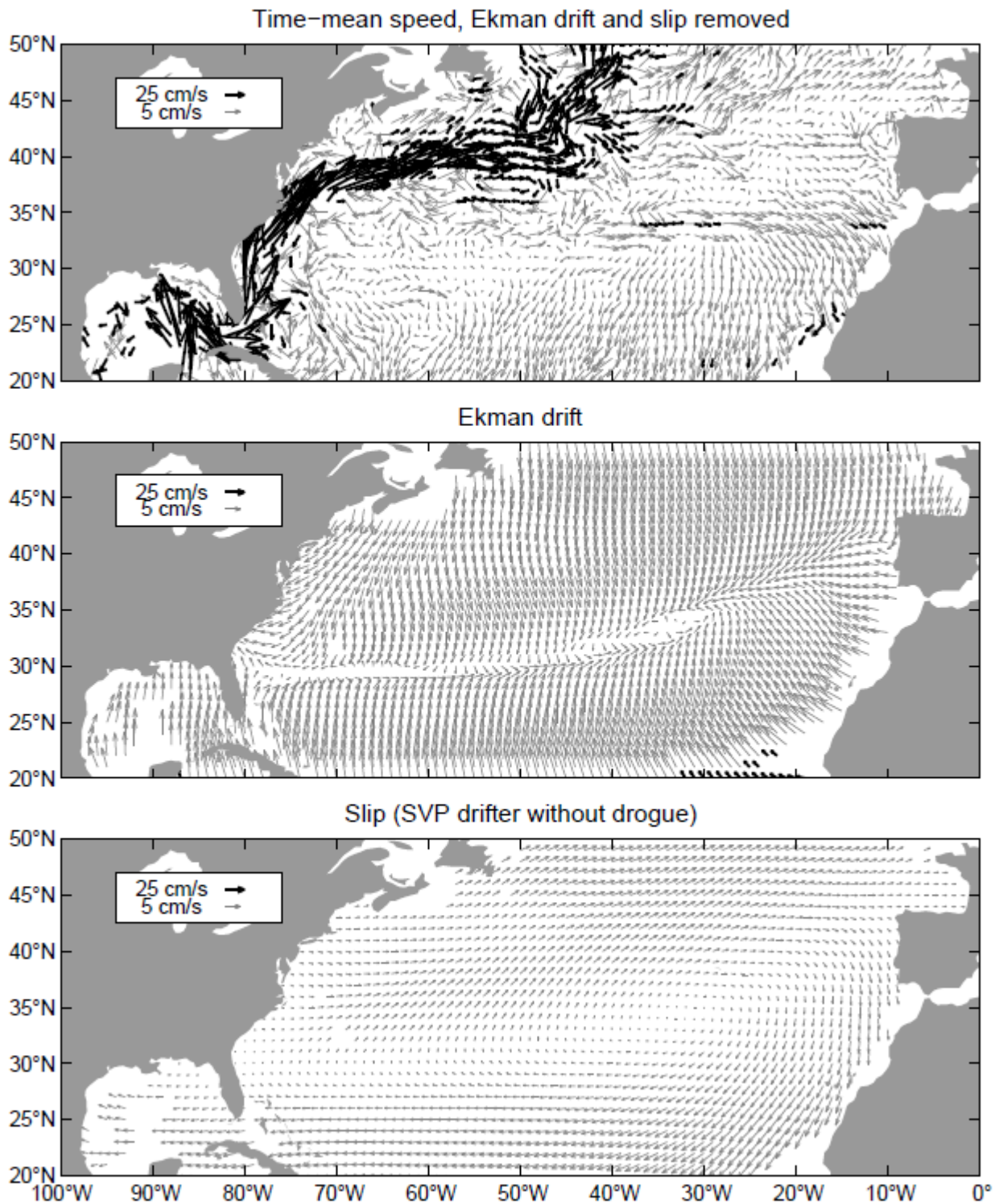


Abbildung 2.5: Oben: Zeitgemittelte Geschwindigkeiten aller SVP-Drifter im subtropischen Nord-Atlantik vom Oktober 1989 bis zum April 2004, die Fehler aufgrund der Ekman-Drift and des windgetriebenen Schlupfes sind entfernt, Mitte: Zeitgemittelte Ekman-Drift aufgrund der wind-getriebenen Konvergenz, welche durch den subtropischen Wirbel angetrieben wird, Unten: Zeitgemittelter windgetriebener Schlupf der SVP-Bojen ohne Treibanker, Abbildung LUMPKIN und PAZOS (2005) entnommen

3. Oberflächenbojen des Surface Velocity Programms (SVPs)

Das SVP treibt seit 1982 die Verbreitung standardisierter Oberflächenbojen, die im englischen als Near Surface Drifter (NSD) bezeichnet werden, voran. Zur Zeit sind weltweit über 1000 dieser Bojen in den Weltmeeren verteilt. Das Verständnis für die Art und Weise der Verarbeitung der von dem SVP vertriebenen Daten ist ein wichtiger Baustein nachfolgender Analysen.

Beschrieben werden neben Verteilung und Anzahl der Drifter im Nordatlantik, die Datenakquise, die zur Verfügung gestellten Datenformate, sowie deren mathematische Grundlage in Hinblick auf die Vorverarbeitung. Letzte Literaturhinweise zur Verarbeitung der Daten sind aus dem Jahr 2005 (LUMPKIN und PAZOS 2005) welche auf eine Auswertung des Jahres 1996 verweisen (HANSEN und POULAIN 1996).

Die Rohdaten werden zunächst, zur Bereinigung von Ausreißern, editiert und anschließend zur besseren Weiterverarbeitung interpoliert. Dieses Kapitel beschreibt die für die Positionsdaten durchgeführte Interpolation auf Grundlage des Ordinary Kriging, die Bestimmung der Kriging-Varianz, welche eine Genauigkeitsangabe der interpolierten Positionen ermöglicht, sowie die verwendeten Variogramme zur Beschreibung des zeitlichen Bewegungsverhaltens der Bojen. Zudem wird die Bestimmung der Bojen-Geschwindigkeit aus den interpolierten Positionen beschrieben.

3.1 Datenerfassung

Es gibt zwei Bauarten der standardisierten Oberflächenbojen, die in dem SVP eingesetzt werden. Der in Abbildung 3.1 dargestellte, sogenannte „original SVP Drifter“ besteht aus einem Schwimmkörper, einem unterirdischen Schwimmer und einem Nylon-Zylinder. Alle Elemente sind über ein Halteseil verbunden. Insgesamt ergibt sich somit eine 45 kg schwere Boje, die sich bis zu einer Tiefe von 18 m erstreckt. Der faltbare Stoff-Zylinder ist, mit einem Durchmesser von 1 m, in einer Tiefe von 15 m zentriert. Er ist ein sogenannter Treibanker, der sich durch kreisrunde durchgängige Aussparungen, welche längs des Zylinders um 90° versetzt sind, auszeichnet. Im Englischen auch als „holey sock drogue,“ bezeichnet, soll er dafür sorgen, dass eine daran befestigte Oberflächenboje mit der Wasserströmung in der Tiefe des Treibankers schwimmt. Die Bojen sollen pseudo-Lagrange Beobachtungen liefern. Dies bedeutet eine nahezu kontinuierliche Positionserfassung nur über die Horizontalbewegung des tiefergelegenen Wasservolumens. Das Präfix pseudo steht für die vernachlässigte vertikale Strömungskomponente, die durch die Bojen nicht erfasst werden können, die aber zu einer regulären Lagrange Beobachtung gehören.

Der Schwimmer ist an der Oberfläche auch Wind und Oberflächenwellen ausgesetzt, die einen windgetriebenen Schlupf, im Englischen mit „Slip“ bezeichnet, bewirken können. Um eine Minimierung dieser Abweichung von der eigentlichen tieferen Wasserströmung zu erreichen, muss der Treibanker

einen Strömungswiderstand herstellen, der deutlich über dem Wind- und Wellenwiderstand der gesamten Boje liegt.

Der mit 20 kg deutlich leichtere „mini SVP Drifter“ besitzt keinen unterirdischen Schwimmer. Zudem ist der Stoff-Zylinder schmaler, das Halteseil dünner und der Schwimmkörper kleiner. Er wurde neu eingeführt, da er auch bei schwierigen Bedingungen von nur einer Person ausgesetzt und wieder eingeholt werden kann.

Die NSD sind mit leichten und kostengünstigen Sendern ausgestattet. Diese senden an einen passierenden ARGOS ein Signal. Der ARGOS berechnet via Doppler-Frequenz-Messung die Position und speichert sie ab. Dazu müssen jeweils mindestens vier Doppler-Verschiebungen gemessen werden, was nur möglich ist, solange die Boje für den Satelliten sichtbar ist. Der Satellit sendet anschließend beim Überflug der Bodenstation die aufgezeichneten Positionen an eine Auswertestation.

Aufgrund der Abhängigkeit von erreichbaren ARGOS sind die Positionsdaten sehr unregelmäßig verteilt. Es treten häufig Daten-Lücken von bis zu zwei Tagen auf (HANSEN und POULAIN 1996, S.907). Die Daten-Lücken können aber auch durch das Untertauchen der Bojen verursacht werden, was bei rauen Umweltbedingungen möglich ist.

Durch eine empirische Messung von 1156 Bojen nach BUSHNELL (1996, S. 135) sind Aussagen bezüglich der Genauigkeit möglich, welche in Tabellen 3.1 zusammengefasst sind. Das ARGOS-System gibt Qualitätsklassen bezüglich der Güte der Positionierung an. Die empirisch bestimmten Werte zeigen jedoch, dass die Standardabweichungen für die Klassen 1 und 2 über denen durch das ARGOS-System ausgegebenen, liegen. Bereits früher, zwischen 1982 und 1989 wurde ein ähnlicher aber nicht so differenzierter Test durchgeführt BITTERMAN et al. 1990. In diesem wurden mit 14000 Bojenpositionen jedoch deutlich mehr Positionen in die empirische Bestimmung der Standardabweichungen einbezogen, was aber zu vergleichbaren Ergebnissen führte.

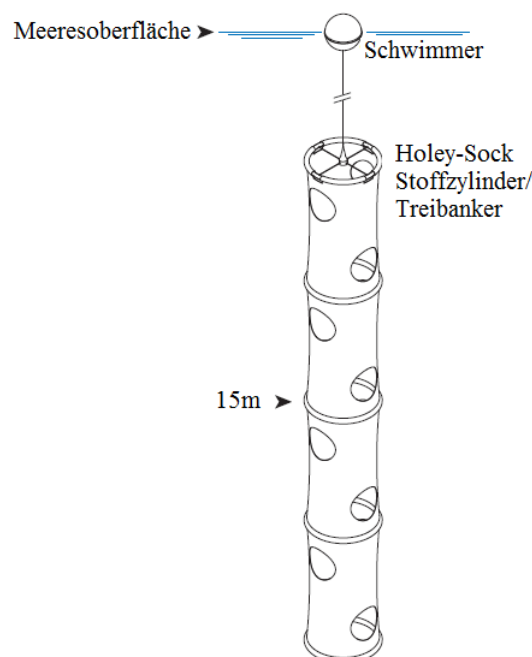


Abbildung 3.1: Aufbau einer Oberflächenboje des Surface Velocity Programms aus PACIFIC GYRE (2013)

ARGOS Qualitätsklasse	Anzahl beobachteter Bojen	Empirisch bestimmte Standardabweichung [m]		Geschätzte Genauigkeit nach ARGOS [m]
		Breite	Länge	
3	411	233	244	150
2	413	377	400	350
1	326	755	978	1000
0	9	1645	6589	>1000

Tabelle 3.1: Empirisch bestimmte und durch das ARGOS-System ausgegebene Genauigkeiten für die Längen und Breitengrade der Bojenpositionen in Metern, Ausschnitt der Daten von BUSHNELL (1996)

Diese Qualitätsklassen des ARGOS-Systems werden nicht in den zur Verfügung gestellten Roh-Daten verzeichnet. Auch HANSEN und POULAIN (1996, S. 901) binden die ausgegebenen Qualitätsangaben nicht in die Vorverarbeitung der Daten ein, stattdessen wird eine Editierung der Daten bevorzugt.

3.2 Datenverteilung

Um einen Anhaltspunkt bezüglich der Verbreitung der Bojen in dem hier untersuchten Nord-Atlantik zu erhalten, wurden bivariate Histogramme angelegt. Dazu werden für jede $1^\circ \times 1^\circ$ -Gitterbox die Anzahl der darin liegenden Bojenpositionen je ausgewählten Zeitraum bestimmt. So entstehen die in Abbildung 3.2 dargestellten Histogramme einmal für je einen Monat (links, mittig) und für ein Jahr (rechts). Die festzustellende inhomogene Abdeckung kommt zum einen durch die begrenzte Anzahl der Bojen, aber auch durch die Aussetzungsstrategien im Ozean und auf Grund der Lagrange-Eigenschaften neigen die Treibbojen dazu, sich in den Konvergenzbereichen zu sammeln und divergente Bereiche, wie der Kap Verde-Bereich, wo Passatwinde eine starke Oberflächen-Divergenz induzieren, zu meiden (RIO und HERNANDEZ 2004, S. 6).

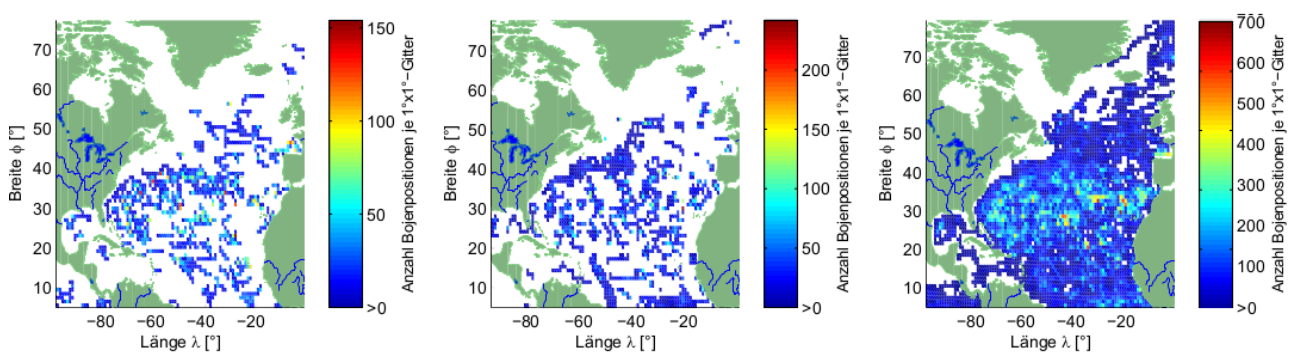


Abbildung 3.2: Anzahl der Bojenpositionen je $1^\circ \times 1^\circ$ -Gitterbox, links: Januar 2009, mittig: Juli 2009, rechts: gesamtes Jahr 2009

Um einen Eindruck über die Verteilung und die Geschwindigkeitsmessungen in Monatsauflösung über ein Jahr zu erhalten, liegen dieser Arbeit zwei Videos bei. Eines zeigt die meridionalen und zonalen Geschwindigkeiten, das andere die Gesamtgeschwindigkeit und ein sich ergänzendes Bild der Gesamtgeschwindigkeit über das gesamte Jahr hinweg.

Das Aussetzen von Bojen begann im Jahre 1979 zunächst hauptsächlich im Pazifik. Es ist daher zu erwarten, dass im Nord-Atlantik erst zu einem späteren Zeitpunkt eine für Analysen hinreichende Anzahl an Bojen vorhanden ist. Die Abbildung 3.3 zeigt, dass bis zum Jahre 1990 die Anzahl der Bojen unter 50 pro Jahr lag.

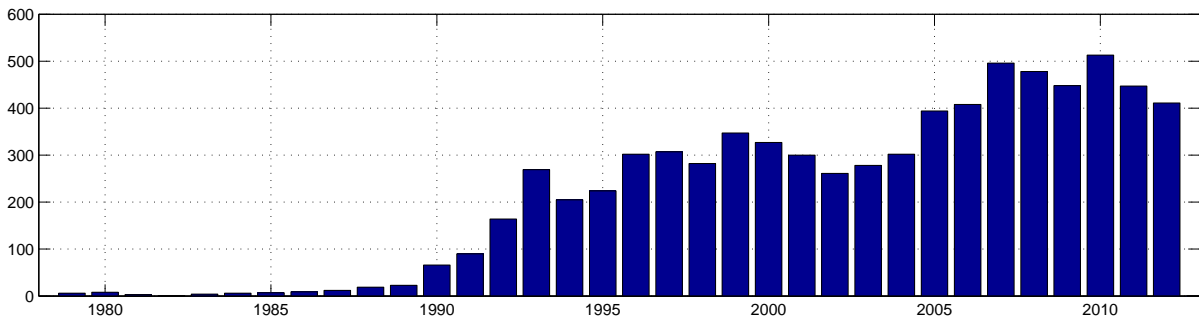


Abbildung 3.3: Anzahl der Bojen pro Jahr im Nordatlantik (1979-2012)

3.3 Datenformate

Sämtliche nachfolgend beschriebenen Datensätze werden von dem kanadischen Fischerei- und Ozeanministerium online zur Verfügung gestellt (www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/drib-bder/svp-vcs/index-eng.asp). Auch die NOAA AOML (<http://www.aoml.noaa.gov/phod/dac/index.php>) stellt die Daten, sowie weiter Informationen bezüglich der Missionen zur Verfügung. Die dort angebotenen Datensätze sind leicht anders untergliedert und können auch für einzelne Drifter abgerufen werden. Ihr Inhalt stimmt allerdings mit dem hier erläuterten Datenformat überein.

Es existieren drei wichtige Datenformate der Bojen-Daten: das Rohdaten-, das Metadaten- und das Kriging-Datenformat. Die Kriging-Daten sind aus den Rohdaten abgeleitet. Zudem existieren sogenannte Positions- und Sensor-Datensätze, welche nur als von Ausreißern befreite verkürzte Rohdatensätze zu verstehen sind.

Die Rohdaten enthalten neben der von ARGOS für die Boje vergebenen Bojen-ID, eine Experimenten-Nummer, eine World Meteorological Organization (WMO) ID, welche die Boje global identifiziert, Datums- und Zeitstempel zu den Positionsangaben in Länge und Breite, die Positionsangaben selbst, sowie eine Genauigkeitsabschätzung derer. Außerdem enthalten sie die gemessenen Sensordaten, den dazugehörigen Datums- und Zeitstempel, die Batterieladung und die Nummer des Stoff-Zylinders.

Die Metadaten sind deutlich überschaubarer und geben jeweils für jede Boje die Bojen-ID, die WMO ID, die Experimenten-Nummer und den Typ der Boje an. Das Datum und die Uhrzeit der Aussetzung und Einholung der Boje mit Positionsangabe sind ebenfalls festgehalten. Auch der Verlust des Stoff-Zylinders ist mit Datums- und Zeitstempel, sowie, falls vorhanden, die Art des Verlusts der gesamten Boje verzeichnet.

Die Kriging-Daten sind anwendungsfreundlich aus den Rohdaten vorverarbeitet. Sie enthalten die Bojen-ID und ein Datum. Da die Positions- und Temperaturangaben der Kriging-Daten auf 6-Stunden-Intervalle interpolierte Rohdaten (s. dazu auch Abschnitt 3.4.2) sind, wird das Datum in Viertel-Tagen hochgezählt und beinhaltet somit bereits einen Zeitstempel. Die interpolierte Länge und Breite und ihre Standardabweichung, sowie die Temperatur werden aufgelistet. Auch die durch den zentralen Differenzenquotienten (s. Abschnitt 3.5) der interpolierten Positionen bestimmten Geschwindigkeiten in Nord- und Ost-Richtung, sowie die sich aus dem Pythagoras errechnete Gesamtgeschwindigkeit sind ebenfalls in einem solchen Datensatz erfasst. Alle Positionsangaben sind geographische Koordinaten bezogen auf das WGS84 (LOPEZ und MALARDÈ 2011, S. 3).

3.4 Übergang von Roh- zu Kriging-Daten

Es ist aus zwei Gründen sinnvoll die Positions-Daten zu interpolieren. Zum Einen entstehen so Positions-Daten aus denen die Geschwindigkeiten über den zentralen Differenzenquotienten abgeleitet werden können und zum Anderen werden Daten zu festen regelmäßigen Zeitpunkten erzeugt. Dadurch sind Analysen von Daten mehrerer Bojen zu bestimmten Zeitpunkten möglich. Unregelmäßig verteilte Daten sind für viele Analysen ungeeignet und liefern meist kein zufriedenstellendes Ergebnis (HANSEN und POULAIN 1996, S. 901).

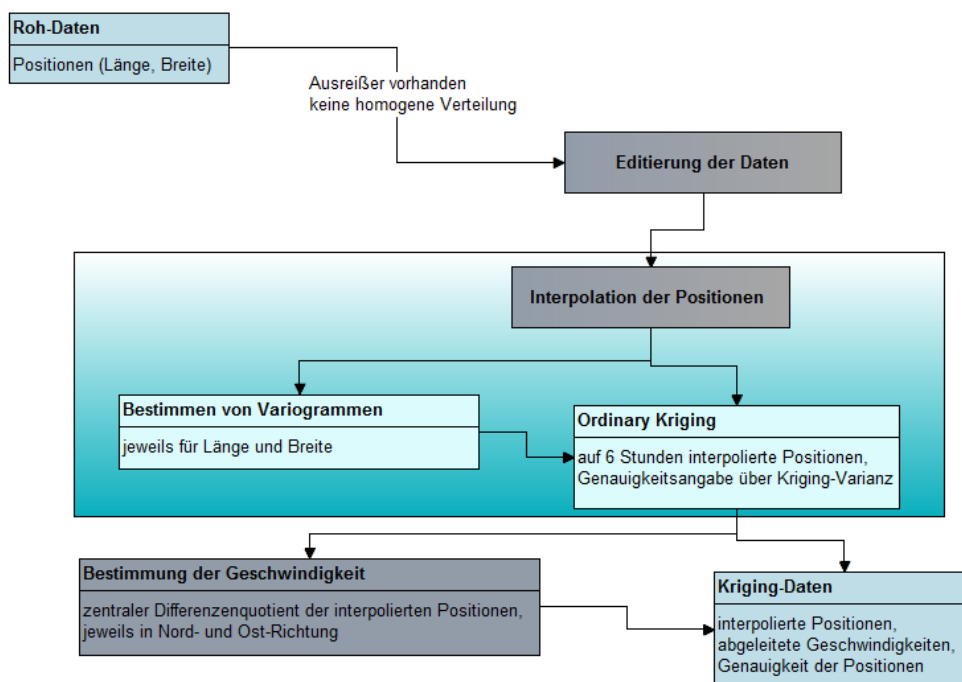


Abbildung 3.4: Ablaufdiagramm zum Übergang von den Roh- zu den Kriging-Daten des SVP

Es wurde die Interpolation auf 6-Stunden-Intervalle gewählt, da diese Intervall-Länge der mittleren Intervall-Länge der Rohdaten in den Tropen ähnelt und anhand der so interpolierten Positionen eine nachvollziehbare Trajektorie erzeugt werden kann (HANSEN und HERMAN 1989, S. 601).

Durch eine zuvor durchgeführte „Editierung“ fließt ein von Ausreißern befreiter Datensatz in die Interpolation ein. Die nachfolgenden Abschnitte beschreiben den Übergang von den Roh- zu den Kriging-Daten. Der Ablauf wird in dem Fließdiagramm in Abbildung 3.4 dargestellt. Für den in diesem Kapitel erläuterten Vergleich werden die Kriging-Daten als vor-verarbeiteter Datensatz verwendet.

3.4.1 Editierung der Daten

Bevor die Daten interpoliert werden, wird von HANSEN und POULAIN 1996 eine „Editierung“ der Rohdaten vorgeschlagen. Durch die Interpolation werden nur zehn Positionsdaten gewichtet eingebracht, sodass einzelne Positionen, die als Ausreißer zu behandeln wären, stark in die interpolierte Position einfließen würden, im Besonderen desto näher der zu interpolierende Punkt dem Ausreißer ist. Zur Interpolation wird daher der editierte Datensatz verwendet.

Das Editieren beinhaltet eine Detektion von Ausreißern und soll so zu einem verlässlicheren Datensatz führen. Betrachtet werden dazu die jeweils für ein Positionspaar über die Zeitinformation gerechneten Geschwindigkeiten, welche einmal über den Vorwärts- und einmal über den Rückwärts-Differenzenquotienten bestimmt werden.

Je Position werden insgesamt vier Kennzeichnungen angebracht, die jeweils mit „Gut“ oder „Schlecht“ ausgedrückt werden. Eine Kennzeichnung für das Vorwärts-Durchlaufen, eine für das Rückwärts-Durchlaufen, eine für die Globale-Blockkennzeichnung und eine finale Kennzeichnung. Zunächst wird der Datensatz **vorwärts** durchlaufen. Mit Hilfe des Vorwärts-Differenzenquotienten wird solange fortlaufend die Geschwindigkeit bestimmt, bis diese eine Geschwindigkeitsgrenze von $257 \frac{\text{cm}}{\text{s}}$ bzw. 5 Knoten überschreitet. Ist das der Fall, so wird die zur Geschwindigkeitsbestimmung neu hinzugekommene Position als „Schlecht“ gekennzeichnet. Die nächste Geschwindigkeit wird von der zuvor noch als „Gut“ gekennzeichneten Position und der nach der als „Schlecht“ gekennzeichneten Position berechnet. So werden alle Positionen mit „Gut“ oder „Schlecht“ gekennzeichnet. Die gleiche Prozedur wird im **rückwärtigen** Durchlauf mit dem Rückwärts-Differenzenquotienten durchgeführt.

Anschließend werden die Positionen **global** gekennzeichnet. Ist die Kennzeichnung in beide Richtungen „Gut“, so ist die globale Kennzeichnung „Gut“, ist eine oder beide Kennzeichnungen „Schlecht“, so lautet die globale Kennzeichnung für diese Position „Schlecht“.

Für jeden global als „Schlecht“ gekennzeichneten Block, wird diejenige Richtung mit mehr, mit „Gut“ gekennzeichneten Positionen zur Entscheidung genutzt (s. Abb. 3.5, oben und mitte). Des weiteren wird diese Richtung als Entscheidungsrichtung betitelt. Ein Block bezeichnet dabei eine einzelne oder aufeinander folgende Positionen, die global mit „Schlecht“ gekennzeichnet sind. Die in dem betrachteten Block in der Entscheidungsrichtung als „Schlecht“ gekennzeichneten Positionen werden als Ausreißer aus dem Datensatz eliminiert beziehungsweise in der **finalen** Kennzeichnung mit „Schlecht“ beschrieben. Ist in beiden Richtungen über den globalen Block die Anzahl der als „Gut“ gekennzeichneten Positionen gleich, so werden alle in dem Block enthaltenen Positionen aus dem Datensatz entfernt (s. Abb. 3.5, unten). Die finale Kennzeichnung ist also „Schlecht“ für den gesamten Block. Durch das Verfahren wird die Ablehnung der gültigen Daten und die Akzeptanz der fehlerhaften Daten minimiert (HANSEN und POULAIN 1996, S. 902).

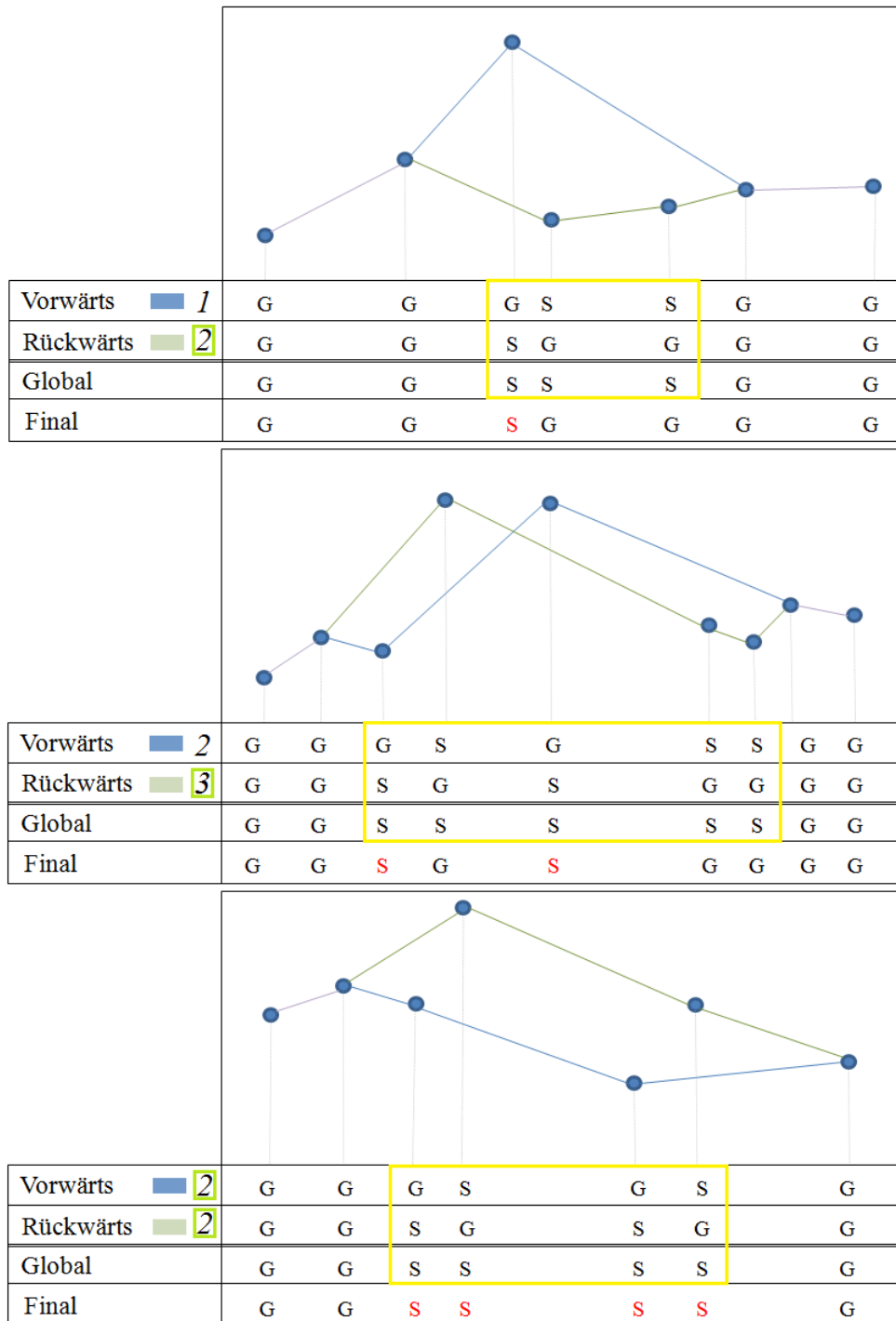


Abbildung 3.5: Beispiele der Editierung nach HANSEN und POULAIN (1996), detektierte Ausreißer: rot, G = Gut, S = Schlecht, Block: gelber Kasten, Entscheidungsrichtung: grüner Kasten.

3.4.2 Interpolation über „Ordinary Kriging“

Das Ordinary Kriging ermöglicht eine beste lineare erwartungstreue Schätzung, es wird kurz als ein BLUE bezeichnet. Ziel ist es, die Kriging-Varianz $\sigma_{\epsilon_s}^2$, welche die Varianz des Interpolationsfehlers beschreibt, zu minimieren. Dies geschieht unter der Bedingung, dass der Erwartungswert des Interpolationsfehlers $\mathbf{E}\{\epsilon_s\}$ Null ist.

Der erwähnte Interpolationsfehler ϵ_s beschreibt die Differenz zwischen dem geschätzten interpolierten Wert \hat{s} und dessen wahren Wert s . Mathematisch ausgedrückt bedeutet das

$$\epsilon_s = \hat{s} - s. \quad (3.1)$$

Als BLUE sind die folgende Eigenschaften zu erfüllen

$$\underbrace{\sigma_{\epsilon_s}^2 \dots Min}_{\text{Best}}, \quad \underbrace{\hat{s} = \mathbf{h}^T \mathbf{l}}_{\text{linear}}, \quad \underbrace{\mathbf{E}\{\epsilon_s\} \doteq 0}_{\text{unbiased}}. \quad (3.2)$$

\mathbf{l} steht für den Beobachtungsvektor der Länge oder Breite. In die Interpolation fließen für diese Anwendung immer nur fünf Positionswerte vor und fünf nach dem zu interpolierenden Punkt ein. Es ergibt sich so ein Beobachtungsvektor der Länge 10 (HANSEN und POULAIN 1996, S.904). Dieser wird mit dem Gewichtsvektor \mathbf{h} multipliziert und anschließend aufsummiert, sodass ein linearer Zusammenhang vorliegt.

Der zufällige Beobachtungsvektor \mathcal{L} kann als Addition der wahren interpolierten Werte \mathcal{S} zu den Zeitpunkten der Beobachtungen und dem Anteil weißen Rauschens \mathcal{N}

$$\mathcal{L} = \mathcal{S} + \mathcal{N} \quad \text{mit } \mathcal{N} \sim \mathcal{N}(0, \sigma_n^2) \quad (3.3)$$

verstanden werden. Die Varianz der Beobachtung setzt sich demnach wie folgt zusammen

$$\sigma_{\mathcal{L}}^2 = \sigma_{\mathcal{S}}^2 + \sigma_n^2. \quad (3.4)$$

Damit der Schätzer erwartungstreu ist, muss eine Bedingung eingeführt werden. Diese ergibt sich, wenn der Erwartungswert des Interpolationsfehlers zu Null gesetzt wird.

$$\begin{aligned} \mathbf{E}\{\epsilon_s\} \doteq 0 &\leftrightarrow \mathbf{E}\{\hat{\mathcal{S}} - \mathcal{S}\} \doteq 0 \\ &\leftrightarrow \mathbf{E}\{\mathbf{h}^T \mathcal{L}\} - \mathbf{E}\{\mathcal{S}\} \doteq 0 \\ &\leftrightarrow \mathbf{h}^T \mathbf{E}\{\mathcal{L}\} - \mathbf{E}\{\mathcal{S}\} \doteq 0 \\ &\leftrightarrow \mathbf{h}^T \mathbf{1} (\mathbf{E}\{\mathcal{S}\} + \underbrace{\mathbf{E}\{\mathcal{N}\}}_{=0}) - \mathbf{E}\{\mathcal{S}\} \doteq 0 \\ &\leftrightarrow \mathbf{h}^T \mathbf{1} \mathbf{E}\{\mathcal{S}\} - \mathbf{E}\{\mathcal{S}\} \doteq 0 \\ &\leftrightarrow (\mathbf{h}^T \mathbf{1} - 1) \mathbf{E}\{\mathcal{S}\} \doteq 0 \end{aligned} \quad (3.5)$$

Möglich wäre es, den Erwartungswert der wahren Werte $\mathbf{E}\{\mathcal{S}\}$ zu Null zu setzen, dies gilt im allgemeinen als „Simple Kriging“. Das Ordinary Kriging setzt als Bedingung hingegen

$$\mathbf{h}^T \mathbf{1} - 1 \doteq 0 \leftrightarrow \mathbf{h}^T \mathbf{1} \doteq 1. \quad (3.6)$$

Um die Gewichtung \mathbf{h} bestimmen zu können, ist die Lösung der Extremwertaufgabe mit Nebenbedingungen notwendig. Dazu wird zunächst die Varianz des Interpolationsfehlers soweit wie möglich aufgelöst.

$$\begin{aligned} \sigma_{\epsilon_s}^2 &= \mathbf{E}\{(\underbrace{\epsilon_s - \mathbf{E}\{\epsilon_s\}}_{=0})(\underbrace{\epsilon_s - \mathbf{E}\{\epsilon_s\}}_{=0})^T\} \\ &= \mathbf{E}\{(\mathbf{h}^T \mathcal{L} - \mathcal{S})(\mathbf{h}^T \mathcal{L} - \mathcal{S})^T\} \\ &= \mathbf{h}^T \mathbf{E}\{\mathcal{L}^T, \mathcal{L}\} \mathbf{h} - \mathbf{h}^T \mathbf{E}\{\mathcal{L}, \mathcal{S}^T\} \\ &\quad - \mathbf{E}\{\mathcal{L}, \mathcal{S}\} \mathbf{h} + \mathbf{E}\{\mathcal{S}, \mathcal{S}^T\} \\ &= \mathbf{h}^T \Sigma\{\mathcal{L}\} \mathbf{h} - 2\mathbf{h}^T \Sigma\{\mathcal{L}, \mathcal{S}\} + \Sigma\{\mathcal{S}\} \end{aligned} \quad (3.7)$$

Die zu lösende Extremwertaufgabe mit Nebenbedingung formuliert sich wie folgt

$$\begin{aligned} \Phi(\mathbf{h}, k) &= \sigma_{\epsilon_s}^2 - 2k(\mathbf{h}^T \mathbf{1} - 1) \dots Min \\ &= \mathbf{h}^T \Sigma\{\mathcal{L}\} \mathbf{h} - 2\mathbf{h}^T \Sigma\{\mathcal{L}, \mathcal{S}\} \\ &\quad + \Sigma\{\mathcal{S}\} - 2k(\mathbf{h}^T \mathbf{1} - 1) \dots Min \end{aligned} \quad (3.8)$$

mit k als Langrange-Multiplikator. Abgeleitet nach der unbekanntem Gewichtung ergibt sich der folgende Ausdruck

$$\begin{aligned} \frac{\partial \Phi}{\partial \mathbf{h}^T} &= 2\Sigma\{\mathcal{L}\} \mathbf{h} - 2\Sigma\{\mathcal{L}, \mathcal{S}\} - 2k\mathbf{1} \doteq 0 \\ &\leftrightarrow \Sigma\{\mathcal{L}\} \mathbf{h} - \Sigma\{\mathcal{L}, \mathcal{S}\} - k\mathbf{1} \doteq 0 \\ &\leftrightarrow \Sigma\{\mathcal{L}\} \mathbf{h} - k\mathbf{1} \doteq \Sigma\{\mathcal{L}, \mathcal{S}\}. \end{aligned} \quad (3.9)$$

Wird die Extremwertaufgabe nach dem Langrange-Multiplikator abgeleitet, so ergibt sich wieder die Bedingung (3.6). Die Bedingungen können in einem Gleichungssystem formuliert und in Matrixschreibweise dargestellt werden

$$\underbrace{\begin{bmatrix} \Sigma\{\mathcal{L}\} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix}}_{\text{Kriging-Matrix}} \begin{bmatrix} \mathbf{h} \\ -k \end{bmatrix} = \begin{bmatrix} \Sigma\{\mathcal{L}, \mathcal{S}\} \\ 1 \end{bmatrix}. \quad (3.10)$$

Für die Bestimmung der Kriging-Varianz (s. Abschnitt 3.4.3) wird der Langrange-Parameter benötigt. Es ist so nach dem Vektor, welcher die Gewichtung und den Parameter enthält, aufzulösen

$$\begin{bmatrix} \mathbf{h} \\ -k \end{bmatrix} = \begin{bmatrix} \Sigma\{\mathcal{L}\} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma\{\mathcal{L}, \mathcal{S}\} \\ 1 \end{bmatrix}. \quad (3.11)$$

Ziel der Interpolation mit Ordinary Kriging ist der Erhalt der geschätzten interpolierten Koordinate \hat{s} , welche die interpolierte Länge beziehungsweise Breite zu einem bestimmten Zeitpunkt beinhaltet. Dies geschieht, indem der lineare Zusammenhang zwischen Beobachtungsvektor und Interpolationspunkt ergänzt wird. So ist es möglich, den in Gleichung (3.11) berechneten Vektor der Gewichtung und des Lagrange-Parameters einzusetzen

$$\begin{aligned}\hat{s} &= \mathbf{h}^T \mathbf{l} \\ &= \begin{bmatrix} \mathbf{h}^T & -k \end{bmatrix} \begin{bmatrix} \mathbf{l} \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\Sigma}\{\mathcal{L}, \mathcal{S}\} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}\{\mathcal{L}\} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{l} \\ 0 \end{bmatrix}.\end{aligned}\tag{3.12}$$

3.4.3 Die Kriging-Varianz

Um die Genauigkeit der interpolierten Positionsangaben in den Kriging-Daten angeben zu können, müssen die Werte für die zuvor angestrebte minimale Varianz des Interpolationsfehlers berechnet werden. Wird die Bedingung (3.9) nach dem Ausdruck $\boldsymbol{\Sigma}\{\mathcal{L}\}\mathbf{h}$ aufgelöst und in die Gleichung (3.7) eingesetzt, so folgt

$$\begin{aligned}\sigma_{\epsilon_s}^2 &= \mathbf{h}^T (\boldsymbol{\Sigma}\{\mathcal{L}, \mathcal{S}\} + \mathbf{1}k) - 2\mathbf{h}^T \boldsymbol{\Sigma}\{\mathcal{L}, \mathcal{S}\} + \boldsymbol{\Sigma}\{\mathcal{S}\} \\ &= \mathbf{h}^T \boldsymbol{\Sigma}\{\mathcal{L}, \mathcal{S}\} + \mathbf{h}^T \mathbf{1}k - 2\mathbf{h}^T \boldsymbol{\Sigma}\{\mathcal{L}, \mathcal{S}\} + \boldsymbol{\Sigma}\{\mathcal{S}\} \\ &= -\mathbf{h}^T \boldsymbol{\Sigma}\{\mathcal{L}, \mathcal{S}\} + k + \boldsymbol{\Sigma}\{\mathcal{S}\} \\ &= \boldsymbol{\Sigma}\{\mathcal{S}\} - \mathbf{h}^T \boldsymbol{\Sigma}\{\mathcal{L}, \mathcal{S}\} + k.\end{aligned}\tag{3.13}$$

Die Varianz des interpolierten Wertes als Skalar entspricht $\sigma_{\mathcal{S}}^2$. Eingesetzt in Gleichung (3.13) ergibt sich so die Formulierung der Kriging-Varianz über die Autokovarianzmatrix

$$\sigma_{\epsilon_s}^2 = \sigma_{\mathcal{S}}^2 - \mathbf{h}^T \boldsymbol{\Sigma}\{\mathcal{L}, \mathcal{S}\} + k.\tag{3.14}$$

Der Eintrag in einem Variogramm lässt sich von demjenigen in einer Autokovarianzfunktion ableiten. Ein Variogramm bestimmt sich über den halben Erwartungswert der quadratischen Differenz eines Beobachtungspaares. Der Zusammenhang zu der Autokovarianzmatrix kann nach ISAACS und SRIVASTAVA (1989, S. 289) wie folgt hergestellt werden

$$\begin{aligned}\gamma_{ij} &= \frac{1}{2} \mathbf{E}\{[\mathcal{L}_i - \mathcal{L}_j]^2\} \\ &= \frac{1}{2} \mathbf{E}\{\mathcal{L}_i^2\} + \frac{1}{2} \underbrace{\mathbf{E}\{\mathcal{L}_j^2\}}_{=\mathbf{E}\{\mathcal{L}_i^2\}} - \mathbf{E}\{\mathcal{L}_i \cdot \mathcal{L}_j\} \\ &= \mathbf{E}\{\mathcal{L}_i^2\} - \mathbf{E}\{\mathcal{L}_i \cdot \mathcal{L}_j\} \\ &= [\mathbf{E}\{\mathcal{L}_i^2\} - \mathbf{E}\{\mathcal{L}\}^2] - [\mathbf{E}\{\mathcal{L}_i \cdot \mathcal{L}_j\} - \mathbf{E}\{\mathcal{L}\}^2] \\ &= \mathbf{C}_0 - \mathbf{C}_{i,j} \\ &= (\sigma_{\mathcal{S}}^2 + \sigma_n^2) - \mathbf{C}_{i,j}.\end{aligned}\tag{3.15}$$

Die Kovarianzmatrix zwischen den Beobachtungen und den interpolierten Werten ergibt sich demnach mit

$$\begin{aligned} \gamma(\mathcal{L}, \mathcal{S})^2 &= \mathbf{1}(\sigma_S^2 + \sigma_n^2) - \Sigma\{\mathcal{L}, \mathcal{S}\} \\ \leftrightarrow \Sigma\{\mathcal{L}, \mathcal{S}\} &= \mathbf{1}(\sigma_S^2 + \sigma_n^2) - \gamma(\mathcal{L}, \mathcal{S}). \end{aligned} \quad (3.16)$$

Wird der Ausdruck (3.16) in die Gleichung (3.14) eingesetzt, so folgt

$$\begin{aligned} \sigma_{\epsilon_s} &= \sigma_S^2 - \mathbf{h}^T(\mathbf{1}(\sigma_S^2 + \sigma_n^2) - \gamma(\mathcal{L}, \mathcal{S})) + k \\ &= \sigma_S^2 - \mathbf{h}^T\mathbf{1}(\sigma_S^2 + \sigma_n^2) - \mathbf{h}^T\gamma(\mathcal{L}, \mathcal{S}) + k \\ &= \sigma_S^2 - \sigma_S^2 - \sigma_n^2 + \mathbf{h}^T\gamma(\mathcal{L}, \mathcal{S}) + k \\ &= -\sigma_n^2 + \mathbf{h}^T\gamma(\mathcal{L}, \mathcal{S}) + k \\ &= \mathbf{h}^T(\gamma(\mathcal{L}, \mathcal{S}) - \mathbf{1}\sigma_n^2) + k. \end{aligned} \quad (3.17)$$

Da die Variogramme des SVPs bereits um den Nugget-Effekt (σ_n^2) bereinigt sind, ergibt sich als endgültige Formel für die Kriging-Varianz (HANSEN und POULAIN 1996, S. 903)

$$\sigma_{\epsilon_s} = \mathbf{h}^T\tilde{\gamma}(\mathcal{L}, \mathcal{S}) + k. \quad (3.18)$$

3.4.4 Verwendete Variogramme

Um eine Interpolation von Positions-Daten oder aber auch die hier nicht behandelte Interpolation von Temperatur-Werten durchzuführen, sind zunächst Variogramme, welche die zeitliche Abhängigkeiten der Daten voneinander beschreiben, zu erstellen. Für die Interpolation der Daten der Oberflächenbojen wurden dazu Daten im tropischen Pazifik aus dem Zeitraum zwischen 1988 und 1993 ausgewertet (HANSEN und POULAIN 1996, S. 904).

Die empirische Berechnung der Variogrammeinträge zu dem Lag τ erfolgt nach CRESSIE (1991, S.75) über

$$\gamma(\tau) = \frac{1}{2N_\tau} \sum_{i=1}^{N_\tau} (l(t_i) - l(t_i + \Delta t))^2 \quad \text{mit} \quad -6 < \Delta t \leq 6. \quad (3.19)$$

$N(\tau)$ entspricht der Anzahl der pro Lag einfließenden Differenzen. Der Lag ist der zeitliche Abstand zwischen den Beobachtungen und wird hier nur bis zu 6 Tagen ausgewertet, da in der Interpolation keine größeren Datenlücken überbrückt werden. Sollten größere Datenlücken auftreten, so wird die Interpolation an dieser Stelle ausgesetzt. Das Variogramm wird zu Lags in 6-Stunden-Intervallen bestimmt. Die für die Differenzbildung genutzten Beobachtungspaare gehören je zu einer Bojen-Trajektorie. Die Differenz wird nicht Bojen-übergreifend bestimmt.

HANSEN und POULAIN (1996) untersuchten die Unterschiede zwischen regionalisiert eingeteilten Variogrammen. Der Datensatz wurde in drei Zonen eingeteilt. Die erste Zone zwischen 7°N und 23.5°N beschreibt die nördlichen Tropen, die zweite zwischen 3°S und 7°N beschreibt die äquatoriale Region und die dritte zwischen 23.5°S und 3°S beschreibt die südlichen Tropen. Für die asymmetrische meridionale Einteilung findet sich keine Begründung. Sie soll angelehnt an die Untersuchung der kinetischen Energiedichteverteilung nach P. Niiler et al. sein. Allerdings ist die Quelle nicht nachzuvollziehen.

Die angesetzte Funktion lautet

$$\tilde{\gamma}(\tau) = \alpha\tau^\beta. \quad (3.20)$$

Die Parameter α und β wurden zur Länge und Breite jeweils zu jeder Zone bestimmt, indem der Datensatz in 11 sechsmonatige Intervalle zerlegt wurde. Außerdem wurde der empirische Variogramm-Wert bei dem 3- und 6-tägigem Lag abgegriffen. Anschließend wurden die Parameter zu jedem Datensatz über

$$\beta^{(l)} = \frac{1}{\ln 2} \ln \left(\frac{\tilde{\gamma}_6^{(l)}}{\tilde{\gamma}_3^{(l)}} \right), \quad \alpha^{(l)} = \frac{\tilde{\gamma}_3^{(l)}}{3^{\beta^{(l)}}} \quad (3.21)$$

mit $l = 1, 2, \dots, 11$ bestimmt und anschließend der Median jener genommen (HANSEN und POULAIN 1996, S. 906). Setzt man die Parametersätze der jeweiligen Zonen in die Gleichung (3.20) ein, so ergibt sich die Abbildung 3.6.

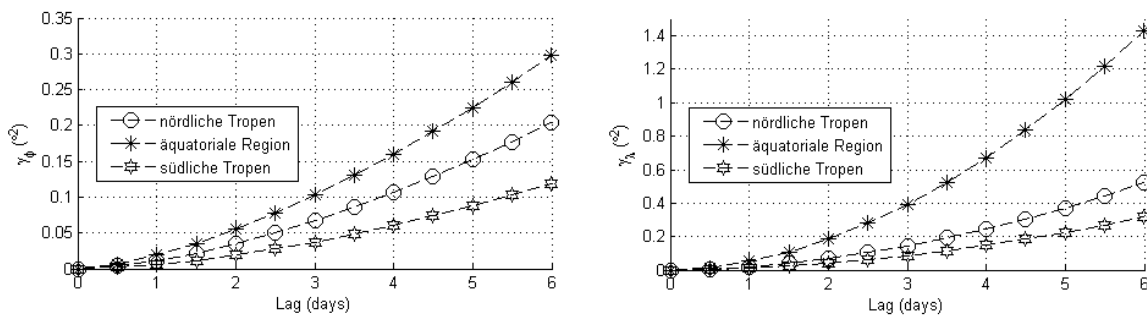


Abbildung 3.6: Funktionale Darstellung der regionalisierten Variogramme für die Länge (links) und die Breite (rechts), bestimmt durch die empirischen Variogramm-Werte (6h-Intervall) des 3- und 6-tägigen Lags, nach HANSEN und POULAIN (1996). Genutzte Daten: 1988-1993, tropischer Pazifik.

Dass der Abgriff der empirischen Variogramm-Werte nur bei dem Lag von 3 und 6 Tagen getätigt wird, begründet sich in der Tatsache, dass dort besonders viele Differenzpaare an Beobachtungen einfließen, was wiederum an dem Beobachtungsplan des WOCE und TOGA SVPs liegt (HANSEN und POULAIN 1996, S. 905–906). Die ermittelten Werte für die Parameter sind HANSEN und POULAIN (1996, S. 906) entnommen und in Tabelle 3.2 aufgelistet.

Die Funktion (3.20) der für das Kriging angesetzten Variogramme folgt nach HANSEN und POULAIN (1996) für entsprechende β bestimmten Bewegungsschemata. Für den Parameter gilt nach den für Bojen sinnvollen physikalischen Bewegungen zunächst

$$1 \leq \beta < 2. \quad (3.22)$$

Für den Fall $\beta = 1$ würden sich die Bojen gemäß der puren Diffusion nach EINSTEIN (1905) wie Atome durch eine Flüssigkeit bewegen, was der Brownschen Bewegung entspricht. Die Grenze $\beta = 2$ ist daher gesetzt, da hierbei von einer geradlinigen Bewegung auszugehen ist, die eine Boje nur in sehr kurzen Zeitintervallen verfolgt. Zudem wäre das Variogramm nicht mehr negativ definit. Ein negativ definites Variogramm ist allerdings für das damit durchzuführende Kriging notwendig. Die hier verwendete Größenordnung liegt, wie aus Tabelle 3.2 ersichtlich, bei etwa 1.6 für die Breite und 1.8 für die Länge.

Nach HANSEN und POULAIN (1996, S. 906) wirkt der Parameter α wie ein Faktor, da der Messfehler gering ist. Daher zeigt die interpolierte Koordinate geringe Empfindlichkeit gegenüber der Variation von α , variiert aber mit β . Die Kriging-Varianz ist nahezu proportional zu α , genauso, wie sie empfindlich gegenüber der Variablen β ist.

HANSEN und POULAIN (1996) weisen auch darauf hin, dass die von Ihnen ausgewerteten Datensätze unkorrelierte zonale und meridionale Geschwindigkeiten beinhalten, für höhere Breiten soll dies aufgrund von Schleifen-Bewegungen in mesoskaligen Wirbeln und Trägheitsschwingungen nicht der Fall sein.

Region	ϕ		λ	
	$10^2\alpha(^{\circ}\phi)^2$	β	$10^2\alpha(^{\circ}\lambda)^2$	β
nördliche Tropen	1.152	1.604	1.912	1.846
äquatoriale Region	1.896	1.538	5.151	1.854
südliche Tropen	0.596	1.673	1.129	1.861

Tabelle 3.2: Regionalisiert berechnete Parameter aus HANSEN und POULAIN (1996), beruhend auf Daten des Pazifiks zwischen 1988-1993

3.5 Berechnung der Geschwindigkeit

Die Geschwindigkeit der Oberflächenbojen wird über den zentralen Differenzen-Quotienten der interpolierten Positionsdaten bestimmt. Die Formeln

$$u^{(i)} = \frac{\hat{s}_{\lambda}^{(i+1)} - \hat{s}_{\lambda}^{(i-1)}}{t^{(i+1)} - t^{(i-1)}} = \frac{\hat{s}_{\lambda}^{(i+1)} - \hat{s}_{\lambda}^{(i-1)}}{12[h]} \quad (3.23)$$

$$v^{(i)} = \frac{\hat{s}_{\phi}^{(i+1)} - \hat{s}_{\phi}^{(i-1)}}{t^{(i+1)} - t^{(i-1)}} = \frac{\hat{s}_{\phi}^{(i+1)} - \hat{s}_{\phi}^{(i-1)}}{12[h]} \quad \text{mit } i = 2, 3, \dots, n-2, n-1, \quad (3.24)$$

bestimmen somit die Geschwindigkeit in zonaler (3.23) und meridionaler (3.24) Richtung (LUMPKIN und PAZOS 2005, S.20). Die Größe n entspricht der Anzahl der gesamten interpolierten Positionen, $t^{(i)}$ steht für den Zeitpunkt, der für die i -te Position gilt. Es ist ersichtlich, dass für die erste und letzte interpolierte Position keine Geschwindigkeit bestimmt werden kann. Diese Werte werden, wie auch die von den Sensoren nicht erfassten Werte, auf 999.999 gesetzt. Die Einheit der Geschwindigkeit in den Kriging-Daten ist $\frac{\text{cm}}{\text{s}}$.

Die Gesamtgeschwindigkeit der Boje berechnet sich über den Pythagoras der Richtungs-Geschwindigkeiten

$$w^{(i)} = \sqrt{u^{(i)2} + v^{(i)2}} \quad \text{mit } i = 2, 3, \dots, n-2, n-1 \quad (3.25)$$

und wird ebenfalls in den Kriging-Daten aufgezeichnet.

3.6 Kritische Betrachtung der Kriging- Daten des Surface Velocity Programms

3.6.1 Die Positionsinterpolation über Ordinary Kriging

Wie in Abschnitt 3.4 beschrieben, werden die durch das ARGOS-System gemessenen Positionen via Ordinary Kriging zu jeder sechsten Stunde interpoliert. Für die Interpolation werden Variogramme verwendet, welche empfindlich bezüglich der untersuchten Breitengrade zu sein scheinen. Eine nähere Untersuchung, welche Auswirkungen eine vorhandene Variation in den geschätzten Parametern der empirisch ermittelten Variogramm-Werte letztendlich auf die Positionsdaten haben, wurde allerdings nicht durchgeführt. Für den Übergang von den Roh- zu den von dem Global Drifter Programm (GDP) ausgegebenen Kriging-Daten werden nur Variogramme angewendet, welche auf Grundlage von Daten aus dem Pazifik berechnet wurden. Für die dort angewendete meridionale Einteilung lässt sich, wie bereits in Abschnitt 3.4.4 erwähnt, keine Begründung finden. Zudem ist keine gemeinsame Auswertung von Längen- und Breitengraden möglich, da diese immer getrennt betrachtet werden. Dies führt dazu, dass die von LUMPKIN und PAZOS (2005) auf Grundlage von HANSEN und POULAIN (1996) angemerkten Korrelation zwischen der zonalen und meridionalen Strömungen, nicht über die richtungsspezifischen Variogramme berücksichtigt werden kann. In hohen Breiten werden diese Geschwindigkeitskomponenten oft durch Schleifenbewegungen in mesoskaligen Wirbel- und Trägheitsschwingungen korreliert. Die von HANSEN und POULAIN (1996) ausgewerteten Daten wiesen keine solche Korrelation auf, da zum Zeitpunkt der Auswertung 1995 keine Datensätze im ausreichendem Umfang für andere Regionen beziehungsweise höhere Breiten vorlagen. Die Daten zeigten daher auch keine fehlerhafte Parametrisierung durch die geschätzten Variogramm-Größen in der von Ihnen durchgeführten Kreuzkorrelation.

3.6.1.1 Bestimmung neuerer Variogramme bezüglich des Atlantiks

Die interpolierten Positionen der Kriging-Daten werden auf Grundlage der Variogramme bestimmt, welche auf Daten des tropischen Pazifiks zwischen 1988 bis 1993 beruhen. Fraglich bleibt, warum bis jetzt keine Auswertung von Variogrammen mit neueren Datensätzen getätigt werden. LUMPKIN und PAZOS (2005, S. 35) verweisen bereits darauf, dass revidierte Interpolationen zu einer verbesserten Kriging-Varianz und Geschwindigkeit und somit zu einer verbesserten Schätzung der Lagrangen Beschleunigung führen würden. Die neueren Datensätze seit 2006 weisen pro Jahr ungefähr so viele Beobachtungen auf, wie der alte Datensatz insgesamt enthält und folgen einem neuen Beobachtungsplan des ARGOS Systems. So fließen mehr Differenzen in die Berechnung ein und die Anzahl der einfließenden Differenzen pro Lag sind außerdem deutlich regelmäßiger verteilt, denn sie zeigen keine deutlichen Spitzen bei 3 und 6 Tagen. Einzig die nördlichen Breiten zwischen 23.5° und 60° , sowie die südlichen zwischen 23.5° und 75° zeigen in der Anzahl der einfließenden Differenzen tägliche Spitzen. Diese sind allerdings auch nicht so ausgeprägt wie die der älteren Datensätze.

Verwendet werden die Datensätze des Atlantiks zwischen 2008 und 2011. Die Funktion (3.20) für das Variogramm wurde auch hier angesetzt und die Parameter nach den Gleichungen (3.21) bestimmt. Es ergeben sich die Parameter der nachfolgenden Tabelle 3.3. Die graphische Darstellung der in der Tabelle aufgelisteten Parameter ist in Abbildung 3.7 zu sehen.

Region	ϕ		λ	
	$10^2\alpha(^{\circ}\phi)^2$	β	$10^2\alpha(^{\circ}\lambda)^2$	β
nördliche Breite	2.089	1.391	3.883	1.545
nördliche Tropen	2.148	1.301	1.813	1.746
äquatoriale Region	2.482	1.484	4.627	1.814
südliche Tropen	0.498	1.617	1.187	1.877
südliche Breite	1.925	1.367	4.591	1.691

Tabelle 3.3: Neu-bestimmte Variogramm-Parameter beruhend auf Daten des Atlantiks zwischen 2008-2011

Die geringste Anzahl an Differenzen fließen bei der Kombination dieser Datensätze bei dem Lag von 6 Tagen für die äquatoriale Region mit $1.17 \cdot 10^6$ ein.

Wie HANSEN und POULAIN (1996, S. 908) auf Grundlage der von SANDERSON und BOOTH (1991) bestimmten Variogramm-Parameter für den nordöstlichen Atlantik behaupten, ist dort in den höheren Breiten eine diffusere Bewegung der Drifter festzustellen. Dies lässt sich auch durch die hier berechneten Größen für β bestätigen. Aus den in Tabelle 3.3 dargestellten Werten für β lässt sich auch für die nördlichen Tropen und die südlichen Breiten dieses diffusere Bewegungsverhalten der Drifter ablesen. Die Werte für α schwanken in ähnlicher Weise wie diejenigen von Tabelle 3.2. Bezüglich α zeigen sich für die äquatoriale Region und die nördlichen Tropen höhere Werte für die Variogramme der Länge und geringere für die Variogramme der Breite. Die Variogramm-Werte für die südlichen Tropen in Länge und Breite sind für beide Parameter in beiden Tabellen vergleichbar.

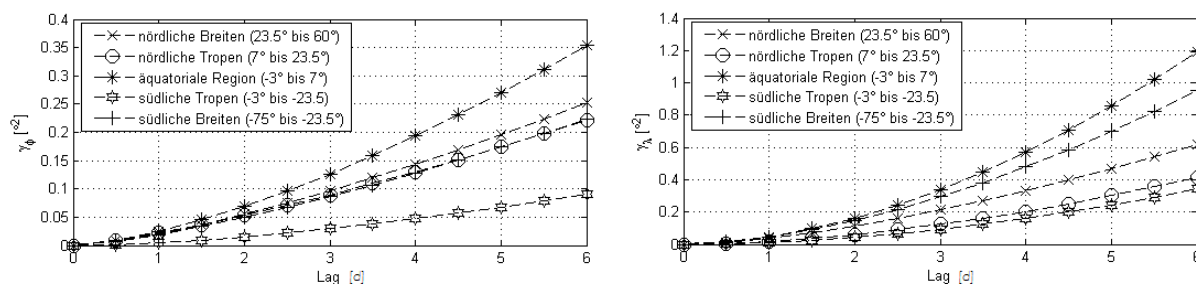


Abbildung 3.7: Funktionale Darstellung der regionalisierten Variogramme für die Länge (oben) und die Breite (unten), bestimmt durch die empirischen Variogramm-Werte (6h-Intervall) des 3- und 6-tägigen Lags. Genutzte Daten: 2008-2011, Atlantik.

Die Variogramm-Werte, außer für die südlichen Tropen, schwanken relativ zu den in Tabelle 3.2 aufgelisteten Werten, wie die Werte dieser Tabelle innerhalb der meridionalen Aufteilung. Ist demnach die meridionale Aufteilung von Bedeutung so ist auch die nach Ozeanen getrennte Betrachtung wichtig.

3.6.2 Die Geschwindigkeitsbestimmung über den zentralen Differenzenquotienten

Die Geschwindigkeit bestimmt sich über ein 12-Stunden-Intervall mit Hilfe des zentralen Differenzenquotienten nach Abschnitt 3.5. Bei einer Geschwindigkeit von 50 cm/s würde so die Boje eine Strecke von 21.6 km zurücklegen. Dass sich die Boje auf einer solchen Strecke geradlinig bewegt, was die Annahme des zentralen Differenzenquotienten ist, bleibt fraglich.

Es lässt sich aus der Kriging-Varianz, welche die Genauigkeitsangabe der Längen- beziehungsweise Breite ist, eine Genauigkeit für die Geschwindigkeiten berechnen. Auch wenn diese die Ungenauigkeit aufgrund der vereinfachten Annahme der geradlinigen Bewegung nicht abbildet, gibt sie dennoch einen Eindruck für die Gültigkeit der über das Kriging und den zentralen Differenzenquotienten berechneten Geschwindigkeiten. Die Varianz für die beiden Geschwindigkeitskomponenten wird über

$$\sigma_u^{(i)2} = \left(\frac{1}{43\,200\text{ s}} \right)^2 \left(\sigma_{\varepsilon_s}^{\phi (i-1)2} + \sigma_{\varepsilon_s}^{\phi (i+1)2} \right) \quad [\text{cm}^2/\text{s}^2], \quad (3.26)$$

$$\sigma_v^{(i)2} = \left(\frac{1}{43\,200\text{ s}} \right)^2 \left(\sigma_{\varepsilon_s}^{\lambda (i-1)2} + \sigma_{\varepsilon_s}^{\lambda (i+1)2} \right) \quad [\text{cm}^2/\text{s}^2] \quad (3.27)$$

gemäß dem Varianzfortpflanzungsgesetz berechnet. Die Varianz entspricht der quadrierten Genauigkeit.

Aus der Abbildung 3.8 ist erkennbar, dass die einfache Standardabweichung der meridionalen Geschwindigkeiten nur durchschnittlich bei gerundet 2 cm/s und für die zonalen Geschwindigkeiten bei 6 cm/s liegen. Insgesamt widersprechen die abgeleiteten Genauigkeiten der Bojengeschwindigkeiten für diesen beispielhaften Zeitraum der Aussage von RIO und HERNANDEZ (2004, S. 6), dass die Genauigkeit der Bojengeschwindigkeit zwischen 2 und 3 cm/s liegt insoweit, dass Geschwindigkeiten mit einer sehr hohen Standardabweichung nicht getrennt betrachtet werden beziehungsweise mit einer entsprechenden Gewichtung in die Analysen eingebracht werden. Wenn die Bojen ohne Treibanker weggelassen werden, so verschiebt sich die durchschnittliche Genauigkeit der meridionalen Geschwindigkeiten zu 8 cm/s. Es hilft also nicht, die Bojen ohne Treibanker aus dem Datensatz zu entfernen, um eine bessere durchschnittliche Genauigkeit zu erhalten. Auch Bojenpositionen, welche mit Treibanker ausgestattet sind, weisen teilweise höhere Standardabweichungen auf. Demnach ist es sinnvoll vor allem die Bojen ohne und die mit Treibanker, welche hohe Standardabweichungen aufweisen aus dem Datensatz zu entfernen. Vor allem die Bojenpositionen ohne Treibanker sind in höheren Standardabweichungen dominant und es liegt ein höherer Slip vor (s. Abschnitt 3.1). Das Weglassen all dieser Bojenpositionen ohne Treibanker würde allerdings zu einem Verlust von circa einem Drittel der Daten führen. Der Slip wird in RIO et al. (2011) zur Generierung des in Abschnitt 4.1 beschriebenen Strömungsfeldes modelliert und von den Geschwindigkeiten der Bojen ohne Treibanker subtrahiert. Trotzdem sollte dort auch eine Sortierung der verwendeten Drifterpositionen nach der Geschwindigkeitsgenauigkeit stattfinden. Geschwindigkeiten mit Genauigkeiten schlechter als 10 cm/s sollten aufgrund des Verhältnisses zu den tatsächlichen Geschwindigkeiten zwischen 0 und 257 cm/s (s. Abschnitt 3.4.1), zur Strömungsmodellierung nur bedingt eingebracht

werden.

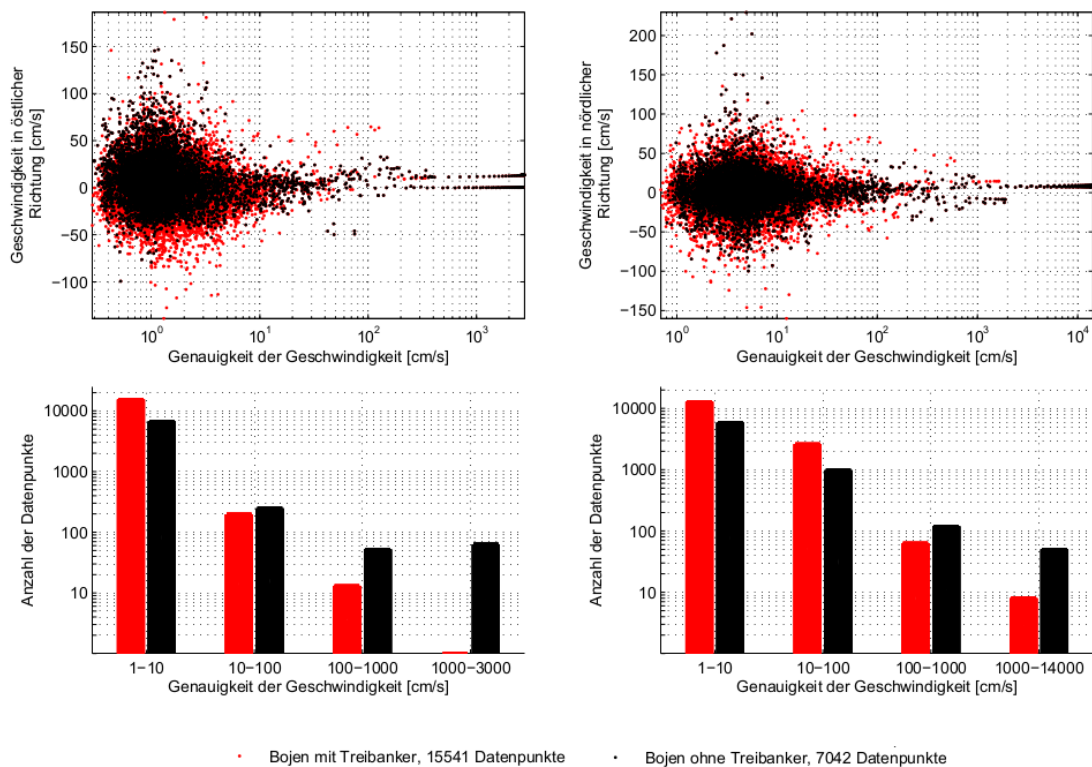


Abbildung 3.8: Oben: Genauigkeit der Geschwindigkeit gegen die Geschwindigkeit selbst aufgetragen, Unten: Histogramme für die Dichteverteilung der Datenpunkte über die Geschwindigkeitsgenauigkeiten für den Januar 2009 im Nord-Atlantik, links: meridionale Geschwindigkeiten, rechts: zonale Geschwindigkeiten, separate Betrachtung der Bojen mit und ohne Treibanker

3.6.3 Fazit bezüglich der vorprozessierten Positions- und Geschwindigkeitsdaten

Gemäß der vorläufigen Analyse müsste die Notwendigkeit geprüft werden, neue Variogramme zu bestimmen und ob diese zur Interpolation angewendet werden sollten. Beispielsweise sollte die von HANSEN und POULAIN (1996) durchgeführte Kreuzvalidierung auch auf Daten des Atlantiks angewendet werden. Sinnvoll wäre sicherlich auch eine Analyse des tatsächlichen Einflusses der Variation der Parameter auf die interpolierte Position und die Kriging-Varianz, sowie eine damit abschätzbare Signifikanz dieser Variation.

Eine Spline-Approximation könnte, da sie keine Variogramme benötigt, aber ebenfalls eine Genauigkeitsabschätzung zulässt, eine denkbare Alternative sein. Zudem wäre die aufwendige Analyse der richtigen meridionalen Aufteilung bezüglich der Variogramme überflüssig, wodurch eventuell eingegangene Interpolationsfehler vermieden werden. Auch eine direkte Schätzung der Geschwindigkeiten parallel zur Interpolation ist möglich, wodurch die stark vereinfachte Geschwindigkeitsbestimmung über den zentralen Differenzenquotienten vermieden wird.

4. Vergleich zwischen SVP- und den aus der ADT abgeleiteten Geschwindigkeiten

Ziel dieses Kapitels soll der Vergleich von Geschwindigkeiten von bereits vorverarbeiteter Oberflächendrifterdaten des SVP mit den aus der absoluten dynamischen Topographie abgeleiteten Geschwindigkeiten sein. Die vorverarbeiteten Oberflächendrifterdaten stellen bereits auf regelmäßige Zeitraster vorprozessierte Geschwindigkeitsfelder dar, deren räumliche Auflösung allerdings sehr inhomogen ist. Hingegen bieten gegitterte ADT- und MDT-Modelle eine sowohl räumlich als auch zeitlich homogene Abdeckung des Nord-Atlantiks. Das CLS09-Modell (RIO et al. 2011) liefert geostrophische Geschwindigkeitsfelder, welche aus der absoluten dynamischen Topographie abgeleitet sind, während die Bojen-Geschwindigkeiten auch ageostrophische Komponenten wie Tiden, Ekman- und Trägheitsströme beinhalten.

Aus dem Vergleich sollen Rückschlüsse über eine mögliche Strömungsmodellierung rein aus den Oberflächendrifterdaten, gewonnen werden.

4.1 Modell der dynamischen Topographie

Die hier als Modell-Geschwindigkeiten bezeichneten Daten entsprechen den aus dem CNES-CLS09 Modell nach RIO et al. (2011) über das geostrophische Gleichgewicht bestimmten geostrophischen Geschwindigkeitsfeldern mit der Aktualität vom 6. Januar 2014. Die geostrophischen Geschwindigkeiten sind in zonaler und meridionaler Richtung gegeben und aus der gegitterten ADT abgeleitet. Diese setzt sich aus der dem Modell entnommenen MDT und der aus der Altimetermessung stammenden SLA zusammen. Das Modell stellt ein zu jedem vollen Tag (0.00 Uhr) auf 1/4-Grad-Gitter gerechnetes Strömungsabbild. Es soll als Referenzdatensatz für den nachfolgenden Vergleich genutzt werden.

Die Daten im netCDF-Format sind durch AVISO über <http://www.avis0.altimetry.fr/en/data/products/sea-surface-height-products/global/madt.html> abrufbar. Eine nähere Beschreibung der Datenverarbeitung findet sich in RIO und HERNANDEZ (2004), RIO et al. (2011) und CNES/AVISO (2014). Für die Erstellung des Modells werden zunächst zwei Methoden verfolgt. Die direkte und die synthetische Methode.

Bei der direkten Methode wird die aus Altimeterdaten bestimmte MSSH so gefiltert, dass sie die gleiche Auflösung wie das verwendete Geoid-Modell besitzt. Anschließend wird von der MSSH die Geoid-Höhe subtrahiert, wodurch eine erste Schätzung für die MDT erhalten wird. Als Geoid-Modell wird das nur aus Satellitendaten (4.5 Jahre GRACE-Daten) gewonnene EIGEN-GRGS.RL02 (BRUNSMAN et al. 2010) Geoid bis zum Grad/Ordnung 160 sphärisch-harmonischer Entwicklung mit einer theoretischen räumlichen Auflösung von 125 km genutzt (RIO et al. 2011). Zudem wurden in sich homogene Datensätze für die MSSH und SLA aus Altimeter-Daten des SSALTO-DUACS Projekts verwendet. Diese sind wöchentlich in „Multimissions-Karten“ in einem 1/3-Grad-Gitter erhältlich.

Die synthetische Methode nutzt in situ Beobachtungen, welche allerdings nicht direkt die ADT abbilden. So werden ergänzend zu den Altimeter-Daten in situ Beobachtungen in Form der im Abschnitt 3.5 erläuterten Geschwindigkeiten der Oberflächenbojen des SVP und der Argo-Floats, zur Ableitung von hydrologischen Tiefenprofilen, genutzt.

Argo-Floats sind autonome batteriebetriebene Tiefenschwimmer zur Temperatur und Salzgehaltprofilierung bis in Tiefen von 2 km, deren Daten vom Coriolis Argo GDAC (<http://www.coriolis.eu.org/>) ausgegeben werden. Die Argo-Floats bilden über Temperatur- und Salzgehaltprofile die barokline Höhe in einer bestimmten Referenztiefe ab (RIO und HERNANDEZ 2004, S. 6). Zur hydrologischen Messung der Argo-Floats ist der barokline Teil der Ozeanzirkulation aufgrund der Dichteveriation von der Profiltiefe bis zum Boden und der barotrope Teil der Ozeanzirkulation zu addieren.

Die Oberflächenbojen messen die totale, also geostrophische und ageostrophische Ozeanzirkulation, sodass aus den Messwerten der Bojen der ageostrophische Anteil herauszurechnen ist. Durch die Einbringung der NSD-Daten in diesen Datensatz sind die Signale des nachfolgenden Vergleichs nicht vollständig unabhängig. Auch für dieses Modell wurden die bereits vorverarbeiteten auf 6-Stundenintervalle ausgegebenen Daten der Bojen, welche mit Treibanker versehen sind, genutzt. Außerdem werden die Differenzen zwischen den Bojen-Geschwindigkeiten und den aus der Summe von MDT des RIO05 Modells (RIO et al. 2005) und der SLA aus Altimeter-Messungen gewonnenen auf die Bojenpositionen zeitlich und räumlich interpolierten Geschwindigkeiten genutzt, um die Parameter des Ekman-Modells zu verbessern. Das dadurch verbesserte Modell wird dann von den Bojen-Geschwindigkeiten subtrahiert (RIO et al. 2011, S. 11). Zusätzlich werden durch Filterung mit einer Filterlänge von drei Tagen, andere ageostrophische Komponenten eliminiert (RIO und HERNANDEZ 2004, S.6). Es werden Windgeschwindigkeits-Daten zur Berechnung der Ekman-Strömung verwendet. Genutzt werden Windkarten der ERA-Interim des European Centre for Medium-Range Weather Forecasts (ECMWF) mit einer drei-stündigen und 80 km räumlichen Auflösung (http://data-portal.ecmwf.int/data/d/interim_daily/).

Durch die angesprochenen Korrekturen der in situ Beobachtungen kann die ADT bestimmt werden und durch Subtraktion mit der SLA das Endprodukt der synthetischen Methode, die auf den Bojenpositionen lokal bestimmte MDT.

Durch eine „Multivariate Objective Analyse“ welche auf einer Interpolationsmethode basiert (RIO und HERNANDEZ 2004, S. 18), werden anschließend die direkt geschätzte und die synthetische MDT zu der oben bereits bezeichneten CNES-CLS09 MDT. Die hier als gegitterter Datensatz verwendeten geostrophische Geschwindigkeitsfelder werden zuletzt über ein 9-Punkte-Matrizen-Verfahren nach ARBIC et al. (2012), welches die Gradienten einer Fläche bestimmt, aus den gegitterten ADT berechnet (CNES/AVISO 2014, S. 17).

4.2 Untersuchungsgebiet

Für den Vergleich wird nur der Nord-Atlantik in einem Bereich von 4.5° bis 80° nördlicher Breite und 0° bis 100° westlicher Länge betrachtet. Neben der gesamtheitlichen Betrachtung wird dieses Gebiet in vier weitere Regionen unterteilt. Zusätzlich wird so die Golf-Strom-Region, die der karibischen

Strömung, die der ruhigeren mittleren Breiten und Teile des Nord-Äquatorstromes betrachtet. Die Aufteilung ist in Abbildung 4.1 dargestellt. Die genaue räumliche Aufteilung ist in Verbindung mit der Kennzeichnung in Abbildung 4.1 in Tabelle 4.1 aufgelistet.

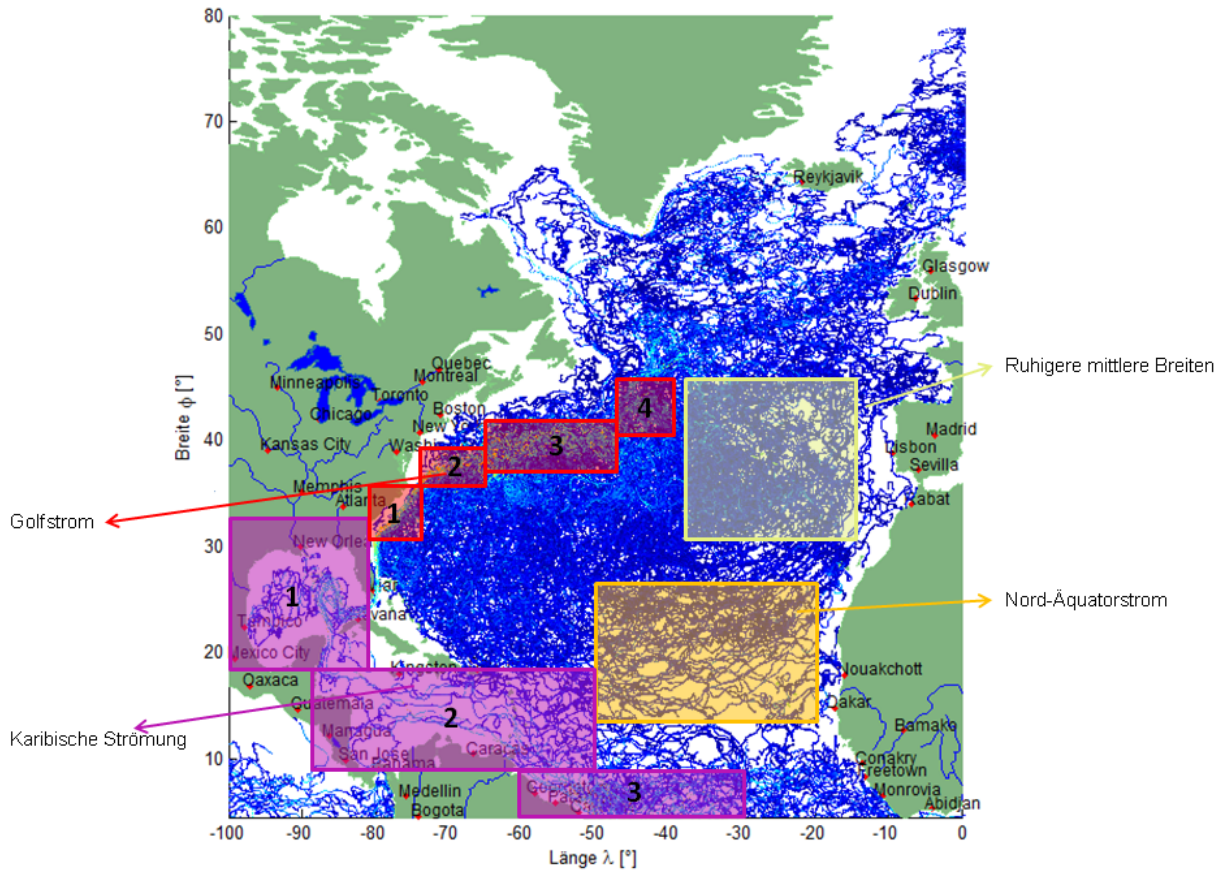


Abbildung 4.1: Aufteilung des Nord-Atlantiks in vier Regionen, blau: Drifterpositionen des Jahres 2008

Bezeichnung der (zusammengesetzten) Begrenzungsbereiche (BB)	Nummer des BB	nördliche Breite von [°] bis [°]		westliche Länge von [°] bis [°]	
Nord-Atlantik		4.5	80.0	100.0	0.0
Golfstrom	1	41.0	46.0	46.0	40.0
	2	38.0	43.0	65.0	46.0
	3	36.0	40.0	75.0	65.0
	4	30.0	35.0	80.0	75.0
ruhigere mittlere Breiten		30.0	46.0	37.0	15.0
karibische Strömung	1	8.0	18.0	90.0	50.0
	2	18.0	32.0	100.0	80.0
	3	4.5	8.0	60.0	30.0
Nord-Äquatorstrom		13.0	25.0	50.0	20.0

Tabelle 4.1: Räumliche Definition des Nord-Atlantik und der ausgewählten Unterregionen

4.3 Datenhomogenisierung

Die Daten der Modell-Geschwindigkeiten werden auf einem $1/4^\circ$ -Gitter ausgegeben. Ein Datensatz mit je einem globalen Gitter für die zonalen, sowie einem für die meridionalen Geschwindigkeiten liegt zu 0.00 Uhr eines jeden Tages vor. Dieses Modell bietet somit eine in zeitlicher und räumlicher Hinsicht homogene Strömungsdarstellung.

Die Daten, welche aus den Positionsdaten der Oberflächenbojen abgeleitet werden, sind entsprechend der Dichte und Verteilung der Bojen zu dem jeweiligen Zeitpunkt sehr inhomogen. Die Positionen beziehungsweise Geschwindigkeiten der Bojen sind alle 6-Stunden gegeben.

Um nun einen Vergleich zu ermöglichen, werden zunächst die Modell-Geschwindigkeiten auf die Positionen der Bojen interpoliert. Dies geschieht, da die Abdeckung durch die Bojen sehr lückenhaft ist, die Modell-Daten hingegen ein dichtes Netz bilden und sich daher besser für eine lineare Interpolation eignen.

Da nicht zu allen Zeitpunkten die Modell-Geschwindigkeiten vorliegen, sondern nur zu den entsprechend vollen Tagen, werden diese nach der räumlichen Interpolation auch zeitlich linear interpoliert.

4.4 Varianzanalyse

Die Varianzanalyse soll einen Eindruck über die Kongruenz der Datensätze geben. Es soll abgeleitet werden, wie viel der in dem durch das CLS09-Modell abgebildeten Strömung durch die Bojen-Geschwindigkeiten der vorverarbeiteten Daten beschrieben wird. Dazu soll die Streuung der Signale in den Vergleich gesetzt werden. Dies lässt sich durch die einfache Varianz der Geschwindigkeiten bestimmen.

Zunächst berechnet sich der Mittelwert der beiden Signale über

$$\bar{V}^{CLS09} = \frac{1}{n} \sum_{i=1}^n V_i^{CLS09}, \quad \bar{V}^{NSD} = \frac{1}{n} \sum_{i=1}^n V_i^{NSD} \quad \text{mit } V \in u, v. \quad (4.1)$$

Die Varianz der einzelnen Signale, also einmal der interpolierten Modell-Geschwindigkeiten σ_{CLS09}^2 und der Bojen-Geschwindigkeiten σ_{NSD}^2 jeweils in zentraler (u) und meridionaler (v) Richtung ergeben sich über

$$\sigma_V^{CLS09^2} = \frac{1}{n-1} \sum_{i=1}^n \left(V_i^{CLS09} - \bar{V}^{CLS09} \right)^2, \quad \sigma_V^{NSD^2} = \frac{1}{n-1} \sum_{i=1}^n \left(V_i^{NSD} - \bar{V}^{NSD} \right)^2, \quad (4.2)$$

wobei n die Anzahl sämtlicher Bojen-Positionen in einem Zeitrahmen sind (KREYSZIG 1991, S. 38). Die Varianz der Signaldifferenz zwischen den Modell- und Bojen-Geschwindigkeiten ergibt sich wie folgt

$$\sigma_V^{CLS09,NSD^2} = \frac{1}{n-1} \sum_{i=1}^n \Delta_V^{(i)^2} \quad \text{mit} \quad \Delta_V^{(i)} = \left(V_i^{NSD} - \bar{V}^{NSD} \right) - \left(V_i^{CLS09} - \bar{V}^{CLS09} \right). \quad (4.3)$$

Diese Größe gibt ein Maß dafür, wie viel Signal durch die interpolierten Modell-Daten aus den Drifter-Datensatz herausgenommen wird. Damit die Größe besser zu interpretieren ist, wird nachfolgend die Standardabweichung angegeben, welche die Wurzel der Varianz ist.

Der Bravais-Pearson-Korrelationskoeffizient als standardisierte Kovarianz wird mit

$$\rho_V = \frac{\frac{1}{n-1} \sum_{i=1}^n (V_i^{NSD} - \bar{V}^{NSD}) \cdot (V_i^{CLS09} - \bar{V}^{CLS09})}{\sigma_V^{CLS09} \cdot \sigma_V^{NSD}} \quad (4.4)$$

(FAHRMEIR et al. 1999, S. 139) jeweils für die zonale und meridionale Richtung bestimmt. Er gibt unabhängig von der Größenordnung der Geschwindigkeiten einen zwischen Null und Eins normierten Wert für die Kongruenz der Signale.

Die für vier Zeitrahmen bestimmten Werte werden in der Tabelle 4.2 aufgelistet. Jeweils zwei Zeitrahmen überspannen jeweils vierzehn aufeinanderfolgende Tage im Januar 2009, einer den gesamten Januar 2009 und einer den Juli 2009. Die Ergebnisse der Tabelle zeigen eine in allen vier Zeitrahmen deutlich höhere Varianz in den Bojen-Geschwindigkeiten. Dies lässt auf Faktoren schließen, die zusätzlich auf die Bewegung der Bojen wirken und die Geschwindigkeit verstärken. Vorstellbar sind dazu der windgetriebene Schlupf (s. Abschnitt 3.1) sowie die nicht eliminierte Ekman-Drift. Zudem sind alle Bojen in die Auswertung aufgenommen. Einige Bojen besitzen jedoch keine Treibanker und werden so deutlich stärker von den lokalen Winden (0.6% der Windgeschwindigkeit) beeinflusst. Der standardisierte NSD mit Treibanker hingegen wird nur mit ungefähr 0.07% der Windgeschwindigkeiten bis zu $10 \frac{m}{s}$ beeinflusst (NILER et al. 1995, POULAIN et al. 2009). Auf die gemessene Geschwindigkeit wirkenden Faktoren, die nicht der Strömungserfassung dienen, sind in den Daten des Modells bereits eliminiert, beziehungsweise nicht vorhanden.

Der Korrelationskoeffizient ist in allen Fällen positiv und verweist damit auf einen gleichsinnigen linearen Zusammenhang. In nahezu allen Regionen ist eine schwache ($\rho_V < 0.5$) bis mittlere ($0.5 \leq \rho_V < 0.8$) Korrelation nach FAHRMEIR et al. (1999, S. 139) festzustellen. Einzig für den Golfstrom ist eine mittlere bis starke ($\rho \geq 0.8$) Korrelation registriert, was darauf verweist, dass desto stärkere Strömungen vorliegen, die Signale eine höhere Korrelation aufweisen.

Im Golfstrom werden allgemein höhere Geschwindigkeiten als in den anderen Regionen erwartet. Dies zeigen auch die Varianzen in beiden Signalen, den interpolierten Modell-Daten und den Bojen-Geschwindigkeiten. Die karibische Strömung weist ebenfalls erhöhte Varianzen auf. Außerdem weisen Golfstrom und karibische Strömung höhere Korrelationen auf. Dies ist durch die hohen Geschwindigkeiten innerhalb des Wasservolumens zu begründen, sodass die Bojen weniger von den Winden beeinflusst werden. Im Juli zeigt sich der Korrelationskoeffizient, im Vergleich zu den im Januar gemessenen Werten, deutlich höher. Dies kann durch geringere Windgeschwindigkeiten, höheren Strömungsgeschwindigkeiten oder einem ruhigeren Wellengang begründet werden, was hierfür aber nicht weiter verfolgt werden soll.

Auch die Genauigkeiten der Bojen-Geschwindigkeiten blieben unbeachtet. Diese liegen jedoch teilweise über der Größenordnung der erfassten Geschwindigkeit selbst. Diese Tatsache und die geringere Auflösung der Modell-Daten auf 1/4-Grad-Gitter, welche nur linear interpoliert werden sind denkbare Begründungen für die hohen Varianzen in den Differenzen der Signale. Diese sind in der

	Zonale Geschwindigkeiten u				Meridional Geschwindigkeiten v			
	σ_u^{CLS09} [$\frac{cm}{s}$]	σ_u^{NSD} [$\frac{cm}{s}$]	ρ_u [-]	$\sigma_u^{CLS09,NSD}$ [$\frac{cm}{s}$]	σ_v^{CLS09} [$\frac{cm}{s}$]	σ_v^{NSD} [$\frac{cm}{s}$]	ρ_v [-]	$\sigma_v^{CLS09,NSD}$ [$\frac{cm}{s}$]
	gesamter Nordatlantik							
01.01 - 14.01	13.24	22.57	0.56	18.75	15.36	20.34	0.64	15.89
15.01 - 28.01	14.50	23.44	0.57	19.26	14.24	19.73	0.63	15.38
01.01 - 31.01	13.84	22.91	0.57	18.92	14.67	20.10	0.62	15.96
01.07 - 31.07	18.19	26.38	0.75	17.40	15.63	21.96	0.69	15.88
	Golfstrom							
01.01 - 14.01	31.42	43.01	0.80	25.84	30.08	41.96	0.74	28.13
15.01 - 28.01	26.64	35.41	0.76	23.03	29.29	36.81	0.81	21.58
01.01 - 31.01	28.57	39.47	0.78	24.70	30.05	39.56	0.77	25.18
01.07 - 31.07	39.01	49.02	0.85	26.13	33.94	41.57	0.81	24.56
	Ruhigere mittlere Breiten							
01.01 - 14.01	8.20	14.26	0.50	12.41	9.37	14.17	0.55	11.92
15.01 - 28.01	7.14	16.24	0.53	13.86	7.44	13.07	0.43	11.97
01.01 - 31.01	7.90	16.15	0.53	13.67	8.31	14.55	0.46	12.99
01.07 - 31.07	10.35	16.66	0.63	13.00	9.37	14.18	0.58	11.64
	karibische Strömung							
01.01 - 14.01	14.43	21.62	0.59	17.59	15.96	21.28	0.73	14.48
15.01 - 28.01	19.81	31.01	0.65	23.54	17.63	22.01	0.68	16.25
01.01 - 31.01	16.56	25.06	0.62	19.59	267.87	21.69	0.71	15.36
01.07 - 31.07	28.07	41.32	0.76	26.99	561.41	34.67	0.70	24.81
	Nord-Äquatorstrom							
01.01 - 14.01	8.05	13.62	0.40	12.82	5.61	12.29	0.32	11.79
15.01 - 28.01	8.76	13.10	0.49	11.66	5.74	10.81	0.30	10.60
01.01 - 31.01	8.47	13.72	0.45	12.44	5.83	11.70	0.29	11.46
01.07 - 31.07	8.03	13.27	0.60	10.60	8.13	12.06	0.62	9.53

Tabelle 4.2: Varianzen und Korrelationen der Modell- und Bojen-Geschwindigkeitssignale entsprechend der fünf Regionen und vier Zeitrahmen (zwei Zwei-Wochen-Intervalle, darunter jeweils zwei Monats-Intervalle)

Größenordnung der Varianzen der Einzel-Signale. Durch die lineare Interpolation der Modell-Daten in dem 1/4-Grad-Gitter werden die feiner strukturierten Strömungen nicht abgetastet. Dazu wurden Positionen im Golfstrom in einem kleinen Ausschnitt betrachtet (s. Abbildung 4.2). Während die Signale im direkten Vergleich ähnliche Gesamt-Geschwindigkeitsstrukturen aufweisen (links und mittig unten), zeigen sich in dem Differenzbild (rechts unten) dennoch hohe Differenzen zwischen den beiden Signalen. Besonders im nordöstlichen Bereich fällt auf, dass die Modell-Daten nicht so viel Variation in den Geschwindigkeiten aufweisen wie die Bojen-Geschwindigkeiten, wodurch sich dort hohe Differenzen zeigen. Die Differenzen in der Strömungserfassung scheint demnach nicht nur durch zusätzliche Faktoren, welche auf die Bojen wirken, sondern auch an der eingeschränkten Auflösung der Modell-Daten zu liegen. Die Mittelwerte der beiden Signale sollten allerdings trotz dieser Unterschiede bei größeren Datensätzen und eingeschränkten Gebieten nicht unabhängig sein. Dies lässt mit einem statistischen Test, einfach oder einfaktorielle Varianzanalyse genannt, prüfen.

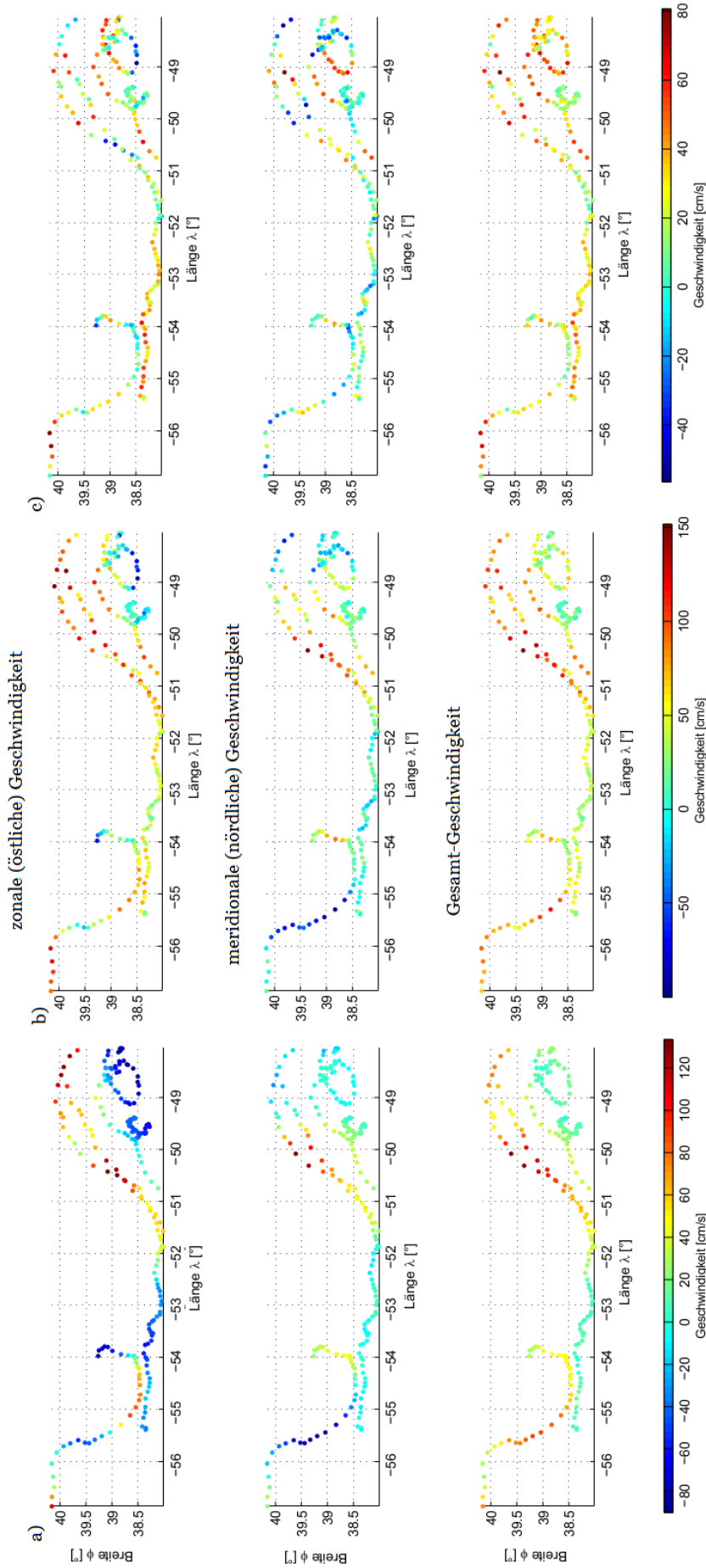


Abbildung 4.2: Zonale, meridionale und Gesamt-Geschwindigkeiten, sowie deren Differenzen für den Juli 2009, Ausschnitt aus der Golfstrom-Region, Spalten: a) lineare interpolierte Geschwindigkeiten des CLS09-Modells, b) Drifter-Geschwindigkeiten, c) Differenz zwischen Spalte a) und b)

Die **einfache Varianzanalyse** (ANOVA) nach KREYSZIG (1991, S. 240), setzt normalverteilte Geschwindigkeiten voraus. Dies ist bei den Geschwindigkeiten nur bei diffusen Verhalten der Fall. Da die Bojen- beziehungsweise Modell-Geschwindigkeiten der Strömung folgen, ist dies nicht mehr gegeben. Dennoch kann der hier angestrebte Vergleich der Mittelwerte der beiden Signale einen Eindruck über die Kongruenz dieser vermitteln. Einfach bedeutet in diesem Fall, dass von einem in einem der beiden Signalen eingebrachten Faktor ausgegangen wird. In diesem Fall wird erwartet, dass die Bojen durch mehr als die Strömung in ihrem Bewegungsmuster beeinflusst werden. Dieser Test ermöglicht es, eine Entscheidung darüber zu fällen, ob die beobachteten Unterschiede in den berechneten Mittelwerten der einzelnen Signale zu groß sind, „um davon auf Unterschiede in den zugehörigen Grundgesamtheiten schließen zu können“ (FAHRMEIR et al. 1999, S. 498).

Hierbei wird nun neben der separaten Auswertung der zonalen und meridionalen Geschwindigkeiten auch die horizontale Gesamtgeschwindigkeit w in den Bojen-Positionen betrachtet. Es gilt also $V \in u, v, w$. Der signalübergreifende Mittelwert bestimmt sich über

$$\bar{V} = \frac{1}{2} \left(\bar{V}^{CLS09} + \bar{V}^{NSD} \right). \quad (4.5)$$

Zudem berechnet sich die „Quadratsumme zwischen den Mittelwerten der Signale“ q_1 und die „Quadratsumme innerhalb der Signale“ q_2 mit

$$q_{1,V} = n \cdot \left(\left(\bar{V}^{CLS09} - \bar{V} \right)^2 + \left(\bar{V}^{NSD} - \bar{V} \right)^2 \right), \quad q_{2,V} = (n-1) \cdot \left(\sigma_V^{CLS09^2} + \sigma_V^{NSD^2} \right).$$

Die Fisher-verteilte Testgröße ergibt sich dann durch

$$t_V = \frac{q_{1,V}/(r-1)}{q_{2,V}/(n-r)} \sim \mathcal{F}_{(r-1),(n-r)} \quad \text{mit } r = 2 \dots \text{Anzahl der Signale.} \quad (4.6)$$

Die Ergebnisse dieser einfachen Varianzanalyse sind in Tabelle 4.3 für die ausgesuchten Zeitrahmen aufgelistet. Liegt die Testgröße betragsmäßig über dem Quantil der Fisher-Verteilung, so ist von einer Unabhängigkeit der Mittelwerte auszugehen. In der Tabelle ist zunächst eine starke Zeitabhängigkeit festzustellen. Bestes Beispiel bildet dazu die karibische Strömung in den meridionalen Geschwindigkeiten. Während der Januar eine Abhängigkeit der Mittelwerte voneinander aufweist, zeigt der Juli diese keineswegs. Auch innerhalb des Januars ist in den ersten beiden Wochen keine Abhängigkeit der Mittelwerte der beiden Signale, also der interpolierten Modell-Geschwindigkeiten und der Bojen-Geschwindigkeiten festzustellen. Die Regionen des Nord-Äquatorstroms und des gesamten Nordatlantik ausgenommen, zeigen die übrigen Regionen in nahezu allen Zeitrahmen Abhängigkeiten der Mittelwerte in den meridionalen Geschwindigkeiten. Dies weist auf Effekte die in dieser Richtung in den Bojengeschwindigkeiten enthalten sind, welche im zweiwöchigen oder Monatsmittel um Null schwanken. Auch möglich ist, dass die Amplitude dieser zusätzlichen Effekte gering ist oder keine zusätzlichen Effekte vorliegen. Unabhängige Mittelwerte in den meridionalen Geschwindigkeiten liegen für den gesamten Nordatlantik vor. Für die karibische Strömung und die mittleren Breiten liegt die Unabhängigkeit nur für den Juli vor.

Die Abhängigkeit der Mittelwerte für die zonalen Geschwindigkeiten im gesamten Nordatlantik lässt

sich durch die in Abbildung 4.3 dargestellten Windgeschwindigkeiten beziehungsweise Richtungen erklären. Werden diese im Mittel zusammengenommen, würden sich die ost- und westwärts gerichteten Vektoren nahezu aufheben, was zu keiner Lageveränderung des Mittelwertes führen würde. Da die Windgeschwindigkeiten größtenteils zonal wirken, sollte sich der dadurch großflächige Unterschied zu den Modell-Daten in einer Autokorrelationsfunktion (s. Abschnitt 4.5) über eine konstante Korrelation wiederfinden lassen. In der Region des Nord-Äquatorstroms liegt in der zonalen Geschwindigkeit die größte Verschiebung des Mittelwertes vor, was sich an dieser Stelle durch hier deutlich größere Windgeschwindigkeiten erklären lässt. Diese sind in den höheren Breiten ausgeprägter und im Januar deutlich stärker als im Juli gleichsinnig gerichtet (s. Abb. 4.3).

Die Gesamtgeschwindigkeiten lassen sich schlechter interpretieren, da sie das Zusammenspiel von meridionaler und zonaler, sowie gleichsinniger und gegensinniger Verschiebung ausdrücken.

Zeitraumen	Anzahl Datenpunkte	Testgrößen			Quantil der Fisher- Verteilung
		zonale	meridionale	gesamte	
		Geschwindigkeiten			
gesamter Nordatlantik					
01.01 - 14.01	9777	1.48	19.22	1594.83	6.64
15.01 - 28.01	9494	3.08	60.19	2054.92	6.64
01.01 - 31.01	21975	5.39	63.26	4236.74	6.64
01.07 - 31.07	24794	88.61	99.05	2258.88	6.64
Golfstrom					
01.01 - 14.01	311	39.45	1.50	30.64	6.68
15.01 - 28.01	338	14.18	1.50	26.74	6.67
01.01 - 31.01	730	46.97	3.58	61.98	6.65
01.07 - 31.07	1745	24.06	0.00	67.45	6.64
Ruhigere mittlere Breiten					
01.01 - 14.01	1157	120.59	4.63	297.03	6.65
15.01 - 28.01	1241	223.03	0.14	706.45	6.65
01.01 - 31.01	2800	430.27	5.99	1203.37	6.64
01.07 - 31.07	3371	0.92	39.23	818.62	6.64
karibische Strömung					
01.01 - 14.01	482	65.19	15.52	107.75	6.66
15.01 - 28.01	366	20.82	0.04	56.86	6.67
01.01 - 31.01	961	92.16	5.74	185.93	6.65
01.07 - 31.07	2013	22.80	156.63	252.42	6.64
Nord-Äquatorstrom					
01.01 - 14.01	1435	528.60	18.20	1028.92	6.64
15.01 - 28.01	1286	934.08	393.61	893.72	6.64
01.01 - 31.01	3070	1532.80	283.88	2102.60	6.64
01.07 - 31.07	3068	1047.58	84.88	1598.36	6.64

Tabelle 4.3: Ergebnisse der ANOVA für vier Zeitraumen (zwei Zwei-Wochen-Intervalle, darunter jeweils zwei Monats-Intervalle) und die drei Geschwindigkeitskomponenten (zonal, meridional und gesamt).

Zudem sei bemerkt, dass durch mehrere Einflussfaktoren, die zusätzlich zu der Strömung auf die Bojen-Geschwindigkeiten wirken, die einfache Varianzanalyse nicht zu dem gewünschten Ergebnis

führen muss. Diese Analyse geht nur von einem Faktor aus. Dies bedeutet, dass bei einer Wechselwirkung zwischen den Einflussfaktoren nur bedingt von einem Erfolg des Tests auszugehen ist. In der Interpretation wurde beispielsweise der Faktor des windgetriebenen Schlupfes (s. Abbildung 2.5) nicht betrachtet, welcher zwar einen kleineren Einfluss auf die Bojengeschwindigkeit hat und hauptsächlich auf die Bojen ohne Treibanker wirkt. Um eine Interpretation bezüglich dieses Faktors zu ermöglichen, müsste ein zwei- oder mehrfaktorieller Test angewendet werden (FAHRMEIR et al. 1999, S. 507). Dazu müssten allerdings mehr Vergleichssignale pro Zeitrahmen vorliegen, was bei dem hier beschriebenen Vergleich nicht möglich ist.

4.5 Empirische Autokorrelationsfunktion

Um einen Eindruck über die räumliche Relation der verglichenen Datensätze zu erhalten, werden empirische Autokorrelationsfunktionen mit der Distanz als Lag τ bestimmt. Es wird für die beiden Geschwindigkeitsrichtungen u und v jeweils die Differenz zwischen den Bojen- und Modell-Geschwindigkeiten Δ_u beziehungsweise Δ_v betrachtet. Der Mittelwert dieser Differenzsignale ergibt sich über

$$\bar{\Delta}_u = \frac{1}{n} \sum_{i=1}^n \Delta_u^{(i)}, \quad \bar{\Delta}_v = \frac{1}{n} \sum_{i=1}^n \Delta_v^{(i)}. \quad (4.7)$$

Die Formel zur Bestimmung der Autokovarianz zum Lag τ ergibt sich gemäß JOURNAL und HUIJBREGTS (1978, S. 194) zu

$$C_u(\tau) = \frac{1}{N_\tau} \sum_{j=1}^{N_\tau} (\Delta_u(\lambda_j, \phi_j) - \bar{\Delta}_u) (\Delta_u(\lambda_j + \lambda_\tau, \phi_j + \phi_\tau) - \bar{\Delta}_u), \quad (4.8)$$

für die zonalen Geschwindigkeiten und

$$C_v(\tau) = \frac{1}{N_\tau} \sum_{j=1}^{N_\tau} (\Delta_v(\lambda_j, \phi_j) - \bar{\Delta}_v) (\Delta_v(\lambda_j + \lambda_\tau, \phi_j + \phi_\tau) - \bar{\Delta}_v), \quad (4.9)$$

für die meridionalen Geschwindigkeiten. Dabei gilt

$$\arccos(\sin(\phi_j + \phi_\tau) \sin(\phi_j) + \cos(\phi_j + \phi_\tau) \cos(\phi_j) \cos(\lambda_\tau)) \cdot 6378 \text{ km} < d_\tau. \quad (4.10)$$

Die Distanz d_τ beschreibt alle Lags des zu τ gehörigen Lag-Intervalls. Die Position (ϕ_j, λ_j) der einzelnen Geschwindigkeit wird im metrischen Maß betrachtet. N_τ entspricht der Anzahl an Multiplikationen innerhalb des jeweiligen Lag-Intervalls. Die Lag-Intervalle wurden mit 25 km gewählt. Die Auswertung erfolgt bis zu einem Lag von 500 km. Bei einem Lag von Null ist die Varianz der Signal-Differenz verzeichnet. Bei $\tau = 12.5 \text{ km}$ ist der erste Wert bezüglich der Kovarianz verzeichnet. Er überspannt die Lags $0 < \tau \leq 25 \text{ km}$. Die nachfolgenden Kovarianz-Werte schließen mit den entsprechenden Lag-Intervallen an.

Da die Varianzen des Differenzsignals bereits in Form der Standardabweichung in den Spalten 5 und

9 der Tabelle 4.2 abzulesen sind, bietet es sich an, nur die Autokorrelationen darzustellen. Diese geben einen besseren Eindruck der sich über die Lags ergebenden Korrelationen, da die Kovarianzen skaliert sind. Die Autokorrelationwerte berechnen sich über

$$K_u(\tau) = C_u(\tau)/C_u(0) \quad \text{und} \quad K_v(\tau) = C_v(\tau)/C_v(0).$$

Die Abbildung 4.4 zeigt die berechneten Korrelationswerte für den Januar 2009, Abbildung 4.5 diejenigen für den Juli. Schwankt die Differenz der beiden Signale, welche um ihre Mittelwerte bereinigt sind, um Null, so sollten die empirischen Autokorrelationswerte, welche dieses Differenzsignal abbilden, bei einem kleinen Lag ebenfalls um Null schwanken. Dies entspräche einem normal-verteiltem Rauschen. Wie aber bereits bekannt, sollte sich nach dem Abzug des Modell-Signals ein Eindruck über die in den Bojen-Geschwindigkeiten zusätzlich zur geostrophischen Strömungserfassung mitgetragenen Einflüssen und deren räumliche Ausdehnung ergeben. Außerdem wird bis ungefähr 25 km der Fehler durch die über 1/4-Grad-Gitter interpolierten Modell-Geschwindigkeiten sichtbar sein und der Einfluss der tatsächlichen räumlichen und zeitlichen Auflösung dieser Daten, denn auch die Modell-Geschwindigkeiten wurden auf dieses 1/4-Grad-Gitter durch verschiedene Interpolations- und Filter-Methoden prozessiert (s. Abschnitt 4.1). Genaue Angaben zur räumlichen und zeitlichen Auflösung oder deren Genauigkeiten bezüglich der gegitterten Geschwindigkeiten sind nicht gegeben. Bekannt ist jedoch, dass die synthetische Methode, welche zur Generierung der Modell-Daten verwendet wird, mit einer Auflösung von $1^\circ \times 0.5^\circ$ Gitterboxen eingebunden wird, und das verwendete Geoid eine theoretische Auflösung von 125 km aufweist.

Betrachtet werden hier mit einem Lag bis maximal 500km Distanzen von etwas weniger als 5° . Damit werden nach Tabelle 2.1 Hauptströme noch mit erfasst, großskalige Wirbel oder Gezeiten zeigen sich in dieser Längenskala jedoch nicht. Zudem werden Differenzen zu Zeitschritten von sechs Stunden betrachtet, was die Erfassung von internen Wellen und Mikroturbulenzen ausschließt.

Bei den zonalen Geschwindigkeiten zeigt sich, der Golfstrom ausgenommen, ein anhaltend gleichbleibender Korrelationswert der Signale über sämtliche Lags. Hingegen scheinen in den meridionalen Geschwindigkeiten nach circa 100 km in den meisten Regionen keine Korrelationen mehr vorhanden. Eine Erklärung für die gleichbleibend hohen Korrelationen der zonalen Geschwindigkeiten für den Januar innerhalb des gesamt-nordatlantischen Signals lassen sich nicht direkt interpretieren. Wetterdaten bezüglich der Windverhältnisse müssten untersucht werden. Im Juli ist es vor allem die karibische Strömung, welche eine anhaltend starke Korrelation in höheren Lags vor allem in der zonalen Geschwindigkeit aufweist. Dies lässt sich weder über durchschnittliche Windgeschwindigkeiten noch über eine langzeitlich gemittelte Ekman-Drift erklären.

Die anhaltende Korrelation des Nord-Atlantik in den zonalen Geschwindigkeiten lässt sich hingegen über große Distanzen unter anderem durch einen großräumigen windgetriebenen Effekt erklären. Teil der windgetriebenen Effekte, die auf die Bojengeschwindigkeiten beschrieben werden, sind der Ekman-Drift zuzuschreiben. Sie ist an der Oberfläche um ungefähr 45° zur rechten der Windgeschwindigkeit gedreht und würde sich so vor allem in den nördlicheren Breiten zonal auswirken. Die Ekman-Drift wirkt sich allerdings für Tiefen von 15m, in welcher Bojen mit Treibanker zentriert sind in einem größeren Winkel zur rechten der Windgeschwindigkeit aus. So kommt es, dass die gemittel-

te Ekman-Drift in den mittleren Breiten eine rein meridional gerichtete Geschwindigkeitsänderung bewirkt. Nach RIO und HERNANDEZ (2004, S.6) stehen in dem atlantischen subpolaren Wirbel 30-35% der Bojen-Geschwindigkeit für die Ekman-Strömung. Dies lässt hier für die Breiten über 50° für die Windgeschwindigkeiten bezüglich der zonalen Geschwindigkeiten im gesamten Nordatlantik-Datensatz vor allem im Januar bestätigen. Es ist zu beachten, dass RIO und HERNANDEZ (2004) diese Werte aufgrund eines von ihnen bestimmten Ekman-Modells ableiten, was wiederum aus den gleichen Bojendaten abgeleitet wird und anschließend zur Einbringung der Bojendaten in das hier zu vergleichende Modell verwendet wird. Im Juli, in dem die Windgeschwindigkeiten geringer sind, zeigen sich auch kleinere Korrelationswerte. Die Ekman-Drift wirkt über einen langen Zeitraum gemittelt, in den Breiten zwischen 20° und 50° hauptsächlich meridional, was in der Abbildung 2.5 nachzuvollziehen ist. In diesem Bereich ist die Windgeschwindigkeit allerdings mit durchschnittlich 4 m s^{-1} geringer als in den höheren Breiten (s. Abbildung 4.3). In Breitengraden darunter sind auch in den übrigen Regionen besonders in den zonalen Geschwindigkeiten die anhaltend hohen Korrelationswerte festzustellen. Gerade bei Strömungen mit sehr hohen Geschwindigkeiten sinkt dieser relative Anteil jedoch (RIO und HERNANDEZ 2004, S.6). Dies würde eine Erklärung für die im Golfstrom nach ungefähr 75 km auf Null abfallende Korrelationen geben. Zudem ist hier die Ekman-Drift vor allem meridional ausgeprägt, was auch für die ruhigeren mittleren Breiten gilt.

Den Windgeschwindigkeit ist auch deshalb ein so großer Anteil an den zusätzlich auf die Bojengeschwindigkeit wirkenden Effekte zuzuschreiben, da zur Auswertung der Daten auch Bojen genutzt wurden, die keinen Treibanker mehr besitzen und dadurch verstärkt von dem Wind über den windgetriebenen Schlupf beeinträchtigt werden.

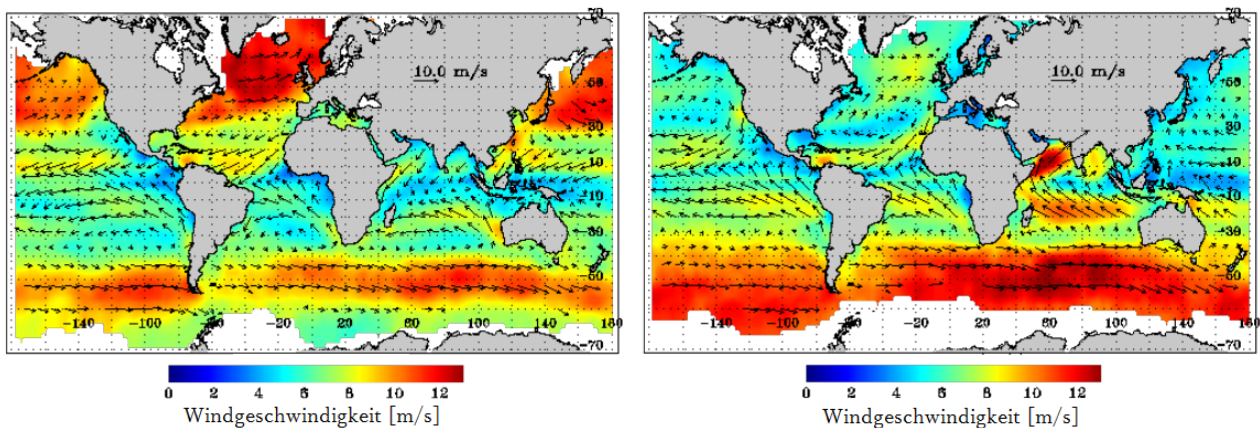


Abbildung 4.3: Durchschnittliche Windgeschwindigkeiten für den Januar (links) und Juli (rechts) von EARSEL (2014)

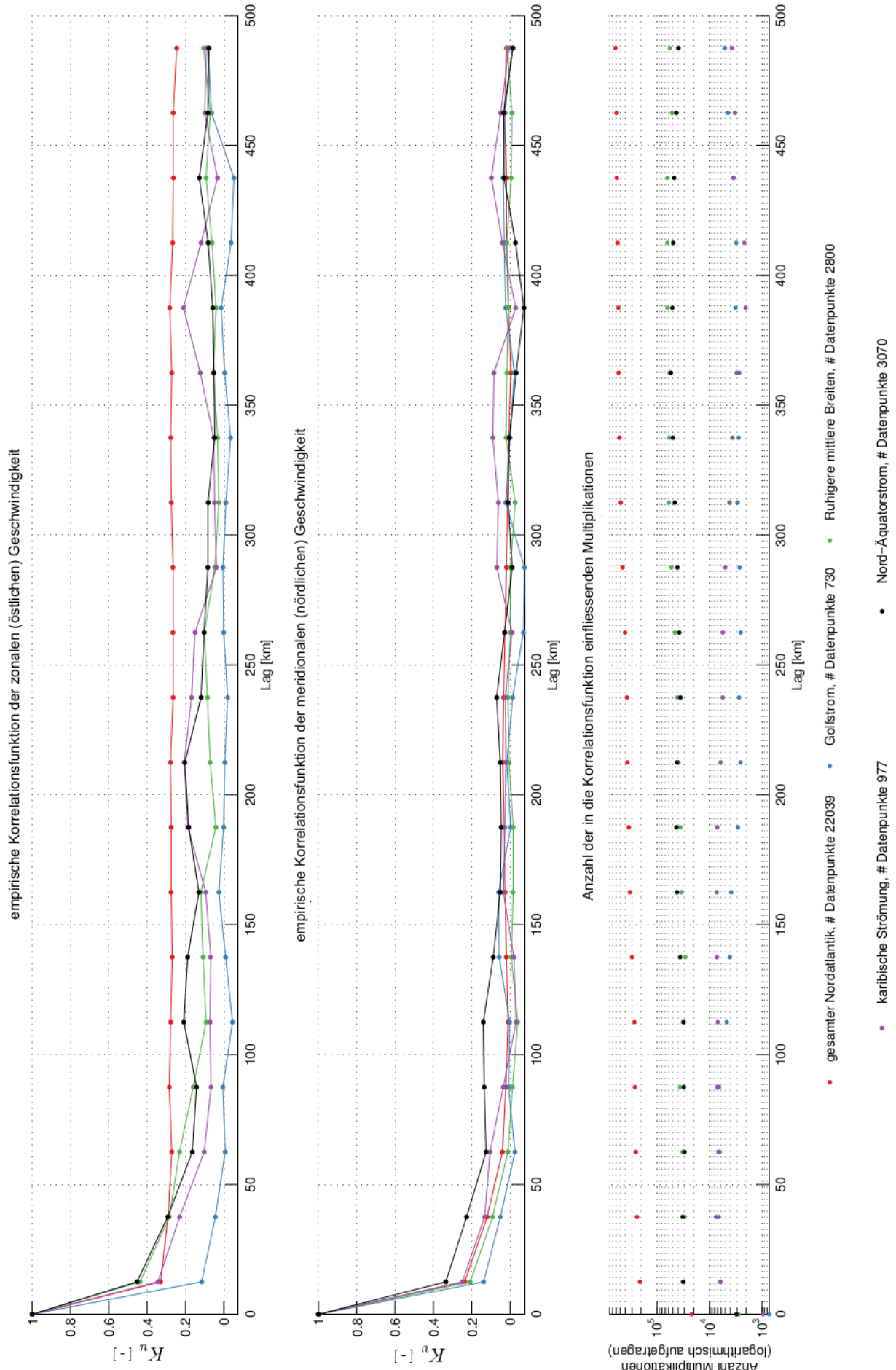


Abbildung 4.4: Empirische Korrelationswerte der fünf Regionen für den Januar 2009

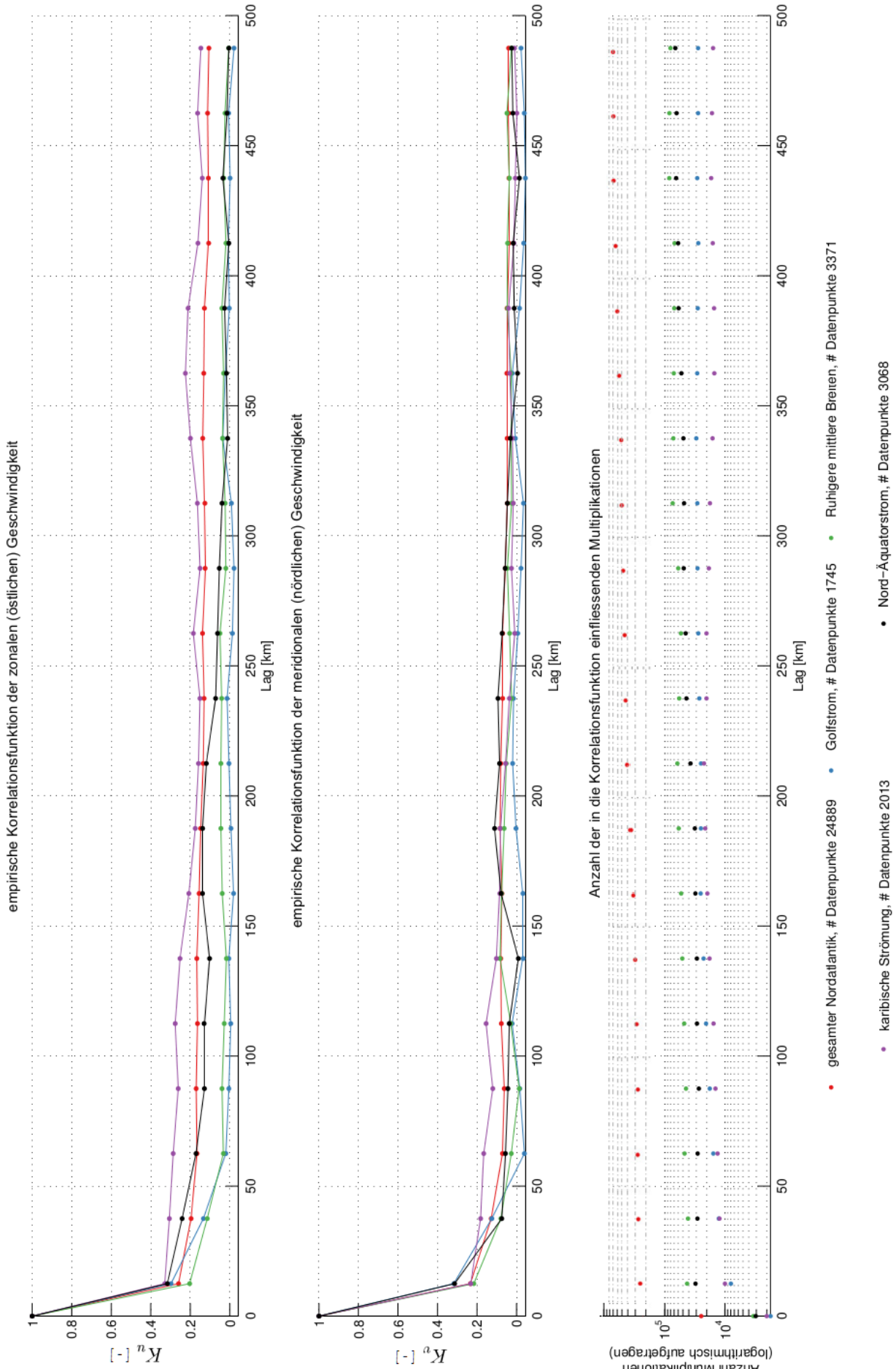


Abbildung 4.5: Empirische Korrelationswerte der fünf Regionen für den Juli 2009

5. Alternative Positionsinterpolation und Geschwindigkeitsbestimmung

Ziel dieses Kapitels soll es sein, ein adäquates Verfahren zu entwickeln, welches erlaubt, Positionen und Geschwindigkeiten zu gesetzten Zeitpunkten aus den verrauschten Positionen der Rohdaten der SVP-Bojen zu bestimmen. Es soll die Nachteile des SVP-Standardverfahrens zur Positions- und Geschwindigkeitsinterpolation, welche in Abschnitt 3.6 dargestellt werden, aufheben.

Anders als das Verfahren zur Erzeugung der von der SVP ausgegebenen Kriging-Daten soll statt des Ordinary Krigings eine Approximation der Roh-Daten mittels stückweiser kubischer Polynome stattfinden. Der Vorteil liegt zum Einen in der direkt aus den geschätzten Funktionswerten abzuleitenden Geschwindigkeiten, wodurch kein Genauigkeitsverlust hinzunehmen ist. Dieser Genauigkeitsverlust tritt im Standardverfahren aufgrund des zentralen Differenzenquotienten zur Geschwindigkeitsableitung aus den interpolierten Positionen auf. Ein weiterer Vorteil liegt in der Tatsache, dass keine empirisch bestimmten Variogramme zur Interpolation genutzt werden müssen. Dadurch müssen Vernachlässigungen bezüglich der höheren Diffusität in den höheren Breiten, also die Sensitivität der Variogrammparameter mit den Breitengraden, sowie die Korrelation der Variogramme in den höheren Breiten, nicht hingenommen werden.

In den nachfolgenden Abschnitten soll eine numerisch günstige und zugleich den Bewegungsschemata der Bojen bestmöglich angepasste Zeitreihenapproximation vorgestellt werden. Sie soll für die Weiterverarbeitung der Daten angepasst werden können und die volle Kovarianzinformation bezüglich der abgeleiteten Positionen und Geschwindigkeiten besitzen.

5.1 Stückweise kubische Polynome über trunktierte Potenzen

Nach FAHRMEIR et al. (2009, S. 296) ist ein Polynom-Spline eine Funktion vom Grad $l \geq 0$, falls sie $(l - 1)$ -mal stetig differenzierbar ist und in den lokalen Intervallen ein Polynom vom Grad l ist. Das globale Intervall $[a, b]$, in dem die Funktion f lebt, definiert die lokalen Intervalle über die Knoten $a = \kappa_1 < \dots < \kappa_m = b$. Ein Polynom-Spline, auch Regressions-Spline genannt, ergibt sich im allgemeinen Modell mit den Parameter α_l mit $l = 0, 1, \dots, 3$ und den Stützstellen t_i ohne Unterteilung in lokale Intervalle zu

$$f(t_i) = \alpha_0 + \alpha_1 t_i + \alpha_2 t_i^2 + \alpha_3 t_i^3. \quad (5.1)$$

Um für den im weiteren erläuterten Algorithmus eine günstige Parametrisierung zu erhalten, wird der von FAHRMEIR et al. (2009, S. 295) vorgeschlagene Funktionsaufbau bezüglich der lokalen Intervalle mit dem erweiterten Parametersatz α_l mit $l = 0, 1, \dots, 2 + m$ verwendet

$$f(t_i) = \alpha_0 + \alpha_1 t_i + \alpha_2 t_i^2 + \alpha_3 t_i^3 + \alpha_4 (t_i - \kappa_2)_+^3 + \dots + \alpha_{2+m} (t_i - \kappa_{m-1})_+^3 \quad (5.2)$$

mit $(t_i - \kappa_j)_+^3 = \begin{cases} (t_i - \kappa_j)^3 & t_i \geq \kappa_j, \\ 0 & \text{sonst,} \end{cases}$ mit $j = 2, \dots, m - 1$.

Dieser Funktionsaufbau erlaubt die Bildung lokaler Polynome in den lokalen Intervallen und erfüllt zugleich die globale Glattheitsanforderung in dem globalen Intervall. Im Nachfolgenden werden die Breite und Länge als einzelne Zeitreihen betrachtet. Sie bilden jeweils den Beobachtungsvektor \mathbf{l} und die dazugehörigen Zeitwerte bilden den Stützstellenvektor \mathbf{t} . Für die Stützstellen wird eine Translation durchgeführt, sodass $t_1 = 0$ ist. Aufgrund des linearen Zusammenhangs zwischen Beobachtungswert beziehungsweise Funktionswert und den Parametern kann ein lineares Gauß-Makoff-Modell (KOCH 2004, S.166f.) zur Durchführung der Ausgleichung angesetzt werden. Die Beobachtungsgleichungen ergeben sich zu

$$\mathcal{L} + \mathbf{v} = \mathbf{A}\mathbf{x} \quad \text{mit} \quad \Sigma\{\mathcal{L}\} = \mathbf{I}, \quad (5.3)$$

wobei diese mit Zufallsvariablen geschrieben ist, welche mit \mathcal{L} die Beobachtungen beschreibt, \mathbf{v} die Verbesserungen und mit \mathbf{x} die Parameter. Die Matrix \mathbf{A} beschreibt den funktionalen Zusammenhang zwischen den Beobachtungen und den Parametern indem sie die jeweilige Ableitung der Beobachtungsgleichung nach den Parametern abbildet. Die Genauigkeit und Korrelation der Beobachtungen wird durch die Kovarianzmatrix $\Sigma\{\mathcal{L}\}$ beschrieben. Da für die Breiten- und Längengrade, welche über das ARGOS-System bestimmt werden, keine Genauigkeits- und Korrelationsangaben existieren, wird diese in den nachfolgenden Berechnungen durch eine Einheitsmatrix ersetzt.

Für den Ausgleich gilt die Minimierungsbedingung der kleinsten Quadrat-Schätzung

$$\mathbf{v}^T \mathbf{v} \dots \text{Min.} \quad (5.4)$$

Die Designmatrix \mathbf{A} , welche die Basisfunktionen der Polynom-Splines enthält, setzt sich über

$$\mathbf{A}_{[n \times m+2]} = \left[\mathbf{1} \quad \mathbf{t} \quad \mathbf{t}^2 \quad \mathbf{t}^3 \quad (\mathbf{t} - \kappa_2)_+^3 \quad \dots \quad (\mathbf{t} - \kappa_{m-1})_+^3 \right] \quad (5.5)$$

zusammen, mit n der Anzahl an Beobachtungen und $m + 2$ der Anzahl an Parametern. Die Spalten beschreiben einen d -dimensionalen Vektorraum, sodass die Bezeichnung als Basis zur Abbildung des Spline-Raums nach HÄMMERLIN und HOFFMANN (1994) gerechtfertigt ist. Die Basen der hier vorgeschlagenen Parametrisierung werden als Basen trunkierter Potenzen, kurz TP-Basen bezeichnet, um eine Abgrenzung zu anderen Spline-Repräsentationen zu schaffen (FAHRMEIR et al. 2009, S. 298). Der Parametervektor, welchen es zu schätzen gilt, um den ausgeglichenen Funktionsverlauf der stückweise kubischen Polynome abbilden zu können, ergibt sich zu

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{l}. \quad (5.6)$$

5.2 Wahl der Knoten

Die Approximation mittels stückweise kubischer Polynome basiert auf der Idee, das globale Intervall, welches sich über die gesamte Zeitreihe erstreckt, in lokale Intervalle zu unterteilen, in welchen dann die kubischen Polynome existieren. Der zu approximierende Datensatz ist aber inhomogen, was seine Bewegungsschemata betrifft. Wie in den Variogrammen bereits festgestellt wurde, zeigen die Bojen je nach Breitengrad unterschiedliche Diffusität in der Trajektorie. Vorstellbar ist, dass Bojen sich in Wirbeln verfangen und dadurch in der Breite und Länge kleine Auf- und Abwärtsbewegungen vollziehen oder ein eher geradliniges Bewegungsmuster aufweisen, weil sie beispielsweise

durch eine Strömung fortgetragen werden oder an einer Stelle ruhen. Polynom-Splines zeigen sich sehr empfindlich bezüglich der Knotenwahl. So kann es in Zeitabschnitten mit viel Bewegung passieren, dass der Spline zu glatt ist, weil die Knoten mit einem zu großen Mindestabstand gewählt wurden. In Bereichen mit einem eher glatten Verlauf kann es passieren, dass der Spline aufgrund zu nahe aneinanderliegender Knoten dem Rauschen oder zu vernachlässigenden Frequenzen folgt. Er könnte auch oszillieren aufgrund weniger oder mit größeren Lücken belegten Stützstellen. Es gilt demnach zwei Dinge zu beachten, wenn es um die Festlegung der Knoten geht. Zum Einen muss ein **Mindestabstand** zwischen den Knoten festgelegt werden, sodass sich die Polynome nicht von dem Rauschen in den Funktionswerten beeinflussen lassen und speziell für diese Zeitreihen zudem Frequenzen herausnehmen, die unter sechs Stunden liegen. Der Wert von sechs Stunden wurde gewählt, da im Späteren aus diesem Datensatz andere Effekte herausgerechnet werden sollen, die mit einer maximalen Auflösung von sechs Stunden verfügbar sind. Dazu gehören der Schlupf und die Ekman-Drift, welche beide durch den Wind beeinflusst werden, dessen Daten nur in diesem Intervall aus der ERA-Interim des ECMWF verfügbar sind. Der Mindestabstand berechnet sich demnach wie folgt

$$\Delta\kappa = \frac{3 \cdot 6}{2} [\text{h}] = 9 [\text{h}]. \quad (5.7)$$

Dies ergibt sich aus der Tatsache, dass sich ein Polynom dritten Grades und der damit verbundenen vier zu bestimmenden Parameter pro Intervall, aus vier Punkten rekonstruieren lässt. Vier Punkte der Winddaten liegen in einem Intervall von $3 \cdot 6$ Stunden und definieren in diesem Zeitrahmen ein kubisches Polynom. Durch die Forderung der Stetigkeit in der zweiten Ableitung entfällt ein zu bestimmender Parameter pro Knoten, wodurch nur noch halb so viele Punkte das Polynom innerhalb eines Intervalls festlegen müssen, sodass auch die Mindestbreite halbiert werden muss.

Als weiterer Aspekt, der die Knotenwahl betrifft, ist die **Aussonderung von Knoten** in glatten Bereichen und in solchen, mit zu wenigen Beobachtungen. Liegen in einem lokalen Intervall weniger als sechs Beobachtungen vor, so wird der Knoten zur rechten entfernt und damit das lokale Intervall doppelt so groß wie zuvor. Die Anzahl von sechs Beobachtungen wurde deshalb gewählt, da so auch in den Anfangs- und Endintervallen, in welchen es je drei Parameter zu bestimmen gilt, das Verhältnis von Parametern zu Beobachtungen $1 : 2$ ist. In den inneren Intervallen müssen nur zwei Parameter bestimmt werden, sodass dadurch auch der Einfluss von einzelnen Ausreißern auf den zu approximierenden Funktionsverlauf deutlich sinkt. Die Vergrößerung der lokalen Intervalle wird solange durchgeführt, bis jedes lokale Intervall nicht weniger als sechs Beobachtungen besitzt. Eine Zeitreihe wird dann unterbrochen, wenn innerhalb von zwölf Tagen weniger als sechs Beobachtungen liegen. Dies ist angelehnt an die Interpolation über Ordinary Kriging von HANSEN und HERMAN (1989, S. 602), welche die Interpolation unterbrechen, wenn weniger als vier Beobachtungen innerhalb von zwölf Tagen vorliegen. Da das Polynom im Gegensatz zur gewichteten Interpolation keine Überbestimmung bei vier Beobachtungen pro Intervall aufweisen würde und so die Lage der Beobachtungen (s. Anhang A.1) und deren Rauschen großen fehlleitenden Einfluss auf die geschätzte Funktion hätte, wird das Minimum an Beobachtungen für den Ausgleich mit Splines auf sechs Beobachtungen innerhalb von zwölf Tagen gesetzt. Durch die Unterbrechung der Zeitreihe kann es sein, dass eine Bojenspur mehrere Zeitreihen beschreibt.

Zuletzt gilt es noch Knoten auszusondern, die in glatten Bereichen der Zeitreihe liegen und zu ungewolltem Oszillieren in der ausgeglichenen Spur führen. Zu diesem Zweck wird im Nachfolgenden zur Erstellung eines multivariaten adaptiven Regression-Splines der gleich benannte Algorithmus MARS beschrieben. Genutzt werden die Positions- und Sensordaten, da von Ihnen die größten Datensätze zur Verfügung stehen (<http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/drib-bder/svp-vcs/res/types-eng.htm>).

5.2.1 Multivariater adaptiver Regressions-Spline (MARS)-Algorithmus

Dieser Algorithmus dient der bestmöglichen Auswahl an Knoten zur Approximation mit Splines. Er wird erstmals von FRIEDMAN (1991) erläutert und in FAHRMEIR et al. (2009, S. 360) ist die hier verwendete Variante skizziert. Im Wesentlichen ist der Algorithmus in zwei Schritte zu unterteilen. Der erste Schritt beschreibt eine Vorwärtssuche. Ausgehend vom minimalen Modell (lokales Intervall entspricht dem globalen Intervall) wird je ein Knoten je Durchlauf zum Modell hinzugenommen, der die beste Anpassung an die Daten gegenüber seinen Mitbewerbern verspricht. Dies geschieht bis das volle Modell erreicht ist. Der zweite Schritt beschreibt die Rückwärtssuche und geht von dem vollen Modell aus, welches sämtliche Knoten einer Startsequenz enthält. Hierbei wird je Durchlauf je ein Knoten herausgenommen, welche den geringsten Einfluss auf die Anpassungsgüte des Modells an die Daten hat. Zuletzt wird über ein Gütekriterium, welches durch einen Wert sowohl eine Aussage über die Güte der Datenanpassung, als auch eine zu hohe Komplexität des Modells bestraft, entschieden, ob das beste Modell der Vorwärtssuche oder das der Rückwärtssuche als insgesamt bestes Modell zurückgegeben wird.

Im Gegensatz zu FAHRMEIR et al. 2009, der die maximale Anzahl der Knoten festlegt um das volle Modell zu limitieren und alle Stützstellen der Beobachtungen als mögliche Positionen für die Knoten setzt, soll hier eine Startsequenz von Intervallgrenzen erzeugt werden, welche die Kriterien zur Wahl der Knoten des letzten Abschnitts 5.2 erfüllen. Die Abbildung 5.1 zeigt eine skizzenhafte Darstellung einer solchen Startsequenz.

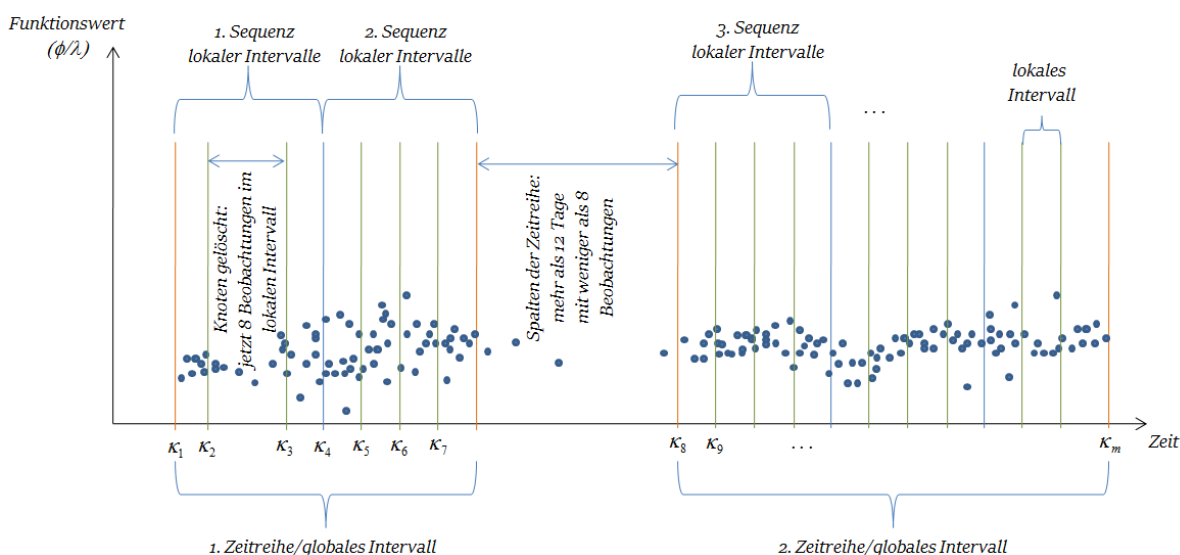


Abbildung 5.1: Schematische Darstellung der Startsequenz, welche alle möglichen Knotenkandidaten κ_j mit $j = 1, \dots, m$ enthält

Die initialisierten Zeitreihen werden in 15-Tagessequenzen unterteilt, da die verwendete Parametrisierung über trunkeerte Potenzen bei größeren Zeitreihen ein numerisch schlecht gestelltes Problem bezüglich der Lösung der dazu gehörigen Normalgleichungen (5.6) ergibt. Betragsmäßig große Werte in den Stützstellen beziehungsweise Knoten können zum Gleitkommaüberlauf bei der numerischen Auswertung führen (FAHRMEIR et al. 2009, S. 303), da bei den trunkeerten Potenzen durch das globale Polynom in den ersten vier Spalten der Designmatrix das globale Intervall überspannt wird, kommt es so in der vierten Spalte zu den maximalen und minimalen Werten. Die so schlecht konditionierte Matrix würde bei der numerischen Auswertung dazu führen, dass die Rechengenauigkeit gängiger Rechensysteme bei Datensätzen von einem bis zwei Jahren deutlich überstiegen wird. Zudem begrenzt die Auswertung von 15 Tagessequenzen innerhalb des MARS-Algorithmus auch die Anzahl der Durchläufe bei der Vorwärts- und Rückwärtssuche. Die Anzahl an Durchläufen je Suche d_s ergibt sich gemäß der Gaußschen Summenformel zu

$$d_s = \frac{(m-2)^2 + (m-2)}{2}. \quad (5.8)$$

Für maximal 41 Knoten in einer 15-Tagessequenz, was sich durch den Mindestabstand zwischen den Knoten (vgl. Formel (5.7)) ergibt, liegt die Anzahl der Durchläufe bei 780.

Die Einteilung in Sequenzen wird allerdings nur zur Bestimmung der für die adaptive Modellanpassung relevanten Knoten verwendet. Um anschließend die Positionen und die Geschwindigkeiten zu bestimmen und die dafür vollständige Kovarianzmatrix zu erhalten werden Basis-Splines (B-Splines) angesetzt, welche den numerischen Nachteil der Polynom-Splines nicht besitzen. Sie beschreiben aber die gleiche mathematischen Grundlage und lassen sich in die gleiche Parametrisierung umschreiben. Somit bilden sie damit über die gleichen Knoten die bestmöglichen Anpassung.

Warum werden also die Polynom-Splines innerhalb des MARS-Algorithmus genutzt? Von FAHRMEIR et al. 2009 wird die Wahl der entsprechenden Knoten über die sogenannte Score- beziehungsweise bei der Rückwärtssuche über die Wald-Statistik ausgeführt. Diese nutzen die Eigenschaften der Polynom-Splines auf effiziente Weise, sodass nicht in den maximal 780 Durchläufen je Suche ein Ausgleich nach (5.6) gerechnet werden muss, sondern nur maximal $m-1=40$ -mal.

Ein innerer Knoten ist bei der Parametrisierung durch Polynom-Splines nur in einer einzelnen Spalte der Designmatrix integriert. Sollte ein Knoten wegfallen oder hinzukommen, so muss die Spalte nur angehängt oder gestrichen werden, was bei den B-Splines oder anderen numerisch effizienteren Parametrisierungen nicht der Fall ist. Zudem ist der zu schätzende Parameter nur dann signifikant für die approximative Darstellung, wenn er sich von Null unterscheidet. Für die Score-Statistik, welche in der Vorwärtssuche das entscheidende Kriterium zur Wahl des als nächstes in das Teilmodell aufzunehmenden Knotens ist, ergibt sich die Score-Funktion zu

$$\mathbf{s} = \mathbf{A}^{+T} \mathbf{l} - \mathbf{A}^{+T} \mathbf{A}^+ \mathbf{x}^+ = \mathbf{A}^{+T} (\mathbf{l} - \mathbf{A}^+ \mathbf{x}^+) = \mathbf{A}^{+T} (\mathbf{l} - \mathbf{A} \mathbf{x}) = \mathbf{A}^{+T} \mathbf{v} \quad (5.9)$$

(FAHRMEIR et al. 2009, S. 201). Der Parametervektor \mathbf{x}^+ ist dabei der um einen Null-Eintrag ergänzte Vektor des Teilmodells. Die Designmatrix ist ebenfalls die Designmatrix des Teilmodells, mit der ergänzten Spalte bezüglich des zu testenden Knotens. Die Score-Funktion wird demnach

umso größer, desto sensitiver die zu minimierende Funktion bezüglich des zu testenden Parameters ist. Die Testgröße der Score-Statistik ergibt sich zu

$$t_{score} = \mathbf{s}^T (\mathbf{N}^+)^{-1} \mathbf{s} \quad \text{mit} \quad \mathbf{N}^+ = \mathbf{A}^{+T} \mathbf{A}^+ \quad (5.10)$$

(FAHRMEIR et al. 2009, S. 204). Die Score-Statistik zeigt nur den Vorteil, dass \mathbf{x}^+ und \mathbf{A}^+ mit geringem Aufwand erstellt werden können, da auch hier die inverse der Normalgleichungsmatrix \mathbf{N}^+ bestimmt werden muss. Die Wald-Statistik zeigt sich effizienter. Diese wird in der Rückwärtssuche verwendet, da sie die Signifikanz eines Knotens durch das Testen eines Teilmodells gegenüber einem übergeordneten Modell ausgibt. Die Testgröße bezüglich der Wald-Statistik ergibt sich über

$$t_{wald} = \frac{\mathbf{x}_i^2}{\mathbf{N}_{i,i}^{-1}} \quad (5.11)$$

(FAHRMEIR et al. 2009, S. 204), wobei \mathbf{x}_i den geschätzten Parameter bezüglich des zu testenden Knotens bezeichnet. Die Normalgleichung \mathbf{N} wird hier nur einmal invertiert, die Einträge $\mathbf{N}_{i,i}$ können dann herausgegriffen werden. Beide Statistiken sind asymptotisch Chi-Quadrat verteilt mit je einem Freiheitsgrad.

Der MARS-Algorithmus wird in Algorithmus 1 abgebildet. FRIEDMAN (1991, S. 17) notiert den Algorithmus zweigeteilt und bestimmt zu ergänzende und zu eliminierende Knoten jeweils durch das generalisierte Kreuzvalidierungskriterium GCV, nutzt aber auch die hier verwendeten trunkierten Potenzen zur Parametrisierung. Der Grund für die Anwendung der Score- und Wald-Statistik in den inneren Schleifen (Zeile 6-7 und 18-19), wurde bereits erläutert. Der Grund für ein Gütekriterium wie das Akaikes Informationskriterium (AIC) oder das generalisierte Kreuzvalidierungs-Kriterium (GCV) in den äußeren Schleifen (Zeile 5-12 und 17-24) und nach der Initialisierung des Ausgangsmodells (Zeile 4 und 16) soll nun erörtert werden. Zum Einen können sich die Teststatistiken Wald und Score nach FAHRMEIR et al. (2009, S.205) für kleine Stichprobenumfänge ($n \leq 50$) deutlich voneinander unterscheiden. Zudem muss in dieser äußeren Schleife ohnehin das neue Teil- beziehungsweise übergeordnete Modell geschätzt werden, sodass die Effizienz der beiden Statistiken entfällt. Ergänzend soll in dieser äußeren Schleife ein Gütekriterium gewählt werden, das eine Aussage bezüglich der Datenanpassung und der Modellkomplexität tätigt. Die Statistiken geben allerdings nur eine Aussage bezüglich der Datenanpassung, was in den inneren Schleifen auch nicht von Bedeutung ist, da dort die Anzahl der Parameter konstant ist. Das GCV, welches von FRIEDMAN (1991) als Gütekriterium in allen Schleifen genutzt wurde, lässt sich wie folgt berechnen

$$GCV = \frac{\frac{1}{n} \mathbf{v}^T \mathbf{v}}{1 - \frac{C}{n}} \quad \text{mit} \quad \mathbf{v} = \mathbf{l} - \mathbf{A} \mathbf{x} \dots \text{Verbesserungsvektor} \quad (5.12)$$

wobei er die Kostenfunktion C nicht rein über die effektive Parameterzahl beziehungsweise äquivalenten Freiheitsgrade bestimmt, sondern eine zusätzliche anwachsende Kostenfunktion ansetzt, welche die zusätzlich eingebrachten Parameter besonders bestraft.

Algorithmus 1 : Variation des MARS-Algorithmus zur optimalen Knotenwahl**Eingabe** : $\mathbf{l}_{[n \times 1]}$, $\mathbf{t}_{[n \times 1]}$, $\boldsymbol{\kappa}_{[m \times 1]}$ **Ausgabe** : $\hat{\boldsymbol{\kappa}}$ 1 Initialisiere $\mathbf{A}_{[n \times m+2]}$ mit $(\mathbf{t}, \boldsymbol{\kappa})$.

// Schritt 1: Vorwärtssuche beginnend mit dem minimalen Modell

2 Setze $\mathbf{A}^f \leftarrow \mathbf{A}_{(:,1:4)}$ (Designmatrix zum Startmodell), $\mathbf{A}^R \leftarrow \mathbf{A}_{(:,5:end)}$ (übrige Spalten)3 Setze $\hat{\boldsymbol{\kappa}} \leftarrow \boldsymbol{\kappa}_{([1,m])}$ (Knoten des Startmodells), $\boldsymbol{\kappa}^R \leftarrow \boldsymbol{\kappa}_{(2:m-1)}$ (übrige Knoten), $m^* \leftarrow m - 2$ (Anzahl zu testender Knoten)4 Bestimme $\mathbf{x} \leftarrow (\mathbf{A}^f, \mathbf{l})$ und den Wert des Gütekriteriums $\mathbf{c}^{GK}(1) \leftarrow (\mathbf{x}, \mathbf{A}^f, \mathbf{l})$ 5 **for** $i = 1, \dots, m-2$ **do**6 **for** $k = 1, \dots, m^*$ **do**7 Bestimme die Testgröße der Score-Statistik $\mathbf{c}^{Score}(k) \leftarrow ([\mathbf{x}; 0], [\mathbf{A}^f, \mathbf{A}_{(:,k)}^R], \mathbf{l})$ 8 Setze i_c als Index des Maximums in \mathbf{c}^{Score} 9 Ergänze $\mathbf{A}^f \leftarrow [\mathbf{A}^f, \mathbf{A}_{(:,i_c)}^R]$ und $\hat{\boldsymbol{\kappa}} \leftarrow [\hat{\boldsymbol{\kappa}}; \boldsymbol{\kappa}_{(i_c,1)}^R]$ 10 Lösche $\mathbf{A}_{(:,i_c)}^R \leftarrow []$ und $\boldsymbol{\kappa}_{(i_c,1)}^R \leftarrow []$ 11 Bestimme $\mathbf{x} \leftarrow (\mathbf{A}^f, \mathbf{l})$ und den Wert des Gütekriteriums $\mathbf{c}^{GK}(i+1) \leftarrow (\mathbf{x}, \mathbf{A}^f, \mathbf{l})$ 12 Setze $m^* \leftarrow m^* - 1$ 13 Setze i_c als Index des Minimums in \mathbf{c}^{GK} 14 Setze $\mathbf{c}^f \leftarrow \mathbf{c}^{GK}(i_c)$ und $\hat{\boldsymbol{\kappa}} \leftarrow \hat{\boldsymbol{\kappa}}_{(1:i_c+1,1)}$

// Schritt 2: Rückwärtssuche beginnend mit dem maximalen Modell

15 Setze $\hat{\boldsymbol{\kappa}}^R \leftarrow []$ (Vektor für die gelöschten Knoten), $m^* \leftarrow m - 2$ (Anzahl zu testender Knoten)16 Bestimme $\mathbf{x} \leftarrow (\mathbf{A}, \mathbf{l})$ und den Wert des Gütekriteriums $\mathbf{c}^{GK}(1) \leftarrow (\mathbf{x}, \mathbf{A}, \mathbf{l})$ 17 **for** $i = 1, \dots, m-2$ **do**18 **for** $k = 1, \dots, m^*$ **do**

19 Bestimme die Testgröße der Wald-Statistik

20 $\mathbf{c}^{Wald}(k) \leftarrow (\mathbf{x}_{([1:3+k, k+5:m^*-2])}, \mathbf{A}_{(:, [1:3+k, k+5:m^*-2])}, \mathbf{l})$ 21 Setze i_c als Index des Minimums in \mathbf{c}^{Wald} 22 Hänge den aus dem Modell gelöschten Knoten an $\boldsymbol{\kappa}^R \leftarrow [\boldsymbol{\kappa}^R; \boldsymbol{\kappa}_{(i_c+1)}]$ 23 Lösche $\mathbf{A}_{(:, i_c+4)} \leftarrow []$ und $\boldsymbol{\kappa}_{(i_c+1)} = []$ 24 Bestimme $\mathbf{x} \leftarrow (\mathbf{A}, \mathbf{l})$ und den Wert des Gütekriteriums $\mathbf{c}^{GK}(i+1) \leftarrow (\mathbf{x}, \mathbf{A}, \mathbf{l})$ 25 Setze $m^* \leftarrow m^* - 1$ 26 Setze i_c als Index des Minimums in \mathbf{c}^{GK} 27 Setze $\mathbf{c}^b \leftarrow \mathbf{c}^{GK}(i_c)$

// Schritt 3: Auswahl des besten Modells

28 **if** $\mathbf{c}^b < \mathbf{c}^f$ **then**29 Setze $\hat{\boldsymbol{\kappa}} \leftarrow \text{sortiere}[\boldsymbol{\kappa}; \boldsymbol{\kappa}_{(i_c:m-1,1)}^R]$

So ergibt sich für die erweiterte Kostenfunktion

$$C = \text{spur} \left(\mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A} \right) + 1 + d(m + 2) \quad (5.13)$$

wobei die Berechnung der Spur in linearen Modellen ohne Strafterme durch die Anzahl der Parameter, also in dem hier genutzten Modell durch $m + 2$ ersetzt werden kann.

Für die Konstante d nennt FRIEDMAN (1991, S. 21) einen Wertebereich zwischen $2 \leq d \leq 4$ und gibt einige Anhaltspunkte, wie diese gewählt werden kann. Aufgrund der Unsicherheit bezüglich der Wahl von d soll in der hier vorgestellten Variante des MARS-Algorithmus, das AIC eine Alternative zum GCV als Gütekriterium sein.

Nach BURNHAM und ANDERSON (1998, S. 48) lässt sich für den Ansatz der kleinsten-Quadrate-Schätzung der Wert für das Gütekriterium mit

$$AIC = n \log(s_{MLE}^2) + 2((m + 2) + 1), \quad (5.14)$$

berechnen. Zu beachten ist, dass nicht die geschätzte a posteriori Varianz der Kleinsten-Quadrate-Schätzung in den Logarithmus eingesetzt wird. Hier ist die Varianz der Maximum-Likelihood-Schätzung (MLE) einzusetzen, wodurch sich

$$s_{MLE}^2 = \frac{\mathbf{v}^T \mathbf{v}}{n} \quad (5.15)$$

ergibt. Akaike fand eine Beziehung zwischen der relativen Kullback-Leibler-Information und der maximierten Log-Likelihood (BURNHAM und ANDERSON 1998, S. 46). Die Kullback-Leibler-Information bezeichnet ein Maß für die Unterschiedlichkeit zweier Wahrscheinlichkeitsverteilungen, wobei hier eine der Verteilungen empirische Beobachtungen und die andere ein Modell repräsentiert. Das daraus abgeleitete Informationskriterium nannte Akaike 1973 zunächst „an information criterion“ (AKAIKE 1973). Seine Erkenntnisse führten zu praktischen und theoretischen Fortschritten in der Modellwahl und der Analyse komplexer Datensätze (BURNHAM und ANDERSON 1998; STONE 1982), wodurch das Informationskriterium für den hier aufgeführten Anwendungsfall die richtige Wahl sein sollte. Dies ist aber in Verbindung mit der gewünschten Art der Schätzung abzuwägen, was in der Ergebnisdiskussion weiter erörtert wird. Gesagt sei, dass durch die im AIC nicht zu verstärkende Kostenfunktion, Korrelationen beziehungsweise zu eliminierende Signale nicht gut ignoriert werden.

Im Gegensatz zum rekursiven Partitionieren, welches den Ursprung des MARS-Algorithmus bildet, setzt der MARS-Algorithmus zwei statt einem Suchlauf ein. Um dies zu begründen muss kurz die rekursive Partitionierung betrachtet werden. Diese teilt den Raum Ω in zwei Unterräume Ω_1 und Ω_2 an der Stelle des minimalsten Gütekriteriums. Anschließend werden Ω_1 und Ω_2 jeweils einzeln für eine weitere Teilung betrachtet und so weiter. Die Unterräume schneiden sich dabei nicht. Die im MARS-Algorithmus verwendeten trunkierten Potenzen teilen den Raum bei der Hinzugabe neuer Knoten jedoch nicht, sodass sich die entstehenden Unterräume überlappen und diese, solange der Startknoten nicht gelöscht wird, nicht als getrennte Unterräume zu betrachten sind. Aus diesem Grund muss keine komplizierte und rechenaufwendige rückwärtig schrittweise Lösstrategie ange-

wandt werden um die Verbindung mit dem getrennten Unterraum zu betrachten (FRIEDMAN 1991, S. 13). Deshalb wird hier eine einfache Betrachtung über eine Rückwärtssuche angewandt. Während die Vorwärtssuche gelegentlich zu einer Überanpassung (Overfitting) des Modells führt, zielt die Rückwärtssuche auf ein Vereinfachen oder Kürzen (Pruning) des Modells, was aus der Erstellung von Entscheidungsbäumen bekannt ist (HASTIE et al. 1994, S. 1264). Dadurch kommt es teilweise zu redundanten Suchabläufen, es zeigt sich aber, dass gerade bei komplexen Modellen signifikante Unterschiede in den Ergebnissen des ersten und zweiten Schritts des Algorithmus auftreten.

5.2.1.1 Alternativen zum MARS-Algorithmus

FAHRMEIR et al. (2009, S. 306) nennt zwei wesentliche Strategien, Splines bezüglich ihrer Sensitivität in Hinblick auf die Knotenwahl, bestmöglich zu schätzen. Zum Einen nennt er die adaptive Wahl, welche auf Modellwahlstrategien basiert und hier durch den MARS-Algorithmus umgesetzt wurde. Zudem beschreibt er die Regularisierung des Schätzproblems durch Penalisierungsansätze, die zu den sogenannten Polynom-Splines führen. Grundlage ist die Schätzung eines Strafterms, der die Minimierungsbedingung der Kleinsten-Quadrate-Schätzung für Polynom-Splines mit trunkeerten Potenzen wie folgt erweitert

$$\mathbf{v}^T \mathbf{v} + \lambda^p (\mathbf{x}^T \mathbf{x}) \quad \dots \text{Min.} \quad (5.16)$$

Der Glättungsparameter $\lambda^p \geq 0$ steuert die Stärke der Bestrafung für zu raue Funktionen. Durch Variation von λ^p zwischen den Extremen $\lambda \rightarrow 0$ und $\lambda \rightarrow \infty$ kann man eine eher an der Kleinsten-Quadrate-Schätzung gelegene oder eine rein durch die Modellwahl dominierte Schätzung erhalten, was wiederum durch ein Informationkriterium wie dem GCV oder dem AIC durch eine Raster-suche objektiv ausgewählt werden kann. Auch wenn FAHRMEIR et al. (2009, S. 308) den Vorteil dieses Verfahrens in der Schätzung nur eines Parameters sieht, was im Gegensatz zu der mühsamen Wahl von Anzahl und Lage der Knoten sicherlich effizienter zu bestimmen ist, so wirkt der Parameter jedoch global über die gesamte Zeitreihe. Ist eine Zeitreihe nun überwiegend mit hohen Frequenzen belastet und hat nur kurze Abschnitte mit einem eher glatteren Verlauf, so wird dieser durch ungewolltes Oszillieren des so geschätzten Polynom-Spline überlagert. Umgekehrt werden hohe Frequenzen zu stark geglättet. Demzufolge ist eine günstige Wahl des Mindestabstands zwischen den Knoten gleichbedeutend mit der Einführung eines Strafterms. Ist der Mindestabstand zwischen den Knoten allerdings nicht sinnvoll zu begründen, so bietet sich die von FAHRMEIR et al. (2009) vorgeschlagene Strategie an. So kann die Lage der Knoten beispielsweise von der in dem Intervall enthaltenen Anzahl an Beobachtungen abhängen, was aber auch bei dem MARS-Algorithmus denkbar ist (FRIEDMAN 1991, S. 26-27).

In Verbindung mit dem MARS-Algorithmus wird mehrfach der sogenannte BRUTO-Algorithmus nach Hastie genannt. Als eine Abwandlung des TURBO-Algorithmus nach FRIEDMAN und SILVERMAN 1989, wird er in der Diskussion des den TURBO-Algorithmus vorstellenden Artikels erstmals erwähnt. In FRIEDMAN (1991, S. 26) wird herausgearbeitet, dass der MARS-Algorithmus für Zeitreihen eine Version des TURBO darstellt. Er schätzt jeweils einen Glättungsparameter pro Knoten

über eine Rückwärtssuche. Die Minimierungsbedingung bezüglich der trunkierten Potenzen ergibt sich demnach zu

$$\mathbf{v}^T \mathbf{v} + \sum_{i=1}^{m+2} \lambda^p x_i^2 \quad \dots \text{Min.} \quad (5.17)$$

Gelöst wird diese Problemstellung durch einen iterativen Gauß-Seidel-Algorithmus, was aber an dieser Stelle nicht weiter ausgeführt werden soll. Nach HASTIE et al. (1994, S. 1265) ist der BRUTO-Algorithmus additiv und daher einschränkender als der multivariater adaptiver Regressions-Spline (MARS)-Algorithmus, andererseits arbeitet er mit Schätzern in einer glättenden Art und Weise (muss einzelne Parameter gegen die Vielzahl anderer einzeln testen) und kann daher in Situationen mit vielen Parametern effektiver sein.

Neben diesen Straftermen in der L2-Schätzung ist es auch möglich, in der L1-Schätzung Strafterme zur Regularisierung der Modellparameter zu integrieren. Diese geschieht beispielsweise in LASSO, dem „least absolute shrinkage and selection operator“, dieser wird hier aber nicht weiter betrachtet.

5.3 Bestimmung der Zielgrößen und deren Kovarianzinformation

Nachdem die Wahl der Knoten erfolgt ist, können die Positionswerte und dazugehörigen Geschwindigkeiten, sowie deren vollen Kovarianzinformation bestimmt werden. Die Zeitpunkte, zu denen diese Werte geschätzt werden ist dabei innerhalb der Zeitreihe variabel.

Da die trunkierten Potenzen der Polynom-Splines aufgrund der schlechten numerischen Eigenschaften nicht dazu geeignet sind, lange Zeitreihen zu approximieren, wird eine verwandte Darstellung genutzt. Die B-Splines beschreiben ebenfalls stückweise kubische Polynome und setzen zudem die gleiche Anzahl an Parametern pro Knoten. So sind die selektierten Knoten des MARS-Algorithmus genauso für B-Splines gültig. Die Basisfunktion für die Parametrisierung eines zweimal stetig differenzierbaren B-Splines mit dem Grad $p = 3$ lässt sich rekursiv über

$$b_i^p(t_j) = \frac{t_j - \kappa_i}{\kappa_{i+p} - \kappa_i} b_i^{p-1}(t_j) + \frac{\kappa_{i+p+1} - t_j}{\kappa_{i+p+1} - \kappa_{i+1}} b_{i+1}^{p-1}(t_j) \quad \text{mit } j = 1, \dots, n \quad (5.18)$$

ausdrücken. Da sich die Basisfunktion eines B-Spline höheren Grades immer durch denjenigen niederen Grades definiert, muss der B-Spline zum Grad $p = 0$ als Initialisierung definiert werden mit

$$b_i^0(t_j) = \begin{cases} 1 & \kappa_i \leq t_j < \kappa_{i+1}, \\ 0 & \text{sonst,} \end{cases} \quad i = 1, \dots, m + 2. \quad (5.19)$$

Die Knotenmenge muss zur Durchführung dieser rekursiven Berechnungsweise um drei irrelevante Knoten vor dem Definitionsbereich $a = \kappa_1$ und drei irrelevante Knoten hinter den Definitionsbereich $b = \kappa_m$ erweitert werden. Diese werden jeweils mit dem Mindestabstand zum nachfolgenden oder vorherigen angehängt.

Die Designmatrix enthält die Basisfunktionen spaltenweise gemäß 5.18, wobei die Auswertung bei

dem Knoten κ_{-2} begonnen wird und bei dem Knoten κ_m endet. So ergeben sich die $m + 2$ Spalten der Designmatrix der B-Splines zu

$$\mathbf{A}^B = \begin{bmatrix} \mathbf{b}_{-2}^3(t) & \mathbf{b}_{-1}^3(t) & \mathbf{b}_0^3(t_1) & \mathbf{b}_1^3(t) & \dots & \mathbf{b}_{m-1}^3(t) & \mathbf{b}_m^3(t) \end{bmatrix}. \quad (5.20)$$

Die Designmatrix weist eine sich überlappende Blockstruktur auf, deren Blöcke jeweils vier Spalten breit sind, sodass sich je 6 Basisfunktionen überschneiden (s. Abbildung 5.2). Die zu schätzenden Parameter ergeben sich über

$$\mathbf{x}^B = \left(\mathbf{A}^{B^T} \mathbf{A}^B \right)^{-1} \mathbf{A}^{B^T} \mathbf{l} \quad \text{mit} \quad \Sigma\{\boldsymbol{\chi}^B\} = \left(\mathbf{A}^{B^T} \mathbf{A}^B \right)^{-1} \quad (5.21)$$

der dazugehörigen Kovarianzmatrix.

Um nun die ausgeglichenen Positionen und Geschwindigkeiten zu den gewählten Zeitpunkten $t_k^{(G)}$ mit $k = 1, \dots, n_g$ zu bestimmen muss eine Matrix \mathbf{G} aufgestellt werden, die abwechselnd den funktionalen Zusammenhang für die Position und für die Geschwindigkeit darstellt. Die Einträge lassen sich wie folgt darstellen

$$\mathbf{G} = \begin{bmatrix} b_{-2}^3(t_1^{(G)}) & b_{-1}^3(t_1^{(G)}) & b_0^3(t_1^{(G)}) & b_1^3(t_1) & 0 & \dots & 0 \\ b_{-2}^3(t_1^{(G)}) & b_{-1}^3(t_1^{(G)}) & b_0^3(t_1^{(G)}) & b_1^3(t_1^{(G)}) & 0 & \dots & 0 \\ & & & \ddots & & & \\ 0 & \dots & & 0 & b_{m-3}^3(t_{n-1}) & b_{m-2}^3(t_{n-1}) & b_{m-1}^3(t_{n-1}) & b_m^3(t_{n-1}) \\ 0 & \dots & & 0 & b_{m-3}^3(t_n) & b_{m-2}^3(t_n^{(G)}) & b_{m-1}^3(t_n^{(G)}) & b_m^3(t_n^{(G)}) \\ 0 & \dots & & 0 & b_{m-3}^3(t_n^{(G)}) & b_{m-2}^3(t_n^{(G)}) & b_{m-1}^3(t_n^{(G)}) & b_m^3(t_n^{(G)}) \end{bmatrix}, \quad (5.22)$$

wobei sich die Ableitung der Basisfunktion nach FAHRMEIR et al. (2009, S. 305) zu

$$b_i^p(t_j) = 3 \left(\frac{1}{\kappa_{i+p} - \kappa_i} b_i^{p-1}(t_j) - \frac{1}{\kappa_{i+p+1} - \kappa_{i+1}} b_{i+1}^{p-1}(t_j) \right) \quad (5.23)$$

ergibt. Der a posteriori Varianzfaktor ergibt sich über

$$\tilde{s}^2 = \frac{\mathbf{v}^T \mathbf{v}}{n - (m + 2)}. \quad (5.24)$$

Nun können die Zielgrößen, also der ausgeglichene Funktionswert- und Geschwindigkeitsvektor bestimmt werden

$$\tilde{\mathbf{l}}_g = \mathbf{G} \mathbf{x} \quad \text{mit} \quad \Sigma\{\tilde{\mathcal{L}}_G\} = \tilde{s}^2 \mathbf{G} \Sigma\{\boldsymbol{\chi}^B\} \mathbf{G}^T, \quad (5.25)$$

der zugehörigen Kovarianzmatrix.

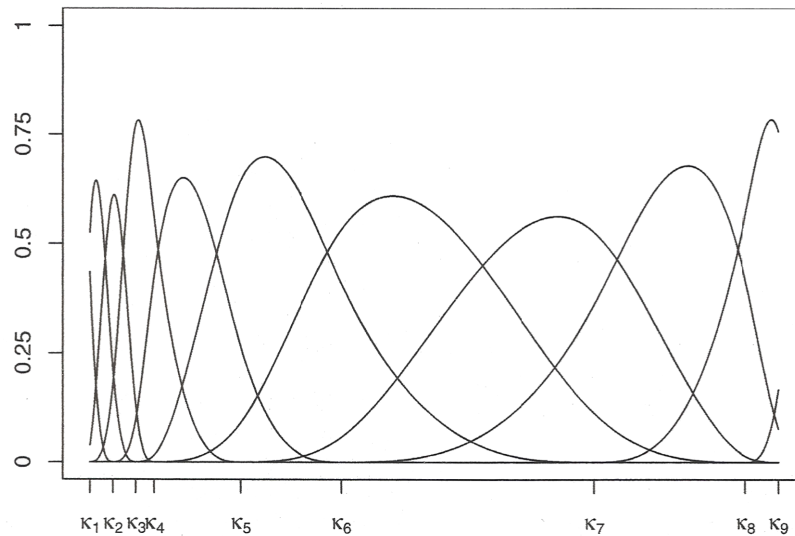


Abbildung 5.2: B-Splines mit unregelmäßig Knotenabständen nach FAHRMEIR et al. (2009)

5.4 Implementierung

In diesem Abschnitt soll der Algorithmus zur Ableitung der Breiten beziehungsweise Längenangabe und der richtungsspezifischen Geschwindigkeit, sowie deren Kovarianzmatrix skizziert werden. Die Abschnitte 5.1 bis 5.3 beschreiben jene in die Funktion `MARSapprox` integrierten Rechenschritte. Die Funktion ist der Daten-CD beigelegt. Der Algorithmus 2 zeigt schematisch den Ablauf dieser Funktion. Neben dem Beobachtungsvektor und dem zugehörigen Stützstellenvektor, welche zu einer Bojenspur gehören, enthält die Eingabe die Variable `Case` vom Struktur-Datentyp, welche sich wie folgt mit den beispielhaften Einträgen aufbaut

```
Case.Interp: 'step'      Case.minDelt: 0.375      Case.LOFcase: 1
Case.Step: 0.25         Case.maxGap: 12
```

und damit dem Verwender bezüglich der zu bestimmenden Zielgrößen und der Wahl der Knotenpositionen einige Einstellmöglichkeiten bietet. Zum Einen ist über den Eintrag `Case.Interp` entweder über 'step' das Abgreifen der Zielgrößen zu äquidistanten Zeitpunkten innerhalb der Zeitreihe möglich. `Case.Step` bestimmt den zeitlichen Abstand zwischen diesen Zeitpunkten. Wird 'observ' in `Case.Interp` eingesetzt, so werden die Zielgrößen zu den Zeiten der Stützstellen der Beobachtungen bestimmt. Der Mindestabstand zwischen den Knoten wird in `Case.minDelt` und der maximale Knotenabstand mit weniger als sechs Beobachtungen wird in `Case.maxGap` festgelegt und dient der Unterbrechung der Zeitreihe bei zu geringer Abtastung. Die Variable `Case.LOFcase` legt das „loss-of-fit“-Kriterium, also das Gütekriterium fest. Eine Eins steht für das GCV, eine zwei für das AIC. Die beispielhaften Inhalte der Struktur `Case` entsprechen den Werten, deren Begründung in dem Abschnitt 5.2 nachzulesen ist.

Der Bezeichnung `MARS` in der Zeile 4 bezeichnet die Implementierung des in Algorithmus 1 beschriebenen Pseudocodes.

Algorithmus 2 : Skizzierung der Funktion MARSapprox**Eingabe** : $\text{Case}, \mathbf{l}_{[n \times 1]}, \mathbf{t}_{[n \times 1]}$ **Ausgabe** : $\tilde{\mathbf{l}}, \Sigma\{\tilde{\mathcal{L}}\}$

- 1 Initialisiere $\overline{\kappa}_i \leftarrow (\mathbf{t}, \text{Case})$ mit $i = 1, \dots, p$ und p der Anzahl an Zeitreihen und \mathbf{s}_i einem Indexvektor, der die Sequenzgrenzen in $\overline{\kappa}_i$ definiert.
- 2 **for** $i = 1, \dots, p$ (*Laufe über die Zeitreihen*) **do**
- 3 **for** $k = 1, \dots, s_i$ (*Laufe über die Sequenzen*) **do**
- 4 Suche die signifikanten Knoten $\hat{\kappa}_i^{(k)} \leftarrow \text{MARS}(\mathbf{l}, \mathbf{t}, \text{Case}, \overline{\kappa}_i(\mathbf{s}_i(k):s_i(k+1)))$
- 5 $\hat{\kappa} \leftarrow [\hat{\kappa}, \hat{\kappa}_i^{(k)}]$
- 6 Lege $\mathbf{t}^{(G)} \leftarrow (\text{Case}, \hat{\kappa}, \mathbf{t})$ an
- 7 Bestimme $\mathbf{A}^B \leftarrow (\hat{\kappa}, \mathbf{t})$ und anschließend $\mathbf{x}^B \leftarrow (\mathbf{A}^B \mathbf{l})$
- 8 Bestimme $\mathbf{G} \leftarrow (\hat{\kappa}, \mathbf{t}^{(G)})$ und $\tilde{\mathbf{l}}_g \leftarrow (\mathbf{x}^B, \mathbf{G})$, sowie $\Sigma\{\tilde{\mathcal{L}}_g\} \leftarrow (\mathbf{G})$
- 9 Ergänze $\tilde{\mathbf{l}} \leftarrow \begin{bmatrix} & \tilde{\mathbf{l}} & \\ \tilde{\mathbf{l}}_g(1:2:end) & \tilde{\mathbf{l}}_g(2:2:end) & \mathbf{t}^{(G)} \end{bmatrix}$; $\Sigma\{\tilde{\mathcal{L}}\} \leftarrow \begin{bmatrix} \Sigma\{\tilde{\mathcal{L}}\} & \mathbf{0} \\ \mathbf{0} & \Sigma\{\tilde{\mathcal{L}}_g\} \end{bmatrix}$;

5.5 Ergebnisse am synthetischen und praktischen Beispiel

Um die Eigenschaften des in diesem Kapitel vorgestellten Verfahrens zu analysieren, zeigt sich eine synthetische Zeitreihe als sinnvoll. Diese wurde durch eine Verkettung von Fourierreihen erzeugt (s. Abbildung 5.3(a)). Sie besteht aus drei Komponenten: der Referenzspur, welche möglichst gut approximiert werden soll, eine überlagernde Spur mit einer Frequenz die ignoriert werden soll, sowie einer Störfrequenz, welche nur stückweise in die Zeitreihe integriert wird und lokal auftretende stark verrauschte und korrelierte Beobachtungen simulieren soll. Durch die Addition der drei Komponenten und 800 zufällig gleichverteilt erzeugten Stützstellen können so Beobachtungswerte abgegriffen werden (s. Abbildung 5.3(b)). Die Beobachtungen wurden außerdem mit einer Standardabweichung von $0.3 \text{ [}^\circ\text{]}^2$ über die gesamte Zeitreihe normalverteilt verrauscht (der Amplitudenausschlag ist der Abbildung zu entnehmen). Zudem wurde die Zeitreihe bezüglich der Referenzspur in zwei Hälften geteilt, die jeweils für ein Extrem bezüglich der Ausprägung der Bojenspur stehen. Die erste Hälfte der Zeitreihe wurde mit einer Frequenz belegt, die nahe an der maximal noch zu modellierenden liegt. Sie steht für einen hohen Grad an Bewegung der Boje in der betrachteten Richtung. Die zweite Hälfte zeigt eine Frequenz deren Periode sich über diese komplett erstreckt. Sie steht für wenig Bewegung der Boje in der betrachteten Richtung.

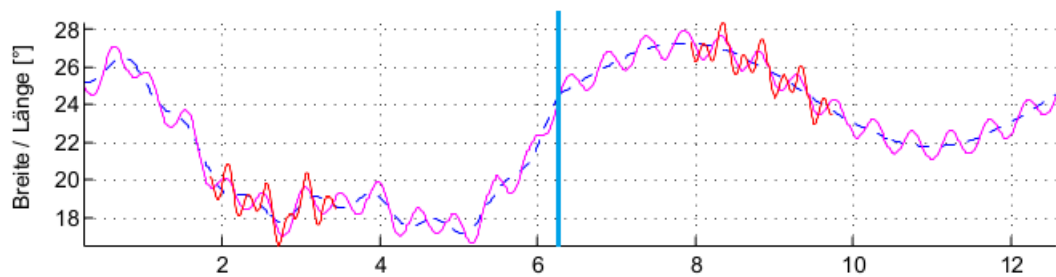
Die Zeitreihe wird einmal komplett geschätzt und einmal die beiden Hälften separat. Die Einstellung des MARS-Algorithmus via **Case** wurden wie in Abschnitt 5.4 aufgeführt, festgelegt. Während sich im Ergebnis des Ausgleichs kein sichtbarer Unterschied zeigt, weshalb auch nur das Ergebnis der kombinierten Zeitreihe dargestellt ist (s. Abbildung 5.3(c)), so ist vor allem in der rechten Hälfte der Zeitreihe, also in jener mit der niedrigen Frequenz, die Wahl der Knoten leicht unterschiedlich. Dies lässt sich vorrangig durch die etwas andere Lage der initialisierten Knoten begründen (s. um den Tag 9). Unterschiede bezüglich der Knotenwahl zeigen sich in dem Bereich

vom 11. bis 13. Tag zwischen der Einzel- und Kombinationsauswertung. Da das Ergebnis eine leichte Oszillation der geschätzten Spur zeigt, scheint es nicht unwesentlich, ob die Beobachtungen global beziehungsweise über einen größeren Bereich korreliert verrauscht sind. Es ist vor allem von der Lage der Knoten und dem Verhältnis vom stochastischen Signal zum deterministischen Signal abhängig, ob der MARS-Algorithmus dieses additive Signal (so wird es von FRIEDMAN 1991 genannt) ignoriert oder findet. Der MARS-Algorithmus ist eigentlich für die Schätzung von Zeitreihen, deren Beobachtungen mit normalverteiltem Rauschen belegt sind, ausgelegt. FRIEDMAN (1991, S. 37) zeigt, dass eine Anpassung des Algorithmus an additive Daten, also solche mit einem zusätzlichen Signal wie einer mit Korrelation belegten Spur, durch die Anpassung des Parameters d in der Kostenfunktion des GCV-Kriteriums gelingen kann. Der Parameter müsste mit $d = 2$ gewählt werden. Die Wahl des GCV mit $d = 2$ ist voreingestellt und kann durch den Verwender durch `Case.LOFcase = 1` angewählt werden. Tests zeigten, dass durch die Verwendung des GCV mit $d = 0$ das gleiche Ergebnis erhalten wird, wie mit dem AIC. Bereits mit $d = 1$ sortieren sich die Knoten sechs bis acht der zweiten Hälfte der synthetischen Zeitreihe aus. Mit der empfohlenen Einstellung $d = 2$ eliminieren sich zudem der sechste Knoten der ersten Hälfte und der dritte beziehungsweise vierte Knoten der zweiten Hälfte, also jene in den stark durch die Störfrequenz beeinflussten Bereichen.

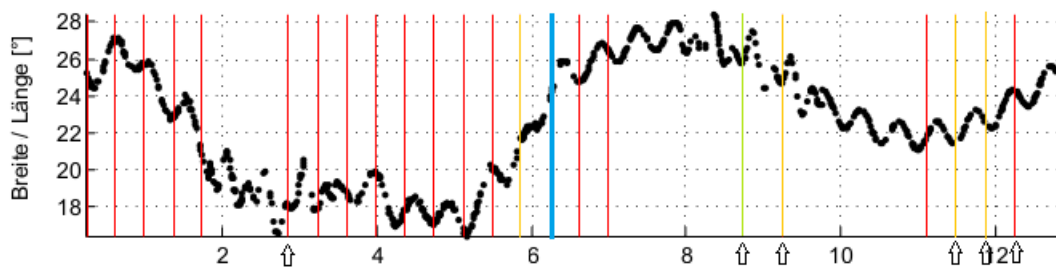
Das ARGOS-System, welches die Bojenspuren misst, hat eine Umlaufzeit von 100 Minuten (JOSEPH 2014, S. 191), sodass die entstehende Frequenz korrelierter Beobachtungen deutlich unter der maximal zu erfassenden liegt und damit vergleichbar mit der hier angebrachten Störfrequenz oder global höheren Frequenz (magenta) ist. Die Genauigkeit der durch das ARGOS-System erfassten Positionen wird mit einer einfachen Standardabweichung von 240m angegeben, sodass die Amplitude der Störfrequenz stark überspitzt ist, wenn die Beobachtungswerte in geografischen Breiten- oder Längenangaben verstanden werden. Einzig die Tatsache, dass der Algorithmus in solchen Bereichen signifikante Knoten findet, führt dazu, dass das Modell nicht richtig minimiert wird. Die Steifheit der Funktion, welche sich aus der zweifach stetigen Differenzierbarkeit ergibt, sowie eine relativ hohe Zahl an Beobachtungen in den Intervallen, verhindert hier eine starke Oszillation des geschätzten Verlaufs.

Die Störfrequenz wurde eingebracht um die Robustheit des Verfahrens gegenüber lokal stark verrauchten und korrelierten Daten überprüfen zu können. Auch die höhere Frequenz die über das globale Zeitintervall gelegt ist, zeigt eine hohe Amplitude. Trotz dieser, den Verlauf stark verfälschender Frequenzen zeigt die geschätzte Spur eine sehr gute Anpassung an die Referenzspur. Der MARS-Algorithmus eliminiert an den richtigen Stellen Knoten und der steife Verlauf der Funktion schafft es, in doch zu viel gesetzten Knoten ein geringes Maß an Oszillation zu zeigen.

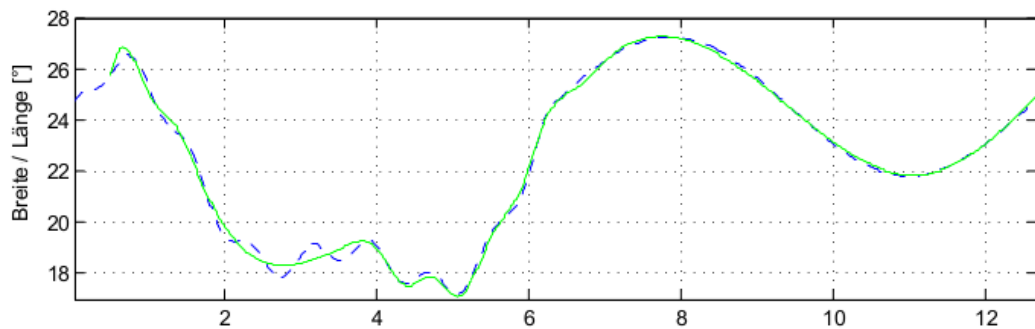
Werden wenige Beobachtungen angesetzt, so werden die Abstände zwischen den Knoten aufgrund der mindestens sechs Beobachtungen pro Intervall vergrößert (s. Abbildung 5.3(d)). Es zeigt sich die Wirkweise des Informationskriteriums, hier des AIC. In Bereichen in denen sich die Verbesserungsquadratsumme nicht soweit verringern lässt, dass sich die Einbringung eines weiteren Knotens rechtfertigen würde, werden geglättet. Die Hinzunahme des Knotens würde zum Anstieg des Gütekriteriums führen, was durch die Minimalanforderung vermieden wird.



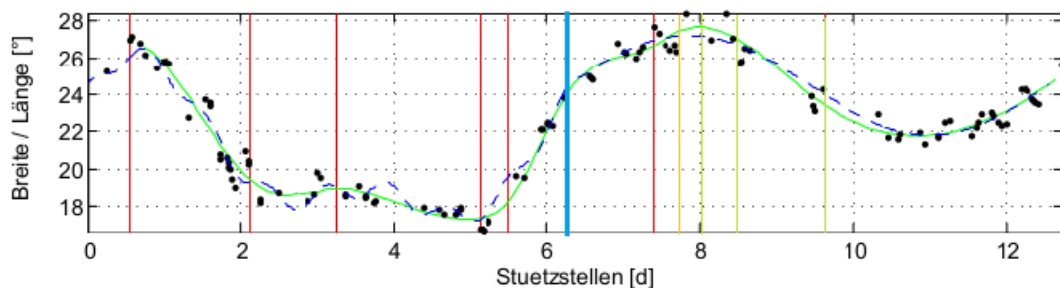
(a) Addition vor Fourierreihen (Referenzspur/Periode $\tau = 0.78/6.28$:blau , hohe Frequenz/Periode $\tau = 0.48$: magenta, Störfrequenz/Periode $\tau = 0.25$: rot)



(b) 800 simulierte Beobachtungen: schwarze Punkte, und gesetzte Knoten mit einem Mindestabstand von 0.375d (noch zu erfassende Periode $\tau = 0.75$)(nur in der Einzelauswertung der beiden Hälften: grün, nur in der Kombinationsauswertung: orange, in beiden Auswertungen: rot, durch GCV-Kriterium verworfene Knoten: Pfeile)



(c) ausgeglichene Bojenspur: grün, Referenzspur: blau gestrichelt



(d) Spurschätzung durch eine geringere Anzahl an Beobachtungen (120 Beobachtungen: schwarz, ausgeglichene Bojenspur: grün, Referenzspur: blau gestrichelt, aufgenommene Knoten: rot, gesetzte Knoten nur in der Einzelauswertung der beiden Hälften: grün, nur in der Kombinationsauswertung: orange)

Abbildung 5.3: Synthetisches Beispiel einer Bojenspur zur Approximation mittels MARS-Algorithmus und B-Splines, blauer Strich teilt die Extrema der Referenzspuren (links hohe, rechts kleine Frequenz)

Das Verhältnis, wie oft die Vorwärtssuche und wie oft die Rückwärtssuche zur Wahl der signifikanten Punkte genutzt wird, ist stark von den ausgewerteten Spuren abhängig. Ein Testdurchlauf zeigte, dass bei 150 Entscheidungen nur 6-mal die Vorwärtssuche gewählt wurde. Um herauszustellen, inwieweit die Rückwärtssuche die Anzahl der Knoten beschneidet und damit die Überanpassung durch die Vorwärtssuche verhindert, wird eine beispielhafte Spurschätzung in Abbildung 5.4 dargestellt. In der Vorwärtssuche wurden in diesem Teil der Zeitreihe zwei Knoten mehr als in der Rückwärtssuche als signifikant gesetzt. Der AIC zeigt sich für die Rückwärtssuche als minimal, sodass die Verkürzung der Knotenanzahl durch die Rückwärtssuche Wirkung zeigt.

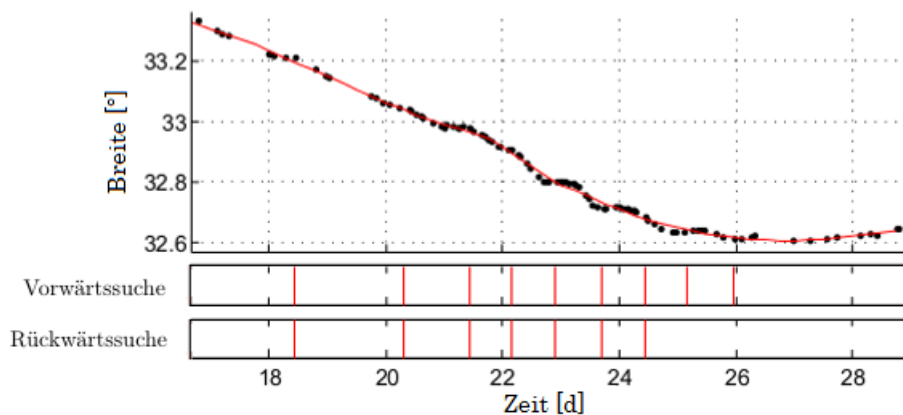


Abbildung 5.4: Bojenpositionen in der Breite aus dem Jahr 2009 (schwarze Punkte), die approximierte Bojenspur (rot) und deren Knotendefinition (Rückwärtssuche zeigt kleineren AIC)

Die Darstellung zeigt auch einen weiteren Effekt. Die Mindestanzahl an Beobachtungen je Intervall wurde hier auf acht statt der als Standard empfohlenen sechs Beobachtungen gesetzt. So verschwinden Frequenzen, welche eigentlich noch gefunden werden sollten, wie zwischen dem 22. und 24. Tag. Es sei darauf hingewiesen, dass diese Spur relativ wenige Beobachtungen pro Zeiteinheit aufweist. Die Abbildung 5.5 zeigt die sich verbessernde Anpassung an die Bojenpositionen, je weniger Beobachtungen als Mindestanzahl pro lokalem Intervall gesetzt werden. Das Heruntersetzen der Beobachtungsanzahl auf vier würde die Spur sehr empfindlich im Bezug auf Ausreißer machen.

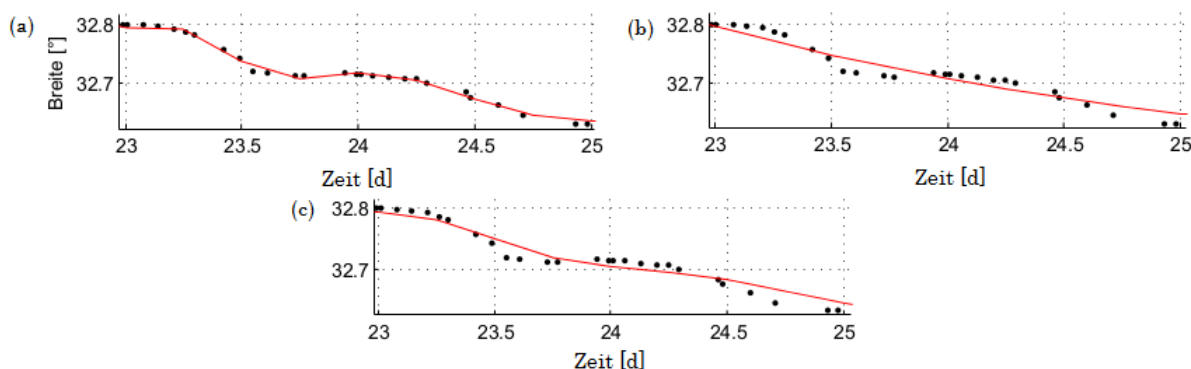


Abbildung 5.5: Approximation der Bojenspur durch eine unterschiedliche Mindestanzahl an Beobachtungen pro lokalem Intervall, a) 4, b) 8, c) 6 Beobachtungen

Es sei damit bewiesen, dass bei solchen Verläufen nicht der MARS-Algorithmus zu einer falschen Entscheidung kommt, was die Aussonderung von Knoten angeht, sondern eine Sicherungsschranke.

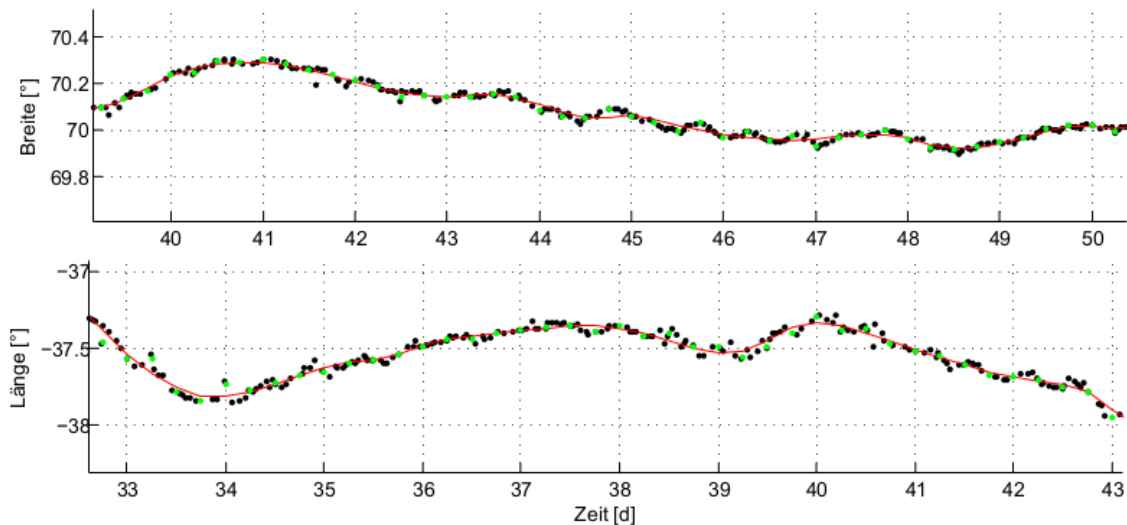


Abbildung 5.6: Ausschnitt einer geglätteten Trajektorie über die Schätzung von B-Splines (rot), Roh-Datenpunkte (schwarz) und die durch das SVP-Kriging interpolierten Punkte (grün), oben: Breitengrade, unten: Längengrade

Diese setzt eine Mindestanzahl von sechs Beobachtungen je Intervall an, was zu einem Informationsverlust führen kann. Ist davon auszugehen, dass der verwendete Datensatz von Ausreißern befreit ist, so ist es natürlich möglich diese Schranke herunterzusetzen. Es zeigt sich jedoch, dass auch editierte Daten des SVP Ausreißer enthalten, welche die ausgeglichene Spur bei wenigen Beobachtungen (< 6) innerhalb eines Intervalls verziehen können.

Die Abbildung 5.6 zeigt die geglättete Trajektorie durch Schätzung mit B-Splines, sowie die Roh-Daten-Positionen und die durch das SVP-Kriging bestimmten geographischen Koordinatenwerte. Zum Einen fallen die ignorierten Frequenzen beispielsweise in der Breitenspur zwischen dem 44. und 46. Tag auf. Außerdem bildet sich ab, in wie weit die Kriging-Daten die Spur glätten, denn zum Beispiel in der Längenspur zwischen dem 33. und 34. Tag folgen diese vor allem der verrauschten Spur. Die Lage der Roh-Datenposition, welche der zu interpolierenden Position (Kriging-Position) am nächsten liegt, hat größten Einfluss auf den interpolierten Koordinatenwert, sodass die übrigen neun Beobachtungswerte, welche zur Interpolation eines Koordinatenwertes einfließen, kaum Bedeutung zugerechnet wird. Die Interpolation durch das Kriging, welche durch das SVP angewandt wird, zeigt sich demnach anfällig für Ausreißer. Gerade diese fehlerhaft bestimmten Positionen und die ungenaue Bestimmung der Geschwindigkeiten über den zentralen Differenzenquotienten führen zu den in Abbildung 5.7 dargestellten Geschwindigkeiten in grün. Es sind die östlichen Geschwindigkeiten, welche von den Längengraden der darüber abgebildeten Spur abgeleitet werden. Dass die Geschwindigkeiten, welche über die beiden Verfahren bestimmt werden, einen solch unterschiedlichen Verlauf nehmen, ist durch die darin jeweils enthaltenen unterschiedlichen maximalen Frequenzen zu erklären, aber im Bereich vom 23. bis 24. Tag auch durch das Rauschen. Die Kriging-Daten folgen hier ebenfalls eher den verrauschten Positionen, anstatt diese zu glätten. Besonders im Zeitraum zwischen dem 18. und 20. Tag ist die zu gering bestimmte Geschwindigkeit aufgrund des zentralen Differenzenquotienten auffällig, denn dieser bestimmt die Geschwindigkeit mit einer Fehlerordnung

von $O(\Delta\lambda^2)$ beziehungsweise $O(\Delta\phi^2)$, was ihn besonders fehleranfällig in Bereichen hoher Dynamik werden lässt.

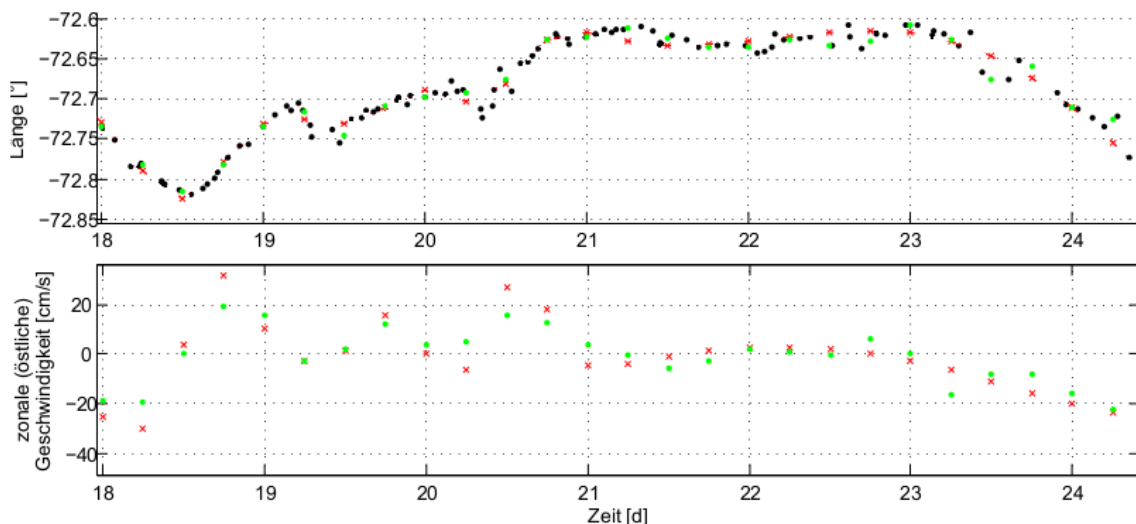


Abbildung 5.7: Ausschnitt einer geglätteten Trajektorie über die Schätzung von B-Splines (rot, mit roter Linie verbunden), Roh-Datenpunkte (schwarz) und die durch das SVP-Kriging interpolierten Punkte (grün), oben: Längengrade, unten: östliche Geschwindigkeiten

Wie die Abbildung 3.8 zeigt, weisen die Kriging-Daten des SVP teilweise unrealistische Genauigkeitsangaben auf, welche zu Geschwindigkeitsgenauigkeiten größer als $1000 \frac{\text{cm}}{\text{s}}$ führen. Der gleiche Zeitraum wurde auch bezüglich der Geschwindigkeiten und deren Genauigkeit durch das hier vorgestellte alternative Verfahren untersucht, wobei sich das Ergebnis in Abbildung 5.8 zeigt. Maximale Werte für die Geschwindigkeitsgenauigkeit finden sich bei $26 \frac{\text{cm}}{\text{s}}$ in den zonalen Geschwindigkeiten und dies nur für Bojen, die ihren Treibanker verloren haben. Dies lässt sich durch das mit 12 Tagen sehr groß gesetzte Fenster für minimal 6 Beobachtungen begründen. Lücken in den Daten sind meist bei beschädigten Bojen, also solchen ohne Treibanker zu finden. Ein kleineres Fenster würde zum schnelleren Aussetzen der Schätzung der Zeitreihe in diesen Bereichen führen und so diese Spurabschnitte für den endgültigen Datensatz nicht interpolieren. Als zweite Begründung für die schlechten Genauigkeiten der Bojenpositionen ohne Treibanker ist eventuell auch ein erhöht verrauschter Datensatz. Insgesamt zeigt das alternative Verfahren im Vergleich zu den Daten des SVP-Krigings verlässlichere Genauigkeiten bezüglich der Geschwindigkeiten. Hier ist der empfohlene vorsichtige Umgang beziehungsweise die notwendige Einbeziehung der Genauigkeitsinformation in Anwendungen mit SVP-Daten aufgrund Genauigkeiten größer als $1000 \frac{\text{cm}}{\text{s}}$ nicht mehr so bedeutsam.

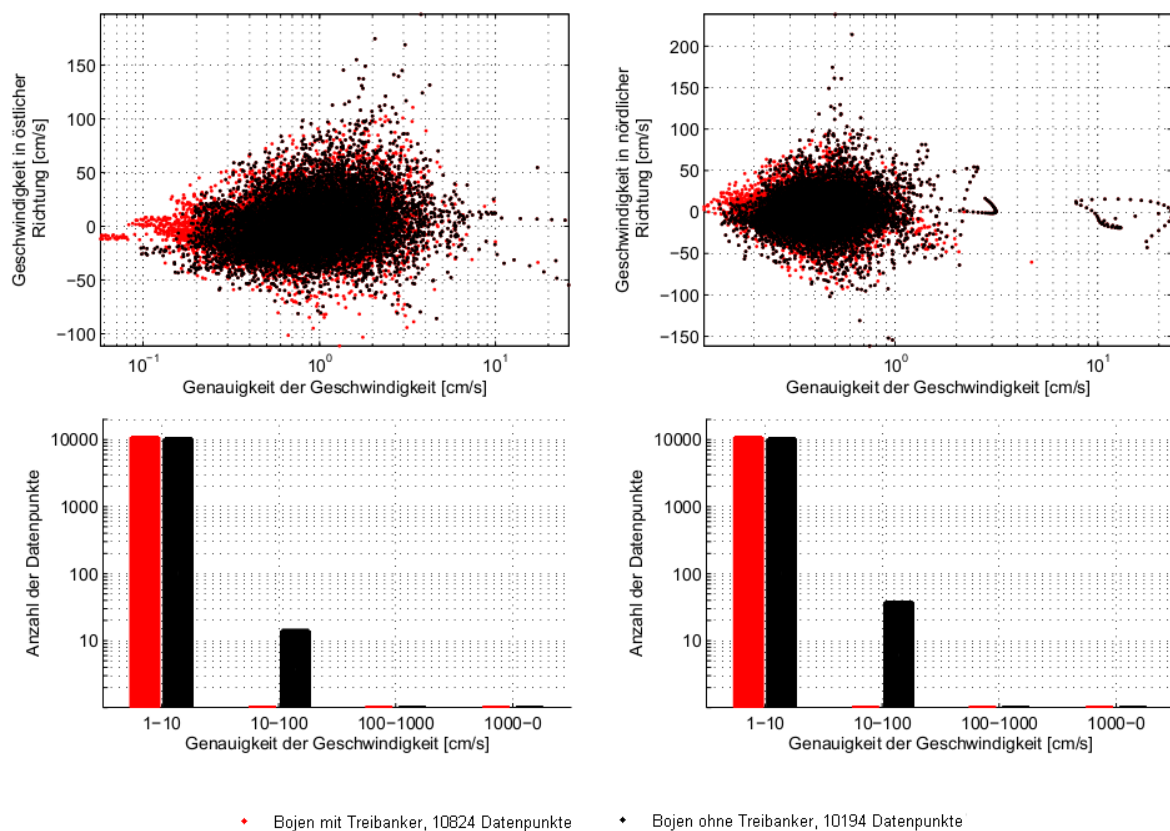


Abbildung 5.8: Oben: Genauigkeit der Geschwindigkeit gegen die Geschwindigkeit selbst aufgetragen, Unten: Histogramme für die Dichteverteilung der Datenpunkte über die Geschwindigkeitsgenauigkeiten für den Januar 2009 im Nord-Atlantik, links: meridionale Geschwindigkeiten, rechts: zonale Geschwindigkeiten, separate Betrachtung der Bojen mit und ohne Treibanker

5.6 Diskussion

Aufgrund der Einteilung in 15-Tagessequenzen werden maximal 24 Knoten in einer Zeitreihe von einem Jahr nicht auf Signifikanz untersucht. In einer solchen Jahreszeitreihe werden maximal 973 Knoten angelegt, sodass dieser Kompromiss bezüglich der Zeitreihe als nicht schwerwiegend erscheint. Dadurch wird hingenommen, dass zu diesen 24 Zeitpunkten der B-Spline mehr oszilliert, als die eigentlich zugrundeliegende Trajektorie. Eine Verbesserung könnte das Einbringen einer weiteren Variante des MARS-Algorithmus bringen, der nur über diese 24 Knoten läuft. Die Variante müsste bedingt durch die Länge der Zeitreihe auf Grundlage der B-Splines arbeiten, also auch in den inneren For-Schleifen (s. Algorithmus 1) mittels AIC Entscheidungen treffen und aufgrund dessen einen vollständigen Ausgleich je Durchlauf ausführen. Dies würde zu maximal 276 Ausgleichen je Suche führen, was aus Fragen der Effizienz und des nur bedingten Informationsgewinns hier unterlassen wurde. Eine weitere Möglichkeit stellt das Lösen der Normalgleichungen durch leichte Modifikation nach FRIEDMAN (1991, S. 32), wodurch bessere numerische Eigenschaften bezüglich der trunkierten Potenzen erreicht wird und die Einteilung in Sequenzen überflüssig werden könnte. Das Verfahren kann außerdem durch die Lösung der Normalgleichungen über die QR-Zerlegung weiter optimiert werden (FRIEDMAN 1991, S. 29). Zudem bietet der MARS-Algorithmus einen hohen Grad an Parallelisierung, sodass eine schnellere Auswertung trotz großer Datensätze möglich ist (FRIEDMAN 1991, S. 31).

Für die Schätzung der Positionen und Geschwindigkeiten können Roh-Daten genutzt werden. Die in Abschnitt 3.4.1 beschriebene Vorverarbeitung der Daten wird überflüssig, da Ausreißer durch die meist hohe Zahl an Beobachtungen je lokalem Intervall bei einer zeitlichen Ausdehnung von 9 Stunden beispielsweise im Jahr 2009 bei durchschnittlich 15 Positionen liegt, aber auch mindestens bei sechs, sodass ein einzelner Ausreißer den Verlauf nicht stark verzieht. Zudem sind auch in den editierten Daten des SVP noch Ausreißer zu finden. Pro Knoten, ausgenommen derjenigen am Rand, ist nur je ein Parameter zu bestimmen, sodass es zu gut kontrollierten Beobachtungen gemäß der Redundanzanteile kommt (s. Anhang A.1). Probleme bereiten Ausreißer in den Rändern der Zeitreihe, da sie dort am ehesten in den Modellverlauf greifen. FRIEDMAN (1991, S. 64) schlägt bezüglich Ausreißern robustere Verfahren anstelle der Kleinste-Quadrate-Schätzung vor. Vorteil der stückweise kubischen Polynome ist es aber auch, dass Ausreißer nur lokale Auswirkung auf die geschätzte Spur zeigen.

Genutzt werden dennoch die Positions- und Sensordaten, da von diesen trotz Editierung mehr Datensätze zugänglich sind (<http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/drib-bder/svp-vcs/res/types-eng.htm>). Außerdem liegen Roh-Datenspuren vor, deren ARGOS-ID nicht in den Meta-Daten vorliegen. Sie sind auch nicht in der speziell ausgegebenen Text-Datei enthalten, welche den Wechsel von ID-Nummer aufzeichnet. Dadurch liegen in dem Roh-Datensatz Spuren vor, deren Treibankerverlust nicht verzeichnet ist, sodass sie für den hiesigen Anwendungsfall verworfen werden würden.

Durch dieses Auswerteverfahren können rein rechnerisch verlässlichere Geschwindigkeiten geschätzt werden. Die Positionsbestimmung ist zudem unabhängig von den empirisch bestimmten und nur bedingt global anwendbaren Variogrammen. Zudem gelingt es auf diese Weise die Genauigkeit und Korrelation der Positionen und Geschwindigkeiten durch eine beides umfassende Kovarianzmatrix in die nachfolgenden Auswertungen zu übernehmen. Da durch das Satellitensystem ARGOS bestimmten Positionsdaten von korreliert verrauschten Beobachtungen auszugehen ist, wird nachfolgend das GCV-Kriterium mit $d = 2$ (s. letzten Abschnitt) über `Case.LOFcase = 1` als Standard zur Interpolation genutzt.

6. Reduktion der Bojendaten um windgetriebene Geschwindigkeiten

Eine Boje erfasst mit ihrer Geschwindigkeit die großskalige geostrophische Strömung in 15 m Tiefe, zusätzlich aber auch ageostrophische Komponenten, wie die hochfrequenten Strömungen beispielsweise durch Gezeiten und die windgetriebenen Effekte, wie den Ekman-Transport und den Schlupf, welcher im Besonderen die Bojen ohne Treibanker betrifft. .

Die Geschwindigkeit \mathbf{u}_b , welche über die Bojenpositionen bestimmt wird, setzt sich vereinfacht wie folgt zusammen

$$\mathbf{u}_b = \mathbf{u}_g + \mathbf{u}_s + \mathbf{u}_{ek} + \mathbf{u}_a, \quad (6.1)$$

mit \mathbf{u}_g der geostrophischen Geschwindigkeit, der Geschwindigkeit \mathbf{u}_s , welche sich direkt durch den windgetriebenen Schlupf ergibt, sowie der hochfrequenten ageostrophischen Geschwindigkeit \mathbf{u}_a und dem Ekman-Transport \mathbf{u}_{ek} .

Die geostrophische Geschwindigkeit bildet die Zielgröße. Sie beschreibt die Zirkulationsmechanismen der Ozeane, zudem ist es möglich über das geostrophische Gleichgewicht die absolute dynamische Topographie abzuleiten, was einen Vergleich mit den Modellen der Satellitenaltimetrie ermöglicht. Um die weiteren auf die Bojengeschwindigkeit wirkenden Effekte eben daraus zu eliminieren, müssen Annahmen getroffen werden. Zum einen bezüglich dem windgetriebenen Schlupf welcher der Windgeschwindigkeit entsprechend anteilmäßig von der Bojengeschwindigkeit abgezogen wird. Außerdem der Effekt des Ekman-Transports, welcher eine Modellierung benötigt, da er nicht in Windrichtung wirkt und von den Breitengraden abhängig eine unterschiedliche Korrektur benötigt. Die hochfrequente ageostrophische Strömung benötigt hier keine spezielle Behandlung, da die spätere Strömungsmodellierung

Dieses Kapitel soll eine Möglichkeit aufzeigen, auf Grundlage von Literaturangaben und eigens geschätzten Parametern, windgetriebenen Effekte aus den Bojengeschwindigkeiten zu eliminieren. Die Modellierung zielt dabei speziell auf den Nordatlantik. Eine Übertragbarkeit auf andere Regionen ist nur bedingt gegeben.

6.1 Der windgetriebene Schlupf

Der windgetriebene Schlupf, im Englischen als Slip bezeichnet, ist ein die Boje in Windrichtung treibender Effekt. Er resultiert aus einer durch den Wind erzeugten Reibung an dem kugelförmigen Schwimmer und eine auf das Halteseil des Trichters und eben auch auf den Treibanker wirkende Scherkraft. Mit einem Reibungsflächenverhältnis von 40 : 1 besitzt eine Boje des SVP, welche mit einem Treibanker versehen ist, einen Schlupf von 0.07 % der Windgeschwindigkeit bis $10 \frac{\text{m}}{\text{s}}$ (SIEDLER et al. 2013, S. 285; LUMPKIN und PAZOS 2005, S. 4; NIILER und PADUAN 1995, S. 2821; NIILER et al. 1995, S. 1964). Dies ist verschwindend gering und wird daher hier vernachlässigt. Das

Reibungsflächenverhältnis setzt die Fläche des Schwimmers in das Verhältnis mit der Fläche aller anderen unter der Wasseroberfläche liegenden Objekte der Boje, was bei einer SVP-Boje vor allem der Treibanker ist. Verliert eine Boje den Treibanker, so wird das Reibungsflächenverhältnis deutlich schlechter und die Boje wird anfälliger für den windgetriebenen Schlupf. RIO 2012 benennt bei Bojen ohne Treibanker ein Verhältnis von 0.6 % der Windgeschwindigkeit bis $10 \frac{\text{m}}{\text{s}}$. Dies ist nicht zu vernachlässigen. Bei Windgeschwindigkeiten über $10 \frac{\text{m}}{\text{s}}$ ist der Effekt größer, sodass er im Antarktischer Zirkumpolarstrom (engl. ACC) sein Maximum von 1.3 % der Windgeschwindigkeit erreicht. Die nachfolgende Betrachtungen und Modellierungen sind auf den Nordatlantik zugeschnitten, sodass spezielle Anpassungen für den Äquator oder für die Südhalbkugel entfallen und der Wert von 0.6 % der Windgeschwindigkeit zur Modellierung des Effekts verwendet wird. Dies ist kongruent mit dem Vorgehen für die Verarbeitung des neuen Modells CNES-CLS13 der MDT.

Zur Schlupf-Korrektur der Geschwindigkeiten von Bojen ohne Treibanker werden die ERA-Interim Windgeschwindigkeiten in 10 m über der Meeresoberfläche \mathbf{u}_{10} mit einer zeitlichen Auflösung von sechs Stunden des ECMWF genutzt (http://apps.ecmwf.int/datasets/data/interim_full_daily/). Wird der Schlupf mit

$$\mathbf{u}_s = 0.006\mathbf{u}_{10} \quad (6.2)$$

von den Bojengeschwindigkeit subtrahiert, so wird die Änderung der Windgeschwindigkeit aus 10 m Höhe zur Ozeanoberfläche hin vernachlässigt. Dies ist eine gängige Methode, da die Windgeschwindigkeiten relativ zu den Strömungsgeschwindigkeiten des Ozeans sehr groß sind (DUHAUT und STRAUB 2005). Eine Änderung der Windgeschwindigkeit in Richtung der Ozeanoberfläche ist abhängig von der Schubspannung, die durch die relative Bewegung zwischen dem Ozean und der Atmosphäre entsteht (PACANOWSKI 1987) und da dieses Verhältnis mit circa 1 % (SCHÖNWIESE 2003, S. 194) wie erwähnt vom Ozean zur Atmosphäre schwach ist, kann die Änderung weitestgehend vernachlässigt werden.

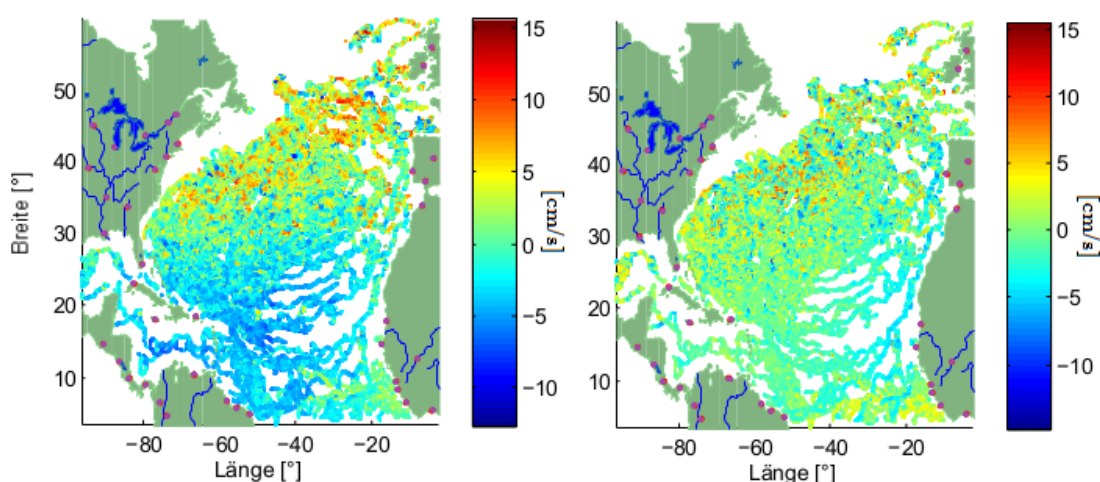


Abbildung 6.1: Anteil des windgetriebenen Schlupfes an den Bojengeschwindigkeiten für den Januar 2009, links für die zonalen Geschwindigkeiten, rechts für die meridionalen Geschwindigkeiten

Die Abbildung 6.1 zeigt den Anteil des windgetriebenen Schlupfes an den Bojengeschwindigkeiten für den Januar 2009. Nur in seltenen Fällen erreichen die Windgeschwindigkeiten mehr als $10 \frac{\text{m}}{\text{s}}$,

sodass auch nur an vereinzelt Stellen der Effekt des Schlupfes größer als $6 \frac{\text{cm}}{\text{s}}$ ist. Da der Schlupf bei Windgeschwindigkeiten über $10 \frac{\text{m}}{\text{s}}$ nicht linear zur Windgeschwindigkeit steigt, wird in diesen Positionen eine zu geringe Korrektur hingenommen.

6.2 Modellierung des Ekman-Transports

Die durch die Ekman-Strömung induzierte Geschwindigkeit in der Tiefe z wird als Ekman-Transport bezeichnet. Auch als Korkenzieherströmung bekannt, wirkt Sie mit der durch den Wind induzierten Scherkraft auf die Oberfläche spiralförmig mit der Tiefe. In mittleren Breiten hat Sie bis in Tiefen von 45 bis 180 Metern je nach Windgeschwindigkeit einen Einfluss auf die Strömung (STEWART 2008, S. 139). Die Bojengeschwindigkeit repräsentiert die gemittelte Bewegung in einer Oberflächenschicht des Meeres mit der Dicke h , wodurch sich nach LAGERLOEF et al. (1999, S. 23315)¹ über die lineare stetige Impulsbilanz folgende Gleichung bezüglich der Bojengeschwindigkeit (u_b, v_b) aufstellen lässt

$$-fhv_b = -gh \frac{\partial \zeta}{\partial x} + \frac{\tau_x}{\rho_w} - ru_{ek} \quad \Rightarrow \quad -fh(v_g + v_{ek}) = -gh \frac{\partial \zeta}{\partial x} + \frac{\tau_x}{\rho_w} - ru_{ek}, \quad (6.3)$$

$$fhu_b = -gh \frac{\partial \zeta}{\partial y} + \frac{\tau_y}{\rho_w} - rv_{ek} \quad \Rightarrow \quad fh(u_g + u_{ek}) = -gh \frac{\partial \zeta}{\partial x} + \frac{\tau_y}{\rho_w} - rv_{ek}. \quad (6.4)$$

Der erste Summand stimmt mit der geostrophischen Geschwindigkeit übereinstimmt, der zweite entspricht dem Anteil des Ekman-Transports an der Bojengeschwindigkeit mit der Dichte von Wasser ρ_w , τ_x und τ_y , den Einträgen des Windschubspannungsvektors, sowie einem linearen Reibungskoeffizienten r . Wird nun die geostrophische Komponente herausgenommen (s. Gleichung (2.6) und 2.7) und die Gleichung (6.4) mit der komplexen Zahl i multipliziert und zu Gleichung (6.3) addiert, so ergibt sich für den Anteil des Ekman-Transports an der Bojengeschwindigkeit u_{ek}^c

$$u_{ek}^c = \frac{1}{ifh + r} \tau^c \quad (6.5)$$

wobei das c für die komplexe Schreibweise steht mit

$$u_{ek}^c = u_{ek} + iv_{ek}, \quad \text{und} \quad \tau^c = \frac{\tau_x + i\tau_y}{\rho_w} \quad (6.6)$$

mit τ^c der komplexen kinetischen Windschubspannung.

Der Ekman-Transport wirkt mit der Tiefe spiralförmig (s. Abbildung 2.3), sodass sein der Einfluss mit der Tiefe abnimmt und die Richtung ausgehend von 45° zur Rechten des Windes mit der Tiefe immer weiter abgelenkt wird.

¹Eine Alternative mathematische Beschreibung wird in dem technischen Vermerk, der der Daten-CD beigelegt ist aufgeführt. Da er in einer Kleinsten-Quadrat-Schätzung nicht konvergiert, wurde die Schätzung auf Grundlage dessen verworfen.

6.2.1 Eingangsgrößen des Ekman-Transports

Der Vektor der Schubspannung des Windes an der Meeresoberfläche $\boldsymbol{\tau}$ lässt sich über die Windgeschwindigkeit in 10 m über der Meeresoberfläche bestimmen

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_x \\ \tau_y \end{bmatrix} = \rho_l C_D \mathbf{u}_{10} |\mathbf{u}_{10}| \quad (6.7)$$

mit ρ_l der Dichte der Luft und C_D dem Reibungskoeffizient der Luft. Die Schubspannung ist wie die Windgeschwindigkeit dem ERA-Interim des ECMWF mit einer halbtäglichen Auflösung auf einem 0.75° -Gitter zu entnehmen und stellt die skaliert quadrierte Windgeschwindigkeit unter Beibehalten der Richtung dar.

Die Dichte des Wassers ρ_W ergibt sich über dessen Temperatur und Salzgehalt. Da Wasser nahezu inkompressibel ist und Tiefen bis maximal 15 m betrachtet werden, kann die Druckänderung über die Tiefe ignoriert und die Dichte für konstant 1013.25 hPa bestimmt werden. Eine Boje im offenen Ozean, die in einer Tiefe von 15 m zentriert ist, befindet sich vollkommen in der horizontalen gemischten Schicht. In dieser Schicht ist die Temperatur und der Salzgehalt nahezu konstant (SPINTALL und CRONIN 2001 S.3119). Sie lassen sich dem Weltozeanatlant WOA entnehmen, in dem beide Größen mit einer monatliche Auflösung vorliegen und in 24 Leveln für die Tiefe vorliegen. Diese Standard-Tiefen-Level erlauben es die Werte an der Meeresoberfläche für Bojen ohne Treibanker direkt abzugreifen (Level 1) und für die in 15 m zentrierten Bojen zwischen den Leveln 2 und 3, welche für die Tiefen 10 m und 20 m stehen, zu interpolieren. Für die hier vorgestellte Modellierung werden die Daten des WOA09 verwendet. Dieser wird von dem National Oceanographic Data Center (NODC) der Vereinigten Staaten zur Verfügung gestellt (www.nodc.noaa.gov/OC5/WOA09/pr_woa09.html) und nutzt Daten verschiedenster Beobachtungsarten unterschiedlichster Datenzentren, die hier nicht alle aufgelistet sein sollen. Durch das Aussortieren von doppelten Beobachtungssätzen und einem objektiven Analyseschema nach BARNES (1964) sowie weiterer Analysen (LOCARNINI et al. 2010) werden so jeweils auf die Monatsmitte zentrierte Datensätze gewonnen, die in Abbildung 6.2 beispielhaft für den Januar dargestellt sind.

Die Temperatur zeigt ein den Breitengraden entlang gestaffeltes Profil und besitzt einen größeren Einfluss auf die Variation der Dichte als der Salzgehalt. Sie wird, falls vorhanden, vorrangig direkt von der Bojenmessung genutzt, welche aus interpolierten Temperaturmessungen der Sensordaten abgeleitet werden. Die geglättete Temperaturkurve wird auf die gleiche Weise geschätzt wie die Positionsdaten durch das hier als Alternative zum SVP-Kriging vorgestellte Verfahren.

Der WOA gibt auch Standardabweichungen heraus, welche allerdings nur lückenhaft abgebildet werden. Es lässt sich aber abschätzen, dass die Genauigkeit dieser durchschnittlichen Werte mit einer räumlichen Auflösung von $1^\circ \times 1^\circ$ -Grad für die Salinität deutlich kleiner als 0.1PSU und für die Temperatur ebenfalls kleiner als 0.1°C ist. Die geglätteten Temperaturmessungen der Boje sind bezüglich der Standardabweichung etwas schlechter, aber durchgehend besser als 0.3°C und aufgrund der besseren räumlichen und zeitlichen Auflösung zu bevorzugen. Die Bestimmung des Salzgehaltes wird sich in den nächsten Jahren aufgrund der Satelliten-Mission SMOS sicherlich noch verbessern lassen (http://www.esa.int/Our_Activities/Observing_the_Earth/SMOS/).

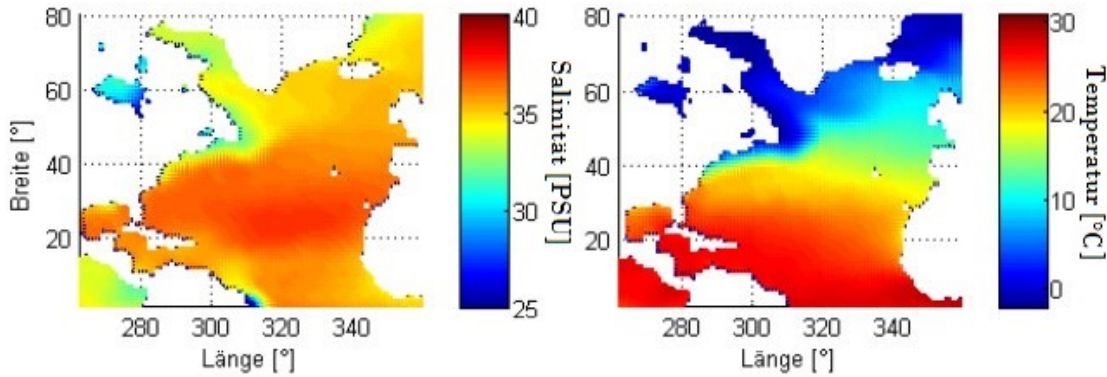


Abbildung 6.2: Verteilung des durchschnittlichen Salzgehaltes (links) und der Temperatur (rechts) im Nordatlantik nach dem WOA09 für den Januar

Die Dichte lässt sich nach MILLERO et al. (1980, S.257) über

$$\begin{aligned}
 \rho_w = & (0.999841594 + 6.793\,952 \times 10^{-5}t - 9.095\,290 \times 10^{-6}t^2 \\
 & + 1.001\,685 \times 10^{-7}t^3 - 1.120\,083 \times 10^{-9}t^4 + 6.536\,332 \times 10^{-12}t^5) \\
 & + (8.259\,17 \times 10^{-4} - 4.4490 \times 10^{-6}t + 1.0485 \times 10^{-7}t^2 \\
 & - 1.2580 \times 10^{-9}t^3 + 3.315 \times 10^{-12}t^4) s \\
 & + (-6.337\,61 \times 10^{-6} + 2.8441 \times 10^{-7}t - 1.6871 \times 10^{-8}t^2 \\
 & + 2.832\,58 \times 10^{-10}t^3) s^{3/2} + (5.4705 \times 10^{-7} - 1.979\,75 \times 10^{-8}t \\
 & + 1.6641 \times 10^{-9}t^2 - 3.1203 \times 10^{-11}t^3) s^2 \text{ [g cm}^{-3}\text{]}
 \end{aligned} \tag{6.8}$$

berechnen. Bei einer Temperatur t mit $0 \leq t \leq 40$ °C und einem Salzgehalt s von $0 \leq s \leq 40$ PSU liegt die Standardabweichung dieser empirisch bestimmten Formel bei $3.3 \times 10^{-6} \frac{\text{g}}{\text{cm}^3}$.

Da Salzwasser keine Dichteanomalie wie Süßwasser aufweist gilt diese Formel mit geringem Genauigkeitsverlust auch für leicht abweichende Temperaturen.

LAGERLOEF et al. (1999, S. 23316) schätzt nun auf Grundlage der Gleichung (6.6) folgendes lineares Modell

$$u_e^c = a_e \tau^c \quad \text{mit } a_e = \frac{1}{ifh + r}, \tag{6.9}$$

mit dem komplexen Parameter a_e , dessen realer und imaginärer Anteil sich wie folgt ausdrücken

$$re(a_e) = \frac{1}{r}, \quad im(a_e) = \frac{1}{fh}. \tag{6.10}$$

In das dort geschätzte Modell fließt im Unterschied zu dem hier vorgestellten allerdings ein konstanter Wert für die Dichte von 1025 hPa ein.

6.2.2 Schätzung des Parameters a_e nach LAGERLOEF et al. (1999)

Zu Bestimmung des Parameters komplexen Parameters a_e aus den Bojengeschwindigkeiten müssen daraus zunächst die geostrophischen und hochfrequenten ageostrophischen Geschwindigkeiten, sowie der windgetriebene Schupf subtrahiert werden (siehe Gleichung (6.1)).

Zunächst müssen die Geschwindigkeiten der ADT, welche von der AVISO ausgegeben werden, auf die Bojenpositionen zeitlich und räumlich interpoliert werden, was über eine lineare Interpolation erfolgt (s. Abschnitt 4.1). Anschließend müssen diese von der Bojengeschwindigkeit subtrahiert werden. So reduziert sich die Bojengeschwindigkeit um den Anteil \mathbf{u}_g der geostrophischen Geschwindigkeiten. Die Geschwindigkeiten der Bojenpositionen ohne Treibanker werden zudem um den windgetriebenen Schlupf \mathbf{u}_s vermindert (s. Abschnitt 6.1). Zuletzt bleibt die hochfrequente ageostrophische Strömung übrig. RIO und HERNANDEZ (2004, S. 7) schlägt diesbezüglich eine Bandpass-Filterung vor, sodass das resultierende Signal maximale Information bezüglich dem Ekman-Transport enthält und die ageostrophischen Geschwindigkeiten in ihrer Amplitude vermindert sind. Eine von ihnen durchgeführte Spektralanalyse zeigte, dass der Fokus auf Perioden größer als das Maximum der täglichen Gezeiten (24 Stunden) und der Trägheitsperiode $2\pi f$ zu legen ist. Zudem lassen Sie Perioden bis 20 Tage passieren, um den hochfrequenten Anteil zu erhalten und den maximalen Zusammenhang zu den Winddaten herstellen zu können. Die Bandpassfilterung erfolgt über einen Tschebyscheff-Filter. Sowohl die um \mathbf{u}_g und \mathbf{u}_s reduzierten Bojengeschwindigkeiten, als auch die dazugehörigen Windschubspannungen werden als Zeitreihen der einzelnen Bojen so gefiltert.

LAGERLOEF et al. (1999, S. 23315) schätzt den komplexen Parameter a_e in 1° Breitenstreifen. Hier wird noch die zusätzliche Einteilung bezüglich der Bojen mit und solchen ohne Treibanker betrieben. Genauigkeiten werden nicht in die Bestimmung des Parameters einbezogen, da für die ADT sowie für den Vektor der Schubspannung keine Varianz-Kovarianz-Informationen vorliegen. Diese Stellgrößen haben allerdings einen deutlich größeren Einfluss auf die Bestimmung des Ekman-Transports als die Temperatur oder der Salzgehalt, sodass in der Integration der Genauigkeiten von Salinität und Temperatur nur wenig Sinn liegt.

Der Parameter a_e wird über ein lineares Gauß-Markoff-Modell KOCH (2004, S. 166f) über die Kleinste-Quadrate-Schätzung bestimmt, wobei die Designmatrix dem zu den reduzierten Bojengeschwindigkeiten \mathbf{u}_{ek}^c entsprechenden komplexen Wert der kinetischen Windschubspannung $\boldsymbol{\tau}^c$ zu ersetzen ist. Der geschätzte komplexe Parameter a_e ergibt sich so mit

$$\tilde{a}_e = \left(\boldsymbol{\tau}^{cT} \boldsymbol{\tau}^c \right)^{-1} \boldsymbol{\tau}^{cT} \mathbf{u}_{ek}^c \quad \text{mit} \quad \sigma_{a_e} = \left(\boldsymbol{\tau}^{cT} \boldsymbol{\tau}^c \right)^{-1} \quad (6.11)$$

6.2.3 Modell des Ekman-Transports

Für die Schätzung der Parameter a_e werden durch das hier vorgestellte alternative Verfahren Bojendaten der Jahre 2007 bis 2010 genutzt, welche Datensätze in 1° Breitengürtel unterteilt werden. Warum keine weitere Unterteilung in die Längengrade stattfindet, soll die Abbildung 6.3 zeigen. Die Variation in den Anteilen des Parameters a_e in den Längengrade ergibt sich hauptsächlich aufgrund sehr großer Standardabweichungen. Zu beachten ist, dass die Farbskalen zur Abbildung der

Standardabweichung nach oben abgeschnitten sind, sodass einige Standardabweichungen mit Größenordnungen von bis zu $15\,000 \frac{\text{s}}{\text{m}}$ vorliegen. Die Einteilung diesbezüglich erscheint als nicht sinnvoll, sodass eine räumliche Unterteilung der Daten nur in Breitengürtel stattfindet. Dazu werden die geschätzten Parameter der Längengrade in einem Ausgleich unter Einbringung der Genauigkeit in die Kovarianzmatrix je 1° Breitengürtel gemittelt.

In der Abbildung 6.4(a) sind die für die einzelnen Breitengrade geschätzten komplexen Parameter a_e für die reduzierten Bojengeschwindigkeiten der Jahre 2007 bis 2010 dargestellt. Aufgeteilt in imaginären und realen Anteil lassen sich die Werte mit denen von LAGERLOEF et al. (1999) (s. Abbildung 6.4(b)) vergleichen, welche den tropischen Pazifik in einem Breitenband von 25°S bis 25°N untersuchten. Während dazu alle Bojen in die Auswertung aufgenommen wurden, hat hier eine zusätzliche Einteilung in Bojen mit Treibanker und in solche ohne stattgefunden. Die Parameterschätzung sollte Unterschiede bezüglich dieses Faktors aufweisen, da die Bojen ohne Treibanker in einer abweichenden Art und Weise von der Ekman-Transports beeinflusst werden. Aufgrund der Zentrierung kurz unter der Ozeanoberfläche, anders als bei Bojen mit Treibanker, welche in 15 m zentriert sind, ist die Richtungsänderung geringer, die Amplitude jedoch größer.

Bojenpositionen nahe der Küste werden ausgeblendet, da für diese Positionen die Bestimmung der Geschwindigkeiten der MADT, sowie der Salzgehalte nicht möglich ist.

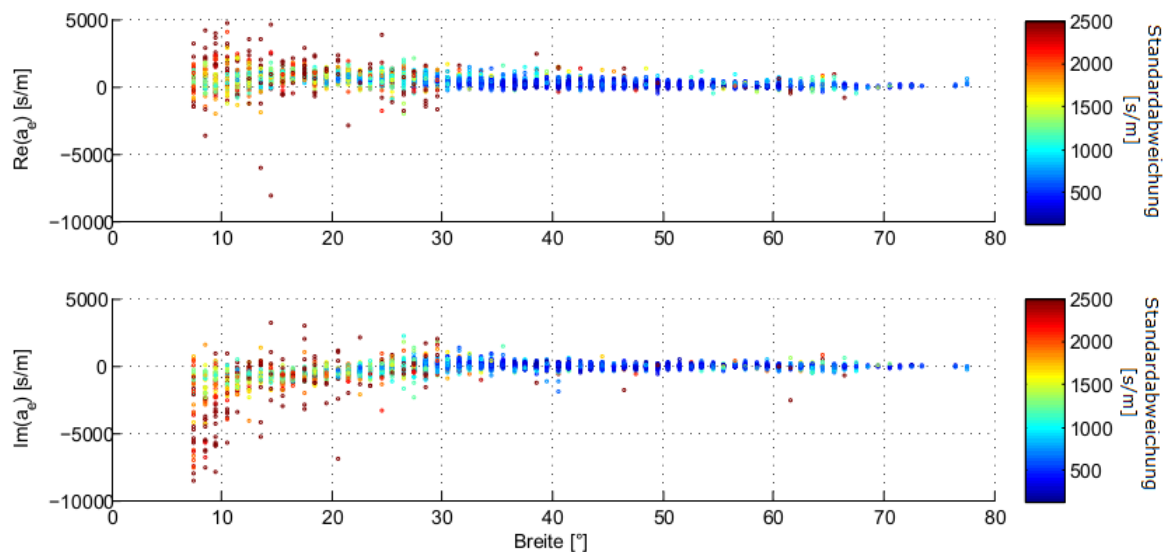


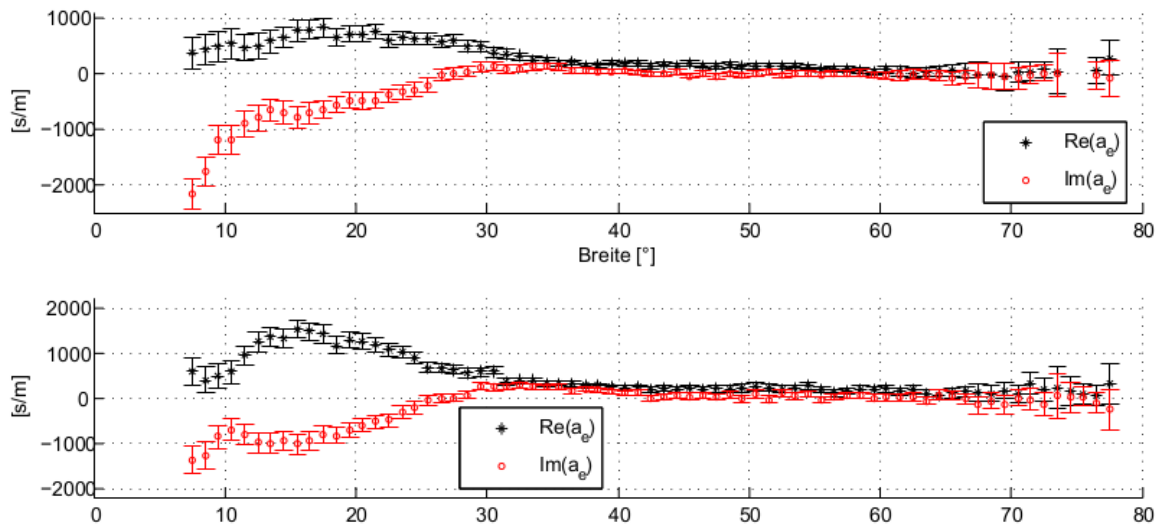
Abbildung 6.3: Geschätzter realer (oben) und imaginärer (unten) Anteil des Parameters a_e

Da die geschätzten Werte trotz unterschiedlicher Untersuchungsgebiete und Zeiträume und der unterschiedlich eingebrachten Dichte des Wassers augenscheinlich vergleichbare Ergebnisse liefern, wird, aufgrund des ausführlichen Vergleichs von LAGERLOEF et al. (1999, S. 23317) mit anderen Modellen, auf weitere Vergleiche an dieser Stelle verzichtet und die Parameter nicht noch in anderen Darstellungsformen wie der Amplitude und Phase dargestellt. Es wird auch darauf verzichtet ein Modell durch diese Parameter zu schätzen, sodass die hier dargestellten Werte des Parameters a_e in der nachfolgenden Strömungsmodellierung direkt zur Ableitung des Ekman-Transports genutzt werden. LAGERLOEF et al. (1999, S. 23323) gibt an, dass das von ihnen entwickelte Modell, welches

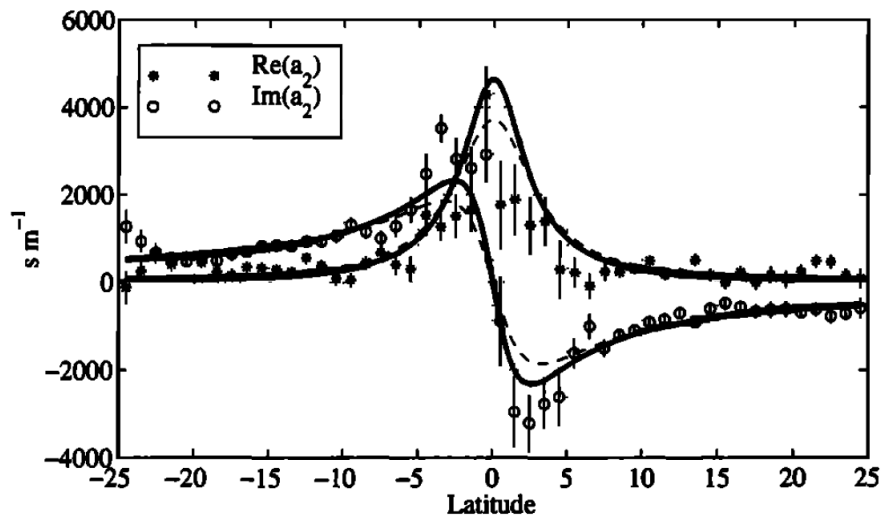
von den zwei Parametern (realer und imaginären Anteil) abhängig ist, sehr sensitiv bezüglich des realen Anteils ist. Für höherer Breiten ($> 25^\circ$) soll das Modell nur eingeschränkt nutzbar sein, auch weil der konstante Wert für den linearen Reibungskoeffizienten in das Modell einfließt und dieser sich höheren Breiten nicht durch einen linearen Verlauf vereinfachen lässt. Für den hier vorgestellten Anwendungsfall des Nordatlantiks müsste dementsprechend ein Modell entwickelt werden, welches mehr Information aus den Anteilen von a_e aufnimmt als nur die Steigung und eine Konstante. Die Abbildungen verweisen vor allem im realen Anteil auf einen abweichenden Verlauf, der sich nicht durch die Breitenabhängigkeit des Coriolisparameter beschreiben lässt.

Die Unterteilung in Bojen mit Treibankern und solchen ohne zeigt deutliche Unterschiede besonders für den realen Anteil, sodass für die nachfolgende Modellierung die Unterteilung beibehalten wird. Jahreszeitliche Unterschiede in der Parameterschätzung werden nur im geringem Maße erwartet, da die in die Bestimmung der kinetischen Windschubspannung bereits die Temperatur und der Salzgehalt in Form der Dichte eingeflossen sind. Und auch der zeitlich variable Windschubspannungsvektor enthält die zeitvariablen Effekte der Atmosphärischen Strömung. Die Abbildung 6.5 zeigt die Schätzergebnisse bezüglich der Einteilung in Winter- (Oktober bis April) und Sommermonate (April bis Oktober). Die geschätzten Werte für a_e der Wintermonate weisen größere Standardabweichungen auf, da hier weniger Datenpunkte verfügbar sind. Zudem liegt ein größeres Rauschen bezüglich der Korrelation von reduzierter Bojengeschwindigkeit und dem Windschubspannung vor. Werden die Abweichungen des imaginären Anteils für die Bojen mit Treibanker bei 7° und 8°N vernachlässigt, so lassen sich in der Tat keine signifikanten Unterschiede zwischen den Parametrisierungen feststellen, sodass eine Unterteilung in Winter- und Sommermonate als überflüssig unterstellt wird.

Werden die aus diesem Modell gewonnenen Parameter auf reale Windschubspannungsvektoren angewendet, so ergibt sich der Anteil des Ekman-Transports an der Bojengeschwindigkeit beispielsweise für das Jahr 2007 zu dem in Abbildung 6.6 dargestellten Strömungsbild. Der Ekman-Transport lenkt die Boje ungefähr 90° zur Rechten des Windes ab, sodass durch den subtropischen Wirbel im nordatlantischen Ozean ein charakteristisches Muster einstellt. Der subtropische Wirbel dreht im Uhrzeigersinn in einer elliptischen Bahn über den Nordatlantik. Der Ekman-Transport wirkt dementsprechend so, dass die Geschwindigkeitsvektoren auf das Zentrum dieser Ellipse weisen. Eine vollständige Übereinstimmung mit der Abbildung 2.5 ist nicht zu erreichen, da zum Einen ein viel kürzerer Zeitraum in die hier dargestellten Daten und damit auch deutlich weniger Windschubspannungsvektoren zur Berechnung einfließen. Zum Anderen wurde für die Abbildung 2.5 ein anderes Modell angesetzt (RALPH und NIILER 1999, S. 2123 ff.), welches globale Parameter enthält, sodass dafür keine Unterteilung in Breitengrade durchgeführt wurde.

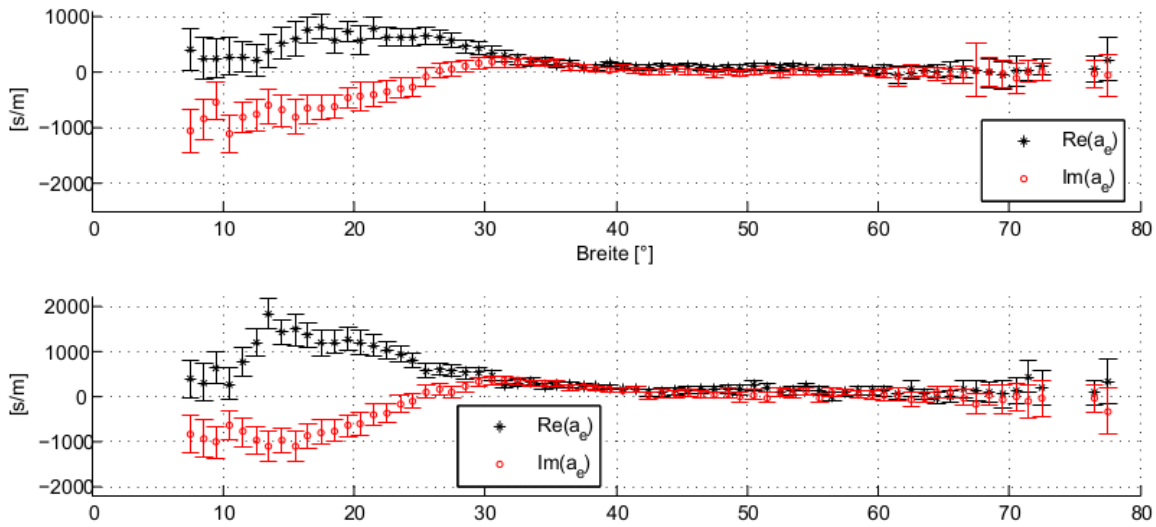


(a) Geschätzter realer und imaginärer Anteil des Parameters a_e für den tropischen Atlantik mit Bojendaten zwischen dem Januar 2007 und dem Dezember 2010, oben: Bojen mit Treibanker, unten: Bojen ohne Treibanker

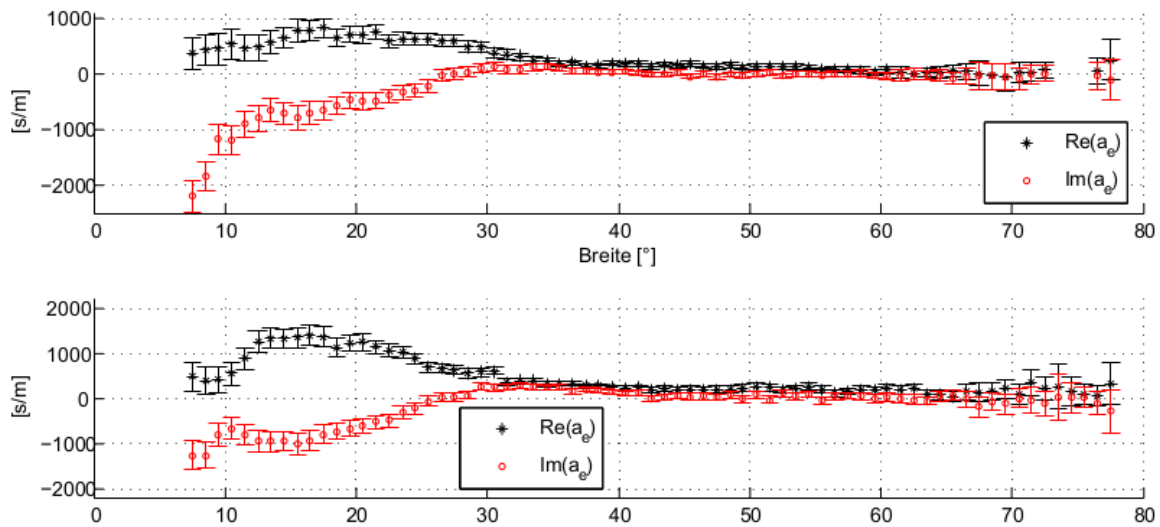


(b) Geschätzter realer und imaginärer Anteil des Parameters a_e für den tropischen Pazifik mit Bojendaten zwischen dem Januar 2007 und dem Dezember 2010, Linie: geschätzte Kurve nach dem Modell von LAGERLOEF et al. (1999), gestrichelte Linie bei einer Skalierung des konstanten Anteils um 1.25 zur Verdeutlichung der Sensitivität diesbezüglich (Modellparameter: $h = 32.5m$ und $r = 2.15 \times 10^{-4} m s^{-1}$)

Abbildung 6.4: Ergebnisse der ganzjährigen Schätzung des Parameters a_e , vertikale Striche stehen für die einfache Standardabweichung



(a) Schätzergebnisse durch die Auswertung mit Bojenpositionen in den Wintermonaten Oktober bis April, oben: Bojen mit Treibanker, unten: Bojen ohne Treibanker



(b) Schätzergebnisse durch die Auswertung mit Bojenpositionen in den Sommermonaten April bis Oktober, oben: Bojen mit Treibanker, unten: Bojen ohne Treibanker

Abbildung 6.5: Geschätzter realer und imaginärer Anteil des Parameters a_e für den tropischen Atlantik mit Bojendaten zwischen dem Januar 2007 und dem Dezember 2010

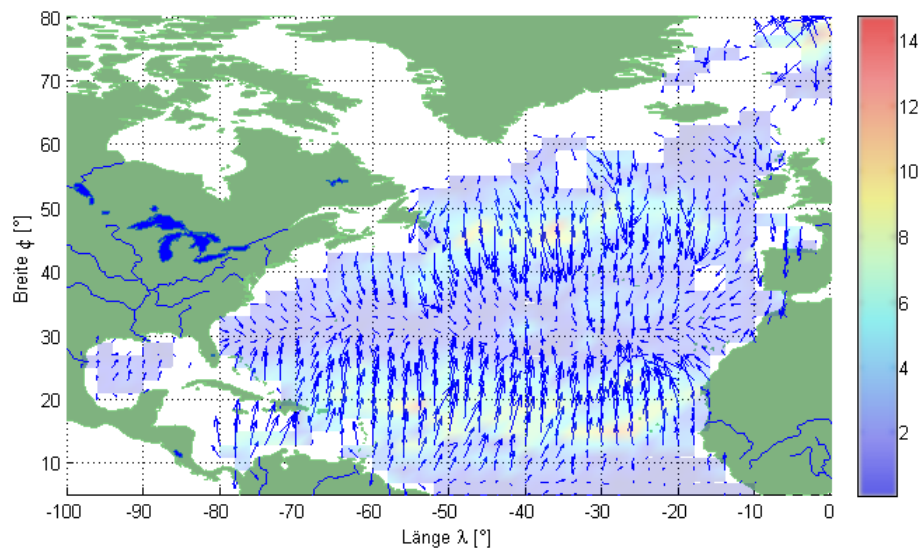


Abbildung 6.6: Strömungsbild des Ekman-Transports für das Jahr 2007 für Bojen mit Treibanker

7. Strömungsmodellierung über stückweise bikubische Polynome

Die Strömungsmodellierung zur Abbildung der ADT durch Bojen nutzt hier die um den windgetriebenen Schlupf und die Ekman-Drift reduzierten Bojengeschwindigkeiten nach Kapitel 6. In den Bojengeschwindigkeiten sind so allerdings noch die hochfrequenten Anteile der ageostrophischen Strömung mit einer Periode von etwa einem Tag oder weniger enthalten. Die Gezeiten mit halbmonatlicher und monatlicher, sowie halbjährlicher Periode sind aufgrund ihrer geringen Amplitude zu vernachlässigen. Die ageostrophischen Anteile werden durch die Darstellung der Strömung in Monatsscheiben heraus gemittelt, sodass die Anwendung eines Tiefpassfilters überflüssig wird. Eine zeitlich höhere Auflösung bietet sich aufgrund der räumlich geringen Abdeckung durch die Bojen nicht an (s. Abbildung 3.2). Als Beobachtungen werden nur die Geschwindigkeiten der Bojen eingeführt, sodass bezüglich der Höhe der ADT ein Randdefekt entsteht.

Die Strömungsmodellierung soll auf Grundlage bikubischer Polynome in einem quadratischen Gitter stattfinden. Durch die angesetzte Parametrisierung sind nach der Schätzung direkt die Geschwindigkeit und Höhe definiert. Zudem wird die volle Kovarianzinformation erhalten, mit der Einschränkung, dass die Ekman-Drift und der windgetriebene Schlupf von den Geschwindigkeiten subtrahiert wurden ohne eine Genauigkeitsaussage zuzulassen.

Die mathematischen Zusammenhänge der Modellierung sollen in den nachfolgenden Abschnitten dieses Kapitels erläutert werden.

7.1 Stückweise bikubische Polynome

Die hier angewendeten stückweise bikubischen Polynome mit quadratischer oder wahlweise rechteckiger Basis entsprechen in ihrer Parametrisierung dem sogenannten Bogner-Fox-Schmidt-Element (BOGNER et al. 1965). Die Parametrisierung ergibt sich wie folgt

$$\begin{aligned} \zeta(\lambda, \phi) = & \sum_{i=0}^1 \sum_{j=0}^1 \chi(\xi_i) \chi(\eta_j) \zeta^{m,n} + (-1)^i \psi(\xi_i) \chi(\eta_j) l \zeta_{\xi}^{m,n} \\ & + (-1)^j \chi(\xi_i) \psi(\eta_j) w \zeta_{\eta}^{m,n} + (-1)^{i+j} \psi(\xi_i) \psi(\eta_j) l w \zeta_{\xi\eta}^{m,n}, \end{aligned} \quad (7.1)$$

mit $\xi_i = \left| \frac{\lambda - \lambda_m}{l} \right|$, $m = e + i$

$\eta_j = \left| \frac{\phi - \phi_n}{w} \right|$, $n = f + j$

$\chi(t) = 1 - 3t^2 + 2t^3$ und $\psi(t) = t - 2t^2 + t^3$.

Die Variablen ξ_i und η_j stehen für die skalierten lokalen Koordinaten in einem Rechtekelement begrenzt durch die Knoten $(0, 0)$, $(0, 1)$, $(1, 0)$ und $(1, 1)$, welche hier durch die Summen über (i, j) durchlaufen werden. Die Variablen e und f bezeichnen die globale Knotenbezeichnung in λ - und

ϕ -Richtung beginnend links unten im Gitter. $\zeta(\lambda, \phi)$ steht für die Beobachtung der Höhe mit den Koordinaten λ und ϕ , welche der Lage in Länge und Breite entsprechen. Die Parameter des Modells werden durch $\zeta^{m,n}$, $\zeta_\xi^{m,n}$, $\zeta_\eta^{m,n}$ und $\zeta_{\xi\eta}^{m,n}$ in dem Knoten (m, n) festgelegt, also dem Höhenwert, der Ableitung der Fläche in östlicher und der in nördlicher Richtung, sowie der gemischten Ableitung. In der Summe sind somit für ein Rechteck 16 Parameter zu bestimmen. Die durch die Bojen gemessenen Geschwindigkeiten sind nach Gleichung (2.6) und (2.7) in die Ableitungen der Höhe ζ in die entsprechenden Richtung umzuformen. In globalen Koordinaten gesehen entspricht somit ζ_ξ der Ableitung der Höhe der Wassersäule über der Äquipotentialfläche in östlicher Richtung $\frac{\partial \zeta}{\partial x}$ und ζ_η der in nördlicher Richtung $\frac{\partial \zeta}{\partial y}$. Da auch so nicht die Höhenwerte sondern die Ableitungen in östlicher und nördlicher Richtung als Beobachtungswerte vorliegen, muss der funktionale Zusammenhang in die entsprechende Richtungen abgeleitet werden. Es ergibt sich

$$\begin{aligned} \zeta_\xi(\lambda, \phi) = \frac{1}{l_M} \sum_{i=0}^1 \sum_{j=0}^1 & (-1)^i \chi'(\xi_i) \chi(\eta_j) \zeta(i, j) + \psi'(\xi_i) \chi(\eta_j) l_M \zeta_\xi(i, j) \\ & + (-1)^{i+j} \chi'(\xi_i) \psi(\eta_j) w_M \zeta_\eta(i, j) + (-1)^j \psi'(\xi_i) \psi(\eta_j) l_M w_M \zeta_{\xi\eta}(i, j) \end{aligned} \quad (7.2)$$

$$\begin{aligned} \zeta_\eta(\lambda, \phi) = \frac{1}{w_M} \sum_{i=0}^1 \sum_{j=0}^1 & (-1)^j \chi(\xi_i) \chi'(\eta_j) \zeta(i, j) + (-1)^{i+j} \psi(\xi_i) \chi'(\eta_j) l_M \zeta_\xi(i, j) \\ & + \chi(\xi_i) \psi'(\eta_j) w_M \zeta_\eta(i, j) + (-1)^i \psi(\xi_i) \psi'(\eta_j) l_M w_M \zeta_{\xi\eta}(i, j). \end{aligned} \quad (7.3)$$

Um die Höhenwerte in Metern zu erhalten sind die Breite und Länge der Gitterquadrate in ihre Länge in Metern bezüglich des WGS84-Ellipsoids umzurechnen, in denen die Koordinatenwerte der Bojen aufgezeichnet sind, dies gilt allerdings nicht für die Normierung auf das Einheitsquadrat zur Bestimmung der lokalen Koordinaten. So ergeben sich die Größen l_M für die Länge der Ausdehnung eines Gitterquadrats in östliche Richtung und w_M für diejenige in nördliche Richtung in der Umrechnung auf dem Breitengrad des betrachteten Knotens bezogen.

Unter Einbringung von Restriktionen nach FRANKEN (2012, S. 18), sodass sich eine in der zweiten Ableitung stetige Fläche ergibt, wird so in einem linearen Ausgleich eine die Strömungsmodellierung ganzheitlich beschreibende Parametrisierung abgeleitet. Die Restriktionen werden zu je zwei für die inneren Knoten und zu je einer an den äußeren Knoten in einem Gesamtausgleich eingebracht. Die an den Knoten in den Ecken des Gitters eingebrachten Restriktionen, welche die zweite Ableitung in östliche und nördliche Richtung zu Null setzen, führen zu einem natürlichen Spline.

Übrig bleiben zwei Problematiken. Da ein regelmäßiges rechteckiges Gitter über die Bojenpositionen gespannt wird, um global den Rahmen für die lokalen Intervalle festzulegen, gibt es Gitterelemente beziehungsweise Rechtecke, in denen keine Beobachtungen vorliegen. Außerdem entsteht durch das Fehlen von Beobachtungen bezüglich der Höhe ein Datumsdefekt.

Der ersten Problematik wird durch das Einbringen von Pseudobeobachtungen abgeholfen. Diese Form der Beobachtungen werden in den Ausgleich mit etwas niedrigerem Gewichtung relativ zu den wahren Beobachtungen eingebracht, sodass sie den zu approximierenden Verlauf, dort wo Beobachtungen vorliegen, nicht stören. Zudem werden sie nur an den Knoten angebracht, wo weniger als 16 Beobachtungen vorliegen. An den übrigen Knoten bestimmen die Beobachtungen den Ver-

lauf in ausreichender Form, sodass das Anbringen von Pseudobeobachtungen überflüssig wird. Die hier eingebrachten Pseudobeobachtungen nach FRANKEN (2012, S. 7) werden bezüglich der vier Parameter je Knoten eingeführt. So ergeben sich je Knoten mit weniger als 16 Beobachtungen vier Beobachtungsgleichungen, deren Beobachtungswert mit Null belegt wird. Den funktionale Zusammenhang bildet das Achtfache des betrachtete Parameterwertes, welcher um die entsprechenden Parameterwerte der acht umliegenden Knoten vermindert wird. Diese Achter-Nachbarschaft setzt sich aus den die angrenzenden Rechtecke beschreibenden Knoten zusammen.

Die Datumsproblematik wird über das Einführen einer zusätzlichen Restriktion gelöst. Diese setzt das Mittel aller Höhenpunkte von Knoten mit mehr als 16 Beobachtungen zu Null, sodass eine auf Null zentrierte Fläche geschätzt wird.

Der Gesamtausgleich gemäß dem Gaußschen Minimierungsprinzip ergibt sich somit wie folgt

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{A}^T \boldsymbol{\Sigma}\{\mathcal{L}\} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{l}_v \\ \mathbf{0} \end{bmatrix} \quad \text{mit } \mathbf{B} = \begin{bmatrix} \mathbf{B}_r \\ \mathbf{B}_h \end{bmatrix} \quad \text{und } \mathbf{A} = \begin{bmatrix} \mathbf{A}_l \\ \mathbf{A}_{ps} \end{bmatrix}. \quad (7.4)$$

Wobei sich die Restriktionsmatrix \mathbf{B} aus derjenigen für die Restriktionen der Stetigkeit und des natürlichen Splines \mathbf{B}_r und derjenigen für die mittlere Höhe \mathbf{B}_h zusammensetzt. Die Designmatrix \mathbf{A} entspricht der Ergänzung der Designmatrix zu den beobachteten Geschwindigkeiten \mathbf{A}_l mit der Designmatrix der Pseudobeobachtungen \mathbf{A}_{ps} . Der Beobachtungsvektor entspricht den Vektor den Geschwindigkeiten und einem Nullvektor, dessen Länge sich über die Anzahl der Pseudobeobachtungen und der eingebrachten Restriktionen definiert, da sich diese alle zu Null ergeben sollen. Die Kovarianzmatrix der Beobachtungen $\boldsymbol{\Sigma}\{\mathcal{L}\}$, kann aufgrund der Tatsache, dass diese keinen vollen Rang besitzt, nur über die Hauptdiagonaleinträge eingeführt werden. Innerhalb des alternativen Verfahrens zur Bestimmung der Positionen und Geschwindigkeiten aus der geglätteten Spur werden sehr viel mehr Werte abgegriffen als Parameter geschätzt werden, sodass auch wenn nur die Kovarianzmatrix der Geschwindigkeiten einfließt, ein Rangdefekt innerhalb dieser gekürzten Kovarianzmatrix entsteht. Der ausgeglichene Parametervektor $\tilde{\mathbf{x}}$ enthält alle für die Modellierung der ADT und der dazugehörigen Gechwindigkeiten relevanten Informationen, dem Höhenwert, der Ableitung in nördlicher und östlicher, sowie in der gemischter Richtung und das für jeden Knoten. Für die entgültige Ausgabe werden solche Knoten auf *NaN* gesetzt, dessen Anzahl einfließender Beobachtungen kleiner als 16 ist, sodass sie das Bild der zeitlichen Interpretation nicht verfälschen. Knoten zu deren Bestimmung keine oder nur sehr wenige Geschwindigkeitsbeobachtungen einfließen, werden in ihrem Hochwert hauptsächlich durch die Pseudobeobachtungen, deren Anwendung auch als Krümmungsminimierung bekannt ist und der die Fläche starrer machenden Restriktionen definiert, sodass der Eintrag der tatsächlichen dynamischen Topographie sehr gering ist. Die entstehenden Lücken sollen durch die im Nachfolgenden eingeführte zeitliche Modellierung dieser räumlichen Modelle aufgefüllt werden.

7.2 Ergebnisse der Strömungsmodellierung über Oberflächenbojengeschwindigkeiten

Für die Modellierung einer Monatsdarstellung der absoluten dynamischen Topographie des Nordatlantiks wurde jeweils der volle Datensatz der durch das alternative Verfahren bestimmten Positionen und Geschwindigkeiten genutzt. Bojengeschwindigkeiten von Bojen ohne Treibanker sind um den windgetriebenen Schlupf korrigiert. Zudem sind, durch Anwendung des in Abschnitt 6.2 beschriebenen Modells die Geschwindigkeiten aller Bojen bezüglich der Ekman-Drift korrigiert. Die Ergebnisse bezüglich des Jahres 2009 sind nachfolgend abgebildet. Die Abbildungen 7.1 weisen eine über das Jahr sich stark verändernde Topographie, was in der Form nicht in der tatsächlichen absoluten dynamischen Topographie festzustellen ist. Versuche zeigten, dass gegitterte Geschwindigkeiten zu der gewünschten Topographie führen, hingegen solche nur teilweise aneinandergrenzende Areale, welche durch die Geschwindigkeiten der Bojen abgedeckt werden, in den nicht abgedeckten Gebieten eine variable Lagerung besitzen. Diese verändert die Topographie aufgrund der sich über das Jahr verändernden Abdeckung und der Bedingung, dass die mit Beobachtungen belegten Knoten bezüglich der Höhe in Summe zu Null gesetzt werden, führt hauptsächlich zu den in der Abbildung sich über das Jahr stark variierenden Topographie.

Da die Meeresforschung aber an den Zirkulationsmechanismen und damit den geostrophischen Strömungsgeschwindigkeiten interessiert ist, gilt das Interesse eher der Validierung der aus der Fläche abgeleiteten Geschwindigkeiten. Die Höhen dienen höchstens dem direkten Vergleich mit den aus der Altimetermessung gewonnenen Beobachtungsgrößen. Bezüglich der aus der Fläche abgeleiteten Geschwindigkeitsfelder, welche genauso in dem geschätzten Parametervektor enthalten sind wie die Höhe, wurde die Auswertung einmal für die Bojen mit Treibanker und einmal für die ohne durchgeführt. Da so weniger Beobachtungen pro Monat vorliegen, wurde der Beobachtungszeitraum auf drei Monate ausgeweitet. Die Geschwindigkeitsfelder der Monate Juli bis September des Jahres 2009 (s. Abbildung 7.2) zeigen für die Bojen ohne Treibanker eine erhöhte Geschwindigkeit in Windrichtung (s. Abbildung 4.3). Dies kann ein Indiz dafür sein, dass der windgetriebene Schlupf nicht ausreichend korrigiert wurde, denn im direkten Vergleich mit den Geschwindigkeiten, welche aus der ADT abgeleitet sind, tauchen diese nicht auf. Die Bojen ohne Treibanker weisen das gleiche Muster auf, allerdings in abgemilderter Amplitude. Auch hier sollte demnach über die Korrektur des windgetriebenen Schlupfes nachgedacht werden. Die Darstellung der Genauigkeiten der bestimmten Geschwindigkeitsfelder macht nur wenig Sinn, da sie keine Genauigkeiten in Hinblick auf die eingebrachten Windschubspannungsvektoren und Windfelder aufzeigen würde. Zudem würden die Genauigkeiten vor allem den Verlust der räumlichen Auflösung abbilden, da durch die Schätzung auf $2^\circ \times 2^\circ$ -Gitter eine großräumige Glättung der Bojengeschwindigkeiten durchgeführt wird. Diese Arbeit sollte aufzeigen, dass aus den Bojengeschwindigkeiten geostrophische Strömungen abgeleitet werden können, sodass in Kombination mit anderen Beobachtungsgrößen eine besseren räumlichen Auflösung der Geschwindigkeitsfelder erhalten werden kann.

Dass die aus der ADT abgeleiteten Geschwindigkeiten einen solch diffusen Richtungswechsel aufweisen ist durch deren erhöhte räumliche Auflösung zu erklären, welche hier in diesem Größenformat nicht in geeigneter Art und Weise dargestellt werden kann.

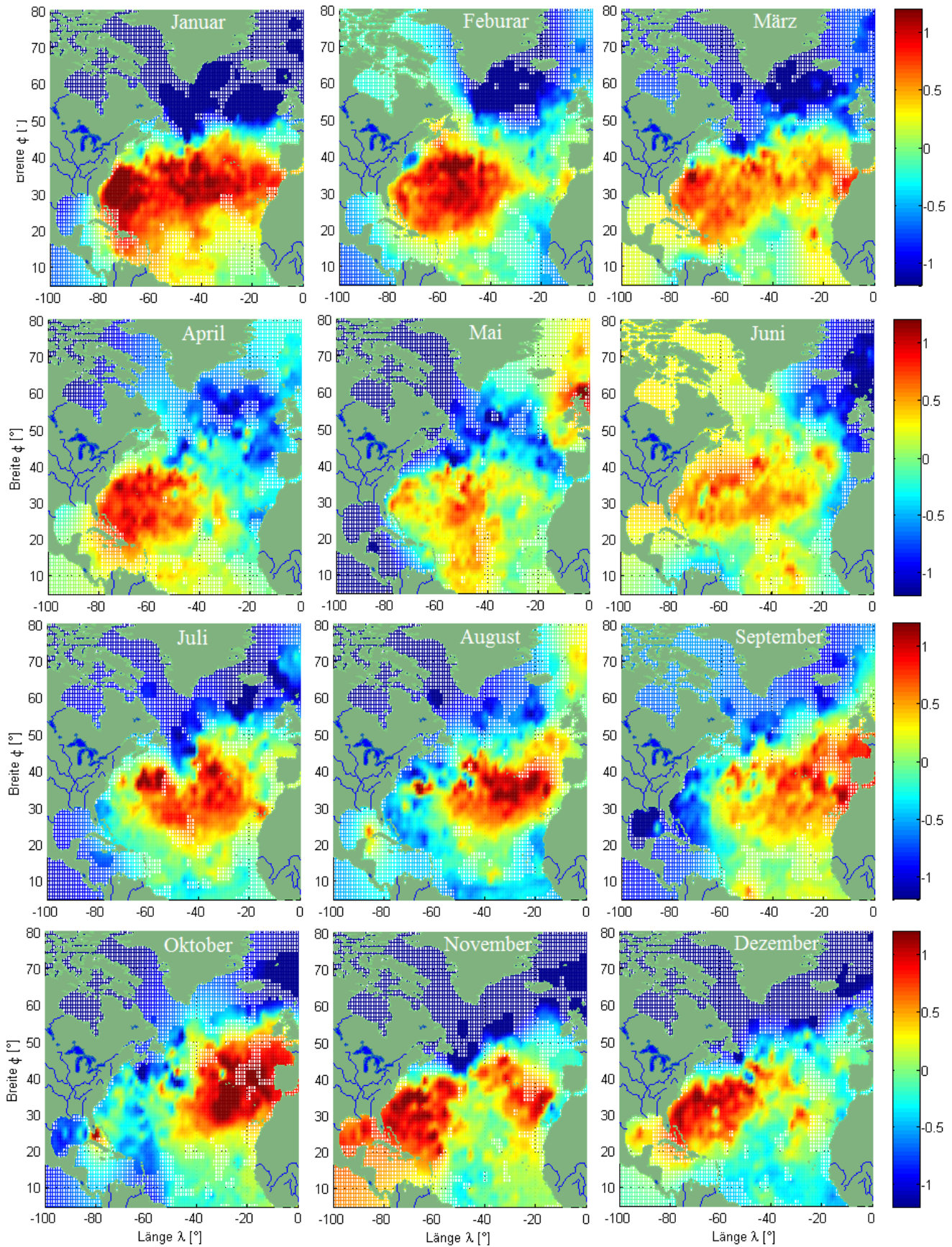


Abbildung 7.1: Räumliche Modellierung der ADT in einem 2°-Gitter für das Jahr 2008 aus Oberflächendrifterdaten im Nordatlantik, Bereiche, in denen mehr als 16 Beobachtungen vorliegen, sind flächenmäßig ausgefüllt, sonst gittert.

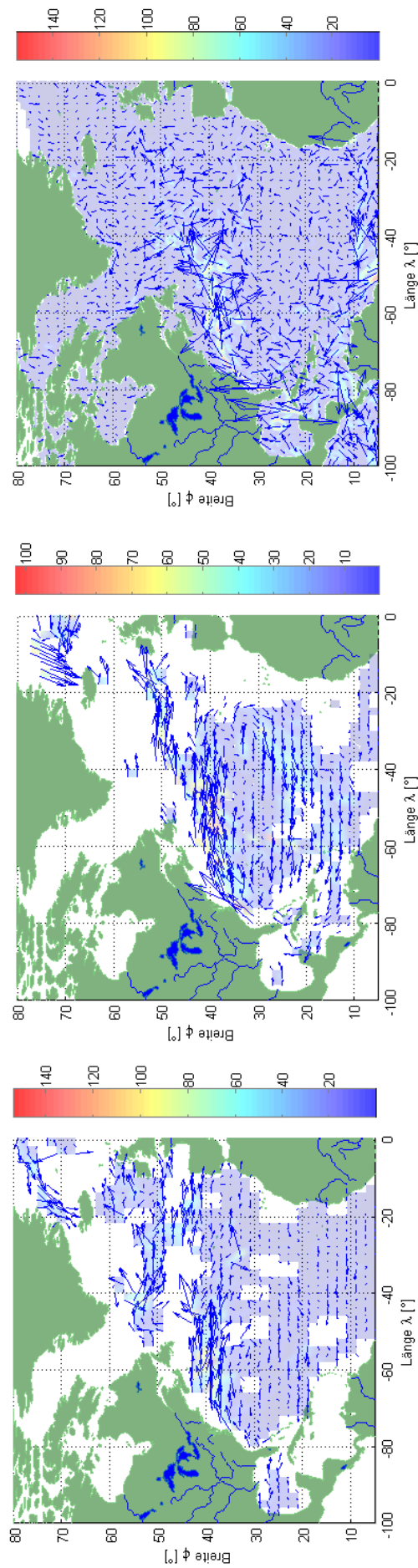


Abbildung 7.2: Gemittelte Geschwindigkeitsfelder der Monate Juli bis September 2009, für die Bojengeschwindigkeiten der Bojen mit Treibanker (links), für die Bojengeschwindigkeiten ohne Treibanker (Mitte) und die aus der ADT abgeleiteten Geschwindigkeiten nach Rio et al. (2011) (rechts)

8. Schlussbemerkungen

8.1 Fazit

Diese Arbeit zeigt ein alternatives Verfahren zu der durch das Surface Velocity Programm durchgeführten Spurglättung und der daraus abgeleiteten Geschwindigkeitsbestimmung, auf. Dieses benötigt keine a priori Informationen bezüglich der Genauigkeit der Daten und setzt zweifach stetige stückweise kubische Polynome zur Glättung der Spur an. Durch die datenadaptive Knotenwahl über den MARS-Algorithmus wird eine statistische Anpassung der Spur an die Daten geschaffen und die Anwendung empirisch bestimmter Variogramme umgangen. Die Knoten legen die zeitliche Ausdehnung der Träger der Splines fest. Durch das Ansetzen von kubischen Polynomen und verschiedenen Stellgrößen ist es nun dem Anwender möglich bereits in der Vorverarbeitung eine zeitliche Filterung der Rohdaten durchzuführen. Auch die Zeitpunkte, in denen die Koordinatenwerte und Geschwindigkeiten abgegriffen werden sind von dem Anwender frei zu wählen. Durch das direkte Ableiten der Geschwindigkeit aus dem funktionalen Zusammenhang wird die vereinfachte Geschwindigkeitsableitungen über den zentralen Differenzenquotienten vermieden. Die Geschwindigkeiten erhalten für spätere Anwendungen ein kompaktes Genauigkeitsspektrum und weisen keine den Amplitudenwert deutlich übersteigende Genauigkeiten wie bei dem SVP-Verfahren auf. Da mehr Positionen und Geschwindigkeiten aus der geglätteten Spur abgegriffen als Parameter geschätzt werden, entsteht ein Rangdefekt bezüglich der Kovarianzmatrix, sodass die volle Kovarianzinformation nicht in die Auswertung der Strömungsmodellierung aufgenommen werden kann.

Das Verfahren zeigt sich unabhängig gegenüber Ausreißern und ist so auch auf Roh-Daten ohne vorherige Editierung anwendbar. Da aber deutlich mehr editierte Positionsdaten über die Datenbanken abzurufen sind, werden diese bereits editierten Daten genutzt.

Die Untersuchung der Unterschiede zwischen den Geschwindigkeiten der absoluten dynamischen Topography aus dem CNES-CLS09 Modell (Rio et al. 2011) und den Bojengeschwindigkeiten ließen die Interpretation zu, dass zusätzliche, auf die Bojengeschwindigkeiten wirkende Effekte wie die der ageostrophischen Strömungen und damit vor allem windgetriebenen Effekte für die geostrophische Strömungsmodellierung zunächst entfernt werden müssen.

Der windgetriebene Effekt wurde als der Literatur entnommener prozentualer Anteil an der Windgeschwindigkeit von der Bojengeschwindigkeit subtrahiert. Ausgewertete Geschwindigkeitsfelder zeigen jedoch, dass noch ein Anteil dieser in den reduzierten Bojengeschwindigkeiten enthalten sein muss. Zudem wurde eine Modellierung der Ekman-Drift erfolgreich implementiert und auf die Bojen und ihren Eigenschaften zugeschnitten.

Durch diese Vorbereitung können die Bojengeschwindigkeiten zur Strömungsmodellierung in die Normalgleichungen kombinierter Modelle aufgenommen werden. Die Einzelauswertung, in der nur die Bojengeschwindigkeiten als Beobachtungsgrößen vorliegen, gestaltet sich jedoch als schwierig, da die Bojen zwar im Vergleich zu Altimetermessungen eine zeitlich und räumlich hohe Abtastung versprechen, allerdings räumlich eine inhomogene Verteilung aufweisen. So kommt es zu Lagerungsproblemen was die Bestimmung der Höhe angeht, wodurch wiederum kein Vergleich mit der ADT

die aus Satellitendaten berechnet wurde, möglich ist. Realisierbar ist aber die Bestimmung von Geschwindigkeitsfeldern mit den hier angesetzten stückweise bikubischen Polynomen der Bogner-Fox-Schmit-Elemente. Die Geschwindigkeitsfelder erlauben die direkte Analyse der Strömung und bilden damit bereits die Zielgröße zur Analyse von Zirkulationsmechanismen in der Meeresforschung.

8.2 Ausblick

Die Ergebnisse in den Geschwindigkeitsfeldern zeigten, dass der windgetriebene Schlupf vor allem in den Bojengeschwindigkeiten von Bojen ohne Treibanker nicht erfolgreich reduziert wurde. Dieser Aspekt erfordert eine weitere Betrachtung.

Bezüglich des Ekman-Transports gibt SUDRE et al. 2013 an, dass die Einteilung in Breitengrade zu vereinfacht ist. Es ist demnach weiter zu prüfen, dass trotz der schlechten Genauigkeiten der bestimmten Parameter in der äquatorialen Region bei einer Unterteilung in Längengrade, dies zu gewünschten Ergebnissen führt. Der Ekman-Transport wird über die Differenz der Bojengeschwindigkeit mit den Geschwindigkeiten der ADT bestimmt, sodass die reduzierten Bojengeschwindigkeiten kein vollständig unabhängiges Modell in einer Kombinationsauswertung darstellen. Eine Überprüfung in welchem Ausmaß die um den Ekman-Transport reduzierten Geschwindigkeiten von der ADT abhängig sind, scheint notwendig um die Bereitstellung unabhängiger Beobachtungswerte durch die Bojengeschwindigkeiten in kombinierten Modellen gewährleisten zu können.

Bezüglich der um die windgetriebenen Effekte reduzierten Bojengeschwindigkeit ist eine kombinierte Auswertung mit Satellitenverfahren und anderen in situ Beobachtungsgrößen möglich. Dazu sollte die zu ignorierende Frequenz innerhalb des MARS-Algorithmus so gewählt werden, dass keine hochfrequenten ageostrophischen Strömungen in der Bojengeschwindigkeiten enthalten sind. Die Frequenz liegt nach RIO und HERNANDEZ (2004) bei dem Maximum der inertialen Periode oder einem Tag. Natürlich sollten nicht die räumlich modellierten Geschwindigkeiten in die Kombinationsauswertung aufgenommen werden, sondern direkt die geostrophischen Geschwindigkeiten abbildenden Geschwindigkeiten. Die räumliche Modellierung sollte aufgrund ihrer stark geglätteten Struktur begründet durch die inhomogene Verteilung der Bojenspuren, rein zur Validierung der Daten genutzt werden.

Während aus der Spurglättung realistische Genauigkeiten bezüglich der Geschwindigkeiten abgeleitet werden können ist dies im weiteren aufgrund fehlender Standardabweichungen beispielsweise bezüglich der Winddaten nicht möglich. Die Einbringung von Informationen zu diesen Punkten ist wünschenswert.

A. Anhang

A.1 Redundanz in einer Spline-Approximation

Dieser Abschnitt soll die Verteilung von Redundanzanteilen in einer B-Spline Approximation analysieren und anhand dessen Aussagen bezüglich der Gewichtung einzelner Beobachtungen in der Schätzung formulieren. Zu diesem Zweck werden äquidistante Stützstellen auf eine drei Intervalle breite Zeitreihe gesetzt. Die Designmatrix kann diesbezüglich aufgestellt werden und die Redundanzanteile über

$$r_i = \left(\mathbf{I} - \mathbf{A}^B \left(\mathbf{A}^{BT} \mathbf{A}^B \right)^{-1} \mathbf{A}^B \right)_{ii} \quad \text{mit} \quad 0 \leq r_i \leq 1 \quad (\text{A.1})$$

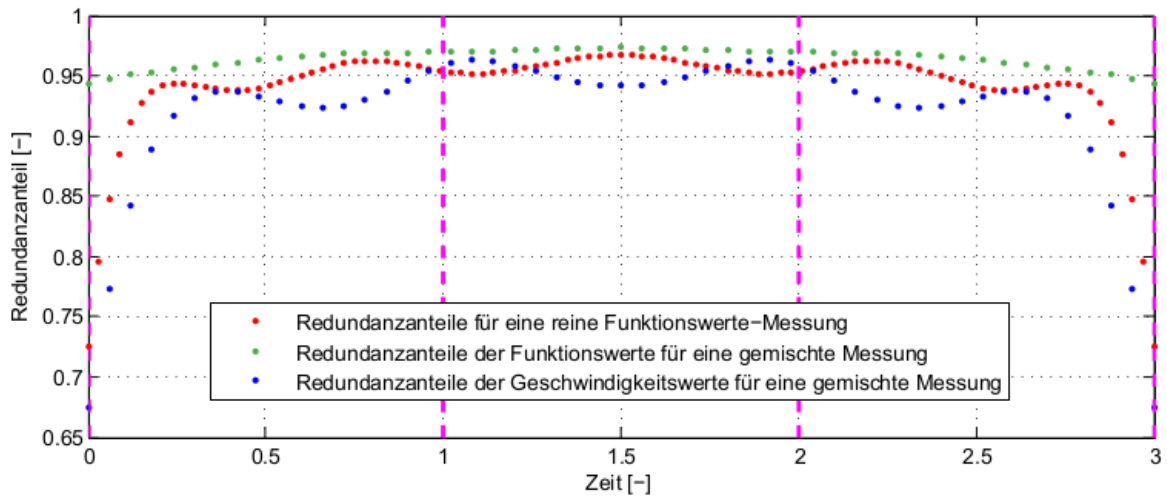
berechnet werden. Die nachfolgende Abbildung A.1 zeigt Redundanzanteile von verschiedenen Beobachtungstypen. Zunächst wurde eine Designmatrix erzeugt, in der nur Funktionswerte als Messwerte einfließen. Die Redundanzanteile dieser Messung werden in der Abbildung als rote Punkte abgebildet und zeigen einen wellenartigen Verlauf. Besonders hohe Redundanzanteile zeugen von geringem Einfluss der einzelnen Beobachtung auf den Verlauf der geschätzten Funktion. Dementsprechend sind die Beobachtungen in der Mitte der Intervalle deutlich besser kontrolliert als an den Intervallgrenzen, an denen der Einfluss einzelner Beobachtungen besonders groß ist. Im Besonderen gilt dies an den nicht an weitere Intervalle angeschlossenen Grenzen.

Ein Polynom dritten Grades ist zweifach stetig differenzierbar, was auch für die hier genutzten B-Splines gilt. Es kann somit über den Funktionswert, die Geschwindigkeit, die Beschleunigung und die sogenannte Ruckfunktion definiert werden. Bis zur Ruckfunktion zeigen sich der Funktionsverlauf und die daraus abgeleiteten Funktionsverläufe des Polynoms als stetig.

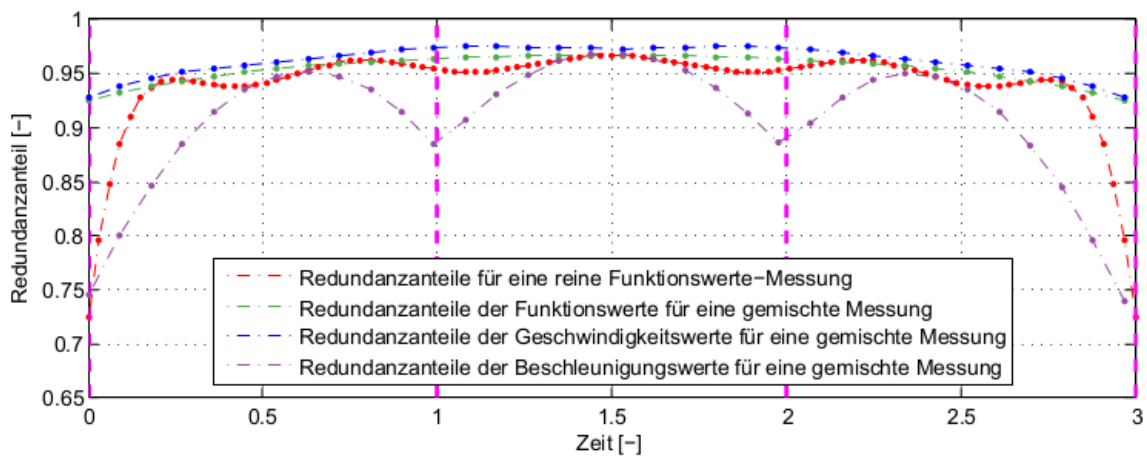
Nun werden zwei weitere Situationen geschaffen. Zum Einen werden Redundanzanteile für eine gemischte Messung bestimmt, die auch Geschwindigkeitswerte als Beobachtungen annimmt. In einer dritten Situation werden zu den Funktions- und Geschwindigkeitswerten auch noch Beschleunigungswerte angenommen. Für alle Situationen ist die Größenordnung der Redundanzanteile ähnlich, weil die Anzahl der Freiheitsgrade gleich gehalten wurde.

Aus den Redundanzanteilen durch Einbringung der Geschwindigkeit als Beobachtungswert ist zu erkennen, dass der Geschwindigkeit im Inneren der Intervalle mehr Bedeutung bezüglich der Funktionsdefinition zuzurechnen ist. Ihr Bedeutung nahe der Grenzen ist geringer als die der Funktionswerte. Für die Beschleunigung gilt der umgekehrte Fall. Ihre Bedeutung ist deutlich größer auf den Intervallgrenzen und ähnlich gering wie die der Funktionswerte für das Intervallinnere.

Abschließend lässt sich bezüglich einer praxisnahen Funktionswertemessung sagen, dass Funktionswerte auf den Intervallgrenzen vor allem den Funktionswert und die Beschleunigung definieren. Die Geschwindigkeit wird maßgeblich durch Funktionswertemessungen im inneren eines Intervalls bestimmt. Für Intervalle in den Randbereichen der Zeitreihe gilt, dass die Bedeutung des Funktionswertes für den Funktionsverlauf nahezu exponentiell ansteigt, desto näher er der äußeren Grenze kommt. Dementsprechend wächst auch die Varianz in den an den äußeren Grenzen bestimmten Parametern.



(a) Gemischte Messung entspricht der Messung von Funktionswert und Geschwindigkeit



(b) Gemischte Messung entspricht der Messung von Funktionswert, Geschwindigkeit und Beschleunigung

Abbildung A.1: Redundanzanteile bei einer Einteilung einer Zeitreihe in drei Intervalle und dem Ausgleich über stückweise kubische Polynome, einmal für eine reine Funktionswerte-Messung und einmal für eine gemischte Messung, beide verglichenen Modelle enthalten nahezu die gleiche Anzahl an Beobachtungen (in b) eine Beobachtung in der gemischten Messung weniger).

B. Bezeichnungen

Physikalische Größen und Konstanten

ϕ	geozentrische Breite
λ	geozentrische Länge
u	Geschwindigkeit in östlicher Richtung
v	Geschwindigkeit in nördlicher Richtung
ϖ	vertikale Geschwindigkeit
w	horizontale Gesamt-Geschwindigkeit
ρ	Dichte
p	Druck
g	Schwerebeschleunigung der Erde
f	Coriolisparameter
Ω	Erdrotationsrate
A_v	Wirbelviskosität
\mathbf{u}	Geschwindigkeitsvektor
\mathbf{u}_g	Geschwindigkeitsvektor der geostrophischen Geschwindigkeit
\mathbf{u}_{ek}	Geschwindigkeitsvektor der Ekman-Drift
\mathbf{u}_a	Geschwindigkeitsvektor der hochfrequenten ageostrophischen Strömung
$\boldsymbol{\tau}$	Vektor der Windschubspannung
\mathbf{u}_{10}	Vektor der Windgeschwindigkeit in 10m über der Meeresoberfläche

Statistische Größen und Variablen der Schätzverfahren

\mathbf{l}	Beobachtungsvektor
\hat{s}	geschätzter Interpolationspunkt
s	wahrer Interpolationspunkt
$\mathcal{L}, \mathcal{S}, \mathcal{N}$	Zufallsvektor (fett) bzw. Zufallsvariable (dünn) der entsprechenden Größe
\hat{s}_ϕ	geschätzter Interpolationspunkt der Breite
\hat{s}_λ	geschätzter Interpolationspunkt der Länge
n	Anzahl der Beobachtungen (definiert auch die Anzahl geschätzter Interpolationspunkte der Länge oder Breite)
m	Anzahl der Parameter

\mathbf{A}	Designmatrix
\mathbf{N}	Normalgleichungsmatrix
\mathbf{G}	um die Beobachtungswerte der Geschwindigkeiten erweiterte Designmatrix
\mathbf{x}	Parametervektor
\mathbf{v}	Verbesserungsvektor
$\tilde{\mathbf{l}}$	ausgeglichenen Beobachtungsvektor
r_i	Redundanzanteil der Beobachtung i
$\Sigma\{\}$	Kovarianzmatrix
$\mathbf{E}\{\}$	Erwartungswert
$\gamma()$	Variogramm-Matrix
$\tilde{\gamma}()$	Variogramm-Matrix um Nugget-Effekt bereinigt
$\mathbf{C}_{i,j}$	Wert der Kovarianzfunktion zum Lag zwischen den Datenwerten i und j
$\gamma_{i,j}$	Wert des Variogramms zum Lag zwischen den Datenwerten i und j
k	Lagrange-Parameter
\mathbf{h}	Gewichtsvektor
σ	Varianz
ϵ_s	Interpolationsfehler
$\Phi()$	Extremwertaufgabe mit Nebenbedingungen
$t^{(i)}$	Zeit zur Beobachtung i
τ	Lag (Zeitabstand zwischen zwei Beobachtungen)
$N(\tau)$	Anzahl der zur Berechnung des Variogrammwertes zum Lag τ bestimmten Differenzen
α, β	Parameter der Variogrammfunktion
d	Konstante der Kostenfunktion C
d_s	Anzahl der Durchläufe je Suchrichtung im MARS-Algorithmus
κ	Knoten, welche die Grenzen lokaler Intervalle definieren
$\Delta\kappa$	Minimalabstand zwischen den Knoten
\bar{V}^{CLS09}	Differenz des CLS09-Signals
\bar{V}^{NSD}	Differenz des NSD(Oberflächenbojen)-Signals
$\sigma_V^{CLS09^2}$	Varianz des CLS09-Signals
$\sigma_V^{NSD^2}$	Varianz des NSD(Oberflächenbojen)-Signals
σ_V^{CLS09,NSD^2}	Varianz der Signaldifferenz zwischen dem CLS09-Signal und dem NSD-Signals
\bar{V}	signalübergreifender Mittelwert zwischen CLS09- und NSD-Signal
ρ_V	Bravais-Pearson-Korrelationskoeffizient zur Angabe der standardisierten Kovarianz zwischen dem CLS09-Signal und dem NSD-Signals
t_V	Fisherverteilte Testgröße
t_{Score}	Testgröße der Score-Statistik
t_{Wald}	Testgröße der Wald-Statistik
\mathbf{s}	Score-Funktion

C. Abkürzungen

AIC Akaiikes Informationskriterium

ADT absolute dynamische Topographie (Absolut Dynamic Topography)

ANOVA Analysis of Variance

ARGOS Advanced Research and Global Observation Satellite

AVISO Archiving, Validation and Interpretation of Satellite Oceanographic data

BLUE Best Linear Unbiased Estimator

CNES Centre National d'Etudes Spatiales

CLS Collecte de Localisation Satellite

DUACS Data Unification and Altimeter Combination System

ECMWF European Centre for Medium-Range Weather Forecasts

ERA Europäische Re-Analyse

GCV generalisiertes Kreuzvalidierung-Kriterium

GDP Global Drifter Programm

GDAC Global Data Assembly Center

GRACE Gravity Recovery And Climate Experiment

netCDF Network Common Data Format

NODC National Oceanographic Data Center

NSD Near Surface Drifter

MARS multivariater adaptiver Regressions-Spline

MDT mittlere dynamische Topographie Mean Dynamic Topography

MSSH mittlere Meeresoberflächenhöhe Mean Sea Surface Height

SLA Meeresspiegelanomalien Sea Level Anomalies

SSALTO Segment Sol multimissions d'ALTimétrie, d'Orbitographie et de localisation précise

SSH Meeresoberflächenhöhe Sea Surface Height

SVP Surface Velocity Programm

TOGA Tropical Ocean Global Atmosphere Studie

WCRP World Climate Research Program

WOA World Ocean Atlas

WOCE World Ocean Circulation Experiment

WMO World Meteorological Organization

D. Verzeichnisse

Abbildungsverzeichnis

1.1	Das System der globalen Meeresströmungen nach RAHMSTORF (2002)	1
2.1	Schematische Darstellung des sich einstellenden Gleichgewichts zwischen Corioliskraft und Druckgradienten	4
2.2	Berechnungsgrößen zur Höhenkartierung der Ozeanoberfläche	5
2.3	Schematische Darstellung der windgetriebenen Ozeanströmung und der Ekman-Spirale	7
2.4	Längen- und Zeitskala verschiedener Ozeanprozesse nach CUSHMAN-ROISIN und BECKERS (2011)	10
2.5	Zeitgemittelte Bojen-Geschwindigkeiten, Ekman-Drift und Slip der Bojen ohne Treibanker	11
3.1	Aufbau einer Oberflächenboje des Surface Velocity Programms aus PACIFIC GYRE (2013)	13
3.2	Anzahl der Bojenpositionen je $1^\circ \times 1^\circ$ -Gitterbox, links: Januar 2009, mittig: Juli 2009, rechts: gesamtes Jahr 2009	14
3.3	Anzahl der Bojen pro Jahr im Nordatlantik (1979-2012)	15
3.4	Ablaufdiagramm zum Übergang von den Roh- zu den Kriging-Daten des SVP	16
3.5	Beispiele der Editierung	18
3.6	Funktionale Darstellung der regionalisierten Variogramme aus Daten des tropischen Pazifik von 1988-1993	23
3.7	Funktionale Darstellung der regionalisierten Variogramme aus Daten aus Daten des Nord-Atlantiks von 2008-2011	26
3.8	Genauigkeit der Geschwindigkeit gegen die Geschwindigkeit selbst aufgetragen, Histogramm zur Geschwindigkeitsgenauigkeit	28
4.1	Aufteilung des Nord-Atlantiks in vier Regionen, blau: Drifterpositionen des Jahres 2008	31
4.2	Zonale, meridionale und Gesamt-Geschwindigkeiten, sowie deren Differenzen (interpolierte CLS09- vs. NSD-Geschwindigkeiten) für den Juli 2009	35

4.3	Durchschnittliche Windgeschwindigkeiten für den Januar (links) und Juli (rechts) von EARSEL (2014)	40
4.4	Empirische Korrelationswerte der fünf Regionen für den Januar 2009	41
4.5	Empirische Korrelationswerte der fünf Regionen für den Juli 2009	42
5.1	Schematische Darstellung der Startsequenz an Knoten für den MARS-Algorithmus .	46
5.2	B-Splines mit unregelmäßig Knotenabständen nach FAHRMEIR et al. (2009)	54
5.3	Synthetisches Beispiel einer Bojenspur zur Approximation mittels MARS-Algorithmus	57
5.4	Bojenpositionen in der Breite aus dem Jahr 2009 und ihr approximierter Verlauf . .	58
5.5	Approximation der Bojenspur durch eine unterschiedliche Mindestanzahl an Beobachtungen pro lokalem Intervall	58
5.6	Vergleich der geglättete Trajektorie über die Schätzung von B-Splines und der interpolierten SVP-Werte	59
5.7	Vergleich der Geschwindigkeiten über B-Splines und Kriging	60
5.8	Genauigkeit der Geschwindigkeit gegen die Geschwindigkeit selbst aufgetragen, Histogramm zur Geschwindigkeitsgenauigkeit zum alternativen Verfahren via B-Splines	61
6.1	Anteil des windgetriebenen Schlupfes an den Bojengeschwindigkeiten für den Januar 2009	64
6.2	Verteilung des durchschnittlichen Salzgehaltes (links) und der Temperatur (rechts) im Nordatlantik nach dem WOA09 für den Januar	67
6.3	Geschätzter realer (oben) und imaginärer (unten) Anteil des Parameters a_e	69
6.4	Ergebnisse der ganzjährlichen Schätzung des Parameters a_e , vertikale Striche stehen für die einfache Standardabweichung	71
6.5	Geschätzter realer und imaginärer Anteil des Parameters a_e für den tropischen Atlantik mit Bojendaten zwischen dem Januar 2007 und dem Dezember 2010	72
6.6	Strömungsbild des Ekman-Transports für das Jahr 2007 für Bojen mit Treibanker . .	73
7.1	Räumliche Modellierung der ADT in einem 2° -Gitter für das Jahr 2008	78
7.2	Gemittelte Geschwindigkeitsfelder der Monate Juli bis September 2009	79
A.1	Redundanzanteile einer fiktiven Messung zum Ausgleich stückweise kubischer Polynome	ii

Tabellenverzeichnis

2.1	Längen-, Geschwindigkeiten- und Zeitskalen für Strömungsphänomene im Ozean, nach CUSHMAN-ROISIN und BECKERS (2011, S. 15)	10
3.1	Empirisch bestimmte und durch das ARGOS-System ausgegebene Genauigkeiten für die Längen und Breitengrade der Bojenpositionen in Metern	14
3.2	Regionalisiert berechnete Parameter aus HANSEN und POULAIN (1996), beruhend auf Daten des Pazifiks zwischen 1988-1993	24
3.3	Neu-bestimmte Variogramm-Parameter beruhend auf Daten des Atlantiks zwischen 2008-2011	26
4.1	Räumliche Definition des Nord-Atlantik und der ausgewählten Unterregionen	31
4.2	Varianzen und Korrelationen der Modell- und Bojen-Geschwindigkeitssignale entsprechend der fünf Regionen und vier Zeitrahmen	34
4.3	Ergebnisse der ANOVA für vier Zeitrahmen und die drei Geschwindigkeitskomponenten	37

Liste der Algorithmen

1	Variation des MARS-Algorithmus zur optimalen Knotenwahl	49
2	Skizzierung der Funktion MARSapprox	55

Literatur

- AKAIKE, H. (1973) Information theory and an extension of the maximum likelihood principle. In: PETROV, B. N. und F. CSAKI (Hrsg.), *2nd International Symposium on Information Theory*, 267–281.
- ALBERTELLA, A., R. SAVCENKO, T. JANJIC, R. RUMMEL, W. BOSCH und J. SCHRÖTER (2012) High resolution dynamic ocean topography in the Southern Ocean from GOCE. *Geophysical Journal International*, 190(2):922–930.
- ANDERSEN, OLE B. und PER KNUDSEN (2009) DNSC08 mean sea surface and mean dynamic topography models. *Journal of Geophysical Research: Oceans*, 114(C11).
- ARBIC, B. K., R. B. SCOTT, D. B. CHELTON, J. G. RICHMAN und J. F. SHRIVER (2012) Effects of stencil width on surface ocean geostrophic velocity and vorticity estimation from gridded satellite altimeter data. *Journal of Geophysical Research*, 117.
- BARNES, S. L. (1964) A Technique for Maximizing Details in Numerical Weather Map Analysis. *Journal of applied Meteorology*, 3:396–409.
- BITTERMAN, D. S., P. P. NIILER, Y. Aoustin und A. DU CHAFFAUT (1990) Drift Buoy Inter-comparison Test Results. Technischer Bericht 17, NOAA Data Report ERL AOML, Miami, FL, USA.
- BOGNER, F. K., R. L. FOX und L. A. SCHMIT (1965) The generation of inter-element-compatible stiffness and mass matrices by the use of interpolation formulas. *Proceedings Conference on Matrix Methods in Structural Mechanics*, 397–444.
- BRUINSMA, S., J.-M. LEMOINE, R. BIANCALE und N. VALÈS (2010) CNES/GRGS 10-day gravity field models (release 2) and their evaluation. *Advances in Space Research*, 45:587–601.
- BURNHAM, K.P. und D.R. ANDERSON (1998) *Model Selection and Inference: A Practical Information-theoretic Approach*. Springer New York.
- BUSHNELL, M. (1996) Location Accuracy of Global Lagrangian Drifters. In: PANEL, DATA BUOY COOPERATION (Hrsg.), *Developments in Buoy and Communication Technologies*, 10. NOAA/ AOML-Global Drifter Center. technisches Dokument.
- CNES/AVISO. (2014) , SSALTO/DUACS User Handbook: (M)SLA and (M)ADT Near-Real Time and Delayed Time Products. http://www.avisio.altimetry.fr/fileadmin/documents/data/tools/hdbk_duacs.pdf, Mai 2014.
- CRESSIE, N. A. C. (1991) *Statistics for Spatial Data*. John Wiley and Sons Inc.

- CUSHMAN-ROISIN, B. (1987) Subduction. In: MULLER, P. und D. HENDERSON (Hrsg.), *Dynamics of Oceanic Surface Mixed Layer*, International Geophysical Research. Hawaii Institute of Geophysics Special Publication, University of Hawaii, 181 – 196.
- CUSHMAN-ROISIN, B. und J.M. BECKERS (2011) *Introduction to Geophysical Fluid Dynamics – Physical and Numerical Aspects*, Band 101 der Reihe International Geophysical Research. Academic Press, 2. Ausgabe.
- DAUBRAWA, J. (2007) Bahnstörungen durch Ozeangezeiten. Diplomarbeit, Geodätisches Institut, Universität Stuttgart.
- DUHAUT, T. H. A. und D. N. STRAUB (2005) Wind Stress Dependence on Ocean Surface Velocity: Implications for Mechanical Energy Input to Ocean Circulation. *Journal of Physical Oceanography*, 36:202 – 211.
- EARSEL. (2014) , SEOS Project. <http://www.seos-project.eu/modules/oceancurrents/oceancurrents-c02-p02.de.html>, September 2014. European Association of Remote Sensing Laboratories.
- EINSTEIN, A. (1905) Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8).
- FAHRMEIR, L., R. KÜNSTLER, I. PIGEOT und G. TUTZ (1999) *Statistik – Der Weg zur Datenanalyse*. Springer Verlag: Berlin, Heidelberg, verbesserte 2. Ausgabe.
- FAHRMEIR, L., T. KNEIB und S. LANG (2009) *Regression – Modelle, Methoden und Anwendungen*. Statistik und ihre Anwendungen. Springer Verlag: Berlin, Heidelberg, 2. Ausgabe.
- FRANKEN, J. (2012) *Flächenhafte Modellierung orts- und zeitabhängiger Höhenänderungen*. Bachelorarbeit. Institut für Theoretische Geodäsie der Universität Bonn, Bonn.
- FRIEDMAN, J. H. (1991) Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1):1–141.
- FRIEDMAN, J. H. und B. W. SILVERMAN (1989) Flexible parsimonious smoothing and additive modeling. *Technometrics*, 31:3–39.
- HANSEN, D. V. und A. HERMAN (1989) Temporal Sampling Requirements for Surface Drifting Buoys in the Tropical Pacific. *American Meteorological Society*, 599–607.
- HANSEN, D. V. und P.-M. POULAIN (1996) Quality Control and Interpolation of WOCE-TOGA Drifter Data. *Journal of Atmospheric and Oceanic Technology*, 13:900–909.
- HASTIE, T., R. TIBSHIRANI und A. BUJA (1994) Flexible Discriminant Analysis by Optimal Scoring. *Journal of the American Statistical Association*, 89:1255–1270.
- HÄMMERLIN, G. und K. H. HOFFMANN (1994) *Numerische Mathematik*. Springer-Lehrbuch. Springer Berlin Heidelberg.

- ISAAKS, E. H. und R. M. SRIVASTAVA (1989) *An Introduction to Applied Geostatistics*. Oxford University Press, Oxford.
- JELIC, S. (2012) Klimaforschung im Ozean – Veränderte Ozeanströmungen und zukünftiges Klimageschehen. Bundesministerium für Bildung und Forschung – Referat System Erde, Bonn.
- JOSEPH, A. (2014) *Measuring Ocean Currents: Tools, Technologies, and Data*. Elsevier, San Diego, USA, 1.. Ausgabe.
- JOURNEL, A.G. und CH. J. HUIJBREGTS (1978) *Mining Geostatistics*. Academic Press, London, San Diego.
- KANTHA, L. H. und C. A. CLAYSON (2000) *Numerical Models of Oceans and Oceanic Processes*. International Geophysics. Elsevier Science.
- KOCH, K. R. (2004) *Parameterschätzung und Hypothesentests in linearen Modellen*. 4. Auflage. Dümmler, Bonn.
- KREYSZIG, E. (1991) *Statistische Methoden und ihre Anwendung*. Vadenhoek und Ruprecht, Göttingen, 7. Ausgabe.
- LAGERLOEF, G. S. E., G. T. MITCHUM, R. B. LUKAS und P. P. NIILER (1999) Tropical Pacific near-surface currents estimated from altimeter, wind, and drifter data. *Journal of Geophysical Research*, 104(C10):23313 – 23326.
- LOCARNINI, R. A., A. V. MISHONOV, J. I. ANTONOV, T. P. BOYER, H. E. GARCIAA, O. K. BARANOVA, M. M. ZWENG und D. R. JOHNSON (2010) . LEVITUS, S. (Hrsg.), *World Ocean Atlas 2009*, Band 1 und 2, NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 Seiten.
- LOPEZ, R. und J.-P. MALARDÈ (2011) Improving ARGOS Doppler Location Using Kalman Filtering. Technischer Bericht, CLS - Collecte Localisation Satellites.
- LUMPKIN, R. und M. PAZOS (2005) Measuring surface currents with Surface Velocity Program drifter: the instrument, its data, and some recent results. *National Oceanographic an Atmospheric Administration, Atlantic Oceanographix and Meteorological Laboratory, Miami, FL USA*.
- MAXIMENKO, N., P. NIILER, M. H. RIO, O. MELNICHENKO, L. CENTURIONI, D. CHAMBERS, V. ZLOTNICKI und B. GALPERIN (2009) Mean dynamic topography of the ocean derived from satellite and drifting buoy data using three different techniques. *Journal of Atmospheric and Oceanic Technology*, 26(9):1910–1919.
- MENARD, H.W. und S.M. SMITH (1966) Hypsometry of ocean basin provinces. *Journal of Geophysical Research*, 71:4305 – 4325.
- MILLERO, F. J., C.-T. CHEN, A. BRADSHAW und K. SCHLEICHER (1980) A new high pressure equation of state for seawater. *Deep Sea Research Part A. Oceanographic Research Papers*, 27:255 – 264.

- NIILER, P. P. und J. D. PADUAN (1995) Wind-Driven Motions in the Northeast Pacific as Measured by Lagrangian Drifters. *Journal of Physical Oceanography*, 25:2819 – 2830.
- NIILER, P. P., A. SYBRANDY, K. BI, P. POULAIN und D. BITTERMAN (1995) Measurements of the water-following capability of holey-sock and TRISTAR drifters. *Deep Sea Research*, 42:1951 – 1964.
- OLBERS, D., J. WILLEBRAND und C. EDEN (2012) *Ocean Dynamics*. Springer, Berlin, Heidelberg.
- PACANOWSKI, R. C. (1987) Effect of Equatorial Currents on Surface Stress. *Journal of Physical Oceanography*, 17:833 – 838.
- PACIFIC GYRE. (2013) , SVP Lagrangian Drifter. www.pacificgyre.com/svp-gps-data-buoys.aspx, November 2013. SVP Specification Sheet.
- POULAIN, P.-M., R. GERIN, E. MAURI und R. PENNEL (2009) Wind effects on drogued and undrogued drifters in the eastern Mediterranean. *Journal of Atmospheric and Oceanic Technology*, 26:1144 – 1156.
- RAHMSTORF, S. (2002) Ocean circulation and climate during the past 120,000 years. *Nature*, 419(6903):207–214.
- RALPH, E. und P. NIILER (1999) Wind-driven Currents in the Tropical Pacific. *Journal of Physical Oceanography*, 29:2121–2129.
- RIO, M.-H. (2012) Use of Altimeter and Wind Data to Detect the Anomalous Loss of SVP-Type Drifters Drogue. *Journal of Atmospheric and Oceanic Technology*, 29:1663–1674.
- RIO, M. H. und F. HERNANDEZ (2004) A mean dynamic topography computed over the world ocean from altimetry, in situ measurements, and a geoid model. *Journal of Geophysical Research*, 109:C12032.
- RIO, M.-H., P. SCHAEFFER, J.-M. LEMOINE und F. HERNANDEZ (2005) Estimation of the ocean Mean Dynamic Topography through the combination of altimetric data, in-situ measurements and GRACE geoid: From global to regional studies.
- RIO, M. H., S. GUINEHU und G. LARNICOL (2011) New CNES-CLS09 global mean dynamic topography computed from the combined of GRACE data, altimetry and in situ measurements. *Journal of Geophysical Research*, 116:C07018.
- SANDERSON, B. und D. A. BOOTH (1991) The fractal dimension of drifter trajectories and estimates of horizontal eddy-diffusivity. *Tellus*, 43A:334–349.
- SCHÖNWIESE, C.D. (2003) *Klimatologie*. UTB für Wiessenschaft: Uni-Taschenbücher. Ulmer, 3. Ausgabe.
- SIEDLER, G., S. M. GRIFFIES, J. GOULD und J. A. CHURCH (Hrsg.) (2013) *Ocean Circulation and Climate: A 21st Century Perspective*. International Geophysics. Academic Press.

-
- SPINTALL, J. und M. F. CRONIN (2001) Upper Ocean vertical structure. *Encyclopedia of Ocean Sciences*, 6:3118 – 3126.
- STEWART, R. H. (2008) *Introduction to Physical Oceanography*. Department of Oceanography, Texas A & M University.
- STONE, C. J. (1982) Local asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Annals of the Institute of Statistical Mathematics*, 34(Part A):123–133.
- SUDRE, J., C. MAES und V. GARÇON (2013) On the global estimates of geostrophic and Ekman surface currents. *Limnology and Oceanography: Fluid and Environment*, 3:1–20.