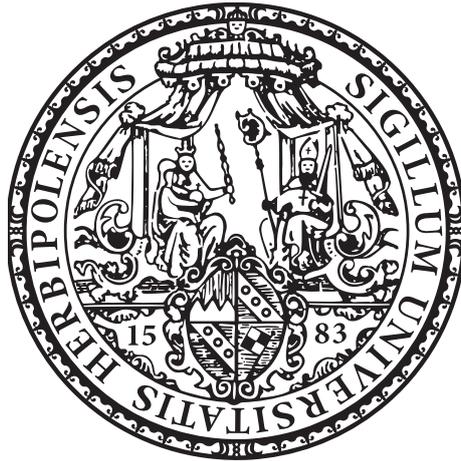


Dissertation zur Erlangung des  
naturwissenschaftlichen Doktorgrades der  
Bayerischen Julius-Maximilians-Universität  
Würzburg



**Bewertung und Auswirkungen  
der Simulationsgüte führender  
Klimamoden in einem  
Multi-Modell Ensemble**

vorgelegt von  
Felix Pollinger  
aus Erlenbach a. Main

Würzburg 2013



Eingereicht am:

13. November 2013

1. Gutachter:

Prof. Dr. Heiko Paeth

2. Gutachter:

Prof. Dr. Dieter Scherer

der Dissertation

1. Prüfer:

Prof. Dr. Heiko Paeth

2. Prüfer:

Prof. Dr. Hartwig Frimmel

der mündlichen Prüfung

Tag der mündlichen Prüfung:

9. April 2014

Doktorurkunde ausgehändigt am:



*People look down on stuff like geography and meteorology, and not only because they're standing on one and being soaked by the other. They don't look quite like real science. But geography is only physics slowed down and with a few trees stuck on it, and meteorology is full of excitingly fashionable chaos and complexity.*

*– Terry Pratchett*



# Zusammenfassung

Der rezente und zukünftige Anstieg der atmosphärischen Treibhausgaskonzentration bedeutet für das terrestrische Klimasystem einen grundlegenden Wandel, der für die globale Gesellschaft schwer zu bewältigende Aufgaben und Herausforderungen bereit hält. Eine effektive, frühzeitige Anpassung an diesen Klimawandel profitiert dabei enorm von möglichst genauen Abschätzungen künftiger Klimaänderungen. Das geeignete Werkzeug hierfür sind Gekoppelte Atmosphäre Ozean Modelle (AOGCMs). Für solche Fragestellungen müssen allerdings weitreichende Annahmen über die zukünftigen klimarelevanten Randbedingungen getroffen werden. Individuelle Fehler dieser Klimamodelle, die aus der nicht perfekten Abbildung der realen Verhältnisse und Prozesse resultieren, erhöhen die Unsicherheit langfristiger Klimaprojektionen. So unterscheiden sich die Aussagen verschiedener AOGCMs im Hinblick auf den zukünftigen Klimawandel insbesondere bei regionaler Betrachtung, deutlich. Als Absicherung gegen Modellfehler werden üblicherweise die Ergebnisse mehrerer AOGCMs, eines Ensembles an Modellen, kombiniert. Um die Abschätzung des Klimawandels zu präzisieren, wird in der vorliegenden Arbeit der Versuch unternommen, eine Bewertung der Modellperformance der 24 AOGCMs, die an der dritten Phase des Vergleichsprojekts für gekoppelte Modelle (CMIP3) teilgenommen haben, zu erstellen. Auf dieser Basis wird dann eine numerische Gewichtung für die Kombination des Ensembles erstellt.

Zunächst werden die von den AOGCMs simulierten Klimatologien für einige grundlegende Klimaelemente mit den betreffenden Klimatologien verschiedener Beobachtungsdatensätze quantitativ abgeglichen. Ein wichtiger methodischer Aspekt hierbei ist, dass auch die Unsicherheit der Beobachtungen, konkret Unterschiede zwischen verschiedenen Datensätzen, berücksichtigt werden. So zeigt sich, dass die Aussagen, die aus solchen Ansätzen resultieren, von zu vielen Unsicherheiten in den Referenzdaten beeinträchtigt werden, um generelle Aussagen zur Qualität von AOGCMs zu treffen. Die Nutzung der Köppen-Geiger Klassifikation offenbart jedoch, dass die prinzipielle Verteilung der bekannten Klimatypen im kompletten CMIP3 in vergleichbar guter Qualität reproduziert wird.

Als Bewertungskriterium wird daher hier die Fähigkeit der AOGCMs die großskalige natürliche Klimavariabilität, konkret die hochkomplexe gekoppelte El Niño-Southern Oscillation (ENSO), realistisch abzubilden herangezogen. Es kann anhand verschiedener Aspekte des ENSO-Phänomens gezeigt werden, dass nicht alle AOGCMs hierzu mit gleicher Realitätsnähe in der Lage sind. Dies steht im Gegensatz zu den dominierenden Klimamoden der Außertropen, die modell-

übergreifend überzeugend repräsentiert werden. Die wichtigsten Moden werden, in globaler Betrachtung, in verschiedenen Beobachtungsdaten über einen neuen Ansatz identifiziert. So können für einige bekannte Zirkulationsmuster neue Indexdefinitionen gewonnen werden, die sich sowohl als äquivalent zu den Standardverfahren erweisen und im Vergleich zu diesen zudem eine deutliche Reduzierung des Rechenaufwandes bedeuten. Andere bekannte Moden werden dagegen als weniger bedeutsame, regionale Zirkulationsmuster eingestuft. Die hier vorgestellte Methode zur Beurteilung der Simulation von ENSO ist in guter Übereinstimmung mit anderen Ansätzen, ebenso die daraus folgende Bewertung der gesamten Performance der AOGCMs. Das Spektrum des Southern Oscillation-Index (SOI) stellt somit eine aussagekräftige Kenngröße der Modellqualität dar.

Die Unterschiede in der Fähigkeit, das ENSO-System abzubilden, erweisen sich als signifikante Unsicherheitsquelle im Hinblick auf die zukünftige Entwicklung einiger fundamentaler und bedeutsamer Klimagrößen, konkret der globalen Mitteltemperatur, des SOIs selbst, sowie des indischen Monsuns. Ebenso zeigen sich signifikante Unterschiede für regionale Klimaänderungen zwischen zwei Teilensembles des CMIP3, die auf Grundlage der entwickelten Bewertungsfunktion eingeteilt werden. Jedoch sind diese Effekte im Allgemeinen nicht mit den Auswirkungen der anthropogenen Klimaänderungssignale im Multi-Modell Ensemble vergleichbar, die für die meisten Klimagrößen in einem robusten multivariaten Ansatz detektiert und quantifiziert werden können. Entsprechend sind die effektiven Klimaänderungen, die sich bei der Kombination aller Simulationen als grundlegende Aussage des CMIP3 unter den speziellen Randbedingungen ergeben nahezu unabhängig davon, ob alle Läufe mit dem gleichen Einfluss berücksichtigt werden, oder ob die erstellte numerische Gewichtung verwendet wird. Als eine wesentliche Begründung hierfür kann die Spannbreite der Entwicklung des ENSO-Systems identifiziert werden. Dies bedeutet größere Schwankungen in den Ergebnissen der Modelle mit funktionierendem ENSO, was den Stellenwert der natürlichen Variabilität als Unsicherheitsquelle in Fragen des Klimawandels unterstreicht. Sowohl bei Betrachtung der Teilensembles als auch der Gewichtung wirken sich dadurch gegenläufige Trends im SOI ausgleichend auf die Entwicklung anderer Klimagrößen aus, was insbesondere bei letzterem Vorgehen signifikante mittlere Effekte des Ansatzes, verglichen mit der Verwendung des üblichen arithmetischen Multi-Modell Mittelwert, verhindert.

# Abstract

The recent and future increase in atmospheric greenhouse gases will cause fundamental change in the terrestrial climate system, which will lead to enormous tasks and challenges for the global society. Effective and early adaptation to this climate change will benefit hugely from optimal possible estimates of future climate change. Coupled atmosphere-ocean models (AOGCMs) are the appropriate tool for this. However, to tackle these questions, it is necessary to make far reaching assumptions about the future climate-relevant boundary conditions. Furthermore there are individual errors in each climate model. These originate from flaws in reproducing the real climate system and result in a further increase of uncertainty with regards to long-range climate projections. Hence, concerning future climate change, there are pronounced differences between the results of different AOGCMs, especially under a regional point of view. It is the usual approach to use a number of AOGCMs and combine their results as a safety measure against the influence of such model errors. In this thesis, an attempt is made to develop a valuation scheme and based on that a weighting scheme, for AOGCMs in order to narrow the range of climate change projections. The 24 models that were included in the third phase of the coupled model intercomparison project (CMIP3) are used for this purpose.

First some fundamental climatologies simulated by the AOGCMs are quantitatively compared to a number of observational data. An important methodological aspect of this approach is to explicitly address the uncertainty associated with the observational data. It is revealed that statements concerning the quality of climate models based on such hindcastig approaches might be flawed due to uncertainties about observational data. However, the application of the Köppen-Geiger classification reveals that all considered AOGCMs are capable of reproducing the fundamental distribution of observed types of climate.

Thus, to evaluate the models, their ability to reproduce large-scale climate variability is chosen as the criterion. The focus is on one highly complex feature, the coupled El Niño-Southern Oscillation. Addressing several aspects of this climate mode, it is demonstrated that there are AOGCMs that are less successful in doing so than others. In contrast, all models reproduce the most dominant extratropical climate modes in a satisfying manner. The decision which modes are the most important is made using a distinct approach considering several global sets of observational data. This way, it is possible to add new definitions for the time series of some well-known climate patterns, which prove to be equivalent to

the standard definitions. Along with this, other popular modes are identified as less important regional patterns. The presented approach to assess the simulation of ENSO is in good agreement with other approaches, as well as the resulting rating of the overall model performance. The spectrum of the timeseries of the Southern Oscillation Index (SOI) can thus be regarded as a sound parameter of the quality of AOGCMs.

Differences in the ability to simulate a realistic ENSO-system prove to be a significant source of uncertainty with respect to the future development of some fundamental and important climate parameters, namely the global near-surface air mean temperature, the SOI itself and the Indian monsoon. In addition, there are significant differences in the patterns of regional climate change as simulated by two ensembles, which are constituted according to the evaluation function previously developed. However, these effects are overall not comparable to the multi-model ensembles' anthropogenic induced climate change signals which can be detected and quantified using a robust multi-variate approach. If all individual simulations following a specific emission scenario are combined, the resulting climate change signals can be thought of as the fundamental message of CMIP3. It appears to be quite a stable one, more or less unaffected by the use of the derived weighting scheme instead of the common approach to use equal weights for all simulations. It is reasoned that this originates mainly from the range of trends in the SOI. Apparently, the group of models that seems to have a realistic ENSO-system also shows greater variations in terms of effective climate change. This underlines the importance of natural climate variability as a major source of uncertainty concerning climate change. For the SOI there are negative trends in the multi-model ensemble as well as positive ones. Overall, these trends tend to stabilize the development of other climate parameters when various AOGCMs are combined, whether the two distinguished parts of CMIP3 are analyzed or the weighting scheme is applied. Especially in case of the latter method, this prevents significant effects on the mean change compared to the arithmetic multi-model mean.

---

# Inhaltsverzeichnis

<b>Zusammenfassung</b>	<b>vii</b>
<b>Abstract</b>	<b>ix</b>
<b>Tabellenverzeichnis</b>	<b>xv</b>
<b>Abbildungsverzeichnis</b>	<b>xvi</b>
<b>Abkürzungsverzeichnis</b>	<b>xviii</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Der Weltklimarat und die Erforschung von Klimaänderungen . . .	1
1.2 Unsicherheit, Modellgüte und Klimaprojektionen . . . . .	3
1.3 Eigenmoden und Telekonnektionen . . . . .	11
<b>2 Daten</b>	<b>17</b>
2.1 Beobachtungsdaten . . . . .	17
2.1.1 Felddaten . . . . .	17
2.1.2 Zeitreihen . . . . .	22
2.1.2.1 Zirkulationsindices . . . . .	22
2.1.2.2 Die globale Mitteltemperatur . . . . .	23
2.2 Modelldaten . . . . .	25
2.2.1 Randbedingungen . . . . .	25
2.2.2 AOGCMs . . . . .	27
2.3 Datentransformation (räumliche Interpolation) . . . . .	29
2.3.1 Inverse-Distanz Wichtung . . . . .	31
2.3.2 Flächengewichtete Interpolation . . . . .	31
2.3.3 Bivariate quadratische Interpolation . . . . .	32
<b>3 Statistische Methoden</b>	<b>34</b>
3.1 Grundlegende Datenaufbereitung und Inferenz . . . . .	34
3.1.1 Univariate Kenngrößen . . . . .	35
3.1.2 Inferenz . . . . .	36
3.1.2.1 Zweistichproben- <i>t</i> -Test . . . . .	36
3.1.2.2 <i>F</i> -Test für Varianzen . . . . .	38

3.2	Korrelation . . . . .	38
3.3	Lineare Regression . . . . .	40
3.3.1	Mathematische Grundlagen . . . . .	40
3.3.2	Modellgüte und Inferenz . . . . .	42
3.3.3	Polynomregression . . . . .	44
3.4	ANOVA . . . . .	44
3.4.1	Mathematische Durchführung . . . . .	44
3.4.2	Spektrale ANOVA . . . . .	46
3.5	EOF-Analyse . . . . .	47
3.5.1	Begriffe und Anwendung . . . . .	47
3.5.2	Grundgleichungen . . . . .	48
3.5.3	Varimax-Rotation . . . . .	50
3.5.4	Verwendete Skalierung und Darstellung . . . . .	51
3.6	Zeitreihentechniken . . . . .	52
3.6.1	Additives Zeitreihenmodell . . . . .	53
3.6.2	Kenngößen stationärer Zeitreihen . . . . .	54
3.6.2.1	Auto- und Kreuzkorrelation . . . . .	55
3.6.2.2	Periodogramm und Spektralschätzer . . . . .	56
<b>4</b>	<b>Grundlegende Modellvalidation</b>	<b>59</b>
4.1	Separative Betrachtung . . . . .	60
4.1.1	Beobachtungsdaten: kombinierte Klimatologien . . . . .	60
4.1.1.1	Bodennahe Temperatur . . . . .	61
4.1.1.2	Niederschlag . . . . .	62
4.1.1.3	Luftdruck auf Meeressniveau . . . . .	63
4.1.2	Methodik . . . . .	65
4.1.2.1	Konfidenzintervalle . . . . .	65
4.1.2.2	Root-Mean-Square Error . . . . .	66
4.1.3	Ergebnisse . . . . .	67
4.2	Klimatypen in Klimamodellen . . . . .	72
4.2.1	Bedeutung der effektiven Klimaklassifikation nach Köppen	73
4.2.2	Ergebnisse . . . . .	75
<b>5</b>	<b>Evaluation der großskaligen Klimavariabilität</b>	<b>80</b>
5.1	Kriteriumsmoden: Definition und Begründung . . . . .	80
5.2	Globale Detektion: Methodik . . . . .	82
5.3	Ergebnisse . . . . .	84
<b>6</b>	<b>Index-Definitionen</b>	<b>89</b>
6.1	Außertropische Zirkulationsmuster . . . . .	89
6.1.1	Raummuster . . . . .	89

6.1.2	Anwendung im CMIP3 . . . . .	91
6.1.2.1	Äquivalenz zu Standardindices . . . . .	92
6.1.2.2	Modellinterne Repräsentation . . . . .	94
6.2	ENSO und die Monsune . . . . .	97
6.2.1	SOI aus regionalem EOF-Ansatz . . . . .	98
6.2.2	Monsune . . . . .	101
6.3	Indexentwicklung im CMIP3 . . . . .	104
<b>7</b>	<b>Das ENSO-Phänomen im Multi-Modell Ensemble</b>	<b>107</b>
7.1	ENSO-Charakteristika in AOGCMs . . . . .	107
7.1.1	Raummuster der Variabilität . . . . .	108
7.1.2	Varianzspektrum . . . . .	111
7.1.3	Telekonnektionen . . . . .	115
7.1.4	ENSO und die globale Temperatur . . . . .	119
7.2	ENSO als Aspekt der Modell-Performance . . . . .	123
7.2.1	Erläuterung des Ansatzes . . . . .	123
7.2.2	Bewertungs- und Gewichtungsfunktion . . . . .	124
7.3	Vergleich mit anderen Ansätzen . . . . .	127
7.3.1	Das ENSO-Problem des CMIP3 . . . . .	127
7.3.2	Qualität der Modelle: Rangplatzfolgen . . . . .	128
<b>8</b>	<b>Erwärmung und ENSO (modellspezifisch)</b>	<b>131</b>
8.1	Globale Mitteltemperatur und SOI: Projektionen im CMIP3 . . . . .	131
8.2	SOI-Signal: Detektion und Entfernung . . . . .	135
<b>9</b>	<b>Klimaänderungen: Signalanalyse der Zeitreihen</b>	<b>141</b>
9.1	Methodik . . . . .	142
9.2	Ergebnisse . . . . .	144
9.2.1	SOI . . . . .	145
9.2.2	Globale Mitteltemperatur . . . . .	147
9.2.3	EIMR . . . . .	147
9.2.4	AMR . . . . .	149
9.2.5	NAM . . . . .	150
9.2.6	SAM . . . . .	152
9.2.7	NPI . . . . .	153
9.2.8	SCA . . . . .	155
<b>10</b>	<b>ENSO-Performance und Klimaänderungen: Zeitreihenentwicklung</b>	<b>157</b>
10.1	Systematische Differenzen: Existenz und Persistenz . . . . .	157
10.1.1	Methodik . . . . .	157
10.1.2	Ergebnisse . . . . .	158

---

10.2	Transiente Varianzattribu- tion . . . . .	163
10.2.1	Methodik . . . . .	163
10.2.2	Ergebnisse . . . . .	164
10.2.2.1	SOI . . . . .	164
10.2.2.2	Globale Mitteltemperatur . . . . .	166
10.2.2.3	EIMR . . . . .	168
10.2.2.4	NPI . . . . .	170
<b>11</b>	<b>ENSO-Performance und Klimaänderungen: Regionale Betrachtung</b>	<b>172</b>
11.1	Methodik . . . . .	172
11.2	Ergebnisse . . . . .	174
11.2.1	Effekte auf die bodennahe Temperatur . . . . .	174
11.2.2	Effekte auf die Entwicklung des Niederschlags . . . . .	177
11.2.3	Effekte auf die Luftdruckverteilung . . . . .	180
<b>12</b>	<b>Gewichtete Ensemble-Projektionen</b>	<b>183</b>
12.1	Methodik . . . . .	183
12.2	Ergebnisse . . . . .	185
12.2.1	Zeitreihen: Die transiente Entwicklung . . . . .	186
12.2.2	Zeitreihen: Der effektive Wandel . . . . .	189
12.2.3	Regionaler Wandel . . . . .	194
<b>13</b>	<b>Diskussion und Ausblick</b>	<b>198</b>
	<b>Literaturverzeichnis</b>	<b>205</b>
	<b>Danksagung</b>	<b>237</b>

---

# Tabellenverzeichnis

2.1	Beobachtungsdaten (Felddaten) . . . . .	18
2.2	Beobachtungsdaten (Zeitreihen) . . . . .	24
2.3	Überblick AOGCM-Daten (SRES A1b, CMIP3) . . . . .	28
5.1	Global erfassbare Klimamoden . . . . .	85
6.1	Definitionen der Zeitreihen $I_S$ . . . . .	92
6.2	Korrelation NCAR-SOI und 1. PC aus regionalem Ansatz für Beobachtungen . . . . .	99
6.3	Korrelationen Beobachtungsdaten (EIMR und AMR) . . . . .	103
7.1	CMIP3: Gruppierung nach ENSO . . . . .	128
7.2	Korrelation Modellbewertungen . . . . .	129
8.1	Globale Erwärmung und ENSO-Einfluss im SRES A1b . . . . .	139
9.1	Klimatische Kenngrößen in CMIP3: Trends und Varianzzerlegung . . . . .	144
12.1	Räumlicher Effekt von Gewichtung und Ensembletrennung . . . . .	195

# Abbildungsverzeichnis

4.1	Beobachtungsdaten für 1960–1999 . . . . .	60
4.2	Kombination Beobachtungen: Temperatur . . . . .	61
4.3	Kombination Beobachtungen: Niederschlag . . . . .	63
4.4	Kombination Beobachtungen: Luftdruck . . . . .	64
4.5	Komposite Modell- und Beobachtungsdaten . . . . .	68
4.6	Übersicht zur Evaluation des Multi-Modell Ensembles . . . . .	71
4.7	Klimatypen für Beobachtungen und AOGCMs . . . . .	76
4.8	Klimatypen im Multi-Modell Ensemble . . . . .	78
5.1	Raummuster in Beobachtungen und Modellen . . . . .	87
6.1	Ektropische, global relevante Zirkulationsmuster . . . . .	90
6.2	Deduktive und induktive Indices . . . . .	93
6.3	Modellinterne Erzeugung von Zirkulationsmoden . . . . .	97
6.4	Raummuster SOI . . . . .	98
6.5	PCs als modelleigener SOI . . . . .	100
6.6	Flächenindices: EIMR und AMR . . . . .	102
6.7	Klimamoden: Entwicklung im SRES A1b . . . . .	106
7.1	Modelle: 1. EOF im tropischen Pazifik . . . . .	109
7.2	Varianzspektrum NCAR-SOI . . . . .	111
7.3	CMIP3 SOI-Varianzspektren (1) . . . . .	113
7.4	CMIP3 SOI-Varianzspektren (2) . . . . .	114
7.5	El Niño und La Niña: Temperatur . . . . .	116
7.6	El Niño und La Niña: Niederschlag . . . . .	118
7.7	NCAR-SOI und Gisstemp- $T_E$ . . . . .	120
7.8	SOI und $T_E$ im SRES A1b . . . . .	121
7.9	Kumulierte Bewertungsfunktion für CMIP3 . . . . .	126
8.1	Erwärmung je Teilensemble . . . . .	132
8.2	SOI- und Temperaturentwicklung in CMIP3 . . . . .	134
8.3	Zusammenhang $\Delta T$ und $\Delta I$ . . . . .	135
8.4	Erwärmung je Teilensemble ohne SOI . . . . .	137
8.5	$\Delta T_E$ im SRES A1b . . . . .	138

9.1	Ausgewählte Ergebnisse der Spektralen ANOVA für SOI, $T_E$ , EIMR und AMR. . . . .	146
9.2	Ausgewählte Ergebnisse der Spektralen ANOVA für NAM, SAM, SCA und NPI. . . . .	151
10.1	Systematische Unterschiede zwischen den Teilensembles für vier Klimagrößen . . . . .	160
10.2	SOI: Transiente Varianzattribuion . . . . .	166
10.3	$T_E$ : Transiente Varianzattribuion . . . . .	168
10.4	EIMR: Transiente Varianzattribuion . . . . .	169
10.5	NPI: Transiente Varianzattribuion . . . . .	171
11.1	Räumliche Differenzierung von $\Delta T$ . . . . .	175
11.2	Räumliche Differenzierung von $\Delta R$ . . . . .	179
11.3	Räumliche Differenzierung von $\Delta P$ . . . . .	181
12.1	Einfluss pro AOGCM auf Ensemblekenngrößen . . . . .	184
12.2	SOI und EIMR: Verlauf gewichtet/ungewichtet. . . . .	188
12.3	Varianzverhältnisse einiger Klimagrößen. . . . .	189
12.4	PDFs Klimaänderungen: $T_E$ , SOI, EIMR, AMR . . . . .	191
12.5	PDFs Klimaänderungen: SAM, NAM, NPI, SCA . . . . .	193
12.6	Gewichtete Streuung der Ensembleprojektionen. . . . .	197

# Abkürzungsverzeichnis

AMR	West African Monsoon Rainfall
ANOVA	Varianzanalyse
(AO)GCM	(Coupled Atmosphere Ocean) Generell Circulation Model
BAS	British Antarctic Survey
CMIP3	Coupled Model Intercomparison Project; Phase 3
CPCC	Global Precipitation Climatology Centre
CPCP	Global Precipitation Climatology Project
COWL	Cold Ocean–Warm Land
CRU	Climate Research Unit
EA	East Atlantic Pattern
EA/WP	East Atlantic/West Russia Pattern
EP/NP	East Pacific/North Pacific Pattern
EIMR	Extended Indian Monsoon Rainfall
ENSO	El Niño–Southern Oscillation
IDW	Inverse-Distanz Wichtung
IPCC	Intergovernmental Panel on Climate Change
JAXA	Japan Aerospace Exploration Agency
NAO	Nordatlantische Oszillation
NAM	Northern Annular Mode/Arctic Oscillation
NASA	National Aeronautics and Space Administration
NCAR	National Center of Atmospheric Research
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
NPI	Nordpazifik Index
PNA	Pacific/North America Pattern
POL	Polar/Eurasia Pattern
$r$	Korrelationskoeffizient der Produkt-Moment-Korrelation
(R)EOF	(Rotierte) Empirische Orthogonal Funktion

(R)PC	(Rotierte) Hauptkomponente
RMSE	Root-Mean-Square Error
SAM	Southern Annular Mode/Antarctic Oszillation
SCA	Scandinavian Pattern
SLP	Sea Level Pressure
SOI	Southern Oscillation Index
SRES	Special Report on Emissions Scenarios
$T_E$	Global gemittelte oberflächennahe Temperatur
TRMM	Tropical Rainfall Measuring Mission
WG	Working Group
WCRP	World Climate Research Programme
$x(t)$	Zeitreihe $x$
$x(t, s)$	Feldgröße $x$
$x(s)$	regionalisierte Größe $x$



# 1 Einleitung

Die vorliegende Arbeit beschäftigt sich mit einer aktuellen und relevanten Fragestellung im Kontext der wissenschaftlichen Erforschung des anthropogenen Klimawandels. Ein bedeutender Teil der Abschätzung des künftigen Klimawandels beruht auf den Ergebnissen von mathematischen Klimasimulatoren, Klimamodellen. Durch das Verfolgen unterschiedlicher Ansätze zur realistischen Abbildung des terrestrischen Klimasystems, weisen auch die modernsten und aufwendigsten Klimamodelle Unterschiede zueinander auf. Hieraus resultieren unterschiedliche Abschätzungen der klimatischen Reaktion auf ein vorgegebenes, identisches Forcing durch Treibhausgase. So ergibt sich häufig die Notwendigkeit, verschiedene und in einigen, insbesondere regionalen, Aspekten auch widersprüchliche Ergebnisse unterschiedliche Klimasimulationen zu bewerten und konkretere Abschätzungen der tatsächlichen Unsicherheit vorzunehmen. Hierbei sollte die Qualität und Zuverlässigkeit der betrachteten Klimamodelle berücksichtigt werden, um die bestmögliche Information aus den Modellrechnungen extrahieren zu können.

Es existiert jedoch kein anerkanntes oder auch nur verbreitetes Verfahren, um eine solche Abschätzung der Modellqualität, auch etwa beim direkten Vergleich verschiedener Modelle, vorzunehmen. In dieser Arbeit wird versucht, ein solches zu entwickeln. In einem induktiven Ansatz wird ein Bewertungskriterium erschlossen und eine Bewertungsmetrik erstellt. Ebenso werden die Effekte dieser Beurteilung der Modellqualität auf langfristige Klimaprojektionen überprüft. In diesem Kapitel wird knapp die Tätigkeit des Weltklimarats als zentrale Organisation bei der Koordination und Konzentration der aktuellen klimatologischen Forschung vorgestellt. Weiterhin erfolgt ein Überblick über bisherige Ansätze, die Unsicherheit in langfristigen Projektionen von Multi-Modell Ensembles zu reduzieren. Anschließend werden in allgemeiner Form großskalige Klimamoden oder Zirkulationsmuster als bedeutsame Aspekte des Klimasystems thematisiert, da sich die Bewertung der Modelle auf diese stützen wird. Zudem werden in dieser Arbeit auch weitere ertragreiche Analysen zu diesem Themenbereich präsentiert.

## 1.1 Der Weltklimarat und die Erforschung von Klimaänderungen

Der Zwischenstaatliche Ausschuss für Klimaänderungen (Intergovernmental Panel on Climate Change, IPCC), auch Weltklimarat, wurde 1988 gemeinsam von

der World Meteorological Organisation und dem Umweltprogramm der Vereinten Nationen gegründet. Von Anfang an bestand der Auftrag des IPCCs darin, die aktuellen Entwicklungen in der wissenschaftlichen Erforschung des Klimawandels zu erfassen und dokumentieren. Die Struktur ist hierbei dreigeteilt, um die vielfältigen Aspekte dieser Thematik adäquat zu repräsentieren. Die Arbeitsgruppe I (WG1) thematisiert die naturwissenschaftlichen Grundlagen von Klimaänderungen, wobei der Fokus sehr weit gefasst ist und sich nicht allein auf den rezenten Klimawandel und die Entwicklung des Klimasystems in der relativ nahen Zukunft beschränkt, sondern etwa auch paläoklimatische Klimaänderungen einschließt. Die Arbeitsgruppen II (WG2) und III (WG3) sind in ihrer Ausrichtung deutlich stärker auf den anthropogenen Klimawandel orientiert. WG2 befasst sich mit den Auswirkungen des Klimawandels auf den Menschen und dessen Umwelt. Wichtige Aspekte sind hierbei die Identifikation und das Aufzeigen von Vulnerabilitäten der Gesellschaft und auch einzelner, bedeutender Ökosysteme, sowie von Möglichkeiten zur Adaption an den Klimawandel. WG3 ist ebenfalls stark handlungsorientiert ausgerichtet, mit Fokus auf der Entwicklung und dem Aufzeigen von Strategien zur Mitigation, also Abmilderung des Klimawandels und seiner Auswirkungen. Alle Bereiche sind von ausgesprochener Interdisziplinarität geprägt, ebenso wie von starker thematischer Verflechtung zwischen den Arbeitsgruppen. So stellen etwa die Berichte der WG1 eine wichtige Grundlage für die Berichte von WG2 und WG3 dar. Zusammenfassend werden die Berichte der drei Arbeitsgruppen als Sachstandsberichte, Assessment Reports (AR), bezeichnet.

Der erste Sachstandsbericht, der mittlerweile First Assessment Report (FAR) genannt wird, erschien 1990 (IPCC 1990a,b,c). In ihm finden sich die wesentlichen und grundlegenden Aussagen zum rezenten Klimawandel, wenn auch die Datengrundlage nicht mit den späteren Berichten vergleichbar war. Der FAR war eine wichtige Grundlage für die Klimarahmenkonvention der Vereinten Nationen, die auf der Konferenz von Rio 1992 beschlossen wurde. Mitte der 1990er Jahre erschien der Nachfolgebericht SAR, in dem neue Erkenntnisse aus aktueller Forschung zusammengefasst wurden (IPCC 1996a,b,c). Die Aussagen des FAR konnten hier in aller Regel bestätigt werden. Für die Klimaprojektionen standen dem SAR bereits deutlich verbesserte und aufwendigere globale Klimamodelle zur Verfügung, was sich etwa in einer höheren räumlichen Auflösung und der Berücksichtigung zusätzlicher klimarelevanter Prozesse niederschlägt.

Der TAR, der Dritte Sachstandsbericht, konnte nicht nur die Befunde seiner Vorgänger bestätigen und viele neue quantitative Abschätzungen zum anthropogenen Klimawandel liefern, sondern sich auch auf eine überlegene Datenbasis stützen (IPCC 2001a,b,c). Einen wesentlichen Anteil hieran hatte das 1995 ins Leben gerufene CMIP, das Vergleichsprojekt für gekoppelte Klimamodelle. Diese Modelle mit gekoppelten Atmosphäre- und Ozeanmodulen (AOGCMs) wurden zudem mit neuen, etwas weniger idealisierten Szenarien über die zukünftigen atmosphärischen Treibhausgaskonzentrationen angetrieben.

Im Jahr 2007 erschien der AR4, der die Befunde der vorherigen Sachstandsberichte bestätigte und konkretisierte (IPCC 2007a,b,c). Als wichtiges Ergebnis kann hier die weitere Verringerung der Unsicherheit über den Anteil des Menschen am beobachteten Klimawandel festgehalten werden. In diesem Jahr erhielt der Weltklimarat in Anerkennung seiner Verdienste bei der Bündelung des Wissens um den Klimawandel für die wissenschaftliche Fachwelt und dessen Kommunikation für die Öffentlichkeit gemeinsam mit dem ehemaligen US-Vizepräsident Gore den Friedensnobelpreis. Ein entscheidender Fortschritt im Vergleich zum TAR war die verbesserte Datengrundlage. Dies gilt vor allem für die Klimamodelle. Die AOGCMs wurden im CMIP3 zusammengefasst und die Experimente koordiniert. Diese Modelle, die auch die Datengrundlage für die vorliegende Arbeit darstellen, sind häufig qualitativ verbesserte Weiterentwicklungen bestehender Modelle, die bereits im TAR Verwendung fanden. Daneben stießen aber auch gänzlich neue Modelle zum CMIP hinzu. In Kombination mit Fortschritten in der benötigten Hardware konnte sich der AR4 so auf eine qualitativ verbesserte und quantitativ vergrößerte Datenbasis stützen. Die CMIP3-Modelle wurden auch in den Jahren nach der Veröffentlichung des AR4 als Grundlage zahlreicher und vielfältiger analytischer Studien herangezogen.

Der fünfte Sachstandsbericht AR5 wird im Jahr 2014 erscheinen. Die Projektionen und Abschätzungen dieses Berichtes stützen sich auf den, wiederum deutlich umfangreicheren, Nachfolger von CMIP3, CMIP5. Dessen Schlüsse und Erkenntnisse konnten in dieser Arbeit noch nicht berücksichtigt werden. Anhand der am 27.9.2013 akzeptierten Zusammenfassung für politische Entscheidungsträger des Berichts der WG1 lässt sich jedoch bereits erkennen, dass die in den früheren Berichten aufgezeigten Tendenzen wiederum bestätigt und weiter konkretisiert werden konnten. Desweiteren konnten die Befunde verstärkt gegen Fehlinterpretation abgesichert werden (IPCC 2013).

## 1.2 Unsicherheit, Modellgüte und Klimaprojektionen

Die Sachstandsberichte des IPCC symbolisieren den aktuellen Schwerpunkt der klimatologischen Forschung, die sicherlich im rezenten und künftigen Klimawandel ihren gemeinsamen Dreh- und Angelpunkt hat. Insbesondere zur Abschätzung der konkreten klimatischen Veränderungen bilden dreidimensionale, physikalisch basierte AOGCMs das leistungsfähigste und grundlegende Werkzeug. Anhand möglicher Entwicklungsleitlinien für die globale wirtschaftliche, technologische und bevölkerungsmäßige Entwicklung erstellen diese Projektionen für die klimatische Zukunft (NAKIĆENOVIĆ et al. 2000, MOSS et al. 2010, VAN VUUREN et al. 2011). Die zugrundeliegenden Annahmen, die Szenarien, werden hierfür in das

anthropogene Forcing, den treibhausgasbedingten Strahlungsantrieb, übersetzt. Variationen in diesen Leitlinien stellen eine wesentliche Quelle der Unsicherheit aller quantitativen Abschätzungen der zukünftigen Klimaentwicklung dar (z. B. NEW und HULME 2000, KATZ 2002, WEBSTER et al. 2002). Über die Veränderung klimabestimmender Parameter, für die hier verwendeten AOGCMs im Wesentlichen der chemischen Zusammensetzung der Atmosphäre, erhalten die Modelle neue Randbedingungen, auf die sie sich einstellen. Angesichts der Tatsache, dass AOGCMs streng deterministischen Prinzipien folgen, kann es hierbei durchaus verwundern, dass die Reaktionen verschiedener Modelle auf das gleiche Forcing auch in wichtigen Fragen divergieren können. Dieses scheinbar unsinnige Ergebnis erklärt sich dadurch, dass alle Modelle nur Approximationen der Realität darstellen und über individuelle Fehler verfügen (z. B. LAMBERT und BOER 2001, SCHMIDT 2007).

Deutlich weniger basal als diese Erkenntnis ist die Frage, was welche Fehler für die jeweilige Modellsimulation konkret bedeuten und wie auf der Grundlage solcher fehlerbehafteten Modellergebnisse wirtschafts- und gesellschaftspolitische Empfehlungen ausgesprochen werden können. Ein Ansatz hierfür ist eine möglichst große Datenbasis. Ergebnissen, die auf verschiedenen, unabhängigen Modellen beruhen wird sinnvoller- und üblicherweise stärker vertraut als solchen, die sich nur in einem einzigen Modell finden. Letztere Ergebnisse sind kaum von ungewollten Nebeneffekten, die von der spezifischen Modellphysik herrühren, zu trennen (etwa RÄISÄNEN und PALMER 2001, KHARIN und ZWIERS 2002, TEBALDI und KNUTTI 2007). Aus diesem Grund baut das IPCC, ebenso wie viele Einzelstudien die sich anhand von Modelldaten mit Aspekten des Klimawandels befassen, auf die Verwendung einer möglichst großen Anzahl von AOGCMs, die mit identischen Randbedingungen angetrieben werden. Auf diese Weise werden probabilistische Abschätzungen für das zukünftige Klima möglich. Solche systematischen Ansätze, die ihren Ursprung in Verfahren zur numerischen Wettervorhersage haben, werden als (Multi-Modell) Ensemble bezeichnet. Zu dem in dieser Arbeit verwendeten CMIP3 Ensemble siehe Kapitel 2.2.

Gerade angesichts der enormen wirtschaftlichen und sozialen Konsequenzen, die aus nicht angepassten Mitigations- oder Adaptionsmaßnahmen resultieren können, ergibt sich neben der Aufdeckung von Unsicherheiten in der Klimaentwicklung aber auch ein Bedarf an besten Schätzern für diese, um ein möglichst realistisches und vertrauenswürdigen Bild der Zukunft zu erhalten (LEMPERT und SCHLESINGER 2000). Die Kombination von möglichst vielen verschiedenen Simulationen soll hierfür, so die implizite Logik, Projektionen mit möglichst hoher Qualität erzeugen. Fragestellungen bei denen neben der Modellunsicherheit auch noch die natürliche Variabilität stark ins Gewicht fällt sollten eventuell sogar ausschließlich mit Ensembles angegangen werden (KENDON et al. 2008). Die individuellen Unzulänglichkeiten der verschiedenen AOGCMs sollen dabei durch die größere und diverse Datenbasis ausgeglichen werden. Der einfachste und gebräuchlichste Ansatz bei

der Kombination der Ergebnisse verschiedener AOGCMs ist die Bildung des arithmetischen Mittelwerts über verschiedene Klimasimulationen hinweg. Der hierbei verfolgte Ansatz wird dann häufig unter dem Schlagwort der Modelldemokratie oder „one model, one vote“ zusammengefasst (z. B. KNUTTI 2010). Tatsächlich werden, entgegen dem griffigen Wortlaut, jedoch häufiger alle  $N$  Einzelläufe  $X_i$  gleich gewichtet, wenn umfangreiche Modellensembles miteinander kombiniert werden (z. B. IPCC 2007a). Die Mittelung der Klimasimulationen erfolgt dann nach:

$$M_a = \frac{1}{N} \sum_{i=1}^N X_i \quad (1.1)$$

In diesem Fall stellt das ungewichtete Mittel der Einzelläufe  $M_a$  ein mit dem jeweiligen Ensembleumfang gewichtetes Mittel der Modelle dar. Dieses ist üblicherweise gemeint, wenn in dieser Arbeit vom arithmetischen bzw. ungewichteten Multi-Modell oder auch Ensemble Mittelwert gesprochen wird. Ein bemerkenswerter Aspekt dieser Größe ist es, bei den üblichen Überprüfungen der Modell-Performance normalerweise bessere Ergebnisse zu erzielen als die besten individuellen Modelle (LAMBERT und BOER 2001, REICHLER und KIM 2008, GLECKLER et al. 2008, PINCUS et al. 2008, PIERCE et al. 2009). Insofern scheint sie durchaus als legitimer Schätzer für das Verhalten des realen Klimasystems geeignet.

Dennoch ist es naheliegend zu erwarten, dass das arithmetische Mittel zu diesem Zweck nicht die ideale Form der Kombination von Klimasimulationen darstellt (vgl. GREENE et al. 2006). Die Performance des Multi-Modell Mittels sollte sich verbessern, wenn die individuelle Qualität der Modelle bei deren Kombination berücksichtigt wird. Symbolisch würde in diesem Fall Gleichung 1.1 durch einen gewichteten Mittelwert  $M_g$  ersetzt, wobei die Gewichtungsfaktoren  $G(i)$  die Hierarchie der Modellgüte wiedergeben.

$$M_g = \frac{1}{\sum_{i=1}^N G(i)} \sum_{i=1}^N G(i) \cdot X_i \quad (1.2)$$

Die Suche nach einer stabilen und „richtigen“ Gewichtung ist also ein aktuelles Thema, dem Potential zu weitreichenden und nachhaltigen Konsequenzen im Hinblick auf die künftige Klima- und vor allem auch Klimafolgenforschung innewohnt. Die Strategie, „guten“ Modellen stärkeren Einfluss auf die kombinierte Projektion des Ensembles zuzugestehen als „schlechten“, ist sicherlich eine sinnvolle und naheliegende Methode, die Unsicherheit dieser Projektion zu vermindern und bessere Abschätzungen für den konkret zu erwartenden Klimawandel zu erhalten (z. B. FORDHAM et al. 2011, PIERCE et al. 2009). Weder für die Theorie, auf die man sich in diesem Zusammenhang berufen sollte, noch für die praktische Umsetzung dieser Überlegung besteht allerdings ein tragfähiger Konsens (z. B. CHANDLER 2013). Für

die numerische Wettervorhersage zeigt sich die Wichtung verschiedener Modelle auf der Basis empirischer Vorhersageleistungen als dem einfachen Mitteln überlegen (KRISHNAMURTI et al. 2000, STEPHENSON et al. 2005). Auf Klimasimulationen ist dieses Vorgehen wegen der langen Zeiträume, für die die Projektionen erstellt werden, jedoch nicht übertragbar.

Als Alternative bietet sich an, die Reproduktion des beobachteten Klimas, das Hindcasting, als Grundlage von Wichtungen heranzuziehen. Jedoch ist bei diesem Vorgehen eine folgenreiche Grundproblematik zu berücksichtigen. Modell- und Beobachtungsdaten sind, schon vom Anspruch der AOGCMs her, nicht direkt vergleichbar. Die meisten transienten Klimasimulationen zum Klimawandel beginnen in der Mitte des 19. Jahrhunderts. Die Startbedingungen der Integration sind somit nicht hinreichend genau bekannt, um die tatsächliche Entwicklung des Klimasystems in der Zeitdomäne realistisch zu reproduzieren. Vergleiche zwischen Modell- und Beobachtungsdaten in dieser Domäne sind dagegen etwa möglich im Bezug auf Effekte des vorgegebenen Forcings durch beobachtete Treibhausgaskonzentrationen (vgl. BARNETT et al. 2005). So finden sich auf dekadischer Ebene zum Teil Änderungsraten die mit den beobachteten durchaus gut übereinstimmen. Ein prominentes Beispiel ist die Nachbildung der globalen Mitteltemperatur durch das CMIP3 Multi-Modell Ensemble (IPCC 2007a, PACHAURI 2008). Üblicherweise zur Attribution der Erwärmung im 20. Jahrhundert auf anthropogene Emissionen herangezogen, lässt sich auch argumentieren, dass die verwendeten Modelle eine Grundtatsache des Klimasystems in guter Übereinstimmung mit den Beobachtungsdaten abbilden. Einschränkend für diese Interpretation gilt jedoch, dass diese Reaktion auch von deutlich primitiveren Klimamodellen nachvollzogen wird (z. B. MEINSHAUSEN et al. 2011). Weiterhin wird häufig die Argumentation bemüht, dass die Nachbildung der globalen Mittelwerttemperatur eine Mindestanforderung ist. Daneben gibt es auch relativ kurzfristige Reaktionen im Klimasystem auf verschiedene Formen des externen Forcings, die entsprechend mit den Reaktionen des realen Klimasystems verglichen werden können. Ein Beispiel hierfür wäre die simulierte Reaktion auf explosive vulkanische Aktivität (z. B. HANSEN et al. 1992, MILLER et al. 2006).

Ein anderer Zusammenhang von Modell und realem Klimasystem findet sich bei Größen mit periodischem oder zyklischem Verhalten sowie bei den Zeitreihen, die diese Prozesse beschreiben. Diese sind wichtige Komponenten des realen Klimageschehens und sollen über AOGCMs realistisch abgebildet werden. Unabhängig von einem konkreten Zeitpunkt  $t$  weisen physikalische Prozesse oftmals eine zeitliche Struktur auf, die beispielsweise auf den Reaktionsgeschwindigkeiten der beteiligten Systemkomponenten beruht. Viele dominierende Größen des Klimasystems weisen stringentes zyklisches bis periodisches Verhalten auf. Üblicherweise ist dies eine Folge des abhängig vom konkreten Ort unterschiedlich ausgeprägten Jahresgangs dieser Größe. Ebenso organisiert der Tagesgang viele klimatische Phänomene. Solche Aspekte, wie der thermale Jahresgang in den Ektropen oder der hygrische in

den wechselfeuchten Subtropen, verdeutlichen, dass in der Frequenzdomäne der direkte Vergleich unterschiedlicher Formen von Beobachtungen und Modelldaten möglich ist (z. B. TAYLOR 2001, HALL und QU 2006).

Verbreiteter ist jedoch die Bewertung der Realitätsnähe von räumlichen Mustern, über die alle Phänomene im Klimasystem verfügen. Auf diesem Prinzip beruhen unter anderem Verfahren zur Detektion und Attribution des anthropogenen Klimawandels (HASSELMANN 1979, HEGERL et al. 1996, HEGERL und ZWIERS 2011). Häufig handelt es sich bei den Raummustern dann um langjährige Klimatologien wichtiger Feldgrößen  $X(t, s)$ , wie etwa der bodennahen Temperatur oder dem Niederschlag. Die simulierten Klimatologien werden mit den beobachteten abgeglichen und die Ähnlichkeit der Muster bewertet. Häufig werden dabei die Beobachtungsdaten als sicher und eindeutig angenommen, was angesichts bestehender Unterschiede zwischen verschiedenen Beobachtungsdaten kritisch betrachtet werden sollte (z. B. bei CHRISTENSEN et al. 2008, GLECKLER et al. 2008, vgl. auch Kapitel 4).

Auch ohne die Berücksichtigung solch fundamentaler Bedenken, bietet der Vergleich von Modellen und Beobachtungen keinen eindeutigen Ansatz zur Bewertung der hochkomplexen Klimamodelle. So kann die Übertragung einer Vorgehensweise auf unterschiedliche Räume zu qualitativ unterschiedlichen Ergebnissen im Hinblick auf die Aussagen zum künftigen Klimawandel und der Verlässlichkeit der einzelnen Modelle führen (z. B. GIORGI und MEARNs 2003, TEBALDI et al. 2005). Auch ist schon bei grundlegenden Variablen keine eindeutige Hierarchie der Performance eines Ensembles von AOGCMs feststellbar (REICHLER und KIM 2008, GLECKLER et al. 2008), was eine, wiederum im Wesentlichen subjektive, Abstufung der Variablen in ihrer Bedeutung erforderlich macht. Auch hierzu siehe Kapitel 4. Dieser Effekt tritt selbst bei einer starken Einschränkung des betrachteten Raums, wie der ausschließlichen Betrachtung des Australischen Kontinents, auf (SUPPIAH et al. 2007). Andererseits können bei derart regionalem Fokus durchaus mit einfachen statistischen Methoden AOGCMs identifiziert werden, deren Performance über verschiedene klimatische Kenngrößen hinweg soweit abfällt, dass ihre Berücksichtigung für die zukünftige Entwicklung im Multi-Modell Ensemble nicht sinnvoll erscheint (PERKINS et al. 2007, IRVING et al. 2012, PERKINS et al. 2012). Diese Ansätze verzichten jedoch auf die Bewertung der Modelle im Hinblick auf die übrigen ca. 98 % der Erdoberfläche.

Daneben beeinflusst die gewählte raumzeitliche Betrachtungsebene die Bewertung der Realitätsnähe von Klimamodelldaten, selbst bei Beibehaltung der Kriteriumsvariable (z. B. KJELLSTRÖM et al. 2010, REIFEN und TOUMI 2009, BLENKINSOP und FOWLER 2007, FOWLER et al. 2007). Die häufig vorkommende Nutzung von Anomalien und auch die Wahl des konkreten Zeitabschnitts für deren Berechnung beeinträchtigt die resultierenden Rangfolgen (MACADAM et al. 2010). Problematisch sind solche Uneindeutigkeiten insbesondere dann, wenn eine Wichtung tatsächlich effektiv ist, also sich die resultierenden Abschätzungen des anthropogenen

Klimawandels substantiell vom einfachen Multi-Modell Mittelwert unterscheiden und somit zumindest potentiell völlig andere Reaktionen oder Dringlichkeiten von Adaption implizieren können (IPCC 2007a, KNUTTI 2010). Ohne Zweifel können gewichtete Ensemble Mittelwerte, die die tatsächliche Unsicherheit nicht korrekt beschreiben, dem ungewichteten Multi-Modell Mittelwert als Schätzer für die realen Bedingungen deutlich unterlegen sein. Ob dies jedoch bei einem konkreten Ansatz der Fall ist, ist schwer zu entscheiden (WEIGEL et al. 2010) .

Eine naheliegende Überprüfung des Erfolgs eines Gewichtungsschemas ist der Vergleich der Güte des Hindcastings durch gewichtete und ungewichtete Ensembledmittel. So kann argumentiert werden, dass eine bessere Übereinstimmung mit den Beobachtungen den Nutzen der Gewichtung belegt. COPPOLA et al. (2010) verwenden für die Erstellung einer Gewichtung für ein Ensemble regionaler Klimasimulationen für Europa, das im Zuge des EU-Projekts ENSEMBLES (VAN DER LINDEN und MITCHELL 2009) erstellt wurde, mesoskalige Strukturen von Temperatur und Niederschlag. In diesem Fall ist das gewichtete Mittel den beobachteten Strukturen ähnlicher als das ungewichtete. Es sollte jedoch als Einschränkung angemerkt werden, dass hier, durchaus dem typischen Vorgehen entsprechend, das Gewichtungskriterium eben genau jenes ist, an dem im Nachhinein die Qualität der Gewichtung gemessen wird. Eine Überprüfung anhand anderer Klimagrößen wäre sicherlich aussagekräftiger und weniger von der Gefahr des Zirkelschlusses belegt.

Für die ENSEMBLES Modelle wurden in einem koordinierten Verfahren unterschiedliche Ansätze zur Bewertung der Modellqualität durchgeführt. Diese basieren auf dem Hindcasting unterschiedlicher Klimaelemente und resultieren in unterschiedlichen Wichtungen. Diese wurden nicht nur entwickelt, sondern die resultierenden Gewichtungen auch derart veröffentlicht, dass ein Abgleich der unterschiedlichen Ergebnisse möglich ist. Schon die Rangfolgen der Modelle, ganz abgesehen von den numerischen Beträgen der Gewichtungsfaktoren, weisen in den unterschiedlichen Ansätzen eine hohe Variabilität auf, stimmen also nicht überein. Dies unterstreicht den Einfluss der Wichtungskriterien und der -metrik ebenso wie die allgemein anerkannte Tatsache, dass die Modelle einer Generation tendenziell nicht zu große Diskrepanzen aufweisen (CHRISTENSEN et al. 2010, für AOGCMs vgl. REICHLER und KIM (2008)). Die Randbedingungen für die Regionalmodelle wurden hierbei durch die ERA 40-Reanalysen vorgegeben, um einen direkten Vergleich mit Beobachtungen zu ermöglichen (hierzu siehe Kapitel 2.1). Daher konnte hier auch die regionale Temperaturentwicklung in der Zeitdomäne mit der realen Erwärmung verglichen werden. Hierbei ergaben sich nur geringe Differenzen zwischen den Modellen in den betrachteten Regionen. Die tatsächliche Bewertung der Modellleistungen erwies sich jedoch als sensibel gegenüber der Wahl des empirischen Skalierungsparameters für die Erstellung der Gewichtungsfaktoren (LORENZ und JACOB 2010). KJELLSTRÖM et al. (2010) konnten auf Basis täglicher Temperatur- und monatlicher Niederschlagsdaten Gewichtungen erzeugen, die

insgesamt eine Verbesserung des Hindcastings für die betrachteten Klimavariablen erreichten. Dafür mussten allerdings für sämtliche Jahreszeiten separate Wichtungen entwickelt werden, so dass nur noch bedingt von einer Rangfolge der beteiligten Modelle gesprochen werden kann. Die gewichteten Ensemblemittel anderer Ansätze konnten dagegen keine signifikanten Verbesserungen im Vergleich zu den ungewichteten erzeugen (DÉQUÉ und SOMOT 2010). Ebenso wenig erfolgreich erwies sich auch der Versuch, die individuellen Gewichtungen für die Verbesserung des Hindcastings zu kombinieren (CHRISTENSEN et al. 2010).

Im Endeffekt ist auch ein Erfolg in Form eines verbesserten Hindcastings noch keine Garantie dafür, dass auch die Projektionen des Modell Ensembles realistischer werden (KLOCKE et al. 2011). Die tatsächliche Überprüfung kann hier immer nur über die Empirie erfolgen, was dem Sinn der Gewichtung, der Konkretisierung des zukünftigen Klimawandels, zuwiderläuft. Ein gewisses Substitut hierfür stellen Intermodellvergleiche dar (z. B. GIORGI und MEARNs 2002). So zeigen SMITH und CHANDLER (2010), dass die AOGCMs, die den beobachteten Niederschlag über dem Murray-Darling Becken am realistischsten reproduzieren, auch zu ähnlichen Ergebnissen im Bezug auf die künftige Niederschlagsentwicklung neigen, was als Bestätigung des Ansatzes gewertet wird. Bei globaler Betrachtung sind auf vergleichbare Art gewonnene Resultate jedoch weniger überzeugend (ABE et al. 2009) und am ehesten, wenn überhaupt, als erster Schritt bei der Erstellung einer Gewichtung zu nutzen (WHETTON et al. 2007). Schon bei der alleinigen Berücksichtigung von Australien ergeben sich zudem leicht unterschiedliche Rangfolgen der Performance der CMIP3 Modelle bei der Berücksichtigung unterschiedlicher Jahreszeiten (WATTERSON 2008), was durchaus Zweifel daran aufkommen lassen kann, inwieweit eine solche Methode dann geeignet ist, die Güte von AOGCMs zu erfassen und realistisch wiederzugeben. Es muss weiterhin berücksichtigt werden, dass sich systematische Fehler der Modelle mit dem Klima verändern könnten (CHRISTENSEN et al. 2008, BUSER et al. 2009). In die Wichtungsschemata von RÄISÄNEN et al. (2010) und RÄISÄNEN und YLHÄISI (2012) gehen sowohl der Abgleich mit beobachteten Klimatologien als auch Intermodellvergleiche ein. Im Hinblick auf den simulierten Klimawandel unterscheiden sich die so erzeugten gewichteten Ensemblemittel jedoch kaum vom ungewichteten Mittel. Daneben muss auch in Zusammenfassung der aufgeführten Ansätze angemerkt werden, dass keine Form der Gewichtung oder Bewertung tatsächlich als objektiv gelten kann. Dies wird durch die notwendigen und unvermeidlichen Entscheidungen bei der Umsetzung eines Schemas, sowohl im Hinblick auf die Auswahl der Kriterien als auch die verwendete Metrik, verhindert (z. B. Diskussion bei GLECKLER et al. 2008).

Diesen immer wieder auftauchenden Problematiken wird hier so begegnet, dass sich die Bewertung der Modellperformance nicht auf die einfache Bewertung des Hindcasting bezieht, sondern auf bekannte und für die langfristige Projektion des Klimas relevante Quellen der Unsicherheit, soweit diese mit den Modellen

selbst verbunden sind (vgl. etwa VISSER et al. 2000, GIORGI und FRANCISCO 2000, COVEY et al. 2003, HO et al. 2012). Somit wird ein Faktor isoliert, dessen Auswirkungen eventuell nicht von allen Modellen im gleichen Maße realistisch abgebildet wird, was die Zuverlässigkeit der Projektionen eines Modellensembles herabsetzt. Die Unsicherheit, die so reduziert werden soll, betrifft jedoch nicht nur den eigentlichen Faktor sondern ebenso andere, mit diesem gekoppelte Aspekte des Klimasystems. Als konkretes Kriterium wird ihre Fähigkeit herangezogen, die großskalige natürliche Klimavariabilität realistisch zu repräsentieren. Zu deren Bedeutung siehe Kapitel 1.3. Im Hinblick auf den anthropogenen Klimawandel erscheint diese Wahl als sinnvoll, da die Klimavariabilität sowohl von diesem betroffen ist, als auch dessen Ausprägung, zumindest auf dekadischer Ebene, modulieren kann.

Die Konzentration auf einzelne, wichtige Aspekte des Klimasystems, um Unterschiede zwischen hochkomplexen Modellen zu identifizieren und herauszuarbeiten, ist ein durchaus vielversprechender Ansatz, insbesondere bei den CMIP Modellen, denen allgemein eine relativ hohe Modellgüte unterstellt werden kann (MEEHL et al. 2007, RÄISÄNEN 2007, REICHLER und KIM 2008). Dennoch werden solche erfassbaren Modellunterschiede zur Erstellung von Wichtungen nur relativ selten herangezogen. So zeigen HALL und QU (2006) deutliche Differenzen innerhalb des CMIP3 auf Basis des Vergleichs der simulierten mit der beobachteten Schnee-Albedo Rückkopplung, allerdings ohne hieraus eine Rangfolge oder gar Gewichtung abzuleiten. Diese Größe beeinflusst die Gesamtreaktion des Klimasystems auf das anthropogene Forcing. OLDENBORGH et al. (2005) zeigen unterschiedliche Performance im CMIP3 im Hinblick auf das ENSO-System auf. Der Abgleich mit dem beobachteten Verhalten wurde, ebenfalls für einen Teil des CMIP3, auch dazu genutzt, den für das 21. Jahrhundert zu erwartenden Rückgang des arktischen Meereises zu konkretisieren (BOÉ et al. 2009). Ein anderer Aspekt, der potentiell die Projektionen von langfristigen Klimaänderungen für das Multi-Modell Ensemble beeinflussen kann, ist die Beziehung der Amplitude des Jahresgangs und der Klimasensitivität. Anhand einiger CMIP3 Modelle und eines großen Ensembles weniger aufwendiger Klimamodelle konnte der Realismus des simulierten Jahresgangs bewertet werden und anhand diesem die Wahrscheinlichkeit für die tatsächliche Gleichgewichtsreaktion des Klimasystems auf eine Verdopplung des atmosphärischen CO<sub>2</sub>-Gehalts neuartig abgeschätzt werden (KNUTTI et al. 2006). Auch die Reaktion der globalen Oberflächentemperatur auf das Treibhausgasforcing im 20. Jahrhundert eignet sich für gewichtete Abschätzungen dieser Größe (PIANI et al. 2005, LOPEZ et al. 2006).

FOWLER und EKSTRÖM (2009) nutzen die Simulation von Extremereignissen im Niederschlag als Kriterium für die Erstellung einer Gewichtung für Regionalmodelle. In diesem Fall liegt deutlich ein Fokus auf einer sehr spezialisierten Fragestellung, die allerdings aufgrund ihrer möglicherweise dramatischen Auswirkungen durchaus als Unsicherheitsfaktor von Klimasimulationen betrachtet werden kann. An

diesem Beispiel zeigt sich besonders anschaulich die allgemeine Motivation für die Erstellung von Wichtungen: die Erzeugung einer realistischeren Projektion zur Abschätzung der Zukunft, auch wenn sie sich in diesem Fall kaum von der ungewichteten unterscheidet. LENDERINK (2010) fand in einem vergleichbaren Ansatz jedoch deutliche Unterschiede zwischen den Leistungen der einzelnen Regionalmodelle, weist aber auch auf den Einfluss der Qualität der Beobachtungsdaten hin. Das sich selbst bei deutlich abstufenden Gewichtungsfaktoren nur geringe Unterschiede zwischen den gewichteten und ungewichteten Mittelwerten eines Ensembles ergeben können findet sich in der Untersuchung von SCHMITTNER et al. (2005). Die Autoren erstellen eine Gewichtung für eine, relativ kleine, Auswahl der hier verwendeten Modelle, mit dem Ziel zuverlässigere Projektionen für die thermohaline Zirkulation im Nordatlantik zu erhalten. Das Hindcasting konzentriert sich hierbei entsprechend auf die ozeanische Komponente der AOGCMs, mit Fokus auf dem Nordatlantik. Dieses stark spezialisierte Wichtungsschema führt zu einer Reduktion der Standardabweichung der simulierten Veränderung der thermohalinen Zirkulation im Multi-Modell Ensemble, während sich der gewichtete und ungewichtete Mittelwert kaum unterscheiden. Somit wird über die Wichtung zwar die mittlere zu erwartende Klimaänderung nicht verändert, aber das Vertrauen in die Zuverlässigkeit dieser Projektion erhöht.

Die bereits erwähnte Unsicherheit, die aus den Annahmen über die Randbedingungen resultiert, ist in derartige Abschätzungen nicht miteinbezogen (etwa WIGLEY und RAPER 2001). Ebenso bleibt der, potentiell dramatische, kombinierte Einfluss von eventuell unrealistisch justierten Modellparametern unberücksichtigt. Derartige Abschätzungen sind mit einem Multi-Modell Ensemble von AOGCMs nicht realisierbar und erfordern ein sog. Perturbed-Physics Ensemble (hierzu MURPHY et al. 2004, 2007).

## 1.3 Eigenmoden und Telekonnektionen

Die interne Klimavariabilität lässt sich zu weiten Teilen in typischen räumlichen Mustern erfassen und ordnen. Über die Kombination einiger immer wiederkehrender präferierter Muster lässt sie sich daher effektiv beschreiben. Die wichtigsten und bedeutsamsten dieser Zirkulationsmuster können als einflussreiche Klimagrößen aufgefasst werden. Übliche Bezeichnungen für diese raumzeitlichen Phänomene sind etwa Klimamoden, Eigenmoden des Klimasystems oder Klimamuster (z. B. STEPHENSON et al. 2004). Die Terminologie ist allerdings nicht kodifiziert, so dass die betreffenden Größen auch mit anderen Bezeichnungen tituliert werden. So der Kontext keine Fehlinterpretation zulässt, werden auch in dieser Arbeit weitere Begriffe verwendet.

Klimamoden äußern sich in physikalischen Feldgrößen, den Klimaelementen, über Telekonnektionen, synchronen Klimaschwankungen an geographisch weit entfernten

Orten (nach ÅNGSTRÖM 1935). Anhand von räumlich aufgelösten Daten lassen sich Telekonnektionen am einfachsten über die Korrelationsrechnung erfassen. Die herausgehobene Stellung der statistischen Methodik in diesem Teilgebiet der Klimatologie geht bereits auf frühe, wegweisende Arbeiten, etwa von WALKER (1924) oder WALKER und BLISS (1932), zurück. Um die zeitliche Variation von Klimamoden zu beschreiben, werden für die einzelnen Phänomene auf unterschiedliche, empirische Weise Indexreihen definiert. In der Folge bedeutet das, dass die Charakterisierung eines räumlichen Phänomens, eines typischen Zustands der Atmosphäre und die damit verbundenen Variationen in unterschiedlichsten Klimaelementen, anhand einer einzigen Zeitreihe möglich ist. Dadurch verfügen die Indices wichtiger Klimamoden über ein enormes Potential bei der Erforschung von natürlichen und quasi-natürlichen, klimatologisch beeinflussten Systemen (vgl. HALLETT et al. 2004). Somit stellt die zukünftige Entwicklung dieser Indexreihen einen wichtigen Aspekt bei allen Abschätzungen der Folgen des künftigen Klimawandels dar. Auf ihre Darstellung und Interpretation soll daher hier kurz eingegangen werden.

Üblicherweise werden für Klimamoden zwei gegenläufige Phasen unterstellt. Dieses Vorgehen entspricht zunächst einer linearen Betrachtungsweise, die sicherlich durch die häufig verwendeten Analysemethoden – EOF-Analyse, Korrelations- und Regressionsrechnung – begünstigt wird. Die Ergebnisse aus alternativen, nicht-linearen Ansätzen bestätigen dieses prinzipielle Muster jedoch (z. B. CAS-SOU und TERRAY 2001). In der Indexreihe wird dieser qualitative Unterschied im Zustand der Atmosphäre mit dem Vorzeichen gekennzeichnet. Die Intensität des Zustands, also die Frage, wie deutlich sich das charakteristische Muster in den klimatischen Feldgrößen widerspiegelt, ist über den Zahlenwert erfassbar (vgl. WANNER et al. 2001). Dieser einfachen und scheinbar eindeutigen Unterscheidung zum Trotz muss aber betont werden, dass sowohl die Definition der Raummuster als auch der Indexreihen rein empirisch erfolgt. Daher sind auch die gegensätzlichen Phasen immer nur als Anomalien zu einer auf Basis verfügbaren Datenmaterials festgelegten Normalperiode aufzufassen. Je nach der konkreten Definition verfügen einige Klimaindices über physikalische Einheiten. Üblicher, und auch hier praktiziert, ist es dagegen, Zeitreihen aus standardisierten Anomalien zu verwenden. Vorteile solcher Zeitreihen sind der einfache und direkte Vergleich der transienten zeitlichen Entwicklung unterschiedlicher Klimamoden und insbesondere auch die vereinfachte Nutzung unterschiedlicher Beobachtungsdaten und Klimasimulationen. Zudem bedeutet ein standardisierter Indexwert neben der Kennzeichnung des Zustandes des Zirkulationsmusters auch eine Aussage über die Wahrscheinlichkeit, oder genauer relative Häufigkeit, der konkreten Ausprägung.

Die Kombination einer geringen Zahl solcher Raummuster und Zeitreihen erlaubt es dann, weite Teile der klimatischen Dynamik zu erfassen. Hierbei ist ein hervorzuhebender Aspekt, dass der so erklärbare Anteil der Variation im Klimasystem umso größer wird, je längere Zeiträume betrachtet werden. Von der saisonalen Variabilität lassen sich dementsprechend größere relative Anteile auf Zirkulations-

muster zurückführen als bei der Betrachtung von Monaten oder Wochen. Auf der Skala von Dekaden sind diese relativen Beiträge der Eigenmoden derart ausgeprägt, dass sie eine vergleichbare Größenordnung wie die anthropogene Beeinflussung des Klimas haben (HURRELL 1995). Somit bieten die Zirkulationsmuster eine elegante Möglichkeit, die Entwicklungen des Klimasystems auf dekadischen oder klimatologischen Zeitskalen zu erfassen und zu beschreiben. Auch das begründet ihre wichtige Stellung in der Klimaforschung (QUADRELLI und WALLACE 2004). Neben diesen langfristigen klimatologischen Aspekten besitzen Zirkulationsmoden allerdings durchaus auch auf der synoptischen Skala Relevanz und sind, über die durch sie vorgegebenen Randbedingungen, beispielsweise wichtige Faktoren für die Häufigkeit und Stärke von Sturmereignissen oder dem Auftreten außertropischer Zyklonen (vgl. MAILIER et al. 2006, SEIERSTAD et al. 2007). Selbst starke Schwankungen in regionalen Witterungsverhältnissen können erfolgreich unter Bezug auf solche makroskaligen Zirkulationsmuster erklärt werden (z. B. CATTIAUX et al. 2010). Als Einschränkung sei bei diesem Punkt jedoch angemerkt, dass die Telekonnektionen einer Klimamode nicht zeitlich konstant ausgeprägt sind, was die Abschätzung der klimatischen Auswirkungen von Indexausprägungen für konkrete Orte erschwert (z. B. GERSHUNOV und BARNETT 1998, POWER et al. 1999, LU und GREATBATCH 2002, VINTHER et al. 2003). Gleiches gilt für die Übertragung von Modellergebnissen auf das reale Klimasystem. Nichtsdestotrotz ermöglichen die Fernwirkungen und die durch sie hervorgerufene Klimazeugnisse die Rekonstruktion einiger Klimamuster für historische Perioden und sogar für weite Teile des Holozäns. Zumindest für das gegenwärtige Interglazial ist ihre klimatologische Relevanz belegt und ihre regionalen Effekte sind, wenn auch nicht quantitativ, so doch wenigstens qualitativ, als weitgehend stabil zu bezeichnen (z. B. BRIFFA 2000, APPENZELLER et al. 1998, FRAPPIER et al. 2002, COBB et al. 2003, WOODROFFE et al. 2003, LUTERBACHER et al. 1999, 2001).

Die prominentesten Beispiele dieser großskaligen Klimaphänomene sind das globale Monsunsystem, von dem in dieser Arbeit zwei regionale Ausprägungen berücksichtigt werden, sowie das ENSO-Phänomen (El Niño-Southern Oscillation). Beide stellen Moden der niederen Breiten dar. Ihre tatsächliche Bedeutung ist jedoch eindeutig globaler Natur (z. B. IPCC 2007a,b). Die Monsune sind die, aus klimatologischer, aber insbesondere auch anthropozentrischer, Perspektive, wichtigsten Eigenmoden. Sie entstehen letztlich als Folge der jahreszeitlichen Änderung der Lage der Innertropischen Konvergenzzone. Die Landwirtschaft in den wechselfeuchten Tropen ist in weiten Teilen abhängig von der mit ihnen verbundenen Regenzeit. In den klassischen Monsunregionen der Erde leben insgesamt etwa 2 Milliarden Menschen, so dass weite Teile der klimatologischen Literatur zu den Monsunen auch sozioökonomische Aspekte thematisieren, und umgekehrt. Als allgemeine Referenz für die Monsune können z. B. WEBSTER und YANG (1992), SUPPIAH (1992), WEBSTER et al. (1998), SULTAN et al. (2003), WANG (2006) und insbesondere CHANG et al. (2011) herangezogen werden.

ENSO wiederum ist die dominierende Mode der interannuellen Klimavariabilität und lässt sich in vielen klimatischen Feldgrößen, insbesondere im Fall eines El Niño- oder La Niña-Ereignisses, global detektieren (z. B. YUAN 2004, ROPELEWSKI und HALPERT 1986, 1987, MCBRIDE und NICHOLLS 1983, GRIMM et al. 2000, BRÖNNIMANN 2007, WANG et al. 2000, ZHANG et al. 1999, WANG et al. 2003, CAMBERLIN et al. 2001, POWER et al. 1999). Die Vorhersage solcher folgenreicher Ereignisse ist ein entsprechend wichtiger Aspekt in der Erforschung von ENSO (z. B. LATIF et al. 1998, IPCC 2007a). Daneben besteht eine Koppelung zwischen der globalen Temperatur und dem ENSO-System, die sich in modernen Datensätzen deutlich identifizieren lässt und eine Besonderheit unter den Eigenmoden darstellt (zu diversen Aspekten von ENSO vgl. PHILANDER 1990, DIAZ und MARKGRAF 1992, 2000).

Die Klimamoden der Außertropen sind im Vergleich weniger bekannt und werden auch mit weniger umfassenden Auswirkungen assoziiert. Am prominentesten ist hier sicherlich die Nordatlantische Oszillation (NAO), die auch als Northern Annular Mode (NAM) bezeichnet wird. Die Mechanismen und Auswirkungen der NAO wurden und werden intensiv erforscht und beschrieben (z. B. THOMPSON und WALLACE 2001). Ein wesentlicher Grund hierfür ist, dass für Europa, dessen Klimavariabilität die NAO entscheidend prägt (z. B. TRIGO et al. 2002), viele und langfristige Messungen von meteorologischen Variablen vorliegen. Eine ausführliche Diskussion dieses Zirkulationsmusters und seiner Auswirkungen gibt HURRELL et al. (2003). Andere außertropische Muster erhielten und erhalten dagegen weniger Beachtung, was unterschiedliche Ursachen hat.

Vor allem besteht trotz der langen Forschungstradition über einige fundamentale Fragen im Umfeld der Klimamoden weiterhin Unsicherheit. Schon die Anzahl der durch verschiedene Autoren mit unterschiedlicher Methodik und Datenbasis identifizierten Zirkulationsmuster weist deutliche Variationen auf (z. B. WALKER und BLISS 1932, WALLACE und GUTZLER 1981, HOREL 1981, ESBENSEN 1984, BARNSTON und LIVEZEY 1987, KUSHNIR und WALLACE 1989, PAVAN et al. 2000, TRENBERTH et al. 2005). Die meisten dieser Studien mussten sich aus Gründen der Datenverfügbarkeit auf die Nordhemisphäre, und hier die außertropischen Gebiete, beschränken. Insbesondere hier divergiert dementsprechend die Zahl der benannten Zirkulationsmuster. Für die Außertropen der Südhemisphäre hat dagegen nur ein dominantes Muster breitere Beachtung gefunden, der Southern Annular Mode (SAM), der als Gegenstück des NAM betrachtet wird (etwa GONG und WANG 1999, THOMPSON und WALLACE 2000, THOMPSON et al. 2000). Die Unübersichtlichkeit erhöht sich zusätzlich dadurch, dass nicht in allen Fällen Einigkeit in der Nomenklatur herrscht. Für einige Muster existieren mehrfache Benennungen (z. B. Diskussion bei MANTUA und HARE 2002), ebenso werden einige Bezeichnungen für verschiedene Muster herangezogen. Die Anzahl der Moden mag zudem von vielen der genannten Autoren schlicht überschätzt worden sein. Diese Gefahr ergibt etwa sich aus der zu unkritischen Interpretation von durch

statistische Verfahren gewonnenen Mustern, in deren Folge numerische Artefakte als physikalische Phänomene gedeutet werden. Desweiteren können verschiedene traditionelle Indices, die auf zu kurzer oder nicht flächendeckend vorliegender Datenbasis erstellt wurden, einfach regionale Aspekte einer übergeordneten Mode darstellen (z. B. KERR 2004). Dieser Problematik wird hier dadurch begegnet, dass in Kapitel 5.1 diejenigen Zirkulationsmoden detektiert werden, die sich auf globaler Ebene als relevant für die Organisation der atmosphärischen Variabilität erweisen. Für diese werden dann umfangreiche Untersuchungen zu ihrer Repräsentation durch die Klimamodelle sowie zudem durch das CMIP3-Ensemble projizierten Veränderungen im Verhalten der Indexreihen durchgeführt.

Neben der individuellen, meist regionalen Bedeutung einzelner Zirkulationsmuster ist auch deren Interaktion, zumindest potentiell, folgenreich und von wissenschaftlichem Interesse. Synchrone Entwicklungen im Verhalten mehrerer Klimaindices können eventuell Umschwünge des gesamten Klimasystems einleiten, was globale Auswirkungen impliziert. Hiervon ist zum einen die globale Mitteltemperatur betroffen, daneben werden mit solchen Umschwüngen auch Veränderungen in der Interaktion zwischen unterschiedlichen Eigenmoden assoziiert. Als wichtiger Aspekt gilt zudem die Veränderung der Variabilität von ENSO sowie der Intensität der verbundenen Telekonnektionen (TSONIS et al. 2007, WANG et al. 2009). Diese Perspektive verdeutlicht die Bedeutung der Klimamoden und legt auch nahe, dass sich Betrachtungen für deren zukünftige Entwicklung nicht auf zu wenige Zirkulationsmuster konzentrieren sollten.

Außer der Repräsentation dieser Zirkulationsmuster durch die AOGCMs wird in dieser Arbeit auch deren Reaktion auf das anthropogene Forcing untersucht. Die möglichen Veränderungen im statistischen Verhalten in Folge des rezenten und künftigen Klimawandels sind ein Gegenstand intensiver Forschung (z. B. MEEHL und WASHINGTON 1996, TRENBERTH und HOAR 1997, TIMMERMANN et al. 1999, PAETH et al. 1999, HARTMANN et al. 2000, HU et al. 2000, HOERLING et al. 2004, COLLINS 2005, RIND et al. 2005, HANSEN et al. 2006, GILLETT und THOMPSON 2003, TURNER et al. 2005, IPCC 2007a, PAETH et al. 2008, COLLINS et al. 2010, PAETH und POLLINGER 2010, TURNER und ANNAMALAI 2012, u. v. a.). Als Phänomene der Dynamik ist für die Klimamoden eine Veränderung in Folge der globalen Erwärmung mit einer allgemeinen Intensivierung der Zirkulation zunächst zu erwarten. Jedoch ist die Detektion des anthropogenen Signals sowohl in Beobachtungsdaten als auch im Output von Zirkulationsmodellen wegen des stochastischen Verhaltens der Moden deutlich aufwendiger und unsicherer als etwa bei der globalen Oberflächentemperatur.

Durch Veränderungen im atmosphärischen Transport von Wärme und Feuchte, die sich in der Veränderung der Häufigkeiten verschiedener Ausprägungen einer Mode äußern, wird das globale, anthropogene Signal regional modifiziert. Infolgedessen kann die regionale Erwärmung im globalen Vergleich intensiver ausfallen, aber auch eine Reduktion des Erwärmungstrends ist dadurch möglich. Typischerweise

unterscheiden sich diese Effekte in verschiedenen Regionen und gleichen sich im globalen Maßstab nahezu aus. Weiterhin wird das typische Niederschlagsverhalten beeinflusst, was abhängig vom Standort deutlich schwerwiegendere Folgen als eine Temperaturveränderung haben kann. Verbunden mit dem Einfluss auf solche grundlegenden klimatischen Variablen, stellen Zirkulationsmuster zudem relevante Faktoren für den Wandel in der traditionellen Tier- und Pflanzenwelt, dem hydrologischen Haushalt, der gesamten Raumausstattung betroffener Regionen dar. Auch daher findet das Konzept der Klimamoden, und besonders die Nutzung von Indexzeitreihen, breite Anwendung etwa in der ökologischen Forschung (hierzu z. B. STENSETH et al. 2002, 2003, HALLETT et al. 2004, FORCHHAMMER und POST 2004). Ebenso besteht von Seiten der Agronomie und Landwirtschaft großes Interesse, besonders natürlich im Hinblick auf zuverlässige und frühzeitige Prognosen, an diesen Maßzahlen der Klimavariabilität (z. B. MJELDE und GRIFFITHS 1998). Konkret fassbar ist, als ein weiteres Beispiel, der Zusammenhang von Klimamoden und der Stromerzeugung aus Wasserkraft (z. B. CHERRY et al. 2005).

Wie sich hier andeutet, stellen die effektiven Auswirkungen von Zirkulationsmoden, in Vergangenheit, Gegenwart und Zukunft, ein noch deutlich weiteres Feld dar als die Erforschung der Moden und ihres Verhaltens selbst. Dementsprechend stellen sie auch eine wesentliche Motivation für die Beschäftigung mit den Eigenmoden in dieser Arbeit dar. Allerdings werden solche, häufig interdisziplinären, Aspekte hier nicht aufgegriffen. Für einen Überblick wird daher lediglich auf IPCC (2007b) und die dort aufgeführte Literatur verwiesen.

## 2 Daten

In diesem Kapitel wird die Datengrundlage für die durchgeführten Untersuchungen vorgestellt. Es wurde darauf geachtet, die jeweils aktuellen Versionen der Datensätze zu verwenden. Jedoch werden einige der Beobachtungsdaten so häufig aktualisiert – im Wesentlichen verlängert – dass davon ausgegangen werden kann, dass es numerische Abweichungen zwischen den in dieser Arbeit verwendeten Daten und den aktuell verfügbaren Versionen der betreffenden Datensätze gibt. Insbesondere basieren einige Zeitreihen auf den Ergebnissen von EOF-Analysen oder werden über ihre eigene Standardabweichung normiert. In solchen Fällen beeinflusst eine einfache Verlängerung der Beobachtungsperiode sämtliche Einzelwerte. Dieses Vorgehen sollte jedoch, angesichts der hier durchgeführten Analysen, die eher auf das große Ganze als auf Nuancen zielen, keine Auswirkungen haben.

Im Fall der AOGCM-Daten, dem eigentlichen Ziel der Untersuchungen, entfällt diese Unsicherheitsquelle. Sämtliche Dateien wurden auf eventuelle Unstimmigkeiten oder Datenfehler überprüft, um Veränderungen nach dem Herunterladen auszuschließen. Fehlerhafte Datensätze wurden nicht in die Untersuchungen miteinbezogen, so dass die Qualität der Daten als weitgehend gesichert betrachtet werden kann.

### 2.1 Beobachtungsdaten

An dieser Stelle werden alle Daten beschrieben, die Aspekte des realen Klimasystems abbilden. Der Begriff „Beobachtung“ wird relativ weit gefasst, da alle Datensätze das Ergebnis aufwendigen Post-Processings sind.

Allerdings sind sämtliche Datensätze in der wissenschaftlichen Literatur beschrieben und bereits in vielfältiger Weise verwendet worden. Eine stringente Anwendung von „Beobachtung“ würde weite Teile der Arbeit unmöglich machen. Der Charakter dieser Daten wird jedoch in Kapitel 4 beim Abgleich mit den Modelldaten berücksichtigt.

#### 2.1.1 Felddaten

Felddaten sind Flächendaten und liegen in dieser Arbeit als gridded data, also auf einem regelmäßigen Gitternetz verortete Daten vor. Als raumzeitliche Variablen werden sie mit  $X(t, s)$  bezeichnet. Der Index  $t$  repräsentiert die Zeit, während  $s$

**Tab. 2.1:** Beobachtungsdaten (Felddaten). Res: ungefähre räumliche Auflösung ( $^{\circ}$ Latitude x  $^{\circ}$ Longitude), T: 2 m-Temperatur, R: Niederschlag, P: Luftdruck auf Meeresniveau, Z<sub>5</sub>: Höhe des 500 hPa Niveaus in gpm, Z<sub>10</sub> wie Z<sub>5</sub> aber 1000 hPa Niveau

Name	Res	Var	Zeit	Referenz
ERA 40	2.5 x 2.5	T,R,P,Z <sub>5</sub> ,Z <sub>10</sub>	09/1957–08/2002	UPPALA et al. (2005)
NCEP/NCAR	1.9 x 1.9	T,R	01/1948–12/2010	KALNAY et al. (1996)
	2.5 x 2.5	P,Z <sub>5</sub> ,Z <sub>10</sub>		
CRU	0.5 x 0.5	T,R	01/1901–12/2002	MITCHELL und JONES (2005)
E-OBS	0.5 x 0.5	T,R	01/1950–12/2006	HAYLOCK et al. (2008)
HadSLP2	5 x 5	P	01/1850–12/2004	ALLAN und ANSELL (2006)
GPCP	2.5 x 2.5	R	01/1979–12/2007	ADLER et al. (2003)
GPCC	0.5 x 0.5	R	01/1901–12/2007	SCHNEIDER et al. (2008)
TRMM	0.25 x 0.25	R	01/1998–12/2007	HUFFMAN et al. (2007)

den Ort kennzeichnet. In dieser Arbeit wird um eine einfachere Notation zu ermöglichen, mit  $s$  ein konkreter, zumindest theoretisch lokalisierbarer Standort symbolisiert. Der Raum wird also nicht explizit als Vektorgroße dargestellt. Damit implizieren sie eine so nicht vorhandene räumliche und zeitliche Abdeckung durch Messdaten. Tatsächlich beruhen sie jedoch weitgehend auf Punktmessungen oder zeitlich nicht konstanten, Satellitenmessungen. Kontinuierliche Datensätze entstehen aus den lückenhaften Rohdaten grundsätzlich über räumliche und zeitliche Interpolationen. Dabei kommen zwei unterschiedliche Ansätze zur Anwendung. Einer ist die dynamische Interpolation mithilfe von physikalischen Zirkulationsmodellen für die Atmosphäre. Diese Atmosphärenmodelle entsprechen denen für die numerische Wettervorhersage und sind somit auch mit den AOGCMs verwandt. Diese Methode wird als Reanalyse bezeichnet. Der zweite Ansatz ist die, wie auch immer konkret durchgeführte, statistische Interpolation der Messungen. Tabelle 2.1 enthält alle Felddaten, die in dieser Arbeit verwendet wurden. Ergänzend erfolgen einige Anmerkungen zu den einzelnen Datensätzen. Ausführlich werden diese in den angebenen Referenzen (Tab. 2.1) vorgestellt. Da hier eine große Anzahl von Beobachtungsdatensätzen gemeinsam genutzt wird und keiner als alleiniger Repräsentant des tatsächlichen realen Klimasystems oder auch nur der Entwicklung einer seiner Größen betrachtet wird, wird auf eine umfassende Diskussion der individuellen Datenqualität verzichtet.

Die größte Vielfalt an Variablen bieten, entsprechend ihrer Generierung über Zirkulationsmodelle, die beiden globalen Reanalysedatensätze. Sie stellen eine unkomplizierte Möglichkeit dar, vielfältige Beobachtungen, die in sich eine physikalisch konsequente Struktur aufweisen zu nutzen. Hier werden sowohl die amerikanischen Reanalysen, ein Gemeinschaftsprodukt der National Centers for Environmental

Prediction (NCEP) und dem National Center for Atmospheric Research (NCAR), als auch das europäische Produkt ERA 40 vom European Centre for Medium-Range Weather Forecasts (ECMWF) eingesetzt.

Reanalysen assimilieren eine sehr große Anzahl von Beobachtungsdaten. Das Prinzip welches hierbei verfolgt wird, ist dass möglichst viele Messungen von den Modellen genutzt werden. Es werden jeweils Stationsdaten, Schiffsmessungen, die Ergebnisse von Radiosondenaufstiegen sowie die Messungen von Flugzeugen und, soweit diese verfügbar sind, auch Satellitendaten bei der Erstellung der Reanalysen berücksichtigt. Dieses Prinzip soll die bestmögliche Erfassung des Zustandes der Atmosphäre ermöglichen. Durch die AGCM-basierte Assimilierung werden diese Beobachtungen zu physikalisch konsistenten Datensätzen. Eine Folge hieraus ist allerdings eine zeitliche Inhomogenität, da die Anzahl der Messungen während der Reanalyseperiode schwankt und insbesondere ab 1979 mit dem Aufkommen operativer Wettersatelliten, stark ansteigt. Daneben liegen für einige Weltregionen, gerade vor 1979, nur in sehr geringem Umfang Messungen vor. Hieraus folgt auch eine räumliche Inhomogenität im Hinblick auf die Qualität und Realitätsnähe beider Reanalysen. Diese irreversiblen Einschränkungen der Zuverlässigkeit gelten jedoch natürlich im vollen Umfang für alle hier verwendeten Beobachtungsdatensätze. Sowohl NCEP/NCAR als auch das ECMWF haben mittlerweile neue Reanalysen erstellt (SIMMONS et al. 2007, SAHA et al. 2010). Diese sind generell besserer Qualität als ihre Vorgänger, klammern die Prä-Satelliten Ära jedoch weitgehend aus, so dass sie für die hier vorgenommenen Untersuchungen weniger geeignet erschienen. Das Gleiche gilt für die japanischen Reanalysen (ONOGI et al. 2007).

Wegen der Assimilierung sind hier jedoch sämtliche atmosphärischen Größen auch vom verwendeten Modell beeinflusst. Dadurch können zwischen tatsächlichen Messungen und den hier genutzten Daten zum Teil erhebliche Diskrepanzen bestehen (z. B. INOUE und MATSUMOTO 2004). Solche Differenzen können, wie etwa beim außertropischen, südhemisphärischen SLP in den NCEP/NCAR Reanalysen derartig persistent und großflächig auftreten, dass die alleinige Verwendung dieser Reanalysen in manchen Fragestellungen zu Verfälschung der tatsächlichen Entwicklungen im Klimasystem führt (MARSHALL 2003, GILLET et al. 2006). Die ERA 40 Daten sind zudem durch die kürzere abgedeckte Periode im Vergleich zu den anderen Daten, insbesondere den qualitativ besseren Daten in den Jahren nach der weitgehend flächendeckenden Erfassung des Klimasystems durch Beobachtungssatelliten, etwas stärker in ihrer Verwendbarkeit eingeschränkt (etwa SIMMONS et al. 2004). Da aber die AOGCMs ohnehin nur bis etwa zur Jahrtausendwende mit beobachteten Treibhausgaskonzentrationen angetrieben werden, stellt dies hier keinen echten Hindernisgrund dar. Da den ERA 40 Daten allgemein eine durchaus hohe Qualität unterstellt wird, nehmen sie eine sehr wichtige Rolle beim Vergleich der Reproduktion der beobachteten Klimatologien durch die AOGCMs ein (vgl. Kapitel 4). Insgesamt kann sowohl für ERA 40 als auch den Datensatz von NCEP/NCAR festgehalten werden, dass ihre Verwendung in manchen Aspekten

auch dieser Arbeit durchaus als alternativlos bezeichnet werden kann. Die bekannten Schwachpunkte dieser Daten sollten durch die Kombination verschiedener Beobachtungen gemildert werden.

Mit Ausnahme der geopotentiellen Höhe werden für alle hier verwendeten Feldgrößen daneben allerdings auch andere Datensätze verwendet, die einen stärkeren Beobachtungscharakter aufweisen.

Ein weiteres verbreitetes Klimadatenprodukt welches Verwendung findet sind die hochauflösenden Gitterdaten der Climate Research Unit (CRU) der Universität von East Anglia. Diese basieren auf Messungen der bodennahen Temperatur und des Niederschlags, aus denen über statistische Interpolationsverfahren Gitternetzdaten erzeugt werden. Die CRU Daten decken sämtliche Landmassen, inklusive einer großen Anzahl an Inseln und mit Ausnahme der Antarktis, ab. Natürlich ist auch in diesem Datensatz die tatsächliche Beeinflussung der angegebenen Werte durch wirkliche Messungen von einer starken raumzeitlichen Variation betroffen.

Für die europäischen Landgebiete steht mit den E-OBS Daten eine weitere Quelle für Beobachtungsdaten zur Verfügung. Sie basieren auf den Messungen von etwa 2000 Stationen, die auf verschiedene Gitternetze interpoliert wurden (HOFSTRA et al. 2010). Die Anzahl der Stationen schwankt hierbei. Aus den täglichen Daten wurden nur dann Monatsmittel gebildet, wenn für den Gitterpunkt im betreffenden Monat Werte für mindestens 2/3 der Tage vorlagen. Diese monatlichen Daten bilden die Basis für die weitere Verwendung von E-OBS. Ursprünglich wurden diese Beobachtungen generiert, um sie als Grundlage für die Validation von Regionalen Klimamodellen zu dienen. Sie sollen ausdrücklich für den direkten Abgleich nutzbar sein, wobei in dieser Arbeit, wie bei allen anderen Beobachtungsdatensätzen, nur klimatologische Werte zu diesem Zweck herangezogen werden. Daher sollten auch die dokumentierten Inhomogenitäten, die sowohl weite Teile der Temperatur- als auch der Niederschlagszeitreihen aufweisen keine tatsächliche Einschränkung hierfür darstellen (BEGERT et al. 2008, HOFSTRA et al. 2009).

Einige Datensätze liegen nur für eine einzige Feldgröße vor. Als Ergänzung für die Größen aus den Reanalysen wird für den auf Meeresniveau reduzierten Luftdruck (sea level pressure, SLP) der Datensatz HadSLP2 herangezogen. Dieser wird vom britischen Hadley Centre zur Verfügung gestellt. In ihm werden die Beobachtungen von 2228 Stationen an Land mit marinen Messungen zu einem global flächendeckenden Datensatz kombiniert. Der HadSLP2 wird später in weiten Teilen als Referenz für die großskaligen Zirkulationsmuster sowohl in den Beobachtungen als auch in den AOGCMs genutzt.

Zusätzlich werden Niederschlagsdaten aus weiteren Quellen herangezogen. Der Deutsche Wetterdienst DWD betreibt in Abstimmung mit der World Meteorological Organisation das Global Precipitation Climatology Centre (GPCC), das einen Teil des Weltklimaforschungsprogramms WCRP darstellt. Das GPCC stellt den auf Sammlung und Aufbereitung von Niederschlagsdaten spezialisierten Teil des weltweiten Klimabeobachtungssystems GCOS dar. Das GPCC stellt verschiedene

Datensätze zur Verfügung, wobei als Datengrundlage jeweils eine große Zahl von weltweiten Niederschlagsmessungen zur Erzeugung der Felddaten herangezogen werden. Die Konzentration auf Stationsdaten bedingt die räumliche Abdeckung der GPCC-Daten. Für die Landmassen, die Kontinente sowie einige Inseln, jedoch ohne den Antarktischen Kontinent, liegen flächendeckend Daten vor. Verwendung findet hier der auf vollständiger Reanalyse beruhende Datensatz, der für klimatologische Fragestellungen wie sie hier adressiert werden von den verfügbaren Produkten am besten geeignet ist. Sein Rückgrat stellt die umfangreiche Niederschlagsklimatologie dar. Die Anzahl der Stationen als Berechnungsgrundlage für die monatlichen Anomalien schwankt hierbei von unter 10 000 bis über 45 000 (RUDOLF und SCHNEIDER 2005).

Das GPCC ist zugleich Teil des Global Precipitation Climatology Project (GPCP). Im GPCP werden die in situ Messungen von 7 000 bis 8 000 Stationen, das Monitoring Product vom GPCC, mit den Messungen von verschiedenen Satelliten kombiniert (ICSU und WMO 1990). Die Daten liegen weltweit flächendeckend vor. Dieser Ansatz hat den Vorteil, dass für weite Teile der Erdoberfläche mehrere unabhängige Datenquellen zur Verfügung stehen. Zur weiteren Stabilisierung der Datenbasis fließen hier auch die Analyseergebnisse von atmosphärischer Zirkulationsmodelle ein. Dieses als „satellite-gauge-model“ charakterisierte GPCP-Produkt wird hier verwendet (HUFFMAN et al. 1995).

Die Felddaten mit der räumlich höchsten Auflösung stammen von der Tropical Rainfall Measuring Mission (TRMM). Diese Niederschlagsdaten werden im Rahmen des Erdbeobachtungssystems gemeinsam von NASA und JAXA erstellt und verfügbar gemacht. Von den diversen TRMM Produkten wird hier konkret „TRMM and Other-GPI Calibration Rainfall“ mit der Produkt-ID 3B42, entsprechend dem dahinterstehenden Algorithmus, genutzt. Polwärts von 50° Nord bzw. Süd liegen keine Daten vor. Auch dieser Datensatz basiert auf indirekten satellitengestützten Abschätzungen, kombiniert mit realen Messungen von Niederschlagsstationen. Die Fernerkundungsprodukte entstehen als Zwischenschritte als Ergebnis von Messungen im Mikrowellenbereich und im Infrarot. Diese Abschätzungen werden anschließend mit monatlichen Stationsdaten skaliert. Die eigentliche zeitliche Auflösung liegt bei drei Stunden, wobei der geschätzte Niederschlag für die Summe innerhalb dieses Zeitintervalls repräsentativ sein soll. Dies ist mit großen Unsicherheiten behaftet, die aber durch die hier verwendete monatliche Zeitskala weitgehend behoben sein sollten. Neben den Unsicherheiten, die sich durch die hohe raumzeitliche Auflösung ergeben, muss ein schwerwiegendes Defizit der TRMM-Daten hervorgehoben werden. Dieser aufwendig erzeugte Datensatz ist nur für einen relativ kurzen Zeitraum verfügbar (vgl. Tab. 2.1). Somit ist seine Verwendung bei der Bildung von Klimatologien und somit zur Validation der AOGCMs nicht möglich. Stattdessen werden die TRMM-Daten im Wesentlichen für die Beurteilung der Zuverlässigkeit der anderen Beobachtungsdaten bei der Erfassung des Afrikanischen und Indischen Monsuns eingesetzt. Hierfür ist jedoch

die Überlappungsperiode insbesondere mit CRU und ERA 40 nicht lange genug.

### 2.1.2 Zeitreihen

Einige wichtige Kenngrößen des Klimasystems können am schnellsten und unkompliziertesten über einfache Zeitreihen oder Indices erfasst werden. Diese Daten, die einfach über  $X(t)$  dargestellt werden, können sämtlich aus in Kapitel 2.1.1 beschriebenen Datensätzen über Standardverfahren extrahiert werden. Dennoch wurden die meisten der grundlegenden Zeitreihen separat aus etablierten Quellen bezogen. Deren Vorteile liegen zum einen in der hohen Datenqualität, zum anderen aber auch im oftmals längeren Verfügbarkeitszeitraum.

#### 2.1.2.1 Zirkulationsindices

Bei der Identifikation der global bedeutsamen großskaligen Zirkulationsmuster wird auf eine Reihe frei verfügbarer Indices zurückgegriffen. Diese dienen zur Absicherung der Befunde in den Flächendaten. Über Korrelationsanalysen der Zeitreihen wird sichergestellt, dass Zirkulationsmoden, welche nicht nach den üblichen Standardverfahren aus den Beobachtungsdaten gewonnen werden, richtig interpretiert werden. Dies könnte zwar im Prinzip ebenso mit klassisch berechneten Indexreihen aus den oben beschriebenen Flächendaten für den Luftdruck geschehen, allerdings ist angesichts des unkonventionellen Vorgehens in dieser Arbeit die Verwendung unabhängiger Daten vorzuziehen, um unerwünschte Nebeneffekte auszuschließen. Solche eventuell einflussreichen Störgrößen können beispielsweise durch Inhomogenitäten oder aber einfach durch räumliche Korrelationen in den Datensätzen hervorgerufen werden. Eine andere potentielle Quelle für derartige Effekte sind schwer detektierbare Artefakte der numerischen Berechnung. Die stattdessen herangezogenen Zeitreihen sind hingegen vielfach überprüft und erfolgreich für vielfältige Fragestellungen eingesetzt worden. Sie können somit als zuverlässig eingestuft werden. Die hier beschriebenen Zeitreihen werden in der Arbeit üblicherweise mit  $I(t)$  bezeichnet. So es erforderlich scheint werden ergänzende Indices genutzt.

Hauptquelle für diese Zeitreihen ist das Archiv des Climate Prediction Centers der NOAA. Hier finden sich Indexreihen für die grundlegenden außertropischen Variabilitätsmoden der Nordhemisphäre. Hierbei handelt es sich um aktualisierte Versionen der detaillierten Untersuchung von BARNSTON und LIVEZEY (1987). Berücksichtigt werden nur die Zeitreihen der ganzjährig als führende Moden identifizierten Variabilitätsmuster.

Daneben werden über den Climate Data Guide von NCAR zur Verfügung gestellte Zeitreihen genutzt. Sie basieren entweder auf Berechnungen anhand von Stationsdaten oder auf den rekonstruierten Luftdruckfeldern von TRENBERTH und

PAOLINO JR (1980). Daher sind hier deutlich längere Zeitreihen verfügbar als bei den NOAA-Daten.

Der SAM stellt angesichts der langfristig betrachteten schlechten Datenlage für die außertropische Südhemisphäre einen Sonderfall dar. Anders als bei den übrigen Moden divergieren hier zudem auch die nach der Standarddefinition berechneten Indexreihen je nach verwendeter Datengrundlage, insbesondere vor den 1980er Jahren, deutlich. Als am zuverlässigsten kann hierbei wohl der HadSLP2r Datensatz gesehen werden. Jedoch zeigt auch diese Zeitreihe deutliche Abweichungen von dem, allerdings nur auf einigen Stationen im beruhenden Index von der British Antarctic Survey (BAS), der hier als am Referenz verwendet wird (MARSHALL 2003). Für die Zeit nach 1979 wird zudem die SAM-Zeitreihe der NOAA herangezogen.

Das ENSO Phänomen schließlich kann über sehr viele verschiedene Indices erfasst werden. Auch in der aktuellen Literatur kommen komplexe multivariate Definitionen neben solchen, die auf einer einzelnen Größe beruhen, vor. In dieser Arbeit wird eine aktualisierte Version des klassischen SOIs verwendet (TRENBERTH 1984), da für die AOGCMs die ansonsten benötigte Meeresoberflächentemperatur nur eingeschränkt verfügbar war. Für die Identifikation der Indexzeitreihen wäre ein SST-basierter Index angesichts der hohen Korrelationen zwischen diesen und dem SOI äquivalent verwendbar. Allerdings weisen die Spektren von solcherart unterschiedlich berechneten Zeitreihen einige, wenn auch nur quantitative, Unterschiede auf (RASMUSSEN und CARPENTER 1982). Da die Frequenzdomäne als wichtige Gemeinsamkeit von Modell- und Beobachtungsdaten in dieser Arbeit angesehen wird, liegt hierin ein weiterer Grund, auf diesen SLP-basierten vom NCAR Climate Data Guide bereitgestellten Datensatz zurückzugreifen.

### 2.1.2.2 Die globale Mitteltemperatur

Der rezente Klimawandel lässt sich durchaus begründet unter dem Schlagwort der Globalen Erwärmung zusammenfassen. Im beobachteten und prognostizierten Anstieg der globalen Temperatur resultiert die verringerte Fähigkeit des Klimasystems, aus der solaren Einstrahlung erhaltene Energie über Strahlungskühlung an das All abzugeben (z. B. HANSEN et al. 1997, 2006). Die mindestens seit der Industrialisierung steigenden Treibhausgaskonzentrationen bewirken die Verlagerung der effektiven Ausstrahlung in höhere, kältere Regionen der Atmosphäre. Dadurch wird, dem Gesetz von Stefan-Boltzmann folgend, die Ausstrahlung der Erde reduziert. Das hieraus resultierende Ungleichgewicht im terrestrischen Strahlungshaushalt kann nicht auf Dauer bestehen und muss ausgeglichen werden. Dies kann nur über die Erhöhung der Ausstrahlung geschehen. Dementsprechend müssen sich die oberen Atmosphäreschichten erwärmen. Die konvektive Dynamik der Troposphäre bedingt wiederum, dass diese Erwärmung nicht isoliert stattfinden kann, sondern sich bis in die untersten Regionen der Atmosphäre und zur Erdoberfläche fortsetzt (vgl. etwa TAYLOR 1991, 2002). Dieser anthropogene Treibhauseffekt ist

**Tab. 2.2:** Beobachtungsdaten (Zeitreihen). Diskussion vieler Zeitreihen z. B. in BARNSTON und LIVEZEY (1987) und BARRY und CARLETON (2001) sowie IPCC (2007a)

Quelle	Zeitreihe(n) mit Zeitraum und hier verwendeten Codes
NOAA (Climate Prediction Centre)	je 01/1950–12/2011: North Atlantic Oscillation (NAO), East Atlantic Pattern (EA), West Pacific Pattern (WP), East Pacific/North Pacific Pattern(EP/NP), Pacific/North American Pattern (PNA), East Atlantic/West Russia Pattern (EA/WR), Scandinavia Pattern (SCA), Polar/Eurasia Pattern (POL), Arctic Oscillation (NAM); 01/1079–12/2011: Antarctic Oscillation (SAM)
NCAR (Climate Data Guide)	01/1864–12/2011: North Atlantic Oscillation (NAO, Stationsdaten), je 01/1899–12/2011: North Atlantic Oscillation (NAO, EOF-basiert), North Pacific Index (NPI), 01/1866–12/2011: Southern Oscillation Index (SOI)
BAS	03/1957–12/2011: Southern Annular Mode (SAM, Stationsdaten)
NASA	01/1880–12/2011: GISSTEMP3 ( $T_E$ )

der grundlegende Mechanismus des globalen Klimawandels.

Die gemittelte Temperatur der Erdoberfläche kann zudem als Maß für die dem Klimasystem zur Verfügung stehende Energie betrachtet werden (PIERREHUBERT 2010). Jede Veränderung in dieser Größe kann daher weitreichende Auswirkungen auf die anderen Komponenten und Kenngrößen des Klimasystems haben (HANSEN 1988, LENTON et al. 2008). Die globale Mitteltemperatur kann somit häufig dazu genutzt werden, die komplexe klimatische Wirklichkeit auf eine greifbare Zahl zu reduzieren.

Daher erfährt die Zeitreihe die diese Größe abbildet besonderes, durchaus nicht unbegründetes Interesse, sowohl von Seiten der Wissenschaft als auch seitens der Öffentlichkeit (etwa JONES et al. 1999). Hierbei gilt das Augenmerk weniger der tatsächlichen Temperatur als vielmehr den Temperaturanomalien, da die exakte Globaltemperatur aufgrund von mangelhafter flächendeckender Datenverfügbarkeit nicht bekannt ist (HANSEN et al. 1981). Als Orientierungswert kann die jährliche Mitteltemperatur für die Periode 1951–1980 mit  $14^\circ\text{C}$  veranschlagt werden. Die jahreszeitlichen oder monatlichen globalen Temperaturen sind mit höheren Unsicherheiten verbunden. Doch auch für die jährliche Temperatur muss ein Schwankungsbereich von einigen  $1/10$  K angenommen werden.

Dagegen sind die Anomalien dieser Feldgröße wegen ihrer hohen Erhaltungsneigung im Raum auch in Gegenden mit nur spärlichen Messnetzen relativ genau erfassbar. Die räumliche Korrelation von Temperaturanomalien ist derart ausge-

prägt, dass die GISS-Analysen von einem Einflussbereich von 1200 km für diese Größe ausgehen (HANSEN und LEBEDEFF 1988, HANSEN et al. 2010).

Die Entscheidung für die Nutzung von GISTEMP als Zeitreihe der globalen Temperatur ist letztlich durch seine Zielgröße begründet, die ihn von einigen vergleichbaren Datensätzen unterscheidet. Teilweise wird die tatsächliche Oberflächentemperatur verwendet, die der häufig genutzten Meeresoberflächentemperatur entspricht. Diese Temperatur ist etwa mithilfe von Satellitendaten relativ gut messbar. Für den Menschen und die Biosphäre ist diese Temperatur allerdings nicht unbedingt die ausschlaggebende. Hier ist die bodennahe Lufttemperatur, die standardmäßig in zwei Metern Höhe definiert wird, bedeutender. In absoluten Werten können sich diese beiden Größen durchaus deutlich unterscheiden. Für Landgebiete hat insbesondere die Vegetation einen großen Einfluss, was sich beispielsweise in Temperaturgradienten zwischen Kronendach und Waldboden widerspiegelt. Vergleichbare Unterschiede zwischen diesen beiden Definitionen der Oberflächentemperatur können auch im Fall von Meereisbedeckung auftreten. Anomalien der genannten Varianten und Definitionen einer Oberflächentemperatur sind jedoch miteinander üblicherweise hochkorreliert, sodass auf globaler Ebene andere verfügbare Datensätze ebenso verwendet werden können. Für die Modelle waren zum Zeitpunkt des Downloads allerdings bedeutend mehr Datensätze für die bodennahe Temperatur vorhanden, so dass diese Variable gewählt wurde.

Somit werden die Eigenschaften der GISTEMP-Zeitreihe auch auf die Modelldaten übertragen. Es werden grundsätzlich monatliche Anomalien der globalen Mitteltemperatur betrachtet, wobei die monatlichen Mittelwerte jeweils für die Periode 1960–1999 gebildet wurden. Diese so erzeugte Größe, die eben nicht die globale Mitteltemperatur sondern nur deren Anomalie darstellt, wird hier mit  $T_E$  bezeichnet, die Zeitreihe entsprechend mit  $T_E(t)$ .

## 2.2 Modelldaten

Alle Klimamodelldaten wurden im Zuge des 3. Coupled Model Intercomparison Project (CMIP3) des World Climate Research Programme (WCRP) koordiniert erstellt. Trotz des großen Umfangs der hier verwendeten Daten, sind diese doch nur ein kleiner Teil des Outputs der unterschiedlichen Experimente, die zwischen Ende 2003 bis 2007 in diesem Kontext durchgeführt wurden (MEEHL et al. 2007). Sie bildeten eine tragende Säule des 4. Sachstandberichts des IPCC (IPCC 2007a).

### 2.2.1 Randbedingungen

Das 20. Jahrhundert enthält an Forcing zum einen die beobachteten Treibhausgaskonzentrationen sowie einige weitere Größen mit bekanntem Einfluss auf den terrestrischen Strahlungshaushalt. Im CMIP3 lautet die Bezeichnung für die vorge-

schriebenen Randbedingungen 20C3m. Eine der zusätzlichen Einflussgrößen sind explosive tropische Vulkanausbrüche, die je nach Modell über eine Erhöhung der Aerosolkonzentrationen oder aber durch die Erhöhung der Stratosphärentemperatur entsprechend der empirischen Befunde simuliert werden. Diese, zu großen Teilen auf echten Daten beruhenden Randbedingungen sind für sämtliche Läufe grundsätzlich gleich, jedoch werden bei einigen Modellen zusätzlich weitere Faktoren, etwa die Entwicklung des stratosphärischen Ozons, berücksichtigt, nicht jedoch in allen. Die Randbedingungen für die Klimamodelle ändern sich also im Laufe der Simulation, so dass es sich bei den hier untersuchten Läufen nicht um Zeitscheibenexperimente, sondern um transiente Simulationen handelt, die prinzipiell in der Lage sind, auch den beobachteten Zeitverlauf von Klimaänderungen im 20. Jahrhundert nachzuvollziehen. Jenseits der globalen Temperatur, deren Veränderungen jenseits der interannuellen Variabilität nahezu ausschließlich auf die vorgeschriebenen atmosphärischen Randbedingungen zurückzuführen sind, sind die Ergebnisse jedoch wenig befriedigend. Begründet ist dies zum einen in individuellen Modelldefiziten, zum anderen in den Startbedingungen mit denen AOGCM-Läufe initialisiert werden. Da der tatsächliche globale Zustand des Klimasystems zu diesen, typischerweise in der zweiten Hälfte des 19. Jahrhunderts liegenden Zeitpunkt nicht bekannt ist, können nur plausible Zustände hierfür verwendet werden, die in den zeitlichen Verlauf der Modellsimulationen eingreifen. Ab der Jahrtausendwende werden die CMIP3 Läufe nicht mehr mit beobachteten Treibhausgaskonzentrationen angetrieben, sondern von solchen, die über ein Szenario vorgegeben werden. Hieraus ergeben sich dann transiente Simulationen bis zum Ende des 21. Jahrhunderts.

Im CMIP3 wurden verschiedene Entwicklungen der Treibhausgaskonzentrationen als Randbedingung für die Simulation des künftigen Klimas vorgegeben. Hierfür wurden unterschiedliche Szenarien entwickelt, die auf einer Vielzahl von idealisierten Annahmen beruhen (NAKIĆENović et al. 2000). Steuernde Faktoren der Szenarien sind die Entwicklung der Weltbevölkerung, des technischen Fortschritts, die Basis der Energieerzeugung, die Ausrichtung des globalen Wirtschaftssystems sowie die regionalen Disparitäten in der wirtschaftlichen Entwicklung. Unterschiedliche Kombinationen dieser Parameter finden sich in storylines wieder. Anhand derer wird der globale Ausstoß an anthropogenen Emissionen – Treibhausgase, Aerosole – über entsprechende Modelle berechnet. Diese Konzentrationsentwicklungen werden dann als Randbedingungen für die AOGCMs verwendet. Somit sind alle Szenarien mit Unsicherheiten behaftet und es kann wohl festgestellt werden, dass keines der 40 so entstandenen Szenarien der tatsächlichen Zukunft entspricht. Nichtsdestotrotz bilden sie in ihrer Gesamtheit zumindest eine gewisse Spannbreite ab, innerhalb der sich die Entwicklung des Klimasystems, je nach der Intensität des menschlichen Einflusses, abspielen wird. So ermöglicht etwa der Vergleich von Szenarien mit aktiver bzw. fehlender Emissionsverminderung von Seiten der Energiewirtschaft die Abschätzung des Nutzens der selbigen und somit auch eine rudimentäre Kosten-Nutzen-Analyse für eine aktive Klimapolitik.

In dieser Arbeit liegt der Fokus jedoch nicht auf solcherart differenzierten Analysen, sondern die AOGCMs selbst sollen beurteilt werden. Daher wird von den drei grundlegenden Szenarien, die von den meisten Modellen im Rahmen von CMIP3 gerechnet wurden, nur das SRES A1b betrachtet. Hierbei handelt es sich, bei der langfristigen Entwicklung der Treibhausgaskonzentrationen, um das mittlere Szenario. Man ist daher häufig geneigt, es auch als das realistischste zu betrachten. Angesichts der vielen, de facto nicht vorhersehbaren Einflussfaktoren, die in die storylines eingehen muss aber immer betont werden, dass es sich nur um mögliche Versionen der Zukunft handelt. Daher stellen Klimasimulationen keine Prognosen dar, sondern werden als Projektionen bezeichnet. Für die erste Dekade der Szenarien liegen mittlerweile empirische Befunde vor. Demnach übertrifft der tatsächliche CO<sub>2</sub>-Ausstoß die meisten Modellrechnungen die zur Festlegung des SRES A1b geführt haben, wenn auch nicht den letztlich zum Antrieb der AOGCMs verwendeten gewählten illustrativen Verlauf (LE QUÉRÉ et al. 2009, MANNING et al. 2010).

Ausschlaggebend für die Wahl des A1b-Szenarios war allerdings schlussendlich nicht die Bewertung der Zuverlässigkeit des vorgegebenen Forcings und somit des Strahlungsantriebs, sondern die Datenbasis. Mit diesem Szenario stehen für die grundlegenden Variablen etwa 50 Läufe zur Verfügung. Demgegenüber wären bei der Verwendung des B1 oder A2 Szenarios nur etwa 40 bzw. 30 transiente Simulationen zu verwenden gewesen. Da sich ein großer Teil der Arbeit zudem dem Vergleich von Modell und Beobachtungsdaten widmet und sich somit auf das 20. Jahrhundert bezieht würde die Berücksichtigung dieser Läufe auch die Datenbasis für weite Teile der Untersuchungen nicht vergrößern.

### 2.2.2 AOGCMs

Die Randbedingungen innerhalb des Multi-Modell Ensembles sind somit weitgehend homogenisiert. Dies gilt indes nicht für die Modelle selbst. Diese sind schon in augenfälligen Parametern wie der horizontalen und vertikalen Auflösung durch deutliche Unterschiede gekennzeichnet. Während die grundlegenden Gleichungen, der dynamische Kern des Klimamodells, in allen AOGCMs die gleichen sind, ist deren numerische Umsetzung Bestandteil der spezifischen Modellphysik (hierzu z. B. TRENBERTH 1992, WASHINGTON und PARKINSON 2005). Dies gilt im noch viel stärkeren Maß für diejenigen Prozesse, die aufgrund von Skaleneffekten nur über Parametrisierungen abgebildet werden können. Hierzu zählen durchaus wichtige Größen wie die Wolkenbildung oder auch der Niederschlag. Auch die Kopplung des Atmosphären- mit dem Ozeanmoduls ist ein wichtiger Punkt in dem sich die diversen AOGCMs unterscheiden.

Es ist allerdings zu beachten, dass längst nicht alle in Tabelle 2.3 aufgeführten Modelle komplett eigenständige, im Sinn von unabhängige, Entwicklungen sind. Dagegen spricht schon ganz generell der wissenschaftliche Prozess, der die Entwick-

**Tab. 2.3:** Überblick AOGCM-Daten (SRES A1b, CMIP3).  $Res_A$ : ungefähre Auflösung des Atmosphärenmoduls in °Latitude x °Longitude;  $L_A$ : Anzahl Atmosphärenlevel;  $L_O$ : Anzahl Ozeanlevel; Jeweils Angabe der Anzahl der Einzelläufe pro AOGCM für drei Klimavariablen (T: bodennahe Temperatur (2 m über Grund); R: Niederschlag; P: Luftdruck auf Meeresniveau); die Referenz bezieht sich auf das Gesamtmodell oder, wenn keine derartige Publikation zur Verfügung stand, das Atmosphärenmodul.

Name	$Res_A$	$L_A/L_O$	T	R	P	Referenz
BCCR-BCM2.0	1.9 x 1.9	31/35	1	1	1	DÉQUÉ et al. (1994)
CGCM3.1(T47)	2.8 x 2.8	31/29	1	5	5	McFARLANE et al. (1992)
CGCM3.1(T63)	1.9 x 1.9	31/29	1	1	1	McFARLANE et al. (1992)
CNRM-CM3	1.9 x 1.9	45/31	1	1	1	DÉQUÉ et al. (1994)
GFDL-CM2.0	2.0 x 2.5	24/50	1	1	1	FREIDENREICH et al. (2004)
GFDL-CM2.1	2.0 x 2.5	24/50	1	1	1	FREIDENREICH et al. (2004)
GISS-AOM	3.0 x 4.0	12/16	2	2	2	RUSSELL et al. (1995)
GISS-EH	3.0 x 5.0	20/16	3	3	3	SCHMIDT et al. (2006)
GISS-ER	4.0 x 5.0	20/13	3	3	3	SCHMIDT et al. (2006)
FGOALS-g1.0	2.8 x 2.8	26/16	2	2	2	WANG et al. (2004)
INM-CM3.0	4.0 x 5.0	21/33	-	1	1	ALEXEEV et al. (1998)
IPSL-CM4	2.5 x 3.8	19/31	1	1	1	HOUDIN et al. (2006)
INGV-SXG	1.1 x 1.1	56/47	1	1	1	SCOCIMARRO et al. (2007)
MIROC3.2(hires)	1.1 x 1.1	56/47	1	1	1	ABE-OUCHI et al. (2004)
MIROC3.2(medres)	2.8 x 2.8	20/43	3	3	3	ABE-OUCHI et al. (2004)
MRI-CGCM2.3.2	2.8 x 2.8	30/23	5	5	5	SHIBATA et al. (1999)
ECHO-G	3.9 x 3.9	19/20	3	3	3	ROECKNER et al. (1996)
CSIRO-MK3.0	1.9 x 1.9	18/31	1	1	1	GORDON et al. (2002)
CSIRO-MK3.5	1.9 x 1.9	18/31	1	1	1	GORDON et al. (2002)
ECHAM5-MPI/OM	1.9 x 1.9	31/15	4	4	4	ROECKNER et al. (2003)
NCAR-CCSM3	1.4 x 1.4	26/40	7	7	7	COLLINS et al. (2006)
NCAR-PCM	2.8 x 2.8	26/40	3	3	3	KIEHL et al. (1998)
UKMO-HadCM3	2.5 x 3.8	19/20	1	1	1	POPE et al. (2000)
UKMO-HadGEM1	1.2 x 1.9	38/40	1	1	1	MARTIN et al. (2004)

lung von Klimamodellen immer begleitet hat. In besonderem Maße gilt dies aber für solche Modelle, die beispielsweise über identische Teilmodule oder Parametrisierungsschemata verfügen. So unterscheiden sich das GISS-EH und das GISS-ER nur im verwendeten Ozeanmodul. Das Modell ECHO-G wiederum verwendet, ebenso wie INGV-SXG, für die Atmosphäre den dynamischen Kern von ECHAM4, dem direkten Vorläufer des ebenfalls am CMIP3 beteiligten ECHAM5-MPI/OM. Die Modelle MIROC3.2(H) und MIROC3.2(M) unterscheidet nur die unterschiedliche

Auflösung, ebenso liegt das CGCM3.1 in zwei räumlichen Auflösungen vor.

Hier wird jedoch die gängige Praxis übernommen diese Modelle, trotz der offensichtlichen Überschneidungen, einfach als nicht identisch zu betrachten. Somit stehen für die weiteren Untersuchungen 53 Einzelläufe von 24 verschiedene AOGCMs zur Verfügung, von denen elf als Modellensembles mit mehrfachen Realisierungen vorliegen. Leider gilt dies nur für die Variablen Luftdruck auf Meeresniveau und Niederschlag. Für die bodennahe Temperatur reduziert sich die Datenbasis auf 48 Läufe. Für INM-CM3.0 war diese Größe ebenso wenig wie für den größten Teil des CGCM3.1(T47)-Ensembles in der Datenbank verfügbar, was die Zahl der Modelle auf 23 reduziert. Für diese Variable liegen 10 Modelle als Modellensembles vor. Tabelle 2.3 gibt einen knappen Überblick über alle hier verwendeten Läufe, inklusive der jeweiligen Schlüsselreferenzen.

## 2.3 Datentransformation (räumliche Interpolation)

Wie die Tabellen 2.1 und 2.3 zeigen, liegen die Raumdaten, sowohl Beobachtungen als auch Modelldaten, auf einer Vielzahl unterschiedlicher Gitternetze vor. Für quantitative Vergleiche dieser Datensätze, so sich diese nicht auf die in Kapitel 2.1.2 vorgestellten raumunabhängige Zeitreihen beschränken sollen, ist die Transformation aller Daten auf ein gemeinsames Gitternetz unvermeidlich. Formal durchgeführt werden diese Transformationen als Interpolation der Originaldaten auf ein gemeinsames Gitternetz. Auf welches Netz hierbei mit welcher Methode interpoliert wird ist abhängig von der betrachteten Variable, den verschiedenen Auflösungen der Originaldaten sowie dem konkreten Verwendungszweck, also der durchgeführten Untersuchung.

Für kontinuierliche Feldgrößen wie Temperatur und Luftdruck lässt sich ein weiteres Gitternetz rechtfertigen als für den räumlich hochvariablen Niederschlag. Grundsätzlich ist es aber vorteilhaft, wenn sich die Auflösungen von Originalgitter und Interpolationsgitter nicht zu stark unterscheiden. Für den direkten Abgleich der Raumdaten bietet sich im globalen Maßstab ein regelmäßiges  $2.5^\circ \times 2.5^\circ$ -Gitter an, wie es die ERA 40- und NCEP/NCAR-Reanalysedaten nutzen. Die räumliche Auflösung der meisten AOGCMs ist vergleichbar, sodass ungewollte Effekte durch die Interpolation, die sich als störende Artefakte äußern können, weitgehend ausgeschlossen werden können und rein distanzorientierte, statische Verfahren verwendet werden können. So wird etwa auf den Ausgleich der unterschiedlichen, den Daten zugrundeliegenden Topographien verzichtet. Den in der Arbeit gezeigten Karten liegt üblicherweise dieses Gitter zugrunde, jedoch mussten für einige Analysen andere räumliche Auflösungen verwendet werden.

Trotz leistungsfähiger Rechner und optimierter Rechenverfahren stellen die

technischen Anforderungen und die tatsächliche Rechenzeit einiger numerischer Verfahren immer noch limitierende Faktoren bei der klimatologischen Forschung dar. Das gilt natürlich im verstärkten Maße, wenn eine ganze Vielzahl von Datensätzen untersucht wird. Insbesondere die hier häufig praktizierte Durchführung von globalen EOF-Analysen (vgl. Kapitel 3.5) ist unter diesen Umständen problematisch. Im einfachsten Fall, bei der Untersuchung einer einzelnen Feldgröße, ist die Diagonalisierung einer räumlichen Kovarianzmatrix erforderlich, die in der  $2.5^\circ \times 2.5^\circ$ -Auflösung über  $10\,512 \cdot 10\,512 = 110\,502\,144$  Elemente verfügt. Wird stattdessen ein  $5^\circ \times 5^\circ$ -Gitter verwendet so reduzieren sich die Dimensionen dieser Matrix auf  $2\,664 \times 2\,664$  oder  $7\,096\,896$  Elemente.

Neben diesen Aspekten der technischen Praktikabilität besteht allerdings auch einer der thematischen, der für die Durchführung von EOF-Analysen auf diesem relativ groben Gitter spricht. Die hier betrachteten Zirkulationsmuster sind allesamt großskalige Phänomene, deren Existenz und Ausprägung nicht, oder allenfalls gering, von der räumlichen Auflösung abhängt. Der Grund für diese Eigenschaft ist, dass atmosphärische Daten, insbesondere im Raum, „rot“ sind, und die klimatologischen Aktionszentren Ausdehnungen von mehreren tausend Kilometern aufweisen. Besser aufgelöste Daten bedeuten daher üblicherweise ein Vielfaches an Rechenaufwand ohne echten Zugewinn an Information. EOF-Analysen für Felddaten werden daher üblicherweise auf relativ groben Gitternetzen ausgeführt (vgl. MOBLEY und PREISENDORFER 1988). Dementsprechend werden solche Berechnungen in dieser Arbeit auf einem regelmäßigen  $5^\circ \times 5^\circ$ -Gitter, wie es der HadSLP2r Datensatz verwendet, durchgeführt. Die in Kapitel 6 besprochenen Untersuchungen wurden jedoch mit unterschiedlichen Auflösungen auf Stabilität getestet, um den Einfluss einer speziellen Interpolation auch empirisch auszuschließen. Die dort vorgestellten und im weiteren Verlauf der Arbeit verwendeten Ansätze haben sich als unabhängig vom konkreten Gitternetz erwiesen.

Im Folgenden werden die verwendeten Interpolationsverfahren vorgestellt. Dabei wird auch erklärt, für welche Teilaufgaben und warum die jeweilige Methode genutzt wurde. Es handelt sich durchgehend um lokale Verfahren, weswegen grundsätzlich nur Teilmengen der Ausgangsdaten, nämlich jene in der unmittelbaren Nachbarschaft der neu zu bestimmenden Gitterboxen, berücksichtigt werden.

Regionen innerhalb der Originaldatensätze die keine Daten enthalten wurden während der Interpolation nicht berücksichtigt. Dies ist insbesondere bei Beobachtungsdaten relevant, die beispielsweise keine Abdeckung der Ozeane aufweisen. Die transformierten Daten enthalten entsprechend vergleichbare Datenlücken wie der Ausgangsdatensatz. Dieses Vorgehen reduziert die Anforderungen an die Verfahren erheblich, so dass auf aufwendige Methoden, wie sie v. a. für die Interpolation des Niederschlags existieren, verzichtet werden kann (z. B. JEFFREY et al. 2001).

### 2.3.1 Inverse-Distanz Wichtung

Die Methodik der Inversen-Distanz Wichtung (IDW) bietet sich insbesondere an, wenn Ziel- und Ausgangsgitternetz eine vergleichbare Auflösung haben (PRICE et al. 2000). Unter dieser Bedingung liefert sie bei der Interpolation von Klimavariablen sehr ähnliche Ergebnisse wie aufwendige Verfahren der Geostatistik (vgl. SAMANTA et al. 2012). Hierbei ergibt sich  $Y(t, s_y)$  als der auf einen Punkt  $s$  interpolierte Wert einer Feldgröße  $X$  über eine deterministische Beziehung aus den Ursprungsdaten.  $Y(t, s_y)$  ist somit eine Funktion der Originaldaten  $X(t, s_x)$ , deren Funktionswerte über die relativen Entfernungen bestimmt werden. Allgemein lässt sich die IDW-Interpolation, lokal durchgeführt unter der Berücksichtigung von  $m$  Gitterpunkten  $s_x$ , schreiben als:

$$Y(t, s_y) = \frac{1}{\sum_{s_x=1}^m \omega(s_x)} \sum_{s_x=1}^m \omega(s_x) X(t, s_x) \quad (2.1)$$

Hier bezeichnet  $X(t, s_x)$  den Wert der Feldgröße  $X$  am Gitterpunkt  $s_x$  und  $Y(t, s_y)$  den interpolierten Wert am Punkt  $s_y$  im transformierten Gitternetz. Immer, wenn die IDW hier genutzt wird, wird  $m = 6$  gesetzt. Die verwendeten Punkte weisen die geringste euklidische Distanz  $d(s_y, s_x)$  zum gesuchten Datenpunkt des Zielnetzes auf. Die Gewichte in Gleichung 2.1 ergeben sich über:

$$\omega(s_x) = \left( \frac{1}{d(s_y, s_x)} \right)^2 \quad (2.2)$$

Hierbei werden die Koordinaten im Himmelsgradnetz verwendet. Da die Interpolation lokal erfolgt und die Daten keine relevanten Lücken aufweisen, wäre die explizite Berücksichtigung der tatsächlichen Geometrie der Erde ohne relevante Auswirkungen.

Über die IDW wurden auch die geringer aufgelösten Beobachtungsdatensätze auf das  $2.5^\circ \times 2.5^\circ$ -Gitter interpoliert. Da nicht alle diese Datensätze global vorliegen (siehe Kapitel 4), hat das zusätzliche Quadrieren nicht nur einen glättenden Einfluss auf die resultierende Datenoberfläche, sondern zusätzlich einen screening effect, analog zu Techniken der Geostatistik (z. B. OLEA 1999). Ansonsten sind die Unterschiede zwischen 2.2 und einer linearen IDW für die hier verwendeten Felddaten gering, da die Datenpunkte insgesamt sehr regelmäßig verteilt sind.

### 2.3.2 Flächengewichtete Interpolation

Einige Beobachtungsdaten liegen in deutlich höheren Auflösungen vor als die Modellsimulationen. Daten aus GCMs werden üblicherweise als Punktdaten betrachtet und auch als solche interpoliert (vgl. Diskussion bei SKELLY und HENDERSON-SELLERS 1996). Gerade im Fall von Beobachtungen, die idealerweise über die

Kombination mehrerer Punktmessungen pro Gitterbox generiert werden, ist die alternative Betrachtung von Gitternetzdaten als Flächendaten aber nicht unbegründet.

So ist eine höhere Auflösung in der Horizontalen mit einem ausgeprägteren Relief verbunden, was speziell bei der 2 m-Temperatur starke Gradienten auf einer vergleichsweise geringen Distanz bewirken kann. Bei einer lokalen Interpolation ist es somit möglich, eine große Gitterzelle mit nicht repräsentativen Werten auszustatten, wenn die zum zentralen Punkt der Gitterzelle nächstgelegenen Gitterpunkte im Originalgitter über eine deutlich andere Höhenlage verfügen. Diese Situation ist etwa im Alpenraum anhand des CRU-Datensatzes einfach nachzuvollziehen. Solche Effekte verursachen nicht nur Artefakte in Gebieten mit starken Reliefunterschieden, sondern beeinflussen sogar die globale Mitteltemperatur eines Datensatzes.

Daneben ist auch der Niederschlag anfällig für Verzerrungen infolge einer lokalen Interpolation zwischen zu verschiedenen Gitternetzen. So können Luv- und Leeeffekte zu einer Über- aber auch Unterschätzung des Niederschlags innerhalb der größeren Zelle führen, wenn die Interpolation nicht auf diese Rücksicht nimmt. Um solche Probleme weitestgehend auszuschließen, wird beim Upscaling der hochaufgelösten CRU und E-OBS Daten eine Interpolationsmethode angewendet, die sicherstellt, dass der Energiegehalt der interpolierten Gitterzelle, repräsentiert durch die mittlere 2 m Temperatur, sich nicht zu stark von dem des entsprechenden Gebietes in der Originalauflösung unterscheidet. Das gleiche Verfahren bewahrt bei der Anwendung auf Niederschlagsdaten entsprechend die interessierende Größe  $\text{mm/m}^2$ . Hierfür wird der interpolierte Wert  $Y(t, s_y)$  als flächengewichteter Mittelwert von  $m$  Gitterzellen im Originalgitter gebildet (vgl. Kapitel 3.1.1). Die Anzahl  $m$  ist die Summe der Gitterboxen für die  $d(s_y, s_x)$  minimal ist.

### 2.3.3 Bivariate quadratische Interpolation

Die Transformation von SLP- und Geopotential-Daten auf das  $5^\circ \times 5^\circ$ -Gitter wurde über bivariate quadratische Interpolation durchgeführt. Da sich alle Daten auf einem Druckniveau befinden, ist diese Operation vergleichsweise unproblematisch (JEFFREY et al. 2001).

An jeweils drei Gitternetzpunkten in der (topologischen) Nachbarschaft werden sowohl in zonale als auch meridionale Richtung quadratische Regressionsgleichungen (vgl. Kapitel 3.3) aufgestellt. Aus der Kombination dieser beiden, ihre Stützwerte exakt interpolierenden Hyperbeln ergibt sich die gesuchte Ausprägung von  $Y(t, s_y)$ .

Allgemein erzielt diese Interpolation, bei vergleichbaren Gitternetzauflösungen, sehr gute, glatte Ergebnisse. Jedoch ist sie nicht so universell einsetzbar wie die IDW-Methodik, schon weil Fehlwerte die Bildung der Regressionsgleichung verhindern. Außerdem befinden sich die  $Y(t, s_y)$  nicht unbedingt im Wertebereich der umgebenden  $X(t, s_x)$ . Diese Eigenschaft kann durchaus dazu führen, dass so gewon-

nene Ausprägungen von  $Y(t, s_y)$  außerhalb des Definitionsbereiches von  $X$  liegen. Solche physikalisch inkonsistenten Artefakte können die statistische Auswertung der Daten verfälschen oder zumindest erschweren. Bei den Datensätzen, für die dieses Verfahren hier angewendet wurde, kann diese Gefahr jedoch ausgeschlossen werden.

## 3 Statistische Methoden

Die Ergebnisse in dieser Arbeit basieren auf einer Vielzahl verbreiteter statistischer Methoden und deren datenadaptiven Nutzung. Im Sinne eines kompakten Überblicks werden diese hier knapp erläutert. Auf Methoden, die weniger Verbreitung in der Geographie aufweisen oder deren Grundlagen fundamental für die Interpretation der dargelegten Ergebnisse erscheinen, wird dabei etwas ausführlicher eingegangen. Die Basis der Diskussion dieser Methoden bilden verschiedene Lehrbücher zur Statistik. Die zentrale Referenz für die folgenden Ausführungen ist das Werk von STORCH und ZWIERS (1999). Daneben wurden weitere Lehrbücher sowohl bei der Ausarbeitung dieses Kapitels als auch bei der Implementation der Methoden herangezogen (v. a. BORTZ und SCHUSTER 2010, WILKS 2011, SCHÖNWIESE 2006, PRUSCHA 2007, BOHLEY 2000).

Ein wichtiger Aspekt dieses Kapitels ist die Einführung der in der Arbeit verwendeten Terminologie, die sich teilweise in verschiedenen Quellen, auch bei zentralen Begrifflichkeiten, durchaus unterscheidet. Weiterhin wird die verwendete Symbolik eingeführt. An dieser Stelle wird neutralen Standardbezeichnungen für Daten und Variablen der Vorzug gegenüber sinntragenden Bezeichnungen mit Bezug zum klimatologischen Themenkomplex gegeben. Ausnahmen von diesem Vorgehen werden an Stellen gemacht, an denen der Nutzen reiner Anschaulichkeit für die Erläuterung der Methodik höher erscheint als der von Allgemeingültigkeit. In den inhaltlichen Kapiteln der Arbeit wird dieses Vorgehen dagegen umgekehrt, so dass dort üblicherweise Variablenbezeichnungen verwendet werden, die an die sie repräsentierenden Größen angelehnt sind. Eigene methodische Ansätze, auch wenn sie auf den in diesem Kapitel vorgestellten Grundlagen beruhen, werden an der Stelle ihres Auftretens erläutert, um ihren Nutzen für die Aufgabenstellung zu verdeutlichen.

### 3.1 Grundlegende Datenaufbereitung und Inferenz

In dieser Arbeit werden nahezu ausschließlich metrisch skalierte Variablen untersucht. Unabhängig von der betrachteten Größe lässt sich jeder konkrete Datensatz als Vektor  $\mathbf{x}$  mit  $n$  Elementen  $x_i$  darstellen. Wenn tatsächlich die Variable oder auch Zufallsvariable gemeint ist, die durch diese Realisationen repräsentiert wird, so wird diese als  $X$  bezeichnet. In diesem Kapitel wird bevorzugt die für empirische

Daten übliche Notation in lateinischen Kleinbuchstaben verwendet. Für deren Beschreibung werden dann in der Regel nur einige wenige statistische Kenngrößen herangezogen. Diese Aggregierungsmethoden werden hier genannt und definiert. So gewonnene anschauliche Parameter besitzen auch ohne Bezug zur stochastischen Theorie und Wahrscheinlichkeitsrechnung eine gewisse Aussagekraft.

### 3.1.1 Univariate Kenngrößen

Die Bestimmung eines Mittelwerts ist die erste Form der Datenaufbereitung. Allgemein erfolgt dies durch die Formel

$$x_\omega = \sum_{i=1}^n \omega_i x_i \quad (3.1)$$

Es gilt hierbei immer  $\sum_{i=1}^n \omega_i = 1$ . Die konkreten Ausprägungen der Gewichte ist jedoch abhängig von der jeweiligen Fragestellung, insbesondere bei der Bearbeitung geowissenschaftlicher Daten. So können die  $\omega_i$  etwa unterschiedliche Flächenanteile repräsentieren. Dieses Vorgehen ist bei der Verwendung von Gitternetzdaten ein üblicher Weg, um die Konvergenz der Meridiane in Richtung der Pole zu berücksichtigen, z. B. bei der Berechnung der globalen Mitteltemperatur. Sie verhindert in diesem Fall die Unterrepräsentation der Tropen gegenüber den flächenmäßig kleinen, aber mit ebenso vielen Gitterboxen pro Breitenwall ausgestatteten Polarregionen. Die Gewichte  $\omega_i$  sind in diesem Fall somit  $\omega(s)$  und ergeben sich aus dem Kosinus der geographischen Breite der Realisation einer Feldgröße  $X(t, s)$ . Diese Flächengewichtung ist ein grundlegendes Verfahren und kommt hier auch in weiteren Methoden zur Anwendung.

Eine andere häufige Ursache für die Vergabe unterschiedlicher Gewichte ist die Bestimmung eines Mittelwerts für Daten unterschiedlicher Qualität. Im dennoch üblichen Fall der Gleichgewichtung mit  $\omega_1 = \dots = \omega_n = n^{-1}$  resultiert aus Gleichung 3.1 der in der Statistik überragend wichtige arithmetische Mittelwert  $\bar{x}$ . Die Abweichungen der Elemente  $x_i$  von diesem werden als Anomalien

$$x'_i = x_i - \bar{x} \quad (3.2)$$

bezeichnet. Sie werden immer dann herangezogen, wenn das Interesse an einem Datensatz eher auf dessen Dynamik und Variation oder auch Veränderung und Entwicklung, beruht als auf seinen absoluten Ausprägungen. Außerdem stellt die Bildung von Anomalien eine effektive Methode dar, um systematische Unterschiede, im Fall von Fehlern den Bias, aus Datensätzen zu entfernen. So kann etwa die zeitliche Entwicklung einer Größe in unterschiedlichen Klimamodellen, oder aber auch an unterschiedlichen geographischen Orten im Fall von Messdaten, ohne den Einfluss des spezifischen Mittelwerts dargestellt und untersucht werden.

Ein Maß für die Streuung innerhalb eines Datensatzes ist die Varianz  $s^2$ :

$$s^2 = \sum_{i=1}^n \omega_i (x_i - x_\omega)^2 \quad (3.3)$$

Wird dabei eine Gleichgewichtung der Daten zugrunde gelegt, lässt sich die Varianz auch als  $s^2 = \overline{x'_i{}^2}$  schreiben. Wird explizit die Varianz einer Stichprobe berechnet, wird die Quadratsumme der Anomalien mit  $n - 1$  statt  $n$  normiert. Dieses Vorgehen ist jedoch eher von theoretischer Relevanz (PRESS et al. 1992). Mit  $s = \sqrt{s^2}$  ergibt sich Standardabweichung, die einen Schätzer für die mittlere Abweichung innerhalb eines Datensatzes darstellt und die gleiche Einheit wie  $X$  aufweist. Sie wird, wie auch  $\bar{x}$ , für viele weitere aufwendigere statistische Verfahren benötigt. Dabei wird  $s$  häufig verwendet, um Variablen mit unterschiedlichen Einheiten zu standardisieren. Dies geschieht, wenn Datensätze mit systematischen Unterschieden in der Streuung untersucht werden, und diese Unterschiede bei der Analyse vernachlässigt werden können oder aber müssen. Soll gleichzeitig auch der Einfluss unterschiedlicher  $\bar{x}$  ignoriert werden, verwendet man standardisierte Anomalien.

$$\tilde{x}_i = \frac{x'_i}{s} \quad (3.4)$$

Diese sogenannte  $z$ -Transformation überführt jede beliebige Normalverteilung in die Standardnormalverteilung. Sie erzeugt, unabhängig von den Ausgangswerten und deren Verteilung, einen Datensatz  $\tilde{\mathbf{x}}$  mit  $\bar{\tilde{x}} = 0$  und  $s_{\tilde{x}}^2 = 1$ . Ein numerischer Wert  $\tilde{x}_i = 1$  entspricht somit dem Betrag der Standardabweichung von  $X$ .

### 3.1.2 Inferenz

Die hier aufgeführten Inferenzmethoden setzen ein theoretisches Verteilungsmodell voraus. Dieses besagt, dass die empirischen Variablen  $\mathbf{x}$  Realisationen einer normalverteilten Zufallsvariablen  $X \sim NV(\mu, \sigma^2)$  darstellen. In der Regel lässt sich diese Annahme bei einer ausreichend großen Datenbasis und insbesondere bei der Wahl geeigneter Zufallsvariablen rechtfertigen.

#### 3.1.2.1 Zweistichproben- $t$ -Test

Dieses Verfahren, oder besser die unter diesem Überbegriff zusammengefassten Verfahren, dienen zum Vergleich der arithmetischen Mittelwerte zweier unabhängiger Stichproben. In dieser Arbeit wird diese grundlegende Methodik der Inferenzstatistik sehr häufig angewendet. Hier wird in allen Fällen die ungerichtete  $H_0 : \mu_1 = \mu_2$  überprüft. Der arithmetische Mittelwert  $\bar{x}_1$  der Stichprobe  $\mathbf{x}_1$  ist der Schätzer für  $\mu_1$  und  $\bar{x}_2$  schätzt entsprechend  $\mu_2$ . Wenn die Prüfgröße  $t$  die Verwerfung dieser Aussage ermöglicht, können die Unterschiede in den beiden Gruppen als

signifikant zum gewählten Signifikanzniveau  $\alpha$  betrachtet werden. In diesem Fall wird die  $H_1 : \mu_1 \neq \mu_2$  angenommen. Die Alternativhypothese besagt dann, dass die Differenzen zwischen den Mittelwerten so deutlich sind, dass diese als Schätzer für unterschiedliche Erwartungswerte interpretiert werden müssen. Dieses Resultat des  $t$ -Tests bedeutet dann, dass die Gruppen aus unterschiedlichen Grundgesamtheiten stammen. Voraussetzungen des Verfahrens sind jeweils die Unabhängigkeit der Stichproben und dass das untersuchte Merkmal in den Grundgesamtheiten normalverteilt vorliegt. Gegen die Verletzung dieser Annahmen reagiert das Verfahren allerdings robust.

Der statistische Test betrachtet also die Differenzen zwischen den Gruppenmitteln. So lässt sich beim ungerichteten Test die  $H_0$  äquivalent als  $H_0 : \mu_1 - \mu_2 = 0$  schreiben. Um diese Hypothese zu testen wird noch mit  $s_d$  die Streuung der Mittelwertsdifferenz  $\bar{x}_1 - \bar{x}_2$  eingeführt. Die Prüfgröße berechnet sich dann über:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_d} \quad (3.5)$$

Sind die Varianzen von  $\mathbf{x}_1$  und  $\mathbf{x}_2$  gleich, kann  $s_d$  mit Hilfe der gepoolten Stichprobenvarianz  $s_p^2$  geschätzt werden. Diese ergibt sich als Mittel von  $s_1^2$  und  $s_2^2$ , die Gewichtungsfaktoren ergeben sich aus den jeweiligen Freiheitsgraden.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad (3.6)$$

Schließlich wird  $s_d$  über

$$s_d = \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) s_p^2} \quad (3.7)$$

berechnet und in Gleichung 3.5 eingesetzt. Die Prüfgröße in diesem Test ist  $t$ -verteilt mit  $df = n_1 + n_2 - 2$  Freiheitsgraden.

Muss allerdings davon ausgegangen werden, dass die beiden Stichproben heterogene Varianzen aufweisen, so sollte dieser Test nicht verwendet werden. In diesem Fall wird stattdessen ein WELCH-Test durchgeführt. Dieser unterscheidet sich im Vergleich zum bisher beschriebenen Vorgehen zunächst im Bezug auf die Schätzung von  $s_d$ . Diese ergibt sich hier als

$$s_d = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}. \quad (3.8)$$

Wird  $s_d$  auf diese Weise berechnet, ist die durch Gleichung 3.5 gegebene Prüfgröße nicht mehr exakt  $t$ -verteilt. Daher muss für die Durchführung der Prüfung auf Signifikanz der Differenzen die Anzahl der Freiheitsgrade modifiziert werden. Der

übliche und auch in dieser Arbeit verwendete Ansatz hierfür ist:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}} \quad (3.9)$$

Welche der beiden Varianten des  $t$ -Tests zu verwenden ist, wurde jeweils über einen zweiseitigen  $F$ -Test zum Signifikanzniveau  $\alpha = 0.1$  entschieden. Wenn für diesen die  $H_0 : \sigma_1^2 = \sigma_2^2$  verworfen werden konnte, so wurde der Welch-Test durchgeführt, ansonsten der übliche Zweistichproben- $t$ -Test.

Wenn im Folgenden einfach vom  $t$ -Test gesprochen wird, sind daher jeweils beide beschriebene Verfahren gemeint.

### 3.1.2.2 $F$ -Test für Varianzen

Um die Varianzen zweier Datensätze zu vergleichen, wird der  $F$ -Test verwendet. Signifikante Unterschiede im Streuungsparameter zeigen an, dass die Datensätze verschiedene Verteilungen aufweisen, bzw. Stichproben aus Grundgesamtheiten mit unterschiedlichen Varianzen  $\sigma_1^2$  und  $\sigma_2^2$  darstellen. Hiervon kann ausgegangen werden, wenn die  $H_0 : \sigma_1^2 = \sigma_2^2$  zum Signifikanzniveau  $\alpha$  verworfen wird. Die Prüfgröße für diesen Test ist

$$F = \frac{s_1^2}{s_2^2} \quad (3.10)$$

und folgt unter der  $H_0$ , und bei unabhängigen, normalverteilten Größen, einer  $F$ -Verteilung mit  $df_1 = n_1 - 1$  und  $df_2 = n_2 - 1$ .

## 3.2 Korrelation

In einem bivariaten Datensatz stellt die Kovarianz ein Maß für die Synchronizität der Variabilität beider Variablen dar. Werden zwei metrische Variablen  $X_1$  und  $X_2$  in den Datensätzen  $\mathbf{a}$  und  $\mathbf{b}$  erfasst, so erfolgt die Berechnung über:

$$s_{ab} = \overline{a_i' b_i} \quad (3.11)$$

Die Kovarianz ist ein Maß für eine bivariate Wenn-Dann-Beziehung zwischen  $X_1$  und  $X_2$  das Mittelwertdifferenzen ignoriert, jedoch durch unterschiedliche Streuungsparameter der Variablen, etwa auch Einheiten, beeinflusst wird. Die Produkt-Momenten Korrelation ist die übliche und häufig genutzte Methode, um den linearen Zusammenhang zweier, metrisch skaliertes, Variablen zu quantifizieren. Wenn in dieser Arbeit von Korrelation oder Korrelationskoeffizienten ohne weitere Spezifizierung gesprochen wird, so ist immer die Korrelation nach PEARSON

gemeint. Eine Deutung des Korrelationskoeffizienten  $r$  ist die Kovarianz zweier  $z$ -transformierter Variablen.

$$r = \overline{\tilde{a}_i \tilde{b}_i} \quad (3.12)$$

Er ist normiert auf den Bereich  $[-1, \dots, 1]$ . Bei  $r = 0$  liegt für zwei Variablen kein linearer Zusammenhang vor. Ein anderer Zusammenhang kann allerdings durchaus existieren, der über diese Methode nicht erfasst werden kann. Daneben muss hier noch einer weiteren Problematik der Korrelationsrechnung begegnet werden. Sie ist, besonders bei geringem Datenumfang, anfällig gegenüber Ausreißern, seien es Messfehler oder Extremwerte. Dadurch können tatsächlich vorhandene Korrelationen maskiert, aber ebenso  $r$  in die Höhe getrieben werden. Letzteres würde dazu führen, dass fälschlicherweise Zusammenhänge vermutet werden, wo keine bestehen.

Um insbesondere solche Fehlinterpretation zu vermeiden, wird in dieser Arbeit an besonders kritischen Stellen die Rangkorrelation nach SPEARMAN genutzt, wenn die Möglichkeit besteht, dass eine Korrelation auf dem Vorhandensein eines oder mehrerer Ausreißer oder Extremwerte beruht. Die eigentliche Berechnung erfolgt analog zu Gleichung 3.12, mit dem Unterschied, dass nicht die Variablen als solche sondern deren Rangplatzfolge  $z$ -transformiert werden. Die Rangkorrelation erfasst durch dieses Vorgehen auch nichtlineare, monotone Zusammenhänge. Daneben muss der Rangkorrelationskoeffizient  $r_{sp}$  herangezogen werden, wenn mindestens eine ordinal skalierte Variable bei der Berechnung der Korrelation berücksichtigt werden soll.

Wird der Korrelationskoeffizient  $r$ , ebenso auch  $r_{sp}$ , als Schätzer für den Parameter  $\rho$  verwendet, so kann dieser mit der  $H_0 : \rho = \rho_0$  auf Signifikanz geprüft werden. Üblicherweise wird hierbei  $\rho_0 = 0$  gesetzt, es ist aber auch jeder andere theoretisch mögliche Wert zulässig, anders als bei Nutzung des verbreiteten  $t$ -Tests zur  $H_0 : \rho = 0$ . Die hier betrachtete Nullhypothese kann verworfen werden, wenn die Prüfgröße  $z$  sich als signifikant zum Niveau  $\alpha$  erweist. Die Berechnung erfolgt nach:

$$z = \sqrt{n - 3} \cdot (Z - Z_0) \quad (3.13)$$

Die Prüfgröße folgt der Standardnormalverteilung. Hier bezeichnen  $Z$  und  $Z_0$  den Fisher  $Z$ -transformierten Korrelationskoeffizienten  $r$  und  $\rho_0$ .

$$Z = \frac{1}{2} \cdot \left( \frac{1 + r}{1 - r} \right) \quad (3.14)$$

Diese Transformation bewirkt, dass die Korrelationskoeffizienten in eine Intervallskala überführt werden. Außer bei der Durchführung von Signifikanztests kommt diese linearisierende Transformation von  $r$  in  $Z$  in dieser Arbeit daher auch dann

zum Einsatz, wenn der Mittelwert über verschiedene Korrelationskoeffizienten gebildet wird.

### 3.3 Lineare Regression

Die Regression ist eventuell das bedeutendste statistische Basiswerkzeug. Daher finden sich unter diesem Namen eine ganze Reihe unterschiedlicher, aber miteinander verwandter Methoden. Die in allen quantitativen Wissenschaften gebräuchliche lineare Regression wurde in verschiedenen Anwendungen auch zur Erstellung der hier präsentierten Ergebnisse verwendet. Die wesentliche Motivation für ihre etwas eingehendere Erläuterung an dieser Stelle ist, im weiteren Verlauf der Arbeit kompaktere Erläuterungen der konkreten Anwendungen zu ermöglichen. Ausführliche Darstellungen der Methodik finden sich in den meisten der genannten Werke zur Statistik sowie, vertieft und mit Berücksichtigung weiterer und alternativer Verfahren, etwa bei FAHRMEIR et al. (2009) oder BINGHAM et al. (2010).

Grundsätzlich ist es Ziel und Aufgabe der diversen Formen der Regression, ein parametrisches Modell für quantitative Zusammenhänge zwischen verschiedenen Zufallsvariablen zu erstellen. Die univariate Variante erklärt das Verhalten einer Zufallsvariablen  $Y$ , dem Prädiktand oder Regressand, durch eine oder mehrere sogenannte unabhängige Zufallsvariablen  $X_i$ . Die  $X_i$  werden entsprechend als Prädiktoren oder Regressoren bezeichnet. Beim Sonderfall der Einfachregression wird nur ein Prädiktor durch das Modell in Betracht gezogen, ansonsten handelt es sich um eine multiple Regressionsrechnung. Terminologisch abzugrenzen ist die multivariate Regression, die mehr als einen Prädiktand zu erklären versucht. Im Gegensatz zu vielen anderen statistischen Verfahren erlaubt die Regression nicht nur die Analyse vorhandener Daten sondern auch die Extrapolation. Liefert die Regression ein stabiles Modell, können theoretische Werte für die abhängige Variable erzeugt werden, so die Prädiktoren für diese Situation vorliegen.

#### 3.3.1 Mathematische Grundlagen

Die Reaktion des Prädiktand auf ein Forcing, welches aus Veränderungen einer unabhängigen Variablen resultiert, wird über den Regressionskoeffizienten  $\beta_i$  quantifiziert. Jener Anteil der Variation von  $Y$ , der sich nicht über das Regressionsmodell erklären lässt, wird als Fehler betrachtet und findet sich im Residuum  $\epsilon$  wieder. Für diese Störvariablen wird eine zentrierte Normalverteilung angenommen und vorausgesetzt. Mit  $n$  Prädiktoren lässt sich das zugrundeliegende Modell dann wie folgt formulieren:

$$Y_j = \beta_0 + \sum_{i=1}^n \beta_i X_{ji} + \epsilon_j \quad (3.15)$$

Dieses unbekannte Modell ist abhängig von den zugrundeliegenden Einheiten. Die Regressionskonstante  $\beta_0$  hat die gleiche Einheit wie  $Y$ , für die übrigen Koeffizienten gilt  $[\beta_i] = [Y] / [X]$ . Für  $m$  Realisationen von  $Y$  stellt 3.15 ein Gleichungssystem dar, mit dem die unbekannt Parameter  $\beta_i$  bestimmt werden. Zu diesem Zweck wird eine Kombination von Regressionskoeffizienten gesucht, die den Prädiktand möglichst gut erklärt. Das übliche, wenn auch nicht einzig mögliche Kriterium (vgl. FAHRMEIR et al. 2009) ist die Minimierung der Fehlerquadratsumme  $SQE$  und somit eine Parameterschätzung nach der Methode der kleinsten Quadrate. Zur Darlegung des Vorgehens eignet sich in diesem Fall die Matrixnotation. Die  $m$  Elemente des Vektors  $\mathbf{y}$  stellen die Realisierungen des Prädiktand dar. Die Prädiktorenmatrix  $\mathbf{X}$  hat  $m \times (n + 1)$  Elemente.

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1i} & \dots & x_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{j1} & \dots & x_{ji} & \dots & x_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{m1} & \dots & x_{mi} & \dots & x_{mn} \end{pmatrix} \quad (3.16)$$

Die Spalten der Matrix beinhalten die einzelnen Prädiktoren, die erste Spalte repräsentiert dabei die Regressionskonstante. Im Vektor  $\mathbf{b}^T = (b_0, b_1, \dots, b_n)$  werden die Parameter des Modells zusammengefasst und der  $m$ -dimensionale Vektor  $\mathbf{e}$  enthält dessen stochastische Komponente. Das zu bestimmende Modell lässt sich nun kompakt schreiben als:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e} \quad (3.17)$$

Die  $e_j$  schwanken zufällig um den Schätzer des Prädiktand aus den Prädiktoren und Regressionskoeffizienten, dem deterministischen Teil des Modells. Mit  $\hat{\mathbf{y}} = \mathbf{Xb}$  erhält man die Definition dieser Modellfehler durch einfaches Umstellen von 3.17:

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{Xb} \quad (3.18)$$

Die Fehlerquadratsumme  $SQE$  des Modells ist das Skalarprodukt  $\mathbf{e}^T \mathbf{e}$ . Unter Verwendung von Gleichung 3.18 lassen sich dann folgende äquivalente Darstellungen dieser Größe schreiben:

$$\begin{aligned} \mathbf{e}^T \mathbf{e} &= (\mathbf{y} - \mathbf{Xb})^T (\mathbf{y} - \mathbf{Xb}) \\ &= \mathbf{y}^T \mathbf{y} - \mathbf{y}^T \mathbf{Xb} - \mathbf{b}^T \mathbf{X}^T \mathbf{y} + \mathbf{b}^T \mathbf{X}^T \mathbf{Xb} \end{aligned} \quad (3.19)$$

Durch den Ersatz des Skalars  $\mathbf{y}^T \mathbf{X} \mathbf{b}$  durch dessen Transponierte  $(\mathbf{y}^T \mathbf{X} \mathbf{b})^T = \mathbf{b}^T \mathbf{X}^T \mathbf{y}$  lässt sich 3.19 umformen zu:

$$\mathbf{e}^T \mathbf{e} = \mathbf{y}^T \mathbf{y} - 2\mathbf{b}^T \mathbf{X}^T \mathbf{y} + \mathbf{b}^T \mathbf{X}^T \mathbf{X} \mathbf{b} \quad (3.20)$$

Über Differentiation dieses Ausdrucks lässt sich das System der Normalgleichungen bestimmen.

$$\begin{aligned} \frac{\partial \mathbf{e}^T \mathbf{e}}{\partial \mathbf{b}} &= -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{0} \\ \Rightarrow \mathbf{X}^T \mathbf{y} &= \mathbf{X}^T \mathbf{X} \mathbf{b} \end{aligned} \quad (3.21)$$

Diese liefern schließlich ein lineares Gleichungssystem für die Schätzer  $b_i$  der Koeffizienten  $\beta_i$ .

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.22)$$

Bedingung für dessen Lösbarkeit ist, dass die Inverse von  $\mathbf{X}^T \mathbf{X}$  existiert. Dafür muss  $m > n$  und dabei die Kollinearität innerhalb des Datensatzes nicht zu ausgeprägt sein. Konkret darf keiner der Prädiktoren über eine Linearkombination der Übrigen darstellbar sein.

### 3.3.2 Modellgüte und Inferenz

Eine Grundlage für die Abschätzung der Qualität der Regression liefert das Bestimmtheitsmaß oder Determinationskoeffizient  $B$ . Hohe Werte bedeuten in diesem Fall, dass die Variation der  $\hat{y}_j$  die der  $y_j$  abbildet.  $B$  lässt sich mit dem – im Fall von mehr als einem Prädiktor multiplen – Korrelationskoeffizienten über  $B = r_{y\hat{y}}^2$  berechnen. Dieser wiederum lässt sich berechnen und veranschaulichen über bivariate Korrelation zwischen den  $y_j$  und den  $\hat{y}_j$ . Es ist nur zu beachten, dass die multiple Korrelation keine negativen Zusammenhänge beschreiben kann und somit gilt:

$$r_{y\hat{y}} = |r_{y\hat{y}}| \quad (3.23)$$

Häufig ist ein hoher Wert von  $B$  ein ausreichendes Gütekriterium für die Regression, insbesondere wenn diese nicht als eigentliche Analyse eingesetzt wird sondern nur eine Prozedur darstellt, um einen Datensatz für weitere Untersuchungen vorzubereiten. Beispielsweise wenn ein bekannter, aber für weitere Untersuchungen unwichtiger oder störender Einfluss einer Variable  $X$  auf  $Y$  entfernt werden soll. Insbesondere zur Filterung von Zeitreihendaten mit quasi-deterministischen Trends ist dieses Vorgehen üblich und wird im Regelfall ohne gesonderte Prüfung auf statistische Signifikanz des Trends durchgeführt.

Darüber hinaus existieren aber klassische Signifikanztests für die verschiedenen Fragestellungen, die sich bei Anwendung der Regressionsrechnung als Analysewerkzeug ergeben. Wie alle derartigen Verfahren beruhen diese auf bestimmten, teilweise nur schwer zu überprüfenden Annahmen oder Modellvoraussetzungen. Die Annahmen des klassischen linearen Modells betreffen die Verteilung der Residuen  $\epsilon$ . Für diese wird eine Normalverteilung unterstellt. Deren Parameter sind der Erwartungswert  $E(\epsilon) = \mathbf{0}$ , was eine zusätzliche Annahme, die Zentriertheit der  $\epsilon_j$ , bezeichnet. Für die Varianz setzt die Methode der kleinsten Quadrate in der beschriebenen Form Homoskedastizität, also Varianzgleichheit unabhängig von der konkreten Realisation der Prädiktorvariablen, voraus. Weiterhin sollen die Residuen echte Zufallsvariablen ohne Autokorrelation darstellen. Diese beiden Forderungen sind bei der Analyse von Geodaten normalerweise kaum zu überprüfen und häufig nicht erfüllt. Jedoch können verschiedene Filtertechniken eventuelle Verstöße gegen diese Annahmen ausgleichen. Formal lassen sich beide Forderungen, unter Verwendung der Einheitsmatrix  $\mathbf{I}$  und der Kovarianzmatrix  $\mathbf{\Sigma}$ , schreiben als:

$$\mathbf{\Sigma} = \sigma_\epsilon \mathbf{I} \quad (3.24)$$

Ob diese Modellannahmen erfüllt sind, ist immer dann entscheidend, wenn inhaltliche Aussagen getroffen werden. Relevant ist zunächst die Frage, ob die Kombination der Prädiktorvariablen im Regressionsmodell in der Lage ist, die Variabilität des Prädiktand ausreichend gut zu erklären, um als überzufällig zu gelten. Bei der Beurteilung dieser fundamentalen Frage wird der globale  $F$ -Test eingesetzt. Mit diesem wird die globale Nullhypothese:

$$H_0 : \beta_i = 0 \quad (3.25)$$

mit  $i = 1, n$  überprüft. Die Verwerfung dieser  $H_0$  bedeutet dann, dass mindestens einer der Regressionskoeffizienten von 0 verschieden ist. Eine andere Interpretation des globalen  $F$ -Tests wäre die Überprüfung des durch  $r_{y\mathbf{X}}$  geschätzten Korrelationskoeffizienten mit der  $H_0 : \rho = 0$ . Die Prüfgröße  $F$  errechnet sich unter Verwendung von  $B$  zu:

$$F = \frac{(m - n + 1)B}{n(1 - B)} \quad (3.26)$$

Unter der Bedingung, dass die Voraussetzungen der Regression erfüllt sind, kann der  $F$ -Test dann mit  $df_1 = n$  und  $df_2 = m - n + 1$  Freiheitsgraden durchgeführt werden. In dieser Arbeit werden keine Ergebnisse gezeigt, bei denen Tests auf die Signifikanz einzelner Prädiktoren durchgeführt wurden.

### 3.3.3 Polynomregression

Liegt zwischen dem Prädiktor und einem Prädiktor ein nichtlinearer Zusammenhang vor, so ist dessen Modellierung über die oben beschriebene Vorgehensweise üblicherweise zu ungenau, eventuell lässt sich sogar überhaupt kein Zusammenhang feststellen. Die Anwendung des beschriebenen Modells ist in diesem Fall dennoch möglich, wenn die Beziehung zwischen  $Y$  und  $X_i$  über geeignete Transformationen linearisiert wird. Ein wichtiger Sonderfall ist die Verwendung eines polynomialen Modells. Um dieses entsprechend dem beschriebenen Verfahren zu bewerkstelligen, muss lediglich die Matrix  $\mathbf{X}$  mit den Polynomen der Prädiktorvariable  $X$  besetzt sein. Damit wird die polynomiale Regression  $n$ -ten Grades formal zu einer multiplen mit  $n$  Prädiktoren  $X^k$ , wobei  $k = 1, \dots, n$ .

Solche Kurven werden hier zur Modellierung der Trendkomponenten von Zeitreihen verwendet (vgl. Kapitel 3.6.1). Sie eignen sich zur Beschreibung nichtlinearer Zusammenhänge, die sich im Kontext dieser Arbeit bei entscheidenden Größen des Klimasystems in Reaktion auf die anthropogene Emission von Treibhausgasen feststellen lassen. So lässt sich die Entwicklung der globalen Mitteltemperatur seit dem Beginn des 20. Jahrhunderts sehr gut durch einen kubischen Trend der Zeitreihe  $T_E(t)$  beschreiben.

## 3.4 ANOVA

Die Varianzanalyse, kurz ANOVA (analysis of variance), ist ein statistisches Werkzeug zur Untersuchung des Einflusses von einer oder mehreren unabhängigen Variablen  $X$  auf eine Variable  $Y$ . Im Unterschied zur Regression muss hierbei nur  $Y$  metrisch skaliert sein, die unabhängigen Daten können dagegen als kategoriale Variablen vorliegen. Das hauptsächliche Anwendungsgebiet dieser Methode ist die Auswertung von Experimenten. Ziel ist die Identifikation und Quantifizierung systematischer Auswirkungen eines oder mehrerer Markierungsmaterialien auf die Ausprägung von  $Y$  innerhalb einer Stichprobe. Diese Gruppierungsmerkmale werden als Faktoren oder Effekte bezeichnet. Hier werden verschiedene Varianten der zweifaktoriellen ANOVA eingesetzt.

### 3.4.1 Mathematische Durchführung

Das zugrundeliegende Modell berücksichtigt zur Erklärung von  $Y$  zwei Hauptfaktoren  $A$  und  $B$ . Deren Faktorstufen werden im folgenden mit  $i = 1, \dots, n_a$ , bzw.  $j = 1, \dots, n_b$  gekennzeichnet. Es wird nicht a priori angenommen, dass  $A$  und  $B$  unabhängig voneinander sind. Daher umfasst das Modell weiterhin die Interaktion  $C$ . Da die Daten in dieser Arbeit keinen balancierten Modellversuch darstellen, treten die Faktorstufenkombinationen unterschiedlich häufig auf, weshalb der Index

$k$  immer bis  $n_c(a, b)$  läuft. Dies entspricht einer Datenanordnung in einer Tabelle mit  $n_a$  Spalten,  $n_b$  Zeilen und einer variablen Anzahl von Elementen  $y_{ijk}$  in den Zellen. Diese zu erklärenden Daten lassen sich dann, in Anlehnung an die Regression, mithilfe der Effekte der Faktoren auf den entsprechenden Faktorstufen und ihrer Kombination sowie eines stochastischen Terms  $\epsilon_{ijk}$  darstellen, der die Abweichung von diesem Modell repräsentiert und mit  $NV(0, \sigma^2)$  angenommen wird. Das statistische Modell der ANOVA lautet somit

$$Y_{ijk} = \bar{\mu}_T + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \quad (3.27)$$

Die Schätzer für dieses Modell ergeben sich dann wie folgt: der Erwartungswert  $\bar{\mu}_T$  wird über den Gesamtmittelwert  $\bar{y}_T = \frac{1}{n_T} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \sum_{k=1}^{n_c(a,b)} y_{ijk}$  der  $n_T$  Daten geschätzt. Für die Effekte der Faktorstufen verwendet die ANOVA verschiedene Anomalien von  $\bar{y}_T$ , d.h.  $\alpha_i$  wird über  $\hat{a}_i = \bar{a}_i - \bar{y}_T$  und  $\beta_j$  über  $\hat{b}_j = \bar{b}_j - \bar{y}_T$  geschätzt. Der Effekt der Interaktion  $\gamma_{ij}$  aus der Kombination von  $A$  und  $B$  auf den entsprechenden Faktorstufen ist  $\hat{c}_{ij} = \bar{c}_{ij} - \hat{a}_i - \hat{b}_j + \bar{y}_T$ . Wenn diese Effekte nicht 0 sind, tragen sie zur Erklärung der Gesamtvariation, die über die Quadratsumme  $QS_T = \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \sum_{k=1}^{n_c(a,b)} (y_{ijk} - \bar{y}_T)^2$  gekennzeichnet ist, bei.

$$QS_T = QS_A + QS_B + QS_C + QS_E \quad (3.28)$$

Ob diese Quadratsummen statistisch signifikant sind, also die damit verbundenen Effekte tatsächlich varianzerzeugend sind, wird jeweils mithilfe eines  $F$ -Tests ermittelt. Als Prüfgrößen werden die Quotienten der mittleren Quadrate aus den beiden Faktoren und der Interaktion als Zähler und der Fehlerquadratsumme als Nenner gebildet. Die mittleren Quadrate ergeben sich aus der Division der Quadratsummen mit ihren Freiheitsgraden. Für  $QS_T$  entspricht  $MQ_T$  mit  $df_T = n_T - 1$  der Stichprobenvarianz von  $Y$ . Mit den mittleren Fehlerquadraten

$$MQ_E = \frac{QS_E}{df_E} = \frac{\sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \sum_{k=1}^{n_c(a,b)} (y_{ijk} - \bar{c}_{ij})^2}{n_T - n_a \cdot n_b} \quad (3.29)$$

und den jeweiligen Freiheitsgraden ergeben sich für die restlichen Größen auf der rechten Seite der Gleichung 3.28 unter Beibehaltung der Effektschreibweise die

Prüfgrößen

$$F_A = \frac{MQ_A}{MQ_E} = \frac{\frac{QS_A}{df_A}}{MQ_E} = \frac{\frac{\sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \sum_{k=1}^{n_c(a,b)} \hat{a}_i^2}{n_a - 1}}{MQ_E} \quad (3.30)$$

$$F_B = \frac{MQ_B}{MQ_E} = \frac{\frac{QS_B}{df_B}}{MQ_E} = \frac{\frac{\sum_{j=1}^{n_b} \sum_{i=1}^{n_a} \sum_{k=1}^{n_c(a,b)} \hat{b}_j^2}{n_b - 1}}{MQ_E} \quad (3.31)$$

$$F_C = \frac{MQ_C}{MQ_E} = \frac{\frac{QS_C}{df_C}}{MQ_E} = \frac{\frac{\sum_{k=1}^{n_c(a,b)} \sum_{j=1}^{n_b} \sum_{i=1}^{n_a} \hat{c}_{ij}^2}{(n_a - 1)(n_b - 1)}}{MQ_E} \quad (3.32)$$

Diese Tests prüfen die  $H_0 : \alpha_1 = \dots = \alpha_{n_a} = 0$ ,  $H_0 : \beta_1 = \dots = \beta_{n_b} = 0$  und  $H_0 : \gamma_{11} = \dots = \gamma_{n_a n_b} = 0$ . Für die durch die Effekte erklärten Anteile an der Gesamtvariation, die erklärte Varianz  $B$  je Effekt, werden die konservativen Schätzer

$$B_A = \frac{df_A}{df_T} \frac{MQ_A - MQ_E}{MQ_T} \quad (3.33)$$

$$B_B = \frac{df_B}{df_T} \frac{MQ_B - MQ_E}{MQ_T} \quad (3.34)$$

$$B_C = \frac{df_C}{df_T} \frac{MQ_C - MQ_E}{MQ_T} \quad (3.35)$$

verwendet. Dieses Vorgehen dient als Anpassung dafür, dass die Signifikanztests, die für die ANOVA in dieser Arbeit durchgeführt werden, nicht exakt sondern nur approximativ gültig sind. Grund hierfür ist der durch die Datenlage gegebene unbalancierte Versuchsaufbau mit unterschiedlich besetzten Faktorstufenkombinationen (STORCH und ZWIERS 1999).

### 3.4.2 Spektrale ANOVA

Als Modifikation der beschriebenen zweifaktoriellen Varianzanalyse wird eine hier als Spektrale ANOVA bezeichnete Variante genutzt. Hierbei handelt es sich um die wiederholte Ausführung der standardisierten Analyse, wobei die  $y_{ijk}$  jeweils die Realisationen von transienten Zeitreihen darstellen. Diese werden mit sukzessive anwachsenden Tiefpassfiltern nach der Methode des gleitenden Mittels geglättet (vgl. Kapitel 3.6.1). Dieses Vorgehen stellt bei hochvariablen Daten eine effektive Varianzreduktion dar. Da diese Art des Filterns die hochfrequente Variabilität aus den Zeitreihen entfernt, während die roten Komponenten erhalten bleiben, ermöglicht die Spektrale ANOVA die Detektion langfristiger Signale, die auf die Hauptfaktoren zurückzuführen sind und ohne Glättung vom Rauschen der simulierten natürlichen Variabilität verdeckt würden. Üblicherweise konvergieren die

durch die Haupteffekte und die Interaktion dieser beiden erklärten Varianzanteile mit zunehmender Glättung gegen einen festen Wert, der deren langfristige Auswirkungen realistisch quantifiziert. Mit der Glättung erfolgt eine Reduzierung der Freiheitsgrade, die entsprechend Formel 3.55 vorgenommen wird.

## 3.5 EOF-Analyse

Die grundlegende Konzeption der EOF-Analyse besteht darin, die Dimensionen eines Datensatzes zu reduzieren, wobei die Variabilität der Ausgangsdaten durch diesen reduzierten Datensatz möglichst vollständig erhalten bleiben soll. Dies ist immer dann möglich, wenn die metrisch skalierten Daten nicht unabhängig voneinander sind, also Korrelationen zwischen den Variablen bestehen. Es handelt sich um ein induktives Verfahren, das schlussendlich auf der Überführung der Ausgangsdaten in ein anderes Bezugssystem basiert.

### 3.5.1 Begriffe und Anwendung

Die EOF-Analyse ist von einer reichhaltigen Terminologie geprägt. Synonym für die hier gewählte Bezeichnung wird auch principle component(s) analysis, abgekürzt zu PCA oder Hauptkomponentenanalyse verwendet. Es existieren weitere Bezeichnungen, die allerdings häufig im Kern ein anderes statistisches Verfahren beschreiben. Teilweise ist dies darauf zurückzuführen, dass viele Entwicklungen in Theorie und Anwendung dieser Methodik parallel im Bereich der Atmosphärenwissenschaft und der multivariaten Statistik erfolgten. Detaillierte Darstellungen aus der ersteren Perspektive finden sich bei MOBLEY und PREISENDORFER (1988) und aktueller bei HANNACHI et al. (2007). Aber auch das allgemeine Buch von JOLLIFFE (2002) hebt die Bedeutung der EOF-Analyse für die Klimatologie und Meteorologie, und umgekehrt, hervor. Hieran lässt sich durchaus die Stellung dieser Methode in der klimatologischen Forschung verdeutlichen.

Eine Folge dieser begrifflichen Vielfalt ist es, dass nicht alle verwendeten Termini eindeutig sind. Daher erfolgt die Erläuterung hier stärker mit Bezug zur tatsächlichen Anwendung in dieser Arbeit als bei den anderen Methoden in diesem Kapitel. In der hier verwendeten Form der s-modalen EOF-Analyse ist es ganz grundsätzlich das Ziel, die Variabilität einer raumzeitlichen Größe derart zu zerlegen, dass Zeit und Raum getrennt in Variablen gefasst werden können.

$$X(t, s) = \sum_{i=1}^m c_i(t)v_i(s) \quad (3.36)$$

Die  $i = 1, \dots, m$  Zeitreihen  $c_i(t)$  werden als Hauptkomponenten oder Prinzipale Komponenten, kurz PCs, bezeichnet. Sie stellen die Anregungskoeffizienten der

Raummuster  $v_i(s)$  dar. Diese werden im Folgenden als Empirische Orthogonalfunktionen oder EOFs angesprochen. Die EOFs bilden ein neues Koordinatensystem, das die ursprünglichen Daten enthält. In diesem stellen die PCs dann neue, sogenannte latente Variablen dar, die aus Linearkombinationen der Ausgangsdaten hervorgehen.

Die EOF-Analyse hat zum Ziel, die Variation der Ausgangsdaten auf möglichst effiziente Weise zu erfassen. Die EOF/PC-Paare werden daher derart konstruiert, dass für die paarweisen Korrelationen sowohl der  $v_i(s)$  als auch der  $c_i(t)$  jeweils gilt  $r = 0$ . Dagegen bilden klimatologische Felddaten typischerweise Phänomene mit großer räumlicher Ausdehnung ab. Es bestehen daher zwischen den verschiedenen Zeitreihen  $x(t, s)$  normalerweise bedeutende Korrelationen, womit innerhalb von  $X(t, s)$  redundante Informationen vorhanden sind. Die EOF-Analyse ermöglicht es in diesem Fall, das raumzeitliche Datenfeld auf ein bedeutend kleineres zu reduzieren, indem große Teile der Variabilität von  $X(t, s)$  oder äquivalent  $X'(t, s)$  mit einigen wenigen PCs erfasst werden. Neben dieser rein mathematischen Reduktion der Dimension besteht teilweise die Möglichkeit, die EOF/PC-Paare als Repräsentanten von physikalischen Phänomenen zu identifizieren. Somit stellen die PCs nicht zwangsläufig nur nach einem bestimmten Schema transformierte Daten dar, sondern können tatsächlich als Variablen betrachtet werden. In klimatologischen Daten lassen sich auf diese Weise bevorzugte Moden der Klimavariabilität, Eigenmoden oder Zirkulationsmuster identifizieren. Zu diesem Zweck wird die EOF-Analyse in dieser Arbeit genutzt (vgl. Kapitel 5.1). Im Folgenden wird die Konstruktion von EOFs und PCs knapp erläutert.

### 3.5.2 Grundgleichungen

Da die EOF-Analyse das Ziel hat, die Variabilität von  $X$  zu erfassen und der Mittelwert hierzu nichts beiträgt, werden im Folgenden statt den  $s = 1, 2, \dots, m$  Zeitreihen  $x(t, s)$  jeweils die Anomalien  $x'(t, s)$  betrachtet. Die Verwendung standardisierter Zeitreihen  $\tilde{x}(t, s)$  ist ebenfalls möglich, wenn auf die physikalische Einheit von  $X$  verzichtet werden kann oder, etwa bei der Analyse mehrerer Feldgrößen zugleich, muss.

Diese Zeitreihen werden dann in der Ausgangsdatenmatrix  $\mathbf{X}$  zusammengefasst. Sie bilden dort die  $n$  Spaltenvektoren  $\mathbf{x}_i$  mit jeweils  $m$  Elementen, die den Zeitschritten entsprechen. Die Kovarianzmatrix  $\Sigma$  ergibt sich aus:

$$\Sigma = n^{-1} \mathbf{X}^t \mathbf{X} \quad (3.37)$$

Bilden  $z$ -transformierte Daten die Ausgangsmatrix, so handelt es sich um die Korrelationsmatrix. Die Diagonalelemente von  $\Sigma$  sind die Varianzen der  $\mathbf{x}_i$ . Die EOF-Analyse identifiziert nun einen Vektor  $\mathbf{v}_1$ , der den größten Anteil der Varianz von  $\mathbf{X}$  erklärt und eine Linearkombination der Ausgangsdaten darstellt. Dies

entspricht der Maximierung der Varianz der Projektion von  $\mathbf{X}$  auf  $\mathbf{v}_1$ :

$$s_{\mathbf{X}\mathbf{v}_1}^2 = n^{-1}(\mathbf{X}\mathbf{v}_1)^t(\mathbf{X}\mathbf{v}_1) = \mathbf{v}_1^t \Sigma \mathbf{v}_1 \quad (3.38)$$

Entsprechend muss zur Bestimmung des gesuchten Eigenvektors  $\mathbf{v}_1$  der letzte Ausdruck maximiert werden. Er ergibt sich dann über die Lösung des Eigenwertproblems, mit  $i = 1$ :

$$\Sigma \mathbf{v}_i = \lambda_i \mathbf{v}_i \quad (3.39)$$

Der Eigenvektor, der diese Forderung erfüllt, ist definiert als die 1. EOF. Nachdem auf diese Weise  $\mathbf{v}_1$  bestimmt ist, wird das gleiche Verfahren zur Identifikation von weiteren Eigenvektoren angewendet. Hierbei wird dann derjenige von  $\mathbf{v}_1$  linear unabhängige Eigenvektor  $\mathbf{v}_2$  gesucht, der den größten Anteil am Residuum der Varianz der Ausgangsdaten erklärt. Entsprechend ist die  $i$ -te EOF die Lösung von Gleichung 3.39 zum  $i$ -ten Eigenwert  $\lambda_i$ . Der durch  $\mathbf{v}_i$  erklärte Anteil der Varianz ergibt sich über

$$Var(i) = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} \cdot 100\% \quad (3.40)$$

Die  $\lambda_i$  werden ihrer Größe entsprechend absteigend geordnet. Zur Vereinfachung der folgenden Ausführungen werden nur normierte Vektoren mit Einheitslänge  $|\mathbf{v}| = 1$  berücksichtigt. Somit ergibt sich, dass die orthogonalen Eigenvektoren eine Orthonormalbasis bilden:

$$\mathbf{v}_i^t \mathbf{v}_{i'} = \begin{cases} 1 & \text{für } i = i' \\ 0 & \text{für } i \neq i' \end{cases} \quad (3.41)$$

Dieses Muster kann solange fortgeführt werden bis die gesamte Varianz von  $\mathbf{X}$  erfasst ist. Häufig wird hierauf allerdings verzichtet, insbesondere wenn die ersten  $p$  EOF mit  $p \ll m$  bereits große Anteile der Varianz von  $\mathbf{X}$  erklären.

Aufgrund der Definition von  $\mathbf{X}$  handelt es sich bei den EOFs um räumliche Muster, die als Karte dargestellt werden können. Ihre Elemente  $v_i(s)$  sind die Ladungen dieser Muster. Zur vollständigen Zerlegung der Varianz von  $X(t, s)$  nach Gleichung 3.36 fehlen nun noch die Anregungskoeffizienten. Die  $i$ -te PC ergibt sich aus der Projektion der Ausgangsdaten auf die entsprechende EOF:

$$\mathbf{c}_i = \mathbf{X}\mathbf{v}_i \quad (3.42)$$

Für diese PCs gilt  $s_{\mathbf{c}_i}^2 = \lambda_i$ . Weiterhin sind sie paarweise unkorreliert. In der  $s$ -modalen Variante stellen die PCs den dynamischen Part dar. Sie regen die, einmal gefundenen starren, EOFs entsprechend der zeitlichen Variation im Datenfeld an. Das

bedeutet auch, dass sich die zeitliche Veränderung von räumlichen Mustern der Variabilität ebenfalls in den Amplituden der PCs, nicht in veränderten Ladungen der EOFs abbildet.

### 3.5.3 Varimax-Rotation

Die lineare Unabhängigkeit der EOFs und PCs ermöglicht die effektive Abbildung der Varianz. Allerdings bedeutet dies auch, dass die höheren EOFs zunehmend von der Voraussetzung der Orthogonalität geprägt und beeinflusst sind. Das kann darin resultieren, dass tatsächlich vorhandene physikalische Phänomene nicht mehr über die EOF-Analyse erfasst werden können, da diese nicht vollkommen linear unabhängig voneinander sind. Angesichts der komplexen Kopplungen im Klimasystem kann vielfach von dieser Situation ausgegangen werden. Zudem erweist sich die EOF-Analyse bei diesem Unterfangen durchaus als empfindlich gegenüber der betrachteten Datendomäne. Sollen also reale Variabilitätsmoden stabil erfasst werden, kann dies durch die Struktur der EOFs selbst erschwert werden. Es existiert eine Vielzahl von Ansätzen, um diese rein mathematischen Einschränkungen des wissenschaftlichen Nutzens der EOF-Analyse zu überwinden und die Interpretation von EOFs zu erleichtern. Ziel ist jeweils die Erzeugung einer klaren Einfachstruktur für die EOF/PC-Paare. Hierfür werden, in unterschiedlichem Umfang, die starren Orthogonalitätsvoraussetzungen der EOF-Analyse zumindest teilweise aufgegeben. Welches Kriterium hierbei ausschlaggebend dafür ist, ob es sich um eine Einfachstruktur handelt, unterscheidet sich bei zwischen den diversen Verfahren deutlich. Letztlich handelt es sich um die subjektive Entscheidung für ein Verfahren (siehe RICHMAN 1986).

Die Herbeiführung der Einfachstruktur wird als Rotation bezeichnet. Entsprechend werden die so erzeugten Raummuster hier zur Abgrenzung als REOFs und die Zeitreihen der Anregungskoeffizienten dieser Muster als RPCs bezeichnet, so eine Abgrenzung zu den normalen EOF/PC-Paaren notwendig erscheint. Für ein Set von  $p$  Eigenvektoren  $\mathbf{v}_i$  wird hierzu eine Rotationsmatrix  $\mathbf{R}$  bestimmt, um die  $p$  REOFs  $\mathbf{q}_i$  zu erzeugen:

$$\mathbf{Q} = \mathbf{R}\mathbf{V} \quad (3.43)$$

Wieviele EOFs rotiert werden, wird in dieser Arbeit davon abhängig gemacht, dass mindestens 80 % der Varianz der Ausgangsdaten erklärt werden. Als konkrete Methode wird die weitverbreitete Variante der Varimax-Rotation verwendet. Die Einfachstruktur wird hier auf die Weise erzeugt, dass das Kriterium  $Q$ , die Varianz der quadrierten Ladungen der rotierten Eigenvektoren  $\mathbf{q}_i$ , maximiert wird.

$$Q = \sum_{i=1}^p s_{*i}^2 = \sum_{i=1}^p m^{-1} \sum_{s=1}^m ((q_i(s))^2 - \overline{(q_i(s))^2})^2 \quad (3.44)$$

Für eine einzelne REOF lässt sich diese spezielle Form der Varianz mit dem Verschiebungssatz schreiben als

$$s_{*i}^2 = \frac{1}{m} \sum_{s=1}^m (q_i(s))^4 - \frac{1}{m^2} \left( \sum_{s=1}^m (q_i(s))^2 \right)^2 \quad (3.45)$$

**R** wird iterativ über die paarweise Maximierung der Varianzsummen zweier REOFs entwickelt. Idealerweise werden so REOFs erzeugt mit entweder sehr hohen Beträgen von  $q_i(s)$  oder aber solchen gleich oder nahe 0. Bei der Varimax-Methode handelt es sich um eine orthogonale Rotation. Daher bilden die REOFs, wie die unrotierten Eigenvektoren, weiterhin eine Orthonormalbasis. Die RPCs, die nach Gleichung 3.42 über die Projektion von **X** auf die REOFs gewonnen werden, sind dagegen nicht mehr unkorreliert. Entsprechend ihrer Nutzung in dieser Arbeit rückt die Rotation zur Identifikation von realen physikalischen Phänomenen die EOF-Analyse näher an die Faktorenanalyse, wenn auch auf die problematische Kommunalitäten Schätzung verzichtet wird.

### 3.5.4 Verwendete Skalierung und Darstellung

Neben der Terminologie und den Verfahren zur Vereinfachung der gewonnenen Muster ist auch bei der Skalierung von (R)PCs und (R)EOFs, insbesondere für die graphische Darstellung, eine Vielzahl von Praktiken in Gebrauch. Ein Grund hierfür ist, dass jedes beliebige skalare Vielfache des Eigenvektors  $\mathbf{v}_i$  zusammen mit  $\lambda_i$  eine Lösung für Gleichung 3.39 liefert. Daneben gibt es aber auch inhaltliche Gründe die zur Diversifizierung beitragen. Für eine einfache und eindeutige Interpretation von Ergebnissen der EOF-Analyse ist es daher notwendig, das hier verwendete Vorgehen kurz zu erläutern und begründen. Es ergibt sich aus der konkreten Anwendung der EOF-Analyse in der s-modalen Variante.

Da die hier untersuchten Datensätze allesamt auf mehr oder weniger regelmäßigen Gradnetzen vorliegen, repräsentieren die Gitterpunkte unterschiedliche Flächenanteile (siehe Kapitel 3.1). Diese resultierende Verteilung der Daten beeinflusst, insbesondere bei der Berücksichtigung der globalen Domäne, die EOFs, da die tropischen Regionen unterrepräsentiert sind. Um dies zu verhindern erfolgt die Multiplikation der  $n$   $\mathbf{x}_s$  mit dem passenden Gewichtungsfaktor

$$\omega(s) = \sqrt{\frac{\cos(s)}{\sum_{s=1}^n \cos(s)}} \quad (3.46)$$

Die Wurzelbildung ist notwendig, da die Gewichtung für die Varianz der  $\mathbf{x}_s$  gilt, die in die Kovarianzmatrix eingeht. Bei der EOF-Analyse der Korrelationsmatrix wird entsprechend  $\tilde{\mathbf{x}}_s$  mit  $\omega_s$  gewichtet. Neben unregelmäßigen Flächenanteilen können auch andere physikalische oder inhaltliche Gründe für eine unterschiedliche

Gewichtung der Daten sprechen. In jedem Fall ist diese nach obigem Muster durchzuführen.

Die tatsächliche Extraktion der EOFs und PCs erfolgt gemäß der Konvention, dass jeder Eigenvektor  $\mathbf{v}_i$  die Länge 1 hat. Somit entspricht die Varianz der PC  $\mathbf{c}_i$  dem Eigenwert  $\lambda_i$ . Dies liefert jedoch schwer interpretierbare Raummuster, die zudem wesentlich von der, bei den hier verwendeten Daten notwendigen, Gewichtung der Kovarianzmatrix beeinflusst werden. Da die physikalische Interpretation ein wichtiger Aspekt bei der Nutzung der EOF-Analyse ist, wird von diesen, mathematisch korrekten und gebräuchlichen, Skalierungen Abstand genommen.

Stattdessen wird eine in der Klimatologie verbreitete Konvention angewendet. Diese zielt zunächst darauf ab, die Gewichtungsfaktoren nach Gleichung 7.2 in den  $\mathbf{e}_i$  wieder zu beseitigen. Solche EOFs zeigen dann in der kartographischen Darstellung nicht mehr die tatsächlichen Ladungen  $v_i(s)$  der  $\mathbf{v}_i$  an einem Gitterpunkt  $s$ , sondern die effektive, lokale Variation der Feldgröße  $X(t, s)$  bzw.  $\tilde{X}(t, s)$ , verbunden mit der Variation der PC. Diese Art der Darstellung erleichtert die Erfassung der unterschiedlichen räumlichen Schwerpunkte, die ein EOF/PC-Paar aufweist und somit ihre Interpretation. Es handelt sich also bei diesen Karten strenggenommen nicht mehr um Eigenvektoren oder EOFs, auch wenn sie gelegentlich als solche angesprochen werden, sondern um Darstellungen der Kovariabilität der Ausgangsdaten und der PCs (z. B. ZHANG et al. 1997, THOMPSON und WALLACE 1998, 2000). Als Grundlage einer solchen Darstellung ist die  $z$ -Transformation der  $c_i$  das einfachste Mittel. Da in dieser Arbeit viele über EOF-Analyse gewonnene Zeitreihen aus verschiedenen Datensätzen verglichen werden, bietet sich dieses Vorgehen schon deswegen an, um die Indexentwicklungen zuverlässig vergleichen zu können. Die Ladungen  $\hat{v}_i(s)$  ergeben sich dann nach Formel 3.11 als:

$$\hat{v}_i(s) = s_{\mathbf{x}_s} \tilde{\mathbf{c}}_i = \overline{x'(t, s) \tilde{c}(t)} \quad (3.47)$$

Eine einsichtsreiche äquivalente Interpretation ist es, die  $\hat{v}_i(s)$  mit den Regressionskoeffizienten  $\beta_1(s)$  gleichzusetzen, die sich durch bivariate Regression der Ausgangsdaten  $X(t, s)$  auf  $\tilde{c}_i(t)$  ergeben, und somit als  $\Delta[X]/s_{\tilde{c}}$ .

### 3.6 Zeitreihentechniken

Zeitreihen sind dadurch gekennzeichnet, dass eine Größe  $Y$  zu unterschiedlichen Zeiten, normalerweise diskreten Zeitpunkten  $t$ , realisiert ist. Dies ist bei den in dieser Arbeit untersuchten, wie bei nahezu allen klimatologischen, Daten der Fall. Dem wurde bereits bei einigen Erläuterungen in der Notation Rechnung getragen. Grundsätzlich sind solche  $Y(t)$  und Zufallsvariablen nicht dasselbe, da erstere eben nicht nur eine Funktion des Zufalls sind, sondern auch eine der Zeit. Für die statistische Analyse von Zeitreihen bedeutet dies, dass häufig Anpassungen

der statistischen Modelle vorgenommen werden müssen, um auch mit solchen Daten  $Y(t)$  zu belastbaren Aussagen zu gelangen. Daneben besteht jedoch die Möglichkeit, von den Ausprägungen von  $Y(t)$  auf die diesen zugrundeliegenden Prozesse und ihre physikalischen Charakteristika zu schließen. Insbesondere zu diesem Zweck existieren spezielle Methoden, jedoch werden auch die zuvor erläuterten Methoden für die Analyse von Zeitreihen herangezogen. Mit klimatologischem Fokus bietet MUDELSEE (2010) einen aktuellen Überblick über die vielfältigen Aspekte der Zeitreihenanalyse. Hier finden nur wenige, durchaus aber grundlegende, Techniken aus diesem Gebiet Anwendung, die im Folgenden erläutert werden.

### 3.6.1 Additives Zeitreihenmodell

Einen üblichen und praktischen Ausgangspunkt für die Analyse von Zeitreihen bildet das additive Zeitreihenmodell. Hierbei wird eine konkrete metrische Ausprägung von  $Y(t)$  als Summe verschiedenster deterministischer Einflussfaktoren  $\kappa_i(t)$ , die kontinuierlich oder nur sporadisch wirksam sein können, und einer stochastischen Komponente  $\epsilon(t)$  betrachtet.

$$Y(t) = \sum_{i=1}^{\infty} \kappa_i(t) + \epsilon(t) \quad (3.48)$$

Somit handelt es sich zunächst um ein konzeptionelles Modell, das die Möglichkeit zur gezielten und zielführenden Manipulation von Zeitreihen eröffnet. In der Klimatologie haben die  $\kappa_i(t)$  meist physikalische Ursachen. Daher geht ihre Identifikation üblicherweise mit einem wachsenden Prozessverständnis einher, so dass schon dieses einen wichtigen Aspekt der statistischen Klimatologie darstellt. Da diese  $\kappa_i$  die Verteilungsfunktion der  $Y(i)$  beeinflussen, müssen sie in der Regel vor der Durchführung weiterer statistischer Analysen entfernt werden, so sie für diese von keiner inhaltlichen Bedeutung sind.

In dieser Arbeit, wie im Kontext des anthropogenen Klimawandels allgemein, weisen viele Zeitreihen eine niederfrequente Komponente auf, die das langfristige Verhalten von  $Y(t)$  dominiert, den Trend. Zwar existieren verschiedene Ansätze, die Trendkomponente zu bestimmen, die üblichste und auch hier praktizierte Methode ist es jedoch, diese über eine Regression als Funktion der Zeit zu schätzen. Hierbei können, je nach der Struktur des Trends, unterschiedliche Polynomgrade der Regression zugrunde gelegt werden. Das additive Modell erlaubt dann die einfache Extraktion des auf den Trend  $\hat{Y}(t)$  zurückzuführenden Anteils der Variation von  $Y(t)$  über

$$\check{Y}(t) = Y(t) - \hat{Y}(t) \quad (3.49)$$

Korrelationsrechnungen und andere Analysen liefern vielfach erst nach dieser

Operation sinnvolle und verlässliche Ergebnisse, jedoch können sich auch solche Residuen  $\check{Y}(t)$  als nicht-stationär (vgl. Kapitel 3.6.2) erweisen, wenn das Regressionsmodell die wahre Trendkomponente von  $Y(t)$  nicht ausreichend genau nachgebildet. In diesem Fall kann die Trendkomponente auch nach der Methode des gleitenden Durchschnitts gebildet werden.

$$\hat{Y}(t) = \frac{1}{m} \sum_{k=t-\frac{m-1}{2}}^{t+\frac{m-1}{2}} Y(k) \quad \text{für } m = 1, 3, \dots \quad (3.50)$$

Dieses Vorgehen wird hier nicht angewandt, allerdings stellt ein so berechneter Trend zugleich eine gefilterte Version der Zeitreihe  $Y(t)$  dar. Von diesem Tiefpassfilter wird in den folgenden Kapiteln häufig Gebrauch gemacht, wenn es gilt, die hochfrequenten Varianzanteile von Zeitreihen zu unterdrücken, um auf diese Weise das Signal-Rausch Verhältnis von Zeitreihen zugunsten des Signals zu verändern.

Weitere bedeutende Komponenten klimatologischer Zeitreihen haben häufig astronomische Ursachen, wie etwa der Jahres- oder auch der, für die hier verwendeten Daten allerdings nicht relevante, Tagesgang. Der Jahresgang verschiedener Klimatelemente und -größen dagegen ist, insbesondere bei der Bestimmung der Eigenmoden des Klimasystems, hinderlich und muss daher entfernt werden. Eine Vorgehensweise hierfür ist es, diese hochfrequente Störgröße über Mittelung zu beseitigen, indem die betreffenden Analysen nur für bestimmte Jahreszeiten oder aber Jahresmittel durchgeführt werden. Alternativ hierzu werden, so denn Monatsdaten benötigt werden, die Zeitreihen über

$$Y'_i(t) = Y_i(t) - \bar{Y}_i \quad \text{mit } i = 1, 2, \dots, 12 \quad (3.51)$$

in monatliche Anomalien  $Y'(t)$  umgewandelt. Über die Entfernung solcherart identifizierbarer Komponenten können Zeitreihen in stationäre Prozesse umgewandelt werden.

### 3.6.2 Kenngrößen stationärer Zeitreihen

Hier werden spezielle Momente von Zeitreihen in der Zeitdomäne und Frequenzdomäne und ihre Identifikation in den Daten erläutert. Die hier verwendeten, nicht-parametrischen Methoden gehen jeweils davon aus, dass die Zeitreihen die Realisationen von ergodischen, stationären Prozessen darstellen.

Als Merkmal des hier verwendeten schwachen Stationaritätsbegriffs gilt, dass die Erwartungswerte  $E(Y(t))$  und  $E((Y(t) - E(Y(t)))(Y(t+h) - E(Y(t+h))))$  nicht vom konkreten Zeitpunkt  $t$  abhängen. Somit ist in einem solchen stationären Prozess  $\mu$  konstant und die Autokovarianz  $\gamma(h)$  nur eine Funktion des lags  $h$ , dem Abstand zwischen den Zeitpunkten. Die Autokovarianzfunktion ist gerade, somit gilt  $\gamma(-h) = \gamma(h)$ . Sie gilt für stetige Funktionen, hier werden aber nur deren

Realisierungen in Form von diskreten Zeitreihen betrachtet.

### 3.6.2.1 Auto- und Kreuzkorrelation

Der Schätzer für  $\gamma(h)$  ist die empirische Autokovarianzfunktion  $c(h)$ . Für  $n$  Zeitschritte  $t$  erfolgt die Berechnung zum lag  $h$  über:

$$c(h) = \frac{1}{n} \sum_{t=1^{*}}^{n^{*}} x'(t+h)x'(t) \quad \text{mit} \quad 1^{*}, n^{*} = \begin{cases} 1, n-h & \text{für } h \geq 0 \\ 1+|h|, n & \text{für } h < 0 \end{cases} \quad (3.52)$$

Wegen der vorausgesetzten und entsprechend herbeigeführten Stationarität werden die Anomalien jeweils zum arithmetischen Mittel  $\bar{x} = \sum_{t=1}^n x(t)$  gebildet. Somit ist  $c(h)$  zwar nicht mehr erwartungstreu, dafür wird aber die Varianz insbesondere bei größeren lags nicht künstlich erhöht. Für die Zeitreihenanalyse wird diese Abwägung üblicherweise zugunsten des geringen Bias entschieden (STORCH und ZWIERS 1999). Über Normierung von  $c(h)$  mit der Varianz der Zeitreihe ergibt sich dann mit  $r_A(h)$  ein Schätzer für die Autokorrelationsfunktion  $\rho_A(h)$ .

$$r_A(h) = \frac{c(h)}{c(0)} \quad (3.53)$$

Sie ist wie  $c(h)$  ebenfalls gerade und zusätzlich wie der gewöhnliche Korrelationskoeffizient normiert, so dass  $-1 \leq r(h) \leq 1$ . Die Autokorrelation von Zeitreihen wird in dieser Arbeit vor allem dafür genutzt, die Freiheitsgrade für statistische Tests von Zeitreihendaten anzupassen. Auch wenn Trend- und Saisonkomponente entfernt sind, können sich signifikante Werte von  $r_A(h)$  bzw.  $c(h)$  auf verschiedene Analysen störend auswirken. Dies gilt vor allem bei der, häufig vorgenommenen, Berechnung der Korrelation zwischen zwei Zeitreihen. Um die hierbei auftretenden Korrelationskoeffizienten sinnvoll interpretieren und beurteilen zu können, werden bei dieser Form der Korrelation die Freiheitsgrade für den Signifikanztest angepasst, also reduziert. Dadurch ändert sich der konkrete Wert von  $r$  nicht, es wird jedoch die Verwerfung der  $H_0$  erschwert. Bei der bivariaten Korrelation von  $a(t)$  und  $b(t)$  werden hierfür jeweils über

$$df_r = df \cdot \frac{1 - r_{A,a}(1)r_{A,b}(1)}{1 + r_{A,a}(1)r_{A,b}(1)} \quad (3.54)$$

die Autokorrelationen der Zeitreihen zum lag  $h = 1$  berücksichtigt (SCHÖNWIESE 2006). Müssen die Freiheitsgrade einer einzelnen Zeitreihe geschätzt werden, etwa bei Mittelwertvergleichen, so wird der effektive Datenumfang über

$$df_r = df \cdot \frac{1 - r_A(1)}{1 + r_A(1)} \quad (3.55)$$

etwas konservativer angepasst (BRETHERTON et al. 1999). In beiden Fällen werden die Zeitreihen als rotes Rauschen betrachtet. Für bivariate Zeitreihen, mit  $a(t)$  und  $b(t)$  als Realisationen stationärer Prozesse  $Y_1(t)$  und  $Y_2(t)$ , existiert die Kreuzkorrelationsfunktion  $\rho_{K,ab}(h)$ . Diese beschreibt den linearen Zusammenhang zweier Prozesse, zu einem zeitlichen Versatz zum lag  $h$ . Empirisch ermittelt sie sich über

$$r_{K,ab}(h) = \frac{s_{K,ab}(h)}{s_a s_b} \quad (3.56)$$

Im Zähler steht hier die Kreuzkovarianzfunktion  $s_{K,ab}(h) = \frac{1}{n} \sum_{i=1}^{n*} a'(t+h)b'(t)$ , wobei die Notation entsprechend Formel 3.52 verwendet wird. Der Nenner ist das Produkt Standardabweichungen der stationären Zeitreihen. Es gilt  $r_K(0) = r$ , jedoch muss die normale Korrelation nicht den maximalen Wert der Kreuzkorrelationsfunktion darstellen. Die Kreuzkorrelation ist weder eine gerade Funktion, noch hat sie andere Symmetrieeigenschaften. Daher ist es in diesem Fall wichtig, die Reihenfolge der Variablen zu beachten. Hier wird das lag jeweils auf die zuerst aufgeführte Zeitreihe bezogen. Wenn  $h < 0$ , bedeutet dies in Formel 3.56, dass  $Y_1(t)$  die Zeitreihe  $Y_2(t)$  führt, bei positiven  $h$  beeinflussen die Realisationen  $b(t_1)$  dagegen  $a(t_2)$  mit  $t_1 < t_2$ . Wenn  $r_{K,ab}$  eine tatsächliche physikalische Beziehung erfasst und nicht nur eine Scheinkorrelation beschreibt, so ist davon auszugehen, dass die führende Komponente das Agens, die folgende das Reagenz des bivariaten Prozesses darstellt.

### 3.6.2.2 Periodogramm und Spektralschätzer

Ausführlich werden spektrale Verfahren bei KAY (1988) und mit Konzentration auf analytische Anwendungen in der Geographie und Klimatologie bei RAYNER (1971) behandelt. Dieses weite Feld wird hier nur in einem Aspekt genutzt, der Spektralen Varianzanalyse von Zeitreihen. Einen themenbezogenen Überblick über diese und weitere verwandte und weiterführende Methoden geben GHIL et al. (2002). Bei diesen Methoden, auch der grundlegenden Spektralanalyse, werden Zeitreihen in die Frequenzdomäne überführt, um ein tieferes Verständnis für die zugrundeliegenden Prozesse und deren Charakteristik zu erhalten. Dementsprechend wird die Zeitreihe als Überlagerung von harmonischen Schwingungen mit unterschiedlichen Frequenzen oder Perioden betrachtet, die in der Summe ihre Ausprägung in der Zeitdomäne hervorrufen. Wird der Prozess den  $Y(t)$  abbildet von einer oder mehreren Oszillationen auf speziellen Zeitskalen dominiert, so lässt sich dies anhand der Spektraldichtefunktion, auch einfach Spektrum, identifizieren.

Ausgangspunkt für die Bestimmung des Spektrums ist die Diskrete Fouriertransformation von  $Y(t)$ , die diese in harmonische Schwingungen umwandelt. Dies entspricht der Überführung der Zeitreihe mit  $n$  Realisationen in ein neues Orthogonalsystem mit den  $n$  Funktionen  $\frac{1}{2}$ ,  $\cos(\omega_j t)$  mit  $j = 1, \dots, \frac{n}{2}$  und  $\sin(\omega_j t)$  mit  $j =$

$1, \dots, \frac{n}{2} - 1$ . Anstelle der Kreisfrequenzen  $\omega_k = \frac{2\pi k}{n}$  mit  $k = 0, 1, \dots, \frac{n}{2}$  können äquivalent die Frequenzen  $\nu_k = \frac{k}{n}$  oder die Perioden  $P_k = \frac{n}{k}$  genutzt werden. In dieser Repräsentation lassen sich die Elemente von  $Y(t)$  schreiben als

$$y(t) = \frac{1}{2}a_0 + \sum_{j=1}^{\frac{n}{2}} a_j \cos(\omega_j t) + \sum_{j=1}^{\frac{n}{2}} b_j \sin(\omega_j t) \quad (3.57)$$

Die insgesamt  $n$  Fourierkoeffizienten  $a_j$  und  $b_j$  können als Regressionskoeffizienten über einen linearen Ansatz geschätzt werden (vgl. Kapitel 3.3). Da die hier untersuchten Zeitreihen äquidistant sind, lassen sich die Koeffizienten auch über ein weniger aufwendiges Verfahren bestimmen. Dann gilt  $a_0 = 2\bar{y}$ ,  $a_{n/2} = n^{-1} \sum_{t=1}^n (-1)^t y(t)$  und, weil ausschließlich gerade  $n$  vorkommen,  $b_{n/2} = 0$ . Für  $j = 1, \dots, \frac{n}{2} - 1$  sind die restlichen Koeffizienten  $a_j$  und  $b_j$  aufgrund der Orthogonalität der harmonischen Funktionen jeweils unabhängig über lineare Regression mit zwei Prädiktoren  $x_1(t) = \cos(\omega_j t)$  und  $x_2(t) = \sin(\omega_j t)$  zu ermitteln. Es handelt sich hierbei zunächst um eine reine Transformation ohne Informationsverlust oder -zugewinn (vgl. BUTZ 2009). Mit diesen Koeffizienten lässt sich nun die Varianz von  $Y(t)$  zerlegen. Sie ergibt sich aus:

$$s^2 = \frac{1}{2} \sum_{j=1}^{\frac{n}{2}} a_j^2 + b_j^2 \quad (3.58)$$

In dieser Darstellung lässt sich erschließen, wie viel Variabilität von  $Y(t)$  auf die Schwingungen der einzelnen Frequenzen zurückgeführt werden kann. Aus Gleichung 3.58 ergibt sich auch eine Überführung in die relativen Anteile der erklärten Varianz pro Schwingung. Die Summanden in der Varianzzerlegung, multipliziert mit dem Faktor  $\frac{n}{4}$ , bilden das Periodogramm  $P(\omega_j)$ , das die Grundlage für die Schätzung der Spektraldichtefunktion bildet. Ein anderer Weg das Periodogramm zu bestimmen ist die Fourier-Transformation der Autokovarianzfunktion von  $Y(t)$ . Numerisch identische Ergebnisse wie im aufgezeigten Fall werden erreicht, indem  $c(h)$  für  $h = 1, \dots, n - 1$  geschätzt wird. In diesem Fall werden nicht existierende Realisierungen,  $y(t)$  wenn  $t + h > n$ , über die symmetrische Verlängerung von  $Y(t)$  ersetzt.

Das diskrete Periodogramm stellt einen erwartungstreuen Schätzer für das kontinuierliche Spektrum der Zeitreihe dar. Die direkte Verwendung als solcher ist allerdings aufgrund seiner fehlenden Konsistenz und starken Streuung, sowie der Anfälligkeit gegenüber Leakage- und Aliasingeffekten problematisch. Stattdessen werden Schätzer für das Spektrum über unterschiedliche Ansätze der Filterung des Periodogramms gewonnen.

Hierfür existieren verschiedene Fensterfunktionen, die sich allerdings in den konkreten Auswirkungen auf das damit erzeugte Spektrum nur wenig unterscheiden

(PRESS et al. 1992), wenn sich auch manche Methoden für spezielle Anwendungen, insbesondere bei relativ kleinen  $n$ , als nützlicher als andere erwiesen haben. So können einige Fensterfunktionen spektrale Eigenschaften detektieren, die bei der Nutzung anderer Fenster nicht auftreten. Die Entscheidung, ob solche Effekte real oder Artefakte spezieller Fensterfunktionen sind, ist nicht trivial (YIOU et al. 1996). Da hier ein möglichst stabiler und allgemeiner Schätzer für das Spektrum angestrebt wird, wird das gängige Daniell-Fenster genutzt. In der Frequenzdomäne ist das Daniell-Fenster  $W_D$  einfach der gleitende Durchschnitt über das Periodogramm. Der Schätzer für das Spektrum  $S$  und seine Ausprägungen  $S(\omega_j)$  ist somit:

$$S(\omega_j) = \frac{1}{m} \sum_{k=j-\frac{m-1}{2}}^{j+\frac{m-1}{2}} P(\omega_k) \quad \text{mit } m = 1, 3, \dots \quad (3.59)$$

Alternativ wäre auch die Filterung der Autokovarianzfunktion in der Zeitdomäne möglich.

## 4 Grundlegende Modellvalidation

Der erste Zugang zur Beurteilung der Performance eines AOGCMs ist das Hindcasting. Damit wird die Fähigkeit der Klimamodelle getestet, beobachtete Zustände des Klimasystems realistisch zu simulieren, was eine logische Anforderung an Modelle darstellt, die zum Studium von Prozessen und Phänomenen im Klimasystem verwendet werden.

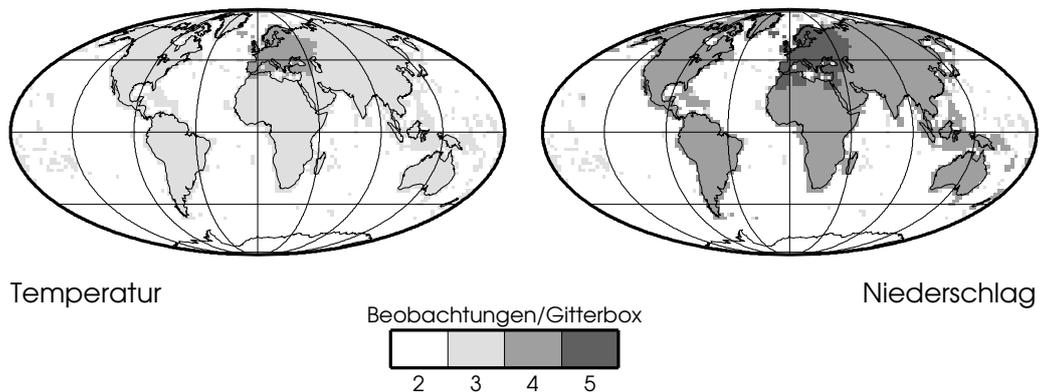
Allerdings muss die Aussagekraft des Hindcastings, speziell bei der Erforschung des künftigen Klimawandels, als beschränkt angesehen werden. So kann ein Klimamodell, das für die Vergangenheit realistische Bedingungen simuliert, nicht per se als für die Zukunft zuverlässig eingestuft werden. Die erfolgreiche Simulation der Vergangenheit, also eines bekannten Zustandes, kann viele Gründe haben, und die realistische Abbildung des dynamischen Klimasystems ist nur einer davon. Nicht alle diese Gründe bedeuten jedoch, dass ein Modell auch unter den völlig anderen Randbedingungen im späten 21. Jahrhundert realistische Ergebnisse liefert. Dieser Ansatz wird in Kapitel 1.2 diskutiert. Als Grundlage für die Gewichtung oder Bewertung von AOGCMs wird er hier nicht verfolgt.

Natürlich stellt sich aber auch die Frage, ob ein AOGCM ohne jede erkennbare Fähigkeit das beobachtete Klima zu reproduzieren dazu in der Lage dazu sein kann, eine realistische Abschätzung der Reaktion des Klimasystems auf den anthropogenen Einfluss zu ermöglichen. Die erfolgreiche Simulation der Vergangenheit wird hier daher als notwendige, aber nicht hinreichende Bedingung für ein gutes AOGCM verstanden (dazu z. B. auch ORESKES et al. 1994). Ein deutlicher Abfall eines Modells von der durchschnittlichen Leistung des übrigen Ensembles könnte durchaus dessen Ausschluss aus weiteren Analysen rechtfertigen.

Hierfür muss aber ein besonderer Aspekt in den Entscheidungsprozeß mit einbezogen werden. Eine, häufig ignorierte, grundsätzliche Problematik des Hindcastings ist die Verfügbarkeit an qualitativ hochwertigen Beobachtungsdaten. Sie ist natürlich in unterschiedlichen Weltregionen unterschiedlich stark ausgeprägt. Für einige Gebiete (etwa West- und Mitteleuropa), lassen sich leicht zuverlässige Messreihen bis ins 19. Jahrhundert in so großer Zahl finden, dass zumindest die Temperatur als bekannt betrachtet werden kann. In anderen Gebieten, im Extremfall in der Antarktis, finden sich bis zum Einsatz von Wettersatelliten allenfalls kurze Messreihen an wenigen Stationen. Daher ist es bei globaler Betrachtung angeraten, sich auf langjährige Mittelwerte zu konzentrieren, wenn ein Abgleich mit Modelldaten nicht schon aufgrund der Qualität der Beobachtungsdaten als sinnlos erscheinen muss.

## 4.1 Separative Betrachtung

Da sich die folgenden Untersuchungen auf die langjährigen Klimatologien der grundlegenden Klimaelemente konzentrieren, werden die relativ kurzen Beobachtungsdatensätze TRMM und GPCP nicht berücksichtigt. Alle Beobachtungen, die die Periode 1960–1999 abdecken, wurden auf das Gitternetz der ERA-40-Daten interpoliert. Abbildung 4.1 zeigt die Anzahl der Beobachtungen pro Gitterpunkt für bodennahe Temperatur und Niederschlag. Die drei Datensätze für den Luftdruck auf Meeresniveau liegen jeweils in globaler Abdeckung vor.



**Abb. 4.1:** Beobachtungsdaten für 1960–1999. Bodennahe Temperatur und Niederschlag, interpoliert auf das ERA 40-Gitternetz

Betrachtet werden die Jahresmittel (Januar bis Dezember) und üblichen saisonalen Mittelwerte. Demnach berechnet sich der Mittelwert für den borealen Winter im Jahr 1960 über:

$$\bar{x}_{djf}(1960) = \frac{x_d(1959) + x_j(1960) + x_f(1960)}{3} \quad (4.1)$$

Die unterschiedlichen Beobachtungsdaten werden durch einfache Mittelung kombiniert, wobei aber insbesondere auch die Streuung zwischen den Daten von besonderem Interesse ist. Die Beurteilung, wann ein AOGCM die untersuchte Klimatologie zufriedenstellend reproduziert, kann somit vor dem Hintergrund der tatsächlichen Datenqualität getroffen werden.

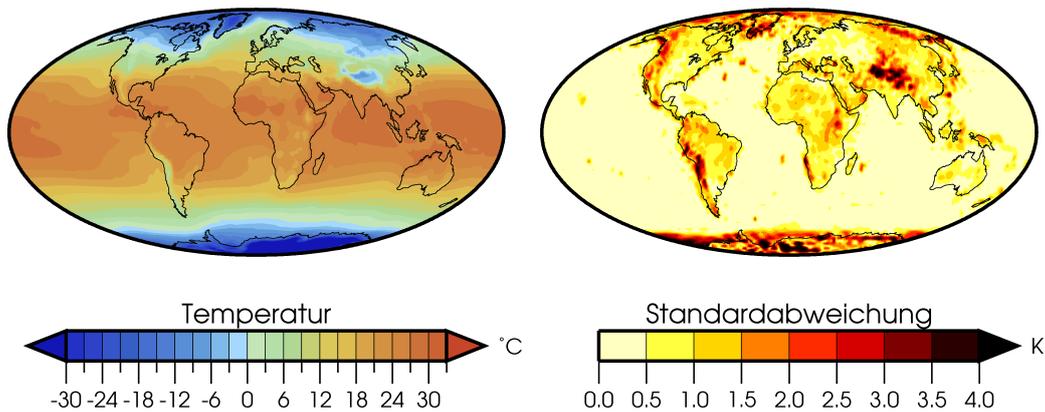
### 4.1.1 Beobachtungsdaten: kombinierte Klimatologien

Klimatologische Mittelwerte für die bodennahe Temperatur  $\bar{T}(s)$ , den Niederschlag  $\bar{R}(s)$  und den auf Meeresniveau reduzierten Luftdruck  $\bar{P}(s)$  bilden die Zielgrößen des Hindcastings. Sie werden hier knapp vorgestellt, wobei Regionen mit

auffälligen Abweichungen zwischen den Datensätzen, die Standardabweichungen  $s_T(s)$ ,  $s_R(s)$  und  $s_P(s)$  hervorgehoben werden, da diese die Begründung für die gewählte Vorgehensweise liefern.

#### 4.1.1.1 Bodennahe Temperatur

Bei kleinskaliger Betrachtung unterscheiden sich grundsätzliche Muster und Amplituden der Klimatologien der verschiedenen Beobachtungsdatensätze nur geringfügig. Entsprechend liefert die Kombination dieser Datensätze durchweg vertraute und plausible Ergebnisse. Diese Aussage relativiert sich etwas bei der Untersuchung der auftretenden Standardabweichungen. Der systematische Unterschied über dem Ozean, d. h. zwischen ERA 40 und NCEP/NCAR, ist nur gering und nicht von der betrachteten Jahreszeit abhängig. Über den Landmassen, und mit dem Einfluss der CRU-Daten, erhöht sich die Streuung allgemein. Am geringsten ist sie im europäischen Bereich, der zusätzlich durch E-OBS abgedeckt wird.



**Abb. 4.2:** Kombination Beobachtungen: Temperatur. Klimatologie für den Zeitraum 1960–1999. Jahresmittelwerte  $\bar{T}(s)$  in  $^{\circ}\text{C}$  und Standardabweichung  $s_T(s)$  in K für die bodennahe Temperatur.

Die Unsicherheit ist in einigen Gebieten besonders erhöht. Zunächst ist hier die Antarktis zu nennen. Im Südwinter werden hier, ebenso wie über dem Südpolarmeer, Standardabweichungen von bis zu 6 K erreicht. Auch die Klimatologien für die anderen Jahreszeiten sowie die Jahresmittelwerte sind mit 3 bis 4 K deutlich über dem globalen Mittel.

Ebenfalls über alle Jahreszeiten hinweg finden sich über den Hochgebirgen der Erde deutliche Unterschiede zwischen den Beobachtungen. Begründet ist dies einerseits in der geringen Anzahl verfügbarer Messdaten, auf die sich die Flächendaten stützen müssen, sowie im Einfluss des Reliefs, der statistische wie dynamische Interpolationen erschwert. Es ist nicht verwunderlich, dass dies insbesondere für

das Hochland von Tibet gilt. Hier finden sich sowohl in der jährlichen Betrachtung als auch über alle Saisons hinweg Standardabweichungen von bis zu 4 K, wobei die Unterschiede im borealen Winter und Herbst sowohl flächenmäßig als auch vom Betrag her am größten sind. Aus der gleichen Problematik resultierend und zeitlich ähnlich konsistent, wenn auch betragsmäßig nicht so stark ausgeprägt, sind die Unterschiede über den Kordilleren sowie im Bereich des afrikanischen Grabenbruchs. Andere Faltengebirge, etwa die Nordalpen oder der Atlas, sind wohl flächenmäßig zu klein um derartige Effekte hervorzurufen, zudem kann hier die Datengrundlage als besser angenommen werden.

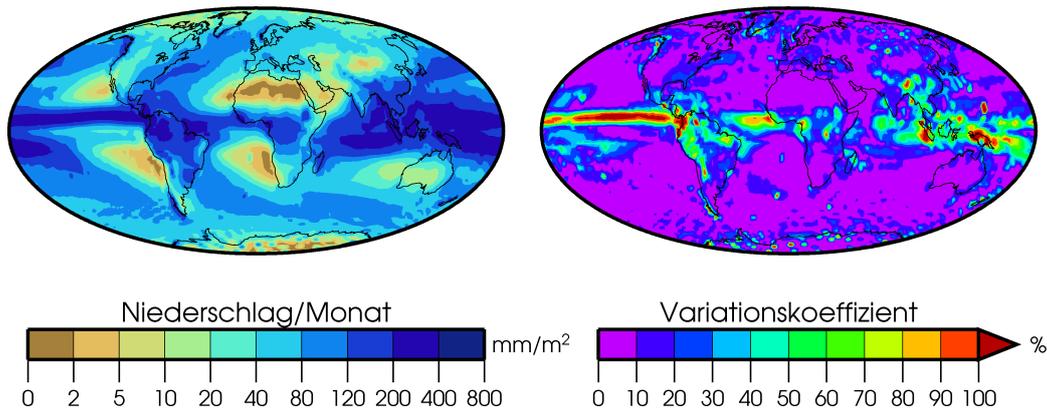
Für die Arktis finden sich ebenfalls deutliche Differenzen in den Klimatologien der Temperatur. Über das Jahr hinweg gilt dies für Grönland. Im Nordsommer sind die Unterschiede zwischen den beiden Reanalyse-Datensätzen über dem arktischen Ozean und seinen Nebenmeeren weniger stark ausgeprägt. Diese Annäherung der Datensätze im Sommer findet sich auch für den Bereich ansonsten erhöhter Unsicherheit in Nordasien. Auffällig sind zudem die hohen Standardabweichungen für die nordamerikanische Ostküste, den Mittelmeerraum und Neuseeland im jeweiligen hemisphärischen Winter.

#### 4.1.1.2 Niederschlag

Der Niederschlag ist eine räumlich hochvariable Größe. Diese Variabilität findet sich auch in seiner Klimatologie wieder. Daher sind bei globaler Betrachtung Standardabweichungen weniger aussagekräftig als für  $barT(s)$  und  $barP(s)$ . Angesichts der Problematik bei der Bestimmung des gefallenen Niederschlags für eine bestimmte Fläche ist grundsätzlich davon auszugehen, dass in ariden Regionen die Standardabweichung  $s_R(s)$  einen geringeren Wert annimmt als in humiden. Über die Qualität der Daten ist damit allerdings noch keine Aussage zu treffen. Daher wird hier in der kartographischen Darstellung sowie in der Erläuterung der Variabilitätskoeffizient  $vk(s)$  verwendet. Dieser setzt die Standardabweichung zum Mittelwert  $\bar{R}(s)$ , in diesem Fall zur mittleren Monatssumme in mm, in Beziehung.

$$vk(s) = \frac{s_R(s)}{\bar{R}(s)} \cdot 100\% \quad (4.2)$$

Wegen des geringen absoluten Niederschlags nimmt  $vk(s)$  für die Trockengebiete der Erde durchaus Werte von über 100 % an. Die absoluten Abweichungen sind dagegen allerdings gering, was bei der Interpretation der Karten beachtet werden muss. Die Unterschiede in den Beobachtungen sind auf individuelle Probleme und Fehler in den verschiedenen Interpolationsverfahren zurückzuführen. Aber auch in feuchteren Regionen treten Werte von über 50 % auf. Insbesondere der tropische Atlantik und Ostpazifik weisen zu allen Perioden zwei lokale Maxima nördlich und südlich des Äquators auf. Die Intensität, die Ausdehnung und teilweise



**Abb. 4.3:** Kombination Beobachtungen: Niederschlag. Wie Abb. 4.2, aber durchschnittlicher Niederschlag  $\bar{R}(s)$  pro Monat im  $\text{mm}/\text{m}^2$  und dessen Variationskoeffizient  $vk(s)$  in %

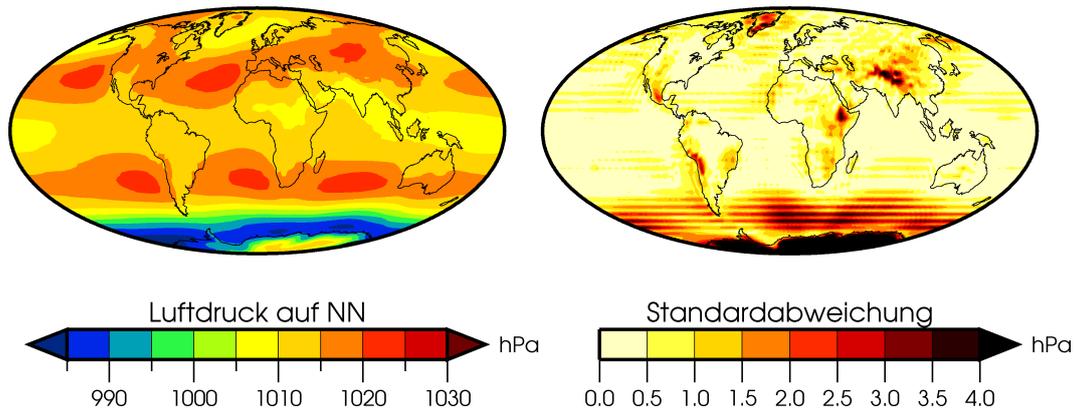
auch die Lage der klimatologischen innertropische Konvergenzzone unterscheidet sich in diesen Regionen also zwischen den Reanalysen. Die Standardabweichung der Monatssummen überschreitet hier teilweise 130 mm, was für die jährliche Klimatologie einen Unterschied von knapp 1600 mm bedeutet.

Für die Landmassen lässt sich, mit wenigen Ausnahmen, eine Streuung von mindestens 15–30 % ausmachen, wobei die Werte von  $vk(s)$  in einigen Regionen, etwa Australien aber auch Teilen Europas, abhängig von der Jahreszeit auch 70 % und mehr betragen können. Besonders ausgeprägt sind die Unterschiede in den Beobachtungen über Grönland und auch im zentral- bis ostasiatischen Raum, inklusive dem – und besonders markant für – das Hochland von Tibet. Hier werden höhere Werte von  $vk(s)$  erreicht als in den sonstigen Festlandregionen. In absoluten Werten sind die Abweichungen der Beobachtungen für Indien und Südostasien, insbesondere an den Küsten, während der Monsunzeit vergleichbar mit denen im tropischen Pazifik, trotz des größeren Datenumfangs (Abb. 4.1).

#### 4.1.1.3 Luftdruck auf Meeresniveau

Für die Nordhalbkugel und die Tropen sind die Beobachtungsdatensätze für den SLP über den Meeren recht ähnlich. Relativ hohe Standardabweichungen  $s_P(s)$  finden sich in den Randbereichen der Zonen bevorzugter zyklonaler Aktivität sowie der großen Antizyklonen des subtropisch-randtropischen Hochdruckgürtels.

Die im Vergleich zu den sonstigen Gebieten enormen Abweichungen und offensichtlichen Unterschiede zwischen den Beobachtungen in den südhemispherischen Außertropen sind auf den NCEP/NCAR Datensatz zurückzuführen, der hier, vor allem vor der Satellitenära aber auch in späteren Daten, einen deutlichen Bias im



**Abb. 4.4:** Kombination Beobachtungen: Luftdruck. Wie Abb. 4.2 aber für Luftdruck auf Meeresebene. Jahresmittelwerte  $\bar{P}(s)$  mit Standardabweichung  $s_P(s)$  in hPa.

Vergleich mit in situ Messungen aufweist (SMITH et al. 2001). Da die SLP-Daten von NCEP/NCAR auf dem gleichen räumlichen Gitter vorliegen wie das hier verwendete ERA 40-Gitter, ist auch die bemerkenswerte räumliche Struktur der Standardabweichungen nicht auf hier vorgenommene Berechnungen zurückführbar, sondern stellen Artefakte aus dem Assimilierungsprozess selbst dar.

Für den antarktischen Kontinent, einem Gebiet mit Werten von  $s_P(s) > 10$  hPa, finden sich hingegen auch zwischen den beiden anderen Datensätzen deutliche Unterschiede. Für die Westwindzone der Nordhalbkugel lassen sich keine vergleichbaren quantitativen Unterschiede feststellen. Jedoch zeigt sich in Regionen, die ungefähr mit Grönland, der Beringsee sowie dem sibirischen Raum übereinstimmen, während der Wintermonate eine bedeutende Spannbreite innerhalb der verschiedenen Klimatologien. Dagegen weisen die drei Datensätze in den Klimatologien über den übrigen Landmassen relativ geringe Unterschiede auf. Die Ausnahmen hiervon lassen sich in Abb. 4.4 leicht ausmachen und ungefähr als Hochland von Tibet, als afrikanischer Grabenbruch mit Schwerpunkt auf dem Hochland von Äthiopien und als das Altiplano identifizieren. Alle drei geographischen Einheiten sind also durch starke Reliefgradienten gekennzeichnet. Die hohen Standardabweichungen in diesen Gebieten sind in allen betrachteten Saisons und im Jahresmittel festzustellen und sicherlich auf die gleichen Effekte zurückzuführen wie im Fall der Temperatur. Allerdings nimmt sowohl der Betrag als auch die räumliche Ausdehnung dieser Zonen der erhöhten Unsicherheit zwischen den Beobachtungen generell im Winter der jeweiligen Hemisphäre zu.

## 4.1.2 Methodik

Das Hindcasting der beobachteten Klimatologien wird über zwei unterschiedliche Methoden bewertet. Die so gewonnenen Resultate werden einander gegenübergestellt. Die Nutzung von Konfidenzintervallen, die über verschiedene Beobachtungsdatensätze hinweg gebildet werden, ist ein ungewöhnlicher und weitgehend ungebräuchlicher Ansatz hierfür. Dieses also durchaus neue Vorgehen wird durch die parallele Analyse der Unterschiede zwischen den Modellklimatologien und denen der Beobachtungsdaten anhand des Root-Mean-Square-Error überprüft und abgesichert.

### 4.1.2.1 Konfidenzintervalle

Sämtliche Messungen unterliegen grundsätzlich einer gewissen Unsicherheit. Während diese bei modernen und sachgerecht bedienten Geräten zur in situ Erfassung der grundlegenden Klimaelemente Temperatur und Luftdruck sehr gering ist, kann dies nicht grundsätzlich auch für ältere Messungen oder solche des räumlich hochvariablen und von diversen anderen Parametern beeinflussten Niederschlags gelten (FOKEN 2006). Dennoch kann bei klimatologischer Betrachtung im Fall einer konkreten Messung grundsätzlich eine hohe Ähnlichkeit des Messwertes und der tatsächlichen Größe angenommen werden.

Die Beobachtungsdatensätze in diesem Kapitel stellen jedoch keine realen Messungen dar. Stattdessen wurden sie über statistische oder dynamische Verfahren in den Raum interpoliert. Diese Gitterdaten weichen für eine Region zu einem bestimmten Zeitpunkt teilweise deutlich voneinander ab. Ein offensichtlicher Grund hierfür ist die räumliche Auflösung der Datensätze, stellen die jeweiligen Elemente doch repräsentative Werte für die jeweilige Gitterbox dar. Weiterhin entscheidend ist die räumliche Verteilung der ursprünglichen Messungen, aus denen der Datensatz erstellt wurde, bzw. die Dichte der Referenzpunkte, mit denen die dynamische Interpolation initialisiert wurde.

Somit ist für einen Datensatz die Qualität der Daten im Hinblick auf die Ähnlichkeit mit dem tatsächlichen Zustand des Klimasystems abhängig von der geographischen Lage der Gitterpunkte. Es kann und muss angenommen werden, dass keiner der verwendeten Datensätze die tatsächlichen Bedingungen fehlerfrei wiedergibt. Die unterschiedlichen Generierungsprozeduren könnten allerdings dazu beitragen, dass einige Datensätze in bestimmten Regionen zuverlässigere Ergebnisse erzeugen als andere.

Um dieser Problematik Rechnung zu tragen, werden bei der Beurteilung des Hindcastings der AOGCMs keine Punkt-, sondern Intervallschätzungen für die Klimatologie der Beobachtungen betrachtet. Für alle  $j = 1, \dots, n_B$  Datensätze, die die Periode 1960–1999 abdecken wird zunächst separat der Mittelwert  $\bar{x}_{B_j}(s)$  für jeden der  $s = 1, \dots, n$  Gitterpunkte berechnet. Anschließend wird aus diesen Schätzungen

das interessierende, symmetrische Konfidenzintervall  $KI(s)$  für  $\bar{x}_B(s)$  bestimmt.

$$KI(s) = \left[ \bar{x}_B(s) - t_{n_B-1;1-\frac{\alpha}{2}} \frac{s_B}{\sqrt{n_B}}; \bar{x}_B + t_{n_B-1;1-\frac{\alpha}{2}} \frac{s_B}{\sqrt{n_B}} \right] \quad (4.3)$$

Dies bedeutet, dass zur Bestimmung der Unsicherheit innerhalb der Beobachtungen nur sehr kleine Stichproben – zwischen zwei und fünf Klimatologien, je nach Variable und Gitterpunkt – zur Verfügung stehen. Allerdings ist dies ein Ausdruck der grundsätzlichen Problematik der Datenverfügbarkeit und als solcher bei dem gewählten Ansatz unvermeidbar. Da es sich bei diesen Variablen um langjährige Mittelwerte handelt kann mit großer Sicherheit von einer Normalverteilung der Werte ausgegangen werden und die Intervallschätzung unter Verwendung der  $t$ -Verteilung statistisch korrekt erfolgen. Bei der Bestimmung von  $KI(s)$  wird jeweils eine Irrtumswahrscheinlichkeit von  $\alpha = 0.05$  vorgegeben. Damit gilt für die so erzeugten Konfidenzintervalle eine Überdeckungswahrscheinlichkeit des über  $\bar{x}_B(s)$  geschätzten wahren Parameters  $\theta(s)$  von 95 %.

Um die Performance der AOGCMs zu beurteilen, werden die entsprechenden Klimatologien  $\bar{x}(s, i)$  der  $n_M$  Läufe des Multi-Modell Ensembles dahingehend überprüft, ob sie entsprechend der Intervallschätzung als plausibel zu betrachten sind. Dies ist der Fall, wenn die simulierte Klimatologie an dem Gitterpunkt innerhalb von  $KI(s)$  liegt. Anschließend wird für den  $i$ -ten Modelllauf der Anteil der plausiblen Fläche  $A_i$  berechnet.

$$A_i = \sum_{s=1}^n \omega(s) f(\bar{x}(s, i)) \cdot 100\% \quad \text{mit} \quad f(\bar{x}_i(s)) = \begin{cases} 1 & \text{für } \bar{x}(s, i) \in KI(s) \\ 0 & \text{für } \bar{x}(s, i) \notin KI(s) \end{cases} \quad (4.4)$$

Die  $\omega(s)$  stellen eine Flächengewichtung dar. Hiermit steht eine Maßzahl zur Verfügung, um die globale Performance eines einzelnen Modells zu beurteilen. Daneben ist der schnelle Vergleich mehrerer Modelle über diese Größe möglich.

#### 4.1.2.2 Root-Mean-Square Error

Der schnelle und direkte Vergleich zwischen zwei oder mehr Klimamodellen ist auch bei der Nutzung des RMSE möglich. Diese Methodik eignet sich daher gut zum Vergleich mit den Ergebnissen der Methode aus Kapitel 4.1.2.1. Da die Berechnung des RMSE ein gebräuchliches und häufig verwendetes Verfahren zur Quantifizierung von Abweichungen zwischen Modellen und Beobachtungen darstellt, kann man dieses Vorgehen auch als Kontrolle des Ansatzes über die Intervallschätzungen über die  $KI(s)$  betrachten (z. B. FOX 1981, WILLMOTT et al. 1985).

An jedem Gitterpunkt wird zunächst die Differenz zwischen dem klimatologischen Wert des  $i$ -ten Modelllaufs  $\bar{x}(s, i)$  und dem Mittelwert der diversen Beobachtungs-

datensätze  $\bar{x}_B(s)$  gebildet:

$$e(s, i) = \bar{x}(s, i) - \bar{x}_B(s) \quad (4.5)$$

Hieraus ergibt sich eine Komposit-Karte. Positive Differenzen zeigen die Überschätzung des Klimaelementes an dieser Gitterbox durch das GCM im Vergleich zu den Beobachtungsdaten an, negative dementsprechend eine Unterschätzung. Diese Differenzen werden quadriert und unter Berücksichtigung der Flächenkonvergenz in Richtung der Pole gemittelt um den RMSE zu erhalten, der dann eine Maßzahl für die mittlere Abweichung des Modells von der beobachteten Klimatologie darstellt.

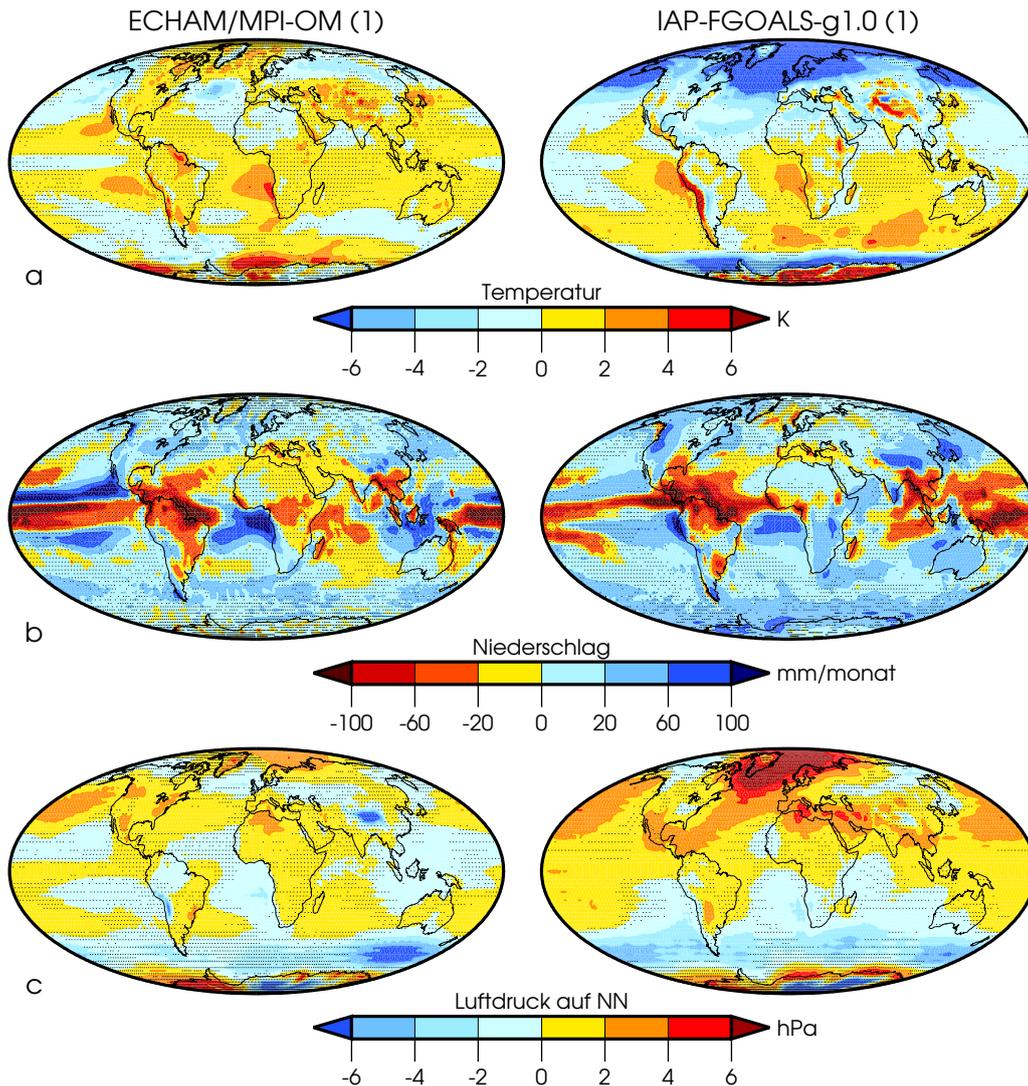
$$RMSE_i = \sum_{s=1}^n \omega(s) \cdot (e(s, i))^2 \quad (4.6)$$

Diese Größe ist für sich nicht einfach zu interpretieren. Dies liegt zum einen an den Gewichtungen, zum anderen daran, dass sie abhängig von der verwendeten Einheit des Klimaelements ist. Daher ist es üblich eine Normierung vorzunehmen und somit diese absolute Größe in eine relative zu überführen. Hier wird für jedes Klimaelement der Lauf mit dem höchsten Wert für den RMSE auf 100 % gesetzt und die übrigen entsprechend skaliert. Diese Normierung bietet sich an, da sie sowohl den schnellen Vergleich der einzelnen Werte erlaubt, als auch die Rücktransformation in absolute Werte. Natürlich wären auch andere Normierungen möglich, etwa gitterboxweise über ein Streuungsmaß für die Beobachtungsdaten. Hiermit würde man die Unsicherheit der Beobachtungen stärker berücksichtigen als durch die bloße Verwendung des Mittelwerts. Im Gegensatz zur Verwendung von Konfidenzintervallen würde dies aber keine Wahrscheinlichkeitsaussage ermöglichen, sondern nur die lokalen Fehler relativieren, was die Interpretation des RMSE eher weiter erschweren würde.

### 4.1.3 Ergebnisse

In Abb. 4.5 sind einige Ergebnisse der Analysen in diesem Kapitel für zwei Modellläufe dargestellt. Hier werden jeweils Jahresmittelwerte gezeigt, generell sind die Beträge  $|e(s, i)|$  für die Jahreszeiten größer und der Anteil  $A_i$  entsprechend kleiner. ECHAM5/MPI-OM (1) und IAP-FGOALS-g1.0 (1) können hier einen guten Eindruck von den Modellsread, also den Spannbreiten, die die Simulationen im Ensemble aufweisen, geben. Der erste Lauf kann in nahezu allen Jahreszeiten und für sämtliche Klimavariablen mit die besten Ergebnisse in beiden Analyseverfahren aufweisen. Für IAP-FGOALS-g1.0 (1) muss dagegen festgestellt werden, dass dieses Modell sich durchweg am unteren Ende einer entsprechenden Rangfolge der Performance bewegt. Allerdings verfügt auch dieses Modell über durchaus nicht zu vernachlässigende Anteile mit als plausibel geltenden Ausprägungen der

Klimavariablen. Dies verdeutlicht die große Streuung innerhalb der Beobachtungsdatensätze. Es muss jedoch festgehalten werden, dass die Abweichungen zwischen Beobachtungen und Modelldaten teilweise außerordentlich sind. Als ein Beispiel von vielen sei hier nur darauf hingewiesen, dass IAP-FGOALS-g1.0 (1) im Bereich des Nordatlantiks und weiten Teilen Europas  $\bar{T}(s)$  um rund 10 K unterschätzt, was auf eine fehlende oder äußerst unrealistische thermohaline Zirkulation hinweist.



**Abb. 4.5:** Komposite Modell- und Beobachtungsdaten. Jährliche Mittel der 1960–1999 Klimatologien für bodennahe Temperatur (a), Niederschlag (b, unregelmäßige Skalierung) und Luftdruck auf NN (c). Schattierte Bereiche: Gitterboxen mit Modelldaten innerhalb des 95 %-Konfidenzintervalls der Beobachtungen

Besonderes Interesse gilt bei der Bewertung der Modell-Performance den elf in mehrfachen Realisierungen vorliegenden Modellen. Ausgeprägte Unterschiede zwischen individuellen Läufen eines GCMs müssten als Hinweis dafür gewertet werden, dass das angewandte Verfahren deutlich von den Anfangsbedingungen, und somit vom Zufall, beeinflusst wird. Jedoch geben die Ergebnisse hierfür keinerlei Anlass. Für alle berücksichtigten Variablen sind die Flächenanteile mit plausiblen Werten für die einzelnen Modellläufe sehr ähnlich, vor allem verglichen mit den Unterschieden zwischen verschiedenen Modellen. Beispielhaft ist dies in Abb. 4.6 für die jährlichen Mittelwerte dargestellt. Zudem ist auch die jahreszeitliche Variation der Flächenanteile nahezu identisch für die Klimaelemente. Die einzige Ausnahme innerhalb des Multi-Modell Ensembles findet sich beim Lauf NCAR-CCSM (5). Hier ist der als plausibel bewertete Flächenanteil der bodennahen Temperatur für den borealen Winter um etwa 1 % größer als für die SON-Saison, während dies bei den übrigen sechs Läufen des Modells umgekehrt ist. Dieser geringfügige Unterschied kann wohl als die regelbestätigende Ausnahme aufgefasst werden. Somit wird die Methode als grundsätzlich geeignet aufgefasst, Modellunterschiede zu erfassen und abzubilden.

Der relative Anteil der plausiblen Flächen ist für alle Modelle bei der Betrachtung saisonaler Ergebnisse erwartungsgemäß geringer als bei Betrachtung der Jahresmittel. Auf diesem niedrigeren Niveau sind die Rangfolgen der Modelle allerdings insgesamt als stabil zu bezeichnen. Somit vermittelt Abb. 4.6 auch für die Betrachtung der Jahreszeiten einen recht guten Überblick. Die Leistung der Modelle scheint also nicht, oder nur in geringem Umfang, von der jeweiligen Jahreszeit mit ihren spezifischen Prozessen abzuhängen. Dies gilt zumindest im globalen Maßstab, wenn die räumliche Verteilung der plausiblen Flächen nicht in Betracht gezogen wird.

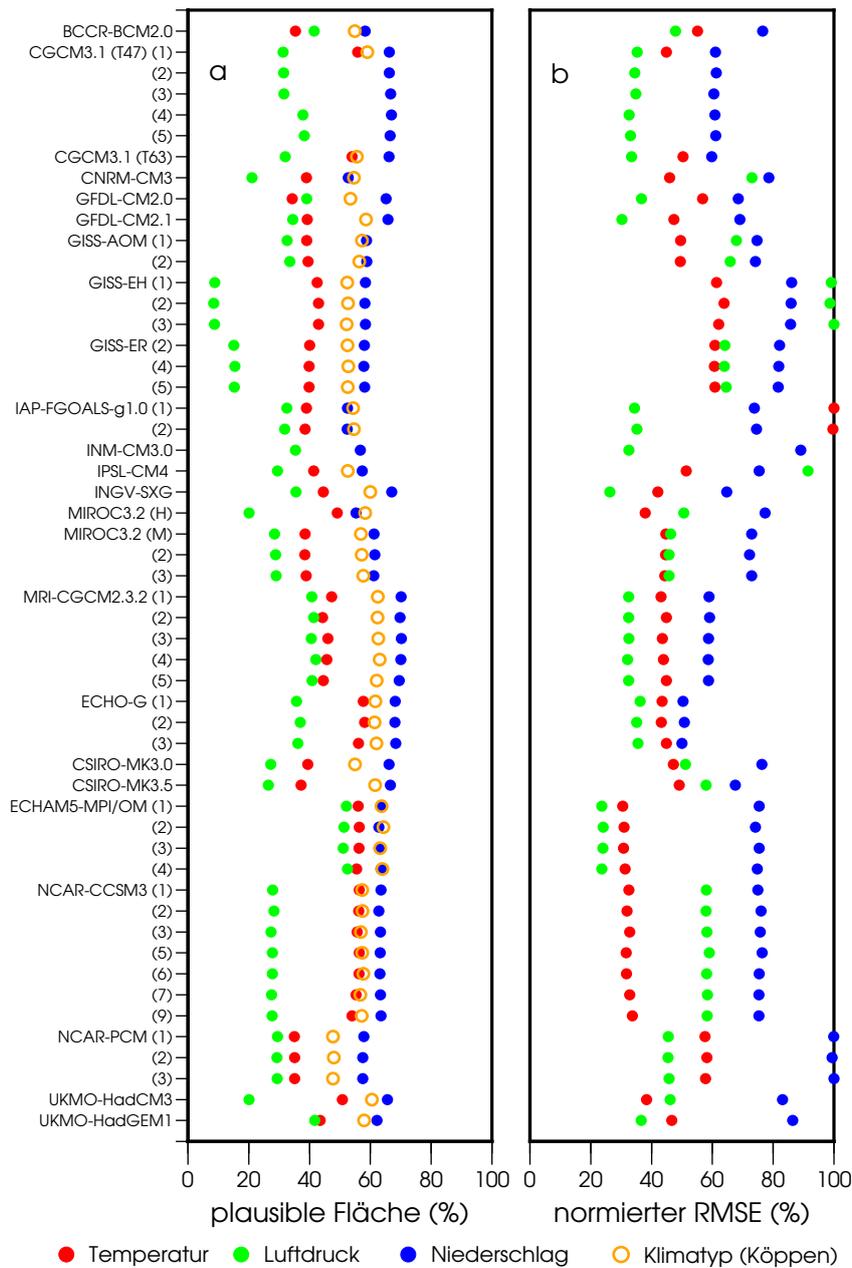
Im Bezug auf die plausiblen Flächen lassen sich durchaus Unterschiede zwischen den verschiedenen Modellen feststellen. So erscheint der als korrekt klassifizierte Anteil des Niederschlags im Intermodellvergleich deutlich robuster als die Simulation des Luftdrucks. Ebenfalls sind die plausiblen Flächenanteile, der AOGCMs für die einzelnen Größen deutlich unterschiedlich. So trennen den erfolgreichsten Lauf bei der Reproduktion der globalen Niederschlagsklimatologie nur 10 Prozentpunkte. Die Spannbreite beim SLP ist mit etwa 40 % der Erdoberfläche deutlich größer.

Die Performance zeigt keine klar erkennbare Clusterung über die Variablen hinweg. Ausnahmen hiervon sind z. B. ECHAM5/MPI-OM, das grundsätzlich sehr hohe plausible Flächenanteile aufweist, sowie, wenn auch weniger deutlich, NCAR-PCM und IAP-FGOALS-g.1.0, die variabelübergreifend eher geringe Werte von  $A_i$  erhalten. Beim Hindcasting ergeben sich somit bei modellübergreifender Betrachtung durchaus Ausreißer nach oben und unten. Die Berücksichtigung von mehrerer Feldgrößen und der globalen Domäne zeigt jedoch insgesamt starke Variationen in der Performance der Modelle auf. Diese teilweise widersprüchlichen Ergebnisse lassen somit eine Aufteilung des CMIP3 als nicht zielführend erschei-

nen. Die Variabilität der Resultate für Temperatur, Niederschlag und Luftdruck lässt es zudem nicht unplausibel erscheinen, dass eine Erweiterung der Analyse um zusätzliche Variablen identifizierbare Tendenzen eher weiter verwischen als verstärken würde.

Etwas ernüchternd muss zudem festgestellt werden, dass in den im Allgemeinen als am zuverlässigsten geltenden Daten – den Datensätzen für den Luftdruck auf Meereshöhe –  $A_i$  modellübergreifend die geringsten Werte annimmt. Hingegen finden sich durchgehend die höchsten Flächenanteile mit plausiblen Werten beim von den Beobachtungsdaten als am unzuverlässigsten abgebildeten Niederschlag. In diesem Fall liefert die Validation des CMIP3 somit quantitative Hinweise auf die enorme Ungenauigkeit, die mit der Nutzung von verbreiteten Beobachtungsdaten für den Niederschlag verbunden ist. Deren hier aufgezeigte Streuung für die fundamentale Klimatologie lässt ihre Verwendung für die Beurteilung des Realismus der AOGCMs als fragwürdig erscheinen.

Diese Feststellung wäre mit dem RMSE alleine nicht zu treffen gewesen, da dieser sich allein auf den Mittelwert der Beobachtungen bezieht. Trotz dieser im Vergleich geringeren Information, die zur Erstellung von Abb. 4.6 b) herangezogen wurde, sind die Ergebnisse der beiden Ansätze durchaus vergleichbar.  $A_i$  und  $RMSE_i$  einer Variable weisen deutlich eine negative Korrelation auf, weiterhin sind die Beträge für unterschiedliche Läufe eines AOGCMs als quasi identisch einzustufen. Auch der RMSE identifiziert den SLP als die Variable, die die deutlichsten, relativen, Unterschiede für das CMIP3 aufweist. Die Simulation der Niederschlagsklimatologie dagegen ist, mit Ausnahme von ECHO-G welches hier die beste Performance zeigt, durch  $RMSE_i$  am einfachsten mit „vergleichbar schlecht“ zu beurteilen. Hier ist allerdings der für diese Größe besonders starke Effekt mitzubedenken, der sich z. B. aus der Interpolation durch eine Verschiebung der Innertropischen Konvergenzzone ergeben kann. So erzeugte große quantitative Unterschiede wirken sich deutlich stärker auf  $RMSE_i$  aus als auf  $A_i$  mit seiner dichotomen Bewertung. Der RMSE berücksichtigt damit auch nicht die tatsächliche Ungewissheit der Beobachtungsdaten, die insbesondere im Bezug auf Lage und Intensität der ITC als tatsächlich ausschlaggebend für die Beurteilung der Modellklimatologie herangezogen werden muss (vgl. Abb.4.3). Für die Feldgröße der bodennahen Temperatur liefern beide Verfahren modellübergreifend betrachtet keine wesentlichen Ergänzungen im Vergleich zu den Informationen und Interpretationen, die die beiden anderen Variablen eröffnen. Der aufgezeigte Bias von IAP-FGOALS-g.1.0 4.5 summiert sich im RMSE soweit auf, dass dieses Modell hier einen deutlichen Ausreißer darstellt. Die auf dieses Modell folgenden AOGCMs weisen nur Werte von knapp 60 % seines RMSEs auf. Dagegen zeigt der zwar niedrige, aber doch nicht vergleichbar abfallende Wert von  $A_i$  für die beiden Läufe, dass es sich hierbei nur um einen regionalen Effekt, wenn auch an prominenter Stelle und beachtlichen Ausmaßes, handelt.



**Abb. 4.6:** Übersicht zur Evaluation des Multi-Modell Ensembles. Flächenanteile innerhalb des 95 %-Konfidenzintervalls der Beobachtungsdaten für drei Klimaelemente (a) bzw. normierter RSME im Vergleich zum Mittel der Beobachtungen (flächengewichtet, b). Jeweils jährliche Klimatologien. In a) zusätzlich die Flächenanteile, deren Klimatyp bei der Anwendung der kompletten Klimaformel nach Köppen mit den kombinierten Beobachtungsdaten übereinstimmt.

Diese Diskrepanz in der Interpretation der beiden Verfahren zeigt zugleich einen Vorteil der hier etwas explorativ angewendeten Methode im Vergleich zum RMSE auf. Die Unterscheidung in plausible oder nicht-plausible Werte liefert eine, theoretisch nicht weiter interpretationsbedürftige, Aussage über das Verhalten eines Modells an einer bestimmten Gitterzelle, im Gegensatz zur Verwendung von reinen Distanzmaßen, die nur im Vergleich mit anderen Modellen eine, dann nur relative, Aussage erlauben. Problematisch an diesem Vorgehen ist jedoch der geringe Umfang der Beobachtungsdaten. Die Bestimmung von Konfidenzintervallen ist kaum statistisch abzusichern. Allerdings konnte dargelegt werden, dass unter den verfügbaren Datensätzen selbst auf der klimatologischen Ebene deutliche quantitative Unterschiede bestehen. Diese sind derart ausgeprägt, dass die Beschränkung auf einen einzigen Datensatz, im globalen Maßstab, als willkürlich aufgefasst werden kann. Auch eine bloße Mittelung löst dieses Problem nicht, da die Streuung der Beobachtungen nicht mehr berücksichtigt wird.

Insgesamt muss so wohl als ein Ergebnis dieser Analyse festgehalten werden, dass die verfügbaren Beobachtungsdaten nicht ausreichend genau für die vorgenommene Untersuchung sind. Anbieten würde sich dieses Vorgehen jedoch für eine andere Fragestellung. Auf maximal kontinentaler Ebene könnte die entwickelte und präsentierte Methodik als Entscheidungshilfe bei der Auswahl eines Teilensembles, das für den betreffenden Raum besonders überzeugende Ergebnisse liefert, angewendet werden, etwa der Argumentation von PERKINS et al. (2007) folgend. Abschließend erscheint auf globaler Ebene der rein quantitative Abgleich von Beobachtungs- und Modelldaten jedoch nicht als besonders gut geeignet für die Beurteilung der Qualität eines Modells.

## 4.2 Klimatypen in Klimamodellen

Das Klima eines realen Ortes oder einer Gitterbox in einem Klimamodell ist durch bestimmte Ausprägungen einzelner Klimaelemente gekennzeichnet, die separat betrachtet werden können. In der Regel, abgesehen etwa von allgemeiner Lebensfeindlichkeit in Folge extremer Temperaturen wie in der zentralen Antarktis, lässt sich jedoch mit einzelnen Klimaelementen die für den Menschen ausschlaggebende Charakteristik nicht befriedigend erfassen. Ebenso wenig ist das rein physikalische Klima auf diese Weise wiederzugeben. Stattdessen ist die Kombination verschiedener Größen erforderlich, um sich dem Klima als solches anzunähern, was sich auch in vielen Definitionen des Begriffs „Klima“ widerspiegelt (SCHÖNWIESE 2008, BLÜTHGEN und WEISCHET 1980). Die Klimaelemente Temperatur und Niederschlag ermöglichen bereits die Charakterisierung aussagekräftiger Klimatypen. Im Folgenden wird auf eine einfache Weise überprüft inwieweit moderne AOGCMs das Zusammenspiel der grundlegenden thermischen und hygrischen Größen abbilden.

### 4.2.1 Bedeutung der effektiven Klimaklassifikation nach Köppen

Klimaklassifikationen sind eine effektive Methode, kleinmaßstäblich das Klima unterschiedlicher Regionen systematisch zu vergleichen und klimatisch unterschiedliche bzw. ähnliche Regionen zu identifizieren und auszuweisen. Grundsätzlich bietet sich hierfür entweder deduktives oder induktives Vorgehen an. Der ältere Ansatz ist deduktiver Natur: anhand unterschiedlicher Kriterien werden eine Reihe von Klimatypen definiert, anhand derer sich Regionen mit vergleichbaren klimatischen Bedingungen ausweisen lassen. Traditionell wird zwischen „genetischen“ und „effektiven“ Klimaklassifikationen unterschieden. Ziel genetischer Klimaklassifikationen ist die Abbildung der Ursachen beobachteter klimatischer Unterschiede. Betrachtet werden vor allem der Strahlungs- und Wärmehaushalt und die atmosphärische Zirkulation. Die Abgrenzungskriterien für die einzelnen Klimaregionen stammen dementsprechend auch aus diesen Bereichen, bei Anwendungen auf das reale Klimasystem ergänzt durch tellurische und orographische Einflüsse (FLOHN 1957, HENDL 1991).

Bei dem bekannten und hier verwendeten System nach Köppen bzw. Köppen und Geiger handelt es sich dagegen um eine effektive Klimaklassifikation. Deren grundsätzliches Merkmal ist es, dass sie nicht die Differenzierung des Klimas als solches, sondern die Erklärung von klimagesteuerten Phänomenen und Prozessen zum Ziel haben. Zur Parametrisierung des Klimas können so auch durchaus unterschiedliche Klimatelemente für die räumliche Verbreitung der betrachteten Größen herangezogen werden. Köppens Klassifikation begründete diesen Ansatz, indem er das Ziel verfolgte quantitative Schwellwerte so zu bestimmen, dass diese nicht einfach quasi willkürliche Abgrenzungen darstellen sondern dass sich „erkennbare Beziehungen“ zu anderen Naturerscheinungen ergeben (KÖPPEN 1918). In diesem Fall wird die Vegetation und ihre typischen Verbreitung als Grundlage für die Schwellwertfestlegung herangezogen.

Die großskalige (natürliche oder potentiell-natürliche) Vegetation ist für viele Klimaklassifikationen als zu erklärende Größe gebräuchlich (etwa TROLL 1964, LAUER et al. 1996). Kartographische Darstellungen in kleinen Maßstäben, wie in diesen Fällen üblich, solcher Systeme entsprechen der von Köppen in weiten Teilen, was als Hinweis auf die Relevanz dieser Klassifikation gewertet werden kann. Das Gleiche gilt für topische Systematisierungen mit anderen thematischen Schwerpunkten (Diskussion bei SCHULTZ 2000). Neuere Ansätze verfolgen häufig einen anderen, induktiven Weg. Dabei werden über statistische Verfahren, insbesondere die Clusteranalyse, anhand verschiedener Klimaparameter „objektive Klimatypen“ ausgewiesen. Die Trennung der einzelnen Klimatypen ist, einem gewählten Kriterium folgend, neutral und exakt. Solcherart aus Daten abgeleitete Klimatypen verfügen jedoch über keine sinntragende Terminologie und somit Interpretation. Diese Verfahren sind erst durch die Entwicklung der Rechnerleistung und natürlich

durch die Verfügbarkeit flächendeckender Klimadaten – seien es ausreichend guter Beobachtungen oder Modelldaten – praktikabel geworden (z. B. PAETH et al. 2005, GERSTENGARBE et al. 1999).

Der Begriff Objektivität sollte im Zusammenhang mit solchen Methoden allerdings nicht überstrapaziert werden, da schon die Variablenauswahl mit einem gewissen Vorwissen getroffen wird und das Ergebnis entscheidend beeinflusst. Gleiches gilt für die Aggregierungsmethode, die bei den üblichen Verfahren das Ergebnis beeinflussen kann (WILKS 2011). Zudem werden in der Regel solche Klassifikationen als erfolgreich gewertet, die über eine gewisse Ähnlichkeit mit traditionellen Ansätzen verfügen.

Sowohl Köppen selbst als auch andere Autoren überarbeiteten die ursprüngliche Klassifikation im Laufe der Zeit mehrmals. Hierbei wurden einerseits Sonderklimate (etwa die tropischen Hochgebirge) in das System eingearbeitet, andererseits offenbarte die wachsende Anzahl zur Verfügung stehender Daten einige Schwachstellen. Eine Diskussion zur Geschichte und den verschiedenen Überarbeitungen findet sich bei HENDL (1991).

Obwohl mit der monatlichen Durchschnittstemperatur bzw. Niederschlagssumme die Datengrundlage für diese Klassifikation recht gering ist und mittlerweile andere Ansätze existieren, ist das System nach Köppen-Geiger immer noch das verbreitetste und gebräuchlichste seiner Art. Zum einen wird diesem Ansatz ein gewisser didaktischer Wert unterstellt, so dass der überwiegende Teil an einführenden Lehrbüchern zur Physischen Geographie (etwa HENDL 1997, BAUMHAUER 2011, STRAHLER 2004) das Köppen-Geiger System behandelt. Insbesondere in der angelsächsischen Grundlagenliteratur wird es häufig, wie in MCKNIGHT und HESS (2008) oder HOLDEN (2008), als einzige Systematik der globalen Klimate diskutiert, teilweise organisiert die Köppen-Geiger Klassifikation sogar sämtliche klimageographische Betrachtungen (etwa DE BLIJ et al. 2004, MARSH und KAUFMAN 2012). Außerdem findet sie sich natürlich ebenso in Literatur mit explizit klimatologischen Charakter (z. B. LAUER und BENDIX 2006, SCHÖNWIESE 2008, HANTEL 2001) und Lehrbüchern zur Meteorologie (z. B. KRAUS 2004, AHRENS 2007, MALBERG 2007).

Das Interesse an dieser aussagekräftigen Klassifikation zeigt sich aber auch durch der aktuellen Versionen der klassischen Karte auf neuer Datengrundlage (KOTTEK et al. 2006, PEEL et al. 2007). Ihre quantitative Natur ermöglicht zudem vielfältige Studien zum Klimawandel (z. B. GREGORY 1954, FRAEDRICH et al. 2001, BORN et al. 2008, RUBEL und KOTTEK 2010). Auch in solchen Anwendungen ist die Anschaulichkeit der Köppen-Klassifikation wohl der ausschlaggebende Grund für ihren Einsatz. Graduelle, quantitative Änderungssignale werden durch die Klimaformel in fassbare, qualitative Veränderungen übersetzt. Ähnlich ist ihr Nutzen als ergänzendes Verfahren bei der Validation von Klimamodellen (BORN et al. 2008, GUETTER und KUTZBACH 1990) begründet. Findet sich eine Übereinstimmung der simulierten und beobachteten Klimate, so ist dies eine

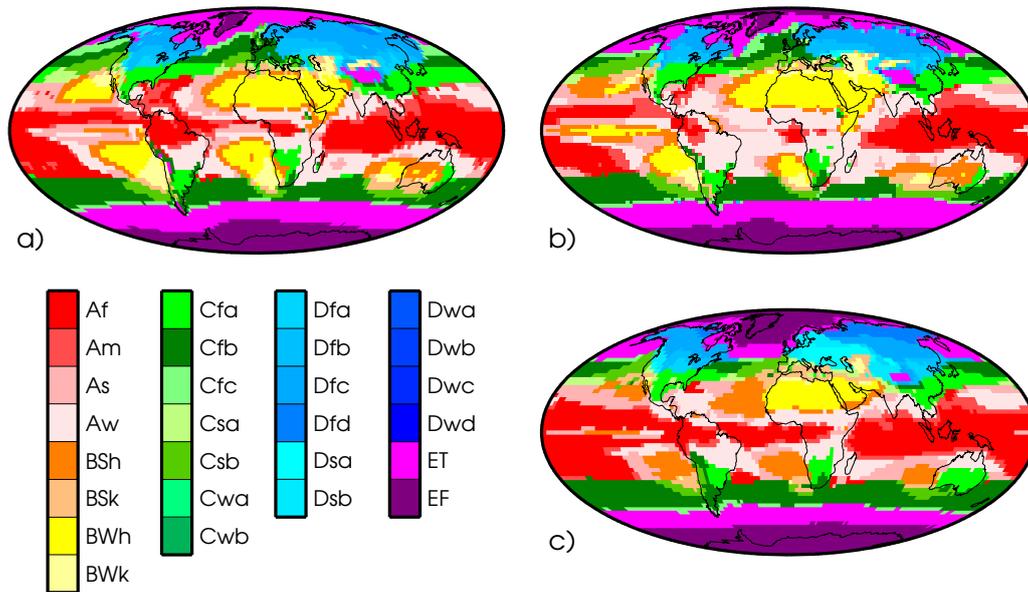
eindeutige, scheinbar einfache Aussage. In diesen kategorialen Variablen ist es, anders als im Fall der metrischen Klimatelemente, zudem möglich, nicht nur den tatsächlichen Zustand mit hoher Sicherheit bestimmen zu können, sondern auch eine Übereinstimmung von Beobachtungen und simuliertem Klima zu erzeugen. In diesem Fall lässt sich der Schluss ziehen, dass ein Modell die wesentlichen Aspekte des Klimas einer Region ausreichend gut abbildet.

Es ist durchaus üblich, nicht die komplette Klassifikation zu verwenden, sondern sich auf bestimmte Aspekte zu beschränken. Das können etwa die sogenannten Hauptklimatypen sein (zwei Buchstaben), teilweise werden aber auch nur die Hauptklimazonen (ein Buchstabe) oder beispielsweise nur der Ariditätsindex (GAO und GIORGI 2008) verwendet. In Anbetracht ihrer immer noch andauernden Verwendung lässt sich wohl feststellen, dass die Köppen-Geiger Klassifikation den Test der Zeit bestanden hat und immer noch besteht. Wegen ihrer weiten Verbreitung ist die Wiedergabe der Definition der Klimaformel an dieser Stelle sicher nicht nötig. Die hier verwendeten Schwellwerte finden sich in vielen der angegebenen Quellen, konkret wurde HENDL (1991) als Referenz herangezogen.

### 4.2.2 Ergebnisse

Mithilfe der kombinierten Daten für Niederschlag und bodennahe Temperatur wurde für jeden Gitterpunkt der Klimatyp bestimmt (Abb. 4.7). Obwohl die Klassifikation, wegen ihrer Bindung an die Vegetation, ursprünglich auf das Festland beschränkt ist, ist dies wegen den verwendeten Parametern problemlos möglich und wurde schon in frühen Auflagen so gehandhabt. Für das Ziel dieser Untersuchung erscheint das Ausklammern von 70 % der Erdoberfläche nicht ratsam. Trotz der deutlich niedrigeren räumlichen Auflösung und der leicht unterschiedlichen Basisperiode lässt sich eine gute Übereinstimmung mit der Karte von KOTTEK et al. (2006) feststellen. Auffällig ist lediglich das geringe Vorkommen des Steppenklimas. Da es sich bei diesem um eine schmale Übergangszone zwischen Wüstenklimaten und Regen- bzw. Schneeklimaten handelt, sollte angenommen werden, dass dies eine Folge der Interpolation darstellt. Die Klassifikation scheint insgesamt aber stabil gegenüber den Abweichungen innerhalb der Beobachtungsdaten, und die Nutzung der so erstellten Karte zum Abgleich mit den Modellen legitim. Hierbei wurde die komplette Klimaformel in der Version von KÖPPEN (1936) angewendet, auch wenn dies zu einer sehr großen Anzahl insbesondere von nur selten realisierten, aber theoretisch möglichen D-Klimaten führt (vgl. KRAUS 2004).

In Abbildung 4.7 werden Karten auf der Basis der kombinierten Beobachtungen und zweier Modellläufe gezeigt. Bei den beiden Vertretern des CMIP3 handelt es sich wiederum um ECHAM5/MPI-OM (1) und IAP-FGOALS-g1.0 (1). Der ECHAM5/MPI-OM Lauf weist im Vergleich mit den kombinierten Beobachtungen den größten übereinstimmenden Flächenanteil auf, IAP-FGOALS-g1.0 (1) liegt im Bezug auf diese Größe im unteren Quartil des Multi-Modell Ensembles. So-



**Abb. 4.7:** Klimatypen für Beobachtungen und AOGCMs. Jeweils komplette Klimaformel nach Köppen-Geiger für den Zeitraum 1960–1999. a) die kombinierten Beobachtungsdaten, b) ECHAM5/MPI-OM (1) und c) IAP-FGOALS-g1.0 (1).

mit können diese Läufe auch hier als eine Art Modellsread betrachtet werden (vgl. Abb. 4.6). Obwohl durchaus markante Unterschiede bestehen, kann festgehalten werden, dass die grundsätzliche Verbreitung der Klimatypen in allen drei Karten gut übereinstimmt. Für ECHAM5/MPI-OM (1) lässt sich lediglich im Bereich des Atlantiks und Afrikas die Unterschätzung des Af-Klimas, beruhend auf zu geringen Niederschlägen in diesen Regionen als qualitative Schwäche des Modells, feststellen. Dagegen zeigt Abb. 4.7 c) hier größere Ähnlichkeit mit den Beobachtungen. Der starke Temperatur-Bias (vgl. Abb. 4.5) des Modells führt allerdings zu einer übermäßigen Ausdehnung der Schneeklimare in West- und Mitteleuropa, also Abweichungen bereits auf der Ebene der Klimazonen. Statt Wüstenklimaten simuliert IAP-FGOALS-g1.0 (1) vor allem über den Ozeanen an den Ostseiten der Kontinente bevorzugt Steppenklimate, eine Folge der systematischen Überschätzung des Niederschlags in diesen Gebieten.

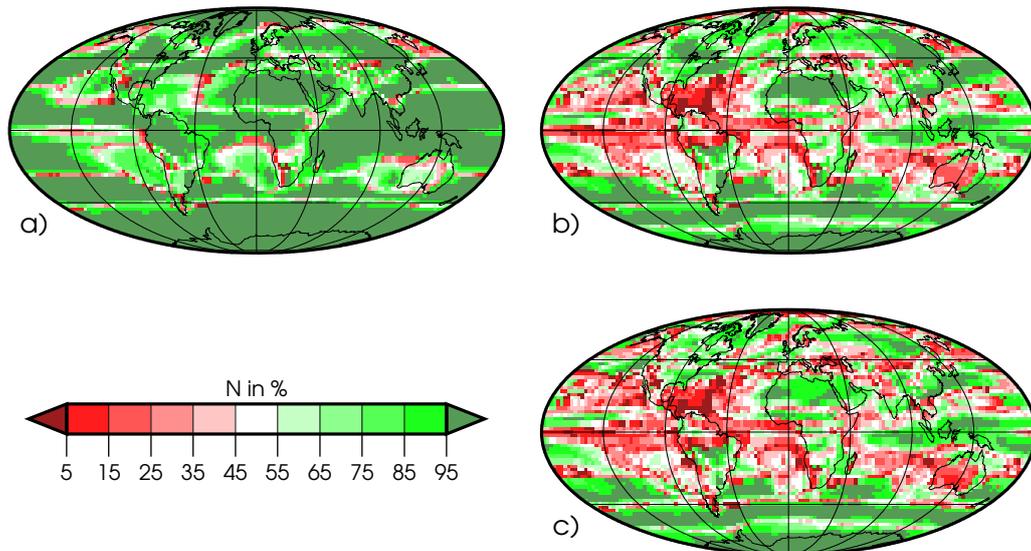
Offen bleibt hierbei jedoch die Frage, inwieweit diese Unterschiede zur Beurteilung der Qualität der Modelle herangezogen werden können. Der Eindruck der sich beim Vergleich der Leistungen dieser beiden Modellläufe ergibt gilt auch für das gesamte Multi-Modell Ensemble. In jedem Lauf finden sich Differenzen zu den Beobachtungen während in globaler Betrachtung die Verteilung der Klimatypen durchaus befriedigend wiedergegeben wird. Damit werden die klimatischen Rahmenbedingungen für die Biosphäre durch alle Modelle im Wesentlichen er-

folgreich reproduziert (Abb. 4.6). Die Bildung von Konfidenzintervallen oder die Quantifizierung der Differenzen im Vergleich zu den Beobachtungen ist wegen der kategorialen Skala der Klimaklassifikation nicht möglich. Stattdessen wurden die  $n = 48$  AOGCM-Läufe darauf untersucht, ob sie für die betreffende Gitterbox den gleichen Klimatyp  $T_i(s)$  wie die Beobachtungsdaten  $T_B(s)$  erzeugen. Für jede Gitterbox wird so die Größe  $N(s)$  berechnet.

$$N(s) = \sum_{i=1}^n f(T_i(s)) \quad \text{mit} \quad f(T_i(s)) = \begin{cases} 1 & T_i(s) = T_B(s) \\ 0 & T_i(s) \neq T_B(s) \end{cases} \quad (4.7)$$

Der Abgleich erfolgt für die drei möglichen Detaillierungsgrade. Werden nur die Klimazonen berücksichtigt zeigt sich, dass das Multi-Modell Ensemble für die nahezu gesamte Erdoberfläche mit den Beobachtungsdaten übereinstimmt. Merkliche Unsicherheiten innerhalb der AOGCMs treten nur in den Übergangsbereichen der Trockenklimate auf, was sich vermutlich durch Unterschiede in der Modellphysik, insbesondere aber in der räumlichen Auflösung erklären lässt. Der Bereich mit geringen Werten für  $N(s)$  im äquatorialen Ostpazifik ist an dieser Stelle wegen der großen Streuung in den Beobachtungsdaten (siehe Kapitel 4.1.1.2) ausgeklammert. Die sonstigen Klimazonen werden allein über die Temperatur definiert. In Anbetracht der geringen Auflösung von  $2.5^\circ \times 2.5^\circ$  kann das Relief in dieser Untersuchung nur reduzierten Einfluss aufweisen, was die geographische Breite in ihrem Einfluss weiter verstärkt. Somit sind hohen Übereinstimmungen für weite Teile der Erdoberfläche in Abb. 4.8 durchaus zu erwarten.

Die weniger ausgeprägte Übereinstimmung an den Übergängen der Klimazonen, schmale Bereiche mit  $N(s) \leq 50 \%$ , kann hier vernachlässigt werden. Zum einen ist der Flächenanteile dieser Bereiche nur gering, zum anderen handelt es sich um ein Artefakt aus der Definition der Köppen-Klassifikation: den abrupten Wechsel der Klimazonen aufgrund der vorgegebenen Schwellenwerte. Da diese Übergangsbereiche nur schmal sind, kann wohl festgehalten werden, dass das Multi-Modell Ensemble diese Problematik, die aus der Diskretisierung des Klimas erwächst, zufriedenstellend löst. Allerdings finden sich in Abb. 4.8 a) auch Hinweise auf Abweichungen zwischen Beobachtungen und Modellen, die durchaus auf Unterschiede in der Qualität der Modelle hinweisen könnten. So wird das Tibet-Plateau von vielen AOGCMs durchaus als Hochgebirge dargestellt, und hier finden sich entsprechend Unterschiede innerhalb des Multi-Modell Ensembles, die in der geringeren Anzahl an Übereinstimmungen mit den Beobachtungen resultieren. Die schmalen Zonen mit geringen Ähnlichkeiten im Vergleich zu den Beobachtungsdaten im Bereich von Nordatlantik und -pazifik lassen sich ebenfalls auch physikalisch interpretieren. In den Beobachtungsdaten findet sich hier der Einfluss von Meeresströmungen (Nordatlantikstrom bzw. Oyashio und Kuroshio im Pazifik), die Teilmglieder der thermohalinen Zirkulation darstellen. Die Abbildung dieses komplexen Systems



**Abb. 4.8:** Klimatypen im Multi-Modell Ensemble. Vergleich mit den kombinierten Beobachtungen für 1960–1999. Dargestellt ist  $N(s)$  (siehe Gleichung 4.7), für a) Klimazonen, b) Klimatypen und c) die gesamte Klassifikation

in dynamischen Modellen ist generell nicht so weit fortgeschritten wie die der Atmosphäre.

Wird die Untersuchung auf die Hauptklimatypen ausgeweitet, ergibt sich ein differenzierteres Bild. Dies ist wegen der Schwierigkeiten in der Modellierung des Niederschlags so auch zu erwarten. Unsicherheiten zeigen sich nun zusätzlich im Bereich des tropischen und subtropischen Atlantiks, aber auch für das Amazonasbecken und weite Teile Australiens werden nur noch geringe Werte für  $N(s)$  erreicht. Die Niederschlagscharakteristik der außertropischen Gebiete wird dagegen in besserer Übereinstimmung mit den Beobachtungsdaten abgebildet. Die Anwendung der kompletten Klimaformel hat in der globalen Betrachtung nur noch geringe Auswirkungen auf die Größe  $N$ . Zum einen liegt dies daran, dass eine solche Erweiterung für die A- und E-Klimate ohnehin keine Auswirkungen hat, zum anderen aber sicher auch an der Tatsache, dass der dritte Buchstabe in der Klimaformel wieder thermisch definiert ist. Somit sind die quantitativen Unterschiede von Modellen und kombinierten Beobachtungen zu gering, um sich angesichts der Schwellenwerte der Klassifikation noch entscheidend niederzuschlagen.

Nur in der, im Bezug auf dieses Merkmal, sehr heterogenen Gruppe der Schneeklimate, zeigt sich ein flächenmäßig relevanter Abfall von  $N(s)$  im Bereich Osteuropas und des westlichen Zentralasiens. Aus Abb. 4.6 geht hervor, dass, mit Ausnahme von NCAR-PCM alle Modelle das Klima von über 50 % der Erdoberfläche im Vergleich mit den kombinierten Beobachtungsdaten korrekt abbilden. Auch bei

dieser Größe wird der geringe Unterschied zwischen verschiedenen Läufen eines Klimamodells offensichtlich, wie auch im Zuge der bisherigen Ergebnisse nicht anders zu erwarten war.

Wegen der Definition von  $N(s)$  spielt in diesem Ansatz die Streuung innerhalb der Beobachtungsdatensätze kaum eine Rolle. Wie aus Abb. 4.8 ersichtlich, lassen sich einige Regionen identifizieren, in denen im Multi-Modell Ensemble eine besonders hohe Übereinstimmung mit den Beobachtungsdaten besteht. Insbesondere die großen Trockengebiete sind in jedem Modell gut abgebildet. Dagegen zeigen sich in den tropischen Regenklimate deutliche Unterschiede der Modellperformance. Auffällig ist zudem der geringe Grad an Übereinstimmung mit den Beobachtungsdaten im Bereich der Arktis, der auf der systematischen Unterschätzung der Temperatur im 20. Jahrhundert in diesem Gebiet durch die AOGCMs beruht. Für das 21. Jahrhundert ergibt sich jedoch, unabhängig vom vorgeschriebenen Emissionsszenario, in allen Klimamodellen ein überdurchschnittlicher Erwärmungstrend (IPCC 2007a). Obwohl es sich hier um eine relativ geringe Fläche handelt ist dies durchaus beachtenswert. Schließlich handelt es sich bei der Arktis um eine Region mit eigenständigen, äußerst empfindlichen Ökosystemen und enormer Bedeutung für das Klimageschehen der Nordhemisphäre sowie des globalen Klimasystems (ACIA 2005). Dennoch, und dies ein wichtiges Ergebnis für das Folgende, zeigt der Übergang von klimatischen Feldgrößen hin zu aussagekräftigen Klimatypen, dass das CMIP3 Ensemble die grundsätzliche beobachtete Klimadiversität, sowohl in ihrer räumlichen Anordnung als auch in ihrer ökologischen Relevanz, überzeugend abbildet. Dieser qualitative Befund liefert ein robustes Argument dafür, im Folgenden weiterhin sämtliche AOGCMs zu nutzen. Für den kompletten Ausschluss eines oder mehrerer GCMs erscheinen die Differenzen innerhalb des CMIP3 auch in dieser grundlegenden Qualität nicht ausreichend.

## 5 Evaluation der großskaligen Klimavariabilität

Statt dem einfachen Reproduzieren von bekannten Klimatologien wird die überzeugende Abbildung von bedeutsamen, charakteristischen Prozessen als Anforderung an die AOGCMs formuliert. Die realistische Erzeugung der natürlichen internen Variabilität des Klimasystems ist eine nicht-triviale Anforderung, die das Zusammenspiel vielfältiger Aspekte der gekoppelten Atmosphäre-Ozean Dynamik erfordert. Hier soll versucht werden, die Betrachtung auf die wichtigsten Moden im weltweiten Maßstab zu reduzieren, die allerdings erst erfasst und bestimmt werden müssen. Neben der Beurteilung der Modell-Performance erlaubt eine solche Auswahl auch die gezielte Untersuchung der Reaktion dieser Moden auf das anthropogene Forcing. Es wird im weiteren Verlauf dieser Arbeit zudem untersucht, ob ein Zusammenhang zwischen der Qualität der Erzeugung und dem Verlauf der Entwicklung der Moden der internen Klimavariabilität besteht. Es erscheint legitim zu unterstellen, dass solche Effekte, so sie existieren, potentielle Unsicherheitsquellen bei langfristigen Klimaprojektionen darstellen können. Somit kann die Analyse der internen Klimavariabilität auf vielfältige Weise einen Beitrag zur Untersuchung relevanter Fragestellungen leisten. Daher steht sie im Folgenden im Fokus.

### 5.1 Kriteriumsmoden: Definition und Begründung

Ein bedeutender Teil der beobachteten Variabilität des Klimasystems lässt sich durch verschiedene Zirkulationsmoden erklären. Derartige Größen existieren für Zeitskalen von einigen Wochen, wie etwa die Madden-Julian Oszillation im äquatorialem Pazifik, aber auch Aspekte der multidekadischen Variabilität werden teilweise auf diese Klimamoden zurückgeführt. Ein Beispiel für derartig niederfrequente Moden ist die multidekadische Variation der Meeresoberflächentemperaturen im Nordatlantik. Ebenso unterschiedlich sind die räumlichen Maßstäbe der einzelnen Moden, die bis hin zu globalen Auswirkungen reichen. Diese als Zirkulationsmoden identifizierten Phänomene beschreiben vielfältige Kopplungen innerhalb des Klimasystems und seiner Subsysteme. Die erfolgreiche Simulation dieser Moden kann somit zur Überprüfung der Fähigkeit zur Abbildung der komplexen Beziehungen

und Interaktionen, die das reale Klimasystem auszeichnen, durch AOGCMs dienen.

Es erscheint durchaus folgerichtig Modellen, die hierbei überzeugende Ergebnisse nachweisen können, auch im Hinblick auf die zukünftige Entwicklung des Klimas realistische Abschätzungen, etwa auch bei der regionalen Differenzierung des Klimasystems, zuzubilligen. Solche Modelle bilden schließlich wesentliche Bereiche der beobachteten Dynamik realistisch ab, was sie auch zu einer realistischen Reaktion auf veränderte Randbedingungen befähigen sollte. Hingegen könnten Modelle, denen diese wesentlichen Aspekte des Klimasystems fehlen, auch Schwierigkeiten dabei haben, das anthropogene Signal auf angemessene Weise umzusetzen. Diese Überlegung erhält dadurch Unterstützung, dass erwartet werden kann, dass weite Teile des künftigen Klimawandels durch eine relativ kleine Anzahl von Zirkulationsmoden umgesetzt werden.

Somit wird hier überprüft ob, und wenn ja wie, sich die Qualität von AOGCMs über deren Fähigkeit zur Abbildung der organisierten, großskaligen Variabilität beurteilen und vergleichen lässt. Ein bedeutender Vorteil dieses Ansatzes ist die relativ geringe Zahl von Größen die hierbei beachtet werden müssen. Dieser Aspekt ist im Hinblick auf die Praktikabilität, insbesondere bei der Analyse von Multi-Modell Ensembles, von tatsächlicher Relevanz. Daneben ist im Hinblick auf diese Phänomene des Klimasystems das Phänomen des Tunings der Modelle auf eine bestimmte Klimatologie kaum vorhanden.

Ein bedeutendes Problem beim Versuch, die Leistungen der CMIP3-Modelle über die Simulation der bedeutsamsten Zirkulationsmoden zu bewerten, besteht darin, eben jene zu identifizieren. Hierfür lässt sich eine Vielzahl an Kriterien finden, je nachdem welche Aspekte als besonders bedeutsam befunden werden. Eine erste Einschränkung an in Frage kommenden Moden ergibt sich aus den zur Verfügung stehenden Daten. So werden alle intrasaisonalen oder noch höherfrequenten Zirkulationsmuster ausgeschlossen, da deren zufriedenstellende Erfassung anhand von monatlichen Daten als nicht unbedingt gesichert erscheint. Daneben liegen für die, teilweise nur vermuteten oder konstatierten, multidekadischen Moden zu wenige Beobachtungsdaten vor, um eine tatsächliche Überprüfung der AOGCMs zur ermöglichen. Bestenfalls könnten bei einem Modell niederfrequente Varianzanteile festgestellt werden, die mit einer solchen Mode assoziiert sein könnten. Eine quantitative Erfassung einer solchen Aussage und insbesondere der Vergleich zwischen verschiedenen derartigen Feststellungen erscheint als schwer möglich und wenig zielführend.

Diese beiden Beschränkungen bedeuten, dass die Zirkulationsmuster auf die sich die weitergehende Untersuchung stützen wird, unter jenen zu finden sein müssen, deren Auswirkungen sich auf der interannuellen Skala gut erfassen lassen. In diese Gruppe fallen eine große Anzahl von verschiedenen Moden mit weitreichenden und im Wesentlichen als stabil befundenen und gut dokumentierten Auswirkungen auf weite Gebiete der Erdoberfläche (vgl. BARRY und CARLETON 2001, BRIDGMAN et al. 2006). Das Wissen über die Entwicklung dieser Moden im Klimawandel

würde bereits eine bedeutende Reduktion der Unsicherheit in Fragen des regionalen Klimawandels bedeuten. Insofern ist die Konzentration auf diese Gruppe von Zirkulationsindices sicher gerechtfertigt.

Deutlich problematischer ist es dagegen auszuwählen, welche dieser Moden zur Überprüfung der AOGCMs herangezogen werden können. Sicher erscheint allein, dass nicht jeder jemals identifizierte und publizierte Index zur Beschreibung und Erfassung der internen Variabilität regionaler Aspekte des Klimasystems hierfür herangezogen werden kann und sollte. Die aktive Auswahl dieser Moden ist daher unumgänglich. Gleichzeitig stellte sie einen ebenso wichtigen wie aufwendigen Teil dieser Arbeit dar. Zu diesem Zweck wurden einige Kriterien formuliert, um die Entscheidung für oder gegen verschiedene Muster zu fällen. Die Moden sollten stabil und reproduzierbar in den Beobachtungsdaten zu erfassen sein. Insbesondere sollte keine Abhängigkeit von der Verwendung eines bestimmten Datensatzes bestehen. Daneben sollten die verwendeten Eigenmoden möglichst große Anteile der Variabilität des globalen Klimasystems repräsentieren.

## 5.2 Globale Detektion: Methodik

Hier wird die Vorgehensweise zur Bestimmung der wichtigsten Zirkulationsmoden im Klimasystem erläutert. Das Kriterium, möglichst viel Variabilität mithilfe der Moden erklären zu können, legt den Einsatz der EOF-Analyse nahe, welche zudem die übliche Methode zur Identifikation von Zirkulationsmustern darstellt. In vielen Fällen können die führenden EOFs direkt als verschiedene Klimamoden identifiziert werden, sofern die räumliche Domäne es zulässt (z. B. PAXIAN et al. 2013). Eine hier zudem als besonders wichtig erachtete Eigenschaft dieses Verfahrens ist es auch, dass sich so, über die resultierenden Raummuster, die Möglichkeit zum quantitativen Vergleich von Beobachtungs- und Modelldaten, ebenso wie zum Vergleich verschiedener AOGCMs ergibt.

Sowohl die routinemäßige Nutzung zur Identifikation von bedeutenden Zirkulationsmoden als auch die Möglichkeit zum Abgleich von Modell- und Beobachtungsdaten erklären sich dadurch, dass mit der EOF-Analyse sowohl die zeitliche auch die räumliche Struktur der Variabilität der Ausgangsdaten erfasst werden kann (vgl. Kapitel 3.5). So ist zwar die Zeitdomäne bei der Identifikation der Klimamoden in einem GCM völlig nutzlos, das charakteristische Raummuster eines Prozesses ist jedoch sowohl bei der Betrachtung von Beobachtungsdaten als auch von Modelldaten unabhängig vom konkreten Verhalten innerhalb einer Periode eindeutig zu erfassen (vgl. Kapitel 6). Voraussetzung hierfür ist allerdings, dass ein Prozess, hier ein Zirkulationsmuster, tatsächlich existiert und im betrachteten Maßstab relevanten Einfluss auf die Variabilität der Ausgangsdaten hat.

Klimamoden, auf die diese Voraussetzungen zutreffen, sind entsprechend der erste Schritt bei der Beurteilung der Modellqualität auf diese Weise. Zur Identifikation der

wichtigsten Variabilitätsmoden des Klimasystems wurde eine Vielzahl verschiedener Datensätze mit unterschiedlichen Varianten der EOF-Analyse untersucht.

Für eine, was die Spannbreite der verfolgten Ansätze betrifft durchaus repräsentative, Auswahl finden sich im Folgenden kurze Erläuterungen für den Aufbau der Analyse und deren Ergebnisse. Aus den hierbei verwendeten Beobachtungsdaten wurde jeweils vor der Analyse der lineare Trend entfernt, um zu verhindern, dass bereits realisierte anthropogene Einflüsse die Detektion der Muster der natürlichen Klimavariabilität verhindern.

Da die Raummuster in der globalen Domäne bestimmt werden, ist der direkte Vergleich mit verbreiteten und allgemein anerkannten Referenzen nur eingeschränkt möglich, der quantitative überhaupt nicht. TRENBERTH et al. (2005) nutzen EOF-Analysen von ERA 40 (Felddaten der atmosphärischen Masse) in einem globalen Ansatz und erfassen damit SAM, NAM und ENSO. Diese Größe stand für die unternommen Analysen allerdings nicht zur Verfügung. Stattdessen musste das Raummuster subjektiv bewertet werden. Hierbei wurden die (R)EOFs als mögliche Repräsentationen bekannter Klimamoden im jeweiligen Datenfeld betrachtet, wenn diese ein plausibles Raummuster aufwiesen. Angesichts dieser wenig konkreten Vorgehensweise im Fall des charakteristischen Raummusters ist ersichtlich, dass die Zeitdomäne bei der Identifikation der Klimamoden einen wichtigen Anteil hat. Die Entscheidung, ob eine hierdurch gewonnene Hauptkomponente die zeitliche Entwicklung eines bekannten Zirkulationsmusters beschreibt und dieser somit im globalen Feld detektiert werden konnte, erfolgt auf der Basis von Korrelationsanalysen. Hierfür werden sowohl die in Kapitel 2.1.2.1 beschriebenen Zeitreihen als auch solche, die über klassische Indexdefinitionen aus Felddaten gewonnen wurden, verwendet. Letzteres betrifft vor allem die Monsune, für die keine Zeitreihen aus unabhängigen Daten zur Verfügung standen. Die hierbei entstehende besondere Problematik wird in Kapitel 6.2.2 thematisiert. Es wurden hohe Korrelationskoeffizienten angestrebt, da die Bewertung der AOGCMs letztlich auf einer physikalischen Begründung fußen sollte und nicht auf überinterpretierten Effekten einer Hauptachsentransformation. Weiterhin ist die zukünftige Entwicklung dieser Indexreihen von großem Interesse. Daher muss für folgende Analysen sichergestellt sein, dass die verwendeten Zeitreihen das entsprechende Phänomen auch ausreichend zuverlässig erfassen.

Vom technischen Standpunkt aus ist die einzige Größe, die bei den in diesem Kapitel präsentieren Untersuchungen verändert wird, die Datenmatrix  $\mathbf{X}$ . Die einzige Variation in der Methodik ist die Entscheidung, ob die Eigenvektoren im Anschluss an ihre Extraktion noch nach der Varimax-Methode rotiert werden. Wenn dies der Fall war, wurden jeweils so viele der führenden EOFs rotiert, wie erforderlich waren um mindestens 80 % der Gesamtvarianz der Ausgangsdaten zu repräsentieren. In Tab. 5.1 sind für einige aussagekräftige Experimente Ergebnisse aufgeführt. Alle dort aufgeführten Analysen wurden mit monatlichen Anomalien durchgeführt. Diesen wurde sowohl der Jahresgang als auch ein eventuell vor-

handener linearer Trend entfernt, um sicherzustellen, dass kein deterministisches externes Signal die interne Variabilität in den Daten maskiert. Es wurde jeweils 1960–1999 als Basisperiode analysiert. Bei weniger eindeutigen Ergebnissen wurden die EOF-Analyse für die beiden Teilzeiträume 1960–1979 und 1980–1999 separat erneut durchgeführt. Saisonale Werte der PCs wurden durch einfache Mittelung der entsprechenden monatlichen Amplituden der (R)PCs berechnet. Die EOF-Analyse von saisonalen Werten brachte insgesamt nicht genug Vorteile, um diesen Ansatz schließlich für die weiteren Untersuchungen heranzuziehen. Aus diesem Grund wird hier auf die Darstellung so gewonnener Ergebnisse verzichtet.

Die Experimente 1.1, 1.2, 1.3 und 1.4 wurden jeweils mit SLP-Daten durchgeführt. Bei 1.1 und 1.3 wurde die Kovarianzmatrix, bei 1.2 und 1.4 jeweils die Korrelationsmatrix betrachtet. Weiterhin ist bei den Experimenten 1.3 und 1.4 eine Varimaxrotation im Anschluss an die EOF-Analyse durchgeführt worden.

In Experiment 2.2 wurde die relative Topographie der Reanalysedaten über eine Korrelationsmatrix untersucht. Die verwendete Schichtdicke ergibt sich aus der Differenz der geopotentiellen Höhen des 500 hPa- und des 1000 hPa-Niveaus.

Bei 4.2 und 4.4 wurden die Felddaten für Luftdruck und bodennahe Temperatur gemeinsam untersucht. Die EOF-Analyse wurde mit der Korrelationsmatrix durchgeführt. Im Fall von Experiment 4.4 wurden die führenden EOFs zusätzlich nach der Varimax-Methode rotiert.

Die Experimente 5.1 und 5.2 wurden mit der Höhe des 500 hPa-Niveaus durchgeführt. Sowohl die EOF-Analyse der Kovarianzmatrix (5.1) als auch der Korrelationsmatrix (5.2) konnte hierbei vergleichbar gut Klimamoden identifizieren wie im SLP-Feld. Auch konnte hier ebenso in einigen Fällen über die Rotation der Eigenvektoren (in Tab. 5.1 nicht gezeigt) die Trennung zwischen den verschiedenen Moden verbessert werden.

Diese gute Übereinstimmung der Experimente 1.1–1.4 und der entsprechenden Verfahren für die Höhe des 500 hPa-Niveaus ließen die Betrachtung der Kombination der beiden Datengrundlagen als eine Methode erscheinen, die Ergebnisse weiter zu verbessern. Um die Verwendung der Kovarianzmatrix zu ermöglichen, wurde hierfür anstelle des SLP die geopotentielle Höhe des 1000 hPa-Niveaus mit der für 500 hPa kombiniert. Diese Untersuchungen haben in der Tabelle die Nummern 6.1–6.4, wobei die Benennung die Details der Analyse entsprechend der Luftdruckdaten beschreibt. Hierbei wurden allerdings tatsächlich weniger eindeutige Ergebnisse erreicht als bei der Beschränkung auf nur einen Datensatz.

### 5.3 Ergebnisse

Der Großteil der in Betracht gezogenen Moden konnte bei keiner der durchgeführten Analysen erfasst werden. Für diese Zirkulationsmuster wird daher angenommen, dass sie bei globaler Betrachtung keine der wesentlichen Komponenten des Klima-

**Tab. 5.1:** Global erfassbare Klimamoden. Klimamoden, die in der globalen Domäne über verschiedene Ansätze detektiert werden konnten. Für Erläuterungen und Definition der nummerierten Experimente siehe Text. Alle Daten entweder HadSLP2, Era 40 Reanalysen oder NCEP/NCAR Reanalysen. **H/E/N**: eine (R)EOF  $v_k(s)$  mit plausibler räumlicher Struktur identifizierbar und für Korrelation der (R)PC  $c_k(t)$  mit dem Standardindex  $I(t)$  für die monatlichen Werte gilt  $r \geq 0.6$  und für mindestens zwei Jahreszeiten  $r \geq 0.8$ . **H/E/N**:  $v_k(s)$  mit plausibler räumlicher Struktur identifizierbar und Korrelation von  $c_k(t)$  mit  $I(t)$  für die monatlichen Werte  $r \geq 0.4$  und für mindestens eine Jahreszeit  $r \geq 0.8$ ; **h/e/n**: mehrere (R)EOFs mit plausiblen Raummuster und (oder) mind. eine (mehrere) (R)PCs mit Korrelationen mit den Referenzreihen von  $0.4 \leq r < 0.6$  bei monatlichen Daten und mindestens einer Jahreszeit mit  $r \geq 0.7$ ; **h/e/n**: mehrere plausible Raummuster und (oder) mind. eine (mehrere) (R)PCs korrelieren mit der Referenzreihe mit  $r > 0.4$  für Monatswerte oder mindestens eine Jahreszeit mit  $r > 0.7$ . -: Analyse nicht möglich wg. Datenverfügbarkeit; x: Klimamode nicht detektierbar;

Exp.	SOI	SAM	NAM	SCA	NPI
1.1	H e n	H E N	H E N	h x x	h E N
1.2	h e n	H E N	H E N	h e n	H E N
1.3	H e n	H E N	H E N	H E N	H E N
1.4	h e n	H E N	H E N	h E n	h E N
2.2	- e n	- E N	- e N	- x x	- e n
3.1	- e n	- E N	- E N	- x n	- E N
4.2	- e n	- E N	- E N	- E x	- e n
4.4	- e n	- e N	- e n	- x x	- x x
5.1	- e n	- E N	- E N	- E N	- E N
5.2	- e n	- E N	- E N	- e n	- E N
6.1	- e n	- E N	- E N	- e n	- E N
6.2	- e n	- E N	- E N	- e N	- e n
6.3	- e n	- E N	- E N	- x n	- E N
6.4	- e n	- E N	- E N	- E N	- E N

systems darstellen. Grundsätzlich abgesprochen werden sollen ihnen aber weder ihre physikalische Natur noch ihre Klimawirksamkeit. Auch die Zirkulationsmuster, auf die sich die folgenden Untersuchungen beschränken, sind keinesfalls in sämtlichen durchgeführten Ansätzen zu detektieren gewesen. Angesichts der unbeschränkten Möglichkeit, die grundsätzliche Methode von globalen EOF-Analysen

zu variieren, kann sicherlich nicht ausgeschlossen werden, dass einige der hier nicht erfassten Zirkulationsmuster bei Verwendung einer anderen Datenbasis auch in der globalen Domäne zu erfassen wäre. Daneben ist natürlich die EOF-Analyse zwar ein wichtiger und, seit der ubiquitären Verfügbarkeit von leistungsstarken Rechner und flächendeckenden Datensätzen, der gebräuchlichste Ansatz zur Detektion und Analyse von Zirkulationsmoden, aber sicher nicht der einzige.

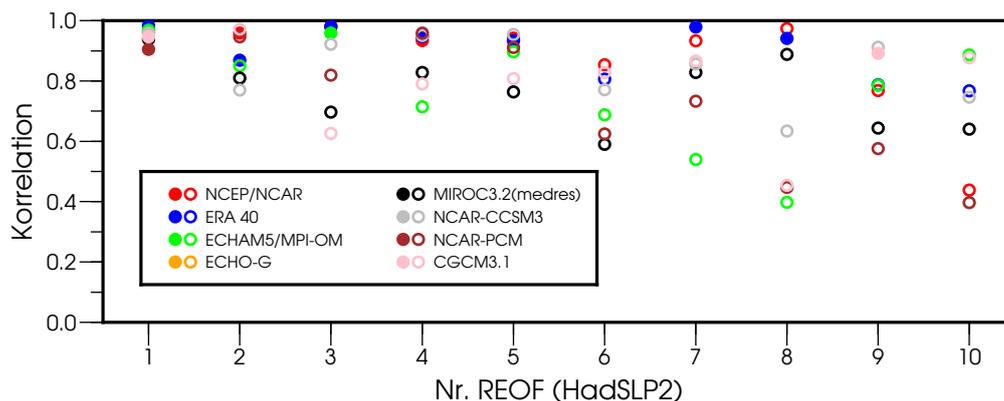
Als nicht erfassbar, weder durch die in Tab. 5.1 aufgeführten Ansätze noch bei der Analyse anderer Ausgangsdaten oder der Abkehr von der Verwendung von Monatswerten, erwiesen sich EAP, WPP, EP-NPP, PEP, sowie die Monsune. Für die weitere Verringerung der Anzahl der zu betrachteten Muster wurden einige Indices, die offensichtlich das gleiche physikalische Phänomen beschreiben, zusammengefasst. Die hiervon betroffenen Indices sind zum einen, wie zu erwarten war (WALLACE 2000), der NAM und die NAO. Ebenfalls im Verhalten der Indices einander sehr ähnlich sind NPI und PNA, die beide vom Aleutentief dominiert werden (TRENBERTH und HURRELL 1994). SCA und EA/WR in den aus globalen EOF-Analysen resultierenden Raummustern zu unterscheiden ist in den meisten Fällen kaum möglich und insbesondere nicht objektiv begründbar.

Alle diese Indexpaare weisen in den Beobachtungsdaten, die durch die jeweilige Analyse erfasst werden, ausgeprägte signifikante Korrelationen mit der gleichen PC auf. Auch die Raummuster sind für diese Moden kein eindeutiges Unterscheidungsmerkmal. Gerade die physikalische Einheit von NAM und NAO ist zudem im wesentlichen Konsens (THOMPSON et al. 2003, HURRELL und DESER 2009, TRENBERTH et al. 2005). Als durchaus subjektive Entscheidung wurden schließlich die Bezeichnungen NAM, NPI und SCA übernommen. Begründet ist das damit, dass bei letztlich verwendeten Indexdefinition die Korrelationen für diese Zeitreihen mit den PCs der rotierten EOFs des SLP höhere Beträge annahmen, insbesondere während des Winters, wenn die Indices ihre stärkste und deutlichste Ausprägung haben. Auch das für diesen Ansatz resultierende Raummuster kann durchaus stärker in Richtung der jeweils gewählten Zirkulationsmode interpretiert werden (vgl. Kapitel 6.1.1). Hier wird allerdings dennoch der Standpunkt vertreten, dass in allen Fällen die Entscheidung nicht zwingend war und sicherlich auch anders hätte ausfallen können, ohne die übrigen Schritte zu beeinträchtigen. Der entscheidende Aspekt der Zirkulationsmuster, die zugrunde liegenden physikalischen Prozesse, sind von der verwendeten Terminologie natürlich in keiner Weise betroffen. Die insgesamt besten Resultate im Hinblick auf die vier zu berücksichtigenden Aspekte

- Anzahl der detektierten bekannten Klimamoden,
- Interpretierbarkeit des Raummusters,
- Korrelation mit den Referenzindices sowie insbesondere
- Reproduzierbarkeit über verschiedene Beobachtungsdatensätze

konnten für den auf Meeresebene reduzierten Luftdruck erzielt werden, wenn das Analyseschema des Experiments 1.3 angewendet wurde. Es kann festgehalten werden, dass insbesondere die ersten fünf rotierten Eigenvektoren der Beobachtungsdaten räumlich sehr gut korrelieren (Abb. 5.1). Die teilweise ebenfalls sehr hohen Korrelationen für höhere REOFs werden im Folgenden nicht mehr beachtet, da sich für diese aus der vorherigen Untersuchung keine Interpretation als Eigenmode des Klimasystems ergab.

Die führenden Raummuster finden sich zudem nicht nur in den Beobachtungsdatensätzen sondern sind auch in den AOGCMs vorhanden, wie sich anhand einer zufällig ausgewählten Stichprobe erkennen lässt. Für diese Auswahl an Modellen konnte bei keinem der in Tabelle 5.1 aufgeführten Ansätze bessere Übereinstimmung mit einer so großen Anzahl an Raummustern erfasst werden. Ebenso wenig gelang dies bei anderen, hier nicht vorgestellten Ansätzen. Durch die modellübergreifende Reproduktion der als Klimamoden identifizierten Raummuster wird deren Interpretation als Ausdruck von dominanten physikalischen Prozessen gestärkt. Ein weiteres Indiz hierfür ist, dass die Spannbreite der Korrelationen bei höheren REOFs deutlich zunimmt (Abb. 5.1).



**Abb. 5.1:** Raummuster in Beobachtungen und Modellen. Jeweils paarweise Korrelation der führenden  $i = 1, 2, \dots, 10$  REOF vom HadSLP2 ( $v_{i,H}(s)$ ) mit ERA 40 und NCEP/NCAR sowie einigen AOGCMs (jeweils Lauf (1) der angegebenen Modelle). Gefüllte Kreise bedeuten, dass die angegebene Korrelation zwischen  $v_{i,H}(s)$  und  $v_{j,X}(s)$  mit  $X$ =Vergleichsdatensatz für  $i = j$  gilt. Leere Kreise zeigen eine höhere Korrelation für  $i \neq j$  mit  $j = 1, 2, \dots, 20$  an.

Für höhere REOFs ist zwar die Reihenfolge nicht mehr identisch, dennoch finden sich die im HadSLP2 erfassten Muster in NCEP/NCAR und ERA 40 in ganz wesentlichem Maße wieder. Allerdings ist hier ein wichtiger Punkt festzuhalten. Die 2. RPC, für HadSLP2r findet sich hier ein wesentlicher Teil der Dynamik des SOI (für Monatswerte  $r^2 \approx 0.4$ ; für den borealen Winter  $r^2 \approx 0.65$ ), kann für die

beiden anderen Datensätze sicher nicht analog als Zeitreihe des SOI interpretiert werden. Das ENSO-Phänomen ist in den Reanalysen nicht über eine einzelne (R)EOF/(R)PC-Kombination zu beschreiben. FELDSTEIN (2000) berichtet, allerdings in Folge eines andersartigen Ansatzes der zudem auf die Nordhemisphäre beschränkt ist, ebenfalls von wenig eindeutigen Ergebnissen bei der Detektion des ENSO-Musters anhand von REOFs und RPCs der geopotentiellen Höhe des 300 hPa Feldes.

Insgesamt lieferte dieser konzeptionell sicher nachvollziehbare, im Wesentlichen quantitative und reproduzierbare Weg zur Bestimmung der besonders relevanten Klimamoden somit nur teilweise befriedigende Ergebnisse. Im globalen Maßstab gelang es nicht, unzweifelhaft bedeutsame Moden des Klimasystems in den Tropen zu erfassen. Im Fall des ENSO-Phänomens waren die Ergebnisse der Beobachtungsdaten zu unterschiedlich, die Monsune ließen sich auch mit sehr gezieltem Vorgehen nicht befriedigend erfassen. Da die Auswirkungen eben dieser Klimamoden jedoch in ganz besonderem Maße die Menschen der betreffenden Regionen in ihrer Lebens- und Wirtschaftsweise beeinflussen, erschien ihre Ausklammerung aus der Analyse wenig ratsam. Stattdessen wurden für die tropischen Moden, letztlich pragmatisch, separate Ansätze zur Erfassung entwickelt und genutzt (Kapitel 6.2).

---

## 6 Index-Definitionen

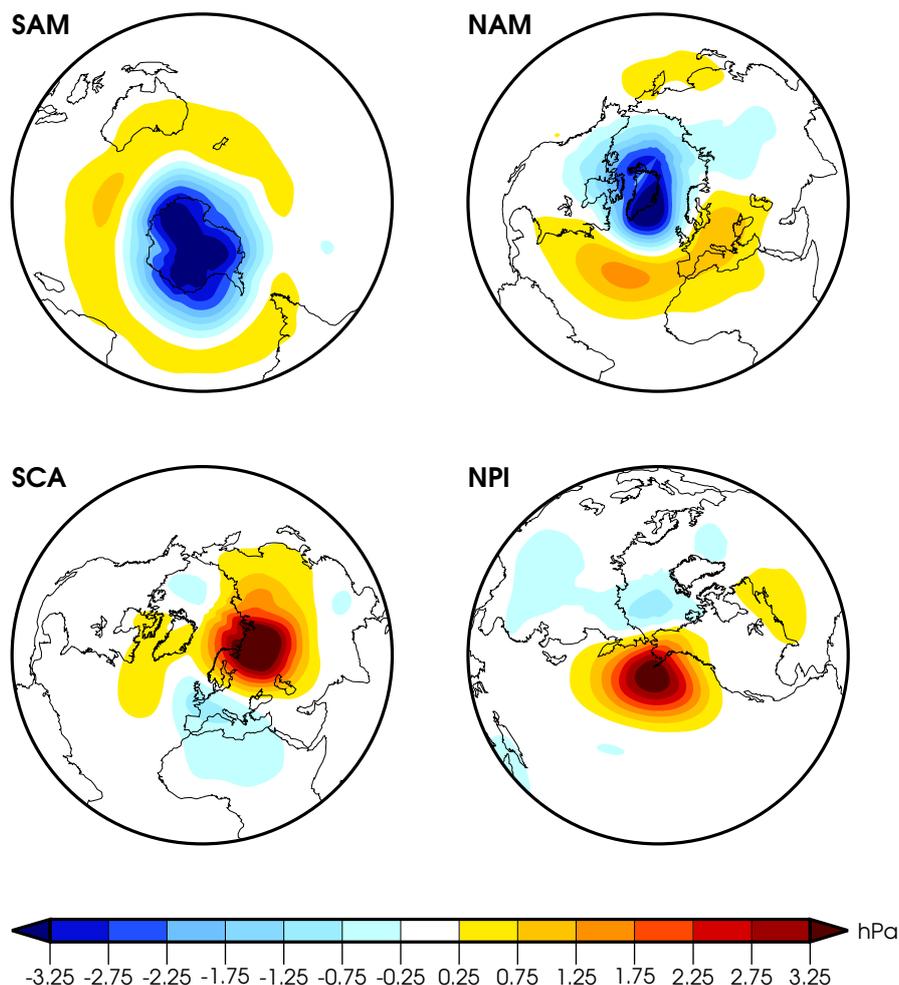
In diesem Kapitel werden die im Weiteren verwendeten Indices vorgestellt und ihre Definition bzw. Berechnung erläutert.

### 6.1 Außertropische Zirkulationsmuster

Die Ergebnisse aus Kapitel 5 zeigen, dass die globalen REOFs der monatlichen Anomalien des SLP die wesentlichen ektropischen Zirkulationsmoden gut erfassen. Die räumlichen Variationsmuster sind über die verschiedenen Beobachtungsdatensätze und auch Modelldaten derart stabil, dass die Verwendung eines gemeinsamen Raummusters zur Berechnung der Zeitreihen für die verschiedenen Datensätze möglich erscheint. Eine gemeinsame Basis ist auch hilfreich bei der Detektion von eventuell bedeutsamen Modellunterschieden. Hierbei wird angenommen, dass die individuellen Modellunterschiede durch die Wahl der globalen Domäne nur einen relativ geringen Einfluss auf die Struktur des Raummusters haben sollten. Angesichts des Aufwandes der zur Erzeugung von REOFs und RPCs nötig ist, ist ein solches Vorgehen zudem von praktischer Seite aus vorteilhaft und naheliegend.

#### 6.1.1 Raummuster

Die Berechnung der vier global relevanten Moden der außertropischen Zirkulation erfolgt unter Verwendung des Raummusters des HadSLP2-Datensatzes. Dieser Datensatz wurde ausgewählt, da er, im Gegensatz zu den Reanalysen, nicht von einem Zirkulationsmodell erzeugt wurde. Somit können Artefakte, die sich durch die dynamische Modellierung ergeben könnten, ausgeschlossen werden. Jedoch ergab eine ausführliche Prüfung, dass die Indexdefinitionen nicht sensitiv gegenüber dem Ursprung der Raummuster sind. Werden die SLP-Felder der beiden Reanaly-sedatensätze auf die hier verwendeten REOFs projiziert, ändert sich in keinem Fall die Korrelation der so erzeugten Zeitreihen mit den Referenzindices signifikant. Das Gleiche gilt für die Korrelationskoeffizienten mit den nach der Standarddefinition aus den SLP-Daten (siehe Tab. 6.1) berechneten Indexzeitreihen. Unabhängig davon welche globale Raummuster, die REOFs von HadSLP2, dem SLP von Era 40 oder dem SLP von NCEP/NCAR, konkret verwendet wird um die Zeitreihen zu generieren sind die Korrelationskoeffizienten mit diesen unterschiedlichen Referenzzeitreihen sehr hoch. Für die Periode 1960–1999 betragen sie, sowohl für



**Abb. 6.1:** Ektropische, global relevante Zirkulationsmuster. Die Muster sind globale monatliche REOFs der HadSLP2-Daten, vgl. Text. Die Darstellung folgt im Schema den üblichen Darstellungen dieser Moden, um die gute Abbildung der bekannten Raummuster zu betonen. Sie fokussiert daher jeweils die Aktionszentren. In dieser Skalierung zeigen die Muster jeweils die Anomalien für einen Wert der entsprechenden, standardisierten RPC von  $\tilde{c}_i(t) = +1$ .

Monatsdaten als auch saisonale Zeitreihen, immer  $r > 0.88$ . Diese Werte liegen deutlich über den von TRENBERTH et al. (2005) erreichten.

Nachdem bislang der Fokus der Darstellung auf dem quantitativen Abgleich der Zeitdomäne lag, werden hier die schlussendlich ausschlaggebenden Raummuster knapp vorgestellt. Die 1. REOF ist das Raummuster des SAM. Der kreisförmige Charakter dieser Mode ist in Abb. 6.1 klar ersichtlich. Die dort gezeigte Struktur ist mit der der Standarddefinition quasi identisch. Intensiviert sich der Druckgegensatz

zwischen den polaren Regionen und den Subtropen, so äußert sich dies im SLP-Feld mit sinkenden Werten über der Polarkalotte bis in die subpolaren Gebiete und entsprechend steigenden Luftdruck im Bereich der Subtropenhochs. Angezeigt wird dies durch einen positiven Indexwert der RPC, die die Zeitreihe des SAMs darstellt. Es ist deutlich erkennbar, dass hierbei die polwärtige Druckveränderung der dynamischere Part ist.

Das nordhemisphärische kreisförmige Muster, der NAM, findet sich in der 3. REOF wieder. Die Struktur ist weniger klar und schematisch wie die auf der Südhalbkugel, jedoch gilt dies auch für das Raummuster der Standarddefinition. Es zeigt sich sehr deutlich, dass die dominierenden Aktionszentren des NAM, Islandtief und Azorenhoch, im Muster klar erfasst sind. Insofern ist es sicherlich auch vertretbar, dieses Muster als NAO anzusprechen. Hier wird allerdings die Bezeichnung NAM für dieses physikalische Phänomen beibehalten, um die deutlich erkennbare zirkumpolare Struktur zu betonen.

Ähnlich wie die 4. RPC trennt auch die entsprechende REOF das Skandinavien Muster vom Ostatlantik/Westrussland Muster nicht ganz eindeutig. Somit ist die Interpretation dieser Struktur als SCA eventuell weniger abgesichert als die drei anderen hier präsentierten Moden, dennoch kann davon ausgegangen werden, dass die REOF ein physikalisches Zirkulationsmuster erfasst. Weitere Untersuchungen dieser Mode sind somit gerechtfertigt. Die Benennung als SCA wiederum rechtfertigt sich im Hinblick auf das Raummuster aus der Beobachtung, dass diese, wie auch in der Standarddefinition, ein Aktionszentrum über Skandinavien und ein weiteres, weniger dynamisches über dem Mittelmeerraum aufweist. Die Aktionszentren des EA/WR sind dagegen eher zonal orientiert.

Die 5. REOF ist das Raummuster des NPI. Das dominierende Aktionszentrum (vgl. Abb. 6.1) stimmt sehr gut mit der relevanten Region im Nordpazifik gemäß der Standarddefinition überein. Der starke Fokus, den diese REOF auf dem Aleutentief hat, bestätigt die Interpretation als NPI. Zwar finden sich auch Hinweise auf die übrigen Aktionszentren des PNA, diese treten jedoch in ihrem Ausmaß und ihrem Einfluss auf die RPC deutlich zurück. Dass diese Aktionszentren im SLP weniger markant ausgeprägt sind als bei der Betrachtung der relativen Topographie ist ein bekannter Aspekt des Zirkulationsmusters (WALLACE und GUTZLER 1981).

### 6.1.2 Anwendung im CMIP3

Hier werden die neuen Indexdefinitionen auf die Modelldaten angewendet. Es ergeben sich hieraus zwei Fragestellungen, die im Folgenden bearbeitet werden. Zum einen hat das SRES A1b Ensemble sicherlich einen ausreichenden Umfang und somit das Potential, um die Allgemeingültigkeit der Neudefinitionen zu überprüfen. Daneben ermöglicht es der induktive Charakter der vorgestellten Index-Versionen auch zu überprüfen, ob die AOGCMs in der Lage sind, diese tatsächlich in realistischer Art und Weise zu reproduzieren.

### 6.1.2.1 Äquivalenz zu Standardindices

Die neuen Indexdefinitionen werden hier auf ihre Zulässigkeit und Effektivität bei der Erfassung der Zirkulationsmuster geprüft. Dafür wird in einem ersten Schritt getestet, ob sie in der Lage sind, das Verhalten der auf klassischen Definitionen basierenden Indexreihen zuverlässig abzubilden. Somit wird sichergestellt, dass Zeitverlauf und Entwicklung der Indices unter anthropogenem Forcing, die im weiteren Verlauf der Arbeit untersucht werden, nicht durch die neue Definition der Zeitreihen beeinflusst werden. Im Folgenden werden die Indexreihen, die über die gebräuchlichen Standardverfahren berechnet wurden (vgl. Tabelle 6.1), allgemein mit  $I_S(t)$  bezeichnet, die Zeitreihen, die über die Projektion der Modelldaten auf die REOFs des HadSLP2 generiert wurden, dagegen mit  $I_P(t)$ .

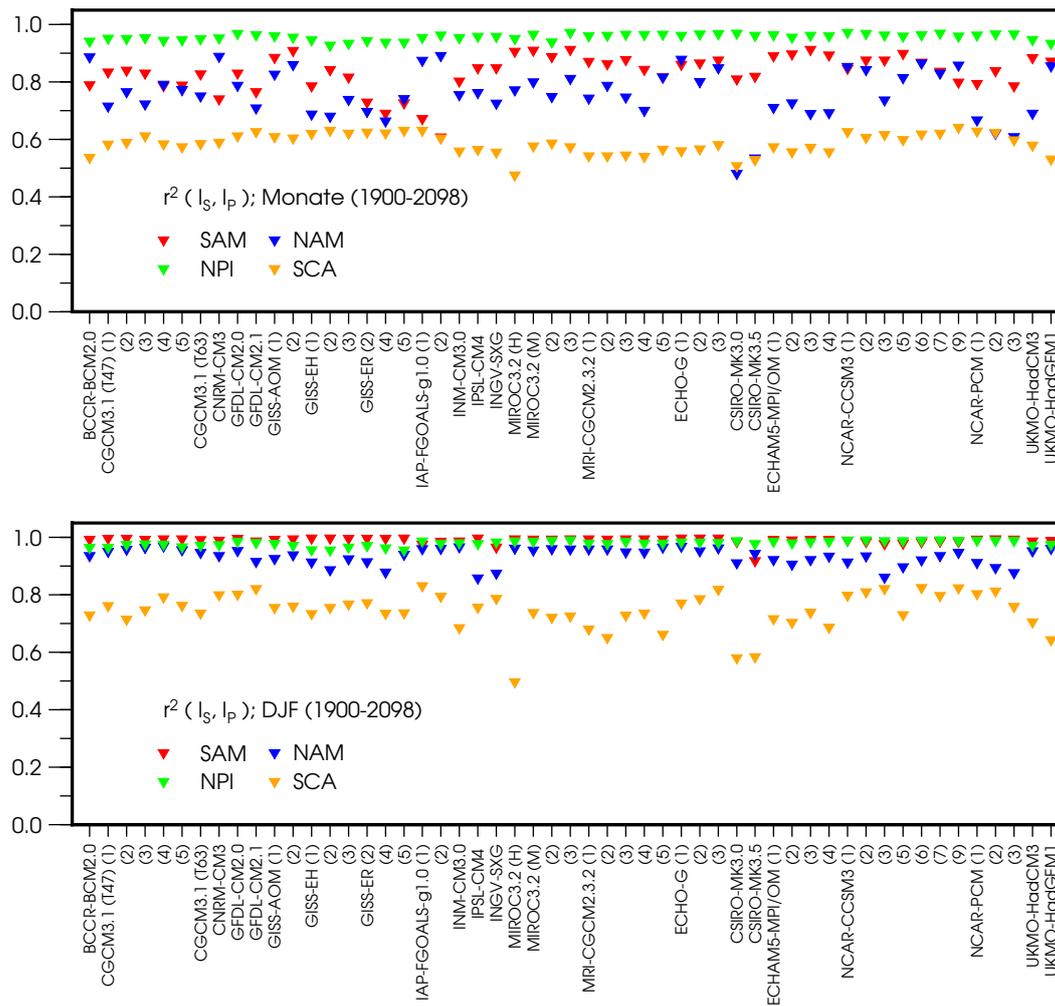
**Tab. 6.1:** Definitionen der Zeitreihen  $I_S$ . Konventionelle Berechnungen für die Indices der außertropischen Eigenmoden.

Muster	Berechnung $I_S$
SAM	1. PC SLP-Anomalien südlich $\Phi = 20^\circ\text{S}$
NAM	1. PC SLP-Anomalien nördlich $\Phi = 20^\circ\text{N}$
NPI	gemittelter SLP Nordpazifik ( $30^\circ\text{N}$ – $65^\circ\text{N}$ , $160^\circ\text{E}$ – $140^\circ\text{W}$ )
SCA	standardisierte Differenz $P(t, 45^\circ\text{N}/5^\circ\text{W}) - 2 \cdot P(t, 65^\circ\text{N}/20^\circ\text{E})$

Die Überprüfung basiert im wesentlichen auf der bivariaten Korrelationsrechnung. Alle  $I_S$  und  $I_P$  wurden für den Zeitraum 1960–1999 standardisiert. Die Korrelationskoeffizienten, die die Grundlage für die erklärten Varianzen in Abb. 6.2 darstellen, wurden jedoch jeweils für den hier betrachteten Gesamtzeitraum von 1900–2098 berechnet. Auf diese Weise soll sichergestellt werden, dass eventuelle Unterschiede in den langfristigen Trends oder Veränderungen in der Variabilität, also den wichtigen Aspekten in der Erforschung des Klimawandels, durch die Analyse erfasst werden.

Die  $I_P$  der unterschiedlichen Jahreszeiten werden über einfache Mittelung der Monatswerte berechnet. Dagegen wurde die Berechnung der Zeitreihen der Perioden im Fall der  $I_S$  jeweils einzeln durchgeführt. Die präsentierten Ergebnisse sind allerdings auch für den Fall reproduzierbar, dass auch die  $I_S$  einfach aus der monatlichen Zeitreihe berechnet werden.

In Abb. 6.2 sind für die  $n = 53$  transienten Simulationen, für die SLP-Daten verfügbar waren, die quadrierten Korrelationskoeffizienten von  $I_S(t)$  und  $I_P(t)$  dargestellt. Es werden jeweils die  $r^2(I_S, I_P)$  für den borealen Winter und die Reihen der Monatswerte abgebildet. Diese Auswahl begründet sich dadurch, dass, jeweils bezogen auf alle vier Zirkulationsmuster und alle Einzelläufe, die Summe der erklärten Varianzen für DJF maximal und bei der Betrachtung der monatlichen



**Abb. 6.2:** Deduktive und induktive Indices. Quadrierte Korrelationen zwischen den nach Standardverfahren berechneten Zeitreihen  $I_S(t)$  der vier global relevanten außertropischen Zirkulationsmoden und den Zeitreihen  $I_P(t)$ , die durch Projektion der Modelldaten auf die HadSLP2-REOFs erzeugt wurden. Oben: Monatswerte, unten für den borealen Winter.

Werte minimal wird. Somit findet sich in diesen Bildern die Spannbreite der linearen Zusammenhangs der verschiedenen Indexdefinitionen. Grundsätzlich ist  $r_{mon}$  geringer als  $r_{djf}$ . Am wenigsten macht sich dieser Unterschied beim NPI bemerkbar. In beiden Fällen liegen sämtliche Korrelationen deutlich über 0.95. Dies ist nicht verwunderlich angesichts der Tatsache, dass die REOF in diesem Fall quasi ausschließlich hohe, nahezu konstante Ladungen in dem Bereich aufweist, der für die Berechnung der Zeitreihe  $I_S(t)$  herangezogen wird. Dagegen sind

die Raummuster von SAM und NAM im Ensemble weniger stabil, wegen ihrer induktiven Generierung über die EOF-Analyse. Die  $I_S$  und  $I_P$  des SAM sind bei Betrachtung sämtlicher Jahreszeiten nahezu perfekt korreliert (jeweils  $r > 0.95$ ). Dass die Korrelationen der Monatswerte dagegen in einigen Läufen auf nur etwa 0.8 absinken, dürfte auf die größere Domäne, für die  $I_P$  repräsentativ ist, zurückzuführen sein. Hier werden vermutlich kleinskalige oder hochfrequente Aspekte der Variabilität des Druckfeldes nicht auf die betreffende REOF attribuiert, die im Standardverfahren erfasst werden. Diese Effekte sollten aber an der Erfassung des physikalischen Phänomens, welches der SAM repräsentiert durch  $I_P$ , nichts ändern. Das gleiche gilt in vollem Umfang auch für den NAM, auch wenn die Korrelationen hier etwas geringer sind als beim SAM. Beim SCA sind jedoch deutlich geringere erklärte Varianzen feststellbar. Für die monatlichen Indexreihen ergeben sich für das Bestimmtheitsmaß nur noch Werte von etwa 50–60 %. Jedoch existieren für die Indexreihe des Skandinavien Musters alternative Berechnungsformen, die mit dem hier verwendeten Referenzindex von NOAA auch nur relativ geringe Korrelationen von 0.66 auf Monatsbasis aufweisen, so dass dennoch davon ausgegangen wird, dass  $I_P(t)$  eine brauchbare Version des SCA-Index darstellt (BUEH und NAKAMURA 2007). Dementsprechend steigen die Korrelationen zwischen den beiden Zeitreihen in allen Einzelläufen auf der saisonalen Skala auch deutlich an. Im Nordwinter werden, bei einem Ausreißer mit  $r = 0.73$ , Korrelationskoeffizienten von  $r = 0.83$ – $0.91$  erreicht.

Als Ergebnis kann also festgehalten werden, dass die Verwendung der  $I_P$  folgende Analysen nicht beeinflussen wird. Entsprechend werden sich die Analysen der außertropischen Eigenmoden im Folgenden auf die  $I_P$  stützen. Tatsächlich sind die Ergebnisse der folgenden Kapitel bei der Verwendung der  $I_S$  nahezu mit den hier präsentierten, auf den  $I_P$  basierenden, identisch. Damit wurde hier auch für vier wichtige Größen des Klimasystems eine neue und effektive Berechnungsmethoden dargelegt. Der massiv verringerte Aufwand, insbesondere bei den über EOF-Analysen generierten Indices ist v.a. bei der Bearbeitung von Multi-Modell Ensembles ein sicherlich hochgradig nützlicher Aspekt.

### 6.1.2.2 Modellinterne Repräsentation

Die Resultate aus Kapitel 6.1.2.1 zeigen, dass die Projektion von globalen SLP-Anomalien auf die REOFs aus Beobachtungsdaten unabhängig vom AOGCM Indexreihen erzeugt, die die statistischen Eigenschaften der gebräuchlichen Standardindices aufweisen. Jedoch bedeutet dies noch nicht, dass AOGCMs diese großskaligen Zirkulationsmuster auch eigenständig erzeugen können, dass also die Organisation der atmosphärischen Variabilität realistisch erfolgt. Es ist somit unklar, ob die so generierten Zeitreihen tatsächlich einen korrekt simulierten physikalischen Aspekt des Klimasystems beschreiben oder ob es sich hierbei nur um Artefakte der Analyseverfahren handelt. Obwohl diese Frage auch im Zusammenhang mit dem

realen Klimasystem immer wieder diskutiert wird (DESER 2000, AMBAUM et al. 2001), wird hier davon ausgegangen, dass die betrachteten Zirkulationsmuster reale Phänomene darstellen.

Ob sich die Zirkulationsmuster nun in realistischer Form in den GCMs wiederfinden, wird über einen mehrstufigen Ansatz geprüft. Wie bei den Beobachtungsdatensätzen werden globale REOFs für die Anomalien des SLP berechnet. In die Varimax-Rotation werden ebenfalls jeweils so viele EOFs miteinbezogen, dass mindestens 80 % der Gesamtvariabilität der Ausgangsdaten repräsentiert sind. Für alle folgenden Schritte werden nur solche REOFs herangezogen, die mindestens 3 % der ursprünglichen Varianz der SLP-Daten erklären.

Diese Raummuster werden mit den REOFs aus dem HadSLP2-Datensatz über gewöhnliche Korrelation verglichen. Hierdurch ergibt sich für jeden Modelllauf pro Zirkulationsmuster ein Kandidat  $K$ , der sich einfach dadurch auszeichnet, dass der Korrelationskoeffizient zwischen den Raummustern den höchsten Wert annimmt. Anschließend wird die RPC von  $K$  mit  $I_P$  korreliert. Die Analyse zielt somit auf den Anteil der Variabilität von  $I_P$ , der durch  $K$  erklärt werden kann. Entsprechend werden hohe Beträge so gedeutet, dass der betreffende GCM-Lauf die jeweilige Mode in einer raumzeitlich realistischen Art und Weise eigenständig erzeugt. In Abb. 6.3 ist das Ergebnis dieser Analyse überblicksartig, in Anlehnung an die klassische Lorenzkurve, dargestellt. Grundlage dieser Abbildung sind die  $n$  quadrierten Korrelationen  $r_i^2(K, I_P)$  zwischen den RPCs  $K$  und den jeweiligen Index-Zeitreihen. Diese werden aufsteigend geordnet und zu

$$\check{r}_i^2(K, I_P) = \frac{r_i^2(K, I_P)}{\sum_{i=n}^1 r_i^2(K, I_P)} \cdot 100 \% \quad (6.1)$$

normiert. Anschließend werden diese  $\check{r}_i^2(K, I_P)$  sukzessive aufsummiert. Die Flächen  $F_1$  in Abb. 6.3 repräsentieren somit die durch die entsprechenden  $K$ s der auf der Abszisse aufgetragenen Anzahl der Läufe des Ensembles (in %) erklärte Anteil der Gesamtvarianz der  $I_P$ , der jeweils auf 100 % festgelegt wird. Nehmen sämtliche  $r_i^2(K, I_P)$  den Wert 1 an, so bedeutet dies mit diesen Konventionen, dass mit 100 % der Läufe auch 100 % der Gesamtvarianz im Ensemble erklärt werden. Dies ist beim SAM in sehr guter Näherung der Fall, der somit als ein reales Phänomen in sämtlichen AOGCMs betrachtet werden kann.

Abweichungen von diesem Idealergebnis zeigen sich in der Fläche  $F_3$ . Der entsprechende Bereich zeigt den nicht erklärten Varianzanteil im Ensemble. Ihre untere Begrenzung ergibt sich durch ein rechtwinkliges Dreieck mit der Höhe  $h = \sum_{i=1}^n \check{r}_i^2(K, I_P)$ , so dass die dadurch definierte Fläche die sukzessive Akkumulation von erklärter Gesamtvarianz bei identischen Werten von  $\check{r}_i^2(K, I_P)$  anzeigt. Für die hier definierten Index-Reihen werden über die modellinternen RPCs  $K$  mindestens 80 % der Gesamtvarianz erklärt. Es kann also davon ausgegangen werden, dass die  $I_P$  durchweg die entsprechenden Phänomene in den Modellen

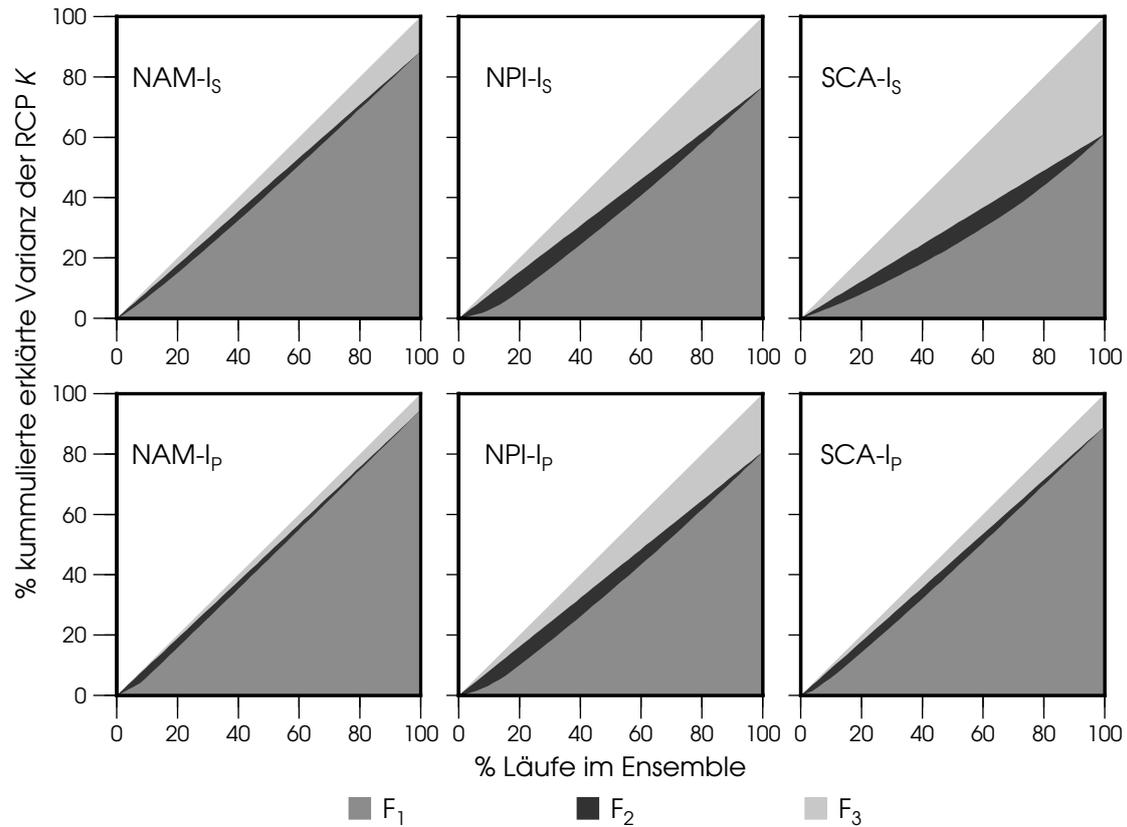
erfassen. Somit kann als ein Ergebnis dieser Analyse festgehalten werden, dass es grundsätzlich möglich ist, mithilfe des CMIP3-Ensembles Untersuchungen über die betrachteten Zirkulationsmoden durchzuführen, da diese durch die Modelle offensichtlich realitätsnah simuliert werden. Diese Feststellung ist etwa für die Beurteilung von Multi-Modell Ansätzen zur zukünftigen Entwicklung dieser Phänomene von praktischer Bedeutung.

Wie ähnlich die modellinternen Versionen der Zirkulationsindices den Referenzzeitreihen im Vergleich zum Ensembledurchschnitt ist, lässt sich an der Fläche  $F_2$  beurteilen. Sie zeigt die Abweichungen der kumulierten  $\tilde{r}_i^2(K, I_P)$  von der theoretischen Gleichverteilung der tatsächlich durch die  $K$  erklärten Varianz. Somit zeigt sie auch die Unterschiede hierbei zwischen den einzelnen Läufen auf, was für die Beurteilung der Zuverlässigkeit und Qualität der individuellen Modelle genutzt wird. Offensichtlich unterscheiden sich die Einzelläufe im Hinblick auf ihre Fähigkeit, die außertropischen Moden zu generieren, nur gering voneinander. Lediglich im Fall des NPI lassen sich einige Modelle identifizieren, die sich deutlicher von der durchschnittlichen erklärten Varianz im Ensemble unterscheiden.

Wird die Analyse mit  $I_S$  statt  $I_P$  durchgeführt, so zeigt sich, dass die RPCs  $K$ , mit Ausnahme von jeweils einem Einzellauf für NPI und SCA, durchweg höhere Korrelationen mit den  $I_P$  aufweisen. Entsprechend sind die Anteile der erklärten Gesamtvarianz im Multi-Modell Ensemble geringer, aber immer noch durchgehend hoch (vgl. Abb. 6.3). Es lässt sich daraus schließen, dass bei Betrachtung der gleichen räumlichen Domäne die Organisation der Variabilität in den AOGCMs jener in den Beobachtungsdaten sehr ähnlich ist. Die deutlichsten Unterschiede zwischen  $I_S$  und  $K$  treten im CMIP3 für das Skandinavien Muster auf. Allerdings lässt sich auch in diesem Fall der größere Teil der Gesamtvarianz im Ensemble durch die modellinternen RPCs erklären. Dies kann dennoch als Hinweis dafür gedeutet werden, dass die tatsächliche Entwicklung in den AOGCMs über die  $I_P$  besser erfasst wird als über die für Beobachtungsdaten definierten Standarddefinitionen.

Als Fazit dieser Untersuchungen ist also festzuhalten, dass die wichtigen Moden der Außertropen in sämtlichen Modellen gut und realistisch abgebildet werden. Da alle  $K$  führende RPCs der Modelle darstellen, kann auch davon ausgegangen werden, dass die relative Bedeutung dieser induktiven Versionen der Zeitreihen sich nicht dramatisch von jener in den Beobachtungsdaten unterscheidet.

Ein weiterer wichtiger Punkt, der unter den Ergebnissen dieser Analyse hervorgehoben werden kann, ist, dass, wenn auch mit geringen Abstrichen im Vergleich zu den  $I_P$ , auch die Standardversionen der Indices die wichtigen Moden der Außertropen in sämtlichen Modellen gut zu erfassen scheinen. Das bedeutet, dass auch bei regionaler Betrachtung die räumlichen Muster der Variabilität dieser Phänomene in sehr guter Übereinstimmung mit den Verhältnissen im realen Klimasystem stehen. Somit ist für diese wichtigen Glieder des Klimasystems nicht von gravierenden Unterschieden innerhalb der CMIP3-Läufe auszugehen. Damit wird das Vertrauen in Untersuchungen dieser Zirkulationsphänomene mithilfe dieses

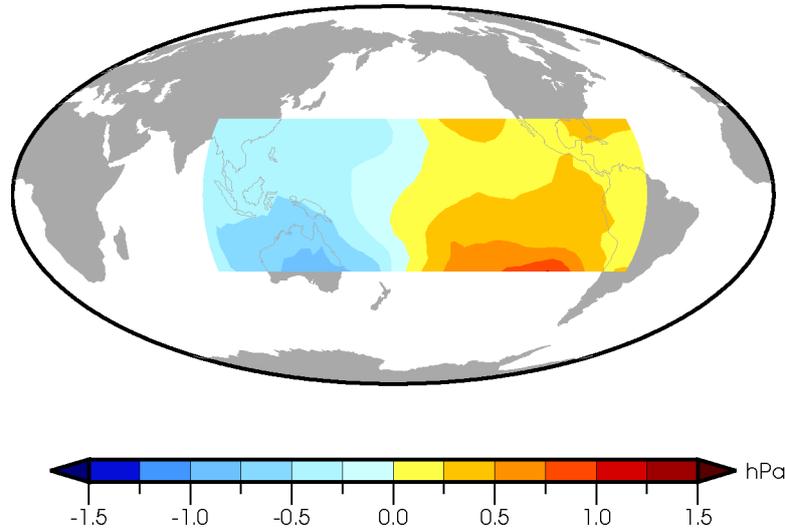


**Abb. 6.3:** Modellinterne Erzeugung von Zirkulationsmoden. Vergleich der beiden Indexdefinitionen mit der RPC  $K$  pro Modelllauf. Jeweils kummulierte erklärte Varianz für die Einzelläufe (aufsteigend geordnet).  $F_2$  zeigt die Abweichungen im SRES A1b von der mittleren erklärten Varianz über alle Ensemblemitglieder,  $F_3$  die Abweichungen von der Ideallinie, wenn für alle  $n$   $r_i^2(K, I_{P/S}) = 1$  angenommen wird. Für SAM- $I_{S/P}$  gilt jeweils  $F_2 \approx F_3 \approx 0$ , daher keine Abbildung. Weitere Erläuterungen im Text

Ensembles gestärkt. Für die Identifikation von qualitativen Unterschieden zwischen den AOGCMs sind diese Variabilitätsmoden angesichts der modellübergreifend guten Performance allerdings nicht geeignet.

## 6.2 ENSO und die Monsune

Im globalen Maßstab hat sich die Erfassung von Zirkulationsmustern der Tropen als nicht auf vergleichbare Art wie in den Außertropen möglich erwiesen. Dennoch kann sicherlich nicht konstatiert werden, dass auf dieser Ebene die bekannten tropischen



**Abb. 6.4:** Raummuster SOI. Wie Abb. 6.1, jedoch 1. EOF aus regionalem Ansatz.

Variabilitätsmoden keine Bedeutung hätten. Sie werden daher in gesonderten Ansätzen berücksichtigt.

### 6.2.1 SOI aus regionalem EOF-Ansatz

Mit ENSO assoziierte EOF/PC-Paare konnten in den Experimenten, die zur Bestimmung der global bedeutenden Zirkulationsmoden durchgeführt wurden, häufig detektiert werden. Allerdings ist ein Ergebnis dieser Experimente, dass diese Zuordnungen nicht stabil bei der Übertragung einer Methode auf unterschiedliche Datensätze sind. So kann im HadSLP2, angesichts einer Korrelation mit der Referenzzeitreihe von 0.64, die 2. RPC durchaus als Indexreihe des SOI interpretiert werden. Bei ERA 40 und NCEP/NCAR dagegen sinkt der Korrelationskoeffizient auf  $r_E = 0.35$  bzw.  $r_N = 0.21$ . Diese Zahlen gelten für monatliche Indexwerte, allerdings sind die Korrelationen bei der Betrachtung der borealen Wintersaison mit  $r_E = 0.51$  (ERA 40) und  $r_N = 0.43$  (NCEP/NCAR) ebenfalls nicht mit der des HadSLP2 ( $r_H = 0.81$ ) zu vergleichen. Diese Unterschiede sind bemerkenswert, angesichts hohen Korrelation der räumlichen Muster in den drei Datensätzen (vgl. Abb. 5.1). Sie illustrieren die Problematik, dass die mit ENSO assoziierte Variabilität bei vielen globalen Ansätzen zwar erfasst wird, sich aber auf zwei oder mehr EOF/PC-Paare verteilt. In Anbetracht dieser Ergebnisse für die Beobachtungsdaten erscheint es wenig sinnvoll, einen solchen Ansatz zur Beurteilung der Qualität der AOGCMs heranzuziehen.

Stattdessen wird ein regionaler Ansatz zur Erfassung von ENSO herangezogen.

**Tab. 6.2:** Korrelation NCAR-SOI und 1. PC aus regionalem Ansatz für Beobachtungen. EOF und PC wurden für monatliche Anomalien bestimmt, Jahreszeiten und Jahresmittel als Mittel der Monatswerte berechnet. Die Tabelle enthält keine Werte die zum Niveau  $\alpha = 0.01$  nicht signifikant wären. Allen Zeitreihen wurde eine als Polynom 3. Grades geschätzte Trendkomponente vor der Korrelation entfernt.

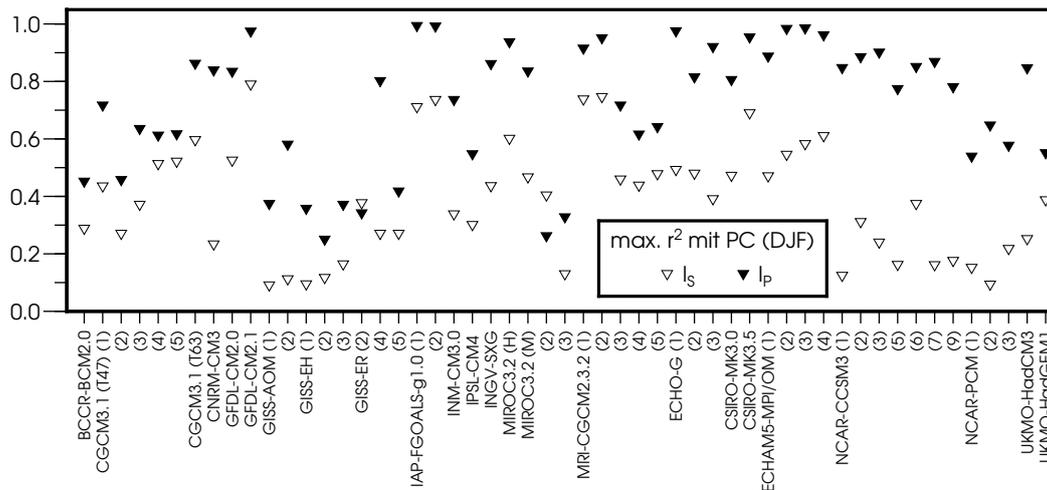
	Monate	DJF	MAM	JJA	SON	Jahr
HadSLP2	.83	.96	.93	.81	.93	.95
NCEP/NCAR	.83	.95	.93	.77	.91	.93
ERA 40	.86	.96	.92	.84	.93	.94

Dieser konzentriert sich auf die betreffende Kernregion, hier abgegrenzt als der tropische Pazifik von  $90^\circ$  Ost und  $120^\circ$  West sowie jeweils  $30^\circ$  Nord und Süd. Für diesen Bereich wird eine Hauptkomponenten Analyse durchgeführt. Es werden monatliche SLP-Anomalien als Ausgangsdaten verwendet, wobei auch in diesem Ansatz die einzelnen Werte entsprechend der durch sie repräsentierten Fläche gewichtet wurden. In Abb. 6.4 ist die führende EOF des HadSLP2 für den Zeitraum 1960–1999 dargestellt. Eindeutig wird hierbei die Southern Oscillation, die charakteristische atmosphärische Wippe zwischen West- und Ostpazifik, erfasst (TRENBERTH und CARON 2000). Die Reanalysedatensätze zeigen in ihrer 1. EOF ein nahezu identisches Dipol-Muster. Auch die zu diesen Mustern gehörenden PCs können mit großer Sicherheit als Zeitreihen des SOI betrachtet werden (vgl. Tabelle 6.2). Die Rotation der Eigenvektoren oder die Verwendung der Korrelations- statt der Kovarianzmatrix ändert weder das Erscheinungsbild der jeweiligen Raummuster noch die dazugehörigen Zeitreihen in einer relevanten Art und Weise. Daher wird im Folgenden die 1. PC der Kovarianzmatrix als SOI verwendet. Es ist auch bei diesem Index möglich, die Raummuster der Beobachtungsdaten weitgehend ohne Informationsverlust auszutauschen. Werden die Zeitreihen der Reanalysedatensätze entsprechend über die Projektion der gewichteten, monatlichen Anomalien auf die 1. EOF des HadSLP2 berechnet, so unterscheiden sich die Korrelationen dieser SOIs mit dem NCAR-SOI nicht signifikant von denen in Tabelle 6.2.

Das gleiche gilt allerdings auch, wenn die ersten Eigenvektoren NCEP/NCAR oder ERA 40 als Projektionsbasis verwendet werden. Insofern ist die Entscheidung, welcher Eigenvektor als Grundlage für die Modellindices  $I_P$  verwendet wird, einigermaßen zufällig. Hier wird, schon um der Konsistenz mit dem Vorgehen bei den außertropischen Zirkulationsindices willen, die 1. EOF des HadSLP2 (Abb. 6.4) verwendet.

Die Beträge der Korrelationen zwischen diesen  $I_P$  und den  $I_S$  sind insgesamt

geringer als bei den außertropischen Moden. Qualitativ scheinen sie bei Betrachtung des CMIP3 in der Gesamtschau mit denen des Skandinavien Musters vergleichbar, allerdings existieren hierbei deutliche Ausnahmen mit nur geringen Korrelationen zwischen den beiden Indexreihen. Ein Grund hierfür ist sicherlich, dass hier als  $I_S$  der einfache SOI verwendet wurde. Die erforderlichen SLP-Werte der Stationen Darwin und Tahiti wurden über die nächsten Gitterboxen der Modelle geschätzt. Wird ein solcher SOI aus den hier verwendeten Beobachtungsdaten berechnet, so ist die Korrelation dieses Index mit der Zeitreihe von NCAR grundsätzlich geringer als die Werte in Tabelle 6.2. Diese Feststellung spricht auch dafür, dass diese Form der Berechnung als weniger zuverlässig betrachtet werden kann. Das gilt in verstärktem Maße bei der Anwendung der Indexdefinition auf Modelldaten. Zwar wird durch den  $I_S$  auch in diesem Fall, ein einigermaßen realistisches Verhalten der Atmosphäre in dieser Region vorausgesetzt, die grundsätzliche Dynamik der Southern Oscillation, die die EOF sicher repräsentiert, auf diese Weise erfasst. Dennoch wird im Hinblick auf die AOGCMs der  $I_P$  für den SOI, aufgrund der ungleich größeren Datengrundlage seiner Berechnung, als zuverlässiger eingestuft als der Referenzindex.



**Abb. 6.5:** PCs als modelleigener SOI. Maximale erklärte Varianz durch lineare Korrelationen zwischen den ersten 1–5 regionalen PCs der AOGCMs und den beiden Versionen für den SOI. Je 1900–2098 (DJF).

Für die meisten Modelle repräsentiert nicht die führende EOF die Southern Oscillation. In diesen Fällen findet sich ein entsprechender Dipol in einer weniger wichtigen EOF (vgl. Kapitel 7.1). Solche modelleigenen SOIs werden über die  $I_P(t)$  deutlich besser repräsentiert als über die  $I_S(t)$ , wie Abb. 6.5 verdeutlicht. Dargestellt sind die maximalen Werte des Bestimmtheitsmaß aus bivariater Korrelation

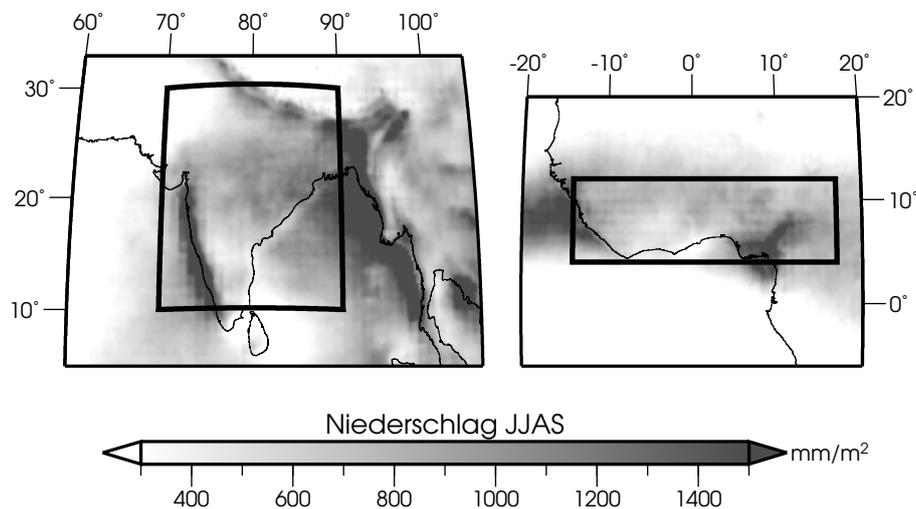
zwischen den beiden SOI-Versionen und den führenden fünf PCs  $c_i(t)$  mit  $i = 1, \dots, 5$ . In allen Fällen weist  $I_P(t)$  höhere Korrelationen mit den  $c_i(t)$  auf. Das gilt auch, wenn die Analyse mit rotierten PCs durchgeführt wird. Für die Indexreihen im borealen Winter lassen sich im Regelfall Korrelationskoeffizienten von  $r > 0.8$  erzielen. Einige Modelle weisen allerdings mit  $r = 0.55\text{--}0.63$  deutlich geringere Werte auf. Die beobachtete Dynamik scheint im Bezug auf ihre raumzeitliche Struktur in diesen Modellen weniger gut simuliert zu werden. Insgesamt sinken die Korrelationen zwischen  $I_P(t)$  und  $c_i(t)$  bei Betrachtung der monatlichen Werte ebenso wie im Fall der außertropischen Zirkulationsmoden geringfügig ab, ohne dass dies jedoch die grundsätzliche Interpretation der PCs als modelleigene SOIs beeinträchtigt. Wie Abb. 6.5 deutlich zeigt, ist dagegen die einfache Implementation des SOIs über  $I_S(t)$  weniger erfolgreich darin, die tatsächliche Dynamik, die die Atmosphäre der AOGCMs für den tropischen Pazifik aufweist, zu erfassen. Eine Schlussfolgerung aus dieser Beobachtung ist, dass dieser Index weniger geeignet erscheint, um die Entwicklung des ENSO-Systems im CMIP3 zu beurteilen. Zudem lassen sich hier, anders als bei den Außertropischen Moden, innerhalb des CMIP3 deutliche Unterschiede bei der Simulation der natürlichen Variabilität feststellen.

### 6.2.2 Monsune

Weder der westafrikanische, noch der indische Monsun, noch irgendein anderer bekannter regionaler Monsun konnte bei der Analyse der globalen Felder der Beobachtungsdatensätze detektiert werden. Hier erwies sich unter anderen die Anwendung der EOF-Analyse auf die Niederschlagsfelder der Reanalysedatensätze als ungeeignet die einzelnen Monsune zu separieren oder den globalen Monsun zu erfassen (vgl. auch CHANG et al. 2011). Zurückzuführen ist dies wohl auf die zu hohe und insbesondere zu kleinskalige Variabilität des Niederschlagsfeldes, die hohe Ladungen in einzelnen EOFs verhindert. Auch Experimente mit anderen Feldgrößen blieben hierbei im Wesentlichen erfolglos. Eine raumzeitlich stabile und plausible Erfassung war mit der beschriebenen Vorgehensweise für die hier in Betracht gezogenen Monsune nicht möglich.

Da es sich bei den Monsunen allerdings um die Zirkulationsphänomene handelt, die den mit Abstand größten Einfluss auf den wirtschaftenden Menschen in den durch sie geprägten Regionen ausüben, soll im Rahmen dieser Arbeit nicht darauf verzichtet werden, diese in die folgenden Analysen miteinzubeziehen. Als zwei Vertreter der verschiedenen regionalen Monsunsysteme wurden der indische Monsun und der westafrikanische Monsun hierfür ausgewählt. Für beide lässt sich eine überragende Prägung der sozio-ökonomischen Systeme in den betroffenen Regionen konstatieren (WASSMANN et al. 2009). Damit einher geht eine ebenfalls enorme Vulnerabilität gegenüber Schwankungen und klimatologischen Veränderungen (WEBSTER et al. 1998, SIVAKUMAR et al. 2005, GADGIL und GADGIL 2006). Die Alternative eines regionalen, ebenfalls EOF-basierten Index ergab ebenfalls

keine stabilen Erfolge. Ein wesentliches Problem hierbei war fehlende Konsistenz in der Variabilität der Beobachtungsdaten, was die Reproduktion von Ergebnissen bei der Verwendung anderer Daten verhinderte. Sicherlich ist hierfür die geringe interannuelle Variabilität in den Feldern von Luftdruck und Temperatur in den Tropen ein entscheidender Punkt. Eventuell sind auch die Aggregierungsmethoden insbesondere der Reanalysen in den Tropen weniger zuverlässig als für die polwärtsigen Regionen, was in deutlichen Differenzen zu anderen Datensätzen in den niederen Breiten resultiert (z. B. KISTLER et al. 2001). Westafrika weist zudem einen Mangel an kontinuierlichen Messreihen mit ausreichender Länge für die hier angewandte Methodik auf (HULME 1992). Die Kreuzvalidierung einer eigenen Definition der Monsun-Indices wäre demzufolge nicht im gleichen Umfang wie für die übrigen Zirkulationsmuster möglich.



**Abb. 6.6:** Flächenindices: EIMR und AMR. Jeweils mittlerer Niederschlag in der Monsunsaison für die Jahre 1998–2007 (TRMM Daten). Die Indexreihen werden über den gemittelten Niederschlag innerhalb der markierten Bereiche ermittelt.

So wurde in diesen beiden Fällen eine pragmatische Lösung gewählt und jeweils auf einen Standardindex als Zeitreihe zurückgegriffen. Für den indischen Monsun wird hier der EIMR, für den westafrikanischen entsprechend der AMR, zur Operationalisierung der zeitlichen Entwicklung verwendet. Diese Indices ergeben sich als mittlerer Gesamtniederschlag in den betreffenden Definitionsregionen (vgl. Abb. 6.6) während der Monsunsaison von Juli bis September. Der Niederschlag über den Gebieten ergibt sich bei der Verwendung diskreter Gitterdaten als flächengewichtetes Mittel  $R_{\bar{A}}$  des Niederschlags in  $\text{mm}/\text{m}^2$  über die  $n_A$  Gitterboxen

der Region:

$$R_{\bar{A}}(t) = \sum_{s=1}^{n_A} \omega(s) R(t, s) \quad (6.2)$$

Die  $\omega(s)$  sind die normierten Werte des Kosinus der geographischen Breite der Gitterbox von  $s$ .

In den Modelldaten wird diese Zeitreihe jeweils für den Zeitraum 1960–99 standardisiert, um den starken Bias abzumildern, den die AOGCMs für den Niederschlag aufweisen (vgl. Kapitel 4). Der besondere Vorteil dieser Indexdefinitionen ist sicherlich, dass eine Fehlinterpretation von EIMR und AMR auch in ihrer zeitlichen Entwicklung quasi ausgeschlossen ist, da sie sich allein auf die interessierende Größe, den Niederschlag zur Monsunsaison, begründen. Das bedeutet aber auch einschränkend, dass Änderungen in der Monsunzirkulation als Reaktion auf das anthropogene Forcing im SRES A1b nur in ihren effektiven Auswirkungen erfasst werden können. Dagegen könnte die Verwendung einer dynamisch basierten Indexvariante (z. B. WEBSTER und YANG 1992) hierbei auch die beteiligten und verantwortlichen Prozesse offenbaren. Auf die Nutzung einer solchen wurde allerdings verzichtet, was im Wesentlichen durch die mangelnde Verfügbarkeit der benötigten Daten für die AOGCMs begründet ist. Wie Tabelle 6.3 zeigt, ist die interannuelle Variabilität in diesen beiden einfachen Flächenindices insbesondere in den Reanalysedatensätzen durchaus als individuell zu bezeichnen. Diese teilweise nur geringen Korrelationen der Beobachtungsdaten verdeutlichen die unbefriedigende Qualität der verbreiteten Datenprodukte in diesen so bedeutsamen, bevölkerungsreichen und vulnerablen Weltregionen.

**Tab. 6.3:** Korrelationen Beobachtungsdaten (EIMR und AMR). Jeweils Monsunsaison (JJAS) für den maximalen Überschneidungsbereich der beiden Datensätze. Wenn dieser kürzer ist als 10 Jahre, ist keine Korrelation angegeben (-). Die Korrelationskoeffizienten rechts und oberhalb der Hauptdiagonalen (leer) beziehen sich auf den indischen, die Werte links und unterhalb auf den westafrikanischen Monsun.

	ERA 40	NCEP/NCAR	CRU	GPCP	CPCC	TRMM
ERA 40		.17	.32	.61	.36	-
NCEP/NCAR	-.14		.71	.23	.62	.53
CRU	-.25	.63		.76	.94	-
GPCP	.36	.48	.76		.71	.97
GPCC	-.19	.64	.89	.82		.98
TRMM	-	.53	-	.28	.84	

### 6.3 Indexentwicklung im CMIP3

Grundsätzlich von Interesse ist die Entwicklung der so erzeugten Indexzeitreihen im CMIP3. Die überragende Bedeutung, die die hier erfassten Phänomene für das Klimasystem haben, erlauben anhand dieser wenigen Zeitreihen bereits eine recht genaue Charakterisierung der herrschenden Zustände. Somit besteht auch an ihrer Entwicklung als Reaktion auf anthropogenes Forcing ein vitales Interesse. Dies gilt insbesondere, da es mithilfe der dominanten Klimamoden möglich ist, die Betrachtung des Klimawandels auf eine räumliche Ebene unterhalb der globalen Skala auszuweiten. Die geringe Auflösung globaler Klimamodelle sowie eine Vielzahl damit verbundener Schwierigkeiten bei konkreten Aussagen zur künftigen klimatischen Entwicklung auf subkontinentaler Basis können durch den Fokus auf diese großskaligen Phänomene teilweise ausgeglichen werden. Da kaum eine bevölkerte Weltregion von keinem der hier identifizierten Indices beeinflusst wird, ist anhand der Zeitreihen in Abb. 6.7 eine erste Näherung an die klimatischen Verhältnisse auf regionaler Ebene, wie sie durch das CMIP3 Ensemble im Verlauf des 21. Jahrhunderts simuliert werden, möglich. Es wird hierbei davon ausgegangen, dass Entwicklungen, die sich in ähnlicher Form in verschiedenen Modellen finden, als wahrscheinlicher im Hinblick auf ihr Eintreten im realen Klimasystem bei entsprechender Veränderung der Randbedingungen beurteilt werden können als solche, die sich nur in einem einzigen Modell finden. Solche Entwicklungen könnten eher auf die spezifische Modellphysik des AOGCMs zurückzuführen sein als auf eine physikalisch realistische Reaktion auf das angesetzte Forcing. So werden jeweils kombinierte Zeitreihen für sämtliche Modellläufe dargestellt. Hierfür wurden jeweils Multi-Modell Mittelwertszeitreihen  $M(t)$  gebildet.

$$M(t) = \frac{1}{n} \sum_{i=1}^n I(t, i) \quad (6.3)$$

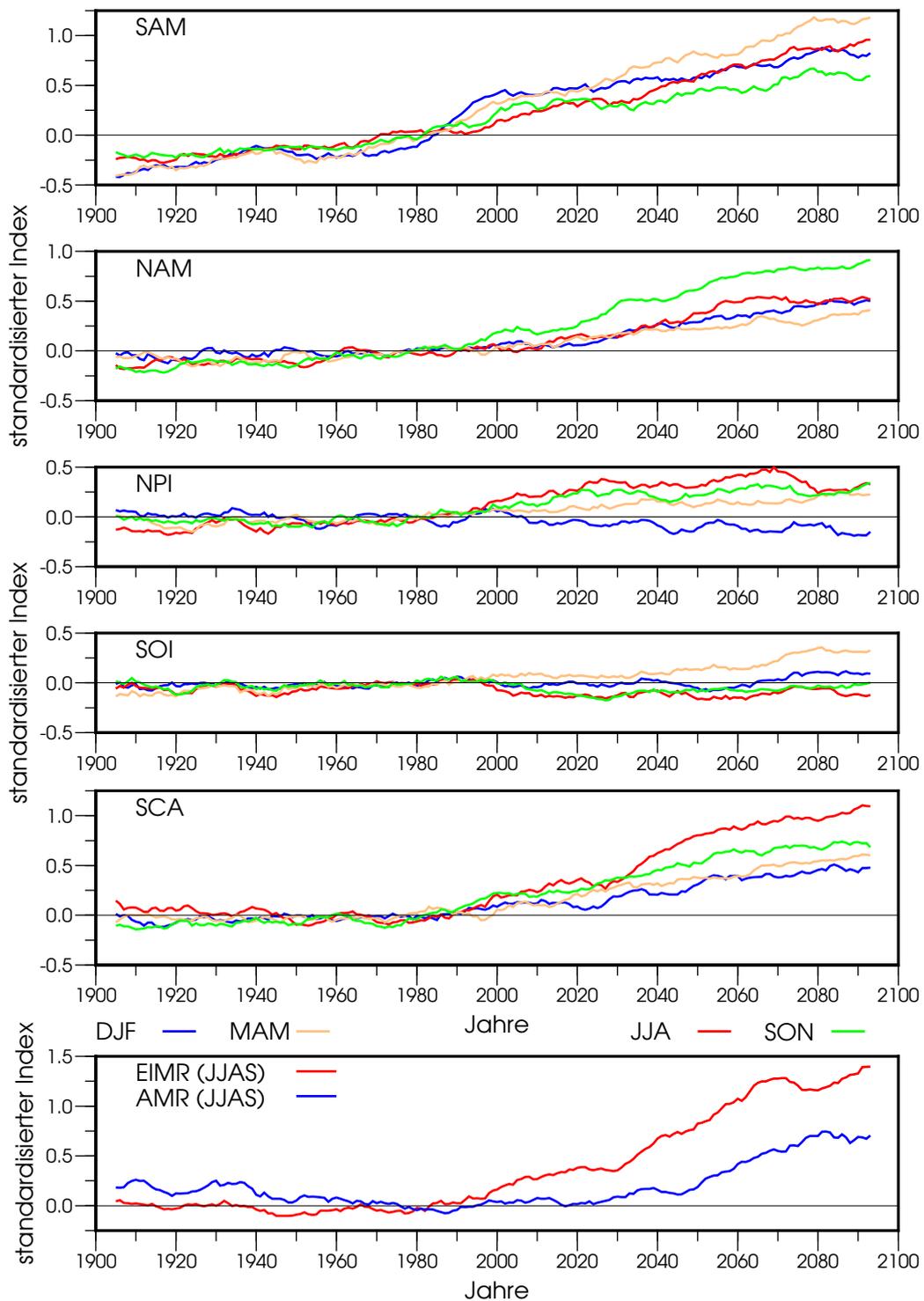
Als Indices  $I(t, i)$  werden jeweils die beschriebenen  $I_P(t, i)$  bzw. die beiden Monsunindices genutzt. Um die in der Frage des zu erwarteten, langfristigen Verhaltens der Zirkulationsmoden irrelevante kurzfristige interne Variabilität weiter zu reduzieren, werden in Abb. 6.7 jeweils  $M_{11}(t)$  gezeigt, bei denen diese durch einen 11-jährigen Tiefpassfilter entfernt wurde.

Sicherlich erlauben diese Zeitreihen die Benennung einiger Grundtendenzen im CMIP3. Für sämtliche Jahreszeiten sind hier für den SAM und den NAM positive Trends in  $M_{11}(t)$  erkennbar. Damit ist grundsätzlich die Intensivierung der Westwinde der jeweiligen Halbkugel verbunden. Dies ist zunächst Ausdruck der allgemeinen Tendenz der Luftdruckentwicklung, die durch Abnahme in den Außertropen in Kombination mit der Zunahme in den Tropen, speziell aber den subtropischen Hochdruckgebieten gekennzeichnet ist. Die vielfältigen Auswirkungen, die mit den positiven Trends dieser Zeitreihen verbunden sind, sollen hier nicht

näher diskutiert werden, da dies an anderer Stelle bereits für die beobachteten und simulierten Indexreihen geschehen ist (z. B. IPCC 2007a). Die intensivsten Veränderungen hin zur positiven Indexphase sind jeweils während der Herbstmonate feststellbar, während der Frühling die geringsten Veränderungen aufweist. Die Trends für die solistalen Jahreszeiten, in denen SAM und NAM z. Zt. die größte klimatische Bedeutung haben, sind jeweils von vergleichbaren mittleren Niveau. Der NPI zeigt dagegen im CMIP3 in den Jahreszeiten unterschiedliches Verhalten. So lässt sich im Winter im 21. Jahrhundert ein leicht negativer Trend für  $M_{11}(t)$  feststellen. Mit dieser Intensivierung des Aleutentiefs verbunden sind ebenfalls verstärkte Westwinde in dieser Region und Jahreszeit. Dagegen werden in den übrigen Saisons, insbesondere im Sommer, für den NPI positive Trends angezeigt.

Für den Nordpazifik ist also eine andere Dynamik des Klimawandels als in der atlantischen Region oder auf der Südhalbkugel, zumindest im Hinblick auf die Westwinde, festzuhalten. Das SCA weist positive Trends in vergleichbarem Umfang mit dem SAM auf. Diese Veränderung im CMIP3 ist zwar weniger gut dokumentiert als die des NAM, kann aber für die betroffenen Regionen durchaus vergleichbare Effekte bedeuten. Die Sommer-Zeitreihe zeigt für den SCA den intensivsten Trend auf. Der besonders bedeutende SOI weist für alle Jahreszeiten nur geringe Trends im Multi-Modell Mittelwert. Diese sind zudem von unterschiedlichem Vorzeichen. Für den borealen Sommer und Herbst ist eine leichte Tendenz hin zur negativen Phase feststellbar. Für die Monate von Dezember bis Mai, also einschließlich des für die klimatische Bedeutung des ENSO-Phänomens wichtigen Nordwinters, wird dagegen in der Mehrzahl der Modellläufe eine Tendenz zur positiven Phase sichtbar. Die beiden Monsunzeitreihen weisen ausgeprägte positive Trend auf. Sie repräsentieren damit zunächst den intensivierten hydrologischen Zyklus, über den eine wärmere Erde verfügen wird (z. B. ALLEN und INGRAM 2002). Da die Zeitreihen hier nur den Niederschlag miteinbeziehen, kann allein aus diesen allerdings noch nicht auf die Veränderungen der tatsächlichen Wasserverfügbarkeit geschlossen werden. Die Zunahme des Monsuns ist im CMIP3 für den EIMR deutlich stärker ausgeprägt als für den westafrikanischen Monsun. Dennoch ist auch für diesen der Trend in der Zeitreihe ein deutliches Signal der Klimaänderung unter dem steigenden Strahlungsantrieb. Am Beispiel der Monsune werden die regionalen Aspekte des Klimawandels, hier in der einfachsten Version der regionalen Umsetzung des globalen Forcings, besonders deutlich.

Neben den gemeinsamen Entwicklungen der Zeitreihen im Ensemble und somit der Grundtendenz in der klimatischen Entwicklung für die von den dargestellten Zirkulationsmustern betroffenen Regionen, lassen sich auf diese Weise auch die Differenzen zwischen den Modellen im Hinblick auf diese Größen gebündelt erfassen. Entsprechend können mit diesen Indexreihen die tatsächlichen Unsicherheiten in den Multi-Modell Ensemble Prognosen konzentriert betrachtet, abgeschätzt und bewertet werden. In Kapitel 9 werden die hier durch die Mittelwertszeitreihen aufgezeigten Klimaänderungssignale auf ihre Stabilität getestet.



**Abb. 6.7:** Klimamoden: Entwicklung im SRES A1b. Jeweils Multi-Modell Mittelwerte aller verfügbaren Modellläufe. Alle Einzelläufe sind für die betrachteten Jahreszeiten auf die Periode 1960–1999 standardisiert, konstante Skalierung.

## 7 Das ENSO-Phänomen im Multi-Modell Ensemble

Die Simulation von ENSO erfordert, im Gegensatz etwa zu den beiden Annular Moden der Außertropen, beide Komponenten eines gekoppelten Atmosphäre-Ozean Modell (z. B. WANG und PICAUT 2004, DELCROIX 1998). Es handelt sich um eine nicht triviale Aufgabe für die Modellperformance, die nicht von allen CMIP3-Modellen auf gleichermaßen befriedigende Art und Weise gelöst wird (OLDENBORGH et al. 2005, COLLINS et al. 2010). Die Ergebnisse aus Kapitel 6 legen zudem nahe, dass es mit der verwendeten Methodik und der zur Verfügung stehenden Datenauswahl möglich ist, die bekannten Differenzen zwischen den AOGCMs in der Repräsentation dieser Zirkulationsmode zu erfassen. Diese Unterschiede werden in diesem Kapitel weiter herausgearbeitet und konkretisiert. Es liegt die Vermutung nahe, dass Modelle mit realistischem ENSO-System besser in der Lage sind das hochkomplexe Klimasystem abzubilden als solche, deren ENSO-Dynamik deutlich von der beobachteten abweicht oder scheinbar über keine solche verfügen. Daneben lässt sich aus der Klimageschichte ableiten, dass ENSO bei früheren Veränderungen des Strahlungshaushaltes eine wichtige Rolle bei der Umsetzung dieses externen Signals in die Reaktion des Klimasystems gespielt hat (MANN et al. 2005, D'ARRIGO et al. 2005). Die Fähigkeit das ENSO-Phänomen in der freien Integration zu reproduzieren wird im Folgenden als Gütekriterium für die untersuchten AOGCMs herangezogen. In diesem Kapitel wird mit  $I(t)$  jeweils die SOI-Variante  $I_P(t)$  nach der Definition in Kapitel 6.2.1 bezeichnet. Beobachtungsdaten werden mit dem Index  $B$  gekennzeichnet, die  $n$  Modelle mit dem Index  $i$  unterschieden, so dies notwendig ist.

### 7.1 ENSO-Charakteristika in AOGCMs

Hier werden anhand einiger Kriterien die offensichtlichen Differenzen zwischen den verschiedenen AOGCMs bei der Simulation des ENSO-Phänomens aufgezeigt. In diesem Kapitel ist es erforderlich, die individuellen Modellergebnisse im Hinblick auf einzelne, bekannte Aspekte von ENSO zu verdeutlichen.

Deren Umsetzung stellt offensichtlich auch im Fall der AOGCMs charakteristische Kenngrößen der individuellen Modelle dar, so dass die Unterschiede zwischen verschiedenen Realisationen eines Modells für dieses Kapitel zu vernachlässigen sind.

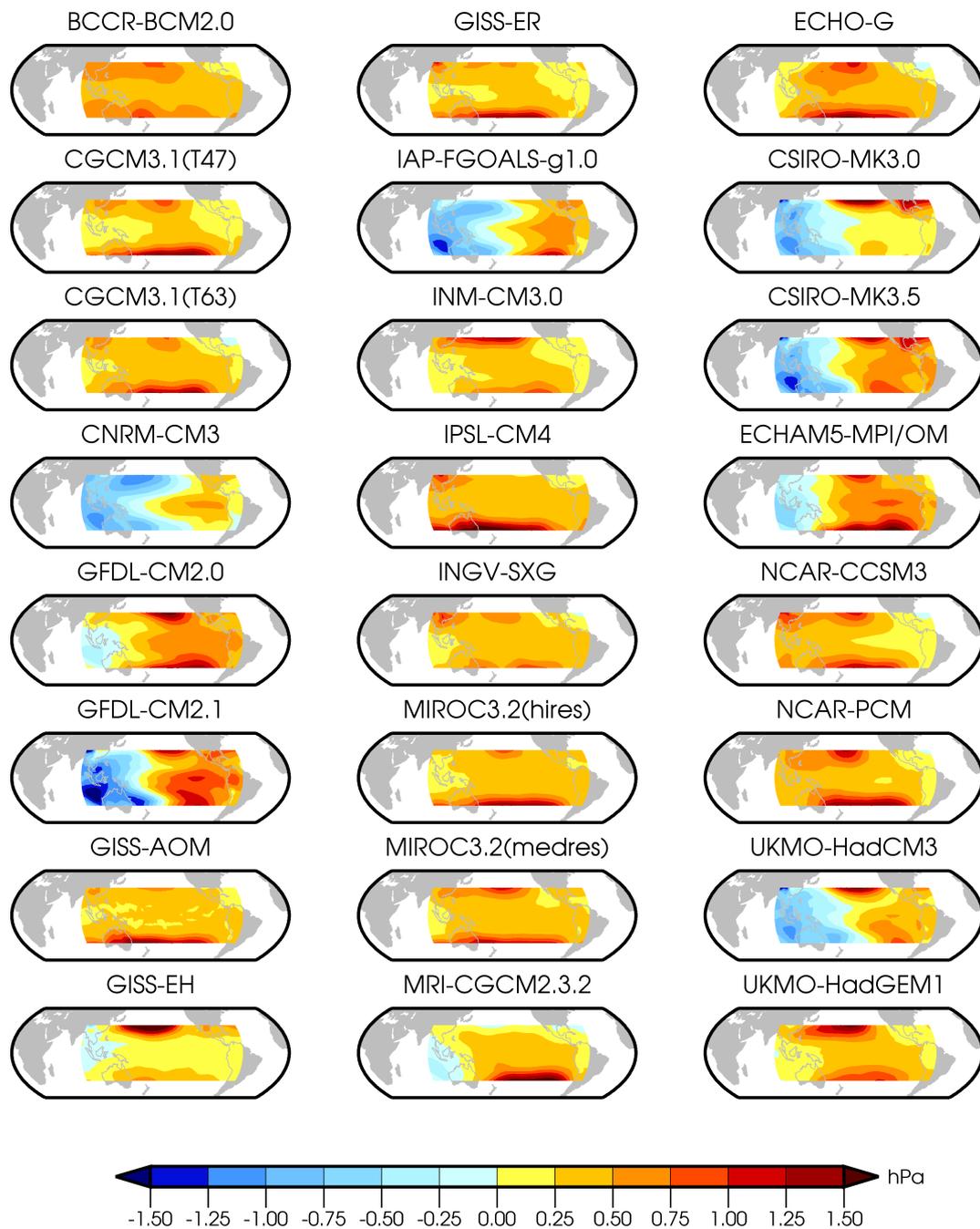
Daher kann im Folgenden davon abgesehen werden im Fall der elf Modell Ensembles die Befunde jeweils für sämtliche Einzelläufe darzustellen. Stattdessen werden, für diese GCMs jeweils die Ergebnisse für den Lauf (1) präsentiert. Die Ausnahme hiervon bildet GISS-ER, da dieser für dieses GCM nicht vorliegt. Stattdessen wird auf (2) zurückgegriffen. Diese geringfügige Einschränkung der verfügbaren Information ist durch den damit verbundenen Zugewinn an Übersichtlichkeit zu rechtfertigen. Die grundsätzliche Aussage würde durch die mehrfache Abbildung von nahezu identischen Ergebnissen im Fall von mehreren Einzelläufen nicht verändert, da hier der Fokus auf Unterschieden, nicht Gemeinsamkeiten oder dominanten Strukturen, im CMIP3 liegt. Die Differenzen zwischen den verschiedenen AOGCMs sind dagegen derart prononciert, dass eine schematische Darstellung für deren qualitativen Erfassung ausreicht. Ebenso ermöglicht dies den Abgleich mit der gleichen ENSO-bezogenen Kenngröße im realen Klimasystem.

Die Ergebnisse sind bei diesen unterschiedlichen Aspekten durchaus variabel. Einige AOGCMs produzieren sämtliche hier berücksichtigten Aspekte deutlich, andere keinen. Die Mehrzahl der CMIP3-Modelle jedoch liefert ein gemischtes Resultat, bei dem die Eigenschaften des ENSO-Systems unterschiedlich realistisch nachvollzogen werden.

### 7.1.1 Raummuster der Variabilität

Es ist zunächst naheliegend, zu überprüfen, ob die raumzeitliche Variabilität der Beobachtungen durch die AOGCMs entsprechend nachvollzogen wird. In Kapitel 6 konnte gezeigt werden, dass dies im Fall der außertropischen Moden in hohem und überzeugendem Maße der Fall ist. Entsprechend kann überprüft werden, ob die EOF-Analyse des tropischen Pazifik ebenfalls die charakteristischen Strukturen enthüllt, die sich in den Beobachtungsdaten finden lassen. Die Analyse erfolgt über jeden Einzellauf mit den monatlichen Anomalien für die Modelljahre 1960–1999. Wiederum wurden eventuelle Trends der Zeitreihen an den Gitterpunkten über lineare Regression entfernt. Angesichts des regionalen Ansatzes sollte hierbei die Southern Oscillation sicherlich das dominierende Muster der Domäne sein, so die Bedingungen im Modell mit denen in der Realität vergleichbar sind. Hiervon kann ausgegangen werden, so das Modell über einen realitätsnahen ENSO-Mechanismus verfügt.

Im Multi-Modell Ensemble ergibt sich bei diesem Vorgehen jedoch ein differenziertes Bild. Nicht einmal die Hälfte der Modelle zeigt in ihrer führenden EOF das charakteristische, mit der Southern Oscillation assoziierte, Dipolmuster. Dieser Befund ändert sich auch nicht, wenn statt den Modelljahren 1960–1999 die Perioden 1900–1930 oder 2040–2070 zur Extraktion der Eigenvektoren herangezogen wird. Er kann also durchaus als stabil für die einzelnen Modelle befunden werden. Hieran ändert auch die Verlängerung der untersuchten Periode nichts, so dass dies als eine tatsächliche Eigenschaft dieser Modelle aufgefasst werden muss. Somit las-



**Abb. 7.1:** Modelle: 1. EOF im tropischen Pazifik. Für Modell-Ensemble wird nur ein Lauf gezeigt. Flächengewichtungen entfernt. Die gezeigten Raummuster entsprechen dem Wert der standardisierten PC  $\tilde{c}_1(t, i) = 1$ . Datenbasis sind monatliche Anomalien (linearer Trend entfernt)

sen sich an dieser Stelle durchaus Unterschiede zwischen den Modellen konstatieren, zumal die Basisregion so gewählt ist, dass die Southern Oscillation, und damit das ENSO-Phänomen, als dominierende Oszillation als gesichert angenommen werden kann. Sicherlich ist dieser Befund in der Art zu bewerten, dass in vielen AOGCMs die raumzeitliche Variabilitätsstruktur von der beobachteten abweicht. Es steht zu vermuten, dass dies eine Folge von nicht ausreichend realitätsnah abgebildeten Prozessen darstellt. Eine extremere Form der Deutung wäre es, AOGCMs, deren erste EOF keinen Dipol aufweist, ein realistisches ENSO-System generell abzuspochen. Andere Befunde sind mit einer solchen Interpretation von Abb. 7.1 jedoch nicht vereinbar.

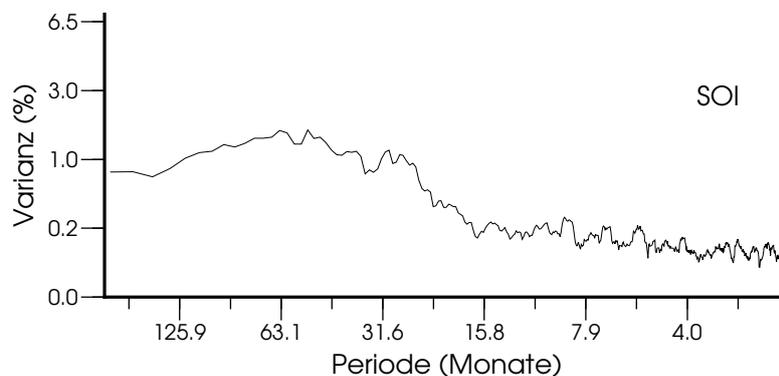
So findet sich in allen Modellen zumindest eine Variabilitätsmode die das Raummuster grundsätzlich abbildet, wodurch die Korrelationen zwischen den  $c_i(t)$  und  $I(t)$  hervorgerufen werden (vgl. Abb. 6.5). Die beobachtete Bedeutung für die Organisation der atmosphärischen Dynamik in der Indexregion wird häufig aber nicht realistisch wiedergegeben. Teilweise findet sich erst in der fünften EOF ein zonaler Dipol im Luftdruckfeld. Die Ladungen dieser EOFs sind dann insbesondere für den äquatornahen Bereich sehr gering. Somit wird die Rolle der randtropischen Druckgebilde für die Southern Oscillation relativ überschätzt. Für diese AOGCMs weist der einfache, klassische SOI die geringsten Korrelationen mit den modelleigenen PCs auf. Angesichts der Bedeutung, die die atmosphärische Komponente für das gesamte ENSO-System hat, wäre dagegen anzunehmen, dass auch in einem AOGCM die einfache Differenz zweier SLP-Zeitreihen die zuverlässige Bestimmung seines Zustands erlauben sollte.

Neben diesen qualitativen Unterschieden zwischen Beobachtungen und AOGCMs finden sich auch weniger augenfällige quantitative. Tendenziell weisen die AOGCMs, die die Southern Oscillation in der führenden EOF repräsentieren, größere Amplituden im SLP als die drei Beobachtungsdatensätze auf. Dagegen wird die Dynamik insbesondere von jenen Modellen, die diese erst in der vierten oder fünften EOF abbilden, deutlich unterschätzt. Damit ist zu vermuten, dass in diesen AOGCMs der SOI weniger Auswirkungen auf das simulierte Wettergeschehen hat. Einige Modelle zeigen zudem Abweichungen sowohl in der Position der dynamischen Zentren des Dipols, als auch in der zonalen Lage des Vorzeichenwechsels der EOF.

Die hier aufgezeigten Differenzen innerhalb des CMIP3 stellen einen eindrücklichen Nachweis dafür dar, dass die interne Klimavariabilität der Tropen modellübergreifend weniger überzeugend abgebildet wird als die der Außertropen. Dies impliziert die grundlegende Frage, ob die Reaktionen des Klimasystems auf externes Forcing durch AOGCMs realistisch abgebildet werden, denen die Fähigkeit fehlt, dessen bedeutsamstes Variabilitätsmuster zu reproduzieren.

### 7.1.2 Varianzspektrum

Das ENSO-Phänomen stellt einen vielfach gekoppelten Mechanismus dar, an dessen Entstehung und Aufrechterhaltung Prozesse der Atmosphäre und des Ozeans gleichermaßen beteiligt sind. Die ersten Analysen und Erkenntnisse in diesem Zusammenhang gehen auf BJERKNES (1969) zurück. Im SOI als einer Realisation der rückgekoppelten Prozesse, prägen sich diese ab. Dadurch verfügt er, im Gegensatz zu allen anderen hier betrachteten Zeitreihen der Klimavariabilität, über typisches, über weißes oder auch rotes Rauschen hinausgehendes spektrales Verhalten. Dieses resultiert aus dem Ineinandergreifen der unterschiedlichen Komponenten des ENSO-Systems mit ihren spezifischen unterschiedlichen Zeitskalen, wobei letztlich der trägere Ozean die ausschlaggebende Rolle spielt. Dagegen stellen die rein atmosphärischen Moden im wesentlichen Markov-Prozesse dar, die zu beschreiben anhand der hier verwendeten Monatsdaten nicht möglich ist (vgl. z. B. FELDSTEIN 2000). In Abb. 7.2, wie auch in den übrigen Darstellungen in dieser



**Abb. 7.2:** Varianzspektrum NCAR-SOI. Erklärte Varianz pro Periode für den Zeitraum 1900–2000 auf Basis monatlicher Anomalien zum Zeitraum 1960–1999 (kubische Trendkomponente entfernt). Spektrum mit Daniell 11 Fenster gefiltert.

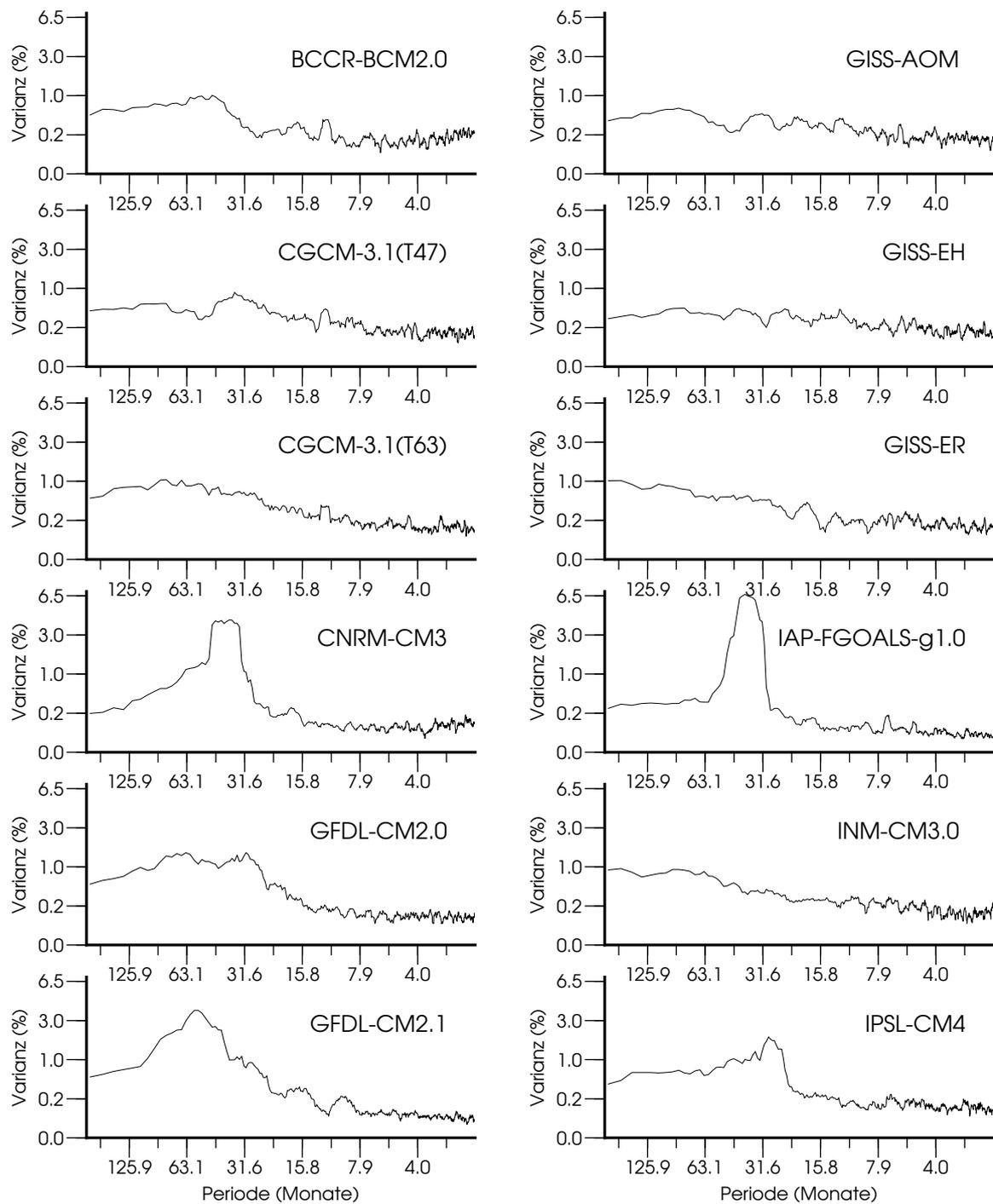
Arbeit, ist das Spektrum des SOI in erklärter Varianz pro Periode dargestellt. Allen Zeitreihen wurde vor der Berechnung des Spektrums sowohl Jahresgang als auch die über ein kubisches Polynom geschätzte Trendkomponente entfernt. Das auffälligste Merkmal und die typische Charakteristik des beobachteten SOI in der spektralen Darstellung ist der breite Peak im Spektrum, der die Dominanz von Oszillationen zwischen etwa zwei und acht Jahren anzeigt. Die Angaben zu diesen Perioden schwanken in der Literatur geringfügig (z. B. LAU und SHEU 1988, SALISBURY und WIMBUSH 1999). Hierbei handelt es sich um die typische, dem Klimasystem immanente Dauer eines ENSO-Zyklus. Neben dieser zyklischen, aber keinesfalls periodischen Komponente zeigt das Spektrum des SOI allerdings auch Variabilität auf nahezu allen erfassten Wellenlängen, wobei vor allem auf längeren

Zeitskalen weitere Zyklizität vermutet wird (DIAZ und MARKGRAF 2000). Somit stellt auch der SOI eine im wesentlichen chaotische, nicht vorhersagbare Zeitreihe dar. Das Muster dieser speziellen Version von Chaos ist allerdings so spezifisch, dass es, wie auch das räumliche Dipol im Luftdruck, für die Identifikation des ENSO-Systems herangezogen werden kann.

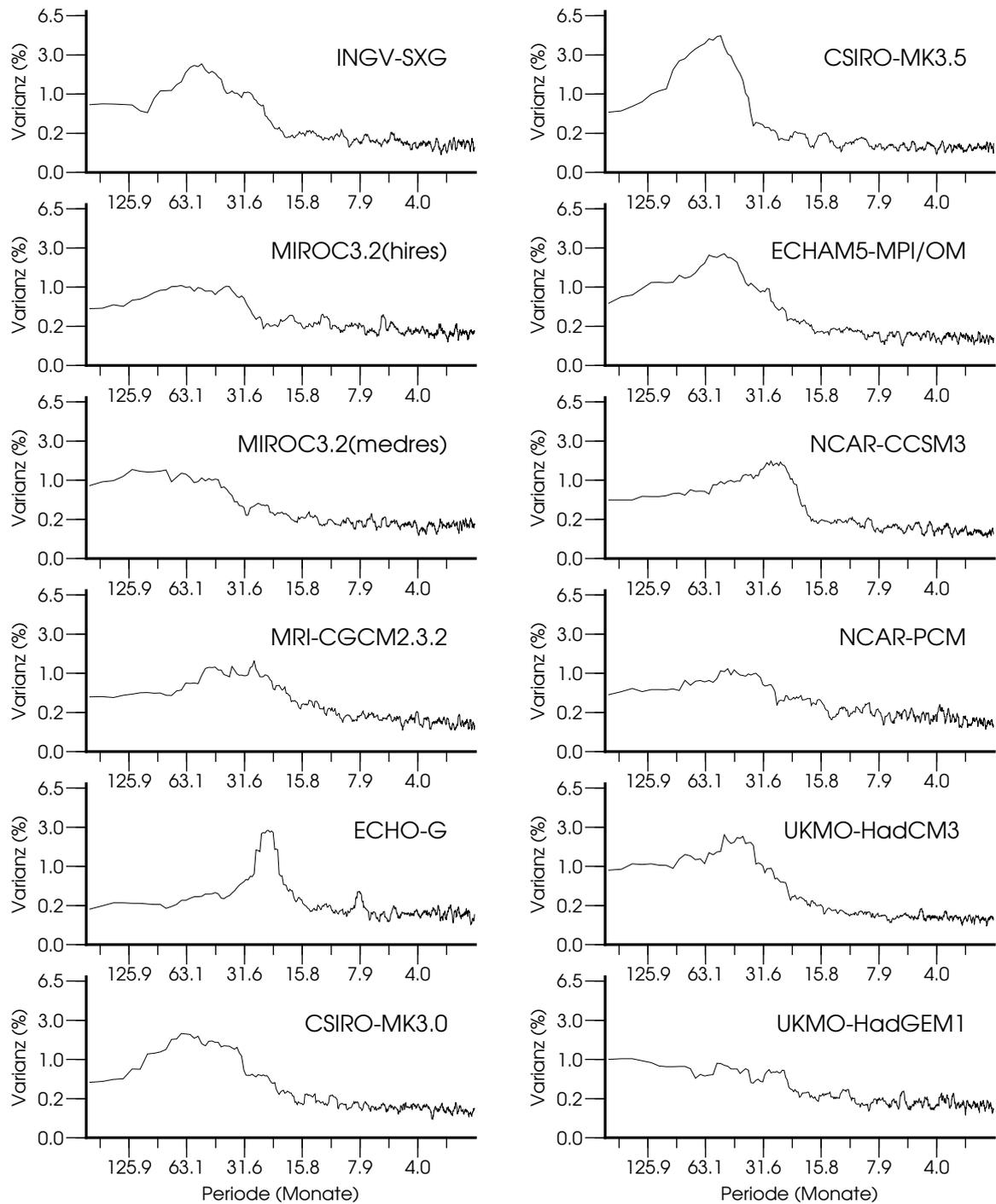
Für die AOGCMs erweist sich das Spektrum des SOI als eine charakteristische Größe, die in ihrer grundsätzlichen Form nicht vom zugrunde liegenden Zeitraum abhängt. Zudem stellt es für verschiedene Läufe eines Modells eine nahezu konstante Größe dar, während viele sonstige Kenngrößen durch variierende Anfangsbedingungen durchaus beeinflusst werden. Als Beispiel sei hier angeführt, dass die  $I(t, i)$  der vier Läufe des ECHAM5-MPI/OM de facto keinerlei linearen Zusammenhang aufweisen (die paarweisen Korrelationen befinden sich zwischen  $r = -0.06$ – $0.07$ ). Die Spektren  $S(P, i)$  sind jedoch quasi identisch, mit  $r = 0.91$ – $0.96$  für den Zeitraum 1900 bis 2000. Das Spektrum stellt also einen Aspekt der GCMs dar, der ganz wesentlich abhängig von der individuellen Modellphysik ist. Ein realistisches Spektrum wird hierbei dann so gedeutet, dass auch die Prozesse, die zu dessen Ausprägung führen realistisch durch das Modell nachvollzogen werden. Hier wird somit die Frequenzdomäne für den direkten Vergleich von Modell- und Beobachtungsdaten genutzt.

Insofern sind die Abb. 7.3 und Abb. 7.4 durchaus zum direkten Vergleich mit dem Spektrum des SOI vom CGD geeignet. Die weite Bandbreite, die das CMIP3 in dieser Größe auszeichnet, ist hierbei ein augenfälliges Merkmal. Grundsätzlich lassen sich die  $S(P, i)$  in drei Gruppen differenzieren. Eine dieser Gruppen ist dem  $S_B$  sehr ähnlich. Die Spektren weisen ein breites, aber deutliches Maximum an erklärter Varianz für Oszillationen mit Perioden zwischen ca. 20 und ca. 90 Monaten auf, was dem typischen Zyklus des ENSO-Phänomens gut entspricht. Dennoch sind die Amplituden in diesen Perioden nicht so stark ausgeprägt, dass das Verhalten der  $I(t, i)$  als zu regelmäßig angesehen werden muss. Somit lässt sich für diese Modelle nur eine grundsätzliche Tendenz zur Organisation der zeitlichen Varianz auf den bekannten Perioden feststellen. Die Spektren bilden eindeutig Zyklizität und keine Periodizität ab. Diese Spektren sind dem des NCAR-SOI in Abb. 7.2 somit in den entscheidenden Charakteristika sehr ähnlich.

Die zweite Gruppe von Modellen weist scheinbar überhaupt keinen bevorzugten Zyklus auf. Zwar verfügen auch diese Modelle gelegentlich über relative Maxima in ihren Spektren, diese sind jedoch im Vergleich zum  $S_B$  nur durch geringe Amplituden gekennzeichnet. Weiterhin unterscheiden sich diese typischerweise dadurch, dass sie sich auf eine geringere Anzahl von Perioden konzentrieren, was einen realistischen ENSO-Mechanismus als zugrundeliegenden Prozess ebenfalls wenig wahrscheinlich erscheinen lässt. Insgesamt erinnern diese Spektren am ehesten an jene, die durch weißes Rauschen hervorgerufen werden, was zwar die atmosphärischen Moden der Außertropen, nicht jedoch das gekoppelte System für das der SOI steht zufriedenstellend beschreibt.



**Abb. 7.3:** CMIP3: SOI-Varianzspektren (1). Wie Abb. 7.2 aber ein Lauf pro AOGCM in CMIP3.



**Abb. 7.4:** CMIP3 SOI-Varianzspektren (2). Wie Abb. 7.2 aber ein Lauf pro AOGCM in CMIP3.

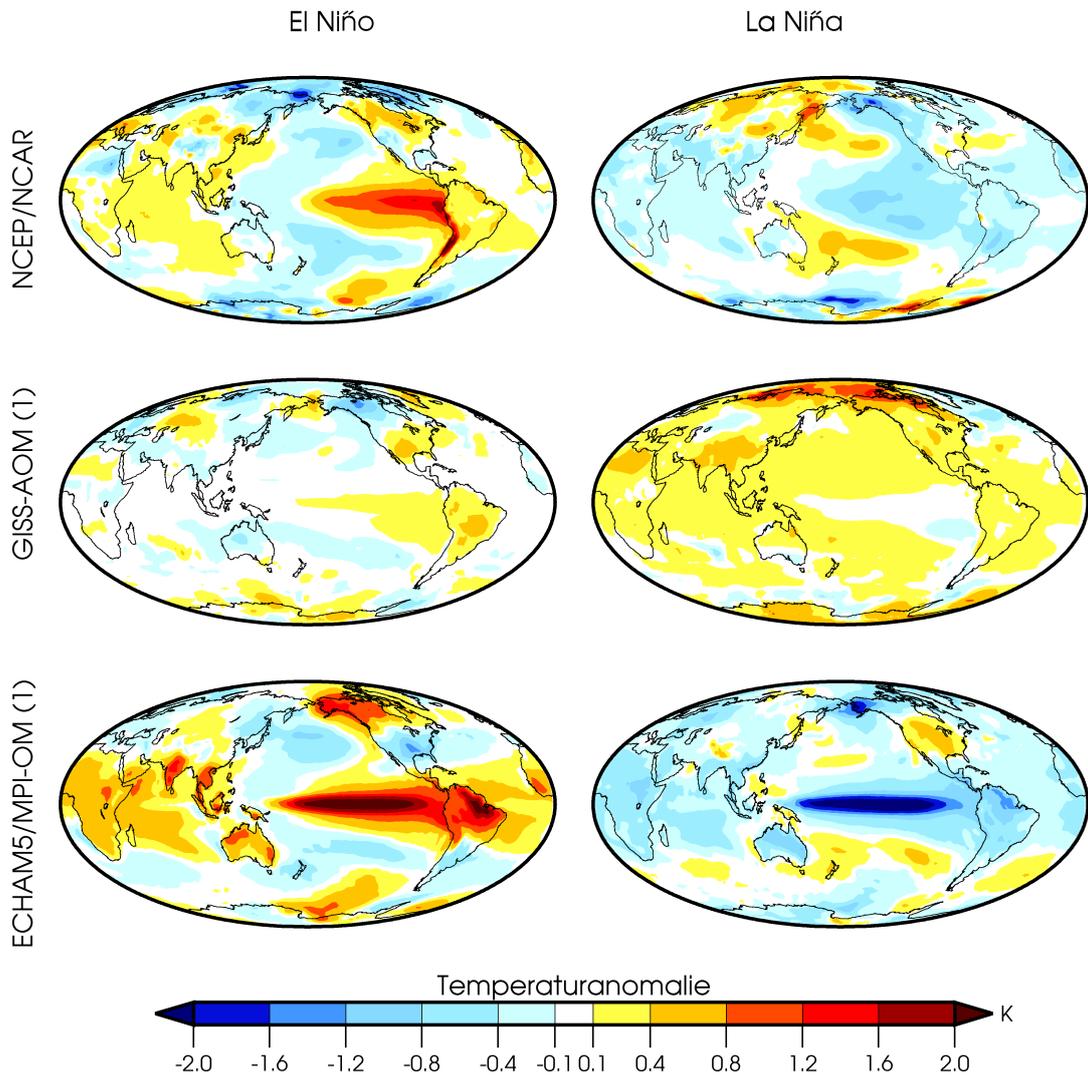
Neben diesen beiden, zumindest qualitativ recht eindeutig zu beurteilenden Gruppen existiert eine weitere, einfach zu identifizierende jedoch schwierig zu interpretierende, Gruppe. Es handelt sich hierbei um Modelle, deren SOI-Spektren auf einen sehr regelmäßigen ENSO-Zyklus hinweisen. Diese verfügen über steile Peaks für wenige Perioden. Angesichts der Anwendung des Daniell-Filters muss hierbei festgehalten werden, dass dieser Effekt in den Abbildungen eher abgeschwächt als prononciert erscheint. Diese enorme Konzentration der Variabilität der Zeitreihe  $I(t, i)$  lässt deren Verhalten im Vergleich mit Beobachtungen als zu zyklisch erscheinen. So offenbart etwa das Spektrum von IAP-FGOALS-g1.0 ein quasi-periodisches Verhalten mit einem zu regelmäßigen und im Vergleich auch zu kurzen Zyklus um etwa drei Jahre. In dieser Gruppe scheint also ein dominierender, ENSO-artiger Prozess im tropischen Pazifik zu existieren. Dieser weist allerdings deutliche Abweichungen von dem im realen Klimasystem auf.

In einigen Fällen unterscheiden sich die Spektren der AOGCMs Beobachtungsdaten dahingehend, dass der Jahresgang über das beschriebene Vorgehen weniger vollständig entfernt werden konnte. Diese auffällige Besonderheit äußert sich in geringen relativen Maxima der erklärten Varianz um die Periode von 12 Monaten. In diese Läufe ändert innerhalb des 20. Jahrhunderts die Variabilität der verschiedenen Monatswerte des SOI, was in den Beobachtungsdaten weniger intensiv stattfindet. Da dieser Effekt jedoch auf sämtliche Gruppen gleichermaßen verteilt ist, scheint er nicht originär mit dem modelleigenem ENSO in Beziehung zu stehen.

### 7.1.3 Telekonnektionen

Das ENSO-Phänomen hat zwar seinen Ursprung im tropischen Pazifik, ist aber über ein weites Netz an Telekonnektionen ein global wirksames Klimamuster. Am ausgeprägtesten sind seine Effekte während der extremen Zustände, den El Niño Warm- und den La Niña Kalt ereignissen. Die hierdurch hervorgerufenen Muster sind das Thema einer Vielzahl unterschiedlichster Untersuchungen (DIAZ und MARKGRAF 1992, 2000). Bei einer realistischen Erzeugung des ENSO-Mechanismus sollten sich diese speziellen und prominenten Effekte auch in den klimatischen Feldgrößen der AOGCMs abbilden. Einschränkend ist hierbei allerdings grundsätzlich festzuhalten, dass die Effekte vor allem in den Tropen ein typisches Muster aufweisen, während die Telekonnektionen der Southern Oscillation in den Außertropen im Allgemeinen nicht oder zumindest weniger stabil sind (z. B. BRÖNNIMANN et al. 2007). Dies bedeutet auch, dass der quantitative Abgleich von Modell- und Beobachtungsdaten, der schon angesichts des in diesem Fall zu berücksichtigten Datenumfanges als problematisch zu beurteilen ist, unter Umständen zu fehlerhaften Aussagen führt, die letztlich auf dem beschränkten Umfang an flächendeckenden Beobachtungsdaten beruhen. Entsprechend werden hier in einer eher qualitativen Weise die Differenzen zwischen AOGCMs und Beobachtungen in diesem Punkt demonstriert

So liegt hier der Fokus auf den Extremereignissen, die die Fernwirkungen am deut-



**Abb. 7.5:** El Niño und La Niña: Temperatur. Effekte und Telekonnektionen. Gemittelte Ereignisse für Beobachtungsdaten und zwei Modellläufe. Jeweils Anomalien der extremen Ereignisse im ENSO-Zyklus vom neutralen Zustand in Kelvin (Definition und Berechnung siehe Text). Bodennahe Temperatur, über den Ozeanen sind die Muster denen in der Meeresoberflächentemperatur sehr ähnlich. Langfristige Trends der  $T(t, s)$  entfernt.

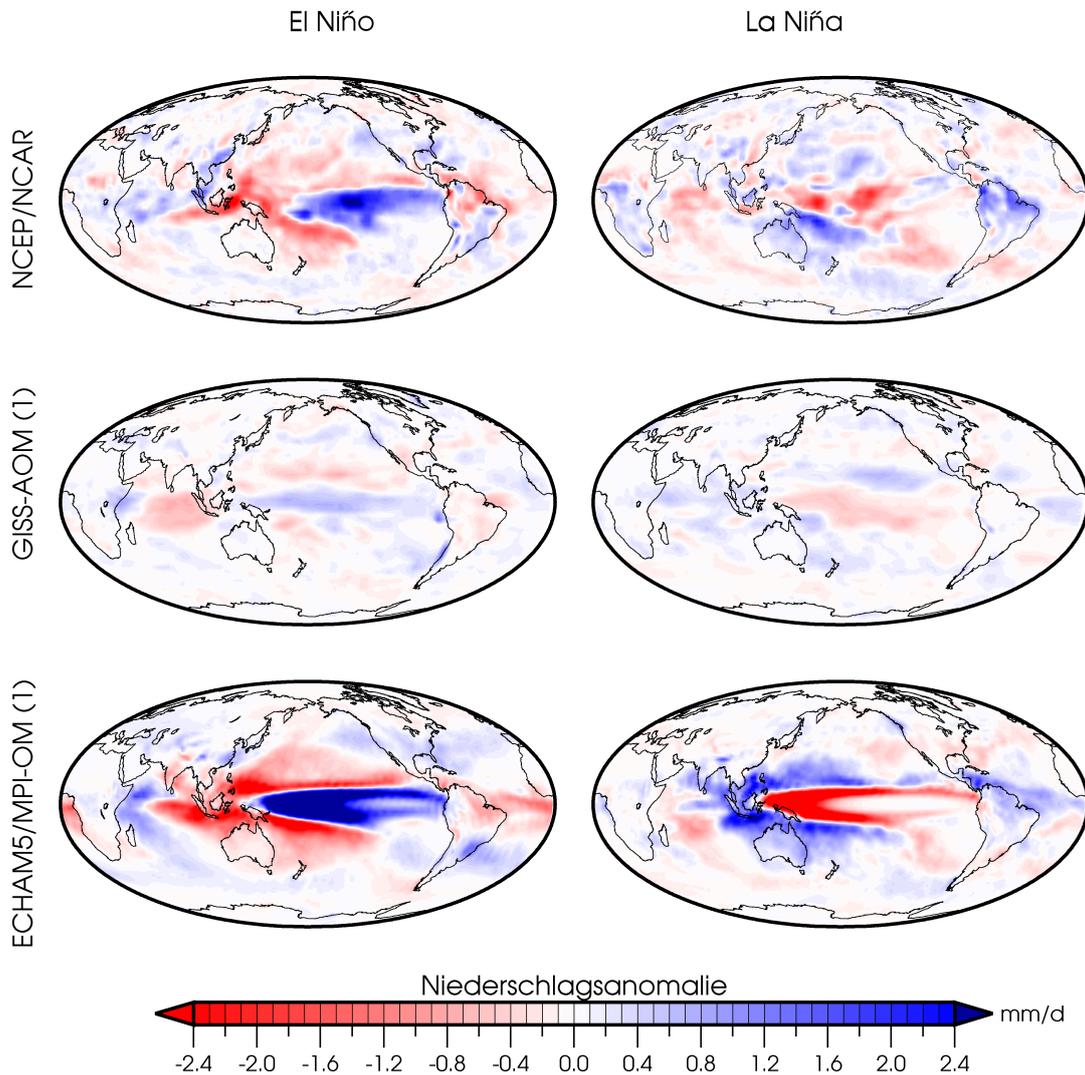
lichsten herausarbeiten sollten. ENSO als Ganzes, aber insbesondere die El Niño und La Niña Episoden, stellen keinen linearen Prozess dar. Aus diesem Grund wird hier deren Effekt nicht über Korrelations- oder Regressionsansätze zu erfassen versucht, sondern über die Bildung von Kompositen. Damit ist sichergestellt, dass beide Phasen gleichermaßen identifiziert werden können. Als Kriterium, wann ein

Ereignis in den Beobachtungen oder Modellen stattfindet, wurde die Ausprägung des SOI herangezogen. In Anlehnung an TRENBERTH (1997) wird die Zeitreihe aus standardisierten monatlichen Anomalien zur Basisperiode 1960–1999 mit einem 5-Monate Tiefpassfilter geglättet. Episoden von mindestens sechs aufeinanderfolgenden Monaten mit  $I_5(t) \leq -1$  werden als El Niño gewertet. La Niña Ereignisse werden über eben solange Episoden mit  $I_5(t) \geq 1$  definiert. Da hier, anders als bei TRENBERTH (1997) der atmosphärische Part zur Diagnose eines Extremereignisses herangezogen wird und zudem auch die Fernwirkungen betrachtet werden sollen, werden die beiden, auf ein nach diesen Regeln detektiertes Ereignis folgenden Monate noch in die entsprechende Phase miteinbezogen. Im SOI von NCAR und der 1. PC von HadSLP2 werden so in sehr guter Übereinstimmung auch mit anderen Definitionen die bekannten Ereignisse in der zweiten Hälfte des 20. und dem Beginn des 21. Jahrhunderts detektiert (vgl. KILADIS und VAN LOON 1988). Sowohl für die Beobachtungen, als auch für die Modelldaten, wurden sämtliche Ereignisse gemittelt und Komposite zu den neutralen Phasen mit  $|I_5(t)| < 1$  gebildet. Sämtliche Berechnungen wurden mit monatlichen Anomalien durchgeführt, die Zeitreihen der Indices und Felddaten wurden jeweils linear enttrentet, um den anthropogenen Einfluss auszuschließen.

Beispielhaft sind in Abb. 7.5 für die NCEP/NCAR Reanalysen und für zwei Modelle die räumliche Response auf die extremen ENSO-Ereignisse dargestellt. Anhand der Beobachtungsdaten ist offensichtlich, dass die beiden Phasen keine spiegelverkehrten Versionen des gleichen Musters darstellen. Sowohl in der Temperatur, als auch bei der Betrachtung des Niederschlags, ist das El Niño-Muster, vor allem in der Kernregion, deutlicher und prononcierter als die Effekte des Kaltereignisses. Die beiden Modellläufe wurden ausgewählt, weil sie die bestehenden Unterschiede im CMIP3 offensichtlich machen.

ECHAM5/MPI-OM reproduziert das beobachtete Muster räumlich überzeugend, Effekte und Telekonnektionen werden dabei jedoch überschätzt. Sicherlich ist insbesondere die La Niña-Phase in diesem AOGCM im Vergleich mit den Reanalysen als zu stark von der neutralen Phase abweichend zu bezeichnen. Gekoppelt an die heftigen Effekte auf die äquatoriale Temperatur, die mit bis zu -4 K die beobachtete Amplitude um etwa den Faktor 2 überschätzen, sind auch die Effekte auf den Niederschlag zu intensiv und wohl auch räumlich zu klar abgegrenzt. Grundsätzlich sind aber die regionalen und auch globalen Auswirkungen der Extremereignisse durchaus mit den Beobachtungen im Einklang, und es kann von einem wirkungsvollen und eventuell auch realistischen ENSO-Mechanismus ausgegangen werden.

GISS-AOM weist dagegen bereits bei der qualitativen Betrachtung größere Differenzen zu NCEP/NCAR in den Effekten auf beide Feldgrößen auf. Die El Niño-Phase weist eine deutlich geringere Temperaturerhöhung im Westpazifik auf als zu erwarten wäre. Die bekannten und relativ stabilen Effekte auf die Temperatur in Nordamerika sind im Vergleich mit den Beobachtungen geradezu



**Abb. 7.6:** El Niño- und La Niña: Niederschlag. Wie Abb. 7.5, aber für Niederschlagszeitreihen  $R(t, s)$  in mm/d

umgekehrt. Die La Niña-Phase in diesem Modell ist ebenfalls nur äußerst schwach an der zu erwarteten Stelle des Modells zu identifizieren. Die relative Abkühlung im äquatorialen Pazifik ist nur gering, dagegen scheint sich diese Phase im globalen Hinblick eher als relativ warm im Vergleich zum neutralen Zustand darzustellen. Diese Feststellungen müssen bei der Beurteilung des Realismus von ENSO in diesem Modell berücksichtigt werden. Die Effekte auf den hydrologischen Zyklus sind dagegen in ihrem räumlichen Muster mit den Beobachtungen besser zu vergleichen. Allerdings ist hierbei festzuhalten, dass die Beträge der Anomalien sowohl während

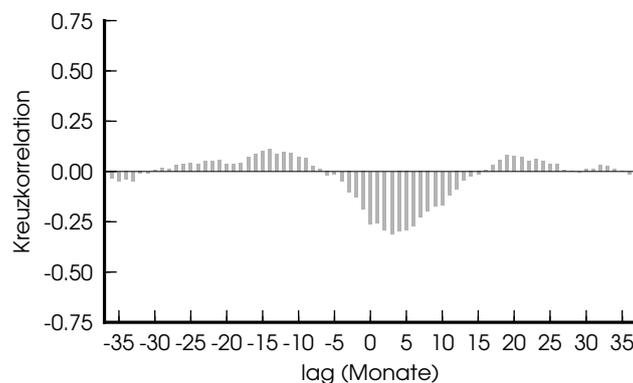
der El Niño- als auch der La Niña-Phasen deutlich unterschätzt werden. Jedoch kann dies auf die generell geringere räumliche Auflösung von GISS-AOM zurückgeführt werden. Dennoch ist sicherlich davon auszugehen, dass das ENSO-Phänomen in der Modellphysik dieses AOGCMs eine weniger prominente Rolle spielt als im realen Klimasystem.

Diese Beispiele illustrieren, dass auch im Hinblick auf Telekonnektionen und regionale Effekte schon in den beiden fundamentalen Feldgrößen Temperatur und Niederschlag merkliche, sicherlich auch qualitative, Unterschiede innerhalb des CMIP3 Ensembles bestehen. Die ENSO-typischen Effekte auf die regionalen und globalen Muster und Ausprägungen der Klimaelemente werden in vielen, auch hier nicht aufgezeigten Aspekten von einer bedeutenden Anzahl von Modellen nicht realitätsnah wiedergegeben (HAM und KUG 2012). Die hierfür gezeigten Abbildungen verdeutlichen zudem eindrucksvoll, dass diese Unterschiede auch bedeuten, dass ohne realistisches ENSO-System dem simulierten Klimasystem im Vergleich zum beobachteten ein wichtiger und aktiver Faktor fehlt. Diese Feststellung sollte etwas, aber auch nicht zu stark, abgemildert werden durch den Hinweis, dass diese Quelle der internen Variabilität des Klimasystems im CMIP3 angesichts des zweifelsfrei dominanten anthropogenen Einfluss auf die Klimaentwicklung des 21. Jahrhunderts nur als zweitrangig angesehen werden kann. Sicherlich wird der Fokus auf ENSO bei der Beurteilung der Modelle hierdurch eher gerechtfertigt als in Frage gestellt.

### 7.1.4 ENSO und die globale Temperatur

Das ENSO-Phänomen ist die einzige Mode im Klimasystem mit Auswirkungen auf die globale Oberflächentemperatur. Entsprechend der Definition des SOIs verfügt die Korrelation die diese Beziehung kennzeichnet über ein negatives Vorzeichen. Der eigentliche Grund hierfür ist, dass sich mit dem Nachlassen der Passatwinde der Warmpool im Westpazifik nach Osten ausdehnen kann (HALPERT und ROPELEWSKI 1992). Mit der Erhöhung der Meeresoberflächentemperatur verbunden ist auch die verstärkte Abgabe von Wärme an die Atmosphäre. Dagegen bedeutet die Verstärkung der Passate auch die Intensivierung des Upwellings von kalten Tiefenwasser vor der Westküste Südamerikas. Zudem drängen die stärkeren Ostwinde auch den Warmpool weiter gegen Indonesien, und dessen verringerte Ausdehnung bewirkt einen zusätzlichen Abfall der globalen Temperatur. In Jahren mit El Niño Bedingungen steigt so die globale Temperatur um bis zu 0.2 K an (TRENBERTH et al. 2002). Ebenso weisen La Niña Jahre entsprechend verringerte Temperaturen auf, ebenfalls als globale Auswirkungen der regionalen Temperaturreaktion in den Tropen (YULAEVA und WALLACE 1994). Aber auch abseits dieser Extremjahre und auf kürzeren, auch monatlichen Zeitskalen wirken die entsprechenden Mechanismen, die sich über eine Kreuzkorrelation zwischen der Zeitreihe des SOI oder anderer ENSO-Indexreihen und  $T_E$  beschreiben lassen (z. B. THOMPSON et al. 2008).

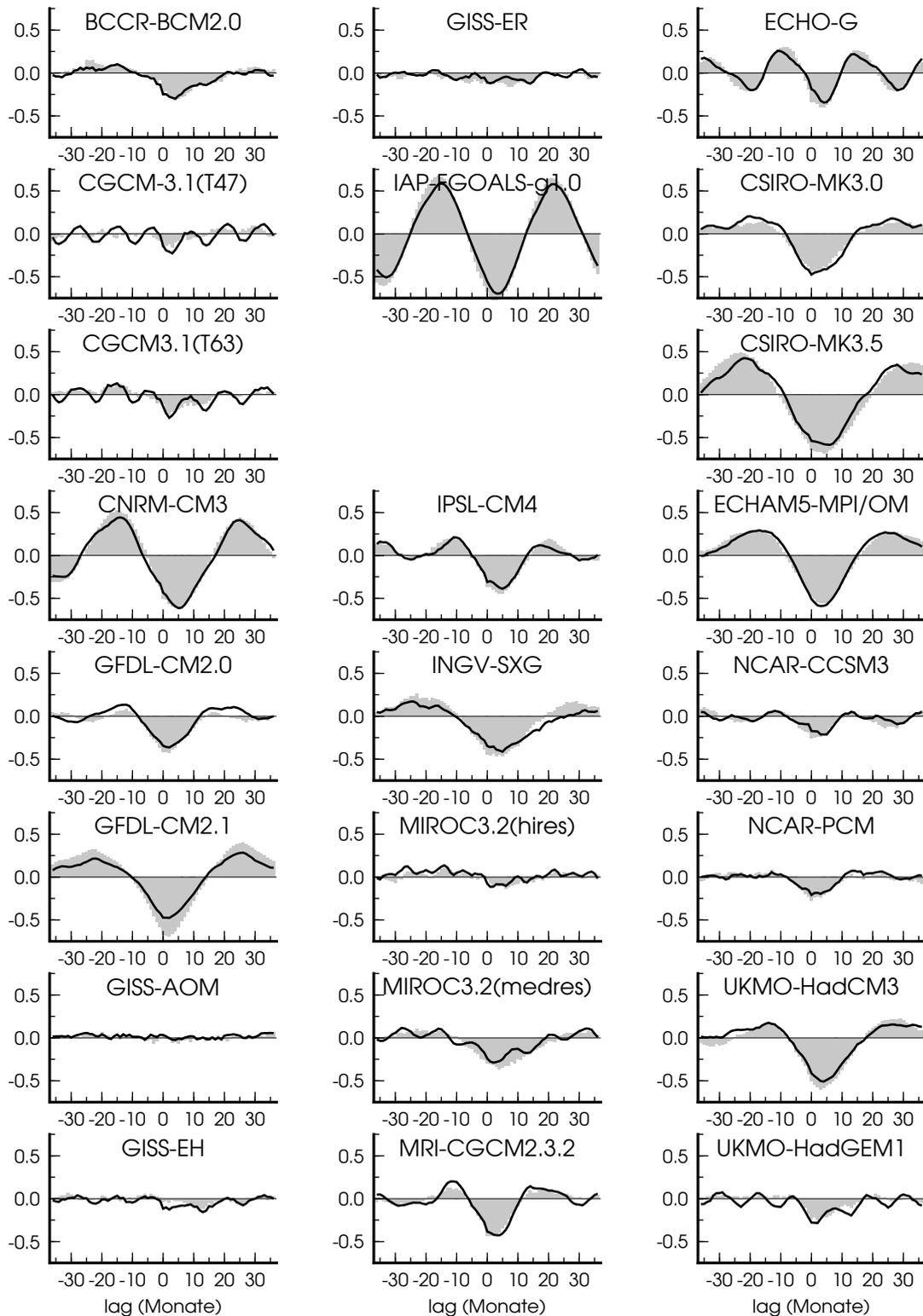
Diese Kopplung wurde mehrfach beschrieben, und verfügt bei der statistischen Modellierung der Globaltemperatur über ein signifikantes Erklärungspotential (vgl. LEAN und RIND 2008, FOSTER und RAHMSTORF 2011). Ebenso kann die dynamische Modellierung der globalen Temperatur durch die künstliche Herbeiführung des beobachteten Verhaltens von ENSO, also die Überführung eines internen Variabilitätssignals in ein externes Forcing wie die Treibhausgasemissionen oder die Effekte des explosiven Vulkanismus, dramatisch verbessert werden (KOSAKA und XIE 2013). In Kombination mit seinem quasi-zyklischen Verhalten bildet dies eine nutzbare Grundlage für mittelfristige, etwa dekadische Prognosen für  $T_E$  (z. B. SMITH et al. 2007). Der maximale Temperatureffekt folgt dabei dem SOI um einige



**Abb. 7.7:** NCAR-SOI und Gistemp- $T_E$ . Die Kreuzkorrelationsfunktion  $r_K(h)$  mit  $I(t)$  und  $T_E(t+h)$  hat den maximalen Betrag bei  $h = 3$ . Zeitreihen der monatliche Anomalien (kubische Trends entfernt) zur Referenzperiode 1960–1999. Analyse für den Zeitraum 1900–2000.

Monate versetzt. In diesem Kapitel wird jeweils die Kreuzkorrelation  $r_K(h)$  für die Zeitreihen  $T_E(t+h)$  und  $I(t)$  gezeigt. Abb.7.7 weist den maximalen Einfluss für ein lag von  $h = 3$  aus, jedoch zeichnet sich der Einfluss des SOI auch ohne Verschiebung der Zeitreihen ab. Die, wegen der fehlenden perfekten Periodizität des SOIs deutlich schwächere, Korrelation um  $h = 20$  kann als das Umschlagen des ENSO-Systems in den jeweils anderen Zustand interpretiert werden. Die scheinbar unsinnigen Korrelationen bei negativem lag, wenn  $T_E$  den SOI führt, sind in der mehrmonatigen Erhaltungsneigung des ENSO-Phänomens begründet. Die Kreuzkorrelationsfunktion in Abb. 7.7 stellt somit eine Zusammenfassung vielfältiger, komplexer Vorgänge im Klimasystem dar. Sichtbar wird nicht nur die zeitliche Struktur der Dynamik des SOI, sondern auch die Umsetzung dieses internen Forcings in ein Temperatursignal durch Atmosphäre und Ozean.

Klimamodelle die über einen realistischen ENSO-Mechanismus verfügen, sollten diese Beziehung entsprechend abbilden. Abb. 7.8 gibt einen Überblick über die



**Abb. 7.8:** SOI und  $T_E$  im SRES A1b. Kreuzkorrelationsfunktionen für 1900–2000 und 1900–2098, ein Lauf pro Modell. Erläuterungen siehe Text.

unterschiedlichen Ausprägungen der Kreuzkorrelation im Multi-Modell Ensemble. Die Berechnungen erfolgten ebenfalls mit kubisch enttrenteten Anomalien zur Periode 1960–1999. Gezeigt wird die Funktion jeweils für das 20. Jahrhundert und den Gesamtzeitraum, um eventuelle Veränderungen der Beziehung zwischen diesen Größen unter dem anthropogenen Forcing zu verdeutlichen. Auch bei dieser ENSO-bezogenen Kenngröße sind grundlegende Unterschiede zwischen den einzelnen Modellen augenfällig. Qualitativ wird die beobachtete Kreuzkorrelation vom größeren Teil der AOGCMs durchaus nachvollzogen. Nur wenige Modelle, etwa GISS-AOM und GISS-EH, weisen im 20. Jahrhundert kein betragsmäßiges Maximum mit negativen Vorzeichen für ein lag von  $0 < h < 7$  Monaten auf, was als Quantifizierung die grundsätzliche Beziehung zwischen den beiden Größen sinnerfassend beschreibt.

Das Umschlagen des ENSO-Systems in den jeweils anderen Zustand, also den Vorzeichenwechsel des SOIs, das sich in  $r_K(h)$  durch die positiven Korrelationen um die lag  $h = \pm 20$  abbildet, wird dagegen bereits von mehreren Modellen nicht mehr realitätsnah abgebildet. In Abb. 7.8 ist ebenfalls eindeutig zu erkennen, dass in vielen AOGCMs die Beziehung zwischen dem SOI und  $T_E$  sogar deutlich ausgeprägter ist als in den Beobachtungsdaten. Dies könnte darauf zurückzuführen sein, dass die Modellrandbedingungen einige reale Einflussgrößen auf die globale Temperatur, insbesondere die Sonnenzyklen oder auch den Einfluss der Aerosole, nicht realistisch beinhalten. Somit würde der relative Effekt des ENSO-Systems als Modulator von  $T_E(t)$  nicht von anderen Einflüssen überlagert und dementsprechend zunehmen. Zu ausgeprägte Korrelationen im Vergleich mit denen in Abb. 7.7 müssen also nicht zwangsläufig als ein unrealistisches Element im AOGCMs gewertet werden, sondern werden durchaus erwartet. In einigen Fällen ist  $r_K(h)$  auch als sehr regelmäßig im Vergleich mit den Beobachtungen zu bezeichnen. Ein Beispiel für eine relativ regelmäßige Kreuzkorrelation ist etwa die im ECHAM5-MPI/OM, welches den Zusammenhang mit  $r_K(3) \approx -0.5$  auch deutlich stärker als beobachtet simuliert. Dieses Verhalten ist teilweise ebenfalls über die nicht perfekten Randbedingungen erklärbar. Daneben zeichnen sich hierbei allerdings auch die Spektren der modelleigenen SOIs ab. Dies ist augenfällig bei ECHO-G, dessen ausgeprägter und zu kurzer SOI-Zyklus auch  $r_K(h)$  dominiert. Noch extremer findet sich dieses Phänomen allerdings bei IAP-FGOALS-g1.0, dessen quasi-periodischer SOI mit  $T_E$  Kreuzkorrelationen zwischen 0.7 und -0.7 bei auffälliger Taktung erzeugt. Die Realitätsnähe der verschiedenen  $r_K(h)$  ist somit sicherlich als variabel zu bewerten. Ebenso kann hier festgehalten werden, dass einige Modelle scheinbar keinerlei systematische Beziehung für die Zeitreihen  $I(t, i)$  und  $T_E(t, i)$  aufweisen. Dies ist umso erstaunlicher, als das auch in diesen Modellen der Einfluss des ENSO-Phänomens durch das Fehlen sonstiger externer Forcingfaktoren eher hervorgehoben werden sollte. Bei einigen Modellen ändert sich auch merklich die Struktur der Kreuzkorrelation, wenn der Gesamtzeitraum statt dem 21. Jahrhundert betrachtet wird. Dies mag auf individuelle Reaktionen auf den steigenden Strahlungsantrieb bewertet werden,

stellt im CMIP3-Ensemble allerdings eine Ausnahme dar, die bevorzugt bei solchen GCMs auftritt, die einen eher wenig ausgeprägten Zusammenhang zwischen  $I(t, i)$  und  $T_E(t, i)$  aufweisen.

## 7.2 ENSO als Aspekt der Modell-Performance

Die aufgezeigten Differenzen rechtfertigen sicherlich die Aussage, dass mit ENSO ein Aspekt des Klimasystems identifiziert wurde, für den sich innerhalb der CMIP3-Modelle deutliche Unterschiede in seiner Umsetzung aufzeigen lassen. Das Heranziehen dieser Unterschiede um die Gesamtperformance der AOGCMs zu bewerten erfährt seine Berechtigung durch die prominente Rolle der ENSO-Mode im realen Klimasystem, der vielfältigen gekoppelten Prozesse die zur Erzeugung eines realistischen ENSO-Systems notwendig ist und der unklaren Rolle, die ihr in den aktuellen Langzeitprognosen zum anthropogenen Klimawandel zukommt. Weiterhin spricht, nicht nur im Hinblick auf die sonstigen hier betrachteten Größen, auch das Ausschlussprinzip für dieses Vorgehen.

Hier wird nun auf dieser Basis eine Funktion entwickelt, die die Abbildung des ENSO-Systems durch die GCMs bewerten soll. Diese Funktion soll metrisch skaliert sein. Sie dient im Folgenden unterschiedlichen Zwecken

- der Erfassung einer Rangfolge der Modellperformance im Hinblick auf den Realismus des modelleigenen ENSO-Systems
- der Bildung zweier Teilensembles  $A$  und  $B$ , die sich durch ihre Abbildung des ENSO-Phänomens unterscheiden
- als Basis für die numerische Gewichtung der einzelnen GCMs bei der Auswertung des Multi-Modell Ensemble

Es bestehen nahezu unbegrenzte Möglichkeiten die vorgestellten Resultate, sonstige Analysen in diesem Kontext, sowie Verfahren und Erkenntnisse aus der Literatur bei diesem Unterfangen zu nutzen. Hier wird ein originäres Vorgehen gewählt, welches sicherlich nicht das einzig- oder auch nur best-mögliche darstellt, sich jedoch als stabil gegenüber Änderungen der Basisperiode oder des Umfang des Ensembles erwiesen hat. Die Wahl stellt jedoch eine subjektive Entscheidung dar. Der konkrete Ansatz wird im Folgenden vorgestellt und begründet.

### 7.2.1 Erläuterung des Ansatzes

Die Grundlage der Bewertung der Modellperformance bildet der SOI, genauer dessen Spektraldichtefunktion. Für diese Auswahl sprechen einige wichtige Gründe. Zunächst existieren für diese Zeitreihe sicherlich deutlich zuverlässigere Daten als für die Feldgrößen, die in Kapitel 7.1 im Fall der Beobachtungsdaten herangezogen

wurden. Somit können die Referenzdaten als Fehlerquelle weitestgehend ausgeschlossen werden. Daneben stellt das spektrale Verhalten des SOI einen klaren Hinweis darauf dar, ob die wesentlichen Dynamiken und Aspekte der gekoppelten Klimamode tatsächlich befriedigend reproduziert werden. Zudem eröffnet das Spektrum die Möglichkeit, nachvollziehbare und einfach reproduzierbare Bewertungsschemata zu entwickeln, da nur eine Größe berücksichtigt wird und keine, letztlich wiederum subjektive, Abwägung darüber getroffen werden muss, welche Aspekte von höherer Bedeutung für die Qualität der ENSO-Abbildung sind. Weiterhin bedingt die Berechnung des SOIs über die Projektion der simulierten SLP-Daten auf die 1. EOF des HadSLP2, dass der Index, so das Modell ein ENSO-System besitzt welches er erfassen kann, ein realistisches Raummuster aufweist.

Ob die Organisation der Variabilität in den GCMs derart ist, dass dieses Raummuster für den tropischen Pazifik auch die dominierende Mode darstellt wird, konsequent mit der Definition der außertropischen Indices, als weniger wichtig eingestuft als die grundsätzliche Tatsache, ob eine solche Variabilitätsmode in den Modellen überhaupt existiert. Insofern wird die Organisation der Klimavariabilität in dieser Region, und somit die Raumdomäne, durch den SOI als Kriterium implizit durchaus mitberücksichtigt.

Die Telekonnektionen werden dagegen nicht für die Bildung der Bewertungsfunktion verwendet. Technisch spricht hier zunächst der hohe Aufwand zur Validation sowie teilweise fehlende Variablen für einzelne Läufe gegen dieses Vorgehen. Ebenfalls muss die Qualität der Beobachtungsdaten als Einschränkung in Betracht gezogen werden. Insbesondere aber die unklare Beurteilung der Wichtigkeit der diversen Telekonnektionen sowie deren bekannte Nicht-Stationarität im realen Klimasystem ließen die Erstellung einer hierauf basierenden Rangfolge oder Gewichtung im Rahmen dieser Arbeit, gerade angesichts des Datenumfangs, als nicht umsetzbar erscheinen.

Die Auswirkungen des ENSO-Phänomens auf die globale Oberflächentemperatur schließlich wird als unabhängige Kontrolle verwendet, ob die Bewertungsfunktion die Aussagen des SRES A1b Ensembles im Hinblick auf die Entwicklung von  $T_E$  beeinflusst. Ist dies der Fall, kann der Bewertungsfunktion sicherlich eine gewisse Relevanz zugesprochen werden.

### 7.2.2 Bewertungs- und Gewichtungsfunktion

Im üblichen Multi-Modell Ansatz werden alle Läufe als gleich realistisch angenommen. Diesen Fall kann man als Realisation einer Bewertungsfunktion  $f_\omega(i)$  betrachten, mit  $f_\omega(1) = f_\omega(2) = \dots = f_\omega(n) = 1$ . Seinen Niederschlag bei der Kombination verschiedener Klimamodellsimulationen findet dieses Modell in der Verwendung des arithmetischen Mittelwerts entsprechend den Gewichtungsfaktoren  $\omega_1 = \omega_2 = \dots = \omega_n = n^{-1}$ . Dies verdeutlicht die direkte Überführbarkeit einer Bewertungs- in eine Gewichtungsfunktion, die im Folgenden praktiziert wird.

Die eigentliche Metrik dieser Funktion basiert auf der Forderung, dass der simulierte SOI die gleichen charakteristischen Periodizitäten in seiner zeitlichen Entwicklung aufweisen soll, wenn er die gleichen physikalischen Prozesse repräsentiert. Ist dies der Fall bedeutet dies in der Frequenzdomäne eine große Ähnlichkeit zwischen dem beobachteten Spektrum  $S_B$  und dem Spektrum des  $i$ -ten Modelllaufs im Ensemble  $S(i)$ . Als Referenz dient der Index vom CPC, die Ergebnisse sind jedoch nicht durch diese Wahl beeinflusst und mit anderen SOI-Zeitreihen in guter Übereinstimmung reproduzierbar. Der Zusammenhang wird quantifiziert, wofür der normale Korrelationskoeffizient  $r(i)$  als Proximitätsmaß verwendet wird. Als Wert der Bewertungsfunktion für die Ähnlichkeit der beiden Spektren wird, mit der Eulerschen Zahl  $e$ , die Exponentialfunktion

$$f_\omega(i) = e^{r(i)} \quad (7.1)$$

verwendet. Da der Wertebereich von  $r(i)$  zwischen -1 und +1 liegt gilt  $f_\omega(i) \in [e^{-1}, e]$ . Ein Modelllauf, dessen Spektrum als perfekt linear anhängig mit  $S_B$  erscheint, erfährt hierbei eine um den Faktor  $e^2 \approx 7.3891$  höhere Bewertung als ein Lauf, dessen Verhalten in der Frequenzdomäne perfekt gegensätzlich zu dem des NCAR-SOI erscheint. Hier liegt also der Fokus, anders als bei der Korrelationsrechnung, auf Ähnlichkeit, nicht dem linearen Zusammenhang der zur Erklärung der Variabilität der Datensätze herangezogen werden kann. Die Berechnung der  $r(i)$  erfolgt mit den erklärten Varianzanteilen für Perioden zwischen 18 und 96 Monaten, um den Fokus auf den prominenten ENSO-Zyklus zu richten.

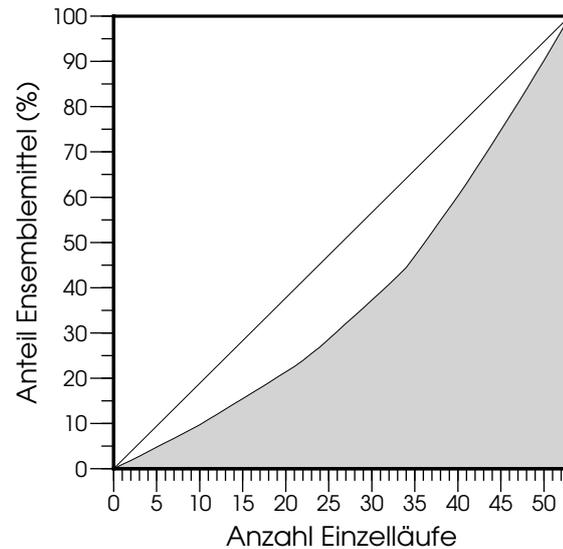
Da das Ziel aber die Bewertung der Modelle und nicht der Modellläufe ist, werden die, allerdings geringen, numerischen Unterschiede für diese Korrelation zwischen verschiedenen Läufen eines Modell Ensembles ignoriert. Statt für jeden Lauf eines Modells  $f_\omega(i)$  zu bestimmen wird der Wert für die Bewertungsfunktion für das AOGCM berechnet und dann für alle Läufe gleichermaßen übernommen. Hierfür wird die mittlere Korrelation  $\overline{r(i)}$  in Gleichung 7.1 verwendet. Hierfür werden die betroffenen  $r(i)$  zunächst über die Fisher Z-Transformation umgewandelt und das so errechnete Mittel anschließend wieder in einen Korrelationskoeffizienten umgerechnet (vgl. Kapitel 3.2). Für alle Läufe des betreffenden GCMs wird  $f_\omega(i) = e^{\overline{r(i)}}$  gesetzt. Über Normierung dieser Funktionen werden dann die Gewichtungsfaktoren  $\omega_i$  festgelegt.

$$\omega_i = \frac{f_\omega(i)}{\sum_{i=1}^n f_\omega(i)} \quad (7.2)$$

Abb. 7.9 zeigt diese Gewichtungsfaktoren für 53 Läufe von CMIP3als Lorenzkurve im Vergleich zur Verwendung des arithmetischen Mittelwerts. Trotz dieser offensichtlichen Unterschiede in den Gewichtungen stellt dieses Verfahren einen eher konservativen Bewertungsansatz dar.

Ein Grund für dieses Vorgehen ist, dass der SOI nur ein, eventuell auch ungeeigneter, Proxy für die physikalischen Eigenschaften eines AOGCMs ist. Weiterhin erwies sich die Bewertungsfunktion als stabil bei der Veränderung des Zeitraumes für den diese aufgestellt wurde, sowohl im Hinblick auf die Rangfolge als auch die konkreten Werte von  $f_\omega(i)$ . Andere Proximitäts- oder auch Distanzmaße, konkret der RMSE und Metriken auf Grundlage des  $\chi^2$ -Tests, zeigten dagegen eine deutliche Sensitivität gegenüber

den den Spektren zugrunde liegenden Zeiträumen. Das gilt vor allem im Hinblick auf die Ausprägungen von  $f_\omega(i)$ , aber durchaus auch bei der resultierenden Rangfolge. Somit werden diese Methoden als stark durch methodische Artefakte geprägt beurteilt und nicht weiter betrachtet. Obwohl eine stärkere Ungleichgewichtung der einzelnen Modelle als durchaus erstrebenswert erscheint, wird diese also zugunsten der Stabilität des Verfahrens verhindert. Außer bei der numerischen Gewichtung der Einzelläufe im SRES A1b werden die  $\omega_i$  auch zur Abgrenzung zweier Teilensembles genutzt. Diese werden insbesondere zur Überprüfung der Sensitivität des durch die Modelle simulierten Klimawandels gegenüber dem ENSO-System verwendet. Im Teilensemble *A* werden jene Modelle gebündelt, für die gilt  $\omega_i > n^{-1}$ , die übrigen Modelle bilden das Teilensemble *B*. Als Ergebnis der hier vorgenommenen Untersuchungen werden diese derart interpretiert, dass in *A* der Einfluss der dominierenden internen Variabilitätsmode realistischer abgebildet wird als in *B*. Eventuell findet sich in einigen Modellen von *B* auch überhaupt kein realitätsnahes ENSO-Phänomen, dieser Frage wird allerdings hier nicht weiter nachgegangen. Der Grund für diese relativ vorsichtige Charakterisierung von *A* und *B* ist, dass die in Kapitel 7.1 aufgezeigten Charakteristika des ENSO-System eindeutig nicht sämtlich durch alle Läufe von *A* realitätsnah abgebildet werden, während alle Modelle die Teilensemble *B* bei sämtlichen Aspekten scheitern. Stattdessen erweisen sich einige Modelle als hochgradig ambivalent.



**Abb. 7.9:** Kumulierte Bewertungsfunktion für CMIP3. Lorenzkurve: ermöglicht den Vergleich mit dem üblichen Ansatz alle Läufe mit dem gleichen Gewicht bei der Mittelbildung heranzuziehen.

## 7.3 Vergleich mit anderen Ansätzen

Das vorgestellte Bewertungsverfahren, sowohl das Kriterium als auch die Metrik, verfügt über einen heuristischen Charakter. Solche Verfahren sind grundsätzlich vor ihrer Anwendung auf ihre Zulässigkeit zu prüfen, wenn die auf ihnen basierenden Ergebnisse mit einer irgendwie gearteten Sicherheit und Legitimität interpretiert werden sollen. An dieser Stelle wird daher geprüft, inwieweit sich die aus ihr resultierenden Bewertungen sowohl speziell im Hinblick auf das ENSO-System, als auch die Einschätzung der gesamten Modellqualität, mit den Ergebnissen aus anderen Ansätzen vergleichen lassen.

### 7.3.1 Das ENSO-Problem des CMIP3

In einer umfangreichen Untersuchung haben OLDENBORGH et al. (2005) ENSO für einen Teil der CMIP3-Modelle anhand verschiedener physikalischer Feldgrößen sowie einer SST-basierten Indexzeitreihe analysiert. Hierbei konnten nicht nur vielfältige Unterschiede zwischen den diversen Modellen offengelegt werden, sondern auch deren Unterschiede im Vergleich zu den beobachteten Charakteristika festgestellt werden. Im Resultat steht keine komplette Rangfolge der Qualität der ENSO-Repräsentation, jedoch wird einigen Modelle eine deutlich höhere Qualität attestiert. Hierbei handelt es sich um GFDL-CM2.0, GFDL-CM2.1 ECHAM5/MPI-OM und UKMO-HadCM3. Etwas weniger gut, aber ebenfalls als überdurchschnittlich, werden MIROC3.2(hires) und MIROC3.2(medres) beurteilt. Mit Ausnahme des britischen Modells werden alle diese AOGCMs in das Teilensemble *A* eingeteilt und somit ebenso durch die hier entwickelte Metrik als im Vergleich mit dem gesamten CMIP3 überdurchschnittlich gut beurteilt. Es sei darauf hingewiesen, dass das Teilensemble *A* daneben auch vier Modelle umfasst, die in OLDENBORGH et al. (2005) nicht berücksichtigt wurden und für die somit kein Vergleich der beiden Bewertungsmethoden möglich ist. Die Übertragung des Verfahrens auf diese dort nicht betrachteten Modelle war nicht möglich. Als prohibitiv erschien zum einen der Datenbedarf, zum andern wäre dies aufgrund von fehlenden objektiven Entscheidungskriterien in dieser Studie auch nicht möglich gewesen. Diese beiden Gründe verhinderten folglich auch die Erstellung einer Rangfolge nach der Methode von OLDENBORGH et al. (2005). Es kann aber festgehalten werden, dass in wesentlichen Teilen die durch  $f_{\omega}(i) < 1$  dem Teilensemble *B* zugeordneten AOGCMs auch in der angeführten Studie als Modelle mit weniger realistischen ENSO eingestuft werden. Dem hier bereits als Beispiel für ein wenig realistisches ENSO-System aufgeführten GISS-AOM wird in der angeführten Studie sogar die elementare, typische Variabilität abgesprochen.

Trotz der angesprochenen Differenzen zwischen den Ergebnissen, der unterschiedlichen Anzahl der analysierten Modelle und natürlich der vollkommen anderen Fragestellung ist das Urteil der beider Verfahren im Wesentlichen von einer sehr

**Tab. 7.1:** CMIP3: Gruppierung nach ENSO. In den Spalten E wird angegeben ob das AOGCM dem Teilensemble *A* oder *B* zugeordnet wird. In den O-Spalten die Beurteilung der für ENSO-Performance durch OLDENBORGH et al. (2005): n.b.: nicht berücksichtigt; k.a.: berücksichtigt, aber keine explizit positive Beurteilung; m: mittleres Vertrauen in Realismus von ENSO formuliert; h: wie m, aber hohes Vertrauen. COLLINS et al. (2010) fassen die mit m und h gekennzeichneten Modelle als jene mit den besten Simulationen von ENSO zusammen.

	E	O		E	O		E	O
BCCR-BCM2.0	A	n.b.	GISS-ER	A	k.a.	ECHO-G	B	n.b.
CGCM3.1(T47)	B	k.a.	FGOALS-g1.0	B	k.a.	CSIRO-MK3.0	A	k.a.
CGCM3.1(T63)	A	n.b.	INM-CM3.0	B	k.a.	CSIRO-MK3.5	A	n.b.
CNRM-CM3	B	k.a.	IPSL-CM4	B	k.a.	ECHAM5-MPI/OM	A	h
GFDL-CM2.0	A	h	INGV-SXG	A	n.b.	NCAR-CCSM3	B	k.a.
GFDL-CM2.1	A	h	MIROC3.2(hires)	A	m	NCAR-PCM	B	k.a.
GISS-AOM	B	k.a.	MIROC3.2(medres)	A	m	UKMO-HadCM3	B	h
GISS-EH	B	k.a.	MRI-CGCM2.3.2	B	k.a.	UKMO-HadGEM1	A	k.a.

guten Übereinstimmung gekennzeichnet (ca. 80 %; vgl. Tab. 7.1. In den wenigen Punkten, in denen sich tatsächlich Unterschiede zwischen den beiden Methoden ergeben, ist dies vermutlich auf die hier natürlich stark verknappte Bewertungsfunktion zurückzuführen. Dennoch wurde auf eine Anpassung der Ergebnisse verzichtet. Dies geschah gemäß der Überlegung, dass hier die Betrachtung auf einer großen Anzahl von Modellen beruht und das letztliche Interesse an der Kombination derer Ergebnisse beruht. Hierbei erschien eine nachvollziehbare, einfache und konsequente Gewichtungsfunktion sinnvoller als ein nachträgliches Vermischen der methodischen Ansätze. Bei der Betrachtung des kompletten CMIP3, bzw. der Teilensembles *A* und *B*, sollten die wenigen Abweichungen hierbei ohnehin keinen zu großen Einfluss ausüben. Somit kann in dieser Arbeit auch eine in sich geschlossene, eigenständige Methodik präsentiert werden. IRVING et al. (2012) erstellen, für den tropischen Pazifik eine Liste mit weniger geeigneten AOGCMs für die Abschätzung des regionalen Klimawandels. Von dieser Negativ-Liste finden sich  $\frac{2}{3}$  im Teilensemble *B* wieder. Die Hälfte der Nichtübereinstimmung lässt sich dadurch erklären, dass die Liste auch auf OLDENBORGH et al. (2005) basiert, und eine der wenigen Differenzen zwischen den beiden ENSO-Bewertungen zum Ausschluss des AOGCMs führt.

### 7.3.2 Qualität der Modelle: Rangplatzfolgen

Allerdings wird hier nicht nur der Realismus des modelleigenen ENSO beurteilt, sondern dieses zur Abstufung der Modellperformance herangezogen. Dies ist ein neuer, noch nicht verfolgter Ansatz, so dass hierfür auch keine direkten Vergleichsmöglichkeiten bestehen. Es existieren allerdings für das CMIP3 zwei veröffentlichte Bewertungen der Modellperformance. Diese beruhen auf verschiedenen Metriken für

das Verhalten der diversen Felddaten im Vergleich mit den entsprechenden Größen aus Beobachtungs-, genauer Reanalysedatensätzen. Hier werden der Model Climate Performance Index (MCPI) von GLECKLER et al. (2008) sowie der Performance Index  $I^2$  nach REICHLER und KIM (2008) als Referenz für die aus Kapitel 7.2 resultierende Rangfolge der Modellperformance herangezogen. Der MCPI liegt jeweils für Tropen und Außertropen getrennt vor, da die Unterschiede der Modelle ansonsten nicht deutlich genug erfasst werden würden. Diese Entscheidung der Autoren unterstützt die hier getroffene, aber auf völlig andere Weise abgeleitete, Feststellung, dass die Dynamiken der Tropen und Außertropen der AOGCMs nicht gleichzusetzen sind. Der Abgleich mit diesen Studien wird durch die leicht unterschiedliche Ensemblezusammenstellung erschwert. Eine entsprechende Rangfolge für das komplette CMIP3 wird nur in dieser Arbeit erstellt. Die Berechnung von MCPI und  $I^2$  für den hier angestrebten Vergleich war aufgrund des enormen Datenbedarfs, den diese Kenngrößen haben, nicht möglich. In Tab. 7.2 wird daher jeweils die Rangkorrelation für die gemeinsame Schnittmenge der unterschiedlichen Quellen aufgezeigt. Die Berechnungen und folgenden Ausführungen beruhen also nur auf der Bewertung von 19 der 24 Modelle.

**Tab. 7.2:** Korrelation Modellbewertungen. Rangkorrelationen der Performance der CMIP3-Modelle. a: MCPI (Außertropen); b: MCPI (Tropen); c:  $I^2$  Performance Index ; d: Rangfolge nach  $f_\omega(i)$ . Signifikanz zum Niveau  $\alpha = 0.1$  wenn  $r_{sp} \geq 0.30$  (wenn  $r_{sp} > 0.38$  signifikant zu  $\alpha = 0.05$ )

	a	b	c	d
a	1	.70	.86	.46
b		1	.78	.40
c			1	.34
d				1

Die hier erstellte Bewertungsfunktion  $f_\omega(i)$  ist mit den anderen Bewertungsverfahren durchaus vereinbar und damit sicherlich nicht un schlüssig. Die höhere Korrelation zwischen den übrigen Rangfolgen dürfte darüber hinaus zumindest teilweise durch die Tatsache begünstigt werden, dass außer  $f_\omega(i)$  alle hier präsentierten Verfahren auf dem Abgleich von Datenfeldern beruhen. Somit sind diese Ansätze methodisch einander sehr viel ähnlicher als sie es dem hier entwickelten sind. In Anbetracht dieser Tatsache sollten die geringeren Werte von  $r_{sp}$  in der Spalte d nicht verwundern. Dennoch sind alle Rangkorrelationen in Tab. 7.2 nicht nur positiv, sondern auch signifikant zum Niveau  $\alpha = 0.1$ . Trotz einiger Abweichungen im Detail lässt sich also, zumindest für die Rangfolgen der 19 Modelle für die alle notwendigen Daten verfügbar waren, festhalten, dass sie eine gemeinsame Grundtendenz aufweisen. So wird auch gezeigt, dass das Verhalten des SOIs einen

brauchbaren Proxy für die Qualität eines AOGCMs darstellen kann. Diese Zeitreihe ist das Produkt so vielfältiger physikalischer Prozesse, dass sie ein AOGCM zu einem nicht trivialen Teil repräsentiert. Dies gilt insbesondere vor dem Hintergrund des hohen Qualitätsniveaus, dass für alle AOGCMs im Bezug auf die Dynamik der Außertropen festgestellt werden konnte.

Insofern kann der Vergleich der diversen Messungen und Betrachtungen zur Performance sowohl im Hinblick auf das ENSO-Phänomen als auch die gesamte Modellqualität als Bestätigung der Sinnhaftigkeit und Nützlichkeit der hier entwickelten Bewertungsfunktion angeführt werden. Für die hier und im Folgenden vorgenommenen Untersuchungen ist die tatsächliche Rangfolge maximal von untergeordneter Bedeutung, weswegen sie hier auch nicht gesondert aufgeführt wird. Angesichts der teilweise nur sehr geringen Unterschiede zwischen den Werten von  $f_{\omega}(i)$  sollte der Rangfolge zudem nicht zu viel Aussagekraft zugemessen werden. Auch in den beiden Referenzpublikationen finden sich nur sehr geringe Unterschiede in der Bewertung der Modellperformance. Das Gleiche gilt für WATTERSON (2008), obwohl sich letztere Untersuchung nur auf das australische Klima bezieht und somit zumindest theoretisch zu konkreteren Aussagen in der Lage sein sollte als bei Betrachtung der globalen Domäne. Wichtig ist dagegen der Einfluss den ein Modell auf das Verhalten des gesamten Ensembles ausübt, da der kombinierten Projektion des Multi-Modell Ensembles besondere Bedeutung für die klimatische Zukunft beigemessen wird. Dieser ist in Abb. 12.1 erkennbar.

## 8 Erwärmung und ENSO (modellspezifisch)

Der Fokus liegt im folgenden auf der Frage, ob die hier getroffene Bewertung einen originären Erklärungswert in Fragen des durch das CMIP3-Ensemble projizierten langfristigen Klimawandels besitzt. Hierfür wird der beobachtete Zusammenhang zwischen dem SOI und der globalen Mitteltemperatur genutzt. Für dieses Vorgehen spricht die Tatsache, dass diese Größe, als Maß für die inhärente Energie des Klimasystems, einen zentralen Stellenwert für sämtlichen Betrachtungen zum globalen Klimawandel einnimmt. Effekten die sich in  $T_E$  niederschlagen kann somit eine Wirksamkeit auf das Klimasystem als Ganzes unterstellt werden. So ist es das Ziel im Folgenden herauszufinden ob sich zwischen den Läufen der Teilensembles  $A$  und  $B$  Unterschiede in der Entwicklung von  $T_E$  detektieren lassen, und mit der – unterstellten – unterschiedlichen Qualität der Repräsentation des ENSO-Phänomens assoziiert werden können. In diesem Fall kann davon ausgegangen werden, dass die gewählte Bewertung der Modelle von einer gewissen Relevanz und ihre Anwendung auf andere Klimagrößen und Aspekte begründet ist.

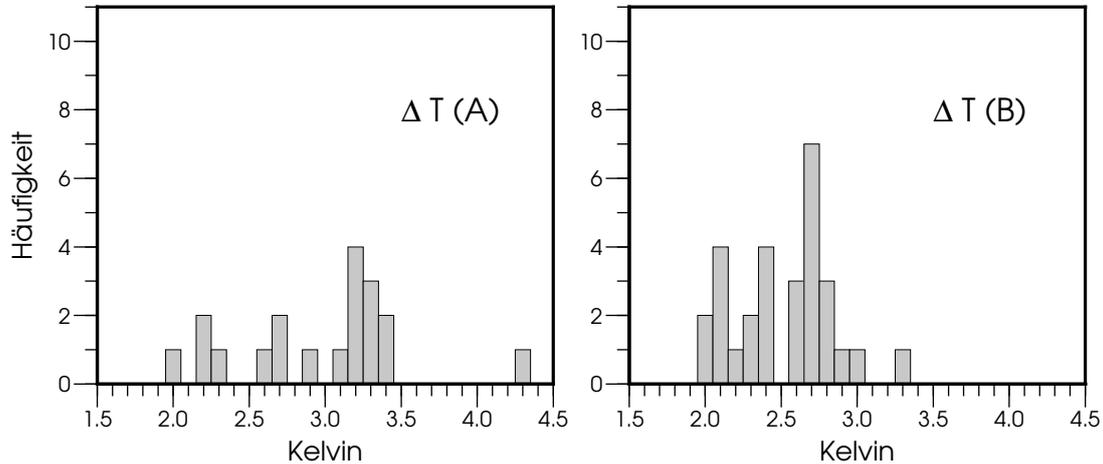
### 8.1 Globale Mitteltemperatur und SOI: Projektionen im CMIP3

Zu diesem Zweck werden die klimatologischen Veränderungen von  $T_E$  unter dem anthropogenen Forcing betrachtet. Als Maß hierfür wird für jeden Einzellauf jeweils die Verschiebung  $\Delta T(i)$  berechnet. Das ist die Differenz der gemittelten Globaltemperatur im Zeitraum 2080–2098 ( $P_2$ ) zur Normalisierungsperiode  $P_1$  in der zweiten Hälfte des 20. Jahrhunderts (1960–1999). Diese Größe ist eng mit der Klimasensitivität der Modelle verbunden (RAPER et al. 2002).

$$\Delta T(i) = \bar{T}_E(P_2, i) - \bar{T}_E(P_1, i) \quad (8.1)$$

Abbildung 8.1 zeigt die Verteilung von  $\Delta T(i)$  im Multi-Modell Ensemble, getrennt für Teilensembles  $A$  und  $B$  dargestellt. Die Histogramme von  $A$  und  $B$  unterscheiden sich deutlich. Die am häufigsten besetzte Klasse für  $A$  weist eine Erwärmung von etwa 3.2 K aus. Dagegen liegt der Modalwert von  $\Delta T$  im Teilensemble  $B$  nur um 2.7 K. Dieser Betrag von 0.5 K liegt nur knapp unter der seit der Industriali-

sierung beobachteten globalen Erwärmung, was die tatsächliche Bedeutung dieser Differenz veranschaulicht. Zudem fällt die Erwärmung bei einem Drittel der Läufe von  $A$  noch deutlicher aus, wohingegen sich in  $B$  lediglich ein einzelner Lauf mit einer vergleichbar intensiven Erwärmung findet.



**Abb. 8.1:** Erwärmung je Teilensemble. Histogramme für  $\Delta T(i)$ , separat für die beiden Teilensembles. Beide enthalten Modell Ensembles, dennoch ist  $B$  deutlich kompakter und homogener als  $A$ . Parameterschätzer siehe Tab. 8.1, Erläuterungen im Text

Neben dem Schwerpunkt ist auch die Struktur der Verteilung der Einzelläufe auf die Klassen unterschiedlich. Das Histogramm von  $B$  weist eine kompaktere Struktur auf als das von  $A$  (Abb. 8.1) und zeigt die geringere Streuung der  $\Delta T(i)$ s in  $B$  an. Sowohl die Mittelwerte als auch die Varianzen beider Teilensembles unterschieden sich signifikant zum Niveau  $\alpha = 0.05$ . Während die Ergebnisse des  $t$ -Tests sich über andere Unterschiede in der Modellphysik erklären lassen können, können die Unterschiede in den Varianzen durchaus als Auswirkung des ENSO-System interpretiert werden. Realistisch abgebildet, stellt es eine Variabilitätsquelle für  $T_E$  dar, was bedeutet, dass auch die Abschätzungen der anthropogenen Veränderungen von  $T_E$  über höhere Variabilität verfügen sollten. In Abb. 8.1 sticht der extrem hohe Wert von  $\Delta T$  für MIROC 3.2(hires) im Histogramm für  $A$  heraus. Bei diesem individuellen Lauf liegt die Erwärmung bei rund 4.3 K und somit etwa ein 1 K höher als bei sämtlichen Modellläufen des Teilensembles  $B$ .

Es könnte durchaus in Betracht gezogen werden, diesen Lauf als Ausreißer zu betrachten. In dem Fall sollte er bei der Abschätzung der Folgen des globalen Klimawandels eventuell ignoriert werden und auch nicht für weitere Analysen in dieser Arbeit herangezogen werden. Gegen dieses Vorgehen spricht zum einen die relativ gute Bewertung, die diesem Modell nach Kapitel 7.2 zukommt. Auch lässt

sich anhand der Modellvalidation keine Begründung erkennen, dass dieses Modell im Vergleich zu den Übrigen unrealistischere Ergebnisse erzielt (vgl. Kapitel 4). Entsprechend spricht die übliche Praxis bei der Verwendung der CMIP3-Modelle gegen dessen Ausschluss. Stattdessen wird im Hinblick auf die bestehenden Unsicherheiten, Probleme und Defizite der AOGCMs auch dieser Lauf als eine mögliche Realisation der Zukunft betrachtet.

Aus statistischer Sicht besteht jedoch die Gefahr, dass alleine MIROC 3.2(hires) in der folgenden Analyse signifikante Ergebnisse maskiert oder auch erzeugt. Dem Ziel dieser Untersuchung würden solche Effekte eindeutig zuwider laufen. Aus diesem Grund wurden alle weiteren Analysen in diesem Kapitel mit drei leicht unterschiedlichen Datengrundlagen durchgeführt und zwar

- mit dem kompletten SRES A1b,
- mit dem SRES A1b, reduziert um den Lauf MIROC 3.2(hires) und
- mit einem robustifizierten Multi-Modell Ensemble.

Das robustifizierte Ensemble ist um insgesamt sechs Läufe verringert. Es handelt sich jeweils um die drei Läufe mit den höchsten Beträgen bei den positiven bzw. negativen Abweichungen von der Zeitreihe der Multi-Modell Mittelwerte. Die wichtigsten Kenngrößen, auf die sich die folgenden Ausführungen stützen, finden sich für alle drei Versionen in Tabelle 8.1. Hier werden ansonsten, insbesondere in Abbildungen, die Ergebnisse für die Untersuchung des kompletten SRES A1b wiedergegeben. Obwohl sich die konkreten Werte in den drei Varianten der Analyse unterscheiden, gilt dies weder für die grundsätzlichen Aussagen noch für die Signifikanz der Ergebnisse. Die hier präsentierten Befunde können also als stabil betrachtet werden.

Nach Abb. 8.1 und Tab. 8.1 weisen die beiden Teilensembles bei der Verteilung von  $\Delta T$ , einer oder der zentralen Kenngröße des Klimawandels, deutliche Unterschiede auf. Im Folgenden wird untersucht ob, und gezeigt dass, der Rückgriff auf die Erstellte Bewertung durchaus einen Ansatzpunkt bietet diese Differenzen zu erklären. Klimatologisch bedeutet dies, dass eine Beziehung mit der Entwicklung des ENSO-Phänomens im jeweiligen Einzellauf und auf seine Repräsentation durch die unterschiedlichen GCMs naheliegt. Aufbauend auf diesen Erkenntnissen wird der Versuch unternommen, diese Differenzen zu reduzieren.

Analog zur Verschiebung der globalen Temperatur wird als  $\Delta I(i)$  die für den Index  $I(t, i)$  der Southern Oscillation der einzelnen Modellläufe berechnet.

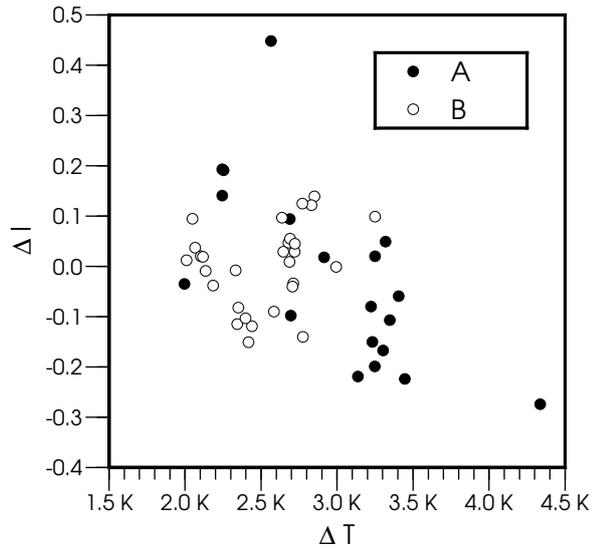
$$\Delta I(i) = \bar{I}(P_2, i) - \bar{I}(P_1, i) \quad (8.2)$$

Bei der Interpretation von  $\Delta I$  ist zu beachten, dass es sich um zeitliche Mittelwerte handelt. Daher kann ein negativer Wert  $\Delta I$  für einen Lauf tendenziell häufigere oder stärkere El Niño-Episoden zum Ende des 21. Jahrhunderts anzeigen. Ebenso

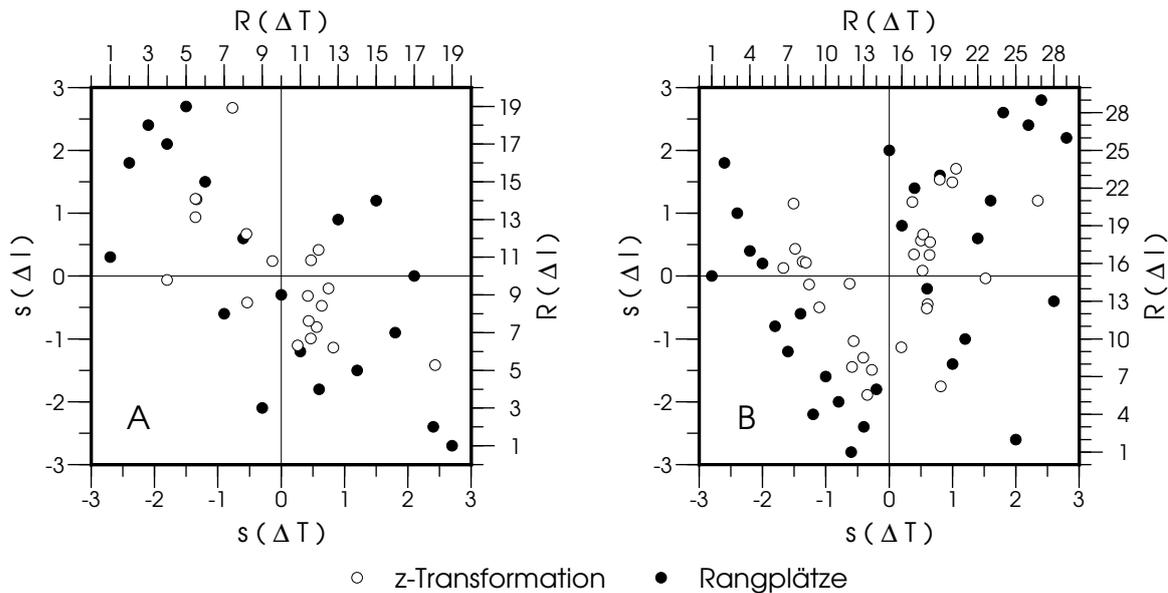
ist es aber auch möglich, dass eine negative Verschiebung im SOI auf seltenere oder auch auf weniger starke bzw. schwächer ausgeprägte La Niña-Phasen im Vergleich zur Referenzperiode hindeutet. Die globalen Auswirkungen und klimatologischen Folgen für das reale Klimasystem dieser beiden Szenarien sind klar zu trennen. Im Bezug auf  $T_E$  würden jedoch beide Formen der Änderung des ENSO-Systems zur einer Erhöhung führen.

In Abb. 8.2 ist für sämtliche Einzelläufe  $\Delta I(i)$  als Scatterplot gegen  $\Delta T(i)$  aufgetragen. Hier wird deutlich erkennbar, dass insbesondere im Teilensemble  $B$  für die Wertepaare einige Cluster identifiziert werden können. Obwohl auch in  $A$  mehrere Modelle als Ensembles vertreten sind, sind die Unterschiede zwischen den einzelnen Läufen deutlicher ausgeprägt. In  $A$  ist zudem die Spannweite der  $\Delta I(i)$  sichtbar größer als in  $B$ . Entsprechend finden sich hier stärkere Trends im SOI. Diese sind jedoch nahezu ausgeglichen verteilt und in dieser komprimierten Darstellung in vielen Läufen auch nicht deutlich ausgeprägter als in der Mehrzahl der Läufe von  $B$ . Auch in diesem Teilensemble finden sich sowohl positive als auch negative Trends im SOI. Die Korrelation dieser Größen bei Berücksichtigung des gesamten SRES A1b beträgt  $-0.41$  bzw.  $-0.30$  bei Verwendung der Rangkorrelation. Beide Maße sind statistisch signifikant und lassen sich derart interpretieren, dass die grundsätzliche Entwicklung von  $T_E$  zwar durch das Forcing bestimmt und somit in allen Läufen quasi deterministisch festgelegt ist, die letztendlich realisierte Erwärmung jedoch durch den Trend in der internen Variabilität modelliert wird.

Allerdings begründet sich diese physikalisch sinnvolle, also negative Korrelation von  $\Delta I$  und  $\Delta T$  quasi nur auf den Läufen in  $A$ , während im Teilensemble  $B$  scheinbar keine oder wenn dann umgekehrte Beziehung herrscht. Zur Verdeutlichung der Unterschiede zwischen  $A$  und  $B$  werden beide in Abb. 8.3 getrennt betrachtet. Für beide Teilensembles sind hier die Rangplatzpaare für  $\Delta T(i)$  und  $\Delta I(i)$  abgebildet. Weiterhin sind beide Größen als standardisierte Anomalien abgebildet. Diese  $z$ -Transformation wurde für jedes Teilensemble separat vorgenommen. Unabhängig von der gewählten Methode lässt sich für  $A$  eine ausgeprägte, statistisch signifikante



**Abb. 8.2:** SOI- und Temperaturentwicklung in CMIP3. Jeweils absolute Veränderung am Ende des 21. Jahrhunderts.



**Abb. 8.3:** Zusammenhang  $\Delta T$  und  $\Delta I$ . Jeweils Wertepaare der standardisierten Anomalien und Rangplätze für  $A$  und  $B$ . Es resultieren inhaltlich sinnvoll interpretierbare negative Korrelationskoeffizienten für  $A$  und positive für  $B$ .

te und physikalisch sinnvolle Beziehung zwischen der langfristigen Entwicklung des SOI und der der globalen Temperatur feststellen ( $r = -0.67$ ). Im Gegensatz hierzu fehlt eine solche Beziehung im Teilensemble  $B$ , in dem der betreffende Korrelationskoeffizient mit  $r = 0.26$  zwar knapp nicht signifikant, aber doch positiv und somit im Hinblick auf die Beobachtungen nicht sinnvoll interpretierbar ist (vgl. Tab. 8.1). Dagegen ist für  $B$  die, ebenfalls nicht schlüssige Beziehung der Rangplätze die  $r_{sp}$  beschreibt signifikant. Diese Ergebnisse geben einen deutlichen Hinweis darauf, dass die Beziehung zwischen SOI und  $T_E$  in beiden Teilensembles unterschiedlich ist. Allerdings kann an dieser Stelle nur das Vorhandensein im Fall von  $A$ , bzw. für das Teilensemble  $B$  das Fehlen eines sinnvollen systematischen Zusammenhangs aufgezeigt werden. Es ergibt sich so aber keine quantitative Abschätzung des Einflusses der Entwicklung des SOIs auf die globale Temperatur.

## 8.2 SOI-Signal: Detektion und Entfernung

So sich die Unterschiede der Histogramme in Abb. 8.1 durch das ENSO-Phänomen erklären lassen, sollte sich die Temperaturentwicklung in den beiden Teilensembles angleichen, wenn der Einfluss von ENSO aus der globalen Temperatur entfernt wird.

Grundlage für diese Überlegung ist die grundsätzliche Beziehung von ENSO und globaler Temperatur (vgl. Kapitel 7.1). Mit ENSO wird der einzige interne Faktor des Klimasystems mit nennenswertem Einfluss auf  $T_E$  entfernt, so dass nun das vorgeschriebene Forcing aus dem Strahlungsantrieb für alle Modellläufe die alleinige Einflussgröße ist. Schon hieraus sollte sich eine Annäherung der einzelnen Modelle einstellen, wenn die Unterschiede in den konkreten  $\Delta T$  nicht allein auf individuellen Modelldefiziten beruhen, sondern auch ein klimatologisch erklärbarer Einfluss besteht. Abb. 8.3 und Tab. 8.1 legen nahe, dass dies zumindest bei  $A$  gegeben ist.

Um den Einfluss des SOI aus der  $T_E$  zu entfernen wird von einem additiven Zeitreihenmodell (siehe Kapitel 3.6.1) ausgegangen.

$$T_E(t) = \hat{T}_E(t) + T_I(t) \quad (8.3)$$

Mit  $T_I(t)$  wird der Effekt des SOI auf die globale Temperatur zu einem Zeitpunkt  $t$  bezeichnet, im Residuum  $\hat{T}_E(t)$  werden die Effekte sämtlicher anderer Einflussgrößen inklusive der stochastischen Komponente zusammengefasst. Für jeden Lauf  $i$  ergibt sich Gleichung 8.3 direkt zu

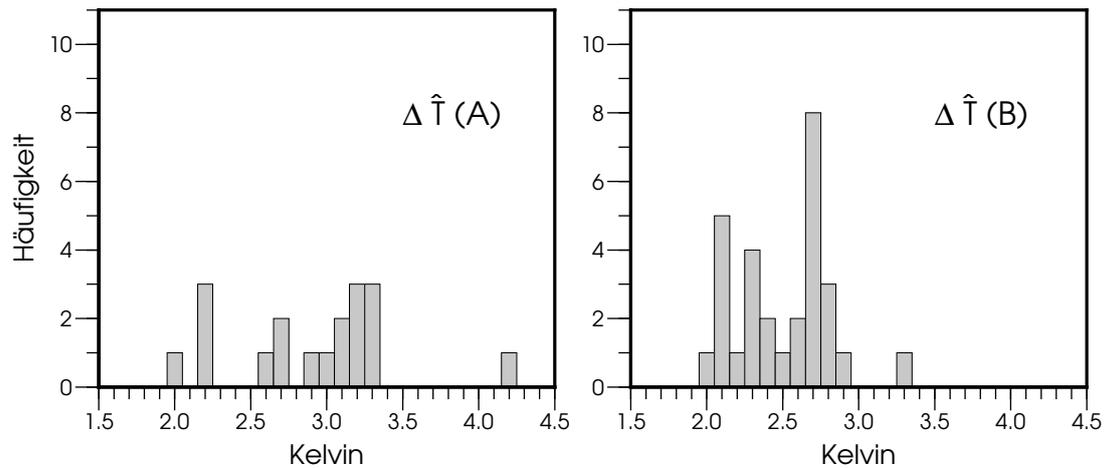
$$T_E(t, i) = \hat{T}_E(t, i) + T_I(t, i) \quad (8.4)$$

Der konkrete Wert  $T_I(t, i)$  ist im doppelten Sinn modellabhängig. Zum einen ist es angesichts der in Abb. 7.8 gezeigten Ergebnisse vernünftig, ein spezifisches lag  $h_i$  für jeden Lauf anzunehmen, zudem der Einfluss des SOI auf  $T_E$  am stärksten ist, und das lag nicht einfach entsprechend der Beobachtungsdaten festzusetzen. Es wird jeweils das lag  $h_i$  verwendet, für das der Betrag der Kreuzkorrelationsfunktion maximal wird. Bei führenden SOI bedeutet dies, dass jeweils gelten muss  $h_i \leq 0$ , um den beobachteten Zusammenhang mit der folgenden globalen Temperatur qualitativ korrekt abzubilden (vgl. Abb. 7.8).

Ebenfalls modellabhängig ist das Ausmaß der Sensitivität der globalen Temperatur gegenüber dem SOI. Für jeden Lauf wird dieser Faktor in Form eines Anregungskoeffizienten  $k_i$  daher separat durch lineare Regression von  $T_E(t, i)$  auf  $I(t, i)$  bestimmt. Problematisch sind bei diesem Vorgehen Trends, die massiv bei sämtlichen Läufen im Fall von  $T_E(t, i)$  auftreten, teilweise aber auch in Indexreihen  $I(t, i)$  des SOIs vorhanden sind. Daher wird für sämtliche Zeitreihen die Trendkomponente entfernt, bevor  $k_i$  bestimmt wird. Entsprechend dem Vorgehen in Kapitel 7.1 wird diese als kubischer Trend über Polynomregression geschätzt.

Mit den resultierenden Residuen  $\check{T}_E(t, i)$  und  $\check{I}(t, i)$  ergibt sich für jeden Lauf der lineare Regressionsansatz, der auf Basis der, im Fall von  $I(t, i)$  um  $h$  verschobene, Jahresmittel erstellt wird:

$$\check{T}_E(t, i) = \beta_0 + \beta_1 \cdot \check{I}(t + h, i) + \epsilon(t) \quad (8.5)$$



**Abb. 8.4:** Erwärmung je Teilensemble ohne SOI. Wie Abb. 8.1 jedoch ENSO-Einfluss über additives Zeitreihenmodell entfernt, damit  $\Delta\hat{T}(i)$ . Effekte v.a. in  $A, B$  ist gegen diese Maßnahme weitgehend stabil.

Der Anregungskoeffizient  $k_i$  wird mit  $\beta_1$  gleichgesetzt. Hiermit wird dann der Temperatureffekt des SOI pro Modelllauf abgeschätzt. Dieser kommt im Bezug auf die Zielgröße allerdings nur bei einem tatsächlichen Trend im SOI zum Ende des 21. Jahrhunderts hin zum tragen.

$$T_I(t, i) = k_i \cdot I(t + h, i) \quad (8.6)$$

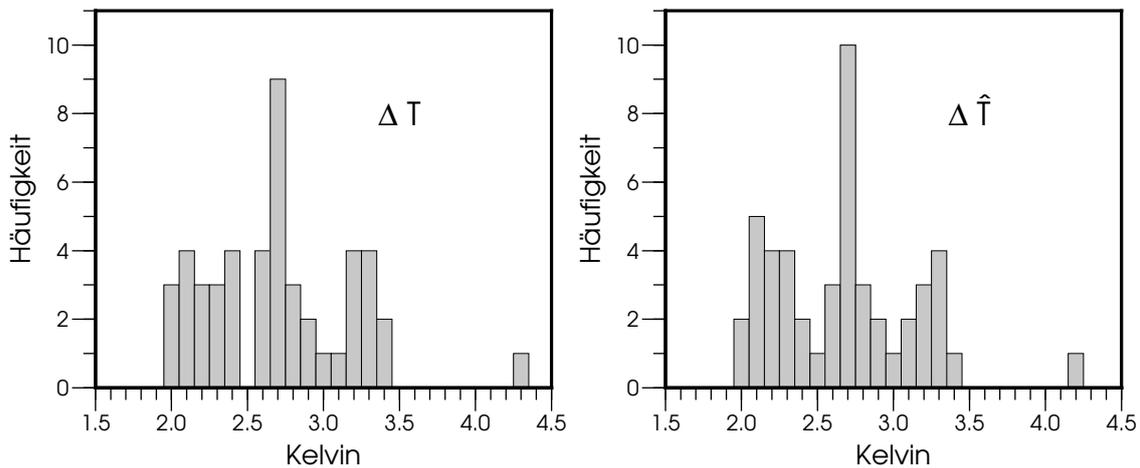
Schließlich kann in Gleichung 8.4  $\hat{T}_E(t, i)$  und somit eine ENSO-bereinigte Zeitreihe der Entwicklung der globalen Mitteltemperatur bestimmt werden:

$$\hat{T}_E(t, i) = T_E(t, i) - T_I(t, i) \quad (8.7)$$

Für die so gewonnenen Residuumszeitreihen wird, entsprechend Gleichung 8.1, die Verschiebung  $\Delta\hat{T}(i)$  bestimmt. Hierbei handelt es sich damit um die langfristige Veränderung der globalen Mitteltemperatur ohne den Einfluss von El Niño und La Niña. Die Verteilung dieser Größe in den Teilensembles ist in Abb. 8.4 dargestellt. Die Effekte dieser Maßnahme für  $B$  sind als gering zu bezeichnen, obwohl sich in einigen Fällen  $\Delta T(i)$  durchaus so stark ändert, dass Läufe in andere Histogrammklassen wechseln. Diese individuellen Anzeichen für die Auswirkungen, die ENSO auf  $T_E$ , hat verändern die statistischen Eckdaten der globalen Erwärmung in  $B$  jedoch nicht. Sowohl der Schwerpunkt als auch die Spannweite der Verschiebungen bleiben unverändert und nur wenige Klassen weisen eine andere Häufigkeit auf als in Abb. 8.1. Obwohl mehrere Läufe dieses Teilensembles durchaus relevante Trends für den SOI aufweisen, muss dessen Einfluss auf die globale Temperatur somit als

sehr gering angesehen werden.

Dagegen bewirkt die ENSO-Bereinigung bei  $A$  eine deutlichere Veränderung im entsprechenden Histogramm (Abb. 8.4). Der Schwerpunkt der empirischen Verteilung liegt zwar weiterhin um  $\Delta\hat{T} \approx 3.2$  K, jedoch lässt sich für diese Größe eine stärkere Tendenz hin zu einer geringeren Erwärmung feststellen. Dagegen ist die Anzahl der Läufe mit größeren Beträgen von  $\Delta T$  höher. Weiterhin ist die Spannweite der Erwärmung reduziert. Dies resultiert aus der geringfügigen Abmilderung des sehr hohen Wertes von MIROC 3.2(hires). Die Entfernung des ENSO-Signals aus der Temperatur ist allerdings nicht immer mit einer Minderung der globalen Erwärmung verbunden. So ist der kleinste Wert von  $\Delta\hat{T}$  in diesem Teilensemble etwas höherer Wert als der von  $\Delta T$ , was durch die Entfernung der Auswirkungen von ENSO bei dominierendem La Niña-Einfluss erklärbar ist. Für  $A$  ergibt sich so ein etwas kompakteres Histogramm als in Abb. 8.1. Jedoch ist die Abnahme der Streuung zwischen den Läufen des Ensembles auch in diesem Fall nicht statistisch signifikant.



**Abb. 8.5:**  $\Delta T_E$  im SRES A1b. Histogramme für das gesamte Ensemble, vor der Entfernung des Einflusses von ENSO und danach.

Es kann also festgehalten werden, dass das ENSO-Phänomen, wenn es vom AOGCM realistisch abgebildet wird und über einen entsprechenden Trend verfügt, durchaus auch auf klimatologischen Zeitskalen für die quantitative Ausprägung der fundamentalen Größe  $T_E$  von Bedeutung sein kann. Auch im SRES A1b Ensemble als Ganzes lässt sich diese durchaus bedeutungsvolle Feststellung erkennen. Werden die Histogramme für  $\Delta T$  und  $\Delta\hat{T}$  gegenübergestellt, zeigt sich für das ENSO-bereinigte Histogramm eine etwas kompaktere Form. Das Zentrum des Histogramm erscheint hierbei gestärkt, während die sowohl extrem geringen als auch die extrem hohen Erwärmungen etwas seltener und schwächer ausgeprägt

**Tab. 8.1:** Globale Erwärmung und ENSO-Einfluss im SRES A1b.  $r$ : Korrelation zwischen  $\Delta T$  und  $\Delta I$ ;  $r_{sp}$ : Rangkorrelation;  $\bar{x}/s^2$ : je Mittelwerte/Varianzen für  $\Delta T$  bzw.  $\Delta \hat{T}$  (vor bzw. nach der Entfernung des ENSO-Signals. Jeweils für alle Läufe (ges.) und die Teilensembles separat,  $n$ : jeweilige Anzahl der Läufe. Drei Versionen des CMIP3: komplett, reduziert um MIROC 3.2(hires) und robustifiziert durch Entfernung der drei Läufe mit der höchsten und geringsten Erwärmung.

CMIP3	komplett			reduziert			robustifiziert		
	ges.	A	B	ges.	A	B	ges.	A	B
$r$	-.41	-.67	.26	-.31	-.63	.26	-.25	-.71	.38
$r_{sp}$	-.30	-.67	.29	-.25	-.61	.29	-.19	-.64	.38
$\bar{x}_{\Delta T}$	2.71	2.99	2.52	2.67	2.92	2.52	2.69	2.91	2.56
$s^2_{\Delta T}$	.23	.31	.10	.18	.22	.10	.14	.17	.08
$\bar{x}_{\Delta \hat{T}}$	2.68	2.94	2.51	2.64	2.87	2.51	2.66	2.87	2.53
$s^2_{\Delta \hat{T}}$	.22	.29	.10	.17	.20	.10	.13	.15	.09
$n$	48	19	29	47	18	29	42	15	27

erscheinen. Diese Unterschiede bedeuten, dass in diesem Fall die Entfernung der simulierten natürlichen Variabilität das Intervall der Vorhersagen des Ensembles verringert.

Entsprechend dem qualitativen Eindruck, den Abb. 8.5 vermittelt, sind die numerischen Unterschiede in den statistischen Kenngrößen nur gering. So ist kein signifikanter Unterschied der Mittelwerte  $\bar{x}_{\Delta \hat{T}}$  und  $\bar{x}_{\Delta T}$  feststellbar. Auch führt das Verfahren nicht zur signifikanten oder relevanten Reduktion der Variabilität. Allerdings ist die, geringe, Reduktion der Varianz der Verschiebungen in  $A$  ausgeprägter (vgl. Tab: 8.1). Dennoch, bei Berücksichtigung der individuellen Entwicklung der Einzelläufe, zeigt die hier vorgenommene Untersuchung, dass die ungewisse Entwicklung des SOIs, und damit von ENSO, durchaus einen nicht zu vernachlässigenden Unsicherheitsfaktor im Hinblick auf die Globale Erwärmung darstellt. Auch im rezenten, seit dem Jahrtausendwechsel verringerten Trend der globalen Erwärmung, lässt sich der Einfluss von ENSO feststellen (KOSAKA und XIE 2013). Für das Klimasystem scheint mit Blick auf die hier präsentierten Ergebnisse ein, intensivierender oder auch mildernder, Effekt von einigen Zehntel K auf die Entwicklung der globalen Temperatur auf klimatologischen Zeitskalen durchaus nicht unrealistisch. Diese Aussage ist vor dem Hintergrund der eventuell hieran gekoppelten Auswirkungen auf Klimasystem und Erdbevölkerung zu beurteilen (IPCC 2007b, LENTON et al. 2008, HANSEN et al. 2008, SMITH et al. 2009).

Angesichts der dargelegten Ergebnisse kann im Bezug auf die Bewertungs- und Gewichtungsfunktionen festgehalten werden, dass sowohl das gewählte Kriterium als auch seine quantitative Erfassung sinnvoll und nicht zufällig erscheint. Dieses

Fazit legen etwa die aufgezeigten, konsequent interpretierbaren Unterschieden zwischen den Teilensembles im Bezug auf die in Abb. 8.2 erkennbaren Kovarianzen nahe. Die bei Betrachtung der Teilensembles insgesamt geringen Reaktionen auf die vorgenommene Entfernung des SOI-Einflusses aus der Entwicklung von  $T_E$  mag darauf zurückzuführen sein, dass die eingesetzte Methode hierfür nicht sensitiv genug ist. Es kann aber festgehalten werden, dass die Anzahl der Läufe die ein im Hinblick auf die zentrale Größe  $T_E$  relevantes ENSO-Phänomen aufweisen im Teilensemble  $A$  deutlich höher ist als in  $B$ . In diesem Teilensemble, in dem die spektralen Charakteristika des SOI weniger gut mit den Beobachtungen übereinstimmen, sind auch die entsprechenden Effekte auf  $T_E$  insgesamt geringer. Im Folgenden wird daher mithilfe elaborierterer statistischer Verfahren untersucht, ob sich die Entwicklung des Klimas grundsätzlich oder auch nur in geringer quantitativer, aber systematischer Form unter anthropogenem Forcing unterscheidet, wenn die qualitativen Unterschiede im modelleigenen ENSO-System berücksichtigt werden.

## 9 Klimaänderungen: Signalanalyse der Zeitreihen

Die folgende Untersuchung konzentriert sich auf die Entwicklung des SOI, der Indices global relevanter Modi der außertropischen Zirkulation sowie der globalen Mitteltemperatur unter dem anthropogenen Forcing. Hierbei werden zwei unterschiedliche Ziele verfolgt. Zunächst sollen Klimaänderungssignale, so sie im CMIP3 unter den Randbedingungen vom SRES A1b modellübergreifend detektierbar, sind quantifiziert werden. Hierzu werden vielfältige Aspekte des Klimawandels in CMIP3 auf systematische und ausführliche Weise analysiert. Der besondere Nutzen liegt sicherlich in der direkten Vergleichbarkeit der erfassten und abgesicherten Klimaänderungen, die sich durch die Kontinuität der Methodik ergibt. Zudem stellt die verwendete Datenbasis die maximal mögliche dar, die aus CMIP3 gewonnen werden kann.

Der zweite Aspekt der hier behandelt werden soll ist die für diese Arbeit wichtige Frage, ob sich aus der Gruppierung in die Teilensembles *A* und *B* Effekte ergeben, die sich in der Entwicklung dieser Klimagrößen so deutlich und systematisch niederschlagen, dass ihre Detektion gegen den Einfluss des Strahlungsantriebs sowie den der internen Variabilität der Einzelläufe möglich ist. Dies entspricht der Überprüfung, ob die qualitativen Unterschiede der Modelle bei der Abbildung des ENSO-Systems einen Beitrag für die Erklärung von unterschiedlichen Ausprägungen der Klimaänderungssignale ermöglichen, die sich in CMIP3 unter gleichen Randbedingungen detektieren lassen. Da es sich jeweils um Multi-Modell Ensembles handelt, sind systematische Unterschiede zwischen den beiden Gruppen als Effekt des Kriteriums zu bewerten.

Die Interpretation solcher Ergebnisse kann auf zweierlei Arten erfolgen, beide sowohl berechtigt als auch folgenreich. In jedem Fall wird eine Unsicherheitsquelle im CMIP3 aufgezeigt und konkretisiert, die potentiell die gesamte Entwicklung im Ensemble beeinflussen kann und die die Zuverlässigkeit der Ensemble Projektionen herabsetzt. So ist es zunächst möglich, dass signifikante Unterschiede zwischen den Teilensembles direkt auf ENSO zurückzuführen sind. In diesem Fall ist die Umsetzung des anthropogenen Forcings durch das Klimasystem abhängig von einem realistischen ENSO-Mechanismus. Hier stellt also die Klimamode ENSO einen wichtigen Akteur in der zukünftigen Entwicklung der untersuchten Klimagröße dar. Für alle hier untersuchten Größen sind zumindest Fernwirkungen mit ENSO beschrieben, so dass dies eine mögliche Interpretation darstellt. Mit den verfügbaren

Daten im Endeffekt, vor allem für die große Anzahl an Klimasimulationen, nicht von dieser Interpretation zu unterscheiden ist eine alternative Betrachtungsweise, die weniger auf die klimatologische Bedeutung des SOI abzielt, als auf die Qualität der Modelle. Aus dieser Perspektive sind die Modelle, die im Teilensemble *A* zusammengefasst sind, insgesamt besser und somit vertrauenswürdiger. Signifikante Unterschiede zwischen den Modellen sind dann Unsicherheiten in den Multi-Modell Projektionen, die durch die Entfernung oder zumindest Reduktion des Einflusses der als schlechter erkannten Modelle beseitigt werden können.

Hier wird eine in diesem Sinne neutrale Position eingenommen, indem die Analysen zu diesem Themenkomplex zunächst als Sensitivitätsstudie betrachtet werden (vgl. KATZ 2002). Damit werden zwei Multi-Modell Ensembles betrachtet, deren systematisches Unterscheidungsmerkmal es ist, dass eines der beiden, *A*, bei der Simulation von ENSO realistischere Ergebnisse erzeugt als das andere. Wo sich eine weitergehende Interpretation anbietet, etwa wegen bekannter Beziehungen zwischen ENSO und einer anderen Größe, wird diese an der betreffenden Stelle gegeben.

## 9.1 Methodik

Eine Vielzahl von Studien beschäftigt sich mit der Entwicklung der hier betrachteten fundamentalen Zeitreihen in den CMIP3-Modellen. Obwohl sich diese Studien durchaus unterscheiden – etwa in der Indexdefinition, der Anzahl der betrachteten Läufe, den Methoden der Erfassung des anthropogenen Einflusses — lässt sich festhalten, dass für einige Zeitreihen die Ergebnisse im Hinblick auf diesen Faktor zumindest qualitativ bereits bekannt und im Rahmen der Möglichkeiten von Modellstudien abgesichert sind. Beispielsweise sind die Ergebnisse zur zukünftigen Entwicklung der globalen Temperatur aber auch des SAMs unter den Bedingungen des SRES A1b eindeutig (z. B. MILLER et al. 2006, PAETH und POLLINGER 2010, SIMPKINS und KARPECHKO 2012). Andere Zirkulationsmuster, wie SCA oder AMR, sind dagegen weniger gut untersucht und insbesondere für den SOI oder andere ENSO-Indexreihen ergaben sich keine eindeutigen Aussagen (COLLINS et al. 2010). Um belastbare Aussagen über einen eventuellen Einfluss des ENSO-Phänomens treffen zu können muss außerdem für alle Zeitreihen der anthropogene Anteil an der Entwicklung festgestellt werden. Beides lässt sich über eine spektrale Varianzanalyse bewerkstelligen. Dieses Verfahren kann als stabile Form der Detektion von systematischen Signalen in Multi-Modell Ensembles betrachtet werden.

In dieser Arbeit wird dieses Verfahren erstmals für eine so große Anzahl von Klimamoden sowie die Zeitreihe der globalen Temperatur auf der Basis sämtlicher verfügbarer Läufe des SRES A1b durchgeführt. Ermöglicht wird dies durch die wegen der Fragestellung, wohingegen üblicherweise nur Modelle mit mehreren

Realisationen berücksichtigt werden können. Aus diesem Grund zeigen die Abbildungen 9.1 und 9.2 auch eine äußerst umfangreiche und methodisch konsistente Analyse und Detektion von Klimaänderungen im CMIP3. In Tabelle 9.1 sind einige aussagekräftige Ergebnisse zusammengestellt. Alle Einzelanalysen beginnen mit dem Jahr 1970 und enden 2098.

Derjenige Anteil der Variabilität innerhalb des Multi-Modell Ensembles, der sich durch gemeinsame Veränderung als Reaktion auf steigende Treibhausgaskonzentrationen äußert, wird als anthropogenes Signal interpretiert. Diese Größe untersucht also den gemeinsamen Trend aller Zeitreihen. Zur Bestimmung der Richtung des Zusammenhangs kann jeweils der für diesen Zeitraum über lineare Regression bestimmte Trend des Multi-Modell Mittelwerts nach Gleichung 1.1 herangezogen werden. Hohe Werte bei dieser Größe bedeuten, dass für die betrachtete Größe ein modellübergreifender Konsens im Hinblick auf ihre zukünftige Entwicklung besteht.

Solche Ergebnisse zeigen somit konkrete Aspekte des Klimawandels auf, auf die sich die Bewohner betroffener Gebiete einstellen müssen. Somit ergibt sich auf langen Zeitskalen ein an das Forcing gekoppeltes Prognosepotential. Sicherlich liegt auf diesem Aspekt das Hauptinteresse bei der Auswertung von Klimasimulationen. Daneben, und insbesondere unter der hier vorgenommenen Einteilung der Modelle, sind allerdings auch die systematischen Unterschiede zwischen den Gruppen von Bedeutung. Diese äußern sich in der ANOVA als Blockeffekt. Im dargestellten Experiment repräsentieren Varianzanteile, die durch diese Kategorie erklärt werden, die Auswirkung der Gruppenbildung und somit der Gewichtung. Eine Fehlinterpretation dieser Größe ist zwar prinzipiell nicht auszuschließen, es handelt sich aber jeweils um Läufe aus verschiedenen Klimamodellen, was diese Gefahr sicher reduziert. Es lässt sich jedoch nicht entscheiden, ob diese Effekte tatsächlich auf das realistischere ENSO-Phänomen zurückgeführt werden können oder aber aus einer insgesamt anderen, der bisherigen Argumentation folgend besseren, Modellphysik resultieren.

Diese beiden Haupteffekte der Varianzanalyse können nicht von vornherein als statistisch unabhängig betrachtet werden. Tatsächlich erscheint es durchaus nicht unrealistisch, dass die Unterschiede in der Modellphysik unter den wechselnden Randbedingungen, die aus dem Einfluss des massiven Forcings des SRES A1b resultieren, nicht konstant sind. Diese Interaktion generiert Varianz, die sich in Unterschieden des Verhaltens der Teilensemble äußert. Hier können etwa unterschiedliche Vorzeichen in den Trends und somit qualitative Unterschiede detektiert werden. Um die beiden Gruppen zu vergleichen ist diese Größe daher von besonderem Interesse.

**Tab. 9.1:** Klimatische Kenngrößen in CMIP3: Trends und Varianzzerlegung. Regressionskoeffizient für linearen Trend ( $\beta$  in  $s/\Delta t$  für Indices und  $K/a$  für  $T_E$ ). Die erklärten Varianzanteile sind signifikant zum 5 %-Niveau und gelten für mit 11-jährig gleitenden Mittel geglättete Zeitreihen. Zeitraum 1970–2089.

Typ	Periode	$10^2 \cdot \beta$	Strahlungsantrieb	Gruppierung	Interaktion
SOI	djf	.05	-	10.7	3.9
	mam	.28	2.5	.2	-
	jja	-.09	-	-	-
	son	-.05	-	1.5	-
$T_E$	Jahr	2.55	91.4	.3	.5
EIMR	jjas	1.24	32.3	6.4	3.5
AMR	jjas	.68	12.1	.6	-
NAM	djf	.47	17.1	.2	-
	mam	.35	9.3	.1	-
	jja	.55	20.8	1.1	-
	son	.86	39.7	.3	-
SAM	jja	.86	31.1	.7	-
	son	.52	15.9	.6	-
	djf	.73	27.3	1.2	-
	mam	1.04	46.4	.5	-
NPI	djf	-.13	.2	-	-
	mam	.19	.8	-	-
	jja	.31	4.2	3.8	.8
	son	.28	3.2	-	-
SCA	djf	.45	13.0	-	-
	mam	.55	19.7	-	-
	jja	1.04	34.6	1.3	-
	son	.70	28.2	-	-

## 9.2 Ergebnisse

Hier werden die spezifischen Effekte des Strahlungsantriebs und der Einfluss des ENSO-Systems auf die Entwicklung der Klimagrößen aufgezeigt und diskutiert. Um eine gewisse Übersichtlichkeit zu wahren, werden die Ergebnisse für die einzelnen Größen separat diskutiert.

### 9.2.1 SOI

Die transiente Entwicklung des SOIs weist im Multi-Modell Ensemble kaum modellübergreifende Aspekte auf. Schon die langfristigen Trends unterscheiden sich quantitativ, aber auch schon qualitativ, derart, dass für das Ensembledittel kaum konkrete Aussagen zu treffen sind. Große Unterschiede zwischen den Einzelläufen kennzeichnen hierbei insbesondere das Teilensemble *A*, welches im Vergleich zu *B* die deutlich höhere Variabilität aufweist. Die Entwicklung in letzterem ist dagegen zwischen den Einzelläufen weniger stark verschieden, dafür weisen die Indexreihen selbst kaum Trends auf.

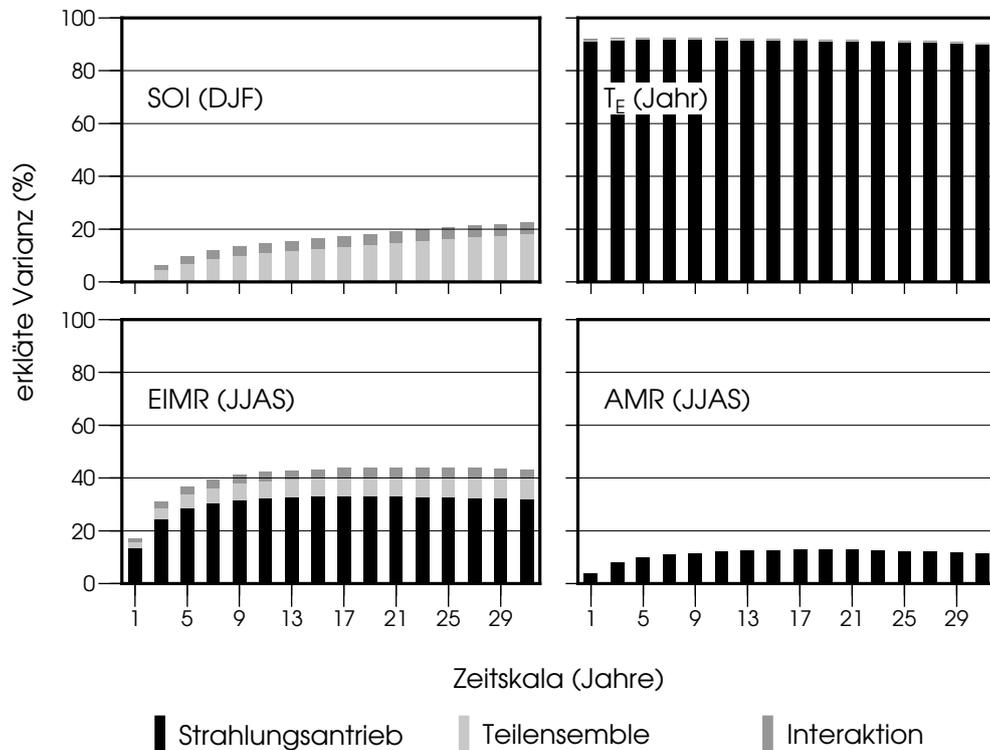
Daher findet sich lediglich für die Monate März, April und Mai ein nennenswerter Trend, und zwar hin zu positiven Indexwerten im Ensembledittel. Die Spektrale ANOVA kann allerdings auch für diese Jahreszeit nur einen sehr geringen Anteil, je nach betrachteter Zeitskala bis maximal knapp 3 %, der Variabilität durch die externen Effekte erklären. Signifikant ist hierbei nur der Strahlungsantrieb. Die Gruppenbildung hat dagegen keine systematische Auswirkungen.

Den übrigen Jahreszeiten fehlen solche Trends, dementsprechend erklärt in diesen Fällen der Strahlungsantrieb auch keinen signifikanten Anteil an der Gesamtvarianz. Für den Nordsommer ist die Indexentwicklung derart durch die hohe Variabilität der individuellen Zeitreihen geprägt, dass keinerlei systematische Varianz detektiert werden kann. Für die Monate September bis November findet sich ein geringer Einfluss der Gruppierung. Die hierdurch erklärten Varianzanteile sind zwar im Vergleich zu der internen Variabilität mit etwa 1 %–2 % nur als marginal zu bezeichnen, überschreiten allerdings bei dekadischer oder längerer Glättung das Signifikanzniveau. Die geringen Werte in diesem Blockeffekt sind im Fall des SOI auf das oben beschriebene, unterschiedliche Verhalten der Einzelläufe der Teilensembles zurückzuführen. Für den borealen Winter, der Jahreszeit in der der SOI seine stärkste Klimawirksamkeit aufweist und auch die typischerweise mit ENSO assoziierten Telekonnektionen am ausgeprägtesten sind (z. B. LANGENBRUNNER und NEELIN 2013), weist das Ensembledittel keinen Trend auf. Dafür erklärt die Gruppierung hier signifikante Anteile der Varianz. Das ist auch ein Beleg für die Sinnhaftigkeit des Wichtungskriteriums. Obwohl Monatswerte des SOI hierfür herangezogen wurden ist der Effekt in der Jahreszeit mit der stärksten ENSO-Dynamik deutlich, in den Monaten für die die Kopplung von Ozean und Atmosphäre weniger dominant in der Region ist dagegen kaum auszumachen.

Für die entscheidende Jahreszeit lassen sich also systematische Unterschiede zwischen den Teilensembles identifizieren. Dies gilt für alle in Abb. 9.1 dargestellten Zeitskalen. Die relative Bedeutung des Blockeffekts erhöht sich hierbei von knapp 2 % bei der Analyse der Originalzeitreihen auf deutlich über 10 %, wenn diese mit einem 11-jährigen Tiefpassfilter geglättet wurden. Bis zur maximalen Filterlänge von 31 Jahren lässt sich hierbei keine Konvergenz feststellen. Noch längere Filterlängen, bis zu 41 Jahren, ergeben sogar Werte bis zu 25 % der ver-

bliebenen Gesamtvarianz. Noch stärkere Glättung erhöht diesen Anteil nicht mehr, die Reduktion der Freiheitsgrade verhindert zudem signifikante Ergebnisse.

Obwohl das Forcing für sich keinen signifikanten Einfluss auf die Variabilität im Ensemble hat, identifiziert die Spektrale ANOVA signifikante Varianzanteile die aus der Kombination der beiden Haupteffekte hervorgehen. Hierbei handelt es sich um ein wichtiges Analyseergebnis. Anthropogenes Forcing löst bei den Mittelzeitreihen der beiden Teilensembles unterschiedliche Reaktionen aus. Das Fehlen des entsprechenden Haupteffekts deutet an, dass sich die transiente Entwicklung von  $A$  und  $B$  gegenseitig neutralisiert. Während in  $A$  eher eine Tendenz in Richtung El Niño besteht, verfügt die gemittelte Zeitreihe der Einzelläufe von  $B$  über einen leichten positiven Trend. Dieses Ergebnis wird weiter unten (Kapitel 10) vertieft.



**Abb. 9.1:** Ausgewählte Ergebnisse der Spektralen ANOVA für SOI,  $T_E$ , EIMR und AMR. Erläuterungen siehe Text. Zeitskala (Jahre) entsprechen der vor der ANOVA vorgenommenen Tiefpassfilterung. Zeitskala (1) steht für die ungefilterte Zeitreihe. Alle dargestellten Varianzanteile sind signifikant zu  $\alpha = 0.05$ . Die Summenbalken werden jeweils durch die nicht dargestellte interne Variabilität zum Wert 100 % ergänzt

### 9.2.2 Globale Mitteltemperatur

Für die grundsätzliche Entwicklung der globalen Temperatur im Ensemble ist wenig Überraschendes bei der Spektralen ANOVA zu erwarten. Der Strahlungsantrieb dominiert in diesem Fall die Zeitreihen aller individuellen Läufe. Entsprechend lassen sich, unabhängig von der Glättung durch verschiedene Tiefpassfilter, jeweils über 90 % der Variabilität erklären. Dieses eindeutige Ergebnis ist auch auf die besonderen Eigenschaften dieser Zeitreihe zurückzuführen.

So weist  $T_E$  nur eine äußerst geringe natürliche interne Variabilität auf. In diesem Punkt sind somit die Zeitreihen der anderen Klimagrößen, die hier betrachtet werden, nicht mit der Globaltemperatur vergleichbar. Zudem werden natürliche Quellen der externen Variabilität, angesichts des relativ langen Zeitraums der hier betrachtet wird, gegenüber dem kontinuierlich wachsenden und ständig präsenten anthropogenen Forcing im Ensemble marginalisiert. Dennoch können natürlich in der realen Zukunft solche Faktoren, insbesondere explosiver tropischer Vulkanismus, weiterhin die globale Mitteltemperatur kurzfristig beeinflussen. Angesichts der enormen Bedeutung des Strahlungsantriebs für  $T_E$  ist hier durchaus ein reales Prognosepotential gegeben. Dieses ist allerdings immer an die getroffenen Annahmen in Bezug auf die Emission im Verlauf des 21. Jahrhunderts gekoppelt. Die Möglichkeit der direkten Übertragung der Modellergebnisse auf die Realität ist hierdurch eingeschränkt, es bestätigt allerdings auch die Potentiale, über die aktive Klimaschutzpolitik zur Mitigation dramatischer Klimafolgen verfügt (IPCC 2007b).

Alle übrigen Effekte, inklusive der internen Variabilität, treten neben dem Strahlungsantrieb in den Hintergrund. Nichtsdestotrotz finden sich auf allen Zeitskalen sowohl Varianzanteile, die aus systematischen Unterschieden zwischen den Gruppen resultieren als auch solche, die als Interaktion der beiden Haupteffekte klassifiziert werden. Auch für diese beiden Effekte ist die relative erklärte Varianz, wenn auch auf einem geringen Niveau, nahezu unabhängig von der Glättung der Zeitreihen. Die hiermit verbundene Reduktion der Freiheitsgrade bedeutet, dass Signifikanz für diese beiden Effekte auf relativ kurzen Zeitskalen festgestellt werden kann.

Neben der Dominanz des anthropogenen Strahlungsantriebs sind die geringen Effekte der Gruppierung auf die unklaren Trends in der langfristige Entwicklung im SOI zurückzuführen. Eine tatsächliche Beeinflussung dieser fundamentalen Klimagröße durch die Gruppierung lässt sich jedoch nicht widerlegen und steht im Einklang mit den Ergebnissen aus Kapitel 8.

### 9.2.3 EIMR

Der indische Monsun zeigt sich sowohl gegenüber dem anthropogenen Forcing als auch der Einteilung der Modelle sensitiv. Der EIMR verfügt zudem über signifikante Interaktionen der beiden Haupteffekte. Werden die jeweiligen Zeitreihen

der Regenzeiten ohne Glättung analysiert sind 82 % der Variabilität nicht über das statistische Modell erklärbar. Allerdings lassen sich mit der Spektralen ANOVA auf der dekadischen Zeitskala bereits etwa 42 % der Gesamtvariabilität als extern identifizieren. Weitere Glättung der Indexreihen erhöht diesen Anteil nur geringfügig.

Die Niederschlagszunahme über Indien zwischen Juni und September ist ein Charakteristikum, das sich in den meisten AOGCMs für das 21. Jahrhundert findet. Der Einfluss des anthropogenen Forcings weist zunächst einen Anteil von knapp 14 % auf und konvergiert bei zunehmender Glättung zu ungefähr 1/3 der Gesamtvarianz. Auch ist der hohe Betrag dieses positiven Trends im Ensemblemittel hervorzuheben (vgl. Tab. 9.1).

Daneben kann die Einteilung der Teilensembles auf allen Zeitskalen signifikante Anteile der Gesamtvarianz erklären. Dies erscheint, angesichts der bekannten und viel diskutierten Beziehung zwischen dem SOI und dem indischen Monsun (z. B. SHUKLA und PAOLINO 1983, WEBSTER und YANG 1992, HARZALLAH und SADOURNY 1997, KRISHNAMURTHY und GOSWAMI 2000), erwartbar. Das gilt allerdings nicht unbedingt für das Ausmaß. Auf dekadischer Skala können über 6 % der Varianz der Indexentwicklung auf diesen Blockeffekt zurückgeführt werden, bei weiterer Glättung erhöht sich der Wert auf etwa 8 %. Angesichts der Fragestellung, welche Effekte eine realistische ENSO-Dynamik auf die Ausprägungen des künftigen Klimawandels hat, sind diese Anteile durchaus als bedeutend zu bewerten. Ebenfalls für die Sensitivitätsstudie relevant ist die Tatsache, dass zwischen den beiden Haupteffekten signifikante Interaktionen vorhanden sind. Es finden sich also Unterschiede in den Trends der Teilensemlemittel, die bis zu 4 % der Gesamtvarianz in der transienten Zeitreihenentwicklung erklären. Die beiden Gruppen setzen das anthropogene Forcing signifikant unterschiedlich in Reaktionen der Indexreihen um. Dies deutet auf Sensitivitätsunterschiede hin, die in Kapitel 10 eingehender thematisiert werden.

Der indische Monsun muss in Folge der Analyse als äußerst empfindlich gegenüber Veränderungen des Strahlungshaushaltes gelten. Die starke Übereinstimmung zwischen den verschiedenen Modellen ist angesichts der Tatsache, dass der hier verwendete Index auf deren simulierten Niederschlägen basiert, besonders hervorzuheben. Diese Variable ist nicht nur durch hohe raumzeitliche Variabilität gekennzeichnet, sondern wird in GCMs zudem im Wesentlichen über Parametrisierungen realisiert. Beides kann eine Ursache für hohe unsystematische Variabilität im Multi-Modell Ensemble sein. Um eine ungefähre Abschätzung der relativen Bedeutung der beiden Komponenten dieser Kategorie zu erhalten, wurde eine ähnliche Untersuchung, allerdings mit anderer Gruppenbildung durchgeführt. Der Ansatz der Spektralen ANOVA wurde nicht auf sämtliche verfügbaren CMIP3-Läufe angewendet, sondern nur auf jene Modelle, die über mehrere Simulationen des 21. Jahrhunderts verfügen. So kann der Einfluss der individuellen Modellphysiken erfasst werden (PAETH und POLLINGER 2010). Bei geringer Glättung ist noch der als interne

Variabilität bezeichnete Einfluss unterschiedlicher Startbedingungen ausgeprägter als der Blockeffekt der Modelle. Die Bedeutung der systematischen Unterschiede zwischen den verschiedenen AOGCMs nimmt allerdings schnell zu. Der hierdurch erklärte Varianzanteil übertrifft den der Startbedingungen, sobald die Zeitreihen mit 9-jährigen Tiefpassfiltern geglättet werden. Diese Analyse zeigt des Weiteren, dass die Reaktion des Index auf den Strahlungsantrieb nicht unabhängig vom jeweiligen Klimamodell ist. Da hierbei sowohl Modelle die Teilensemble *A* als auch solche die Teilensemble *B* zugeordnet wurden untersucht wurden, lässt sich nicht ausschließen, dass dies ursächlich auf das Gruppierungskriterium zurückzuführen ist.

### 9.2.4 AMR

Der westafrikanische Monsun weist im Zuge der Spektralen ANOVA eine hohe Variabilität innerhalb der Gruppen auf. Während der Regenzeit müssen auf allen Zeitskalen mehr als 85 % der Varianz in diese Kategorie eingestuft werden. Entsprechend ist das unterstellte statistische Modell scheinbar nicht besonders gut geeignet, um die Entwicklungen im SRES A1b zu erfassen.

Da für den AMR als aus den simulierten Niederschlagsfeldern berechneten Index, die beim EIMR angestellten Überlegungen ebenso gelten wie im Fall des indischen Monsuns, wurde die gleiche zusätzliche Untersuchung durchgeführt, um einen etwas tieferen Einblick in die „black-box“ der internen Variabilität zu erhalten (siehe Kapitel 9.2.3). Nach der Entfernung der hochfrequenten Variabilität durch Filterlängen von über elf Jahren ist in dieser Analyse der systematische Unterschied zwischen den Modellen von größerer Bedeutung als die Unterschiede zwischen den verschiedenen Läufen eines Modells. Für den AMR zeigen sich bei dieser Betrachtung auch relativ hohe Beträge bei der Interaktion der beiden Haupteffekte, die sich wegen der nur geringen Besetzung der Gruppen in dieser Analyse aber nicht als signifikant erwiesen haben (keine Abbildung). Nichtsdestotrotz kann dies als Hinweis auf unterschiedliche Umsetzungen des Antriebs aus dem Strahlungshaushalt durch verschiedene Modelle gewertet werden.

Dennoch lassen sich auch für den AMR bei der Berücksichtigung aller Modelle einige modellübergreifende Entwicklungen festhalten. Das Multi-Modell Ensemble zeigt durchaus Sensitivität gegenüber dem anthropogenen Forcing. Der lineare Trend ist hierbei deutlich positiv. Allerdings ist die durch den entsprechenden Haupteffekt erklärte Varianz relativ gering, entsprechend der hohen Variabilität im Ensemble. So werden hier auch auf längeren Zeitskalen, bei dekadischer oder stärkerer Tiefpassfilterung der Zeitreihen, Beträge bis zu maximal 12.5 % erreicht. Für dekadische Zeitskalen erscheint daher mithilfe der verwendeten AOGCMs kein reales Prognosepotential zu bestehen.

Erschwerend kommt bei dieser Problematik außerdem hinzu, dass mit der Vegetationsdynamik ein wichtiger und hochkomplexer Baustein für die Genese und

Erklärung der Variabilität des afrikanischen Monsuns in den verwendeten Modellen nicht enthalten ist. Eventuell zeigt sich hier aber auch eine Möglichkeit auf, aussagekräftigere Simulationen mit künftigen Modellgenerationen zu erzeugen (PAETH und THAMM 2007).

Das Fehlen dieses fundamentalen Aspekts sollte sich auch hinderlich auf die Abschätzung der Bedeutung eines realistischen ENSO-Systems auswirken. Daher wird auch dies eine Aufgabe für neue Modellgenerationen darstellen. Für die CMIP3-Modelle leistet die Gruppierung der Modelle, ebenso wie die Interaktion, nur geringe Beiträge zur Gesamtvariabilität. Sie erklären jeweils nicht einmal 1 % der Gesamtvarianz und sind zudem nur auf einigen Zeitskalen statistisch signifikant. Der AMR ist, zumindest bei den verwendeten Modellen, dementsprechend als unempfindlich gegenüber den Unterschieden in der ENSO-Dynamik zu bewerten.

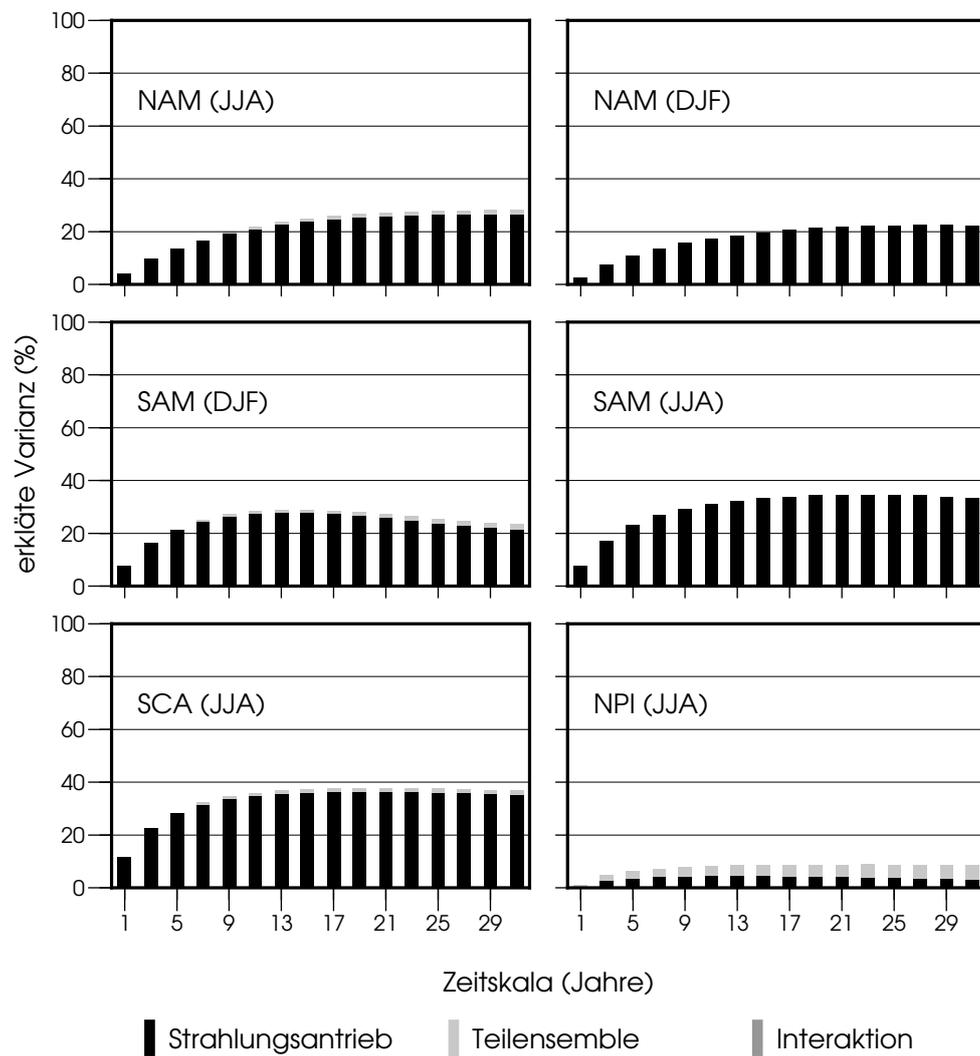
### 9.2.5 NAM

Die modellübergreifende Entwicklung des NAM ist zu sämtlichen Jahreszeiten von einem positiven Trend geprägt. Somit zeigt sich die dominierende Mode der Nordhemisphäre grundsätzlich als sensitiv gegenüber dem anthropogenen Forcing. Der ausgeprägteste Trend findet sich im borealen Herbst. Das Verhalten der einzelnen Modellläufe weist hier für einen derart dynamischen Index eine bemerkenswert hohe Übereinstimmung auf. Schon bei einer geringfügigen Glättung mit einem 3-jährigen Tiefpassfilter lassen sich etwa 20 % der Varianz im Multi-Modell Ensemble auf die zeitliche Entwicklung zurückführen. Dieser Anteil erhöht sich bei zunehmender Glättung der Zeitreihen weiter. Bei dekadischen Filtern werden bereits 40 % derartig attribuiert. Ab einer Filterlänge von etwa 21 Jahren stellt sich eine Sättigung bei einem Anteil um etwa 47 % ein.

Der anthropogene Einfluss zu den anderen Jahreszeiten wird durch die Analyse ebenfalls erfasst. Qualitativ ist er dann ähnlich ausgeprägt wie im Herbst, allerdings sind die Varianzanteile, ebenso wie die linearen Trends, geringer. Es lässt sich bei zunehmender Glättung, bis zu etwa einer dekadischen Filterlänge, jeweils ein deutlicher Anstieg der Bedeutung des Forcings für die Zeitreihenentwicklung feststellen. Eine noch weitergehende Glättung hat weniger Konsequenzen und schlussendlich konvergieren die Varianzanteile gegen einen saisonal unterschiedlichen Maximalwert. Für den Winter können so auf dekadischer Skala knapp 19 % und maximal etwa 22 % der Varianz auf das Forcing zurückgeführt werden. Da zu dieser Jahreszeit die Auswirkungen des NAMs auf die betroffenen Kontinente am größten sind, liegt auf dieser Jahreszeit auch besonderes Interesse. Im borealen Sommer sind diese Varianzanteile um etwa 1 %–2 % höher. Die Bedeutung des NAMs für die Ausprägungen des regionalen Klimas der betroffenen Gebiete sind jedoch im Sommer weniger stark ausgeprägt als im Winter und auch in den Beobachtungsdaten weniger konstant als in den Wintermonaten. Für den Frühling lässt sich zwar, sobald die hochfrequenten Varianzanteile durch eine geringe Glättung

herausgefiltert sind, der anthropogene Einfluss signifikant detektieren, dieser bleibt allerdings mit kaum mehr als 10 % erklärter Varianz quantitativ deutlich hinter dem Effekt zu den anderen Jahreszeiten zurück.

Insgesamt stellt der systematische Effekt des Forcings im NAM also ein ganzjähriges, stabil erfassbares Phänomen dar. Gleiches gilt auch für die grundsätzliche Tendenz der Indexentwicklung, den Trend zu positiven NAM-Phasen. Es zeigt sich hier also eine fundamentale großskalige, aber auch regional bedeutsame Konkretisierung des globalen Klimawandels. Allerdings muss, ebenfalls für alle Jahreszeiten, festgehalten werden, dass sich, trotz der teilweise sehr hohen Übereinstimmung im



**Abb. 9.2:** Ausgewählte Ergebnisse der Spektralen ANOVA für NAM, SAM, SCA und NPI. Wie Abb. 9.1.

SRES A1b, selbst auf dekadischem Maßstab kein zuverlässiges Prognosepotential für den NAM zu ergeben scheint.

Dagegen ist kaum ein Einfluss aus der Gruppierung in die Teilensembles feststellbar. Lediglich im Sommer und im Herbst lassen sich signifikante, jedoch nur geringfügige, Varianzanteile aus diesem Effekt heraus erklären. Maximale Werte, von jedoch nur 1.5 %, nimmt dieser Effekt im Sommer bei dekadischer Glättung an. Diese könnten, angesichts des bestehenden zeitlichen Lags, auf Telekonnektionen mit ENSO zurückgeführt werden (etwa BRÖNNIMANN 2007, BRÖNNIMANN et al. 2007). Interaktionseffekte konnten nicht festgestellt werden. Somit scheint der anthropogen bedingte Trend unabhängig von der Repräsentation des ENSO-Systems im GCM zu sein.

Die präsentierten Ergebnisse implizieren, dass im Fall des NAMs die interne Variabilität als bedeutendste Einflussgröße identifiziert werden kann. Da das Experiment keine weitere Differenzierung dieser Größe erlaubt, kann nicht geklärt werden, ob für den NAM die unterschiedlichen Startbedingungen der individuellen Modellläufe oder aber systematische Effekte der Modellphysiken dominieren. Jedoch sind Modellprojektionen des NAM durch hohe Sensitivität gegenüber wechselnden Startbedingungen gekennzeichnet, was sich durchaus auch bei starker Glättung der Zeitreihen feststellen lässt (RAUTHE et al. 2004).

### 9.2.6 SAM

Der SAM zeigt zu allen Jahreszeiten ausgeprägte Sensitivität gegenüber dem wachsenden Strahlungsantrieb. Die Ensemblemittel weisen jeweils positive Trends auf, die sich nicht von denen der Teilensembles unterscheiden. Für alle Jahreszeiten ist der entsprechende Haupteffekt auf allen betrachteten Zeitskalen von signifikanter Bedeutung, wobei der konkrete Anteil der erklärten Varianz sich aber systematisch unterscheidet.

Der stärkste Einfluss auf die Entwicklung im Ensemble ist für die Monate März bis Mai im Herbst der Südhalbkugel feststellbar. Schon ohne Glättung erklärt der Strahlungsantrieb hier knapp 14 % der Gesamtvarianz. Infolge der Reduktion der hochfrequenten Variabilität erhöht sich der relative Anteil auf dekadischen Zeitskalen auf etwa 47 %, bei noch stärkerer Filterung können sogar über 50 % der verbliebenen Varianz im Ensemble durch diesen Anteil erklärt werden. Eine Besonderheit im Analyseergebnis des SAMs ist die Tatsache, dass für die in Abb. 9.2 dargestellten Tiefpassfilter lokale Maxima der erklärten Varianz bei Filterlängen zwischen 11 und 19 Jahren existieren. Am deutlichsten ist dieses zu allen Jahreszeiten feststellbare Verhalten während des Sommers ausgeprägt.

Diese Verringerung der erklärten Varianzanteile korreliert mit dem Anwachsen der externen Variabilität. Als Interpretation dieses ungewöhnlichen Ergebnisses kann eine Zunahme der systematischen Unterschiede zwischen den einzelnen Modellen bei der Betrachtung sehr langer Zeitskalen angenommen werden, da sich

der Anteil an Variabilität, die sich durch unterschiedliche Anfangsbedingungen ergibt, üblicherweise bei zunehmender Glättung der Zeitreihen effektiv reduzieren lässt. Insbesondere die Unterschiede zwischen verschiedenen Einzelläufen eines Klimamodells werden daher durch längere Tiefpassfilter reduziert.

Als inhaltliche Erklärung für diese strukturelle Besonderheit der Varianzspektren des SAMs könnten Differenzen in der zukünftigen Entwicklung des stratosphärischen Ozons herangezogen werden. Diese ist im SRES A1b nicht homogenisiert. Für diese These spricht, dass die Auswirkungen des anthropogen bedingten Ozonabbaus in der Stratosphäre auf die Entwicklung des SAMs während des Südsommers am ausgeprägtesten ist (vgl. THOMPSON und SOLOMON 2002).

Auch die Gruppierung der Modelle erklärt Teile der Variabilität im Multi-Modell Ensemble. Ganzjährig lassen sich schon bei relativ geringer Glättung, entsprechend einer Filterlänge von drei oder fünf Jahren, signifikante, allerdings geringe, Effekte feststellen. Die höchsten Anteile erreicht dieser Effekt bei starker Glättung während des Sommers. Jedoch kann auch in diesem Fall das anthropogene Forcing etwa zehnmal soviel Varianz erklären. Der Einfluss der Gruppierung ist im Herbst am geringsten, für Winter und Frühling sind geringfügig größere Varianzanteile erklärbar. Interaktionen der Haupteffekte lassen sich beim SAM nicht signifikant erfassen. Somit zeigt sich die Entwicklung der dominierende Mode der südhemisphärischen Außertropen als unempfindlich gegenüber dem ENSO-System, wohingegen anthropogene Emissionen massive Effekte hervorrufen, die weitreichende Folgen für die atmosphärische Zirkulation im 21. Jahrhundert implizieren.

### 9.2.7 NPI

Das Verhalten des NPI im SRES A1b gibt keinen Hinweis auf ein modellübergreifendes, die regionale Entwicklung prägendes Klimaänderungssignal. Die linearen Trends sind zu allen Jahreszeiten nur gering ausgeprägt. Das Ensemblemittel weist für die wichtige Wintersaison einen negativen Trend auf, der die Intensivierung des Aleutentiefs und somit der zonalen Komponente der atmosphärischen Zirkulation beschreibt. Dagegen deuten die positiven Trends der kombinierten Indexreihen während der anderen Jahreszeiten eine Abnahme der zyklonalen Aktivität im Nordpazifik an. Die Unsicherheit über die zukünftige Entwicklung des NPI äußert sich allerdings nicht nur in diesen schwachen und gegenläufigen Trends. Die Spektrale ANOVA offenbart, dass sich beim NPI kaum systematische Effekte identifizieren lassen, so dass auch die langfristige Entwicklung des Ensemblemittels eher als Ausdruck von zufälligen Fluktuationen als als konsistente Reaktion auf ein physikalisches Forcing angesehen werden sollte.

Im Winter kann die Analyse keinerlei systematischen Effekt identifizieren. Das heißt unabhängig von der verwendeten Zeitskala entfallen 100 % der Varianz auf die interne Variabilität. Für den NPI ist bei einem kleineren Teil des CMIP3 bereits eine starke Abhängigkeit vom verwendeten Modell nachgewiesen worden (PAETH

und POLLINGER 2010). Insbesondere die Auflösung des AOGCMs dürfte hierbei eine Rolle spielen. Diese bestimmt die tatsächliche Höhe des Himalaya und des Hochlands von Tibet, was für die Beeinflussung des Jetstreams und damit die genetischen Voraussetzungen für die Ausprägung des Aleutentiefs entscheidend ist. Es ist daher davon auszugehen, dass dieser Einfluss auch bei der hier untersuchten größeren Modellanzahl, neben den Effekten unterschiedlicher Startbedingungen, von großer Bedeutung ist.

Ebenfalls allein von der internen Variabilität geprägt ist die Indexentwicklung im Frühling. Somit muss auch in diesem Fall der geringe positive Trend der gemittelten Indexreihe auf unsystematische, zufällige Effekte innerhalb des Multi-Modell Ensembles zurückgeführt werden. Für den etwas deutlicheren Trend dieser Zeitreihe im borealen Herbst detektiert die Analyse dagegen einen signifikanten systematischen Effekt des anthropogenen Forcings. Dieser beläuft sich auf maximal 3 % der Varianz der Zeitreihe, so dass auch hier die unterschiedlichen Modellphysiken, unabhängig von ENSO, sowie quasi-natürliche Effekte als die dominierenden Faktoren betrachtet werden müssen.

Im Nordsommer lässt sich ein geringer Anteil an der Gesamtvarianz über den Strahlungsantrieb erklären. Je nach verwendeter Glättung können bis zu 5 % der Variabilität im Ensemble auf diesen Einfluss zurückgeführt werden. Der lineare Trend entspricht dem des NPIs Herbst. Im Sommer lässt sich ein signifikanter Einfluss der Gruppenbildung nachweisen. Der Blockeffekt zwischen den beiden Teilensembles erklärt durchweg größere Varianzanteile als der Strahlungsantrieb, wobei deren Beträge bei zunehmender Filterlänge gegen etwa 7 % der Gesamtvarianz konvergieren.

Schon angesichts der geographischen Position des Aktionszentrums des NPIs erscheint bei diesem Index eine Reaktion auf ein El Niño oder La Niña Ereignis erwartbar (z. B. auch SHABBAR und KHANDEKAR 1996, BONSAI et al. 2001, NEWMAN et al. 2003). Ein Einfluss der Gewichtung in diesem Fall daher wenig überraschend. Die hohe Variabilität des SOI, insbesondere in A, dürfte aber auch hier den Blockeffekt zugunsten der internen Variabilität reduzieren.

Die Ergebnisse der Spektralen ANOVA deuten zudem für die Sommerzeitreihe auf eine Interaktion zwischen den Haupteffekten hin. Diese erklärt aber nur etwa 1 % der Gesamtvarianz und ist nur in einigen Fällen auf dem 5 %-Niveau signifikant. Jede Interpretation dieses Effekts muss daher vor dem Hintergrund einer mangelnden statistischen Stabilität gesehen werden. Trotzdem erscheint sie für diese Zeitreihe, insbesondere da sich beide Haupteffekte als signifikant erwiesen haben, physikalisch plausibel. Die hier identifizierten Varianzanteile deuten in diesem Fall auf gegenläufige Trends in den gemittelten Teilensembles hin. Der geringe Prozentsatz an erklärter Varianz lässt sich wiederum auf die großen Unsicherheiten bei der Entwicklung des SOI zurückführen.

### 9.2.8 SCA

Für das SCA zeigt die Analyse für alle Jahreszeiten im Multi-Modell Ensemble einen positiven Trend als Reaktion auf die steigende Treibhausgaskonzentration. Der wachsende Strahlungsantrieb erklärt zudem auf allen Zeitskalen signifikante Anteile an der Gesamtvarianz. Strukturell zeigen sich hierbei keine Auffälligkeiten. Dementsprechend steigt der betreffende Varianzanteil bei zunehmender Glättung von anfänglich geringen Werten zunächst rasch an. Dieser Anstieg schwächt sich ab der Verwendung dekadischer Filterlängen ab. Bei weiterer Glättung der Zeitreihen konvergiert die erklärte Varianz. Für diese Zirkulationsmode muss daher für die Zukunft mit deutlichen, systematischen Veränderungen gerechnet werden, wobei allerdings, gemäß dem Spektrum, hohe Variabilität auf kurzen Zeitskalen weiterhin zu erwarten ist.

Der maximale Trend des Ensemblemittels und der stärkste Einfluss des Haupteffekts findet sich in den Sommermonaten. Hier kann schon ohne Filterung etwa 10 % der Varianz im Ensemble auf diese Weise erklärt werden. Dieser Anteil steigt bei dekadischen Zeitskalen auf über 1/3. Im Winter ist der Effekt schwächer. Bei 11-jähriger Filterung können etwa 13 % der Gesamtvariabilität über den Strahlungsantrieb erklärt werden, maximale Anteile von knapp 16 % werden bei weiterer Glättung erreicht. Auch im Winter ist der anthropogene Einfluss auf das SCA also auf der dekadischen Zeitskala gut erfasst.

Diese Feststellung gilt auch für Frühling und Herbst. Für die Saison März bis Mai beträgt die erklärte Varianz aus dem Forcing bei der Verwendung des elfjährigen Tiefpassfilters knapp 20 %, bei weitergehender Glättung stagniert dieser Anteil um 23 %. Im Herbst ist der Einfluss des Strahlungsantriebs auf die zeitliche Entwicklung der Indexreihen größer, aber strukturell vergleichbar. So ermöglichen es dekadische Filter, etwa 28 % der Varianz auf diesen Effekt zurückzuführen. Weitere Glättung der Zeitreihen erhöht diesen Anteil nur noch auf etwa 30 %. Gegen diesen Wert scheint der Anteil auch zu konvergieren.

Trotz weitgehender Reduktion der Variabilität im Ensemble steigt die relative Bedeutung des anthropogenen Signals in allen Jahreszeiten nicht weiter an. Stattdessen bleibt der nicht erklärbare Anteil der Varianz nahezu konstant. Ob dies auf Modellunterschiede zurückzuführen ist oder auf den großen Einfluss der simulierten natürlichen Variabilität kann nicht abschließend beantwortet werden. Einen Hinweis gibt jedoch wiederum die Durchführung der Analyse von PAETH und POLLINGER (2010). Diese zeigt keine signifikante Zunahme des Blockeffektes zwischen verschiedenen AOGCMs bei den entsprechenden Tiefpassfiltern, so dass die natürliche Variabilität als dominant in allen Jahreszeiten betrachtet werden muss.

Die geographische Lage der Aktionszentren des SCA lassen eine große Abhängigkeit vom ENSO-Phänomen unwahrscheinlich erscheinen. Jedoch existieren speziell für den Sommer Fernwirkungen in den Regionen, die durch das SCA besonders

beeinflusst werden (etwa MARES et al. 2003). Zu dieser Ausgangslage passend weist die Gruppierung der Modelle nur im Sommer signifikante Effekte auf das Ensemble auf. Auf keiner betrachteten Zeitskala überschreiten diese allerdings 1.5 % der Gesamtvariabilität, so dass dieser Haupteffekt für den SCA, auch im Vergleich zum anthropogenen Einfluss, als nur geringfügig relevant gelten kann.

---

## 10 ENSO-Performance und Klimaänderungen: Zeitreihenentwicklung

Im Folgenden wird überprüft, ob sich signifikante Unterschiede in der transienten Entwicklung der beiden Teilensembles feststellen lassen. Hierbei liegt ein Schwerpunkt der Betrachtung auf der konsistenten und plausiblen Entwicklung während der Integrationsperiode. Damit stellen die hier gezeigten Ergebnisse wichtige Ergänzungen zu Kapitel 9.1 dar. Die Existenz solcher Differenzen zeigt dann eine in CMIP3 existierende Unsicherheitsquelle auf, die in ihrer Bedeutung geeignet ist, die langfristigen Abschätzungen des anthropogenen Klimawandels als weniger gesichert erscheinen zu lassen als es die neutrale Betrachtung der diversen AOGCM-Experimente nahelegt. Entsprechend des Aufbaus dieser Sensitivitätsstudie ist der einzige infrage kommende Kandidat für diese Unsicherheitsquelle die Repräsentation des SOI in Raum- und Frequenzdomäne.

### 10.1 Systematische Differenzen: Existenz und Persistenz

Zunächst werden die Zeitreihen der Multi-Modell Mittelwerte der Teilensembles auf signifikante Differenzen untersucht. Durch die zeitliche Allokation und Beurteilung der Persistenz dieser Effekte wird die eventuelle Attribution auf das Gruppierungsmerkmal unterstützt. Ziel ist hierbei insbesondere die Absicherung gegen die Überinterpretation von Ergebnissen, die auf numerischen Artefakten beruhen. Auf diese Weise sollen die Effekte der Gruppierung als stabil gekennzeichnet werden, bevor weitere Analysen erfolgen.

#### 10.1.1 Methodik

Um eventuelle systematische Abweichungen zwischen den Teilensembles zu detektieren, werden für alle betrachteten Klimagrößen die Zeitreihen der Mittelwerte der beiden Multi-Modell Teilensembles  $\bar{a}(t)$  und  $\bar{b}(t)$  untersucht. Hier gilt, mit  $n_A$

als Anzahl der Läufe in  $A$ :

$$\bar{a}(t) = \frac{1}{n_A} \sum_{i=1}^{n_A} a(t, i) \quad (10.1)$$

Die Berechnung von  $\bar{b}(t)$  erfolgt dementsprechend über alle  $n_B$  Einzelläufe von  $B$ . Welche Klimagröße diese Zeitreihen jeweils repräsentieren ist durch den Aufbau des Kapitels eindeutig zu erfassen.

Die Analyse nutzt den normalen  $t$ -Test für unabhängige Stichproben. Geprüft wird, mit separaten Tests für jedes Modelljahr  $t$ , ob  $\bar{a}(t) = \bar{b}(t)$  verworfen werden kann. Diese  $H_0$  wird beidseitig zum Signifikanzniveau  $\alpha = 0.05$  getestet. Bei diesem Vorgehen ist es angesichts der großen Anzahl von Testdurchführungen wahrscheinlich, dass bei jeder der Zeitreihen Fehler der 1. Art auftreten, die eventuell schon durch ein einziges Extremereignis in einem der Einzelläufe hervorgerufen werden können. Dies ist eine unmittelbare Folge der hohen Variabilität, die das ganze Multi-Modell Ensemble während der gesamten Analyseperiode aufweist. Insbesondere Extremereignisse, aber auch weniger dramatische Effekte der simulierten natürlichen Dynamik der Zeitreihen, sind allerdings ebenso geeignet, modellübergreifende systematische Entwicklungen zu maskieren (siehe Kapitel 9.1), was statistisch dem Fehler 2. Art entspricht. Es ist somit erforderlich, die Analyse und vor allem die Interpretation ihrer Ergebnisse gegen diese beiden ernstzunehmenden Problematiken abzusichern. Da letztlich die hochfrequenten Varianzanteile die potentielle Basis für beide Arten der Ergebnisverfälschung darstellen, wird diese analog zum Vorgehen in Kapitel 9.1 sukzessiv aus den Temperatur- und Indexreihen der Einzelläufe durch Glättung mit anwachsenden Filterlängen entfernt. Der statistische Test wird dann mit den Teilensemblemitteln  $\bar{a}_k(t)$  und  $\bar{b}_k(t)$  durchgeführt. Diese berechnen sich für  $A$ , und entsprechend für  $B$ , als Multi-Modell Mittelwerte der gefilterten Einzelläufe:

$$a_k(t, i) = \frac{1}{k} \sum_{j=t-\frac{k}{2}}^{t+\frac{k}{2}} a(j, i) \quad (10.2)$$

Auch die benötigte Stichprobenstandardabweichung bezieht sich dann auf die gefilterten Werte. Die so erzeugte spektrale Struktur der Differenzen von  $A$  und  $B$  erlaubt die zuverlässige Detektion von tatsächlichen Unterschieden. Diese haben das Potential, das bestehende Bild vom zukünftigen Klimawandel zu beeinträchtigen.

### 10.1.2 Ergebnisse

Konsistent mit den vorherigen Untersuchungen zeigt sich die Entwicklung der außertropischen Moden nur als wenig sensitiv im Hinblick auf die Gruppierung. Zwar erzeugen die  $t$ -Tests für jeden der Indices und zu allen Jahreszeiten durchaus

signifikante Ergebnisse, deren häufig zufällig erscheinendes Auftreten in den verschiedenen Modelljahren spricht allerdings dafür, dass es sich hierbei nicht um echte, von der Fragestellung gesuchte Differenzen zwischen den Teilensembles handelt. Des Weiteren ist die Anzahl dieser Treffer üblicherweise sehr gering. Entsprechend dieser Eigenschaften entfernt die Glättung die meisten diese hier als Artefakte gewerteten Effekte. Somit kann für diese Zeitreihen konstatiert werden, dass sich die Entwicklung in  $A$  und  $B$  nicht grundlegend unterscheidet. Das bedeutet allerdings nicht, dass das Gruppierungskriterium keinen Einfluss auf die Entwicklung der Klimagrößen nimmt. Die divergente Entwicklung des SOIs in  $A$  könnte im ungünstigsten Fall dazu führen, dass sich realistische Reaktionen auf langfristige Veränderungen im ENSO-System in den individuellen Läufen bei der Mittelbildung gegenseitig ausgleichen. Somit besteht grundsätzlich die Möglichkeit, dass auch für diese Zirkulationsmuster eine größere Unsicherheit besteht als sich über das SRES A1b über traditionelle Methoden abschätzen lässt.

Trotzdem ist die Methode auch bei der scheinbar ungünstigen Struktur der Variabilität in den Teilensembles durchaus in der Lage, signifikante Ergebnisse zu erzeugen. Eindeutige Unterschiede zwischen  $A$  und  $B$  werden zwar nur bei einigen wenigen Klimagrößen aufgezeigt, unter diesen befinden sich allerdings drei, die unzweifelhaft zu bedeutsamsten im Hinblick auf ihre Rolle im Klimasystem und ihre Auswirkung auf den Menschen gehören. Es handelt sich um:

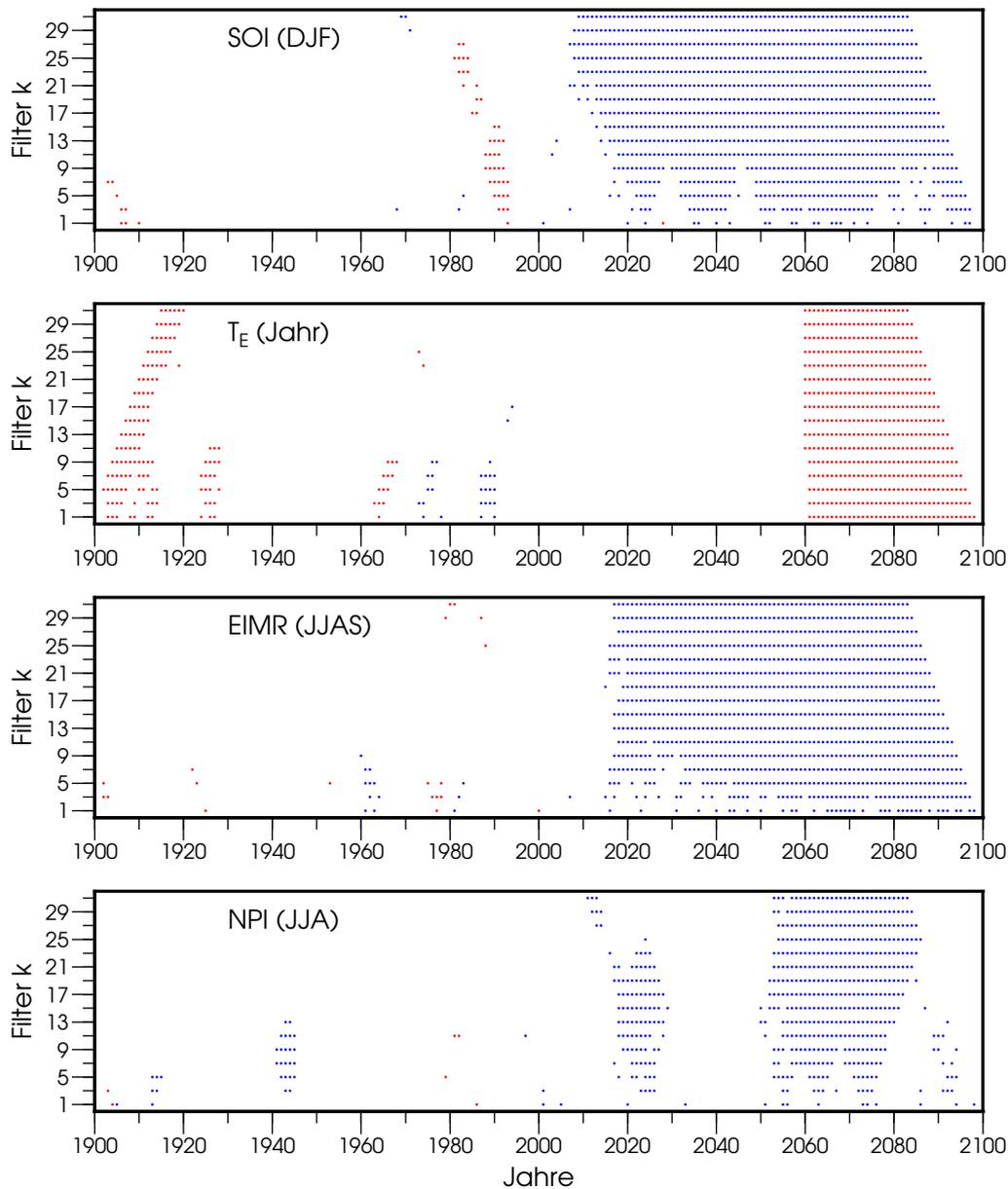
- den SOI während der entscheidenden Monate Dezember, Januar und Februar
- den Niederschlag über dem indischen Subkontinent während der Monsunsaison
- die Entwicklung der globalen Temperatur

Weiterhin unterscheidet sich die Zeitreihe des NPIs in den Sommermonaten signifikant zwischen den Teilensembles.

Bei allen diesen Größen konzentrieren sich die signifikanten Testergebnisse auf das Ende des Beobachtungszeitraums, wenn das anthropogene Forcing am ausgeprägtesten ist. Weiterhin zeigt Abb. 10.1, dass die Differenzen der jeweiligen Zeitreihen  $\bar{a}_k(t)$  und  $\bar{b}_k(t)$  kaum Vorzeichenwechsel aufweisen, sondern eines der Teilensembles das andere systematisch unterschätzt. Diese Eigenschaften sind stabil bei der Analyse der gefilterten Zeitreihen. Sie werden sogar durch die Entfernung der hochfrequenten Varianz und, vermutlich, einiger Artefakte die auf Fehlern der 1. Art beruhen, hervorgehoben.

Die Resultate sind für diese Klimagrößen statistisch also unbedingt als robust zu bezeichnen. Daneben macht das Ergebnis dieser Analyse auch physikalisch eindeutig Sinn, wie im Folgenden dargelegt wird.

Das ENSO-Phänomen hat im borealen Winter die größte Klimawirksamkeit. Die Southern Oscillation ist in dieser Jahreszeit am stärksten ausgeprägt und



**Abb. 10.1:** Systematische Unterschiede zwischen den Teilensembles für vier Klimagrößen. Markiert sind signifikante Differenzen. Rot bedeutet  $\bar{a}_k(t) > \bar{b}_k(t)$ , blau das Gegenteil. Wenn  $k > 1$  zeigt die Markierung das mittlere Jahr an. Berechnet aus zweiseitigen  $t$ -Tests zum 5 %-Signifikanzniveau.

die Kopplung von Ozean und Atmosphäre für die Dynamik entscheidend (z. B. RASMUSSEN und CARPENTER 1982). Somit ist der Effekt der Gruppierung, wenn überhaupt, in eben dieser Zeitreihe zu vermuten. Die Ergebnisse lassen hierbei

systematische Unterschiede der Teilensembles, insbesondere im Hinblick auf die Reaktion auf das anthropogene Forcing, erkennen.

So lassen sich im 20. Jahrhundert kaum signifikante Differenzen der Mittelwerte detektieren (Abb. 10.1). Diese wiederum erscheinen ohne auffällige zeitliche Konzentration. Dementsprechend sind sie bei der Glättung mit längeren Tiefpassfiltern in der Regel nicht stabil. Bei der Betrachtung der dekadischen Zeitskala reduziert sich die Anzahl der signifikanten Differenzen deutlich. Vereinzelt liefern auch die Tests der noch stärker geglätteten Daten signifikante Ergebnisse, die allerdings keine physikalisch interpretierbare Struktur aufweisen und somit als Artefakte betrachtet werden.

Dagegen zeigt der SOI im 21. Jahrhundert für die beiden Teilensembles derartig konstante Differenzen zwischen  $\bar{a}_k(t)$  und  $\bar{b}_k(t)$ , dass reine Zufallseffekte ausgeschlossen werden können. Schon auf Basis der Jahresdaten wird die grundsätzliche Tendenz in  $A$  zu, relativ zu  $B$ , negativen Indexwerten des SOI deutlich. Somit lässt sich, trotz der hohen interannuellen Variabilität dieser Größe, zumindest festhalten, dass die Grundbedingungen für El Niño-Ereignisse in  $A$  häufiger gegeben sind als in  $B$ . Diese Aussage ist stabil auf sämtlichen betrachteten Zeitskalen. Damit ist der Schluss zulässig, dass die Teilensembles hier klimatologisch relevante Unterschiede in der Umsetzung des anthropogenen Strahlungshaushalts aufweisen. Die Reaktion des SOI ist, zumindest während des borealen Winters, abhängig von der Repräsentation dieser Mode durch das jeweilige Klimamodell.

Der Zusammenhang von SOI und  $T_E$  wurde bereits diskutiert (vgl. Kapitel 7) und als relevanter Einfluss auf das tatsächliche Ausmaß der globalen Erwärmung im CMIP3 identifiziert (Kapitel 8). Dementsprechend erscheinen die signifikanten Unterschiede zwischen  $A$  und  $B$  während der zweiten Hälfte des 21. Jahrhunderts folgerichtig und wenig überraschend. Dass diese Größe jedoch in jedem einzelnen Jahr ab 2060 für  $A$  signifikant größer ist als in  $B$  verdient dennoch Beachtung. Die interne Variabilität der Modelle ist bei dieser Größe nicht ausgeprägt genug, um die Auswirkungen der klimatologischen Tendenzen des SOI in den beiden Teilensembles, insbesondere in  $A$ , auszugleichen. Angesichts der Konstanz dieser Unterschiede und der Bedeutung von  $T_E$  für das Klimasystem als Ganzes, scheint es zulässig den beiden Teilensembles unterschiedliche Sensitivität für das Forcing des SRES A1b zu unterstellen. Die Wahl des Gruppierungsmerkmals wird hierdurch bestätigt.

Zusätzlich finden sich noch zwei weitere Phasen, in denen die beiden Multi-Modell Mittelwerte signifikante Differenzen aufweisen. Eine lässt sich ungefähr zwischen den Jahren 1960–1990 lokalisieren. Unterschiedliche Vorzeichen der Differenzen in diesem Zeitabschnitt legen es jedoch nahe, hier Effekte der interannuellen Variabilität der Klimagröße zu vermuten. Auf der dekadischen Zeitskala sind diese Effekte nicht mehr feststellbar.

Die signifikanten Ergebnisse zu Beginn des 20. Jahrhunderts sind dagegen auch in den gefilterten Zeitreihen feststellbar. Zudem handelt es sich hier um systematische

Differenzen ohne Vorzeichenwechsel. Eine mögliche Erklärung innerhalb dieser Sensitivitätsstudie für diese Periode findet sich, wenn der SOI während der anderen Jahreszeiten betrachtet wird. Für diese Periode zeigt die Analyse in diesen Fällen ebenfalls einige signifikante Ergebnisse an.

Daneben bestehen für NPI und EIMR bekannte und komplexe Telekonnektionen mit ENSO. Auch wenn teilweise nicht abschließend geklärt ist welcher physikalische Prozess hierbei tatsächlich den Ausschlag gibt (etwa WEBSTER et al. 1998, TORRENCE und WEBSTER 1999, NEWMAN et al. 2003, LAU et al. 2004, YU und ZWIERS 2007), bieten diese für die Ergebnisse dieses Abschnitts doch eine durchaus tragfähige Interpretationsbasis. AOGCMs ohne überzeugende ENSO-Performance, nach dem hier vertretenen Standpunkt, bilden diese Telekonnektionen insgesamt schlechter ab als die des Teilensembles *A* (vgl. Kapitel 7.1.3). Entsprechend sollten sich auch die Auswirkungen des anthropogenen Klimawandels auf den SOI nicht so deutlich in den Fernkopplungen niederschlagen.

Beide Teilensembles zeigen keinerlei systematische Unterschiede bei der Entwicklung des indischen Monsun bis etwa 2010. Vereinzelt statistisch signifikante Unterschiede, jeweils in beide Richtungen, können auf die interne Variabilität der Einzelläufe zurückgeführt werden. Auf längeren Zeitskalen nähern sich die beiden Zeitreihen wieder an. Dagegen sind die Zeitreihen im 21. Jahrhundert, insbesondere bei der Untersuchung gefilterter Daten, in weiten Teilen signifikant voneinander unterschiedlich. Die Niederschlagsentwicklung über Indien im Teilensemble *A* unterschätzt die in *B* systematisch. Da mit wärmeren Bedingungen auch erhöhte Evapotranspiration einher geht, kann diese Differenz massive Folgen für die dicht besiedelte Großregion haben.

Hier wird an dieser Stelle eine äußerst ungünstige Kombination im CMIP3 detektiert: der Monsun stellt ein hochkomplexes natürliches System dar, dass für eine enorme Anzahl von Menschen von entscheidender Bedeutung für ihre Wirtschafts- und Lebensweise ist. Angesichts der hieraus resultierenden gesellschaftlichen Vulnerabilität gegenüber jedweder Form von mittel- oder langfristigen Veränderungen müssen im Hinblick auf den Klimawandel Adaptionenmaßnahmen durchgeführt werden. Offensichtlich ist jedoch die Fähigkeit das ENSO-System zu erfassen entscheidend bei der Abschätzung der Reaktion auf steigende Treibhausgaskonzentrationen. Die Aussagekraft von Projektionen die auf dem kompletten CMIP3-Ensemble beruhen ist daher in diesem wichtigen Detail als nicht optimal zu bezeichnen. Entsprechend sind Adaptionenmaßnahmen, denen derartige Projektionen zugrunde liegen, unter Umständen nicht ideal für die tatsächlichen Herausforderungen des Klimawandels. Die Anzahl der vom indischen Monsun direkt und indirekt abhängigen Menschen rechtfertigt es sicherlich, möglichst präzise und verlässliche Prognosen zu dessen zukünftiger Entwicklung als von globalem Interesse zu bezeichnen.

Neben diesen auf die Qualität der Klimamodelle zielenden Argumenten, kann auch die unterschiedliche Grundtendenz in den Teilensembles als Erklärung her-

angezogen werden. Hierbei muss allerdings die hohe Variabilität der SOI-Reihen insbesondere in  $A$  als beschränkender Faktor beachtet werden.

## 10.2 Transiente Varianzattribuion

Für die Zeitreihen erfolgt nun der letzte Schritt der Sensitivitätsstudie. Ziel ist die Erfassung der Entwicklung der varianzgenerierenden Größen im Zeitverlauf. Die Auswirkungen des Strahlungsantriebs wurden in Kapitel 9.1 bereits eingehend analysiert. Bestehende Sensitivitäten wurden dort erläutert. Auf die umfassende Darstellung für sämtliche Zeitreihen im Hinblick auf diesen Einfluss wird an dieser Stelle daher verzichtet.

Stattdessen werden hier die Ergebnisse der Untersuchung für jene Klimagrößen präsentiert, die im Kapitel 10 als systematisch von der Gruppierung beeinflusst identifiziert wurden. Hierbei liegt der Fokus auf dem Einfluss der Gruppierung, um einen tieferen Einblick in die durch unterschiedliche ENSO-Performance generierte Unsicherheit im CMIP3 zu erlangen. Dieser Effekt wird durch die verwendete Methodik betont.

### 10.2.1 Methodik

Als Analysewerkzeug wird wiederum eine Variante der zweifaktoriellen ANOVA genutzt. Um die zeitliche Entwicklung der Einflussgrößen zu detektieren, werden nicht alle verfügbaren Modelljahre, sondern nur jeweils ein 30-jähriger Abschnitt der Zeitreihen untersucht. Dieser bewegt sich als gleitendes Fenster über den Zeitraum 1900–2098 hinweg. Dieses Vorgehen offenbart die Dynamik der Beeinflussung der Modellläufe durch die Faktoren.

In den Abbildungen 10.2, 10.3, 10.4 und 10.5 sind die Varianzanteile, die sich aus dem Strahlungsantrieb, der Gruppierung der GCMs in die beiden Teilensembles sowie die Interaktion dieser beiden Faktoren ergeben jeweils als Zeitreihen dargestellt. Zur besseren Übersicht werden sie jeweils zu Summenkurven kombiniert. Varianzanteile, die nicht durch einen  $F$ -Test als signifikant klassifiziert werden konnten, wurden hierfür der internen Variabilität zugerechnet, die die Kurven jeweils auf den Wert 100 % ergänzt. Die Werte sind so eingetragen, dass die in einem 30 Jahresabschnitt gefundene Varianzstruktur jeweils im Zentraljahr dieser Zeitspanne zu finden ist. Für ungefilterte Daten ist das Zentraljahr der ersten ANOVA 1915, das der letzten 2084.

Die relativ kurzen Zeiträume bedingen, dass insbesondere Trends, hier der Einfluss des Strahlungsantriebs, durch die interne Variabilität der Zeitreihen leichter maskiert werden können. Zwar kann auch in diesem Fall über die Glättung der Ausgangszeitreihen dieser Problematik entgegen gewirkt werden, die Reduktion der Freiheitsgrade führt allerdings schnell an den Punkt, an dem keinerlei signifikante

Ergebnisse mehr erzeugt werden können. Dennoch ist dieses Vorgehen aufgrund der relativ großen Datenbasis zu einem gewissen Grad möglich.

Um die Überinterpretation von vereinzelt Artefakten zu vermeiden, wurden Tiefpassfilter mit Längen von bis zu 21 Jahren verwendet. Der Vergleich dieser Resultate zeigt, dass die Verwendung von dekadischen Filtern einen guten Kompromiss zwischen Robustheit und Sensitivität darstellt. Daher werden in den entsprechenden Abbildungen derartig gefilterte Daten und die Ergebnisse der Analyse dieser Daten gezeigt. Diese Beschränkung stellt einen konservativen Ansatz dar, gewählt um die Interpretation der Ergebnisse abzusichern. Es soll dennoch explizit erwähnt werden, dass in mehreren Fällen bei der Verwendung weniger starker Filter durchaus weitere signifikante Ergebnisse zusätzlich zu den hier dargestellten erzielt werden konnten.

## 10.2.2 Ergebnisse

Die Detektion des anthropogenen Signals durch die vielfältig beeinflusste Varianz innerhalb des Multi-Modell-Ensembles ist über die Spektrale ANOVA zuverlässiger und effektiver möglich. Im Hinblick auf die Unterschiede zwischen den Teilensembles bietet die hier durchgeführte Analyse jedoch weitere Einsichten, die die bisheigen Ergebnisse dieses Kapitels sinnvoll ergänzen.

### 10.2.2.1 SOI

Die beiden Zeitreihen  $\bar{a}_{11}(t)$  und  $\bar{b}_{11}(t)$  weisen im Fall des SOI eindeutig unterschiedliche Entwicklungen auf (Abb. 10.2). Während die Modelle mit guten Bewertungen ihrer ENSO-Performance, allen Unwägbarkeiten der Variabilität in den Einzelläufen zum Trotz, insgesamt klar in Richtung El Niño tendieren, deutet sich für das Teilensemble  $B$  das Gegenteil an. Dessen SOI hat im Mittel für den gesamten betrachteten Zeitraum einen positiven Trend. Diese gegenläufige Entwicklung setzt bereits im 20. Jahrhundert, etwa um 1960, also unter noch relativ geringem anthropogenen Strahlungsantrieb ein. Die scheinbare Annäherung der beiden Teilensembles ist lediglich auf die Standardisierung der Indexreihen zurückzuführen. Die in Abb. 10.2 wiedergegebene Struktur ist jedoch unabhängig von der Wahl der hierfür verwendeten Periode.

Ein besonderes Merkmal, mit weitreichenden Konsequenzen sowohl für die hier unternommenen Analysen als auch für die Interpretation der Ergebnisse, ist die völlig unterschiedliche Entwicklung der Varianz in den Teilensembles. Für  $A$  lässt sich, unabhängig von der Glättung der Ausgangsdaten, unter anthropogenem Einfluss eine deutliche Erhöhung der Varianz im Teilensemble,  $s_a^2$ , erkennen.

Diese hier quantifizierte und sichtbar gemachte Unsicherheit über die Entwicklung der dominierenden Zirkulationsmode dürfte entscheidenden Anteil daran haben, dass die Gewichtung nicht in der Lage war, signifikante Ergebnisse zu

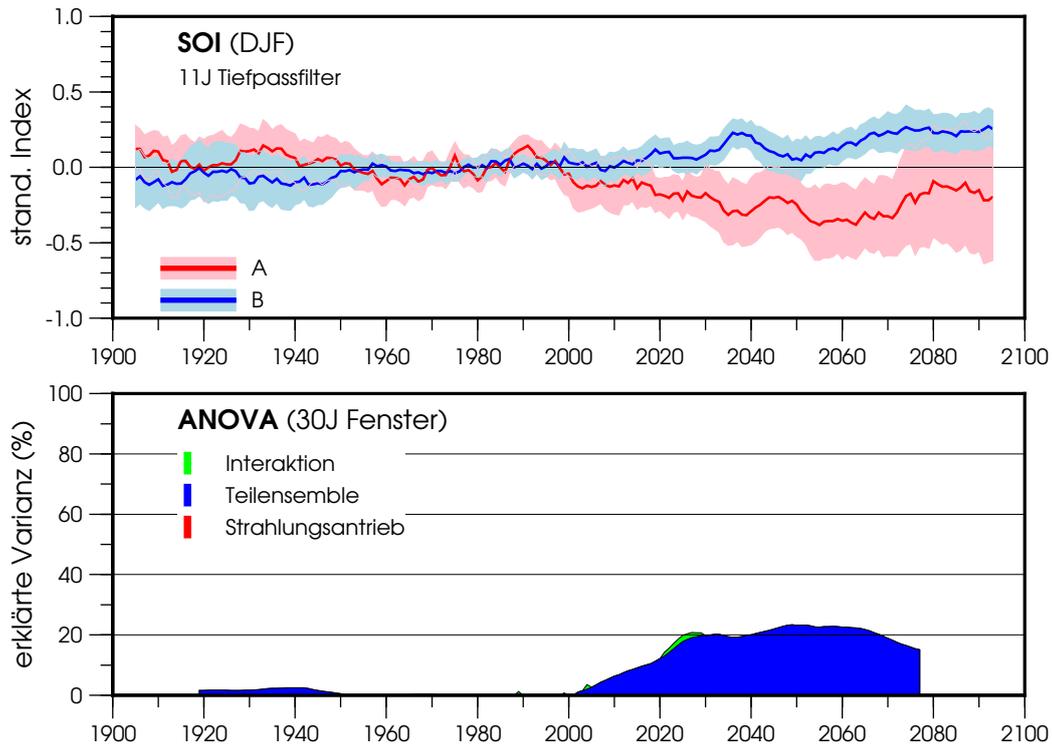
erzielen. Konkret verantwortlich gemacht werden können die vielfachen Abweichungen individueller Modellläufe von der mittleren Tendenz von  $A$ . Da sich  $s_a^2$  unter steigendem Forcing immer weiter erhöht ist es sicher angebracht, diese Varianzerhöhung auch kausal auf die steigenden Treibhausgase zurückzuführen. Es ist also festzuhalten, dass eben jene Modelle, denen hier, in wesentlicher Übereinstimmung mit OLDENBORGH et al. (2005), ein realistisches ENSO-System zugebilligt wird, teilweise direkt gegenläufige Reaktionen im Klimawandel zeigen. Als Ersatz für konkrete Zukunftsprognosen müssen die üblichen Verfahren zur Auswertung zumindest der CMIP3-Läufe somit als noch ungeeigneter erscheinen.

Im Teilensemble  $B$  hingegen streuen die Indexwerte des SOI nicht nur deutlich geringer als in  $A$ , es lässt sich sogar eine Reduktion von  $s_a^2$  unter dem anthropogenen Forcing feststellen. Zwar wird die Verringerung der natürlichen Variabilität des ENSO-Phänomens als eine Möglichkeit für die Zukunft durchaus diskutiert (etwa COLLINS et al. 2010), üblicherweise wird dieses Szenario allerdings nicht mit der in  $B$  ebenfalls bestehenden Tendenz hin zu häufigeren oder stärkeren La Niña-Bedingungen in Verbindung gebracht.

Die transiente ANOVA liefert mit Blick auf die bisherigen Ergebnisse durchweg plausible Ergebnisse. Bei dekadischer Filterung ist für keines der Zeitfenster ein allgemeiner, modellübergreifender Trend der Indices feststellbar. Stattdessen hat die Gruppierung der Modelle bereits im 20. Jahrhundert signifikanten Einfluss auf die Ausprägungen des SOI. Diese werden durch die beschriebenen unterschiedlichen Grundtendenzen bedingt. Werden Zeitfenster mit Zentraljahren zwischen etwa 1970 und 2000 analysiert, maskiert die Standardisierung diesen Effekt. Anschließend steigt der durch diesen Haupteffekt erklärte Anteil der Gesamtvarianz allerdings schnell an und erreicht ungefähr bei dem durch das Zentraljahr 2030 markierte Zeitfenster einen Wert von rund 20 %. Um dieses Niveau herum hält sich der durch die Gruppierung erklärte Varianzanteil bis zum Ende des 21. Jahrhunderts. Ab dem Zentraljahr 2070 lässt sich allerdings eine geringe abnehmende Tendenz dieses Wertes feststellen.

Von besonderer Wichtigkeit ist jedoch der signifikante Beitrag den die Interaktion der beiden Haupteffekte hier leistet. In Abbildung 10.2 finden sich zwar nur geringe Anteile von etwa 2–3 % die dieser Effekt an der Gesamtvarianz hat, jedoch stellen diese signifikante, qualitativ unterschiedliche Entwicklungen in den Teilensembles dar. Sie zeigen die unterschiedlichen Trends der beiden Ensemblemittel, die andersartige Reaktion auf den Strahlungsantrieb. Die Interaktion der Haupteffekte ist ein Hinweis darauf, dass sich, immer nur im Hinblick auf den hier untersuchten SOI, die Simulationen der CMIP3-Modelle und der reale Klimawandel in grundsätzlicher Weise unterscheiden können. Verbunden wäre dies dann aber natürlich mit den für den SOI angemessenen globalen Konsequenzen und Auswirkungen.

Wird statt eines Tiefpassfilters mit 11 Jahren als Basis für den gleitenden Durchschnitt ein auf dem 5 jährigen gleitenden Mittel basierender Filter verwendet, so finden sich für deutlich mehr Zeitabschnitte signifikante, durch die Interaktion



**Abb. 10.2:** SOI: Transiente Varianzattribuion. Multi-Modell-Mittelwerte der beiden Teilensembles (elfjähriger gleitender Durchschnitt), jeweils mit 95 %-Konfidenzintervall (oben). Unten: zeitlich aufgelöste zweifaktorielle ANOVA der gefilterten Daten (11 Jahre gleitender Durchschnitt; Erläuterungen im Text). Alle dargestellten Varianzanteile sind signifikant für  $\alpha = 0.05$ .

erklärte Varianzanteile. Sie finden sich zeitlich konzentriert in der ersten Hälfte des 21. Jahrhunderts. Diese Ergebnisse lassen sich derart interpretieren, dass sich das Klimasystem im Hinblick auf den SOI und somit auf ENSO derzeit in einer entscheidenden Phase befindet.

### 10.2.2.2 Globale Mitteltemperatur

Für  $T_E$  weisen  $\bar{a}_{11}(t)$  und  $\bar{b}_{11}(t)$  in Abb. 10.3 eindeutig mehr Gemeinsamkeiten als Unterschiede auf. In beiden Fällen ist das dominierende Element der Zeitreihen der massive, positive Trend. Auch ist die Temperaturentwicklung schon ab einer geringfügigen Glättung hochkorreliert. Die grundsätzlichen Effekte des anthropogenen Forcings sind in beiden Teilensembles über einen langen Zeitabschnitt hinweg quasi identisch. Etwa zwischen den Modelljahren 1920 und 2020 lassen sich nicht nur keine signifikanten Differenzen feststellen, die Abweichungen der beiden

Mittelwertreihen sind auch betragsmäßig kaum zu unterscheiden. Hinsichtlich der Vorzeichen lässt sich keinerlei Präferenz feststellen, so dass es für diesen Zeitraum keine offensichtlichen Argumente gibt, die beiden Teilensembles als nicht identisch zu betrachten.

Diese synchrone Entwicklung lässt sich durch die Ergebnisse der Varianzanalyse quantifizieren. Spätestens ab den 1960er Jahren verliert die interne Variabilität der Modellläufe ihre dominante Bedeutung für die Entwicklung von  $T_E$ . Frühere Schwankungen, ausgelöst durch die simulierte natürliche Variabilität, und systematische Unterschiede zwischen den AOGCMs treten in den Hintergrund. Es folgt eine vom modellübergreifenden Trend geprägte Phase. Die deutlichste Übereinstimmung der Modelle stellt die ANOVA eindeutig für jene gleitenden Fenstern fest, die das Ende des 20. Jahrhunderts abdecken. Somit bilden alle Modelle für diese Zeit einen sehr ähnlichen Trend ab. Der Haupteffekt erklärt für diese Zeitabschnitte, unabhängig von der Glättung durch Tiefpassfilter, 80 % und mehr der Gesamtvarianz. Diese hohe Übereinstimmung dürfte an den für diesen Zeitraum zur Verfügung stehenden guten Beobachtungsdaten liegen. Sicherlich ist hier zu beobachten, dass im Entwicklungsprozess der AOGCMs eben die befriedigende Nachbildung der globalen Mitteltemperatur in diesem Zeitabschnitt wichtiges Kriterium zur Beurteilung der Modellqualität ist. Entsprechend sinkt der Grad der Übereinstimmung zwischen den Modellen mit dem Beginn der Szenariodaten ab dem Jahr 2000 ab. Die interne Variabilität der Einzelläufe gewinnt wieder an Bedeutung, der Trend der gemittelten Zeitreihen in Abb. 10.3 bleibt aber weiter nahezu identisch. Mit deutlich über 55 % dominiert der Strahlungsantrieb die Entwicklung von  $T_E$  bis in die Modelljahre um etwa 2040 weiterhin deutlich.

Noch in der ersten Hälfte des 21. Jahrhunderts jedoch trennen sich die beiden Zeitreihen. Die Erwärmungsrate ist in *A* durchgehend höher. Die Unterschiede der Teilensemblesmittel, auch wenn die entsprechenden Differenzen erst zum Ende des betrachteten Zeitraums hin signifikant werden, sind eindeutig und systematisch.

Die größere Variabilität in *A* ist ebenfalls deutlich erkennbar. Sie wird hier als Effekt der höheren Sensitivität von  $T_E$  in *A* gegenüber ENSO gewertet. In der zweiten Hälfte des 21. Jahrhunderts lässt sich schließlich auch ein signifikanter Anteil der Variabilität im SRES A1b aus der Gruppierung erklären. Dieser Anteil erhöht sich kontinuierlich und beträgt zum Ende des Szenarios etwa 13 %. Angesichts der Tatsache, dass der durch das anthropogene Forcing erklärte Varianzanteil für diese Periode nur noch knapp 18 % beträgt, kann der Gruppierung durchaus ein ernstzunehmender Einfluss zuerkannt werden.

Die signifikanten Effekte der Gruppierung zu Beginn der Analyseperiode dürften ebenfalls durch die Unterschiede in den Auswirkungen der ENSO-Variabilität auf  $T_E$  begründet sein (siehe Kapitel 10.1.2).

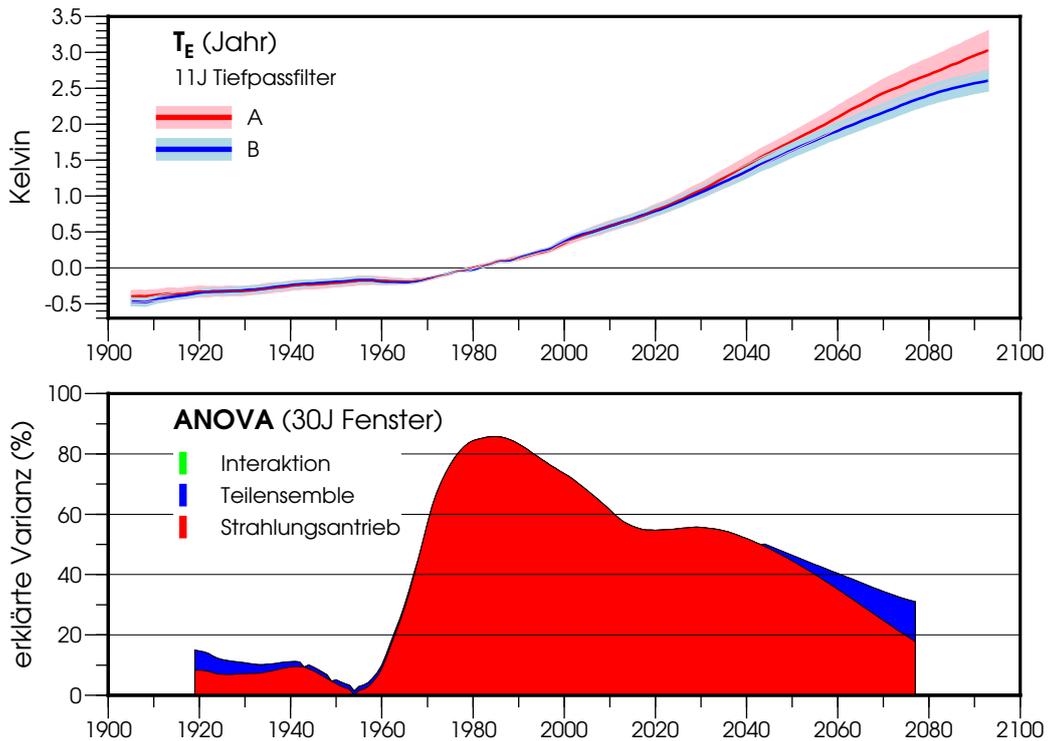


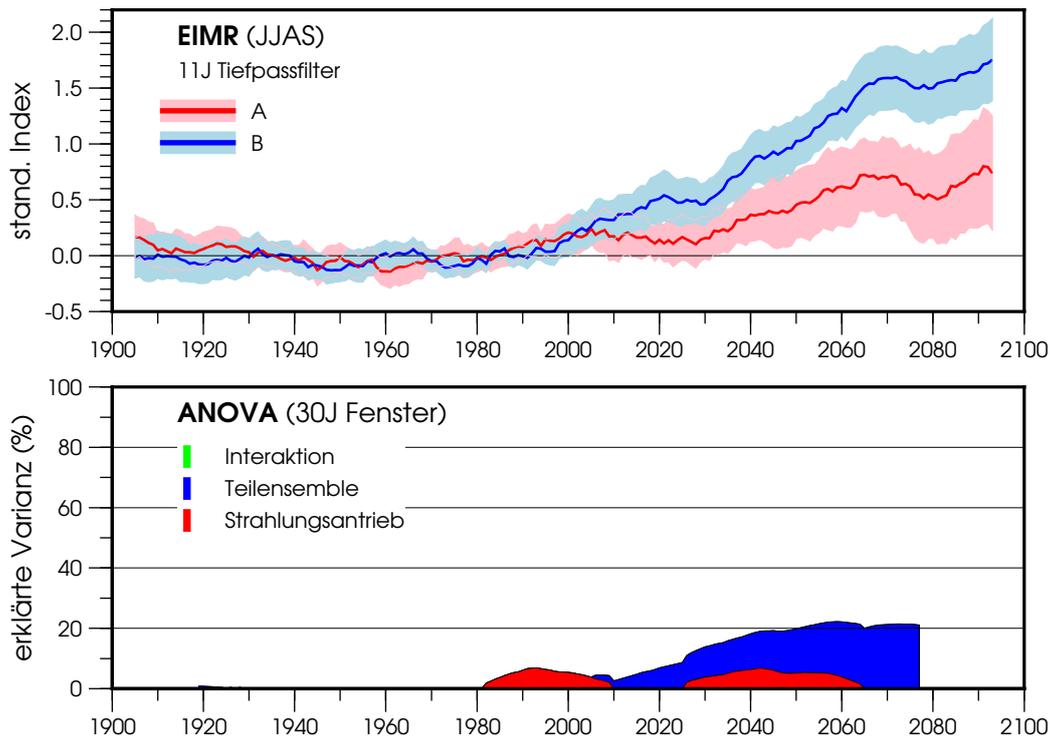
Abb. 10.3:  $T_E$ : Transiente Varianzattribuion. Wie Abb. 10.2.

### 10.2.2.3 EIMR

Die Entwicklung des Monsuns ist in beiden Teilensembles durch einen positiven Trend im Verlauf des 21. Jahrhunderts gekennzeichnet. Dieser Anstieg der Indexwerte des EIMR ist jedoch in beiden Mittelwertzeitreihen von dekadischen Schwankungen überlagert. Diese Variabilität erschwert die Detektion der langfristigen Veränderung durch die transiente ANOVA. Dennoch kann für das Ende des 20. Jahrhunderts und weite Teile des 21. Jahrhunderts, bei moderater Glättung der Indexreihen, eine gewisse modellübergreifende Entwicklung festgestellt werden. Für die relativ kurzen Zeitfenster von 30 Jahren lassen sich jedoch kaum mehr als 10 % der Varianz hierdurch erklären, in Anbetracht der hohen Variabilität des Niederschlags ein durchaus hoher Wert.

Sowohl innerhalb von *A* als auch von *B* sind bei dieser Größe auf allen Zeitskalen starke Unterschiede zwischen den Einzelläufen feststellbar. Dies äußert sich nicht nur in der großen Bedeutung der internen Variabilität bei der Erklärung der Gesamtvarianz, sondern auch in hohen Werten von  $s_a^2$  und  $s_b^2$ . Trotz der hieraus resultierenden Unsicherheit, ist die quantitative Entwicklung in *A* und *B* ab der Mitte des 21. Jahrhundert klar voneinander getrennt. Der Trend im EIMR führt

in *A* nur zu einer etwa halb so starken Erhöhung wie in *B*. Der Beginn der unterschiedlichen Verläufe lässt sich etwa um das Jahr 2000, mit dem Übergang von den Randbedingungen vom 20C3m hin zum SRES A1b feststellen. Signifikante Varianzanteile lassen sich fast im gesamten 20. Jahrhundert auf die Gruppenzugehörigkeit attribuieren. Es handelt sich hierbei durchaus um hohe Beträge. Zum Ende des betrachteten Zeitraums, wenn neben der internen Variabilität auf allen untersuchten Zeitskalen nur noch die Gruppenzugehörigkeit signifikante Beiträge zur Variabilität im Ensemble leistet, werden für diese Größe Werte von über 20 % erreicht.



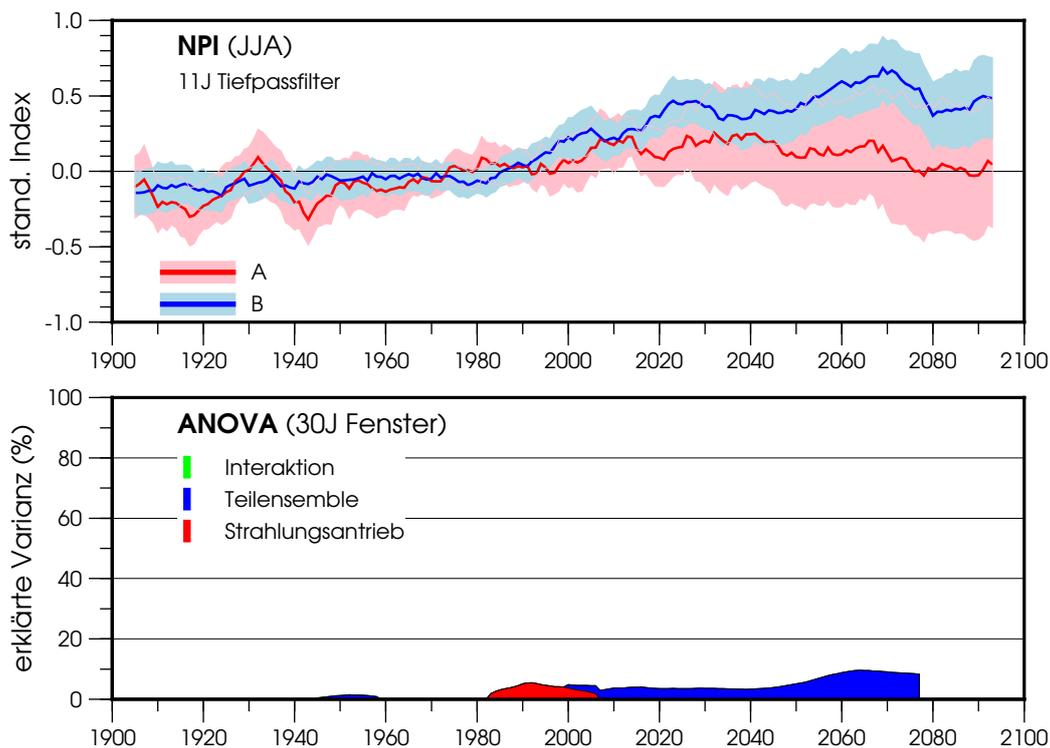
**Abb. 10.4:** EIMR: Transiente Varianzattribuion. Wie Abb. 10.2.

Beim EIMR erscheint die zeitliche Kopplung des Auftretens der Unterschiede zwischen den Teilensembles an die Differenzen im SOI besonders auffällig. Die transiente ANOVA bestätigt in diesem Fall die Ergebnisse aus Kapitel 10. Die deutliche El Niño-Tendenz, die sich in *A* ab etwa dem Jahr 2000 findet hat sicherlich das Potential, die synchron stattfindende Trennung der Reihen  $\bar{a}(t)$  und  $\bar{b}(t)$  im Fall des EIMR zumindest in Ansätzen zu erklären. Schließlich könnte auch das deutliche Anwachsen von  $s_a^2$  in den letzten Dekaden des 21. Jahrhunderts auf nahezu sprunghafte Erhöhung der entsprechenden Größe im SOI zurückgeführt werden. Der grundsätzliche Zusammenhang ist hier als wahrscheinlich zu bewerten,

die vielfältigen Möglichkeiten der Beeinflussung zwischen Monsun und ENSO verhindern im Rahmen dieser Analyse jedoch die Konkretisierung der hierfür verantwortlichen Mechanismen und Kopplungen. Für das Teilensemble  $B$  scheint dagegen, angesichts der tatsächlich qualitativen Unterschiede im langfristigen Verhalten der Mittelwertzeitreihen von SOI und EIMR, für diese höchstens ein geringer Zusammenhang zu bestehen. Die Kombination beider Erkenntnisse führt zu dem Ergebnis, dass ein bedeutender Teil der Unsicherheit über die Beeinflussung des EIMR im Multi-Modell Ensemble auf ENSO zurückführbar ist. Dessen realistische Repräsentation ist die Voraussetzung für einen merklichen Zusammenhang zwischen den beiden Phänomenen. Die konkreten Auswirkungen dieses Zusammenhanges lassen sich zwar in den Einzelläufen feststellen, die ungewisse Entwicklung des ENSO-Systems erhöht jedoch trotz dieser Erkenntnis die Unsicherheit über die Zukunft des indischen Monsuns.

#### 10.2.2.4 NPI

Positive Indexwerte des NPI sind mit verringerter Zyklonenaktivität im Nordpazifik verbunden. Weiterhin weisen positive Trends im NPI auf die Zunahme blockierender Hochdruckzellen im klimatologischen Mittel hin. Der Vergleich von  $\bar{a}(t)$  und  $\bar{b}(t)$  in Abb. 10.5 zeigt einige Auffälligkeiten in der Zeitreihenentwicklung, die auch im Fall dieser weniger prominenten Klimagröße die Auswirkungen der vorgenommenen Einteilung der Klimamodelle aufzeigen. Mit Ausnahme einer markanten Schwankung von  $\bar{a}_{11}(t)$  zwischen 1920 und 1940 zeigen die gemittelten Indexreihen beider Teilensembles bis zum Ende des 20. Jahrhunderts einen ruhigen Verlauf. Jedoch ist, unabhängig von der Glättung durch Tiefpassfilter, die  $s_a^2$  deutlich größer als die Varianz in  $B$ . Entsprechende  $F$ -Tests liefern, mit wenigen Ausnahmen zum Anfang des 20. Jahrhunderts, durchweg signifikante Ergebnisse. Diese hohe Variabilität in  $A$  verhindert auch die Identifikation von modellübergreifenden Trends. Lediglich in der relativ kurzen Periode von etwa 1980 bis 2010 können signifikante Anteile der Varianz auf den Strahlungsantrieb zurückgeführt werden. Hierbei ist ein positiver Trend im SRES A1b feststellbar. Ein solcher prägt, überlagert von einigen dekadischen Schwankungen, auch die weitere Entwicklung von  $\bar{b}(t)$ . Der Mittelwert von  $A$  zeigt dagegen die meiste Zeit im 21. Jahrhundert eine Seitwärtsbewegung, die zum Ende dieses Zeitraums jedoch eine deutliche Komponente hin zur Nulllinie aufweist. Im Bezug auf die Mittelwertzeitreihen findet sich also in  $B$  das deutlich stärkere anthropogene Signal. In  $A$  ist dagegen ein starkes und für das 21. Jahrhundert kontinuierliches Anwachsen von  $s_a^2$  feststellbar. Daher beinhalten die 95 %-Konfidenzintervalle  $KI(t)$ , unabhängig von vorheriger Glättung der Indexreihen, quasi durchgehend die 0, so dass keine signifikante Änderung konstatiert werden kann. Die Unterschiede zwischen den Teilensembles sind bedeutend genug, um im 21. Jahrhundert durchgehend signifikante Anteile bei der Erklärung der Gesamtvarianz aufzuweisen. Bis zu 12 % hiervon lassen sich



**Abb. 10.5:** NPI: Transiente Varianzattribuion. Wie Abb. 10.2.

auf die Gruppierung der Modelle zurückführen.

In Anbetracht der offensichtlichen Unterschiede der Entwicklung des NPI in den Teilensembles erscheint die Gruppierung auch in diesem Fall als relevant. Insbesondere die aufgezeigte Entwicklung der Variabilität in *A* legt die Vermutung nahe, dass in diesem Fall die Variabilität des SOI in diesem Teilensemble einen bedeutenden Anteil an dieser Entwicklung hat. Wiederum stellen die unbekanntenen Auswirkungen des veränderten Strahlungsantriebs auf die ENSO-Mode einen Faktor bei der Reaktion des Klimasystems auf das anthropogene Signal dar. Dagegen könnte die Entwicklung des NPI in *B* sowohl auf der geringeren Variabilität des SOI in diesem Teilensemble zurückgeführt werden als auch auf das Fehlen einer realistischen Kopplung der beiden Zirkulationssysteme. Für die zweite Interpretation spricht in diesem Fall, dass die Trendentwicklung der beiden Moden sich schon auf dekadischer Ebene deutlich unterscheidet.

# 11 ENSO-Performance und Klimaänderungen: Regionale Betrachtung

Der Einfluss des anthropogenen Klimawandels auf die drei in dieser Arbeit betrachteten Feldgrößen und insbesondere das hiermit verbundene räumliche Muster werden in diesem Kapitel im Rahmen der Sensitivitätsstudie thematisiert. Daher und dabei liegt der Fokus des Interesses hier nicht auf den tatsächlichen quantitativen Veränderungen dieser Klimagrößen, die in IPCC (2007a) bereits ausführlich thematisiert wurden. Stattdessen soll das Potential zur Erklärung von unterschiedlichen Ergebnissen der AOGCMs durch deren unterschiedliche ENSO-Systeme betont werden. Wenn dies aufgrund der ungewissen Entwicklung des hierbei ausschlaggebenden SOIs auch derzeit in der Praxis nur wenig zur Reduktion der Unsicherheit der Prognosen beiträgt, so verspricht dieses Vorgehen doch durchaus einen besonderen Nutzen. Dieser wird ersichtlich, wenn ein signifikanter Einfluss der Gruppierung für Regionen identifiziert werden kann, für die die Modelle des SRES A1b stark streuende Ergebnisse aufweisen. In dem Fall könnten, etwa zur Entwicklung von Adaptionsszenarien, bei der Abschätzung des regionalen Klimawandels unterschiedliche Entwicklungspfade von ENSO berücksichtigt werden. Hier ist jede aufgedeckte Unsicherheitsquelle hilfreich.

Daneben kann für Regionen ohne erkennbaren Einfluss der Gruppierung das Vertrauen in Langfristprognosen gestärkt werden oder auch die Suche nach regionalen Quellen der Unsicherheit initialisiert werden.

## 11.1 Methodik

Ziel ist es, die anhand der Entwicklung der fundamentalen Klimagrößen bereits aufgezeigten Differenzen von  $A$  und  $B$  räumlich und quantitativ zu konkretisieren. Hierfür werden die Unterschiede zwischen den Teilensembles für die bodennahe Temperatur, den Niederschlag und den Luftdruck separat betrachtet.

Die Analyse kann in Anbetracht der Datenmenge nicht auch nur ansatzweise so detailliert ausfallen wie bei der Untersuchung der, bereits sehr zahlreichen, Zeitreihen. Hier werden nur die absoluten Auswirkungen des Klimawandels auf die drei Klimatelemente und deren Unterschiede in den Teilensembles betrachtet. Als

Maß hierfür wird, analog zum Vorgehen in Kapitel 8, die Differenz zweier klimatologischer Mittelwerte verwendet. Der Subtrahend ist hier jeweils der klimatologische Mittelwert der Periode  $P_1$  von 1960–1999. Als Minuend wird der Mittelwert der von 2080 bis 2098 definierten Periode  $P_2$  genutzt. Diese Verschiebung erfasst und quantifiziert die grundsätzliche Tendenz in den Teilensembles.

$$\Delta X(s) = \bar{X}(P_2, s) - \bar{X}(P_1, s) \quad (11.1)$$

Die räumliche Verortung der Variablen erfolgt hier über die Gitterbox  $s$ . Es werden jeweils die Jahresmittelwerte und die vier üblichen Jahreszeiten separat untersucht. Die so gewonnenen Größen werden entsprechend ihrer ENSO-Performance und somit ihrer Zugehörigkeit zu den Teilensembles  $A$  und  $B$  zusammengefasst. Für diese werden für jede der  $\Delta X(s)$  Mittelwerte und Standardabweichungen nach den üblichen Verfahren errechnet. Mit  $\bar{a}_X(s)$  wird im Folgenden der Mittelwert der Verschiebungen von  $X$  im Teilensemble  $A$  bezeichnet, die dazugehörige Standardabweichung mit  $s_{a_X}(s)$ . Im Teilensemble  $B$  werden diese Größen entsprechend mit  $\bar{b}_X(s)$  und  $s_{b_X}(s)$  bezeichnet. Dieses Vorgehen macht es wiederum unvermeidbar, sämtliche Datensätze auf ein gemeinsames Gitternetz zu interpolieren. Verwendet wurde das der ERA 40-Reanalysen.

Zur Detektion von systematischen Unterschieden wird an jedem Gitterpunkt ein  $t$ -Test durchgeführt. Geprüft wird letztlich die Signifikanz der Effekte  $E_X(s)$ , die sich aus der Gruppierung ergeben.

$$E_X(s) = \bar{a}_X(s) - \bar{b}_X(s) \quad (11.2)$$

In Gleichung 11.2 steht der Index  $X$  für die jeweilige Feldgröße. In Anbetracht des Maßstabs, der dieser Untersuchung zugrunde liegt, kann und muss dieses Vorgehen als ausreichend betrachtet werden.

Da sowohl  $A$  als auch  $B$  Multi-Modell Ensembles sind, sollten sich sonstige Unterschiede zwischen den GCMs, sofern Sie unabhängig von ENSO sind, ausgleichen. Die hohe Variabilität des SOI könnte sich bei der Analyse als hinderlich erweisen. Jedoch zeigen die Kapitel 9 und 10, dass die typischen Unterschiede zwischen den Teilensembles ausgeprägt genug sind, um mit statistischen Methoden erfasst zu werden. Die vielfältigen Effekte, die die Veränderungen eines realistischen ENSO-Systems als Folge des Klimawandels haben können, wären sicher bei konsistenter Entwicklung des SOI in  $A$  eindeutiger zu bestimmen. Eine nachträgliche Sortierung dieser Läufe würde jedoch nicht nur der Zielsetzung dieser Untersuchung zuwiderlaufen, sondern die über die hier angewendeten Methoden gewonnenen Erkenntnisse entwerten, da die Voraussetzungen für ihre Anwendung nicht mehr gegeben wären.

## 11.2 Ergebnisse

Die Ergebnisse der Analyse werden für jedes Klimaelement separat diskutiert. Hier liegt der Fokus auf dem Aufzeigen der allgemeinen Strukturen, die die Ergebnisse der Analyse kennzeichnen und als stabil betrachtet werden können. Eine zu kleinräumige Betrachtung verbietet sich hierbei zum einen wegen des Datenumfangs, zum anderen aufgrund der Vielzahl an Faktoren, die die konkrete Ausprägung einer Klimavariablen an einem Gitterpunkt beeinflussen können. Die verwendete Interpolation ist hierbei nur ein Beispiel.

### 11.2.1 Effekte auf die bodennahe Temperatur

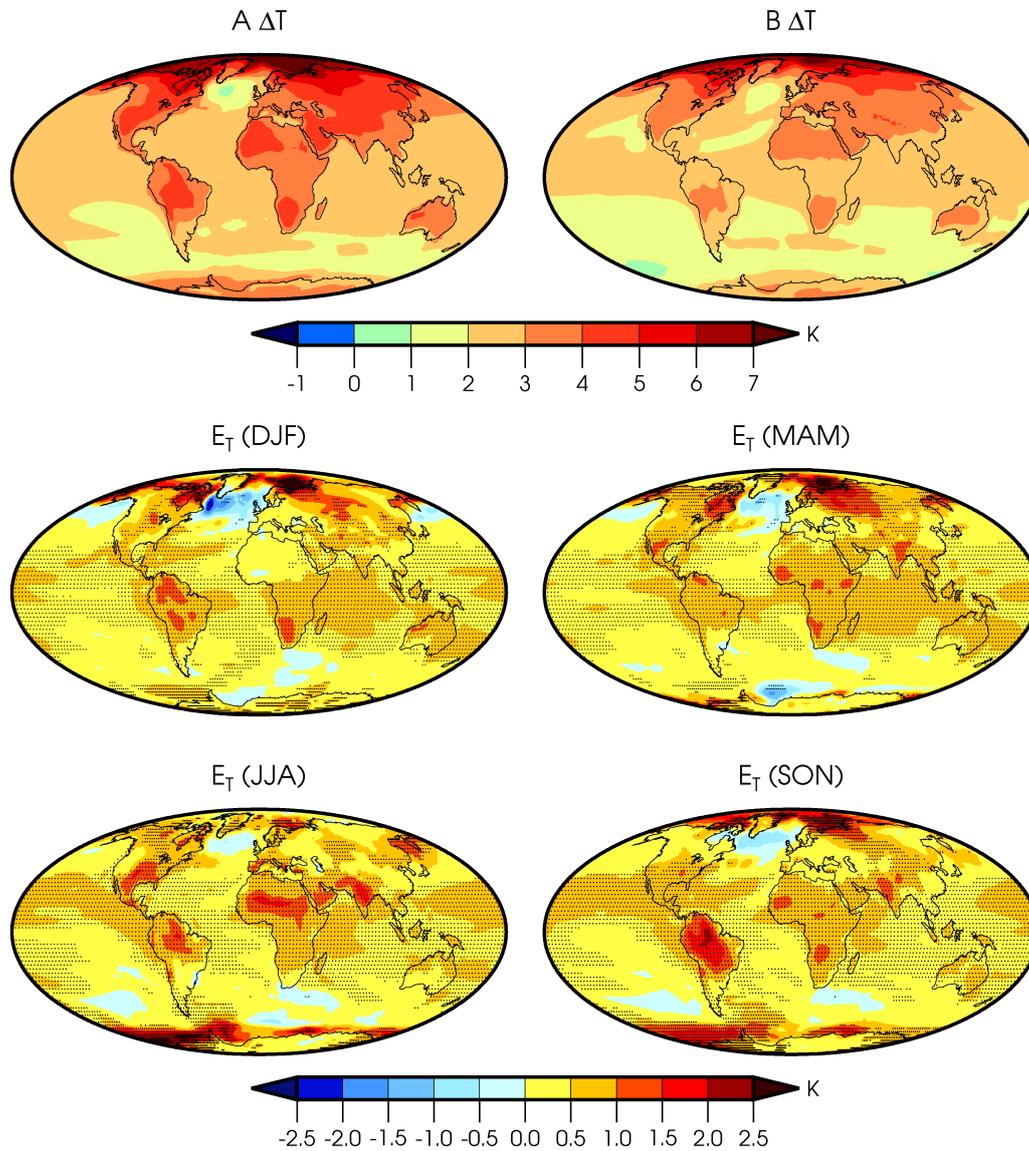
Die Erhöhung der bodennahen Temperatur bis zum Ende des 21. Jahrhunderts ist in *A* für nahezu die gesamte Erdoberfläche stärker ausgeprägt als in *B*. Dies korrespondiert mit den etwas stärkeren Effekten, die sich für die Zeitreihen von  $T_E$  ergeben. Unabhängig von der Jahreszeit ist das COWL-Muster, welches die räumliche Struktur der Temperaturentwicklung kennzeichnet, im Teilensemble *A* stärker ausgeprägt als in *B*. Dies resultiert in signifikant stärkeren Erwärmungen in *A* für weite Bereiche der Landmassen.

Der afrikanische Kontinent und die arabische Halbinsel erfahren, quasi ganzjährig und flächendeckend, im Teilensemble *A* eine um rund 0.8 K stärkere Erwärmung als in *B*. Insbesondere während der Sommermonate in den randtropischen Trockengebieten übersteigen die Differenzen der beiden Teilensembles aber auch auf der klimatologischen Ebene durchaus 2 K.

Für Australien lassen sich im Jahresmittel vergleichbare Unterschiede zwischen *A* und *B* feststellen. Auch hier ist die stärkere Erwärmung in *A* für nahezu den gesamten Kontinent signifikant. Die Differenzen zwischen den Teilensembles sind hier im Bereich der Südwestküste, wo die Beeinflussung des regionalen Klimas durch ENSO weniger stark ist als im Rest des Kontinents, am geringsten.

Die Unterschiede für Europa sind etwas geringer, aber auch differenzierter. Skandinavien und Nordosteuropa weisen sehr große signifikante Unterschiede zwischen *A* und *B*. Mit bis zu 2 K im Jahresmittel finden sich hier die deutlichsten Unterschiede resultiert zwischen den Teilensembles für die Landmassen. Die Höhe der Unterschiede, bei allerdings durchweg hohen Niveau, v.a. aus den Differenzen während des borealen Winters. Dagegen sind für die übrigen Gebiete des europäischen Festlandes hohe und signifikante Unterschiede in der Erwärmung vor allem im Sommer zu finden.

Für Asien sind die höheren Werte der Temperaturveränderung in *A* in weiten Teilen signifikant. Insbesondere der zentralasiatische Bereich sowie der indische Subkontinent weisen deutliche Differenzen auf. Im Fall von Indien sind diese während der Regenzeit besonders ausgeprägt. Dies resultiert zum einen aus der unterschiedlichen Entwicklung des EIMR in den Teilensembles, zum anderen bedeutet dies eine



**Abb. 11.1:** Räumliche Differenzierung von  $\Delta T$ . Sensitivität des Klimawandels gegenüber der Gruppierung nach der Repräsentation des ENSO-Systems. Oben:  $\Delta T(s) = \bar{T}(2080-2098, s) - \bar{T}(1960-1999, s)$  nach Teilensembles, jeweils jährliche Mittelwerte. Mitte und unten:  $E_T(s) = \bar{a}_T(s) - \bar{b}_T(s)$  nach Jahreszeiten. Schattierte Werte  $E_T$  sind signifikant zu  $\alpha = 0.05$  (zweiseitiger Test), die Einheit ist Kelvin

höhere potentielle Verdunstung, was sich wiederum auf die tatsächlich verfügbare Wassermenge auswirkt und die Konsequenzen für die Bevölkerung verschärfen kann. Südostasien und die indonesischen Inseln weisen ebenfalls stärkere Differenzen in

der Erwärmung auf als der Ozean.

Für Nord- und Mittelamerika ist die stärkere Erwärmung in *A* auf das ganze Jahr betrachtet in weiten Teilen statistisch signifikant. Allerdings unterscheiden sich die Muster der Differenzen zwischen den Jahreszeiten recht deutlich. Sie sind in den Wintermonaten sowohl an der Ostküste der USA als auch in weiten Teilen Kanadas nicht signifikant. Für die nördlichen Gebiete gilt dies auch während des Herbsts. Die stärksten Unterschiede in der Erwärmung finden sich für die Sommermonate im mittleren Westen der USA. Die Differenzen sind an der Ostküste für keine Jahreszeit signifikant.

In Südamerika äußern sich die Unterschiede im regionalen Klimawandel vor allem über die unterschiedliche Erwärmung im Bereich des Amazonasbeckens. Mit Ausnahme von Patagonien lassen sich jedoch für den gesamten Kontinent für weite Teile des Jahres signifikante Unterschiede feststellen. Diese sind durchaus im Einklang mit beobachteten Auswirkungen von El Niño-Ereignissen in der Region und könnten somit eventuell eine Folge des negativen Trends im SOI während des borealen Winters in *A* darstellen.

Die Beträge von  $E_T(s)$  sind über den Ozeanen im Allgemeinen geringer. Signifikante Werte werden daher vor allem in den tropischen Regionen erreicht, wo die Streuung der  $\Delta T(s)$  im SRES A1b nur gering ist. Neben dieser breiten Zone zwischen etwa  $\pm 30^\circ$  Breite konzentrieren sich die signifikante Differenzen auch im Bereich der Polarmeere. Hier finden sich die höchsten Werte für  $E_T(s)$ . Es ist dabei zu beachten, dass die Arktis im SRES A1b die Region mit der eindeutig stärksten Erwärmung ist. Für die Barentssee betragen die Unterschiede in den jährlichen Klimatologien bis zu 3 K und die Differenzen zwischen den Teilensembles sind in den übrigen Bereichen des Arktischen Ozeans nur unwesentlich geringer. Am prägnantesten sind sie im Winter, was darauf hindeutet, dass die Meereisverteilung hierbei einen Einfluss ausüben könnte. Gleiches gilt für die signifikanten Unterschiede die sich auch im Jahresmittel im antarktischen Ozean zwischen *A* und *B* in der Temperaturentwicklung feststellen lassen. Auch hier ist die mittlere Temperaturerhöhung im Winter im Teilensemble *A* um bis zu 3 K höher als in *B*. Im Bereich der Westwinde finden sich über den Meeren dagegen die wenigen Gebiete, in denen  $\bar{b}_T(s)$  stärkere Erwärmungen anzeigt als  $\bar{a}_T(s)$ . Insbesondere im Nordatlantik ist dieses Ergebnis für alle Jahreszeiten identisch, allerdings liefert der *t*-Test keine signifikanten Resultate. Dennoch bietet sich zumindest für diese Region eine physikalische Interpretation dieses auffälligen Aspekts an. Da die Entwicklung des NAM in beiden Teilensembles vergleichbar ist, liegt eine Erklärung durch unterschiedliche Entwicklungen in der ozeanischen Zirkulation nahe. So sollte die deutlich stärkere Erwärmung der Arktis in *A* die thermohaline Zirkulation stärker abschwächen als in *B*. Dieser Prozess wirkt der anthropogenen Erwärmung in dieser Region entgegen und stellt einen eigenen Forschungs- und Risikokomplex in diesem Kontext dar (z. B. LEVERMANN et al. 2007, KUHLEBRODT et al. 2009).

Die Differenzen der Teilensembles in der Entwicklung der bodennahen Tem-

peratur und deren räumliche Schwerpunkte lassen sich insgesamt durchaus aus dem Gruppierungsansatz begründen. Die stärksten Effekte in den niederen Breiten finden sich für das Amazonasbecken und Indien während der Sommermonate, also Regionen die über bekannte Telekonnektionen mit ENSO verfügen. Die Effekte in den hohen Breiten können unter dem Stichwort der Arktischen Verstärkung als Resultat der höheren  $T_E$  in  $A$  betrachtet werden, die zumindest teilweise auf das ENSO-System zurückgeführt werden kann.

### 11.2.2 Effekte auf die Entwicklung des Niederschlags

Das räumliche Muster der Niederschlagsentwicklung ist in beiden Teilensembles auf globalem Maßstab sehr ähnlich. Allerdings existieren durchaus nicht nur quantitative Unterschiede in den absoluten Ausprägungen des Niederschlags sondern auch für bedeutende Regionen gegensätzliche Entwicklungen. In Abb. 11.2 sind absolute Werte für den Niederschlag angegeben, die hier getroffenen Aussagen sind aber auch für die Betrachtung von relativen Veränderungen dieser Größe gültig.

Das sowohl im Jahresmittel als auch in allen Jahreszeiten dominierende räumliche Muster lässt sich derart zusammenfassen, dass im SRES A1b für humide Regionen im 21. Jahrhundert eine Niederschlagserhöhung simuliert wird, während für die tropischen und insbesondere subtropischen Gebiete mit ariden bis subhumiden Niederschlagsregimen Niederschlagsabnahmen die regionalen Klimatologien kennzeichnen. Außertropische Trockengebiete sind hiervon nicht betroffen.

Die Unterschiede zwischen den Teilensembles sind beim Niederschlag über Land in weiten Teilen so geartet, dass  $E_R(s)$  negative Werte annimmt. Als Grundtendenz sind somit die negativen Trends in  $A$  stärker und positive schwächer. Diese Regel findet sich über den meisten Regionen Afrikas südlich der Sahara bestätigt. Die seltenen Ausnahmen erweisen sich durchweg als nicht signifikant. Für den subtropisch-tropischen Trockengürtel finden sich zwar einige signifikante Unterschiede zwischen  $A$  und  $B$ , diese sollten aber angesichts des geringen tatsächlichen Regens, den die GCMs realistischerweise für diese Region simulieren, als statistisches Artefakt betrachtet werden. Allerdings ist festzuhalten, dass in  $B$  ein durchaus nennenswerter Anstieg des jährlichen Niederschlags um etwa  $90 \text{ mm/m}^2$  für den Süden der arabischen Halbinsel existiert. Dies entspräche knapp einer Verdopplung des derzeitigen klimatologischen Werts.

Auch für Australien gilt die aufgezeigte Regel für die klimatologischen Veränderungen der Regenmenge für nahezu alle betrachteten Jahreszeiten. Signifikante Werte erreicht  $E_R(s)$  allerdings nicht in flächendeckendem Ausmaß und nur für einzelne Jahreszeiten. Auch hier verhindert ein ohnehin geringer Niederschlag quantitativ stabil erfassbare Veränderungen.

Europa weist in beiden Teilensembles eine divergierende Entwicklung im Niederschlag auf. Der Mittelmeerraum ist dabei durch sinkende Niederschläge gekennzeichnet, während für die nördlichen Gebiete positive Veränderungen im

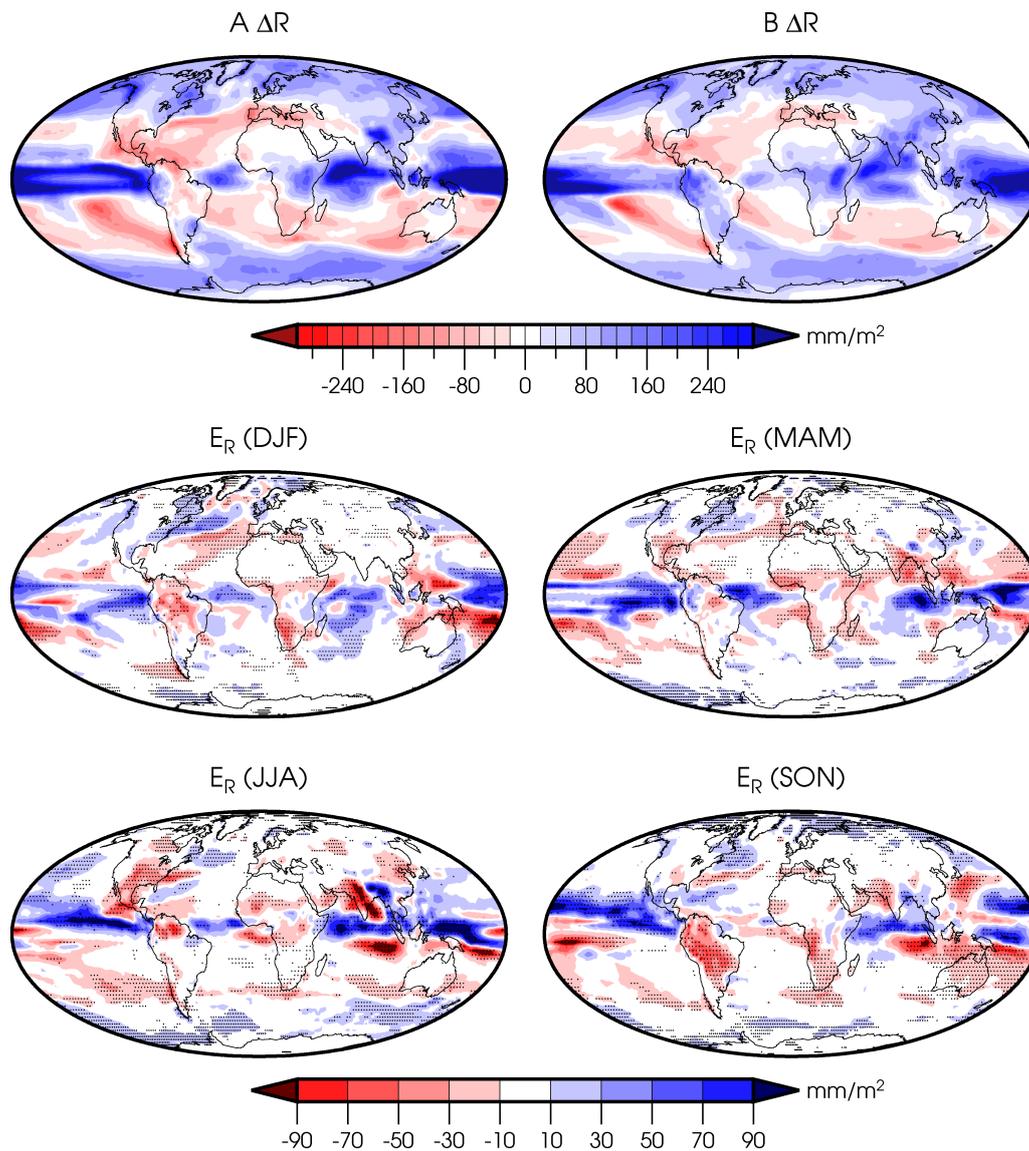
klimatologischen Mittelwert im SRES A1b überwiegen. Für Mitteleuropa sind die Vorzeichen der Entwicklung abhängig von der Jahreszeit. Hier stehen vermehrte Regenmengen während des Winters geringeren im Sommer gegenüber. Alle diese Aspekte der europäischen Niederschlagsentwicklung finden sich so in beiden Teilensembles, wobei *A* wiederum signifikant höheres Wachstum der Niederschläge im Norden und gleichzeitig ebenfalls signifikant stärkere Abnahmen im Süden des Kontinents als *B* anzeigt. Für weite Teile Europas sind die Unterschiede allerdings nur gering.

Das gilt auch für große Gebiete in Asien, insbesondere nördlich des Himalayas. Zwar finden sich hier zu verschiedenen Jahreszeiten durchaus signifikante Unterschiede, auch diese gelten jedoch nur für geringe Werte von  $E_R(s)$ . Dagegen scheinen die Effekte der Gruppierung über Indien und Südostasien durchaus auch rein mengenmäßig signifikant. Von besonderer Bedeutung ist hierbei die Niederschlagsentwicklung während der Monsunsaison über Indien. In *B* findet sich über den gesamten Kontinent eine starke Zunahme der Regenmenge in den entsprechenden Monaten. Im Teilensemble *A* ist dagegen nur die Südspitze Indiens durch steigende Niederschläge im Sommer gekennzeichnet. Für den Norden wird dagegen sogar ein leichter Rückgang des wichtigen Monsunniederschlags simuliert. Diese qualitativen Unterschiede führen zu signifikanten Werten für  $E_R(s)$  von über 100 mm im klimatologischen Mittel während der Monsunsaison.

Nordamerika zeigt ein ähnliches Nord-Süd-Charakteristikum wie Europa. Signifikant werden die Differenzen von *A* und *B* vorwiegend für den trockenen Mittleren Westen der USA sowie Mittelamerika. Im kanadischen Raum und speziell an der Pazifikküste findet sich eine große Übereinstimmung im gesamten SRES A1b.

Die ausgeprägtesten qualitativen Unterschiede zwischen *A* und *B* weist Südamerika auf. Von der Karibik über weite Teile des Amazonasbeckens weist *A* negative Niederschlagstrends auf. Für *B* finden sich hier dagegen zumindest in den zentralen Bereichen des Kontinents Erhöhungen der Regenmenge. Im jährlichen Mittelwert betragen  $E_R(s)$  bis zu 120 mm. Die Trockenheit im Norden des Kontinents ist eine bekannte Fernwirkung von ENSO, die mit El Niño-Ereignissen einhergeht, welche in *A* durch den Trend des SOI wahrscheinlicher und häufiger werden.

Über den Ozeanen ist das Grundmuster von  $E_R(s)$  einfach erfassbar. Die mittlere Entwicklung im CMIP3 ist in *A* etwas intensiver als in *B*. Damit sind die Niederschlagsabnahmen in den Subtropen in *A* ebenso viel stärker im Vergleich zu *B* wie der ansonsten verbreitete Anstieg der Regenmenge. Werden die Landmassen außen vor gelassen ist die Niederschlagsentwicklung im Wesentlichen durch die subtropischen Hochdruckgebiete strukturiert. Im Zuge derer Intensivierung und Ausdehnung findet sich für die für diese Gebilde typischen Regionen eine Niederschlagsreduzierung. Da in diesen Regionen die Niederschlagsentwicklung innerhalb des CMIP3 sehr ähnlich ist, finden sich dennoch signifikante Unterschiede zwischen den Teilensembles, wobei die Ausdehnung der betroffenen Gebiete im Jahresverlauf starken Schwankungen unterliegt. Dies ist im Wesentlichen aus der



**Abb. 11.2:** Räumliche Differenzierung von  $\Delta R$ . Wie Abb. 11.1, aber klimatologische Veränderung des Niederschlags in  $\text{mm}/\text{m}^2$ .

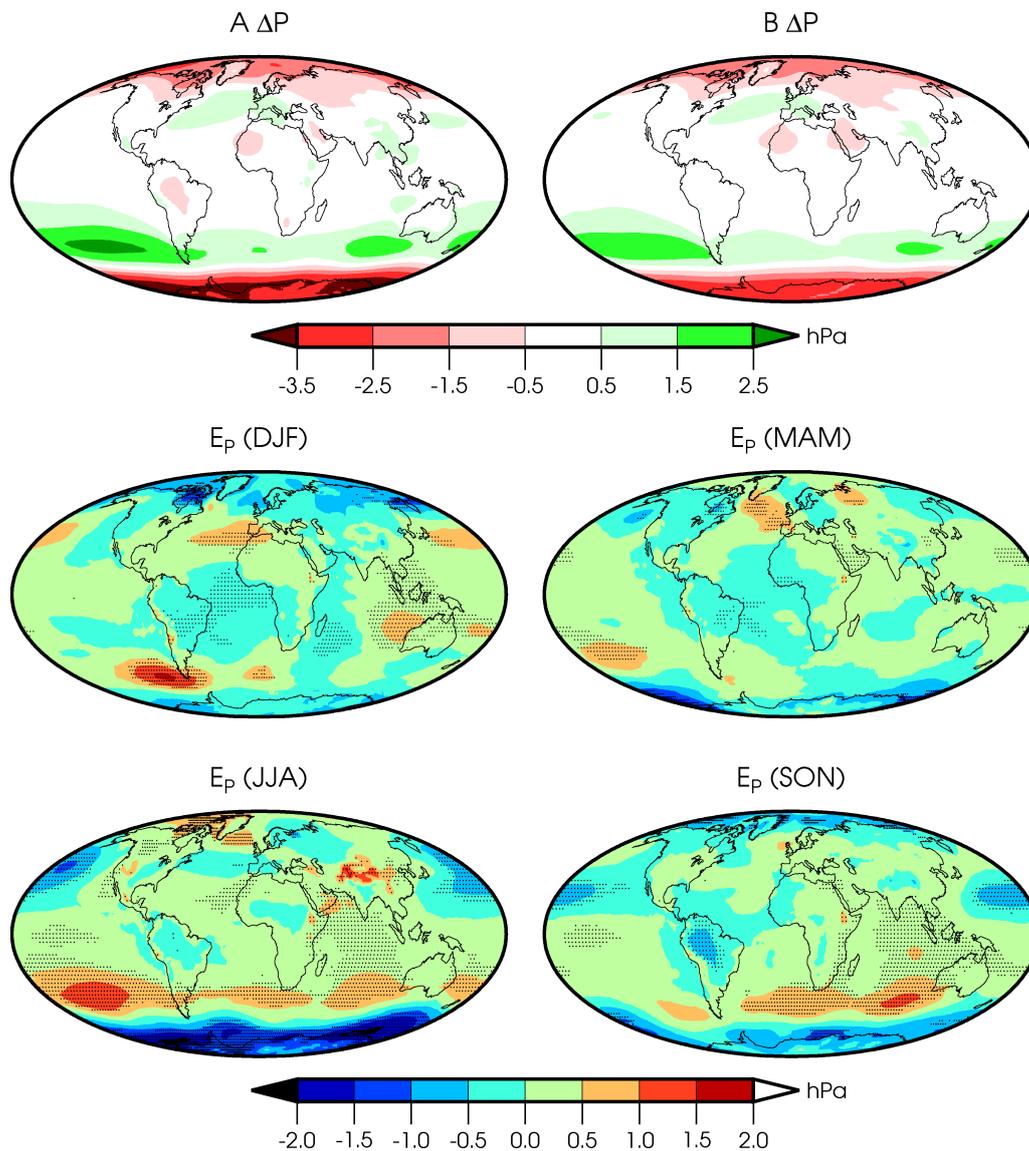
Tatsache heraus erklärbar, dass in den Sommermonaten die Niederschläge in diesen Regionen modellübergreifend realistischerweise so gering sind, dass eine Tendenz zur weiteren Verringerung des Regens keine signifikanten Auswirkungen mehr haben kann. Für den südlichen Ozean finden sich signifikante Werte für  $E_R(s)$ , was insbesondere auf stärkere Niederschlagserhöhungen in *A* während des Süd winters in dieser Regionen zurückgeht. Allein im Bereich der Antarktischen Halbinsel sind

diese Differenzen für alle Jahreszeiten signifikant. Dennoch sind auch in diesem Bereich des Antarktischen Ozeans deren absoluten Beträge nur gering. Im Nordpazifik sind nur für den Nordsommer signifikante Werte für  $E_R(s)$  vorhanden. Dies dürfte auf die unterschiedliche Entwicklung des NPI zu dieser Jahreszeit zurückzuführen sein, die für  $B$  die Randbedingungen für häufigere blockierende Hochdruckgebiete anzeigt. Für den Nordatlantik bedingt die stärkere Ausdehnung des Azorenhochs in  $A$  vor der Iberischen Halbinsel signifikante Werte für  $E_R(s)$  während des gesamten Jahres. Im Bereich des Islandtiefs ist die Niederschlagszunahme in  $A$  im Winter und Frühjahr etwas geringer als in  $B$ . Diese Unterschiede sind aber durchaus signifikant. Der Arktische Ozean ist dagegen weiträumig dadurch gekennzeichnet, dass  $A$  stärkere positive Trends aufweist als  $B$ . Besonders ausgeprägt und konstant im Jahresverlauf sind die signifikanten Werte von  $E_R(s)$  im Bereich der Barentssee, gekoppelt an  $E_T(s)$ . Im äquatorialen Pazifik und Indik findet sich in beiden Teilensembles eine starke Niederschlagszunahme im Bereich der ITC. In  $A$  ist diese für alle Jahreszeiten intensiver, allerdings ist die Zone mit deutlich erhöhten Regenmengen zum Ende des 21. Jahrhunderts in  $B$  breiter. Wegen des ausgeprägten Einflusses den die räumliche Auflösung der Niederschlagsdaten in diesem Bereich hat (vgl. Kapitel 4.1.1.2), müssen die signifikanten  $E_R(s)$  in diesem Bereich vorsichtig interpretiert werden. Dennoch kann, insbesondere da sich in den Differenzen der beiden Teilensembles im pazifischen Westen für den Nordwinter und das Jahresmittel durchaus das typische, mit El Niño verbundene Muster an Niederschlagsanomalien erkennen lässt, eine gewisse Bedeutung der Gruppierung für diese Unterschiede sicher angenommen werden.

### 11.2.3 Effekte auf die Luftdruckverteilung

Die klimatologischen Veränderungen der Luftdruckverteilung finden im Wesentlichen in den Außertropen statt. Auf beiden Hemisphären äußert sich die Entwicklung in der Abnahme des Luftdrucks über den Polarkalotten. Damit einher geht der simultane Anstieg des Drucks im Bereich der subtropischen Hochdruckgürtel. Somit ist in den CMIP3-Läufen als grundsätzliche Tendenz die Intensivierung der Gradienten zwischen den Subtropen und den polaren Regionen festzustellen. Dieses Muster der klimatologischen Entwicklung regt die dominierenden kreisförmigen Moden der Außertropen an. Die positiven Trends der Zeitreihen von SAM und NAM sind der Ausdruck dieser Entwicklung.

Während der jeweiligen Wintermonate sind diese, die Westwinddrift intensivierenden Verschiebungen, in  $A$  etwas ausgeprägter als in  $B$ . Auf der Nordhalbkugel ist sowohl der Luftdruckanstieg im Bereich des Azorenhochs als auch die -abnahme speziell im Bereich von Hudson Bay und Neufundland aber auch Kamtschatka in  $A$  signifikant deutlicher als in  $B$ . Auf der Südhalbkugel sind die  $E_P(s)$  zirkumpolar signifikant, obwohl beide Teilensembles für die Monate Juni bis August im Mittel sehr deutliche Veränderungen in der Luftdruckverteilung simulieren. Diese Un-



**Abb. 11.3:** Räumliche Differenzierung von  $\Delta P$ . Wie Abb. 11.1, aber klimatologische Veränderung im Luftdruck auf Meeressniveau in hPa.

terschiede sind während der sonstigen Jahreszeiten weniger stark ausgeprägt und auf der Nordhalbkugel ist auch das qualitative Muster nicht stabil. Im Frühjahr und Sommer ändern die  $E_p(s)$  hier ihre Vorzeichen. Somit ist das Bild inwieweit sich die Gruppierung auf die Luftdruckverteilung auswirkt hier weniger eindeutig, als auf der Südhalbkugel. So weist A zwar das ganze Jahr hindurch die stärkere Intensivierung des hemisphärischen Druckgradienten auf, das räumliche Muster

der Unterschiede, sowie deren Intensität, verändert sich im Jahresverlauf allerdings deutlich.

Im globalen Druckfeld zeigen sich also durchaus bedeutsame Muster des anthropogenen Forcings, die zudem im wesentlichen durch die beiden dominanten Zirkulationsmoden erfasst werden. Dagegen ist die räumliche Verteilung des Luftdrucks in den Tropen, bei verminderter Wirksamkeit der Corioliskraft und somit wenig stabilen Druckgradienten, als wenig sensitiv gegenüber dem anthropogenen Forcing zu bezeichnen. So sind schon die absoluten Beträge der Entwicklung während des 21. Jahrhunderts nur gering. Da zudem das grundsätzliche Muster der Veränderung des Luftdrucks im SRES A1b auch in den niederen Breiten sehr stabil ist, können die Unterschiede der Mittelwerte der Teilensembles ebenfalls kein besonders großes quantitatives Ausmaß annehmen. Durch die Stabilität dieser Größe auf klimatologischer Zeitskala in den Tropen werden dennoch durchaus großflächige Gebiete mit signifikanten Differenzen zwischen  $A$  und  $B$  durch die Analyse aufgezeigt. Insbesondere für den indischen Ozean mit Schwerpunkt in der Region des westaustralische Beckens und der indonesischen Inseln sind die positiven  $E_P(s)$  im Jahresmittel und in den meisten Jahreszeiten signifikant. Es kann sich hierbei um einen Aspekt der durch den SOI angezeigten unterschiedlichen langfristigen Entwicklung zwischen den Teilensembles handeln, jedoch muss jede Interpretation bei den geringen Beträgen der Differenzen und tatsächlichen Trends in den Modellen als schwierig bezeichnet werden.

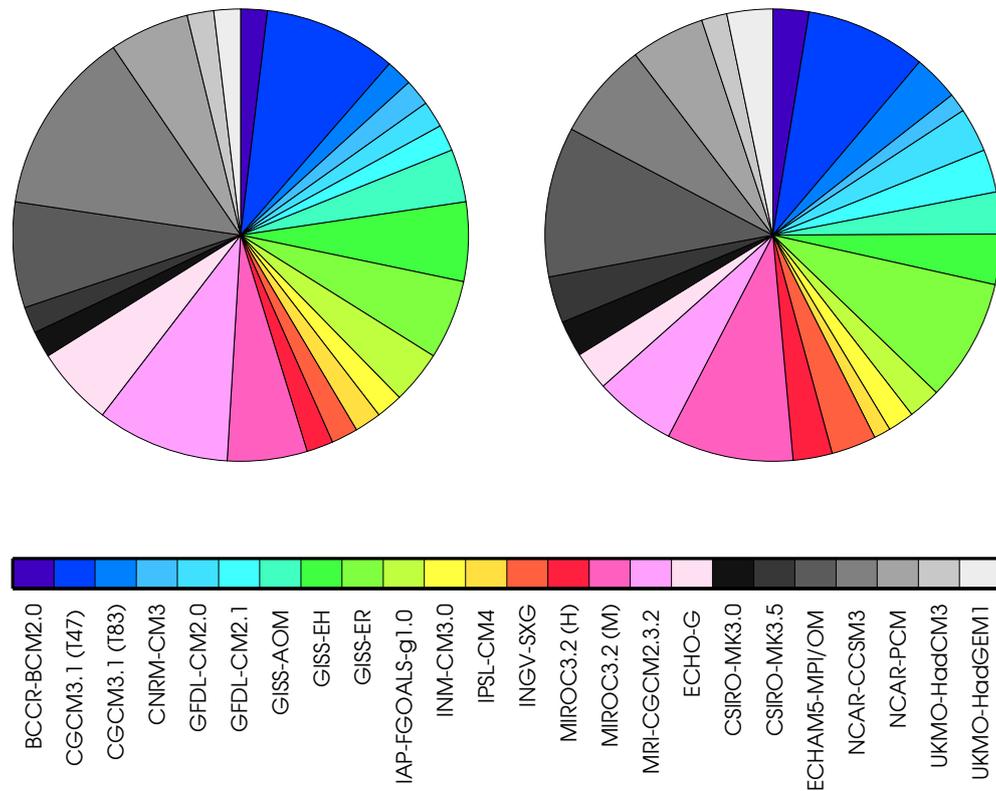
## 12 Gewichtete Ensemble-Projektionen

Neben dem ENSO-Phänomen existiert eine Vielzahl anderer, hier nicht angesprochener Unsicherheitsfaktoren innerhalb der Modelle. Diese betreffen zum einen, wie im hier adressierten Fall, die Fähigkeit der AOGCMs, bekannte Aspekte des Klimasystems zu reproduzieren, zum anderen aber auch technische und numerische Aspekte der Modellierung. Es ist davon auszugehen, dass ein bedeutender Teil dieser individuellen Modelldefizite unabhängig vom spezifischen ENSO-Mechanismus ist. Somit kann die einseitige Fokussierung auf diesen unter Umständen bewirken, dass die ausgleichende Wirkung eines Multi-Modell Ensembles in Bezug auf andere Fehler herabgesetzt wird. Das gilt insbesondere wenn ein spezielles Kriterium zur Verkleinerung der Ensembles herangezogen wird, mit dem neuen Ensemble aber allgemeine Fragestellungen in Angriff genommen werden. Daher sollten die Differenzen der beiden Teilensembles nicht als unterschiedliche Prognosen für den anthropogenen Klimawandel aufgefasst werden, wenn sich hier auch sicherlich auf vielfältige Weise Unterschiede zwischen  $A$  und  $B$  feststellen ließen. Hier wird hingegen untersucht, ob und wie sich die durch diese Arbeit gewonnenen Erkenntnisse auf die grundsätzlichen Tendenzen im Ensemble für den anthropogenen Klimawandel auswirken.

### 12.1 Methodik

Um die Ergebnisse der Überprüfung der ENSO-Performance in die Ensemble-Prognosen mit einzubeziehen, werden im Folgenden die Einzelläufe gemäß der aufgestellten Bewertungsfunktion gewichtet (vgl. Kapitel 7.2). Abb. 12.1 zeigt die Auswirkungen dieses Vorgehens. Dort sind die relativen Anteile, die ein Modell zur entsprechenden Kenngröße des Ensembles beiträgt, für den üblichen Fall der Gleichgewichtung und bei der Nutzung der Gewichtung aufgezeigt. Im Folgenden wird das SRES A1b, so seine Kenngrößen unter Berücksichtigung der Gewichtung berechnet wurden als  $G$  bezeichnet. Die übliche Version dieser Parameterschätzer basierend auf ungewichteten Gleichbehandlung sämtlicher Läufe wird  $U$  genannt.

In beiden Teilabbildungen sind, wie üblich bei der Betrachtung der Gesamtaussage eines Ensembles, sämtliche Realisationen pro Modell bei der Berechnung der Ensemble-Kenngrößen berücksichtigt. Durch die Kombination von jeweils relativ



**Abb. 12.1:** Einfluss pro AOGCM auf Ensemblekenngrößen. Bei Gleichgewichtung (links) und Verwendung der ENSO-basierten Gewichtungsfunktion (rechts)

hoher Laufanzahl und guter Bewertung des ENSO-Systems ist ECHAM5/MPI-OM im Ensemble  $G$ , statt NCAR-CCSM3 in  $U$ , das Modell mit dem größten Einfluss auf die Ensemble-Projektionen. Hier beschränkt sich die Analyse auf den Vergleich von Mittelwerten und Varianzen der Entwicklung der wichtigsten Kenngrößen des Klimasystems zwischen den beiden Versionen des Multi-Modell Ensembles. Jede signifikante Veränderung in diesen Größen bedeutet eine veränderte langfristige Prognose zum Klimawandel.

Unterschiede in den Mittelwerten bedeuten, dass sich die langfristige Projektion des Ensembles verändert, wenn die Qualität der Modelle bei der Kombination der individuellen Projektionen berücksichtigt wird. Somit können auf diese Weise Aspekte des Klimawandels erkannt werden, die mit der natürlichen Variabilität im besonderen Maße verbunden sind. Verringert sich  $s_g^2$  im Vergleich zu  $s_u^2$ , so bedeutet dies, dass die Reduzierung des Einflusses von weniger realitätsnahen AOGCMs dazu führt, dass sich die Ensemble-Projektionen stabilisiert. Die Interpretation hiervon ist, dass die Unterschiede in den Auswirkungen des vorgegebenen Forcings

entscheidend auf eine fehlerhafte Repräsentation des Klimasystems zurückzuführen sind. Dieser Faktor sollte durch ein realistisches ENSO-System abgemildert werden. Steigt dagegen die Varianz einer Größe im Ensemble durch die Gewichtung an, könnte dies darauf zurückzuführen sein, dass die Modelle mit den höheren Gewichten in der zeitlichen Entwicklung des ENSO-Systems große quantitative und sogar qualitative Unterschiede aufweisen. Auch in diesem Fall kann aber ein Erkenntnisgewinn aus den hier präsentierten Analysen gezogen werden: es können Klimagrößen oder Regionen identifiziert werden, für die die langfristigen Projektionen des Klimawandels durch eine bessere Repräsentation des ENSO-Phänomens und dessen zeitlicher Entwicklung verbessert werden kann.

Der Vergleich von  $G$  und  $U$  erfordert statistische Tests. Schon die Verwendung des Gewichtungsschemas schränkt die Zuverlässigkeit und Zulässigkeit der üblichen Verfahren hierbei ein. Dieser Faktor wird noch dadurch verstärkt, dass die beiden Versionen des SRES A1b sicherlich nicht als unabhängige Daten aufgefasst werden können. Um über die veränderte Verteilung der Ensemblegrößen keine Fehlerquelle beim Testentscheid zu generieren, wird dieser an die Situation angepasst. Dies geschieht für jede betrachtete Variable separat über die Anpassung der Freiheitsgrade. Hierzu werden für die gewichteten Kenngrößen von CMIP3 die Freiheitsgrade  $df_g$  nach folgender Formel, die sich an der Berücksichtigung von Autokorrelation bei Zeitreihen orientiert, reduziert

$$df_g = df_u \cdot \frac{1 - (\sum_{i=1}^n \omega(i)r_g(i) - \bar{r}_u)}{1 + (\sum_{i=1}^n \omega(i)r_g(i) - \bar{r}_u)} \quad (12.1)$$

In dieser Gleichung bezeichnet  $df_u$  die Freiheitsgrade bei Gleichgewichtung aller Läufe. Sie werden nach den üblichen Verfahren für den betroffenen Parameterschätzer und den Datenumfang  $n$  ermittelt. Die Korrelationskoeffizienten beschreiben die Korrelation zwischen den Einzelläufen und der gemittelten Zeitreihe der betreffenden Größe im Multi-Modell Ensemble. Mit  $r_g(i)$  wird die Korrelation des  $i$ -ten Einzellaufs mit der Zeitreihe der gewichteten Mittelwerte  $g(t)$ , mit  $r_u(i)$  die mit der Zeitreihe der arithmetische Mittelwerte  $u(t)$  bezeichnet. Somit wird auch berücksichtigt, dass auch bei den üblichen Verfahren zur Kombination von Modellläufen keine tatsächliche Unabhängigkeit der Daten existiert.

## 12.2 Ergebnisse

Hier werden die Anwendung der Gewichtung auf unterschiedliche Fragestellungen präsentiert, beginnend mit der Entwicklung der Klimaindices und  $T_E(t)$  im Zeitverlauf unter anthropogenen Forcing. Hierbei wird die simulierte Entwicklung dieser Größen in den Fokus genommen, und festgestellt inwieweit sich diese im gewichteten Ensemble von der im ungewichteten Ensemble unterscheidet. Das

Hauptaugenmerk der Darstellung liegt aber auf den effektiven Veränderungen, die die Größen zum Ende des 21. Jahrhunderts aufweisen. Bei dieser Betrachtung wird der kurzfristige Einfluss der internen Klimavariabilität der einzelnen Modellläufe weitgehend ausgeschaltet, um den tatsächlichen Effekt der Gewichtung auf die Quantifizierung des Klimawandels zu demonstrieren.

### 12.2.1 Zeitreihen: Die transiente Entwicklung

Hier wird der Verlauf und die Entwicklung der Ensemblemittel betrachtet. Das ungewichtete Mittel wird jeweils durch die Zeitreihe  $u(t)$ , das gewichtete durch  $g(t)$  repräsentiert. Wiederum wurden die Analysen sowohl direkt mit den Zeitreihen, als auch mit tiefpassgefilterten Versionen  $u_k(t)$  und  $g_k(t)$  durchgeführt. Die maximale Filterlänge  $k$  betrug jeweils 31 Jahre, in diesen Fällen bezeichnet der Index  $t$  das Zentraljahr. Welche Größe hierbei durch die Zeitreihen repräsentiert wird geht aus dem jeweiligen Kontext eindeutig hervor.

Abb. 12.1 verdeutlicht, dass die Gewichtung nicht zur Dominanz eines einzelnen Modells führt. Auch im gewichteten Mittel halten sich somit verschiedene Modelle die Waage, was bei eher individuellen Zeitreihenverläufen einzelner AOGCMs, auch wenn diesen ein relativ hohes Gewicht zugestanden wird, eine ausgleichende Wirkung auf den mittleren Zeitreihenverlauf hat. Dieses Verhalten hat sich schon bei der Betrachtung der beiden Teilensembles gezeigt (vgl. Kapitel 10). Im Sinne der Erzielung signifikanter Ergebnisse kommt an dieser Stelle natürlich erschwerend hinzu, dass sämtliche AOGCMs sowohl den Verlauf von  $u(t)$  als auch den von  $g(t)$  beeinflussen.

Für keine der betrachteten Zeitreihen lassen sich entsprechend signifikante Unterschiede im Verlauf feststellen. Zwar ließen sich vereinzelt signifikante Werte detektieren, diese wurden allerdings bei steigender Filterlänge vollständig beseitigt. Auf der dekadischen Zeitskala befindet sich  $g(t)$  für alle Indices und  $T_E$  gänzlich innerhalb des 95 %-Konfidenzintervalls von  $u(t)$ . Es ist erwähnenswert, dass dies auch gilt, wenn statt  $f_\omega(i)$  in Gleichung 7.2 deren 4. Potenz zur Bestimmung der Gewichtungsfaktoren herangezogen wird. Die mittlere Entwicklung der hier betrachteten Zeitreihen im CMIP3 kann also als unabhängig von der konkreten Methode der Kombination der Einzelläufe bezeichnet werden. Angesichts der unterschiedlichen langfristigen Verläufe des SOIs innerhalb des Teilensembles  $A$  und der ausgeprägten Variabilität dieser Größe in den Einzelläufen, ist dieses Ergebnis für das Multi-Modell Ensemble durchaus mit der grundsätzlichen Relevanz des Gewichtungsmerkmals vereinbar.

Dennoch lässt sich in einigen Fällen sehr wohl ein systematischer Einfluss der Gewichtung feststellen. Im Fall des SOI unterschätzt  $g(t)$  im 21. Jahrhundert  $u(t)$  während des borealen Winters konsequent (Abb.12.2). Somit verfestigt sich der Eindruck, dass die aus  $U$  abgeleitete indifferente Reaktion des SOI auf das anthropogene Forcing eine Folge der insgesamt nicht gleichmäßig hohen Qualität

der ENSO-Repräsentation darstellt. Für die außertropischen Moden zeigen sich dagegen nur geringe quantitative Abweichungen zwischen  $u(t)$  und  $g(t)$ , die sich zudem im zeitlichen Verlauf nahezu auf 0 aufsummieren.

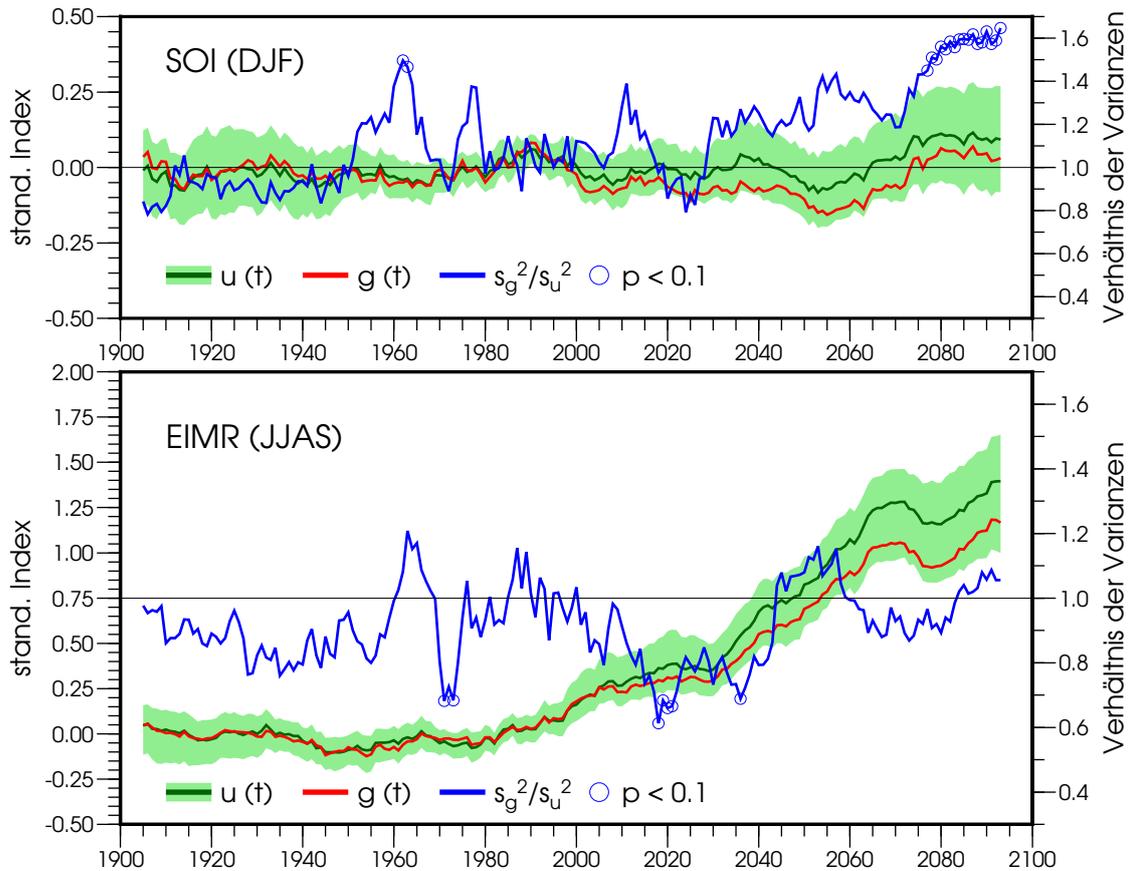
Den deutlichsten numerischen Unterschied weist der EIMR für die beiden Versionen der Ensemble-Mittelwertszeitreihe auf. Die Differenz von  $g(t)$  und  $u(t)$  ist hierbei eine klare Funktion der Zeit, d. h. hier kann eine Sensitivität auf das anthropogene Forcing unterstellt werden. Gegen Ende des 21. Jahrhunderts ist der projizierte Anstieg des EIMRs in  $G$  zwar ebenfalls ein stabiler Aspekt des Klimawandels, jedoch beträgt er nur etwa  $3/4$  des Anstiegs in  $U$ . Auch für den AMR sowie  $T_E$  weist die Analyse systematische, aus der Gewichtung, Veränderungen auf. Allerdings ist sowohl die konstant stärkere Erwärmung, die  $g(t)$  für  $T_E$  anzeigt, als auch der geringere Anstieg des westafrikanischen Monsunindexes auch quantitativ im Vergleich zur Gesamtentwicklung der Ensembledittel nur von geringer Bedeutung.

Die Variabilität der Zeitreihen stellt an dieser Stelle ebenfalls eine interessierende Größe dar. Die Beeinflussung der Varianz durch die Gewichtung erweist sich insgesamt als deutlicher als die der Mittelwerte, dafür lassen sich nur weniger klare qualitative Effekte feststellen. In Abb. 12.2 und Abb. 12.3 ist für eine Auswahl der bedeutsamsten Größen die Entwicklung des Quotienten  $s_g^2/s_u^2$  im Zeitverlauf dargestellt.

Sicherlich ist im Fall des SOI festzuhalten, dass dieser tendenziell eher eine Unterschätzung der Variabilität dieser Größe in  $U$  anzeigt. Dies gilt insbesondere ab etwa dem Jahr 2030. Für das Ende des 21. Jahrhundert erscheinen diese Unterschiede als signifikant zum Niveau  $\alpha = 0.1$ . Zwar ist das Ergebnis angesichts der bisher gezeigten Effekte der modellspezifischen ENSO-Präsentation nur wenig überraschend, die Tatsache, dass sich diese Unterschiede jedoch in signifikanten Effekten der Gewichtung niederschlagen soll aber dennoch ausdrücklich betont werden. Diese Entwicklung zum Ende des Betrachtungszeitraums ist unabhängig von der verwendeten Filterlänge.

Für den EIMR ist das Verhältnis der Varianzen weniger prägnant als für den SOI. Insgesamt kann festgehalten werden, dass  $s_g^2$  kleiner ist als  $s_u^2$ . Dies sollte eine Folge des geringeren Trends sein, der sich für  $G$  einstellt. Der zeitliche Verlauf des Quotienten kann zumindest qualitativ in Beziehung zum Verlauf der entsprechenden Größe des SOI gesetzt werden (vgl. Abb. 12.2). Dies würde einen Zusammenhang zwischen den beiden Größen nahelegen, infolgedessen ausgeprägte Variabilität in SOI und EIMR zeitlich korrelieren. Dieser Zusammenhang würde offensichtlich durch die Gewichtung betont, was wiederum auf die Unterschiede in der Wirksamkeit der vielfältigen Telekonnektionen des ENSO-Systems zurückgeführt werden könnte.

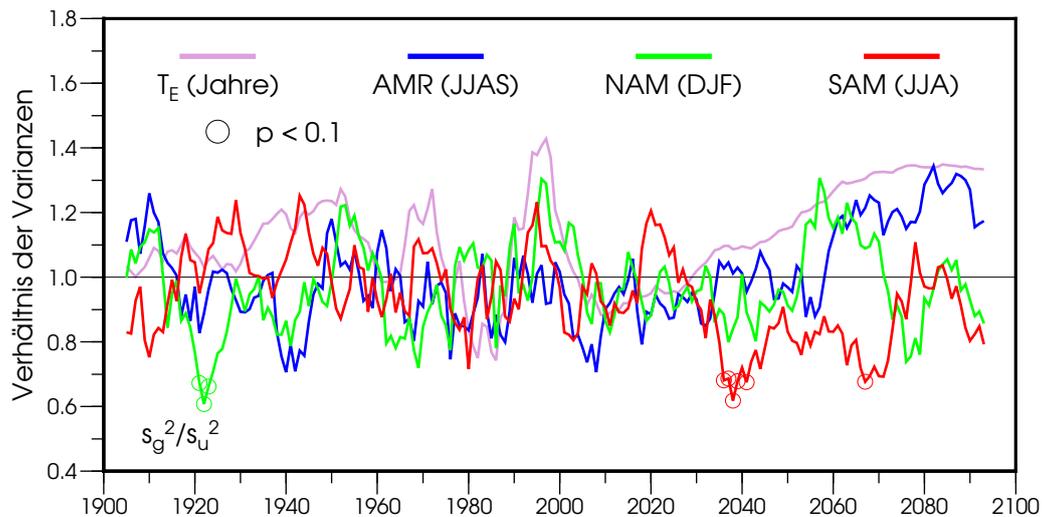
Der Varianzquotient der Verläufe der globalen Mitteltemperatur zeigt ebenfalls an, dass die Varianz dieser Größe von  $U$  scheinbar systematisch unterschätzt wird. Besonders auffällig ist dies ab etwa dem Jahr 2030. Ab diesem Zeitpunkt steigt die



**Abb. 12.2:** SOI und EIMR: Verlauf gewichtet/ungewichtet. Tiefpassgefilterte Zeitreihen, 11-jährige gleitende Mittel. Mit  $u(t)$  ist jeweils das 95 %-Konfidenzintervall abgebildet, das  $g(t)$  nie verlässt. Die blauen Kreise markieren Jahre, für die ein  $F$ -Test zum Niveau  $\alpha = 0.1$  signifikante Unterschiede zwischen den Varianzen ausweist. Unterschiedliche Skalierung,

Varianz von  $T_E$  in  $G$  im Vergleich zu  $s_u^2$  kontinuierlich an, bis sich das Verhältnis ab ca. 2070 auf hohem Niveau stabilisiert. Auch hier kann zeitliche Übereinstimmung mit dem Varianzquotienten des SOI unterstellt werden. Es soll jedoch betont werden, dass die Unterschiede der Varianzen im Fall von  $T_E$  nicht signifikant sind. Für den AMR zeigt Abb. 12.3 einen zunächst wenig systematischen, eher zufälligen Verlauf des Quotienten der Varianzen. Erst zum Ende des 21. Jahrhunderts nimmt auch in diesem Fall  $s_g^2$  relativ zu  $s_u^2$  zu. Eine Synchronizität mit den Varianzquotienten anderer Klimagrößen ist jedoch weniger deutlich auszumachen.

Die außertropischen Moden weisen wiederum keine eindeutige Reaktion auf die Gewichtung auf. Stattdessen oszilliert die Zeitreihe des Varianzquotienten um den Wert 1. Lediglich für den SAM während des Südwinters und -sommers ist



**Abb. 12.3:** Varianzverhältnisse einiger Klimagrößen. Erläuterungen siehe Text und Abb. 12.2

im 21. Jahrhundert eine längere Phase feststellbar, während der  $s_g^2$  systematisch, und vereinzelt sogar signifikant, kleiner ist als  $s_u^2$ . In dieser Periode finden sich allerdings in allen Läufen des CMIP3 ausgeprägte positive Trend des SAM, so dass sich hier sicherlich nur ein quantitativer Unterschied zwischen den Modellen aufzeigt, der die grundsätzliche Aussage des CMIP3, unabhängig ob aus  $U$  oder  $G$  gewonnen, nicht beeinträchtigt.

### 12.2.2 Zeitreihen: Der effektive Wandel

Nun werden die langfristigen effektiven Veränderungen untersucht, die für die globale Mitteltemperatur und die Indexreihe der Klimamuster im Mittel erzeugt werden. Um die mittlere Veränderung dieser Größen in CMIP3, die sich durch die Gewichtung ergibt, ohne die interne Variabilität der Modelle beurteilen zu können, wird an dieser Stelle wieder die klimatische Entwicklung in den Zeitreihen durch die Differenzen der Mittelwerte für die Perioden 2080–2098 und 1960–1999 operationalisiert. Das Mittel dieser klimatologischen Verschiebungen wird im Folgenden als die Grundaussage des Multi-Modell Ensembles interpretiert.

Für jede der Größen liegen damit  $n$   $\Delta X(i)$  vor. Anhand dieser Einzelwerte wird die mittlere Verschiebung jeweils als arithmetisches und gewichtetes Mittel berechnet. Der Mittelwert wird als der wahrscheinlichste Wert interpretiert, den das Ensemble für die Realität simuliert. Somit stellt dieser dann die langfristige Projektion für die betreffende Klimagröße dar. Desweiteren wird aus den  $\Delta X(i)$  die Unsicherheit dieser Projektion abgeleitet. Diese erlaubt dann die Aussage des

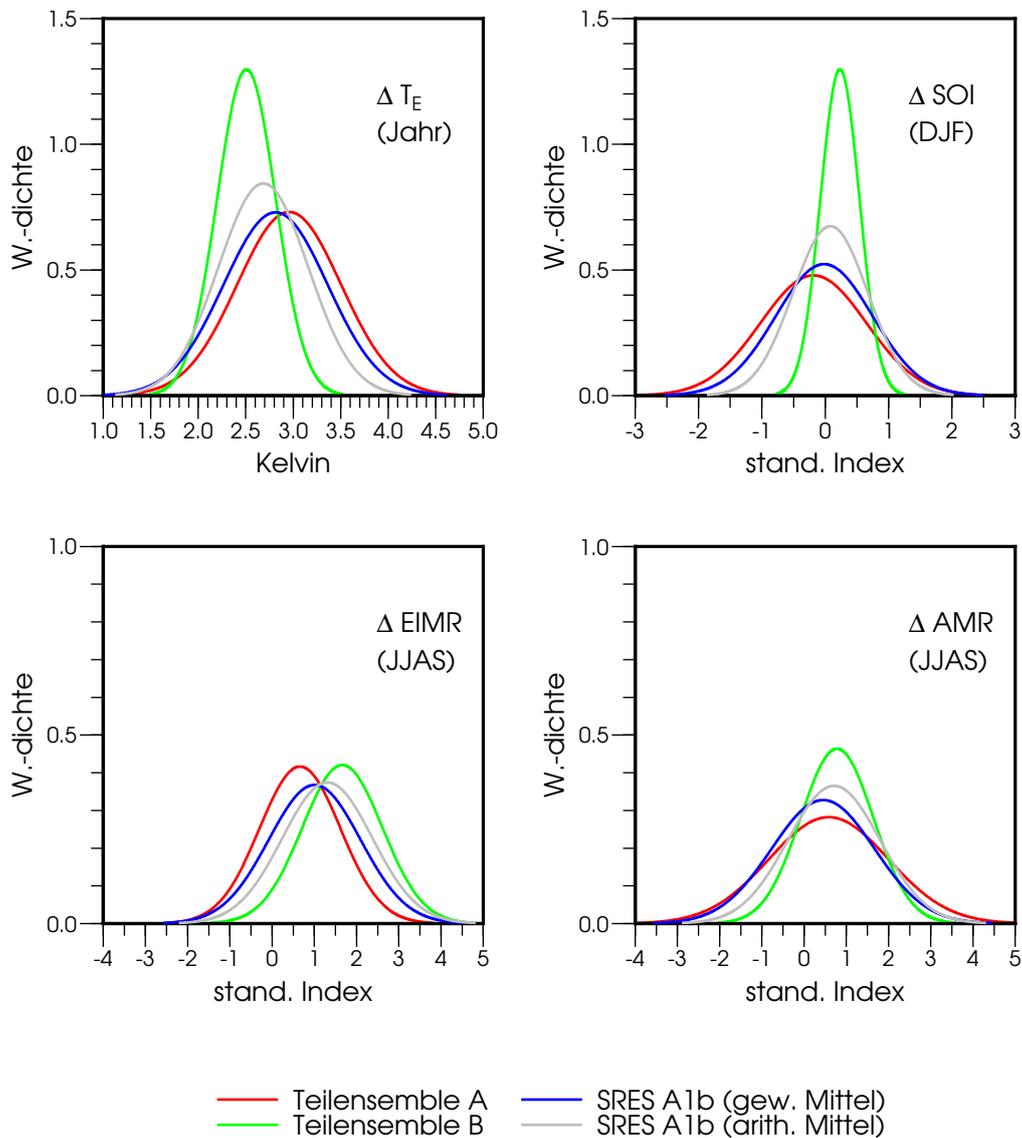
CMIP3-Ensembles dahingehend zu beurteilen, ob sie als im Wesentlichen stabil betrachtet werden kann. Mit dieser soll die tatsächliche Eintrittswahrscheinlichkeit abgeschätzt werden. Hierfür wird angenommen, dass die Veränderung dieser Größen einer Normalverteilung folgt. Angesichts der realisierten Werte von  $\Delta X(i)$  scheint dies durchaus angemessen, wenn auch vor allem im Fall des SOI durch die bisher präsentierten Ergebnisse auch andere Verteilungsmodelle gerechtfertigt erscheinen.

In Abb. 12.4 sind die Wahrscheinlichkeitsdichtefunktionen jener Größen dargestellt, die sich als durch die Gewichtung am stärksten beeinflusst herausgestellt haben. Es handelt sich um fundamentale Klimagrößen: die globale Mitteltemperatur, den SOI während des borealen Winters und die beiden Monsune. Hier wird deutlich, dass sich die langfristige Klimaentwicklung der Modelle vor allem in den Tropen unterscheidet. Dass dies durch die hier gewählte Darstellung herausgearbeitet wird lässt auch den Schluss zu, dass diese Unterschiede zumindest in Teilen auf das modellspezifische ENSO-System zurückgeführt werden können. Um diesen Einfluss zu verdeutlichen, sind zusätzlich die PDFs der Teilensembles in Abb. 12.4 zu finden. Bei allen vier dargestellten Größen lässt sich die geringste Variabilität der  $\Delta X(i)$  im Teilensemble  $B$  feststellen. Dieses weist also insgesamt die homogensten Prognosen für den langfristigen Klimawandel auf. Sicherlich besonders augenfällig ist dies für den SOI und  $T_E$ . Der hier bestehende kausale Zusammenhang konnte in Kapitel 8 aufgezeigt werden. Sowohl die Form der PDFs als auch ihre erkennbar synchrone Transformation je nach verwendeter Datenbasis ( $A$ ,  $B$ ,  $G$  oder  $U$ ) visualisieren diesen Befund eindeutig.

Von besonderer Wichtigkeit ist jedoch die Feststellung, dass für den SOI und  $T_E$  nicht nur zwischen den Teilensembles, sondern auch zwischen  $G$  und  $U$  signifikante Unterschiede zum Niveau  $\alpha = 0.05$  im Bezug auf die Varianz der Prognosen für beide Größen bestehen. In beiden Fällen steigt diese bei der Anwendung der Gewichtung an. Somit kann die Berücksichtigung der Qualität des ENSO-Systems die zu erwartende langfristige Entwicklung nicht konkretisieren, weder für dieses selbst noch, daraus intuitiv wie deduktiv folgerichtig, für die hiervon beeinflusste globale Mitteltemperatur. Tatsächlich ist das Gegenteil der Fall. Die Bedeutung der internen Variabilität als modifizierender Faktor für die Zukunft des Klimasystems wird durch die hier gezeigten Ergebnisse betont. So wird auch der Nutzen der weiteren Verbesserung der AOGCMs an einem konkreten Beispiel hervorgehoben. Um die fundamentale Größe  $T_E$  auch in ihrer langfristigen Entwicklung zuverlässig abbilden zu können, ist es auch notwendig, die Entwicklung des ENSO-System zu simulieren. Damit dies allerdings überzeugend gelingen kann müssen die Prozesse, die ENSO auszeichnen und definieren, durch einen Großteil der AOGCMs besser als im vorliegenden CMIP3 abgebildet werden. Das Teilensemble  $A$  hat zwar eine grundsätzlich andere mittlere Entwicklung des SOI als  $B$ , jedoch findet sich gerade in  $A$  kein Konsens für die langfristige Entwicklung dieser Zeitreihe, wie aus Abb. 12.4 hervorgeht und im Verlauf der Arbeit mehrfach gezeigt werden konnte. In dieser Darstellung wird für den SOI zudem offensichtlich, dass auch die

Grundtendenz im CMIP3 angesichts der breiten Wahrscheinlichkeitsdichtefunktion dieser Größe keinesfalls als gesichert betrachtet werden kann.

Für den EIMR lässt sich anhand der hier gezeigten PDFs erkennen, dass die Variabilität der  $\Delta X(I)$  in sämtlichen betrachteten Ensembles vergleichbar ist.



**Abb. 12.4:** PDFs Klimaänderungen:  $T_E$ , SOI, EIMR, AMR. Verschiebungen für gewichtetes und ungewichtetes CMIP3 und für die beiden Teilensembles. Die Entwicklung aller Größen wird als normalverteilt angenommen. Die Skalierung ist in den in den Teilabbildungen nicht einheitlich

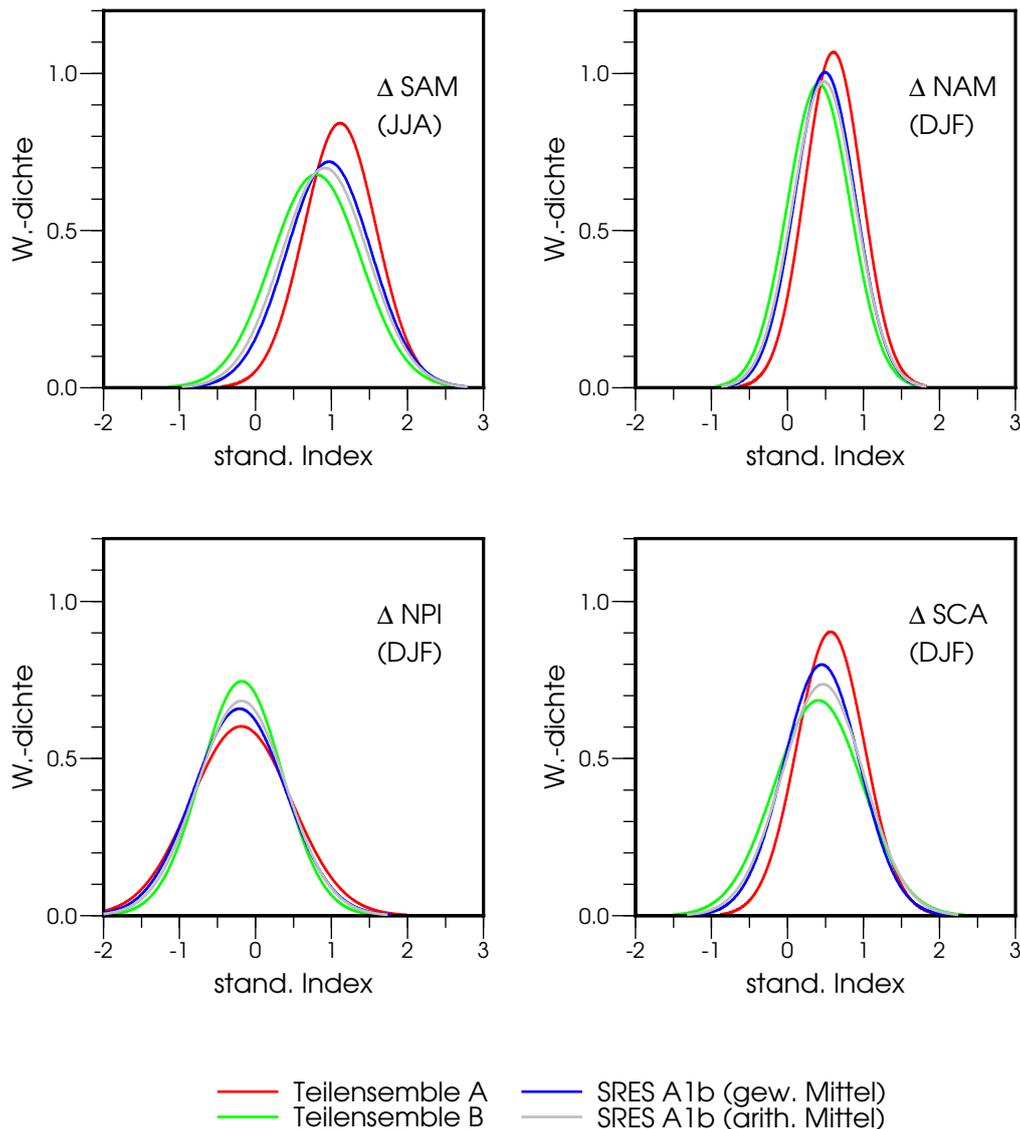
Somit ist der größere Unsicherheitsbereich, der sich in Abb. 10.4 für das Teilensemble  $A$  ergeben hat, auf interne Variabilität der Einzelläufe zurückzuführen. Auch das könnte über die modellspezifischen Telekonnektionen des ENSO-Systems erklärbar sein. Die PDFs von  $G$  und  $U$  lassen einen Unterschied im Mittelwert der Verschiebungen erkennen, der geringere Wert bei  $G$  ist allerdings nicht statistisch signifikant. Die Mittelwerte der Teilensembles sind dagegen signifikant ( $\alpha = 0.05$ ) sowohl voneinander als auch jeweils von den Mittelwerten der beiden Versionen des kompletten SRES A1b verschieden. Es ist abschließend für den EIMR festzuhalten, dass auch die Prognose dieser Größe sehr wahrscheinlich von einer verbesserten Qualität des ENSO-Systems profitieren würde.

Der AMR dagegen hat für keine der vier dargestellten PDFs signifikante Differenzen im Mittelwert. Diese scheinbare Stabilität ist jedoch im Hinblick auf die in allen Fällen recht hohe Varianz der  $\Delta X(i)$  zu relativieren. Die Prognose für die mittlere Erhöhung der Indexwerte im 21. Jahrhundert ist in  $U$  dennoch deutlich höher als in  $G$ . Allerdings sind die Varianzen der  $\Delta X(i)$  für  $G$  und  $U$  signifikant voneinander verschieden. Die Gewichtung identifiziert also auch für dieses vulnerable Klimaphänomen eine im CMIP3 vorhandene Unsicherheitsquelle. Die Unsicherheit in der Entwicklung dieses Klimaindex erscheint allerdings auch ohne die Berücksichtigung des durch das ENSO-Phänomen geleisteten Beitrags derart ausgeprägt, dass eine effektive Adaption auf der Grundlage der CMIP3-Simulationen tatsächlich nicht möglich erscheint.

In Abb. 12.5 sind die PDFs der Verschiebungen der Indexwerte für die außertropischen Zirkulationsmoden dargestellt. Hier werden nur die jeweiligen Wintermonate betrachtet. Diese weisen zwar nicht immer den ausgeprägtesten Trend auf, jedoch ist der Einfluss der Moden auf das Witterungsgeschehen zu dieser Jahreszeit am ausgeprägtesten. Weiterhin erklärt sich die Beschränkung auf den Winter dadurch, dass für diese Jahreszeit die Effekte der Gewichtung am deutlichsten zum tragen kommen. Dies gilt allerdings vor dem Hintergrund eines insgesamt nur äußerst geringen Effekts. Die Außertropen und die Abbildung und Entwicklung ihrer dominanten Klimamoden scheinen vom ENSO-Effekt insgesamt nur in geringem Maße betroffen zu sein. Dies gilt selbst in dieser sehr konzentrierten Darstellung, wenn auch hierbei einige Unterschiede deutlicher herausgearbeitet werden als bei der direkten Analyse der Zeitreihen und ihrer Verläufe. Insbesondere ist die Varianz der  $\Delta X(i)$  für die außertropischen Zirkulationsmoden stabil, was angesichts der hohen Variabilität, die der theoretisch steuernde SOI im Hinblick auf seine Entwicklung im CMIP3 an den Tag legt, einen großen Einfluss seinerseits als unwahrscheinlich erscheinen lässt.

Für den NPI sind weder für die Mittelwerte noch für die Varianzen statistisch signifikante Unterschiede zwischen  $G$  und  $U$  oder den Teilensembles feststellbar. Auch die optische Kontrolle lässt nur leichte Abweichungen der PDFs, beruhend auf numerischen Unterschieden der Standardabweichungen, erkennen. Sicherlich wird jedoch die grundsätzliche Entwicklung des CMIP3 für den NPI nicht durch

die Berücksichtigung des SOI verändert. Im Fall der Sommersaison lassen sich zwischen  $A$  und  $B$  signifikante Unterschiede in den Mittelwerten, nicht jedoch in der Varianz feststellen. Die aufgezeigte deutlich größere Varianz der Mittelwertzeitreihe von  $A$  ist somit auf interne Variabilität zurückzuführen, die eventuell durch die unterschiedlichen Telekonnektionen von ENSO hervorgerufen wird. Die Schätzer für die Parameter von  $G$  und  $U$  sind allerdings nicht statistisch signifikant zu unterscheiden.



**Abb. 12.5:** PDFs Klimaänderungen: SAM, NAM, NPI, SCA. Wie Abb. 12.4 aber für außertropische Zirkulationsmoden.

Ebenso liefern für SAM, NAM und SCA beide Versionen der CMIP3-Projektion für die Verschiebung der Indexwerte keine belastbaren Unterschiede. Das gilt sowohl für Varianzen als auch für Mittelwerte und sowohl für die gezeigte Indexentwicklung im Winter als auch für die während der anderen Jahreszeiten. Dieses ergebnisarme Ergebnis zeigt jedoch, dass die dokumentierten Effekte in den tropischen Moden und der globalen Temperatur tatsächlich auf dem unterstellten Unterscheidungsmerkmal beruhen und nicht einfach das Resultat einer beliebigen Einteilung der Modelle sind. Die Effekte von ENSO sind für die Außertropen einfach weniger stark ausgeprägt und stabil als die im Fall der Monsune oder der globalen Mitteltemperatur. Somit lässt sich hierbei auch festhalten, dass die CMIP3-Prognosen für diese Indices vermutlich nur in geringem Maße durch die unterschiedliche Qualität des ENSO-Systems beeinflusst werden. Das dies dennoch einen gewissen Ausschlag geben kann wird in Abb. 12.5 allerdings auch erkennbar.

Die Mittelwerte der  $\Delta X(i)$  unterscheiden sich zwar nicht für  $G$  und  $U$ , jedoch ist für die Wintermonate diese Größe bei  $A$  von den drei andern Ensembles signifikant unterschieden. Die Modelle dieses Teilensembles weisen im Mittel eine stärkere Verschiebung hin zu positiven Werten auf. Für den SAM gilt das gleiche während der Sommermonate, wobei hier im Fall des NAM jedoch  $B$  die stärkeren positiven Trends besitzt. Da diese Effekte allerdings zumindest im Fall des SAMs durch die Entfernung jeweils nur eines Modells aus den Teilensembles verschwinden, werden sie eher als Artefakte der spezifischen Modellphysik angesehen. Damit könnten sie natürlich weiterhin auf dem modellspezifischen ENSO-System beruhen, andere Quellen hierfür sind allerdings ebenso wenig auszuschließen.

### 12.2.3 Regionaler Wandel

In diesem Kapitel werden die Effekte der Gewichtung auf das räumliche Muster des Klimawandels untersucht. Hierfür werden einerseits, wie in Kapitel 11.1, die Differenzen der simulierten mittleren Klimaänderung im 21. Jahrhundert betrachtet. Entsprechend der bisherigen Nomenklatur werden die Mittelwerte  $\bar{g}_{\Delta X}(s)$  und  $\bar{u}_{\Delta X}(s)$  der Änderung der Klimavariablen  $X$  pro Gitterpunkt  $s$  für das gewichtete und das ungewichtete Multi-Modell Ensemble miteinander verglichen. Die Ergebnisse dieser Untersuchung zeigen, dass die Gewichtung die langfristige Änderung der drei grundlegenden klimatischen Feldgrößen, die sich aus der Kombination der CMIP3 Einzelläufe ergibt, nur in geringem Maße beeinflusst. Signifikante Unterschiede lassen sich nur an wenigen, vereinzelt Gitterpunkten feststellen, was sich teilweise durchaus durch die beschriebene Anpassung der Freiheitsgrade erklären lässt. Die resultierenden Muster sowohl für Niederschlag, Luftdruck als auch die Temperatur besitzen, wie zu erwarten war, große Ähnlichkeit mit denen, die die Effekte  $E_X(s)$  aufweisen (vgl. Kapitel 11). Tab. 12.1 zeigt die räumlichen Korrelationen zwischen diesen beiden Methoden, in den langfristigen Projektionen zum Klimawandel die Performance der AOGCMs mit einzubeziehen.

**Tab. 12.1:** Räumlicher Effekt von Gewichtung und Ensembletrennung. Korrelationskoeffizienten für die räumlichen Muster der Differenzen von  $G$  und  $U$  sowie  $A$  und  $B$ .

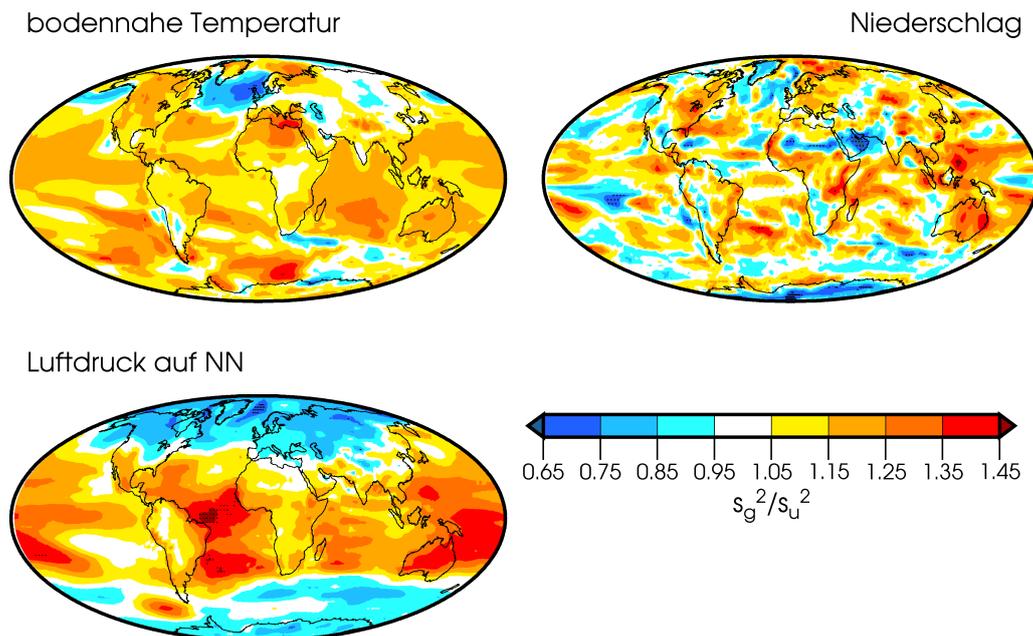
	Jahre	DJF	MAM	JJA	SON
$\Delta T$	.93	.93	.84	.91	.82
$\Delta R$	.87	.74	.67	.73	.76
$\Delta P$	.89	.79	.70	.86	.75

Allerdings sind die absoluten Beträge der Differenzen  $\bar{g}_{\Delta T}(s) - \bar{u}_{\Delta T}(s)$  im Vergleich zu den Effektgrößen, die bei der Betrachtung der Teilensembles resultieren, im Fall der Temperatur um rund den Faktor 5 geringer. Im Fall des Luftdrucks sind die quantitativen Auswirkungen der Gewichtung sogar noch weiter abgeschwächt. Der Niederschlag ist weniger regelmäßig in dieser Hinsicht. An einigen Orten finden sich im gewichteten Mittel der klimatologischen Veränderung jedoch sehr wohl quantitativ bedeutsame Veränderungen. Die Amplitude der Differenzen von  $G$  und  $U$  ist in diesen Fällen etwa halb so groß wie die der Effekte  $E_P(s)$ . Diese Werte sind dann im Wesentlichen auch statistisch signifikant. Eine Region mit solcherart deutlichen Auswirkungen der Verwendung des gewichteten Mittels findet sich im Nordatlantik. Vor Neufundland weist  $G$  signifikant stärkere Erhöhungen der Niederschläge auf. Diese Beobachtung findet sich in sämtlichen Jahreszeiten. Ansonsten sind erwähnenswerte quantitative Effekte der Gewichtung auf den tropischen Pazifik beschränkt. Hierbei halten sich aber Regionen mit positiven Abweichungen von  $G$  und solche mit negativen die Waage. Es liegt damit die Vermutung nahe, dass diese Effekte, die am markantesten während des Nordsommers auftreten, auf unterschiedlichen klimatologischen Lagen der innertropischen Konvergenzzone beruhen. Der Haupteinfluss dürfte in der Originalauflösung der AOGCMs begründet liegen, ein sachlogischer Bezug zum Wichtungskriterium lässt sich allerdings kaum herstellen.

Somit erweist sich die zu erwartende mittlere Veränderung in den drei Feldgrößen für den weitaus größten Teil der Erdoberfläche als stabil gegenüber der neuen Gewichtung des Ensembles. Auch bei regional differenzierter Betrachtung sind die auf Basis des CMIP3 im Hinblick auf diesen schematisierten Klimawandel zu treffenden Aussagen nicht empfindlich gegenüber der quantitativen Veränderung des Einflusses der unterschiedlichen AOGCMs. Anhand des Vergleiches der Varianzen der mittleren Veränderungen, die hier vereinfacht mit  $s_g^2(s)$  und  $s_u^2(s)$  bezeichnet werden, wird nun geprüft, ob sich über dieses Vorgehen stattdessen die Unsicherheiten für diese Klimaänderungen verändern. Abb. 12.6 zeigt für die drei Feldgrößen das Verhältnis der beiden Varianzen. Für jede der Variablen lassen sich regionale Cluster der Effekte der Gewichtung auf die Varianz der betrachteten Größen feststellen,

die sich als im Wesentlichen stabil für sämtliche Jahreszeiten erweisen. Aus diesem Grund ist eine Beschränkung auf die Darstellung der Verhältnisse im Jahresmittel angebracht. Auch in dieser Analyse zeigt sich, dass signifikante Ergebnisse kaum über den hier verfolgten Ansatz der Gewichtung erzielt werden können. Nichtsdestotrotz sind die Unterschiede zwischen den verschiedenen Versionen von Varianzen an einem Gitterpunkt zum Teil so groß, dass ihre vorsichtige Interpretation auch ohne signifikantes Ergebnis der  $F$ -Tests möglich erscheint. Dies gilt insbesondere dann, wenn stabile Quotienten  $s_g^2(s)/s_u^2(s)$  für sehr große Gebiete resultieren. Für die Veränderungen in der bodennahen Temperatur lässt sich eine markante räumliche Struktur dieser Größe erkennen. Abb. 12.6 und auch die Betrachtung der einzelnen Jahreszeiten identifiziert über den außertropischen Ozeanen eine Verringerung der Streuung der Erwärmung. Dies gilt für den Nordatlantik und den Nordpazifik. Auf der Südhalbkugel beschränken sich diese Effekte auf den außertropischen Bereich des Indischen Ozeans. Dieser eigentlich angestrebte Effekt der Konkretisierung der mittleren Klimaveränderung findet sich im Fall der Temperatur nahezu ausschließlich für diese drei ozeanischen Großräume. Ansonsten ist die Berücksichtigung der Qualität des ENSO-System bei der Mittelung der Einzelläufe gleichbedeutend mit der Erhöhung der Streuung. Die Gleichgewichtung der AOGCMs führt somit für einen Großteil der Erdoberfläche zu stabilen Simulationen des Klimawandels.

Der Natur des Niederschlags entsprechend sind auch die Effekte der Gewichtung auf die Grundtendenz der Niederschlagsentwicklung von hoher räumlicher Variabilität geprägt. Diese Effekte sind zudem bei dieser Größe immer vor der enorm hohen Variabilität zu betrachten, die sich auch für die Werte von  $s_u^2(s)$  ergeben. Die jeweiligen  $F$ -Tests liefern, mit wenigen Ausnahmen, keinerlei signifikante Ergebnisse, so dass an dieser Stelle auch nur einige, im Jahresverlauf stabile Ergebnisse aufgezeigt werden. So ist  $s_g^2(s)$  für weite Teile des Nordatlantiks geringer als  $s_u^2(s)$ . Für die humiden Landregionen nimmt die Varianz dagegen zu. Somit zeigt die Gewichtung auch beim tatsächlich vom Menschen nutzbaren Niederschlag, einer für den regionalen Klimawandel sicherlich bedeutsamen Größe, eine ernstzunehmende Problematik auf. Die langfristigen Veränderungen, die sich aus dem CMIP3 ableiten lassen, sind von größerer Unsicherheit betroffen als sich aus neutraler Betrachtung der Daten ergibt. Folgen hiervon ergeben sich auch für die Adaption an den Klimawandel, die dadurch als noch schwieriger zu gestalten erscheint. Allerdings muss bei diesem Punkt angemerkt werden, dass gerade für die dichtbevölkerten, agrarisch gut nutzbaren Landgebiete auch die Streuung der klimatischen Veränderung in  $U$  derartig bedeutend ist, dass kaum von tatsächlich stabil erfassbaren Grundtendenzen im regionalen Niederschlag gesprochen werden kann. Der Einfluss des Gewichtungskriteriums ist in Abb. 12.6 gut erkennbar. Im tropischen Westpazifik nimmt  $s_g^2(s)/s_u^2(s)$  nicht nur hohe, sondern auch signifikante Werte an. Diese finden sich im Jahresmittel und während des borealen Winters. Sie sind sicherlich mit den unterschiedlichen Langzeitrends im SOI verbunden. Folglich zeigt sich hier, dass die Modelle mit den höheren Gewichten tatsächlich auch im



**Abb. 12.6:** Gewichtete Streuung der Ensembleprojektionen. Verhältnis gewichteter und ungewichteter Varianzen der simulierten Klimaänderung im 21. Jahrhundert. Werte  $< 1$  bedeuten Konkretisierung der langfristigen Projektion durch die Gewichtung. Wenn  $s_g^2/s_u^2 > 1$  unterschätzt das Standardverfahren die Schwankung der Klimasimulation. Punktiierte Bereiche zeigen Signifikanz zu  $\alpha = 0.1$ .

Muster des Niederschlags ein dynamisches ENSO-System abbilden. Die Trends bilden sich in  $U$  angesichts der reduzierten Dynamik vieler Läufe weniger dominant ab.

Für den Luftdruck auf Meeressniveau findet sich ein eindeutiges räumliches Muster. Der Quotient der Varianzen trennt die hohen und gemäßigten Breiten auf der einen Seite und die Tropen und Subtropen auf der anderen Seite. Die Außertropen verlieren durch die Gewichtung an Varianz, bei den niederen Breiten dagegen wird durchgehend eine höhere Variabilität in den klimatischen Veränderungen festgestellt. Diese systematischen Raumeffekte sind im Wesentlichen zu allen Jahreszeiten feststellbar, wobei die Varianzreduktion in den Außertropen während der jeweiligen Wintermonate der Halbkugel am deutlichsten ausfällt. Die Quotienten in den Tropen und Subtropen zeigen dagegen keinen auffälligen Jahresgang. Die deutlichste Erhöhung der Varianz im Vergleich zu  $U$  weist  $G$  im Bereich des tropischen Atlantiks auf. Ein zweites Zentrum findet sich im Westpazifik.

## 13 Diskussion und Ausblick

Die vorliegende Arbeit hat, trotz ihrer begrenzten Seitenzahl doch einige Ergebnisse inhaltlicher Art hervorgebracht. Diese werden hier in der gebotenen Kürze zusammenfassend dargelegt und ihre Bedeutung diskutiert. Im Anschluss wird ein knapper Ausblick auf wichtige, noch ausstehende Forschungsschritte im Kontext dieser Arbeit gegeben. Hierbei wird insbesondere auf die durch den gewählten Ansatz zurückzuführenden Problematiken und Einschränkungen eingegangen, und ergänzende Vorgehensweisen werden aufgezeigt.

Die Durchführung des klassischen Hindcastings mit einigen neuartigen Anpassungen in Kapitel 4 konnte durchaus Erkenntnisse bringen. Zunächst ist die teilweise sehr große Streuung zwischen den Beobachtungen zu betonen, die sich sogar in langjährigen Klimatologien deutlich niederschlägt. Diese Unsicherheit über das reale Klimasystem reduziert den Wert des Hindcastings zur Beurteilung von AOGCMs enorm. Folgerichtig wurden und werden die teilweise hohen Abweichungen zwischen den AOGCMs und beliebigen Beobachtungsdaten im Wesentlichen ignoriert. Methodisch wurden hier zwei neue Ansätze vorgestellt. Die Nutzung von Konfidenzintervallen hat sich, im Hinblick auf die Beurteilung der relativen Performance der einzelnen Läufe eines Multi-Modell Ensembles, als dem klassischen RMSE nicht unterlegen herausgestellt. Beide Methoden identifizieren jene Modelle, die sich von den anderen durch größere Nähe zu den Beobachtungen hervortun. Das Gleiche gilt für Modelle mit im Vergleich stärkeren Abweichungen von den Beobachtungen. Die hier nicht ausführlicher gezeigte Analyse der einzelnen Jahreszeiten ergab, dass die Ergebnisse für jährliche Daten repräsentativ für die Jahreszeiten ist, ebenso dass auf der globalen Ebene die Leistung der Modelle für alle betrachteten Größen durch beide Verfahren als schlechter bewertet wird. Die hier vorgestellte Methodik hat den deutlichen Vorteil, dass sie auch auf einzelne Modelle anwendbar ist. Die Validation oder Beurteilung neuer Modelle kann somit anhand eines sinntragenden Kriteriums erfolgen, nämlich der Gebiete für die ein Modell die Klimatologien realistisch simuliert. Dies erscheint aussagekräftiger als ein Fehlermaß, dessen Aussage nur beim Vergleich mit anderen AOGCMs eine Beurteilung der Modellperformance ermöglicht.

Ebenfalls als erfolgreich kann die alternative Nutzung der Köppen-Geiger Klassifikation zur Validation betrachtet werden. Mit den Klimazonen als Entscheidungsmerkmal sind zunächst die meisten Differenzen innerhalb der Beobachtungsdaten zu beseitigen. Somit ist hierbei das Vertrauen in die Referenzdaten deutlich höher, was bedeutet, dass die Ergebnisse mit größerer Sicherheit interpretiert werden

können. Zudem ist die grundsätzlich realistische Darstellung der bestehenden Verhältnisse durch übereinstimmende Klimazonen eine sinnvolle Forderung an Modelle, die kommende Klimate erzeugen sollen. Hier soll betont werden, dass das Modell, welches nach der hier entwickelten Gewichtung den größten Einfluss auf den gewichteten Multi-Modell Mittelwert ausübt, ECHAM5/MPI-OM, auch die besten Resultate bei der Umsetzung der Köppen-Geiger Klassifikation erzielte.

Ein wichtiges Ergebnis dieser Arbeit ist sicherlich die Identifikation der global wichtigsten Eigenmoden. Insbesondere für die außertropische Zirkulation der Nordhemisphäre konnte aus den zahlreichen bekannten und konstatierten Klimamoden eine deutlich verringerte Anzahl in globaler Betrachtung tatsächlich relevanter Moden extrahiert werden.

Die umfangreichen Ansätze die unternommen wurden um Zirkulationsmuster zu detektieren, erlauben sicher die Aussage, dass Zirkulationsmuster, die hierbei berücksichtigt wurden und nicht identifiziert werden konnten, im globalen Maßstab keine wesentliche Rolle in der Organisation der Klimavariabilität spielen. Sie können somit als regionale Muster von nachgeordneter Bedeutung aufgefasst werden. Dagegen sollten die hier präsentierten Ergebnisse die wichtige Rolle von NAM, NPI und SCA im Klimasystem betonen. Damit verbunden wird hier auch die Notwendigkeit der weiteren Erforschung dieser Zirkulationsmoden sowohl im Hinblick auf ihre zukünftige Entwicklung, ihr Verhalten in der Vergangenheit sowie ihre physikalischen Grundlagen verdeutlicht. Das gleiche gilt für den SAM, der sich in nahezu allen betrachteten globalen Feldern von Klimagrößen als klar zu erfassende, dominierende Mode der Südhemisphäre erwies. Für diese wichtigen Bestandteile des globalen Klimasystems konnte zudem eine eigenständige Indexdefinition erstellt werden, die eindeutig als äquivalent zu verbreiteten, klassischen Methoden angesehen werden kann. Auch dies kann als Erfolg dieser Arbeit bewertet werden, zumal die hier entwickelten Indices den numerischen Aufwand bei der Erfassung dieser Größen, insbesondere bei multiplen Datensätzen, dramatisch reduzieren. Schlussendlich lieferte die neue Definition der Indices eine elegante Methode zur Überprüfung der Fähigkeit der AOGCMs, diese Muster mit realistischer raumzeitlicher Variabilität eigenständig zu simulieren. Dass alle Modelle dazu nach den hier gezeigten Ergebnissen in der Lage sind ist eine klare Bestätigung der Modellqualität. Auch scheint es in Folge des hier entwickelten Ansatzes angebracht, dem Multi-Modell Ensemble im Bezug auf die künftige Entwicklung der außertropischen Zirkulation ein deutliches Potential zuzusprechen. Dennoch sollte angesichts der klaren Trends in vielen außertropischen Zirkulationsindices in CMIP3 und der hierzu in einem gewissen Gegensatz stehender beobachteten enormen Variabilität der letzten Jahre darauf hingewiesen werden, dass auch für die Außertropen noch erheblicher Forschungsbedarf besteht. So lassen sich auch für die außertropischen Zirkulationsmoden in Detailfragen teilweise noch deutliche Abweichungen zwischen simuliertem und beobachtetem Verhalten feststellen (z. B. BARNES et al. 2010, RAPHAEL und HOLLAND 2006). Dies sind Fragen der Qualität

der Modelle, daneben muss aber auch die Qualität der Randbedingungen, die zu ihrem Antrieb genutzt werden, ständig kontrolliert und weiter verbessert werden. Die hier gezeigten Ergebnisse legen nahe, dass dies in der näheren Zukunft die vielleicht direkteste Möglichkeit darstellt, die Projektionen für die ektropische Klimavariabilität in qualitativer Hinsicht zu verbessern.

Sicherlich besteht auch die Notwendigkeit, die hier entwickelten Indexdefinitionen auf weitere Datensätze, Klimamodelle und weitere Beobachtungsdatensätze zu übertragen. Kann die Methodik auch diese Tests bestehen, stellen die Indexdefinitionen wegen ihrer einfachen und schnellen Anwendbarkeit auf große Datenensembles sicherlich einen wichtigen und nützlichen Beitrag für die praktische klimatologische Forschung dar.

Dass es nicht gelang ENSO auf diese Weise zu erfassen, legt die Vermutung nahe, dass sich diese Mode weniger eindeutig in globalen Feldgrößen identifizieren lässt. Dies bedeutet ausdrücklich nicht, dass ENSO weniger Bedeutung im globalen Klimasystem aufweist als die außertropischen Moden. Stattdessen werden verschiedene Aspekte von ENSO häufig durch unterschiedliche EOF/PC-Paare repräsentiert. Im Gegensatz dazu ließen sich einige bekannte außertropische Muster global überhaupt nicht erfassen. Die gewählte Methodik scheint für die Detektion von ENSO nicht geeignet. Das Modell, dass eine Mode durch eine EOF repräsentiert wird erwies sich hier als ungeeignet. Eventuell spielt auch die Auswahl der Variablen eine Rolle. Somit zeigt sich hier auch die Limitierung des konzeptionellen Ansatzes.

Es soll aber betont werden, dass dieser Aspekt allein aufgrund der Verwendung von verschiedenen Beobachtungsdaten überhaupt detektiert werden konnte. Bei alleiniger Verwendung des HadSLP2 hätte die 2. EOF als Repräsentation von ENSO aufgefasst werden können. Die Nicht-Nutzung von verfügbaren Daten hätte in diesem Fall wichtige Erkenntnisse verhindert. Die hier aufgezeigten Unterschiede in typischen Charakteristika des ENSO-Phänomens, sowohl die raumzeitliche Dynamik als auch seine Auswirkungen auf andere wichtige Klimagrößen, verdeutlichen dagegen, dass nicht alle Modelle gleichermaßen dazu geeignet sind, diese hochkomplexe Klimamode zu simulieren. Die Konsequenz hieraus muss dementsprechend sein, dass zumindest in Fragen die durch das Verhalten von ENSO betroffen sind, nicht allen AOGCMs des Multi-Modell Ensembles mit dem gleichen Vertrauen begegnet werden sollte.

Die unterschiedliche, und teilweise realitätsferne, Repräsentation der dominierenden Mode der interannuellen Klimavariabilität konnte durch diese Arbeit als eine relevante Unsicherheitsquelle für langfristige Klimaprojektionen identifiziert werden. Detektiert und quantifiziert wurde dies anhand einiger fundamentaler Größen des Klimasystems. So sind Modelle mit wenig realistischem ENSO ursächlich dafür verantwortlich, dass der SOI im CMIP3 keinen modellübergreifenden Trend aufweist. Eine negative Tendenz für den Index findet sich dagegen im Teilensemble A. Die Entwicklung des Indischen Monsuns unterscheidet sich in den beiden Teilensembles zwar nicht qualitativ, die quantitativen Unterschiede im simulierten Anstieg der

Regenmenge könnten in Kombination mit der regionalen und globalen Erwärmung aber durchaus ernstzunehmende Folgen haben. Für eine nähere Analyse dieser durch diese Arbeit aufgezeigten Problematik standen hier nicht die notwendigen AOGCM-Daten, geschweige denn die erforderlichen agrarischen und sozioökonomischen Daten, zur Verfügung. Die Grundlagen für eine derartige Untersuchung sind jedoch sicherlich durch die hier präsentierten Ergebnisse gelegt worden.

Bedeutsam erscheint auch der in Kapitel 8 aufgezeigte Zusammenhang der Entwicklungen von ENSO und der globalen Mitteltemperatur. Dass es möglich war diesen über eine relativ einfache Methodik zu erfassen, ist ein Zeichen für die Relevanz der verstärkten Berücksichtigung der natürlichen Klimavariabilität auch bei den Grundfragen des rezenten Klimawandels. Die Befunde dieses Kapitels sind ganz wesentlich ein Resultat des hier verfolgten Ansatzes. Sicherlich ermöglichte zu weiten Teilen erst die sorgfältige Auswahl des konkreten verwendeten SOIs die in diesem Kapitel gezeigten Ergebnisse. Ebenso wäre die Interpretation der unterschiedlichen Ausprägungen der Beziehung der modellspezifischen Zeitreihen von SOI und  $T_E$  ohne die Aufdeckung der stark unterschiedlichen ENSO-Systeme im CMIP3 nicht möglich gewesen. Somit stellt Kapitel 8 nicht nur ein wichtiges Ergebnis dieser Arbeit dar. Es ist zudem ein Beispiel für die starke Verzahnung der Themenbereiche großskaliger Klimavariabilität, anthropogener Klimawandel sowie Unsicherheitsfaktoren in Klimamodellen, mit denen sich diese Arbeit auseinandersetzte. Ebenso stellt es eine Rechtfertigung dar, auch im Bezug auf bekannte Phänomene im Klimasystem immer wieder auf grundlegende Fragen wie etwa eine Indexdefinition zurückzukommen. Dennoch bietet es einen konkreten Ansatzpunkt für weitere Untersuchungen. Konkret zeigte sich der Bedarf für eine ausreichend sensitive Methodik um die tatsächlichen Auswirkungen des SOIs auf die globale Temperatur robust zu quantifizieren, um dessen tatsächlichen Einfluss zuverlässig abzuschätzen.

Die Detektion von Klimaänderungssignalen in den hier betrachteten Zeitreihen (Kapitel 9) ist die derzeit umfangreichste und systematischste Analyse dieser Art für das CMIP3 Ensemble. Zudem findet sich ansonsten keine explizite Berücksichtigung des Einflusses eines realistischen ENSO-Systems auf die Entwicklung von Klimagrößen. Dessen Einfluss war nur in geringem Umfang erkennbar, was jedoch vermutlich zumindest teilweise durch die ungünstige Verteilung der SOI Variabilität im Teilensemble *A* erklärt werden kann. Einige bedeutsame Klimagrößen zeigen jedoch statistisch signifikante Unterschiede in ihrer Entwicklung unter dem anthropogenen Forcing, wenn das spektrale Verhalten des SOIs zur Unterscheidung von Teilensembles herangezogen wird. Der anthropogene Einfluss dagegen konnte in nahezu allen Zeitreihen und für alle Jahreszeiten stabil detektiert und quantifiziert werden. Die Veränderung der internen Klimadynamik wird dadurch als modellübergreifender Aspekt in CMIP3 und somit als wahrscheinliche Entwicklung für das reale Klimasystem hervorgehoben. Dies ist ein Beitrag dieser Arbeit zur Konkretisierung auch regionaler Aspekte des künftigen Klimawandels. Dennoch

konnte im Hinblick auf die zeitliche Entwicklung der untersuchten Klimagrößen auf sämtlichen Zeitskalen weiterhin eine hohe interne Variabilität für die CMIP3-Läufe festgestellt werden. Diese ist derart ausgeprägt, dass für die meisten Klimamoden kein wirkliches Prognosepotential erkannt werden kann. Einen wichtigen Anteil an dieser internen Unsicherheit dürften vom Gruppierungsmerkmal unabhängige Effekte der Modellphysik haben. Weder die verfügbaren Daten noch die Methodik erlaubte eine Berücksichtigung dieses Aspekts. Hier zeigt sich der Bedarf für ein neues Analyseverfahren, dessen Entwicklung allerdings nicht im Rahmen dieser Arbeit geleistet werden konnte. Da eine statistische Methode die an dieser Stelle hätte genutzt werden können, allerdings für eine Vielzahl ähnlicher Fragestellungen einsetzbar wäre, sollte die Entwicklung einer solchen ein Ziel künftiger Bemühungen darstellen. Die Ergebnisse sind jedoch im Bezug auf die modellübergreifenden Effekte des anthropogenen Forcings konkret genug, um einen direkten Vergleich mit dem CMIP5 Ensemble zu ermöglichen. Die Übertragung der Methode ist ohne Probleme möglich und bietet sicherlich einen sinnvollen Ausgangspunkt für den Vergleich der Entwicklung einiger wichtiger Größen des Klimasystems in verschiedenen Generationen von Klimamodellen.

Eine Grundfrage mit der sich diese Arbeit beschäftigte war die Behandlung von Modellen mit offensichtlich unterschiedlicher Qualität. Hierbei wurde versucht eine Gewichtung zu entwickeln, deren Anwendung die grundlegende Entwicklung des Klimasystems im Multi-Modell Ensemble durch Reduktion der Varianz der mittleren Veränderung einer Größe im Ensemble reduzieren sollte. Die numerische Gewichtung hat sich hierbei als wenig effektives Mittel erwiesen. Bei der Berücksichtigung eines großen Ensembles an Klimasimulationen erwies sich der Einfluss der konkreten Methode der Kombination der individuellen Resultate als kaum relevant für die tatsächliche Entwicklung des Ensembles. Dieses Ergebnis kann so interpretiert werden, dass die grundlegenden Aussagen der Projektionen zum Klimawandel derart stabil sind, dass individuelle Modellfehler auf sie nur marginalen Einfluss haben. Allerdings kann auch das Bewertungskriterium selbst für die wenig beeindruckenden Veränderungen des gewichteten Ensemblemittels im Vergleich zum arithmetischen herangezogen werden. Die Entwicklung des SOI ist gerade in Modellen mit realistischem ENSO-System derart variabel, dass sich gegenläufige Trends in den Indexreihen bei der Kombination unterschiedlicher Einzelläufe ausgleichen können. Deutlich erkennbar ist dieser Effekt bei der Entwicklung der globalen Mitteltemperatur. Zudem ist jede Gewichtung oder Beurteilung von Unterschieden, sogar deren Identifikation, durch die Auswahl von Kriterien und Metriken letztlich subjektiver Art. Der hier verfolgte Ansatz ist dabei durch seine Beschränkung auf einen einzigen Aspekt der AOGCMs sicherlich als ein extremes Beispiel für diese Aussage anführbar. Jedoch zeigte der Vergleich mit GLECKLER et al. (2008) sowie REICHLER und KIM (2008), dass basale zeitreihenanalytische Eigenschaften des SOI tendenziell eine Aussagekraft über die Performance eines AOGCMs haben.

Die hier gezeigten Resultate provozieren unweigerlich die Frage wieweit sich Gewichtungsmetriken im Hinblick auf die langfristigen Klimaprojektionen überhaupt auszahlen. CHRISTENSEN et al. (2010) erstellen etwa für eine geringere Anzahl von regionalen Klimamodellen eine auf einer Vielzahl unterschiedlicher Performance-Indices basierende Gewichtung. Auch für diese scheinbar weniger komplexe Situation mit einem sicherlich weniger dem Vorwurf der Beliebigkeit des Kriteriums ausgesetzten Methodik, sind die Effekte der Gewichtung nur gering und ohne tatsächliche Relevanz für die Aussagen der gemittelten Modelle. Das Fazit der Autoren, dass die Wichtung eine weitere Quelle für Unsicherheit für Ensembles-basierte Klimaprojektionen ist, ist gewiss nicht von der Hand zu weisen.

Signifikante Effekte in nennenswerten und interpretierbaren Maßen ließen sich allein über den kompletten Ausschluss von AOGCMs aus dem Multi-Modell Ensemble erzielen. Dies wurde im Laufe der Sensitivitätsstudie deutlich, wobei sich für die Teilensembles in einigen Punkten signifikante Unterschiede in der simulierten Entwicklung einiger Aspekte des Klimasystems, sowohl Zeitreihen als auch Feldgrößen, ergaben. Es muss jedoch betont werden, dass selbst die Differenzen zwischen den beiden hier ausführlich untersuchten Teilensembles nur als marginal bezeichnet werden können, verglichen mit den Unterschieden zwischen den Entwicklungen, die sich für die CMIP3 Modelle bei Verwendung eines anderen Emissionsszenarios ergeben. Insofern muss bei allen hier vertretenden Aussagen und Ergebnissen als die größte Unsicherheitsquelle das vorgegebene Forcing bedacht werden (WIGLEY und RAPER 2001).

Für zukünftige Studien im Anschluss an die hier präsentierten Ergebnisse ergeben sich einige wichtige Ansatzpunkte. Als erstes sollte die hier gefundene Bewertungsfunktion auf andere Modellensembles angewendet werden. Als mögliche Kandidaten hierfür bieten sich natürlich CMIP2 und CMIP5 an. Auch in dem aktuellsten Ensemble scheint es noch deutliche Unterscheide bei der Realisation von ENSO zwischen den einzelnen Modellen zu geben (LANGENBRUNNER und NEELIN 2013). Die wichtigste Mode der interannuellen Klimavariabilität behält also ihr Potential zur Beeinflussung der langfristigen Modellprojektionen. Sicherlich hat hier die Erweiterung des Prozessverständnisses Potential schlussendlich auch die Qualität künftiger Modellgenerationen zu verbessern. Eventuell würden dies auch die Verringerung der Unsicherheit in langfristigen Projektionen bedeuten, wenn die hier präsentierten Ergebnisse auch eher das Gegenteil andeuten. Erst mit einer höheren Anzahl an AOGCMs mit realistischem ENSO wird sich hierauf eventuell eine Antwort finden lassen.

Bei der Übertragung der grundsätzlichen Bewertungsmethodik sollten auch jeweils die übrigen Charakteristika des ENSO-Phänomens Beachtung finden. Vermutlich kann auf diese Weise eine noch besser geeignete Metrik für den Realismus des ENSO-Systems gefunden werden. Hierbei sollten aber die hier umgesetzten Leitgedanken beibehalten werden. Diese, sowohl die einfache und quantitative Erfassung der Modelleistung als auch die Reduktion der zum Vergleich herangezo-

genen Beobachtungen auf ein Minimum möglichst hochwertiger und gesicherter Daten, ermöglichen ein reproduzier- und diskutierbares Vorgehen, welches die Grundlage auch für die Bewertung von hochkomplexen GCMs darstellen muss. Dies gilt im besonderen Maße bei der sinnvollen und notwendigen Verwendung von Multi-Modell Ensembles.

Zum Abschluss wird hier noch auf die eigentliche Frage, die hinter der Motivation zur Bewertung der Qualität von Klimamodellen, und überhaupt der Verwendung von Multi-Modell Ensembles, steht eingegangen. Dies ist die Frage, inwieweit individuelle Modelldefizite die Fähigkeit der AOGCMs schmälern, das anthropogene Forcing realistisch umzusetzen und somit die Abschätzung von künftigen Erfordernissen zur Anpassung an die veränderte klimatische Realität beeinträchtigen. Hier konnte anhand des prominenten Beispiels ENSO aufgezeigt werden, dass sich solche Unsicherheitsquellen durchaus erfassen lassen und ihre Effekte quantifizierbar sind. Mit Sicherheit ist das zu diesem Zweck entwickelte und eingesetzte Methodenarsenal direkt auf andere Bereiche mit bedeutsamen und erfassbaren Modellunterschieden übertragbar, wofür zudem keine Beschränkung auf die interne Klimavariabilität besteht. Stattdessen könnte durchaus auch die Umsetzung von physikalischen Vorgängen und Prozessen herangezogen werden. Dabei bieten sich besonders solche Aspekte des Klimasystems an, deren Umsetzung im starken Maße auf Parametrisierungen beruht. Ein besonders naheliegendes Beispiel wäre etwa die modellspezifische Umsetzung der Wolkenbildung. Bei diesem Prozess bestehen bekannte Unterschiede zwischen den AOGCMs und alle Modelle weisen deutliche Defizite auf (z. B. SANDERSON et al. 2010). Weiterhin ist die Entwicklung der Bewölkung für die Zukunft des Weltklimas von direkterer Auswirkung als die Entwicklung des ENSO-Systems, so erlaubt die Manipulation der betreffenden Parametrisierungen die gezielte und nahezu unbegrenzte Änderung der Klimasensitivität von AOGCMs (SOKOLOV und MONIER 2012). Anhand dieser Überlegung wird der Bedarf und das Potential für die robuste Identifikation und Bewertung von Unsicherheitsquellen und unrealistischen Prozessen in komplexen Klimamodellen als auch in Zukunft wichtiger Forschungsbereich deutlich. Diese Arbeit sollte einen kleinen Beitrag bei der Erarbeitung dieses thematischen Komplexes geleistet haben.

---

# Literaturverzeichnis

- ABE, MANABU, H. SHIOGAMA, J. C. HARGREAVES, J. D. ANNAN, T. NOZAWA und S. EMORI (2009). *Correlation between inter-model similarities in spatial pattern for present and projected future mean climate*. *Sola*, 5(0):133–136.
- ABE-OUCHI, A, S. EMORI, A. HASEGAWA, H. HASUMI, T. INQUE, M. KIMOTO, S. MATSUMURA, T. NAGASHIMA, H. NAKANO, T. NISHIMURA, T. NOZAWA, R. OHGAI, A. OKA, N. OKADA, K. OGOCHI, T. OGURA, F. SAITO, K. SAITO, T. SAKAMATO, T. SEGAWA, T. SOGA, K. SUDO, A. SUMI, T. SUZUKI, T. SUZUKI, H. TAKAHASHI, K. TAKATA, T. TAKEMURA, M. TAKIGAWA, Y. TSUSHIMA, M. WATANABE, S. WATANABE und T. YOKOHATA (2004). *K-1 Coupled GCM (MIROC) Description*.
- ACIA (2005). *Arctic Climate Impact Assessment - Scientific Report* -. Cambridge University Press, Cambridge.
- ADLER, ROBERT F, G. J. HUFFMAN, A. CHANG, R. FERRARO, P.-P. XIE, J. JANOWIAK, B. RUDOLF, U. SCHNEIDER, S. CURTIS, D. BOLVIN et al. (2003). *The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present)*. *Journal of Hydrometeorology*, 4(6):1147–1167.
- AHRENS, C.D. (2007). *Meteorology Today: An Introduction to Weather, Climate, and the Environment*. International student edition. Brooks/Cole.
- ALEXEEV, VA, E. VOLODIN, V. Y. GALIN, V. DYMNIKOV und V. LYKOSOV (1998). *Modelling of the present-day climate by the INM RAS atmospheric model DNM GCM*.
- ALLAN, ROB und T. ANSELL (2006). *A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850-2004*. *Journal of Climate*, 19(22):5816–5842.
- ALLEN, MYLES R und W. J. INGRAM (2002). *Constraints on future changes in climate and the hydrologic cycle*. *Nature*, 419(6903):224–232.
- AMBAUM, MAARTEN HP, B. J. HOSKINS und D. B. STEPHENSON (2001). *Arctic oscillation or North Atlantic oscillation?*. *Journal of Climate*, 14(16):3495–3507.

- ÅNGSTRÖM, ANDERS (1935). *Teleconnections of climatic changes in present time*. Geografiska Annaler, 17:242–258.
- APPENZELLER, C, T. STOCKER und M. ANKLIN (1998). *North Atlantic Oscillation dynamics recorded in Greenland ice cores*. Science, 282(5388):446–449.
- BARNES, ELIZABETH A, D. L. HARTMANN, D. M. FRIERSON und J. KIDSTON (2010). *Effect of latitude on the persistence of eddy-driven jets*. Geophysical research letters, 37(11).
- BARNETT, T, F. ZWIERS, G. HEGERL, M. ALLEN, T. CROWLEY, N. GILLETT, K. HASSELMANN, P. JONES, B. SANTER, R. SCHNUR et al. (2005). *Detecting and attributing external influences on the climate system: A review of recent advances*. J. Clim, 18(9):1291–1314.
- BARNSTON, ANTHONY G und R. E. LIVEZEY (1987). *Classification, seasonality and persistence of low-frequency atmospheric circulation patterns*. Monthly Weather Review, 115(6):1083–1126.
- BARRY, R.G. und A. CARLETON (2001). *Synoptic and Dynamic Climatology*. Routledge.
- BAUMHAUER, R. (2011). *Physische Geographie 2 ; Klima-, Hydro-, Boden-, Vegetationsgeographie*. WBG (Wissenschaftliche Buchgesellschaft).
- BEGERT, MICHAEL, E. ZENKLUSEN, C. HABERLI, C. APPENZELLER und L. KLOK (2008). *An automated procedure to detect discontinuities; performance assessment and application to a large European climate data set*. Meteorologische Zeitschrift, 17(5):663–672.
- BINGHAM, N.H., N. BINGHAM und J. FRY (2010). *Regression: Linear Models in Statistics*. Springer undergraduate mathematics series. Springer London.
- BJERKNES, JAKOB (1969). *Atmospheric teleconnections from the equatorial pacific 1*. Monthly Weather Review, 97(3):163–172.
- BLENKINSOP, S und H. FOWLER (2007). *Changes in drought frequency, severity and duration for the British Isles projected by the PRUDENCE regional climate models*. Journal of Hydrology, 342(1):50–71.
- BLÜTHGEN, J. und W. WEISCHET (1980). *Allgemeine Klimageographie*. Lehrbuch der Allgemeinen Geographie. W.De Gruyter.
- BOÉ, JULIEN, A. HALL und X. QU (2009). *September sea-ice cover in the Arctic Ocean projected to vanish by 2100*. Nature Geoscience, 2(5):341–343.

- BOHLEY, P. (2000). *Statistik: einführendes Lehrbuch für Wirtschafts- und Sozialwissenschaftler*. Oldenbourg.
- BONSAL, BARRIE R, A. SHABBAR und K. HIGUCHI (2001). *Impacts of low frequency variability modes on Canadian winter temperature*. International Journal of Climatology, 21(1):95–108.
- BORN, KAI, A. H. FINK und H. PAETH (2008). *Dry and wet periods in the northwestern Maghreb for present day and future climate conditions*. Meteorologische Zeitschrift, 17(5):533–551.
- BORTZ, JÜRGEN und C. SCHUSTER (2010). *Statistik für Human-und Sozialwissenschaftler. Lehrbuch mit Online-Materialien*. Springer.
- BRETHERTON, CHRISTOPHER S, M. WIDMANN, V. P. DYMNIKOV, J. M. WALLACE und I. BLADE (1999). *The effective number of spatial degrees of freedom of a time-varying field*. Journal of Climate, 12(7):1990–2009.
- BRIDGMAN, H.A., J. OLIVER und M. H. GLANTZ (2006). *The Global Climate System: Patterns, Processes, and Teleconnections*. Cambridge University Press.
- BRIFFA, KEITH R (2000). *Annual climate variability in the Holocene: interpreting the message of ancient trees*. Quaternary Science Reviews, 19(1):87–105.
- BRÖNNIMANN, S (2007). *Impact of El Niño–Southern Oscillation on European climate*. Reviews of Geophysics, 45(3).
- BRÖNNIMANN, S, E. XOPLAKI, C. CASTY, A. PAULING und J. LUTERBACHER (2007). *ENSO influence on Europe during the last centuries*. Climate Dynamics, 28(2-3):181–197.
- BUEH, CHOLAW und H. NAKAMURA (2007). *Scandinavian pattern and its climatic impact*. Quarterly Journal of the Royal Meteorological Society, 133(629):2117–2131.
- BUSER, CHRISTOPH M, H. KÜNSCH, D. LÜTHI, M. WILD und C. SCHÄR (2009). *Bayesian multi-model projection of climate: bias assumptions and interannual variability*. Climate Dynamics, 33(6):849–868.
- BUTZ, T. (2009). *Fouriertransformation Für Fußgänger*. Vieweg Studium. Vieweg Verlag, Friedr, & Sohn Verlagsgesellschaft mbH.
- CAMBERLIN, P, S. JANICOT und I. POCCARD (2001). *Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic vs. ENSO*. International Journal of Climatology, 21(8):973–1005.

- CASSOU, CHRISTOPHE und L. TERRAY (2001). *Dual influence of Atlantic and Pacific SST anomalies on the North Atlantic/Europe winter climate*. Geophysical research letters, 28(16):3195–3198.
- CATTIAUX, J, R. VAUTARD, C. CASSOU, P. YIOU, V. MASSON-DELMOTTE und F. CODRON (2010). *Winter 2010 in Europe: a cold extreme in a warming climate*. Geophysical Research Letters, 37(20).
- CHANDLER, RICHARD E (2013). *Exploiting strength, discounting weakness: combining information from multiple climate simulators*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371(1991).
- CHANG, C.P., Y. DING, W. M. ORGANIZATION und N. LAU (2011). *The Global Monsoon System: Research and Forecast*. World Scientific series on Asia-Pacific weather and climate. World Scientific.
- CHERRY, JESSIE, H. CULLEN, M. VISBECK, A. SMALL und C. UVO (2005). *Impacts of the North Atlantic Oscillation on Scandinavian hydropower production and energy markets*. Water resources management, 19(6):673–691.
- CHRISTENSEN, JENS H, F. BOBERG, O. B. CHRISTENSEN und P. LUCAS-PICHER (2008). *On the need for bias correction of regional climate change projections of temperature and precipitation*. Geophysical Research Letters, 35(20).
- CHRISTENSEN, JENS HESSELBJERG, F. GIORGI und G. LENDERINK (2010). *Weight assignment in regional climate models*. Climate Research, 44(2-3):179–194.
- COBB, KIM M, C. D. CHARLES, H. CHENG und R. L. EDWARDS (2003). *El Niño/Southern Oscillation and tropical Pacific climate during the last millennium*. Nature, 424(6946):271–276.
- COLLINS, MATTHEW (2005). *El Niño-or La Niña-like climate change?*. Climate Dynamics, 24(1):89–104.
- COLLINS, MATTHEW, S.-I. AN, W. CAI, A. GANACHAUD, E. GUILYARDI, F.-F. JIN, M. JOCHUM, M. LENGAINNE, S. POWER, A. TIMMERMANN et al. (2010). *The impact of global warming on the tropical Pacific Ocean and El Niño*. Nature Geoscience, 3(6):391–397.
- COLLINS, WILLIAM D, C. M. BITZ, M. L. BLACKMON, G. B. BONAN, C. S. BRETHERTON, J. A. CARTON, P. CHANG, S. C. DONEY, J. J. HACK, T. B. HENDERSON et al. (2006). *The community climate system model version 3 (CCSM3)*. Journal of Climate, 19(11):2122–2143.

- COPPOLA, E, F. GIORGI, S. RAUSCHER und C. PIANI (2010). *Model weighting based on mesoscale structures in precipitation and temperature in an ensemble of regional climate models*. Climate research (Open Access for articles 4 years old and older), 44(2):121.
- COVEY, CURT, K. M. ACHUTARAO, U. CUBASCH, P. JONES, S. J. LAMBERT, M. E. MANN, T. J. PHILLIPS und K. E. TAYLOR (2003). *An overview of results from the Coupled Model Intercomparison Project*. Global and Planetary Change, 37(1):103–133.
- D'ARRIGO, ROSANNE, E. R. COOK, R. J. WILSON, R. ALLAN und M. E. MANN (2005). *On the variability of ENSO over the past six centuries*. Geophysical Research Letters, 32(3):L03711.
- DE BLIJ, H.J., P. MULLER und J. RICHARD S. WILLIAMS (2004). *Physical Geography: The Global Environment, Third Edition and Geography*. Oxford University Press.
- DELCROIX, THIERRY (1998). *Observed surface oceanic and atmospheric variability in the tropical Pacific at seasonal and ENSO timescales: A tentative overview*. Journal of Geophysical Research: Oceans (1978–2012), 103(C9):18611–18633.
- DÉQUÉ, M und S. SOMOT (2010). *Weighted frequency distributions express modelling uncertainties in the ENSEMBLES regional climate experiments*. Climate research (Open Access for articles 4 years old and older), 44(2):195.
- DÉQUÉ, MICHEL, C. DREVEYTON, A. BRAUN und D. CARIOLLE (1994). *The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling*. Climate Dynamics, 10(4-5):249–266.
- DESER, CLARA (2000). *On the teleconnectivity of the 'Arctic Oscillation'*. Geophysical Research Letters, 27(6):779–782.
- DIAZ, HENRY F und V. MARKGRAF (1992). *El Niño: historical and paleoclimatic aspects of the Southern Oscillation*. Cambridge University Press.
- DIAZ, HENRY F und V. MARKGRAF (2000). *El Niño and the Southern Oscillation: multiscale variability and global and regional impacts*. Cambridge University Press.
- ESBENSEN, STEVEN K (1984). *A comparison of intermonthly and interannual teleconnections in the 700 mb geopotential height field during the Northern Hemisphere winter*. Monthly weather review, 112(10):2016–2032.
- FAHRMEIR, L., T. KNEIB und S. LANG (2009). *Regression: Statistik und ihre Anwendungen*. Springer.

- FELDSTEIN, STEVEN B (2000). *The timescale, power spectra, and climate noise properties of teleconnection patterns*. *Journal of Climate*, 13(24):4430–4440.
- FLOHN, HERMANN (1957). *Zur Frage der Einteilung der Klimazonen (The Problem of a Classification of Climatic Zones)*. *Erdkunde*, S. 161–175.
- FOKEN, THOMAS (2006). *Angewandte Meteorologie: Mikrometeorologische Methoden*. Springer.
- FORCHHAMMER, MADIS C und E. POST (2004). *Using large-scale climate indices in climate change ecology studies*. *Population Ecology*, 46(1):1–12.
- FORDHAM, DAMIEN A, T. M. WIGLEY und B. W. BROOK (2011). *Multi-model climate projections for biodiversity risk assessments*. *Ecological Applications*, 21(8):3317–3331.
- FOSTER, GRANT und S. RAHMSTORF (2011). *Global temperature evolution 1979–2010*. *Environmental Research Letters*, 6(4):044022.
- FOWLER, HJ und M. EKSTRÖM (2009). *Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes*. *International Journal of Climatology*, 29(3):385–416.
- FOWLER, HJ, M. EKSTRÖM, S. BLENKINSOP und A. SMITH (2007). *Estimating change in extreme European precipitation using a multimodel ensemble*. *Journal of Geophysical Research: Atmospheres* (1984–2012), 112(D18).
- FOX, DOUGLAS G. (1981). *Judging Air Quality Model Performance*. *Bulletin of The American Meteorological Society*, 62:599–609.
- FRAEDRICH, K, F.-W. GERSTENGARBE und P. WERNER (2001). *Climate shifts during the last century*. *Climatic Change*, 50(4):405–417.
- FRAPPIER, AMY, D. SAHAGIAN, L. A. GONZÁLEZ und S. J. CARPENTER (2002). *El Niño events recorded by stalagmite carbon isotopes*. *Science*, 298(5593):565–565.
- FREIDENREICH, S M, S. GARNER, R. G. GUDGEL, C. GORDON und I. M. HELD (2004). *The new GFDL global atmosphere and land model AM2-LM2: Evaluation with prescribed SST simulations*. *Journal of Climate*, 17:4641–4673.
- GADGIL, SULOCHANA und S. GADGIL (2006). *The Indian monsoon, GDP and agriculture*. *Economic and Political Weekly*, S. 4887–4895.
- GAO, XUEJIE und F. GIORGI (2008). *Increased aridity in the Mediterranean region under greenhouse gas forcing estimated from high resolution simulations with a regional climate model*. *Global and Planetary Change*, 62(3):195–209.

- GERSHUNOV, ALEXANDER und T. P. BARNETT (1998). *Interdecadal modulation of ENSO teleconnections*. Bulletin of the American Meteorological Society, 79(12):2715–2725.
- GERSTENGARBE, F-W, P. WERNER und K. FRAEDRICH (1999). *Applying non-hierarchical cluster analysis algorithms to climate classification: some problems and their solution*. Theoretical and applied climatology, 64(3):143–150.
- Ghil, MICHAEL, M. ALLEN, M. DETTINGER, K. IDE, D. KONDRASHOV, M. MANN, A. W. ROBERTSON, A. SAUNDERS, Y. TIAN, F. VARADI et al. (2002). *Advanced spectral methods for climatic time series*. Reviews of Geophysics, 40(1):1003.
- GILLETT, NATHAN P und D. W. THOMPSON (2003). *Simulation of recent Southern Hemisphere climate change*. Science, 302(5643):273–275.
- GILLETT, NP, T. KELL und P. JONES (2006). *Regional climate impacts of the Southern Annular Mode*. Geophysical Research Letters, 33(23).
- GIORGI, F und L. MEARNs (2003). *Probability of regional climate change based on the Reliability Ensemble Averaging (REA) method*. Geophysical Research Letters, 30(12):1629.
- GIORGI, FILIPPO und R. FRANCISCO (2000). *Evaluating uncertainties in the prediction of regional climate change*. Geophysical Research Letters, 27(9):1295–1298.
- GIORGI, FILIPPO und L. O. MEARNs (2002). *Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the 'reliability ensemble averaging'(REA) method*. Journal of Climate, 15(10):1141–1158.
- GLECKLER, PETER J, K. E. TAYLOR und C. DOUTRIAUX (2008). *Performance metrics for climate models*. Journal of Geophysical Research, 113(D6):D06104.
- GONG, DAOYI und S. WANG (1999). *Definition of Antarctic oscillation index*. Geophysical Research Letters, 26(4):459–462.
- GORDON, KB, L. ROTSTAYN, J. MCGREGOR, M. DIX, E. KOWALCZYK, S. OFARRELL, L. WATERMAN, A. HIRST, S. WILSON, M. COLLIER, I. WATERTON und T. ELLIOTT (2002). *The CSIRO Mk3 Climate System Model*. CSIRO Atmospheric Research Technical Paper, 60:130.
- GREENE, ARTHUR M, L. GODDARD und U. LALL (2006). *Probabilistic multimodel regional temperature change projections*. Journal of Climate, 19(17):4326–4343.

- GREGORY, STANLEY (1954). *Climatic Classification and Climatic Change (Klimaklassifikation und Klimaänderung)*. Erdkunde, S. 246–252.
- GRIMM, ALICE M, V. R. BARROS und M. E. DOYLE (2000). *Climate variability in southern South America associated with El Niño and La Niña events*. Journal of climate, 13(1):35–58.
- GUETTER, PETER J und J. E. KUTZBACH (1990). *A modified Köppen classification applied to model simulations of glacial and interglacial climates*. Climatic change, 16(2):193–215.
- HALL, ALEX und X. QU (2006). *Using the current seasonal cycle to constrain snow albedo feedback in future climate change*. Geophysical Research Letters, 33(3).
- HALLETT, TB, T. COULSON, J. PILKINGTON, T. CLUTTON-BROCK, J. PEMBERTON und B. GRENFELL (2004). *Why large-scale climate indices seem to predict ecological processes better than local weather*. Nature, 430(6995):71–75.
- HALPERT, MICHAEL S und C. F. ROPELEWSKI (1992). *Surface temperature patterns associated with the Southern Oscillation*. Journal of Climate, 5(6):577–593.
- HAM, YOO-GEUN und J.-S. KUG (2012). *How well do current climate models simulate two types of El Niño?*. Climate dynamics, 39(1-2):383–398.
- HANNACHI, A, I. JOLLIFFE und D. STEPHENSON (2007). *Empirical orthogonal functions and related techniques in atmospheric science: A review*. International Journal of Climatology, 27(9):1119–1152.
- HANSEN, J, M. SATO und R. RUEDY (1997). *Radiative forcing and climate response*. Journal of Geophysical Research: Atmospheres (1984–2012), 102(D6):6831–6864.
- HANSEN, JAMES, D. JOHNSON, A. LACIS, S. LEBEDEFF, P. LEE, D. RIND und G. RUSSELL (1981). *Climate impact of increasing atmospheric carbon dioxide*. Science, 213(4511):957–966.
- HANSEN, JAMES, A. LACIS, R. RUEDY und M. SATO (1992). *Potential climate impact of Mount Pinatubo eruption*. Geophysical Research Letters, 19(2):215–218.
- HANSEN, JAMES und S. LEBEDEFF (1988). *Global surface air temperatures: Update through 1987*. Geophysical Research Letters, 15(4):323–326.

- HANSEN, JAMES, R. RUEDY, M. SATO und K. LO (2010). *Global surface temperature change*. *Reviews of Geophysics*, 48(4):RG4004.
- HANSEN, JAMES, M. SATO, P. KHARECHA, D. BEERLING, R. BERNER, V. MASSON-DELMOTTE, M. PAGANI, M. RAYMO, D. L. ROYER und J. C. ZACHOS (2008). *Target atmospheric CO<sub>2</sub>: Where should humanity aim?*. arXiv preprint arXiv:0804.1126.
- HANSEN, JAMES, M. SATO, R. RUEDY, K. LO, D. W. LEA und M. MEDINA-ELIZADE (2006). *Global temperature change*. *Proceedings of the National Academy of Sciences*, 103(39):14288–14293.
- HANSEN, JAMES E (1988). *The greenhouse effect: Impacts on current global temperature and regional heat waves*. Testimony before the Committee on Energy and Natural Resources, US Senate, Washington DC, June, 23:1988.
- HANTEL, MICHAEL (2001). *Klimatologie*. In: BERGMANN, L., W. RAITH, S. BAUER und C. SCHÄFER, Hrsg.: *Bergmann/Schaefer:Experimentalphysik Bd 7*, Aktualisierte Auflage, 2. De Gruyter.
- HARTMANN, DENNIS L, J. M. WALLACE, V. LIMPASUVAN, D. W. THOMPSON und J. R. HOLTON (2000). *Can ozone depletion and global warming interact to produce rapid climate change?*. *Proceedings of the National Academy of Sciences*, 97(4):1412–1417.
- HARZALLAH, A und R. SADOURNY (1997). *Observed lead-lag relationships between Indian summer monsoon and some meteorological variables*. *Climate dynamics*, 13(9):635–648.
- HASSELMANN, KLAUS (1979). *On the signal-to-noise problem in atmospheric response studies*. *Meteorology of tropical oceans*, S. 251–259.
- HAYLOCK, M. R., N. HOFSTRA, A. M. G. KLEIN TANK, E. J. KLOK, P. D. JONES und M. NEW (2008). *A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006*. *Journal of Geophysical Research: Atmospheres*, 113(D20).
- HEGERL, GABRIELE und F. ZWIERS (2011). *Use of models in detection and attribution of climate change*. *Wiley Interdisciplinary Reviews: Climate Change*, 2(4):570–591.
- HEGERL, GABRIELE C, H. VON STORCH, K. HASSELMANN, B. D. SANTER, U. CUBASCH und P. D. JONES (1996). *Detecting greenhouse-gas-induced climate change with an optimal fingerprint method*. *Journal of Climate*, 9(10):2281–2306.

- HENDL, MANFRED (1991). *Globale Klimaklassifikation*. In: HUPFER, PETER, Hrsg.: *Das Klimasystem der Erde*, S. 218–266. Akademie Verlag GmbH.
- HENDL, MANFRED (1997). *Allgemeine Klimageographie*. In: HENDL, MANFRED und H. LIEDTKE, Hrsg.: *Lehrbuch der Allgemeinen Physischen Geographie*, S. xx–xx. Klett.
- HO, CHUN KIT, D. B. STEPHENSON, M. COLLINS, C. A. FERRO und S. J. BROWN (2012). *Calibration strategies: a source of additional uncertainty in climate change projections*. *Bulletin of the American Meteorological Society*, 93(1):21–26.
- HOERLING, MARTIN P, J. W. HURRELL, T. XU, G. T. BATES und A. PHILLIPS (2004). *Twentieth century North Atlantic climate change. Part II: Understanding the effect of Indian Ocean warming*. *Climate Dynamics*, 23(3-4):391–405.
- HOFSTRA, NYNKE, M. HAYLOCK, M. NEW und P. D. JONES (2009). *Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature*. *Journal of Geophysical Research: Atmospheres* (1984–2012), 114(D21).
- HOFSTRA, NYNKE, M. NEW und C. MCSWEENEY (2010). *The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data*. *Climate dynamics*, 35(5):841–858.
- HOLDEN, J. (2008). *An Introduction to Physical Geography and the Environment*. Pearson Prentice Hall.
- HOREL, JOHN D (1981). *A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field*. *Monthly Weather Review*, 109(10):2080–2092.
- HOUDIN, FRÉDÉRIC, I. MUSAT, S. BONY, P. BRACONNOT, F. CODRON, J.-L. DUFRESNE, L. FAIRHEAD, M.-A. FILIBERTI, P. FRIEDLINGSTEIN, J.-Y. GRANDPEIX et al. (2006). *The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection*. *Climate Dynamics*, 27(7-8):787–813.
- HU, ZENG-ZHEN, M. LATIF, E. ROECKNER und L. BENGTSSON (2000). *Intensified Asian summer monsoon and its variability in a coupled model forced by increasing greenhouse gas concentrations*. *Geophysical Research Letters*, 27(17):2681–2684.
- HUFFMAN, GEORGE J, R. F. ADLER, B. RUDOLF, U. SCHNEIDER und P. R. KEEHN (1995). *Global precipitation estimates based on a technique for combining*

- satellite-based estimates, rain gauge analysis, and NWP model precipitation information.* Journal of Climate, 8(5):1284–1295.
- HUFFMAN, GEORGE J, D. T. BOLVIN, E. J. NELKIN, D. B. WOLFF, R. F. ADLER, G. GU, Y. HONG, K. P. BOWMAN und E. F. STOCKER (2007). *The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales.* Journal of Hydrometeorology, 8(1):38–55.
- HULME, MIKE (1992). *Rainfall changes in Africa: 1931–1960 to 1961–1990.* International Journal of Climatology, 12(7):685–699.
- HURRELL, JAMES W (1995). *Decadal trends in the North Atlantic oscillation.* Science, 269:676–679.
- HURRELL, JAMES W und C. DESER (2009). *North Atlantic climate variability: the role of the North Atlantic Oscillation.* Journal of Marine Systems, 78(1):28–41.
- HURRELL, JAMES W, Y. KUSHNIR, G. OTTERSEN und M. VISBECK (2003). *The North Atlantic Oscillation: climatic significance and environmental impact*, Bd. 134. American Geophysical Union.
- ICSU und WMO (1990). *Global Precipitation Climatology Project, Implementation and Data Management Plan.* WMO/TD. World Meteorological Organization, International Council of Scientific Unions [1990].
- INOUE, TOMOSHIGE und J. MATSUMOTO (2004). *A comparison of summer sea level pressure over East Eurasia between NCEP-NCAR reanalysis and ERA-40 for the period 1960-99.* Journal of the Meteorological Society of Japan. Ser. II, 82(3):951–958.
- IPCC, WG1 (1990a). *Climate change: The IPCC Scientific Assessment..* Cambridge University Press, Cambridge and , New York.
- IPCC, WG1 (1996a). *Climate change 1995: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge and , New York.
- IPCC, WG1 (2001a). *Climate change 2001: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge and , New York.

- IPCC, WG1 (2007a). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and , New York.
- IPCC, WG1 (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers*.
- IPCC, WG2 (1990b). *Climate change: The IPCC Impacts Assessment*. Cambridge University Press, Cambridge and , New York.
- IPCC, WG2 (1996b). *Climate change 2001: Impacts, Adaptation, and Vulnerability: Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and , New York.
- IPCC, WG2 (2001b). *Climate change 2001: Impacts, Adaptation, and Vulnerability: Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and , New York.
- IPCC, WG2 (2007b). *Climate Change 2007: Impacts, Adaptation and Vulnerability: Contribution of Working Group II to the Fourth Assessment Report of the IPCC Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and , New York.
- IPCC, WG3 (1990c). *Climate change: The IPCC Response Strategies*. Cambridge University Press, Cambridge and , New York.
- IPCC, WG3 (1996c). *Climate change 2001: Mitigation: Contribution of Working Group III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and , New York.
- IPCC, WG3 (2001c). *Climate change 2001: Mitigation: Contribution of Working Group III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and , New York.
- IPCC, WG3 (2007c). *Climate Change 2007: Mitigation of Climate Change: Contribution of Working Group II Contribution to the Fourth Assessment Report of the IPCC Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and , New York.
- IRVING, DAMIEN B, S. E. PERKINS, J. R. BROWN, A. SEN GUPTA, A. F. MOISE, B. F. MURPHY, L. C. MUIR, R. A. COLMAN, S. B. POWER, F. P.

- DELAGÉ et al. (2012). *Evaluating global climate models for the Pacific island region*. *Climate Research*, 49(3):169.
- JEFFREY, STEPHEN J, J. O. CARTER, K. B. MOODIE und A. R. BESWICK (2001). *Using spatial interpolation to construct a comprehensive archive of Australian climate data*. *Environmental Modelling & Software*, 16(4):309–330.
- JOLLIFFE, I.T. (2002). *Principal Component Analysis*. Springer Series in Statistics. Springer.
- JONES, PHIL D, M. NEW, D. E. PARKER, S. MARTIN und I. G. RIGOR (1999). *Surface air temperature and its changes over the past 150 years*. *Reviews of Geophysics*, 37(2):173–199.
- KALNAY, EUGENIA, M. KANAMITSU, R. KISTLER, W. COLLINS, D. DEAVEN, L. GANDIN, M. IREDELL, S. SAHA, G. WHITE, J. WOOLLEN et al. (1996). *The NCEP/NCAR 40-year reanalysis project*. *Bulletin of the American meteorological Society*, 77(3):437–471.
- KATZ, RICHARD W (2002). *Techniques for estimating uncertainty in climate change scenarios and impact studies*. *Climate Research*, 20(2):167–185.
- KAY, S.M. (1988). *Modern Spectral Estimation: Theory and Application/Book and Disk*. Prentice-Hall Signal Processing Series: Advanced monographs. PTR Prentice Hall.
- KENDON, ELIZABETH J, D. P. ROWELL, R. G. JONES und E. BUONOMO (2008). *Robustness of future changes in local precipitation extremes*. *Journal of Climate*, 21(17):4280–4297.
- KERR, RA (2004). *Climate change. A few good climate shifters..* *Science* (New York, NY), 306(5696):599.
- KHARIN, VIATCHESLAV V und F. W. ZWIERS (2002). *Climate predictions with multimodel ensembles*. *Journal of Climate*, 15(7):793–799.
- KIEHL, JT, J. HACK, G. BONAN, B. BOVILLE, D. WILLIAMSON und P. RASCH (1998). *The national center for atmospheric research community climate model: CCM3\**. *Journal of Climate*, 11(6):1131–1149.
- KILADIS, GEORGE N und H. VAN LOON (1988). *The Southern Oscillation. Part VII: Meteorological anomalies over the Indian and Pacific sectors associated with the extremes of the oscillation*. *Monthly weather review*, 116(1):120–136.

- KISTLER, ROBERT, W. COLLINS, S. SAHA, G. WHITE, J. WOOLLEN, E. KALNAY, M. CHELLIAH, W. EBISUZAKI, M. KANAMITSU, V. KOUSKY et al. (2001). *The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation*. Bulletin of the American Meteorological society, 82(2):247–267.
- KJELLSTRÖM, ERIK, F. BOBERG, M. CASTRO, J. H. CHRISTENSEN, G. NIKULIN und E. SÁNCHEZ (2010). *Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models*. Climate research (Open Access for articles 4 years old and older), 44(2):135.
- KLOCKE, DANIEL, R. PINCUS und J. QUAAS (2011). *On constraining estimates of climate sensitivity with present-day observations through model weighting*. Journal of Climate, 24(23):6092–6099.
- KNUTTI, RETO (2010). *The end of model democracy?*. Climatic change, 102(3–4):395–404.
- KNUTTI, RETO, G. A. MEEHL, M. R. ALLEN und D. A. STAINFORTH (2006). *Constraining climate sensitivity from the seasonal cycle in surface temperature*. Journal of Climate, 19(17):4224–4233.
- KÖPPEN, WLADIMIR PETER (1918). *Klassifikation der Klimate nach Temperatur, Niederschlag und Jahreslauf*. Petermanns Geographische Mitteilungen, 62:193–203.
- KÖPPEN, W.P. (1936). *Das geographische System der Klimate*. Handbuch der Klimatologie in fünf Bänden. Borntraeger.
- KOSAKA, YU und S.-P. XIE (2013). *Recent global-warming hiatus tied to equatorial Pacific surface cooling*. Nature.
- KOTTEK, MARKUS, J. GRIESER, C. BECK, B. RUDOLF und F. RUBEL (2006). *World map of the Koppen-Geiger climate classification updated*. Meteorologische Zeitschrift, 15(3):259–263.
- KRAUS, H. (2004). *Die Atmosphäre Der Erde: Eine Einführung in Die Meteorologie*. Springer.
- KRISHNAMURTHY, V und B. N. GOSWAMI (2000). *Indian monsoon-ENSO relationship on interdecadal timescale*. Journal of Climate, 13(3):579–595.
- KRISHNAMURTI, TIRUVALAM NATARAJAN, C. KISHTAWAL, Z. ZHANG, T. LAROW, D. BACHIOCHI, E. WILLIFORD, S. GADGIL und S. SURENDRAN (2000). *Multimodel ensemble forecasts for weather and seasonal climate*. Journal of Climate, 13(23):4196–4216.

- KUHLBRODT, TILL, S. RAHMSTORF, K. ZICKFELD, F. B. VIKEBØ, S. SUNDBY, M. HOFMANN, P. M. LINK, A. BONDEAU, W. CRAMER und C. JAEGER (2009). *An integrated assessment of changes in the thermohaline circulation*. *Climatic Change*, 96(4):489–537.
- KUSHNIR, YOCHANAN und J. M. WALLACE (1989). *Low-frequency variability in the northern hemisphere winter: geographical distribution, structure and time-scale dependence*. *Journal of the Atmospheric Sciences*, 46(20):3122–3143.
- LAMBERT, STEVEN J und G. J. BOER (2001). *CMIP1 evaluation and intercomparison of coupled climate models*. *Climate Dynamics*, 17(2-3):83–106.
- LANGENBRUNNER, BAIRD und J. D. NEELIN (2013). *Analyzing ENSO teleconnections in CMIP models as a measure of model fidelity in simulating precipitation*. *Journal of Climate*, 26.
- LATIF, MOJIB, D. ANDERSON, T. BARNETT, M. CANE, R. KLEEMAN, A. LEETMAA, J. O'BRIEN, A. ROSATI und E. SCHNEIDER (1998). *A review of the predictability and prediction of ENSO*. *Journal of Geophysical Research: Oceans* (1978–2012), 103(C7):14375–14393.
- LAU, K-M und P. SHEU (1988). *Annual cycle, quasi-biennial oscillation, and southern oscillation in global precipitation*. *Journal of Geophysical Research: Atmospheres* (1984–2012), 93(D9):10975–10988.
- LAU, KM, J. LEE, K. KIM und I. KANG (2004). *The North Pacific as a regulator of summertime climate over Eurasia and North America*. *Journal of Climate*, 17(4):819–833.
- LAUER, W und J. BENDIX (2006). *Klimatologie. 2*. Westermann.
- LAUER, WILHELM, M. D. RAFIQPOOR und P. FRANKENBERG (1996). *Die Klimate der Erde: Eine Klassifikation auf ökophysiologischer Grundlage der realen Vegetation (Climates of the Earth: A Classification on the Base of Eco-Physiological Characteristics of the Real Vegetation)*. *Erdkunde*, S. 275–300.
- LE QUÉRÉ, CORINNE, M. R. RAUPACH, J. G. CANADELL, G. MARLAND et al. (2009). *Trends in the sources and sinks of carbon dioxide*. *Nature Geoscience*, 2(12):831–836.
- LEAN, JUDITH L und D. H. RIND (2008). *How natural and anthropogenic influences alter global and regional surface temperatures: 1889 to 2006*. *Geophysical Research Letters*, 35(18):L18701.
- LEMPERT, ROBERT J und M. E. SCHLESINGER (2000). *Robust strategies for abating climate change*. *Climatic Change*, 45(3):387–401.

- LENDERINK, GEERT (2010). *Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations*. Climate research (Open Access for articles 4 years old and older), 44(2):151.
- LENTON, TIMOTHY M, H. HELD, E. KRIEGLER, J. W. HALL, W. LUCHT, S. RAHMSTORF und H. J. SCHELLNHUBER (2008). *Tipping elements in the Earth's climate system*. Proceedings of the National Academy of Sciences, 105(6):1786–1793.
- LEVERMANN, A, J. MIGNOT, S. NAWRATH und S. RAHMSTORF (2007). *The role of northern sea ice cover for the weakening of the thermohaline circulation under global warming*. Journal of Climate, 20(16):4160–4171.
- LINDEN, P VAN DER und J. MITCHELL (2009). *ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project*. Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK, 160.
- LOPEZ, ANA, C. TEBALDI, M. NEW, D. STAINFORTH, M. ALLEN und J. KETTLEBOROUGH (2006). *Two approaches to quantifying uncertainty in global temperature changes*. Journal of Climate, 19(19):4785–4796.
- LORENZ, PHILIP und D. JACOB (2010). *Validation of temperature trends in the ENSEMBLES regional climate model runs driven by ERA40*. Climate research (Open Access for articles 4 years old and older), 44(2):167.
- LU, JIAN und R. J. GREATBATCH (2002). *The changing relationship between the NAO and northern hemisphere climate variability*. Geophysical research letters, 29(7):52–1.
- LUTERBACHER, J, E. XOPLAKI, D. DIETRICH, P. JONES, T. DAVIES, D. PORTIS, J. GONZALEZ-ROUCO, H. VON STORCH, D. GYALISTRAS, C. CASTY et al. (2001). *Extending North Atlantic oscillation reconstructions back to 1500*. Atmospheric Science Letters, 2(1-4):114–124.
- LUTERBACHER, JÜRIG, C. SCHMUTZ, D. GYALISTRAS, E. XOPLAKI und H. WANNER (1999). *Reconstruction of monthly NAO and EU indices back to AD 1675*. Geophysical Research Letters, 26(17):2745–2748.
- MACADAM, I, A. PITMAN, P. WHETTON und G. ABRAMOWITZ (2010). *Ranking climate models by performance using actual values and anomalies: Implications for climate change impact assessments*. Geophysical Research Letters, 37(16):L16704.
- MAILIER, PASCAL J, D. B. STEPHENSON, C. A. FERRO und K. I. HODGES (2006). *Serial clustering of extratropical cyclones*. Monthly weather review, 134(8):2224–2240.

- MALBERG, H. (2007). *Meteorologie und Klimatologie: Eine Einführung*. Springer.
- MANN, MICHAEL E, M. A. CANE, S. E. ZEBIAK und A. CLEMENT (2005). *Volcanic and solar forcing of the tropical Pacific over the past 1000 years*. *Journal of Climate*, 18(3):447–456.
- MANNING, MR, J. EDMONDS, S. EMORI, A. GRUBLER, K. HIBBARD, F. JOOS, M. KAINUMA, R. KEELING, T. KRAM, A. MANNING et al. (2010). *Misrepresentation of the IPCC CO<sub>2</sub> emission scenarios*. *Nature Geoscience*, 3(6):376–377.
- MANTUA, NATHAN J und S. R. HARE (2002). *The Pacific decadal oscillation*. *Journal of Oceanography*, 58(1):35–44.
- MARES, C, I. MARES und M. MIHAILESCU (2003). *Testing of NAO and ENSO signals in the precipitation field in Europe*. In: *Climatic Change: Implications for the Hydrological Cycle and for Water Management*, S. 113–121. Springer.
- MARSH, W.M. und M. KAUFMAN (2012). *Physical Geography: Great Systems and Global Environments*. Cambridge University Press.
- MARSHALL, GARETH J (2003). *Trends in the Southern Annular Mode from observations and reanalyses*. *Journal of Climate*, 16(24):4134–4143.
- MARTIN, GM, C. DEARDEN, C. GREEVES, T. HINTON und P. INNESS (2004). *Evaluation of the Atmospheric Performance of HadGAM/GEM1*. Hadley Centre Technical Note, 54.
- MCBRIDE, JOHN L und N. NICHOLLS (1983). *Seasonal relationships between Australian rainfall and the Southern Oscillation*. *Monthly Weather Review*, 111(10):1998–2004.
- McFARLANE, NORM A, G. BOER, J. BLANCHET und M. LAZARE (1992). *The Canadian Climate Centre second-generation general circulation model and its equilibrium climate*. *Journal of Climate*, 5(10):1013–1044.
- MCKNIGHT, T.L. und D. HESS (2008). *Physical Geography: A Landscape Appreciation*. *Physical Geography: A Landscape Appreciation*. Pearson Prentice Hall.
- MEEHL, GERALD A, C. COVEY, K. E. TAYLOR, T. DELWORTH, R. J. STOFFER, M. LATIF, B. MCAVANEY und J. F. MITCHELL (2007). *The WCRP CMIP3 multimodel dataset: A new era in climate change research*. *Bulletin of the American Meteorological Society*, 88(9):1383–1394.

- MEEHL, GERALD A und W. M. WASHINGTON (1996). *El Niño-like climate change in a model with increased atmospheric CO<sub>2</sub> concentrations*. *Nature*, 382(6586):56–60.
- MEINSHAUSEN, M, T. WIGLEY und S. RAPER (2011). *Emulating atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6–Part 2: Applications*. *Atmospheric Chemistry and Physics*, 11(4):1457–1471.
- MILLER, RL, G. SCHMIDT und D. SHINDELL (2006). *Forced annular variations in the 20th century intergovernmental panel on climate change fourth assessment report models*. *Journal of Geophysical Research*, 111(D18):D18101.
- MITCHELL, TIMOTHY D und P. D. JONES (2005). *An improved method of constructing a database of monthly climate observations and associated high-resolution grids*. *International journal of climatology*, 25(6):693–712.
- MJELDE, JAMES W und J. F. GRIFFITHS (1998). *A review of current evidence on climate forecasts and their economic effects in agriculture*. *American Journal of Agricultural Economics*, 80(5):1089–1095.
- MOBLEY, C.D. und R. PREISENDORFER (1988). *Principal Component Analysis in Meteorology and Oceanography*. *Developments in atmospheric science*, 17. Elsevier.
- MOSS, RICHARD H, J. A. EDMONDS, K. A. HIBBARD, M. R. MANNING, S. K. ROSE, D. P. VAN VUUREN, T. R. CARTER, S. EMORI, M. KAINUMA, T. KRAM et al. (2010). *The next generation of scenarios for climate change research and assessment*. *Nature*, 463(7282):747–756.
- MUDELSEE, MANFRED (2010). *Climate time series analysis: classical statistical and bootstrap methods*, Bd. 42. Springer.
- MURPHY, JAMES M, B. BOOTH, M. COLLINS, G. HARRIS, D. SEXTON und M. WEBB (2007). *A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles*. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1857):1993–2028.
- MURPHY, JAMES M, D. M. SEXTON, D. N. BARNETT, G. S. JONES, M. J. WEBB, M. COLLINS und D. A. STAINFORTH (2004). *Quantification of modelling uncertainties in a large ensemble of climate change simulations*. *Nature*, 430(7001):768–772.
- NAKIĆENOVIĆ, NEBOJŠA, J. ALCAMO, G. DAVIS, B. DE VRIES, J. FENHANN, S. GAFFIN, K. GREGORY, A. GRÜBLER, T. JUNG, T. KRAM et al. (2000). *IPCC special report on emissions scenarios (SRES)*.

- NEW, MARK und M. HULME (2000). *Representing uncertainty in climate change scenarios: a Monte-Carlo approach*. *Integrated Assessment*, 1(3):203–213.
- NEWMAN, MATTHEW, G. P. COMPO und M. A. ALEXANDER (2003). *ENSO-forced variability of the Pacific Decadal Oscillation*. *Journal of Climate*, 16(23):3853–3857.
- OLDENBORGH, GJ VAN, S. PHILIP und M. COLLINS (2005). *El Nino in a changing climate: a multi-model study*. *Ocean Science*, 1(2):81–95.
- OLEA, R.A. (1999). *Geostatistics for Engineers and Earth Scientists*. Kluwer Academic.
- ONOGI, KAZUTOSHI, J. TSUTSUI, H. KOIDE, M. SAKAMOTO, S. KOBAYASHI, H. HATSUSHIKA, T. MATSUMOTO, N. YAMAZAKI, H. KAMAHORI, K. TAKAHASHI et al. (2007). *The JRA-25 reanalysis*. *J. Meteor. Soc. Japan*, 85(3):369–432.
- ORESQUES, NAOMI, K. SHRADER-FRECHETTE, K. BELITZ et al. (1994). *Verification, validation, and confirmation of numerical models in the earth sciences*. *Science*, 263(5147):641–646.
- PACHAURI, RAJENDRA K (2008). *Climate change 2007. Synthesis report. Contribution of Working Groups I, II and III to the fourth assessment report*.
- PAETH, H, A. HENSE, R. GLOWIENKA-HENSE, S. VOSS und U. CUBASCH (1999). *The North Atlantic Oscillation as an indicator for greenhouse-gas induced regional climate change*. *Climate Dynamics*, 15(12):953–960.
- PAETH, HEIKO, K. BORN, R. PODZUN und D. JACOB (2005). *Regional dynamical downscaling over West Africa: model evaluation and comparison of wet and dry years*. *Meteorologische Zeitschrift*, 14(3):349–367.
- PAETH, HEIKO und F. POLLINGER (2010). *Enhanced evidence in climate models for changes in extratropical atmospheric circulation*. *Tellus A*, 62(5):647–660.
- PAETH, HEIKO, A. SCHOLTEN, P. FRIEDERICHS und A. HENSE (2008). *Uncertainties in climate change prediction: El Niño-Southern Oscillation and monsoons*. *Global and Planetary Change*, 60(3):265–288.
- PAETH, HEIKO und H.-P. THAMM (2007). *Regional modelling of future African climate north of 15 S including greenhouse warming and land degradation*. *Climatic Change*, 83(3):401–427.

- PAVAN, V, S. TIBALDI und Č. BRANKOVIĆ (2000). *Seasonal prediction of blocking frequency: Results from winter ensemble experiments*. Quarterly Journal of the Royal Meteorological Society, 126(567):2125–2142.
- PAXIAN, A, E. HERTIG, G. VOGT, S. SEUBERT, J. JACOBET und H. PAETH (2013). *Greenhouse gas-related predictability of regional climate model trends in the Mediterranean area*. International Journal of Climatology.
- PEEL, MURRAY C, B. L. FINLAYSON, T. A. MCMAHON et al. (2007). *Updated world map of the Köppen-Geiger climate classification*. Hydrology and Earth System Sciences Discussions Discussions, 4(2):439–473.
- PERKINS, SARAH E, D. B. IRVING, J. R. BROWN, S. B. POWER, A. F. MOISE, R. A. COLMAN und I. SMITH (2012). *CMIP3 ensemble climate projections over the western tropical Pacific based on model skill*. Climate Research, 51(1):35.
- PERKINS, SE, A. PITMAN, N. HOLBROOK und J. MCANENEY (2007). *Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions*. Journal of climate, 20(17):4356–4376.
- PHILANDER, S GEORGE (1990). *El Niño, La Niña, and the southern oscillation*, Bd. 46. Academic press San Diego.
- PIANI, C, D. FRAME, D. STAINFORTH und M. ALLEN (2005). *Constraints on climate change from a multi-thousand member ensemble of simulations*. Geophysical Research Letters, 32(23).
- PIERCE, DAVID W, T. P. BARNETT, B. D. SANTER und P. J. GLECKLER (2009). *Selecting global climate models for regional climate change studies*. Proceedings of the National Academy of Sciences, 106(21):8441–8446.
- PIERREHUMBERT, RAYMOND T (2010). *Principles of planetary climate*. Cambridge University Press.
- PINCUS, ROBERT, C. P. BATSTONE, R. J. P. HOFMANN, K. E. TAYLOR und P. J. GLECKER (2008). *Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models*. Journal of Geophysical Research, 113(D14):D14209.
- POPE, VD, M. GALLANI, P. ROWNTREE und R. STRATTON (2000). *The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3*. Climate Dynamics, 16(2-3):123–146.

- POWER, S, T. CASEY, C. FOLLAND, A. COLMAN und V. MEHTA (1999). *Interdecadal modulation of the impact of ENSO on Australia*. *Climate Dynamics*, 15(5):319–324.
- PRESS, W.H., S. TEUKOLSKY, W. VETTERLING und F. B.P. (1992). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*. Nr. Bd. 1 in *Fortran Numerical Recipes: The Art of Scientific Computing*. *Numerical Recipes in Fortran 77*. Cambridge University Press.
- PRICE, DAVID T, D. W. MCKENNEY, I. A. NALDER, M. F. HUTCHINSON und J. L. KESTEVEN (2000). *A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data*. *Agricultural and Forest meteorology*, 101(2):81–94.
- PRUSCHA, H. (2007). *Statistisches Methodenbuch: Verfahren, Fallstudien, Programmcodes*. Statistik Und Ihre Anwendungen. Springer London, Limited.
- QUADRELLI, ROBERTA und J. M. WALLACE (2004). *A simplified linear framework for interpreting patterns of Northern Hemisphere wintertime climate variability*. *Journal of Climate*, 17(19):3728–3744.
- RÄISÄNEN, JOUNI (2007). *How reliable are climate models?*. *Tellus A*, 59(1):2–29.
- RÄISÄNEN, JOUNI und T. PALMER (2001). *A probability and decision-model analysis of a multimodel ensemble of climate change simulations*. *Journal of Climate*, 14(15):3212–3226.
- RÄISÄNEN, JOUNI, L. RUOKOLAINEN und J. YLHÄISI (2010). *Weighting of model results for improving best estimates of climate change*. *Climate dynamics*, 35(2-3):407–422.
- RÄISÄNEN, JOUNI und J. S. YLHÄISI (2012). *Can model weighting improve probabilistic projections of climate change?*. *Climate dynamics*, 39(7-8):1981–1998.
- RAPER, SARAH CB, J. M. GREGORY und R. J. STOUFFER (2002). *The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response*. *Journal of Climate*, 15(1):124–130.
- RAPHAEL, MARILYN N und M. M. HOLLAND (2006). *Twentieth century simulation of the Southern Hemisphere climate in coupled models. Part 1: Large scale circulation variability*. *Climate Dynamics*, 26(2-3):217–228.
- RASMUSSEN, EUGENE M und T. H. CARPENTER (1982). *Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño*. *Monthly Weather Review*, 110(5):354–384.

- RAUTHE, MONIKA, A. HENSE und H. PAETH (2004). *A model intercomparison study of climate change-signals in extratropical circulation*. International Journal of Climatology, 24(5):643–662.
- RAYNER, J.N. (1971). *An introduction to spectral analysis*. Monographs in spatial and environmental systems analysis. Pion.
- REICHLER, THOMAS und J. KIM (2008). *How well do coupled models simulate today's climate?*. Bulletin of the American Meteorological Society, 89(3):303–311.
- REIFEN, C und R. TOUMI (2009). *Climate projections: Past performance no guarantee of future skill?*. Geophysical Research Letters, 36(13).
- RICHMAN, MICHAEL B (1986). *Rotation of principal components*. Journal of climatology, 6(3):293–335.
- RIND, D, J. PERLWITZ und P. LONERGAN (2005). *AO/NAO response to climate change: 1. Respective influences of stratospheric and tropospheric climate changes*. Journal of geophysical research, 110(D12):D12107.
- ROECKNER, E, K. ARPE, L. BENGTSSON, M. CHRISTOPH, M. CLAUSSEN, L. DÜMENIL, M. ESCH, M. GIORGETTA, U. SCHLESE und U. SCHULZWEIDA (1996). *The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate*. Max-Planck-Institut für Meteorologie Rep, 218:90.
- ROECKNER, E, G. BÄUML, L. BONAVENTURA, R. BROKOPF, M. ESCH, M. GIORGETTA, S. HAGEMANN, I. KIRCHNER, L. KORNBLUEH, E. MANZINI, A. RHODUN, U. SCHLESE, U. SCHULZWEIDA und A. TOMPHINS (2003). *The Atmospheric General Circulation Model ECHAM5. Part I: Model Description*. MPI Report, 349:127.
- ROPELEWSKI, CF und M. HALPERT (1986). *North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO)*. Monthly Weather Review, 114(12):2352–2362.
- ROPELEWSKI, CHESTER F und M. S. HALPERT (1987). *Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation*. Monthly weather review, 115(8):1606–1626.
- RUBEL, FRANZ und M. KOTTEK (2010). *Observed and projected climate shifts 19012100 depicted by world maps of the Koppen-Geiger climate classification*. Meteorologische Zeitschrift, 19(2):135–141.

- RUDOLF, BRUNO und U. SCHNEIDER (2005). *Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations*. In: *Proceedings of the 2nd workshop of the International Precipitation Working Group IPWG*, S. 231–247.
- RUSSELL, GARY L, J. R. MILLER und D. RIND (1995). *A coupled atmosphere-ocean model for transient climate change studies*. *Atmosphere-ocean*, 33(4):683–730.
- SAHA, SURANJANA, S. MOORTHY, H.-L. PAN, X. WU, J. WANG, S. NADIGA, P. TRIPP, R. KISTLER, J. WOOLLEN, D. BEHRINGER et al. (2010). *The NCEP climate forecast system reanalysis*. *Bulletin of the American Meteorological Society*, 91(8):1015–1057.
- SALISBURY, JI und M. WIMBUSH (1999). *Using modern time series analysis techniques to predict ENSO events from the SOI time series*. *Nonlinear Processes in Geophysics*, 9(3/4):341–345.
- SAMANTA, SAILESH, D. K. PAL, D. LOHAR und B. PAL (2012). *Interpolation of climate variables and temperature modeling*. *Theoretical and Applied Climatology*, 107(1-2):35–45.
- SANDERSON, BENJAMIN M, K. M. SHELL und W. INGRAM (2010). *Climate feedbacks determined using radiative kernels in a multi-thousand member ensemble of AOGCMs*. *Climate dynamics*, 35(7-8):1219–1236.
- SCHMIDT, GAVIN A (2007). *The physics of climate modeling*. *Physics Today*, 60:72.
- SCHMIDT, GAVIN A, R. RUEDY, J. E. HANSEN, I. ALEINOV, N. BELL, M. BAUER, S. BAUER, B. CAIRNS, V. CANUTO, Y. CHENG et al. (2006). *Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data*. *Journal of Climate*, 19(2):153–192.
- SCHMITTNER, ANDREAS, M. LATIF und B. SCHNEIDER (2005). *Model projections of the North Atlantic thermohaline circulation for the 21st century assessed by observations*. *Geophysical Research Letters*, 32(23).
- SCHNEIDER, UDO, T. FUCHS, A. MEYER-CHRISTOFFER und B. RUDOLF (2008). *Global precipitation analysis products of the GPCC*. Global Precipitation Climatology Centre (GPCC), DWD, Internet Publikation, S. 1–12.
- SCHÖNWIESE, C.D. (2006). *Praktische Statistik für Meteorologen und Geowissenschaftler*. Borntraeger Gebrueder.

- SCHÖNWIESE, C.D. (2008). *Klimatologie*. Uni-Taschenbücher M. UTB GmbH.
- SCHULTZ, J. (2000). *Handbuch der Ökozonen*. Uni-Taschenbücher L. UTB GmbH.
- SCOCCIMARRO, ENRICO, S. GUALDI, P. G. FOGLI, E. MANZINI, A. GREZIO und A. NAVARRA (2007). *INGV-SXG: A Coupled Atmosphere Ocean Sea-Ice General Circulation Climate Model*. CMCC Research Paper, 15.
- SEIERSTAD, IA, D. STEPHENSON und N. KVAMSTØ (2007). *How useful are teleconnection patterns for explaining variability in extratropical storminess?*. *Tellus A*, 59(2):170–181.
- SHABBAR, AMIR und M. KHANDEKAR (1996). *The impact of el Nino-Southern oscillation on the temperature field over Canada: Research note*. *Atmosphere-Ocean*, 34(2):401–416.
- SHIBATA, KIYOTAKA, H. YOSHIMURA, M. OHIZUMI, M. HOSAKA und M. SUGI (1999). *A simulation of troposphere, stratosphere and mesosphere with an MRI/JMA98 GCM*. *Papers in Meteorology and Geophysics*, 50(1):15–53.
- SHUKLA, J und D. A. PAOLINO (1983). *The Southern Oscillation and long-range forecasting of the summer monsoon rainfall over India*. *Monthly Weather Review*, 111(9):1830–1837.
- SIMMONS, ADRIAN, S. UPPALA, D. DEE und S. KOBAYASHI (2007). *ERA-Interim: New ECMWF reanalysis products from 1989 onwards*. *ECMWF newsletter*, 110(110):25–35.
- SIMMONS, AJ, P. JONES, V. DA COSTA BECHTOLD, A. BELJAARS, P. KÅLLBERG, S. SAARINEN, S. UPPALA, P. VITERBO und N. WEDI (2004). *Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature*. *Journal of Geophysical Research: Atmospheres* (1984–2012), 109(D24).
- SIMPKINS, GRAHAM R und A. Y. KARPECHKO (2012). *Sensitivity of the southern annular mode to greenhouse gas emission scenarios*. *Climate dynamics*, 38(3-4):563–572.
- SIVAKUMAR, MVK, H. DAS und O. BRUNINI (2005). *Impacts of present and future climate variability and change on agriculture and forestry in the arid and semi-arid tropics*. *Climatic Change*, 70(1-2):31–72.
- SKELLY, WILLIAM C und A. HENDERSON-SELLERS (1996). *Grid box or grid point: What type of data do GCMs deliver to climate impacts researchers?*. *International Journal of Climatology*, 16(10):1079–1086.

- SMITH, DOUG M, S. CUSACK, A. W. COLMAN, C. K. FOLLAND, G. R. HARRIS und J. M. MURPHY (2007). *Improved surface temperature prediction for the coming decade from a global climate model*. *Science*, 317(5839):796–799.
- SMITH, IAN und E. CHANDLER (2010). *Refining rainfall projections for the Murray Darling Basin of south-east Australia – the effect of sampling model results based on performance*. *Climatic Change*, 102(3-4):377–393.
- SMITH, JOEL B, S. H. SCHNEIDER, M. OPPENHEIMER, G. W. YOHE, W. HARE, M. D. MASTRANDREA, A. PATWARDHAN, I. BURTON, J. CORFEE-MORLOT, C. H. MAGADZA et al. (2009). *Assessing dangerous climate change through an update of the Intergovernmental Panel on Climate Change (IPCC) 'reasons for concern'*. *Proceedings of the National Academy of Sciences*, 106(11):4133–4137.
- SMITH, SHAWN R, D. M. LEGLER und K. V. VERZONE (2001). *Quantifying uncertainties in NCEP reanalyses using high-quality research vessel observations*. *Journal of Climate*, 14(20):4062–4072.
- SOKOLOV, ANDREI P und E. MONIER (2012). *Changing the climate sensitivity of an atmospheric general circulation model through cloud radiative adjustment*. *Journal of Climate*, 25(19):6567–6584.
- STENSETH, NILS C, A. MYSTERUD, G. OTTERSEN, J. W. HURRELL, K.-S. CHAN und M. LIMA (2002). *Ecological effects of climate fluctuations*. *Science*, 297(5585):1292–1296.
- STENSETH, NILS CHR, G. OTTERSEN, J. W. HURRELL, A. MYSTERUD, M. LIMA, K.-S. CHAN, N. G. YOCCOZ und B. ÅDLANDSVIK (2003). *Studying climate effects on ecology through the use of climate indices: the North Atlantic Oscillation, El Nino Southern Oscillation and beyond*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1529):2087–2096.
- STEPHENSON, DB, C. COELHO, F. DOBLAS-REYES und M. BALMASEDA (2005). *Forecast assimilation: a unified framework for the combination of multi-model weather and climate predictions*. *Tellus A*, 57(3):253–264.
- STEPHENSON, DB, A. HANNACHI und A. O'NEILL (2004). *On the existence of multiple climate regimes*. *Quarterly Journal of the Royal Meteorological Society*, 130(597):583–605.
- STORCH, HANS VON und F. W. ZWIERS (1999). *Statistical analysis in climate research*. Cambridge University Press.
- STRAHLER, A.H. (2004). *Physical Geography: Science and Systems of the Human Environment*. John Wiley & Sons.

- SULTAN, BENJAMIN, S. JANICOT und A. DIEDHIOU (2003). *The West African monsoon dynamics. Part I: Documentation of intraseasonal variability*. Journal of Climate, 16(21):3389–3406.
- SUPPIAH, R (1992). *The Australian summer monsoon: a review*. Progress in Physical Geography, 16(3):283–318.
- SUPPIAH, R, K. HENNESSY, P. WHETTON, K. MCINNES, I. MACADAM, J. BATHOLS, J. RICKETTS und C. PAGE (2007). *Australian climate change projections derived from simulations performed for the IPCC 4th Assessment Report*. Australian Meteorological Magazine, 56(3):131–152.
- TAYLOR, FW (1991). *The greenhouse effect and climate change*. Reports on Progress in Physics, 54(6):881.
- TAYLOR, FW (2002). *The greenhouse effect and climate change revisited*. Reports on Progress in Physics, 65(1):1.
- TAYLOR, KARL E (2001). *Summarizing multiple aspects of model performance in a single diagram*. Journal of Geophysical Research: Atmospheres (1984–2012), 106(D7):7183–7192.
- TEBALDI, CLAUDIA und R. KNUTTI (2007). *The use of the multi-model ensemble in probabilistic climate projections*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1857):2053–2075.
- TEBALDI, CLAUDIA, R. L. SMITH, D. NYCHKA und L. O. MEARNES (2005). *Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles*. Journal of Climate, 18(10):1524–1540.
- THOMPSON, DAVID WJ, J. J. KENNEDY, J. M. WALLACE und P. D. JONES (2008). *A large discontinuity in the mid-twentieth century in observed global-mean surface temperature*. Nature, 453(7195):646–649.
- THOMPSON, DAVID WJ, S. LEE und M. P. BALDWIN (2003). *Atmospheric processes governing the northern hemisphere annular mode/North Atlantic oscillation*. Geophysical Monograph-American Geophysical Union, 134:81–112.
- THOMPSON, DAVID WJ und S. SOLOMON (2002). *Interpretation of recent Southern Hemisphere climate change*. Science, 296(5569):895–899.
- THOMPSON, DAVID WJ und J. M. WALLACE (1998). *The Arctic Oscillation signature in the wintertime geopotential height and temperature fields*. Geophysical Research Letters, 25(9):1297–1300.

- THOMPSON, DAVID WJ und J. M. WALLACE (2000). *Annular modes in the extratropical circulation. Part I: month-to-month variability\**. Journal of Climate, 13(5):1000–1016.
- THOMPSON, DAVID WJ und J. M. WALLACE (2001). *Regional climate impacts of the Northern Hemisphere annular mode*. Science, 293(5527):85–89.
- THOMPSON, DAVID WJ, J. M. WALLACE und G. C. HEGERL (2000). *Annular modes in the extratropical circulation. Part II: Trends*. Journal of Climate, 13(5):1018–1036.
- TIMMERMANN, AXEL, J. OBERHUBER, A. BACHER, M. ESCH, M. LATIF und E. ROECKNER (1999). *Increased El Niño frequency in a climate model forced by future greenhouse warming*. Nature, 398(6729):694–697.
- TORRENCE, CHRISTOPHER und P. J. WEBSTER (1999). *Interdecadal changes in the ENSO-monsoon system*. Journal of Climate, 12(8):2679–2690.
- TRENBERTH, K.E. (1992). *Climate System Modeling*. Cambridge University Press.
- TRENBERTH, KEVIN E (1984). *Signal versus noise in the Southern Oscillation*. Monthly Weather Review, 112(2):326–332.
- TRENBERTH, KEVIN E (1997). *The definition of el nino*. Bulletin of the American Meteorological Society, 78(12):2771–2777.
- TRENBERTH, KEVIN E und J. M. CARON (2000). *The Southern Oscillation revisited: Sea level pressures, surface temperatures, and precipitation*. Journal of Climate, 13(24):4358–4365.
- TRENBERTH, KEVIN E, J. M. CARON, D. P. STEPANIAK und S. WORLEY (2002). *Evolution of El Niño–Southern Oscillation and global atmospheric surface temperatures*. Journal of Geophysical Research, 107(D8):4065.
- TRENBERTH, KEVIN E und T. J. HOAR (1997). *El Niño and climate change*. Geophysical Research Letters, 24(23):3057–3060.
- TRENBERTH, KEVIN E und J. W. HURRELL (1994). *Decadal atmosphere-ocean variations in the Pacific*. Climate Dynamics, 9(6):303–319.
- TRENBERTH, KEVIN E und D. A. PAOLINO JR (1980). *The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities*. Monthly Weather Review, 108(7):855–872.

- TRENBERTH, KEVIN E, D. P. STEPANIAK und L. SMITH (2005). *Interannual variability of patterns of atmospheric mass distribution*. Journal of climate, 18(15):2812–2825.
- TRIGO, RICARDO M, T. J. OSBORN, J. M. CORTE-REAL et al. (2002). *The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms*. Climate Research, 20(1):9–17.
- TROLL, CARL (1964). *Karte der Jahreszeiten-Klimate der Erde (The Map of the Seasonal Climates of the Earth)*. Erdkunde, S. 5–28.
- TSONIS, ANASTASIOS A, K. SWANSON und S. KRAVTSOV (2007). *A new dynamical mechanism for major climate shifts*. Geophysical Research Letters, 34(13).
- TURNER, ANDREW G und H. ANNAMALAI (2012). *Climate change and the South Asian summer monsoon*. Nature Climate Change, 2(8):587–595.
- TURNER, JOHN, S. R. COLWELL, G. J. MARSHALL, T. A. LACHLAN-COPE, A. M. CARLETON, P. D. JONES, V. LAGUN, P. A. REID und S. IAGOVKINA (2005). *Antarctic climate change during the last 50 years*. International journal of Climatology, 25(3):279–294.
- UPPALA, SAKARI M, P. KÅLLBERG, A. SIMMONS, U. ANDRAE, V. BECHTOLD, M. FIORINO, J. GIBSON, J. HASELER, A. HERNANDEZ, G. KELLY et al. (2005). *The ERA-40 re-analysis*. Quarterly Journal of the Royal Meteorological Society, 131(612):2961–3012.
- VAN VUUREN, DETLEF P, J. EDMONDS, M. KAINUMA, K. RIAHI, A. THOMSON, K. HIBBARD, G. C. HURTT, T. KRAM, V. KREY, J.-F. LAMARQUE et al. (2011). *The representative concentration pathways: an overview*. Climatic Change, 109(1-2):5–31.
- VINTHER, BM, A. HANSEN und H. VON STORCH (2003). *A major deviation from the NAO temperature seesaw pattern*. Geophysical Research Letters.
- VISSER, H, R. FOLKERT, J. HOEKSTRA und J. DE WOLFF (2000). *Identifying key sources of uncertainty in climate change projections*. Climatic Change, 45(3-4):421–457.
- WALKER, GILBERT T und E. BLISS (1932). *World weather*. V. Mem. Roy. Meteor. Soc., 4:53–84.
- WALKER, G.T. (1924). *Correlations in seasonal variations of weather*. IX. Mem. Ind. Meteor. Dept., 24:275–332.

- WALLACE, JOHN M (2000). *North atlantic oscillation/annular mode: Two paradigms—one phenomenon*. Quarterly Journal of the Royal Meteorological Society, 126(564):791–805.
- WALLACE, JOHN M und D. S. GUTZLER (1981). *Teleconnections in the geopotential height field during the Northern Hemisphere winter*. Monthly Weather Review, 109(4):784–812.
- WANG, B. (2006). *The Asian Monsoon*. Springer Praxis Books / Environmental Sciences. Springer.
- WANG, BIN, H. WAN, Z. JI, X. ZHANG, R. YU, Y. YU und H. LIU (2004). *Design of a new dynamical core for global atmospheric models based on some efficient numerical methods*. Science in China Series A: Mathematics, 47(1):4–21.
- WANG, BIN, R. WU und X. FU (2000). *Pacific-East Asian teleconnection: how does ENSO affect East Asian climate?*. Journal of Climate, 13(9):1517–1536.
- WANG, BIN, R. WU und T. LI (2003). *Atmosphere-Warm Ocean Interaction and Its Impacts on Asian-Australian Monsoon Variation\**. Journal of Climate, 16(8):1195–1211.
- WANG, CHUNZAI und J. PICAUT (2004). *Understanding ENSO physics—A review*. Geophysical Monograph Series, 147:21–48.
- WANG, GELI, K. L. SWANSON und A. A. TSONIS (2009). *The pacemaker of major climate shifts*. Geophysical Research Letters, 36(7).
- WANNER, HEINZ, S. BRÖNNIMANN, C. CASTY, D. GYALISTRAS, J. LUTERBACHER, C. SCHMUTZ, D. B. STEPHENSON und E. XOPLAKI (2001). *North Atlantic Oscillation—concepts and studies*. Surveys in geophysics, 22(4):321–381.
- WASHINGTON, W.M. und C. PARKINSON (2005). *An Introduction to Three-dimensional Climate Modeling*. University Science Books.
- WASSMANN, R, S. JAGADISH, K. SUMFLETH, H. PATHAK, G. HOWELL, A. ISMAIL, R. SERRAJ, E. REDONA, R. SINGH und S. HEUER (2009). *Regional vulnerability of climate change impacts on Asian rice production and scope for adaptation*. Advances in Agronomy, 102:91–133.
- WATTERSON, IAN G (2008). *Calculation of probability density functions for temperature and precipitation change under global warming*. Journal of Geophysical Research: Atmospheres (1984–2012), 113(D12).

- WEBSTER, MORT DAVID, M. BABIKER, M. MAYER, J. M. REILLY, J. HARNISCH, R. HYMAN, M. C. SAROFIM und C. WANG (2002). *Uncertainty in emissions projections for climate models*. Atmospheric Environment, 36(22):3659–3670.
- WEBSTER, PETER J, V. O. MAGANA, T. PALMER, J. SHUKLA, R. TOMAS, M. YANAI und T. YASUNARI (1998). *Monsoons: Processes, predictability, and the prospects for prediction*. Journal of Geophysical Research: Oceans (1978–2012), 103(C7):14451–14510.
- WEBSTER, PETER J und S. YANG (1992). *Monsoon and ENSO: Selectively interactive systems*. Quarterly Journal of the Royal Meteorological Society, 118(507):877–926.
- WEIGEL, ANDREAS P, R. KNUTTI, M. A. LINIGER und C. APPENZELLER (2010). *Risks of model weighting in multimodel climate projections*. Journal of Climate, 23(15):4175–4191.
- WHETTON, PENNY, I. MACADAM, J. BATHOLS und J. O’GRADY (2007). *Assessment of the use of current climate patterns to evaluate regional enhanced greenhouse response patterns of climate models*. Geophysical Research Letters, 34(14):L14701.
- WIGLEY, TOM ML und S. C. RAPER (2001). *Interpretation of high projections for global-mean warming*. Science, 293(5529):451–454.
- WILKS, D.S. (2011). *Statistical Methods in the Atmospheric Sciences*. International Geophysics. Elsevier Science.
- WILLMOTT, CORT J, S. G. ACKLESON, R. E. DAVIS, J. J. FEDDEMA, K. M. KLINK, D. R. LEGATES, J. O’DONNELL und C. M. ROWE (1985). *Statistics for the evaluation and comparison of models*. Journal of Geophysical Research: Oceans (1978–2012), 90(C5):8995–9005.
- WOODROFFE, COLIN D, M. R. BEECH und M. K. GAGAN (2003). *Mid-late Holocene El Niño variability in the equatorial Pacific from coral microatolls*. Geophysical Research Letters, 30(7).
- YIOU, P, E. BAERT und M. LOUTRE (1996). *Spectral analysis of climate data*. Surveys in Geophysics, 17(6):619–663.
- YU, B und F. ZWIERS (2007). *The impact of combined ENSO and PDO on the PNA climate: a 1,000-year climate modeling study*. Climate dynamics, 29(7-8):837–851.
- YUAN, XIAOJUN (2004). *ENSO-related impacts on Antarctic sea ice: a synthesis of phenomenon and mechanisms*. Antarctic Science, 16(4):415–425.

- 
- YULAEVA, ELENA und J. M. WALLACE (1994). *The signature of ENSO in global temperature and precipitation fields derived from the microwave sounding unit*. Journal of climate, 7(11):1719–1736.
- ZHANG, RENHE, A. SUMI und M. KIMOTO (1999). *A diagnostic study of the impact of El Niño on the precipitation in China*. Advances in Atmospheric Sciences, 16(2):229–241.
- ZHANG, YUAN, J. M. WALLACE und D. S. BATTISTI (1997). *ENSO-like interdecadal variability: 1900-93*. Journal of climate, 10(5):1004–1020.



# Danksagung

Die vorgestellte Arbeit wurde am Institut für Geographie und Geologie in einem mehrjährigen Forschungsprozess erstellt. In Folge dessen gibt es eine ganze Reihe Personen bei denen ich mich an dieser Stelle bedanken möchte und muss.

An erster Stelle gilt mein Dank Herrn Prof. Dr. Heiko Paeth, der mich bei der Erstellung dieser Arbeit betreut und angeleitet hat. Besonders zu schätzen gelernt habe ich nicht nur seine fachliche Expertise, sondern auch seine Bereitschaft diese auf ungezwungene und eingängige Weise weiterzugeben. Auch bedanke ich mich für die immer vorhandene Bereitschaft und Ermutigung, bisherige Ansätze im Licht neuer Ergebnisse zu diskutieren, sowie neue Befunde weiterzuverfolgen, wo dies lohnenswert und machbar erschien. Weiterhin möchte ich mich hier ausdrücklich für das Vertrauen und die Geduld bedanken, die mir Prof. Dr. Paeth immer entgegengebracht hat.

Weiterhin danke ich Herrn Prof. Dr. Dieter Scherer für die Übernahme des zweiten Gutachtens.

Für das angenehme Arbeitsklima, dass ich während der Erstellung dieser Arbeit genießen konnte bedanke ich mich bei der gesamten, teilweise ehemaligen, Arbeitsgruppe Klimatologie des Instituts für Geographie und Geologie, namentlich meinen Mit-Promovierenden Hervé Awoyé, Tran Nam Binh, Birgit Mannig, Sebastian Mutz, Andreas Paxian, Christian Steger und Gernot Vogt sowie Doro Schill und den DiplomandInnen, M. Sc.-StudentInnen und Hiwis.

Aus der Vorbereitung sowie der Durchführung von Seminaren zur Statistischen Methodik und zum Themenkomplex des anthropogenen Klimawandels konnte ich fruchtbare Anregungen sowohl für den Inhalt dieser Arbeit als auch die Präsentation meiner Ergebnisse ziehen. Für diese wertvolle Hilfe danke ich den SeminarteilnehmerInnen.

Vergleichsweise fruchtlose aber äußerst willkommene Unterstützung habe ich aus meinem Freundeskreis erhalten, was mir auch die Freude an der Promotion immer wieder aufgeladen hat. Den Dank hierfür kann ich auf meine erweiterte Familie erweitern, der ich weiterhin für allgemeine und spezifische moralische und konkrete Unterstützung danken möchte. Hervorgehoben meinen Eltern.

An letzter und prominentester Stelle bedanke ich mich bei, meiner mittlerweile Frau, Katharina für Alles.



# Erklärung

Hiermit versichere ich, Felix Pollinger, geboren am 31.07.1981 in Erlenbach a. Main, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Würzburg, den 13. November 2013



# Curriculum Vitae

Felix Pollinger, geb. am 31.07.1981 in Erlenbach a. Main

## Wissenschaftlicher Werdegang

- Wissenschaftlicher Mitarbeiter am Institut für Geographie und Geologie der Bayerischen Julius-Maximilians-Universität Würzburg (seit 2009)
- Studiengang Diplom-Geographie mit den Nebenfächern Statistik und Geschichte an der Bayerischen Julius-Maximilians-Universität Würzburg (2004–2009; Diplom im Jahr 2009); Titel der Diplomarbeit:“Signalanalyse außertropischer Zirkulationsmodi“
- Studentische Hilfskraft am Institut für Ökologie (Fachbereich Klimatologie) der Technischen Universität Berlin (03/2009–08/2009)
- Studentische Hilfskraft im Umweltbundesamt, Fachgebiet Luftmessnetz (01/2008–02/2008)
- Studiengang Kulturgeographie (M.A.) mit den Nebenfächern Geschichte und Germanistik an der Bayerischen Julius-Maximilians-Universität Würzburg (2003–2004)
- Hermann-Staudinger-Gymnasium Erlenbach (Abitur im Jahr 2001)

## Publikationen

- Paeth, Heiko und Felix Pollinger (2010). Enhanced evidence in climate models for changes in extratropical atmospheric circulation. *Tellus A*, 62(5):647–660.

## Sonstiges

- Praktikum beim Umweltbundesamt, Fachbereich Luftmessnetz (10/2007–12/2007)
- Praktikum beim Landratsamt Schweinfurt, Umweltamt (03/2007–04/2007)
- Zivildienst in der Missionsärztlichen Klinik in Würzburg (2001–2002)