

**Repräsentation atmosphärischer
Dynamik in dekadischen
Vorhersagemodellen**

Dissertation zur Erlangung des
naturwissenschaftlichen Doktorgrades
an der Fakultät für Angewandte Informatik
der Universität Augsburg

vorgelegt von
Benjamin Lang

Augsburg 2018

Erstgutachter:

Prof. Dr. Jucundus Jacobeit

Zweitgutachter:

PD Dr. Christoph Beck

Tag der mündlichen Prüfung:

19.09.2018

Meiner kleinen und
meiner großen Familie

Inhaltsverzeichnis

Inhaltsverzeichnis		I
Abbildungsverzeichnis		IV
Tabellenverzeichnis		X
Abkürzungsverzeichnis		XI
Abstract		XIV
1 Einleitung		1
1.1 Motivation		1
1.2 Fragestellungen und Zielsetzung		2
1.3 Aufbau der Arbeit.....		4
2 Forschungsstand		5
2.1 Forschungsgrundlagen.....		5
2.2 Aktuelle Forschungsaktivitäten		7
3 Datengrundlage		10
3.1 Modelldaten.....		10
3.1.1 Dekadisches Vorhersagesystem (Initialisiert).....		12
3.1.2 Historische Läufe (Uninitialisiert).....		15
3.2 Beobachtungsdaten.....		15
3.2.1 Datensätze		16
3.2.2 Konsistenz der Reanalysen		17
4 Methoden		20
4.1 Untersuchungsgegenstand		20

4.1.1	Telekonnektionsindizes	21
4.1.2	Zirkulationstypen	35
4.2	Evaluation	48
4.2.1	Hindcasts	48
4.2.2	Skillmetriken	50
5	Ergebnisse: Telekonnektionsindizes	54
5.1	Die uninitialisierten Läufe	55
5.1.1	Nordatlantische Oszillation (nach Hurrell)	55
5.1.2	CPC-Telekonnektionsmuster	57
5.1.3	Südliche Oszillation (SO)	60
5.1.4	Antarktische Oszillation (AAO)	61
5.2	Der Einfluss der Initialisierung	63
5.2.1	Differenzierung nach Modellgenerationen	64
5.2.2	Differenzierung nach Telekonnektionsindizes	66
5.2.3	Differenzierung nach Jahreszeiten	67
5.2.4	Differenzierung nach zeitlicher Auflösung	68
5.3	Der Einfluss der Ensemblegröße	70
5.3.1	Ensemblegröße und Modellgenerationen	71
5.3.2	Ensemblegröße und Telekonnektionsindizes	73
5.3.3	Ensemblegröße und Jahreszeiten	77
5.4	Vorhersageleistung des dekadischen Modells	79
5.4.1	Betrachtung ausgewählter Telekonnektionsindizes	79
5.4.2	Überblickender Vergleich der Modellgenerationen	83
5.5	Zusammenfassung	90
6	Ergebnisse: Zirkulationstypen	91
6.1	Die uninitialisierten Läufe	94
6.1.1	Luftdruck auf Meeresspiegel (SLP)	94
6.1.2	500 hPa-Geopotentialfläche (ZG500)	100
6.2	Der Einfluss der Initialisierung	101
6.2.1	Differenzierung nach Modellgenerationen	102

6.2.2	Differenzierung nach Klassifikationsverfahren	105
6.2.3	Differenzierung nach Jahreszeiten	106
6.3	Der Einfluss der Ensemblegröße	107
6.3.1	Ensemblegröße und Modellgeneration	108
6.3.2	Ensemblegröße und Jahreszeiten	111
6.3.3	Ensemblegröße und Variablen	112
6.3.4	Ensemblegröße und Klassenanzahl	114
6.4	Vorhersageleistung des dekadischen Modells	115
6.5	Zusammenfassung	124
7	Zusammenfassung und Ausblick	125
	Anhang	131
	Literaturverzeichnis	140

Abbildungsverzeichnis

Abbildung 3.1: Das Klimamodell MPI-ESM und seine Komponenten (Quelle: Giorgetta et al. 2013).....	10
Abbildung 3.2: Index der Nordatlantischen Oszillation im Winter (DJF) des Zeitraums 1980-2011 (PC-basiert nach Hurrell 1995 (NAO-PC); vgl. Tabelle 4.1).....	17
Abbildung 3.3: Saisonale gemittelte Auftrittshäufigkeiten von acht Zirkulationstypen (Großwettertypen nach Beck; vgl. Kapitel 4.1.2.1) im nordatlantisch-europäischen Raum (90°W-40°O, 20°N-80°N) im Frühjahr (JJA) für verschiedene Reanalysen (CFSR, ERAINT, JRA25, MERRA).	18
Abbildung 3.4: Saisonale Auftrittshäufigkeiten (Zeitraum 1979-2011) von acht Zirkulationstypen (Großwettertypen nach Beck; vgl. Kapitel 4.1.2.1) im nordatlantisch-europäischen Raum (90°W-40°O, 20°N-80°N) im Frühjahr (MAM), dargestellt für die vier Reanalysen (a) und vier zufällig ausgewählte Ensemblemitglieder von b1 (b). Die Farbgebung für die einzelnen Zirkulationstypen orientierte sich an derjenigen von Abbildung 3.3.	19
Abbildung 4.1: Schema des (globalen) Klimasystems (Quelle: Schönwiese 2013).....	21
Abbildung 4.2: Räumliches Muster der Nordatlantischen Oszillation (NAO-PC) im Winter (DJF).	26
Abbildung 4.3: Vergleich der beiden Berechnungsmethoden NAO-GP und NAO-PC für die Jahre 1980-2011 (basierend auf ERAINT) im Winter (DJF) unter Angabe des Korrelationskoeffizienten nach Pearson.	27
Abbildung 4.4: Übersicht über die neun ausgewählten Telekonnektionsmuster nach der CPC-Methode im Winter (Datengrundlage: ERAINT 1980-2011).	30
Abbildung 4.5: Telekonnektionsmuster der Antarktischen Oszillation (AAO-PC) im Sommer (Datengrundlage: ERAINT 1979-2011).	33
Abbildung 4.6: Vergleich der beiden Berechnungsmethoden NAO-GP und NAO-PC für die Jahre 1979-2011 (basierend auf ERAINT) im Sommer (JJA) unter Angabe des Korrelationskoeffizienten nach Pearson.	34
Abbildung 4.7: Lage und Ausdehnung der nordatlantisch-europäischen Untersuchungsregion.	36
Abbildung 4.8: Idealtypische Grundmuster der Großwettertypen nach Beck (Quelle: Beck 2000).....	37
Abbildung 4.9: Zentroide der Zirkulationstypen (basierend auf GWT) in der nordatlantisch-europäischen Region im Winter (DJF) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Auftrittshäufigkeit (Datengrundlage: SLP-Anomalien, ERAINT). ..	39
Abbildung 4.10: Zentroide der Zirkulationstypen (basierend auf SAN) in der nordatlantisch-europäischen Region im Winter (DJF) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Auftrittshäufigkeit (Datengrundlage: SLP-Anomalien, ERAINT). ..	43

<i>Abbildung 4.11: Zentroide der Zirkulationstypen (basierend auf TPCA) in der nordatlantisch-europäischen Region im Winter (DJF) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Aufttrittshäufigkeit (Datengrundlage: SLP-Anomalien, ERAINT). ..</i>	47
<i>Abbildung 4.12: Das Prinzip der Hindcasts für zwei Beispiele (links: decadal_1978; rechts: decadal:1979) mit Angabe der Leadjahre 1-10.</i>	48
<i>Abbildung 5.1: (Standardisierter) Index der Nordatlantischen Oszillation (NAO-PC) im Winter (DJF) des Untersuchungszeitraums, dargestellt für die historischen, uninitialisierten Modellläufe (hist; s. Erklärungen im Text) und die beobachteten Werte (obs). Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. Der Korrelationskoeffizient drückt den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.</i>	56
<i>Abbildung 5.2: Wie Abbildung 5.1, aber für NAO-GP.</i>	57
<i>Abbildung 5.3: (Standardisierte) Indizes der CPC-Telekonnektionsmuster (jeweils für die Jahreszeit mit dem höchsten Varianzerklärungsanteil im Untersuchungszeitraum), dargestellt für die historischen, uninitialisierten Modellläufe (hist; s. Erklärungen im Text) und die beobachteten Werte (obs). Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. Die Korrelationskoeffizienten drücken jeweils den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.....</i>	59
<i>Abbildung 5.4: Standardisierter Index der Südlichen Oszillation (SOI) im Frühjahr (MAM), Sommer (JJA), Herbst (SON) und Winter (DJF) des Untersuchungszeitraums, dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs). Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. Die Korrelationskoeffizienten drücken jeweils den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.</i>	60
<i>Abbildung 5.5: (Standardisierter) Index der Antarktischen Oszillation (AAO-PC) im Winter (DJF) des Untersuchungszeitraums, dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs). Der Korrelationskoeffizient drückt den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.</i>	61
<i>Abbildung 5.6: Wie Abbildung 5.5, aber für AAO-ZM.</i>	62
<i>Abbildung 5.7: CORR (a) und MSE (b) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT hinsichtlich der Telekonnektionsindizes in Abhängigkeit von verschiedenen Leadjahren.</i>	64
<i>Abbildung 5.8: Wie Abbildung 5.7, aber für die unterschiedlichen Telekonnektionsindizes....</i>	67
<i>Abbildung 5.9: Wie Abbildung 5.7, aber für die unterschiedlichen Jahreszeiten.....</i>	68

- Abbildung 5.10:** Wie Abbildung 5.7, aber für die unterschiedlichen zeitlichen Auflösungen. . 69
- Abbildung 5.11:** **CORR** zwischen den unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) und ERAINT hinsichtlich der ausgewählten Telekonnektionsindizes in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (a) bzw. das zweite Leadjahr (b) nach Initialisierung..... 71
- Abbildung 5.12:** **MSE** für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT hinsichtlich der ausgewählten Telekonnektionsindizes in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (a) bzw. das zweite Leadjahr (b) nach Initialisierung. 72
- Abbildung 5.13:** Wie Abbildung 5.11 (**CORR**), aber lediglich für den **Southern Oscillation Index** (alle anderen Indizes gingen nicht in die Mittelung ein). 73
- Abbildung 5.14:** Wie Abbildung 5.12 (**MSE**), aber für den **Southern Oscillation Index** (alle anderen Indizes gingen nicht in die Mittelung ein). 74
- Abbildung 5.15:** Wie Abbildung 5.11 (**CORR**), aber für den **PC-basierten Index der Antarktischen Oszillation** (alle anderen Indizes gingen nicht in die Mittelung ein)..... 75
- Abbildung 5.16:** Wie Abbildung 5.12 (**MSE**), aber für den **PC-basierten Index der Antarktischen Oszillation** (alle anderen Indizes gingen nicht in die Mittelung ein)..... 76
- Abbildung 5.17:** **CORR** für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT hinsichtlich der ausgewählten Telekonnektionsindizes in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für verschiedene Jahreszeiten (a: MAM, b: JJA, c: SON, d: DJF) des ersten Leadjahrs nach Initialisierung..... 77
- Abbildung 5.18:** Saisonale aggregierter **SOI** im Frühjahr (a) bzw. **NAO** (nach Hurrell) im Sommer (b) des ersten Leadjahres für die Beobachtung (obs: ERAINT) und die gemittelte Vorhersage (fc: **pS**) unter Angabe des Korrelationskoeffizienten. Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. 80
- Abbildung 5.19:** Saisonale aggregierter **AAO-PC** im Sommer für das erste (a) bzw. das zweite Leadjahr (b) für die Beobachtung (obs: ERAINT) und die gemittelte Vorhersage (fc: **pS**) unter Angabe des Korrelationskoeffizienten. Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. 82
- Abbildung 5.20:** **MSSS** (a) und **RPSS** (b) jeweils für (von oben nach unten) **b0 – hist**, **b1 – hist**, **pS – hist** und **pG – hist** für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse der saisonal aggregierten **Indizes** (Abkürzungen s. Tabelle 4.1) sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF, ALL). Bei schwarz hinterlegten Flächen sind keine Daten vorhanden. 84

Abbildung 5.21: MSSS (a) und RPSS (b) jeweils für (von oben nach unten) $b1 - b0$, $pS - b0$ und $pG - b0$ für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse der saisonal aggregierten Indizes (Abkürzungen s. Tabelle 4.1) sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF, ALL). Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.	85
Abbildung 5.22: MSSS (a) und RPSS (b) jeweils für (von oben nach unten) $pS - b1$, $pG - b1$ und $pG - pS$ für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse der saisonal aggregierten Indizes (Abkürzungen s. Tabelle 4.1) sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF, ALL). Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.	86
Abbildung 6.1: Auftrittshäufigkeit (a), Persistenz (b) und Intensität (c) von 8 Zirkulationstypen (nach GWT) im nordatlantisch-europäischen Raum über alle Jahreszeiten des Untersuchungszeitraums (1979-2011, gemittelt) basierend auf dem Luftdruck auf Meereshöhe (SLP). Die Ergebnisse sind dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs).....	95
Abbildung 6.2: Auftrittshäufigkeit von 8 Zirkulationstypen (nach GWT) im nordatlantisch-europäischen Raum für den Sommer (JJA) des Untersuchungszeitraums (1979-2011, gemittelt) basierend auf dem Luftdruck auf Meereshöhe (SLP). Die Ergebnisse sind dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs).	97
Abbildung 6.3: Wie Abbildung 6.2, aber für SAN (a) bzw. TPCA (b).	98
Abbildung 6.4: Auftrittshäufigkeit (a), Persistenz (b) und Intensität (c) von 8 Zirkulationstypen (nach GWT) im nordatlantisch-europäischen Raum über alle Jahreszeiten des Untersuchungszeitraums (1979-2011, gemittelt) basierend auf der 500hPa-Geopotentialfläche (ZG500). Die Ergebnisse sind dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs).	100
Abbildung 6.5: MSE hinsichtlich der Auftrittshäufigkeit bei Mittelung über den gesamten Untersuchungszeitraum (a) bzw. bei saisonaler Aggregation (b) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von verschiedenen Leadjahren.	103
Abbildung 6.6: Wie Abbildung 6.5, aber für die unterschiedlichen Klassifikationsverfahren	105
Abbildung 6.7: Wie Abbildung 6.5, aber für die unterschiedlichen Jahreszeiten	106
Abbildung 6.8: MSE hinsichtlich der Auftrittshäufigkeit (bei Mittelung über den gesamten Untersuchungszeitraum) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (a) bzw. das zweite Leadjahr (b) nach Initialisierung.	108
Abbildung 6.9: Wie Abbildung 6.8, aber für das typinterne Charakteristikum der Persistenz	109

- Abbildung 6.10:** MSE hinsichtlich der **Auftrittshäufigkeit** (bei Mittelung über den gesamten Untersuchungszeitraum) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für verschiedene Jahreszeiten des ersten Leadjahrs nach Initialisierung..... 110
- Abbildung 6.11:** Wie Abbildung 6.10, aber für das typinterne Charakteristikum der **Persistenz**. 111
- Abbildung 6.12:** MSE hinsichtlich der **Auftrittshäufigkeit** (bei Mittelung über den gesamten Untersuchungszeitraum) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für SLP (a) und ZG500 (b) des ersten Leadjahrs nach Initialisierung..... 112
- Abbildung 6.13:** Wie Abbildung 6.12, aber für das typinterne Charakteristikum der **Persistenz**. 113
- Abbildung 6.14:** MSE hinsichtlich der **Auftrittshäufigkeit** (bei Mittelung über den gesamten Untersuchungszeitraum) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren mit unterschiedlicher Klassenanzahl; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für ZG500 des ersten Leadjahres nach Initialisierung. 114
- Abbildung 6.15:** MSSS jeweils für (von oben nach unten) **b0 – hist**, **b1 – hist** und **pS – hist** für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse für die über den gesamten Zeitraum (a) bzw. saisonal aggregierten (b) **Auftrittshäufigkeiten** der auf verschiedenen Methoden (s. Überschriften) beruhenden Zirkulationstypen sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarzen Flächen sind keine Daten vorhanden. 117
- Abbildung 6.16:** MSSS jeweils für (von oben nach unten) **b1 – b0**, **pS – b0** und **pG – b0** für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse für die über den gesamten Zeitraum (a) bzw. saisonal aggregierten (b) **Auftrittshäufigkeiten** der auf verschiedenen Methoden (s. Überschriften) beruhenden Zirkulationstypen sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarzen Flächen sind keine Daten vorhanden. 118
- Abbildung 6.17:** MSSS jeweils für (von oben nach unten) **pS – b1**, **pG – b1** und **pG – pS** für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse für die über den gesamten Zeitraum (a) bzw. saisonal aggregierten (b) **Auftrittshäufigkeiten** der auf verschiedenen Methoden (s. Überschriften) beruhenden Zirkulationstypen sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarzen Flächen sind keine Daten vorhanden. 119
- Abbildung 6.18:** MSSS jeweils für den Vergleich mit **hist** (a, b), **b0** (c, d) sowie **b1/pS** (e, f). Die Ergebnisse für über den gesamten Zeitraum (linke Hälfte) bzw. die saisonal (rechte Hälfte) gemittelten **Intensitäten** der auf verschiedenen Methoden (GWT, PCA, SAN) beruhenden jeweils 8 Zirkulationstypen (basierend auf SLP) sind dargestellt für

verschiedene Jahreszeiten. Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.
..... 122

Abbildung 6.19: MSSS jeweils für den Vergleich mit **hist** (a, b), **b0** (c, d) sowie **b1/pS** (e, f). Die Ergebnisse für über den gesamten Zeitraum (linke Hälfte) bzw. die saisonal (rechte Hälfte) gemittelten **Persistenzen** der auf verschiedenen Methoden (GWT, PCA, SAN) beruhenden jeweils 8 Zirkulationstypen (basierend auf SLP) sind dargestellt für verschiedene Jahreszeiten. Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.
..... 123

Tabellenverzeichnis

<i>Tabelle 3.1: Übersicht über die verschiedenen Modellgenerationen des dekadischen Vorhersagesystems (MPI-ESM-LR) und ihre Eigenschaften.</i>	<i>11</i>
<i>Tabelle 3.2: Übersicht über die in der vorliegenden Arbeit verwendeten Reanalysedaten unter Angabe der jeweiligen Referenz.</i>	<i>16</i>
<i>Tabelle 4.1: Übersicht über die Telekonnexionsindizes (unter Angabe ihrer Berechnungsgrundlagen und der jeweiligen Referenz.), die in der vorliegenden Arbeit zur Analyse der dekadischen Vorhersagbarkeit verwendet wurden.</i>	<i>23</i>
<i>Tabelle 5.1: Varianzerklärungsanteile der CPC-Telekonnexionsmuster für die verschiedenen Jahreszeiten (MAM, JJA, SON, DJF).</i>	<i>58</i>
<i>Tabelle 5.2: Korrelationskoeffizienten verschiedener Indizes und Jahreszeiten zwischen Vorhersage (pS) und Beobachtung (ERAINT) der Jahre 1979-2011 (saisonale Aggreg.) für das erste Leadjahr. Fett markierte Werte zeigen Signifikanz auf dem 5 %-Niveau (nach Pearson).</i>	<i>79</i>

Abkürzungsverzeichnis

AAO	Antarktische Oszillation
AMO	Atlantische Multidekaden-Oszillation
AOGCM	Atmosphere-Ocean Global Circulation Model, dt.: globales Atmosphären-Ozean-Modell
b0	Baseline0 (erste Modellgeneration)
b1	Baseline1 (zweite Modellgeneration)
BMBF	Bundesministerium für Bildung und Forschung
CFSR	Climate Forecast System Reanalysis (Reanalyse)
CMIP5	Coupled Model Intercomparison Project Phase 5
CORR	Korrelation
CPC	Climate Prediction Center
CRPSS	Continuous Ranked Probability Skill Score
DLR	Deutsches Zentrum für Luft- und Raumfahrt
dt.	deutsch
EA	East Atlantic (Pattern)
EAWR	East Atlantic/Western Russia (Pattern)
ECHAM5	Atmosphärische Komponente von MPI-ESM (fünfte Generation)
ECHAM6	Atmosphärische Komponente von MPI-ESM (sechste Generation)
ECMWF	European Centre for Medium-Range Weather Forecasts, dt.: Europäisches Zentrum für mittelfristige Wettervorhersage)
engl.	englisch
ENSO	El Niño-Southern Oscillation
EPNP	East Pacific/North Pacific (Pattern)
ERA40	ERA40 (Reanalyse)
ERAINT	ERA-Interim (Reanalyse)
GECCO2	German contribution to the project ECCO (Estimating the Circulation and Climate of the Ocean)
GWT	Großwettertypen (Klassifikationsverfahren)
HAMOCC	Komponenten von MPI-ESM zur Berücksichtigung der Biochemie im Ozean
hist	Historical (uninitialisierte, historische Modellläufe)
JMA	Japan Meteorological Agency, dt.: Japanische Meteorologische Behörde

JRA25	Japanese 25-year Reanalysis (Reanalyse)
JSBACH	Komponenten von MPI-ESM zur Berücksichtigung der Landbiosphäre
MERRA	Modern-Era Retrospective for Research and Applications (Reanalyse)
MiKlip	Mittelfristige Klimaprognose (Titel des Gesamtprojekts)
MPI-ESM(-LR/-MR)	Max Planck Institute Earth System Model (low/mixed resolution), dt.: Erdsystemmodell des Max-Planck-Instituts (geringe/mittlere Auflösung)
MPIOM	Ozeanische Komponente von MPI-ESM
MSE	Mean Squared Error
MSSS	Mean Squared Error Skill Score
NAO	Nordatlantische Oszillation
NASA	National Aeronautics and Space Administration, dt.: Nationale Aeronautik- und Raumfahrtbehörde
NCEP/NCAR	National Centers for Environmental Prediction / National Center for Atmospheric Research)
NWS	National Weather Service, dt.: Nationaler Wetterdienst (der USA)
OASIS	Koppler-Programm von MPI-ESM
ORAS4	Ocean Reanalysis System 4
PC	Principal component, dt.: Hauptkomponente
PCA	Principal component analysis, dt.: Hauptkomponentenanalyse
PDO	Pazifische Multidekaden-Oszillation
pG, pS	Prototype (dritte Modellgeneration, Initialisierung mit <u>GECCO2</u> bzw. <u>ORAS4</u>)
PNA	Pacific/North American (Pattern)
POL	Polar/Eurasia (Pattern)
RCP4.5	Representative Concentration Pathways 4.5
RPC	Ratio of Predictable Components
RPCA	Principal component analysis, dt.: Rotierte Hauptkomponentenanalyse
RPSS	Ranked Probability Skill Score
SAM	Southern Annular Mode (alternative Bezeichnung für AAO)
SAN (SANDRA)	Simulated Annealing and Diversified Randomization (Clusteranalyse)
SAT	Surface air temperature, dt.: bodennahe Lufttemperatur
SCAND	Scandinavia (Pattern)
SLP	Sea level pressure, dt.: Luftdruck auf Meereshöhe

SOI	Southern Oscillation Index
SPCA	S-mode principal component analysis, dt.: s-modale Hauptkomponentenanalyse
SST	Sea surface temperature, dt.: Meeresoberflächentemperatur
TNH	Tropical/Northern Hemisphere (Pattern)
TPCA	T-mode principal component analysis, dt.: t-modale Hauptkomponentenanalyse
VADY	Validierung der Dynamik der Atmosphäre (Projekttitle)
WP	West Pacific (Pattern)
ZG500	Geopotentielle Höhe (hier: geopotentielle Höhe des 500 hPa-Niveaus)

Abstract

This thesis analyses to what extent the dynamic of the atmosphere is represented by decadal climate models. To capture this, two different methodological approaches are used. On the one hand, different teleconnection indices like the North Atlantic Oscillation (NAO) or the Southern Oscillation Index (SOI), in which the internal variability of the atmosphere is expressed, serve as an object of investigation. On the other hand, different classification procedures are used (based on correlation coefficients, cluster or principal component analyses) for the generation of circulation types. For this purpose, a spatial restriction on the North Atlantic and European area (90°W to 40°E, 20°N to 80°N) is carried out.

To examine the decadal prediction skill, the global climate model of the Max Planck Institute (MPI-ESM) is used in different model generations (*historical*, *baseline0*, *baseline1*, *prototype*). These differ concerning their initialization technique as well as the number of available ensemble members. This work examines, based on different model generations, to what extent an initialization of the model as well as an enlargement of the ensemble have an impact on the prediction skill. The methodological approach is oriented towards the hindcast analyses, which are common in the decadal field of research. This implies, that with the aid of the decadal climate model retrospective decadal predictions for the investigation period (1979-2011) are provided and compared with observed values of the different reanalyses. The evaluation is not carried out for single predictions of a certain decadal period within the whole investigation period, but by so-called “lead years”. For this purpose, one year is picked out from all annually started predictions within the investigation period after the initialization of the model (e. g., for “lead year 1” the first year of the prediction). To measure the prediction skill, different skill scores are used: Mean Squared Error Skill score (MSSS) and Ranked Probability Skill score (RPSS), a deterministic one as well as a probabilistic skill score. Moreover, the linear correlation coefficient (CORR) is used for the evaluation of the time series.

The results show, that for the teleconnection indices as well as for the different classification procedures improvements can be observed in the prediction skill. They appear basically only in the first year after the initialization of the model, i.e. in the first lead year. The improvements can especially be observed for the initialized second (*baseline1*) and

third model generation (*prototype*) of the decadal prediction model MPI-ESM, but not for the uninitialized model runs (*historical*) and the first model generation (*baseline0*). However, the different teleconnection indices and classification procedures show big differences. On the one hand, in the spring (MAM) of the first lead year, i.e. for the third, fourth and fifth month after the initialization of the model, correlation coefficients near 0 occur for single teleconnection indices (e.g. NAO). On the other hand, values greater than $r=0,7$ can be observed for certain indices (e.g. SOI).

Moreover, this thesis indicates that not only the implementing of an initialization onto the climate model positively affects the prediction skill. Rather, it can be proved that the rise of the ensemble number has a decisive influence on the improvement of the prediction skill. Nevertheless, it must be ascertained that the improvements can be shown primarily for the first lead year but cannot be observed for the whole decadal period.

1 Einleitung

1.1 Motivation

Wetter und Klima haben einen großen Einfluss auf die Lebensbedingungen des Menschen und seine ihn unmittelbar betreffende Umwelt. Deswegen liegt ein besonderes Augenmerk darauf, zukünftige Entwicklungen des Wettergeschehens sowie Änderungen des Klimas zuverlässig vorherzusagen. Lange Zeit stand die tägliche Wettervorhersage im Mittelpunkt des Interesses, bevor im Zusammenhang mit der Klimawandel-Diskussion und durch die Entwicklung leistungsfähiger Klimamodelle die Praxis langfristiger Klimaprojektionen Einzug in die Wissenschaft gehalten hat.

Seit einigen Jahren gibt es Bestrebungen, die Lücke zu schließen, die zwischen der auf mehrere Tage ausgelegten Wettervorhersage und den Klimaprojektionen klafft, welche für unterschiedliche Emissionsszenarien auf einem Zeitraum bis 2100 angesiedelt sind. Das Untersuchungsgebiet der dekadischen Klimavorhersage hat sich seitdem in der Forschungslandschaft etabliert, was sich an der Publikation einschlägiger Artikel (Smith et al. 2007, Pohlmann et al. 2009, Goddard et al. 2013, Meehl et al. 2014 etc.) und der Einrichtung von nationalen wie internationalen Forschungsprojekten ablesen lässt. Wissenschaftliche Fortschritte auf diesem Gebiet sind von besonderem Interesse, weil sie – im Gegensatz zu den langfristigen Klimaprojektionen – zum einen planerisch relevante Zeiträume in Politik, Wirtschaft und Gesellschaft unmittelbar adressieren. Zum anderen sind

viele natürliche Schwankungen innerhalb des Klimasystems auf dieser Zeitskala angesiedelt (Hurrell et al. 2010; Murphy et al. 2010), die ähnliche Magnituden aufweisen wie anthropogen bedingte Klimaänderungen in den nächsten Dekaden (Solomon et al. 2011).

Um klimainterne Schwankungen in geeigneter Weise in die Vorhersage zu integrieren, ist die Initialisierung dekadischer Vorhersagemodelle, d. h. das Starten des Modells mit bekannten Anfangsbedingungen, von zentraler Bedeutung. Diese Tatsache verdeutlicht die besondere Herausforderung, die dekadische Vorhersagesysteme bewältigen müssen. Während bei Wettervorhersagen auf der einen Seite des Vorhersagezeitraums ein Initialisierungsproblem und bei Klimaprojektionen auf der anderen Seite des Vorhersagezeitraums ein Randwertproblem auftritt, muss bei dekadischen Vorhersagen das Ziel sein, beide Ansätze zu integrieren (Meehl et al. 2009). Das bedeutet, dass zum einen langfristig ablaufende Klimaänderungen (Randwerte, also externe Klimaantriebe wie anthropogene CO₂-Emissionen oder die Solarstrahlung) Berücksichtigung finden müssen. Um die interne Klimavariabilität adäquat abbilden zu können, müssen zum anderen die Anfangsbedingungen zur Initialisierung des Modells möglichst genau bekannt sein, wofür detaillierte Kenntnisse sowohl des aktuellen Klimazustands als auch von geeigneten Initialisierungstechniken vonnöten sind.

Diese knappe Zusammenschau verdeutlicht den enormen Forschungsbedarf auf dem Gebiet der dekadischen Vorhersagen, welches erst am Beginn der Entwicklung steht, in der Forschungslandschaft aber zunehmend an Aufmerksamkeit gewinnt. Die vorliegende Arbeit soll einen Beitrag auf dem Weg zur Entwicklung dekadischer Vorhersagesysteme liefern.

1.2 Fragestellungen und Zielsetzung

Einen wichtigen Baustein bei der Entwicklung von Modellen stellt die Modellvalidierung dar, deren Ziel darin liegt, Modellierung und Beobachtung zu vergleichen, um Aussagen über die Modellgüte treffen und Empfehlungen an die Modellbauer für die Weiterentwicklung des Modellsystems aussprechen zu können. Die vorliegende Arbeit liefert einen Beitrag zu diesem Vorhaben im Kontext der dekadischen Klimamodellierung und konzentriert sich auf einen Teilbereich des Klimasystems: die Atmosphäre. Im Folgenden

soll untersucht werden, inwiefern es basierend auf dem aktuellen Forschungsstand gelingt, die Dynamik der Atmosphäre auf einer dekadischen Zeitskala vorherzusagen. Die Dynamik der Atmosphäre wird in diesem Sinne anhand von zwei unterschiedlichen Ansätzen erfasst: Zum einen werden bekannte Telekonnectionsindizes wie die Nordatlantische Oszillation (NAO) oder der Southern Oscillation Index (SOI) untersucht, zum anderen werden Analysen zu unterschiedlichen Zirkulationstypen im nordatlantisch-europäischen Raum durchgeführt, die mithilfe unterschiedlicher Klassifikationsverfahren ermittelt werden.

Wie bereits erwähnt wurde, befindet sich die Forschung auf dem Gebiet der dekadischen Vorhersage noch in einem initialen Stadium. Bei den fortschreitenden Untersuchungen kommt v. a. den Fragen nach der geeigneten Initialisierung dekadischer Vorhersagemodelle eine entscheidende Rolle zu (Marotzke et al. 2016). Aber auch grundlegende Untersuchungsgegenstände wie die räumliche Auflösung des dekadischen Vorhersagemodells, die Ensemblegröße oder die Möglichkeiten zur Evaluierung des Vorhersagemodells spielen eine wichtige Rolle. Folgende Fragen spannen das Forschungsfeld der vorliegenden Arbeit auf und sollen geklärt werden:

- Sind dekadische Klimamodelle in der Lage, die Dynamik der Atmosphäre adäquat wiederzugeben?
- Wirken sich unterschiedliche Initialisierungstechniken auf die Güte der dekadischen Vorhersage aus?
- Gibt es einen Einfluss der Ensemblegröße auf die Güte der Vorhersage?
- Welche Metriken zur Messung der Güte der Repräsentation im dekadischen Vorhersagemodell sind sinnvoll?
- Sind die beiden methodischen Herangehensweisen (Telekonnectionsindizes, Zirkulationstypen) zur Evaluation der atmosphärischen Dynamik geeignet?
- Ist eine der beiden gewählten Methoden für die Validierung dekadischer Vorhersagemodelle zu bevorzugen?
- Wie wirkt sich die zeitliche Aggregation der Datengrundlage auf die Vorhersageleistung aus?

1.3 Aufbau der Arbeit

Die Arbeit führt im zweiten Kapitel in der gegebenen Kürze in das junge, aber dennoch mittlerweile breit aufgestellte Forschungsfeld der dekadischen Vorhersage ein. Es wird ein Überblick über die Thematik sowie den aktuellen Stand der Forschung gegeben. Daran anknüpfend wird im dritten Kapitel die Datengrundlage, welche die Basis für die Forschungsergebnisse darstellt, thematisiert, wobei sowohl die Modelldaten des dekadischen Vorhersagesystems als auch die für Vergleichszwecke eingesetzten Beobachtungsdaten näher erläutert werden. Im vierten Kapitel werden schließlich die Methoden vorgestellt, die in der vorliegenden Arbeit verwendet wurden. In diesem Kapitel werden zum einen die methodischen Herangehensweisen erläutert, mithilfe derer die dekadischen Vorhersagen untersucht wurden: Telekonnexionsindizes sowie Zirkulationstypen. Zum anderen werden diejenigen Methoden betrachtet, die für die Evaluation der dekadischen Vorhersagen verwendet wurden. Daran anschließend werden die Untersuchungsergebnisse für die beiden methodischen Herangehensweisen präsentiert. Zunächst wird die Vorhersageleistung des dekadischen Modells hinsichtlich ausgewählter Telekonnexionsindizes vorgestellt (Kapitel 5), um schließlich auf die Repräsentation von mithilfe verschiedener Klassifikationsverfahren generierten Zirkulationstypen im nordatlantisch-europäischen Raum einzugehen (Kapitel 6). Die beiden Kapitel sind derart aufgebaut, dass zunächst die Vorhersageleistung der uninitialisierten Modellläufe, also die Läufe, die noch keine Initialisierung aufweisen, vorgestellt werden, bevor anschließend die Einflüsse der Initialisierung sowie der Ensemblegröße analysiert werden. Zum Abschluss erfolgt jeweils eine überblickende Betrachtung der Vorhersageleistung des dekadischen Modells. Im daran anknüpfenden siebten Kapitel werden die Ergebnisse zusammengefasst und eingeordnet, bevor abschließend ein Ausblick auf mögliche zukünftige Forschungsarbeiten geworfen wird.

2 Forschungsstand

2.1 Forschungsgrundlagen

Das Forschungsgebiet der dekadischen Vorhersagen zielt auf einen Zeithorizont ab, der für den Menschen überschaubar ist und ihn unmittelbar betrifft. Um auf dekadischer Zeitskala eine gute Vorhersageleistung zu erzielen, kommen grundsätzlich zwei Quellen in Betracht: externe Klimaantriebe (z. B. CO₂-Emissionen, Änderungen der solaren Aktivität, vulkanische Aktivität, Ausstoß anthropogen bedingter Aerosole) sowie die interne Variabilität des Klimasystems, die sich aus der Interaktion der einzelnen Komponenten, wie z. B. Atmosphäre, Ozean, Landoberfläche oder Meereis, ergibt (Hurrell et al. 2010, Murphy et al. 2010, Smith et al. 2012). Wenngleich schon frühere Forschungsaktivitäten auf dem Gebiet der dekadischen bis multidekadischen Vorhersage beobachtet werden können und die aktuellen Ansätze auf diesen Aktivitäten aufbauen (Müller et al. 2012), beruhen die Fortschritte und die starke Zunahme an Forschungsaktivitäten in den letzten Jahren auf den Erfolgen, die durch die Initialisierung der Modelle (z. B. Smith et al. 2007, Keenlyside et al. 2008, Pohlmann et al. 2009) erzielt wurden. Dadurch, dass die Initialisierung der Klimamodelle Einzug in das dekadische Forschungsgebiet gehalten hat, werden nun auch interne Mechanismen des Klimageschehens besser berücksichtigt. Dies ist umso interessanter, da innerhalb des Klimasystems viele Prozesse auf der dekadischen bis multidekadischen Zeitskala angesiedelt sind, die eine vielversprechende Quelle der Vorhersagbarkeit liefern können.

Ein besonderes Augenmerk der dekadischen Klimaforschung liegt dabei auf zyklisch auftretenden Zirkulationsschwankungen der Ozeanströmung wie der Atlantischen Multidekaden-Oszillation (AMO) oder der Pazifischen Dekaden-Oszillation (PDO), die für mehrere Jahre potentiell vorhersagbar sind. Während die AMO ein großräumiges Muster der Variabilität der Meeresoberflächentemperaturen (SST) im Nordatlantik bezeichnet und auf einer multidekadischen Zeitskala mit einer Periode von 30-80 Jahren angesiedelt ist (Smith et al. 2012), stellt die PDO das dominierende Muster der SST-Variabilität im außertropischen Nordpazifik dar und weist einen Zyklus von einer oder mehreren Dekaden auf (Murphy et al. 2010, Hurrell et al. 2010). Durch ein besseres Verständnis der Zirkulationsschwankungen der Ozeanströmung erhofft man sich eine ähnliche Entwicklung, wie sie auf dem Forschungsgebiet der saisonalen Vorhersage beobachtet werden kann: Eine bessere Vorhersage der Ozeanströmung ermöglicht eine bessere Vorhersage der SST, wodurch eine bessere Vorhersage atmosphärischer Prozesse und Parameter wie dem Bodenluftdruck (SLP) und der bodennahen Lufttemperatur (SAT), die mit der Variabilität der SST zusammenhängen, erwartet wird (Murphy et al. 2010, Müller et al. 2012).

Eine wichtige Schlüsselregion für dekadische Vorhersagen stellt in diesem Kontext der Nordatlantik dar. Zwar wird ein Zusammenhang zwischen AMO und der Nordatlantischen Oszillation (NAO) noch kontrovers diskutiert (Müller et al. 2012) bzw. ist nicht eindeutig (Murphy et al. 2010), dennoch deuten aktuelle Forschungsergebnisse an, dass ein möglicher Zusammenhang zwischen der nordatlantischen SST und der NAO im Winter existiert (Müller et al. 2012). Außerdem konnte ein signifikanter Einfluss für beobachtete Werte von SLP und SAT im Sommer gefunden werden (Hodson et al. 2010). Hier besteht allerdings noch ein gesteigerter Forschungsbedarf. Aus dem Forschungsgebiet der saisonalen Vorhersage ist bekannt, dass die größte Quelle potentieller Vorhersagbarkeit das Phänomen El Niño-Southern Oscillation (ENSO) darstellt, eine weitere Schlüsselregion der dekadischen Vorhersage (u. a. Hurrell et al. 2010, Smith et al. 2012). ENSO ist ein komplex gekoppeltes Zirkulationssystem im tropischen Pazifik, welches sich durch positive Rückkopplungen zwischen SST und Passatwinden aufschaukelt: Eine Abschwächung der Passatzirkulation verursacht positive SST-Anomalien im östlichen Pazifik, was wiederum die zonale Walkerzirkulation abändert und zu einer weiteren Abschwächung der östlichen Passatwinde führt. Für das Forschungsgebiet der dekadischen Vorhersage ist das ENSO-Klimaphänomen auch deswegen von Interesse, weil es nicht wie andere

bereits beschriebene Zirkulationsschwankungen auf einer dekadischen bis multidekadischen Zeitskala angesiedelt ist, sondern mit einer Periode von zwei bis sieben Jahren hochfrequenter auftritt. Während für multidekadische Phänomene wie der AMO oder der PDO aufgrund der geringen Anzahl an Beobachtungsjahren detaillierte Aussagen zur mittleren Periode nur schwer zu treffen sind, bietet ENSO mit seinem subdekadischen Auftreten bessere Möglichkeiten für die dekadische Klimaforschung. Ziel muss es weiterhin sein, zusätzliche Quellen mit einer potentiellen dekadischen Vorhersagbarkeit zu identifizieren und zu verstehen.

2.2 Aktuelle Forschungsaktivitäten

Aktuelle Forschungsaktivitäten auf dem Gebiet der dekadischen Vorhersage sind häufig in großen, international ausgerichteten Forschungsprojekten angesiedelt. Einen wichtigen Anstoß lieferte das Klimamodell-Vergleichsprojekt CMIP5 (Coupled Model Intercomparison Project Phase 5; Taylor et al. 2012), bei dem mit vergleichbaren Modellläufen aktuelle Fragen zu Mechanismen und Charakteristika des Klimawandels und der Klimavariabilität erforscht werden. CMIP5 integriert idealisierte Experimente von 16 verschiedenen Modellierungsgruppen (für eine tabellarische Übersicht s. Meehl et al. 2014), wobei neben den bisher üblichen gekoppelten Atmosphären-Ozean-Klimamodellen (AOGCM; Atmosphere-Ocean Global Circulation Model) erstmals Erdsystemmodelle zum Einsatz kommen, welche zusätzliche Komponenten zur Simulation der Vegetation oder Landbedeckung enthalten. Neben den etablierten Langzeit-Experimenten für die ferne Zukunft (bis 2100 bzw. teilweise bis 2300) bietet CMIP5 einen neuen Rahmen initialisierter Vorhersagen mit zwei Kernexperimente für die nahe Zukunft (Meehl et al. 2009, Murphy et al. 2010). Das erste führte eine Serie von dekadischen Vorhersagen durch, die mit beobachteten Werten des aktuellen Klimazustanden initialisiert werden. Hierbei ist vor allem die Frage nach der besten Initialisierung von Interesse. Das zweite Kernexperiment erweitert den Vorhersagezeitraum auf 30 Jahre und untersucht die Vorhersagbarkeit auf einer Zeitskala, die stärker von externen Mechanismen (v. a. dem steigenden CO₂-Gehalt) beeinflusst wird. Dennoch erhofft man sich, auch auf dem längeren Vorhersagezeitraum einen Einfluss der Initialisierung entdecken zu können (Taylor et al. 2012). Zu diesen zwei idealisierten Ansätzen leisten die verschiedenen Modellierungsgruppen mit ihren

Modellen und unterschiedlichen Initialisierungstechniken einen Beitrag und ermöglichen einen Vergleich der unterschiedlichen Klimamodelle.

Eine andere Herangehensweise wählt das Projekt MiKlip („Mittelfristige Klimaprognosen“), in dessen Rahmen diese Arbeit entstand. Das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Projekt, das auf dem ersten Kernexperiment von CMIP5 aufbaut, vergleicht nicht mehrere Klimamodelle miteinander, sondern bedient sich lediglich eines Klimamodells, des gekoppelten Globalen Klimamodells des Max-Planck-Instituts (MPI-ESM). Außerdem konzentriert sich das Projekt auf dekadische Vorhersagen und lässt Vorhersagen für die weitere Zukunft außer Acht. Ein Vorteil des Ansatzes liegt darin, dass nicht – wie bei CMIP5 – in einem aufwändigen Prozedere mehrere Klimamodelle miteinander verglichen werden, sondern der Fokus auf die Weiterentwicklung eines Klimamodells gelegt werden kann. Zwar bietet CMIP5 mit dem Vergleich von 16 Klimamodellen ein deutlich breiteres Spektrum an Modellierungsansätzen, allerdings kann MiKlip mit der Fokussierung auf ein Klimamodell schnell zugeschnittene Antworten auf Forschungsfragen produzieren, die dabei helfen können, das Vorhersagesystem zu verbessern (Marotzke et al. 2016). So wurden während der ersten Projektphase von MiKlip drei Generationen (*baseline0*, *baseline1*, *prototype*) des dekadischen Vorhersagesystems erstellt, um Antworten auf drängende Fragen der Forschung, z. B. der Wahl der besten Initialisierungstechnik, der horizontalen und vertikalen Auflösung des Klimamodells oder der Anzahl an Ensemblemitgliedern, auf dem Gebiet der dekadischen Vorhersage geben zu können. Zudem legt MiKlip einen besonderen Fokus auf die Operationalisierung der Evaluationssoftware, zu dem viele der beteiligten Projekte beigesteuert haben. In der vorliegenden Arbeit tritt diese Tatsache dadurch zutage, dass die Analysen auch im Hinblick auf einen effizienten Rechenaufwand ausgerichtet sind.

Alle Publikationen, die in den letzten Jahren aus den verschiedenen Forschungsgruppen heraus entstanden sind, dienen dem Zweck, zu untersuchen, ob dekadische Klimavorhersagen prinzipiell möglich und sinnvoll sind oder nicht. Auf dem Forschungsgebiet der dekadischen Vorhersage lassen sich mehrere Felder identifizieren, auf denen Forschungsaktivitäten stattfinden. Ein Hauptaugenmerk liegt allerdings auf der Anwendung verschiedener Techniken zur Initialisierung der dekadischen Klimavorhersagen. Während Smith et al. (2007) demonstrieren, dass die Initialisierung zu besseren Vorhersagen der globalen Durchschnittstemperatur über die letzten drei Dekaden globaler Erwärmung

führt, entdecken Keenlyside et al. (2008) und Pohlmann et al. (2009) im nordatlantischen Raum eine potentielle Quelle der dekadischen Klimavorhersage. Basierend auf initialisierten Klimamodellen zielen weitere Forschungsaktivitäten auf die Identifizierung potentieller Quellen dekadischer Vorhersagbarkeit sowohl für einzelne klimatische Parameter (z. B. untersuchen Kruschke et al. (2015) die Vorhersagbarkeit nordhemisphärischer Winterstürme) als auch in räumlicher Hinsicht ab (z. B. entdecken Pohlmann et al. (2013) eine verbesserte Vorhersageleistung in den Tropen). Ein weiteres Forschungsfeld besteht in der Anwendung geeigneter Metriken zur Messung der Vorhersageleistung (sog. Skillmetriken). Basierend auf der Publikation von Goddard et al. (2013), die sich grundlegend mit der Evaluation dekadischer Vorhersagen beschäftigt, bewerten Kadow et al. (2015) das dekadische Vorhersagesystem des MiKlip-Projekts anhand der (über mehrere Ensemblemitglieder) gemittelten Vorhersage sowie der Bandbreite aller Ensemblemitglieder als Ausdruck der Unsicherheit des Modells. Weitere Forschungsfelder, die nach Murphy et al. (2010) Berücksichtigung finden sollten, bieten die Verbesserung der Klimamodelle (z. B. der horizontalen und vertikalen Auflösung sowie der Physik) und die Verbesserung der Beobachtungsdatensätze, mit denen die Initialisierung der dekadischen Vorhersagesysteme durchgeführt wird, da eine fehlerbehaftete Beobachtung der Atmosphäre, des Ozeans, der Kryosphäre oder der Landoberfläche eine mögliche Fehlerquelle darstellt. In diesem Zusammenhang bereiten z. B. Pattantyús-Ábrahám et al. (2015) Daten von homogenisierten Radiosonden auf, um Datensätze für die Evaluierung dekadischer Klimamodelle zu generieren. Auch die Frage nach der Anzahl der Ensemblemitglieder stellt ein wichtiges Forschungsfeld dar, dem sich Sienz et al. (2016) widmen, indem sie den Einfluss der Ensemblegröße auf die dekadische Vorhersageleistung untersuchen.

Die überblickende Darstellung der aktuellen Forschungsfelder offenbart die vielfältigen Studien, die derzeit im Bereich der dekadischen Vorhersage durchgeführt werden. Ein gemeinsames Resultat all dieser Studien ist, dass sie eine gewisse Vorhersageleistung zeigen, diese jedoch bezüglich der verschiedenen Klimaparameter und Regionen variiert. In dieses Umfeld ist die vorliegende Arbeit mit den Untersuchungen zur Vorhersagbarkeit der Dynamik der Atmosphäre einzuordnen.

3 Datengrundlage

Um Aussagen über die Leistung (den Skill) einer Klimavorhersage treffen zu können, werden die Werte, die von dem Vorhersagesystem prognostiziert werden (Modelldaten), mit der Realität, d. h. mit beobachteten Werten (Beobachtungsdaten), verglichen. In diesem Kapitel werden die Daten vorgestellt, die den vorliegenden Analysen zugrunde liegen: die Modelldaten (Kapitel 3.1) sowie die Beobachtungsdaten (Kapitel 3.2).

3.1 Modelldaten

Für die Modelldaten wurde auf das gekoppelte Klimamodell MPI-ESM (Max-Planck-Institute Earth System Model; Giorgetta et al. 2013) zurückgegriffen. Zentrale Bestandteile des Klimamodells stellen die atmosphärische Komponente ECHAM6 (Stevens et al. 2013) sowie die ozeanische Komponente MPIOM (Jungclaus et al. 2013) dar (s. Ab-

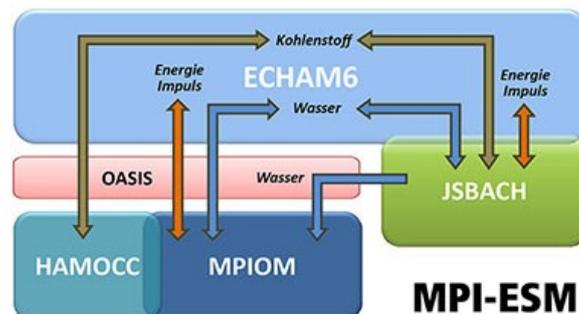


Abbildung 3.1: Das Klimamodell MPI-ESM und seine Komponenten (Quelle: Giorgetta et al. 2013).

bildung 3.1). Zusätzlich ist die Komponente HAMOCC zur Berücksichtigung der Biochemie im Ozean sowie die Subsysteme JSBACH für die Landbiosphäre und das separate Koppler-Programm OASIS integriert, welches den Austausch von Wasser, Energie, Impuls und wichtigen Spurengasen wie CO₂ zwischen den einzelnen Komponenten organisiert. Der größte konzeptionelle Unterschied zu der Vorgängerversion (ECHAM5/MPIOM) besteht darin, dass durch die Integration von HAMOCC und JSBACH der Kohlenstoffkreislauf zum Modellsystem hinzugefügt wurde. Weitere Verbesserungen zwischen ECHAM5 und ECHAM6 wurden u. a. bei der Berechnung der Oberflächenalbedo (z. B. Berücksichtigung von Schneedecken oder Wasserflächen auf Meereis), dem kurzwelligen Strahlungstransport sowie den Konvektionszuständen erzielt. Zudem stehen dem Anwender unterschiedliche Modellauflösungen zur Verfügung: MPI-ESM-LR (geringe Auflösung; low resolution) mit 1,9° horizontaler Auflösung/47 Level in der Atmosphäre (T63/L47) bzw. 1,5° horizontaler Auflösung/40 Level im Ozean und MPI-ESM-MR (mittlere Auflösung; mixed resolution) mit 1,9° horizontaler Auflösung/95 Level in der Atmosphäre (T63/L95) bzw. 0,5° horizontaler Auflösung/40 Level im Ozean (Giorgetta et al. 2013).

Table 3.1: Übersicht über die verschiedenen Modellgenerationen des dekadischen Vorhersagesystems (MPI-ESM-LR) und ihre Eigenschaften.

Experiment	Auflösung (Atmosphäre)	Initialisierung (Atmosphäre)	Initialisierung (Ozean)	Ensemblemitglieder
baseline0 (b0)	T63L47 (LR)	-	anomaly (NCEP/NCAR)	3
baseline1 (b1)	T63L47 (LR) / T63L95 (MR)	full-field (ERA40/ERAINT)	anomaly (ORAS4)	10 / 5
prototype (pG)	T63L47 (LR)	full-field (ERA40/ERAINT)	full-field (GECCO2)	15
prototype (pS)	T63L47 (LR)	full-field (ERA40/ERAINT)	full-field (ORAS4)	15
historical (≤ 2005), rcp45 (> 2006) (hist)	T63L47 (LR)	-	-	3 / 10

Innerhalb des MiKlip-Projekts wurde das Klimamodell MPI-ESM (in beiden Auflösungen) dazu verwendet, um initialisierte dekadische Modellläufe zu generieren. Durch die Initialisierung des Modells mit dem beobachteten Klimazustand wird erwartet, dass die Simulation anschließend in einem gewissen Grad der beobachteten Klimaentwicklung folgt (Taylor et al. 2012). Zudem existiert die Hoffnung, dass die Initialisierung falsche Rückmeldungen vorheriger externer Antriebe korrigiert (Kruschke et al. 2015). Zusätzlich wurde auf die Historischen Läufe von MPI-ESM zurückgegriffen, die im Kontext dekadischer Vorhersage auch als „nicht initialisierte“ bzw. „uninitialisierte“ Läufe bezeichnet werden (im Folgenden wird auf das aus dem Englischen übersetzte Wort „uninitialized“ (engl. *uninitialized*) zurückgegriffen), da sie typischerweise mehrere Dekaden vor dem Validierungszeitraum aus einem vorindustriellen Zustand heraus gestartet werden und mit dem aktuellen Klimazustand nicht in Kontakt kommen.

Im Folgenden werden die verschiedenen Modellgenerationen der initialisierten Läufe sowie die Historischen Läufe mit ihren jeweiligen Eigenschaften, wie sie aus Tabelle 3.1 hervorgehen, erläutert.

3.1.1 Dekadisches Vorhersagesystem (Initialisiert)

Von dem Modellsystem MPI-ESM stehen drei Modellgenerationen mit dekadischen Vorhersagen zur Verfügung: *baseline0* (**b0**), die erste Modellgeneration, *baseline1* (**b1**), die zweite Modellgeneration, sowie *prototype* (**pG/pS**), die dritte Modellgeneration (s. Tabelle 3.1). Alle Modellgenerationen weisen dieselbe Modellkonfiguration auf, unterscheiden sich aber in der Anzahl der Ensemblemitglieder und in der Form der Initialisierung.

3.1.1.1 Baseline0 (Erste Modellgeneration)

Da die beiden Projekte CMIP5 und MiKlip zeitnahe Starttermine aufwiesen, konnte für die erste Modellgeneration (*baseline0*) auf die für das CMIP5 generierten dekadischen Vorhersagen von MPI-ESM zurückgegriffen werden. Wie aus Tabelle 3.1 hervorgeht, weist die Atmosphäre keine Initialisierung auf, sondern lediglich der Ozean. Hierfür wurde die sog. „anomaly“-Initialisierungstechnik verwendet, d. h. eine Initialisierung erfolgte mithilfe von Anomaliefeldern der Klimaparameter. Diese wurden durch ein sog. Assimilations-Experiment gewonnen, das für die Jahre 1948-2012 durchgeführt wurde. Die ozeanische Komponente von MPI-ESM, MPIOM, wurde hierfür täglich mit Werten

des Impuls-, des Wärme- und des Süßwasserflusses angetrieben (engl. *forced*), wobei dafür auf die Reanalyse des NCEP/NCAR (National Centers for Environmental Prediction / National Center for Atmospheric Research) zurückgegriffen wurde. Die Anomalien der Ozeantemperatur und des Salzgehalts wurden der Klimatologie des gekoppelten Modells hinzugefügt (= „anomaly“-Initialisierungstechnik). Die als Ergebnis des Assimilations-Experiments resultierenden, dreidimensionalen Felder des Ozeans und der Temperatur wurden als initiale Bedingungen für die dekadischen Vorhersagen verwendet (Müller et al. 2012, Pohlmann et al. 2013, Marotzke et al. 2016).

Im Vergleich zum Klimamodell-Vergleichsprojekt CMIP5 wurde die Anzahl der Startjahre deutlich erhöht, indem die dekadischen Läufe nicht mehr nur alle fünf Jahre, sondern für jedes Jahr des Zeitraums 1961-2010 (später verlängert bis 2012) initialisiert wurden. Dadurch stehen mehr dekadische Vorhersagen für Analysezwecke zur Verfügung. Wie aus Tabelle 3.1 hervorgeht, wurden die dekadischen Vorhersagen für jeweils drei Ensemblemitglieder durchgeführt. Die Ensemblemitglieder wurden durch die Initialisierung des Assimilations-Experiments mit einem Tag Zeitversatz folgend auf den 1. Januar des jeweiligen Jahres gewonnen.

3.1.1.2 Baseline1 (Zweite Modellgeneration)

Die dekadischen Vorhersagen der zweiten Modellgeneration (**b1**) wurden mit demselben Modell (MPI-ESM) und derselben Modellkonfiguration generiert. Der Unterschied zu **b0** liegt in der Initialisierung und der Anzahl an Ensemblemitgliedern (s. Tabelle 3.1).

Die ozeanische Komponente MPIOM wurde nach dem Prinzip, wie es für **b0** erläutert wurde, initialisiert. D. h. es wurden die aus einem Assimilationslauf gewonnen dreidimensionalen Felder der Ozeantemperatur und des Salzgehalts der Modellklimatologie hinzugefügt, welche wiederum aus einem historischen Lauf der Jahre 1958-2005 gewonnen wurde. Allerdings wurde nicht wie bei **b0** auf den NCEP/NCAR-Reanalysedatensatz zurückgegriffen, sondern die ORAS4-Ozeanreanalyse (Ocean Reanalysis System 4; Balmaseda et al. 2013) des ECMWF (Europäisches Zentrum für mittelfristige Wettervorhersage) verwendet. Wie aus Tabelle 3.1 hervorgeht, wurde für die dekadischen Vorhersagen von **b1** nicht nur der Ozean, sondern zusätzlich auch die Atmosphäre (ECHAM6) initialisiert. Hierfür wurden dreidimensionale Felder der Temperatur, der Vorticity, der Divergenz und des Bodenluftdrucks von Reanalysedatensätzen des ECMWF (ERA40 von

1960-1989 und ERAINT von 1990-2013; s. Uppala et al. 2005, Dee et al. 2011) verwendet (Pohlmann et al. 2013, Kadow et al. 2015, Marotzke et al. 2016). Allerdings erfolgte die Initialisierung nicht mittels der „anomaly“-Technik, sondern die Felder wurden ohne Anomalienbildung verwendet („full-field“-Initialisierung). Während demnach bei einer „anomaly“-Initialisierung das Vorhersagemodell mit den zum Klimamittel hinzugefügten Anomalien initialisiert wird, bringt die „full-field“-Initialisierung den Modellzustand nahe an die Beobachtung heran. Allerdings driftet das dekadische Vorhersagemodell bei der „full-field“-Initialisierung zu seinem systematischen Fehlerzustand, der eigenen, bevorzugten Klimatologie, während der Vorhersage ab, was eine Biaskorrektur erforderlich macht (Meehl et al. 2014).

Im Vergleich zu **b0** wurde für **b1** der Zeitraum leicht verlängert und jährlich gestartete dekadische Vorhersagen für die Jahre 1961-2013 erzeugt. Um die Unsicherheiten des Modells zu ermitteln, wurden wiederum mehrere Ensembleläufe generiert, die – ebenso wie bei **b0** – mit einem Zeitversatz von jeweils einem Tag nach dem 1. Januar gestartet wurden. Die Anzahl wurde allerdings von drei auf zehn vergrößert. Neben der Modellvariante mit geringer Auflösung (MPI-ESM-LR) wurde zudem eine Modellversion mit einer höheren Auflösung (MPI-ESM-MR) und fünf Ensemblemitgliedern erstellt.

3.1.1.3 Prototype (Dritte Modellgeneration)

Für die dritte Modellgeneration (prototype) wurden mehrere Verbesserungsvorschläge übernommen, die von Seiten der Community an die Modellbauer herangetragen wurden (s. Tabelle 3.1). Zum einen wurde konsequent die „full-field“-Initialisierungstechnik angewendet, d. h. auch der Ozean wurde auf diese Weise initialisiert, was eine generelle Tendenz auf dem Gebiet der dekadischen Vorhersage widerspiegelt (Meehl et al. 2014, Marotzke et al. 2016). Zum anderen wurde die Anzahl der Ensemblemitglieder deutlich von zehn auf 30 erhöht, wobei 15 mit ORAS4, dem Ozean-Reanalysedatensatz des ECMWF, und 15 mit GECCO2 (Deutscher Beitrag des Projekts ECCO (Estimating the Circulation and Climate of the Ocean) in der zweiten Version; vgl. Köhl 2015) initialisiert wurden. Diese wurden wieder durch Initialisierung mit einem Zeitversatz von einem Tag folgend auf den 1. Januar für den Zeitraum 1961-2013 (jährlich gestartet) erzeugt. Damit entspricht der Ansatz zur Erhöhung der Anzahl an Ensemblemitgliedern der Annahme, dass die Korrelation des Ensemblemittels zur Referenz steigt und das Risiko sinkt, zweifelhafte Vorhersageleistung zu ermitteln. Da ein Hauptaugenmerk auf der Erhöhung der

Ensemblemitgliederanzahl lag, wurden aus rechentechnischen Gründen dekadische Vorhersagen lediglich in der Modellversion mit einer geringen Auflösung (MPI-ESM-LR) erzeugt.

3.1.2 Historische Läufe (Uninitialisiert)

Die uninitialisierten Läufe, die zunächst lediglich drei Ensemblemitglieder umfassten, wurden erstellt, um im Vergleich mit den initialisierten den Effekt der Initialisierung zu ermitteln (s. Tabelle 3.1). Nachträglich wurde die Anzahl auf zehn Ensemblemitglieder erhöht, um eine äquivalente Anzahl zu den initialisierten Läufen (insbesondere **b1**) zu erhalten. Grundsätzlich besitzen die Historischen Läufe dieselbe Modellkonfiguration wie die initialisierten Läufe, es wurde lediglich auf eine Initialisierung verzichtet. Gestartet werden sie ausgehend von einer vorindustriellen Kontrollsimulation und berücksichtigen externe Strahlungsantriebe (Solarstrahlung, Aerosol- und Treibhausgaskonzentrationen, Einfluss von Vulkanen, Landnutzungsänderungen) für den Zeitraum 1850-2005. Folglich werden bei den uninitialisierten Modellläufen lediglich externe Klimaantriebe beachtet, die Initialisierung mit beobachteten Klimazuständen bleibt aus.

Da die Historischen Läufe 2005 enden, wurden sie mit dem CMIP5-Experiment erweitert, welches sich aus dem Emissionsszenario RCP4.5 (Representative Concentration Pathways 4.5; d. h. der Strahlungsantrieb beträgt zusätzlich 4,5 W/m² im Jahr 2100, wobei die Treibhausgasemissionen bis 2040 ein Maximum erreichen und danach abnehmen) zusammensetzt (Müller et al. 2012). Auch wenn die Erweiterung einen Bruch in der Kontinuität darstellt, wurde sie durchgeführt, um eine möglichst lange Vergleichszeitreihe mit uninitialisierten Modelldaten bereitstellen zu können.

3.2 Beobachtungsdaten

Für die Evaluation der Vorhersageleistung der dekadischen Vorhersagen ist es notwendig, die vorhergesagten Werte mit der Realität, d. h. Beobachtungsdaten, zu vergleichen. Zu diesem Zweck kamen in der vorliegenden Arbeit Reanalysedaten zum Einsatz. Reanalysen sind keine direkten Beobachtungsdaten, sondern mithilfe von Modellen berechnete Daten, in die qualitätsgeprüfte Beobachtungsdaten assimiliert wurden. Der Vorteil von Reanalysen (z. B. im Gegensatz zu Einzelbeobachtungen) besteht darin, dass sie für

verschiedene Parameter eine räumlich vollständige und kohärente Aufzeichnung der atmosphärischen Zirkulation enthalten, d. h. für verschiedene Parameter in unterschiedlichen Höhengniveaus globale, zeitlich und räumlich hoch aufgelöste homogene Felder über einen langen Zeitraum liefern (Dee et al. 2011). Da für die Erstellung von Reanalysedaten unterschiedliche Beobachtungswerte sowie Modelle zum Einsatz kommen, ist es durchaus möglich, dass sich diese stärker voneinander unterscheiden. Aus diesem Grund sollen im Folgenden die einzelnen Reanalysedatensätze, die in der vorliegenden Arbeit zum Einsatz kamen, kurz vorgestellt und hinsichtlich ihrer Konsistenz verglichen werden.

3.2.1 Datensätze

Tabelle 3.2 gibt einen Überblick über die in der vorliegenden Arbeit verwendeten Reanalysedatensätze. Die Reanalysen wurden für die Untersuchungen zur dekadischen Vorhersagbarkeit für mehrere Variablen (SLP, ZG500), Höhenlevels und in verschiedenen zeitlichen Auflösungen (6-stündlich, monatlich) verwendet – *monatlich* aufgelöst für Untersuchungen der Telekonnectionsindizes (s. Kapitel 5), *6-stündlich* für Analysen der Zirkulationstypen (s. Kapitel 6).

Tabelle 3.2: Übersicht über die in der vorliegenden Arbeit verwendeten Reanalysedaten unter Angabe der jeweiligen Referenz.

Reanalyse (Abk.)	Reanalyse	Organisation	Referenz
CFSR	Climate Forecast System Reanalysis	NCEP	Saha et al. (2010)
ERAINT	ERA-Interim	ECMWF	Dee et al. (2011)
JRA-25	Japanese 25-year ReAnalysis	JMA	Onogi et al. (2007)
MERRA	Modern Era Retrospective-analysis for Research and Applications	NASA	Rienecker et al. (2011)

Die Reanalysen besitzen unterschiedliche räumliche Auflösungen (z. B. MERRA die höchste mit $0,5^\circ \times 0,5^\circ$). Sie werden deshalb nicht in der Tabelle aufgeführt, weil die räumliche Auflösung der Beobachtungsdaten mittels eines einfachen Interpolationsverfahrens (*First-Order Conservative Remapping*, s. Jones 1999) an die Auflösung der Modelldaten angepasst wurde.

3.2.2 Konsistenz der Reanalysen

Vergleicht man modellierte Vorhersagewerte mit Beobachtungsdaten, ist es von Vorteil, wenn die eingesetzten Beobachtungsdaten keine große Diskrepanz aufweisen. Andernfalls würden Unterschiede in der Vorhersageleistung von Klimamodellen aufgrund von Unterschieden in den Beobachtungsdaten resultieren, was nicht im Interesse einer Modellevaluation stehen kann.

In Abbildung 3.2 sind für den Untersuchungsgegenstand der Telekonnectionsindizes, der in Kapitel 4.1.1 ausführlich erläutert wird, die Zeitreihen (Zeitraum: 1980-2011) der Nordatlantischen Oszillation (NAO) im Winter (DJF) für verschiedene Reanalysen dargestellt. Die Grafik offenbart eine starke Übereinstimmung zwischen den Reanalysen, was an der starken Überlagerung der Zeitreihen ersichtlich ist. Die Korrelationskoeffizienten zwischen den einzelnen Reanalysen untereinander weisen dementsprechend einen Wert nahe 1 auf.

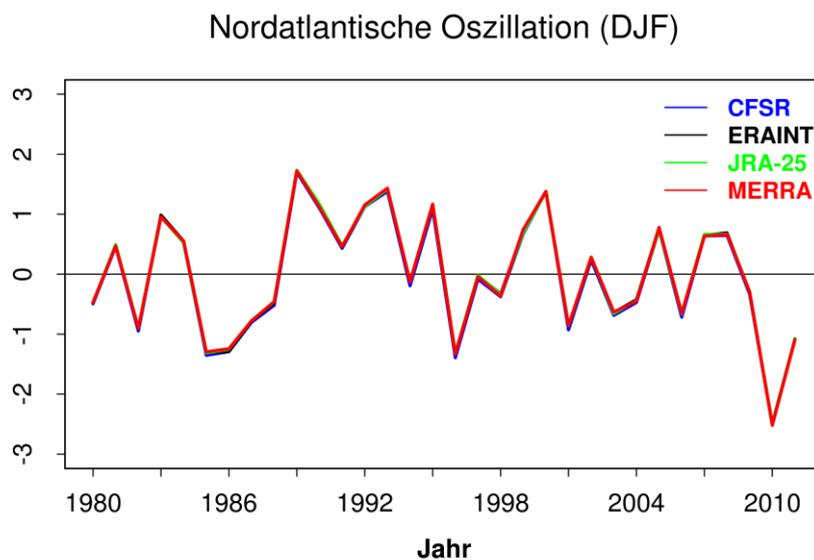


Abbildung 3.2: Index der Nordatlantischen Oszillation im Winter (DJF) des Zeitraums 1980-2011 (PC-basiert nach Hurrell 1995 (NAO-PC); vgl. Tabelle 4.1)

Auch für den zweiten methodischen Ansatz der unterschiedlichen Zirkulationstypen, der in Kapitel 4.1.2 näher beschrieben wird, zeigt sich für die verschiedenen Reanalysen eine große Homogenität. Exemplarisch wird dieser Befund für die saisonalen Auftretshäufigkeiten von acht Zirkulationstypen (nach GWT) im nordatlantisch-europäischen Raum im Winter (MAM) in Abbildung 3.3 dargestellt (Näheres zur Methode in Kapitel 4.1.2.1). Zwar sind einzelne geringfügige Abweichungen zwischen den Reanalysen zu erkennen (z. B. JRA25 mit geringerem Auftreten der Klasse „W“, dafür erhöhtem Auftreten der Klassen „NW“ und „NE“ im Vergleich mit den anderen Reanalysen), dennoch drückt sich die große Übereinstimmung auch mit einem p-Wert von 0,999 beim Chi-Quadrat-Homogenitätstest auf Kontingenztafeln (nach Pearson) aus.

Wenn man die Auftretshäufigkeiten nicht, wie in Abbildung 3.3, über alle Jahre gemittelt darstellt, sondern die einzelnen Jahre des Zeitraums 1979-2011 saisonal aggregiert analysiert (vgl. Abbildung 3.4; Farbgebung in Anlehnung an Abbildung 3.3), gelangt man ebenfalls zu dem Ergebnis, dass sich die in das Projekt einbezogenen Reanalysen – trotz

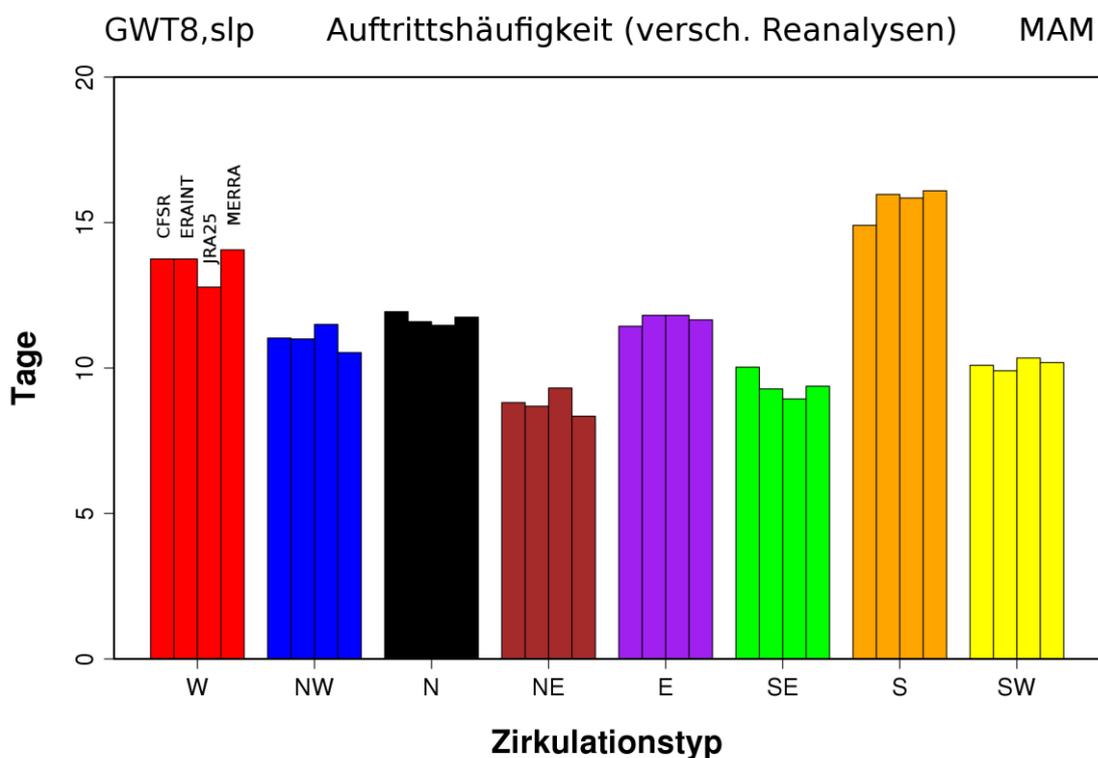


Abbildung 3.3: Saisonal gemittelte Auftretshäufigkeiten von acht Zirkulationstypen (Großwettertypen nach Beck; vgl. Kapitel 4.1.2.1) im nordatlantisch-europäischen Raum (90°W-40°O, 20°N-80°N) im Frühjahr (MAM) für verschiedene Reanalysen (CFSR, ERAINT, JRA25, MERRA).

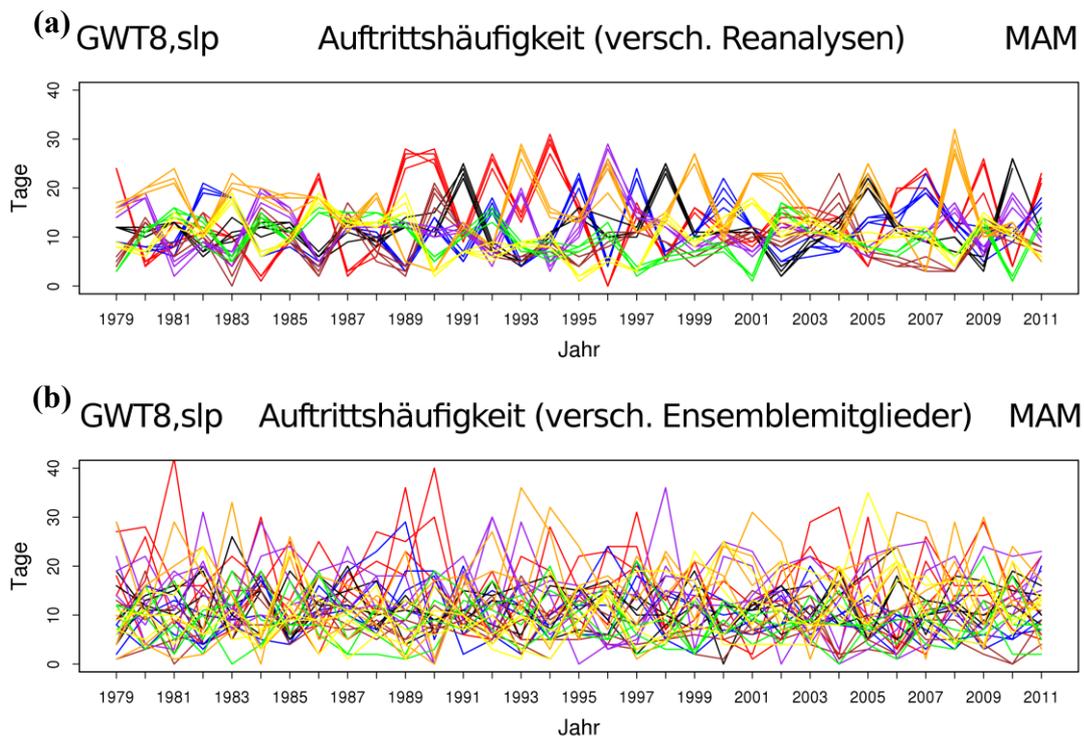


Abbildung 3.4: Saisonale Auftrittshäufigkeiten (Zeitraum 1979-2011) von acht Zirkulationstypen (Großwettertypen nach Beck; vgl. Kapitel 4.1.2.1) im nordatlantisch-europäischen Raum (90°W - 40°O , 20°N - 80°N) im Frühjahr (MAM), dargestellt für die vier Reanalysen (a) und vier zufällig ausgewählte Ensemblemitglieder von **b1** (b). Die Farbgebung für die einzelnen Zirkulationstypen orientiert sich an derjenigen von Abbildung 3.3.

leichter Abweichungen – stark ähneln (Abbildung 3.4a). Als Kontrast dazu ist ein Beispiel mit vier zufällig ausgewählten Ensemblemitgliedern der zweiten Generation des dekadischen Vorhersagemodells gezeigt (Abbildung 3.4b). Das Modell ist scheinbar nicht in der Lage, den zeitlichen Verlauf der saisonal aggregierten Häufigkeitsverteilungen von Zirkulationstypen aufzulösen, was sich an der chaotischen Struktur zeigt, die an dieser Stelle allerdings nicht näher erläutert wird.

Aufgrund der Ähnlichkeit der einzelnen Reanalysen wird in der vorliegenden Arbeit die Reanalyse ERAINT für weitere Untersuchungszwecke verwendet, was darin begründet ist, dass diese auch beim MiKlip-Projekt, innerhalb dessen diese Arbeit entstand, als Referenz verwendet wurde.

4 Methoden

Ein Ziel der Arbeit ist es, zu evaluieren, wie die Dynamik der Atmosphäre in den dekadischen Vorhersagen repräsentiert wird. Wie bereits erwähnt wurde, wird die Dynamik der Atmosphäre mit zwei unterschiedlichen Ansätzen erfasst: zum einen mittels Telekonnectionsindizes, zum anderen mithilfe von Zirkulationstypen. Beide Ansätze werden in Kapitel 4.1 näher erläutert. Die Evaluation (Kapitel 4.2), d. h. die Frage nach der Repräsentation der atmosphärischen Dynamik im dekadischen Klimamodell, folgt einem Ansatz, der sich auf dem Gebiet der dekadischen Vorhersage etabliert hat: die Analyse mittels sog. Hindcasts (Kapitel 4.2.1). Außerdem stellt sich bei der Evaluation eines dekadischen Vorhersagesystems die Frage nach geeigneten Möglichkeiten zur Messung der Vorhersageleistung. Diese Metriken, die sog. Skillmetriken, werden in Kapitel 4.2.2 erläutert.

4.1 Untersuchungsgegenstand

Die Atmosphäre ist neben der Biosphäre, der Kryosphäre, der Geosphäre, die wiederum in Pedosphäre und Lithosphäre aufgeteilt werden kann, und der Hydrosphäre Teil des globalen Klimasystems (s. Abbildung 4.1). Sie unterliegt dem Einfluss von externen Faktoren (z. B. solarer Strahlungsantrieb, Vulkanismus etc.) und steht zudem im intensiven wechselseitigen Austausch mit den anderen Teilsystemen. Weil Menschen und alle an der Erdoberfläche lebenden Organismen vom Zustand der Atmosphäre in ihrem täglichen Leben stark beeinflusst werden, ist eine valide Vorhersage zukünftiger Entwicklungen der Atmosphäre von besonderem Interesse.

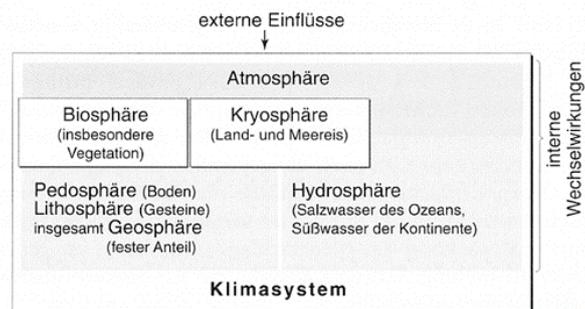


Abbildung 4.1: Schema des (globalen) Klimasystems (Quelle: Schönwiese 2013).

Möchte man analysieren, inwiefern die Atmosphäre und ihre dynamische Entwicklung durch ein dekadisches Vorhersagemodell repräsentiert werden, eröffnet sich ein weites Feld möglicher Untersuchungsgegenstände. Aktuelle Studien beschäftigen sich auf dem Gebiet der dekadischen Vorhersage z. B. mit der Repräsentation von Winterstürmen auf der Nordhemisphäre (Kruschke et al. 2015), der globalen, oberflächennahen Temperatur (Pohlmann et al. 2013), der winterlichen Situation in Europa und Nordamerika gemessen an verschiedenen atmosphärischen Parametern (Scaife et al. 2014) oder dem Zusammenhang zwischen der globalen Lufttemperatur und der nordatlantischen multidekadischen Variabilität (García-Serrano & Doblas-Reyes 2012). Analysen zur Dynamik der Atmosphäre haben v. a. die Repräsentation der Nordatlantischen Oszillation (NAO) in dekadischen Vorhersagesystemen zum Thema (z. B. Smith et al. 2014).

Genau wie die erwähnten Studien konzentriert sich die vorliegende Arbeit auf lediglich einen Aspekt der Atmosphäre, nämlich die Wiedergabe ihrer Dynamik. Die beiden Ansätze, die dazu angewendet wurden, werden im Folgenden dargestellt. Zunächst wird auf die verschiedenen Telekonnectionsindizes (Kapitel 4.1.1) eingegangen, anschließend werden die Zirkulationstypen, welche mithilfe unterschiedlicher Klassifikationsverfahren ermittelt werden, erläutert (Kapitel 4.1.2).

4.1.1 Telekonnectionsindizes

In Telekonnectionsindizes kommt die interne Variabilität der Atmosphäre zum Ausdruck, weshalb sie ein interessantes Forschungsfeld für dekadische Vorhersagen darstellen. Telekonnectionen werden von Wallace & Gutzler (1981, S. 784) als „signifikante, gleichzeitig auftretende Korrelationen zwischen zeitlichen Schwankungen klimatologischer Pa-

parameter an weit entfernten Punkten der Erde“ bezeichnet. Es handelt sich also um Fernkopplungen im Klimasystem, deren zeitliche Ausprägungen in Telekonnectionsindizes zum Ausdruck kommen. In der Klimawissenschaft haben sich mehrere bedeutende Telekonnectionsindizes etabliert. Mit der Nordatlantischen Oszillation (NAO) bzw. dem Southern Oscillation Index (SOI) sollen nur zwei Beispiele an dieser Stelle genannt werden. Tabelle 4.1 gibt eine Übersicht über alle untersuchten Indizes, die Methode ihrer Berechnung sowie die Variable, die der jeweiligen Berechnung zugrunde liegt. In der letzten Spalte der Tabelle wird auf die Referenz verwiesen, in der das jeweilige Berechnungsverfahren ausführlich erläutert wird.

Zwar weisen die einzelnen Indizes unterschiedliche Berechnungsverfahren auf, die im nächsten Kapitel einzeln erläutert werden, allerdings basieren alle Berechnungen auf Anomaliefeldern des Bodenluftdrucks (engl. *sea level pressure*, SLP) oder der geopotentiellen Höhe (engl. *geopotential height*, ZG) des 500 (ZG500) bzw. 700 hPa-Niveaus (ZG700). Um den Einfluss der Aggregierung zu untersuchen, wurden sowohl monatlich als auch saisonal aggregierte Anomaliefelder (mit Bezug auf den Untersuchungszeitraum der Beobachtungsdaten) gebildet. Weisen Modell- und Beobachtungsdaten eine unterschiedliche räumliche Auflösung auf, wurden die Beobachtungsdaten mit ihren unterschiedlichen Auflösungen mittels eines einfachen Interpolationsverfahrens (vgl. Kapitel 3.2) an die Modelldaten angepasst. Die Indizes wurden für die vier Standardjahreszeiten (MAM, JJA, SON, DJF) berechnet. Bei monatlicher Aggregierung der Daten wurde der monatliche Einfluss jeweils entfernt, um eine falsche Vorhersageleistung basierend auf der intraannuellen Variabilität der Indizes zu eliminieren. Zudem wurde eine Korrektur des Bias‘, d. h. des systematischen Fehlers zwischen Vorhersage und Beobachtung (vgl. Hawkins et al. 2014), vorgenommen. Diese wurde auf die Datenfelder des Luftdrucks bzw. der geopotentiellen Höhe vor der Berechnung der Telekonnectionsindizes angewendet. Der Biaskorrektur liegt dabei eine vereinfachte Annahme eines konstanten Bias‘ über die Zeit zugrunde, sodass die Standardprozedur (z. B. Goddard et al. 2013) mit der Subtraktion des mittleren Bias‘ von allen Vorhersagen angewendet werden konnte.

Tabelle 4.1: Übersicht über die Telekonnexionsindizes (unter Angabe ihrer Berechnungsgrundlagen und der jeweiligen Referenz.), die in der vorliegenden Arbeit zur Analyse der dekadischen Vorhersagbarkeit verwendet wurden.

Index (Abk.)	Index	Variable	Methode	Referenz
AAO-PC	Antarktische Oszillation, PC-basiert	ZG700 (700 hPa)	Zonale Mittel	Mo (2000)
AAO-ZM	Antarktische Oszillation, zonale Mittel	ZG700 (700 hPa)	Zonale Mittel	Gong & Wang (1999), Nan & Li (2003)
EA	East Atlantic (Pattern)	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)
EAWR	East Atlantic/ Western Russia (Pattern)	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)
EPNP	East Pacific/ North Pacific (Pattern)	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)
NAO	Nordatlantische Oszillation	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)
NAO-GP	Nordatlantische Oszillation, Gitterpunkt-basiert (nach Hurrell)	SLP	Gitterpunkt-werte	Hurrell (1995)
NAO-PC	Nordatlantische Oszillation, PC-basiert (nach Hurrell)	SLP	PCA	Hurrell (1995)
PNA	Pacific/ North American (Pattern)	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)
POL	Polar/Eurasia (Pattern)	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)
SCAND	Scandinavia (Pattern)	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)
SOI	Southern Oscillation Index	SLP	Gitterpunkt-werte	Trenberth (1984)
TNH	Tropical/ Northern Hemisphere (Pattern)	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)
WP	West Pacific (Pattern)	ZG500 (500 hPa)	RPCA	Barnston & Livezey (1987)

4.1.1.1 Nordatlantische Oszillation (nach Hurrell)

Einer der bedeutendsten und auch bekanntesten Telekonnektionsindizes ist die Nordatlantische Oszillation (NAO), die eine Hauptquelle interannueller (und längerer) Variabilität auf der Nordhemisphäre darstellt (Hurrell 1995). Sie bezieht sich auf die Verteilung von Luftmassen über dem nördlichen („Islandtief“) und dem subtropischen Atlantik („Azorenhoch“) und schwingt von der einen Phase zu der anderen, was mit großräumigen Veränderungen der bodennahen Temperatur, der Windsituation und dem Niederschlag über dem Nordatlantik und den angrenzenden Landflächen einhergeht (Hurrell & Deser 2009): Sind die beiden Drucksysteme deutlich ausgeprägt, d. h. gibt es eine starke Druckdifferenz zwischen Islandtief und Azorenhoch, spricht man von einer positiven Phase der NAO (NAO+). Es ergeben sich starke Westwinde, die v. a. im Winter zu einer milden und feuchten Witterung in Mitteleuropa und – durch Verstärkung des Kanarenstroms – zu eher trockenen und kühleren Bedingungen im westlichen Mittelmeerraum führen. Demgegenüber tritt während einer negativen Phase der NAO (NAO-), bei der die Druckdifferenz zwischen Islandtief und Atlantikhoch gering ausgeprägt ist, für gewöhnlich eine Abschwächung der Westwinde ein. Typischerweise stellt sich in Mitteleuropa ein mit östlichen Winden verbundenes Hochdruckgebiet ein, was zu einer Strengwinterkonstellation führen kann. Im Mittelmeerraum ist demgegenüber mit zunehmender Niederschlagstätigkeit zu rechnen (Lauer & Bendix 2006, Schönwiese 2013).

Die Berechnung des Index‘ der Nordatlantischen Oszillation kann nach verschiedenen Vorgehensweisen erfolgen. Für die vorliegende Arbeit wurde auf die zwei Berechnungsverfahren von HURRELL (Hurrell 1995, Hurrell et al. 2003) zurückgegriffen. Zum einen ist der NAO-Index nach Hurrell als Differenz der standardisierten Luftdruckanomalien zwischen den beiden Klimastationen Lissabon (Portugal) und Reykjavik (Island) definiert. Da in der vorliegenden Arbeit Luftdruckfelder und keine Stationsdaten verwendet wurden, wurde auf ausgewählte Gitterpunkte, die um die Klimastationen liegen, zurückgegriffen. Für Lissabon ($38^{\circ} 43' \text{N}$, $9^{\circ} 20' \text{W}$) gingen diejenigen Gitterpunkte gemittelt in die Berechnung ein, die zwischen 37°N und 41°N bzw. 7°W und 11°W liegen, für Reykjavik ($64^{\circ} 9' \text{N}$, $21^{\circ} 56' \text{W}$) die Gitterpunkte, die zwischen 64°N und 67°N bzw. 20°W und 24°W liegen. Der Gitterpunkt-basierte NAO-Index, der im Folgenden als NAO-GP bezeichnet wird, wurde für die Analyse der dekadischen Vorhersagbarkeit jeweils separat

für die Beobachtungsdaten und die einzelnen Ensemblemitglieder der Modelldaten ermittelt. Der Vorteil der Methode liegt darin, dass sie relativ einfach zu berechnen und zu verstehen ist. Ein Nachteil liegt allerdings darin, dass die Gitterpunkte, die zur Berechnung des NAO-Index⁴ verwendet werden, räumlich fixiert sind und so die Bewegung der Luftdruckzentren nicht nachvollzogen wird, d. h. die Luftdruckzentren somit nicht korrekt erfasst werden können.

Eine Methode, die diesen Nachteil ausgleicht und eine räumlich vollständige Abbildung der Luftdruckzentren garantiert, ist die Berechnung des NAO-Index mittels einer Hauptkomponentenanalyse (engl. *Principal Component Analysis*, PCA). Die Methode der Hauptkomponentenanalyse ist nach Wilks (2006) das multivariate statistische Verfahren, das in den Atmosphärenwissenschaften am häufigsten zum Einsatz kommt. Es findet sich in allen einschlägigen Lehrwerken (z. B. Bahrenberg, Giese & Nipper 2003, Jolliffe 2002, von Storch & Zwiers 1999 oder Wilks 2006) und wird u. a. verwendet, um eine Reduktion der Datendimension zu erreichen: Aus einer Vielzahl an Variablen sollen deutlich weniger neue, unkorrelierte Größen, sog. Hauptkomponenten (PC, Principal Components), gewonnen werden, die die wesentlichen Informationen enthalten und vom Hintergrundrauschen (engl. *noise*) befreit sind. Die Generierung der Hauptkomponenten erfolgt auf Basis einer aus $i = 1, \dots, n$ Variablen und $j = 1, \dots, m$ Ereignisfällen aufgebauten Datenmatrix. Die Hauptkomponenten ergeben sich schließlich als Linearkombinationen der Variablen und sind hinsichtlich ihrer Bedeutsamkeit hierarchisch angeordnet. D. h. die erste extrahierte Hauptkomponente vereint die größte Varianz des ursprünglichen Datenfeldes auf sich. Alle weiteren werden so gewählt, dass die Linearkombinationen die maximal mögliche Varianz aufweisen und statistisch unkorreliert zu den anderen sind. Als Resultat einer Hauptkomponentenanalyse ergeben sich die Hauptkomponentenwerte (engl. *scores*) sowie die Hauptkomponentenladungen (engl. *loadings*), die sich als Korrelation bzw. Kovarianz zwischen Hauptkomponentenwerten und Originalvariablen berechnen.

Die Hauptkomponentenanalyse kann verwendet werden, um die dominanten Muster der atmosphärischen Variabilität zu entdecken. Für die Berechnung des NAO-Index⁴ nach Hurrell wurde eine s-modale PCA (SPCA) verwendet, die im Allgemeinen für die Ermittlung von Telekonnektionsindizes geeignet ist (Compagnucci & Richman 2008). Bei diesem Modus der Hauptkomponentenanalyse gehen die Gitterpunkte des Datenfeldes als

Variablen und die Zeitpunkte als Ereignisfälle in die Berechnung ein. Ziel der SPCA ist es, Zusammenfassungen von Raumeinheiten, sog. Modi der Variabilität bzw. Variabilitätszentren (Jacobbeit 2010), zu erhalten, die in ihrer zeitlichen Variabilität möglichst optimal durch eine gemeinsame Zeitreihe repräsentiert werden. Beim s-modalen Modus der PCA wird die gemeinsame Zeitreihe durch die Hauptkomponentenwerte bereitgestellt, deren räumliche Verortung erfolgt mit den Hauptkomponentenladungen (Rathmann 2009). Die SPCA wurde auf ausgeschnittene Anomaliefelder der nordatlantisch-europäi-

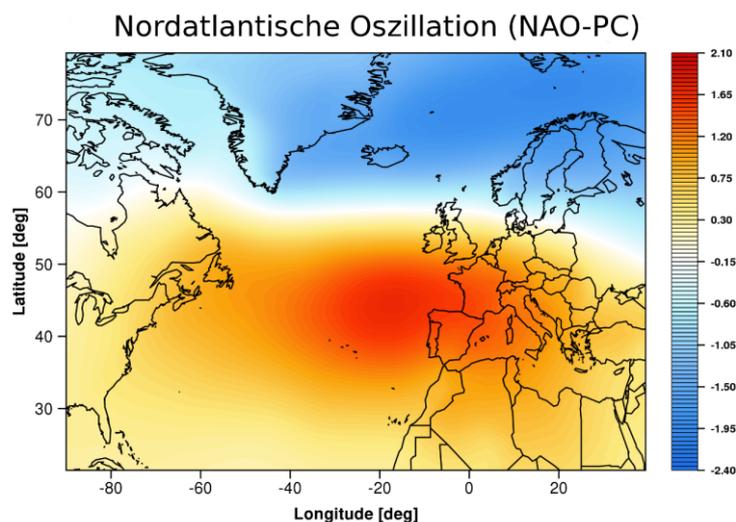


Abbildung 4.2: Räumliches Muster der Nordatlantischen Oszillation (NAO-PC) im Winter (DJF).

schen Region (20°N - 80°N , 90°W - 40°E) angewendet, eine anschließende Rotation der Daten wurde nicht durchgeführt. Der PC-basierte NAO-Index, im Folgenden als NAO-PC bezeichnet, ist definiert als die erste, führende Hauptkomponente (Hurrell et al. 2003). Die Hauptkomponentenladungen zeigen in Abbildung 4.2. für den Winter (DJF) das charakteristische räumliche Muster mit den beiden Variationszentren über Island und den Azoren.

Für die Analysen der dekadischen Vorhersagen wurde die PCA jeweils für die Beobachtung berechnet und anschließend auf die Modelldaten projiziert. D. h. um die Zeitreihen für die Modelldaten zu erhalten, wurden für jeden vorhandenen Zeitpunkt die Datenfelder der Modelldaten mit den Hauptkomponentenladungen der Beobachtung multipliziert. Diese Projektionsmethode wurde gewählt, um zum einen die Rechenzeiten zu verkürzen, da bei jeder Analyse nur eine PCA (für die Beobachtung) und nicht viele PCAs (für die Beobachtung und alle Ensemblemitglieder der Modelldaten) berechnet werden müssen.

Zum anderen wird die Konsistenz der Datenauswertung gewahrt, wenn sich alle Ensemblemitglieder auf die Hauptkomponentenanalyse der Beobachtung beziehen und nicht jeweils eine eigene Berechnung des NAO-Index‘ einbringen.

Auch wenn die beiden Berechnungsmethoden (NAO-GP, NAO-PC) völlig unterschiedliche Ansätze verwenden, kommen sie dennoch zu ähnlichen Resultaten. Abbildung 4.3 zeigt den NAO-Index der Jahre 1980-2011 im Winter (DJF) für die beiden Verfahren im

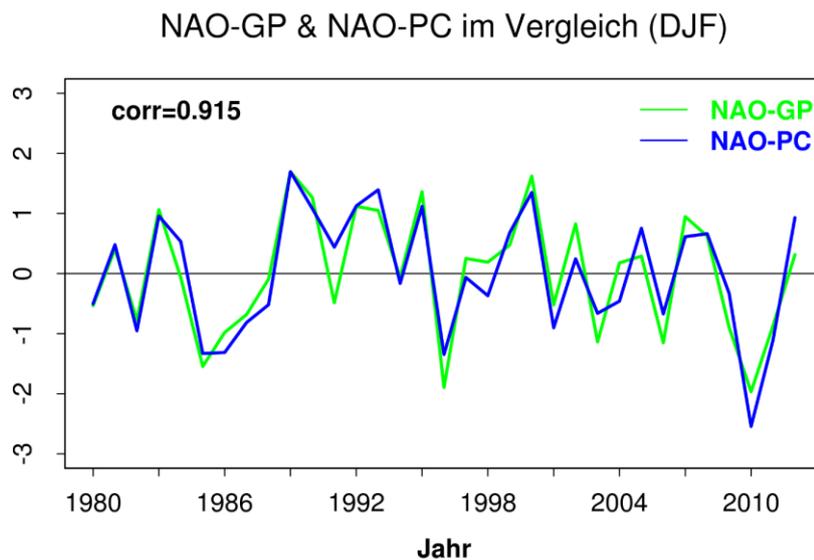


Abbildung 4.3: Vergleich der beiden Berechnungsmethoden NAO-GP und NAO-PC für die Jahre 1980-2011 (basierend auf ERAINT) im Winter (DJF) unter Angabe des Korrelationskoeffizienten nach Pearson.

Vergleich. Es ist eine große Übereinstimmung zwischen den beiden Zeitreihen mit nur geringen Abweichungen ersichtlich. Der Korrelationskoeffizient nach Pearson spiegelt den Zusammenhang mit einem hohen Wert von 0,915 wider. Auch in den sonstigen Jahreszeiten zeigen sich hohe Korrelationskoeffizienten, die bei 0,592 (SON) bzw. 0,563 (MAM) liegen. Lediglich der Sommer (JJA) fällt durch einen insignifikanten Zusammenhang mit $r=-0,081$ aus der Reihe, d. h. die beiden Verfahren resultieren in völlig unterschiedlichen Zeitreihen. Das liegt darin begründet, dass bei NAO-PC ein anderes räumliches Muster, der sog. Sommer-NAO, als in den übrigen Jahreszeiten entsteht.

4.1.1.2 CPC-Telekonnektionsmuster

Einen umfassenderen Ansatz zur Ermittlung von Telekonnektionsindizes auf der nördlichen Hemisphäre verwendet das CPC (Climate Prediction Center), eine Abteilung des amerikanischen Wetterdienstes NWS (National Weather Service), die sich mit der operativen Vorhersage klimatischer Variabilität und der Ermittlung der Herkunft wichtiger Klimaanomalien beschäftigt. Im Gegensatz zum Ansatz nach Hurrell wird bei den Untersuchungen zur klimatischen Variabilität auf der nördlichen Hemisphäre nicht nur eine Hauptkomponente extrahiert, sondern es werden mehrere Telekonnektionsmuster identifiziert. Die Berechnungsmethodik beruht auf den Ausführungen von Barnston & Livezey (1987), die zur Untersuchung der interannuellen Variabilität einen hauptkomponentenanalytischen Ansatz benutzen, der auf monatlich aufgelöste, nordhemisphärische Felder der geopotentiellen Höhe auf dem 500 hPa-Niveau angewendet wird. Im Gegensatz zu dem Ansatz von Wallace & Gutzler (1981) führen die Autoren eine Rotation der Hauptkomponenten durch (RPCA; Rotated Principal Component Analysis), wodurch die Hauptkomponenten besser meteorologisch interpretierbar sowie statistisch stabiler sind, wenn verschiedene Untersuchungszeiträume verglichen werden. Zudem nehmen sie eine kleinere Fläche der Hemisphäre ein (Barnston & Livezey 1987).

In der vorliegenden Arbeit wird auf das vom CPC angewendete Berechnungsverfahren zurückgegriffen, das auf den Ausführungen von Barnston & Livezey (1987) beruht, sich von diesem aber in kleinen Details unterscheidet. Bei dem Verfahren wird eine SPCA auf standardisierte Anomaliefelder der geopotentiellen Höhe auf dem 500 hPa-Niveau angewendet und zehn Hauptkomponenten extrahiert. Als räumlicher Ausschnitt geht die gesamte nördliche Hemisphäre unter Ausschluss der Tropen (20°N-90°N, 180°W-180°E) in die Berechnung ein. Anschließend erfolgt eine Rotation der zehn Hauptkomponenten nach dem Varimax-Kriterium, dem am häufigsten verwendeten Ansatz für die orthogonale Rotation (Wilks 2006). Die extrahierten Hauptkomponenten sind orthogonal so zu rotieren, dass die Varianz der Ladungen s_i^2 der 1...q Hauptkomponenten

$$\sum_{i=1}^q s_i^2 \quad (1)$$

ein Maximum wird, was sich durch ein rechenaufwändiges Iterationsverfahren, welches zu einer Näherungslösung führt, lösen lässt (Bahrenberg, Giese & Nipper 2003). Neben der besseren meteorologischen Interpretierbarkeit und der größeren statistischen Stabilität hat die Rotation zur Folge, dass sich die Anteile der erklärten Varianz gleichmäßiger über die extrahierten Hauptkomponenten verteilen, d. h. die erste Hauptkomponente verliert, die nachfolgenden gewinnen an Bedeutung. Am Gesamtvarianzerklärungsanteil ändert sich nichts. Während bei der Methode nach Hurrell bei der Berechnung des Index‘ der Nordatlantischen Oszillation (s. Kapitel 4.1.1.1) ohne eine Rotation der PCs angestrebt wird, möglichst viel Varianz auf die erste PC zu vereinen, werden bei der Methode nach dem CPC durch die Rotation der zehn PCs die Varianzerklärungsanteile gleichmäßiger verteilt und mehrere stabile Telekonnectionsindizes ermittelt.

Während das CPC die Ermittlung der Telekonnectionsindizes für alle zwölf Monate durchführt, wurden diese in der vorliegenden Arbeit für die vier Jahreszeiten (MAM, JJA, SON, DJF) ermittelt. Die Berechnung erfolgte auf Basis monatlicher bzw. saisonaler Aggregation der Daten, d. h. es gehen drei Werte bzw. ein Wert in die jahreszeitliche Analyse ein, wobei bei monatlicher Aggregation wiederum der monatliche Einfluss eliminiert wurde. Die RPCA wurde (wie in Kapitel 4.1.1.1 für die NAO beschrieben) bei der Analyse der dekadischen Vorhersagen auf die Beobachtungsdaten angewendet und auf die Modelldaten zur Ermittlung der Zeitreihen projiziert. Neben dem geringeren Rechenaufwand liegt ein Vorteil der Methode des Projizierens darin, dass das Problem der Zuweisung der einzelnen PCs zueinander, das sich ergibt, wenn man die RPCA sowohl auf die Beobachtung als auch alle Ensemblemitglieder der dekadischen Vorhersage (bzw. auf die über alle Ensemblemitglieder gemittelte Vorhersage) anwendet, gelöst wird.

Von den jeweils zehn extrahierten Hauptkomponenten, die zwischen 75,05 % (JJA) und 82,54 % (DJF) der Originalvarianz erklären, treten acht in nahezu allen Jahreszeiten stabil auf. Diese definieren acht Telekonnectionsmuster (EA, EAWR, EPNP, NAO, PNA, POL, SCAND, WP, s. Tabelle 4.1), die für die Analysen zur dekadischen Vorhersagbarkeit verwendet wurden. Außerdem wurde noch das Telekonnectionsmuster TNH (Tropi-

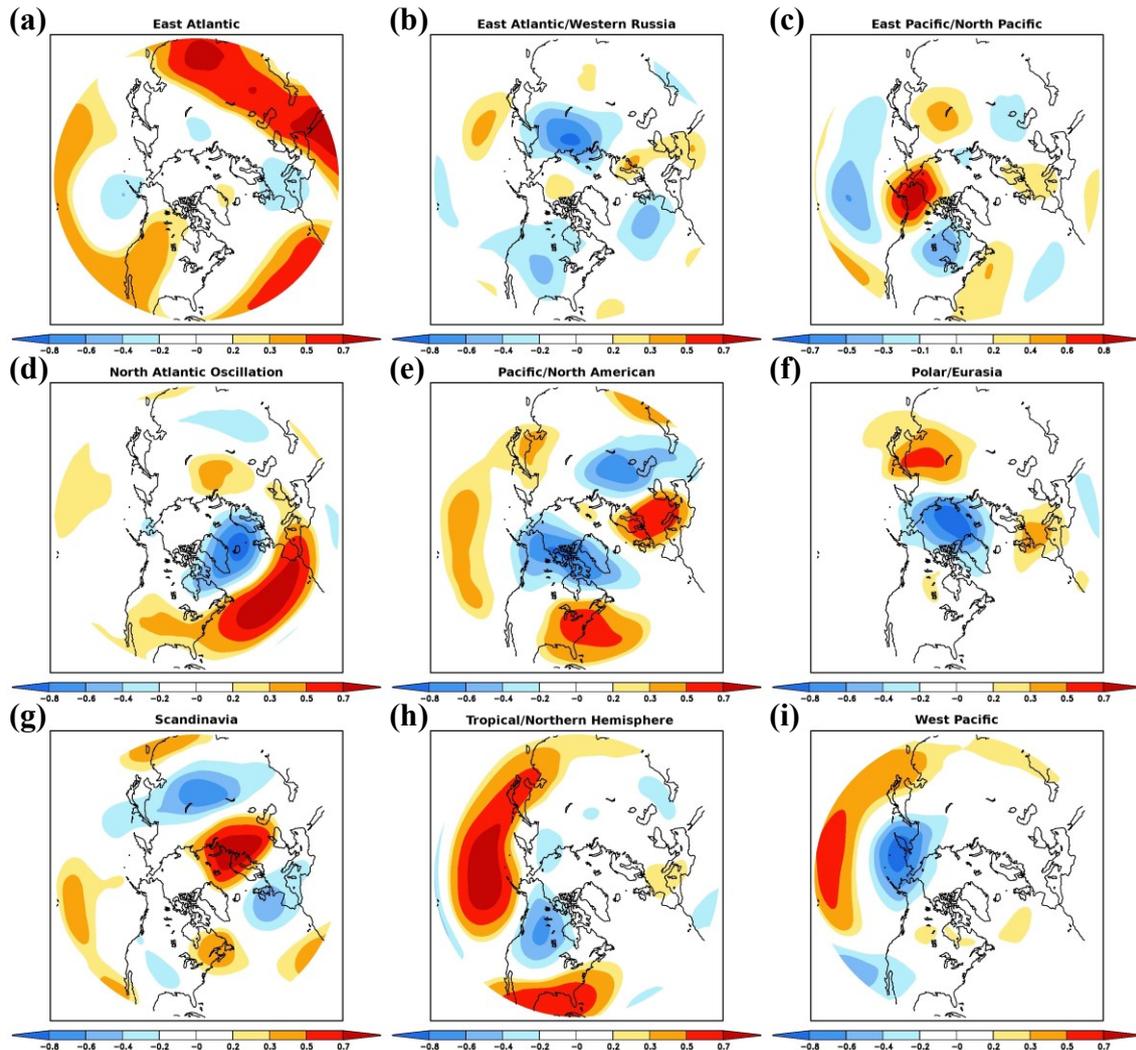


Abbildung 4.4: Übersicht über die neun ausgewählten Telekonnektionsmuster nach der CPC-Methode im Winter (Datengrundlage: ERAINT 1980-2011).

cal/Northern Hemisphere), welches lediglich im Winter zu beobachten ist, zu den Analysen hinzugefügt. Abbildung 4.4 zeigt eine Übersicht über die zugehörigen Telekonnektionsmuster, die auf Basis saisonal aggregierter Werte von ERAINT der Jahre 1980-2011 erstellt wurde. Dargestellt sind die Telekonnektionsmuster in alphabetischer Anordnung jeweils für den Winter (DJF), da in dieser Jahreszeit der Gesamtvarianzerklärungsanteil am höchsten ist. Lediglich das Muster von East Pacific/North Pacific (EPNP; Abbildung 4.4c) entstammt dem Frühjahr (MAM), da es im Winter nicht beobachtet werden kann. Zu sehen sind in Abbildung 4.4 die jeweiligen Variationszentren in roter bzw. blauer Farbe. Für die bereits im vorherigen Kapitel beschriebene Nordatlantische Oszillation (NAO; Abbildung 4.4d), die eine wichtige Telekonnektion zur Erklärung der nordhemi-

sphärischen, klimatischen Variabilität darstellt und in allen Jahreszeiten identifiziert werden kann, sind die beiden Variationszentren über Island und den Azoren gut erkennbar. Diese sind ebenfalls bei den übrigen Telekonnektionsmustern (mehr oder weniger) deutlich zu beobachten. Bei der Karte von Pacific/North American (PNA; Abbildung 4.4e) ist ein Zentrum südlich der Aleuten sowie eines mit gegensätzlichem Vorzeichen im Norden des amerikanischen Kontinents zu finden, wie es von Barnston & Livezey (1987) in ihren Ausführungen beschrieben wird. Ebenso werden von Barnston & Livezey (1987) die beiden Telekonnektionsmuster West Pacific (WP; Abbildung 4.4i) bzw. Tropical Northern Hemisphere (TNH; Abbildung 4.4h) mit Variationszentren über dem nördlichen Pazifik (östlich von Kamtschatka) und dem südlicheren Teil bzw. über dem Pazifik (westlich des amerikanischen Kontinents) und dem Bereich der Großen Seen charakterisiert. Als weiteres Telekonnektionsmuster mit Zentren im Bereich des Pazifiks (eines über Alaska, das zweite mit gegensätzlichem Vorzeichen weiter südlich davon) ist EPNP (East Pacific/North Pacific; von Barnston & Livezey (1987) als East Pacific Pattern bezeichnet; Abbildung 4.4c) zu nennen. Zudem treten weitere Beispiele im Bereich Eurasiens bzw. des Atlantiks auf, die von Barnston & Livezey (1987) nicht mehr eindeutig zugeordnet werden können: East Atlantic (EA; Abbildung 4.4a) mit Zentren über Westeuropa und einem Band vom Mittelmeerraum bis nach Asien sowie die beiden Telekonnektionsmuster East Atlantic/Western Russia (EAWR; Abbildung 4.4b) und Scandinavia (SCAND; Abbildung 4.4g), die beide Variationszentren über Skandinavien und (mit entgegengesetztem Vorzeichen) westlich und östlich davon aufweisen. Zu guter Letzt soll Polar/Eurasia (POL; Abbildung 4.4f) erwähnt werden, was zwei räumlich konzentriert auftretende Zentren über der Arktis sowie dem Osten des eurasischen Kontinents aufweist.

4.1.1.3 Südliche Oszillation (SO)

Eine der bekanntesten und bedeutendsten Luftdruckschwankungen der Südhemisphäre ist die Südliche Oszillation (SO, Southern Oscillation), die die atmosphärische Komponente des komplex zwischen Ozean und Atmosphäre gekoppelten ENSO-Phänomens (El Niño-Southern Oscillation) charakterisiert, wobei El Niño die ozeanischen Zusammenhänge erfasst (u. a. Schönwiese 2013). Aufgrund ihres deutlichen Signals interannueller Variabilität (Trenberth 1984), die sie vom Hintergrundrauschen kurzfristiger Klimaent-

wicklungen abhebt, bietet die Südliche Oszillation eine der bzw. die Hauptquelle saisonaler Vorhersagbarkeit (Smith et al. 2012). Dieser Befund prädestiniert die Südliche Oszillation, sie hinsichtlich ihrer dekadischen Vorhersagbarkeit zu untersuchen.

Die räumlich-zeitliche Ausprägung der Südlichen Oszillation erfährt ihre Wiedergabe im Southern Oscillation Index (SOI) als Maß für die Intensität oder Stärke der Walker-Zirkulation. In der Literatur lassen sich mehrere Definitionen für den SOI finden. Für die vorliegende Arbeit wird auf die von Trenberth (1984) zurückgegriffen, der den SOI als bodennahe Luftdruckdifferenz zwischen Tahiti und Darwin definiert, zwei Klimastationen, die in der Nähe der Variationszentren liegen. Aus den Gitterpunktdaten wurden die bodennahen Luftdruckwerte (SLP, s. Tabelle 4.1) um die Insel Tahiti ($17^{\circ} 41'S$, $149^{\circ} 27'W$), die zwischen $15^{\circ}S$ und $19^{\circ}S$ bzw. $147^{\circ}W$ und $151^{\circ}W$ liegt, und die Werte um Darwin ($12^{\circ} 27'S$, $130^{\circ} 50'O$) ausgewählt, die sich zwischen $10^{\circ}S$ und $14^{\circ}S$ bzw. $128^{\circ}O$ und $132^{\circ}O$ befinden. Die über die jeweils ausgewählten Gitterpunkte gemittelten Anomalien gingen – bei monatlicher Auflösung wieder um den monatlichen Einfluss eliminiert – in die Berechnung des SOI ein. Die Indexwerte wurden (wie beim NAO-GP / NAO-PC) separat für die Beobachtungsdaten bzw. die Modelldaten berechnet, wobei die Anomalienbildung wiederum mit Bezug auf den Untersuchungszeitraum der Beobachtung erfolgte.

4.1.1.4 Antarktische Oszillation (AAO)

Ein weiterer auf der südlichen Hemisphäre angesiedelter Modus der Variabilität ist die Antarktische Oszillation (AAO), die auch als Southern Annular Mode (SAM), bezeichnet wird. Die AAO beschreibt die Verteilung von Luftmassen zwischen den mittleren sowie den hohen Breiten und bezieht sich somit auf die Nord-Süd-Bewegung des Westwindgürtels, der die Antarktis umgibt. Die wechselnde Position des Westwindgürtels beeinflusst die Lage von Kaltfronten und Windsystemen der Mittelbreiten. Während positiver Phasen der AAO kontrahiert sich das Band der Westwinde gen Antarktis und sorgt für hohen Luftdruck und eher stabile, trockene Bedingungen im südlichen Australien (v. a. im Herbst/Winter). Demgegenüber dehnt sich das Band der Westwinde während negativer Phasen der AAO gen Äquator aus, was in ausgeprägten Tiefdruckbedingungen mit verstärktem Niederschlag und Sturmaktivität über dem südlichen Australien resultieren kann (BOM 2012). Das verdeutlicht die Bedeutung der Antarktischen Oszillation für die südliche Hemisphäre, um die klimatische Variabilität zu erklären (Gong & Wang 1999;

Thompson & Wallace 2000). Sie besitzt somit eine vergleichbare Bedeutung wie die Arktische Oszillation für die Nordhemisphäre.

Für die Berechnung der AAO finden sich in der Literatur verschiedene Ansätze, die sich in einem oder mehreren Details mehr oder weniger voneinander unterscheiden. Für die vorliegende Arbeit wurde auf zwei Berechnungsverfahren zurückgegriffen. Zum einen erfolgt die Berechnung des Index der Antarktischen Oszillation als Differenz der zonalen Mittel von 35°S bzw. 65°S (in die Berechnung gingen alle zonalen Werte gemittelt ein, für die gilt $34^{\circ}\text{S} \leq zM_{35} \leq 36^{\circ}\text{S}$ bzw. $64^{\circ}\text{S} \leq zM_{65} \leq 66^{\circ}\text{S}$) basierend auf Feldern der geopotentiellen Höhe des 700 hPa-Niveaus. Der Ansatz folgt damit in leicht variierender Form den Methoden von Nan & Li (2003) bzw. Gong & Wang (1999), die die AAO als Differenz der zonalen Mittel von 40°S und 70°S (Nan und Li 2003) bzw. 40°S und 65°S (Gong und Wang 1999) basierend auf dem Luftdruck auf Meereshöhe (SLP) berechnen. Dieser Index wird im Folgenden mit AAO-ZM bezeichnet.

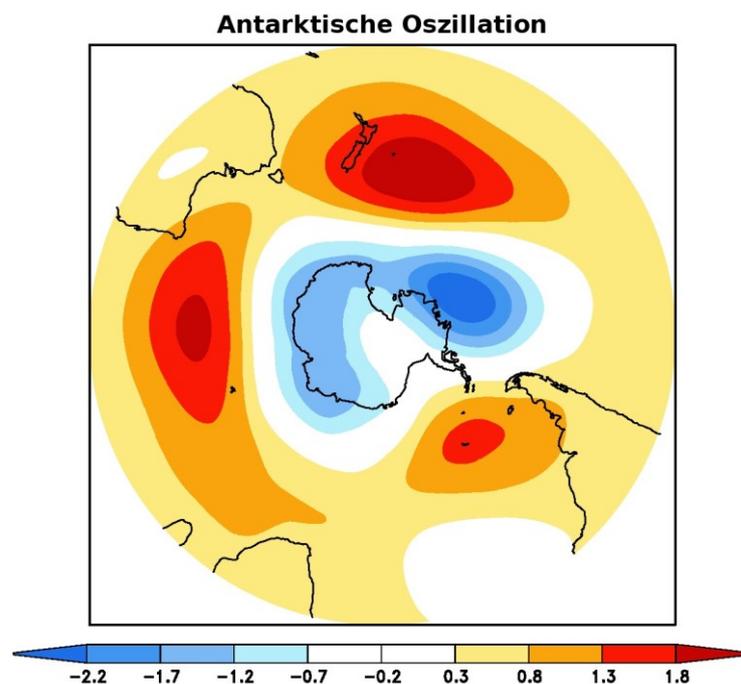


Abbildung 4.5: Telekonnectionsmuster der Antarktischen Oszillation (AAO-PC) im Sommer (Datengrundlage: ERAINT 1979-2011).

Die zonalen Mittel wurden in der vorliegenden Arbeit basierend auf Feldern der geopotentiellen Höhe für die beschriebenen Breitengrade gewählt, da hier die größten Korrelationen zu der zweiten Berechnungsmethode auftreten. Bei dieser Methode wurde auf ein

Verfahren des CPC zurückgegriffen (als Referenz s. Mo 2000). Vergleichbar zur Berechnung des Index‘ der Nordatlantischen Oszillation (NAO-PC, s. Kapitel 4.1.1.1) wurde auf monatlich/saisonal gemittelte, ausgeschnittene Anomaliefelder (20°S - 90°S , 180°W - 180°E) der geopotentiellen Höhe auf dem 700 hPa-Niveau (ZG700) eine s-modale PCA (ohne anschließende Rotation der extrahierten Hauptkomponenten) angewendet. Die führende Hauptkomponente, die zwischen 32,29 % (SON) und 38,25 % (DJF) der ursprünglichen Varianz erklärt, definiert den PC-basierten Index der Antarktischen Oszillation, der im Folgenden mit AAO-PC abgekürzt wird. Die PCA wurde bei den Analysen der dekadischen Vorhersagen wiederum für die Beobachtung durchgeführt und die Hauptkomponentenladungen anschließend auf die Modelldaten projiziert. Abbildung 4.5 veranschaulicht das Telekonnektionsmuster der Antarktischen Oszillation im Sommer (JJA) mit einem Variationszentrum im zentralen Bereich der Antarktis sowie einem bandförmig angeordneten, äquatorwärts an die Antarktis anschließenden Anomaliebereich mit umgekehrtem Vorzeichen, was die Sinnhaftigkeit der Verwendung von zonalen Mitteln bei der alternativen Berechnung der AAO (AAO-ZM) unterstreicht.

Abbildung 4.6 veranschaulicht den Zusammenhang zwischen den beiden Berechnungsverfahren für saisonal gemittelte Werte. Dargestellt ist sowohl der auf zonalen Mitteln basierende (AAO-ZM) als auch der PC-basierte Index der Antarktischen Oszillation (AAO-PC) in der kalten Jahreszeit der südlichen Hemisphäre (JJA). Vergleichbar zum

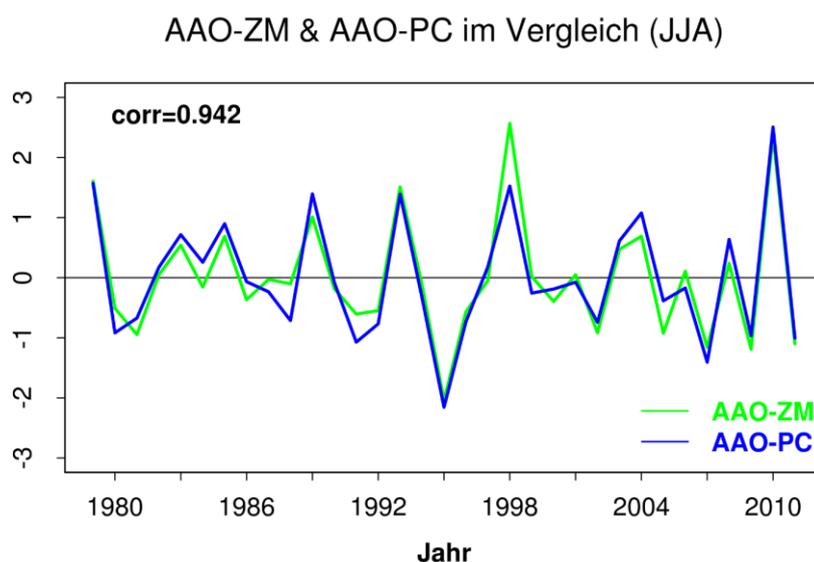


Abbildung 4.6: Vergleich der beiden Berechnungsmethoden AAO-ZM und AAO-PC für die Jahre 1979-2011 (basierend auf ERA-Interim) im Sommer (JJA) unter Angabe des Korrelationskoeffizienten nach Pearson.

Index der NAO ist die kalte Jahreszeit von besonderem Interesse, da – in Übereinstimmung mit den Ergebnissen des CPC – beim PC-basierten Verfahren (AAO-PC) am meisten Variabilität erfasst wird, d. h. die führende Hauptkomponente die höchsten Werte aufweist. Zwischen AAO-PC und AAO-ZM zeigt sich eine hohe Übereinstimmung der beiden Zeitreihen (s. Abbildung 4.6), was auch der Korrelationskoeffizient nach Pearson von $r=0,942$ ausdrückt. Auch in anderen Jahreszeiten nimmt der Korrelationskoeffizient Werte nahe 1 an: Der stärkste Zusammenhang zwischen den beiden Berechnungsmethoden lässt sich im südhemisphärischen Sommer (DJF) beobachten ($r=0,970$). Die geringsten Werte treten in den Übergangsjahreszeiten auf (MAM mit $r=0,881$, SON mit $r=0,871$), wobei dennoch ein hochsignifikanter Zusammenhang besteht. Wie bei der Nordatlantischen Oszillation (vgl. Kapitel 4.1.1.1) lässt sich auch für die Antarktische Oszillation beobachten, dass die beiden Verfahren, die völlig unterschiedliche Ansätze verwenden, zu ähnlichen Ergebnissen kommen.

4.1.2 Zirkulationstypen

Für die Untersuchungen zur dekadischen Vorhersagbarkeit der atmosphärischen Dynamik wird – neben den Analysen basierend auf Telekonnectionsindizes – auf einen zweiten Ansatz zurückgegriffen. In der vorliegenden Arbeit soll überprüft werden, inwiefern Resultate von atmosphärischen Klassifikationsverfahren, d. h. Zirkulationstypen, hinsichtlich ihres Auftretens als auch ihrer typinternen Charakteristika (wie z. B. ihrer Persistenz) repräsentiert werden. Klassifikationsverfahren haben eine weit zurückreichende Geschichte in der Meteorologie und Klimatologie (Huth et al. 2008). Mit ihrer Hilfe wird versucht, die klimatische Variabilität der Atmosphäre (als kontinuierliche Funktion vieler Variablen) anhand einer überschaubaren Anzahl an repräsentativen Mustern (Klassen/Typen/...) der Zirkulation als getrennte, diskrete Zustandsbedingungen zu erfassen (Yarnal 1993, Huth 1996). Aufgrund der vielseitigen Verwendungsmöglichkeit der Resultate einer Klassifikation, z. B. Anwendung von Zirkulationsmustern zur Abschätzung der Oberflächentemperatur mittels Downscaling (Enke et al. 2005), sind Fortschritte in der dekadischen Vorhersagbarkeit diesbezüglich von besonderem Interesse.

In der Literatur findet sich eine Vielzahl an unterschiedlichen Ansätzen zur Klassifizierung der atmosphärischen Zirkulation (vgl. Huth et al. 2008, Philipp et al. 2010). Grundsätzlich lässt sich zwischen „manuellen“ oder „subjektiven“, wie z. B. dem Katalog der Großwetterlagen Europas nach Hess & Brezowsky (1977), „automatisierten“ oder „objektiven“ sowie „hybriden“ oder „gemischten“ Verfahren unterscheiden (Huth et al. 2008). Subjektive Klassifikationen basieren auf vordefinierten Katalogen mit Mustern der Zirkulation. Eine Zuordnung individueller, stündlich/täglich/monatlich/... aufgelöster Felder erfolgt anhand des Expertenwissens erfahrenerer Meteorologen. Demgegenüber benutzen objektive Klassifikationen (wie z. B. das clusteranalytische Verfahren SANDRA in Kapitel 4.1.2.2) statistische Methoden und Computertechnologie (Philipp et al. 2007). Sie besitzen den Vorteil, nicht vom Kenntnisstand und Urteilsvermögen des Bearbeiters abzuhängen. Außerdem sind sie deutlich schneller durchführbar und zudem objektiv nachvollziehbar, wenngleich auch bei objektiven Verfahren subjektive Entscheidungen (z. B. hinsichtlich der Klassenanzahl) getroffen werden müssen. Hybride Verfahren nutzen wie die subjektiven Ansätze vordefinierte Klassen, die Zuordnung der individuellen Felder zu diesen geschieht allerdings automatisiert nach objektiven Kriterien (z. B. Großwettertypen nach Beck in Kapitel 4.1.2.1).

In der vorliegenden Arbeit wurden drei verschiedene objektive Klassifikationsverfahren vergleichend angewendet. Sie unterscheiden sich grundlegend in ihrer mathematischen Herangehensweise und werden im Folgenden näher erläutert. Unter den Klassifikationsverfahren verbirgt sich ein schwellenwertbasierter (GWT, Kapitel 4.1.2.1), ein clusteranalytischer (SANDRA/SAN, Kapitel 4.1.2.2) sowie ein eigenvektorbasierter Ansatz (TPCA, Kapitel 4.1.2.3). Die unterschiedlichen Verfahren wurden mit Fokus auf die

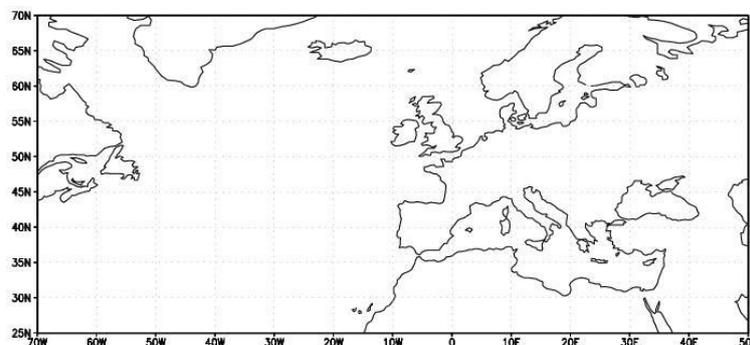


Abbildung 4.7: Lage und Ausdehnung der nordatlantisch-europäischen Untersuchungsregion.

nordatlantisch-europäische Region angewendet (90°W - 40°O , 20°N - 80°N , s. Abbildung 4.7), einer Schlüsselregion der dekadischen Vorhersage. Wie beim Ansatz mit Telekonnectionsindizes (s. Kapitel 4.1.1) basieren alle Berechnungen auf Anomaliefeldern von SLP oder ZG (für unterschiedliche Höhengniveaus). Bei einem solchen Vorgehen, d. h. bei der Berücksichtigung von lediglich die Zirkulation charakterisierenden Parametern, sprechen Huth et al. (2008) von „Zirkulationsklassifikationen“ und bezeichnen das Resultat der Klassifikation als „Zirkulationstyp“. D. h. im Gegensatz zu Wetterklassifikationen wurden keine klimatologischen Parameter wie Temperatur, Luftfeuchte oder Bedeckungsgrad berücksichtigt. Die Analysen wurden für unterschiedliche Jahreszeiten (MAM, JJA, SON, DJF) mit und ohne Korrektur des mittleren Bias‘ durchgeführt. Sie basieren auf 6-stündlich aufgelösten Beobachtungs- und Modelldaten, wobei bei den Berechnungen des jeweiligen Klassifikationsverfahrens nur die 12 Uhr-Werte Berücksichtigung fanden. Wie bei den Untersuchungen zu den Telekonnectionsindizes wurde die räumliche Auflösung der Beobachtungsdaten mittels eines Interpolationsverfahrens an diejenige der Modelldaten angepasst (s. Kapitel 4.1.1).

4.1.2.1 GWT – Großwettertypen nach Beck

Mit den Großwettertypen nach Beck wird kein vollständig neuer methodischer Ansatz, sondern ein Objektivierungsversuch eines bereits bestehenden subjektiven Klassifikationsverfahrens gewählt. In Anlehnung an die „Großwetterlagen“ nach Hess/Brezowsky entwickelt Beck eine eigene Klassifikation mit der Definition prototypischer Strömungsmuster (Beck 2000, Beck et al. 2007). Die prototypischen Ausprägungen der Großwettertypen lassen sich aufgrund ihres räumlichen Korrelationskoeffizienten zu drei idealen Grundmustern der Zirkulation definieren: ein zonales (W-E-Lage), ein meridionales (S-

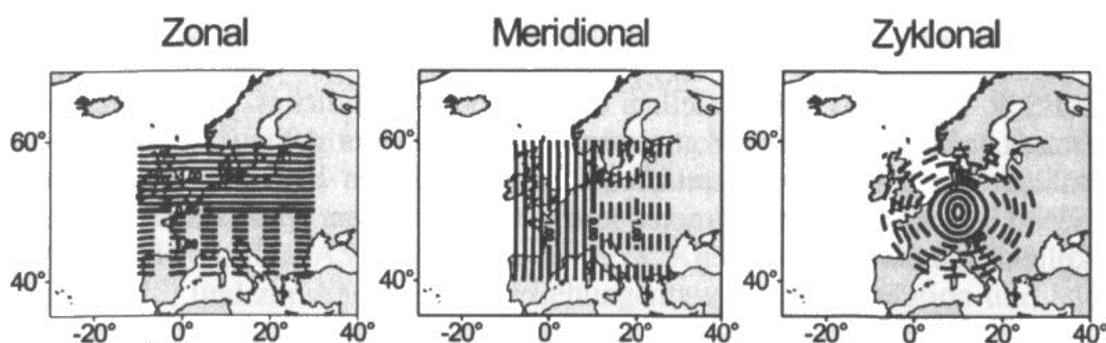


Abbildung 4.8: Idealtypische Grundmuster der Großwettertypen nach Beck (Quelle: Beck 2000).

N-Lage) und ein zyklonales (Zentraltieflage) Muster (s. Abbildung 4.8). Durch die Definition von Prototypen wird das häufig beobachtbare Problem umgangen, dass die aus automatisierten, computergestützten Klassifikationsverfahren resultierenden Muster die aus den subjektiven Verfahren bekannten, wichtigsten Zirkulationstypen nur in ungenügender Art und Weise widerspiegeln (Beck et al. 2007).

Aufgrund der räumlichen Korrelation mit diesen drei idealisierten Grundmustern der Zirkulation leiten sich die zentraleuropäischen Großwettertypen (GWT) als Koeffizienten der Zonalität (Z), der Meridionalität (M) bzw. der Zyklonalität (V, Vorticity) ab. Die Zuordnung eines individuellen Feldes zu einer Klasse geschieht aufgrund vorgegebener Schwellenwerte: Weist der Koeffizient V (bei der Korrelation des individuellen Felds mit den prototypischen Grundmustern) den betragsmäßig höchsten Wert auf, resultieren die beiden Großwettertypen mit zentralem Hoch (**H**; bei negativem V-Wert) bzw. Tief (**L**; bei positivem V-Wert) über Mitteleuropa. Die verbleibenden acht prototypischen Großwettertypen sind benannt nach der Anströmrichtung über Zentraleuropa und ergeben sich aus der Kombination der beiden Koeffizienten Z und M. So weist beispielsweise der Großwettertyp W (mit westlicher Anströmrichtung) eine perfekte Korrelation ($r=1,0$) zum zonalen Grundmuster ($Z=1,0$) und keine Korrelation ($r=0,0$) zum meridionalen Grundmuster ($M=0,0$) auf, d. h. **W** (1,0/0,0). Dementsprechend sind die übrigen Großwettertypen definiert: **NW** (0,7/-0,7), **N** (0,0/-1,0), **NE** (-0,7/-0,7), **E** (-1,0/0,0), **SE** (-0,7/0,7), **S** (0,0/1,0), **SW** (0,7/0,7). Die finale Zuordnung eines individuellen Feldes zu einem der prototypischen Großwettertypen (ausgenommen der Fall einer vorherigen Zuordnung zu **H** oder **L**) erfolgt auf Grundlage der geringsten euklidischen Distanz seiner Ausprägungen von Z und M zu einem der Prototypen.

Neben der an Hess und Brezowsky (1977) angelehnten Anzahl von zehn Großwettertypen ermöglicht die Methode eine Vielfalt an unterschiedlichen Klassenzahlen (8, 10, 11, 16, 18, 19, 24, 26, 27). Für die Anzahl von acht Großwettertypen werden lediglich die acht Hauptwindsektoren (N, NE, E, SE, S, SW, W, NW) verwendet, der V-Koeffizient und damit die Großwettertypen H und L werden nicht ausgewiesen. Durch die Hinzunahme eines indifferenten Zirkulationstypen (X) ergibt sich eine Variante mit elf Großwettertypen. Unterscheidet man die Hauptwindsektoren noch zwischen einer zyklonalen (z. B. Nz) und antizyklonalen Ausprägung (z. B. Na), ergeben sich 16 bis 19 Möglichkeiten. Fügt man zudem eine Vorticity-indifferente Ausprägung der Hauptwindrichtungen (z. B.

Nx) hinzu, ergibt sich die Möglichkeit von 24 bis 27 sehr fein differenzierten Zirkulationstypen.

Um einen einfachen Überblick über die untersuchten Großwettertypen zu gewährleisten und einen Vergleich zu den anderen Klassifikationsverfahren zu ermöglichen, wurde in der vorliegenden Arbeit vorwiegend auf die Modifikation mit acht Klassen zurückgegriffen. Für die Analysen zur dekadischen Vorhersagbarkeit wurde jedes individuelle Feld,

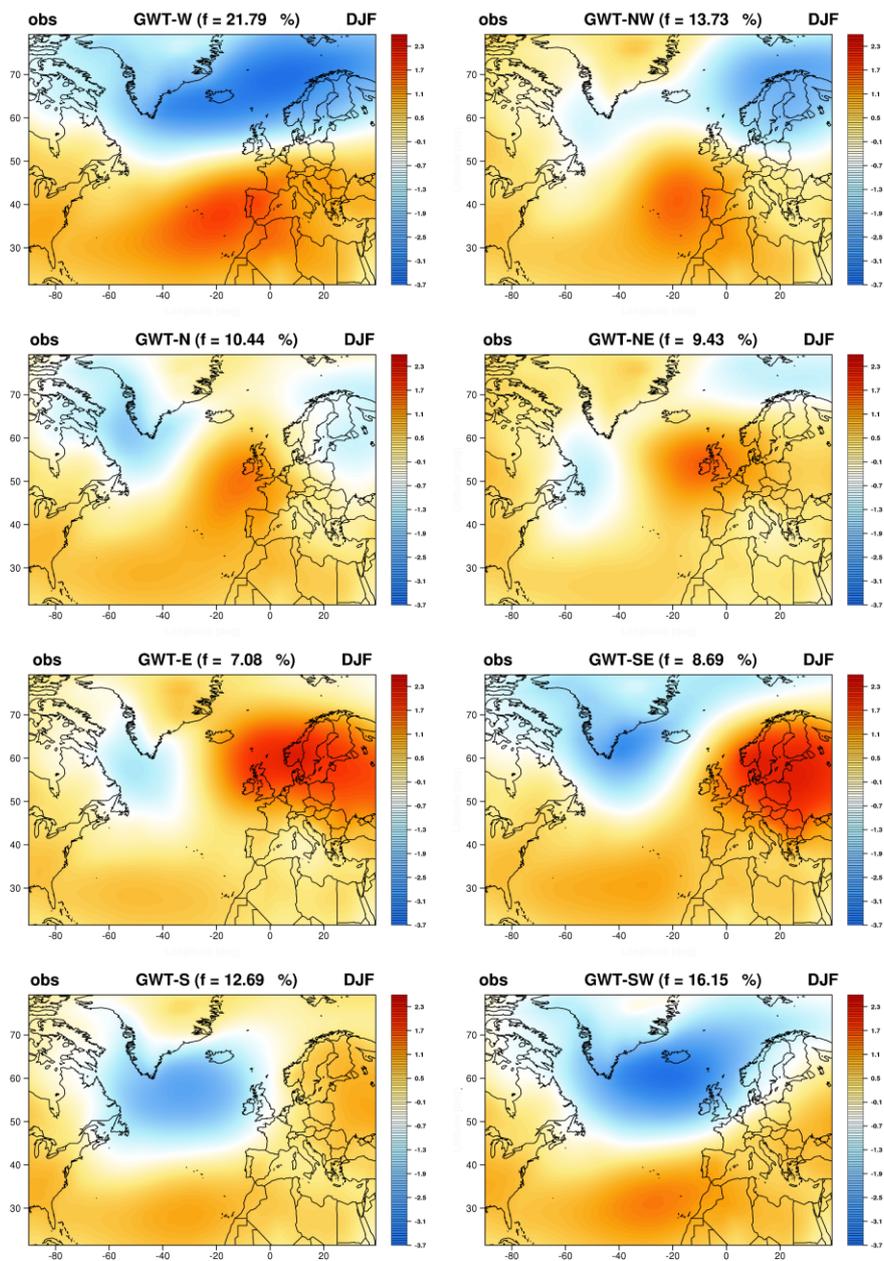


Abbildung 4.9: Zentroide der Zirkulationstypen (basierend auf GWT) in der nordatlantisch-europäischen Region im Winter (DJF) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Auftretshäufigkeit (Datengrundlage: SLP-Anomalien, ERAINT).

d. h. jeder einzelne Tag des Untersuchungszeitraums, einem Großwettertyp zugewiesen. Die Zuordnung der Einzelfelder der nordatlantisch-europäischen Region (90°W - 40°E , 20°N - 80°N) beruht auf dem bereits erwähnten zentraleuropäischen Gitterfeldausschnitt (10°W - 30°E , 40°N - 60°N). Die Mittelung aller zu einer Klasse zugewiesenen Einzelfelder, die jeweils mit dem Faktor 1 in die Berechnung eingehen, ergeben die klassenspezifischen Zentroide (von z. B. SLP oder ZG500), die nicht nur für den Gitterfeldausschnitt, sondern für das gesamte Datenfeld visualisiert werden können. Durch die Definition von Prototypen a priori ist es möglich, die Zuordnung für die Beobachtungsdaten sowie die Modelldaten separat vorzunehmen, die resultierenden Großwettertypen jedoch direkt miteinander (z. B. hinsichtlich ihrer Auftrittshäufigkeit oder ihrer Persistenz) zu vergleichen.

Abbildung 4.9 zeigt beispielhaft die Zentroide des Luftdrucks (SLP) für acht Zirkulationstypen im Winter (DJF) unter Angabe der jeweiligen relativen Auftrittshäufigkeit. Durch die Herangehensweise mit der Definition von idealtypischen Prototypen ergeben sich deutlich voneinander unterscheidbare Muster. Die Anströmungsrichtung aufgrund der Lage der Druckgebiete zeigt sich sowohl bei dem am häufigsten auftretenden Großwettertyp W (GWT-W) mit tiefem Luftdruck über Island und hohem über den Azoren als auch bei dem am seltensten auftretenden Großwettertyp E (GWT-E) mit einer ausgeprägten Antizyklone über dem nördlichen Europa und einer schwach ausgeprägten Zyklone im Mittelmeerraum.

4.1.2.2 SAN – SANDRA-Clusteranalyse

Neben den schwellen- (vgl. das vorherige Kapitel 4.1.2.1) und eigenwertbasierten (vgl. das folgende Kapitel 4.1.2.3) Klassifikationsverfahren zählen die auf Clusteranalysen beruhenden Verfahren zu den am häufigsten verwendeten objektiven Methoden zur Klassifizierung großräumiger Muster der atmosphärischen Zirkulation. Das Prinzip von Clusteranalysen besteht in der Zusammenlegung von Objekten (hier: einzelne Zeitpunkte) in einem von Variablen (hier: einzelne Gitterpunkte) aufgespannten Variablenraum auf Grundlage eines zu wählenden Ähnlichkeitskriteriums. Sie haben die Bildung von Clustern zum Ziel, die eine hohe interne Homogenität aufweisen, sich aber zugleich stark voneinander unterscheiden. Hinsichtlich des Vorgehens lassen sich grundsätzlich zwei verschiedene Ansätze unterscheiden: hierarchische Verfahren und nicht-hierarchische Verfahren der Clusteranalyse (Backhaus et al. 2003).

Die SANDRA-(*Simulated Annealing and Diversified Randomization*) Clusteranalyse (Philipp et al. 2007) lässt sich den nicht-hierarchischen Verfahren der Clusteranalyse zuordnen. Diese versuchen, dem Nachteil der fehlenden Flexibilität von Clustern, den hierarchische Verfahren dadurch aufweisen, dass eine Zuordnung eines Objekts zu einem Cluster nicht mehr rückgängig gemacht werden kann, entgegenzuwirken. Hierarchische Verfahren gehen von der feinsten Partition, d. h. jedes Objekt stellt im übertragenen Sinn ein eigenes Cluster dar, aus und fassen beim Fusionierungsprozess Cluster bis zu einer gewählten finalen Clusteranzahl in jedem Schritt aufgrund eines gewählten Fusionierungskriteriums (basierend auf einem Distanzmaß) zusammen (Bahrenberg, Giese & Nipper 2003). Nicht-hierarchische Verfahren hingegen gehen von einer bereits gegebenen Gruppierung der Objekte aus (*Anfangspartition*) und ordnen die einzelnen Elemente unter Zuhilfenahme eines bestimmten Austauschalgorithmus zwischen den einzelnen Gruppen [Clustern] so lange um, bis eine definierte Zielfunktion ein Optimum erreicht (Backhaus et al. 2003). Als Zielfunktion wird nach Philipp et al. (2007) häufig das Varianzkriterium WSS (*within sum of square of deviations*) als Maß der clusterinternen Unähnlichkeit verwendet:

$$WSS = \sum_{j=1}^k \sum_{i \in C_j} D(X_i, \bar{X}_j)^2, \quad (2)$$

wobei k die Anzahl der Cluster C , i die Objektnummer und \bar{X}_j den Clustermittelwert (Zentroid) angeben. D ist die euklidische Distanz zwischen den Objekten und ihren entsprechenden Cluster-Zentroiden (mit m als Anzahl der Parameter; hier: Gitterpunkte).

Die Aufgabe eines iterativen Optimierungsverfahrens besteht in der Minimierung der WSS. Die Optimierung der Zielfunktion ist bei geringer Objektzahl einfach, stellt aber bei großer Objektzahl ein Problem dar, das selbst hochleistungsfähige Rechner nicht bewältigen können. Für die Lösung des Problems wurden verschiedene Austauschalgorithmen entwickelt, von denen das *k-means*-Verfahren (Hartigan & Wong 1979) weit verbreitet ist (Jacobeit et al. 1998). Es besitzt allerdings den Nachteil, oftmals frühzeitig in lokale Optima (eine nur scheinbar beste Anordnung) zu konvergieren, ohne das globale Optimum (die beste Anordnung, die nur bei Kenntnis aller Kombinationsmöglichkeiten

bekannt ist) zu approximieren. Indem sie zwei Verfahren (*simulated annealing* & *diversified randomization*) verbindet, gelingt der SANDRA-Clusteranalyse (SAN) eine stärkere Annäherung des Klassifikationsergebnisses an das globale Optimum als dem gängigen *k-means*-Algorithmus. D. h. einzelne Objekte können mit einer bestimmten Wahrscheinlichkeit, die zu Beginn groß ist und – gesteuert durch einen „Abkühlungsfaktor“ – im Verlauf des Verfahrens langsam abnimmt, zwischen den Clustern wandern (*simulated annealing*), wobei in einzelnen Schritten des iterativen Optimierungsverfahrens – im Gegensatz zum *k-means*-Algorithmus – eine der Zielfunktion widersprechende, „falsche“ Zuordnung möglich ist. Dieses Verfahren wird mehrmals mit zufälligen Partitionen gestartet, wobei während jedem Durchlauf auch die Reihenfolge der Objekte und der Cluster zufällig geändert wird (*diversified randomization*). Aus den unterschiedlichen Endpartitionen wird schließlich die mit der am besten optimierten Zielfunktion ausgewählt.

Für die Untersuchungen der vorliegenden Arbeit wurde den SANDRA-Clusteranalysen jeweils eine s-modale, korrelationsbasierte PCA mit anschließender Rotation der Hauptkomponenten (nach Varimax-Kriterium) vorgeschaltet. Die PCA hat eine Datenreduktion zum Ziel, d. h. die Information einer großen Anzahl an Variablen (hier: Gitterpunkte) soll durch eine geringere Anzahl an Hauptkomponenten wiedergegeben werden, wobei 99 % der ursprünglichen Varianz erklärt werden müssen. Mit diesem Vorgehen konnte eine deutliche Verkürzung der Rechenzeit erzielt werden. Bei der Anwendung der SANDRA-Clusteranalysen auf die erhaltenen Hauptkomponenten wurde eine Abkühlungswahrscheinlichkeit von $p=0,95$ gewählt, d. h. 95 % aller Objekte können während der ersten Iteration zwischen den Clustern wandern, und das Verfahren jeweils 100 Mal mit zufälligen Partitionen durchgeführt. Zwar wird bei Philipp et al. (2007) die Variante mit $p=0,99$ bzw. $p=0,90$ und 1000 Läufen (der erste Lauf als Referenzlauf mit $p=0,99$, die 999 weiteren mit $p=0,90$) vorgeschlagen, allerdings wurde in der vorliegenden Arbeit aufgrund der Vielzahl an durchgeführten SANDRA-Clusteranalysen und im Hinblick auf eine mögliche Operationalisierung des Vorhersagesystems auf eine akzeptable Rechenzeit geachtet.

Die SANDRA-Clusteranalyse wurde für die Untersuchungen in der vorliegenden Arbeit jeweils auf die Beobachtungsdaten angewendet, um die prototypischen Zirkulationstypen zu erhalten. Die Zuordnung der Modelldaten zu den Prototypen erfolgte aufgrund der geringsten euklidischen Distanz. Zwar kommen durchaus auch andere Vorgehensweisen

in Betracht, wie z. B. die Anwendung der SANDRA-Clusteranalyse auf einen kombinierten Datensatz aus Beobachtungs- und Modelldaten bzw. die Anwendung der SANDRA-Clusteranalyse auf Beobachtungs- und (alle Ensemblemitglieder der) Modelldaten getrennt voneinander, allerdings wurde davon aufgrund eines sehr starken Anstiegs der Rechenzeiten (in beiden Fällen) bzw. einer Verkomplizierung der Zuweisung (im zweiten Fall, da unterschiedliche Zirkulationstypen resultieren, die sich nicht immer eindeutig

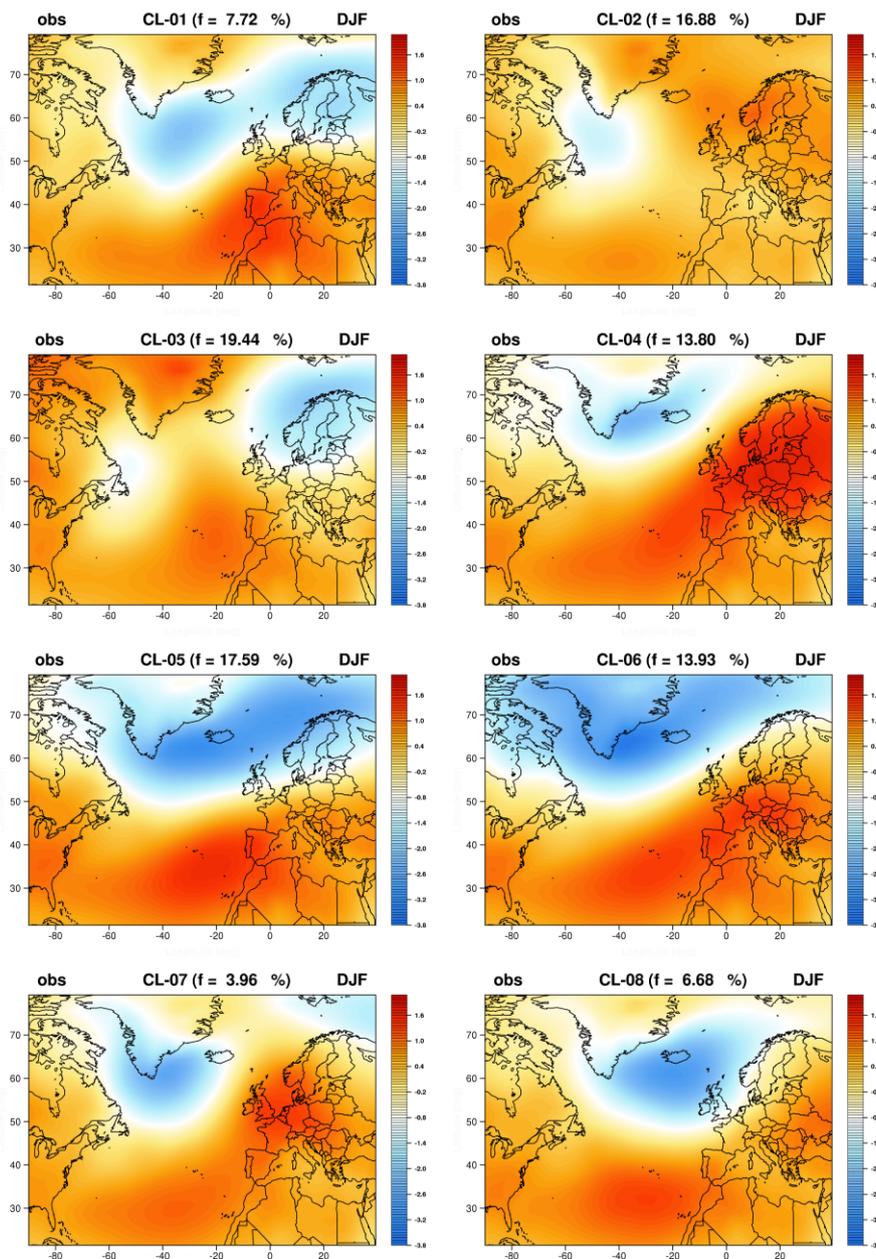


Abbildung 4.10: Zentroiden von 8 Zirkulationstypen (basierend auf SAN) in der nordatlantisch-europäischen Region im Winter (DJF) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Auftretshäufigkeit (Datengrundlage: SLP-Anomalien, ERAINT).

miteinander vergleichen lassen) abgesehen. Durch Mittelung aller zu einem Cluster zugeordneten Felder – jedes Feld geht mit einer Gewichtung von 1 ein – erhält man die klassenspezifischen Zentroide, die in Abbildung 4.10 wie in Anlehnung an die Großwettertypen nach Beck für den Winter (DJF) des Untersuchungszeitraums dargestellt sind. Abgebildet sind acht Zirkulationstypen des bodennahen Luftdrucks. Die Bestimmung der Klassenanzahl stellt bei der nicht-hierarchischen Clusteranalyse einen subjektiven Eingriff in das objektive, automatisierte Ermittlungsverfahren dar. Um für die Analysen zur dekadischen Vorhersagbarkeit einen einfachen Vergleich zu den acht Großwettertypen nach Beck (vgl. Abbildung 4.9) zu ermöglichen und einen Einfluss der Klassenanzahl auf das Untersuchungsergebnis, wie es bei Beck & Philipp (2010) nachgewiesen werden konnte, auszuschließen, wurde auch bei der SANDRA-Clusteranalyse standardmäßig auf die Klassenanzahl von acht zurückgegriffen.

Die acht Zirkulationstypen, für die in Abbildung 4.10 jeweils die prozentualen Auftretshäufigkeiten angegeben sind, weisen auf den ersten Blick eine ähnliche Drucksituation mit tiefem Druck im nördlichen und hohem Druck im südlichen Bereich auf. Bei genauer Betrachtung zeigen sich allerdings erkennbare Unterschiede hinsichtlich der Anströmrichtung und den damit verbundenen Witterungsbedingungen über Mitteleuropa. Beispielsweise zeigt Cluster 5 (CL-05) mit einer ausgeprägten Tiefdrucksituation über Island und hohem Druck über den Azoren eine ausgeprägte Westwindwetterlage über Mitteleuropa. Bei Cluster 4 (CL-04) hat sich der Hochdruckeinfluss ausgedehnt und verursacht einen antizyklonalen Einfluss über Mitteleuropa. Bei Cluster 3 (CL-03) ist demgegenüber eine Verlagerung des tiefen Drucks über Skandinavien zu erkennen, was in Mitteleuropa eine nordwestliche Anströmrichtung zur Folge hat.

4.1.2.3 TPCA – T-modale Hauptkomponentenanalyse

Wie bereits erwähnt wurde, ist die Methode der Hauptkomponentenanalyse ein in den Atmosphärenwissenschaften häufig angewendetes multivariates Verfahren (Wilks 2006). Auch zur Charakterisierung der atmosphärischen Zirkulation hat sie sich seit längerer Zeit etabliert und wurde nach den Recherchen von Richman (1981) erstmals von Fukuoka (1951) zu diesem Zweck eingesetzt. Mittlerweile besitzt sie ein weites Verbreitungsgebiet und es finden sich in der Literatur zahlreiche Anwendungsbeispiele (z. B. Huth 1993, Huth 1996, Bartzokas & Metaxas 1996, Compagnucci & Salles 1997, Jacobeit et al. 2003a, Jacobeit et al. 2003b). Richtungsweisend waren die Arbeiten von Huth, der zur

Klassifizierung der atmosphärischen Zirkulation eine t-modale Hauptkomponentenanalyse (TPCA) verwendet. Im Gegensatz zur s-modalen Variante, die in Kapitel 4.1.1 bereits rudimentär beschrieben wurde, wird die Eingangsdatenmatrix bei der TPCA um 90° gedreht, d. h. die Zeiteinheiten stellen die Variablen und die Gitterpunkte die Ereignisse dar. Während bei der Erstellung von Telekonnektionen die s-modale Hauptkomponentenanalyse (SPCA) Anwendung findet, die charakteristische Zeitreihen für räumliche Variationszentren repräsentiert, ist die Verwendung der t-modalen Variante mit der Dominanz von räumlichen Mustern zu bestimmten Zeitpunkten bei der Berechnung atmosphärischer Zirkulationsmuster geeignet (Compagnucci & Richman 2008, Jacobeit 2010). Als Ergebnis einer t-modalen PCA resultieren Zirkulationsmuster (Prototypen) mit variierender Ähnlichkeit zu Originalfeldern (ausgedrückt durch die Ladungen). Jacobeit (2010) bezeichnet diese treffenderweise mit dem Begriff „Dynamische Modi“.

Zur Evaluierung der dekadischen Vorhersagbarkeit der Dynamik der Atmosphäre können die Ladungen, die sich aus den t-modalen Hauptkomponenten ergeben, für die Auswertungen zur dekadischen Vorhersagbarkeit herangezogen werden. Allerdings ergeben sich bei diesem Ansatz Probleme für das vorliegende Forschungsvorhaben. Zum einen lassen sich typinterne Charakteristika, wie z. B. die Intensität des Dynamischen Modus als Ausdruck des maximalen Luftdruckgegensatzes, in ihrem zeitlichen Verlauf nicht darstellen, wodurch Vergleiche zu den anderen Klassifikationsverfahren nicht möglich sind. Außerdem ergeben sich dadurch, dass die t-modale PCA getrennt für Beobachtungs- und Modelldaten durchgeführt werden muss, Schwierigkeiten: Zum einen ist eine Zuweisung der erhaltenen Dynamischen Modi zu den beobachteten aufgrund eines bestimmten Kriteriums notwendig, die sich v. a. bei den nicht dominanten Dynamischen Modi, d. h. den Modi mit nur einem geringen Varianzerklärungsanteil, schwierig gestaltet. Zum anderen resultiert, da die TPCA getrennt auf Beobachtungs- und die Modelldaten mit ihren unterschiedlichen Ensemblemitgliedern angewendet wird, eine enorme Rechenzeit, die bei häufiger Anwendung des Verfahrens (z. B. zur Untersuchung der Vorhersageleistung der verschiedenen Modellgenerationen in verschiedenen Jahreszeiten) zu Buche schlägt und sich bei einer möglichen Operationalisierung problematisch gestaltet.

Aus den genannten Gründen wurde auf das Konzept der Zirkulationstypen in Anlehnungen an die bisher erläuterten Klassifikationsverfahren zurückgegriffen. Die resultierenden Zirkulationsmuster (*scores*) einer TPCA, die auf die Beobachtungsdaten angewendet

wurde, dienen dabei als Prototypen, denen die individuellen Einzelfelder (SLP oder ZG) von Beobachtungs- und Modelldaten getrennt voneinander zugeordnet wurden. Die Zuordnung zu den Prototypen erfolgt hierbei auf Grundlage der führenden (maximalen) Ladung, die als Korrelation zwischen den individuellen Feldern der Ausgangsdatenmatrix und den Zirkulationsmustern berechnet werden können. Aus der Mittelung aller zu einem Prototyp zugeordneten Felder können – vergleichbar zu dem schwellenwert- (s. Kapitel 4.1.2.1) und dem clusteranalysebasierten Verfahren (s. Kapitel 4.1.2.2) – die klassenspezifischen Zentroide berechnet werden. Der Vorteil der TPCA gegenüber den anderen Klassifikationsverfahren besteht darin, dass jeder Tag nicht mit der Gewichtung von 1 in die Analyse einfließt, sondern mit dem Wert der führenden Ladung als Ausdruck der Ähnlichkeit zu den Prototypen. Abschließend wurden – in Analogie zu den Auftrittshäufigkeiten bei den bisher geschilderten Klassifikationsverfahren – die führenden Ladungen für die einzelnen Zirkulationstypen aufsummiert, um Aussagen über die zirkulationsdynamische Relevanz jeder Klasse in zeitlicher Hinsicht zu ermöglichen. Die aufsummierten Ladungen werden im Folgenden ebenfalls als „Auftrittshäufigkeiten“ wie bei den anderen Klassifikationsverfahren bezeichnet, obwohl es sich inhaltlich nicht um dasselbe handelt.

Abbildung 4.11 zeigt die Zirkulationstypen für den Winter (DJF) des Untersuchungsgebiets. Die relativen Auftrittshäufigkeiten spiegeln die Tatsache wider, dass die Hauptkomponenten, denen die Einzeltage zugeordnet wurden, ihrem Varianzklärungsanteil nach angeordnet sind, d. h. die erste PC erklärt am meisten Varianz und ihr werden in der Regel auch am meisten Einzeltage zugeordnet. Demzufolge dominiert PC-01, deren Zirkulationsmuster sich als eine klassische zonale Westwindwetterlage über Mitteleuropa interpretieren lässt, mit einer relativen Auftrittshäufigkeit von 46,17 %. Die nachgeordneten Zirkulationstypen, die mit einer relativen Auftrittshäufigkeit von knapp über 5 % (PC-04 mit 5,44 % und PC-05 mit 6,35 %) bzw. unter 5 % (PC-06 mit 1,46 %, PC-07 mit 0,93 % und PC-08 mit 0,66 %) eine viel geringere Relevanz besitzen, deuten allesamt – lediglich mit Ausnahme von PC-04 – eine Strömungssituation mit einem ausgeprägten

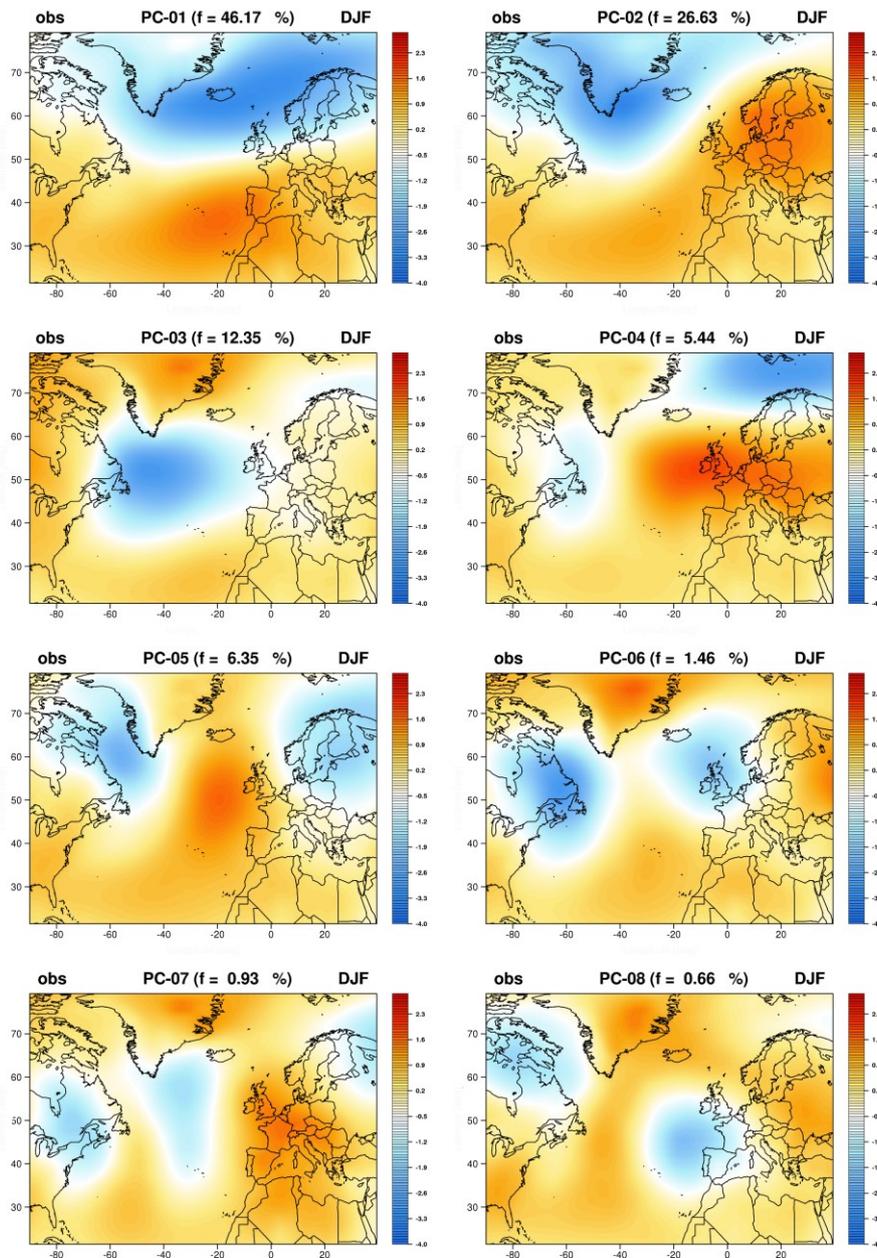


Abbildung 4.11: Zentroide der Zirkulationstypen (basierend auf TPCA) in der nordatlantisch-europäischen Region im Winter (DJF) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Auftretshäufigkeit (Datengrundlage: SLP-Anomalien, ERAINT).

meridionalen Charakter an. Betrachtet man alle abgebildeten Zirkulationstypen im Vergleich, lässt sich ein klar erkennbarer Vorteil des methodischen Vorgehens erkennen. Es offenbaren sich infolge der Orthogonalität der zugrundeliegenden Hauptkomponenten deutlich stärker voneinander unterscheidbare Muster als z. B. es bei den Zirkulationstypen als Resultate der SANDRA-Clusteranalyse (vgl. Abbildung 4.10) der Fall ist.

4.2 Evaluation

4.2.1 Hindcasts

Alle Analysen zur Evaluation der dekadischen Vorhersagen, auf denen die Ergebnisse der vorliegenden Arbeit beruhen, wurden mit sog. Hindcasts durchgeführt. Das bedeutet, dass mithilfe des Klimamodells retrospektiv dekadische Vorhersagen für jedes Jahr des Kernzeitraums 1961-2013 (*baseline0*: 1961-2012) erstellt wurden. Abbildung 4.12 veranschaulicht anhand zweier ausgewählter Jahre dieses Zeitraums das Prinzip der retrospektiven Hindcasts und ihre Auswertung. Das Beispiel „decadal_1978“ Abbildung 4.12a

(a) decadal_1978										(b) decadal_1979										...
1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	..
1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	...

Abbildung 4.12: Das Prinzip der Hindcasts für zwei Beispiele (links: decadal_1978; rechts: decadal:1979) mit Angabe der Leadjahre 1-10.

zeigt die dekadische Vorhersage für die Jahre 1979-1988. Benannt ist der Zeitraum mit „decadal_1978“, weil der Assimilationslauf zur Gewinnung der Initialisierungsbedingungen 1978 gestartet wurde, wobei die einzelnen Ensemblemitglieder mit einem zeitlichen Versatz von einem Tag gewonnen werden (1. Januar 1978: erstes Ensemblemitglied, 2. Januar 1978: zweites Ensemblemitglied, 3. Januar 1978: drittes Ensemblemitglied etc.). Mit den aus dem Assimilationslauf resultierenden Feldern wurde schließlich das Klimamodell initialisiert, um dekadische Vorhersagen zu machen. In diesem Fall für den Zeitraum 1979-1988. In Abbildung 4.12b ist die dekadische Vorhersage für den Zeitraum 1980-1989 dargestellt, welche dem beschriebenen Vorgehen folgend mit „decadal_1979“ bezeichnet ist.

Auf dem Forschungsgebiet der dekadischen Vorhersage hat sich etabliert, dass nicht einzelne dekadische Vorhersagen eines bestimmten Startjahres analysiert werden, sondern sog. Leadjahre betrachtet werden. Das bedeutet, dass aus allen dekadischen Vorhersagen eines Untersuchungszeitraums beispielsweise jeweils das erste Jahr der Vorhersage herausgegriffen wird, was als „Leadjahr 1“ bzw. „erstes Leadjahr“ bezeichnet wird (s. rot markierte Jahre in Abbildung 4.12). Demensprechend wird für „Leadjahr 2“, „Leadjahr

3“ etc. verfahren. Führt man saisonale Analysen wie in der vorliegenden Arbeit durch, besteht das Frühjahr (MAM) des ersten Leadjahres jeweils aus dem dritten (März), dem vierten (April) und dem fünften Monat (Mai) nach der Initialisierung. Für den Winter (DJF) werden für das erste Leadjahr der zwölfte (Dezember), der 13. (Januar) und der 14. Monat (Februar) nach Initialisierung herangezogen, d. h. es werden die ersten zwei Monate aus dem zweiten Leadjahr verwendet. Da nach diesem Prinzip für das zehnte Leadjahr keine Berechnung des Winters erfolgen kann (Leadjahr 11 existiert nicht), werden Untersuchungen auf dem Gebiet der dekadischen Klimavorhersage i. d. R. für die Leadjahre 1 bis 9 durchgeführt.

Diesem Prinzip folgend lassen sich für die einzelnen Leadjahre Zeitreihen (für Auswertungen zur Repräsentation von Telekonnektionsindizes; s. Kapitel 5) bzw. kumulierte Auftrittshäufigkeiten (für Auswertungen zur Repräsentation von Zirkulationstypen; s. Kapitel 6) für den gesamten Untersuchungszeitraum erstellen. In der vorliegenden Arbeit bilden die Jahre 1979-2011 standardmäßig den bereits in den vorherigen Kapiteln mehrfach erwähnten Untersuchungszeitraum. Zwar liegen die dekadischen Vorhersagen (vgl. Kapitel 3.1) schon bereits ab 1961 für alle Modellgenerationen vor, allerdings sind die verwendeten Beobachtungsdaten (vgl. Kapitel 3.2) erst ab 1979 verfügbar. Somit können die Hindcast-Analysen zur dekadischen Vorhersagbarkeit erst ab 1979 verwirklicht werden, d. h. die dekadische Vorhersage entstammt „decadal_1978“ für das erste Leadjahr. Zur Durchführung der Hindcast-Analysen für alle Leadjahre werden jedoch auch dekadische Vorhersagen mit einem früheren Startpunkt als 1979 benötigt. Z. B. wird die dekadische Vorhersage für das zweite Leadjahr des Jahres 1979 aus „decadal_1977“ entnommen, die für das neunte Leadjahr aus „decadal_1970“. Lediglich für den Winter (DJF) beginnt der Untersuchungszeitraum nicht 1979, sondern 1980 (siehe z. B. Abbildung 3.2), da für den Winter 1979 (setzt sich aus Dezember 1978 sowie Januar und Februar 1979 zusammen) der Dezember 1978 in den Beobachtungsdaten nicht verfügbar ist. Zudem wurde der Untersuchungszeitraum in der vorliegenden Arbeit auf das Jahr 2011 begrenzt. Dieses Jahr wurde gewählt, da für die meisten Reanalysen die verwendeten Variablen (SLP, ZG) in den verschiedenen zeitlichen Auflösungen (*monatlich*, *6-stündlich*) zur Verfügung stehen. Lediglich bei CFSR reichen die Daten bis 2010, weshalb bei Analysen, bei denen CFSR-Reanalysedaten verwendet wurden, der Untersuchungszeitraum bis zum Jahr 2010 begrenzt ist.

4.2.2 Skillmetriken

Vergleicht man eine Vorhersage mit beobachteten Werten, stellt sich immer die Frage nach geeigneten Möglichkeiten zur Beurteilung der Qualität der Vorhersage. Ein Blick in die Literatur offenbart für die Evaluation dekadischer Vorhersagen „a dizzying array of different climate metrics“ (Hurrell et al. 2010, S. 8). Während Smith et al. (2007) für die Evaluation ihrer dekadischen Vorhersagen den RMSE (Root-Mean-Square Error) verwenden, greifen Keenlyside et al. (2008) auf Korrelationen für die Untersuchung von Zeitreihen zurück, womit auch die am häufigsten verwendeten Skillmetriken erwähnt wären. Daneben sind in der Literatur Werkzeuge der Evaluation zu finden, die auf dem Gebiet der dekadischen Vorhersage eher selten Verwendung finden, wie z. B. der CRPSS (*Continuous Ranked Probability Skill Score*) bei Kadow et al. (2015) oder der RPC (*Ratio of Predictable Components*) bei Eade et al. (2014). Allen gemein ist jedoch, dass sie eine Vorhersage mit einer Beobachtung vergleichen.

Grundsätzlich kann es unterschiedliche Ansichten geben, was eine gute Vorhersage ausmacht (Wilks 2006). Der lineare Korrelationskoeffizient (s. Kapitel 4.2.2.1) zwischen Anomalien der Vorhersage und der Beobachtung vermag ein sehr nützliches und leicht zu interpretierendes Verifikationsmaß sein, allerdings kann ein hoher Wert auf dekadischer Zeitskala bereits durch die Vorhersage des Erwärmungstrends in Folge der Erhöhung der Treibhausgase erzielt werden (Meehl et al. 2009). Murphy et al. (2010) schlagen deshalb vor, mehrere Skillmetriken zu verwenden, um die Vorhersageleistung eines (dekadischen) Modellsystems abschätzen zu können. Insbesondere bei dekadischen Vorhersagen, die von Natur aus probabilistisch sind, d. h. mehrere mögliche Realisierungen eines Modells (Ensemblemitglieder) beinhalten, ist dieses Vorgehen notwendig. Deswegen sollten bei der Evaluation sowohl deterministische Maße zum Abschätzen der mittleren Vorhersage als auch probabilistische Maße für die Abschätzung der Bandbreite der Vorhersage, des sog. Spreads, als Ausdruck der Unsicherheit des Modells in Betracht gezogen werden. In der vorliegenden Arbeit kommen – neben dem linearen Korrelationskoeffizienten – mit dem MSSS (Mean Squared Error Skill Score; s. Kapitel 4.2.2.2) und dem RPSS (Ranked Probability Skill Score; s. Kapitel 4.2.2.3) sowohl ein deterministisches als auch ein probabilistisches Gütemaß zum Einsatz.

Alle Skillmetriken werden angewendet, um die Qualität einer bestimmten dekadischen Vorhersage gegenüber einer Referenz zu beurteilen. Als Referenz wird häufig die Klimatologie oder eine zufällige Vorhersage eingesetzt. In der vorliegenden Arbeit bietet es sich allerdings an, die verschiedenen Modellgenerationen (mit den unterschiedlichen Initialisierungstechniken) direkt miteinander zu vergleichen, also z. B. zu untersuchen, inwiefern **b1** (als zu untersuchende Vorhersage) Verbesserungen gegenüber **b0** (als Referenz) aufweist.

4.2.2.1 Korrelationskoeffizient (nach Pearson)

Der lineare Korrelationskoeffizient (nach Pearson) wird in der vorliegenden Arbeit für die Analyse der Telekonnektionsindizes verwendet. Der Zusammenhang der Zeitreihe von über alle Ensemblemitglieder gemittelter dekadischer Vorhersage (*Forecast*) und Beobachtung (*Observation*) wird folgendermaßen berechnet:

$$Corr_F = \frac{s_{FO}}{s_F * s_O} = \frac{\sum_{j=1}^n (F_j - \bar{F})(O_j - \bar{O})}{\sqrt{\sum_{j=1}^n (F_j - \bar{F})^2 * \sum_{j=1}^n (O_j - \bar{O})^2}} \quad (3)$$

Der Korrelationskoeffizient ($Corr_F$) berechnet sich aus dem Verhältnis von Kovarianz (s_{FO}) und den Standardabweichungen der Vorhersage (s_F ; *Forecast*) und der Beobachtung (s_O ; *Observation*) und untersucht die gemeinsame Varianz der beiden Zeitreihen. Zur Evaluation der dekadischen Vorhersage wird zudem der Korrelationskoeffizient für eine Referenz ($Corr_R$) berechnet (wie bereits erwähnt z. B. **b1** als „Vorhersage“, **b0** als „Referenz“). Anhand der Differenz der beiden Werte ($Corr_F - Corr_R$) lässt sich ermitteln, ob die Vorhersage eine Verbesserung (bei positiven Werten) oder eine Verschlechterung (bei negativen Werten) gegenüber der Referenz zeigt (vgl. die Anwendungsbeispiele in Goddard et al. 2013).

4.2.2.2 Mean Squared Error Skill Score (MSSS)

Der MSSS ist ein deterministisches Gütemaß und kann eine Antwort auf die Frage geben, ob eine Vorhersage eine höhere Genauigkeit (*Accuracy*; vgl. Wilks 2006) zeigt als eine Referenz (also z. B.: Zeigen initialisierte dekadische Vorhersagen eine höhere Genauigkeit als uninitialisierte?). Der MSSS basiert auf der Berechnung des Mean Squared Error (MSE),

$$MSE_F = \frac{1}{n} \sum_{j=1}^n (F_j - O_j)^2, \quad (4)$$

der für die über alle Ensemblemitglieder gemittelte Vorhersage F_j bzw. Referenz R_j jeweils im Vergleich mit der Beobachtung O_j berechnet wird (vgl. Goddard et al. 2013).

Der MSSS ist als Verhältnis zwischen Vorhersage (MSE_F) und Referenz (MSE_R) für eine bestimmte Leadzeit τ folgendermaßen definiert:

$$MSSS_\tau = 1 - \frac{MSE_{F,\tau}}{MSE_{R,\tau}}, \quad (5)$$

Wie aus (5) hervorgeht, zeigen Werte über 0 eine bessere Performanz der Vorhersage (d. h. $MSE_F < MSE_R$), Werte unter 0 eine schlechtere gegenüber der Referenz (d. h. $MSE_F > MSE_R$) an.

4.2.2.3 Ranked Probability Skill Score (RPSS)

Der RPSS ist im Gegensatz zum MSSS ein probabilistisches Gütemaß. Analog zum MSSS stellt der Ranked Probability Score (RPS) die Grundlage des RPSS dar (vgl. Wilks 2006), allerdings gehen zur Berechnung – im Gegensatz zum MSE – alle Ensemblemitglieder und nicht das Ensemblemittel ein:

$$RPS_{F,\tau,M} = \frac{1}{I} \sum_{i=1}^I \sum_{k=1}^K (F_{\tau,i,k} - O_{t(\tau,i),k})^2 - \frac{M - m_i}{M(m_i - 1)} F_{\tau,i,k} (1 - F_{\tau,i,k}) \quad (6)$$

τ	= Leadzeit	F	= kumulat. Wahrsch. (Vorhersage)
i	= Initialisierungszeitpunkt	O	= kumulat. Wahrsch. (Beobachtung)
k	= Ereignisklasse		
M	= fiktive Ensemblegröße		
m_i	= existierende Ensemblegröße		
$t(\tau,i)$	= Wert der Beobachtung zum Zeitpunkt τ,i		

Die Berechnung basiert auf dem von Ferro et al. (2008) entwickelten Verfahren. $F_{\tau,i,k}$ bezeichnet die kumulative Wahrscheinlichkeit der Klasse k (hier: mit $K = 3$: mittlere, hohe, niedrige Werte) für eine bestimmte Leadzeit τ einer bestimmten Initialisierung i der

Vorhersage. Von diesem Wert wird $O_{t(\tau,i),k}$ subtrahiert, die kumulative Wahrscheinlichkeit der Klasse k zum entsprechenden Zeitpunkt $t(\tau,i)$ der Beobachtung. $O_{t(\tau,i),k}$ nimmt den Wert 1 an, wenn Klasse k beobachtet wird, ansonsten gilt $O_{t(\tau,i),k} = 0$ (= Heaviside-Funktion). Explizit sei auf den zweiten Term in Formel (6) hingewiesen, der dem Problem begegnet, dass der RPS für eine endliche Anzahl an Ensemblemitgliedern verzerrt ist. Dadurch ist es möglich, Modellversionen mit einer unterschiedlichen Anzahl an Ensemblemitgliedern direkt zu vergleichen, was im Fall des MSSS nicht möglich ist.

Äquivalent zum MSSS erfolgt die Berechnung des RPSS als Verhältnis zwischen Vorhersage (RPS_F) und Referenz (RPS_R) zu einer bestimmten Leadzeit τ :

$$RPSS_\tau = 1 - \frac{RPS_{F,\tau}}{RPS_{R,\tau}} \quad (7)$$

Wiederum zeigen Werte über 0 Verbesserungen (Werte unter 0 Verschlechterungen) der untersuchten Vorhersage gegenüber der Referenz an.

5 Ergebnisse:

Telekonnektionsindizes

In Kapitel 5 werden die Ergebnisse hinsichtlich der dekadischen Vorhersagbarkeit von Telekonnektionsindizes unter verschiedenen Gesichtspunkten, d. h. Einflüsse auf die Vorhersageleistung, dargestellt. Zunächst wird die Vorhersageleistung der uninitialisierten, historischen Läufe analysiert (Kapitel 5.1), bevor zwei zentrale Fragen dekadischer Forschung untersucht werden: Welchen Einfluss auf die dekadische Vorhersageleistung von Telekonnektionsindizes hat die Initialisierung des Klimamodells (vgl. Kapitel 5.2) bzw. welchen Einfluss hat die Ensemblegröße (vgl. Kapitel 5.3)?

Nach Betrachtung der einzelnen Einflussfaktoren auf die Vorhersageleistung des dekadischen Vorhersagesystems wird in Kapitel 5.4 eine überblickende Betrachtung gewährt, in der die Ergebnisse der einzelnen beeinflussenden Faktoren zu einem Gesamtbild zusammengefügt und mehrere Fragen (z. B. Welche Initialisierungstechnik liefert die besten Resultate?) aufgeworfen und abschließend beantwortet werden. In Kapitel 5.5 erfolgt schließlich eine knappe Zusammenfassung der Ergebnisse zu der dekadischen Vorhersagbarkeit von Telekonnektionsindizes.

5.1 Die uninitialisierten Läufe

Die Darstellung der uninitialisierten, historischen Läufe (vgl. Kapitel 3.1.2) soll einen grundlegenden Einblick in die Vorhersagbarkeit interannueller Variabilität geben, um mögliche Verbesserungen beurteilen zu können, die durch die Initialisierung bzw. die Vergrößerung der Ensemblegröße des dekadischen Vorhersagemodells erzielt werden. Die Telekonnektionsindizes werden in diesem Kapitel bei saisonaler Aggregation der Daten für den Untersuchungszeitraum 1979-2011 präsentiert. Es wird jeweils lediglich die Jahreszeit angesprochen, in der das entsprechende Telekonnektionsmuster in den Beobachtungsdaten den größten Varianzerklärungsanteil (= prozentualer Anteil der erklärten Varianz an der Gesamtvarianz) besitzt. Anzumerken ist an dieser Stelle, dass das Frühjahr (MAM) durch den dritten, vierten und fünften Monat, der Sommer (JJA) durch den sechsten, siebten und achten Monat und der Herbst (SON) durch den neunten, zehnten und elften Monat berechnet wird, während der Winter (DJF) anhand des zwölften sowie des ersten und zweiten Monat des Folgejahres abgebildet und diesem zugerechnet wird, weshalb der Untersuchungszeitraum im Winter im Folgenden 1980, in allen anderen Jahreszeiten allerdings bereits 1979 beginnt.

Die Vorhersagen der historischen Modellläufe (**hist**; blaue Farbe in den folgenden Grafiken) basieren jeweils auf drei Ensemblemitgliedern, für die beobachteten Werte (**obs**; grau) wurde auf die Reanalysen von ERA-Interim (ERAINT) zurückgegriffen.

5.1.1 Nordatlantische Oszillation (nach Hurrell)

Erfolgt die Berechnung des NAO-Index‘ auf Grundlage der Methode nach Hurrell (NAO-PC; vgl. Kapitel 4.1.1.1, Tabelle 4.1) basierend auf einem nordatlantisch-europäischen Untersuchungsgebiet, weist der Winter (DJF) mit einem Wert von 55,75 % den höchsten Varianzerklärungsanteil aller Jahreszeiten auf, d. h. die erste Hauptkomponente, die für die Reanalyse ERAINT durchgeführt wurde, erklärt 55,75 % der Originalvarianz. Die zeitliche Ausprägung des Telekonnektionsmusters ist in Abbildung 5.1 dargestellt. Für den Untersuchungszeitraum zeigt die Grafik lediglich eine geringe Übereinstimmung im Verlauf zwischen den beobachteten Werten (**obs**) und den Werten der über die drei En-

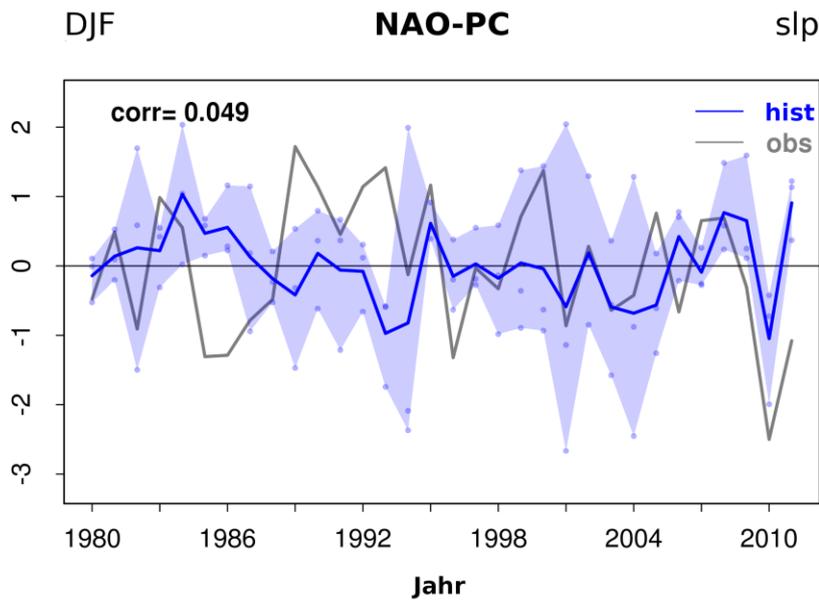


Abbildung 5.1: (Standardisierter) **Index der Nordatlantischen Oszillation (NAO-PC)** im Winter (DJF) des Untersuchungszeitraums, dargestellt für die historischen, uninitialisierten Modellläufe (**hist**; s. Erklärungen im Text) und die beobachteten Werte (**obs**). Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. Der Korrelationskoeffizient drückt den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.

semblemitglieder gemittelten historischen Modellläufe (**hist**), die sich in einem Korrelationskoeffizienten von $r=0,049$ manifestiert. Häufig befinden sich die Werte der Beobachtungszeitreihe außerhalb des hellblau markierten Bereichs, der durch die unterschiedlichen Ausprägungen der drei Ensemblemitglieder aufgeworfen wird. Insbesondere extreme Werte des *NAO-Index*‘ vermögen die historischen Modellläufe nur unzureichend wiederzugeben, was z. B. für die niedrigen Werte Mitte der 1980er-Jahre sowie die stark positiven Werte am Ende der 1980er-Jahre bzw. zu Beginn der 1990er-Jahre beobachtet werden kann. Ende der 1990er-Jahre und zu Beginn der 2000er-Jahre fällt in der Vorhersage zudem die große Heterogenität der einzelnen Ensemblemitglieder auf, was in einer sehr breiten, hellblau hinterlegten Fläche resultiert. An dieser Stelle sei aber auf das Jahr 2010 verwiesen, in dem die uninitialisierten Modellläufe zwar nicht die stark negative Ausprägung der Beobachtung wiederzugeben vermögen, ihnen allerdings in großer Übereinstimmung aller Ensemblemitglieder eine gute Annäherung gelingt.

Ein vergleichbares Ergebnis ist für den *NAO-Index* zu beobachten, dessen Berechnung auf Gitterpunkten (NAO-GP, vgl. Kapitel 4.1.1.1, Tabelle 4.1) basiert (Abbildung 5.2).

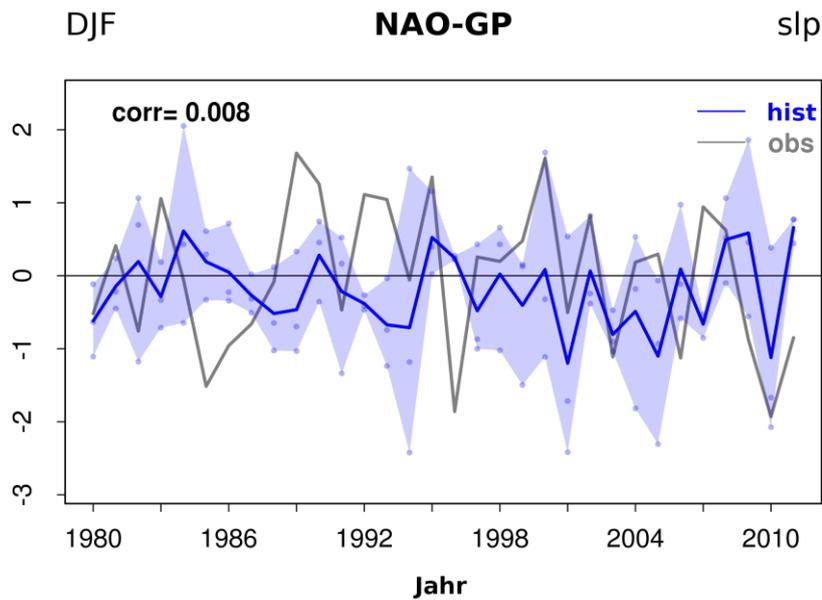


Abbildung 5.2: Wie Abbildung 5.1, aber für NAO-GP.

Mit einem Korrelationskoeffizienten von $r=0,008$ zeigt sich auch hier für die Wintermonate (DJF) ein sehr geringer Wert nahe 0 zwischen der Beobachtung und den Werten der über die drei Ensemblemitglieder gemittelten Vorhersage des uninitialisierten Modells. Häufig liegen die Werte wiederum außerhalb des hellblau markierten Bereichs, d. h. die Beobachtung wird durch keine Realisation des Modells repräsentiert.

5.1.2 CPC-Telekonnektionsmuster

Für die CPC-Telekonnektionsmuster (vgl. Kapitel 4.1.1.2, Tabelle 4.1) wurde jeweils die Jahreszeit herausgegriffen, für die der höchste Varianzerklärungsanteil ermittelt werden konnte. Da sich das Untersuchungsgebiet über die gesamte Nordhemisphäre erstreckt und in jeder Jahreszeit zehn Hauptkomponenten extrahiert werden, weisen die Varianzerklärungsanteile deutlich geringere Werte auf als z. B. bei der Berechnung des NAO-Index‘ auf Grundlage der Methode nach Hurrell. Dementsprechend zeigt das Telekonnektionsmuster der Nordatlantischen Oszillation nach der CPC-Berechnungsmethode ebenfalls im Winter (DJF) den höchsten Varianzerklärungsanteil, allerdings weist dieser lediglich einen Wert von 9,60 % auf (vgl. Tabelle 5.1). Ebenfalls im Winter offenbaren die Telekonnektionsmuster *EA* und *TNH* mit Werten von 9,92 % bzw. 8,40 % die höchsten Varianzerklärungsanteile, wobei hinzugefügt werden muss, dass *TNH* lediglich im Winter stabil zu beobachten ist. Im Frühjahr (MAM) hingegen haben die Telekonnektionsmuster

EAWR mit 7,69 %, *WP* mit 7,42 % und *POL* mit 18,86 % ihre jeweiligen höchsten Werte hinsichtlich des Varianzerklärungsanteils, wobei *POL* damit den insgesamt höchsten Wert offenbart. Mit 16,88 % weist außerdem das Telekonnektionsmuster *PNA* einen zweistelligen Wert auf, der im Sommer (JJA) erreicht wird, wohingegen *EPNP* und *SCAND* im Herbst (SON) mit 7,21 % bzw. 9,29 % am meisten Gesamtvarianz erklären.

Tabelle 5.1: Varianzerklärungsanteile der CPC-Telekonnektionsmuster für die verschiedenen Jahreszeiten (MAM, JJA, SON, DJF). Fett markierte Werte heben jeweils die Jahreszeit mit dem höchsten Varianzerklärungsanteil hervor.

	MAM	JJA	SON	DJF
EA	5,32 %	8,55 %	4,87 %	9,92 %
EAWR	7,69 %	4,90 %	5,07 %	4,24 %
EPNP	6,51 %	5,21 %	7,21 %	-
NAO	8,34 %	6,80 %	7,60 %	9,60 %
PNA	7,53 %	16,88 %	6,49 %	4,24 %
POL	18,86 %	7,59 %	15,19 %	8,77 %
SCAND	7,69 %	4,93 %	9,29 %	4,51 %
TNH	-	-	-	8,40 %
WP	7,42 %	5,21 %	6,21 %	4,51 %

In Abbildung 5.3 ist die Vorhersageleistung – wiederum gekennzeichnet anhand des Korrelationskoeffizienten zwischen beobachteten Werten (**obs**) und den gemittelten uninitialisierten, historischen Modellläufen (**hist**) – für die verschiedenen CPC-Telekonnektionsmuster jeweils für die Jahreszeit mit dem höchsten Varianzerklärungsanteil dargestellt, wie sie aus Tabelle 5.1 hervorgeht. Insgesamt kann im Untersuchungszeitraum für alle Indizes der Telekonnektionsmuster nur ein geringer Zusammenhang zwischen der Beobachtung und der uninitialisierten Vorhersage beobachtet werden, was in geringen Korrelationskoeffizienten zum Ausdruck kommt. Die Werte bewegen sich entweder nahe 0 und weisen somit keinen Zusammenhang auf (z. B. *PNA* mit $r=-0,017$, *EA* mit $r=0,062$ oder *SCAND* mit $r=-0,074$) oder es ist ein leicht positiver (*WP* mit $r=0,148$) bzw. negativer Zusammenhang (z. B. *EPNP* oder *TNH* mit $r=-0,184$ bzw. $r=-0,249$) zu beobachten.

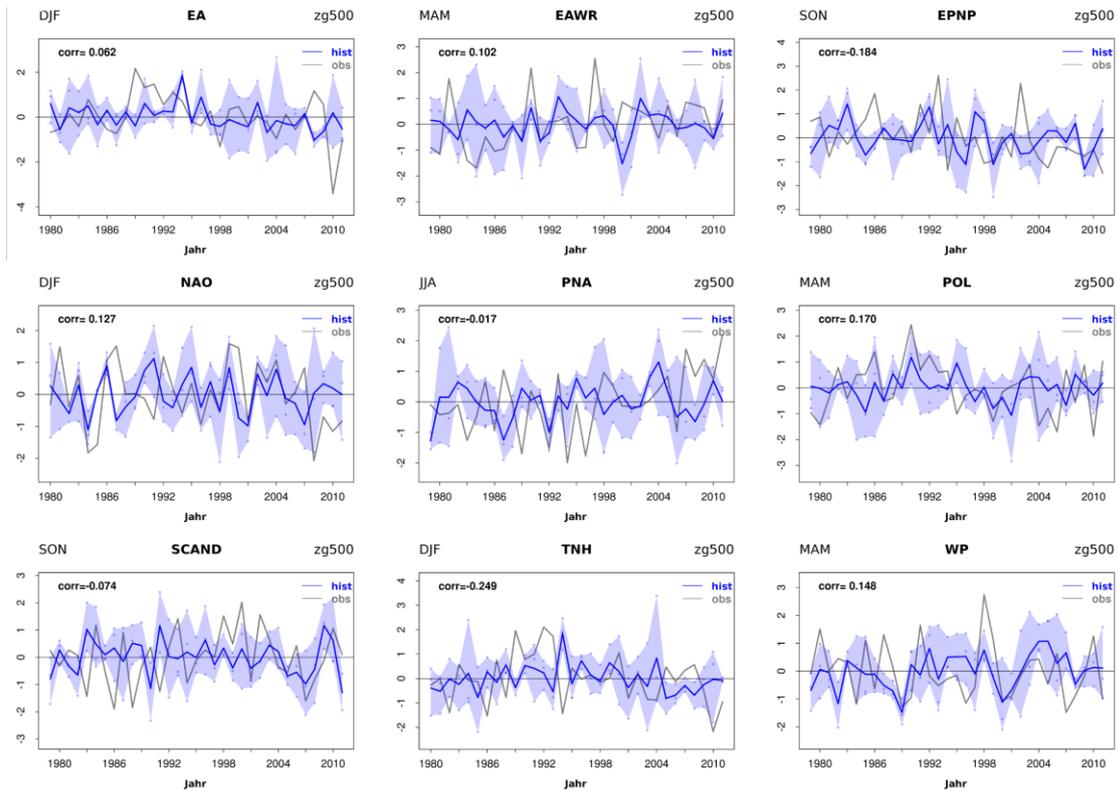


Abbildung 5.3: (Standardisierte) **Indizes der CPC-Telekonnectionsmuster** (jeweils für die Jahreszeit mit dem höchsten Varianzerklärungsanteil im Untersuchungszeitraum), dargestellt für die historischen, uninitialisierten Modellläufe (*hist*; s. Erklärungen im Text) und die beobachteten Werte (*obs*). Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. Die Korrelationskoeffizienten drücken jeweils den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.

Für jeden Index der CPC-Telekonnectionsmuster lassen sich (wie für die Indizes der Nordatlantischen Oszillation nach Hurrell) mehr oder weniger viele Jahre feststellen, in denen sich die beobachteten Werte (graue Linie) außerhalb der hellblau markierten Fläche befinden, die Beobachtung also durch keine der drei Ensemblemitglieder der uninitialisierten Modellläufe erfasst wird. Insbesondere die Repräsentation von extremeren Beobachtungswerten scheint den uninitialisierten Modellläufen Probleme zu bereiten (vgl. die extremeren Werte z. B. bei *EA* oder *NAO* Ende der 1980er-Jahre oder bei *EAWR* in den 1990er-Jahren). Allerdings können auch Zeiträume identifiziert werden, in denen sich die beobachteten Werte mehrere Jahre innerhalb der hellblau markierten Fläche befinden und die beiden Linien zudem einen ähnlichen Verlauf aufweisen (z. B. Ende der 1990er- bzw. Anfang der 2000er-Jahre bei *TNH* oder die anfänglichen 2000er-Jahre bei *NAO*).

5.1.3 Südliche Oszillation (SO)

Da die Berechnung des Index‘ der Südlichen Oszillation (*SOI*) auf Grundlage von Gitterpunkten und nicht mithilfe einer Hauptkomponentenanalyse erfolgt (vgl. Kapitel 4.1.1.3; Tabelle 4.1), ist eine Angabe des Varianzerklärungsanteils und daraus resultierend eine Auswahl der Jahreszeit wie in den beiden vorherigen Kapiteln nicht möglich. Aus diesem Grund sind in Abbildung 5.4 die Ergebnisse für alle Jahreszeiten dargestellt, was einen

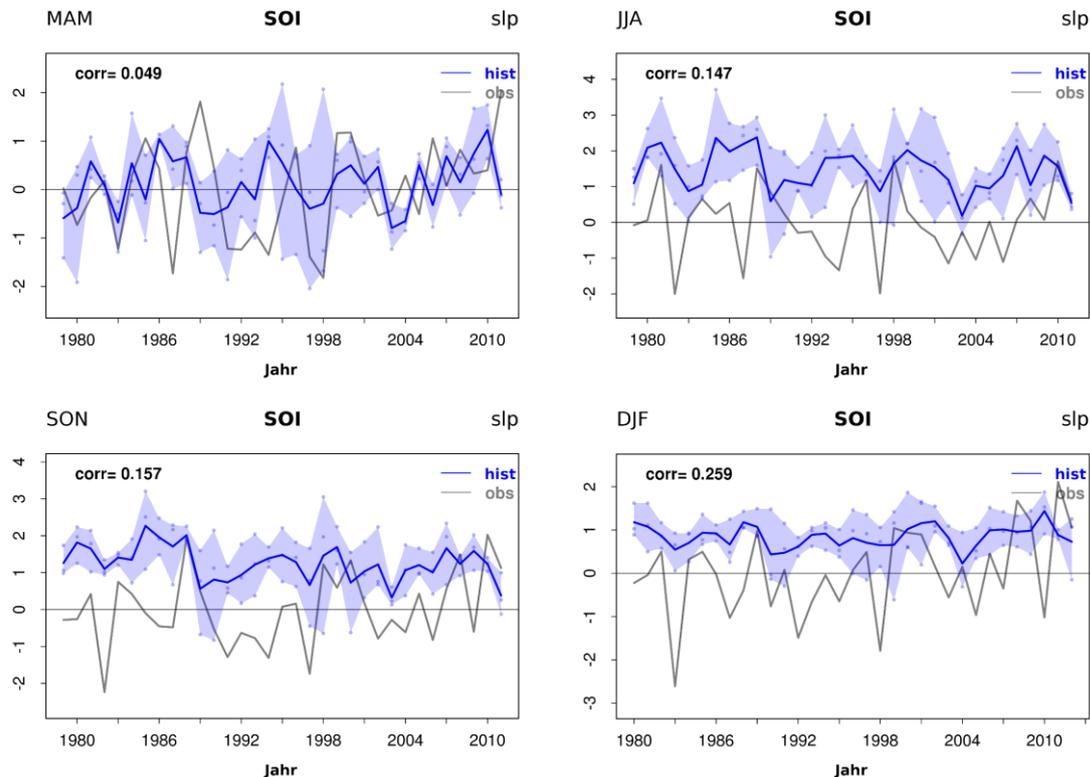


Abbildung 5.4: Standardisierter *Index der Südlichen Oszillation (SOI)* im Frühjahr (MAM), Sommer (JJA), Herbst (SON) und Winter (DJF) des Untersuchungszeitraums, dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs). Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. Die Korrelationskoeffizienten drücken jeweils den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.

differenzierten Einblick in die Vorhersage des *SOI* gewährt. Zudem erlaubt es die Berechnungsmethode, die auf Gitterpunktdifferenzen basiert, eine nachträgliche Standardisierung der Zeitreihe über die beobachteten und die historischen Werte hinweg vorzunehmen, wodurch eine Analyse nicht nur des gemeinsamen Verlaufs der Zeitreihen (wiederum ausgedrückt durch den Korrelationskoeffizienten), sondern auch des Niveaus ermöglicht wird.

Grundsätzlich lässt sich für alle Monate feststellen, dass das uninitialisierte Modell die Intensität der Walker-Zirkulation, gemessen als Luftdruckdifferenz aus den Gitterpunktdaten um die Insel Tahiti und Darwin, überschätzt, was daran zu erkennen ist, dass die Werte der Beobachtung (graue Linie) in allen Jahreszeiten mehr oder weniger deutlich unter denen der Modelldaten (blaue Linie) liegen (s. Abbildung 5.4). Insbesondere im Winter (DJF), Sommer (JJA) und Herbst (SON) ist dieses Ergebnis zu beobachten, lediglich im Frühjahr (MAM) liegen die Werte von Beobachtung und historischem Modell auf einem Niveau. Betrachtet man die Korrelationskoeffizienten als Maß für die Ähnlichkeit des Verlaufs der beiden Zeitreihen, lässt sich wie für die in den vorherigen Kapiteln untersuchten Telekonnektionsmuster nur ein geringer Zusammenhang feststellen. Im Frühjahr liegt der Wert mit $r=0,049$ nahe 0 und auch die anderen Jahreszeiten weisen mit $r=0,147$ (Sommer) bzw. $r=0,157$ (Herbst) geringe Werte auf. Auch im Winter ist mit $r=0,259$ nur ein schwach ausgeprägter Zusammenhang zu beobachten, der allerdings den stärksten der bisher untersuchten Telekonnektionsmuster darstellt.

5.1.4 Antarktische Oszillation (AAO)

Das Telekonnektionsmuster der AAO, dessen Berechnung auf dem 700 hPa-Geopotential (ZG700) beruht (vgl. Kapitel 4.1.1.4, Tabelle 4.1), weist mit einer erklärten Varianz von

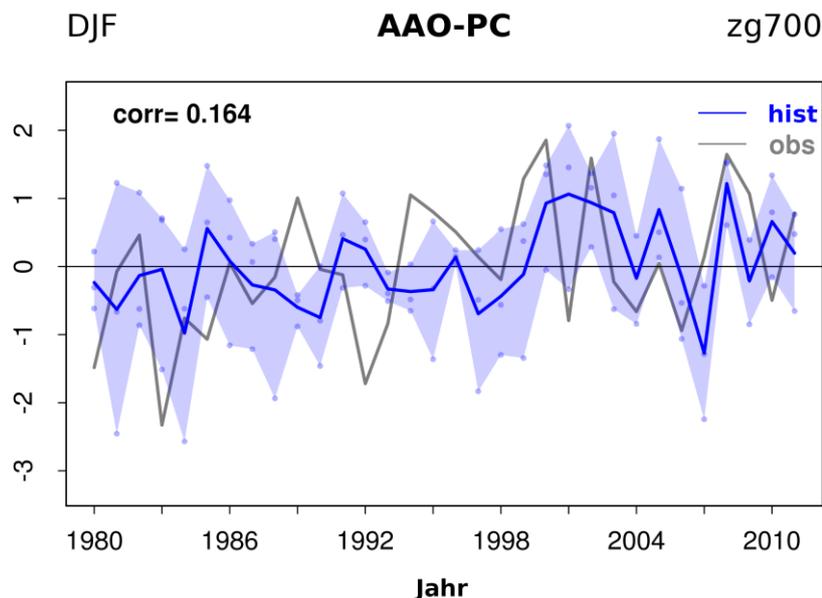


Abbildung 5.5: (Standardisierter) *Index der Antarktischen Oszillation (AAO-PC)* im Winter (DJF) des Untersuchungszeitraums, dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs). Der Korrelationskoeffizient drückt den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.

39,15 % (nicht dargestellt) im nordhemisphärischen Winter (DJF) den höchsten Varianzklärungsanteil auf – gegenüber 35,40 % im Frühjahr (MAM), 30,76 % im Sommer (JJA) und 31,86 % im Herbst (SON). Wie in Abbildung 5.5 für das auf einer Hauptkomponente basierende Berechnungsverfahren (*AAO-PC*) zu erkennen ist, gelingt es den uninitialisierten, historischen Modellläufen (**hist**) nicht, die beobachteten Werte (**obs**) adäquat zu repräsentieren, was sich zum einen an einem geringen Korrelationskoeffizienten von $r=0,164$ manifestiert. Zum anderen lässt sich wiederum beobachten, dass die Beobachtung häufig außerhalb des durch die drei Ensemblemitglieder hellblau markierten Bereichs liegt, was insbesondere auf die 1990er-Jahre zutrifft.

Für das Verfahren, das für die Berechnung der Antarktischen Oszillation die Differenz zonaler Mittel verwendet (*AAO-ZM*), kann ein vergleichbares Ergebnis konstatiert werden. Die beobachteten Werte liegen wiederum häufig außerhalb der hellblau markierten Fläche, d. h. die Beobachtung wird durch keine der drei Ensemblemitglieder angemessen repräsentiert, was auch daran liegt, dass das Modell die Differenz der zonalen Mittel generell unterschätzt. Außerdem weist der Korrelationskoeffizient zwischen Beobachtung und gemittelter Vorhersage einen geringen Wert von $r=0,063$ auf.

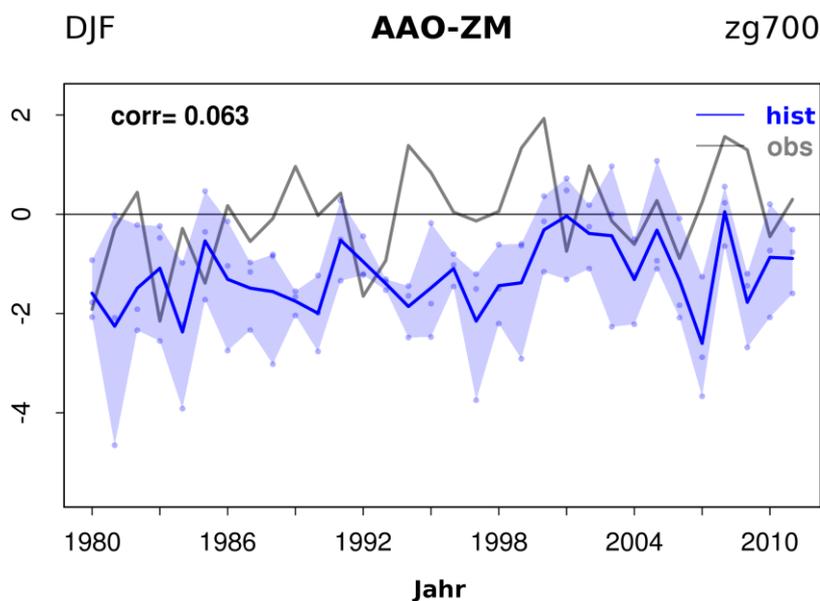


Abbildung 5.6: Wie Abbildung 5.5, aber für *AAO-ZM*.

Zusammenfassend lässt sich somit festhalten, dass es den uninitialisierten, historischen Modellläufen nicht gelingt, die Telekonnektionsmuster, die in der vorliegenden Arbeit untersucht werden, auf zufriedenstellende Art und Weise zu repräsentieren, was sich wiederholt in geringen Werten hinsichtlich der Korrelation manifestiert. Deshalb soll in den folgenden Kapiteln untersucht werden, inwiefern sich strukturelle Veränderungen am Vorhersagesystem (z. B. Implementierung einer Initialisierung, Änderung der Ensemblegröße) auf die Vorhersageleistung auswirken.

5.2 Der Einfluss der Initialisierung

Wie bereits mehrfach erwähnt wurde, stellen Fragen zur Initialisierung von Klimamodellen einen zentralen Forschungsaspekt auf dem Gebiet der dekadischen Vorhersage dar, da mithilfe verschiedener Initialisierungstechniken versucht wird, die klimainternen Schwankungen in geeigneter Weise in die Vorhersage zu integrieren. Aufgrund der Fülle an durchgeführten Einzelanalysen, die sich durch die verschiedenen Generationen des initialisierten Vorhersagemodells (s. Kapitel 3.1.1) ergibt, soll im Folgenden nicht auf die einzelnen Telekonnektionsindizes eingegangen werden. Vielmehr wird versucht, eine Überblicksdarstellung über die Vorhersageleistung – gemittelt und differenziert nach verschiedenen Untersuchungsparametern (Modellgenerationen, Telekonnektionsindizes, Jahreszeiten und zeitliche Auflösung) – zu ermöglichen.

Um dies zu erreichen, wurde eine möglichst große Datenbasis in die Analyse integriert. Außer TNH sowie EPNP, die nach Definition des CPC nicht für alle Jahreszeiten zur Verfügung stehen (vgl. Tabelle 5.1), gingen sowohl alle verfügbaren Telekonnektionsindizes als auch alle Jahreszeiten (MAM, JJA, SON, DJF) in die Mittelung ein. Die verschiedenen Modellgenerationen wurden jeweils mit der maximal verfügbaren Anzahl an Ensemblemitgliedern (vgl. Tabelle 3.1) in die Analyse aufgenommen. Außerdem wurde auf beide zeitliche Aggregationen (saisonal, monatlich) zurückgegriffen, allerdings erfolgte eine Beschränkung auf die geringere räumliche Auflösung des Klimamodells (MPI-ESM-LR), da die höhere Auflösung (MPI-ESM-MR) nur für **b1** und nicht für die anderen Modellgenerationen zur Verfügung stand. Aus diesem Grund werden im Folgenden auch keine Aussagen über Unterschiede in der Vorhersageleistung bezüglich der räumlichen Auflösung getroffen.

5.2.1 Differenzierung nach Modellgenerationen

Abbildung 5.7 informiert über die unterschiedlichen Generationen des dekadischen Vorhersagemodells von MPI-ESM. Abgebildet sind sowohl die Korrelation (CORR; Abbil-

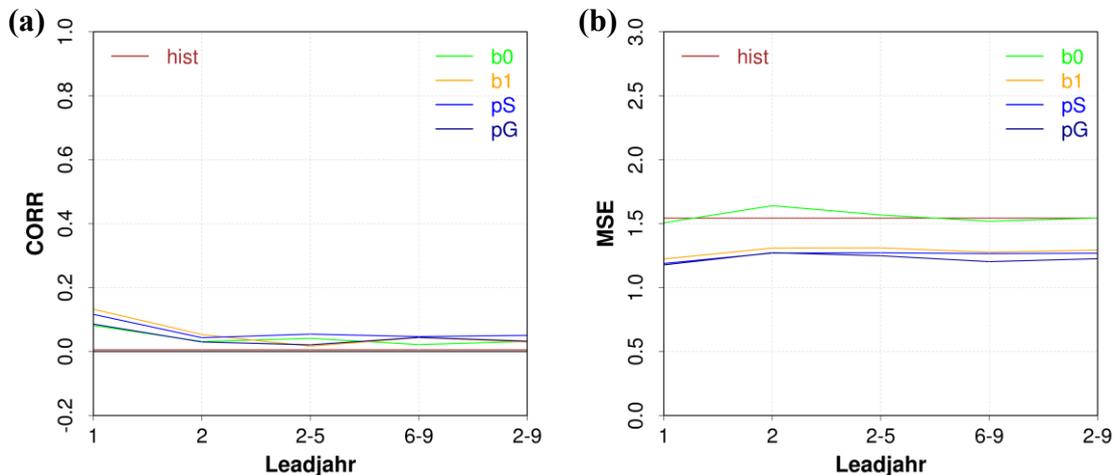


Abbildung 5.7: CORR (a) und MSE (b) für die unterschiedlichen **Modellgenerationen** des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT hinsichtlich der Telekonnektionsindizes in Abhängigkeit von verschiedenen Leadjahren.

dung 5.7a) als Maß für den Zusammenhang zwischen Vorhersage und Beobachtung (vgl. Formel (3)) als auch der MSE (Abbildung 5.7b) als Maß für den Unterschied zwischen Vorhersage und Beobachtung (vgl. Formel (4)), wobei als Beobachtung die Reanalyse von ERAINT verwendet wurde. MSE und CORR sind angegeben für verschiedene Leadjahre. Die ersten beiden Leadjahre nach Initialisierung des Vorhersagemodells sind gesondert ausgewiesen, da hier der Einfluss der Initialisierung von besonderem Interesse ist. Außerdem sind die Ergebnisse für die erste Hälfte der Dekade (gemittelt über Leadjahre 2-5) sowie die zweite Hälfte der Dekade (gemittelt über Leadjahre 6-9) dargestellt. Schließlich ist noch der gesamte Zeitraum der dekadischen Vorhersage unter Ausschluss des ersten Jahres nach der Initialisierung (Leadjahre 2-9 abgebildet). Dieses auf dem dekadischen Forschungsgebiet übliche Leadjahr-Schema wird auch in den folgenden Kapiteln verwendet.

Abbildung 5.7 zeigt für die verschiedenen Modellgenerationen – historical (**hist**), baseline0 (**b0**), baseline1 (**b1**), prototype mit ORAS4 (**pS**) prototype mit GECCO2 (**pG**) – deutliche Unterschiede auf. Der MSE als Maß für den Unterschied zwischen Beobachtung und dekadischer Vorhersage (s. Abbildung 5.7b) weist für die die uninitialisierten Läufe (**hist**) die höchsten Werte auf, d. h. hier liegen (die über alle Ensemblemitglieder gemittelte) Vorhersage und die Beobachtung (ERAINT) am weitesten auseinander. Ein ähnliches Ergebnis offenbart die erste Modellgeneration des dekadischen Vorhersagesystems (**b0**), die ähnliche Werte wie die uninitialisierten Läufe aufweist. Demgegenüber zeigen die folgenden Modellgenerationen (**b1**, **pS/pG**) deutlich geringere Werte für den MSE, sie liegen also näher an der Beobachtung. Die Ergebnisse für CORR (als Maß für die gemeinsame Varianz) bestätigen dieses Resultat (s. Abbildung 5.7a). Die geringsten Werte weisen die uninitialisierten, historischen Läufe mit einem Korrelationskoeffizienten nahe 0 auf. Leicht höhere Werte zeigt die erste Modellgeneration (**b0**). Für die beiden folgenden Modellgenerationen sind höhere Korrelationskoeffizienten zu erkennen, wobei prototype (**pS/pG**) mit Ausnahme für das erste Leadjahr die beste Vorhersageleistung zeigt. **B1** weist im ersten Jahr nach der Initialisierung zwar den höchsten Wert auf, allerdings ist der Wert im ersten Teil der Dekade (Leadjahre 2-5) ähnlich gering wie bei den uninitialisierten Läufen.

Der genauere Blick auf die einzelnen Leadjahre zeigt für die initialisierten Modellläufe eine Verbesserung in der Vorhersageleistung bezüglich des ersten Leadjahres. Beim MSE (s. Abbildung 5.7b) liegen alle Modellgenerationen des dekadischen Vorhersagesystems unter dem Wert der uninitialisierten, historischen Läufe. Die historischen Läufe, die für den Vergleich mit den initialisierten Läufen verwendet werden und keine dekadischen Vorhersagen liefern, weisen deshalb bei allen Leadjahren den gleichen Referenzwert auf. Betrachtet man die Werte der Korrelationskoeffizienten (s. Abbildung 5.7a), wird das Ergebnis bestätigt, da alle initialisierten Modelle im ersten Leadjahr im Schnitt einen deutlich höheren Korrelationskoeffizienten aufweisen als die uninitialisierten Modellläufe, die im Mittel einen Wert nahe 0 offenbaren. Allerdings zeigt sich auch, dass die Verbesserung in der Vorhersageleistung, die im ersten Leadjahr beobachtet werden kann, nicht für die anderen Leadjahre gilt. Lediglich für das zweite Leadjahr lässt sich eine Verbesserung in der Vorhersageleistung erahnen, wenn man, wie in diesem Fall, nicht einzelne Telekonnektionsindizes für einzelne Jahreszeiten untersucht, sondern über eine

Vielzahl der verfügbaren Analysen mittelt. Zudem muss konstatiert werden, dass sich die Verbesserungen insgesamt in einem sehr begrenzten Rahmen bewegen.

Ein scheinbarer Widerspruch offenbart sich, wenn man in Abbildung 5.7 die beiden Varianten der dritten Modellgeneration (**pS/pG**) miteinander vergleicht. Hinsichtlich des MSE (s. Abbildung 5.7b) zeigen die beiden Varianten von prototype, die sich lediglich in der Verwendung des Reanalysedatensatzes zur Initialisierung des Ozeans (vgl. Kapitel 3.1.1.3) unterscheiden, die geringsten Werte, d. h. sie weisen die geringsten Abweichungen zur Beobachtung auf. Die mit GECCO2 initialisierte Variante (**pG**) liegt in allen Leadjahren leicht unterhalb der mit ORAS4 initialisierten Version (**pS**) und offenbart somit eine bessere Vorhersageleistung. Demgegenüber liegen die Werte des Korrelationskoeffizienten von **pG** in allen Leadjahren leicht unterhalb derer von **pS** (s. Abbildung 5.7a). Demzufolge weist **pG** in dieser Hinsicht einen geringeren Zusammenhang mit der Beobachtung auf als **pS**. Dieser scheinbare Widerspruch wird aufgelöst, wenn man sich in Erinnerung ruft, dass der MSE den mittleren Abstand zwischen den beobachteten und vorhergesagten Werte berechnet, wohingegen der Korrelationskoeffizient den gemeinsamen Verlauf der Zeitreihen von Beobachtung und Vorhersage abbildet, ohne den Abstand der einzelnen Werte voneinander zu berücksichtigen.

5.2.2 Differenzierung nach Telekonnektionsindizes

Schaut man sich nun die Vorhersageleistung nicht für die verschiedenen Modellgenerationen, sondern aufgeschlüsselt für die einzelnen Telekonnektionsindizes an (Abbildung 5.8; vgl. Tabelle 4.1), zeigt sich ein ähnliches Bild zu dem soeben beschriebenen. Die Berechnungsgrundlage ist dieselbe wie bei Abbildung 5.7, jedoch wurden im Unterschied dazu keine historischen Läufe berücksichtigt. Eine Differenzierung hinsichtlich der verschiedenen Initialisierungstechniken ist nicht mehr möglich, da bei den dargestellten Werten über alle initialisierten Modellgenerationen gemittelt wurde. Allerdings zeigt sich bei der Grafik ebenfalls eine Verbesserung in der Vorhersageleistung für das erste (und tlw. das zweite) Leadjahr, was in höheren Korrelationskoeffizienten (Abbildung 5.8a) bzw. geringeren MSE-Werten (Abbildung 5.8b) Ausdruck findet.

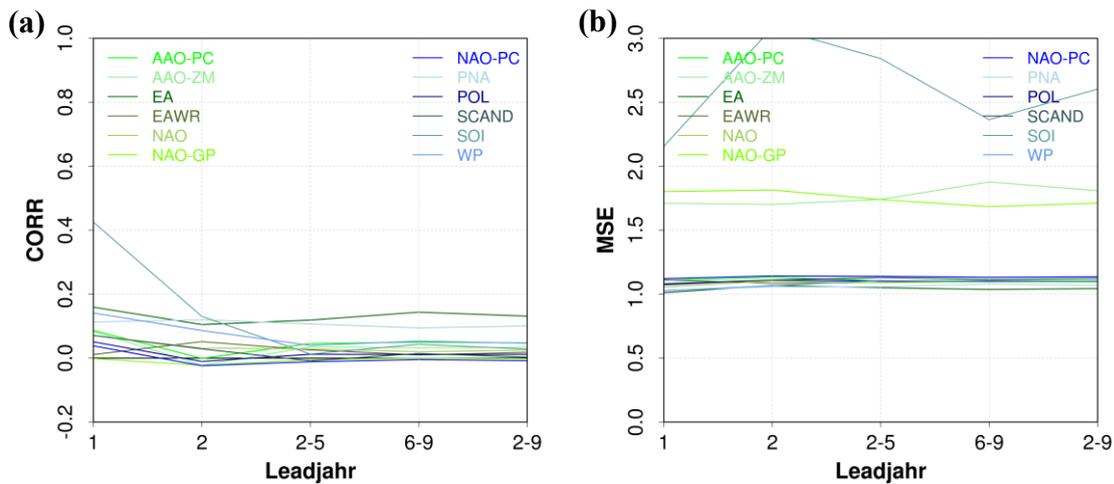


Abbildung 5.8: Wie Abbildung 5.7, aber für die unterschiedlichen **Telekonnektionsindizes**.

Außerdem offenbart Abbildung 5.8 eine ausgeprägte Heterogenität, wobei einige Indizes eine deutlich bessere Repräsentation (z. B. EA, PNA) aufweisen als andere (z. B. die verschiedenen Indizes der NAO). Eine besondere Stellung nimmt in diesem Zusammenhang SOI ein, da – gemittelt über alle initialisierten Modellgenerationen, Jahreszeiten und zeitliche Aggregationsniveaus – für die ersten beiden Leadjahre ein Korrelationskoeffizient von über 0,4 (erstes Leadjahr) bzw. 0,2 (zweites Leadjahr) beobachtet werden kann. Weil in den anderen Leadjahren aber deutlich geringe Werte nahe 0 zu erkennen sind, scheint sich die Initialisierung des dekadischen Vorhersagesystems beim Index der Südlichen Oszillation besonders deutlich bemerkbar zu machen.

5.2.3 Differenzierung nach Jahreszeiten

Ein vergleichbares Ergebnis tritt auf, wenn man dieselbe Berechnungsgrundlage wie bei Abbildung 5.8 verwendet, die aufgetragenen Werte allerdings nach Jahreszeiten (MAM, JJA, SON, DJF) differenziert (Abbildung 5.9). Auch hier ist aufgrund der Initialisierung – wie für die Erstellung von Abbildung 5.8 gingen nur die initialisierten Modellläufe in die Mittelung ein – eine deutliche Verbesserung in der Vorhersageleistung im ersten Leadjahr zu erkennen. Betrachtet man die einzelnen Jahreszeiten genauer, zeigen sich unabhängig vom Leadjahr und dem Einfluss der Initialisierung Unterschiede. Besonders auffallend ist die Vorhersageleistung im Sommer (JJA) mit vergleichsweise hohen MSE-Werten (s. Abbildung 5.9b), was eine relativ große Abweichung zwischen Vorhersage

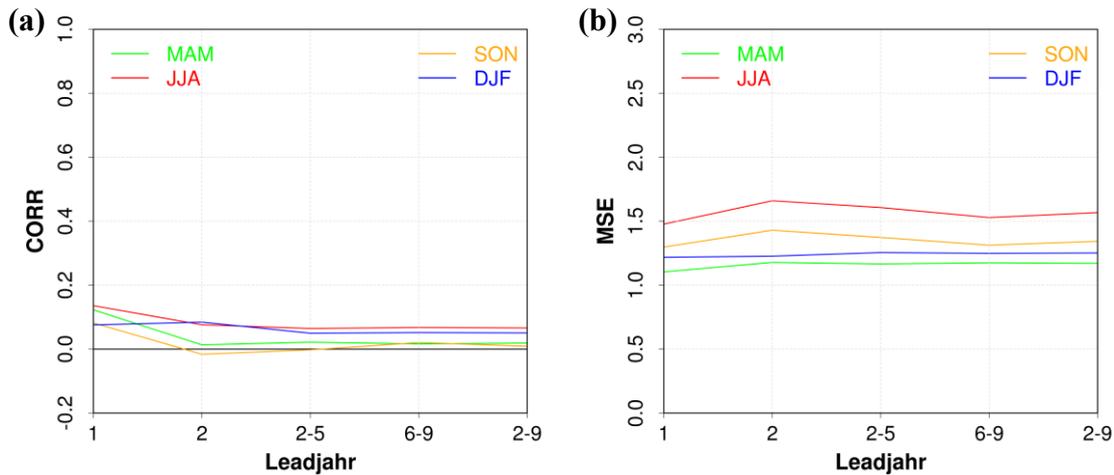


Abbildung 5.9: Wie Abbildung 5.7, aber für die unterschiedlichen Jahreszeiten.

und Beobachtung aufzeigt. Allerdings weist der Sommer auch die höchsten Werte hinsichtlich des Korrelationskoeffizienten auf (s. Abbildung 5.9a), was, wie bereits erläutert wurde, nur scheinbar einen Widerspruch darstellt. Demgegenüber zeigt z. B. der Herbst (SON) sowohl mit relativ hohen MSE-Werten als auch niedrigen Korrelationskoeffizienten eine im Vergleich mit den anderen Jahreszeiten (MAM, DJF) schlechte Vorhersageleistung.

Kritisch anzumerken ist an dieser Stelle, dass Telekonnektionsindizes der beiden Hemisphären in die Mittelung eingeflossen sind. Dennoch ist auch bei der nach Jahreszeiten differenzierten Betrachtungsweise deutlich der Einfluss der Initialisierung erkennbar, was anhand erhöhter Werte des Korrelationskoeffizienten (Abbildung 5.9a) bzw. erniedrigter MSE-Werte (Abbildung 5.9b) im ersten Leadjahr belegt werden kann. Dieser Effekt ist in nahezu allen Jahreszeiten deutlich zu beobachten, lediglich im Winter (DJF) ist er nicht so stark ausgeprägt, erstreckt sich dafür aber nicht nur über das erste, sondern auch über das zweite Leadjahr. In den anderen Jahreszeiten fällt hingegen auf, dass häufig die höchsten MSE-Werte und die geringsten Werte der Korrelation, d. h. eine vergleichsweise schlechte Vorhersageleistung, im zweiten Leadjahr auftreten, was auch in Abbildung 5.7 beobachtet werden kann.

5.2.4 Differenzierung nach zeitlicher Auflösung

Lässt man wiederum alle Untersuchungsergebnisse der initialisierten Läufe einfließen und differenziert diese nach der zeitlichen Auflösung, zeigt sich – wie in Abbildung 5.10

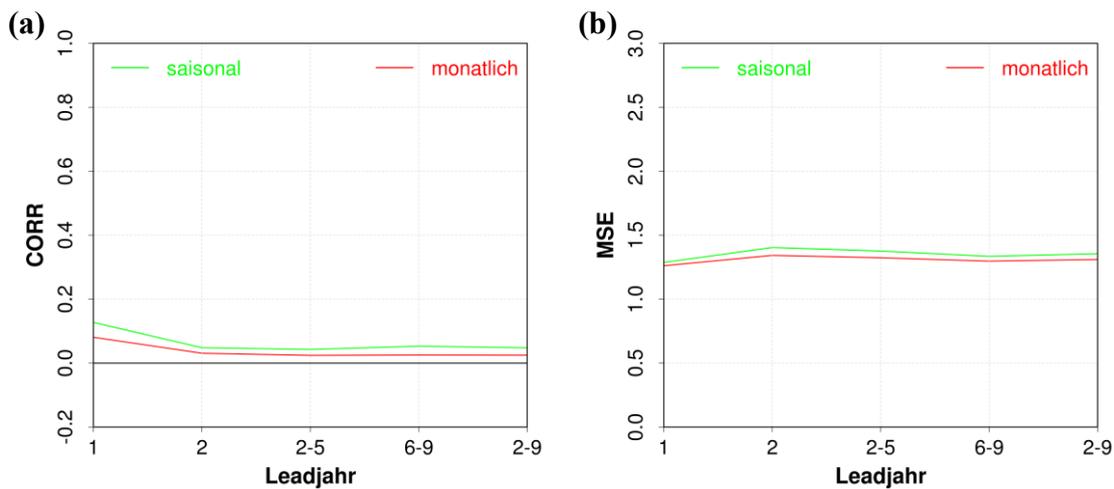


Abbildung 5.10: Wie Abbildung 5.7, aber für die unterschiedlichen zeitlichen Auflösungen.

zu erkennen ist – ein bereits bekanntes Bild. Sowohl bei monatlicher Aggregation der Daten als auch bei einer saisonalen Aggregation ist eine Verbesserung der Vorhersageleistung für das erste Leadjahr zu erkennen, was sich an den leicht erhöhten Korrelationswerten (s. Abbildung 5.10a) bzw. den etwas niedrigeren MSE-Werten (s. Abbildung 5.10b) erkennen lässt. Abbildung 5.10 zeigt zudem nur geringe Unterschiede zwischen den beiden unterschiedlichen zeitlichen Auflösungen. Es ist wiederum ein scheinbar widersprüchliches Ergebnis zu erkennen, da bei saisonaler Auflösung durchgängig höhere Korrelationswerte zu erkennen sind, allerdings auch höhere MSE-Werte festgehalten werden können, d. h. größere mittlere Abweichungen von der Beobachtung. Dieses scheinbar widersprüchliche Ergebnis lässt sich wieder mit den unterschiedlichen Messverfahren erklären.

Betrachtet man die dekadische Vorhersageleistung des MiKlip-Modellsystems überblickend und gemittelt über viele Analysen und differenziert nach den unterschiedlichen Modellgenerationen, Telekonnektionsindizes und Jahreszeiten, lässt sich zusammenfassend festhalten, dass die initialisierten Läufe durchaus eine Verbesserung in der Vorhersageleistung offenbaren, was durch geringere Werte bezüglich des MSE bzw. durch höhere Korrelationskoeffizienten belegt wird. Wie aus Abbildung 5.7 hervorgeht, zeigt sich diese Verbesserung v. a. bei der zweiten und den beiden Versionen der dritten Modellgeneration des dekadischen Vorhersagesystems (**b1**, **pG/pS**), d. h. den Modellvarianten mit einer „full-field“-Initialisierung. Demgegenüber fällt die erste Modellgeneration (**b0**), die noch keine ausgereifte Initialisierungstechnik aufweist (vgl. Tabelle 3.1), deutlich ab.

5.3 Der Einfluss der Ensemblegröße

Neben einer Initialisierung bietet die Erhöhung der Anzahl an Ensemblemitgliedern eine weitere potentielle Quelle für die Verbesserung der dekadischen Vorhersageleistung. Im vorherigen Kapitel wurden die Modellgenerationen lediglich mit ihrer maximal verfügbaren Anzahl an Ensemblemitgliedern in die Analyse integriert. Um den Einfluss der Ensemblegröße zu ermitteln, wird in diesem Kapitel untersucht, wie sich die Anzahl der Ensemblemitglieder auf die Vorhersage auswirkt, d. h. es wird die Vorhersageleistung des dekadischen Klimamodells mit unterschiedlichen Ensemblegrößen untersucht. Analysen hierzu wurden bereits für die saisonale Vorhersagbarkeit des NAO-Index‘ (Scaife et al. 2014) bzw. für die Vorhersagbarkeit zentraleuropäischer Sommertemperaturen auf dem dekadischen Forschungsgebiet (Sienz et al. 2016) durchgeführt, wobei letztere generell Ensemblegrößen kleiner als 10 nicht für die Evaluation dekadischer Modellvorhersagen empfehlen. In Anlehnung an die erwähnten Analysen wurde in der vorliegenden Arbeit ein Verfahren entwickelt, das zum einen versucht, eine möglichst breite Datenbasis zu ermöglichen, zum anderen die Rechenzeit der Computer aber nicht ins Unermessliche steigen lässt. Die Auswahl einer geringeren Anzahl an benötigten Ensemblemitgliedern aus dem größeren Ensemble (z. B. die Auswahl von sieben Ensemblemitgliedern aus den maximal möglichen 15 Ensemblemitgliedern bei **pG**) wurde mittels eines zufälligen Ziehens vorgenommen. Um die Robustheit der Ergebnisse zu garantieren, wurde dieser Zufallsprozess zehnmal durchgeführt und die Werte gemittelt.

Als Gütemaße wurden wiederum CORR als Maß für den Zusammenhang zwischen Vorhersage und Beobachtung sowie MSE als Maß für den Unterschied zwischen Vorhersage und Beobachtung verwendet. Aufgrund der hohen Rechenzeiten konnten nicht alle Telekonnektionsindizes in die Analyse aufgenommen werden, weshalb auf die Verwendung der CPC-Telekonnektionsmuster verzichtet wurde, sodass mit *AAO-PC*, *AAO-ZM*, *NAO-GP*, *NAO-PC* und *SOI* fünf Telekonnektionsindizes in die Analyse integriert wurden. Aus demselben Grund wurden lediglich die ersten beiden Leadjahre der dekadischen Vorhersage berücksichtigt. Zunächst sollen die Ergebnisse im Überblick differenziert nach den unterschiedlichen Modellgenerationen dargestellt werden, bevor auf einzelne Aspekte näher eingegangen wird.

5.3.1 Ensemblegröße und Modellgenerationen

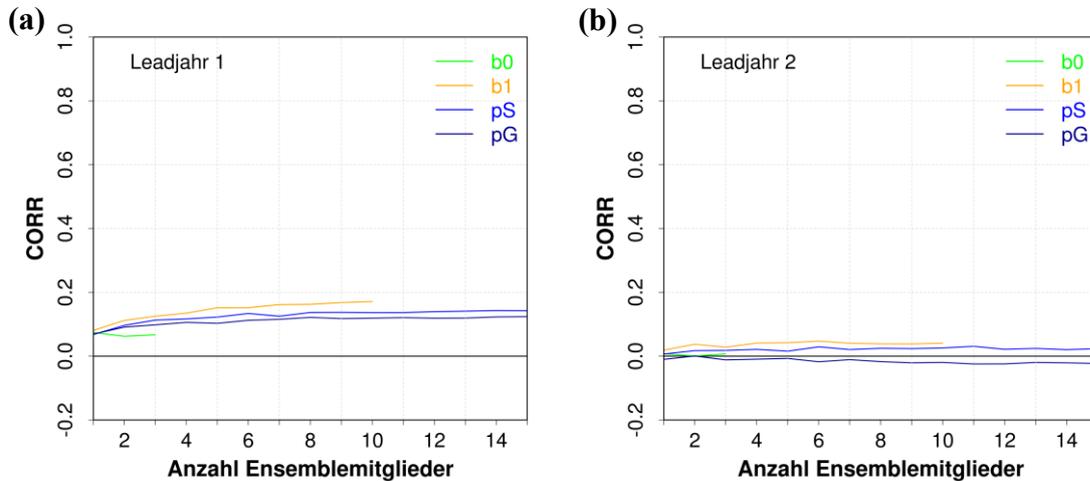


Abbildung 5.11: *CORR* zwischen den unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) und ERAINT hinsichtlich der ausgewählten Telekonnectionsindizes in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (a) bzw. das zweite Leadjahr (b) nach Initialisierung.

In Abbildung 5.11 ist für die verschiedenen Modellgenerationen des MiKlip-Vorhersagesystems der über alle ausgewählten Telekonnectionsindizes, alle Jahreszeiten sowie die beiden zeitlichen Auflösungen gemittelte Korrelationskoeffizient in Abhängigkeit von der Anzahl der Ensemblemitglieder für das erste Leadjahr (Abbildung 5.11a) bzw. das zweite Leadjahr (Abbildung 5.11b) abgebildet. Auf den ersten Blick gut erkennbar ist, dass die Korrelationskoeffizienten bezüglich aller Modellgenerationen für das erste Leadjahr höhere Werte aufweisen als für das zweite Leadjahr. Während sie im ersten Leadjahr (s. Abbildung 5.11a) zwischen $r=0$ und $r=0,2$ liegen, befinden sich diese im zweiten Leadjahr (s. Abbildung 5.11b) mehr oder weniger um $r=0$, was sowohl bei einer geringen Ensemblemitgliederanzahl zu beobachten ist als auch für ein größeres Ensemble gilt. D. h. auch bei Erhöhung Ensemblemitgliederanzahl lässt sich für das zweite Leadjahr auf Grundlage der gewählten Datenbasis kein Zusammenhang zwischen Vorhersage und Beobachtung erkennen. Ein anderes Bild zeigt sich für das erste Leadjahr, wo eine leichte Zunahme des Korrelationskoeffizienten mit einer zunehmenden Anzahl an Ensemblemitgliedern beobachtet werden kann. Am deutlichsten ist dieses Ergebnis bei der zweiten Modellgeneration (**b1**) zu beobachten, die zudem die höchsten Werte des Korrelationskoeffizienten aufweist. Aber auch die Resultate der dritten Modellgeneration (**pS**, **pG**) weisen diese Tendenz auf. Ausgenommen davon ist allerdings die erste Modellgeneration

b0, für die lediglich drei Ensemblemitglieder zu Verfügung stehen und die zudem noch die geringsten Werte des Korrelationskoeffizienten aufweist.

Eine deutliche Verbesserung in der Vorhersageleistung mit einer zunehmenden Anzahl an Ensemblemitgliedern lässt sich registrieren, wenn man als Gütemaß den MSE und nicht den Korrelationskoeffizienten heranzieht (vgl. Abbildung 5.12). Die Grafik ergibt sich wiederum als Mittelung der ausgewählten Telekonnektionsindizes, aller Jahreszeiten sowie der zeitlichen Aggregationen. Sie zeigt sowohl für das erste Leadjahr (s. Abbildung 5.12a) als auch für das zweite Leadjahr (s. Abbildung 5.12b) eine deutliche Ab-

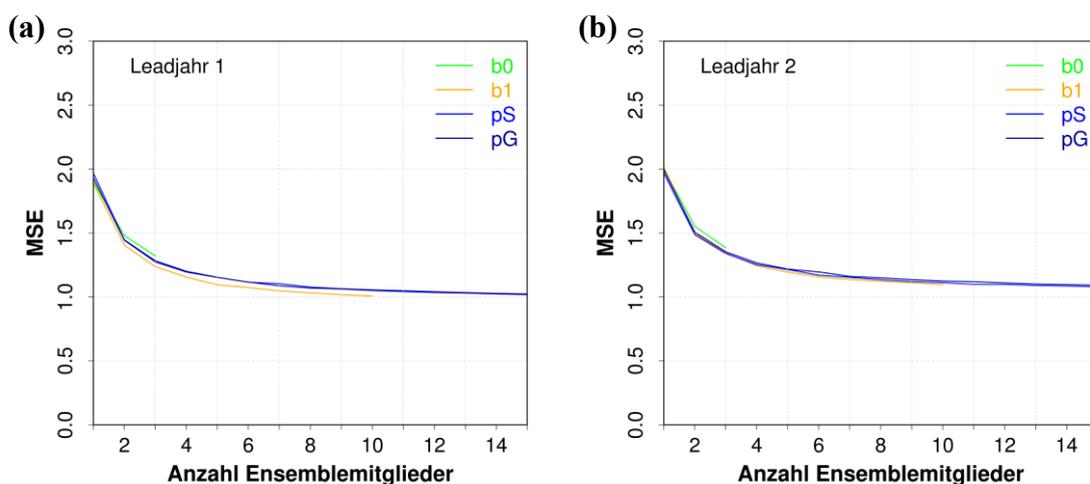


Abbildung 5.12: MSE für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT hinsichtlich der ausgewählten Telekonnektionsindizes in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (a) bzw. das zweite Leadjahr (b) nach Initialisierung.

nahme der Unähnlichkeit zwischen Beobachtung und der Vorhersage. Alle Modellgenerationen zeigen für beide Leadjahre den höchsten MSE-Wert bei einem Ensemblemitglied und den niedrigsten Wert bei der jeweiligen maximalen Anzahl an Ensemblemitgliedern. Auch hier fällt in Einklang mit den Ergebnissen der Korrelationsanalyse auf, dass die zweite Modellgeneration (**b1**) v. a. im ersten Leadjahr die stärkste Verbesserung in der Vorhersageleistung bei steigender Anzahl an Ensemblemitgliedern offenbart. Wiederum weist das erste Leadjahr bessere Ergebnisse auf, erkennbar an dem durchgehend niedrigeren Niveau der Werte bei jeglicher Ensemblegröße für alle Modellgenerationen. Insgesamt muss allerdings festgehalten werden, dass trotz der beobachteten Verbesserungen ein großer Unterschied zwischen Vorhersage und Beobachtung existiert, was daran zu

erkennen ist, dass die Werte deutlich von einer perfekten Vorhersage (MSE=0 bzw. CORR=1) abweichen.

5.3.2 Ensemblegröße und Telekonnektionsindizes

Differenzierte Erkenntnisse über und tiefere Einblicke in die Vorhersageleistung des dekadischen Vorhersagesystems erhält man, wenn man die Ergebnisse der einzelnen Telekonnektionsindizes betrachtet, anstatt eine Mittelung über alle hinweg vorzunehmen. Dabei zeigen sich große Unterschiede zwischen den einzelnen Indizes, auf die im Folgenden eingegangen werden soll.

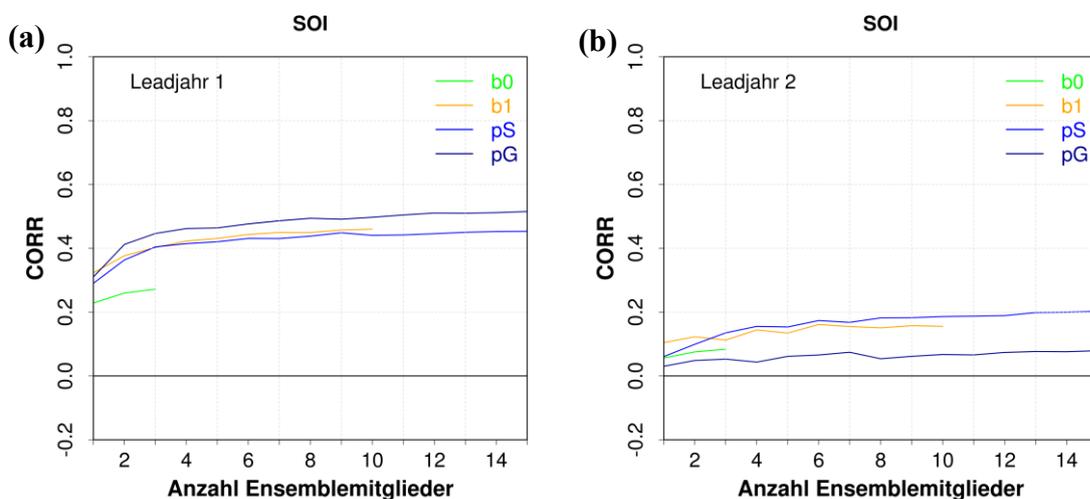


Abbildung 5.13: Wie Abbildung 5.11 (CORR), aber lediglich für den **Southern Oscillation Index** (alle anderen Indizes gingen nicht in die Mittelung ein).

Besonders bemerkenswert sind die Resultate für den Index der Southern Oscillation (SOI). Wie in Kapitel 5.2.2 bereits angedeutet wurde, nimmt der SOI aufgrund der vergleichsweise hohen Werte des Korrelationskoeffizienten eine Sonderstellung ein, was Abbildung 5.13, in dem lediglich die Resultate für den SOI veranschaulicht sind, zu bestätigen vermag. Basieren die Vorhersagen lediglich auf einem Ensemblemitglied, nimmt die Korrelation für das erste Leadjahr (s. Abbildung 5.13a) Werte von $r \approx 0,3$ an, was durch die Erhöhung der Anzahl an Ensemblemitgliedern auf $r = 0,4$ und höher gesteigert werden kann, wobei **pG** bei der maximal verfügbaren Ensemblegröße Werte von über 0,5 aufweist. Lediglich die erste Modellgeneration (**b0**) fällt diesbezüglich deutlich ab. Im zweiten Leadjahr (Abbildung 5.13b) ist dieser Effekt ebenfalls zu beobachten, allerdings in

abgeschwächter Form: Zwar lässt sich eine Verbesserung der Vorhersageleistung mit zunehmender Anzahl an Ensemblemitgliedern erkennen, allerdings ist diese nicht so stark ausgeprägt. Zudem sind die Korrelationswerte als Ausdruck des Zusammenhangs zwischen Beobachtung und Vorhersage nicht so hoch wie im ersten Leadjahr. Im Gegensatz zum ersten Leadjahr zeigt **pS** sowohl die höchsten Werte der Korrelation als auch die höchsten Steigerungsraten bei Vergrößerung des Ensembles, wohingegen das Modell **pG**, das im ersten Leadjahr am besten performt, für das zweite Leadjahr die geringsten Werte aufweist und sogar bei der maximal verfügbaren Anzahl an Ensemblemitgliedern keine besseren Vorhersageleistungen offenbart als die erste Modellgeneration (**b0**).

Betrachtet man für den SOI die Entwicklung des MSE in Abhängigkeit von der Anzahl an Ensemblemitgliedern (Abbildung 5.14), lassen sich die Ergebnisse von CORR bestätigen. Die Werte für den MSE liegen zum einen deutlich unter denen bei Mittelung über alle in die Analyse einbezogenen Telekonnektionsindizes (s. Abbildung 5.12), d. h. ein geringerer Unterschied zwischen Vorhersage und Beobachtung besteht. Zum anderen lässt sich bei beiden Leadjahren eine deutliche Verbesserung in der Vorhersageleistung bei größer Anzahl an Ensemblemitgliedern erkennen, wobei wiederum **pG** im ersten Leadjahr (Abbildung 5.14a) bzw. **pS** im zweiten Leadjahr (Abbildung 5.14b) die besten Ergebnisse aufweisen. Die erste Modellgeneration (**b0**) bildet hierbei wiederum eine deutlich erkennbare Ausnahme.

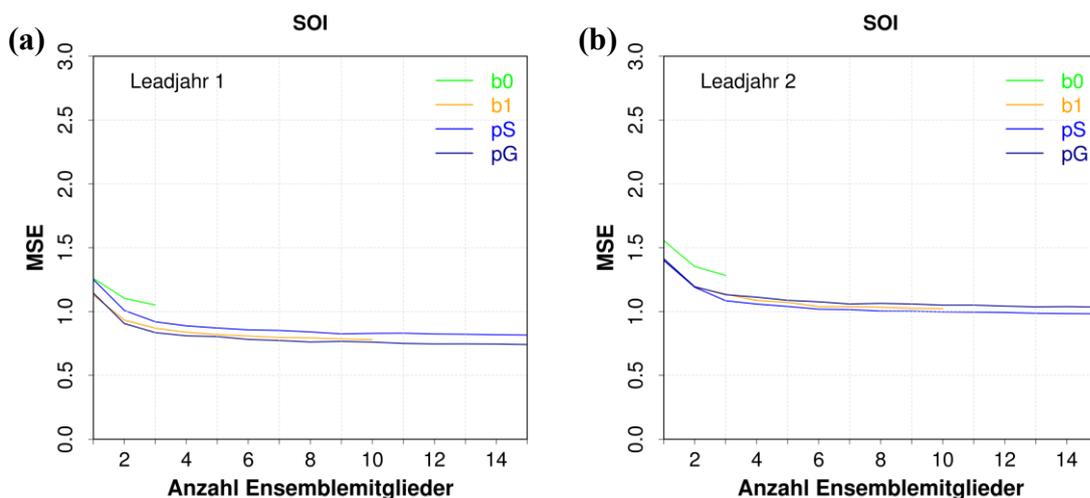


Abbildung 5.14: Wie Abbildung 5.12 (MSE), aber für den *Southern Oscillation Index* (alle anderen Indizes gingen nicht in die Mittelung ein).

Dass eine Erhöhung der Anzahl an Ensemblemitgliedern nicht immer zu einer deutlichen und eindeutigen Verbesserung in der Vorhersageleistung führt, zeigt das Beispiel für den PC-basierten Index der Antarktischen Oszillation (AAO-PC). Im ersten Leadjahr (Abbildung 5.15a) liegen die Werte des Korrelationskoeffizienten nahe 0, weisen also keinen Zusammenhang zwischen Vorhersage und Beobachtung auf. Lediglich die Werte für die zweite Modellgeneration des MPI-ESM-Vorhersagesystems (**b1**) liegen im leicht posi-

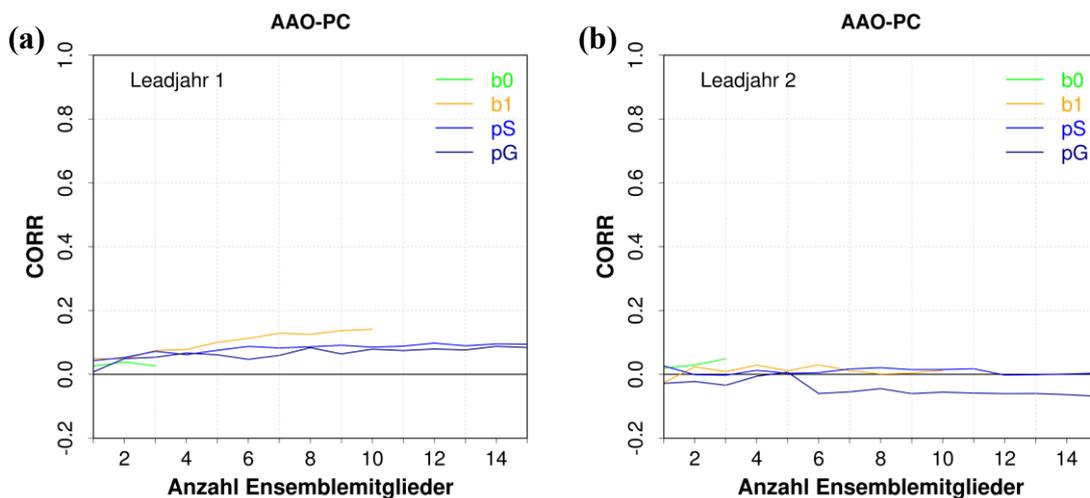


Abbildung 5.15: Wie Abbildung 5.11 (*CORR*), aber für den *PC-basierten Index der Antarktischen Oszillation* (alle anderen Indizes gingen nicht in die Mittelung ein).

ven Bereich und deuten damit einen schwachen Zusammenhang an. Zudem ist für **b1** eine geringe Verbesserung bei steigender Ensemblegröße zu beobachten, was für die anderen Modellgenerationen nicht behauptet werden kann. Bei **pS** und bei **pG** zeigt die Erhöhung der Anzahl an Ensemblemitgliedern lediglich einen geringen Einfluss, bei **b0** ist sogar eine geringe Verschlechterung in der Vorhersageleistung auf allerdings auch niedrigem Niveau zu beobachten. Für das zweite Leadjahr (s. Abbildung 5.15b) kann für alle Modellgenerationen keine Verbesserung konstatiert werden, d. h. eine Vergrößerung des Ensembles hat keinen Einfluss auf die Vorhersageleistung, wenn man den Korrelationskoeffizienten als Gütemaß zurate zieht.

Demgegenüber zeigt die Grafik mit dem MSE als Gütemaß ein bekanntes Bild. Sowohl im ersten Leadjahr (s. Abbildung 5.16a) als auch im zweiten Leadjahr (s. Abbildung 5.16b) zeigt sich für alle Modellgenerationen die eindeutige Tendenz mit abnehmenden

MSE-Werten bei zunehmender Ensemblegröße. Für das erste Leadjahr weist wiederum **b1** die besten Vorhersageleistungen, d. h. die geringsten MSE-Werte, auf.

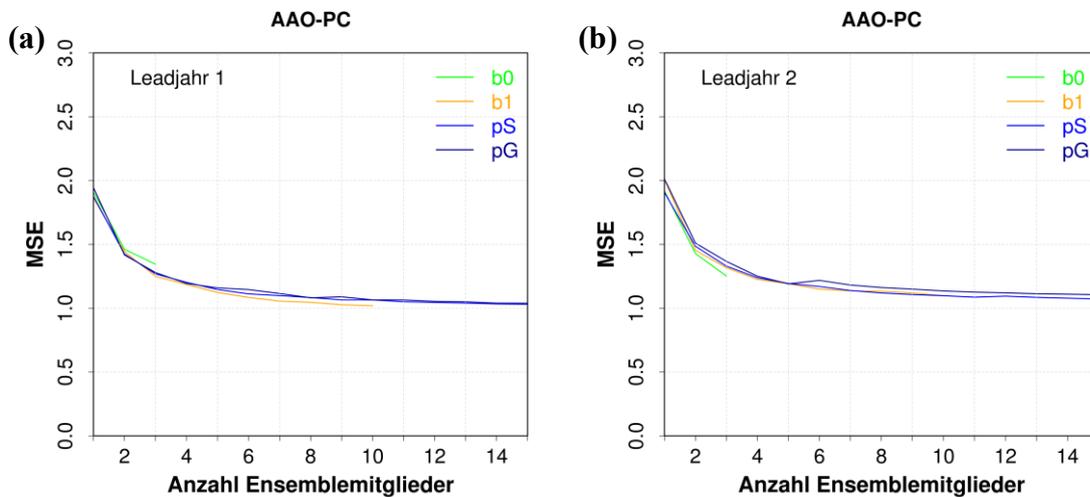


Abbildung 5.16: Wie Abbildung 5.12 (MSE), aber für den PC-basierten Index der Antarktischen Oszillation (alle anderen Indizes gingen nicht in die Mittelung ein).

Wenn man die Ergebnisse für die zwei ausgewählten Telekonnectionsindizes überblickt, fällt auf, dass bei der Vorhersageleistung keine Einheitlichkeit zwischen den einzelnen Modellgenerationen besteht. Während im ersten Leadjahr bei SOI die dritte Modellgeneration mit der GECCO2-Initialisierung (**pG**) die größten Verbesserungen bei Vergrößerung des Ensembles zeigt, weist diese Eigenschaft die zweite Modellgeneration (**b1**) bei AAO-PC auf. Schließt man die weiteren Telekonnectionsindizes, die für die Untersuchungen zum Einfluss der Ensemblegröße ausgewählt wurden, in die Überlegungen mit ein, wird dieses heterogene Bild bestätigt. Beim anhand des zonalen Mittels berechneten Index der Antarktischen Oszillation (AAO-ZM, Anhang 1) lässt **b1** die höchsten Steigerungsraten erkennen, beim Gitterpunkt-basierten Index der Nordatlantischen Oszillation (NAO-GP, Anhang 2) hingegen die dritte Modellgeneration mit ORAS4-Initialisierung (**pS**). Beim PC-basierten Index der Nordatlantischen Oszillation (NAO-PC, Anhang 3) weist diese Eigenschaft wiederum die zweite Modellgeneration (**b1**) auf. Eine Verbesserung in der Vorhersageleistung durch eine Vergrößerung des Ensembles lässt somit nicht mit einer bestimmten Modellgeneration in Verbindung bringen.

5.3.3 Ensemblegröße und Jahreszeiten

Betrachtet man den Einfluss der Ensemblegröße auf die Vorhersageleistung differenziert für die unterschiedlichen Jahreszeiten, zeigt sich ein bekanntes Bild. In Abbildung 5.17 sind die Grafiken für die unterschiedlichen Jahreszeiten (MAM, JJA, SON, DJF) mit CORR als Gütemaß für das erste Leadjahr dargestellt. Im Allgemeinen ist eine Zunahme der Vorhersageleistung, ausgedrückt durch höhere Korrelationskoeffizienten, bei Vergrößerung des Ensembles zu erkennen. Dennoch treten auch deutliche Unterschiede zwischen den einzelnen Jahreszeiten auf. Besonders der Sommer (s. Abbildung 5.17b) sticht hervor und nimmt in mehrfacher Hinsicht eine Sonderstellung ein. Die beiden Varianten

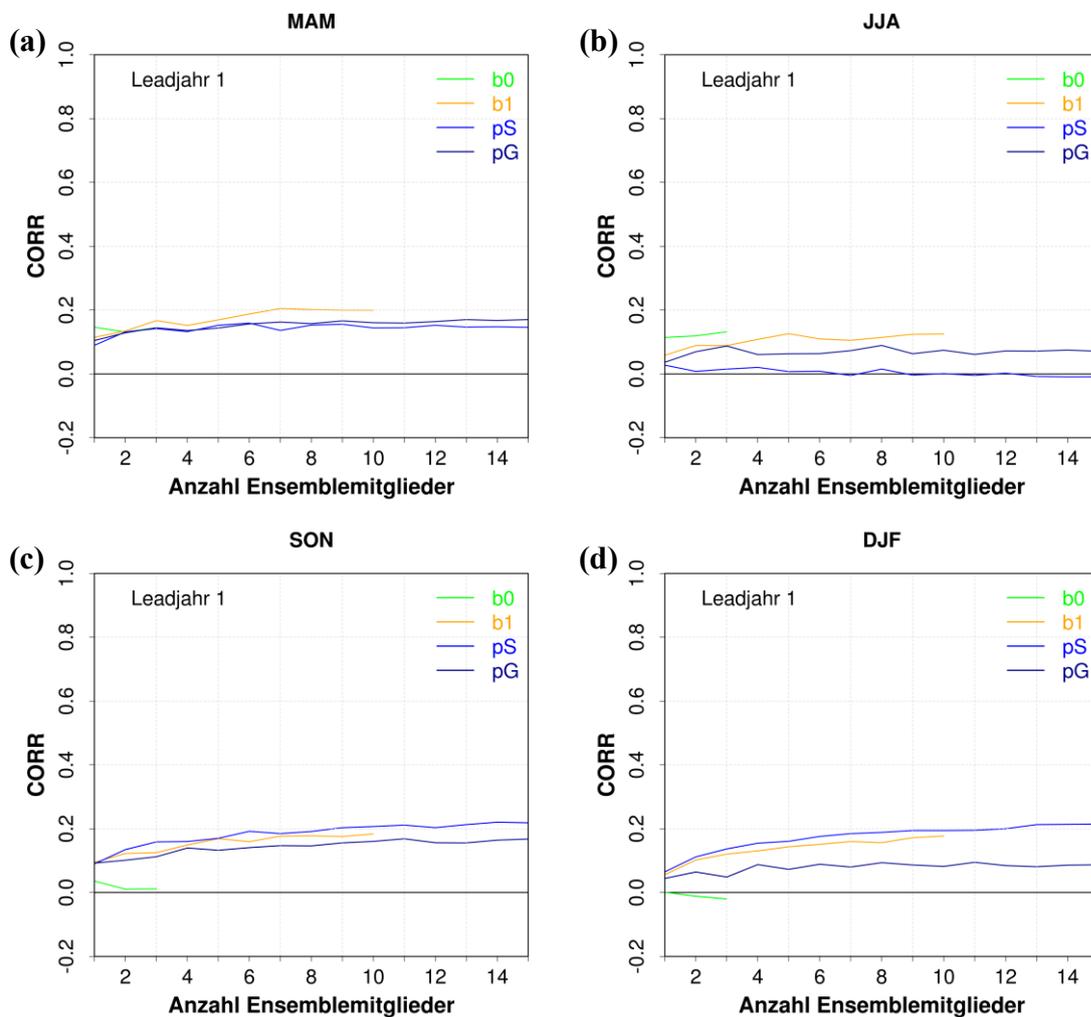


Abbildung 5.17: CORR für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT hinsichtlich der ausgewählten Telekonnektionsindizes in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für verschiedene Jahreszeiten (a: MAM, b: JJA, c: SON, d: DJF) des ersten Leadjahrs nach Initialisierung.

der dritten Modellgeneration (**pS**, **pG**) weisen im Gegensatz zu allen anderen Jahreszeiten keine steigenden Werte des Korrelationskoeffizienten bei zunehmender Ensemblegröße auf. Zudem ist die Steigerungsrate bei der zweiten Modellgeneration (**b1**) nicht so stark ausgeprägt wie in den anderen Jahreszeiten und der höchste Wert liegt nicht bei der maximal verfügbaren Anzahl an Ensemblemitgliedern, sondern bei einer Ensemblegröße von fünf. Auffallend ist zudem, dass die erste Modellgeneration (**b0**), die in den übrigen Jahreszeiten den schlechtesten Skill offenbart und auch bei Vergrößerung des Ensembles keine Verbesserung zeigt, im Sommer positiv hervorsteht. Allerdings muss die geringe Anzahl an verfügbaren Ensemblemitgliedern bei der Bewertung der Aussagekraft berücksichtigt werden. Vergleicht man die einzelnen Modellgenerationen in ihrer Vorhersageleistung über die Jahreszeiten hinweg, erkennt man, dass die verschiedenen Modellgenerationen unterschiedlich performen. Während im Herbst (s. Abbildung 5.17c) und Winter (s. Abbildung 5.17d) **pS** die höchsten Steigerungsraten zeigt, offenbart dies **b1** im Frühjahr (s. Abbildung 5.17a) des ersten Leadjahres. Es ergibt sich also wiederum kein einheitliches Bild. Die äquivalenten Grafiken mit dem MSE als Gütemaß, die die getroffenen Aussagen im Wesentlichen bestätigen und an dieser Stelle deswegen nicht ausführlich erläutert werden, befinden sich in Anhang 4.

Zusammenfassend lässt sich für die Untersuchungen zum Einfluss der Ensemblegröße auf die Vorhersageleistung des MPI-ESM-Vorhersagesystems konstatieren, dass eine Vergrößerung des Ensembles im Allgemeinen eine Verbesserung in der Vorhersageleistung bewirkt, was sich sowohl beim Korrelationskoeffizienten (als Maß des Zusammenhangs) als auch beim MSE (als Maß der Unterschiedlichkeit) insbesondere für das erste Leadjahr feststellen lässt. Diese Aussage muss im Einzelnen allerdings differenziert getätigt werden. Einige Indizes (wie z. B. der SOI) zeigen für mehrere Modellgenerationen und unabhängig vom Gütemaß eine deutliche Steigerung bei der Vorhersage, bei anderen Indizes (z. B. AAO-PC) kann dies nicht bestätigt werden. Außerdem sind Unterschiede in der Verbesserung der Vorhersageleistung für die unterschiedlichen Jahreszeiten erkennbar. Insbesondere der Sommer (JJA) fällt in diesem Zusammenhang auf. Da für die verschiedenen Telekonnektionsindizes und Jahreszeiten auch jeweils unterschiedliche Modellgenerationen des Vorhersagesystems am besten performen, lässt sich keine Empfehlung für eine bestimmte Modellgeneration aussprechen. Lediglich die erste Modellgeneration (**b0**) kann in diesem Zusammenhang ausgeschlossen werden.

5.4 Vorhersageleistung des dekadischen Modells

Wie in den vorherigen Kapiteln erläutert wurde, spielen sowohl die Initialisierungstechnik als auch die Ensemblegröße eine zentrale Rolle für die Vorhersageleistung des dekadischen Modells. Außerdem differieren die verschiedenen Telekonnectionsindizes miteinander deutlich in ihrer Vorhersagbarkeit, wie in den bisherigen Ausführungen bereits angedeutet wurde. Deshalb soll in Kapitel 5.4.1 detailliert analysiert werden, welche Telekonnectionsindizes durch eine gute Vorhersageleistung auffallen, um Quellen der dekadischen Vorhersagbarkeit identifizieren zu können. Darüber hinaus soll untersucht werden, wie die einzelnen Modellgenerationen, die sich in ihrer jeweiligen Initialisierungstechnik und ihrer Ensemblegröße voneinander unterscheiden, im Vergleich abschneiden. Um potentielle Verbesserungen zwischen den einzelnen Modellgenerationen messen zu können, kommen in Kapitel 5.4.2 mit dem Mean Squared Error Skill Score (MSSS, vgl. Kapitel 4.2.2.2) und dem Ranked Probability Skill Score (RPSS, vgl. Kapitel 4.2.2.3) zwei Gütemaße zur Anwendung, die in ihrer kombinierten Anwendung auch Aussagen darüber erlauben, worauf die Verbesserungen im Einzelnen zurückzuführen sind.

5.4.1 Betrachtung ausgewählter Telekonnectionsindizes

Aus den bisherigen Untersuchungsergebnissen lässt sich bereits ableiten, dass die Vorhersageleistungen für die einzelnen Telekonnectionsindizes ein heterogenes Bild offenbaren. Tabelle 5.2, die für das erste Leadjahr für verschiedene Indizes und Jahreszeiten den Korrelationskoeffizienten zwischen **pS** und der Beobachtung (ERAINT) der Jahre 1979-2011 bei saisonaler Aggregation zeigt, bestätigt das heterogene Bild.

*Tabelle 5.2: Korrelationskoeffizienten verschiedener Indizes und Jahreszeiten zwischen Vorhersage (**pS**) und Beobachtung (ERAINT) der Jahre 1979-2011 (saisonale Aggreg.) für das erste Leadjahr. Fett markierte Werte zeigen Signifikanz auf dem 5 %-Niveau (nach Pearson).*

	AAO-PC	AAO-ZM	EA	EAWR	NAO	NAO-GP	NAO-PC	PNA	POL	SCAND	SOI	WP
MAM	0,275	0,170	0,219	-0,152	0,174	-0,093	-0,108	-0,011	0,099	0,100	0,782	0,454
JJA	-0,313	-0,209	0,591	0,011	0,164	-0,083	0,001	0,284	0,125	-0,167	0,350	0,108
SON	0,334	0,196	0,111	-0,058	0,111	0,249	-0,034	0,025	0,084	0,252	0,558	-0,204
DJF	0,239	0,118	0,068	0,098	-0,048	0,235	0,409	-0,048	0,203	0,084	0,396	0,462

Das für die zweite und dritte Modellgeneration ausgewählte repräsentative Beispiel verdeutlicht das Auftreten relativ niedriger Werte nahe 0 bei vielen Telekonnectionsindizes (z. B. EAWR, POL). Im Speziellen sei an dieser Stelle auf den für die Klimatologie der nordatlantisch-europäischen Region im Winter (DJF) besonders relevanten NAO-Index verwiesen, der für die PCA-basierte Methode nach CPC (NAO) einen Korrelationskoeffizienten von $r=-0,048$ aufweist, bei der PCA-basierten Methode nach Hurrell (NAO-PC) allerdings einen Wert von $r=0,409$ zeigt. Höhere Korrelationskoeffizienten treten nicht durchgängig, sondern für die meisten Indizes nur in einzelnen Jahreszeiten auf, wie z. B. die auf dem 5%-Niveau signifikanten Werte von EA im Sommer ($r=0,591$), NAO-PC im Winter ($r=0,409$) oder WP im Frühling ($r=0,454$) bzw. Winter ($r=0,462$)

Die Werte des SOI bestätigen die Sonderstellung, wobei lediglich für das in Tabelle 5.2 gezeigte erste Leadjahr durchgängig signifikante Korrelationswerte auftreten. Mit $r=0,782$ offenbart der SOI insbesondere im nordhemisphärischen Frühjahr (MAM) des ersten Leadjahres, d. h. dem dritten, vierten und fünften Monat nach Initialisierung des dekadischen Modells, einen hohen Wert, der sich auch für die anderen Modellgenerationen (**b1**: $r=0,776$, **pG**: $r=0,751$; nicht dargestellt) zeigt. Abbildung 5.18a veranschaulicht die großen Übereinstimmungen im Verlauf des Index' zwischen der über alle 15 Ensemblemitglieder gemittelten Vorhersage (**pS**) und der Beobachtung (ERAINT). Zudem ist erkennbar, dass sich die Vorhersage nahezu konstant innerhalb der schattierten Fläche, d. h. der Spannweite aller Ensemblemitglieder, befindet und diese zudem in einem engen

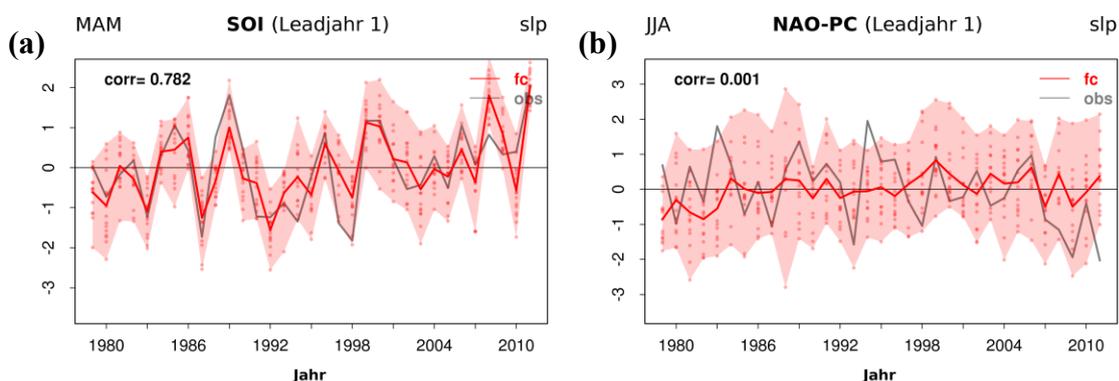


Abbildung 5.18: Saisonale aggregierter **SOI** im Frühjahr (a) bzw. **NAO** (nach Hurrell) im Sommer (b) des ersten Leadjahres für die Beobachtung (*obs*: ERAINT) und die gemittelte Vorhersage (*fc*: **pS**) unter Angabe des Korrelationkoeffizienten. Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage.

Bereich um die Beobachtung liegt. Die Homogenität zwischen den einzelnen Ensemblemitgliedern in der Vorhersage lässt sich quantifizieren durch das sog. Signal-Rausch-Verhältnis (engl. *signal-to-noise ratio*, SNR), welches als Relation von SD_{mean} und SD_{all} berechnet wird

$$SNR = 1 - \frac{SD_{\text{mean}}}{SD_{\text{all}}}, \quad (8)$$

wobei SD_{mean} die Standardabweichung des Ensemblemittels (d. h. der über alle Ensemblemitglieder gemittelten Vorhersage) und SD_{all} die Standardabweichung aus allen Ensemblemitgliedern darstellen (vgl. Kumar, 2009). Das Signal-Rausch-Verhältnis nimmt beim SOI einen hohen Wert von $SNR=0,812$ an. Einen Gegensatz dazu stellt NAO-PC im Sommer (JJA) des ersten Leadjahres dar, d. h. dem sechsten, siebten und achten Monat nach Initialisierung (s. Abbildung 5.18b). Bei einem Korrelationskoeffizienten von $r=0,001$ zeigen die beiden Linien keine Gemeinsamkeiten in ihrer Variation. Zwar liegt die Beobachtung wie beim vorherigen Beispiel nahezu konstant innerhalb der schraffierten Fläche, allerdings liegt diese in einem sehr breiten Bereich um die Beobachtung, was die Unsicherheit der Vorhersage offenbart. Diese Tatsache mache sich auch an einem niedrigen Wert des Signal-Rausch-Verhältnisses von $SNR=0,410$ bemerkbar. Die beiden repräsentativ ausgewählten Beispiele von SOI und NAO-PC zeigen die gesamte Spannweite an Vorhersagbarkeit von Telekonnektionsindizes durch das dekadische Modell, ohne auf alle Einzelfälle einzugehen.

Zu berücksichtigen gilt aber, dass sie lediglich das erste Leadjahr repräsentieren. Aus diesem Grund ist in Abbildung 5.19 der auf einer Hauptkomponentenanalyse basierende Index der Antarktischen Oszillation (AAO-PC) für das erste (Abbildung 5.19a) bzw. das zweite (Abbildung 5.19b) dargestellt. An dem konkreten Beispiel soll repräsentativ für die anderen Telekonnektionsindizes veranschaulicht werden, was bereits in Kapitel 5.2 angedeutet wurde. Während im ersten Leadjahr ein Korrelationskoeffizient von $r=0,334$ für den Herbst (SON), d. h. den neunten, zehnten und elften Monat nach Initialisierung des Modells ausgewiesen werden kann (s. Abbildung 5.19a), vermindert sich dieser Wert im zweiten Leadjahr auf $r=-0,022$, d. h. dem 21., 22. und 23. Monat nach der Initialisierung (s. Abbildung 5.19b). Für das zweite Leadjahr kann also quasi kein

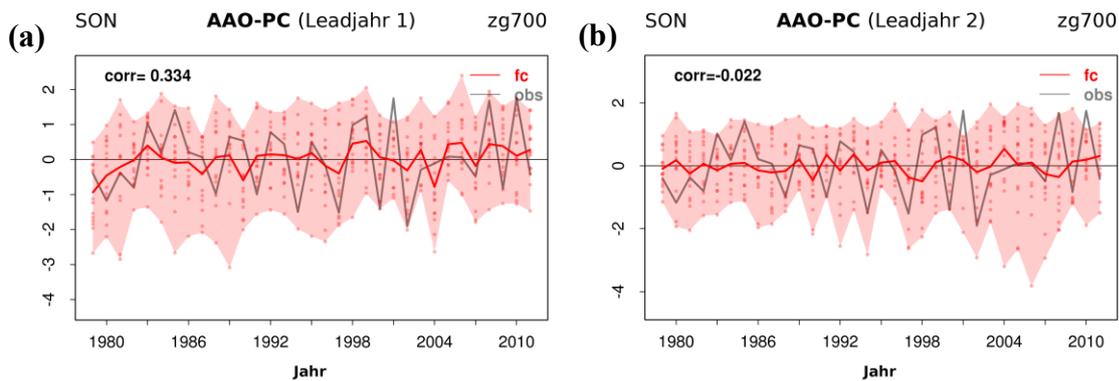


Abbildung 5.19: Saisonale aggregierter AAO-PC im Herbst für das erste (a) bzw. das zweite Leadjahr (b) für die Beobachtung (obs: ERAINT) und die gemittelte Vorhersage (fc: pS) unter Angabe des Korrelationskoeffizienten. Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage.

Zusammenhang zwischen Vorhersage und Beobachtung ermittelt werden. Einher geht diese Verschlechterung in der Vorhersageleistung mit einer Veränderung im Trend-Rausch-Verhältnis: Während für das erste Leadjahr ein Wert von $\text{SNR}=0,423$ zu verzeichnen ist, reduziert sich dieser Wert auf $\text{SNR}=0,301$ im zweiten Leadjahr. Zudem liegen die beobachteten Werte, obwohl sich das Rauschen erhöht, im zweiten Leadjahr häufiger außerhalb des gefärbten Vorhersagebereichs, der durch alle Ensemblemitglieder aufgeworfen wird.

Wie mit dem Hintergrundwissen aus Kapitel 5.2 zum Einfluss der Initialisierung zu vermuten ist, lässt sich diese Beobachtung auch auf die anderen Telekonnektionsindizes in graduellen Abweichungen übertragen. Geringfügige Abweichungen treten beispielsweise beim SOI dergestalt auf, dass im Frühjahr (MAM) des zweiten Leadjahres, d. h. dem 15., 16. und 17. Monat nach Initialisierung, ein Korrelationskoeffizient von $r=0,621$ beobachtet werden kann, was nur eine geringe Abnahme von $r=0,782$ im ersten Leadjahr (s. Abbildung 5.18) und einen Wert deutlich über 0 bedeutet, obwohl das Trend-Rausch-Verhältnis sich auf einen Wert von $\text{SNR}=0,421$ verringert (Leadjahr 1: $\text{SNR}=0,812$; s. o.). Für das dritte Leadjahr ergibt sich allerdings für den SOI im Frühjahr, d. h. dem 27., 28. und 29. Monat nach Initialisierung, ein Bild, das bei anderen Indizes bereits häufig im zweiten Leadjahr beobachtet werden kann: Der Korrelationskoeffizient nimmt einen Wert von $r=0,026$ an und es kann kein Zusammenhang mehr zwischen Vorhersage und Beobachtung festgestellt werden. Zudem verringert sich der Wert für das Trend-Rausch-Verhältnis auf $\text{SNR}=0,398$.

Ähnliches lässt sich für den SOI auch im Herbst (SON) bzw. im Winter (DJF) des zweiten Leadjahres beobachten, wo sich die Werte des Korrelationskoeffizienten bei $r=0,155$ bzw. $r=0,393$ bewegen. Lediglich im nordhemisphärischen Sommer (JJA), der im ersten Leadjahr bereits den geringsten Korrelationskoeffizienten ($r=0,350$) aufweist, liegt der Wert mit $r=-0,023$ nahe 0. Daraus lässt sich schließen, dass scheinbar für einzelne Indizes in bestimmten Jahreszeiten ein Potenzial an dekadischer Vorhersagbarkeit vorhanden ist, wenngleich andere Telekonnektionsindizes, wie z. B. der für Mitteleuropa v. a. im Winter (DJF) relevante NAO-Index (NAO-PC), dieses Ergebnis nicht offenbart, da der hohe Wert des Korrelationskoeffizienten im ersten Leadjahr ($r=0,409$; vgl. Tabelle 5.2) im zweiten Leadjahr ($r=0,025$) nicht beobachtet werden kann.

Zusammenfassend lässt sich konstatieren, dass die beschriebenen, repräsentativen Beispiele eine ausgeprägte Heterogenität in der Vorhersageleistung erkennen lassen. Trotz dieser Heterogenität sollen im folgenden Kapitel anhand eines Vergleichs verschiedener Modellgenerationen in einem überblickenden Ansatz die Fragen aufgeworfen werden, ob generell für einzelne Indizes Verbesserungen in der Vorhersageleistung beim Vergleich der verschiedenen Modellgenerationen ermittelt werden können und wo diese gegebenenfalls herrühren.

5.4.2 Überblickender Vergleich der Modellgenerationen

In den drei folgenden Grafiken (Abbildung 5.20 bis Abbildung 5.22) werden in einer überblickenden Gesamtschau für alle an den Analysen beteiligten Telekonnektionsindizes jeweils die Werte für MSSS bzw. PRSS für die unterschiedlichen Leadjahre (1, 2, 2-5, 6-9, 2-9) und die unterschiedlichen Jahreszeiten (MAM, JJA, SON, DJF) in saisonaler Aggregation präsentiert, um Aussagen über mögliche Verbesserungen in der Vorhersageleistung gegenüber einer Referenz treffen zu können. Als Referenz dienen in Abbildung 5.20 die historischen, uninitialisierten Modellläufe (**hist**), d. h. die Vorhersageleistung der übrigen Modellgeneration (**b0**, **b1**, **pS**, **pG**) wird mit der der historischen Modellläufe verglichen. In Abbildung 5.21 bildet die erste Modellgeneration (**b0**) des dekadischen Modells die Referenz, in Abbildung 5.22 die zweite Modellgeneration (**b1**), wobei darauf zu achten ist, dass im letzten Analyseblock der Grafik die beiden Varianten der dritten Modellgeneration, die mit unterschiedlichen Datengrundlage initialisiert werden, miteinander verglichen werden (**pG–pS**).

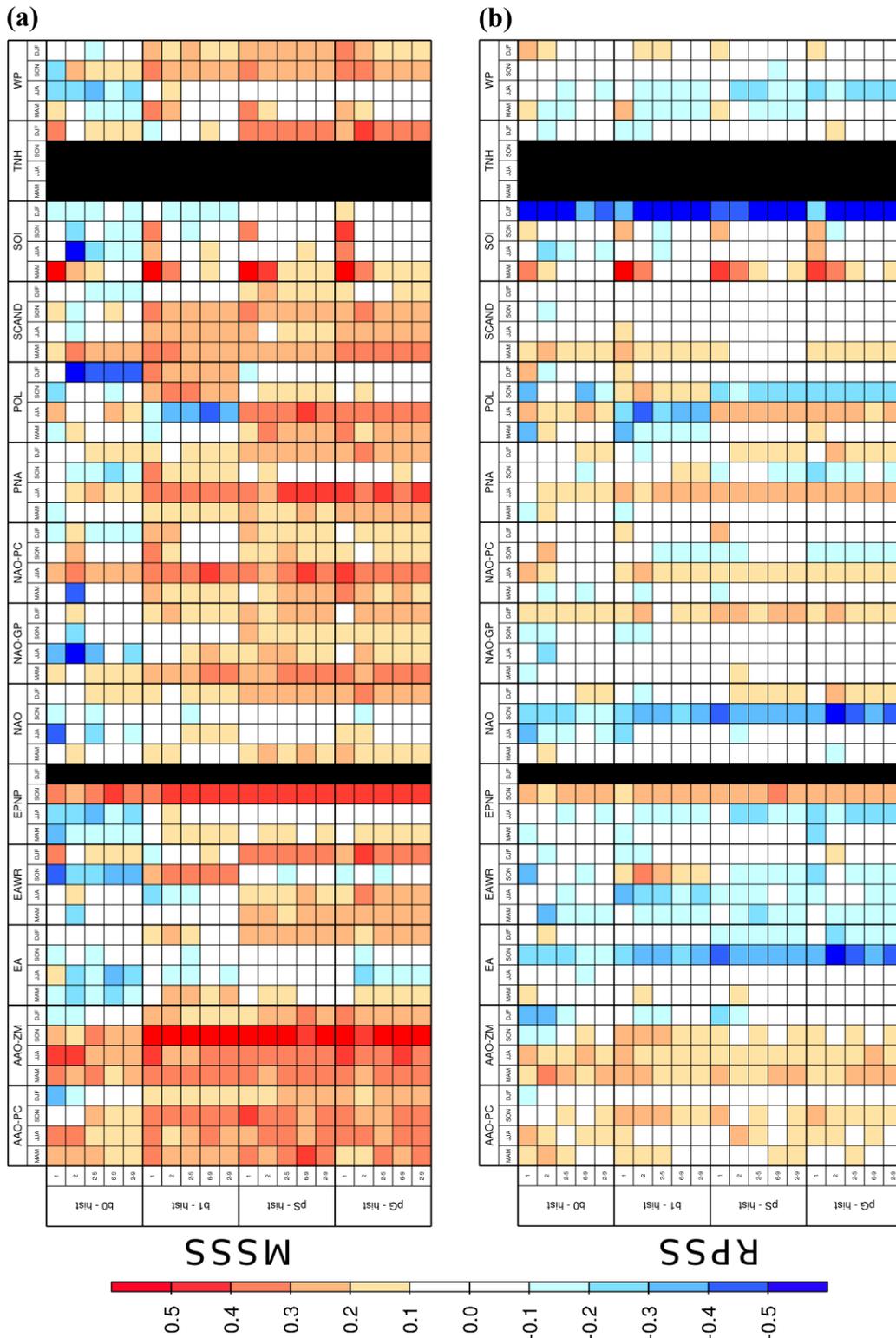


Abbildung 5.20: MSSS (a) und RPSS (b) jeweils für (von oben nach unten) *b0 – hist*, *b1 – hist*, *pS – hist* und *pG – hist* für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse der saisonal aggregierten Indizes (Abkürzungen s. Tabelle 4.1) sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.

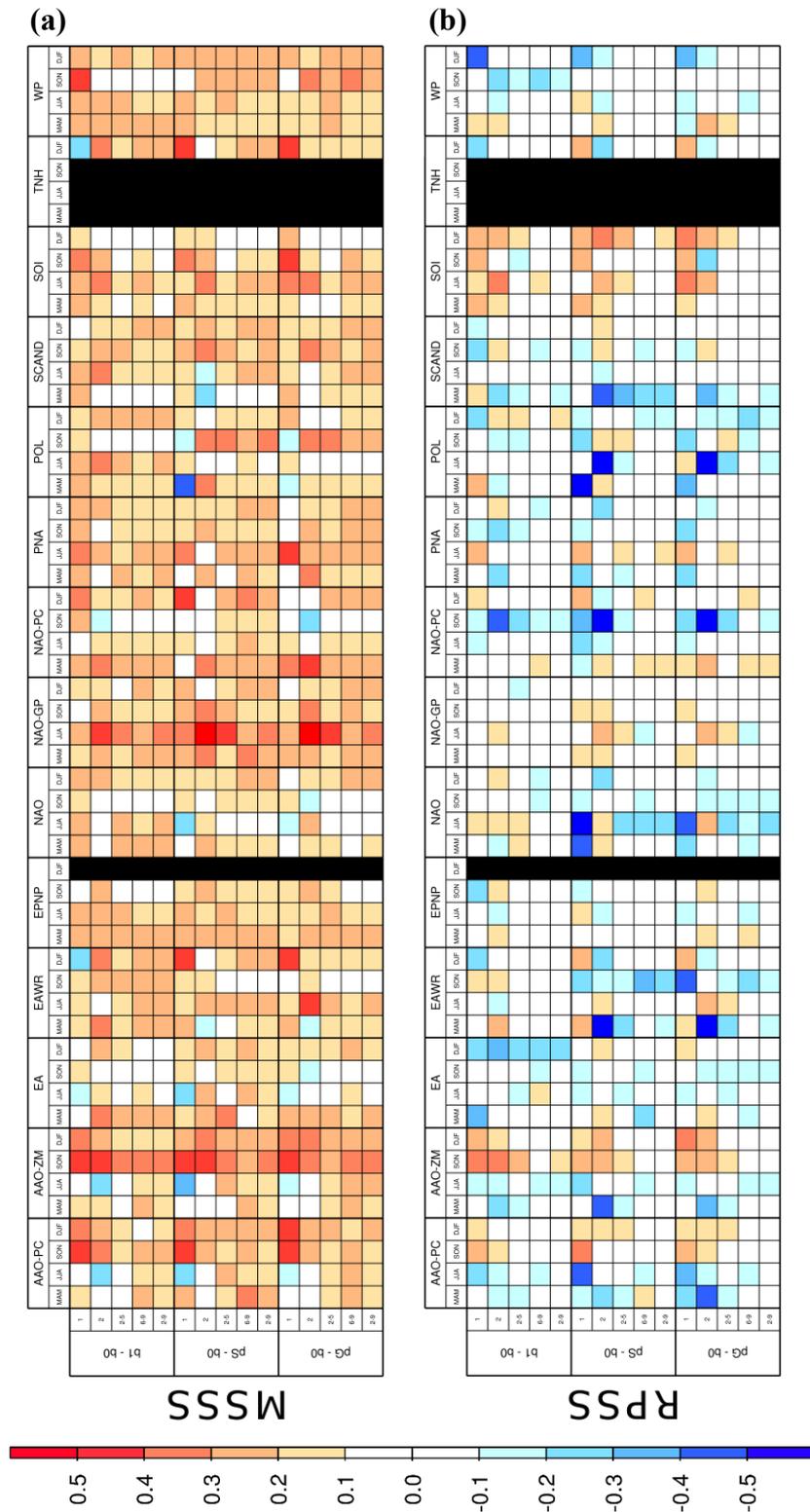


Abbildung 5.21: MSSS (a) und RPSS (b) jeweils für (von oben nach unten) $b1 - b0$, $pS - b0$ und $pG - b0$ für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse der saisonal aggregierten Indizes (Abkürzungen s. Tabelle 4.1) sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.

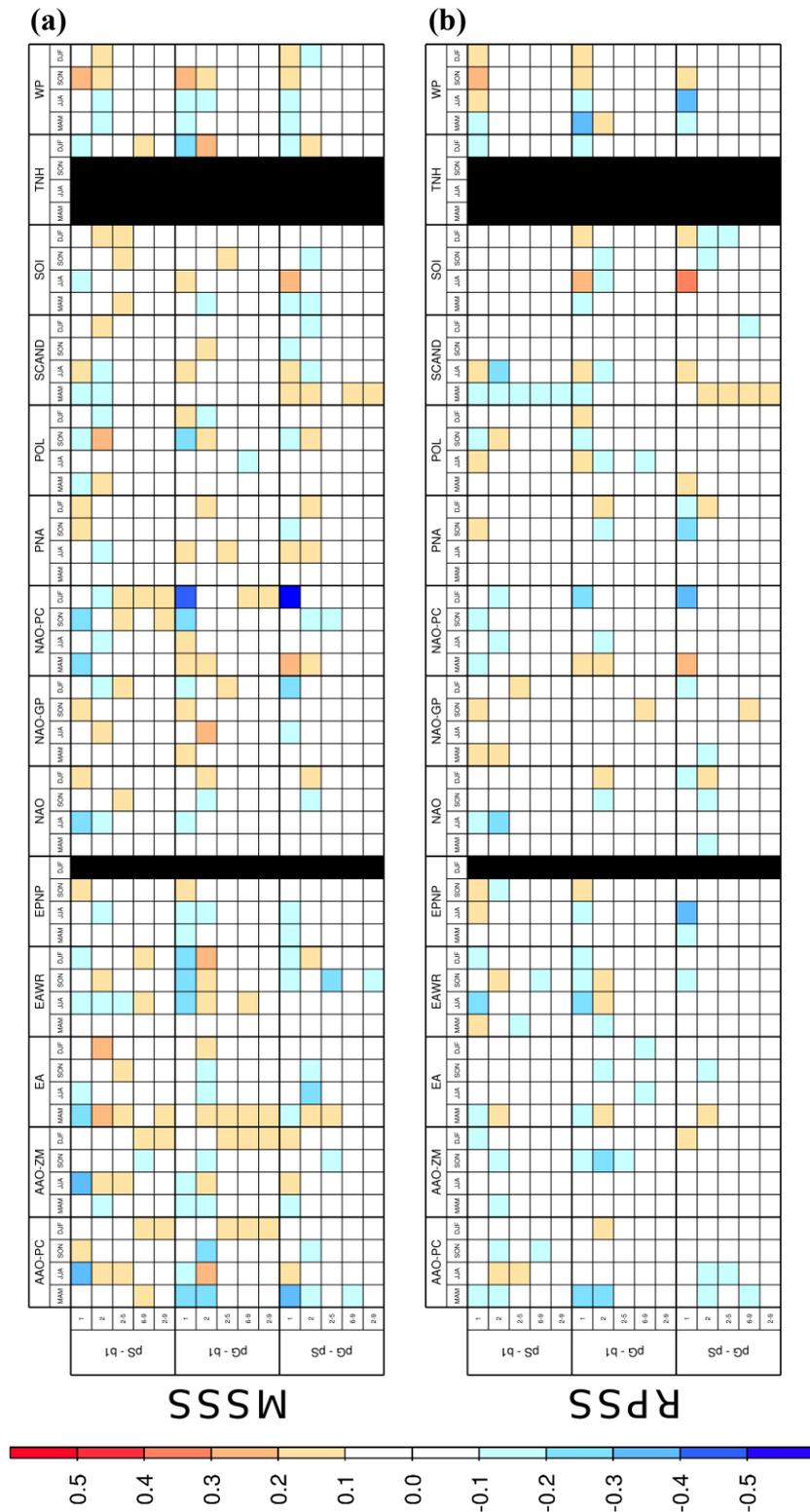


Abbildung 5.22: MSSS (a) und RPSS (b) jeweils für (von oben nach unten) pS – b1, pG – b1 und pG – pS für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse der saisonal aggregierten Indizes (Abkürzungen s. Tabelle 4.1) sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.

Um überblickend über alle Telekonnektionsindizes die Frage zu beantworten, ob die initialisierten Läufe bessere Ergebnisse gegenüber den uninitialisierten liefern, werden in Abbildung 5.20 die initialisierten, saisonal aggregierten Vorhersagen der verschiedenen Modellgenerationen des dekadischen Vorhersagemodells (**b0**, **b1**, **pS**, **pG**) einer uninitialisierten Referenz, den historischen Läufen (**hist**), gegenübergestellt. Wie für den MSSS (s. Abbildung 5.20 a) erkennbar ist, treten Verbesserungen in der Vorhersage, d. h. Werte größer als 0, verbreitet für die zweite (**b1**) und dritte Modellgeneration (**pS**, **pG**) auf. Die besseren Repräsentationen der Indizes gegenüber den uninitialisierten Läufen sind dabei nicht an einzelne Jahreszeiten, Leadjahre oder Telekonnektionsindizes gebunden, sondern treten nahezu durchgängig auf. Deutlich erkennbar sind die überdurchschnittlich starken Verbesserungen durch die Initialisierung des Modells bei den auf unterschiedlichen Berechnungsverfahren beruhenden Varianten der Antarktischen Oszillation (AAO-PC, AAO-ZM) und mit Abstrichen bei SCAND, PNA sowie den beiden Varianten des Index der Nordatlantischen Oszillation, die nach der Methode von Hurrell berechnet wurden (NAO-GP, NAO-PC). Auffallend ist an dieser Stelle zudem, dass die Variante für den Index der Nordatlantischen Oszillation, die nach der Methode des CPC berechnet wurde (NAO) und demnach die gesamte Nordhemisphäre einbezieht, diese Verbesserungen in der Vorhersageleistung nicht zeigt. Außerdem treten beim SOI bei allen Modellen mit die höchsten positiven Werte auf, insbesondere im Frühjahr (MAM) des ersten Leadjahres. Demgegenüber zeigt der Vergleich der ersten Modellgeneration (**b0**) mit den historischen Läufen keine eindeutige Tendenz und es treten sowohl Werte über als auch unter 0 auf. Lediglich bei den bereits erwähnten Indizes AAO-PC und AAO-ZM sowie mit deutlichen Abstrichen bei NAO-PC und SCAND können verbreitet Verbesserungen auch für **b0** beobachtet werden.

Betrachtet man in Abbildung 5.20 anstatt des deterministischen Gütemaßes MSSS das probabilistische Gütemaß RPSS (s. Abbildung 5.20 b), ist ein anderes Bild zu erkennen. Im Gegensatz zum MSSS dominieren nicht die roten Farbtöne, die auf eine Verbesserung des dekadischen Vorhersagemodells gegenüber den historischen Läufen hindeuten, vielmehr halten sich Blau- und Rottöne in etwa die Waage. Lediglich bei AAO-PC, AAO-ZM sowie wieder mit Abstrichen bei NAO-GP, NAO-PC, PNA und SCAND treten nahezu durchgängig Verbesserungen auf. Allerdings sind auch deutliche Verschlechterungen zu erkennen, wie z. B. bei EA, EAWR oder NAO. Wie bereits erwähnt, wird bei der

Berechnung des RPSS der Einfluss der Ensemblemitgliederanzahl eliminiert (vgl. Kapitel 4.2.2.3), was beim MSSS aufgrund der Mittelung über alle Ensemblemitglieder nicht möglich ist. Somit kann aus den Beobachtungen geschlussfolgert werden, dass die beim MSSS beobachteten verbreiteten Verbesserungen vorwiegend auf die Erhöhung der Ensemblemitgliederanzahl – von drei Ensemblemitgliedern bei **hist** und **b0** auf 10 bei **b1** bzw. 15 bei **pS** und **pG** (vgl. Tabelle 3.1) – und nicht auf die Initialisierung des dekadischen Vorhersagesystems zurückzuführen ist. Auch wenn der Einfluss der Initialisierung auf die Verbesserung in der Vorhersageleistung scheinbar eine untergeordnete Rolle einnimmt, kann er dennoch bei einzelnen Telekonnektionsindizes angenommen werden. Besonders auffallend lässt sich dies bei SOI beobachten. Betrachtet man die Werte des RPSS (s. Abbildung 5.20b), lässt sich im Frühjahr des ersten (und teilweise zweiten) Leadjahres eine deutliche Verbesserung mit Werten um bzw. über 0,5 feststellen. Außerdem können – wie bereits erwähnt – Werte über 0 auch für AAO-PC und AAO-ZM konstatiert werden.

Wenn also, wie beschrieben, eine Verbesserung in der Vorhersageleistung durch die Initialisierung des Modells angenommen werden kann, stellt sich die Frage, welche Initialisierungstechnik die besten Resultate liefert. Um dieser Frage angemessen nachzugehen, wurde in einem ersten Schritt untersucht, ob die „full-field“-Initialisierung Vorteile gegenüber der „anomaly“-Initialisierungstechnik aufweist. In Abbildung 5.22 werden die „full-field“-initialisierten Läufe (**pS** / **pG** mit „full-field“-Initialisierung von Atmosphäre und Ozean, **b1** mit „full-field“-Initialisierung der Atmosphäre, vgl. Tabelle 3.1) den „anomaly“-initialisierten der ersten Modellgeneration (**b0**) gegenübergestellt. Die Grafik weist denselben Aufbau wie Abbildung 5.20 auf, d. h. es werden die Werte sowohl für den MSSS (Abbildung 5.21a) als auch den RPSS (Abbildung 5.21b) präsentiert.

Die Werte für den MSSS zeigen wiederum nahezu durchgängig für alle Telekonnektionsindizes, Leadjahre und Jahreszeiten rote Farbtöne, was auf eine Verbesserung in der Vorhersageleistung von **b1**, **pS** sowie **pG** jeweils gegenüber **b0** hindeutet. Lediglich vereinzelt sind blaue Farbtöne, d. h. Verschlechterungen zu erkennen. Um abschätzen zu können, inwiefern die gegenüber der Referenz **b0** geänderte Initialisierungstechnik einen Einfluss darauf hat, können wieder die Resultate des RPSS, die den Einfluss der Vergrößerung des Ensembles eliminieren, hinzugezogen werden (s. Abbildung 5.21b). Es ist deutlich zu erkennen, dass nun nicht die Rottöne überwiegen, sondern in etwa gleichermaßen Rot- wie Blautöne auftreten. Bei manchen Indizes

(NAO, NAO-PC oder POL) überwiegen sogar die blauen Farbtöne, d. h. sie lassen sogar eine deutliche Verschlechterung in der Vorhersageleistung erkennen, insbesondere wenn man die dritte Modellgeneration (**pS**, **pG**) mit **b0** vergleicht. Lediglich SOI weist wiederum überwiegend positive Werte (v. a. im ersten Leadjahr) auf und lässt somit auf den Einfluss der „full-field“-Initialisierungstechnik bezüglich der Verbesserung in der Vorhersageleistung schließen. Zudem offenbart AAO-ZM im Herbst (SON) bzw. dem nordhemisphärischen Winter (DJF) der ersten beiden Leadjahre teilweise positive Werte. Letztendlich kann somit geschlussfolgert werden, dass die Verbesserungen in der Vorhersage von Telekonnektionsindizes größtenteils nicht in der Umstellung von einer „anomaly“-Initialisierung auf eine „full-field“-Initialisierung zu finden sind, sondern vorwiegend auf die Erhöhung der Anzahl an Ensemblemitgliedern zurückgeführt werden können.

Der anschließende dritte Analyseschritt bestand darin, zu untersuchen, welche „full-field“-Initialisierung die besten Ergebnisse liefert, wenngleich der Einfluss auf die Vorhersage von Telekonnektionsindizes nach den bisher erläuterten Ergebnissen gering scheint. Zu diesem Zweck wurde Abbildung 5.22, die denselben Aufbau wie die beiden vorherigen Grafiken aufweist, erstellt. Allerdings lässt die Grafik keine eindeutige Tendenz erkennen, welche „full-field“-Initialisierung die besten Ergebnisse bei der Vorhersage von Telekonnektionsindizes liefert. Sowohl beim Vergleich der dritten mit der zweiten Modellgeneration (**pS** – **b1**, **pG** – **b1**) als auch beim Vergleich zwischen GECCO2 und ORAS4 der dritten Generation (**pG** – **pS**) treten für den MSSS (Abbildung 5.22a) verbreitet Werte um 0 (gekennzeichnet durch weiße Farbgebung) auf. Zwar lassen sich in der Vorhersageleistung mitunter leichte Verbesserungen bei der dritten (**pS** / **pG**) gegenüber der zweiten Modellgeneration (**b1**) erkennen, allerdings sind diese nicht so deutlich und durchgängig wie in den vorherigen Fällen. Auch beim RPSS (Abbildung 5.22b) treten vorwiegend Werte um 0 auf. Somit ist im Gegensatz zu den bisherigen Erkenntnissen auch kein markanter Einfluss der Ensemblegröße (zehn Ensemblemitglieder bei **b1** gegenüber 15 Ensemblemitgliedern bei **pG** / **pS**) erkennbar, woraus geschlossen werden kann, dass die Erhöhung von drei (**b0**) auf zehn bzw. 15 Ensemblemitglieder eine größere Auswirkung auf die Vorhersageleistung hat als die Erhöhung von zehn auf 15 Ensemblemitglieder.

5.5 Zusammenfassung

Zusammenfassend lässt sich festhalten, dass bezüglich der dekadischen Vorhersage von Telekonnektionsindizes durch das MPI-ESM-Vorhersagesystem eine ausgeprägte Heterogenität beobachtet werden kann. Verbesserungen in der Vorhersageleistung durch die Initialisierung des Modells und die Vergrößerung des Ensembles können, wenn überhaupt, vorwiegend nur für das erste Leadjahr beobachtet werden. Während bestimmte Telekonnektionsindizes diese Verbesserung in der Vorhersageleistung nicht zeigen, wobei v. a. der für das mitteleuropäische Wettergeschehen relevante Index der Nordatlantischen Oszillation (NAO) genannt werden muss, weisen einige Indizes positive Ergebnisse auf. An dieser Stelle kann insbesondere SOI hervorgehoben werden, was im Einklang mit anderen Ergebnissen aus dem Forschungsgebiet der saisonalen Vorhersage steht (z. B. Smith 2012). Die Überblicksdarstellung für den SOI im ersten Leadjahr (Anhang 5), in welcher die Indizes für die unterschiedlichen Jahreszeiten und Modellgenerationen unter Angabe des Korrelationskoeffizienten dargestellt und Verbesserungen in der Vorhersageleistung belegt sind, untermauern die Sonderrolle.

Zieht man den RPSS als Analysewerkzeug hinzu, lassen sich Verbesserungen in der Vorhersageleistung vorwiegend auf eine Erhöhung der Ensemblemitgliederanzahl und nur vereinzelt auf veränderte Initialisierungsverfahren (wie z. B. beim SOI im ersten Leadjahr bzw. mit deutlichen Abstrichen bei AAO-ZM und AAO-PC in den ersten beiden Leadjahren) zurückführen. Anders ausgedrückt bedeutet dieser Befund, dass es für die Vorhersage der betrachteten Telekonnektionsindizes keinen entscheidenden Unterschied ausmacht, ob man das Modell mithilfe der „full-field“ oder der „anomaly“-Initialisierungstechnik startet. An dieser Stelle muss nochmals daran erinnert werden, dass sich die Änderungen in der Initialisierungstechnik von **b1** zu **pS/pG** lediglich auf den Ozean beschränken (vgl. Tabelle 3.1).

Trotz der großen Heterogenität, die für das erste bzw. für die ersten beiden Leadjahre beobachtet werden kann, zeichnen sich alle Telekonnektionsindizes durch eine Gemeinsamkeit aus: Für die restlichen Leadjahre, d. h. die Leadjahre 3-9, kann (wie v. a. aus den Grafiken in Kapitel 5.2 hervorgeht) keine nennenswerte Vorhersageleistung beobachtet werden. Deshalb kann dem Modell der Anspruch der dekadischen Vorhersage nicht zugesprochen werden.

6 Ergebnisse:

Zirkulationstypen

Um die Analysen der Repräsentation der atmosphärischen Dynamik in dekadischen Klimamodellen mithilfe von Zirkulationstypen nicht ins Uferlose laufen zu lassen und um eine Übersichtlichkeit der Analysen und der Ergebnisse zu bekommen, mussten Entscheidungen auf subjektiver Basis getroffen werden. Neben der räumlichen Beschränkung auf den nordatlantisch-europäischen Raum als Untersuchungsgebiet betrifft dies v. a. die Anzahl der Klassen, die – wie bereits in Kapitel 4.1.2 erläutert wurde – standardmäßig auf acht Klassen festgelegt wurde. Zwar wäre es auch möglich gewesen, für jedes Verfahren und jede Jahreszeit eine nach einem bestimmten Kriterium ermittelte optimale Anzahl an Klassen anzuwenden. Allerdings muss bei der Analyse zur dekadischen Vorhersagbarkeit der atmosphärischen Dynamik die Vergleichbarkeit über die unterschiedlichen Verfahren und Jahreszeiten hinweg im Vordergrund stehen, um Aussagen über Verbesserungen in der Vorhersage z. B. durch die Initialisierung des Modells oder die Vergrößerung der Ensemblemitgliederanzahl treffen zu können. Dies wäre bei einem Vorgehen mit vielen unterschiedlichen Klassenanzahlen nicht möglich bzw. würde die Analyse unnötig erschweren. Die Anzahl von acht Klassen wurde ausgehend von den Großwettertypen nach Beck (GWT) getroffen, bei welcher die acht Klassen die acht Hauptwindsektoren (N, NE, E, SE, S, SW, W, NW) repräsentieren, und als Standard festgelegt, da auch bei den beiden anderen Verfahren (SAN, TPCA) sinnvolle Ergebnisse mit

interpretierbaren Mustern resultieren (vgl. Kapitel 4.1.2). Um den Einfluss der Klassenanzahl im Zusammenhang mit der dekadischen Vorhersagbarkeit zu untersuchen (s. Kapitel 6.3.4), wurde bei einzelnen Untersuchungen zudem auf eine größere Anzahl (18 Klassen) und eine kleinere Anzahl an Klassen (vier Klassen) zurückgegriffen, was wiederum subjektiv festgelegt wurde.

Ein weiterer subjektiver Eingriff liegt in der Entscheidung, auf welcher Datengrundlage die Ermittlung der Zirkulationstypen erfolgt. Grundsätzlich kommen dazu der Luftdruck auf Meereshöhe (SLP) sowie die geopotentielle Höhe (ZG) in Frage, die auch beide in der vorliegenden Arbeit zur Anwendung kamen. Für ZG standen für das dekadische Vorhersagemodell mehrere Höhenniveaus zur Auswahl. Da keine vergleichbaren Vorgaben wie bei den einzelnen Telekonnektionsindizes existieren, bedurfte es wiederum einer subjektiven Entscheidung. In der vorliegenden Arbeit wurde auf das 500hPa-Geopotential zurückgegriffen (ZG500), das in anderen Untersuchungen (z. B. Jacobeit 2003b) bereits verwendet wurde.

Um die Repräsentation der dynamischen Zirkulation der Atmosphäre durch das dekadische Vorhersagemodell einschätzen zu können, stehen mehrere Parameter zur Verfügung. Zum einen wurden in der vorliegenden Arbeit die **Auftrittshäufigkeiten** der Zirkulationstypen analysiert. D. h. eine dekadische Vorhersage ist umso besser, je näher die Häufigkeitsverteilung der Vorhersage an der Beobachtung liegt. Neben der Häufigkeitsverteilung wurden in der vorliegenden Arbeit zum anderen typinterne Charakteristika der einzelnen Zirkulationstypen hinsichtlich ihrer dekadischen Vorhersagbarkeit untersucht. Dazu wurde für jeden Zirkulationstyp aus allen zugewiesenen Einzeltagen Zentroide, d. h. Mittelungen über alle Einzeltage, berechnet. Um die vorhergesagten Zentroide mit den beobachteten hinsichtlich ihrer Ähnlichkeit vergleichen zu können, wurde die **Intensität** der Zentroide, die sich aus der Differenz zwischen maximalem Wert und minimalem Wert (von SLP bzw. ZG500) ergibt, als typinternes Charakteristikum analysiert. Zu guter Letzt wurde die **Persistenz** jedes Zirkulationstyps, d. h. das durchschnittliche Andauern in aufeinanderfolgenden Tagen, ermittelt und als weiteres typinternes Charakteristikum hinsichtlich seiner Repräsentation untersucht.

Zur Untersuchung der atmosphärischen Zirkulation wurde, wie bereits erwähnt, jeder einzelne Tag des Untersuchungszeitraums einem vordefinierten Zirkulationstyp zugewiesen.

Ein Ansatz der Analysen bestand nun darin, die Auftrittshäufigkeiten bzw. typinternen Charakteristika für Beobachtung und Vorhersage über den gesamten Untersuchungszeitraum aufzusummieren und zu mitteln (Methode „Untersuchungszeitraum“). Dieses Vorgehen steht im Vordergrund der vorliegenden Arbeit. Da bei diesem Ansatz durch die Mittelung über den gesamten Untersuchungszeitraum die zeitliche Information verloren geht, bestand ein anderer Ansatz darin, die Auftrittshäufigkeiten bzw. typinternen Charakteristika nicht über den gesamten Untersuchungszeitraum, sondern saisonal aggregiert zu untersuchen (Methode „Saisonal“). In Abbildung 3.4 wurde bereits visuell veranschaulicht, dass das dekadische Vorhersagemodell dazu nach optischem Ermessen scheinbar nicht in der Lage ist. Zudem besteht ein Problem darin, dass v. a. bei hoher Klassenanzahl viele Klassen unbesetzt sind, was zwar für die Untersuchungen bezüglich der Auftrittshäufigkeiten kein Problem darstellt, dafür aber umso mehr bei der Berechnung typinterner Charakteristika, da unbesetzte Klassen nicht in die Berechnung einbezogen werden können. Auch wenn das Vorgehen problematisch erscheint, führte es dennoch zu aussagekräftigen Ergebnissen und wurde bei einzelnen Analyseschritten angewendet, worauf im Folgenden eingegangen wird.

Die Präsentation der Ergebnisse orientiert sich an derjenigen zu den Telekonnektionsindizes. Zunächst wird in Kapitel 6.1 die Vorhersageleistung der uninitialisierten, historischen Läufe präsentiert und anschließend der Einfluss der Initialisierung des dekadischen Vorhersagesystems (Kapitel 6.2) sowie der Ensemblegröße (Kapitel 6.3) auf die Vorhersageleistung untersucht. Nach Untersuchung der einzelnen Einflussgrößen erfolgt in Kapitel 6.4 wiederum eine überblickende Darstellung zur Vorhersageleistung des dekadischen Modells hinsichtlich der Dynamik der Atmosphäre. Abschließend werden die Resultate in Kapitel 6.5 zusammengefasst.

6.1 Die uninitialisierten Läufe

Vergleichbar zum vorherigen Kapitel zu den Telekonnektionsindizes sollen auch in diesem Kapitel zunächst die Ergebnisse für die uninitialisierten, historischen Modellläufe gemittelt über den gesamten Untersuchungszeitraum 1979-2011 gezeigt werden, um anschließend mögliche Verbesserungen einfacher beurteilen zu können, die durch die Initialisierung des Modells (Kapitel 6.2) bzw. die Erhöhung der Ensemblemitgliederanzahl (Kapitel 6.3) hervorgerufen werden.

6.1.1 Luftdruck auf Meeresspiegel (SLP)

In Abbildung 6.1 sind nach dem Verfahren der Großwettertypen nach Beck (GWT) mit acht Zirkulationstypen die Ergebnisse für die Auftrittshäufigkeit (Abbildung 6.1a) sowie für die zwei typinternen Charakteristika der Intensität (Abbildung 6.1b) und Persistenz (Abbildung 6.1c) basierend auf dem Luftdruck auf Meereshöhe (SLP) dargestellt. Alle Werte sind Mittelungen über alle Jahreszeiten, repräsentieren also das gesamte Jahr und wurden, wie bereits erläutert wurde, über den gesamten Untersuchungszeitraum gemittelt. Die Betrachtungsweise für das gesamte Jahr ist nur bei der GWT-Methode möglich, da es sich um vordefinierte Zirkulationsmuster handelt, denen die einzelnen Tage zugeordnet werden (vgl. Kapitel 4.1.2.1).

Die Auftrittshäufigkeiten in Abbildung 6.1 zeigen, dass zwischen ERAINT als Beobachtung („obs“) und den uninitialisierten, historischen Modellläufen („hist“) generell lediglich geringe Abweichungen zu erkennen sind. So stellt beispielsweise der Zirkulationstyp „W“, der eine westliche Anströmrichtung aufweist, den über das gesamte Jahr gesehen am häufigsten auftretenden Zirkulationstyp sowohl bei der Beobachtung als auch dem uninitialisierten Modell dar. Zudem tritt der Zirkulationstyp „SE“ in beiden Fällen am seltensten auf. Auch die übrigen Klassen werden durch das historische Modell gut repräsentiert. Auffallend ist jedoch, dass zonale Wetterlagen (insbesondere der Zirkulationstyp „W“) häufiger als in den beobachteten Werten auftreten, wohingegen eher meridionale Wetterlagen („N“, „NE“, „SE“) unterrepräsentiert sind, wengleich der meridionale Zirkulationstyp „S“ in den Modelldaten häufiger auftritt als in der Beobachtung.

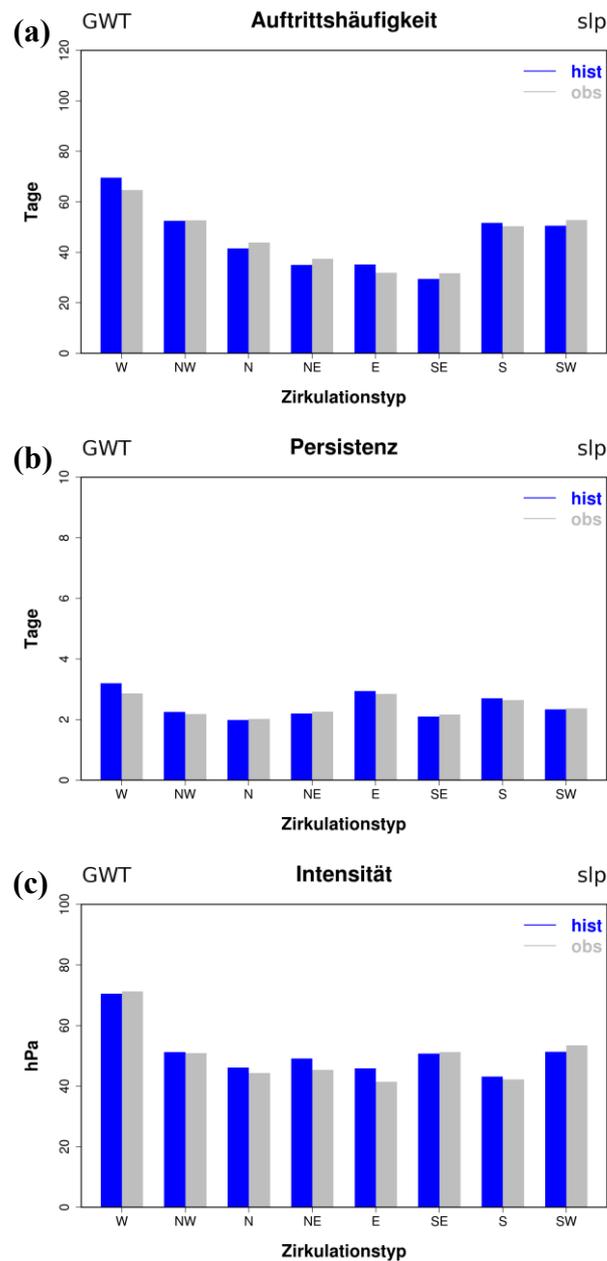


Abbildung 6.1: Aufttrittshäufigkeit (a), Persistenz (b) und Intensität (c) von 8 Zirkulationstypen (nach GWT) im nordatlantisch-europäischen Raum **über alle Jahreszeiten** des Untersuchungszeitraums (1979-2011, gemittelt) basierend auf dem Luftdruck auf Meereshöhe (SLP). Die Ergebnisse sind dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs).

Einher geht diese Beobachtung mit einer typinternen Überbetonung der Persistenz beim Zirkulationstyp „W“ (s. Abbildung 6.1b). D. h., wenn die Klasse mit westlicher Anströmungsrichtung auftritt, hält sie sich im Durchschnitt zu lang, was mit dem insgesamt zu häufigen Auftreten Hand in Hand geht und lediglich bei diesem Zirkulationstyp in allen

Jahreszeiten zu beobachten ist. Überblickt man alle Klassen, gelangt man zu dem Entschluss, dass die historischen, uninitialisierten Modellläufe die Persistenz – über das gesamte Jahr betrachtet – gut repräsentieren und große Abweichungen nicht zu erkennen sind. Tendenziell kann jedoch beobachtet werden, dass die historischen, uninitialisierten Modellläufe die Persistenz der Zirkulationstypen überschätzen: Bei Unterrepräsentation der Persistenz liegen die Werte nur geringfügig unterhalb der beobachteten (z. B. bei den Zirkulationstypen „N“, „NE“ und „SE“), es überwiegt jedoch die Überrepräsentation der Persistenz (z. B. neben dem Zirkulationstyp „W“ bei „NW“, „E“ oder „S“).

Auch bei der Betrachtung eines weiteren typinternen Charakteristikums, der Intensität (s. Abbildung 6.1c) als Luftdruckgegensatz zwischen maximalem und minimalem Druck im Zentroid des Untersuchungsgebiets, kann festgestellt werden, dass das dekadische Vorhersagesystem die Beobachtung bei Mittelung der Daten über den gesamten Untersuchungszeitraum hinweg gut nachbildet. Sowohl der stark ausgeprägte Luftdruckgegensatz beim Zirkulationstyp „W“ als auch die Zirkulationstypen mit einem gering ausgeprägten Luftdruckgegensatz (z. B. Klassen „E“ und „S“) werden durch das Modell gut repräsentiert. Hierbei fällt jedoch auf, dass die historischen, uninitialisierten Modellläufe die Intensität bei den Klassen mit eher gering ausgeprägtem Luftdruckgegensatz überschätzen, d. h. sie fällt zu stark aus.

Um die Erkenntnisse, die für das GWT-Verfahren ermittelt wurden, für die beiden anderen Methoden – die SANDRA-Clusteranalyse (SAN) sowie die t-modale Hauptkomponentenanalyse (TPCA) – überprüfen zu können, muss eine einzelne Jahreszeit herausgegriffen werden. Dies ist notwendig, weil bei SAN bzw. TPCA die Betrachtungsweise über das gesamte Jahr hinweg nicht möglich ist, da die Klassen sich in den einzelnen Jahreszeiten nicht entsprechen. Wie bereits bei GWT erläutert wurde, tritt das Phänomen der Überbetonung von zonalen Wetterlagen in allen Jahreszeiten auf, lässt sich bei GWT aber besonders deutlich im Sommer (JJA) beobachten, weshalb diese Jahreszeit ausgewählt wurde. Diese Tatsache wird in Abbildung 6.2 widergespiegelt, in der die Auftrittshäufigkeiten für die acht Großwettertypen nach Beck im Sommer (JJA) dargestellt sind. Im Vergleich zur Betrachtungsweise über alle Jahre hinweg (s. Abbildung 6.1a) ist noch deutlicher die Überbetonung des Zirkulationstyps mit westlicher Anströmrichtung (Zirkulationstyp „W“) zu erkennen. Das dazugehörige Zentroid in Anhang 6a, welches über

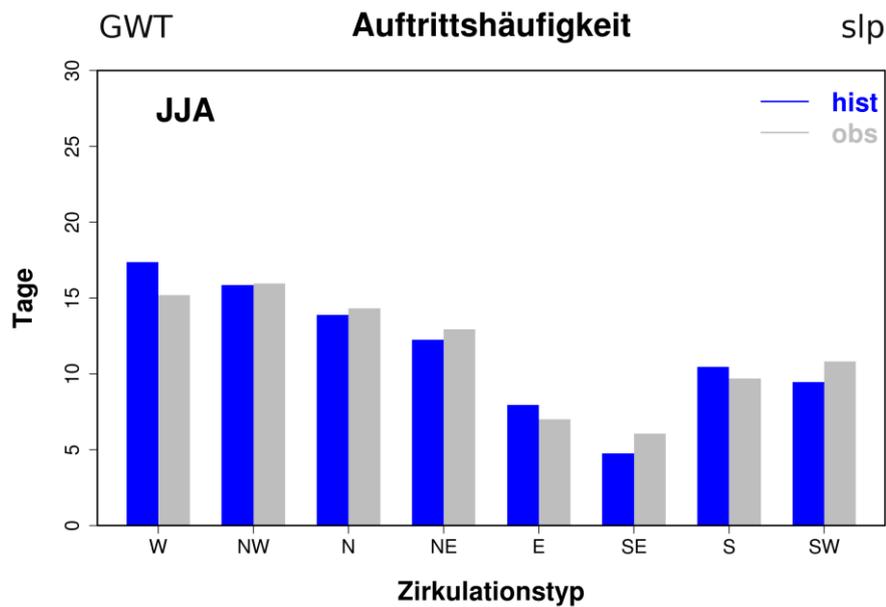


Abbildung 6.2: Auftrittshäufigkeit von 8 Zirkulationstypen (nach GWT) im nordatlantisch-europäischen Raum für **den Sommer (JJA)** des Untersuchungszeitraums (1979-2011, gemittelt) basierend auf dem Luftdruck auf Meereshöhe (SLP). Die Ergebnisse sind dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs).

alle zugewiesenen Tage gemittelt wurde, zeigt tiefen Luftdruck über Island, der sich bis nach Skandinavien erstreckt, und hohen Luftdruck über den Azoren, sodass in Mitteleuropa eine westliche Anströmrichtung resultiert. Zudem tritt der Zirkulationstyp „E“, bei dem sich eine Hochdruckbrücke von den Azoren bis nach Skandinavien erstreckt (vgl. Anhang 6e), in den historischen Modellläufen häufiger auf als in der Beobachtung. Wiederum sind die eher meridional orientierten Zirkulationstypen („N“, „NE“, SE“, „SW“) unterrepräsentiert, wobei der Großwettertyp „S“ aus dieser Beobachtung herausfällt.

In Abbildung 6.3 sind die äquivalenten Auftrittshäufigkeiten der SANDRA-Clusteranalyse (Abbildung 6.3a) und der t-modalen Hauptkomponentenanalyse (Abbildung 6.3b) dargestellt. Wiederum zeigt sich, dass das historische, uninitialisierte Modell das Auftreten der einzelnen Klassen im Allgemeinen gut repräsentiert. Zirkulationstyp 3, der an knapp einem Drittel aller Tage auftritt, sticht in Abbildung 6.3a besonders hervor. Das dazugehörige Zirkulationsmuster (Anhang 7c; „Zirkulationstyp 3“ entspricht „CL-03“) lässt ein ausgeprägtes Gebiet tiefen Luftdrucks rund um Island und ein Gebiet höheren Luftdrucks über den Azoren erkennen. Es handelt sich also um eine Wetterlage, bei der

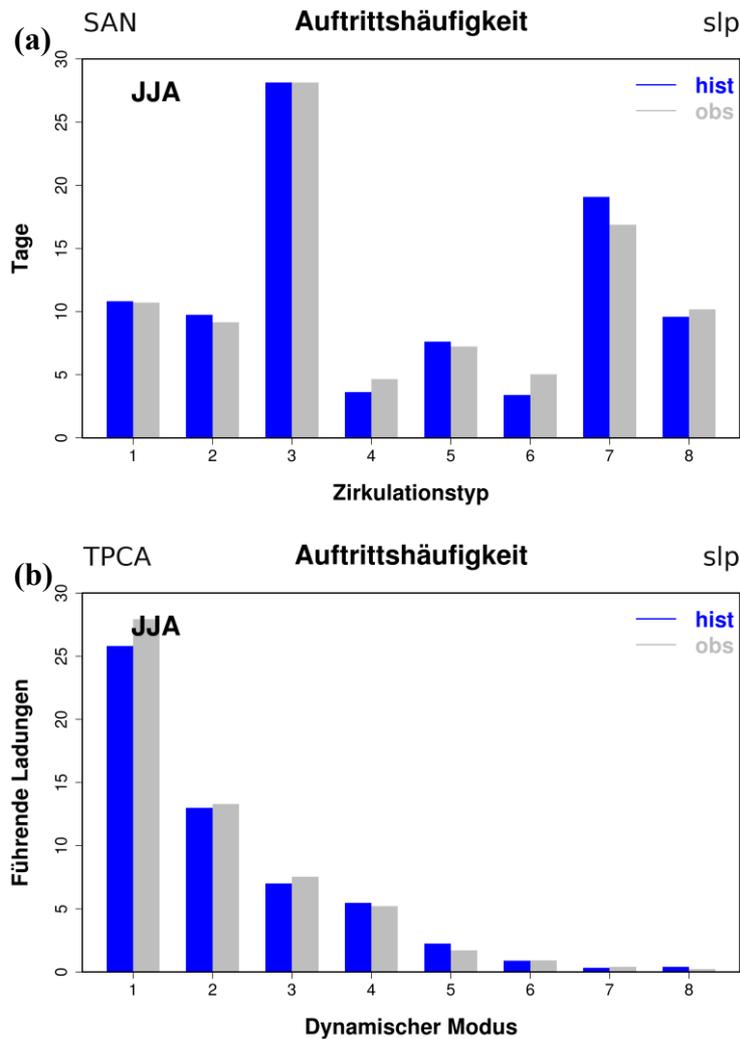


Abbildung 6.3: Wie Abbildung 6.2, aber für SAN (a) bzw. TPCA (b).

Luftmassen aus westlicher Richtung nach Mitteleuropa geführt werden. Eine Überrepräsentation dieses Zirkulationstyps lässt sich aus Abbildung 6.3a allerdings nicht erkennen. Diese offenbart sich allerdings bei den Zirkulationstypen 2, 5 und 7. Während Zirkulationstyp 7 aufgrund der Anordnung der Druckgebiete mit Einfluss von hohem Luftdruck über Mitteleuropa am ehesten dem Großwettertyp „S“ entspricht (vgl. Anhang 7g), der in den historischen Modellläufen ebenfalls zu häufig auftritt, können die Zirkulationstypen 2 (vgl. Anhang 7b) und insbesondere 5 (vgl. Anhang 7e) als Abwandlungen des Großwettertyps „W“ gedeutet werden, da bei beiden Zirkulationsmustern eine zonale Anströmrichtung überwiegt. Demgegenüber werden die eher meridional geprägten Zirkulationstypen, wie z. B. Zirkulationstyp 4 mit einem ausgeprägten Hochdruckrücken über

dem Pazifik (vgl. Anhang 7d), aber auch die Zirkulationstypen 6 und 8 in ihrer Auftrittshäufigkeit durch die historischen Modellläufe nicht adäquat wiedergegeben, sodass der durch die GWT-Methode gewonnene Eindruck gestärkt wird.

Ein anderes Bild zeigt sich, wenn man die Auftrittshäufigkeiten der Zirkulationstypen basierend auf der t-modalen Hauptkomponentenanalyse betrachtet (s. Abbildung 6.3b). Es muss beachtet werden, dass hierbei nicht die über den gesamten Untersuchungszeitraum (1979-2011) zugewiesenen und gemittelten Tage für den Sommer (JJA) dargestellt werden, sondern die gemittelten führenden Ladungen. Die führenden Ladungen drücken, wie bereits in Kapitel 4.1.2.3 beschrieben wurde, den Grad der Ähnlichkeit zu dem beobachteten Prototyp aus und gehen deshalb gewichtet in die Berechnung ein. Einerseits lässt sich erkennen, dass das historische Modell die beobachteten Werte im Allgemeinen gut repräsentiert. Der Zirkulationstyp 1, der am häufigsten auftritt, zeigt mit einem über dem nördlichen Atlantik und einem Hochdrucksystem im Bereich der Azoren (vgl. Anhang 8a; „Zirkulationstyp 1“ entspricht „PC-01“) Tiefdrucksystem eine klassische zonal geprägte Wetterlage. Im Gegensatz zum Klassifikationsverfahren der SANDRA-Clusteranalyse sind deutlich unterscheidbare Zirkulationsmuster zu erkennen. So strömen beispielsweise beim Zirkulationsmuster 2 des entsprechenden Zirkulationstyps mit tiefem Druck über Skandinavien und hohem Druck über den Azoren bzw. Grönland Luftmassen aus nordwestlicher Richtung nach Mitteleuropa (vgl. Anhang 8b). Markant sind auch die Hochdruckbrücken bei den Zirkulationsmustern 4 (vgl. Anhang 8d) und 8 (vgl. Anhang 8h), die sich von den Azoren bis nach Skandinavien erstrecken.

Andererseits lässt sich beobachten, dass die gemittelten Werte der führenden Ladungen für die historischen Modellläufe im Allgemeinen geringer ausfallen als für die Beobachtung. Das lässt sich damit erklären, dass die Modellwerte den mithilfe der Beobachtungsdaten berechneten Prototypen weniger ähnlich sind als die beobachteten Werte selbst. Dass der Zirkulationstyp 1, der die größte Relevanz bezüglich der Zirkulationsdynamik besitzt, auch die stärksten Abweichungen zwischen Modell und Beobachtung aufweist, ist demzufolge logisch, lässt aber keine Aussage über eine mögliche tatsächliche Unterrepräsentation durch die historischen Modellläufe zu.

6.1.2 500 hPa-Geopotentialfläche (ZG500)

In gleicher Weise wie für den Luftdruck auf Meereshöhe (SLP) sind in Abbildung 6.4 die Auftrittshäufigkeit (Abbildung 6.4a), die Persistenz (Abbildung 6.4b) sowie die Intensität (Abbildung 6.4c) für die einzelnen Großwettertypen nach Beck basierend auf der 500 hPa-Geopotentialfläche dargestellt. Betrachtet man die Auftrittshäufigkeiten, fällt

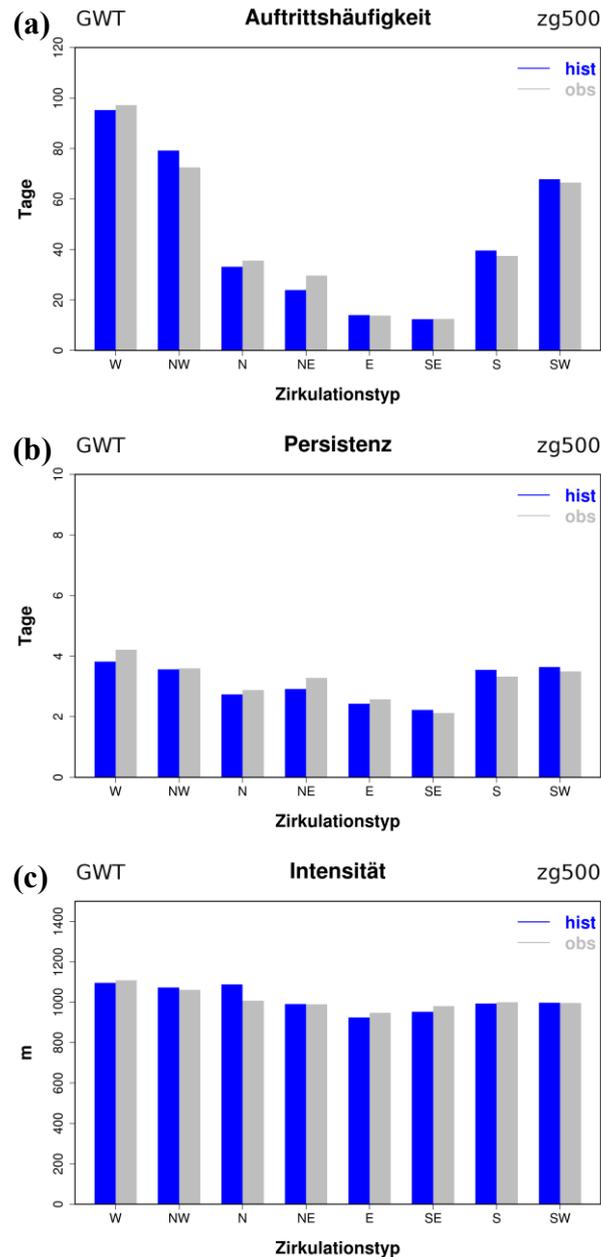


Abbildung 6.4: Auftrittshäufigkeit (a), Persistenz (b) und Intensität (c) von 8 Zirkulationstypen (nach GWT) im nordatlantisch-europäischen Raum **über alle Jahreszeiten** des Untersuchungszeitraums (1979-2011, gemittelt) basierend auf der 500hPa-Geopotentialfläche (ZG500). Die Ergebnisse sind dargestellt für die historischen, uninitialisierten Modellläufe (hist) und die beobachteten Werte (obs).

zunächst auf, dass das historische, uninitialisierte Modell die beobachteten Werte gut wiederzugeben vermag und sowohl die häufig als auch die eher selten auftretenden Klassen gelungen repräsentiert werden. Dasselbe gilt auch für die typinternen Charakteristika der Persistenz bzw. der Intensität. In beiden Fällen sind große Abweichungen von den beobachteten Werten selten (wie z. B. die Abweichungen zwischen Modell und Beobachtung bezüglich der Intensität beim Zirkulationstyp „N“; s. Abbildung 6.4c).

Betrachtet man die einzelnen Parameter im Detail, fällt bei der Auftrittshäufigkeit (s. Abbildung 6.4a) auf, dass der Zirkulationstyp „W“ im Gegensatz zu dem auf dem Luftdruck auf Meereshöhe basierenden Klassifikationsverfahren (vgl. Abbildung 6.1) im Modell nicht häufiger auftritt, sondern sogar eher unterrepräsentiert ist. Betrachtet man die gesamte Häufigkeitsverteilung kann von einer Überbetonung zonaler Wetterlagen, wie sie im vorherigen Kapitel für den Luftdruck auf Meereshöhe (SLP) aufgezeigt werden konnte, nicht gesprochen werden. Zudem weist der Zirkulationstyp „W“ für die historischen Modellläufe keine zu hohen Werte bezüglich der Persistenz wie bei SLP auf. Vielmehr scheint das Modell die Persistenz insgesamt zu unterschätzen, wenn man überblickend alle Zirkulationstypen analysiert.

6.2 Der Einfluss der Initialisierung

Nachdem aufgezeigt wurde, welche Vorhersageleistung das uninitialisierte, historische Modell zeigt, soll nun überprüft werden, ob und inwiefern die Initialisierung des Modells eine Verbesserung hervorruft. Vergleichbar zu den Telekonnektionsindizes soll das Augenmerk nicht auf einzelnen Klassifikationsergebnissen liegen, sondern wiederum eine zusammenfassende Analyse der verschiedenen Modellgenerationen erstellt werden. Als Datenbasis gingen die drei verschiedenen methodischen Ansätze (GWT, SAN, TPCA), alle Jahreszeiten (MAM, JJA, SON, DJF) und die beiden unterschiedlichen Variablen (SLP, ZG500) in die Analyse ein. Die verschiedenen Modellgenerationen wurden jeweils anhand der maximal verfügbaren Anzahl an Ensemblemitgliedern (vgl. Tabelle 3.1) – mit Beschränkung auf die geringere räumliche Auflösung des Klimamodells (MPI-ESM-LR) – in die Analysen aufgenommen.

Im vorherigen Kapitel 6.1 wurden für das historische Modell nur die über den gesamten Untersuchungszeitraum (1979-2011) gemittelten Ergebnisse präsentiert. Bei der Mittelung über den gesamten Untersuchungszeitraum geht allerdings die zeitliche Information verloren. Für eine Vorhersage, insbesondere auf dekadischer Zeitskala, ist die zeitliche Ausprägung bestimmter Parameter allerdings von großem Interesse. Wie in Abbildung 3.4 für die saisonale Auftrittshäufigkeit verschiedener Ensemblemitglieder gezeigt wurde, vermag das dekadische Vorhersagemodell den zeitlichen Verlauf nach visuellem Eindruck nur ungenügend wiederzugeben. Dennoch soll in diesem Kapitel untersucht werden, ob die Initialisierung des Modells eine Verbesserung in der Vorhersage hervorruft, oder ob ein Einfluss tatsächlich nicht erkennbar ist. Um einen zeitlichen Verlauf zu erhalten, wurden die Auftrittshäufigkeiten nicht über den gesamten Untersuchungszeitraum gemittelt, sondern saisonal aggregiert, sodass für jedes Jahr Werte für die acht Klassen vorhanden sind. Eine monatliche anstatt einer saisonalen Aggregation wäre ebenso möglich, ist aufgrund der vielen Fehlwerte – hervorgerufen durch leere Klassen – jedoch nicht sinnvoll.

Als Messinstrument wurde wie bei den Telekonnektionsindizes wieder der Mean Squared Error (MSE) angewendet. Ein Einsatz des Korrelationskoeffizienten ist bei der Mittelung über den gesamten Untersuchungszeitraum mit Verlust der zeitlichen Information nicht möglich. Bei der Betrachtungsweise des zeitlichen Verlaufs ist er außerdem nicht sinnvoll, da keine einzelne Zeitreihe wie bei den Telekonnektionsindizes (vgl. Kapitel 5) vorhanden ist, sondern zu jedem Zeitpunkt Häufigkeiten verschiedener Zirkulationstypen vorliegen.

6.2.1 Differenzierung nach Modellgenerationen

In Abbildung 6.5 sind die berechneten MSE-Werte (auf der Ordinate) gegenüber den bekannten Leadjahre-Kategorien (auf der Abszisse) hinsichtlich der Auftrittshäufigkeiten für das historische, uninitialisierte Modell (**hist**) sowie für die unterschiedlichen Modellgenerationen (**b0**, **b1**, **pS**, **pG**) dargestellt. Betrachtet man die Ergebnisse, wenn man die Auftrittshäufigkeiten über das gesamte Jahr mittelt (Abbildung 6.5a), fällt auf, dass bei der überblickenden Darstellung – d. h. Mittelung über die unterschiedlichen Ansätze

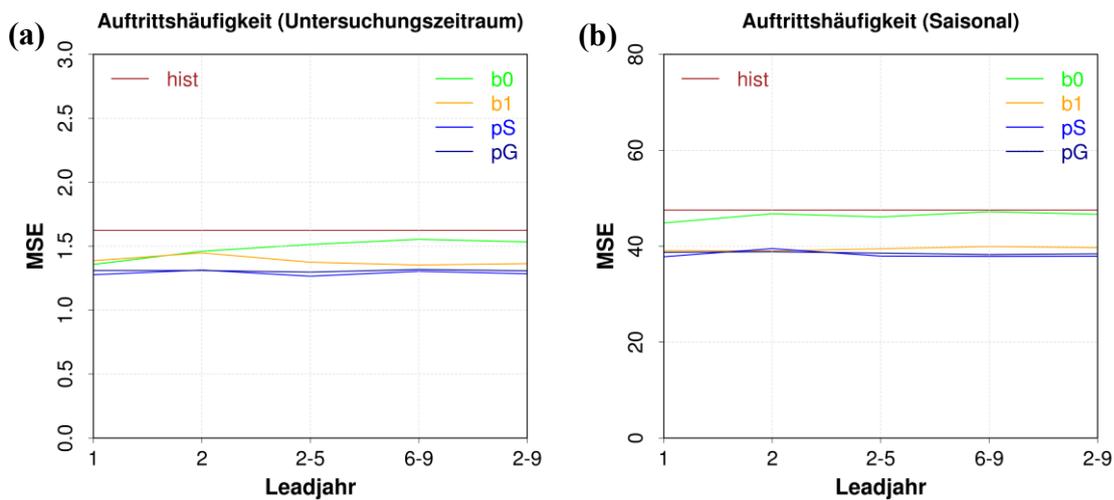


Abbildung 6.5: MSE hinsichtlich der Auftrittshäufigkeit bei Mittelung über den gesamten Untersuchungszeitraum (a) bzw. bei saisonaler Aggregation (b) für die unterschiedlichen **Modellgenerationen** des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von verschiedenen Leadjahren.

(GWT, SAN, TPCA), alle Jahreszeiten und unterschiedlichen Variablen (SLP, ZG500) – die historischen Modellläufe die stärksten Abweichungen zu den beobachteten Werten aufweisen und diese sich für die unterschiedlichen Leadjahre aufgrund der fehlenden Initialisierung nicht ändern. Demgegenüber zeigen die initialisierten Läufe Verbesserungen in der Vorhersageleistung, auch wenn diese insgesamt gering ausfallen. Die Linien der initialisierten Modelle liegen in allen Leadjahren unterhalb derjenigen des historischen Modells, insbesondere diejenigen der dritten Modellgeneration (**pS**, **pG**) mit einer „full-field“-Initialisierung von Atmosphäre und Ozean. Der Einfluss der Initialisierung ist aber auch dahingehend sichtbar, dass im ersten Leadjahr zumeist die geringsten MSE-Werte auftreten und diese für die folgenden Leadjahre ansteigen, was insbesondere bei der ersten Modellgeneration (**b0**) zu beobachten ist.

Ein ähnliches Bild zeigt sich, wenn man die Auftrittshäufigkeiten nicht gemittelt über den gesamten Untersuchungszeitraum, sondern saisonal aggregiert betrachtet (Abbildung 6.5b). Wiederum ist zu erkennen, dass das uninitialisierte, historische Modell die schlechteste Vorhersageleistung offenbart, diese sich aber durch die Initialisierung in geringem Umfang über alle Leadjahre hinweg verbessert. Zwar liegt die erste Modellgeneration (**b0**) nah an den historischen Modellläufen, jedoch weisen die weiteren Modellgenerationen (**b1**, **pS**, **pG**), die sich alle auf einem ähnlichen Niveau befinden, eine bemerkbare

Verbesserung in der Vorhersageleistung auf. Zu beachten ist, dass die Abweichungen von der Beobachtung bei der saisonalen Aggregierung (s. Abbildung 6.5b) um ein Vielfaches größer ausfallen als bei Mittelung über den gesamten Untersuchungszeitraum (s. Abbildung 6.5a). Dieser Tatsache wird in Abbildung 6.5 durch eine unterschiedliche Skalierung der y-Achsen Rechnung getragen, wodurch die beiden Grafiken allerdings nicht direkt miteinander vergleichbar sind. Dennoch lässt sich festhalten, dass bei beiden Herangehensweisen der Einfluss der Initialisierung beobachtet werden kann.

Neben den Auftrittshäufigkeiten können die verschiedenen Modellgenerationen hinsichtlich der typinternen Charakteristika untersucht werden. Die Ergebnisse sollen an dieser Stelle nur knapp erläutert werden, da sie sich von denen der Auftrittshäufigkeiten nicht wesentlich unterscheiden. In Anhang 9 offenbart sich hinsichtlich des typinternen Charakteristikums der Persistenz sowohl bei Mittelung über den Untersuchungszeitraum (Anhang 9a) als auch bei saisonaler Aggregierung (Anhang 9b) ein bekanntes Bild. Das uninitialisierte, historische Modell offenbart in der Regel die schlechteste Vorhersageleistung, wohingegen die Initialisierung des Modells zu geringeren MSE-Werten führt (v. a. bei der dritten Modellgeneration). Oftmals ist zudem zu erkennen, dass die geringsten Abweichungen im ersten Jahr nach der Initialisierung auftreten. Auffallend ist die Vorhersageleistung der ersten Modellgeneration (**b0**), die bei Mittelung über den gesamten Untersuchungszeitraum für die ersten beiden Leadjahre sogar bessere Ergebnisse als die zweite Modellgeneration (**b1**) liefert, ansonsten aber hinter die historischen Modellläufe zurückfällt, und bei saisonaler Aggregierung im Vorhersagebereich der historischen Modellläufe liegt. Auch für das typinterne Charakteristikum der Intensität zeigt sich in Anhang 9 ein ähnliches Bild, wobei sowohl bei Mittelung über den Untersuchungszeitraum (Anhang 9c) als auch bei saisonaler Aggregierung (Anhang 9d) erneut die dritte Modellgeneration (**pS**, **pG**) die beste Vorhersageleistung abliefert. Anzumerken ist wiederum, dass ein direkter Vergleich zwischen den einzelnen typinternen Charakteristika wie auch den unterschiedlichen Aggregierungsarten aufgrund der unterschiedlichen Skalierungen der y-Achse nicht möglich ist.

6.2.2 Differenzierung nach Klassifikationsverfahren

Differenziert man die Vorhersageleistung des dekadischen Vorhersagesystems nicht nach den unterschiedlichen Modellgenerationen, sondern nach den verschiedenen Klassifikationsverfahren, ergibt sich ein Gesamtbild, das Abbildung 6.6 zeigt. In den beiden Grafiken ist zudem noch eine Unterscheidung hinsichtlich der jeweils verwendeten Variable getroffen worden: Während die helleren Farbtöne auf die Verwendung des Luftdrucks auf Meereshöhe (SLP) hindeuten, zeigen dunklere Farbtöne die Verwendung der geopotentiellen Höhe auf dem 500hPa-Niveau (ZG500) an.

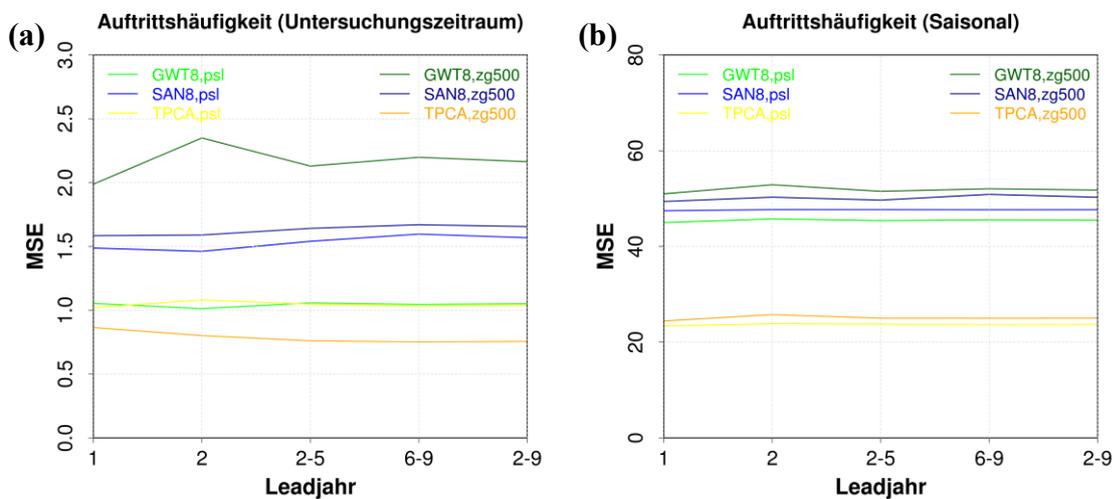


Abbildung 6.6: Wie Abbildung 6.5, aber für die **unterschiedlichen Klassifikationsverfahren**.

Betrachtet man die Grafiken im Detail, weisen die über den gesamten Untersuchungszeitraum gemittelten Auftrittshäufigkeiten (Abbildung 6.6a) deutliche Unterschiede für die einzelnen Klassifikationsverfahren auf. Dabei zeigt das Verfahren der t-modalen PCA die besten Vorhersageleistungen, egal ob man es auf den Luftdruck auf Meeresspiegel (SLP) oder auf die 500 hPa-Geopotentialfläche (ZG500) anwendet, was mitunter auch daran liegt, dass bei dieser Methode die erste Klasse sehr dominant auftritt. Ein vergleichbar gutes Ergebnis zeigt auch das GWT-Klassifikationsverfahren, allerdings nur für SLP. Wendet man das Verfahren demgegenüber auf ZG500 an, werden die Auftrittshäufigkeiten durch das dekadische Vorhersagemodell am schlechtesten repräsentiert. Zwischen diesen beiden Bereichen liegt die SANDRA-Clusteranalyse (SAN), wobei diese bei der Anwendung auf SLP leicht besser ausfällt als auf ZG500. Der Einfluss der Initialisierung ist wiederum durch geringere MSE-Werte im ersten Leadjahr ersichtlich, er ist aber nicht

stark ausgeprägt und lässt sich auch nicht durchgehend für alle Klassifikationsverfahren festhalten.

Für die saisonale Aggregation (Abbildung 6.6b) liegen die MSE-Werte wiederum in einem deutlich höheren Wertebereich, es zeigt sich jedoch ein vergleichbares Bild. Das Verfahren der t-modalen Hauptkomponentenanalyse (TPCA) weist für beide Variablen die beste Vorhersageleistung bezüglich der Auftrittshäufigkeiten auf. Schlechtere Ergebnisse offenbaren die SANDRA-Clusteranalysen (SAN) und die Großwettertypen nach Beck (GWT) in derselben Reihenfolge wie bei der Mittelung über den gesamten Untersuchungszeitraum.

6.2.3 Differenzierung nach Jahreszeiten

Abbildung 6.7 weist mit der Darstellung der Ergebnisse für die über den gesamten Untersuchungszeitraum gemittelten Auftrittshäufigkeiten (Abbildung 6.7a) bzw. für die saisonal aggregierten Auftrittshäufigkeiten (Abbildung 6.7b) denselben Aufbau auf wie die beiden vorherigen Abbildungen. Zwischen den einzelnen Jahreszeiten sind für beide zeitliche Auflösungen Unterschiede zu erkennen, wobei es dem dekadischen Vorhersagemodell in beiden Fällen überraschenderweise im Winter am schlechtesten gelingt, die Auftrittshäufigkeiten abzubilden. Demgegenüber weist der Sommer (JJA) bei der Mittelung über den Untersuchungszeitraum bzw. der Herbst (SON) bei saisonaler Aggregation die beste Vorhersageleistung auf.

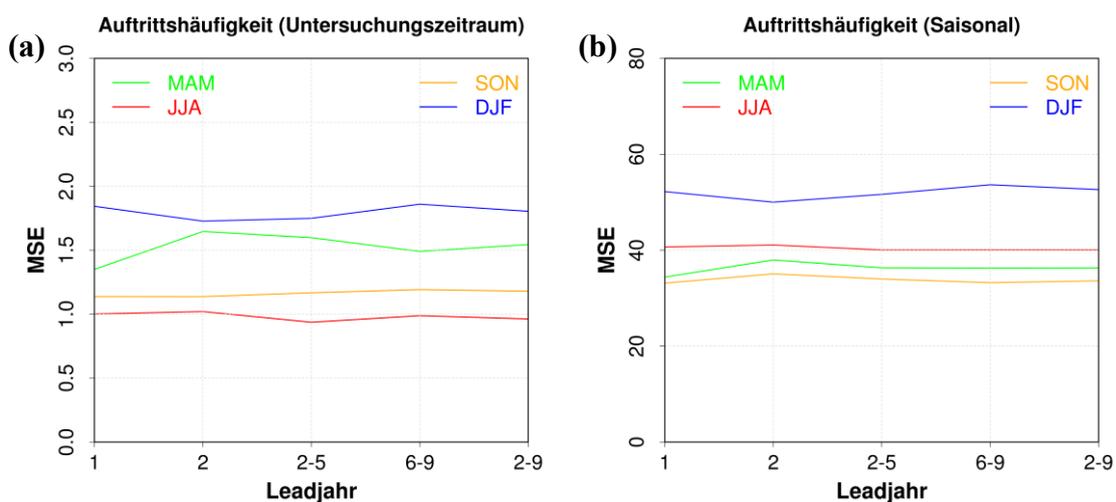


Abbildung 6.7: Wie Abbildung 6.5, aber für die unterschiedlichen Jahreszeiten.

Das wirklich Interessante und Erkenntnisreiche an den beiden Grafiken ist allerdings, dass sich der Einfluss der Initialisierung klarer herausbildet als in den vorherigen Abbildungen. Dieser ist für beide zeitliche Auflösungen im Frühjahr (MAM) daran zu beobachten, dass die MSE-Werte im ersten Leadjahr deutlich geringer ausfallen als in den übrigen Leadjahren. In anderen Jahreszeiten ist sogar eine Verschlechterung in der Vorhersageleistung vom ersten zum zweiten Leadjahr zu erkennen, wie z. B. im Sommer (JJA) oder Winter (DJF) bei Mittelung über den gesamten Untersuchungszeitraum (Abbildung 6.7a). Aus diesem Ergebnis kann geschlussfolgert werden, dass die Initialisierung des Modells bei Mittelung über alle Klassifikationsverfahren und Modellgenerationen nur im Frühjahr (MAM) des ersten Leadjahres, d. h. dem dritten, vierten und fünften Monat, und somit lediglich für die ersten Monate nach der Initialisierung einen erkennbaren Einfluss besitzt.

6.3 Der Einfluss der Ensemblegröße

Neben der Initialisierung des MPI-ESM-Vorhersagemodells wurde in der vorliegenden Arbeit der Einfluss der Ensemblegröße untersucht. Es wurde auf dasselbe Verfahren zurückgegriffen, das bereits für die Untersuchungen zu den Telekonnektionsindizes angewendet wurde (vgl. Kapitel 5.3). Aufgrund der hohen Rechenzeiten mussten wiederum Einschränkungen vorgenommen werden, sodass lediglich die ersten beiden Leadjahre in die Analyse integriert wurden. Zudem wurde die Vorhersageleistung des dekadischen Vorhersagemodells hinsichtlich Häufigkeit sowie der Persistenz als typinternes Charakteristikum untersucht. D. h. die Analyse der Intensität blieb außen vor und zudem fand lediglich die Mittelung über den gesamten Untersuchungszeitraum und nicht die saisonale Aggregation Berücksichtigung. Als Maß für den Unterschied zwischen Vorhersage und Beobachtung wurde der MSE verwendet.

Zunächst werden die Ergebnisse im Überblick differenziert nach den unterschiedlichen Modellgenerationen dargestellt, bevor auf weitere erwähnenswerte Aspekte (Jahreszeiten, Variablen, Klassenanzahl) näher eingegangen wird.

6.3.1 Ensemblegröße und Modellgeneration

In Abbildung 6.8, in der – mit dem MSE als Maß für den Unterschied zwischen Vorhersage und Beobachtung – die Vorhersageleistung für die unterschiedlichen Modellgenerationen hinsichtlich der Auftrittshäufigkeit der Zirkulationstypen dargestellt ist, offenbart sich ein Bild, das auch für die Telekonnectionsindizes beobachtet werden konnte (vgl. Kapitel 5.3). Sowohl im ersten (Abbildung 6.8a) als auch im zweiten Leadjahr (Abbildung 6.8b) sinken die MSE-Werte mit zunehmender Anzahl an Ensemblemitgliedern, d. h. die Vorhersageleistung der verschiedenen Modellgenerationen verbessert sich demnach.

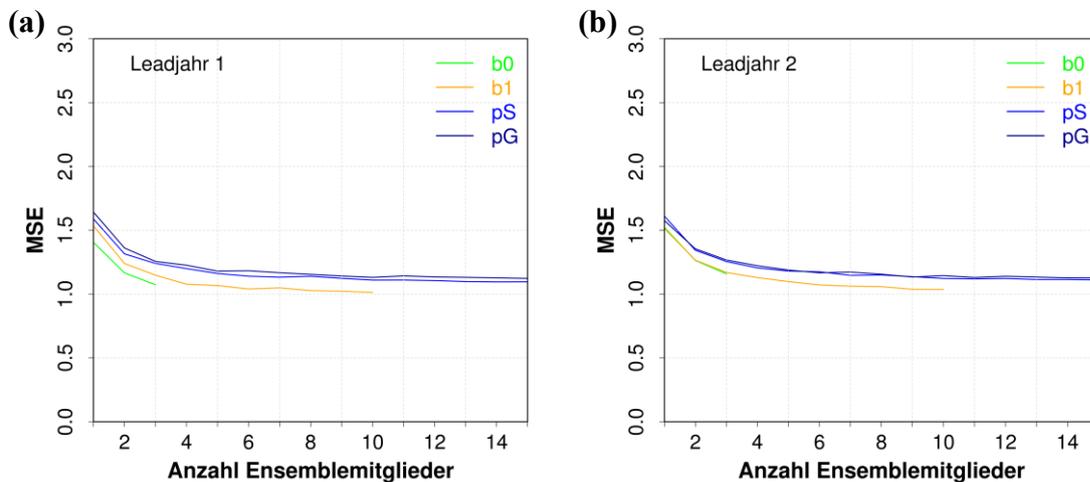


Abbildung 6.8: MSE hinsichtlich der *Auftrittshäufigkeit* (bei Mittelung über den gesamten Untersuchungszeitraum) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (a) bzw. das zweite Leadjahr (b) nach Initialisierung.

Überraschend ist allerdings, dass die erste Modellgeneration (**b0**), die für die Telekonnectionsindizes in Kapitel 5 insgesamt die schlechteste Vorhersageleistung aller Modellgenerationen offenbarte, im ersten Leadjahr (Abbildung 6.8a) für eine geringe Anzahl an Ensemblemitgliedern die geringsten MSE-Werte aufweist, wohingegen die dritte Modellgeneration (**pS**, **pG**) am schlechtesten abschneidet. Erst für eine Anzahl von 15 Ensemblemitgliedern ist für die dritte Modellgeneration ein ähnlich geringer MSE-Wert erkennbar wie für die erste Modellgeneration (**b0**) bei drei Ensemblemitgliedern. Die zweite Modellgeneration (**b1**) lässt für eine steigende Anzahl an Ensemblemitgliedern niedrigere

MSE-Werte erkennen und weist für die maximal mögliche Anzahl von zehn Ensemblemitgliedern die beste Vorhersageleistung auf.

Im zweiten Leadjahr (Abbildung 6.8b) zeigt sich ein ähnliches Ergebnis, wenngleich **b0** im Vergleich zum ersten Leadjahr schlechter abschneidet. Im Gegensatz zu den Ergebnissen für die Telekonnexionsindizes (vgl. Kapitel 5.3.1) liegen die MSE-Werte im zweiten Leadjahr nicht deutlich höher als im ersten Leadjahr, sondern heben sich – wenn überhaupt – nur geringfügig von denen des ersten Leadjahres ab, was auf eine eher geringe Bedeutung der Initialisierung für die Verbesserung der Vorhersageleistung schließen lässt. Insgesamt kann aber für beide Leadjahre festgehalten werden, dass die Vergrößerung der Ensemblemitgliederanzahl zwar die Vorhersage verbessert, das Modell aber von einer perfekten Vorhersage, d. h. einem MSE-Wert nahe 0, noch weit entfernt ist.

Das typinterne Charakteristikum der Persistenz bestätigt die Ergebnisse, die für die Auftrittshäufigkeit beschrieben wurden. In Abbildung 6.9 sind äquivalent zu Abbildung 6.8 für verschiedene Modellgenerationen die MSE-Werte zwischen Vorhersage und Beobachtung hinsichtlich der Persistenz (bei Mittelung über den gesamten Untersuchungs-

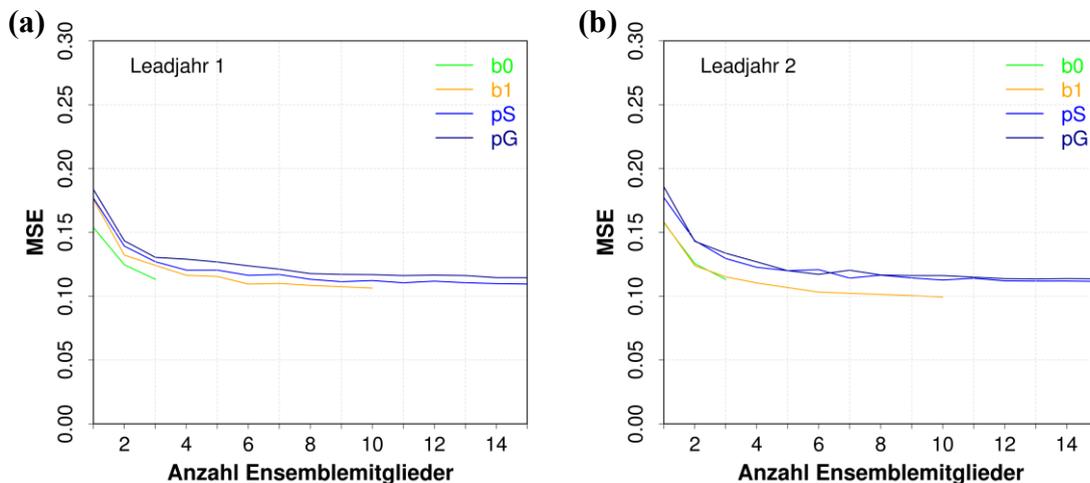


Abbildung 6.9: Wie Abbildung 6.8, aber für das typinterne Charakteristikum der *Persistenz*.

zeitraum) in Abhängigkeit von der Anzahl an Ensemblemitgliedern aufgetragen. Sowohl für das erste Leadjahr (Abbildung 6.9a) als auch das zweite Leadjahr (Abbildung 6.9b) lässt sich – wie für die Auftrittshäufigkeiten – eine Verbesserung in der Vorhersage bei Vergrößerung des Ensembles beobachten. Auffallend ist wiederum die vergleichsweise gute Vorhersageleistung der ersten (**b0**) und die relativ schlechte Vorhersageleistung der

dritten Modellgeneration (**pS**, **pG**) in beiden Leadjahren. Zudem fällt auf, dass insbesondere die zweite Modellgeneration (**b1**) für das zweite Leadjahr niedrigere MSE-Werte offenbart als für das erste Leadjahr, was im Vergleich zu den bisherigen Ergebnissen ungewöhnlich ist. Somit lässt sich für das typinterne Charakteristikum der Persistenz festhalten, dass die MSE-Werte aufgrund der geringeren Absolutwerte im Vergleich zu den Auftrittshäufigkeiten zwar niedriger ausfallen. Aber vergleichbar zu den Auftrittshäufigkeiten lässt sich auch für die Persistenz nur eine geringe Verbesserung in der Vorhersageleistung durch die Erhöhung der Anzahl an Ensemblemitgliedern beobachten, sodass das dekadische Modell zu einer perfekten Vorhersage mit MSE-Werten nah Null noch eine große Differenz aufweist.

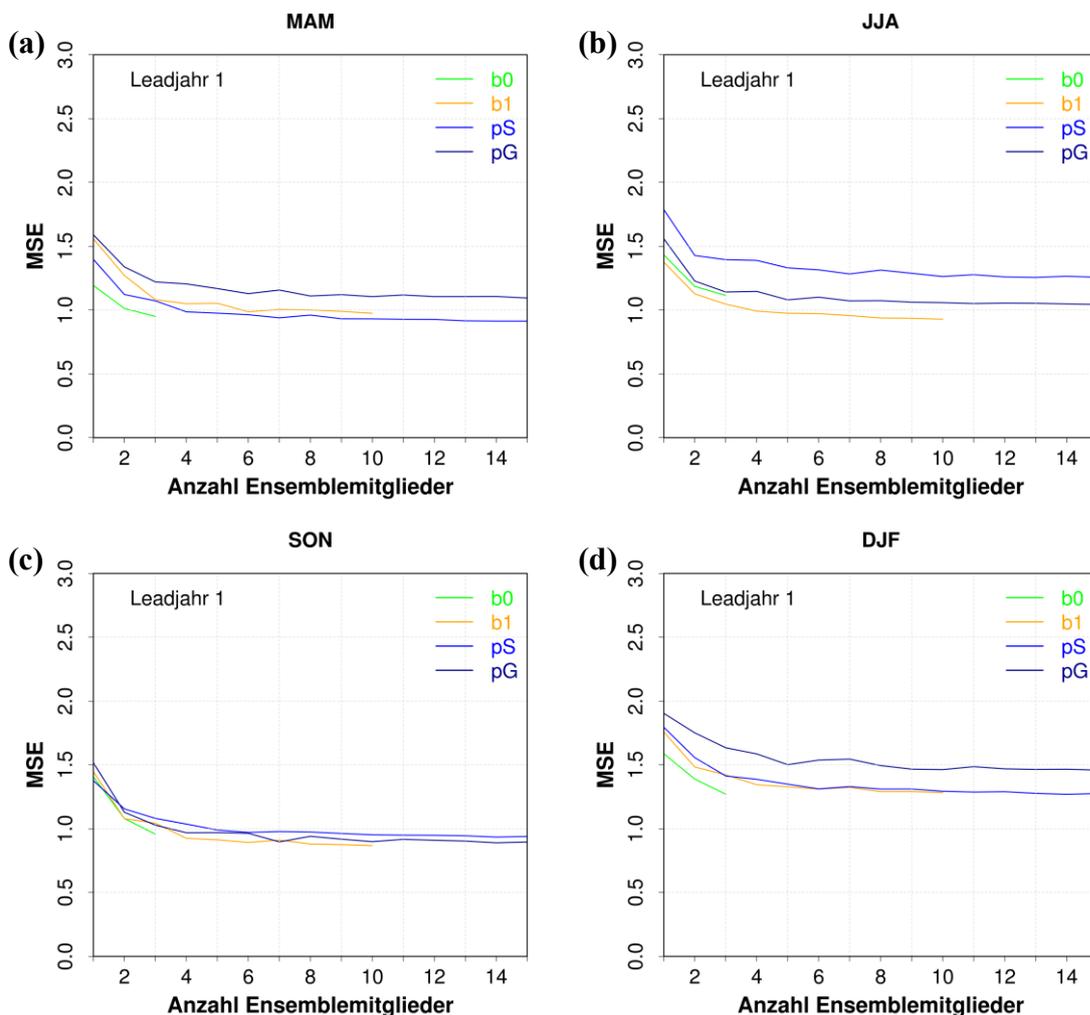


Abbildung 6.10: MSE hinsichtlich der *Auftrittshäufigkeit* (bei Mittelung über den gesamten Untersuchungszeitraum) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für verschiedene Jahreszeiten des ersten Leadjährrs nach Initialisierung.

6.3.2 Ensemblegröße und Jahreszeiten

Die Differenzierung der Vorhersageleistung für die unterschiedlichen Jahreszeiten (Abbildung 6.10) bestätigt das Bild, dass eine Vergrößerung der Anzahl an Ensemblemitgliedern bei allen Modellgenerationen im ersten Leadjahr zu einer Annäherung zwischen dekadischer Vorhersage und Modell hinsichtlich der Auftrittshäufigkeiten sorgt, was sich an niedrigeren MSE-Werten manifestiert. Auffallend sind jedoch die Unterschiede zwischen den einzelnen Jahreszeiten. So zeigt im Frühjahr (MAM) des ersten Leadjahres, d. h. dem dritten, vierten und fünften Monat nach Initialisierung, die erste Modellgeneration (**b0**) die beste Vorhersageleistung, die durch die nachfolgenden Modellgenerationen (**b1**, **pS**) nur bei deutlich höherer Ensemblemitgliederanzahl bzw. gar nicht (**pG**) erreicht wird. Dass mit **b0** die Modellgeneration, die als einzige keine Initialisierung der Atmosphäre besitzt, in den ersten Monaten nach der Initialisierung die niedrigsten MSE-Werte

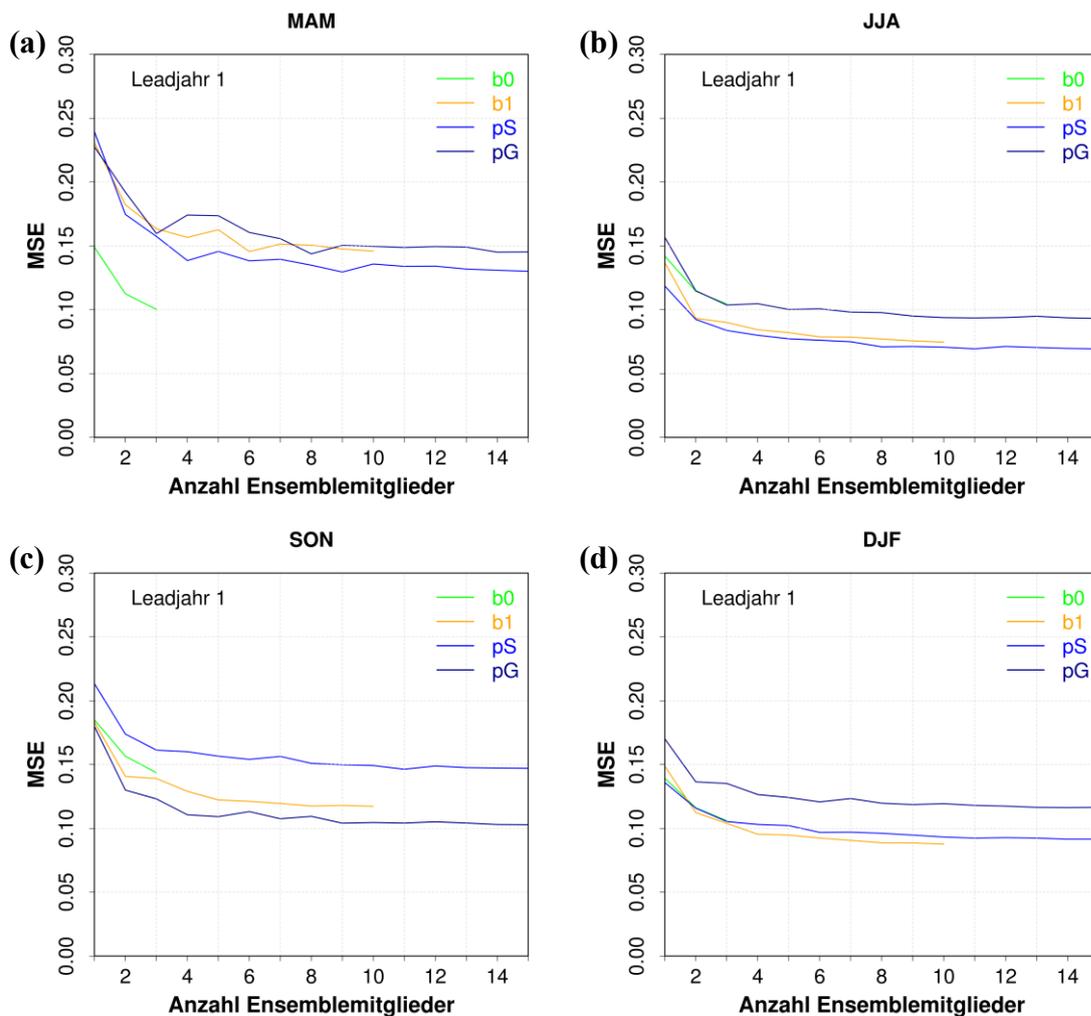


Abbildung 6.11: Wie Abbildung 6.10, aber für das typinterne Charakteristikum der *Persistenz*.

aufweist, offenbart sich noch deutlicher für die Persistenz (s. Abbildung 6.11), bei der die nachfolgenden Modellgenerationen (**b1**, **pS**, **pG**) selbst bei ihrer jeweils maximal möglichen Anzahl an Ensemblemitgliedern nicht die Vorhersageleistung von **b0** erreichen. Die Überlegenheit der ersten Modellgeneration gegenüber den nachfolgenden zeigt sich für die Persistenz (s. Abbildung 6.11) jedoch in keiner weiteren Jahreszeit, bezüglich der Auftrittshäufigkeiten lediglich im Winter (DJF), wie aus Abbildung 6.10 zu ersehen ist.

In Abbildung 6.10 fällt zudem auf, dass im Winter (DJF) und einschränkend auch im Sommer (JJA) im Vergleich zu den anderen Jahreszeiten höhere MSE-Werte zu beobachten sind, wohingegen die Übergangsjahreszeiten Frühling (MAM) und Herbst (SON) die geringsten Werte aufweisen. Bezieht man in die Auswertung allerdings noch die Ergebnisse für die Persistenz (s. Abbildung 6.11) ein, lässt sich diesbezüglich kein Muster erkennen und der Befund kann nicht bestätigt bzw. verallgemeinert werden.

6.3.3 Ensemblegröße und Variablen

Des Weiteren wurde der Einfluss der Ensemblegröße differenziert nach den in den Analysen verwendeten Variablen – dem Luftdruck auf Meereshöhe (SLP) bzw. der geopotentiellen Höhe auf dem 500 hPa-Niveau (ZG500) – betrachtet (vgl. Abbildung 6.12).

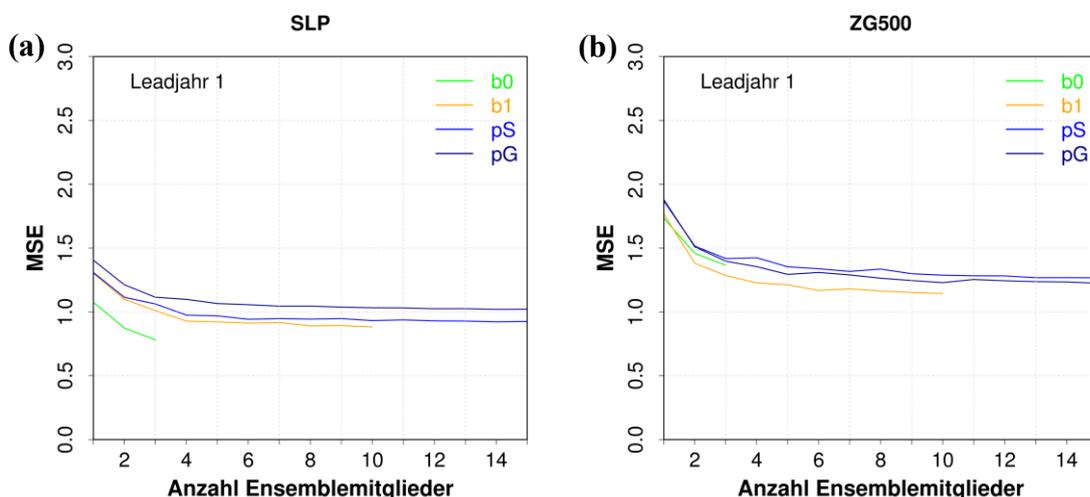


Abbildung 6.12: MSE hinsichtlich der *Auftrittshäufigkeit* (bei Mittelung über den gesamten Untersuchungszeitraum) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERA-Interim in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für SLP (a) und ZG500 (b) des ersten Leadjahrs nach Initialisierung.

Hierbei sticht wiederum die Überlegenheit der ersten Modellgeneration (**b0**) bei der Vorhersage der Auftrittshäufigkeiten hervor, die für die Variable SLP (Abbildung 6.12a) deutlich ausgeprägt ist, allerdings für die Variable ZG500 (Abbildung 6.12b) nicht behauptet werden kann, wo die zweite Modellgeneration (**b1**) die beste Vorhersageleistung offenbart.

Wie in den vorherigen Kapiteln zeigt sich logischerweise auch bei der Differenzierung nach den verwendeten Variablen eine Abnahme der MSE-Werte bei zunehmender Zahl an Ensemblemitgliedern, d. h. bei einer Vergrößerung des Ensembles nähert sich die Vorhersage der Beobachtung an. Davon abgesehen fällt beim Vergleich der Variablen auf, dass dem dekadischen Modell die Vorhersage der über den gesamten Untersuchungszeitraum gemittelten Auftrittshäufigkeiten für den Luftdruck auf Meereshöhe (SLP) besser gelingt als für die geopotentielle Höhe (ZG500). Noch stärker ausgeprägt, lässt sich diese Beobachtung für die Persistenz machen (vgl. Abbildung 6.13), wo die MSE-Werte für ZG500 (Abbildung 6.13b) deutlich über denen für SLP (Abbildung 6.13a) liegen.

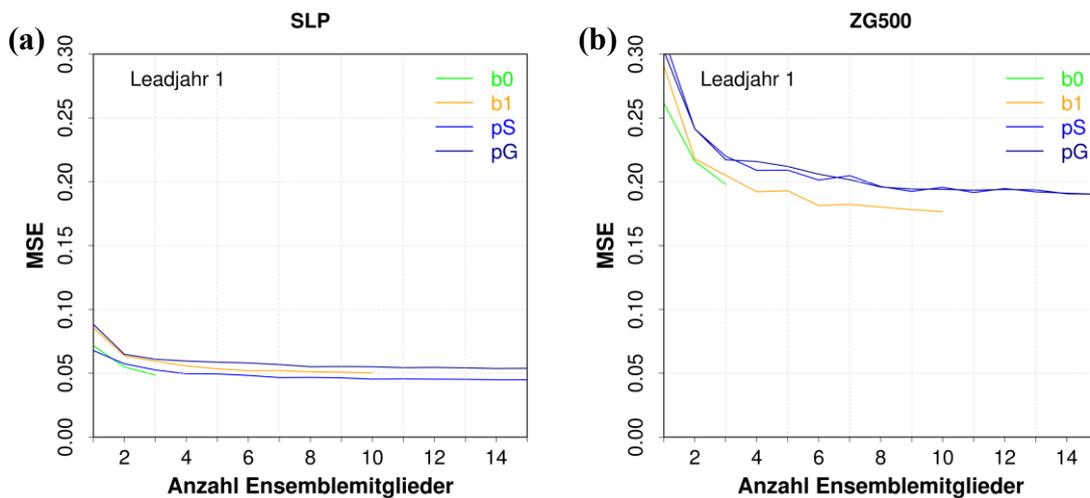


Abbildung 6.13: Wie Abbildung 6.12, aber für das typinterne Charakteristikum der **Persistenz**.

Abschließend kann für ZG500 festgehalten werden, dass sich sowohl für die Auftrittshäufigkeiten (s. Abbildung 6.12) als auch für das typinterne Charakteristikum der Persistenz (s. Abbildung 6.13) eine deutlich stärker ausgeprägte Verbesserung in der Vorhersageleistung durch die Vergrößerung des Ensembles beobachten lässt, allerdings weicht die Vorhersage bei ZG500 auch deutlich stärker von der Beobachtung ab als bei SLP.

6.3.4 Ensemblegröße und Klassenanzahl

Zu guter Letzt wurde im Zusammenhang mit der Ensemblegröße untersucht, ob die Klassenanzahl (d. h. die Anzahl an Zirkulationstypen) einen Einfluss auf die Vorhersageleistung hat und inwiefern sich diese durch eine Vergrößerung des Ensembles verändert. Die Untersuchungen wurden mit einer Anzahl von vier, acht und 18 Klassen durchgeführt. Aus diesem Grund gingen lediglich die beiden Klassifikationsverfahren SAN und TPCA in die Analyse ein, da sich die Methode GWT nicht mit einer Klassenanzahl von vier verwirklichen lässt. In Abbildung 6.14 werden die Ergebnisse für die Variable ZG500 für das erste Leadjahr bezüglich der Vorhersage der Auftrittshäufigkeiten in gewohnter Art präsentiert. Abbildung 6.14 offenbart große Unterschiede hinsichtlich der MSE-Werte zwischen den drei Klassenzahlen, die sich anhand der unterschiedlichen Linientypen unterscheiden lassen. Die Unterschiede fallen deutlich größer aus als derjenige zwischen den unterschiedlichen Modellgenerationen, die farblich differenziert gestaltet sind und innerhalb einer Klassenanzahl ähnliche Ergebnisse aufweisen.

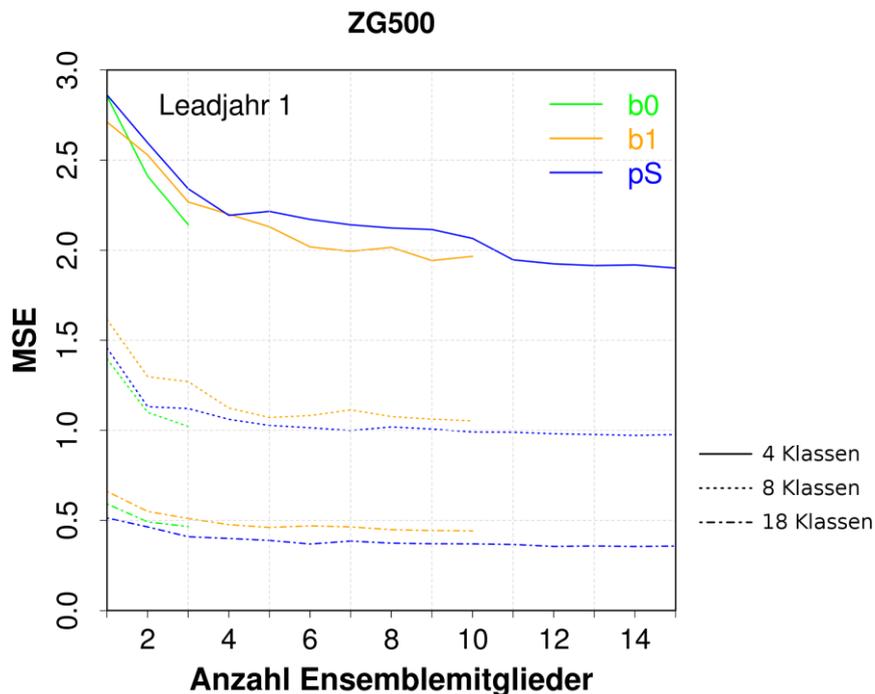


Abbildung 6.14: MSE hinsichtlich der *Auftrittshäufigkeit* (bei Mittelung über den gesamten Untersuchungszeitraum) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren mit unterschiedlicher Klassenanzahl; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für ZG500 des ersten Leadjahres nach Initialisierung.

Wenig überraschend weist die Variante mit vier Klassen für alle Modellgeneration die höchsten Werte auf, was sich damit erklären lässt, dass die Unterschiede in den Auftrittshäufigkeiten bei geringer Klassenanzahl besonders hoch ausfallen, wodurch höhere MSE-Werte resultieren als bei geringeren Unterschieden in den Auftrittshäufigkeiten. Dementsprechend weist die Variante mit der höchsten Klassenanzahl, d. h. mit 18 Klassen, für alle Modellgenerationen die geringsten Werte auf. Auch in diesem Fall lässt sich, wie Abbildung 6.14 zeigt, wiederum eine Abnahme der MSE-Werte, d. h. eine Verbesserung in der Vorhersageleistung, mit einer zunehmenden Anzahl an Ensemblemitgliedern beobachten. Diese Verbesserung fällt bei der Variante mit vier Klassen, d. h. der Variante mit dem höchsten Verbesserungspotenzial, absolut am höchsten aus. Insgesamt weichen jedoch alle Varianten im Verhältnis noch deutlich von einer perfekten Vorhersage ab.

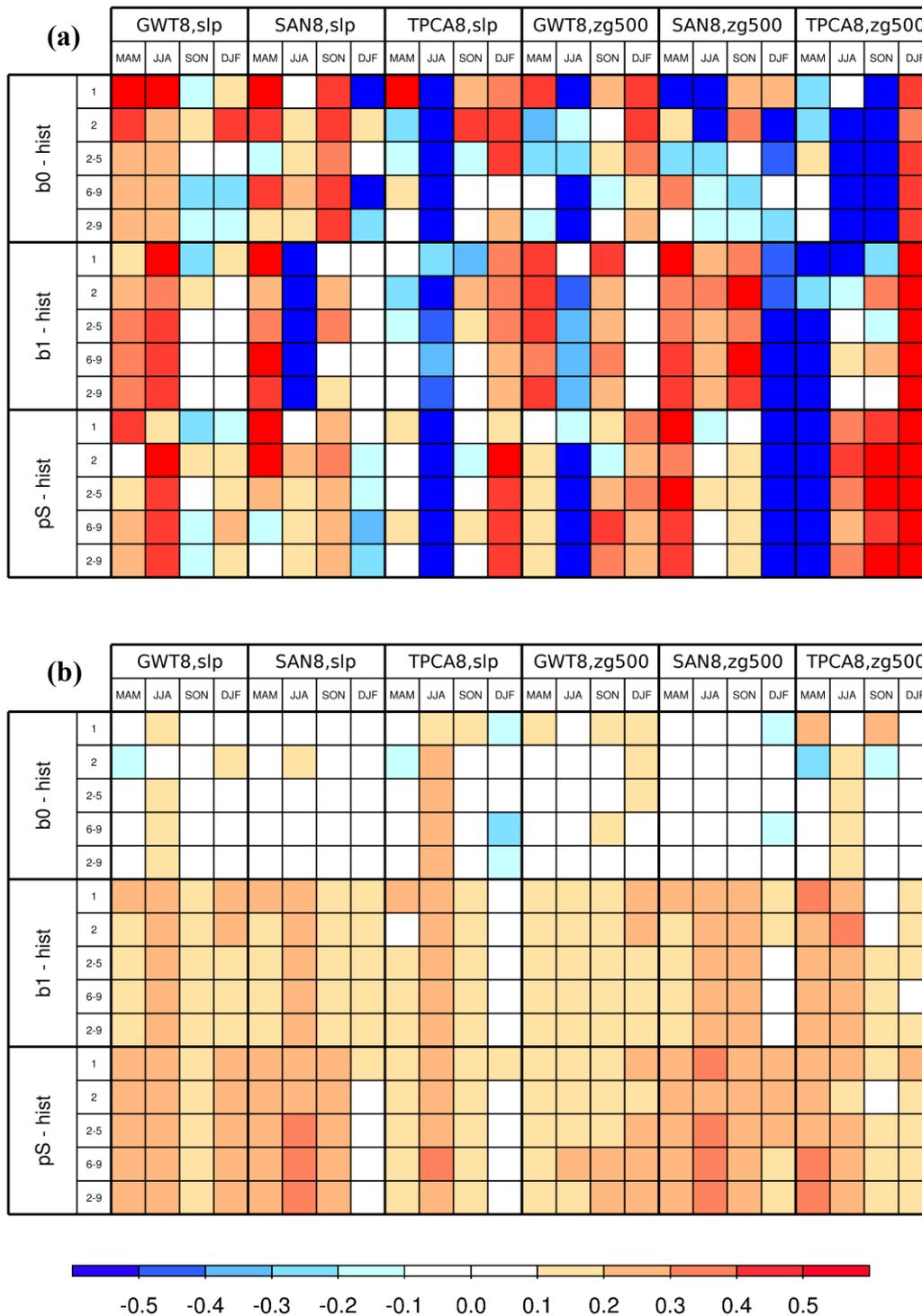
Zusammenfassend lässt sich für den Ansatz der Zirkulationstypen ein ähnliches Fazit über den Einfluss der Ensemblegröße auf die Vorhersageleistung des MPI-ESM-Vorhersagesystems ziehen, wie es bereits in Kapitel 5.4 für die Telekonnektionsindizes erfolgte: Die Vergrößerung des Ensembles führt in der Regel zu einer Verbesserung der Vorhersage, was sich in niedrigeren Werten des Mean Squared Errors (MSE) offenbart. Zwar nähert sich die Vorhersage der Beobachtung an, dennoch lassen sich trotz der Erhöhung der Anzahl an Ensemblemitglieder große Abweichungen sowohl hinsichtlich der untersuchten Parameter (Klassenhäufigkeit, Intensität, Persistenz) als auch der Klassifikationsverfahren (GWT, SAN, PCA) konstatieren. Als größter Unterschied zu den Ergebnissen der Telekonnektionsindizes kann die relativ gute Performanz der ersten Modellgeneration (**b0**) festgehalten werden.

6.4 Vorhersageleistung des dekadischen Modells

Nach der separaten Betrachtung der beiden Einflussgrößen – Initialisierungstechnik und Ensemblegröße – steht in diesem Kapitel die integrative Betrachtungsweise der Vorhersageleistung des dekadischen Modells im Vordergrund. Als Gütemaß kommt der Mean Squared Error Skill Score (MSSS, vgl. Kapitel 4.2.2.2) zum Einsatz, der Ranked Probability Skill Score (RPSS, vgl. Kapitel 4.2.2.3) eignet sich hingegen nicht.

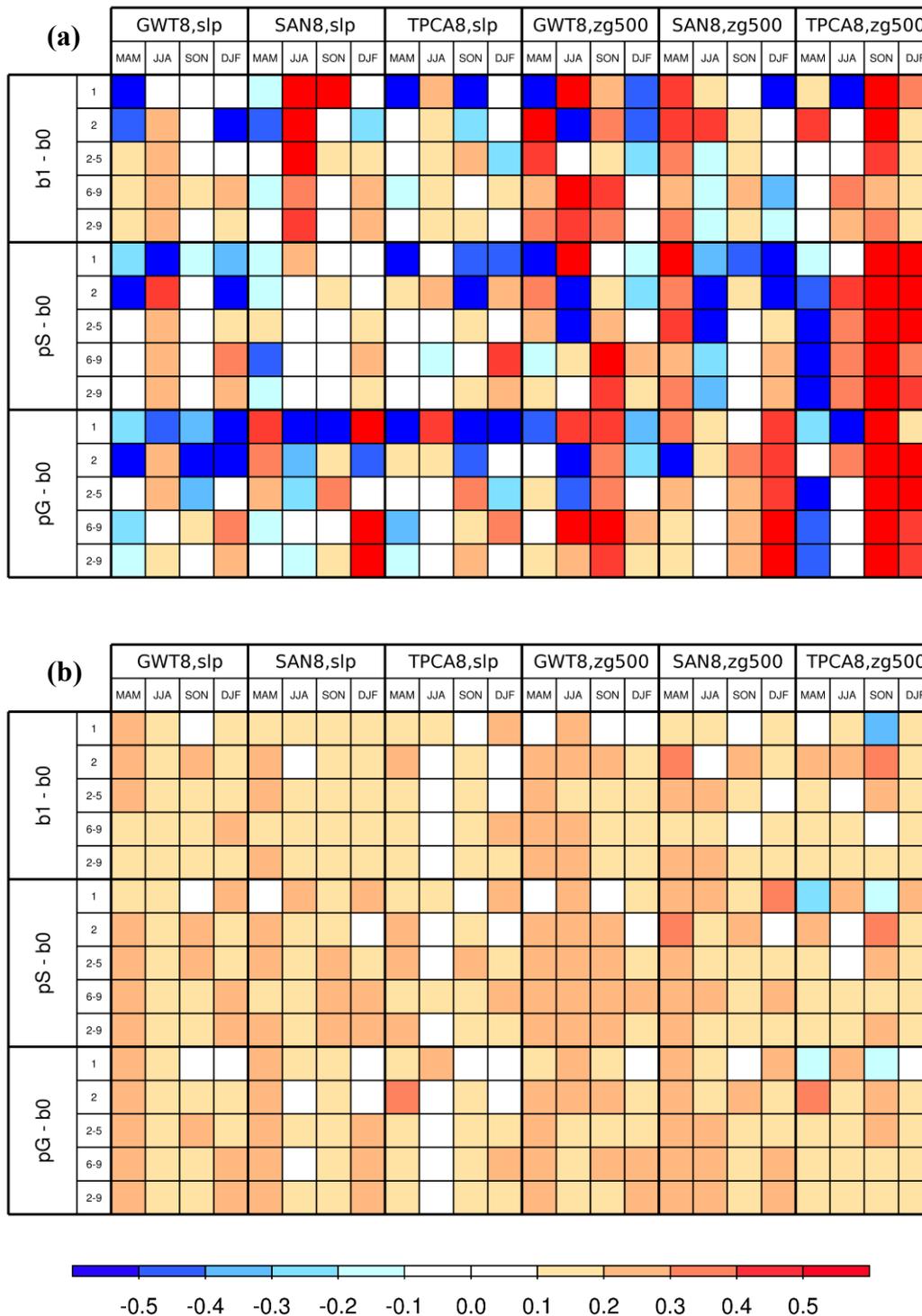
Um Aussagen über mögliche Verbesserungen in der Vorhersageleistung treffen zu können, werden in einer überblickenden Gesamtschau die Ergebnisse für die verschiedenen Klassifikationsverfahren (aus Gründen der besseren Vergleichbarkeit jeweils mit einer Klassenanzahl von acht) präsentiert. Dafür werden in den drei folgenden Abbildungen (Abbildung 6.15 bis Abbildung 6.17) für die drei verwendeten Klassifikationsverfahren (GWT, SAN, TPCA) und die beiden Variablen (SLP, ZG500) jeweils die MSSS-Werte für die Auftrittshäufigkeiten bei Aggregation über den gesamten Zeitraum (jeweils obere Grafik) bzw. bei saisonaler Aggregation (jeweils untere Grafik) dargestellt.

Die Grafiken weisen einen ähnlichen Aufbau auf wie für die Telekonnektionsindizes (vgl. Kapitel 5.4.2), d. h. die Ergebnisse werden für die unterschiedlichen Leadjahre (1, 2, 2-5, 6-9, 2-9) und die unterschiedlichen Jahreszeiten (MAM, JJA, SON, DJF) hinsichtlich einer bestimmten Referenz präsentiert. Als Referenz dienen wiederum **hist** (Abbildung 6.15), **b0** (Abbildung 6.16) bzw. **b1** (Abbildung 6.17), wobei bei Abbildung 6.17 darauf zu achten ist, dass im untersten Analyseblock der Grafik nicht **b1** als Referenz dient, sondern die beiden Varianten der dritten Modellgeneration miteinander verglichen werden (**pG–pS**). Abbildung 6.15 gibt Auskunft darüber, ob die initialisierten Läufe bessere Ergebnisse gegenüber den uninitialisierten aufweisen. Betrachtet man den MSSS für die über den gesamten Zeitraum gemittelten Auftrittshäufigkeiten (Abbildung 6.15a) ist ein heterogenes Bild zu erkennen. Zwar sind für die initialisierten Modellgenerationen (**b0**, **b1**, **pS**) deutliche Verbesserungen gegenüber den historischen Läufen (**hist**) feststellbar, v. a. bei den auf die Variable SLP angewendete Methoden GWT8 und SAN8 im ersten Leadjahr sowohl im Frühjahr (MAM) als auch im Sommer (JJA). Allerdings treten auch blaue Farbtöne, d. h. negative MSSS-Werte, beim Vergleich mit den uninitialisierten Läufen auf. Besonders markant sind schlechtere Vorhersageleistungen vor allem bei der Methode TPCA8 zu beobachten, egal ob die Methode auf SLP oder ZG500 angewendet wird. Dennoch lohnt sich ein genauerer Blick, denn verbreitet sind für die Methode TPCA8 deutliche Verbesserungen in der Vorhersageleistung zu erkennen, wie z. B. für die Variable ZG500 beim Vergleich der dritten Modellgeneration mit den historischen Läufen (**pS–hist**), wenn man das Frühjahr (MAM) außer Acht lässt. Insgesamt scheinen die Jahreszeiten einen größeren Einfluss auf die Vorhersage zu haben, was sich in der vertikal angeordneten Farbgebung offenbart, insgesamt prägt jedoch eine starke Heterogenität das Gesamtbild.



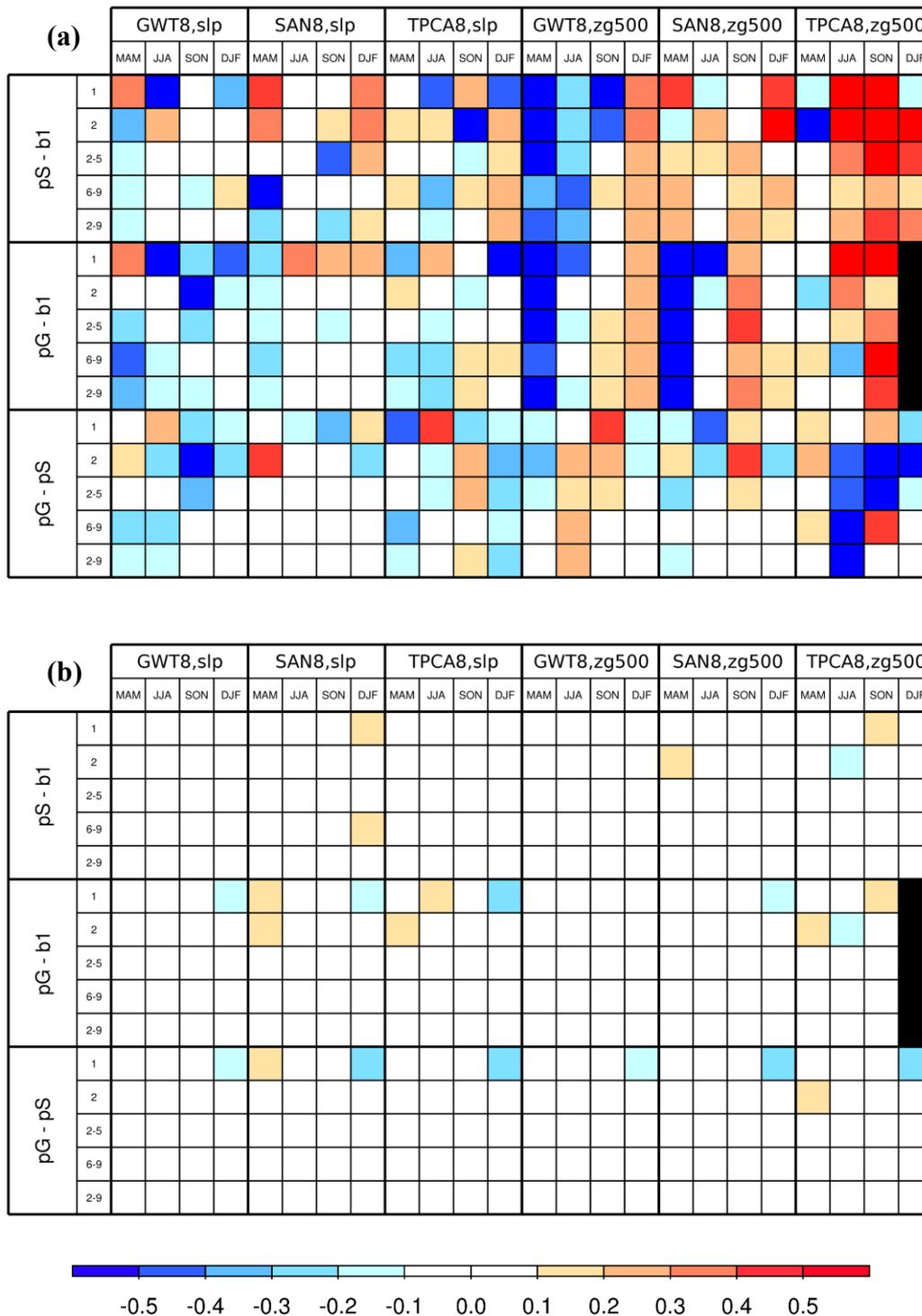
MSSS

Abbildung 6.15: MSSS jeweils für (von oben nach unten) *b0 – hist*, *b1 – hist* und *pS – hist* für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse für die über den gesamten Zeitraum (a) bzw. saisonal aggregierten (b) **Auftrittshäufigkeiten** der auf verschiedenen Methoden (s. Überschriften) beruhenden Zirkulationstypen sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarzen Flächen sind keine Daten vorhanden.



MSSS

Abbildung 6.16: MSSS jeweils für (von oben nach unten) $b1 - b0$, $pS - b0$ und $pG - b0$ für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse für die über den gesamten Zeitraum (a) bzw. saisonal aggregierten (b) Auftretshäufigkeiten der auf verschiedenen Methoden (s. Überschriften) beruhenden Zirkulationstypen sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarzen Flächen sind keine Daten vorhanden.



MSSS

Abbildung 6.17: MSSS jeweils für (von oben nach unten) *pS – b1*, *pG – b1* und *pG – pS* für verschiedene Leadjahre (1, 2, 2-5, 6-9, 2-9). Die Ergebnisse für die über den gesamten Zeitraum (a) bzw. saisonal aggregierten (b) **Auftrittshäufigkeiten** der auf verschiedenen Methoden (s. Überschriften) beruhenden Zirkulationstypen sind dargestellt für verschiedene Jahreszeiten (MAM, JJA, SON, DJF). Bei schwarzen Flächen sind keine Daten vorhanden.

Zu einem anderen Ergebnis gelangt man, wenn man nicht die über den gesamten Zeitraum gemittelten, sondern die saisonal aggregierten Auftrittshäufigkeiten (Abbildung 6.15b) betrachtet. Im Vergleich mit dem uninitialisierten, historischen Modell offenbaren die initialisierten Läufe insbesondere der zweiten (**b1**) und dritten Modellgeneration (**pS**) durchgängig Verbesserungen in der Vorhersageleistung, was an den rötlichen Farbtönen zu erkennen ist. Lediglich bei der ersten Modellgeneration (**b0**) treten die rötlichen Farbtöne deutlich seltener auf und es zeigen sich mitunter auch blaue Farbtöne, die negative MSSS-Werte implizieren, d. h. eine Verschlechterung in der Vorhersageleistung aufzeigen. Große Unterschiede zwischen den unterschiedlichen Klassifikationsverfahren sind genauso wenig zu erkennen wie hinsichtlich der zugrundeliegenden Variablen SLP bzw. ZG500. Somit lässt sich für die saisonal aggregierten Auftrittshäufigkeiten insgesamt festhalten, dass in den meisten Fällen Verbesserungen im Vergleich mit den uninitialisierten Läufen zu erkennen sind, diese allerdings nicht so stark ausfallen, was die nicht so stark ausgeprägten Rottöne (im Vergleich zu Abbildung 6.15a) mit MSSS-Werten von überwiegend unter 0,4 anzeigen.

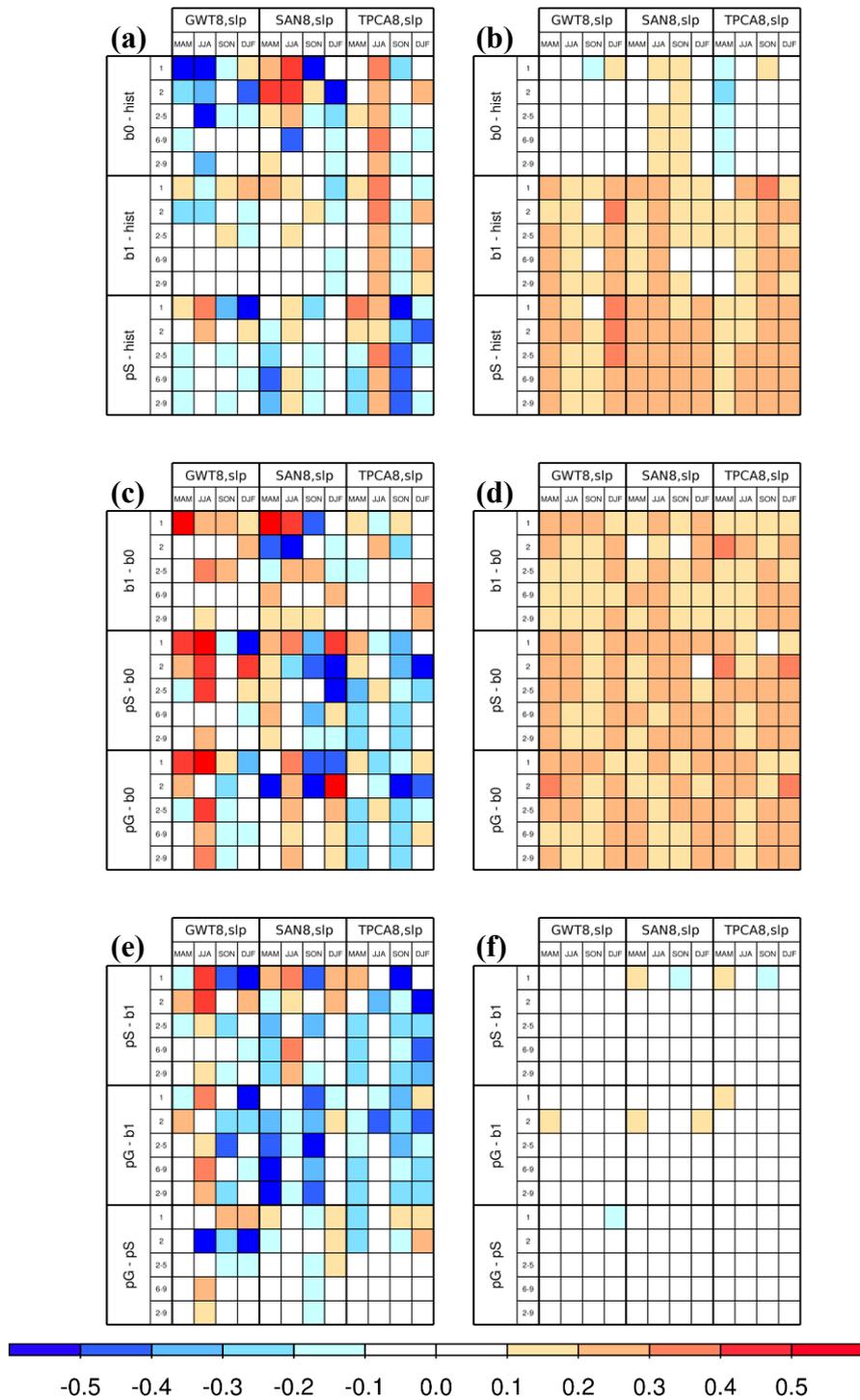
Wenn durch die Initialisierung vereinzelt (Abbildung 6.15a) bzw. nahezu durchgängig (Abbildung 6.15b) Verbesserungen in der Vorhersageleistung beobachtet werden können, drängt sich die Frage auf, ob die „full-field“-Initialisierung (der Atmosphäre) Verbesserungen gegenüber einer „anomaly“-Initialisierung offenbart, d. h. welche Initialisierungstechnik die besten Resultate liefert. Darauf gibt Abbildung 6.16 eine Antwort, in der die erste Modellgeneration (**b0**) als Referenz dient. Betrachtet man die über den gesamten Zeitraum gemittelten Auftrittshäufigkeiten (Abbildung 6.16a), zeigt sich wiederum ein heterogenes, ja sogar chaotisches Bild. Tendenzen in der Vorhersageleistung lassen sich nicht erkennen und somit können allgemein gültige Aussagen nur schwer getroffen werden. Ansatzweise lässt sich erkennen, dass die „full-field“-initialisierten Läufe – und dabei insbesondere **pG** – gegenüber den „anomaly“-initialisierten (**b0**) im ersten Leadjahr häufig negative MSSS-Werte aufweisen, was v. a. für die Variable SLP beobachtet werden kann. Das würde bedeuten, dass die „full-field“-Initialisierung sich negativ auf die Vorhersageleistung auswirkt, allerdings dominieren bei den folgenden Leadjahren (2, 2-5, 6-9 bzw. 2-9) tendenziell Rottöne. Demgegenüber offenbaren die saisonal aggregierten Auftrittshäufigkeiten (Abbildung 6.16b) erneut ein völlig anderes Bild. Im Vergleich mit

der „anomaly“-initialisierten ersten Modellgeneration des dekadischen Vorhersagemodells (**b0**) weisen die „full-field“-initialisierten Modellgenerationen (**b1**, **pS**, **pG**) nahezu durchgehend positive MSSS-Werte auf, was nach den Erkenntnissen aus Abbildung 6.15b zu erwarten war.

In einer abschließenden Untersuchung wurde überprüft, welche „full-field“-Initialisierung die besten Ergebnisse liefert. Wie aus Abbildung 6.17 zu entnehmen ist, zeigt sich weder bei Mittelung über den gesamten Zeitraum (Abbildung 6.17a) noch beim saisonalen Ansatz (Abbildung 6.17b) eine Präferenz hinsichtlich der Frage, welche Initialisierung die besten Ergebnisse liefert. Während im ersten Fall ein Durcheinander von Rot- und Blautönen zu erkennen ist, liegen im zweiten Fall Werte des MSSS nahezu durchgehend um 0.

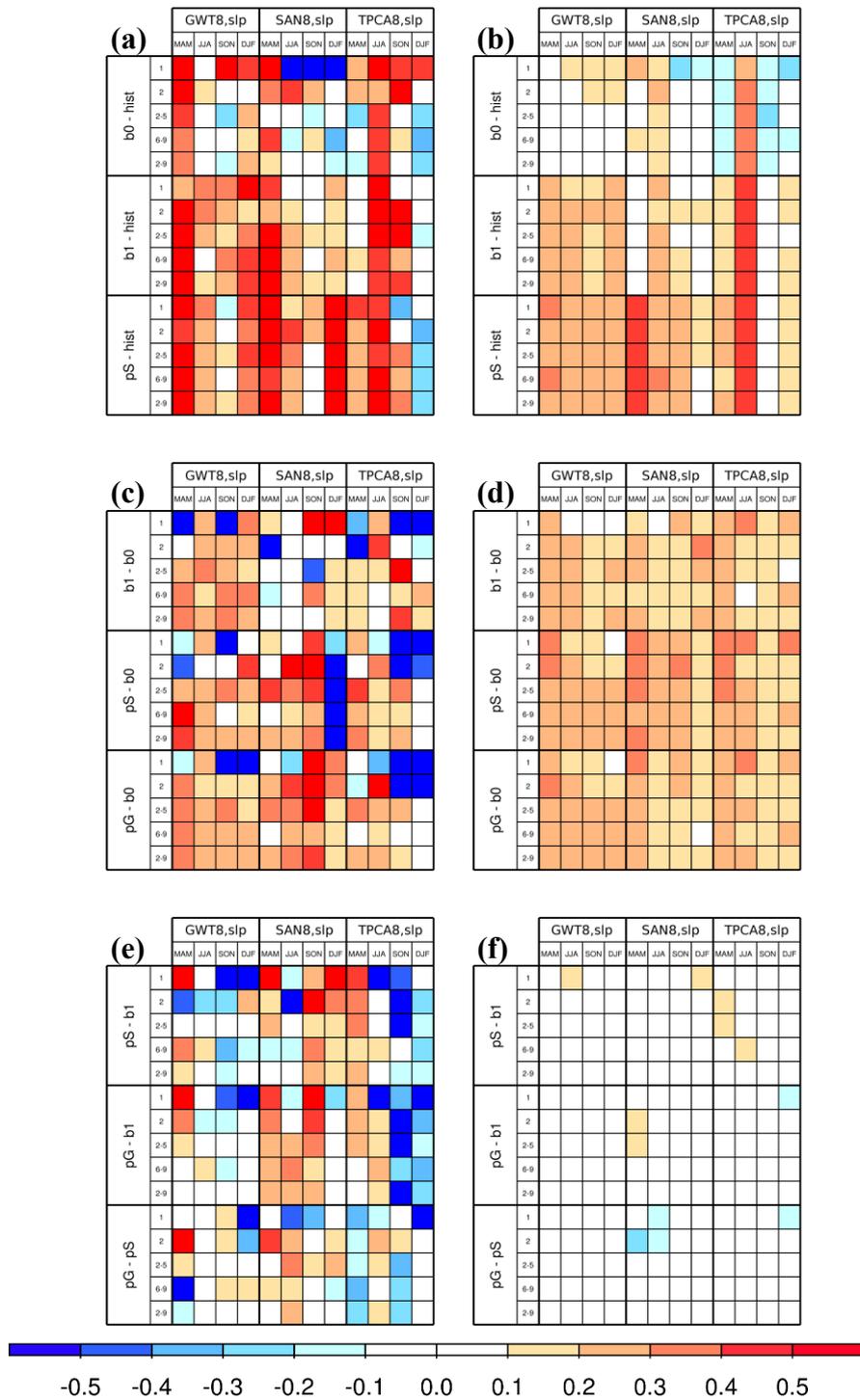
Zu ähnlichen Ergebnissen gelangt man, wenn man anstatt der Auftretshäufigkeiten die Intensität (Abbildung 6.18) bzw. die Persistenz (Abbildung 6.19) betrachtet. Aus Gründen der Übersichtlichkeit berücksichtigen die beiden Grafiken lediglich die Ergebnisse für den Luftdruck auf Meereshöhe (SLP), d. h. die geopotentielle Höhe (ZG500) bleibt außen vor. Die Abbildungen sind derart angelegt, dass jeweils die obere Grafik **hist**, die mittlere Grafik **b0** und die untere Grafik **b1** als Referenz aufweisen. Die Grafiken auf der linken Seite der Abbildung (a, c, e) zeigen die Ergebnisse für die über den gesamten Zeitraum gemittelten Werte, die auf der rechten Seite (b, d, f) die Ergebnisse für die saisonal aggregierten. Sowohl für die Intensität (s. Abbildung 6.18) als auch für die Persistenz (s. Abbildung 6.19) fallen die Farbtöne bei Mittelung über den gesamten Zeitraum insgesamt deutlich kräftiger aus, was auf ausgeprägtere Verbesserungen bzw. Verschlechterungen hinweist, ergeben aber insgesamt ein unstrukturierteres Bild als bei den saisonal aggregierten Werten, wo rote Farbtöne und damit Verbesserungen in der Vorhersage dominieren, diese aber schwächer ausfallen.

Allerdings sind markante Unterschiede zwischen den beiden typinternen Parameter sichtbar. Deutlich ist zu erkennen, dass die Verbesserungen für die Persistenz stärker ausfallen als für die Intensität. Zum einen ist das für die über den gesamten Zeitraum gemittelte Persistenzen (Abbildung 6.19a, c, e) zu beobachten, wo zwar auch blaue Farbtöne im un-



MSSS

Abbildung 6.18: MSSS jeweils für den Vergleich mit *hist* (a, b), *b0* (c, d) sowie *b1/pS* (e, f). Die Ergebnisse für über den gesamten Zeitraum (linke Hälfte) bzw. die saisonal (rechte Hälfte) gemittelten **Intensitäten** der auf verschiedenen Methoden (GWT, PCA, SAN) beruhenden jeweils 8 Zirkulationstypen (basierend auf SLP) sind dargestellt für verschiedene Jahreszeiten. Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.



MSSS

Abbildung 6.19: MSSS jeweils für den Vergleich mit *hist* (a, b), *b0* (c, d) sowie *b1/pS* (e, f). Die Ergebnisse für über den gesamten Zeitraum (linke Hälfte) bzw. die saisonal (rechte Hälfte) gemittelten **Persistenzen** der auf verschiedenen Methoden (GWT, PCA, SAN) beruhenden jeweils 8 Zirkulationstypen (basierend auf SLP) sind dargestellt für verschiedene Jahreszeiten. Bei schwarz hinterlegten Flächen sind keine Daten vorhanden.

strukturierten Gesamtbild zu erkennen sind, diese aber im Vergleich zu den Ergebnissen der Intensität (Abbildung 6.18a, c, e) deutlich seltener auftreten. So treten bei der Persistenz für alle Modellgenerationen nahezu durchgehend rote Farbtöne beim Vergleich mit den historischen Modellläufen (**hist**) auf (Abbildung 6.19a) und auch gegenüber der ersten Modellgeneration **b0** (Abbildung 6.19c) dominieren die Verbesserungen. Aber auch bei den saisonal aggregierten Werten sind die Verbesserungen in der Vorhersageleistung für die Persistenz (Abbildung 6.19b, d, f) ausgeprägter als für die Intensität (Abbildung 6.18b, d, f), wo die Rottöne schwächer ausgeprägt sind.

Auch wenn die beiden typinternen Charakteristika Unterschiede aufweisen, bestätigen sie insgesamt das Bild, das anhand der Ergebnisse über die Auftrittshäufigkeiten gewonnen wurde.

6.5 Zusammenfassung

Genauso wie die Ergebnisse für die Telekonnektionsindizes (vgl. Kapitel 5) offenbaren diejenigen für die Zirkulationstypen in ihrer Gesamtschau Verbesserungen für die weiterentwickelten Modellgenerationen. Diese sind bei allen Untersuchungsgegenständen zu beobachten, d. h. sowohl hinsichtlich der Auftrittshäufigkeiten als auch den typinternen Charakteristika der Persistenz und Intensität. Allerdings fallen die Verbesserungen in der Vorhersageleistung nicht so eindeutig aus, wie dies für die Telekonnektionsindizes beobachtet werden konnte. Zum einen liegt das daran, dass die über den gesamten Zeitraum gemittelten Daten häufig ein indifferentes, chaotisches Gesamtbild offenbaren, aus dem sowohl Verbesserungen als auch Verschlechterungen hinsichtlich der Vorhersageleistung durch die Initialisierung des Modells und die Erhöhung der Ensemblemitgliederanzahl hervorgehen. Zum anderen ist keine kontinuierliche Verbesserung für die weiterentwickelten Modellgenerationen – ausgehend von **hist** über **b0**, **b1** und die beiden prototype-Varianten (**pG**, **pS**) – wie bei den Ergebnissen zu den Telekonnektionsindizes zu erkennen. Vielmehr konnte gezeigt werden, dass auch die erste Modellgeneration (**b0**), die keine Initialisierung der Atmosphäre aufweist und lediglich drei Ensemblemitglieder besitzt, mitunter bessere Ergebnisse aufweist als die nachfolgenden Entwicklungsstufen. Dennoch kann abschließend eine Verbesserung aller initialisierten Modellgenerationen gegenüber den uninitialisierten festgehalten werden.

7 Zusammenfassung und Ausblick

Die vorliegende Arbeit untersuchte anhand zweier unterschiedlicher methodischer Herangehensweisen (Telekonnektionsindizes, Zirkulationstypen) die Vorhersagbarkeit der Dynamik der Atmosphäre auf dekadischer Zeitskala. Dafür wurde auf das dekadische Vorhersagesystem des Max-Planck-Instituts (MPI-ESM) zurückgegriffen. Eine zentrale Fragestellung lag nicht nur darin, herauszufinden, ob das dekadische Klimamodell in der Lage ist, die Dynamik der Atmosphäre adäquat wiederzugeben, sondern auch zu untersuchen, inwiefern verschiedene Initialisierungstechniken eine Auswirkung auf die Güte der Vorhersage haben.

Sowohl für die Telekonnektionsindizes als auch für die unterschiedlichen Klassifikationsverfahren konnte beobachtet werden, dass Verbesserungen in der Vorhersageleistung häufig im ersten Jahr nach Initialisierung des Modells, d. h. im ersten Leadjahr, und insbesondere für das Frühjahr, d. h. den dritten, vierten und fünften Monat nach Initialisierung des Modells, ermittelt werden konnten. V. a. durch den Vergleich mit uninitialisierten Modellläufen wurde dieser Einfluss auf die Vorhersageleistung sichtbar gemacht. Hierbei zeigt sich, dass die Läufe, deren atmosphärische Komponente eine „full-field“-Initialisierung aufweisen (**b1**, **pS**, **pG**), häufig bessere Ergebnisse liefern als die erste Modellgeneration (**b0**) mit einer „anomaly“-Initialisierung des Ozeans und keiner der At-

mosphäre. Eine Präferenz für eine bestimmte der verschiedenen „full-field“-Initialisierungstechniken ist jedoch nicht erkennbar. Dieses Resultat ergibt sich aus der Mittelung über verschiedene Telekonnektionsindizes bzw. Klassifikationsverfahren, Modellgenerationen, Jahreszeiten und zeitliche Aggregationsstufen (saisonal, monatlich, über den gesamten Untersuchungszeitraum).

Betrachtet man die Ergebnisse jedoch im Einzelnen, treten Abweichungen sowohl im positiven als auch im negativen Sinn auf, wodurch sich ein heterogenes Gesamtbild ergibt. Als besonders positives Beispiel gilt es, den Index der Südlichen Oszillation (SOI) hervorzuheben, für den in der vorliegenden Arbeit im ersten und teilweise auch im zweiten Leadjahr deutliche Verbesserungen in der Vorhersage aufgezeigt werden konnten. Insbesondere die hohen Korrelationskoeffizienten im Frühjahr (MAM) des ersten Leadjahrs mit Werten von über $r=0,7$ lassen eine potentielle Quelle dekadischer Vorhersagbarkeit erkennen. Demgegenüber konnte für die Nordatlantische Oszillation (NAO), deren Berechnung nach verschiedenen methodischen Ansätzen vorgenommen wurde, durch die Initialisierung nur eine äußerst geringfügige Verbesserung in der Vorhersageleistung und Werte nahe 0 für den Korrelationskoeffizienten sogar im Frühjahr (MAM) des ersten Leadjahres ermittelt werden.

Anhand der Ergebnisse in der vorliegenden Arbeit konnte zudem aufgezeigt werden, dass nicht nur die Initialisierung, sondern auch die Vergrößerung des Ensembles einen entscheidenden Einfluss auf die Vorhersageleistung des dekadischen Klimamodells besitzt. Dieser Einfluss konnte sowohl für die Telekonnektionsindizes als auch für den Ansatz der Zirkulationstypen nachgewiesen werden. Grundsätzlich zeigten die verschiedenen Abbildungen eine Verbesserung in der Vorhersageleistung mit einer zunehmenden Anzahl an Ensemblemitgliedern, wobei dieser Effekt für den niedrigen Zahlenbereich (z. B. bei Erhöhung der Anzahl von drei auf fünf Ensemblemitglieder) stärker ausfällt als für den höheren Zahlenbereich (z. B. bei Erhöhung der Anzahl von 13 auf 15 Ensemblemitglieder). Auch wenn sich grundsätzlich eine Verbesserung in der Vorhersageleistung durch die Vergrößerung des Ensembles einstellt, ergab sich wiederum ein heterogenes Bild für die unterschiedlichen Telekonnektionsindizes bzw. Klassifikationsverfahren, Jahreszeiten oder Variablen.

Beim Vergleich der unterschiedlichen Einflüsse auf die Vorhersageleistung ergab sich, dass zwar sowohl die Initialisierung des Modells als auch die Erhöhung der Ensemblemitgliederanzahl zu Verbesserungen führen, diese aber vorwiegend auf eine Erhöhung der Ensemblemitgliederanzahl und nur in geringerem Maße auf die unterschiedlichen Initialisierungsverfahren zurückzuführen sind. Dieses Ergebnis konnte insbesondere durch den Einsatz des Ranked Probability Skill Score (RPSS) als Gütemaß sichtbar gemacht werden, da bei dieser Metrik zur Beurteilung der Vorhersageleistung Modellversionen mit einer unterschiedlichen Anzahl an Ensemblemitgliedern direkt miteinander vergleichbar sind. Aber obwohl Verbesserungen in der Vorhersage sowohl für die Telekonnectionsindizes als auch für die Zirkulationstypen zu beobachten sind, muss dennoch als Fazit festgehalten werden, dass weder die Initialisierung des Modells noch die Vergrößerung des Ensembles sich auf die gesamte Dekade auswirken. Von einer dekadischen Vorhersagbarkeit kann somit nicht gesprochen werden, sondern lediglich von Verbesserungen in der Vorhersageleistung im ersten (und teilweise im zweiten) Jahr nach der Initialisierung, was insbesondere für das Frühjahr (MAM) gilt.

Neben der Frage nach der Initialisierung des Klimamodells bzw. dem Einfluss der Vergrößerung des Ensembles auf die Vorhersageleistung bezog sich eine weitere Fragestellung auf die Methodik der Analyse. Ziel der Arbeit war es nicht nur herauszufinden, welcher methodische Ansatz (Telekonnectionsindizes, Zirkulationstypen) am besten geeignet ist, um die Dynamik der Atmosphäre hinsichtlich ihrer dekadischen Vorhersagbarkeit zu untersuchen, sondern auch zu klären, welche Skillmetriken in diesem Zusammenhang besonders nützlich sind und welche zur Messung der Vorhersageleistung nicht geeignet sind. Für die unterschiedlichen Klassifikationsverfahren zeigte sich in der vorliegenden Arbeit, dass nicht alle Herangehensweise gleich gut geeignet sind. Ein grundlegender problematischer Schritt bei allen Klassifikationsverfahren besteht in der Festlegung der Klassenanzahl. Dieser subjektive Eingriff wirkt sich, wie gezeigt wurde, auf die Vorhersageleistung des dekadischen Klimamodells aus. Um Vergleiche anstellen und Aussagen zu möglichen Verbesserungen in der Vorhersageleistung treffen zu können, ist es deshalb notwendig, dieselbe Klassenanzahl für alle Klassifikationsverfahren und Jahreszeiten zu verwenden. Diese generalisierende Festlegung widerspricht dem herkömmlichen Vorgehen, anhand verschiedener Kriterien die am besten geeignete Klassenanzahl individuell festzulegen. Ausgenommen hiervon ist die Methode der Großwettertypen (GWT), bei der

die Klassenanzahl aufgrund der Definition von Prototypen bereits vor der Durchführung der Klassifikation festgelegt ist. Außerdem sind dadurch Beobachtungs- und Modelldaten direkt miteinander vergleichbar, wohingegen bei SAN und TPCA, in der Form wie sie in der vorliegenden Arbeit angewandt wurden, eine Zuordnung der Modelldaten zu den Beobachtungsdaten aufgrund eines bestimmten Kriteriums (z. B. führende Ladung) notwendig ist, um die Vorhersageleistung des dekadischen Modells beurteilen zu können.

Ein weiterer Nachteil des Ansatzes der Zirkulationstypen besteht darin, dass keine einzelne auswertbare Zeitreihe zur weiteren Untersuchung generiert wird. Vielmehr wird eine der gewählten Klassenanzahl entsprechende Anzahl an Zeitreihen erzeugt, die beispielsweise die Auftrittshäufigkeiten der einzelnen Klassen in ihrem zeitlichen Verlauf in Abhängigkeit vom Aggregationsniveau (monatlich, saisonal) darstellen. Dieses methodische Vorgehen ist unüblich auf dem dekadischen Forschungsgebiet, wo gewöhnlich einzelne Zeitreihen evaluiert werden, um Aussagen über die dekadische Vorhersagbarkeit eines bestimmten Parameters treffen zu können. Zudem erschwert es die Auswertung (z. B. hinsichtlich typinterner Charakteristika), wenn v. a. bei hoher Klassenanzahl einzelne Klassen unbesetzt sind. Ein möglicher Ansatz, um diesem Problem zu begegnen, besteht in der Aggregation der Daten über den gesamten Untersuchungszeitraum, wie es in der vorliegenden Arbeit durchgeführt wurde (Methode „Untersuchungszeitraum“). Dieses Vorgehen steht allerdings wiederum im Gegensatz zu den Zielen der dekadischen Forschung, das Klima und seine Variabilität auf dekadischer Zeitskala aufzudecken, da durch die Mittelung über den gesamten Untersuchungszeitraum die zeitliche Information verlorengeht. Zudem erschwert es den Einsatz von geeigneten Skillmetriken bzw. macht deren Einsatz unmöglich, da Metriken wie der RPSS oder der Korrelationskoeffizient nicht sinnvoll angewendet werden können.

Aus den genannten Gründen hat sich der Ansatz der Telekonnectionsindizes zur Evaluierung der dekadischen Vorhersagbarkeit der atmosphärischen Dynamik am hilfreichsten und sinnvollsten erwiesen. Genau wie bei TPCA basiert die Berechnung vieler Telekonnectionsindizes auf der Verwendung einer Hauptkomponentenanalyse zur Reduzierung der Dimensionalität, allerdings wird häufig, wie bei AAO-PC und NAO-PC, nur auf die erste Hauptkomponente mit dem höchsten Varianzerklärungsanteil zurückgegriffen oder die einzelnen Indizes werden wie bei den CPC-Telekonnectionsindizes prototypischen Vertretern zugewiesen. Genau wie bei den auf Gitterpunktwerten (NAO-GP, SOI)

bzw. zonalen Mitteln (AAO-ZM) beruhenden Berechnung ergibt sich bei der Evaluation zur dekadischen Vorhersagbarkeit der Vorteil, dass im Gegensatz zu den Zirkulationstypen nur eine Zeitreihe generiert wird, die die zeitliche Ausprägung des jeweiligen Telekonnexionsindex‘ zum Ausdruck bringt und mithilfe der unterschiedlichen Skillmetriken (RPSS, MSSS, Korrelationskoeffizient) analysiert werden kann. Zusammenfassend lässt sich somit festhalten, dass Telekonnexionsindizes aus den genannten Gründen besser zur Evaluierung der Dynamik der Atmosphäre hinsichtlich ihrer dekadischen Vorhersagbarkeit geeignet sind als die mit unterschiedlichen Klassifikationsverfahren generierten Zirkulationstypen.

Zu guter Letzt sollte mithilfe der vorliegenden Arbeit untersucht werden, welche Metriken sich zur Messung der Vorhersageleistung des dekadischen Modells besonders eignen. Aus der Vielzahl der zur Verfügung stehenden Metriken auf dem Forschungsgebiet der saisonalen und dekadischen Vorhersage wurde vorwiegend auf drei zurückgegriffen: Korrelationskoeffizient, MSSS (basierend auf MSE) und RPSS (basierend auf RPS). MSE erwies sich als das Gütemaß, das am flexibelsten einsetzbar ist und sich auch für die Evaluierung der Zirkulationstypen eignet. Für die Evaluierung einer Zeitreihe haben sich alle drei Skillmetriken als sinnvoll und hilfreich erwiesen, auch deswegen, weil sie Unterschiedliches abbilden. Während der MSE den mittleren Abstand zwischen den beobachteten und vorhergesagten Werte berechnet, bildet der Korrelationskoeffizient den gemeinsamen Verlauf der Zeitreihen von Beobachtung und Vorhersage ab, ohne den Abstand der einzelnen Werte voneinander zu berücksichtigen. Beide Verfahren können zu unterschiedlichen Ergebnissen führen und machen eine vergleichende Betrachtungsweise lohnend. Der große Vorteil des RPSS liegt hingegen in der Fähigkeit, Modellversionen mit einer unterschiedlichen Anzahl an Ensemblemitgliedern direkt zu vergleichen, was im Fall des MSSS nicht möglich ist. Somit eignen sich alle Skillmetriken zur und ergänzen sich in der Evaluierung der dekadischen Vorhersageleistung.

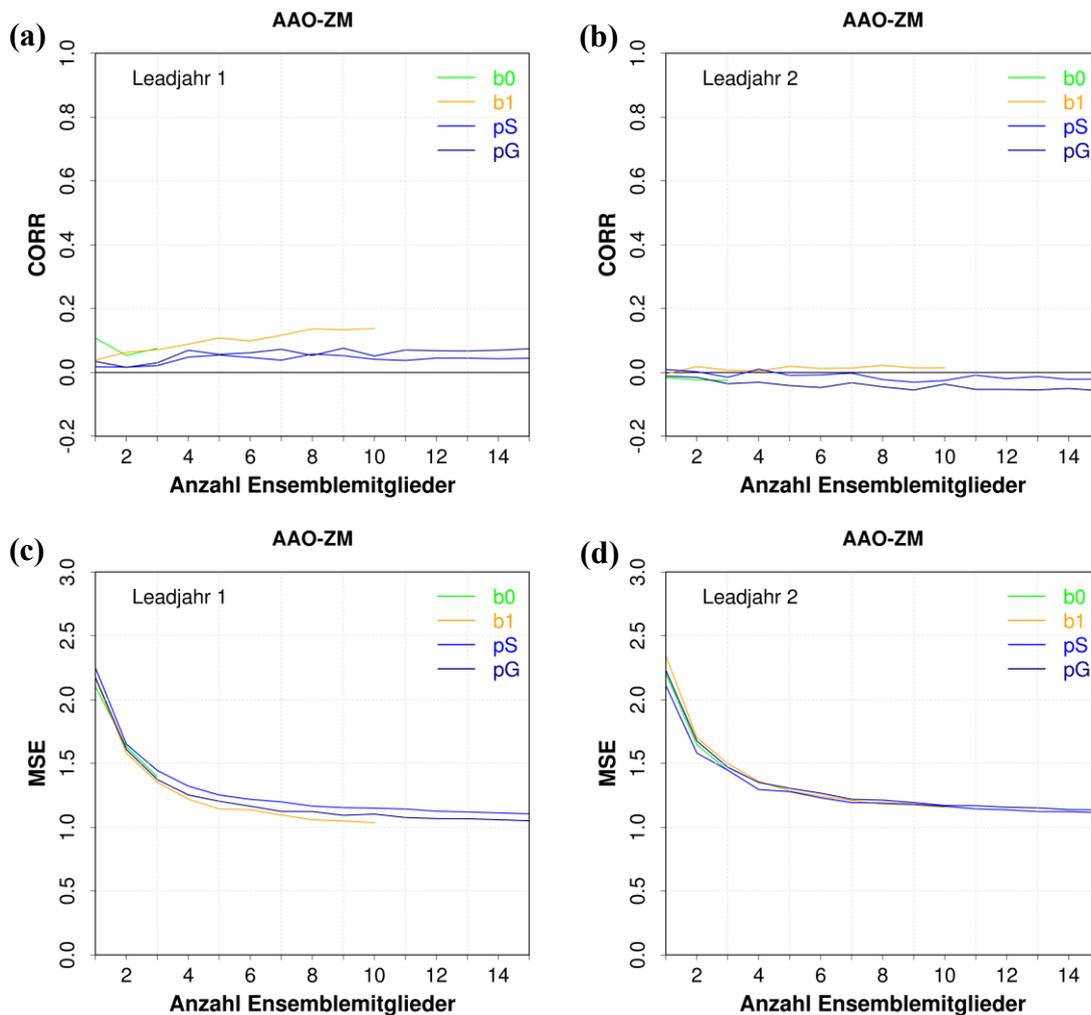
Ausblick

Die vorliegende Arbeit konnte aufzeigen, dass eine dekadische Vorhersagbarkeit der Dynamik der Atmosphäre nach derzeitigem Stand und mit den aktuellen Modellversionen noch nicht möglich ist. Um dieses Ziel zu erreichen, bedarf es intensiver Anstrengungen auf dem Gebiet der dekadischen Forschung. Zum einen müssen bestehende Initialisierungsverfahren verfeinert und verbessert werden, zum anderen bedarf es einer weiteren Erhöhung der Ensemblemitgliederanzahl, was in neueren Modellversionen bereits realisiert wurde. Ebenso kann eine Verbesserung der räumlichen Auflösung der Klimamodelle sowie der Modellphysik dazu beitragen, genauere Vorhersageergebnisse auf dekadischer Zeitskala zu erreichen.

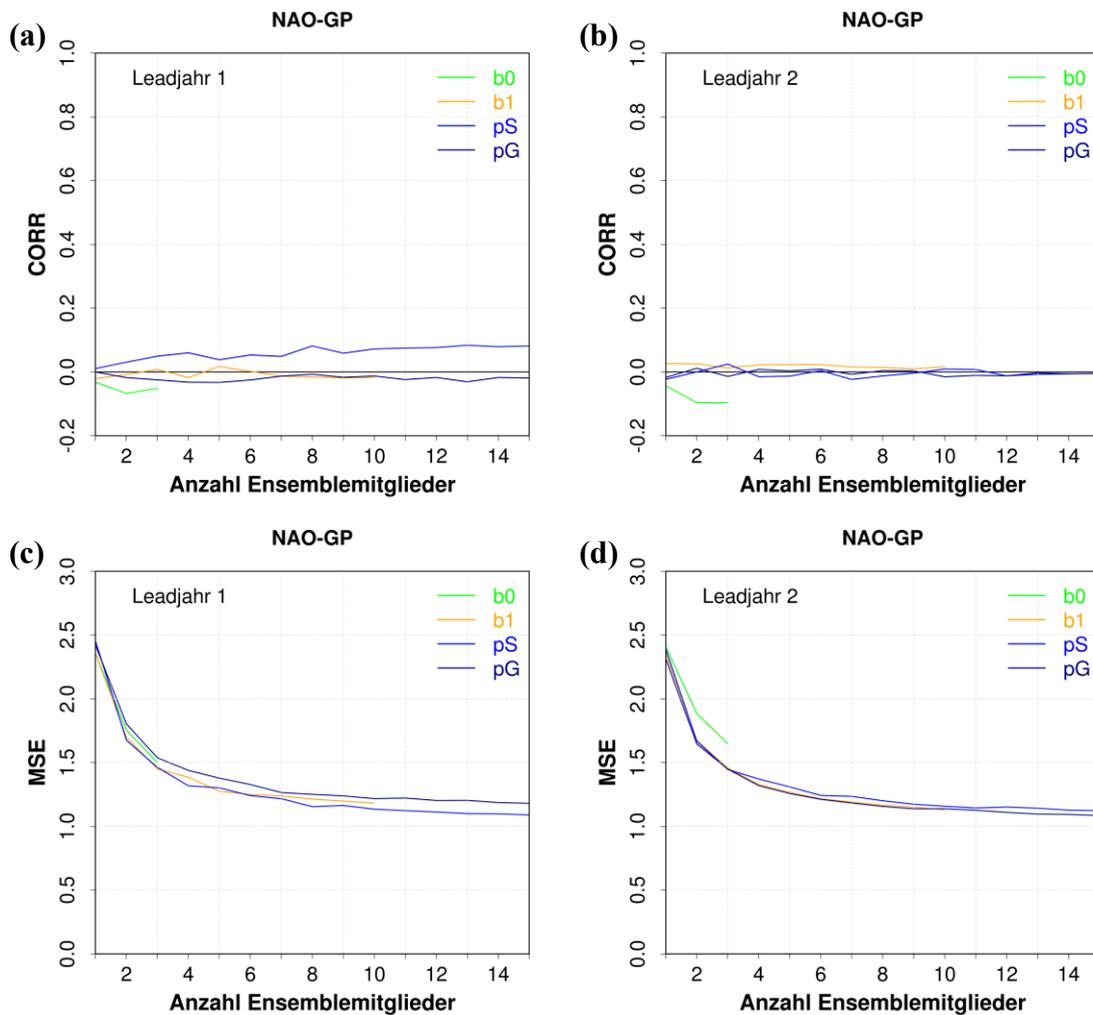
Im Hinblick auf die vorliegende Arbeit hat sich gezeigt, dass die Analyse einer Vielzahl an Telekonnektionsindizes und mit unterschiedlichen Klassifikationsverfahren ermittelten Zirkulationstypen zwar eine überblickende Darstellung zur dekadischen Vorhersagbarkeit der Dynamik der Atmosphäre erlaubt. Allerdings sind detaillierte Analysen und differenzierte Aussagen zu einzelnen Untersuchungsgegenständen nicht möglich. Außerdem konnte in der vorliegenden Arbeit keine genaue Ursachenforschung betrieben werden. In einem nächsten Schritt gilt es deshalb, potentielle Quellen der dekadischen Vorhersagbarkeit (wie z. B. SOI oder mitunter auch AAO) zu identifizieren und diese genauer zu analysieren. Es muss untersucht werden, worin die Ursachen dafür liegen und ob die guten Ergebnisse auch bei anderen dekadischen Klimamodellen beobachtet werden können. Außerdem können beispielsweise einzelne Modellläufe herausgegriffen werden, die eine besonders gute dekadische Vorhersageleistung offenbaren, und Untersuchungen zum Modellbias vorangetrieben werden, um eine weitere Ursachenanalyse zu betreiben.

Die skizzierten Untersuchungsansätze offenbaren ein weites Feld für anschließende Untersuchungsmöglichkeiten und zeigen auf, dass sich die Forschungsanstrengungen auf dem Gebiet der dekadischen Vorhersage erst im initialen Stadium befinden.

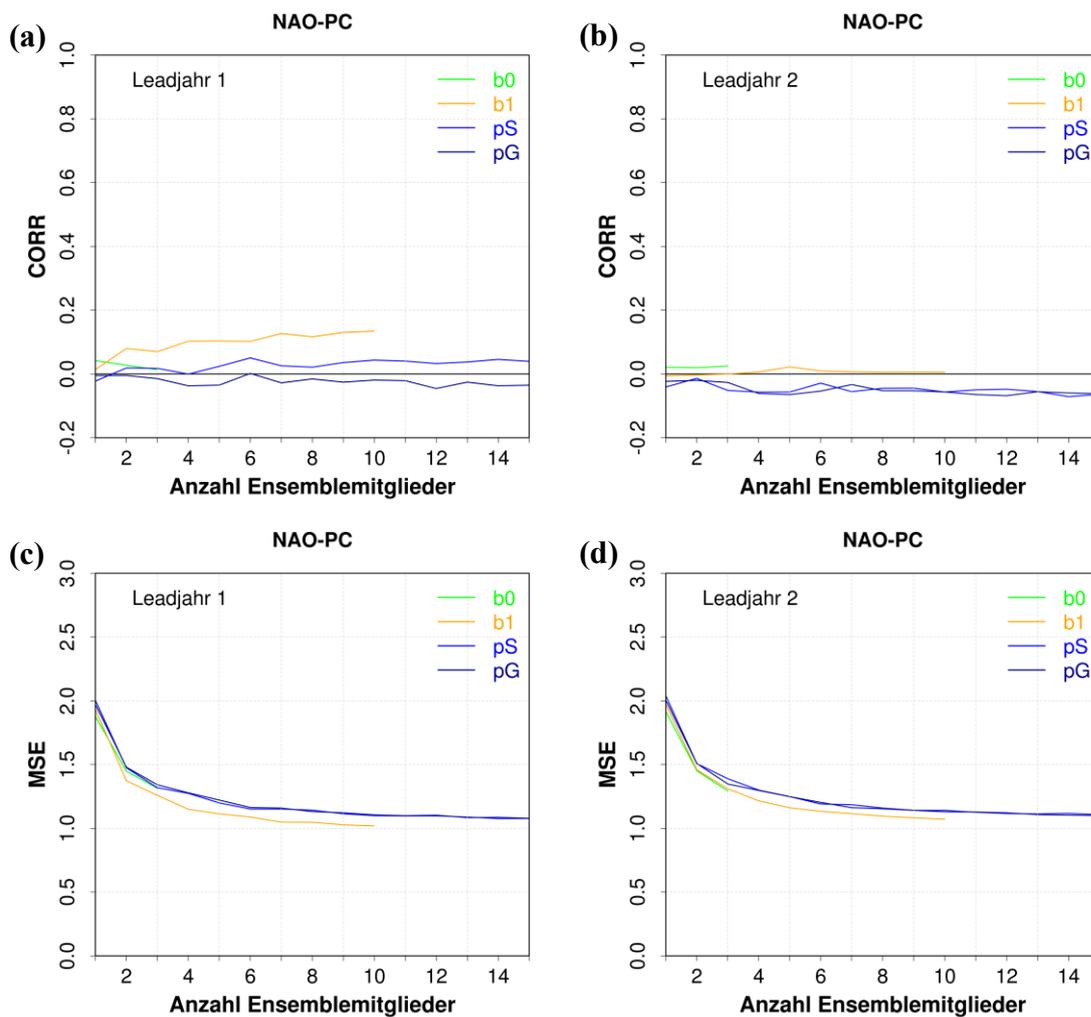
Anhang



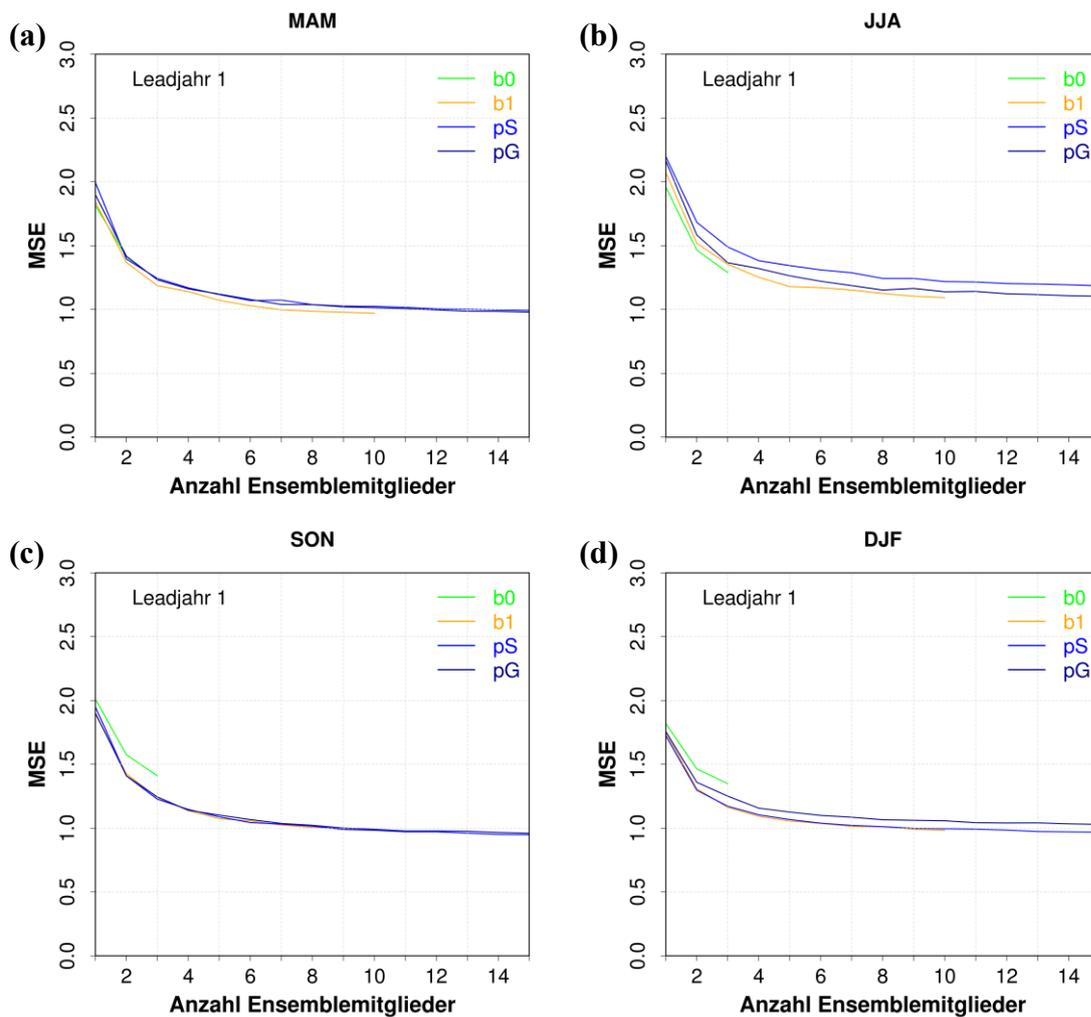
Anhang 1: Korrelation (obere Reihe: a, b) bzw. MSE (untere Reihe: c, d) zwischen den unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) und ERAINT hinsichtlich **AAO-ZM** in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (linke Hälfte) bzw. das zweite Leadjahr (rechte Hälfte) nach Initialisierung.



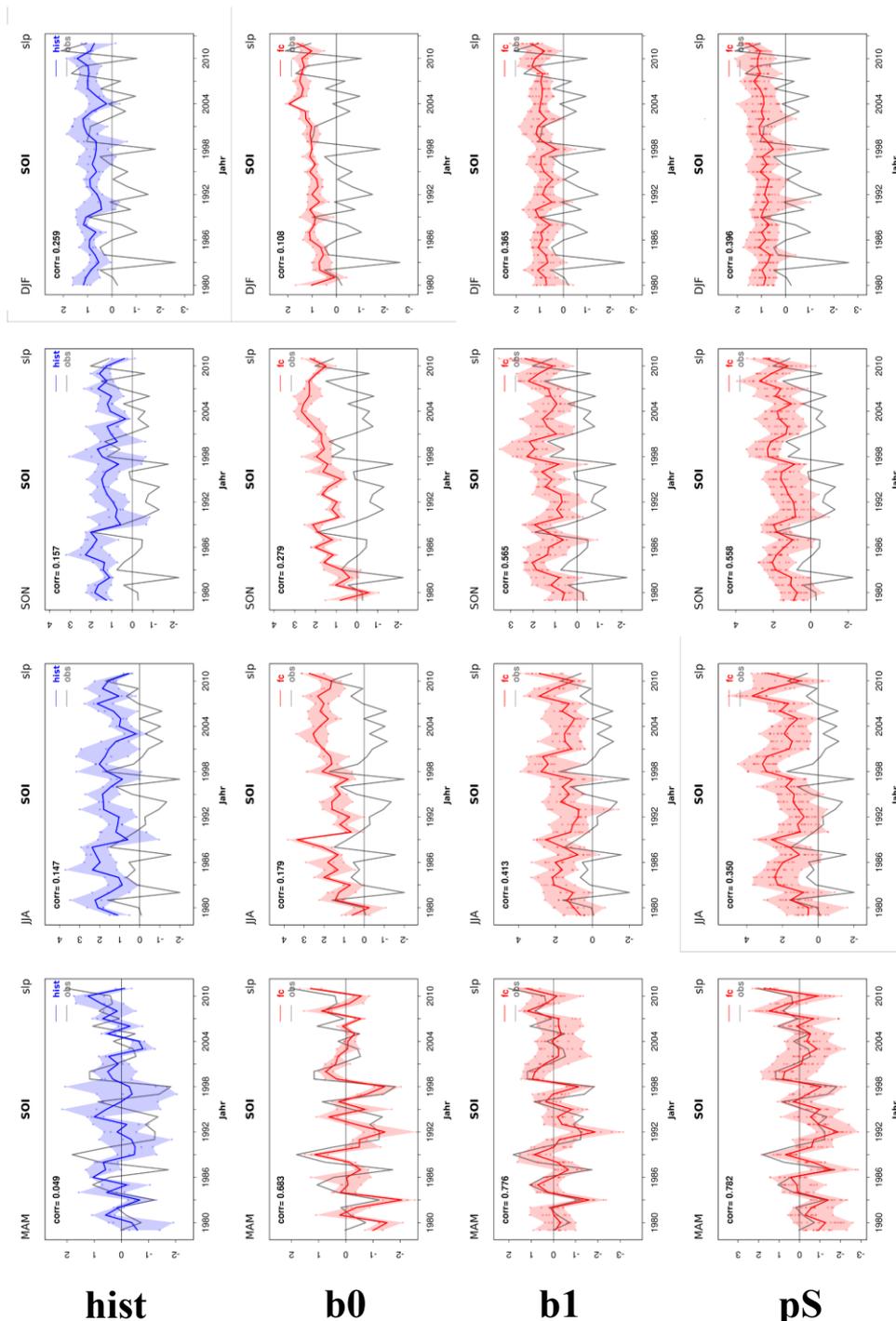
Anhang 2: Korrelation (obere Reihe: a, b) bzw. MSE (untere Reihe: c, d) zwischen den unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) und ERAINT hinsichtlich **NAO-GP** in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (linke Hälfte) bzw. das zweite Leadjahr (rechte Hälfte) nach Initialisierung.



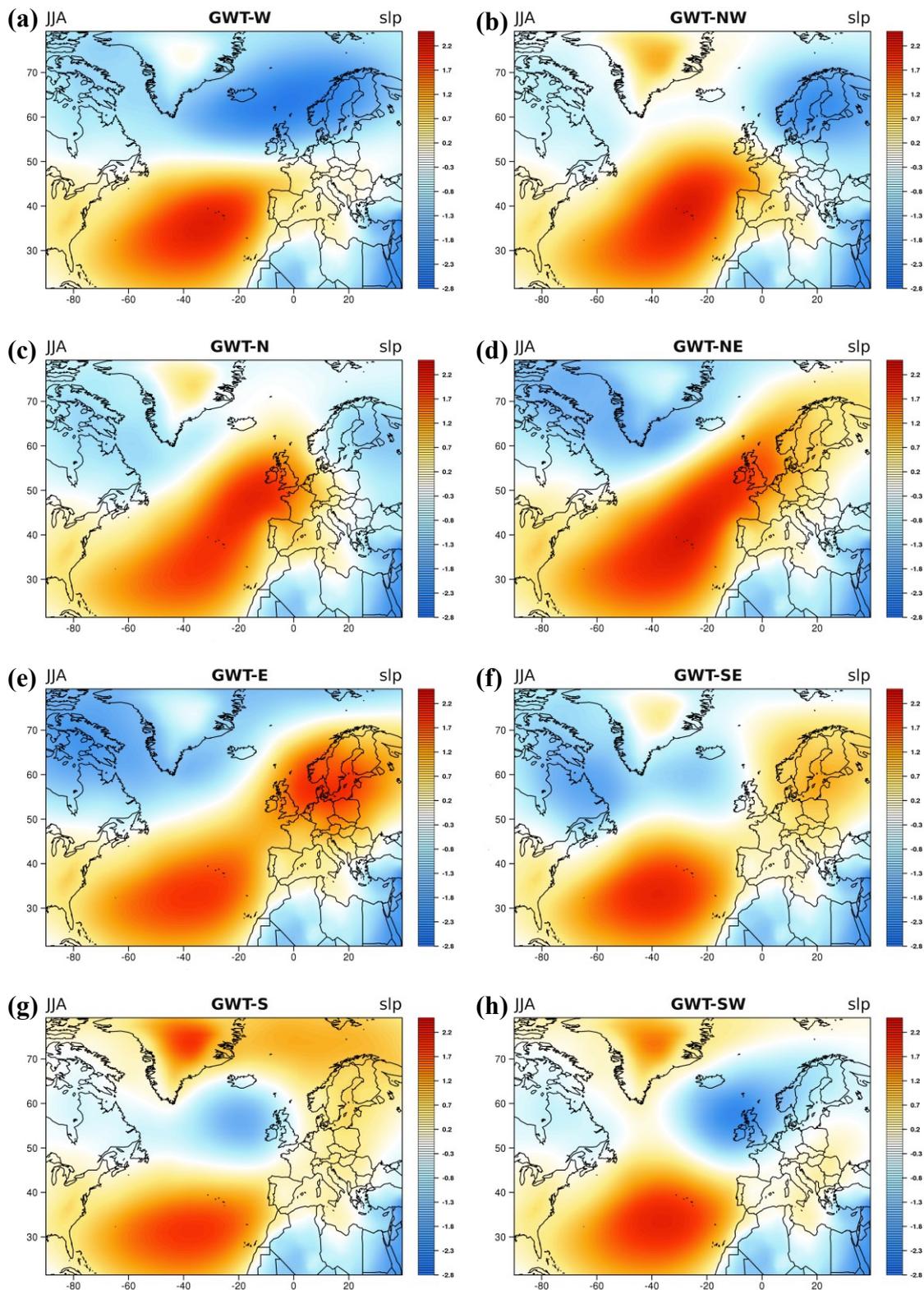
Anhang 3: Korrelation (obere Reihe: a, b) bzw. MSE (untere Reihe: c, d) zwischen den unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) und ERAINT hinsichtlich **NAO-PC** in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für das erste (linke Hälfte) bzw. das zweite Leadjahr (rechte Hälfte) nach Initialisierung.



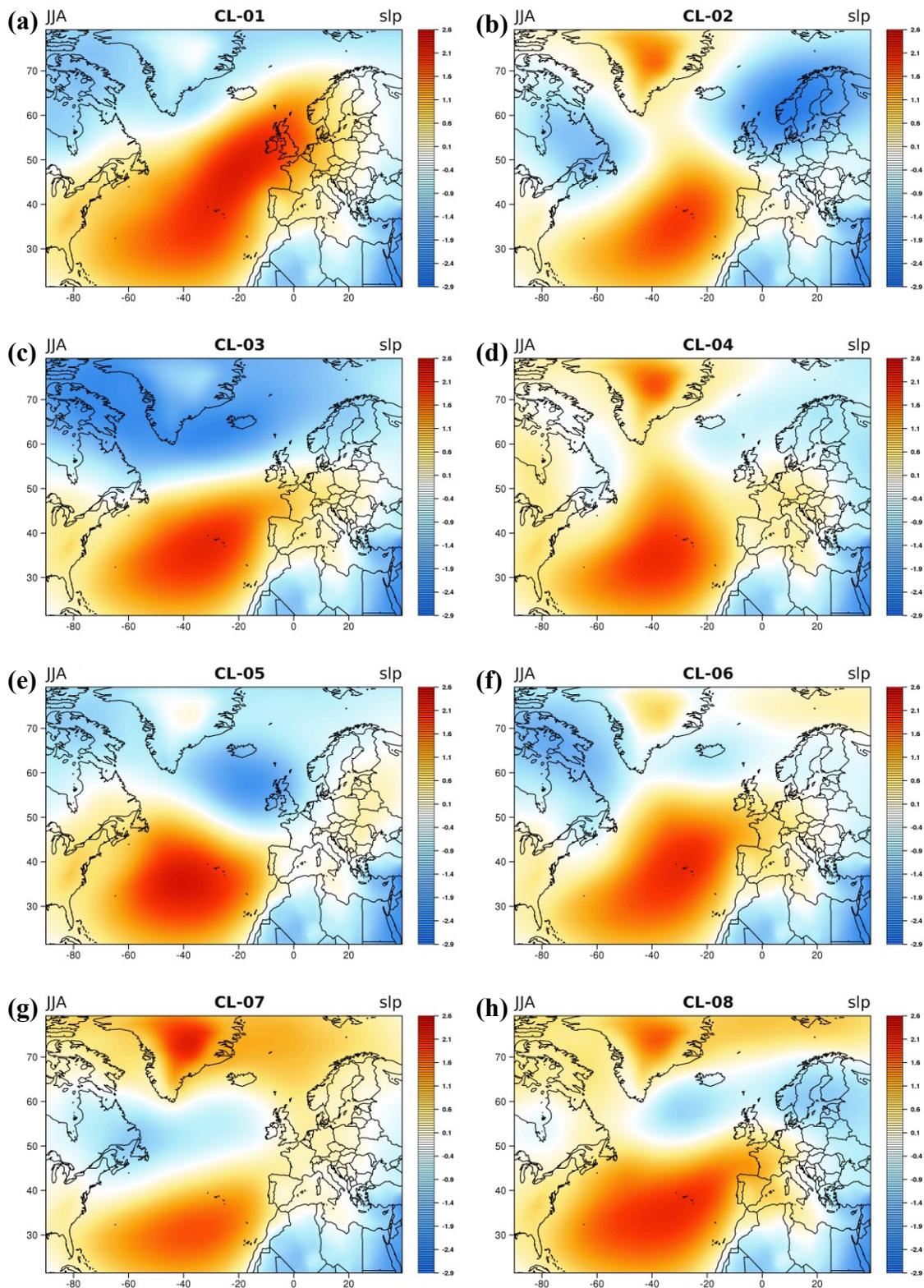
Anhang 4: MSE für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT hinsichtlich der ausgewählten Telekonnectionsindizes in Abhängigkeit von der Anzahl an Ensemblemitgliedern. Die Ergebnisse sind dargestellt für verschiedene Jahreszeiten (a: MAM, b: JJA, c: SON, d: DJF) des ersten Leadjahrs nach Initialisierung.



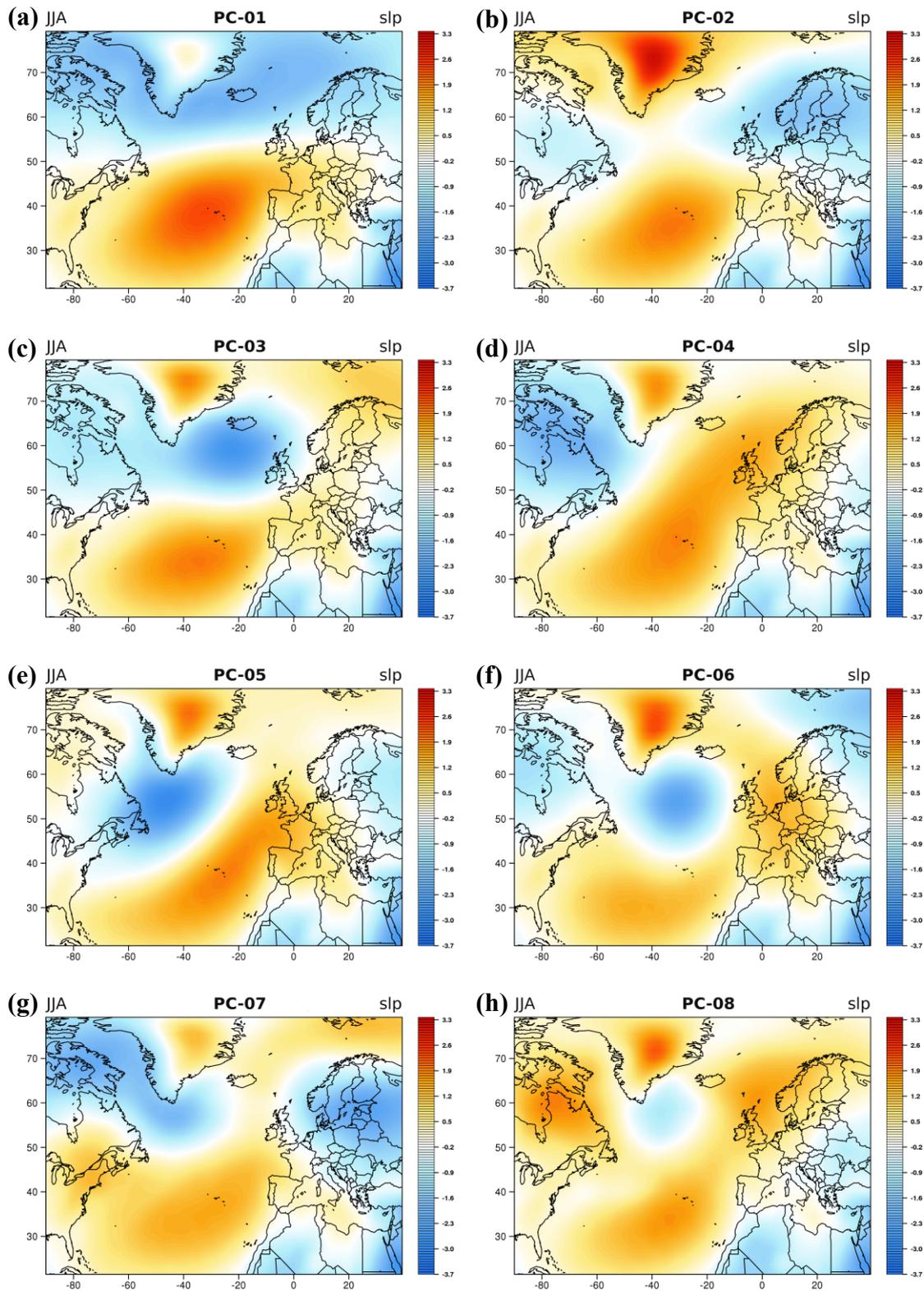
Anhang 5: Saisonaler aggregierter Southern Oscillation Index (SOI) des ersten Leadjahres im (von links nach rechts) Frühjahr (MAM), Sommer (JJA), Herbst (SON) und Winter (DJF) des Untersuchungszeitraums. Dargestellt sind die Ergebnisse für (von oben nach unten) **hist**, **b0**, **b1** und **pS** ggü. ERAINT. Die Punkte symbolisieren die einzelnen Ensemblemitglieder, die gefärbte Fläche die Spannweite aller Ensemblemitglieder, die durchgezogene Linie die über alle Ensemblemitglieder gemittelte Vorhersage. Der Korrelationskoeffizient drückt den Zusammenhang zwischen den gemittelten Werten der uninitialisierten Modellläufe und der Beobachtung aus.



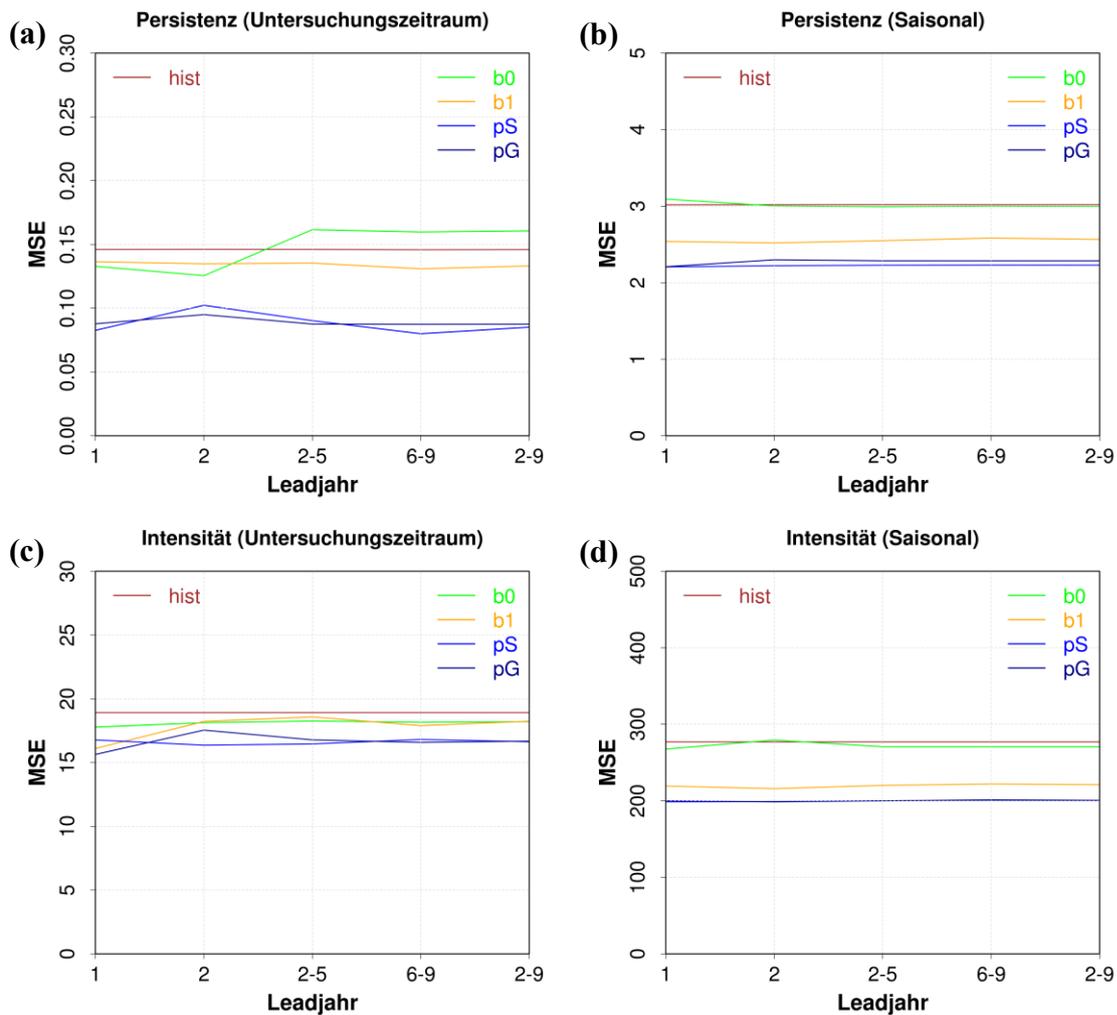
Anhang 6: Zentroide der Zirkulationstypen (basierend auf GWT) in der nordatlantisch-europäischen Region im Sommer (JJA) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Auftrittshäufigkeit (Datengrundlage: SLP-Anomalien, *hist*).



Anhang 7: Zentroide der Zirkulationstypen (basierend auf SAN) in der nordatlantisch-europäischen Region im Sommer (JJA) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Auftrittshäufigkeit (Datengrundlage: SLP-Anomalien, *hist*).



Anhang 8: Zentroide der Zirkulationstypen (basierend auf TPCA) in der nordatlantisch-europäischen Region im Sommer (JJA) des Untersuchungszeitraums (1979-2011) unter Angabe der relativen Auftrittshäufigkeit (Datengrundlage: SLP-Anomalien, *hist*).



Anhang 9: MSE hinsichtlich der Persistenz (obere Reihe: a, b) bzw. Intensität (untere Reihe: c, d) bei Aggregation über den gesamten Untersuchungszeitraum (linke Hälfte) bzw. bei saisonaler Aggregation (rechte Hälfte) für die unterschiedlichen Modellgenerationen des MPI-ESM-Modellsystems (gemittelt über verschiedene Analyseverfahren; Erläuterungen im Text) im Vergleich mit ERAINT in Abhängigkeit von verschiedenen Leadjahren.

Literaturverzeichnis

Backhaus, K., B. Erichson, W. Plinke und R. Weiber (2003): Multivariate Analysemethoden. *Springer*, Berlin.

Bahrenberg, G., E. Giese und J. Nipper (1999): Statistische Methoden in der Geographie. Band 1 – Univariate und bivariate Statistik. *Teubner*, Stuttgart.

Bahrenberg, G., E. Giese und J. Nipper (2003): Statistische Methoden in der Geographie. Band 2 – Multivariate Statistik. *Borntraeger*, Berlin/Stuttgart.

Balmaseda, M.A., K. Mogensen und A.T. Weaver (2013): Evaluation of the ECMWF ocean reanalysis system ORAS4. *Q.J.R. Meteorol. Soc.* **139**, 1132–1161.

Barnston, A.G. und R.E. Livezey (1987): Classification, Seasonality and Persistence of Low Frequency Atmospheric Circulation Patterns. *Mon. Weather Rev.* **115**, 1083-1126.

Bartzokas A. und D.A. Metaxas (1996): Northern Hemisphere gross circulation types. Climatic change and temperature distribution. *Meteorologische Zeitschrift* **5**, 99-109.

Beck C. (2000): Zirkulationsdynamische Variabilität im Bereich Nordatlantik-Europa seit 1780. *Würzburger Geographische Arbeiten*, Heft **95**.

- Beck, C., J. Jacobeit und A. Philipp (2001): Variability of North-Atlantic-European circulation patterns since 1780 and corresponding variations in Central European climate. – In: Detecting and Modelling regional climate change (ed. by Brunet-India, M., and D. López-Bonillo). *Springer*, 321-333.
- Beck, C., J. Jacobeit und P. D. Jones (2007): Frequency and within-type variations of large-scale circulation types and their effects on low-frequency climate variability in Central Europe since 1780. *Int. J. Climatol.* **27**, 473-491.
- Beck, C. und A. Philipp (2010): Evaluation and comparison of circulation type classifications for the European domain. *Physics and Chemistry of the Earth* **35**, 374-387.
- BOM (2012): Record-breaking La Niña events <http://www.bom.gov.au/climate/enso/history/La-Nina-2010-12.pdf> (04.08.2016).
- Compagnucci, R.H. und M.A. Salles (1997): Surface pressure patterns during the year over southern South America. *Int. J. Climatol.* **17**, 635–653.
- Compagnucci, R.H. und M.B. Richman (2008): Can principal component analysis provide atmospheric circulation or teleconnection patterns? *Int. J. Climatol.* **28**, 703-726.
- Dee, D.P., S.M. Uppala, A. Simmons, P. Berrisford, P. Poli et al. (2011): The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc* **137**, 553-597.

- Eade, R., D. Smith, A. Scaife, E. Wallace, N. Dunstone et al. (2014): Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys. Res. Lett.* **41**, 5620–5628.
- Enke, W., F. Schneider, T. Deutschländer (2005): A novel scheme to derive optimized circulation pattern classifications for downscaling and forecast purpose. *Theor. Appl. Climatol.* **82**, 51-63.
- Ferro, C.A.T., D.S. Richardson und A.P. Weigel (2008): On the effect of ensemble size on the discrete and continuous ranked probability scores. *Meteorol. Appl.* **15**, 19-24.
- Fukuoka, A. (1951): A study of 10-day forecast (A synthetic report). *Geophys. Mag.* **22**, 177-208.
- García-Serrano, J. und F.J. Doblas-Reyes (2012): On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. *Clim. Dyn.* **39**, 2025-2040.
- Giorgetta, M.A., J.H. JungCLAUS, C.H. Reick, S. Legutke, J. Bader et al. (2013): Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the coupled model intercomparison project phase 5. *J. Adv. Model. Earth Syst.* **5**, 572-579.
- Goddard, L., A. Kumar, A. Solomon, D. Smith, G. Boer et al. (2013): A verification framework for interannual-to-decadal predictions experiments. *Clim. Dyn.* **40**, 245-272.

Gong, D. und S. Wang (1999): Definition of Antarctic Oscillation Index. *Geophys. Res. Lett.* **26**, 459-462.

Hartigan, J. A. und M. A. Wong (1979): A K-means clustering algorithm. *Appl. Statistics* **28**, 100-108.

Hawkins, E., B. Dong, J. Robson, R. Sutton und D. Smith (2014): The Interpretation and Use of Biases in Decadal Climate Predictions. *J. Climate* **27**, 2931–2947.

Hess, P. und Brezowsky, H. (1977): Katalog der Großwetterlagen Europas 1881-1976. *Ber. d. Deutschen Wetterdienstes*, Offenbach a. M.

Hodson, D.L.R., R. T. Sutton, C. Cassou, N. Keenlyside, Y. Okumura, T. Zhou (2010): Climate impacts of recent multidecadal changes in Atlantic Ocean Sea Surface Temperature: a multimodel comparison. *Climate Dynamics* **34**, 1041-1058.

Hurrell, J.W. (1995): Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation. *Science* **269**, 676-679.

Hurrell, J. W., Y. Kushnir, G. Ottersen und M. Visbeck, M. (2003): The North Atlantic Oscillation: Climatic Significance and Environmental Impact. *American Geophysical Union*, Washington.

Hurrell, J. W., C. Deser (2009): North Atlantic climate variability: The role of the North Atlantic Oscillation. *Journal of Marine Systems* **78**, 28-41.

- Hurrell, J.W., T.L. Delworth, G. Danabasoglu, H. Drange, K. Drinkwater et al. (2010): Decadal Climate Prediction: Challenges and opportunities. *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society*.
- Huth, R. (1993): An example of using obliquely rotated principal components to detect circulation types over Europe. *Meteorologische Zeitschrift* **2**, 285–293.
- Huth, R. (1996): An intercomparison of computer assisted circulation classification methods. *Int. J Climatol* **16**, 893-922.
- Huth, R. (2000): Disaggregating Climatic trends by classification of circulation patterns. *Int. J Climatol* **21**, 135-153.
- Huth, R., C. Beck, A. Philipp, M. Demuzere, Z. Ustrnul et al. (2008): Classifications of atmospheric circulation patterns: recent advances and applications. *Ann. NY Acad. Sci.* **1146**, 105-152.
- Jacobeit, J., C. Beck und A. Philipp (1998): Annual to decadal variability in climate in Europe – objectives and results of the German contribution to the European climate research project ADVICE. *Würzburger Geographische Manuskripte* **43**, 163 pp.
- Jacobeit, J., P.D. Jones, T. Davies und C. Beck (2001a): Circulation changes in Europe since the 1780s. – In: *History and Climate: Memories of the future?* (ed. by P.D. Jones, A. Ogilvie, T. Davies und K. Briffa). Kluwer Academic/Plenum Publishers, 79-99.

Jacobeit, J., P. Jönsson, L. Bärring, C. Beck und M. Ekström (2001b): Zonal indices for Europe 1780-1995 and running correlations with temperature. *Climatic Change* **48**, 219-241.

Jacobeit, J., H. Wanner, J. Luterbacher, C. Beck, A. Philipp und K. Sturm (2003a): Atmospheric circulation variability in the North-Atlantic-European area since the mid-seventeenth century. *Climate Dynamics* **20**, 341-352.

Jacobeit, J., R. Glaser, J. Luterbacher und H. Wanner (2003b): Links between flood events in Central Europe since AD 1500 and large-scale atmospheric circulation modes. *Geophys. Res. Lett.* **30**, 1172-1175.

Jacobeit, J., A. Philipp und M. Nonnenmacher (2006): Atmospheric circulation dynamics linked with prominent discharge events in Central Europe. *Hydrological Sciences Journal* **51**, 946-965.

Jacobeit, J., J. Rathmann, A. Philipp und P.D. Jones (2009a): Central European precipitation and temperature extremes in relation to large-scale atmospheric circulation types. *Meteorologische Zeitschrift*, **18(4)**, 397-410.

Jacobeit, J., A. Philipp, J. Rathmann und A. Walther (2009b): European and North Atlantic daily to multidecadal climate variability - General Overview and Final Reports for the German contribution to the European Climate Research Project EMULATE. *Geographica Augustana* **7**, 65 pp.

- Jacobeit, J. (2010). Classifications in Climate Research. *Physics and Chemistry of the Earth* **35**, 411-421.
- Jolliffe, I.T. (2002): Principal Component Analysis. *Springer*, New York.
- Jones, P.W. (1999): First-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Mon. Weather Rev.* **127**, 2204-2210.
- Jungclaus, J.H., N. Fischer, H. Haak, K. Lohmann, J. Marotzke et al. (2013): Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *J. Adv. Model. Earth Syst.* **5**, 422-446.
- Kadow, C., S. Illing, O. Kunst, H.W. Rust, H. Pohlmann et al. (2015): Evaluation of forecasts by accuracy and spread in the MiKlip decadal climate prediction system *Meteorologische Zeitschrift* **25**, 631–643.
- Keenlyside, N., M. Latif, J. Jungclaus, L. Kornblueh und E. Roeckner (2008): Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature* **453**, 84–88.
- Köhl, A. (2015): Evaluation of the GECCO2 ocean synthesis: transports of volume, heat and freshwater in the Atlantic. *Q.J.R. Meteorol. Soc.* **141**, 166–181.
- Kruschke, T., H.W. Rust, C. Kadow, W.A. Müller, H. Pohlmann, G.C. Leckebusch und U. Ulbrich (2015): Probabilistic evaluation of decadal prediction skill regarding

Northern Hemisphere winter storms. *Meteorologische Zeitschrift*, doi:
10.1127/metz/2015/0641.

Kumar, A. (2009): Finite samples and uncertainty estimates for skill measures for seasonal prediction. *Mon. Weather Rev.* **137**, 2622–2631.

Lauer, W und J. Bendix (2006): *Klimatologie*. Westermann, Braunschweig.

Marotzke, J., W. Müller, F. Vamborg, P. Becker, U. Cubasch, H. Feldmann et al. (2016): MiKlip – a National Research Project on Decadal Climate Prediction. *Bull. Am. Meteorol. Soc.*, in press.

Meehl, G.A., L. Goddard, J. Murphy, R.J. Stouffer, G. Boer et al. (2009): Decadal Prediction. Can it be skilful? *Bull. Am. Meteorol. Soc.* **90**, 1467-1485.

Meehl, G.A., L. Goddard, G. Boer, R. Burgman, G. Branstator et al. (2014): Decadal climate prediction: An update from the trenches. *Bull. Am. Meteorol. Soc.* **95**, 243-267.

Mo, K.C. (2000): Relationships between Low-Frequency Variability in the Southern Hemisphere and Sea Surface Temperature Anomalies. *J. Climate* **13**, 3599-3610.

Müller, W.A., J. Baehr, H. Haak, J.H. Jungclaus, J. Kröger et al. (2012): Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology. *Geophys. Res. Lett.* **39**, doi: 10.1029/2012GL053326.

- Murphy, J., V. Kattsov, N. Keenlyside, M. Kimoto, G. Meehl et al. (2010): Towards Prediction of Decadal Climate Variability and Change. *Procedia Environmental Sciences* **1**, 287-304.
- Nan, S. und J. Li (2003): The relationship between the summer precipitation in the Yangtze River valley and the boreal spring Southern Hemisphere annular mode. *Geophys. Res. Lett.* **30**, 2266-2269.
- Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi et al. (2007): The JRA-25 reanalysis. *Journal of the Meteorological Society of Japan* **85**, 369-432.
- Pattantyús-Ábrahám, M und Steinbrecht, W. (2015): Temperature Trends over Germany from Homogenized Radiosonde Data. *J. Climate* **28**, 5699–5715.
- Philipp, A., P. Della Marta, J. Jacobeit, D. Fereday, P.D. Jones, A. Moberg und H. Wanner (2007): Long-term variability of daily North-Atlantic–European Pressure Patterns since 1850 classified by Simulated Annealing Clustering. *J. Climate* **20**, 4065-4095.
- Philipp A. (2009): Comparison of principal component and cluster analysis for classifying circulation pattern sequences for the European domain. *Theoretical and Applied Climatology* **96**, 31-41.
- Philipp A. et al. (2010): COST733CAT – a database of weather and circulation type classifications. *Physics and Chemistry of the Earth* **35**, 360-373.

Philipp A., C. Beck, R. Huth und J. Jacobeit (2014): Development and comparison of circulation type classifications using the COST 733 dataset and software. *Int. J. Climatol*, doi: 10.1002/joc.3920.

Philipp A., C. Beck, R. Huth und J. Jacobeit (2016): Development and comparison of circulation type classifications using the COST 733 dataset and software. *Int. J. Climatol.* **36**, 2673–2691.

Pohlmann, H., J.H. Jungclaus, A. Köhl, D. Stammer und J. Marotzke (2009): Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic. *J. Climate* **22**, 3926-3938.

Pohlmann, H., W.A. Müller, K. Kulkarni, M. Kameswarrao, D. Matei et al. (2013): Improved forecast skill in the tropics in the new MiKlip decadal climate predictions. *Geophys. Res. Lett.* **40**, 5798–5802

Rathmann, J. (2009): Klima-und Zirkulationsvariabilität im südhemisphärischen Afrika seit Beginn des 20. Jahrhunderts. *Dissertation*, Universität Augsburg.

Richman, M.B. (1981): Obliquely rotated principal components: An improved meteorological map typing technique? *J. Appl. Meteor.* **20**, 1145–1159.

Richman, M.B. (1986): Rotation of principal components. *Journal of Climatology* **6**, 293-335.

- Rienecker, M.M., M.J. Suarez, R. Gelaro, R. Todling, L. Bacmeister et al. (2011): MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate* **24**, 3624-3648.
- Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang et al. (2010): The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.* **91**, 1015-1057.
- Scaife, A.A., A. Arribas, E. Blockley, R.T. Clark, N. Dunstone et al. (2014), Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.* **41**, 2514–2519.
- Schönwiese, C.-D. (2013): *Klimatologie*. Ulmer, Stuttgart.
- Sienz, F., W.A. Müller und H. Pohlmann (2016): Ensemble size impact on the decadal predictive skill assessment. *Meteorologische Zeitschrift* **25**, 645-655.
- Smith, D.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**, 796-799.
- Smith, D.M., A.A. Scaife und B.P. Kirtman (2012): What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? *Environ. Res. Lett.* **7**.
- Smith, D. M., Scaife, A. A., Eade, R. und Knight, J. R. (2014): Seasonal to decadal prediction of the winter North Atlantic Oscillation: emerging capability and future prospects. *Q.J.R. Meteorol. Soc.*, doi:10.1002/qj.2479.

- Solomon, A., L. Goddard, A. Kumar, J. Carton, C. Deser et al. (2011): Distinguishing the roles of natural and anthropogenically forced decadal climate variability. *Bull. Amer. Meteor. Soc.* **92**, 141-156.
- Stevens, S., M. Giorgetta, M. Esch, T. Mauritsen, T. Crueger, S. Rast et al. (2013): Atmospheric component of the MPI-M Earth System Model: ECHAM6. *J. Adv. Model. Earth Syst.* **5**, 146-172.
- Taylor, K.E., R.J. Stouffer und G.A. Meehl (2012): An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485-498.
- Thompson, D., und J. M. Wallace (2000): Annular modes in the extratropical circulation. Part I: Month-to-month variability. *J. Climate* **13**, 1000-1016.
- Trenberth, K.E. (1984): Signal Versus Noise in the Southern Oscillation. *Mon. Weather Rev.* **112**, 326-332.
- Uppala, S.M., P.W. Kallberg, A.J. Simmons, U. Andrae, V. Bechtold et al. (2005): The ERA-40 re-analysis. *Q.J.R. Meteorol. Soc.* **131**, 2961–3012.
- von Storch, H. und W. Zwiers (1999): Statistical Analysis in Climate Research. *Cambridge University Press*, Cambridge.
- Wallace, J.M. und David S. Gutzler (1981): Teleconnections in the Geopotential Height Field during der Northern Hemisphere Winter. *Mon. Weather Rev.* **109**, 784-812.

Wilks, D.S. (2006): *Statistical Methods in the Atmospheric Sciences*. Elsevier, Amsterdam.

Yarnal, B. (1993): *Synoptic climatology in environmental analysis*. Belhaven Press, London.