

Analyse von klimatologischen Tagesdaten mithilfe von Markov Ketten

Diplomarbeit

von

Heiko Steiner

vorgelegt beim Institut für Atmosphäre und Umwelt
der Johann Wolfgang Goethe-Universität in Frankfurt am Main

April 2007

Zusammenfassung

In der vorliegenden Arbeit wurden so genannte Markov Ketten auf 301 deutsche Stationszeitreihen, die tägliche Niederschlagswerte liefern, im Zeitraum 1951 bis 2000 angewendet. Markov Ketten werden charakterisiert durch eine Übergangsmatrix, deren Elemente die einzelnen Übergangswahrscheinlichkeiten sind, und durch ihre Ordnung, die mittels dem Bayesian Information Criterion (BIC) sowie dem Akaike Information Criterion (AIC) festgestellt werden kann. Die Übergangswahrscheinlichkeiten beschreiben, mit welcher Wahrscheinlichkeit der Wechsel zwischen Zuständen auftritt, und legen somit stochastisch zusammen mit der Ordnung der Markov Kette den zeitlichen Verlauf der Markov Kette fest. Die Ordnung beinhaltet die Aussage, ob bestimmte Zustände (un-)abhängig von vergangenen Zuständen ist. Beispielsweise sind bei der Ordnung 0 die Tage unabhängige Ereignisse, während bei der Ordnung 1 der vergangene Zustand Informationen für den nächsten Tag beinhaltet. Es werden nur zeit-diskrete Markov Ketten betrachtet, d. h. jedem Tag wird ein einziger Zustand zugeordnet. Dafür müssen die Stationszeitreihen in eine Ordinalskala transformiert werden. Die Niederschlagsdaten sind zum einen mit zwei Zuständen untersucht worden, wobei zwischen Regen- und Trockentag unterschieden wird, und zum anderen mit vier Zuständen, die durch die Quantile der angepassten Weibullverteilung ermittelt wurden. Die Anpassung fand für den Gesamtzeitraum (1951-2000) stationsweise und für jeden Monat einzeln statt. Die vier Zustände sind ein Trockentag sowie Tage mit wenig, mittlerer und hoher Niederschlagsmenge. Dabei wurde die maximal mögliche Ordnung auf $m = 3$ begrenzt. Die Auswertung fand dabei saisonal nach den meteorologischen Jahreszeiten getrennt statt.

In einem ersten Schritt sind die Übergangswahrscheinlichkeiten und Ordnungen für den Gesamtzeitraum sowohl für jede Station einzeln als auch für alle 301 Stationen gemeinsam berechnet worden. Anschließend wurde der Zeitraum halbiert und für beide Hälften die Ordnungen bestimmt. Die Ordnung der Markov Kette hängt von der geografischen Lage, der Jahreszeit und den verwendeten Auswahlkriterien ab und kann nur die mittlere Verteilung von Zuständen beschreiben. Daher ist auch die Länge der Zeitreihe ein wichtiger Faktor für die Auswahl der Ordnungen. Darüber hinaus sind die Übergangswahrscheinlichkeiten für jedes Jahr einzeln berechnet worden. Damit ist es möglich, Aussagen über die zeitliche

Entwicklung der Übergangswahrscheinlichkeiten in Form von linearen Trends zu treffen.

Die Analyse der Niederschlagsdaten bei Verwendung von zwei Zuständen ergibt, dass nach dem BIC für fast alle Stationen Markov Ketten der Ordnung 2 angepasst werden. Auch die Halbierung des Zeitraums erbringt dasselbe Ergebnis und damit keine Veränderungen in den Ordnungen. Nach dem BIC zeigt sich kein saisonaler Unterschied in den Ordnungen. Ebenso können in den Übergangswahrscheinlichkeiten 2. Ordnung kaum Trends ausgemacht werden. Jedoch hat im Sommer die Wahrscheinlichkeit, dass nach zwei Trockentagen wieder ein Trockentag folgt, vor allem im Nordwesten Deutschlands um bis zu 20% signifikant zugenommen. Im Winter wird die Serie von drei Regentagen an vielen Stationen wahrscheinlicher. Andere Stationen zeigen aber auch signifikante Abnahme.

Die Ergebnisse der Analyse mit vier Zuständen sind vielfältiger. Zunächst wird nach dem BIC für den Gesamtzeitraum im Winter die dritte Ordnung detektiert und in den übrigen Jahreszeiten jedoch nur die zweite Ordnung. Bei Halbierung des Gesamtzeitraums ist die wichtigste Erkenntnis, dass im Winter ein Ordnungswechsel von der dritten zur zweiten Ordnung und im Frühling von der ersten zur zweiten Ordnung stattgefunden hat. Der Sommer und Herbst zeigen ein gemischtes Bild. Hier treten Wechsel von der ersten zur zweiten Ordnung und umgekehrt auf. Trends konnten nur für die nullte und erste Ordnung berechnet werden, da bei höheren Ordnungen ein Großteil der Übergangswahrscheinlichkeiten Null ist. Die Ergebnisse der Trendanalyse fallen aufgrund der vergleichsweise hohen Niederschlagsvariabilität sehr unterschiedlich aus. So treten im Sommer hauptsächlich Abnahmen und im Winter Zunahmen der sehr feuchten Tage auf. Stärkste Zunahmen betreffen die Wahrscheinlichkeit, dass nach einem sehr feuchten Tag ein Trockentag folgt. Einzelne Stationen können aber auch davon abweichende Trends aufweisen.

Inhaltsverzeichnis

Zusammenfassung	III
Abbildungsverzeichnis	VII
1 Einleitung	1
2 Theorie der Markov Ketten	3
2.1 Grundbegriffe: Stochastischer Prozess, Zeitreihe und Markov-Urnenmodell	3
2.2 Markov Kette	7
2.3 Transformation in eine Ordinalskala	9
2.4 Klassifizierung der Zustände	11
2.5 Ergodische Markov Ketten	12
2.6 Selektionskriterien für die Modell-Auswahl	13
2.7 Beispielrechnung für eine Markov Kette 1. Ordnung mit zwei Zuständen . .	14
3 Datengrundlage	19
4 Methodik	23
5 Ergebnisse und Interpretation	25
5.1 Ergebnisse der Analyse von den Niederschlagsdaten mit zwei Zuständen . .	25
5.1.1 Zur stationären Wahrscheinlichkeit, dem Persistenzparameter und den Trends der Regentage	25
5.1.2 Detektierte Ordnungen der Markov Ketten	29
5.1.3 Übergangswahrscheinlichkeiten der Markov Ketten	31
5.1.4 Zeitliche Veränderungen der Übergangswahrscheinlichkeiten	41
5.2 Ergebnisse der Analyse für Niederschlagsdaten mit vier Zuständen	47
5.2.1 Ergebnisse der Anpassung der Weibullverteilung	47
5.2.2 Detektierte Ordnungen der Markov Ketten	50
5.2.3 Übergangswahrscheinlichkeiten der Markov Ketten	54
5.2.4 Zeitliche Veränderungen der Übergangswahrscheinlichkeiten	64

6 Wertung und Ausblick	71
Literaturverzeichnis	73
A Statistische Grundlagen	75
A.1 Axiome der Wahrscheinlichkeitsrechnung	75
A.2 Theoretische Verteilungen	76
A.2.1 Weibull Verteilung	76
A.2.2 Anpassung einer Weibullverteilung	77
A.3 Statistische Tests	79
A.3.1 Test des Regressionskoeffizienten	79
A.3.2 Kolmogorov-Smirnov Test	79
A.3.3 Mann-Kendall Trendtest	79
B Ergänzende Karten	81
B.1 Niederschlag mit zwei Zuständen	81
B.1.1 Ordnung der Markov Ketten nach dem AIC	81
B.1.2 Ordnung der Markov Ketten nach dem BIC	83
B.1.3 Karten zu den Übergangswahrscheinlichkeiten zu Kap. 5.1.2	85
B.1.4 Trends der Übergangswahrscheinlichkeiten	92
B.2 Niederschlag mit vier Zuständen	97
B.2.1 Ordnung der Markov Ketten nach dem AIC	97
B.2.2 Karten zu den Übergangswahrscheinlichkeiten zu Kap. 5.2.3	100
B.2.3 Trends der Übergangswahrscheinlichkeiten	103
Danksagung	105
Erklärung	106

Abbildungsverzeichnis

2.1	Realisation einer Zufallsvariable X_t	5
2.2	Venn Diagramm für den Zustandsraum	6
2.3	Prozessdiagramm einer reduziblen Markov Kette	13
2.4	Prozessdiagramm für eine ergodische Markov Kette	13
3.1	geografische Verteilung der Niederschlagsstationen	19
3.2	mittlerer Jahresniederschlag (1951-2000) in [mm] generiert von den verwendeten 301 Niederschlagsstationen für das ganze Jahr (oberes Bild) und die vier Jahreszeiten	20
3.3	Karte der Beispiel-Stationen	21
5.1	stationäre Wahrscheinlichkeit π_1 für das Eintreten von Regentagen	26
5.2	Persistenz Parameter r_1	27
5.3	Trend nach Mann-Kendall Test für die Anzahl n_1 der Trockentage	28
5.4	Ordnungen nach dem BIC für zwei Zustände	29
5.5	Vergleich der Ordnungen nach AIC mit zwei Zuständen für den Sommer	30
5.6	Vergleich der Ordnungen nach AIC mit zwei Zuständen für den Herbst	31
5.7	Wahrscheinlichkeit p_1 (1951-2000)	32
5.8	Wahrscheinlichkeit p_{01} (1951-2000)	33
5.9	Wahrscheinlichkeit p_{11} (1951-2000)	34
5.10	mittlerer Jahresgang für alle 301 Niederschlagsstationen	35
5.11	Jahresgang von p_0, p_1, p_{00} und p_{11}	39
5.12	linearer Trend und Häufigkeiten für p_{000} für Bremen im Sommer	43
5.13	Steigung der Regressionsgeraden für p_{000}	44
5.14	Steigung der Regressionsgeraden für p_{111}	45
5.15	Verteilungsfunktionen der Weibullverteilung für die Station Altenahr im März und im Juni für den Gesamtzeitraum 1951-2000	47
5.16	Signifikanz der Anpassung der Weibullverteilung	48

5.17 Trends der Anzahl der Zustände 1 (obere Reihe), 2 (mittlere Reihe) und 3 (untere Reihe) jeweils für Frühling, Sommer, Herbst und Winter (von links nach rechts) berechnet für 1951-2000	50
5.18 Ordnung der Markov Ketten nach BIC mit vier Zuständen (1951-2000)	51
5.19 Vergleich der Ordnung zwischen 1951-1975 und 1976-2000 für Frühling und Sommer	52
5.20 Vergleich der Ordnung zwischen 1951-1975 und 1976-2000 für Herbst und Winter	53
5.21 Wahrscheinlichkeit p_{03} (1951-2000) für vier Zustände	56
5.22 Wahrscheinlichkeit p_{30} (1951-2000) für vier Zustände	56
5.23 Wahrscheinlichkeit p_{33} (1951-2000) für vier Zustände	57
5.24 Wahrscheinlichkeit p_{010} (1951-2000) für vier Zustände	58
5.25 Wahrscheinlichkeit p_{013} (1951-2000) für vier Zustände	58
5.26 Wahrscheinlichkeit p_{3330} (1951-2000) für vier Zustände	59
5.27 Wahrscheinlichkeit p_{3333} (1951-2000) für vier Zustände	60
5.28 Jahresgang für vier Zustände	61
5.29 zeitlicher Verlauf von p_{3x}	65
5.30 Steigung von p_{03}	66
5.31 Steigung von p_{30}	67
5.32 Steigung von p_{33}	68
5.33 Steigung von p_{003}	69
5.34 Steigung von p_{0003}	70
A.1 Wahrscheinlichkeitsdichtefunktion der Weibull Verteilung	76
B.1 Ordnung nach dem AIC für zwei Zustände (1951-2000)	81
B.2 Vergleich der Ordnungen nach AIC für den Frühling	82
B.3 Vergleich der Ordnungen nach AIC für den Winter	82
B.4 Vergleich der Ordnungen nach dem BIC für zwei Zustände für Frühling und Sommer	83
B.5 Vergleich der Ordnungen nach dem BIC für zwei Zustände für Herbst und Winter	84
B.6 Wahrscheinlichkeit p_0 (1951-2000)	85
B.7 Wahrscheinlichkeit p_{000} (1951-2000)	86
B.8 Wahrscheinlichkeit p_{011} (1951-2000)	87
B.9 Wahrscheinlichkeit p_{101} (1951-2000)	88
B.10 Wahrscheinlichkeit p_{111} (1951-2000)	89
B.11 Wahrscheinlichkeit p_{0000} (1951-2000)	90

B.12	Wahrscheinlichkeit p_{1111} (1951-2000)	91
B.13	linearer Trend für p_{001} , p_{011} , p_{101} und p_{111} berechnet für alle Stationen	93
B.14	linearer Trend für p_{001} , p_{011} , p_{101} und p_{111} für Bremen	94
B.15	linearer Trend für p_{001} , p_{011} , p_{101} und p_{111} berechnet für Frankfurt	95
B.16	linearer Trend für p_{001} , p_{011} , p_{101} und p_{111} für Zugspitze	96
B.17	Ordnung der Markov Ketten mit vier Zuständen nach dem AIC (1951-2000)	97
B.18	Vergleich der Ordnung nach dem AIC zwischen 1951-1975 und 1976-2000 für Frühling und Sommer	98
B.19	Vergleich der Ordnung nach AIC zwischen 1951-1975 und 1976-2000 für Herbst und Winter	99
B.20	Wahrscheinlichkeit p_1 (1951-2000) für vier Zutände	100
B.21	Wahrscheinlichkeit p_2 (1951-2000) für vier Zutände	101
B.22	Wahrscheinlichkeit p_3 (1951-2000) für vier Zutände	102

1 Einleitung

Nicht alle Vorgänge in der Natur laufen immer in festen, geordneten Bahnen ab. Jedoch lassen sich einzelne, grundlegende Eigenschaften eines Systems oftmals mit einfachen Gesetzmäßigkeiten beschreiben, aber das System selbst verhält sich nicht deterministisch sondern chaotisch. Wenn ein System untersucht werden soll, interessiert man sich dafür, in welchem Zustand es sich zum Zeitpunkt der Beobachtung befindet [5]. Oder aber man interessiert sich dafür, mit welcher Wahrscheinlichkeit es in einem bestimmten Zustand in naher Zukunft sein wird. Als Beispiel kann man ein Wettersystem mit den Zuständen $\{\text{Auftreten von Niederschlag}\}$ und $\{\text{Nicht-Auftreten von Niederschlag}\}$ betrachten. Die Reihenfolge beider Zuständen ist nicht zwangsläufig festgelegt, folgt aber sehr wohl physikalischen Eigenschaften. Zur Beschreibung solcher Systeme bieten sich statistische Modellierungen an wie die der **Markov Modelle**, die auch **Markov Ketten** genannt werden. Die Theorie der Markov Ketten geht auf MARKOV¹ zurück. Er untersuchte die Buchstabenabfolge der russischen Sprache und erkannte dabei bestimmte stochastische Strukturen, die die Reihenfolge von Vokalen und Konsonanten beschreiben [10].

Markov Ketten werden in der stationsbezogenen Vorhersage der täglichen Niederschlagswahrscheinlichkeiten aber auch anderer meteorologischen Messgrößen wie Wolkenart und Windextrema angewendet (z. B. [15, 23]). Darüberhinaus werden Markov Ketten heute eingesetzt, um verschiedene Prozesse stochastisch zu modellieren wie z. B. die Brownsche Molekularbewegung, Diffusionserscheinungen und demografische Entwicklungen (Geburts-, Absterbe- und Einwanderungsprozesse) [17].

Gabriel und Neumann [7] gehören zu den ersten, die ein Markov Modell einsetzten, um den täglichen Niederschlag an einer Station in Tel Aviv zu modellieren. Sie stellten fest, dass eine Markov Kette 1. Ordnung gut geeignet ist, den täglichen Niederschlag an der besagten Station zu modellieren. Das bedeutet, dass die Kenntnis des heutigen (Wetter-)Zustands ausreicht, um den morgigen Zustand statistisch zu prognostizieren.

¹A. A. Markov (1856-1922), russ. Mathematiker, Arbeiten auf dem Gebiet der Wahrscheinlichkeitstheorie und Analysis

In der vorliegenden Arbeit wird untersucht, wie sich tägliche Niederschlagsreihen mithilfe von Markov Ketten im Zeitraum 1951-2000 in Deutschland beschreiben lassen. Analysiert wird auch, ob sich regionale und/oder saisonale Unterschiede zeigen und ob sich signifikante Veränderungen im Beobachtungszeitraum ergeben haben. Im nächsten Kapitel (Kap. 2) wird die Theorie der Markov Ketten vorgestellt, ehe dann in Kapitel 3 die Datengrundlage behandelt wird. Daran anschließend wird in Kapitel 4 Analysemethodik etwas näher erläutert und abschließend werden in den Kapiteln 5 und 6 die Ergebnisse präsentiert und interpretiert.

2 Theorie der Markov Ketten

In diesem Kapitel werden die Grundlagen für die Markov Ketten beschrieben. Hierfür werden zentrale Begriffe aus der Wahrscheinlichkeitstheorie und der Theorie der Markov Ketten eingeführt. Darauf aufbauend wird die Ordnung einer Markov Kette sowie Kriterien für die Modell-Auswahl vorgestellt. Um die Theorie zu verdeutlichen, wird in Kapitel 2.7 eine Beispielrechnung für eine Markov Kette durchgeführt. Auf grundlegende statistische Größen und Verfahren (wie z. B. theoretische Verteilungen) wird in diesem Kapitel aus Gründen der Übersichtlichkeit nicht eingegangen. Diese sind im Anhang dargestellt.

2.1 Grundbegriffe: Stochastischer Prozess, Zeitreihe und Markov-Urnenmodell

In der Wahrscheinlichkeitstheorie versucht man mittels stochastischer Modelle gegenwärtige Ereignisse zu beschreiben und Entscheidungen über zukünftige Entwicklungen zu treffen [17]. Diese Ereignisse lassen sich als Durchführung eines Zufallsexperiments in Zusammenhang mit der (Eintritts-)Wahrscheinlichkeit¹ eines Ereignisses verstehen [16]. Man spricht von einem **stochastischen Prozess**. Der stochastische Prozess wird durch eine Funktion $X(t)$ mit $t \in \mathbb{T}$ beschrieben, wobei \mathbb{T} als Parameterraum bezeichnet wird und die Menge der Zeitpunkte t beinhaltet. Der Zustandsraum $\mathbb{S} = \{X(t) \mid t \in \mathbb{T}\}$ definiert die Menge der möglichen Ereignisse.

Als Beispiel eines stochastischen Prozesses kann man sich ein Urnenmodell vorstellen. Das **Bernoulli Urnenmodell** besteht aus einer einzigen Urne, die beispielsweise eine bestimmte Anzahl roter und schwarzer Kugeln erhält. Nun wird eine Kugel aus der Urne gezogen, ihre Farbe notiert und wieder in die Urne zurück gelegt. Die Wahrscheinlichkeit, mit der eine rote bzw. eine schwarze Kugel gezogen, bleibt im zeitlichen Verlauf bei diesem Modell konstant. Werden nun die gezogenen Kugeln vor der nächsten Entnahme nicht mehr in die Urne zurückgelegt, handelt es sich um ein **Urnenmodell ohne Zurücklegen**. Hierbei ändert sich bei jeder Ziehung aus der Urne die Zusammensetzung der Restgesamt-

¹Eine mathematisch formale Definition von Wahrscheinlichkeiten befindet sich im Anhang (A.1)

heit und damit die Wahrscheinlichkeit, dass weitere rote bzw. schwarze Kugeln gezogen werden. Bei diesem Modell tritt eine **Wahrscheinlichkeitsverkettung** ein [17], d. h. mit welcher Wahrscheinlichkeit eine Kugel einer bestimmten Farbe gezogen, hängt davon ab, welche Kugeln schon vorher aus der Urne entfernt worden sind. Das führt zu dem Begriff der **bedingten Wahrscheinlichkeit** $P(A | B)$, die die Wahrscheinlichkeit für das Ereignis A unter der Bedingung angibt, dass das Ereignis B vorher bereits eingetreten ist:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}. \quad (2.1)$$

Eine mögliche und simple Realisation eines **Markov Urnenmodells** besteht aus drei Urnen, die alle rote und schwarze Kugeln enthalten [27]. Aus Urne I wird nun eine Kugel gezogen mit der Vereinbarung, dass, wenn diese Kugel rot ist, die nächste Kugel aus Urne II gezogen wird. Ist die Kugel jedoch schwarz, wird die nächste Kugel aus Urne III entnommen. Alle gezogenen Kugeln werden in ihre ursprüngliche Urne zurück gelegt. Zwei weitere Annahmen werden gemacht: (1) Nur Urne I enthalte die gleiche Anzahl an roten und schwarzen Kugeln. Urne II habe mehr rote Kugeln als schwarze und Urne III mehr schwarze Kugeln als rote. (2) Nach zwei Zügen wird die Ziehung bei Urne I fortgesetzt. Das bedeutet, dass es beim ersten Zug aus Urne I gleich wahrscheinlich ist, ob es sich um eine rote oder schwarze Kugel handelt. Beim nächsten Zug aber wird nach den oben gemachten Annahmen eine Farbe präferiert. Denn wenn die erste Kugel eine rote (schwarze) ist, ist es beim nächsten Zug wahrscheinlicher, dass eine rote (schwarze) Kugel gezogen wird. Auf diese Weise entsteht eine Erhaltungsneigung, die begrenzt ist auf das Persistenzintervall von einem zum nächsten Zug. Die Eintrittswahrscheinlichkeiten bleiben in diesem Markov Urnenmodell zeitlich konstant. Modelle dieser Art können durch **Markov Ketten** beschrieben werden [17].

Da im Mittelpunkt dieser Arbeit die Modellierung von Zeitreihen mit Markov Ketten steht, soll im Folgenden der Zusammenhang zwischen einem stochastischem Prozess inklusive Zustandsraum und einer Zeitreihe erläutert werden. In der Meteorologie werden Klimalelemente, wie z. B. Niederschlag und Temperatur, durch Messung, Schätzung und Phänomenbeobachtung (z. B. Auftreten und Häufigkeit von Gewittern) erfasst und in Zeitreihen zusammengefasst [18]. Eine Zeitreihe ist eine zeitlich geordnete Folge von Beobachtungswerten X_t mit $t \in \mathbb{T}$, die in den meisten Fällen in äquidistanten Zeitschritten gemessen werden.

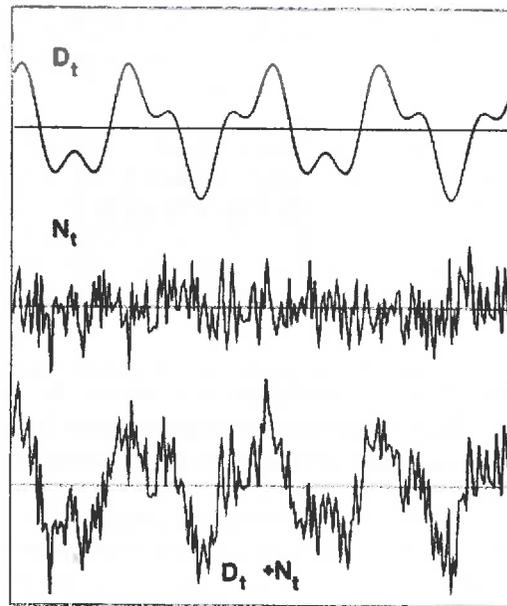


Abbildung 2.1: Realisation einer Zufallsvariable X_t (unten) als Überlagerung der determinierten Komponente D_t (oben) und der stochastischen Komponente N_t (mitte) jeweils zum Zeitpunkt t (nach Storch [22])

In Bezug auf die Zeitreihenanalyse kann eine Zeitreihe allgemein als **Realisation eines Zufallsprozesses** verstanden werden, der durch Überlagerung von verschiedenen determinierten und stochastischen Komponenten entsteht (z. B. [9, 19, 22]), welche für gewöhnlich noch weiter zerlegt werden können. Nach Storch [22] gilt für ein Element X_t der Zeitreihe zum Zeitpunkt t

$$X_t = D_t + N_t, \quad (2.2)$$

wobei D_t die determinierte Komponente und N_t die stochastische Komponente (Rauschen) bezeichnet (vgl. auch Abb. 2.1). Bei einer Zeitreihe kann der Fall auftreten, dass die Daten der Zeitreihe voneinander abhängig sind. Im Falle stochastischer Abhängigkeit (Unabhängigkeit) ist das Ereignis X_t zu einem beliebigen Zeitpunkt t abhängig (unabhängig) von den vorhergehenden Ereignissen. Meteorologische Zeitreihen - hier vor allem Temperatur- und Niederschlagsreihen - sind im Allgemeinen autokorreliert. Das bedeutet, dass zwischen aufeinander folgenden Zeitreihenelementen eine mehr als zufällige Verbindung besteht, solange der Zeitabstand zwischen diesen Elementen nicht zu groß ist [22].

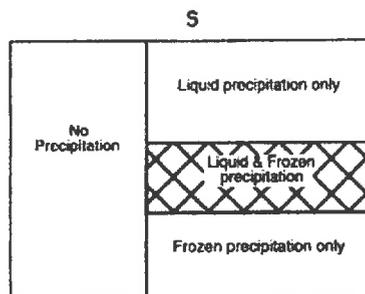


Abbildung 2.2: Venn Diagramm für den Zustandsraum S , der die möglichen auftretenden Niederschlagsereignisse beinhaltet: kein Niederschlag; flüssiger oder gefrorener Niederschlag; flüssiger und gefrorener Niederschlag (vgl. Text). (verändert nach Wilks [24])

Der Zustandsraum hat eine wesentliche Bedeutung: Denn er beinhaltet die möglichen Ereignisse, die eintreten können. In Abbildung 2.2 ist ein so genanntes **Venn Diagramm** dargestellt, das die Beziehung von Zustandsraum und Ereignissen herstellt. Diese Unterscheidung soll vorerst als Beispiel dienen. Später wird ersichtlich werden (Kap. 2.3), dass der Zustandsraum noch stärker eingeschränkt wird. Die einzelnen Rechtecke in Abbildung 2.2 unterscheiden die Art des Niederschlags. Demnach tritt entweder Niederschlag auf oder es tritt kein Niederschlag auf, wobei der gefallene Niederschlag nochmal unterschieden wird in flüssig, flüssig und gefroren oder nur gefroren. In einem gewissen Zeitintervall wird immer eines dieser Ereignisse realisiert. Aus diesem Grund füllen die Rechtecke in Abbildung 2.2 auch den gesamten Zustandsraum aus, da immer ein Ereignis eintreten muss.

Mit Markov Ketten, die in den folgenden Abschnitten mathematisch genauer betrachtet werden, können solche Prozesse, die in regelmäßigen Zeitpunkten beobachtet werden und von Beobachtung zu Beobachtung mit konstanten Wahrscheinlichkeiten zwischen verschiedenen Zuständen (z. B. Niederschlagsereignisse, vgl. Abb. 2.2) wechseln, modelliert werden. In der Realität - bei meteorologischen Systemen - sind die bedingten Wahrscheinlichkeiten bei solchen Prozessen in der Regel von Beobachtung zu Beobachtung zeitlich nicht konstant [11]. Unterstellt man bei der Modellierung jedoch konstante Wahrscheinlichkeiten, so lassen sich trotzdem viele Strukturen in Zeitreihen aufdecken [24].

2.2 Markov Kette

Die **Markov Kette** ist ein stochastisches Modell, dass durch einen **Markov Prozess** definiert wird [5]. Der Markov Prozess ist ein spezieller stochastischer Prozess [3], bei dem nur eine bestimmte Anzahl vergangener Zustände wichtig für die Modellierung des weiteren zeitlichen Verlaufs der Zustände ist. Für eine zeit-diskrete Markov Kette (DTMC, discrete-time Markov Chain) kann für ein diskretes Set an Zeitpunkten, das repräsentiert wird durch den Parameterraum $\mathbb{T} = \{0, 1, 2, \dots, N\}$, jeweils genau eine Beobachtung X_n gemacht werden, die der Zustand zum Zeitpunkt n genannt wird [21]. Jeder Zustand tritt gleich lange auf und der Wechsel zwischen Zuständen geschieht instantan, d. h. mit der Zeit Null. Bei einer zeit-kontinuierlichen Markov Kette (CTMC, continuous-time Markov Chain) treten die Zustände unterschiedlich lange auf und der Wechsel zwischen ihm kann ebenso unterschiedliche Zeit in Anspruch nehmen. In dieser Arbeit werden nur DTMC betrachtet. Mathematisch wird der Markov Prozess wie folgt definiert [3]:

Ein stochastischer Prozess mit der Zufallsvariable X_t mit $t \in \mathbb{T}$, dem Zustandsraum $\mathbb{S} = \{i_1, \dots, i_{t-1}, i, j\}$ heißt Markov Prozess, wenn für die bedingeten Wahrscheinlichkeiten gilt:

$$P(X_{t+1} = j \mid X_t = i, X_{t-1} = i_{t-1}, \dots, X_2 = i_2, X_1 = i_1) = P(X_{t+1} = j \mid X_t = i) \quad (2.3)$$

Aus der Kenntnis von X_{t-1} und weiter zurückliegenden Zuständen werden keine Informationen erlangt, die für die Prognose des Zustands zum Zeitpunkt $t + 1$ relevant sind. Angewendet auf eine Beobachtungsreihe bedeutet das, dass der nächste Zustand ($t+1$) bei Kenntnis des jetzigen Zuständen (t) statistisch prognostizierbar ist [24]. Diese bedingeten Wahrscheinlichkeiten

$$P(X_t = j \mid X_s = i) = p_{ij}(s, t) \text{ mit } j - s = n \quad (2.4)$$

werden auch Übergangswahrscheinlichkeiten n -ter Ordnung genannt. In diesem Zusammenhang spricht man von der **Markov Eigenschaft** oder auch von einem **Prozess mit begrenztem Gedächtnis**. Die Markov Eigenschaft lässt sich auch auf weiter zurückliegende Zustände erweitern [11]. Eine Markov Kette hat die **Ordnung r** , wenn für $t \geq r + 1$ mit $r \in \mathbb{N}$ folgendes gilt:

$$P(X_{t+1} \mid X_t, X_{t-1}, X_{t-2}, \dots, X_1) = P(X_{t+1} \mid X_t, X_{t-1}, X_{t-2}, \dots, X_{t-r}) \quad (2.5)$$

Wenn die Übergangswahrscheinlichkeiten zeitlich konstant sind, spricht man von **homo-**

genen Markov Ketten. Die Übergangswahrscheinlichkeiten erhält man durch Schätzen der relativen Häufigkeiten aus der Stichprobe, welche eine gute Approximation für die Maximum-Likelihood Schätzer der Wahrscheinlichkeiten darstellen [1]. Die relativen Häufigkeiten gewinnt man durch Auszählen der Zustandskombinationen, deren Anzahl abhängig von der Ordnung der Markov Kette ist. Beispielsweise hat man bei einer Markov Kette 1. Ordnung mit den zwei Zuständen „0“ und „1“ vier mögliche Zustandskombinationen: nämlich 00, 01, 10 und 11 (vgl. Abb. 2.4 auf Seite 13). Die dazugehörigen Übergangswahrscheinlichkeiten berechnen sich mit folgenden Gleichungen

$$p_{00} = P\{X_{t+1} = 0 \mid X_t = 0\} = \frac{n_{00}}{n_{00} + n_{01}} = 1 - p_{01} \quad (2.6)$$

$$p_{01} = P\{X_{t+1} = 0 \mid X_t = 1\} = \frac{n_{01}}{n_{00} + n_{01}} = 1 - p_{10} \quad (2.7)$$

$$p_{10} = P\{X_{t+1} = 1 \mid X_t = 0\} = \frac{n_{10}}{n_{10} + n_{11}} = 1 - p_{11} \quad (2.8)$$

$$p_{11} = P\{X_{t+1} = 1 \mid X_t = 1\} = \frac{n_{11}}{n_{10} + n_{11}} = 1 - p_{10} , \quad (2.9)$$

wobei n_{ij} die Anzahl der ij -ten Zustandskombination aus der Beobachtungsreihe angibt. Für Übergangswahrscheinlichkeiten höherer Ordnung gelten analoge Gleichungen. Bezeichnet N die Anzahl der Zustände, so lassen sich die Übergangswahrscheinlichkeiten einer Markov Kette 1. Ordnung in der sogenannten Übergangsmatrix \mathcal{P} mit N Spalten und N Zeilen anordnen.

$$\mathcal{P} = \begin{pmatrix} p_{11} & \cdots & p_{1N} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NN} \end{pmatrix} \quad (2.10)$$

Aus den Übergangswahrscheinlichkeiten von Markov Ketten 1. Ordnung, die nur zwei Zustände beinhalten, lassen sich weitere Größen wie die **stationäre Wahrscheinlichkeit** π_1 und den **Persistenz Parameter** r_1 berechnen. Die stationäre Wahrscheinlichkeit wird nach

$$\pi_1 = \frac{p_{01}}{1 + p_{01} - p_{11}} \quad (2.11)$$

berechnet und beschreibt die generelle klimatologische Eintrittswahrscheinlichkeit von Ereignis „1“. Die Berechnung sollte daher über einen großen Zeitraum geschehen [24]. Die entsprechende Wahrscheinlichkeit π_0 ergibt sich aus $\pi_0 = 1 - \pi_1$. Ebenso impliziert die Übergangswahrscheinlichkeit p_{11} einen gewissen Grad an Autokorrelation [24], der durch den Persistenz Parameter r_1 beschrieben wird:

$$r_1 = p_{11} - p_{01} . \quad (2.12)$$

Je größer r_1 ist, desto weiter liegen p_{11} und p_{01} auseinander. Das wiederum ist ein Maß dafür, wie wahrscheinlicher es ist, dass eine „1“ nach einer „1“ folgt anstatt nach einer „0“. Deswegen ist Persistenz Parameter vergleichbar mit dem lag-1 Autokorrelationskoeffizient. Für die Zeilensummen der Übergangsmatrix gilt aufgrund der Gleichungen (2.6) bis (2.9) $\sum_{j=1}^N p_{ij} = 1$, da ein Zustand immer eintreten muss. Mit Gleichung (2.4) und dem Gesetz der totalen Wahrscheinlichkeit² lässt sich die Wahrscheinlichkeit für homogene Markov Ketten ausrechnen, mit der sich das System nach k Schritten im Zustand i befindet, wobei zu beachten ist, dass die hochgestellten Klammerausdrücke in Gleichung (2.13) keine Exponenten sind und deswegen in Klammern gesetzt sind.

$$p_i^{(k+1)} = P(X_{k+1} = i) = \sum_j P(X_{(k+1)} = i \mid X_k = j) \cdot P(X_k = j) = \sum_j p_{ij} \cdot p_i^{(k)} \quad (2.13)$$

Diese Gleichung lässt sich auch in Matrixschreibweise umschreiben, so dass man abschließend Folgendes erhält:

$$(p_1^{(k)}, p_2^{(k)}, \dots, p_N^{(k)}) = (p_1, p_2, \dots, p_N) \mathcal{P}^k \quad (2.14)$$

$$\vec{\gamma}_k = \vec{\gamma}_0 \mathcal{P}^k \quad (2.15)$$

Will man also ausrechnen, welche Wahrscheinlichkeitsverteilung $\vec{\gamma}_k$ nach k Schritten vorliegt, wird die Übergangsmatrix k -mal mit sich selbst und anschließend mit dem Zeilenvektor $\vec{\gamma}_0 = (p_1, p_2, \dots, p_N)$ multipliziert, der die Startverteilung der Wahrscheinlichkeiten p_i angibt, mit denen sich der Prozess zum Startzeitpunkt im Zustand i befindet. Durch den Zustandsraum \mathbb{S} , der Übergangsmatrix \mathcal{P} und dem Zeilenvektor γ_0 , der die Anfangsverteilung der Wahrscheinlichkeiten festlegt, ist die Markov Kette vollständig definiert [3].

2.3 Transformation in eine Ordinalskala

Im vorigen Kapitel ist erwähnt worden, dass die Zustände zur Unterscheidung in der Regel mit Nummern versehen werden. Nun muss geklärt werden, wie die Daten einer Zeitreihe behandelt werden, um eine Einteilung und damit eine Unterscheidung der Zustände zu erzielen. Für die Auswertung mit Markov Ketten ist nicht der absolute Wert entscheidend, sondern vielmehr eine Klassenbildung, durch welche die Originaldaten aufgeteilt werden. Die ursprünglich kontinuierliche Skala wird für die Analyse in eine diskrete Skala, die Ordinalskala, durch Klassenbildung transformiert. Die Klasse fungiert nach der Transformation als Zufallsvariable X_t . Man erhält eine transformierte Zeitreihe, in der nicht

² $P(A) = \sum_t P(A \mid B_t) \cdot P(B_t)$

mehr die Niederschlagshöhe steht, sondern die Klasse, in der die entsprechende Niederschlagshöhe eingeordnet wurde.

Wie viele Klassen berücksichtigt werden, ist durch die Fragestellung der Analyse bestimmt, bleibt aber letztenendes willkürlich. Bei der Niederschlagsanalyse wird zunächst zwischen zwei Zuständen unterschieden, da dies die konventionelle Vorgehensweise³ ist (z. B. [7]). Daran anschließend werden vier Zustände eingeführt, um eine größere Differenzierung der Zustände zu erhalten und damit eine Information über die Intensität des Niederschlags mit einzubauen, welches mit nur zwei Zuständen nicht möglich ist. Mehr als vier Zustände einzuführen, hat sich als nicht sinnvoll erwiesen, da zu wenige Daten für die Schätzung der Übergangswahrscheinlichkeiten zur Verfügung stehen⁴. Tabelle 1 gibt Definition der Zustände wieder.

Tabelle 1: Bedeutung der Zustände für Niederschlagsdaten. Dabei sind die Zustände 2 und 3 für dichotome Analyse des Niederschlags nicht definiert (-/-).

Klimaelement	Zustand X_t (ordinales Merkmal)			
	0	1	2	3
Niederschlag (dichotom)	Trockentag	Regentag	-/-	-/-
Niederschlag	Trockentag	wenig feucht	feucht	sehr feucht

Dabei versteht man unter einem *Trockentag* einen Tag mit Niederschlagsmenge $0mm$ sowie auch einen Tag mit keinem messbaren Niederschlag. Dementsprechend ist ein *Regentag* das komplementäre Niederschlagsereignis zu einem Trockentag (Niederschlagsmenge $> 0mm$). Für die Analyse mit mehreren Zuständen wird an den jeweiligen Datensatz eine theoretische Verteilung angepasst, um über die Quantile die Grenzen der Zustände bestimmen zu können. Genaueres zur Anpassung und zu den verwendeten Quantilen folgt im Kapitel 4 „Methodik“.

³Diese Vorgehensweise wird auch dichotome Analyse genannt, da hier nur zwischen ja/nein bzw. 0/1 Ereignissen unterschieden wird. (0 bedeutet „Niederschlag ist nicht aufgetreten“ und 1 „Niederschlag ist aufgetreten“)

⁴Einer statistischen Faustregel zur Folge sollten 30 Werte pro Schätzer zur Verfügung stehen. Bei n Zuständen und bei k . Ordnung hat man n^{k+1} Parameter (Übergangswahrscheinlichkeiten).

2.4 Klassifizierung der Zustände

Aus der Theorie der Markov Ketten sind verschiedene Typen an Zuständen bekannt (z. B. [3, 11]), die im Folgenden anhand von Übergangswahrscheinlichkeiten einer Markov Kette 1. Ordnung festgelegt werden, analog aber auch für höhere Ordnungen gelten.

absorbierende Zustände

Ein Zustand i wird absorbierend genannt, wenn er nicht mehr verlassen werden kann und damit den Endzustand des Systems bildet. Das bedeutet, dass sich in allen weiteren Zeitschritten dieser Zustand wiederholt. Für die Übergangswahrscheinlichkeit gilt:

$$p_{ii} = 1$$

Die nicht-absorbierenden Zustände werden als **innere** Zustände bezeichnet.

transiente Zustände

Ein Zustand j ist transient, wenn er von i aus erreicht werden kann, aber die Rückkehr nach i nicht möglich ist. Für die entsprechenden Übergangswahrscheinlichkeiten gilt:

$$p_{ij} > 0 \text{ und } p_{ji} = 0$$

rekurrente Zustände

Ein Zustand j ist rekurrent, wenn er von i aus potenziell erreichbar ist, und auch wieder zu i zurückkehren kann. Für die entsprechenden Übergangswahrscheinlichkeiten gilt:

$$p_{ij} > 0 \text{ und } p_{ji} > 0$$

periodische und aperiodische Zustände

Einen Zustand i nennt man periodisch, wenn eine Rückkehr nach i nicht nach beliebig vielen Schritten möglich ist. Ist eine Rückkehr prinzipiell nach einem Zeitschritt möglich, handelt es sich um einen aperiodischen Zustand mit der Periode 1.

2.5 Ergodische Markov Ketten

Nachdem die Zustände genauer betrachtet wurden, lässt sich eine interessante Eigenschaft einer Markov Kette herausarbeiten: die **Aperiodizität** und **Irreduzibilität**. Diese Art von Markov Ketten werden in der Literatur auch **ergodische** Markov Ketten genannt (z. B. [3, 11]). Die Aperiodizität bedeutet, dass jeder beliebiger Zustand von jedem beliebigen anderen Zustand aus in einem Zeitschritt erreichbar ist. Bei Irreduzibilität sind alle Übergangswahrscheinlichkeiten der Übergangsmatrix größer Null. Daraus folgt, dass die Markov Kette im zeitlichen Verlauf niemals auf ein Subsystem (engl. closed subset) an Zuständen reduziert wird, das dann nicht mehr verlassen werden kann. Alle Zustände können wieder im Verlauf des Prozesses erreicht werden. Bei reduzierbaren Markov Ketten ist die Wahrscheinlichkeit zum Verlassen des Subsystems gleich Null.

Bei ergodischen Markov Ketten konvergiert die Übergangsmatrix \mathcal{P} komponentenweise gegen die Grenzmatrix \mathcal{W} , in der alle Zeilen gleich sind (Theorem 7.4, in Behrends [3]). Das bedeutet, dass für die Übergangswahrscheinlichkeiten von \mathcal{P} gilt

$$p_{ji}^{(k)} \rightarrow \pi_i \text{ für } k \rightarrow \infty .$$

Für den Zeilenvektor π von \mathcal{W} gilt

$$\pi^T \mathcal{P} = \pi^T \text{ mit } \pi = (\pi_1, \pi_2, \dots, \pi_N), \pi_i > 0 \text{ und } \sum_i \pi_i = 1 . \quad (2.16)$$

Der Fehler zwischen \mathcal{P}^k und \mathcal{W} lässt sich durch den **maximalen relativen Fehler** $\delta(k)$ angeben:

$$\delta(k) := \frac{\max_{ij} |p_{ij}^{(k)} - \pi_j|}{\pi_j} . \quad (2.17)$$

Aus der Konvergenz der Übergangsmatrix lässt sich ableiten, dass man für ein nicht zu kleines k aus der Kenntnis von $p_{ji}^{(k)}$ nicht mehr auf k zurückschließen kann. Die Kette vergisst die Länge ihrer Geschichte. Da der Grenzwert nicht von der Startposition abhängt, vergisst die Kette ebenso ihre Startposition. Wenn man als Anfangsverteilung direkt π^T anwendet, geschieht dies sofort. Daher wird π^T auch **stationäre Verteilung** (engl. equilibrium distribution) genannt [3].

In der Regel sind Markov Ketten 1. Ordnung in meteorologischen Anwendungen irreduzibel und aperiodisch [11]. Wäre dies nicht der Fall, bedeutete das, dass im Verlauf des Markov Prozesses z. B. nur noch Regentage auftreten und der anderer Zustand, ein Trock-

entag, nicht mehr erreichbar ist. Das würde eine sehr schlechte Modellierung bedeuten. Markov Ketten werden mit sogenannten **Prozessdiagrammen** dargestellt. Abbildung 2.3 zeigt das Prozessdiagramm einer reduziblen Markov Kette. In Abbildung 2.4 ist ein Diagramm einer ergodischen Markov Kette dargestellt. Die Knoten repräsentieren die zwei verschiedenen Zustände und die Pfeile die Übergangswahrscheinlichkeiten.



Abbildung 2.3: Prozessdiagramm einer reduziblen Markov Kette. Der grüne Zustand ist transient. Ist ein lila Zustand erreicht, ist es nunmehr unmöglich in einen blauen Zustand zu wechseln. Die Markov Kette reduziert sich in diesem Fall auf das lila Subsystem (nach Trömel [23]).

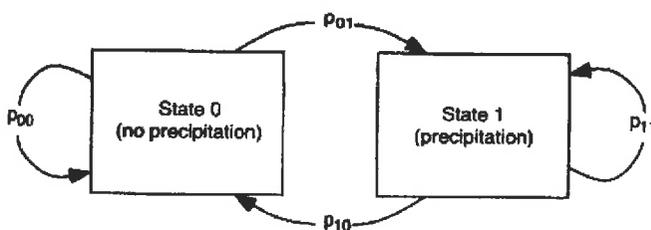


Abbildung 2.4: Prozessdiagramm für eine ergodische Markov Kette 1. Ordnung mit den Zuständen „0“ für einen Trockentag und „1“ für einen Regentag. Die Knoten repräsentieren die Zustände und die Pfeile die Übergangswahrscheinlichkeiten P_{00}, P_{01}, P_{10} und P_{11} , die alle größer Null sind. (nach Wilks [24])

2.6 Selektionskriterien für die Modell-Auswahl

Welches Modell ausgewählt wird, entscheidet sich durch die Ordnung r der Markov Kette, die den Datensatz am besten charakterisiert. Nun muss a priori festgelegt werden, wieviele Ordnungen überhaupt zugelassen werden. Die Anzahl der Ordnungen muss genügend groß gewählt werden, damit es möglich ist, eine kleinere Ordnung auszuwählen, als die maximal zugelassene Ordnung vorgibt [6]. Wieviele Zustände in Betracht kommen, ist ebenso entscheidend für die Modell-Auswahl (vgl. Kap. 2.2). Eine größere oder kleinere Anzahl an Zuständen kann sicher auch die Ordnung der Markov Kette verändern. Es existieren

verschiedene Methoden, wie die geeignete Ordnung der Markov Kette bestimmt wird. Gabriel und Neumann [7] haben dazu den χ^2 -Test verwendet, der aber nur hilfreich ist, wenn nur zwei Alternativen zur Auswahl stehen [25]; also beispielsweise wird die 0. gegen die 1. Ordnung getestet. Zwei weitere Methoden, die mehrere Ordnungen vergleichen können, sind das **Akaike Information Criterion (AIC)** und das **Bayesian Information Criterion (BIC)**, die die Log-Likelihood Funktionen L_m der Übergangswahrscheinlichkeiten der angepassten Markov Kette verwenden [24].

$$AIC(m) = -2L_m + 2s^m(s-1) \quad (2.18)$$

$$BIC(m) = -2L_m + s^m(\ln n) \quad (2.19)$$

Dabei bezeichnet $s \in \mathbb{N}$ die Anzahl der Zustände und $m \in \mathbb{N}$ die jeweilige zu testende Ordnung. Es wird die Ordnung ausgewählt, bei der das AIC und BIC minimal wird. Beide Kriterien basieren auf einem Ausgleich zwischen der Güte der Anpassung und einer Bestrafung, die mit größerer Anzahl der angepassten Parameter steigt [24]. Das BIC tendiert zur Auswahl von niedrigeren Ordnungen als das AIC und wird von verschiedenen Autoren (z. B. [11, 23]) als präferiertes Kriterium erachtet, wenn es sich vor allem um große Datenmengen handelt [11]. Das BIC wird auch in dieser Arbeit als bevorzugtes Kriterium eingesetzt. Die Log-Likelihood Funktionen L_m für die Ordnungen 0 bis 3 lauten:

$$L_0 = \sum_{j=0}^{s-1} n_j \ln p_j \quad (2.20)$$

$$L_1 = \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} n_{ij} \ln p_{ij} \quad (2.21)$$

$$L_2 = \sum_{h=0}^{s-1} \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} n_{hij} \ln p_{hij} \quad (2.22)$$

$$L_3 = \sum_{g=0}^{s-1} \sum_{h=0}^{s-1} \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} n_{ghij} \ln p_{ghij} \quad (2.23)$$

2.7 Beispielrechnung für eine Markov Kette 1. Ordnung mit zwei Zuständen

Um einen Eindruck davon zu gewinnen, wie eine Markov Kette praktisch berechnet wird, wird eine Beispielrechnung der Einfachheit halber für eine (ergodische) Markov Kette 1. Ordnung mit zwei Zuständen durchgeführt. Dazu wird der Datensatz der Niederschlagsstation Frankfurt am Main aus dem Juni 2000 verwendet, der in Tabelle 2 zusammen mit

der durch Transformation entstandenen Ordinalskala aufgelistet ist (vgl. Kapitel 2.3). Für jeden Zeitpunkt t ist der Markov Prozess entweder im Zustand „0“ für einen Trockentag oder „1“ für einen Regentag. Im nächsten Zeitschritt $t + 1$ kann der Markov Prozess im selben Zustand verharren oder den Zustand wechseln.

Tabelle 2: Tabelliert ist der tägliche Niederschlagsverlauf in [mm] in Frankfurt am Main mit dem dazugehörigen Zustand i_t für den jeweiligen Tag t im Juni 2000.

Tag, t	1	2	3	4	5	6	7	8	9	10
Niederschlag [mm], X_t	0	0	8.4	7.0	3.1	0.1	0	0	0	0
Zustand, i_t	0	0	1	1	1	1	0	0	0	0
Tag, t	11	12	13	14	15	16	17	18	19	20
Niederschlag [mm], X_t	1.3	0	0	0	0	0	0	0	0	0
Zustand, i_t	1	0	0	0	0	0	0	0	0	0
Tag, t	21	22	23	24	25	26	27	28	29	30
Niederschlag [mm], X_t	0.3	0.2	5.1	1.9	0.2	0	0	0	0	0
Zustand, i_t	1	1	1	1	1	0	0	0	0	0

Durch Auszählen der möglichen Zustandskombinationen aus Tabelle 3 erhält man $n_0 = 20$ und $n_1 = 10$ und damit $p_0 = 0.667$ und $p_1 = 0.333$. Ebenso ist $n_{00} = 16$, $n_{01} = 3$, $n_{10} = 3$ und $n_{11} = 7$. Damit kann man mithilfe der Gleichungen (2.6) bis (2.9) die Übergangswahrscheinlichkeiten ausrechnen und es ergibt sich für die Übergangsmatrix:

$$P = \begin{pmatrix} 0.84 & 0.16 \\ 0.3 & 0.7 \end{pmatrix}$$

Die Übergangswahrscheinlichkeiten der zweiten und dritten Ordnung sind in Tabelle 4 einzusehen. Die Ergebnisse für die die Berechnung der Modell-Selektionskriterien (Gleichungen 2.18 und 2.19) sind in Tabelle 3 wiedergegeben. Beide Kriterien erachten in diesem Rechenbeispiel die erste Ordnung als geeignete Wahl, da die entsprechenden Werte für das AIC und BIC für die erste Ordnung minimal sind.

Tabelle 3: Ergebnisse der Beispielrechnung für die Log-Likelihood Funktionen L_m , das Akaike Kriterium (AIC(m)) und das Bayes'sche Kriterium (BIC(m)) für die Ordnungen $m = \{0, 1, 2, 3\}$.

Ordnung m	L_m	AIC(m)	BIC(m)
0	-19.39	40.78	42.18
1	-14.39	32.78	35.98
2	-13.63	35.26	40.86
3	-9.37	34.74	45.94

Die Anfangsverteilung ist offenbar

$$\vec{\gamma}_0 = (p_0^{(0)}, p_1^{(0)}) = (1, 0) ,$$

da am ersten Tag der Beispielfolge (vgl. Tab. 3) der Zustand 0 auftritt und $p_0 = 1$ gesetzt wird. Damit kann man nach Gleichung (2.15) die Wahrscheinlichkeitsverteilung $\vec{\gamma}_k$ mit den Wahrscheinlichkeiten $p_0^{(k)}$ und $p_1^{(k)}$ ausrechnen, mit denen nach k Schritten die Zustände 0 und 1 eintreten. Dieses ist bis zu $k = 17$ durchgeführt worden. Man erkennt deutlich den Konvergenzcharakter (vgl. Kapitel 2.5).

$$\begin{aligned} \vec{\gamma}_0 &= (1, 0) \\ \vec{\gamma}_1 &= (0.84, 0.16) \\ \vec{\gamma}_2 &= (0.7536, 0.2464) \\ \vec{\gamma}_5 &= (0.6681, 0.3319) \\ \vec{\gamma}_{10} &= (0.6535, 0.3465) \\ \vec{\gamma}_{15} &= (0.6521, 0.3479) \\ \vec{\gamma}_{16} &= (0.6522, 0.3478) \\ \vec{\gamma}_{17} &= (0.6522, 0.3478) \end{aligned}$$

Der maximale relative Fehler (Gleichung 2.17) wird erst für $k = 17$ vernachlässigbar klein.

$$\mathcal{P}^{17} = \begin{pmatrix} 0.6522 & 0.3478 \\ 0.6522 & 0.3478 \end{pmatrix} \text{ mit } \delta(17) = 5.75\text{E} - 5$$

Tabelle 4: Auflistung der berechneten Übergangswahrscheinlichkeiten anhand der aus den Niederschlags-Tagesdaten (Juni 2000 in Frankfurt/Main, vgl. Tabelle 3) empirisch gefundenen Häufigkeiten der Zustandskombinationen.

Ordnung	Zustandskombination	empirische Häufigkeit	Übergangswahrscheinlichkeit
0	1	10	0.33
	0	20	0.67
1	11	7	0.7
	10	3	0.3
	01	3	0.16
	00	16	0.84
2	111	5	0.71
	110	2	0.29
	011	2	0.67
	010	1	0.33
	101	0	0
	100	3	1
	001	3	0.2
	000	12	0.8
3	1111	3	0.6
	1110	2	0.4
	0111	2	1
	0110	0	0
	1011	0	0
	1010	0	0
	0011	2	0.67
	0010	1	0.33
	1101	0	0
	1100	2	1
	0101	0	0
	0100	1	1
	1001	0	0
	1000	3	1
	0001	2	0.18
	0000	9	0.82

3 Datengrundlage

In dieser Arbeit wurden Niederschlagsdaten in täglicher Auflösung von 301 Stationen verwendet, die aus einem Pool von 1700 Stationen aus dem Messnetz des Deutschen Wetterdienstes ausgewählt wurden, so dass sie in etwa in äquidistanten Abstand zueinander liegen. In den neuen Bundesländern ist die Lage der Stationen weniger dicht, gilt aber noch als ausreichend gut [9]. Alle diese Stationen liefern 50-jährige Datenreihen für den Zeitraum 1951-2000. Die Abbildung 3.1 zeigt die geografischen Stationsverteilungen.

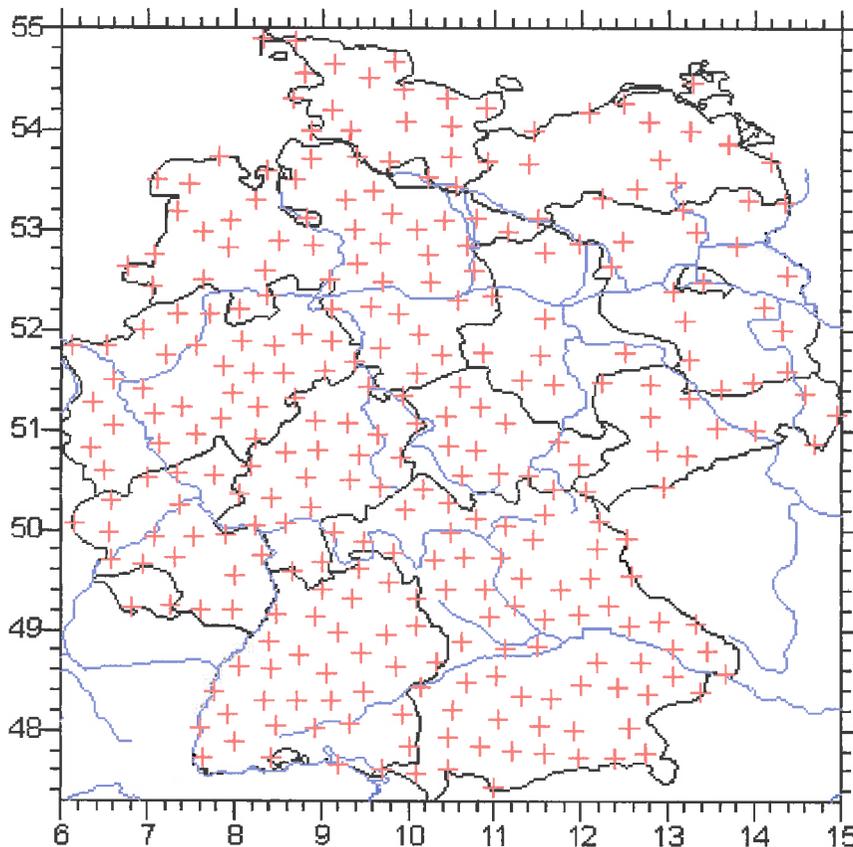


Abbildung 3.1: geografische Verteilung der Niederschlagsstationen

In Abbildung 3.2 ist die die mittlere Niederschlagssumme für das ganze Jahr und für jede Jahreszeit dargestellt. Diese wurde durch Aufsummieren der Tagesdaten gebildet und über die 50 Jahre gemittelt.

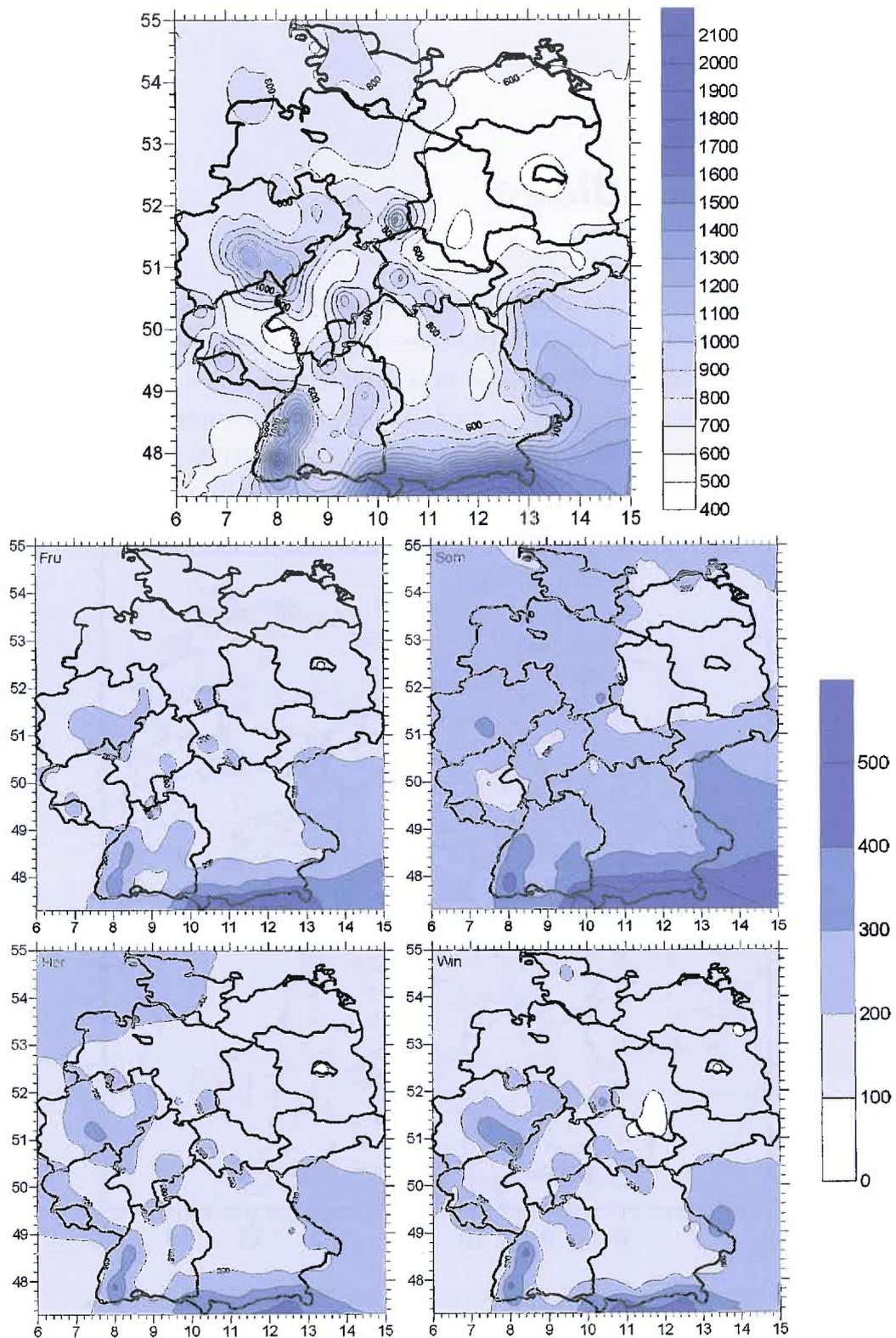


Abbildung 3.2: mittlerer Jahresniederschlag (1951-2000) in [mm] generiert von den verwendeten 301 Niederschlagsstationen für das ganze Jahr (oberes Bild) und die vier Jahreszeiten

Später in dieser Arbeit wird anhand der Ergebnisse aus der Schätzung der Wahrscheinlichkeiten für Bremen, Sankt Peter-Ording, Frankfurt/Main, Potsdam, Feldberg (Schwarzwald) sowie für die Zugspitze der Jahresgang der Übergangswahrscheinlichkeiten erläutert. Die Lage dieser und weiterer Stationen, die an gesonderter Stelle in Erscheinung treten, ist in Abbildung 3.3 dargestellt.

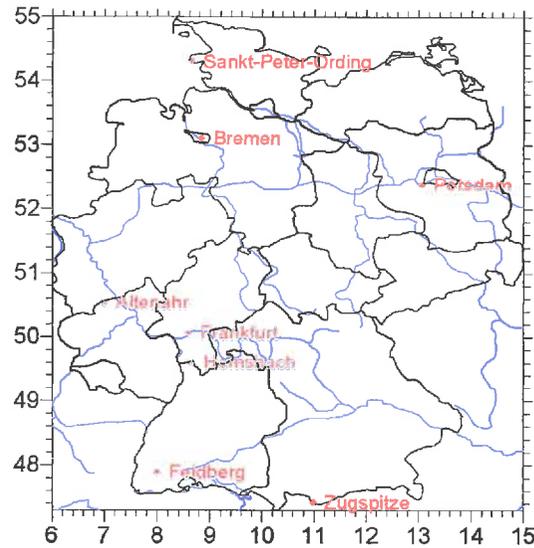


Abbildung 3.3: Die Karte zeigt Stationen, die später an gesonderter Stelle Erwähnung finden.

4 Methodik

Bevor die Analyse mit Markov Ketten durchgeführt werden kann, werden die einzelnen Zeitreihen der Niederschlagsstationen nach Entfernung des Schaltjahres in eine Ordinalskala umgewandelt. Die Niederschlagsdaten werden mit zwei und mit vier Zuständen untersucht. Bei der dichotomen Analyse der Niederschlagsdaten können die beiden Zustände direkt aus dem Datensatz bestimmt werden. Werden mehr als zwei Zustände betrachtet, wird die Weibullverteilung (WV) verwendet (vgl. Anhang A.2.1), um über die in Tabelle 5 einzusehenden Quantile die Zustände zu unterscheiden.

Tabelle 5: Verwendete Quantile der Weibullverteilung (WV) zur Unterscheidung der Zustände für den jeweiligen Datensatz.

Datensatz (Tagesdaten)	Verteilung	Quantil 1	Quantil 2
Niederschlag	WV	0.3	0.7

Die Anpassung der WV wird dabei monatlich nach der Methode der Momente (MOM) und der Maximum Likelihood Schätzung (MLE)(vgl. Anhang A.2.2) vorgenommen, wobei aber alle Daten gleicher Monate aus dem Gesamtzeitraum 1951-2000 verwendet werden. Somit kann auf die Entfernung des Jahresgangs verzichtet werden. Bei Niederschlagsdaten existiert im Allgemeinen ein Jahresgang in der Varianz, der aber unberücksichtigt bleibt.

Die Verteilungsparameter aus der MOM dienen der MLE als Startwerte. Die Durchführung MLE hat das Ziel, die Verteilungsparameter iterativ weiter zu verbessern, um so die best mögliche Anpassung zu erhalten. Des Weiteren wird die Anpassung der WV nur an die Regentage vorgenommen. Denn typischerweise kommen in Niederschlags-Tagesdaten Trockentage am häufigsten vor gefolgt von Tagen mit wenig Niederschlag. Deswegen ist es sinnvoll, Trockentage direkt als eine eigene Klasse zu betrachten.

Diese Vorgehensweise kann man kritisch betrachten, da davon auszugehen ist, dass einige Zeitreihen instationär sind und daher eine Anpassung unmöglich bzw. nicht signifikant sein sollte. Aber um Änderungen im Beobachtungszeitraum zu detektieren, ist es notwendig, eine gewisse Instationarität zuzulassen. Damit ist es möglich, über die

Übergangswahrscheinlichkeiten Änderungen der zeitlichen Verteilung der Zustände zu erfassen. Die Güte der Anpassung wird mit dem Kolmogorov-Smirnov Test festgestellt (vgl. Anhang A.3.2).

Nach Auszählen der Zustandskombinationen werden die Übergangswahrscheinlichkeiten für den Gesamtzeitraum (1951-2000), für die beiden Hälften des Gesamtzeitraums (1951-1975 und 1976-2000) und darüber hinaus für jedes Jahr getrennt nach den meteorologischen Jahreszeiten berechnet. Daran anschließend werden die Log-Likelihood Funktionen mithilfe den Gleichungen (2.20) bis (2.23) berechnet. Die maximal mögliche Ordnung der Markov Ketten wird auf $m = 3$ begrenzt und mit dem AIC und dem BIC (vgl. Kap. 2.6) für jede Station berechnet. Die Ordnung wird für den Gesamtzeitraum sowie für dessen Hälften bestimmt. So ist es möglich die Ordnungen aus der ersten Hälfte mit denen aus der zweiten Hälfte zu vergleichen.

Darüber hinaus wird die zeitliche Entwicklung der Übergangswahrscheinlichkeiten auf Trends untersucht. Dazu wird die lineare Regressionsgerade berechnet und deren Steigung mit dem t-Test (vgl. Anhang A.3.1) geprüft, ob die Steigung signifikant von Null verschieden ist.

Bei der Auswertung der Niederschlagsdaten wird zusammenfassend folgendermaßen vorgegangen:

1. Entfernung des Schaltjahres; der Jahresgang wird nicht eliminiert
2. Anpassung der Weibullverteilung an die Regentage jeweils monatlich für den Gesamtzeitraum 1951-2000
3. Transformation in die Ordinalskala für zwei und für vier Zustände über die Quantile der Weibullverteilung (vgl. Tab. 5); Trockentage werden direkt als eigenständige Klasse betrachtet
4. Auszählen der Zustandskombinationen, Berechnung der Übergangswahrscheinlichkeiten und der Loglikelihood Funktionen nach Jahreszeiten getrennt
5. Bestimmung der Ordnung der Markov Kette durch das AIC und BIC (vgl. Kap. 2.6) für den Gesamtzeitraum und für die beiden Hälften des Gesamtzeitraums; BIC ist bevorzugtes Kriterium
6. Berechnung der linearen Regressionsgeraden der Übergangswahrscheinlichkeiten; Steigung wird mit t-Test beurteilt.

5 Ergebnisse und Interpretation

In diesem Kapitel werden die Ergebnisse von der Anwendung der Markov Ketten auf die täglichen Niederschlagswerte mit zwei und mit vier Zuständen in jeweils einem Unterkapitel präsentiert. Dabei wird zuerst auf Voruntersuchungen eingegangen und danach auf die festgestellten Ordnungen im Gesamtzeitraum sowie in den beiden Hälften des Gesamtzeitraums. Anschließend werden anhand von Karten einige Übergangswahrscheinlichkeiten vorgestellt und ihre zeitliche Entwicklung mittels linearer Trends untersucht.

5.1 Ergebnisse der Analyse von den Niederschlagsdaten mit zwei Zuständen

5.1.1 Zur stationären Wahrscheinlichkeit, dem Persistenzparameter und den Trends der Regentage

Die stationäre Wahrscheinlichkeit π_1 ist mithilfe von Gleichung (2.11) für den Gesamtzeitraum 1951-2000 berechnet worden und ist in Abbildung 5.1 für die einzelnen Jahreszeiten für die Regentage kartografisch dargestellt. In allen Jahreszeiten treten in den Gebieten nahe der Nordsee und zum Teil auch an der Ostsee sowie in den Mittelgebirgen und Alpenregionen Regentage häufiger auf als Trockentage. Im Frühling findet man im Nordwesten und Süden Deutschlands Wahrscheinlichkeiten über 50%. Die höchsten Werte treten im Westerwald, Schwarzwald und in den Voralpen auf. Minimale Werte, die teilweise unter 40% liegen, zeigen sich in Brandenburg sowie in Teilen Mecklenburg-Vorpommerns und Bayerns. Die geografische Verteilung im Sommer ist der im Frühling sehr ähnlich, wobei im Frühling die Flächen mit Wahrscheinlichkeiten über 50% etwas größer sind. Im Herbst tritt der Nordwesten deutlich hervor mit Werten um 62%. Im Winter ist in praktisch ganz Deutschland das Auftreten von Regentagen wahrscheinlicher als das der Trockentage. Hier befinden sich ebenfalls die Gebiete mit den höchsten Wahrscheinlichkeiten (um 66%) im Nordwesten. Die Wahrscheinlichkeiten im Alpengebiet sind im Herbst und Winter nicht mehr ganz so ausgeprägt wie in den beiden anderen Jahreszeiten.

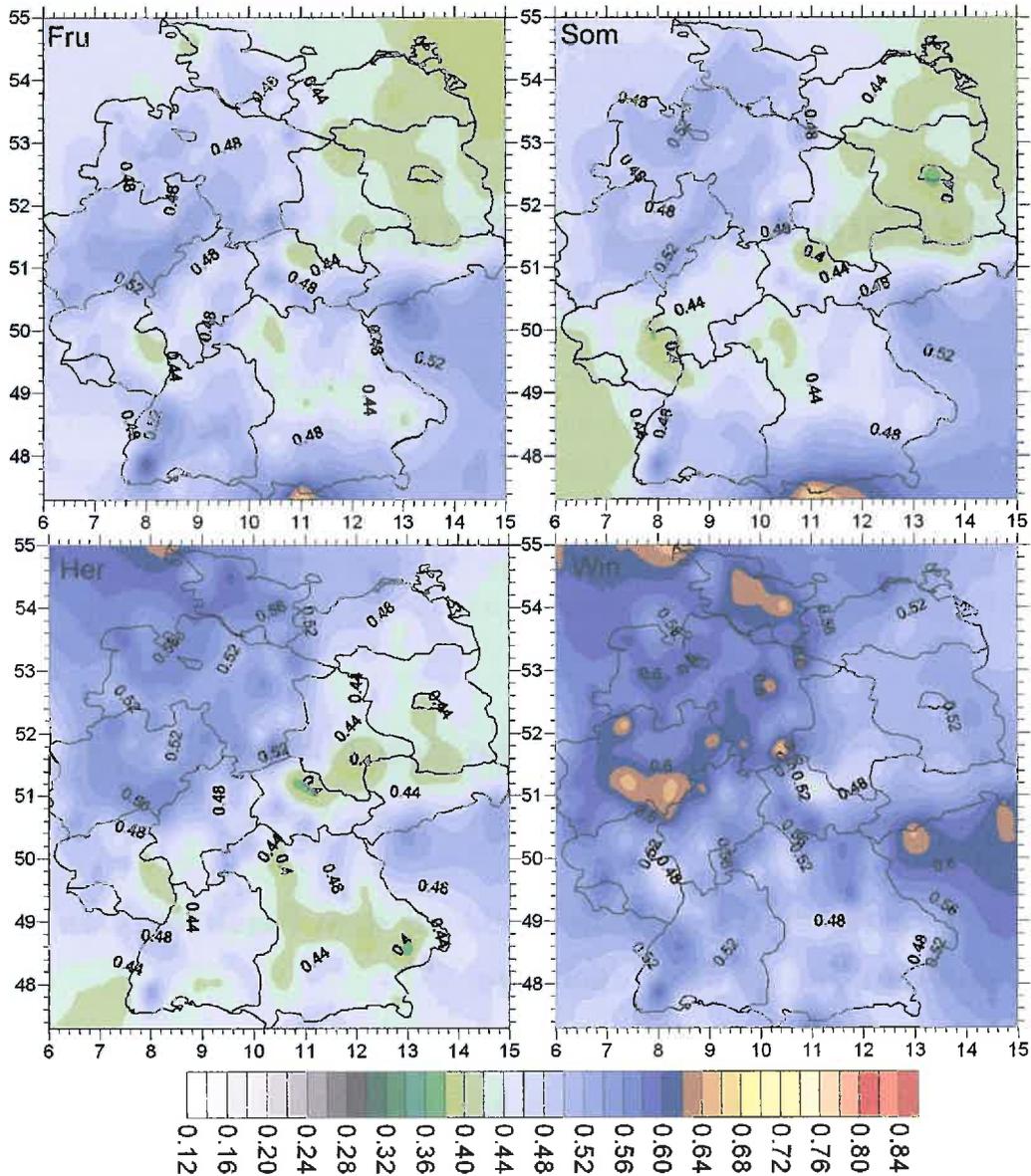


Abbildung 5.1: stationäre Wahrscheinlichkeit π_1 für das Eintreten von Regentagen berechnet für den Zeitraum 1951-2000 jeweils für den Frühling, den Sommer, den Herbst und den Winter

In Abbildung 5.2 ist der Persistenzparameter r_1 , der nach Gleichung 2.12 ebenso für den Gesamtzeitraum berechnet wurde, für die einzelnen Jahreszeiten aufgetragen. Der Persistenzparameter gibt an, um wieviel p_{11} größer ist als p_{01} . Im Sommer ergeben sich für die Differenz dieser beiden Wahrscheinlichkeiten die niedrigste Werte, die abgesehen von den Küstengebieten an der Nordsee und von Nordrhein-Westfalen bei etwa 30% liegen. Im Frühling und Herbst ergibt sich etwa das gleiche Bild. Jedoch ist im Frühling der Persistenzparameter im Nordwesten bei Werten um 45% stärker ausgeprägt und auch im Nordosten finden sich größere Werte als im Herbst.

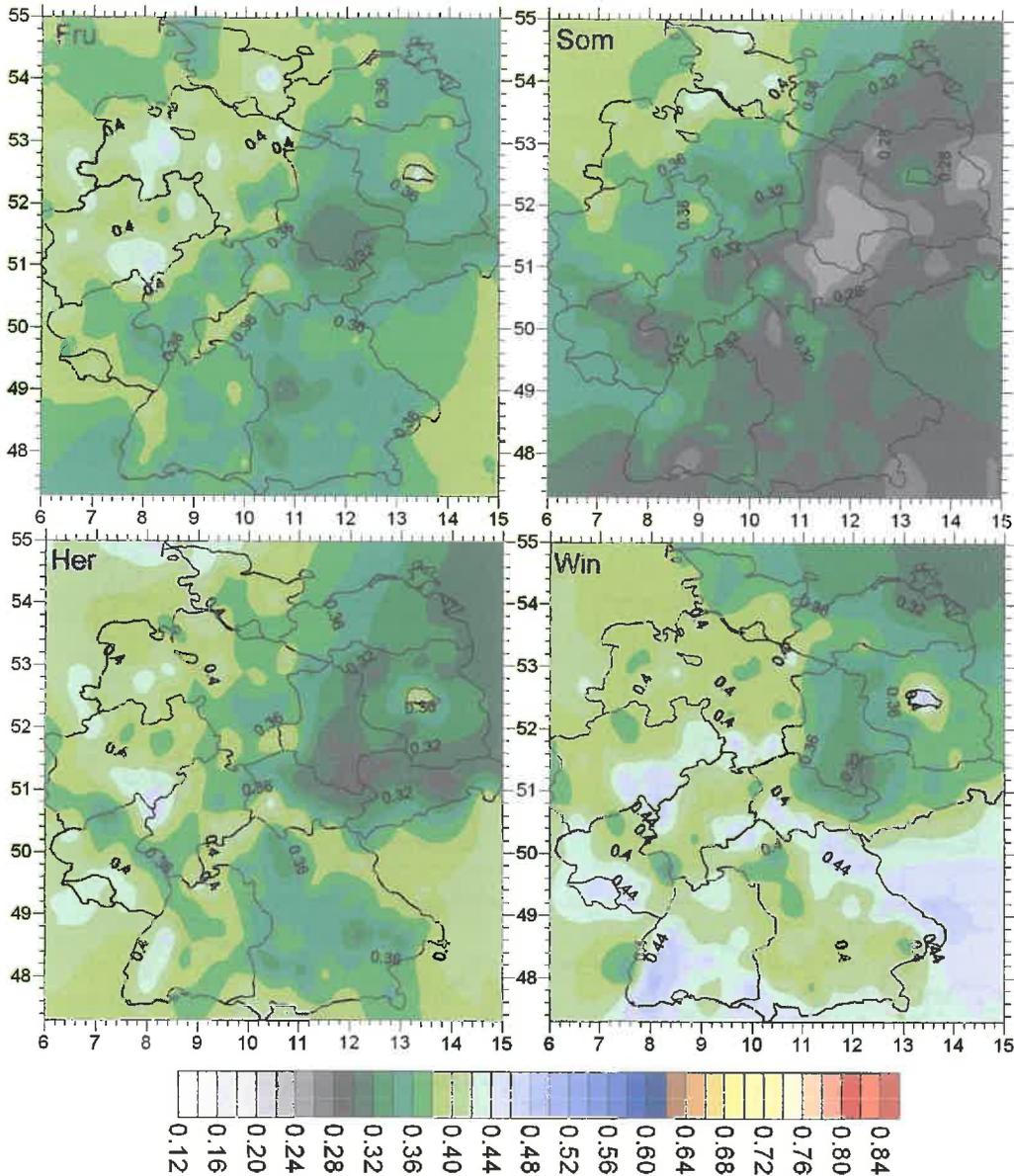


Abbildung 5.2: Persistenz Parameter r_1 berechnet für den Zeitraum 1951-2000 jeweils für den Frühling, den Sommer, den Herbst und den Winter

Im Winter erkennt man Höchstwerte bis zu 48%, die sich stark an der Orographie orientieren (Westerwald, Schwarzwald und Erzgebirge). Auch ist die Tatsache bemerkenswert, dass in allen vier Jahreszeiten Berlin mit höheren Werten als im Umland in Erscheinung tritt.

In Abbildung 5.3 ist abschließend der Trend für die Anzahl der Regentage nach Mann Kendall Test (vgl. Anhang A.2.3) dargestellt wiederum für die einzelnen Jahreszeiten und den Gesamtzeitraum 1951-2000. Dabei ist der Test so zu interpretieren, dass dieser keine

Aussage über die exakte Art¹ des Trends erlaubt. Vielmehr wird generell eine Zu- oder Abnahme nachgewiesen. In diesem Sinne zeigt sich im Sommer bis auf ein einige Stationen in Bayern und Baden-Württemberg, wo entweder kein oder negativer Trend vorliegt, in ganz Deutschland signifikante Abnahme der Anzahl der Regentage im Bezugszeitraum. In den übrigen Jahreszeiten liegt vorwiegend Zunahme der Anzahl der Regentage vor, wobei aber auch einige Stationen keinen signifikanten Trend aufweisen. Dies gilt hauptsächlich für Stationen in den neuen Bundesländern und für die Winter Jahreszeit.

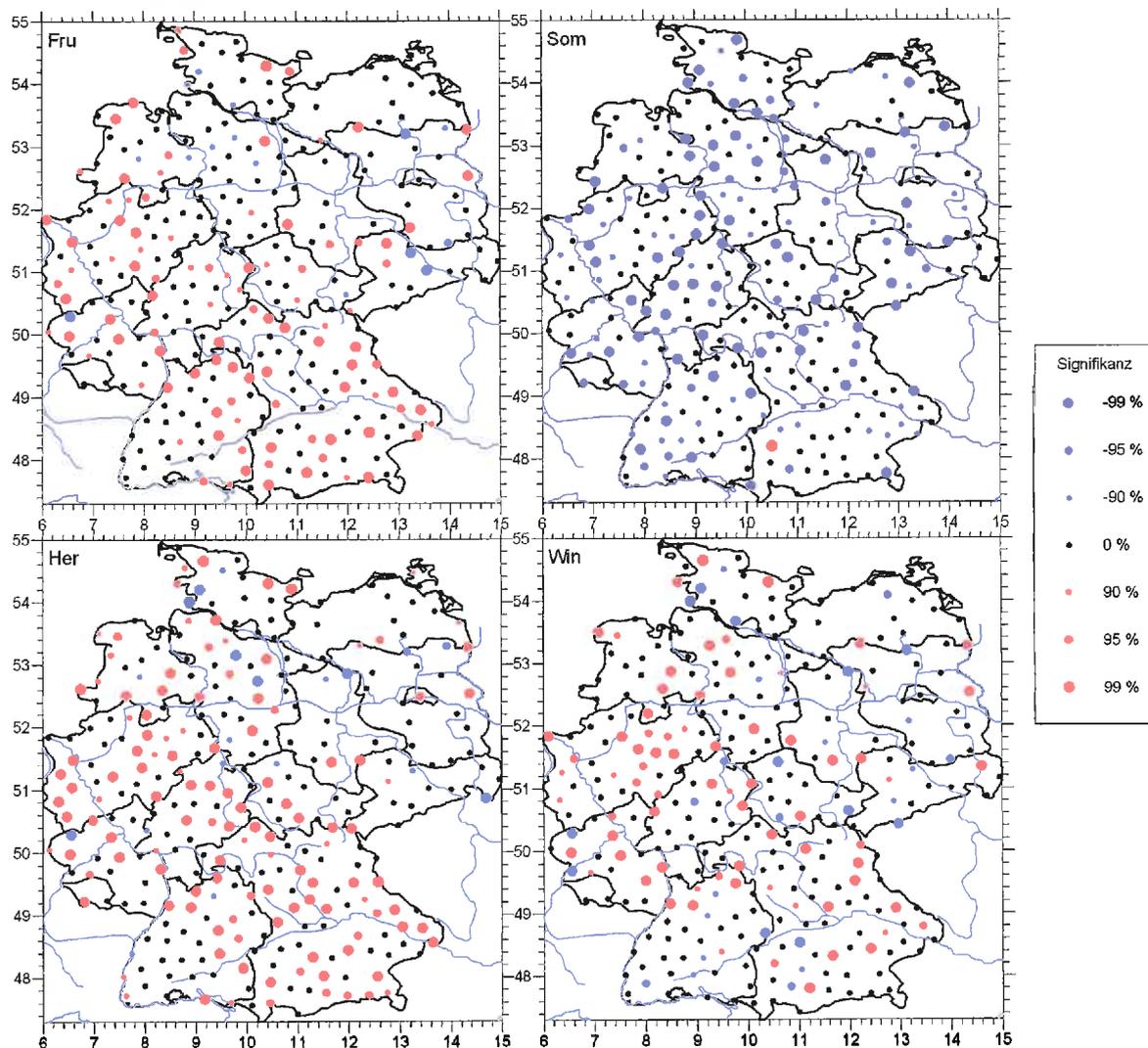


Abbildung 5.3: Trend mit Signifikanzniveau nach Mann-Kendall Test für die Anzahl n_1 der Regentage im Frühling, Sommer, Herbst und Winter für 1951-2000. Blaue Punkte bedeuten signifikante Abnahme und rote Punkte signifikante Zunahme der Regentage.

¹Trends können linear steigend bzw. fallend oder auch progressiv/degressiv steigend bzw. fallend geartet sein. Es ist auch möglich, dass sich verschiedene Trendtypen gegenseitig überlagern.

5.1.2 Detektierte Ordnungen der Markov Ketten

In diesem Unterkapitel werden nun die Ergebnisse gezeigt, die nach dem BIC (Gleichung 2.19) für den Gesamtzeitraum und dessen Hälften auf Basis der Log-Likelihood Funktionen (Gleichungen 2.20 bis 2.23) berechnet wurden. Entsprechende Karten für das AIC und weitere Karten für das BIC, die an dieser Stelle nicht präsentiert werden, befinden sich im Anhang B.1.1 und B.1.2.

Wie in Abbildung 5.4 ersichtlich, werden für alle Stationen und für alle Jahreszeiten - wohl gemerkt für den Gesamtzeitraum - Markov Ketten mit der Ordnung 2 als beste Modellwahl detektiert. Es zeigen sich also keine jahreszeitliche und regionale Unterschiede.

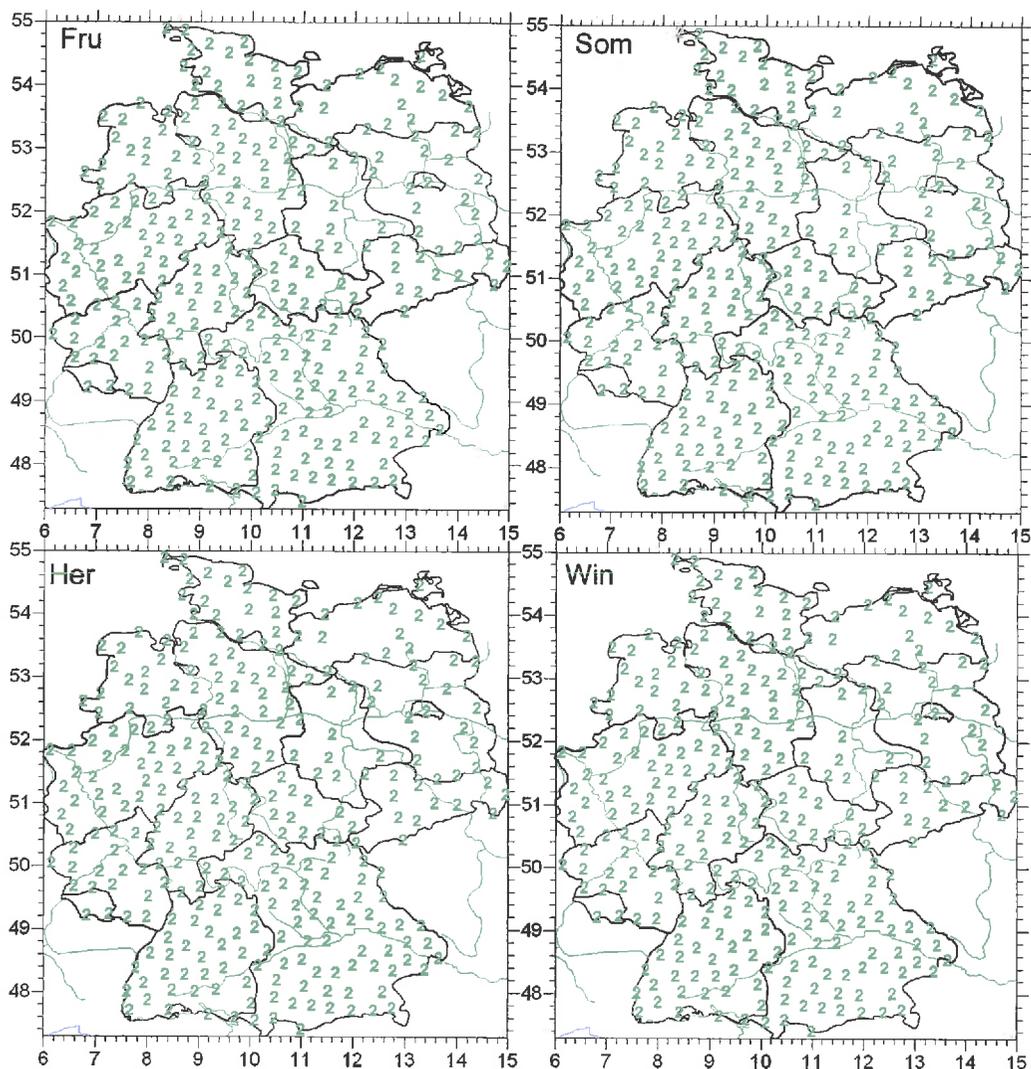


Abbildung 5.4: Ordnungen nach dem BIC für zwei Zustände für Frühling, Sommer, Herbst und Winter (1951-2000)

Für die Modellierung des nächsten, täglichen Niederschlagswertes (Zeitpunkt $t + 1$ in der Zeitreihe) ist damit die Information des jetzigen und des dazu vorherigen Tages (Zeitpunkte t und $t - 1$) ausreichend. Diese Information schließt nur mit ein, ob es an diesen beiden Tagen, an einem der beiden Tage oder ob es an keinem von beiden Tagen geregnet oder nicht. Die Niederschlagsintensität spielt keine Rolle bei der Analyse mit nur zwei Zuständen. Versucht man nun Veränderungen in dem 50-jährigen Beobachtungszeitraum auszumachen, ist es sinnvoll, die in die Ordinalskala transformierte Zeitreihe in zwei Hälften zu zerlegen. Das hat offenbar den Vorteil, dass man zwei gleich lange Zeitreihen hat. Denn der Vergleich zwischen Ordnungen von Markov Ketten, der auf Basis zweier verschiedener Zeitreihen beruht, ist nicht unproblematisch. Verschiedene Autoren (z. B. [8, 11, 12]) weisen darauf hin, dass die Ordnung neben geografischen Gegebenheiten u. a. auch von der Länge als auch von der Jahreszeit abhängt. Darüber hinaus charakterisiert die Ordnung immer nur die mittlere Verteilung der Zustände. Um die Verteilung richtig einzuschätzen, sind daher lange Zeitreihen notwendig. Betrachtet man einzelne Monate - also sehr kurze Reihen - erhält man von der Analyse des gesamten zur Verfügung stehenden Zeitraums abweichende Ergebnisse. Beispielsweise ergibt die Rechnung, die in Kapitel 2.7 für den Monat Juni 2000 für Frankfurt durchgeführt wurde, dass die Ordnung 1 ausgewählt wird. Das Gleiche ist mit allen anderen Stationsreihen und Monaten der Fall. Erst wenn man etwa fünf bis zehn Monate (also einen Zeitraum von fünf bis zehn Jahren) - die Zahl ist von der jeweiligen Station abhängig - zusammenfasst, springt die Ordnung von 1 auf 2 und bleibt dann auch stabil, wenn noch mehr Monate hinzu genommen werden.

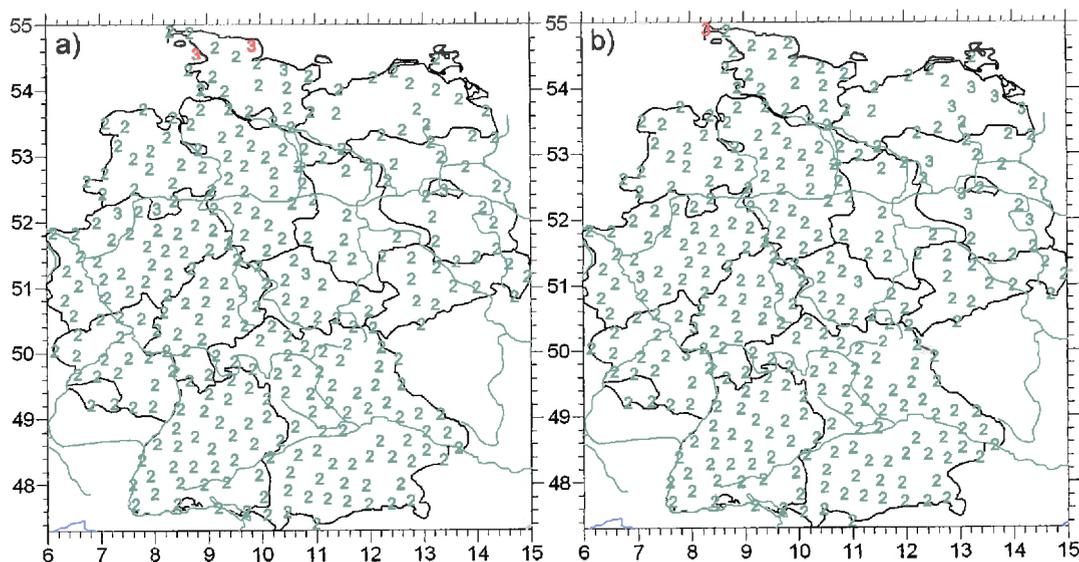


Abbildung 5.5: Vergleich der Ordnungen nach AIC mit zwei Zuständen für den Sommer:
a) 1951-1975 und **b)** 1976-2000

Die Abbildung 5.5 und 5.6 zeigen für den Sommer und den Herbst den Unterschied zwischen die Ordnungen gemäß des AICs zwischen 1951-1975 und 1976-2000, denn hier sind im Gegensatz zum BIC überhaupt Unterschiede, wenn auch nur wenige und regional begrenzt, erkennbar. Im Sommer haben einige Stationen in Mecklenburg-Vorpommern und Brandenburg von der Ordnung 2 in der ersten Hälfte zur Ordnung 3 in der zweiten Hälfte gewechselt, wohingegen im Herbst ein Wechsel von 3 zu 2 vor allem in Südbayern stattgefunden hat. In den übrigen Jahreszeiten haben nur sehr wenige und vereinzelt liegende Stationen die Ordnung der Markov Ketten verändert (Abb. B.2 und B.3). Gleiches gilt für das BIC (Abb. B.4 und B.5).

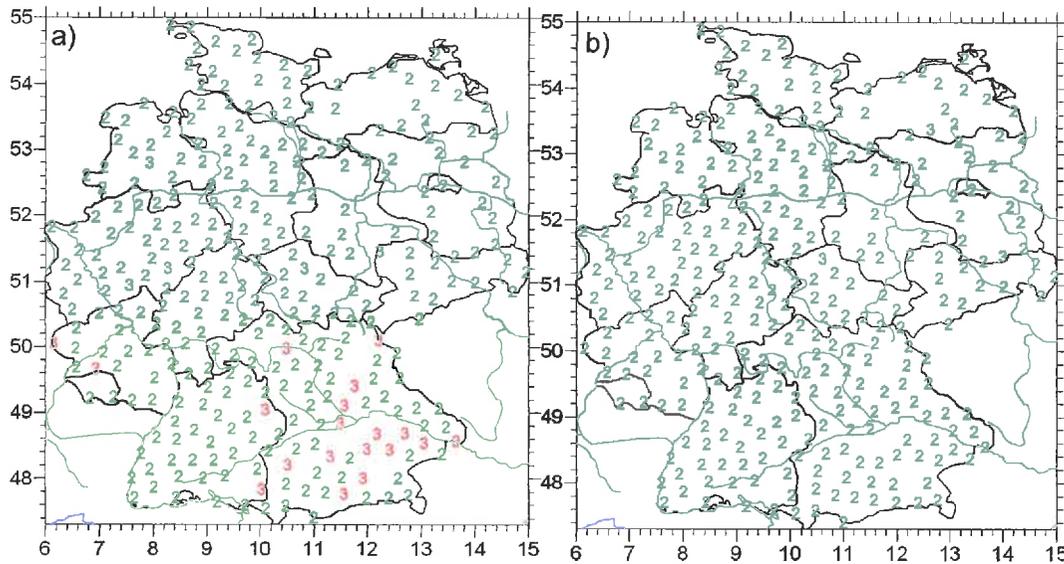


Abbildung 5.6: Vergleich der Ordnungen nach AIC mit zwei Zuständen für den Herbst: a) 1951-1975 und b) 1976-2000

Ebenso ist für die Gesamtstatistik, die sich durch Aufsummieren aller Zustandskombinationen von allen Stationen ergibt, die Ordnung berechnet worden. Sowohl das AIC als auch das BIC finden für 1951-2000, 1951-1975 und für 1976-2000 für alle Jahreszeiten Markov Ketten zweiter Ordnung.

5.1.3 Übergangswahrscheinlichkeiten der Markov Ketten

Zunächst soll in diesem Kapitel auf die Übergangswahrscheinlichkeiten eingegangen werden, wie sie für den Gesamtzeitraum (1951-2000) berechnet wurden (vgl. Kap. 2.6). Dabei werden nur die höchsten Wahrscheinlichkeiten vorgestellt. In Abbildung 5.7 ist die Übergangswahrscheinlichkeit p_1 nach Jahreszeiten getrennt für den Gesamtzeitraum in Kartenform dargestellt. Frühling und Sommer sind strukturell ähnlich. Höchste Werte mit über 48% treten im Nordwesten und im Süden Deutschlands auf. Die neuen Bun-

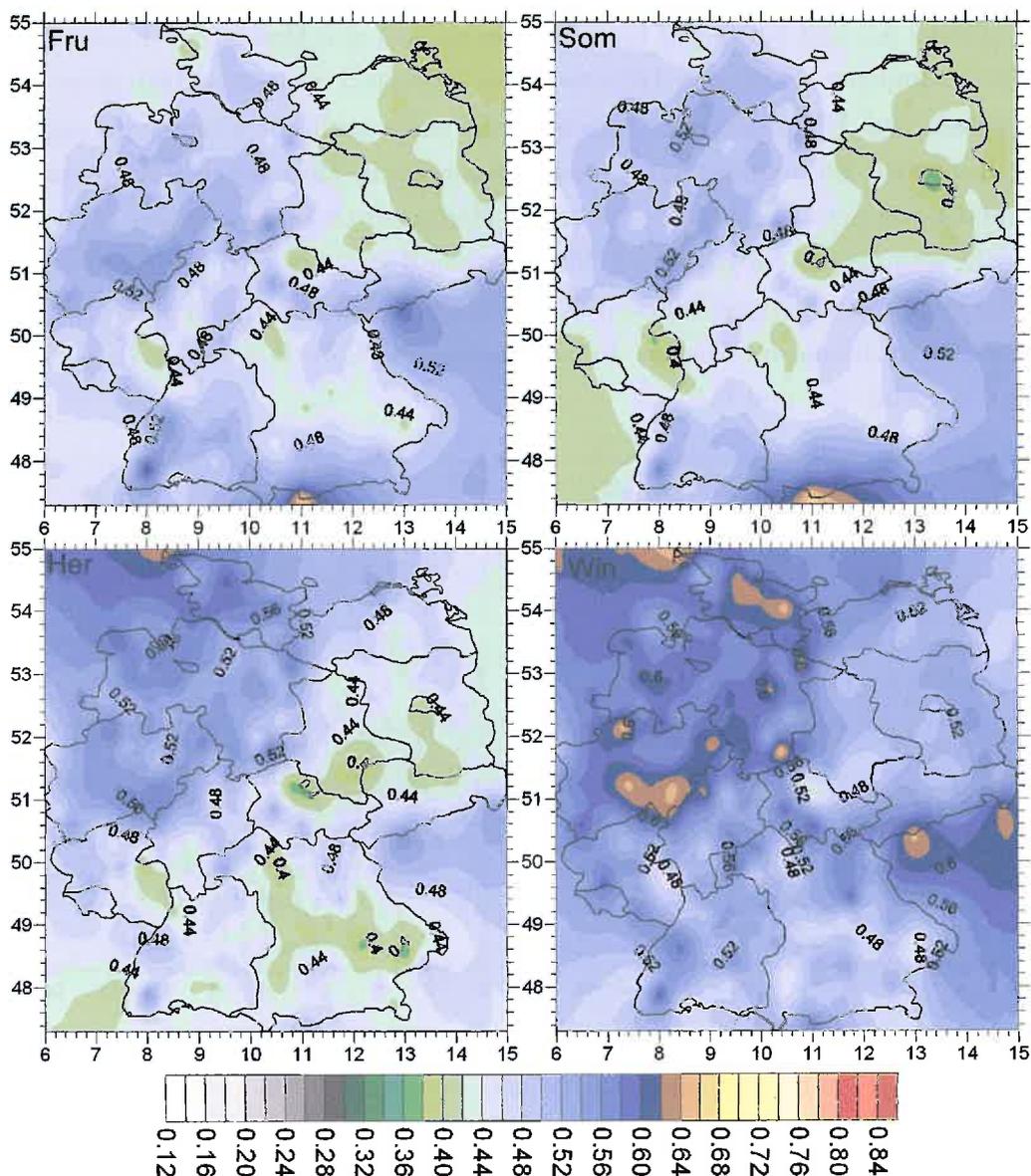


Abbildung 5.7: Wahrscheinlichkeit p_1 für 1951-2000 im Frühling, Sommer, Herbst und Winter (von oben links nach unten rechts)

desländer sind mit ungefähr 40% die Gebiete mit der geringsten Wahrscheinlichkeit für p_1 . Auch im Herbst zeigt sich ein ähnliches Bild. Aber hier liegen die minimalen Werte in Bayern. Das minimalen Flächen im Osten sind dagegen weniger stark ausgeprägt als noch im Sommer.

In Abbildung 5.8 ist die Übergangswahrscheinlichkeit p_{00} gezeigt. Die gelben Gebiete sind die mit der höchsten Wahrscheinlichkeit (um 70%), die blauen die mit der geringsten Wahrscheinlichkeit (um 58%). Wegen der Verteilung von p_1 - und der damit zusam-

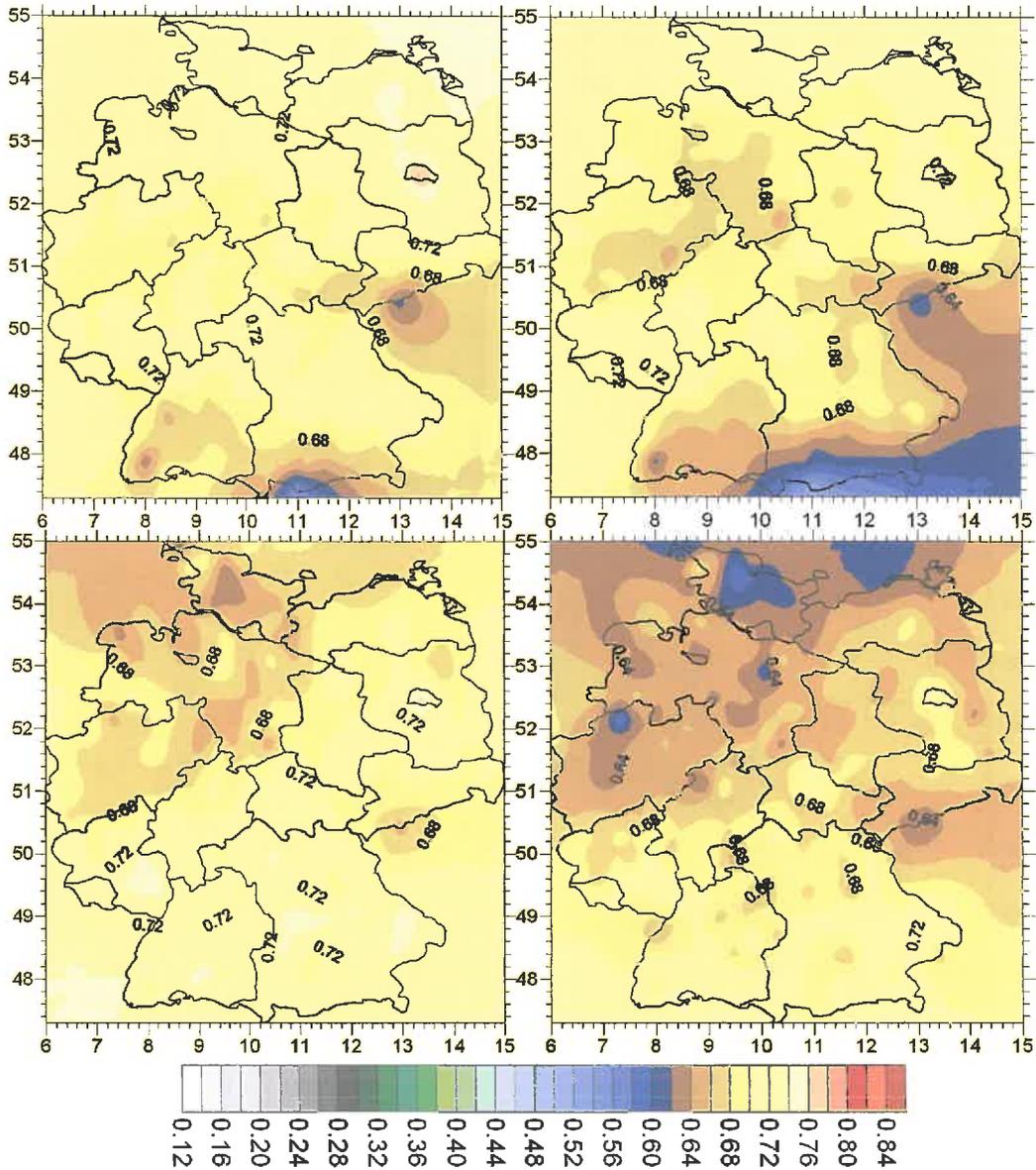


Abbildung 5.8: Wahrscheinlichkeit p_{00} für 1951-2000 im Frühling, Sommer, Herbst und Winter (von oben links nach unten rechts)

menhängenden, komplementären Verteilung von p_0 - zeigt sich auch für p_{00} im Frühling und im Sommer ein ähnliches Bild, wobei im Frühling in Nordrhein-Westfalen und in Niedersachsen höhere Wahrscheinlichkeiten (70% bis 72%) berechnet wurden als im selben Gebiet im Sommer. Über das Jahr gesehen findet man die niedrigsten Wahrscheinlichkeiten (64% und weniger) im Sommer im Süden und im Winter im Nordwesten Deutschlands. Im Herbst ist ebenfalls im Nordwesten die Wahrscheinlichkeit geringer, dass ein Trockentag auf einen Trockentag folgt.

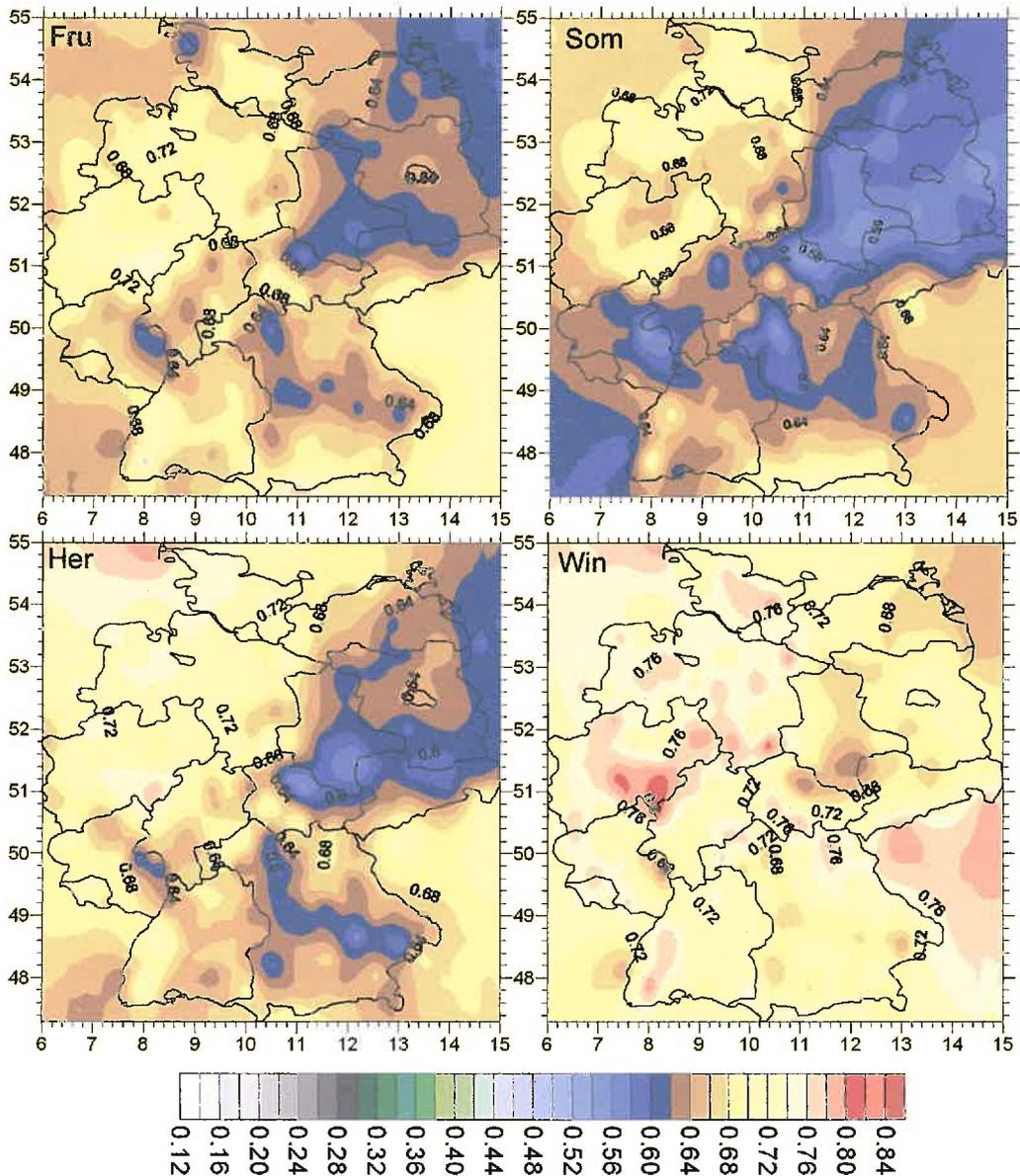


Abbildung 5.9: Wahrscheinlichkeit p_{11} für 1951-2000 im Frühling, Sommer, Herbst und Winter (von oben links nach unten rechts)

In Abbildung 5.9 ist die Wahrscheinlichkeit p_{11} aufgetragen. Hier erkennt man Ähnlichkeiten zwischen Frühling und Herbst. Im Herbst allerdings ist es im Nordwesten mit 76% deutlich wahrscheinlicher, dass auf einen Regentag ein Regentag folgt. Minimale Werte finden sich zum wiederholten Mal im Osten und im Süden Deutschlands - hier vor allem in Bayern - bei Werten um 64% und weniger. Im Sommer tritt die Wahrscheinlichkeit für p_{11} in diesen Gebieten teils mit Werten unter 58% noch großflächiger auf als noch im Frühling. Insgesamt bleibt also festzuhalten, dass es in allen Jahreszeiten im Westen und Nordwesten - besonders aber im Winter - am wahrscheinlichsten ist, dass nach einem Regentag

wieder ein Regentag auftritt. Zusätzliche Karten für p_0 , p_{000} , p_{011} , p_{101} und p_{111} sowie für p_{0000} und p_{1111} befinden sich im Anhang B.1.2 (Abb. B.4 bis B.10).

Jahresgang der Übergangswahrscheinlichkeiten

Obwohl die Niederschlagswerte zeitlich und räumlich stark variabel sind, lohnt es sich einen Blick auf den Jahresgang einiger Übergangswahrscheinlichkeiten zu werfen, weil hier gewisse Strukturen erkennbar sind. Dazu wurde zum einen der mittlere Jahresgang aller 301 Niederschlagsdaten in monatlicher Auflösung berechnet und zum anderen der Jahresgang von den einzelnen Stationen ebenfalls in monatlicher Auflösung, wovon sechs aufgrund ihrer geografischen Lage ausgewählt wurden und hier präsentiert werden.

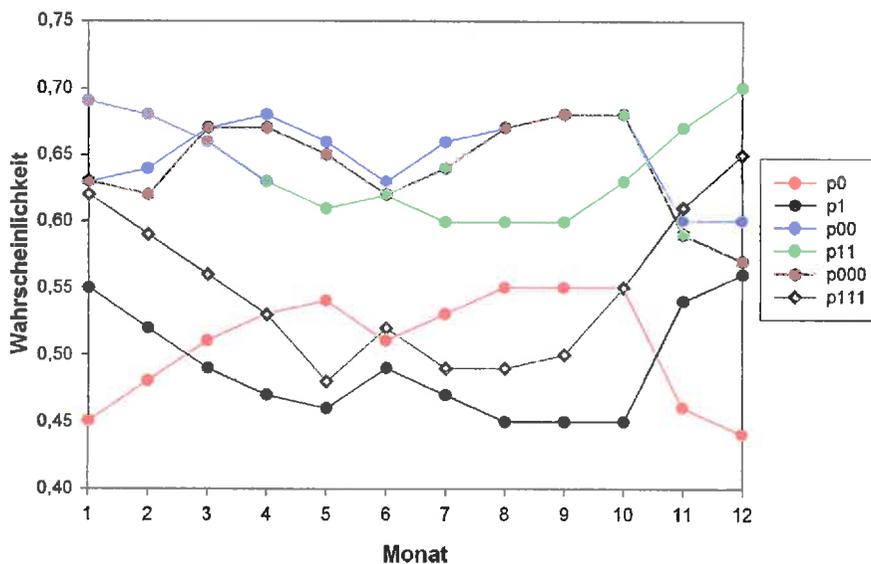


Abbildung 5.10: mittlerer Jahresgang einiger Übergangswahrscheinlichkeiten berechnet auf Basis von allen 301 Niederschlagsstationen im Zeitraum 1951-2000.

In Abbildung 5.10 sind die mittleren Übergangswahrscheinlichkeiten 0., 1., und 2. Ordnung für ganz Deutschland (1951-2000) gegen die Monate aufgetragen. Man erkennt, dass in allen Monaten - bis auf die Wintermonate - für p_0 höhere Werte registriert werden als für p_1 . Der Verlauf von p_0 zeigt ein Maximum im Herbst und ein relatives Maximum im Monat Mai. Trockentage sind - vor allem im Spätsommer - um bis zu 8% wahrscheinlicher als Regentage. Von Oktober auf November fällt p_0 drastisch ab, um im Dezember bei weniger als 45% das Minimum zu erreichen. Des Weiteren ist p_{11} - also die Wahrscheinlichkeit, dass nach einem Regentag erneut ein Regentag auftritt - im Spätsommer am geringsten und im Winter am größten.

In Tabelle 6 sind alle mittleren Übergangswahrscheinlichkeiten allerdings nur in jahreszeitlicher Auflösung aufgelistet. Betrachtet man nicht mehr einzelne Tage bzw. Zustände sondern Kombinationen aus diesen, so können die Wahrscheinlichkeiten nicht mehr direkt miteinander verglichen werden. Hier ist vielmehr von Interesse, wie wahrscheinlich es ist, dass ein Zustand auf einen oder auch auf mehrere andere Zustände folgt.

Das bedeutet z. B. im Winter: Ein Regentag tritt mit $p_1 = 0.54$ auf (vgl. Tab. 6). Ein Regentag folgt einem Regentag mit $p_{11} = 0.73$. Ebenso entnimmt man Tabelle 6 $p_{111} = 0.75$. Das bedeutet, dass es bei Eintritt eines Regentages sehr viel wahrscheinlicher ist, dass ein weiterer Regentag folgt (73%) als ein Trockentag (27%) und auf diese dann nunmehr zwei Regentage wiederum ein Regentag folgt (75%). Hat man also die Kombination „11“ wird in 3 von 4 Fällen eine weitere „1“ folgen. Irgendwann wird der unwahrscheinlichere Fall eintreten und ein Trockentag auf einen Regentag folgen. Ebenso erkennt man (vgl. Tab. 7), dass auch die Trockentage dazu neigen, sich in Clustern aneinanderzureihen ($p_{00} = 0.68$, $p_{000} = 0.72$).

Insgesamt sollten im Winter die Serien von Regentagen aufgrund der höheren Eintrittswahrscheinlichkeit dominant sein. Im Sommer ist das Bild umgekehrt. Hier sind die Serien von Trockentagen vorherrschend.

Tabelle 6: Aufgelistet ist die Gesamtstatistik der Übergangswahrscheinlichkeiten für alle 301 Niederschlagsstationen nach den Jahreszeiten getrennt (1951-2000) für 0., 1., 2. und 3. Ordnung. Die jeweils größten Wahrscheinlichkeiten sind hellgrau unterlegt.

Wahrscheinlichkeit	Frühling	Sommer	Herbst	Winter
p_0	0.53	0.53	0.52	0.46
p_1	0.47	0.47	0.48	0.54
p_{00}	0.71	0.69	0.70	0.68
p_{01}	0.29	0.31	0.30	0.32
p_{10}	0.33	0.35	0.32	0.27
p_{11}	0.77	0.65	0.68	0.73
p_{000}	0.74	0.71	0.74	0.72
p_{001}	0.26	0.29	0.26	0.28
p_{010}	0.34	0.38	0.36	0.31
p_{011}	0.66	0.62	0.64	0.69
p_{100}	0.64	0.63	0.61	0.59
p_{101}	0.36	0.37	0.39	0.41
p_{110}	0.32	0.34	0.30	0.25
p_{111}	0.68	0.66	0.70	0.75
p_{0000}	0.75	0.73	0.76	0.73
p_{0001}	0.25	0.27	0.24	0.27
p_{0010}	0.35	0.40	0.38	0.32
p_{0011}	0.65	0.60	0.62	0.68
p_{0100}	0.56	0.52	0.53	0.59
p_{0101}	0.44	0.48	0.47	0.47
p_{0110}	0.35	0.38	0.34	0.29
p_{0111}	0.65	0.62	0.66	0.71
p_{1000}	0.70	0.68	0.68	0.67
p_{1001}	0.30	0.32	0.32	0.33
p_{1010}	0.32	0.35	0.34	0.30
p_{1011}	0.68	0.65	0.66	0.70
p_{1100}	0.64	0.62	0.61	0.59
p_{1101}	0.36	0.38	0.39	0.41
p_{1110}	0.31	0.32	0.29	0.24
p_{1111}	0.69	0.68	0.71	0.76

Nun ist zu erwarten, dass der Jahresgang einzelner Stationen mitunter erheblich vom mittleren Jahresgang aller Stationen abweichen kann. Deshalb ist in Abbildung 5.11 für eine Auswahl an Stationen der lokale Jahresgang dargestellt. Gezeigt sind Bremen und Sankt Peter-Ording als küstennahe Stationen, Frankfurt und Potsdam als Stationen, die mehr oder weniger weder maritim noch orografisch beeinflusst sind, und der Feldberg (Schwarzwald) sowie die Zugspitze aufgrund der orografischen Lage.

Sankt Peter-Ording ist dabei ein Beispiel für die höchste Wahrscheinlichkeit p_0 im Frühling und Frühsommer (knapp 60%) und die niedrigste im Winter (um 35%). Hier ist der Jahresgang stark ausgeprägt. In Bremen halten sich Trocken- und Regentage im April und Mai die Waage, in den übrigen Monaten - vor allem im Winter - sind Regentage bis zu 20% wahrscheinlicher als Trockentage.

Der Jahresgang in Frankfurt und Potsdam ist in etwa vergleichbar. Von April bis Oktober sind die Wahrscheinlichkeiten auf ähnlichem Niveau, wobei in Frankfurt Trockentage im Herbst um etwa 8% wahrscheinlicher sind als im Frühling. Der Hauptunterschied in den in Abbildung 5.11 gezeigten Wahrscheinlichkeiten liegt zwischen dem Winter und den übrigen Monaten des Jahres.

Die Zugspitze ist die Station mit der höchsten Regenwahrscheinlichkeit im Sommer (Maximum im Juni über 70%) und hat einen ausgeprägten Jahresgang mit 30% Unterschied in der Wahrscheinlichkeit p_1 zwischen Juni und Oktober. Einzig im Oktober sind Trockentage wahrscheinlicher als Regentage. Am Feldberg (Schwarzwald) ist die Wahrscheinlichkeit für Trockentage im Herbst um etwa 10% größer als im Frühling, liegt aber insgesamt unter 50%.

In Tabelle 7 sind noch weitere Übergangswahrscheinlichkeiten nach Jahreszeiten getrennt für die oben genannten sechs Stationen aufgelistet. Im Frühling werden die höchsten Werte für $p_{111} > 0.8$ für den Feldberg und die Zugspitze registriert. Solche Werte treten in der Analyse der Niederschlagsdaten mit zwei Zuständen nur noch im Westerwald und an vereinzelt Stationen in Niedersachsen und Schleswig-Holstein auf (vgl. Abb. B.8). Potsdam, als Station in den Gebieten mit minimaler Wahrscheinlichkeit p_1 gelegen (vgl. Abb. 5.7), erweist sich auch in dieser Auflistung (Tab. 8) als Ort mit den geringsten Wahrscheinlichkeiten, und dies vor allem im Sommer. Die Wahrscheinlichkeiten p_{00} , p_{000} und p_{0000} sind in den meisten Fällen größer als p_{11} , p_{111} und p_{1111} . Ausnahme bilden hier die Bergstationen vom Feldberg und der Zugspitze.

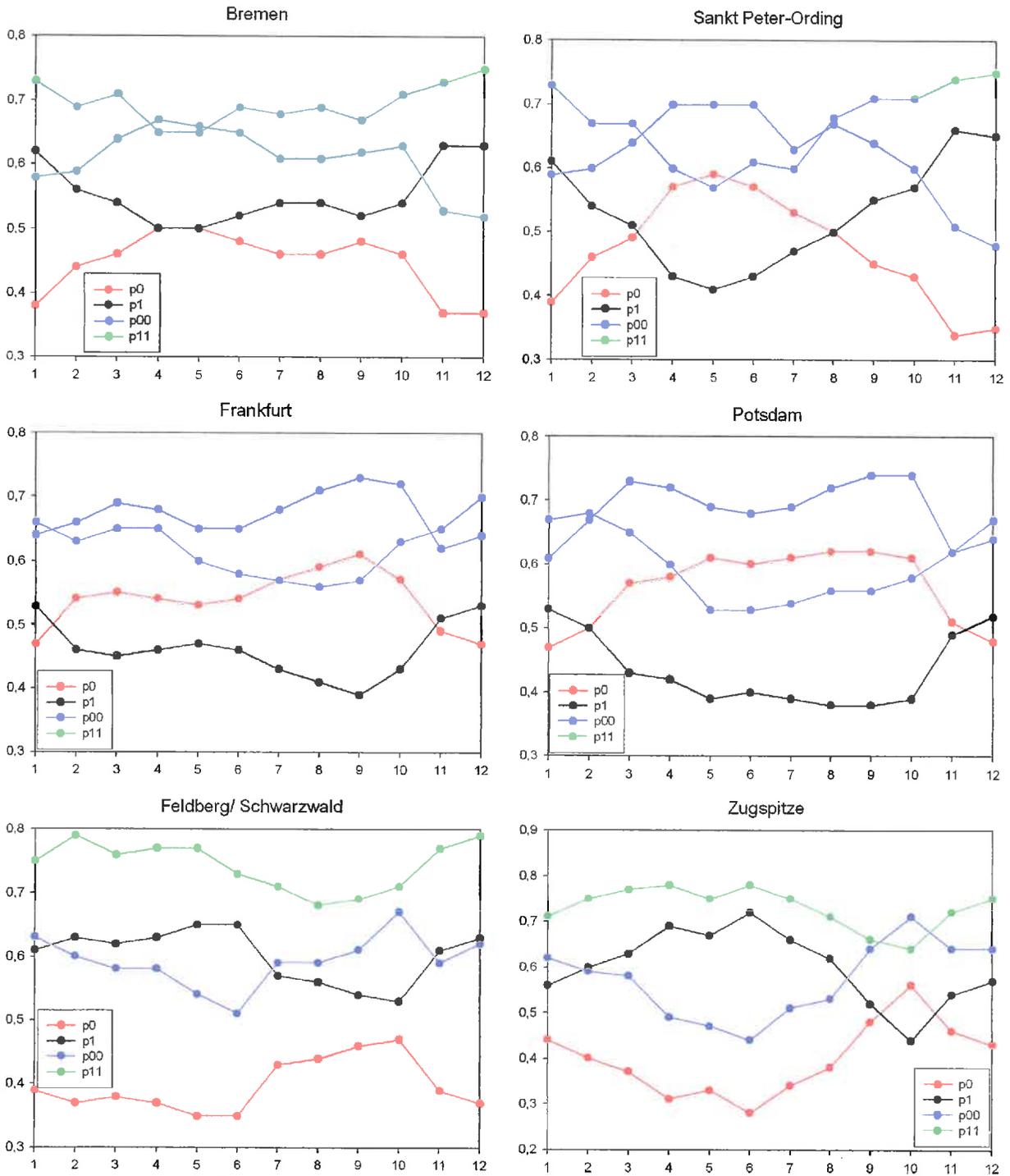


Abbildung 5.11: Jahrgang von p_0 , p_1 , p_{00} und p_{11} für Bremen, Sankt Peter-Ording, Frankfurt, Potsdam, Feldberg (Schwarzwald) und Zugspitze berechnet auf Basis von 1951-2000. Die Ordinate zeigt die Wahrscheinlichkeit an und die Abszisse den jeweiligen Monat.

Tabelle 7: Übersicht der größten Übergangswahrscheinlichkeiten getrennt nach den Jahreszeiten (1951-2000) von ausgewählten Stationen mit Höhenangabe bezogen auf Meeresspiegelniveau (NN)

Frühling								
Station	NN [m]	p_1	p_{00}	p_{11}	p_{000}	p_{111}	p_{0000}	p_{1111}
Sankt Peter-Ording	4	0.45	0.72	0.67	0.75	0.62	0.64	0.66
Bremen	2	0.51	0.70	0.71	0.74	0.70	0.60	0.71
Potsdam	81	0.41	0.74	0.64	0.77	0.61	0.68	0.65
Frankfurt	113	0.46	0.70	0.65	0.74	0.68	0.63	0.63
Feldberg/Schwarzwald	1486	0.63	0.62	0.78	0.65	0.80	0.56	0.77
Zugspitze	2960	0.66	0.56	0.78	0.60	0.82	0.52	0.77
Sommer								
Station	NN [m]	p_1	p_{00}	p_{11}	p_{000}	p_{111}	p_{0000}	p_{1111}
Sankt Peter-Ording	4	0.46	0.71	0.67	0.75	0.64	0.61	0.68
Bremen	2	0.53	0.67	0.72	0.70	0.64	0.61	0.74
Potsdam	81	0.39	0.73	0.59	0.76	0.58	0.68	0.60
Frankfurt	113	0.43	0.71	0.61	0.73	0.61	0.65	0.61
Feldberg/Schwarzwald	1486	0.59	0.61	0.72	0.63	0.73	0.57	0.72
Zugspitze	2960	0.67	0.52	0.76	0.53	0.79	0.52	0.75
Herbst								
Station	NN [m]	p_1	p_{00}	p_{11}	p_{000}	p_{111}	p_{0000}	p_{1111}
Sankt Peter-Ording	4	0.59	0.64	0.75	0.69	0.68	0.53	0.77
Bremen	2	0.57	0.73	0.73	0.70	0.70	0.56	0.74
Potsdam	81	0.42	0.74	0.64	0.76	0.61	0.67	0.65
Frankfurt	113	0.44	0.73	0.66	0.77	0.64	0.61	0.67
Feldberg/Schwarzwald	1486	0.56	0.68	0.75	0.73	0.74	0.59	0.76
Zugspitze	2960	0.50	0.71	0.71	0.75	0.71	0.63	0.70
Winter								
Station	NN [m]	p_1	p_{00}	p_{11}	p_{000}	p_{111}	p_{0000}	p_{1111}
Sankt Peter-Ording	4	0.60	0.63	0.75	0.67	0.70	0.56	0.77
Bremen	2	0.60	0.63	0.76	0.66	0.70	0.57	0.78
Potsdam	81	0.52	0.69	0.71	0.72	0.71	0.62	0.71
Frankfurt	113	0.51	0.69	0.71	0.73	0.67	0.61	0.72
Feldberg/Schwarzwald	1486	0.62	0.68	0.81	0.71	0.78	0.61	0.82
Zugspitze	2960	0.58	0.68	0.77	0.72	0.75	0.59	0.77

5.1.4 Zeitliche Veränderungen der Übergangswahrscheinlichkeiten

Um die zeitliche Entwicklung zu beurteilen, wird der lineare Trend berechnet und die Steigung der Regressionsgeraden auf Signifikanz geprüft (vgl. Anhang A.3.1). Die lineare Regression durch die Methode der kleinsten Quadrate setzt voraus, dass die Übergangswahrscheinlichkeiten normalverteilt und unabhängig voneinander sind. Die Unabhängigkeit ist dadurch sichergestellt, dass die jeweilige Übergangswahrscheinlichkeit pro Jahr (bezogen auf die jeweilige Jahreszeit) getrennt berechnet wird. Dass die Wahrscheinlichkeiten normalverteilt sind, wird im Allgemeinen - insbesondere bei instationären Reihen - nicht erfüllt sein. Des Weiteren ist aufgrund der Natur der täglichen Niederschlagsdaten zu erwarten, dass starke interannuäre Variationen in der Verteilung der Regen- und Trockentage existieren, die vom linearen Trend nicht erfasst werden.

Dennoch soll der lineare Trend verwendet werden, um Aussagen anhand der Steigung der Regressionsgerade über zeitliche Veränderungen im Zeitraum von 1951 bis 2000 machen zu können. Da für alle Stationen in allen Jahreszeiten Markov Ketten 2. Ordnung angepasst wurden (vgl. Kapitel 5.1.2) - insbesondere auch in den beiden Hälften des Beobachtungszeitraums und ebenso für die Gesamtstatistik aller Stationen - , werden die Trends in grafischer Form nur für die 2. Ordnung vorgestellt und zwar nur die, die auf eine „1“ enden. Denn es reicht aus, nur eine von zwei miteinander gekoppelten Übergangswahrscheinlichkeiten zu präsentieren. Diese Abbildungen, die die linearen Trends für die Gesamtstatistik sowie für die Stationen Bremen, Frankfurt und Zugspitze zeigen, befinden sich im Anhang B.1.4.

In den Tabellen 8 bis 9 sind darüber hinaus die Werte für den absoluten Trend (Differenz zwischen dem letzten und dem ersten Funktionswert der Regressionsgerade) der Übergangswahrscheinlichkeiten 0., 1. und 2. Ordnung sowie die Signifikanz aufgelistet, die die Wahrscheinlichkeit angibt, mit der die jeweilige Steigung - statistisch gesehen - identisch mit Null ist. Tabellen für Frankfurt und die Zugspitze befinden sich im Anhang B.1.4. Die 3. Ordnung wird nicht gezeigt, weil zum einen nach dem BIC diese Ordnung nicht ausgewählt wird und zum anderen treten die entsprechenden Zustandskombinationen pro Station und pro Jahr nur sehr selten oder gar nicht auf, was die Trendanalyse problematisch macht. Anschließend werden Karten für die geografische Verteilung der Steigungen der Regressionsgeraden von p_{000} sowie von p_{111} präsentiert. Karten für die Steigungen von p_1 , p_{00} und p_{11} sind mit denen für p_{000} und p_{111} qualitativ vergleichbar und werden aus diesem Grund hier nicht gezeigt.

Nun wird mit der Betrachtung der Trends aus der Gesamtstatistik begonnen (Tab. 8).

Zunächst sind alle Trends im einstelligen Prozentbereich und damit im Vergleich zu den drei exemplarisch ausgewählten Stationen nur schwach ausgeprägt. Eine Folge der Aufsummierung der Zustandskombinationen aller Stationen. So nimmt die Wahrscheinlichkeit für Regentage im Sommer um 4% ab und in den anderen Jahreszeiten insbesondere im Herbst zu. Des Weiteren existiert im Sommer ein Zunahme von p_{00} um 5.6%, aber kein bemerkenswerter Trend für p_{11} (+0.5%). Ansonsten zeigen sich für p_{11} in allen Jahreszeiten positive Trends, die Trends für p_{00} hingegen weisen im Frühling (-4%), im Herbst (-3.2%) negatives Vorzeichen auf. Im Winter gibt es keine Veränderungen (-0.6%). Die stärkste Zunahme findet man im Sommer für p_{000} mit 7.2%. Aber auch p_{111} zeigt im Sommer eine leichte Zunahme von 1.5%. Größte Veränderungen für p_{111} ergeben sich im Herbst (+4.1%) und im Frühling (+3.6%).

Tabelle 8: lineare Wahrscheinlichkeitstrends in Prozent für die Übergangswahrscheinlichkeiten 0., 1. und 2. aufgeschlüsselt nach den Jahreszeiten; die Zahl in Klammern gibt die Signifikanz der Steigung der Regressionsgerade an (vgl. Text)

Wahrscheinlichkeit	Frühling	Sommer	Herbst	Winter
p_1	+3.6 0	-4 0	+6.1 0	+2.4 0
p_{00}	-4 0	+5.6 95	-3.2 0	-0.6 0
p_{11}	+3.8 90	+0.5 90	+3.4 90	+2.5 0
p_{000}	+0.3 0	+7.2 0	-0.3 0	-1.7 0
p_{011}	+3.3 90	-1.8 0	+1.4 0	+2.1 0
p_{100}	+0.2 0	+3 0	-2.2 0	+0.5 0
p_{111}	+3.6 0	+1.5 0	+4.1 95	+2.2 0

Tabelle 9: wie Tabelle 8 jedoch für Bremen

Wahrscheinlichkeit	Frühling	Sommer	Herbst	Winter
p_1	-1.3 0	-11.9 99	-2.9 0	-7.1 0
p_{00}	+1.4 0	+12.2 99	+1.6 0	8.2 0
p_{11}	-0.3 0	-3.7 0	-4.3 0	-3.4 0
p_{000}	+1.3 0	+20.9 0	+4.6 0	+5.4 0
p_{011}	0 0	-5.1 0	-7.4 90	-1.3 0
p_{100}	-3.2 0	-3.4 0	-1.6 0	+7.1 0
p_{111}	-1.5 0	-1.6 0	+2 0	-2.8 0

Als ein Beispiel für eine große Veränderung im Beobachtungszeitraum sind, wie schon in Tab. 8 für die Gesamtstatistik, in Tabelle 9 die Trends der Übergangswahrscheinlichkeiten für Bremen aufgelistet. Vergleichsweise stark ausgeprägte Trendwerte sind im Sommer

auszumachen. Hier nimmt p_1 um 11.9% ab. Große Zunahmen ergeben sich für p_{00} (+12.2%) und p_{000} (+20.9%). p_{11} nimmt in allen Jahreszeiten ab und p_{111} nimmt mit Ausnahme vom Herbst in allen übrigen Jahreszeiten leicht ab. Der Winter zeigt ähnliche Tendenzen wie der Sommer, nur ist die Zunahme von p_{00} mit 8.2% und von p_{000} mit 5.4% weniger stark ausgeprägt.

In der Abbildung 5.12 (linkes Bild) der zeitliche Verlauf der Übergangswahrscheinlichkeit p_{000} , die für Bremen im Sommer berechnet wurde, mit entsprechender linearer und quadratischer Regressionskurve dargestellt. In den 1950er Jahren hat im zeitlichen Verlauf von p_{000} ein starker Anstieg von etwa 40% auf 70% stattgefunden, überlagert von starken Jahr zu Jahr Schwankungen. Der quadratische Trend reproduziert den Anstieg im ersten Jahrzehnt besser als der lineare, zeigt dann aber ab etwa 1985 einen Abfall. Ab etwa 1960 variiert p_{000} zwischen etwa 60% und 80%, danach ist kaum noch ein Trend auszumachen.

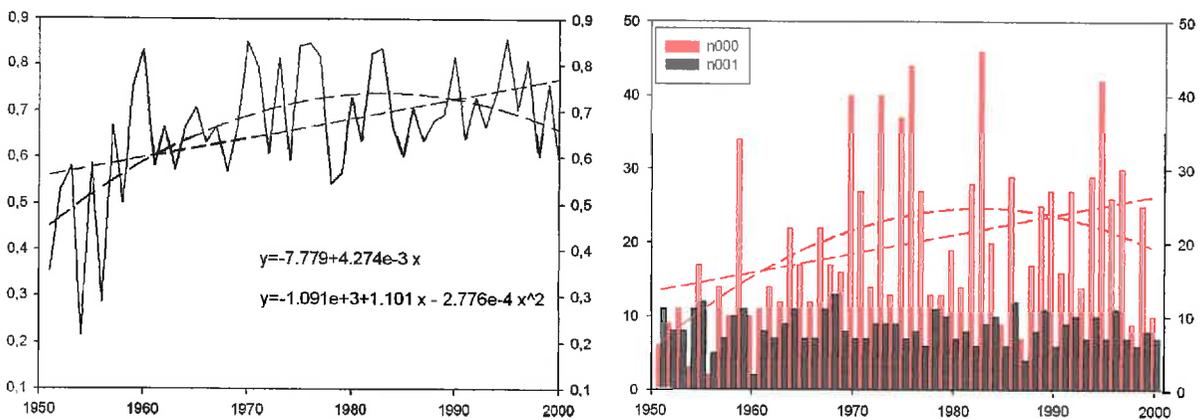


Abbildung 5.12: jährliche Wahrscheinlichkeit p_{000} (links) für Bremen im Sommer (1951-2000) mit linearem und quadratischem Trend; Anzahl der Zustandskombinationen von n_{000} und n_{001} (rechts); vgl. Text

Rechts in Abbildung 5.12 ist der zeitlicher Verlauf von n_{000} (rote Balken) und n_{001} (schwarze Balken) gezeigt, aus denen p_{000} berechnet wurde. Im Gegensatz zu Anzahl n_{001} , bei der keine Veränderungen erkennbar sind, zeigt n_{000} eine Zunahme zwischen 1951 und 2000. Dadurch spiegelt sich der Trend von n_{000} qualitativ in p_{000} wider. Vor allem im ersten Jahrzehnt mit Ausnahme von 1959 liegen beide Häufigkeiten dicht beieinander, was zu großen Schwankungen in den Wahrscheinlichkeiten führt. Das verdeutlicht nochmals die Problematik bei der Beurteilung der Trends der Wahrscheinlichkeiten: Denn insgesamt bei 92 Tagen in den Sommermonaten sind genau 90 Übergangswahrscheinlichkeiten 2. Ordnung möglich. Welche aber in welcher Anzahl realisiert werden, unterliegt sehr großen Jahr zu Jahr Schwankungen. Aus den Wahrscheinlichkeiten kann man prinzipiell nicht auf die Anzahl der zugrunde liegenden Zustandskombinationen schließen (ausgenommen

bei der 0. Ordnung). Umgekehrt können Schwankungen in diesen Häufigkeiten aber zu relativ großen Veränderungen in den Wahrscheinlichkeiten führen. Das führt zur Notwendigkeit, die Zustandskombinationen selbst auf Trends zu untersuchen, was im Rahmen dieser Arbeit nicht möglich ist und daher nur ansatzweise für die Anzahl der Trockentage geschehen konnte (vgl. Abb. 5.3). Jedoch die Zunahme von n_{000} und die Tatsache, dass n_{000} zumindest ab 1960 immer größer ist als n_{001} , impliziert, dass es insgesamt ein Zunahme von einer Serie von drei Trockentagen gibt.

Weil sich für p_{000} und p_{111} die größten und sicherlich auch die interessantesten Veränderungen zeigen - Wahrscheinlichkeiten wie z. B. p_{100} oder p_{011} beschreiben im Prinzip denn Wechsel zwischen Serien von Trocken- bzw. Regentagen - , wird im Folgenden die geografische Verteilung der Steigungen für diese beiden Wahrscheinlichkeiten in den Abb. 5.13 und 5.14 dargestellt. Die Zunahme von p_{000} ist gerade im Sommer im Westen und Nordwesten stark ausgeprägt. Die Steigung beträgt in diesem Gebiet ungefähr 0.004,

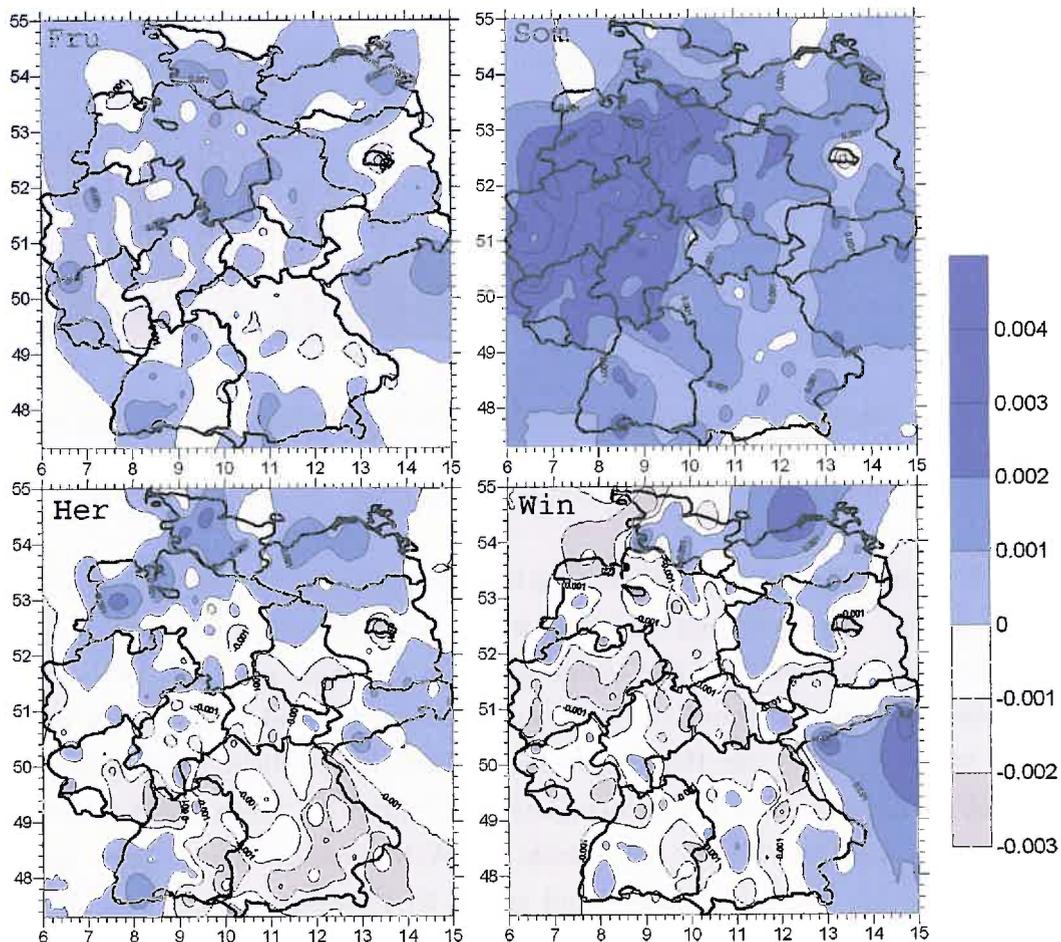


Abbildung 5.13: Steigung der Regressionsgeraden im Frühling, Sommer, Herbst und Winter für p_{000} berechnet für 1951-2000

was einer Zunahme von 20% über den 50-jährigen Zeitraum entspricht. In Berlin tritt hingegen die stärkste Abnahme auf (-15%). In den meisten übrigen Gebieten wird ebenfalls eine Zunahme von p_{000} im Bereich von etwa 5% bis 15% registriert. Im Winter nimmt p_{000} in den meisten Gebieten bis auf den Nordosten ab.

Die Steigungen von p_{111} im Winter zeigen ein gemischtes Bild (Abb. 5.14). Hier treten positive wie auch negative Trends auf. Negative Werte sind vor allem in Baden-Württemberg, Schleswig-Holstein und Sachsen zu finden. Die stärksten Zunahmen, die teilweise über 20% liegen, zeigen sich bei ein paar Stationen im Süden von Sachsen-Anhalt, Rheinland-Pfalz und Thüringen. Im Sommer sind vor allem in Bayern und im Nordwesten Zunahmen auszumachen.

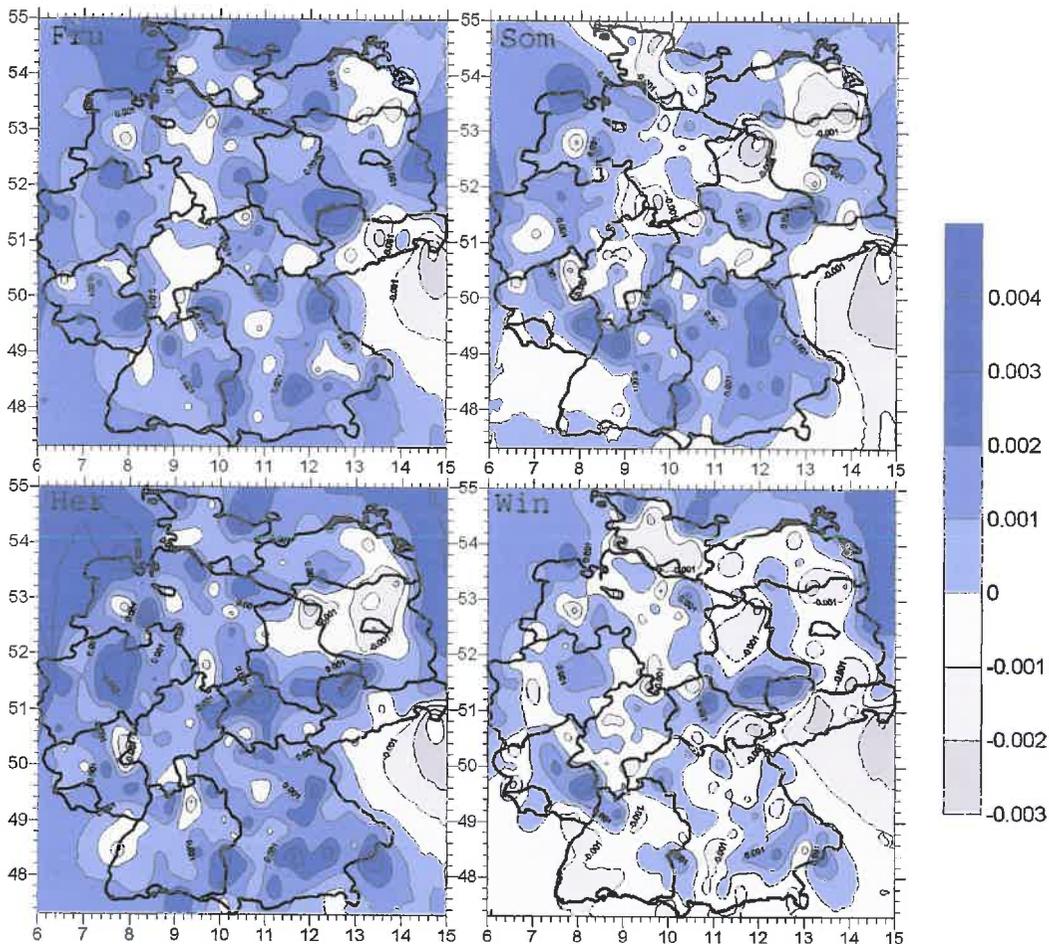


Abbildung 5.14: Wie Abb. 5.13 jedoch für p_{111} im Winter

5.2 Ergebnisse der Analyse für Niederschlagsdaten mit vier Zuständen

5.2.1 Ergebnisse der Anpassung der Weibullverteilung

Bei der Anpassung der Weibullverteilung an die Niederschlagsdaten hat sich gezeigt, dass in einigen Fällen die Anpassung nach der Methode der Momente (MOM) besser gelingt, als diejenige, die nach der Maximum Likelihood Schätzung (MLE). Hier scheint das Newton Raphson Verfahren *aus dem Ruder* zu laufen. Auf die Problematik iterativer Verfahren wird u. a. bei Press et al. [13] hingewiesen (z. B. fehlende Konvergenz gegen die Nullstelle). Jedoch sehen Wilks [26] und Wong [28] diese Vorgehensweise als prinzipiell geeignet an. Für die Auswertung bedeutet das, dass die durch MLE errechneten Verteilungsparameter nur dann verwendet werden können, wenn diese die Startwerte auch wirklich verbessern und nicht verschlechtern. Die MLE ist für vergleichsweise kleine Streuparamater besser geeignet. Die MOM hingegen liefert für größere Streuparamater bessere Anpassungen. Dieser Befund ist auch in Übereinstimmung mit Wilks [26]. In Abbildung 5.15 ist die Anpassung für die Station Altenahr für den März und Juni (1951-2000) gezeigt. Im Juni wird die Anpassung nach der MOM mit dem Komogorov-Smirnov Test auf dem Niveau von 99% als signifikant eingestuft. Die Verbesserung durch die MLE schlägt aber fehl. Im März hingegen ist die Anpassung per MOM schon recht gut, wird aber mit der MLE weiter verbessert (Signifikanzniveau 99%).

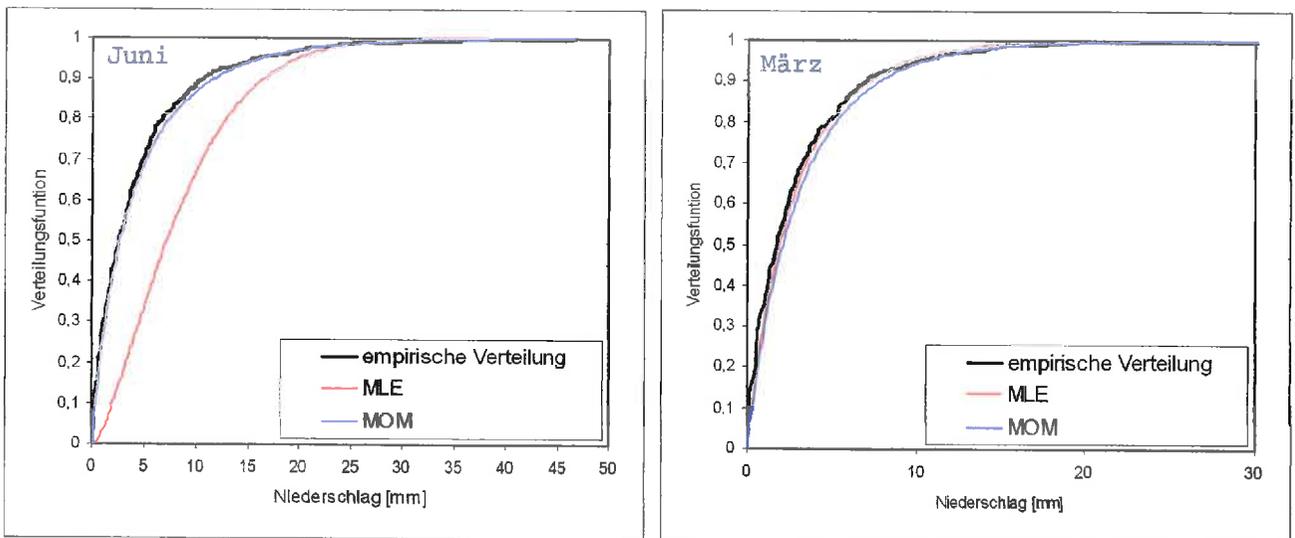


Abbildung 5.15: Verteilungsfunktionen der Weibullverteilung für die Station Altenahr im März (rechts) und im Juni (links) im Zeitraum 1951-2000, die bestimmt wurden durch die empirischen Niederschlagsdaten und durch Berechnung der theoretischen Verteilungen nach MOM und MLE.

Für die Einteilung der Zustände wird nun die am besten angepasste Verteilung benutzt, gleichgültig durch welche Methode sie erreicht worden ist. In Abbildung 5.16 ist für alle zwölf Monate das Signifikanzniveau nach dem Kolmogorov-Smirnov Test zeigt, wobei die signifikanteste Anpassung - also entweder durch MOM oder MLM - verwendet wurde. In

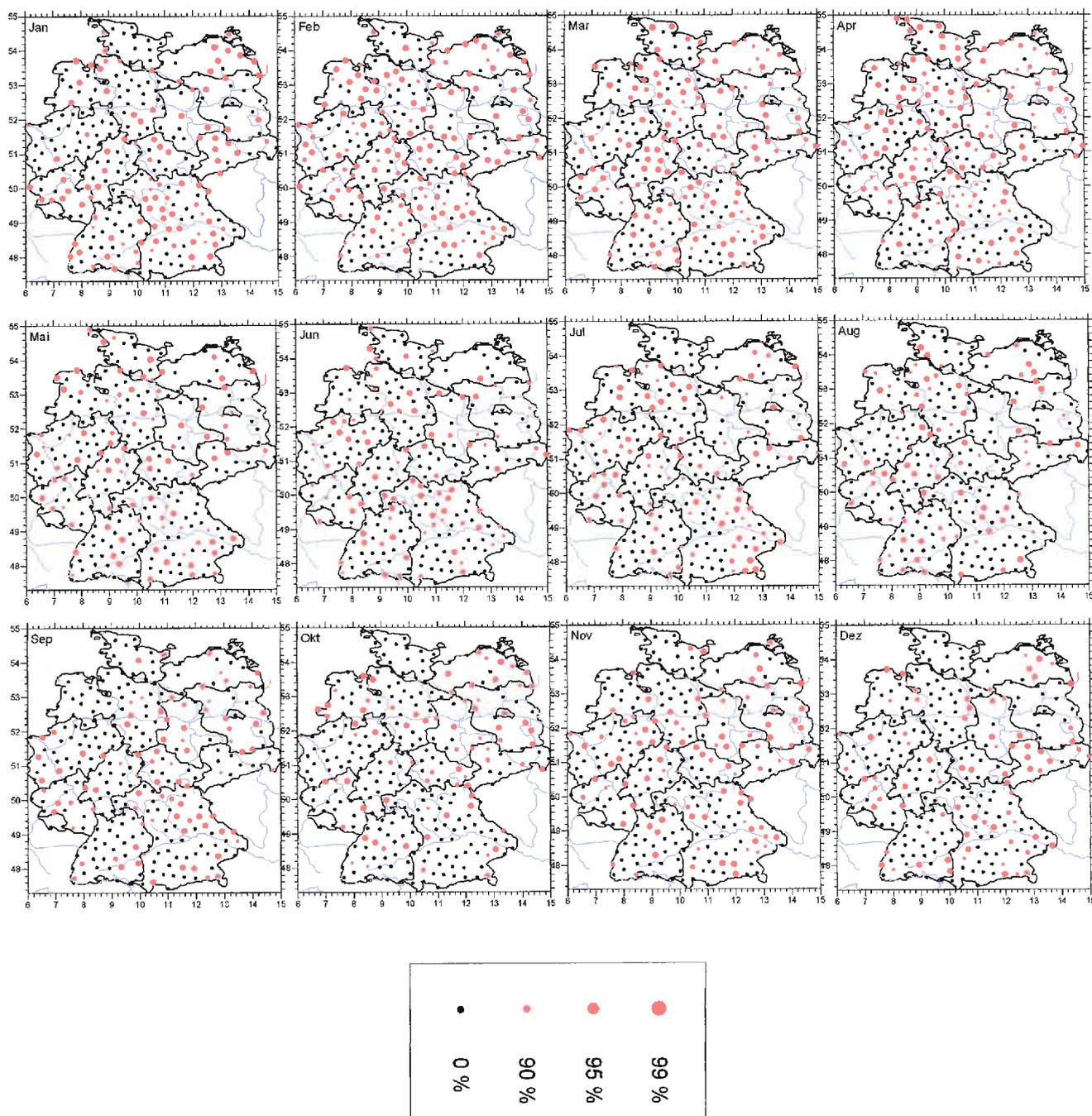


Abbildung 5.16: Signifikanz nach Kolmogorov-Smirnov Test für die Anpassung der Weibullverteilung jeweils monatlich im Zeitraum 1951-2000. Die beste Anpassung (MOM oder MLE) für eine Station wurde ausgewählt.

den Monaten Februar bis April gelingt für die Mehrheit der Stationen ein Anpassung mit mindestens 90% Signifikanz. Jedoch für Oktober und Dezember ist die theoretische Verteilung für weniger als ein Drittel der Stationen signifikant. Insgesamt ist die Anpassung der Weibullverteilung an die täglichen Niederschlagsdaten problematisch und daher kritisch zu beurteilen.

Wie schon in Kapitel 5.1.1 in Abbildung 5.3 für die Anzahl der Regentage n_1 bei den Markov Ketten mit zwei Zuständen erfolgt, so soll auch hier bei der Analyse mit vier Zuständen auf die Trends der drei Zustände eingegangen werden, also die Anzahl der Tage mit wenig, mittel und viel Niederschlag. Die Anpassung der Weibullverteilung erfolgte stationsweise an die Daten aller gleicher Monate über den Gesamtzeitraum. Die Zustände wurden aufgrund dieser Anpassung über die vorgegebenen Quantile eingeteilt. Nun kann man für jede Jahreszeit die Anzahl des jeweiligen Zustands auszählen und diese in eine neue Datenreihe zusammensetzen. Mit dem Mann-Kendall Test wurde schließlich versucht, eine Aussage über eine signifikante Zu- oder Abnahme innerhalb der Dateihe zu gewinnen. Das Ergebnis ist in Abbildung 5.17 dargestellt und wird im Folgenden kurz zusammengefasst:

1. Die Trends für Zustand 1 (wenig Niederschlag) sind in allen Jahreszeiten unterschiedlich. Es treten positive wie negative signifikante und nicht signifikante Veränderung auf. Insgesamt sind sie schwer interpretierbar.
2. Für Zustand 2 (mittlerer Niederschlag) werden hauptsächlich keine signifikanten Trends detektiert. Hier erwartet man auch aus Gründen, die bei der Anpassung der Weibullverteilung zu suchen sind, eher wenig zeitliche Veränderungen.
3. Ein interessantes Bild ergibt sich bei Zustand 3 (hoher Niederschlag). So sind im Winter vor allem im nördlichen Nordrhein-Westfalen, in Niedersachsen und in Schleswig-Holstein positive Trends (zum Teil mit 99% Signifikanz) zu finden. Im Sommer sind viele Stationen im Westen und Nordosten Deutschlands mit negativen Trends vertreten. Hier nimmt also die Anzahl von Tagen mit extremem Niederschlag ab. Daraus kann aber nicht auf die Intensität geschlossen werden. Im östlichen Teil Deutschlands sind hingegen keine signifikanten Trends auszumachen. Im Herbst weisen einige Stationen in Südbayern und im Norden von Baden-Württemberg signifikant positive Trends auf. Auch im Frühling nimmt die Anzahl der Tage mit hohen Niederschlagswerten im Südwesten von Deutschland signifikant zu.

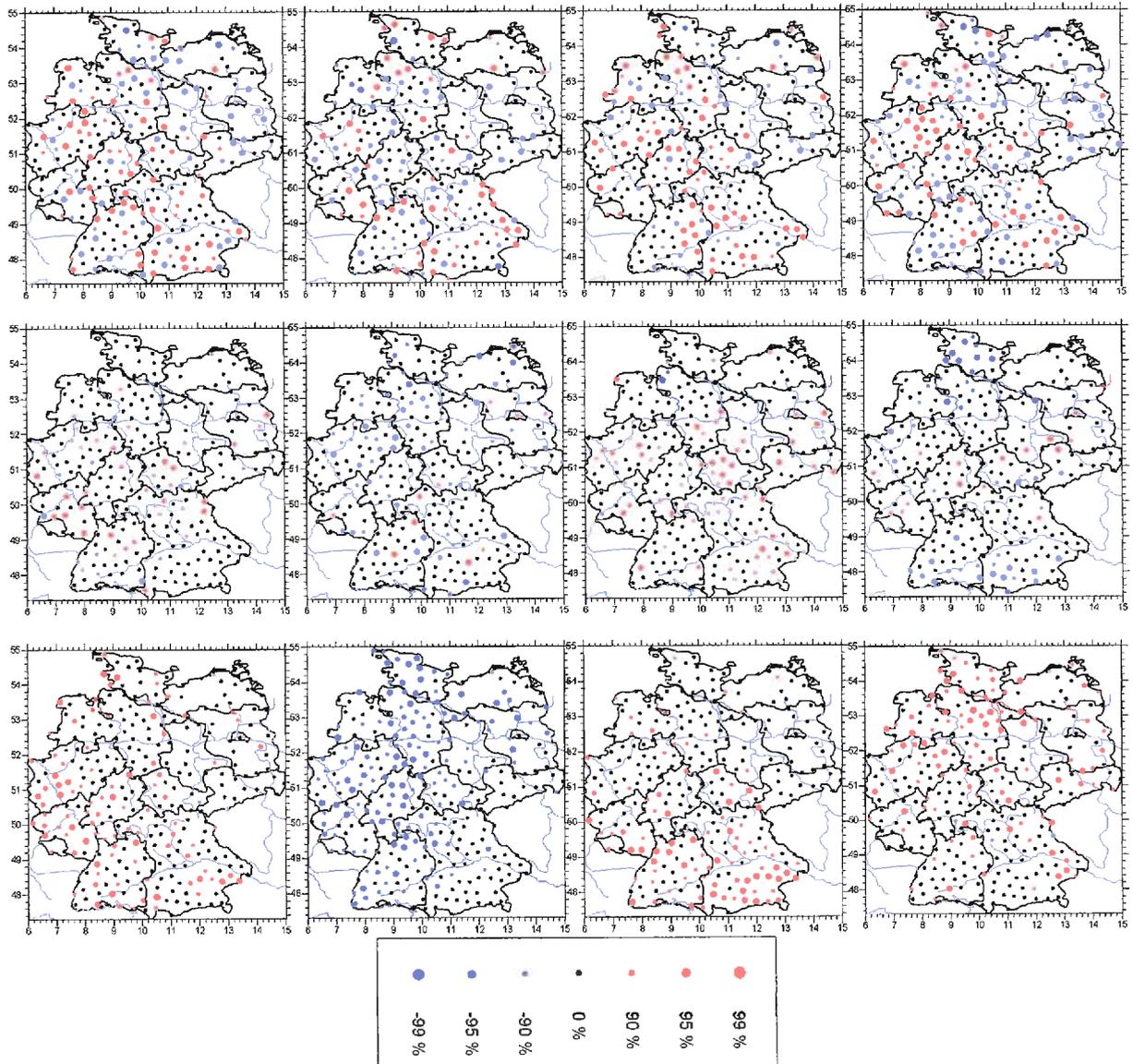


Abbildung 5.17: Trends der Anzahl der Zustände 1 (obere Reihe), 2 (mittlere Reihe) und 3 (untere Reihe) jeweils für Frühling, Sommer, Herbst und Winter (von links nach rechts) berechnet für 1951-2000

5.2.2 Detektierte Ordnungen der Markov Ketten

Hier werden die Ergebnisse aus der Berechnung der Ordnung nach dem BIC (Gleichung 2.19) für den Gesamtzeitraum und dessen beiden Hälften vorgestellt. Bei Betrachtung von Abbildung 5.18, die die detektierte Ordnung für 1951-2000 jahrezeitlich zeigt, erkennt man, dass im Frühling, Sommer und Herbst jeweils die Ordnung 2 ausgewählt wird. Im Winter jedoch wird mit Ausnahme einiger, weniger Stationen im Süden von Baden-Württemberg die dritte Ordnung als am besten geeignet eingestuft, die entsprechenden

Stationsreihen zu modellieren. Innerhalb einer Jahreszeit zeigen sich ansonsten aber keine regionalen Unterschiede.

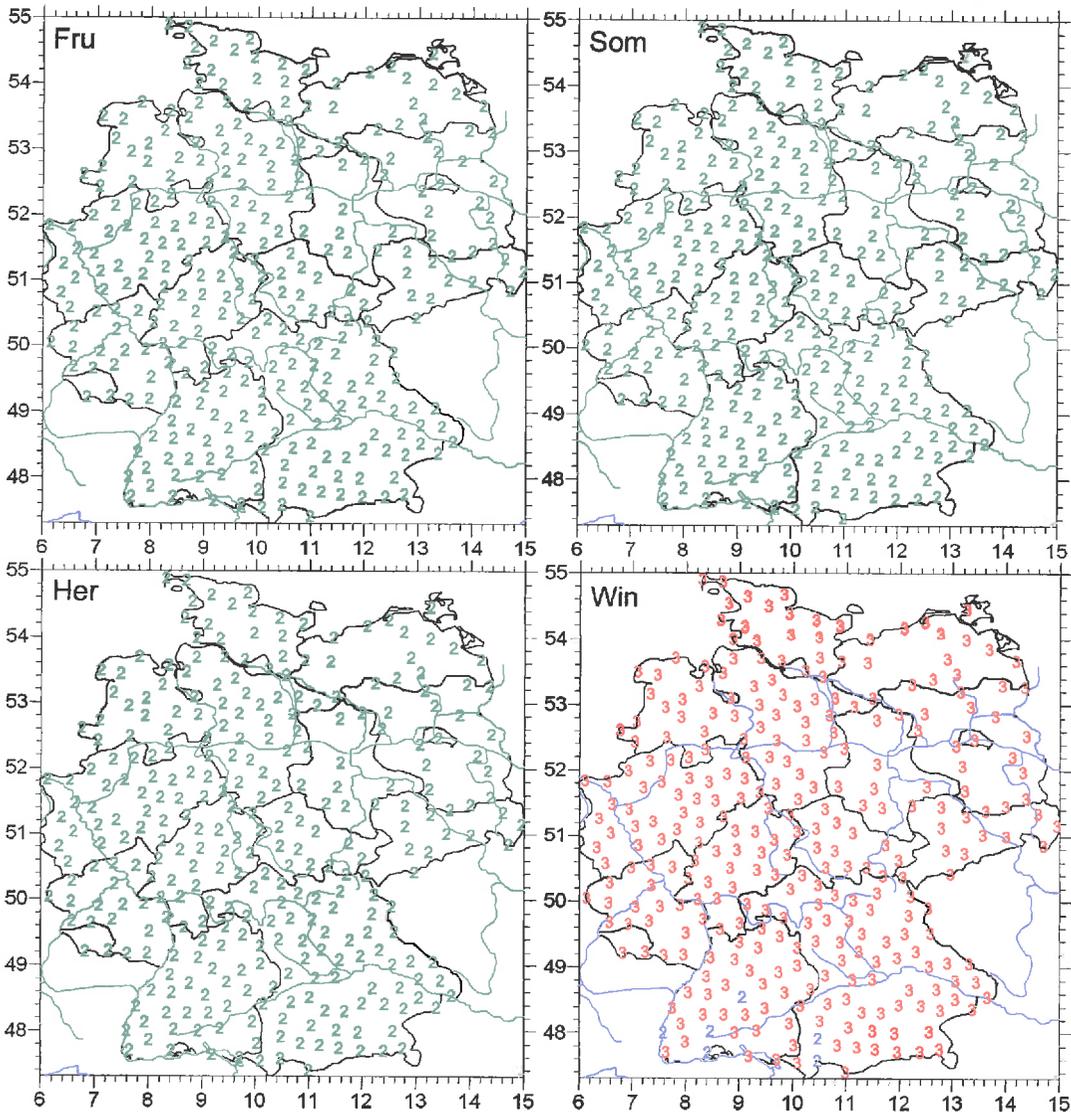


Abbildung 5.18: Ordnung der Markov Ketten nach BIC mit vier Zuständen (1951-2000) für Frühling, Sommer, Herbst und Winter

Betrachtet man nun die Ergebnisse, die sich aus Teilung des gesamten Beobachtungszeitraums ergeben (Abb. 5.19 und 5.20), so werden deutliche jahreszeitliche Unterschiede registriert.

Im Frühling weisen fast alle Stationen im süddeutschen Raum zwischen 1951 und 1975 die Ordnung 1 auf. Für Stationen, die im Westen und Nordenwesten liegen, zeigt sich ein gemischtes Bild. Im Westen sind mehr Stationen mit der Ordnung 1 vertreten als

im Osten. Das kann mit der größeren Entfernung zur Nordsee zusammenhängen und ein Hinweis auf ein anderes, regionales Klima sein. Im Zeitraum 1976 bis 2000 wird bis auf wenige Ausnahmen die Ordnung 2 detektiert.

Im Sommer sieht man einen diagonal durch Deutschland verlaufenden Unterschied zwischen Nord-West und Süd-Ost. Treten in der ersten Hälfte des Gesamtzeitraums vorwiegend Stationen im Nordwesten mit der Ordnung 1 und im Südosten mit der Ordnung 2 auf, so ist es in der zweiten Hälfte umgekehrt. Hier sind alle Stationen im Nordwesten mit der Ordnung 2 vertreten.

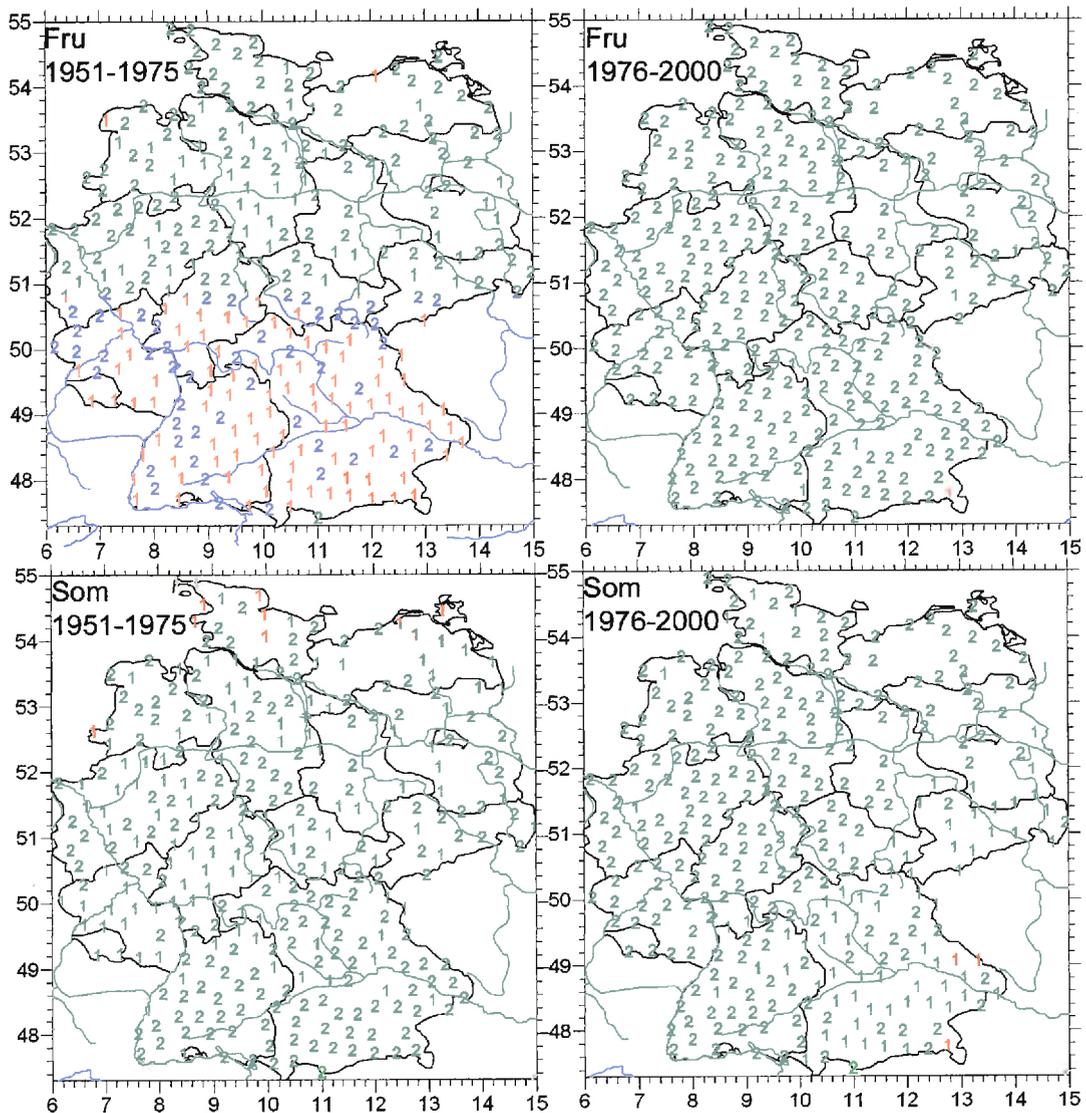


Abbildung 5.19: Ordnung der Markov Ketten mit vier Zuständen nach dem BIC: Vergleich zwischen 1951-1975 und 1976-2000 für den Frühling (oben) und den Sommer (unten)

Im Herbst (1951-1975) sind in Mitteldeutschland einige Stationen, die die Ordnung 1 aufweisen, zu sehen. Zwischen 1976 und 2000 sind im Süden Deutschlands vorwiegend Markov Ketten 1. Ordnung vertreten im Gegensatz zum nördlich Teil, wo viele Stationen die Ordnung 2 aufweisen.

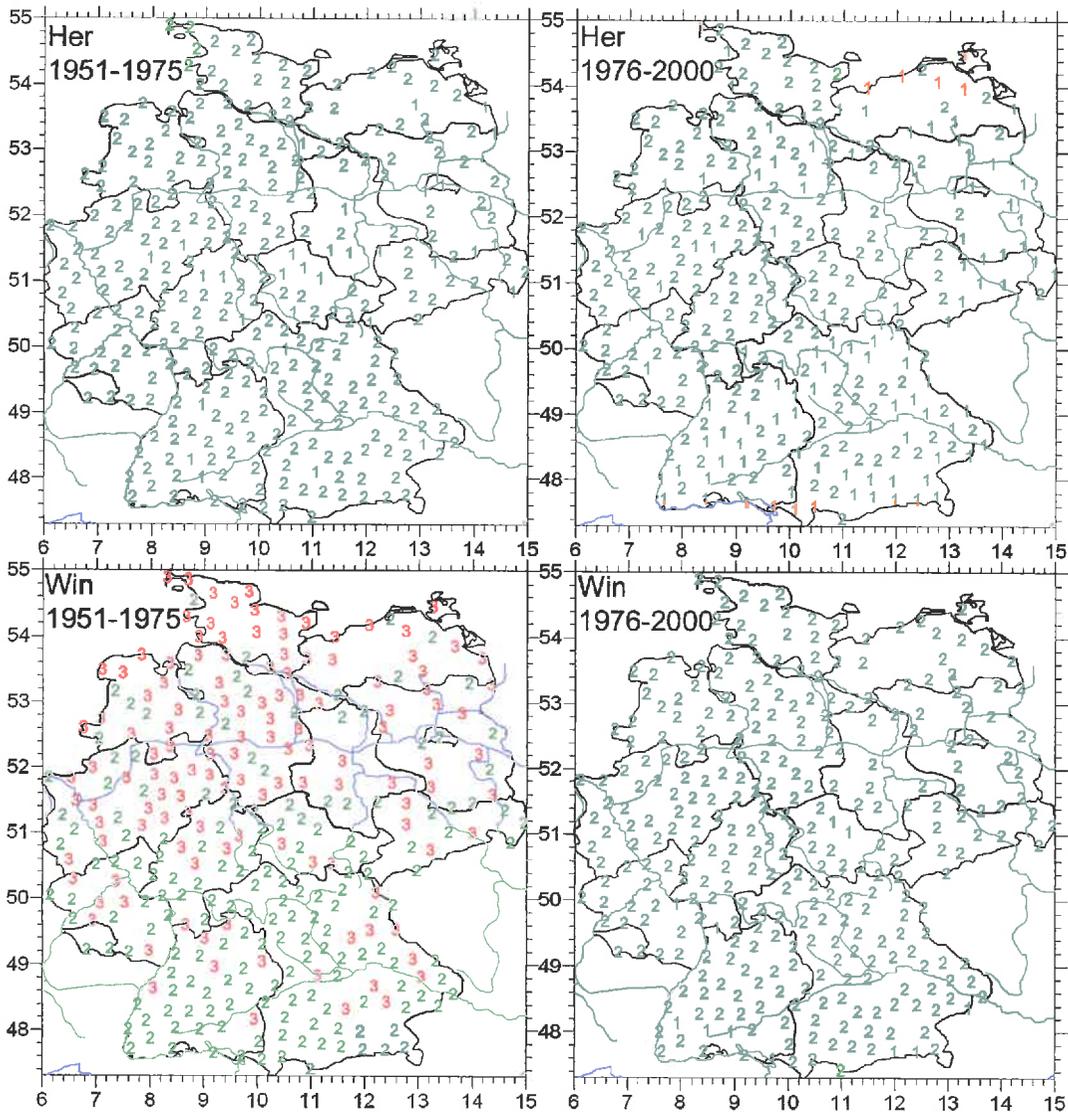


Abbildung 5.20: Ordnung der Markov Ketten mit vier Zuständen nach dem BIC: Vergleich zwischen 1951-1975 und 1976-2000 für den Herbst (oben) und den Winter (unten)

Im Winter erkennt man, dass die überwiegende Mehrheit der Stationen mit Ausnahme von denen in Baden-Württemberg, Bayern und Thüringen die Ordnung 3 zeigen. Hier sieht man ein deutliches Nord-Süd Gefälle, wobei im Norden die Ordnung 3 und im Süden die Ordnung 2 dominiert. Jedoch zwischen 1977 und 2000 - der Winter von 1976 zählt nicht

mit - wird hauptsächlich die Ordnung 2 ausgewählt in einigen wenigen Fällen auch die Ordnung 1. Interessant ist die südhessische Station Hemsbach (vgl. auch Abb. 3.3). Hier findet einen Wechsel von der höchsten Ordnung zur ersten Ordnung statt.

Werden die Ordnung nach dem AIC berechnet, stellt man fest, dass diese insgesamt größer liegen. So wird im Gesamtzeitraum nicht nur im Winter (Abb. B.13) die Ordnung 3 detektiert sondern auch in den anderen Jahreszeiten, wobei in diesen - vor allem im Frühling - aber auch die Ordnung 2 vertreten ist. Vergleicht man darüber hinaus die Ordnungswahl der zwei Hälften des Gesamtzeitraums (Abb. B.14 und B.15) miteinander, so lassen sich außer im Winter keine Unterschiede ausmachen. Der Winter (1952-1975) tritt mit der Ordnung 3 in Erscheinung, in der zweiten Hälfte jedoch mit der Ordnung 2. Insgesamt wird im Gegensatz zum BIC nur sehr selten eine Markov Kette 1. Ordnung angepasst.

Des Weiteren ist auch untersucht worden, welche Ordnungen sich aus der Gesamtstatistik ergeben, wenn alle 301 Niederschlagsstationen gemeinsam betrachtet werden. Sowohl das AIC als auch das BIC finden für 1951-2000, 1951-1975 und für 1976-2000 für alle Jahreszeiten Markov Ketten dritter Ordnung.

5.2.3 Übergangswahrscheinlichkeiten der Markov Ketten

In diesem Kapitel werden die Übergangswahrscheinlichkeiten der Markov Ketten vorgestellt, die sich unter Verwendung von vier verschiedenen Regenintensitäten von trocken bis sehr feucht (vgl. Tab. 1) ergeben. Wie die Analyse der Ordnung im vorigen Kapitel (Abb. 5.18) ergeben hat, wird für den Winter hauptsächlich die Ordnung 3 und für die übrigen Jahreszeiten die Ordnung 2 ausgewählt jeweils für den Gesamtzeitraum 1951-2000. Daher soll hier auch auf Übergangswahrscheinlichkeiten höherer Ordnung eingegangen werden. Des Weiteren sind die Übergangswahrscheinlichkeiten, die die Übergänge von Trockentagen betreffen, naturgemäß identisch mit denen aus der Untersuchung mit nur zwei Zuständen und müssen hier nicht mehr wiederholt werden. Karten, die die geografische Verteilung der Eintrittswahrscheinlichkeiten von wenig feuchten, feuchten und sehr feuchten Tagen zeigen, sind im Anhang B.2.2 (Abb. B.16, B.17 und B.18) dargestellt.

Aufgrund der großen Anzahl der Übergangswahrscheinlichkeiten für Markov Ketten aller Ordnungen mit vier Zuständen kann nur eine kleine Auswahl präsentiert werden. Daher soll zunächst versucht werden, einen qualitativen Einblick zu geben, wie die Verteilung der Zustände aussehen kann. Solch einen Einblick gibt Tabelle 10, in der - wie schon in

Tabelle 2 für Frankfurt - die Niederschlagswerte zusammen mit dem ordinalen Merkmal der Station Achberg aus dem Mai 2000 aufgelistet sind.

Tabelle 10: Tabelliert ist der tägliche Niederschlagsverlauf in [mm] in Achberg mit dem dazugehörigen Zustand i_t für den jeweiligen Tag t im Mai 2000.

Tag, t	1	2	3	4	5	6	7	8	9	10
Niederschlag [mm], X_t	0	0.3	0.2	0.1	0	0	0	0.6	0	0
Zustand, i_t	0	1	1	1	0	0	0	1	0	0
Tag, t	11	12	13	14	15	16	17	18	19	20
Niederschlag [mm], X_t	8.2	17.3	0.2	0	0	0	15.5	12.6	2.6	0.2
Zustand, i_t	1	3	1	0	0	0	3	3	2	1
Tag, t	21	22	23	24	25	26	27	28	29	30 / 31
Niederschlag [mm], X_t	12.3	0	0	0	1.7	0	0.7	8.4	0.9	27.8 / 14.7
Zustand, i_t	3	0	0	0	1	0	1	2	1	3 / 3

Der Übergang von einem Trockentag zu einem sehr feuchten Tag findet in diesem Beispiel nur ein einziges Mal statt: nämlich vom 16. auf den 17. Mai. Das hat zur Folge, dass die Wahrscheinlichkeit für den Übergang „03“ sehr gering ist im Vergleich zu der des Übergangs „01“, der insgesamt fünfmal eintritt. Die Kombination „02“ tritt überhaupt nicht ein. Meistens ist es so, dass es erst wenig regnet und dann viel regnet. Dann gehen entweder die Niederschlagswerte langsam zurück oder die Regentage werden abrupt beendet. Das bedeutet, dass Kombinationen wie „0231“, „2320“ und „3000“ etc. in ihrer Anzahl stark vertreten sind und daher die entsprechenden Übergangswahrscheinlichkeiten im Vergleich zu den anderen drei möglichen Realisationen - beispielsweise muss „0231“ mit „0230“, „0232“ und „0233“ verglichen werden - eher groß sind. Des Weiteren hängen die Übergangswahrscheinlichkeiten verschiedener Ordnungen miteinander zusammen. Wenn beispielsweise wie oben am Beispiel Achberg die Kombination „02“ nicht realisiert ist, gilt das Gleiche auch für „002“ etc. . Wird diese Kombination jedoch realisiert, fällt der Vergleich zwischen Wahrscheinlichkeiten verschiedener Ordnungen jedoch deutlich schwerer aus.

Bevor auf Übergangswahrscheinlichkeiten 3. und 2. Ordnung eingegangen wird, soll die 1. Ordnung diskutiert werden und zwar mit Augenmerk auf den dritten, niederschlagreichsten Zustand. Darum sind in den Abbildungen 5.21, 5.22 und 5.23 die Wahrscheinlichkeiten dargestellt, die die Übergänge von „0“ nach „3“, „3“ nach „0“ und von „3“ nach „3“ beschreiben. Die letzten beiden können direkt miteinander verglichen werden.

In den folgenden Karten wird immer Sommer mit Winter verglichen, da zwischen diesen Jahreszeiten die Unterschiede am besten erkennbar sind.

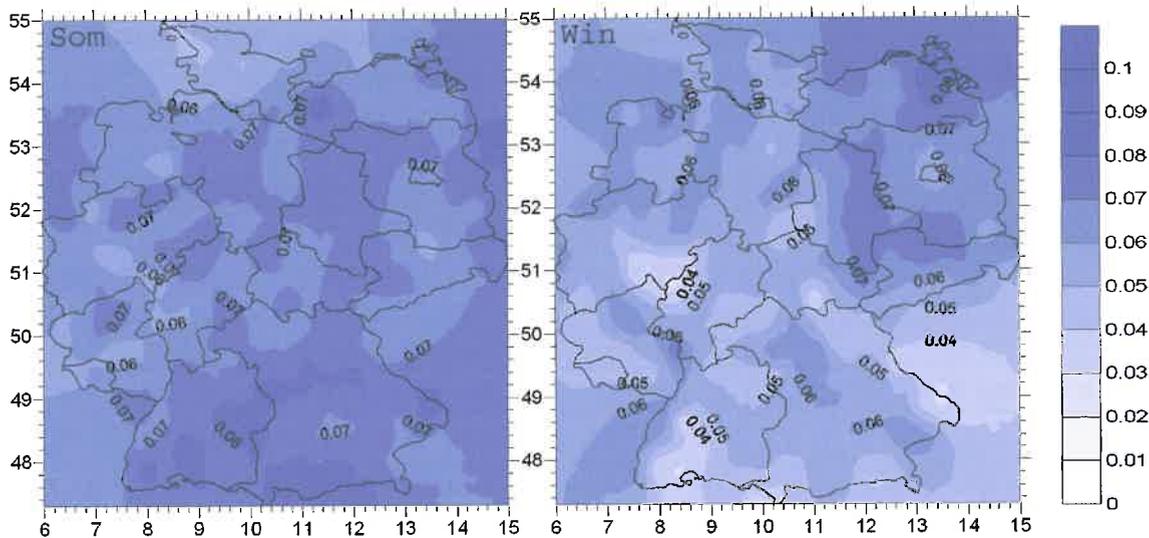


Abbildung 5.21: Wahrscheinlichkeit p_{03} für 1951-2000 im Sommer (links) und Winter (rechts)

Die Übergangswahrscheinlichkeit p_{03} (Abb. 5.21) ist insgesamt mit Werten unter 10% nur schwach ausgeprägt, ist aber im Sommer in den meisten Bereichen größer als im Winter.

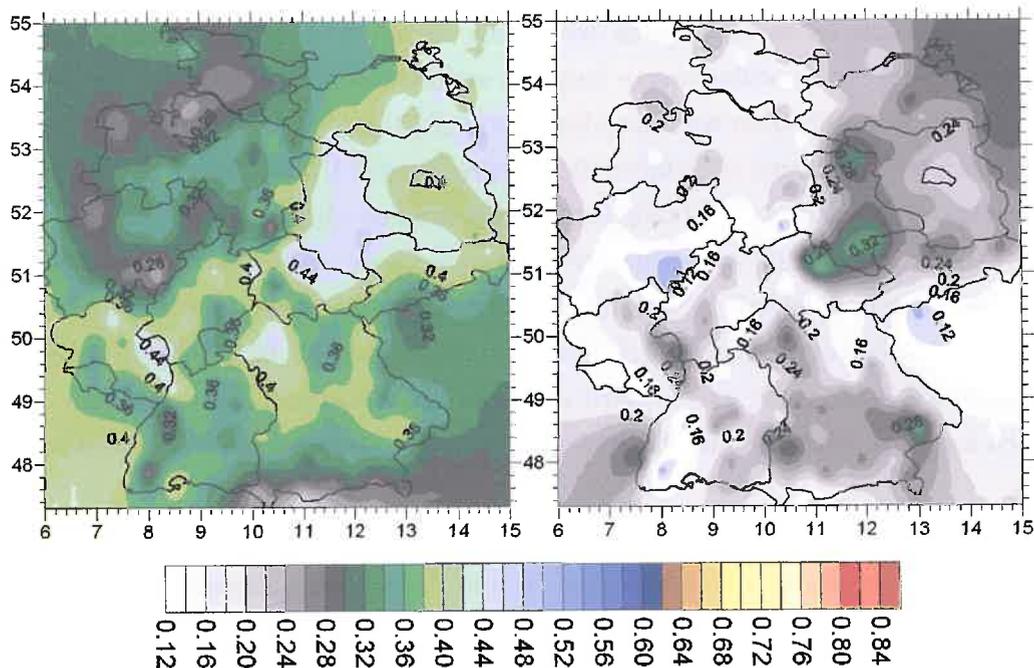


Abbildung 5.22: Wahrscheinlichkeit p_{30} für 1951-2000 im Sommer (links) und Winter (rechts)

Hier dominiert ganz klar p_{00} (Abb. 5.9). Interessant ist die Tatsache, dass Bergstationen wie z. B. am Feldberg (Schwarzwald) im Vergleich zur näheren Umgebung im Sommer maximale Werte aufweisen und im Winter minimale Werte. Vergleicht man p_{30} mit p_{33} (Abb. 5.22 und 5.23) für Sommer und Winter, so stellt man fest, dass im Sommer p_{30} dominant ist und im Winter p_{33} . Das spricht für kurze konvektive Niederschläge im Sommer und eher lange anhaltender Niederschlag im Winter. Kleine Werte um 28% für p_{30} findet man vor allem an der Nordseeküste, im Alpengebiet und nordwestlich von Hessen. Ebenso halten sich beide Übergangswahrscheinlichkeiten in diesen Gebieten die Waage, d. h. es ist in etwa gleich wahrscheinlich, ob nach einem sehr feuchten Tag ein Trockentag oder wieder ein sehr feuchter Tag folgt.

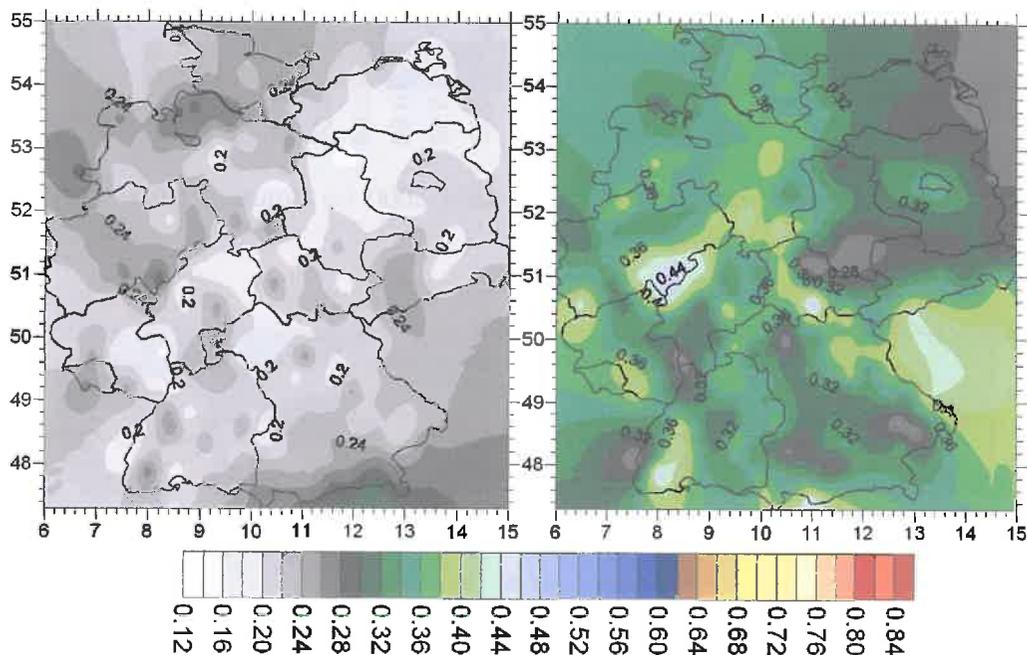


Abbildung 5.23: Wahrscheinlichkeit p_{33} für 1951-2000 im Sommer (links) und Winter (rechts)

In den neuen Bundesländern ist die Kombination „30“ im Sommer mindestens doppelt so wahrscheinlich als „33“. Im Winter ist p_{33} bei Werten knapp über 30% um etwa 8 bis 10% größer als p_{30} .

Bei den Übergangswahrscheinlichkeiten 2. und 3. Ordnung sind in den aller meisten Fällen immer diejenigen am größten, die eine Serie von Trockentagen enthalten wie z. B. p_{100} und p_{200} oder aber Serien von gleichen Zuständen wie z. B. p_{222} und p_{333} . Eine neben diesen genannten größte Übergangswahrscheinlichkeit 2. Ordnung ist p_{010} . Diese und p_{013} sind für Sommer und Winter in den Abbildungen 5.24 und 5.25 zu sehen.

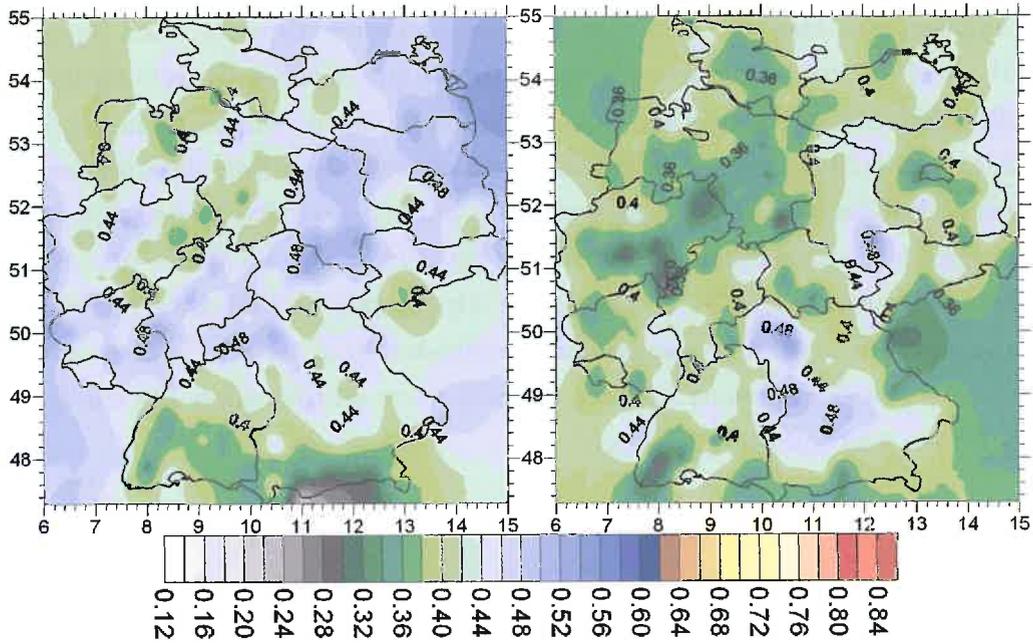


Abbildung 5.24: Wahrscheinlichkeit p_{010} für 1951-2000 im Sommer (links) und Winter (rechts)

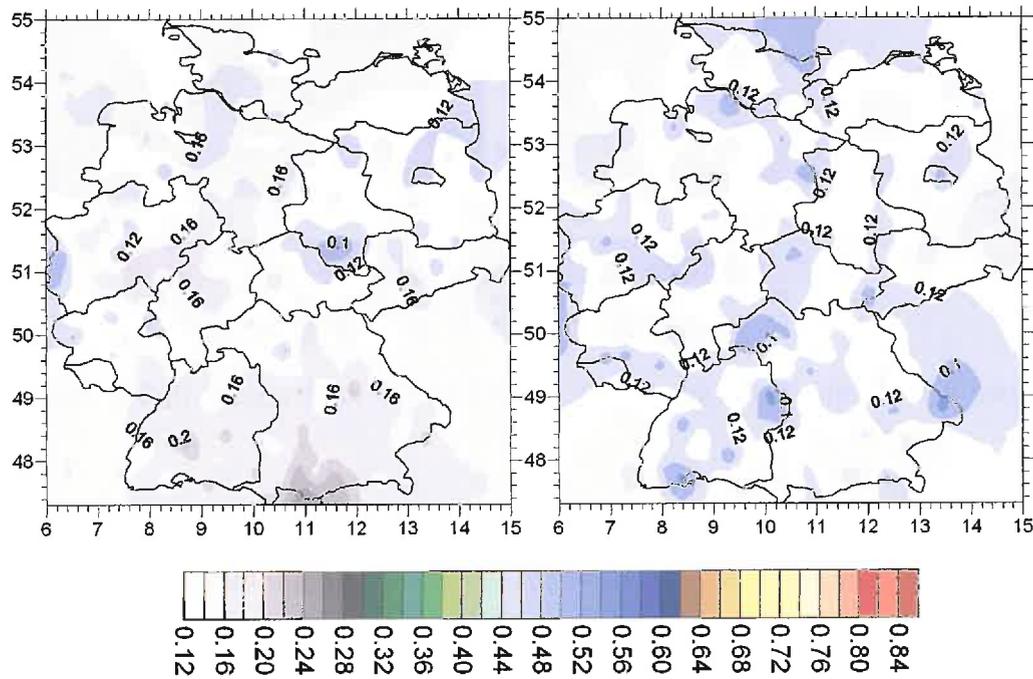


Abbildung 5.25: Wahrscheinlichkeit p_{013} für 1951-2000 im Sommer (links) und Winter (rechts)

Beide Wahrscheinlichkeiten stehen in Konkurrenz zu p_{011} , die in diesem vierer Tupel die größte ist. In der geografischen Verteilung von p_{010} (Abb. 5.24) wird wieder die besondere

Rolle der Bereiche nahe zur Nordsee und den Alpen sowie des Schwarzwalds deutlich, die sich deutlich von den restlichen Gebieten abheben. So sind Sommers wie Winters dort die niedrigsten Werte zu finden, die im Winter aber schwächer ausgeprägt sind. Insgesamt ist wieder der Niederschlagscharakter mit eher mehr einzelnen Regentagen im Sommer als im Winter erkennbar. Der Übergang von „01“ nach „3“ spielt indes eine untergeordnete Rolle (Abb. 5.25). Diese Wahrscheinlichkeit ist im Sommer im Vergleich mit Winter am größten und dies vor allem in Westdeutschland.

Als letztes Beispiel für eine Übergangswahrscheinlichkeit soll p_{3330} und p_{3333} (Abb. 5.26 und 5.27) gezeigt werden; also die Wahrscheinlichkeit, dass auf drei sehr feuchte Tage wieder ein sehr feuchter Tag bzw. ein Trockentag folgt. Die Serie von 4 feuchten Tagen ist im Sommer im Großen und Ganzen zwei- bis viermal (bsp. in Sachsen-Anhalt) so wahrscheinlich als im Winter. Das ist ein Indiz für den vergleichsweise hohen Sommerniederschlag (vgl. Abb. 3.2). Von der anderen Seite aus betrachtet, muss die Übergangswahrscheinlichkeit, dass die Serie von feuchten Tagen nach dem dritten abgebrochen wird (Abb. 5.26), im Winter höher sein als im Sommer, was auch der Fall ist, jedoch mit geringer ausgeprägten Unterschied zwischen Sommer und Winter als für p_{3333} .

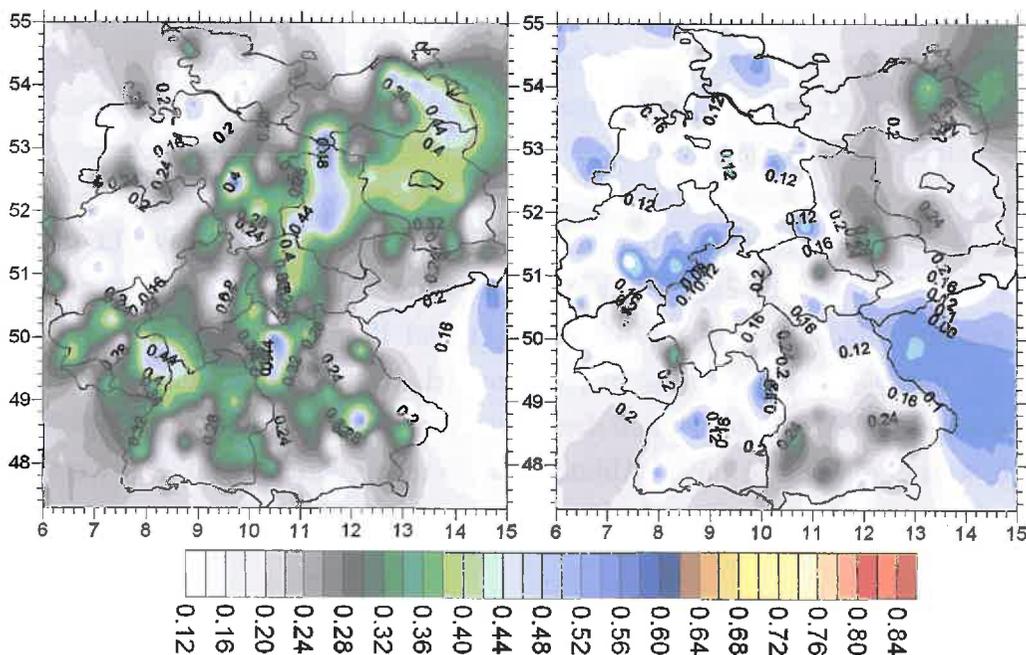


Abbildung 5.26: Wahrscheinlichkeit p_{3330} für 1951-2000 im Sommer (links) und Winter (rechts)

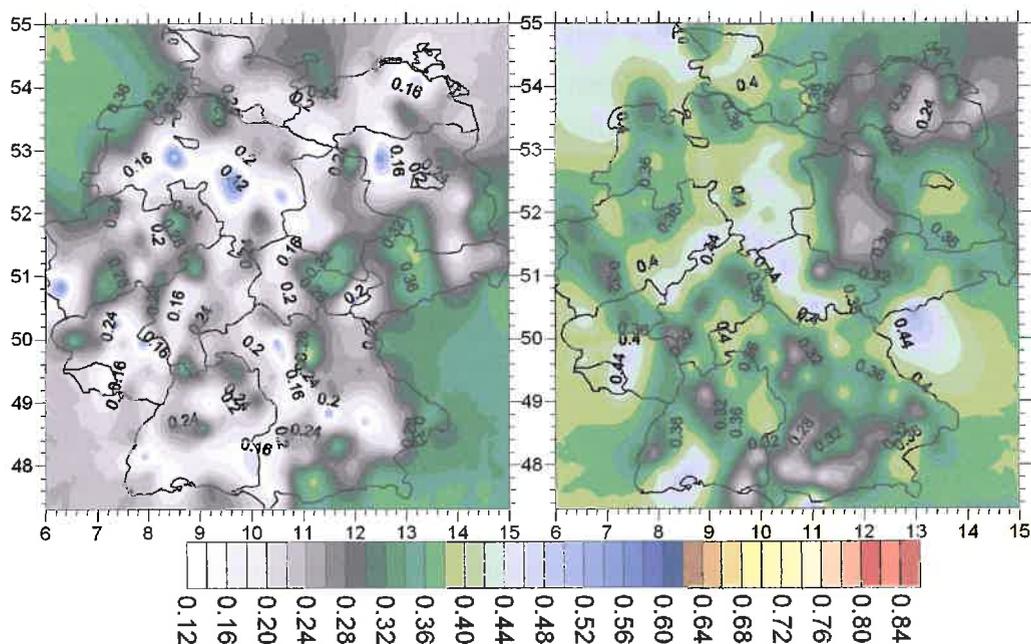


Abbildung 5.27: Wahrscheinlichkeit p_{3333} für 1951-2000 im Sommer (links) und Winter (rechts)

Jahresgang der Übergangswahrscheinlichkeiten

Auch bei der Analyse mit vier Zuständen soll anhand einiger Übergangswahrscheinlichkeiten der Jahresgang erläutert werden. Als Beispiel werden die Eintrittswahrscheinlichkeiten p_1 , p_2 und p_3 für einen wenig feuchten, einen feuchten und einen sehr feuchten Tag gezeigt. Darüber hinaus werden auch die Wahrscheinlichkeiten präsentiert, die das Eintreten eines Trockentages nach einem wenig feuchten, feuchten und sehr feuchten Tag beschreiben sowie eines Trockentages nach einem sehr feuchten Tag; also p_{10} , p_{20} , p_{30} und p_{33} . Dazu ist in Abbildung 5.28 der Verlauf dieser Wahrscheinlichkeiten für die Station Bremen, Frankfurt und die Zugspitze sowie wieder für die Gesamtstatistik aller Stationen in monatlicher Auflösung gezeigt. In Tabelle 11 und 12 sind diese Wahrscheinlichkeiten zusammen mit p_{03} , p_{003} und p_{0003} in jahreszeitlicher Auflösung zusammengestellt. Bei Betrachtung der Werte dieser letzten drei genannten Wahrscheinlichkeiten wird auch ersichtlich, warum diese nicht mit in Abbildung 5.28 untergebracht sind. Die Werte sind zu klein und weisen nur einen schwach ausgeprägten Jahresgang auf.

Betrachtet man den Jahresgang der sich aus allen 301 Stationen ergibt, so stellt man fest, dass der Verlauf der Übergangswahrscheinlichkeiten sehr glatt ist und damit kaum repräsentativ für andere Stationen ist. Die Eintrittswahrscheinlichkeiten der Grundzustände sind auf etwa dem selben Niveau. Diese sind im Sommer aber niedriger als im Winter. Die

Wahrscheinlichkeit p_{30} zeigt im Sommer ein Maximum bei etwa 35%. Stärkste Fluktuationen zeigt der Verlauf an der Zugspitze. In Frankfurt ist es im Sommer wahrscheinlicher (über 40%) als im Winter, dass ein Trockentag auf einen Tag mit sehr feuchtem Niederschlag folgt. Auf der Zugspitze hat diese Wahrscheinlichkeit im Herbst ihr Maximum. Des Weiteren ist auf der Zugspitze p_{33} im ganzen Jahr höher als in Bremen und Frankfurt.

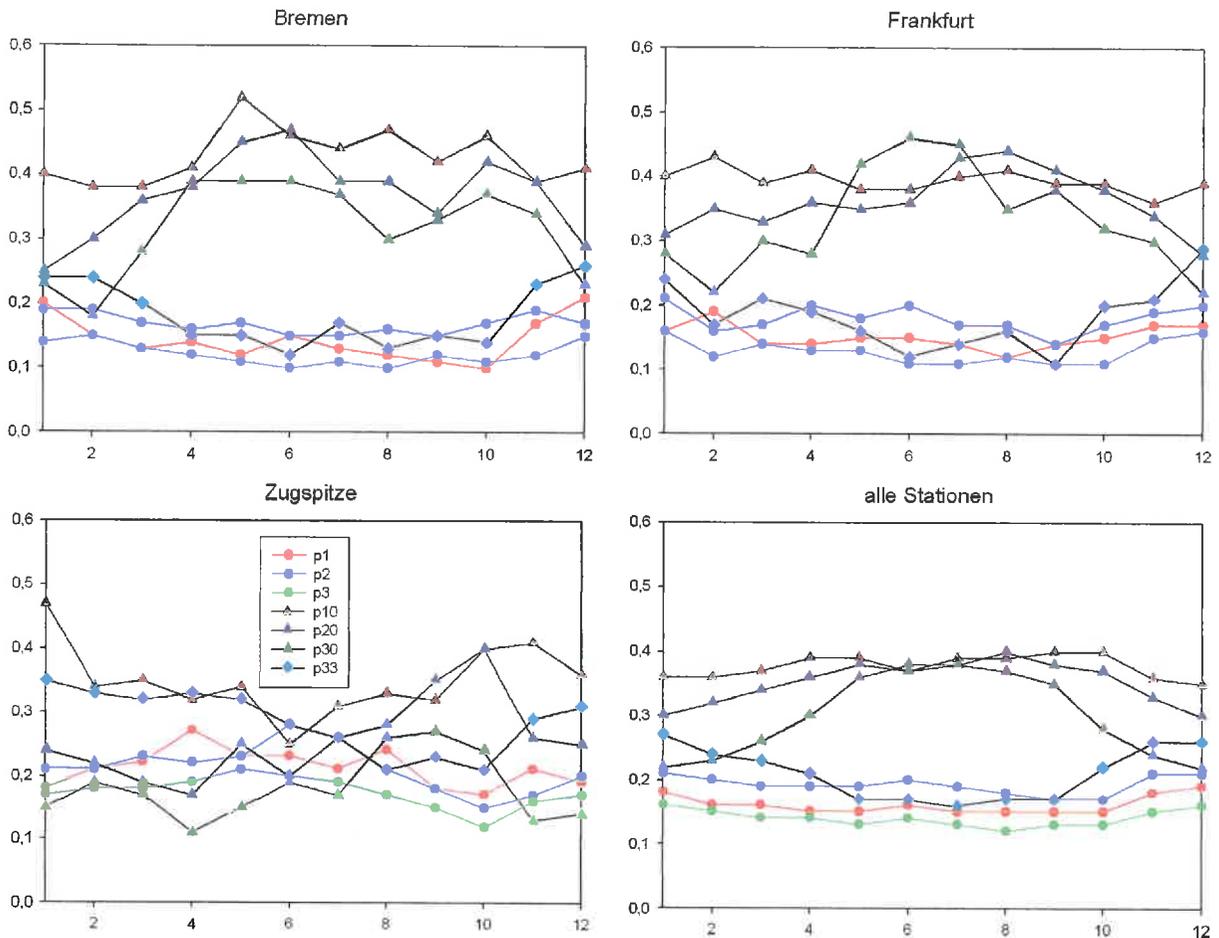


Abbildung 5.28: mittlerer Jahrgang von Bremen, Frankfurt, Zugspitze und mittlerer Jahrgang aller Stationen für ausgewählte Übergangswahrscheinlichkeiten für vier Zustände berechnet für 1951-2000

Die Wahrscheinlichkeiten, dass ein sehr feuchter Tag nach einem Trockentag folgt, sind nur schwach ausgeprägt (um 10%), haben aber die Tendenz, im Sommer höher zu sein als im Winter (Tab. 11). Des Weiteren sind die Wahrscheinlichkeiten, dass nach einem sehr feuchtem Tag eine Serie von Trockentagen folgt, im Sommer größer als im Winter. Umgekehrt betrachtet, ist die Wahrscheinlichkeit einer Serie von sehr feuchten Tagen im Winter wahrscheinlicher als im Sommer.

Tabelle 11: Aufgelistet ist die Gesamtstatistik der einiger Übergangswahrscheinlichkeiten für alle 301 Niederschlagsstationen nach den Jahreszeiten getrennt (1951-2000)

Wahrscheinlichkeit	Frühling	Sommer	Herbst	Winter
p_1	0.15	0.15	0.16	0.18
p_2	0.14	0.13	0.13	0.15
p_3	0.19	0.18	0.19	0.21
p_{03}	0.09	0.11	0.08	0.09
p_{30}	0.30	0.35	0.27	0.21
p_{33}	0.37	0.30	0.38	0.43
p_{003}	0.09	0.11	0.08	0.09
p_{300}	0.61	0.62	0.59	0.21
p_{333}	0.38	0.33	0.42	0.43
p_{0003}	0.09	0.11	0.08	0.08
p_{3000}	0.69	0.67	0.65	0.65
p_{3333}	0.40	0.35	0.43	0.48

In Tabelle 12 sind die Übergangswahrscheinlichkeiten der sechs Beispielstationen in jahreszeitlicher Auflösung gezeigt. Hier zeigt sich ein ähnliches Bild wie auch schon zuvor beschrieben. Im Sommer ist ein sehr feuchter Tag nach einer Serie von Trockentagen wahrscheinlicher als im Frühling, Herbst und Winter. Eine Serie von sehr feuchten Tagen ist im Winter (um 48%) aber wahrscheinlicher als im Sommer (35%).

Tabelle 12: Übersicht einiger Übergangswahrscheinlichkeiten, die mindestens einen sehr feuchten Tag enthalten; von ausgewählten Stationen für 1951-2000 mit Höhenangabe bezogen auf Meeresspiegelniveau (NN)

Frühling									
Station	NN [m]	p_1	p_2	p_3	p_{03}	p_{30}	p_{33}	p_{003}	p_{0003}
Sankt Peter-Ording	4	0.15	0.17	0.13	0.06	0.34	0.22	0.064	0.059
Bremen	2	0.16	0.18	0.16	0.06	0.26	0.3	0.073	0.071
Potsdam	81	0.13	0.16	0.12	0.06	0.32	0.26	0.05	0.051
Frankfurt	113	0.14	0.18	0.13	0.06	0.33	0.26	0.07	0.066
Feldberg/Schwarzwald	1486	0.22	0.24	0.18	0.05	0.14	0.36	0.057	0.054
Zugspitze	2960	0.24	0.22	0.19	0.07	0.13	0.4	0.059	0.059
Sommer									
Station	NN [m]	p_1	p_2	p_3	p_{03}	p_{30}	p_{33}	p_{003}	p_{0003}
Sankt Peter-Ording	4	0.16	0.18	0.12	0.05	0.31	0.22	0.058	0.049
Bremen	2	0.17	0.21	0.15	0.07	0.28	0.26	0.091	0.086
Potsdam	81	0.13	0.15	0.10	0.06	0.40	0.21	0.064	0.059
Frankfurt	113	0.13	0.18	0.11	0.06	0.39	0.21	0.075	0.066
Feldberg/Schwarzwald	1486	0.21	0.20	0.17	0.08	0.25	0.28	0.097	0.093
Zugspitze	2960	0.22	0.25	0.18	0.09	0.19	0.28	0.11	0.12
Herbst									
Station	NN [m]	p_1	p_2	p_3	p_{03}	p_{30}	p_{33}	p_{003}	p_{0003}
Sankt Peter-Ording	4	0.21	0.20	0.17	0.04	0.15	0.38	0.057	0.058
Bremen	2	0.19	0.21	0.15	0.05	0.18	0.31	0.052	0.036
Potsdam	81	0.13	0.16	0.12	0.06	0.32	0.25	0.056	0.048
Frankfurt	113	0.15	0.16	0.12	0.06	0.31	0.23	0.053	0.039
Feldberg/Schwarzwald	1486	0.19	0.19	0.16	0.04	0.17	0.38	0.058	0.044
Zugspitze	2960	0.18	0.17	0.14	0.04	0.20	0.34	0.044	0.040
Winter									
Station	NN [m]	p_1	p_2	p_3	p_{03}	p_{30}	p_{33}	p_{003}	p_{0003}
Sankt Peter-Ording	4	0.21	0.21	0.18	0.06	0.18	0.35	0.071	0.067
Bremen	2	0.20	0.22	0.18	0.06	0.16	0.35	0.072	0.067
Potsdam	81	0.18	0.18	0.15	0.06	0.22	0.34	0.054	0.046
Frankfurt	113	0.17	0.19	0.14	0.05	0.21	0.31	0.04	0.045
Feldberg/Schwarzwald	1486	0.21	0.23	0.18	0.03	0.08	0.45	0.029	0.024
Zugspitze	2960	0.20	0.20	0.18	0.05	0.10	0.11	0.033	0.033

5.2.4 Zeitliche Veränderungen der Übergangswahrscheinlichkeiten

Beim Übergang von der Betrachtung von zwei Zuständen zu vier Zuständen ist die Trendanalyse generell als problematisch einzustufen, denn die Übergangswahrscheinlichkeiten 2. und 3. Ordnung treten in ihrer Gesamtheit - 64 Wahrscheinlichkeiten für die zweite und 256 Wahrscheinlichkeiten für die dritte Ordnung - pro Jahr nur sehr selten oder überhaupt nicht auf. Ausnahmen bilden diejenigen Wahrscheinlichkeiten, die drei Trockentage in Serie enthalten wie z. B. p_{0003} . Wenn nun der lineare Trend angewendet wird, erwartet man nur sehr kleine Steigungen, die in den meisten Fällen statistisch signifikant identisch mit Null sein sollten. Hinzu kommt die Tatsache, dass - wenn man Ordnungen pro Jahr berechnet und nicht wie geschehen über den Gesamtzeitraum und dessen Hälften - , die zweite und dritte Ordnung nicht zugelassen werden können. Denn es sind keine ergodischen Markov Ketten, die eine Grundvoraussetzung für meteorologische Anwendungen sind (vgl. Kap. 2.5 und [11]). Aus diesem Grund werden nur die Trends der 0. und 1. Ordnung vorgestellt.

In Tabelle 13 sind die Übergangswahrscheinlichkeiten für die Gesamtstatistik aller 301 Stationen aufgelistet. In Abbildung 5.29 sind die entsprechenden Wahrscheinlichkeiten

Tabelle 13: lineare Wahrscheinlichkeitstrends in Prozent für einige Übergangswahrscheinlichkeiten getrennt nach den Jahreszeiten; die Zahl in Klammern gibt die Signifikanz der Steigung der Regressionsgerade an (vgl. Text)

Wahrscheinlichkeit	Frühling	Sommer	Herbst	Winter
p_1	+0.6 0	-0.3 0	+1.8 99	-0.3 0
p_2	+0.5 0	-0.1 0	+1.3 0	-0.7 0
p_3	+2.4 0	-2.6 90	+2.9 90	+3.6 90
p_{01}	+0.4 0	-1.3 90	+2.1 99	+0.3 0
p_{02}	-0.1 0	-2.9 0	+2.1 99	-0.4 0
p_{03}	-2.4 0	-2.5 0	-2.7 0	+0.4 0
p_{13}	+1.8 0	-0.5 0	+1.8 0	+2.6 0
p_{23}	+1.5 0	+0.1 0	+2.7 0	+3.4 90
p_{30}	-7 90	+0.1 0	-4.3 95	-4.6 95
p_{31}	+0.7 0	+1.3 0	+1.2 0	-1 90
p_{32}	+1.5 0	+1 0	-0.8 0	-0.5 0
p_{33}	+4.8 95	-2.3 0	+3.8 0	+6 95
p_{003}	-0.2 0	-3.4 0	-2.7 0	+0.7 0
p_{0003}	-0.3 0	-3.5 0	+0.6 0	+0.8 0

erster Ordnung veranschaulicht, dass nach einem sehr niederschlagsreichen Tag Tage verschiedener Niederschlagsintensität folgen. Im Frühling und Herbst haben die Eintrittswahrscheinlichkeiten eines wenig feuchten, eines feuchten und eines sehr feuchten Tages zugenommen. Im Winter hat nur die Wahrscheinlichkeit des niederschlagsreichsten Zustands zugenommen (+3.6%). Die Wahrscheinlichkeit, dass nach einem Trockentag ein sehr feuchter Tag folgt, hat mit Ausnahme des Winters abgenommen. Die Wahrscheinlichkeit, dass nach einem sehr feuchten Tag wieder ein sehr feuchter Tag folgt, hat im Winter am größten (+6%), im Frühling immerhin noch um (+4.8%) zugenommen und im Sommer hingegen um -2.3% abgenommen. Die Wahrscheinlichkeiten, dass nach zwei bzw. drei Trockentagen der niederschlagsreichste Zustand folgt, haben im Sommer am stärksten abgenommen (-3.4% bzw. -3.5%).

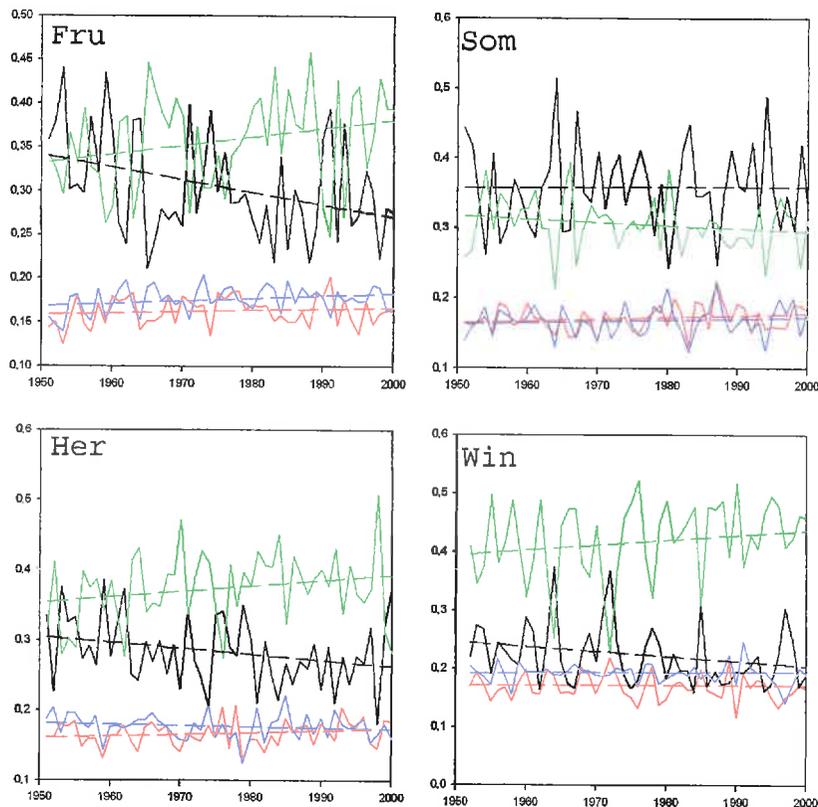


Abbildung 5.29: zeitlicher Verlauf von p_{3x} für Frühling, Sommer, Herbst und Winter für alle 301 Niederschlagsstationen; $x=0$ schwarz, $x=1$ rot, $x=2$ blau, $x=3$ grün

Nun soll die Gesamtstatistik verlassen werden und auf die geografische Verteilung der Steigungen der linearen Regressionsgerade einiger Übergangswahrscheinlichkeiten eingegangen werden. Dazu sind in den nächsten Abbildungen die Trends gezeigt, die - wie schon in Kapitel 5.2.2 - den Übergang von einem oder mehreren Trockentagen zu einem

sehr feuchten Tag bzw. von einem sehr feuchten Tag zu Trockentagen beschreiben. Im Anhang B.2.3 befinden sich Tabellen für die Trendwerte von Bremen, Frankfurt und der Zugspitze. In Abbildung 5.30 ist die Steigung der Übergangswahrscheinlichkeit p_{03} für alle vier Jahreszeiten gezeigt. Diese Zustandskombination ist nach Berechnung über den Gesamtzeitraum (vgl. Tab. 11 und 12) mit einer Eintrittswahrscheinlichkeit vergleichsweise niedrig bei etwa 10% und darum sind, wenn augenscheinlich auch nur geringfügige Änderungen auftreten, diese umso gewichtiger zu beurteilen. So erkennt man (Abb. 5.30), dass der absolute Trend im Großen und Ganzen in allen Jahreszeiten nicht mehr als +5% - eine Steigung von 0.001 entspricht über 50 Jahre gesehen einen Anstieg von 5% - und ebenso nicht weniger als 5% beträgt. Im Sommer sind praktisch über dem gesamten deutschen Gebiet Abnahmen festzustellen mit Ausnahme von Berlin plus der Ostseeküste sowie auch vereinzelt Stationen, wo eine Zunahme stattfindet. Frühling und Herbst zeigen wieder ein ähnliches Bild. Aber im Herbst sind die Flächen, die eine Abnahme

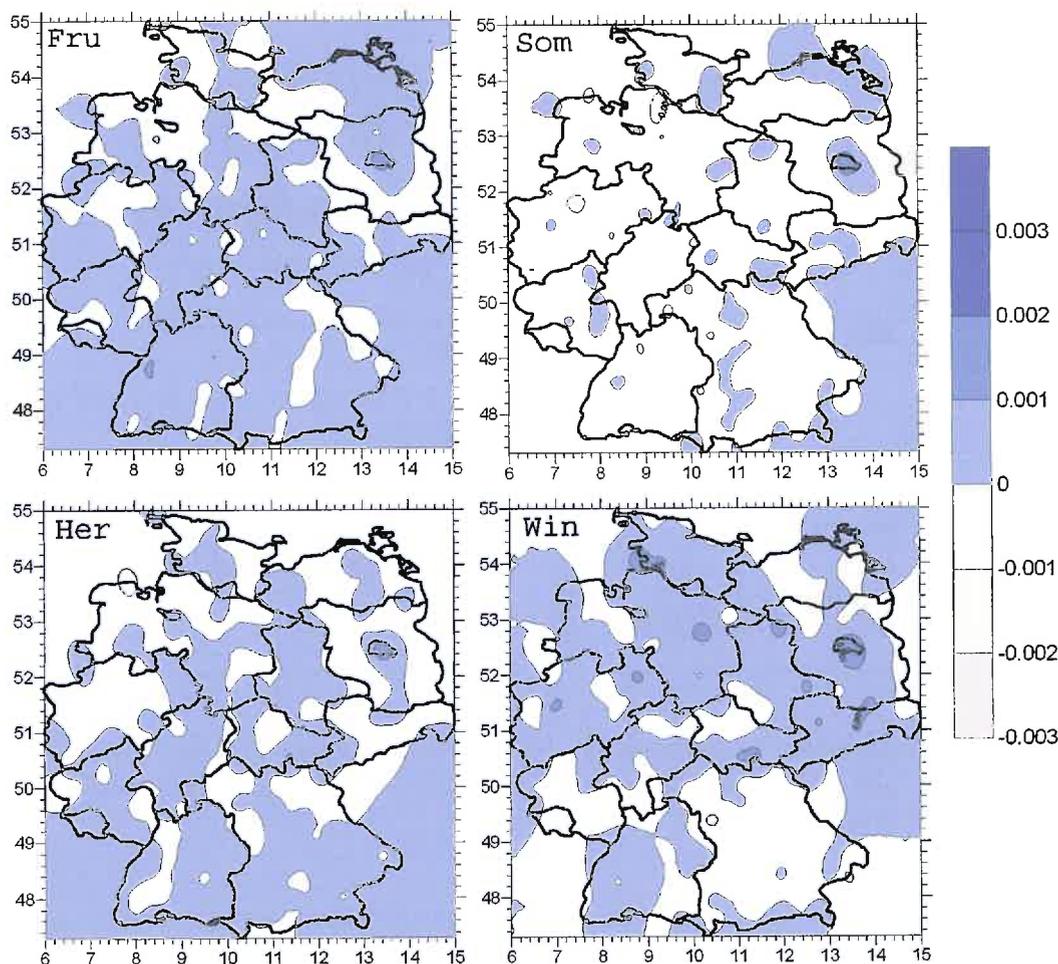


Abbildung 5.30: Steigung von p_{03} für Frühling, Sommer, Herbst und Winter berechnet für 1951-2000

zeigen größer als im Frühling. Ansonsten tritt in beiden Jahreszeiten eine Zunahme der Wahrscheinlichkeit p_{03} auf. Im Winter sind vor allem in Norddeutschland großflächige Gebiete mit Zunahmen auszumachen. Negative Steigungen befinden sich vor allem in Bayern und Rheinland-Pfalz.

In den nächsten beiden Abbildungen 5.31 und 5.32 sind die Steigungen der Regressionsgeraden gezeigt, die die Übergangswahrscheinlichkeiten von einem sehr feuchten Tag zu einem Trockentag sowie von einem sehr feuchten Tag zu wiederum einem sehr feuchten Tag beschreiben. Hier sind große Veränderung ersichtlich. Im Frühling hat eine sehr starke Abnahme von p_{30} (um -10 bis -20%) vor allem im Nordwesten stattgefunden und dies zugunsten einer Zunahme von p_{33} . Des Weiteren sind im Herbst und im Winter großflächige Abnahmen festzustellen, die vor allem Mittelddeutschland und Süddeutschland betreffen. Im Herbst sind allerdings in Brandenburg und Teilen von Schleswig-Holstein auch Zunahmen auszumachen, die bis zu +15% betragen. Im Sommer ist in Norddeutschland

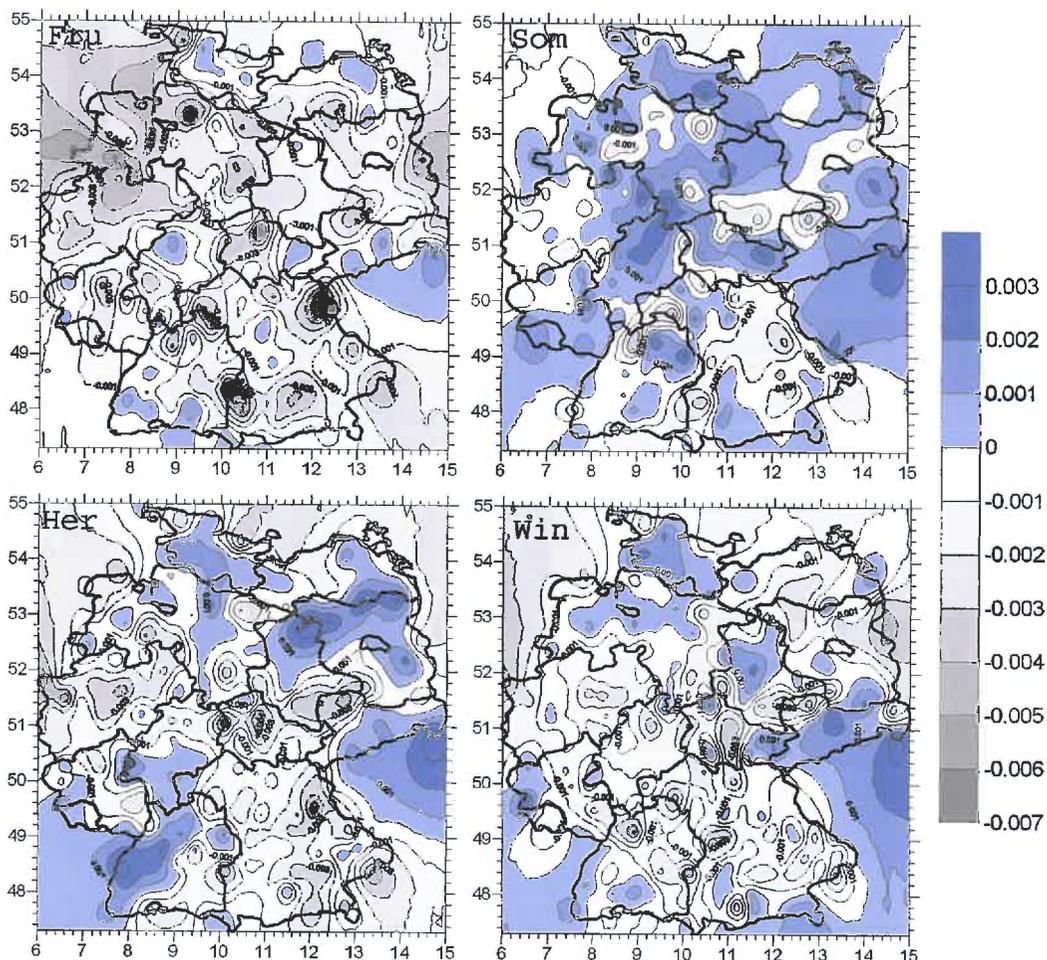


Abbildung 5.31: wie Abb. 5.30 jedoch für p_{30}

Hessen und Thüringen die Wahrscheinlichkeit, dass ein Trockentag nach einem Tag sehr starken Niederschlags folgt, angestiegen.

p_{33} (Abb. 5.32) zeigt im Winter den stärksten Anstieg und dies insbesondere im Norden (Werte um +15%). Aber auch im Frühling und im Herbst sind Zunahmen auszumachen und dies insbesondere nahe der Nordseeküste und in Westdeutschland. Im Herbst weisen einige Stationen in Sachsen-Anhalt und Brandenburg auch negative Trends auf, die teilweise bei -15% liegen. Im Sommer hat in großen Gebieten in Westdeutschland und Ostdeutschland eine Abnahme der Übergangswahrscheinlichkeit p_{33} stattgefunden, wohingegen in Bayern und in den Alpenregionen Zunahmen zu registrieren sind.

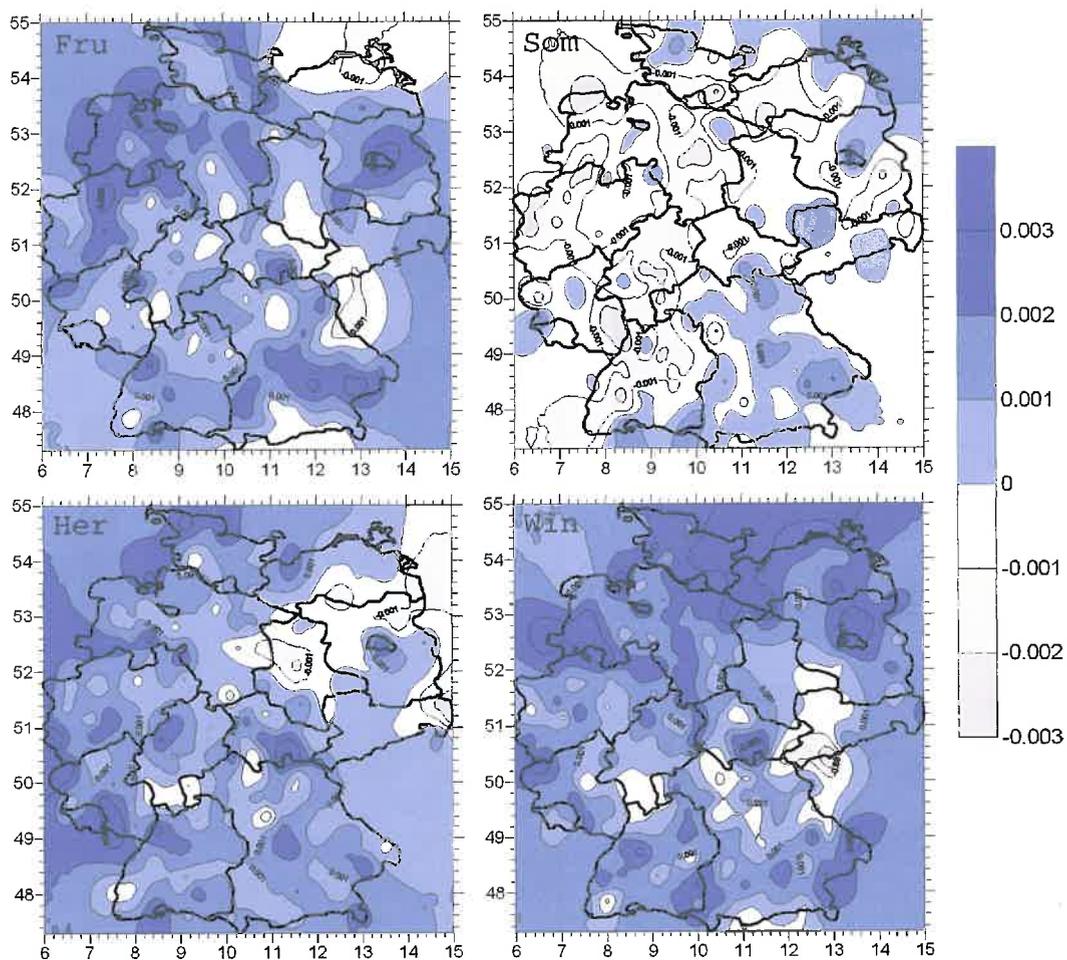


Abbildung 5.32: wie Abb. 5.30 jedoch für p_{33}

Als letztes soll untersucht werden, welche Veränderungen in den Wahrscheinlichkeiten aufgetreten sind, dass ein feuchter Tag nach einer Serie von Trockentagen folgt. In Abbildung 5.30 wurde schon die Wahrscheinlichkeit p_{03} untersucht. Diese Trends sollten erwartungs-

gemäß qualitativ mit denen von p_{003} und p_{0003} übereinstimmen, wobei bei letzteren sicher ein noch größerer Zusammenhang zu erwarten ist. Für p_{003} (Abb. 5.33) finden sich die stärksten Zunahmen um bis zu +5% im Winter und an der Nordseeküste um mehr als +5%. Dies stimmt in etwa auch mit den Veränderungen von p_{0003} (Abb. 5.34) überein, wobei aber auch Abnahmen von p_{0003} erkennbar sind. Im Sommer existieren bei beiden Wahrscheinlichkeiten Abnahmen, die in den meisten Gebieten bei bis zu -5% nur leicht ausgeprägt sind, in Westdeutschland aber mitunter auch bis zu -15% betragen können. Im Herbst sind vor allem in Süddeutschland Zunahmen auszumachen. Der Frühling zeichnet sich in den meisten Gebieten mit einer leichten Abnahme von p_{003} sowie auch von p_{0003} von bis zu -5% aus. Insgesamt treten bei diesen eher klein ausgeprägten Wahrscheinlichkeiten also, wie beschrieben, erhebliche Zu- oder Abnahmen auf.

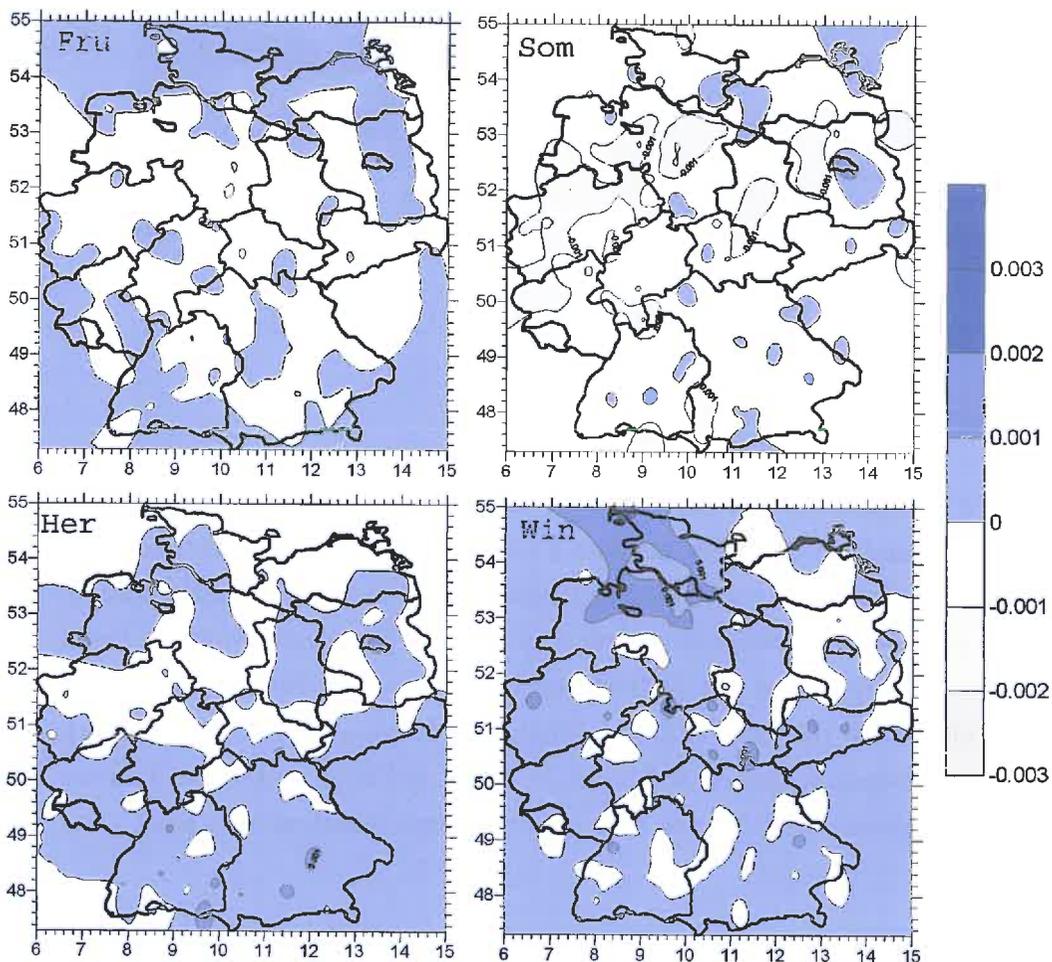


Abbildung 5.33: wie Abb. 5.30 jedoch für p_{003}

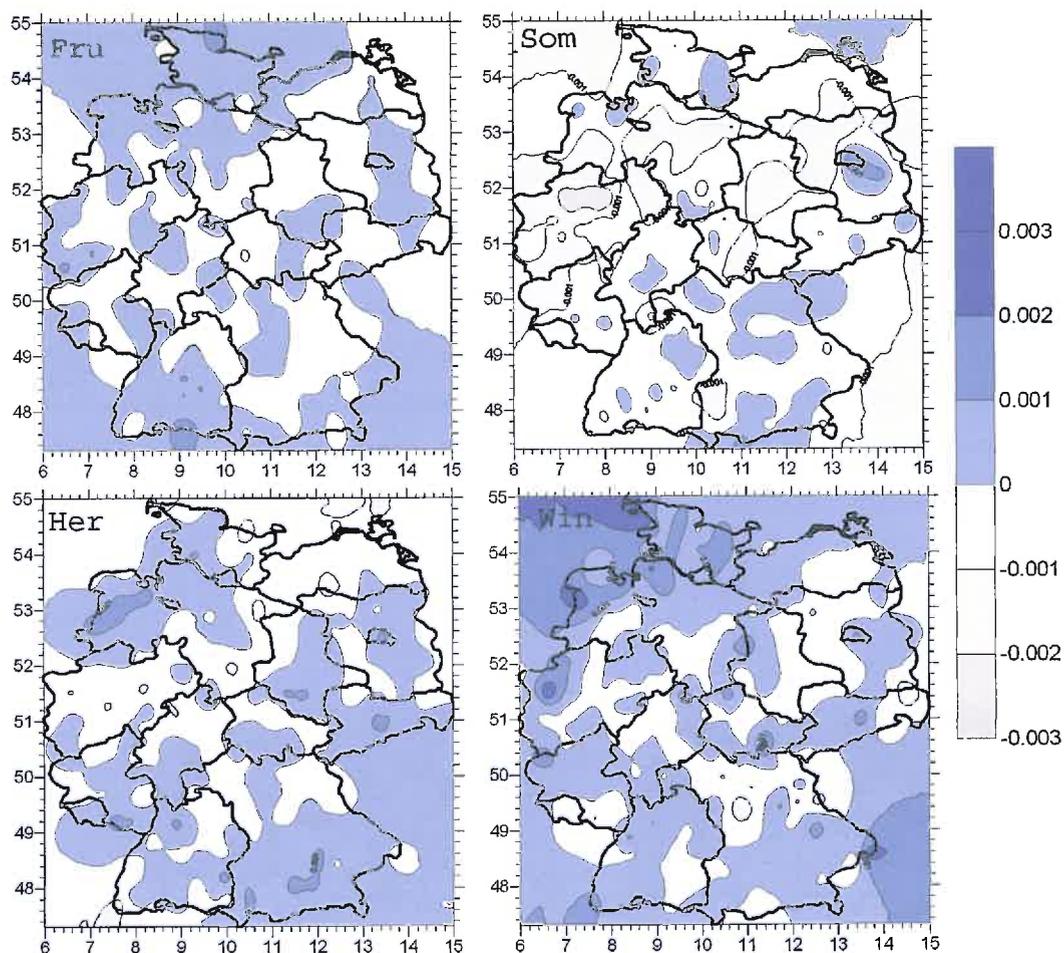


Abbildung 5.34: wie Abb. 5.30 jedoch für p_{0003}

Bemerkung zur Signifikanz

Wie in den Tabellen 13, B3, B4 und B5 ersichtlich ist, unterscheidet sich die Steigung des linearen Trends in den aller meisten Fällen signifikant von Null. Dasselbe gilt auch für die überwiegende Mehrheit der Stationen. Jedoch treten vereinzelt auch Steigungen auf - wie auch bei den drei exemplarisch gewählten Stationen Bremen, Frankfurt und Zugspitze - , die signifikant mit Null identisch sind wie z. B. für p_{30} . In diesem aber eher seltenen Fall betrifft das ebenso die Mehrheit der Stationen. Der zeitliche Verlauf dieser Übergangswahrscheinlichkeiten fluktuiert in der Regel sehr stark von Jahr zu Jahr, so dass der Test negativ ausfällt, was aber nicht verallgemeinert werden kann.

6 Wertung und Ausblick

Die Ordnung der Markov Kette gibt darüber Auskunft, welche Anzahl der vorherigen Zustände in Betracht gezogen werden, um den jetzigen Zustand im Rahmen des Markov Modells zu prognostizieren. Bei der Analyse mit zwei Zuständen hat sich in den Ergebnissen gezeigt, dass dafür zwei vorherige Tage ausreichen. Insgesamt zeigen sich keine regionalen und saisonalen Unterschiede, obwohl auf Grund der Natur der Niederschlagscharakteristik ein jahreszeitlicher Unterschied zu erwarten wäre. Das hängt damit zusammen, dass bei zwei Zuständen keine Niederschlagsintensität betrachtet wird. Wird nun der Zustandsraum erweitert auf vier Zustände, so zeigen sich sehr wohl jahreszeitliche Unterschiede. Im Winter werden höhere Ordnungen detektiert als im Sommer. Höhere Ordnungen deuten auf abwechslungsreichere Niederschlagswerte hin. Es sind einfach mehr Übergangswahrscheinlichkeiten in Betracht zu ziehen. Je niedriger die Ordnung, umso monotoner wird das Niederschlagsverhalten. Gerade im Sommer, da an weniger Tagen Niederschlag auftritt, dieser dafür aber umso stärker ausgeprägt ist, zeigen sich die kleinsten Ordnungen. Im Winter regnet es häufiger mit wechselnden Intensitäten, so dass mehr Übergangswahrscheinlichkeiten höherer Ordnung ungleich Null sind.

Um Veränderungen in den Ordnungen der Markov Ketten zu analysieren, müssen die Log-Likelihood Funktionen herangezogen werden. In diesen gehen nicht nur die Übergangswahrscheinlichkeiten ein sondern auch die Anzahl der aufgetretenden Zustandskombinationen. In dieser Arbeit wurden nur Veränderungen in den Übergangswahrscheinlichkeiten verfolgt. Wenn nun Ordnungswechsel stattfinden, beispielsweise von der dritten zur zweiten Ordnung, müssen im Prinzip 256 Übergangswahrscheinlichkeiten dritter Ordnung mit 64 Übergangswahrscheinlichkeiten zweiter Ordnung verglichen werden. Dieses ist direkt aber nicht möglich. So müssen über den gesamten Beobachtungszeitraum alle Übergangswahrscheinlichkeiten und Häufigkeiten der Zustandskombinationen auf Veränderungen untersucht werden, was im Rahmen dieser Arbeit nicht möglich ist.

Ein generelles Problem ist auch die Einteilung des Niederschlags in die Ordinalskala. In dieser Arbeit wurden Trockentage definiert als Tage ohne Niederschlag. Man kann natürlich andere Definitionen wählen gerade im Hinblick auf den hydrologisch wirksamen

Niederschlag, welcher definiert ist mit Werten höher als 1 mm Niederschlagshöhe.

Ein weiteres Problem ist die Anpassung der Weibullverteilung, auf dessen Grundlage die Zustandseinteilung erfolgte. Obwohl einige signifikante Anpassungen durchgeführt werden konnten, betrifft dies nicht alle Stationszeitreihen. Naturgemäß sind tägliche Niederschlagsdaten nur schwer mit theoretischen Verteilungen beschreibbar. In dieser Hinsicht könnte auch versucht werden, andere Verteilungen wie die Gammaverteilung zu verwenden. Mit diesem Hintergrund sind die Ordnungen, die Übergangswahrscheinlichkeiten und die Trends, die auf Grundlage der Weibullverteilung berechnet worden sind, auch nur hinsichtlich dieser Problematik verwertbar.

Des Weiteren kann man sich überlegen, ob der lineare Trend das geeignete Mittel ist, um Veränderungen in den Übergangswahrscheinlichkeiten auszumachen. Denkbar wäre auch eine zeitlich gleitende Analyse sowie der Test, ob die jeweilige Wahrscheinlichkeit signifikant von der Eintrittswahrscheinlichkeit abweicht [14, 23]. Bei der großen Fluktuation der Tagesdaten aber hat sich diese Methode als nicht durchführbar erwiesen. Darüber hinaus lässt dieser Test nur zu, ob Gleichheit oder Ungleichheit der Übergangswahrscheinlichkeiten besteht. Ob größer oder kleiner kann statistisch nicht festgestellt werden. Aus diesem Grund und wegen der Tatsache, dass oftmals die Steigung der Regressionsgerade signifikant von Null abweicht, wurde der lineare Trend mit entsprechendem Test der Signifikanz verwendet.

Die Anwendung von Markov Ketten in der Meteorologie ermöglicht noch ein weites Aufgabenfeld. Das Kerngebiet der Theorie der Markov Ketten aber umfasst wohl nur die Betrachtung von zwei Zuständen sowie höchstens die erste Ordnung (z. B. [3, 7, 11, 12, 23, 24]).

Literaturverzeichnis

- [1] **Anderson**, T. W. und L. A. Goodmann, 1957: Statistical Inference about Markov Chains, *The Annals of Math. Stat.*, Vol. 28, No. 1, pp. 89-110
- [2] **Akaike**, H., 1974: A new look at the statistical model identification, *IEEE Transactions on automatic control*, Vol. 19, pp. 716-723
- [3] **Behrends**, E., *Introduction to Markov Chains*, Vieweg, 2000
- [4] **Bosch**, K. , *Elementare Einführung in die angewandte Statistik*, Vieweg, 2005
- [5] **Doob**, J. L. , *Stochastic Processes*, John Wiley & Sons, 1953
- [6] **Fraedrich**, K. und K. Müller, 1983: On Single Station Forecasting: Sunshine and Rainfall Markov Chains, *Betr. Phys. Atmosph.*, Vol. 56, No. 1
- [7] **Gabriel**, K. R. und J. Neumann, 1962: A Markov chain model for daily rainfall occurrence at Tel Aviv, *Quart. J. Roy. Meteor. Soc.*, Vol. 88, pp. 90-95
- [8] **Jimoh**, O. D. und P. Webster, 1996: The optimum order of a Markov chain model for daily rainfall in Nigeria, *Journal of Hydrology*, 185, pp. 45-69
- [9] **Jonas**, M., Staeger, T. und C.-D. Schönwiese, Bericht Nr.1, Institut für Atmosphäre u. Umwelt, Univ. Frankfurt/Main, 2005
- [10] **Markov**, A. A., 1913: Example of a statistical investigation of the text of „Eugene Onegin“ illustrating the dependence between samples in chain, übersetzt von Morris Halle, *Bulletin de l'Académie Impériale des Sciences de St. Pétersbourg*, Bd. 7
- [11] **Murphy**, A. H. und R. W. Katz, *Probability, Statistics and Decision Making in the Atmospheric Science*, Westview Press, 1985
- [12] **Murphy**, A. H., 1981: On Some Criteria for Estimating the Order of a Markov Chain , *Technometrics*, Vol. 23, No. 3, pp. 243-249
- [13] **Press**, W. H. et al., *Numerical Recipes in Fortran 77*, Second Edition, Cambridge University Press, 2. Auflage, 2003

- [14] **Radivojevic**, K. , 2003: Markov Ketten, Fortgeschrittenen Praktikum der Meteorologie durchgeführt am Institut für Atmosphäre und Umwelt
- [15] **Raible**, C. C., Kirk, E. und K. Fraedrich, 2001: Kurzfristvorhersagen von Niederschlagswahrscheinlichkeit und Temperatur, *Promet* 27, pp. 80-86, Selbstverlag des Deutschen Wetterdienstes
- [16] **Rinne**, H., *Taschenbuch der Statistik*, Verlag Harri Deutsch, 2003
- [17] **Sachs**, L., *Angewandte Statistik*, Springer Verlag, 2004
- [18] **Schönwiese**, C.-D., *Klimatologie*, Ulmer Verlag, 2. Auflage, 2003
- [19] **Schönwiese**, C.-D., *Praktische Statistik für Meteorologen und Geowissenschaftler*, Gebrüder Bornträger, 4. Auflage, 2006
- [20] **Schwarz**, G., 1978: Estimating the dimension of a model, *Annals of Statistics*, Vol. 6 (2), pp. 461-464
- [21] **Stewart**, J. S., *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, 1994
- [22] **Storch**, H. von und F. W. Zwiers, *Statistical Analysis in Climate Research*, Cambridge University Press, 2. Auflage, 2002
- [23] **Trömel**, S., 2001: Statistik des Auftretens von Temperatur-, Niederschlag- und Windextrema sowie Schätzungen von Übergangswahrscheinlichkeiten mit Hilfe von Markovketten, Final report of a Fellowship with the Insurance company Munich Re.,
- [24] **Wilks**, D. S., *Statistical Methods in the Atmospheric Science*, International Geophysics Series, Vol. 59, Academic Press, 1995
- [25] **Wilks**, D. S. und R. L. Wilby, 1999: The weather generation game: a review of stochastic weather models, *Progress in Physical Geography*, Vol. 23 (3), pp. 329-357
- [26] **Wilks**, D. S., 1989: Rainfall Intensity, the Weibull Distribution, and Estimation of Daily Surface Run Off, *J. Appl. Met.*, Vol. 28 , pp. 52-58
- [27] **Wiser**, E. E., 1965: Modified Markov Probability Models of Sequences of Precipitation Events, *Monthly Weather Review*, Vol. 23 (8), pp. 511-516
- [28] **Wong**, R. K., 1977: Weibull Distribution, Iterative Likelihood Techniques and Hydrometeorological Data, *J. Appl. Met.*, Vol. 16, pp. 1360-1364

A Statistische Grundlagen

A.1 Axiome der Wahrscheinlichkeitsrechnung

Die Wahrscheinlichkeit $P(E)$ eines Ereignisses E gilt als nach der statistischen Methode bestimmt, wenn das zugehörige Zufallsexperiment n -mal wiederholt und $P(E)$ als $\frac{n(E)}{n}$ berechnet wird [17]. Dabei bezeichnet $n(E)$ die Anzahl der Experimente, in denen E eintritt. Es existieren drei Axiome, die die Wahrscheinlichkeiten definieren:

1. **Nichtnegativitätsaxiom:** Einem Ereignis E ist die Wahrscheinlichkeit $P(E)$ zugeordnet:

$$0 \leq P(E) \leq 1 . \quad (\text{A.1})$$

Bei einer Wahrscheinlichkeit von Null ist das betreffende Ereignis „fast unmöglich“, bei einer Wahrscheinlichkeit von 1 „fast sicher“.

2. **Normierungsaxiom:** Das sichere Ereignis S hat die Wahrscheinlichkeit Eins:

$$P(S) = 1 . \quad (\text{A.2})$$

3. **Additivitätsaxiom:** Die Wahrscheinlichkeit dafür, dass von mehreren, paarweise einander sich ausschließenden Ereignissen eines eintritt, ist gleich der Summe der Wahrscheinlichkeiten der Ereignisse:

$$P(E_1 \cup E_2 \cup \dots) = P(E_1) + P(E_2) + \dots \text{ wenn } E_i \cap E_j = \emptyset \text{ gilt für } i \neq j . \quad (\text{A.3})$$

A.2 Theoretische Verteilungen

A.2.1 Weibull Verteilung

Die Wahrscheinlichkeitsdichtefunktion $f_{wv}(x)$ der Weibullverteilung ($WV(\alpha, \beta, \gamma)$) ist gegeben durch Gleichung (A.4) mit dem Formparameter α , dem Streuparameter β und dem Lageparameter γ . Die Wahrscheinlichkeitsdichtefunktion der Weibullverteilung kann elementar integriert werden und man erhält die Verteilungsfunktion $F_{wv}(x)$.

$$f_{wv}(x) = \frac{\alpha}{\beta} \left(\frac{x - \gamma}{\beta} \right)^{\alpha-1} \exp \left[- \left(\frac{x - \gamma}{\beta} \right)^\alpha \right] \quad \text{für } x \geq \gamma \quad (\text{A.4})$$

$$F_{wv}(x) = 1 - \exp \left[- \left(\frac{x - \gamma}{\beta} \right)^\alpha \right] \quad (\text{A.5})$$

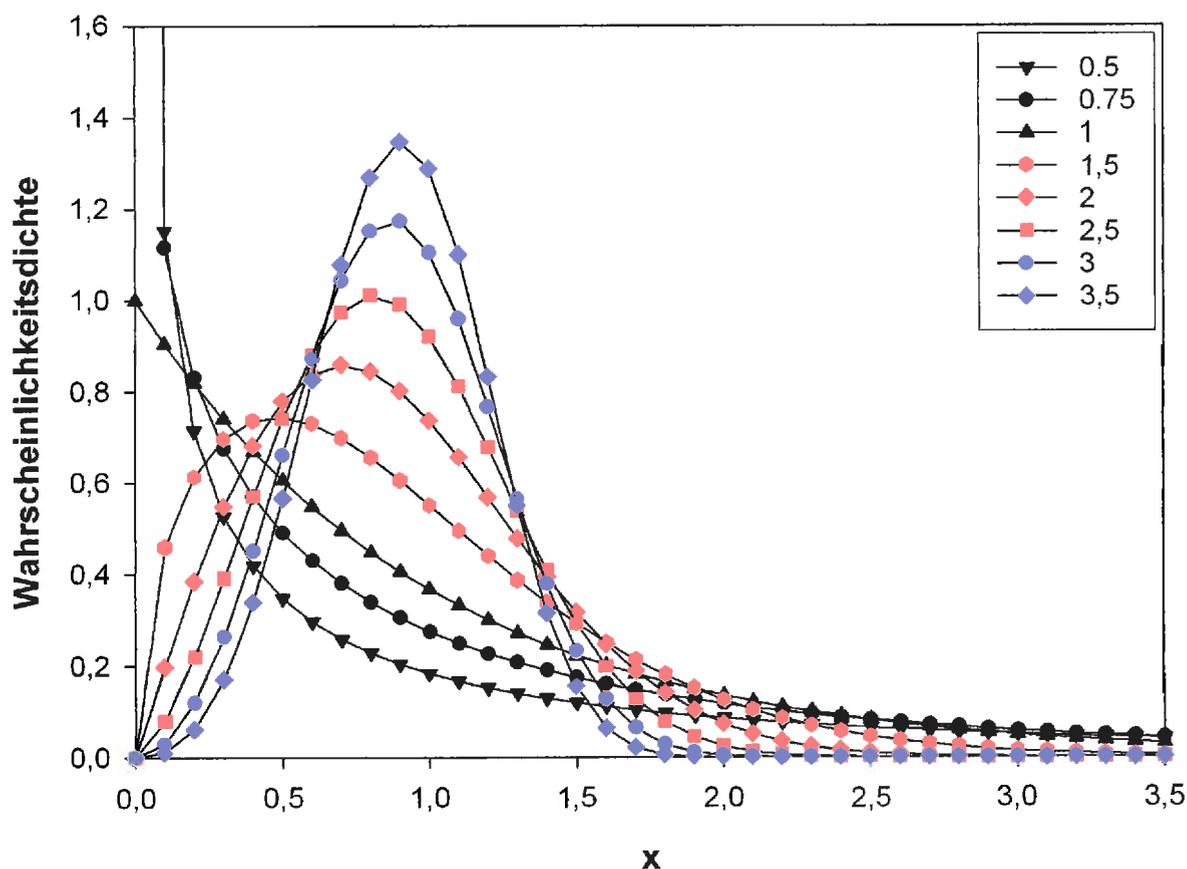


Abbildung A.1: Weibull Dichte für verschiedene Werte des Formparameters α , berechnet mit $\gamma = 0$ und $\beta = 1$

A.2.2 Anpassung einer Weibullverteilung

Nach Wong [28] und Wilks [24] werden zuerst die Parameter der Weibullverteilung $WV(\alpha, \beta, \gamma)$ durch die Methode der Momente berechnet. Diese werden dann mit Hilfe der Maximum Likelihood Methode und durch das Newton-Raphson Iterationsverfahren sukzessive verbessert. Da dieses Verfahren anfällig für „ungeeignete“ Startwerte ist [13], werden keine beliebigen Startwerte vorgegeben, sondern die geeigneteren Werte für die Verteilungsparameter aus der Methode der Momente verwendet.

Berechnung von α und β nach der Methode der Momente

Die WV hat den Mittelwert

$$\mu = \beta\tau \left(1 + \frac{1}{\alpha}\right) \quad (\text{A.6})$$

und die Varianz

$$\sigma^2 = \beta^2 \left[\tau \left(1 + \frac{2}{\alpha}\right) - \tau \left(1 + \frac{1}{\alpha}\right)^2 \right] \quad (\text{A.7})$$

wobei $\tau(t) = \int_0^\infty \xi^{t-1} \exp(-\xi) d\xi$ die Gammafunktion ist [19].

Allgemein gilt für die Schätzung der Momente aus der Stichprobe [17]:

$$\mu_k = \frac{1}{n} \sum_{i=1}^N X_i^k \quad (\text{A.8})$$

Über die ersten beiden empirische Momente, kann $\Psi(\alpha)$ berechnet werden. Dieser Wert für $\Psi(\alpha)$ wird mit einem theoretisch berechnetem über die rechte Seite von Gleichung (A.9) verglichen.

$$\frac{\mu_2}{\mu_1^2} = \Psi(\alpha) = \frac{\tau \left(1 + \frac{2}{\alpha}\right)}{\tau \left(1 + \frac{1}{\alpha}\right)^2} \quad (\text{A.9})$$

Daraus wird ein Wert für α und $\chi(\alpha)$ bestimmt (für Details vgl. [28]). Über Gleichung (A.10) wird β berechnet.

$$\beta = \frac{\mu_2}{\mu_1} \chi(\alpha) \quad (\text{A.10})$$

Auf diese Weise erhält man α und β . Der Schätzer für den Lageparameter γ ist das Stichprobenminimum [16].

 Anwendung der Maximum Likelihood Schätzung und des Newton-Raphson Verfahrens

Die Maximum Likelihood Schätzung für α und β sind Lösungen von

$$\frac{\partial \ln L}{\partial \alpha} = 0 \text{ und } \frac{\partial \ln L}{\partial \beta} = 0 \quad (\text{A.11})$$

mit der Likelihood Funktion der WV

$$\ln L = N \ln \alpha - N \alpha \ln \beta + (\alpha - 1) \sum_{i=1}^N \ln x_i - \sum_{i=1}^N \left(\frac{x_i}{\beta} \right)^\alpha \quad (\text{A.12})$$

Beim Newton-Raphson Verfahren muss das Gleichungssystem (A.13) gelöst werden, das durch Taylorentwicklung von (A.11) hervorgeht. Nach jedem Iterationsschritt erhält man neue Werte für $\tilde{\alpha}$ und $\tilde{\beta}$, die auf der rechten Seite wieder eingesetzt werden für den nächsten Iterationsschritt. Wenn der Fehler zwischen α und $\tilde{\alpha}$ bzw. β und $\tilde{\beta}$ kleiner als 0.01 ist, wird das Iterationsverfahren abgebrochen und $\tilde{\alpha}$ und $\tilde{\beta}$ als Schätzer für die Weibullparameter verwendet.

$$\begin{pmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \mathcal{H}^{-1} \begin{pmatrix} -\frac{\partial \ln L}{\partial \alpha} \\ -\frac{\partial \ln L}{\partial \beta} \end{pmatrix} \text{ mit } \mathcal{H} = \begin{pmatrix} \frac{\partial^2 \ln L}{\partial \alpha^2} & \frac{\partial^2 \ln L}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ln L}{\partial \beta \partial \alpha} & \frac{\partial^2 \ln L}{\partial \beta^2} \end{pmatrix} \quad (\text{A.13})$$

Für die einzelnen partiellen Ableitungen gilt nach [28]:

$$\frac{\partial \ln L}{\partial \alpha} = -N \ln L + \sum_{i=1}^N \ln x_i - \sum_{i=1}^N \left(\frac{x_i}{\beta} \right)^\alpha \ln \frac{x_i}{\beta} \quad (\text{A.14})$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{\alpha}{\beta} \left[\sum_{i=1}^N \left(\frac{x_i}{\beta} \right)^\alpha - N \right] \quad (\text{A.15})$$

$$\frac{\partial^2 \ln L}{\partial \alpha^2} = -\frac{N}{\alpha^2} - \sum_{i=1}^N \left(\frac{x_i}{\beta} \right)^\alpha \ln^2 \left(\frac{x_i}{\beta} \right) \quad (\text{A.16})$$

$$\frac{\partial^2 \ln L}{\partial \beta^2} = \frac{\alpha}{\beta^2} \left[N - (\alpha + 1) \sum_{i=1}^N \left(\frac{x_i}{\beta} \right)^\alpha \right] \quad (\text{A.17})$$

$$\frac{\partial^2 \ln L}{\partial \beta \partial \alpha} = \frac{1}{\beta} \sum_{i=1}^N \left(\frac{x_i}{\beta} \right)^\alpha + \frac{\alpha}{\beta} \sum_{i=1}^N \left(\frac{x_i}{\beta} \right)^\alpha \ln \left(\frac{x_i}{\beta} \right) - \frac{N}{\beta} \quad (\text{A.18})$$

A.3 Statistische Tests

A.3.1 Test des Regressionskoeffizienten

Die Signifikanz des Regressionskoeffizienten β_{xy} eines linearen Regressionsmodells $y = \beta_0 + \beta_{xy}x$ kann mit dem t-Test bestimmt werden. Bei zweiseitiger Fragestellung wird die Nullhypothese ($H_0 : \beta_{xy} = 0$) gegen die Alternativhypothese ($H_A : \beta_{xy} \neq 0$) getestet, indem die Prüfgröße \tilde{t} mit einem tabellierten Wert $t_{n-2,\alpha}$ der Irrtumswahrscheinlichkeit α der t-Verteilung verglichen wird. Die Prüfgröße wird berechnet mit folgender Gleichung und nachstehender Hilfsgrößen (vgl. Sachs [17]):

$$\tilde{t} = \frac{|\beta_{xy}|}{s_{b_{xy}}} \quad (\text{A.19})$$

$$\begin{aligned} s_{b_{xy}} &= \frac{s_{x,y}}{\sqrt{Q_x}} \\ s_{y,x} &= \sqrt{\frac{Q_y - \frac{Q_{xy}^2}{Q_x}}{n-2}} \\ Q_x &= \sum (x - \bar{x})^2 \\ Q_y &= \sum (y - \bar{y})^2 \\ Q_{xy} &= \sum xy - \frac{\sum x \sum y}{n} \end{aligned}$$

A.3.2 Kolmogorov-Smirnov Test

Beim Kolmogorov-Smirnov Test wird der maximale Abstand zwischen der empirischen und der theoretischen Verteilungsfunktion als Kriterium für die Güte der Anpassung verwendet. Der Test ist verteilungsfrei und kann deshalb bei jeder theoretischen Verteilung angewendet werden (z. B. [4, 19]). Die Prüfgröße \tilde{d} berechnet sich wie folgt:

$$\tilde{d} = \frac{\text{Max}\{KH_k(\text{emp}) - KH_k(\text{theor})\}}{n} \quad (\text{A.20})$$

wobei n die Grösse der Stichprobe, $KH_k(\text{emp})$ bzw. $KH_k(\text{theor})$ die empirische bzw. kumulative klassenorientierte Häufigkeitsverteilung angeben.

A.3.3 Mann-Kendall Trendtest

Der Mann-Kendall Test wertet das relative Ansteigen oder Abfallen der Datenwerte aus, um eine Aussage über die Signifikanz eines Trends machen. Dieser Test ist verteilungsfrei.

Folgende Gleichungen werden verwendet:

Prüfgröße nach MANN:

$$Q_* = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^N \operatorname{sgn}(x_j - x_i)}{\frac{1}{2}N(N-1)} = \frac{S}{\frac{1}{2}N(N-1)} \quad (\text{A.21})$$

normierte Prüfgröße:

$$Q = \frac{Q_* - \mu}{\sigma} = \frac{S}{\sqrt{\frac{1}{18}(N-1)(2N+5)}} \quad (\text{A.22})$$

korrigierte Prüfgröße nach KENDALL:

$$Q = \frac{S}{\sqrt{\frac{1}{18} \left[(N-1)(2N+5) - \sum_l b_l (b_l - 1) (2b_l + 5) \right]}} \quad (\text{A.23})$$

(b_l : Anzahl der Bindungen des Wertes l)

B Ergänzende Karten

B.1 Niederschlag mit zwei Zuständen

B.1.1 Ordnung der Markov Ketten nach dem AIC

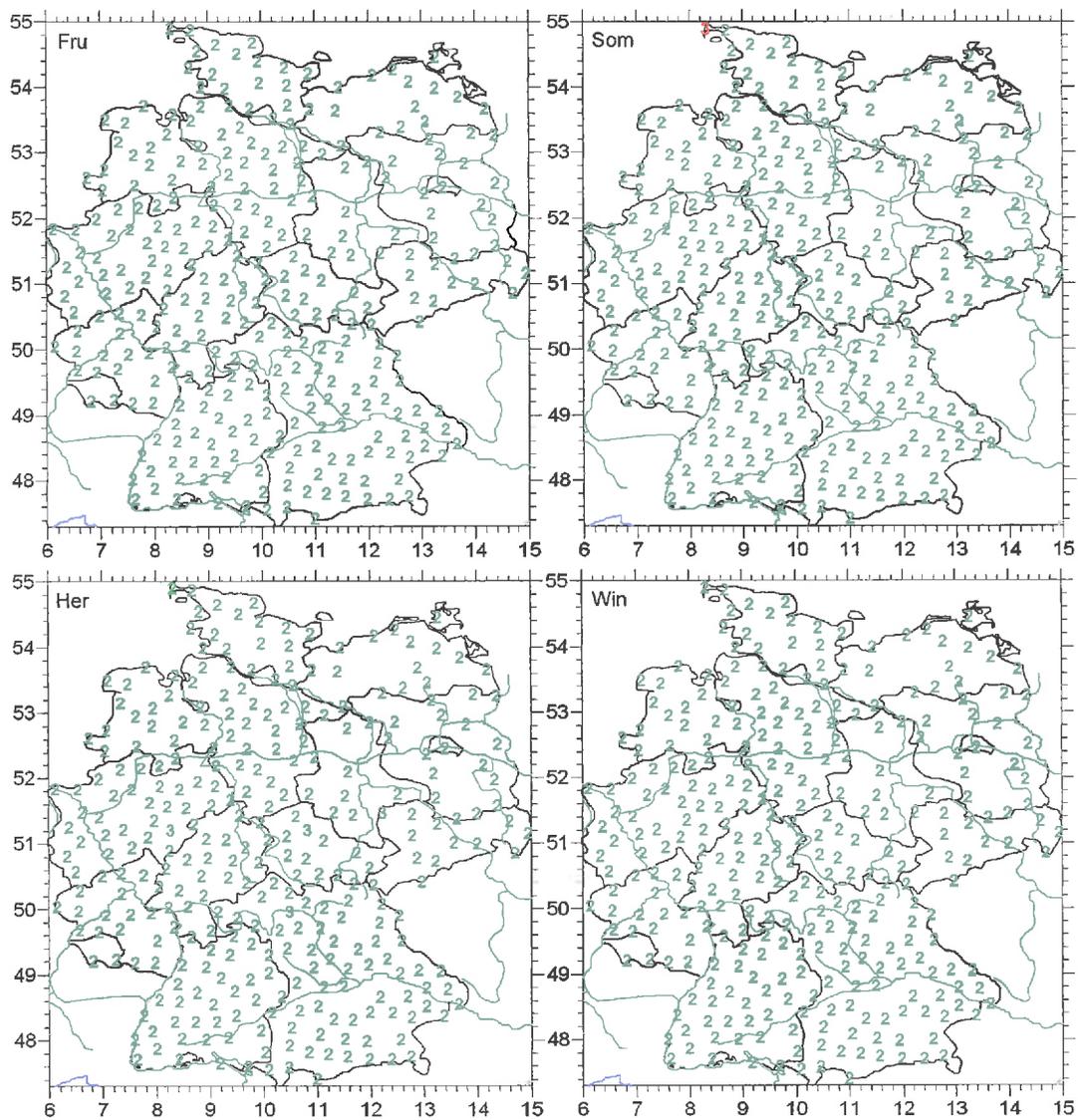


Abbildung B.1: berechnete Ordnung nach dem AIC für zwei Zustände (1951-2000)

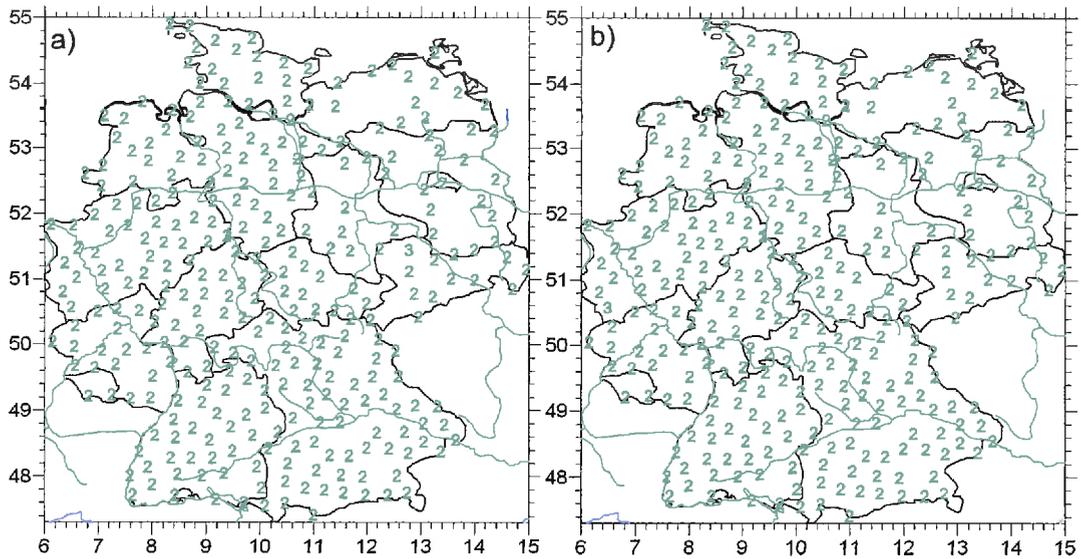


Abbildung B.2: Vergleich der Ordnungen nach AIC mit zwei Zuständen für den Frühling:
a) 1951-1975 und b) 1976-2000

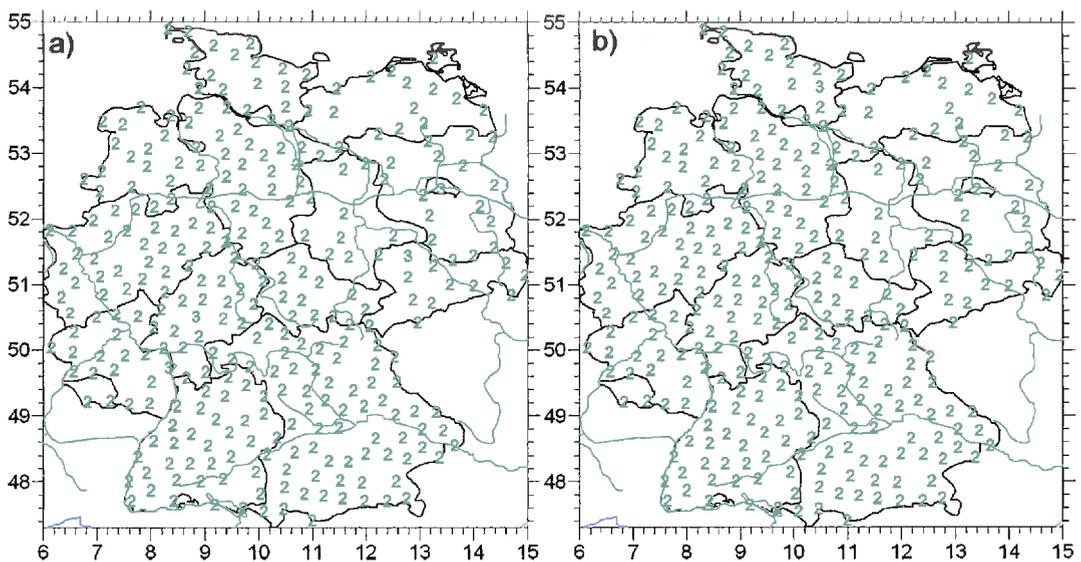


Abbildung B.3: Vergleich der Ordnungen nach AIC mit zwei Zuständen für den Winter:
a) 1951-1975 und b) 1976-2000

B.1.2 Ordnung der Markov Ketten nach dem BIC

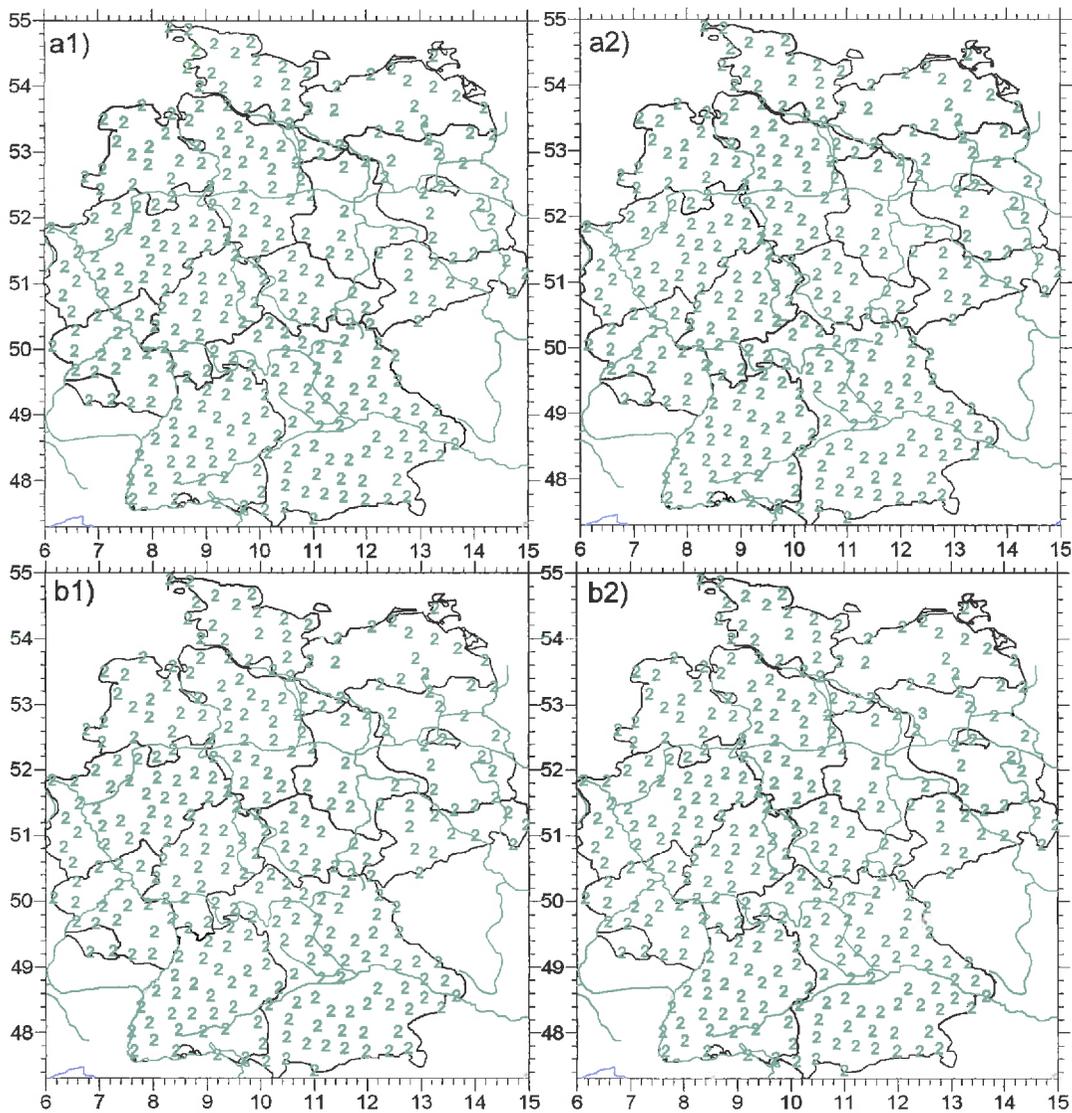


Abbildung B.4: Vergleich der Ordnungen nach dem BIC für zwei Zustände: **a1)** Frühling 1951-1975 und **a2)** Frühling 1976-2000, **b1)** Sommer 1951-1975 und **b2)** Sommer 1976-2000

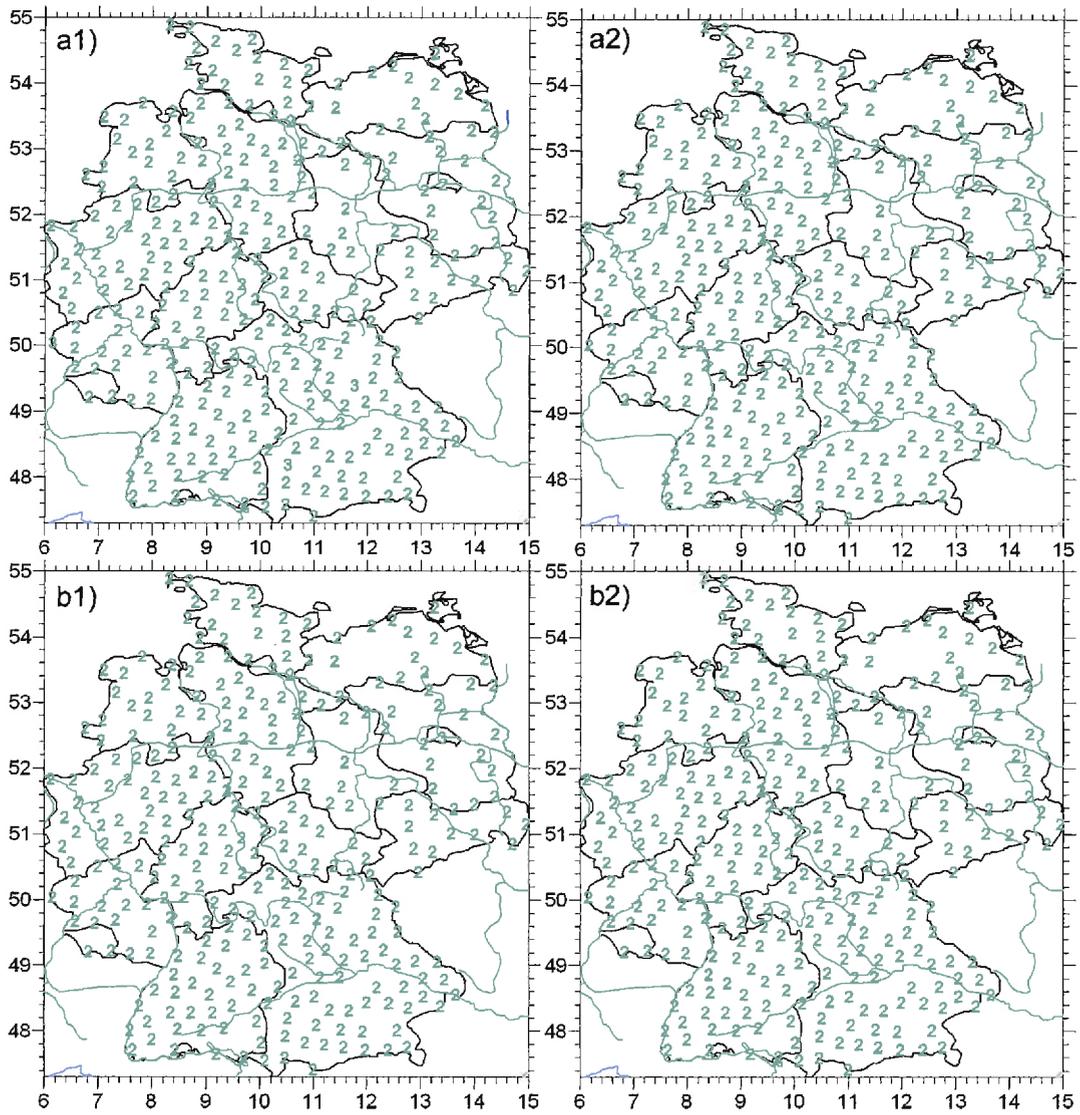


Abbildung B.5: Vergleich der Ordnungen nach dem BIC für zwei Zustände: **a1)** Herbst 1951-1975 und **a2)** Herbst 1976-2000, **b1)** Winter 1952-1975 und **b2)** Winter 1976-2000

B.1.3 Karten zu den Übergangswahrscheinlichkeiten zu Kap. 5.1.2

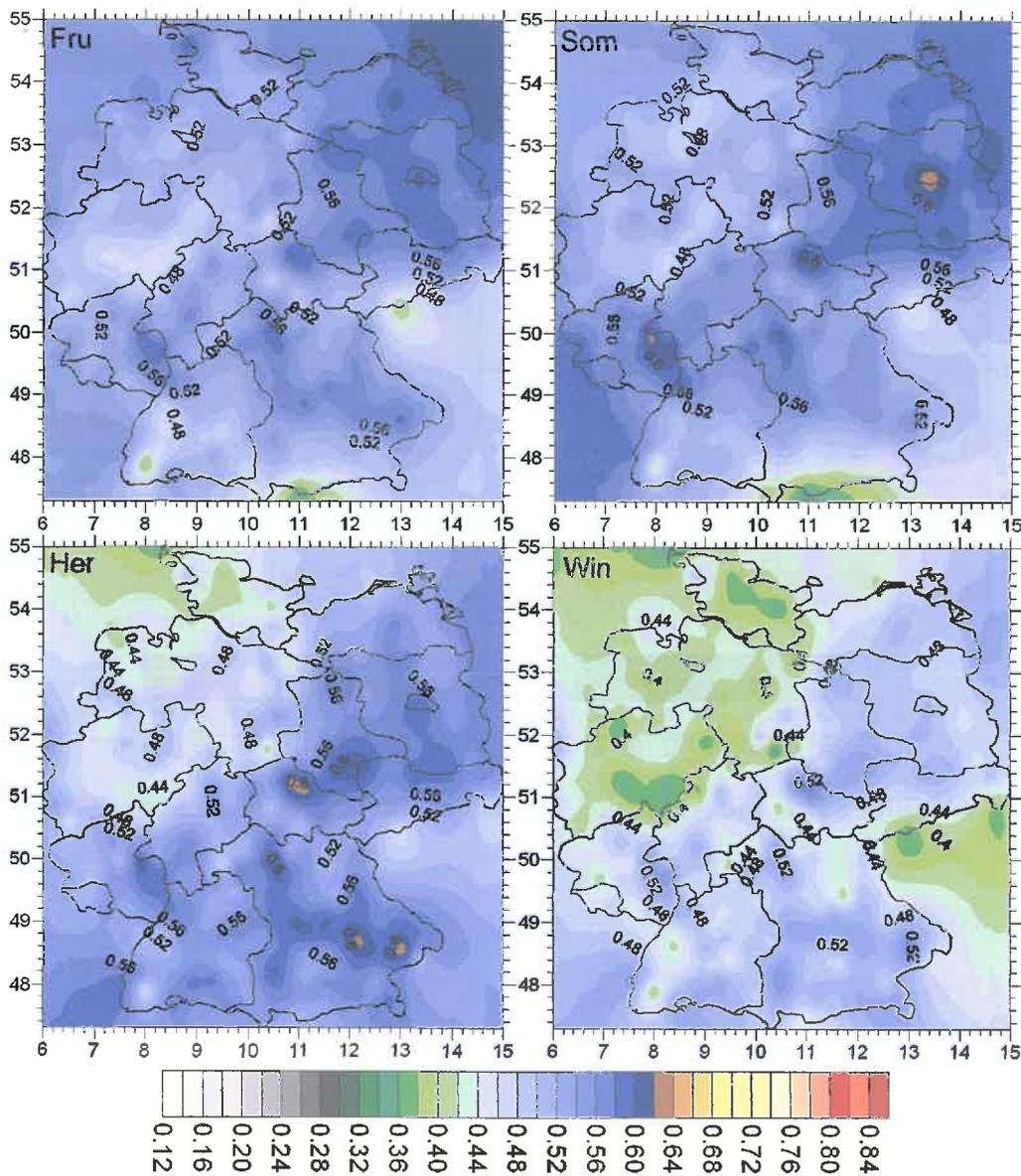


Abbildung B.6: Wahrscheinlichkeit p_0 für 1951-2000 im Frühling, Sommer, Herbst und Winter (von oben links nach unten rechts)

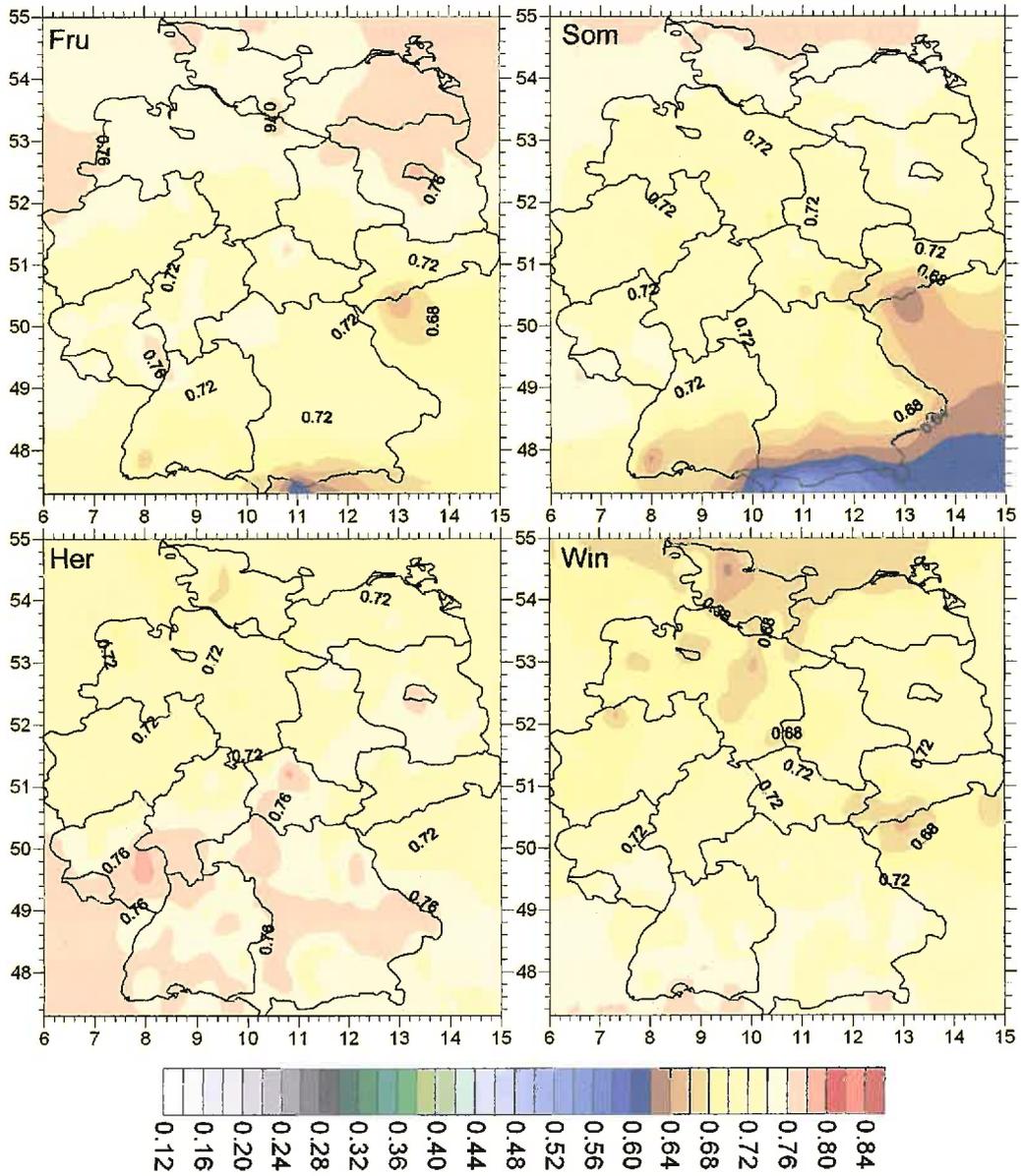


Abbildung B.7: wie Abbildung B.1 jedoch für Wahrscheinlichkeit p_{000}

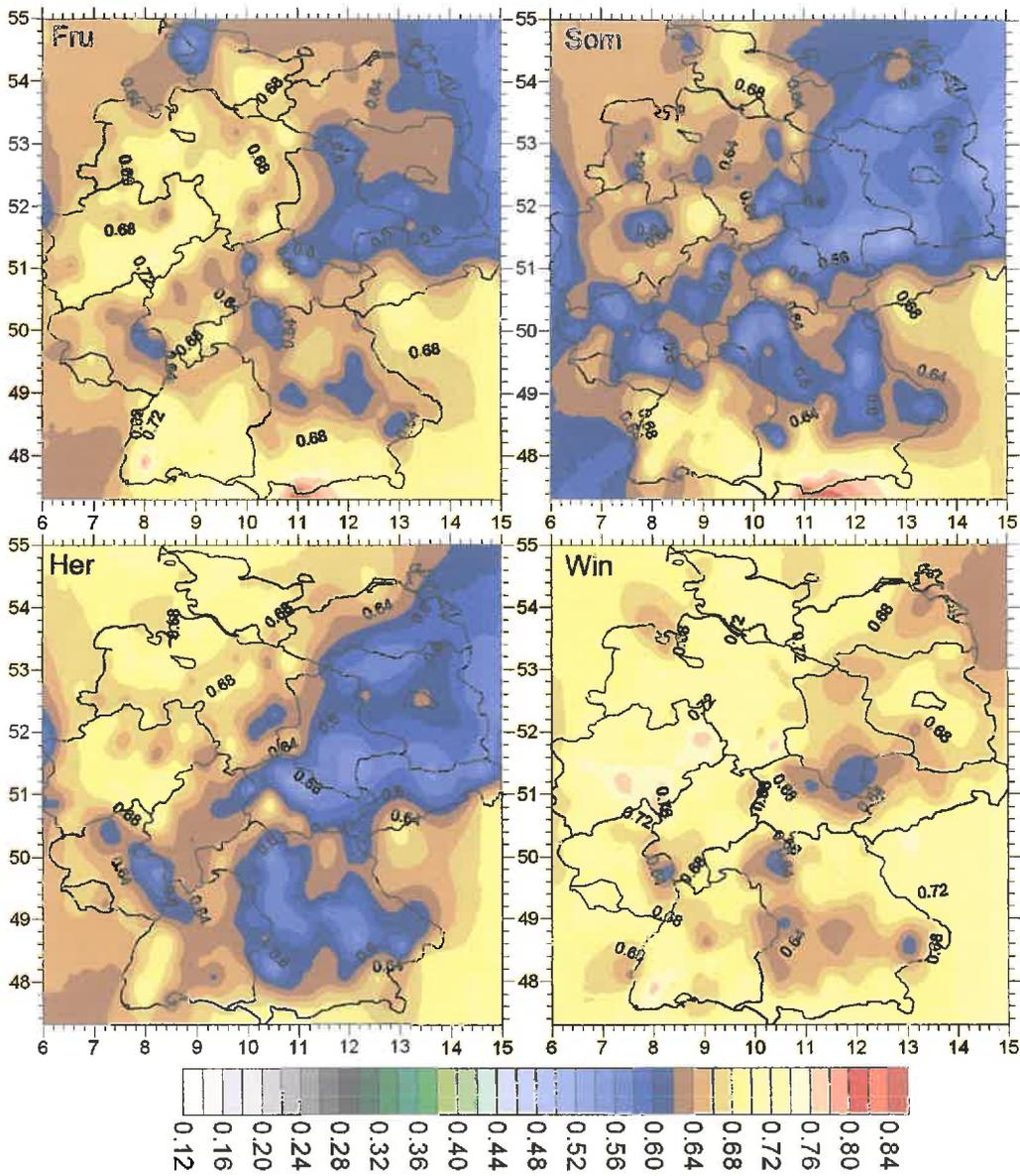


Abbildung B.8: wie Abbildung B.1 jedoch für Wahrscheinlichkeit p_{011}

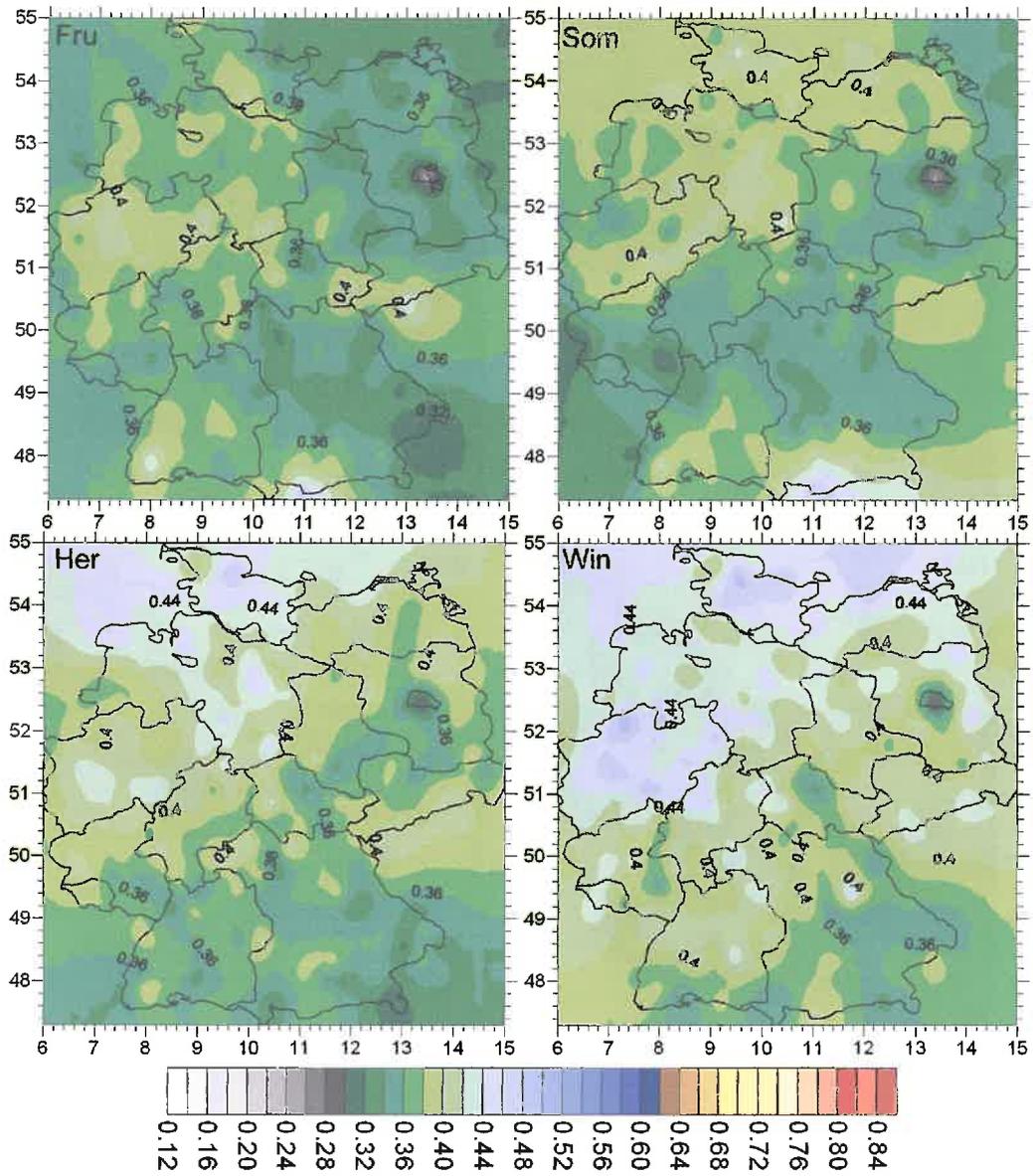


Abbildung B.9: wie Abbildung B.1 jedoch für Wahrscheinlichkeit p_{101}

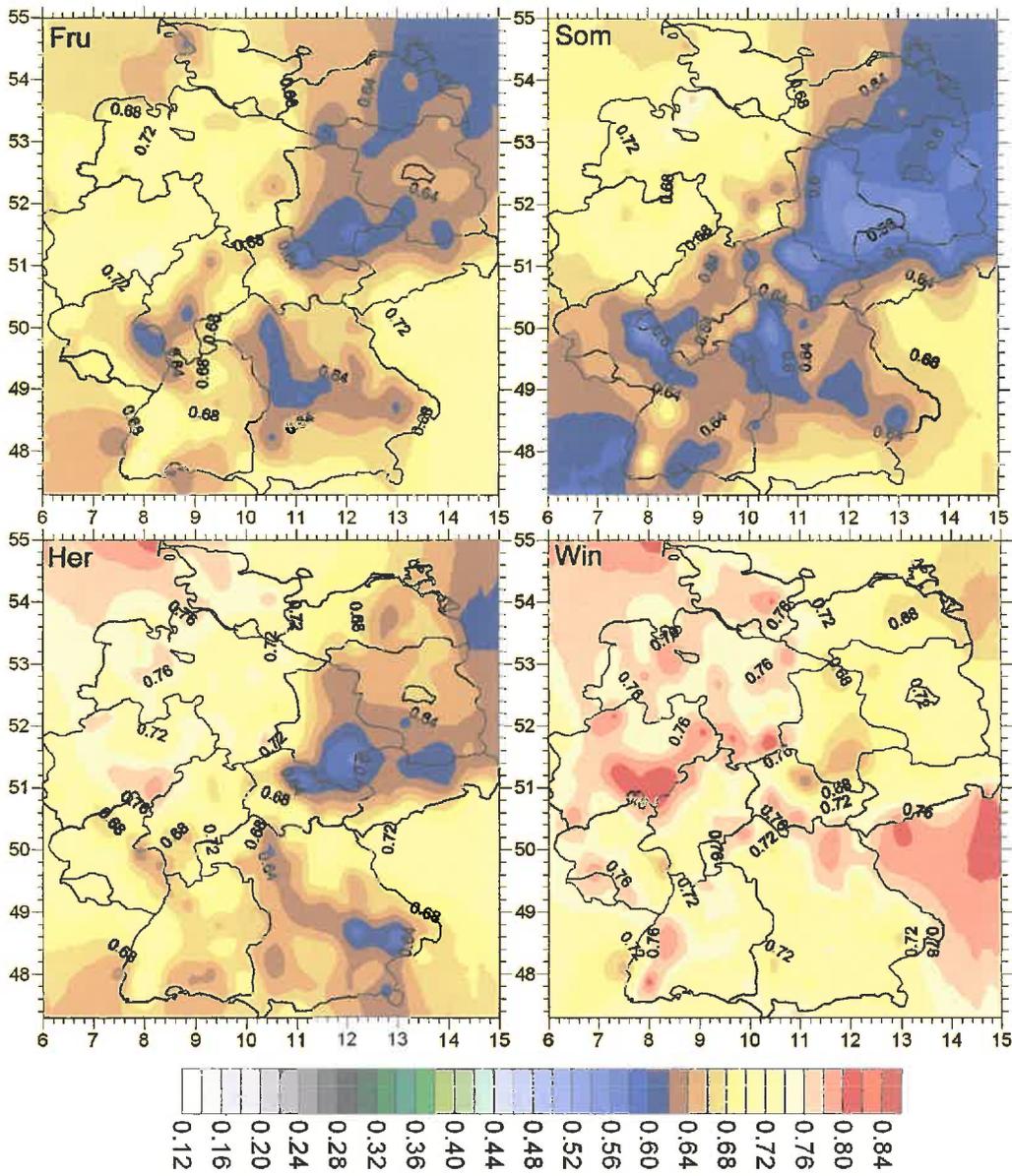


Abbildung B.10: wie Abbildung B.1 jedoch für Wahrscheinlichkeit p_{111}

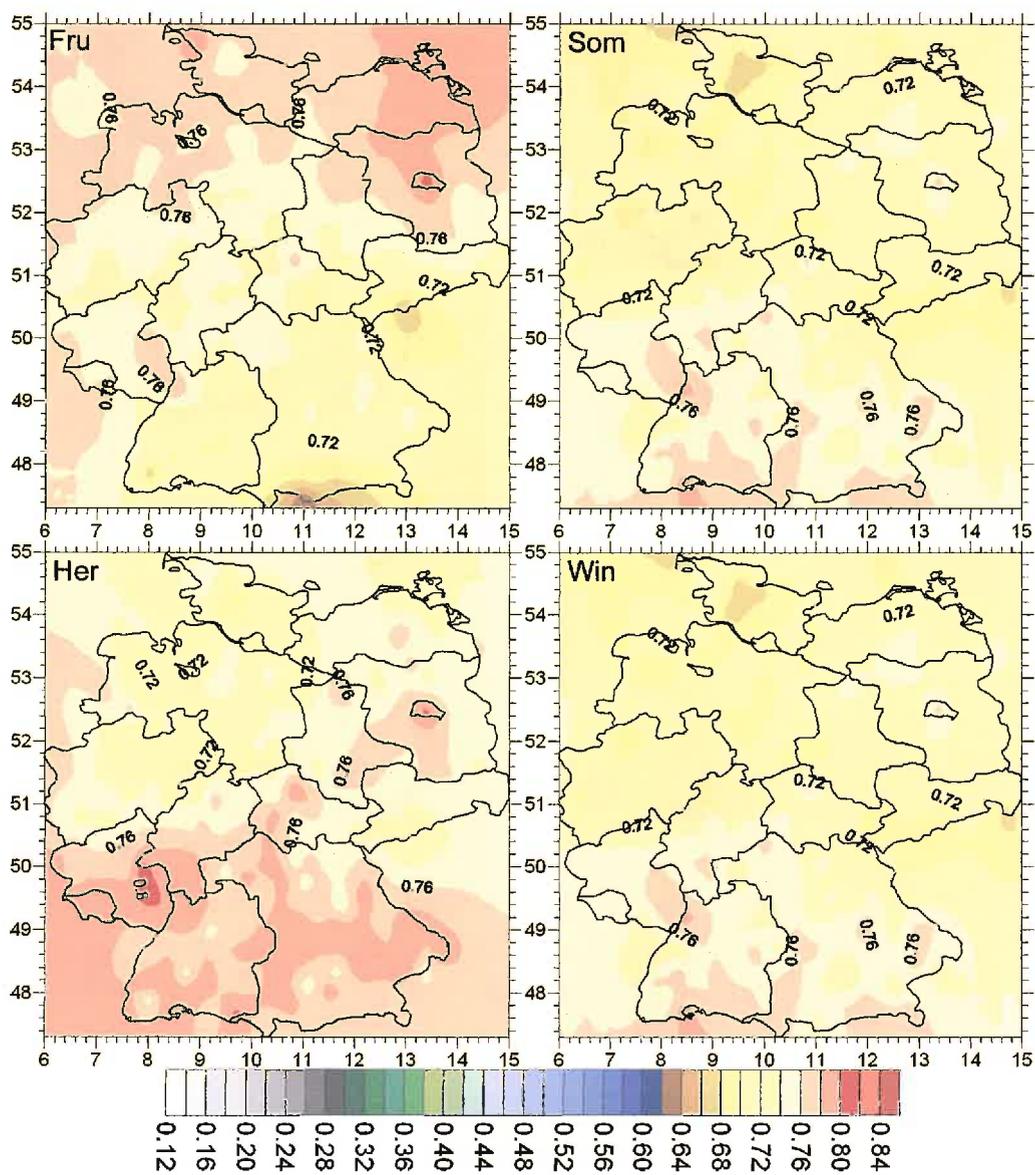


Abbildung B.11: wie Abbildung B.1 jedoch für Wahrscheinlichkeit $p=0000$

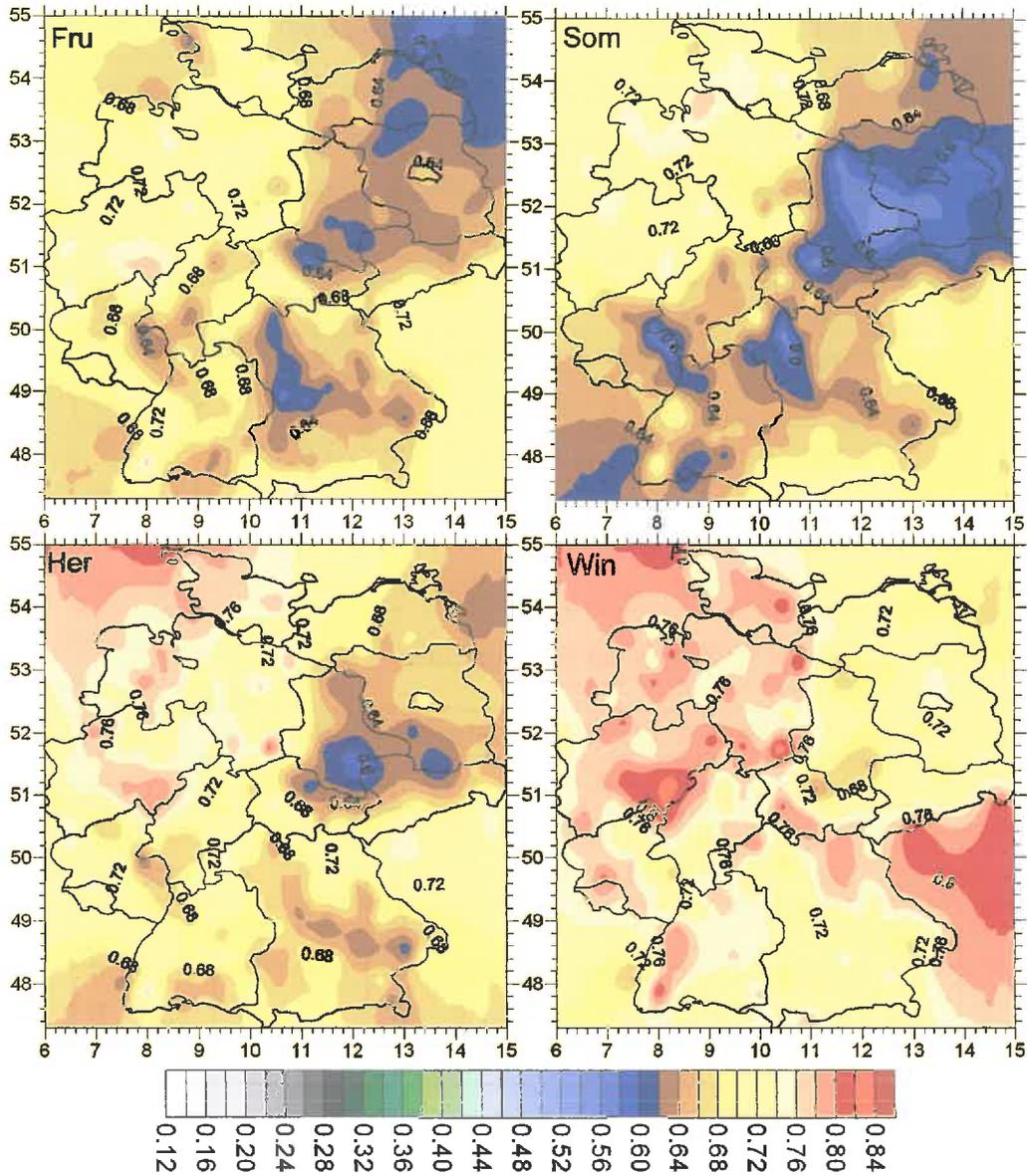


Abbildung B.12: wie Abbildung B.1 jedoch für Wahrscheinlichkeit p_{1111}

B.1.4 Trends der Übergangswahrscheinlichkeiten

Tabelle B1: lineare Wahrscheinlichkeitstrends in Prozent für die Übergangswahrscheinlichkeiten 0., 1. und 2. aufgeschlüsselt nach den Jahreszeiten für Frankfurt; die Zahl in Klammern gibt die Signifikanz der Steigung der Regressionsgerade an

Wahrscheinlichkeit	Frühling	Sommer	Herbst	Winter
p_1	+2.1 0	-6.2 0	+3.7 0	-7 90
p_{00}	-1.6 0	+6.4 0	-4 0	2.4 0
p_{11}	+0.8 90	-1.2 0	+1.9 0	-5.1 0
p_{000}	-1.9 0	+8.3 99	-2.6 0	+3.6 0
p_{011}	+1.5 0	-10.2 95	-0.8 0	-12.4 90
p_{100}	-1.2 0	+2.5 0	-6.3 0	0 0
p_{111}	+0.7 0	+4.1 0	+6.5 0	-3.3 0

Tabelle B2: wie Tabelle B1 jedoch für die Zugspitze

Wahrscheinlichkeit	Frühling	Sommer	Herbst	Winter
p_1	+4.4 0	-1 0	+5.4 0	-4.7 0
p_{00}	+2.6 0	-3.4 95	-4.2 0	+2.2 0
p_{11}	+4.9 90	-2.6 0	+1.1 0	-1.4 0
p_{000}	+6.3 0	-5.9 0	-2.6 0	-1.4 0
p_{011}	+3.5 0	-10.6 0	-9.6 90	+1.3 0
p_{100}	+1.2 0	-1.5 0	-1.8 0	+3.4 0
p_{111}	+5.3 95	-0.6 0	+6.1 90	-2 0

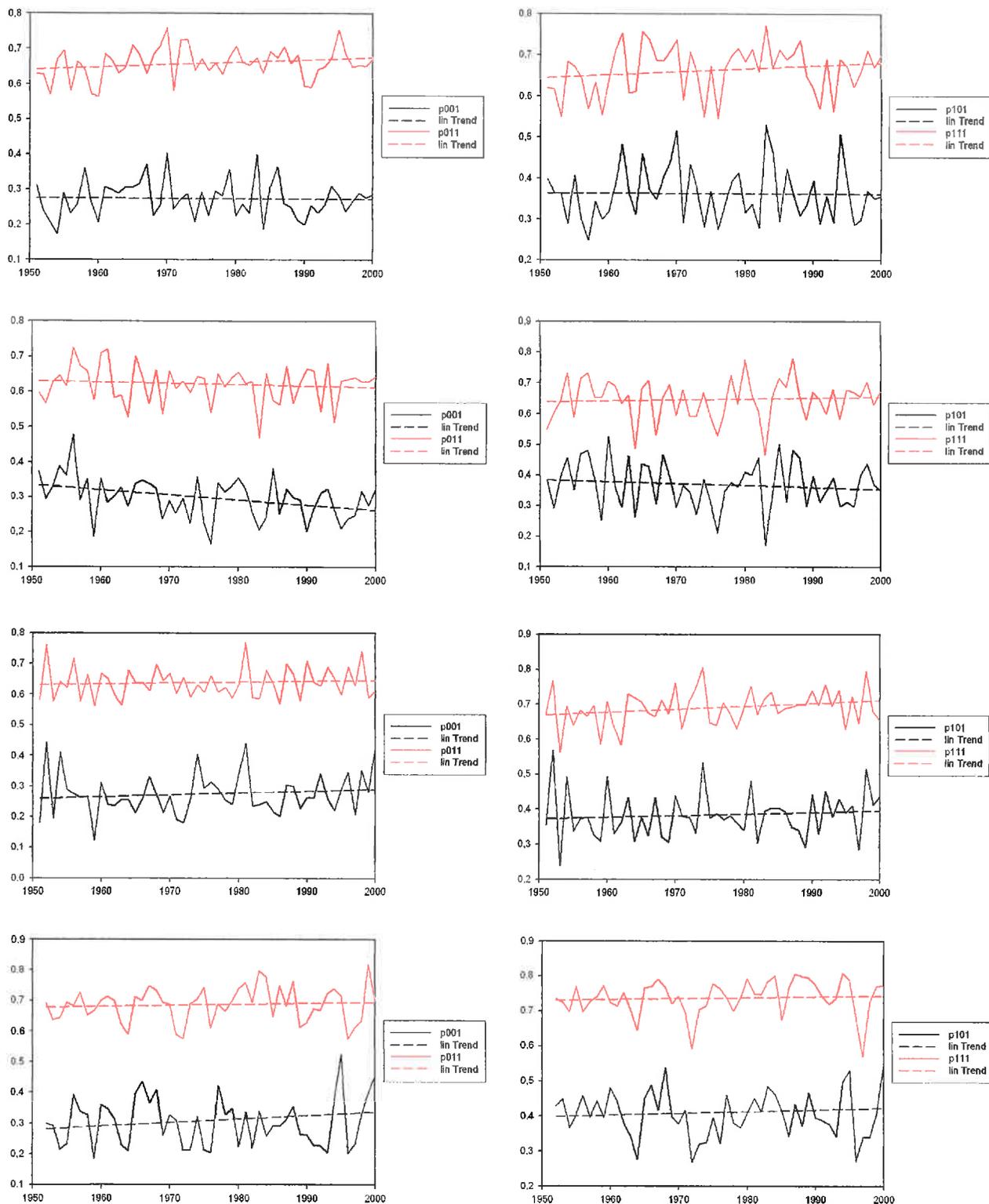


Abbildung B.13: jährliche Wahrscheinlichkeiten p_{001} und p_{011} (jeweils links) sowie p_{101} und p_{111} (jeweils rechts) berechnet alle Stationen mit linearem Trend: Frühling (oberer Reihe), Sommer (2. Reihe), Herbst (3. Reihe) und Winter (untere Reihe), vgl. Tab. 9

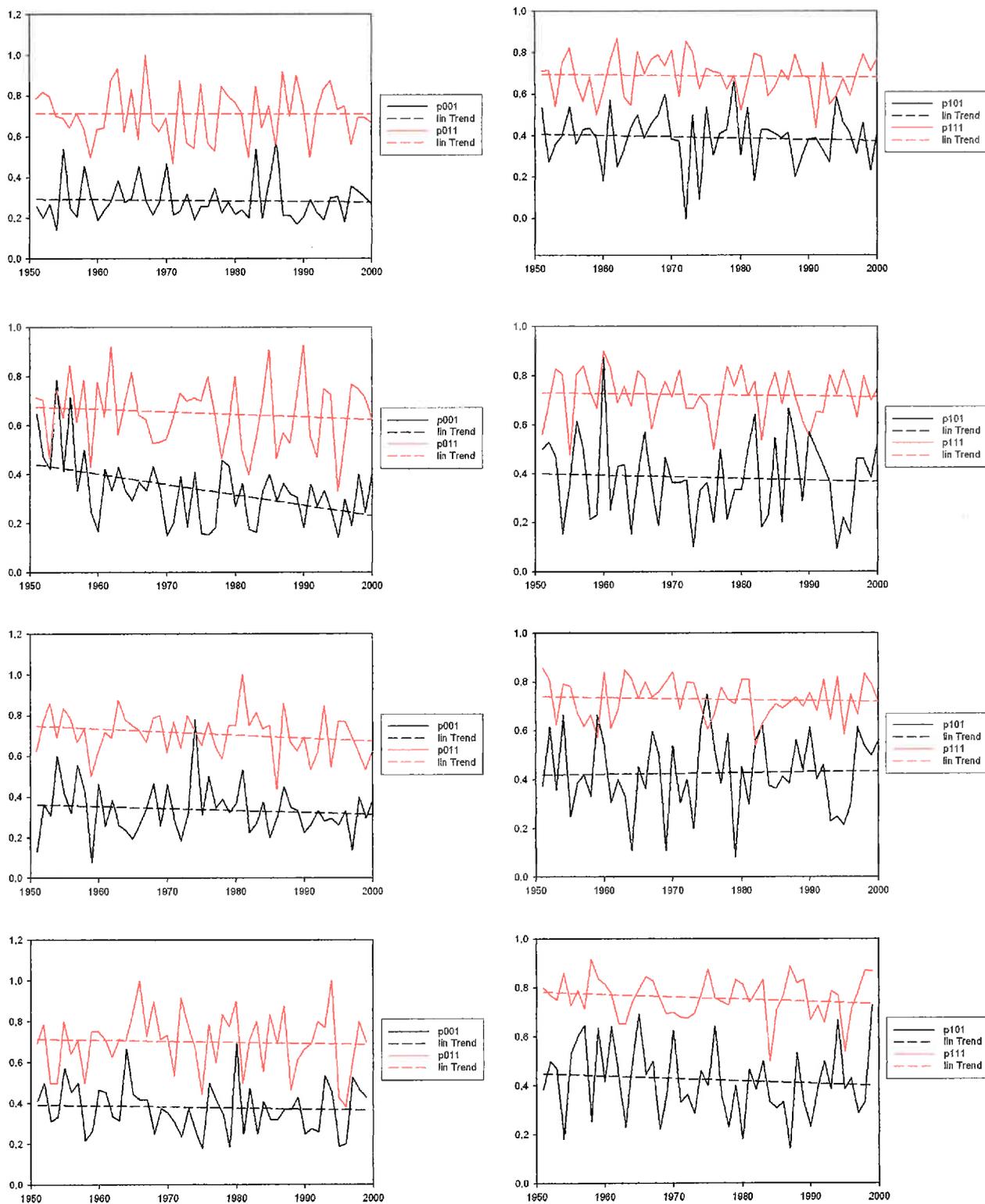


Abbildung B.14: wie Abb. B.13 jedoch für Bremen, vgl. Tab. 10

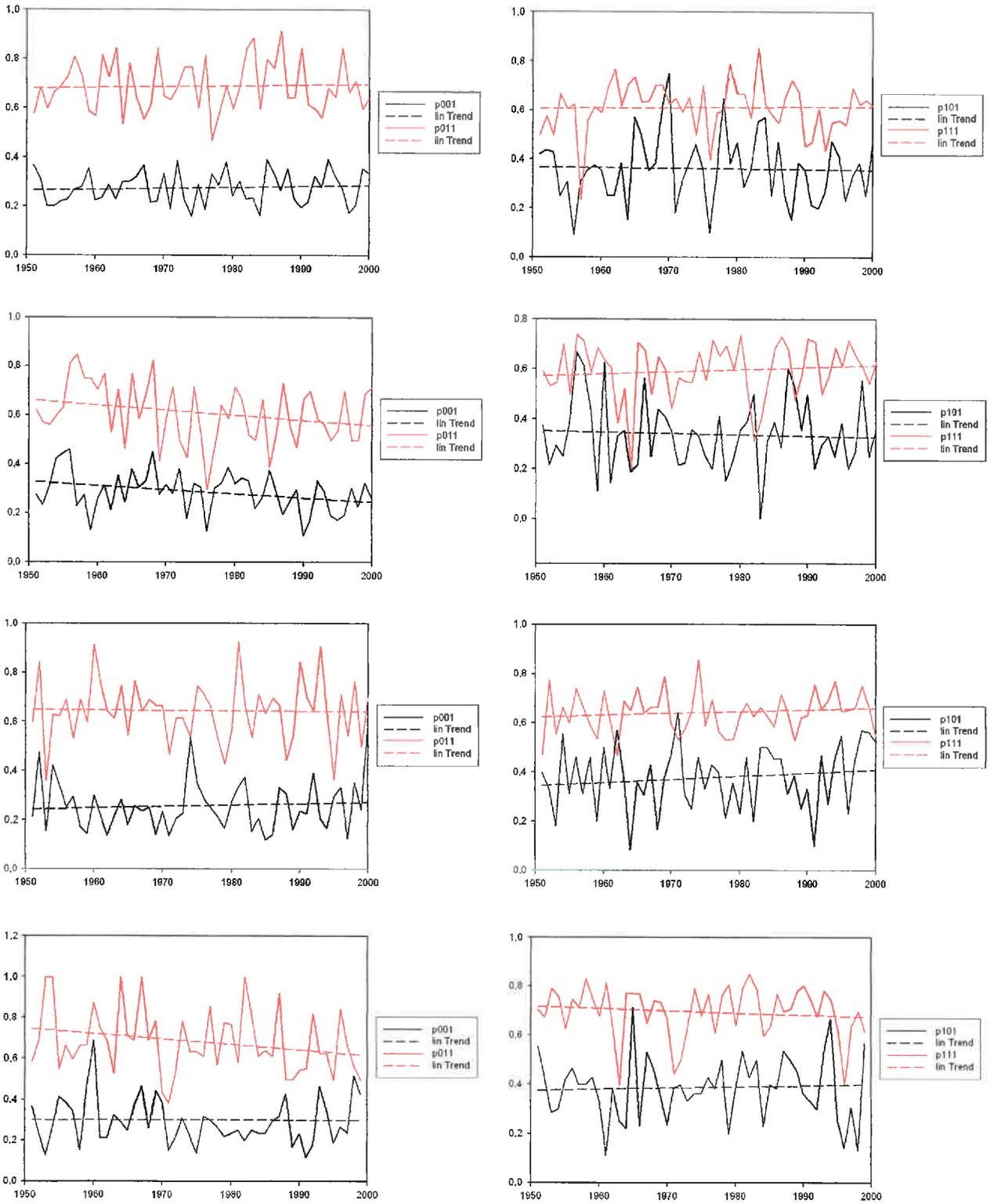


Abbildung B.15: wie Abb. B.13 jedoch für Frankfurt), vgl. Tab. B1

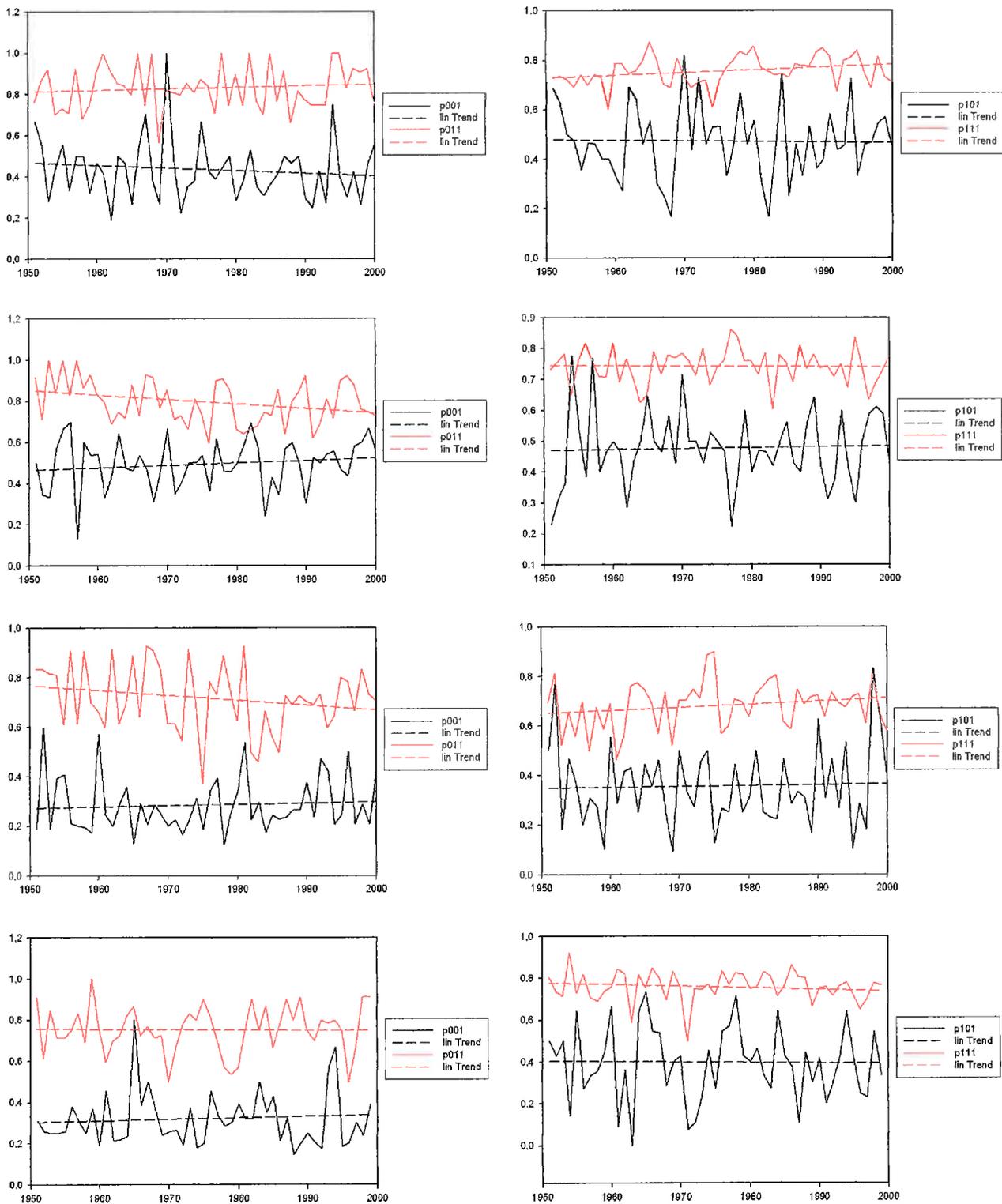


Abbildung B.16: wie Abb. B.13 jedoch für die Zugspitze, vgl. Tab. B2

B.2 Niederschlag mit vier Zuständen

B.2.1 Ordnung der Markov Ketten nach dem AIC

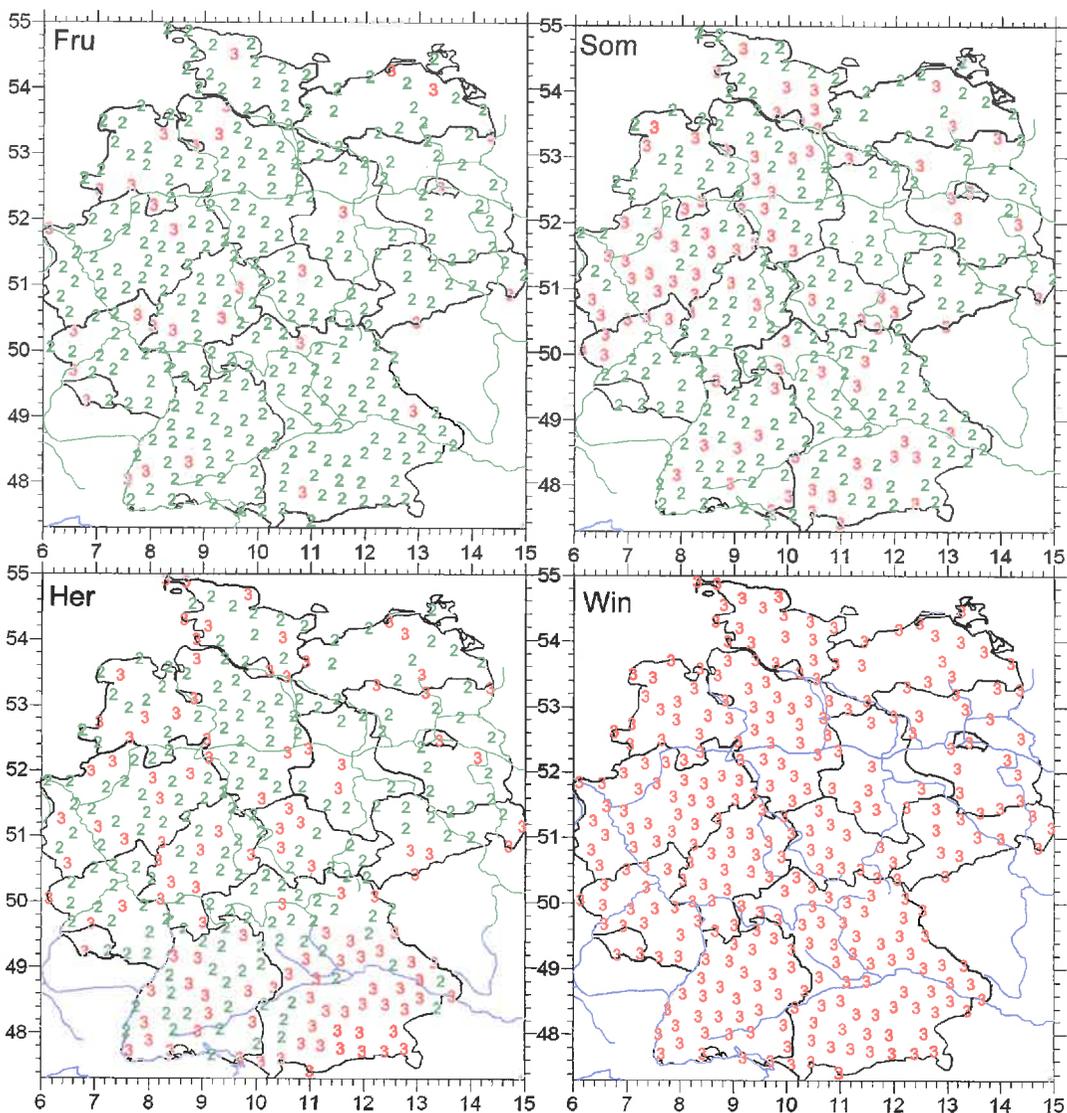


Abbildung B.17: Ordnung der Markov Ketten mit vier Zuständen nach dem AIC (1951-2000) für Frühling, Sommer, Herbst und Winter

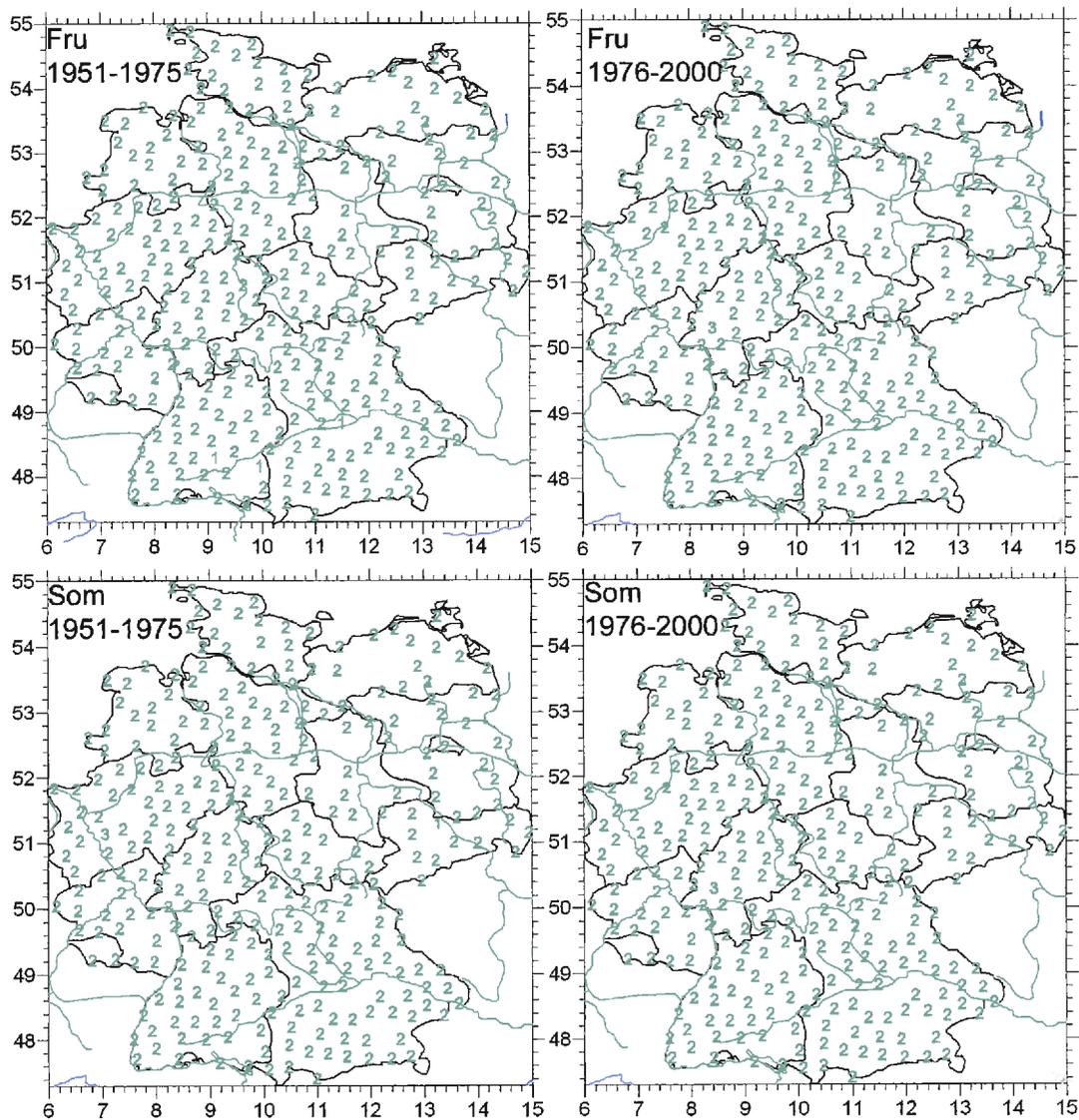


Abbildung B.18: Ordnung der Markov Ketten mit vier Zuständen nach dem AIC: Vergleich zwischen 1951-1975 und 1976-2000 für den Frühling (oben) und den Sommer (unten)

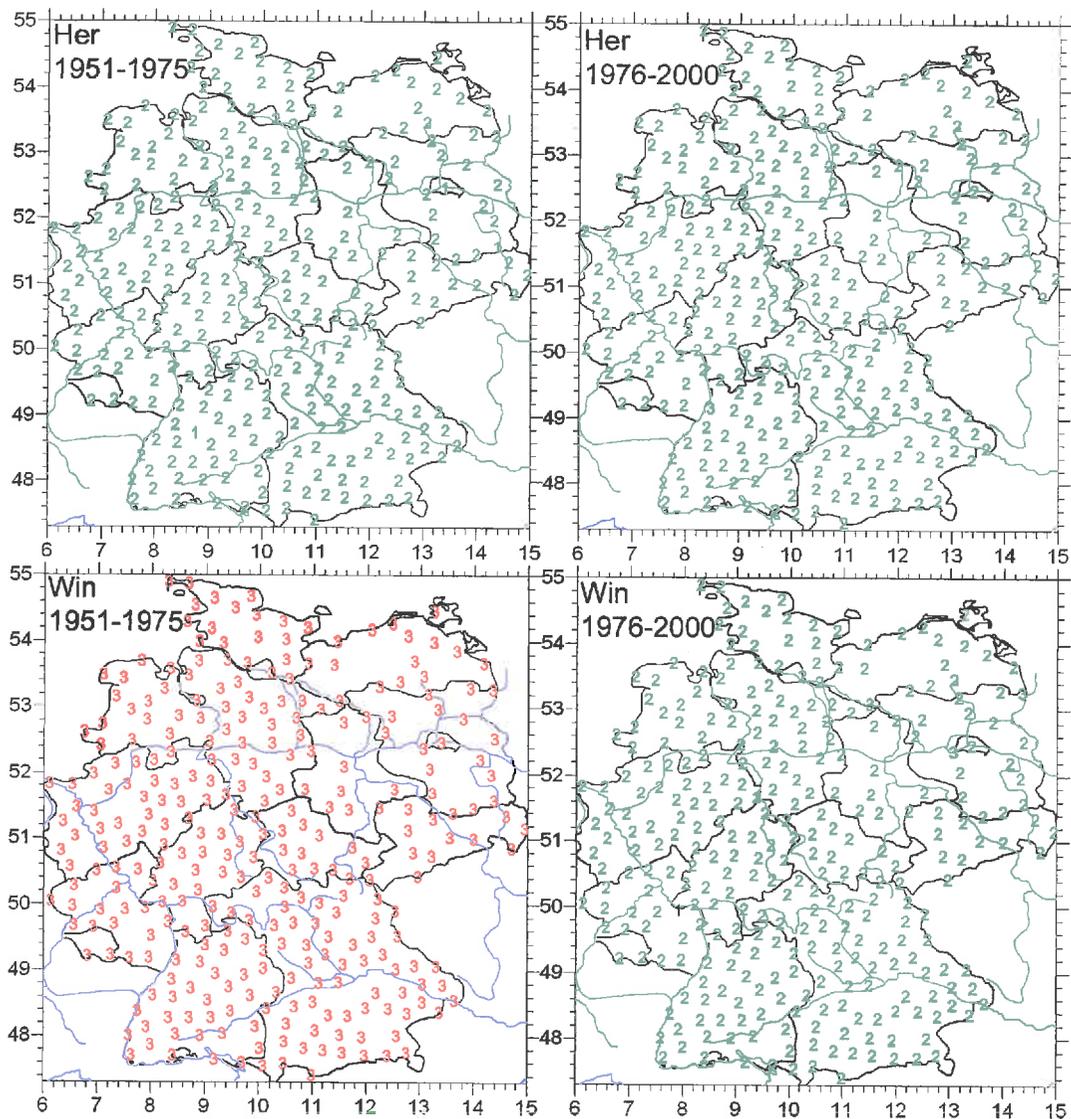


Abbildung B.19: Ordnung der Markov Ketten mit vier Zuständen nach dem AIC: Vergleich zwischen 1951-1975 und 1976-2000 für den Herbst (oben) und den Winter (unten)

B.2.2 Karten zu den Übergangswahrscheinlichkeiten zu Kap. 5.2.3

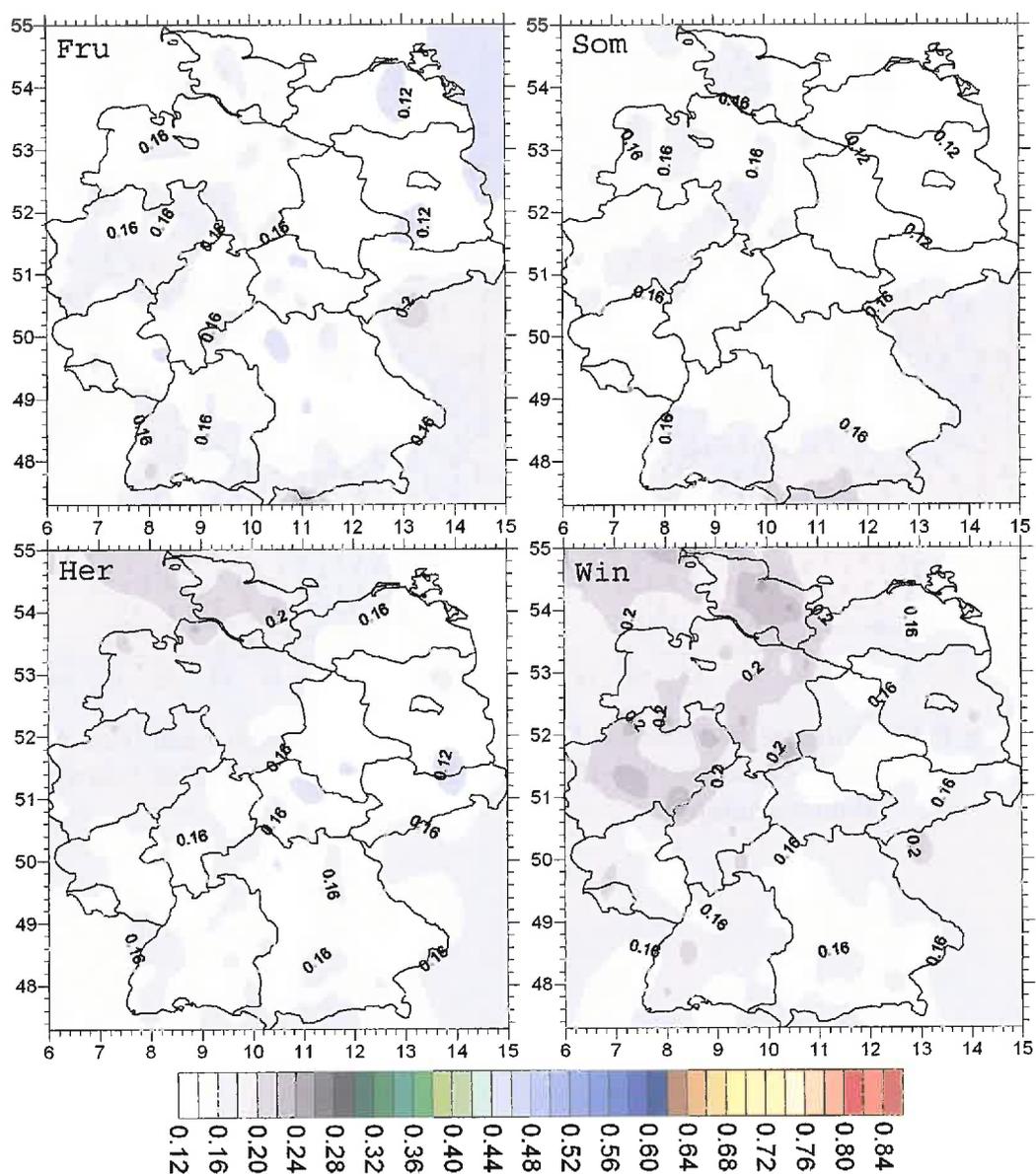


Abbildung B.20: Wahrscheinlichkeit p_1 für 1951-2000 im Frühling, Sommer, Herbst und Winter (von oben links nach unten rechts)

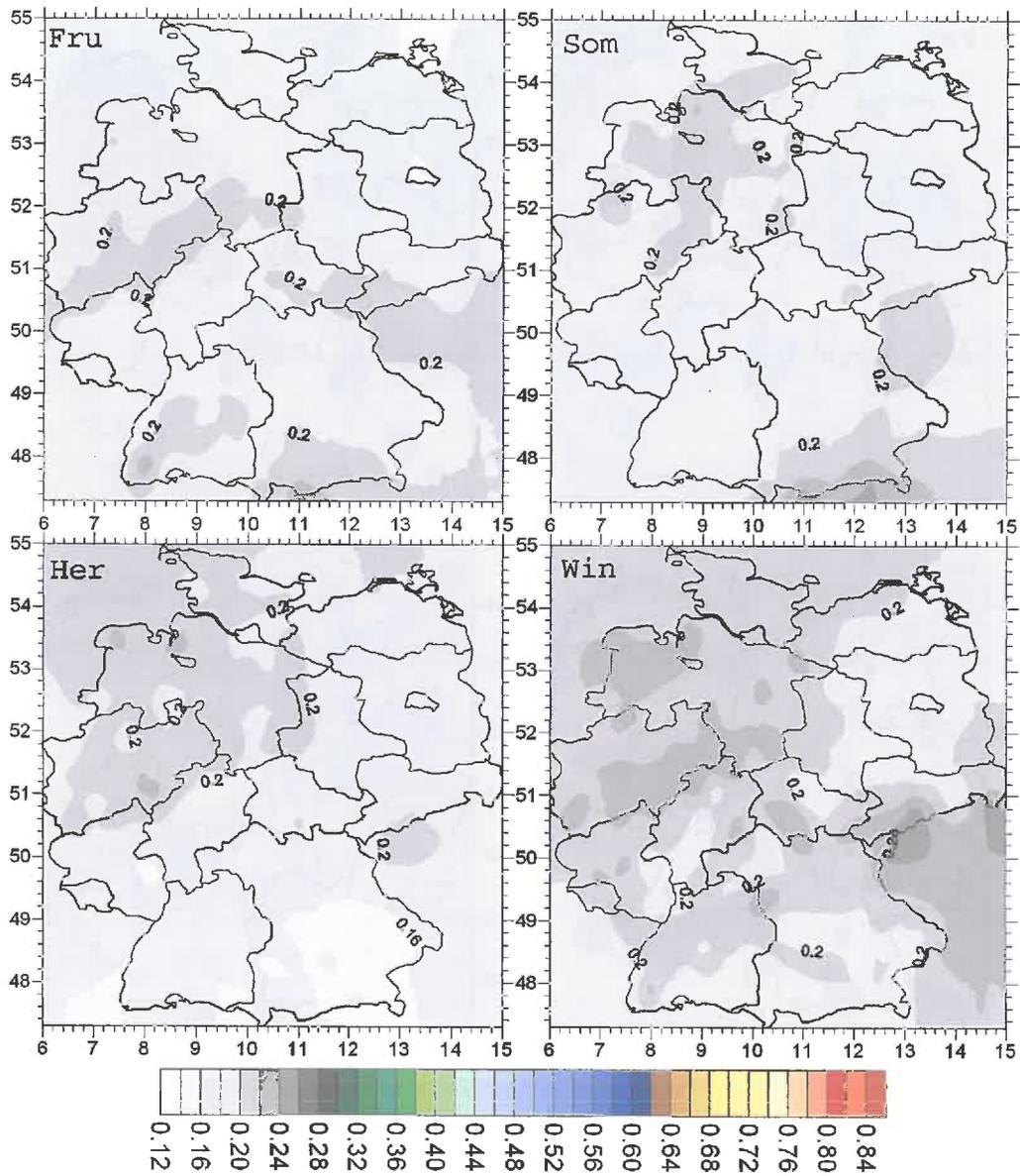


Abbildung B.21: Wahrscheinlichkeit p_2 für 1951-2000 im Frühling, Sommer, Herbst und Winter (von oben links nach unten rechts)

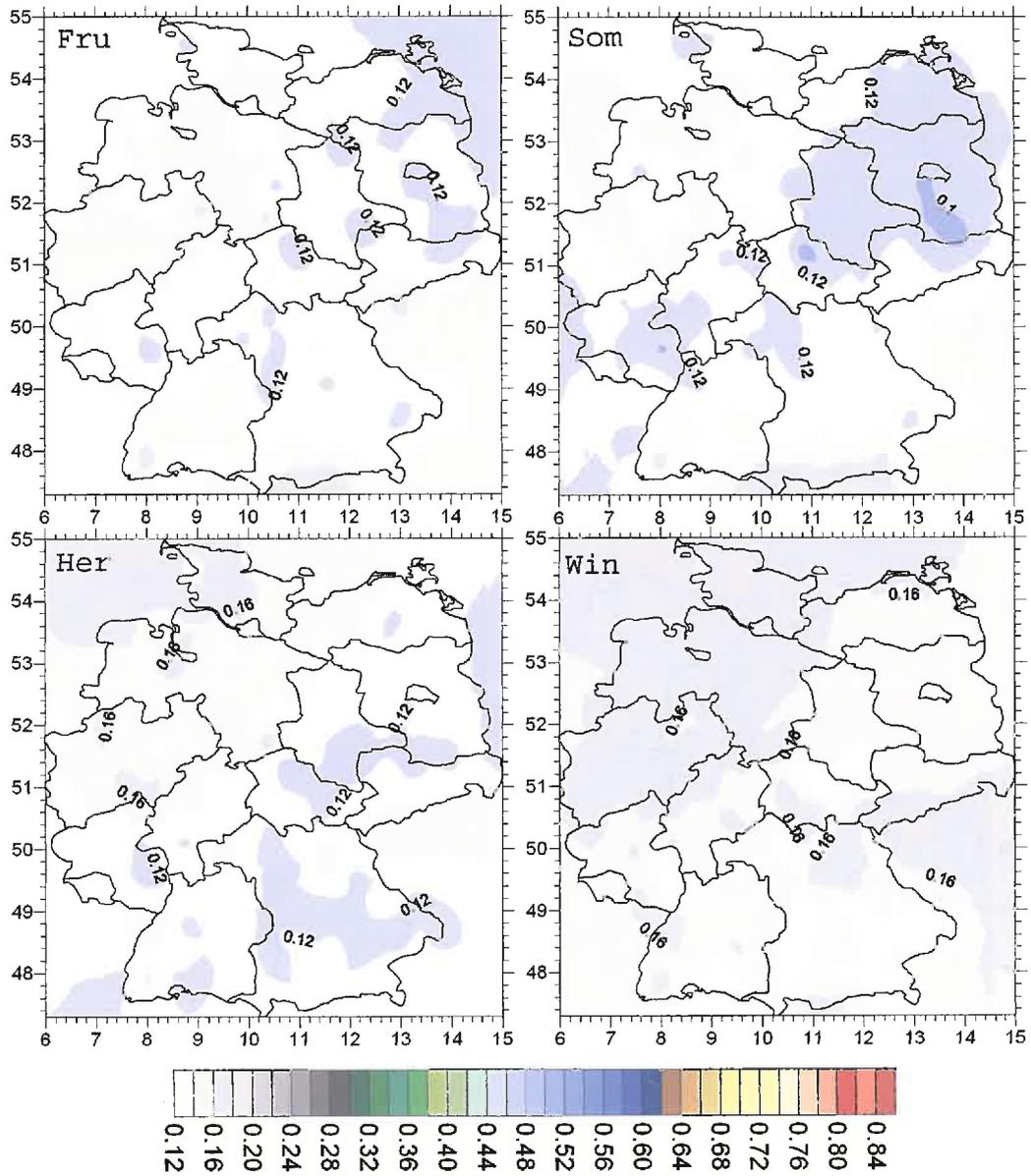


Abbildung B.22: Wahrscheinlichkeit p_3 für 1951-2000 im Frühling, Sommer, Herbst und Winter (von oben links nach unten rechts)

B.2.3 Trends der Übergangswahrscheinlichkeiten

Tabelle B3: lineare Wahrscheinlichkeitstrends in Prozent für einige Übergangswahrscheinlichkeiten getrennt nach den Jahreszeiten für Bremen; die Zahl in Klammern gibt die Signifikanz der Steigung der Regressionsgerade an

Wahrscheinlichkeit	Frühling		Sommer		Herbst		Winter	
p_1	-1.7	0	-7.3	0	-8.1	0	-10.9	0
p_2	-2	0	-2.7	0	+0.7	0	-3	0
p_3	+2.4	0	-1.8	90	+4.5	95	+6.7	90
p_{01}	-2.6	0	-9.4	99	-4.4	99	-12.3	0
p_{02}	+1.6	0	-3.4	0	+2.3	0	-0.8	0
p_{03}	-0.2	0	+1	0	+1.5	0	+3.9	0
p_{13}	-5	0	-0.7	90	+2.5	0	+4.5	0
p_{23}	+6.9	0	+2.5	0	+6.8	90	+7.4	90
p_{30}	-13.9	99	+7.6	0	+5.7	95	+3	
p_{31}	-2.8	0	-10.4	95	-1.3	0	-12.5	99
p_{32}	+0.5	0	+0.2	0	-0.1	0	-6.3	0
p_{33}	+16.2	0	+2.6	0	+10.7	0	+15.8	95
p_{003}	-1.3	0	-7.3	99	+1.5	0	+5.8	0
p_{0003}	+3.2	0	-7.9	99	+0.3	0	+2.5	0

Tabelle B4: wie Tabelle B3 jedoch für die Zugspitze

Wahrscheinlichkeit	Frühling		Sommer		Herbst		Winter	
p_1	-0.3	0	+3.7	90	-0.1	0	-2.1	99
p_2	+3.5	0	-6.3	0	+1.5	0	-2.5	0
p_3	+1.2	0	+1.5	0	+4	95	0	0
p_{01}	+3.6	0	+4.6	99	+6.4	99	-4.9	0
p_{02}	-5.1	95	-2.4	0	+0.4	0	-1.1	0
p_{03}	-0.4	0	+0.3	0	-1.5	0	+3.1	90
p_{13}	+4.4	0	+0.7	0	+7	0	+3.3	0
p_{23}	-0.7	0	+6.9	0	+5.9	90	-4.2	0
p_{30}	+0.4	90	+0.2	0	-5.6	95	+4.2	0
p_{31}	-2.4	0	+3	95	+2.8	0	+2.6	0
p_{32}	+5.5	0	-1.9	0	-4.1	0	-1.3	0
p_{33}	-3.4	0	-1.3	0	+7	0	-5.5	0
p_{003}	+3.3	0	+1.4	95	+2.6	0	+0.5	0
p_{0003}	+2.3	0	+6.4	0	+1.6	0	-2.1	0

Tabelle B5: wie Tabelle B3 jedoch für Frankfurt

Wahrscheinlichkeit	Frühling		Sommer		Herbst		Winter	
p_1	-0.3	0	-2.3	00	+0.8	0	-4.1	0
p_2	+2.7	0	-0.4	0	+2.7	0	-2.2	0
p_3	-0.2	0	-4.4	0	+0.2	0	-0.7	90
p_{01}	-0.4	0	-4.9	0	+0.7	99	-3.1	0
p_{02}	+2.3	95	+0.1	0	+3.2	90	-0.9	0
p_{03}	-0.1	0	-1.5	0	+0.1	0	+1.1	0
p_{13}	-1.8	0	+7.9	0	-2.2	0	-5.5	0
p_{23}	-0.9	0	-5.2	0	-1.4	90	+11	95
p_{30}	-11.2	90	-1.1	0	-5.6	95	-1.5	
p_{31}	+6.3	0	+6.1	0	-8.8	0	+9.2	0
p_{32}	+9.2	0	-0.4	0	+2.7	0	-3.4	0
p_{33}	-4.3	0	-4.6	0	+2.3	0	-4.3	0
p_{003}	-3.6	0	-4.8	0	+0.8	0	+1.6	0
p_{0003}	-3.8	0	-2.3	0	+1.8	0	-1.2	0

Danksagung

Ohne große Worte zu verlieren, möchte ich mich an dieser Stelle bei allen Personen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Im Besonderen danke ich Herrn Prof. Dr. Schönwiese und Frau Dr. Silke Trömel für die Vergabe und Betreuung dieser Arbeit.

Insbesondere möchte ich mich bei meinen Eltern, Kornelia und Friedrich Steiner, für die fortwährende materielle und immaterielle Unterstützung bedanken, ohne derer mir das Studium wohl kaum möglich gewesen wäre.

Erklärung

Heiko Steiner
Güldensöllerweg 47
61350 Bad Homburg

Erklärung

Hiermit erkläre ich, die vorliegende Arbeit mit dem Thema „Analyse von klimatologischen Tagesdaten mithilfe von Markov Ketten“ selbstständig und unter ausschließlicher Verwendung der angegebenen Literatur erstellt zu haben.

Bad Homburg , den 10.04.2007



Heiko Steiner