

9. EOF-Analyse und „principal-components regression“

Wie bereits mehrfach darauf hingewiesen wurde, liegt im Falle der im 7. Kapitel beschriebenen multiplen linearen Regression ein aus statistischer Sicht schwerwiegendes Problem vor. Auf Grund der großen Zahl potentieller Einflußgrößen ist es nicht auszuschließen, daß die Vorhersagemodelle von den Besonderheiten der Stichproben abhängen und somit instabil sind. Wesentlich günstiger ist es, die zugrundeliegenden Feldverteilungen des Luftdrucks mit Hilfe von nur sehr wenigen Variablen zu beschreiben, welche dann zur Bestimmung der linearen Zusammenhänge zwischen Zirkulation und Prediktanden herangezogen werden. Die neuen Variablen müssen dabei natürlich das Gros an Information des ursprünglichen Datensatzes enthalten. Diese Vorgehensweise ist als „principal-components regression (PCR)“ bekannt.

Zur Reduktion der Daten bedient man sich der „empirischen Orthogonalfunktions-Analyse (EOF-Analyse)“. Sind die Eingangsvariablen substantiell untereinander korreliert, so erhält man durch ihre Anwendung einen neuen, kleineren Datensatz, welcher dennoch einen Großteil der Variabilität des Originalen enthält. Je mehr redundante Informationen vorhanden sind, desto weniger der neuen Variablen, den sogenannten Hauptkomponenten (engl. principal components), werden zur Beschreibung benötigt. Dementsprechend hoch ist die Effektivität im Falle der EOF-Analyse klimatologischer Feldverteilungen. Hier werden neben der schieren Datenkompression zusätzlich Einblicke in die räumlichen und zeitlichen Variationen der analysierten Felder ermöglicht. Mit gewissen Einschränkungen können somit die Moden der atmosphärischen Zirkulation bestimmt werden.

Auf eine ausführliche Darstellung der theoretischen Grundlagen der EOF-Analyse kann an dieser Stelle verzichtet werden. Vor allem in den geophysikalischen Disziplinen ist die Methodik weit verbreitet und daher Gegenstand einer großen Zahl wissenschaftlicher Veröffentlichungen. Es soll lediglich dasjenige vorgestellt werden, was im Rahmen dieser Arbeit zur Anwendung kam und zum Verständnis der Untersuchungen notwendig ist. Eine umfassende Abhandlung findet man z.B. bei *Preisendorfer* (1988), kompaktere Darstellungen u.a. bei *Wilks* (1995) oder *Fechner* (1983). Welche Schwierigkeiten sich bei der physikalischen Interpretation der berechneten Strukturen ergeben können, ist bei *Dommenget und Latif* (2000) nachzulesen.

9.1 Mathematisch-statistische Grundlagen

Liegen n zeitlich aufeinanderfolgende Beobachtungen einer k Gitterpunkte bzw. Stationsmeldungen umfassenden Feldverteilung einer atmosphärischen Zustandsvariablen vor, so läßt sich diese Stichprobe in die Form eines eindimensionalen Vektors $\bar{x}(t)$ mit $t=1,2,\dots,n$ bringen:

$$\bar{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_k(t) \end{bmatrix}. \quad (9.1)$$

Jedem Beobachtungspunkt wird also eine Ziffer zugeteilt, die sich auf die entsprechende Komponente des Vektors $\mathbf{x}(t)$ bezieht. Die geographische Verteilung ist dabei für die EOF-Analyse unerheblich, lediglich die Zuordnung der Indizes ist stets einzuhalten. Ebenso wenig ist es notwendig, daß es sich um eine vollständige Feldverteilung handelt, d.h. räumliche Lücken sind ohne Bedeutung. Im Gegensatz dazu dürfen keine zeitlichen Fehlwerte vorhanden sein. Zusätzliche Voraussetzung ist, daß die Daten in Form von Anomalien vorliegen.

Zur Berechnung der Hauptkomponenten werden zunächst die Eigenvektoren entweder der Matrix aller Varianzen und Kovarianzen \mathbf{B} oder der Korrelationsmatrix \mathbf{R} , die sich durch Division sämtlicher Elemente von \mathbf{B} durch die Standardabweichungen der jeweils betroffenen Variablen (σ_{ii} und σ_{ij}) ergibt, benötigt:

$$\mathbf{B} = \begin{pmatrix} s_1^2 & s_{12} & s_{13} & \cdots & s_{1k} \\ s_{21} & s_2^2 & s_{23} & \cdots & s_{2k} \\ s_{31} & s_{32} & s_3^2 & \cdots & s_{3k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{k1} & s_{k2} & s_{k3} & \cdots & s_k^2 \end{pmatrix}. \quad (9.2)$$

Da es sich sowohl bei \mathbf{B} als auch bei \mathbf{R} um symmetrische Matrizen handelt, besitzen deren Eigenvektoren und Eigenwerte einige für die statistischen Anwendungen sehr wertvolle Eigenschaften. Allgemein gilt: Ist die Gleichung

$$\mathbf{A} \bar{\mathbf{e}} = \lambda \bar{\mathbf{e}} \quad (9.3)$$

erfüllt, so heißen λ Eigenwert und \mathbf{e} Eigenvektor der Matrix \mathbf{A} . Ist der Rang der Matrix k , so existieren auch k zusammengehörige Paare von Eigenwerten λ_l und Eigenvektoren \mathbf{e}_l , welche die Gleichung 9.3 erfüllen. Normiert man die Eigenvektoren derart, daß

$$|\bar{\mathbf{e}}_l| = 1 \quad (9.4)$$

gilt, so sind sie *orthonormal*. Neben der Einheitslänge bedeutet dies insbesondere, daß sie senkrecht aufeinander stehen. Stellt man daher die Matrix der Eigenvektoren \mathbf{E} auf, deren l . Spalte \mathbf{e}_l enthält, so läßt sich die ursprüngliche Matrix \mathbf{A} diagonalisieren:

$$\mathbf{E}^{-1} \mathbf{A} \mathbf{E} = \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{pmatrix}. \quad (9.5)$$

Dabei ist die Summe der Eigenwerte gleich der Summe aller Elemente der Diagonalen der Matrix \mathbf{A} , d.h. es gilt:

$$\sum_{l=1}^k \lambda_l = \sum_{i=1}^k a_{ii} . \quad (9.6)$$

Liegen die Eigenvektoren der Kovarianz- (\mathbf{B}) bzw. Korrelationsmatrix (\mathbf{R}) vor, so ergeben sich die Hauptkomponenten $\mathbf{h}(t)$ durch Projektion des Datenvektors $\mathbf{x}(t)$ auf die Eigenvektoren \mathbf{e}_l . In Komponentenschreibweise lautet die Beziehung daher:

$$h_l(t) = \sum_{i=1}^k e_{il} x_i(t), \quad \text{mit } l = 1, 2, \dots, k . \quad (9.7)$$

Geometrisch gesehen handelt es sich bei der EOF-Analyse um eine Drehung des Koordinatensystems im k -dimensionalen Raum. Dabei zeigt der erste Eigenvektor \mathbf{e}_1 in die Richtung der größten gemeinsamen Variabilität des analysierten Datensatzes. Der Zweite (\mathbf{e}_2) weist entsprechend in die Richtung der zweitgrößten Varianz, wobei er jedoch der Beschränkung der Orthogonalität unterliegt. Die übrigen $k-2$ Eigenvektoren folgen diesem Muster. Zusammen bilden die Eigenvektoren die empirischen Orthogonalfunktionen. Da sie ausschließlich auf statistischen Meßreihen basieren und nicht theoretisch berechnet werden können, bezeichnet man sie als empirisch bzw. natürlich. Dementsprechend erhält man für jede Stichprobe einen individuell verschiedenen Satz EOFs.

Folge der Orthogonalität der Eigenvektoren ist die Eigenschaft der Hauptkomponenten, untereinander vollständig unkorreliert zu sein. Dies drückt sich auch in Gleichung 9.5 aus, denn die Kovarianzmatrix der Hauptkomponenten ist nichts anderes als die diagonalisierte Kovarianzmatrix der Ausgangsdaten \mathbf{B} . Berücksichtigt man zusätzlich Gleichung 9.6, so läßt sich der Anteil der l . Hauptkomponente an der Gesamtvarianz leicht berechnen.

Sollen die Originaldaten aus den Hauptkomponenten zurückgewonnen werden, so ist dies ebenfalls möglich:

$$x_i(t) = \sum_{l=1}^k e_{il} u_l(t), \quad \text{mit } i = 1, 2, \dots, k . \quad (9.8)$$

Stellt man die aus Feldverteilungen resultierenden Eigenvektoren einzeln graphisch dar, so erhält man die geographische Verteilung simultaner Datenanomalien, die von den zugehörigen Hauptkomponenten repräsentiert werden. Sie werden daher auch als die unkorrelierten Moden der Variabilität bezeichnet (*Wilks, 1995*). Auf diese Weise können Telekonnektionen wie z.B. die NAO oder die PNA aufgezeigt werden. Die physikalische Interpretation ist jedoch auf Grund der Orthogonalitätsbedingung immer problematisch. Lediglich der erste Eigenvektor stellt hier zumeist eine Ausnahme dar.

9.2 Beschneidung der Hauptkomponenten

Sind die Daten, welche einer EOF-Analyse unterworfen werden, untereinander hoch korreliert, so nimmt die Varianz der Hauptkomponenten rasch ab. D.h. schon die ersten Hauptkomponenten enthalten einen Großteil der Gesamtvarianz. Es ist genau dieser Effekt, den man sich bei der Datenkompression mittels EOF-Analyse zunutze macht. Nun gilt es jedoch zu entscheiden, welcher Anteil der ursprünglich vorhandenen Information beibehalten werden soll, d.h. wie viele Hauptkomponenten zur Repräsentation verwendet werden sollen.

Hierfür existiert eine Vielzahl unterschiedlicher Auswahlregeln, welche speziell auf den geophysikalischen Bereich zugeschnitten sind. Die meisten von ihnen sind konzipiert, das Signal vom Rauschen zu trennen, d.h. die Moden der Zirkulation und die Variabilität hochfrequenter Prozesse zu differenzieren. Welche Regel zur Beschneidung der Hauptkomponenten herangezogen werden sollte, hängt von der Beschaffenheit der Daten und vom Zweck der EOF-Analyse ab. Insbesondere die Anzahl der in Folge der verschiedenen Kriterien beibehaltenen Hauptkomponenten unterscheidet sich dabei z.T. erheblich voneinander.

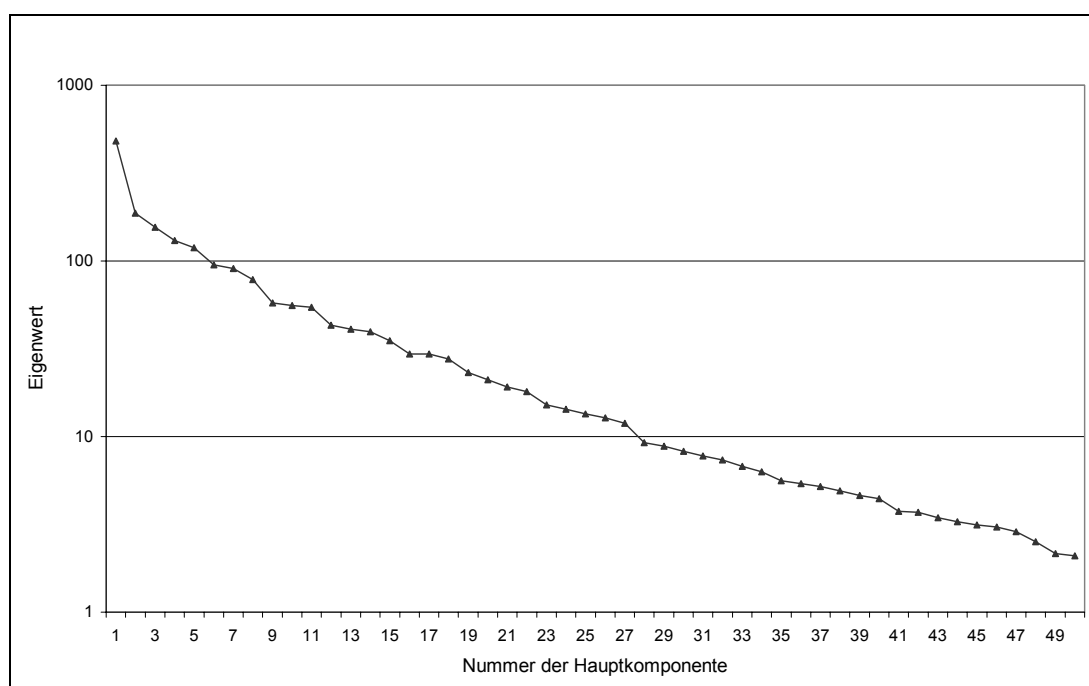


Abbildung 9.1: Logarithmische Darstellung der Magnituden der ersten 50 Eigenwerte der Korrelationsmatrix der Feldverteilungen des Bodendrucks und des 500-hPa-Geopotentials der Monate März des Zeitraums 1946-1999. Die Gitterpunkte lagen dabei in Abständen von 5° zu 5° für den Bereich 20° - 90° N vor, d.h. für die Dimension der EOF-Analyse gilt: $n=2018$.

Eine Möglichkeit zur Entscheidungsfindung bietet das sogenannte „log-eigenvalue diagram“ (LEV-Diagramm), bei dem es sich um eine logarithmische Darstellung der Eigenwerte λ_l gegen die jeweilige Nummer l der zugehörigen Hauptkomponente h_l handelt. Ein Beispiel eines solchen Diagramms ist in Abbildung 9.1 dargestellt. Hierbei handelt es sich um die Verteilung der ersten 50 Eigenwerte der Korrelationsmatrix, wie sie sich im Rahmen dieser Untersuchung für die Druckverteilung des Monats März ergeben hat (siehe dazu auch die Abschnitte 9.3 sowie 9.4.1). Die Methode basiert auf

der Annahme, daß die Varianz des Rauschanteils exponentiell abnimmt, was sich in Form einer Geraden im rechten Teil des LEV-Diagramms ausdrücken müßte. Beibehalten werden daher sämtliche Hauptkomponenten, deren Eigenwerte oberhalb der nach links extrapolierten Geraden liegen. An Hand der Abbildung 9.1 wurden hier die ersten acht Hauptkomponenten ausgewählt, wobei sich jedoch fast immer mehrere Lösungen anbieten, da die Methode subjektiv ist.

Eine objektive Entscheidungshilfe stellt u.a. die sogenannte **Kaiser's rule** dar, die sich am Durchschnitt aller Eigenwerte orientiert. Hiernach werden diejenigen Hauptkomponenten beibehalten, welche einen überdurchschnittlichen Varianzanteil repräsentieren. Gilt demnach

$$\lambda_l > \frac{1}{k} \sum_{j=1}^k \lambda_j, \quad (9.9)$$

so ist h_l beizubehalten, andernfalls fallen zu lassen. Wurde die EOF-Analyse auf der Basis der Korrelationsmatrix \mathbf{R} durchgeführt, so ergibt sich für die rechte Seite der Ungleichung 9.9 stets der Wert Eins, da die Elemente der Diagonalen r_{ii} ebenfalls alle gleich Eins sind.

9.3 Vorgehensweise

Im Gegensatz zur multiplen linearen Regression lag das Hauptaugenmerk im Falle der PCR auf der vollständigen Ausnutzung aller zur Verfügung stehenden Daten, welche den Zustand der atmosphärischen Zirkulation wiedergeben. Daher sollte neben dem Bodendruck unbedingt auch das 500-hPa-Geopotential mit in die Untersuchung einbezogen werden. Aus diesem Grund mußte auf die Verwendung eines vollständig unabhängigen Examinationskollektivs verzichtet werden, da diese Daten nur bis ins Jahr 1946 zurückreichen. Daher umfaßt die analysierte Stichprobe lediglich den 54-jährigen Zeitraum 1946-1999 (siehe auch Kapitel 2).

Für jeden Monat des Jahres wurde zunächst eine individuelle EOF-Analyse der mittleren atmosphärischen Zirkulation durchgeführt. Dabei wurden Bodendruck und 500-hPa-Geopotential in einer Matrix zusammengefaßt. Da beide Feldverteilungen für den gesamten Bereich von 20° bis 90° nördlicher Breite im Abstand von 5° zu 5° lückenlos vorlagen, haben die Matrizen die Dimension 2018 (1009 Gitterpunkte je Druckniveau). Um zu gewährleisten, daß sämtliche Eingangsvariablen gleichgewichtet berücksichtigt werden, wurden die Korrelationsmatrizen verwendet. Die Benutzung der Varianzmatrizen hätte zu einer Überdeckung der verhältnismäßig kleinen Varianzen des Bodendruckfeldes durch die wesentlich größeren Varianzen des 500-hPa-Geopotentials geführt. Ganz allgemein ist die Korrelationsmatrix stets dann vorzuziehen, wenn die Analyse für Variablen durchgeführt wird, die unterschiedliche Einheiten haben.

Zur Beschneidung der Hauptkomponenten wurden die LEV-Diagramme verwendet (siehe Abschnitt 9.2). Da die EOF-Analyse zum Zwecke der Datenkompression durchgeführt wurde, wurde das Auswahlkriterium eher stringent ausgelegt, wodurch stets die kleinstmögliche Anzahl an Hauptkomponenten beibehalten wurde. Im Falle des Monats März wäre es z.B. auch möglich gewesen, 11 oder gar 27 statt lediglich 8 Hauptkomponenten zu verwenden (vergl. Abb. 9.1). Daher sei an dieser Stelle darauf

hingewiesen, daß Detailinformationen einiger Gitterpunkte, welche möglicherweise einen prognostischen Wert für die Langfristprognose besitzen, durch die scharfe Beschneidung der Hauptkomponenten verloren gegangen sein könnten. Der Vorteil, den die drastische Verringerung der potentiellen Prediktoren mit sich bringt, ist jedoch aus statistischer Sicht höher einzuschätzen.

Die ausgewählten Hauptkomponenten dienen nun als Prediktoren linearer Regressionsanalysen der Witterung der kommenden Monate. Wie bereits zuvor wurden die Regressionen dabei für jeden Folgemonat einzeln und unabhängig voneinander durchgeführt. Der Monat, dessen Druckfelder durch die Hauptkomponenten repräsentiert wird, ist dementsprechend jeweils der Vormonat t_0 . Die maximale Vorlaufzeit wurde wieder auf elf Monate ausgeweitet (vergl. Abschnitt 7.3). Im Gegensatz zur multiplen linearen Regression der monatlichen Gitterpunktmittel wurde die PCR lediglich in Form einfacher linearer Regressionen durchgeführt, d.h. für alle Hauptkomponenten wurden die linearen Zusammenhänge mit den Prediktanden einzeln bestimmt.

Neben den grundsätzlich verwendeten monatlichen Mitteltemperaturen der Station Berlin-Dahlem diente auch die monatliche Anzahl der Tage mit meßbarem Niederschlag erneut als zusätzlicher Prediktand (siehe ebenfalls Kapitel 2). Motivation und Vorgehensweise der vollzogenen Trendbereinigung sämtlicher Zeitreihen sind in Abschnitt 4.2 erläutert.

Zur Abschätzung der Vorhersageleistung der einzelnen Regressionen wurde wiederum die Cross-Validation eingesetzt. Dabei wurde der Datensatz in jeweils 18 Jahre umfassende Kollektive unterteilt, d.h. gedrittelt. Wie schon im Falle der Cross-Validation des „Vielfach-Analog-Modells“ muß theoretisch von einer etwas zu optimistischen Schätzung der Prognosengüte ausgegangen werden. Ebenso wie die Berechnung der langjährigen Klimamittel zur Bestimmung der Anomalie-Korrelationen wurde die EOF-Analyse der Korrelationsmatrizen für jeden Kalendermonat nur ein einziges Mal durchgeführt. Dadurch wurden stets auch Informationen derjenigen Jahre mit verwendet, welche gerade dem (künstlich geschaffenen) Verifikationskollektiv zugehören. Die Folge ist der Verlust der vollständigen Unabhängigkeit dieser Jahre und somit eine Überschätzung der Vorhersagequalität (vergl. Abschnitt 8.5.1). Dieser Umstand war leider nicht zu vermeiden, da die Orthogonalisierung nur einer Matrix allein schon ca. fünf Stunden Rechenzeit in Anspruch nahm. Eine Verdreifachung wäre dementsprechend nicht mehr tragbar gewesen. Zudem trat der Effekt bei der praktischen Berechnung der Niederschlagsmengen des indischen Monsunregens basierend auf den Hauptkomponenten der Vorzirkulationen durch *Cannon und McKendry* (1999) überhaupt nicht auf – im Gegensatz zu den Untersuchungen von *Gruza und Rankova* (1993), die jedoch ein Analogverfahren verwendeten.

9.4 Ergebnisse

9.4.1 EOF-Analysen

Vor der eigentlichen Auswertung sollen zunächst die wichtigsten statistischen Eckdaten der individuellen EOF-Analysen der nordhemisphärischen Zirkulation aller Kalendermonate dargelegt werden, die bei der PCR als Basismonate t_0 dienen. Der Tabelle 9.1 können sowohl die Anzahl L^* der beibehaltenen Hauptkomponenten h_1 als auch der durch sie repräsentierte Anteil der Gesamtvarianz der analysierten Felder entnommen werden. In Anhang A sind die Feldverteilungen der zugehörigen

Eigenvektoren e_i , also die wichtigsten Moden der Variabilität, graphisch dargestellt. Es sei noch einmal daran erinnert, daß Bodendruckfeld und 500-hPa-Geopotential in jeweils einer Matrix zusammengefaßt wurden. Hierdurch wurden nicht nur die horizontalen, sondern auch die vertikalen räumlichen Zusammenhänge erfaßt, was insbesondere bei der Interpretation der geographischen Verteilung der empirischen Orthogonalfunktionen zu berücksichtigen ist. Sämtliche Vektorkomponenten, sowohl die den Bodendruckvariablen als auch die dem 500-hPa-Geopotential zugeordneten, sind als Einheit und nicht unabhängig voneinander zu betrachten.

Der Erwartung entsprechend reichen je nach Kalendermonat vier bis zehn Hauptkomponenten aus, um zwischen 37,5% und 66,7% der Gesamtvarianz zu beschreiben (siehe Tabelle 9.1). Auf Grund der hohen Interkorrelationen ist es demnach in der überwiegenden Zahl der Fälle gut möglich, die Zahl der potentiellen Prediktoren extrem zu reduzieren, ohne allzuviel Information zu vernachlässigen. Eine Ausnahme bilden hier die Monate Mai und Juni, bei denen die ausgewählten Hauptkomponenten nur 43,1% bzw. 37,5% der ursprünglichen Varianz enthalten. In diesen Monaten ist demnach die Gefahr besonders groß, daß prognostisch wertvolle Informationen verloren gegangen sind (vergl. Abschnitt 9.3).

Bemerkenswert sind vor allem die Resultate der Hochwintermonate. Im Januar und Februar entfallen 23,4% bzw. 21,9% der gesamten Variabilität allein auf die erste Mode, zusätzlich werden nur sechs bzw. vier Hauptkomponenten durch das Auswahlkriterium nicht als Rauschen identifiziert. Die atmosphärische Zirkulation wird also von nur wenigen dominanten Wetterregimen beherrscht.

| h_i | Jan | Feb | Mär | Apr | Mai | Jun | Jul | Aug | Sep | Okt | Nov | Dez |
|------------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1 | 23,4 | 21,9 | 23,8 | 16,4 | 17,0 | 15,3 | 13,1 | 14,0 | 12,6 | 13,3 | 15,8 | 16,9 |
| 2 | 12,0 | 10,8 | 9,3 | 8,2 | 7,8 | 9,1 | 10,1 | 9,2 | 8,5 | 11,1 | 10,8 | 11,1 |
| 3 | 8,2 | 8,9 | 7,7 | 7,7 | 6,9 | 6,6 | 6,5 | 7,3 | 7,3 | 7,3 | 8,2 | 9,7 |
| 4 | 7,5 | 8,4 | 6,5 | 7,0 | 5,9 | 6,5 | 6,3 | 6,1 | 6,6 | 6,0 | 7,4 | 7,6 |
| 5 | 6,4 | | 5,9 | 5,6 | 5,5 | | 5,1 | 5,4 | 5,9 | 5,8 | 5,4 | 7,2 |
| 6 | 5,0 | | 4,7 | 4,9 | | | 5,0 | 4,8 | 5,4 | 5,4 | 5,3 | 5,5 |
| 7 | | | 4,5 | 4,2 | | | 4,3 | 4,4 | 4,6 | 4,8 | | 4,7 |
| 8 | | | 3,9 | 4,1 | | | 3,9 | 4,0 | | | | 4,0 |
| 9 | | | | 3,9 | | | 3,6 | 3,8 | | | | |
| 10 | | | | | | | 3,3 | | | | | |
| $\Sigma 1$ | 62,5 | 50,0 | 66,3 | 62,0 | 43,1 | 37,5 | 61,2 | 59,0 | 50,9 | 53,7 | 52,9 | 66,7 |

Tabelle 9.1: Durch die 1. Hauptkomponente h_1 repräsentierter Anteil der Gesamtvarianz der Feldverteilungen des Bodendrucks und des 500-hPa-Geopotentials für die verschiedenen Kalendermonate im Zeitraum 1946-1999 in Prozent.

9.4.2 PCR

In Tabelle 9.2 sind zunächst wieder die für den Untersuchungszeitraum bestimmten RV-Werte der Temperaturprognosen in Abhängigkeit des Vormonats t_0 und der Vorlaufzeit aufgeführt (obere Zeile). Dargestellt ist jeweils die abgeschätzte Vorhersageleistung des „besten“ Prediktors, d.h. der Hauptkomponente, deren simple lineare Regression den höchsten RV-Wert erzielen konnte. Bei der Beurteilung ist daher trotz der durchgeführten Cross-Validation auch hier zu beachten, daß es sich um das

beste Resultat mehrerer Möglichkeiten handelt. Die Auswirkungen dieses Umstandes sollten jedoch eher gering sein, da lediglich zwischen vier und zehn Prediktoren getestet wurden, so daß der Spielraum für Zufälle klein sein dürfte. In der unteren Zeile ist zusätzlich jeweils die Nummer 1 der betreffenden Hauptkomponente eingetragen.

| t ₀ | t+1 | t+2 | t+3 | t+4 | t+5 | t+6 | t+7 | t+8 | t+9 | t+10 | t+11 |
|----------------|------------------|------------------|------------------|-------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Januar | 0,02 1 | -0,01 2 | -0,05 1 | -0,06 1 | 0,05 4 | -0,06 3 | -0,04 1 | -0,01 1 | -0,04 2 | 0,04 1 | 0,00 1 |
| Februar | 0,13 1 | -0,13 2 | -0,06 4 | 0,00 2 | -0,07 2 | -0,02 3 | -0,09 1 | -0,01 3 | 0,04 4 | 0,07 4 | -0,13 1 |
| März | -0,11 7 | 0,06 1 | 0,04 2 | -0,04 1 | -0,02 7 | -0,06 8 | 0,10 6 | 0,07 6 | 0,04 7 | 0,02 8 | 0,05 3 |
| April | 0,03 9 | 0,01 1 | 0,20 9 | 0,06 4 | -0,01 9 | 0,05 6 | 0,01 7 | -0,01 6 | 0,02 6 | 0,02 1 | 0,00 7 |
| Mai | 0,05 3 | -0,03 5 | -0,04 4 | -0,04 3 | -0,04 2 | 0,02 3 | -0,04 3 | 0,03 3 | 0,03 2 | 0,05 3 | -0,16 4 |
| Juni | -0,06 2 | -0,07 2 | -0,06 1 | 0,00 4 | 0,04 4 | 0,00 1 | 0,08 2 | 0,00 4 | 0,02 4 | -0,10 3 | -0,03 1 |
| Juli | -0,02 8 | 0,16 4 | 0,00 5 | 0,01 10 | 0,00 8 | 0,07 9 | 0,00 6 | -0,01 3 | -0,09 7 | 0,01 3 | 0,05 6 |
| August | -0,03 8 | 0,04 5 | 0,03 2 | -0,13 3 | 0,02 6 | 0,01 4 | -0,01 8 | -0,17 2 | 0,01 5 | 0,02 6 | -0,01 9 |
| September | 0,08 3 | 0,08 6 | 0,08 4 | 0,04 7 | -0,03 2 | 0,05 5 | -0,15 7 | -0,01 3 | 0,07 2 | -0,02 2 | 0,07 7 |
| Oktober | -0,01 3 | 0,03 3 | 0,01 3 | 0,01 3 | 0,04 2 | -0,13 1 | -0,05 4 | 0,06 5 | -0,01 6 | 0,07 5 | -0,02 4 |
| November | -0,02 2 | 0,00 3 | -0,02 4 | -0,04 3 | -0,17 6 | -0,03 3 | 0,05 2 | -0,01 4 | -0,01 4 | -0,01 3 | -0,03 6 |
| Dezember | -0,03 4 | 0,01 4 | 0,03 2 | -0,12 5 | 0,01 4 | -0,01 3 | -0,05 2 | -0,03 6 | 0,05 8 | 0,03 6 | 0,15 1 |

Tabelle 9.2: RV-Werte der durch PCR erstellten Temperaturprognosen im Zeitraum 1946-1999 (obere Zeile) sowie Nummer 1 der verwendeten Hauptkomponente (untere Zeile). Positive Vorhersageleistung ist durch Fettdruck gekennzeichnet.

Ließen die Resultate aller bislang getesteten Vorhersagemodelle, für welche kein gesondertes Examinationskollektiv verwendet werden konnte (Analogverfahren), stets kaum Zweifel daran, daß eine Reduktion der Varianz durch langfristige Temperaturvorhersagen höchstens im Einzelfalle, d.h. für wenige spezielle Monatskombinationen, möglich ist, so gilt dies für die PCR nicht. Statt dessen ist eine Einschätzung äußerst schwierig, da die vorliegenden Ergebnisse keine eindeutige Schlußfolgerung zulassen.

Einerseits hat die multiple lineare Regression gezeigt, daß selbst im Mittel noch höhere RV-Werte im Entwicklungskollektiv keine Garantie für eine reale Vorhersageleistung an vollständig unabhängigen Daten darstellen (vergl. Abschnitt 7.5).

Andererseits wurde hier nicht die U-Methode angewandt, sondern der Datensatz wurde gedrittelt. Im Gegensatz zur multiplen linearen Regression läßt sich daher eine durch die Form der Cross-Validation bedingte Überschätzung der Vorhersagequalität praktisch

ausschließen (vergl. Abschnitt 4.1). Zusätzlich liegt die Zahl der potentiellen Prediktoren in einem völlig anderen Bereich. Daher ist es fraglich, ob es überhaupt zu einem mit dem Einbruch bei der mehrfachen Regression vergleichbaren Abfall der Vorhersageleistung käme.

Orientiert man sich daher an den Ergebnissen derjenigen Modelle, bei denen das Phänomen der „multiplicity“ keinen Einfluß hat, d.h. an den Resultaten der Analogverfahren, so ist festzustellen, daß sich die PCR überlegen präsentiert. Die Anzahl positiver und negativer RV-Werte hält sich nahezu exakt die Waage, und der Betrag des höchsten RV-Wertes liegt mit 0,20 bei der 3-Monats-Prognose des Juli über dem des niedrigsten (-0,17 sowohl bei der 5- als auch bei der 8-Monats-Prognose des April). Beides sind Resultate, die selbst vom vielversprechendsten Analogverfahren klar verfehlt werden. Kritisch stimmt lediglich, daß weder die verschiedenen Vorlaufzeiten noch die einzelnen Vor- bzw. Zielmonate eine unterschiedliche Vorhersagegüte aufweisen. Ganz gleich, welcher Zielmonat vorhergesagt wird, welcher Monat als Basismonat t_0 dient und welcher Zeitraum zwischen ihnen liegt, eine dauerhafte Vorhersageleistung ist stets etwa gleich wahrscheinlich. Dennoch erscheint es in Anbetracht aller Umstände schwer vorstellbar, daß es sich bei sämtlichen positiven RV-Werten um bloße Zufälle handelt. Für die in Tabelle 9.3 aufgeführten Ergebnisse der Niederschlagsprognosen gilt grundsätzlich Ähnliches.

| t_0 | t+1 | t+2 | t+3 | t+4 | t+5 | t+6 | t+7 | t+8 | t+9 | t+10 | t+11 |
|-----------|------------------|------------------|------------------|------------------|------------------|-------------------|------------------|------------------|------------------|------------------|------------------|
| Januar | 0,00 3 | -0,04 2 | 0,09 1 | 0,00 5 | 0,00 5 | -0,01 2 | -0,05 2 | -0,01 3 | 0,01 6 | 0,02 4 | -0,07 1 |
| Februar | 0,03 1 | -0,05 4 | 0,00 2 | 0,03 3 | -0,01 3 | 0,01 2 | 0,05 1 | 0,02 3 | 0,03 1 | -0,02 2 | -0,01 2 |
| März | -0,02 6 | 0,02 1 | 0,02 2 | 0,03 1 | 0,03 7 | 0,01 2 | 0,00 6 | -0,05 2 | -0,06 2 | 0,00 3 | -0,01 8 |
| April | 0,03 3 | 0,03 2 | 0,05 1 | -0,02 5 | 0,02 9 | 0,05 2 | -0,03 1 | 0,02 8 | -0,08 4 | 0,08 9 | -0,02 1 |
| Mai | 0,08 3 | -0,01 2 | -0,07 2 | -0,02 2 | 0,03 3 | -0,06 2 | -0,04 2 | 0,04 4 | -0,06 5 | -0,04 3 | 0,01 4 |
| Juni | 0,00 4 | -0,01 2 | 0,10 3 | -0,06 2 | -0,02 1 | -0,08 3 | 0,03 3 | -0,05 1 | 0,07 4 | 0,03 3 | -0,03 2 |
| Juli | 0,00 2 | 0,04 7 | 0,08 6 | -0,03 10 | -0,06 10 | 0,04 10 | -0,04 6 | 0,00 1 | 0,01 1 | 0,00 1 | 0,01 9 |
| August | -0,01 8 | 0,03 9 | 0,03 1 | -0,08 9 | 0,03 4 | 0,03 9 | 0,12 1 | 0,01 1 | 0,01 4 | 0,00 8 | 0,00 6 |
| September | -0,01 7 | -0,13 3 | -0,02 4 | 0,03 6 | -0,01 3 | -0,02 4 | 0,02 7 | 0,04 3 | 0,03 5 | -0,01 3 | -0,04 3 |
| Oktober | 0,02 2 | -0,04 4 | -0,25 3 | 0,07 3 | -0,02 5 | 0,00 1 | -0,08 6 | 0,05 1 | 0,00 2 | -0,07 1 | 0,08 2 |
| November | -0,09 4 | 0,05 2 | 0,02 1 | 0,01 1 | -0,01 1 | 0,00 2 | 0,05 5 | 0,00 3 | 0,02 2 | 0,02 1 | 0,02 1 |
| Dezember | 0,04 8 | -0,03 6 | 0,01 2 | 0,01 4 | 0,09 6 | 0,01 3 | 0,01 7 | 0,11 8 | 0,07 8 | -0,03 8 | 0,04 7 |

Tabelle 9.3: RV-Werte der durch PCR erstellten Niederschlagsprognosen im Zeitraum 1946-1999 (obere Zeile) sowie Nummer 1 der verwendeten Hauptkomponente (untere Zeile). Positive Vorhersageleistung ist durch Fettdruck gekennzeichnet.

Da die sicherste Methode zur Stabilitätsprüfung eines Vorhersagemodells, die Verwendung eines Examinationskollektivs, nicht möglich war, mußte hier ein anderer Weg beschritten werden. Es wurde ein Algorithmus erstellt, der die Erwartungstreue sämtlicher Prediktoren nach strengsten Maßstäben auswertet, um so diejenigen zu isolieren, welche mit hoher Sicherheit stabile Regressionen garantieren. Dabei wurden sowohl das Kriterium der Regularität als auch das Kriterium der Unverzerrtheit berücksichtigt. Auch dieser Schritt wurde für sämtliche Monatskombinationen einzeln und unabhängig voneinander durchgeführt.

Zunächst wurde für jedes der $j=1, \dots, 3$ künstlichen Verifikationskollektive eine gesonderte Rangfolge der jeweils ausgewählten Hauptkomponenten an Hand der von ihnen erzielten Vorhersageleistung im betreffenden Zeitraum festgelegt. D.h. diejenige Hauptkomponente, deren Vorhersagen während der ersten 18 Jahre (1946-1963) den höchsten RV-Wert liefern, wird im Zeitraum $j=1$ auf den ersten Rang gesetzt, diejenige, deren Vorhersagen den zweithöchsten RV liefern, auf den zweiten Rang usw. Wie bei der Cross-Validation üblich, wurden diese Modelle an den verbliebenen 36 Jahren des Untersuchungszeitraumes (1964-1999) entwickelt. Der Index $j=2$ bezieht sich dementsprechend auf die mittleren 18 Jahre (1964-1981) als Verifikationszeitraum, der Index $j=3$ auf die letzten 18 Jahre (1982-1999). Anschließend wurden die Ergebnisse der drei Unterkollektive $j=1, \dots, 3$ zu einer Gesamtrangfolge zusammengefaßt. Hierfür wurde die Summe der drei Ränge $j=1, \dots, 3$ gebildet. Diejenige Hauptkomponente, für welche sich die kleinste Summe ergab, wurde dann auf den ersten Rang der Gesamtrangfolge gesetzt usw.

Als erwartungstreu wurden nur solche Hauptkomponenten eingestuft, welche die folgenden Kriterien erfüllen, wobei die Anzahl der ausgewählten Prediktoren innerhalb der einzelnen Monatskombinationen keiner Begrenzung unterlag:

- Für die Summe der $j=1, \dots, 3$ Ränge der Hauptkomponente muß die Ungleichung

$$\sum j \leq 3(i-1) + 5 + x \quad (9.10)$$

erfüllt sein, worin i den Gesamtrang der betreffenden Hauptkomponente angibt. Die Variable x sorgt dabei für einen Ausgleich der unterschiedlichen Anzahl beibehaltener Hauptkomponenten. Wurden mehr als sieben ausgewählt, so gilt $x=1$, ansonsten $x=0$.

- Der Mittelwert der in den drei Zeiträumen $j=1, \dots, 3$ erzielten RV-Werte darf den Wert 0,035 nicht unterschreiten, wenn $i=1$ gilt. Zusätzlich darf keiner der Einzelwerte kleiner Null sein. Dies gilt für alle i .
- Das Verhältnis der $j=1, \dots, 3$ zugehörigen Steigungen der Regressionsmodelle darf in keinem Fall den Wert 0,700 (1,429) unterschreiten (überschreiten).

Zwar wurden diese Kriterien in gewisser Weise willkürlich erstellt, sie sollten jedoch dennoch geeignet sein, ihren Zweck zu erfüllen, da sie genau auf die bei der Auffindung stabiler Modelle entscheidenden Aspekte abzielen. Während die Ungleichung 9.10 und der zweite Punkt bewirken, daß nur Prediktoren ausgewählt werden, welche zu jeder Zeit verlässliche und überdurchschnittlich genaue Vorhersagen liefern, stellt der letzte Punkt die Ähnlichkeit der Regressionskoeffizienten der einzelnen Modelle sicher.

In den Tabellen 9.4 und 9.5 sind die Ergebnisse der Stabilitätsprüfung für die Temperatur bzw. den Niederschlag festgehalten. Aufgeführt sind die Nummern 1 derjenigen Hauptkomponenten h_1 , die als erwartungstreu identifiziert wurden. In beiden Fällen erwiesen sich die Verhältnisse längst nicht immer als „eindeutig stabil“, wenn für die betreffende Monatskombination ein positiver RV-Wert für den gesamten Zeitraum (Tabellen 9.2 und 9.3) vorlag. Dies ist ein Anhaltspunkt dafür, daß es sich zumindest in einigen Fällen um Scheingüte, d.h. um eine zufällig gute Anpassung handelt. Andererseits sind insbesondere im Falle der Temperatur derart viele Monatskombinationen mit mindestens einem (18 Fälle) z.T. sogar zwei (2 Fälle) erwartungstreuen Prediktoren vorhanden (Niederschlag 10 bzw. 1), daß die Existenz eines realen Vorhersagepotentials nach wie vor nicht auszuschließen ist.

| t_0 | t+1 | t+2 | t+3 | t+4 | t+5 | t+6 | t+7 | t+8 | t+9 | t+10 | t+11 |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|
| Januar | - | - | - | - | 4 | - | - | - | - | 1 | - |
| Februar | - | - | - | - | - | - | - | - | 4,1 | 4 | - |
| März | - | - | - | - | - | - | - | 6 | 7 | - | - |
| April | - | - | 9 | 4 | - | 6 | - | - | - | - | - |
| Mai | - | - | - | - | - | - | - | - | - | 3 | - |
| Juni | - | - | - | - | 4 | - | - | - | - | - | - |
| Juli | - | 4 | - | - | - | - | - | - | - | - | - |
| August | - | - | - | - | - | - | - | - | - | - | - |
| September | 3 | - | 4 | - | - | - | - | - | 2 | - | 7 |
| Oktober | - | - | - | - | - | - | - | 5 | - | - | - |
| November | - | - | - | - | - | - | - | - | - | - | - |
| Dezember | - | - | - | - | - | - | - | - | - | - | 1,4 |

Tabelle 9.4: Nummern 1 der Hauptkomponenten h_1 , welche mit hoher Wahrscheinlichkeit stabile Temperaturprognosen mittels PCR liefern.

| t_0 | t+1 | t+2 | t+3 | t+4 | t+5 | t+6 | t+7 | t+8 | t+9 | t+10 | t+11 |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|
| Januar | - | - | 1 | - | - | - | - | - | - | - | - |
| Februar | - | - | - | - | - | - | - | - | - | - | - |
| März | - | - | - | - | - | - | - | - | - | - | - |
| April | - | - | - | - | - | 2 | - | - | - | 9 | - |
| Mai | - | - | - | - | - | - | - | - | - | - | - |
| Juni | - | - | - | - | - | - | - | - | - | - | - |
| Juli | - | 7 | - | - | - | - | - | - | - | - | - |
| August | - | - | - | - | - | - | 1 | - | - | - | - |
| September | - | - | - | - | - | - | - | 3,7 | - | - | - |
| Oktober | - | - | - | - | - | - | - | - | - | - | 2 |
| November | - | - | - | - | - | - | 5 | - | - | - | - |
| Dezember | - | - | - | - | 6 | - | - | 8 | - | - | - |

Tabelle 9.5: Wie Tabelle 9.4, jedoch für den Niederschlag als Prediktand.

In einem letzten Schritt wurde eine sich von der Cross-Validation unterscheidende „resampling“-Technik angewendet, um Aussagen zur Signifikanz der ermittelten Vorhersageleistung der PCR zu ermöglichen. Es handelt sich um eine weitverbreitete Methode zur Erzeugung künstlicher Vertrauensgrenzen, die als „bootstrap“ bezeichnet

wird (vergl. z.B. *Wilks*, 1995). Sie wird immer dann verwendet, wenn für die Prüfgröße keine theoretische Wahrscheinlichkeitsdichteverteilung existiert, d.h. die Anwendung eines klassischen Hypothesentests nicht möglich ist. Da die Methodik parameterfrei ist und lediglich das Vorliegen einer großen Menge gleichverteilter Zufallszahlen voraussetzt, sind ihrer Anwendungsvielfalt praktisch keine Grenzen gesetzt.

Das Grundprinzip ist einfach: Mit Hilfe der Zufallszahlen wird aus den Originaldaten eine Vielzahl künstlicher Stichproben gleichen Stichprobenumfanges erzeugt, für welche die Prüfgröße jeweils unabhängig von den anderen ermittelt wird. Dadurch erhält man eine empirische Verteilung der Prüfgröße, welche die theoretische Wahrscheinlichkeitsdichteverteilung ersetzt. Die Bestimmung von Vertrauensgrenzen geschieht dann in Analogie zu klassischen Hypothesentests, d.h. es werden die Regionen der Verteilung festgelegt, die als ausreichend unwahrscheinlich angesehen werden. Hierzu bietet sich die Verwendung von Quantilen an.

Bei der Erzeugung der künstlichen Stichproben ist es wichtig, daß diese eine unterschiedliche Zusammensetzung aufweisen. Dies ist nur möglich, wenn jede einzelne Beobachtung der Originaldaten den individuellen künstlichen Stichproben unbegrenzt häufig zugeordnet werden kann. Vergleicht man die Situation mit einer Urne, aus welcher man Kugeln zieht, so ist dies derart zu verstehen, daß die soeben gezogene Kugel stets wieder in die Urne zurückgelegt wird, bevor eine weitere entnommen wird. Auf diese Weise ist es möglich, daß eine Beobachtung der Originaldaten einer speziellen künstlichen Stichprobe mehrfach angehören kann, wodurch andere Beobachtungen natürlich gänzlich fehlen. Diese Vorgehensweise stellt sicher, daß sich der Wert der Prüfgröße von Stichprobe zu Stichprobe unterscheidet. Eine in der Literatur typische Größenordnung für die Anzahl künstlicher Stichproben stellen 1000 Realisierungen dar.

Um nun die Signifikanz der Vorhersageleistung der PCR zu testen, wurde die Berechnung sämtlicher Regressionen für alle 132 Monatskombinationen 1000fach wiederholt. Dabei fungierten jedoch nicht die Berliner Monatsmitteltemperaturen als Prediktanden, sondern mittels „bootstrap“ erzeugte künstliche Neuankordnungen dieser Temperaturreihen. Als Gegenstand des Signifikanztests, sozusagen als Prüfgröße, wurde die Anzahl verschiedener Monatskombinationen, für die unterschiedlich viele erwartungstreue Hauptkomponenten ausfindig gemacht werden konnten (s.o.), ausgewählt. Die zugehörige Nullhypothese lautet daher: Die Anzahl der Monatskombinationen mit erwartungstreuen Prediktoren, die sich für den Prediktand Temperatur bzw. Niederschlag ergeben hat, ist nicht überzufällig. Um die empirische Verteilung der „Prüfgröße“ zu erhalten, wurden die Ergebnisse aller 1000 Realisierungen nach denselben Stabilitätskriterien ausgewertet, welche auch bei der Temperatur und dem Niederschlag verwendet wurden (s.o.).

Die hier gewählte Vorgehensweise ist insofern ungewöhnlich, daß üblicherweise die Prediktoren durch zufällig angeordnete Stichproben bzw. durch Zufallszahlen ersetzt oder ergänzt werden. Es ist jedoch nicht von Bedeutung, ob die Einfluß- oder aber die Zielgröße bzw. sogar beide Reihen aus Zufallszahlen zusammengesetzt sind. Eine Kausalität besteht in keinem Fall. Dies hat zur Konsequenz, daß ein aus realen Beobachtungen bestehender Prediktor für den Prediktand Zufall ebenfalls zur Zufallsfolge wird.

In Abbildung 9.2 sind die für die realen Prediktanden Temperatur und Niederschlag ermittelten Ergebnisse im Vergleich zu ausgesuchten Quantilen der durch „bootstrap“ erzeugten empirischen Verteilung graphisch veranschaulicht. Dargestellt ist die Anzahl derjenigen Monatskombinationen (von insgesamt 132), für die mindestens ein, zwei oder drei erwartungstreue Hauptkomponenten ausfindig gemacht wurden. So fanden

sich z.B. bei zehn der künstlichen Realisierungen drei oder mehr Monatskombinationen, bei der mindestens zwei erwartungstreue Hauptkomponenten existierten (0,99stes Quantil). Genauso fanden sich bei 50 Realisierungen 14 oder mehr Monatskombinationen, bei welcher mindestens eine Hauptkomponente stabile Prognosen liefern konnte (0,95stes Quantil).

Wie sich zeigt, ist die zuvor formulierte Nullhypothese im Falle der Temperatur mit einer Irrtumswahrscheinlichkeit von 2% (zweiseitiger Signifikanztest) zu verwerfen, nimmt man die Anzahl der Monatskombinationen, für die zumindest eine stabile Regression gefunden wurde, als Maßstab. Bei zwei Hauptkomponenten liegt die Irrtumswahrscheinlichkeit bei 10%, darüber läßt sich keine sinnvolle Aussage mehr machen. Es ist also mit hoher statistischer Sicherheit davon auszugehen, daß ein reales – wenngleich geringes – Vorhersagepotential für langfristige Temperaturvorhersagen existiert, welches mittels PCR genutzt werden kann.

Im Falle des Niederschlags werden die üblichen Signifikanzschwellen hingegen deutlich verfehlt. Die Anzahl der Monatskombinationen mit mindestens einem erwartungstreuen Prediktor liegt mit zehn Fällen nur wenig über dem Median der empirischen Verteilung und ist daher keinesfalls als überzufällig anzusehen. Demnach ist es nicht auszuschließen, daß die Ergebnisse der Niederschlagsprognosen rein zufälliger Natur sind und ein Vorhersagepotential nicht vorhanden ist.

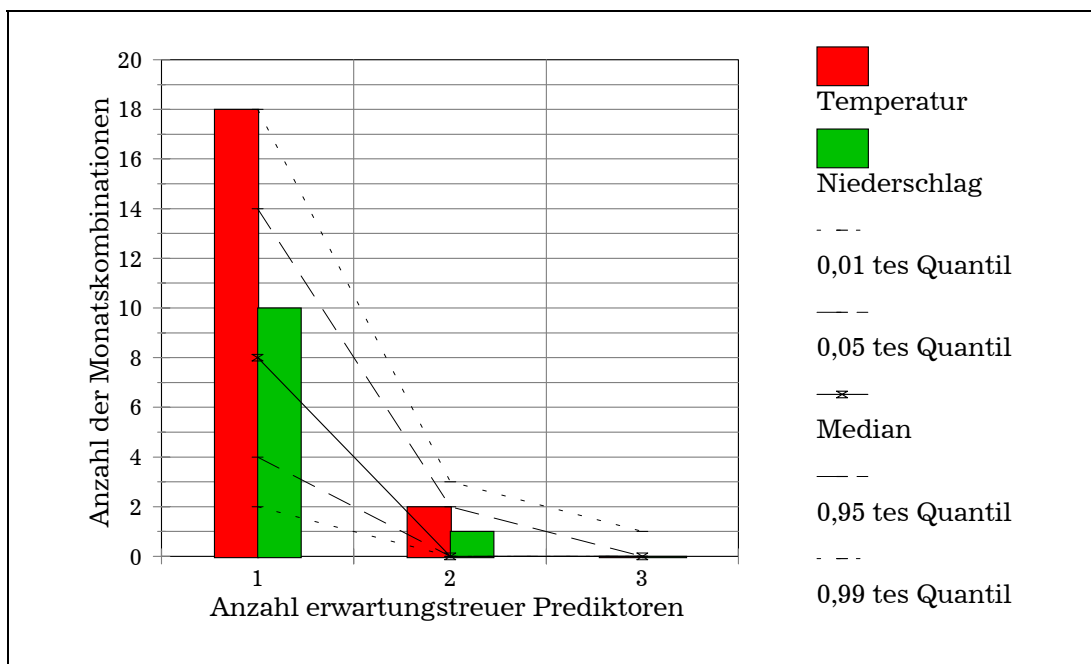


Abbildung 9.2: Anzahl der Monatskombinationen mit mindestens einem, zwei bzw. drei erwartungstreuen Prediktoren (Hauptkomponenten) für die Prediktanden Temperatur bzw. Niederschlag im Vergleich zu ausgewählten Quantilen einer durch „bootstrap“ erzeugten empirischen Verteilung. Zur Definition eines erwartungstreuen Prediktors siehe Text.

Im Falle der Temperatur ist nun die zu erwartende Vorhersagequalität im dauerhaften operationellen Betrieb (vergl. Abschnitt 7.5 sowie 8.2.2) von besonderem Interesse. In Abbildung 9.3 sind die im gesamten Untersuchungszeitraum erzielten RV-Werte der 1-, 2- und 3-Monats-Prognosen sowie der jeweils „besten“ Prognose für die verschiedenen Jahreszeiten dargestellt. Für die Schätzungen der „besten“ Prognosen gilt hier dasselbe wie in Abschnitt 8.2.2, d.h. sie dürften nennenswert zu optimistisch sein. Zusätzlich gilt

es, den an vollständig unabhängigen Daten stets zu verzeichnenden Abfall der Vorhersageleistung zu berücksichtigen. Dennoch sollte zumindest im Sommer und Herbst eine Reduktion der Varianz um bis zu 10% möglich sein. Ob auch die „beste“ Winterprognose (5,1% RV) bzw. die 1-Monats-Prognose im Frühling (3,5%) in einem Examinationskollektiv noch eine positive Vorhersageleistung erzielen würden, ist dagegen fraglich. Hier steht zu befürchten, daß die „Macht des Zufalls“ (Balzer, 1989) für die positiven Ergebnisse im Untersuchungszeitraum verantwortlich ist. Zudem ist auch im Falle der PCR weder ein kontinuierlicher Anstieg der Vorhersageleistung mit abnehmender Vorlaufzeit noch eine zu allen Jahreszeiten einheitlich besonders günstige Vorlaufzeit zu erkennen. Diese Tatsachen sind kritisch zu betrachten und dämpfen daher den Optimismus hinsichtlich der maximal möglichen Vorhersageleistung.

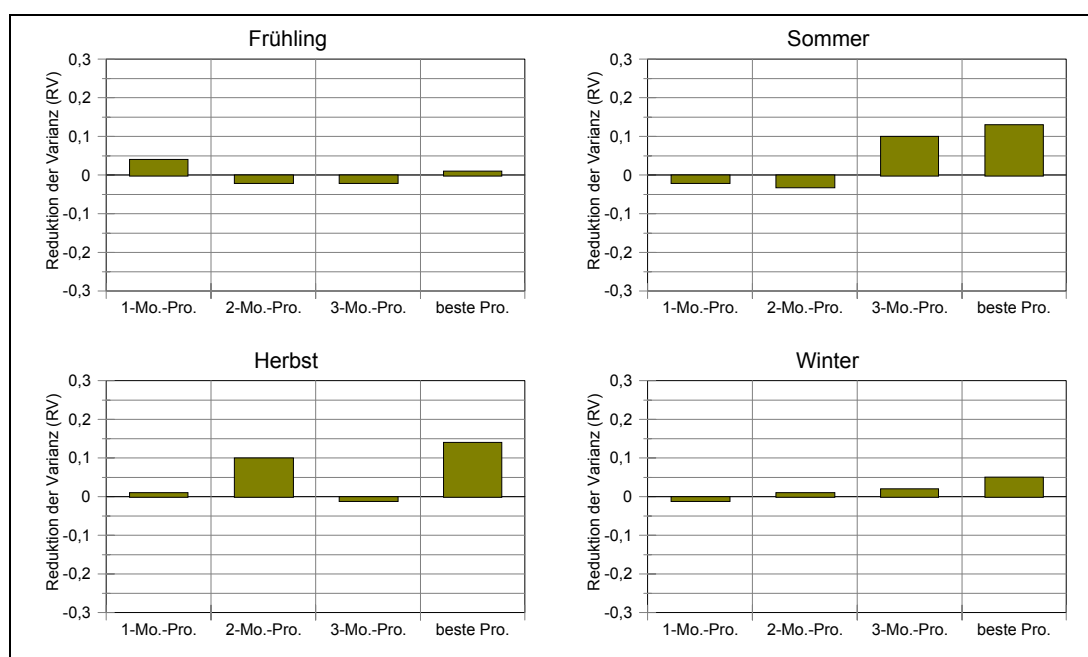


Abbildung 9.3: Mittels Cross-Validation bestimmte RV-Werte der durch PCR ermittelten Temperaturprognosen für unterschiedliche Jahres- und Vorlaufzeiten.

Im Anteil der „korrekt“ erstellten Prognosen innerhalb fester Fehlergrenzen spiegelt sich die Überlegenheit der PCR hingegen kaum wider. Vergleicht man die hier erzielten Resultate mit denen des einfachsten Analogverfahrens (Abschnitt 8.2.2), so sind keine wesentlichen Unterschiede zu erkennen. Der Anteil der Prognosen, deren Vorhersagefehler $0,674\sigma$ nicht überschreitet, liegt hier je nach Jahres- und Vorlaufzeit zwischen 45% und 53% (siehe Abb. 9.4). Wie schon im Falle des Analogverfahrens ist also keine Abweichung von der theoretischen Erwartung bei Annahme einer normalverteilten Grundgesamtheit festzustellen.

Der Vollständigkeit halber sei noch darauf hingewiesen, daß die im Rahmen der PCR berechneten Fehlergrenzen von denen im Falle des Analogverfahrens geringfügig abweichen. Auf Grund des unterschiedlichen Testzeitraums liegen sie hier bei ca. $1,1^{\circ}\text{C}$ im Frühling, bei ca. $1,0^{\circ}\text{C}$ im Sommer und Herbst sowie bei ca. $1,8^{\circ}\text{C}$ im Winter. Wie schon zuvor handelt es sich bei diesen Werten um Mittel der jeweiligen drei Kalendermonate. Bei der Bestimmung der prozentualen Anteile wurden hingegen die exakten Standardabweichungen der einzelnen Kalendermonate zugrunde gelegt.

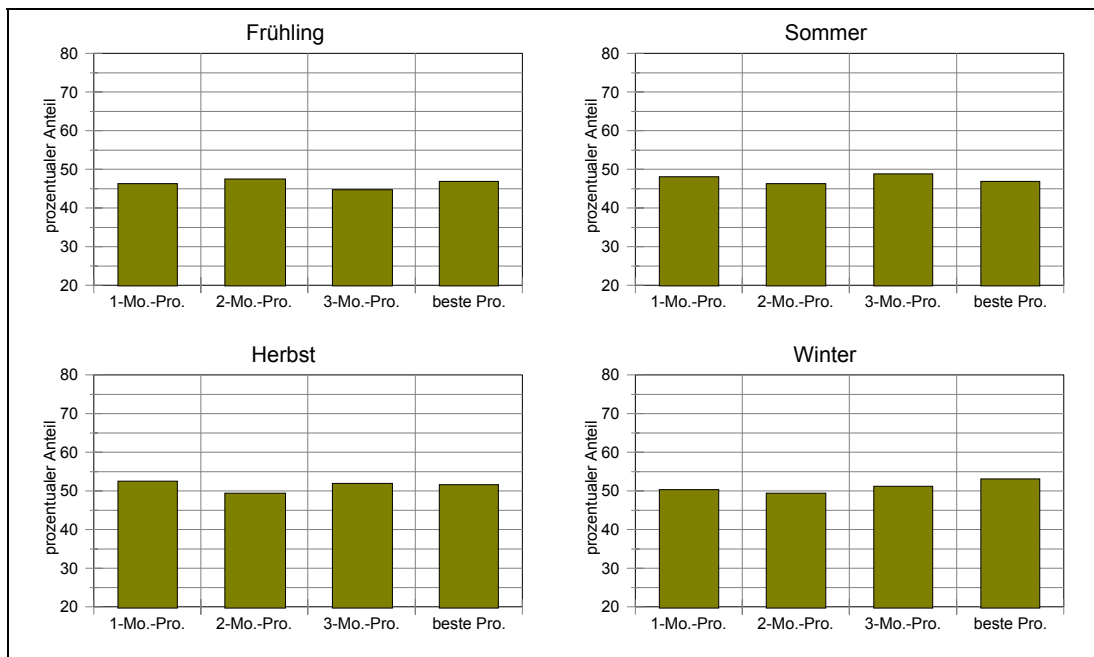


Abbildung 9.4: Mittels Cross-Validation bestimmter prozentualer Anteil der durch PCR ermittelten Temperaturprognosen, deren Prognosefehler im Bereich 0,674 Standardabweichungen liegt. Je nach Jahreszeit entspricht das einer Abweichung von ca. 1°-2°C (siehe Text).

