

# 1. Einführung

Es wird allgemein davon ausgegangen, daß sich das dynamische System Atmosphäre chaotisch verhält. Derartige Systeme zeichnen sich durch eine sehr hohe Sensitivität zu den Anfangsbedingungen aus – schon geringfügige Abweichungen führen zu einem völlig unterschiedlichen Verlauf der Zustandsänderungen. Dabei hängt lediglich der Zeitpunkt, ab welchem sich die Entwicklung des Systems unterscheidet, von der Größe der Abweichungen in den Anfangsbedingungen ab. Zu einer unterschiedlichen Evolution kommt es aber in jedem Fall. Diese Eigenschaft der Atmosphäre führt dazu, daß sich derzeitige erfolgreiche numerische Wettervorhersagen nur bis etwa zehn Tage im voraus erstellen lassen. Selbst die wesentlich genauere Kenntnis der atmosphärischen Zustände würde diesen Zeitraum vermutlich nur um wenige Tage erweitern, da die Verwendung präziserer Anfangsbedingungen eine höhere Auflösung der dynamischen Modelle erfordert. Zudem ist eine Verfeinerung der Modellphysik notwendig, um die Parametrisierung der neu erfaßten kleinräumigen Prozesse durch eine tatsächliche Berechnung zu ersetzen. Es hat sich aber im Laufe der Forschungsentwicklungen herausgestellt, daß die dynamischen Modelle auf derartige Modifikationen mit einer Verringerung der sogenannten Fehlerverdoppelungszeit reagieren (*Lorenz, 1993*). Dadurch geht der zeitliche Gewinn, den die Anwendung noch detaillierterer Anfangsbedingungen mit sich bringt, zumindest teilweise wieder verloren. Somit muß z. Zt. davon ausgegangen werden, daß auch die Ausnutzung der stetig steigenden Computerkapazitäten nicht zu verlässlichen Wettervorhersagen über mehr als zwei Wochen hinaus führen wird.

Scheinbar diesen Erkenntnissen zum Trotz hat sich das Forschungsgebiet langfristige Witterungsvorhersage weltweit etabliert. In früheren Jahren wurden dabei ausschließlich statistische Modelle verwendet, die auf folgender Annahme basierten: Wenn zwei beobachtete atmosphärische Zustände ursprünglich sehr nahe beieinander liegen, d.h. sich ähneln, so ist zu erwarten, daß ihre Entwicklung noch eine Zeitlang vergleichbar verläuft (z.B. *Barnston und Livezey, 1989a*; *Van den Dool, 1994*). Heute werden neben den statistischen Verfahren auch dynamische Modelle zur Langfristprognose, die in diesem Bereich häufig auch als kurzfristige Klimasimulation bezeichnet wird, verwendet. Die zugrundeliegende Philosophie hat sich dabei jedoch nicht wesentlich geändert: Im saisonalen Zeitskalenbereich existiert nur eine begrenzte Anzahl an Zirkulationsmustern, welche die Atmosphäre regelmäßig über längere Zeit hinweg anzunehmen vermag (siehe dazu Abschnitt 1.1). Durch den Vergleich mit der Vergangenheit gibt der momentane Zustand der Atmosphäre darüber Auskunft, welche dieser Attraktoren in der Zukunft bevorzugt angesteuert werden. Eine Aussage über die genauen Zeitpunkte, zu welchen das System diese wahrscheinlichsten Zirkulationsanomalien einnimmt, ist jedoch nicht möglich. Es können nicht Zugbahn und Intensität einzelner Druckgebilde vorhergesagt werden, sondern nur das Verhalten großskaliger, barotroper Strukturen, d.h. die Lage der planetaren Wellen (*Vautard et al., 1996*). Somit ist zwar keine detaillierte Wettervorhersage mehr möglich, jedoch kann der grundsätzliche Witterungscharakter vorausgesagt werden. In jüngster Vergangenheit haben dabei das Klima regulierende Grenzbedingungen, auch Forcing-Parameter genannt (engl. lower boundary conditions), zunehmend an Bedeutung gewonnen (siehe auch Abschnitt 1.1). Da sie nur langsamen Veränderungen unterworfen sind, beeinflussen sie den Witterungsverlauf über einen längeren Zeitraum hinweg auf die gleiche Art und Weise (z.B. *Carson, 1998*) – sie lenken die Atmosphäre immer wieder in bestimmte Richtungen. Beide, sowohl die atmosphärischen Zustandsvariablen als auch die unteren Grenzbedingungen, spiegeln den Zustand des gekoppelten

Gesamtsystems Ozean/Erdoberfläche – Atmosphäre wider. Daher ist intuitiv davon auszugehen, daß ihre Verwendung äquivalent ist. Jedoch gehen einige Autoren mittlerweile davon aus, daß selbst zeitliche Mittel atmosphärischer Parameter einen zu großen Rauschanteil hochfrequenter Prozesse enthalten. *Kumar et al.* (2001) kommen sogar zu dem Fazit, daß der saisonale Witterungsablauf völlig unabhängig von den atmosphärischen Anfangsbedingungen ist.

Eine entscheidende Rolle bei der Frage, inwieweit eine erfolgversprechende Langfristprognose realistisch ist, spielt das Verhältnis zwischen der langjährigen und der internen Variabilität der Atmosphäre. Dieses Verhältnis wird auch als „signal-to-noise“ ratio bezeichnet. Die langjährige Variabilität ist schlicht die beobachtete Varianz eines Klimaparameters an einem bestimmten Ort der Erde, z.B. der Temperatur. Die interne Variabilität hingegen ist ein Maß für den Rauschanteil. Sie gibt den Teil der Gesamtvarianz an, der auf die langfristig nicht vorhersagbaren hochfrequenten Prozesse zurückzuführen ist (z.B. Hoch- und Tiefdruckgebiete, kurze Wellen). Ist das Verhältnis beider Größen klein, so wird das Klimasignal vom Rauschen überdeckt und eine Langfristprognose ist praktisch nicht möglich. In den anderen Fällen überwiegt der „Wettereinfluß“ die klimatischen Einflüsse nicht und eine Vorhersage ist sinnvoll.

Aufgrund der Beschaffenheit des Erdklimas kann man davon ausgehen, daß das Verhältnis von langjähriger zu interner Variabilität auf der Erdoberfläche sehr unterschiedlich ist. Insbesondere in den mittleren Breiten, im Bereich der Westwinde, dort wo Zyklonen einen Großteil des Wetterablaufs prägen, ist von einem hohen Rauschanteil auszugehen. Im tropischen und subtropischen Bereich hingegen ist zu vermuten, daß ein wesentlicher Teil des Witterungsablaufs von den Forcing-Parametern geprägt wird.

Daher ist es nicht verwunderlich, daß es neben den Untersuchungen zur langfristigen Vorhersage relevanter Klimaparameter einige Autoren gab, die versucht haben, den „signal-to-noise“ ratio zu bestimmen. *Madden* (1976) berechnete unter Zuhilfenahme der spektralen Dichtefunktionen von täglichen sowie monatlichen Werten die natürliche Variabilität (sie entspricht der internen Variabilität) des Bodenluftdrucks für den Bereich von 20° bis 80° nördlicher Breite. Er fand heraus, daß sie die langjährige Variabilität im Bereich zwischen 40° und 60°N vollständig erklärt. *Madden und Shea* (1978) führten eine vergleichbare Untersuchung für die Temperatur an 107 in den USA gelegenen Stationen durch. Im Gegensatz zum Luftdruck verblieb hier an mehreren Stationen eine Restdifferenz bei Subtraktion der natürlichen Variabilität von der Gesamtvarianz. Man schlußfolgerte demnach, daß ein Potential für die langfristige Temperaturvorhersage in diesem Gebiet existiert. *Feldstein* (2000) hingegen bestimmte den Anteil des Rauschens für einige der wichtigsten Zirkulationsmuster der Atmosphäre. Er fand heraus, daß nur im Falle der sogenannten PNA (siehe ebenfalls Abschnitt 1.1) Signal und Rauschen in der gleichen Größenordnung liegen.

Auch mittels dynamischer Modelle wurde in den letzten Jahren wiederholt die saisonale Vorhersagbarkeit der wichtigsten meteorologischen Parameter abgeschätzt. Das Prinzip, nach dem dabei vorgegangen wird, ist im wesentlichen immer gleich: Bei einem aus mehreren Integrationen bestehenden Ensemble werden für jeden Lauf leicht unterschiedliche Anfangsbedingungen verwendet. Der Starttermin jedes Mitgliedes variiert um einige Tage, so daß die atmosphärischen Parameter in allen Fällen kleine Abweichungen zueinander aufweisen. Die Forcing-Parameter bleiben jedoch für alle Mitglieder unverändert, so daß ihr Einfluß in allen Fällen gleich bleibt. Die Varianz der Ensemblemitglieder untereinander stellt dann ein Maß für die interne Variabilität der Atmosphäre dar (*Kumar et al.*, 2001). Eine Untersuchung von *Martineu et al.* (1999)

zeigte auf, daß die Modulation der Klimavariabilität europäischer Winter durch Forcing im Vergleich zur internen Variabilität in diesem Gebiet nur gering ist. Gar eine quantitative Abschätzung für das Verhältnis zwischen Klimasignal und Rauschen nahmen *Fennesy und Shukla* (2000) für das Gebiet der Vereinigten Staaten vor. Sie errechneten Werte im Bereich zwischen 0,5 und 3,0 für die Temperatur und die Niederschlagsmenge, wobei für die Temperatur großräumig die höheren Werte ermittelt wurden.

Obwohl solche Resultate auf Annahmen und Vereinfachungen basieren, deuten sie an, wie schwierig erfolgreiche Langfristprognosen vor allem für die außertropischen Gebiete zu realisieren sind. Zumindest für den Parameter Temperatur scheint jedoch ein gewisses Vorhersagepotential vorhanden zu sein. Nur praktische Untersuchungen können aufzeigen, inwieweit die mit hoher Sicherheit nicht perfekt den natürlichen Bedingungen entsprechende Physik der dynamischen Modelle die tatsächlich vorhandene Vorhersagbarkeit falsch einschätzt. Insbesondere der Umstand, daß die Modelle der genannten Untersuchungen lediglich die Meeresoberflächentemperatur (SST) als einzige der unteren Grenzbedingungen verwenden und dabei andere vernachlässigen, gibt Anlaß zur Hoffnung, daß das Klimasignal unterschätzt wird.

## 1.1 Quellen der atmosphärischen Vorhersagbarkeit

Betrachtet man die großräumigen Feldverteilungen meteorologischer Parameter zu den unterschiedlichsten Zeitpunkten, so fallen bestimmte, häufig vorhandene Strukturen in der Zirkulation auf. Erweitert man dabei den zeitlichen Rahmen, d.h. erstellt man monatliche, saisonale oder gar jährliche Mittelwerte, so kristallisieren sich diese Zirkulationsanomalien immer deutlicher heraus. Es zeigt sich, daß sie sich an ganz bestimmten Orten befinden, während in anderen Gebieten der Erde nur die jeweiligen zeitlichen Mittelwerte vorzufinden sind. Man bezeichnet sie als „Telekonnektionen“ (Fernwirkungen), da sich ihre Zentren nicht unabhängig voneinander verhalten, obwohl sie teilweise viele 1000 km voneinander entfernt sind. *Silverman und Dracup* (2000) definieren sie als „wiederkehrende und andauernde Beziehungen zwischen klimatischen Phänomenen in zwei diskreten Gebieten der Erde.“ Diese Eigenschaften der Telekonnektionen erleichtern die Identifikation des atmosphärischen Zustandes – sie ermöglichen das Auffinden bestimmter Moden.

Die Auswirkungen solcher dominierenden Strukturen in der Zirkulation auf den langfristigen Witterungscharakter untersuchten *Rodionov und Assel* (2000). Sie fanden einen Zusammenhang zwischen der Strenge des nordamerikanischen Winters und den individuellen Zuständen einiger Telekonnektionen. Bei ihnen findet man auch eine detaillierte Beschreibung nahezu sämtlicher bereits bekannter Fernwirkungen (u.a. Pacific/North American pattern [PNA], Nordatlantische Oszillation [NAO], Polar/Eurasian pattern [POL]). *Malberg und Bökens* (1997) fanden ebenfalls einen signifikanten Zusammenhang zwischen dem Zustand der NAO und den Wintermitteltemperaturen Berlins. In deutlich abgeschwächter Form hat diese Verknüpfung auch in den Sommermonaten Bestand. Zwar handelt es sich in beiden Fällen um rein diagnostische Zusammenhänge, jedoch wird die Weiterentwicklung der Telekonnektionen von Änderungen der tropischen SSTs und der tropischen Konvektion gesteuert (*Barnston und Livezey*, 1991), somit auch die Lage und Intensität der Strahlströme, die Zugbahn von Zyklonen und andere Charakteristika der Großwetterlage. *Palmer und Anderson* (1994) fanden Hinweise dafür, daß die nordhemisphärische Zirkulation in quasistationären Moden (Wetterregimen) abläuft,

d.h. sprunghaft von einem Mode in den nächsten übergeht. Dabei kann ein solcher Wechsel zwar auch von den chaotischen Wetterabläufen verursacht werden, jedoch haben die Forcing-Parameter vermutlich einen Einfluß darauf, welche Moden über längere Zeit hinweg häufiger angenommen werden.

Neben der Meeresoberflächentemperatur werden Grad und Temperatur der Eisbedeckung, die Landoberflächentemperatur, die Albedo, die Bodenfeuchte sowie die Schneebedeckung zu den die Entwicklung der Atmosphäre beeinflussenden Grenzbedingungen gezählt, denen jedoch unterschiedliche Bedeutung beigemessen wird (u.a. *Palmer und Anderson*, 1994; *Carson*, 1998). Für die Bodenfeuchte, z.B., konnten *Huang et al.* (1996b) einen signifikanten Zusammenhang mit den monatlichen Temperaturabweichungen in den USA nachweisen. Die übergeordnete Rolle der SSTs hängt mit dem sogenannten ENSO-Phänomen zusammen:

Der Begriff ENSO (El Niño/Southern Oscillation) umfaßt sowohl großräumige SST-Anomalien im östlichen tropischen Pazifik als auch eine atmosphärische Telekonnektion, deren Zentren sich im südostpazifischen bzw. indonesischen Raum befinden. Beide Anteile sind eng miteinander gekoppelt. Mit El Niño bezeichnet man das unregelmäßige Auftreten ungewöhnlich warmer Wassermassen an der Westküste Südamerikas; bei der Southern Oscillation handelt sich um eine periodische Schwingung der Luftdruckdifferenz zwischen Tahiti und Darwin. Die Änderung der Meeresoberflächentemperatur führt nahezu im gesamten tropischen Raum zu einer Verlagerung jener Gebiete, in denen regelmäßig starke konvektive Niederschläge niedergehen bzw. jenen, in denen Niederschlagsarmut herrscht. Die Auswirkungen des ENSO-Phänomens reichen jedoch weit über die Tropen hinaus. Insbesondere im Bereich der PNA (Nordamerika sowie der angrenzende nördliche Pazifik) beeinflussen die tropischen SST-Anomalien die Zirkulation der mittleren Breiten (*Horel und Wallace*, 1981; *Ropelewski und Halpert*, 1986). Aber auch für den nordatlantischen und europäischen Sektor konnten Zusammenhänge nachgewiesen werden: *Fraedrich und Müller* (1992) wiesen eine Verbindung zwischen den Wintermitteltemperaturen Skandinaviens und tropischen SST-Anomalien nach. *Fraedrich* (1994) stellte eine Verlagerung der Zugbahn atlantischer Tiefdruckgebiete hin zu mehr nördlichen Routen im El Niño-Fall, bzw. zu mehr zonalen Routen im La Niña-Fall (abnormal kalte Wassermassen vor den Küsten Perus und Ecuadors) fest. Eine detaillierte Beschreibung der Vorgänge, Wechselwirkungen und Telekonnektionen des ENSO-Phänomens findet sich z.B. bei *Dettmann* (2000).

Wesentlich umstrittener ist die Frage, ob auch Prozesse der oberen Atmosphäre sowie der Strahlung einen nachweisbaren Einfluß auf das saisonale Klima der Troposphäre haben. Über einen längeren Zeitraum beständige Vorgänge sind hier die sogenannte „quasi-biennial oscillation“ (QBO) der Stratosphäre und der 11-jährige solare Zyklus. *Van Loon und Labitzke* (1988) wiesen einen signifikanten Zusammenhang zwischen der Zirkulation der nördlichen Hemisphäre und dem solaren Fluß für die Monate Januar und Februar nach. Dabei wurden die verwendeten Daten nach der Phase der QBO (West- bzw. Ostwinde in der äquatorialen Stratosphäre) differenziert. *Barnston und Livezey* (1989b) überprüften die Feldsignifikanzen der von *van Loon und Labitzke* aufgestellten Beziehungen und stellten fest, daß diese insbesondere für die Westphase der QBO ein hohes Potential für die langfristige Witterungsvorhersage in den USA enthalten. Die eingetretene winterliche Zirkulation des Jahres 1989 stand jedoch im Widerspruch zu den Ergebnissen obengenannter Untersuchungen. *Barnston und Livezey* (1991) suchten nach Ursachen für das Scheitern der aufgestellten Beziehungen. Sie kamen zu dem

Schluß, daß vermutlich zwar ein Zusammenhang zwischen der Sonnenaktivität und der winterlichen Witterung der Nordhemisphäre besteht, dieser jedoch von den Auswirkungen der Southern Oscillation überlagert werden kann. Da beide Beziehungen voneinander unabhängig sind, kann es sowohl zur Aufhebung als auch Verstärkung ihrer Effekte auf das saisonale Klima kommen.

## 1.2 Methoden zur langfristigen Witterungsvorhersage

Die vielleicht älteste Form der Langfristprognose sind die Bauernregeln. Schon vor Jahrhunderten stellte man aus den langjährigen Beobachtungen des saisonalen Klimaverlaufs sogenannte Witterungsregeln auf. Man ging davon aus, daß zu mehreren fixen Zeitpunkten im Verlaufe des Jahres auf eine ganz bestimmte Witterung (zumeist) auch eine ganz bestimmte Wetterentwicklung folgt. D.h. das aktuelle Wetter dient den Bauernregeln als Vorbote für den zukünftigen Wetterzustand. *Malberg* (1999) unterzog derartige Regeln einer wissenschaftlichen Prüfung und fand heraus, daß zumindest bei einem Teil von ihnen die Eintreffwahrscheinlichkeit deutlich über 50% liegt.

Die weltweit ersten Langfristprognosen auf wissenschaftlicher Basis wurden schon Ende des vorletzten Jahrhunderts gemacht (*Bagrov*, 1993). In Deutschland wurde 1920 eine „staatliche Forschungsstelle für langfristige Witterungsvorhersage“ gegründet, deren Leiter Franz Baur war. Seit Anfang der siebziger Jahre wächst das Forschungsinteresse an saisonalen Wettervorhersagen deutlich an. Zunächst wurde mittels statistischer Verfahren versucht, das weitere Wettergeschehen erfolgreich vorherzusagen. Die Verfügbarkeit leistungsfähiger Großrechner führte später dazu, daß auch im Bereich der dynamischen Modellierung häufiger der Versuch unternommen wurde, den Vorhersagezeitraum auszudehnen. Aber auch für statistische Methoden wird die gesteigerte Computerkapazität genutzt: Es werden stets kompliziertere Verfahren verwendet und die Datensätze, auf die man zurückgreifen kann, werden immer länger. Jedoch erweist sich selbst die Leistungsfähigkeit modernster Methoden als begrenzt. Die Wissenschaftler kommen zumeist zu dem Fazit, daß eine zwar reale, jedoch nur recht geringe Prognosenleistung (engl. skill) existiert (*Livezey*, 1990). Das dennoch stetig steigende öffentliche Interesse erklärt sich durch den hohen wirtschaftlichen Nutzen akkurater Langfristprognosen. Man geht sowohl seitens der Verbraucher als auch seitens der Meteorologen davon aus, daß selbst kleinste Verbesserungen der Vorhersagegüte einen großen wirtschaftlichen Gewinn nach sich ziehen könnten (*American Meteorological Society*, 1983). Einen Überblick darüber, welche Wirtschaftszweige von saisonalen Vorhersagen profitieren könnten, findet man z.B. in *Nicholls* (1980).

In den folgenden Abschnitten sollen die wichtigsten Arbeiten zur langfristigen Witterungsvorhersage vorgestellt und ihre Vorgehensweise kurz skizziert werden.

### 1.2.1 Statistische Verfahren

Empirische Methoden zur Langfristprognose setzen den aktuellen Zustand der Atmosphäre auf unterschiedliche Art und Weise in Beziehung zu den Beobachtungen der Vergangenheit. Verfahren, welche die Feldverteilungen (selten auch monovariate Datenreihen) eines oder mehrerer klimatologischer Parameter (Prediktoren) auf Übereinstimmungen in der Vergangenheit hin überprüfen, bezeichnet man als Analogverfahren. Dabei wird stets vorausgesetzt, daß die Entwicklung zweier diskreter

Zustände einen ähnlichen Verlauf nimmt, wenn die Anfangsbedingungen möglichst gut übereinstimmen. Die Prognose ergibt sich aus den tatsächlich eingetroffenen Werten der Zielgröße (Prediktand) der ähnlichen (analogen) Fälle. Zumeist werden mehrere analoge Fälle aus dem Datenarchiv ausgewählt, deren Mittelwert dann vorhergesagt wird. Die größte Schwierigkeit der Analogverfahren besteht darin, wirklich gute Analogfälle überhaupt zu finden. *Van den Dool* (1994) leitete eine Beziehung zwischen der Größe einer Bibliothek historischer Daten, der Distanz zwischen einem willkürlich herausgesuchten Fall und seinem ähnlichstem Nachbarn sowie der Größe des untersuchten Gebietes in Form von Freiheitsgeraden ab. Dabei zeigte sich, daß es statistisch etwa  $10^{30}$  Jahre dauert, bis ein analoger Fall auftritt, dessen Abweichungen in derselben Größenordnung wie die heutigen Beobachtungsfehler liegen. Die derzeitigen Archivgrößen (10-100 Jahre) lassen nur die Verwendung von etwa zwei Freiheitsgeraden zu, um noch relativ sicher gute analoge Fälle zu finden. Die notwendige Reduktion der Freiheitsgerade läßt sich vor allem durch Datenkompression, aber auch durch zeitliche Mittelung realisieren.

In der ehemaligen Sowjetunion wurden Mitte der siebziger Jahre u.a. verschiedene Analogverfahren zur Vorhersage monatlicher bzw. saisonaler (teils noch längerer) Temperatur- und Niederschlagsanomalien verwendet (vergl. *Livezey und Jamison*, 1977). Die monatliche Prognose ergab sich durch die Auswahl eines einzelnen analogen Jahres, wobei die Entwicklung der atmosphärischen Zirkulation bis zu fünf Monate in die Vergangenheit zurück berücksichtigt wurde. Als Prediktoren wurden der Bodendruck, das 500- sowie 1000-hPa-Geopotential, die Temperatur, die Eisbedeckung in der Barentssee, ein Index des zonalen Flusses sowie nordatlantische SSTs verwendet. Die Häufigkeit bestimmter Zirkulationsmuster und die Reihenfolge ihrer Übergänge spielten bei der Wahl des Analogons zusätzlich eine wichtige Rolle, da die sowjetischen Wissenschaftler vom zyklischen Verhalten sogenannter „Makroprozesse“ (übergeordnete Zirkulationsanomalien) ausgingen. „Makroprozesse“ kamen u.a. auch bei Vorhersagen über einen Monat hinaus zur Anwendung: Die hemisphärische Zirkulation wurde in neun verschiedene Typen eingeteilt, wobei es jeweils drei Typen im atlantisch-eurasischen Sektor (W,C,E) und im pazifisch-amerikanischen Sektor (Z,M<sub>1</sub>,M<sub>2</sub>) gab. In einem 70 Jahre umfassenden Datensatz fand man genau sieben Zyklen mit charakteristischen Häufigkeiten und Übergängen der Zirkulationstypen in dem 15-monatigen Zeitraum von August bis Oktober des darauffolgenden Jahres. Die Vorhersage erfolgte durch Identifikation des aktuellen Zyklus in den Monaten von August bis Dezember, wodurch die atmosphärische Entwicklung der folgenden zehn Monate als bekannt vorausgesetzt werden konnte.

Bei der Anwendung analoger Techniken haben insbesondere zwei Kriterien entscheidenden Einfluß auf die finale Vorhersage und somit auf die Prognosenleistung: Zum einen die Anzahl der ausgewählten analogen Fälle, zum anderen die verwendeten Ähnlichkeitsmaße, d.h. die Kriterien, durch welche die Analogie von Feldverteilungen bestimmt wird. *Gruza und Rankova* (1981) untersuchten die Frage, ob zur Prognose monatlicher Temperaturanomalien der Nordhemisphäre eher die Verwendung nur eines Analogons oder aber mehrerer analoger Fälle geeignet ist. Die auf einer Gruppe analoger Jahre basierenden Vorhersagen zeigten sich überlegen. Auch *Gruza et al.* (1981) stellten fest, daß die Benutzung nur eines „besten“ analogen Jahres zumeist erfolglos verlief. Die Prognosenleistung erwies sich jedoch auch bei der Verwendung von Kollektiven als gering. *Golubev* (1983) suchte nach Möglichkeiten, eine Grenze für die analogen Fälle zu bestimmen. Ziel war es, statt einer vorab festgelegten Anzahl an Fällen einen Schwellenwert einzuführen, um somit die Anzahl der Mitglieder des

Kollektivs zu optimieren. Hierfür wertete er die Analogien zwischen meteorologischen Feldern zunächst rein subjektiv aus. Diese Ergebnisse verglich er später mit jenen, welche a priori festgelegte Ähnlichkeitsmaße lieferten. *Gruza und Rankova* (1993) weiteten ihre Untersuchungen auf die Frage aus, welche Algorithmen zur Bestimmung von Analogien, zumeist in Form eines skalaren Wertes, die geeignetsten sind. Sie resümieren, daß zur optimalen Vorhersage zwischen sieben und sechzehn Fälle notwendig sind, deren Auswahl an Hand mehrerer Distanzmaße vorgenommen werden sollte. Damit ein Analogon hinsichtlich des Vorhersagepotentials von guter Qualität ist, müssen verschiedene statistische Kriterien kleine Differenzen, d.h. eine gute Ähnlichkeit, ergeben. Für die drei untersuchten Kontinente Europa, Asien und Nordamerika zeigte sich, daß für Europa das größte Potential für langfristige Temperaturprognosen vorhanden ist.

Praktisch ein Meilenstein der Analogtechnik ist das von *Barnett und Preisendorfer* (1978) entwickelte Verfahren zur Vorhersage saisonaler Mittel der Temperaturen Nordamerikas. Sie führten den „climate state vector“ (CSV) ein, einen Vektor, dessen Bewegung (Orbit) im Phasenraum die zeitliche Evolution des Klimasystems repräsentiert. Die Parameter, welche zur Bestimmung des CSV herangezogen werden, stellen ein Maß für die Energie des Klimasystems dar. Um die Zahl der Freiheitsgrade nicht zu groß werden zu lassen, wurde jedes Prediktorfeld einer separaten EOF-Analyse (empirische Orthogonal funktions-Analyse) unterworfen. Anschließend wurden die wichtigsten Regionen der Variabilität jedes Feldes ermittelt und die resultierenden Datenreihen mittels einer zweiten EOF-Analyse nochmals komprimiert. *Livezey und Barnston* (1988) nahmen gewisse Modifikationen an diesem Modell vor, um es in die routinemäßigen Vorhersagen des „Climate Analysis Center (CAC)“ mit einzubinden (siehe auch *Barnston und Livezey*, 1989a). Sie führten eine Gewichtung der analogen Fälle in Abhängigkeit von ihrem Abstand zum aktuellen Fall sowie die Verwendung von sogenannten Antianaloga ein. Zusätzlich führten sie umfangreiche Sensibilitätstests zur Bestimmung der besten Prediktoren durch. Insbesondere für den nordamerikanischen Bereich überraschend war dabei das Ergebnis, daß die Verwendung eines auf Luftdruckdifferenzen basierenden ENSO-Indexes bessere Resultate brachte als diejenige von SST-Anomalien. *Livezey et al.* (1990) verbesserten die Vorhersagen des bestehenden Analogverfahrens durch die Hinzunahme der Persistenz, d.h. der Erhaltungsneigung der Atmosphäre. Mittels bivariater Regression wurden die zwei unabhängigen Prognosen zu einer finalen zusammengefügt.

Ein wesentlich einfacheres Analogverfahren zur Vorhersage von Temperaturanomalien im Bereich der USA entwickelten *Bergen und Harnack* (1982). Sie verwendeten ausschließlich die Anomalie-Korrelation zur Bestimmung der analogen Fälle. Bei der Anomalie-Korrelation handelt es sich um einen simplen Algorithmus zur Aufdeckung von Ähnlichkeiten in den Strukturen klimatologischer Felder. Dabei bezieht sie sich immer auf die Abweichungen von mittleren Feldern, also auf Anomalien (vergl. *Wilks*, 1995). Es zeigte sich, daß dieses Verfahren mit dem viel komplexeren Ansatz des CSV konkurrieren kann. Beide Methoden erreichen nur eine recht geringe Prognosenleistung und schlagen die Persistenz nicht während des gesamten Jahresverlaufes. Während *Barnett und Preisendorfer* insbesondere für die Sommermonate eine signifikante Prognosenleistung ermittelten, fand sich das Gütemaximum bei *Bergen und Harnack* im Januar und Februar.

In den letzten Jahren haben die Analogverfahren an Bedeutung verloren. In den Mittelpunkt des Forschungsinteresses sind statt dessen hochentwickelte Verfahren wie die PCR und vor allem die CCA gerückt (s.u.). Daß zumindest das weitere

Miteinbeziehen analoger Techniken jedoch noch immer vorteilhaft sein kann, zeigten *Vautard et al.* (1999). Sie schlagen ein zweistufiges Modell zur Vorhersage nordamerikanischer Temperaturanomalien vor: In einem ersten Schritt werden sogenannte space-time principal components (ST PCs), aus einem Prediktorfeld extrahierte, vorhersagbare Komponenten, mittels linearer Regression in den vorherzusagenden Zeitraum extrapoliert. Die ST PCs werden zuvor durch eine spezielle Form der erweiterten EOF-Analyse, der MSSA („multichannel singular spectrum analysis“), bestimmt. Im Gegensatz zur einfachen EOF-Analyse ist in diesem Fall eine räumliche und zeitliche Auflösung der Eingangsdaten möglich, so daß der evolutionäre Verlauf der Atmosphäre ebenfalls Berücksichtigung findet. Im zweiten Schritt wird dann durch ein auf den prognostizierten ST PCs basierendes Analogverfahren die Temperaturprognose erstellt. Dieses Modell erwies sich im Vergleich zur CCA als konkurrenzfähig, für einige Jahreszeiten lieferte es sogar leicht überlegene Vorhersagen.

Zu den am häufigsten verwendeten Verfahren gehören die Regressionen, zu denen auch die Kanonische Korrelationsanalyse (engl. „canonical correlation analysis [CCA]“) gehört. Im Gegensatz zu den Analogverfahren erstellen diese eine konkrete, meist lineare mathematische Beziehung zwischen den Einfluß- und Zielgrößen eines Datensatzes, welche dann zur Schätzung zukünftiger Werte verwendet wird. Da im Bereich der Langfristprognose eine Vielzahl potentieller Prediktoren (möglicherweise Vorhersagepotential enthaltende Einflußgrößen) existiert, wird praktisch immer eine Datenreduktion durchgeführt, um somit die größtmögliche Stabilität der ermittelten statistischen Beziehungen zu gewährleisten. Üblicherweise bedient man sich hier entweder der Vorselektion oder aber der Datenkompression durch EOF-Analyse. Verfahren, bei denen die aus einer EOF-Analyse resultierenden Hauptkomponenten (engl. principal components [PCs]) als Einflußgrößen für eine lineare Regression verwendet werden, bezeichnet man als „principal-components regression (PCR)“. Die CCA ist ein statistisches Verfahren, welches Linearkombinationen zweier multivariater Datensätze derart berechnet, daß der Korrelationskoeffizient zwischen ihnen jeweils maximal wird. Dabei gibt es, in Analogie zur EOF-Analyse, mehrere verschiedene Moden. Vereinfacht dargestellt wird nach einem Zusammenhang zwischen den Strukturen der Einflußgrößen und denen der Zielgrößen gesucht.

Eine außergewöhnliche Form der PCR zur Vorhersage des Temperaturverlaufs im kommenden Monat stellten *Batyreva et al.* (1995) vor: Sie bestimmten den Zusammenhang zwischen den einzelnen Kalendertagen des jeweiligen Monats mittels einer EOF-Analyse, d.h. sie suchten nach vorhandenen Strukturen in der Dimension Zeit statt nach räumlichen Mustern. Durch simple lineare Regression werden aus den von dynamischen Modellen prognostizierten Temperaturmitteln der ersten zehn Tage des Monats diese zeitlichen Hauptkomponenten geschätzt. Durch Rekonstruktion aus den Hauptkomponenten läßt sich dann ein geglätteter Temperaturverlauf für den gesamten Monat bestimmen. Auf diese Art und Weise konnte eine Prognosenleistung für die Monate Januar und Juli an 14 sowjetischen Stationen nachgewiesen werden. *Chuvashina und Margasova* (1997) nutzten einen vergleichbaren Ansatz zur Prognose von 5-Tages-Mitteln der Temperatur, wobei sie jedoch klimatologische Parameter wie SSTs, den Southern Oscillation Index (SOI) und Niederschlagsverteilungen als Prediktoren verwendeten.

Einen Vergleich zwischen den Vorhersagepotentialen der CCA und der multiplen linearen Regression führten *Vilfand und Rudicheva* (1993) durch. Beide Methoden wurden mit den räumlichen PCs einer großen Anzahl potentieller Prediktoren



durchgeführt. Dazu zählten das 500-hPa-Geopotential, SST-Gradienten, ein NAO-Index, der SOI, Temperaturanomalien der Vormonate, die mittlere troposphärische Zirkulation sowie ein stratosphärischer Index. Es zeigte sich, daß bei gleicher Güte der Vorhersagen (erklärte Varianz  $r^2$  zwischen 0,1 und 0,2) die CCA mit einer deutlich kleineren Anzahl an Einflußgrößen auskommt, so daß eine wesentlich höhere Stabilität am unabhängigen Material die Konsequenz ist – die CCA ist daher der multiplen Regression vorzuziehen. Auch *Batyreva et al.* (1994) fanden eine Bestätigung für die Überlegenheit der CCA im Vergleich zu anderen Verfahren. Sie erstellten Vorhersagen der monatlichen Temperaturanomalien sowohl mittels der CCA als auch mittels Diskriminanzanalyse. Dabei zeigte sich die CCA deutlich überlegen, wobei die höchste Prognosenleistung durch die Einteilung der monatlichen Temperaturanomalien in fünf vorherzusagende Klassen erzielt wurde. Die Verwendung von Quantilen anstelle diskreter Werte zur langfristigen Vorhersage und der Verifikation klimatologischer Parameter ist insbesondere unter den Autoren der ehemaligen Sowjetunion weit verbreitet.

Während sich ein Großteil der Untersuchungen, welche Meeresoberflächentemperaturen als Prediktoren verwenden, auf tropische SSTs (ENSO) beschränkt, entdeckte *Colman* (1997) einen Zusammenhang zwischen nordatlantischen SSTs im Winter und der Temperatur Zentral-Englands im darauffolgenden Sommer. Das sich aus dieser Beziehung ergebende Korrelationsmuster ist vermutlich auf den Einfluß der NAO zurückzuführen. *Colman und Davey* (1999) weiteten diese Untersuchung auf das gesamte Europa aus, wobei sie die Niederschlagsmenge und den Luftdruck als Prediktanden hinzufügten. Für die Monate Juli und August konnte für große Teile Europas eine signifikante Prognosenleistung nachgewiesen werden. Mit Korrelationen über 0,5 in einem Gebiet von Frankreich bis nach England erwies sich die Temperatur als am besten vorhersagbar. Ebenfalls der PCR bedienten sich *Cannon und McKendry* (1999) zur Vorhersage des gesamten jährlichen indischen Monsunregens. Sie verglichen die Leistungsfähigkeit dieses Modells mit derer nichtlinearer neuronaler Netze (s.u.). Die Höhe des zu erwartenden Niederschlages wurde dabei aus den Feldverteilungen mehrerer atmosphärischer Druckniveaus ermittelt, welche bis zu fünf Monate in die Vergangenheit zurückreichten. Jedoch enthalten nur die Feldverteilungen unmittelbar vor dem Einsetzen des Monsuns ein nennenswertes Vorhersagepotential. Keines der beiden getesteten Verfahren zeigte sich dem anderen überlegen.

*Unger* (2000) reduzierte die Anzahl potentieller Prediktoren nicht durch eine vorgeschaltete EOF-Analyse, sondern indem er nur Daten aus solchen Regionen zuließ, deren Relevanz bei Vorhersagen amerikanischer Temperatur- und Niederschlagsanomalien schon bekannt ist. Somit gelang es ihm, aus den Feldverteilungen der Prediktanden selbst, jenen des nordhemisphärischen 700-hPa-Geopotentials sowie aus quasi globalen SSTs 25 potentielle Prediktoren auszuwählen. Ein Screening-Verfahren (eine schrittweise multiple lineare Regression) sucht anschließend die Prediktoren mit dem größten Informationsgehalt heraus. Eine weitere Möglichkeit, das Überangebot an Einflußgrößen zu vermeiden, ist die Verwendung monovariater Datenreihen, welche den Zustand großräumiger Feldverteilungen repräsentieren. Häufig werden Indizes, z.B. der SOI, oder räumliche Mittelwerte als derartige Maße benutzt. Auf diese Weise berechneten *Lloyd-Hughes und Saunders* (2002) die zu erwartenden saisonalen Niederschlagsmengen Europas. Sie verwendeten einen winterlichen Niño 3 (ein Gebiet, das sich von 5° S bis 5° N sowie 90° bis 150° W erstreckt) SST-Index sowie die mittlere Abweichung der Temperatur des Seegebietes westlich von Irland zur Prognose der Niederschlagsanomalien im Frühjahr. Dabei stellten sie fest, daß die Güte der

Vorhersagen in hohem Maße vom ENSO-Phänomen abhängig ist. Nur in den Jahren ausgeprägter El Niño- bzw. La Niña-Ereignisse waren hinreichend genaue Vorhersagen möglich. Auf ein weiteres Ergebnis der Untersuchung von *Lloyd-Hughes und Saunders* sei an dieser Stelle ebenfalls hingewiesen: Die Autoren rechtfertigen die Verwendung einer bivariaten linearen Regression mit der Tatsache, daß sie Hinweise auf eine Symmetrie zwischen dem Vorzeichen von ENSO und den Niederschlägen in Europa gefunden hätten. Andere Autoren im Bereich der Langfristprognose gehen hingegen zumeist von der Nichtlinearität der untersuchten Zusammenhänge aus (vergl. z.B. *Silverman und Dracup*, 2000; *Folland et al.*, 2001; *Hsieh*, 2001). Obgleich bei keiner der aufgeführten Untersuchungen exakt dieselben klimatologischen Parameter in derselben Region der Erde untersucht wurden, ist dieses Ergebnis bemerkenswert.

In guter Übereinstimmung mit den Resultaten von *Lloyd-Hughes und Saunders* zeigte sich dagegen eine Arbeit von *Landman und Mason* (1999). Bei dem Versuch, die Gesamtmenge südafrikanischer Sommerniederschläge mittels CCA vorherzusagen, stellten auch sie fest, daß eine tatsächliche Prognosenleistung nur bei ENSO-Ereignissen zu erzielen ist. Alleinige Kenntnis über die Verteilung tropischer SST-Anomalien ist nicht ausreichend. Den dominanten Einfluß der ENSO-Telekonnektionen bestätigt auch die Untersuchung von *Shabbar und Barnston* (1996). Auch sie wendeten die CCA zur Prognose saisonaler Temperatur- und Niederschlagsanomalien Kanadas aus den globalen SST-Feldverteilungen sowie einigen atmosphärischen Parametern an. Dem Standard moderner CCA-Untersuchungen entsprechend wurden dabei vier aufeinanderfolgende 3-monatige Prediktor-Perioden verwendet, um sowohl das evolutionäre Verhalten als auch den momentanen Zustand der Atmosphäre zu berücksichtigen. Das Maximum der Prognosengüte liegt im Winter, wobei es im Falle der Temperatur bis in den Frühling hinein reicht. Eine Abhängigkeit vom Status des ENSO-Phänomens stellten sie jedoch nicht fest. Die wohl umfassendste Arbeit zum Vorhersagepotential der CCA stammt von *Barnston* (1994). Seine Untersuchung schließt quasi sämtliche Größen von Interesse sowohl seitens der Prediktoren als auch seitens der Prediktanden ein. In einigen Regionen der Nordhemisphäre lassen sich, so das Fazit, zu unterschiedlichen Jahreszeiten bestimmte Parameter (v.a. die Temperatur) derart gut prognostizieren, daß die Vorhersagen für einen potentiellen Verbraucher von Nutzen sein könnten. Die erfolgreichsten Prognosen sind für das Gebiet der PNA im Zeitraum Januar bis April zu erwarten, während im Spätsommer eher die subtropischen Regionen der Kontinente bevorzugt sein dürften. Speziell für europäische Temperaturprognosen gilt: Die höchste Güte versprechen Vorhersagen für den Juli mit etwa einem Monat Vorlaufzeit (engl. lead time bzw. lead) in einem Gebiet, welches das westliche Mittelmeer, das südöstliche Frankreich, Italien und das ehemalige nördliche Jugoslawien umfaßt. Ein zweites Gebiet reicht vom nördlichen Großbritannien über die Niederlande und den Nordwesten Deutschlands bis zum südlichen Dänemark. Für die Wintermonate fand *Barnston* generell nur ein geringes Potential für die Langfristprognose. Demgegenüber ergab eine vergleichbare Untersuchung von *Johansson et al.* (1998), daß das Maximum der Prognosenleistung für Vorhersagen der Temperaturanomalien in Europa im Winter liegt; ein sekundäres Maximum tritt im Sommer auf. Als größte Quelle der Vorhersagbarkeit stellte sich das 700-hPa-Geopotential, also die atmosphärische Zirkulation, heraus. Tropische SST-Anomalien enthielten hingegen keine wesentlichen Informationen für den europäischen Raum.

Ein vergleichsweise simples Verfahren zur langfristigen Temperaturprognose in den USA, die „optimal climate normals (OCN)“, entwickelten *Huang et al.* (1996a). Sie sind definiert als der Durchschnittswert einer bestimmten Anzahl vorangegangener

Jahre, wobei die Zahl der heranzuziehenden Fälle derart bestimmt wird, daß ihr Mittel die beste Vorhersage für das Folgejahr darstellt. Die Größe des Kollektivs variiert dabei sowohl räumlich als auch saisonal. Die Prognosengüte der OCN erreicht ihre Maxima im Spätwinter und im Sommer. Von *Wilks* (1996) wurden die OCN einer strengen statistischen Überprüfung unterzogen. Er führte sogenannte „resampling tests“ durch, um den Einfluß räumlicher Korrelationen und jenen der Vielfalt der vorhandenen Möglichkeiten (die Kollektivgröße ist a priori beliebig) zu bestimmen. Er bestätigte die Feldsignifikanz der OCN, stellte jedoch gleichermaßen fest, daß ein Großteil der Vorhersagbarkeit allein auf allmähliche Temperaturänderungen, also Trends, zurückzuführen ist.

Von den ebenfalls zur Langfristprognose genutzten statistischen Methoden, zu ihnen zählen u.a. die Zeitreihenanalyse (insbesondere die Autoregressionen) sowie die Diskriminanzanalyse, seien hier nur noch die sogenannten neuronalen Netze erwähnt. Sie versuchen, der Nichtlinearität des dynamischen Systems Atmosphäre gerecht zu werden, da vielfach auch im Bereich der Langfristprognose von der Nichtlinearität der relevanten Prozesse ausgegangen wird (s.o.). Ihre Funktionsweise „basiert auf biologischen Modellen des menschlichen Gehirns und auf der Art und Weise, wie es Muster erkennt und am Beispiel lernt“ (*Silverman und Dracup*, 2000). *Yuval* (2000) nahm einen Vergleich der Leistungsfähigkeit neuronaler Netze und dynamischer Modelle am Beispiel der Lorenz-Gleichungen vor, denen ein synthetisches rotes Rauschen hinzugefügt wurde, um die atmosphärischen Bedingungen zu simulieren. Die numerischen Integrationen des Modells zeigten sich dabei den Anfangsbedingungen des Systems gegenüber als wesentlich empfindlicher als das neuronale Netz. Auch bei der Vorhersage tropischer SST-Anomalien (Niño 3.4) erzielte das Verfahren mit Werten zwischen 0,8 und 0,5 (3 bis 12 Monate lead) hohe Korrelationen zwischen den Vorhersagen und den Beobachtungen.

## 1.2.2 Dynamische Verfahren

Dynamische Methoden zur kurzfristigen Klimasimulation bedienen sich der Gesetze der Physik. Dabei ist es nicht ausreichend, lediglich die Integrationen der atmosphärischen Kurz- bzw. Mittelfristmodelle zeitlich auszudehnen, sondern das Modell muß mit den Forcing-Parametern dynamisch gekoppelt sein (*Palmer und Anderson*, 1994). Zumeist beschränkt man sich auf SST-Anomalien, d.h. es werden globale Ozeanzirkulationsmodelle zur Kopplung verwendet. Bei vielen Untersuchungen wird die Kopplung durch die simple Vorgabe von Randwerten (für das Atmosphärenmodell) ersetzt. Speziell dann können auch andere Parameter, wie z.B. die Eisbedeckung oder die Bodenfeuchte, hinzugefügt werden. Die benötigten Randwerte werden dabei entweder durch über den gesamten Integrationszeitraum unveränderte Anfangsbedingungen oder durch die klimatische Erwartung approximiert (vergl. *Roads et al.*, 2001a). Eine andere Möglichkeit zur Vereinfachung des gesamten Klimasystems stellen hybride Verfahren dar. Sie verbinden dynamische Modelle des Ozeans mit statistischen Modellen der Atmosphäre. Nur die steigenden Computerkapazitäten ließen es zu, daß stets aufwendigere Ansätze entwickelt werden konnten. *Barnston et al.* (1994) stellen die bis zu diesem Zeitpunkt verwendeten Ansätze auf dem Gebiet der Modellierung vor: 1) einfache Modelle des Ozeans, welche mit statistischen Atmosphären gekoppelt sind; 2) globale Ozeanzirkulationsmodelle („general circulation models“ [GCMs]), die mit statistischen Atmosphären gekoppelt sind; 3) einfache

gekoppelte Modelle des Ozeans und der Atmosphäre; 4) globale Ozeanzirkulationsmodelle, die mit einfachen Modellen der Atmosphäre gekoppelt sind; 5) globale Ozeanzirkulationsmodelle, welche mit globalen atmosphärischen Zirkulationsmodellen (GCMs) gekoppelt sind. Einfache Modelle des Ozeans und der Atmosphäre verwenden gezielt gekürzte Versionen der vollständigen physikalischen Gleichungen, welche die Dynamik beider Teilsysteme sowie ihre Wechselwirkungen beschreiben. Einerseits vernachlässigen sie durch ihre Einfachheit für die Prognose möglicherweise entscheidende Zusammenhänge, andererseits vermeiden sie die problematische Simulation fehlerhafter Details, wie sie bei GCMs vorkommen kann (vergl. *Barnston et al.*, 1994; *Balzer et al.*, 1998). Selbst die komplexen gekoppelten Ozean-Atmosphären-GCMs stellen jedoch nur ein primitives Abbild der natürlichen Verhältnisse dar. Daher wird in überwiegender Mehrheit von einem in zwei Schleifen arbeitenden System Gebrauch gemacht, in deren erster Schleife zunächst die Meeresoberflächentemperaturen berechnet werden, die dann im zweiten Schritt das atmosphärische Modell antreiben (*Anderson et al.*, 1998). *Roads et al.* (2001a) gehen gar davon aus, daß alleinige Atmosphärenmodelle, angetrieben von unverändert belassenen SST-Anomalien, bis zu einem Zeitraum von mindestens zwölf Wochen bessere Vorhersagen liefern können als gekoppelte Systeme.

Der bei dynamischen Modellen ohnehin erhebliche Aufwand an Rechenzeit wird durch die Notwendigkeit mehrerer Integrationen noch deutlich erhöht. Nur durch die Verwendung eines Ensembles numerischer Integrationen lassen sich die vorhersagbaren saisonalen Klimafluktuationen wirksam von den chaotischen Wettereinflüssen trennen. Die eigentliche Vorhersage beruht in der Regel auf dem Ensemblemittel; die Streuung aller Läufe dient als ein Maß für die Stabilität der Ergebnisse. Je geringer sie ist, desto sicherer kann davon ausgegangen werden, daß die einzelnen Läufe nicht ausschließlich hochfrequentes Rauschen beschreiben.

*Davies et al.* (1997) führten eine Studie zur saisonalen Vorhersagbarkeit für ein Gebiet durch, das sowohl den Nordatlantik als auch Europa umfaßt. Mit einem sechs Mitglieder starken Ensemble des atmosphärischen Klimamodells HADAM1 des „Hadley Centre“ wurden sogenannte „hindcasts“ (der Untersuchungszeitraum liegt in der Vergangenheit) durchgeführt, angetrieben von tatsächlich beobachteten SST-Anomalien und dem Grad der Eisbedeckung. „Das Modell reproduzierte die Strukturen vieler wichtiger Moden der Varianz des Bodenluftdrucks dieser Region, inklusive der NAO, mit einer beträchtlichen Prognosengüte.“ Der Einfluß von ENSO erwies sich laut den Autoren in Europa als generell schwach, über dem Atlantik während El Niño-Ereignissen hingegen als etwas stärker ausgeprägt.

*Branković und Palmer* (1997) nahmen eine vergleichbare Untersuchung mit neun Mitgliedern des am „European Centre for Medium-Range Weather Forecasts (ECMWF)“ verwendeten Modells weltweit vor. Zu diesem Zweck wurde die horizontale Auflösung des Modells reduziert (T63L19). Wie auch bei einigen statistischen Verfahren zur Langfristprognose der Fall, hing die Güte der Vorhersagen stark von ENSO-Ereignissen ab. In den Jahren hoher Amplituden tropischer SST-Anomalien erwiesen sich sowohl der gesamte tropische Raum als auch einige außertropische Gebiete als gut prognostizierbar. Zu den außertropischen Regionen zählen der Bereich der PNA im Winter sowie der europäische Raum im Frühjahr. Fehlten ausgeprägte ENSO-Signale, so war weltweit praktisch kein Vorhersagepotential vorhanden. Eine weitere Frage von Bedeutung für dynamische Vorhersagen, die Anzahl der notwendigen Ensemblemitglieder, wurde ebenfalls untersucht: Zur Vorhersage saisonaler Temperatur- und Niederschlagsanomalien werden ungefähr 20 Mitglieder

benötigt, in den Tropen teils weniger. Wesentlich weniger differenzierte Ergebnisse erhielten *Stockdale et al.* (1998) unter Verwendung eines vollständig gekoppelten Modells des Ozeans und der Atmosphäre. Sie schränken ihr Fazit, daß das Modell eine Vorhersageleistung erbringen kann, weder geographisch noch saisonal ein.

Eine Methode, die sich in der lokalen Kurz- und Mittelfristvorhersage bewährt hat, wurde von *Fennessy und Shukla* (2000) getestet. Sie betteten ein hochauflösendes Regionalmodell in ein atmosphärisches GCM ein, um die Güte langfristiger Temperatur- und Niederschlagsprognosen für die USA zu steigern. Das globale Modell wird durch die wöchentliche Vorgabe von SST-Anomalien angetrieben und übergibt zwölfstündlich Randwerte an das Regionalmodell. Die Vorhersagen des Regionalmodells zeigten sich verlässlicher, wenngleich nur eine moderate Güte erzielt werden konnte.

Durch Kombination der Ergebnisse möglichst vieler verschiedener Modelle wurde in letzter Zeit versucht, das Leistungsvermögen dynamisch erstellter Langfristprognosen zu steigern. Hierfür werden die Resultate der einzelnen Modelle mittels statistischer Nachbearbeitung (engl. statistical postprocessing), im Regelfall wird die multiple lineare Regression verwendet, zu einem „Superensemble“ zusammengefaßt. Somit lassen sich die individuellen Stärken bzw. Schwächen der Modelle bei zukünftigen Vorhersagen optimal ausnutzen bzw. umgehen. Gleichzeitig können systematische Fehler korrigiert werden. Untersuchungen haben gezeigt, daß auf diesem Wege die Prognosegüte der Ensembles einzelner Modelle teils deutlich übertroffen werden kann. Aufgrund der Kürze der zur Verfügung stehenden Datenreihen ist die Stabilität der ermittelten Beziehungen jedoch eingeschränkt. Eine detaillierte Beschreibung der unterschiedlichen Vorgehensweisen sowie eine Aufstellung der zur Ermittlung der statistischen Beziehungen herangezogenen Modelle findet man bei *Krishnamurti et al.* (2000), *Kharin und Zwiers* (2002) sowie bei *Stefanova und Krishnamurti* (2002).

Betrachtet man die zahlreichen Arbeiten auf dem Gebiet der Langfristprognose eingehend, so gelangt man zu der Erkenntnis, daß sich die Güte dynamisch erzeugter Langfristprognosen derzeit kaum von der statistischer Ansätze unterscheidet. Beiderseits berichten die Autoren zumeist nur von einer geringen Prognosenleistung (in den außertropischen Regionen), teils von saisonalen bzw. räumlichen Einschränkungen. Auch die Möglichkeit, daß das Vorhersagepotential auf Jahre ausgeprägter ENSO-Ereignisse beschränkt ist, ziehen beide Seiten in Betracht. Eine Bestätigung für das in etwa ebenbürtige Leistungsvermögen beider Methoden liefern zwei Untersuchungen, welche einen direkten Vergleich anstellen: *Anderson et al.* (1999) simulierten die Reaktion außertropischer 700-hPa-Geopotentiale auf Änderungen tropischer SSTs. Dabei zeigte sich die CCA zwei numerischen Modellen leicht überlegen. In einem zweiten Schritt wurde versucht, die Grenze der Leistungsfähigkeit der dynamischen Modelle zu bestimmen: In Analogie zur Vorgehensweise bei der Abschätzung der internen Variabilität der Atmosphäre wird hierfür ein beliebiges Ensemblemitglied bestimmt, das als „Beobachtung“ fungiert. Das Mittel der übrigen Ensemblemitglieder wird dann an jener „Beobachtung“ verifiziert. Je geringer die Streuung der Integrationen, desto genauer könnte ein perfektes Modell prognostizieren. Unter diesen Umständen wären die Vorhersagen beider Modelle und die der CCA von gleicher Qualität. Eine Arbeit von *Folland et al.* (2001) ergab ähnliche Resultate, jedoch mit umgedrehten Vorzeichen. Sie sagten die zu erwartenden Regenschichten für den Nordosten Brasiliens mittels multipler Regression bzw. Diskriminanzanalyse sowie unter Verwendung des Modells HADAM2 des „Hadley Centre for Climate Prediction and Research (Bracknell, England)“ vorher. Dabei lagen die höchsten Korrelationen

zwischen den Vorhersagen und Beobachtungen bei 0,85 für das Modell, während durch die statistischen Methoden Werte um 0,7 erreicht werden konnten.

Solange es nicht gelingt, numerische Modelle derart zu verbessern, daß sie statistische Verfahren nennenswert übertrumpfen, hat die Anwendung empirischer Methoden ihre Berechtigung. Sie benötigen wesentlich geringere Computerkapazitäten und ermöglichen zusätzlich Einblicke in die dem Klimasystem zugrundeliegende Physik (*Johansson et al.*, 1998). Ein solches tiefgreifenderes Verständnis der Morphologie und der Physik regionalen Klimas könnte entscheidende Impulse bei der Vorhersage der Auswirkungen der globalen Erwärmung mit sich bringen (*Livezey*, 1990).

### 1.3 Verifikation operationeller Langfristprognosen

Nur die Auswertung operationeller, d.h. routinemäßig veröffentlichter Vorhersagen läßt eine Aussage darüber zu, inwieweit die am Entwicklungszeitraum abgeschätzte Prognosenleistung beim Übergang zum vollständig unabhängigen Kollektiv erhalten bleibt. Weder die Anwendung ausgefeilter statistischer Verfahren noch die Durchführung von „hindcasts“ zur Validierung numerischer Modelle können einen Einbruch vollständig verhindern.

*Livezey und Jamison* (1977) nahmen eine umfangreiche Verifikation sowjetischer Vorhersagen kommender Temperatur- und Niederschlagsanomalien vor. Die Ergebnisse wurden sowohl räumlich als auch saisonal unterteilt. Die Temperaturanomalien erwiesen sich für die arktische Region bzw. die Monate März und April als am besten vorhergesagt. Allgemein konnte die Güte von Persistenzvorhersagen jedoch nicht übertroffen werden. Ein Trend zur Verbesserung innerhalb des Verifikationszeitraumes war ebenfalls nicht feststellbar. Die Niederschlagsvorhersagen dagegen waren der Persistenz grundsätzlich überlegen, wenngleich auch sie nur eine moderate Prognosenleistung erzielen konnten.

*Wilks* (2000a) verifizierte die operationellen Langfristvorhersagen des „Climate Prediction Center (CPC)“ des nationalen Wetterdienstes der USA. Zu ihrer Erstellung werden die Ergebnisse einer CCA, der OCN sowie eines gekoppelten Modells verwendet (vergl. *Carson*, 1998). Jeweils zur Mitte eines jeden Monats werden Prognosen für den Mittelwert des kommenden Monats sowie für die 13 folgenden, sich überlappenden 3-Monats-Perioden herausgegeben, so daß der gesamte Vorhersagezeitraum 15 Monate umfaßt. Veröffentlicht werden diese Prognosen in Form von Wahrscheinlichkeiten dafür, daß die zu erwartenden Temperatur- bzw. Niederschlagsanomalien in die Klassen „warm“, „normal“ oder „kalt“ bzw. „naß“, „normal“ oder „trocken“ fallen werden. Dabei wird für jede Klasse die Eintreffwahrscheinlichkeit angegeben. Die Verifikation nahm *Wilks* auf zwei unterschiedliche Arten vor: Einerseits verwendete er „klassische“ skalare Größen, andererseits betrachtete er die vollständige gemeinsame Häufigkeitsverteilung der Vorhersagen und Beobachtungen mittels „diagnostischer Verifikation“ (engl. diagnostic verification). Sie ermöglicht die Aufdeckung spezieller Stärken und Schwächen der Vorhersagen und erlaubt dem Meteorologen somit, die gewonnenen Erkenntnisse zur Verbesserung der Vorhersagen einzusetzen (*Wilks*, 1995). Während die Verwendung der skalaren Größen nur eine minimale Prognosenleistung ergab, deckte die „diagnostische Verifikation“ das Vorhandensein einer (positiven) Auflösung für sämtliche Vorlaufzeiten der Temperatur auf. Die Niederschlagsprognosen wiesen hingegen wenig brauchbare Ansätze auf.

Eine weitere Problematik der langfristigen Vorhersagen des CPC zeigte ebenfalls *Wilks* (2000b) auf: Er berechnete den sich aus allen vorhergesagten saisonalen Wahrscheinlichkeiten notwendigerweise ergebenden Verlauf der monatlichen Temperaturabweichungen beispielhaft für eine Region im westlichen Wisconsin (USA). Dieser wies mehrere unrealistisch große Amplituden und Sprünge auf. Der Autor geht ganz allgemein davon aus, daß sich überlappende 3-Monats-Vorhersagen nicht mathematisch konsistent erstellt werden können.

Ausschließlich auf numerischen Integrationen beruhen die globalen Vorhersagen des „Scripps Experimental Climate Prediction Center (ECPC)“, die seit September 1997 regelmäßig verbreitet werden. Grundlage ist das globale Spektralmodell des „National Centers for Environmental Prediction (NCEP)“, welches von den Forcing-Parametern Eis- und Schneebedeckung, Bodenfeuchte und Meeresoberflächentemperaturen angetrieben wird. *Roads et al.* (2001b) werteten 104 dieser wöchentlich erstellten Vorhersagen aus, die bis zu zwölf Wochen im voraus gemacht werden. Neben dem systematischen Fehler wurden die Korrelationen zwischen Vorhersagen und Beobachtungen zur Beurteilung der Prognosen von Temperatur- und Niederschlagsverteilungen, Windgeschwindigkeiten, Bodenfeuchte sowie einem Feuer-Index herangezogen. Die Temperatur erwies sich in den Regionen des östlichen äquatorialen Pazifiks, des südlichen indischen Ozeans und des Atlantiks als am besten vorhersagbar. Auch für weite Teile Europas ergaben sich beachtliche Korrelationen. Praktisch keine Prognosenleistung konnte für das äquatoriale Afrika sowie für die mittleren Breiten der Südhalbkugel erzielt werden. Die Niederschlagsprognosen erbrachten nur in jenen tropischen Gebieten, die mit dem ENSO-Phänomen in Verbindung stehen, eine reale Prognosenleistung.

Durch die Zusammenfügung der wöchentlichen Vorhersagen zu 3-Monats-Perioden konnte z.B. in Europa die Prognosenleistung noch einmal gesteigert werden. Mit die ungünstigsten Voraussetzungen für die Langfristprognose in Europa scheinen in Deutschland gegeben zu sein, wo die Korrelationen weitverbreitet kleinere Werte aufwiesen als in den angrenzenden Staaten (*Roads et al.*, 2001a).

## 1.4 Langfristige Temperaturprognosen für Berlin

*Kuglin* (1992) entwickelte eine Methode, die Berliner Hochwintertemperaturen aus der vorwinterlichen Zirkulation im nordatlantisch-europäischen Raum vorherzusagen. Er entdeckte einen signifikanten Zusammenhang zwischen der Temperaturentwicklung in Berlin und dem herrschenden Druck in bestimmten, räumlich begrenzten Regionen des Untersuchungsgebietes, die er als „signifikante Gebiete“ bezeichnete. So läßt z.B. hoher Bodendruck über Labrador im November auf einen kalten Januar schließen. Als Prediktoren verwendete er den Bodendruck, das 500-hPa-Geopotential sowie die relative Topographie 500/1000 hPa, wobei die Signifikanzbetrachtungen für sämtliche Feldverteilungen unabhängig voneinander durchgeführt wurden. Später wurde das Verfahren zwecks regelmäßiger Veröffentlichung auf das gesamte Kalenderjahr ausgeweitet (*Malberg und Kuglin*, 1992).

Basierte die Identifikation der „signifikanten Gebiete“ zunächst nur auf dem Vergleich zweier Mittelwerte und den dazugehörigen Streuungen, so fügte ihr *Dettmann* (2000) die Bestimmung der Korrelationen zwischen den Druckwerten an den einzelnen Gitterpunkten der Vorzirkulationen und den Monatsmitteltemperaturen der Vorhersagemonate hinzu. Das bestehende Verfahren wurde zusätzlich durch die Ausweitung des Untersuchungsgebietes auf die Nordhemisphäre und durch die

Verlängerung des verwendeten Datensatzes von 38 auf 49 Jahre verbessert. Eine weitere wesentliche Neuerung stellte die Verwendung der linearen Regression zur Prognosenbestimmung dar, die an die Stelle von Intervallabschätzungen trat. Aufgrund der Ergebnisse eines fünf Jahre umfassenden Examinationskollektivs geht *Dettmann* davon aus, daß langfristige Temperaturprognosen für Berlin bis zu fünf Monate im voraus möglich sind. Die Anzahl der effektiv verifizierten Fälle reicht jedoch in noch keinem Fall aus, eine statistische Signifikanz der ermittelten Prognosenleistung (Reduktion der Varianz um bis zu 0,26) nachzuweisen.

## 1.5 Zu dieser Arbeit

Diese Arbeit stellt die Fortführung der bisherigen Forschungsbemühungen (vergl. Abschnitt 1.4) auf dem Gebiet der Langfristprognose für die Station Berlin-Dahlem vor. Ziel war es festzustellen, ob die bislang erzielte Prognosenleistung durch die Anwendung alternativer statistischer Methoden gesteigert werden kann. Dabei wurden z.T. Wege beschritten, die für die Station Berlin-Dahlem bislang noch nicht dokumentiert sind. Es kamen die PCR, die Spektralanalyse (eine spezielle Form der Zeitreihenanalyse), die Clusteranalyse sowie mehrere Varianten eines Analogverfahrens zur Anwendung. Zusätzlich wurde die schon bestehende Methode erweitert. Bei einigen dieser Verfahren wurde neben der Temperatur auch der Niederschlag als Prediktand mit in die Untersuchungen einbezogen.

Große Beachtung wurde insbesondere der Abschätzung der Prognosenleistung geschenkt. Hierfür wurden Verfahren angewendet, welche die zu erwartende Güte der Vorhersagen bereits bei der Erstellung der statistischen Beziehungen eingrenzen. Somit ist eine Bewertung der Leistungsfähigkeit der Vorhersagen auch ohne Verifikation am vollständig unabhängigen Kollektiv recht gut möglich. Dieses stellt einen auf dem Gebiet der Langfristprognose wünschenswerten Umstand dar, weil hier die Anzahl der unabhängigen Fälle (tatsächliche Vorhersagen) nur sehr langsam anwächst und die vorhandenen Datensätze zumeist zu kurz für die Verwendung eines Examinationskollektivs sind. Ob eine dauerhafte Vorhersageleistung erzielt werden kann, ließe sich daher ohne eine Abschätzung der Prognosengüte vermutlich erst nach vielen Jahren der Vorhersagepraxis mit Sicherheit sagen – die geringe Prognosenleistung langfristiger Witterungsvorhersagen verschärft diese Problematik zusätzlich. Es war jedoch das Bestreben vorhanden, ad hoc aussagekräftige Indizien für die Machbarkeit bzw. Unmöglichkeit statistischer Langfristprognosen für Berlin zu finden. Um möglichst Aussagen zur Signifikanz der erhaltenen Ergebnisse machen zu können, wurde stichprobenartig in einem Fall ein speziell auf die Situation zugeschnittener Test durchgeführt, welcher einen Vergleich mit dem Zufall anstellt.

Auf die angestrebte Ergänzung der verwendeten Prediktorensätze (z.B. 850-hPa-Temperatur) mußte leider verzichtet werden, da diese Daten nur für sehr begrenzte Zeiträume zur Verfügung stehen. Verfahren zur Abschätzung der Prognosenleistung sind jedoch um so genauer, desto umfangreicher das Entwicklungskollektiv ist. Um die Gefahr, daß sich die vorab ermittelte Güte später als nicht zutreffend erweist zu minimieren, wurde daher – wie schon bei *Dettmann* (2000) – allein auf die Zirkulation in Form des Bodendrucks und des 500-hPa-Geopotentials zurückgegriffen.