

11. Schlußbetrachtung und Ausblick

Ziel dieser Arbeit war es, das Vorhersagepotential für langfristige Witterungs- und insbesondere Temperaturprognosen an der Station Berlin-Dahlem abzuschätzen. Dabei ging es nicht so sehr darum, ein spezielles Verfahren vollständig auszuarbeiten und dessen operationellen Betrieb zu installieren, sondern aus der Vielzahl der zur Verfügung stehenden statistischen Methoden diejenige herauszufiltern, mit welcher auf Dauer die beste Vorhersagegüte zu erzielen ist. Zusätzlich war es das Bestreben, trotz der zu erwartenden geringen Vorhersageleistung Aussagen zur statistischen Signifikanz der erhaltenen Resultate machen zu können. Beide Ziele konnten zumindest teilweise erreicht werden.

Die unter Verwendung von Cross-Validation bzw. eines Examinationskollektivs ermittelten RV-Werte deuten darauf hin, daß die als „principal-components regression (PCR)“ bezeichnete Methode den anderen getesteten Verfahren überlegen ist. Zu diesen gehören mehrere Varianten sogenannter Analogverfahren, die Clusteranalyse sowie die multiple lineare Regression. Lediglich die Vorhersagen des von der Konstruktion simpelsten Analogverfahrens mit einer Vorlaufzeit von einem Monat erreichen eine zur PCR vergleichbare Vorhersagequalität. Der Grund für die Überlegenheit der PCR ist vermutlich die durch die Methode vorgenommene Zusammenfassung aller potentiellen Prediktoren zu nur wenigen neuen Variablen, was neben der reinen Datenkompression auch eine Glättung der Eingangsfelder zur Folge hat. Hierdurch fallen die Zufälligkeiten an den einzelnen Gitterpunkten nicht mehr ins Gewicht, sie können nicht modelliert werden. Speziell im Vergleich zur multiplen linearen Regression spielt auch die Anzahl der potentiellen Prediktoren eine entscheidende Rolle. Während im Falle der PCR lediglich zwischen vier und zehn Hauptkomponenten verwendet wurden, umfaßt der Pool der potentiellen Prediktoren im Falle der multiplen linearen Regression insgesamt 389 Gitterpunkte. Die Wahrscheinlichkeit einer rein zufällig guten Anpassung ist daher ungleich höher. Eine andere Möglichkeit stellen die Unterschiede der verwendeten Datensätze dar. Während die Untersuchung der PCR an Hand der vollständigen Feldverteilungen des nordhemisphärischen Bodendrucks sowie 500-hPa-Geopotentials des 54-jährigen Zeitraums 1946-1999 vorgenommen wurde, mußte bei den übrigen Verfahren auf eine lückenhafte Feldverteilung ausschließlich des Bodendrucks der Jahre 1909-1998 zurückgegriffen werden. Im Vergleich zur PCR standen also möglicherweise entscheidende Informationen nicht zur Verfügung. Dies würde auch die Diskrepanz der Vorhersageleistung zwischen PCR und Clusteranalyse erklären. Im Falle letztgenannter entfällt nämlich der ursprünglich genannte Vorteil, da hier ebenfalls eine vorgeschaltete EOF-Analyse durchgeführt wurde. Dennoch ist eher ein methodischer Vorteil der PCR anzunehmen, da das Vorhandensein hoher Interkorrelationen, d.h. die Existenz atmosphärischer Telekonnektionen, das Fehlen des 500-hPa-Geopotentials und einzelner Gitterpunkte des Bodendruckfeldes zumindest teilweise ausgleichen sollte. Ein Erklärungsansatz für das schlechte Abschneiden der Clusteranalyse könnte z.B. die mangelnde Differenzierung der einzelnen Clustermitglieder sein, die sich durch die notwendige Nebenbedingung einer bearbeitungsfähigen Gesamtanzahl von Clustern ergibt.

Wie erwartet verfehlen die RV-Werte sowohl der Niederschlags- als auch der Temperaturprognosen des besten Vorhersagemodells (PCR) ein statistisch signifikantes Niveau deutlich. Die im Falle der Temperatur maximal erzielten ca. 10% Reduktion der Varianz wären selbst bei Verwendung eines vollständig unabhängigen Examinationskollektivs erst bei einer über 1100 Fälle umfassenden Stichprobe mit einer Irrtumswahrscheinlichkeit von 5% signifikant von Null verschieden. Hier standen

hingegen pro Jahreszeit lediglich etwa 160 Fälle zur Verfügung, die zusätzlich noch durch Cross-Validation, d.h. nicht an Hand eines Examinationskollektivs, ermittelt wurden. Dennoch konnten durch die Anwendung einer speziellen „resampling“-Technik Anhaltspunkte dafür gefunden werden, daß zumindest im Falle der Temperatur ein reales Vorhersagepotential vorhanden ist. Lediglich zehn von insgesamt 1000 „bootstrap“-Realisierungen zur exakten Simulation der Untersuchungsmethode und der gegebenen Rahmenbedingungen wiesen eine höhere Anzahl von Monatskombinationen (Kombinationen zwischen den Monaten, für die der Prediktand vorhergesagt wird, und jenen Monaten, aus denen die verwendeten Prediktoren stammen) mit erwartungstreuen Einzelprediktoren auf, als der reale Prediktand Temperatur selbst. Die Stabilität der einzelnen Vorhersagemodelle wurde dabei sowohl an Hand des Kriteriums der Regularität als auch des Kriteriums der Unverzerrtheit gemessen und beurteilt. Daher läßt sich in Analogie zu den zweiseitigen klassischen Hypothesentests feststellen, daß die hier ermittelten Resultate mit einer statistischen Sicherheit von 98% als überzufällig angesehen werden können. Im Falle des Niederschlags fielen die Testergebnisse hingegen wesentlich weniger positiv aus. Hier ist nicht auszuschließen, daß es sich bei den vorliegenden positiven RV-Werten um reine Scheingüte handelt. Ob auch diese Größe auf Dauer erfolgreich vorherzusagen ist, kann nur an einem ausreichend langen Examinationskollektiv bestimmt werden. Gesichert scheint hingegen, daß die im Falle der Temperatur zu bestimmten Jahreszeiten vermutlich zu erreichenden 5-10% RV bei Niederschlagsprognosen nicht möglich sind. Im Bestfalle kann von nur wenigen Prozentpunkten ausgegangen werden. Es hat sich also bestätigt, daß sich langjährige und interne Variabilität der Atmosphäre im mitteleuropäischen Raum kaum voneinander unterscheiden. Hier wird das Klimasignal zu einem großen Teil vom sogenannten Rauschen überdeckt.

Um statistische Langfristprognosen in Zukunft verbessern zu können, empfiehlt es sich, den momentanen Gesamtzustand der Atmosphäre mit möglichst wenigen Variablen zu beschreiben und diese als Prediktoren für den weiteren Witterungsverlauf zu verwenden. Methodisch bietet sich hier vor allem die Kanonische Korrelationsanalyse (CCA) an. Wesentlich interessanter, da auf dem Gebiet der langfristigen Wettervorhersage bislang noch nicht erprobt, erscheint hingegen die Möglichkeit, die aktuelle Großwetterlage einer Region mit lediglich einem Parameter zu erfassen, dem sogenannten dynamischen Zustandsindex (engl. dynamic state index, DSI). Im räumlichen Mittel charakterisiert diese Größe die Abfolge von Hoch- und Tiefdruckgebieten, d.h. sie beschreibt die Wetterdynamik (*Névir und Brand, 2002*). Testreihen, bei denen der für das Gebiet Mitteleuropas bestimmte DSI als Prediktor für die Berliner Folgetemperaturen verwendet wurde, verliefen sehr vielversprechend. Für unterschiedliche Zeitverschiebungen konnten hier zu bestimmten Jahreszeiten lineare Kreuzkorrelationskoeffizienten ermittelt werden, deren Betrag eine Größenordnung von 0,5 erreichte. Dabei kam es jedoch in den Einzeljahren zu erheblichen Schwankungen. Für tiefgreifendere Untersuchungen reichte die vorliegende, die Jahre 1990-2001 umfassende Zeitreihe des DSI leider nicht aus. Dennoch geben die bislang erzielten Ergebnisse Anlaß zu der Hoffnung, daß sich die durch die PCR (mit den hier verwendeten Eingangsvariablen) zu erreichende Vorhersageleistung steigern läßt.