

# Kapitel 7

## Eine wetterlagenbasierte Klassifikation mit nachfolgender Regression

### 7.1 Einführung und Zielsetzung

Seit vielen Jahrzehnten werden Wetterlagen für die Einteilung der synoptischen Situation eines Tages genutzt. Der folgende Abschnitt 7.2 beschäftigt sich mit der Definition von Wetterlagen und deren historischer Entwicklung. Danach wird die hier genutzte Methode zur Ableitung von Wetterlagen beschrieben und deren Unterschiede zu den anderen Klassifikationen aufgezeigt (Abschnitt 7.3). Mit den vorhandenen Daten konnten die mittleren Felder für alle Klassen bestimmt werden. Nach der Durchführung der Diskriminanzanalyse konnte beschrieben werden, welche Felder den entscheidenden Einfluss auf die Wiedererkennung haben (Abschnitt 7.4). Nachdem ein jeder Tag in einer Region einer Wetterlage zugeordnet wurde, wurde für jede Station innerhalb der Region eine Screening-Regressions-Analyse durchgeführt. Dieses erfolgte mit dem Ziel, die bestmögliche Vorhersage der maximalen Ozonkonzentration zu erreichen (Abschnitt 7.5).

Die Erfahrungen von [Eder u. a. 1994] und [Spichtinger u. a. 1996] zeigen, dass es möglich ist, die großräumige meteorologische Situation eines Tages als Prädiktor für die Ozonsituation heranzuziehen. [Eder u. a. 1994] nutzen die *principal-components analysis* mit nachgeschalteter Clusteranalyse zur Einteilung der Tage. Alternativ nutzen [Spichtinger u. a. 1996] die Zuordnung der Tage zu Großwetterlagen. Mit der Entwicklung einer neuen Wetterlagenklassifikation für Ozonsituationen sollen die Erfahrungen der beiden oben genannten Arbeitsgruppen weiterentwickelt werden. Die neue Klassifikation soll erlauben, eine bestmöglich angepasste nachgeordnete Regressionsanalyse durchführen zu können. Es besteht die Hoffnung, somit im gesamten Spektrum der möglichen Ozonwerte eine gute Vorhersageleistung zu erzielen.

## 7.2 Historische Entwicklung

Die deutsche Sprache unterscheidet die Begriffe *Wetter* und *Witterung*. Hierbei bezeichnet der Begriff *Wetter* den unmittelbar wahrnehmbaren Zustand der Atmosphäre oder auch die hervorstechenden Merkmale einiger Stunden, bis zu einem Tag. Die *Witterung* bezieht sich dagegen auf längere Zeiträume mit typischen Wetterzuständen. Beide Begriffe beziehen sich auf das Geschehen an einem Ort. Der Begriff der *Großwetterlage* ist an die Definition der *Witterung* angelehnt. Beide beziehen sich auf ein atmosphärisches Geschehen mehrerer Tage. Im Gegensatz zur *Witterung* betrachtet eine *Großwetterlage* aber größere Gebiete, zum Beispiel den Raum Europa [Baur 1947, Baur 1948].

**Zitat 3** Unter *Großwetterlage* verstehen wir heute die mittlere Luftdruckverteilung (am Boden) über einem größeren Gebiet (etwa von der Größe Europas mit angrenzenden Meeresteilen) während eines Zeitraumes, in welchem die Luftdruckverteilung und die *Steuerung* (d. h. die Bahnen aufeinander folgender 24stündiger Drucksteig- und -fallgebiete) im wesentlichen unverändert blieben. [Baur 1947]

Die Grundlagen der Wetterlagenforschung sind in der Hauptsache von Baur gelegt worden. Die Hauptaufgabe in der Anfangszeit der *Großwetterforschung* von Baur war das Suchen nach der tatsächlichen Existenz solcher *Großwetterlagen* und der Möglichkeit, diese zu klassifizieren. Er war dabei nicht der Erste. Wie Baur in seinem Buch [Baur 1948] schreibt, fanden solche Versuche auch schon vorher statt (Hoffmeyer 1878 und Teisserence de Bort 1883; beide erwähnt in [Baur 1948], aber ohne ausführliche Zitatangabe). Die im Zitat erwähnte mittlere Luftdruckverteilung am Boden stimmt für die aktuellen Wetterlagenklassifikationen so nicht mehr. Bei den heutigen *Großwetterlagen* werden die Geopotentialflächen im 1000 und 500 hPa Druckniveau genutzt. [Dittmann u. a. 1995] nutzt beispielsweise neben den Druckfeldern auch die Windrichtung und Informationen zur Luftfeuchte. Auch sind heute Wetterlagenklassifikationen für kleinere Gebiete üblich [Fuentes Hutfilter 1998]. Eine Bemerkung ist notwendig: die von [Van Bebber 1891] vorgestellten Zugbahnen barometrischer Minima sind keine *Wetterlagen*.

### 7.2.1 Einteilung der Wetterlagen

Heutzutage können *Wetterlagen* nach zwei Methoden definiert werden. Das sind zum Ersten die *subjektiven* *Wetterlagen*, zu welchen auch die *Großwetterlagen* nach Baur gehören und zum Zweiten die *objektiven* *Wetterlagen*.

Die *subjektiven* *Wetterlagen* werden von Wissenschaftlern (z.B. Synoptikern) durch vergleichende Betrachtung klassifiziert. Inhomogenitäten sind somit sehr wahrscheinlich. Im Gegensatz dazu werden *objektive* *Wetterlagen* nach vorgegebenen Algorithmen klassifiziert.

## 7.2 Historische Entwicklung

---

### 7.2.2 Beispiele für Wetterlagenklassifikationen

In den vergangenen Jahrzehnten ist eine große Anzahl von Wetterlagenklassifikationen entwickelt worden. Die Bekanntesten sind hierbei die Großwetterlagen von Baur [Baur u. a. 1943, Baur 1947, Baur 1948]. Diese wurden weitergeführt und überarbeitet von [Hess und Brezowsky 1952] bis hin zu [Gerstengarbe u. a. 1999].

Eine ähnliche Klassifikation wurde von [Lamb 1972] für den Raum der Britischen Inseln entwickelt. Für Deutschland wurde vom Deutschen Wetterdienst [Dittmann u. a. 1995] eine neue Klassifikation entwickelt. Die letzten beiden Klassifikationen führen zu objektiven Wetterlagen. In der Doktorarbeit von [Fuentes Hutfilter 1998] wird eine für den Alpenraum geltende Klassifikation vorgestellt. In dieser Arbeit ist auch ein Vergleich zwischen mehreren multivariaten statistischen Methoden zu finden. Eine Übersicht über weitere Klassifikationen ist in der Arbeit von [Bissolli und Dittmann 2001] gegeben.

Neben den oben aufgezählten und in [Bissolli und Dittmann 2001] aufgelisteten Klassifikationen, werden auch von Enke und Mitarbeitern seit vielen Jahren objektive Wetterlagenklassifikationen entwickelt. Ausgehend von der Entwicklung von 10 Wetterlagen je Jahreszeit, zur Erklärung der für den europäischen Raum wesentlichen synoptischen Muster [Enke und Spekat 1997, Enke u. a. 1999], wurden für den Raum Sachsen zwei Klassifikationen entwickelt, die die mittleren Temperatur- und Feuchteverhältnisse, jeweils bis in zu zehn Klassen pro Jahreszeit, auftrennen [Enke 2001c].

### 7.2.3 Nutzen von Wetterlagen

Wetterlagen sind eine Möglichkeit, innerhalb der ausgewählten Bearbeitungsperiode, alle Tage z.B. anhand ihrer synoptischen Struktur in Klassen einzuteilen. Für die Betrachtung von realen Wetterereignissen sind dafür (unter Umständen) auch subjektive Verfahren nutzbar. Wenn die Wetterlagen im Rahmen des statistischen Downscalings von Wetter- oder Klimasimulationen genutzt werden sollen, sind subjektive Verfahren nicht mehr nutzbar. Gleiches gilt auch für automatisierte Prognoseverfahren, wie die hier vorgestellte Ozonprognosemethode.

### 7.2.4 Der Unterschied zwischen Wetterlagen und Ozonwetterlagen

Der in der meteorologischen Gemeinschaft bekannte Begriff Wetterlage bezieht sich auf die Einteilung von großräumigen meteorologischen Situationen in Klassen. Jeder Wetterlage können meteorologische Parameter zugeordnet werden. Der Begriff Ozonwetterlage lehnt sich an diese Vorstellung an. Im Gegensatz zu den bekannten Wetterlagen sind die Ozonwetterlagen für die mittleren Ozonverhält-

nisse optimiert. Die Idee dabei ist, dass bei bestimmten atmosphärischen Situationen die Ozonverhältnisse ähnlich sind.

### **7.3 Methode zur Bestimmung von Ozonwetterlagen**

Grundlage für die hier genutzte Methode zur Bestimmung von Wetterlagen ist das Vorliegen einer Einstiegsklassifikation einzelner Tage. Diese Klassifikation auf Basis der gemessenen Ozonkonzentrationen wird im Unterabschnitt 7.3.1 vorgestellt. Daneben ist das Vorhandensein eines Archivs numerischer Analysen meteorologischer Felder (Unterabschnitt 7.3.2) notwendig. Für die nächsten Schritte werden gemittelte Felder aller Klassen (Unterabschnitt 7.3.3) als Vergleichsgrundlage benötigt. Danach können mittels Screening-Diskriminanz-Analyse, aus dem Angebot potentieller Felder (Tabelle 7.1), die für die Objektivierung der Wetterlagen geeigneten Felder selektiert werden (Unterabschnitt 7.3.4). Die im Folgenden beschriebene Methode kann als eine Weiterentwicklung der Wetterlagenklassifikationen nach Enke angesehen werden.

#### **7.3.1 Die Clusterung der Tage entsprechend der Ozonsituation**

Das Umweltbundesamt (UBA) hat für die Arbeit sein Datenarchiv der in Deutschland gemessenen Ozondaten zur Verfügung gestellt. Basierend auf diesem Datensatz wurde für jeden Tag und für jede Station das Maximum der stündlichen Mittelwerte im Zeitraum von 6 bis 20 Uhr ermittelt.

Ausgehend davon, dass sich die Ozonsituation in den einzelnen Teilen Deutschlands unterscheiden kann, wurde Deutschland in drei Regionen unterteilt. Grundlage hierfür sind die Informationen der Stationsdatenbank des UBA. Als Erstes wurde Deutschland in eine Nord- und eine Südregion aufgeteilt. Hierbei liegt die Grenze bei  $\geq 50^{\circ}18' N$  (geografischer Breite) für die Zuordnung in die Nordregion. Diese wurde im Weiteren in eine Ost- und eine Westregion unterteilt. Hierbei liegt die Grenze bei  $\geq 11^{\circ} O$  (geografischer Länge) für die Zuordnung zur Ostregion (Abbildung: 7.1).

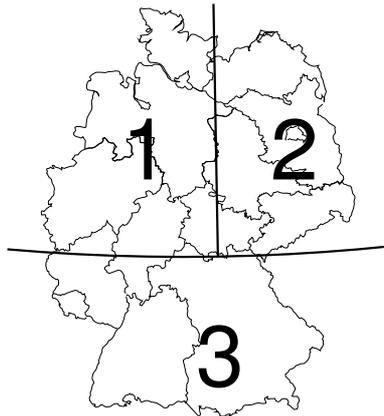
Innerhalb der drei Regionen wurden dann die Ozonmaxima eines Tages der Größe nach sortiert. Im nächsten Schritt wurden die 10 höchsten Tagesmaxima innerhalb der Region gemittelt. Die Anzahl 10 ist das bestmögliche Ergebnis von Versuchen und Abschätzungen. Diese hatten zur Aufgabe eine optimale Anzahl zu finden, die einen Kompromiss zwischen der Angabe des repräsentativen regionalen Ozonmaximums<sup>1</sup> und der Sicherheit nicht nur ein Einzelereignis zu betrachten,

---

<sup>1</sup>Die Nutzung der höchsten Werte der Region ergab gegenüber der Nutzung mittleren Situation der Region bessere Ergebnisse bei der nachfolgenden stationsbezogenen Prognose.

### 7.3 Methode zur Bestimmung von Ozonwetterlagen

---



**Abbildung 7.1.** Darstellung der 3 Regionen der Ozonwetterlagenklassen

bildet. Der Mittelwert der 10 höchsten Messwerte stellt dann den charakterischen Wert des Tages für die Region dar. Nach der Vorgabe der Klassenzahl (5) und der Grenzen ( $0 < K1 \leq 100$ ,  $100 < K2 \leq 120$ ,  $120 < K3 \leq 150$ ,  $150 < K4 \leq 180$ ,  $180 < K5 [\mu\text{g}/\text{m}^3]$ ) zwischen den Klassen, kann jeder Tag einer Klasse zugeordnet werden.

#### 7.3.2 Die meteorologischen Felder des DWD

Vom Deutschen Wetterdienst (DWD) wurden für diese Arbeit Datensätze vom Output numerischer Wetterprognosemodelle<sup>2</sup> zur Verfügung gestellt. Diese Datensätze umspannen ein Gebiet von  $10^\circ$  W bis  $29^\circ 45'$  O (geografischer Länge) und  $40$  bis  $70^\circ$  N (geografischer Breite), mit einer Auflösung von  $0,75^\circ$ . Für den Zeitraum vom 1.1.1997 bis 10.1.2002 liegen die diagnostischen Daten (Analysen) zu den Zeitpunkten 0, 6, 12, 18 Uhr UTC vor. Ab dem 11.1.2002 ist es eine Analyse von 0 Uhr UTC und die Prognosen für 0 UTC +6, +12, +18, +24, +30, +36, +42, +48, +54, +60 und +66 Stunden. Der Datensatz umfasst insgesamt bis zu 49 unterschiedliche Parameter (siehe Tab. F.1).

Aus dem vorliegenden Archiv wurden die Geopotential-Felder in 1000, 850, 700 und 500 hPa, die Temperatur-Felder in 850 und 500 hPa und die Felder der relativen Feuchte in 850 und 500 hPa des 12 Uhr UTC Termins gewählt. Diese Felder wurden im nächsten Schritt auf ein annähernd äquidistantes Gitter<sup>3</sup> interpoliert (siehe Abbildung 7.2). Die Wettersituation des 12 UTC Termins repräsentiert am besten die beobachteten Wettererscheinungen des jeweiligen Tages. Aus diesen acht ausgewählten Feldern wurden weitere zwölf, daraus abgeleitete Fel-

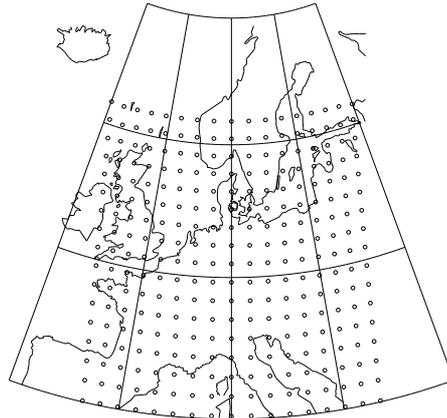
---

<sup>2</sup>Im Zeitraum 1997 bis Dezember 1999 liefen das GM- und EMD-Modell im operationellen Dienst des DWD. Im Dezember 1999 wurden beide Modelle vom GME-Modell abgelöst.

<sup>3</sup>Nach der Interpolation der Daten auf ein äquidistantes Gitter, können alle Werte auf dem Gitter mit gleicher Gewichtung genutzt werden.

## Eine wetterlagenbasierte Klassifikation mit nachfolgender Regression

---



**Abbildung 7.2.** Gitterpunkte des äquidistanten Gitters; die Konstruktion erfolgte von einem Zentralpunkt aus (51° N Breite und 10° O Länge) mit einem Gitterpunktsabstand von ca. 150 km

der, berechnet. Diese sind die horizontalen Geopotentialdifferenzen (HD) in N-S und W-O Richtung in 850 und 500 hPa, die Vorticity (VOR) in 1000, 850, 700 und 500 hPa, die Schichtdicke zwischen 1000 hPa und jeweils 850, 700 und 500 hPa und die Temperaturdifferenz (TD) zwischen den Druckflächen 850 und 500 hPa. In der Summe waren somit 20 meteorologische Felder verfügbar. In den Gleichungen 7.1, 7.2 und 7.3 sind die Formeln zur Berechnung der horizontalen Geopotentialdifferenzen und der Vorticity  $\zeta$  angegeben.  $GP_{i,j}$  bezeichnet das Geopotential am jeweiligen Gitterpunkt im zu bearbeitenden Druckniveau.

$$HD1_{i,j} = GP_{i-1,j} - GP_{i+1,j} \quad (7.1)$$

$$HD2_{i,j} = GP_{i,j-1} - GP_{i,j+1} \quad (7.2)$$

$i, j \dots$  Indizes der Gitterpunkte

$$\zeta = \vec{k} \cdot \vec{\nabla} \times \vec{v} \quad (7.3)$$

mit dem k-Vektor  $\vec{k}$  und  $\vec{v}$  dem Geschwindigkeitsvektoren. Die als potentielle Klassifikatoren zur Verfügung stehenden Felder sind in Tabelle 7.1 aufgelistet.

### 7.3 Methode zur Bestimmung von Ozonwetterlagen

---

**Tabelle 7.1.** Liste der potentiellen Prädiktorenfelder

| Nummer | Feld  | Kürzel      |
|--------|---|-------------|
| 1      | Geopotential 1000 hPa                           | GP 1000     |
| 2      | Geopotential 850 hPa                            | GP 850      |
| 3      | Geopotential 700 hPa                            | GP 700      |
| 4      | Geopotential 500 hPa                            | GP 500      |
| 5      | Temperatur 850 hPa                              | TP 850      |
| 6      | Temperatur 500 hPa                              | TP 500      |
| 7      | relative Feuchte 850 hPa                        | RH 850      |
| 8      | relative Feuchte 500 hPa                        | RH 500      |
| 9      | horizontale Geopotentialdifferenzen N-S 850 hPa | HD1 850     |
| 10     | horizontale Geopotentialdifferenzen W-O 850 hPa | HD2 850     |
| 11     | horizontale Geopotentialdifferenzen N-S 500 hPa | HD1 500     |
| 12     | horizontale Geopotentialdifferenzen W-O 500 hPa | HD2 500     |
| 13     | Vorticity 1000 hPa                              | VOR 1000    |
| 14     | Vorticity 850 hPa                               | VOR 850     |
| 15     | Vorticity 700 hPa                               | VOR 700     |
| 16     | Vorticity 500 hPa                               | VOR 500     |
| 17     | Schichtdicke 1000/850 hPa                       | RT 1000/850 |
| 18     | Schichtdicke 1000/700 hPa                       | RT 1000/700 |
| 19     | Schichtdicke 1000/500 hPa                       | RT 1000/500 |
| 20     | Temperatur-Differenz 850 - 500 hPa              | TD 850-500  |

#### 7.3.3 Die Mittlung der Felder

Mit den zwei genannten Datenarchiven (vgl. Unterabschnitte 7.3.1, 7.3.2) wurden, getrennt für jede Klasse, die mittleren Felder berechnet. Das heißt zum Beispiel, dass für jede Ozonklasse, in jeder der drei Regionen alle Tage mit der Klasse 1 gemittelt wurden. Am Ende sind für jede Ozonklasse in jeder Region 20 mittlere meteorologische Felder vorhanden. Beispielhaft können drei der 20 mittleren Felder der Region 1 in Abbildung 7.3 betrachtet werden.

Diese mittleren Felder sind der Ausgangspunkt für die Objektivierung. Sie können als die Musterozonwetterlagen angesehen werden. Bei der Wiedererkennung werden diese Felder als Bezugspunkt herangezogen.

#### 7.3.4 Objektivierungsverfahren

Die Grundfrage der Objektivierung ist, wie die Ähnlichkeit zwischen vorgegebenen Mustern (hier der jeweils aktuelle Tag in Bezug auf die gemittelten Felder) bestimmt werden kann. Die Ähnlichkeit bzw. Unähnlichkeit wird über ein Distanzmaß beschrieben. Hierfür gibt es eine große Anzahl von Ansätzen. In [Enke und Spekat 1997] wird ein normalisiertes euklidisches Distanzmaß  $D$  beschrieben (Gleichung 7.4). Mit Hilfe dieses Maßes ist es möglich, die Distanzen zwischen

## Eine wetterlagenbasierte Klassifikation mit nachfolgender Regression

Feldern zu berechnen; im konkreten Fall zwischen mittleren Druckmustern oder daraus abgeleiteten Feldern (Tab. 7.1) und aktuellen (d.h. täglichen) Feldern. Die kleinste Distanz zeigt an, zu welcher Klasse der jeweilige Tag gehört.

$$D = \frac{\sum_{i=1}^m (|dp_i - d_i|)}{\sum_{i=1}^m \max(|dp_i|, |d_i|)} \quad (7.4)$$

wo  $dp_i$  = der Gitterpunktwert eines gegebenen Feldes (wetterlagenspezifische gemittelte Felder) und  $d_i$  = der Gitterpunktwert des Vergleichsfeldes (des aktuellen Feldes) ist. In Gleichung 7.4 ist eine Normierung der unterschiedlichen Ausgangsfelder durch die Division durch  $\sum_{i=1}^m \max(|dp_i|, |d_i|)$  eingearbeitet, so dass beliebige Ausgangsfelder zur Diskriminierung verwendet werden können.

Für die hier vorliegende Klassifizierung wurde nicht die Gleichung 7.4 sondern der RMSD (Root Mean Squared Distance) als Distanzmaß genutzt (siehe Gleichung 7.6). Der RMSD hat gegenüber der obigen Methode den Vorteil, dass durch die Quadrierung der Gitterpunktabstände große Differenzen stärker gewichtet werden<sup>4</sup>. Durch diese Änderung musste die in Gleichung 7.4 implizite Normierung der Felder durch eine vorgeschaltete Normierung ersetzt werden.

### Die Normierung

Die Skalen der meteorologischen Felder, die für die Klassifikation potentiell zur Verfügung stehen, haben sehr unterschiedliche Größenordnungen. So haben zum Beispiel die Werte des Geopotentialfeldes in 500 hPa die Größenordnung von  $10^3$ , während die Vorticity eine Größenordnung von  $10^1$  hat. Damit die Felder vergleichbar werden, müssen sie normiert werden. Hierzu wird für das jeweilige Feld Mittelwert  $\bar{x}$  und Spread  $S$  (Distanz zwischen Minimum und Maximum) berechnet. Zur Normierung werden die mittelwertbereinigten Gitterwerte durch den Spread dividiert (Formel 7.5).

$$x_{n_{i,j}} = \frac{x_{i,j} - \bar{x}}{S} \quad (7.5)$$

### Screeningprozedur zur Selektion der verwendeten Felder

Nach der Normierung aller Felder mittels Gleichung 7.5 wurde eine Screening-Diskriminanz-Analyse durchgeführt. Für jeden Tag wurden die Differenzen zwischen den mittleren Feldern und den Feldern des aktuellen Tages (für jeden Gitterpunkt) berechnet und daraus das Distanzmaß  $D$  nach Gleichung 7.6 bestimmt.

$$D = RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n R_i^2} \quad (7.6)$$

---

<sup>4</sup>Dieses entspricht somit dem verwendeten Fehlermaß RMSE.

### 7.3 Methode zur Bestimmung von Ozonwetterlagen

---

Hierin ist  $R_i$  die Differenz zwischen den normierten mittleren Vergleichsfeldern (Gl. 7.5) und den Feldern des zu diskriminierenden Tages an den Gitterpunkten und  $n$  die Anzahl der Gitterpunkte. Wird für mehr als ein Feld das Distanzmaß berechnet, werden die einzelnen Distanzmaße der Felder addiert. Dieses ist durch die vorgeschaltete Normierung möglich.

Der jeweilige Tag wird der Klasse zugeordnet, die das geringste Distanzmaß aufweist. Danach wird geprüft, ob die Zuordnung mit der Originalklassifikation übereinstimmt. Im ersten Schritt wird jedes Feld aus Tabelle 7.1 einzeln genutzt. Das Feld, das die beste Wiedererkennungsrates hat, wurde als erster Diskriminator ausgewählt. Im zweiten Schritt wurde das Feld gesucht, welches in Kombination mit dem ersten Diskriminator (meteorologischen Feld) die beste Wiedererkennungsrates aufweist. Die Screening-Prozedur wurde solange durchgeführt, bis entweder eine vorgegebene maximale Anzahl von vier Feldern erreicht wurde oder sich die Wiedererkennungsrates nur weniger als ein vorgegebenes Maß verbessert hat.

#### Die endgültige Zuordnung der Einzeltage zu den Wetterlagen

Mit der oben beschriebenen Screening-Prozedur wurde für jede Jahreszeit eine Kombination an Feldern gefunden, mit der im nächsten Schritt die endgültige Zuordnung durchgeführt wurde. Hierbei wurden für jeden Tag die aufsummierten Distanzen (Gleichung 7.6), der maximal vier gewählten Felder, zwischen den normierten mittleren Feldern und den normierten Feldern des zu untersuchenden Tages berechnet. Jeder Tag wurde der Wetterlagenklasse zugeordnet, die die geringste euklidische Distanz aufwies.

#### 7.3.5 Die Fuzzifizierung

Die bisher vorgestellte Methode der objektiven Wetterlagen Klassifikation führt mit Hilfe des Distanzmaßes RMSD eine eindeutige Zuordnung der Wettersituation eines jeden Tages zu einer konkreten Ozonklasse durch. Diese Vorgehensweise hat sowohl Vor- als auch Nachteile. Durch die Zuordnung kann für jeden Tag eine eindeutige Aussage getroffen werden. Andererseits ist die Wettersituation eines Tages immer ein Gemisch aus verschiedenen Wetterlagen. Deshalb führt eine eindeutige Zuordnung zu einer bestimmten Klasse zu Ungenauigkeiten bei deren Anwendung. Es erweist sich als sinnvoll, zusätzlich eine unscharfe Zuordnung der Wetterlagen vorzunehmen. Zu diesem Zweck wird die Zuordnung jedes Tages zu einer Klasse nicht eindeutig gewählt, sondern es wird der approximierter prozentuale Anteil der Wetterlagen an dem jeweiligen zu klassifizierenden Tag berechnet. Im ersten Schritt wird für jeden Tag, zu jeder Wetterlagenklasse das Distanzmaß ( $D$ ) berechnet (Gleichung 7.6). Aus allen Distanzmaßen eines Tages wird das Maximum bestimmt.

$$D_{max} = \max(D_i); \text{ für } i = 1, n \text{ mit } n = \text{Klassenanzahl} \quad (7.7)$$

## Eine wetterlagenbasierte Klassifikation mit nachfolgender Regression

Im nächsten Schritt wird die Summe der Differenzen der einzelnen Distanzmaße zum Maximum ermittelt.

$$D_{sum} = \sum_{i=1}^n (D_{max} - D_i); \text{ mit } n = \text{Klassenanzahl} \quad (7.8)$$

Mit  $D_{max}$  und  $D_{sum}$  und einem Fuzzifizierungsexponenten  $F$  erfolgt eine Normierung der einzelnen Distanzmaße.

$$D_{nor_i} = \left( \frac{D_{max} - D_i}{D_{sum}} \right)^F; \text{ für } i = 1, n \text{ mit } n = \text{Klassenanzahl} \quad (7.9)$$

Der Fuzzifizierungsexponent  $F$  ist dabei ein Maß für die Zuordnungsschärfe. Je größer er gewählt wird, umso eindeutiger wird die Zuordnung zur dominierenden Klasse. Mit  $F \rightarrow \infty$  steigt der Anteil der dominierenden Klasse auf 100%. Die Berechnung des prozentualen Anteils  $r_i$  der Klassen zur untersuchten Wetterlage erfolgt nach Gleichung 7.10 und 7.11. In der Arbeit wurde  $F$  gleich 6 gesetzt.

$$D_{sumnor} = \sum_{i=1}^n D_{nor_i}; \text{ mit } n = \text{Klassenanzahl} \quad (7.10)$$

$$r_i = \frac{D_{nor_i}}{D_{sumnor}} \times 100\%; \text{ für } i = 1, n \text{ mit } n = \text{Klassenanzahl} \quad (7.11)$$

### 7.3.6 Unterschiede zu anderen Klassifikationen

Jede Klassifikation wurde mit einer bestimmten Absicht entwickelt. Baur suchte zum Beispiel nach einer Möglichkeit, die Vorhersageleistungen bei der Kurz-, Mittel- und Langfristprognose zu verbessern [Baur 1956, Baur 1958]. Bei den neueren Klassifikationen steht meist die Nutzung für das statistische Downscalen von Klimasimulationen (GCM) im Vordergrund [Enke und Spekat 1997, Dittmann u. a. 1995, Fuentes Hufilter 1998, Enke u. a. 1999, Enke 2001c].

Entgegen der üblichen Einteilung der Wettersituation im Allgemeinen wird bei [Enke 2001c] eine Einteilung für einzelne Wetterelemente durchgeführt. Dies sind das *Temperaturregime* für die Temperaturverhältnisse und das *Feuchteregime* für die Niederschlags- und Luftfeuchteverhältnisse. Die hier entwickelte Klassifikation soll der Beschreibung der Ozonverhältnisse in den Regionen dienen. Daher ist die Klassifikation für andere Elemente, wie zum Beispiel für die Wetterelemente Temperatur oder Niederschlag, nicht nutzbar! Einschränkend muss auch erwähnt werden, dass keine Klassifikation eine perfekte Auftrennung zwischen den Klassen erreichen kann. Zur Verbesserung der primären Klassifikation wird hier nachträglich eine klassenspezifische Regression durchgeführt. Um den Ozonwert für einen Tag in der gewählten Region genauer vorherzusagen, nutzt man durch eine Screening-Regressions-Analyse gefundene Korrelationen. Dieser Ozonwert wird dann für die endgültige Einordnung des Tages in eine Klasse genutzt.

### 7.4 Ergebnisse bei der Ableitung von Ozonwetterlagen

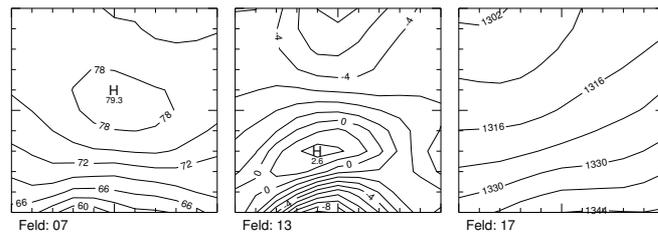
Mit den oben genannten Vorgaben konnte nun die Diskriminanz-Analyse durchgeführt werden. Das Qualitätskriterium war die bestmögliche Anpassung an die vorgegebene Clusterung. In allen Regionen wurden die gleichen Felder selektiert. Dieses sind die relative Feuchte in 850 hPa (7), die Schichtdicke 1000/850 hPa (17) und die Vorticity in 1000 hPa (13).

In der Abbildung 7.3 werden drei atmosphärische Felder für alle Klassen, beispielhaft an der Region 1, dargestellt. Dadurch lassen sich die Unterschiede zwischen den Klassen gut erläutern. Eine kleine Anmerkung zu der Darstellung ist nötig. Die Isolinienbilder zeigen, auf ein Rechteckformat verzerrte, Felddarstellungen der  $9 \times 9$  inneren Gitterpunkte der Abbildung 7.2. Diese wurden via Glättung (Tiefpassfilter) aus den  $11 \times 11$  inneren Gitterpunkten der mittleren Felder berechnet.

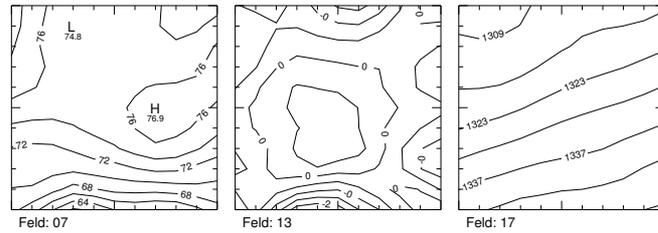
Die relative Feuchte nimmt von der Klasse 1 (niedrige Ozonwerte) zur Klasse 5 (sehr hohe Ozonwerte) ab. Das ist mit den Ergebnissen früherer Untersuchungen vergleichbar. Auch die Zunahme der Schichtdicke von Klasse 1 zu Klasse 5 ist durch die physikalische Erklärung der Schichtdicke als eine direkt proportionale Größe zur Temperatur sehr einfach in Verbindung zu den Ozonwerten zu bringen. Schwieriger ist die Interpretation der Informationen des Vorticityfeldes. Dort nehmen die Werte von Klasse 1 (positive Werte) zu Klasse 5 (negative Werte) ab. Das bedeutet, dass bei Klasse 1 aufsteigende Vertikalbewegungen stattfinden, während bei Klasse 5 absinkende Vertikalbewegungen vorherrschend sind. In Analogie zu den Vertikalbewegungen bei Hoch- und Tiefdruckgebieten, sind diese Bewegungen gut verständlich. Die Informationen aller drei Felder sind als Komplex für die Zuordnung zu einer Klasse entscheidend.

Wie aus Tabelle 7.2 ersichtlich, können über die einfache Wetterlagenklassifikation zwischen 42 und 50 % der Tage richtig zugeordnet werden. Mit einer erlaubten Fehlklassifikation vom  $\pm$  einer Klasse sind es zwischen 86 und 88 %.

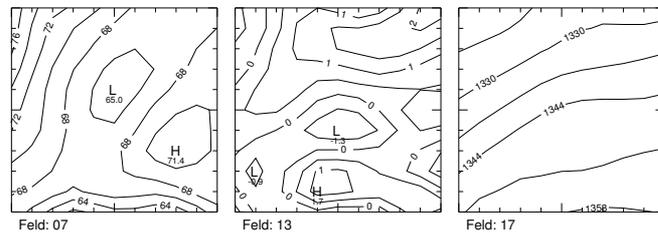
## Eine wetterlagenbasierte Klassifikation mit nachfolgender Regression



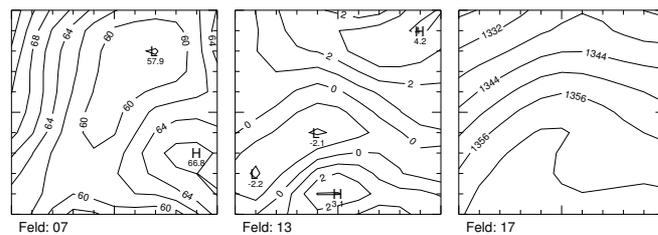
(a) Klasse 1



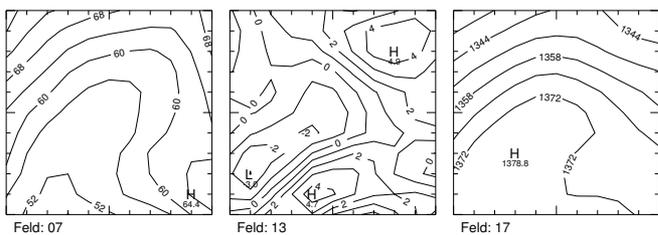
(b) Klasse 2



(c) Klasse 3



(d) Klasse 4



(e) Klasse 5

**Abbildung 7.3.** Vergleich der Musterozonwetterlagen der Region 1; dargestellt sind die drei zur Wiedererkennung selektierten Felder: relative Feuchte 950 hPa (Feld 7), Vorticity 1000 hPa (Feld 13) und Schichtdicke 1000/850 hPa (Feld 17) (Monate Mai bis August)

## 7.5 Eine stationsbezogene Screening-Analyse

---

**Tabelle 7.2.** Übersicht zu den gewählten Feldern und die Qualität der Anpassung (für die Feldkürzel siehe Tabelle 7.1). Die Trefferquote zeigt, wie viel Prozent der Tage richtig klassifiziert wurden (direkt) bzw. an wie viel Prozent der Tage bei einer Zuordnung mit einem Fehler von  $\pm$  einer Klasse richtig klassifiziert wurde (indirekt). Dabei bezieht sich der Bereich WL auf die reine Ozonwetterlagenklassifikation und der Bereich Reg auf die Klassifikation mit nachgeordneter multipler Regression.

| Region | Felder    | Trefferquote WL |          | Trefferquote Reg |          |
|--------|-----------|-----------------|----------|------------------|----------|
|        |           | direkt          | indirekt | direkt           | indirekt |
| 1      | 7, 17, 13 | 42 %            | 88 %     | 59 %             | 98 %     |
| 2      | 7, 17, 13 | 50 %            | 88 %     | 61 %             | 99 %     |
| 3      | 7, 17, 13 | 42 %            | 86 %     | 62 %             | 99 %     |

Dieser vorläufigen Klassifikation wird dann eine Regression für jede Region und Klasse nachgeschaltet. Eine Screening-Regressions-Analyse musste dafür aus einem Pool von 59 potentiellen Prädiktoren die besten Prädiktorenkombinationen wählen (siehe Tabelle F.3). Diese Prädiktoren sind die innerhalb der jeweiligen Region gemittelten Gitterpunktsvorhersagen des GME-Modells des Deutschen Wetterdienstes und der Vortagswert der Maxima der Ozonkonzentration (die Persistenz). Neben diesen 59 Prädiktoren wurde die fuzzifizierte Wahrscheinlichkeit, der durch die Wetterlagenbestimmung gewählten Klasse, als möglicher Prädiktor zugelassen. Die Screening-Regressions-Analyse wählte als Prädiktoren die Temperatur in 2 m Höhe, die Taupunkttemperatur, die relative Topografie 850/1000 hPa und die prozentuale Wahrscheinlichkeit der bestimmenden Wetterlage (Gleichung 7.11) aus. Die Optimierung erfolgte in Richtung der Minimierung des Vorhersagefehlers des regionalen Ozonwertes.

Mit dieser gefundenen multiplen Regression wird für jede Region ein Vorhersagewert berechnet. Anhand dieses Wertes erfolgte die entgeltliche Einordnung des jeweiligen Tages in eine Ozonklasse. Hierdurch konnte die Wiedererkennung auf 59, 61 und 62 % in den Regionen bei der genauen Zuordnung und auf 98-99 % bei einer erlaubten Fehlklassifikation von  $\pm$  einer Klasse gesteigert werden.

## 7.5 Eine stationsbezogene Screening-Analyse

Ziel der Arbeit ist eine stationsbezogene Vorhersage der maximalen Ozonkonzentrationen für die Zeithorizonte aktueller Tag, 1. und 2. Folgetag. Mit dem Aufbau einer Ozonklassifikation für drei Regionen konnte eine bedingte Regressions-Analyse (*stratified regression technique*) durchgeführt werden. Mit Hilfe einer Screening-Regressions-Analyse wurde der potentiell mögliche Satz von 60 Prädiktoren (siehe Tabelle F.3) auf zehn Prädiktoren reduziert. Aus diesen konnten dann in einer zweiten Screeningprozedur die besten Prädiktoren, maximal 4 je Klasse,

## Eine wetterlagenbasierte Klassifikation mit nachfolgender Regression

für die stationsbezogene Vorhersage gewählt werden. Für diese Prädiktoren wurde im Folgenden für die Jahre 1997, 1998, 2000 und 2001 die Screening-Regressions-Analyse durchgeführt.

**Tabelle 7.3.** Reduzierte Liste der möglichen Prädiktoren für die stationsbezogene Vorhersage

| Nr. | Name / Beschreibung                              |                            |
|-----|--|----------------------------|
| 1   | Persistenz                                       | $[\mu\text{g}/\text{m}^3]$ |
| 5   | Temperatur (2m), 12 UTC                          | [K]                        |
| 8   | Temperatur (2m), Tagesmaximum                    | [K]                        |
| 14  | kurzw. Strahlungsbilanz (TAO)                    | $[\text{W}/\text{m}^2]$    |
| 27  | rel. Feuchte in 850 hPa, 12 UTC                  | [%]                        |
| 28  | rel. Feuchte in 950 hPa, 12 UTC                  | [%]                        |
| 44  | pot. Temperatur in 950 hPa, 12 UTC               | [K]                        |
| 46  | Auslösetemperatur, 12 UTC                        | [K]                        |
| 58  | Differenz Temperatur (2m) 12 UTC Gestern - Heute | [K]                        |
| 60  | regionale Prognose Ozonmaximum                   | $[\mu\text{g}/\text{m}^3]$ |

## 7.6 Ergebnisse der stationsbezogenen Prognose

Die abgeleiteten Regressionsfunktionen konnten am Jahr 2002 getestet werden. Die Ergebnisse sind in der Tabelle 7.4 aufgelistet. Zum Vergleich wurden in Tabelle 7.5 die Ergebnisse des im Jahre 2002 am UBA eingesetzten Ozonprognose-systems aufgelistet.

Erkennbar ist, dass das hier vorgestellte System eine Verbesserung der Vorhersageleistung mit sich bringt. Dies gilt sowohl für die Prüfung ohne Schwellenwert (allen Prognosen werden geprüft) als auch für die schwellenwertabhängige Prüfung (wenn Prognose oder Messung den Schwellenwert überschreitet). Systembedingt unterscheidet sich die Anzahl der Fälle bei beiden Prognosen und daher ist der RMSE der Persistenzvorhersage nicht exakt gleich. Grund dafür ist eine Änderung der Stationsliste, für die die Prognose durchgeführt wurde. Einige Stationen wurden im Jahr 2002 geschlossen und sind daher nicht mehr in der neuen Prognosenliste, dafür wurden einige neue Stationen in die Prognosenliste aufgenommen.

In der letzten Spalte wird die Anzahl der Fälle in Prozent dargestellt. So treten an mehr als einem Viertel der Tage Werte mit über  $120 \mu\text{g}/\text{m}^3$  auf, aber nur an ca. 0,8 % der Tage Werte mit über  $180 \mu\text{g}/\text{m}^3$ . Die prozentualen Angaben beziehen sich auf die Summe aller Tage der einzelnen Stationen. Es muss angemerkt werden, dass die Vorhersageleistung (dargestellt durch den RV) sich bei der am UBA eingesetzten Routine im Jahr 2002 deutlich gegenüber den Vorjahren verschlechtert hat (siehe dafür Unterabschnitt 5.2.1 und Tabelle 5.2).

## 7.6 Ergebnisse der stationsbezogenen Prognose

Die Entwicklung erfolgte anhand der Daten für den ersten Tag. Eine spezielle Anpassung für die Vorhersage der beiden Folgetage wurde nicht durchgeführt. Vielmehr wird der Idee gefolgt, dass die Vorhersage des aktuellen Tages die beste Information für den aktuellen Tag liefert und daher als Persistenz für den Folgetag genutzt werden kann. So wird dann auch die Vorhersage für den 1. Folgetag als Persistenzwert für den 2. Folgetag genutzt. Erkennbar ist, dass die wetterlagenbasierte Prognose eine deutliche Verbesserung der langfristigen Kurzfristprognose (2. Folgetag) erbringt.

**Tabelle 7.4.** Verifikation der Vorhersageleistung der stationsbezogenen Ozonvorhersage gemittelt über alle Stationen (321); die Daten beziehen sich auf die Monate Mai bis Juli 2002; Ergebnisse der bedingten Regression nach der regionalen Ozonwetterlagenbestimmung (WL)

| Schwellenwert: |       | ohne                             |       |          |      |       |  |
|----------------|-------|----------------------------------|-------|----------|------|-------|--|
|                | Bias  | RMSE Prog                        | Bias  | RMSE Per | RV   | %     |  |
| aktueller Tag  | 1.0   | 17.0                             | -0.43 | 29.0     | 65.5 | 100.0 |  |
| 1. Folgetag    | 2.9   | 20.3                             | -0.58 | 36.3     | 68.7 | 100.0 |  |
| 2. Folgetag    | 3.6   | 21.2                             | -1.14 | 39.2     | 70.8 | 100.0 |  |
| Schwellenwert: |       | 120 [ $\mu\text{g}/\text{m}^3$ ] |       |          |      |       |  |
|                | Bias  | RMSE Prog                        | Bias  | RMSE Per | RV   | %     |  |
| aktueller Tag  | -5.8  | 21.8                             | -14.0 | 33.7     | 58.1 | 25.7  |  |
| 1. Folgetag    | -6.3  | 25.1                             | -25.2 | 44.6     | 68.3 | 26.5  |  |
| 2. Folgetag    | -2.8  | 25.7                             | -29.6 | 49.6     | 73.2 | 28.5  |  |
| Schwellenwert: |       | 150 [ $\mu\text{g}/\text{m}^3$ ] |       |          |      |       |  |
|                | Bias  | RMSE Prog                        | Bias  | RMSE Per | RV   | %     |  |
| aktueller Tag  | -10.3 | 27.7                             | -22.4 | 39.6     | 51.1 | 6.3   |  |
| 1. Folgetag    | -10.8 | 30.6                             | -44.0 | 56.5     | 70.7 | 6.7   |  |
| 2. Folgetag    | -7.6  | 31.6                             | -53.5 | 65.7     | 76.9 | 7.0   |  |
| Schwellenwert: |       | 180 [ $\mu\text{g}/\text{m}^3$ ] |       |          |      |       |  |
|                | Bias  | RMSE Prog                        | Bias  | RMSE Per | RV   | %     |  |
| aktueller Tag  | -20.6 | 33.5                             | -40.9 | 54.7     | 62.3 | 0.8   |  |
| 1. Folgetag    | -26.7 | 38.4                             | -78.8 | 86.4     | 80.3 | 0.8   |  |
| 2. Folgetag    | -26.0 | 45.4                             | -90.9 | 99.2     | 79.0 | 0.8   |  |

Bei den hohen Schwellenwerten tritt mit jeder höheren Prüfschwelle ein steigender negativer Bias auf. Dieser lässt sich folgendermaßen erklären. Bei der Entwicklung erfolgte eine Optimierung auf die mittleren Bedingungen innerhalb der Ozonwetterlagenklasse. Extremereignisse werden dadurch geglättet. Somit ist die Vorhersage seltener Ereignisse am oberen Rand einer Klasse sehr schwierig. Nachteilig wirkt sich dabei die geringe Zahl an Fällen aus. Von den ca. 30 000 (100%) untersuchten Fällen sind nur 240 (0,8%) in der höchsten Schwellenwertklasse. Einzelne Unterschätzungen können sich daher deutlich auswirken.

## Eine wetterlagenbasierte Klassifikation mit nachfolgender Regression

**Tabelle 7.5.** Verifikation der Vorhersageleistung der stationsbezogenen Ozonvorhersage gemittelt über alle Stationen (329); die Daten beziehen sich auf die Monate Mai bis Juli 2002; Ergebnisse des im Jahre 2002 am UBA eingesetzten Prognosesystems (WE)

| Schwellenwert: |       | ohne                             |      |       |      |       |      |       |
|----------------|-------|----------------------------------|------|-------|------|-------|------|-------|
|                | Bias  | RMSE                             | Prog | Bias  | RMSE | Per   | RV   | %     |
| aktueller Tag  | 1.4   |                                  | 22.4 | 0.8   |      | 31.4  | 49.4 | 100.0 |
| 1. Folgetag    | 3.1   |                                  | 25.3 | 0.9   |      | 36.8  | 52.6 | 100.0 |
| 2. Folgetag    | 3.0   |                                  | 24.9 | 0.2   |      | 39.0  | 59.3 | 100.0 |
| Schwellenwert: |       | 120 [ $\mu\text{g}/\text{m}^3$ ] |      |       |      |       |      |       |
|                | Bias  | RMSE                             | Prog | Bias  | RMSE | Per   | RV   | %     |
| aktueller Tag  | -4.0  |                                  | 25.3 | -13.0 |      | 35.5  | 48.9 | 24.6  |
| 1. Folgetag    | -6.3  |                                  | 29.7 | -23.7 |      | 44.8  | 56.0 | 25.1  |
| 2. Folgetag    | -6.6  |                                  | 28.4 | -28.2 |      | 49.7  | 67.4 | 25.8  |
| Schwellenwert: |       | 150 [ $\mu\text{g}/\text{m}^3$ ] |      |       |      |       |      |       |
|                | Bias  | RMSE                             | Prog | Bias  | RMSE | Per   | RV   | %     |
| aktueller Tag  | -8.7  |                                  | 29.8 | -22.8 |      | 40.2  | 45.2 | 6.2   |
| 1. Folgetag    | -11.3 |                                  | 34.3 | -45.6 |      | 57.4  | 64.3 | 6.2   |
| 2. Folgetag    | -15.4 |                                  | 37.1 | -52.7 |      | 65.9  | 68.3 | 6.1   |
| Schwellenwert: |       | 180 [ $\mu\text{g}/\text{m}^3$ ] |      |       |      |       |      |       |
|                | Bias  | RMSE                             | Prog | Bias  | RMSE | Per   | RV   | %     |
| aktueller Tag  | -15.1 |                                  | 36.4 | -35.8 |      | 52.8  | 52.4 | 0.8   |
| 1. Folgetag    | -19.6 |                                  | 42.5 | -77.3 |      | 84.7  | 74.8 | 0.8   |
| 2. Folgetag    | -42.8 |                                  | 56.4 | -92.7 |      | 102.7 | 69.8 | 0.7   |

Beim direkten Vergleich der neuen wetterlagenbasierten Methode mit der im Jahr 2002 am UBA eingesetzten Methode ergibt sich folgendes Bild bei der Berechnung ohne Schwellenwerte:

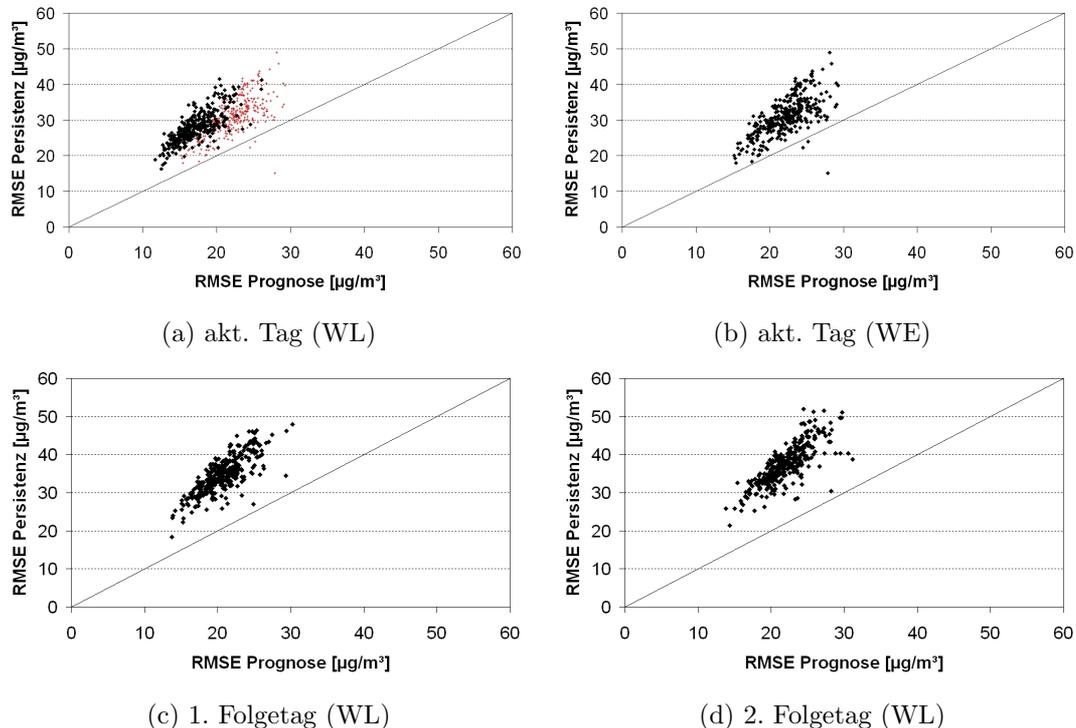
$$RV = 1 - \frac{RMSE_{WL}^2}{RMSE_{WE}^2} = 1 - \frac{17,0^2}{22,4^2} = 42,5\%$$

beim Schwellenwert 120 [ $\mu\text{g}/\text{m}^3$ ] ist der RV = 25,8%, beim Schwellenwert 150 [ $\mu\text{g}/\text{m}^3$ ] ist der RV = 13,5% und beim Schwellenwert 180 [ $\mu\text{g}/\text{m}^3$ ] ist der RV = 15,3%. Somit ist die Steigerung der Vorhersageleistung außer beim Schwellenwert 180 [ $\mu\text{g}/\text{m}^3$ ] signifikant (siehe Gleichung 4.13 und Abbildung 4.1). Beim Schwellenwert 180 [ $\mu\text{g}/\text{m}^3$ ] müsste der RV aufgrund der geringen Anzahl der Fälle (240) mindestens 19,2% betragen, um mit  $\alpha = 5\%$  signifikant zu sein.

In der Abbildung 7.4 werden die Vorhersageleistung (RMSE) der Prognose und der Persistenz gegenübergestellt. Die Abbildungen können wie folgt interpretiert werden. Je linearer die Punktwolke entlang der 1:1 Geraden verläuft umso höher ist der Zusammenhang zwischen Prognoseleistung und Persistenz. Je weiter

## 7.6 Ergebnisse der stationsbezogenen Prognose

die Punktwolke von der 1:1 Geraden nach links entfernt ist, umso höher ist der Einfluss anderer Parameter. Weiterhin zeigt eine Verschiebung der Punktwolke nach oben eine höhere Vorhersageleistung.



**Abbildung 7.4.** Gegenüberstellung der Vorhersageleistung (RMSE) der Prognose und der Persistenz für alle Stationen für den Zeitraum 1.5.2002 - 31.7.2002. Für den aktuellen Tag die Routinen WL und WE und für die beiden Folgetage nur die Routine WL. Zur Verdeutlichung der Unterschiede zwischen den beiden Prognosen wurden in der Teilabbildung (a) als kleine Kreuze die Ergebnisse der alten Routine eingefügt. Die Auswertung erfolgte unabhängig vom Schwellenwert.

Wie aus der Abbildung 7.4a ersichtlich, besteht weiterhin ein sehr starker Zusammenhang zwischen der Erhaltungsneigung (Persistenz) und der Vorhersageleistung der wetterlagenbasierten Routine (WL) an den Stationen. Somit ist die Vorhersageleistung, die durch die Routine erreicht werden kann, durch die auftretende Persistenz limitiert. Beim Vergleich der Abbildungen 7.4a und 7.4b zeigt sich aber eine deutlich verringerte Abhängigkeit der wetterlagenbasierten Prognoseroutine (WL) gegenüber der alten Routine (WE).

## 7.7 Zusammenfassung und Diskussion

Die in diese Methode gesetzten Erwartungen wurden voll erfüllt. Es ist möglich, basierend auf den großräumigen meteorologischen Bedingungen eine Einteilung der Tage in Ozonwetterlagenklassen durchzuführen. Dadurch kann eine speziell auf diese Situation angepasste Regressionsfunktion zur Vorhersage genutzt werden. Die Ergebnisse in den Tabellen 7.4 und 7.5 zeigen, dass diese Vorhersage deutliche Verbesserungen gegenüber der bisher am UBA eingesetzten Methode bringt. Für die operationelle Prognose muss die Umsetzung in mehreren Teilroutinen erfolgen:

- Bestimmung der Ozonwetterlage entsprechend der 12 UTC Wettersituation,
- Berechnung des mittleren maximalen Ozonwertes über eine wetterlagenspezifische Regressionsfunktion,
- Zuordnung des Tages zu der Klasse in der der berechnete Ozonwert liegt, wenn der berechnete Ozonwert sich außerhalb der durch die Wetterlagenbestimmung vorgegebenen Klasse befindet und
- Berechnung der stationsbezogenen Prognosen in Abhängigkeit von der regionalen Ozonklasse.

Ein entscheidender Vorteil der Methode ist, dass für die Klassifikation die großräumigen meteorologischen Bedingungen ausschlaggebend sind. Mit der stationsbezogenen Regression werden im zweiten Schritt die stationsspezifischen Eigenschaften (z.B. Verkehrs- oder Berglage) in die Prognose einbezogen.

Ein Nachteil ist auch weiterhin in den Regressionsbeziehungen enthalten. Sie beinhalten die Korrelationen der Jahre 1997 bis 2001 und sind damit statisch. Im Kapitel 4 wird im Unterabschnitt 4.2.2 eine Möglichkeit der Dynamisierung der Korrelationen beschrieben. Die Umsetzung war im vorgegebenen Zeitrahmen nicht möglich. Damit eine Umsetzung möglich ist, müssen umfangreiche Untersuchungen, mit dem Ziel einen oder mehrere Adaptionszeiträume zu ermitteln, durchgeführt werden. Aus früheren Untersuchungen [Enke 1984] ist bekannt, dass eine Aufteilung in zwei Summenmatrizen sinnvoll sein könnte, beispielsweise in eine Summenmatrize mit einem kurzen Adaptionszeitraum von einigen Tagen und eine weitere mit einem deutlich längeren Adaptionszeitraum von mindestens 100 Tagen. Das entscheidende Problem dabei ist, dass eine Regel gefunden werden muss, wann zwischen den beiden Summenmatrizen gewechselt wird. Eine Idee wäre, die Vorhersage mit der Summenmatrize mit dem kurzen Adaptionszeitraum zu starten. Somit kann der Beginn einer Ozonepisode gut erfasst werden. Erkennt die Regel zum Wechseln zwischen beiden Summenmatrizen das ein Wetterumschwung und damit ein Ende der Ozonepisode in Aussicht ist, wird für einige Tage auf die Summenmatrix mit dem langen Adaptionszeitraum umgeschaltet. Wie schon angesprochen wäre eine Untersuchung dieser Problematik

## 7.7 Zusammenfassung und Diskussion

---

eine Möglichkeit, eine Verbesserung der Prognoseleistungen zu erreichen. Es muss aber auch darauf hingewiesen werden, dass mit der Dynamisierung auch Fehler wahrscheinlicher werden, denn trotz aller Kontrollen können problematische Werte in die Summenmatrix aufgenommen werden.

Im Anhang E wird eine Anwendungsmöglichkeit für die neuentwickelte Methode zur Ableitung von Wetterlagen vorgestellt.

## Eine wetterlagenbasierte Klassifikation mit nachfolgender Regression