

Kapitel 4

Statistische Routinen

Das Kapitel stellt die Methoden vor, die eine Basis für die nachfolgenden Kapitel bilden. Begonnen wird mit dem Abschnitt 4.1 über die methodischen Grundlagen für die Entwicklung von Regressionsmodellen. Der Abschnitt geht darauf ein, auf welche Arbeitsschritte geachtet werden muss, damit eine spätere Nutzung des entwickelten Modells erfolgreich ist. Nachdem erklärt worden ist, was alles bedacht werden muss, werden im Abschnitt 4.2 die zwei Vorgehensweisen vorgestellt mit denen die Entwicklung durchgeführt wurde. Im Abschnitt 4.3 werden die Prüfmaße vorgestellt, mit denen die Güte der Entwicklung und die Vorhersageleistung kontrolliert wurden.

Ein weiterer wichtiger Punkt ist die Bereinigung der Eingangsdaten von Jahres- und Tagesgang. Im Abschnitt 4.4 werden daher Beispielverläufe dargestellt und die Entfernung von Jahres- und Tagesgang aus den Zeitreihen erläutert.

4.1 Methodische Grundlagen

Das Hauptproblem der hier genutzten Statistik ist die Suche nach stabilen und erwartungstreuen Parametern für die Regressionsmodelle. Durch die vorgegebenen Randbedingungen ist das nicht ganz einfach. Das Ziel dabei ist, aus den Daten eine möglichst optimale Zahl an potentiellen Prädiktoren zu finden. Mit diesen soll dann ein Modell mit möglichst optimaler Kompliziertheit entwickelt werden [Ivachnenko und Müller 1984, Enke 1988].

Damit das gewählte Modell nicht zur Schwachstelle der gesamten Arbeit wird, müssen eine Vielzahl von Stabilitätskriterien eingehalten werden. Im Folgenden wird eine Übersicht an Anforderungen aus einer ausgewählten Gruppe von Publikationen dargestellt [Enke 1988, Balzer 1989, Wilks 1995]. Alle drei Autoren stellen sehr hohe Anforderungen, die bei der Entwicklung eines statistischen Prognosesystems eingehalten werden müssen. Somit ist bei Einhaltung der Kriterien mit einem stabilen und erwartungstreuen Modell zu rechnen.

4.1.1 Auswahl der Prädiktoren

Üblicherweise ist eine große Anzahl von Prädiktoren vorhanden. So stehen zum Beispiel für jeden Gitterpunkt der hier genutzten Wetterprognose des DWD 60 Parameter zur Verfügung (siehe Tabelle F.3). Im ersten Schritt muss diese Anzahl reduziert werden. Ein Grund dafür ist die Gefahr des Overfittings. Dieses kann dann auftreten, wenn das Verhältnis der Anzahl der Fälle im Entwicklungskollektiv zu der Anzahl der Prädiktoren ungünstig ist.

Ist beispielsweise die Anzahl der Fälle gleich zwei und die Zahl der Prädiktoren eins, so ist eine perfekte lineare Regression möglich. Dabei ist es völlig unerheblich ob zwischen den Prädiktoren und dem Prädiktanden ein Zusammenhang besteht oder nicht. Wie im Beispiel 6.5 von [Wilks 1995] auf Seite 185ff. dargestellt, gilt dieses für jede Kombination, bei der die Anzahl der Fälle gleich der Anzahl der Prädiktoren plus eins ist. Daraus können mehrere Lehren gezogen werden:

- es sollten nur physikalisch begründbare Prädiktoren für die Regression herangezogen werden,
- jede Regression muss an einem unabhängigen Kollektiv getestet werden
- und die Anzahl der Fälle soll in einem sinnvollen Verhältnis zu der Anzahl der Prädiktoren stehen.

4.1.2 Screening-Regressions-Analyse

Die Auswahl der Prädiktoren erfolgt üblicherweise in Form einer Screening-Regressions-Analyse. Am bekanntesten ist hier die *forward selection* oder *stepwise regression*. Hierbei werden schrittweise erst einer, dann jeweils einen Prädiktoren hinzunehmend, der beste bzw. die beste Kombination an Prädiktoren gesucht. Ist dieser bzw. sind diese gefunden, so wird im nächsten Schritt die jeweils beste Kombination der gefundenen Prädiktoren mit einem neu hinzugefügten geprüft.

Alternativ dazu kann die *backward elimination* genutzt werden. Diese ist analog zur *forward selection* nur mit einer umgekehrten Arbeitsrichtung. Beginnend mit allen Prädiktoren wird schrittweise immer der schlechteste Prädiktor herausgenommen.

Mit dem Screening wird der Beitrag der einzelnen Prädiktoren für die zu entwickelnde Routine untersucht. Hierbei ist auch die Erwartungstreue ein wichtiger Aspekt. Unter Erwartungstreue wird verstanden, dass die bei der Entwicklung auftretenden statistischen Zusammenhänge zwischen Prädiktand und Prädiktor auch bei unabhängigen Daten weiterhin Bestand haben. Dieses wird mit der Crossvalidation (Unterabschnitt 4.1.4) kontrolliert. Am Ende der Screening-Analyse werden nur noch die Prädiktoren genutzt, die einen hohen Beitrag bei der Berechnung liefern. Alle anderen Prädiktoren wurden durch die Screening-Analyse herausgenommen. Erfolgt eine multiple Regression mit allen Prädiktoren im Block, fehlt diese Information und die Erwartungstreue ist nicht mehr gegeben.

4.1 Methodische Grundlagen

4.1.3 Abbruchkriterien

Beide oben genannten Methoden benötigen ein Abbruchkriterium. Eine solche Regel soll bestimmen, nach der Hinzunahme welches Prädiktors das Screeningverfahren beendet wird. Dazu werden Prüfmaße herangezogen, mit denen die Güte der Regressionsfunktion geprüft wird. Bei der *forward selection* wird dann abgebrochen, wenn sich die Qualität der Vorhersage nur noch weniger als eine vorgegebene Schranke verbessert oder sogar schlechter ist. Umgekehrt wird bei der *backward elimination* aufgehört, wenn die Qualität der Vorhersage stärker abnimmt.

Theoretisch können unendlich viele Prädiktoren gewählt werden. Es muss aber mit jeder Hinzunahme eines weiteren Prädiktoren mit einer erhöhten Wahrscheinlichkeit für ein Overfitting gerechnet werden. Daher sollten [Wilks 1995] zu Folge nicht mehr als zwölf Prädiktoren gewählt werden. Auch sollte möglichst eine Mindestanzahl von 50 Fällen je Prädiktor eingehalten werden.

4.1.4 Crossvalidation

Eine der Lehren, aus der oben genannten Problematik des Overfittings, ist die Notwendigkeit der Prüfung der abgeleiteten Regressionsfunktion an einem unabhängigen Kollektiv. Diese erfolgt durch die Crossvalidation.

Die simpelste Form ist die Herausnahme eines Wertepaares bei der Erstellung der Regressionsfunktion und die Prüfung der Funktion an genau diesem Wertepaar. Dieses erfolgt dann im Rotationsprinzip für jeden einzelnen Fall. Neben der Möglichkeit einen Fall gegen die Gesamtheit zu prüfen, ist es auch möglich eine Gruppe von Fällen (bis zur Hälfte des gesamten Kollektives) herauszunehmen und diese als Prüfkollektiv zu nutzen. Auch dann wird mit dem Rotationsprinzip gearbeitet. Das heißt es wird wechselseitig an einem Kollektiv entwickelt und am anderen getestet und dann umgekehrt. Damit mögliche Trends im Kollektiv keinen Einfluss haben, sollte in einem vorbereitenden Schritt das Datenkollektiv vom Trend bereinigt werden. Es muss festgestellt werden, dass fast immer die Qualität der Vorhersage im Prüfkollektiv schlechter ist, als im Entwicklungskollektiv. Überschreitet der Unterschied zwischen den Ergebnissen eine vorgegebene Schwelle, so sind die gefundenen Regressionsfunktionen als nicht nutzbar anzusehen!

Neben der bisher beschriebenen einfachen Crossvalidierung wurde bei der Entwicklung der Routinen eine weitergehende Crossvalidierung genutzt. Hierbei werden die gefundenen Regressionsfunktionen an einem für die Entwicklung nicht genutzten Zeitraum getestet. Nur wenn die Prüfergebnisse für diesen unabhängigen Zeitraum nicht deutlich schlechter sind als die Ergebnisse im Entwicklungszeitraum, wurden die ermittelten Korrelationen genutzt. Dieses in der Arbeit genutzte Prinzip wird doppelte Crossvalidation genannt.

Im Gegensatz zur hier vorliegenden guten Datensituation sollten bei schlechter Datenlage die Ausführungen in den Artikeln von [Michaelsen 1987, Barnston und Van de Dool 1993] studiert werden. Dort werden Probleme, die bei schlechter Datenlage auftreten können, erläutert und Vorschläge gemacht, wie trotzdem gute Vorhersagen erstellt werden können.

4.1.5 Eine Selbstlernende Bias-Eliminierung

Für den Fall von vollständig bekannten und korrekten Relationen sowie Eingangsdaten ohne systematischen Fehler, müssten biasfreie Prognosen erzeugt werden können. In der Realität kann der Bias jedoch je nach Witterungsabschnitt recht beachtlich sein. Bei der Prognose von Immissionskonzentrationen sind neben den sich ändernden Relationen (siehe Abbildung 3.4) oft auch externe Störungen (z.B. die Qualität der Wettervorhersage) die Ursache für diese Fehler. Eine Möglichkeit solche Fehler zu korrigieren beschreibt [Balzer 1986].

$$Cor_{neu} = Cor_{alt} + \beta \times (Bias_{aktuell} - Cor_{alt}) \quad (4.1)$$

Jede Prognose wird dann um den Korrekturwert Cor geändert. $Bias_{aktuell}$ ist der bei der letzten Prognose (am Vortag) aufgetretene Bias. Die Glättungskonstante β beschreibt den Grad der Anpassung an die aktuellsten Daten. Bei kleinem β lernt das System langsam bei großem schneller. Die Bestimmung des Wertes von β muss subjektiv mittels probieren erfolgen, da es keinen mathematischen Algorithmus dafür gibt. Die Ermittlung des Korrekturwertes erfolgt für jede Station und wird täglich entsprechend Gleichung 4.1 neu angepasst.

4.2 Screeningmethoden in der Arbeit

Bei der Suche nach der besten Kombination von Prädiktoren für die Regression wurden zwei alternative Screeningverfahren angewendet. Das sind die *stufenweise einfache lineare Regression* (Unterabschnitt 4.2.1) und die *multiple lineare Regression* (Unterabschnitt 4.2.2). Mit beiden Wegen sind in etwa gleiche Ergebnisse bei der Regressionsfunktionsentwicklung zu erwarten. Die stufenweise einfache lineare Regression liefert dann die gleichen Ergebnisse wie die multiple lineare Regression, wenn sie mit einer sehr großen Zahl an Schritten, unter möglicher mehrfacher Verwendung der Prädiktoren, erfolgt ist.

4.2.1 Stufenweise einfache lineare Regression

Das im Folgenden vorgestellte Verfahren ist in der Literatur unter dem Begriff *Restverfahren* nach [Panovskij und Brajer 1967] bekannt.

4.2 Screeningmethoden in der Arbeit

Begonnen wird mit der Suche nach dem besten linearen Prädiktor. Dieser zeichnet sich dadurch aus, dass er den größten RV^1 -Wert, berechnet aus dem Vorhersagefehler der Referenzvorhersage (in dieser Arbeit wurde die Persistenz als Referenzprognose genutzt) und dem Vorhersagefehler der Regressionsfunktion, hat. Unter Nutzung der Regressionskoeffizienten dieses Prädiktors, wird eine Reihe der Residuen, aus der Differenz eingetroffene Werte zur Vorhersage, ermittelt. Jene ist die Ausgangsreihe für die nächste Iteration. Bei dieser wird der beste Prädiktor zur Vorhersage der vorliegenden Residuenreihe gesucht. Auch hier wird äquivalent zum ersten Schritt der berechnete RV -Wert als Richtmaß herangezogen. Je nach Höhe der vorgegebenen maximalen Prädiktorenanzahl wird diese Iteration entsprechend oft wiederholt.

Damit die Prädiktorenselktion erwartungstreu ist, wurde die H-Methode angewendet. Hierfür wird das Datenkollektiv in zwei Teile aufgeteilt (z.B. Daten mit einer geraden Feldnummer als Teil A und die anderen als Teil B). Im nächsten Schritt wird der Datenkollektivteil A als Entwicklungskollektiv angesehen und die dafür gefundenen Koeffizienten am Datenkollektivteil B (Testkollektiv) getestet. Anstelle der Standardabweichung selbst wurden jetzt die Teilsummen für die Berechnung der Standardabweichung bzw. des RMSE gespeichert ($\sum R^2$, $\sum R$; siehe Gleichungen 4.2, 4.3).

$$\sigma = \sqrt{\frac{1}{n-1} \left(\sum_{i=1}^n R_i^2 - \frac{1}{n} \sum_{i=1}^n R_i \sum_{i=1}^n R_i \right)} \quad (4.2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n R_i^2} \quad (4.3)$$

Mit R_i = registrierter Wert - vorhergesagter Wert. Danach wurden Entwicklungs- und Testkollektiv ausgetauscht. Auch hierfür wurden die Teilsummen berechnet. Durch die Nutzung der Teilsummen ist es möglich, beide Standardabweichungen zu mitteln. Der Prädiktor, bei dem der höchste mittlere RV gefunden wurde, wird für die Ermittlung der Residuenreihe und die darauf folgenden Schritte verwendet. Neben einem guten mittleren RV müssen beide Teilsummen in etwa gleich groß sein, damit die Erwartungstreue erfüllt ist. Beendet wurde die Prädiktorenselktion wenn

- die maximale Anzahl der Prädiktoren erreicht wurde oder
- die Steigerung der Vorhersageleistung kleiner war als ein vorgegebener Schwellenwert (zum Beispiel $\Delta RV \leq 0,5\%$).

¹Reduktion der Varianz; siehe Unterabschnitt 4.3.3

4.2.2 Multiple lineare Regression

Alternativ zu der schrittweisen Regressionsanalyse, bei der jedem weiteren Prädiktor die Residuenreihe angeboten wird, kann die multiple lineare Regression genutzt werden. Wird bei dieser mehr als ein Prädiktor verwendet, muss die Residuenreihe nicht mehr berechnet werden. Einen Berechnungsansatz liefert die Publikation von [Enke 1984]. Die dort beschriebene Methode hat neben der Entwicklung von multiplen linearen Regressionsfunktionen noch einige andere im Folgenden vorgestellte Spezifikationen. Die Entwicklung beginnt mit dem datensatzweisen auffüllen (Prädiktand plus n Prädiktoren) einer Summenmatrix.

$$S = \begin{pmatrix} \sum x_0 & \sum x_1 & \sum x_2 & \dots & \sum x_n \\ \sum x_0x_0 & \sum x_0x_1 & \sum x_0x_2 & \dots & \sum x_0x_n \\ & \sum x_1x_1 & \sum x_1x_2 & \dots & \sum x_1x_n \\ & & \sum x_2x_2 & \dots & \sum x_2x_n \\ & & & & \vdots \\ & & & & \vdots \\ & & & & \sum x_mx_n \end{pmatrix} \quad (4.4)$$

Dabei stellt der Index 0 den Prädiktanden dar. Alle weiteren Indizes (1 .. n,m) sind die Prädiktoren. Die Anzahl der Prädiktoren und damit die Größe der Matrix ist variabel.

Es kann in der Summenmatrix ein ständiger Informationsaustausch durchgeführt werden. Dazu muss ein Verfallswert eingeführt werden. Dabei bedeutet ein Wert von 100, dass ab dem 101sten zu lernenden Datensatz erst $\frac{1}{100}$ der Summenmatrix gelöscht wird und danach der aktuelle Datensatz hinzugefügt wird. Somit wird eine Anpassung der Regression an Änderungen im Datenkollektiv erreicht.

$$\sum (x_n, x_m)^T = \sum (x_n, x_m)^{T-1} - \frac{1}{a} \sum (x_n, x_m)^{T-1} + (x_n, x_m)^T \quad (4.5)$$

mit: $n, m =$ Ordnungsindizes der Prädiktoren; $T, T - 1$ als Kennzeichen für den Speicherinhalt vom Tag bzw. Vortag und a als Lernintervall in Tagen (Verfallswert).

Bei der Nutzung zur Berechnung der multiplen linearen Regressionsfunktionen muss der Verfallswert größer sein als die Anzahl der Trainingsfälle. Dadurch wird kein Informationsaustausch durchgeführt. Nachdem die Summenmatrix *trainiert* wurde, kann sie für die Bestimmung der Koeffizienten genutzt werden. Dabei werden mehrere Schritte durchlaufen. Basierend auf der Formel 4.6 wird aus der Summenmatrix die Kovarianzmatrix erzeugt. Hierbei können die einzelnen Koeffizienten an- bzw. ausgeschaltet werden. Zum Ausschalten muss der entsprechende Feldwert Null gesetzt werden.

$$C_{n,m} = \frac{\sum_{i=1}^N x_{ni}x_{mi} - \frac{1}{N} \sum_{i=1}^N x_{ni} \sum_{i=1}^N x_{mi}}{N - 1} \quad (4.6)$$

4.2 Screeningmethoden in der Arbeit

Mit der oben erzeugten Kovarianzmatrix wird zuerst die einfache Korrelation zwischen dem Prädiktanden und den Prädiktoren berechnet. Im zweiten Schritt wird die multiple Korrelation errechnet. Vorher erfolgt eine Prüfung darauf, welche Prädiktoren für die Regression signifikant sind. Unterschreitet die Korrelation eines der Prädiktoren einen Grenzwert, so wird der Prädiktor aus der Kovarianzmatrix gestrichen. Die zum Vergleich genutzten Größen werden nach den Gleichungen 4.7 und 4.8, die aus [Taubenheim 1969] entnommen wurden, berechnet:

$$t = \frac{|r|}{\sqrt{1-r^2}} \times \sqrt{n-2} \quad (4.7)$$

mit r als Korrelationskoeffizienten und n als die Anzahl der Fälle.

$$t_{0,05} \approx 1,96 + \frac{2,44}{n-1} \quad (4.8)$$

gültig für $n \geq 3$, $t_{0,05}$ ist die Schranke mit 5% Irrtumswahrscheinlichkeit. Wenn $t < t_{0,05}$, dann wird auch dieser Prädiktor abgeschaltet. Dieses erfolgt durch Nullsetzung der entsprechenden Spalte und Zeile. Der Schnittpunkt beider wird auf Eins gesetzt. Für den Fall, dass ein Prädiktor in Bezug auf einen anderen eine höhere Korrelation als 0,95 aufweist (Interkorrelation), löscht die vorliegende Routine den Prädiktor mit der kleineren Korrelation zum Prädiktanden.

Zur Berechnung der Koeffizienten der multiplen Regressionsgleichung wird ein in [Taubenheim 1969] beschriebenes Verfahren eingesetzt. Hierzu werden unter Nutzung der Determinantenberechnung die einzelnen Koeffizienten ermittelt.

Ein Vorteil der beschriebenen Methode ist die Robustheit gegenüber Datenausfällen. Fehlt einer der Prädiktoren, so wird dieser bei der Koeffizientenberechnung abgeschaltet. Das Modul berechnet dann die Koeffizienten so, dass eine Vorhersage ohne den fehlenden Prädiktoren möglich ist. Dadurch ist eine Verschlechterung der Vorhersageleistung zu erwarten. Die Prädiktorenselktion wurde auch hier beendet, wenn:

- die maximale Anzahl der Prädiktoren erreicht wurde oder
- die Steigerung der Vorhersageleistung kleiner war als ein vorgegebener Schwellenwert (zum Beispiel $\Delta RV \leq 0,5\%$).

4.3 Prüfmaße

In der Arbeit werden in der Hauptsache drei Prüfmaße genutzt. Das ist der RMSE (Unterabschnitt 4.3.1), der Bias (Unterabschnitt 4.3.2) und der RV-Wert (Unterabschnitt 4.3.3). Zur Einschätzung der Vorhersagegüte wird in der Literatur eine umfangreiche Anzahl an Prüfgrößen beschrieben. Durch die Quadrierung der Fehler ist der RMSE besonders großen Fehlvorhersagen gegenüber empfindlich. Dieses ist auf dem Gebiet der Ozonprognose sehr wichtig. Der Unterabschnitt 4.3.4 geht auf Prüfmaße ein, die bei der Prüfung von Binäreignissen genutzt werden.

4.3.1 RMSE - Root Mean Squared Error

Eine wesentliche Größe zur Einschätzung der Qualität einer Prognose ist der Root Mean Squared Error. Er gibt die Streuung der prognostizierten Datenreihe als Differenz zu den tatsächlich eingetretenen Werten an. Somit ist der RMSE ein Fehlerbewertungsmaß. Im Gegensatz zur Standardabweichung ist der RMSE nicht mittelwertbereinigt und damit ein strengeres Prüfmaß.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n R_i^2} \quad (4.9)$$

mit R = vorhergesagter Wert - registrierter Wert
und n = Anzahl der Prognosen.

4.3.2 Bias

Mit dem Prüfmaß Bias können systematische Fehler aufgedeckt werden. Treten einseitige Verzerrungen zwischen Prognose und Wirklichkeit auf, so zeigt sie der Bias an.

$$Bias = \sum_{i=1}^n \frac{R_i}{n} \quad (4.10)$$

mit R = vorhergesagter Wert - registrierter Wert
und n = Anzahl der Prognosen.

4.3.3 RV - Reduktion der Varianz

Von einer neuen Vorhersage wird gefordert, dass sie besser/genauer ist als andere schon bekannte Prognoseroutinen. Hierzu muss ein Vergleich erfolgen. Der RV ist eine Größe, die zwei Fehlermaße miteinander vergleicht.

4.3 Prüfmaße

$$RV_{RMSE} = 1 - \frac{RMSE_V^2}{RMSE_R^2} \quad (4.11)$$

mit

$RMSE_V$ = RSME der Vorhersagereihe

$RMSE_R$ = RSME der Referenzreihe.

Möglich ist auch ein Berechnungsansatz über die Standardabweichung.

$$RV_\sigma = 1 - \frac{\sigma_V^2}{\sigma_R^2} \quad (4.12)$$

mit

σ_V = Standardabweichung der Vorhersagereihe

σ_R = Standardabweichung der Referenzreihe.

Bei einem $Bias = 0$ sind RV_{RMSE} und RV_σ gleich. Ist der Bias ungleich Null so berechnet der RV_{RMSE} ein strengeres Prüfmaß da er nicht mittelwertbereinigt ist. Nur ein perfektes Vorhersageverfahren mit $RMSE_V = 0$ ergibt $RV = 100\%$. Ein positiver RV zeigt eine Verbesserung der Vorhersageleistung an. Hierbei ist ein vom Untersuchungskollektiv abhängiger Mindestwert zu überschreiten. Damit der RV signifikant größer als Null ist, wird ein mit $\alpha = 5\%$, nach Gleichung 4.13 zu berechnender, in Abbildung 4.1 dargestellter Mindestwert gefordert [Balzer 1997].

$$RV_{Min} \approx \frac{186}{(N - 2)^{0,415}} \quad (4.13)$$

mit N als Stichprobenumfang.

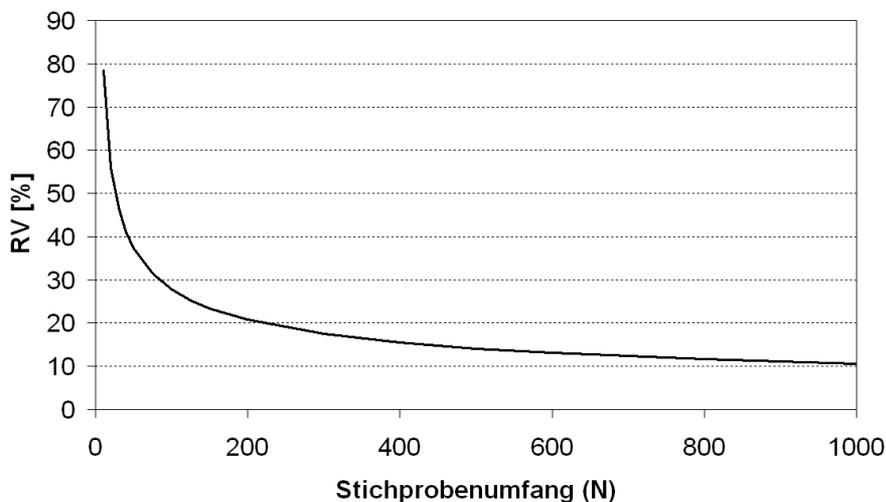


Abbildung 4.1. Mindest - RV in Abhängigkeit vom Stichprobenumfang nach Gleichung 4.13

Negative RV-Werte weisen darauf hin, dass eine Verbesserung der Vorhersageleistung, mit dem aktuellen Prognoseansatz, nicht gegeben ist. Als Referenz wurde in dieser Arbeit die Persistenzvorhersage genutzt.

Die beiden statistischen Kenngrößen RMSE und RV haben jeweils Vor- und Nachteile. Der RMSE ist ein Maß der Vorhersagegüte. Mit ihm allein ist aber keine Wichtung der Qualitätsunterschiede zweier Prognosen möglich. Die vergleichende Vorhersageleistung kann durch den RV dargestellt werden. Er ist aber immer ein Relativmaß zur Vorhersageleistung der Referenzprognose.

4.3.4 Prüfmaße für Binärereignisse

Eine weitere Prüfmöglichkeit sind kategorische Prüfmaße. Dazu müssen die zu prüfenden Ereignisse binarisiert werden. Es wird dann zwischen eingetroffen (1) und nicht eingetroffen (0) und danach ob das Ereignis so vorhergesagt wurde unterschieden. Als erster Schritt muss dafür eine 2×2 Prüfmatrix aufgestellt werden (siehe Tabelle: 4.1).

Tabelle 4.1. 2×2 Prüfmatrix für Binärereignisse; EIN für eingetroffene Ereignisse und VOR für vorhergesagte Ereignisse

		EIN		
		0	1	Σ
VOR	0	a	b	g
	1	c	d	h
		Σ	e	f
		N		

Mit dieser Prüfmatrix können dann mehrere Prüfgrößen berechnet werden.

- TSS = true skill statistic = $\frac{d}{f} - \frac{c}{e}$; (ideal = 100% = 1)
- HSS = Heidke skill score = $\frac{2*(a*d-b*c)}{(a+c)*(c+d)+(a+b)*(b+d)}$; (ideal = 1)
- POD = probability of detection = $\frac{d}{f}$; (ideal = 100% = 1)
- FAR = false alarm rate = $\frac{c}{h}$; (ideal = 0%)
- Bias = $\frac{h}{f}$; (ideal = 1)

Eine genaue Beschreibung dieser Größen und was sie unterscheidet ist in [Wilks 1995] zu finden. Ein Einsatzfeld für diese Prüfmaße ist die Arbeit mit Schwellenwerten.

4.4 Zyklische Schwankungen - Jahres- und Tagesgang

Viele Zeitreihen, wie auch Ozonreihen, weisen einen signifikanten Jahres- und Tagesgang auf. Möchte man solche Zeitreihen statistisch bearbeiten, so ist es sinnvoll, diesen zu entfernen. Hierzu wurde aus dem vorliegenden Datenarchiv der mittlere Jahres- und Tagesgang für jede Station berechnet.

Je nach Aufgabengebiet sind für jede Stunde in jedem Monat die mittleren (oder die maximalen) Ozonwerte ermittelt worden. Die vorliegende Zeitreihe dieser Mittelwerte wurde dann mittels einer Cubic-Spline-Interpolation auf Tagesbasis geglättet. Die vorliegenden Jahres- und Tagesgänge können dann zur Zeitreihenanalyse genutzt werden. Nach der Anwendung auf die Originalzeitreihe erhält man eine neue Zeitreihe, die die Abweichung des jeweiligen Messwerts vom Jahres- und Tagesgang aufzeigt.

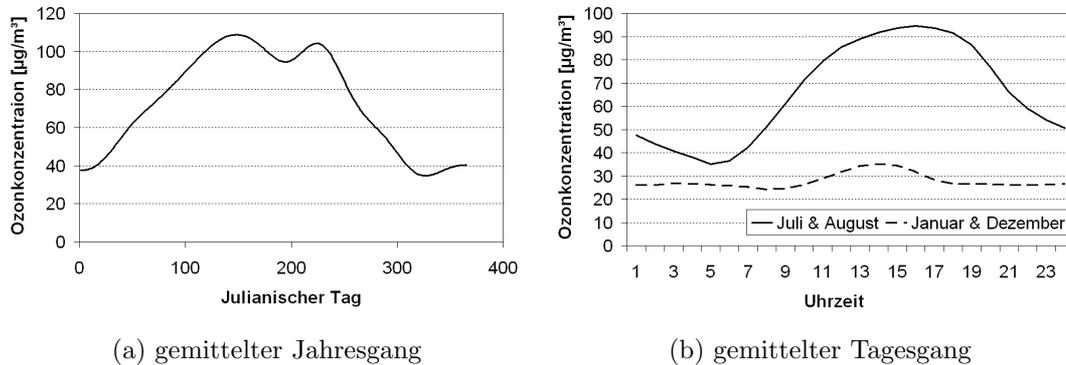


Abbildung 4.2. Darstellung des über sechs Jahre (1997-2002) gemittelten Jahresganges der Tagesmaxima der Ozonkonzentration und der über jeweils zwei Monate gemittelten Tagesgänge der Ozonkonzentration an der Station Burg (DEBB001).

Die Entfernung der zyklischen Schwankungen hat einen entscheidenden Vorteil. Wird anstatt der Ozonkonzentration die Abweichung vom Jahresgang als Prädikand gewählt, so tritt eine deutliche Verbesserung der Vorhersageleistung ein. Daher wird in den nachfolgenden Routinen fast immer die Abweichung vom Jahresgang prognostiziert und dann der vom Tag abhängende Jahresgangwert aufaddiert.

