

Was für eine Bullenhitze!

Analyse einiger Temperaturdatensätze

von A. Kimber – Übersetzt und bearbeitet von G. König

Kurzfassung: Über längere Zeiträume notierte Monatsmittelwerte der bodennahen Lufttemperatur werden nach verschiedenen Fragestellungen aufbereitet und mit Hilfe der Zeitreihenanalyse ausgewertet. Wir gehen bei der Analyse von einem additiven Modell aus und benutzen zur Bestimmung der glatten Komponente bzw. des Trends die Methode der gleitenden Durchschnitte.

Einleitung

In seinem ausgezeichneten Werk über graphische Methoden diskutiert Tufto das Ergebnis einer Untersuchung über die Verwendung von graphischen Darstellungen in 15 der weltweit bekanntesten Zeitungen und Zeitschriften im Zeitraum von 1970-1980. Eine Zufallstichprobe aus 4000 dieser graphischen Darstellungen ergab zu 75% Diagramme von Zeitreihen. Aus diesem Grunde sollte es im Interesse jedes Bürgers liegen, einige Kenntnisse über Zeitreihen zu besitzen und die statistischen Methoden zu verstehen, die auf Zeitreihen angewendet werden.

In diesem Beitrag werden einige einfache Untersuchungsmethoden bei Zeitreihen diskutiert und anhand von Datensätzen illustriert. Die Daten können leicht erweitert werden und sollten für den Einsatz im Unterricht geeignet sein, insbesondere wenn Computer oder Taschenrechner für die zum Teil mühseligen Rechnungen zur Verfügung stehen. Der englische Autor benutzte das Minitab-Programm bei den hier dargestellten Rechnungen (Ryan et al;). Der deutsche Benutzer sei u.a. auf das Statistikprogrammpaket GSTAT verwiesen (s. Böker 1989 bzw. 1991).

Die Daten

In letzter Zeit werden häufig Umweltprobleme in den Medien diskutiert; ein Aspekt hierbei ist die Möglichkeit einer Klimaänderung. Viele Statistiken zu verschiedenen Klimaelementen, wie z.B. Niederschlag, Wind oder Lufttemperaturen wurden bereits angefertigt. Aus der Analyse dieser Klimadaten versucht man Trends und Hinweise auf anthropogen bedingte Veränderungen des heuti-

gen Klimas abzuleiten. Da hier das zeitliche Verhalten der Beobachtungsreihen von besonderem Interesse ist, liegt es nahe, Methoden bei Zeitreihen zu benutzen. Zeitreihen sind Werte einer Variablen, die über eine Reihe von Zeitpunkten beobachtet werden. Zeitreihen enthalten also Längsschnittdaten.

Wir wollen in diesem Beitrag die Zeitreihen der Tabelle 1 analysieren. Es handelt sich um Mittelwerte der Lufttemperaturen in England und Wales für die Monate Januar 1981 bis Dezember 1989 (Hier wurde also der Wert der Variablen mittlere Lufttemperatur in jedem Monat des beobachteten Zeitraums festgestellt). Als Hintergrundinformation wurden die entsprechenden mittleren Monatswerte für die 30Jahres-Zeiträume 1941-1970 bzw. 1951-1980 hinzugefügt; alle Daten entstammen öffentlichen Statistiken.

Tab.1: Mittlere monatliche Lufttemperaturen in England und Wales für 1981-1989 und die dreißigjährigen Perioden 1941-1970 und 1951-1980

Monat	Jahr										
	81	82	83	84	85	86	87	88	89	41-70	51-80
Jan	5,2	3,6	7,0	3,8	1,3	4,1	1,4	5,8	6,7	4,0	4,0
Feb	3,7	5,5	2,6	3,5	2,8	-0,4	4,1	5,2	6,4	4,2	4,1
Mar	8,3	6,6	7,0	4,6	5,0	5,4	4,6	6,6	7,8	6,2	5,9
Apr	8,2	8,9	7,3	8,5	8,7	6,1	10,3	8,5	6,9	8,8	8,2
Mai	11,6	11,9	10,6	10,1	11,0	11,3	10,3	12,0	13,2	11,6	11,3
Jun	13,8	15,7	14,6	14,7	12,9	14,8	13,1	14,6	14,7	14,7	14,3
Jul	16,0	16,8	19,2	16,6	16,4	16,1	16,2	15,1	18,3	16,3	16,0
Aug	16,6	16,4	17,6	17,8	15,0	14,1	16,0	15,7	17,0	16,1	15,9
Sep	15,1	14,8	14,2	14,1	14,9	11,9	14,2	13,8	15,1	14,3	14,0
Okt	8,9	10,7	11,0	11,9	11,4	11,4	10,3	11,1	12,3	11,2	11,0
Nov	8,1	8,5	8,1	8,6	4,7	8,3	7,2	6,2	7,2	7,2	7,1
Dez	1,4	4,9	6,3	5,7	6,7	6,4	6,1	7,7	5,5	5,5	5,1

Es wäre interessant zu überprüfen, inwieweit Temperaturänderungen über längere Zeiträume nachweisbar sind, wie z.B. steigende Temperaturen aufgrund von Wetterlagenänderungen oder wegen des Treibhauseffekts (Erwärmung der Erde durch in die Atmosphäre freigesetzte Gase). Hier muß auf der Grundlage realer Daten und nicht nur subjektiver Empfindungen argumentiert werden. Wer erinnert sich nicht an scheinbar gute oder mäßigen Sommer oder glaubte festzustellen, daß der Winter 1990 der dritte milde Winter in Folge war. Wie kann man solche Aussagen aber mit echten Daten statistisch belegen?

Ein Modell

Bevor wir die Lufttemperaturreihen genauer analysieren, ist es nützlich, die Datensätze in einem größeren Zusammenhang zu sehen. In der Zeitreihenanalyse geht man davon aus, daß sich die zeitlich geordneten Werte einer Datenreihe auf bestimmte Komponenten zurückführen lassen. Ein einfacher Zugang zur

zugrundeliegenden zeitlichen Struktur ist die Zerlegung einer Zeitreihe y_1, y_2, \dots, y_n in eine Trendkomponente T_t und eine saisonale bzw. periodische Komponente S_t . Trend und periodische Schwankung werden im allgemeinen von einer unregelmäßigen Restschwankung R_t überlagert, die zufällige Schwankungen, außergewöhnliche Ereignisse, usw. umfaßt. Einfachste Formen des Zusammenwirkens dieser Komponenten sind additive Verknüpfungen.

Geht man von einem additiven Grundmodell aus, ergibt sich damit, daß sich die Zeitreihe $y_t, t=1, \dots, n$, aus drei Teilprozessen zusammensetzt:

$$y_t = T_t + S_t + R_t \quad (1)$$

Unserer Aufgabe entsprechend wollen wir uns in den nächsten Abschnitten mit der Ermittlung des Trends T_t beschäftigen und prüfen, inwieweit T sich mit der Zeit änderte.

Gleitende Durchschnitte

Die Trendkomponente T_t gibt die allgemeine langfristige Bewegung einer Zeitreihe wieder. Um Trends zu erkennen, werden Glättungsverfahren durchgeführt, bei denen aus der ursprünglichen Zeitreihe systematisch Unebenheiten herausgeglättet werden. Das einfachste Verfahren der Trendermittlung ist die Bestimmung sogenannter gleitender Durchschnitte. Bei der Bestimmung gleitender Durchschnitte berechnet man aus k Zeitreihenwerten das arithmetische Mittel und ordnet diesen Mittelwert dem mittleren der bei der Durchschnittsbildung berücksichtigten Zeitpunkte bzw. Zeitintervalle zu. Wir wählen $k=12$ und bestimmen den sog. gleitenden Durchschnitt 12. Ordnung:

$$T_t = (0,5y_{t-6} + y_{t-5} + y_{t-4} + \dots + y_{t+4} + y_{t+5} + 0,5y_{t+6}) / 12 \quad (2)$$

(Zur Ausschaltung periodischer Schwankungen in einem gleitenden Durchschnitt werden so viel Werte einbezogen, wie die Periodenlänge der Schwankung umfaßt.) Alle Monate gehen mit dem gleichen Gewicht in die Berechnung ein, und da sich die Werte beim Übergang von t zu $t+1$ nur durch einen Wert unterscheiden, sollten sich die berechneten gleitenden Durchschnitte nicht sehr stark unterscheiden, die Reihe sollte glatte "Jahrestemperaturen" andeuten. Man sieht, daß es dabei nicht möglich ist, für die ersten 6 und die letzten 6 Werte der Zeitreihe gleitende Durchschnitte zu berechnen; die dazugehörigen Glättungswerte müßten gesondert festgelegt werden.

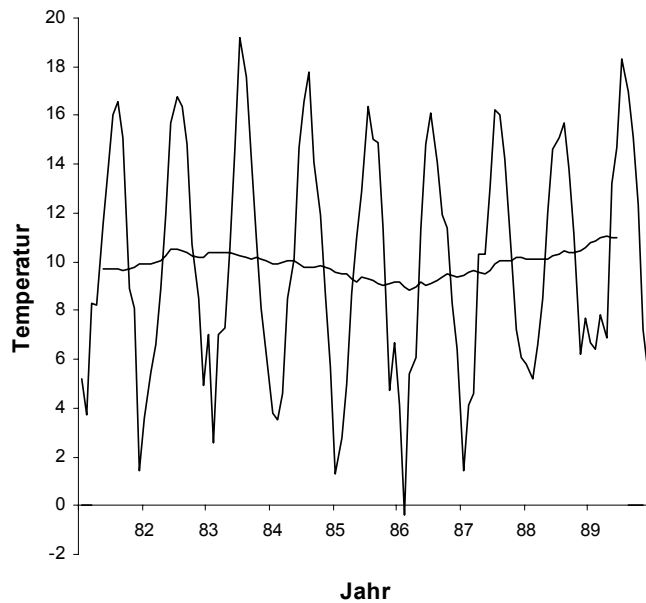


Abb. 1: Mittlere monatliche Temperaturen 1981–1989; Urdaten und gleitender Zwölfmonatsdurchschnitt

In Abb. 1 sind die Temperaturdaten (die Fieber-Sinuskurve) sowie die gleitenden Durchschnitte 12. Ordnung (die relativ glatte Kurve) graphisch dargestellt. Aus dieser Abbildung ist ersichtlich, daß im ersten Teil der Dekade ein Abwärtstrend zu verzeichnen ist, dem ab ca. 1986 ein Aufwärtstrend folgt und daß die saisonale Schwankung periodisch mit Periode 12 ist: $S_t = S_{t+12}$. Dies bedeutet, daß die saisonale Komponente für jeden Monat eines jeden Jahres gleich ist. Tabelle 2 gibt die Werte von T_t wieder. Um das Prinzip der Berechnung der gleitenden Durchschnitte zu verdeutlichen, folgt als Beispiel die Berechnung von $T_{82\text{Oktober}}$:

$$\begin{aligned} T_{82\text{Okto.}} &= (0,5 \times 8,9 + 11,9 + 15,7 + 16,8 + 16,4 + 14,8 + 10,7 + 8,5 + 4,9 + 7 + 2,6 + 7 + 0,5 \times 7,3) / 12 \\ &= 10,366\dots \end{aligned}$$

Saisonale Komponenten

Nach der Berechnung von T_t können wir aus der Zeitreihe den Trend eliminieren und betrachten die neue Reihe:

$$Z_t = y_t - T_t \quad (3)$$

Nach dem Modell in Abschnitt 3 sollten die Werte von Jahr zu Jahr etwa die gleiche Größenordnung haben und sich nur durch Zufallsschwankungen unterscheiden. Alle Januarwerte von Z_t sollten also um den wahren Wert der saisonalen Komponente des Januars, die wir S_{Jan} bezeichnen wollen, schwanken. Dies müßte für alle Monate gelten. Das Mittel der Januarwerte von Z_t würde eine Schätzung \hat{S}_{Jan} von S_{Jan} sein, entsprechend würde der Mittelwert der Februarwerte von Z_t eine Schätzung \hat{S}_{Febr} von S_{Febr} sein, usw. $S_{\text{Jan}}, \dots, S_{\text{Dez}}$ werden i. allg. saisonale Indizes genannt.

Tab. 2: Gleitende Zwölfmonatsdurchschnitte für die Temperaturdaten (* nicht verfügbar)

81	Jan	*	84	Jän	10,0250	87	Jän	9,3375
81	Feb	*	84	Feb	9,9250	87	Feb	9,4208
81	Mar	*	84	Mär	9,9292	87	Mär	9,5958
81	Apr	*	84	Apr	9,9625	87	Apr	9,6458
81	Mai	*	84	Mai	10,0208	87	Mai	9,5542
81	Jun	*	84	Jun	10,0167	87	Jun	9,4958
81	Jul	9,6750	84	Jul	9,8875	87	Jul	9,6667
81	Aug	9,6833	84	Aug	9,7542	87	Aug	9,8958
81	Sep	9,6875	84	Sep	9,7417	87	Sep	10,0250
81	Okt	9,6458	84	Okt	9,7667	87	Okt	10,0333
81	Nov	9,6875	84	Nov	9,8125	87	Nov	10,0292
81	Dez	9,7792	84	Dez	9,7750	87	Dez	10,1625
82	Jan	9,8917	85	Jän	9,6917	88	Jän	10,1792
82	Feb	9,9167	85	Feb	9,5667	88	Feb	10,1208
82	Mar	9,8958	85	Mär	9,4833	88	Mär	10,0917
82	Apr	9,9583	85	Apr	9,4958	88	Apr	10,1083
82	Mai	10,0500	85	Mai	9,3125	88	Mai	10,1000
82	Jun	10,2125	85	Jun	9,1917	88	Jun	10,1250
82	Jul	10,5000	85	Jul	9,3500	88	Jul	10,2292
82	Aug	10,5208	85	Aug	9,3333	88	Aug	10,3167
82	Sep	10,4167	85	Sep	9,2167	88	Sep	10,4167
82	Okt	10,3667	85	Okt	9,1250	88	Okt	10,4000
82	Nov	10,2458	85	Nov	9,0292	88	Nov	10,3833
82	Dez	10,1458	85	Dez	9,1208	88	Dez	10,4375
83	Jän	10,2000	86	Jän	9,1875	89	Jän	10,5750
83	Feb	10,3500	86	Feb	9,1375	89	Feb	10,7625
83	Mär	10,3750	86	Mär	8,9750	89	Mär	10,8708
83	Apr	10,3625	86	Apr	8,8500	89	Apr	10,9750
83	Mai	10,3583	86	Mai	9,0000	89	Mai	11,0667
83	Jun	10,4000	86	Jun	9,1375	89	Jun	11,0167
83	Jul	10,3250	86	Jul	9,0125	89	Jul	*
83	Aug	10,2292	86	Aug	9,0875	89	Aug	*
83	Sep	10,1667	86	Sep	9,2417	89	Sep	*
83	Okt	10,1167	86	Okt	9,3833	89	Okt	*
83	Nov	10,1458	86	Nov	9,5167	89	Nov	*
83	Dez	10,1292	86	Dez	9,4042	89	Dez	*

Die graphische Darstellung der Z_t -Werte für die einzelnen Monate kann zur Überprüfung unseres Modells dienen. In Abb. 2 sind die Z_t -Werte für die Monate August, Oktober und Dezember dargestellt. Die Dezember-Punkte scheinen anzusteigen, was den Schluß nahelegt, daß die Dezembertage immer wärmer werden (Weiße Weihnachten ade?) Alle Z_t -Werte, zusammen mit den saisonalen Indizes und den Standardabweichungen, (innerhalb der Monate) sind in Tab. 3 dargestellt.

Reihe der Residuen (Residuale)

Nachdem Trend und Saisonschwankungen der Datenmengen analysiert wurden, bleibt die Reihe der Residuen, die "zufällig" sein kann oder auch nicht. Die Residuen sind bestimmt durch:

$$R_t = y_t - T_t - S_t, \quad (4)$$

also die Differenzen zwischen beobachteten und angepaßten Werten. Eine graphische Darstellung der R_t über t könnte auf atypische Datenpunkte hinweisen oder bei der Suche nach anderen Gesetzmäßigkeiten helfen. Die Residuen sollten sich völlig unregelmäßig verhalten und keine Muster aufweisen. Im Gegensatz zu ihrer Bedeutung bei der linearen Regression, ist eine Interpretation bei Zeitreihen wegen des Glättungsvorgangs nicht so selbstverständlich.

Tab. 3: Z_t -Werte, Mittelwerte und Standardabweichungen (SD) für Z_t innerhalb jeden Monats

Monat	Jahr									MEAN	SD
	81	82	83	84	85	86	87	88	89		
Jan	*	-6,29	-3,20	-6,23	-8,39	-5,09	-7,94	-4,38	-3,88	-5,67	1,87
Feb	*	-4,42	-7,75	-6,43	-6,77	-9,54	-5,32	-4,92	-4,36	-6,19	1,81
Mar	*	-3,30	-3,38	-5,33	-4,48	-3,58	-5,00	-3,49	-3,07	-3,95	0,86
Apr	*	-1,06	-3,06	-1,46	-0,80	-2,75	0,65	-1,61	-4,07	-1,77	1,48
Mai	*	1,85	0,24	0,08	1,69	2,30	0,75	1,90	2,13	1,37	0,88
Jun	*	5,49	4,20	4,68	3,71	5,66	3,60	4,48	3,68	4,44	0,80
Jul	6,33	6,30	8,88	6,71	7,05	7,09	6,53	4,87	*	6,72	1,11
Aug	6,92	5,88	7,37	8,05	5,67	5,01	6,10	5,38	*	6,30	1,05
Sep	5,41	4,38	4,03	4,36	5,68	2,66	4,18	3,38	*	4,26	0,98
Okt	-0,75	0,33	0,88	2,13	2,28	2,02	0,27	0,70	*	0,98	1,08
Nov	-1,59	-1,74	-2,05	-1,21	-4,33	-1,22	-2,83	-4,18	*	-2,40	1,26
Dez	-8,38	-5,25	-3,83	-4,07	-2,42	-3,00	-4,06	-2,74	*	-4,22	1,91

Abb. 3 zeigt die zeitliche Darstellung der Residuen für unsere Temperaturdaten. Es gibt keine auffällige Struktur mit der einen Ausnahme, daß die größten Residuen Wintermonaten entsprechen. Wenn man zur Interpretation die Standardabweichungen der Tab. 3 hinzuzieht, kann man schließen, daß die Varianz der R_t für die Wintermonate größer ist. Dies würde intuitiv auch Sinn machen; denn man beobachtet in den Wintermonaten größere Temperaturspannen als im Sommer.

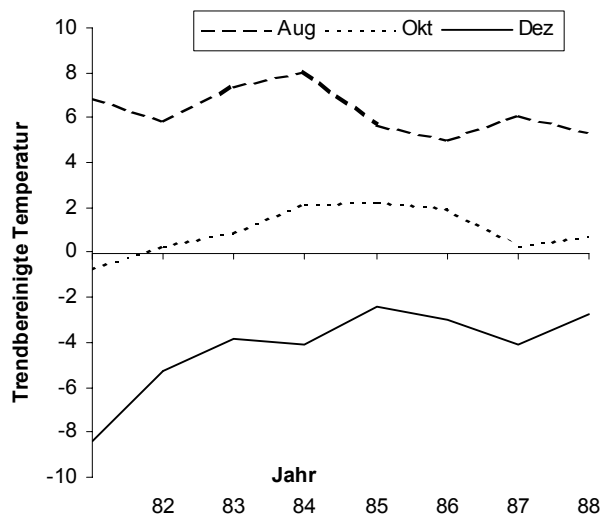


Abb. 2: Saisonabweichungen vom Trend für August, Oktober und Dezember

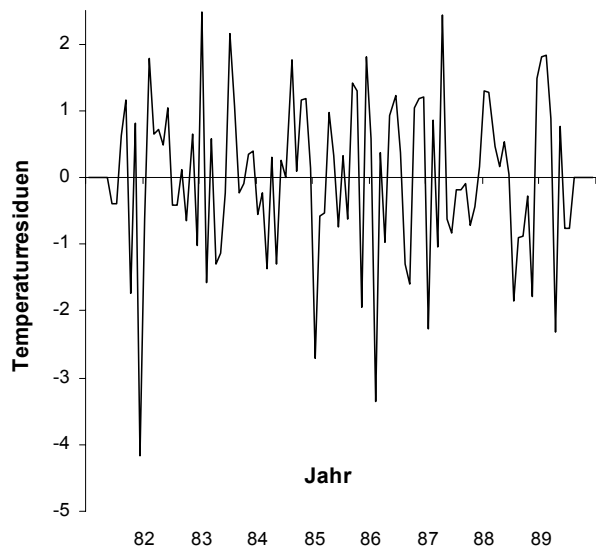


Abb. 3: Graphische Darstellung der Residuen für die Temperaturdaten

Schlußfolgerungen und Ergänzungen

Ab 1986 zeigt das vorgelegte Datenkollektiv der mittleren monatlichen Temperaturwerte einen leichten Aufwärtstrend. Kann daraus globale Erwärmung gefolgert werden? Neuere Daten, wenn verfügbar, werden von großer Bedeutung sein. Hält der Aufwärtstrend an? Zusätzlich wären ältere Daten zum Vergleich von Interesse, weil nur größere Zeiträume in der Meteorologie zuverlässige Aussagen erlauben. Statistischen Verifikationsuntersuchungen von Meteorologen kann z.B. entnommen werden, daß die mittlere bodennahe Lufttemperatur in den letzten 100 Jahren um ca. 0,6 bis 0,7 Grad Celsius zunahm (Schönwiese). In der Nordhemisphäre ist dieser langfristige Temperaturanstieg allerdings zwischen 1949 und 1970 durch eine vorübergehende Abkühlung unterbrochen.

Die vorgestellten Daten stammen aus England. Deutschen Lehrern stehen für die Behandlung in ihrem Unterricht langjährige Datenreihen aus Deutschland oder aus bestimmten Regionen zur Verfügung. Diese Datenreihen betreffen nicht nur Lufttemperaturen, sondern auch Niederschlagshöhen, Ozonwerte oder Höhenwindverlauf. Solche Beobachtungsreihen liefern: Bergs/Geiss/Polster, Hoeschele/Kalb oder Schäfer.

Klimadiagramme werden im Erdkundeunterricht ab der Jahrgangsstufe 6 an vielen Stellen im Unterricht eingesetzt. Unter diesem Aspekt bietet sich eine fachübergreifende Behandlung des Themas an. Erleben wir wirklich eine Klimakatastrophe? Im Zusammenspiel von Geographie- und Mathematikunterricht kommen wir den Dingen auf die Spur. Auf der einen Seite sollen Einsichten über normale und außergewöhnliche Temperaturschwankungen gewonnen, soll Verständnis für das Arbeiten mit Mittelwerten zur Klimabeschreibung erreicht werden. Auf der anderen Seite sollen Kenntnisse der Zeitreihenanalyse vermittelt werden. Ferner können Kenntnisse über Zufallsgrößen, ihre Häufigkeits- bzw. Wahrscheinlichkeitsverteilung und statistische Kenngrößen vermittelt werden. Dies zeigt genau Hilsberg 1991. Im Chemieunterricht könnten zusätzlich Ursachen und mögliche Auswirkungen des anthropogenen Treibhauseffekts diskutiert werden (s. z.B. Adelheim/Höhn 1991).

Literatur

- Adelhelm, M.; Höhn, E.-G. (1991): Zu Behandlung des Treibhauseffekts im Chemieunterricht. *MNU* 44/7, 417-421
- Bergs, W.; Geiss, H.; Polster, G. (1986): *Klimawerte der meteorologischen Station der Kernforschungsanlage Jülich*. Jülich: KfA Zentralbibliothek (Spezielle Berichte der Kernforschungsanlage Jülich, Bd 351)

- Boeker, F. (1991): *Mehr Statistik lernen am PC. Programmbeschreibungen, Übungen und Lernziele zum Statistikprogrammpaket GSTAT2*. Göttingen: Vandenhoeck R Ruprecht
- Boeker, F. (1989): *Statistik lernen am PC. Programmbeschreibungen, Übungen und Lernziele zum Statistikprogramm GSTAT*. Göttingen: Vandenhoeck und Ruprecht
- Frankenberg, P.; Kappas, M. (1991): *Temperatur- und Wetterlagentrends in Westdeutschland*. Mannheim: Selbstverlag des Geographischen Instituts der Universität Mannheim
- Hilsberg, I. (1991): Für die Jahreszeit zu warm? Ein fachübergreifendes Projekt für Geographie u. Mathematik. *Mathematik in der Schule* 29/9, 629-642
- Hoeschele, K.; Kalb, M. (1988): *Das Klima ausgewählter Orte der Bundesrepublik Deutschland, Karlsruhe*. Offenbach: Selbstverlag des Deutschen Wetterdienstes
- Ryan, B.F.; Joiner, B.L.; Ryan, T.A. (1985): *Minitab Handbook*. Boston: Duxbury Press
- Schäfer, P.J. (1982): *Das Klima ausgewählter Orte der Bundesrepublik Deutschland, München*. Offenbach: Deutsches Wetteramt (Berichte d. Dt. Wetterd. Nr. 159)
- Schutz der Erdatmosphäre (1988): *Eine internationale Herausforderung; Zwischenbericht der Enquete-Kommission des 11. Deutschen Bundestages "Vorsorge zum Schutz der Erdatmosphäre"*. Hrsg.: Dr. Bundestag, Referat Öffentlichkeitsarbeit, Bonn, 1988, Zu Sache; 88, 5; ISBN 3-924521-27-1
- "Schutz der Erde" (1990): *Dritter Bericht der Enquete-Kommission "Vorsorge zum Schutz der Erdatmosphäre"*, Drucksache 11/8030 (02.10.1990), Sachgebiet 2129
- Schönwiese, C.D. (1989): Klima in Gefahr. Anthropogene Spurengase verändern das Weltklima. *Siemens-Zeitschrift* 63/1, 32-36
- Schönwiese, C.D.; Malcher, J.; Hartmann, C. (1986): *Globale Statistik langer Temperatur- und Niederschlagsreihen*. Frankfurt: Universität Frankfurt, Inst. f. Meteorologie und Geophysik
- Tufte, E.R. (1983): *The visual display of quantitative information*. Cheshire, Connecticut: Graphics Press