

19

## **Die statistische Auswertung von langjährigen Temperaturreihen**

Von  
Theo Ginsburg

*Bisher erschienen*

Nr. 1 a	Uttinger H.,	Die Niederschlagsstunden in Zürich. 22 Seiten, 1962	Fr. 5.50
Nr. 1 b	Ambrosetti Fl.,	Die Niederschlagsstunden in Locarno-Monti. 12 Seiten, 1965	Fr. 4.50
Nr. 2	Thams J. C.,	unter Mitarbeit von A. Aufdermaur, P. Schmid und E. Zenone. Die Ergebnisse des Grossversuches III zur Bekämpfung des Hagels im Tessin in den Jahren 1957–1963. 32 Seiten, 1966	Fr. 5.—
Nr. 3	Grütter M.,	Die bemerkenswertesten Niederschläge der Jahre 1948–1964 in der Schweiz. 20 Seiten, 1966	Fr. 4.—
Nr. 4	Schram K. und Thams J. C.,	[Redaktion], 9. Internationale Tagung für Alpine Meteorologie in Brig und Zermatt, 14.–17. September 1966, 366 Seiten, 1967	Fr. 30.—
Nr. 5	Ambrosetti Fl. und Thams J. C.,	Die direkte Sonnenstrahlung auf die Flächen eines nach Süden orientierten Würfels ohne Grundfläche in Locarno-Monti. 16 Seiten, 1967	Fr. 3.50
Nr. 6	Schram K. und Thams J. C.,	Der Tagesgang der Abkühlungs- und Aufwärmungsgrösse in Locarno-Monti. 20 Seiten, 1968	Fr. 4.—
Nr. 7	Ambrosetti Fl., Schram K. und Thams J. C.,	Die Intensität der direkten Sonnenstrahlung in verschiedenen Spektralbereichen in Locarno-Monti. 13 Seiten, 1968	Fr. 3.50
Nr. 8	Uttinger H.,	Die Zahl der Tage mit Windspitzen von mindestens 20 Metern pro Sekunde in Zürich (1934–1967). 22 Seiten, 1968	Fr. 5.—
Nr. 9	Mäder F.,	Untersuchung über die Windverhältnisse in Bodennähe bei verschiedenen Wetterlagen. 42 Seiten, 1968	Fr. 7.—
Nr. 10	Schram K.,	Die Windverhältnisse in der bodennahen Luftschicht an einem Hang von etwa 25 Grad Neigung. 13 Seiten, 1968	Fr. 4.—
Nr. 11	Schüepp M.,	Kalender der Wetter- und Witterungslagen von 1955 bis 1967. 44 Seiten, 1968	Fr. 7.—
Nr. 12	Ackermann P.,	Die neue Radiosondenstation Payerne der Schweizerischen Meteorologischen Zentralanstalt. 36 Seiten, 1968	Fr. 6.—
Nr. 13	Junod A.,	Contribution à la méthodologie granulométrique des aérosols amicroscopiques. 70 Seiten, 1969	Fr. 9.—

**MZA-19**  
**JUNI 1970**

**DIE STATISTISCHE AUSWERTUNG  
VON LANGJÄHRIGEN TEMPERATURREIHEN**

**THE STATISTICAL EVALUATION  
OF LONG-RANGE TEMPERATURE SERIES**

**THEO GINSBURG**

Die vorliegenden Untersuchungen wurden  
im Rahmen des Forschungsprojekts No. 2.91.69 des  
«Schweizerischen Nationalfonds» über

**Die Anwendung moderner statistischer Methoden zur  
Auswertung von umfangreichen klimatologischen Daten**  
durchgeführt.

## ABSTRACT

Das Ziel der vorliegenden Untersuchungen besteht in der Entwicklung von Methoden, die erlauben, das umfangreiche Datenmaterial, welches die während Jahrzehnten gesammelten Termin- und Extremwerte der Temperatur einer Station darstellen, auf einen minimalen Umfang zu komprimieren, ohne aber wesentliche Information über deren klimatologisches Verhalten zu verlieren. Zu diesem Zwecke muss ein mathematisches Modell des Temperaturfeldes gefunden werden, welches den tatsächlich gemessenen Verlauf mit Hilfe einiger weniger charakteristischer Zahlen in guter Näherung wiederzugeben erlaubt.

Das im Kap. III.2. kurz beschriebene Markov-Modell eines Zufallsprozesses genügt diesen Bedingungen weitgehend. Zwei notwendige Voraussetzungen für die Brauchbarkeit des Modells werden anhand der Temperaturreihen von Basel aus den Jahren 1901 – 1963 untersucht: einmal wird gezeigt, dass die im Kap. II.5. definierte normierte Temperatur in einem weiten Bereich als normal verteilt angesehen werden kann, und als weiteres wird im Kap. III.5. nachgewiesen, dass das Persistenzverhalten des Temperaturfeldes durch das Markov-Modell in guter Näherung erfasst wird.

Die so verifizierte jahreszeitliche Temperaturstruktur bietet die Möglichkeit, die in langjährigen Messreihen gespeicherte klimatologische Information auf knappstem Raum mittels der Fourierkoeffizienten des Temperaturmittelwertes  $\bar{T}$ , der Standardabweichung  $s$  und des Persistenzkoeffizienten  $a$  wiederzugeben. Dabei werden für die Darstellung der Mittelwerte zwei Harmonische, für die Darstellung der Standardabweichung und des Persistenzkoeffizienten drei Harmonische für die Fouriersynthese herangezogen. Jede Temperaturreihe kann auf diese Weise durch 19 Werte charakterisiert werden: durch fünf Fourierkoeffizienten für den Mittelwert und je sieben Fourierkoeffizienten für die Standardabweichung und den Persistenzkoeffizienten. Die gesamte klimatologische Information der Termin-, Extrem- und Vierermittelwerte ist somit in 114 Werten konzentriert. Mit einfachen Verfahren (Kap. II.4.) lassen sich daraus Monats- und Saisonmittelwerte berechnen.

Um den Informationsgewinn durch Markov-Modelle gegenüber der gebräuchlichen klimatologischen Information (Mittelwerte und evt. Streuung) auch quantitativ erfassen zu können, müssen gewisse Grundbegriffe der Informationstheorie bekannt sein; eine kurze Zusammenfassung derselben findet sich in Kap. III.4. zusammen mit einem praktischen Anwendungsbeispiel. Für die Beurteilung von Prognosen spielt die Informationstheorie eine ausschlaggebende Rolle.

Es ist bisher kein Versuch unternommen worden, die in diesem Bereich ermittelten Resultate in klimatologischem Sinne zu interpretieren. Es ist jedoch zu hoffen, dass im Rahmen unseres vom Schweizerischen Nationalfonds unterstützten Forschungsprojektes eine verstärkte Zusammenarbeit zwischen Klimatologen und Statistikern einsetzen wird, welche erst eine optimale Auswertung der mit statistischen Methoden ermittelten Ergebnisse ermöglichen wird.

## ABSTRACT

The object of the present investigation is to develop methods by which an extensive collection of data representing the daily values of temperature at a station (three measurements at standard times together with maximum and minimum values) over a period of several decades can be compressed into as compact a form as possible without significant loss of climatological information. For this purpose a mathematical model of the temperature series must be found which can reproduce the measured values with good accuracy but which requires only a few characteristic numbers.

Chapter III.2. briefly describes the Markov-model of a random process which is suitable for this purpose. Two assumptions required for this model are proved, using the temperature series at Basel for the period 1901 — 1963. It is first shown that the normalized temperature defined in Chapter II.5. follows a normal Gaussian distribution over a rather wide range; then in Chapter III.5. the behaviour of the persistence of the temperature series is shown to be well represented by the Markov-model.

This satisfactory reproduction of the annual temperature structure leads to the possibility of representing a long series of climatological data in the most economical manner by means of the Fourier coefficients of the mean temperature  $\bar{T}$ , the standard deviation  $s$  and the coefficient of persistence  $a$ . The number of harmonic terms required are 2 for  $\bar{T}$ , 3 for  $s$  and 3 for  $a$ . Thus any temperature series can be represented in this way by 19 characteristic values: 5 for the mean value, and 7 each for the standard deviation and coefficient of persistence. The complete climatological information on the three actual, maximum, minimum and mean values can thus be condensed into 114 numbers. A simple calculation (Chapter II.4.) enables monthly and seasonal means to be derived.

In order to appreciate the gain in information obtained by the Markov-model as compared with the customary climatological mean and standard deviation, a basic knowledge of information theory is required; a short summary is given in Chapter III.4., accompanied by a worked example. Information theory is also of decisive importance in forecast verification.

So far no attempt has been made at a climatological interpretation of the results obtained by these methods. It is hoped, however, that in our research projects supported by the Swiss National Foundation active collaboration between climatologists and statisticians will lead to an optimal application of the results obtained by statistical methods.

## INHALTSVERZEICHNIS

I.	EINLEITUNG	1
I. 1.	Statistische Probleme in Geophysik und Meteorologie	1
I. 2.	Forschungsprojekt zur Auswertung meteorologischer Daten	1
I. 3.	Problemstellung	2
II.	DARSTELLUNG DES TEMPERATURVERLAUFS	3
II. 1.	Glättung der Daten durch Fourieranalyse	3
II. 2.	Die Auswertung der Temperaturdaten von Basel	4
II. 3.	Die Variabilität der Fourierkoeffizienten	4
II. 4.	Bestimmung von Monatsmittelwerten	7
II. 5.	Verteilungsfunktion der normierten Temperatur	9
III.	PERSISTENZ DES TEMPERATURFELDES	12
III.1.	Stochastische Prozesse (Grundbegriffe)	12
III.2.	Das Markov-Modell	13
III.3.	Verallgemeinerte Zeitreihen (Markov-Ketten)	14
III.4.	Der Informationsgewinn durch Markov-Ketten	18
III.4.1.	Einige Grundbegriffe der Informationstheorie (Entropie)	18
III.4.2.	Bedingte Entropie und Information	19
III.4.3.	Beispiel	19
III.5.	Temperaturreihe und Markov-Modell	21
IV.	ZUSAMMENFASSUNG	23
IV.1.	Modell des Temperaturfeldes	23
IV.2.	Ausblick	23
	Literatur	25
	Appendix A: ALGOL-Prozedur für die Fourier-Analyse und -Synthese	26
	Appendix B: Tabelle der Grössen $-p \cdot \ln p$	28
	Verzeichnis der Figuren	31
	Figuren	32–41

## CONTENTS

I.	INTRODUCTION	1
I. 1.	Statistical problems in geophysics and meteorology	1
I. 2.	Research project for the evaluation of meteorological data	1
I. 3.	Outline of problems	2
II.	PRESENTATION OF THE TEMPERATURE DISTRIBUTION	3
II. 1.	Smoothing of data by Fourier analysis	3
II. 2.	The evaluation of the temperature data of Basle	4
II. 3.	The variability of the Fourier coefficients	4
II. 4.	The computation of monthly mean values	7
II. 5.	Frequency distribution of the normalized temperature	9
III.	PERSISTENCE OF THE TEMPERATURE FIELD	12
III.1.	Stochastic processes (basic concepts)	12
III.2.	The Markov model	13
III.3.	Markov Chains	14
III.4.	The information gain by Markov chains	18
III.4.1.	Some basic concepts of information theory (entropy)	18
III.4.2.	Conditional entropy and information	19
III.4.3.	An example	19
III.5.	Temperature series and the Markov model	21
IV.	CONCLUSIONS	23
IV.1.	Model of the temperature field	23
IV.2.	Future developments	23
	Literature	25
	Appendix A: ALGOL procedure for the Fourier analysis and synthesis	26
	Appendix B: Tabulation of the values — p · lb p	28
	List of figures	31
	Figures	32–41

## I. EINLEITUNG

### I. 1. Statistische Probleme in Geophysik und Meteorologie

Die klassischen Methoden der Statistik gehen im allgemeinen von der Voraussetzung aus, dass die beobachteten Ereignisse voneinander unabhängig sind, d.h. dass die Reihenfolge der Messdaten keine Rolle spielt. Bei der statistischen Behandlung meteorologischer und geophysikalischer Daten ist diese Voraussetzung jedoch in den meisten Fällen nicht erfüllt, da die hierbei verarbeiteten Messgrößen gewöhnlich Funktionen des Ortes und der Zeit und somit nicht vertauschbar sind. Zwei Messdaten, deren Ablesung zeitlich oder räumlich nahe beieinander liegen, können deshalb nicht als voneinander unabhängig betrachtet werden. In diesem Sinne liefern beispielsweise die 30 Tagestemperaturen eines Monats an einem bestimmten Ort keinesfalls 30 voneinander unabhängige Informationen, da die Temperatur jedes Tages weitgehend von dem Temperaturverlauf der vorangehenden Tage abhängt: die Temperatur weist eine **Erhaltensneigung** (Persistenz, Autokorrelation) auf. Es kann beispielsweise gezeigt werden, dass sechs an jeweils fünf auseinanderliegenden Tagen gemessene Temperaturwerte etwa denselben statistischen Informationsgehalt aufweisen wie die Gesamtheit aller 30 Messdaten, wenn die Autokorrelation genügend gross ist und man die Information zur Ermittlung des Mittelwertes und der Varianz benutzen will.

Mit den konventionellen statistischen Methoden allein können daher die in der Meteorologie und in der Geophysik auftretenden Probleme nicht gelöst werden. In den vergangenen Jahren wurde nun mit der **Theorie der Zufallsprozesse** ein statistisches Werkzeug entwickelt, welches auf diesen Problemkreis zugeschnitten ist und für alle Arbeiten auf diesem Gebiete nicht mehr ausser Acht gelassen werden kann. Während in der klassischen Wahrscheinlichkeitstheorie ein Versuch definiert wird als Realisation eines Ereignisses, deren Ergebnis durch einen einzigen Messwert  $x$  charakterisiert ist (z.B. die Temperatur  $T$  zu einem bestimmten Zeitpunkt), wird in der Theorie der Zufallsprozesse bei einem Versuch eine ganze kontinuierliche Zeitfunktion  $x(t)$  erzeugt, welche im ganzen Wertebereich des Parameters  $t$  definiert ist. Diese Kurve wird aufgefasst als eine von unendlich vielen Realisierungen eines Zufallsprozesses, dessen Gesamtheit alle möglichen Zeitkurven enthält. Eine spezielle Bedeutung haben hierbei die sog. **stationären Zufallsprozesse**, bei welchen die Verteilungsfunktion  $f(x)$  zu jedem Zeitpunkt immer dieselbe Form mit denselben Parametern hat; sie können für gewisse geophysikalischen Ereignisfolgen mit Persistenz als Modell dienen und sind mathematisch in der Form von Markov-Ketten relativ einfach zu behandeln. — Eine ausgezeichnete, auf die geophysikalische Praxis ausgerichtete Einführung in die Theorie der Zufallsprozesse findet sich in der Monographie von J. Taubenheim [9].

Die stationären Zufallsprozesse sind jedoch nicht nur für Zeitreihen anwendbar. Betrachtet man anstatt zeitlicher Registrierungen an einer festen Station die räumliche Verteilung eines meteorologischen Elements, so kann man auch hier eine Autokorrelation einführen, welche Information liefert über die Abhängigkeit der Messdaten benachbarter Stationen. Handelt es sich um viele in einem grösseren Gebiet verteilte Stationen, so erlaubt die speziell von russischen Forschern entwickelte Methode der "Objektiven Analyse", welche auf dem Prinzip der stationären Zufallsprozesse beruht, eine Glättung der grossräumigen Verteilung. Diese Verteilung kann dann beispielsweise die Ausgangsdaten liefern für numerische Wettervorhersagen, indem die Werte in bestimmten Gitterpunkten berechnet werden. Für eine ausführliche Darstellung dieses Problems sei auf das Werk von L.S. Gandin [2] verwiesen.

Ein anderes Teilgebiet der Statistik, welches für geophysikalische Fragestellungen von speziellem Interesse ist, betrifft Untersuchungen über **Extremalwerte**. Hier sind es vor allem die Anwendungen auf technischem und wirtschaftlichem Gebiet, die eine durch die Theorie und langjährige Messreihen untermauerte Angabe über die Wahrscheinlichkeit des Eintreffens extremer Bedingungen wünschbar machen. Anstoss zu diesen Untersuchungen gaben Probleme im Zusammenhang mit Ueberschwemmungen; in der Folge fand die Theorie — gekoppelt mit den Methoden des kalkulierten Risikos — auch Eingang in viele andere technischen Bereiche (starke Erdbeben für Gebäudestrukturen, Niederschläge für die Kanalisation, extreme Windspitzen für Brücken und Kamine, etc.). Die damit zusammenhängenden Probleme wurden von A. Court [1] diskutiert, welcher sich auf die theoretischen Arbeiten von E.J. Gumbel stützt, dessen Monographie [6] wohl als Standardwerk für dieses Gebiet angesehen werden kann.

### I. 2. Forschungsprojekt zur Auswertung meteorologischer Daten

Die vorliegende Arbeit ist Teil eines vom Schweizerischen Nationalfonds unterstützten Forschungsprojekts, welches sich zum Ziele setzt, moderne statistische Methoden zur Auswertung meteorologischer Daten anzuwenden. Der Zeitpunkt hierzu scheint günstig gewählt: dank der zukunftsweisenden Vorarbeit von PD Dr. M. Schüepp von der Meteorologischen Zentralanstalt in Zürich ist heute ein Grossteil des meteorologischen Datenmaterials der vergangenen 70 Jahre auf Magnetbändern gespeichert; gleichzeitig stehen heute leistungsfähige elektronische Rechenanlagen zur Verfügung, welche eine Auswertung dieser enormen Datenmenge überhaupt erst ermöglichen. Zudem wurden in den vergangenen Jahren neue statistische Methoden entwickelt, deren Anwendung



auf so grosse Datenmengen interessante Resultate erhoffen lässt. – Die bisherigen Untersuchungen wurden auf dem CDC-1604 Computer des Rechenzentrums der ETH durchgeführt; für das Jahr 1970 ist der Einsatz des grösseren Modells CDC-6400/6500 vorgesehen.

Die vorläufige Zielsetzung unseres Forschungsprojekts beschränkt sich auf die Untersuchung von einzelnen Wetterelementen, wie Temperatur, Niederschlag, Wind, Sonnenscheindauer etc. Angestrebt wird eine konzentrierte Darstellung des Klimaverlaufs dieser Elemente einer Station mit Hilfe einiger weniger Zahlenwerte, welche aber ein Maximum an Information aus dem vorhandenen Datenmaterial representieren sollen. Dies bedingt für jedes Element die Wahl eines **mathematischen Modells**, welches nicht nur Mittelwerte und Varianzen, sondern auch den zeitlichen Prozess charakterisieren kann. Dabei ist zu untersuchen, inwieweit diese Daten als Unterlage für die statistische Prognose dienen können.

Neben dem klimatologischen Verlauf der Wetterelemente sollen auch deren Extremwerte untersucht werden. Auch hier handelt es sich um statistische Vorhersagen, jedoch für viel grössere Zeiträume.

Die gleichzeitige Behandlung verschiedener meteorologischer Elemente, wie Temperatur und Feuchtigkeit, Regen und Wind etc., ist einer späteren Phase des Forschungsprojektes vorbehalten, wenn die Charakteristika der einzelnen Wetterelemente festgelegt sind. Ebenfalls für einen späteren Zeitpunkt ist eine Untersuchung der räumlichen Korrelation der Wetterelemente für verschiedene schweizerische Stationen geplant.

### I. 3. Problemstellung

Die Struktur des Temperaturfeldes einer Station soll untersucht werden. Dazu steht die Temperaturreihe von Basel aus den Jahren 1901 – 1963 zur Verfügung: von jedem Tag sind die drei Terminwerte  $T_7, T_{13}, T_{21}$  wie auch die Maximum- und Minimumtemperaturen  $T_{\max}, T_{\min}$  bekannt. Als Tagesmittel wird der Vierermittelwert

$$(1) \quad T_{av} = \frac{1}{4} (T_7 + T_{13} + 2 \cdot T_{21})$$

berechnet und ausgewertet.

Für jeden Tag des Jahres werden für alle sechs Temperaturreihen der Mittelwert und die Standardabweichung ermittelt und der Jahresverlauf durch Fourieranalyse mit anschliessender Synthese geglättet. Die Anzahl der dazu notwendigen Harmonischen ergibt sich aus den Untersuchungen; um für zukünftige vergleichende Untersuchungen eine gemeinsame Grundlage zu schaffen, wird vorgeschlagen, für den Mittelwert die beiden ersten Harmonischen und für die Streuung die drei ersten Harmonischen zur Glättung heranzuziehen. Die mit Hilfe des Mittelwertes  $\bar{T}$  und der Streuung  $s$  (beide von der Jahreszeit abhängig) normierte Temperatur

$$(2) \quad X_j = \frac{T(j) - \bar{T}(j)}{s(j)}$$

kann für die gesamte Dauer der untersuchten Periode von 60 Jahren in erster Näherung als quasistationärer Zufallsprozess betrachtet werden, wobei die Verteilung  $F(X_j)$  den Mittelwert 0 und die Varianz 1 aufweist. Eine Untersuchung der normierten Tagesmittel-Temperaturen  $X$  zeigt, dass diese in guter Näherung als normalverteilt betrachtet werden kann. Wesentliche Abweichungen treten lediglich in den Wintermonaten bei den tiefen Werten auf, sodass sich eine spezielle Analyse der Extremalwerte aufdrängt. Über diese, im besonderen für technische Anwendungen interessante Resultate wird in weiteren Berichten [3, 4] referiert werden.

Eine Untersuchung über das Persistenzverhalten der Temperatur zeigt, dass das in Kap. III. 2. kurz beschriebene **Markov-Modell** der Autokorrelation eine gute Beschreibung der Wirklichkeit liefert. Es genügt deshalb zur Charakterisierung des Modells, den Autokorrelationskoeffizienten  $a = \rho_1$  zweier aufeinanderfolgender Tage zu bestimmen. Einen besseren Schätzwert für  $a$  erhält man allerdings, wenn man eine Funktion der vier ersten Autokorrelationskoeffizienten  $\rho_1, \rho_2, \rho_3, \rho_4$  zur Bestimmung von  $a$  heranzieht. Dieser Persistenzkoeffizient weist auch wieder für alle sechs Temperaturreihen einen charakteristischen Jahresverlauf auf, der mittels der Fourieranalyse mit drei Harmonischen geglättet werden kann.

Ein mathematisches Modell der klimatologischen Temperaturverteilung einer bestimmten Station, welches auf die Normalverteilung der normierten Temperatur und dem Markov-Modell beruht, erlaubt die Charakterisierung der Temperaturstruktur mit einigen wenigen, aber aussagekräftigen Zahlen. Durch die Darstellung des jahreszeitlichen Verlaufs des Mittelwertes, der Standardabweichung und des Persistenzkoeffizienten mittels einiger weniger Harmonischer gelingt es, die in langjährigen Messreihen gespeicherte Information auf knappstem Raum zu komprimieren.

## II. DARSTELLUNG DES TEMPERATURVERLAUFS

### II. 1. Glättung der Daten durch Fourieranalyse

Um bei der Bestimmung der Mittelwerte und der Standardabweichungen der Temperatur trotz der Persistenz die gewohnten statistischen Methoden anwenden zu können, müssen die Daten so zusammen gefasst werden, dass sie in jeder Gruppe unter sich voneinander unabhängig sind. Da in erster Linie der jahreszeitliche Verlauf interessiert, liegt es nahe, bei einer langjährigen Messreihe jeweils die Temperaturwerte eines bestimmten Datums gemeinsam zu verarbeiten. So sind die in Fig. 1 als Beispiel aufgeführten 63 Temperaturdaten der Vierermittelwerte  $T_{av}$  des 9. April der Jahre 1901–1963 stochastisch voneinander unabhängig, wie sich mit einem nichtparametrischen Test leicht nachweisen lässt. Die auf der rechten Seite von Fig. 1 dargestellte Verteilungsfunktion weist dabei auf eine normale Verteilung des Temperaturfeldes hin, doch erlaubt sie wegen der geringen Stichprobengröße natürlich noch keine schlüssige Aussage. Erst die Verarbeitung des gesamten Datenmaterials wird zeigen, dass die (normierte) Temperatur in guter Näherung normalverteilt ist (siehe Kap. II. 5.).

Man kann nun für jedes Datum den Mittelwert  $\bar{T}$  und die Standardabweichung  $s$  der Stichprobe von  $m$  Temperaturwerten berechnen ( $m$ : Anzahl Jahre der Messreihe). Bezeichnen wir mit  $i$  [1:m] den Laufindex der Jahre und mit  $j$  [1:366] den Laufindex der Tage (\*), so ist

$$(3) \quad \bar{T}(j) = \frac{1}{m} \sum_{i=1}^m T_i(j)$$

$$(4) \quad s^2(j) = \frac{1}{m-1} \sum_{i=1}^m [T_i(j) - \bar{T}(j)]^2$$

Alle diese und im folgenden ausgeführten Berechnungen wurden sowohl für die Terminwerte  $T_7, T_{13}, T_{21}$  wie auch für die Extremtemperaturen und den durch (1) definierten Vierermittelwert  $T_{av}$  durchgeführt.

Trägt man die so gefundenen Werte von  $T$  und  $s$  über einer Zeitskala auf, so erhält man wegen der geringen Stichprobengröße ( $m=60$ ) der einzelnen Tage naturgemäss keinen glatten Verlauf (Fig. 2). Als Glättungsmethode zeigt die Fourieranalyse mit nachfolgender Fouriersynthese wesentliche Vorteile gegenüber dem Verfahren der binomischen Glättung, welches einen Spezialfall der überlappenden Mittel mit besonderen Gewichtsfunktionen darstellt (siehe auch [9] S.311 ff). Für den Computer wurde die Fourieranalyse zu diesem Zweck nach dem von C. Runge und R. Terebesi vorgeschlagenen Verfahren als Prozedur programmiert, in Anlehnung an das von Amer entwickelte Bibliotheksprogramm des Rechenzentrums der ETH; das bei den vorliegenden Untersuchungen verwendete ALGOL-Programm ist im Anhang wiedergegeben. Auf die Theorie der Fourieranalyse wird hier nicht eingegangen, da diese in der Fachliteratur ausgiebig behandelt wird (siehe z.B. [8]).

Jede in den 366 Tagen des Jahres definierte Funktion  $F(j)$  kann mit Hilfe der Fourieranalyse geglättet und in folgender Form dargestellt werden:

$$(5) \quad f(j) = a_0 + \sum_{k=1}^N a_k \cos(k \cdot \varphi_j) + \sum_{k=1}^N b_k \sin(k \cdot \varphi_j)$$

$$(6) \quad \varphi_j = j \cdot \frac{2\pi}{366}$$

Die Anzahl  $N$  der in die Glättung miteinbezogenen Harmonischen hängt weitgehend ab von der Art, mit welcher die Amplituden

---

(\*) Ganz allgemein wird bei allen Berechnungen das Jahr in 366 Tage eingeteilt. Die Daten des Schalttages (29. Februar) werden dabei für alle Zwischenschaltjahre durch den Mittelwert der entsprechenden Daten des 28. Februar und des 1. März simuliert. Dies hat den Vorteil, dass jede Messerie gleich viele Werte aufweist und keine Information verloren geht.

$$(7) \quad c_k = \sqrt{a_k^2 + b_k^2} \quad k = 1, 2, \dots, N$$

der k-ten Harmonischen mit wachsendem k abfallen. Ausserdem geben die quadratischen Abweichungen S der geglätteten Werte f(j) von den Stützwerten F(j)

$$(8) \quad S = \sqrt{\frac{1}{366} \sum_{j=1}^{366} [f(j) - F(j)]^2}$$

oder die mittlere Abweichung

$$(9) \quad \Delta = \frac{1}{366} \sum_{j=1}^{366} |f(j) - F(j)|$$

welche natürlich beide von der Anzahl N der zur Glättung herangezogenen Harmonischen abhängen, einen Hinweis auf die beste Wahl von N.

## II. 2. Die Auswertung der Temperaturdaten von Basel

Damit die in Kap. II. 1. erläuterte Darstellung des Temperaturverlaufs für alle Stationen verwendet und auch für Vergleichszwecke herangezogen werden kann, wäre es sehr zu begrüssen, wenn man sich von Anfang an einheitlich auf die Wahl von N für die verschiedenen Funktionen einigen könnte. Um einen Hinweis zu erhalten auf die günstigste Wahl von N für die Mittelwerte und die Standardabweichungen, sind in **Tabelle I** die ausgewerteten Daten von S und  $\Delta$  wiedergegeben für alle Termin-, Extrem- und Vierermittelwerte, wobei N zwischen 1 und 6 variiert wird. Auf Grund dieses Zahlenmaterials und dem Verhalten von  $c_k$  mit zunehmendem k (siehe Tabellen XV, XVI), wird hier folgender Vorschlag gemacht:

(10)	Temperaturmittelwert	$\bar{T}$ : N = 2
	Standardabweichung	s : N = 3
	Persistenzkoeffizient	a : N = 3 (siehe Kap. III. 5.)

Die derart mit N Harmonischen ermittelten Kurven für die Mittelwerte und die Standardabweichungen der Termin-, Extrem- und Vierermittelwerte sind in **Fig. 3** dargestellt, wobei zusätzlich als Streuungsmass noch die mittlere Differenz der Extremwerte

$$(11) \quad \Delta T = \frac{1}{2} (\bar{T}_{\max} - \bar{T}_{\min})$$

eingetragen ist. Charakteristisch scheint der Verlauf der Streuung: für das Temperaturmaximum und  $T_{13}$  ist die jahreszeitliche Variation gering; dies geht auch aus den Werten der Fourierkoeffizienten hervor, wo derjenige der ersten Harmonischen in der Grössenordnung von 10% des Mittelwertes  $c_0$  liegt (siehe Tabellen XV, XVI). Lediglich im Spätherbst sinkt die Standardabweichung auf Werte unter  $4^\circ\text{C}$  ab. Hingegen weisen die vier anderen Temperaturen einen ganz spezifischen Jahresgang der Streuung auf: diese sinkt von über  $5^\circ\text{C}$  im Winter stellenweise fast auf die Hälfte im Sommer, wobei  $T_{\min}$  und  $T_7$  den grössten Schwankungsbereich aufweisen. Inwieweit dieses Verhalten auch bei anderen Stationen anzutreffen oder für Basel charakteristisch ist, wird sich zeigen, wenn auch andere schweizerische Stationen mittels dieser Methode untersucht worden sind. In Zusammenarbeit mit den Klimatologen sollten dann auch die physikalischen Gründe für diesen Jahresgang ermittelt werden.

## II. 3. Die Variabilität der Fourierkoeffizienten

Um die Vertrauenswürdigkeit der Fourierkoeffizienten zu testen, wurden verschiedene Jahresgruppen untersucht (**Tabelle II**). Die Gruppe I enthält die ganze Periode 1901–1960, während die Gruppen II und III die ungeraden, resp. die geraden Jahre dieser Periode umfassen. Ein Vergleich der Fourierkoeffizienten für die Mittelwerte und die Standardabweichungen dieser Jahresgruppen (**Tabelle III**) zeigt, dass die Abweichungen erstaunlich gering sind und sich in der Grössenordnung von einigen wenigen Zehntelsgrad bewegen.

Bedeutend grösser sind die Abweichungen der Fourierkoeffizienten zwischen den Gruppen IV und V, welche die erste, resp. die zweite Hälfte der Gesamtperiode umfassen. Die Trennung dieser beiden Gruppen wurde auf den Jahreswechsel 1928/29 festgelegt, weil zu diesem Zeitpunkt die Messtation verlegt wurde und die zwischen diesen beiden Gruppen auftretenden Differenzen sowohl auf die Verlegung als auch auf eine eventuelle Klimaänderung

Mittelwert							
	N	T <sub>7</sub>	T <sub>13</sub>	T <sub>21</sub>	T <sub>min</sub>	T <sub>max</sub>	T <sub>av</sub>
Quadratische Abweichung S	1	0.541	0.898	0.579	0.484	0.884	0.549
	2	0.447	0.598	0.473	0.424	0.587	0.459
	3	0.437	0.590	0.459	0.420	0.577	0.448
	4	0.423	0.574	0.438	0.390	0.565	0.429
	6	0.392	0.511	0.396	0.361	0.501	0.381
Mittlere Abweichung Δ	1	0.415	0.724	0.466	0.391	0.723	0.438
	2	0.358	0.480	0.379	0.341	0.479	0.369
	3	0.349	0.473	0.367	0.338	0.471	0.356
	4	0.334	0.457	0.354	0.317	0.460	0.345
	6	0.303	0.403	0.320	0.290	0.393	0.302
Standardabweichung							
Quadratische Abweichung S	1	0.519	0.446	0.465	0.500	0.448	0.465
	2	0.452	0.438	0.400	0.459	0.434	0.398
	3	0.393	0.408	0.369	0.373	0.403	0.358
	4	0.384	0.401	0.356	0.358	0.399	0.348
	6	0.380	0.400	0.353	0.351	0.395	0.344
Mittlere Abweichung Δ	1	0.397	0.360	0.355	0.370	0.357	0.363
	2	0.353	0.356	0.306	0.357	0.349	0.307
	3	0.296	0.330	0.283	0.277	0.320	0.277
	4	0.284	0.324	0.267	0.260	0.320	0.266
	6	0.284	0.323	0.264	0.258	0.320	0.263

Tabelle I: Quadratische Abweichung S (8) und mittlere Abweichung Δ (9) der geglätteten Werte von den ursprünglichen Werten bei der Glättung der Temperatur-Vierermittel und der Standardabweichungen durch Fouriersynthese mit N Harmonischen.

No	Jahresgruppe	Anzahl Jahre	Charakterisierung
I	1901 (1) 1960	60	Ganze Periode
II	1901 (2) 1959	30	Ungerade Jahre
III	1902 (2) 1960	30	Gerade Jahre
IV	1901 (1) 1928	28	Erste Hälfte
V	1929 (1) 1960	32	Zweite Hälfte
VI	1901 (1) 1914	14	Erstes Viertel
VII	1915 (1) 1928	14	Zweites Viertel
VIII	1929 (1) 1944	16	Drittes Viertel
IX	1945 (1) 1960	16	Viertes Viertel

Tabelle II: Jahresgruppen der verschiedenen Fourier-Berechnungen.

	I	II	III	IV	V	VI	VII	VIII	IX
Vierermittelwert $T_{av}$									
$a_0$	9.20	9.24	9.16	8.89	9.47	8.72	9.07	9.23	9.71
$a_1$	-8.81	-8.95	-8.66	-8.53	-9.04	-8.63	-8.44	-9.09	-9.00
$a_2$	-0.19	-0.18	-0.21	-0.02	-0.34	-0.15	0.10	-0.38	-0.31
$a_3$	0.12	0.14	0.09	0.13	-0.10	-0.03	0.29	0.01	0.20
$b_1$	-2.56	-2.70	-2.41	-2.43	-2.67	-2.59	-2.27	-2.90	-2.44
$b_2$	0.36	0.40	0.33	0.44	0.29	0.26	0.63	0.26	0.32
$b_3$	-0.10	-0.10	-0.10	0.03	-0.21	-0.31	0.38	-0.15	-0.27
$c_0$	9.20	9.24	9.16	8.89	9.47	8.72	9.07	9.23	9.71
$c_1$	9.17	9.35	8.99	8.88	9.43	9.01	8.74	9.54	9.32
$c_2$	0.41	0.44	0.39	0.44	0.45	0.30	0.63	0.46	0.45
$c_3$	0.15	0.17	0.13	0.14	0.24	0.31	0.48	0.15	0.34
Standardabweichung $s$ von $T_{av}$									
$a_0$	3.64	3.63	3.62	3.53	3.65	3.45	3.47	3.60	3.67
$a_1$	0.51	0.41	0.60	0.46	0.55	0.32	0.51	0.66	0.42
$a_2$	0.30	0.23	0.33	0.27	0.34	0.21	0.31	0.35	0.29
$a_3$	0.13	0.09	0.14	0.13	0.15	0.10	0.14	0.26	0.02
$b_1$	0.38	0.37	0.39	0.30	0.44	0.40	0.16	0.30	0.54
$b_2$	0.14	0.09	0.18	0.03	0.21	0.17	-0.12	0.18	0.24
$b_3$	0.18	0.12	0.24	0.08	0.27	0.08	0.04	0.33	0.21
$c_0$	3.64	3.63	3.62	3.53	3.65	3.45	3.47	3.60	3.67
$c_1$	0.63	0.56	0.72	0.55	0.71	0.51	0.53	0.73	0.69
$c_2$	0.33	0.25	0.38	0.28	0.40	0.27	0.33	0.40	0.37
$c_3$	0.22	0.15	0.28	0.15	0.31	0.12	0.15	0.42	0.21

Tabelle III: Fourierkoeffizienten der Vierermittelwerte  $T_{av}$  und ihrer Standardabweichungen  $s$  für neun verschiedene Jahresgruppen.

	I	II	III	IV	V	VI	VII	VIII	IX
Mittelwerte									
$T_7$	7.00	7.01	6.99	6.69	7.27	6.56	6.83	7.04	7.50
$T_{13}$	12.62	12.70	12.53	11.67	13.44	11.53	11.82	13.18	13.69
$T_{21}$	8.60	8.63	8.56	8.60	8.59	8.38	8.82	8.36	8.82
$T_{min}$	5.55	5.53	5.57	5.98	5.18	5.73	6.22	5.04	5.32
$T_{max}$	14.63	14.72	14.54	13.99	15.20	13.82	14.16	14.92	15.48
$T_{av}$	9.20	9.24	9.16	8.89	9.47	8.72	9.07	9.23	9.71
Standardabweichungen									
$T_7$	3.70	3.70	3.69	3.55	3.72	3.53	3.44	3.67	3.70
$T_{13}$	4.69	4.70	4.66	4.32	4.77	4.24	4.27	4.68	4.82
$T_{21}$	3.67	3.65	3.65	3.59	3.67	3.50	3.54	3.62	3.68
$T_{min}$	3.50	3.51	3.48	3.31	3.56	3.26	3.23	3.56	3.50
$T_{max}$	4.62	4.62	4.59	4.33	4.70	4.23	4.31	4.61	4.75
$T_{av}$	3.64	3.63	3.62	3.52	3.65	3.46	3.47	3.60	3.67

Tabelle IV: Jahresmittel der Termin-, Extrem- und Vierermittelwerte und ihrer Standardabweichungen für neun verschiedene Jahresgruppen.

zurückgeführt werden können. Besonders ist zu beachten, dass der Wert:  $c_0$  von  $T_{av}$ , welches dessen Jahresmittel darstellt, von  $8.89^{\circ}\text{C}$  auf  $9.47^{\circ}\text{C}$  steigt, also um mehr als ein halbes Grad. Um diesen Effekt noch genauer zu untersuchen, wurden die Halbperioden nochmals unterteilt (Gruppen VI-IX). Dabei wird dieser Trend noch besser sichtbar: das Jahresmittel steigt kontinuierlich von  $8.72^{\circ}\text{C}$  im ersten Viertel auf  $9.71^{\circ}\text{C}$  im letzten Viertel, also um ein ganzes Grad an. Hingegen scheint bei der Standardabweichung eher ein Sprung vorzuliegen, welcher von der Standortverlegung herrührt. Allerdings ist die Messreihe nicht genügend lang, um diesen Sprung mit statistischen Tests als signifikant zu deklarieren.

Die hier für den Vierermittelwert  $T_{av}$  dargestellten Berechnungen wurden auch für die drei Terminwerte und die zwei Extremaltemperaturen durchgeführt. Um die oben gewonnene Erkenntnisse zu untermauern, sind in **Tabelle IV** die Jahresmittel der Termin-, Extrem- und Vierermittelwerte wie auch ihrer Standardabweichungen wiedergegeben. Es scheint, dass trotz der wegen der Standortverlegung durchgeführten Korrektur der Terminwerte und von  $T_{av}$  bei  $T_{13}$  und  $T_{21}$  zwischen dem zweiten und dritten Viertel ein Sprung auftritt. Die Extremwerte wurden nicht korrigiert; es hat jedoch den Anschein, dass sie durch die Verlegung im Mittel um mindestens ein Grad erhöht wurden.

Durch die Verlegung werden aber nicht nur die Jahresmittel verändert, sondern auch die Standardabweichungen, welche für die Struktur des Temperaturfeldes eine wesentliche Rolle spielen. Aus den in **Tabelle IV** zusammengestellten Daten scheint klar hervorzugehen, dass die Standardabweichungen bei allen Temperaturreihen durch die Standortverlegung im Jahre 1928 zwischen  $0.1^{\circ}\text{C}$  und  $0.4^{\circ}\text{C}$  - d.h. bis zu 10% - angestiegen sind. Welche Faktoren für diese doch immerhin beträchtliche Modifikation einer meteorologischen Variablen verantwortlich sind, bleibt noch abzuklären.

Während die ersten Harmonischen der Mittelwerte durch die Verlegung kaum beeinflusst werden, weisen in allen Temperaturreihen die ersten Harmonischen der Streuung eine zum Teil beträchtliche Erhöhung auf. Aus **Tabelle III** ist für  $T_{av}$  ersichtlich, dass  $c_1$  vom zweiten zum dritten Viertel (Jahresgruppen VII und VIII) von 0.53 auf 0.73 ansteigt. Bei den anderen Messreihen ist der Sprung teilweise noch ausgeprägter: so springt  $c_1$  bei  $T_{21}$  von 0.50 auf 0.90.

Zusammenfassend lässt sich hier feststellen, dass bei der Verlegung einer Messtation durch eine Korrektur der Temperaturwerte für die Temperatur selbst ein Sprung wohl vermieden werden kann, dass aber unter Umständen die Streuung signifikant verändert wird, was ebenfalls zu einer Modifikation der Temperaturstruktur führt.

#### II. 4. Bestimmung von Monatsmittelwerten

Die Darstellung des jahreszeitlichen Verlaufes mit Hilfe einiger weniger Fourierkoeffizienten erlaubt auch eine ganz einfache Berechnung von Monats- und Saisonmittelwerten, da sich die trigonometrischen Funktionen leicht integrieren lassen. So ergibt sich für den Mittelwert der Funktion  $f$  zwischen den Tagen  $j_1$  und  $j_2$  sofort

$$(12) \quad \overline{f(j_1, j_2)} = a_0 + \sum_{k=1}^N a_k A_k + \sum_{k=1}^N b_k B_k$$

wobei sich die Funktionen  $A_k(j_1, j_2)$  und  $B_k(j_1, j_2)$  nach folgenden Formeln berechnen lassen

$$(13) \quad A_k(j_1, j_2) = \frac{\sin(k \cdot \varphi_{j_2}) - \sin(k \cdot \varphi_{j_1})}{k(\varphi_{j_2} - \varphi_{j_1})}$$

$$(14) \quad B_k(j_1, j_2) = \frac{\cos(k \cdot \varphi_{j_1}) - \cos(k \cdot \varphi_{j_2})}{k(\varphi_{j_2} - \varphi_{j_1})}$$

Diese Werte gelten natürlich ein für allemal für die zwölf Monate und sind in **Tabelle V** zusammengestellt. Für die Bestimmung der Mittelwerte sind nur die vier ersten Kolonnen ( $k=1:2$ ), für die Streuung nur die ersten sechs Kolonnen ( $k=1:3$ ) erforderlich. Die restlichen Kolonnen mit den Werten von  $A_4, B_4, A_5, B_5$  sind lediglich aufgeführt für den Fall, dass die fünf ersten Harmonischen zur Glättung einer meteorologischen Grösse herangezogen werden sollen.

Die Anwendung von Formel (12) auf den Temperatur- und Streuungsverlauf der Jahresgruppe I (ganze Periode) ergibt die in **Tabelle VI** aufgeführten Monats- und Saison-Mittelwerte.

Monat	A <sub>1</sub>	B <sub>1</sub>	A <sub>2</sub>	B <sub>2</sub>	A <sub>3</sub>	B <sub>3</sub>	A <sub>4</sub>	B <sub>4</sub>	A <sub>5</sub>	B <sub>5</sub>
Januar	.954	.260	.822	.484	.626	.642	.398	.718	.174	.709
Februar	.703	.697	.008	.959	-.635	.651	-.843	.014	-.550	-.526
März	.268	.951	-.813	.498	-.658	-.609	.374	-.732	.716	.143
April	-.244	.958	-.840	-.457	.614	-.663	.453	.698	-.706	.239
Mai	-.690	.708	-.024	-.953	.658	.609	-.820	.042	.482	-.548
Juni	-.952	.268	.815	-.499	-.614	.663	.378	-.741	-.146	.731
Juli	-.958	-.243	.838	.455	-.658	-.609	.447	.689	-.234	-.692
August	-.702	-.696	.008	.953	.626	-.642	-.822	.014	.527	.505
September	-.260	-.954	-.824	.485	.647	.631	.403	-.728	-.724	-.178
Oktober	.252	-.956	-.830	-.470	-.626	.642	.423	.704	.701	-.204
November	.696	-.702	-.008	-.956	-.647	-.631	-.832	.014	-.516	.538
Dezember	.954	-.260	.822	-.484	.626	-.642	.398	-.718	.174	-.709
Dez-Feb	.873	.222	.562	.306	.224	.208	.003	.005	-.057	-.168
März-Mai	-.222	.872	-.556	-.302	.200	-.216	-.003	-.005	.173	-.059
Juni-Aug	-.870	-.229	.551	.312	-.211	-.206	-.003	-.005	.051	.176
Sept-Nov	.230	-.872	-.557	-.315	-.213	-.219	.003	.005	-.170	.050

Tabelle V: Monats- und Saison-Koeffizienten A<sub>k</sub> und B<sub>k</sub>, berechnet nach (13) und (14).

Monat	T <sub>av</sub>	s	Monat	T <sub>av</sub>	s	Monate	T <sub>av</sub>	s
Januar	0.16	4.74	Juli	18.27	3.18	Dez-Feb	0.94	4.45
Februar	1.56	4.43	August	17.51	3.12	März-Mai	8.92	3.64
März	4.74	3.77	September	14.26	3.16	Juni-Aug	17.46	3.25
April	8.89	3.52	Oktober	9.41	3.12	Sept-Nov	9.40	3.23
Mai	13.13	3.61	November	4.52	3.39			
Juni	16.56	3.39	Dezember	1.13	4.17			

Tabelle VI: Monats- und Saisonmittelwerte der Temperatur und der Streuung, berechnet nach (12) für die ganze Periode 1901 - 1960 (siehe auch Fig. 2).

## II. 5. Verteilungsfunktion der normierten Temperatur

Mit Hilfe der für jeden Tag des Jahres festgelegten Mittelwertes  $\bar{T}(j)$  und der Standardabweichung  $s(j)$  ist es möglich, die Temperatur  $T_i(j)$  eines beliebigen Tages wie folgt zu normieren:

$$(15) \quad X_i(j) = \frac{T_i(j) - \bar{T}(j)}{s(j)}$$

Für diese normierte Temperaturvariable  $X$  ist der Einfluss der Jahreszeit eliminiert; bezogen auf die gesamte Periode, für welche  $T(j)$  und  $s(j)$  bestimmt wurde, ergeben sich folgende Werte für den Erwartungswert und die Varianz:

$$(16) \quad E(X) = 0; \quad \text{Var}(X) = E(X^2) = 1$$

Die gesamte Messreihe  $X_i(j)$  kann somit in erster Näherung als stationärer Zufallsprozess aufgefasst werden, wobei allerdings beachtet werden muss, dass die Autokorrelation einen gewissen Jahresgang aufweisen kann (siehe Kap. III).

Diese Darstellung der Temperatur erlaubt auch eine sehr übersichtliche Untersuchung über deren Verteilungsfunktion, da die Daten aller Tage der 63 Jahre verarbeitet werden können, was eine Stichprobe von über 20'000 Temperaturwerten ergibt. Auch können die Verteilungsfunktionen von einzelnen Monaten oder Jahreszeiten ermittelt werden, wie dies aus den Tabellen VII und VIII ersichtlich ist. In Tabelle VII sind die absoluten Häufigkeitsverteilungen des normierten Vierermittelwertes  $T_{av}$  von Basel aus 63 Jahren aufgeführt, während der Tabelle VIII die daraus ermittelten Summenhäufigkeitsfunktionen entnommen werden können. Der Bereich von  $X = -3.1$  bis  $X = 3.1$  ist dabei in 31 Klassen eingeteilt, wozu noch die Extrem-Klassen  $X < -3.1$  ( $k = -16$ ) und  $X > 3.1$  ( $k = 16$ ) kommen. In Tabelle VII ist zudem noch für jede Verteilung die Stichprobengrösse und der Wert

$$(17) \quad \chi^2 = \sum_{k=-16}^{16} \frac{(d_k - e_k)^2}{e_k}$$

angegeben, wobei  $d_k$  die praktisch ermittelte,  $e_k$  die bei einer Normalverteilung zu erwartende Belegung der  $k$ -ten Klasse bedeutet. Bei einer Signifikanzzahl  $\alpha = 1\%$  beträgt der kritische Wert für die Verwerfung der Hypothese (Normalverteilung der normierten Temperaturwerte)

$$(18) \quad \chi_{kr}^2 = \chi^2 (\alpha=1\%) \approx 53$$

Dieser Wert wird sowohl für die Verteilung des ganzen Jahres und der Jahreszeiten wie auch für die meisten Monate überschritten; dies lässt sich dadurch erklären, dass grössere Abweichungen von der Normalverteilung speziell bei den Extremalwerten ( $|X| > 2$ ) auftreten. Beschränkt man sich auf eine Intervalleinteilung von 21 Klassen zwischen  $X = -2.1$  und  $X = 2.1$  mit zwei Extrem-Klassen  $X < -2.1$  und  $X > 2.1$ , so fallen die  $\chi^2$ -Werte stark ab und unterschreiten zum Teil wesentlich die kritischen Grenzen.

Prüft man die Normalität der Verteilung ausserdem mit dem Kolmogorov-Test (Fig. 4), so erkennt man, dass bei diesem – mehr die mittleren Werte der Verteilung berücksichtigenden – Verfahren die Hypothese der Normalverteilung nicht verworfen werden muss.

In Fig. 5 ist die aus den Daten von Basel ermittelte Verteilung des ganzen Jahres und des Herbstes der theoretischen Normalverteilung gegenübergestellt. In beiden Fällen ist die Häufigkeit der Werte um den Nullpunkt etwas geringer als theoretisch erwartet, während sie dafür an den Flanken leicht erhöht ist. Bei anderen Jahreszeiten und Monatsverteilungen ist aber auch das gegenteilige Verhalten zu beachten: leicht erhöhte Frequenz bei den Mittelwerten bei leicht reduzierten Flanken. So darf wohl daraus geschlossen werden, dass die normierte Temperatur im Bereich zwischen  $X = -2$  und  $X = 2$  in guter Näherung normalverteilt ist. Hingegen müssen die Extreme einer speziellen Untersuchung unterzogen werden, da deren Frequenz – speziell in den Wintermonaten – weit über den theoretisch zu erwartenden Werten liegt. Da gerade die Extremwerte von meteorologischen Daten im Hinblick auf technische Anwendungen von besonderem Interesse sind, wurden spezifische Methoden zu deren statistischer Erfassung entwickelt, welche in einem anderen Bericht ausführlich behandelt werden sollen [3].

Die normierte Darstellung der Temperatur durch (15) erlaubt die Berechnung eines generalisierten Temperaturindex, welcher für verschiedene Zeitperioden definiert werden kann. Ein derartiger Index wäre sehr nützlich für den Vergleich der Temperaturverläufe in verschiedenen Klima- und Höhegebieten. Ein Vorschlag zur Einführung eines Temperaturindex auf dieser Grundlage wird in einem gegenwärtig in Bearbeitung stehenden Bericht [5] gemacht, worin auch einige Resultate diskutiert werden sollen.



Intervall $X_{k-1} \dots X_k$	k	Jahr	Winter 12 - 2	Frühling 3 - 5	Sommer 6 - 8	Herbst 9 - 11	Jan.	Feb.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
$-\infty \dots -3.10$	-16	37	33	1	0	3	6	22	0	0	1	0	0	0	0	2	1	5
$-3.10 \dots -2.90$	-15	24	19	1	1	3	2	10	1	0	0	1	0	0	0	2	1	7
$-2.90 \dots -2.70$	-14	36	25	4	1	6	5	11	2	2	0	1	0	0	2	3	1	9
$-2.70 \dots -2.50$	-13	50	33	11	3	3	9	11	5	2	4	2	0	1	1	1	1	13
$-2.50 \dots -2.30$	-12	86	37	21	11	17	9	15	99	77	55	4	7	0	3	7	7	13
$-2.30 \dots -2.10$	-11	165	52	47	19	47	22	16	12	21	14	9	8	2	14	22	11	14
$-2.10 \dots -1.90$	-10	254	52	62	71	69	23	13	18	22	22	29	28	14	26	25	18	16
$-1.90 \dots -1.70$	-9	385	91	97	93	104	29	30	29	34	34	26	42	25	25	43	36	32
$-1.70 \dots -1.50$	-8	558	107	141	140	170	40	34	38	61	42	55	42	43	56	69	45	33
$-1.50 \dots -1.30$	-7	740	158	206	200	176	57	54	57	80	69	68	60	72	62	56	58	47
$-1.30 \dots -1.10$	-6	940	166	241	281	252	59	53	89	70	82	90	91	100	78	87	87	54
$-1.10 \dots - .90$	-5	1055	197	266	296	296	84	61	76	86	104	109	90	97	98	82	116	52
$- .90 \dots - .70$	-4	1314	285	337	364	328	96	100	110	120	107	103	130	131	92	100	136	89
$- .70 \dots - .50$	-3	1504	349	378	429	348	122	106	112	135	131	122	149	158	103	123	122	121
$- .50 \dots - .30$	-2	1688	418	387	450	433	115	146	122	123	142	144	146	160	156	135	142	157
$- .30 \dots - .10$	-1	1695	467	387	433	408	146	136	118	119	150	139	151	143	147	132	129	185
$- .10 \dots .10$	0	1787	506	415	436	430	184	138	151	126	138	160	137	139	145	138	147	184
$.10 \dots .30$	1	1738	479	418	420	421	182	142	154	141	123	138	142	140	117	158	146	155
$.30 \dots .50$	2	1659	467	441	385	366	180	149	147	151	143	124	122	139	102	142	122	138
$.50 \dots .70$	3	1471	369	384	337	381	128	105	135	133	116	114	106	117	134	133	114	136
$.70 \dots .90$	4	1360	324	374	303	359	118	101	136	117	121	98	111	94	112	133	114	105
$.90 \dots 1.10$	5	1188	304	311	304	269	101	106	127	91	93	107	107	90	87	98	84	97
$1.10 \dots 1.30$	6	967	259	249	195	264	84	83	91	63	95	68	67	60	94	90	80	92
$1.30 \dots 1.50$	7	775	170	225	187	193	65	55	89	60	76	60	63	64	83	60	50	50
$1.50 \dots 1.70$	8	579	116	154	157	152	38	29	60	42	52	46	50	61	51	55	46	49
$1.70 \dots 1.90$	9	372	82	102	99	89	19	28	32	38	32	29	35	35	38	32	19	35
$1.90 \dots 2.10$	10	219	47	55	66	51	9	12	16	18	21	18	21	27	26	11	14	26
$2.10 \dots 2.30$	11	160	29	45	49	37	12	6	10	12	23	11	20	18	17	7	13	11
$2.30 \dots 2.50$	12	80	14	17	26	23	6	1	2	6	9	8	9	9	13	1	9	7
$2.50 \dots 2.70$	13	48	14	10	11	13	2	3	1	6	3	1	5	5	6	1	6	9
$2.70 \dots 2.90$	14	41	9	6	15	11	0	3	2	3	1	2	9	4	1	2	8	6
$2.90 \dots 3.10$	15	22	3	3	10	6	1	0	2	1	0	3	3	4	1	2	3	2
$3.10 \dots \infty$	16	13	4	0	4	5	0	0	0	0	0	1	2	1	0	1	4	4
Total		23010	5685	5796	5796	5733	1953	1779	1953	1890	1953	1890	1953	1953	1890	1953	1890	1953
$\chi^2$		89.5	315.3	89.4	120.2	77.7	74.5	360.2	78.7	52.2	44.6	42.7	59.6	80.7	75.7	68.7	49.5	112.0

Tabelle VII: Absolute Häufigkeitsverteilung des normierten Temperatur-Vierermittelwertes  $T_{av}$  bei einer Intervalleinteilung von 33 Klassen. Ausgewertet sind die Daten der Jahre 1901 – 1963. In der ersten Kolonne ist das gesamte Datenmaterial verarbeitet (Jahr), während in den folgenden vier Kolonnen eine Aufteilung nach Jahreszeiten und in den restlichen Kolonnen eine Aufteilung nach Monaten vorgenommen wird. Die Zeile "Total" gibt für jede Kolonne die Anzahl der verarbeiteten Daten und in der letzten Zeile sind die nach Formel (17) berechneten Werte von  $\chi^2$  aufgeführt.

Die grossen Werte von  $\chi^2$  – speziell im Februar und im Winter – sind im wesentlichen bedingt durch die Abweichungen der Extremwerte  $|X| > 2$  von der theoretischen Normalverteilung. Bei einer Zusammenfassung der Klassen  $X > 2.1$  und  $X < -2.1$  zu Extremklassen fallen die  $\chi^2$ -Werte stark ab und unterschreiten zum Teil wesentlich die kritischen Grenzen  $\chi^2 (\alpha=1\%) = 41.6$  oder  $\chi^2 (\alpha=5\%) = 35.2$ .

$X_k$	k	Theorie	Jahr	Winter 12 - 2	Frühling 3 - 5	Sommer 6 - 8	Herbst 9 - 11	Jan.	Feb.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
-3.1	-16	0.10	.16	0.58	0.02	0.00	0.05	0.31	1.24	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.10	0.05	0.25
-2.9	-15	0.19	.26	0.91	0.03	0.02	0.10	0.41	1.80	0.05	0.00	0.05	0.05	0.00	0.00	0.00	0.20	0.11	0.61
-2.7	-14	0.35	.42	1.35	0.10	0.03	0.21	0.66	2.42	0.15	0.11	0.05	0.11	0.00	0.00	0.10	0.36	0.16	1.08
-2.5	-13	0.62	.64	1.93	0.29	0.09	0.26	1.13	3.04	0.41	0.21	0.26	0.21	0.00	0.05	0.16	0.41	0.21	1.74
-2.3	-12	1.07	1.01	2.58	0.66	0.28	0.56	1.59	3.88	0.87	0.58	0.51	0.42	0.36	0.05	0.32	0.77	0.58	2.41
-2.1	-11	1.79	1.73	3.50	1.47	0.60	1.38	2.71	4.78	1.48	1.69	1.23	0.90	0.77	0.15	1.06	1.89	1.16	3.12
-1.9	-10	2.87	2.83	4.42	2.54	1.83	2.58	3.89	5.51	2.41	2.86	2.36	2.43	2.20	0.87	2.44	3.17	2.12	3.94
-1.7	-9	4.46	4.51	6.02	4.21	3.43	4.40	5.38	7.19	3.89	4.66	4.10	3.81	4.35	2.15	3.76	5.38	4.02	5.58
-1.5	-8	6.68	6.93	7.90	6.64	5.85	7.36	7.42	9.10	5.84	7.88	6.25	6.72	6.50	4.35	6.72	8.91	6.40	7.27
-1.3	-7	9.68	10.15	10.68	10.20	9.30	10.43	10.34	12.14	8.76	12.12	9.78	10.32	9.58	8.04	10.00	11.78	9.47	9.68
-1.1	-6	13.57	14.23	13.60	14.36	14.15	14.83	13.36	15.12	13.31	15.82	13.98	15.08	14.24	13.16	14.13	16.23	14.07	12.44
-0.9	-5	18.41	18.82	17.06	18.94	19.26	19.99	17.66	18.55	17.20	20.37	19.30	20.85	18.84	18.13	19.31	20.43	20.21	15.11
-0.7	-4	24.20	24.53	22.08	24.76	25.54	25.71	22.58	24.17	22.84	26.72	24.78	26.30	25.50	24.83	24.18	25.55	27.41	19.66
-0.5	-3	30.85	31.06	28.21	31.28	32.94	31.78	28.83	30.13	28.57	33.86	31.49	32.75	33.13	32.92	29.63	31.85	33.86	25.86
-0.3	-2	38.21	38.40	35.57	37.96	40.70	39.33	34.72	38.33	34.82	40.37	38.76	40.37	40.60	41.12	37.88	38.76	41.38	33.90
-0.1	-1	46.02	45.77	43.78	44.63	48.17	46.45	42.19	45.98	40.86	46.67	46.44	47.72	48.34	48.44	45.66	45.52	48.20	43.37
0.1	0	53.98	53.53	52.68	51.79	55.69	53.95	51.61	53.74	48.59	53.33	53.51	56.19	55.35	55.56	53.34	52.58	55.98	52.79
0.3	1	61.79	61.08	61.11	59.01	62.94	61.29	60.93	61.66	56.48	60.79	59.80	63.49	62.62	62.72	59.52	60.67	63.70	60.73
0.5	2	69.15	68.30	69.32	66.62	69.58	67.68	70.15	70.04	64.00	68.78	67.13	70.05	68.87	69.84	64.92	67.94	70.16	67.80
0.7	3	75.80	74.69	75.82	73.24	75.40	74.32	76.70	75.94	70.92	75.82	73.07	76.09	74.30	75.83	72.01	74.76	76.19	74.76
0.9	4	81.59	80.60	81.51	79.69	80.62	80.59	82.74	81.62	77.88	82.01	79.26	81.27	79.98	80.64	77.94	81.56	82.22	80.14
1.1	5	86.43	85.76	86.86	85.06	85.87	85.28	87.92	87.58	84.38	86.82	84.03	86.93	95.46	85.25	82.54	86.58	86.67	85.10
1.3	6	90.32	89.96	91.42	89.36	89.23	89.88	92.22	92.24	89.04	90.16	88.89	90.53	88.89	88.32	87.52	91.19	90.90	89.81
1.5	7	93.32	93.33	94.41	93.24	92.46	93.25	95.55	95.33	93.60	93.33	92.78	93.70	92.12	91.60	91.91	94.26	93.55	92.37
1.7	8	95.54	95.85	96.45	95.89	95.17	95.90	97.49	96.96	96.67	95.55	95.44	96.14	94.68	94.72	94.60	97.08	95.98	94.88
1.9	9	97.13	97.46	97.89	97.65	96.88	97.45	98.46	98.54	98.31	97.57	97.08	97.67	96.47	96.52	96.62	98.72	96.98	96.68
2.1	10	98.21	98.42	98.72	98.60	98.02	98.34	98.93	99.21	99.13	98.52	98.16	98.62	97.54	97.90	97.99	99.28	97.73	98.01
2.3	11	98.93	99.11	99.23	99.38	98.86	98.99	99.54	99.55	99.64	99.15	99.34	99.21	98.57	98.82	98.89	99.64	98.41	98.57
2.5	12	99.38	99.46	99.47	99.67	99.31	99.39	99.85	99.61	99.74	99.47	99.80	99.63	99.03	99.28	99.58	99.69	98.89	98.93
2.7	13	99.65	99.67	99.72	99.84	99.50	99.62	99.95	99.78	99.79	99.79	99.95	99.68	99.28	99.54	99.90	99.74	99.21	99.39
2.9	14	99.81	99.85	99.88	99.95	99.76	99.81	99.95	99.94	99.90	99.95	100.00	99.79	99.75	99.74	99.95	99.84	99.63	99.70
3.1	15	99.90	99.94	99.93	100.00	99.93	99.91	100.00	100.00	100.00	100.00	100.00	99.95	99.90	99.95	100.00	99.95	99.79	99.80
∞	16	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Tabelle VIII: Prozentuale Summenhäufigkeitwerte  $100 \cdot p(X < X_k)$  des normierten Vierermittelwertes, ausgewertet nach den Daten von Tabelle VII.  
(Diese Funktionen bilden die Ausgangswerte für den Kolmogorov-Test).

### III. PERSISTENZ DES TEMPERATURFELDES

#### III. 1. Stochastische Prozesse (Grundbegriffe)

Sei  $x(t)$  eine Variable, welche für diskrete Zeitwerte  $t$  definiert ist und eine vorgegebene Verteilungsfunktion besitzt, wobei im allgemeinen die Parameter der Verteilung — die Mittelwerte  $\mu_t$ , die Streuungen  $\sigma_t$  usw. — von der Zeit abhängig sind. Dann ist durch die diskreten Werte  $x_i$  eine **Zeitreihe** oder ein **stochastischer Prozess** definiert, welcher mit dem Symbol  $\{x_t\}$  bezeichnet wird:

$$(19) \quad \{x_t\} = \dots x_{i-2}, x_{i-1}, x_i, x_{i+1}, x_{i+2}, \dots$$

Für die meteorologischen Anwendungen werden die diskreten Zeitwerte  $t$ , für welche die Funktion  $x(t)$  definiert ist, in den meisten Fällen äquidistant sein. Als Beispiel mag eine Temperaturreihe gelten, wobei für jeden Tag die Minimaltemperatur charakteristisch ist. Im folgenden werden nur **Prozesse mit äquidistanten Stützstellen** behandelt. Als weiteres setzen wir voraus, dass die **Verteilung von  $x(t)$  normal** sei und daher durch die im allgemeinen zeitabhängigen Parameter  $\mu(t)$  und  $\sigma(t)$  eindeutig charakterisiert ist.

Als **stationärer Prozess** wird eine Zeitreihe bezeichnet, für welche die Mittelwerte  $\mu_i = \mu$  und die Streuungen  $\sigma_i = \sigma$  unabhängig von der Zeit sind. Dabei gilt:

$$(20) \quad E(x_i) = E(x) = \mu$$

$$(21) \quad \text{Var}(x_i) = \text{Var}(x) = E([x - \mu]^2) = E(x^2) - \mu^2 = \sigma^2$$

Eine weitere Bedingung muss für die Autokorrelation  $\rho(x_i, x_j)$  zwischen den Variablen  $x_i$  und  $x_j$

$$(22) \quad \rho(x_i, x_j) = \frac{E([x_i - \mu][x_j - \mu])}{\sqrt{E([x_i - \mu]^2)} \sqrt{E([x_j - \mu]^2)}} = \frac{E(x_i \cdot x_j) - \mu^2}{\sigma^2}$$

erfüllt sein: sie darf nur vom Zeitintervall  $k = |j - i|$  und nicht vom Zeitpunkt  $t$  abhängig sein:

$$(23) \quad \rho(x_i, x_j) = f(|j - i|) = \rho_{|j - i|} = \rho_k$$

Durch eine geeignete lineare Transformation ist es immer möglich, den Erwartungswert  $E(x)$  auf 0 zu normieren, sodass die Gleichungen (20), (21), (22) folgende Form annehmen:

$$(20') \quad E(x) = 0$$

$$(21') \quad \text{Var}(x) = E(x^2) = \sigma^2$$

$$(22') \quad \rho(x_i, x_j) = \frac{E(x_i \cdot x_j)}{\sigma^2}$$

Ein **Zufallsprozess** oder ein **stochastischer Prozess** ist eine stationäre Zeitreihe, bei welcher jede Autokorrelation verschwindet:

$$(24) \quad \rho(x_i, x_j) = \rho_k = 0$$

Ein solcher Zufallsprozess wird mit  $\{\epsilon_t\}$  bezeichnet.

Bei einem **autoregressiven Prozess** (häufig auch als **Zufallsprozess in erweitertem Sinne** bezeichnet), lässt sich der Wert von  $x_i$  aus den  $n$  vorangehenden Werten  $x_{i-1}, x_{i-2}, \dots, x_{i-n}$  nach folgender Formel berechnen:

$$(25) \quad x_i = a_1 x_{i-1} + a_2 x_{i-2} + \dots + a_n x_{i-n} + \epsilon_i$$

Dabei sind die Koeffizienten  $a_k$  Konstanten und  $\epsilon_t$  stellt einen Zufallsprozess dar, dessen Wert zur Zeit  $t=i$  vollständig unabhängig ist von jedem vorangehenden  $x_{i-k}$ . Sei die Variable  $x$  durch eine lineare Transformation bereits auf den Mittelwert 0 normiert, so ergeben sich folgende Erwartungswerte und Varianzen:

$$(26) \quad E(x_i) = E(x_{i-k}) = E(\epsilon_i) = E(\epsilon_i \cdot x_{i-k}) = 0$$

$$(27) \quad E(x_i^2) = E(x_{i-k}^2) = \sigma^2$$

### III. 2. Das Markov-Modell

Wird in der Regressionsformel (25)  $n=1$  gesetzt, so bedeutet dies, dass der Wert von  $x_i$  — bis auf einen stochastischen Anteil — aus dem Wert von  $x_{i-1}$  bestimmt werden kann; hingegen haben die weiter zurückliegenden Werte  $x_{i-k}$  ( $k \geq 2$ ) keinen direkten Einfluss auf  $x_i$ .

Multipliziert man beide Seiten der Gleichung

$$(25') \quad x_i = a x_{i-1} + \epsilon_i \quad (|a| \leq 1)$$

mit  $x_{i-k}$  und bildet den Erwartungswert, so ergibt sich

$$(28) \quad E(x_i \cdot x_{i-k}) = a E(x_{i-1} \cdot x_{i-k})$$

Nach der Division durch  $\sigma^2$  findet man folgende Rekursionsformel für die Autokorrelationskoeffizienten:

$$(29) \quad \rho_k = a \rho_{k-1} \quad \text{für } k \leq 1$$

Wegen  $\rho_0 = 1$  folgt daraus sofort

$$(30) \quad \rho_1 = a; \quad \rho_k = a^k$$

Die Varianz  $\sigma_\epsilon^2$  des Zufallsprozesses  $\epsilon_t$  lässt sich aus (25) bestimmen:

$$(31) \quad E(x_i^2) = a^2 E(x_{i-1}^2) + 2 a E(x_{i-1} \cdot \epsilon_i) + E(\epsilon_i^2)$$

Wegen (26) und (27) folgt daraus

$$(32) \quad \sigma_\epsilon^2 = E(\epsilon_i^2) = (1 - a^2) \sigma^2 = (1 - \rho_1^2) \sigma^2$$

Dies lässt sich folgendermassen interpretieren: Betrachten wir (25) als Prognosegleichung, so setzt sich der prognostizierte Wert  $x_i$  aus zwei Beiträgen zusammen: einem **persistenten** Anteil  $x_p = a x_{i-1}$ , welcher lediglich von dem vorhergehenden bekannten Wert  $x_{i-1}$  abhängt, und einem **stochastischen** Anteil  $x_s$  mit dem Erwartungswert  $E(x_s) = 0$  und der Varianz  $E(x_s^2) = (1 - a^2) \sigma^2$ . Dieses sog. Markov-Modell ist schematisch in Fig. 6 dargestellt, wo auch noch die zwei Extremfälle:

$a = 0$ : rein stochastische Prognose

$a = 1$ : rein persistente Prognose

eingetragen sind. Der durch die Persistenz erklärte Informationsanteil für die Variation von  $x_i$  beträgt somit  $a^2 = \rho_1^2$ , während sich der Anteil der unerklärten (stochastischen) Variation auf  $1 - \rho_1^2$  beläuft.

Soll mit Hilfe des Markov-Modells der Wert von  $x_i$  über  $k$  Zeiteinheiten aus  $x_{i-k}$  prognostiziert werden, so ergibt sich aus (25) und (30) folgender Zusammenhang:

$$(33) \quad x_i = a^k x_{i-k} + \epsilon_i = \rho_k x_{i-k} + \epsilon_i$$

Entsprechend reduziert sich der Informationsanteil aus  $x_{i-k}$  für  $x_i$  auf  $\rho_k^2 = a^{2k}$ , während der stochastische Anteil entsprechend auf  $1 - \rho_k^2 = 1 - a^{2k}$  ansteigt. Zur Illustration sind die Anteile der stochastischen Variation in Tabelle IX für  $a = 0.8$  aufgeführt. (Tabelle IX siehe nächste Seite).

Während bei der Prognose für eine Zeiteinheit ( $k=1$ ) immerhin 64% der Variation durch  $x_{i-1}$  erklärt werden können und nur 36% durch zufällige Prozesse verursacht sind, steigt der stochastische Anteil nach fünf Zeiteinheiten

(k=5) bereits auf 90% an. Dies gibt einen Hinweis darauf, wie stark der Prognosenwert mit wachsender Zeitdauer abnimmt.

k	$\rho_k = a^k$	$\rho_k^2 = a^{2k}$	$1 - \rho_k^2$
1	0.80	0.64	0.36
2	0.64	0.41	0.59
3	0.51	0.26	0.74
4	0.41	0.18	0.82
5	0.33	0.11	0.89

Tabelle IX: Autokorrelationskoeffizient und stochastischer Anteil der Variation von  $x_i$  nach k Zeiteinheiten.

Geht man von der Voraussetzung aus, dass das Markov-Modell den effektiven Vorgang wirklich beschreibt — dass also das theoretische Maximum an erklärter Variation gerade  $\rho_k^2 = a^{2k}$  beträgt —, dann lassen sich auch andere Prognosenmodelle auf ihre Wirksamkeit hin untersuchen, indem der Anteil der unerklärten Variation mit dem theoretisch nicht unterbietbaren Wert  $1 - \rho_k^2 = 1 - a^{2k}$  verglichen wird. Wir beschränken uns auf zwei Spezialfälle:

(34) Klimatologische Prognose :  $x_i = 0 + \epsilon_i$

(35) Persistente Prognose :  $x_i = x_{i-k} + \epsilon_i$

Im ersten Fall lautet die Prognose immer  $x_i = 0$ , unabhängig von dem im vorangehenden Zeitpunkt gemessenen Wert  $x_{i-1}$ , während im zweiten Fall eine vollständige Persistenz angenommen wird. Die Anteile der stochastischen Variation sind dann:

(36)  $x_i = 0$  :  $\sigma_\epsilon^2 = E(\epsilon_i^2) = E(x_i^2) = \sigma^2$

(37)  $x_i = x_{i-1}$  :  $\sigma_\epsilon^2 = E(x_i^2) + E(x_{i-1}^2) - 2 E(x_i \cdot x_{i-1}) = 2(1 - \rho_k) \sigma^2 = 2(1 - a^k) \sigma^2$

Definiert man als Güte G der Prognose das Verhältnis des theoretischen Minimalwertes von  $\sigma_\epsilon^2 = (1 - a^{2k}) \sigma^2$  zu der bei den beiden Prognosenmethoden auftretenden stochastischen Variation, so ergibt sich:

(38) Klimatologische Prognose:  $G_k = \frac{(1 - a^{2k}) \sigma^2}{\sigma^2} = 1 - a^{2k}$

(39) Persistente Prognose:  $G_p = \frac{(1 - a^{2k}) \sigma^2}{2(1 - a^k) \sigma^2} = \frac{1 + a^k}{2}$

Diese Gütefunktionen sind in Fig. 7 in Abhängigkeit von a und k dargestellt. Da man es bei Temperaturverläufen bei Zeiteinheiten von einem Tag im allgemeinen mit Persistentkoeffizienten zwischen  $a = 0.7$  und  $a = 0.85$  zu tun hat (siehe Kap. III. 5.), erkennt man sofort, dass die Güte der klimatologischen Prognose für k=1 sehr schlecht ist gegenüber der gemischten Prognose nach (25). Hingegen ist für eine eintägige Temperaturvorhersage die persistente Prognose nicht viel schlechter als die gemischte Prognose.

### III. 3. Verallgemeinerte Zeitreihen (Markov-Ketten)

Für viele meteorologische Probleme ist es oft zweckmässig, kontinuierliche Variable (wie z.B. die Temperatur) durch bestimmte Einteilungen in Klassen zu diskretisieren. Eine Beobachtung kann dann verschiedene — einander ausschliessende Ereignisse  $A_1, A_2, \dots, A_k, \dots, A_m$  zeitigen, die entsprechend der Einteilung der Klassen mit den Wahrscheinlichkeiten  $p_1, p_2, \dots, p_k, \dots, p_m$  auftreten, wobei wegen der Ausschliessungsbedingung gilt:

$$(40) \quad \sum_{k=1}^m p_k = 1$$

Als Beispiel mit Anwendungen wird im folgenden die Einteilung einer normalverteilten Variablen  $X(\mu, \sigma^2)$  in fünf Klassen betrachtet (Fig. 8):

l	$A_l$	$p_l$	untere Grenze	obere Grenze	Klassenbezeichnung
1	$A_1$	0.125	$-\infty$	$\mu - 1.15\sigma$	extrem unternormal
2	$A_2$	0.250	$\mu - 1.15\sigma$	$\mu - 0.32\sigma$	unternormal
3	$A_3$	0.250	$\mu - 0.32\sigma$	$\mu + 0.32\sigma$	normal
4	$A_4$	0.250	$\mu + 0.32\sigma$	$\mu + 1.15\sigma$	übernormal
5	$A_5$	0.125	$\mu + 1.15\sigma$	$+\infty$	extrem übernormal

Tabelle X: Spezielle Klasseneinteilung einer normalverteilten Variablen.

Betrachten wir wiederum eine Zeitserie  $\{x_t\}$ , so entspricht jedem  $x_i$  die Realisation eines Ereignisses  $A_i$ . Das in Kap. III. 2. diskutierte Markov-Modell geht von der Voraussetzung aus, dass  $x_i$  nur von dem Ergebnis der unmittelbar vorangehenden Beobachtung  $x_{i-1}$ , nicht aber von den Ergebnissen noch früherer Beobachtungen abhängig ist. Übertragen auf das diskrete Modell bedeutet dies, dass die Wahrscheinlichkeit des  $i$ -ten Ereignisses  $A_i$  lediglich abhängig ist von dem  $i-1$ -ten Ereignis  $A_k$ . Eine Ereignisfolge dieser Art wird eine **Markov'sche Kette** genannt. Sie ist charakterisiert durch alle Übergangswahrscheinlichkeiten  $p_{ki}$ , welche angeben, mit welcher Wahrscheinlichkeit das Ereignis  $A_i$  eintritt, wenn die vorangehende Beobachtung das Ereignis  $A_k$  gebracht hat. Die Gesamtheit aller Übergangswahrscheinlichkeiten  $p_{ki}$  kann in der Übergangsmatrix

$$(41) \quad P_{ki} = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1l} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2l} & \dots & p_{2m} \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ p_{k1} & p_{k2} & \dots & p_{kl} & \dots & p_{km} \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ p_{m1} & p_{m2} & \dots & p_{ml} & \dots & p_{mm} \end{pmatrix}$$

zusammengefasst werden. Sind die Übergangswahrscheinlichkeiten von der Zeit unabhängig, so spricht man von **homogenen Markovschen Ketten**.

Das kontinuierliche Markov-Modell, wie es in Kap. III. 2. beschrieben ist, lässt sich für eine normalverteilte Variable  $X(0,1)$  mit der in Tabelle VIII angegebenen Einteilung für einen bestimmten Wert der Autokorrelation  $a = \rho$  nach der folgenden Formel in eine Markov-Kette umwandeln:

$$(42) \quad p_{ki} = \frac{\int_{x_k}^{x_{k+1}} \left\{ \Phi \left[ \frac{y_{i+1} - a \cdot x}{\sqrt{1-a^2}} \right] - \Phi \left[ \frac{y_i - a \cdot x}{\sqrt{1-a^2}} \right] \right\} \varphi(x) dx}{\int_{x_k}^{x_{k+1}} \varphi(x) dx}$$

$$(42) \quad p_{kl} = \frac{\int_{x_k}^{x_{k+1}} \left\{ \Phi \left[ \frac{y_{l+1} - a \cdot x}{\sqrt{1-a^2}} \right] - \Phi \left[ \frac{y_l - a \cdot x}{\sqrt{1-a^2}} \right] \right\} d \Phi(x)}{\Phi(x_{k+1}) - \Phi(x_k)}$$

Dabei bedeuten  $x_k$  und  $x_{k+1}$  die Begrenzungen der k-ten Klasse und analog dazu  $y_l$  und  $y_{l+1}$  die Begrenzungen der l-ten Klasse, während mit  $\varphi(x)$  resp.  $\Phi(x)$  die Verteilungsfunktion resp. die Summenhäufigkeit der normierten Normalverteilung bezeichnet wird.

Die Formel (42) wurde für den Fall  $a = 0.8$  ausgewertet und ergibt folgende Übergangsmatrix:

k \ l	l				
	extrem unter-normal	unter-normal	normal	über-normal	extrem über-normal
extrem unternormal	0.585	0.346	0.063	0.006	0
unternormal	0.174	0.474	0.271	0.078	0.003
normal	0.032	0.271	0.394	0.271	0.032
übernormal	0.003	0.078	0.271	0.474	0.174
extrem übernormal	0	0.006	0.063	0.346	0.585
meteorologisches Mittel	0.125	0.250	0.250	0.250	0.125

Wenn also — um ein Beispiel auszuführen — bei einer Autokorrelation  $a = 0.8$  die Temperatur am Vortage in dem durch Tabelle IX genau festgelegten Bereich "unternormal" gelegen hat, so lauten die Wahrscheinlichkeiten für den folgenden Tag:

extrem unternormale Temperatur:	17.4%	(12.5%)
unternormale Temperatur:	47.4%	(25.0%)
normale Temperatur:	27.1%	(25.0%)
übernormale Temperatur:	7.8%	(25.0%)
extrem übernormale Temperatur:	0.3%	(12.5%)

Vergleichsweise sind in Klammern immer die Wahrscheinlichkeiten angegeben, welche lediglich auf der Kenntnis der meteorologischen Verteilung beruhen. Man erkennt leicht, dass die Angabe der Autokorrelation eine wesentliche zusätzliche Information liefert, welche eine genauere und bessere Prognose erlaubt. Eine Masszahl für diese zusätzliche Information wird in Kap. III. 4. diskutiert.

Die zur Definition einer Markov-Kette gehörige Feststellung, dass das Ergebnis  $x_i$  der i-ten Beobachtung nur von  $x_{i-1}$ , nicht aber von den vorangehenden Beobachtungen  $x_{i-n}$  ( $n \geq 2$ ) abhängig ist, darf nicht zu dem Fehlschluss verleiten, dass zwischen  $x_i$  und  $x_{i-2}$  keine stochastische Beziehung besteht. Denn selbstverständlich beeinflusst das Ergebnis  $x_{i-2}$  den Wert von  $x_{i-1}$  und damit auch  $x_i$ . Man kann somit eine Übergangswahrscheinlichkeit zweiter Ordnung  $p_{kl}^{(2)}$  definieren, welche die Wahrscheinlichkeit des Ergebnisses  $A_l$  beim i-ten Versuch angibt, wenn der i-2-te Versuch das Resultat  $A_k$  gezeitigt hat. Es gilt, wie sich leicht zeigen lässt:

$$(44) \quad p_{kl} = \sum_{\nu=1}^m p_{k\nu} \cdot p_{\nu l}$$

oder in Matrixschreibweise

$$(45) \quad P_{kl}^{(2)} = P_{kl}^2$$

Ganz allgemein lässt sich die Übergangsmatrix n-ter Ordnung  $P_{kl}^{(n)}$  welche die Übergangswahrscheinlichkeiten vom i-n-ten Beobachtungsergebnis auf das i-te Ergebnis charakterisiert, als n-te Potenz der Grundmatrix  $P_{kl} = P_{kl}^{(1)}$  bestimmen:

$$(46) \quad P_{kl}^{(n)} = P_{kl}^n$$

Diese Formel ist ganz analog zu der entsprechenden Beziehung (30) für die Autokorrelationskoeffizienten höherer Ordnung.

Bei zunehmenden zeitlichen Abständen zwischen den Beobachtungen (d.h. mit wachsendem n) wird das Ergebnis von  $x_i$  immer weniger vom Beobachtungswert von  $x_{i-n}$  abhängen, und im Grenzfall  $n \rightarrow \infty$  schliesslich wird (nach sog. **Ergodensatz der Markov-Ketten**) eine vollständige Unabhängigkeit zwischen den Ergebnissen des i-ten und des i-n-ten Versuches erreicht:

$$(47) \quad \lim_{n \rightarrow \infty} P_{kl}^{(n)} = P_l$$

In **Tabelle XI** sind, basierend auf der Markov-Kette  $P_{kl} = P$  für  $a = 0.8$ , die Matrizen  $p^2, p^4, p^8, p^{16}$  aufgeführt; der Angleichungsprozess an die klimatologische Wahrscheinlichkeitsverteilung ist daraus klar ersichtlich. Neben den Matrizen ist jeweils der zugehörige Autokorrelationskoeffizient  $a^n$  angegeben, welcher natürlich mit zunehmendem n gegen 0 strebt.

P :	0.585	0.346	0.063	0.006	0.000	a = 0.8
	0.174	0.474	0.271	0.078	0.003	
	0.032	0.271	0.394	0.271	0.032	
	0.003	0.078	0.271	0.474	0.174	
	0.000	0.006	0.063	0.346	0.585	
p <sup>2</sup> :	0.404	0.385	0.157	0.050	0.004	a = 0.64
	0.193	0.365	0.268	0.149	0.025	
	0.079	0.268	0.306	0.268	0.079	
	0.025	0.149	0.268	0.365	0.193	
	0.004	0.050	0.157	0.385	0.404	
p <sup>4</sup> :	0.251	0.346	0.229	0.139	0.035	a = 0.41
	0.173	0.303	0.254	0.200	0.070	
	0.115	0.254	0.262	0.254	0.115	
	0.070	0.200	0.254	0.303	0.173	
	0.035	0.139	0.229	0.346	0.251	
p <sup>8</sup> :	0.160	0.282	0.249	0.216	0.092	a = 0.168
	0.142	0.266	0.250	0.234	0.108	
	0.125	0.250	0.250	0.250	0.125	
	0.108	0.234	0.250	0.266	0.142	
	0.092	0.216	0.249	0.282	0.160	
p <sup>16</sup> :	0.128	0.254	0.250	0.246	0.122	a = 0.028
	0.127	0.251	0.250	0.249	0.123	
	0.125	0.250	0.250	0.250	0.125	
	0.123	0.249	0.250	0.251	0.127	
	0.122	0.246	0.250	0.254	0.128	

Tabelle XI: Übergangsmatrizen höherer Ordnung mit zugehörigen Autokorrelationskoeffizienten



### III. 4. Der Informationsgewinn durch Markov-Ketten

Wie im vorangehenden Kapitel ausgeführt wurde, liefert die Angabe der Autokorrelation eine wesentliche zusätzliche Information für die Prognose, welche für Markov-Ketten mit Hilfe des in der Informationstheorie geläufigen Begriffes der **Entropie** auch zahlenmässig erfasst werden kann. Es seien hier nur einige wenige grundlegende Eigenschaften der Entropie ohne Beweise aufgeführt; für eine leicht lesbare Einführung in den ganzen Problemkreis ist auf das Buch von Jaglom/Jaglom [7] verwiesen.

#### III. 4. 1. Einige Grundbegriffe der Informationstheorie (Entropie)

Wir gehen aus von einem Versuch  $\alpha$ , dessen Resultat in  $m$  verschiedenen, einander ausschliessenden Ereignissen  $A_1, A_2, \dots, A_m$  bestehen soll, welche mit den Wahrscheinlichkeiten  $p_1, p_2, \dots, p_m$  eintreffen. Versucht man, die Unbestimmtheit des Versuchsergebnisses durch eine Masszahl auszudrücken, so wird man davon ausgehen, dass für  $p=0$  (unmögliches Ereignis) und für  $p=1$  (sicheres Ereignis) keine Unbestimmtheit vorliegt und somit diese Masszahl für diese Fälle verschwinden muss. Diese Forderung wird erfüllt durch den Ausdruck

$$(48) \quad - p \cdot \log p$$

Gestützt auf dieses elementare Unbestimmtheitsmass wird der Grad der Unbestimmtheit des Versuches  $\alpha$  definiert durch den Ausdruck

$$(49) \quad H(\alpha) = - \sum_{i=1}^m p_i \cdot \log p_i \quad \left( \sum_{i=1}^m p_i = 1 \right)$$

wobei  $H(\alpha)$  auf Grund gewisser physikalischer Analogien in Anlehnung an die thermodynamische Terminologie als **Informations-Entropie** oder kurz als **Entropie** bezeichnet wird.

Die Wahl der Basis des Logarithmensystems ist unwesentlich, da dadurch nur die Masseinheit für den Grad der Unbestimmtheit beeinflusst wird. In den meisten Fällen wird die Basis 2 gewählt (mit der Bezeichnung  $\log_2 x = \text{lb } x$ ); die Masseinheit der Entropie ist dann diejenige Unbestimmtheit, welcher ein Versuch mit zwei gleichwahrscheinlichen Ausgängen ( $p_1 = p_2 = 0.5$ ) enthält und wird als **bit** (Abkürzung für "binary digit") bezeichnet. Wählt man als Grundlage die dekadischen Logarithmen, so erhält man als Masseinheit das **dit** ("decimal digit"), welches etwa  $3^{1/3}$  mal so gross ist wie ein bit ( $\text{lb } 10 = 3.32 \approx 3^{1/3}$ ). Im Anhang ist eine Tabelle der binären Werte von  $p \cdot \text{lb } p$  wiedergegeben, welche dem Buch von Jaglom/Jaglom [7] entnommen ist. Alle Berechnungen von Entropien in diesem Bericht werden mit der Einheit "bit" durchgeführt.

Wie leicht ersichtlich, ist die Entropie immer positiv, da alle  $p$  zwischen 0 und 1 liegen und somit gilt

$$(50) \quad - p \cdot \text{lb } p \geq 0$$

Das Gleichheitszeichen ist lediglich für den Fall  $p=0$  und  $p=1$  erfüllt, wie eine kleine Grenzwertbetrachtung sofort zeigt. Die Entropie  $H(\alpha)$  eines Versuches als Mass für die Unbestimmtheit verschwindet dann und nur dann, wenn eine der Wahrscheinlichkeiten  $p_i = 1$  ist und alle anderen Wahrscheinlichkeiten gleich Null sind. Die Aussage  $H(\alpha) = 0$  bedeutet somit, dass der Ausgang des Versuches eindeutig bestimmt ist.

Unter allen Versuchen  $\alpha$  mit  $m$  Ausgängen weist derjenige Versuch  $\alpha_0$  die grösste Unsicherheit auf, bei welcher alle Ereignisse gleich wahrscheinlich sind, d.h. wo

$$(51) \quad p_i = \frac{1}{m} \quad (i = 1, 2, \dots, m)$$

In diesem Falle ist es somit am schwersten, das Resultat des Versuchs vorherzusagen. Für die Entropie ergibt sich dabei

$$(52) \quad H(\alpha_0) = + \sum_{i=1}^m \frac{1}{m} (\text{lb } m) = \text{lb } m = H_{\max}$$

Für einen Versuch mit 8 Ausgängen wird die Entropie somit - je nach Vorgabe der Wahrscheinlichkeiten  $p_i$  - Werte zwischen 0 bit und 3 bit annehmen können.

### III. 4. 2. Bedingte Entropie und Information

Bei der durch (41) definierten Übergangsmatrix können die Übergangswahrscheinlichkeiten auch als bedingte Wahrscheinlichkeiten aufgefasst werden:

$$(53) \quad p_{kl} = p(A_l | A_k)$$

$p_{kl}$  gibt somit die Wahrscheinlichkeit des Eintreffens des Ereignisses  $A_l$  an, wenn der vorangehende Versuch das Ergebnis  $A_k$  gezeitigt hat.

Als **bedingte Entropie** des Versuches  $A_l$  bei Kenntnis des vorangehenden Versuches  $A_k$  definiert man

$$(54) \quad H(A_l | A_k) = \sum_{k=1}^m p_k \left\{ - \sum_{l=1}^m p_{kl} \cdot (\log p_{kl}) \right\} = \sum_{k=1}^m p_k \cdot H(A_k)$$

Die so gewonnene Entropie stellt somit das gemäss der Wahrscheinlichkeit  $p_k$  gewichtete Mittel der einzelnen Entropien bei den  $m$  verschiedene Ergebnissen  $A_k$  des vorangehenden Versuchs dar.

Durch die Kenntnis der Übergangsmatrix wird die Unsicherheit über das Resultat des neuen Versuchs offensichtlich verringert, da das bekannte Resultat des vorangehenden Versuchs eine bessere Prognose erlaubt. Diese zusätzliche Kenntnis kann quantitativ durch den Wert der **Information  $I$**  charakterisiert werden, welche folgendermassen definiert ist:

$$(55) \quad I(A_l | A_k) = H(A) - H(A_l | A_k)$$

Dabei bedeutet

$$(56) \quad H(A) = - \sum_{l=1}^m p_l \cdot (\log p_l)$$

die Entropie des Versuchs, wenn lediglich die totalen (klimatologischen) Wahrscheinlichkeiten und keine Übergangswahrscheinlichkeiten bekannt sind.

Der **relative (oder spezifische) Informationsgewinn  $i$**  ist definiert durch

$$(57) \quad i(A_l | A_k) = \frac{I(A_l | A_k)}{H(A)} = 1 - \frac{H(A_l | A_k)}{H(A)}$$

Sein Bereich umfasst die Werte zwischen 0 und 1.

### III. 4. 3. Beispiel

Die in diesem Kapitel definierten Begriffe werden anhand der Übergangsmatrix (43) erläutert und die Resultate kurz interpretiert.

Als einfacher Versuch wird die in Tabelle X aufgeführte Klasseneinteilung mit den dazugehörigen Wahrscheinlichkeiten betrachtet. Die jeweils zu dem Wert von  $p$  gehörige Entropie wird der Tabelle im Anhang entnommen.

Versuchsausgänge	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$
Wahrscheinlichkeiten ( $p$ )	.125	.250	.250	.250	.125
Entropie [bit]	.375	.500	.500	.500	.375

Tabelle XII: Wahrscheinlichkeiten und Entropien der in Tabelle X gewählten Klasseneinteilung.

Gemäss (56) ergibt sich somit für die Entropie des ganzen Versuchs ein Wert von  $H(A) = 2.25$  bit. Der Maximalwert der Entropie eines Versuchs mit 5 Ausgängen würde bei gleichwahrscheinlichen Ergebnissen ( $p_i = 0.2$ ) mit  $H_{\max} = \log 5 = 2.32$  bit erreicht.

Die durch die Übergangsmatrix (43) verfügbare zusätzliche Information kann mit Hilfe der bedingten Entropie berechnet werden. In der folgenden Matrix (Tabelle XI) sind die zu den Übergangswahrscheinlichkeiten  $p_{kl}$  gehörigen Entropien eingetragen und die bedingten Entropien  $H(A_k)$  durch Summation berechnet.

$k \backslash l$	1	2	3	4	5	$H(A_k)$ [bit]
1	.4525	.5298	.2513	.0443	.0000	1.2779
2	.4390	.5105	.5105	.2871	.0251	1.7722
3	.1589	.5105	.5294	.5105	.1589	1.8682
4	.0251	.2871	.5105	.5105	.4390	1.7722
5	.0000	.0443	.2513	.5298	.4525	1.2779

Tabelle XIII: Entropiewerte zu den Übergangswahrscheinlichkeiten der Übergangsmatrix (43) und aufsummierte bedingte Entropien.

Mit Hilfe der nach (54) berechneten bedingten Entropie

$$(54') \quad H(A_l | A_k) = \sum_{k=1}^m p_k H(A_k) = 1.6726 \text{ bit}$$

lässt sich die Information

$$(55') \quad I(A_l | A_k) = 2.25 - 1.6726 = 0.5774 \text{ bit}$$

und der relative Informationsgewinn

$$(57') \quad i(A_l | A_k) = 1 - \frac{1.6726}{2.25} = 0.2566$$

berechnen.

Auch für die bedingten Entropien kann der relative Informationsgewinn bestimmt werden. Wie aus Tabelle XIV ersichtlich ist, wird er am grössten für die extremen Klassen, während er für die mittlere (normale) Klasse am geringsten ist.

Die Matrix (43) entspricht einer Autokorrelation von  $a = 0.8$ . Es ist klar, dass für einen Autokorrelationswert von  $a = 1$  keine Unsicherheit vorhanden ist und deshalb alle bedingten Entropien verschwinden. Die zugehörige Übergangsmatrix nimmt dann die Form einer Einheitsmatrix an und weist nur Wahrscheinlichkeiten von 0 und 1 auf. Die Information einer solchen Matrix ist maximal und erreicht den Wert der unbedingten Entropie  $H(A)$ , während der relative Informationsgewinn 1 wird.

Um den Zusammenhang zwischen Autokorrelation  $a$  und Information  $I$  aufzuzeigen, werden auch die anderen Matrizen der Tabelle XII mit den Autokorrelationen 0.64, 0.41, 0.168 und 0.028 ausgewertet. Das Resultat ist aus Tabelle XIV ersichtlich, wo die Information und (in Klammern) der relative Informationsgewinn für die verschiedenen Klassen und die Gesamtmatrix bei verschiedenen Werten von  $a$  aufgeführt sind. In Fig. 9 schliesslich ist der Zusammenhang zwischen relativem Informationsgewinn und Autokorrelation graphisch aufgetragen. Geht man von einer von Tag zu Tag geltenden Autokorrelation von  $a = 0.8$  aus, so stellen die weiteren Kolonnen in Tabelle XIV gemäss dem Markov-Modell die Information der Übergangsmatrix  $n$ -ter Ordnung (für einenn-tägige Prognose) dar. Man erkennt den raschen Abfall des Informationsgewinns, welcher bereits nach  $n = 4$  Tagen in allen Klassen praktisch vernachlässigbar ist.

a	1	0.8	0.64	0.41	0.168	0.028
n	0	1	2	4	8	16
lb n	-	0	1	2	3	4
extrem unter normal	2.2500 (1)	0.9721 (.4320)	0.5241 (.2329)	0.1676 (.0745)	0.0183 (.0081)	0.0000 (0)
über normal	2.2500 (1)	0.4778 (.2123)	0.2099 (.0933)	0.0549 (.0244)	0.0048 (.0021)	0.0000 (0)
normal	2.2500 (1)	0.3818 (.1697)	0.1304 (.0579)	0.0217 (.0096)	0.0000 (0)	0.0000 (0)
gesamter Versuch	2.2500 (1)	0.5774 (.2566)	0.2686 (.1194)	0.0748 (.0332)	0.0070 (.0031)	0.0000 (0)

Tabelle XIV: Information und (in Klammern) relativer Informationsgewinn für verschiedene Klassen und verschiedene Werte der Autokorrelationskoeffizienten (siehe auch Fig. 9).

### III. 5. Temperaturreihe und Markov-Modell

Wenn es gelingt nachzuweisen, dass das Persistenzverhalten des Temperaturfeldes durch das Markov-Modell mit guter Näherung erfasst wird, dann kann dieses Verhalten durch eine einzige Masszahl – die Autokorrelation  $a = \rho$  – charakterisiert werden. Natürlich wird  $a$  im allgemeinen noch von der Jahreszeit abhängen, sodass man es nicht mit einem reinen stationären Zufallsprozess zu tun haben wird. Doch zeigt es sich, dass die zeitlichen Änderungen von  $a$  so langsam sind, dass von einem **quasistationären Zufallsprozess** gesprochen werden kann.

Gemäss den in Kap. III. 1. hergeleiteten Beziehungen lässt sich mit Hilfe der bereits ermittelten geglätteten Mittelwerte  $T(j)$  und Varianzen  $s^2(j)$  die Autokorrelation  $\rho_k$  zwischen den Temperaturwerten von zwei Tagen, welche  $k$  Tage auseinanderliegen, folgendermassen bestimmen:

$$(58) \quad \rho_k(j) = \frac{\frac{1}{m} \sum_{i=1}^m [T_i(j) - \bar{T}(j)] [T_i(j+k) - \bar{T}(j+k)]}{s(j) \cdot s(j+k)} = \frac{1}{m} \sum_{i=1}^m X_i(j) \cdot X_i(j+k)$$

Dabei bezeichnen wir mit  $i[1:m]$  wiederum den Laufindex der Jahre und mit  $j[1:366]$  den Laufindex der Tage. Mit  $X$  wird die normierte Temperaturvariable bezeichnet, wie sie durch (15) definiert ist. Für den Grenzfall  $n=0$  ergibt sich direkt die Varianz, welche wegen der normierten Darstellung von  $X$  den Wert 1 annimmt:

$$(59) \quad \rho_0(j) = 1$$

Ganz analog mit dem Vorgehen zur Bestimmung der geglätteten Mittelwerte und Varianzen werden auch hier zunächst für jeden Tag  $j$  mit einer Stichprobe von  $m$  Werten aus  $m$  Jahren die Autokorrelationen berechnet und dann mit Hilfe der Fourieranalyse und -synthese die so entstehenden Zeitreihen geglättet. Die numerischen Untersuchungen zeigen, dass  $N = 3$  Harmonische für die Darstellung des jahreszeitlichen Verlaufs der Autokorrelationen als optimal angesehen werden können.

Die charakteristische Beziehung für das Markov-Modell ist durch die Verknüpfung der Autokorrelationskoeffizienten gegeben:

$$(30) \quad \rho_1 = a_1 ; \quad \rho_k = a_k = a_1^k$$

Die Prüfung des Modells kann nun auf die Art durchgeführt werden, dass der jahreszeitliche Verlauf von  $\rho_1, \rho_2, \rho_3, \rho_4$  aus dem 60-jährigen Datenmaterial von Basel für die Termin-, Extrem- und Vierermittelwerte der Temperatur berechnet und dann die Kurven

$$(60) \quad a^{(k)} = k \sqrt{\rho_k(j)} \quad (k = 1, 2, 3, 4)$$

miteinander verglichen werden.

Die Resultate dieser Untersuchungen sind in den **Figuren 10–15** dargestellt. Obwohl natürlich keine vollständige Übereinstimmung der Kurven auftritt, kann doch aus dem ähnlichen Verlauf der Werte von  $a^{(k)}$  geschlossen werden, dass in **erster Näherung das Persistenzverhalten der Temperatur mit dem Markov-Modell gut angenähert werden kann.**

Zur Bestimmung des charakteristischen Persistenzkoeffizienten  $a$  wird das gewichtete Mittel aus den Anteilen  $a^{(1)}$ ,  $a^{(2)}$ ,  $a^{(3)}$ ,  $a^{(4)}$  nach folgender Formel bestimmt:

$$(61) \quad a = a^{(1)} + \frac{a^{(2)}}{2} + \frac{a^{(3)}}{3} + \frac{a^{(4)}}{4}$$

In **Fig. 16** ist der jahreszeitliche Verlauf von  $a$  für die Terminwerte wie für die Extremal- und Vierermittelwerte zusammengestellt. Auch hier zeigt sich wieder wie bei den Mittelwerten und den Varianzen ein charakteristischer Jahresgang, indem im Sommer die Persistenz mit Werten von  $a \approx 0.7$  im allgemeinen wesentlich geringer ist als in den Wintermonaten ( $a \approx 0.8$ ). Auffallend sind die extrem tiefen Werte von  $a = 0.6$  für die Minimaltemperaturen im Sommer.

Alle Kurven in **Fig. 16** sind durch die Fourierkoeffizienten ihrer drei ersten Harmonischen bestimmt; deren Werte sind in den **Tabellen XV und XVI** aufgeführt.

## IV. ZUSAMMENFASSUNG

### IV. 1. Modell des Temperaturfeldes

Das Ziel der vorliegenden Untersuchungen besteht darin, Methoden zu entwickeln, um das ausgedehnte Datenmaterial, welches die während Jahrzehnten gesammelten Termin- und Extremwerte der Temperatur einer Station darstellen, auf einen minimalen Umfang zu komprimieren, ohne aber wesentliche Informationen über deren klimatologisches Verhalten zu verlieren. Dazu muss in erster Linie ein mathematisches Modell des Temperaturfeldes gefunden werden, welches den tatsächlich gemessenen Verlauf mit Hilfe einiger weniger charakteristischer Zahlen in guter Näherung wiederzugeben erlaubt.

Das in Kap. III. 2. kurz beschriebene Markov-Modell eines Zufallsprozesses genügt diesen Bedingungen weitgehend. Als erste Voraussetzung muss abgeklärt werden, ob die sechs untersuchten Temperaturreihen (drei Terminwerte, zwei Extremalwerte und der Vierermittelwert) normal verteilt sind. Zu diesem Zwecke wird mittels Fourieranalyse und -synthese zunächst der geglättete jahreszeitliche Verlauf der Mittelwerte  $\bar{T}$  und der Streuungen  $s$  ermittelt (Kap. II. 1.), womit sich dann für jeden Temperaturwert gemäss (15) die normierte Temperatur  $X$  bestimmen lässt. Wie sich aus den im Kap. II. 5. beschriebenen Untersuchungen ergibt, kann die normierte Temperatur im Bereich zwischen  $X = -2$  und  $X = +2$  als normal verteilt betrachtet werden. In einem zweiten Schritt muss daraufhin nachgewiesen werden, dass das Persistenzverhalten des Temperaturfeldes durch das Markov-Modell in guter Näherung erfasst wird. Dieser Nachweis wird – jedenfalls für die Temperaturreihen der Station Basel – im Kap. III. 5. erbracht.

Die so verifizierte Temperaturstruktur, welche einer Normalverteilung und dem Markov-Modell genügt, bietet nun die Möglichkeit, die in langjährigen Messreihen gespeicherte klimatologische Information auf knappstem Raum durch die Angabe der Fourierkoeffizienten, des Temperaturmittelwertes  $\bar{T}$ , der Standardabweichung  $s$  und des Persistenzkoeffizienten wiederzugeben. Dabei werden für die Darstellung der Mittelwerte zwei Harmonische, für die Darstellung der Standardabweichung und der Persistenzkoeffizienten drei Harmonische für die Fouriersynthese herangezogen. Diese Regelung wurde auch bei der graphischen Darstellung in Fig. 3 und in den Figuren 9–15 angewendet. In den Tabellen XV und XVI sind die wesentlichen Fourierkoeffizienten durch Fettdruck hervorgehoben. Man erkennt, dass für jede Temperaturreihe fünf Koeffizienten für den Mittelwert und je sieben Koeffizienten für die Standardabweichung und den Persistenzkoeffizienten notwendig sind, total also 19 Werte. Für die gesamte klimatologische Information der Termin-, Extrem- und Vierermittelwerte werden somit 114 Werte benötigt. Wie daraus mit einfachsten Mitteln Monats- und Saisonmittelwerte ermittelt werden können, ist aus den Ausführungen in Kap. II. 4. sofort ersichtlich.

Lediglich zur Information sind in den Tabellen XIII und XIV die Koeffizienten je einer zusätzlichen Harmonischen und die Werte von

$$(7) \quad c_k = \sqrt{a_k^2 + b_k^2}$$

aufgeführt, welche Hinweise liefern sollen über den Abfall der Amplituden mit wachsendem  $k$ .

### IV. 2. Ausblick

Es wurde bisher kein Versuch unternommen, die hier gefundenen Resultate in klimatologischem Sinne zu interpretieren. Es ist jedoch zu hoffen, dass im Verlaufe der Durchführung des geplanten Forschungsprojektes eine verstärkte Zusammenarbeit zwischen Klimatologen und Statistikern einsetzen wird, welche erst eine optimale Auswertung der mit statistischen Methoden ermittelten Ergebnisse ermöglichen kann.

Weitere Untersuchungen können auf den Ergebnissen dieses Berichts aufgebaut werden. Es wurde bereits erwähnt, dass die normierte Darstellung der Temperatur durch (15) die Berechnung eines generalisierten Temperaturindex erlaubt, welcher für verschiedene Zeitperioden definiert werden kann; in einer in Vorbereitung stehenden Publikation [5] soll darüber berichtet werden. Ebenfalls sind die Resultate einiger Untersuchungen über das Verhalten der Extremwerte gegenwärtig in Bearbeitung und sollen in zwei weiteren Berichten publiziert werden, wobei einer speziell auf technische Bedürfnisse ausgerichtet sein wird [3, 4].

Das hier dargestellte Markov-Modell für den zeitlichen Ablauf der Temperatur kann dazu benutzt werden, Zeitreihen mit Hilfe von Zufallsprozessen auf den Computer zu generieren, um dann anhand von so erzeugten künstlichen Temperaturreihen weitaus kompliziertere stochastische Zusammenhänge ermitteln zu können. Es sei in diesem Zusammenhang lediglich auf die Arbeiten von I. I. Gringorten verwiesen (z.B. [10]), welcher mit Hilfe derartiger Ansätze in Frequenz und Zeitdauer gewisser Wetterelemente untersucht hat.

k		T <sub>7</sub>			T <sub>13</sub>			T <sub>21</sub>		
		a <sub>k</sub>	b <sub>k</sub>	c <sub>k</sub>	a <sub>k</sub>	b <sub>k</sub>	c <sub>k</sub>	a <sub>k</sub>	b <sub>k</sub>	c <sub>k</sub>
Mittelwert	0	7.01		7.01	12.62		12.62	8.60		8.60
	1	-8.31	-2.73	8.75	-10.06	-2.58	10.39	-8.41	-2.43	8.76
	2	0.42	0.11	0.43	-0.76	0.57	0.95	-0.23	0.41	0.47
	3	0.13	0.03	0.13	0.13	-0.06	0.14	0.08	0.15	0.17
Standard- Abweichung	0	3.68		3.68	4.67		4.67	3.64		3.64
	1	0.94	0.32	0.99	-0.28	0.36	0.45	0.60	0.33	0.69
	2	0.32	0.17	0.36	0.02	0.11	0.11	0.32	0.09	0.33
	3	0.05	0.31	0.31	0.21	0.08	0.23	0.17	0.14	0.22
	4	0.05	0.11	0.12	0.09	0.05	0.10	0.05	0.13	0.14
Persistenz- koeffizient	0	0.708		0.708	0.727		0.727	0.719		0.719
	1	0.016	0.020	0.026	0.053	-0.008	0.053	0.038	0.005	0.038
	2	0.037	0.015	0.040	0.011	0.028	0.030	0.021	0.025	0.033
	3	0.009	0.019	0.021	0.004	0.014	0.014	0.007	0.024	0.025
	4	-0.005	0.004	0.006	0.005	0.001	0.005	-0.010	0.003	0.010

Tabelle XV: Fourierkoeffizienten der Terminwerte (Basel 1901 - 1960).

k		T <sub>min</sub>			T <sub>max</sub>			T <sub>av</sub> = $\frac{1}{4} (T_7 + T_{13} + 2T_{21})$		
		a <sub>k</sub>	b <sub>k</sub>	c <sub>k</sub>	a <sub>k</sub>	b <sub>k</sub>	c <sub>k</sub>	a <sub>k</sub>	b <sub>k</sub>	c <sub>k</sub>
Mittelwert	0	5.56		5.56	14.64		14.64	9.21		9.21
	1	-7.35	-2.87	7.90	-10.53	-2.37	10.80	-8.80	-2.54	9.16
	2	-0.10	0.31	0.33	-0.70	0.62	0.94	-0.20	0.38	0.43
	3	0.04	-0.07	0.08	0.12	-0.08	0.15	0.10	-0.10	0.14
Standard- Abweichung	0	3.48		3.48	4.59		4.59	3.61		3.61
	1	0.92	0.34	0.98	-0.09	0.33	0.34	0.49	0.34	0.59
	2	0.20	0.20	0.28	0.09	0.13	0.16	0.32	0.10	0.34
	3	0.14	0.35	0.38	0.21	0.08	0.23	0.17	0.17	0.25
	4	0.03	0.15	0.15	0.08	0.02	0.08	0.05	0.11	0.12
Persistenz- koeffizient	0	0.708		0.708	0.737		0.737	0.765		0.765
	1	0.083	0.022	0.086	0.043	-0.008	0.044	0.037	-0.006	0.037
	2	0.018	0.004	0.019	0.011	0.030	0.032	0.011	0.026	0.028
	3	0.009	0.042	0.043	0.005	0.014	0.015	0.003	0.017	0.017
	4	-0.003	-0.003	0.004	0.001	-0.001	0.002	-0.004	0.002	0.005

Tabelle XVI: Fourierkoeffizienten der Extrem- und Vierermittelwerte (Basel 1901 - 1960).

Zum Schluss sei noch auf die Bedeutung der Informationstheorie (Kap. III. 4.) und der von ihr entwickelten Begriffe (Entropie, Information etc.) im Hinblick auf die Beurteilung von Prognosen hingewiesen. Mit den heute üblichen Angaben des prozentualen Anteils richtiger Vorhersagen ist keine objektiv wertbare Aussage verbunden. Lediglich die Reduktion der "Unsicherheit" eines zukünftig eintretenden Ereignisses – was in der Informationstheorie als Information bezeichnet wird – liefert ein vernünftiges Mass für die Güte einer Prognose. So kann beispielsweise bei einer Autokorrelation  $a = 0.8$  mit Hilfe der Übergangsmatrix ein spezifischer Informationsgewinn von rund 25% gegenüber der gewöhnlichen, auf rein klimatologischen Daten beruhenden Vorhersagen erzielt werden (Kap. III. 4. 2. und III. 4. 3.). Weitere Studien auf diesem Gebiete drängen sich auf.

---

### Literatur

- [1] Court, A.: Some new statistical Techniques in Geophysics; Advances in Geophysics, Vol. 1, Academic Press, New York, 1952.
- [2] Gandin, L. S.: Objectiv Analysis of meteorological Fields; Leningrad, 1963; translated from Russian, Israel Program of Scientific Translations, Jerusalem, 1965.
- [3] Ginsburg, Th.: Statistische Erfassung von Kälteperioden im Hinblick auf technische Bedürfnisse; (in Vorbereitung).
- [4] – Die Anwendung der Extremwert-Statistik auf meteorologische Daten; (in Vorbereitung).
- [5] – Vorschlag zur Einführung eines generalisierten Temperaturindex; (in Vorbereitung).
- [6] Gumbel, E. J.: Statistics of Extremes; Columbia University Press, New York, 1968.
- [7] Jaglom, A. M. & Jaglom, I. M.: Wahrscheinlichkeit und Information; VEB Deutscher Verlag der Wissenschaften, Berlin, 1965.
- [8] Stiefel, E.: Numerische Mathematik; Teubner, Stuttgart, 1965.
- [9] Taubenheim, J.: Statistische Auswertung geophysikalischer und meteorologischer Daten; Akademische Verlagsgesellschaft GEEST & Portig, Leipzig, 1969.
- [10] Gringorten, I. I.: A stochastic Model of the Frequency and Duration of Weather Elements; Air Force Cambridge Research Laboratories, AFCRL-67-0125, February 1967.



### Appendix A

#### ALGOL-Prozedur für die Fourier-Analyse und Synthese

```
'PROCEDURE' FOURANSY(F,N,H,HH) RESULT:(A,B,G,SQR,SAB);
  'VALUE' N,H,HH;   'INTEGER' N,H,HH;   'REAL' SQR,SAB;
  'ARRAY' F,A,B,G;

'COMMENT'   FOURIER-ANALYSE NACH 'RECHENSCHABLONEN FUER HARMONISCHE
            ANALYSE' NACH C.RUNGE UND P.TEREBESI.
            VEREINFACHTE FORM DES BIBLIOTHEKPROGRAMMS 'FOURAS' (AMER).
            ANALYSE UND SYNTHESE VON H HARMONISCHEN.

F[I] = A[0] + S( A[J]*COS(J*Y[I]) ) + S( B[J]*SIN(J*Y[I]) )
Y[I] = I/N * 2*PI .
N      ANZAHL EINTEILUNGEN DER PERIODE (MUSS GERADE SEIN)
F[I]   URSPRUENGLICHE FUNKTIONSWERTE   I = [0:N-1]
      (GEGEBEN WENN H≥0)
G[I]   BERECHNETE FUNKTIONSWERTE MIT H HARMONISCHEN
H      ANZAHL DER ZU BERECHNENDEN FOURIERKOEFFIZIENTEN
      H≤0: ANALYSE WIRD NICHT DURCHGEFUEHRT
      H=0: DIE URSPRUENGLICHEN FUNKTIONSWERTE F[I] SIND
            GEGEBEN, SQR UND SAB WERDEN BERECHNET
HH     ANZAHL DER HARMONISCHEN FUER DIE SYNTHESE
      HH≤0: SYNTHESE WIRD NICHT DURCHGEFUEHRT
A,B    FOURIERKOEFFIZIENTEN IN DER OBIGEN DARSTELLUNG
      J = [0:H]           H≤10
SQR    MITTLERE QUADRATISCHE ABWEICHUNG (BERECHNET WENN H≥0)
SAB    MITTLERE ABSOLUTE ABWEICHUNG   (BERECHNET WENN H≥0);

'BEGIN'
'REAL' P,Q,PHI,S1,S2,SA,SQ;
'INTEGER' I,J,K,M,MA,MB,SIG;

M:=N/2;  MA:=MB:=H;   'IF' H=M 'THEN' MB:=M-1;

'BEGIN' 'ARRAY' X,Y[0:N], U,V[0:M];

            COSINUS SINUS TABELLE:
X[0]:=1;  Y[0]:=0;  PHI:=3.1415926536/M;
P:=X[1]:=COS(PHI);  Q:=Y[1]:=SIN(PHI);
'FOR' I:=1 'STEP' 1 'UNTIL' N-1 'DO'
'BEGIN' X[I+1]:=P*X[I]-Q*Y[I];
      Y[I+1]:=Y[I] +Q*X[I+1])/P;
'END' COS-SIN-TABELLE;
```

```

                                ANALYSE:
'IF' H<=0 'THEN' 'GOTO' SYNTHESE;
U(0):=F(0); U(M):=F(M); V(0):=V(M):=0;
'FOR' I:=1 'STEP' 1 'UNTIL' M-1 'DO'
  'BEGIN' U(I):=F(I)+F(N-I); V(I):=F(I)-F(N-I); 'END' VORBER;

S1:=S2:=0; SIG:=1;
'FOR' I:=0 'STEP' 1 'UNTIL' M 'DO'
  'BEGIN' S1:=S1+U(I); S2:=S2+SIG*U(I); SIG:=-SIG;
  'END' I;
A(0):=S1/N; 'IF' MA=M 'THEN' A(M):=S2/N;

'FOR' K:=1 'STEP' 1 'UNTIL' MB 'DO'
'BEGIN' S1:=S2:=0;
  'FOR' I:=0 'STEP' 1 'UNTIL' M 'DO'
    'BEGIN' J:=K*I-ENTIER(K*I/N)*N;
      S1:=S1+U(I)*X(J); S2:=S2+V(I)*Y(J);
    'END' I;
  A(K):=S1/M; B(K):=S2/M;
'END' K;

                                SYNTHESE:
'IF' HH<=0 'THEN' 'GOTO' FIN;
U(0):=A(0); U(HH):=A(HH); V(0):=V(HH):=0;
'FOR' I:=1 'STEP' 1 'UNTIL' HH 'DO'
  'BEGIN' U(I):=A(I); V(I):=B(I); 'END' VORBER;

S1:=S2:=0; SIG:=1;
'FOR' I:=0 'STEP' 1 'UNTIL' HH 'DO'
  'BEGIN' S1:=S1+U(I); S2:=S2+SIG*U(I); SIG:=-SIG; 'END' I;
SA := ABS(S1-F(0)) + ABS(S2-F(M));
SQ := (S1-F(0))^2 + (S2-F(M))^2;
G(0):=S1; G(M):=S2; G(N):=S1;
'FOR' K:=1 'STEP' 1 'UNTIL' M-1 'DO'
'BEGIN' S1:=S2:=0;
  'FOR' I:=0 'STEP' 1 'UNTIL' HH 'DO'
    'BEGIN' J:=K*I-ENTIER(K*I/N)*N;
      S1:=S1+U(I)*X(J); S2:=S2+V(I)*Y(J);
    'END' I;

  'IF' H>=0 'THEN'
  'BEGIN'
    SA:=SA+ABS(S1+S2-F(K))+ABS(S1-S2-F(N-K));
    SQ:=SQ+(S1+S2-F(K))^2+(S1-S2-F(N-K))^2;
  'END' H>=0;
  G(K):=S1+S2; G(N-K):=S1-S2;
'END' K;

'IF' H>=0 'THEN' 'BEGIN' SQR:=SQRT(SQ/N); SAB:=SA/N; 'END' H>=0;
'END';
FIN: 'END' FOURANSY;
```

Appendix B  
TABELLE DER GRÖSSEN - p · lb p

p	0	1	2	3	4	5	6	7	8	9
0,00	—	0,0100	0,0179	0,0251	0,0319	0,0382	0,0443	0,0501	0,0557	0,0612
0,01	0,0664	0,0716	0,0766	0,0815	0,0862	0,0909	0,0955	0,0999	0,1043	0,1086
0,02	0,1129	0,1170	0,1211	0,1252	0,1291	0,1330	0,1369	0,1407	0,1444	0,1481
0,03	0,1518	0,1554	0,1589	0,1624	0,1659	0,1693	0,1727	0,1760	0,1793	0,1825
0,04	0,1858	0,1889	0,1921	0,1952	0,1983	0,2013	0,2043	0,2073	0,2103	0,2132
0,05	0,2161	0,2190	0,2218	0,2246	0,2274	0,2301	0,2329	0,2356	0,2383	0,2409
0,06	0,2435	0,2461	0,2487	0,2513	0,2538	0,2563	0,2588	0,2613	0,2637	0,2661
0,07	0,2686	0,2709	0,2733	0,2756	0,2780	0,2803	0,2826	0,2848	0,2871	0,2893
0,08	0,2915	0,2937	0,2959	0,2980	0,3002	0,3023	0,3044	0,3065	0,3086	0,3106
0,09	0,3127	0,3147	0,3167	0,3187	0,3207	0,3226	0,3246	0,3265	0,3284	0,3303
0,10	0,3322	0,3341	0,3359	0,3378	0,3396	0,3414	0,3432	0,3450	0,3468	0,3485
0,11	0,3503	0,3520	0,3537	0,3555	0,3571	0,3588	0,3605	0,3622	0,3638	0,3654
0,12	0,3671	0,3687	0,3703	0,3719	0,3734	0,3750	0,3766	0,3781	0,3796	0,3811
0,13	0,3826	0,3841	0,3856	0,3871	0,3886	0,3900	0,3915	0,3929	0,3943	0,3957
0,14	0,3971	0,3985	0,3999	0,4012	0,4026	0,4040	0,4053	0,4066	0,4079	0,4092
0,15	0,4105	0,4118	0,4131	0,4144	0,4156	0,4169	0,4181	0,4194	0,4206	0,4218
0,16	0,4230	0,4242	0,4254	0,4266	0,4277	0,4289	0,4301	0,4312	0,4323	0,4335
0,17	0,4346	0,4357	0,4368	0,4379	0,4390	0,4400	0,4411	0,4422	0,4432	0,4443
0,18	0,4453	0,4463	0,4474	0,4484	0,4494	0,4504	0,4514	0,4523	0,4533	0,4543
0,19	0,4552	0,4562	0,4571	0,4581	0,4590	0,4599	0,4608	0,4617	0,4626	0,4635
0,20	0,4644	0,4653	0,4661	0,4670	0,4678	0,4687	0,4695	0,4704	0,4712	0,4720
0,21	0,4728	0,4736	0,4744	0,4752	0,4760	0,4768	0,4776	0,4783	0,4791	0,4798
0,22	0,4806	0,4813	0,4820	0,4828	0,4835	0,4842	0,4849	0,4856	0,4863	0,4870
0,23	0,4877	0,4883	0,4890	0,4897	0,4903	0,4910	0,4916	0,4923	0,4949	0,4935
0,24	0,4941	0,4947	0,4954	0,4960	0,4966	0,4971	0,4977	0,4983	0,4989	0,4994
0,25	0,5000	0,5006	0,5011	0,5016	0,5022	0,5027	0,5032	0,5038	0,5043	0,5048
0,26	0,5053	0,5058	0,5063	0,5068	0,5072	0,5077	0,5082	0,5087	0,5091	0,5096
0,27	0,5100	0,5105	0,5109	0,5113	0,5118	0,5122	0,5126	0,5130	0,5134	0,5138
0,28	0,5142	0,5146	0,5150	0,5154	0,5158	0,5161	0,5165	0,5169	0,5172	0,5176
0,29	0,5179	0,5182	0,5186	0,5189	0,5192	0,5196	0,5199	0,5202	0,5205	0,5208
0,30	0,5211	0,5214	0,5217	0,5220	0,5222	0,5225	0,5228	0,5230	0,5233	0,5235
0,31	0,5238	0,5240	0,5243	0,5245	0,5247	0,5250	0,5252	0,5254	0,5256	0,5258
0,32	0,5260	0,5262	0,5264	0,5266	0,5268	0,5270	0,5272	0,5273	0,5275	0,5277
0,33	0,5278	0,5280	0,5281	0,5283	0,5284	0,5286	0,5287	0,5288	0,5289	0,5290
0,34	0,5202	0,5293	0,5294	0,5295	0,5296	0,5297	0,5298	0,5299	0,5299	0,5300
0,35	0,5301	0,5302	0,5302	0,5303	0,5304	0,5304	0,5305	0,5305	0,5305	0,5306
0,36	0,5306	0,5306	0,5307	0,5307	0,5307	0,5307	0,5307	0,5307	0,5307	0,5307
0,37	0,5307	0,5307	0,5307	0,5307	0,5307	0,5306	0,5306	0,5306	0,5305	0,5305
0,38	0,5304	0,5304	0,5303	0,5303	0,5302	0,5302	0,5301	0,5300	0,5300	0,5299
0,39	0,5298	0,5297	0,5296	0,5295	0,5294	0,5293	0,5292	0,5291	0,5290	0,5289
0,40	0,5288	0,5286	0,5285	0,5284	0,5283	0,5281	0,5280	0,5278	0,5277	0,5275
0,41	0,5274	0,5272	0,5271	0,5269	0,5267	0,5266	0,5264	0,5262	0,5260	0,5258
0,42	0,5256	0,5255	0,5253	0,5251	0,5249	0,5246	0,5244	0,5242	0,5240	0,5238
0,43	0,5236	0,5233	0,5231	0,5229	0,5226	0,5224	0,5222	0,5219	0,5217	0,5214
0,44	0,5211	0,5209	0,5206	0,5204	0,5201	0,5198	0,5195	0,5193	0,5190	0,5187
0,45	0,5184	0,5181	0,5178	0,5175	0,5172	0,5169	0,5166	0,5163	0,5160	0,5157

p	0	1	2	3	4	5	6	7	8	9
0,46	0,5153	0,5150	0,5147	0,5144	0,5140	0,5137	0,5133	0,5130	0,5127	0,5123
0,47	0,5120	0,5116	0,5112	0,5109	0,5105	0,5102	0,5098	0,5094	0,5090	0,5087
0,48	0,5083	0,5079	0,5075	0,5071	0,5067	0,5063	0,5059	0,5055	0,5051	0,5047
0,49	0,5043	0,5039	0,5034	0,5030	0,5026	0,5022	0,5017	0,5013	0,5009	0,5004
0,50	0,5000	0,4996	0,4991	0,4987	0,4982	0,4978	0,4973	0,4968	0,4964	0,4959
0,51	0,4954	0,4950	0,4945	0,4940	0,4935	0,4030	0,4026	0,4921	0,4916	0,4911
0,52	0,4906	0,4901	0,4896	0,4891	0,4886	0,4880	0,4875	0,4870	0,4865	0,4860
0,53	0,4854	0,4849	0,4844	0,4839	0,4833	0,4828	0,4822	0,4817	0,4811	0,4806
0,54	0,4800	0,4795	0,4789	0,4784	0,4778	0,4772	0,4767	0,4761	0,4755	0,4750
0,55	0,4744	0,4738	0,4732	0,4726	0,4720	0,4714	0,4708	0,4702	0,4697	0,4691
0,56	0,4684	0,4678	0,4672	0,4666	0,4660	0,4654	0,4648	0,4641	0,4635	0,4629
0,57	0,4623	0,4616	0,4610	0,4603	0,4597	0,4591	0,4584	0,4578	0,4571	0,4565
0,58	0,4558	0,4551	0,4545	0,4538	0,4532	0,4525	0,4518	0,4512	0,4505	0,4498
0,59	0,4491	0,4484	0,4477	0,4471	0,4464	0,4457	0,4450	0,4443	0,4436	0,4429
0,60	0,4422	0,4415	0,4408	0,4401	0,4393	0,4386	0,4379	0,4372	0,4365	0,4357
0,61	0,4350	0,4343	0,4335	0,4328	0,4321	0,4313	0,4306	0,4298	0,4291	0,4283
0,62	0,4276	0,4268	0,4261	0,4253	0,4246	0,4238	0,4230	0,4220	0,4215	0,4207
0,63	0,4199	0,4192	0,4184	0,4176	0,4168	0,4160	0,4153	0,4145	0,4137	0,4129
0,64	0,4121	0,4113	0,4105	0,4097	0,4089	0,4080	0,4072	0,4064	0,4056	0,4048
0,65	0,4040	0,4032	0,4023	0,4015	0,4007	0,3998	0,3990	0,3982	0,3973	0,3965
0,66	0,3957	0,3948	0,3940	0,3931	0,3922	0,3914	0,3905	0,3897	0,3888	0,3880
0,67	0,3871	0,3862	0,3854	0,3845	0,3836	0,3828	0,3819	0,3810	0,3801	0,3792
0,68	0,3784	0,3775	0,3766	0,3757	0,3748	0,3739	0,3730	0,3721	0,3712	0,3703
0,69	0,3694	0,3685	0,3676	0,3666	0,3657	0,3648	0,3639	0,3630	0,3621	0,3611
0,70	0,3602	0,3593	0,3583	0,3574	0,3565	0,3555	0,3546	0,3536	0,3527	0,3518
0,71	0,3508	0,3499	0,3489	0,3480	0,3470	0,3461	0,3451	0,3441	0,3432	0,3422
0,72	0,3412	0,3403	0,3393	0,3383	0,3373	0,3364	0,3354	0,3344	0,3334	0,3324
0,73	0,3314	0,3304	0,3295	0,3285	0,3275	0,3265	0,3255	0,3245	0,3235	0,3225
0,74	0,3215	0,3204	0,3194	0,3184	0,3174	0,3164	0,3154	0,3144	0,3133	0,3123
0,75	0,3113	0,3103	0,3092	0,3082	0,3071	0,3061	0,3051	0,3040	0,3030	0,3019
0,76	0,3009	0,2999	0,2988	0,2978	0,2967	0,2956	0,2946	0,2935	0,2925	0,2914
0,77	0,2903	0,2893	0,2882	0,2871	0,2861	0,2850	0,2839	0,2828	0,2818	0,2807
0,78	0,2796	0,2785	0,2774	0,2763	0,2753	0,2741	0,2731	0,2720	0,2709	0,2698
0,79	0,2687	0,2676	0,2664	0,2653	0,2642	0,2631	0,2620	0,2609	0,2598	0,2587
0,80	0,2575	0,2564	0,2553	0,2542	0,2531	0,2519	0,2508	0,2497	0,2485	0,2474
0,81	0,2462	0,2451	0,2440	0,2428	0,2417	0,2405	0,2394	0,2382	0,2371	0,2359
0,82	0,2348	0,2336	0,2324	0,2313	0,2301	0,2290	0,2278	0,2266	0,2255	0,2243
0,83	0,2231	0,2220	0,2208	0,2196	0,2184	0,2172	0,2160	0,2149	0,2137	0,2125
0,84	0,2113	0,2101	0,2089	0,2077	0,2065	0,2053	0,2041	0,2029	0,2017	0,2005
0,85	0,1993	0,1981	0,1969	0,1957	0,1944	0,1932	0,1920	0,1908	0,1896	0,1884
0,86	0,1871	0,1859	0,1847	0,1834	0,1822	0,1810	0,1797	0,1785	0,1773	0,1760
0,87	0,1748	0,1735	0,1723	0,1711	0,1698	0,1686	0,1673	0,1661	0,1648	0,1635
0,88	0,1623	0,1610	0,1598	0,1585	0,1572	0,1560	0,1547	0,1534	0,1522	0,1509
0,89	0,1496	0,1484	0,1471	0,1458	0,1445	0,1432	0,1419	0,1407	0,1394	0,1381
0,90	0,1368	0,1355	0,1342	0,1320	0,1316	0,1303	0,1200	0,1277	0,1264	0,1251
0,91	0,1238	0,1225	0,1212	0,1199	0,1186	0,1173	0,1159	0,1146	0,1133	0,1120
0,92	0,1107	0,1094	0,1080	0,1067	0,1054	0,1040	0,1027	0,1014	0,1000	0,0987
0,93	0,0974	0,0960	0,0947	0,0933	0,0920	0,0907	0,0893	0,0880	0,0866	0,0853
0,94	0,0839	0,0826	0,0812	0,0798	0,0785	0,0771	0,0758	0,0744	0,0730	0,0717
0,95	0,0703	0,0689	0,0676	0,0662	0,0648	0,0634	0,0621	0,0607	0,0593	0,0579
0,96	0,0565	0,0552	0,0538	0,0524	0,0510	0,0496	0,0482	0,0468	0,0454	0,0440
0,97	0,0426	0,0412	0,0398	0,0384	0,0370	0,0356	0,0342	0,0328	0,0314	0,0300
0,98	0,0286	0,0271	0,0257	0,0243	0,0230	0,0214	0,0201	0,0186	0,0172	0,0158
0,99	0,0140	0,0129	0,0115	0,0101	0,0086	0,0072	0,0058	0,0043	0,0029	0,0014

## Verzeichnis der Figuren

- Fig. 1. Temperaturwerte  $T_{av}$  des 9. April für Basel während der Periode 1901 – 1963. Häufigkeitsverteilung, Mittelwert und Standardabweichung der Temperaturwerte vom 9. April.
- Fig. 2. Ungeglätteter und geglätteter klimatologischer Jahresgang des Mittelwertes und der Standardabweichung von  $T_{av}$ .
- Fig. 3. Klimatologischer Jahresgang der Mittelwerte und der Standardabweichungen der Termin-, Extrem- und Vierermittelwerte.
- Fig. 4. Prüfung der Normalität der Temperaturverteilung mit Hilfe des Kolmogorov-Tests.
- Fig. 5. Verteilungsfunktion der normierten Temperatur für Basel während der Periode 1901 – 1960.
- Fig. 6. Schematische Darstellung des Markov-Modells.
- Fig. 7. Gütefunktionen  $G_K$  (Klimatologische Prognose) und  $G_P$  (persistente Prognose) in Abhängigkeit vom Persistenzkoeffizienten für eine Zeiteinheit  $a$  und der Anzahl der Zeiteinheiten  $k$ .
- Fig. 8. Graphische Darstellung der in Tabelle X definierten Klasseneinteilung einer normalverteilten Variablen.
- Fig. 9. Zusammenhang zwischen relativem Informationsgewinn und der Autokorrelation  $a$  bei der in Tabelle X definierten Klasseneinteilung.
- Fig. 10. Prüfung des Markov-Modells durch Ermittlung des klimatologischen Jahresverlaufs der Autokorrelationskoeffizienten von  $T_{av}$ .
- Fig. 11. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Terminwerte  $T_7$ .
- Fig. 12. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Terminwerte  $T_{13}$ .
- Fig. 13. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Terminwerte  $T_{21}$ .
- Fig. 14. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Extremwerte  $T_{min}$ .
- Fig. 15. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Extremwerte  $T_{max}$ .
- Fig. 16. Klimatologischer Jahresverlauf der Persistenzkoeffizienten  $a$  für die Termin-, Extrem- und Vierermittelwerte.

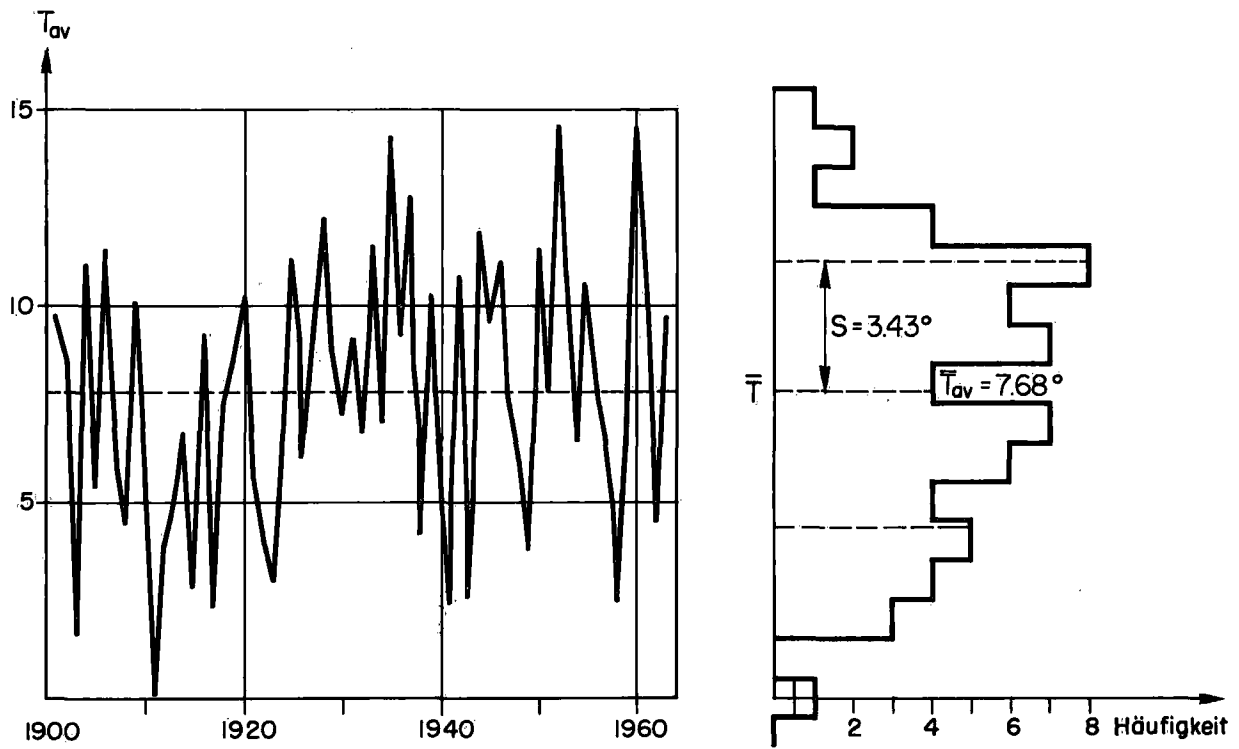


Fig. 1. Links: Temperaturwerte  $T_{av}$  des 9. April für Basel während der Periode 1901–1963  
Rechts: Häufigkeitsverteilung, Mittelwert und Standardabweichung der Temperaturwerte vom 9. April

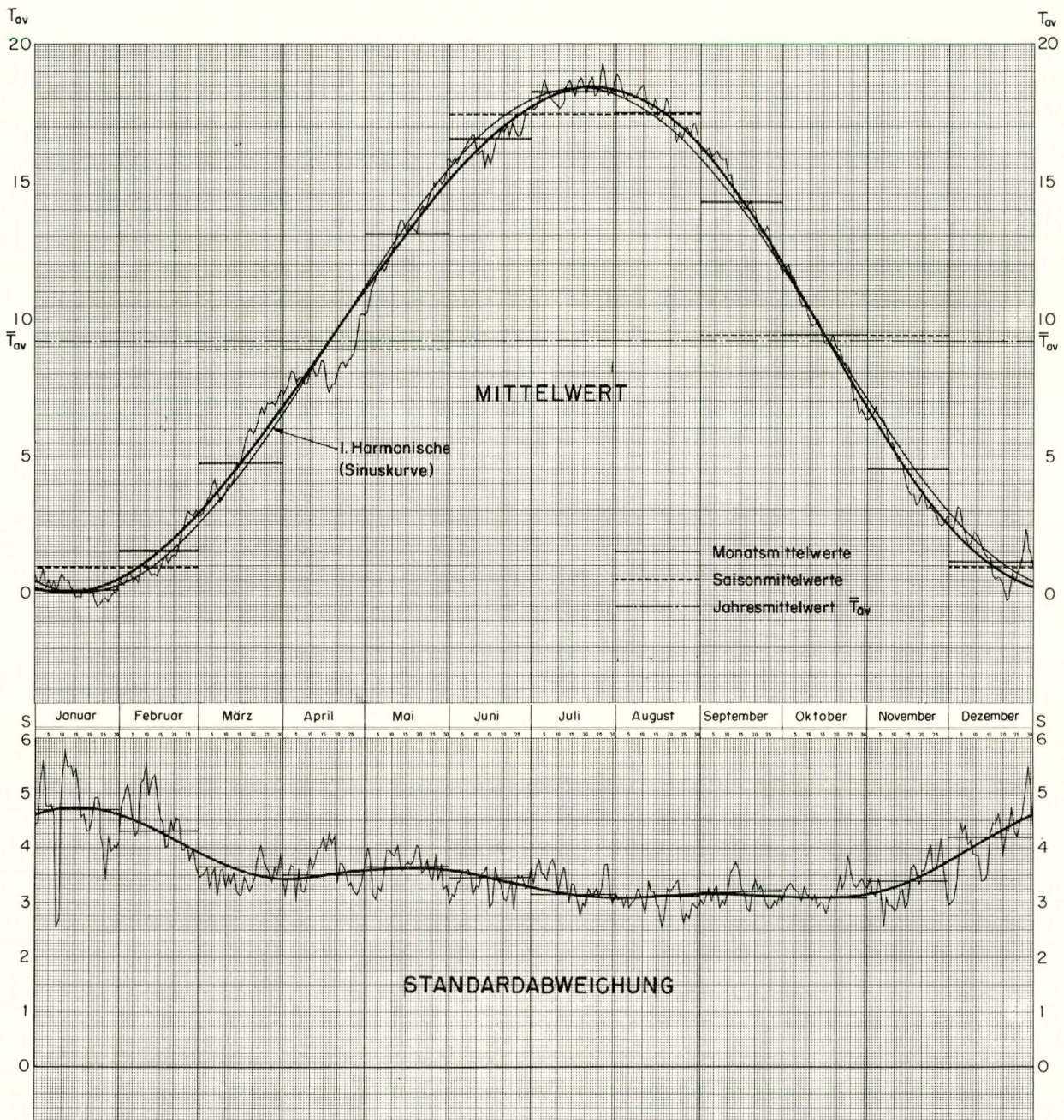


Fig. 2. Ungeglätteter und geglätteter klimatologischer Jahresgang des Mittelwertes und der Standardabweichung von  $T_{av}$ . Zur Glättung werden für den Mittelwert zwei, für die Standardabweichung drei Harmonische herangezogen. Die zur Glättung benötigten Fourierkoeffizienten sind in Tabelle XVI aufgeführt. Zur Illustration des Einflusses der zweiten Harmonischen ist noch der reine Sinusverlauf als erste Näherung für den Mittelwert eingezeichnet.

Erstaunlich sind der stark verzögerte Temperaturanstieg im April wie auch – in geringerem Masse – die aus dem Rahmen fallenden Temperaturspitzen Mitte Juni (negativ) und Ende Dezember (positiv). Es ist klar, dass derartige Spitzen nur mit höheren Harmonischen erfasst werden können.

Die eingetragenen Monats- und Saisonmittel wurden nach der in Kap. II. 4. beschriebenen Methode berechnet (siehe auch Tabelle VI).

# Klimatologischer Jahresverlauf der Temperatur Basel 1901-1960

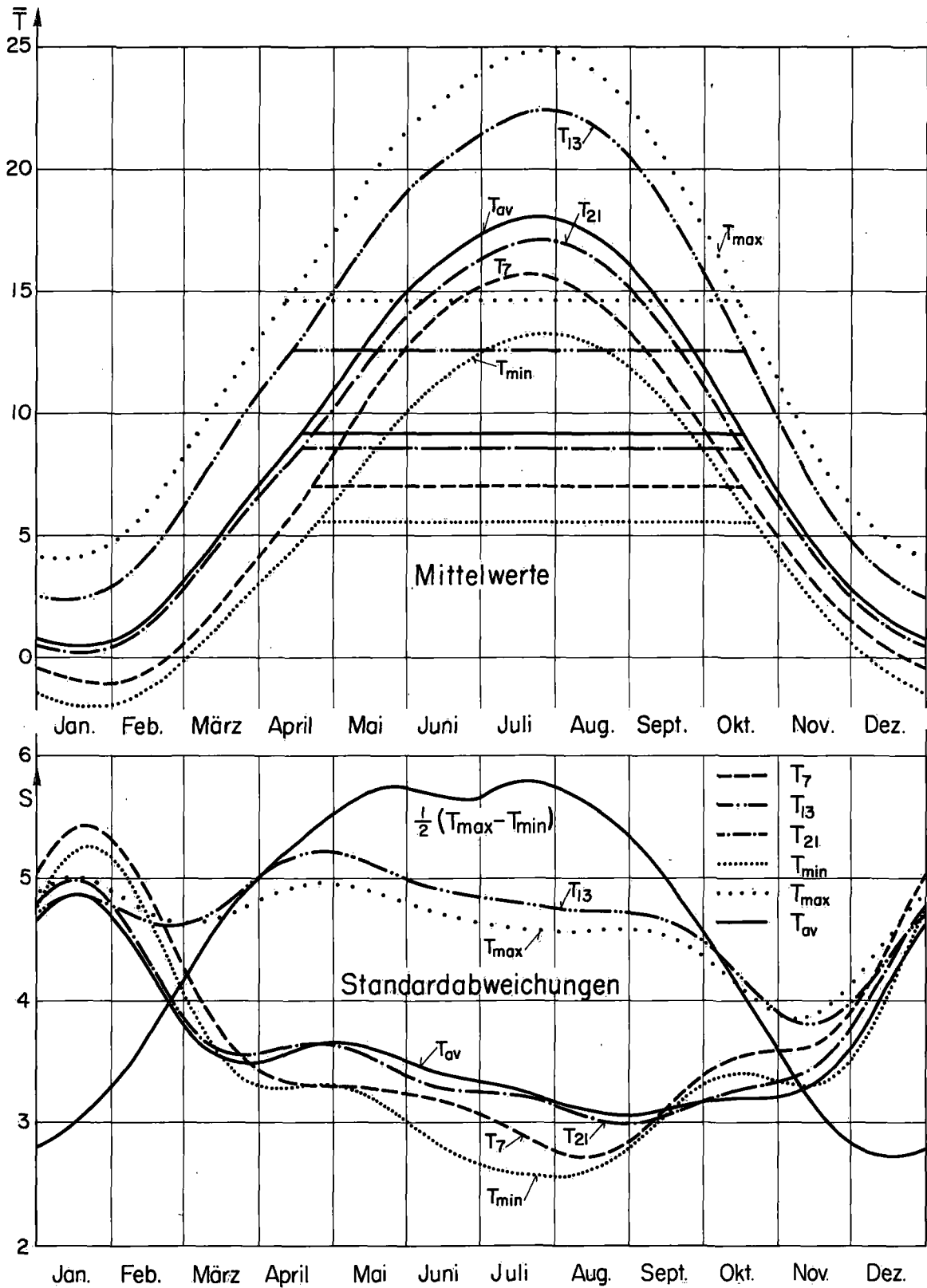


Fig. 3. Klimatologischer Jahresgang der Mittelwerte (oben) und der Standardabweichungen (unten) der Termin-, Extrem- und Vierermittelwerte.

In der oberen Graphik sind noch die Jahresmittel eingezeichnet, während unten als zusätzliches Streuungsmaß die mittlere Differenz der Extremwerte eingetragen ist.



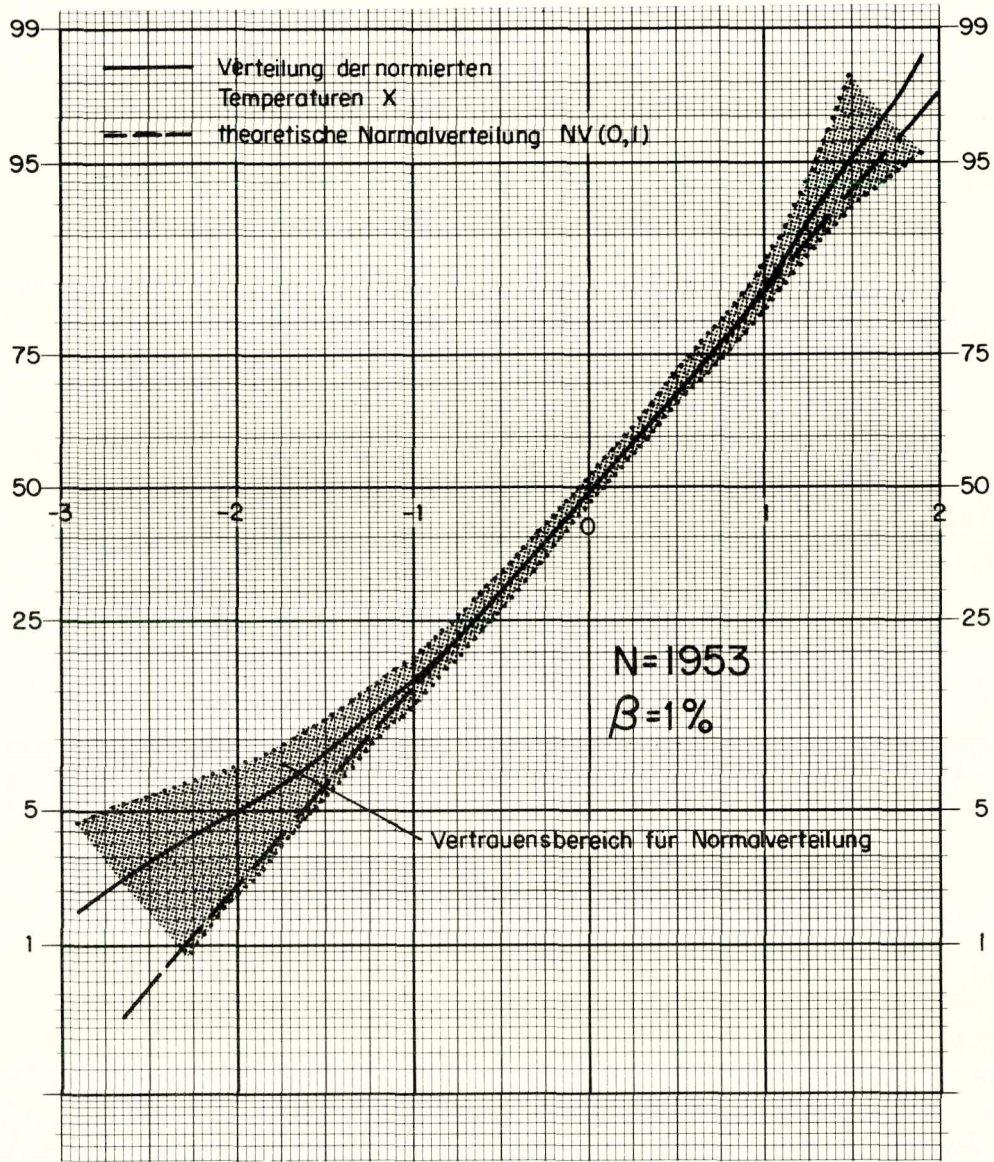


Fig. 4. Prüfung der Normalität der Temperaturverteilung mit Hilfe des Kolmogorov-Tests.

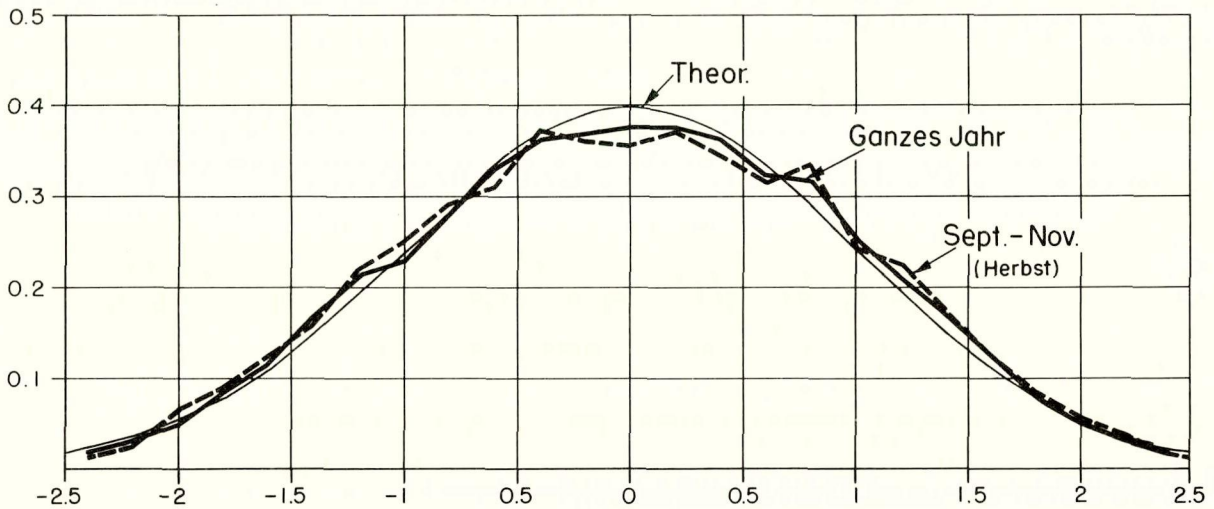


Fig. 5. Verteilungsfunktion der normierten Temperatur für Basel während der Periode 1901–1960:

- ausgezogen: ganzes Jahr
- gestrichelt: September–November (Herbst)
- dünne Kurve: theoretische Normalverteilung

$$X_i = \underbrace{a \cdot X_{i-1}}_{\text{persistenter Anteil}} + \underbrace{\sqrt{(1-a^2)} \cdot \varepsilon_t}_{\text{stochastischer Anteil}}$$

$\rho_0 = 1$
$\rho_k = a^k$

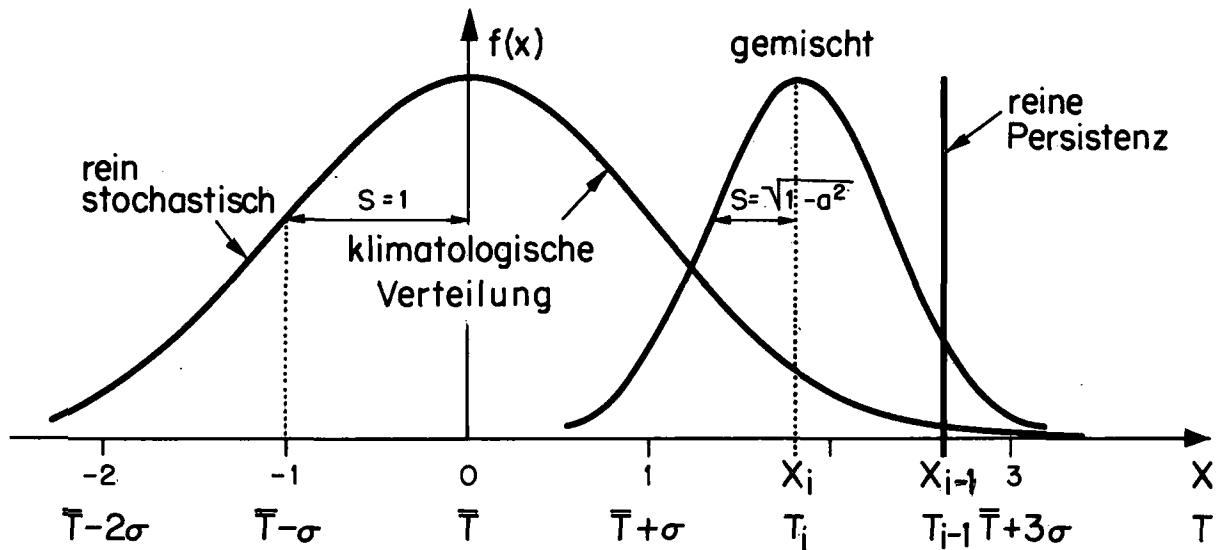


Fig. 6. Schematische Darstellung des Markov-Modells.

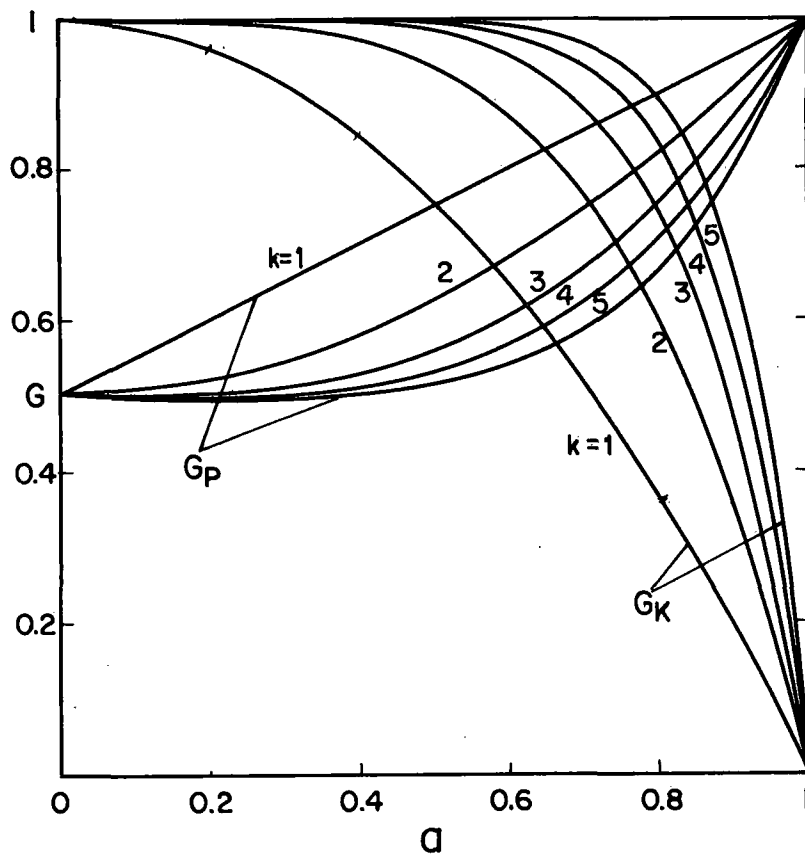


Fig. 7. Gütefunktionen  $G_K$  (Klimatologische Prognose) und  $G_P$  (persistente Prognose) in Abhängigkeit vom Persistenzkoeffizienten für eine Zeiteinheit  $a$  und der Anzahl der Zeiteinheiten  $k$ .

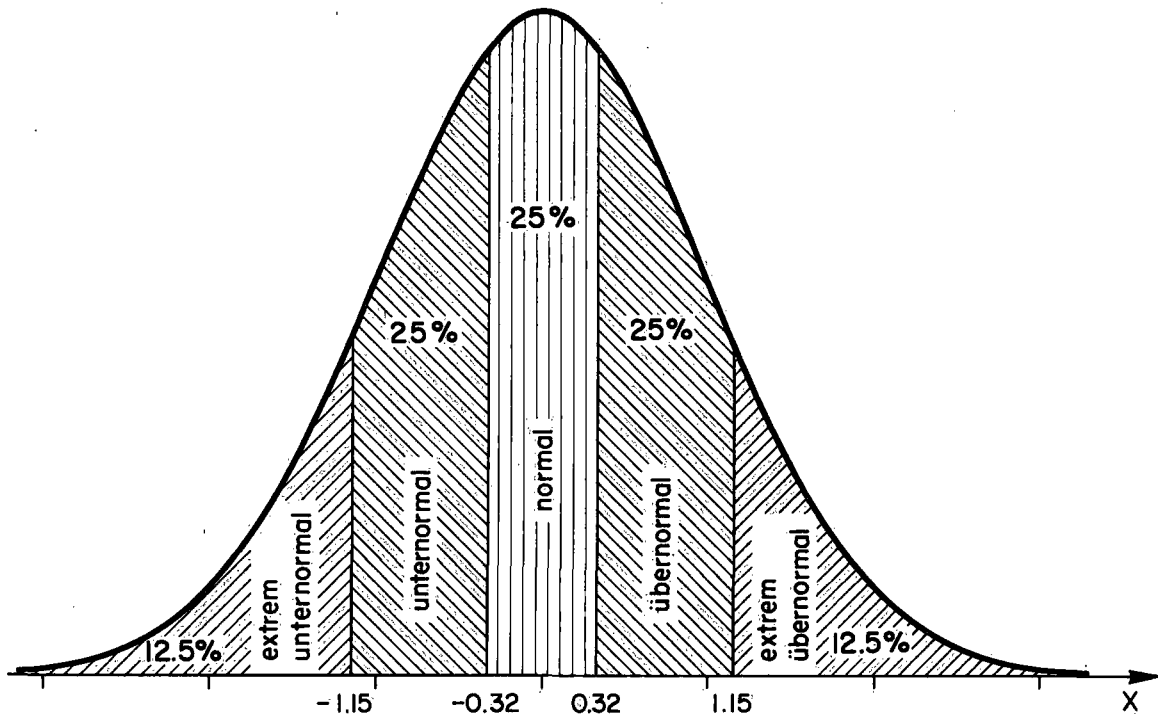


Fig. 8. Graphische Darstellung der in Tabelle X definierten Klasseneinteilung einer normalverteilten Variablen.

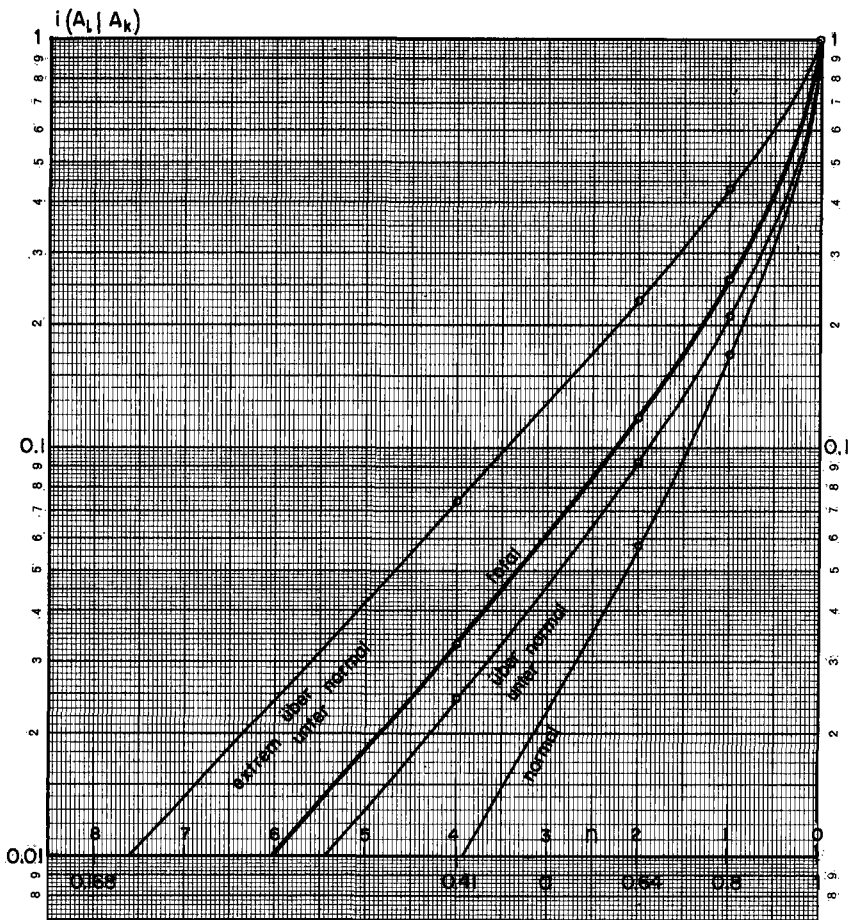


Fig. 9. Zusammenhang zwischen relativem Informationsgewinn und der Autokorrelation  $a$  bei der in Tabelle X definierten Klasseneinteilung.

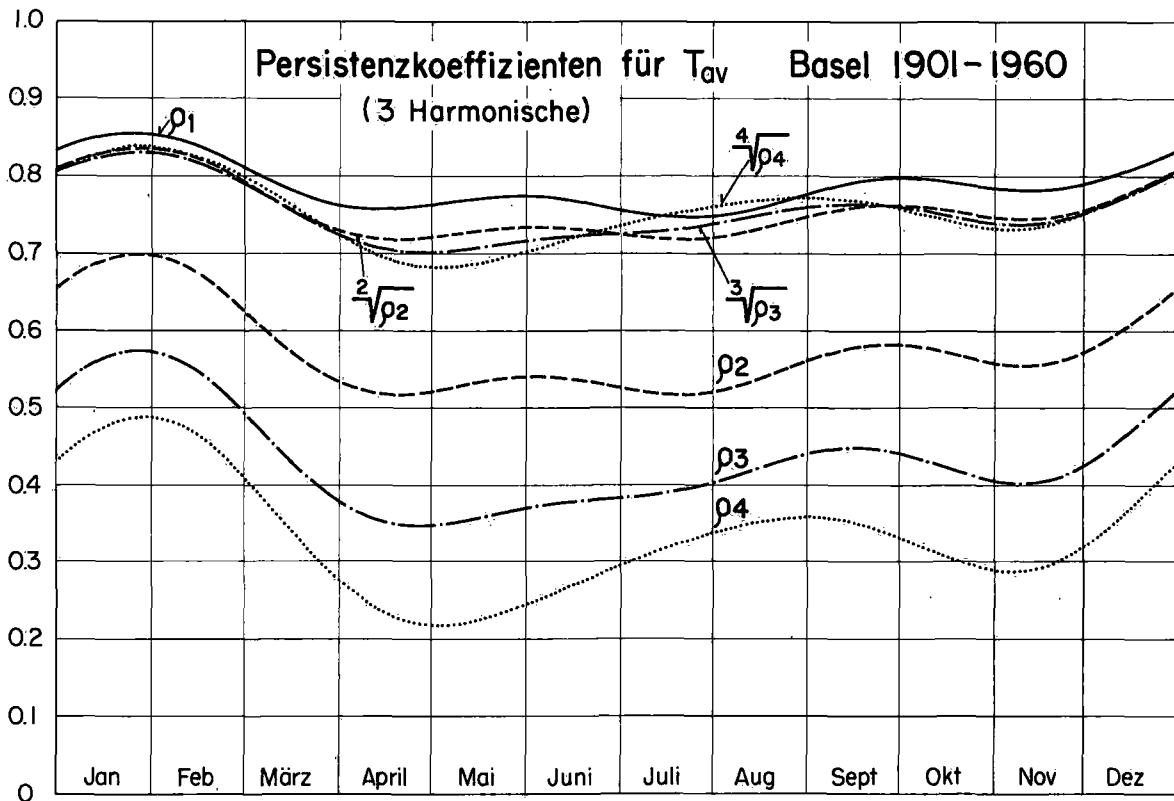


Fig. 10. Prüfung des Markov-Modells durch Ermittlung des klimatologischen Jahresverlaufs der Autokorrelationskoeffizienten  $\rho_1, \rho_2, \rho_3, \rho_4$  der Vierermittelwerte und Vergleich der Kurven  $a^{(k)} = k\sqrt{\rho_k}$

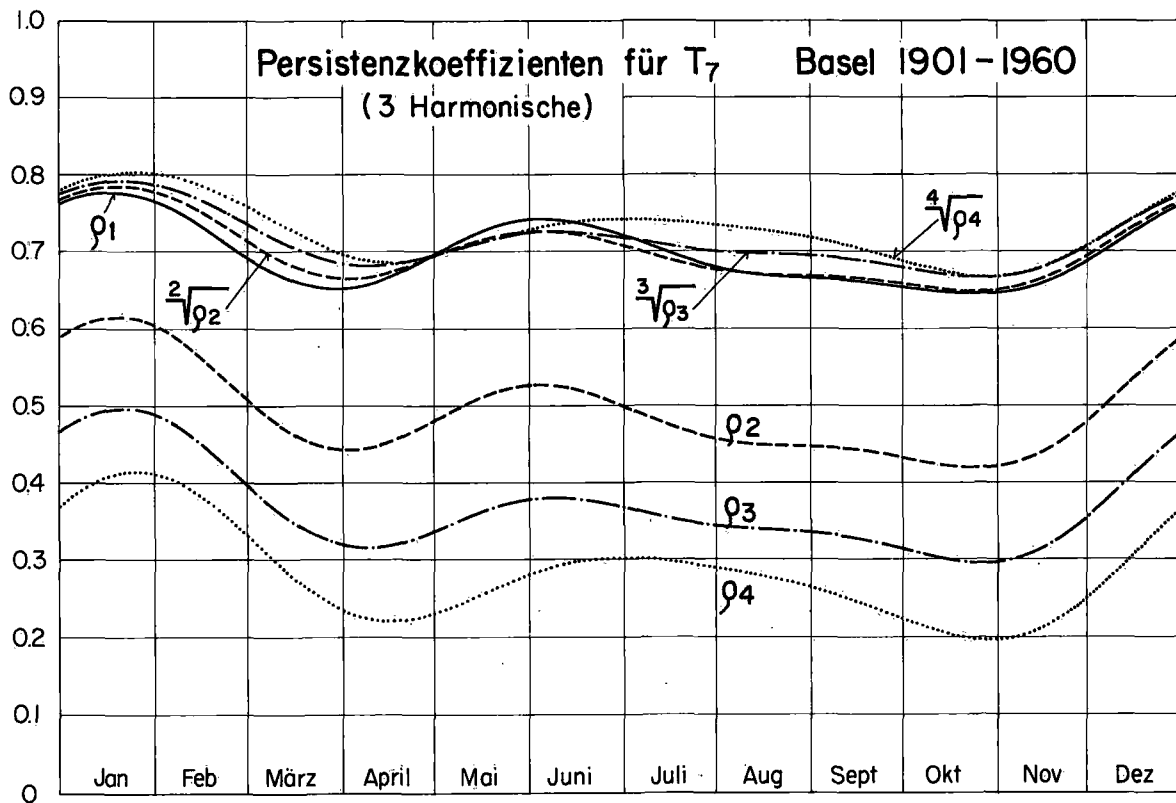


Fig. 11. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Terminwerte  $T_7$

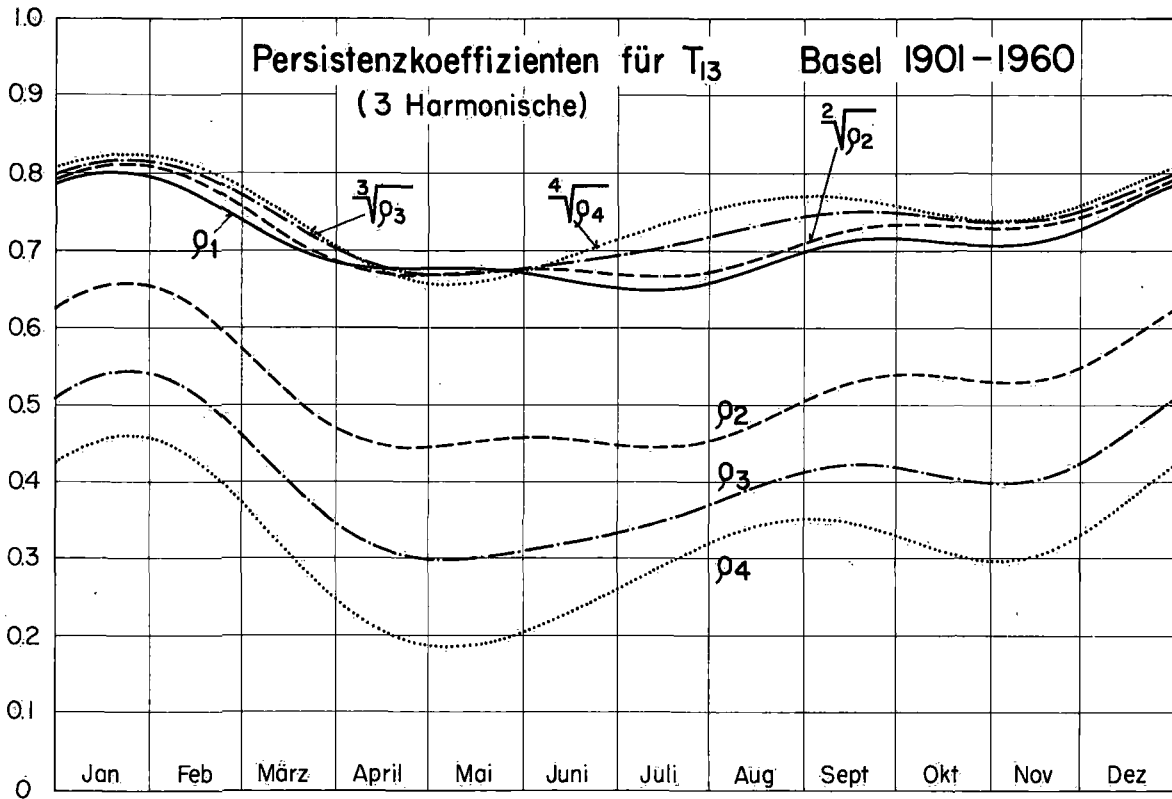


Fig. 12. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Terminwerte  $T_{13}$

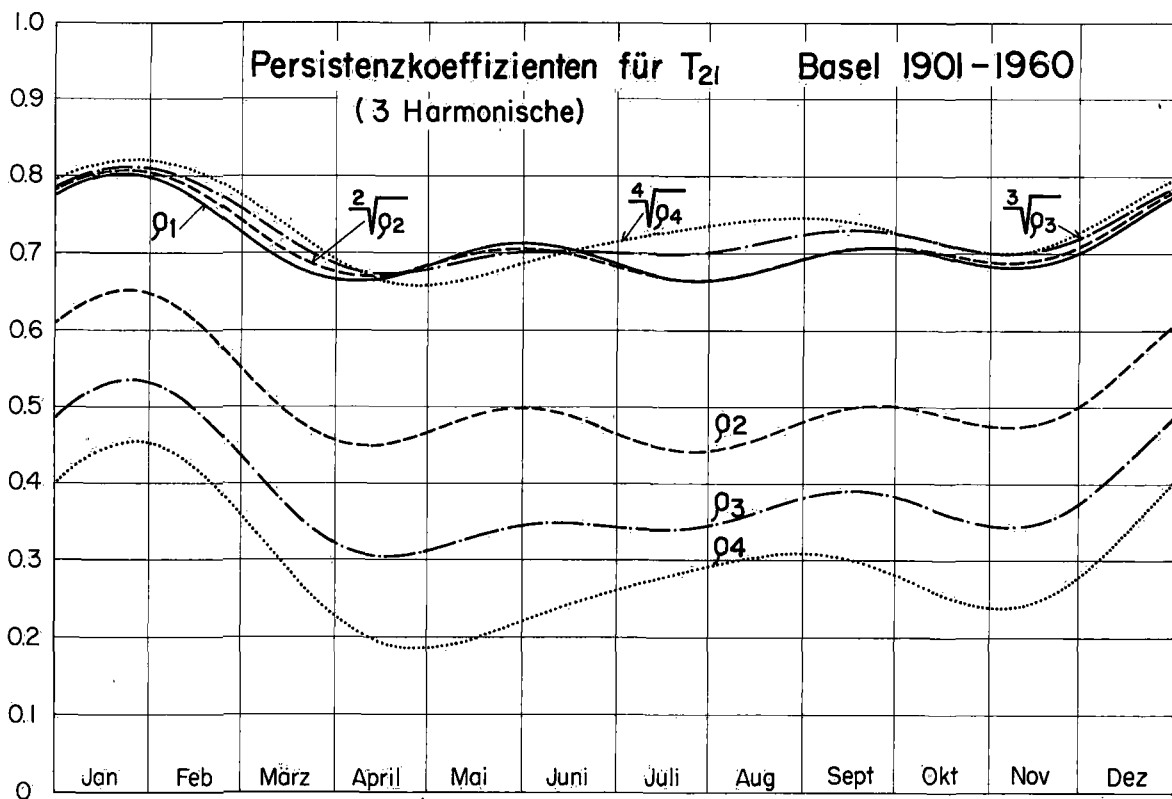


Fig. 13. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Terminwerte  $T_{21}$

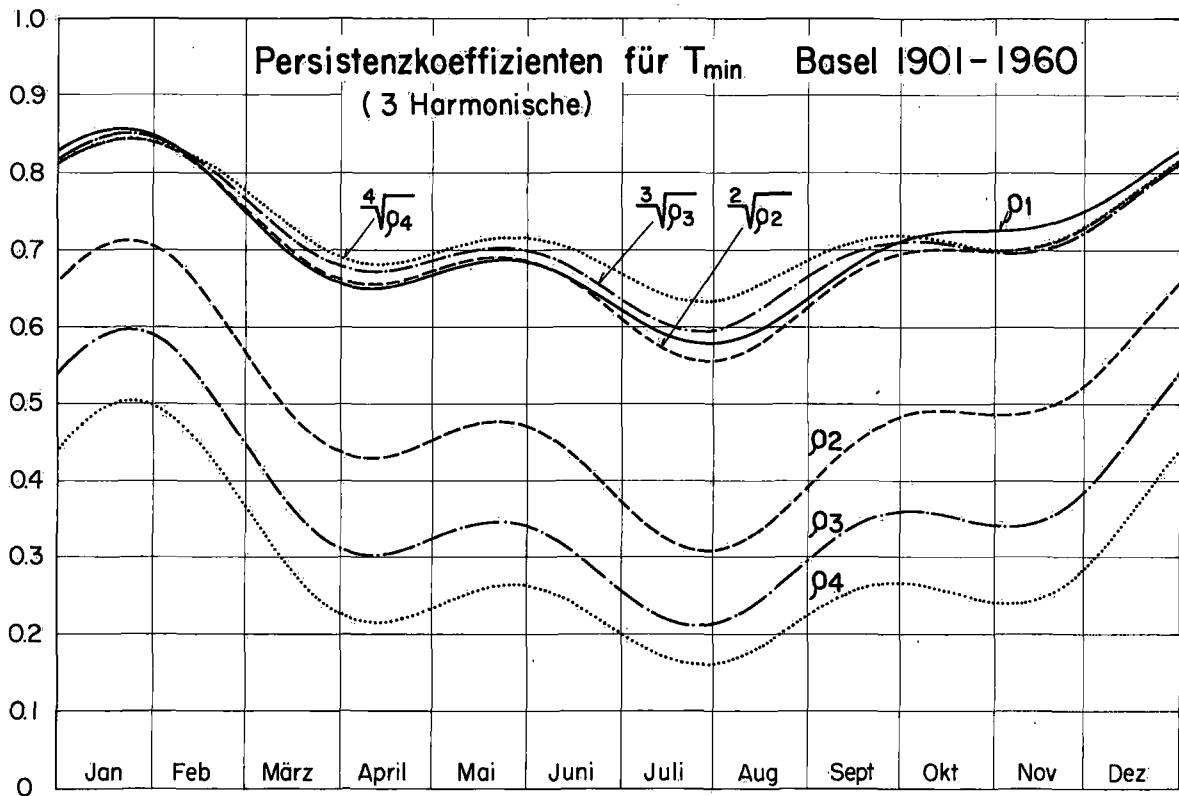


Fig. 14. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Extremwerte  $T_{\min}$

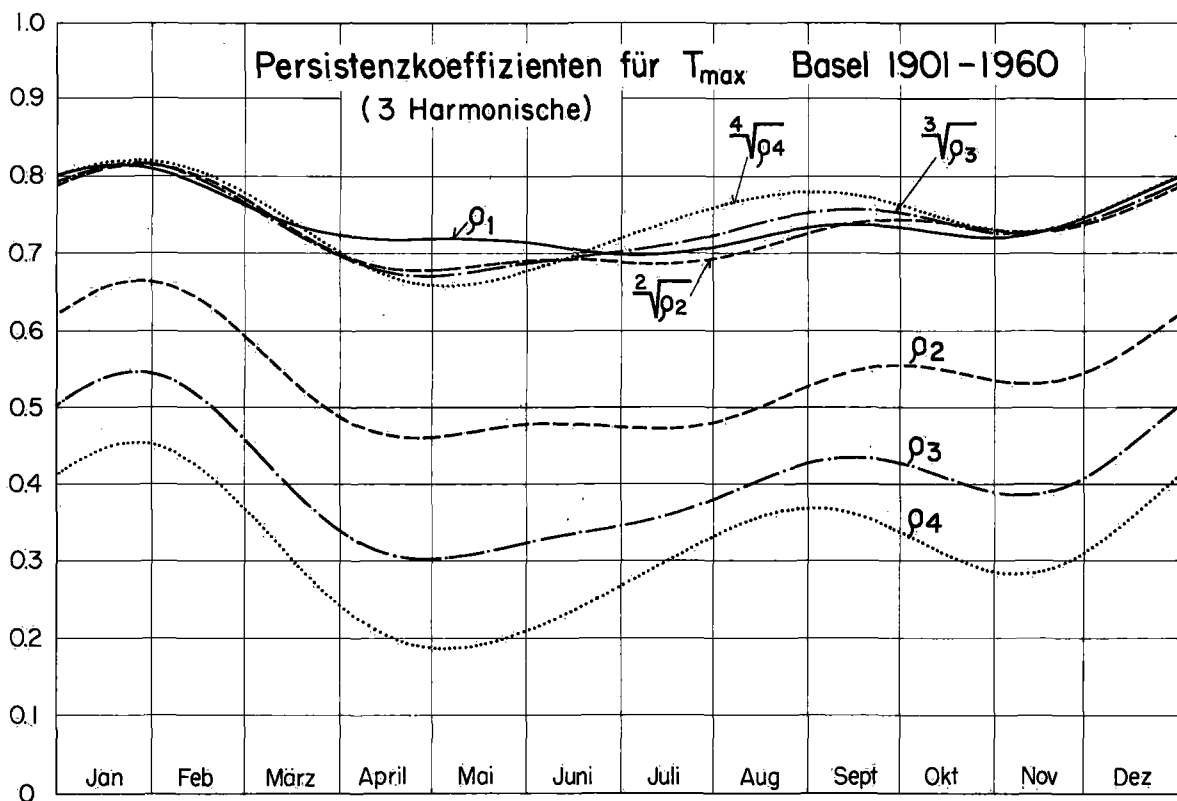


Fig. 15. Analoge Prüfung des Markov-Modells wie in Fig. 10: für Extremwerte  $T_{\max}$

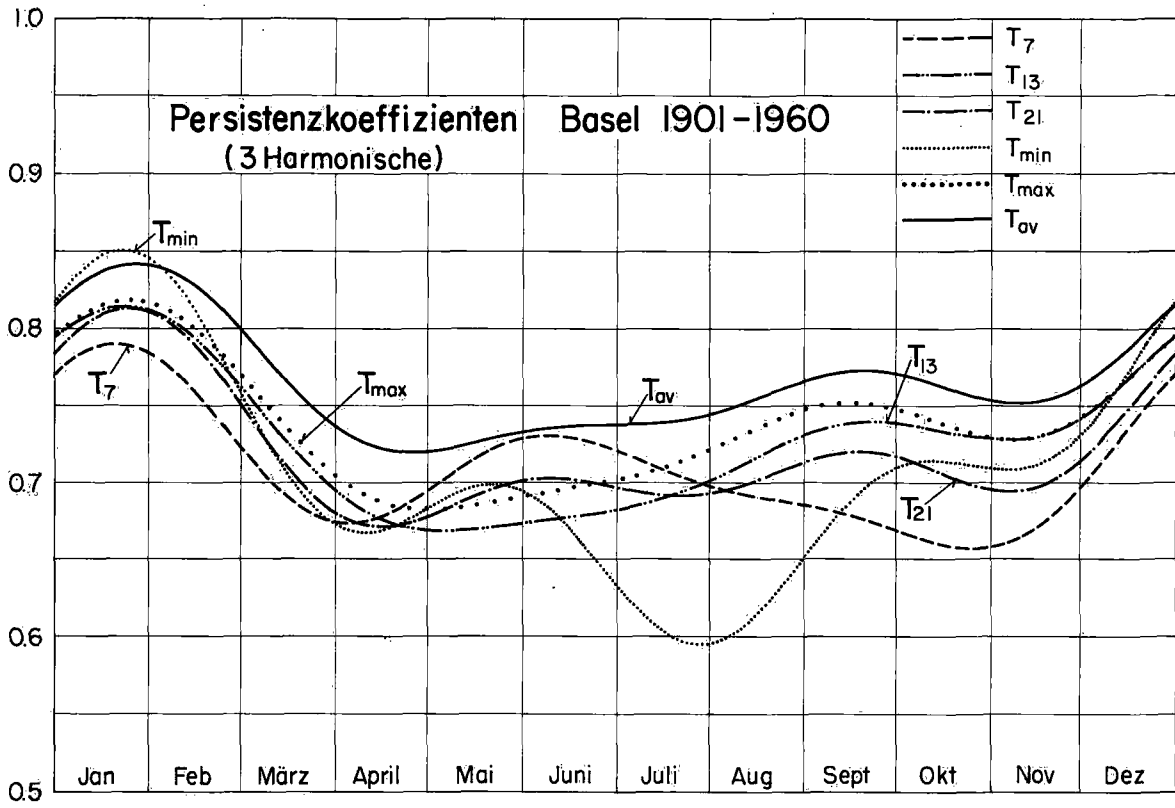


Fig. 16. Klimatologischer Jahresverlauf der Persistenzkoeffizienten  $a$  für die Termin-, Extrem- und Vierermittelwerte.

- Nr. 14 Joss J., Schram K., Thams J.C., Waldvogel A., Untersuchungen zur quantitativen Bestimmung von Niederschlagsmengen mittels Radar.  
37 Seiten, 1969 Fr. 7.—
- Nr. 15 Courvoisier H.W., Die quantitative Niederschlagsprognose winterlicher zyklonaler Witterungslagen auf der Alpennordseite der Schweiz.  
15 Seiten, 1970 Fr. 4.—
- Nr. 16 Schram Karin und Thams J.C., Die kurzweilige Globalstrahlung und die diffuse Himmelsstrahlung auf dem Flugplatz Zürich-Kloten.  
18 Seiten, 1970 Fr. 4.—
- Nr. 17 Kasser P., Schram Karin und Thams J.C., Die Strahlungsverhältnisse im Gebiet der Baye de Montreux.  
46 Seiten, 1970 Fr. 7.—
- Nr. 18 Gutermann Th., Vergleichende Untersuchungen zur Föhnhäufigkeit im Rheintal zwischen Chur und Bodensee.  
68 Seiten, 1970 Fr. 9.—
- Nr. 19 Ginsburg Theo., Die statistische Auswertung von langjährigen Temperaturreihen.  
42 Seiten, 1970 Fr. 7.—



