

Teubner-Reihe UMWELT

D. Stoyan/H. Stoyan/U. Jansen  
Umweltstatistik

# Teubner-Reihe UMWELT

Herausgegeben von

Prof. Dr. Dr. Müfit Bahadır, Braunschweig

Prof. Dr. Hans-Jürgen Collins, Braunschweig

Prof. Dr. Bertold Hock, Freising

Diese Buchreihe ist ein Forum für Veröffentlichungen zum gesamten Themenbereich Umwelt. Es erscheinen einführende Lehrbücher, Monographien und Forschungsberichte, die den aktuellen Stand der Wissenschaft wiedergeben.

Das inhaltliche Spektrum reicht von den naturwissenschaftlich-technischen Grundlagen über umwelttechnische Fragestellungen bis hin zu juristisch, sozial- und gesellschaftswissenschaftlich ausgerichteten Titeln. Besonderer Wert wird dabei auf eine allgemeinverständliche, dennoch exakte und präzise Darstellung gelegt. Jeder Band ist in sich abgeschlossen.

Die Autoren der Reihe wenden sich vorwiegend an Studierende, Lehrende sowie in der Praxis tätige Fachleute.

# Umweltstatistik

Statistische Verarbeitung und Analyse  
von Umweltdaten

Von Prof. Dr. Dietrich Stoyan  
Helga Stoyan  
Dr. Uwe Jansen  
TU Bergakademie Freiberg



B. G. Teubner Verlagsgesellschaft  
Stuttgart · Leipzig 1997

**Prof. Dr. Dietrich Stoyan**

Geboren 1940 in Berlin. Von 1959 bis 1964 Mathematikstudium an der TU Dresden. 1967 Promotion an der Bergakademie Freiberg. Rektor der TU Bergakademie Freiberg von 1991 bis 1997. Mitglied der Academia Europaea. Fellow of IMS.

**Dipl.-Math. Helga Stoyan**

Geboren 1941 in Liegau-Augustusbad. Von 1959 bis 1964 Mathematikstudium an der TU Dresden.

**Dr. Uwe Jansen**

Geboren 1949 in Bordesholm. Von 1967 bis 1971 Mathematikstudium an der Bergakademie Freiberg, dort 1975 Promotion zum Dr. rer. nat.

Gedruckt auf chlorfrei gebleichtem Papier.

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

**Stoyan, Dietrich:**

Umweltstatistik : statistische Verarbeitung und Analyse von  
Umweltdaten / von Dietrich Stoyan ; Helga Stoyan ; Uwe Jansen. –  
Stuttgart ; Leipzig : Teubner, 1997  
(Teubner-Reihe Umwelt)

ISBN 978-3-8154-3526-7

ISBN 978-3-322-99480-6 (eBook)

DOI 10.1007/978-3-322-99480-6

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt besonders für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

© B. G. Teubner Verlagsgesellschaft Leipzig 1997

Umschlaggestaltung: E. Kretschmer, Leipzig

# Vorwort

Die Statistik ist ein unentbehrliches Hilfsmittel der Umweltforschung. Sie wird benutzt, um die massenhaft anfallenden, oft stark schwankenden, heterogenen und problembeladenen Daten aus der Umwelt zu analysieren und zu interpretieren. Viele wichtige Aussagen über die Umwelt wurden und werden auf statistischem Wege erhalten. Das sind zum Beispiel diejenigen über eventuelle Klimaveränderungen, aber die Statistik ermöglicht auch zahlreiche weniger spektakuläre (und weniger umstrittene) Erkenntnisse. Das gilt besonders für Situationen, in denen das Verhalten zahlreicher Menschen eine Rolle spielt oder wo den eigentlichen Zusammenhängen erhebliche „zufällige“ Schwankungen überlagert sind, die auf vielen verschiedenen und unübersichtlichen Einflüssen beruhen. Auch im Rahmen der Überwachung der Umwelt (des „Umweltmonitoring“) ist die Statistik ein wichtiges Element. Schließlich sind statistische Methoden für die Planung und Auswertung von Experimenten in der Umweltforschung erforderlich.

Natürlich gibt es keine eigentliche wissenschaftliche Disziplin „Umweltstatistik“. Tatsache ist aber, dass für die Untersuchung der Umwelt bestimmte statistische Verfahren besonders wichtig sind, und zwar wegen der dabei auftretenden speziellen Fragestellungen. Dabei müssen insbesondere räumliche und zeitliche Aspekte berücksichtigt werden, oder es sind die Einflüsse verschiedener, gleichzeitig wirkender Größen simultan zu untersuchen. Verschiedene Verfahren, die das leisten, werden in diesem Buch behandelt, das natürlich auch den Titel „Ausgewählte Kapitel der Statistik, dargestellt an Beispielen aus der Umweltforschung“ hätte haben können. Insbesondere geht es in diesem Buch um

multivariate Statistik,  
Zeitreihenanalyse,  
Geostatistik und  
zufällige Punktfolgen.

Zusätzlich werden eine Reihe anderer statistischer Methoden kurz erläutert, zum Beispiel Probennahmeverfahren und Extremwertstatistik.

Das vorliegende Buch beruht auf Lehrveranstaltungen der Autoren für Ingenieure und Naturwissenschaftler an der TU Bergakademie Freiberg. Es setzt lediglich den Stoff einer Einführungsvorlesung Statistik voraus, zu der Themen wie beschreibende Statistik, Grundlagen der Wahrscheinlichkeitsrechnung und Tests und Konfidenzintervalle gehören sollten. (Das Buch Stoyan, 1993, dessen Symbolik hier benutzt wird, enthält diesen Stoff.) Im übrigen sind gewisse mathematische Grundkenntnisse aus einer Vorlesung zur höheren Mathematik erforderlich. Hauptanliegen des Buches ist es, einen Einstieg in die Beherrschung der statistischen Verfahren zu ermöglichen, ihre Anwendungen zu demonstrieren und ihre Möglichkeiten aufzuzeigen. Wegen mathematischer Einzelheiten und Begründungen wird jeweils auf die Fachliteratur hingewiesen.

Der Leser wird erfahren, dass die Statistik ein Werkzeug ist, mit dem man in großen, unübersichtlichen Datenmengen Zusammenhänge sichtbar machen kann, natürlich nur, sofern solche überhaupt vorhanden sind. Er wird lernen, dass die Ergebnisse durchaus davon abhängen, welche Methode man benutzt; insoweit ist angewandte Statistik eine Kunst, die man durch Beispiele und Vorbilder erlernen kann. Für die Wahl bestimmter Verfahrensparameter, die die Ergebnisse wesentlich beeinflussen können, gibt es mitunter keine festen Regeln, sondern lediglich Rezepte; manchmal wird dem Statistik-Anwender sogar empfohlen, selbst durch Probieren „vernünftige“ Werte zu suchen. Da hierzu oft umfangreiche Rechnungen erforderlich sind, ist der Personalcomputer für statistische Analysen ein unschätzbares Hilfsmittel.

Die moderne statistische Software spielt in diesem Buch eine wichtige Rolle. Es wird erwartet, dass der Leser Zugriff zu einem Statistikprogrammpaket hat oder demnächst haben wird. Wünschenswert sind eigene Experimente mit den Daten dieses Buches oder eigenen, ähnlichen Datensätzen. Die in dem Buch analysierten Datensätze sind auf dem www-Server der Fakultät für Mathematik und Informatik der TU Bergakademie Freiberg abgelegt und über die Adresse

<http://www.mathe.tu-freiberg.de/Stoyan/umwdat.html>

abrufbar. Die Verfasser haben sich bewusst nicht auf ein einziges Statistikprogrammpaket beschränkt. Vielmehr wendeten sie ganz verschiedene solcher Programme an. (Das Fehlen des Zeichens ® an den kommerziellen Namen soll übrigens nicht zeigen, dass diese Namen frei verwendet werden dürfen.) Bei der Auswahl der Programme spielten der Zufall und die Situation an der TU Bergakademie Freiberg eine gewisse Rolle. Die meisten Beispiele wurden den Verfassern nämlich von Wissenschaftlern dieser Universität überlassen, die sich systematisch, aus der Sicht der geschlossenen Stoffkreisläufe, den Umweltproblemen widmet.

Fragen der amtlichen Umweltstatistik kommen in diesem Buch nur am Rande vor, nämlich dann, wenn von ihr gelieferte Daten statistisch analysiert

werden. Auch die verschiedenen, vorhandenen und im Entstehen begriffenen Informationssysteme (Geographische Informationssysteme, Umweltinformationssysteme) werden nicht behandelt.

Es ist den Autoren eine angenehme Pflicht, einer Reihe von Kolleginnen und Kollegen zu danken, die sie mit Ideen, Ratschlägen, Informationen und Daten unterstützt haben. Hier seien in alphabetischer Reihenfolge und unter Weglassung der Titel genannt: H. Bandemer, A. Bellmann, S. Berndt, J. W. Einax, I. Gugel, W. Härdle, T. Hillmann, T. Jonsson, A. Kluge, B. Markert, J. Menz, B. Merkel, W. Näther, U. Neu, O. Nitzsche, R. Pohlink, W. Rasemann, R.-D. Reiß und H. von Storch. Ebenfalls sei herzlich Herrn J. Weiß vom Teubner-Verlag in Leipzig für die sehr angenehme Zusammenarbeit gedankt.

Freiberg, 31. Mai 1997

Die Autoren

Die in diesem Buch verwendeten Daten stammen aus der Umweltforschung verschiedener Institute der TU Bergakademie Freiberg, des Geographischen Instituts der Universität Bern sowie vom Isländischen Meteorologischen Büro. Die Autoren danken für die Möglichkeit ihrer statistischen Analyse und für fachliche Beratung.

## **Symbole für Mittelwert und Streuung**

**E** Erwartungswert, Mittelwert

**var** Varianz, Streuung



# Inhaltsverzeichnis

<b>1</b>	<b>Umweltdaten – Visualisierung – Monitoring</b>	<b>13</b>
1.1	Besonderheiten von Umweltdaten und Umweltstatistik . . . . .	13
1.2	Strukturierungsprobleme . . . . .	17
1.2.1	Allgemeines . . . . .	17
1.2.2	Grundstrukturen . . . . .	17
1.2.3	Verarbeitungsprobleme . . . . .	19
1.3	Dateneigenschaften . . . . .	20
1.3.1	Darstellungs- und Verarbeitungseigenschaften . . . . .	20
1.3.2	Messtechnische Konsequenzen . . . . .	24
1.3.3	Zuverlässigkeit . . . . .	24
1.3.4	Relevanzfragen . . . . .	25
1.3.5	Datenschutzprobleme . . . . .	26
1.3.6	Zeitprobleme . . . . .	27
1.4	Graphiken und Visualisierungstechniken . . . . .	27
1.4.1	Einleitung . . . . .	27
1.4.2	Explorative Zielstellung . . . . .	28
1.4.3	Argumentative Zielstellung . . . . .	30
1.4.4	Graphiken und Programme . . . . .	32
1.4.5	Summarische Statistiken . . . . .	33
1.4.6	Kastendiagramme . . . . .	38
1.4.7	Häufigkeitsdiagramme . . . . .	39
1.4.8	Ähnlichkeiten von Datensätzen . . . . .	43
1.4.9	Darstellung von Abhängigkeiten . . . . .	46
1.5	Umweltstatistik und Umweltüberwachung . . . . .	49
1.6	Einige Ratschläge für Anfänger bei der Benutzung von Statistik- programm Paketen . . . . .	52

<b>2</b>	<b>Multivariate Statistik</b>	<b>53</b>
2.1	Einleitung . . . . .	53
2.2	Vorbereitungen für die multivariate Statistik . . . . .	55
2.2.1	Vorbemerkungen . . . . .	55
2.2.2	Vereinheitlichung qualitativ verschiedener Messwerte . . . . .	55
2.2.3	Die $Z$ -Transformation . . . . .	56
2.2.4	Logarithmus-Transformation . . . . .	57
2.3	Korrelationsanalyse . . . . .	61
2.3.1	Der Korrelationskoeffizient . . . . .	61
2.3.2	Rangkorrelationskoeffizienten . . . . .	63
2.3.3	Die Korrelationsmatrix . . . . .	65
2.3.4	Multipler und partieller Korrelationskoeffizient . . . . .	71
2.3.5	Kanonische Korrelationsanalyse . . . . .	75
2.3.6	Literatur zur Korrelationsanalyse . . . . .	76
2.4	Regressionsanalyse . . . . .	76
2.4.1	Einleitung . . . . .	76
2.4.2	Ausgleichsgeraden . . . . .	76
2.4.3	Nicht parametrische Regression . . . . .	86
2.4.4	Multiple lineare Regression . . . . .	89
2.4.5	Nicht lineare Regression . . . . .	99
2.4.6	Literatur zur Regressionsanalyse . . . . .	104
2.5	Hauptkomponenten- und Faktorenanalyse . . . . .	106
2.5.1	Einleitung . . . . .	106
2.5.2	Die Hauptkomponentenanalyse . . . . .	107
2.5.3	Interpretation der Ergebnisse der Hauptkomponenten- analyse . . . . .	108
2.5.4	Faktorwerte bei der Hauptkomponentenanalyse . . . . .	114
2.5.5	Die Faktorenanalyse . . . . .	115
2.5.6	Literatur und Programme zur Hauptkomponenten- und Faktorenanalyse . . . . .	126
2.6	Clusteranalyse . . . . .	126
2.6.1	Einleitung . . . . .	126
2.6.2	Proximitäten und Fusionsalgorithmen . . . . .	127
2.7	Weitere multivariate Verfahren . . . . .	136
2.7.1	Varianzanalyse . . . . .	136
2.7.2	Klassifizierungsverfahren . . . . .	140
<b>3</b>	<b>Zeitreihenanalyse</b>	<b>147</b>
3.1	Einleitung . . . . .	147
3.2	Glättung von Zeitreihen . . . . .	154

3.2.1	Drei Glättungsmethoden . . . . .	154
3.2.2	Trendschätzungen . . . . .	158
3.2.3	Laufende Glättung und kurzfristige Vorhersage . . . . .	168
3.3	Grundbegriffe aus der Theorie der stochastischen Prozesse . . . . .	171
3.3.1	Stochastische Prozesse und Zeitreihen . . . . .	171
3.3.2	Mittelwerts- und Kovarianzfunktion . . . . .	172
3.3.3	Formeln, Beziehungen und Eigenschaften für Kovarianz- und Korrelationsfunktionen . . . . .	181
3.4	Einige Modelle für stationäre Zeitreihen . . . . .	185
3.4.1	Einleitung . . . . .	185
3.4.2	MA( $q$ )-Prozess oder Gleitmittelprozess . . . . .	187
3.4.3	AR( $p$ )-Prozess oder autoregressiver Prozess . . . . .	189
3.4.4	ARMA( $p, q$ )-Prozess . . . . .	191
3.5	Literatur und Programme zur Zeitreihenanalyse . . . . .	192
<b>4</b>	<b>Geostatistik</b>	<b>193</b>
4.1	Einleitung . . . . .	193
4.2	Homogene und isotrope Zufallsfelder . . . . .	195
4.3	Variogramme . . . . .	200
4.4	Theoretische Variogramme . . . . .	209
4.5	Räumliche Interpolation und Kriging . . . . .	213
4.5.1	Kriging des Mittelwertes . . . . .	214
4.5.2	Punktkriging . . . . .	215
4.5.3	Weitere Kriging-Verfahren . . . . .	222
4.6	Interpolation bei Inhomogenität . . . . .	223
4.7	Weitere Probleme und Anwendungen der Geostatistik . . . . .	230
4.8	Literatur und Programme zur Geostatistik . . . . .	232
<b>5</b>	<b>Folgen von Ereignissen, Punktprozesse und Punktfelder</b>	<b>233</b>
5.1	Einleitung . . . . .	233
5.2	Grundlagen der Theorie der Punktprozesse . . . . .	234
5.2.1	Anzahlverteilungen . . . . .	234
5.2.2	Punktprozesse . . . . .	238
5.2.3	Intensitätsgrößen . . . . .	242
5.2.4	Zeittransformation . . . . .	244
5.2.5	Größen zweiter Ordnung für stationäre Punktprozesse . . . . .	245
5.2.6	Abstandsverteilungen . . . . .	247
5.2.7	Markierte Punktprozesse . . . . .	254
5.3	Poisson-Prozesse . . . . .	262
5.3.1	Der stationäre Poisson-Prozess . . . . .	262

5.3.2	Instationärer Poisson-Prozess . . . . .	267
5.4	Punktfelder . . . . .	273
5.4.1	Einleitung . . . . .	273
5.4.2	Das homogene Poisson-Punktfeld . . . . .	279
5.5	Statistische Beschreibung von Wechselwirkungen in Punktfeldern	282
5.6	Weitere Anwendungen von Punktprozessen . . . . .	290
5.7	Literatur über Punktprozesse und -felder . . . . .	292
<b>6</b>	<b>Weitere statistische Methoden</b>	<b>293</b>
6.1	Probennahme . . . . .	293
6.2	Ideen der statistischen Versuchsplanung . . . . .	300
6.3	Umwelt-Indizes . . . . .	313
6.4	Extremwertstatistik . . . . .	319
6.4.1	Einleitung . . . . .	319
6.4.2	Fakten aus der Extremwertstatistik . . . . .	321
6.5	Klimamodelle und Klimawechsel . . . . .	325
6.6	Vorhersage-Intervalle . . . . .	327
	<b>Literaturverzeichnis</b>	<b>329</b>
	<b>Sachwortverzeichnis</b>	<b>345</b>