

QIXIANG TONGJI FENXI YU YUBAO FANGFA

气象统计分析与预报方法

黄嘉佑 编著

(第三版)

气象出版社
China Meteorological Press



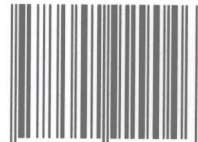
责任编辑：陶国庆

装帧设计：阳光图文工作室

气象统计分析与预报方法

QIXIANG TONGJI FENXI YU YUBAO FANGFA

ISBN 978-7-5029-2904-6



9 787502 929046

02 >



定价：39.00元

气象统计分析与预报方法

(第三版)

黄嘉佑 编著



气象出版社

内 容 提 要

本书主要介绍气象学中有关天气统计分析预报方面的基本理论及计算方法,系统地阐述了目前国内外常用的有关方法,如多元分析中的回归分析、主分量分析、因子分析、判别分析、聚类分析及时间序列分析中的自回归滑动平均模型、谱分析及马尔可夫概型分析等。本书着重讲授这些方法的基本原理、计算步骤以及它们在天气分析及动力预报中的应用。

本书经中国气象局高等学校大气科学类教材编审领导小组审查,并经教育部高等学校教学指导委员会确认,作为大学本科通用教材。此外也可作为大专院校有关专业教学参考书,对气象业务人员也有参考价值。

图书在版编目(CIP)数据

气象统计分析与预报方法/黄嘉佑编著. —3版. 北京:气象出版社,2004.3

(2007.4重印)

ISBN 978-7-5029-2904-6

I. 气… II. 黄… III. ①气象资料-统计分析②气候变化-预测 IV. P468.0

中国版本图书馆 CIP 数据核字(2004)第 009324 号

出 版 者:气象出版社

地 址:北京市海淀区中关村南大街 46 号

网 址:<http://cmp.cma.gov.cn>

邮 编:100081

E - mail:qxcbs@263.net

电 话:总编室:010-68407112

发行部:010-62175925

责任编辑:陶国庆

终 审:陈云峰

封面设计:李勤学

版式设计:王丽梅

责任校对:王丽梅

印 刷 者:北京昌平环球印刷厂

发 行 者:气象出版社

开 本:787mm×960mm 1/16

印 张:19.5

字数:392千字

版 次:2004年3月第3版 2007年4月第2次印刷

印 数:5001~10000

定 价:39.00元

本书如存在文字不清、漏印以及缺页、倒页、脱页等,请与本社发行部联系调换

再版前言

气象统计分析与预报方法,作为气象学中三大分析与预报方法之一,是高等学校大气科学类专业普遍开设的课程。本书是在作者多年来为本科生讲授的同名课程讲义的基础上编写而成的,着重介绍大气科学中统计分析与预测方法的基础理论,也涉猎一些新技术和新方法,除具有较强的理论性外,还有相当丰富的例子和应用内容。本书第一版曾经中国气象局高等学校大气科学类教材编审领导小组确认为大学本科生通用教材,并获1996年全国第三届大气科学类优秀教材一等奖。本书出版后受到读者的欢迎和好评,应广大读者的要求,加上本教材也已使用多年,为了适应学科的发展和教学上的需要,这次作者对本教材的内容进行了必要的增删,使之更臻完善。1999年10月本教材再次经教育部高等学校大气科学教学指导委员会确认,作为大学本科通用教材。

气象数据是描述和记录气象现象的性质和变化的符号,是信息的形式化的表示,它可以表现为文字或数字。在自然科学和社会科学中广泛使用气象数据(定性的或定量的、直接的或间接获得的描述、观测或实验气象数据)。目前我们处在信息时代,一切信息都可以用文字或数字型的数据来表现,甚至还包含无线电波或电视等媒体信息,它们也能用电脑多媒体技术转化为数据。因此,我们的社会中各种现象的描述都可以用数字型的数据来表现。

由此,我们就可以通过气象数据来研究大气现象,甚至研究自然、人类社会与大气的关系等。因为气候是包含五个圈层,它们的相互关系包含在大量的气象数据中,从中可以找出其演变规律性,或从描述不同对象的气象数据之间寻找它们存在的相互依赖关系。

对气象数据进行分析的方法常常是建立描述气象数据关系和演变的模

式,并达到预测的目的。本书所介绍的方法是常用的统计分析与预测方法,在统计方法中是一些基础的方法。

本书所介绍的统计分析方法是处理这些大量数据和各种信息资料的有力工具。随着计算条件的改善,许多新方法和新理论正在发展,这些众多的内容几乎无法用一本书加以概括和讲述,也不可能在一个学期内讲授完毕。因此,在重版中我们着重选择了最基本的统计量和方法做了一些补充,作为学生今后进一步学习的基础。同时为方便读者使用该书介绍的方法,还补充了一些基础方法的计算机源程序供参考使用。

编著者

2004年2月于北京大学

第一版前言

近些年来,由于电子计算机的广泛使用,使得大量资料有可能得到迅速而有效的处理,近代统计方法亦随之迅速发展,这些方法也相继引入到气象学中,使用它们可以及时处理和分析来自地面、高空、甚至由卫星、雷达等先进大气探测工具所得到的大量资料,使得气象中的天气分析和预报也得到很快的发展。例如在传统天气学基础上发展起来的天气气候学、动力气候学,以及传统动力学及数值预报基础上发展起来的统计动力预报,已成为现代气象分析与预报的新内容。事实上,大气运动不仅具有随机性的特征,也具有确定性的运动规律,只有把这两者有机地结合进行分析研究,才能使大气运动得到全面深刻的认识。因此,天气-统计以及动力-统计的结合是近年来气象分析预报发展的必然趋势。

在气象分析与预报的业务中,多元分析以及时间序列分析等统计方法使用十分广泛,其中一些基本方法已成为气象预报的基本方法。本书主要介绍涉及天气统计分析和预报方面的近代统计方法。这些方法除了在这个领域上广泛使用外,还在大气探测、大气湍流、大气污染、人工控制天气、雷达气象、卫星气象、农业气象、以及水文气象等方面都有着广泛的应用。

本书在内容上大致分如下四个方面:

第一部分介绍气象资料的整理,基本统计量的求法及其在气象中的应用(第一章)。

第二部分内容从第二章到第七章,主要介绍目前常用于气象要素场的分析和预报上的多元分析方法。

第三部分介绍时间序列分析、谱分析及马尔可夫概型分析(第八章至第十章),其中涉及如何分析气象要素随时间变化的规律性,如持续性、周期性及状态转移可能性大小等规律,以及如何利用它们作气象要素未来时刻的

预报。

第四部介绍预报模式的评价及综合预报方法(第十一章)。

本书是为大学及专科气象专业学生编写的教材,在学习本书之前最好具备有一定的概率论和数理统计学及线性代数等方面必要的知识,自修学习的读者亦可参阅有关方面的书籍作必要的准备。

本书是在北京大学地球物理系“气象统计预报试用教材”基础上编写的。由于编者实践经验不足,理论水平有限,资料和方法收集也不够齐全,错误和不当之处在所难免,希望读者批评指正。

编著者

1989年1月于北京大学

目 录

再版前言

第一版前言

第一章 气象资料的整理	(1)
§ 1.1 气象资料的表示	(1)
1.1.1 单个变量	(1)
1.1.2 多个变量	(2)
§ 1.2 基本统计量	(3)
1.2.1 平均值	(3)
1.2.2 标准差与方差	(5)
1.2.3 协方差与相关系数	(8)
1.2.4 分级相关系数与列联表	(11)
1.2.5 自协方差与自相关系数	(16)
1.2.6 落后交叉协方差与相关系数	(17)
1.2.7 峰度系数与偏度系数	(18)
§ 1.3 统计量的检验与应用	(18)
1.3.1 平均值的显著性检验	(19)
1.3.2 两组样本平均值差异的显著性检验	(20)
1.3.3 方差的显著性检验	(21)
1.3.4 相关系数的检验	(23)
1.3.5 变量的分布检验	(24)
1.3.6 气象中的应用	(26)
参考文献	(27)
第二章 回归分析	(28)
§ 2.1 一元线性回归	(28)
2.1.1 一元线性回归模型	(28)
2.1.2 回归问题的方差分析	(31)
2.1.3 相关系数与线性回归	(32)
2.1.4 回归方程的显著性检验	(33)

2.1.5	回归系数的显著性检验	(34)
2.1.6	预报值的置信区间	(36)
§ 2.2	多元线性回归	(36)
2.2.1	多元线性回归模型	(36)
2.2.2	向量 β 的最小二乘估计	(37)
2.2.3	回归系数向量 b 的统计性质	(39)
2.2.4	线性回归模型的其它形式	(40)
2.2.5	回归问题的方差分析	(43)
2.2.6	复相关系数	(44)
2.2.7	回归方程的显著性检验	(46)
2.2.8	预报值的置信区间	(47)
2.2.9	利用回归方程进行预报	(47)
§ 2.3	事件概率回归(REEP)	(50)
2.3.1	事件概率回归方程的建立	(50)
2.3.2	概率回归方程的显著性检验	(52)
2.3.3	变量为原值的事件概率回归方程	(53)
§ 2.4	因子数目	(54)
§ 2.5	逐步回归	(58)
2.5.1	回归系数的显著性检验	(58)
2.5.2	“最优”回归方程的选择	(61)
2.5.3	逐步剔除方案	(62)
2.5.4	逐步引进方案	(63)
2.5.5	双重检验的逐步回归方案	(64)
2.5.6	逐步回归例子	(70)
§ 2.6	残差分析	(73)
2.6.1	残差散布分析	(73)
2.6.2	预报残差方差分析	(73)
§ 2.7	非线性回归	(76)
2.7.1	多项式回归	(76)
2.7.2	可化为线性的曲线回归	(79)
2.7.3	一般的非线性回归模型	(79)
2.7.4	Logit 模型	(81)
§ 2.8	回归分析在气象中的应用	(83)
参考文献		(86)

第三章 判别分析	(89)
§ 3.1 费歇判别准则.....	(89)
3.1.1 费歇判别准则的概念	(89)
3.1.2 判别系数的确定	(90)
3.1.3 多因子二级判别	(92)
3.1.4 判别方程的显著性检验	(93)
§ 3.2 多级判别.....	(95)
3.2.1 判别函数离差平方和的分解	(95)
3.2.2 多级判别的费歇准则	(97)
3.2.3 判别函数的性质	(98)
3.2.4 判别函数的显著性检验	(99)
3.2.5 决策规则.....	(101)
3.2.6 多级判别计算步骤.....	(102)
§ 3.3 贝叶斯判别准则	(105)
3.3.1 贝叶斯判别准则概念.....	(105)
3.3.2 正态总体的判别.....	(106)
§ 3.4 逐步判别	(107)
3.4.1 单个因子判别能力的显著性检验.....	(108)
3.4.2 统计量 Λ 的计算方法	(110)
3.4.3 逐步判别计算步骤.....	(111)
3.4.4 非线性逐步判别.....	(115)
3.4.5 回归逐步判别.....	(116)
§ 3.5 判别分析在气象中的应用	(118)
参考文献.....	(120)
第四章 主分量分析	(121)
§ 4.1 两个变量的主分量	(121)
4.1.1 主分量的导出.....	(122)
4.1.2 主分量的性质.....	(123)
4.1.3 主分量的几何意义.....	(124)
4.1.4 标准化主分量.....	(125)
§ 4.2 多个变量的主分量	(126)
4.2.1 主分量的表示.....	(126)

4.2.2	主分量的导出	(126)
4.2.3	主分量的性质	(127)
4.2.4	主分量的其它形式	(128)
4.2.5	主分量的计算	(129)
§ 4.3	经验正交函数分解	(130)
4.3.1	气象要素场分解	(130)
4.3.2	要素场的拟合	(132)
4.3.3	计算中的时空转换	(133)
4.3.4	不同形式的经验正交函数分解	(134)
§ 4.4	主分量分析的应用	(134)
4.4.1	对气象要素场的气候分析	(134)
4.4.2	主分量个数的选取	(138)
4.4.3	利用主分量分析作预报	(138)
4.4.4	主分量分析与其它方法	(139)
参考文献		(140)
第五章 因子分析		(142)
§ 5.1	因子分析的一般模型	(142)
§ 5.2	主要因子	(144)
§ 5.3	特殊因子的考虑	(147)
5.3.1	用相关系数估计特殊方差	(148)
5.3.2	用迭代法估计特殊因子方差阵	(148)
§ 5.4	因子轴的转动	(149)
5.4.1	因子轴转动过程	(149)
5.4.2	极大方差转动	(150)
5.4.3	斜交转动	(153)
§ 5.5	对应分析	(155)
5.5.1	R 型与 Q 型因子分析	(155)
5.5.2	对应分析	(155)
5.5.3	对应分析在聚类中的应用	(157)
§ 5.6	因子分析在气象中的应用	(159)
参考文献		(160)
第六章 典型相关分析		(161)
§ 6.1	典型因子的表示	(161)

§ 6.2	协方差极大原则	(165)
§ 6.3	典型因子的性质及典型相关系数的检验	(168)
§ 6.4	典型因子的回归	(171)
§ 6.5	典型相关分析在气象中的应用	(179)
参考文献		(180)
第七章 聚类分析		(181)
§ 7.1	相似性度量	(181)
7.1.1	相关距离系数	(181)
7.1.2	相似系数	(181)
§ 7.2	逐级归并法	(183)
§ 7.3	平均权重串组法	(184)
§ 7.4	最近矩心串组法	(186)
7.4.1	步骤与原理	(186)
7.4.2	实际例子	(187)
§ 7.5	最优分割法	(188)
7.5.1	变差的计算	(188)
7.5.2	最优分割步骤	(188)
7.5.3	多要素的最优分割	(190)
7.5.4	最优分割的显著性检验	(190)
§ 7.6	聚类分析的应用	(191)
参考文献		(192)
第八章 时间序列分析		(193)
§ 8.1	随机序列的基本概念	(193)
§ 8.2	自回归模型(AR)	(195)
8.2.1	一阶自回归模型	(195)
8.2.2	二阶自回归模型	(196)
8.2.3	p 阶自回归模型	(197)
8.2.4	自回归模型的参数估计	(197)
8.2.5	自回归预报方程	(198)
§ 8.3	滑动平均模型(MA)	(199)
8.3.1	一阶滑动平均模型	(199)
8.3.2	二阶滑动平均模型	(200)

8.3.3	q 阶滑动平均模型	(200)
8.3.4	滑动平均模型的系数估计	(201)
§ 8.4	自回归滑动平均模型(ARMA)	(202)
8.4.1	ARMA(p, q)模型	(202)
8.4.2	ARMA(p, q)模型的自相关函数	(203)
8.4.3	ARMA(p, q)模型的系数估计	(203)
8.4.4	用逆函数估计模型的系数	(205)
8.4.5	ARMA(p, q)模型的预报	(206)
8.4.6	模型阶数选择	(209)
§ 8.5	非平稳时间序列的处理	(211)
§ 8.6	气象中的时间序列分析应用	(212)
	参考文献	(213)
第九章	谱分析	(215)
§ 9.1	谱的概念	(215)
§ 9.2	功率谱	(217)
9.2.1	功率谱的概念	(217)
9.2.2	功率谱估计	(219)
§ 9.3	利用功率谱作周期分析	(223)
9.3.1	离散功率谱检验	(223)
9.3.2	连续功率谱检验	(223)
§ 9.4	滤波	(226)
9.4.1	低通过滤	(227)
9.4.2	高通过滤	(228)
9.4.3	带通过滤	(229)
§ 9.5	交叉谱	(231)
9.5.1	交叉谱概念	(231)
9.5.2	交叉谱的估计	(234)
9.5.3	凝聚谱与位相差谱	(235)
§ 9.6	谱分析的应用	(236)
	参考文献	(237)
第十章	马尔可夫概型分析	(239)
§ 10.1	马尔可夫链	(239)

§ 10.2 转移概率	(239)
§ 10.3 绝对概率	(242)
§ 10.4 转移概率矩阵的谱分解	(243)
§ 10.5 马尔可夫性质的检验	(245)
§ 10.6 马尔可夫概型在气象中的应用	(246)
参考文献	(248)
第十一章 预报的评分与集成	(249)
§ 11.1 离散型变量的预报评分	(249)
11.1.1 预报成功率	(249)
11.1.2 Hedike 评分	(249)
11.1.3 χ^2 评分法	(251)
11.1.4 两类事件列联表的评分	(251)
§ 11.2 连续型变量的预报评分	(252)
11.2.1 平均绝对误差和均方误差	(252)
11.2.2 降水概率评分	(252)
§ 11.3 预报的集成	(253)
11.3.1 权重线性集成	(253)
11.3.2 线性回归集成	(253)
11.3.3 主分量集成	(254)
§ 11.4 统计方法的使用	(254)
参考文献	(256)
习题	(257)
附录 A 矩阵和向量的微分	(259)
附录 B 用消去求逆紧凑方案解非齐次线性方程组	(260)
B-1 求解求逆紧凑变换法	(260)
B-2 紧凑求解求逆的几个性质	(262)
附录 C 求函数的条件极值	(264)
附录 D 求矩阵的特征值及特征向量	(265)
D-1 矩阵的特征值问题	(265)
D-2 用雅可比法求矩阵的特征值及特征向量	(266)
D-3 用幂法求矩阵最大特征值	(269)
D-4 求实非对称阵的特征值及特征向量	(271)
附录 E 常用气象数据处理软件	(275)

E-1	序列分析	(275)
E-2	多要素相关与回归分析	(279)
E-3	要素场分析	(284)
E-4	频谱分析	(288)
附录 F	气象统计常用数表	(294)
F-1	正态分布(密度函数)	(294)
F-2	正态分布函数	(294)
F-3	χ^2 分布	(295)
F-4	F 分布	(296)
F-5	t 分布	(298)

1.1.2 多个变量

气象要素观测是三维空间的,有各种等压面上的要素资料,既有空间,又有时间变化。这时就可以把多个要素在某一段时间收集的资料看作为多个变量的样本,每个变量的样本可看成为一个向量。 p 个变量 n 次观测的样本可看成为 n 维空间中 p 个向量,每个向量可用行向量($1 \times n$ 矩阵)表示

$$\begin{cases} \mathbf{x}_1 = (x_{11} \ x_{12} \cdots x_{1n}) \\ \mathbf{x}_2 = (x_{21} \ x_{22} \cdots x_{2n}) \\ \vdots \\ \mathbf{x}_p = (x_{p1} \ x_{p2} \cdots x_{pn}) \end{cases} \quad (1.2)$$

对第 i 个变量、第 j 个时刻的观测值,可表示为

$$x_{ij} \quad (i = 1, \cdots, p; \quad j = 1, \cdots, n)$$

有时为处理方便对多个要素的资料(样本),可用一个矩阵来表示,记 p 个气象要素的 $p \times n$ 个资料(设 $p < n$) 的样本为

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{pmatrix} \quad (1.3)$$

把 $\mathbf{X}$ 称为资料阵,它是把每个变量资料作为行的形式排列,称为横资料阵。对 p 个要素 $p \times n$ 个资料的样本也可以把每一个变量资料以列的形式排列,写为如下矩阵形式称为竖资料阵。

$$\mathbf{X}^* = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \quad (1.4)$$

上面(1.2) ~ (1.4) 式仅是表示 p 个要素 $p \times n$ 个资料的样本表示方式。它们形式上不同,只是为在不同问题上处理方便而设的记号,本质是一样的。为了不致使读者混淆,我们尽量采用(1.3) 式的形式来表示 p 个要素 n 个资料的样本。例如我们取 12 月、1 月、2 月平均气温的 1951 ~ 1980 年资料(对 1 月、2 月是下一年资料)(见表 1.1 中第 2 至 4 列)。

变量记为 x_1, x_2, x_3 , 对应的各变量的向量表示为

$$\mathbf{x}_1 = (1.0 \quad -5.3 \cdots -3.9)$$

$$\mathbf{x}_2 = (-2.7 \quad -5.9 \cdots -4.8)$$

$$\mathbf{x}_3 = (-4.3 \quad -3.5 \cdots -1.4)$$

由上面三个向量的分量亦可构成一个资料阵

$$\mathbf{X} = \begin{pmatrix} 1.0 & -5.3 & \cdots & -3.9 \\ -2.7 & -5.9 & \cdots & -4.8 \\ -4.3 & -3.5 & \cdots & -1.4 \end{pmatrix}$$