

# 气象数据统计 分析方法

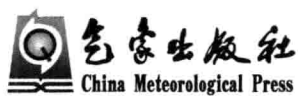
黄嘉佑 李庆祥 编著

QIXIANG SHUJU TONGJI FENXI FANGFA

 气象出版社  
China Meteorological Press

# 气象数据统计分析方法

黄嘉佑 李庆祥 编著



## 内容简介

本书主要介绍了近代数据处理与统计方法及其在各类气象分析中的应用,并加入了许多 21 世纪以来国际上在气象应用统计领域的一些最新成果。此外,还将一些常用的数据处理方法的基本原理和计算步骤编写在附录中,希望有助于气象工作人员参考使用。

本书可以作为具有一定气象学基础的气象及相关行业技术人员从事相关研究及业务工作的参考书,也可以作为气象学硕士、博士研究生开展相关研究的参考书或教材。

### 图书在版编目(CIP)数据

气象数据统计分析方法/黄嘉佑,李庆祥编著.

—北京:气象出版社,2014.12

ISBN 978-7-5029-5792-6

I. ①气… II. ①黄… ②李… III. ①气象数据-统计分析-分析方法 IV. ①P416

中国版本图书馆 CIP 数据核字(2015)第 002453 号

## Qixiang Shuju Tongji Fenxi Fangfa 气象数据统计分析方法

出版发行:气象出版社

地址:北京市海淀区中关村南大街 46 号

总编室:010-68407112

网址:<http://www.cmp.cma.gov.cn>

责任编辑:陈红 黄海燕

封面设计:博雅思企划

印刷:北京中新伟业印刷有限公司

开本:787 mm×1092 mm 1/16

字数:828 千字

版次:2015 年 3 月第 1 版

定价:100.00 元

邮政编码:100081

发行部:010-68409198

E-mail: [qxcbs@cma.gov.cn](mailto:qxcbs@cma.gov.cn)

终审:黄润恒

责任技编:吴庭芳

印张:32.5

印次:2015 年 3 月第 1 次印刷



## 前 言

无论是人类社会或自然科学中,均存在大量的资料数据。巨量资料(big data),或称大数据、海量资料,是人们了解社会和自然的基础。大数据具有四大特点:第一,数据体量巨大;第二,数据类型繁多,包括网络日志、视频、图片、地理位置信息等;第三,价值密度低,在连续不间断观测和监控过程中,可能有用的数据很少;第四,处理速度快,依靠计算机可以快速处理。“大数据”是由数量巨大、结构复杂、类型众多数据构成的数据集合,是基于云计算的数据处理与应用模式,通过数据的整合共享,交叉复用,形成的智力资源和知识服务能力。对大数据处理、挖掘技术、数据分析是时代发展的需要。在大数据时代已经到来的时候需要用大数据思维去发掘大数据的潜在价值。

在现代气象学科的发展中,人们认识天气现象和天气过程,总是先从局地的天气现象的观测开始,了解局地的天气气候变化特征。这种了解过程是利用天基(卫星)和陆基(探空、雷达、观测站和自动气象站等)的全球气象观测网,了解全球范围的天气气候现象,这些数据是大数据。此外,在气象业务和科研中,很多是通过数值模式对大气运动的变化过程进行模拟和预报,其中也存在大量的气象要素和相关物理量的数据。但是,由于过去没有电脑的帮助,无法从大量的数据中总结大气运动的规律性。近年来,随着计算机的发展,利用数据处理和统计学方法来研究天气和气候现象发生的过程和成因,利用气象数据探索大气过程演变的规律成为可能。

近代发展起来的天气气候学、以研究原理为主的物理气候学或动力气候学均离不开气象数据。它们使用气象数据来总结天气概念模式、解释和总结大气动力过程,形成以统计观点和方法为主体又与气象学结合的新分支;它们补充和加强传统气候学的研究,并形成—个气象学中崭新的学科——天气统计学和气候统计学;它们均是利用资料数据,运用数理统计和数据处理方法研究天气和气候的特征、变化和预测的学科。

从现代气候学的观点来看,气候不仅是局地气象要素的平均状态的描述,从空间上扩展到全球范围,即所谓的全球气候系统,扩展到地球—大气—外层空间,即所谓的地球气候系统,即把局地的气候变化看成是地球五大圈——岩石圈、水圈、冰雪圈、大气圈和生物圈,甚至包括地球以外的天文圈相互作用的结果。事实上,以气象与其他学科相互关系为研究对象的学科业已形成,如农田气候学、森林气候学、航空气候学、海洋气候学、水面气候学、山地气候学、水文气候学、城市气候学、建筑气候学、室内气候学、污染气候学、生态气候学、生物气候学和医疗气候学等,均以气象数据处理方法为主要研究手段。当然在气象中,数据处理和统计方法作为研究手段不仅在气候学、天气学和动力气象学上,还广泛用于大气探测、卫星遥感、大气物理和大气化学等学科中。总而言之,以气象资料为基础和以气象数据处理为主要研究方法的学科可以称为气象统计学,它在边缘学科和交叉学科的研究中将起重要作用。

与国内以往气象统计类著作不同之处在于,本书将侧重介绍近代数据处理与统计方法如



何实际应用到各类气象分析中。例如,如何描述气象中的数据,如何对气象数据进行基础处理,如何对气象要素的时间演变序列做规律性分析,如何研究不同气象要素之间的变化关系,如何分析一个地区的气象要素和气候要素的时空变化,如何做气象要素和气候变化的预报和预测,如何进行气象资料均一化处理,如何对气象资料进行插补,等等。本书正文内容主要介绍数据处理方法如何应用在气象中;并将一些常用的数据处理方法的基本原理和计算步骤作为参考编写在附录中,希望有助于从事气象研究和业务的工作人员参考使用。

由于水平有限,本书错误和不当之处在所难免,敬请读者指正。

编著者

2014年10月



# 目 录

## 前 言

<b>第 1 章 气象资料与数据</b> .....	( 1 )
1.1 气象数据的特征 .....	( 1 )
1.2 气象数据的类型 .....	( 1 )
1.3 气象数据的描述 .....	( 2 )
1.4 数据整理的方法论 .....	( 6 )
参考文献.....	( 6 )
<b>第 2 章 大气基本状态</b> .....	( 7 )
2.1 大气平均状态的描述 .....	( 7 )
2.2 大气状态的异常 .....	( 8 )
2.3 大气平均状态的代表性 .....	( 10 )
2.4 大气平均状态的差异性 .....	( 11 )
2.5 大气状态出现的频率 .....	( 12 )
2.6 大气变量的分布 .....	( 14 )
2.7 大气状态的分级 .....	( 18 )
2.8 大气异常的极端状态 .....	( 21 )
2.9 大气变量的数据变换 .....	( 24 )
参考文献.....	( 26 )
<b>第 3 章 大气变量的相互关系</b> .....	( 28 )
3.1 大气状态的关联性 .....	( 28 )
3.2 信息关联 .....	( 30 )
3.3 列联表 .....	( 32 )
3.4 级别变量的相关 .....	( 33 )
3.5 连续变量的相关 .....	( 34 )
3.6 偏相关 .....	( 37 )
3.7 不同时刻的交叉相关 .....	( 38 )
3.8 不同时段的相关 .....	( 39 )
3.9 变量的瞬时相关 .....	( 40 )



3.10	变量相关程式关系	(43)
3.11	多个变量的相关	(44)
3.12	变量变化的相似性	(45)
	参考文献	(47)
<b>第4章</b>	<b>大气变量的时间演变特征</b>	<b>(49)</b>
4.1	离散变量持续性	(49)
4.2	连续变量的持续性	(51)
4.3	变量的变化趋势	(54)
4.4	变量变化趋势的检验	(60)
4.5	变量变化的突变	(61)
4.6	变量变化的周期性	(70)
4.7	两个变量变化的交叉周期性	(82)
4.8	变量的时间变化滤波	(84)
	参考文献	(86)
<b>第5章</b>	<b>大气变量场基本特征</b>	<b>(89)</b>
5.1	大气变量场基本状态	(89)
5.2	变量场的变化特征	(92)
5.3	变量的条件平均场	(94)
5.4	变量信号场	(99)
5.5	条件差值场	(104)
5.6	外力影响特征	(108)
	参考文献	(113)
<b>第6章</b>	<b>大气变量场中的相关性</b>	<b>(115)</b>
6.1	遥相关	(115)
6.2	高度场时空特征模态	(119)
6.3	大气涛动	(124)
6.4	地面要素场时空特征模态	(128)
6.5	风场时空特征模态	(133)
6.6	变量场中的关联性	(136)
6.7	变量场中气候分类	(139)
6.8	变量场中的波动特征	(146)
	参考文献	(150)
<b>第7章</b>	<b>两个变量场的关系</b>	<b>(154)</b>
7.1	两个变量场的差异性	(154)
7.2	两个变量场的相似性指标	(155)



7.3	两个风场相似性 .....	(157)
7.4	两个变量场的关联性 .....	(158)
7.5	两个变量场的回归关系 .....	(165)
7.6	两个变量场的耦合关系 .....	(168)
	参考文献 .....	(179)
<b>第 8 章</b>	<b>多个变量场耦合分析 .....</b>	<b>(181)</b>
8.1	多变量场的综合模态 .....	(181)
8.2	风场综合模态 .....	(182)
8.3	多变量场模态 .....	(184)
8.4	变量场时间演变模态 .....	(185)
8.5	联合耦合模态 .....	(186)
	参考文献 .....	(189)
<b>第 9 章</b>	<b>变量场的时间演变特征 .....</b>	<b>(190)</b>
9.1	变量场的趋势分析 .....	(190)
9.2	变量场的平均序列演变特征 .....	(192)
9.3	变量场的时间演变特征提取 .....	(194)
9.4	变量场的周期变化特征 .....	(198)
9.5	变量场时间演变模态 .....	(203)
9.6	风场的时间演变特征 .....	(207)
9.7	多变量场时间演变特征 .....	(210)
9.8	周期变化外力成因 .....	(214)
	参考文献 .....	(217)
<b>第 10 章</b>	<b>大气变量的预报 .....</b>	<b>(219)</b>
10.1	外因子预报模型 .....	(219)
10.2	持续性预报模型 .....	(224)
10.3	周期模式预报模型 .....	(225)
10.4	时间外延预报模型 .....	(227)
10.5	大气变量场的预报 .....	(229)
10.6	潜在可预报性 .....	(236)
10.7	预报的稳定性 .....	(243)
10.8	预报效果评价 .....	(247)
	参考文献 .....	(256)
<b>第 11 章</b>	<b>大气变量动力统计预报 .....</b>	<b>(259)</b>
11.1	数值预报产品释用 .....	(259)
11.2	随机气候模式 .....	(267)





11.3	动力系统的可预报性	(274)
11.4	集合预报	(276)
11.5	降尺度预报	(279)
11.6	数值预报的统计订正	(285)
	参考文献	(289)
<b>第 12 章</b>	<b>气象数据修正、插补和融合</b>	<b>(292)</b>
12.1	气象资料数据的质量评估与控制	(292)
12.2	气象资料数据的均一性处理	(295)
12.3	城市化对气候序列均一性的影响	(302)
12.4	资料恢复与插补	(305)
12.5	大气变量场的空间插值	(309)
12.6	气象数据融合、同化及再分析	(315)
	参考文献	(317)
<b>附录 A</b>	<b>回归分析</b>	<b>(321)</b>
A1	单个因子的回归模型	(321)
A2	多因子线性回归模型	(325)
A3	逐步回归模型	(331)
A4	事件概率回归(REEP)	(335)
A5	Logit 回归模型	(337)
A6	最佳子集回归模型	(339)
A7	预报残差最小逐步回归	(342)
A8	权重回归	(344)
A9	卡曼滤波回归	(346)
A10	岭回归	(348)
A11	贝叶斯回归	(349)
A12	支持向量机回归	(351)
<b>附录 B</b>	<b>判别分析</b>	<b>(354)</b>
B1	费歇判别方程	(354)
B2	贝叶斯判别方程	(359)
B3	逐步判别	(360)
B4	回归逐步判别	(363)
<b>附录 C</b>	<b>变量场的分解</b>	<b>(365)</b>
C1	主分量分析	(365)
C2	经验正交函数分解	(368)
C3	多变量场经验正交函数分解	(373)

C4	复向量经验正交函数分解 .....	(374)
C5	扩展经验正交函数分解 .....	(377)
C6	联合经验正交函数分解 .....	(378)
C7	复经验正交函数分解 .....	(379)
C8	主振荡模态分析 .....	(382)
C9	独立分量分析 .....	(384)
<b>附录 D</b>	<b>聚类分析 .....</b>	<b>(388)</b>
D1	因子分析的一般模型 .....	(388)
D2	主因子分析模型 .....	(390)
D3	因子轴的转动 .....	(393)
D4	对应分析 .....	(396)
D5	串组法 .....	(398)
<b>附录 E</b>	<b>变量场的耦合分析 .....</b>	<b>(402)</b>
E1	典型相关分析 .....	(402)
E2	奇异值分解 .....	(413)
E3	偏最小二乘回归 .....	(416)
<b>附录 F</b>	<b>大气变量时域分析 .....</b>	<b>(419)</b>
F1	自回归模型 .....	(419)
F2	滑动平均模型 .....	(423)
F3	自回归滑动平均模型 .....	(425)
F4	方差分析模型 .....	(428)
F5	均生函数模型 .....	(429)
F6	经验模态分解 .....	(430)
F7	去趋势的涨落分析 .....	(430)
<b>附录 G</b>	<b>大气变量频域分析 .....</b>	<b>(432)</b>
G1	变量的频谱 .....	(432)
G2	功率谱 .....	(434)
G3	非整谱 .....	(439)
G4	最大熵谱 .....	(440)
G5	双谱分析 .....	(442)
G6	多窗口谱分析 .....	(443)
G7	滤波 .....	(444)
G8	交叉谱 .....	(451)
G9	奇异谱 .....	(454)
G10	交叉奇异谱 .....	(456)



G11	小波分析	(457)
G12	交叉小波谱	(459)
<b>附录 H</b>	<b>大气变量场的谱分析</b>	<b>(460)</b>
H1	纬向谐波分析	(460)
H2	高度场的物理量谱	(462)
H3	高度场的球谐分析	(464)
H4	时空谱	(465)
H5	二维空间谱分析	(467)
H6	变量场中的交叉谱分析	(467)
H7	多窗口—奇异值分析	(468)
H8	循环平稳经验正交函数分解	(469)
<b>附录 I</b>	<b>马尔科夫概型分析</b>	<b>(471)</b>
I1	马尔科夫链	(471)
I2	转移概率	(471)
I3	绝对概率	(473)
I4	转移概率矩阵的谱分解	(474)
I5	马尔科夫性质的检验	(475)
<b>附录 J</b>	<b>神经网络</b>	<b>(477)</b>
J1	神经元模型	(477)
J2	神经网络结构	(479)
J3	网络学习	(481)
J4	前馈型神经网络	(482)
J5	径向基函数网络	(486)
J6	自组织映射网络	(487)
<b>附录 K</b>	<b>统计检验</b>	<b>(490)</b>
K1	假设检验	(490)
K2	平均值检验	(491)
K3	两组样本平均值差异的检验	(491)
K4	方差检验	(492)
K5	相关系数的检验	(493)
K6	变量的分布检验	(495)
K7	频率的检验	(496)
K8	趋势检验	(497)
K9	突变检验	(498)
K10	蒙特卡洛检验	(501)

附录 L 气象统计常用数表 .....	(502)
L1 正态分布函数 .....	(502)
L2 $\chi^2$ 分布 .....	(503)
L3 $F$ 分布 .....	(504)
L4 $t$ 分布 .....	(506)



# 第1章 气象资料与数据

## 1.1 气象数据的特征

目前,在气象业务系统中拥有大量的气象资料和各种其他有关的信息资料(直接或间接获得的观测数据),它们是来自地面、高空、雷达和卫星的观测资料,隐含着丰富的地球环境和大气环流信息,这些信息既包含地区天气变化和气候环境状况,也包含外界影响局地天气和气候的因素。人类和大部分生物生存在地球大气中,大气圈是人类对气候变化感受的主要空间环境,人类生存环境与大气、下垫面生态环境有关,也与人类居住环境有关,如果能够从数据中找出其演变规律,就能够对人类环境的未来状态进行预测。因此,如何利用这些资料数据,分析其基本特征及其演变规律,是开发利用气象资源的重要内容。

地球系统包含岩石圈、冰雪圈、大气圈、水圈和生物圈,圈层之间存在相互作用。气象资料主要来自大气圈观测资料,诊断大气中的天气和气候状态及变化是通过气象要素和气候要素来描述的,即通过气候要素的量化描述,来反映气候各子系统的状态和变化过程。大气系统的气候要素常用的有温度、降水、气压,此外还有水汽含量、风向风速、二氧化碳含量、臭氧含量等。考虑到大气圈与水圈相互作用,还常常使用海洋和河流的各种物理量的测量资料。例如,海洋系统常用的有海表海温、深层海温和海洋热含量等,它们是描述海洋向大气子系统传输热量多少的要素。陆地中常用河流流量、水库水位等资料。冰雪系统常用积雪厚度、冰雪范围来描述冷源作用的程度和大小。陆面系统中常用火山喷发和火山灰强度指数来描述。生物圈的作用目前还无公认的描述要素。上述的资料数据都可以用于气象规律性的分析,均可以称为气象资料数据。

## 1.2 气象数据的类型

地球系统中的气象资料类型大体有两种,一种是定性资料。例如,天气现象,雷暴、霜冻、雾霾和沙尘暴等现象,观测资料中使用“有”和“无”来记载。对地区温度的气候状态,一般分为特冷、冷、正常、暖和特暖5个级别进行描述。有时为描述简便,也可以把温度的气候状态分为两级,即冷和暖。对地区降水的气候状态,也有特旱、旱、正常、涝和特涝5个级别的状态描述,或旱和涝两个级别的状态描述。在我国历史上,气温或降水的状态描述还使用文字来记载。例如,某地的县志等资料记录该地某年某月“城门水深三尺”“颗粒无收”等,均是历史文献中有关水旱状态的文字描述。这些文字记载也是一种定性资料。

另一种是定量资料。即由现代仪器观测(各种在地面观测和卫星中使用的仪器)得到的资料数据,常常表现为连续变化的数值资料。



在气象观测资料数据中有元数据,即“数据的数据”,或者叫作用来描述数据的数据,或者叫作信息的信息。元数据可以简单地理解成最小的数据单位,可以为数据说明其元素或属性(如名称、大小、数据类型等),或其结构(如长度、字段、数据列),或其相关数据(如位于何处、如何联系、拥有者)等。

非气象领域的专业人员也需要从观测中提取最准确的数据,并且通常和从不同地方或以不同次数采集的数据相比较。同时需要对测量条件有完整的了解。数据的准确性可能受到仪器类型、仪器安装位置、记录程序以及许多其他因素的影响。尽量保留所有元数据的记录,从而可以最大可能地利用这些数据。全球气候观测系统(GCOS)气候监测的原则是,“当地条件、仪器、操作程序、数据处理算法以及其他与解释数据相关的因素(例如元数据)的历史和细节都应该与数据本身得到同等重要的重视”(WMO,2002)。

为了计算处理方便,需要把定性的气象资料转化为数据。例如,为了数量化描述,定性资料常常以离散数值描述。例如,把“有”和“无”描述的天气现象,以及“冷(旱)”和“暖(涝)”的两种级别的记载,记为“1”和“0”的数据资料。对气温(降水量)划分的5级气候状态,记为“1”“2”“3”“4”和“5”的数据资料(中央气象局气象科学研究院,1981),即可以把定性的气象资料转变为定量的数据资料。但是,这种数据与气象仪器观测的连续型数据不同,它们是自然数列,是离散型的数据。因此,在气象中的数据类型有两种:离散型数据、连续型数据。

不同数据类型,根据研究分析需要,可以相互转化。例如,对连续型降水量数据,根据降水量数值大小,把降水量分为小雨、中雨、大雨、暴雨、大暴雨和特大暴雨等6个级别,其中小雨指1 d(或24 h)降雨量小于10 mm者;中雨指1 d(或24 h)降雨量10~25 mm者;大雨指1 d(或24 h)降雨量25~50 mm者;暴雨指1 d(或24 h)降雨量50~100 mm者;大暴雨指1 d(或24 h)降雨量100~200 mm者;特大暴雨指1 d(或24 h)降雨量在200 mm以上者。这样,连续型数据就转化为1~6个级别的离散型数据。

对离散型气象数据,有时为了处理方便,把它们看成连续型中不同时刻的观测值,也可以作为连续型数据处理。

### 1.3 气象数据的描述

为了便于数学的研究,常把各种大气中描述天气过程的气象要素,或描述气候变化的气候要素,以及与大气层有相互作用的各气候系统圈层的物理量(如海洋表面和深层的海温、极地的冰盖、高原的雪盖等)统称为大气变量,由于它们是影响大气层变化的物理量,是随时间变化的变量,有时为了方便,也简称为变量。

大气变量的数据类型,是离散型变量和连续型变量。根据研究对象不同,研究的变量可以是单变量,如某一个测站的气温或降水量;也可以是多变量,例如,中国范围所有观测站某时刻的气温或降水量,它们的观测值组成一个空间分布图像,常称为气温场或降水量场,统称为某时刻的大气变量场。场中测站分布可以是不规则的,如中国气象观测站点的分布、某月中国降水量的分布(中国气象局气候变化中心,2012);也可以是规则的,如高度场中的格点、某一等压面(图1.3.1)。常常把这些变量称为大气变量场。例如,温度场、降水量场和高度场等。

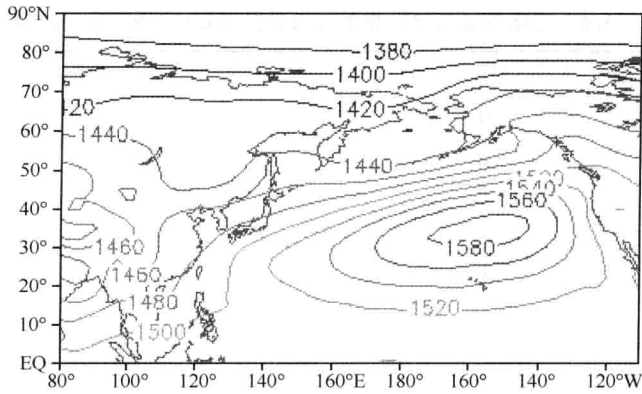


图 1.3.1 850 hPa 高度场(单位:gpm)

气象中观测数据是随时间变化的,常常把变量按时间顺序排列的一组数据,称为时间序列,或称为该变量时间序列,或简称为序列(图 1.3.2)。如果研究两个变量以上的数据,则称为多变量时间序列。例如,某一等压面(如 850 hPa)中有很多网格点,某个格点都有随时间变化的观测数据,它们一起称为多变量时间序列数据,也称为变量场数据,它们随空间和时间变化。

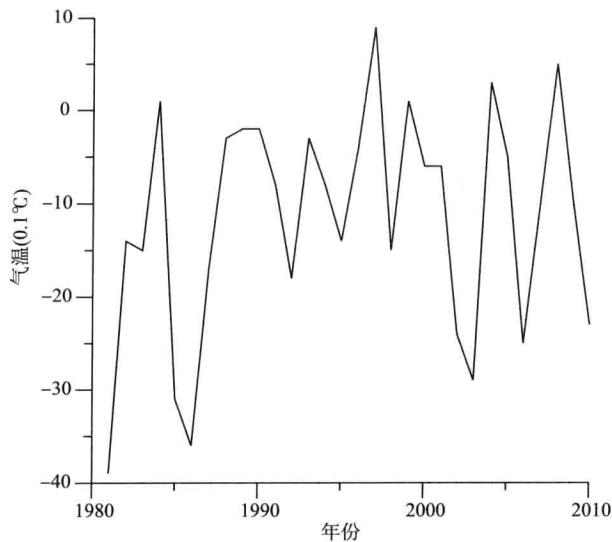


图 1.3.2 北京 12 月气温逐年变化序列(1981—2010 年)

对气象数据进行研究是为了了解研究对象整体的变化规律。研究对象的全体也称为总体,我们所取得的气象数据是大气变量总体的部分,称为样本。数据采样的数量称为样本容量。

气象数据除了用变量样本数据的数值变化描述外,常以列表形式给出,例如,北京近 1 个气候阶段(30 年)冬季(12 月—次年 1 月)气温和夏季(6—8 月)降水量 1981—2010 年逐年变化序列(表 1.3.1)。



表 1.3.1 北京冬季气温和夏季降水量数据资料(1981—2010 年)

年份	冬季气温(0.1℃)			夏季降水量(0.1 mm)		
	12 月	1 月	2 月	6 月	7 月	8 月
1981	-39	-48	-14	194	1709	1025
1982	-14	-43	4	1402	2074	1194
1983	-15	-29	-21	665	265	2252
1984	1	-45	-25	441	600	2775
1985	-31	-47	-19	320	2895	2977
1986	-36	-37	-18	2032	1629	1427
1987	-17	-36	1	912	1309	2465
1988	-3	-29	-14	618	2787	2040
1989	-2	-20	16	390	1248	1042
1990	-2	-49	-6	40	2230	1570
1991	-8	-23	1	2363	1980	1247
1992	-18	-11	18	694	1539	1414
1993	-3	-37	16	392	2064	1585
1994	-8	-16	8	236	4592	2142
1995	-14	-7	21	689	1956	1199
1996	-4	-22	-4	551	3074	2500
1997	9	-38	13	355	1398	832
1998	-15	-39	24	1429	2479	1144
1999	1	-16	22	240	596	570
2000	-6	-64	-15	190	615	1505
2001	-6	-54	-15	458	1286	497
2002	-24	1	34	1035	549	743
2003	-29	-31	9	661	577	342
2004	3	-23	29	696	1820	507
2005	-5	-28	-29	664	961	1234
2006	-25	-19	-9	351	1550	475
2007	-10	-15	37	461	1162	1036
2008	5	-30	6	1253	793	1321
2009	-10	-30	10	955	1966	609
2010	-23	-48	-10	887	340	1778

数据还可以用变量值随时间的变化曲线表现(图 1.3.2)。在数学上,变量常用字母来表示,例如,以  $x$  表示某大气变量,它的一组观测数据称为该变量的样本,用式(1.3.1)表示:

$$x_i (i = 1, 2, \dots, n) \quad (1.3.1)$$

式中,  $i$  表示数据顺序,把样本中某个时刻观测的数据称为样品;  $n$  表示数据的总个数,也称为样本容量。变量数据也可以表示为函数形式,为

$$x(i) (i = 1, 2, \dots, n) \quad (1.3.2)$$

气象中的变量数据是随时间变化的,也常把  $i$  改为  $t$  来表示不同时间所取得的样本。记为



$$x_t (t = 1, 2, \dots, n) \quad (1.3.3)$$

变量数据的变化可以用图表示,例如,图 1.3.2 给出了北京 12 月气温 1981—2010 年的变化情况,通过曲线可以相互比较,其变化趋势也一目了然。

离散型的变量数据也可以表示为时间序列形式。例如,海冰结冰的冰情,按轻、较轻、正常、较重、重分为 5 级描述。图 1.3.3 给出了 1932—2000 年黄、渤海海冰冰级的年际变化曲线。

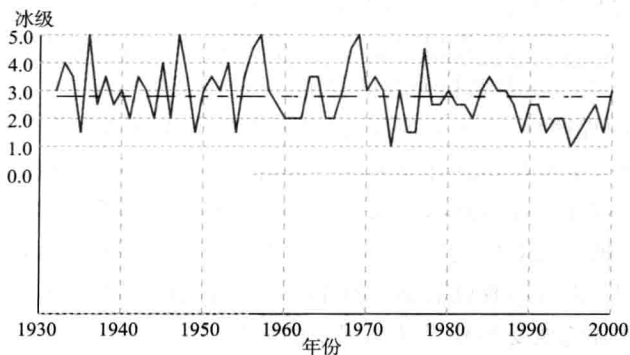


图 1.3.3 1932—2000 年黄、渤海海冰冰级年际变化(引自李剑等,2005)

对单个变量数据,有时为了矩阵计算方便,也可以使用向量表示。即表示为

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \quad (1.3.4)$$

该向量称为行向量,记为  $\mathbf{x} (1 \times n)$  或  $\mathbf{x}_{1 \times n}$ 。

当然,也可以表示为列向量,即

$$\mathbf{x}_{n \times 1} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (1.3.5)$$

记为  $\mathbf{x} (n \times 1)$  或  $\mathbf{x}_{n \times 1}$ 。

对包含多变量的大气变量场数据,通常表示为矩阵。例如,含  $p$  个变量(或格点)的变量场的数据矩阵,记为

$$\mathbf{x}_{p \times n} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{pmatrix} \quad (1.3.6)$$

矩阵中每一行的数据是某变量的  $n$  个数据。当然  $p$  个变量的数据矩阵也可以按列记录场中某变量的数据,表示为

$$\mathbf{x}_{n \times p} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \quad (1.3.7)$$