

东北地区降水分布特性的主成分分析*

赵 焕 宸

(中国科学院长春地理研究所)

主成分分析是一种多变量数学分析方法之一。这种方法,在气象现象解析上的应用,最早是 Loretz, E.N. 在1956年对北美海面上的空间气压场的解析。之后,如陆续有气压 (Kutzbach, J.E., 1967, 1970)、气温(三上, 1975)、风(Barnett, T.P., 1977)、降水(北岛, 1981)等各种气象要素的主成分分析,论述了各种气象要素水平分布的优势型及其时间变动特性。

本文以历年月平均降水量为例,进行降水变动的主成分分析,划分成各主成分的多年降水偏差分布型(简称降水分布型)。然后,通过各主成分逐月得分的时间序列,各主成分年间得分的时间序列,探讨了主要降水分布型近20年来的变动倾向和周期性。特别是通过年间变化的解析,查明由于季节变动而引起的差异。另外,还通过月平均气压的主成分分析,说明月平均降水的优势水平分布型形成的原因和过程。为了进一步揭示东北地区内的降水差异,依据主成分聚类分析,划分了降水分布型的气候区,阐明了区域气候特性,从而为合理配置生产力,振兴区域经济提供科学依据。

一、主成分分析的资料和研究方法

1. 主成分分析的必要性 过去,在研究区域气候特性和主要气候因子的长期变动时,都用某个或某几个典型站平均值的时间序列和等值线图。事实上,某地气候要素的时间变化是和其它地点的相同要素以及其它要素相互制约、相互依存的。而且,东北地区面积辽阔,地形各异,境内气候差异明显,有独特性。为了客观地揭示本地区的降水偏差及其时间变动特性,有必要对降水进行只含观测点数的多次元时间序列的主成分分析。

2. 资料和研究方法 本文选取有代表性的35个气象台站,1961~1978年逐月平均降水量资料。研究方法,是把观测点数作为 p ,各观测点的资料数作为 n , m 个($m < p$)主成分,用 p 个特性值(本文中指各站月降水量的平均偏差)的一次结合式(或称一次线性组合)表示之,即:

$$\begin{cases} Z_1 = l_{11}x_1 + l_{12}x_2 + \cdots + l_{1p}x_p \\ Z_2 = l_{21}x_1 + l_{22}x_2 + \cdots + l_{2p}x_p \\ \cdots \cdots \\ Z_m = l_{m1}x_1 + l_{m2}x_2 + \cdots + l_{mp}x_p \end{cases}$$

* 东北地区包括黑龙江、吉林、辽宁三省,以及内蒙古自治区的呼伦贝尔、哲里木、昭乌达三个盟。

本文计算工作是在日本北海道大学大型计算中心完成的。计算过程中,得到导师高桥英纪先生、加藤央之博士的指导,并承刘兴士、郝凌云副研究员审阅,提出宝贵意见,插图由齐秀兰同志清绘,在此表示感谢。

在这里

$$\sum_{i=1}^p l_{ki}^2 = 1 \quad (k = 1, 2, \dots, m)$$

依上式计算结果,使第 1 主成分的方差成为最大,而第 k 主成分 Z_k 同其它主成分 $Z_1, Z_2, \dots, Z_{k-1}$ 成为无相关。公式中 l_{ki} 是从 $[x]$ 的方差-协方差矩阵算出的特征值 λ_k 的特征向量成分。它的意义如由 l_{ki} 规定的空间变动是月降水量多年偏差中的最优势分布型,它的盛衰由 Z_1 的大小来表示。

关于各主成分 Z_k 的变动对于全变动的比例,即贡献率 V_k ($k = 1, 2, \dots, m$) 以及 m 个主成分的累积贡献率 V_{ck} , 由下式求出,用百分率表示。

$$V_k = (\lambda_k / \sum_{i=1}^p \lambda_i) \times 100$$

$$V_{ck} = \sum_{k=1}^m V_k$$

在这里,各主成分的方差之和 ($\sum_{i=1}^p \lambda_i$) 等于全部观测点的 x_i 的方差之总和。因此,第 K 主成分 Z_k 便能说明降水偏差的全地域变动的 $\alpha\%$ 。

另外,关于 n 个资料,其主成分的得分 Z_{ak} ($a = 1, 2, \dots, n; k = 1, 2, \dots, m$) 由下式计算:

$$\begin{cases} Z_{a1} = l_{11}x_{a1} + l_{12}x_{a2} + \dots + l_{1p}x_{ap} \\ Z_{a2} = l_{21}x_{a1} + l_{22}x_{a2} + \dots + l_{2p}x_{ap} \\ \dots\dots\dots \\ Z_{ak} = l_{k1}x_{a1} + l_{k2}x_{a2} + \dots + l_{kp}x_{ap} \end{cases}$$

文中, $p = 35$, $n = 12$ (月) 乘以 18 (年) 等于 216, 对 35 乘 216 的资料行列进行主成分分析。由 35 乘 216 的降水偏差的资料行列 (降水量的偏差值, 是取各地点每月降水量正规化后, 即立方根后, 所求得的降水量的偏差值。) 算出 35 乘 35 的相关矩阵, 再依 Jacobi 法^[1] 求出特征值和特征向量。而贡献率、累积贡献率和各主成分的得分则由前述公式算出。

另外, 由降水的主成分分析结果, 求各地点的降水时间变动特性值 (因子负荷量) 因子负荷量由下式求出:

$$f_{ij} = \sqrt{\lambda_i} \cdot l_{ij} / \sqrt{S_j}$$

S_j ——表示在第 j 地点的降水 (x_j) 的方差

λ_i ——表示第 i 主成分的特征值

l_{ij} ——表示第 i 主成分的特征向量

f_{ij} ——表示 Z_i 和 x_i 的因子负荷量

最后, 用各主成分 x_i 和 Z_i 的因子负荷量 (计算到第 4 主成分) 的值, 由群平均法^[5] (Group average method) 求出 4 次元空间的聚类间的类似度 (或称距离)。由类似度作成树形图, 表示聚类过程。进而对东北地区的多年降水偏差分布型, 进行降水的气候分区。

二、主成分分析结果解析

1. 降水偏差分布特性 对历年逐月平均降水量偏差，进行主成分分析，把得到的前 6 个主成分的特征值、贡献率和累积贡献率组列为表 1。从表上看出，前 4 个主成分，虽然只能大约说明全变动的 60%，但是，降水现象同气温等气候要素相比，它是局地性很强的气候要素，因此，可认为前 4 个主成分在较大程度上能表现出本区的降水偏差变动特性，或者说 5 成分以下是不重要的。所以，只对前 4 个主成分进行解析。

表1 降水的主成分分析的示性数

Table 1 The indicative value of the principal components for precipitation

各主成分	特征值	贡献率 (%)	累积贡献率 (%)
1	5.77	33.0	33.0
2	1.96	11.2	44.2
3	1.42	8.1	52.3
4	1.30	7.4	59.7
5	0.70	4.0	63.7
6	0.59	3.4	67.1

2. 多年降水偏差特征向量的地理分布和解析 对上述表 1 抽出的前 4 个主成分，分别作多年降水偏差的特征向量分布图 (图 1)。各主成分主要表示各有向量分布图上绝

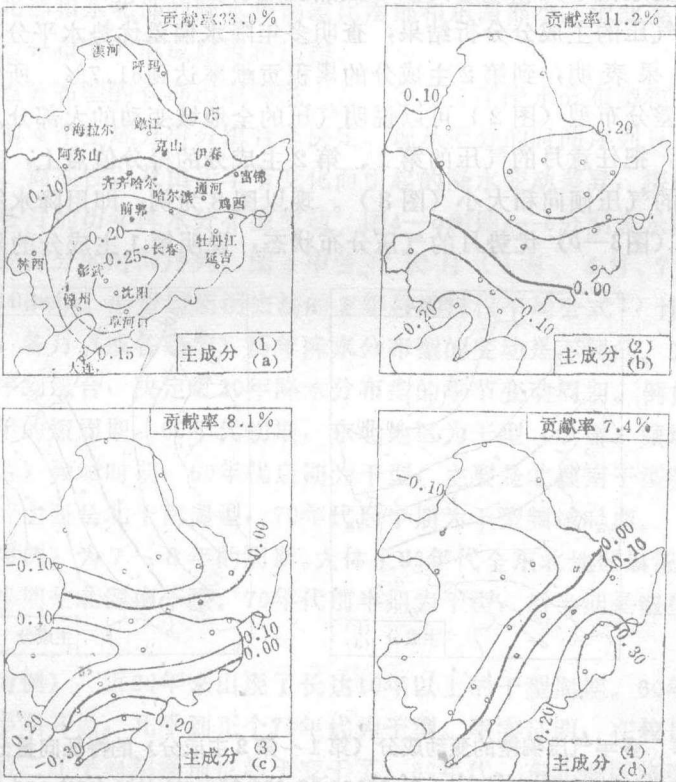


图1 多年降水偏差的变动成分 (第 1~第 4 成分) 的特征向量分布

Fig.1 The eigenvector distribution of changing components (from the first to the fourth principal components) of perennial precipitation deviation

对值大的地域的降水量变动。由于各固有向量分布图的符号也包括有逆符号,所以,每个固有向量分布图可解释为分别表示两个降水偏差分布型。现将各主成分的特征叙述如下。

第 1 主成分说明全变动的 33.0% (图 1-a)。全东北地区是正的同符号,最大值分布在东北地区的东南部,中心在四平、沈阳一带,最低值分布在黑龙江沿岸的呼玛一带。第 1 主成分可以看作是全东北地区湿型或全东北地区干型¹⁾的降水偏差分布型。

第 2 主成分说明全变动的 11.2% (图 1-b)。东北地区北部和南部的符号相反,北部为正,南部为负。正域以小兴安岭山地为中心,负域以辽东湾附近为中心。第 2 主成分可以看作是北湿南干型或北干南湿型的降水偏差分布型。

第 3 主成分说明全变动的 8.1% (图 1-c)。这个主成分具有呈东西走向的相互交错分布的正、负域,它表示大、小兴安岭山脉、长白山脉一带(负偏差),同东北平原(包括三江平原)一带(正偏差)的干、湿型,呈相反的降水偏差分布型。第 3 主成分是明显地表现出局地降水偏差分布型的主成分。

第 4 主成分说明全变动的 7.4% (图 1-d)。东北地区东部为正,西部为负。第 4 主成分可以看作东湿西干型或东干西湿型的降水偏差分布型。从降水偏差分布型来看,东北地区的降水受东南季风和地形两因子强烈支配着。

现依月平均气压的主成分分析结果,查明多年降水偏差优势水平分布型的形成原因和过程。计算结果表明,到第 2 主成分的累积贡献率达到 61.7%,所以用前两个主成分的多年气压偏差分布型(图 2)可以说明气压的全地域变动的大部分。另外,如果在 $P_{z1} \sim P_{z2}$ 平面上,把任意月的气压的第 1、第 2 主成分的得分值点上,其所在位置便能表示该月的概略的气压倾向和大小(图 3)。现以图 3 为例,即用降水第 1 主成分的正(图 3-a)、负(图 3-b)优势月的气压分布状态,说明第 1 主成分的降水偏差分布型

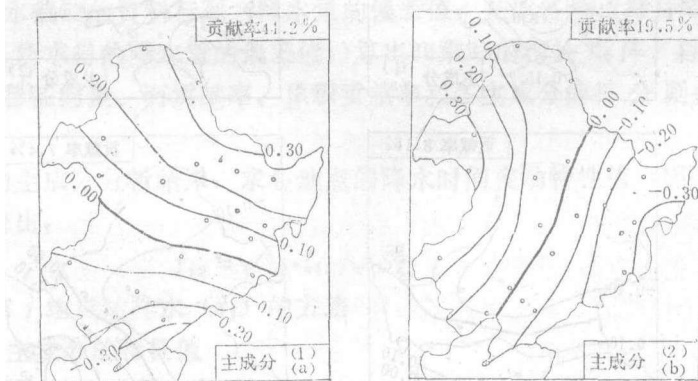


图 2 多年气压偏差的变动成分(第 1~第 2 主成分)的特征向量分布

Fig. 2 The eigenvector distribution of changing components (the first and the second principal components) of perennial air pressure deviation

1) 湿型或干型,指比平均降水量多(即正偏差)或少(即负偏差)。

的形成原因和过程如下:

根据图3—a, 在正的降水偏差优势分布型时, 北部气压比南部相对较高。这种情况, 是由于蒙古高压向西北偏北方向移动而变弱, 印度低压从西南向东北延伸, 东部的太平洋高压势力逐渐向西增强, 而形成全东北地区湿型(多雨型)的降水偏差分布型。此时, 从东南部海面吹向内陆的湿润而温暖的东南季风是盛行风向。所以, 在东南部形成多雨中心。

根据图3—b, 在负降水偏差的优势分布型时, 西部气压比东部相对较高。这表明, 西部且向南扩张的蒙古高压, 和南部且向东北偏移的鄂霍次克海低压同时存在。此时, 东北地区盛行西风和西北风, 形成全东北地区干型(少雨型)的降水偏差分布型。但是, 东北部由于鄂霍次克海低压的影响, 使北部和东北部的降水负偏差比南部和东南部小。依此类推, 可解释其它主成分的降水偏差分布型的形成原因和过程。

3. 降水偏差分布型的长期出现特性 关于各月历年降水分布型和年间降水分布型, 分别由上述前4个主成分得分组合来决定。现从得分的时间序列只对前2个主成分的多年变动倾向、周期性以及由于季节变化而引起的降水变动差异, 探讨于下:

(1) 关于各月的历年降水分布型变动 图4—a和图4—b分别表示降水的第1主成分和第2主成分的得分的时间序列。图4中各代表月(1月、4月、7月、10月)的虚线, 是用神山(1968年)的短周期消去法的5年移动加权平均公式¹⁾计算的周期变动曲线。图4表明, 各月(或各季节)历年降水分布型的变动是不同的。这里用前两个主成分的得分时间序列组合, 决定近20年降水分布型的季节变动周期。例如, 冬季(1月为例)为3~5年的短周期。60年代初期, 东北地区为干型(少雪)频域时期, 60年代中期为湿型(多雪)频域时期, 60年代后期为干型, 主要是北湿南干型频域时期。70年代上半期为湿型, 主要是北干南湿型, 70年代后半期为干型频域时期。

春季(4月为例)为7~8年的周期。大体上60年代全东北地区属湿型, 而前半期是北干南湿型, 后半期是北湿南干型。70年代前半期为干型, 后半期是湿型, 且北湿南干型。

夏季(7月为例), 近20年来出现了长达10年以上的干型周期。60年代前半期为湿型, 从60年代后半期开始, 几乎到正个70年代为干型。事实证明, 在较长期的干型时期控制下, 连湿润的三江平原沼泽地也趋于变干了。80年代, 东北地区将进入湿型周期。

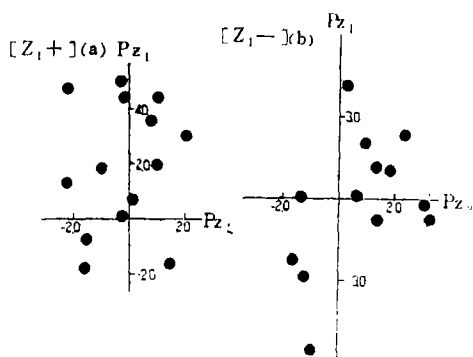


图3 降水第1主成分的正(a)、负(b)优势月的气压分布(Pz_1, Pz_2 分别表示气压的第1、第2主成分)

Fig. 8 The air pressure distribution of the positive (a) and negative (b) of the first principal component for precipitation in dominant months

1) $x'_3 = 0.1102(x_1 + x_5) + 0.2374(x_2 + x_4) + 0.3048x_3$

秋季(10月为例),近20年来也出现了10年左右的较长的干、湿周期。60年代虽然有干湿波动,但总的说来是属于干型时期,其中,60年代后半期,是北湿南干型。70年

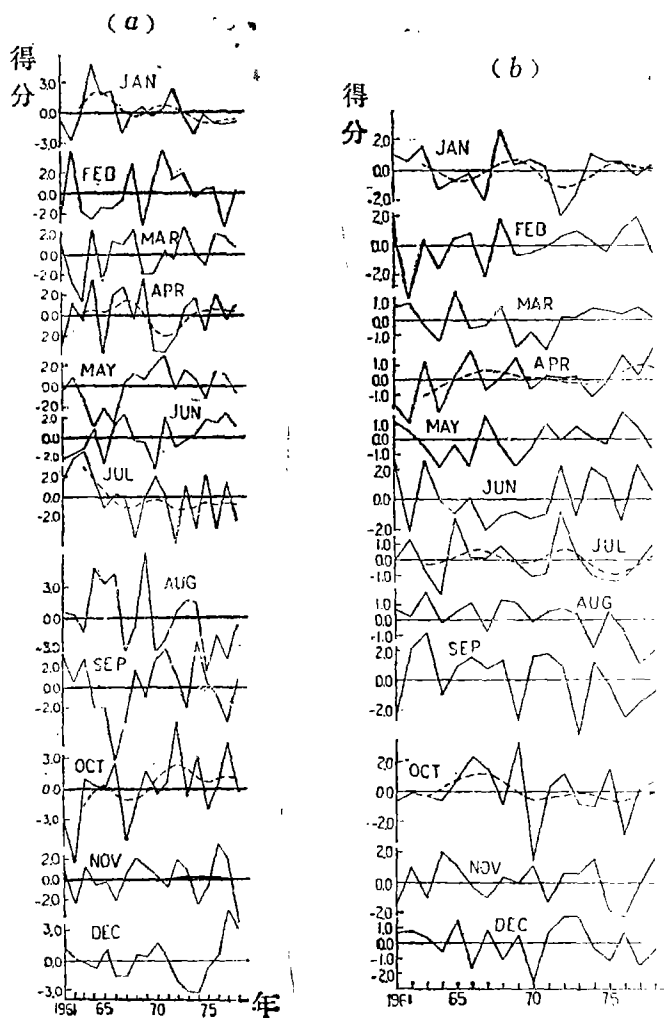


图4 第1主成分(a)第2主成分(b)的得分的时间序列

Fig. 4 The scoring time-series of the first principal component (a) and the second principal component (b)

代是湿型时期。

(2) 关于年间降水分布型的长期出现特性 图5—a和图5—b分别表示第1主成分和第2主成分的得分的历年逐月时间序列。图5的虚线也是用神山的短周期消去法的5年移动加权平均公式计算的周期变动曲线。关于年间降水分布型的长期出现特性的一般情况,如图5表示那样,由于降水这个气候要素不同于气温,即它的月间降水量差别很

大。虽然在计算过程中对月平均降水量作了正规化（立方根变换）处理，但从图上仍可看到年间降水分布型长期出现波动较大的特点。不过年间降水分布型的变动趋势，仍有 3~4 年的短周期变化。

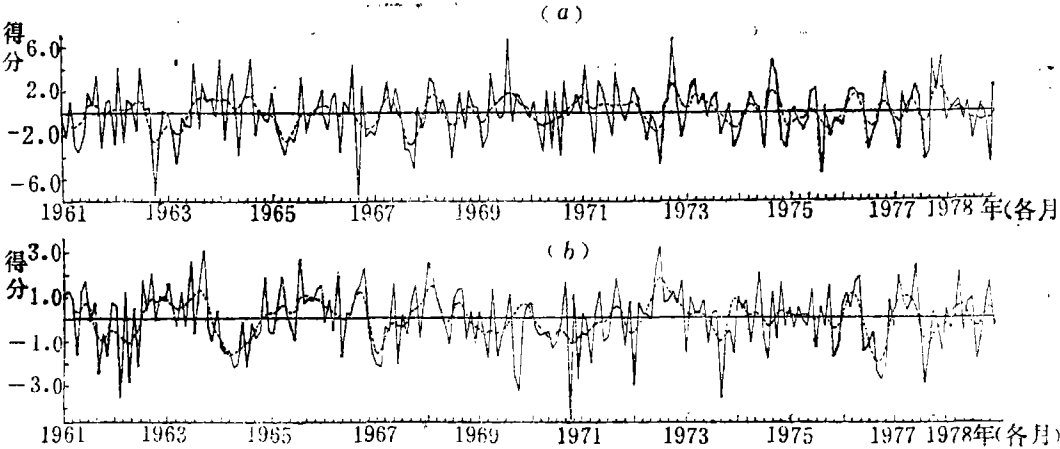


图 5 第 1 (a)、第 2 (b) 主成分得分的年间时间序列
Fig. 5 The yearly scoring time-series of the first (a) and the second (b) principal component

三、主成分聚类分析的气候分区

表 2 是依据前述的降水主成分分析结果和公式 $(f_{ij} = \sqrt{\lambda_i} \cdot l_{ij} / \sqrt{S_j})$ ，计算出各

表 2 各地点降水时间变动的特性值
Table 2 Eigenvalue of changing time of precipitation various stations

因子主成分 站号	降 水				因子主成分 站号	降 水			
	第 1 主成分	第 2 主成分	第 3 主成分	第 4 主成分		第 1 主成分	第 2 主成分	第 3 主成分	第 4 主成分
1	0.1209	0.3289	-0.1523	-0.1017	19	0.6133	-0.0235	0.3686	-0.2768
2	0.3142	0.4703	-0.2467	-0.1455	20	0.7379	0.1095	0.1603	0.0460
3	0.3832	0.2198	-0.2665	-0.2732	21	0.7062	0.0080	0.3378	-0.1510
4	0.3876	0.3851	-0.3785	-0.2262	22	0.8000	0.0286	0.1891	0.0364
5	0.3315	0.3628	-0.3008	-0.3100	23	0.5193	0.0402	-0.0626	0.5291
6	0.3654	0.5555	-0.3037	-0.3088	24	0.5735	-0.0402	-0.1383	0.3450
7	0.4927	0.5725	-0.2489	-0.2762	25	0.5684	-0.1893	0.4639	-0.2917
8	0.4724	0.4764	-0.2607	-0.2433	26	0.5151	-0.1834	0.3705	-0.2911
9	0.4473	0.4850	-0.0866	0.2462	27	0.7812	-0.2378	0.0234	-0.1090
10	0.5176	0.4420	-0.0447	-0.2596	28	0.4693	-0.2212	0.3647	-0.2445
11	0.4849	0.4886	0.1206	0.2926	29	0.7836	-0.2378	-0.1784	-0.0207
12	0.5578	0.4685	0.2370	0.0212	30	0.6614	-0.3746	0.2019	-0.1793
13	0.5666	0.1942	0.2061	0.5915	31	0.6558	-0.4714	-0.0125	-0.1703
14	0.6039	0.4017	0.1591	0.3138	32	0.6801	-0.3121	-0.3578	0.1383
15	0.6226	0.1803	0.2179	0.5472	33	0.6472	-0.4628	-0.2650	-0.0032
16	0.5229	0.1074	0.1767	0.6098	34	0.5331	-0.2026	-0.5501	0.2288
17	0.3823	0.2890	-0.1793	-0.3166	35	0.3437	-0.4573	-0.5334	0.0208
18	0.6390	0.2404	0.0657	-0.0833					

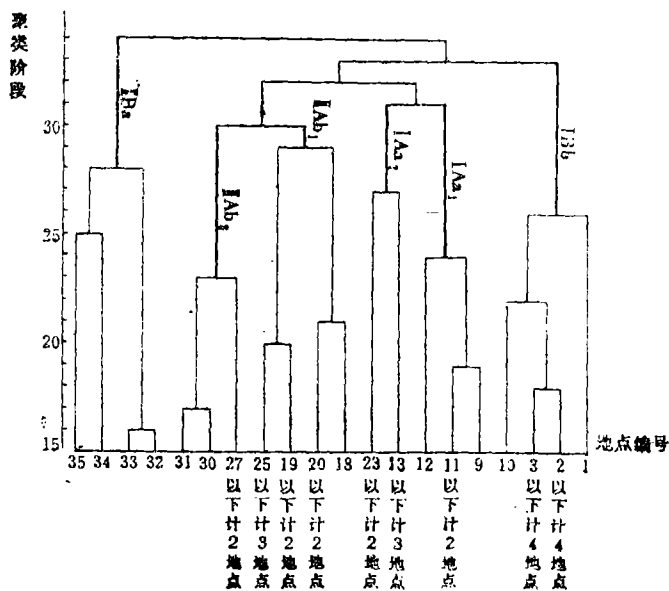


图 6 东北地区降水年变化型的树形图

Fig. 6 The dendrogram of the changing patterns of yearly precipitation of the Northeast China

主成分同各地点的降水时间变动特性值（因子负荷量），这是用主成分聚类分析进行气候分区的必要资料。用因子负荷量的值，由群平均法求出聚类间的类似度（或称距离），用类似度编制东北地区降水年变化型的树形图（图 6）。

图 6 的分区符号（如 IAa_1 等），I、II 是由降水的第 1、第 2 主成分的因子负荷量的符号，A、B 由降水的第 3 主成分的因子负荷量的符号，a、b 由降水的第 4 主成分的因子负荷量的符号决定的，下端的 1、2 是表示进一步分级的记号。这样一来，这个树形图便可以表明其聚类过程。

最后，由树形图进一步编制降水年变化型的气候分区图（图 7）。这就是依主成分聚类分析编制的气候分区图（本文是 4 级分区），各区的范围和特征归纳概括如表 3。

上述研究，客观地揭示了东北地区降水偏差的时间变动特性和气候分区。但是，有关东北地区主要气候要素的更长期的变动特性，以及综合气候要素的气候分区问题，仍需进行深入研究。

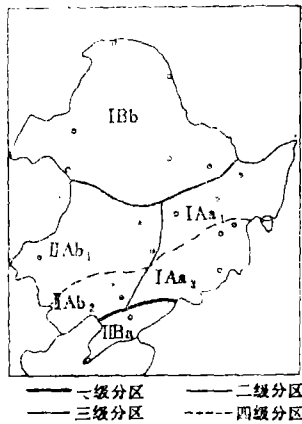


图 7 降水年变化型的气候分区

Fig. 7 The climatic regionalization of the changing patterns of yearly precipitation

表3 东北地区降水年变化型的各气候分区特征

Table 3 Characteristics of climatic regions on yearly precipitation changing patterns in Northeast China

分区名	范 围	代表 性站	因 子 负 荷 量				各 区 主 要 降 水 特性值 ^[6]
			第 1 主成分	第 2 主成分	第 3 主成分	第 4 主成分	
IBb	大、小兴安岭北端、呼伦贝尔高原以及松嫩平原西北部	呼玛	0.38	0.41	-0.24	-0.25	年平均降水量300~600mm 4月降水变率40~80% 7月降水变率20~40%
IAa ₁	松花江中段和三江平原	富锦	0.52	0.46	0.11	0.22	年平均降水量500~700mm 4月降水变率约50% 7月降水变率约30%
IAa ₂	长白山和张广才岭南端	鸡西	0.56	0.10	0.08	0.52	年平均降水量600~800mm 4月降水变率40~50% 7月降水变率30%以下
IAb ₁	大兴安岭南端和松辽分水岭	前郭尔罗斯	0.63	-0.03	0.29	-0.16	年平均降水量350~600mm 4月降水变率40~80% 7月降水变率30~40%
IAb ₂	辽西山地和辽河平原	彰武	0.72	-0.33	0.01	-0.12	年平均降水量400~700mm 4月降水变率40~60% 7月降水变率30~40%
IBa	辽东半岛和长白山南部	草河口	0.55	-0.36	-0.43	0.10	年平均降水量600~1000mm以上 4月降水变率50~60% 7月降水变率30~40%

参 考 文 献

- [1] 三上岳彦, 主成分による日本の夏季気温分布型の表現とその動気候学の考察, 《地理学评论》48—11, 784—780, 1975年。
- [2] 高桥润二郎, 地域区分のための主成分分析, 《三田学会志》60, 221—236, 1969年。
- [3] 加藤央之, 日照时间、日平均気温の变动形态から見た北海道の地域特性, 《地理学评论》56—1, 1—16, 1983年。
- [4] 日野藤雄, スペクトル解析, 184—236, 朝仓书店, 1977年。
- [5] 河口至商, 多变量解析入门Ⅰ, 35—51, 多变量解析入门Ⅱ, 26—44, 森北出版, 1973年。
- [6] 中央气象局, 中华人民共和国气候图集, 地图出版社, 1978年。

AN ANALYSIS OF PRINCIPAL COMPONENTS FOR DISTRIBUTIVE CHARACTERISTICS OF PRECIPITATION IN NORTHEAST CHINA

Zhao Huanchen

(Changchun Institute of Geography, Academia Sinica)

ABSTRACT

The basic patterns of the deviation distribution of annual precipitation in the Northeast China were defined objectively through the principal component analysis of monthly precipitation from January to December (1961—1978) in 35 meteorological observatories or stations of the Northeast. The eigenvalues and distributive features of precipitation deviation are presented in Table 1 and Fig. 1. The first, second and fourth principal components represented wet or dry patterns of the whole area, the pattern of north wet—south dry or north dry—south wet, the pattern of east wet—west dry or east dry—west wet, respectively. The third principal component represented different local precipitation deviation. The author made a dynamic investigation on the formations reasons and processes of dominant patterns of precipitation deviation (Fig. 2 and Fig. 3) based on the result of principal component analysis of monthly mean air pressure, and also approached the changing time characteristics of precipitation deviation and reasonably climatic changing differences in the Northeast China (Fig. 4 and Fig. 5) through the analysis of scoring time-series each month and each year. In order to reveal fully the precipitation differences within the region further, the climatic regionalization of precipitation distributive pattern was carried out with the principal component cluster analysis (Fig. 7) and the regional climatic characteristics were clear. Therefore the scientific basis could be provided for rational arrangement of agriculture, forestry, animal husbandry, production of by-product and fishery.