

# 我国东部夏季雨带的气候分类 及其环流特征

赵汉光

张先恭

(国家气象中心,北京 100081) (中国气象科学研究院,北京 100081)

## 提 要

利用主分量分析和K-均值聚类分析的方法,对我国东部夏季雨带进行客观的气候分类。分析表明:当聚类因子( $N=4$ ),聚类数( $K=2$ )时,第2主分量成分是在所有聚类因子中起决定性的因素,它不仅反映出两类不同雨带类型及其强弱的变化,而且在环流特征上具有显著性的差异。

**关键词:** 主分量 聚类因子 雨带类型

## 引 言

随着东亚夏季风的向北推进,我国东部夏季风雨带从低纬向中纬度的移动过程中,经历了三个相对停滞阶段和两次突然北跳过程。第一次雨带停滞阶段在华南一带,表现为华南的前汛期;第二次雨带停滞阶段在江淮一带,表现为江淮的梅雨期;第三次雨带停滞阶段在华北一带,表现为北方的雨季。夏季风雨带的两次突然北跳过程为:一次在6月中旬,雨带北跳到江淮一带;另一次在7月中旬,雨带北跳到华北一带。夏季风雨带在向北移动过程的年际变化有明显的差异,这种差异主要表现为历年夏季风雨带两次突然北跳过程在时间上有早有晚。所以夏季风雨带在各个相对停滞阶段的时间有所不同。从而导致了我国东部夏季雨带位置的空间分布有显著的差异。

在以往的长期预报业务工作中,对我国东部夏季主要雨带位置的气候分类<sup>[1]</sup>,一般是以地区分成3类:1类雨带位于长江流域及其以南地区;2类雨带位于淮河流域;3类雨带位于黄河流域及其以北地区。尽管这种雨带的气候分类方法比较简便,一目了然,但毕竟是定性的,而且在某种程度上带有一定的主观任意性。因此需要一种较为客观定量的分

类方法。为此,本文采用一种K-均值聚类分析的方法,对我国东部夏季雨带进行气候分类,并对不同类型的雨带位置所产生的环流背景作了分析。

## 1 资料和方法

本文在我国东部( $20^{\circ}$ — $45^{\circ}$ N,  $105^{\circ}$ — $122^{\circ}$ E)范围内,选取50个代表站,利用1951—1991年6—8月夏季降水量的资料。首先对我国东部50个站的夏季降水量场,采用主分量分析方法<sup>[2]</sup>,进行降水量场的展开,并取前5个特征向量场和前5个主分量成分。然后根据选取的主分量成分作为聚类因子,进行K-均值聚类分析。

现将K-均值聚类分析的基本思路<sup>[3]</sup>,作简要的概述:

设有M个样本  $X_i(x_{i1}, x_{i2}, \dots, x_{iN})$   $i=1, 2, 3, \dots, M$ , 其中N是表征事物特征的聚类因子数。

定义: $i, k$  两样本的欧氏距离为

$$D(X_i, X_k) = \sqrt{\sum_{j=1}^N w_j^2 (x_{ij} - x_{kj})^2}$$

其中  $w_j$  为权重系数,表示第j个因子在反映事物特征方面作用的大小。

把各类事物的均值规定为该类的“类中

心”即均值向量  $B_0$ , 对于  $L_k$  类有

$$B_k' = (b_{k1}, b_{k2}, \dots, b_{kN}) \quad k = 1, 2, \dots, K$$

其中  $b_{kj} = \frac{1}{m_k} \sum_{z \in L_k} z_j, j = 1, 2, \dots, N$ . 上式求和

是对属于  $L_k$  的所有事物  $z$  求的,  $m_k$  是  $L_k$  中

的样本数  $\sum_{k=1}^K m_k = M$ , 有了距离和类均值后,

规定一个样本与一个类之间距离为

$$D_L(X, L_k) = D(X, B_k)$$

则对于一种分类方案或分解  $P(X, L)$  分离总误差  $ED$  是

$$ED = E[P(X, L)] = \sum_{i=1}^M \{D_L[X_i, L_L(x_i)]\}^2$$

取定  $K$  类值并给出初始分解  $P_0(X, L)$  后, 根据距离最小原则不断调整样本分类, 使每个样本都处在与某类中心距离最短的类中, 分解总误差达到最小值, 便得到一定的  $K$  和  $N$  最优的分类结果。

把事物分类后, 进行  $F$  检验, 判定各类事物的差异是否显著, 分类是否适当。

## 2 降水量场的主分量分析

本文对我国东部夏季降水量场的展开, 采用主分量分析方法, 目的是寻找降水量场的空间函数模型, 同时把主分量成分作为划分雨带类型的聚类因子。在进行降水量场主分量分析之前, 对降水量场进行资料预处理有几种方案, 一是用实际降水量; 二是用距平值; 三是用标准化值。本文对三种方案分别进行了主分量分析并作了比较, 发现用实际降水量进行主分量分析后, 主分量成分与实际降水量的相关性最高, 特征向量场的空间模型与实际降水量场的空间分布类型最符合。所以, 本文采用第一种方案, 在50个主分量中仅前5个主分量已描述了该降水量场总方差的94.5%。

夏季降水场展开后的空间函数模型(即特征向量场)如下: 第一特征向量场(图略), 反映的是夏季平均降水量场的空间分布, 所以符号都是正的, 对雨带的气候分类并不起作用。我们感兴趣的是第二、三、四、五特征向量场的正负符号的分布。图1给出了第二、三特征向量场。从图中可见, 第二、三特征向量

场, 自南向北大体呈“+ - +”的空间分布, 但正负中心区的位置有差异, 反映了我国夏季降水量场往往是中间多(少)南北少(多), 呈相反的变化趋势。这种空间模型正是我国夏季降水分布的基本类型。第四特征向量场(图略)反映了我国夏季降水的空间分布南多(少)北少(多)符号相反的变化趋势。第五特征向量场(图略), 反映了我国夏季降水的空间分布东多(少)西少(多)符号相反的变化趋势。

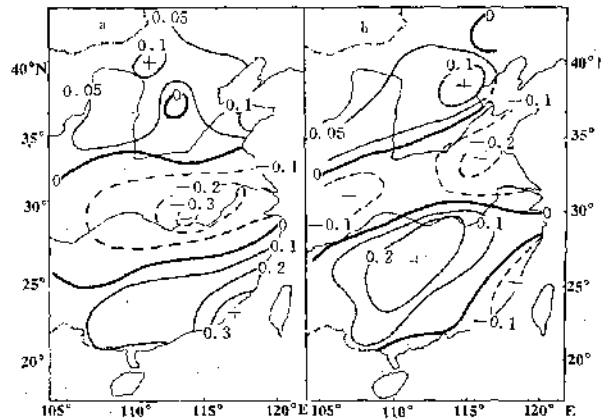


图1 第二、三特征向量场的分布图

从以上展开的几种夏季降水量场变化趋势的空间模型来看, 究竟哪种空间模型与夏季降水量的关系最密切呢? 图2a 为第二主分量成分与夏季降水量的相关场, 最大负相关中心区在江淮一带, 负相关系数最大为-0.72, 而华南及河套地区为正相关区, 其中华南一带最大正相关系数为0.62; 河套一带最大正相关系数为0.43。这种相关场呈“+ - +”的分布, 说明了我国夏季雨带位置的分布情况有两种, 当第二主分量成分为负值时, 雨带位于江淮一带, 而华南及河套一带少雨; 反之, 当第二主分量成分为正值时, 雨带分布情况恰好相反。图2b 为第三主分量成分与夏季降水量的相关场, 最大负相关中心在黄淮地区, 负相关系数最大为-0.52, 而相应的正相关区分别在江南和华北一带, 其中江南正相关系数最大为0.65, 华北的正相关系数最大为0.53。这种相关场的分布说明了我国夏季

雨带位置的空间分布,当第三主分量成分为负值时,主要雨带位于淮河流域一带,而江南及华北一带少雨;反之,当第三主分量成分为正值时,淮河流域少雨,而江南及华北一带分别为一条雨带,从图2a和图2b比较来看,虽然两者的正负相关区的空间分布都呈“+ - +”的分布,但在正负中心区的位置上有所不同,而且相关区的相关系数达到信度0.01(即 $r = \pm 0.40$ )的面积,前者远比后者大。第四主分量成分与夏季降水量的相关场(图略),符号大致呈北正南负的相关分布,反映了夏季降水趋势存在着两种情况:当第四主分量成分为正值时,北方为雨带,南方则少雨;当第四主分量成分为负值时,则恰好相反。这种雨带分布类型,从相关场的信度检验来看,南方的负相关区虽最大相关系数达到 $-0.70$ ,可是信度达到0.01的相关区( $r = -0.40$ )的面积却很小,仅位于江南东部沿海地区;同样北方的正相关区,相关系数最大仅0.42,相关系数达到信度0.01( $r = 0.40$ )的仅仅是个别地区。第五主分量与夏季降水量的相关场(图略),符号大致呈东正西负的相关空间分布,它反映了我国东西部之间的降水呈相反的变化趋势,但值得注意的是东部正相关区,最大相关系数仅0.32;而西部的负相关区,最大相关系数虽为 $-0.52$ ,但达到信度0.01( $r = -0.40$ )的相关区却很小,仅位于川北一带。

通过相关场的分析表明:我国东部夏季

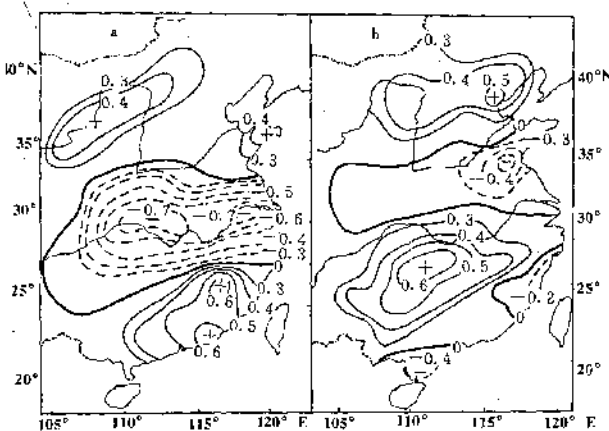


图2 第二、三主分量成分与夏季降水量的相关场分布

雨带位置的空间分布,以中间多(少)南北少(多)的雨带分布类型为主,而南多(少)北少(多)或东多(少)西少(多)的雨带分布类型较少。

### 3 雨带的气候分类

通过对我国东部夏季降水场的主分量展开后,获得了我国东部夏季降水的几种空间分布类型。下面我们将利用K-均值聚类的分析方法,以第2—5主分量的时间函数作为聚类因子,进行雨带的气候分类,对各主分量成分给予一定的权重系数。本文给出两种方案,一种按各主分量占总方差的贡献大小排列,分别以权重系数为 $1:0.5:0.36:0.36$ ;另一种是各主分量的权重系数均为1。

表1给出了聚类数 $K=2$ 和聚类因子数 $N=4$ 时,历年夏季雨带的分类结果,表中符号I为一类雨带分布类型,雨带位于黄河至长江流域之间,而黄河以北和江南至华南一带为少雨区,即中间多南北少的雨带分布类型。符号II为二类雨带分布类型,表示黄河至长江之间为少雨区,而黄河以北和江南至华南一带各有一条雨带,即中间少南北多的雨带分布类型。

表1 我国东部夏季雨带的气候分类(当 $K=2$ , $N=4$ 时)

年代	雨带类型	年代	雨带类型	年代	雨带类型	年代	雨带类型
1951	I	1962	I	1973	I	1984	I
1952	I	1963	I*	1974	I	1985	I
1953	I	1964	I	1975	I	1986	I*
1954	I	1965	I	1976	I	1987	I
1955	I	1966	I	1977	I	1988	I
1956	I	1967	I	1978	I	1989	I
1957	I	1968	I	1979	I	1990	II
1958	II	1969	I	1980	I	1991	I
1959	I	1970	I	1981	II		
1960	II	1971	II	1982	I		
1961	II	1972	II	1983	I		

\* 当相同权重系数为1, $K=2$ , $N \geq 3$ ,聚类结果为I类

从表1中可见:当各主分量给予不同或相同权重系数时,雨带分类结果大体相同,仅当聚类因子数 $N \geq 3$ 时,个别年份有所差异。但从分类的总误差结果分析表明:上述两种方案的总误差有明显的差异,当各主分量按总

方差贡献大小给予一定比例的权重系数时,分类总误差比较小,相反各主分量按相同的权重系数为1时,则分类总误差比较大。例如当聚类因子数  $N=4$  时,后者的总误差为前者的3倍;当  $N=3$  时,后者为前者的2.2倍;当  $N=2$  时,则后者为前者的1.6倍。由此可见,利用各主分量作聚类因子进行雨带的气候分类时,按各主分量占方差贡献的大小给予一定比例的权重系数进行分类为最佳方案。

其次,分析一下聚类数  $K=2$  时,各聚类因子的分类平均值及  $F$  值的情况。表2给出了 I 类和 II 类雨带类型,各主分量的分类平均值及  $F$  值的检验结果。

表2 聚类数  $K=2$  时,各聚类因子 ( $N=4$ ) 的分类平均值及  $F$  值

	相同权重系数			不同权重系数		
	I类 平均	II类 平均	$F$ 值	I类 平均	II类 平均	$F$ 值
$P_2$	-147.5	137.2	76.7	-164.4	125.5	79.5
$P_3$	-8.3	-0.5	0	-0.2	-6.9	0
$P_4$	18.3	-12.3	0.8	0.8	2.6	0
$P_5$	14.4	-15.9	0.9	7.4	-8.5	0.2
$F_{总}$			19.60			19.95

$P_2, P_3, P_4, P_5$  分别代表第二、第三、第四、第五主分量

从表2中可见:当各主分量给予相同或不同权重系数时, I 类和 II 类雨带类型之间,分类平均值差异最大的是第二主分量,其中 I 类平均值分别为  $-147.5\text{mm}$  和  $-164.4\text{mm}$ ; II 类平均值分别为  $137.2\text{mm}$  和  $125.5\text{mm}$ 。显著性检验结果,统计量  $F$  值分别为 76.7 和 79.5。这说明上述两种方案对第二主分量的分类平均值检验结果是显著的。而第三、四、五主分量的分类平均值之间的差异都很小,  $F$  检验结果不显著。因此,进行聚类数  $K=2$  时,第二主分量成分在所有的聚类因子数中起着决定性的作用。

最后,我们还分析了表1中的历年雨带类型与第二主分量成分之间的变化关系。从图3中可见,凡第二主分量成分为负的年份,聚类分析结果都是 I 类雨带类型;相反第二主分量为正的年份,都是 II 类雨带类型。值得注意的是:第二主分量成分的变化不仅能定性地指出雨带类型,而且还能反映出雨带类型的强弱。一般来说,当第二主分量成分绝对值愈大,反映两类雨带类型更典型,所以它具有明

显的指示性。

此外,谱分析表明,第二主分量成分具有 7—8 年的振荡周期。

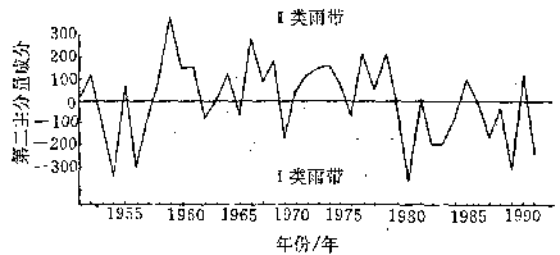


图3 雨带类型和第二主分量成分之间的关系

#### 4 两类雨带类型的环流特征

通过前面  $K$ -均值聚类分析的结果表明:当  $K=2$  分成两类雨带类型时,第二主分量成分的变化能较好的反映出我国东部夏季两种不同雨带类型及其强弱的变化。为了进一步分析两类不同雨带类型的环流特征,下面我们分别选取 I 类和 II 类的典型年份来分析。凡第二主分量成分  $\leq -150$  的作为 I 类雨带的典型年份,有 1954、1956、1969、1980、1982、1983、1987、1989、1991 年共 9 年;凡第二主分量成分  $\geq 150$  的作为 II 类雨带的典型年份,有 1959、1960、1961、1966、1968、1972、1973、1976、1978 年共 9 年。图4给出了夏季 I 类雨带减 II 类雨带典型年份的 500hPa 高度距平差值图。由图中可见:从东亚低纬经日本海到鄂霍次克海直至美洲东岸,出现的正负距平分布呈一串波列状结构。这种距平分布结构表明:当我国东部夏季出现典型的 I 类雨带分布时,其环流特征在东亚上空表现为“+ +”的距平场分布结构。相反当出现 II 类雨带分布时,则表现为“- + -”的距平场分布结构。那末,上述两类不同雨带类型在环流距平场上的特征是否具有显著性差异呢?我们首先分别计算了①鄂霍次克海阻隔区 ( $55-65^\circ\text{N}$ ,  $130-150^\circ\text{E}$ ), ②日本海低槽区 ( $35-40^\circ\text{N}$ ,  $125-140^\circ\text{E}$ ) 和 ③副热带高压区 ( $15-25^\circ\text{N}$ ,  $110-130^\circ\text{E}$ ) 夏季 500hPa 高度距平和,然后用①区减②区再加③区的高度距平,综合成一个东亚遥相关距平指数,并对综合后的东亚遥相关距平指数进行两类雨带典型年的

环流特征差异作  $t$  检验, 计算结果为:  $t = 2.82$ , 当  $\alpha = 0.05$  时,  $t_{\alpha} = 2.12$ , 所以  $t > t_{\alpha} (\alpha = 0.05)$ 。这表明上述两类雨带类型在环流距平场上的差异是显著的。

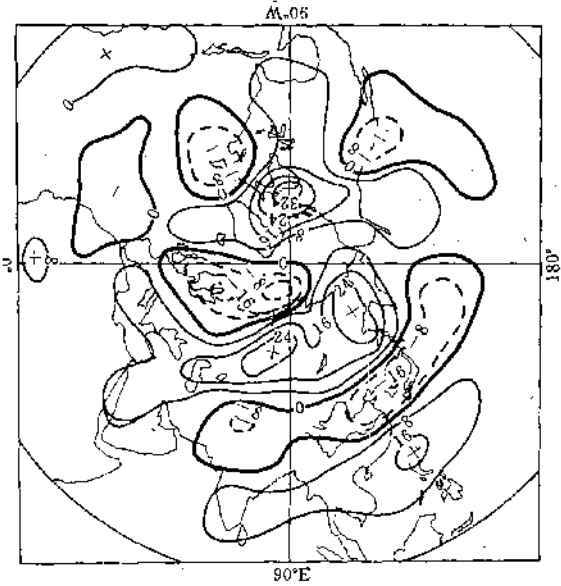


图4 I类雨带减II类雨带夏季500hPa高度距平差值图

上述分析仅仅表明两类雨带典型年份在环流特征上的差异是显著的。那末第二主分量成分的变化与环流变化之间的关系如何呢?为此,我们还计算了第二主分量成分和夏季500hPa高度场的相关(图略)。图中在东亚副热带地区为负相关区,相关系数可达-0.56,日本海附近为正相关区,相关系数为0.36,鄂霍次克海附近为负相关区,相关系数达-0.41。分析发现,相关图上的负相关中心区与图4中的正距平中心区的位置相对应。而相关图上的正相关中心区与图4中的负距平中心区的位置相对应。同样在东亚低纬经日本海到鄂霍次克海直至美洲东岸,形成一串正负相间的相关区,呈波列状的结构。这种东亚夏季遥相关的结构,与黄荣辉<sup>[4]</sup>利用准地转模式模拟出的夏季东亚遥相关结构十分相似。由此可见,第二主分量成分的变化,不仅反映了两类不同雨带类型及其强弱的变化,而且与环流有明显的遥相关关系。这种联系

实质上是反映了两类不同雨带类型的环流背景,即与东亚高纬是否有阻塞形势,中纬度西风带锋区位置以及低纬副热带高压位置三者之间的相互联系。当我国东部夏季出现I类雨带类型时,其环流特征往往表现为鄂霍次克海附近有阻塞高压维持,在距平场上呈现正距平分布,相应在日本海附近多低槽活动,西风带位置较常年偏南,在距平场上呈现为负距平分布。低纬副热带高压位置比常年偏南,赤道辐合带不活跃,所以反映在距平场上为正距平分布。相反当出现II类雨带类型时,其环流特征则表现为鄂霍次克海附近一般无阻高活动,而有低压活动,在距平场上呈现负距平的分布,与此同时,在日本海附近由于西风带位置偏北和副热带位置亦偏北,在距平场上出现正距平的分布,低纬地区由于副热带高压脊线位置偏北,相应赤道辐合带的位置亦较偏北,台风、热带风暴活动较频繁,所以在距平场上呈现为负距平分布。

## 5 小结

5.1 我国东部夏季主要雨带位置的空间分布类型,以中间多(少)南北少(多)的分布类型为主,而南多(少)北少(多)或东多(少)西少(多)的分布类型较少。

5.2 当聚类因子( $N=4$ ),分成两类( $K=2$ )时,第二主分量成分是所有聚类因子中起决定性的因素,它不仅反映出两类不同雨带类型,而且能表示出雨带类型的强弱。其中I类雨带类型是雨带位于江淮一带,而华南及河套一带为少雨区分布;II类雨带类型则恰好相反。

5.3 两类不同雨带类型存在着7—8年左右的振荡周期。

5.4 I类雨带类型其环流特征在东亚呈现为“+—+”的距平场分布;II类雨带类型则为“—+—”的距平场分布,两者之间具有显著性的差异。

## 参考文献

- 1 廖益荪等,北半球西风带环流和我国夏季降水,长期天气预报文集,北京:气象出版社,1981:103—114.
- 2 黄嘉佑,气象统计分析与预报方法,北京:气象出版社,

- 1990:170—188.
- 3 郑祖光. 变 K 变 N 聚类方案在气象场分析中的应用. 气象, 1984, 10(2): 7—11.
- 4 黄荣辉. 影响我国短期气候变化的北半球遥相关现象及其物理机制. 长期天气预报论文集, 北京: 气象出版社, 1990: 38—51.
- 5 谭冠日. 气象站数理统计预报方法, 北京: 科学出版社, 1980.

## The Climatic Classification of Summer Rain Belt in Eastern China and Their Circulation Features

Zhao Hanguang

Zhang Xiangong

(National Meteorological Center)

(Chinese Academy of Meteorological Sciences)

### Abstract

Utilizing the principal component and K-mean cluster analysis, the summer rain belts in eastern China have been classified. The results show that the second principal component plays an important role in all cluster factors when the cluster factor of N is equal to 4 and the cluster number of K is equal to 2. It can not only identify two different rain belt types and their strength but also have significant difference in their circulation features.

**Key Words:** principal component cluster factor rain belt type