

用频率分布数字特征对中国近海水文气候区划的验证*

陈上及 李炳兰 姚澍予 杜兵

(国家海洋信息中心, 天津 300171)

提 要 利用国家海洋信息中心国内海洋文档资料(1953—1988)和日本气象厅《东中国海表层月平均水温 30 年(1953—1982)报》等资料,将中国近海按 153 个统计区统计了水、气温频率分布的数字特征值,经主因子分析,求得累积贡献率达 95.29% 的 3 个主要因子。再用模糊聚类软划分方法将中国近海划分为 3 个气候带和 10 个气候区,本方法所取特征值能完整地反映数列的频率分布特征,避免了选取气候特征值时的主观随意性。

关键词 中国近海 海洋水文气候区划 频率分布数字特征 主因子分析 模糊聚类软划分

在海洋水文气候区划中,影响区划结果可靠性的一个重要问题,是如何选取主要影响因子。通常是取对水文气候最有影响的各要素气候特征值作为区划的主要因子。Chen Shangji 等 (1994)曾取 19 个海洋水文要素的特征值进行了中国近海区划,取得了满意的结果。但这样选取气候特征值却有一定的主观随意性。由此所得的气候区划结果,其代表性和可靠性究竟如何?尚须作进一步论证。本文从数列频率分布数字特征出发,以各水文要素的均值、均方差(σ)、离差系数(C_v)、偏态系数(C_s)和峰度系数(C_k)为基本变量,进行主因子分析和模糊聚类软划分,这可反映数列频率分布的总体特征,避免入选因子的主观随意性。

1 水文要素频率分布的数字特征

本文由国家海洋信息中心的海洋资料文档(1952—1988)¹⁾和日本长崎海洋气象台《东中国海表层月平均水温 30 年(1953—1982)报》等资料中,选用水气温的实测资料,进行频率分布数字特征的计算,其他要素因资料不足,未予考虑。

1.1 水温频率分布的数字特征 表层水温多年年平均值(\bar{T})的分布,反映海洋受热总量的分布格局。如图 1 所示,中国近海年平均水温的分布,主要受纬度和黑潮暖流的影响,南高(约 29℃)北低(约 12℃),随纬度向北递减。在同一纬度上相比,东海以黑潮流域为最高,高温脊恰与黑潮流轴相吻合,在其左侧,水平梯度较大,等温线与岸线平行;黄海高温脊与黄海暖流流轴相吻合。南海水温终年较高,表层水温年均值在 25℃ 以上,随纬度向南递增。总之,水温年均值的分布反映了中国近海热量分布的格局。

均方差是表征数列离散度的统计量。月平均表层水温均方差的分布系反映水温年较

* 国家自然科学基金资助项目,49070249 号。陈上及,男,出生于 1924 年 11 月,研究员。

收稿日期: 1994 年 1 月 31 日,接受日期: 1994 年 7 月 6 日。

1) 来自国家海洋信息中心数据库南森站资料文档。

差的区域特征。如图 2a 所示,表层水温均方差由北向南递减(由 9°C 至 1°C),表明水温年较差向南减小,尤其在黑潮流轴上,水温均方差显著小于两侧(尤其是左侧),说明黑潮流轴水温比左侧稳定得多,使等均方差线呈 SW—NE 走向,平行于岸线。

离差系数 (C_v) 是均方差与均值的比值。水温 C_v 的分布比 $\sigma_{(T)}$ 更能揭示水温季节变化的区域特征。比较图 2b 与图 2a 可知, C_v 的区域特征比 $\sigma_{(T)}$ 更为明朗,在黑潮流域以北, C_v 值向北向岸剧增,渤海 C_v 可达 0.8, 年变化极大。而在西沙群岛—东沙群岛以南的南海, C_v 值很小(仅为 0.05), 而且很均匀,说明南海水温季节变化极小。

表层水温偏态系数

$$C_s \left(= \frac{\sum(T_i - \bar{T})^3}{(n - 3)\bar{T}^3 C_v^3} \right)$$

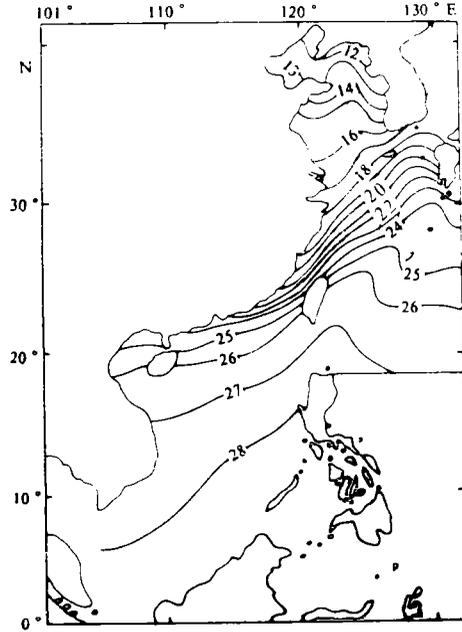


图 1 中国近海表层多年年平均水温 (\bar{T}) 的分布
Fig. 1 Distribution of annual mean sea surface temperature (\bar{T}) over the years in the China Seas

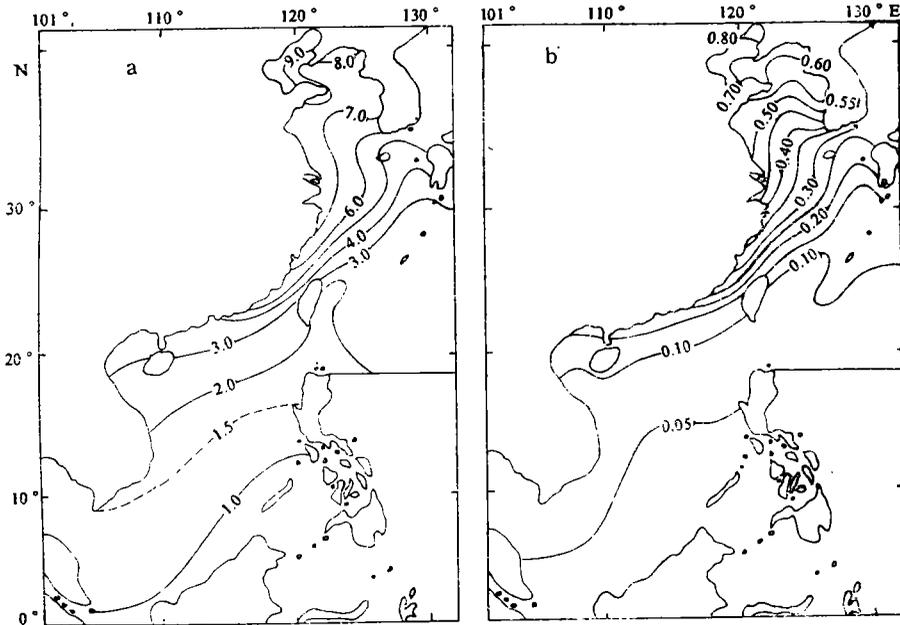


图 2 中国近海月平均表层水温均方差和离差系数的分布

Fig. 2 Distribution of variance (σ) of monthly mean and dispersion coefficient (C_v) of sea surface temperature in the China Seas

表示水温数列频率分布的偏度。 $C_s = 0$, 为正态分布; $C_s > 0$, 为正偏; $C_s < 0$, 为负偏。由图 3a 可知, $C_s = 0$ 线横贯台湾北端之东西, 将中国近海分成南、北两部分。由此以南的台湾海峡和南海为负偏区, 加里曼丹岛近海 C_s 达 -0.7 ; 由此以北的东、黄、渤海均为正偏区, 济州岛东南近海, $C_s \geq 0.25$ 。可见, 从偏态系数的分布来看, 南海与东、黄、渤海的水文气候特征是截然不同的。

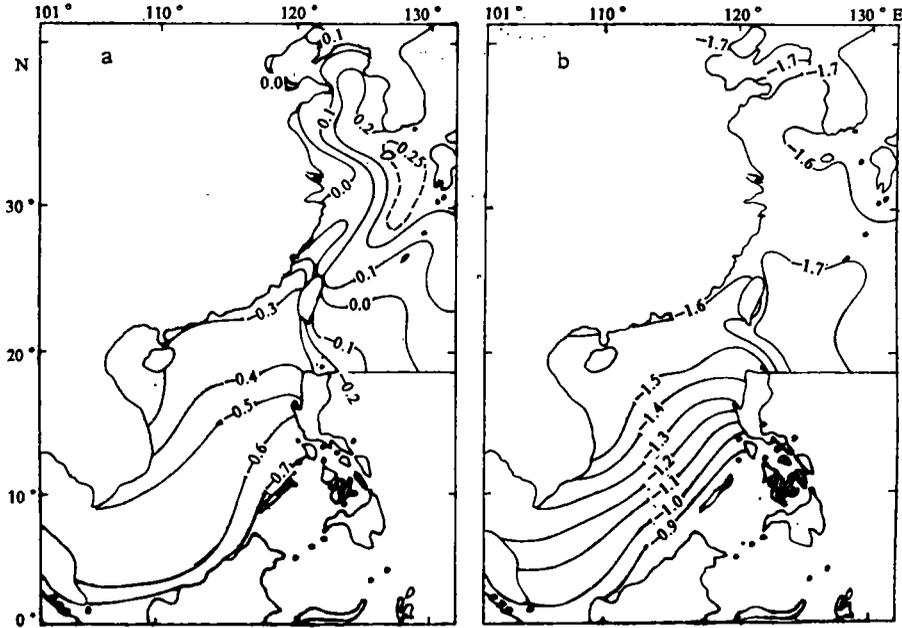


图 3 中国近海表层水温偏态系数 (a) 和峰度系数 (b)

Fig. 3 Distribution of deviation (a) and kurtosis coefficients (b) of sea surface temperature in the China Seas

频率分布的另一特征是以峰度系数 (C_k) 的大小为标志。 C_k 由下式求得 (金光炎, 1964):

$$C_k = \frac{\sum (T_i - \bar{T})^4}{(n-4)\bar{T}^4} - 3$$

$C_k = 0$, 为正态分布; $C_k > 0$, 为高峰态; $C_k < 0$, 为低峰态。由图 3b 可知, 中国近海表层水温的频率曲线均属低峰态, 以渤、黄、东海为最低, C_k 值均在 -1.7 至 -1.6 之间。南海则稍高, C_k 值向东南方向增高 (自 -1.6 至 -0.9), 这与南海水温年较差较小并向东南向递减相一致, 反映了南海的热带海洋气候特征。

1.2 气温数列数字特征的分布 中国近海海面气温数字特征的分布趋势大致与水温相似, 但有以下不同点: (1) 气温数字特征 $\bar{t}, \sigma(t), C_s(t)$, 分布的纬度差异较水温更为显著, 特征值等值线与纬度的交角较水温有所减小。(2) 黑潮对气温数字特征分布的影响, 不如对水温那么显著, 使黑潮左侧气温水平梯度大为减弱。(3) 气温的 $C_s = 0$ 线较水温 $C_s = 0$ 线显著北移, 使气温频率曲线正偏范围退缩至温州外海以北的东、黄、渤海, 由此以南的海区及南海均属负偏区。(4) 南海气温频率曲线的峰度普遍高于水温的峰度, 反映

了南海气温年较差略大于水温年较差的气候特征。

总之,上述数字特征的分布,集中反映了中国近海各海区的气候特征,为水文气候区划提供了主要依据。

2 数字特征的主因子分析

根据上述所得 153 个统计区的 $\bar{T}, C_v(T), C_i(T), C_e(T), \sigma(T)$ 和 $\bar{t}, C_v(t), C_i(t), C_e(t), \sigma(t)$ 值建立相关矩阵 R , 用 Jacobi 方法解得 R 阵的特征值及相应的特征向量。由表 1 可知, 只需取前 3 个主要综合因子, 其累积方差贡献率达 95.29%, 误差率仅 4.71%, 已足以反映中国近海水文气候区分布格局的基本特征。至于这 3 个主因子 F_1, F_2 和 F_3 的物理意义, 主要取决于各变量因子载荷的大小。表 1 指出, 第一个主因子 F_1 的因子载荷均在 $|0.187|$ 与 $|0.352|$ 之间, 彼此十分接近, 很难判断其主次关系。第二、三个主因子 F_2, F_3 的各变量因子载荷, 也有类似情况, 相互差别不大。要突出它们间的主次关系, 必须使因子载荷向 0 与 1 两极分化, 故须对因子载荷矩阵 $A_{p \times m} = (a_{ij})$ 作方差最大正交旋转(王学仁, 1982), 这里 $m = 3, p = 10$ 。经过第 5 次循环的正交旋转后, 因子载荷平方的方差 V_i 达最大 ($V_i = 0.246\ 403\ 6$), 此时, 所得矩阵各列因子载荷已向两极分化。正交旋转结果如表 2 所示, 第一个主因子 F_1 中, $C_v(T), C_v(t), \sigma(T)$ 和 \bar{T} 的因子载荷绝对值均 ≥ 0.9176 , 尤其是 $C_v(T), C_v(t)$ 对 F_1 起着主导作用, 说明 F_1 是综合反映水、气温年较差和年均值的主要因子。同样可由表 2 看出 F_2, F_3 的物理意义, F_2 是综合反映气温、水温频率峰度特征的主要因子, 其 $C_i(t)$ 和 $C_e(T)$ 分别为 0.9689 和 0.6775; F_3 是综合反映水、气温频率偏态系数特征的主要因子, $C_i(T)$ 和 $C_e(t)$ 分别为 0.8680 和 0.7289。

表 1 各主因子的特征值和特征向量

Tab. 1 Eigenvalues and eigenvectors of principal factors

主因子			F_1	F_2	F_3
特征值 (λ_i)			7.8891	1.1637	0.5296
贡献率(%)			78.45	11.57	5.27
累积贡献率(%)			78.45	90.02	95.29
累积误差率(%)			21.55	9.98	4.71
特征向量(因子载荷)	气 温	\bar{t}	0.3492	-0.1553	-0.0010
		$C_v(t)$	-0.3411	0.2083	-0.1513
		$C_i(t)$	-0.3135	-0.0054	0.4276
		$C_e(t)$	0.1870	0.6591	0.5760
		$\sigma(t)$	-0.3522	0.0437	-0.0078
	水 温	\bar{T}	0.3448	-0.1954	0.0736
		$C_v(T)$	-0.3321	0.2607	-0.2629
		$C_i(T)$	-0.2947	-0.2932	0.5839
		$C_e(T)$	0.2668	0.5257	-0.1160
		$\sigma(T)$	-0.3424	0.1664	-0.1832

3 水文气候区划

由上述结果可知, 取 F_1, F_2, F_3 3 个主因子, 就能可靠地进行水文气候区划(误差

表 2 经方差最大正交旋转后的因子载荷矩阵

Tab. 2 Matrix of factor loading after varimax orthogonal rotation

主 因 子			F_1	F_2	F_3
各 变 量 的 因 子 载 荷	气 温	\bar{i}	0.888 2	0.179 6	-0.422 9
		$C_v(i)$	-0.934 1	-0.186 9	0.304 0
		$C_s(i)$	-0.672 6	-0.127 5	0.728 9
		$C_e(i)$	0.182 2	0.968 9	-0.167 1
		$\sigma(i)$	-0.838 8	-0.282 2	0.465 6
	水 温	\bar{T}	0.917 6	0.170 7	-0.359 0
		$C_v(T)$	-0.962 5	-0.176 8	0.205 6
		$C_s(T)$	-0.401 9	-0.291 6	0.868 0
$C_e(T)$		0.313 9	0.677 5	-0.665 2	
	$\sigma(T)$	-0.922 8	-0.237 1	0.303 6	

率 $\leq 4.71\%$)。算得这 3 个主因子在中国近海 153 个统计点的得分数作为初始矩阵,进行模糊聚类软划分。取最佳分类数 $C = 10$,模糊参数 $a = 1.8$,经过 68 次迭代计算,求得最佳划分矩阵

$$U = \begin{pmatrix} 0.003 & 0.007 \cdots 0.005 & 0.008 \\ 0.062 & 0.434 \cdots 0.001 & 0.001 \\ \vdots & \vdots & \vdots \\ 0.001 & 0.004 \cdots 0.046 & 0.125 \\ 0.001 & 0.003 \cdots 0.931 & 0.834 \end{pmatrix}_{10 \times 153}$$

从矩阵 U 各列中挑选出隶属度 u_{ij} 最大元所在行的序号,即为该统计点所属气候区的区号。把相同区号的点标在同一平面上,即可绘出海洋水文气候区划图。如图 4 所示,中国近海可划分为下列 3 个气候带和 10 个气候区:

温带气候带

渤海和北黄海气候区 (I_1)

南黄海气候区 (I_2)

东海低盐混合水气候区 (I_3)

亚热带气候带

东海过渡性亚热带气候区 (II_1)

黑潮流域亚热带气候区 (II_2) 及南海北部近海亚热带气候区 (II'_2)

黑潮右侧冷、热涡亚热带气候区 (II_3)

热带气候带

南海中东部热带气候区 (III_1)

南海中西部热带气候区 (III_2)

南海南部赤道热带气候区 (III_3)

4 结果检验

将图 4 与陈上及等(1994)曾用气候特征值所得结果相比较可得如下结论。

4.1 应用水文要素气候特征值与应用频率分布数字特征进行中国近海水文气候区划, 两

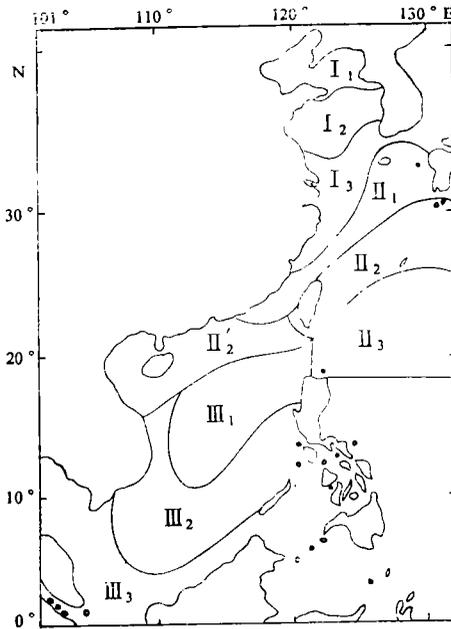


图4 根据数字特征对中国近海的水文气候区划

Fig. 4 Division of hydroclimatic area over the China Seas on the basis of frequency characteristics

者结果十分吻合。尤其是3个气候带的分界线,恰与水文季节划分(陈上及等,1992)中的“无冬线”和“全夏线”相吻合。证明所得中国近海水文气候区划是完全符合水文气候规律,反映其气候特征的。

4.2 对温带3个气候区(I₁,I₂,I₃)及亚热带4个气候区(II₁,II₂,II₂'和II₃)的划分,两种方法所得结果完全相同。只是对南海3个热带气候区(III₁,III₂及III₃)的划分稍有差别。这是由于南海水、气温分布较均匀,气候区划时必须同时考虑其它要素。但计算数字特征时,因别的要素资料不足,未能予以应有的考虑。尽管如此,两种方法的结果差别不大,区划较为合理。

4.3 从分析方法上来说,气候特征值法,直接选取各要素气候特征值作为统计区划的依据,所需资料较易搜集到,方法也较直观方便。但从频率统计意义上说,缺乏频率分布特征的总体概念,不免有主观随意性。本文所用的频率分布数字特征方法,充分反映了频率分布的总体概念,较为合理。虽对资料容量要求过高,不易普遍应用,可作为结果的验证。根据上述两种方法的对比验证结果,两者颇为符合,说明中国近海水文气候分为3个气候带和10个气候区是符合实情的,能充分反映中国近海的水文气候特征,是合理可靠的。

参 考 文 献

- 王学仁,1982,地质数据多变量统计分析,科学出版社(北京),269—283。
 陈上及、何维焕、姚混予等,1992,中国近海海洋水文气候季节的划分,海洋学报,14(6): 1—11。

金光炎,1964,水文统计原理和方法,中国工业出版社(北京),13—46。

长崎海洋气象台,1985,東シナ海の海面水温30年報(1953—1982),気象庁技術報告,第1報。

Chen Shangji, Yao Shiyu, 1994, Division of hydroclimatic area over China seas I. Principal component analysis, *Acta Oceanol. Sin.*, **13** (1):31—46.

Chen Shangji, Yao Shiyu, 1994, Division of hydroclimatic area over China seas II. Cluster analysis and fuzzy ISODATA, *Acta Oceanol. Sin.*, **13**(2):213—224.

A VERIFICATION ON DIVISION OF HYDROCLIMATIC AREA IN CHINA SEAS USING DIGITAL CHARACTERISTICS OF FREQUENCY DISTRIBUTION

Chen Shangji, Li Binglan, Yao Shiyu, Du Bing

(*National Marine Data and Information Service, S. O. A., Tianjin 300171*)

ABSTRACT

1953—1988 oceanographic data files on the China Seas prepared by the National Marine Data and Information Service, S. O. A., China and the “30-years (1953—1982) Reports of Sea Surface Monthly Mean Temperature in the East China Sea by the Meteorological Agency, Japan”, were used to calculate the digital characteristics of frequency distribution of sea and air temperature in 153 areas over the China Seas. Principal factor analysis and fuzzy cluster ISODATA were used to divide the China Seas hydroclimatic area into three climatic belts including ten climatic areas. It is concluded that the characteristic values derived by this method may completely reflect the characteristics of frequency distribution of sea and air temperature in the studied area and the resulting division of hydroclimatic area is fully coincident with the author's former result (Chen and Yao, 1994).

Key words China Seas Division of marine hydroclimatic area Digital characteristic of frequency distribution Principal factor analysis Fuzzy cluster ISODATA