

水团分析中的模糊数学方法*

杨冬梅 李凤岐 苏育嵩 于 迎

(青岛海洋大学海洋系, 266003)

提要 使用多年平均温盐资料,应用模糊数学的一系列方法,对渤、黄、东海整个海域逐月进行了水团划分与分析。用移动格域法和模糊积分法确定水团的聚类中心,依 Fuzzy-F 统计判定海区内水团的个数;用模糊 ISODATA 聚类法调整软划分矩阵,依模糊贴适度衡量水团之间的相近程度,用模糊熵讨论水团的不均匀程度,从而给出了在水团分析中系统地应用模糊数学方法的实例和计算流程。

关键词 水团 渤海黄海东海 模糊数学 模糊积分 模糊-F 判别

许多学者深入分析了中国近海的某些水团,如黄海冷水团^[10],黑潮水团^[11,13],对马暖流水和台湾暖流水等^[9]。变性水团概念的提出^[7],聚类分析和模糊数学方法应用于水团研究^[12,13],为浅海水团分析增添了新内容。用模糊数学方法讨论水团的核心、混合带与变性,以及软划分调整等,已取得一些成果^[2,4]。本文则进一步用移动格域法和模糊积分,较好地确定了水团聚类中心和初始软划分矩阵,用 Fuzzy-F 统计判断水团个数,继而对水团的性质和相互关系进行模糊分析,从而使水团分析的模糊数学方法形成了较完整的体系,并得以在电算机上快速而简便地实现。

一、模糊数学方法的应用

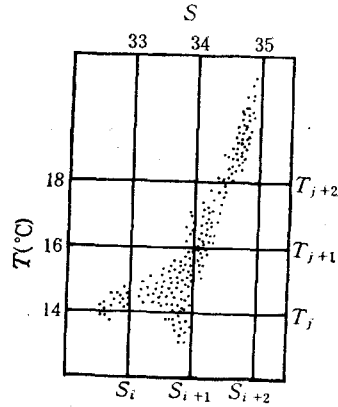
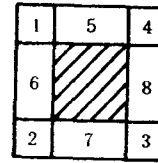
1. 水团聚类中心的确定——移动格域法和模糊积分

水团的核心是其特征最强的部分,亦即隶属度极大值的所在^[9]。因而用模糊积分法,依温盐及其均方差指标来确定初始聚类中心,是合理而可行的。移动格域法是作者为应用模糊积分而提出的准备。在 $t-S$ 平面上先划分格域(考虑本海区的特点,温度间隔取 2°C ,盐度取 1,见图 1),统计落入格域内温盐点的个数、温盐均值及均方差。对均值明显偏离格域中心(图 2 的影线区)者,将格域平移,以便使格域成为包络点比较密集的区域。点子较密集之处,成为水团核心的可能性也较大,测站及层次等距时更如此。

格域 i 的温盐均值及均方差分别为:

$$T_i = \frac{1}{m_i} \sum_{j=1}^{m_i} t_{ij}, S_i = \frac{1}{m_i} \sum_{j=1}^{m_i} s_{ij} \quad (1.1)$$

* 国家自然科学基金资助项目。
接受日期: 1990 年 11 月 17 日。

图1 t - S 图上的格域划分图2 格域 i 的放大Fig.1 Dividing lattice areas in the t - S diagramFig.2 The extension of lattice area i

$$\delta_i = \frac{1}{m_i} \sum_{j=1}^{m_i} [(t_{ij} - T_i)^2 + E(s_{ij} - S_i)^2]^{\frac{1}{2}} \quad (1.2)$$

式中, m_i 为格域 i 内温盐点的个数; E 为权重系数, 一般取为 1。当两个格域 i, j 的距离

$$d_{ij} = [(T_i - T_j)^2 + E(S_i - S_j)^2]^{\frac{1}{2}} \quad (1.3)$$

小于给定的 r (本文取 $r = 0.8$) 时, 则将 i, j 合并为 i' , 并重新计算格域 i' 内的均值和均方差。这样就增加了 i' 内的点数, 使 Fuzzy 积分值加大, 也符合“性质相近的水型集合为水团”的过程^[4]。经合并后所剩格域数记为 N' , 它们所对应的温盐均值, 成为水团聚类中心的可能性孰大孰小, 则可用模糊积分予以判断。

据文献[8],[11], 首先在可测空间上定义模糊测度函数 $g(b)$ 。对命题“(T_i, S_i) 是水团聚类中心”, $g(b)$ 表示 b 决定命题是否成立所起的重要性。考虑到均方差是描述点集疏密程度的重要参数, 盐度在本海区水团水系划分中也起重要作用^[4], 对比不同计算方案后, 取:

$$g(T_i) = 0.25, \quad g(S_i) = 0.35, \quad g(\delta_i) = 0.4$$

对于评价指标 b 的满意程度, 用可测函数 $h(b)$ 予以描述。据已有研究成果并结合本海区特点, 定义:

$$h(T_i) = |T_i - T| \cdot (T_{\max} - T_{\min})^{-1} \quad (1.4)$$

$$h(S_i) = |S_i - S| \cdot (S_{\max} - S_{\min})^{-1} \quad (1.5)$$

$$h(\delta_i) = \left(1 - \frac{\delta_i}{\delta_{\max}}\right) \left(\frac{2m_i}{M_{\max}} - \frac{m_i^2}{M_{\max}^2}\right) \quad (1.6)$$

式中, T, S 为所有测站的温、盐总平均值; 下标 \max 表示所有格域中对应指标的最大者, \min 为最小; M_{\max} 为各格域中包络点子的最大数。显见式(1.4), (1.5)可体现“外异性”, 而(1.6)可体现“内同性”^[4]。

对 $h(T_i), h(S_i), h(\delta_i)$ 依大小重排序, 得新顺序 b_1, b_2, b_3 , 满足 $h(b_1) \leq h(b_2) \leq h(b_3)$, 于是, 定义模糊积分:

$$f_B h(b) \circ g_\lambda(\cdot) = \bigvee_{i=1}^3 [h(b_i) \wedge g(B_i)] \quad (1.7)$$

式中, $B_i = \{b_1, b_2, \dots, b_i\}$ 。模糊分布函数 H , 记为 $H(b_i) = g(B_i)$, 其计算公式为:

$$H(b_1) = g(B_1) \quad (1.8)$$

$$H(b_i) = g(B_i) = \frac{1}{\lambda} \left[\prod_{b_j \in B_i} (1 + \lambda g_j) - 1 \right] \quad (1.9)$$

$$H(B) = 1, (B = \{b_1, b_2, b_3\}) \quad (1.10)$$

由式(8),(9)可得:

$$g_1 g_2 g_3 \lambda^2 + (g_1 g_2 + g_1 g_3 + g_2 g_3) \lambda + g_1 g_2 g_3 = 1 \quad (1.11)$$

解此方程可得 λ 值, 然后再由式(1.9)计算 H 。

分别计算 N' 个格域的模糊积分值, 积分值越大者, 成为水团聚类中心的可能性就越大。

2. 水团个数的判定——Fuzzy-F 统计及其应用

文献[2]曾用经典 F -统计判断水团的个数。依模糊集合观点, 每个水样对论域中所有水团各有不同程度的隶属^[4], 故用 Fuzzy-F 统计确定水团个数更为合理。

设水样 i 对水团 i 的隶属度为 u_{ij} , 则可构成隶属矩阵 $U = [u_{ij}]$ 。定义:

$$Q_1 = \sum_{i=1}^M \left(\sum_{l=1}^n u_{il}^2 \right) (V_i - V)(V_i - V)' \quad (1.12)$$

$$Q_2 = \sum_{i=1}^M \sum_{l=1}^n u_{il}^2 (X_l - V_i)(X_l - V_i)' \quad (1.13)$$

式中, M 为水团的个数; n 是样本的总个数; A 是控制水团模糊程度的参数; V_i 是第 i 个水团的聚类中心, 写成矢量形式:

$$V_i = (z_i, S_i, \dots) \quad (1.14)$$

V 是全部水样性质的总平均值, 亦写为矢量:

$$V = (T, S, \dots) \quad (1.15)$$

X_l 为第 l 个水样的测值:

$$X_l = (z_l, S_l, \dots) \quad (1.16)$$

而 $(\dots)'$ 为 (\dots) 的转置。显然, Q_1 描述外异性, Q_2 体现内同性。于是, 可构造 Fuzzy-F 统计量

$$FF = \frac{Q_1(n-M)}{Q_2(M-1)} \quad (1.17)$$

FF 服从 $F[P(M-1), P(n-M)]$ 分布。其中 P 为式(1.16)所示矢量的维数, 当只用温盐资料时, $P = 2$ 。

在文献[3],[6]中, 将 Fuzzy-F 统计称为 Fuzzy 伪 F 统计, 并定义

$$FF = \frac{\text{tr}(Q_1)(n-M)}{\text{tr}(Q_2)(M-1)} \quad (1.18)$$

我们证实: (1)经典 F -统计是 Fuzzy-F 统计的一个特例, 故后者根本不必冠以“伪”字; (2)式(1.17)与(1.18)计算结果一致, 而前者计算量小得多(详细推证待另文发表)。

于是,在已知 M 个聚类中心时,依 ISODATA 方法(见后文3.)确定隶属矩阵 U ,并依式(1.12),(1.13)计算 Q_1, Q_2 , 求出对应于个数 M 的 FF 值 $FF(M)$ 。因渤、黄、东海水团个数不小于 5^{11-5} ,故令分组数从 5 浮动到 N ,同时计算各对应的 FF 值。若

$$FF(M') = \max_{5 \leq M \leq N} \{FF(M)\},$$

则 M' 为合理的水团个数。

3. 水团划分的调整——模糊 ISODATA 聚类法

由模糊积分和 Fuzzy- F 统计确定 M' 个初始聚类中心后,可直接依下述步骤迭代调整:

(1) 初始迭代次数 l 置 0,取 $\varepsilon = 0.05$,

(2) 计算隶属矩阵 $U^{(l)} = [u_{ij}^{(l)}]$

$$u_{ij}^{(l)} = \left[\sum_{k=1}^{M'} \left(\frac{X_j - V_i^{(l)}}{X_j - V_k^{(l)}} \right)^{\frac{2}{A-1}} \right]^{-1} \quad (1.19)$$

式中, A, X, V 的含意同前,角标 (l) 表示迭代次数。若恰遇 $X_j = V_k^{(l)}$, 即 X_j 完全属于水团 k , 为防计算溢出,可置 $\|X_j - V_k^{(l)}\|$ 以小值(如 10^{-6})。

(3) 修改聚类中心 $V_i^{(l)}$:

$$V_i^{(l+1)} = \frac{\sum_{k=1}^n (u_{ik}^{(l)})^A X_k}{\sum_{k=1}^n (u_{ik}^{(l)})^A} \quad (1.20)$$

式中, n 为水样的总个数,其它符号见前。

(4) 计算范数 $\|\cdot\| = \|U^{(l+1)} - U^{(l)}\| = \max_{i,j} \|u_{ij}^{(l+1)} - u_{ij}^{(l)}\|$ 。若 $\|\cdot\| \geq \varepsilon$ 则将迭代次数加 1,返回步骤(3)。否则即停止迭代,得到调整后的 $U^{(l+1)}$ 和 $V_i^{(l+1)}$ 。

4. 水团性质及相互关系的分析——模糊贴近度及模糊熵的应用

贴近度可定量描述水团之间的相似程度^[4,13], 本文试算过三种贴近度,较好的是格贴近度

$$f(A, B) = \frac{1}{2} (A \cdot B + A \odot B) \quad (1.21)$$

其中 $A \cdot B$ 为内积,而 $A \odot B$ 为外积

$$A \cdot B = \bigvee_{k=1}^n (A(X_k) \wedge B(X_k)) \quad (1.22)$$

$$A \odot B = \bigwedge_{k=1}^n (A(X_k) \vee B(X_k)) \quad (1.23)$$

式中, $A(X_k), B(X_k)$ 是水样 X_k 对水团 A, B 的隶属度,可以从软划分矩阵 $U^{(l+1)}$ 中方便地摘出,故计算很方便。

模糊熵可用以描述水团的不均匀性和变性的程度^[13],其定义为:

$$H(A) = \frac{1}{n \ln 2} \sum_{k=1}^n S(A(X_k)) \quad (1.24)$$

式中, $S(x)$ 为仙农函数。由于从 $U^{(l+1)}$ 中能方便地取出 $d(X_k)$, 故亦可由电子计算机一并完成。

二、计算结果与讨论

1. 资料

在渤、黄、东海区内取经纬线整数度交点(图 3), 统计各月的多年温盐平均值¹⁾, 水深大于 200m 的站, 以 200m 层代底层。

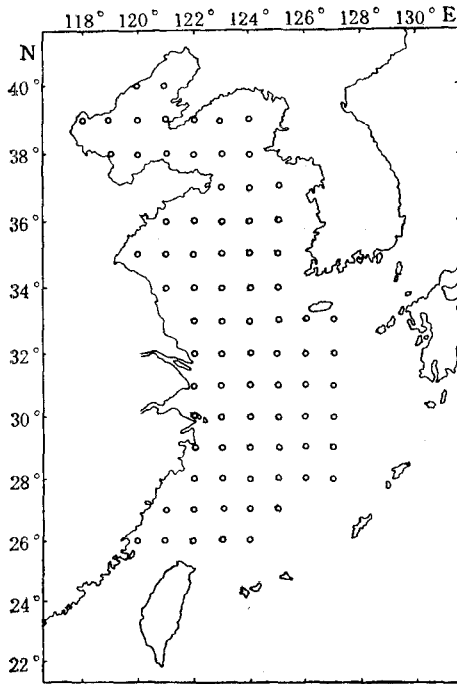


图 3 站位图

Fig.3 The distribution of stations

2. 水团的个数与名称

依 Fuzzy-F 统计判定水团个数, 如图 4 表明, 2 月份分 8 个水团统计意义最强。该月份表层各水团的分布见图 5。

本文用上述方法划分了各月的水团。限于篇幅, 为便于讨论, 只将名称和符号开列如下: 渤海水团 (B), 黄海沿岸水 (YS), 东海沿岸水 (ES), 径流冲淡水 (F), 黄海水

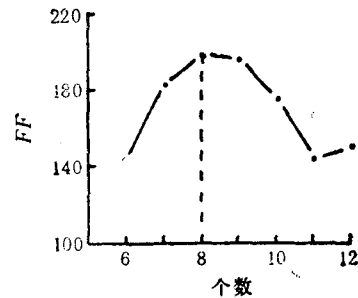


图 4 FF 值随水团个数的变化(2 月)

Fig.4 The change of FF with the number of water masses(Febr.)

团 (Y), 黄海冷水团 (YC), 东海表层水团 (E), 东海次表层水团 (EC), 黄-东海混合水团 (YE), 黑潮表层水 (K), 台湾暖流水 (TW) 和对马暖流水 (T)。

3. 水团核心的调整

经迭代后, 水团的聚类中心稍有变动, 平均调整幅度温度为 0.56°C , 盐度为 0.12。这说明本文的方法是有效可行的。尤其是对不太熟悉的海区, 用本文的方法可以简便而合理地确定水团的核心。变性较复杂的水团 Y, 在四季代表月的初始及调整后的聚类中心, 列于表 1 中。

4. 水团之间相互关系的分析

逐月计算了各水团间的格贴近度, 表 2 列出了 2 月份的结果。表中最大的贴近度是

1) 东海水产研究所刘树勋副研究员等提供。

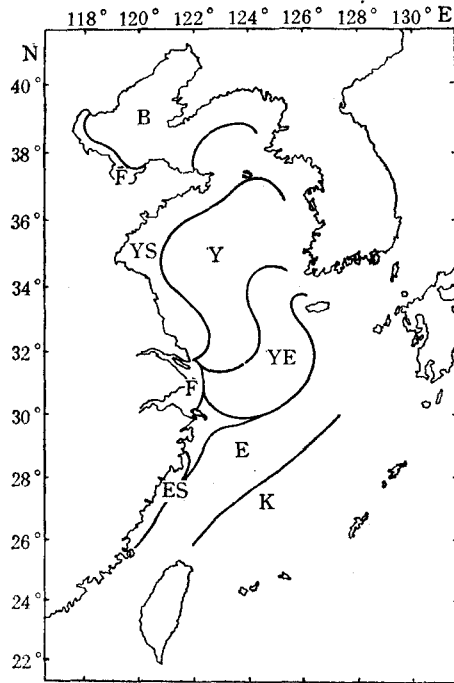


图 5 2 月份表层水团分布

Fig.5 The distribution of water masses at sea surface in Feb.

表 1 水团 Y 初始及调整后的聚类中心

Tab. 1 Original and adjusted cluster centres of water masses Y

月 份 指 标	2		5		8		11	
	$t(^{\circ}\text{C})$	S	$t(^{\circ}\text{C})$	S	$t(^{\circ}\text{C})$	S	$t(^{\circ}\text{C})$	S
初 始	9.33	32.84	14.58	32.50	27.47	31.17	16.43	31.35
调 整 后	9.30	32.35	14.63	32.71	26.03	31.23	16.36	31.79

表 2 2 月份各水团之间的贴近度

Tab. 2 The approach degrees between water masses in Feb.

水 团	K	B	YS	Y	F	YE	ES	E
1	1							
B	0.524	1						
YS	0.536	0.667	1					
Y	0.541	0.540	0.649	1				
F	0.541	0.559	0.554	0.548	1			
YE	0.541	0.528	0.536	0.696	0.550	1		
ES	0.541	0.544	0.544	0.563	0.599	0.640	1	
E	0.681	0.524	0.536	0.548	0.551	0.704	0.642	1

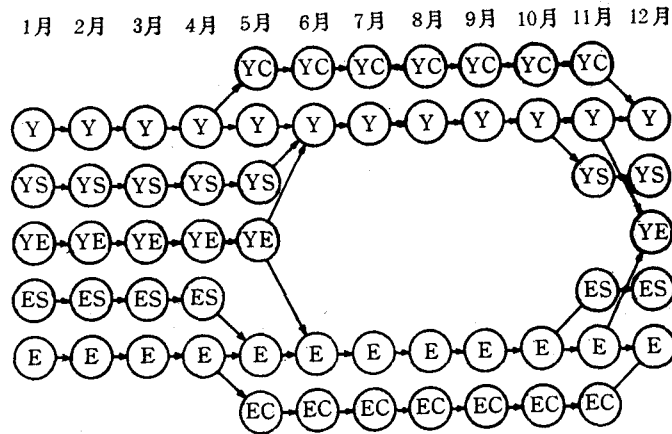


图 6 几个主要变性水团的消长过程及相互关系

Fig.6 Growing and decline process and mutual relationship among basic modified water masses

表 3 2 月份各水团的模糊熵

Tab. 3 Fuzzy entropies of water masses in Feb.

水团	YE	E	Y	YS	B	ES	K	F
熵	0.235	0.226	0.206	0.171	0.153	0.146	0.082	0.080

YE 与 E, 稍次是 YE 与 Y, 可见 YE 是名“符”其实的“黄-东海混合水团”。 $f(E, K)$, $f(B, YS)$ 位居三、四, 则体现了由 K 向 E 及由 YS 向 B 逐级变性的规律。 $f(K, B)$ 及 $f(E, B)$ 最小, 显然是地理位置之远和变性之甚所致。

5. 水团变性程度的讨论

模糊熵值的悬殊, 说明各水团的变性程度大有径庭。仍以 2 月为例(表 3), 高熵水团是 YE, E 和 Y, 这与它们所跨纬距较大、冬季降温变性强烈有关。K, F 熵值很低, 文献 [4] 已有讨论, 此略。

6. 主要变性水团的消长过程及相互关系

变性水团中 E 和 Y 是长年性水团, 其它是季节性水团, 即有生、消、分、合的过程。综合分析几个主要变性水团在全年的消长过程及相互关系, 可给出图 6。

参 考 文 献

[1] 毛汉礼等, 1964, 应用 T-S 关系定量地分析浅海水团的初步研究, 海洋与湖沼, 6(1): 1—22。
 [2] 王凤钦等, 1986, 浅海变性水团的软划分, 海洋学报, 8(4): 408—415。
 [3] 卢中发, 1989, 模糊目标函数聚类法及其在东海黑潮水团分析中的应用, 海洋学报, 11(3): 265—274。
 [4] 李凤岐等, 1986, 用模糊集合观点讨论水团的有关概念, 海洋与湖沼, 17(2): 102—110。
 [5] 苏育嵩, 1980, 划分变性水团边界的温盐点聚对照法与东海西部海区变性水团的分析, 海洋学报, 2(1): 1—15。
 [6] 张伟, 1987, Fuzzy 聚类的一个新算法, 模糊数学, 7(3—4): 51—56。
 [7] 汪培庄、韩立岩, 1989, 应用模糊数学, 北京经济学院出版社, 305—328。
 [8] 赵汝怀, 1981, (N) 模糊积分, 数学的研究与评论, 2: 55—72。

- [9] 翁学传,1984,台湾暖流水夏季 T-S 特征和来源的初步分析,海洋科学集刊,21: 113—132。
[10] 赫崇本等,1959,黄海冷水团的形成及其性质的初步探讨,海洋与湖沼,2(1): 11—15。
[11] 浅居喜代治等著(赵汝怀译),1988,模糊系统理论入门,北京师范大学出版社,173—186。
[12] Li Fengqi, Su Yusong and Yu Zuxiang, 1984, Application of cluster analysis method to modified water masses in the shallow sea, *Acta Oceanol. Sin.*, 3(4):451—461.
[13] Li Fengqi, Su Yusong and Fan Liqun, 1988, Application of method of fuzzy sets to the analysis of water masses in the northern South China Sea, *Acta Oceanol. Sin.*, 7(2):170—185.

FUZZY SETS METHODS IN ANALYSIS OF WATER MASSES

Yang Dongmei, Li Fengqi, Su Yusong and Yu Ying

(Dept. of Physical Oceanology and Marine Meteorology, Ocean University of Qingdao, 266003)

ABSTRACT

Several methods of Fuzzy Sets are applied systematically to the classification and analysis of water masses month by month in all the areas of Bohai Sea, Yellow Sea, and the East China Sea where temperature and salinity values which are averaged every month. The cluster centres of water masses are determined by method of moving lattice areas and Fuzzy Integral; the number of water masses in studied sea area is judged with the Fuzzy-*F* discrimination; the elastic classification matrix is adjusted using the method of fuzzy ISODATA; approximate degrees are weighed with the fuzzy approach degrees between water masses; inhomogeneity of every water mass is discussed using fuzzy entropy. Technical process of computation is presented in this paper, an example applied systematically the Fuzzy Sets method to analysis of water masses is given.

Key words Water mass, Bohai Sea, Yellow Sea, the East China Sea, Fuzzy Sets, Fuzzy Integral, Fuzzy-*F* discrimination.