

沿岸海水温、盐度序列正态性检验

卢婉娴 何国民

(中国水产科学研究院南海水产研究所, 广州 510300)

摘 要 本文应用偏峰度法和适合度法对南海北部涠洲、莺歌海、碇洲、闸坡、遮浪、云澳、万山等七个海洋站的温、盐度时间序列进行了正态性检验。检验项目包括月平均值、月极大值、月极小值、旬平均值、旬极大值、旬极小值的温、盐度多年距平序列, 对 84 个序列分别用两种方法检验, 共取得 168 个检验结果, 其中符合正态分布的序列不及三分之一。检验结果揭示了南海北部海水温、盐度的变化可能有潜在趋势或长周期振荡, 应作进一步研究。

关键词 正态性检验, 海水, 温度, 盐度

决定海水温、盐度变化的因素很多, 除了确定型的动力学和热力学因素以外, 还有随机因素。它们的随机特性在沿岸水域尤为突出。但是国内外学者对于海水温、盐度统计特性研究甚少。其实, 这项基础性工作不仅对海洋预报方法的选择有其实际应用价值, 同时在学术上有一定意义。

统计学对气象诸要素的统计特性研究得颇为充分。例如, 比较普遍的结论是: 气压、气温、年降水量等长期气候要素能近似地看成遵从正态分布, 而短期降水量、雪厚度、云量、风速、日照时数等都不能认为遵从正态分布^[4-11]。这些研究成果是确定气象要素的经验分布函数, 建立预报模型的重要依据。国内对定点水文要素预报研究得不多, 尤其对于沿岸水温、盐度的预报研究更少。因此, 逐步开展海洋站水文要素的统计特性研究是必要的。

本文介绍了南海北部七个海洋观测站温度、盐度月平均、旬平均、月极大、月极小、旬极大和旬极小值的距平值序列, 用正态适合度检验和正态偏度、峰度检验的结果。

检 验 方 法

序列正态分布具有概率密度函数

$$\Psi(x) = \frac{1}{2\pi\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

式中 x 为样本值, μ 为数学期望, σ 为标准差。样本 x 落入任意区间 (x_1, x_2) 的概率为

$$p(x_1 < x < x_2) = \int_{x_1}^{x_2} \psi(x) dx \quad (2)$$

收稿日期: 1992-06-24。

正态检验方法很多,有传统的偏、峰度检验、 X^2 平方检验和用正态概率纸的检验等^[1,2]。本文系选用正态偏度、峰度检验法和正态适合度检验法^[7]。

1. 正态偏度、峰度检验法

偏度指数为

$$g1 = \frac{m_3}{m_2^{3/2}} \quad (3)$$

峰度指数为

$$g2 = \frac{m_4}{m_2^2} - 3 \quad (4)$$

(3)、(4)两式中 m_k 为样本的 k 阶中心距

$$m_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^k \quad (5)$$

这里 $k=2,3,4$; x_i 为实测序列; n 为样本容量; $i=1,2,\dots,n-1,n$ 。

在样本容量很大的情况下,如果序列遵从正态分布,其偏度系数和峰度系数亦遵从正态分布^[5,8],而且数学期望为 0,均方差分别为

$$\sigma_1 = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}} \quad (6)$$

$$\sigma_2 = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)(n+3)(n+5)}} \quad (7)$$

假设序列遵从正态分布,取置信度 $\alpha=0.05$,当:

$$\begin{cases} |g1| > 1.95\sigma \\ |g2| > 1.96\sigma \end{cases} \quad (8)$$

时拒绝假设,否则假设成立。

2. 正态适合度检验

(2)式可改写为

$$p(x_1 < x < x_2) = \int_{-\infty}^{x_2} \phi(x) dx - \int_{-\infty}^{x_1} \phi(x) dx \quad (9)$$

令 $T = \frac{x-\mu}{\sigma}$, 则上式可写为

$$p(x_1 \leq x \leq x_2) = \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right) \quad (10)$$

其中

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt \quad (11)$$

如用样本均值 $\bar{x}$ 估计 μ ,用样本标准差 s 估计 σ ,则可算出任何 x 值域的理论频数(v)。把检验序列按大小分成 m 个组,并分别计算 $P(x_j)$ ($j=1,2,\dots,m-1,m$)。在实际计算中,(11)可查表代替。根据每组理论频数与实测频数之间的差异,由 F 检验确定序列是否符合

合正态分布。

在 F 检验中,由样本检验

$$H: F(x) = \Phi_0(x) \quad (12)$$

根据文献[1],当 $n \rightarrow \infty$ 时

$$X^2 = \sum_{j=1}^m \frac{(v_j - np_j)^2}{np_j} \quad (13)$$

的极限分布自由度为 $m-1$ 的 χ^2 分布,并与 $\Phi_0(x)$ 的分布类型无关。因此,只要算出上式 χ^2 的值,当

$$X^2 < \alpha \quad (14)$$

便可近似地认为序列服从正态分布。式中

$$\alpha = \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} \int_{\chi^2}^{\infty} e^{-\frac{u}{2}} \cdot u^{\frac{n}{2}-1} du \quad (15)$$

在实际计算中可查 χ^2 值表而得。另外,在计算理论频数时,由于要用样本的平均值和标准差估计总体的数学期望和方差,所以自由度应取 $m-2-1=3$ 。

数据及其预处理

样本数据取自国家海洋局海洋站逐日观测资料中的温度和盐度值。共取七个站(表 1)。

表 1 检验序列站名表
Table 1 List of Stations and Sequences

站 名(Station name) 代 码(Code name)	涠洲 A	莺歌海 B	碇洲 C	闸坡 D	遮浪 E	云澳 F	万山 G
起始日期(Start date)	60.1	62.2	60.1	60.1	60.1	60.1	73.9
终止日期(End date)	81.12	81.12	81.12	81.12	81.12	81.12	81.12
样本容量(N) 旬(For 10-day)	792	717	792	792	792	792	300
(Length of sequence) 月(For month)	264	239	264	264	264	264	100

由于海洋要素具有明显的年周期性,故在进行数据预处理时,旬序列和月序列均取多年平均值的距平值。

在进行适合度检验时,取 $m=6$ (表 2)。

检 验 结 果

按上述两种检验方法,应用微型计算机,分别对七个站的 12 个项目共 168 个序列作正态性检验,由广东沿海增养殖环境数据库支持运行(程序设计详见附件)。检验结果见表 3。

表 2 正态适合度检验序列分组情况

Table 2 Groups for Fitness Test

组号(Group)	水温(°C)(Temperature)	盐度(‰)(Salinity)
1	≤ 1.00	≤ -2.00
2	1.01—1.50	-2.01—1.00
3	1.51—2.00	-1.01—0.00
4	2.01—2.50	0.01—1.00
5	2.51—3.00	1.01—2.00
6	> 3.00	> 2.00

表 3 正态性检验结果

Table 3 Results of the Test

序列 (Sequence)	方法 (Method)	项 目 (Item)													
		水温(Temperature)							盐度(Salinity)						
		A	B	C	D	E	F	G	A	B	C	D	E	F	G
月平均值距平 ^J	适合度 ^H	×	○	×	×	○	×	○	○	×	×	×	×	×	×
	偏峰度 ^I	○	×	×	○	○	○	○	○	○	×	○	×	×	○
月极大值距平 ^K	适合度	×	×	×	×	×	○	○	×	×	×	○	×	○	×
	偏峰度	×	×	×	×	○	×	○	×	×	×	○	×	○	×
月极小值距平 ^L	适合度	×	×	○	○	×	×	○	×	×	×	×	×	×	×
	偏峰度	○	×	○	○	○	○	×	×	×	×	×	×	×	○
旬平均值距平 ^M	适合度	×	×	○	×	○	○	○	×	×	×	×	×	×	×
	偏峰度	×	×	○	○	○	○	×	×	×	×	×	×	×	×
旬极大值距平 ^N	适合度	×	×	×	×	×	○	○	×	×	×	×	×	×	×
	偏峰度	×	×	×	○	○	○	○	×	×	×	×	×	○	○
旬极小值距平 ^P	适合度	×	×	×	×	×	×	○	×	×	×	×	×	×	○
	偏峰度	○	×	○	○	○	×	○	○	×	×	×	×	×	○

“○”表示符合正态分布条件,“×”表示不符合正态分布条件。

H: Fitness method I: Offset peak method J: Anomalies of month means

K: Anomalies of month maximums L: Anomalies of month minimums

M: Anomalies of 10-day means N: Anomalies of 10-day maximums

P: Anomalies of 10-day minimums

检验结果表明,信度 $\alpha=0.05$ 水平时,海水盐度序列普遍不遵从正态分布。两种方法检验结果几乎完全一致。水温的检验结果与盐度有所不同。两种方法共得 84 个水温序列检验结果,符合正态分布者 40 个,小于总检个数的一半。从测站的情况来说,只有万山海洋站在 12 个检验结果中有 10 个符合正态分布条件,其余为不明显或明显非正态分布。从这个角度来说,被检验的六项水温距平序列,基本上是不符合正态分布的。但两种检验方法之间存在一定的差异,适合度检验结果,在 42 个总检个数中仅有 15 个结果符合正态分布,其余为不符合。用偏度、峰度检验法,其结果有 25 个符合正态分布,仅占 60%。可见,定点海水温度时间序列未经检验,不宜假设为正态过程。

讨 论

1. 两种不同的检验方法对水温的检验结果略有差异。从方法本身来看,适合度检验法中,组距划分问题有可能影响最后的结果,这有待进一步研究试验。偏度、峰度检验是传统的检验方法,在理论上较为成熟,结论比较可取。

2. 序列符合与不符合正态分布,在气象要素检验中都是常有的,主要是长周期变化和趋势项引起。因此,南海北部的温、盐度的变化可能有潜在趋势或长周期振荡,应作进一步研究。

参 考 文 献

- [1] 复旦大学数学系,1962。概率论与数理统计。上海科学技术出版社。
- [2] 中国科学院数学所概率组,1972。常用数理统计方法。科学出版社。
- [3] 克拉美,H.,1966。统计学的数学方法。上海科技出版社。
- [4] 徐尔瀛,1960。论年雨量的常态性。气象学报,21:17—34。
- [5] 王宗浩,李麦村,1974。天气预报中的概率统计方法。科学出版社。
- [6] 中央气象局气象科学研究所,1974。数值预报和数理统计预报文集。科学出版社。
- [7] 黄嘉佑,1984。气象统计预报试用教材。北京大学地球物理系气象专业。
- [8] 铃木荣一,1968。气象统计学。地人书馆。
- [9] Anderson, T. W., 1958. An Introduction to multivariate statistical analysis.
- [10] Seber, G. A. F., 1977. Linear regression analysis. John Wiley & Sons.
- [11] Dyer, T. G. J., 1977. On the application of some stochastic models to precipitation forecasting, Quart. J. R. Met. Soc. 103:177—189.

A NORMAL DISTRIBUTION TEST ON THE SEQUENCES OF SEAWATER TEMPERATURE AND SALINITY

Lu Wanxian He Guomin

(South China Sea Fishery Research Institute, Chinese Academy Of Fishery Sciences, Guangzhou 510300)

ABSTRACT The article is on using the offset peak method and the fitness method to make the normal distribution test on the sequences of seawater temperature and salinity. The data were collected by 7 observation stations of the northern South China Sea. The sequences tested were the anomalies of means and maximums as well as minimums of month and 10—day. Amount 168 results, tested by the two methods, less than 1/3 were in the conformity with the normal distribution. It revealed that a latent tendency or a long periodic vibration would exist in the changes of the seawater temperature and salinity of the northern South China Sea.

KEYWORDS Normal distribution test, Seawater, Temperature, Salinity

欢迎订阅《中国水产科学》

《中国水产科学》是由中国水产科学研究院主办的水产科学技术领域的学术性刊物,国内外公开发行。主要刊登水产系统和有关单位科技工作者在水产资源、海淡水捕捞、水产养殖与增殖、水产品保鲜与加工、渔业水域环境保护、渔船、渔业机械与仪器以及渔业基础研究的论文报告、研究简报、综述和学术动态等文稿。《中国水产科学》主要服务对象为水产科学研究、教学、科技管理人员以及大专院校师生。它面向水产业,为水产经济建设服务。

《中国水产科学》为季刊,16开本84页,每期定价8.00元(包括邮费),全年四期,共收费32.00元。欢迎广泛订阅。

欲订购者,可随时与《中国水产科学》编辑部联系办理手续。订费请通过银行信汇,开户银行为:北京市海淀区永定路分理处(工商行) 帐号:891195-90;或者通过邮局将订费汇寄 100039 北京市永定路南青塔村150号《中国水产科学》编辑部。