

# 功率谱分析中几个重要问题的讨论

王 强

(中国科学院大气物理研究所, 北京)

么 杭 生

(南京大学大气科学系)

**提要** 本文从理论上推导了谱分析中最大后延 $M$ 的取值, 并且从理论和实际计算中得出了样本长度和其隐含周期之间的关系。只要抽样是平稳的, 样本长度并非一定要求太长, 仅略为超过其隐含周期的长度即可。但谱分析绝非是万能的。如果样本中仅含有两个波动, 且当这两个波动的周期相差一个整数倍或更大时, 这两个波动才可以通过谱分析分辨出, 反之就会歪曲了其隐含周期。由此可见, 在具体应用时, 应作一些滤波和去掉趋势是必要的。因而这将对广大气象、地理科学工作者以及其它相邻学科的科研工作者具有重要的参考价值。

**关键词** 最大后延 样本序列长度 周期

## 一、引言

近四十年来, 谱分析方法已广泛应用于气象学、地理学的各个领域, 并且除预报外已取得很大成功, 谱分析技术已应用于湍流与风结构的研究(包括小气候)、大气环流、物理气候(包括大尺度湍流)、大气物理、气候学、卫星云图、雷达观测、天气动力学以及气候资料的研究与处理。这些详细介绍可参见 Bath 著作(1974)<sup>[1]</sup>。谱分析在我国也有很多著述<sup>[2]</sup>。

谱分析技术目前正在气象、气候问题的具体应用中不断得到改进和发展, 本文正是在这个宗旨下对下列几个重要问题进行探讨:

1. 如何从理论上确定最大后延 $M$ 的取值?
2. 所取样本长度与该样本中所隐含的周期究竟存在怎样的关系?
3. 谱分析所揭示的周期现象究竟有多大的可信度?

对上述问题所进行的探讨, 无疑将有助于更广泛地应用和发展谱分析方法。

## 二、从理论上探讨最大后延 $M$ 的取值

在功率谱计算中, 最大后延 $M$ 是人为的选择, 其取值不同, 可以使估计的功率谱曲线发生变化。 $M$ 的取值愈小时, 谱曲线的平滑效果就愈大。但若 $M$ 取值过小, 谱估计曲线十分光滑, 就不容易确定主要峰点, 峰点分辨不好会对寻找其对应的主要周期造成困难。 $M$ 的取值比较大时, 谱曲线本身的波动就会出现较多峰值, 而这些峰值并不意味着有对应的周期现象存在, 不过是由于平滑过程, 使真实谱的估计量出现偏差, 造成虚假峰值。

来稿日期: 1988年12月

上述仅是从计算中所求得的经验性直观说明,下面我们着手从谱分析的统计特性来佐证上述结论,使对最大后延 $M$ 的取值有其理论背景。

按统计学的观点,我们希望得到功率谱估计量的方差尽可能小,同时又希望它是无偏估计。

将功率谱估计量表示为:

$$I_N(f) = \sum_{k=-(N-1)}^{N-1} \hat{R}(k) \cos 2\pi k f \quad (1)$$

其中  $f$  为频率,  $N$  为样本长度,  $\hat{R}(k)$  表示如下:

$$\hat{R}(k) = \frac{1}{N} \sum_{i=1}^{N-|k|} x_i x_{i+|k|} \quad (2)$$

通常当  $|k| \rightarrow \infty$  时,  $\hat{R}(k) \rightarrow 0$ , 这样可把(1)式写为:

$$\hat{G}(f) = \sum_{k=-M}^M \hat{R}(k) \cos 2\pi k f \quad (3)$$

这里  $M$  为最大后延。(3)式实际上又可看作是  $\hat{R}(k)$  乘以矩形函数

$$q(k) = \begin{cases} 1 & |k| \leq M \\ 0 & |k| > M \end{cases} \quad (4)$$

以后再进行傅氏变换,(4)式所表示的函数称为“迟后窗”。

一般变量为复数的情况下,带有迟后窗的功率谱估计式写为:

$$\hat{G}(f) = \sum_{k=-M}^M q(k) \hat{R}(k) e^{-i2\pi f k} \quad (5)$$

利用卷积关系,(5)式可写成

$$\begin{aligned} \hat{G}(f) &= \int_{-1/2}^{1/2} \theta(f) I_N(f - \lambda) d\lambda \\ &= \int_{-1/2}^{1/2} \theta(f - \lambda) I_N(\lambda) d\lambda \end{aligned} \quad (6)$$

其中  $\theta(f)$  是  $q(k)$  的傅氏变换:

$$\theta(f) = \sum_{k=-M}^M q(k) e^{-i2\pi f k} \quad (7)$$

对(6)式求期望:

$$E[\hat{G}(f)] = \int_{-1/2}^{1/2} \theta(\lambda) E[I_N(f - \lambda)] d\lambda \quad (8)$$

因  $I_N(f)$  对于  $G(f)$  是渐近无偏的,故当  $N$  很大时,则渐近有:

$$E[\hat{G}(f)] = \int_{-1/2}^{1/2} \theta(\lambda) G(f - \lambda) d\lambda \quad (9)$$

如果谱窗  $\theta(f)$  足够窄,使得  $G(f)$  在比  $\theta(\lambda)$  的主瓣稍宽范围内变化很小视为常数,则近似为

$$E[\hat{G}(f)] = G(f) \int_{-1/2}^{1/2} \theta(\lambda) d\lambda = G(f) q(0) \quad (10)$$

显然,  $G(f)$  为  $\hat{G}(f)$  的无偏估计的必要条件是:

$$q(0) = \int_{-1/2}^{1/2} \theta(f) df = 1 \quad (11)$$

同时考虑(7)式和(11)式,  $M$  越大, (11)式越能成立, 因此为了使  $G(f)$  能满足无偏估计, 那么应该尽可能取  $M$  大些。

又  $\hat{G}(f)$  的协方差为:

$$\begin{aligned} \text{Cov}[\hat{G}(f_1), \hat{G}(f_2)] &= E[(\hat{G}(f_1) - E[\hat{G}(f_1)])(\hat{G}(f_2) - E[\hat{G}(f_2)])] \\ &= \iint_{-1/2}^{1/2} \theta(\lambda)\theta(\mu)\{E[I_N(f_1 - \lambda)I_N(f_2 - \mu)] \\ &\quad - E[I_N(f_1 - \lambda)]E[I_N(f_2 - \mu)]\}d\mu d\lambda \end{aligned}$$

上式最后{……}内可以证明是  $\text{Cov}[I_N(f_1 - \lambda), I_N(f_2 - \mu)]$ , 故得:

$$\text{Cov}[\hat{G}(f_1), \hat{G}(f_2)] = \iint_{-1/2}^{1/2} \theta(\lambda)\theta(\mu)\text{Cov}[I_N(f_1 - \lambda), I_N(f_2 - \mu)]d\mu d\lambda \quad (12)$$

假定观测数据是来自功率谱  $G(f) = \sigma_x^2$  的正态白噪声, 这时(12)式可以写成:

$$\begin{aligned} \text{Cov}[I_N(f_1 - \lambda), I_N(f_2 - \mu)] &= \frac{\sigma_x^4}{N} [2\pi F_N(f_1 + f_2 - \lambda - \mu) \\ &\quad + 2\pi F_N(f_1 - f_2 - \lambda + \mu)] \end{aligned} \quad (13)$$

其中 Fejer 核<sup>[3]</sup>为

$$\begin{aligned} F_N(f_1 + f_2 - \lambda - \mu) &= \frac{1}{2\pi N} \frac{\sin^2[\pi N(f_1 + f_2 - \lambda - \mu)]}{\sin^2[\pi(f_1 + f_2 - \lambda - \mu)]} \\ F_N(f_1 - f_2 - \lambda + \mu) &= \frac{1}{2\pi N} \frac{\sin^2[\pi N(f_1 - f_2 - \lambda + \mu)]}{\sin^2[\pi(f_1 - f_2 - \lambda + \mu)]} \end{aligned}$$

如果  $N$  足够大, 使得在  $-1/2 \leq \lambda \leq 1/2$  范围内  $2\pi F_N(f_1 + f_2 - \lambda - \mu)$  和  $2\pi F_N(f_1 - f_2 - \lambda + \mu)$  可以近似看作为出现在  $\lambda = f_1 + f_2 - \mu$  和  $\lambda = f_1 - f_2 + \mu$  处的脉冲, 则(12)式为:

$$\begin{aligned} \text{Cov}[\hat{G}(f_1), \hat{G}(f_2)] &\simeq \frac{\sigma_x^4}{N} \int_{-1/2}^{1/2} \theta(\mu)[\theta(f_1 + f_2 - \mu) \\ &\quad + \theta(f_1 - f_2 + \mu)]d\mu \end{aligned} \quad (14)$$

进一步设  $\theta(f)$  的主瓣足够窄以致可忽略  $\theta(\mu)\theta(f_1 + f_2 - \mu)$ , 由(14)式得:

$$\text{Var}[\hat{G}(f)] \simeq \frac{\sigma_x^4}{N} \int_{-1/2}^{1/2} [\theta(\mu)]^2 d\mu \quad (15)$$

由于  $N$  很大,  $N$  个观测值的周期图方差近似为  $\sigma_x^4$ , 所以经过加窗平滑和未经加窗平滑的方差之比为:

$$\begin{aligned} \Lambda &= \frac{\text{Var}[\hat{G}(f)]}{\sigma_x^4} \\ &= \frac{1}{N} \int_{-1/2}^{1/2} [\theta(\mu)]^2 d\mu \\ &= \frac{1}{N} \sum_{k=-M}^M q^2(k) \end{aligned} \quad (16)$$

从(16)式可以看出, 当  $M$  相对于  $N$  减小时, 谱估计的方差也随之减小。

综上所述,考虑到功率谱估计量的无偏性希望最大后延 $M$ 尽可能取得大,而又为了减小功率谱估计量的方差, $M$ 又希望取得尽可能小,鉴于此,最大后延 $M$ 的选取宜大小适当,通过计算等试验,一般取为样本长度 $N$ 的 $1/3$ 到 $1/10$ 比较好。表1a即是一个 $N=60$ 的理想样本序列所计算的周期(在各种不同的 $M$ 之下)

表1a 时间序列所含周期随最大后延 $M$ 的变化情况

Tab. 1a The variations of period  $T$  with the Maximize Lag  $M$

$M$	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$T$	6.316	6.364	6.400	6.207	6.250	6.286	6.316	6.341	6.222	6.250	6.275	6.296	6.316	6.333	6.250

Chatfield, C.<sup>[4]</sup> 曾给出经验公式为  $M = 2\sqrt{N}^D$ , 这个式子的计算结果由表1b所示。

表1b 截断点 $M$ 的选取和记录长度 $N$ 的关系

Tab. 1b Choice of truncation point  $M$  depends upon data length

$N$	$M$	$M/N$
30	11	0.37
50	14	0.28
100	20	0.20
200	28	0.14
300	35	0.12
400	40	0.10
500	45	0.09
1000	63	0.063

### 三、样本长度与其隐含周期的关系

在气象序列的研究中很可能出现这种情况,即记录的区间只包含重要周期现象的一部分或大部分。如果这种情况在分析中是有意义的,自然会提出这样的问题:当仅含有不到一个周期的观测记录时,我们能以多大可靠性确定周期值?

假定连续函数:

$$f(t) = \begin{cases} \sin(\omega_0 t + \varphi) & 0 \leq t \leq t_0 \\ 0 & \text{其它情况下} \end{cases} \quad (17)$$

其傅氏变换为:

$$F(\omega) = \int_0^{t_0} \sin(\omega_0 t + \varphi) e^{i\omega t} dt \quad (18)$$

通过推算可得:

$$F(\omega) = \frac{e^{i\varphi}}{2(\omega_0 - \omega)} [e^{i(\omega_0 - \omega)t_0} - 1] - \frac{e^{-i\varphi}}{2(\omega + \omega_0)} [e^{-i(\omega_0 + \omega)t_0} - 1], \quad (19)$$

1) 经验公式  $M = 2\sqrt{N}$  的理论证明, 么枕生将以论文“功率谱中的实际与理论问题”在《气候学研究——统计气候学》论文集发表。

其功率谱为:

$$I(\omega) = |F(\omega)|^2 \quad (20)$$

其自相关函数为:

$$\begin{aligned} R(k) &= \int_{-\infty}^{\infty} f(t)f(t+k)dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2 e^{i\omega k} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} I(\omega) e^{i\omega k} d\omega \end{aligned} \quad (21)$$

将(19)式代入(20)式得:

$$\begin{aligned} I(\omega) &= \frac{1}{4(\omega_0 - \omega)^2} [e^{-i(\omega_0 - \omega)t_0} - 1][e^{i(\omega_0 - \omega)t_0} - 1] \\ &\quad + \frac{e^{-2\varphi i}}{4(\omega_0^2 - \omega^2)} [e^{-i(\omega_0 - \omega)t_0} - 1][e^{-i(\omega_0 + \omega)t_0} - 1] \\ &\quad + \frac{1}{4(\omega_0 + \omega)^2} [e^{i(\omega_0 + \omega)t_0} - 1][e^{-i(\omega_0 + \omega)t_0} - 1] \\ &\quad + \frac{e^{2\varphi i}}{4(\omega_0^2 - \omega^2)} [e^{i(\omega_0 + \omega)t_0} - 1][e^{-i(\omega_0 - \omega)t_0} - 1] \end{aligned} \quad (22)$$

时间函数与它的复谱的关系本来可以作为时间序列分析的工具,只是由于复谱使用不便,况且对于时间序列的各波动来说,我们感兴趣的仅是哪个周期波动起主要作用,为此对(22)式求导并命等于零:

$$\frac{d|I(\omega)|}{d\omega} = 0 \quad (23)$$

对频率方程(23)式求解。即可求得其主要周期波动所对应的频率。当然这仅是从理论上而言,所给的函数又是极其简单的,实际情况又将是怎样呢?通过实际计算,我们能否寻求一些有用的信息呢?回答是肯定的。

首先,举一简单函数  $y = \sin x$ , 在此不妨令  $x = 1, 2, \dots, 60$ 。此函数离散成一组样本长度  $N = 60$  的序列,然后对此序列进行功率谱估计,求出其相应的周期  $T$ 。不断地变化  $N$ ,重复上述计算过程,得到一系列的  $T$ ,列于表 2。

表 2 功率谱的周期估计 ( $\Delta x = 1, y = \sin x$ )

Tab. 2 Periodic estimation of power spectra ( $\Delta x = 1, y = \sin x$ )

$N$	60	50	40	30	20	15	10	9	8	7	6
$T$ 理论	6.283	6.283	6.283	6.283	6.283	6.283	6.283	6.283	6.283	6.283	6.283
$T$ 计算	6.286	6.286	6.286	6.286	6.286	6.286	6.154	7.000	6.667	6.250	7.272

上述是在抽样区间  $\Delta x = 1$  的情况下,所得的计算结果。如果取  $\Delta x = 0.1$ , 其计算结果列于表 3。从表 2、表 3 中可以清楚地发现,在以单个周期振动的时间序列中,样本长度

表 3 功率谱的周期估计 ( $\Delta x = 0.1, y = \sin x$ )Tab. 3 Periodic estimation of power spectra ( $\Delta x = 0.1, y = \sin x$ )

$T \backslash N$	70	60	55	50	45
$T$ 理论	6.283	6.283	6.283	6.283	6.283
$T$ 计算	6.667	7.272	6.667	5.714	5.333

$N$  小于或接近于其周期  $T$  时, 谱分析能够反映出其原有的周期特征, 抽样区间小些, 其抽样长度可以相应地短些。需要指出的是尽管表 2、表 3 中所计算的周期, 其最大相对误差达 15.7%。不过我们认为这很大部分归结于计算误差。由于计算的周期  $T = \frac{2M}{l}$ , 按通常的方法,  $l = 0, 1, 2, \dots, M$ , 这势必会带来相当大的计算误差, 作者曾对此方法作了改进<sup>[9]</sup>。

上述仅是对单周期振动的序列而言, 下面就双周期的序列进行考察。假定一个函数为:

$$y = \sin\left(\frac{10}{7}x\right) + \cos\left(\frac{2}{3}x\right)$$

同样, 抽样区间  $\Delta x = 1$ , 计算结果列于表 4。

表 4 功率谱的周期估计 [ $\Delta x = 1, y = \sin\left(\frac{10}{7}x\right) + \cos\left(\frac{2}{3}x\right)$ ]Tab. 4 Periodic estimation of power spectra [ $\Delta x = 1, y = \sin\left(\frac{10}{7}x\right) + \cos\left(\frac{2}{3}x\right)$ ]

$T \backslash N$	60	50	40	30	20	15	14	13	12	11	10
$T_1$ 理论	9.425	9.425	9.425	9.425	9.425	9.425	9.425	9.425	9.425	9.425	9.425
$T_1$ 计算	9.565	9.565	9.565	9.565	9.565	9.167	8.8	8.8	9.474	10.588	10.588
$T_2$ 理论	4.398	4.398	4.398	4.398	4.398	4.398	4.398	4.398	4.398	4.398	4.398
$T_2$ 计算	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.286	4.286	4.286

从表 4 中同样可以看到抽样长度在接近那个较长的周期时, 还是能很好地反映出其周期特征的。

虽然上述计算是在理想情形下进行的, 所得的信息却是令人鼓舞的。气象、气候序列的样本长度一般都是仅仅几十年而已, 这就使得人们在作周期分析时, 忧心重重, 但是, 从我们的论述中, 不仅在理论上, 而且在实际计算中有力地佐证了功率谱分析的实用性。只要原序列是平稳的, 那就并非一定需要太长的样本序列, 只要其峰值达到显著性限度即可, 这既克服了气象序列不够长的弱点, 又节省了大量的计算时间。

#### 四、谱分析对所含周期的可辨性

谱分析在气象学中的应用和发展是有目共睹的, Roden(1966)<sup>[6]</sup> 曾分析了美国西部气温的周期情况, 高仕功(1970)<sup>[7]</sup>曾分析了温度的时空分布, Nicholson 等(1986)<sup>[8]</sup>曾分析了非洲干旱情况。这种方法在理论上似乎是无懈可击的, 而实际应用时又受到一定的限制。我们仅从一些实例中揭示一些规律及注意事项。

采用谱方法不一定能把序列中所隐含的各种周期都反映出来, 从下面几个例子, 可充分说明这一点。假定一个理想函数:

$$y = \sin\left(\frac{10}{7}x\right) + \cos\left(\frac{2}{3}x\right)$$

抽样区间  $\Delta x = 1$ , 计算结果列于表 4, 由表 4 所示, 原序列中的两个周期确是分开了, 且相当精确。又假定一个理想函数:

$$y = \sin\left(\frac{4}{3}x\right) + \cos\left(\frac{2}{3}x\right)$$

抽样区间  $\Delta x = 1$ , 则同样可以精确地表现出原序列中的两个周期来, 计算结果列于表 5。

不过, 如果假定一个理想函数:

$$y = \sin\left(\frac{10}{8}x\right) + \cos\left(\frac{2}{3}x\right)$$

同样在抽样区间  $\Delta x = 1$  的情况下, 虽然分出了两个计算周期, 却与序列原来的两个真实周期具有差异, 其结果列于表 6。

表 5 功率谱的周期估计

Tab. 5 Periodic estimation of power spectra  $[\Delta x = 1, y = \sin\left(\frac{4}{3}x\right) + \cos\left(\frac{2}{3}x\right)]$

$T \backslash N$	60	50	40	30	20	15
$T_1$ 理论	9.425	9.425	9.425	9.425	9.425	9.425
$T_1$ 计算	8.696	9.091	9.091	9.524	9.524	8.333
$T_2$ 理论	4.712	4.712	4.712	4.712	4.712	4.712
$T_2$ 计算	4.762	4.762	4.762	4.762	4.651	4.878

在上面所举的例子中, 两周期尽管不正确, 却还能分离出来, 而在下面例子中, 将出现两个周期的合并现象。假定一个理想函数:

$$y = \sin\left(\frac{10}{9}x\right) + \cos\left(\frac{2}{3}x\right)$$

表 6 功率谱的周期估计

Tab. 6 Periodic estimation of power spectra  $\left[\Delta x = 1, y = \sin\left(\frac{10}{8}x\right) + \cos\left(\frac{2}{3}x\right)\right]$ 

$T \backslash N$	70	60	50	40
$T_1$ 理论	9.425	9.425	9.425	9.425
$T_1$ 计算	8.148	8.148	7.586	7.857
$T_2$ 理论	5.027	5.027	5.027	5.027
$T_2$ 计算	5.366	5.366	5.366	5.366

抽样区间  $\Delta x = 1$ 。当样本长度  $N = 60$  时,  $T_{\text{计算}} = 6.667$ 。介于原来两个周期之间, 当  $N = 50$  时,  $T_{\text{计算}} = 6.875$  同样如此, 又假定一个理想函数:

$$y = \sin x + \cos \frac{2}{3}x$$

取抽样区间  $\Delta x = 1$ , 其计算结果, 列于表 7。

表 7 功率谱的周期估计  $\left[\Delta x = 1, y = \sin x + \cos\left(\frac{2}{3}x\right)\right]$ Tab. 7 The Periodic estimation of power spectra  $\left[\Delta x = 1, y = \sin x + \cos\left(\frac{2}{3}x\right)\right]$ 

$T \backslash N$	150	120	110	100	90	80	70	60	45	40	30	20	15
$T$ 计算	7.407	7.407	7.407	7.407	7.407	7.407	7.407	7.407	7.407	7.407	7.407	7.407	7.407

从表 7 中发现, 谱分析更侧重于表现原序列的整体波动状况, 其计算结果相当稳定, 不管样本长度的长短, 在由两个不同周期起伏所组成的新周期起伏时, 谱分析结果就不再表现出原来的固有周期的波动情况, 而是表现出其合成周期的波动情况, 这在理论上也是可以理解的。因为谱分析的先决条件是对任一以周期  $T$  变化的时间函数, 在满足狄氏条件下, 可以展开成傅里叶级数, 然后通过一系列推导出谱密度函数。么枕生<sup>[9]</sup>曾用谱分析的参数处理方法在理论上证明了周期叠加的事实。

从上述计算中, 发现这样一个有趣的事实, 在原序列中的两个固有周期相差一个整数倍时, 或大于一个整数倍时, 能正确地分辨出原有的两个周期。反之, 就比较难, 甚至分辨不出原来的固有周期。

由此可见, 谱分析并非一定能把原序列中的各个周期分离出来, 这种分离定受到一定的条件限制, 此外, 气象、气候序列中, 一般不可能由单一周期所组成。所以, 为了更好地利用谱分析技术, 在对原序列作谱分析之前, 进行一番考证, 作必要的滤波处理。以及去掉趋势是必要的。



## 五、结 论

通过上述讨论,可得出以下结论:

1. 谱分析中最大后延 $M$ 取作为样本长度 $N$ 的 $1/3$ 到 $1/10$ ,因样本长度而各异,有其理论背景,并且得到实际计算的验证。
2. 当样本平稳时,样本长度 $N$ 并非一定要求太长,仅略为超过其隐含周期的长度即可。由于通常气象、气候序列都不是很长的,这就显得尤其重要,一方面人们可以放心地去使用谱分析技术,另一方面又可以节省大量的计算时间。
3. 谱分析也是有其弱点的,如果样本中仅含有两个波动,且当这两个波动的周期相差一个整数倍或更大,这两个波动的周期是可以通过谱分析求出,反之,就会歪曲了其隐含周期。由此可见,在具体应用谱分析技术时,作一些滤波和去掉趋势是十分必要的。本文的工作对广大气象、地理工作者以及其它相邻学科的科研工作者具有十分重要的参考价值。

## 参 考 文 献

- [1] Bath, M., Spectral Analysis in Geophysics, Elsevier, 1974. (郑治真等译,地震出版社,1978。)
- [2] 么枕生,我国气候科学四十年来的蓬勃发展,地理研究, 9(1),1990。
- [3] Priestley, M. B., Spectral Analysis and Time Sieries, Academic Press. 1, 1981.
- [4] Chatfield, C., The Analysis of Time Series: An Introduction, Chapman and Hall, 1980.
- [5] 王强,气候功率谱的一种改进的计算方法,南京大学学报, 25(1), 1989, 142—149。
- [6] Roden, G. I., A Modern Statistical Analysis and Documentation of Historical. Temperature Records in California, Oregon and Washington, 1821—1964, J. Appl. Meteorol., 5, 1966, 3—24.
- [7] Kao, S.K., Wavenumber-Frequency Spectra of Temperature in the Free Atmosphere, J. Atmos. Sci., 27, 1970, 1000—1007.
- [8] Nicholson, S.E., and Entekhabi, D. The Quasi-Periodic Behavior of Rainfall Variability in Africa and its Relationship to the Southern Osillation, Arch. Met. Geoph. Biocl. Ser. A 34, 1986, 311—348.
- [9] 么枕生,自回归模式在气候学研究中的应用,气候学研究——“天、地、生”相互影响问题,气象出版社,1989,1—9。

## DISCUSSIONS ON SOME KEY PROBLEMS IN POWER SPECTRA ANALYSIS

Wang Qiang

*(Institute of Atmospheric Physics, Chinese Academy of Sciences Beijing)*

Yao Zhengsheng (C. S. Yao)

*(Department of Atmospheric Sciences, Nanjing University)*

**Key words** Maximum lag; Length of sample series; Period

### Abstract

The present paper include three aspects of question. (1) how to determine the value of Maximum Lag  $M$  in theory? (2) what is the relationship between the length of sample series  $N$  and the period  $T$  which is contained in this sample series? (3) how much is reliability of the period  $T$  which is disclosed by spectra analysis?

The value of Maximum Lag  $M$  is expected to be high if spectra unbiased estimation is considered. At the same time, however, this value is expected limited under the consideration of reduction variance of spectra estimation. It is advisable then that the value of Maximum Lag  $M$  be selected as  $1/3$  to  $1/10$  of the length of sample series  $N$  in both theory and practical calculations.

Here lies the relationship between the length of sample series  $N$  and the period  $T$  which is contained in sample series with calculation and deduction. As long as the sample series is stable, sample series  $N$  needn't be too long, only a bit longer than the longest period  $T$  which is included in this sample series.

Meanwhile, spectra analysis is not omnipotent. If one sample series has two waves, when one wave period is one times larger than another or much more, we can get these two wave periods with spectra analysis. Otherwise, there will appear "aliasing" and the attempt to get these two wave periods will be a failure.

So it is necessary to filter and re-trend for sample series, before we use this spectra analysis. These speculations have great relevance to the research of meteorologists and other fields.