水电水利

Hydropower and Water Resources

第 9 卷◆第 8 期◆版本 1.0◆2025 年

文章类型：论文|刊号（ISSN）：2529-7821 /（中图刊号）：868GL002

基于功率谱的湛江市降雨序列周期分析

张胜华

广东省水文局湛江水文分局

DOI:10.12238/hwr.v9i8.6536

[摘　要] 雷州半岛地形三面环海,中部较高而沿海较低,为典型的台地。受此影响,形成众多以独流入海的面积在200km²以下的中小河流。为了开展水文测验工作,设立了相应的雨量站,多年以来,累积了相当长序列的降雨资料。为分析各站资料序列的周期性问题,利用离散傅立叶变换、正交滤波函数等方法,对逐月降雨进行了功率谱密度分析。分析结果显示连续长序列显著性水平较高,虽然中间有较长缺测时段显著性水平较低,但主要显著性周期均为1年,对应于约一年一度的台风周期。

[关键词] 雷州半岛；降雨；周期性；功率谱密度；PSD；DFT；正交滤波函数；Multitaper Method

中图分类号：P481　文献标识码：A

Routing Analysis of the Tutangshui Watershed Based on geomorphological Characteristics

Shenghua Zhang

Guangdong Provincial Bureau of Hydrology,Zhanjiang Hydrological Branch

[Abstract] The terrain of Leizhou Peninsula is surrounded by the sea on three sides, the central part is higher and the coastal part is lower, which is a typical platform. Affected by this, many small and medium-sized rivers with an area below 200km² flow into the sea alone. Corresponding rainfall stations have been set up for the purpose of hydrological surveys, and a fairly long sequence of rainfall data has been accumulated over the years. In order to analyze the periodicity of the data series of each station, the power spectral density analysis of monthly rainfall is carried out by using methods such as discrete Fourier transform and orthogonal filter function. The analysis results show that the significance level of the continuous long series is relatively high, and the significance level of the long series with a long missing period in the middle is relatively low, but the main significance period is 1 year, corresponding to the typhoon cycle about once a year.

[Key words] Leizhou Peninsula; rainfall; periodicity; power spectral density; DFT; Multitaper Method

前言

湛江市面积13263平方千米,进行降雨观测历史比较悠久,梅菉站自1951年起进行记录,1963年,降雨站点达到40个；2009年,站数超过100个；2012年,站数超过150个。目前湛江市有173个降雨站点,进行数据整编出版年鉴的有52个站,资料序列长短不一,站点降雨是典型的随机过程,可以用功率谱方法探讨各站数据的周期性。

1 谱密度函数的估计

1.1周期图(Periodogram)法

周期图(Periodogram)又称直接法,通过对全部样本数据进行傅立叶变换得到,如式(1)

$$\hat{S}_x(k\Delta f)=\frac{1}{T}\left|\sum_{j=0}^{N-1} x_j e^{-i\frac{2\pi ijk}{N}}\right|^2, k=0,1,2,\cdots,N/2 \tag{1}$$

其中,T为N个样本x_j的长度,Δf=1/T。然而这种估计有很大缺陷：方差有与估计一样大的不确定性、会产生某种偏差等。由于每次FFT变换只能对有限长度的时域数据进行变换,因此,需要对时域信号进行信号截断。如果截断的时间长度不是周期的整数倍(周期截断),那么,截取后的信号将会存在频谱泄漏。周期图法得到的功率谱随着数据点数N的增大,分辨率变大、方差也变大。

1.2间接法

间接法又称Welch法。样本序列长度是有限的,对现有数据长度N,如果能获得更多的段数分割,将会得到更小的方差。允许数据段间有重叠部分,来得到更多的段数。首先将长度为N的数据xn,n=0,1,2,...,N-1分成L段,每段有M个数据,第i段数据表示为：

$$x_i(n)=x(n+iM-M),0\leqslant n\leqslant M,1\leqslant i\leqslant L \tag{2}$$

然后把窗函数 $w(n)$ 加到每一个数据段上,求出每一段的周期图,第i段的周期图为：

$$I_i(\omega)=\frac{1}{U}\left|\sum_{n=0}^{M-1} x_i(n)\,w(n)\,e^{j\omega n}\right|^2, i=1,2,...,M-1 \tag{3}$$

式中,U称为归一化因子,

$$U=\frac{1}{M}\sum_{n=0}^{M-1} w^2(n) \tag{4}$$

将每一段的周期图之间近似看成互不相关,最后功率谱估计为：

$$P_{xx}(e^{j\omega})=\frac{1}{L}\sum_{i=1}^{L} I_i(\omega) \tag{5}$$

常用的加窗函数有矩形、Bartlett、blackman、hamming、hanning、kaiser,在常用的信号处理软件包中均有。Welch法仍是参数化方法,需指定傅立叶变换长度(NFFT)、重叠长度,NFFT越长,则计算结果分辨率越精细。

1.3信度检验

对于得到的功率谱结果,应进行显著性检验。检验方法称为红白噪声检验。给定置信度α=0.05,有判据

$$R_\alpha=\frac{-1+t_\alpha\sqrt{n-2}}{n-1} \tag{6}$$

t_α为α信度水平和v=n-2时t分布值。如果原序列的一阶自相关系数R(1)大于R_α,则为红噪声,不同周期红噪声谱95%置信度上限为：

$$S_r^{0.05}=S_k\times\frac{1-R(1)^2}{1+R(1)^2-2R(1)cos\left(\frac{\pi k}{M}\right)}\times\frac{x^2}{(2n-\frac{M}{2})/M} \tag{7}$$

其中,x^2为卡方检验值。相应的,白噪声水平为：

$$S_w^{0.05}=S_k\times\frac{x^2}{(2n-\frac{M}{2})/M} \tag{8}$$

如果S_k超过红噪声水平,说明该周期是显著的,反之,则为不显著的。

1.4多窗口谱估计(Multitaper Method)法

由于几个原因,Welch法不能令人满意。首先,由于有效记录缩短了M倍,因此存在严重的低频分辨率损失,特别是在需要大量平均时。其次,没有好的理论来选择M,段数;此参数设置了整个光谱的分辨率和方差减少之间的固有权衡。第三,所有频率的方差减少和分辨率是固定的。

多窗口谱估计对序列首先进行数字滤波,为此,对于序列长度T,构造一组正交滤波函数

$$\int_0^T\prod_{k=1}^{M}\phi_k(t)\,dt=0 \tag{9}$$

产生一组统计独立周期图谱,可进行平均。Slepian法构造称为离散延长球体(discrete prolate spheroidal)的序列作为正交滤波函数；正弦函数法(sine)构造一组正弦函数,

$$\phi_k(t)=\sqrt{\frac{2}{T}}\sin\left(\frac{k\pi t}{T}\right), 0\leq t\leq T, k=1,2,3,... \tag{10}$$

通过检验每个频率上的最小均方误差(MSE),来决定应取多少滤波函数。构造一个MSE的近似函数

$$L=\beta^2+\upsilon\{\hat{S}(f)\}=\frac{S''(f)^2K^4}{576T^4}+\frac{S(f)^2}{K} \tag{11}$$

其中β为偏态系数,υ为方差,K是滤波后周期图谱的平均数,$S''(f)$是功率谱密度PSD对频率f的二阶导数。对于最小可能MSE,

$$K_{opt}(f)=\left(\frac{12T^2S(f)}{|S''(f)|}\right)^{2/5} \tag{12}$$

式(12)在每个频率上计算最小MSE,从而避免了选择平衡分辨率和合适方差的难度,这被称为“自适应”估计。

2 数据计算与分析

以资料序列最长的梅菉为例,从1951年4月开始观测,至2022年总计序列长度71年849月。当采用不同的nfft长度时,频率的分辨率越精细,相应的,噪声也增大,间接法计算结果见图1,x轴坐标转换为log10,y轴转换为20log10。自适应法计算成果见图2,红色线显示了估计的成果,蓝色为使用的sine滤波器个数。

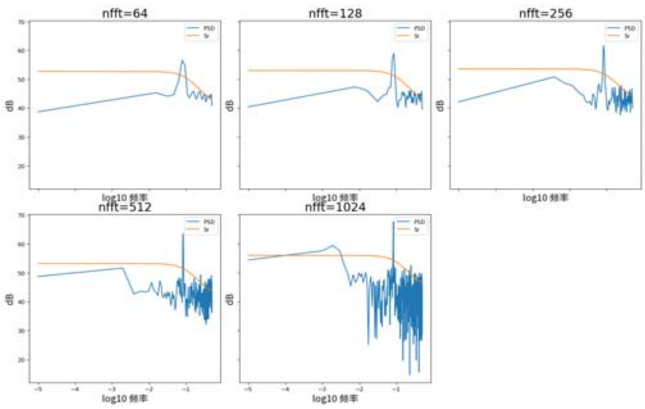

图1 梅菉站间接法计算PSD

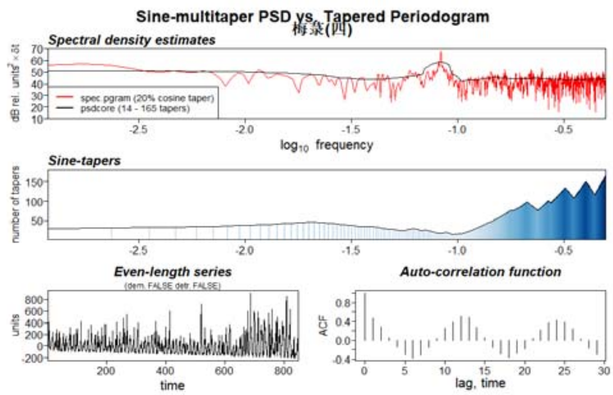

图2 自适应法计算PSD成果

部分站点序列进行计算的结果见表1和表2。

表1 间接法计算部分结果

站名	序列长度	不同nfft长度最显著频率				
		64	128	256	512	1024
梅菉(四)	849	0.078125	0.078125	0.078125	0.082031	0.081055
黄坡	825	0.078125	0.078125	0.082031	0.082031	0.082031
新联村	65	0.078125	0.078125		0.009766	
源水	65	0.078125	0.078125		0.009766	
良垌	65	0.078125	0.078125		0.009766	
桔仔树	65	0.078125			0.009766	
石滩仔	65	0.078125	0.078125		0.009766	
湛江港	624	0.078125	0.078125	0.078125	0.082031	0.081055
志满	156	0.078125	0.078125	0.078125	0.082031	
恭坑	156	0.078125	0.078125	0.082031	0.082031	
文章	65	0.078125	0.078125		0.009766	
东吴	131	0.078125	0.078125	0.078125		0.008789
滨洋	156	0.078125	0.078125	0.078125	0.082031	
南渡(二)	804	0.078125	0.078125	0.078125	0.082031	0.081055

表2 自适应法计算结果

站名	序列长度	Taper 个数	最显著频率
梅菉(四)	849	424	0.083924
黄坡	825	412	0.083942
新联村	65	32	0.096774
源水	65	32	0.096774
良垌	65	32	0.096774
桔仔树	65	32	0.080645
石滩仔	65	32	0.080645
湛江港	624	312	0.083601
志满	156	78	0.084416
恭坑	156	78	0.084416
文章	65	32	0.080645
东吴	131	65	0.085938
滨洋	156	78	0.084416
南渡(二)	804	402	0.083541

两种算法的结果,可以看出对于较短序列用间接法不能采用较大的nfft长度,并且由于批量计算的关系,很难针对单个序列进行nfft长度与结果的细微调整；自适应法也会有显著性不强的问题,查看序列,应与序列中间缺测有关,并且总体序列偏短。

3 结语

本文探讨了针对逐月降雨序列的周期性问题,并使用两种方法、不同软件包平台进行计算。对于各个站点,不考虑小数问题,均有显著的约1年的周期,这在实际上对应了雷州半岛地区一年一度的台风雨成因。由于水文气象数据序列相对于信号处理的高频数据偏短,所以在成果可靠性上不一定很显著,但进行这方面的探讨还是很有必要的。

[参考文献]

[1]中华人民共和国水利部.中华人民共和国水文年鉴:Vol.第8卷(第9册:珠江流域水文资料),2019.

[2]J.D.Hunter,Matplotlib:A 2D Graphics Environment[J].Computing in Science & Engineering,vol.9,no.3,pp.90-95,2007.

[3]A.V.Oppenheim and R.W.Schafer,Discrete-Time Signal Processing[M].Prentice-Hall,1999,pp.468-471.

作者简介：

张胜华(1986--),女,汉族,湖南汉寿人,硕士研究生,高级工程师,研究方向为水文水资源。