

文章编号:1001-4179(2013)S1-0022-04

# EEMD 在三峡水库入库量多尺度分析中的应用

王 坤<sup>1</sup>, 胡煜斌<sup>2</sup>

(1. 长江工程职业技术学院 水利工程系, 湖北 武汉 430020; 2. 长江工程职业技术学院 信息工程系, 湖北 武汉 430020)

**摘要:**三峡库区入库量作为一项重要的水文资料,其信号具有随时间变化的非线性和非平稳性特征。应用 EEMD 及其相关方法,对所获取的三峡库区入库量数据进行多尺度分解,提取了不同周期的入库量信号。分析结果表明:短周期 100 d 范围内的入库量信号在汛期变化剧烈,而在枯水期相对较小且平稳,且自 2003 年开始,入库量频率有略微增大的趋势;年变化周期占据了整个三峡库区入库量的主要成分,并且存在 5.7 a 周期的变化特征;分解信号的趋势项显示,入库量自 1998 年以来有下降趋势。

**关键词:**入库量; EEMD; 多尺度分析; 三峡水库

**中图分类号:** P337      **文献标志码:** A

三峡工程是目前世界上最大的水利枢纽工程,大坝位于湖北宜昌市三斗坪,库岸线全长约 5 300 km,库区蓄水后库容可达 390 亿  $m^3$ 。2003 年,三峡水库开始围堰蓄水至 135 m 发电,2008 年汛末,三峡水库初步建成并开始试验性蓄水,2010~2012 年正式完成 175 m 蓄水目标。实践证明,三峡工程建设对社会经济发展产生了显著带动和促进作用。与此同时,如此巨大的人类工程的建设 and 建成后的运行将会对环境形成何种程度的扰动,倍受社会各界关注。三峡水库在 175 m 蓄水水位下,回水长度达 600 余千米,平均库宽仅约 1 km,属于典型的河道型水库<sup>[1]</sup>,由此形成的长流流域水文系统变得更为复杂与多变。

中国长江三峡集团公司自建坝以来,提供了包括入库量、出库量、坝前上游及下游水位等水文信息。然而,随着长江沿岸经济的发展,除三峡大坝建成外,长江干、支流上相继建成了不少的大、中、小型水电站。这些水电站的建立使天然河道改变为人工控制的河道,因此,各监测点的水位信息所包含人工控制(蓄水)的因素过多,不能完整的体现库区上游及长江各支流的水情信息<sup>[2]</sup>。而对于具有区域性特征的长江流域,施晓晖和徐祥德研究认为,三峡库区入库量与长江流域上游前期降水有直接的关系,并且研究入库量

特征变化对于安全而有效地完成三峡水库蓄水任务具有重要的实用意义。因此,全面的认识与分析入库量的特征,对长江中下游防洪、三峡水库和葛洲坝水库的调度以及流域其他水利工程建设、工农业生产、航运等非常重要<sup>[3]</sup>。

傅里叶变换、小波多尺度分析在的水文信号处理和分析中得到了广泛应用。对于一些非线性、非平稳信号而言,傅里叶变换、小波分析等方法具有一定的局限性<sup>[4-5]</sup>。1998 年, N. E. Huang 等人提出了一种信号分析技术<sup>[4]</sup>,称为希尔伯特-黄变换(Hilbert-Huang Transform, HHT)。该方法包括两部分算法,即,经验模态分解(Empirical Mode Decomposition, EMD)和 Hilbert 谱分析,二者有机结合定量地刻画出非线性和非平稳过程。该方法不依赖于核函数的先验选择,而是分解信号起源于信号连续性极值的本征振荡模。王永文等采用 EMD 实现了对桂江流域年径流的多尺度分解,获得了简单且平稳性较好的分量,建立了基于 EMD 的年径流组合预测模型,预测结果稳定可靠、精度较高<sup>[6-7]</sup>。曹丽青和林振山应用基于 EMD 的 HHT 技术,对长江宜昌站 1900~2000 年来天然年平均流量序列进行了分析研究,并在此基础上建立了长江三峡水库年平均流量多尺度统计动力预报模型,预报效果

收稿日期:2013-01-07

作者简介:王 坤,女,硕士研究生,主要从事水文资料处理与分析、水利水电安全工程应用方面的研究。E-mail:wkun100@

163.com

较好。集成经验模态分解(Ensemble Empirical Mode Decomposition, EEMD)方法是在EMD基础上的改进,它继承了EMD的优点,通过向原始信号中添加白噪声,成功解决了由间歇性成分(间断)导致的模态混叠问题<sup>[8]</sup>,邵俊等通过对黄河三门峡水文站实测天然年径流序列进行分析,发现EEMD方法能够较好地解决模态混叠现象<sup>[9]</sup>。同时,与小波分析方法对比,得出的年径流量主要周期基本相同,并且该方法较之传统的EMD具有更高的精度,能够应用于水文时间序列多尺度分析研究中。由此可见,EMD及EEMD的创立与发展为基于非平稳变化的水文事件及过程开辟了一个新的领域。

三峡库区入库量作为一项水文系统的重要资料,它随时间变化的序列具有非线性和非平稳性特征。所涵盖的频率信息不仅能够体现长江流域年变化尺度以上的周期特点,而且在每年汛期与枯水期的频率变化也有所不同。本文应用EEMD及其相关方法,对所获取的三峡库区入库量数据进行处理和分析,研究水文信号分析方法,表述三峡库区流量的连续变化,挖掘并揭示其中所包含的各种特征信息。

## 1 EMD及EEMD方法原理

EMD是从复杂信号里分离出固有模态函数(Intrinsic Mode Function, IMF)的过程,也称之为筛选过程(The Sifting Process)。EMD方法假设任何信号都由不同的IMF组成,每个IMF既可以是线性的,也可以是非线性的,并且具有相同数目的极值点与过零点,即要求在时间轴上下对称分布。IMF需要满足以下2个条件<sup>[4]</sup>:①极值点数目和过零点数目相等或者最多相差1个;②在整个序列(任意一点),由局部极大值点和局部极小值点构成的两条包络线平均值为0。其分解过程主要步骤如下。

(1) 首先把原始数据序列 $x(t)$ 的局部极大值点与局部极小值点用三次样条插值(或其他插值方法)分别拟合成 $x(t)$ 的上下包络线 $u(t)$ 和 $v(t)$ ,则上下包络线的平均曲线

$$m(t) = [u(t) + v(t)]/2 \quad (1)$$

用 $x(t)$ 减去 $m(t)$ 后剩余部分

$$h_1(t) = x(t) - m(t) \quad (2)$$

(2) 由于包络线样条的过冲和欠冲影响,会产生新的极值并影响原来极值的位置与大小,因此,分解得到的 $h_1(t)$ 并不能完全满足IMF条件。用 $h_1(t)$ 代替 $x(t)$ ,再求出上下包络线 $u_1(t)$ 和 $v_1(t)$ ,重复进行,即:

$$m_1(t) = [u_1(t) + v_1(t)]/2 \quad (3)$$

$$h_2(t) = h_1(t) - m_1(t) \quad (4)$$

⋮

$$m_{k-1}(t) = [u_{k-1}(t) + v_{k-1}(t)]/2 \quad (5)$$

$$h_k(t) = h_{k-1}(t) - m_{k-1}(t) \quad (6)$$

(3) 当 $h_k(t)$ 满足IMF条件时,分解得到第一个IMF,  $c_1(t)$ 和信号剩余部分 $r_1(t)$ ,即:

$$c_1(t) = h_k(t) \quad (7)$$

$$r_1(t) = x(t) - c_1(t) \quad (8)$$

对信号的剩余部分 $r_1(t)$ 继续进行EMD分解,直到所得到的剩余部分为单一信号或者其值小于设定终止条件时,分解完成。在实际信号的处理过程中,完全满足IMF第二个条件是不现实的,所以只要二者的平均值小于一个预先确定的最小值即可。可以利用限制标准差 $I_{SD}$ 的值来判断每次筛选结果是否为IMF分量。

$$I_{SD} = \sum_{k=0}^T \left[ \frac{|h_{k-1}(t) - h_k(t)|^2}{h_{k-1}(t)^2} \right] \quad (9)$$

式中, $t$ 为时间信号长度; $I_{SD}$ 的取值决定了筛选过程停止条件。通常需要避免过于严格的准则,以免使IMF分量变成纯粹的频率调制信号,造成幅值恒定。而过宽的准则会使IMF分量与要求的相差太远。实际过程中通过对信号的反复筛选来决定 $I_{SD}$ 的值。经验表明, $I_{SD}$ 取0.2~0.3时,既可以保证IMF的线性和稳定性,又可使IMF具有相应的物理意义。

分解最终得到所有的IMF及残余量:

$$r_2(t) = r_1(t) - c_2(t), \dots, r_n(t) = r_{n-1}(t) - c_n(t) \quad (10)$$

而原始 $x(t)$ 是所有IMF及残余量之和:

$$x(t) = \sum_{i=1}^n c_i + r_n \quad (11)$$

由于每个IMF分量是代表一组特征尺度的数据序列,因此整个过程实际上是将原始数据序列分解为各种不同特征波动的叠加。所有的IMF分量经过逆向叠加,最后可以还原至原始数据 $x(t)$ 。由此可见,EMD分解法不需要一个平均值或零参照系,只需要局部极值的位置,对每个分量的零参照系将由筛分过程产生。因此EMD方法是依据原始数据进行的分解,得到的IMF分量个数通常也是有限的,并且各分量具有完备性和正交性,因此这种分解具有唯一性、自适应性<sup>[10-11]</sup>。

原始信号经过EMD分解,得到各个IMF的频率依次降低,剩余项是原始数据分解最后的趋势,代表信号的整体趋势。在实际应用中,可以根据各个分量频率由高到低的特点,对信号进行灵活重构,使其到达高通、带通及低通滤波的效果。

EEMD 方法以 EMD 为基础,继承了 EMD 的优点,利用噪声辅助数据分析,可解决由间歇性成分(间断)导致的模态混叠问题。这种新的方法是基于白噪声特性,能够明确地分离各个尺度的信号,而不需任何预先的主观选择标准,将白噪声应用于 EMD 形成一个有效的、自适应的、动态二进滤波器组,这样有利于在 EMD 中的数据分解。EEMD 的分解过程如下:

- (1) 在目标信号中加入白噪声辅助数据;
- (2) 将加入白噪声后的目标信号分解为 IMF;
- (3) 每次加入不同的白噪声,反复重复步骤(1)与步骤(2);
- (4) 把分解得到的各个 IMF 的均值作为最终的结果。

利用 EEMD 方法的目的是使添加的白噪声相互抵消,使 IMF 的均值保持在一个正常的动态滤波器窗口范围内,能够显著地削弱模式混叠,保存了信号中原有的动态特性。EEMD 利用了噪声中重要的统计特性以及 EMD 的尺度分离原则,使 EMD 能够真正成为任意数据的二进制滤波器组。通过加上有限的噪声,EEMD 自动排除了所有情况下的模式混叠。

## 2 三峡库区入库量

三峡库区入库量信息源于中国三峡总公司,同时,公司网站还提供 2003 年以后的水位、出库量等信息,其中入库量数据每天记录 4 次(汛期为每小时记录一次),水位和出库量数据每小时记录一次。本文采用库区 1992~2011 年近 20 a 来的日入库量(每天 20:00 记录数据)资料作为研究对象。因三峡工程于 2003 年进入围堰发电期,本文所用 1992~2002 年的入库量数据是否与建坝后入库量统计资料的位置(坝前)相同,网站未给出特别说明。而网站水情信息在 1992 年前是没有三峡入库量这项数据记录的,这也是本文数据时间范围为 1992~2011 年的原因之一。在入库量原始数据非平稳、非线性的变化过程中,能够较清晰地看出存在周期性特征,并且汛期与枯水期分界明显,汛期入库量波动大。

## 3 三峡库区入库量多尺度分析

对三峡库区入库量进行 EEMD 处理后,得到 11 个 IMF 分量  $imf1 \sim imf11$  及 1 个残余项( $imf12$ ),根据各个分量中心频率(cycle per day,简称 cpd,即:周期天数 =  $1/cpd$ )由高到低的特点,对信号进行了灵活重构,得到了  $imf1 \sim imf5$ ( $0.3 \sim 0.01$  cpd)、 $imf6 \sim imf8$ ( $0.01 \sim 0.001$  cpd)、 $imf9 \sim imf11$ ( $0.001 \sim 0.0002$  cpd)及  $imf12$ ( $0.0002$  cpd 以上)的结果。为更好地显示不同

周期分量的频率变化特点,将重构后的各个分量求得时频谱,直观地显示了入库量信号频率随时间变化的情况。

$imf1 \sim imf5$  对应的平均周期范围约为 3~100 d,反映了三峡库区在相对短周期内的入库量波动变化特征,其中入库量在汛期时间范围内(6~10 月)变化剧烈,是长江流域上游地区降水量增加后在库区的反映。而在枯水期(12~4 月)上游来水流量相对小且平稳。从频率变化特征来看, $imf1 \sim imf5$  重构分量主要存在约 7~8 d、14~15 d 及 45~50 d 的不同周期,且在 6~10 月的汛期频率有增大的趋势,表明因长江上游区域性降雨导致的三峡库区流量变化剧烈,周期减小到 4~5 d。特别注意的是,随着 2003 年库区水位突然的提升,处于 135~175 m 高水位(2003~2011 年)的库区入库量频率相比 100 m 以下(1992~2002 年)有增加的趋势。

$imf6 \sim imf8$  对应的平均周期范围约为 100~1000 d,体现了三峡库区长周期内的入库量变化特征,其中该分量分别在 1993、1998、1999、2005、2007、2010 年汛期期间表现出较大的入库量变化幅值。在 20 a 时间变化尺度内,年变化周期( $365 d = 1/0.00274$  cpd)占据了整个三峡库区入库量的主要成分。

$imf9 \sim imf11$  对应的平均周期范围约为 1000~5000 d,重构后的该分量存在 5.7 a 周期( $2092 d = 1/0.000478$  cpd)的变化特征,分析 20 a 内出现的 4 个主要波峰可以发现,两个较大或较小的波峰存在 11.4 a 周期特征,其中 1998 年与 2010 年处于两个较大波峰上,该两年洪峰在长江流域历史同期上处于相对较大的水平。由于在分解处理中同样避免不了的数据边界效用,导致在该分量频率在数据的两端产生了较大误差。

而趋势项  $imf12$  代表近 20 a 库区入库量整体趋势,显示自 1998 年以来,库区流量有下跌趋势,长江上游流域周边因经济发展用水量的增加以及各支流水电站的拦水效应可能是造成其入库量减小的主要原因。但是,三峡工程的建成是否对上游地区的气候及降雨量有所影响,从而导致了近年来入库量的下降趋势,因本文未涉及到该区域气候等气象资料,在此不作详细讨论。

## 4 讨论与结论

三峡库区连续入库量信号处理与分析的主要目的是将各种信息进行分离。本文结合国内外信号分析领域的发展趋势,应用 EEMD 方法对三峡库区入库量信号进行了多尺度分解,提取了不同周期的入库量信号,

有效避免了模态混叠问题,给出了重构后不同分量基于时间变化的频率特征,基于该方法得出以下 3 点结论:①短周期 100 d 范围内的入库量在汛期变化剧烈,而在枯水期相对较小且平稳,且自 2003 年开始入库量频率有略微增大的趋势;②年变化周期占据了整个三峡库区入库量的主要成分,并且存在 5.7 a 周期的变化特征;③通过分析趋势项,显示入库量自 1998 年以来有下跌趋势,长江上游流域周边因经济发展用水量的增加以及各支流水电站的拦水效应可能是造成其入库量减小的主要原因。

水文系统研究是一个很复杂的问题,对于大型的河道型水库,影响其入库量、水位等水文信息的因素更是难为全面掌握,因此通过总结与分析入库流量的变化特征,能够为三峡水库蓄水前后水情预报、三峡工程的库区流量控制提供一些有效的参考。本文所采用的 EEMD 方法完全趋于数据本身,因此所获得的多尺度分解信息不仅依赖于入库量统计精度,而且得到的主要周期仅体现原始数据 20 a 内的变化特征。而长江流域因涉及区域大,支流河道多,特别是上游气候变化、汛期降水量及流量统计影响因素等都是需要面对的问题,并且在分解过程中,虽然克服了 EMD 方法中的模态混叠现象,但边界效应同样是不能忽略的影响因素,这些问题也是本文下一步有待补充完善的工作。

#### 参考文献:

- [1] 张俊,闵要武,陈新国. 三峡水库动库容特性分析[J]. 人民长江, 2011, 42(6): 90-93.
- [2] 施晓晖,徐祥德. 三峡库区来水流量与长江流域上游前期降水的

关系研究[J]. 长江流域资源与环境, 2011, 29(9): 1062-1066.

- [3] 黄忠新,韩庆菊,李春香,等. 三峡水库 156 m 蓄水宜昌站水文特性与水情预报[J]. 人民长江, 2008, 39(13): 5-8.
- [4] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. Proceedings of the Royal Society A, 1998, 454(1971): 903-995.
- [5] Huang N E, Chern C C, Huang K, et al. A new spectral representation of earthquake data: Hilbert spectral analysis of station TCU129, Chi-Chi, Taiwan, 21 September 1999 [J]. Bulletin of the Seismological Society of America, 2001, 91(5): 1310-1338.
- [6] 王永文,付娟,金菊良,等. 基于经验模态分解的年径流组合预测模型[J]. 水电能源科学, 2010, 28(10): 16-19.
- [7] 曹丽青,林振山. 基于 EMD 的 HHT 变换技术在长江三峡水库年平均流量预报中的应用[J]. 水文, 2008, 28(6): 21-23.
- [8] Wu Z and Huang N E. Ensemble Empirical Mode Decomposition: A noise assisted data analysis method [J]. Adv. in Adap. Data Analy., 2009, 1(1): 1-41.
- [9] 邵骏,吕孙云,钱晓燕,等. 基于总体经验模态分解的水文序列多尺度分析[J]. 华中科技大学学报:自然科学版, 2011, 39(11): 105-108.
- [10] Flandrin P, Goncalves P. Empirical Mode Decompositions as data driven wavelet-like expansion [J]. International Journal of Wavelets, Multiresolution and Information Processing, 2004, (20): 310-315.
- [11] Huang N E, Shen Z, Long S R. A new view of nonlinear waterwaves: The Hilbert Spectrum [J]. Annual Review of Fluid Mechanics, 1999, (31): 417-457.

(编辑:李慧)

(上接第 2 页)

常情况及时报告相关人员并组织撤离。加强施工人员素质教育,尊重当地民族风俗习惯和宗教信仰,加强与当地政府联系,保障现场工作人员人身及财产安全。

## 4 结语

目前,雅砻江流域已建成的水情自动测报系统测站共计 146 个,包括水文(位)站 48 个、雨量站 80 个、自动气象站 18 个。其中,两河口以上共有水文站 8 个,雨量站 18 个,自动气象站 1 个;两河口至锦屏区间共有水文站 9 个,雨量站 33 个,自动气象站 14 个;锦屏至桐子林区间共有水文站 13 个,雨量站 29 个,自动气象站 3 个。目前雅砻江流域水情站网与 2007 年站

网规划前测站数量相比有大幅提高,基本满足流域中下游梯级水电站开发需要。

流域水文站网规划和建设是一项漫长而复杂的工作,应根据流域水利水电开发计划和需要进行统一规划,分期实施,避免重复建设。而水文站建设是一项过程复杂、专业性强、涉及行业广、建设项目繁多、单项工程量及投资较小的工作。若能在建设前期将施工过程中的各种情况考虑周全,则能够顺利开展,圆满完成水文站建设任务,为流域水电开发提供强有力的技术支持。

(编辑:李慧)