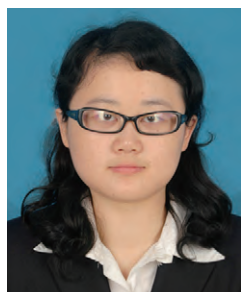


改善EMD端点效应的分析及在负荷预处理中的应用

在负荷预测中,对历史负荷数据进行预处理是十分必要的。电力负荷预处理应该充分考虑负荷曲线的非平稳性和非线性。经验模态分解能较好地平稳化处理信号,但是仍存在端点效应,因此提出改进的负荷预处理方案,先通过分解波形和评价指标对比分析了5种端点处理算法的优缺点,接着使用基于四中点估计法的EMD分解负荷信号,并引入相关系数法,按照合理阈值去除虚假分量并合并相关分量。最后针对某电网实际一周的电力负荷进行实例分析,验证了方案的实用性和有效性。

胡 杨 常鲜戎/华北电力大学电气与电子工程学院



胡 杨/硕士研究生

近几年,电力市场改革不断深入,对负荷预测准确度的要求也越来越高。提高预测准确度是负荷预测的关键,但目前多数预测方法假设负荷是平稳信号,未考虑负荷实际的非平稳性和非线性。采用经验模态分解法(Empirical Mode Function, EMD)对电力负荷进行预处理,能够分解出表现原负荷各局部特征的较平稳分量(Intrinsic Mode Function, IMF)^[1-3],再根据各分量的特点建模来提高预测准确度。

本文对电力负荷的历史数据做预处理,即运用EMD将非平稳的负荷时间序列分解成具有不同特征尺度的平稳IMF分量,并引入端点延拓评价指标,对几种端点效应处理方法进行分析比较,验证了四中点估计法在负荷预处理分解中抑制端点效应的效果。通过对某电网的实际负荷进行分解,引入相关系数 μ 去除分解IMF中的虚假分量,为接下来的负荷预测减小误差,表明所提出的方法具有较高的实用性和准确性。

关键词/Keywords

负荷预处理 ·
端点效应 ·
经验模态分解 ·
相关系数 ·

EMD算法及端点效应

EMD是把一个任意的数据序列 $x(t)$ 分解成为有限个基本模式分量 c_i 和一个趋势项 $r(t)$ 的过程,可表示为

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

电力负荷可表示为周期分量、随机变化量与趋势分量的和。EMD对负荷序列有平稳化的作用,能将负荷按其内在特性自适应地分解为若干个不同频率的平稳IMF。图1为利用EMD对某一天的负荷进行分解,对结果分析能更清楚地把握负荷特性。

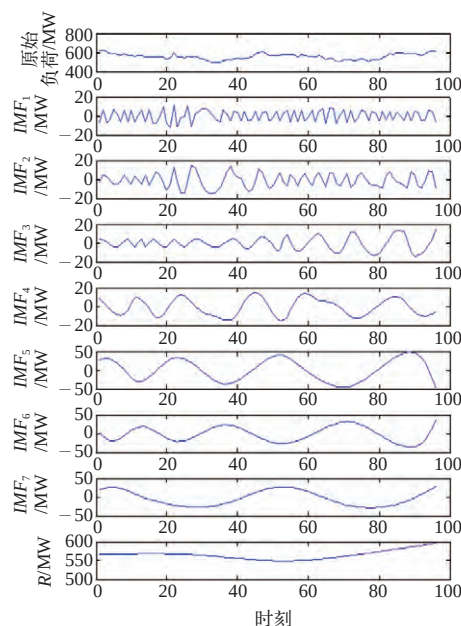


图1 原始负荷及EMD分量

当用EMD拟合上、下包络线时,由于边界处没有端点外信息的约束,信号的边界位置缺少极值点的信息,造成端

点处估计均值不准确,使拟合得到的包络线和真实包络线有差别,从而使信号的两端出现发散现象,产生端点效应。

抑制端点效应方法的仿真分析对比

目前,端点抑制大致可以分为3类:极值延拓法、波形延拓法与人工智能预测法^[4]。本文对5种常用的改善EMD端点效应的方法进行了C++编程并仿真,包括极值延拓法^[5]、包络线性延拓法^[6]、波形匹配法^[6]、半周期半对称延拓法^[7]和四中点估计法^[8]。其中,四中点估计法是在Directly-Mean EMD^[9]基础上引入的,直接从上、下线性包络的中点出发得到上、下包络的平均值,减少了计算量,提高了计算准确度。

同时,为了准确地分析各种方法的延拓效果,引入4种评价标来进行比较。

1) 比较EMD分解前后的能量来评估端点效应的影响程度。

实际计算中,可先计算原信号和EMD得到的各IMF的方根有效值,即

$$RMS = \sqrt{\frac{\sum_{i=1}^n s^2(i)}{n}} \quad (2)$$

式中, RMS 为信号有效值; $s(i)$ 为信号序列; n 为信号的采样点数。

按式(3)比较各IMF有效值的总和与原信号的有效值,得到一个评价指标 θ ,即

$$\theta = \frac{\sqrt{\sum_{i=1}^n RMS_i^2 - RMS_{original}}}{RMS_{original}} \quad (3)$$

式中, $RMS_{original}$ 为原信号有效值; RMS_i 为第*i*个IMF的有效值; n 为IMF总个数,包括EMD的残留项。 θ 值越大,端点效应影响越大。

2) 比较IMF和原始信号间的相似度,即计算相似系数 ρ 。

$$\rho_i[x_i(k), IMF_i(k)] = \frac{\cos[x_i(k), IMF_i(k)]}{\sqrt{\sigma(x_i)} \cdot \sqrt{\sigma(IMF_i)}} \quad (4)$$

ρ 值越大,说明端点效应的抑制越好。

3) 计算各个IMF和相应原信号分量的平均相

对误差,比较IMF的准确度。

$$\text{error_}IMF_i = \frac{\text{sqrt}\left\{\sum_{k=1}^N [x_i(k) - IMF_i(k)]^2\right\}}{N} \quad (5)$$

式中, N 为信号的个数; $x_i(k)$ 为原信号的第*i*个分量; $IMF_i(k)$ 为分解后的相应分量。

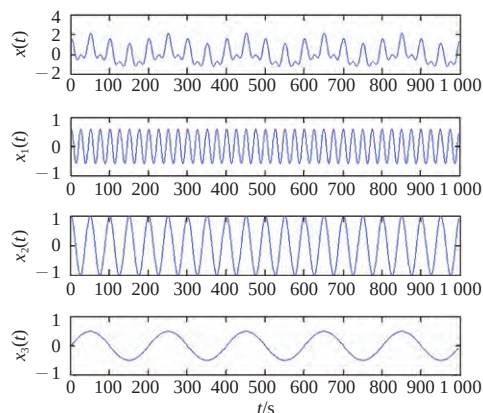
由式(5)可得, $\text{error_}IMF_i$ 越小,说明端点效应的抑制越好。

4) 计算程序运行时间和空间长度,希望在保证延拓质量的同时,还能兼顾实时性的要求。

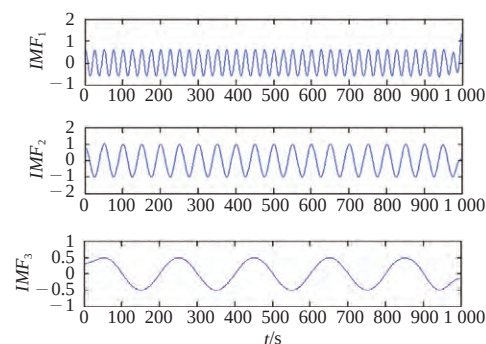
以一组调频调幅非线性信号进行仿真测试,该信号由3种分量叠加而成,频率各不相同,即

$$x(t) = \cos\left(\frac{2\pi}{50}t\right) + 0.6\cos\left(\frac{2\pi}{25}t\right) + 0.5\sin\left(\frac{2\pi}{200}t\right) \quad (6)$$

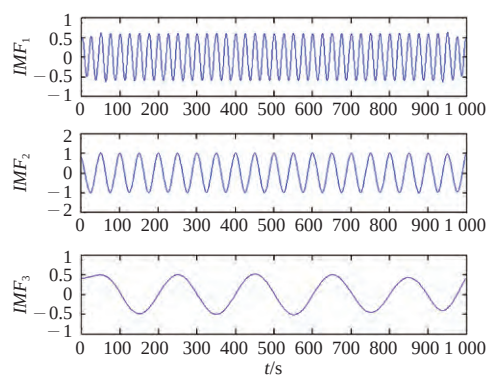
信号波形如图2a所示,通过仿真结果来考察上述5种端点抑制方法的性能,信号延拓后EMD分解的各分量如图2b~f所示。为了更具体地比较这5种端点效应抑制方法,考察这5种方法分解的各分量的评价指标如表1所示。



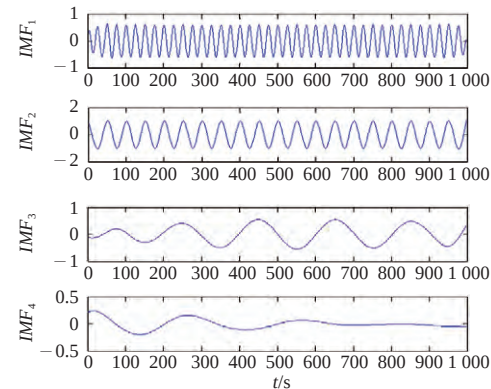
(a) 原信号及其组成成分



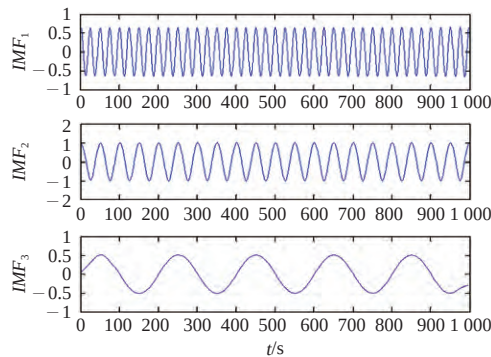
(b) 极值延拓法下的EMD分解



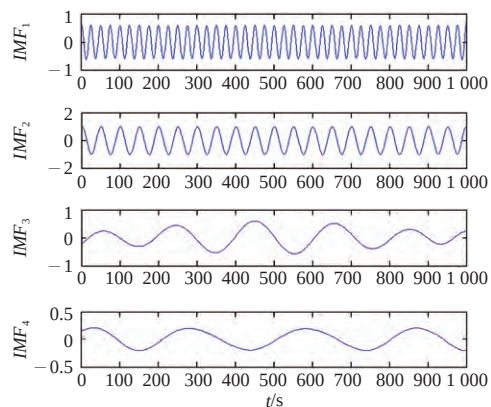
(c) 波形匹配法下的EMD分解



(d) 包络线性延拓法下的EMD分解



(e) 四中点估计法下的局域波分解



(f) 半周期半对称延拓法下的EMD分解

图2 原信号及其组成成分和5种延拓方法的分解结果

表1 信号端点延拓评价指标

延拓方法	能量误差	相似系数	分解误差	响应时间/s	空间长度
极值延拓法	0.008 57	0.987 5	0.002 131	0.52	原数据长度
		0.995 3	0.002 214		+ 左右延拓
		0.995 5	0.001 120		各4点
波形匹配法	0.014 79	0.997 7	0.000 921	0.62	原数据长度
		0.997 6	0.001 622		
		0.968 9	0.002 886		
包络线性延拓法	0.000 720	0.997 2	0.001 012	0.55	原数据长度
		0.996 2	0.001 977		
		0.930 8	0.004 101		
四中点估计法	0.007 389	0.998 3	0.001 245	0.72	原数据长度 +
		0.998 6	0.001 878		延拓2点
		0.998 2	0.000 692		
半周期半对称延拓法	0.006 100	0.999 1	0.000 558	1.42	3倍原数据长度
		0.999 5	0.000 783		
		0.907 5	0.004 724		

观察图2a~f中原始信号和分解的各分量波形以及各算法的评价指标，可以分析出以下几点。

1) 四中点估计法在端点抑制上效果最好，且各IMF分量相似系数最高，各频率分量的分解误差都较小，能量误差小于极值延拓法。

2) 极值延拓法、波形匹配法和四中点估计法在得到的IMF数量上较好，包络线性延拓法和半周期半对称延拓法分解出的IMF存在多余的虚假模式分量，且越是低频IMF分量，相似系数越差。

3) 5种算法的计算时间都在可接受范围内，极值延拓法时间最短，四中点估计法同等条件下计算时间稍长，但分解准确度有所提高，符合预期效果；半周期半对称延拓计算时间最长。

考虑到电力负荷的随机因素较多，非平稳性极强，波形匹配法在实际预测过程中并不适用；针对短期负荷来说，周期性和对称性并不十分显著，所以半周期半对称延拓法也存在不足。因此四中点估计法凭借相似系数最大，分解准确度较高，能量误差较小，运行时间适中，最重要的是对具有非平稳性的电力负荷分解更具实践性，能够满足工程需要而成为较佳选择。

负荷预处理实例分析

图3为某地区2011年9月1~7日连续一周（每日288个点）共2 016个数据点的负荷序列。该地区电力系统负荷序列经EMD分解后自适应地从低频到高频解析出23个IMF分量和一个趋势项 r 。

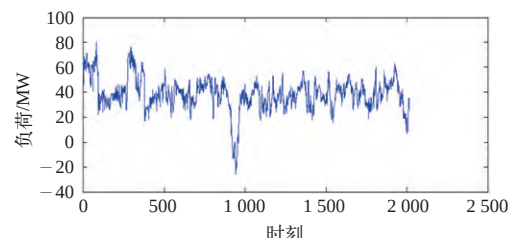


图3 2011年9月1~7日负荷曲线

通过相关系数 μ_i 分析可知，真实的IMF与原始信号的相关性较好，虚假IMF则较差。因此可以由 μ_i 来判断虚假IMF与真实IMF，即

$$\mu_i = \frac{\sum_{j=1}^N IMF_i(j)y(j)}{\left[\sum_{j=1}^N IMF_i^2(j) \sum_{j=1}^N y^2(j) \right]^{1/2}} \quad (7)$$

式中， $i=1,2,\cdots,n$ ； n 为IMF的个数。

虚假IMF与真实IMF的判断阈值 λ 定义为

$$\lambda = \max(\mu_i)/\eta \quad i=1,2,\cdots,n \quad (8)$$

式中，比率因子 η 的取值参考文献[10]中的 $\eta=8$ 。若相关系数 μ_i 大于 λ ，则认为其对应的IMF为真实的，将其保留；反之，则认为是虚假的IMF，将其去除，并加入到残余项中。

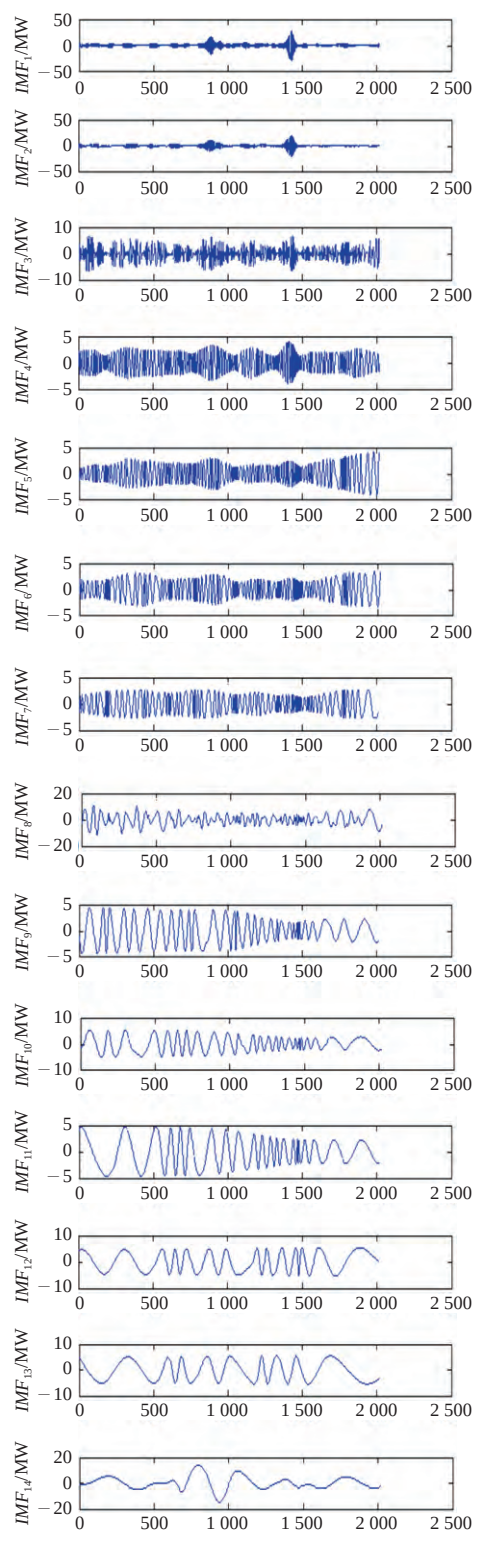
各个IMF分量和原负荷序列的相关系数如表2所示。

根据虚假分量判断准则，做以下处理。去除

表2 某地区一周负荷各IMF分量与负荷信号的相关系数

分量	相关系数	分量	相关系数	分量	相关系数
IMF_1	0.076 8	IMF_9	0.236 6	IMF_{17}	0.246 8
IMF_2	0.019 4	IMF_{10}	0.235 1	IMF_{18}	0.246 7
IMF_3	0.105 6	IMF_{11}	0.241 3	IMF_{19}	0.025 2
IMF_4	0.090 1	IMF_{12}	0.326 5	IMF_{20}	0.246 3
IMF_5	0.170 2	IMF_{13}	0.185 5	IMF_{21}	0.123 6
IMF_6	0.146 1	IMF_{14}	0.369 9	IMF_{22}	0.129 6
IMF_7	0.135 2	IMF_{15}	0.403 3	IMF_{23}	0.120 5
IMF_8	0.163 0	IMF_{16}	0.381 9	r	0.259 6

虚假分量 IMF_2 和 IMF_{19} ，考虑实际负荷序列的过度分解，合并相关系数较相近的分量，即合并 IMF_3 和 IMF_4 ，合并 IMF_9 、 IMF_{10} 和 IMF_{11} ，合并 IMF_{17} 、 IMF_{18} 和 IMF_{20} 。最终负荷分解波形如图4所示。



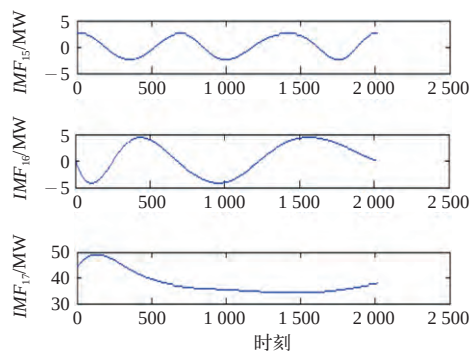


图4 去除虚假分量后的负荷序列EMD分解结果

改进后的电力负荷预处理包括EMD分解和去除虚假分量，将负荷分解为频率不等的分量，较好地解决了端点效应和不真实的分量，方案的基本流程如图5所示。

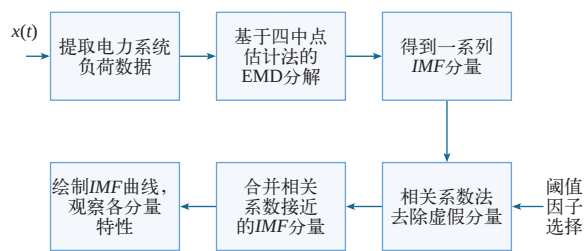


图5 负荷预处理流程图

方案考虑了负荷实际的非平稳性和非线性，先采用EMD对负荷平稳化，再按照合理阈值去除虚假分量并合并相关分量，通过上述实例分析，预处理降低了端点效应并减少了IMF的个数，提高了负荷预测建模效率，同时预测准确度也有所提高，证明方案有一定的实用性。

结束语

在负荷预测中，对具有周期性和非平稳性的历史负荷数据进行预处理是十分必要的。在EMD对负荷序列进行平稳化的处理，引入四中点估计法减弱端点效应带来的误差，将负荷按其内在特性自适应地分解为若干较平稳的IMF，而负荷预处理后的IMF突出了原负荷的局部特征，这些分量更易预测，波动性较小，使预测的负荷曲线较平滑，能提高预测准确度。

参考文献

- [1] 范新桥, 朱永利, 尹金良. 基于经验模态分解和基因表达式程序设计的电力系统短期负荷预测 [J]. 电力系统保护与控制, 2011, 39 (3): 46-51.
- [2] 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]. Mathematical, Physical and Engineering Sciences, 1998, 454 (1971): 903-995.
- [3] Veltcheva A D. Wave and group transformation by a hilbert spectrum [J]. Coastal Engineering Journal, 2002, 44 (4): 283-300.
- [4] 张金良, 谭忠富. 基于EMD计量经济学模型和混沌模型的短期负荷预测 [J]. 电网技术, 2011, 35 (9): 181-187.
- [5] 黄大吉, 赵进平, 苏纪兰. 希尔伯特-黄变换的端点延拓 [J]. 海洋学报, 2003, 25 (1): 1-11.
- [6] 胡维平, 莫家玲, 龚英姬, 等. 经验模态分解中多种边界处理方法的比较研究 [J]. 电子与信息学报, 2007, 29 (6): 1394-1398.
- [7] 高军, 孟涛, 祁继鑫, 等. 半周期半对称延拓的EMD边界处理方法的实现 [J]. 震灾防御技术, 2011, 6 (2): 24-29.
- [8] 刘坦然. 可完全消除 EMD 边界效应的四中点估计方法及风浪信号的希尔伯特谱 [D]. 青岛: 中国海洋大学, 2012.
- [9] Zhong Y, Qin S, Tang B. Research on theoretic evidence and realization of directly-mean EMD method [J]. Chinese Journal of Mechanical Engineering (English Edition), 2004, 17 (3): 399-404.
- [10] 刘畅, 周川, 伍星, 等. 基于广义形态滤波和相关系数的Hilbert-Huang变换方法 [J]. 机械科学与技术, 2011, 30 (1): 71-75.

(收稿日期: 2014-12-11) EA