

# 异常事件对EMD方法的影响 及其解决方法研究\*

赵进平

(国家海洋局第二海洋研究所,杭州,310012)

**摘要** 作者指出异常事件在数据中形成局部的高频信号,运用经验模态分解(EMD)方法分析这种存在异常事件干扰的数据,就会产生本征模函数(IMF)的频率混叠现象,而造成物理过程的重叠,使得难以用时间过程曲线表现特定的物理过程。这一问题是EMD方法中尚未妥善解决的问题。为解决这一问题,作者利用干扰信号极值及其两边的极大与极小值位置与原始数据有明显对应关系的特征,将相关IMF中的异常信息直接滤除,再用Spline插值方法弥补滤除时段的数据,得到重新拟合的该IMF数据。采用这种方法可以提取出异常信号,提取的精度与异常信号的时段长度有关。而且,拟合结果消除了异常干扰,可以将该IMF与其余IMF一起叠加成没有异常干扰的数据。将滤除了异常干扰的数据再次进行EMD分解,可以得到新的IMF系列,而它与不加校正的分解结果有相当大的差别,可靠地反映了真实物理过程。结果表明,只有在有效滤除异常干扰的情况下才能获得可靠的IMF系列,并准确地描述各种尺度的现象;消除了异常干扰的IMF可以任意单独或组合使用,表现各种时间尺度的变化与过程,所讨论的方法只适合异常时段较小的情形。对于异常时段接近或大于正常变化周期的干扰还需要探讨其他方法。

**关键词** 经验模态分解(EMD)方法;异常事件;本征模函数(IMF);高频信号

**中图法分类号** O141;O21;P73

**文章编号** 1001-1862(2001)06-805-10

## 0 导论

经验模态分解(EMD)方法和相应的Hilbert变换是由Huang等发展的一种新的数据分析方法<sup>[1]</sup>。这种方法不仅可以适用于线性过程的分析,而且适用于非线性和非平稳时间序列的分析。采用EMD方法可将信号分解为若干个本征模函数(Intrinsic mode function, IMF),这些IMF是完备的并且正交的<sup>[2]</sup>。由于EMD方法是依据数据本身的信息进行的分解,得到的IMF通常是个数有限的,而且表现了数据内含有的真实物理信息。EMD方法比Fourier级数方法有更多的优势,与小波分析方法有类似的功能,可以给出时间-频率-振幅的三维谱形<sup>[3]</sup>。在过去的几年,EMD方法已经被广泛用于非线性的海洋波动数据分析、地震信号和结构分析、桥梁

\* 国家攀登计划项目(970231003)和国家自然科学基金委员会重点基金项目(49634140)资助;国家海洋局海洋动力过程与卫星海洋学重点实验室科技成果第2001B006号

收稿日期:2001-06-29;修订日期:2001-08-18

赵进平,男,1954年出生,博士,研究员。

和建筑物状况监测的分析、生物医学信号如血压的波动、太阳辐射的变化分析等领域的研究<sup>[4~6]</sup>。

EMD方法的分解过程是:先将原始数据分解成第一个IMF和随时间变化的均值之和,然后将均值考虑为新的数据,将其分解为第二个IMF和新的均值。持续这种分解过程直至获得最后一个IMF。最后一个IMF的均值是一个常数。均值的获得方法是首先用三次样条函数拟合确定数据的上下包络,然后将上下包络的平均确定为均值。为保证均值确定的准确性,需要多次叠代,直至满足给定的判据<sup>[1]</sup>。

由于采用三次样条函数确定上下包络,需要给出边界点的包络值或者数据的一阶导数。事实上,边界上的数据顶多只能确定一条包络在边界上的位置,而对于另一条包络无法有效地确定。Huang等<sup>[1]</sup>采用的是在数据两端增加两组特征波的方式进行数据延拓,这种方法是行之有效的,但在以下几个方面不是很清楚:第一,由于数据本身不包含边界外点的信息,对数据的任何延拓都是没有依据的,增加特征波的方法缺乏物理依据。第二,确定每一个IMF往往需要几百次叠代,原有的方法每次都要增加特征波,会引入新的误差,影响结果的精度。第三,原有方法的一些延拓的结果可能因人而异,难以保证最终结果的惟一性。因此,Huang等<sup>[1]</sup>也认为该方法仍需要进行改进。Zhao and Huang认为,边界延拓的目的不是为了给出准确的端点外数据,而是提供一种条件,使得包络完全由端点以内的数据确定。根据这种观点,他们提出所谓“镜像”延拓方法。该方法在数据两端设置对称性的延拓,使得时间序列的数据成为闭合的环形数据,并证明了镜像延拓是满足物理条件的延拓方法<sup>[7]</sup>。这种方法保证了不同使用者可以得出同一的结果。最为重要的是,镜像延拓只需要在分解之前进行一次性延拓,在分解期间对所有模态和整个叠代过程都不再需要数据延拓,消除了误差传递的影响。因此,镜像延拓克服了原有延拓方法的不足,提高了EMD方法的可用性。

尽管EMD方法正在得到越来越广泛的应用,如果真的应用这种方法就会注意到方法本身还存在一些问题:

EMD方法的分解过程依赖数据本身包含的变化信息,这是EMD方法的主要优点。但是,如果在真实过程中由于某种原因发生异常事件,数据就会在某一段时间内出现高频信号,在分解结果中,第一个模态就会包括正常信号和那段高频信号,在出现高频信号的地方,正常信号被推移到下一个模态。这样就会造成两方面的问题:在第一模态,出现两种频率的混叠现象,使得该模态不能表现正常的频率过程;另一方面,在第二模态中出现与第一模态的正常频率相对应的信号,使第二模态也出现频率混叠。这种混叠可以一直延续到最后一个模态。图1是一个余弦振荡信号 $a$ 与一个异常事件 $b$ 相重叠的信号 $a+b$ 。其中余弦信号为

$$a = \cos\left(\frac{2\pi t}{T_0}\right) \quad (1)$$

而异常事件的脉冲信号为

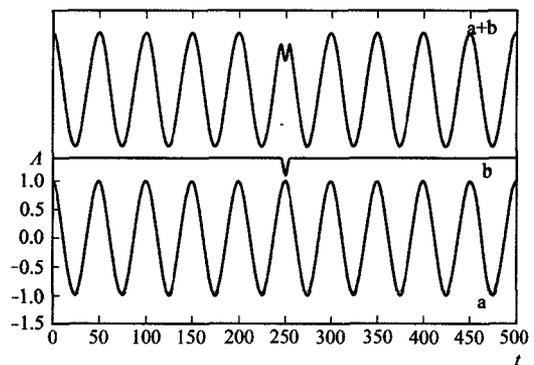


图1 存在异常事件干扰的信号(a. 正常的余弦信号;b. 异常事件;a+b. 合成信号)

Fig. 1 The signal with the interference from an abnormal event., (a. normal cosine signal; b. abnormal signal; a+b. combined signal)

$$b = -0.15 \left[ 1 - \cos \left( \frac{2\pi t_1}{T_1} \right) \right] \quad (t_0 < t_1 < t_0 + T_1) \quad (2)$$

式中,  $a$  为时间  $t$  的余弦信号, 周期为  $T_0$ ;  $b$  为时刻  $t_1$  的脉冲信号, 周期为  $T_1$ ;  $b$  只发生在  $t_0 < t_1 < t_0 + T_1$  的时间段,  $t_0$  为异常信号起始时刻。虽然这种异常信号只产生了很小的干扰, 但是, 正如图 2 所示, 对这种信号进行 EMD 分解, 各个模态都受到严重影响。在第一模态, 受干扰部分的余弦信号几乎无法看出, 第二个模态出现第一模态的频率, 并且不合理地扩展到很宽的范围。后面的模态也都受到影响。为方便起见, 我们将这一问题称为异常干扰问题。

EMD 方法并不认为这种干扰是个问题, 因为通过 Hilbert 变换, 混叠在各个 IMF 中的信号可以出现在时间-频率空间中的相应位置, 从而全面地展现频率信息的正确分布, 能量的分布并没有因分解的不同而发生错误。但是, 事实上, 异常干扰问题对 EMD 方法的使用存在严重影响, 是一个不可回避、而且必须妥善解决的问题。问题的严重性在于:

(1) 时间-频率-能量图展示的主要是能量随时间和频率的分布, 但是人们有时更感兴趣的是在某个频率范围内能量的时间变化过程, 以便于比较不同现象之间的相关性。但是由于异常事件造成每个 IMF 的频率混叠, 任何一个 IMF 或几个 IMF 之和都不能提供这样的变化曲线, 给不同现象的结果比较造成了困难。EMD 方法至今还没有解决这个问题。

(2) EMD 方法的优点是只有有限个 IMF, 而且每个 IMF 都可望与一定的物理过程相对应。但是, 从图 2 的分解明显看到, 频率混叠破坏了这种优点。干扰信号占据了原来物理过程曲线的位置, 使得每个 IMF 都不能反映真实的物理过程。正是这种异常干扰造成了两个不同物理过程的混淆, 妨碍了人们对分析结果的理解。在这种情况下, 虽然模态之间是正交的, 但却无法反映相互独立的过程。

基于这些考虑, 我们认为, 尽管我们可以从时间-频率-能量图上得到需要的变化信息, 我们仍然希望找出相近频率的时间变化过程作为 IMF, 使得 EMD 的结果能够更好地体现现象的物理背景, 也使 EMD 对不同现象的研究结果便于相互比较。因此, 本文集中研究异常事件对 EMD 方法的影响及其解决方法, 进一步拓展 EMD 方法的使用范畴。

## 1 异常干扰问题的解决方法

### 1.1 异常干扰的发生和可识别性

造成异常干扰的原因很多。异常的高频变化在地球科学现象中是普遍存在的。其实, 还有许多因素可以引起异常干扰, 如: 取样间隔较小时引起的局部波动, 数据非等间距时局部的密集或数据漏失等。

事实上, 由不同动力学作用产生的频率相同或相近的变化用 EMD 方法是不能区分的, 而

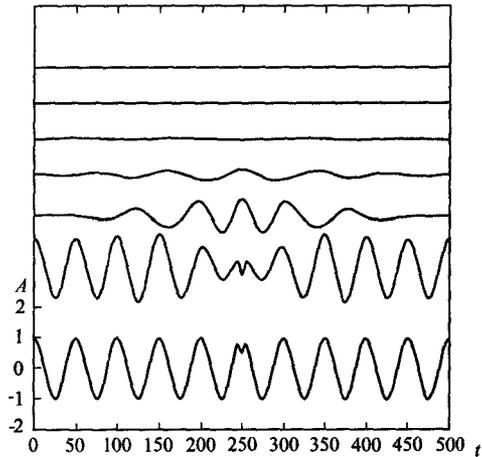


图 2 用 EMD 方法分解的各个模态  
(最下面是原始信号, 向上分别为 1-6 模态)

Fig. 2 IMF's of signal in Fig. 1 decomposed by EMD method. (The lower curve is the original signal and the upper curves are IMF's 1-6)

体现为不同频率的变化才有可能区分开来。如果两个不同频率的信号叠加在一起,可以用EMD方法将这两种信号在一定的精度下区分开来。但是,如果其中一个频率的信号不是连续的,而是体现为单个的脉冲,就会造成图2的问题,我们将其视为异常干扰。因此,可以认为,异常干扰信号只是不连贯的一般信号。

图1给出的信号是在正常信号峰值给出一个反向的脉冲信号,增加了曲线极值的数量,用EMD方法可以分解出图2所示的各个模态。如果异常信号是正向叠加在正常信号上,则不会增加曲线的极值数量,因而也就不能提取出异常信号;若要在正向叠加的信号中提取异常信息,需要采用其它数据处理方法。本文提供的方法只能区分反向叠加的正常信号和异常信号。

消除异常信号的干扰,人们首先想到的是,对原始信号进行一些处理,使异常信号在用EMD方法分解之前就被消除。但是,在实际研究中发现,不可能预先消除异常干扰,因为EMD方法与谐波分析方法不同,EMD方法面对的本身就是非稳态的数据。当我们在原始数据中注意到异常信号时,我们无法确定异常信号的量值和分布。也就是说,我们只能知道异常信号的存在,不能知道究竟信号的哪些部分属于正常的非稳态信号,哪些属于异常干扰信号。另外,有时异常干扰不是存在于第一模态,可以发生在第二模态,甚至更高的模态;这种情况下,无法在分解之前从原始数据中识别出异常信号。因此,我们需要一种方法来滤除干扰信号,这种滤除方法本身也将是提取异常信号的方法。

一般而言,对于任一模态的数据可以采用数字滤波的方法滤除这种不协调的波动,如Thompson设计的滤波器<sup>[8]</sup>。但是数字滤波的缺欠是需要牺牲一定长度的数据。而且,追求的滤波效果越好,滤波器的长度应该越长,牺牲的数据也就越多。有些类别的观测数据时间间隔较大,采用滤波器造成数据损失十分可惜。在此,我们尝试用其它方法消除异常干扰。

**1.2 异常信号的滤除方法** 既然不能直接从原始数据中滤除异常信号,我们需要从分解后的各个模态寻找可用的信息。我们还是用图1的合成信号为例来研究消除异常信号的方法。从图2可以看到,EMD方法分解的第1个IMF具有以下特征:

- (1) 异常信号的极小值与原始数据的极小值位置一致;
- (2) 异常信号极小值两侧的极大值与原始数据的相应极大值位置一致;
- (3) 异常信号极小值两侧的极小值与原始数据的相应极小值位置一致。

这些特征表明,第一IMF中包含了异常信号周期和正常信号周期的全部信息,我们将这种现象称之为位置对应特征。我们所不知道的是它是否也包含了振幅信息。从分解结果看,第2个IMF肯定不包括异常干扰的信息。我们可以期望第1个IMF中包含异常干扰的全部信息。

在图2的第1个IMF中,令干扰信号极小值的位置为 $X_0$ ,两侧极大值位置为 $\pm X_1$ ,两侧极小值的位置为 $\pm X_2$ 。我们以 $X_0$ 为中心,将其两侧 $\pm X_1$ 之间的数据滤除,并用 $(-X_2, -X_1)$ 和 $(X_2, X_1)$ 之间的数据构造Spline函数,对 $\pm X_1$ 之间进行插值,得到一个新的第一IMF,这样做的结果如图3所示。图中,粗实线为原来的第一IMF,细实线为用Spline拟合得到的第一IMF,虚线为细实线与粗实线数值之差,也就是提取出的异常信号。

将订正后的第一IMF与其余所有IMF叠加构成新的数据集,与原来的数据集相比,滤除了异常信号。在图4中,细实线是校正后叠加的结果,与余弦曲线相比,存在可以察觉的误差,最大误差为0.021,大约为1%。Spline拟合的结果略大于正常数据。

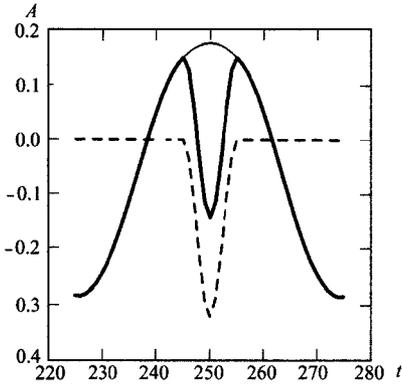


图3 异常信号的提取与模态拟合(对称情形)  
 (粗实线为原来的第一 IMF,细实线为用 Spline 插值拟合的第一 IMF,虚线为异常信号)  
 Fig. 3 extraction for abnormal signal and spline fitting for the first IMF (symmetric composing case). (Thick solid line is the first IMF, thin solid line the corrected first IMF, and dashed line for abnormal signal)

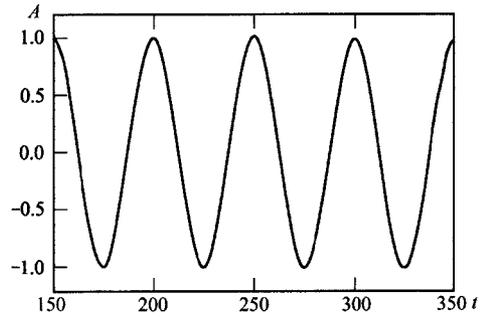


图4 模态拟合后的合成结果  
 (两条实线一条是正常信号的余弦曲线,另一条为拟合结果,结果显示差别很小)  
 Fig. 4 The result for removing abnormal signal and re-superimposing IMFs. One solid line is for the normal cosine curve and the other line is for the re-superimposed result. Both results have little differences

显然,IMF 的拟合精度和异常信号的提取精度是一回事,都取决于 Spline 拟合与正常信号的偏差。对于一般数据,我们无法知道这样做的精度,因为对非调和信号的拟合精度无法检验。由于误差是采用 Spline 插值造成的,如果采用别的插值方法,误差值会不一样。对于一般的数据,Spline 插值是比较可靠的插值方法。插值的精度在很大程度上取决于滤除数据的长度,滤除的时段越长,数据拟合的精度就越差。因此,这种方法过于简单,只能解决小时段脉冲型的异常干扰,对于时段较长,或包含几个周期的频率混叠,不能采用这种方法。然而,这种方法能有效地滤除异常扰动,保证了 EMD 方法在有异常干扰时可以得出有物理意义的结果。将图4 的拟合结果与其他 IMF 叠加后再进行 EMD 分解,就会得出单一模态的结果,因为拟合值偏高不增加极值数目,不能进一步分解。

对于一般的情况,异常信号不一定出现在正常曲线的峰值上,将异常信号向左移动,取

$$b = p_1 + \frac{p_2 - p_1}{T_1} t_1 + 0.15 \left[ 1 - \cos\left(\frac{2\pi t_1}{T_1}\right) \right] \quad (t_0 < t_1 < t_0 + T_1) \tag{3}$$

形成图5 最下面曲线的异常干扰,其中  $p_1$  和  $p_2$  分别为异常信号起迄点的值。图5 的分解结果表明,第一 IMF 不再对称,有些变形,但3 种位置对应特征依然存在。我们仍然采用上述的方法对第1 个 IMF 进行异常信号提取和模态拟合,得到的结果如图6 所示。拟合精度为 0.011,大约为 0.5%。与余弦曲线的比较几乎无法看出,因此不再给出图形。

## 2 应用实例

上面一节使用的正常数据是简单的余弦函数,校正后再作 EMD 分解只有一个模态,不能看到校正的整体效果。这里我们给出应用 EMD 方法滤除异常干扰的例子。

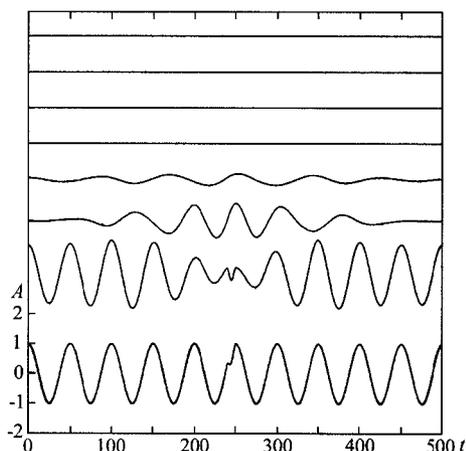


图5 用EMD方法分解的各个模态  
(与图2相同,但异常信号为非对称)  
最下面是原始信号,向上分别为1-6模态

Fig. 5 IMFs of signal

decomposed by EMD method.

The lower curve is the original signal and the upper curves are IMFs 1-6

图7最下面的曲线是美国阿拉斯加州 Barrow 的气球探空记录,数据为 1 000 hpa 的气温,记录长度从 1958 年持续到 1997 年,每月一个月平均温度数据。原始数据中反映出 1979 年 1 月有几天出现持续温度偏高的天气,使得月平均气温数据出现了一个微弱的局部极大值,这种发生在 1 月的极值在其它年份都不存在。这个气温异常现象在 EMD 方法中产生了频率混叠。看起来,这个异常值主要体现在第二模态,一段明显低于年周期的振荡出现在第 2 个 IMF 之中,然后扩展到后续模态。但是,实际上,这个信号也出现在第 1 个 IMF 之中,只是由于数据密集而不易观察到。将 1978 至 1979 年的异常信号放大为图 8 可以看到,第 1 个 IMF 完全表现了异常信号,第 2 个 IMF 的异常信号是次生的。第 2 个 IMF 整体上是年变化的结构。而第 2 个 IMF 的异常信号与年变化完全不协调。应该出现的年变化振荡出现在第 3 个 IMF 上,而且有展宽的趋势。同样,第 3 个模态反映的是准两年周期的变化,但在那段时间里有年变化插入,也产生了严重的不协调,使重要的准两年振荡表现不出来。这种异常干扰递推到后面所有的模态,连最后一个模态都受到这种

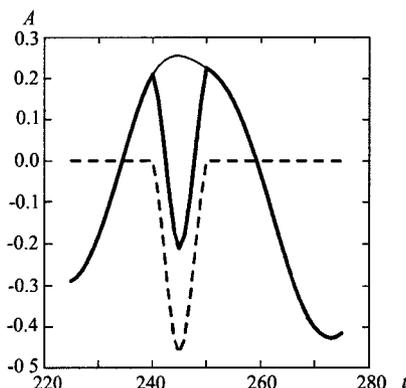


图6 异常信号的提取与模态拟合(非对称情形)  
(粗实线为原来的第一 IMF,细实线为用  
Spline 插值拟合的第一 IMF,虚线为异常信号)

Fig. 6 Extraction for abnormal signal and spline fitting for the first IMF (asymmetric composing case).

Thick solid line is the first IMF, thin solid line the corrected first IMF, and dashed line for abnormal signal

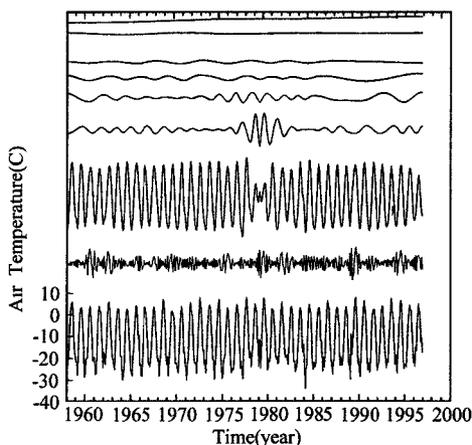


图7 Barrow 站 1 000 hPa 月平均  
气温的 EMD 分解结果  
(最下面的曲线是原始数据,  
其上是个个 IMF)

Fig. 7 Decomposed result by EMD method for 1 000 hPa monthly averaged air temperature in Barrow, USA. The lower curve is the original signal and the upper curves are IMFs

干扰的影响。

图 8 显示,第 1 个 IMF 中含有 1979 年 1 月异常信号的全部位置信息,可以用本文使用的方法进行校正。在这个例子中,因为异常只涉及 1979 年 1 月的单一结点,因此只需对这个点的数据进行校正,将异常值 +6.3 订正为 -4.0,然后将全部 IMF 叠加构成新的数据。新数据在该点的值从 -11.215 变为 -21.523,与第 1 个 IMF 的改变量一致。最后对校正的数据再次进行 EMD 分解,所得结果见图 9。从图 9 可以看到,1979 年 1 月的异常增暖事件已经有效地滤除,在第 2 个 IMF 中体现明显的年周期振荡,在第 3 个 IMF 中,准两年的变化的信号也非常明显,后面的 IMF 同样得到很好的结果。从图中可见,本文只校正了一个点的数据,但 EMD 分解的全部 IMF 都得到了明显改善。

对 1979 年 1 月的数据进行了校正,并不等于我们失去了原来数据的信息。将该点的异常信息提取了出来,即 1979 年 1 月 1 000 hpa 的气温偏高 10.3℃,这个异常信息可以另外使用。另外,出现的异常信号本来就是高频成分,在提取后应该合并到高频模态中,而不应该影响到年变化过程。

从图 8 也可以看出,并不是任何异常信号都将产生频率混叠,1978 年 11 月的信号峰值并没有对分解过程产生影响,只是作为一种高频信号存在于第一个 IMF。从作者对各种数据的分析结果来看,只有出现在峰值或谷值附近的反向叠加信号才有可能对 EMD 分解过程产生多种频率的效应,从而引起异常干扰。对异常事件校正的价值可以用 IMF 合成的办法进行估价。

图 10 给出了图 9 所示分解结果的两组合成结果,一组是周期大于年周期变化的合成,称为年际变化(上图),一组是多年变化的合成结果(下图)。从图中可见,虽然异常信号只涉及一个点,但在年际尺度上却影响数年的分析结果。未校正的分解结果组合存在相当大的不合理成分,出现较高频率的强振荡,组成的时间过程曲线没有实际意义;而经过校正后的 EMD 分解数据可以组成有价值的时间过程曲线。

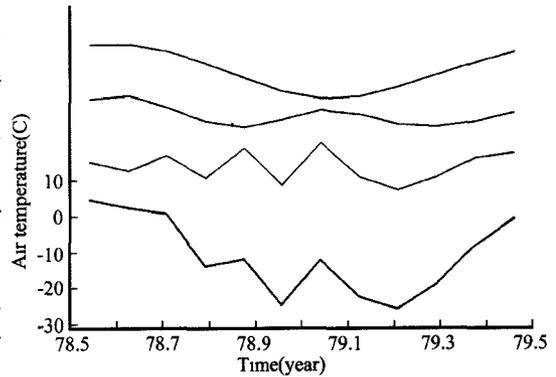


图 8 异常过程信号对 IMF 的影响分析  
(自下而上分别为原始信号,1~3 的 IMF)  
Fig. 8 Effect of abnormal signal on IMFs  
The lower curve is the original signal and the upper curves are IMFs 1~3

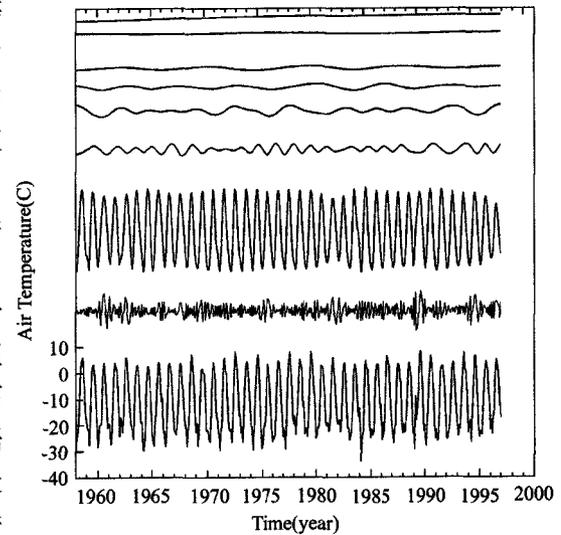


图 9 消除异常干扰后的 EMD 分解结果  
(数据与图 7 相同,但消除了异常干扰)  
Fig. 9 Decomposed result by EMD method for 1 000 hpa monthly averaged air temperature in Barrow  
Similar with Fig. 7 but with the abnormal signal removed

未校正的分解结果组合存在相当大的不合理成分,出现较高频率的强振荡,组成的时间过程曲线没有实际意义;而经过校正后的 EMD 分解数据可以组成有价值的时间过程曲线。

令人震惊的是,单点的数据异常不仅影响高频部分的结果,而且影响多年变化的信号。这是由于,异常信号在EMD分解中影响全部IMF,越高的IMF,影响范围越广,以至对低频结果也造成显著影响。从这个分析实例可以看到,对异常事件的数据校正是绝对必要的,因为异常信号在EMD分解过程中会向各个IMF扩展,导致整个结果失真。

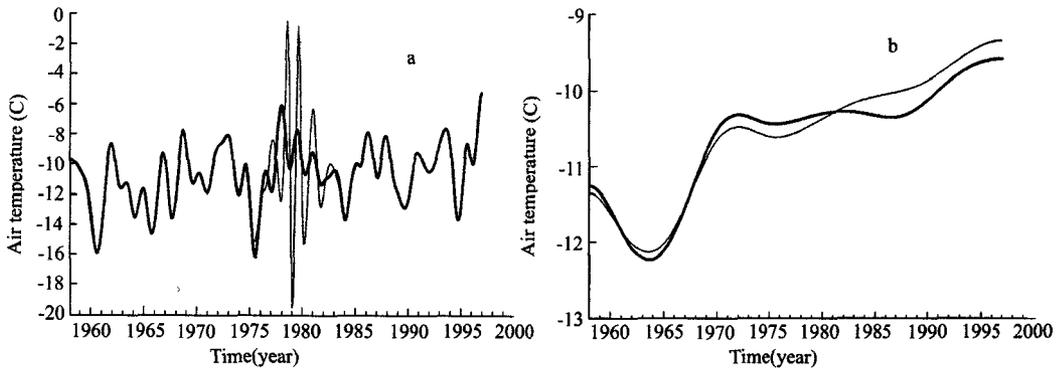


图10 校正前后分解结果的差异

(a. 年际以上尺度变化的合成;b. 多年变化的合成结果。细线——未校正的分解结果;粗线——校正后的分解结果)

Fig. 10 The differences of decomposed results with and without corrections.

Curves in fig. a are superposed with all IMFs longer than one year, and that in fig. b are superposed by IMFs longer than decadal period. Thin lines are for uncorrected results and solid line for corrected results

### 3 结果讨论

本文指出异常事件在数据中造成的干扰,形成发生在局部的高频信号。应用EMD方法分析这种存在异常事件干扰的数据就会形成IMF的频率混叠现象,使得IMF的物理过程不纯,同一频率的信号又混叠在不同的IMF中,形成了对结果分析与认识的困难,也无法用时间过程曲线表现特定的物理过程。这一问题在EMD方法中尚未妥善解决。

由于EMD方法本身是用于非稳态的方法,不能通过对原始数据校正的方法来消除异常事件干扰。本文提出了一个简单的办法来滤除异常干扰。因为相关IMF在干扰信号极值及其两边的极大与极小值与原始数据有明显的对应性,可望从相关IMF提取异常信息。将相关IMF中的异常信息直接滤除,再用Spline插值方法弥补滤除时段的数据,得到重新拟合的该IMF数据。这样做导致两个直接结果,第一,用重新拟合的数据减去原来的数据可以提取出异常信号,提取的精度与异常信号的时段长度有关,异常干扰时间尺度越短,提取的精度越高。第二,拟合结果消除了异常干扰,可以将该IMF与其余IMF一起叠加成新的数据集,成为没有异常干扰的数据。从本文的工作可以看出,所谓的“异常信号”,是指对特定分析方法或结果的特定表现形式产生异常影响的信号,因此,在EMD分解之前没有一般的办法判断异常信号,只有在第一次分解之后才能根据结果有无异常来判断异常信号,并予以校正。由此可见,本文是用分解结果来作为发现异常信号的判据。

将滤除了异常干扰的数据再次进行EMD分解,可以得到新的IMF系列。新的IMF可以与不加校正的分解结果有相当大的差别,显然更为合理和可靠。本文的工作表明,只有在有效滤除异常干扰的情况下才能获得可靠的IMF系列,真正准确地描述各种尺度的现象。消除了异常干扰的IMF可以任意单独或组合使用,表现各种时间尺度的变化与过程。

本文的方法只能用来提取并消除与正常信号反向叠加的异常信号,而不能提取正向叠加的异常信号。读者可能会问,反向叠加的异常信号可以提取出来,而正向叠加的异常信号只是因为发生位相略有差异就不能提取,是不是这种方法有问题。事实上,在数学上只能提取反向叠加的异常信号,在物理上也有重要背景。重视反向叠加信号在自然科学研究中是普遍存在的。例如:在严寒季节,连续5天的暖天气就成为显著的暖事件,但在炎热的夏季,连续5天的气温偏高天气并不认为是特殊事件。这是由于,反向叠加的信号反映了与现状相反的趋势,有“质的不同,受到格外关注;而正向叠加的信号只反映了现状“量”的变化,有时不能得到重视。从这个意义上讲,反向叠加的异常信号很可能代表了所有可能造成影响的异常信号,具有普遍的代表性。

本文讨论的方法只适合异常时段小于正常变化周期的情形,对于异常时段大于正常变化周期的干扰还需要探讨其他方法。

## 参考文献

- 1 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] Proc R soc Lond, 1998, 454:903~995
- 2 Magrin-Chanolleau I, Baraniuk R G. Empirical mode decomposition based time-frequency attributes. [C] Proc 69th SEG Meeting. Houston;TX,US,1999
- 3 Huang N E, Shen Z, Long S R. A new view of nonlinear water waves;the hilbert spectrum. [J] Ann Rev Fluid Mech, 1999,34:417~457
- 4 Zhu X, Shen Z, Eckermann S D, et al. Gravity wave characteristics in the middle atmosphere derived from the empirical mode decomposition method. [J] J Geophys Res, 1997,102(D14):16545~16561
- 5 Xie L, Wu Kejian, Pietrafesa L J. Empirical modes of landfalling tropical cyclones in North Carolina. [C] 24th Conference on Hurricanes and Tropical Meteorology, Fort Lauderdale;[s. n. ], 2000. P1:68
- 6 Vincent B, Hu J, Hou Z. Damage detection using empirical mode decomposition method and a comparison with wavelet analysis. [C] Proceedings of the Second International Workshop on Structural Health Monitoring. Stanford;[s. n. ], 1999. 891~900
- 7 Zhao Jinping, Huang Daji. Mirror extending and circular spline function for empirical mode decomposition method. [J] Journal of Zhejiang University(Science), 2001,2(3):247~252
- 8 Thompson R O R Y. Low-pass filters to suppress inertial and tidal frequencies. [J] J Physic Oceanogr, 1983,13: 1077~1083

## Study on the Effects of Abnormal Events to Empirical Mode Decomposition Method and the Removal Method for Abnormal Signal

Zhao Jinping

*(Second Institute of Oceanography, State  
Oceanic Administration, Hangzhou 310012, China)*

Roger Colony

*(International Arctic Research Center, University of Alaska Fairbanks, USA)*

**Abstract** Empirical Mode Decomposition (EMD) is a newly developed method used to analyzing non-steady and non-linear temporal process data. By EMD method, the original data can be decomposed into several Intrinsic Mode Functions (IMF). In time series data, local signals with higher frequency can often appear, which is usually caused by abnormal events. If EMD method is adopted to analyze this kind of data with abnormal signal, different frequencies will be mixed in each IMF. The frequency mixing causes the difficulty in understanding temporal processes. This problem has not been solved in EMD method.

In this paper, three corresponding characteristics for abnormal signal between the original data and the first IMF are noticed. Based on these characteristics, the start and end positions of an abnormal signal can be determined. Then the abnormal signal within both positions is removed directly. The gap left is then interpolated by Spline function, which is constructed by the known data outside the abnormal interval. By superposing the all IMFs including the revised first IMF, a new data set with the abnormal signal removed is formed. Decomposing the new data set by EMD method again, a set of new IMFs is obtained, which is quite different than the one with abnormal signal. Our study shows that the abnormal signal removal is absolutely important in getting reasonable decomposing result and in displaying the temporal processes with different time scale. This is also a method for information extracting of abnormal signal.

This method is suitable only for the short interval abnormal signal. Other methods are still needed for removing the abnormal phenomena with the interval longer than the time scale of background process.

**Key words** Empirical Mode Decomposition Method; abnormal event; intrinsic mode function; high-frequency signal