

固体潮观测数据的预处理

唐 九 安

(兰州地震研究所)

摘 要

本文介绍了固体潮观测数据的平滑检验和NAKAI法检验之方法原理和具体应用。平滑检验可以有效地监测并改正观测数据中的随机错误, 突跳或坏读数, 经平滑改正后, 可使观测数据的读数误差小于1毫米。用NAKAI(1975)提出的方法检验观测数据可以监测仪器灵敏度变化情况和数据质量好坏, 根据该检验计算获得的有关参数(主要是漂移系数和方差)剔除质量不好的48小时数据组或修正某些质量不好的数据可提高调和分析结果的精度。

一、数 据 平 滑

固体潮实测曲线是一条连续光滑的曲线, 它的谐波成分主要有周日波, 半日波, 三日波和长周期波, 同时还包含了仪器的漂移。要作调和分析, 首先需将记录图象数字化。处理长时间系列的观测资料, 将涉及大批数据; 在这大批数据中, 即使你倍加小心, 反复核对, 最终也难以避免错误数据的存在; 但这些错误数据掺杂到分析计算中去, 就将使结果严重失真。为了寻找这些错误方便, 或者在计算时找出这些错误数据后, 机器能有效地自动加以更正, 不致因有错误数据影响计算结果, 在应用维尼迪科夫数字滤波器以前, 对观测数据实施平滑检验可以有效地做到这一点。

数据平滑的原理和方法如下

1. 用坐标组合公式 $Z_{1/2}^2 (Y_3 - Y_0/2)$ (LECOLAZET, 1959) 对日观测系列作两次组合。两次组合间隔4小时。该组合对连续24小时观测值系列的乘系数依次为 -1, 5, -10, 11, -10, 11, -11, 10, -11, 10, -5, 1, -1, 5, -10, 11, -10, 11, -11, 10, -11, 10, -5, 1。经该组合后, 消除了潮波成分和漂移, 剩余值 R_{24} 即为观测误差的线性函数值, 按误差传布定律有:

$$R_{24}^2 = 1872\sigma^2 \quad (1)$$

式中 σ 为读数误差。一般人工采样其读数误差可取 0.3~0.4 毫米。如取 $\sigma_{max} = 0.7$ 毫米: 则有:

$$R_{24max} = 30 \text{ 毫米}$$

如果这一天的两个组合值均不超过某一预期的限值, 则表明这一天读数中无明显错误。但

是,只要其中任何一个组合值超过限值,即表明这一天数据中有错误,所以用上述坐标组合公式可以判断某一天的数据中是否有错误。

2.当确定某一天的数据中有错误后,就对这一天的数据转作 $Z_{1/2}^i$ 组合(LECOLAZET, 1959)。该组合对连续6小时读数的乘系数为, $-1, 5, -10, 10, -5, 1$ 。其组合值 R_0 也是消除了潮波成分和漂移后的观测误差的线性函数值,按误差传布定律有

$$R_0^2 = 252\sigma^2 \quad (2)$$

显然可得

$$R_{0max} = 0.37R_{24max}$$

对这一天数据多次使用该组合,就可以找出错误数据的具体位置。

下面我们来看如何利用该组合找错误数据的具体位置。

对连续30个读数作 $Z_{1/2}^i$ 组合,滑动2小时作一次所作13次。约定 R_j 为其中绝对值最大的一个组合值, Y_k 为这30个数据中的第一个数,假定 Y_k 和 Y_{k+1} 没有错(有错也在上一天平滑时改正了),还假定,如果错误数据在 Y_{k+2} 以后,即 $j=13$ 时,这一天的平滑过程也不对它施行改正,留待下一天的平滑过程再去改正它。也就是说,当 $1 < j < 13$ 时,就在 j 次组合的基础上,前后滑动1小时各作一次 $Z_{1/2}^i$ 组合,令其组合值为 R_{j-1} 和 R_{j+1} ,比较这两个值之绝对值的大小,如果 $|R_{j-1}| < |R_{j+1}|$ 则错误数据为 Y_{K+2j+2} ,反之为 Y_{K+2j+1} 。

3.利用上述方法找到错误数据的位置后,第三步,就对错误数据施行平滑改正。改正的方法是使用如下两个插值公式

$$Y_1^i = \frac{1}{8} [4(y_{i-24} + y_{i+24}) - y_{i-48} - y_{i+48}] \quad (3)$$

$$y_1^i = \frac{1}{2} [1.364(y_{i+1} + y_{i-1}) - 0.364(y_{i+2} + y_{i-2})] \quad (4)$$

设每段数据的总个数为 N ,错误数据的序号为 I ,当

$$48 < I \leq N - 48$$

时,可同时使用上述两个插值公式,分别求得 Y_1^i 和 Y_1^i ,取 $Z_{1/2}^i$ 组合值之绝对值较小者取代错误数据。当

$$I \leq 48 \text{ 或 } I > N - 48$$

时,只能使用公式(4)对错误数据进行改正。

公式(4)中的系数是基于理论值用最小二乘原理求得的。将上述两个插值公式用于理论值,统计了30个点的插值与原有值之差的绝对值。由(3)式求得插值的平均偏差为0.3微伽,最大不超过0.8微伽;用(4)式求得平均偏差为0.1微伽,最大不超过0.2微伽。

关于 R_{24} 值的限值问题,取决于观测条件。我们的经验是,对于GS₁₆型重力仪记录资料,可取70~100微伽。

为了进一步说明如何寻找错误数据的位置和改正的方法,现举如下一个例子。

设有如下一个观测值系列, 77.6, 74.8, 79.2, 91.0, 106.0, 131.0, 130.2, 132.2, 125.7, 110.7, 91.0, 72.1, 57.1, 49.9, 52.7, 64.8, 83.0, 102.2, 119.0, 129.2, 131.0, 125.0, 112.2, 98.2 (毫米)…。对这24个读数,施以 $Z_{1/2}^i, Z_1^i(Y_3 - Y_{0/2})$ 组合得:

$$R_{24} = 89.2 \text{ (毫米)}$$

显然其中有错误数据。将 $Z_{1/2}^i$ 组合用于该系列,依次得 R_0 的值为

$$15.4, 107.0, 51.2, -2.1 \dots$$

显然 $j=2$ 。再对 $Y_2 \dots Y_7$ 和 $Y_4 \dots Y_9$ 这两个系列作 $Z_{1/2}^i$ 组合得

$$R_{j-1} = -53.6, R_{j+1} = -104.3$$

由此可知是 Y_0 错了。用两种插值公式求得该点的插值分别为124和120.5, 将此二值取代 Y_0 。

$$\text{求得 } R_1^1 = 37.2, R_1^2 = 2.0$$

所以最后取120.5作为 Y_0 的平滑值。对照图纸检查该点的正确读数为121.0。

二、NAKAI检验及其应用

数据平滑对于监测和改正观测数据中的随机错误是行之有效的。数据经过上述方法平滑以后, 随机错误都能得到更正。但是对于记录中带有某种系统性的非潮汐干扰, 如灵敏度波动, 异常漂移等, 要进行监测, 上述平滑方法就做不到。为此, 我们先应用NAKAI提出的方法对相当于维尼迪科夫数字滤波器间隔的每个独立48小时组进行粗略的最小二乘拟合, 求出每个组的视灵敏度因子, 漂移系数和方差。

在该法中, 设固体潮的观测值可以写成如下形式

$$y_t = aR_t + b \frac{dR}{dt} + \sum K_j t^j \quad (5)$$

式中, a 是仪器视灵敏度因子, R_t 是时刻 t 的潮汐理论值, $\frac{dR}{dt}$ 为理论值的变化率, K_j 为多项式漂移系数。显然有

$$b = -a\Delta t \quad (6)$$

假定该地潮汐因子 δ 不随时间变化, 在48小时间隔内, 仪器灵敏度因子为常量, 则有

$$\sigma = a \cdot S \quad (7)$$

S 为记录格值。对于每个48小时组, 可以列出形如(5)式的48个方程。因为 j 取0, 1, 2, 则未知数只有5个, 所以将最小二乘原理用于这48个方程的方程组, 就可解出未知数 a, b, K_0, K_1, K_2 。並可由下式计算方差

$$M_0 = ([vv]/43)^{1/2} \quad (8)$$

式中 $[vv] = \sum_{i=1}^{48} (Y_t^{\text{观}} - Y_t^{\text{计}})^2$ 。由若干个独立48小时组的视灵敏度因子和实测记录格值

就可求得该地的潮汐因子的最可几值 $\bar{\delta}$ 。进而由 $\bar{\delta}$ 和式(7)可求得每个48小时组一个计算格值(或称NAKAI平滑格值)。

如果仪器稳定, 记录正常, 则各个48小时组的视灵敏度因子和漂移变化不大, 则方差也较小。计算结果表明, 当漂移和方差较大时, 其相应的48小时的数据质量也较差, 无疑这些数据对调和分析计算结果影响也较大。例如GS15—218重力仪在兰州的记录, 从1980年3月13日至6月18日共有80天的记录资料, 当这些资料全部参加运算时, 一个 M_i 和 N_i 的均方差为7.9微伽, 而当把漂移和方差过大的几个48小时组去掉以后, 其 M_i 和 N_i 的均方差降到3.2微伽。所以, 我们可以把方差和漂移系数作为衡量数据质量好坏的主要指标。对于用维尼迪科夫法作调和分析, 在正则方程组成之前, 我们主要依据这几个参数的大小来决定各个48小时数据的取舍。对于用莱柯拉兹方法作调和分析, 因为715个小时内数据不能中断, 所以不能舍弃数据, 为了提高分析结果的精度, 则可用(5)式对其不好的数据给予修正。

在我们的维尼迪科夫调和分析程序中, 取舍资料的具体作法如下:

将NAKAI法检验求得每个48小时组的有关参数连同该组中央时刻的儒略日和经纬度迪科夫数字滤波器滤波以后的 M_i 和 N_i 共同存放在一个命名为MN的数组中,视不同精度的观测资料给定舍弃资料的方差,二次漂移和一次漂移的限值。将超过限差的资料统统去掉后,重新排列MN数组,再列观测方程式组,求解最后结果。

关于舍弃资料的有关参数的限值,提供如下几个指标供读者参考。

对于GS15—218重力仪在兰州的记录资料,我们根据某段数据各个48小时组各有关参数的数字平均值的两倍,即取 $K_1 = 3$ 微伽/小时, $K_2 = 0.03$ 微伽/小时², $M_0 = 3.2$ 微伽作为保留资料的限值。对于其它不同精度的记录资料,则应视不同的质量情况灵活改变这些参数限值的大小。这些参数限值之选取的总的原则就是,在充分利用所有观测资料的前提下,尽可能把质量不好的48小时组舍弃,以使最后结果能充分地反映观测结果的精度。

(本文1980年11月10日收到)

参 考 文 献

- [1] Paul Melchior. Harmonic Analysis of Earth Tides. Method in Computational Physics.
- [2] B. Ducarme. The Computation Procedures at the International Center for Earth Tides.
- [3] J. Kaariainen and B. Ducarme. the Finnish Tidal Gravity Registrations in Scandieavia.
- [4] 宋兴黎,对固体潮观测中几个问题的研究,中国科学院测量与地球物理研究所,1980. 6.
- [5] 唐九安,固体潮调和分析程序(维尼迪科夫法),国家地震局兰州地震研究所,1980. 7.

THE PRE-PROCESSING METHOD OF THE OBSERVATION DATA OF EARTH TIDES

Tang Jiu-an

(Seismological Institute of Lanzhou)

Abstract

~~This paper~~ deals with the theory and application of the pre-processing method in the harmonic analysis of earth tides. Before using the ~~Venedikov filters~~, ^{first} we ^{applied} ^{two-step} error detection to observation data. ~~First, the data was smoothed by means of several filters. Then the occidental fault data, jump data or bad reading among the observation data were corrected, and thus the reading errors of observation data were not more than 1 mm. Secondly, a rough least square adjustment was made on each 48 hour interval that would be covered by a Venedikov filter {Nakai, 1975}. Then the bad groups of 48-hour interval could be detected out and a smoothing scale for each group could be given. By adopting this method the bad groups caused by anomalous drift in a short time (for example, 3—5 hours) might be corrected. Finally, the rest of bad groups should be discarded in the computation of final results. Through the two step error detection, the accuracy of the computational results was promoted.~~