

曲线拟合法在地电场资料处理中的应用

王兰炜¹, 赵家骝², 席继楼², 王燕琼²

(1. 中国地震局地壳应力研究所, 北京 100085; 2. 中国地震局分析预报中心, 北京 100036)

摘要: 地电场和电阻率同场地观测时, 电阻率观测会干扰地电场的观测数据发生大幅度变化. 本文主要探讨通过曲线拟合的方法消除这种干扰及其它随机干扰, 并给出了对实际观测数据处理的结果.

关键词: 地电场观测; 电阻率观测; 干扰; 曲线拟合

中图分类号: P319.3 **文献标识码:** A **文章编号:** 1000-0844(2003)02-0131-05

0 引言

目前在许多地震综合前兆观测台站, 地电阻率和大地电场在同一场地观测. 由于地电阻率观测时需要人工向大地供电, 在一定的范围内产生人工电场. 而地电场观测仪器将记录到这个人工电场. 在一些观测环境中, 这种人工电场可能会比正常的地电场变化大几倍甚至十几倍, 严重干扰了地电场的正常观测, 造成地电场的观测数据出现瞬间大幅度的变化, 使原本应该比较平滑的观测数据曲线出现了许多“毛刺”. 这就影响了这些观测台站地电场观测资料的准确性和连续性, 从而影响到观测资料内在质量. 因此消除地电场观测数据中这种干扰就具有十分重要的实际应用价值. 本文利用了最小二乘曲线拟合的方法消除实际观测数据中这种干扰, 并取得了较好的效果.

1 地电场和地电阻率观测

地电场观测是在观测场地某一方向 AB, 测定 AB 间的电位差 U_{AB} , 则 AB 间的电场强度 $E_{AB} = -\frac{U_{AB}}{AB}$. 采样率是 1 次/分.

地电阻率观测是在观测场地某一方向埋设 4 个电极 A、M、N、B(图 1), 从 AB 间供入电流 I , 在 MN 测量由 I 产生的人工电位差 ΔV . 根据公式(1)求得地电阻率值

$$\rho_s = \frac{1}{K} \times \frac{\Delta V}{I} \quad (1)$$

其中:
$$K = \left[\frac{1}{AM} + \frac{1}{BN} - \frac{1}{AN} - \frac{1}{BM} \right]$$

地电阻率是每小时观测一次, 观测时供电电流周期是几秒钟, 整个观测过程中的总供电时间约为五分钟. 因此, 地电场每小时大约有五个, 每天大约有 120 个观测数据是受到供电干扰的. 由于有时地电场观测采样正好在供电间隔时间上, 所以实际被干扰的数据点数要少于 120 个. 由于两种观测仪器不同步独立工作, 被干扰数据出现时间和干扰幅度都具有随机性, 产生的干扰如图 2 所示.

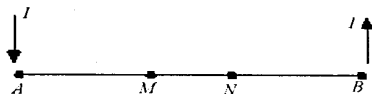


图 1 地电阻率观测示意图

Fig. 1 Sketch of resistivity observation.

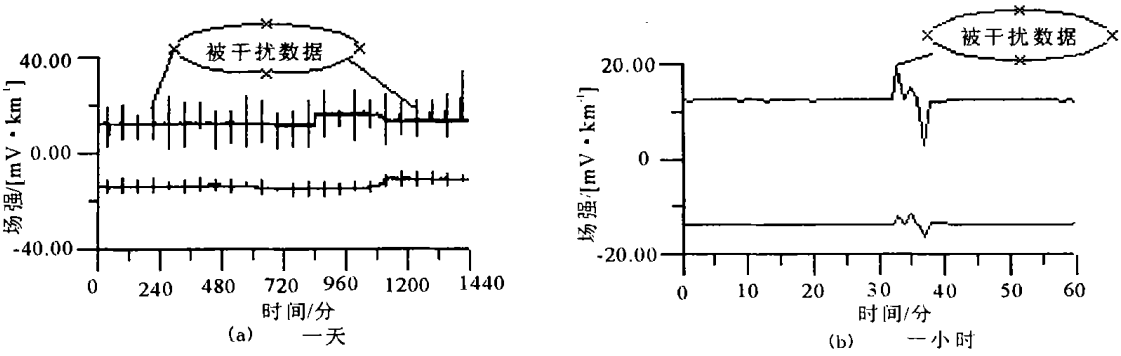


图 2 被干扰的地电场观测数据
Fig. 2 Interfered waveform of observation data of telluric.

如果直接剔除被干扰数据,会造成地电观测数据的不连续,给进一步处理数据造成影响. 所以将这些被干扰数据从原始观测数据中剔除以后,要采用恰当的方法进行插补.

2 被干扰观测数据的判别

目前地电场观测仪器的频带为 10S—DC, 获取的是地电场缓慢变化的部分. 因此在连续的观测中, 观测数据在短时间内不可能出现剧烈的起伏变化. 当受到供电干扰时, 观测数据会出现幅度较大, 速度较快的变化, 出现的时间和幅度的大小都具有随机性, 因此可以利用这一特点来判别观测数据是否受到供电干扰.

首先根据观测数据的一阶差分给出一个阈值 $D(n)$, 作为控制限来判断观测数据是否受到供电干扰. 此阈值并不是一个固定值, 它随着新的观测数据的输入在不断变化, 判别门限为 $\lambda D(n-1)$. 对于第 n 个观测数据 $X(n)$, 利用下式求出它的一阶差分的绝对值

$$D(n) = |X(n) - X(n-1)| \tag{2}$$

如果满足条件: $D(n) \leq \lambda D(n-1)$, 则认为此观测数据是未受供电干扰的数据; 反之则认为此观测数据是受到供电干扰的, 由原始数据中剔除, 并记录相应的数据点. 其中 λ 为一常数, 根据各个台站不同干扰幅度, 选择不同的 λ 值. 一般选择在 14 之间. 图 3 是判别观测数据是否受到供电干扰流程.

为防止 $D(0)$ 过大, 使 $D(1)$ 总是满足条件 $D(1) \leq \lambda D(0)$, 首先要设定 $D(0)$. 方法是: 首先计算整个数据序列的平均值及均方差, 大于 2 倍均方差的数据认为是粗大误差或被干扰数据, 剔除; 再计算其余数据的平均值及均方差; 求出第一个数据与平均值的差, 若差值大于 2 倍的均方差, 则用下一个数据减去平均值, 直到差值小于 2 倍的均方差为止, 以此差值作为初始阈值 $D(0)$.

当 $D(n-1)$ 太小或是等于 0 时会出现误判, 因此必须给 $D(n-1)$ 设定一个最小值, 当实际值小于此最小值时, 就用最小值代替 $D(n-1)$. 最小值可以根据均方差给定.

通过这种方法, 可以判定出受到供电干扰的观测数据, 并记录其相应的观测时间.

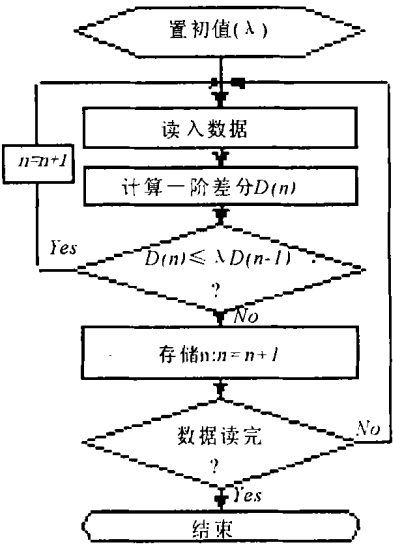


图 3 干扰数据判别流程图
Fig. 3 Flowchart of identifying interfered data.

3 被干扰观测数据的处理

3.1 信号的线性最小二乘拟合

对于一组测量数据 $(X_i, Y_i) (i = 1, 2 \cdots n)$, 假设拟合函数为 $y = f(x, a)$, 其中 $a = (a_1, a_2, a_3, \cdots, a_n)$ 是参数, 求 a 使得拟合函数 $f(x, a)$ 在 X_i 点计算出的函数值与对应数据差值的平方和的平均值最小, 即

$$\sum_{i=1}^n (f_i - y_i)^2 = \min \quad (3)$$

这就是最小二乘数据拟合问题.

对被干扰的地电场观测数据采用最小二乘法进行拟合计算时, 选取多项式 $P(x) = \sum_{i=0}^n a_i x_i^i$ 作为拟合函数. 通过选取适当的系数 a_i , 就可使 $S = \sum_{i=0}^n [p(x_i) - y_i]^2$ 达到最小.

当 $\frac{\partial S}{\partial a_i} = 0$ 时 $(i = 0, 1 \cdots n)$, S 有极小值. 通过推导, 可以得到一个如下的矩阵:

$$\begin{bmatrix} \sum X_i^0 & \sum X_i^1 & \cdots & \sum X_i^n \\ \sum X_i^1 & \sum X_i^2 & \cdots & \sum X_i^{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \sum X_i^n & \sum X_i^{n+1} & \cdots & \sum X_i^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum y_1 x_i^0 \\ \sum y_2 x_i^1 \\ \vdots \\ \sum y_n x_i^n \end{bmatrix} \quad (4)$$

设

$$X = \begin{bmatrix} \sum X_i^0 & \sum X_i^1 & \cdots & \sum X_i^n \\ \sum X_i^1 & \sum X_i^2 & \cdots & \sum X_i^{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \sum X_i^n & \sum X_i^{n+1} & \cdots & \sum X_i^{2n} \end{bmatrix}; \quad A = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}; \quad Y = \begin{bmatrix} \sum y_1 x_i^0 \\ \sum y_2 x_i^1 \\ \vdots \\ \sum y_n x_i^n \end{bmatrix}$$

其中 A 为待估的多项式系数, X, Y 可以根据观测数据算出, 则可以求得相应的拟合多项式系数 A .

$$A = YX^{-1} \quad (5)$$

式中 X^{-1} 为 X 矩阵的逆矩阵.

在实际工程应用中, 一般选择 n 不大于 5.

3.2 模拟干扰数据的产生

模拟干扰数据 $D(i)$ 是由正常的、未受地电阻率观测供电干扰的地电场观测数据 $X(i)$ (图 4) 和模拟实际的供电干扰数据对应相加而得. 为了突出干扰, 把干扰数据的幅度定为 200 mV/km. 其时间序列如图 5.

3.3 数值模拟数据的处理结果及分析

对数值模拟观测数据 $D(i)$ 采用最小二乘拟合方法进行处理; 再计算出处理的数据与原始数据 $X(i)$ 差值的百分比, 即

$$\frac{1}{n} \sum \left| \frac{x_2(i) - x(i)}{x(i)} \right| \quad (6)$$

式中 $x_2(i)$ 代表拟合后的数据系列; $x(i)$ 代表原始数据系列; N 为一天的数据个数 ($N = 1440$).

这个百分比反映了处理后的数据与原始数据的偏离程度. 根据对安丘地震台任选 12 天的正常数据加入干扰, 经过以上方法处理, 得到的相对偏离程度在 0.001 40.010 2 之间, 平均为 0.004 5.

图 6 为采用二次多项式拟合对数值模拟观测数据的处理结果, 与原始曲线 (图 4) 很相近, 消除了被供电干扰的数据.

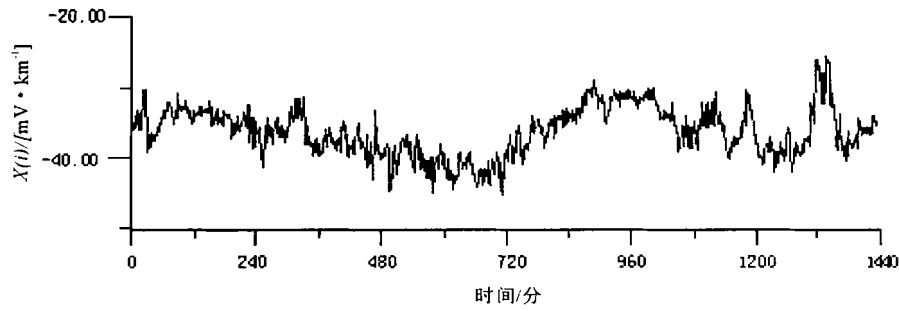


图 4 正常的地电场观测数据曲线

Fig. 4 The normal waveform of observational data of telluric.

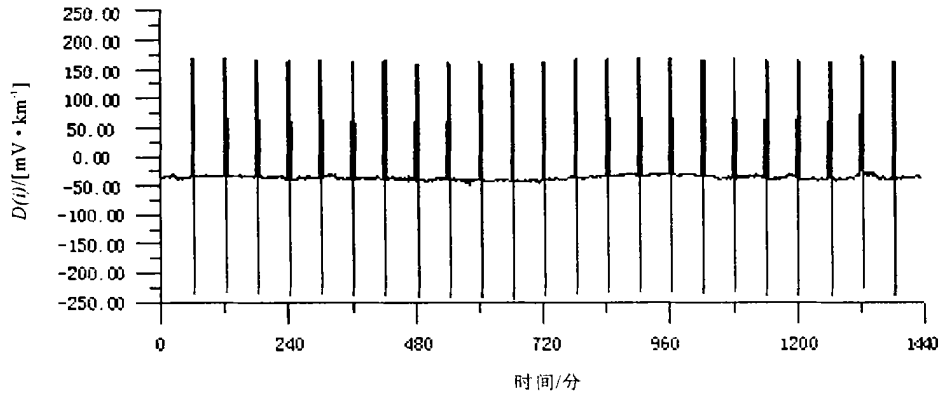


图 5 被干扰的地电场观测数据曲线

Fig. 5 The interfered waveform of observational data of telluric.

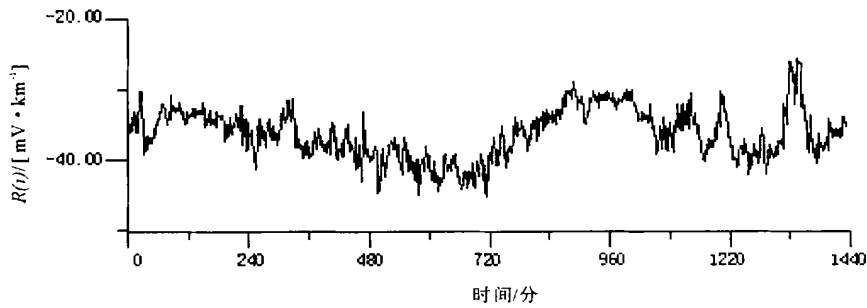


图 6 二次多项式拟合后的数据曲线

Fig. 6 Curve of observational data after the quadratic fitting.

由以上分析可以得出如下结论:对于包含有少量脉冲干扰的地电场观测数据,可以采用最小二乘拟合方法对这些少数被干扰的地电场观测数据进行插补处理. 处理后的数据基本不影响它原有的信息量和变化形态.

4 实际观测数据的处理结果

对宝坻地震台和榆树地震台的实际观测数据采用上述方法的处理结果见图 78. 可以看出, 在干扰的影响下, 数据曲线的趋势变化难以观察; 而在消除干扰之后, 数据曲线的趋势变化就比较清晰了, 干扰消除的效果很理想.

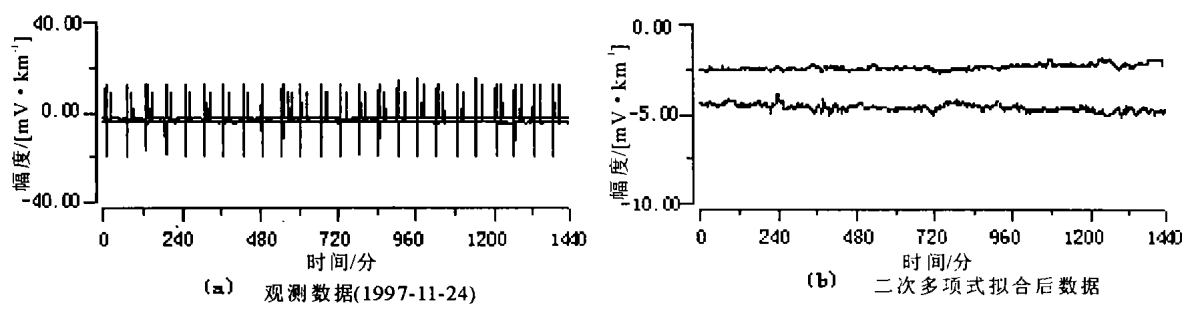


图7 宝坻地震台地电场观测数据和拟合后数据曲线
Fig. 7 Curves of telluric data before and after quadratic fitting at Baodi station.

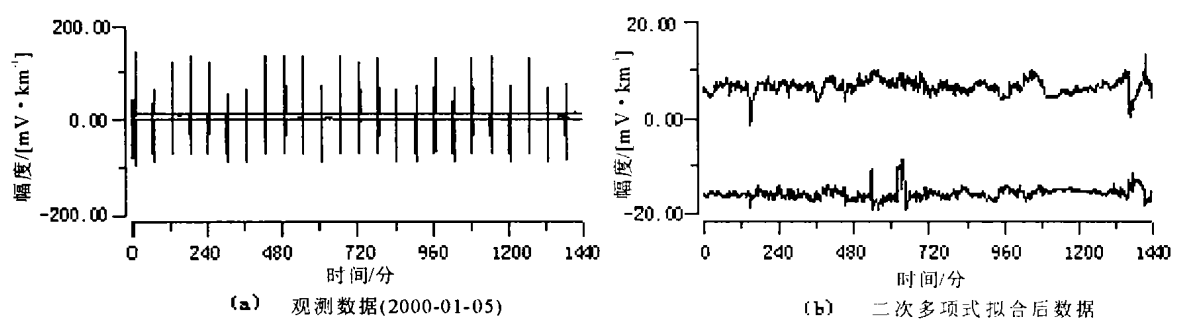


图8 榆树地震台地电场观测数据和拟合后数据曲线
Fig. 8 Curves of telluric data before and after quadratic fitting at Yushu station.

目前, 该方法的应用程序已编写完成, 并已嵌入“地电数据处理软件”在各地电观测台站实际观测中使用.

[参考文献]

[1] 程乾生. 信号处理的数学原理[M]. 北京: 石油工业出版社 1993.
[2] 战同胜. 实用数值计算方法[M]. 大连: 大连理工大学出版社, 1992.

APPLICATION OF LEAST SQUARE FITTING METHOD
TO PROCESSING TELLURIC DATA

WANG Lan-wei¹, ZHAO Jia-liu², XI Ji-lou², WANG Yan-qiong²
(1. Institute of Crustal Dynamics, CSB, Beijing 100085, China;
2. Center for Analysis and Predication, CSB, Beijing 100036, China)

Abstract: To be observed on same field in same time, telluric observation is interfered by electric current in measuring the apparent resistivity, and the data are changed to square impulse. In this paper, the method of least square fitting which can be used to eliminate this interference and other random interference in telluric data is discussed. The practically processed result is given also.
Key word: Telluric observation; Apparent resistivity observation; Interference; Fitting method