

基于主成分分析的航空电磁数据噪声去除方法

朱凯光, 王凌群, 谢 宾, 王 琦, 程宇奇, 林 君

(吉林大学 仪器科学与电气工程学院, 长春 130026)

摘 要: 时间域航空电磁数据经预处理后, 存在残余噪声, 严重影响晚期道数据质量。提出一种基于主成分分析的去噪方法, 通过对航空电磁剖面数据的自相关矩阵进行奇异值分解, 得到特征值及对应的特征向量, 将电磁数据通过旋转矩阵(特征向量矩阵的转置)进行线性映射, 得到其主成分。大特征值对应的低阶主成分反映相关性较强的电磁信号, 而小特征值对应的高阶主成分反映不相关噪声。采用低阶主成分与特征向量重构电磁数据能够去除不相关噪声。采用主成分分析法分别进行了仿真数据与实测数据的主成分特征分析与噪声去除。仿真数据去噪后, 信噪比提高了 13 dB; 而野外飞行实测数据的噪声水平也由 ± 25 nT/s 降低到 ± 8 nT/s。

关键词: 时间域航空电磁; 主成分分析; 特征值; 特征向量; 去噪

中图分类号: P631

文献标志码: A

Noise removal for airborne electromagnetic data based on principal component analysis

ZHU Kai-guang, WANG Ling-qun, XIE Bin, WANG Qi, CHENG Yu-qi, LIN Jun

(School of Instrumentation and Electrical Engineering, Jilin University, Changchun 130026, China)

Abstract: There is still residual noise which affects the quality of later channel data after preprocessing for time domain airborne electromagnetic data. An approach was proposed to remove the residual noise based on principal component analysis. The principal components were computed through the rotation matrix which is the transpose of eigenvectors matrix. The low-order principle components associated with the big eigenvalues reflect the correlated electromagnetic signals, while high-order principle components associated with the small eigenvalues are corresponding to the uncorrelated the noise. Therefore, the electromagnetic data are reconstructed by suitable number of the low-order components to remove uncorrelated noise. The experimental results of the simulation data show that the SNR is improved of 13 dB. The peak to peak value of the latest two channels for the field survey profile data is reduced from ± 25 nT/s to ± 8 nT/s after noise removal.

Key words: time domain airborne electromagnetic method; principal component analysis; eigenvalue; eigenvector; noise removal

时间域航空电磁探测具有勘探深度大、探测范围广、勘查速度快等优势, 广泛用于矿产资源勘查以及地下水、油气探测等。但航空电磁探测的飞行测量方式容易引起发射线圈、接收线圈晃动、震动而导致系统参数变化, 同时飞行速度、飞行高度、飞行环境的变化也能够引起系统参数变化, 从而引入系统测量噪声, 直接影响航空电磁数据有效时间窗以及反演解释

效果, 特别是对于反映地球深部信息的晚期道数据, 影响更为严重。

国外学者针对航空电磁探测过程中的多种噪声, 如线圈运动噪声、天电噪声、电力线噪声、甚低频电磁噪声、地质噪声、仪器内部电子噪声以及人文噪声等, 分析了噪声的时频特性^[1], 研究了噪声的压制算法^[2-6]。在我国, 随着航空电磁探测技术与仪器研究工

基金项目: 国家高技术研究发展计划项目(2013AA063904); 国家自然科学基金资助项目(41274076); 国家重大科研装备研制项目(ZDYZ2012-1-03)

收稿日期: 2013-05-16; 修订日期: 2013-06-30

通信作者: 林 君, 教授; 电话: 0431-88502473; E-mail: lin_jun@jlu.edu.cn

作的深入,我国学者也开展了航空电磁噪声的背景场去除、数据滤波、叠加、抽道等预处理方法的研究^[7-8]。虽然时间域航空电磁探测的原始数据,经过多种预处理与滤波后,剖面数据的信噪比得到了极大提高,但晚期道仍然含有残余噪声,影响地下深部的探测能力。

传统的航空电磁数据处理,采用时域(叠加、抽道等)和频域(小波变换、滤波等)方法,分离包含地下异常信息的电磁信号与噪声。但是,电磁信号与噪声在时域、频域上存在一定的重叠,因此,常规数据处理方法得到的剖面数据仍然存在少量的残余噪声。主成分分析(Principle component analysis, PCA)是一种统计分析方法,利用信号的相关性,通过正交映射,将一组相关性较强的变量转换为一组不相关的主成分。相关性较强信号的能量主要集中在低阶主成分,因此,利用低阶主成分重构电磁数据,压制不相关噪声,解决了数据中噪声与信号的频谱重叠等问题。

20世纪90年代就将数据的主成分分析引入地球物理数据的噪声压制与地下特征显示研究中。JONES等^[9]以及 JACKSON 等^[10]利用主成分分析,成功地压制了地震数据中的不相关噪声;GREEN^[11]提取时间域航空电磁数据的主成分,并利用主成分的 RGB 图反映测区电阻率分布;MINTY 等^[12]对放射数据的频谱提取主成分特征,实现了谱特征的不相关噪声压制;ASTEN^[13]利用主成分的去噪、降维特点,分析时间域电磁数据,进行未爆炸物识别;KASS 等^[14-15]不仅采用主成分压制时间域电磁数据中的噪声,而且将其引入反演计算中。

近年来,我国学者也开展了主成分去噪的相关研究,常威威等^[16]利用主成分分析对 Contourlet 系数进行消噪处理,改善了高光谱图像质量;吴锡等^[17]使用非局部主成分分析,克服了现有局部性去噪方法模糊边界等缺陷,提高了对图像细节信息的复原能力;姚莉丽等^[18]利用主成分分析法有效地去除乘性噪声;傅东浩等^[19]利用主成分分析法的去噪和去冗余特性,实现了无线传感器网络压缩感知算法;朱凯光等^[20]将主成分分析引入时间域航空电磁反演,由于主成分的去噪声特性,在含噪数据的主成分反演中,取得了优于其他方法的反演结果;刘霖等^[21]利用主成分的去冗余特性,进行了湘西喀斯特石漠化地区的生态安全评价,有效地揭示了石漠化形成过程的因子综合作用和主导作用。

本文作者根据主成分分析原理,计算了时间域航空电磁数据的主成分,通过分析剖面数据与残余噪声的主成分特征,确定重构电磁数据的主成分的数目,

实现有效去除剖面数据残余噪声的目的。最后,给出了仿真数据与实测数据的噪声去除结果,以验证本研究算法的有效性。

1 主成分分析法

1.1 主成分分析原理

主成分分析是一种多元统计分析方法,将一组多元数据分解为一系列能量递减的正交基。设一组 m 道时间域电磁数据 $x_i(t_j)$, $i=1, 2, \dots, n$; t_j 为采样时间, $j=1, 2, \dots, m$; 其自相关矩阵 Γ 的矩阵元素 γ_{jk} 由式(1)计算得到:

$$\gamma_{jk} = \sum_{i=1}^n x_i(t_j)x_i(t_k), j, k=1, 2, \dots, m \quad (1)$$

利用奇异值分解法,将自相关矩阵 Γ 分解为特征值与特征向量的形式:

$$\Gamma = \mathbf{R} \mathbf{A} \mathbf{R}^T \quad (2)$$

式中: $\mathbf{A}$ 为由大到小排列的特征值对角矩阵; $\mathbf{R}$ 为各个特征值 j 和 k 对应的特征向量矩阵,代表数据能量递减的各个主成分分量的方向; $\mathbf{R}^T$ 又称为旋转矩阵。

航空电磁数据(矩阵 $\mathbf{X}$)经旋转矩阵 $\mathbf{R}^T$ 进行线性映射,得到能量递减的各个主成分方向的系数,称为第 1, 第 2, $\dots$, 第 m 个主成分(矩阵 Ψ):

$$\Psi = \mathbf{R}^T \mathbf{X} \text{ 或 } \Psi_{ki} = \sum_{j=1}^m r_{jk} x_i(t_j) \quad (3)$$

式中: Ψ 为主成分矩阵; Ψ_{ki} 为第 i 组航空电磁数据的第 k 个主成分; $i=1, 2, \dots, n$; $j, k=1, 2, \dots, m$ 。

1.2 主成分去噪原理

根据矩阵计算方法,电磁数据 $\mathbf{X}$ 可由主成分(Ψ)与特征向量矩阵($\mathbf{R}$)进行重构,即

$$\mathbf{X} = \mathbf{R} \Psi \text{ 或 } x_i(t_j) = \sum_{k=1}^m r_{jk} \Psi_{ki} \quad (4)$$

其中: $i=1, 2, \dots, n$; $j, k=1, 2, \dots, m$ 。

由于电磁数据的能量主要集中在低阶主成分上,而高阶主成分则代表着不相关噪声。因此,通过少量低阶(L 个)主成分重构电磁数据(式(5)),既能保证包含地下异常信息的电磁信号精度,又能达到去除不相关噪声的目的。

$$\tilde{x}_i(t_j) = \sum_{k=1}^L r_{jk} \Psi_{ki} \quad (5)$$

式中: $\tilde{x}_i(t)$ 为采用少量低阶主成分进行重构后的一组

航空电磁数据, 重构数据 $\tilde{x}_i(t)$ 与无噪数据 $x_0(t)$ 的均方根相对误差 ε 为

$$\varepsilon = \sqrt{\frac{\sum_{j=1}^m \left(\frac{x_0(t_j) - \tilde{x}(t_j)}{x(t_j)} \right)^2}{m}} \quad (6)$$

根据主成分的去噪原理, 设计了时间域航空电磁剖面数据的噪声去除步骤如下。

- 1) 计算测线上 n 个测点的 m 道时间域电磁数据 $\mathbf{X}$ ($m \times n$) 的自相关矩阵 $\mathbf{\Gamma}$;
- 2) 利用奇异值分解法计算自相关矩阵 $\mathbf{\Gamma}$ 的特征值矩阵 $\mathbf{\Lambda}$ 和特征向量矩阵 $\mathbf{R}$;
- 3) 分析航空电磁数据的主成分特征, 确定重构电磁数据的主成分数目;
- 4) 由式(5)重构时间域航空电磁数据, 评价去噪效果。

2 仿真数据去噪实例

2.1 仿真数据的主成分特征分析

为研究基于主成分的航空电磁数据去噪方法, 评价去噪效果, 首先以吊舱式直升机电磁探测系统的仿真数据为例, 研究航空电磁数据的主成分特征。

吊舱式直升机时间域航空电磁探测系统, 采用中心回线测量方式, 发射线圈半径为 7.5 m, 发射电流为 25 Hz 的正负方波, 飞行高度为 30 m, 半空间模型的大地电导率为 0.01 S/m, 归一化发射电流, 经正演计算^[22], 得到均匀半空间大地模型的 17 道 off-time 电磁响应, 如图 1 中红色点线所示, 各道中心时间等对数分布于发射电流关断后的 0.2 至 5.3 ms。为仿真均匀半空间模型的一条测线, 将该半空间模型的电磁响应延拓为 200 个测点, 并加入 5% 的高斯白噪声(信噪比 33.8 dB), 仿真野外实测原始数据形成的剖面数据。图 1 中的蓝色点线给出了其中 50 个测点的 off-time 电磁响应, 可以看到噪声严重影响晚期道数据质量。

利用式(2)与式(3), 分别计算了半空间模型的仿真剖面上 200 个测点的正演数据与含噪数据的 17 个主成分(PC1, PC2, ..., PC17), 其中第 1, 第 2 和第 3 主成分分别如图 2(a), (b)和(c)所示。由于是半空间模型, 正演数据的第 1, 第 2, 第 3 主成分(红色点线)沿测线保持一致; 由于噪声的加入, 含噪数据的各个主成分(黑色点线)则沿测线以正演数据主成分为中心, 呈现杂乱变化。可以看到, 正演数据的第 1 主成分值为 4.44 nT/s, 而第 2, 第 3 主成分值却接近 0, 因此, 含噪数

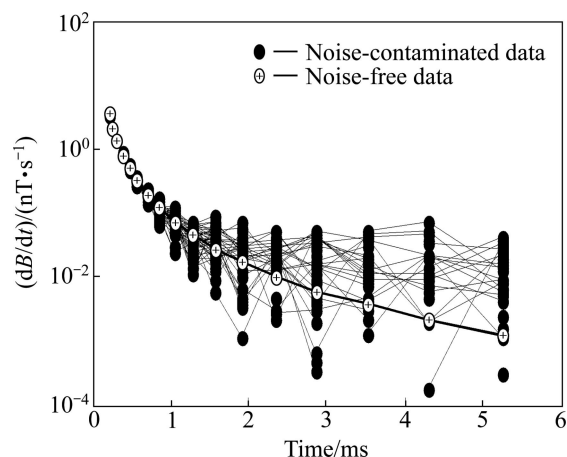


图 1 吊舱式直升机航空电磁探测的仿真数据

Fig. 1 Simulation data for airborne time domain electromagnetic exploration

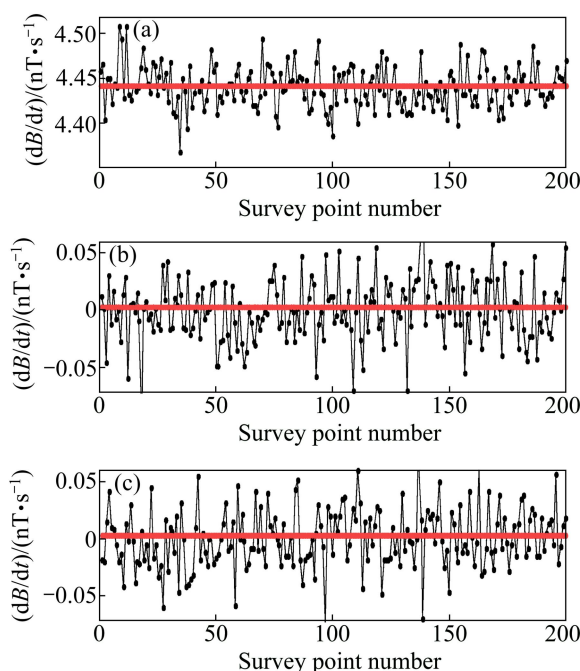


图 2 航空电磁探测仿真数据的第 1、第 2 和第 3 个主成分的剖面曲线

Fig. 2 Profile of PCs for simulation data: (a) Profile of PC1; (b) Profile of PC2; (c) Profile of PC3

据的第 2 和第 3 主成分几乎都是噪声, 数据能量主要集中在第 1 主成分上。为此, 对于含噪数据, 选取第 1 主成分进行电磁数据的重构, 就可以去除包含在第 2、第 3 和第 4 到 17 等主成分中的不相关噪声。

2.2 主成分去噪结果

利用式(5), 采用第 1 主成分重构航空电磁数据, 图 3 中的空心圆点线给出了 50 个测点的电磁数据重构结果。可以看出, 噪声明显减小, 信噪比由 33.8 dB 提

高到 46.3 dB, 重构电磁数据与正演数据的均方根相对误差为 5.24%。若对含噪数据的第 1 主成分进行滤波后, 再进行电磁数据重构, 均方根相对误差达到 1.24%。

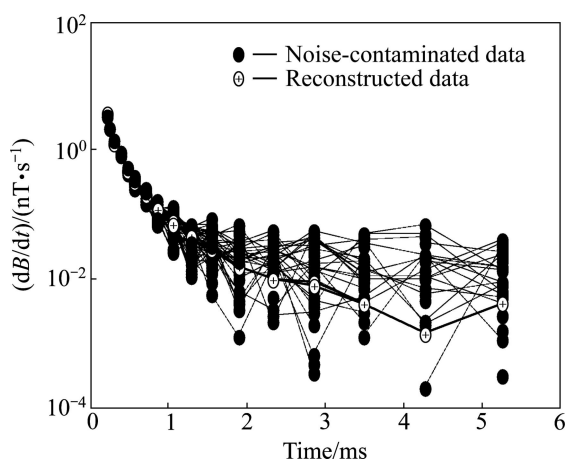


图3 选取1个主成分重构后的电磁数据

Fig. 3 Reconstructed data from PC1

3 实测数据的主成分去噪实例

3.1 时间域航空电磁探测的实测数据

在国家高技术研究发展计划重大项目支持下, 国土资源航空物探遥感中心与吉林大学仪器科学与电气工程学院合作, 研制开发了我国第一套吊舱式直升机时间域航空电磁探测系统 CHTEM(技术指标同上仿真系统)。2012 年 1 月, 该系统在河南省某地首次试飞成功, 并进行了面积性矿场勘查, 发现该地区大地呈现高阻特性, 电阻率为 3 000~8 000 $\Omega\cdot\text{m}$ 。

在此, 以河南野外飞行勘查的某测线为例, 选取长约 8.5 km 测线的 17 道剖面数据(定义测点号为 1~1712), 如图 4 所示, 蓝色-青色-绿色-黄色-红色分别代表第 1 道至第 17 道数据。可以看到, 经过数据滤波、叠加、抽道等预处理后, 信噪比明显提高, 17 道剖面曲线光滑, 大异常明显。该测线在测点号 1 300 附近有明显异常, 在测点号 300 到 500 之间出现正负交替、幅度较大的数据, 表现为高压线的特征, 与飞行测量时飞机上仪器操作员的记录相符。而其余数据段, 由于当地的高电阻率特性, off-time 数据衰减很快, 各数据道的剖面曲线相近。但是, 受残余噪声的影响, 反映数据噪声水平的晚期道数据(第 16 和第 17 道), 沿测线波动较大, 峰峰值约为 ± 25 nT/s, 影响晚期道数据质量, 同时极大地降低了航空电磁探测对小异常的分辨能力。

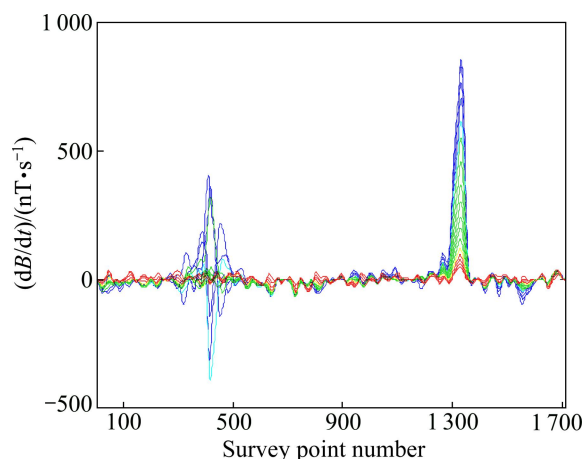


图4 河南野外航空电磁探测的17道剖面数据

Fig. 4 17 channels profile of field data from airborne time domain electromagnetic survey in Henan Province, China

3.2 实测数据的主成分分析

根据式(2)和式(3), 计算了该测线各测点的航空电磁数据的 17 个主成分, 并给出了 3 个主成分沿测线变化的剖面曲线的立体图, 如图 5 所示。为明确测线各段数据对应的主成分, 其中测点号 300~500 之间数据(高压线)的主成分标记为黑色, 测点号 1 200~1 380 数

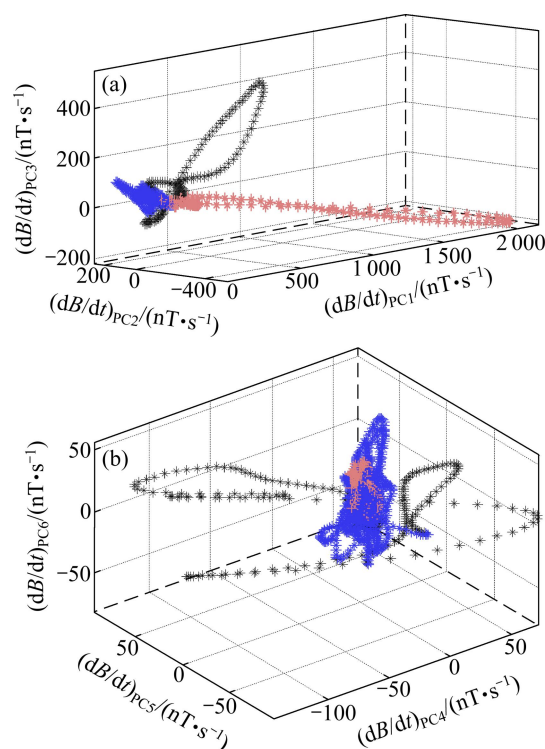


图5 实测数据的主成分剖面曲线

Fig. 5 PCs profile of airborne time domain electromagnetic data collected in Henan Province: (a) Profile of PC1, PC2 and PC3; (b) Profile of PC4, PC5 and PC6

据(异常)的主成分标记为红色,其他数据段的主成分标记为蓝色。图 5(a)所示为实测数据的第 1,第 2,第 3 主成分的剖面曲线,可以看到,红色代表的异常数据的第 1、第 2 和第 3 主成分最大幅值分别为 2 000 nT/s, 400 nT/s 和 200 nT/s,表明主成分代表的数据能量递减;黑色代表的电力线噪声的第 1 主成分较小,第 2 主成分几乎为零;蓝色代表的其他数据段,由于电磁数据非常小,其主成分在图 5(a)的原点附近变化,表明第 1、第 2、第 3 主成分接近于零。图 5(b)所示为实测数据的第 4、第 5、第 6 主成分,明显看到,黑色电力线噪声占主导,红色异常数据以及蓝色无异常数据的第 4、第 5、第 6 主成分基本在图 5(b)的原点附近变化,并且红色大异常的主成分值已经明显小于无异常数据的主成分,表明第 5 和第 6 主成分几乎都是噪声。

3.3 实测数据的主成分去噪结果

通过对实测数据的主成分特性分析,选取前 3 个主成分重构电磁数据,结果如图 6(a)所示。对比图

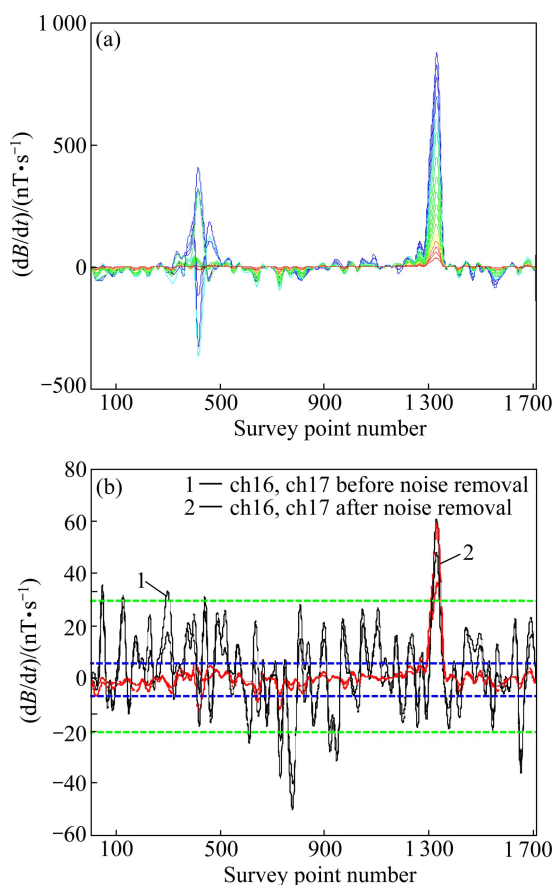


图 6 实测数据的主成分去噪结果

Fig. 6 Results after noise removal for survey data in Henan Province: (a) 17 channels survey profile after noise removal; (b) Comparison of channel 16 and channel 17 profile before and after noise removal

6(a)与图 4, 可以看到, 主成分分析去噪方法有效地保证了异常数据(测点号 1 300)的幅值, 减小了晚期道数据的噪声水平(波动)。图 6(b)给出了第 16, 17 道电磁数据在去噪前后的对比, 可以看出, 第 16, 17 道数据波动的峰峰值由原来的 ± 25 nT/s (绿色虚线)减小到 ± 8 nT/s (蓝色虚线)。同时, 增强了小异常的分辨能力, 如测点号 1 000 到 1 100 段测线上小的异常。可见, 主成分分析法能有效地去除航空电磁数据的残余噪声。

由于本方法主要针对航空电磁数据的残余噪声, 未对高压线噪声做处理, 可将高压线数据的主成分(黑色)通过设置阈值剪切掉, 再进行电磁数据的重构。常规的航空电磁数据处理直接将高压线影响的数据剪切掉, 不做处理与解释。

4 结论

1) 提出基于主成分分析的电磁数据残余噪声去噪方法, 利用信号的相关性, 采用低阶主成分重构电磁数据, 解决了由于电磁信号与噪声之间存在的频域重叠, 采用数字滤波、小波变换等频域滤波技术难以压制残余噪声的问题。

2) 通过分析航空电磁数据的主成分特征, 获取了数据与噪声的主成分分布, 从而确定重构电磁数据的主成分数目, 去除不相关噪声。采用 1 个主成分对仿真数据进行重构, 信噪比提高了 13 dB; 采用 3 个主成分对实测数据进行重构, 代表噪声水平的晚期道数据波动的峰峰值由 ± 25 nT/s 减小到 ± 8 nT/s, 验证了主成分分析法对去除航空电磁数据残余噪声的有效性, 为地球物理数据噪声去除提供了新思路。

REFERENCES

- [1] LANE R, GREEN A, HU Y. Streamed data—A source of insight and improvement for time domain airborne EM [J]. *Exploration Geophysics*, 1998, 29(2): 16–23.
- [2] LANE R, GREEN A, GOLDING C, OWERS M, PLK P, PLUNKETT C, SATTEL D, THORN B. An example of 3D conductivity mapping using the TEMPEST airborne electromagnetic system [J]. *Exploration Geophysics*, 2000, 31(2): 162–172.
- [3] BOUCHEDDA A, CHOUTEAU M, KEATING P, SMITH R. Sferics noise reduction in time-domain electromagnetic systems: application to MegaTEM^{II} signal enhancement [J]. *Exploration Geophysics*, 2010, 41: 225–239.
- [4] MACNAE J C, LAMONTAGNE Y, WEST G F. Noise

- processing techniques for time-domain EM system [J]. *Geophysics*, 1984, 49(7): 934–948.
- [5] RIDSDILL-SMITH T A, DENTITH M C. The wavelet transform in aeromagnetic processing [J]. *Geophysics*, 1999, 64(4): 1003–1013.
- [6] BUSELLI G, HWANG H S. AEM noise reduction with remote referencing [J]. *Exploration Geophysics*, 1998, 29(2): 71–76.
- [7] 李楠. 时间域航空电磁数据预处理技术研究[D]. 长春: 吉林大学, 2009.
- LI Nan. Research on airborne time domain electromagnetic data preprocessing [D]. Changchun: Jilin University, 2009.
- [8] 吕东伟. 吊舱式时间域直升机航空电磁数据处理方法研究[D]. 成都: 成都理工大学, 2011.
- LÜ Dong-wei. Methods study of helicopter-borne towed bird time domain electromagnetic data processing [D]. Chengdu: Chengdu University of Technology, 2011.
- [9] JONES I F, LEVY S. Signal-to-noise ratio enhancement in multichannel seismic data via the Karhunen-Loève transform [J]. *Geophysical Prospecting*, 1987, 35(1): 12–32.
- [10] JACKSON G M, MASON I M, GREENHALGH S A. Principal component transforms of triaxial recordings by singular value decomposition [J]. *Geophysics*, 1991, 56(4): 528–533.
- [11] GREEN A. The use of multivariate statistical techniques for the analysis and display of AEM data [J]. *Exploration Geophysics*, 1998, 29(2): 77–82.
- [12] MINTY B, HOVGGAARD J. Reducing noise in gamma-ray spectrometry using spectral component analysis [J]. *Exploration Geophysics*, 2002, 33(4): 172–176.
- [13] ASTEN M. Use of principal component images for classification of the EM response of unexploded ordnance [C]// The 20th International Geophysical Conference and Exhibition. Adelaide, Australia, 2009.
- [14] KASS M A, LI Y. Use of principal component analysis in the de-noising and signal separation of transient electromagnetic data [C]// Proceedings of The 3rd International Conference on Environmental and Engineering Geophysics (ICEEG). Wuhan, China, 2007.
- [15] KASS M A, LI Y, KRAHENBUHL R, NABIGHIAN M, OLDENBURG D. Enhancement of TEM data and noise characterization by principal component analysis [R]. Douglas Oldenburg: Department of Geophysics Colorado School of Mines, 2010.
- [16] 常威威, 郭雷, 刘坤, 付朝阳. 基于 Contourlet 变换和主成分分析的高光谱数据噪声消除方法[J]. *电子与信息学报*, 2009, 31(12): 2892–2896.
- CHANG Wei-wei, GUO Lei, LIU Kun, FU Zhao-yang. Denoising of hyperspectral data based on contourlet transform and principal component analysis [J]. *Journal of Electronics & Information Technology*, 2009, 31(12): 2892–2896.
- [17] 吴锡, 周激流, 何建新. 采用非局部主成分分析的极大似然估计图像去噪[J]. *光子学报*, 2011, 40(12): 1827–1832.
- WU Xi, ZHOU Ji-liu, HE Jian-xin. Maximum likelihood estimation image denoising using non-local principle component analysis [J]. *Acta Photonica Sinica*, 2011, 40(12): 1827–1832.
- [18] 姚莉丽, 冯象初, 李亚峰. 去除乘性噪声的主成分分析算法[J]. *光子学报*, 2011, 40(7): 1031–1035.
- YAO Li-li, FENG Xiang-chu, LI Ya-feng. Principle component analysis method for multiplicative noise removal [J]. *Acta Photonica Sinica*, 2011, 40(7): 1031–1035.
- [19] 傅东浩, 王平, 邢建春, 张迅. 基于 PCA 的无线传感器网络压缩感知算法[J]. *计算机测量与控制*, 2012, 20(9): 2593–2596.
- FU Dong-hao, WANG Ping, XING Jian-chun, ZHANG Xun. PCA-based compressed sensing algorithm of wireless sensor network [J]. *Computer Measurement & Control*, 2012, 20(9): 2593–2596.
- [20] ZHU Kai-guang, MA Ming-yao, CHE Hong-wei, JI Yan-ju, YU Sheng-bao, LIN Jun. PC-based artificial neural network inversion for airborne time-domain electromagnetic data [J]. *Applied Geophysics*, 2012, 9(1): 1–8.
- [21] 刘霁, 李云. 基于 PCA 法的湘西喀斯特石漠化地区生态安全评价[J]. *中南大学学报: 自然科学版*, 2012, 43(12): 4895–4901.
- LIU Ji, LI Yun. Assessment of ecology security of Xiangxi karst rocky desertification area based on principal component analysis method [J]. *Journal of Central South University: Science and Technology*, 2012, 43(12): 4895–4901.
- [22] 朱凯光, 林君, 韩悦慧, 李楠. 基于神经网络的时间域直升机电磁数据电导率深度成像[J]. *地球物理学报*, 2010, 53(3): 743–750.
- ZHU Kai-guang, LIN Jun, HAN Yue-hui, LI Nan. Research on conductivity depth imaging of time domain helicopter-borne electromagnetic data based on neural network [J]. *Chinese Journal of Geophysics*, 2010, 53(3): 743–750.

(编辑 龙怀中)