

基于自适应差分进化算法的 MT 信号激电信息提取

董莉^{①②} 江沸菠*^{①③} 李帝铨^①

(^①中南大学地球科学与信息物理学院,湖南长沙 410083; ^②湖南涉外经济学院信息科学与工程学院,湖南长沙 410205;

^③湖南师范大学物理与信息科学学院,湖南长沙 410081)

董莉,江沸菠,李帝铨. 基于自适应差分进化算法的 MT 信号激电信息提取. 石油地球物理勘探, 2016, 51(3): 613-624.

摘要 从 MT 信号中提取激电信息的研究有助于提高大深度探测技术的勘探精度、深度和范围。针对 MT 信号激电信息提取中存在的非线性和非凸特征,通过改进混沌自适应差分进化算法中进化参数的自适应策略,提出了一种基于非均匀统计分布的自适应差分进化两阶段最小构造反演方法。一方面该方法利用柯西分布和高斯分布的统计特性自适应获取进化参数 F 和 CR,提高算法的全局搜索能力,通过记忆以往迭代过程中的优秀进化参数提高算法后期的稳定性;另一方面该方法通过引入第二阶段的反演过程强化极化率对观测数据的影响;通过将正则化参数引入差分进化算法的适应度函数解决反演的多解性问题。对含激电效应的 MT 一维模型的反演结果表明,本文算法能够较好地重构地电结构和提取激电信息并在加噪环境下具有较强的鲁棒性。与其他非线性算法(混沌自适应差分进化算法,标准差分进化算法和粒子群优化算法)的反演结果对比表明,本文算法具有更为优越的全局搜索能力和较高的反演精度,适于微弱激电信息的提取。

关键词 自适应差分进化算法 大地电磁测深 非线性反演 激电信息提取

中图分类号:P631 文献标识码:A doi: 10.13810/j.cnki.issn.1000-7210.2016.03.025

1 引言

激发极化(Induced Polarization, IP)是发生在地球介质中因外电流激发而引起介质内部出现电荷分离,产生一种附加的“过电位”(Over voltage)的物理化学现象^[1]。然而随着近年来对深部矿产资源和油气资源需求的不断增加,传统的 IP 法存在着探测深度的瓶颈,越来越多的研究人员开展了从探测深度较大的大地电磁法(Magnetotelluric, MT)资料中提取激发极化参数的研究工作^[2]。

吴汉荣等^[3]早在 20 世纪 70 年代就开展了天然场源激电法测量的可行性研究;Marali 等^[4]认为利用低频天然电场作为频率域激电研究的场源是可行的;Gasperikova 等^[5]对天然场源的 IP 效应检测进

行了深入研究,并得到了肯定的结论;罗延钟等^[6]通过理论分析和数值模拟详细论证了当时已有的天然场源激电法方案的利弊,陈清礼等^[7]基于 Dias 模型,对三层地电模型中间层有无极化情况的正演进行了模拟对比,给出了从 MT 的复视电阻率资料中反演地层真谱参数的算法,但未给出模拟仿真结果;曹中林等^[8]使用 Cole-Cole 模型模拟激发极化效应,采用广义逆方法对 MT 理论模型和实际资料进行反演,获得了较好的效果;岳安平等^[9]在 MT 正演研究的基础上,基于 Dias 模型对含激电参数的 CSAMT 信号展开了正演研究,为提取 CSAMT 信号中所含激电信息提供理论基础;Ghorbani 等^[10]利用 Cole-Cole 模型研究了含电磁效应的频谱域极化数据的正演和反演方法,其中反演采用的是高斯-牛顿法,同时还给出了相应正反演算法的开源软件;

* 湖南省长沙市中南大学地球科学与信息物理学院,410083。Email:jiangfeibo@126.com。

本文于 2014 年 12 月 29 日收到,最终修改稿于 2016 年 3 月 30 日收到。

本研究受国家自然科学基金资助项目(41204054)、中国博士后科学基金资助项目(20110490149)、湖南科技大学页岩气资源利用省重点实验室开发基金资助项目(E21423)、湖南省研究生科研创新项目资助(2015zzts064)、湖南省教育厅科研优秀青年资助项目(15B138)、湖南省科技计划资助项目(2015JC3067)、中国博士后科学基金资助项目(2015M580700)联合资助。

He等^[11]利用三维MT技术在柴达木盆地的三湖地区进行油气藏的勘探,并指出高电阻率和高极化率是寻找油气藏的两个关键指标;罗卫峰等^[12]进行了大地电磁激电效应的油气检测试验;Yu等^[13]采用SVD反演从TEM信号中提取IP信息;冯兵等^[14]分析了Cole-Cole模型参数对CSAMT信号的 E_x 、 H_y 的影响特征,采用最小二乘反演方法,分别利用 H_y 和 E_x 分量反演电阻率参数和极化率参数;Tang等^[15]比较了Dias模型、Cole-Cole模型和Multi-Cole-Cole模型对MT信号的影响,分析了不同IP模型在含激电信息的MT信号中的适用情况。

总的来说,目前对于电磁法信号中激电信息的提取研究多以激电信息提取方法的可行性探讨及验证为主。一方面通过对比有无激电效应的正演结果,总结由极化模型参数变化所引起的极化异常和模型参数之间对应关系的相关规律;另一方面通过线性或拟线性的反演方法,定量分析MT信号中的激电信息。但线性反演方法往往依赖于接近真实模型的初始值才能成功收敛,容易陷入局部极值^[16]。完全非线性反演方法为处理和解释电磁资料提供了新的研究方向,受到地球物理研究者的广泛关注。差分进化(Differential Evolution,DE)算法是一种高效的完全非线性反演方法。该方法采用实数编码,具有控制参数简单,全局搜索能力和鲁棒性强、收敛速度快并且结果精准等特点,非常适合求解不连续、不可微,或者没有函数解析式的全局优化问题,已在地球物理领域的资料解释中获得了广泛的应用^[17,18]。

标准的差分进化算法在解释含激电效应的MT资料时存在着以下不足:①当考虑激电效应时,地下可极化体的电阻率是一个与频率有关的复数,在进行正演迭代时需计算每个频点的地面波阻抗,其运算量有所增加,使得差分进化算法在计算适应度函数时运算效率偏低;②含激电效应的MT反演的模型参数在电阻率、厚度的基础上增加了激电参数,反演参数的增加使得差分进化在反演过程中更加难以收敛;③大地电磁场由于地下导电性不均匀所引起的异常强度远大于激电效应引起的异常,非线性反演算法往往容易收敛于真实的电阻率附近,但在进行极化率搜索时极易陷入局部极值,加大了激电信息提取的难度。

本文在Dias复电阻率模型的MT信号正演基

础上,采用自适应差分进化算法对激电信息以反演的方式进行提取,其具体的研究内容包括:①通过分析混沌自适应差分进化(Chaotic differential evolution, CDE)算法在全局搜索能力和稳定性方面的不足,利用具有非均匀统计特性的柯西随机算子和高斯随机算子自适应获取进化参数 F 和CR,提高算法的全局搜索能力;②通过记忆以往迭代过程中的优秀进化参数来提高自适应差分进化算法后期的稳定性;③通过引入第二阶段的反演过程来加强后期极化率参数对观测数据的影响,提高激电信息提取的准确性;④通过将正则化因子引入DE算法的适应度函数来解决反演的多解性问题,得到满足最小构造条件的反演结果。对含激电效应的MT模型的反演结果表明,本文算法能够较好地重构地电结构和提取激电信息,在加噪环境下具有较强的鲁棒性。与其他非线性算法(CDE,DE和粒子群优化)的反演结果比较表明,本文算法具有更强的全局搜索能力和更高的反演精度,能够对MT信号中的激电信息进行有效提取。

2 方法原理

2.1 基于Dias模型的MT正演理论

以一维水平层状大地模型为研究对象,设有 n 层地电断面,各层厚度分别为 $h_1, h_2, \dots, h_n (h_n \rightarrow \infty)$,各层电阻率分别为 $\rho_1, \rho_2, \dots, \rho_n$,各层极化率分别为 m_1, m_2, \dots, m_n 。在大地电磁测深方法中,地面波阻抗是根据各层厚度及电阻率从下往上递推得到的,具体计算公式如下^[19]

$$Z_i = Z_{0i} \frac{1 + \frac{Z_{i+1} - Z_{0i}}{Z_{i+1} + Z_{0i}} e^{-2k_i h_i}}{1 - \frac{Z_{i+1} - Z_{0i}}{Z_{i+1} + Z_{0i}} e^{-2k_i h_i}} \quad (1)$$

其中

$$Z_{0i} = -\sqrt{i\omega\mu_0\rho_i}$$

$$k_i = -\sqrt{i\omega\mu_0\sigma_i}$$

$$\rho_i = \frac{1}{\omega\mu_0} |Z|^2$$

式中: Z_i 为第 i 层地层顶面的复波阻抗; Z_{0i} 为第 i 层的本征阻抗; k_i 为第 i 层的复波数; ρ_i 为复视电阻率; Z 为地面波阻抗; σ_i 第 i 层的电导率; ω 为角频率; μ_0 为大地导磁率,通常取 $\mu_0 = 4\pi \times 10^{-7} \text{ H/m}$ 。

当考虑激电效应时,相应层的实电阻率用 Dias 模型的复电阻率取代,从而获得激电效应和电磁效应的综合响应。Dias 模型中的复电阻率 $\rho(\omega)$ 的表达式为^[20]

$$\rho(\omega) = \rho \left[1 - m \left[1 - \frac{1}{1 + i\omega\tau' \left(1 + \frac{1}{\mu} \right)} \right] \right] \quad (2)$$

式中: $\rho(\omega)$ 为考虑极化效应后与频率相关的复电阻率; ρ 为未考虑极化效应时的电阻率; m 为极化率; $\mu = i\omega\tau + (i\omega\tau'')^{\frac{1}{2}}$, $\tau' = \frac{\tau(1-\delta)}{\delta(1-m)}$, $\tau'' = \tau^2\eta^2$, τ 为时间常数, η 为电化学参数, δ 为极化电阻率系数。

2.2 差分进化算法

差分进化算法由美国学者 Storn 等^[21]提出,因其形式简洁、控制参数少、收敛速度快等特点成为最优化问题研究方向的一个热点,在许多领域都得到了应用。考虑到反演算法的效率和求解精度,本文采用 DE/rand/1/bin(即标准差分进化算法)策略实现 MT 信号激电信息的提取。算法的基本工作过程如下。

(1) 初始种群: 设种群规模为 NP, 可行解空间的维数为 D , 用 $\mathbf{X}(t)$ 来表示进化到第 t 代时的种群。在问题的可行解空间内随机产生初始种群 $\mathbf{X}(0) = \{\mathbf{x}_1^0, \mathbf{x}_2^0, \dots, \mathbf{x}_{NP}^0\}$, 其中 $\mathbf{x}_i^0 = [x_{i,1}^0, x_{i,2}^0, \dots, x_{i,D}^0]$ 表示第 i 个个体解。个体的各个分量可按下式产生

$$x_{i,j}^0 = x_{j,\min} + \text{rand}(x_{j,\max} - x_{j,\min}) \quad (3)$$

式中: $x_{j,\max}$ 和 $x_{j,\min}$ 分别为解空间第 j 维的上下界; rand 为随机算子, 表示一个 $0 \sim 1$ 之间的独立随机数。

(2) 变异: 对于第 G 代种群中任意一个目标向量 \mathbf{x}_i^G , 根据下列公式生成变异向量

$$\mathbf{v}_i^{G+1} = \mathbf{x}_{r_1}^G + F \cdot (\mathbf{x}_{r_2}^G - \mathbf{x}_{r_3}^G) \quad i = 1, 2, \dots, NP \quad (4)$$

式中: $\{\mathbf{x}_{r_1}^G, \mathbf{x}_{r_2}^G, \mathbf{x}_{r_3}^G\}$ 是在第 G 代种群中随机选择的三个不同个体, 且 $r_1 \neq r_2 \neq r_3 \neq i$; F 被称为缩放因子, 是介于 $[0, 2]$ 间的实常量, 用于控制差分向量 $(\mathbf{x}_{r_2}^G - \mathbf{x}_{r_3}^G)$ 的影响。

(3) 交叉: 交叉操作的目的在于提高种群个体的多样性。新的交叉向量 \mathbf{u}_i^{G+1} 通过以下公式生成

$$\mathbf{u}_{i,j}^{G+1} = \begin{cases} \mathbf{v}_{i,j}^{G+1} & \text{randb} \leq \text{CR 或 } j = \text{rand}_j \\ \mathbf{x}_{i,j}^G & \text{randb} > \text{CR 或 } j \neq \text{rand}_j \end{cases} \quad i = 1, \dots, NP, j = 1, \dots, D \quad (5)$$

式中: randb 是 $[0, 1]$ 间的随机数; CR 被称为交叉因子, 是介于 $[0, 1]$ 间的常数, 取值越大发生交叉的概率越高。rand_j 是在 $[1, D]$ 间随机选择的整数, 它确保交叉向量 \mathbf{u}_i^{G+1} 至少从变异向量 \mathbf{v}_i^{G+1} 中获得一个元素。

(4) 选择: 选择过程通过下式描述

$$\mathbf{x}_i^{G+1} = \begin{cases} \mathbf{u}_i^{G+1} & f(\mathbf{u}_i^{G+1}) < f(\mathbf{x}_i^G) \\ \mathbf{x}_i^G & \text{其他} \end{cases} \quad (6)$$

式中 $f(\mathbf{u}_i^{G+1})$ 和 $f(\mathbf{x}_i^G)$ 分别表示 \mathbf{u}_i^{G+1} 和 \mathbf{x}_i^G 的适应度函数。选择的结果是使得子代个体总是优于或等于父代个体, 从而使种群始终向着最优解的方向进化。

3 基于自适应差分进化的 MT 信号激电信息提取

3.1 影响差分进化反演性能的因素

差分进化算法的反演性能由控制参数缩放因子 F 和交叉因子 CR 决定。缩放因子 F 决定了差分向量对个体的影响, 若其值较小, 则群体的差异度过早下降, 将会引起算法的早熟; 若其值较大, 虽然可以增加算法跳出局部最小的能力, 但是由于群体的差异度不易下降, 算法的收敛速度会明显降低。交叉因子 CR 的取值越大意味着发生交叉的可能性越大, 个体的非线性交互增加, 从而对复杂问题的演化能力增加。在传统的差分进化算法中, 以上参数在整个进化过程中一般是固定不变的, 其值往往根据经验设定。但是针对不同的优化问题, 其参数的设置各不相同, 往往需要根据求解问题的实际情况和多次实验才能确定合适的参数取值。为了解决这个问题, 有学者提出基于混沌自适应参数改进策略的差分进化算法(CDE)^[22,23]。

3.2 基于非均匀统计分布的自适应差分进化算法

由于混沌均匀的概率密度分布和良好的互相关特性, 使得基于混沌的自适应差分进化算法能够在 F 和 CR 的解空间中充分均匀的取值, 提高了差分进化算法跳出局部极小能力, 但是也使得差分进化算法在整个搜索过程中性能过于均衡, 不利于算法前期扩大搜索范围以及后期稳定收敛, 从而影响 MT 信号激电信息提取的准确性和效率。为解决该问题, Islam 等^[24]提出了一种采用非均匀统计分布的自适应进化参数获取方法, 并应用于多种群差分

进化算法中,极大地提高了差分进化算法的收敛速度和计算效率。本文在 Islam 研究的基础上,采用柯西分布和高斯分布的随机算子来自适应获取 F 和 CR 参数,与混沌自适应差分进化算法相比,该方法具有更强的前期全局搜索能力和更稳定的后期收敛性能。自适应参数获取的基本原理如下。

第 G 代 F 参数通过下式自适应产生

$$F_G = \text{Cauchy}(F_{\text{pos},G}, \omega) \quad (7)$$

式中: Cauchy 代表服从柯西分布的随机函数; ω 为柯西分布的尺度参数; $F_{\text{pos},G}$ 为柯西分布第 G 代的位置参数,其定义如下

$$F_{\text{pos},G} = \text{mean}_p(F_{\text{success}}) \quad (8)$$

式中: F_{success} 为进化过程中使得适应度下降的有效参数 F_i 的集合,这些 F 参数能够产生更好的个体, $F_{\text{pos},1}$ 的初始值为 0.5; mean_p 表示幂平均(Power mean)也叫广义平均(Generalized mean),计算公式如下

$$\text{mean}_p(F_{\text{success}}) = \left(\frac{1}{l} \sum_{i=1}^l F_i^p \right)^{\frac{1}{p}} \quad (9)$$

式中: F_i 为 F_{success} 中的历代有效参数; l 为 F_{success} 集合的势; 幂平均参数 $p=1.5$ 。

第 G 代 CR 参数通过下式自适应产生

$$CR_G = \text{Gaussian}(CR_{\text{pos},G}, \omega) \quad (10)$$

式中: Gaussian 代表服从高斯分布的随机函数; ω 为高斯分布的尺度参数; $CR_{\text{pos},G}$ 为高斯分布第 G 代的位置参数,其定义如下

$$CR_{\text{pos},G} = \text{mean}_p(CR_{\text{success}}) \quad (11)$$

式中: CR_{success} 为进化过程中使得适应度下降的有效参数 CR_j 的集合, $CR_{\text{pos},1}$ 的初始值为 0.6, CR_{success} 的广义平均计算公式如下

$$\text{mean}_p(CR_{\text{success}}) = \left(\frac{1}{l} \sum_{j=1}^l CR_j^p \right)^{\frac{1}{p}} \quad (12)$$

式中: CR_j 为 CR_{success} 中的历代有效参数; l 为 CR_{success} 集合的势; $p=1.5$ 。

随机算子的尺度参数 ω 采用混沌振荡形态^[25],第 G 代 ω_G 定义如下

$$\omega_G = \omega_e + (\omega_s - \omega_e)(0.99^G x_G) \quad (13)$$

$$x_{G+1} = \mu \cdot x_G(1 - x_G) \quad (14)$$

式(14)为混沌 Logistic 方程。式中: $\mu=4$ 是控制参数; $\omega_s=0.6$ 和 $\omega_e=0.0001$ 分别为惯性权重的初值和终值。

对于以上自适应过程中的一些关键问题进一步分析和解释如下。

(1) F_{success} 和 CR_{success} 的本质在于记忆以往迭代过程中优秀的 F_G 和 CR_G 参数,从而对下一代进化参数的产生进行指导。

(2) F_G 采用柯西分布的原因是相较于均匀分布的混沌序列而言,柯西分布的曲线形态以位置参数为中心分布更宽,将 $F_{\text{pos},G}$ 作为柯西分布的位置参数,能够获得更加多样化的 F_G 参数,尤其是当全局极值与当前的搜索位置相距较远时,柯西分布能更加有效地避免陷入局部极值。 CR_G 采用高斯分布的原因是 CR 的有效取值范围比 F 窄,因此高斯分布的短尾曲线形态相较柯西分布更加适合 CR 的随机产生。两种非均匀的统计分布能够根据以往迭代过程中的经验(F_{success} 和 CR_{success}),产生比混沌序列更有针对性的进化参数。

(3) 相较于算术平均,幂平均能够获得更大的 $F_{\text{pos},G}$ 和 $CR_{\text{pos},G}$ 参数,从而增大进化参数的影响,增加算法跳出局部极小,快速收敛的能力。

(4) 尺度参数 ω 控制了随机参数和历史有效参数之间的比重,初始迭代时,尺度参数较大,随机参数保证了算法初期的全局搜索能力,有利于避免陷入局部极值,后期尺度参数应该逐渐变小,稳定在历史有效参数的幂平均值上,从而保证算法后期的收敛能力。

(5) 自适应过程中, $F_{\text{pos},1}$, $CR_{\text{pos},1}$ 和 p 值的设定参考了文献[24], ω_s 和 ω_e 的设定通过凑试法确定。

将式(7)和式(10)分别带入式(4)和式(5),就得到了本文提出的基于非均匀统计分布的自适应差分进化算法。

3.3 二阶段最小构造反演方法

在含激电效应的层状地电模型的应用背景下, MT 信号 DE 算法反演的模型参数通常包括各层的厚度 h_i 、电阻率值 ρ_i 以及极化率 m_i ($h_1\rho_1m_1, h_2\rho_2m_2, \dots, h_i\rho_im_i, \dots, h_n\rho_nm_n, n$ 为地电模型层数, $i \in [1, n], h_n \rightarrow \infty$)。

大地电磁场由于地下导电性不均匀所引起的异常强度远大于激电效应引起的异常,在反演过程中,差分进化算法中个体电阻率参数对适应度函数的影响远大于极化率参数对适应度函数的影响。因此差分进化算法提取激电信息的反演过程将明显的分为

两个阶段:在第一阶段,由于电阻率参数对视电阻曲线的主要影响,个体在解空间内将快速收敛至正确的电阻率参数附近;第二阶段,算法开始对电阻率进行微调,虽然此时电阻率参数对适应度函数的影响趋于稳定,但由于极化率参数在数值上远小于电阻率参数,同时对适应度函数的影响也远小于电阻率参数,算法在此时极易陷入局部极值,得到错误的极化率参数,从而加大了激电信息提取的难度。

针对以上激电信息提取过程中的反演特征,很多学者都认为应该首先对电阻率数据进行反演,然后以反演得到的电阻率数据为初始模型进行极化率参数的提取^[11, 12, 14]。

本文在上述思想的启发下,提出一种应用于激电信息提取的二阶段最小构造反演方法,该方法构造的适应度函数为

$$\text{Fit} = E(\mathbf{e}) + \lambda_1 R(\boldsymbol{\rho}) + \lambda_2 R(\mathbf{m}) \quad (15)$$

式中: $R(\boldsymbol{\rho})$ 和 $R(\mathbf{m})$ 分别为对电阻率和极化率的最小构造约束函数; λ_1, λ_2 分别为 $R(\boldsymbol{\rho})$ 和 $R(\mathbf{m})$ 对应的正则化因子(采用两个独立正则化因子的原因是极化率的取值空间($m_i \in [0, 1]$)较电阻率的取值空间有较大差异(一般可认为 $\rho_i \gg m_i$),如果采用统一的正则化因子将无法约束相对较小的极化率参数); $R(\boldsymbol{\rho})$ 和 $R(\mathbf{m})$ 在此均采用下式进行计算^[26]

$$R(\mathbf{M}) = \sum_{i=2}^{n-1} (M_{i+1}^{\text{inv}} - 2M_i^{\text{inv}} + M_{i-1}^{\text{inv}})^2 \quad (16)$$

式中 $\mathbf{M}_i^{\text{inv}}$ 为反演得到的模型参数,包括各层的电阻率 ρ_i^{inv} 以及极化率 m_i^{inv} 。

$E(\mathbf{e})$ 为目标误差函数,在此采用下式进行计算

$$E(\mathbf{e}) = \|\mathbf{e}\|_2 = \left[\sum_i |\rho_{ai}^{\text{obs}} - \rho_{ai}^{\text{pre}}|^2 \right]^{\frac{1}{2}} \quad (17)$$

式中: ρ_{ai}^{obs} 表示第 i 个观测数据; ρ_{ai}^{pre} 表示第 i 个反演预测数据,即

$$\rho_a^{\text{pre}} = \begin{cases} F(\boldsymbol{\rho}^{\text{inv}}, \mathbf{m}^{\text{inv}}, \mathbf{h}^{\text{inv}}) & R^2 < \text{th} \\ F(\mathbf{m}^{\text{inv}}) & R^2 > \text{th} \end{cases} \quad (18)$$

式中: $F(\cdot)$ 为含激电效应的 MT 正演算子; $\boldsymbol{\rho}^{\text{inv}}$ 和 \mathbf{m}^{inv} 分别为反演得到的模型电阻率参数和极化率参数; \mathbf{h}^{inv} 为反演得到的层厚度,其初始化为对数等间距; R^2 为决定系数,是评价反演参数拟合程度的指标,其具体定义参见 4.1 节; th 为设定阈值。由式(18)可知,激电信息提取的反演过程分为两个阶段:在第一阶段中,模型参数 ρ 和 m 同时参与反演,

进行地电模型参数的搜索,此时由于 ρ 对观测数据的影响远大于 m ,差分进化算法中的个体将快速的收敛于正确的 ρ 附近, R^2 的值迅速增加,当 R^2 的值达到设定的阈值 th 时,电阻率参数的搜索结束,反演算法进入第二阶段;在第二阶段中,反演算法固定 ρ 和 h 的值不变,仅针对 m 进行参数搜索。在此阶段中,一方面强化了极化率 m 对观测数据的影响,能够有效克服算法后期的局部收敛问题,当达到设定的搜索代数时,算法将收敛于正确的 m 参数附近;另一方面由于搜索参数的减少,提高了反演算法在第二阶段的执行速度。

3.4 反演步骤及流程

本文提出了一种基于非均匀统计分布的自适应差分进化算法和用于进行激电信息提取的二阶段最小构造反演方法。上述方法进行 MT 信号激电信息提取的步骤如下:

(1)初始化迭代次数、种群规模、自适应算法参数($F_{\text{pos},1}, \text{CR}_{\text{pos},1}, \rho, \omega_s$ 和 ω_e 等)和反演参数(λ_1, λ_2 和 th 等);

(2)根据地电模型参数 m, ρ 和 h 产生初始种群,开始第一阶段反演;

(3)根据种群中个体的电阻率参数和极化率参数构造最小构造约束函数,同时调用含激电效应的 MT 正演算法计算个体的误差函数。由最小构造约束函数和误差函数共同评估种群中个体适应度;

(4)根据适应度更新全局最优解,同时计算全局最优解的决定系数;

(5)根据决定系数与阈值的关系判断是否进入第二阶段反演;如果为否,则执行自适应差分进化算法的自适应参数获取、变异、交叉和选择操作,产生新一代种群,并转至步骤(3);如果为是,则保持 ρ 和 h 参数不变,并转至步骤(6)并进行第二阶段反演;

(6)在当前最优解中 m 参数基础上进一步则执行自适应差分进化算法的自适应参数获取、变异、交叉和选择操作,产生新一代种群;

(7)将第二阶段反演的 m 参数结合第一阶段反演的 ρ 和 h 参数计算种群中的个体适应度,更新全局最优解;

(8)判断是否满足算法的终止条件,如果为否,则转至步骤(6);如果为是,保存当前的全局最优解。

(9)根据全局最优解输出并评估反演结果。

反演算法的流程图如图 1 所示。

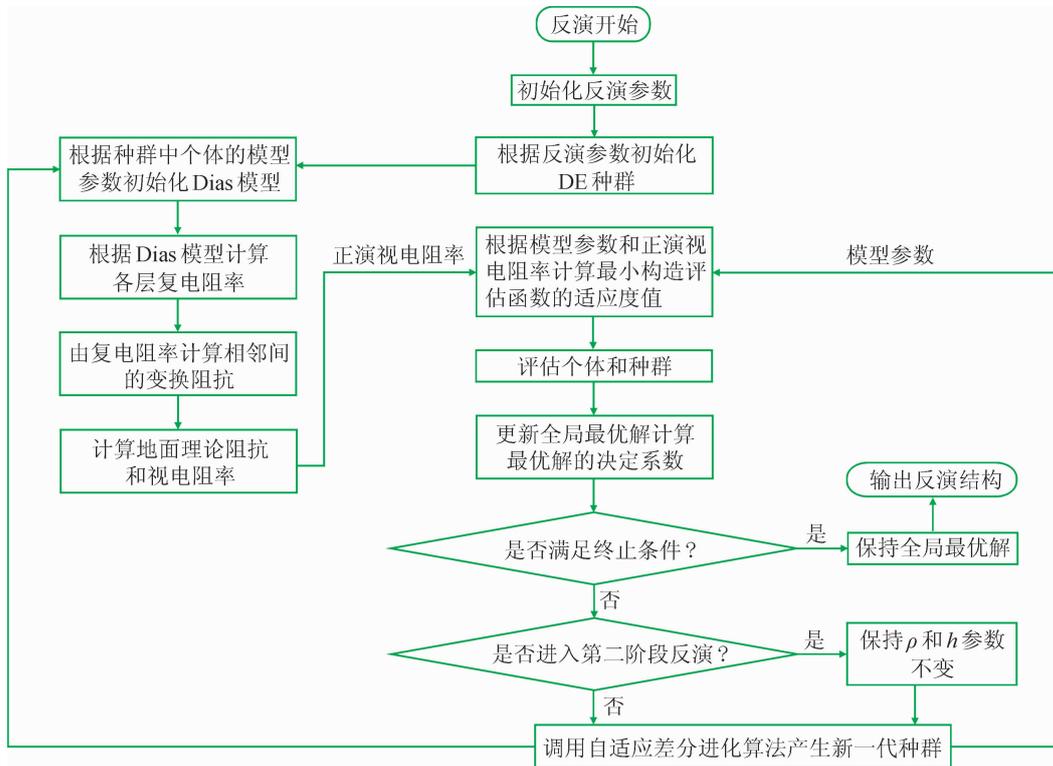


图1 激电信息提取反演算法流程图

4 含激电效应的 MT 反演

4.1 评价指标

本文采用均方误差 (Mean square error, MSE) 和决定系数 (Determination coefficient, R^2) 衡量反演算法的性能, 其相关定义如下

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (Y_i - y_i)^2 \quad (19)$$

$$R^2 = \frac{(N \sum_{i=1}^N Y_i y_i - \sum_{i=1}^N Y_i \sum_{i=1}^N y_i)^2}{[N \sum_{i=1}^N y_i^2 - (\sum_{i=1}^N y_i)^2][N \sum_{i=1}^N Y_i^2 - (\sum_{i=1}^N Y_i)^2]} \quad (20)$$

式中: y_i 为反演结果代入含激电效应的 MT 正演计算后的第 i 个数据; Y_i 为对应的第 i 个测量数据; N 为测量数据的数量; MSE 代表预测误差, 其值越

小, 表示反演模型的预测误差越小; R^2 代表预测值与测量值之间的相关度, 其值越大, 表示两者间存在着越明显的线性相关性。

4.2 三层 (K 型) 地电模型激电信息提取反演

以三层 K 型地电模型为例, 研究极化层位于不同位置时, 本文算法的反演性能。其中地电模型的参数设置为 $\rho_1 = 30 \Omega \cdot \text{m}$, $\rho_2 = 200 \Omega \cdot \text{m}$, $\rho_3 = 10 \Omega \cdot \text{m}$, $h_1 = 500 \text{m}$, $h_2 = 2000 \text{m}$; 极化率参数 $m = 0.2$; 自适应差分进化算法的基本参数设置为种群规模 $NP = 20$, 迭代次数 $t = 1000$, 进化参数 F 和 CR 采用本文所提出的自适应方式获取; 设置正则化参数 $\lambda_1 = 0.001$, $\lambda_2 = 300$, 阈值参数 $\text{th} = 0.98$ 。表 1 给出了当极化层分别位于首层、中间层和底层的三种情况下, 本文算法在不同阶段的反演结果。

由表 1 数据可知, 当极化层位于不同位置时, 本文算法中第二阶段反演均能够获得比第一阶段反演更低的 MSE 和更高的 R^2 。由于在第二阶段中电阻率

表 1 三层地电模型激电信息提取反演结果

模型	第一阶段/(MSE)	第一阶段/ R^2	第二阶段/(MSE)	第二阶段/ R^2	计算时间/s
首层极化	0.2642	0.98	0.2269	0.9984	23.4008
中间层极化	0.2597	0.98	0.2134	0.9997	22.4176
底层极化	0.2683	0.98	0.2226	0.9996	22.4254

参数 ρ 和层厚参数 h 保持不变,反演性能的提升取决于极化率参数 m ,因此证明了二阶段最小构造反演的有效性。由表 1 还可看出,当极化层位于中间层时,反演结果的拟合性能最佳,同时当极化层位于 K 型模型的其他位置时,反演也能够获得较高的决

定系数 R^2 和较低的拟合误差 MSE,说明本文所提出的自适应差分进化算法能够有效地避免局部极值,提高算法的全局搜索能力。图 2 给出了中间层极化的情况下模型反演结果和视电阻率曲线的拟合结果。

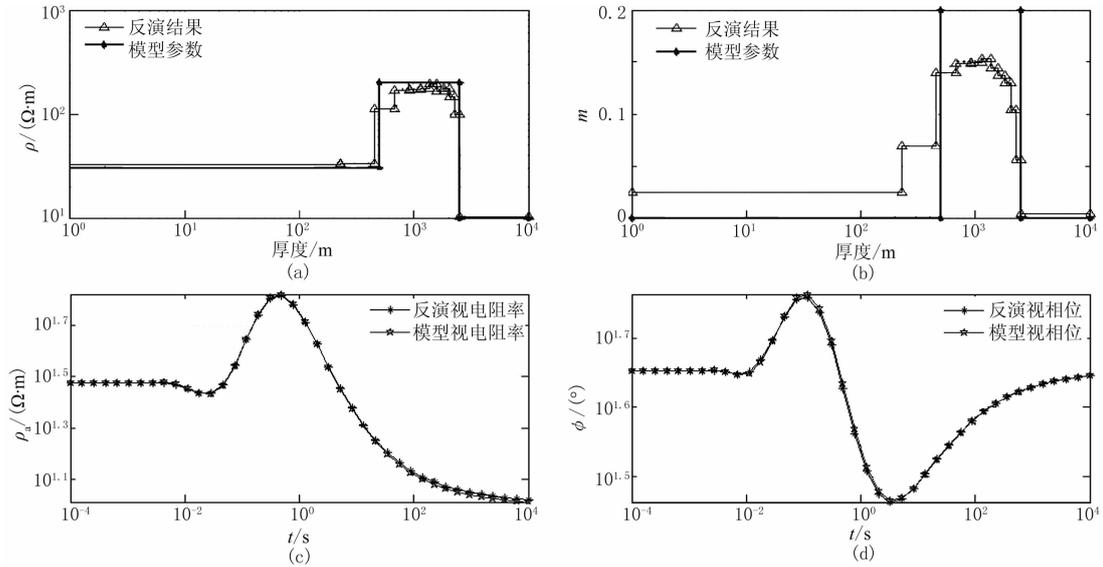


图 2 三层地电模型中间层极化时反演结果

(a)电阻率参数反演结果; (b)极化率参数反演结果; (c)视电阻率拟合曲线; (d)视相位拟合曲线

由图 2 可知,对于三层 K 型地电模型,本文算法能够准确地实现对电阻率参数和极化率参数的反演提取,视电阻率曲线和视相位曲线与观测数据均具有较高的拟合度。相对而言,电阻率反演效果优于极化率反演效果,这是因为大地电磁场地下导电性不均匀所引起的异常强度远大于激电效应引起的异常,因此在反演过程中,电阻率对个体进化的影响大于极化率参数。但总的来说,通过第二阶段的反演方法,极化率的变化趋势也能够较为准确地得到反映。

4.3 四层(KH型)地电模型反演

以四层 KH 型地电模型为例,研究本文算法与 CDE(logistic)^[22]、CDE(tent)^[23]、标准 DE 和 PSO 算法的反演性能比较。其中地电模型的参数设置为 $\rho_1 = 30\Omega \cdot m, \rho_2 = 200\Omega \cdot m, \rho_3 = 10\Omega \cdot m, \rho_4 = 100\Omega \cdot m, h_1 = 1000m, h_2 = 2000m, h_3 = 3000m$; 设置极化层为第 2~3 层,极化率参数 $m = 0.2$; 反演参数 $\lambda_1 = 0.0001, \lambda_2 = 200, th = 0.98$ 。标准 PSO 算法的加速常数为 c_1 和 c_2 ,为便于比较,不同非线性反演算法的其他参数尽量保持一致,具体设置如表 2 所示。

表 2 不同非线性反演方法的参数设置

反演算法	种群规模	迭代次数	演化参数
本文算法	30	1000	$F_G = \text{Cauchy}(F_{\text{pos},G}, \tau)$ $CR_G = \text{Gaussian}(CR_{\text{pos},G}, \tau)$
CDE(logistic)	30	1000	$F_{G+1} = 4F_G(1 - F_G)$ $CR_{G+1} = 4CR_G(1 - CR_G)$
CDE(tent)	30	1000	$F_{G+1} = \begin{cases} 2F_G & 0 \leq F_G \leq 0.5 \\ 2(1 - F_G) & 0.5 < F_G \leq 1 \end{cases}$ $CR_{G+1} = \begin{cases} 2CR_G & 0 \leq CR_G \leq 0.5 \\ 2(1 - CR_G) & 0.5 < CR_G \leq 1 \end{cases}$
DE	30	1000	$F_G = 0.9 \quad CR_G = 0.6$
PSO	30	1000	$c_1 = c_2 = 1.19445 \quad \tau = 0.6$

在上述参数设置条件下,对四层 KH 型地电模型进行反演,表 3 给出了不同非线性反演算法的反演性能。其模型反演结果和视电阻率曲线的拟合情况如图 3 所示。

由图 3 可知,不同的非线性反演方法均能够较好地反演模型中的电阻率参数,但提取极化率信息,各种算法的性能则各不相同。标准的 DE 算法和 PSO 算法分别得到了一个逐渐递减和逐渐递增的

表 3 不同非线性反演方法的反演性能

模型	MSE	R^2	最小适应度	平均适应度
本文算法	0.2631	0.9997	8.0145	28.3093
CDE(tent)	0.2608	0.9991	12.5968	35.3666
CDE(logistic)	0.2621	0.9996	12.2359	34.5480
DE	0.2971	0.9588	29.2047	59.5405
PSO	0.2902	0.9484	34.9362	68.8532

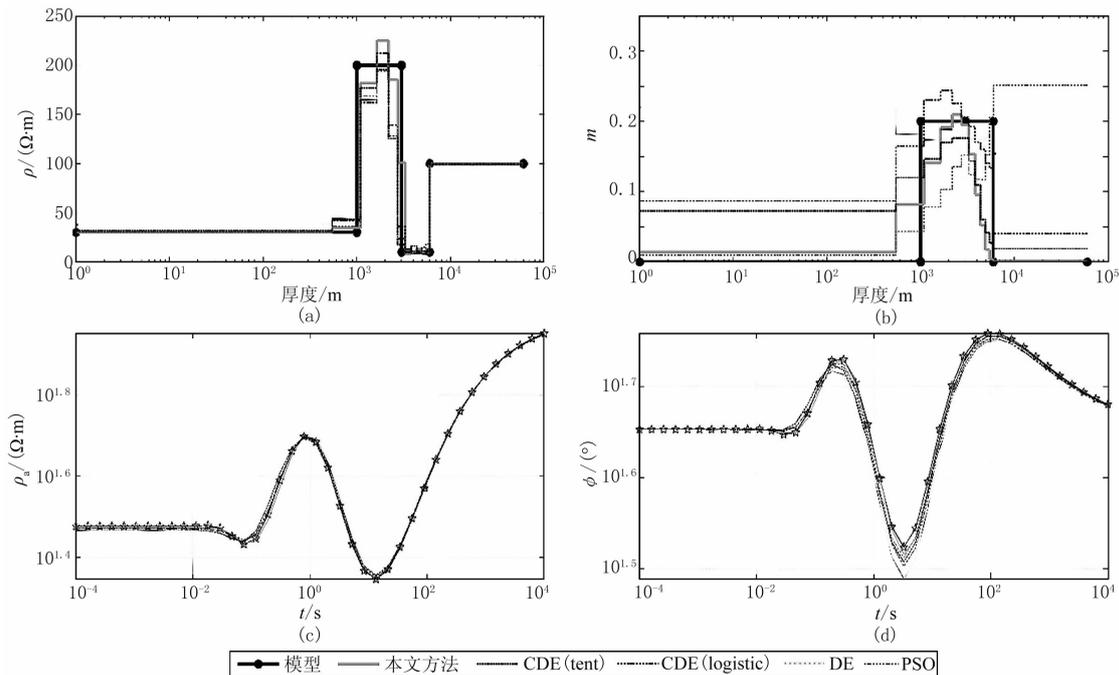


图 3 四层地电模型非线性反演方法比较

(a)电阻率参数反演结果; (b)极化率参数反演结果; (c)视电阻率拟合曲线; (d)视相位拟合曲线

极化率曲线,这与模型中设定的极化参数不符合。从表 3 可知,这是因为标准的 DE 算法和 PSO 算法在反演时获得的 R^2 值分别为 0.9588 和 0.9484,没有达到第二阶段反演的阈值 0.98,所以没有进入第二阶段反演,陷入了局部极值,无法较好地完成极化率提取任务。混沌 DE 算法 CDE(logistic) 和 CDE(tent) 的反演性能优于 DE 算法和 PSO 算法,能够反演出模型中极化率参数的基本形态,但其第一层极化率参数偏高,大于 0.05;本文算法能够准确地实现对电阻率参数和极化率参数的反演提取,尤其是极化率参数的提取比其他非线性反演方法更为准确。从视电阻率和视相位的拟合曲线来看,五种非线性方法均能够较好地拟合观测数据。这一方面显示了非线性反演方法在求解反问题时不依赖初始值和搜索能力强的优点,另一方面也显示了极化率参

数在反演中对适应度函数的影响小于电阻率参数对适应度函数的影响,因此虽然 DE 和 PSO 算法能够较好地拟合适应度曲线,但是无法准确反演极化率参数。为了能够准确提取极化率参数,既要求反演算法具有更强的全局搜索能力,也要求在反演中强化极化率对视电阻率曲线的影响。CDE 算法和本文提出的自适应 DE 算法因为采用自适应参数的进化方案,因此在全局搜索能力上较标准的 DE 算法和 PSO 算法更有优势;而第二阶段反演方案在反演算法搜索的后期,能够保持电阻率参数 ρ 和层厚参数 h 不变,仅针对极化率参数 m 对视电阻率曲线的影响进行搜索,进一步提高了反演算法跳出局部极值的能力,提取了更加准确的极化率参数。为进一步说明本文算法的反演性能,图 4 给出了不同非线性反演算法的适应度收敛曲线。

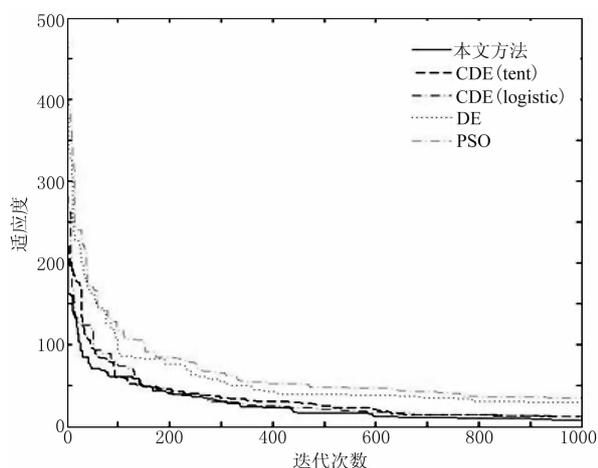


图 4 不同反演算法的适应度收敛曲线

由图 4 可知,所有非线性反演方法在迭代过程中均能够逐渐搜索到更好的模型参数,促使适应度不断下降并收敛。结合表 3 和图 4 可知,与标准的 DE 和 PSO 算法相比,本文算法和混沌算法获得了更小的适应度,说明本文算法和混沌算法的自适应参数获取机制具有更强的全局搜索能力,因此能够在算法搜索的后期提取出有效的极化率信息。CDE (logistic)和 CDE(tent)算法在原理上相近,仅在实现上采用了不同的混沌方程,因此其适应度非常接近,CDE(logistic)算法的适应度略优于 CDE(tent)算法。本文算法采用非均匀统计分布的自适应参数获取方法,比CDE具有更优的全局搜索能力,获得

了最小的适应度函数和最优的激电信息提取结果。同时从平均适应度指标和适应度收敛曲线上可以看出,本文算法的收敛速度更快,效率更高。

4.4 五层(KHA)地电模型

野外采集数据均含有一定噪声,本节以五层 KHA 型地电模型为例,研究本文算法在加入不同噪声情况下的反演稳定性和鲁棒性。其中地电模型的参数设置为 $\rho_1 = 50\Omega \cdot m, \rho_2 = 300\Omega \cdot m, \rho_3 = 50\Omega \cdot m, \rho_4 = 300\Omega \cdot m, \rho_5 = 1000\Omega \cdot m, h_1 = 100m, h_2 = 200m, h_3 = 300m, h_4 = 500m$ 。设置极化层为第 2~3 层,极化率参数 $m = 0.2$ 。自适应差分进化算法的基本参数设置为种群规模 $NP = 30$,迭代次数 $t = 1000$,进化参数 F 和 CR 采用本文所提出的自适应方式获取;反演参数设置为 $\lambda_1 = 0.0003, \lambda_2 = 900$,考虑到噪声对决定系数的影响,设定阈值 $th = 0.96$ 。表 4 给出了加入不同程度噪声时本文算法的反演性能。

图 5~图 7 进一步给出了三种情况下模型反演结果和视电阻率曲线的拟合情况。

表 4 五层加噪地电模型激电信息提取反演结果

模型	MSE	R ²	计算时间/s
0%噪声	0.3590	0.9912	44.3217
10%噪声	0.4328	0.9852	45.2716
20%噪声	0.6898	0.9633	45.7613

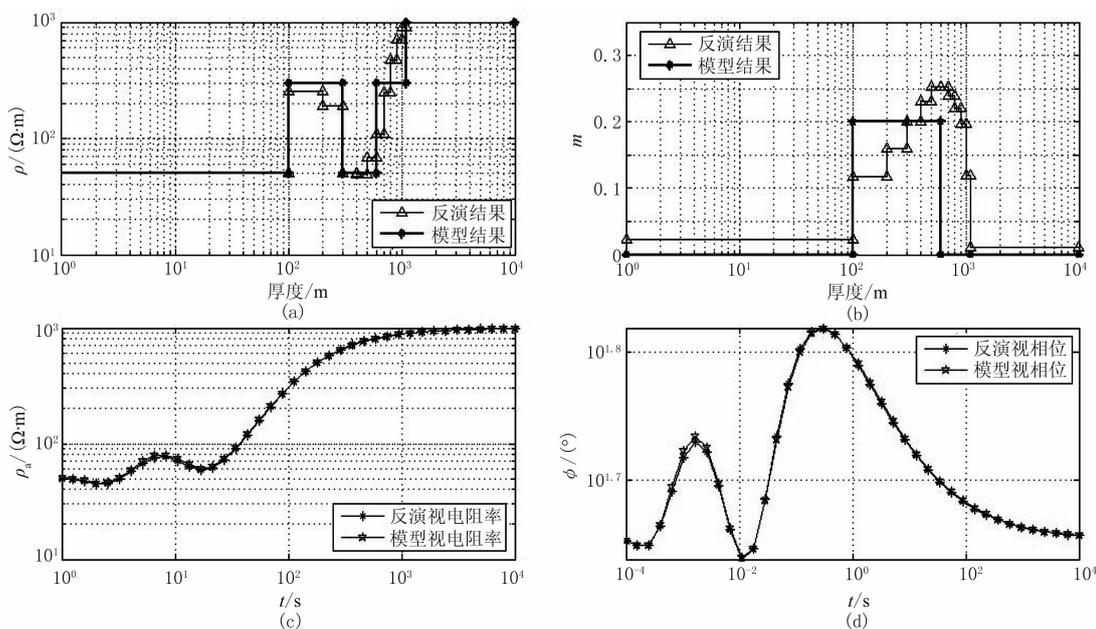


图 5 五层地电模型不加噪声时反演结果

(a)电阻率参数反演结果;(b)极化率参数反演结果;(c)视电阻率拟合曲线;(d)视相位拟合曲线

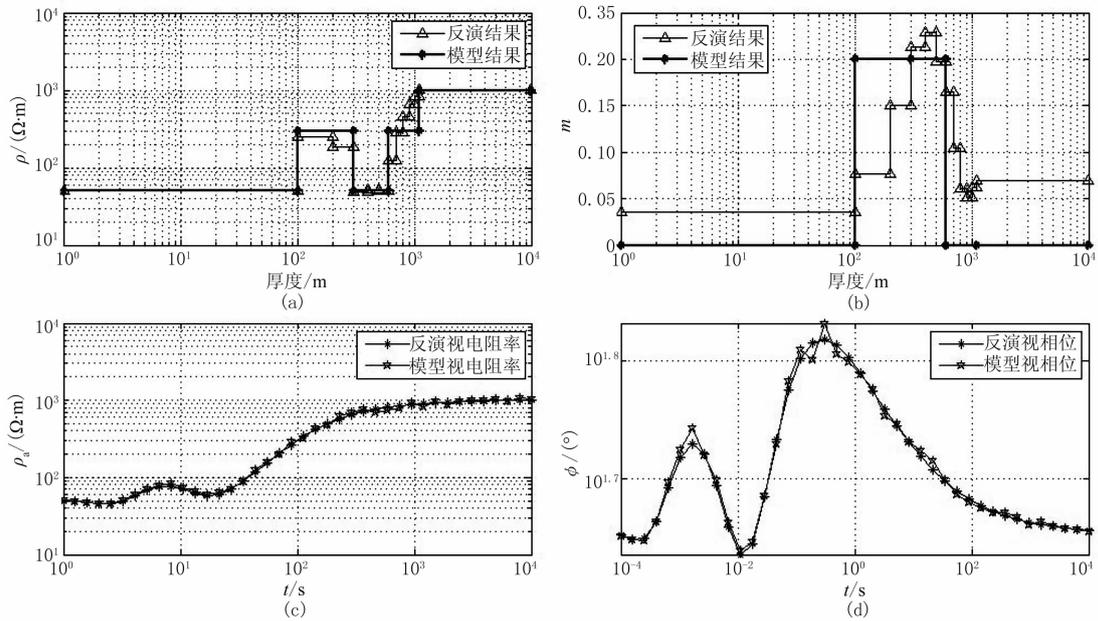


图6 五层地电模型加10%噪声时反演结果

(a)电阻率参数反演结果;(b)极化率参数反演结果;(c)视电阻率拟合曲线;(d)视相位拟合曲线

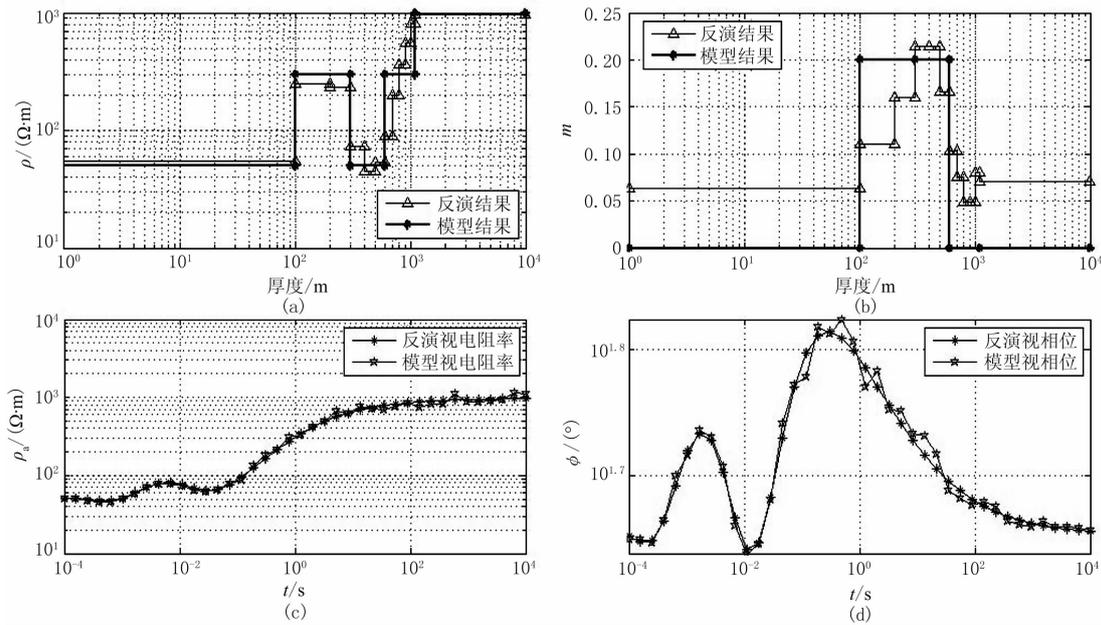


图7 五层地电模型加20%噪声时反演结果

(a)电阻率参数反演结果;(b)极化率参数反演结果;(c)视电阻率拟合曲线;(d)视相位拟合曲线

由图5可知,在无噪声情况下,本文反演算法较为准确地重构了五层地电模型的电阻率参数和极化率参数。由图6和图7可知,反演算法在10%和20%随机噪声的干扰下也能够较好地反演模型的电阻率参数,但对于极化率参数的反演结果吻合程度略有下降,其首层和底层的极化率偏大,但基本能够反演出极化率的变化形态。从数据拟合结果看,反

演模型的视电阻率和视相位均位于加噪后的曲线之间,体现了反演算法的稳定性。由表4可知,随着噪声的增加,反演误差MSE逐渐增大,决定系数 R^2 逐渐下降,但总体来说本文算法保证了较低的反演误差和较高的决定系数,体现了算法的鲁棒性。最后,五层模型反演的计算时间比三层模型反演的计算时间有所增加,这是由于:①正演的计算时间增

加;②地电模型结构复杂,增加了算法的种群规模。

5 结论

本文采用完全非线性方法反演提取 MT 信号的激电参数,实验结果表明:①本文算法能够较为准确地提取层状模型下的极化率参数并重构真实的地电结构。其中中间层极化的提取效果最佳;②相较于同类型的 CDE,DE 和 PSO 等非线性反演算法,本文算法获得了更为优越的适应度以及更为稳定可靠的反演结果;③在 10% 和 20% 的随机噪声干扰下,本文算法的反演结果仍然能较为真实准确地显示实际地电构造和所含激电信息,体现了本文算法较强的抗干扰能力。

由于实际的地电结构非常复杂,实际工程环境中存在的噪声组成也十分复杂。因此上述方法用于实测数据反演时,还存在以下问题:①算法中的正则化因子对不同地电类型以及极化层所在不同位置的反演效果有较大影响,只有合理地设置该系数才能减少虚假构造,取得较好的反演结果。下一步希望通过自适应方法来确定该参数以提高算法的适用性;②阈值参数 th 对激电信息提取的准确性及算法的计算效率有直接影响,本文采用凑试法确定 th 参数。定量分析 th 参数对二阶段反演结果的影响将是一项有意义的研究工作;③本文仅对一维层状地电模型进行了理论上的正反演研究,验证了基于自适应差分进化算法的二阶段最小构造反演在 MT 信号中提取激电信息的可行性,将该方法应用于实际资料的处理是下一步工作的重点。

参 考 文 献

- [1] 何继善. 双频激电法. 北京:高等教育出版社,2006.
- [2] 岳安平,底青云,王妙月等. 油气藏 MT 激电效应一维正演研究. 石油地球物理勘探,2009,44(3):364-370.
Yue Anping, Di Qingyun, Wang Miaoyue et al. 1D forward simulation of MT induced polarization (IP) effect of reservoir. OGP,2009,44(3):364-370.
- [3] 吴汉荣,王式铭. 利用天然电磁场进行激发极化法测量的可能性. 物探与化探,1978,1(1):62-64.
Wu Hanrong, Wang Shiming. The feasibility of nature source induced polarization. Geophysical and Geochemical Exploration,1978,1(1):62-64.
- [4] Murali S, Rao IBR, Bhimasankaram VLS. Comparison of anomalous effects determined using telluric fields and time domain IP technique (test results). Exploration Geophysics, 1980,11(1/2):45-46.
- [5] Gasperikova E, Morrison H. Mapping of induced polarization using natural fields. Geophysics, 2001,66(1):137-147.
- [6] 罗延钟,张胜业,熊彬. 天然场源激电法的可行性. 地球物理学报,2003,46(1):125-130.
Luo Yanzhong, Zhang Shengye, Xiong Bin. Feasibility of natural source induced polarization. Chinese Journal of Geophysics,2003,46(1):125-130.
- [7] 陈清礼,胡文宝,李金铭. 由 MT 资料反演真谱参数的基本原理. 石油天然气学报,2007,28(6):61-64.
Chen Qingli, Hu Wenbao, Li Jinming. Method for inverting real spectral parameters based on magnetotelluric method(MT). Journal of Oil and Gas Technology,2007,28(6):61-64.
- [8] 曹中林,何展翔,昌彦君. MT 激电效应的模拟研究及在油气检测中的应用. 地球物理学进展,2007,21(4):1252-1257.
Cao Zhonglin, He Zhanxiang, Chang Yanjun. A simulation study of induced polarization effect of magnetotelluric and its application in oil and gas detection. Progress in Geophysics,2007,21(4):1252-1257.
- [9] 岳安平,底青云,王妙月等. 含激电效应的 CSAMT 一维正演研究. 地球物理学报,2009,52(7):1937-1946.
Yue Anping, Di Qingyun, Wang Miaoyue et al. 1-D forward modeling of the CSAMT signal incorporating IP effect. Chinese Journal of Geophysics, 2009,52(7):1937-1946.
- [10] Ghorbani A, Camerlynck C, Florsch N. CR1Dinv: A matlab program to invert 1D spectral induced polarization data for the Cole-Cole model including electromagnetic effects. Computers & Geosciences, 2009,35(2):255-266.
- [11] He Zhanxiang, Hu Zuzhi, Luo Weifeng et al. Mapping reservoirs based on resistivity and induced polarization derived from continuous 3D magnetotelluric profiling: Case study from Qaidam basin, China. Geophysics, 2010,75(1):B25-B33.
- [12] 罗卫锋,何展翔,王财富等. 大地电磁激电效应油气检测试验. 石油地球物理勘探,2011,46(6):978-983.
Luo Weifeng, He Zhanxiang, Wang Caifu et al. A test study of induced polarization effects of magnetotelluric in oil and gas detection. OGP,2011,46(6):978-983.
- [13] Yu Chuantao, Liu Hongfu, Zhang Xinjun et al. The analysis on IP signals in TEM response based on SVD. Applied Geophysics, 2013,10(1):79-87.
- [14] 冯兵,王璐,王玉等. 利用 CSAMT 电磁场响应提取激电效应的方法初探. 地球物理学进展,2013,28(4):2116-2122.
Feng Bing, Wang Junlu, Wang Yu et al. The preliminary exploration to extract the IP effect which using CSAMT electromagnetic field response method. Progress in Geophysics,2013,28(4):2116-2122.
- [15] Tang Rui, Yu Peng, Xiang Yang et al. The sensitivity analysis of different induced polarization models

- used in magnetotelluric method. *Acta Geodaetica et Geophysica*, 2014, 49(2): 225-233.
- [16] Shaw R, Srivastava S. Particle swarm optimization: A new tool to invert geophysical data. *Geophysics*, 2007, 72(2): F75-F83.
- [17] 潘克家, 王文娟, 谭永基等. 基于混合差分进化算法的地球物理线性反演. *地球物理学报*, 2009, 52(12): 3083-3090.
Pan Kejia, Wang Wenjuan, Tan Yongji et al. Geophysical linear inversion based on hybrid differential evolution algorithm. *Chinese Journal of Geophysics*, 2009, 52(12): 3083-3090.
- [18] 宋维琪, 高艳珂, 朱海伟. 微地震资料贝叶斯理论差分进化反演方法. *地球物理学报*, 2013, 56(4): 1331-1339.
Song Weiqi, Gao Yanke, Zhu Haiwei. The differential evolution inversion method based on Bayesian theory for micro-seismic data. *Chinese Journal of Geophysics*, 2013, 56(4): 1331-1339.
- [19] 刘国栋, 石应骏. 大地电磁测深法教程. 北京: 地震出版社, 1985.
- [20] Dias C A. Developments in a model to describe low-frequency electrical polarization of rocks. *Geophysics*, 2000, 65(2): 437-451.
- [21] Storn R, Price K. Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. *Journal of Global Optimization*, 1995, 23(4): 341-359.
- [22] Yuan Xiaohui, Cao Bo, Yang Bo et al. Hydrothermal scheduling using chaotic hybrid differential evolution. *Energy Conversion and Management*, 2008, 49(12): 3627-3633.
- [23] Lu Youlin, Zhou Jianzhong, Qin Hui et al. Chaotic differential evolution methods for dynamic economic dispatch with valve-point effects. *Engineering Applications of Artificial Intelligence*, 2011, 24(2): 378-387.
- [24] Islam Sk Minhazul, Das S, Ghosh S et al. An adaptive differential evolution algorithm with novel mutation and crossover strategies for global numerical optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2012, 42(2): 482-500.
- [25] 戴前伟, 江沸波. 基于混沌振荡 PSO-BP 算法的电阻率层析成像非线性反演. *中国有色金属学报*, 2013, 23(10): 2897-2904.
Dai Qianwei, Jiang Feibo. Nonlinear inversion for electrical resistivity tomography based on chaotic oscillation PSO-BP algorithm. *The Chinese Journal of Nonferrous Metals*, 2013, 23(10): 2897-2904.
- [26] Constable S C, Parker R L, Onstale C G. Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, 1987, 52(3): 289-300.

(本文编辑:冯杏芝)

作者简介



董莉 博士研究生, 1982年生; 2004年毕业于湖南师范大学电子信息工程专业, 获学士学位; 2007年毕业于湖南师范大学电路与系统专业, 获硕士学位; 2013年至今在中南大学攻读地球探测与信息技术专业博士学位。目前主要从事电磁法极化信息提取方法研究。

本期广告索引

东方地球物理公司	封2, 插1
东方地球物理公司研究院处理中心	插2~4
环波软件公司	插5, 6
中油油气勘探软件国家工程研究中心有限公司	插7, 8
斯伦贝谢科技服务(北京)有限公司	插10, 11
中国石油勘探开发研究院西北分院	插12, 13
中国石油勘探开发研究院廊坊分院地球物理与信息研究所	插14, 15
中国石化胜利油田物探研究院	插16, 17
大庆钻探工程公司地球物理勘探公司	插18, 19
东方地球物理公司研究院地质研究中心	插20, 21, 封4
东方地球物理公司油藏地球物理研究中心	插22, 23
CNPC 物探重点实验室	插24, 封3