

文章编号:0253-2697(1999)02-00-51-56

# 基于自适应遗传算法的试井分析最优化方法

尹洪军\* 翟云芳

(大庆石油学院)

**摘要:** 如何实现试井资料的自动分析,保证解释结果的唯一性,一直是人们在积极探索的目标。为此,提出了一种基于自适应遗传算法的试井分析最优化方法。遗传算法是一种遵循“适者生存”原则的快速搜索方法。该方法在进化过程中,评价每个解的适应度值以执行交换和变异操作;特别适合于传统优化方法难以解决的非线性和不连续等问题。该方法优越于非线性回归自动拟合试井分析方法。它不需要估计井筒和油藏参数的初值,也不要求目标函数连续可微,并优于简单的遗传算法。它是通过解群体的适应度值,自适应地改变遗传算法的交换概率和变异概率,从而保证群体的多样性和算法的收敛能力,再结合非线性回归自动拟合方法,最终获得最佳的试井分析结果。通过实测资料分析,取得了良好的效果。

**主题词:** 自适应;遗传算法;自动拟合;试井分析;油藏参数

## 1 前言

在利用计算机进行典型曲线最优化自动拟合方面,人们已经做了大量的工作,研究出了多种非线性回归的油藏参数自动估计方法,如 Lenvenberg-Marquardt 方法(L-M)、最小绝对值方法(LAV)以及各种修正的方法<sup>[1~2]</sup>。但研究和应用结果表明,这些最优化方法都是通过求解使目标函数梯度为零的一组非线性方程来进行搜索最优解的。它们要求目标函数连续可微,并且受初值影响很大,是属于寻找局部最优解的方法,无法解决试井分析特别是早期试井分析中解的不唯一性难题。

遗传算法(Genetic Algorithm—GA)是近年来迅速发展起来的一种最优化方法。该算法由美国 Michigan 大学的 J. Holland 教授于 1975 年首先提出的<sup>[3]</sup>。至今已成为 90 年代在国际上受到普遍重视的算法之一<sup>[4]</sup>。由于遗传算法的整体搜索策略和在优化计算时不依赖于梯度信息,所以它的应用范围非常广泛,尤其适合于处理传统搜索方法难以解决的高度复杂的非线性问题。遗传算法以其高效实用的特点已在组合优化和过程控制等领域中得到广泛的应用,并取得了良好的效果。但在油气田勘探开发方面,至今只见到遗传算法在地球物理<sup>[5]</sup>、气举采油<sup>[6]</sup>等优化问题中的应用,而在试井分析方面的应用尚不多见。

本文针对无限大均质油藏试井分析理论,提出了一种基于自适应遗传算法(Adaptive Genetic Algorithm—AGA)的试井分析最优化方法。该方法只需要掌握井筒和油藏参数的值域,而不需要估计井筒和油藏参数的初值,也不要求目标函数连续可微。它是通过解群体的适应度值,自适应地改变遗传算法的交换概率和变异概率,从而保证群体的多样性和算法的收敛能力,再结合非线性回归自动拟合方法,最终获得最佳的试井分析结果。通过实测资料分析,取得了良好的效果。

## 2 自适应的交换概率和变异概率

陈国良等人<sup>[7]</sup>对遗传算法进行了详细的研究,指出遗传算法在参数寻优过程中,基于生物遗传学的基本原理,模拟自然界生物种群的“物竞天择,适者生存”的自然规律。通常,利用 GA 求解问题的关键是对问题的解进

\* 尹洪军,女,1964年7月生。1986年毕业于大庆石油学院,1991年获硕士学位,现正在攻读博士学位。通讯处:黑龙江省安达市大庆石油学院石油工程系。邮政编码:151400。

行编码,构造出适应度函数,并选取 GA 参数—交换概率  $p_c$ 、变异概率  $p_m$  及群体规模  $n$ 。而交换概率  $p_c$  和变异概率  $p_m$  在 GA 中的重要性早已被 GA 研究者所公认<sup>[8~9]</sup>。本文提出的自适应遗传算法(AGA),就是在搜索最优参数过程中,通过解群体的适应度值,自适应地改变遗传算法的交换概率  $p_c$  和变异概率  $p_m$ ,而不必象简单遗传算法那样预先给出交换概率  $p_c$  和变异概率  $p_m$  的确定值。下面讨论在 AGA 寻优过程中,交换概率  $p_c$  和变异概率  $p_m$  的确定方法。

GA 在搜索全局最优解时必须具备两个特点:确定搜索最优解区域并收敛到最优解的能力;在搜索全局最优解时开辟新的解空间的能力。GA 的这些特点是由交换概率  $p_c$  和变异概率  $p_m$  控制的。交换概率  $p_c$  控制着交换的速率。交换概率  $p_c$  越大,新的解个体进入群体的速度就越快,也就是开辟新的解空间的能力就越强。但当交换概率  $p_c$  增加时,高性能的解个体被破坏的可能性增大。交换概率  $p_c$  越小,GA 搜索性能就越可能陷入迟钝状态。变异概率  $p_m$  过大可能将遗传算法变成一种纯粹的随机搜索算法。变异概率  $p_m$  过小又可能引起遗传算法在达到最优解前发生不成熟收敛。

本文的思路是通过解群体的适应度值自适应地改变交换概率  $p_c$  和变异概率  $p_m$ 。当高性能的解趋近于局部最优解时,就增加交换概率  $p_c$  和变异概率  $p_m$  值;当高性能的解分散在解空间时,就减少交换概率  $p_c$  和变异概率  $p_m$  值。为达到这一目的,正确地判断 GA 是否收敛到最优解是至关重要的。本文采用了 Srinivas 等人提出的方法<sup>[9]</sup>,即考察群体的平均适应度值  $\bar{f}$  与最大适应度值  $f_{\max}$  的关系。

许多实测资料分析的结果表明,一般情况下当 GA 收敛到全局最优解时,  $f_{\max} - \bar{f}$  降低。这一特点也可从后面给出的两个实例分析中观察到。于是  $f_{\max} - \bar{f}$  可作为考察 GA 是否收敛的尺度。交换概率  $p_c$  和变异概率  $p_m$  是随  $f_{\max} - \bar{f}$  变化。当 GA 收敛到全局最优解时,交换概率  $p_c$  和变异概率  $p_m$  增加,反之,交换概率  $p_c$  和变异概率  $p_m$  减少,即交换概率  $p_c$  和变异概率  $p_m$  与  $f_{\max} - \bar{f}$  成反比。这样,可以取交换概率  $p_c$  和变异概率  $p_m$  的表达式为

$$p_c = \frac{k_1}{f_{\max} - \bar{f}} \quad (1) \quad p_m = \frac{k_2}{f_{\max} - \bar{f}} \quad (2)$$

由式(1)和式(2)可以看出,交换概率  $p_c$  和变异概率  $p_m$  不依赖于任何解个体的适应度值。对于所有的群体,  $p_c$  和  $p_m$  具有相同的值。高适应度的解个体与低适应度的解个体具有相同频率的交换和变异。正是因为如此,这种方法存在一定的问题,即当一个群体收敛到全局最优解时,交换概率  $p_c$  和变异概率  $p_m$  就增加,并有可能引起在最优解附近的高性能解个体遭到破坏,那么,群体就可能永远收敛不到全局最优解。因此,虽然这种方法可以防止 GA 的早熟现象,但也可能引起 GA 的性能变差,收敛速度降低。

为了克服上述问题,需要保存群体中性能好的解。对高适应度的解个体,取较低的交换概率  $p_c$  和变异概率  $p_m$ ,以保证 GA 的收敛性;而对低适应度的解个体,取较高的交换概率  $p_c$  和变异概率  $p_m$ ,以防止 GA 的不成熟收敛。因此,交换概率  $p_c$  应该不仅取决于  $f_{\max} - \bar{f}$  值,而且还取决于两个父代个体的适应度值,即交换概率  $p_c$  应该随  $f_{\max} - f'$  而变,  $f'$  为两个父代交换个体中较大的适应度值。同样,变异概率  $p_m$  也应该不仅取决于  $f_{\max} - \bar{f}$  值,而且还取决于解个体本身的适应度值  $f$ ,即  $f$  越接近于  $f_{\max}$ ,变异概率  $p_m$  应该越减小。于是,取交换概率  $p_c$  和变异概率  $p_m$  的表达式为

$$p_c = \frac{k_1(f_{\max} - f')}{f_{\max} - \bar{f}} \quad f' \geq \bar{f} \quad (3) \quad p_m = \frac{k_2(f_{\max} - f)}{f_{\max} - \bar{f}} \quad f \geq \bar{f} \quad (4)$$

$$p_c = k_3 \quad f' < \bar{f} \quad (5) \quad p_m = k_4 \quad f < \bar{f} \quad (6)$$

其中  $k_1 \leq 1, k_2 \leq 1, k_3 \leq 1, k_4 \leq 1$ 。一般地,取  $k_2 = 0.5, k_4 = 0.5, k_1 = 1, k_3 = 1$ 。

取  $k_1, k_2, k_3$  和  $k_4$  小于 1 是为了将  $p_c$  和  $p_m$  限制在 0~1 范围内。由式(3)和式(4)可知,当  $f$  接近  $f_{\max}$  时,  $p_c$  和  $p_m$  减少;当  $f$  等于  $f_{\max}$  时,  $p_c$  和  $p_m$  为 0,即不进行交换和变异操作。当  $p_c = k_1$  时,  $f' = f$ ; 当  $p_m = k_2$  时,  $f = \bar{f}$ ;

当  $f < \bar{f}$  时, 式(5)和式(6)可防止  $p_c$  和  $p_m$  大于 1。因此, 这种方法才是使交换概率  $p_c$  和变异概率  $p_m$  自适应变化的理想方法。

### 3 基于自适应遗传算法的最优化试井分析方法

#### 3.1 数学模型及其解

假设地层均质、水平、等厚且各向同性; 流体流动符合达西定律。不考虑重力和毛细管力的影响; 油藏中流体微可压缩。在假设条件下, 建立了无限大均质油藏中的数学模型

$$\frac{\partial^2 p_D}{\partial r_{De}^2} + \frac{1}{r_{De}} \frac{\partial p_D}{\partial r_{De}} = \frac{\partial p_D}{\partial t_{De}} \quad (7) \quad p_D(r_{De}, 0) = 0 \quad (8)$$

$$-\frac{\partial p_D}{\partial r_{De}} \Big|_{r_{De}=1} = 1 - C_{De} \frac{d p_{wD}}{d t_{De}} \quad (9) \quad p_D(\infty, t_{De}) = 0 \quad (10)$$

其中无因次定义为

$$\text{无因次压力} \quad p_D = \frac{Kh}{1.842 \times 10^{-3} q \mu B} (p_i - p)$$

$$\text{无因次时间} \quad t_D = \frac{3.6Kt}{\phi \mu c_i r_w^2}; \quad t_{De} = t_D \cdot e^{2s}$$

$$\text{无因次井筒储存系数} \quad C_D = \frac{C}{2\pi \phi c_i h r_w^2}; \quad C_{De} = C_D \cdot e^{2s}$$

$$\text{无因次半径} \quad r_D = \frac{r}{r_w}; \quad r_{De} = r_D \cdot e^s$$

对式(7)、(8)、(9)和(10)取关于  $t_{De}$  的拉普拉斯变换并求解, 可获得拉普拉斯空间无因次井底压降解

$$\bar{p}_{wD} = \frac{K_0(\sqrt{z})}{z[\sqrt{z} K_1(\sqrt{z}) + C_{De} z K_0(\sqrt{z})]} \quad (11)$$

对式(11)进行拉普拉斯数值反演<sup>[9]</sup>, 得到实空间无因次井底压降解  $p_{wD}\left(\frac{t_D}{C_D}\right)$ 。

利用压力叠加原理可得到压力恢复解, 即

$$p_{wsD}\left(\frac{t_D}{C_D}\right) = p_{wD}\left(\frac{t_D}{C_D}\right) + p_{wD}\left(\frac{t_{pD}}{C_D}\right) - p_{wD}\left(\frac{t_{pD}}{C_D} + \frac{t_D}{C_D}\right) \quad (12)$$

式中  $t_{pD} = \frac{3.6Kt_p}{\phi \mu c_i r_w^2}$ ,  $t_p$  为关井前稳定生产时间。

由式(11)和式(12)就可分别计算出均质油藏的压力降落试井理论曲线和压力恢复试井理论曲线。 $p_{wD}$  或  $p_{wsD}$  与下面介绍的遗传算法的适应度函数密切相关。

#### 3.2 基于自适应遗传算法的最优化试井分析方法

最优化试井分析就是要选择最佳的井筒和油藏参数(如渗透率  $K$ 、表皮系数  $S$  和井筒储存系数  $C$  等)使实测曲线与理论曲线达到最佳拟合, 也就是求下列函数的极小值。

$$E = \sum_{i=1}^n \left| \frac{p_i - F(\vec{\theta}, t_i)}{F(\vec{\theta}, t_i)} \right| \quad (13)$$

式中  $n$  是实测数据的点数;  $\vec{\theta}$  是未知参数矢量;  $t$  是实测时间数据值;  $p$  是实测压力数据值;  $F$  是理论压力值。

对于均质油藏试井分析理论,  $\vec{\theta} = (K, S, C)^T$ ,  $F$  可由式(11)或式(12)和压力的无因次定义计算出来。由于  $F$  是关于油藏参数  $K$ 、 $S$  和  $C$  的非线性函数, 因此  $E$  也是关于油藏参数  $K$ 、 $S$  和  $C$  的非线性函数。而遗传算法是求解非线性函数十分有效的方法。采用的自适应遗传算法的具体实现步骤如下。

(1) 定义适应度函数



$$f = \begin{cases} C_{\max} - E(\vec{\theta}, t) & E(\vec{\theta}, t) < C_{\max} \\ 0 & E(\vec{\theta}, t) \geq C_{\max} \end{cases} \quad (14)$$

式中  $C_{\max}$  是一个合适的输入值,如: $C_{\max}$ 可采用迄今为止进化过程中  $E(\vec{\theta}, t)$  的最大值。

(2) 初始群体的生成 使用随机实验方法产生  $n$  个可行解  $\vec{\theta}_i(k)$  ( $1 \leq i \leq n$ ) 组成初始解群, 然后进行编码,

$$\vec{\theta}_i = \vec{\theta}_{\max} + (\vec{\theta}_{\max} - \vec{\theta}_{\min}) \cdot \text{random}(0, 1) \quad (15)$$

可取  $n=60$ ,  $0 < K \leq 10$ ,  $-10 \leq S \leq 10$ ,  $0 < C \leq 10$ 。random(0, 1) 为 0~1 之间的随机数。

(3) 编码 采取十进制编码方式。

(4) 适应度的计算 对于每一个个体  $\vec{\theta}_i(k)$  (变量  $k$  称作“代数”, 初始时  $k=1$ ), 利用式(14)计算适应度  $f[\vec{\theta}_i(k)]$ , 再由适应度  $f[\vec{\theta}_i(k)]$  确定最大适应度值  $f_{\max}$  和平均适应度  $\bar{f}$ 。

$$\bar{f} = \frac{1}{n} \sum_{i=1}^n f[\vec{\theta}_i(k)] \quad i = 1, 2, \dots, n$$

(5) 选择 由复制概率  $p_i(k)$  计算出每个个体被选择的次数  $M$ ,

$$p_i(k) = \frac{f[\vec{\theta}_i(k)]}{\sum_{i=1}^n f[\vec{\theta}_i(k)]} \quad M = \frac{f[\vec{\theta}_i(k)]}{\bar{f}} = n \cdot p_i(k)$$

复制概率大的最优秀个体将获得最多的生存繁殖机会, 最差的个体将被淘汰。注意, 要保证群体总数  $n$  不变, 选择过程中, 采取最优保存策略, 以防止变异操作对最优解个体的破坏作用。

(6) 交换 随机配对, 确定配对个体  $\vec{\theta}_1$  和  $\vec{\theta}_2$  中较大的适应度值  $f'$ 。由式(3)和式(5)计算交换概率  $p_c$ , 再根据交换概率  $p_c$  执行交换操作, 产生新的个体  $\vec{\theta}_1$  和  $\vec{\theta}_2$ 。

$$\vec{\theta}_1 = \alpha \vec{\theta}_1 + (1 - \alpha) \vec{\theta}_2 \quad \vec{\theta}_2 = \alpha \vec{\theta}_2 + (1 - \alpha) \vec{\theta}_1$$

式中  $\alpha$  可取为 0.8。

(7) 变异 由式(2)和式(4)计算变异概率  $p_m$ , 执行变异操作, 产生新的个体  $\vec{\theta}_i$ 。

$$\vec{\theta}_i = (1 - \beta) + \beta [\text{random}(0, 1) \cdot (\vec{\theta}_{\max} - \vec{\theta}_{\min}) + \vec{\theta}_{\min}]$$

式中  $\beta$  随“代数” $k$  的增加而呈指数下降。

(8) 判断条件 考察  $|f_{\max} - \bar{f}|$  是否达到预先给定的精度  $\epsilon$ , 如果达到就停止, 否则执行第(9)步。

(9) 改选方法 考察“代数” $k$  是否达到预先给定的值, 如果达到就改选方法, 否则转到第(4)步继续。

改选方法是指改用非线性回归自动拟合试井分析等其他优化方法, 如 Levenberg-Marquardt 方法。在实例资料解释中发现, 遗传算法具有很强的全局搜索能力, 而局部搜索能力却不如非线性回归方法。遗传算法能够以极快的速度达到最优解的 90% 左右, 但要达到真正的最优解则需要花费很长的时间。这样, 将遗传算法与非线性回归自动拟合试井分析方法相结合, 把遗传算法第  $k$  “代”的最优个体作为非线性回归方法的初值, 既保证了算法的快速收敛, 又保证了解释结果具有较好的唯一性。

## 4 实例分析

下面给出两口井的压力恢复试井资料的分析实例, 这两口井位于低渗透油藏中。

井 1 稳定生产 8000h 后, 关井进行压力恢复测试了 436h。井和油藏的基本参数为: 产量  $q$  为  $4.5 \text{ m}^3/\text{d}$ , 井半径  $r_w$  为 0.1m, 有效厚度  $h$  为 8.9m, 孔隙度  $\phi$  为 0.14, 流体粘度  $\mu$  为  $4.6 \text{ mPa} \cdot \text{s}$ , 综合压缩系数  $c_i$  为  $0.001306 \text{ MPa}^{-1}$ , 体积系数  $B$  为 1.106。

井 2 稳定生产 4320h 后, 关井进行压力恢复测试了 192h。井和油藏的基本参数为: 产量  $q$  为  $4.9 \text{ m}^3/\text{d}$ , 井半径  $r_w$  为 0.1m, 有效厚度  $h$  为 25.9m, 孔隙度  $\phi$  为 0.114, 流体粘度  $\mu$  为  $4.2 \text{ mPa} \cdot \text{s}$ , 综合压缩系数  $c_i$  为

0.00095MPa<sup>-1</sup>, 体积系数  $B$  为 1.089。

用本文方法对井 1 和井 2 进行分析, 解释结果见表 1。

表 1 实例分析结果  
Table 1 Analysis results of two field examples

	自适应遗传算法(初值)				Levenberg-Marquardt 方法(终值)			
	渗透率 ( $\mu\text{m}^2$ )	表皮系数	井筒储存系数 ( $\text{m}^3/\text{MPa}$ )	代数	渗透率 ( $\mu\text{m}^2$ )	表皮系数	井筒储存系数 ( $\text{m}^3/\text{MPa}$ )	迭代次数
井 1	0.00140486	-3.112	0.7962	10	0.0019343	-2.01121	0.77343	18
	0.00186305	-2.151	0.7575	80	0.0019343	-2.01134	0.77343	13
井 2	0.00061249	-2.216	1.4663	10	0.0007477	-2.23769	1.13957	9
	0.00066572	-2.216	1.3846	80	0.0007473	-2.23955	1.13929	10

从表 1 中可以看出, 在遗传算法中, 代数为 10 和代数为 80 的解释结果相差不太大。不论用第 10 代还是用第 80 代的解释结果作为非线性回归方法—Levenberg-Marquardt 方法的初值, 最终解释结果的唯一性好, 收敛速度快。图 1 和图 2 显示了在 AGA 算法的进化过程中, 最大适应度和平均适应度的变化趋势。图 3 和图 4 显示了自适应遗传算法与非线性回归方法—Levenberg-Marquardt 方法结合后最终的双对数拟合分析图。

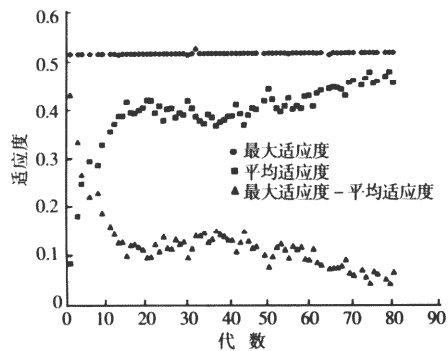


图 1  $f_{\max}$ ,  $\bar{f}$  和  $f_{\max} - \bar{f}$  的变化(井 1)

Fig. 1 Variation of  $f_{\max}$ ,  $\bar{f}$  and  $f_{\max} - \bar{f}$  (Well 1)

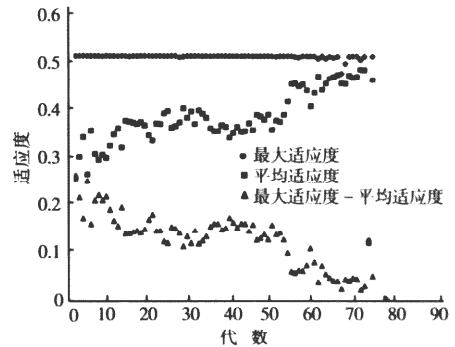


图 2  $f_{\max}$ ,  $\bar{f}$  和  $f_{\max} - \bar{f}$  的变化(井 2)

Fig. 2 Variation of  $f_{\max}$ ,  $\bar{f}$  and  $f_{\max} - \bar{f}$  (Well 2)

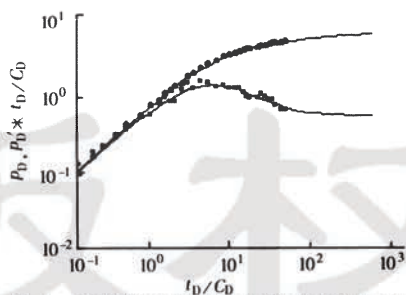


图 3 实测数据与典型曲线拟合(井 1)

Fig. 3 Matching of measured data with type curve (Well 1)

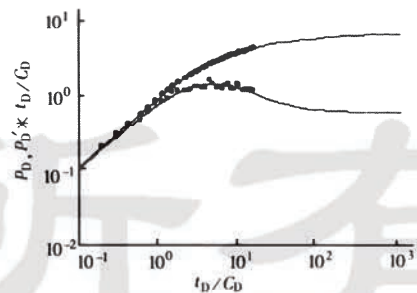


图 4 实测数据与典型曲线拟合(井 2)

Fig. 4 Matching of measured data with type curve (Well 2)

从以上分析情况来看, 本文提出的最优化试井分析方法, 其算法稳定, 效果良好。另外, 这种方法人工干预少, 易于实现试井资料分析的自动化。

## 5 结 论

1. 本文提出了一种基于自适应遗传算法的试井分析最优化方法。该方法主要是通过解群体的适应度值, 自适应地改变遗传算法的交换概率和变异概率, 从而保证群体的多样性和算法的收敛能力。
2. 自适应遗传算法与非线性回归方法—Levenberg-Marquardt 方法相结合, 形成的算法收敛速度快, 稳定性能强, 最终解释结果的唯一性好。
3. 本文方法只需要掌握井筒和油藏参数的值域, 而不需要估计井筒和油藏参数的初值, 也不要求目标函数连续可微, 人工干预少, 易于实现试井资料分析的自动化。
4. 本文只针对无限大均质油藏试井分析理论进行了研究。这种方法还可进一步扩展到其他油藏试井理论分析中去, 如: 天然裂缝油藏、复合油藏等。

### 参 考 文 献

- [1] Nanba T, Horne R N. An improved regression algorithm for automated well test analysis. SPE18161, 1988;543~555.
- [2] Rosa A J et al. Automated well test analysis using robust (LAV) nonlinear parameter estimation. SPE 22679, 1991;191~206.
- [3] Holland J H. Adaptation in nature and artificial systems. The University of Michigan Press, 1975, MIT Press, 1992;1~10.
- [4] 孙艳丰等. 遗传算法在优化问题中的应用研究进展. 控制与决策, 1996, 11(4): 425~431.
- [5] Mrinal K. Sen. et al. Stochastic reservoir modeling using simulated annealing and genetic algorithms. SPEFE, 1995; 49~55.
- [6] Martinez E R. et al. Application of genetic algorithm on the distribution of gas—lift injection. SPE 26993, 1994;11~18.
- [7] 陈国良等. 遗传算法及其应用. 北京: 人民邮电出版社, 1996;28~98.
- [8] Srinivas M et al. Adaptive probabilities of crossover and mutation in genetic algorithms. IEEE Transactions on systems, man and cybernetics, 1994, 24(4): 656~666.
- [9] Stehfest H. Numerical inversion of Laplace transforms. Algorithm 368, Communications of ACM, 1970, 13(1): 47~49.

(本文收到日期 1998-06-09 编辑 杨 苗)

### SJ- I 型分级注水泥器研制成功

由石油勘探开发科学研究院廊坊分院完井所承担的《SJ- I 型分级注水泥器研制》课题是集团公司“深井固井技术研究”的重要组成部分, 所研制的 SJ- I 型分级注水泥器经现场试验证明完全达到了设计和施工要求。

该仪器在江汉油田王新 78 井和江苏油田盐城 1 井、圣科 1 井进行了试验,  $\phi 244.5\text{mm}$  套管最大下深 3338.04m, 分级箍下深 1700.88m。施工顺利, 成功率 100%。

经有关专家评定认为: SJ- I 型分级注水泥器(双接箍)的设计新颖, 结构简单, 性能可靠; 溢流装置的设计完全避免了打开套和关闭套之间水泥浆的存在, 确保了关闭套顺利到位; 锁紧装置的设计保证了上、下内套动作的可靠性, 避免了分级箍的提前关闭。该仪器具有良好的推广应用前景。

(李成岗供稿)