

文章编号:1673-5005(2007)06-0010-03

资源总量随机求和中存在的问题及改进方法

李汉林, 傅晓宁, 翟汝霞

(中国石油大学 地球资源与信息学院, 山东 东营 257061)

摘要:针对现有的资源总量随机求和中存在的资源总量曲线形态有时呈偏态分布的问题,提出了对其随机抽样方法的改进方案,即把随机抽样入口由原来定义在 $(0,1)$ 区间上改为定义在 $(0,k_i)$ 区间上(k_i 为第 i 个局部地质单元的石油资源量保险系数),利用 $(0,k_i)$ 区间上均匀分布的随机数对风险分析后的资源总量分布函数进行随机抽样。改进后的方法克服了传统抽样方法的不足,实现了均匀分布随机抽样,使资源总量趋于正态分布。

关键词:地质单元; 资源总量; 随机抽样; 分布函数

中图分类号:TE 122.35

文献标识码:A

Improvement of problems and methods in random summing of total resources

LI Han-lin, FU Xiao-ning, ZHAI Ru-xia

(Faculty of Geo-Resource and Information in China University of Petroleum, Dongying 257061, Shandong Province, China)

Abstract: In the existing methods of random summing of total resource, the shape of the total resource curve sometimes distributes skewly, in view of which, the new method was put forward to improve random sampling. The entrance of the random sampling was not defined in the interval of $(0, 1)$, but in the interval of $(0, k_i)$, where k_i is the insurance index of the petroleum resource of the i th local geological location. The random numbers distributing in the interval $(0, k_i)$ were used to make a random sampling to the resource distribution function through the risk analysis. The improved method overcomes the flaw of the traditional sampling method, realizes the even distribution random sampling and makes the total resource tend to normal distribution.

Key words: geological location; total resource; random sampling; distribution function

1 风险分析前后资源量的分布

假设已由蒙特卡罗法分别求出了探区 m 个局部地质单元的石油资源量若干估计值,并用频率统计法求得其分布函数 $F(q_i)$ ($i=1, 2, \dots, m$)。经地质风险分析后,若保险系数为 k_i ($i=1, 2, \dots, m$),那么石油资源量分布曲线在纵向上下降为原曲线的 k_i 倍^[1-2](图1)。

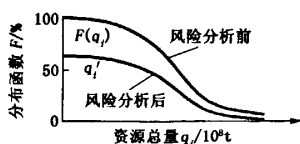


图1 风险分析前、后资源总量分布曲线

2 随机抽样求资源总量中存在的问题

2.1 随机抽样过程

传统的随机抽样求和过程是:用第 i 个局部地质单元的最小石油资源量 q_i^* 和区间 $(0,1)$ 上的随机数 r 在其分布函数的纵轴上确定一个称为随机入口的随机点 (q_i^*, r) ,过该点作横轴的平行线交 $F(q_i)$ 于点 (q_{ir}, r) ,交点的横坐标 q_{ir} 则是该地质单元石油资源量 q_i 对应于随机数 r 的随机抽样值,常称其为随机出口值, m 个随机出口值 q_{ir} ($i=1, 2, \dots, m$)的和为资源总量的一个随机抽样值^[3-5](图2)。该抽样过程是用软件中的插值方法实现的。

收稿日期:2007-05-10

作者简介:李汉林(1947-),男(汉族),山东寿光人,教授,硕士,从事石油数学地质、油气资源评价和计算机语言等方面的教学与科研工作。

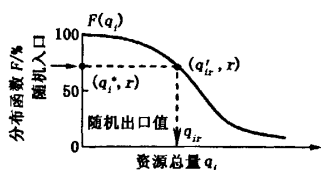


图2 随机抽样过程示意图

2.2 风险分析后求资源总量中的问题

对于地质风险分析后的石油资源量 q'_i ($i=1, 2, \dots, m$) 来说,其分布曲线在纵向上已下降为原曲线的 k_i 倍,若仍用 $(0,1)$ 区间上的随机数 r 对 q'_i 进行随机抽样,必然会出现与 q'_i 分布曲线无交点的情况。在此情况下,传统的随机抽样做法是:当随机入口的纵坐标,即随机数 $r > k_i$ 时,取随机抽样出口值为 0 (图3);当 $r < k_i$ 时,取出口值为 q'_{ir} 。 m 个随机抽样值的和为资源总量的一个估计值^[1-2]。传统抽样处理方法将会导致如下问题:

① 参加求和的局部地质单元数小于 m , 使随机抽样值的和 (资源总量的一次随机抽样值) 减小,因而探区石油资源总量小值一侧抽样和的频率增大,并且大量出现概率为 1 的抽样值,而大值一侧抽样和的频率降低,且出现大量概率为 0 的抽样值。

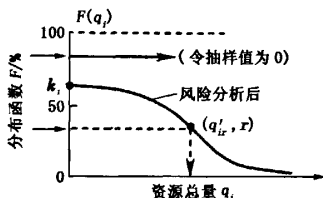


图3 风险分析前、后的随机抽样

② 由若干次资源总量随机抽样值得到资源总量分布函数 $F(q')$, 其曲线的大值一侧缓慢下降, 呈现出偏态的长尾分布。

③ 为使资源总量分布函数正态化, 须采取删除大量概率为 0 和 1 的抽样值并进行内插整理的方

法, 这就改变了抽样的随机性原则。

3 随机抽样方法的改进

在对地质风险分析后的资源量随机抽样求和过程中, 要想不发生分布函数横轴的平行线与分布曲线无交点的情况, 需要改变随机抽样的入口区间, 使随机入口点落在 $(q_i^*, 0)$ 与 (q_i^*, k_i) 两点之间, 即使随机数 r 落在区间 $(0, k_i)$ 内。

假设地质风险分析后随机抽样的随机数为 $(0, k_i)$ 区间上均匀分布的随机数 v , 那么, 令 $v = k_i r$, 则将 $(0,1)$ 区间上均匀分布的随机数 r 变换为 $(0, k_i)$ 区间上均匀分布的随机数 v 。利用随机数 v 实施随机抽样时, 随机入口点必然落在 $(q_i^*, 0)$ 与 (q_i^*, k_i) 两点之间, 由此就避免了资源量随机抽样值为 0 的情况 (图3)。随机抽样方法改进后, 求 m 个局部地质单元的资源总量有如下优点:

① 资源量随机抽样求和过程中, 确保参加求和的局部地质单元数等于 m 。

② 资源总量的每一次随机抽样值都是 m 个局部地质单元资源量随机抽样值的概率和。

③ 由于 v 是 $(0, k_i)$ 区间上均匀分布的随机数, 因此它使探区石油资源总量区间的小值一侧抽样和的频率下降, 大值一侧抽样和的频率升高, 避免了多余的概率为 0 和 1 的抽样值的出现。

④ 由若干次资源总量随机抽样值得到资源总量分布函数 $F(q)$, 其曲线形态趋于正态分布。

4 随机抽样效果对比分析

我国某沉积盆地中的一个地质凹陷有两套生油层系。若用氯仿沥青法估算该凹陷远景石油资源量, 则第 i 套生油层系石油资源量 q_i 的估算公式为

$$q_i = S_i h_i \rho_i w_A k_1 k_2, \quad i = 1, 2.$$

式中的地质参数名称及数据见表1。

表1 生油层地质参数名称及有关数据 (据赵旭东修改)

生油层系	分布面积 S/km^2	生油岩密度 $\rho/(10^3 \text{ t} \cdot \text{km}^{-3})$	排烃 系数 k_1	聚集 系数 k_2	生油岩厚度 h/m		氯仿沥青“A”含量 $w_A/\%$		资源量 保险系数
					数据个数	取值范围	数据个数	取值范围	
第一套	14000	23	0.44	0.111	140	0.1~1.0	37	0.03~1.74	0.75
第二套	7000	23	0.48	0.111	70	0.1~1.0	37	0.02~2.08	0.75

根据凹陷内两套生油层系的资料, 用蒙特卡罗法估算了两套生油层系的石油资源量, 风险分析后, 分别给出了传统随机抽样法资源总量内插整理前、后的分布曲线和随机抽样方法改进后资源总量分布曲线 (图4)。

传统随机抽样资源总量分布曲线 (图4(a)), 其小值一侧有大段平直线段, 即有很多资源总量的概率为 1, 这是当随机数 $r > k_i$ 时, 抽样值取 0 的缘故。对于资源总量分布函数来说, 曲线的这种形态显然是不合理的。删除资源总量中大量概率为 0 和

1 的数据后,再经过内插整理,得到的资源总量分布曲线(图 4(b))形态出现明显的正态化,但小值一侧仍有一段平缓的线段。方法改进后,由于随机抽

样入口点在 $(0, k_i)$ 区间上均匀分布,不再出现多余的概率为 0 和 1 的抽样值,克服了传统抽样法中的不足,得到的资源总量趋于正态分布(图 4(c))。

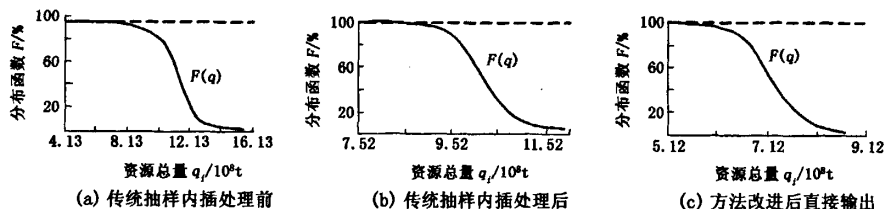


图 4 石油资源总量分布曲线

5 结论

(1)求风险分析后的多个局部地质单元的资源总量时,利用 $(0, k_i)$ 区间上均匀分布的随机数抽样技术是合理的,它使资源总量趋于正态分布。

(2)当 $k_i = 1$ 时,改进前、后的抽样方法是等价的,因为随机抽样入口点在 $(0, 1)$ 区间上都是均匀分布。

(3)改进后的抽样方法优化了数据处理工作,但仍出现部分概率很小、数值相差甚微的数据,使分布曲线的大值一侧平缓延伸。

参考文献:

[1] 赵旭东. 石油数学地质概论[M]. 北京:石油工业出版社

社,1992:194-228.

[2] 李汉林,赵永军. 石油数学地质[M]. 东营:石油大学出版社,1998:261-278.

[3] 《油气资源评价方法研究与应用》编委会. 油气资源评价方法研究与应用[M]. 北京:石油工业出版社,1988:254-265.

[4] 杨通佑,范尚炯. 石油及天然气储量计算方法[M]. 北京:石油工业出版社,1998:218-227.

[5] 武守诚. 石油资源地质评价导论[M]. 北京:石油工业出版社,1994:245-252.

(编辑 刘艳荣)

(上接第 5 页)

[3] 王素华. 下切谷及其油气勘探意义[J]. 地球学报, 2003, 23(增刊):30-32.

WANG Su-hua. Entrenched valley and its significance of the hydrocarbon exploration activity[J]. Acta Geoscientica Sinica, 2003, 23(sup):30-32.

[4] 林畅松,张燕梅,刘景彦,等. 高精度层序地层学和储层预测[J]. 地学前缘,2000,7(3):111-117.

LIN Chang-song, ZHANG Yan-mei, LIU Jing-yan, et al. High resolution sequence stratigraphy and reservoir prediction[J]. Earth Science Frontiers, 2000, 7(3):111-117.

[5] 邓宏文,王红亮,宁宁. 沉积体分配原理——高分辨率层序地层学的理论基础[J]. 地学前缘,2000,7(4):

305-314.

DENG Hong-wen, WANG Hong-liang, NING Ning. Sediment volume partition principle: theory basis for high-resolution sequence stratigraphy[J]. Earth Science Frontiers, 2000, 7(4):305-314.

[6] CROSS T A,杜宁平. 据高分辨率层序地层学认识地层结构对比概念,体积配分,相分异[J]. 国外油气勘探,1996,18(3):285-294.

[7] 刘招君,董清水,王嗣敏,等. 陆相层序地层学导论与应用[M]. 北京:石油工业出版社,2002:36-41.

[8] SANGREE J B, VAIL P R. 应用层序地层学——地震、钻井和露头资料的层序地层解释[M]. 张宏彦,等译. 北京:石油工业出版社,1991:85-97.

(编辑 刘艳荣)