

# 预测油气田数目及规模分布的一种新途径 ——帕雷托截断分布模拟法<sup>①</sup>

刘毅 李军 程学福

(地矿部石油地质中心实验室, 无锡 214151)

本文把油气资源评价的成因法和统计模拟法相结合, 建立了帕雷托截断分布模拟法, 预测大型含油气地质单元的油气田数目及规模分布。本方法提出了对资源评价具有重要意义的“资源集中程度”参数, 采用迭代逼近算法实施模拟运算。最后应用此方法对塔里木盆地石油资源进行了预测, 并就有关参数的选取作了探讨。

关键词 帕雷托截断分布 资源集中程度 油气田数 资源量

第一作者简介 刘毅 男 30岁 工程师 石油地质

油气资源评价的一个重要方面是预测油气田数目及其规模分布, 这方面的信息将有助于油气的勘探部署和经济分析。现有的油气田规模分析方法主要是统计模拟法 (Baroch 等, 1977; L. P. White, 1979; P. J. Lee 等, 1982) 和成因法 (如: 有机碳法、盆地模拟法)。作为相互独立的评价方法, 统计模拟法和成因法对同一地质单元的总资源量的评价结果相差较大, 前者远低于后者。此外, 上述两类方法均没有考虑“资源集中程度”参数, 这是一个地质单元资源经济评价中非常重要的参数, 如果总资源主要集中在少数几个大油气田中, 则其勘探效益必然较好。

鉴于此, 我们尝试将成因法和统计模拟法结合起来, 以成因法总资源量为输入初值, 并增设资源集中程度参数, 建立了一种预测油气田数目和分布的资源评价方法, 其分布概型为帕雷托截断分布, 故称之为“帕雷托截断分布模拟法”, 评价单元为大型区块或含油气盆地。

## 1 基本思路

帕雷托 (Pareto) 截断分布模拟法的基本思路可以概括为以下几个主要环节。

首先, 获取评价区成因法总资源量值, 它是和统计模拟法联系的基础。把总资源量细分成油气田数

目及其规模分布是本方法的显著特色, 有别于传统统计模拟法从油气藏评价到总资源量预测的模式。

其次, 使用评价区已知油气田的储量数据资料, 按小、中、大型油气田的概念, 建立油气田规模分组, 确定各组的界限值。结合评价区实际, 确定经济性的最小油气田地质储量和最大可能的油气田规模。该分组区间反映了现有已知油气田的规模分布, 同时也是预测待发现油气田规模分布的基础。

然后, 在上述分组的基础上, 按资源集中程度参数确定最大一组油气田资源占总资源的比例, 据此求取 Pareto 分布的参数值。

最后, 按 Pareto 分布密度函数, 求取各组资源规模及油气田数目分布的结构表, 用蒙特卡洛模拟方法检验结果, 如得不出可以接受的结果, 修改模拟参数直到得到可以接受的结果为止 (图 1)。

## 2 Pareto 截断分布模拟法中使用的主要技术方法

### 2.1 Pareto 截断分布密度

采用下述的截断帕雷托分布描述地质单元内油气田的规模频率分布, 密度函数为:

$$f(q; c, \lambda) = \begin{cases} c[q^{-\lambda} - (\gamma Q)^{-\lambda}] & \text{当 } q_0 \leq q \leq \gamma Q \\ 0 & \text{其它} \end{cases} \quad (1)$$

① 属“八五”国家重点科技攻关项目 85-101-02-04 的一部分

式中： $0 < q_0 \leq q \leq \gamma Q$ ； $\gamma C(0, 1)$ ； $C > 0, \lambda > 0$ ；

$q$  为油气田的地质储量；

$C$  为标定因子；

$\lambda$  为分布参数；

$Q$  为成因法总资源量

$q_0$  为经济性的最小油气田地质储量；

$[q_0, \gamma Q]$  为分布密度的截断区间

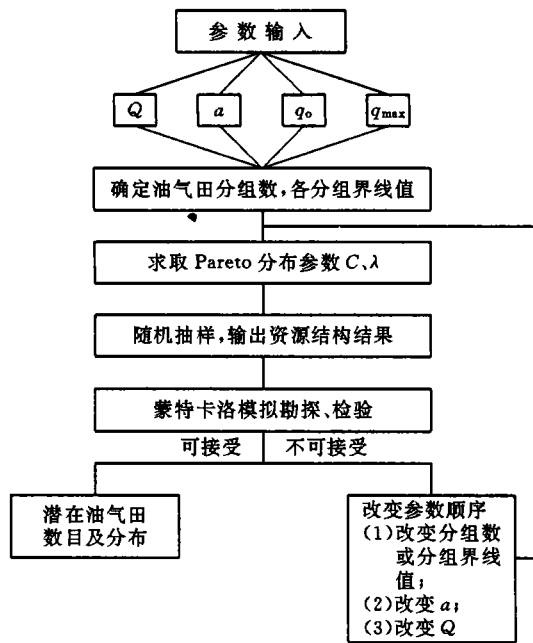


图 1 Pareto 截断分布模拟法流程框图

为了确定上述分布中的参数  $C$  和  $\lambda$ , 除了要知道  $q_0$ 、 $Q$  和  $\gamma$  以外, 还要知道资源集中程度  $a$ , 即集中于  $q \geq q_0$  的油气田资源量。 $q_0$  通常是给定的。 $Q$  是通常用成因法计算得出的总资源量值, 但由于生聚系数较难取准, 因此,  $Q$  要在迭代中逐渐逼近。

通过解如下的超越方程估算参数  $C$  和  $\lambda$ ：

$$\int_{q_0}^{\gamma Q} Q(q; C, \lambda) dq = 1 \quad (2)$$

图 2 是  $\varphi(q; C, \lambda)$  变化图, 在坐标系  $(q, \varphi)$  中, 经过某一点  $A$  总共有两条曲线它们与横轴  $q$  所围成的面积 (广义) 均为 1。

同时,  $C, \lambda$  还应满足以下约束条件：

$$\int_{q_0}^{\infty} Q(q; C, \lambda) dq = T_a \quad (3)$$

即按资源集中程度  $a$  确定的最大一组油气田的数目比  $T_a$ , 通过  $(C, \lambda)$  的搜索进行迭代求解, 具体算法见有关文献<sup>①</sup>。

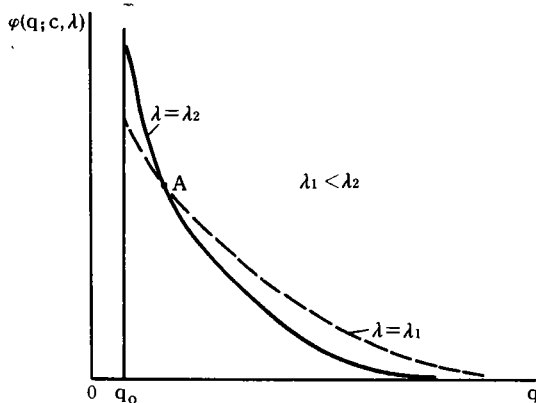


图 2  $\varphi(q; C, \lambda)$  变化图

### 2.2 蒙特卡洛模拟勘探模型

首先确定油气田规模的分组。一般可从  $q=5$  组对  $[q_0, \gamma Q]$  区间进行划分, 将其对应于当地情况的小、中小、中、大中和大型的油气田的概念。根据专家的经验初步确定 5 组的界线, 也可以修改分组数。

根据选定的参数  $C, \lambda$  及分组数, 按一定的随机抽样法选取  $\varphi(q)$  的随机数, 进行蒙特卡洛模拟, 限于篇幅, 其方法从略。采用人机交互式方法, 计算机将自动地把随机抽样的油气田数分别填放在各自的组段内, 并根据总资源量  $Q$  按比例调整各组段的油气田个数和资源量。同时, 把已经发现的油气田个数和资源量分别填入。如果两者相近, 则认为上述分组是合理的, 否则重新分组。

当分组数和分组界线确定后, 可能出现模拟勘探与实际发现不一致的情况。这通常是由于资源集中程度  $a$  取值不合理所致。如果较大型的油气田数目少或是“额外”地“发现”了大量的油气田, 这很可能是  $a$  取值偏小。通过改变  $a$  以使模拟的发现与实际的发现非常接近。

如果在上述条件下, 仍然得不出可以接受的分布形式, 则原因就在于生聚系数的取值不合理导致总资源量  $Q$  的差别 (通常是偏高)。通过改变  $Q$  重复模拟直至与实际相符。

实施本方法的关键是：①分组是否合理；② $\gamma Q$  值是否合理；③ $C, \lambda$  是否合理；④ $Q$  值是否合理。

模拟的结果是资源分布结构表, 用它删除实际的发现即可得到尚未发现的资源结构。

### 2.3 几点说明

(1) 本方法应用于盆地或拗陷、凹陷等在地生、运、

① 程学福等. 天然气资源勘探决策系统研究, “八五”国家科技攻关项目成果报告, 85-102-01-05-02, 1994

聚等方面自成系统的地质单元。对某些区块则应从生成演化方面入手,对形成的资源量加以分配后确定模拟中的  $Q$  值。

(2)本方法采用通常的油(气)田的概念,而不采用油气藏的概念。这是因为发现概率与油气田规模成正比,不见得与油气藏规模成正比。某些小油气藏先发现可能预示着较大的油气田的存在。

(3)本方法在求取合理分布时,是按照对该分布进行不放回的随机抽样结果与实际的勘探结果进行对比判定的。如果在勘探初期,缺乏足够的实际油气田样本,那么可主要依据资源集中程度给出好、中、差 3 个值分别评价。同时,由于现行成因法资源量大多以概率值的形式给出,模拟时也可以作出平均估计(中值概率资源量)、乐观估计(5%概率值)、保守估计(95%概率值)等不同情况下的资源结构分布,以便给决策人员提供某种信息,有助于经济分析和勘探决策。

### 3 应用实例

在“八五”国家攻关课题研究中,我们采用帕雷托截断分布模拟法预测新疆塔里木盆地油气田数及其规模分布。

根据“八五”国家科技攻关的新成果(85-101-02-04-03),全盆油气资源潜量巨大,按成因法全盆油资源量(期望值)为  $145.35 \times 10^6 \text{t}$ 。作者认为以此作为 Pareto 分布模拟初值是合适的。

塔里木盆地属高勘探成本地区,这里取经济性的最小油田规模为  $150 \times 10^4 \text{t}$ 。

塔里木盆地是我国最大的沉积盆地,面积  $56 \times$

$10^4 \text{km}^2$ ,油气勘探程度甚低,目前仅在塔中 402 井发现了控制储量超过亿吨( $100.72 \times 10^6 \text{t}$ )的油气。根据国内外资料,沙特阿拉伯的加瓦尔油田的最终可采储量达  $113.7 \times 10^8 \text{t}$ ,我国大庆萨尔图油田探明储量  $25.69 \times 10^8 \text{t}$ ,克拉玛依油田约  $7.5 \times 10^8 \text{t}$ 。曾有国外学者研究认为,油气盆地的油气储量随沉积速率增大而增大,并按沉积速率将盆地分为 4 类。按其分类,塔里木盆地的体积沉积速率为  $8138 \text{m}^3/\text{Ma}$ ,相当于其第二类,最大油田规模约为  $18 \times 10^8 \text{t}$ 。也有人根据地质分析推断最大油田规模为  $5 \times 10^8 \text{t}$ ,综合考虑目前勘探程度和油气地质概况,我们在模拟时取最大油田规模为  $10 \times 10^8 \text{t}$ 。

这里,我们把油田规模分为 4 组: $1.5 \times 10^6 \sim 10 \times 10^6 \text{t}$ 、 $10 \times 10^6 \sim 40 \times 10^6 \text{t}$ 、 $40 \times 10^6 \sim 100 \times 10^6 \text{t}$ 、大于  $100 \times 10^6 \text{t}$ 。其中最大一组(即超亿吨)油田资源量占总资源量的比例是资源集中程度。国内外对这一参数的取值一般介于 0.25~0.5 之间。为了能提供更多的信息,我们分别按 1/4、1/3、1/2 不同情况作了模拟。

以资源集中程度  $a=1/2$  为例,则用 Pareto 分布模拟得出的塔里木盆地石油资源结构如表 1 所示(表 1 中的潜在资源结构是用模拟分布减去实际发现得到的)。

据此用蒙特卡洛抽样不放回的方法对勘探进行模拟并与实际勘探油气发现结果相比较,结果如表 2 所示。

勘探模拟与实际勘探结果十分接近。这说明上述资源结构是可以接受的。按  $a=1/3$ ,  $a=1/4$  亦得出合理的资源结构形式(图 3、图 4)。图中  $a$  是各分组的油田数,  $b$  表示各组占总资源量的比例。

表 1 塔里木盆地石油资源结构模拟数据表

预 测 值		总 量	模拟参数 $C=1.129$ $\lambda=1.8123$ $a=1/2$			
			规 模 分 组 ( $10^6 \text{t}$ )			
			1.5~10	10~40	40~100	>100
全盆总计 资源结构	油田平均规模( $10^6 \text{t}$ )	15.23	3.48	19.08	63.09	264.49
	油田数(个)	954.56	761.16	132.59	33.41	27.40
	资源量( $10^6 \text{t}$ )	14535.0	2651.68	2529.72	2107.78	7245.82
潜在资源 结构	油田数(个)	936.56	750.16	127.59	33.41	26.40
	资源量( $10^6 \text{t}$ )	14206.45	2592.81	2410.71	2056.75	7145.10

表 2 模拟勘探与实际勘探结果比较

项 目		总 量	规 模 分 组 (10 <sup>6</sup> t)			
			1.5~10	10~40	40~100	>100
实际勘探发现	油田数(个)	18	11	5	1	1
	资源量 10 <sup>6</sup> t	328.55	57.87	118.93	51.03	100.72
模拟勘探发现	油田数(个)	18	11	5	1	1
	资源量 10 <sup>6</sup> t	328.59	58.27	118.92	51.08	100.42

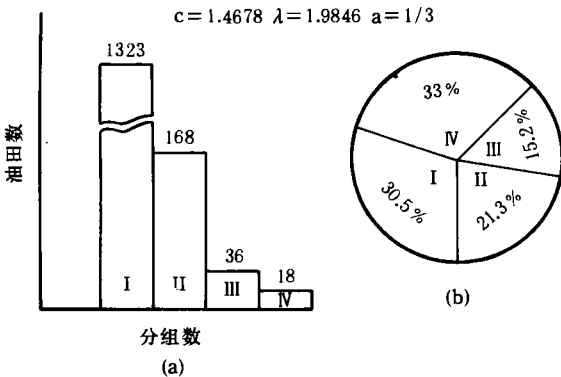


图 3  $\alpha=1/3$  时的资源结构图  
 $c=1.8324 \lambda=2.1497 a=1/4$

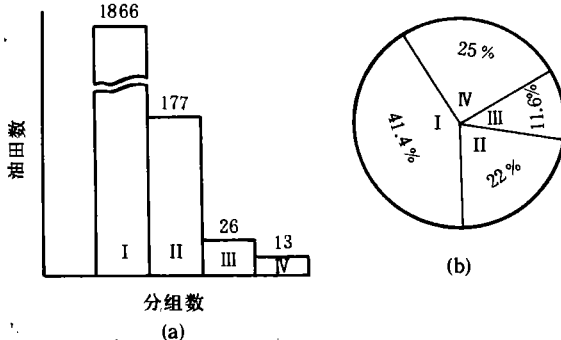


图 4  $\alpha=1/4$  时的资源结构图

模拟结果认为,塔里木盆地具有丰富的油气资源,其  $145.35 \times 10^8$  t 油资源量分布在 954~2085 个油田中,其中大于  $1 \times 10^6$  t 的油田在 12~26 个之间,占总资源量的 49%~24.47%,是已发现石油储量的 14.23~10.8 倍;介于  $40 \times 10^6 \sim 100 \times 10^6$  t 的油田 25~35 个,此外还有大量的小油田。如何寻找期望中的油田,特别是大中型油田,作者认为今后勘探的对象应以非构造圈闭为主。随着勘探工作的进展,发现新的构造圈闭油气藏(田)将会越来越困难。塔里木盆地次级构造单元拗隆相间,盆地内发育着

不同时期的拗陷、斜坡及继承性的古隆起、古斜坡为非构造圈闭的形成创造了条件。因此,只要加强地质综合研究,发育大型和大中型油田是可能的,勘探前景还是诱人的。

#### 4 结束语

上面提到,由于资源评价方法不同,结果互有出入。但 Pareto 截断分布模拟法以成因法资源量值为预测基础是否合理呢?据张抗等(1993)<sup>①</sup>的统计资料,我国各大区中探明程度最高的松辽盆地石油储资比达 65.52%,除南华北以外的江淮区为 69.43%,其中河南油田 94.7%。东部老区为 58.31%,这说明探明程度相对高的老区,成因法资源量是可信的,储资比向中部区、西北区有依次降低的趋势,这与勘探程度有很大的关系。我国天然气的勘探程度甚低,占主要储量的中部气区储资比仅 4.14%,天然气的勘探开发还有很大潜力。尽管资源量的计算受研究程度,主观认识等因素的影响,但我们认为,以有机质热演化理论为基础的成因法是值得信赖的,油气储量的发现将无限地逼近资源量是不容置疑的。此外,如果模拟得不出可接受的资源结构分布时,可以通过迭代产生 Q 值,以修正成因法 Q 值。

模拟所需的另一个重要参数是资源集中程度  $\alpha$ ,其取值的正确与否将直接影响模拟的结果。图 5 表示美国 13985 个油田按储量大小分布的特征(对数座标)。储量大于 50Mbb1 的大油田仅 440 个,占总油田数的 3%,而其储量却占总数 80%。这充分说明大多数的石油储量集中在少数的大油气田中。据统计资料,国内大庆萨尔图油田的储量占全国的

① 张抗等.《我国石油资源的近中期预测》,地矿部石油地质研究所,1993

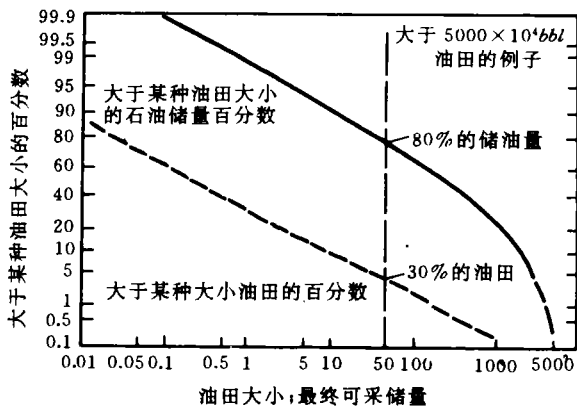


图 5 美国 13985 个油田的大小分布  
(据 R. A. Baker 等, 1992)

下面的曲线表示大于某种油田大小油田数量的百分数;  
上面的曲线表示大于某种油田大小油田中的石油总储量  
例如大于或等于  $5000 \times 10^4 \text{ bbl}$  油田, 仅占总油田数  
量的 3%, 而占石油总储量的 80%

16.16%, 仅松辽盆地的石油储量就占全国的 30.43%。全国  $1 \times 10^8 \text{ t}$  级以上的油田 27 个(1992 年), 占全国总储量的 61.23%。在预测大油气田规模分布时, 建立资源集中程度参数实属必要。然而, 由于油气地质过程的复杂性, 这一参数的取值更多

地是依赖于主观推测。不可否认, 对地质资料的详细了解及选取合适的类比区块(盆地)将有助于这一参数的选取。采用模拟勘探过程的方法, 挑选不同的  $a$  值进行迭代, 也能找到适当的  $a$  值。

此外, 本文对成因法资源量  $Q$  的迭代算法等未进行讨论, 其实现的方法还有待完善, 它们对模拟的结果可能会有影响。同时, 模拟所选的塔里木盆地勘探程度甚低, 已知油气田样本少。因此, 这种模拟对该盆地早期勘探阶段的油气研究具有重要意义。

参 考 文 献

- 1 Barouch E and G Kaufman. Estimation of undiscovered oil and gas; Proceeding of Symposia in Applied Mathematics, 1977, 21:77~91
- 2 lee PJ and Wang PCC. Analysis of pool size distributions for evaluations of petroleum resources. ISPG, 1982
- 3 Betty M. Miller, 翟光明译. 资源评价方法的选择和结果. 油气评价方法与应用, 北京:石油工业出版社, 1992, 1~30

(收稿日期:1996 年 2 月 2 日)

## A NEW APPROACH TO PREDICTING OIL AND GAS FIELD NUMBERS AND THEIR SIZE AND DISTRIBUTION—— A METHOD OF PARETO TRUNCATED DISTRIBUTION MODELLING

Liu Yi Li Jun Cheng Xuefu

(Central Lab of Petroleum Geology, MGMR)

**Abstract**

Based on the methods of origin and statistical modelling in hydrocarbon resource assessment, the method of Pareto truncated distribution modelling is constructed to predict oil and gas field numbers and their size and distribution in a mega-petroliferous region. The degree of resource concentration is also proposed in this paper as an important parameter in resource evaluation, and the modelling procedure is fulfilled through iterative and approximate algorithm. At last, a case application of the oil resource forecasting of Tarim Basin is presented and the selection of related parameters is discussed.