第八卷第一期　　石 油 实 验 地 质　　Vol.8, No.1
1986年3月　　EXPERIMENTAL PETROLEUM GEOLOGY　　Mar., 1986

估算油气资源量的HASP方法浅释

程 学 福

（地质矿产部石油地质中心实验室）

引　　言

HASP（Hydrocarbon Assessment System Processor）方法[1,2),[1]]是加拿大沉积与石油地质研究所建立的一种油气资源评价方法。该方法在统计数学的基础上，把油气资源量看成是多个随机变量（地质参数）的函数，用蒙特卡洛方法模拟其概率分布，并通过勘探层、地质风险分析和油气藏规模的条件概率分布等概念建立起全盆地（或区域）油气资源潜量估算的随机抽样模式。可以说，HASP方法是目前油气资源潜量估算方法中较为精细的一种。1980年以来，我部有较多的单位在应用这个方法。几年来的实践表明，目前在方法应用中出现的问题，仍然是对方法的基本原理和概念缺乏正确的认识。笔者根据近年来学习和实践的体会，对该方法试作一浅释，以期和从事这一工作而又缺乏统计数学基础的地质工作人员共同探讨应用中的问题。

加拿大油气资源量估算方法发展概况简述

油气资源量的估算是油气资源评价的重要组成部分。七十年代初期，由于石油价格的上涨，在西方国家掀起了一场“能源危机”的风波，世界各国均加强了油气资源评价的研究工作，1972年加拿大地质调查局曾对全加拿大的常规油气资源量作了一轮估算。他们根据L.G.Weeks提出的“盆地的分类是估算资源量的基础”的概念，把全加拿大38个盆地（或地区）共分为七个类型，用盆地类比法进行了资源量的估算，并导致加拿大政府作出“立即停止向美国出口石油”的决策。此后他们就把油气资源的评价工作作为他们的一项经常性工作，并着手研究新的评价方法。1976年他们在向政府提交的能源报告中首次提出了HASP方法以及用这个方法对九个不同地区作出的资源量估算。应用HASP方法可以对以下几个问题作出予测：（1）资源有多少？（2）油气藏的规模大小；（3）评价的可靠性；（4）为获得某一资源量应投入的勘探工作量，等等。HASP方法把地质参数用概率分布的形式表示，因此在资料较少的新区或早期勘探阶段，便于地质学家根据经验或用类比的方法作出主观估计；而在资料较多的老区或中、后期勘探阶段，可以直接使用实际的统计资料，使方法能在不同的勘探阶段获得应用，保证估算

1)Oil and Natural Gas Resources of Canada, 1976.

2)R.M.Procter and P.J.Lee，石油资源评价方法，石油地质与实验，1980.10.

结果处于动态的逼近真值的状态。此外，还由于采用了地质风险分析和油气藏的条件概率分布概念，把油气藏是否存在和资源量的多少分别进行分析，从而为从圈闭入手进行资源量的估算奠定了基础。HASP方法设计的随机抽样模式的结构是符合地质学家的逻辑思维和实际的勘探过程的。HASP方法的建立，表明加拿大油气资源评价方法的研究进入一个新的阶段。

1983年，Lee P.J.和Wang P.C.C连续发表了几篇文章[2].[3].[4].[5]，进一步对概率公式作了理论分析，使油气资源评价方法朝理论上更趋完善的方向发展。

油藏规模形成的随机模式

当前，资源量的估算已经从单值估算发展到区间值域和概率分布表示的阶段，大大改善了估算值的合理性。

油气勘探的实践告诉我们，一个盆地的许多油藏在区域和深度上的分布是不均匀的，资源量的多少也各不相同。设盆地A共有N个油藏，它们的资源量分别为 $q_1, q_2, \cdots, q_n$，则可以对油藏规模作出如下的统计学表示。

1.频率分布和累计频率分布

若把各油藏的资源量大小分成若干个等级，并分别统计各个等级区间的油藏个数占总油藏数的百分比，则可以作出油藏规模的频率分布直方图（如图1）。这张分布图表明，特大的和特小的油藏均较少，大部分是在某个区间范围内的“中小型”油藏。这时，盆地A的石油资源总量可以近似地用各区间的平均值与该区间的油藏数的乘积和表示：

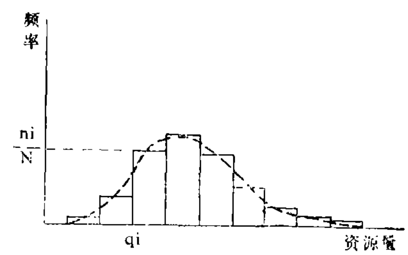

图1.油藏资源量大小的频率分布

$$Q=\sum_i n_i \cdot \frac{1}{2}(q_i+q_{i+1})$$

式中Q——盆地A的油气资源总量；

n_i——第i个区间的油藏数；

q_i——第i个区间的左端点值。

如果我们把N个油藏的资源量按由小到大的顺序加以排列，即 $q_1^* \leqslant q_2^* \leqslant \cdots \leqslant q_K^*$，则可以用统计资源量小于某个值的油藏累计频率的方法作出如图2所示的阶梯形的累计频率分布。它表明，资源量小于 q_1^* 的油藏一个也没有，而小于 q_K^* 的累计频率近似地为1，

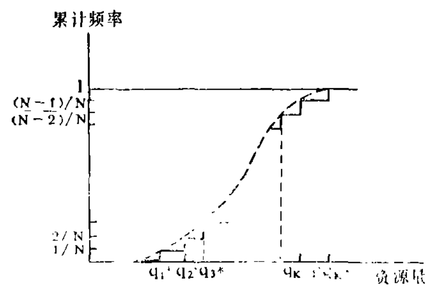

图2.油藏资源量累计频率(小于)分布

即所有的油藏其规模都是小于或等于 q_K^* 的。由于地质上习惯于用“大于”的说法，所以纵坐标改用1减去原累计频率值表示，则累计频率（大于）分布如图3所示。它表示，所有的油藏其规模均大于或等于 q_1^*，而大于 q_K^* 的油藏一个也没有。这时，盆地A的石

油资源总量为$Q=\sum\limits_{j} q_j = q_1 + q_2 + \cdots\cdots + q_n$。

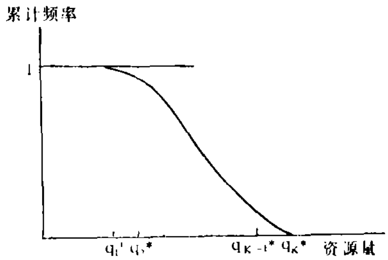

图3.油藏资源量累计（大于）分布

2.油藏规模形成的随机模式

上述油藏规模的概率分布（直方图和阶梯图）是盆地A这个实际的有限总体的分布。分析表明，油藏规模的特征（如含油面积、油层厚度、有效孔隙率、含油饱和度……等等），即使在一个盆地内，甚至在更小的地质单元内差别也是很大的，至今还找不到两个完全一样的油藏。为什么会这样呢？许多学者认为，油藏的形成条件非常复杂，除了那些我们目前能够认识的条件以外，还存在许多难以认识和控制的因素（随机因素）。这就使得形成油藏的确定性规模的条件充分性得不到保证，导致因果律的破缺。因此，同样的地质条件不能产生唯一确定的后果，而是按概率产生多种可能的结果。换句话说：在同样的地质条件下，油藏规模的理论总体是一个具有一定概率分布的无穷总体，这个概率分布就是理论分布（如图1和图2中的虚线分布）。而某一盆地的实际油藏规模分布只是其理论分布的一个随机样本，是对理论分布的一个近似。这是一个很重要的认识或看法，它告诉我们，一个盆地的油气资源在没有完全勘探清楚以前，虽然无法知道它总共有多少个油藏以及每个油藏规模的大小，但却可以通过一些随机样本去寻求它的总体分布。HASP方法是把油藏规模看成一系列地质变量的函数，用蒙特卡洛算法来模拟它的总体分布，

蒙特卡洛（Monte Carlo）方法

在HASP方法中广泛地采用了蒙特卡洛算法。这是一种统计模拟法，早期由于在高能物理中应用较多，所以物理学家首先给这一方法取了个别致的名称——蒙特卡洛法。它的实质就是在计算机上用数学方法产生大量的随机抽样的样本。特别是大容量的高速计算机的问世，大大地拓广了这一方法的应用领域。

1.Buffon的投针试验

法国数学家Buffon曾设计了一个用投针试验求园周率π的方法，可以说是蒙特卡洛方法的雏形。

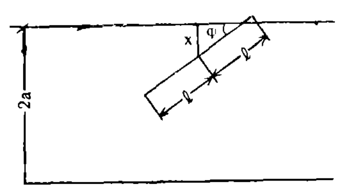

图4.Buffon投针试验

在平面上划出若干条平行线，线距为2a，取针长为2l，a>l>o。若随机地进行投针，投针的结果针线可能相交也可能不相交，判断针线相交的充分必要条件是（图4）：

$$x < l\sin\varphi$$

式中x为针的中点到最近的一条平行线的距离，φ是针线之间的锐夹角。

投针时针线相交或不相交完全是偶然的。在针投下前我们不能予料投针的结果是相交或不相交，但针投下后的结果是确定的，或是相交，或是不相交，二者必居其一且仅居其一。如果我们用上述不等式来表示这一投针试验，就是说在针投下前我们不能予知x将在〔0,a〕区间和φ将在〔0,π〕区间取什么值。但针在投下后，它们的取值就完全确定了。而且x取什么值是和φ取什么值无关的，数学上称它是两个互相独立的随机变量。由于x和φ分别在〔0,a〕和〔0,π〕区间内是均匀分布的，可以证明，针线相交这一随机事件的概率为：

$$P=\int_0^1\frac{1}{\pi}\int_0^1\sin\varphi\frac{1}{a}\,dxd\varphi=\frac{21}{a\pi}。$$

这样就把园周率π和针线相交的概率P联系起来了。因此可以用随机投针的统计试验来取得针线相交的概率的近似估计值$\hat{p}$，将其代入上式就可算得园周率π的近似估计值$\hat{\pi}=\frac{21}{a\hat{p}}$。也就是说，我们作了N次投针试验，若其中有n（n⩽N）次是针线相交的．则$\hat{p}=n/N$，$\hat{\pi}=2Nl/an$。确实有很多人实际做过这样的投针试验并求得园周率的近似估计值。

进一步地分析表明，上述投针试验还可以在计算机上用数学方法对相互独立的两个随机变量x和φ进行随机抽样进行模拟。我们可以设计一个产生〔0，1〕区间上均匀分布的随机数的专用程序（称为伪随机数发生器），并对产生的随机数作如下的变换：

$$\begin{cases}x_i=ar_{2i-1}; \\ \varphi_i=\pi r_{2i},\end{cases}\qquad i=1,\ 2,\ \cdots\cdots$$

作为随机变量x和φ的一次随机抽样。这样，由专用程序每产生两个随机数r_{2i-1}，r_{2i}，就相当于一次投针试验的两个随机变量x和φ的一次随机抽样。同时在计算机上作一次不等式检验，即判不等式

$$x_i<l\sin\phi i\text{或}r_{2i-1}<\frac{1}{a}\sin\pi r_{2i}$$

是否成立。若不等式成立，即表示是一次针线相交的投针试验，否则即为一次针线不相交的试验。因此，由专用程序产生2N个随机数并作N次不等式检验，就代表了N次投针试验。记下不等式成立的次数n，则$\hat{p}=n/N$，$\hat{\pi}=2Nl/an$。

2.模拟随机变量函数概率分布的蒙特卡洛方法

上述在计算机上用专用程序产生随机数来模拟投针试验的方法具有一般性的方法学意义。也就是说，抽去x是表示针的中点到最近的一条平行线的距离和φ是表示针线之间的锐夹角，以及$\hat{p}$表示针线相交这一随机事件出现的概率等这样一些具体的物理意义，方法同样是可行的。那就是只要相互独立的随机变量的概率分布已知，这些随机变量的任意确定型的函数的概率分布都是可以在计算机上用数学的抽样方法来得到，这就是蒙特卡洛法。在一般情况下，如果被研究的对象是一个随机变量y，而它又是m个随机变量x_1，x_2，……，x_m的某种函数，即

$$y=f(x_1,\ x_2,\ \cdots\cdots,\ x_m)．$$

当x_1，x_2，……，x_m是相互独立的，而且它们各自的概率分布又是已知的（如果各变

量不是互相独立的，则要求联合概率分布是已知的），则计算y的概率分布在数学上归结为计算一个多重积分

$$P=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\cdots\int_{-\infty}^{\infty}P\left(x_1,\ x_2,\ \cdots\cdots,\ x_m\right)dx_1dx_2\cdots\cdots dx_m.$$

这时，即使x_1，x_2，……．x_m都遵从正态分布，y的分布一般情况下也不一定是正态分布的。如果x_1,x_2，……,x_m的分布比较复杂,或是随机变量的个数比较多（m较大），或者函数关系$f(x_1,\ x_2,\ \cdots\cdots,\ x_m)$比较复杂，都会使求y的分布的解析解变得很困难，即使采用数值积分的算法,多重积分的计算工作量也是很大的。这时,用上述的在计算机上产生已知概率分布的随机数来模拟对已知分布的随机抽样的算法就显示出极大的优越性。蒙特卡洛算法的步骤可概括如下：

如果随机变量x_1，x_2，……，x_m互相独立，而且它们各自的概率分布$P_1(x_1)$，$P_2(x_2)$，……，$P_m(x_m)$已知,则可以在计算机上用产生随机数的方法，从$P_1(x_1)$中取得一个抽样$x_1^{(1)}$，从$P_2(x_2)$中取得又一个抽样$x_2^{(1)}$，……，从$P_m(x_m)$中取得第m个抽样$x_m^{(1)}$，并按确定的函数关系$y=f(x_1,\ x_2,\ \cdots\cdots,\ x_m)$算得y的一个值为$y^{(1)}=f(x_1^{(1)},\ x_2^{(1)},\ \cdots\cdots x_m^{(1)})$它显然是从分布P（y）中得到的一个抽样值。重覆上述步骤可以得到第二次抽样结果$x_1^{(2)}$，$x_2^{(2)}$，……，$x_m^{(2)}$，并算得y的第二个抽样值$y^{(2)}=f(x_1^{(2)},\ x_2^{(2)}\cdots\cdots,\ x_m^{(2)})$。如此重覆N次就可以得到y的一个容量为N的随机样本（$y^{(1)}$，$y^{(2)}\cdots\cdots$，$y^{(N)}$）。我们可以用样本的统计分布$S_N(y)$来近似地代替y的概率分布，只要样本的容量N足够大（例如N＝5000）就可以有相当的工程精度。

以估算一个油藏的石油潜量为例，我们来阐述一下蒙特卡洛方法的应用过程。

如果一个油藏被证实，那么它有多少石油资源呢？本来一个油藏有多少资源是个确定的数值，但这个确定的数值只有在该油藏的石油被采尽时才能被证实。在这以前我们只能用各种方法对它进行估算。例如我们可以采用最简单的体积法公式：$q=A\cdot H\cdot S\cdot \phi$估算。式中q表示油藏的石油资源量；A为含油面积；H为油层厚度；S为有效孔隙率；ϕ为含油饱和度。上述四个地质参数在油藏空间里的分布可能是不均匀的,而且这种不均匀分布的规律又是很难求得的。若能求得它们的平均值，代入上式可算得油藏资源量的单值估算值。此外，我们还可以通过钻井的测井分析以及取样的实验室分析数据求得这些地质参数的统计频率分布，并把它看成是对总体进行抽样取得的样本值。因此这些地质参数都可以看成是随机变量，而油藏的资源就是这些随机变量的函数（这里的函数关系就是参数的乘积关系）。这样我们就可以用蒙特卡洛方法来模拟油藏资源量的概率分布。

设上述四个地质参数的累计频率分布如图5曲线所示。则计算机产生的第一个随机数r_1代表了对含油面积作的第一次抽样，从累计概率分布曲线上得到样本值A_1。同理，由第二、三、四个随机数r_2、r_3、r_4可得油层厚度、有效孔隙率及含油饱和度的样本值为H_1、S_1和ϕ_1，从而算得油藏资源量的一个样本值$q_1=A_1\cdot H_1\cdot S_1\cdot \phi_1$，这样作五千次就可以得到油藏资源量的五千个可能取得的值q_1，q_2，……，q_{5000}，对这五千个可能取得的值进行统计就可以得到油藏资源量概率分布的渐近估算。整个蒙特卡洛计算

过程可以形象化地表示为一系列概率分布曲线相乘得到一条概率分布曲线(如图6所示)。

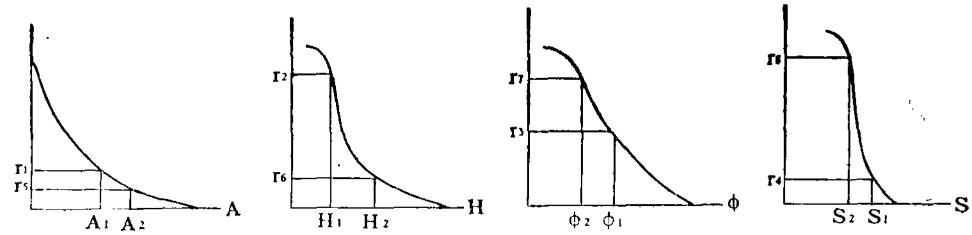

图5.含油面积、油层厚度、有效孔隙率和含油饱和度的累计频率分布

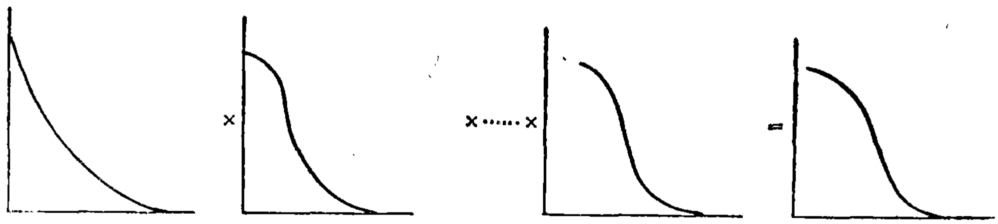

图6. 蒙 特 卡 洛 乘 法 示 意 图

如前所述，本来油藏的资源量是一个确定的数值，但由于我们不能确切地知道油藏各参数的空间不均匀分布的规律，因此用概率统计的观点和方法进行了估算。这种估算给我们提供了油藏资源量取各种可能值的把握性，为开发这个油藏制订技术和经济方案提供了比单值估算或值域估算更多的信息。

3.蒙特卡洛方法的几点说明

由上述可知,蒙特卡洛方法是一个能在广泛领域内获得应用的一般性方法,是处理随机变量函数的概率分布的一个强有力的工具。即使有些问题本身并不是随机性的问题，只要能把它转化成随机变量的函数问题，并构造出各随机变量的恰当的概率分布，都可以用蒙特卡洛方法进行模拟而获得有益的结果。

蒙特卡洛方法是对客观存在的物理实际作的计算机模拟，可以说是一种仿真技术，这一点有着非常重要的现实意义。因为许多破坏性试验是不允许大量进行的，也有许多数据是不能直接得到的（例如油藏资源量）。因此，在计算机上实现物理抽样的模拟，不仅速度快、时间省、经济性好，而且它能获得实际抽样所不可能得到的大数据量。这一点对概率统计的作用也是很重要的。

需要指出的是，一般说来蒙特卡洛模拟结果的精度是比较差的。其误差正比于随机变量y的标准差σ_y，反比于抽样次数的平方根$\sqrt{N}$。计算公式为$\Sigma = c\sigma_y/\sqrt{N}$。所以增大抽样次数N对提高模拟结果的精度效果并不显著，还会增加计算机解题的时间。为了提高模拟结果的精度，常用改进抽样的技巧（如“重要抽样”、“分层抽样”、“控制变数抽样”、“对偶变数抽样”等）来降低方差，加速收敛以提高精度。在一般精度

要求不高的场合，适当增加抽样次数就可以满足一般的工程精度要求。

产生随机数的算法在蒙特卡洛方法中占有重要的地位，所以改进和建立更好的产生随机数的算法始终是改进蒙特卡洛方法的核心问题之一。随机数的物理意义是随机样本的样本值。随着问题的不同，随机变量的取值范围以及概率分布规律也是各式各样的。要想在计算机上建立起适应于如此广泛应用背景的产生随机数的算法显然是不可能的。理论上曾经证明，只要有了一种连续分布的随机数，就可以通过适当的变换而得到任何所要求的概率分布的随机数。由于〔0,1〕区间上均匀分布的随机数是最基本、最简单和最容易得到的随机数，所以一般计算机上均配有这样的随机数发生器。只要针对不同的问题，设计好变换即可。这样就把改进和建立产生随机数的算法集中为〔0,1〕区间均匀分布的随机数的产生上。这里要说明的是，在计算机上用数学方法产生的随机数并不是真正的随机数，因为它在产生之前就被准确地确定了。而且由数学递推公式计算出的随机数序列有一定的周期性，达到一定长度以后会循环出现。所以称之为伪随机数。但是这个弱点可以通过改进算法和进行统计检验加以弥补，使之不产生过大的系统误差。另一方面，由于它速度快、占用内存小、经济性好、可以进行复算检查等优点，所以虽有多种产生随机数的方法，而真正获得广泛应用的还是用数学方法产生的伪随机数。目前最常用的是“同余法”（“加同余法”、“乘同余法”、“混合同余法”等），并根据半经验的办法适当选取参数使之能产生周期长、随机性好的伪随机数序列[6][7]。

在用蒙特卡洛方法进行随机抽样时，我们曾形象化地表示，根据计算机给出的随机数在随机变量的累计频率分布曲线图的纵坐标轴上取点，然后找出曲线所对应的随机数（图5）。这里决不要误解为由概率值去找随机样本的样本值。因为一次随机抽样得到的样本值是个确定值，而概率所对应的仍然是一个区间内的不确定值。这一“作法”的实际含意仍然是由〔0,1〕区间上均匀分布的随机数向给定分布的变换以及对给定分布的随机抽样，这一点是不应有误解的。

HASP方法的抽样模式

基于蒙特卡洛算法，HASP方法在估算全盆地的油气资源量时，较多地考虑了地质因素的影响和对实际勘探过程的模拟。

HASP方法主要是用来对未发现的潜在油气资源进行评价的。资源潜量的估算是以勘探层为单元按圈闭进行累加。资料的收集和地质分析都紧紧围绕着勘探目标数的概率分布、地质风险分析和油气藏规模的条件概率分布这三个基本要素展开。提出的抽样模式模拟了实际的勘探过程。

1.勘探层三要素分析

（1）勘探目标数分布

众所周知，圈闭是形成油藏的必要条件。一般说来，找油要先找圈闭。所以，一个盆地在进行油气勘探时总是要先用地震等物探方法查清构造，圈定一批圈闭。这些圈闭既是我们勘探找油的目标，也是我们进行资源量估算的入口。为此必须根据客观的资料加上主观的判断和估计，作出该勘探层勘探目标数的主观概率分布。这里要求地质学家和

地球物理学家密切配合作出适当的估计。例如某勘探层根据地震资料和构造图已经圈出15个圈闭，则这个勘探层能发现15个以上勘探目标的概率即为1。如果构造图上还有三个圈闭虽未确定，但是可以判断是由于某些原因造成资料上的误差所致，它们是圈闭的可能性很大。例如经地质学家和物探学家共同分析，给出该勘探层能发现18个以上勘探目标的概率估计为0.95。此外，考虑到地震测线的方向和圈闭的形状，若适当选择地震测线的走向，也许还能多发现7个圈闭，估计发现25个以上勘探目标的概率为0.75。由于地震测线较稀，因此适当加密地震测线，有些较小的圈闭也将被发现。估计发现30个以上勘探目标的概率为0.25。各种因素都考虑进去，该勘探层的勘探目标最多不会超过35个，即发现35个以上勘探目标的概率为零。根据这些估计我们可以作出该勘探层勘探目标数概率分布的主观估计曲线（如图7）。

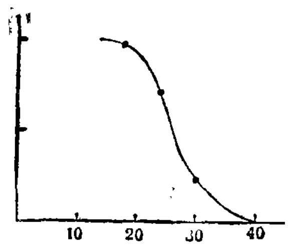

图7.勘探目标数概率分布

（2）地质风险分析

圈闭仅仅是形成油藏的必要条件，而不是充分条件。一个勘探层的许多勘探目标，有的有油，有的没有油。在没有逐个地进行钻探证实以前，我们是不能充分肯定哪些是有油的，哪些是没有油的。只能根据多方面的资料和研究进行判断。这种判断主要来自对地质因素的分析，这也就使勘探工作带有风险性，故称之为地质风险分析，并表现为勘探的成功率。HASP方法强调资源是否存在与资源量的多少是两个根本不同但又相互关联的问题，应分别进行考虑。例如前者考虑的是有否油源、是否有适当的储层和盖层、圈闭形成的时机是否合适、后期构造的破坏等等，它以边缘概率的意义出现。而后者是在确定油气存在的条件下考虑圈闭的大小、进入圈闭的油气数量问题，它以条件概率的意义出现。一个圈闭被判断为有油或没有油，而钻探的证实却相反，其原因可能是非常复杂的。一方面由于逐个地进行圈闭含油的判断是非常困难的，另一方面在作资源潜量估算时还包含了许多推断的圈闭是尚未被发现的。所以HASP方法假设所有的勘探目标都有相同的地质风险因素。油气勘探的实践表明，油藏在盆地中的分布是不均匀的，具有成“群”出现的特点。油藏的许多重要参数（如石油的化学成分、油藏的类型等等）差别是很大的。在“盆地”这样大型的地质单元内是很难找出圈闭含油的地质风险因素的共性。所以HASP方法采用了“勘探层”（Play）作为进行资源评价的地质单元。一个勘探层尽管有很多勘探目标，有的已经发现，有的尚未发现，但是它们的生油层及成熟度、圈闭的成因、构造的历史等是相同的或近似的。这样，两个勘探层可以完全或部分重迭，而虽属同一地层但在盆地的不同部位，也可以是属于不同的勘探层。所以勘探层主要反映了油藏形成的主要地质因素和烃类特征，其地理的和地层的范围也都是限定的，为进行地质风险分析提供了基础。如果我们对某个勘探层选定了风险分析的地质因素为$u_1 u_2$，……，u_k，而且这些因素又都是相互独立的，那么该勘探层的“风险系数”即为$R = P_1(u_1) \cdot P_2(u_2) \cdot \cdots\cdots \cdot P_k(u_k)$。式中$P_i(u_i)$（$i = 1, 2, \cdots\cdots, K$）

为第i个因素对油藏形成所必需的“大于最小值的概率”。R是i个因素同时具备的概率，反映的是该勘探层的勘探成功率。需要指出的是，由于风险系数是各个地质风险因素边缘概率的乘积，因此在资料不充分时，风险因素选得越多，其值越小。所以不宜为了求全而在缺乏充分资料的情况下选用过多的因素。此外，对于不同的勘探层可以选用不同的风险因素，要根据资料有利于进行风险分析而定。地质风险分析是HASP方法的第二个基本要素，也是一个很重要的环节。有些资源量估算方法不考虑风险，只考虑估算的精度，似乎油气的存存是确定无疑的，这当然是不充分的。

地质风险分析目前还没有一个成熟的程式。主要还是根据勘探历史，具体地分析一系列油藏和“空”圈闭形成的地质因素的控制作用。在未经勘探的新区只能靠类比或是建立便于判断的概念性的地质风险模式给出主观概率值。表Ⅰ为某勘探层的示意的地质风险因素及边缘概率值。经分析，该勘探层油源岩成熟，但圈闭分散，在已钻探的圈闭中有一半是由于离油源区太远，横向运移受阻不能形成油藏，所以估计油源条件这个因素的边缘概率为0.5。该区储层发育，仅很少的几个圈闭因孔隙度过低而形不成工业油藏，估计边缘概率为0.8。此外，区域盖层发育，岩性致密，有足够的厚度，已证实的空圈闭中没有一个是因为盖层未能顶封所致，可估计边缘率为1。经地化和构造分析，有近一半的空圈闭是由形成时期晚于高峰排烃期，估计边缘概率为0.5。后期无强烈构造运动，圈闭的保存状况良好，估计边缘概率为1。这样就可算得风险系数$R=0.5\times0.8\times1\times0.5\times1=0.20$，它是勘探层油气存在的概率。这一数据基本上是来源于对该勘探层勘探历史的成功率所作的控制油藏的地质因素的分析。如果是根据理论或经验给出估计，则应和勘探的成功率加以对照，若相差过大，那就应该对估计进行适当的修正。

表1. 油气存在的主观概率分析

地质因素	概率
油源条件	0.5
储层条件	0.8
盖层条件	1.0
圈闭对机	0.5
保存条件	1.0

油气存在的边缘概率$=0.5\times0.8\times1\times0.5\times1=0.20$

（3）油藏规模的条件概率分布和综合概率分布

HASP方法估算油藏资源量的公式为：

油藏潜在量＝圈闭面积×储层厚度×孔隙率×（1－水饱和度）×圈闭充填率×油所占的比例×采收率×常数。

基于条件概率的含意，这里是以勘探层所有已发现的油藏为子样进行统计得出的。也就是说，每个已发现的油藏的上述各个参数均以其均值参加统计得出各参数的条件概率分布，再用蒙特卡洛乘法得出勘探层油藏规模的条件概率分布（图8曲线1）。曲线1表明，如果勘探层发现一个油藏，其资源量必为曲线上某个点所对应的值，或者说当某个勘探

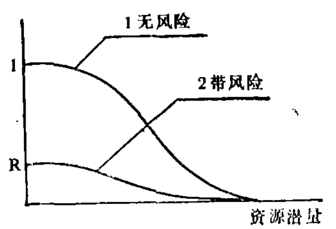

图8.油藏规模的条件概率和综合概率分布

目标存在油气时，其资源潜量估值的把握性如曲线1所示。如果一个勘探目标没有油气存在，也就无所谓资源量的问题。所以曲线1称为条件概率分布曲线。由于HASP方法是以勘探层为单元按圈闭的资源量累加，所以对勘探目标数分布进行随机抽样是资源量估算的入口。勘探目标数本身就是在概率意义下估计的，其中包含着许多尚未发现的勘探目标。这些勘探目标的资源潜量也就只能在估计的油气存在的概率下进行分析。HASP方法提供的油气存在的概率是单值估计，所以图8的曲线1的纵坐标值乘以油气存在的概率值R得出的曲线2就是带风险的潜量分布，称之为综合概率分布。

以上是用HASP方法进行资源潜量估算的三个基本要素，它们和抽样模式密切相关，也告诉我们按怎样的概念和方法去收集和整理资料。所以它是HASP方法应用的准备阶段——资料予处理阶段。

上述三要素都是以概率分布形式表达的，其中充满对客观基础资料的主观判断和估计。客观的基础资料是否充分，是否准确，主观判断和估计的优劣，均直接影响估算结果的可靠性。一般情况下如果不能确定解析概型时，要慎重估计概率为0.95和0.05两点以控制分布曲线的两端以及概率为0.75和0.25两点以控制分布曲线的中段。

2.估算盆地资源潜量的抽样模式

如前所述，HASP方法在估算资源量时是以勘探层为单元逐个圈闭累加的。因此在估算全盆地的油气资源总潜量时，首先要对盆地作勘探层的划分与鉴定。然后按勘探层分别收集和整理“三要素”的资料数据。HASP方法的抽样结构如图9所示，可以看成是一个三级随机抽样的模式：

第一级随机抽样是用蒙特卡洛乘法根据各有关地质参数的条件概率分布求勘探层油藏规模的条件概率分布，并由油气存在的概率得到综合概率分布。

第二级随机抽样是根据随机抽得的勘探目标数，逐个地判断它们油气存在的地质风险，对判断有油的勘探目标在综合概率分布中随机抽样取得潜量值并累加得到勘探层的一个可能的潜量估算值。如此重复两千次则可得到两千个可能的潜量估算值，统计得出这两千个可能值的累计频率分布即为该勘探层的油气资源潜量的累计概率分布估计。

第三级抽样是各勘探层的潜量分布用蒙特卡洛加法求全盆地（或区域）的资源总潜量。

几个问题的讨论

1.关于估算结果的解释

HASP方法用累计概率分布的形式来表达对一个勘探层或盆地（或区域）的油气资源潜量所作的估算。这里“潜量”指的是推断可能发现而目前尚未发现的油气资源，“估算”表明了工作人员根据客观资料、地质理论和经验按某种数学模型（或算法）获得定量予测结果的一种看法。也就是说，这个估算结果决不是“储量”意义下的资源，即使是概率为1情况下的值也不是可以到手的储量。所以在解释这个估算结果必须严格把“储量”（按工业储量方法计算的）和“潜量”从概念上区分开来。为了表明一个勘探层（或盆地、区域）的油气资源情况，可以将已发现的储量分为累计产量和剩余储量

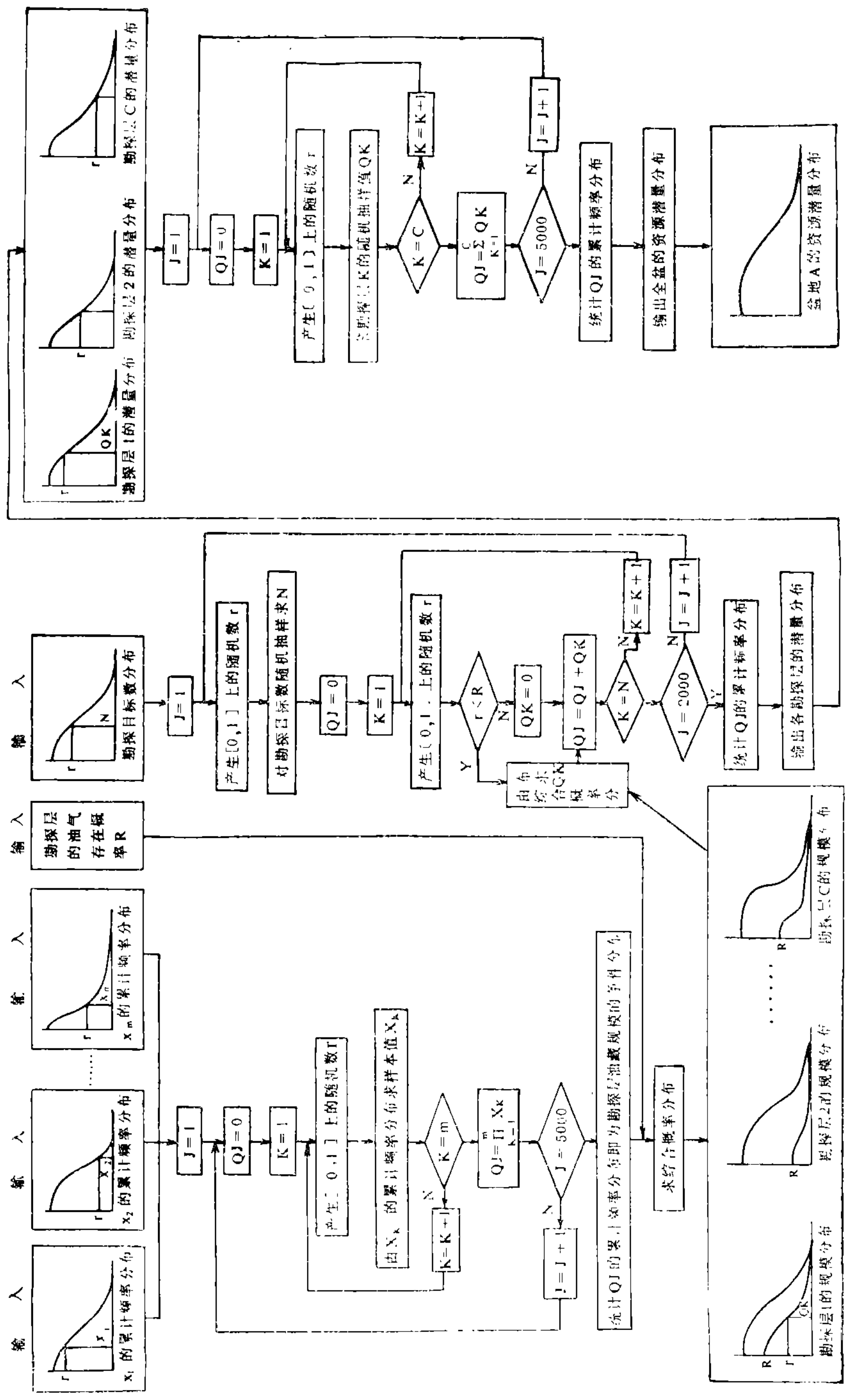

图9.　估算全盆地油气资源潜量的HASP方法结构示意图（引自地矿部访加报告，作者略有改动）

两部分并接在潜量图的左侧加以示意（图10）。

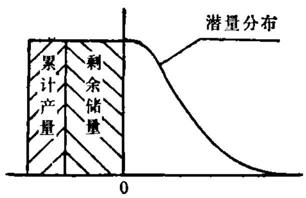

图10．油气资源表示示意图

基于油藏规模形成的随机模式，在确定的地质条件下所能形成的油气资源量其数值是某个连续区间内的无穷多个值。潜量分布曲线表明了对这无穷多个值出现范围的概率估计。显然，估计的范围越宽其可靠性就越小。此外，随着勘探和研究程度的提高，不确定因素逐渐排除（例如某些油藏或空圈闭被证实等），整个分布曲线的值域也在减小，曲线也由较平缓转变为较陡峭。所以，曲线的形态和值域既反映了估算时对勘探层（或盆地、区域）的认识程度，也表明了估算结果的把握程度。使估算保持一种动态逼近的态势。

由于油气资源的勘探既有风险又可能获得较大的盈利，HASP方法的估算结果允许不同的决策人员根据不同的需要使用不同的估算值，为决策提供了更多的信息。

2.关于原始资料数据的统计问题

如前所述，在统计作出油藏规模的条件概率分布时，要求把所有已发现的油藏按单值（平均值）统计各有关参数的条件概率分布。这里油藏规模实际上是由各地质参数分布综合得到的一个无穷总体，一个盆地(或勘探层)的实际油藏的规模分布只是无穷总体的一个随机样本，其中每一个油藏都是一个子样。这一点在估算勘探层的资源潜量分布的第二级随机抽样中看得很清楚。因此那种不管是不是油藏，凡是钻井、测井或实验分析得出的数据，甚至在一个构造上取得的多个数据均作为子样参与统计的作法是不对的。

其次，在勘探目标数分布的估计中，概率为1的勘探目标数不包括已发现的油藏数，所以它们也不参加第二级随机抽样。这是很显然的，因为已发现的油藏数参加统计和抽样，则把一个确定了的事件重新变为不确定的事件是不合理的。但是在最后统计勘探层的潜量累计概率分布时，可以把已发现的油藏的资源量以单值(平均值)加在每次抽样值的上面，或是表示成累计概率分布的形式与潜量分布估算再作蒙特卡洛加法而得出勘探层的资源量累计概率分布。这时曲线如图11所示，其概率为1时的资源量值即为剩余储量。

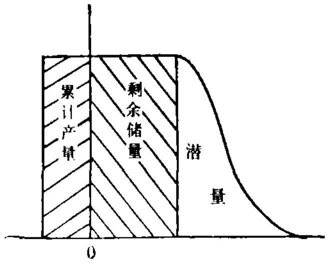

图11．资源量的表示示意图

有些文章把蒙特卡洛乘法和加法简称为“概率乘”和概率加”，没有充分概括出蒙特卡洛算法的特点。因此，把同概率值下的样本值相乘（或相加）的做法是不对的，显然它只是无穷多个值中的一个。

HASP方法的内容是很丰富的，它还包括为作经济分析用的石油勘探的发现模式和马尔柯夫链及抽样不放回总体的机理等内容。限于篇幅，本文只分析了资源量估算的部分，以供参与和关心资源评价工作的同志参考、讨论。

（收稿日期：1985年2月7日）

参 考 文 献

1) Roy K.J., 1979, Hydrocarbon Assessment Using Subjective Probability and Monte Carlo Methods, in M.Grenon, ed.First II ASA Conference on Methods and Models for Assessing Energy Resources: Pergamon Press, New York, P 279-290.

2) Lee P.J. and Wang P.C.C., 1983, Probabilistic Formulation of a Method for The EValuation of Petroleum Resorces, Mathematical Geology, Vol.15, No.1, P 163—181.

3) Lee P.J.and Wang P.C.C., 1983, Conditional Analysis for Petroleum Resources Evaluations, Mathematical Geology, Vol.15, No.2, P 353—365.

4) Lee P.J.and Wang P.C.C., 1983, Prediction of Oil or Gas Pool Sizes When Discovery Method is Available: Manuscript in Preparation.

5) Lee P.J.Wang P.C.C., 1983, Estimation of Superpopulation Parameters in the Presence of Length—Biased Sampling from a Finite Population; Manuscript in Preparation.

6)李惕碚，实验的数学处理，科学出版社，1980年。

7)中国科学院计算中心概率统计组，概率统计计算，科学出版社，1979年。

AN APPLICATION OF HASP IN ASSESSMENT OF OIL AND GAS RESOURCES

Chen Xuefu

(Central Laboratory of Petroleum Geology, Ministry of Geology and Mineral Resources.)

Abstract

This article uses HASP method to analyse the structure of a random sampling model. It introduces the principles of Monte Carlo Method and discusses the faults and mistakes in the present—day applications.