文章编号:1673-5005(2010)01-0070-03

基于模糊概率理论的地层流体精确识别方法

翟应虎[1], 李祖光[1], 张 旭[1], 李 亚[2], 马冬梅[3], 王文广[4]

(1. 中国石油大学 石油工程教育部重点实验室,北京 102249; 2. 大庆钻探工程公司测井二公司,吉林 松原 138003; 3. 中石油吉林油田公司 采油工艺研究院,吉林 松原 138003; 4. 中国石油天然气勘探开发公司,北京 100034)

摘要:充分考虑火山岩地层流体的交互性及模糊性特征,从模糊概率理论出发,应用测井解释基本原理,对选取的火山岩样本层位测井数据进行综合统计;借助离差平方和的概念,构建地层流体的模糊概率识别功能函数;应用模糊数学及统计理论方法,确定模糊概率的概率密度函数和隶属函数;在此基础上建立基于模糊概率理论的松辽盆地火山岩地层流体识别数学模型。实例计算的识别准确率在 80% 以上,表明该方法切实可行,具有推广应用的价值。

关键词:模糊概率; 流体识别; 松辽盆地; 火山岩

中图分类号:TE 142 **文献标志码**:A

Exact identification approach to volcanics formation fluid based on fuzzy probability theory

ZHAI Ying-hu[1], LI Zu-guang[1], ZHANG Xu[1], LI Ya[2], MA Dong-mei[3], WANG Wen-guang[4]

(1. *MOE Key Laboratory of Petroleum Engineering in China University of Petroleum*, *Beijing* 102249, *China*;
2. *2nd Logging Company*, *Daqing Exploratory Drilling Company*, *Songyuan* 138003, *China*;
3. *Oil Production Technology Research Institute*, *Jilin Oilfield Company*, *PetroChina*, *Songyuan* 138003, *China*;
4. *Oil & Gas Exploration and Development Corporation*, *PetroChina*, *Beijing* 100034, *China*)

Abstract: The interactivity and ambiguity of volcanic fluid were adequately considered. Based on fuzzy probability theory and log interpretation axiom, sampled data of volcanics were selected and cluster analyzed. A fuzzy probability function for fluid identification was established by using the conception of sum of squares of deviations. The probability density function and subordinate functions were established by using the theory of fuzzy mathematics and probability. And a mathematical model used for formation fluid identification in Songliao Basin by fuzzy probability was derived. The precision ratio of example identification is above 80%. The calculated results show that the approach is feasible and has popularization value.

Key words: fuzzy probability; fluid identification; Songliao Basin; volcanics

松辽盆地南部深层天然气储量丰富[1-2],正确识别该地区的地层流体不仅为储层含气性评价、井网部署及测井解释图版的制定奠定基础,也为钻井过程中钻井工具及钻井作业方式的选择提供重要依据。目前识别地层流体的主要依据是地球物理测井资料或地震资料,识别方法主要包括交汇图法、人工神经网络法、聚类分析法及模糊灰关联分析法等[3-8]。这些方法在一定条件下可以取得较好的效果,但松南地区深层泉一段下部、登娄库组及营城组地层广泛发育着不同类型的火山岩体系[2],实际计算结果表明,应用以上方法对火山岩地层流体进行识别的计算结果与实际情况符合率较低,且存在着相互矛盾的现象。笔者从模糊概率理论的基本原理出发,构建火山岩地层流体识别的数学模型。

收稿日期:2009 - 05 - 31

基金项目:教育部"长江学者和创新团队发展计划"项目(RT0411);吉林石油集团有限责任公司重点科技攻关项目(jw - 07 - w44 - jz - 1 - 17)

作者简介:翟应虎(1950 -),男(汉族),江苏泰州人,教授,博士,博士生导师,从事岩石破碎与破岩工具的教学及科研工作。

1　基本原理

1.1　模糊概率理论

模糊概率是模糊数学与概率论相结合的产物。如果事件是模糊的,而概率值是普通数值,则称为模糊事件的概率[9]。令 A 表示模糊事件的模糊集,则它发生的概率记为 $P(A)$。

设 A 为随机变量 $\boldsymbol{X}=\{\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n\}$ 上的模糊集,则

$A=\mu_A(x_1)/x_1+\mu_A(x_2)/x_2+\cdots+\mu_A(x_n)/x_n$,

已知 x_i 发生的概率为 P_i,则

$$P(A)=\sum_{i=1}^{n}\mu_A(x)p(x_i)P_i. \tag{1}$$

在论域为连续的情况类似有

$$P(A)=\int_x\mu_A(x)p(x)\mathrm{d}x. \tag{2}$$

式中,$p(x)$ 为概率密度函数;$\mu_A(x)$ 为模糊事件 A 的隶属函数。

式(2) 是具有普遍意义的模糊可靠性分析通用公式。

1.2　利用模糊概率理论识别地层流体的基本思路

设 $A_j(j=1,2,\cdots,n)$ 为论域 U 中的 n 个模糊子集,分别对应于不同类型地层流体的层位构成的集合。用来识别地层流体的测井参数有 k 个,即 $\boldsymbol{X}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_k)$,其中第 i 个变量 $\boldsymbol{X}_i=(x_{1i},x_{2i},\cdots,x_{qi})^{\mathrm{T}}$ 分别对应于用来识别地层流体的声波时差、电阻率等测井参数,$i=1,2,\cdots,k$,q 表示选取测点的个数。

对于任一测点 $\boldsymbol{X}'_i=(x_{i1},x_{i2},\cdots,x_{ik})$,$(i=1,2,\cdots,q)$,该点可能属于模糊子集 A_j 为一模糊事件,其概率记为 $P(A_j)$。显然,$P(A_j)$ 的值反映了该测点属于模糊子集 A_j 的可能性,进而反映了该测点对应于第 j 种地层流体的可能性。

2　计算步骤

2.1　样本数据的选取

标准样本选取的好坏,将直接影响地层流体的最终识别结果,因此标准样本必须严格选择,原始数据要准确,所选取的标准层位要具有连续性及代表性。

在综合分析测井解释原理[10-11] 的基础上,选取自然伽马、声波时差、深侧向电阻率及双侧向电阻率比值 4 个测井参数作为综合分析评价参数。设 A_j 表示由测井参数 $\boldsymbol{X}_i$ 标识的第 j 类层位的模糊子集($j=1,2,\cdots,n$),根据样本选取原则,选取一段由第 j 类层位构成的样本数据为

$$\boldsymbol{X}^j=(\boldsymbol{X}_1^j,\boldsymbol{X}_2^j,\cdots,\boldsymbol{X}_k^j)=(x_{i1}^j,x_{i2}^j,\cdots,x_{ik}^j), \tag{3}$$

取各列平均值,得

$$\bar{\boldsymbol{X}}^j=(\bar{\boldsymbol{X}}_1^j,\bar{\boldsymbol{X}}_2^j,\cdots,\bar{\boldsymbol{X}}_k^j). \tag{4}$$

由于火山岩地层的复杂性,需选取多段测井曲线作为样本数据,这样式(4) 可写成

$$\bar{\boldsymbol{X}}_l^j=(\bar{x}_{l1}^j,\bar{x}_{l2}^j,\cdots,\bar{x}_{lk}^j). \tag{5}$$

式中,$i=1,2,\cdots,q$;$l=1,2,\cdots,p$,p 表示所选取测井曲线的段数。

通过对该地区已完钻井的测井资料及录井资料统计分析,包含同一类别地层流体的层位其各项测井参数基本服从正态分布,即对于任一 $\boldsymbol{X}_i$,有 $\boldsymbol{X}_i\sim N(\mu_i,\sigma_i^2)$。由无偏估计理论[12] 可知,任一组测井样本数据 $\boldsymbol{X}_i$ 的样本均值 $\bar{X}_i$ 和样本方差 S_i 分别为 μ_i 和 σ_i^2 的无偏估计。将其化为标准正态分布即 $\boldsymbol{Y}_i\sim N(0,1)$,则有

$$\boldsymbol{Y}_i=\frac{\boldsymbol{X}_i-\mu_i}{\sigma_i}\approx\frac{\boldsymbol{X}_i-\bar{X}_i}{\sqrt{S_i}}. \tag{6}$$

由于火山岩地层结构复杂,且火山岩种类繁多,即使包含同一地层流体的层位其测井参数也可能比较分散,即同一模糊子集可能存在着多个不同的统计中心。这时为了提高识别的准确性,必须对样本数据进行聚类分析处理。

2.2　计算模型的构建

为综合分析多个随机变量与不同类型的地层流体之间的关系,首先需要借助离差平方和的概念[13] 构建一个包含 k 个随机变量的功能函数,

$$p(\boldsymbol{X})=\sum_{i=1}^{k}\left(\frac{\boldsymbol{X}_i-V_i}{U_i}\right)^2. \tag{7}$$

式中,$U_i>0$,$V_i>0$,且为待定常数。

结合式(6),令 $U_i=\sqrt{S_i}$,$V_i=\bar{X}_i$,将其带入式(7),有

$$p(\boldsymbol{X})=\sum_{i=1}^{k}\left(\frac{\boldsymbol{X}_i-\bar{X}_i}{\sqrt{S_i}}\right)^2=\sum_{i=1}^{k}\boldsymbol{Y}_i^2. \tag{8}$$

由式(8) 可以看出,功能函数 $p(\boldsymbol{X})$ 的值反映了 $\boldsymbol{X}_i$ 与模糊子集 A_j 中心的距离,即 $p(\boldsymbol{X})$ 值越小,其属于 A_j 的可能性就越大。但是,即使当 $p(\boldsymbol{X})=0$ 时,该点仍有可能不属于集合 A_j,这准确地描述了火山岩地层流体的测井响应具有交互性及模糊性的特征。

由多维随机变量函数的分布原理[12] 可知,如果随机变量 $\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_k$ 相互独立,且均服从于 $N(0,1)$,则 $\boldsymbol{Y}$ 的分布称为具有自由度 k 的 χ^2 分布,记作 $\boldsymbol{Y}$

$\sim \chi^2(k)$。χ^2 分布的概率密度为

$$p(y) = \begin{cases} 0, y \leqslant 0; \\ \dfrac{1}{2^{k/2}\Gamma\left(\dfrac{k}{2}\right)} y^{(k/2)-1} \exp\left(-\dfrac{y}{2}\right), \ y > 0. \end{cases} \quad (9)$$

式(9) 即可作为功能函数 $p(\boldsymbol{X})$ 的概率密度函数。

模糊集合是用隶属函数描述的,对模糊对象只有给出切合实际的隶属函数才能应用模糊方法进行计算。目前隶属函数在理论上没有一个普遍适用的方法,往往依赖于人的主观判定或技巧,主要包括示范法、统计法、蕴合解析法、可变模型法、相对选择法、子集比较法、滤波函数法等[14]。笔者综合考虑各项测井参数的正态分布特征,选取的隶属函数形式为

$$\mu_A(y) = \begin{cases} 1, \ 0 < y \leqslant c; \\ \exp\left[-\dfrac{(y-c)^2}{b}\right], \ y > c. \end{cases} \quad (10)$$

式中,b,c 为待定常数,取 $b = 2$,$c = 3$。

式(10) 即为将 $\boldsymbol{X}_i$ 化为标准正态分布 $\boldsymbol{Y}_i$ 后的隶属函数形式。

2.3 概率值计算及流体识别

将式(9),(10) 代入式(2),即可得到模糊事件 A 的概率计算公式。

把已知的测井参数 $\boldsymbol{X}_i$ 按式(6) 化为标准正态分布 $\boldsymbol{Y}_i$,再结合式(7) ~ (10),则任意给定测点 $\boldsymbol{X}'_i$ 属于模糊子集 A_j 的模糊概率计算公式为

$$P(A_j) = \begin{cases} \displaystyle\int_a^3 p(y)\mathrm{d}y + \int_3^\infty \mu_A(y)p(y)\mathrm{d}y, \\ \qquad 0 < a \leqslant 3; \\ \displaystyle\int_a^\infty \mu_A(y)p(y)\mathrm{d}y, \ a > 3. \end{cases} \quad (11)$$

其中

$$a = \sum_{i=1}^{k} [(\boldsymbol{Y}_i - \mu_i^j)/\sigma_i^j]^2.$$

将相关测井数据按照式(11) 进行计算,可得到 $P(A_1)$,$P(A_2)$,…,$P(A_n)$,依据最大模糊概率原则,即若 $P(A_j)|_{x'_i} = \max(P(A_1),P(A_2),\cdots,P(A_n))$,那么就有 $\boldsymbol{X}'_i \in A_j$。

3 计算实例

对松南地区已完钻的 6 口深井泉一段至营城组的不同层位的相关测井参数进行统计分析,并使用 SSPS 统计软件对所选取的具有代表性层位的测井参数进行聚类分析,结果见表 1。

表 1 各层位模糊集合中心的聚类统计

Table 1 Cluster result of collective center of different zone

层位	统计中心	统计项目	声波时差 Δt/($\mu s \cdot m^{-1}$)	自然伽马 GR/API	深侧向电阻率 R/($\Omega \cdot m$)	双侧向电阻率比值
气层	1	样本均值	150.0234	167.9573	62.0557	1.8931
		均方差	10.0045	6.4245	1.2453	0.5343
	2	样本均值	80.0241	180.2983	58.4521	3.7501
		均方差	7.2434	2.4523	1.4533	0.5243
差气层	1	样本均值	206.6412	113.3653	149.5954	0.9682
		均方差	12.1202	5.1241	3.2554	0.0278
	2	样本均值	177.1234	119.7887	120.7313	0.7646
		均方差	8.1783	11.4034	7.5045	0.0445
水层	1	样本均值	66.3744	100.772	15.0552	0.5720
		均方差	2.3453	2.4534	1.2453	0.0202
干层	1	样本均值	62.5434	115.4866	155.4530	1.0407
		均方差	2.2453	11.4523	3.4532	0.0181
	2	样本均值	58.4522	87.5798	65.2045	1.1861
		均方差	2.4554	5.4524	2.5435	0.0345
	3	样本均值	193.5334	99.2156	40.5434	1.0576
		均方差	8.4524	2.5423	2.4510	0.0334

对松南地区几口已完钻井的部分井段进行了地层流体识别,表 2 为预测结果与录井显示结果的对比。可以看出,基于模糊概率理论的地层流体识别结果准确率均超过 80%。表明该方法有效可行,具有推广价值。

表 2 部分井段地层流体识别结果

Table 2 Result of formation fluid identification of several sections

井号	井段 h/km	实际层位	测点个数	正确识别点数	准确率 f/%
CS1-1	3.4201 ~ 3.4273	差气层	112	87	77.68
CS1-1	3.5292 ~ 3.5350	干层	47	39	82.97
CS1-1	3.6680 ~ 3.7014	气层	204	186	91.17
CS1-1	3.7742 ~ 3.7990	气层	189	169	89.41
CS103	3.4610 ~ 3.4704	差气层	147	111	75.51
CS103	3.5480 ~ 3.5553	干层	109	91	83.48
CS103	3.6243 ~ 3.6640	气层	387	353	91.21
CS103	3.8090 ~ 3.8323	水层	183	157	85.79

4 结 论

(1) 模糊概率理论兼具模糊性及随机性的特征,有其自身的优越性,可用于模糊事件的概率计算。

(2) 同常规方法相比,模糊概率识别方法充分考虑了火山地层流体交互性及模糊性的特征,且对不同测井参数加以综合分析,为火山岩地层流体的精确识别开辟了新思路。

(3) 实例结果表明该方法切实可行,具有推广和应用价值。

(下转第 92 页)