

灰色聚类分析在孤东地区油气勘探中的应用

谭河清^{1,2}, 彭存仓², 武国华², 雷利婧², 刘世斌²

(1. 中国科学院兰州地质研究所, 甘肃兰州 730000;

2. 中石化胜利油田有限公司孤东地质研究所, 山东东营 257000)

摘要: 利用灰色系统理论的灰色聚类方法, 详细论述了地质区带的灰色特性, 对灰色聚类分析所需的各项参数进行了深入分析和研究。以济阳拗陷沾化凹陷孤东油田所属范围为例, 对其中的 17 个区带进行了灰色聚类分析, 筛选出了部分有利勘探区带, 经勘探实践证实了该方法具有较高的实用性。

关键词: 灰色聚类; 孤东地区; 区带评价; 勘探应用

第一作者简介: 谭河清, 男, 38 岁, 高级工程师(博士), 石油地质学

中图分类号: TE122.13

文献标识码: A

客观世界是相互联系、相互作用的, 由此组成大大小小的系统。系统就是由若干个相互作用、相互依赖的组成部分构成的一个具有特定功能的整体。任何一个系统, 总包含着若干个子系统, 同时又被别的系统所包围。区带可看成是一个系统, 它包括油源层、储集层、封闭层以及运移通道等多个子系统, 其更高一级的系统为含油气盆地。子系统的功能和子系统间的关系决定着系统的总体性能。

人们常用“黑箱”来形容内部信息缺乏的对象和系统, 用“黑”表示信息缺乏, “白”表示信息完全, 介于二者之间的称为“灰”。相应地, 信息完全不明确的系统称为黑色系统, 信息完全明确的系统为白色系统, 信息不完全的系统为灰色系统。白色、黑色、灰色只是一种相对的概念, 世界上没有绝对的白色系统, 也没有绝对的黑色系统, 而灰色是绝对的, 我们所见到的系统几乎都是灰色系统。所谓白色系统只是对一些不确知的因素忽略不计而已。

地质区带具有明显的灰色特性, 一方面, 通过现有的地质、物探等手段, 能获得它的部分信息, 如有机质类型、沉积相、砂岩百分含量等; 另一方面, 我们对其他一些信息还不明了, 如断层的封堵性、油气运移的过程、形成油气聚集的各因素之间是否匹配等关系还难以完全确定。因此, 区带是一个灰色系统, 区带评价也应按照灰色系统理论方法进行^[1~3]。

1 基本原理与方法

在生产和研究中, 人们往往要把所处理的事物按其性质分为若干类, 聚类分析就是按一定要求和规律对事物进行分类。在区带评价中, 可将一批区带按含油气性进行分类。评价一个区带的含油气性, 可以从油源条件、储集条件、封盖条件、油气运移条件等方面进行分析, 各方面的分析结果很可能不一致, 尽管在每个方面可以得到一个较明确的意见, 但总体上的意见是不确定的。这就需要将各方面的意见进行整理、归纳、综合, 给出一个客观的总体评价。灰色聚类就是将聚类对象对于不同的聚类指标所拥有的白化函数, 归纳成若干个灰色系统类型, 以判断该聚类对象属于哪一类灰色类型。它能给出一个量的评价, 比定性的地质分析更具有客观性。

在灰色系统理论中, 灰数、灰元、灰关系是灰色现象的特征。灰数是信息不完全的数, 灰数常用符号 $\odot$ 来表示。如某生油层有机碳含量在 0.2%~0.8%, 可记为 $\odot(C) \in [0.2\%, 0.8\%]$ 。灰元是指信息不完全的元素, 如有机碳含量, 作为计算生油量的一个重要指标, 它并不能代表有多少有机碳转化成了石油, 它是一个灰元。灰关系是指信息不完全的关系, 如断层与油气运移的关系, 什么时候对运移起通道作用, 什么时候对运移起封堵作用, 都是不完全明确的, 它们之间是灰关系。

灰数不是一个数, 而是一个数集、一个数的区

间。对于一个灰数,我们只知道其大概的范围,而不知道其确切值,灰数的确切值为白化值,灰数通过信息补充,可以转化为白数,区间蜕化为一个点。在灰色系统理论中,为了量化处理的需要,常常要将灰数白化。灰数白化的过程要用到白化权函数,白化权函数就是灰数取值(即白化)的偏爱函数。对属于某个区间的灰数 $\odot$,在该区间内取值时,如果每一个数的取数机会都是均等的,则称为纯灰数或绝对灰数;如果对取数有偏爱,即机会不均等,则称为相对灰数。例如对某个区带,人们

希望其储集层厚度越大越好,厚度越大,偏爱程度越大,且偏爱程度与厚度是成比例增加的,若人们对储集层厚度的愿望是有限制的,则白化函数有上界。如以厚度为横坐标,厚度的“偏爱程度”为纵坐标,可得图1(a),它表示人们希望厚度越大越好,但如果达到50 m也就足够了,该直线的斜率表示人们愿望的强烈程度,这个曲线即为厚度这一灰数的白化权函数。

有些灰数取中间某值时为最佳,过大或过小都是所不希望的,如有机质演化程度的 TTI (一种

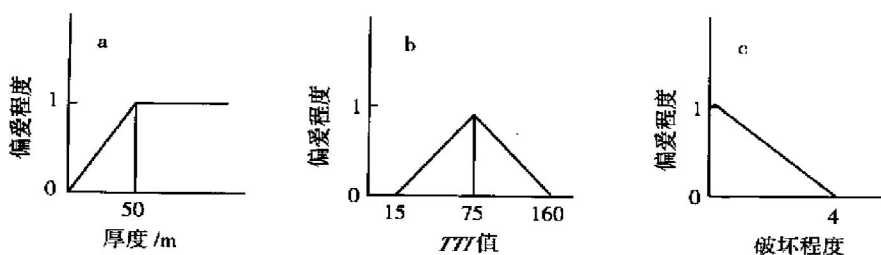


图1 常见的3种白化函数

Fig. 1 Three common graphs of definite function

表示有机质成熟度的时间-温度指数)值,一般 TTI 值在75左右时达到生油高峰,此时为最佳,系数权为1, $TTI < 15$ 时,石油还未开始生成, $TTI > 160$ 时,生油已结束,灰数 TTI 的白化权函数如图1(b)所示。

有些灰数希望越小越好,如希望区带的破坏程度(可用断层条数来描述)越小越好,破坏程度这一灰数的白化权函数如图1(c)所示。

常见的白化权函数有3种(图1a, b, c),白化权函数不一定为直线,有时也可以为曲线。白化权函数转折点的值称为阈值,它是确定白化权函数的关键,阈值可根据已有样本,按某种准则或经验确定^[1]。

一个已给出白化权函数的灰数,将某一具体数值 X_i ,通过白化权函数找出其对应的白化权系数 $f(X_i)$,再乘以该值 X_i ,即 $f(X_i) \cdot X_i$,称为灰数 $\odot$ 的一个白化函数生成值。灰色聚类是建立在白化函数生成基础上的一种数据归纳方法,它是将聚类对象对于不同聚类指标所拥有的白化数,归纳成若干个灰色类型,借以判断各聚类对象属于哪一类。

灰色聚类区带评价是基于灰色聚类原理的一

种新的区带评价方法。例如,有一批有利区带,编号分别为I, II, III, IV, V...,按含油性的优劣程度进行分类(评价),则其灰色聚类问题可作如下处理。

将区带I, II, III, IV, V...等作为聚类对象,将各种地质参数如生油岩厚度 1^* 、有机碳含量 2^* 、有机质演化程度 3^* 、储层的孔隙度 4^* 、盖层的厚度 5^* 等作为聚类指标;将含油性好1、含油性中等2、含油性差3、...作为聚类灰类(即灰类)来对区带进行分类评价。

记 d_{ij} 为聚类白化数, $i \in \{I, II, III, IV, \dots\}$, $j \in \{1^*, 2^*, 3^*, \dots\}$,它表示第 i 个区带,对于第 j 个聚类指标所拥有的白化数,它是实际工作中获得的地质参数。如第I个区带的砂岩厚度,第II个区带储层的孔隙度等等。

灰色聚类区带评价可按下述6个步骤进行^[1,2]。

第一步:给出聚类白化数 d_{ij} ,对于区带评价,就是给出每个区带的各种地质参数。

第二步:确定灰类白化权函数。这一步就是要确定各个聚类指标对于各聚类灰类的白化权函数。

记 f_{jk} 为第 j 个聚类指标对于第 k 个灰类的白化函数。

如 $\odot_1 \in [200, \infty]$, 则表示对于第1个聚类指标(生油岩厚度), 达到200 m就属于“好”; $\odot_2 \in [0.7 - \varepsilon, 0.7 + \varepsilon]$, ε 是足够小的实数, 则表示对于

第2种地质指标(有机碳含量), 在0.7%左右属于“中等”; $\odot_3 \in (0, 0.4)$, 则表示有机碳含量在0.4%以下时属于“差”。上述 $\odot_1$, $\odot_2$ 和 $\odot_3$ 的白化函数如图2所示, 图中 λ_{11} , λ_{22} , λ_{23} 为阈值。

对所有的聚类指标, 都必须确定对于每个聚

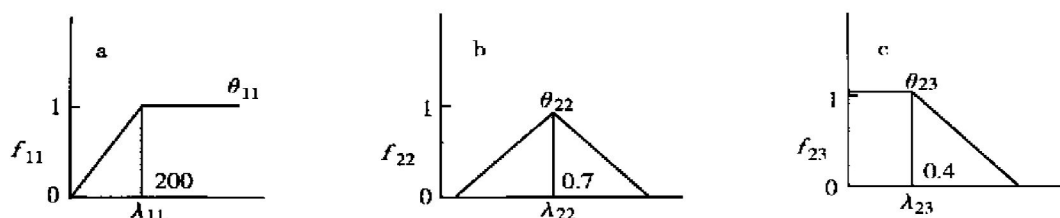


图2 生油岩厚度及有机碳含量的白化函数

Fig. 2 Definite function of source rock thickness and organic carbon content

类灰类的白化权函数, 包括函数的形状、起止点及阈值的确定, 这一步是灰色聚类的关键。聚类的结果在很大程度上取决于白化权函数。白化权函数实质上就是评价准则, 准则不同, 评价结果也不一样。这种解的非唯一性正是灰色系统的特征, 但真正符合客观事实的解只有一个, 这就要求我们根据具体的实际情况和已有的经验给出一个合理的评价准则。

第三步: 求标定聚类权。

设第 j 个聚类指标对于第 k 个灰类的聚类权为 η_k , $j \in \{1^*, 2^*, \dots, n^*\}$; $k \in \{1, 2, 3\}$, 则

$$\eta_k = \frac{\lambda_k}{\sum_{i=1^*}^n \lambda_k}$$

式中 λ_k 为第 j 个聚类指标对于第 k 个灰类的白化权函数 f_{jk} 的阈值。

第四步: 求聚类系数。

记第 i 个聚类对象(即第 i 个区带)关于第 k 个灰类(好、中或差)的聚类系数为 α_k , 则

$$\alpha_k = \sum_{j=1^*}^n f_{jk}(d_{jk}) \eta_k$$

$$i \in \{I, II, III, IV, \dots, m\};$$

$$j \in \{1^*, 2^*, 3^*, \dots, n^*\}; k \in \{1, 2, 3\}$$

第五步: 构造聚类行向量。

对于第 i 个区带的行向量记为 α_i , 则

$$\alpha_i = [\alpha_{i1}, \alpha_{i2}, \alpha_{i3}]$$

即将第 i 个区带分别属于“好”, “中等”, “差”

类的聚类系数作为元素构成一组向量。

第六步: 聚类。即判断第 i 个区带总体上属于“好”、“中等”或“差”中的哪一类。若 α_i 中第 k^* 个聚类系数最大, 即

$$\alpha_{ik^*} > \alpha_{ij} \quad j = 1, 2, 3; j \neq k^*$$

$$\text{或者 } \alpha_{ik^*} = \max_k(\alpha_{ik})$$

则第 i 个区带属于第 k^* 个灰类。

通过灰色聚类分析, 可判别要评价的各个区带总体上分别属于“好”、“中等”、“差”中的哪一类, 进行优选排队, 实现区带评价的目的。

在灰色聚类数据处理中, 若各个灰类的白化值在数量上相差太悬殊时, 应该先作无量纲化处理, 如归一化处理等, 相应地灰类的白化函数也是在无量纲意义下给出的^[4]。

2 灰色聚类区带评价

2.1 参数选取

根据油气的生成、运移、储集、保存及破坏的全过程, 将影响区带含油性的地质因素分为: 生油条件、储集条件、圈闭条件、保存条件、运移条件及配套条件等几类, 其中每一类又包括若干次级因素。在评价一个区带时, 不仅要考虑各地质因素本身的特征及其对含油性的贡献, 更要考虑各地质因素之间相互联系、相互合理的配置关系。

圈闭条件, 包括圈闭的类型、圈闭面积、闭合幅度及圈闭埋深等; 生油条件包括生油岩类型、生油岩厚度、有机质类型、有机碳含量及有机质成熟

度等; 储集条件包括储集岩类型、储集岩厚度、孔隙度、渗透率和储层横向连通性等; 保存条件包括盖层岩性、盖层厚度、断层性质、断层断距、地层水性质及地层水矿化度等; 运移条件包括运移通道、运移距离、运移方式以及是否位于排烃指向等; 圈闭配套条件包括时间和空间配套条件, 只有在大规模油气运移发生以前或与其同时形成的圈闭才是有效的, 另外, 圈闭的生、储、盖在空间上要合理配置, 包括生、储、盖发育旋回次数, 生油层、储集层单层厚度的配置情况等。

区带是个复杂的灰色系统, 上述的参数不一定全部齐备, 有些参数也不一定用得上。这就要求我们首先分析影响含油性的主要因素, 合理地选取聚类指标, 然后对每一个聚类指标确定其白

化权函数, 其中阈值的确定最为关键, 这可以通过统计已有的区带、参考邻区的情况或者根据以往的经验合理确定^[5~7]。对于待评价的区带, 灰色聚类分析能给出一个总体的、客观的和定量的评价。通过信息的不断补充, 评价的精度会逐步提高。

2.2 有利区带评价结果

由于资料所限, 目前仅对流体势分析指出的部分有利油气聚集区带(已知油田区及垦东凸起部分除外), 应用灰色聚类分析进行评价。评价参数的分布面积有限, 仅根据有机质类型、有机碳含量、有机质成熟度、区带所处的沉积相带、地层厚度、砂岩厚度、砂岩百分含量及排烃指向等 8 个地质指标进行评价, 其结果见表 1。

表 1 孤东地区石油聚集区带灰色聚类评价结果
Table 1 Evaluation results of oil accumulation blocks
in Gudong area with gray clustering method

序号	有利区名称	属于“好”类的系数	属于“中”类的系数	属于“差”类的系数	综合评价结果
1	垦东 2 井东南(Ng)	0.35	0.50	0.72	差
2	垦古 1 井南(Ng)	0.35	0.50	0.72	差
3	垦东 11 井(Ng)	0.46	0.56	0.60	差
4	孤东 21 井北(Ng)	0.52	0.55	0.54	差
5	垦利油田南(Ng)	0.35	0.57	0.70	差
6	孤深 1 井南(Ng)	0.59	0.29	0.53	差
7	孤南 34 井北西(Ng)	0.50	0.43	0.60	差
8	孤东 53 井南(Ng)	0.68	0.21	0.46	中
9	垦东 4 井北(Ed)	0.47	0.59	0.48	差
10	孤南 133 井(Ed)	0.78	0.51	0.32	好
11	孤东 34 井西(Ed)	0.72	0.44	0.33	好
12	孤东 34 井南东(Ed)	0.69	0.41	0.37	中
13	孤南 18 井(Ed)	0.82	0.44	0.27	好
14	富林东(Es)	0.48	0.34	0.43	差
15	孤深 1 井南(Es)	0.91	0.27	0.10	好
16	孤南 18 井(Es)	0.86	0.26	0.17	好
17	孤东 34 井北西(Es)	0.72	0.08	0.23	好

从表中可看出, 第 10, 11, 13, 15, 16, 17 号区带含油条件较好, 不仅 $\sigma_1 > \sigma_2 > \sigma_3$, 而且均大于 0.7, 尤其是 15, 16 号区带含油的可能性最好。第 8 和第 12 号区带含油条件中等, 本来它们也有 $\sigma_1 > \sigma_2 > \sigma_3$, 但由于 σ_1 较小, 不是以绝对的优势取胜, 若以 $\sigma_1 > 0.7$ 作为属于“好”一类的及格标准, 则它们的含油条件并不好, 因而将它们归于中等。第 1, 2, 3, 4, 5, 6, 7, 9, 14 号区带含油条件较差, 本来第 6, 14 号区带有 $\sigma_1 > \sigma_2 > \sigma_3$, 第 4, 9 号区带有 $\sigma_1 < \sigma_2 > \sigma_3$, 但由于 σ_1 和 σ_2 并不是以绝对的优势取胜,

若以 $\sigma_1 > 0.6, \sigma_2 > 0.6$ 作为属于中等以上的及格条件, 则上述 4 个区带不能归于中等以上, 因而将其含油性归于较差是合理的。

2.3 应用效果

根据上述油气聚集区灰色聚类评价结果, 优选了含油可能性最好的孤深 1 井南部(表 1 第 15 号)沙河街组进行了勘探, 2000~ 2001 年部署完钻了孤东 281、282 等探井, 获得成功。其中孤东 281 井在沙河街组钻遇油层 8 层 24.7 m。对 3 531~ 3 537 m 沙河街组三段中含油层段进行常规试油,

日产油 13 t, 不含水。对第二层 3 462~ 3 469 m 试油, 日产石油 28 t, 不含水。孤东 282 井在沙河街组钻遇油层 10.1 m/2 层, 油水同层 17.7 m/4 层。对 3 667.1~ 3 671.0 m 层段, 日试采石油 15 t, 不含水; 上返 3 592.7~ 3 600.0 m, 试油抽吸, 日产液体 16.1 t, 日产石油 13.4 t, 含水 17%。通过所做的构造和储层预测图, 上报含油面积为 2.5 km², 储量 278 × 10⁴ t。取得了理想的勘探效果。

3 结语

(1) 区带地质是一个复杂的灰色系统, 越合理地选取含油性的主要因素、合理地选取聚类指标, 越能给出一个总体的、客观的、定量的评价; 信息的不断补充, 评价的精度会逐步提高。

(2) 通过对地质系统的灰色聚类分析, 可以对油气勘探系统管理多提供一层保障, 有助于降低勘探风险, 提高勘探效益。

(3) 通过对孤东地区的灰色聚类分析应用, 此方法具有较高的实用性, 对油气勘探评价系统具有可推广使用的普遍意义。

参 考 文 献

- 1 诸克军, 冯光华, 冯刚顶. 一种石油勘探决策的模糊聚类分析法[J]. 江汉石油学院学报, 1999, 21(3): 15~ 20
- 2 金燕. 测井模糊聚类法在复杂碳酸盐岩气藏含气面积圈定中的应用[J]. 天然气工业, 1997, 17(2): 24~ 27
- 3 陈亮, 聂昌谋. 模糊聚类在胡十二块油藏质量评价中的应用[J]. 断块油气田, 1997, 4(6): 15~ 18
- 4 吴福保, 李奇, 宋文忠. 模糊聚类分析的传递方法[J]. 东南大学学报(自然科学版), 1999, 29(2): 105~ 110
- 5 王俊奇, 徐勇, 郭自新. 投产气井的模糊聚类分析[J]. 钻采工艺, 1998, 21(4): 30~ 34
- 6 王新华, 毕建涛. 基于模糊聚类神经网络的煤炭资源分类研究[J]. 自然资源学报, 2001, 16(3): 275~ 282
- 7 程军. 利用模糊聚类分析的有序谱系图对沉积物的特征进行研究[J]. 中国煤田地质, 1997, 9(1): 30~ 32

GRAY CLUSTERING ANALYSIS AND ITS APPLICATION TO PETROLEUM EXPLORATION IN GUDONG AREA

Tan Heqing^{1,2} Peng Cuncang² Wu Guohua² Lei Lijing² Liu Shibin²

(1. Lanzhou Geological Institute of the Chinese Academy of Sciences, Lanzhou, Gansu;

2. Shengli Oilfield, Dongying, Shandong)

Abstract: The gray clustering method of gray system theory is introduced in this article. The gray characteristics of geological plays have carefully been studied. A thorough analysis and intensive research have also been made on parameters needed in gray clustering analysis. 17 plays which lie in Gudong oilfield, Zhanhua sag, Jiyang depression have been analyzed with the gray clustering method. The results show that some of them are favorable, and exploration practice has confirmed that this method is practicable.

Key words: gray clustering; Gudong area; plays evaluation; exploration application