

# 致密砂岩气藏储层分类新方法

王礼常, 王志章, 陶 果

中国石油大学(北京); 油气资源与探测国家重点实验室, 北京 102249

**摘要** 储层分类是油气勘探开发研究的重要内容之一, 目前致密砂岩储层在低孔低渗储层占有很大比例, 因此研究致密砂岩储层分类具有重要意义。本文结合四川川西新场地区须家河组致密砂岩储层的特点, 综合考虑多种参数, 选用孔隙度、渗透率、流动带指数  $a_{FZ}$ 、泥质含量作为储层特征参数, 采用 K-means 聚类方法将储层分为 5 大类, 并在此基础上, 通过对比判别分析和决策树判别准确率, 采用决策树方法建立储层分类的定量评价模式。实际资料证明, 决策树具有更好的效果, 提高了储层分类评价的效果。

**关键词** 储层分类评价; 聚类分析; 决策树

中图分类号 TE19

文献标识码 A

doi 10.3981/j.issn.1000-7857.2011.24.006

## A Novel Classification Method for the Tight Sandstone Reservoir

WANG Lichang, WANG Zhizhang, TAO Guo

State Key Laboratory of Petroleum Resources and Processing, China University of Petroleum, Beijing 102249, China

**Abstract** Reservoir classification is one of the important parts of the study involving oil and gas exploration and development. Engineering data indicate that reserves of tight sandstone reservoir account for a large proportion of the total reserves in low permeability reservoirs, therefore it is meaningful to work on the tight sandstone reservoir classification. Considering the specialty and various kinds of parameters of the tight sandstone gas field of Xinchang Xu-2 formation in the west part of Sichuan Province, the four parameters, that is, porosity, permeability, flow unit, and shale content are selected as reservoir parameters, and the K-means clustering method is used to divide the formations into five categories. On basis of this, by comparing the accuracy of the discriminant analysis with that of decision tree, the decision tree classification method is used to build the quantitative evaluation of the reservoir model. It is proved by the actual data that decision tree has a better result for enhancing the effect of reservoir classification and evaluation.

**Keywords** reservoir classification and evaluation; cluster analysis; decision tree

### 0 引言

随着油气勘探开发不断深入发展, 油气藏储层分类评价是否合理已经成为勘探开发的关键, 客观地对储层进行分类评价正在向多学科综合性发展。目前国内外许多专家做了大量的工作, 储层分类的参数越来越多<sup>[1-4]</sup>, 方法标准也不断涌现<sup>[5-8]</sup>。近年来, 储层分类评价的参数大多集中在以下参数: 有效厚度、有效厚度钻遇率、有效孔隙度、渗透率、泥质及碳酸岩含量、孔隙结构参数、渗透率变异系数、渗透率突进系数、渗透率级差等。此外, 还有压力、产量、采油指数等参数。研究方法则集中在利用多元逐步回归分析、R 型主因子分析、多种

非线性单相关回归等数学分析方法任选上述各参数作为评价参数, 采用常用的评价方法, 如权重评价法、聚类分析法、模糊综合评判法、模拟试验法等, 根据已知数据段得到储层分类类别, 然后通过神经网络、判别分析等数学方法建立储层分类模式, 推广到整个研究区域。

在目前非常规储层越来越多的情况下, 储层分类方法都具有一定的地区适应性, 要实现储层分类评价的合理性, 必须针对区块储层特点, 合理选择各种参数, 运用多种手段进行分类评价。本文以四川川西新场气田须家河组须二段储层为例, 通过选择储层参数进行聚类分析得到储层分类; 在建

收稿日期: 2011-03-03; 修回日期: 2011-07-15

基金项目: “十二五”国家科技重大专项(2011ZX05008-004)

作者简介: 王礼常, 博士研究生, 研究方向为油气藏开发地质, 电子信箱: wlcwivi@foxmail.com; 王志章(通信作者), 教授, 研究方向为油气田开发地质, 电子信箱: whx3998@vip.sina.com

立分类模式过程中,对比判别分析这种常用的方法,研究引用数据挖掘中的决策树<sup>[9-10]</sup>方法进行致密砂岩储层分类模式识别,通过对输入参数权重判别,自动选择参数进行分类,在达到定量识别的基础上提高识别正确率。

## 1 K-means 聚类分析

K-means 是典型的基于距离的聚类算法,其中  $K$  表示的是最终分类个数。采用距离作为相似性的评价指标,根据分类个数  $K$  随机地选取  $K$  个初始的聚类中心,不断迭代。在每一次迭代中,计算各点与各聚类中心的距离,并将距离最近的类作为该点所属的类,即当目标函数达到最小值时,得到的聚类为最终聚类结果,从而将数据分为  $K$  类。K-means 算法的目的是将一个集合进行等价类划分,即对数据结构相同的记录按照某种分类规则,将其划分为几个同类型的记录集。

目标函数采用平方误差准则,即  $E = \sum \sum |P - m_i|^2$ , 其中  $E$  为各个聚类对象的平方误差之和,  $P$  为聚类对象,  $m_i$  为类  $C_i$  的各聚类对象的平均值<sup>[11]</sup>。

K-means 算法的迭代过程可以描述为:

- (1) 将样本分为  $K$  类;
- (2) 选择  $K$  个点作为“种子”;
- (3) 计算各个样本数据点到种子之间的距离,将样本重新分为  $K$  类;
- (4) 再将每类的平均值作为新的种子,重新聚类预算;
- (5) 重复上述计算过程,当目标函数达到最小值时运算结束,最终将样本分为  $K$  类。

## 2 决策树

决策树(decision tree)是一个类似于流程图的树结构,它从根节点开始对数据样本进行测试,根据不同的结果将数据样本划分成不同的数据样本子集,每个数据样本子集构成 1 个子节点,生成的决策树每个内部节点表示在 1 个属性上的测试,每个分枝代表 1 个测试输出,而每个叶节点代表类或类分布。

决策树算法主要包括 2 个过程:构造过程和分类过程。构造决策树分为建树和剪枝 2 个阶段,前者归纳出决策树,后者用于防止过度拟合。生成最优决策树是一个 NP 难题,现有算法通过启发式属性选择策略完成树的构造<sup>[12]</sup>。

用决策树分类一般分为两步:(1) 利用训练集建立一个决策树模型,这个过程实际上是一个从数据中获取知识,进行机器学习的过程,属于建模过程;(2) 利用生成完毕的决策树对未知的数据样本进行分类,属于应用过程。

## 3 储层分类

### 3.1 储层参数选择

新场气田须二段属于超低孔、超低渗气藏,储层物性差、

孔隙结构复杂、裂缝发育、非均质性极强,实验分析中压汞资料、粒度分析资料极其缺乏,因此通过主因子分析,最终选择孔隙度  $P_{or}$ 、渗透率  $P_{em}$ 、泥质含量  $V_{SH}$ 、流动带指数  $a_{FZI}$  这 4 个参数。其中孔隙度能够反映储层的宏观物性特征,泥质含量  $V_{SH}$  能反映沉积环境,  $a_{FZI}$  反映微观孔渗结构;由于该地区裂缝发育,据经验统计,一般认为该地区渗透率大于  $0.1 \times 10^{-3} \mu m^2$  的为裂缝发育储层,因此引入渗透率参数可以较好地反映储层的裂缝发育情况,以上 4 个参数可以较好地反映储层的多种特征,具有良好的可操作性和可信度。

### 3.2 储层类别划分

通过主因子分析,选择参数为取心井中岩心分析的  $P_{or}$ 、 $P_{em}$  以及对应计算深度得出的  $V_{SH}$ 、 $a_{FZI}$ , 共计 4 个参数,采用 K-means 方法进行聚类分析,经过多次尝试,在分拆 5 类时,储层分类具有较好的效果及地质意义。

K-means 方法对以上 4 个参数聚类的过程较为简单,根据聚类目标分类个数 5 类随机的选取 5 个初始的聚类中心,然后对数据样本不断迭代。在每一次迭代中,计算每个点与各聚类中心的距离,并将距离最近的类作为该点所属的类。当目标函数  $E$  达到最小值时,得到这 5 类的聚类中心数据,将数据样本分为 5 类,如表 1 所示。同时,还应用了模糊聚类分析对数据样本进行聚类,并将结果与 K-means 聚类方法进行对比。须二段气田对无效储层的划分界限为孔隙度小于 3%,渗透率小于  $0.04 \times 10^{-3} \mu m^2$ 。在这两种方法中,K-means 将数据样本聚类成 5 类时,与须二段气田储层分类经验非常接近;而采用模糊聚类时,不管是分 4 类或者 5 类,与油田划分下限均由较大差距,因此本次分类采用 K-means 聚类方法,聚类结果如表 1 所示。

表 1 新场气田须二段储层划分标准

Table 1 Division standard for reservoir type of Xu-2, Xinchang gas field

储层类别	渗透率/( $10^{-3} \mu m^2$ )	孔隙度/%	$a_{FZI}$	泥质含量/%
I	68.21	5.84	31.15	15.43
II	14.88	2.36	75.45	15.73
III	0.06	5.52	1.62	14.01
IV	0.05	4.04	3.19	36.3
V	0.04	2.95	3.92	11.5

由表 1 可以看出,不同的储层之间,除了孔隙度差别较小之外,其余差异较为明显,其中反映储层渗透能力和非均质性的渗透率、流动带指数  $a_{FZI}$  差别最大。由于新场气田须二段大段发育致密砂岩,储层平均厚度约 30m,泥质含量较小,分布区间为 5%—40%之间,且泥质含量主要集中在 8%—20%之间;IV 类储层则是在聚类时,将泥质含量较大的数据聚成了一类,因此较为特殊,除了 IV 类储层外,其余类储层泥质含量差异较小。

I 类储层:储层物性很好,其中渗透率最好,均值达

$68.21 \times 10^{-3} \mu\text{m}^2$ , 泥质含量较低, 而流动带指数  $a_{fz}$  达 31.15, 可见 I 类储层裂缝发育良好, 而且很好地结合了基质储层孔隙发育。I 类储层又可以看作是孔隙-裂缝都发育的储层类型, 是该区投产的最有利区。

II 类储层: 孔隙不发育, 孔隙度均值只有 2.36%, 但是渗透率却较高, 反映出裂缝很发育, 其  $a_{fz}$  值也很高, 达到 75.45, 同样反应该类储层渗透率与孔隙度之间比值较大, 裂缝发育; II 类储层可以看成是裂缝性储层类型。这一类储层往往初期投产为高产, 但由于储层基质孔隙不发育, 高产期很短。

III 类储层: 储层孔隙很发育, 孔隙度均值为 5.52%, 但 III 类储层渗透率较低, 裂缝不发育, 而仅仅是基质孔隙发育; III 类储层可以作为该区投产的次有利区。

IV 类储层: 储层孔隙中等发育, 渗透率很低, 裂缝不发育, 但泥质含量很高, 平均泥质含量为 36.30%, 可见 IV 类储层由于泥质的存在, 严重的影响储层的渗流能力。IV 类储层可以作为该区投产的潜力区。

V 类储层: 储层孔隙度最小, 渗透率同样最小, 可见同样裂缝不发育。V 类储层中砂岩致密, 导致储层几乎丧失渗流能力, V 类储层实际上丧失了作为储集空间的能力, 但是可以作为各类储层的渗流屏障, 如在平面及剖面上规模发育而又无裂缝发育, 则可以作为储层的良好隔夹层, 因此在气田开发射孔时, 应该注意该类储层在空间的分布情况, 合理射孔。

### 3.3 储层类别识别

储层类别确定后, 下一步的工作就是要确定分类的规则, 这里采用决策树方法, 同样选取  $P_{or}$ 、 $P_{em}$ 、 $V_{sh}$ 、 $a_{fz}$  这 4 个参数生成储层分类规则。

在取心井划分储层类别的基础上, 推广到非取心井中, 大多数学者采取的是判别分析方法, 通过建立各类储层的判别函数, 将计算出的非取心井的孔隙度、渗透率(取以 10 为底的对数)及  $a_{fz}$ (取以 10 为底的对数)、泥质含量代入判别函数, 以计算得到最大值的类别作为该数据点的类别。

使用判别分析时, 5 类储层判别函数为

I 类储层:

$$y = 79.44P_{or} - 180.17\lg P_{em} - 0.22V_{sh} + 384.14\lg a_{fz} - 389.29$$

II 类储层:

$$y = 74.05P_{or} - 176.8\lg P_{em} - 0.13V_{sh} + 377.59\lg a_{fz} - 370.32$$

III 类储层:

$$y = 77.35P_{or} - 182.92\lg P_{em} - 0.22V_{sh} + 366.48\lg a_{fz} - 371.9$$

IV 类储层:

$$y = 73.15P_{or} - 178.47\lg P_{em} - 0.24V_{sh} + 358.19\lg a_{fz} - 357.8$$

V 类储层:

$$y = 72.29P_{or} - 177.33\lg P_{em} - 0.18V_{sh} + 356.75\lg a_{fz} - 339.3$$

经检验, 判断分析正判率为 84%。

对储层聚类分析后进行判别分析, 这是大部分学者的普遍思路<sup>[13]</sup>, 实际上, 除了判别分析外, 同样可采取决策树方法识别聚类后的储层类别。

对储层的 5 种分类, 本文采取决策树 C5.0 进行识别, 参与识别的参数与判别分析一样选择了以下 4 个参数: 孔隙度、渗透率(取以 10 为底的对数)及  $a_{fz}$ (取以 10 为底的对数)、泥质含量。

图 1 为决策树对已知分类数据样本的建模过程, 通过孔隙度  $P_{or}$ 、 $\lg P_{em}$ 、 $V_{sh}$ 、 $\lg a_{fz}$  共计 4 个参数进行分类识别, 识别的主要过程如图 1 所示, 部分规则由于数据点所占有的比例小, 形成的规则代表性不强, 予以省去。

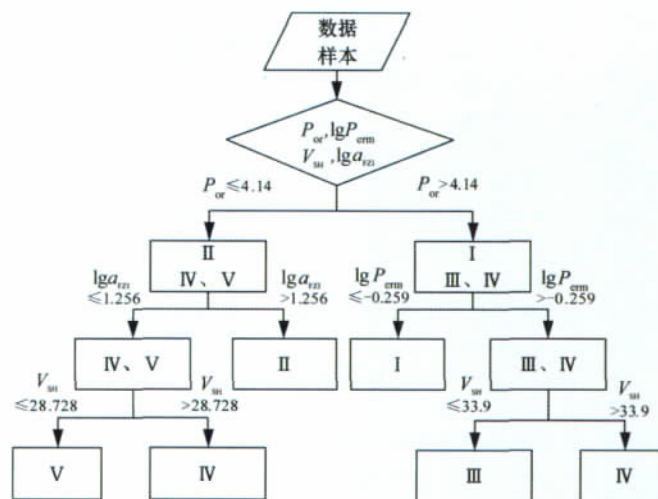


图 1 储层分类决策树识别结果

Fig. 1 Recognition results of reservoir classification by using decision tree

采用决策树对聚类分析后的储层识别正确率可以达到 99.7%, 而判别分析正判率为 84%, 二者采用的相同的参数参与分类识别, 与判别分析正判率相比决策树方法得到提高, 通过决策树分类建模过程得到的规则, 通过程序处理, 对须二段储层新井 A 井进行储层分类, 在 I 类储层较为集中的井段中试气, 获得较好的工业产能。

### 4 结论

目前低孔低渗储层已经成为油气田增加储量和产量的重要组成部分, 而致密砂岩储层在低孔低渗储层中占有很大比例, 因此研究致密砂岩储层分类方法有重要意义。

(1) 对越来越多的非常规储层, 储层分类方法应该多样化, 并且要有合理性, 聚类分析是储层分类的一种有效的方法, 通过对储层参数聚类达到划分储层类别的目的。

(2) 在对储层分类后, 识别分类模式, 进而达到定量识别的目的成为重点, 相比于较为常用的判别分析方法, 决策树能够更好地达到定量识别储层类别的目的, 同时建立起来的识别模式准确度高, 具有很好的实用性。

(3) 在聚类分析和决策树基础上建立起来的分类识别模式, 便于计算机编程处理, 结果可靠, 操作方便, 为非常规油气藏大量的数据处理提供了很好的平台。



# 参考文献 (References)

- [1] 刘宗彦, 王燕. 测井储层分类评价方法的研究[J]. 国外测井技术, 2008, 23(4): 19-22.  
Liu Zongyan, Wang Yan. *World Well Logging Technology*, 2008, 23(4): 19-22.
- [2] 杨秋莲, 李爱琴, 孙燕妮, 等. 超低渗储层分类方法研究 [J]. 岩性油气藏, 2007, 19(4): 51-56.  
Yang Qiulian, Li Aiqin, Sun Yanni, et al. *Lithologic Reservoirs*, 2007, 19(4): 51-56.
- [3] 徐炳高. 川东北地区碳酸盐岩储层分类与油气识别方法研究[J]. 测井技术, 2004, 28(5): 410-413.  
Xu Binggao. *Well Logging Technology*, 2004, 28(5): 410-413.
- [4] 赵靖舟, 吴少波. 论低渗储层的分类与评价标准[J]. 岩性油气藏, 2007, 19(3): 28-31.  
Zhao Jingzhou, Wu Shaobo. *Lithologic Reservoirs*, 2007, 19(3): 28-31.
- [5] 付殿敬, 徐敬领, 王贵文. 基于 Q 型聚类分析和贝叶斯判别算法研究储层分类评价[J]. 科技导报, 2011, 29(3): 29-33.  
Fu Dianjing, Xu Jingling, Wang Guiwen. *Science and Technology Review*, 2011, 29(3): 29-33.
- [6] 姚萌, 徐樟友, 熊琦华, 等. 数理统计分析方法在储层分类中的应用[J]. 石油学报, 1994, 15(S1): 105-109.  
Yao Meng, Xu Zhangyou, Xiong Qihua, et al. *Acta Petrolei Sinica*, 1994, 15(S1): 105-109.
- [7] 杨波, 高清祥. 聚类分析法在城壕油田西 259 井区长 32 储层分类评价中的应用[J]. 石油天然气学报, 2010, 32(6): 22-26.

- Yang Bo, Gao Qingxiang. *Journal of Oil and Gas Technology*, 2010, 32(6): 22-26.
- [8] 王长江, 姜汉桥. 基于支持向量机和聚类理论的储层评价方法研究[J]. 油气田地面工程, 2009, 28(2): 18-20.  
Wang Changjiang, Jiang Hanqiao. *Oil-Gasfield Surface Engineering*, 2009, 28(2): 18-20.
- [9] 李洪奇, 郭海峰. 复杂储层测井评价数据挖掘方法研究[J]. 石油学报, 2009, 30(4): 542-549.  
Li Hongqi, Guo Haifeng. *Acta Petrolei Sinica*, 2009, 30(4): 542-549.
- [10] 石广仁. 数据挖掘在石油勘探数据库中的应用前景 [J]. 勘探管理, 2009(1): 60-64.  
Shi Guangren. *Exploration Management*, 2009(1): 60-64.
- [11] 薛敬桃, 袁恩来, 杨鹏, 等. 用 K-means 聚类分析进行储层分类[J]. 内蒙古石油化工, 2010(20): 34-35.  
Xue Jintao, Yuan Enlai, Yang Peng, et al. *Inner Mongolia Petrochemical Industry*, 2010(20): 34-35.
- [12] 李雄飞, 董元方, 李军, 等. 数据挖掘与知识发现 [M]. 2 版. 北京: 高等教育出版, 2010.  
Li Xiongfei, Dong Yuanfang, Li Jun, et al. *Data mining and knowledge discovery*[M]. 2nd ed. Beijing: High Education Press, 2010.
- [13] 唐海发, 彭仕宓, 赵彦超. 致密砂岩令藏储层流动单元划分方法及随机模式[J]. 吉林大学学报: 地球科学版, 2007, 37(3): 469-474.  
Tang Haifa, Peng Shimi, Zhao Yanchao. *Journal of Jilin University: Earth Science Edition*, 2007, 37(3): 469-474.

(责任编辑 代丽)

## ·学术动态·



## “2011 年盆地动力学与油气储层研讨会”征文

由中国地质学会主办的“2011 年盆地动力学与油气储层研讨会”将于 2011 年 10 月 21—22 日在武汉召开,会议主题为“盆地动力学与油气储层”。

征文范围:层序地层学,重大地质事件、气候和环境演化的沉积记录,沉积盆地形成演化与油气富集机制,碳酸盐岩沉积学与油气藏,深海沉积学及其资源环境效应,新技术、新方法在沉积学中的应用。

论文截止日期 2011 年 10 月 15 日。

联系电话 027- 67848580。

电子信箱 petrolab@cug.edu.cn。

会议网址 [www.cug.edu.cn/new/Show.aspx?ID=518&cid=000103](http://www.cug.edu.cn/new/Show.aspx?ID=518&cid=000103)。