

基于主成分分析法的测井数据识别岩性

——以富县延长组长6地层为例

陈阳¹, 王兴建¹, 何发岐², 齐荣²

(1. 成都理工大学 地球物理学院, 成都 610059;

2. 中国石化华北油气分公司, 郑州 450000)

摘要: 油气藏的分布与岩性息息相关, 所以进行岩性的识别分层, 有助于判断地层的储层分布。取心实验十分昂贵, 同时在钻井过程中获得岩心是不连续的, 所以存在岩心数据较少的问题。为了更好地判断岩性和降低成本, 在本研究中, 将陕北富县延长组长6地层发育较多的四种岩性一共99个岩心取点与其对应的测井数据进行相关性分析。选择的自然伽马测井(GR)、井径测井(CAL)、声波时差测井(AC)和中子补偿测井(CNL)四条测井曲线来分析富县延长组长6岩性特征。随着粒径变大和黏土含量的减少, GR、CAL、AC和CNL值趋于降低, 能够明显划分。利用主成分分析方法, 对数据进行降维, 提取特征, 进行归纳, 解决岩性的多解性问题, 建立了岩性判定指数, 然后将其运用于陕北富县地区多个钻井, 验证其方法和构建的数学模型的适用性和有效性。结果表明, 主成分分析是行之有效的分析岩性的方法, 提高了准确度, 岩性的岩性与测井数据识别的岩性具有良好的一致性。

关键词: 测井数据; 主成分分析; 岩性识别; 延长组长6地层

中图分类号: P618.13 **文献标志码:** A **DOI:** 10.3969/j.issn.1001-1749.2024.06.02

0 引言

油气藏的分布与岩性息息相关, 所以对于岩性的识别, 有助于判断地层的储层分布。随着测井质量的不断提升, 在井中已经有多种地球物理测井响应。具有多元性与高相关性特点的测井数据能反映不同岩性地球物理属性的变化规律, 是岩性分层的重要依据^[1]。不同岩性的粒径和组成成分存在差异, 从而不同岩性的物理和化学特征也会存在差异, 可以理解为地球物理测井响应特征与岩性存在某种物理关系。较高的中子孔隙度对数值通常对应于泥岩^[2]。粉砂岩和粘土岩通常比砂岩具有更大的伽马对数值^[3], 但是在复杂的沉积环境中, 多响应的特征

会干扰仅使用单条对数曲线的岩性判定^[4]。

由于测井资料具有多解性, 所以越来越多的研究者根据地球物理测井资料提出数学统计的方法来提高岩性的识别精度。主成分分析(PCA)作为一种数学统计方法, 也可以视为一种无监督学习方法, 可以对包含各种特征数据集进行降维处理^[5], 对数据进行标签也不会影响数据结构。例如, 刘爱疆等^[6]在碳酸盐岩岩性识别中采用主成分分析, 并运用第一主成分和第二主成分交会图法进行岩性识别。LI等^[7]用主成分回归模型重建临界煤层气测井曲线。任鹏飞等^[8]将地球物理测井资料用主成分分析法来识别中国贵州西北不同煤阶储层的煤结构, 提高了煤结构的识别质量。张强等^[9]基于改进主成分分析法来提高砂岩型铀矿测井曲线岩性分层精度。

收稿日期: 2023-09-03

基金项目: 国家自然科学基金项目(42074163)

第一作者: 陈阳(1999—), 男, 硕士, 从事油气地球物理勘探工作, E-mail: 1326377472@qq.com。

鄂尔多斯盆地陕北富县是进行油气勘探与开采的重要地区,其中长 6 油层组因为受到河湖影响沉积相为三角洲前缘亚相和湖泊相^[10]。由于研究区长 6 地层岩性复杂变化且油气藏的位置分布与岩性存在关系,而地下的取心实验十分昂贵,且取得的岩心是不连续的,导致岩心数据较少,为了更加精确地判断地下岩性分层,笔者根据主成分分析法,将钻井所取岩心对应测井响应资料分为训练集和验证集,建立岩性识别数学模型,从而有助于判断富县延长组长 6 地层岩性分层和储层分布。

1 地质背景





鄂尔多斯盆地是一个大型沉积盆地,富含油气^[11]。根据构造特点主要分为 6 个构造单元,位于

盆地北部的是伊盟隆起,东部的是晋西挠褶带,中部是陕北斜坡,南部为渭北隆起,西部是天环坳陷、西缘冲断带^[12]。

研究区富县位于鄂尔多斯盆地陕北斜坡构造单元的东南部(图 1),发育东西向鼻状隆起^[13-14],其中长 6 地层盆地稳定坳陷,三角洲充填,由 3 个中期旋回构成,油气显示十分丰富。通过对研究区钻井岩心的观察,发现该研究区储层发育多种岩性(表 1),包括细砂岩、粉砂岩、泥质粉砂岩、泥岩、粉砂质泥岩和凝灰岩等,其中砂岩分布广泛,存在较厚的砂体,颜色以浅灰色、灰色为主,含油评价有油斑、油迹、油浸和荧光;泥岩多呈现深灰色、灰色;粉砂岩呈浅灰色,主要以夹层的形式存在于砂泥岩系中,含油评价主要是油迹;泥质粉砂岩和粉砂质泥岩多呈现灰色,泥质粉砂岩含油评价主要是油斑,荧光。

表 1 岩性的特征及分类

Tab. 1 Characteristics and classification of lithology

岩性	岩心照片	描述
泥岩		①深灰色、灰色为主
泥质粉砂岩		①灰色为主 ②含油评价:油斑、荧光
粉砂岩		①主要呈浅灰色 ②含油评价:油迹 ③主要以夹层形式存在于砂泥岩系中
细砂岩		①浅灰色、灰色为主 ②含油评价:油斑、油迹、油浸和荧光

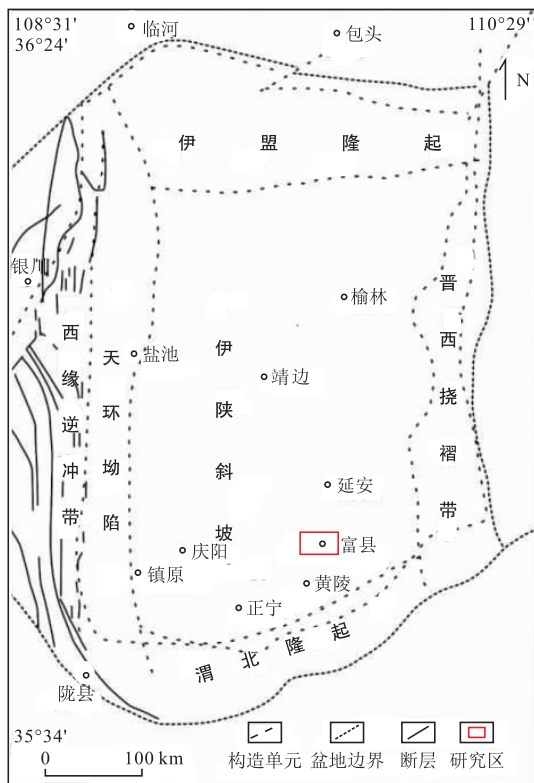


图1 鄂尔多斯盆地构造单元划分及研究区位置图
Fig. 1 Structural unit division and study area location map of the Ordos Basin

2 方法与工作流程

2.1 测井数据选取

在实际钻井过程中,由于钻杆的工作状态,导致钻杆的拉伸变形,因此钻井的深度比实际深度长^[8]。因此,需要校正钻井深度和测井深度^[15],如图2所示,得到:

$$\frac{d'_2 - d'}{d_2 - d} = \frac{d'_2 - d'_1}{d_2 - d_1} = \frac{h'}{h} \quad (1)$$

式中 h 和 h' 分别代表取岩心过程中测量的岩心厚度和用于测井数据解释的岩心厚度; d_1 、 d_2 、 d 代表取心过程中测量的岩心的顶、底和取样点的深度 d'_1 、 d'_2 、 d' 是代表用于测井解释的岩心的顶、底和采样点的深度。

在上面步骤中获得了不同岩心的岩性和深度,采用深度间距为 0.125 m 提取岩心的测井数据,并将这些数据分为训练集和验证集,再将测井数据和岩心深度一一对比。流程图如图3所示。

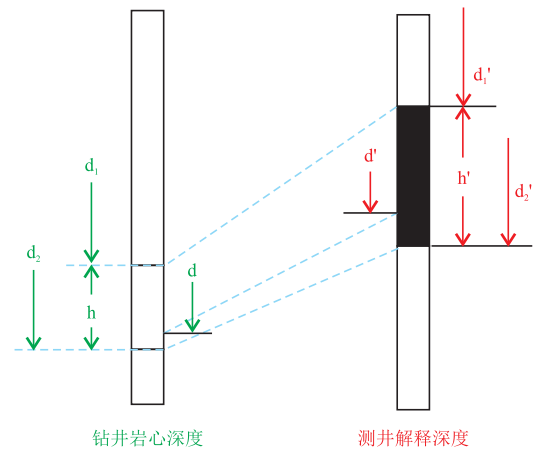


图2 钻井取心和测井解释的深度图

Fig. 2 Depth diagram of the drilled core and the well logging interpretation

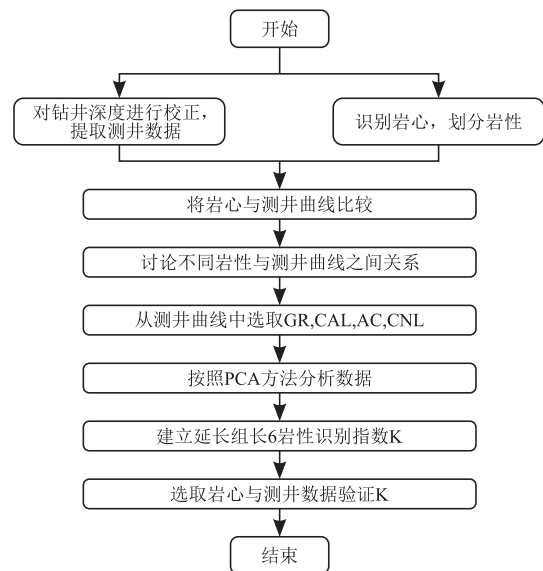


图3 利用测井数据识别岩性岩性的工作流程
Fig. 3 The workflow of identifying core lithology using logging data

根据笔者统计的探区的8口井长6地层的测井数据(洛河1井 1 028.55 m~1 037.17 m,洛河2井 731.11 m~794.85 m,洛河4井 492.56 m~498.32 m,洛河5井 683.28 m~691.32 m,洛河6井 1 006.29 m~1 230.48 m,洛河9井 522.5 m~532.25 m,洛河11井 794.50 m~814.15 m,洛河16井 499.14 m~506.02 m),整理了不同岩性与地球物理测井响应特征,图4表明 SP、DEN 测井响应与岩性没有明显的变化规律,而 GR、CAL、AC、CNL 测井数据大致遵从正态分布,从平均值来看,岩

性对应测井的曲线变化规律还是明显地划分出来,从整体上看,由泥岩、泥质粉砂岩、粉砂岩到细砂岩,随着粒径变小和黏土含量的减少,GR、

CAL、AC、CNL 测井值呈现明显的递减趋势,SP、DEN 测井值递变趋势不明显。这与文献中结果基本相似^[16-17]。

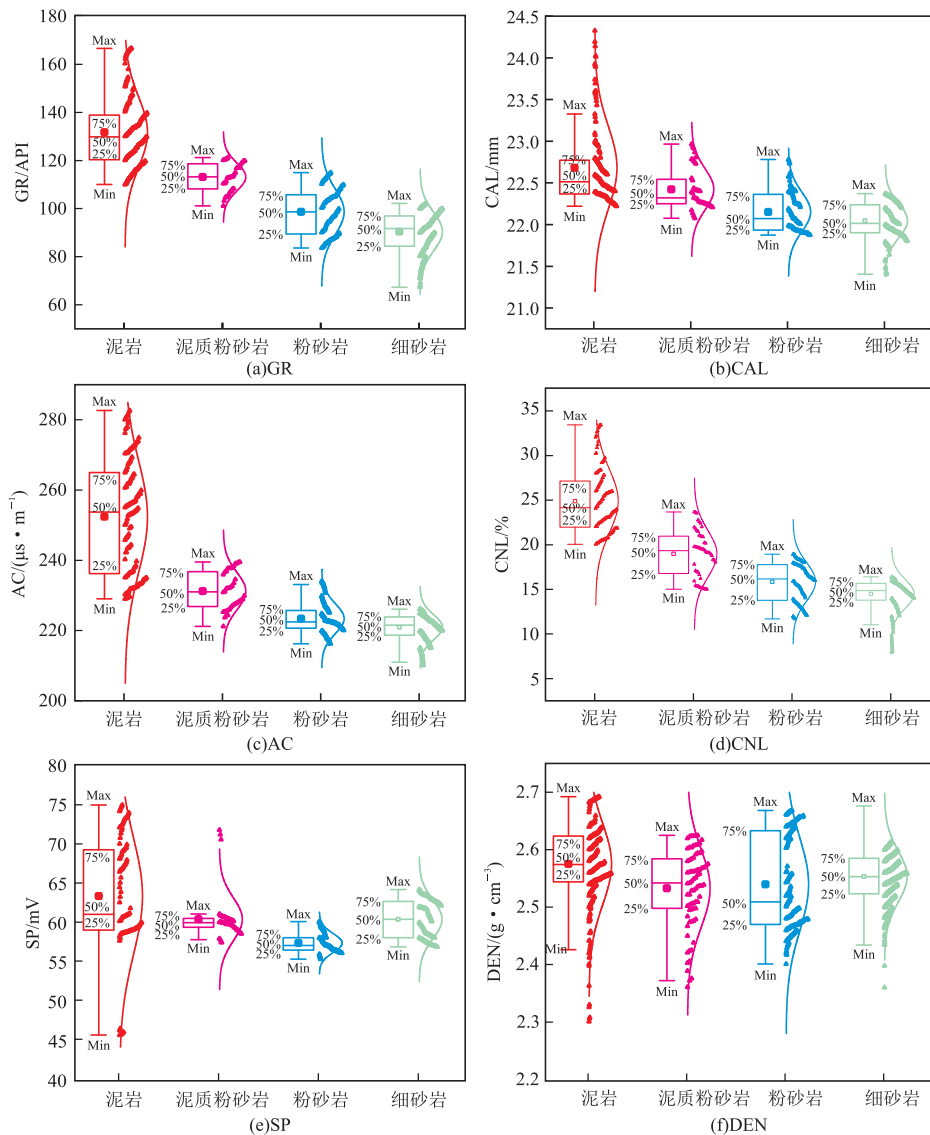


图 4 岩性测井数据的箱线图

Fig. 4 Box plot of lithological logging data

然而不同的岩性之间依然存在着数据的重叠,为了精确的划分岩性,引入了主成分分析方法(PCA)。为了确保训练集的代表性,岩心取样点测井数据的同一性以及训练后模型的适用性,所选择的岩心取样点的 GR、CAL、AC 和 CNL 测井数据(图 5)位于箱线图对应测井数据的上四分位数与下四分位数之间。泥岩、泥质粉砂岩、粉砂岩和细砂岩一共提取了 99 个测井点,从样品中学习特征,并加以总结,统计出规律,这是一个归纳的过程。笔者的目的在于数学模型训练完成后,能够通过归纳出的关系,对训练集以外的岩心的岩性准确判断。所以

根据测井数据的特点选择合适的训练集样本数量是十分必要的。通常在训练的数学模型中,样本数量较小时,增多样本数量可以使构建的模型取得更好的效果,但两者并非呈线性相关性,构建的测井数据数学模型因为模型存在归纳偏好的原因,在样本数据量偏小时,归纳偏好的信息对模型的构建带来有效的帮助,而样本数据量偏大时,反而会产生误差,对模型的学习数据特征产生限制。

2.2 主成分分析方法的步骤

主成分分析方法就是使数据集的可变性更明显,将其缩小到更小的维度,但仍然包含数据信息^[18],

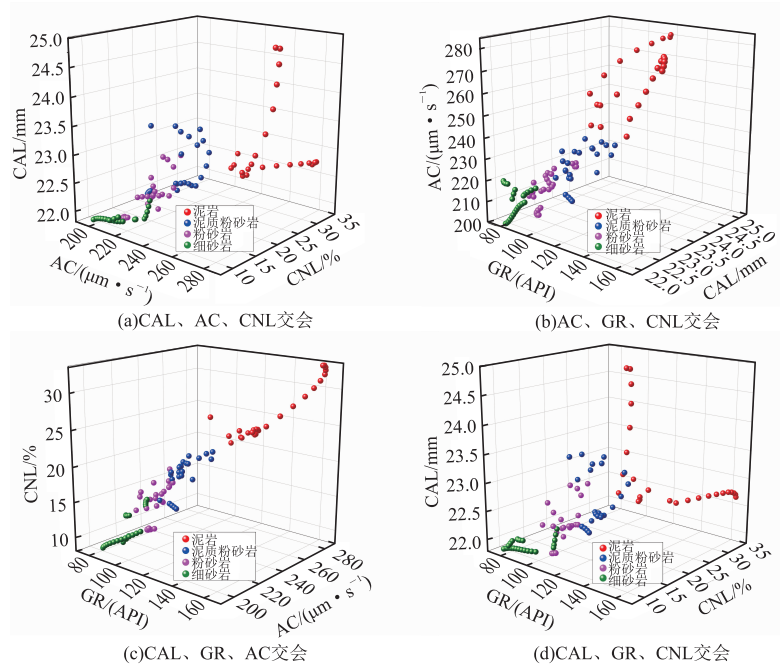


图5 训练集的不同岩性测井数据

Fig. 5 Different lithology logging data from the training set

使数据集在较低维度空间具有最大的方差^[19],最大的方差量的数据方向,就代表了大多数数据的线性方向^[20],从而以综合属性进行解释分析工作^[21]。

主成分分析步骤如下:

1)对于 m 个样本 n 种类型的测井数据,我们可以将原始测井数据表示为矩阵 $\mathbf{X} = [GR, CAL, AC, CNL]^T$ 。有 m 个样本,每个样本有 4 种变量 (X_1, X_2, X_3, X_4),即矩阵 \mathbf{X} 表示为:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{m4} \end{bmatrix} \quad (2)$$

2)数据标准化。为了消除测井曲线量纲的影响,需要对 GR、CAL、AC、CNL 的测井数据做标准化处理(图 6):

$$Z_{ij} = \frac{x_{ij} - u}{S_j} \quad (i=1, 2, \dots, nj=1, 2, 3, 4) \quad (3)$$

式中 u 为平均值, S_j 为标准差。

3)计算出标准化后数据相关系数矩阵 \mathbf{R} (表 2):

$$\mathbf{R} = [r_{ij}]_{4 \times 4} = \frac{\mathbf{Z}^T \mathbf{Z}}{n-1} \quad (4)$$

式中 r_{ij} 是标准化矩阵 \mathbf{Z} 的第 i 行和第 j 列之间的相关系数。

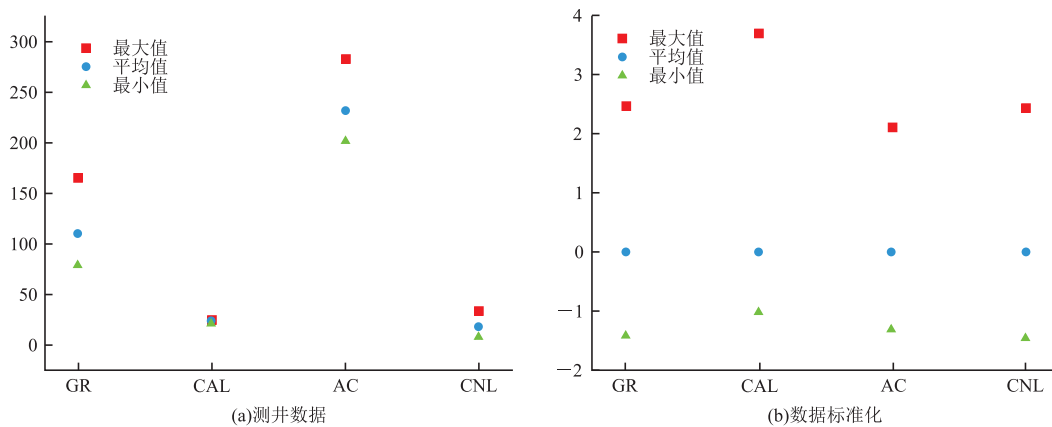


图6 测井数据和数据标准化的分布

Fig. 6 Distribution of well logging data and data standardization

表 2 研究区样本指标的相关系数矩阵

Tab. 2 Correlation coefficient matrix of sample indicators in the research area

样本指标	GR	CAL	AC	CNL
GR	1.000			
CAL	0.477	1.000		
AC	0.855	0.728	1.000	
CNL	0.941	0.584	0.915	1.000

4) 求解矩阵 R 的特征值。采用雅可比行列式法 ($|\mathbf{R} - \lambda \mathbf{I}_p| = 0$) 求取相关系数矩阵的 R 的特征值 (λ_i), 将特征值从大到小排序 ($\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$)。 I_p 是矩阵 R 的特征向量。

5) 计算每个特征值 (图 7) 占全部特征值的总数的比重 (即方差贡献率):

$$\omega_i = \left(\frac{\lambda_i}{\sum_{i=1}^4 \lambda_i} \right) \times 100\% \quad (5)$$

式中 $\sum_{i=1}^4 \lambda_i$ 为累积方差贡献率。

从表 3 可以看出第一主成分 F1 方差贡献率为 81.928%, 贡献信息最多; 第二主成分 F2 的方差贡献率为 14.851%; 两个主成分累计贡献率达到 96.780%, 超过了总贡献率的 85%, 表明尽可能保留了原始变量中的有效信息。所以我们选用第一主成分、第二主成分来表示所有测井数据的信息。

6) 求解第一 (F1)、第二 (F2) 主成分中对应的特征向量 (表 4)。公式如下:

$$R a_i = \lambda_i a_i \quad (6)$$

表 4 主成分 F1 和 F2 特征向量

Tab. 4 Principal component F1 and F2 eigenvectors

	GR	CAL	AC	CNL
F1	0.507	0.414	0.536	0.531
F2	-0.448	0.849	0.040	-0.274

7) 建立主成分的方程并计算其值。其中第 c 个主成分 F_c 表达式为:

$$F_c = a_{c1} x_1 + \dots + a_{cp} x_p \quad (c=1, 2, \dots, p) \quad (7)$$

总的主成分函数表达式为:

$$F = 0.81928 * (0.507 * GR' - 0.414 * CAL' + 0.536 * AC' + 0.531 * CNL') + 0.14851 * (-0.448 * GR' + 0.849 * CAL' + 0.040 * AC' - 0.274 * CNL') \quad (8)$$

根据以上的分析, F 可以更加完整的表示岩性信息, 因此建立岩性的判定指数 $K (K = F)$ 。通过

表 3 方差提取主成分分析表

Tab. 3 Variance extraction principal component analysis table

成分	特征值	贡献率%	累计贡献率%
第一主成分	3.277	81.928	81.928
第二主成分	0.594	14.851	96.780
第三主成分	0.083	2.081	98.861
第四主成分	0.046	1.139	100.000

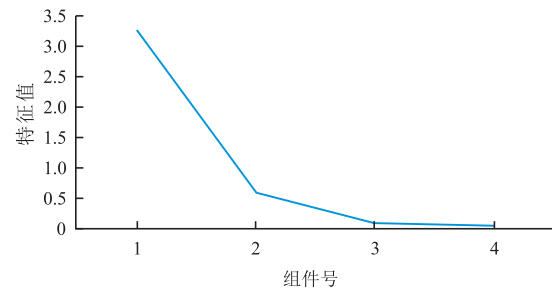


图 7 碎石图

Fig. 7 Scree plot

研究岩性的判定指数 K , 发现不同岩性的 K 有着不同的区间, 还存在明显的分界, 可以用来区分延长组长 6 不同的岩性 (图 8 和表 5)。

基于主成分分析的岩性识别方法通过降维提取了不同岩性的测井数据的主要信息, 提高了准确性, 已经被证明是一种可行的岩性识别工具。在探区中, 泥岩的指数 K 为 1 到 3.11, 泥质粉砂岩的指数 K 为 -0.27 到 1, 粉砂岩的指数 K 为 -1.28 到 -0.27, 细砂岩的指数 K 为 -2.08 到 -1.28。

表5 研究区不同岩性的边界

Tab. 5 Boundary of different lithology in the study area

岩性	泥岩	泥质粉砂岩	粉砂岩	细砂岩
指数	$3.11 > K > 1$	$1 > K > -0.27$	$-0.27 > K > -1.28$	$-1.28 > K > -2.08$

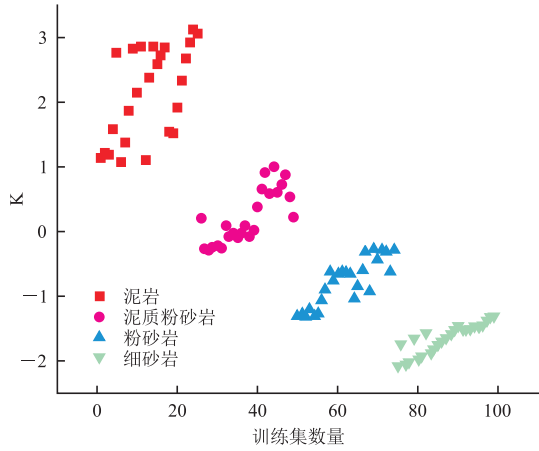


图8 不同岩性的识别指数 K

Fig. 8 Identification index K for different lithologies

3 识别方法的验证

为了保证岩性的判定指数 K 可以直接应用于现场识别岩心的岩性,所以需要验证方法的适用性与有效性。以洛河 10 号井(未参加训练)为例,对岩心描述(图 9)和方法识别结果进行比较。在岩心描述中,从上到下依次是浅灰色油斑细砂岩、深灰色泥岩、浅灰色油斑细砂岩、灰色泥岩、浅灰色细砂岩、灰黑色泥岩和浅灰色粉砂岩(图 10)。

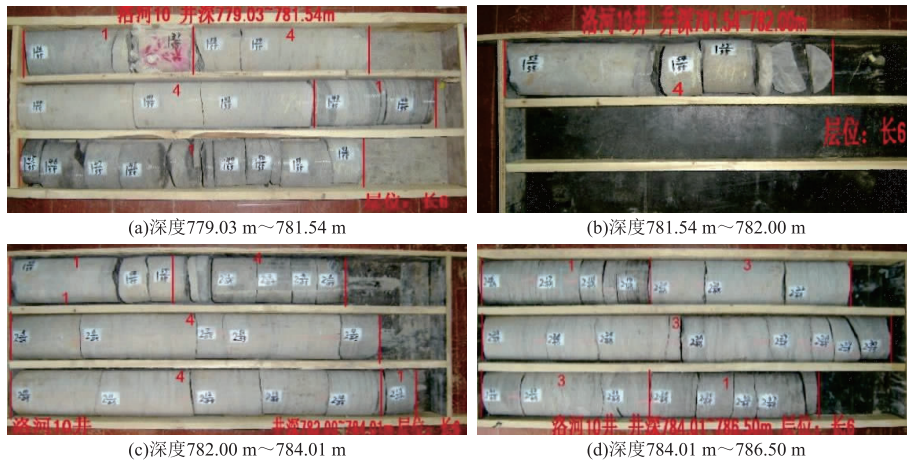


图9 洛河 10 号井岩心

Fig. 9 Core of Luohe well 10

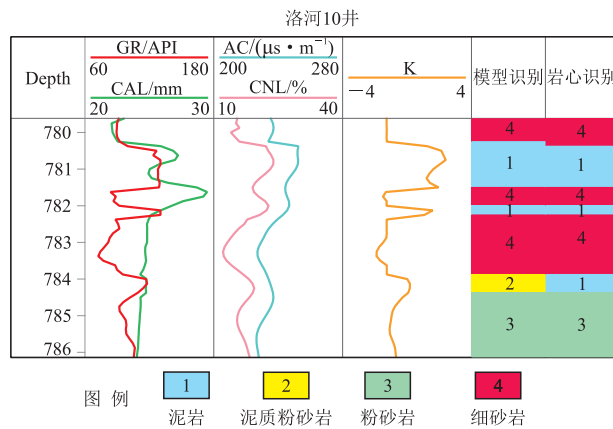


图10 洛河 10 号井岩性识别验证

Fig. 10 Lithological identification and verification of Luohe 10 well

在洛河 7 井(图 11)中,岩心从上到下描述是浅灰色粉砂岩、灰色油斑粉砂岩、深灰色泥岩、浅灰色泥质粉砂岩、深灰色泥岩、浅灰色油斑粉砂岩和浅灰色油斑粉砂岩。

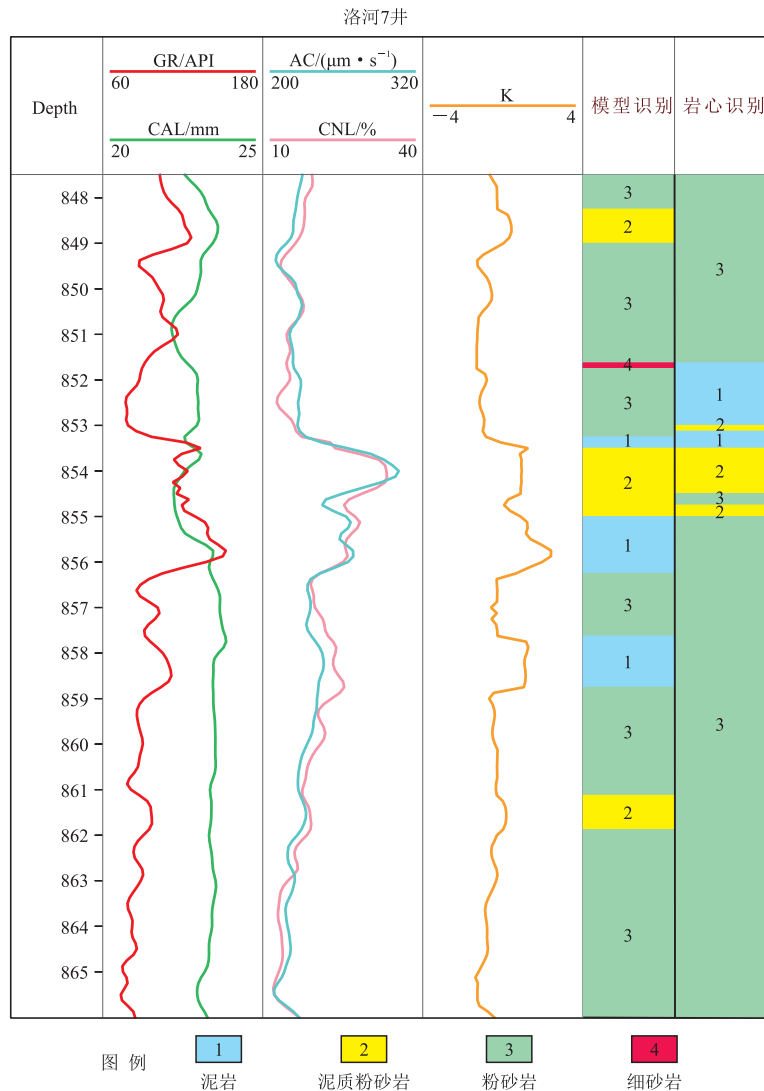


图 11 洛河 7 号井岩性识别验证

Fig. 11 Lithological identification and verification of Luohe 7 well

笔者选择的洛河 10 井和洛河 7 井(图 11)的岩心验证识别结果,表明了该方法是可适用于富县延长组长 6 地层岩性的。从上图中可以发现,测井数据的识别与岩心的识别在总体上能呈现出良好的一致性,除了部分深度。特别是在细砂岩与泥岩的分层效果中比较明显。因为地下环境的复杂多样,且不可避免地受到地下应力和地下温度的影响,影响测井数据的有效性;同时“岩性边界效应”的存在,也会影响测井数据的准确率,所以数据的选择是十分关键的,需要去校正测井曲线。而且因为岩心的破碎情况,也可能导致岩心识别岩性的错误,这也是需要去完善的。

4 结论

1)在识别富县延长组长 6 地层的岩性的实例中,主成分分析能够更好地提取测井数据特征,进行归纳,解决岩性多解性问题,从而进行岩性识别。

2)“岩性边界效应”的存在,影响测井数据的准确率。在我们基于测井数据的提取中,对准岩心是十分重要的,数据不足或不准确会影响岩性的判断,所以校准测井曲线是值得去研究的。

3)在富县延长组的长 6 地层的岩性识别中,基于主成分分析方法建立了一个数学模型,该模型在

岩心和测井对比验证中,呈现了良好的区分性,但是关于岩性分层数学模型的建立有一个更为全面并值得研究的方向,构建的数学模型潜能未充分发挥,需要不断地进行测井数据的选择、训练与验证。

参考文献:

- [1] 王恒,姜亚楠,张欣,等. 基于梯度提升算法的岩性识别方法[J]. 吉林大学学报(地球科学版), 2021, 51(3): 940—950.
WANG H, JIANG Y N, ZHANG X, et al. Lithology identification method based on gradient boosting algorithm[J]. Journal of Jilin University (Earth Science Edition), 2021, 51(3): 940—950. (In Chinese)
- [2] REN Q, ZHANG H B, ZHANG D L, et al. Lithology identification using principal component analysis and particle swarm optimization fuzzy decision tree[J]. Journal of Petroleum Science and Engineering, 2023, 220: 111233.
- [3] REN Q, ZHANG H B, ZHANG D L, et al. A framework of active learning and semi-supervised learning for lithology identification based on improved naive Bayes[J]. Expert Systems with Applications, 2022, 202: 117278.
- [4] DUAN Y J, XIE J, LI B C, et al. Lithology identification and reservoir characteristics of the mixed siliclastic-carbonate rocks of the lower third member of the Shahejie formation in the south of the Laizhouwan Sag, Bohai Bay Basin, China [J]. Carbonates and Evaporites, 2020, 35(2): 55.
- [5] PAREEK J, JACOB J. Data compression and visualization using PCA and T-SNE[M]//Lecture Notes in Networks and Systems. Singapore: Springer Singapore, 2020: 327—337.
- [6] 刘爱疆,左烈,李景景,等. 主成分分析法在碳酸盐岩岩性识别中的应用:以地区寒武系碳酸盐岩储层为例[J]. 石油与天然气地质, 2013, 34(2): 192—196.
LIU A J, ZUO L, LI J J, et al. Application of principal component analysis in carbonate lithology identification: A case study of the Cambrian carbonate reservoir in YH field[J]. Oil & Gas Geology, 2013, 34(2): 192—196. (In Chinese)
- [7] LI W, CHEN T J, SONG X, et al. Reconstruction of critical coalbed methane logs with principal component regression model: A case study[J]. Energy Exploration & Exploitation, 2020, 38(4): 1178—1193.
- [8] REN P F, XU H, TANG D Z, et al. The identification of coal texture in different rank coal reservoirs by using geophysical logging data in northwest Guizhou, China: Investigation by principal component analysis[J]. Fuel, 2018, 230: 258—265.
- [9] 张强,李家金,王毛毛,等. 基于改进主成分分析法的测井曲线岩性分层技术[J]. 吉林大学学报(地球科学版), 2022, 52(4): 1369—1376.
ZHANG Q, LI J J, WANG M M, et al. Logging curve rock layering technology based on improved principal component analysis[J]. Journal of Jilin University (Earth Science Edition), 2022, 52(4): 1369—1376. (In Chinese)
- [10] 王晨,张惠,姚志刚. 鄂尔多斯盆地富县地区延长组长6、长7储层特征及有利区预测[J]. 矿产勘查, 2023, 14(2): 163—173.
WANG C, ZHANG H, YAO Z G. Reservoir characteristics and favorable area prediction of Chang 6 and Chang 7 reservoirs in the Fuxian area of the Ordos Basin[J]. Mineral exploration, 2023, 14(2): 163—173. (In Chinese)
- [11] 党伟,张金川,聂海宽,等. 页岩油微观赋存特征及其主控因素:以鄂尔多斯盆地延安地区延长组7段3亚段陆相页岩为例[J]. 石油学报, 2022, 43(4): 507—523.
DANG W, ZHANG J C, NIE H K, et al. Microscopic occurrence characteristics of shale oil and their main controlling factors: A case study of the 3rd submember continental shale of member 7 of Yanchang formation in Yan'an area, Ordos Basin[J]. Acta Petrologica Sinica, 2022, 43(4): 507—523. (In Chinese)
- [12] 周英男. 鄂尔多斯盆地富县大东沟区延长组烃源岩评价[D]. 西安:西安石油大学, 2020.
ZHOU Y N. Evaluation of source rocks of Yanchang formation in dadonggou district, Fuxian County, Ordos Basin[D]. Xi'an: Xi'an Shiyu University, 2020. (In Chinese)
- [13] 刘化清,袁剑英,李相博,等. 鄂尔多斯盆地延长组湖盆演化及其成因分析[J]. 岩性油气藏, 2007, 19(1): 52—56.
LIU H Q, YUAN J Y, LI X B, et al. Lake basin evolution of Ordos Basin during Middle-Late Triassic and its origin analysis[J]. Lithologic Reservoirs, 2007, 19(1): 52—56. (In Chinese)
- [14] 姚泾利,邓秀芹,赵彦德,等. 鄂尔多斯盆地延长组致密油特征[J]. 石油勘探与开发, 2013, 40(2): 150—158.
YAO J L, DENG X Q, ZHAO Y D, et al. Characteristics of tight oil in Triassic Yanchang formation, Ordos Basin [J]. Petroleum Exploration and Development, 2013, 40(2): 150—158. (In Chinese)

- [15] FU X H, QIN Y, WANG G G X, et al. Evaluation of gas content of coalbed methane reservoirs with the aid of geophysical logging technology[J]. *Fuel*, 2009, 88(11): 2269—2277.
- [16] 游富粮, 柳广弟, 孙明亮, 等. 鄂尔多斯盆地三叠系延长组 7 段高伽马砂岩测井识别及其展布特征[J]. *石油实验地质*, 2023, 45(1): 99—108.
- YOU F L, LIU G D, SUN M L, et al. Logging identification and distribution characteristics of high- γ sandstones in the 7th member of Triassic Yanchang Formation, Ordos Basin[J]. *Petroleum Geology & Experiment*, 2023, 45(1): 99—108. (In Chinese)
- [17] 关叶钦, 徐锦绣, 陆云龙, 等. 利用电成像测井资料定量评价储层分选及渗透率研究[J]. *物探化探计算技术*, 2024, 46(05): 584—590.
- GUAN Y Q, XU J X, LU Y L, et al. Study on quantitative evaluation of reservoir sorting and permeability using Micro-electric imaging logging data[J]. *Computing Techniques For Geophysical And Geochemical Exploration*, 2024, 46(05): 584—590. (In Chinese)
- [18] LEVER J, KRZYWINSKI M, ALTMAN N. *Points of Significance: Principal Component Analysis*[M]. London: Nature Publishing Group, 2017.
- [19] FENG R H, GRANA D, BALLING N. Imputation of missing well log data by random forest and its uncertainty analysis[J]. *Computers & Geosciences*, 2021, 152: 104763.
- [20] ANDREA M, SANDRA V. *Machine Learning*[M]. Amsterdam: Elsevier, 2020.
- [21] 王睿, 李琼, 孙华军, 等. 基于主成分分析的融合方法在断裂识别的应用研究[J]. *物探化探计算技术*, 2021, 43(6): 715—723.
- WANG R, LI Q, SUN H J, et al. Application research of fusion method based on principal component analysis in fracture identification[J]. *Computing Techniques For Geophysical And Geochemical Exploration*, 2021, 43(6): 715—723. (In Chinese)

Lithology identification based on logging data of principal component analysis

——A case study of Chang 6 formation in Yanchang, Fuxian

CHEN Yang¹, WANG Xingjian¹, HE Faqi², QI Rong²

(1. School of Geophysics, Chengdu University of Technology, Chengdu 610059, China;

2. Sinopec North China Oil and Gas Branch, Zhengzhou 450000, China)

Abstract: The distribution of oil and gas reservoirs is closely related to lithology, so identifying and layering lithology can help determine the reservoir distribution of the formation. The coring experiment is , and the core obtained during the drilling process is discontinuous, so there is a problem with the core obtained during the drilling process being limited core data. In order to better determine the lithology and reduce the cost, this study analyzed a total of 99 core points of four lithologies developed more in Chang 6 formation of Yanchang, Fuxian County, northern Shaanxi Province, and their corresponding logging data. Four logging curves, namely natural gamma logging (GR), caliper logging (CAL), acoustic time difference logging (AC), and neutron compensation logging (CNL), were selected to analyze the lithological characteristics of Group 6 in Yanchang, Fuxian. As the particle size increases and the clay content decreases, the GR, CAL, AC, and CNL values tend to decrease, which can be divided. Using the principal component analysis method, the data is dimensionally reduced, features are extracted, and summarised to solve the problem of multiple solutions in lithology. A lithology judgment index is established, which is then applied to multiple drilling wells in the Fuxian area of northern Shaanxi to verify the applicability and effectiveness of the method and the constructed mathematical model. The results indicate that principal component analysis is an effective method for analyzing lithology and improving accuracy, and the lithology of the core is consistent with the lithology identified by logging data.

Keywords: logging data; principal component analysis; lithological identification; Yanchang group leader 6 stratum