

基于MATLAB支持向量机定向拟合在储层孔隙度预测中的应用

徐鹏宇¹, 李响¹, 刘蓉蓉¹, 熊晨皓^{1,2}, 吕正祥^{1,3}

¹成都理工大学能源学院, 四川 成都

²中石化勘探分公司, 四川 成都

³成都理工大学, “油气藏地质及开发工程”国家重点实验室, 四川 成都

Email: 317067647@qq.com, 157145620@qq.com, 971286917@qq.com, 294183271@qq.com

收稿日期: 2020年12月1日; 录用日期: 2020年12月23日; 发布日期: 2020年12月31日

摘要

针对磨溪-高石梯区块灯影组致密碳酸盐岩储层孔隙度难以预测的问题, 本文在使用基于MATLAB支持向量机分类模型的储层类型预测方法的基础上, 总结四种不同储集类型测井电性参数分布规律及主控电性参数, 再根据分布规律划分四个板块以确定回归线走向, 最后采用逐步回归法定向多元回归拟合方程分类预测储层孔隙度。以此方法对致密碳酸盐岩孔隙度进行预测, 不仅预测更加精确, 也能高效地完成预测工作, 同时能方便快速地计算出有效储层的厚度及位置, 提高产油效率。

关键词

分类预测孔隙度, 致密碳酸盐岩, 定向拟合, 磨溪-高石梯

Application of Support Vector Machine Directional Fitting Based on MATLAB in Reservoir Porosity Prediction

Pengyu Xu¹, Xiang Li¹, Rongrong Liu¹, Chenhao Xiong^{1,2}, Zhengxiang Lv^{1,3}

¹College of Energy, Chengdu University of Technology, Chengdu Sichuan

²Sinopec Exploration Company, Chengdu Sichuan

³State Key Laboratory of Oil-Gas Reservoir Geology & Exploitation, Chengdu University of Technology, Chengdu Sichuan

Email: 317067647@qq.com, 157145620@qq.com, 971286917@qq.com, 294183271@qq.com

Received: Dec. 1st, 2020; accepted: Dec. 23rd, 2020; published: Dec. 31st, 2020

Abstract

Aiming at the problem that it is difficult to predict the porosity of tight carbonate reservoir in Moxi-Gaoshiti area, based on the reservoir type prediction method based on MATLAB support vector machine classification model, this paper summarizes the distribution law of logging electrical parameters and major electrical parameters of four different reservoir types, and then divides four plates according to the distribution rules to determine the trend of regression line by using stepwise regression method; directional multiple regression fitting equation is used to classify and predict reservoir porosity. This method can not only predict the porosity of tight carbonate rocks more accurately, but also efficiently complete the prediction work. At the same time, it can calculate the effective reservoir thickness and position conveniently and quickly, and improve the oil production efficiency.

Keywords

Classified Prediction of Porosity, Tight Carbonate Rocks, Directional Fitting, Moxi-Gaoshiti Area

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

储层预测是储层评价研究中的重要问题，而评价储层的好坏的重要岩石物性指标之一为孔隙度。测井解释研究工作之一就是如何利用现有的岩心样本数据预测并解释全井段的地质属性。概率与数理统计强调用样本特性参数去估计总体特性参数，常用的方法有矩估计、极大似然函数及最小二乘法。但由于试验数据存在一定不可避免误差，如：量具仪器超过额定检校周期、人工数据传递失误等。针对该类情况，在二十世纪初，研究概率与数理统计的相关数学家们总结出了采用置信区间消除残差等方法来控制误差确保估计精度。然而现实存在的误差远超理论误差，为了应对致密碳酸盐岩储层这种非均质性极强的地层，目前测井解释模型已经运用了多种非线性方法，如分段最小二乘法[1]、人工神经网络法[2]、数值模拟法[3]等。近几年来，许多研究者在支持向量机对于储层的孔隙度预测方面取得了较大的研究成果。乐友喜等[4]在胜利油田某区块应用支持向量机法，以地震波波幅作为输入向量，进行了砂体孔隙度和含油性预测，预测结果与已知结果吻合较好。张尔华等[5]在地震属性组合的基础上确定支持向量机模型，进而预测火山岩储层厚，并在实际应用中取得了良好效果。熊晨皓等[6]利用支持向量机方法对珠-坳陷恩平组致密砂岩储层某井进行砂体孔隙度预测，证明了采用粒子寻优法下的支持向量机对实现少量样本分类具有显著的准确性。

由于岩心的取芯和测试费用较高，岩心测量数据一般较少且不连续，针对碳酸盐岩储层复杂的岩石架构下难以寻得合适的方程模型对孔隙度进行预测，故在支持向量机分类的基础上，根据岩心柱结合测井成像图片确定储集类型并取样，其后进行样本训练及全井段预测，确定沿深度方向岩层的各储集类型。通过对比分析各类储层电性特征，根据不同岩石的储集类型定向建立多元回归方程对孔隙度进行预测。

川中北部磨溪-高石梯地区是四川盆地天然气的主要产区，其主力储集层为震旦系灯影组和寒武系龙王庙组，其中龙王庙组厚度受古地貌影响显著，呈北西偏薄、南东偏厚的特征，岩石类型主要有砂屑

白云岩、残余砂屑白云岩以及细-中晶白云岩。这些岩石的裂缝、孔洞比较发育，具有良好的储层发育条件，为油气的形成奠定了基础。震旦系灯影组的沉积厚度较大，分布范围较广，主要发育浅灰色微-粉晶白云岩、藻白云岩，其中灯二段和灯四段均为良好的含气层系。储层岩石结构主要以中白云岩为主，孔隙度分布在 10%~20%，储层间的非均质性较强[7]。

2. 原理分析

2.1. 基于 MATLAB 支持向量机分类(LIBSVM)应用

MATLAB 与 Mathematica、Maple 并称为三大数学软件。它在数学类科技应用软件中在数值计算方面首屈一指。MATLAB 可以进行矩阵运算、绘制函数和数据、实现算法、创建用户界面、连接其他编程语言的程序等。用 MATLAB 来解算问题要比用 C, FORTRAN 等语言完成相同的事情简捷得多。同时也可用 MATLAB 软件便利的调用各种工具箱，如：神经网络工具箱、曲线拟合工具箱及支持向量机等[8]。

如果我们在低维空间里找不到一个线性分类面把样本分开，支持向量机就为我们提供了一个思路：将数据从低维空间映射到高维空间后，就很可能使得这堆数据线性可分。比如说，我们要在猫科动物这个特征很局限的“低维空间”里去分猫和老虎，是比较困难的，因为他们很多特征比较相近。但是，如果我们有了更多的参考依据，从生物界的视角，即一个“高维空间”再去区分猫和老虎，我们就有了更多的理由来做出科学的辨别。

2.1.1. 调用 MATLAB 支持向量机分类器(LIBSVM)相关函数

可先通过安装 MATLAB 支持向量机分类器(LIBSVM)安装包[9]，对样本进行训练及其后的预测需调用下列函数：

$$\text{model} = \text{svmtrain}(\text{train_label}, \text{train_matrix}, \text{'libsvm_options'}) \quad (1)$$

式中，train_label——训练集样本标签(储集类型)；

train_matrix——训练集样本；

model——待训练模型；

'libsvm_options'——其他参数选项。

$$[\text{predict_label}, \text{accuracy}] = \text{svmpredict}(\text{test_label}, \text{test_matrix}, \text{model}) \quad (2)$$

式中，test_label——测试集样本标签(储集类型)；

test_matrix——测试集样本；

model——已训练好的模型；

predict_label——预测后样本标签；

accuracy——预测准确度。

2.2. 基于 MATLAB 定向回归拟合原理

在统计学中，回归分析(regression analysis) [10]指的是确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法。回归分析按照涉及的变量的多少，分为一元回归和多元回归分析。按照自变量和因变量之间的关系类型，又可分为线性回归分析和非线性回归分析。广泛的应用于各项研究。针对有时样本过于离散，无论是采用单元线性回归还是多元线性回归，其决定系数 R^2 值依旧不理想，针对该类现象，通常的做法是删除残差较大的样本点，重新拟合回归方程，进而获取较大的 R^2 值。由于该类办法的随机性较大的削减了样本真实性，造成样本失真，进而影响了预测效果。基于前人的研究上，经过大量的试验与数据论证，开发定向多元回归分析处理该类问题。

2.2.1. 定向多元回归分析原理

最小二乘法定义不同的样本集具有唯一且单一的回归方程使其决定系数 R^2 值达到最大, 但往往“最优方程”却不能得到“最优解”, 其原因多重各样无法臆测。需要人工反向修正以修订正确的拟合方向, 得到更合理的预测结果。对于样本集合:

$$X = \{x_1, x_2, \dots, x_i\}, (i = 1, 2, \dots, n) \quad (3)$$

对于各样本的不同属性参数可用矩阵 P 表示:

$$P = [p_1, p_2, \dots, p_i], (i = 1, 2, \dots, n) \quad (4)$$

选取两个主控因素进行二维平面散点图分析, 可调用 MATLAB 绘图 plot 函数[11]进行绘制:

$$\text{plot}(x, y, 'r.', 'MarkerSize', 15) \quad (5)$$

通过 MATLAB ginput()函数[12]可以方便快捷的在图上选取两个坐标点并获取坐标点数据, 再根据两数据点坐标确定直线方程:

$$f(x) = kx + b \quad (6)$$

根据直线方程删除偏离该直线附近的样本点, 最后获取定向回归目标样本集其代码如下:

$$\left\{ \begin{array}{l} \text{for } i = \text{size}(y_i, 1) : -1 : 1 \% \text{删除偏离样本点} \\ \quad \text{if } \text{abs}(y(i, 1) - y_i(i, 1)) \geq \text{rint} \\ \quad \quad y(i, :) = []; \% \text{删除偏离y坐标} \\ \quad \quad x(i, :) = []; \% \text{删除偏离x坐标} \\ \quad \quad SJ(i, :) = []; \% \text{删除样本数据} \\ \quad \text{end} \\ \text{end} \end{array} \right. \quad (7)$$

式中, rint——为残差容许区间, 经过大量试验, 取 1~1.5 为最佳。定向拟合如图 1 所示:

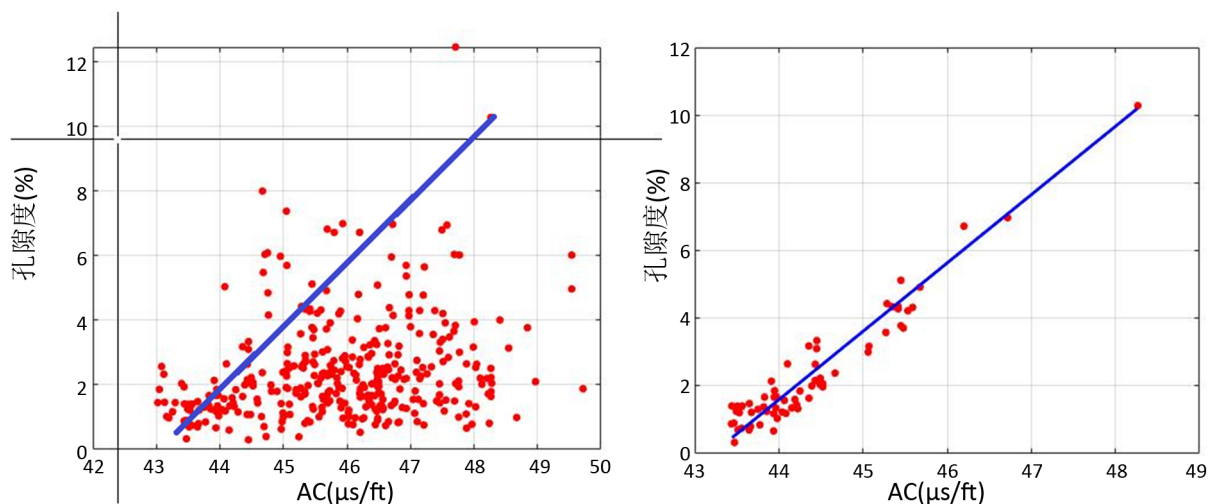


Figure 1. Process chart of directional fitting

图 1. 定向拟合过程图

最后根据定向拟合获取的样本进行多元线性回归拟合, 最终确定最优回归方程[13]。

$$f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \dots + \beta_i x_i, (i = 1, 2, 3, \dots, n) \quad (8)$$

式中, $\beta_i, (i=1,2,3,\dots,n)$ 为回归系数。

3. 孔隙度预测过程

3.1. 数据标准化

由于少部分井测井原始曲线获取条件的差异, 不同井测井仪器型号不统一, 仪器操作人员的熟练程度不同以及井下环境不一样等因素, 使得各井相同的地层、相同的岩性在同种测井曲线上的响应值具有一定的系统误差, 这些差异会直接影响到测井储层参数的计算精度, 给储量参数的准确计算带来一定的不确定性, 要消除这些非地质因素的影响, 使测井资料更真实地反映地层的实际情况, 有必要对测井数据进行标准化, 避免在之后的预测工作中陷入“失之毫厘, 差之千里”的困境, 针对不同井测井曲线深度不一致等问题, 采用线性内插归一化的方式矫正。本文通过 MATLAB 2018a 软件编程实现对 GR、DT、DEN、CNL、RT 等测井曲线的批量标准化及调用 interp1 函数[14]进行深度值线性内插预处理。

3.2. 样本分类与预测

3.2.1. 储层的分类

本文结合岩心样本, 选择具有取心资料的测试进行物性分析, 结合成像测井[15]分别确定出这些典型层位的物性特征, 在此基础上总结出了各类产层的孔隙度大小及储集能力。根据上述方法, 可将 A 地区灯二段储层按照储集物性的好坏依次分为 I、II、III、IV 四类储层, 各类储层的划分标准如表 1 所示。

Table 1. Reservoir classification and evaluation table of Dengger formation in area A

表 1. A 地区灯二段储层分类评价表

分类	储集空间	成像测井响应	储集类型	储层评价
I	溶蚀孔洞为主要储集空间, 裂缝为主要渗流通道	串珠状较大的暗色斑点成群、成带发育, 非定向分布高导裂缝发育。	裂缝-孔洞型为主, 其次是孔洞型	好
II	溶蚀孔洞为主要储集空间	暗色斑点成层性分布和成群分布。	孔洞型为主	较好
III	孔隙为主要储集空间, 裂缝为主要渗流通道	暗色小斑点星点状分布, 水平层理缝为主。	裂缝-孔隙型为主	差
IV	①孔隙为主要储集空间, 裂缝不发育 ②孔隙不发育, 裂缝既是储集空间也是渗流通道	亮色为主, 偶见小黑色斑点。	孔隙型、裂缝型为主	非

其中 I 类为在自然开采条件下能够获得较好产能的储层; II 类为自然开采条件下基本能够获得工业产能的较好储层; III 类储层在自然开采条件下不能达到工业的差储层; IV 类储层为含油性很差的非储层。

3.2.2. 样本选取及储集类型预测

将磨溪-高石梯 13 口重点井灯二段的 960 个实测岩心样本依据照表 1 分为四类。各类样本分布较均匀, 随机选取其中的 70% 作为训练样本, 剩下的 30% 样本作为测试样本, 模型样本与训练样本的分配如表 2 所示。

Table 2. Sample distribution of Dengger formation in area A

表 2. A 地区灯二段样本分配表

储层分级	I	II	III	IV
训练样本数	28	137	247	227
测试样本数	2	9	124	186

所有样本数据根据测井响应识别的原则, 选取代表地层特征的 3 个参数, 其中 GR 为伽马, DT 为声

波, CNL 为中子密度。本次建模使用 MATLAB 2018a 支持向量机(LIBSVM)用训练后的模型对 13 口重点井进行了储集类型预测。

3.2.3. 预测结果分析

最终训练报告显示, 样本集预测正确率为 96.36%, 测试集预测正确率为 92.33%。根据预测结果, 通过对储集类型电性参数的分析发现不同储集类型对 DT 声波时差有显著的相关性。其结果如图 2 所示。

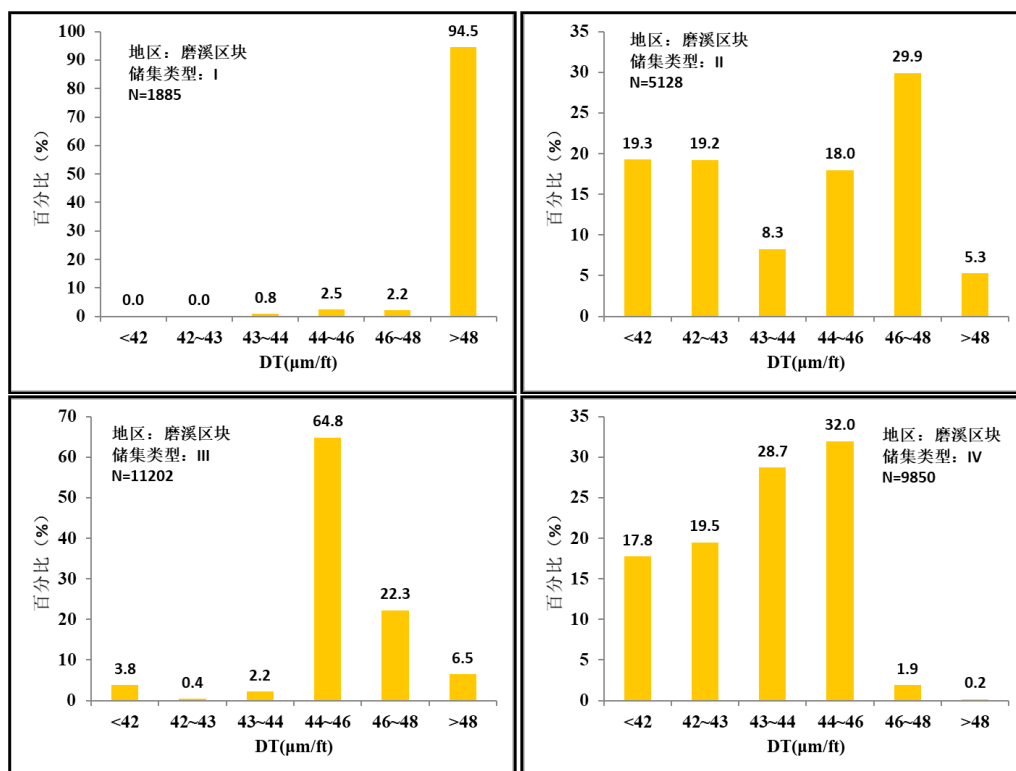


Figure 2. Distribution of DT values of reservoir types in Moxi block

图 2. 磨溪区块储集类型 DT 值分布图

通过对比各类样本的储集类型声波值分布区间, 可以明显的划分出各类储层的差异及特征。通过整理得到结论如表 3 所示:

Table 3. Characteristics of acoustic value of each reservoir type in the Dengger formation in Moxi area

表 3. 磨溪地区灯二段各储集类型声波值分布特点

储层分级	I	II	III	IV
DT 值分布区间	>48	46~48	44~46	<44

3.3. 定向拟合及孔隙度值预测

在储层类型划分的基础上, 确定以声波值为主控因素的前提下, 采用逐步回归法[16]对孔隙度进行预测, 对不同储集类型分别建立不同回归方程定向拟合, 并预测孔隙度。

3.3.1. 定向拟合

选取磨溪区块 A、B、C 井共计样本 362 个, 孔隙度值分布在 5%~13%, 声波值分布在 43~51 μm/ft,

并根据前叙述结论各储集类型声波值的划分特点，分区块定向拟合(图 3)并建立回归方程。

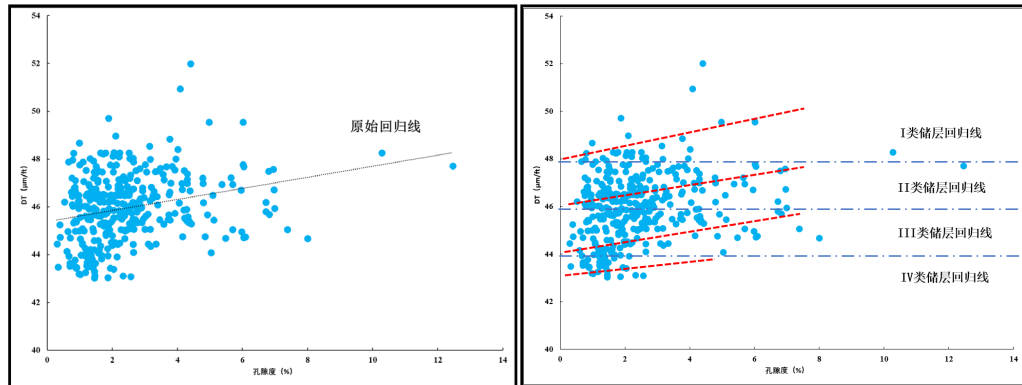


Figure 3. Chart of directional fitting of reservoir type
图 3. 储集类型定向拟合效果图

3.3.2. 回归方程的确定

在定向拟合的基础上，新确立 4 类样本，其中 I 类样本 13 个，II 类样本 14 个，III 类样本 154 个及 IV 类样本 194 个，在此基础上进行逐步回归法建立回归方程对孔隙度进行预测，各类回归方程如下：

$$\Phi_{\text{I类孔隙度}} = -93.9908 + 2.2291 \times DT + 0.052796 \times CNL, (R^2 = 0.8729) \quad (9)$$

$$\Phi_{\text{II类孔隙度}} = -19.4736 + 0.57656 \times DT + 0.1964 \times CNL, (R^2 = 0.7045) \quad (10)$$

$$\Phi_{\text{III类孔隙度}} = -12.6027 + 0.33216 \times DT + 0.061172 \times CNL, (R^2 = 0.6839) \quad (11)$$

$$\Phi_{\text{IV类孔隙度}} = -5.4316 + 0.14886 \times DT + 0.054769 \times CNL, (R^2 = 0.8523) \quad (12)$$

式中， Φ 为孔隙度；

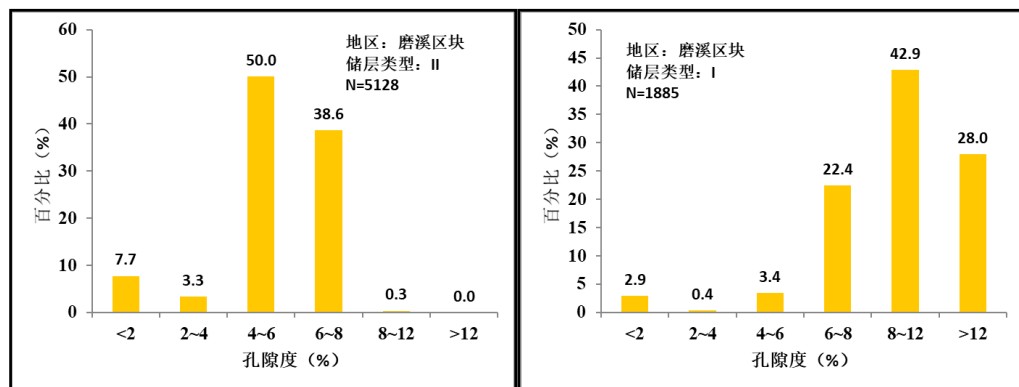
DT ($\mu\text{m}/\text{ft}$) 为声波时差；

CNL (%) 为补偿中子。

3.3.3. 预测结果

在分类对该区段全井储层孔隙度进行预测后，通过分析孔隙度值的分布区间，作如图 4 直方图。

从直方图可得出，I 类储层孔隙度主要分布在 8% 以上，II 类储层主要分布在 4%~8%，III 类储层主要分布在 2%~4%，IV 类储层主要分布在 2% 以下，预测效果良好。



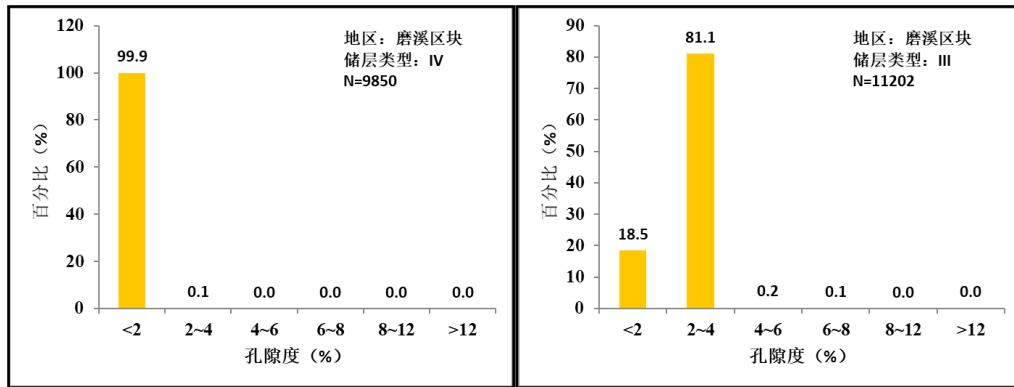


Figure 4. Predicted porosity distribution of each reservoir type
图 4. 各储集类型预测孔隙度分布图

3.4. 预测效果

根据 4 类回归方程分别对 4 类储集类型的孔隙度进行预测, 并将 B 井灯影组在 Resform 软件中成图, 可以得到储层类型的测井解释图如图 5 所示。

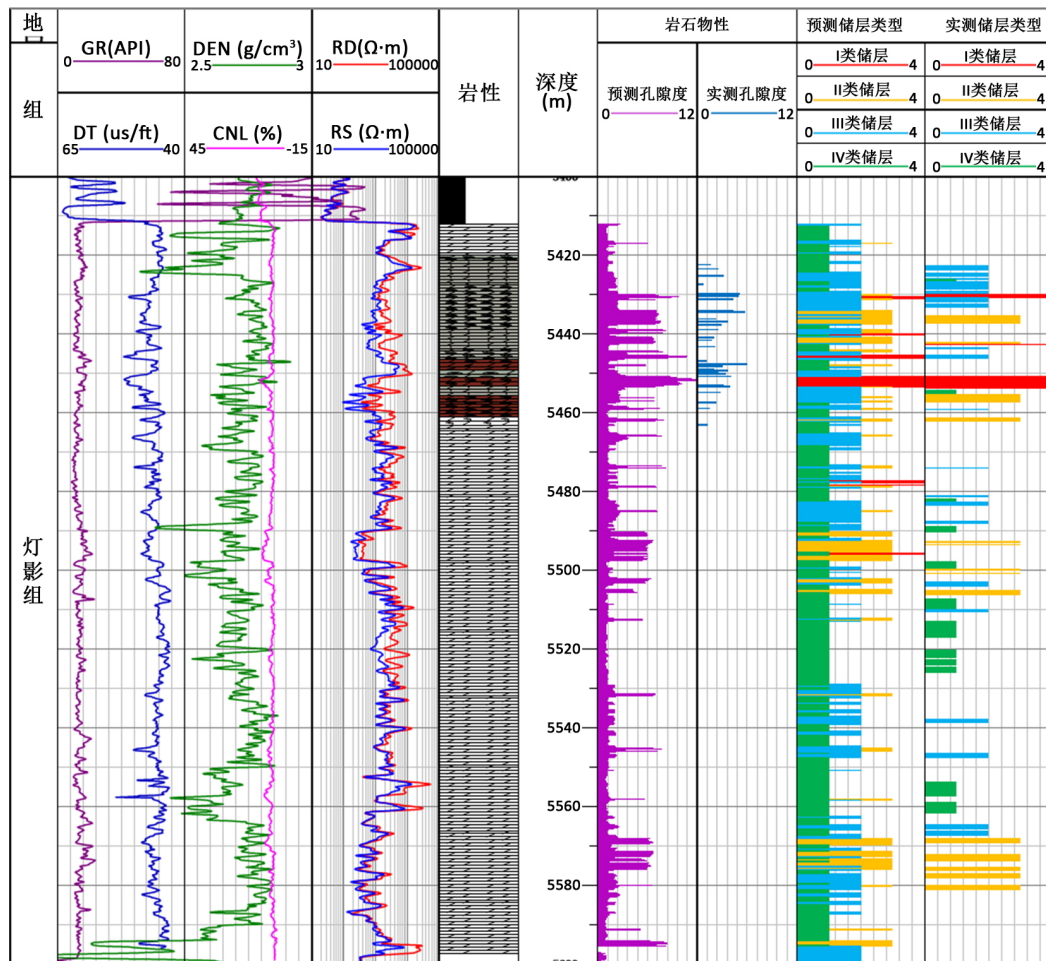


Figure 5. Reservoir type prediction map of Denger formation in well B
图 5. B 井灯二段储层类型预测图

4. 结论

- 1) 支持向量机具有显著的分类能力, 在样本较少的情况下能够精确度较高地进行预测。
- 2) 定向多元线性回归在应对碳酸盐岩储层等非均质性极强的预测工作较实用, 同时有效地解决了样本数据人工误差等问题。
- 3) 单只依照声波曲线值分布特征无法明确地划分 II 类储层与 III 类储层, 还需进一步分析其他电性参数对划分 II 类储层与 III 类储层的贡献率。
- 4) 该预测方法建立在 MATLAB 2018a 软件编程基础上, 结合 Excel VBA 进行统计分析绘图, 全程自动化实现, 将 MATLAB 数据处理功能与 Excel 的绘图功能相结合, 实现统计分析绘图的自动化和批量化, 将烦琐、机械的绘图工作解放出来, 提高工作效率和质量。同时该程序操作简单灵活, 实用性强。

参考文献

- [1] 田垅, 刘宗田. 最小二乘法分段直线拟合[J]. 计算机科学, 2012, 39(B06): 482-484.
- [2] 汪忠德, 王新海, 张汶, 等. BP 神经网络法(ANNS)在塔中地区碳酸盐岩测井评价中的应用[J]. 石油天然气学报, 2008(2): 25+279-281.
- [3] 龙建东, 徐安娜, 等. 缝洞型碳酸盐岩内部局域振幅横向差异性储层预测法[J]. 石油地球物理勘探, 2003, 38(1): 58-61.
- [4] 乐友喜, 袁全社. 支持向量机方法在储层预测中的应用[J]. 石油物探, 2005, 44(4): 388-392.
- [5] 张尔华, 关晓巍, 张元高. 支持向量机模型再火山岩储层预测中的应用——以徐家围子断陷徐东坡带为例[J]. 地球物理学报, 2011, 54(2): 428-432.
- [6] 熊晨皓, 赵福, 覃斌传, 等. 粒子群寻优支持向量机在储层类型预测中的应用[J]. 地球科学前沿, 2020, 10(1): 9.
- [7] 刘冷杉, 胡明毅, 高达, 等. 四川磨溪-高石梯地区龙王庙组层序划分及储层预测[J]. 大庆石油地质与开发, 2016(5): 42-47.
- [8] 杨建强, 罗先香. MATLAB 软件工具箱简介[J]. 水科学进展, 2001, 12(2): 237-242.
- [9] 郭小荟, 马小平. 基于 Matlab 的支持向量机工具箱[J]. 计算机应用与软件, 2007(12): 57-59.
- [10] 冯士雍. 回归分析方法[M]. 北京: 科学出版社, 1974.
- [11] 周明. MATLAB 图形技术: 绘图及图形用户接口[M]. 西安: 西北工业大学出版社, 1999.
- [12] 欧阳联华, 王家林. 一种在等值线图上任意截取剖面的 Matlab 语言算法[J]. 物探化探计算技术, 2003(3): 273-276.
- [13] 范雯. 逐步回归分析方法在储层参数预测中的应用[J]. 西安科技大学学报, 2014, 34(3): 350-355.
- [14] 赵庆乐, 张世红, 王婷婷, 等. 利用 Matlab 函数识别沉积物中的米兰柯维奇旋回信号[J]. 吉林大学学报(地球科学版), 2010, 40(5): 1217-1220.
- [15] 陈义祥, 任小峰, 牟瑜, 等. 基于电成像测井的致密碳酸盐岩储层有效性评价方法[J]. 测井技术, 2020, 2(44): 50-52.
- [16] 杨伍梅. MATLAB 在回归分析中的应用[J]. 高等继续教育学报, 2008, 21(5): 33-34.