

基于支持向量机(SVM)的不同咸化程度烃源岩 总有机碳含量预测模型

初勇志^{1,2}, 刘成林^{1,2}, 太万雪^{1,2}, 阳宏^{1,2}

1. 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;

2. 中国石油大学(北京) 地球科学学院, 北京 102249

摘要: 总有机碳含量(TOC)作为评价烃源岩有机质丰度的重要参数,其精确预测对油气勘探开发具有重要意义。目前总有机碳含量预测以 $\Delta\log R$ 方法、多元回归分析等基于统计分析的方法为主,存在泛化能力弱、主观性强等问题。机器学习方法的引入,可有效解决这类非稳定性、非线性、高复杂性的问题,但当下的研究仍停留在方法的比较与选取层面,没有对优良模型进行深入分析并检验其适用性。采用应用效果更好的支持向量机模型进行总有机碳含量预测,选取渤海湾盆地渤中凹陷古近系东营组淡水湖相和柴达木盆地西部狮子沟地区古近系咸化湖相烃源岩作为研究对象,对模型的效果进行检验与对比。通过相关性和 XGBoost 特征重要性分析,选定声波时差(DT)、体积密度(DEN)、自然电位(SP)、自然伽马(GR)、深度等作为输入层,以总有机碳含量作为输出层,确立 SVM 烃源岩总有机碳含量预测模型。研究表明,模型在应用至差异较大的沉积环境时具有很强的泛化能力以适应不同地区的地质特征;测井曲线对于烃源岩有机质丰度的敏感性由于沉积环境存在差异而有所区别。该模型在渤海湾盆地淡水湖相区域的应用中相关性更高,误差更小。

关键词: 总有机碳含量预测;测井资料;支持向量机;渤海湾盆地;柴达木盆地

中图分类号: TE122.1

文献标识码: A

Prediction model of TOC contents in source rocks with different salinity degrees based on Support Vector Machine (SVM)

CHU Yongzhi^{1,2}, LIU Chenglin^{1,2}, TAI Wanxue^{1,2}, YANG Hong^{1,2}

1. State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum (Beijing), Beijing 102249, China;

2. College of Geosciences, China University of Petroleum (Beijing), Beijing 102249, China

Abstract: The total organic carbon (TOC) content is an important parameter for the evaluation of abundance of organic matter in source rocks, and its predicting accuracy is of great significance to oil and gas exploration and development. At present, TOC prediction is mainly based on statistical analysis methods such as $\Delta\log R$ method and multiple regression analysis, problems such as weak generalization ability and strong subjectivity exist. The introduction of machine learning methods can effectively solve these problems of instability, nonlinearity, and high complexity. However, current research remains at the level of method comparison and selection with no indepth analysis of good models and their applicability. In this paper, a Support Vector Machine (SVM) model with better application effects was used to predict TOC contents of source rocks with different salinity degrees. As source rocks of freshwater lacustrine facies, the Paleogene Dongying Formation in the Bozhong Sag of Bohai Bay Basin and Paleogene source rocks in the Shizigou area of the western Qaidam Basin as saline lacustrine facies source rocks were selected to test and compare the effectiveness of the model. Through correlation analysis and XGBoost feature importance analysis, the logging sonic differential time (DT), volume density (DEN), spontaneous potential (SP), Gamma ray (GR) and depth were selected as the input layer, while the TOC was used as the output layer to establish a TOC prediction model based on SVM. Results show a strong generalization ability when applied to different sedimentary environments. It can adapt to the geological characteristics of different regions.

收稿日期: 2021-12-12; 修订日期: 2022-07-01。

作者简介: 初勇志(2000—), 男, 从事油气成藏和资源评价研究。E-mail: chuyongzhi18@163.com。

通信作者: 刘成林(1970—), 男, 博士, 教授, 博士生导师, 从事油气资源评价与非常规油气地质研究。E-mail: liucl@cup.edu.cn。

基金项目: 国家自然科学基金面上项目“咸化湖盆条件下盐类对地层超压的作用机制研究”(41872127)资助。

The sensitivity of logging curves to the abundance of organic matter in source rocks varies in different sedimentary environments, which makes the model more relevant when applying to the fresh water lacustrine facies area in the Bohai Bay Basin.

Key words: TOC prediction; logging data; Support Vector Machine; Bohai Bay Basin; Qaidam Basin

烃源岩是含油气盆地油气勘探中的基础研究要素,而总有机碳含量(TOC)是其有机质丰度评价的重要指标。自 20 世纪 80 年代测井曲线与有机质丰度的关系被发现后^[1],测井数据被广泛应用于总有机碳含量预测中。现今主要采用 $\Delta\log R$ 、多元回归分析、地震多属性反演等方法进行总有机碳含量预测^[2-3],应用广泛且效果较为理想,但由于其原理大多基于统计分析,往往一种方法或者优化方案仅能应用至一个地区,难以保证其外推到其他地区的准确性。 $\Delta\log R$ 方法是将声波时差曲线和电阻率曲线在非烃源岩层段重合作为总有机碳含量为 0 的基线值,通过 2 条曲线的差值反映有机质丰度^[4];多元回归分析方法是通过建立多个自变量与因变量总有机碳含量之间的数学关系进行预测^[5]。随着人工智能的快速发展,相关学者将大量机器学习方法应用至烃源岩总有机碳含量预测中。赵兴齐等^[6]采用 BP 神经网络法,对东海陆架盆地浙东拗陷西湖拗陷的烃源岩总有机碳含量进行预测;王惠君等^[7]采用卷积神经网络,对鄂尔多斯盆地杭锦旗地区上古生界泥质烃源岩总有机碳含量进行预测;石创等^[8]通过地化—测井—地震联合优选地震属性参数,采用概率神经网络,对珠江口盆地陆丰南区进行预测,以解决海上预测中的部分缺陷和问题;张成龙等^[9]将基于支持向量机(Support Vector Machine, SVM)的模型应用于鄂尔多斯盆地盐池地区,可识别出局部富有机质层。大量研究表明,机器学习方法较常规方法有更高的准确性和适用性。但目前的研究大多停留在与常规方法的比较,即不同方法在单一地区的适用性,没有对应用效果更好的模型在不同盆地的不同环境中进行泛化能力检验。本文基于支持向量机算法,建立烃源岩总有机碳含量预测模型,根据咸化程度选取典型淡水湖相和咸水湖相沉积区域进行预测效果对比,即渤海湾盆地渤中凹陷西南洼和柴达木盆地西部狮子沟地区,以期为预测模型的外推能力检验和烃源岩沉积环境影响效果分析提供一定参考。

1 地质概况

渤海湾盆地是华北地台东部重要的含油气盆

地,其已探明石油储量远大于天然气储量^[10-11]。渤中凹陷是盆地中最大的富生烃凹陷,位于渤海湾盆地渤中拗陷的主体深凹部位,面积约 8 660 km²,夹于石臼坨凸起、沙垒田凸起、渤南低凸起和渤东低凸起之间^[12-17](图 1)。该凹陷自渐新世以来成为渤海湾盆地的沉积和沉降中心,发育古近系沙河街组三段、一段和东营组三段优质烃源岩,总厚度介于 500~2 500 m,有机质类型以 II₁型为主,有机质丰度高,总有机碳含量普遍高于 1%,大部分为超过 2%的优质烃源岩^[18-19]。本文选取东营组作为主要研究对象之一,该烃源岩整体盐度非常低,是典型的淡水—微咸水的沉积环境^[20]。

柴达木盆地是在前侏罗纪柴达木地块基础上发育而成,由于地理位置位于南部祁连山、东部昆仑山脉和阿尔金山之间,盆地干旱少雨,形成了以新生界为主的高原内陆咸化盆地^[21-24]。狮子沟地区处于柴西地区古近系—新近系生油凹陷中(图 2),该生油凹陷油气资源量远远超过东部富油气凹陷标准,有效烃源岩占同期湖盆面积的比例也超过 50%,岩性以泥岩、泥灰岩为主,与淡水湖相烃源岩的明显区别在于总有机碳含量值与生烃潜量的差异性^[25]。有资料显示,该区烃源岩 TOC 大于 0.6%时即为优质烃源岩,在低熟阶段大量生烃,当 $R_o < 0.8\%$ 时即可达到液态烃生烃高峰;其最大液态烃产率可达 450~700 mg/g,是中国其他淡水湖相烃源岩的 1.15~4.58 倍,具有成熟门限低、生烃窗口宽、排烃时间长、烃转化率高等特点^[26]。本文选取该区下干柴沟组作为另一研究对象。

2 测井响应特征

测井曲线可通过对有机质的敏感程度预测烃源岩总有机碳含量,研究表明,自然伽马(GR)、电阻率(RT)、声波时差(DT)、密度(DEN)等曲线均有较好的效果。通常情况下,测井异常可以反映有机质含量,异常值越大对应烃源岩总有机碳含量越高。本文选定的研究对象分别为典型的淡水湖相和咸水湖相烃源岩,渤中凹陷的烃源岩在测井响应中具有高声波时差、高电阻率、高自然伽马和低密度等特征^[27],而柴西狮子沟地区则不具有常规优质烃源岩的测井特征。本文对 2 个

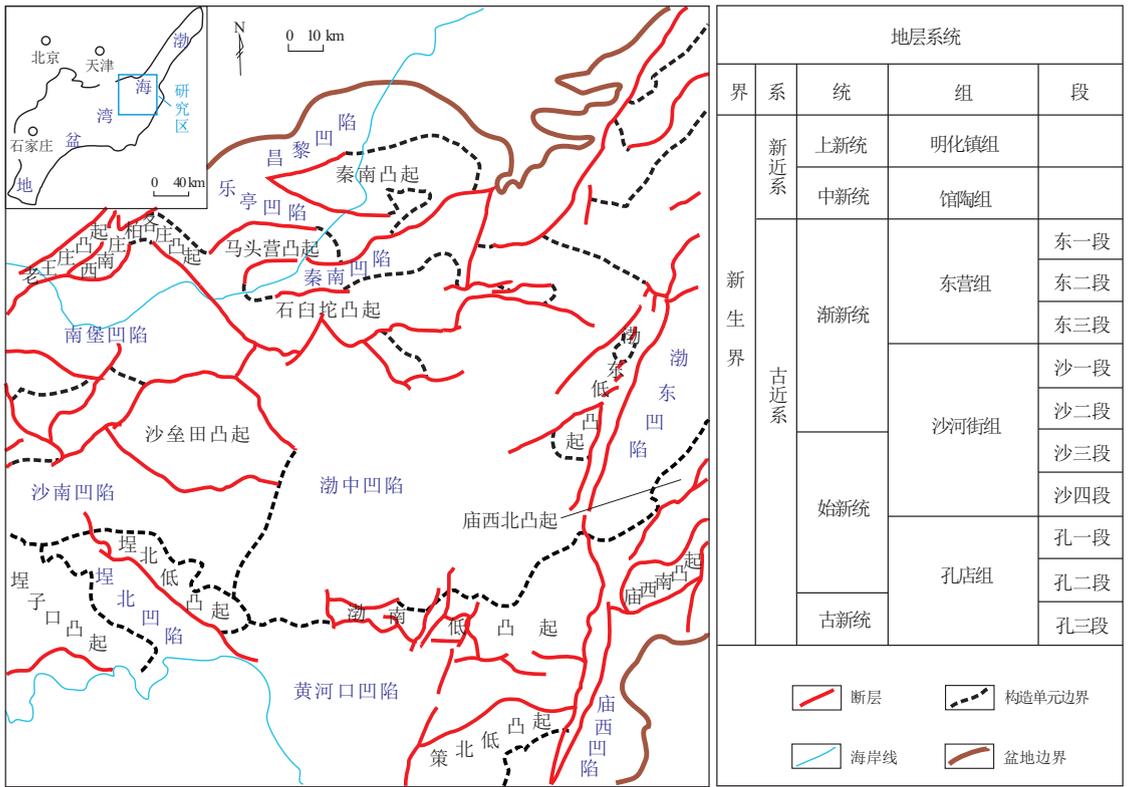


图1 渤海湾盆地渤中凹陷构造位置及新生界地层划分
据文献[11],有修改。

Fig.1 Tectonic location and stratigraphic division of Cenozoic, Bozhong Sag, Bohai Bay Basin

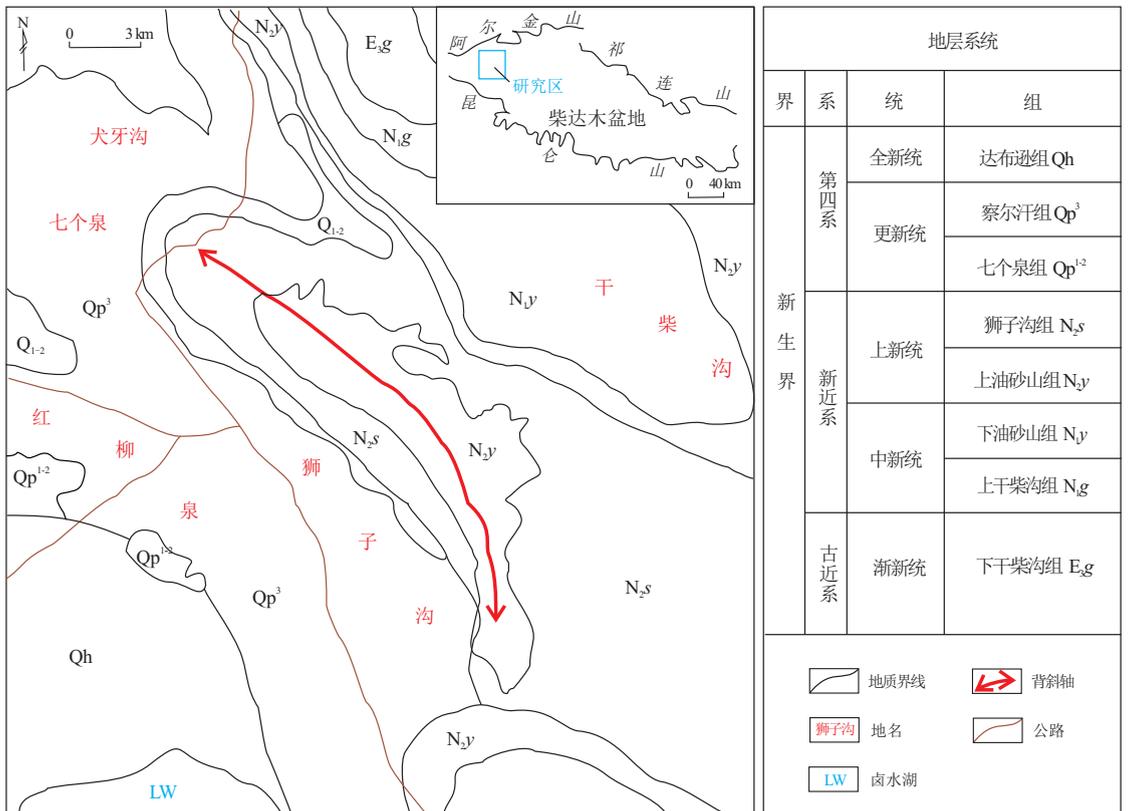


图2 柴达木盆地狮子沟背斜构造位置及新生界地层分布
据文献[21],有修改。

Fig.2 Distribution of Cenozoic in Shizigou anticline and its tectonic location, Qaidam Basin

地区进行测井曲线和实测总有机碳含量相关性分析(图 3),选取渤海湾盆地渤中凹陷 BZ13-2-4 井和柴达木盆地狮子沟地区狮 20 井的 40 个测井及实测总有机碳含量数据进行对比分析。总有机碳含量实测的具体实验方法参照国家标准《沉积岩中总有机碳的测定:GB/T 19145—2003》,将新鲜岩样磨碎,经稀盐酸和去离子水反复冲洗,并在烘干箱干燥处理后,利用在 CS-230 碳硫分析仪中高温下燃烧岩样生成的 CO_2 量换算成碳元素含量,以测定出有机碳含量^[27]。

可以看出,在渤中凹陷,DT 与总有机碳含量相关性最高,决定系数 R^2 为 0.676;其次为 DEN, R^2 为 0.654 7;RD、GR 和 SP 与总有机碳含量的相关性较低(图 3a-e)。在狮子沟地区,DT 与总有机碳含量的相关性仍为最高, R^2 为 0.493 3,其余由高到低依次为 CAL、GR 和 RT(图 3f-i)。对比两地分析可知,由于测井仪器对于酪根的敏感性差异影响着测井响应与有机质丰度的关系^[28],渤中凹陷的相关性显著高于狮子沟地区,其测井参数与实测总有机碳含量的决定系数普遍高于 0.3;电阻率、自

然伽马测井值在渤中凹陷与总有机碳含量相关性较强,而在狮子沟地区很弱;两地的声波时差测井值均与总有机碳含量具备最高的相关性。由于有机质的高声波传播时间,DT 会随总有机碳含量的增大而增大,呈正相关。

3 烃源岩总有机碳含量预测模型

3.1 支持向量机理论

支持向量机在解决复杂的回归和分类问题上有着广泛的应用,其原理可理解为将线性不可分的数据扩展到多维空间中运用超平面进行划分,通过找到最小的结构化风险来增强模型泛化能力,从而实现在统计样本量较少的情况下也可以得到有效统计规律的目的^[29-30],对于解决小样本、非线性的测井相关问题有着很好的应用^[31-33]。

支持向量机是基于统计学习理论的一种机器学习方法,通过给定的训练样本 (x_i, y_i) 构造一个目标函数 $f(x)$,使其与 y 尽可能接近,其中 x_i 为输入向量, y_i 为输出向量。在选取最优回归超平面的过程中需引入非线性映射 $\varphi(x)$:

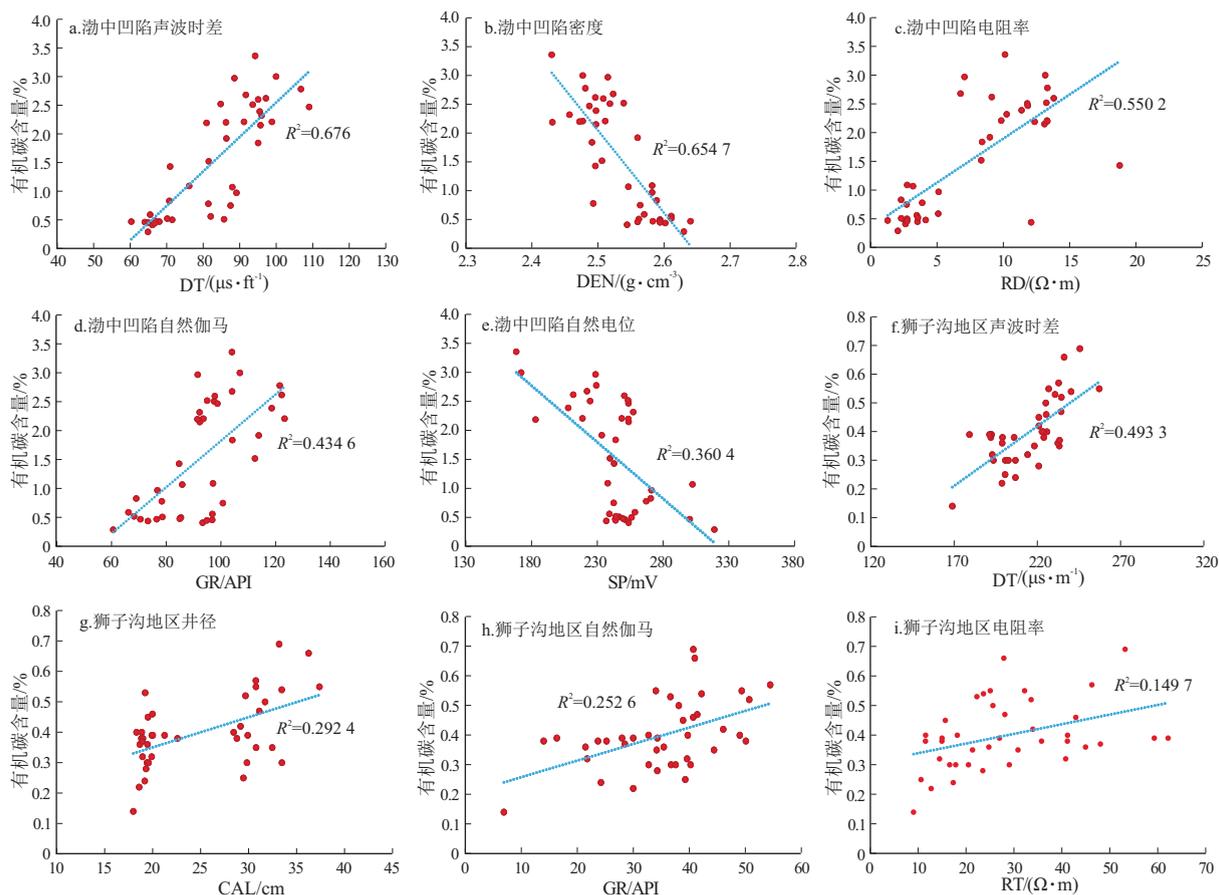


图 3 渤海湾盆地渤中凹陷和柴达木盆地狮子沟地区实测总有机碳含量与测井参数相关图

Fig.3 Measured TOC contents and logging parameters of samples from Bohong Sag of Bohai Bay Basin and Shizigou area of Qaidam Basin

$$f(x) = \omega^T \varphi(x) + b \quad (1)$$

式中: ω 为权重系数; b 为偏差量。

通过引入正则化常数 C 和松弛变量 β 两个参数,可以在一定程度上克服无法完成严格分类、过拟合等问题,从而提高模型泛化能力。最优回归超平面转化为二次规划问题的公式为:

$$\min Q = \frac{1}{2} \|W\|^2 + C \sum_{i=1}^n (\beta_1 + \beta_2) \quad (2)$$

式中: Q 为优化目标; W 为权重系数。

最后通过拉格朗日函数将其转化为对偶形式,其中核函数可以把数据映射到高维,以求取最优分割超平面,得到回归函数为:

$$\begin{cases} f(x) = \sum_{i,j=1}^n (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \\ \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \\ s.t. \begin{cases} 0 \leq \alpha_i \leq C \\ 0 \leq \alpha_i^* \leq C \end{cases} \end{cases} \quad (3)$$

式中: α_i, α_i^* 为拉格朗日因子; $K(x_i, x_j)$ 为核函数,包括线性、多项式、高斯等。

3.2 支持向量机回归预测模型的建立

3.2.1 数据预处理

分别选用渤中凹陷 BZ13-2-4、BZ13-2-5、BZ13-2-6、BZ19-6-12、BZ19-6-14、BZ19-6-15 等 6 口井及狮子沟地区狮 15、狮 20、狮 24、狮 25、狮 27 等 5 口井各 154 个测井及实测总有机碳含量数据,作为模型建立样本。总有机碳含量在采集时仅能采取到毫克量级的样品,测井数据也会受到环境等因素的干扰,这导致实测值可能难以反映地下的真实水平,在进行数据预测时产生一定的误差。基于此,笔者采用人工去除异常点和滤波预处理的方法增强数据的可用性。同时,支持向量机对参数的敏感度较高,且各特征量纲不一致,因此在训练模型时将数据归一化到 $[0, 1]$ 区间,转换公式为:

$$X^* = \frac{X - X_{\max}}{X_{\max} - X_{\min}} \quad (4)$$

式中: X^* 是经归一化计算后的测井曲线数据; X 是原始测井数据; X_{\max} 和 X_{\min} 为该测井曲线数据的极大值和极小值。

3.2.2 模型评价指标

选用决定系数 (R^2)、均方误差 (MSE) 和平均

绝对误差 (MAE) 计算得到相关性及训练误差或测试误差,以评价模型的预测性能。 R^2 也称为拟合优度,是相关系数的平方,用于评价拟合的好坏, R^2 越接近于 1,则拟合回归的效果越好。均方误差是指参数估计值与参数真实值的差平方的期望值,而平均绝对误差是绝对误差的平均值,能更好地反映预测值误差的实际情况,两者的值越小,则预测模型拥有更好的精确度。指标计算公式为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (5)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

式中: \hat{y}_i 为第 i 个样本的总有机碳含量预测值; y_i 为第 i 个样本的总有机碳含量实测值; n 为样本个数。

3.2.3 支持向量机模型

在进行烃源岩总有机碳含量预测时,应选择对有机质更敏感的曲线,研究显示自然伽马、电阻率、声波时差、密度、井径等曲线符合要求。在此基础上,采用 XGBoost 特征重要性分析结合相关性分析的方法进行降维操作,以选取得分较高的特征。前者原理为根据结构分数的增益情况计算得到作为分割点的特征,而特征的重要性即为其在所有树中出现的次数之和。因此本文选出的在渤中凹陷和狮子沟地区的特征分别为 DT, DEN, RD, GR, SP 和 DT, CAL, GR, RT, 能够全面有效反映岩性、孔隙度等烃源岩性质。在实测数据中人工划分 80% 的数据作为训练集,20% 的数据作为测试集,即渤中凹陷测试集为 BZ13-2-5 井和 BZ19-6-14 井的 30 个数据,狮子沟地区测试集为狮 25 井的 31 个数据。在模型的建立过程中,本文采用交叉验证的方法,检验不同情况下的模型预测精度以确定超参数,同时试验了不同的核函数的预测效果(表 1),在对比之后选择了使用线性核函数的支持向量机模型。

4 应用效果对比分析

通过相关性分析和误差检验的方式,可以发现支持向量机烃源岩总有机碳含量预测模型在 2 个研究区的应用都达到了理想效果,实测与预测数据

表 1 渤海湾盆地渤中凹陷烃源岩样品不同核函数预测效果对比

Table 1 Comparison of forecasting effects of different core functions applied on source rock samples from Bozhong Sag, Bohai Bay Basin

预测效果	线性核函数	多项式核函数	高斯核函数
R^2	0.95	0.53	0.92
MSE	0.04	0.41	0.06
MAE	0.17	0.46	0.21

具有很好的一致性。模型在渤海湾盆地渤中凹陷的效果总体表现更好, R^2 达到 0.95, 反映总有机碳含量实测值和预测值具有高相关性(图 4a); MSE 和 MAE 分别为 0.04 和 0.17, 实测值和真实值之间

误差很小, 预测精度很高。模型在柴达木盆地狮子沟地区的预测准确度稍低于前者, 但也具有较强的表达能力, R^2 为 0.72, 在去掉 2 个异常值后达到 0.81, 具有较高的相关性(图 4b), 拟合直线方程的斜率为 0.84, 说明拟合效果随总有机碳含量值的增加而变差; MSE 和 MAE 分别为 0.04 和 0.20。

总体来说, 模型通过大量机器学习均达到了一定的稳定性和泛化能力, 可以应用至不同沉积环境的烃源岩中。同时, 渤中凹陷在相关性和误差等多方面均优于狮子沟地区, 尤其在渤中凹陷可以灵敏地预测出实测数据中出现的高值和低值(图 5)。原因主要是渤中凹陷中测井曲线对于烃源岩有机

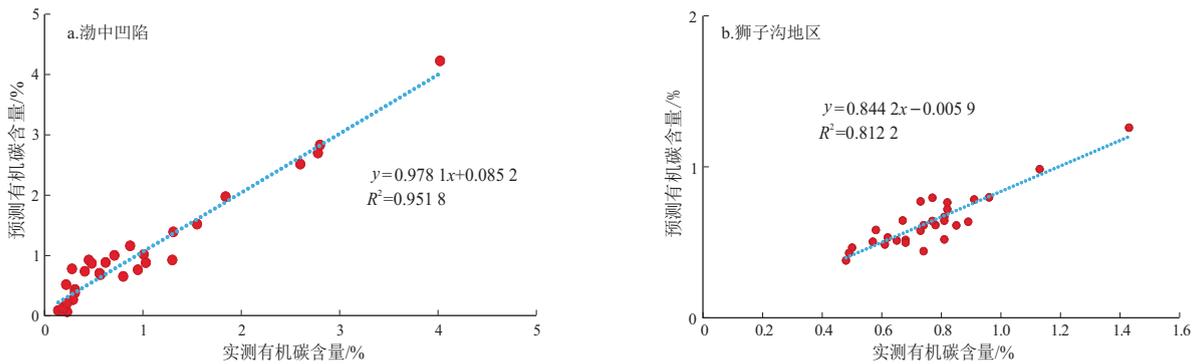


图 4 渤海湾盆地渤中凹陷和柴达木盆地狮子沟地区总有机碳含量实测和预测数据相关性
Fig.4 Correlations of measured and predicted TOC contents of samples from Bozhong Sag of Bohai Bay Basin and Shizigou area of Qaidam Basin

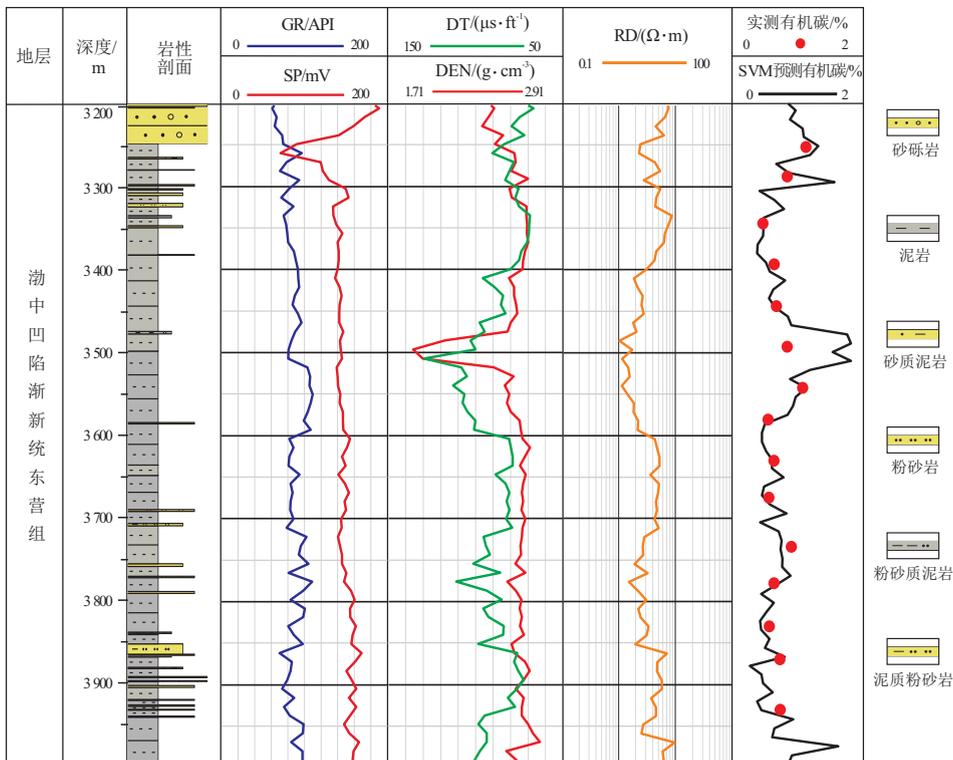


图 5 渤海湾盆地渤中凹陷烃源岩总有机碳含量预测模型 BZ13-2-5 井单井对比
Fig.5 Single well comparison of source rock TOC prediction model in Bozhong Sag, Bohai Bay Basin

质的敏感性更高,机器学习方法根据数据挖掘出测井曲线与总有机碳含量之间的关系,因此更高的相关性能够为模型预测提供更多的参考信息。除此之外,测井参数的有效性会极大影响模型预测结果,在沉积过程中高盐度地区会在一定程度上受到盐度变化影响,从而出现测井曲线变化异常的情况,如石盐段发育有异常高声波时差、高电阻率等,碳酸盐岩与泥岩混积段的测井曲线也易发生震荡^[34]。

5 结论

(1)支持向量机烃源岩总有机碳含量预测模型在应用至典型的淡水湖相盆地和咸水湖相盆地时,可采用不同的超参数而具有很强的泛化能力,能够适应不同地区的地质特征,以实现良好的预测性能。

(2)模型在渤海湾盆地的应用效果显著高于柴达木盆地,相关性检验和误差分析等均存在明显的优劣分异,主要原因在于沉积盆地的咸化程度影响着测井曲线与烃源岩有机质丰度的相关程度,从而成为影响预测性能的主要因素。

(3)本文建立的预测模型可以为烃源岩有机质研究的外推和泛化提供一定参考,后续可以进行更大范围的模型适用性推广,同时采用更多的实测总有机碳含量数据充分训练得到效果卓越的模式。

参考文献:

[1] SCHMOKER J W. Determination of organic-matter content of Appalachian and Evonian shales from Gamma-ray logs[J]. AAPG Bulletin, 1981, 65(7): 1285-1298.

[2] 秦建强,付德亮,钱亚芳,等.烃源岩有机质丰度预测的地球物理研究进展[J].石油物探,2018,57(6):803-812.

QIN Jianqiang, FU Deliang, QIAN Yafang, et al. Progress of geophysical methods for the evaluation of TOC of source rock [J]. Geophysical Prospecting for Petroleum, 2018, 57(6): 803-812.

[3] 区航,杨辉廷,黄晓兵. X 气田飞仙关组礁滩相储层测井解释方法应用[J]. 油气藏评价与开发, 2021, 11(5): 744-752.

OU Fang, YANG Huiting, HUANG Xiaobing. Application of logging interpretation method for reef shoal reservoir in Feixianguan Formation of X Gas Field [J]. Reservoir Evaluation and Development, 2021, 11(5): 744-752.

[4] PASSEY Q R, CREANEY S, KULLA J B, et al. A practical model for organic richness from porosity and resistivity logs[J]. AAPG Bulletin, 1990, 74(5): 1777-1794.

[5] MENDELZON J D, TOKSOZ M N. Source rock characterization using multivariate analysis of log data[C]//SPWLA 26th Annual Logging Symposium. Texas: Society of Petrophysicists & Well-Log Analysts, 1985: 1-21.

[6] 赵兴齐,陈践发,郭望,等. BP 神经网络在西湖凹陷烃源岩评价中的应用[J]. 测井技术, 2013, 37(5): 567-571.

ZHAO Xingqi, CHEN Jianfa, GUO Wang, et al. The application of BP neural network to the source rocks evaluation in Xihu Sag [J]. Well Logging Technology, 2013, 37(5): 567-571.

[7] 王惠君,赵桂萍,李良,等.基于卷积神经网络(CNN)的泥质烃源岩 TOC 预测模型:以鄂尔多斯盆地杭锦旗地区为例[J]. 中国科学院大学学报, 2020, 37(1): 103-112.

WANG Huijun, ZHAO Guiping, LI Liang, et al. TOC prediction model for muddy source rocks based on convolutional neural network (CNN): a case study of the Hangjinqi area of the Ordos Basin [J]. Journal of University of Chinese Academy of Sciences, 2020, 37(1): 103-112.

[8] 石创,朱俊章,龙祖烈,等.基于概率神经网络的烃源岩 TOC 预测:以珠江口盆地陆丰南区为例[J]. 断块油气田, 2019, 26(5): 561-565.

SHI Chuang, ZHU Junzhang, LONG Zulie, et al. Prediction of total organic carbon in source rocks by probabilistic neural network: a case study of southern Lufeng area in Pearl River Mouth Basin [J]. Fault-Block Oil and Gas Field, 2019, 26(5): 561-565.

[9] 张成龙,陶士振,白斌,等.基于支持向量机模型的烃源岩有机碳含量预测:以鄂尔多斯盆地为例[J]. 天然气地球科学, 2019, 30(5): 761-768.

ZHANG Chenglong, TAO Shizhen, BAI Bin, et al. Source rock TOC content prediction based on the support vector machine model: an application in Ordos Basin [J]. Natural Gas Geoscience, 2019, 30(5): 761-768.

[10] 李欣,李建忠,杨涛,等.渤海湾盆地油气勘探现状与勘探方向[J]. 新疆石油地质, 2013, 34(2): 140-144.

LIN Xin, LI Jianzhong, YANG Tao, et al. Oil-gas exploration status and future targets in Bohai Bay Basin [J]. Xinjiang Petroleum Geology, 2013, 34(2): 140-144.

[11] 夏庆龙,徐长贵.渤海海域复杂断裂带地质认识创新与油气重大发现[J]. 石油学报, 2016, 27(S1): 22-33.

XIA Qinglong, XU Changgui. New geological understandings and major hydrocarbon discoveries in the complex fault zone of Bohai Sea [J]. Acta Petrolei Sinica, 2016, 27(S1): 22-33.

[12] 孙永河,漆家福,吕延防,等.渤中坳陷断裂构造特征及其对油气的控制[J]. 石油学报, 2008, 29(5): 669-675.

SUN Yonghe, QI Jiafu, LV Yanfang, et al. Characteristics of fault structure and its control to hydrocarbon in Bozhong Depression [J]. Acta Petrolei Sinica, 2008, 29(5): 669-675.

[13] HAO Fang, ZHOU Xinhui, ZHU Yangming, et al. Charging of the Neogene Penglai 19-3 field, Bohai Bay Basin, China: oil accumulation in a young trap in an active fault zone [J]. AAPG Bulletin, 2009, 93(2): 155-179.

[14] 郭玉新. 渤中凹陷埕岛东段东三段沉积物重力流类型及沉积模式[J]. 油气地质与采收率, 2021, 28(3): 14-24.

GUO Yuxin. Types and sedimentary models of sediment gravity flows of Ed³ member in east slope of Chengdao area, Bozhong Sag [J]. Petroleum Geology and Recovery Efficiency, 2021, 28(3): 14-24.

[15] 牛成民,王飞龙,何将启,等.渤海海域渤中 19-6 潜山气藏成藏要素匹配及成藏模式[J]. 石油实验地质, 2021, 43(2): 259-267.

NIU Chengmin, WANG Feilong, HE Jiangqi, et al. Accumulation

- factor matching and model of Bozhong 19-6 buried hill gas reservoir, Bohai Sea area [J]. *Petroleum Geology & Experiment*, 2021, 43(2): 259-267.
- [16] 林会喜,熊伟,王勇,等.济阳坳陷埕岛潜山油气成藏特征[J].油气地质与采收率,2021,28(1):1-9.
LIN Huixi, XIONG Wei, WANG Yong, et al. Hydrocarbon accumulation in Chengdao buried hill of Jiyang Depression [J]. *Petroleum Geology and Recovery Efficiency*, 2021, 28(1): 1-9.
- [17] 谢玉洪.渤海湾盆地渤中凹陷太古界潜山气藏 BZ19-6 的气源条件与成藏模式[J].石油实验地质,2020,42(5):858-866.
XIE Yuhong. Gas resources and accumulation model of BZ19-6 Archean buried-hill large-scale gas reservoir in Bozhong Sag, Bohai Bay Basin [J]. *Petroleum Geology & Experiment*, 2020, 42(5): 858-866.
- [18] 庞雄奇,郭永华,姜福杰,等.渤海海域优质烃源岩及其分布预测[J].石油与天然气地质,2009,30(4):393-397.
PANG Xiongqi, GUO Yonghua, JIANG Fujie, et al. High-quality source rocks and their distribution prediction in the Bohai Sea waters [J]. *Oil & Gas Geology*, 2009, 30(4): 393-397.
- [19] 江涛,李慧勇,胡安文,等.渤中西洼东三段烃源岩特征与油气成藏模式[J].特种油气藏,2017,24(6):12-17.
JIANG Tao, LI Huiyong, HU Anwen, et al. Source rock characterization and hydrocarbon accumulation of Dong3 Member in west Bozhong Sag [J]. *Special Oil & Gas Reservoirs*, 2017, 24(6): 12-17.
- [20] 王翔宇.渤海湾盆地渤中凹陷渐新统东营组三段烃源岩预测及评价[D].荆州:长江大学,2019.
WANG Xiangyu. Prediction and evaluation of the source rocks of the third member of the Oligocene Dongying Formation in the Bozhong Sag, Bohai Bay Basin [D]. Jingzhou: Yangtze University, 2019.
- [21] 刘溪溪,岳鑫,袁文虎,等.柴达木盆地西部狮子沟背斜构造区深部卤水水化学特征及演化分析[J].盐湖研究,2019,27(1):73-81.
LIU Xixi, YUE Xin, YUAN Wenhua, et al. Hydrochemical characteristics and evolutionary process of deep brines from Shizigou anticline structure in Qaidam Basin, China [J]. *Journal of Salt Lake Research*, 2019, 27(1): 73-81.
- [22] 魏学斌,沙威,沈晓双,等.柴达木盆地油气勘探历程与启示[J].新疆石油地质,2021,42(3):302-311.
WEI Xuebin, SHA Wei, SHEN Xiaoshuang, et al. Petroleum exploration history and enlightenment in Qaidam Basin [J]. *Xinjiang Petroleum Geology*, 2021, 42(3): 302-311.
- [23] 舒豫川,胡广,庞谦,等.柴达木盆地咸湖相烃源岩特征:以英西地区下干柴沟组上段为例[J].断块油气田,2021,(2):179-186.
SHU Yuchuan, HU Guang, PANG Qian, et al. Characteristics of source rocks of salt lake facies in Qaidam Basin; taking upper member of Xiaganchaigou Formation in Yingxi region as an example [J]. *Fault-Block Oil and Gas Field*, 2021, (2): 179-186.
- [24] 王琳霖,于冬冬,浮响,等.柴达木盆地西部构造演化与差异变形特征及对油田水分布的控制[J].石油实验地质,2020,42(2):186-192.
WANG Linlin, YU Dongdong, FU Yun, et al. Tectonic evolution and differential deformation controls on oilfield water distribution in western Qaidam Basin [J]. *Petroleum Geology & Experiment*, 2020, 42(2): 186-192.
- [25] 沈亚,李洪革,管俊亚,等.柴西地区古近系—新近系含油凹陷构造特征与勘探领域[J].石油地球物理勘探,2012,47(S1):111-117.
SHEN Ya, LI Hongge, GUAN Junya, et al. Structure features and exploration potential of the Paleogene-Neogene depressions in western Qaidam Basin [J]. *Oil Geophysical Prospecting*, 2012, 47(S1): 111-117.
- [26] 付锁堂,马达德,陈琰,等.柴达木盆地油气勘探新进展[J].石油学报,2016,37(S1):1-10.
FU Suotang, MA Dade, CHEN Yan, et al. New advance of petroleum and gas exploration in Qaidam Basin [J]. *Acta Petrolei Sinica*, 2016, 37(S1): 1-10.
- [27] 中华人民共和国国家质量监督检验检疫总局.GB/T 19145-2003:沉积岩中总有机碳的测定[S].北京:中国标准出版社,2003.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. GB/T 19145-2003; Determination of total organic carbon in sedimentary rock [S]. Beijing: China Standard Press, 2003.
- [28] SONDERGELD C H, NEWSHAM K E, COMISKY J T, et al. Petrophysical considerations in evaluating and producing shale gas resources [C]//SPE Unconventional Gas Conference. Pittsburgh, Pennsylvania, USA: Society of Petroleum Engineers, 2010.
- [29] VAPNIK V N. Estimation of dependences based on empirical data [M]. New York: Springer-Verlag, 1982.
- [30] VAPNIK V N. The nature of statistical learning theory [M]. New York: Springer-Verlag, 1995.
- [31] 杨斌,匡立春,孙中春,等.一种用于测井油气层综合识别的支持向量机方法[J].测井技术,2005,29(6):511-514.
YANG Bin, KUANG Lichun, SUN Zhongchun, et al. On support vector machines method to identify oil & gas zone with logging and mudlog information [J]. *Well Logging Technology*, 2005, 29(6): 511-514.
- [32] 李新虎.基于不同测井曲线参数集的支持向量机岩性识别对比[J].煤田地质与勘探,2007,35(3):72-76.
LI Xinhua. Lithology identification methods contrast based on support vector machines at different well logging parameter [J]. *Coal Geology & Exploration*, 2007, 35(3): 72-76.
- [33] 牟丹,王祝文,黄玉龙,等.基于SVM测井数据的火山岩岩性识别:以辽河盆地东部坳陷为例[J].地球物理学报,2015,58(5):1785-1793.
MOU Dan, WANG Zhuwen, HUANG Yulong, et al. Lithological identification of volcanic rocks from SVM well logging data; case study in the eastern depression of Liaohe Basin [J]. *Chinese Journal of Geophysics*, 2015, 58(5): 1785-1793.
- [34] 太万雪,刘成林,田继先,等.柴达木盆地西部古近系咸化湖盆烃源岩总有机碳含量预测[J].特种油气藏,2021,28(1):74-80.
TAI Wanxue, LIU Chenglin, TIAN Jixian, et al. Prediction of total organic carbon content of source rocks in Paleogene salinized lake basin in western Qaidam Basin [J]. *Special Oil & Gas Reservoirs*, 2021, 28(1): 74-80.