



DOI: 10.11885/j.issn.1674-5086.2021.10.23.01

文章编号: 1674-5086(2022)04-0081-10

中图分类号: TE33

文献标志码: A

基于可解释机器学习的水平井产能预测方法

马先林^{1,2*}, 周德胜^{1,2}, 蔡文斌^{1,2}, 李宪文³, 何明舫³

1. 西安石油大学石油工程学院, 陕西 西安 710065

2. 西部低渗—特低渗油藏开发与治理教育部工程研究中心, 陕西 西安 710065

3. 中国石油长庆油田公司油气工艺研究院, 陕西 西安 710018

摘要: 准确预测致密气藏分段压裂水平井产能是压裂效果评价和优化设计的关键环节。现有的产能预测方法, 引入了过多的假设和简化, 很难全面反映致密储层流体多尺度的运移机理和复杂物理过程, 导致产能预测误差较大。提出一种基于机器学习的致密气藏分段压裂水平井产能预测方法, 该方法综合利用已收集的地质、压裂水平井产能及钻完井等多类型数据, 通过机器学习算法直接挖掘数据内部规律, 建立产能预测模型。此外, 为解决常规机器学习模型的“黑盒子”问题, 还利用 SHAP (SHapley Additive exPlanations) 方法对建立的机器学习模型进行全局和局部解释, 分析影响产能的主要因素, 增加了模型的可信性和透明度。以苏里格气田苏东示范区为例, 验证了该方法的有效性和实用性。与油气藏数值方法相比, 该方法不仅提高了产能预测的精度, 而且缩短了建模周期, 加快了计算速度。

关键词: 分段压裂水平井; 机器学习; 产能预测; 可解释性; 数据驱动; SHAP 方法

An Interpretable Machine Learning Approach to Prediction Horizontal Well Productivity

MA Xianlin^{1,2*}, ZHOU Desheng^{1,2}, CAI Wenbin^{1,2}, LI Xianwen³, HE Mingfang³

1. School of Petroleum Engineering, Xi'an Shiyou University, Xi'an, Shaanxi 710065, China

2. MOE Engineering Research Center of Development and Management for Low to Ultra-Low Permeability Oil & Gas Reservoirs in West China, Xi'an, Shaanxi 710065, China

3. Oil and Gas Technology Research Institute, Changqing Oilfield Company, PetroChina, Xi'an, Shaanxi 710018, China

Abstract: It is essential to predict multistage fractured horizontal well (MFHW) productivity of tight sand gas reservoirs for evaluation of hydraulic fracturing performance and optimization of hydraulic fracturing design. However, most of the current predictive methods introduce multiple assumptions and simplifications. Therefore, these methods cannot fully account for multi-scale fluid flow mechanisms in the tight formations of well productivity. A machine learning approach for predicting MFHW productivity is proposed. A well productivity model is built by machine learning algorithms to uncover hidden patterns in a data set including geological, fractured well productivity, drilling and completion data. In addition, to solve the “black box” issue of conventional machine learning modelling, the SHAP (SHapley Additive exPlanations) method is used to explain the built ML model globally and locally. The efficiency and practicality of the proposed method is demonstrated by the application to the Eastern Sulige Gas Field. Compared with petroleum reservoir simulation, the method not only improves the prediction performance of the well productivity, but also reduces modelling cycle and improve computational speed.

Keywords: multistage fractured horizontal well; machine learning; productivity prediction; interpretability; data-driven; SHAP method

网络出版地址: <http://kns.cnki.net/kcms/detail/51.1718.TE.20220628.1626.002.html>

马先林, 周德胜, 蔡文斌, 等. 基于可解释机器学习的水平井产能预测方法[J]. 西南石油大学学报(自然科学版), 2022, 44(4): 81-90.

MA Xianlin, ZHOU Desheng, CAI Wenbin, et al. An Interpretable Machine Learning Approach to Prediction Horizontal Well Productivity[J]. Journal of Southwest Petroleum University (Science & Technology Edition), 2022, 44(4): 81-90.

* 收稿日期: 2021-10-23 网络出版时间: 2022-06-30

通信作者: 马先林, E-mail: Xianlinm@xsyu.edu.cn

基金项目: 国家自然科学基金面上项目(51974253); 国家自然科学基金重点项目(51934005); 国家科技重大专项(2016ZX05050-009); 陕西省自然科学基金基础研究计划(2017JM5109); 陕西省教育厅重点实验室科研计划(18JS085)

引 言

传统的多段压裂水平井产能预测模型,大多采用机理驱动的方式建立,建模方法包括解析、半解析及数值模拟方法等。解析法基于均质储层、单相渗流及简单平板状双翼对称裂缝等假设,以油气渗流力学为基础,应用压力叠加原理和位势能理论等得出产能计算公式,计算效率高^[1-5];而对于非均质较强、多相渗流、复杂裂缝网络的储层,应用油气藏数值模拟法进行产能预测,在精细油气藏描述的基础上,应用连续介质或离散模型描述人工裂缝和天然裂缝,通过模拟油气在储层中渗流过程,计算分段压裂水平井的产能^[6-9]。

传统的产能预测方法对储层特征和复杂缝网几何形态表征不但存在较强的不确定性,而且还引入了一些理想化假设和计算上的简化,使得产能预测结果可能偏离实际。另外,在渗流方程求解时,大多数采用网格加密及迭代计算等手段,导致建模周期长,计算效率低下。事实上,分段压裂水平井的生产规律受到地质、钻完井、压裂施工及生产制度等诸多因素的影响,因此,过多的理想化假设以及求解方法的复杂性限制了机理驱动产能预测方法在现场的应用与推广。

随着人工智能的兴起,特别是机器学习技术为

多级压裂水平井产能预测提供了新的途径。与传统的机理驱动方法相比,该方法通过对油气大数据的挖掘,建立压裂井产能与储层、钻完井和压裂施工等参数之间的相关模型,能够迅速地对压裂效果进行评价,从而进行压裂优化设计。基于数据驱动的建模方法已经在致密油气藏开发中得到了初步的应用^[10-14],体现了在产能预测方面独特的优势。但所建立的预测模型大多是“黑盒子”模型,即输入一组相关参数值能够获得产能的预测值,却不能解释模型内部的预测机制,从而降低了模型的可信度与实用性。为此,许多学者对提高机器学习模型可解释性进行了研究,并提出了一些解释方法^[15]。机器学习产能建模是纯数据驱动的方法,预测精度依赖于样本数据的个数和质量,小样本会导致预测失真和泛化能力降低。

本文首先介绍致密气藏水平井产能预测机器学习建模流程和机器学习原理,其次讨论对预测模型进行解释的 SHAP(SHapley Additive exPlanations)方法原理,并以苏里格气田东区分段压裂水平井为例,对方法的有效性和实用性进行验证。

1 水平井产能预测机器学习建模流程

利用机器学习建立致密气藏分段压裂水平井产能预测模型,包括 6 个步骤,如图 1 所示。

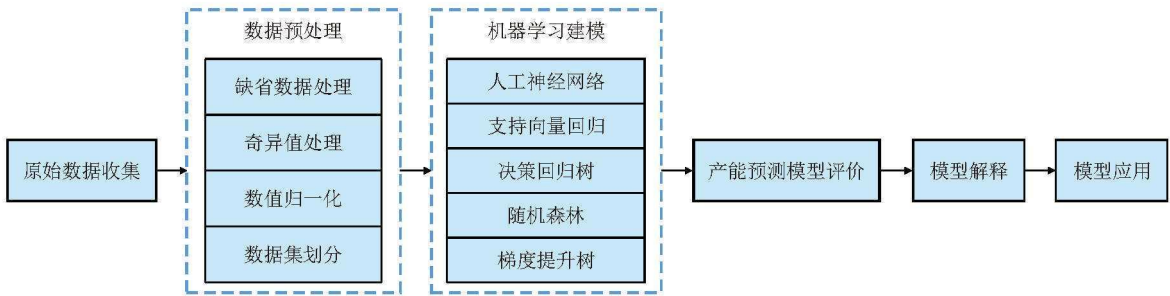


图 1 水平井产能机器学习建模流程图
Fig. 1 Workflow for modelling horizontal well productivity using machine learning

- (1) 原始数据收集。数据集包括主要影响参数及其压裂后水平井的产能评价指标,其中,影响参数包括地质、工程等因素,产能可以是无阻流量或产气量等。
- (2) 数据预处理。先进行数据清洗、缺失数据插补、数据降维及数据转换等^[16],再将预处理后的数据集划分为训练集和测试集,划分比例一般取 70%~30% 或 80%~20%。

- (3) 机器学习建模。应用训练集数据建立产能模型,主要寻找机器学习模型中最优的超参数取值,常用的优化方法包括网格搜索、多折交叉验证及自动学习等。
- (4) 产能预测模型评价。利用测试集数据评估产能预测模型精度,常用精度评价指标包括决定系数(R^2)、平均绝对误差(MAE)及均方误差(MSE)。根据评价指标,选择在测试集上具有最高预测性能

的机器学习算法所建立的产能预测模型。

(5) 模型解释。基于建立的最优产能预测模型,利用 SHAP 方法对产能预测进行全局和局部解释。

(6) 模型应用。应用建立的产能预测模型,评价压裂效果,优化新井的压裂参数设计。

2 机器学习原理

2.1 机器学习方法

机器学习是一门多领域交叉技术,其本质是一种特殊的算法。通过分析大数据,发现数据内部潜在的模式并应用这些模式进行预测^[17]。假设一个训练数据集 D 包含 n 个学习样本,每个样本有 m 个影响参数和一个输出参数,即

$$D = \{x_{1i}, x_{2i}, \dots, x_{mi}, y_i\},$$

$$i = 1, 2, \dots, n \text{ 且 } n \gg m \quad (1)$$

选择相应的机器学习算法在此数据集建立一个预测模型 $\hat{y} = f(x, \theta)$, 优化算法的超参数 θ , 使得以下误差函数值最小

$$f = \min_{\theta} \sum_{j=1}^n L(y_i, f(x_i, \theta)) \quad (2)$$

式中: f —某一机器学习算法建立的模型; θ —该算法的超参数; $L(y_i, f(x_i, \theta))$ —误差函数。

典型的机器学习算法包括人工神经网络、支持向量机、决策树及随机森林及梯度提升树等。

2.2 机器学习模型的可解释性

机器学习模型可解释性为用户提供了理解机器学习模型一个接口,既是机器学习模型的代理,又是一种解释模型的方法^[18-19]。模型可解释性分为

两类:事前可解释性和事后可解释性。事前可解释性是指在建模时,采用可解释性好的模型或设计有解释能力的模型,使模型自身具有解释能力。事后可解释性是对已建立的机器学习模型进行解释,独立于机器学习建模过程,灵活性较强。根据可解释的范围,事后可解释性又分为全局和局部两种,全局可解释用于理解模型内部的工作原理,对模型整体能力的解释^[20];局部可解释用于理解机器学习模型针对单个样本的预测过程和依据,对样本的预测结果进行解释^[21]。

本文是对建好的水平井产能模型预测结果进行解释,属于事后可解释性,应用 SHAP 可解释技术。SHAP 法通过计算每一个输入变量对预测的贡献值进行模型解释,是一种加性解释方法。比如,使用 SHAP 法解释样本 x^* 的机器模型预测值 $\hat{y} = f(x^*)$ 时,预测值 $f(x^*)$ 可以分解成

$$f(x^*) = \phi_0 + \sum_{j=1}^M \phi_j^* \quad (3)$$

式中: ϕ_0 —预测模型 $f(x)$ 在数据集上的平均预测值; ϕ_j^* —第 j 个输入变量对样本 x^* 预测的贡献值,即 SHAP 值; M —输入变量的数量。

对于 3 个输入变量的预测问题,式(3)的示意流程见图 2,红色表示 SHAP 值是正的,蓝色则表示 SHAP 值是负的。图 2 表明,SHAP 值表示某一个样本的预测值各个输入变量的贡献大小^[22], ϕ_1 表示变量 1 用于预测后引起当前预测值的变化,是正影响,导致预测值增加,同理, ϕ_2 为变量 2 加入后当前预测值的增加量,但是,当变量 3 加入预测后,导致当前预测值减少,减少量为 ϕ_3 。

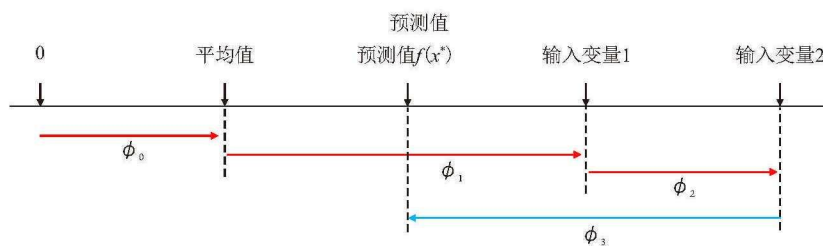


图 2 3 个输入变量 SHAP 方法
Fig. 2 SHAP method with three input variables

3 应用实例

3.1 数据收集

收集的原始数据包括苏里格气田东区 598 口分段压裂水平井,产能影响参数包括 7 个地质和完井

参数以及 15 个压裂施工输入参数,输出参数为无阻流量,具体参数描述见表 1。

由于大部分水平井没有进行测井,不能分析孔隙度、渗透率、杨氏模量及泊松比等参数对产能的影响。

表 1 产能影响参数

Tab. 1 Influential parameters of well productivity

类型	参数
地质和完井	井区、开发层位、垂直井深、完井井深、水平井水平段长、储层水平段长及有效储层水平段长度
压裂施工	改造工艺、压裂段数、砂量、最大排量、入井总液量、破裂压力、最高施工压力及液氮总量

3.2 数据预处理

3.2.1 缺失值处理

在数据集中, 储层水平段长度的缺失值多达 25.00%, 破裂压力有 13.00% 的数值缺失, 远大于 5.00% 安全最大阈值插补的要求, 因此, 这些变量在建模时不予以考虑。其他变量缺失值的分布如图3所示。由图 3a 可以看出, 大部分变量的缺失值低于 2.00%。

图 3b 显示了变量缺失值的分布, 其中, 最上面的红色竖条表示入井总液量有 0.25 % 数值缺少, 第二行的红竖条表示液氮量缺失 0.50 %, 第三行的两个红竖条表示液氮量和施工最高压力同时缺失达 1.26%, 第四行的红竖条表示垂深缺失 2.02 %, 最下行表示 95.97 % 数据没有缺失。由于数据缺失较少, 故决定删除数据表中有缺失值的那些井, 剩余水平井数量为 574 口。

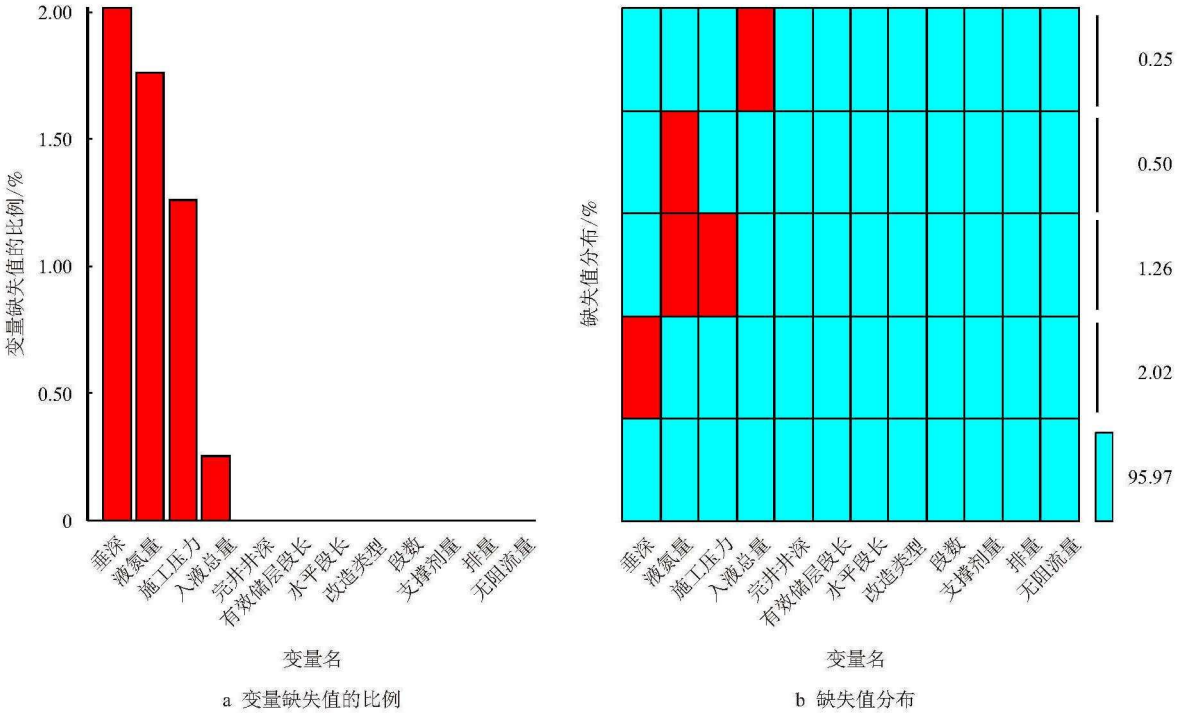


图 3 变量缺失值分析

Fig. 3 Missing data analysis

3.2.2 异常值处理

原始数据中若存在不合理的值, 会影响模型预测精度。采用基于箱线图的异常值检测方法, 利用数据中的上、下四分位数(Q_3 、 Q_1)及四分位距(I_{QR}), 四分位距定义为上四分位数与下四分位数的差值, 即 $I_{QR} = Q_3 - Q_1$, 变量的值大于 $Q_3 + 1.5I_{QR}$ 和小于 $Q_1 - 1.5I_{QR}$ 为异常值。在剔除数据集中的异常值后, 用于机器学习建模的水平井为 532 口。

3.2.3 描述性数据分析

描述性数据分析采用图表等方式对数据进行统

计性描述, 是建模的重要内容。

图 4 显示了 2012—2020 年平均无阻流量和主要压裂施工参数的变化趋势, 随着平均最大排量、加砂量和压裂段数的增加, 水平井分段压裂后的增产效果明显。

图 5 给出了主要压裂参数与无阻流量的箱线图, 从图中可以看出, 无阻流量与水平段长度、压裂段数、加砂量、排量和入井总液量均存在正相关关系, 采用水力桥塞分段压裂工艺后显著地提高了水平井的产能。

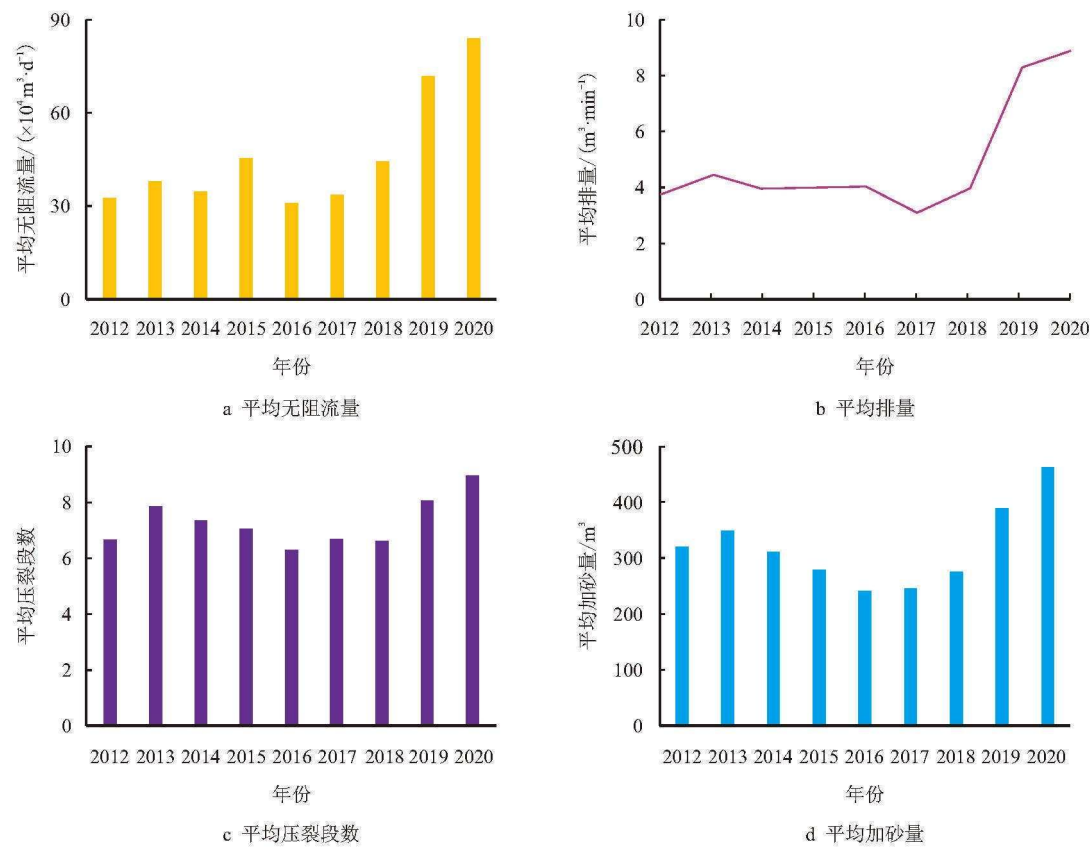


图 4 平均无阻流量和压裂施工参数的变化趋势
Fig. 4 Trend analysis of average absolute open flow potential and hydraulic fracturing treatment

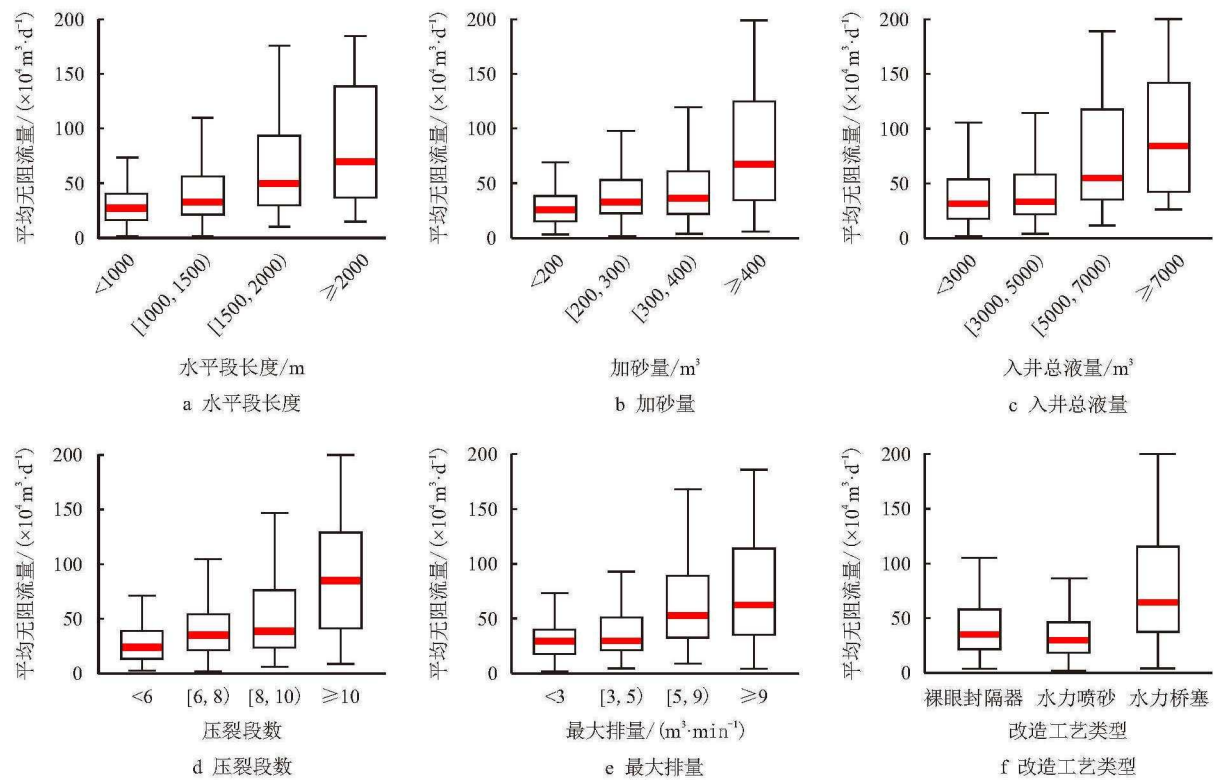


图 5 压裂参数影响无阻流量的箱线图
Fig. 5 Boxplots for effect of hydraulic fracturing treatment parameters on AFOP

3.2.4 相关性分析

若同时使用相关性较强的输入参数,建模不仅会增加模型训练的时间,而且影响模型可解释性。

图 6 给出了变量间的 Pearson 相关系数,与无阻流量正相关最强的参数是加砂量($R=0.50$),其次是有效储层段长度、压裂段数及入井总液量等参数。完井井深与水平段长具有极强的线性正相关性($R=0.92$),表明完井较深的水平井具有较长的水平段长度;水平段长同有效储层水平段长、压裂段数、入井总液量也具有较强的线性正相关;砂量与

入井总液量 also 具有很强的线性相关性($R=0.91$)。

为减少变量间的线性相关性,建立以下输入参数:(1) 段间距 = 水平段长度/压裂段数,不再使用压裂段数;(2) 加砂强度 = 砂量/水平段长度,代替加砂量;(3) 用液强度 = 入井总液量/水平段长度,代替入井总液量。

采用逐步向后回归,对输入参数进行了选择,确定建模的输入参数为水平井水平段长、有效储层水平段长、完井井深、段间距、加砂强度、用液强度、改造工艺及最大排量。

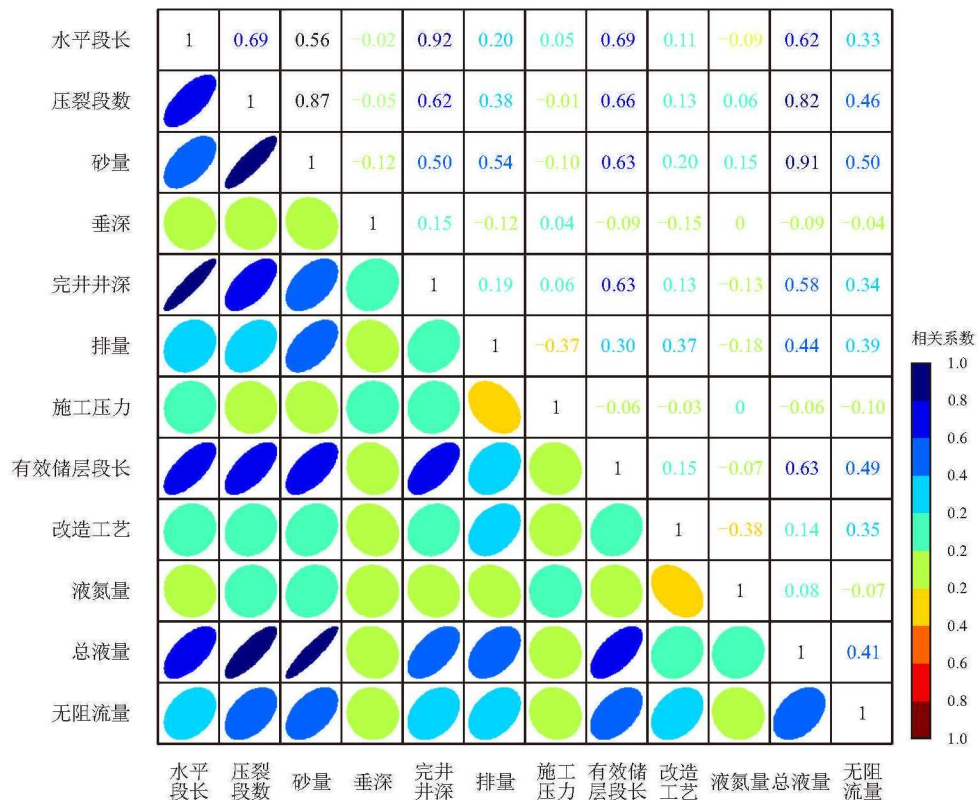


图 6 Pearson 相关系数矩阵
Fig. 6 Pearson correlation matrix

3.3 产能预测模型的构建

经过预处理后的数据集按 80%~20% 的比例随机划分成训练集与测试集,采用人工神经网络(ANN)、支撑向量机(SVM)、随机森林(RF)和梯度提升树(GBDT)算法在训练集学习,构建无阻流量预测模型。

3.3.1 模型超参数优化

采用网格搜索和十折交叉验证方法,在训练集上对 4 种机器学习算法的超参数进行优化。以 GBDT 算法为例,GBDT 中的主要超参数有决策树的棵数、决策树的深度和决策树的深度。首先确定

这些超参数的范围,再应用十折交叉验证方法进行验证,超参数决策树的棵数、决策树的深度及决策树的深度取值分别为 43、8 和 2。GBDT 模型预测结果如图 7 所示。

由图 7 可以产出,无阻流量小于 $100\times10^4\text{ m}^3/\text{d}$ 时,模型预测值比较接近实际值。然而,当无阻流量大于 $100\times10^4\text{ m}^3/\text{d}$ 时,模型预测精度不理想,主要原因是水力桥塞分段压裂水平井只占数据集的 19%,而在这些井中无阻流量大于 $100\times10^4\text{ m}^3/\text{d}$ 的达 57%,增加水力桥塞分段压裂水平井数据有助于改进模型预测精度。

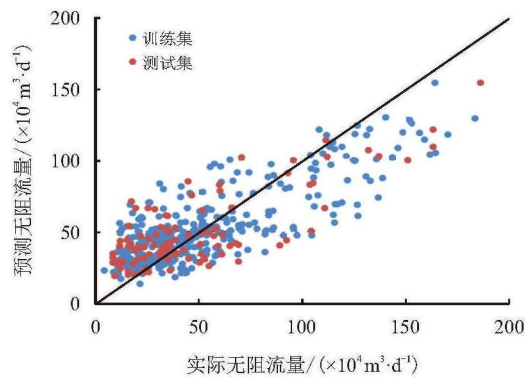


图 7 GBDT 模型预测结果
Fig. 7 Prediction results of GBDT model

3.3.2 模型评价

由于目前没有一个机器算法能够绝对优于其他算法, 主要通过评价机器学习模型在测试数据集的预测性能, 选择最适合的算法, 为此, 应用建立的机器学习模型对测试集中的 106 口水平井的无阻流量进行了预测, 表 2 是这些模型的预测性能对比。

平均绝对误差和均方误差越小, 模型预测值与实际值越接近; 决定系数越接近 1, 模型预测值与实际值越接近。由表 2 可以看出, GBDT 算法在测试集上的预测结果优于其他几种学习算法, 故对用 GBDT 建立的预测模型进行解释。

表 2 机器学习模型性能对比

Tab. 2 Comparison of performances of machine learning model

模型	训练集			测试集		
	决定系数	平均绝对误差	均方误差	决定系数	平均绝对误差	均方误差
梯度提升树(GBDT)	0.67	17.09	481.74	0.58	19.33	566.29
随机森林(RF)	0.73	15.90	119.61	0.45	22.12	783.47
支撑向量机(SVM)	0.49	22.08	817.52	0.46	21.45	780.37
人工神经网络(ANN)	0.52	21.18	798.00	0.38	23.28	890.86

3.4 产能预测模型解释

3.4.1 模型全局解释

图 8 给出了 GBDT 模型的解释, 压裂施工参数中, 改造工艺和最大排量对无阻流量的影响最显著, 2018 年以前, 苏里格气田采用裸眼封隔器和水力喷砂分段压裂改造工艺, 排量大部分集中

在 2~4 m³/min, 施工排量低, 而且施工段数受限。2018 年以后, 改用水力桥塞分段压裂改造方式, 排量提高到 6~10 m³/min, 施工段数也有所增加, 无阻流量大于 100×10⁴ m³/d 的水平井明显多于其他两种改造工艺。有效储层段长度和完井井深对无阻流量也有较大的影响, 井越深, 储层压力越大。

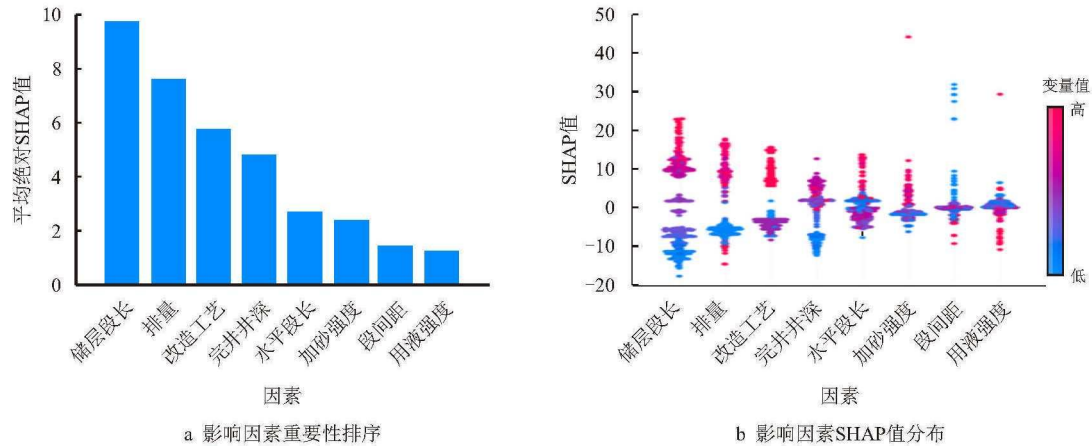


图 8 影响因素重要性全局分析
Fig. 8 Global interpretation of input variables

图 8b 展示了每个影响因素 SHAP 值的分布, 图中, 每个数据点代表一口压裂水平井, 颜色表示变量的值, 从蓝色到红色表示变量数值由低到高的变化。

正(负)的 SHAP 值表示影响参数与无阻流量呈正(负)相关, 例如, 增大排量, 排量的 SHAP 值会变大, 导致无阻流量增大; 而无阻流量则随着段间距增加而减少。

3.4.2 模型局部解释

图 9 给出了一口水平井的局部解释。该井的无阻流量是所有输入参数贡献的总和,基准值为 GBDT 模型预测的无阻流量平均值,等于 $50.88\times 10^4\text{ m}^3/\text{d}$ 。红色导致该平均值增加,而蓝色表示使其下降。有效储层段长、完井井深和用液强

度的 SHAP 值分别为 9.891, 5.294 和 1.578;最大排量、改造工艺和加砂强度的 SHAP 值分别是-6.898, -5.582 和-1.745,段间距和水平段长的影响也是负的,它们的 SHAP 值太小,在图中无法显示。最终,该压裂水平井的无阻流量预测值为 $53.03\times 10^4\text{ m}^3/\text{d}$,接近实际值 $59.93\times 10^4\text{ m}^3/\text{d}$ 。

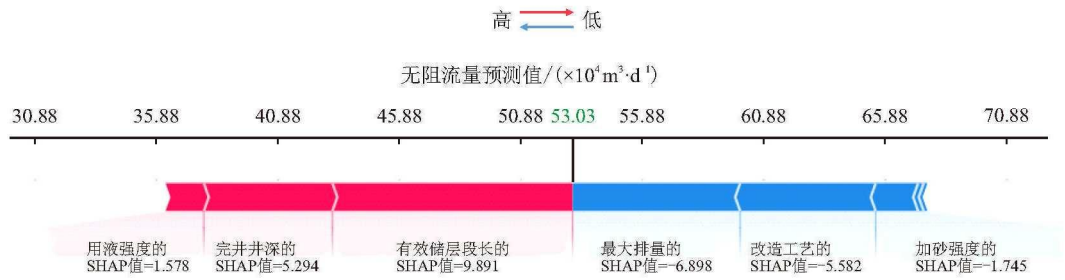


图 9 某水平井无阻流量预测结果解释
Fig. 9 Local interpretation of a fractured horizontal well

3.4.3 变量相关解释

SHAP 方法还能揭示输入变量之间相关性对无阻流量的影响。

如图 10a 所示,当采用水力桥塞分段压裂工艺时,改造工艺的 SHAP 值迅速增加,说明水力桥塞分段压裂工艺与无阻流量呈正相关;而且随着压裂水平段增长,进一步增加水力桥塞分段压裂工艺的

SHAP 值,因此,采用长压裂水平段的桥塞分段压裂能够增加无阻流量。

图 10b 表明,当段间距小于 100 m 时,段间距的 SHAP 值迅速增加,导致无阻流量增大,而且提高加砂强度有助于段间距的 SHAP 值增大,进一步增加无阻流量。但是,随着段间距的增大,却引起其 SHAP 值减少,因此,段间距与无阻流量呈负相关。

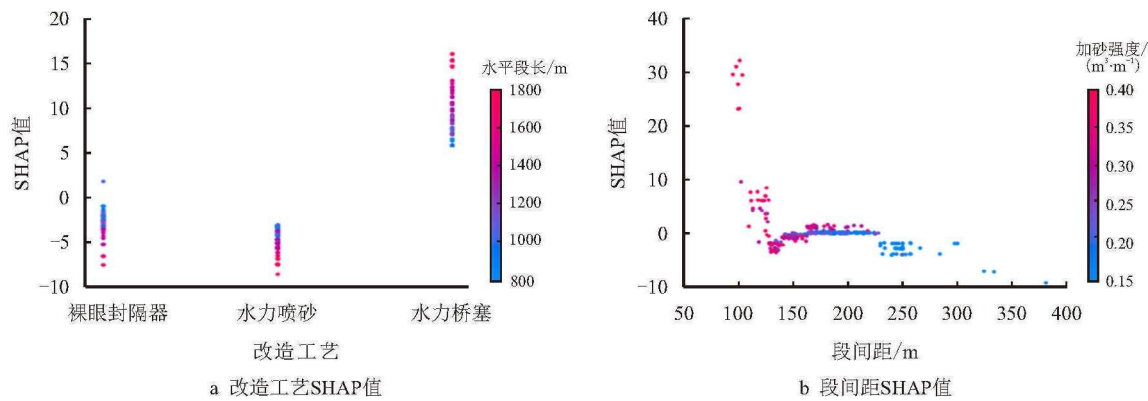


图 10 变量相关性解释
Fig. 10 Interpretation of variable correlation

4 结 论

(1) 提出了一种基于机器学习算法进行致密气藏分段压裂水平井产能预测以及模型解释的方法,该方法是对现有机理驱动产能预测方法的有效补充,具有多类型数据综合和较强的预测能力,能够有效地提高分段压裂水平井产能预测效率和精度,实用性强。

(2) 学习模型全局可解释性有助于理解梯度提升树算法产能预测机制,明确影响苏里格气田东区分段压裂水平井无阻流量主要压裂施工参数是排量和改造工艺。

(3) SHAP 值不仅有助于分析单一压裂水平井产能的影响因素,还能理解输入变量之间的交互作用对产能的影响。

(4) 首次应用可解释机器学习算法对苏里格气

田东区分段压裂水平井产能评价,可为该区的压裂优化设计和开发方案制定提供依据。

参考文献

- [1] RAJAGOPAL R, JOSHI S. Productivity of multiple drainholes or fractured horizontal wells[J]. SPE Formation Evaluation, 1993, 8(1): 11–16. doi: 10.2118/21263-PA
- [2] 郎兆新, 张丽华, 程林松. 压裂水平井产能研究[J]. 石油大学学报(自然科学版), 1994, 18(2): 43–46.
LANG Zhaoxin, ZHANG Lihua, CHENG Linsong. Investigation on productivity of fractured horizontal well[J]. Journal of the University of Petroleum, China, 1994, 18(2): 43–46.
- [3] 孙福街, 韩树刚, 程林松, 等. 低渗气藏压裂水平井渗流与井筒管流耦合模型[J]. 西南石油大学学报, 2005, 27(1): 32–37. doi: 10.3863/j.issn.1000-2634-2005.01.009
SUN Fujie, HAN Shugang, CHENG Linsong, et al. Coupling models between fractured horizontal well fluid flow of low permeability gas reservoir and wellbore pipe flow[J]. Journal of Southwest Petroleum Institute, 2005, 27(1): 32–37. doi: 10.3863/j.issn.1000-2634.2005.01.009
- [4] GUO Boyun, YU Xiance. A simple and accurate mathematical model for predicting productivity of multifractured horizontal wells[C]. SPE 114452-MS, 2008. doi: 10.2118/114452-MS
- [5] 杨兆中, 陈倩, 李小刚, 等. 致密气藏分段多簇压裂水平井产能计算新方法[J]. 大庆石油地质与开发, 2019, 38(1): 147–154. doi: 10.19597/J.ISSN.1000-3754.201806064
YANG Zhaozhong, CHEN Qian, LI Xiaogang, et al. A new method for calculating the productivity of the staged multi-cluster fractured horizontal well in tight gas reservoirs[J]. Petroleum Geology & Oilfield Development in Daqing, 2019, 38(1): 147–154. doi: 10.19597/J.ISSN.1000-3754.201806064
- [6] WU Y, GEORGE M, BAI B, et al. A multi-continuum model for gas production in tight fracture reservoirs[C]. SPE 118944-MS, 2009. doi: 10.2118/118944-MS
- [7] CIPOLLA C L, LOLONE P. Reservoir modeling and production evaluation in shale gas reservoirs[C]. IPTC 13185-MS, 2009. doi: 10.2523/IPTC-13185-MS
- [8] CIPOLLA C L, LOLONE P, MAYERHOFFER M J, et al. Fracture design consideration in horizontal wells drilled in unconventional gas reservoirs[C]. SPE 119366-MS, 2009. doi: 10.2118/119366-MS
- [9] 张烈辉, 刘沙, 雍锐, 等. 基于 EDFM 的致密油藏分段压裂水平井数值模拟[J]. 西南石油大学学报(自然科学版), 2019, 41(4): 1–11. doi: 10.11885/j.issn.1674-5086.2018.11.21.05
- ZHANG Liehui, LIU Sha, YONG Rui, et al. EDFM-based numerical simulation of horizontal wells with multi-stage hydraulic fracturing in tight reservoirs[J]. Journal of Southwest Petroleum University (Science & Technology Edition), 2019, 41(4): 1–11. doi: 10.11885/j.issn.1674-5086-2018.11.21.05
- [10] NEJAD M, SHELUDKO S, HODGSON T, et al. A case history: Evaluating well completions in the Eagle Ford Shale using a data-driven approach[C]. SPE 173336-MS, 2015. doi: 10.2118/173336-MS
- [11] SHELLEY R, NEJAD A, GULIYEV N, et al. Understanding multi-fractured horizontal Marcellus completions[C]. SPE 171003-MS, 2014. doi: 10.2118/171003-MS
- [12] 檀朝东, 贺甲元, 周彤, 等. 基于 PCA-BNN 的页岩气压裂施工参数优化[J]. 西南石油大学学报(自然科学版), 2020, 42(6): 56–62. doi: 10.11885/j.issn.1674-5086.2020.05.12.05
TAN Chaodong, HE Jiayuan, ZHOU Tong, et al. A study on the optimization of fracturing operation parameters based on PCA-BNN[J]. Journal of Southwest Petroleum University (Science & Technology Edition), 2020, 42(6): 56–62. doi: 10.11885/j.issn.1674-5086.2020.05.12.05
- [13] WANG Shuhua, CHEN Shengnan. Insights to fracture stimulation design in unconventional reservoirs based on machine learning modeling[J]. Journal of Petroleum Science and Engineering, 2018, 174(2): 22–45. doi: 10.1016/j.petrol.2018.11.076
- [14] CLAR H, MONACO A. Data-driven approach to optimize stimulation design in Eagle Ford Formation[C]. URTEC 2019-224-MS, 2019. doi: 10.15530/urtec-2019-224
- [15] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[C]. Long Beach: 31st Conference on Neural Information Processing Systems, 2017. doi: 10.48550/arXiv.1705.07874
- [16] 孔钦, 叶长青, 孙赞. 大数据下数据预处理方法研究[J]. 计算机技术与发展, 2018, 28(5): 1–4. doi: 10.3969/j.issn.1673-629X.2018.05.001
KONG Qin, YE Changqing, SUN Yun. Research on data preprocessing methods for big data[J]. Computer Technology and Development, 2018, 28(5): 1–4. doi: 10.3969/j.issn.1673-629X.2018.05.001
- [17] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
ZHOU Zhihua. Machine learning[M]. Beijing: Tsinghua University Press, 2016.
- [18] 纪守领, 李进锋, 杜天宇. 机器学习模型可解释性方法、应用与安全研究综述[J]. 计算机研究与发展, 2019, 56(10): 2071–2096. doi: 10.7544/j.issn1000-1239.2019.20190540
JI Shouling, LI Jinfeng, DU Tianyu. Survey on techniques, applications and security of machine learning interpretability[J]. Journal of Computer Research and Development,

- 2019, 56(10): 2071–2096. doi: 10.7544/issn1000-1239.-2019.20190540
- [19] MOLNAR C. Interpretable machine learning[M/OL]. [2022-03-29]. <https://christophm.github.io/interpretable-ml-book>, 2022.
- [20] GUIDOTTI R, MONREALE A, RUGGIERI S, et al. A survey of methods for explaining black box models[J]. ACM Computing Surveys, 2018, 51(5): 93–100. doi: 10.1145/3236009
- [21] BAEHRENS D, SCHROETER T, HARMELING S, et al. How to explain individual classification decisions[J]. Journal of Machine Learning Research, 2010, 11(6): 1803–1831. doi: 10.1007/s10846-009-9348-4
- [22] 冯德成, 吴刚. 混凝土结构基本性能的可解释机器学习建模[J]. 建筑结构学报, 2022, 43(4): 228–238. doi: 10.14006/j.jzjgxb.2020.0491
- FENG Decheng, WU Gang. Interpretable machine learning-based modeling approach for fundamental properties of concrete structures[J]. Journal of Building Structures, 2022, 43(4): 228–238. doi: 10.14006/j.jzjgxb.2020.0491



周德胜, 1963 年生, 男, 汉族, 四川武胜人, 教授, 主要从事非常规油气水力压裂提高采收率等方面的研究工作。E-mail: dzhou@xsyu.edu.cn



蔡文斌, 1981 年生, 男, 汉族, 甘肃张家川人, 副教授, 主要从事非常规油气水力压裂提高采收率等方面的研究工作。E-mail: caiwenbin@xsyu.edu.cn



李宪文, 1963 年生, 男, 汉族, 河北沧县人, 教授级高级工程师, 主要从事油气井压裂改造及增产技术等方面的研究工作。E-mail: lxw_cq@petrochina.com.cn



何明舫, 1980 年生, 男, 汉族, 江西九江人, 高级工程师, 主要从事油气井压裂改造及增产技术等方面的研究工作。E-mail: hemf_cq@petrochina.com.cn

作者简介



马先林, 1966 年生, 男, 汉族, 浙江东阳人, 教授, 主要从事油气田开发等方面的教学和研究工作。E-mail: Xianlinm@xsyu.edu.cn

编辑: 张云云

编辑部网址: <http://zk.swpuxb.com>