

基于随机森林的超深高压气井出砂程度分类预测方法

杨凤来^{1,2,3}, 陈乃东^{1,2,3}, 任龙⁴, 屈元基^{1,2,3}, 代力^{1,2,3}, 孙健⁴

1. 中国石油天然气股份有限公司塔里木油田公司 新疆库尔勒 841000
2. 中国石油天然气集团有限公司超深层复杂油气藏勘探开发技术研发中心 新疆库尔勒 841000
3. 新疆超深油气重点实验室 新疆库尔勒 841000
4. 西安石油大学石油工程学院 陕西西安 710065

通讯作者: Email: renlong@xsyu.edu.cn

项目支持: 中石油股份重大专项“超深层油气藏开发关键技术研究与应用”(2023ZZ14YJ03); 中石油股份重大专项“超深层油气藏提高采收率关键技术研究与应用”(2023ZZ14YJ04); 国家自然科学基金“基于多源随钻信息与机器学习的储层岩性实时智能识别研究”(52304036); 陕西省自然科学基金基础研究计划项目“基于流固全耦合的致密油体积压裂井工厂立体开发模式探索”(2023-JC-YB-344)

引用: 杨凤来, 陈乃东, 任龙, 等. 基于随机森林的超深高压气井出砂程度分类预测方法[J]. 油气井测试, 2024, 33(6): 9-16.

Cite: YANG Fenglai, CHEN Naidong, REN Long, et al. Classification and prediction method of sand production degree for ultra-deep and high pressure gas wells based on Random Forest machine learning[J]. Well Testing, 2024, 33(6): 9-16.

摘要 针对超深高压气井出砂程度分类预测涉及影响因素多、需要根据经验判别和耗时耗力等问题, 提出了一种基于随机森林模型的超深高压气井出砂程度分类预测方法。首先, 选取塔里木盆地克拉苏构造带裂缝性低孔砂岩气藏 93 口气井的地质、完井和开发特征因素, 建立出砂程度样本数据集; 利用 Pearson 相关系数统计分析方法, 揭示超深高压气井出砂主控因素, 形成机器学习筛选数据集; 在此基础上, 进行随机森林分类模型适用性及可靠性验证, 得到超深高压气井出砂程度随机森林分类决策树。应用结果表明, 地质、完井和开发三大类 8 个特征因素是导致气井出砂程度差异的主要控制因素, 其与出砂程度的相关系数绝对值均大于 0.2; 随机森林融合以基尼指数为特征评价指标的 CART 分类决策算法, 对不同测试集样本数量均具备良好的稳定性和较高的准确率, 适用于超深高压气井出砂程度分类预测问题。相关研究思路及结果可为超深复杂油气储层高效开发及出砂井分类治理提供一定的方法指导。

关键词 超深高压气藏; 出砂程度; Pearson 相关系数; 主控因素; 决策树; 随机森林; 分类预测模型**中图分类号**: TE377 **文献标识码**: A **DOI**: 10.19680/j.cnki.1004-4388.2024.06.002

Classification and prediction method of sand production degree for ultra-deep and high pressure gas wells based on Random Forest machine learning

YANG Fenglai^{1,2,3}, CHEN Naidong^{1,2,3}, REN Long⁴, QU Yuanji^{1,2,3}, DAI Li^{1,2,3}, SUN Jian⁴

1. Tarim Oilfield Company, PetroChina, Korla, Xinjiang 841000, China
2. R & D Center for Ultra-Deep Complex Reservoir Exploration and Development, CNPC, Korla, Xinjiang 841000, China
3. Xinjiang Key Laboratory of Ultra-deep Oil and Gas, Korla, Xinjiang 841000, China
4. College of Petroleum Engineering, Xi'an Shiyou University, Xi'an, Shaanxi 710065, China

Abstract: Aiming at the problems that the classification prediction of sand production degree of ultra-deep and high pressure gas wells involves a variety of influencing factors, which requires expert experience, and is time-consuming and labor-intensive, a classification prediction method of sand production degree based on Random Forest (RF) model is proposed. Firstly, the geological, completion and development characteristics of 93 gas well in Kelasu structural belt of Tarim Basin are selected to establish the sample dataset of sand production degree. Using the Pearson correlation coefficient statistical analysis method, reveal the main controlling factors of sand production in ultra deep and high-pressure gas wells, and form a machine learning screening dataset. On this basis, the applicability and reliability of the RF classification model were verified, and a decision tree for the RF classification of sand production degree in ultra deep and high-pressure gas wells was obtained. The application results show that the eight characteristic factors of geology, well completion, and development are the main controlling factors for the difference in sand production degree of gas wells, and their absolute correlation coefficients with sand production degree are all greater than 0.2. RF, a machine learning method that integrates the CART classification decision algorithm with Gini index as the feature evaluation index, which has good stability and high accuracy for

different sample sizes of test sets, and is suitable for the classification and prediction of sand production degree in ultra deep and high-pressure gas wells. The relevant research ideas and results can provide certain methodological guidance for the efficient development of ultra-deep and complex oil and gas reservoirs and the classification and treatment of sand producing wells.

Keywords: ultra-deep and high pressure gas reservoir; sand production degree; Pearson correlation coefficient; primary controlling factors; decision tree classification; Random Forest (RF); classification prediction model

随着天然气需求量的日益攀升,超深高压裂缝性低孔砂岩气藏因其储量巨大、产能高,已经成为非常规天然气勘探开发的重要领域之一。塔里木盆地库车坳陷克拉苏构造带克深和博孜大北气田具有储层致密、裂缝发育和水体活跃等特点,是典型的超深层高压裂缝性低孔砂岩气藏^[1-3]。目前,该类气藏在开发过程中面临较严峻的气井出砂问题,出砂井比例已超 50%,因砂埋砂堵造成气井减产停产的问题突出。但不同的气井出砂程度对应着完全不同的出砂管理策略^[4-6]。在很多情况下,气井少量出砂能够降低表皮系数改善渗流通道,提高气体的流动效率;只有在出砂程度较严重或明显影响正常生产(如导致产量显著下降、设备磨损加剧等)的情况下,才需要采取主动的防砂措施(如安装防砂筛管、使用化学固砂剂等),来有效地限制砂粒进入井筒,保障气井的稳定运行和生产安全^[7-9]。因此,如何较准确预测超深高压气井出砂程度已成为制约该类气田高效开发的最大瓶颈。

气藏出砂预测方法的研究是一个从简单的矿场统计/经验公式计算到较复杂的物理模拟/数值计算、从利用静态测井/岩石力学资料进行是否出砂的定性判断到基于岩石动态损伤过程开展气井出砂数值模拟定量预测的发展过程^[10-12]。然而,气井出砂是储层地质力学、完井改造方式和生产管理制度的综合表现结果^[13-15]。对于能够综合考虑地质、完井和开发因素的气井出砂数值模拟定量预测方法^[16-18],尽管具有理论上的优势,但其在实际应用中却面临诸多挑战,存在原始地质模型的不确定性、数据输入的繁琐复杂性以及误差随时间累积等问题,这些难题限制了该方法的广泛推广^[19-22]。

为解决上述问题,提出了一种基于随机森林机器学习方法的超深高压气井出砂程度分类预测方法。以塔里木盆地克拉苏裂缝性低孔砂岩气藏的 93 口气井的地质、完井和开发样本数据为例,首先利用 Pearson 相关系数统计分析方法,通过寻找出砂程度与多特征因素之间的相关程度,揭示超深高压气井出砂主控因素;在此基础上,采用随机森林算法构建主控因素出砂程度分类决策树,并通过实际数据验证该方法的可靠性,为高效开发同类超深

非常规油气储层提供技术支撑。

1 地质特征及样本数据集构建

研究区行政隶属于新疆维吾尔自治区阿克苏地区拜城县境内,克深和博孜大北气田储层位于塔里木盆地北部库车坳陷克拉苏构造克深断裂和拜城断裂区带,主力产气层位为白垩系巴什基奇克组,垂深 7 000~8 000 m,地层压力 116~136 MPa。白垩系巴什基奇克组沉积时,扇(或辫状)三角洲沉积体系表现为多期河道砂体垂向叠置、多个扇体平面连接,岩石类型以中、细粒岩屑长石砂岩和长石岩屑砂岩为主,各岩性段沉积相分布稳定,形成了规模巨大的且统一连通的优质储集层,孔隙度 4%~8%、渗透率 0.01~0.1 mD。塔里木盆地超深高压裂缝性低孔砂岩气藏地质条件与世界其他盆地气藏完全不同,受多期复杂构造运动影响,克拉苏构造带微裂缝及断层发育、储层空间类型多样,地质工程条件极为复杂,影响该类气藏气井出砂的地质、完井和开发因素繁多,传统的单一因素分析方法难以明晰主要控制因素及其对气井不同出砂程度的综合影响。因此,采用随机森林机器学习方法从多特征因素中快速对超深高压气井出砂程度类型进行识别。

依据井筒取出砂样成分、油嘴磨损程度和油压/产量波动幅度,对研究区 93 口气井按照出砂程度进行判断,分为未出砂、轻度出砂、中度出砂和重度出砂四种类型。每口井选取的特征参数包含地质、完井和开发三大类特征因素,其中:地质因素包括孔隙度(%)、渗透率(mD)、含气饱和度(%)、原始地层压力(MPa)、天然裂缝密度(条/m)、天然裂缝长度(m)、垂向主应力(MPa)、垂向有效应力(MPa)、岩石泊松比、岩石弹性模量(GPa)、胶结物含量(%);完井因素包括完井方式(0 代表裸眼完井、1 代表套管射孔完井)、储层改造方式(1 代表酸化压裂、2 代表 SRV 体积压裂)、射孔密度(孔/m)、储层打开程度(%);开发因素包括日产气量(10^4 m³/d)、生产压差(MPa)、日产水量(m³/d)、无阻流量(10^4 m³/d)、生产时间(mon)。最终形成由每口气井由 20 个特征因素构成的 93 个超深高压气井特征因素及出砂程度样本数据集,见表 1。

表 1 超深高压气井特征因素及出砂程度样本数据集

Table 1 Sample dataset of characteristic factors and sand production degree for ultra-deep high-pressure gas wells													
特征因素		未出砂			轻度出砂			中度出砂			重度出砂		
出砂程度		最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值
地质因素	孔隙度/%	2.33	8.5	6.16	5.18	9.3	6.9	5.42	11.1	6.86	3.89	8.5	6.34
	渗透率/mD	0.03	0.12	0.08	0.03	0.18	0.09	0.06	0.16	0.08	0.05	0.14	0.07
	含气饱和度/%	53.2	71.6	61.6	57.6	77	66.8	58.6	67.3	63.6	56.3	71.2	64.3
	原始地层压力/MPa	99	134	118	120	134	125	98	128	121	105	133	123
	天然裂缝密度/(条·m ⁻¹)	0.32	8.64	3.66	0.28	9.7	6.11	0.25	8.67	4.31	0.42	9.94	5.39
	天然裂缝长度/m	2.05	5.55	3.12	2.24	6.83	3.63	2.36	5.27	3.23	1.81	6.39	3.24
	垂向主应力/MPa	153	171	161	155	171	161	157	172	162	153	177	162
	垂向有效应力/MPa	52	79	70.7	57	77	73	60	77	71.3	60	77	71.9
	岩石泊松比	0.21	0.35	0.26	0.22	0.33	0.28	0.21	0.34	0.28	0.14	0.33	0.27
	岩石弹性模量/GPa	30.9	50.8	40.6	32.7	44.8	38.8	30.8	37.4	32.9	30.2	44.3	34.5
胶结物含量/%	7.2	10.2	9.37	6.94	9.54	8.37	4.33	7.34	6.47	3.41	8.7	5.31	
完井因素	完井方式	0	1	1	0	1	1	0	1	1	0	1	1
	储层改造方式	1	2	1	1	2	1	1	2	2	1	2	1
	射孔密度/(孔·m ⁻¹)	15	20	17	14	22	17	16	19	17	10	20	18
	储层打开程度/%	0.43	0.92	0.77	0.42	0.99	0.61	0.18	0.88	0.49	0.17	0.98	0.37
开发因素	日产气量/(10 ⁴ m ³ ·d ⁻¹)	12	68.2	30.3	7.6	82.8	58.5	16.4	63.5	41.3	6	87.4	40.3
	生产压差/MPa	3.9	19.1	9.19	10.6	37.9	17.9	17.6	39.6	25.1	12.4	43.7	27.4
	日产水量/(m ³ ·d ⁻¹)	0.3	72.1	18.8	1.4	97.3	24.2	3.15	18.3	10.8	1	140	26.4
	无阻流量/(10 ⁴ m ³ ·d ⁻¹)	45	161	120	53	176	111	28	153	110	35	159	103
	生产时间/mon	1	32	16	4	28	10	10	29	17	8	43	25

2 气井出砂分类预测理论方法

基于 Pearson 相关系数分析方法寻找超深高压气井出砂的主要控制特征因素,在此基础上利用随机森林分类决策机器学习算法建立超深高压气井出砂程度分类预测模型。

2.1 Pearson 相关系数分析

Pearson 相关系数是数据分析中一种常用的统计指标,被广泛应用来度量同一变量集合中的两个变量线性相关性的强弱。两个样本变量之间的 Pearson 相关系数定义为这两个特征变量的协方差与二者标准差积的商,即

$$r_{XY} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (1)$$

式中: X 、 Y 为样本数据集中任意两个特征变量; \bar{X} 、 \bar{Y} 分别为特征变量 X 、 Y 的平均值; n 为样本水平; r_{XY} 代表 Pearson 相关系数,其取值范围为 $[-1,1]$;取 1 时表示两个特征变量之间具有完全线性正相关的关系,取-1 时表示二者之间具有完全线性负相关的关系,0 值表示二者之间没有线性相关关系。

2.2 随机森林分类决策算法

随机森林(Random Forest, RF)是一个用随机方式建立的包含多个决策树的分类器,作为非线性有监督分类模型,属于一种集成机器学习——引导聚集算法(Bootstrap aggregating, 简称 Bagging 算法)。其基本思想是从随机的特征空间中建立多个决策树的集群,每个决策树在对应数据子集上训练并拟合最佳决策边界(最优特征决定节点划分),最后通过对多个决策树的预测结果进行投票来得到最终的预测结果。随机森林算法具有处理高维数据精度高、新鲜样本泛化性好和异常数据鲁棒性强等特点,广泛用于解决实际工程应用的复杂分类问题。

决策树的构建方法包括 ID3(Iterative Dichotomiser 3)、C4.5(Classifier 4.5)和 CART(Classification and Regression Tree)算法,其特征选择的评价指标分别由信息增益、信息增益率和基尼指数根据特征相关性或特征空间筛选出来。由于 ID3 和 C4.5 算法在进行特征选择时使用的信息增益和信息增益率存在特征偏好取值类别过多或过少的问题,笔者选取以基尼指数为评价指标的 CART 算法进行决策树构建。CART 算法用二叉树作为模型结构,通过递归操作不断地对决策属性进行最佳的二元划分,同时利用验证数据对树模型进行优化。

基尼指数 $gini(D)$ 表示在样本集合 D 中一个随机选中的样本被分错的概率。基尼指数值越小,代

表从原始数据集中取出的样本子集被错误分类的概率就越低;反之,错误分类的概率就越高。假设样本共有 k 个类别,样本点属于第 k 类别的概率为 p_k ,即当前样本集合 D 中第 k 类别样本 C_k 所占的比例,其概率分布的基尼指数可表示为^[23]

$$gini(D) = \sum_{k=1}^k p_k(1 - p_k) = 1 - \sum_{k=1}^k p_k^2 = 1 - \sum_{k=1}^k \left(\frac{|C_k|}{|D|} \right)^2 \tag{2}$$

式中: D 表示样本集合; $gini(D)$ 表示集合 D 的基尼指数; k 表示样本类别; p_k 表示样本点属于第 k 类别的概率; C_k 表示集合 D 中第 k 类别的样本数量。

对于多分类问题,如果根据某一特征因素 A^i ($i=1\sim 20$) 是否取某一可能值 A_j ,则样本集 D 可分为 D_1 和 D_2

$$\begin{cases} D_1 = \{ (x,y) \in D | A^i = A_j \} \\ D_2 = D - D_1 \end{cases} \tag{3}$$

基尼指数 $gini(D, A^i=A_j)$ 表示集合 D 中某一特征因素 A^i 经 $A^i=A_j$ 分割后样本被分错的概率。则在特征因素 $A^i=A_j$ 的条件下,集合 D 的基尼指数为

$$gini(D, A^i = A_j) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2) \tag{4}$$

最后选择使得划分后基尼指数最小的候选特

征因素 A^i 作为最优划分特征因素 A^* ,即

$$A^* = \operatorname{argmin} gini(D, A^i = A_j) \quad (i = 1 \sim 20) \tag{5}$$

将超深高压气井特征因素及出砂程度样本数据集采用 Bagging 算法有放回地重复随机抽样,每次抽样后组成含有 N 个样本的独立训练数据集,重复 T 次抽样可得到 T 个训练样本数据集;对于每个训练数据集,从其特征变量中随机选取 M 个特征,按照上述特征选择的节点评价计算指标,将影响出砂的每种特征因素分别进行评估,逐步递归与迭代组成该训练数据集对应的决策树;在每个训练数据集上采用随机森林分类算法独立地训练出 T 个预测结果,保留投票多的决策树,集合后生成由多个分类树组成的随机森林,最终形成符合研究区地质特征及样本数据集内在规律且高效的超深高压气井出砂程度分类预测模型。

3 超深高压气井出砂主控因素分析

基于前述超深高压气井特征因素及出砂程度样本数据集,以气井出砂程度为目标响应变量,利用 Pearson 相关系数统计分析方法,分别计算每个特征因素与目标响应变量之间的相关程度,计算超深高压气井 20 个特征因素与出砂程度的 Pearson 相关系数,见表 2。

表 2 超深高压气井特征因素 Pearson 相关系数

Table 2 Pearson correlation coefficient of characteristic factors for ultra-deep high-pressure gas wells

特征因素		相关程度	特征因素		相关程度
地质因素	孔隙度	0.051	完井因素	完井方式	0.033
	渗透率	-0.129		储层改造方式	0.334
	含气饱和度	0.166		射孔密度	0.215
	原始地层压力	0.169		储层打开程度	-0.727
	天然裂缝密度	0.201	-----		
	天然裂缝长度	0.016	特征因素		相关程度
	垂向主应力	0.114	开发因素	日产气量	0.127
	垂向有效应力	0.064		生产压差	0.731
	岩石泊松比	0.102		日产水量	0.094
	岩石弹性模量	-0.543		无阻流量	-0.192
	胶结物含量	-0.885		生产时间	0.651

在此基础上,进行影响超深高压气井出砂程度的特征因素重要程度排序,结果见图 1。

从影响出砂程度的特征因素重要性排序结果可以看出,超深高压气井出砂程度的主控因素包括地质、完井和开发三大类 8 个特征因素:胶结物含量、生产压差、储层打开程度、生产时间、岩石弹性模量、储层改造方式、射孔密度以及天然裂缝密度,

其与出砂程度的相关系数绝对值均大于 0.2。①地质特征因素是引起出砂程度差异的内在原因,其中胶结物含量影响程度最高,其次为岩石弹性模量和天然裂缝密度。塔里木盆地超深高压裂缝性低孔砂岩气藏储层胶结物类型以方解石、白云石为主,如果天然微裂缝较发育、颗粒的胶结程度及成岩强度较低,将导致岩石骨架颗粒承受的压应力较大,

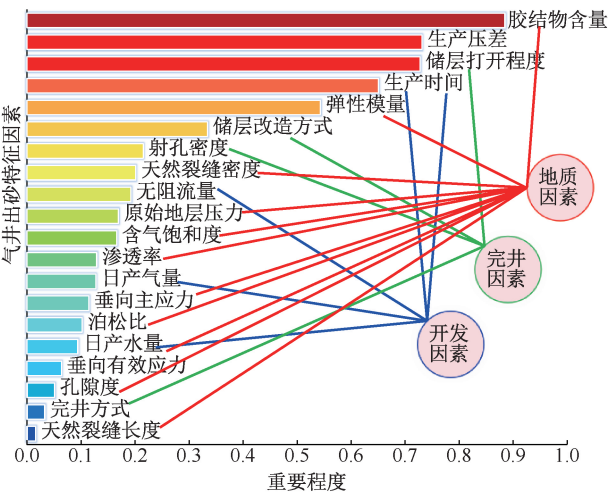


图 1 影响出砂程度的特征因素重要程度排序结果
Fig. 1 Ranking results of the importance of characteristic factors affecting the degree of sand production

近井地层垮塌出砂风险较大。②完井改造程度对出砂差异有一定影响,其中储层打开程度影响较大,其次为改造方式和射孔密度。研究区储层微裂缝较发育,较大程度的压裂改造方式形成的裂缝网络体系会进一步破坏储层力学平衡环境,加之如果储层打开程度和射孔密度较低引起气体流速较高,造成对岩石骨架的剪切破坏更易引发气井出砂。③开发因素是引起出砂程度差异的外在助推,其中生产压差的影响程度最高,其次为生产时间。气井生产压差过大易造成近井区域裂缝内支撑剂破碎、地层微粒运移和过早见水等,随着生产时间的推移,这些不利因素均会加大气井出砂趋势,直观表现为气井生产过程中油压的剧烈波动和产量的不可逆下降,矿场上常利用优化油嘴尺寸来控制气井生产压差,使其保持在出砂临界压差范围内。

4 超深高压气井出砂程度分类决策树

基于出砂主控因素分析结果,将影响出砂程度的 8 个主要特征因素和目标变量(气井出砂程度)组成超深高压气井出砂程度分类的机器学习筛选数据集。

4.1 随机森林分类模型可靠性验证

为了说明随机森林集成机器学习算法在超深高压气井出砂程度分类预测方面的优势,对筛选数据集分别进行决策树和随机森林两种分类预测模型的可靠性验证,两种算法均选择基尼指数作为划分的依据。针对数据集进行机器学习训练前需要人为地对其划分为训练集和测试集的要求,设计测试集样本数目与原始样本数目之比,即测试集占比

分别设置为 0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9 九种方案,对训练数据集分别进行决策树和随机森林的机器学习训练,得到两种分类算法的出砂程度分类决策模型相关参数及其对测试数据集的预测准确率,对比结果见表 3。

表 3 不同数据集划分条件下两种分类模型参数及预测准确率对比
Table 3 Accuracy comparison of two classification models under different dataset partitioning conditions

测试集占比	决策树深度		决策树分支数量/棵		预测准确率/%	
	决策树模型	随机森林模型	决策树模型	随机森林模型	决策树模型	随机森林模型
0.1	6	9	13	16	100.00	100.00
0.2	6	7	13	14	84.21	94.74
0.3	6	7	13	14	82.14	82.14
0.4	6	7	9	13	78.95	89.47
0.5	7	7	9	7	80.85	80.85
0.6	6	5	9	7	75.12	76.79
0.7	5	4	7	4	62.12	83.33
0.8	6	3	5	2	66.67	70.67
0.9	3	4	2	3	60.71	69.05

在不同数据集划分条件下,决策树分类模型和随机森林分类模型对测试数据集的预测准确率对比结果表明:随机森林分类决策算法对于超深高压气井出砂程度分类问题具备明显优势。决策树的数量越多,参与训练的样本重复利用率越高,形成的分类预测模型更加健壮,最终得到的分类结果的纯度(准确率)就越高。不论测试集占比较低(小于 0.4)、训练得到的决策树深度和分支数量较多,还是测试集占比较高(大于 0.4)、训练得到的决策树深度和分支数量较少,随机森林分类决策树均可保持较稳定的分类结果,即对测试数据集的预测准确率可保持在相对较高水平,表现出其对超深高压气井出砂程度分类问题适用性及可靠性。

4.2 随机森林分类决策树构建

运用随机森林 CART 分类决策算法,以基尼指数作为特征因素划分依据,超深高压气井筛选数据集(测试集占比设置为 0.1)经过机器学习训练得到出砂程度分类预测模型,构建超深高压气井出砂程度随机森林分类决策树(预测模型),见图 2。

超深高压气井出砂程度分类决策树以生产压差为根节点,综合考虑了所有 8 个出砂主控因素(胶结物含量、生产压差、储层打开程度、生产时间、岩石弹性模量、储层改造方式、射孔密度以及天然裂缝密度),且在决策过程中部分特征因素选用多

次。决策树中 samples 代表参与该分类步骤的样本数量, value 表示四种出砂类别样本占比, class 为分类结果; 整体深度为 9, 包括 16 个非叶节点(每个非叶节点代表一个特征影响因素)和 17 个叶节点(每

个叶节点代表一个目标变量的判别值); 17 个叶节点对应 17 个决策分支, 其中划分为“未出砂”、“轻度出砂”、“中度出砂”和“重度出砂”的分支分别有 3、3、5、6 个。



图 2 超深高压气井出砂程度随机森林分类决策树

Fig. 2 Classification decision tree of sand production degree of ultra-deep high-pressure gas well based on RF

上述基于主控因素构建的超深高压气井出砂程度随机森林分类决策树, 不仅可以对研究区新投产井进行出砂程度分类预测, 还能快速辅助判断疑似出砂气井的出砂程度。若有条件, 亦可结合出砂在线监测、油嘴磨损程度或取出砂样的分析化验结果, 综合判断修正得到更为精准的气井出砂程度样本数据集。因此, 在进行实际超深高压气井出砂程度分类预测时, 需要注意不同参数来源的可靠性和多种特征因素的影响程度, 避免决策树的过度简化(欠拟合)或复杂化(过拟合); 应综合考虑矿场开发经验、室内化验测试和理论计算结果, 在搞清气井

出砂程度主控因素的基础上, 借助随机森林机器学习方法, 形成操作便捷、高效且具备实用价值的超深高压气井出砂程度分类预测方法。

5 结论

(1) 气井出砂是储层地质力学、完井改造方式和生产管理制度的综合表现, 依据研究区地质特征及完井开发动静态资料, 构建了超深高压气井特征因素及出砂程度样本数据集。

(2) 基于 Pearson 相关系数统计分析, 揭示了影响超深高压气井出砂程度的 8 个主要控制因素; 其

中:地质特征因素(胶结物含量、岩石弹性模量和天然裂缝密度)是引起出砂程度差异的内在原因,完井改造程度(储层打开程度、改造方式和射孔密度)对出砂差异有一定影响,开发因素(生产压差和生产时间)是引起出砂程度差异的外在助推。

(3)进行了随机森林模型的适用性及可靠性验证,形成了一套具备实际应用价值的超深高压气井出砂程度分类决策树(预测模型)。相关研究方法有助于超深气田制定合理的完井方案和生产制度,最大程度发挥储层地质、完井方式和生产制度之间的最佳匹配关系。

致谢:感谢中石油股份重大专项“超深层油气藏开发关键技术研究与应用”和“超深层油气藏提高采收率关键技术研究与应用”项目、国家自然科学基金“基于多源随钻信息与机器学习的储层岩性实时智能识别研究”项目和陕西省自然科学基金基础研究计划项目“基于流固全耦合的致密油体积压裂井工厂立体开发模式探索”对本文的联合资助,以及中国石油塔里木油田分公司勘探开发研究院天然气开发研究部对本文撰写给予的支持和帮助。

参考文献

- [1] 李国欣, 田军, 段晓文, 等. 大幅提高超深致密砂岩气藏采收率对策与实践——以塔里木盆地克拉苏气田为例[J]. 天然气工业, 2022, 42(1): 93-101.
LI Guoxin, TIAN Jun, DUAN Xiaowen, et al. Measures and practice for improving the recovery factor of ultradeep tight sandstone gas reservoirs: a case study of Kelasu gas field, Tarim basin [J]. Natural Gas Industry, 2022, 42(1): 93-101.
- [2] 王振彪, 孙雄伟, 肖香姣. 超深超高压裂缝性致密砂岩气藏高效开发技术——以塔里木盆地克拉苏气田为例[J]. 天然气工业, 2018, 38(4): 87-95.
WANG Zhenbiao, SUN Xiongwei, XIAO Xiangjiao. Efficient development technologies for ultradeep, overpressured and fractured sandstone gas reservoirs: a case study of the Kelasu gas field in the Tarim basin [J]. Natural Gas Industry, 2018, 38(4): 87-95.
- [3] 王珂, 张惠良, 张荣虎, 等. 超深层致密砂岩储层构造裂缝特征及影响因素——以塔里木盆地克深2气田为例[J]. 石油学报, 2016, 37(6): 715-727.
WANG Ke, ZHANG Huiliang, ZHANG Ronghu, et al. Characteristics and influencing factors of ultra-deep tight sandstone reservoir structural fracture: a case study of Keshen-2 gas field, Tarim Basin [J]. Acta Petrolei Sinica, 2016, 37(6): 715-727.
- [4] 孙贺东, 常宝华, 张静楠, 等. 裂缝性致密砂岩气藏出砂原因及对产气量的影响——以塔里木盆地克深气田

- 为例[J]. 天然气工业, 2018, 38(11): 52-58.
SUN Hedong, CHANG Baohua, ZHANG Jingnan, et al. Causes of sand production and its influence on the output of fractured tight sandstone gas reservoirs: a case study on the Keshen gas field, Tarim basin [J]. Natural Gas Industry, 2018, 38(11): 52-58.
- [5] 陈钊, 邹清腾, 贾润元, 等. 昭通太阳区块浅层页岩气井出砂原因分析及防砂治理措施[J]. 天然气勘探与开发, 2022, 45(3): 132-138.
CHEN Zhao, ZOU Qingteng, JIA Runyuan, et al. Causes and controls of sand production in shallow shale gas wells of Taiyang block, Zhaotong demonstration area [J]. Natural Gas Exploration and Development, 2022, 45(3): 132-138.
- [6] 吴燕, 唐斌, 王明, 等. 高压裂缝性致密砂岩气藏出砂影响因素及防治——以塔里木盆地迪那2气藏为例[J]. 石油钻采工艺, 2019, 41(5): 663-669.
WU Yan, TANG Bin, WANG Ming, et al. Influence factors and control of sand production in high-pressure fractured tight sandstone gas reservoirs: a case study on Dina 2 gas reservoir in the Tarim basin [J]. Oil Drilling & Production Technology, 2019, 41(5): 663-669.
- [7] 涂敖, 岳星辰, 曾小军, 等. 长宁构造页岩气井返排出砂规律和精细除砂技术研究[J]. 钻采工艺, 2020, 43(6): 65-67.
TU Ao, YUE Xingchen, ZENG Xiaojun, et al. Study on sand flowback and fine sand removal technology of shale gas well in Changning structure [J]. Drilling & Production Technology, 2020, 43(6): 65-67.
- [8] 董长银. 油气井防砂技术[M]. 北京: 中国石化出版社, 2009: 1-324.
- [9] 李江涛, 项焱伟, 陈汾君, 等. 柴达木盆地涩北气田提高采收率关键技术与发展方向[J]. 天然气工业, 2023, 43(1): 141-152.
LI Jiangtao, XIANG Yiwei, CHEN Fenjun, et al. Key technologies and prospect of EGR in the Sebei gas field in the Qaidam basin [J]. Natural Gas Industry, 2023, 43(1): 141-152.
- [10] 张小飞. 库车区块克深构造气井储层出砂预测研究[D]. 成都: 西南石油大学, 2016: 73-80.
ZHANG Xiaofei. Research on sand production prediction of gas well for Keshen structure in block Kuche [D]. Chengdu: Southwest Petroleum University, 2016: 73-80.
- [11] 袁光杰, 张玉达, 董京楠, 等. 油气井筒出砂理论技术新进展[J]. 科学技术与工程, 2023, 23(7): 2694-2704.
- [12] 罗志锋, 张楠林, 赵立强. 考虑诱导应力的压裂气井出砂预测[J]. 油气藏评价与开发, 2018, 8(1): 38

- 43.
LUO Zhifeng, ZHANG Nanlin, ZHAO Liqiang. Sand prediction of fractured gas wells by consideration of induced stress [J]. *Petroleum Reservoir Evaluation and Development*, 2018, 8(1): 38-43.
- [13] 董长银, 闫切海, 周博, 等. 弱胶结储层微观出砂形态与出砂机理可视化实验模拟研究[J]. *石油钻采工艺*, 2020, 42(2): 227-235.
DONG Changyin, YAN Qiehai, ZHOU Bo, et al. Visual experimental simulation on microscopic sand production morphologies and mechanisms in weakly consolidated reservoirs [J]. *Oil Drilling & Production Technology*, 2020, 42(2): 227-235.
- [14] 刘伟, 闫新江, 林海, 等. 中等固结砂岩出砂机理与规律实验研究[J]. *石油科学通报*, 2021, 6(1): 67-78.
LIU Wei, YAN Xinjiang, LIN Hai, et al. Experimental investigation of sand production in moderately consolidated sandstones [J]. *Petroleum Science Bulletin*, 2021, 6(1): 67-78.
- [15] 张杜杰, 康毅力, 游利军, 等. 超深致密砂岩储层裂缝壁面出砂机理及其对应力敏感性的影响[J]. *油气地质与采收率*, 2017, 24(6): 72-78.
ZHANG Dujie, KANG Yili, YOU Lijun, et al. Mechanisms of sand production from fracture wall and its effect on stress sensitivity in ultra-deep tight sandstone reservoirs [J]. *Petroleum Geology and Recovery Efficiency*, 2017, 24(6): 72-78.
- [16] WU Nengyou, LI Yanlong, CHEN Qiang, et al. Sand production management during marine natural gas hydrate exploitation: review and an innovative solution [J]. *Energy & Fuels*, 2021, 35(6): 4617-4632.
- [17] JIN Yurong, WU Nengyou, LI Yanlong, et al. Characterization of sand production for clayey-silt sediments conditioned to hydraulic slotting and gravel packing: experimental observations, theoretical formulations, and modeling [J]. *SPE Journal*, 2022, 27(6): 3704-3723.
- [18] PERERA M S A, RANJITH P G, RATHNAWEERA T D, et al. 未固结流砂地层采油过程中出砂情况模拟实验[J]. *石油勘探与开发*, 2017, 44(5): 811-816.
PERERA M S A, RANJITH P G, RATHNAWEERA T D, et al. An experimental study to quantify sand production during oil recovery from unconsolidated quicksand formations [J]. *Petroleum Exploration and Development*, 2017, 44(5): 811-816.
- [19] 马亚琴, 张宝, 钟建芳, 等. 高压高产气井出砂机理数值模拟[J]. *钻采工艺*, 2019, 42(5): 43-44.
MA Yaqin, ZHANG Bao, ZHONG Jianfang, et al. Numerical simulation of sand production mechanism in high pressure highproduction gas wells [J]. *Drilling & Production Technology*, 2019, 42(5): 43-44.
- [20] 李占东, 庞鸿, 徐金泽, 等. 基于井周应力作用下气井出砂定量预测[J]. *中国矿业大学学报*, 2020, 49(3): 523-529.
LI Zhandong, PANG Hong, XU Jinze, et al. A quantitative prediction of sand production from gas wells based on the effect of stress field around the borehole [J]. *Journal of China University of Mining & Technology*, 2020, 49(3): 523-529.
- [21] 姚先荣, 杨成, 晏凌, 等. 高温高压天然气深井开采出砂预测模型优选[J]. *钻采工艺*, 2018, 41(3): 37-40.
YAO Xianrong, YANG Cheng, YAN Ling, et al. Determination on optimum sand production prediction model for deep HPHT wells during production stage [J]. *Drilling & Production Technology*, 2018, 41(3): 37-40.
- [22] 薛世峰, 马国顺, 于来刚, 等. 流固耦合模型在定量预测油水井出砂过程中的应用[J]. *石油勘探与开发*, 2007, 34(6): 750-754.
XUE Shifeng, MA Guoshun, YU Laigang, et al. A fluid-solid coupling model and its application in quantitative sand production prediction of oil and water wells [J]. *Petroleum Exploration and Development*, 2007, 34(6): 750-754.
- [23] 聂云丽, 高国忠. 基于随机森林的页岩气“甜点”分类方法[J]. *油气藏评价与开发*, 2023, 13(3): 358-367.
NIE Yunli, GAO Guozhong. Classification of shale gas “sweet spot” based on Random Forest machine learning [J]. *Petroleum Reservoir Evaluation and Development*, 2023, 13(3): 358-367.

编辑 吴志力

第一作者简介: 杨凤来, 男, 1984 年出生, 高级工程师, 硕士研究生, 2012 年毕业于中国石油大学(北京)地质工程专业, 主要从事油气田开发地质方面的工作。电话: 15199911984, Email: yangfl-tlm@petrochina.com.cn。通信地址: 新疆库尔勒市塔指东路中国石油塔里木油田分公司勘探开发研究院, 邮政编码: 841000。