

基于大数据的苏里格砂岩储层产能主控因素分析方法

李风光¹,王思博²,李旭东²,卢伟¹,徐昊垠¹,曾凡辉³

1. 中国石油集团渤海钻探工程有限公司工程技术研究院 天津 300457

2. 中国石油集团渤海钻探工程有限公司井下技术服务分公司 天津 300283

3. 西南石油大学石油与天然气工程学院 四川成都 610500

通讯作者:Email:sdlfg@163.com

项目支持:中国石油集团渤海钻探工程有限公司重大研发项目“‘智能化’压裂方案设计及施工辅助决策系统开发”(2023ZD02F-01)

引用:李风光,王思博,李旭东,等.基于大数据的苏里格砂岩储层产能主控因素分析方法[J].油气井测试,2025,34(3):35-40.

Cite: LI Fengguang, WANG Sibao, LI Xudong, et al. A big data-based method to analyzing key factors controlling productivity in Sulige sandstone reservoirs [J]. Well Testing, 2025, 34(3): 35-40.

摘要 针对目前压裂设计过程中施工参数多依靠经验值、缺乏理论支撑、压裂海量历史数据闲置的问题,提出了基于数据驱动的压裂施工参数智能优化方法。以苏里格气田数十口井为研究对象,采用神经网络、支持向量机和随机森林方法进行产能预测模型构建,结合神经网络和粒子群优化算法,以产量最高为目标推荐液量、砂量和排量等主控施工参数。结果表明,含气饱和度、储层厚度、净液量和加砂量为研究区块的产量主控因素;基于主控因素构建了神经网络产能预测模型,其预测精度可达98.95%。该方法可实现匹配砂岩储层的最优压裂施工参数,助力砂岩储层效益开发。

关键词 苏里格砂岩储层;水力压裂;数据驱动;产能预测模型;施工参数优化;神经网络;粒子群优化

中图分类号:TE357

文献标识码:B

DOI:10.19680/j.cnki.1004-4388.2025.03.006

A big data-based method to analyzing key factors controlling productivity in Sulige sandstone reservoirs

LI Fengguang¹, WANG Sibao², LI Xudong², LU Wei¹, XU Haoyin¹, ZENG Fanhui³

1. Engineering Technology Research Institute, CNPC Bohai Drilling Engineering Co., Ltd., Tianjin 300457, China

2. Downhole Technology Service Company, CNPC Bohai Drilling Engineering Co., Ltd., Tianjin 300283, China

3. Petroleum Engineering School, Southwest Petroleum University, Chengdu, Sichuan 610500, China

Abstract: To address the current challenges in fracturing design—where operational parameters often rely on empirical values with insufficient theoretical support, and vast amounts of historical fracturing data remain underutilized—a data-driven method for intelligent optimization of fracturing parameters was proposed. Taking dozens of wells in the Sulige gas field as the research subject, productivity prediction models were constructed through neural networks, support vector machines, and random forest algorithms. By integrating neural networks with particle swarm optimization algorithms, key operational parameters—such as fluid volume, proppant amount, and displacement rate—were optimized with the objective of maximizing production. The results indicate that gas saturation, reservoir thickness, net fluid volume, and proppant volume are the principal factors controlling productivity in the study area. Based on these factors, a neural network-based productivity prediction model was developed, achieving a prediction accuracy of up to 98.95%. This method enables the identification of optimal fracturing parameters tailored to sandstone reservoirs, thereby supporting the efficient development of such reservoirs.

Keywords: Sulige sandstone reservoir; hydraulic fracturing; data-driven; productivity prediction model; operational parameter optimization; neural networks; particle swarm optimization

水平井多级分段压裂技术使得致密砂岩储层可以被经济有效的开发^[1]。各油田区块地质、工程和生产数据量大复杂,多处于零散闲置状态,数据未得到有效使用。另一方面,随着逐年开发,储层

条件逐渐变差,对于成熟区块开发,压裂工艺参数大多参考以往经验,缺乏理论研究,导致压裂工艺参数与储层条件的不匹配性逐渐增强,压裂效果将越来越差。基于大数据分析的方法目前在各个行业

发挥的作用越来越明显,如何将该方法与压裂工艺参数优化结合,助力油气田效益开发是当前亟需解决的技术难题。数据挖掘方法分析储层特征与改造效果,能够考虑的参数更多,分析效果更具说服力^[2]。智能压裂算法经过多年发展,已被用于参数优化、选段选点、施工调控等方面^[3-5]。宋宣毅等^[6]利用机器学习算法考虑地质工程等方面10项因素建模,适用于小样本预测。宋丽阳等^[7]采用神经网络算法建立模型,以产量最优优化裂缝参数组合。LU等^[8]采用神经网络建立产量预测模型,从而优化施工参数。DONG等^[9]考虑经济价值,对比粒子群优化等算法的应用效果,确定粒子群优化算法适用性最佳。现有研究仅对部分施工参数开展了单方面优化,未考虑地质工程匹配设计。

本文以苏里格气田数十口井为研究对象,采用随机森林、神经网络和支持向量机方法进行产能预测模型构建,结合神经网络和粒子群优化算法,以产能为目标进行砂岩储层主控因素分析方法的研

究,可为油气田效益开发提供一定的技术支撑。

1 砂岩储层压后产能主控因素分析

砂岩储层产能受到多因素综合影响,为后续产能预测模型构建与参数优化应用减少共线性和高维度影响,必须明确制约砂岩产能的主控因素。

1.1 储层主控因素分析方法

产量受到地质-工程等多参数综合影响,地质特征反映了储层基础物性,基于区块现状,选取储层厚度、孔隙度和含气饱和度等关键参数表征地质甜点^[10]。工程参数则保证压裂过程能够形成需要的裂缝形态,因此,选取净液量、用液强度、加砂量、加砂强度、平均砂比以及排量表征工程参数。

目前的产量劈分方法基于储层厚度、孔隙度和含气饱和度之积,称为总参数团,本文采用总参数团方法进行产量劈分,获取劈分后米无阻流量,采用劈分后米无阻流量作为评价指标,以苏里格数十口井为研究对象,构建样本数据库如表1所示。

表1 部分样本数据库
Table 1 Partial sample database

井序号	层位	储层厚度 /m	孔隙度 /%	含气饱和 度/%	用液 强度	加砂 强度	平均砂比 /%	前置液比 例/%	排量 /(m ³ ·min ⁻¹)	米无阻流量 /(10 ⁴ m ³ ·(d·m) ⁻¹)
1	石盒子组	12.70	5.48	31.05	29.75	3.15	19.00	37.8	5.00	0.36
2	石盒子组	15.70	10.42	51.61	29.11	3.18	19.60	37.3	5.00	0.38
3	石盒子组	5.60	9.02	58.48	47.13	4.46	22.90	36.5	4.00	0.90
4	石盒子组	32.10	11.19	57.75	19.71	2.80	23.00	33.1	8.00	0.33
5	石盒子组	11.00	9.51	42.95	34.06	4.09	21.30	37.4	5.10	0.69
6	石盒子组	14.9	14.0	62.4	69.1	6.22	12.8	38.2	8	0.68
...
13	石盒子组	8.3	17.2	66.1	47.4	3.61	16.6	36.6	3.5	1.12
14	石盒子组	31.2	10.9	52.0	15.1	1.99	22	32.8	5	0.32
15	石盒子组	25.7	10.5	58.0	30.9	3.89	20.4	39.0	6	0.47
16	石盒子组	19.3	14.0	62.4	49.8	6.22	12.8	24.7	8	0.53
17	石盒子组	48	10.0	42.0	18.9	2.17	19.5	32.4	8.1	0.31
18	山西组	12.7	5.5	31.1	29.7	3.15	19	22.8	5	0.36
19	山西组	15.7	10.4	51.6	29.1	3.18	19.6	31	5	0.38
20	山西组	5.6	9.0	58.5	47.1	4.46	22.9	34.1	4	0.90
21	山西组	32.1	11.2	57.8	19.7	2.80	23	39.6	8	0.33
22	山西组	11	9.5	43.0	34.1	4.09	21.3	38.7	5.1	0.69
23	山西组	10.3	7.6	51.8	64.1	5.63	22.2	40.5	6.14	0.59
24	山西组	5.7	8.2	28.8	73.3	3.86	13.5	32.5	4.5	1.30
...
32	山西组	48.00	10.02	41.98	18.87	2.17	19.50	38.5	8.10	0.31

1.2 相关性分析

为明确各因素之间的关系,开展斯皮尔曼相关性分析,两组随机变量 $X = (x_1, x_2, \dots, x_n)$ 和 $Y = (y_1, y_2, \dots, y_n)$ 之间的斯皮尔曼等级相关系数为^[11]

$$\rho = \frac{\sum_{i=1}^m (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^m (a_i - \bar{a})^2} \sqrt{\sum_{i=1}^m (b_i - \bar{b})^2}} \quad (1)$$

式中: ρ 为等级相关系数, $0 < \rho \leq 1$ 时呈正相关,

- 1 ≤ ρ < 0 时呈负相关, ρ = 0 时无相关性; m 为样本数量; a_i 和 b_i 分别为 x_i 和 y_i 的排名。

相关性分析结果如图 1 所示, 地质特征中储

层厚度、含气饱和度对产量影响靠前, 为主控因素。施工参数中的加砂量、净液量为产量的主控因素。

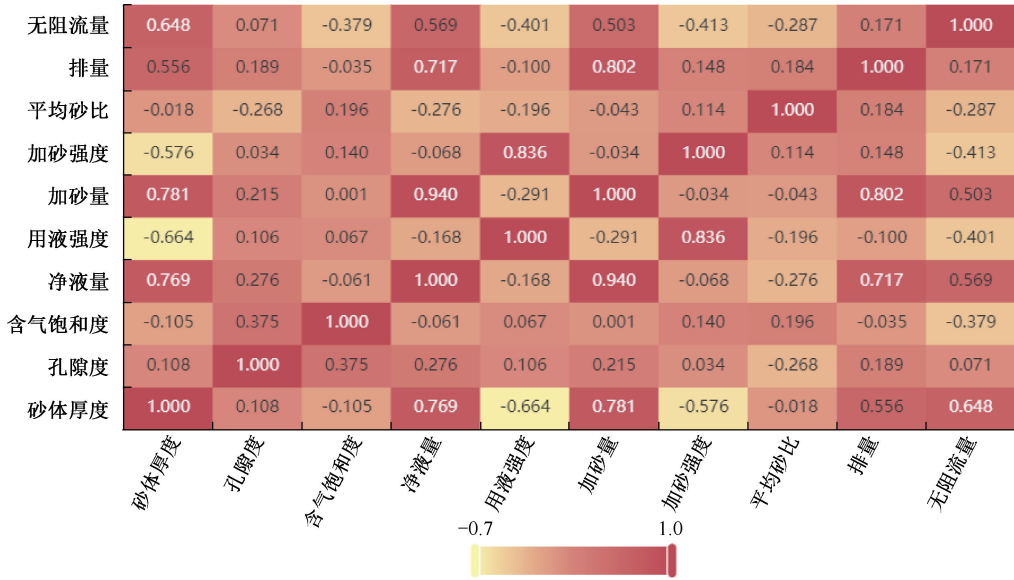


图 1 产量影响因素相关性分析图

Fig. 1 Correlation analysis diagram of production influencing factors

2 储层产量预测

产量预测模型精度会严重影响参数优化的准确性, 传统产量预测模型无法考虑参数与参数之间的共线性影响, 采用机器学习算法学习主控因素与产量之间的关系, 可有效提升精度。

2.1 基于随机森林算法的产量预测

RF 是机器学习中最流行的算法之一。Ho (1995) 提出的思想在 Breiman (2001) 中得到了改进, 至今仍被认为是构造 RF 的参考算法, 具有精确度高、抗噪、不易过拟合等优势, 通常用于基因表达, 预测比特币价格, 工业过程中的故障分类, 数据挖掘, 入侵检测等领域^[12]。

给定分类器的集合 $h_{1(x)}, h_{2(x)}, \dots, h_{k(x)}$, 并且利用从随机向量 Y, X 的分布中随机抽取的训练集, 将函数定义为

$$mg(X, Y) = \underset{j \neq Y}{\text{av}_k} I(h_k(X) = Y) - \max_{j \neq Y} \underset{j \neq Y}{\text{av}_k} I(h_k(X) = j). \quad (2)$$

其中 $I(\cdot)$ 为指示函数。边际越大, 分类的可信度越高。泛化误差由下式给出:

$$PE^* = P_{X, Y}(mg(X, Y) < 0) \quad (3)$$

其中下标 X, Y 指示概率是在 X, Y 空间上。在随机森林中, $hk(X) = h(X, k)$ 的范围内。基于强大数定律、树结构得出:

树数增加到一定程度, 所有序列 $1, \dots, PE^*$ 将会收敛于

$$P_{X, Y}(P_{\Theta}(h(X, \Theta) = Y) - \max_{j \neq Y} P_{\Theta}(h(X, \Theta) = j) < 0). \quad (4)$$

2.2 基于神经网络的产量预测

神经网络本身适用于并行处理信息, 适合非线性数据, 常被用于数据分类处理, 数据拟合, 基本原理如图 2 所示。神经网络算法准确性高, 储存与学习能力强, 还可用于联想记忆, 缺陷在于对数据量要求高, 数据黑箱, 无法查看中间过程, 可信度受到影响, 需要足够长的学习时间。

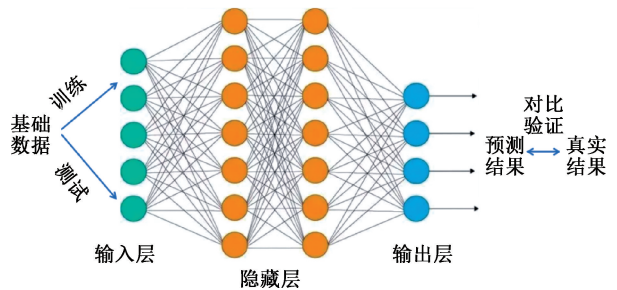


图 2 神经网络基本原理

Fig. 2 Basic principles of neural networks

本文设计的神经网络由输入层、隐含层、输出层三层构成, 每层的节点个数各不相同, 而且会随着进行不同区域的建模时发生改变。

2.3 基于支持向量机的产量预测

支持向量回归(SVR)将高维度数据转换为低维,通过寻找超平面,让集合内所有的数据到该平面的距离最近完成最优化设计,基本原理如图3所示。

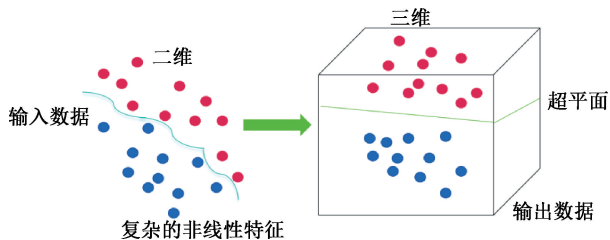


图3 支持向量回归基本原理

Fig. 3 Basic principles of support vector regression (SVR)

当数据完全落在模型平面上时,该平面的损失值为0。对于小样本条件下的机器学习问题,SVR可以解决高维问题,并且可以避免神经网络结构选择和局部极值问题。

3 压裂施工参数优化与现场应用

由于不同的机器学习算法适用条件不同,针对苏里格气田石盒子组和山西组进行产量预测模型构建,测试不同机器学习算法的准确率,并优选精度最高的算法用于后续参数优化,并以X井为例完成参数优化现场应用。

3.1 产量预测模型

考虑储层厚度、孔隙度、渗透率等10项产量影响因素,选取2口井为测试集、其余为训练集,分别对石盒子组和山西组建立三种机器学习方法的产量预测模型。设置神经网络模型隐藏层数量为100,隐藏层大小为50,最大迭代次数为1000;支持向量回归模型核函数为径向基函数和函数,正则化参数为1.5,惩罚因子为0.1;设置随机森林模型的树数量为100,树深度为50,节点随机数为1。石盒子组三种产量预测模型训练集和测试集预测精度如图4所示,其中,神经网络模型训练集平均预测符合率为91.97%,测试集平均预测符合率为94.33%;支持向量回归模型训练集平均预测符合率为72.35%,测试集平均预测符合率为92.41%;随机森林模型训练集平均预测符合率为88.76%,测试集平均预测符合率为94.01%。山西组的三种产量预测模型训练集和测试集预测精度如图5所示,神经网络模型训练集平均预测符合率为98.31%,测试集平均预测符合率为98.95%;支持向量回归模型训练集平均预测符合率为71.36%,测试集平均预测符合率为98.85%;随机森林模型训练集平均预测符合率为88.74%,测试集平均预测符合率为91.25%。由此可以看出神经网络模型预测精度最高,表明该模型更适合开展压裂参数优化。

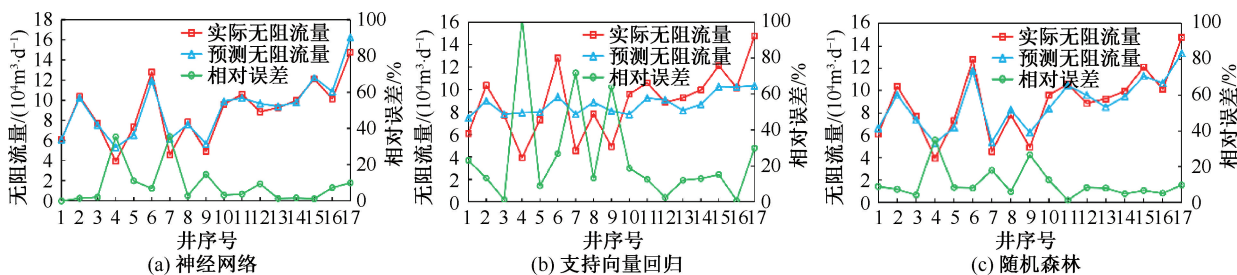


图4 石盒子组产量预测模型精度

Fig. 4 Prediction accuracy of the Shihezi formation production model

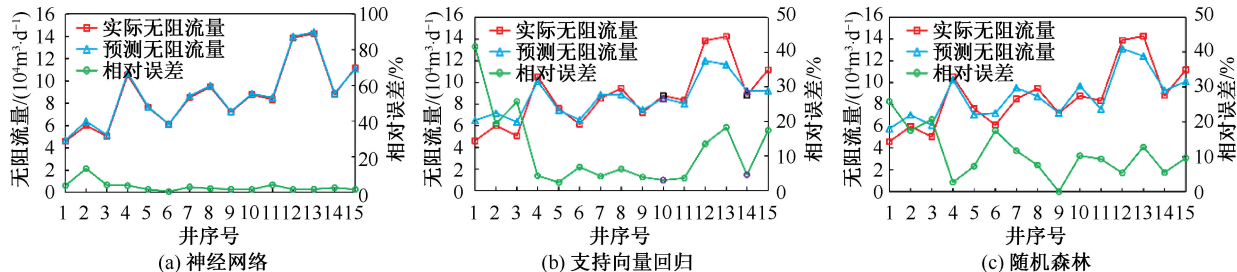


图5 山西组产量预测模型精度

Fig. 5 Prediction accuracy of the Shanxi formation production model

3.2 粒子群优化算法

粒子群优化算法通过随机搜索种群,利用少量

参数即可寻找全局最优。对于包含 N_s 个粒子的 M_s 维搜索空间,每一个粒子包含三个向量信息,分别

为粒子运动速度 $v_i = (v_{i1}, v_{i2}, \dots, v_{im_s})$, 粒子位置 $X_i = (X_{i1}, X_{i2}, \dots, X_{im_s})$ 和其最佳位置 $O_i = (O_{i1}, O_{i2}, \dots, O_{im_s})$, 粒子群全局最优位置为 $O_L = (O_{L1}, O_{L2}, \dots, O_{Lm_s})$ 。粒子速度、位置迭代规则为^[13]

$$v_{im_s}(t+1) = \tau v_{im_s} + \lambda_1 \vartheta_1 [O_{im_s}(t) - X_{im_s}(t)] + \lambda_2 \vartheta_2 [O_{Lm_s}(t) - X_{im_s}(t)] \quad (5)$$

$$X_{im_s}(t+1) = X_{im_s}(t) + v_{im_s}(t)$$

式中: v_{im_s} 为当前速度, $i = 1, 2, \dots, N_s, m_s = 1, 2, \dots,$

$M_s, v_{im_s} \in [v_{\min, M_s}, v_{\max, M_s}]; \tau$ 为惯性因子; λ_1, λ_2 为学习因子; ϑ_1, ϑ_2 为 $[0, 1]$ 中的随机数; O_{im_s} 为当前最佳位置; O_{Lm_s} 全局最佳位置。

3.3 施工参数优化现场应用

以地质参数为基础, 分别采用山西组和石盒子组产量预测模型对各层开展施工参数优化, 各组施工方案与对应的预测米无阻流量如图6所示, 根据各小层厚度设置施工参数搜索范围, 对应地质参数与优化结果如表2所示。

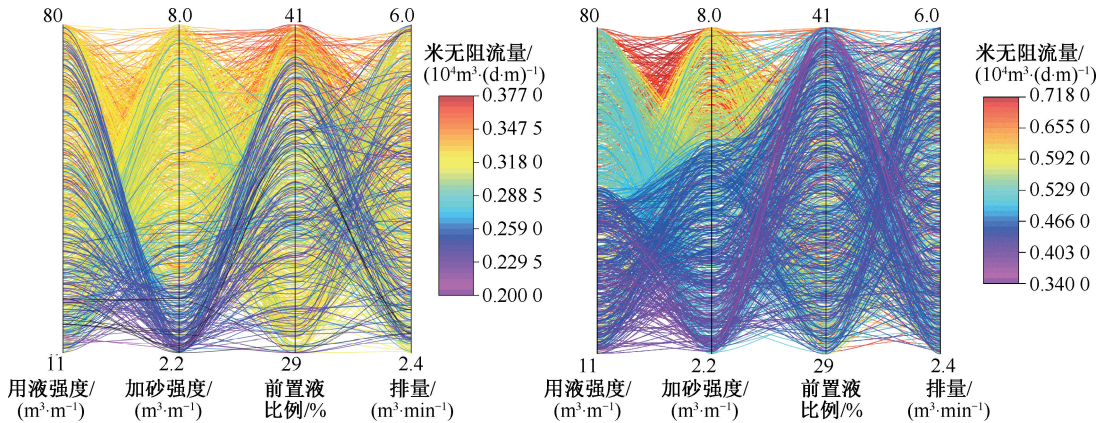


图6 施工参数优化方案及预测无阻流量

Fig. 6 Optimization scheme of fracturing parameters and prediction of absolute open flow potential

表2 应用井施工参数优化结果

Table 2 Optimization results of applied well construction parameters

层号	层位	砂体厚度 /m	孔隙度 /%	含气饱和度 /%	用液强度 /(m ³ · m ⁻¹)	加砂强度 /(m ³ · m ⁻¹)	前置液比例 /%	施工排量 /(m ³ · min ⁻¹)	预测米无阻流量/ (10 ⁴ m ³ · (d · m) ⁻¹)
37、39	盒8上	8.9	9.73	51.18	68.8	5.0	36.0	5.8	0.711
50	山1	9.1	7.74	28.65	62.1	5.0	37.0	5.1	0.371

由图6可知, 山西组米无阻流量对排量和净液量较敏感, 石盒子组米无阻流量仅对净液量敏感。由施工参数与优化结果表2所示, 37、39号小层推荐用液强度 62.1 m³/m, 加砂强度 5.0 m³/m, 前置液比例 37.0%, 施工排量 5.1 m³/min; 50号小层推荐用液强度 68.8 m³/m, 加砂强度 5.0 m³/m, 前置液比例 36.0%, 施工排量 5.8 m³/min; X井在推荐方案下预测无阻流量 9.71 × 10⁴ m³/d。

本文以苏里格气田数十口井的地质、工程等生产数据为基础, 构建了产能预测模型, 结合神经网络和粒子群优化算法, 以产量最高为目标进行了砂岩储层主控因素分析方法的研究, 该方法解决了大量压裂历史数据为施工参数优化提供了可行性, 可为油气田效益开发提供重要的技术支撑。

4 结论

(1) 以苏里格气田某区块单井作为研究对象,

分析了包含地质特征和施工参数的10项因素与无阻流量之间的斯皮尔曼相关性。

(2) 基于数据驱动的产量预测方法中, 各机器学习方法有各自的适用性, 针对本区块井神经网络产能预测模型精度更高, 可达98.95%。

(3) 基于神经网络产量预测模型和粒子群优化算法, 建立数据驱动的施工参数优化模型, 实现产能最大化优化, 完成地质-工程匹配的施工参数差异化设计, 为单井提质增效提供坚实的技术支撑。

致谢: 感谢渤海钻探工程公司重大研发项目“‘智能化’压裂方案设计及施工辅助决策系统开发”(2023ZD02F-01)资金支持; 感谢科技处及项目组研发及现场推广人员对本研究的大力支持。

参考文献

[1] LU C H, JIANG H Q, QU S Y, et al. Hydraulic fracturing

- design for shale oils based on sweet spot mapping: a case study of the Jimusar formation in China [J]. *Journal of Petroleum Science and Engineering*, 2022, 214: 110568.
- [2] YANG R Y, QIN X Z, LIU W, et al. A physics-constrained data-driven workflow for predicting coalbed methane well production using artificial neural network [J]. *SPE Journal*, 2022, 27(3): 1531-1552.
- [3] 张世昆, 陈作. 人工智能在压裂技术中的应用现状及前景展望 [J]. *石油钻探技术*, 2023, 51(1): 69-77.
ZHANG Shikun, CHEN Zuo. Status and prospect of artificial intelligence application in fracturing technology [J]. *Petroleum Drilling Techniques*, 2023, 51(1): 69-77.
- [4] 盛茂, 李根生, 田守增, 等. 人工智能在油气压裂增产中的研究现状与展望 [J]. *钻采工艺*, 2022, 45(4): 1-8.
SHENG Mao, LI Gensheng, TIAN Shouceng, et al. Research status and prospect of artificial intelligence in reservoir fracturing stimulation [J]. *Drilling and Production Technology*, 2022, 45(4): 1-8.
- [5] 李根生, 田守增, 盛茂, 等. 智能压裂技术研究进展与前景展望 [J]. *钻采工艺*, 2025, 48(1): 1-9.
LI Gensheng, TIAN Shouceng, SHENG Mao, et al. Research progress and prospect of intelligent hydraulic fracturing technologies [J]. *Drilling and Production Technology*, 2025, 48(1): 1-9.
- [6] 宋宣毅, 刘月田, 马晶, 等. 基于灰狼算法优化的支持向量机产能预测 [J]. *岩性油气藏*, 2020, 32(2): 134-140.
SONG Xuanyi, LIU Yuetian, MA Jing, et al. Productivity forecast based on support vector machine optimized by grey wolf optimizer [J]. *Lithologic Reservoirs*. 2020, 32(2): 134-140.
- [7] 宋丽阳, 王纪伟, 刘长印. 基于 BP-GA 算法的水平井智能压裂设计方法 [J]. *断块油气田*, 2022, 29(3): 417-421.
SONG Liyang, WANG Jiwei, LIU Changyin. BP-GA algorithm assisted intelligent horizontal well fracturing design [J]. *Fault-Block Oil & Gas Field*, 2022, 29(3): 417-421.
- [8] LU C H, JIANG H Q, YANG J L, et al. Shale oil production prediction and fracturing optimization based on machine learning [J]. *Journal of Petroleum Science and Engineering*, 2022, 217: 110900.
- [9] DONG Z Z, WU L, WANG L J, et al. Optimization of fracturing parameters with machine-learning and evolutionary algorithm methods [J]. *Energies*, 2022, 15: 6063.
- [10] 孙龙德, 赵文智, 刘合, 等. 页岩油“甜点”概念及其应用讨论 [J]. *石油学报*, 2023, 44(1): 1-13.
SUN Longde, ZHAO Wenzhi, LIU He, et al. Concept and application of “sweet spot” in shale oil [J]. *Acta Petrolei Sinica*. 2023, 44(1): 1-13.
- [11] ZHAO G, DING W L, TIAN J, et al. Spearman rank correlations analysis of the elemental, mineral concentrations, and mechanical parameters of the lower cambrian Niutitang shale: a case study in the Fenggang block, Northeast Guizhou Province, South China [J]. *Journal of Petroleum Science and Engineering*, 2022, 208: 109550.
- [12] BREIMAN L. Random forests [J]. *Machine learning*, 2001, 45(1): 5-32.
- [13] FARUK O A. Simultaneous optimization of well count and placement: algorithm, validation, and field testing [J]. *SPE Journal*, 2023, 28(1): 147-172.

编辑 吴志力

第一作者简介:李风光,男,1982年出生,博士,高级工程师,2012年毕业于中国石油大学(北京)化学工程与技术专业,从事油气田储层改造增产技术研究工作。电话:022-25910750,Email:sdlfg@163.com。通信地址:天津市滨海新区大港油田三号院渤海钻探工程技术研究院,邮政编码:300280。