基于机器学习的玛湖地区水平井压裂设计优化

马俊修¹, 石胜男², 陈进¹, 张景臣², 何小东¹, 李雪晨², 郭丁菲²

1) 中国石油新疆油田分公司工程技术研究院, 新疆克拉玛依 834000;

2) 中国石油大学(北京)非常规油气科学技术研究院, 北京 102249

摘要: 中国新疆玛湖致密砂砾岩油藏非均质性强, 各井地质和施工条件参差不齐, 压裂设计方案缺乏针对性, 压裂效果的主控因素也有待进一步深化认识. 为此, 建立了一套基于机器学习的玛湖地区水平井产能预测及压裂设计优化的流程框架. 以玛湖地区75口压裂水平井为例, 针对储层和工程2类16个影响因素, 采用随机森林算法确定了压后产能主控因素, 建立了遗传算法优化的逆传播神经网络产能预测模型, 并基于此对水平井的压裂进行优化设计. 结果表明, 现场2口井在优选压裂参数方案后, 产能分别提高了9.3%和37.3%. 建立的产能预测及压裂设计优化方法可为现场施工提供指导.

关键词: 非常规油气藏; 压裂优化; 玛湖油田; 机器学习; 主控因素; 随机森林; 逆传播神经网络
中图分类号: TE343 **文献标志码:** A **doi:** 10.3724/SP.J.1249.2021.06621

Optimization of fracture design for horizontal wells in Mahu region based on machine learning

MA Junxiu¹, SHI Shengnan², CHEN Jin¹, ZHANG Jingchen²,
HE Xiaodong¹, LI Xuechen², and GUO Dingfei²

1) Institute of Oil and Gas Technology, Research Institute of Engineering and Technology, Xinjiang Oilfield Company, Petro China, Karamay 834000, Xinjiang Uygur Autonomous Region, P. R. China

2) Unconventional Oil and Gas Science and Technology Research Institute, China University of Petroleum (Beijing), Beijing 102249, P. R. China

Abstract: The Mahu tight glutenite reservoir is highly heterogeneous, and the geology and construction conditions of each well are uneven. The fracturing design scheme lacks pertinence. The main control factors of the fracturing effect need to be further understood. Therefore, we establish a set of machine learning-based process frameworks for horizontal well productivity prediction and fracturing design optimization in the Mahu area. Taking 75 fractured horizontal wells in the Mahu area as an example, based on 16 influencing factors of 2 types of geology and engineering, the random forest algorithm is used to determine the main control factors of post-fracturing productivity, and a error back Propagation (BP) neural network productivity prediction model optimized by genetic algorithm is established. Based on this, the optimal design of fracturing for horizontal wells is carried out. The results show that

Received: 2020-12-08; **Accepted:** 2021-02-21; **Online (CNKI):** 2021-06-18

Foundation: Tianshan Youth Program of Science and Technology Department of Xinjiang Autonomous Region (2018Q030); CNPC Strategic Science and Technology Project (ZLZX2020-02-07-03, ZLZX2020-02-07-05); China University of Petroleum (Beijing) Karamay Campus Fund Project (RCYJ2017B-01-003); Major Science and Technology Special Funded Projects in Karamay, Xinjiang (2018ZD001B)

Corresponding author: Associate professor ZHANG Jingchen. E-mail: jingchen120@126.com

Citation: MA Junxiu, SHI Shengnan, CHEN Jin, et al. Optimization of fracture design for horizontal wells in Mahu region based on machine learning [J]. Journal of Shenzhen University Science and Engineering, 2021, 38(6): 621-627. (in Chinese)

the productivity of two wells in the field is increased by 9.3% and 37.3% respectively after the fracturing parameter scheme is optimized. The established productivity prediction and fracturing design optimization methods can be applied to other oil fields and provide targeted guidance for on-site construction.

Key words: unconventional reservoirs; fracturing optimization; Mahu oilfield; machine learning; main control factor; random forest; back propagation (BP) neural network

中国新疆玛湖致密砂砾岩开发普遍采用“水平井+大规模体积压裂”技术,油藏具有砾石粒径变化大、岩性复杂和非均质性强的特征,导致各井的压裂条件参差不齐,且压裂后井间生产效果差异较大.明确玛湖地区水平井的压裂效果主控因素、实现产能预测和定量优化压裂参数一直是玛湖压裂水平井开发的重点和难点^[1].

近年来,机器学习方法在油气田开发中的应用日益广泛,基于机器学习方法的产能预测及压裂工艺参数优化研究成为不少学者研究的重点. WANG 等^[2]采用 GDE 全局优化算法,基于 200 口井的数据建立了经济模型,优化了财务净现值; ANI-FOWOSE 等^[3]结合支持向量机和人工神经网络,成功预测了储层孔隙度和渗透率;叶俊华等^[4]采用灰色关联法对吉 006 断块压后主控因素进行了分析;张杰等^[5]基于神经网络、决策树及聚类等方法,建立了压裂施工效果预测模型;陶亮等^[6]结合层次分析法与灰色理论建立了产能模型,并将其应用于松辽盆地致密油藏,取得了良好的现场效果.综上所述,国内外学者在建立压裂效果预测模型上开展了大量研究,但多侧重于产量预测以及主控因素分析,对进一步的压裂参数优化研究较少.

本研究基于收集的玛湖地区压裂水平井数据,通过缺失值填补、异常值识别和数据缩放等预处理工作提高数据质量,基于随机森林算法分析了影响水平井压裂效果的主控因素,建立各项特征参数与最终可采储量(estimated ultimate recovery, EUR)之间的数学模型,从而实现水平井的压裂参数定量优化,达到针对性地指导压裂方案的目的.

1 特征参数选取及预处理

压裂水平井产能不仅与孔隙度、含油饱和度(S)和储层厚度等储层参数有关,还受压裂规模和生产制度等多因素的影响.为此,本研究收集了新疆玛湖地区 75 口压裂水平井的储层参数、工程参数以及生产数据,共 18 个特征参数,特征参数分为 1 级参数和 2 级参数.其中,储层参数和工程参

数为 1 级参数.储层参数包括杨氏模量、地层压力、孔隙度、最小水平主应力、钻遇率、含油饱和度、泊松比和 I 类储层厚度共 9 个 2 级参数;工程参数包括总液量、平均砂比($\bar{\varphi}$ (砂))、前置液比例、停泵压力、施工排量、裂缝间距、总砂量和水平段长度共 8 个 2 级参数.其中, EUR 是衡量油井产能的重要参数^[7-8].

在数据分析和挖掘的过程中,数据是极其关键的要素,数据的质量对最终获得的模型或结论的影响很大^[9-10].由于各种原因,油气田现场的数据往往存在数据缺失、数错误和异常等现象.因此,在进行产能预测及压裂设计优化前,本研究先对数据进行缺失值和异常值的分析处理,以及数据预处理,提高数据集的质量,保证最后所建模型的准确性.

1.1 缺失值处理

对 75 口井 18 个特征参数的缺失率进行统计分析发现,停泵压力的缺失率为 26.7%,地层压力、泊松比、杨氏模量和最小水平主应力的缺失率为 20%,平均含油饱和度、I 类储层厚度、平均孔隙度、裂缝间距和施工排量的缺失率在 5% 以下,其他参数不存在缺失值.雷蕾等^[11]研究发现,当缺失率超过 5% 时,对分类模型的准确率会有影响,需采取填补等方法进行预处理后再进行建模.

本研究采取精细的缺失值处理策略,当缺失率 > 25%,且属性重要程度低时,直接删除该属性(停泵压力);当缺失率 < 5%,采用中位值简单填补方式(平均含油饱和度、I 类油层厚度、平均孔隙度和裂缝间距);当缺失率 > 5%,且属性重要程度高时,采用随机森林多重插补的方式(地层压力、泊松比、杨氏模量和最小水平主应力),将每个缺失值的特征作为其他特征的函数建立随机森林模型,使用该函数的值作为估算,通过变量之间的关系对缺失数据进行预测,使得填补的缺失值更加真实.

1.2 异常值识别

异常值指远离绝大多数样本点的特殊群体,也称为离群点,通常这样的数据点在数据集中都表现出不合理的特性.如果忽视这些异常值,在某些建模场景下就会导致结论的错误(如神经网络回归模

型等),所以在数据的探索过程中,有必要识别出这些异常值并进行处理^[12-13]。

本研究采用箱型图来识别异常值,即利用数据的四分位数和四分位距识别其中的异常点^[8]。图1为判断平均砂比 $\bar{\varphi}$ (砂)异常值情况的箱型图,位置A和C分别对应数据的上下四分位数(Q_3 和 Q_1)和四分位距(inter-quartile range, IQR) $Q = Q_3 - Q_1$ 。当数据小于 $Q_1 - 1.5Q$ (下限E)或大于 $Q_3 + 1.5Q$ (上限D)时,被识别为异常值^[14],如图2中的数据点f超过上限D,则视为该样本平均砂比的异常值。异常值识别后分析形成的原因,如果为数据输入错误所导致,将其修改为正常值;如果异常值错误程度过于离谱,则视为缺失值进行处理;其他异常值暂且保留,结合整体模型综合分析。

图1 箱型图识别平均砂比

Fig. 1 Box diagram to identify average sand ratio

1.3 数据标准化

由于数据类型复杂,且数据单位、属性和产能的影响方式均不同,在模型优选分析计算时需要进行如式(1)的标准化处理,以消除不同数量级和不同量纲的影响。

$$y_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (1)$$

其中, y_i 为标准化处理后样本*i*的特征输出; x_i 为样本特征; $\min(x_i)$ 为样本特征最小值; $\max(x_i)$ 为样本特征最大值。

2 压裂水平井产能预测模型建立

2.1 基于随机森林的压裂效果主控因素分析

为明确各储层参数、工程参数对产能的非线性影响程度,本研究采用随机森林方法确定产能的主控因素。随机森林是一种集成机器学习方法,在决策树的训练过程中进一步引入随机属性选择而形成

的。相比决策树,该方法泛化性更好,并能够得到各变量的相对重要度。随机森林建模时随机采样未被抽到的数据称为袋外数据集,这些数据没有参与训练集模型的拟合,可以用来检验模型的泛化能力^[15-16]。在对模型进行重要性排序时,使用相应的袋外数据计算它的袋外误差 r_1 。然后袋外数据中的某个特征顺序被随机变换,再次计算袋外误差 r_2 。假设随机森林有 N 棵树,则某个特征的重要性 I 为

$$I = \frac{1}{N} \sum_{i=1}^N (r_1 - r_2) \quad (2)$$

根据上述原理,建模计算得到玛湖地区水平井各项特征因素对产能的重要性评分,各项特征参数重要性总评分为1.0^[17]。储层参数总评分为0.550,2级储层主控参数依次为含油饱和度、孔隙度、钻遇率和地层压力,评分分别为0.175、0.152、0.051和0.046;工程参数的重要性总评分为0.45,2级施工主控参数裂缝间距、总液量、总砂量、施工排量和平均砂比分别为0.152、0.112、0.094、0.042和0.031。由此可见,玛湖地区水平井的产能受储层参数影响较大,但在进行压裂施工参数优化时,应重点考虑裂缝间距、总液量和总砂量等工程参数,以获取最大产能。

2.2 产能预测模型工作流程

逆传播(back propagation, BP)神经网络是一种按误差逆传播算法训练的多层前馈网络,根据输入的训练样本,可自动适应和确定各神经元的连接权重^[18]。BP神经网络基本思想是梯度下降法,利用梯度搜索技术,使网络的实际输出值和期望输出值的误差均方差为最小^[19-20],是目前应用最广泛的神经网络模型之一。遗传算法(genetic algorithms, GA)通过学习模拟生物进化论和自然界遗传机制,形成一种并行随机搜索最优化方法^[21]。

本研究建立了产能预测的BP神经网络模型,并使用遗传算法进行优化,模型的工作流程如图2。使用遗传算法优化神经网络模型可以获取更加准确的初始权值和阈值,神经网络模型进行优化后,可以更好地反应出输入数据与输出数据之间的关系,达到精确预测的目的。

2.3 产能预测模型的训练及验证

本研究以16个储层及施工参数作为输入特征值(表1),以EUR作为输出参数,建立特征参数与产能之间的3层BP神经网络模型,设隐含层的神经元为20个,网络训练的学习率为0.001,期望误差为0.05。通过反误差函数不断调节网络权值与阈

图 2 遗传算法优化 BP 神经网络过程

Fig. 2 BP neural network process optimized by genetic algorithm

值,直到网络输出误差逐步达到期望误差,完成网络训练.本次模型使用遗传算法的个体编码长度应包括中间层到输入层的 16×20 个连接权值,输出层到中间层的 1×20 个连接权值以及 $20 + 1$ 个阈值,遗传算法的个体编码长度最终为 $16 \times 20 + 1 \times 20 + 20 + 1 = 361$.

共收集压裂水平井 75 口,按照数量比为 85:15 划分训练集和测试集,分割后训练集为 64 口,测试集为 11 口,数据集的原始产量分布如图 3.

使用训练集数据训练模型,并基于模型在测试集上的评分评估模型性能.采用决定系数 R^2 评估模型的性能.决定系数 R^2 得分反映因变量的全部

图 3 训练样本与测试样本产量的原始分布情况

Fig. 3 The original distribution of the output of training samples and test samples

变能通过回归关系被自变量解释的比例,如式(3), R^2 越接近 1,样本中预测值和真实值越相近,表示回归分析中自变量对因变量的解释越好.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

其中, y_i 为参数的因变量值; $\hat{y}_i$ 为样本数据对应的预测值; $\bar{y}$ 为样本数据平均值.

模型训练后,为验证优化后的模型性能,将其预测测试集样本的结果与原始数据及未经优化后的 BP 模型预测结果进行对比,结果如图 4. 由图 4 可见,使用遗传算法优化后的神经网络模型(GA-BP)效果更好,该模型测试集的决定系数得分 $R^2 = 0.76$,未经优化的决定系数得分 $R^2 = 0.62$,证明该模型该模型性能优良,泛化能力好,具备一定的适用性.

图 4 测试集样本预测结果

Fig. 4 Test set sample prediction results

3 单井压裂设计优化

3.1 产能预测应用

为了进一步验证玛湖地区水平井体积压裂产能预测模型的适用性,以玛湖地区某两口待压井(井 1 和井 2)为例,开展产量预测.将其邻井数据输入产量预测模型中,验证模型的性能,同时与已投入生产的邻井相比,在储层及工程参数相似的情况下,使用模型验证储层特征相近的临井情况,可以为井 1 和井 2 提供一定参考,井 1、井 2 及邻井的特征参数如表 1,井 1 邻井与井 2 邻井预测的 EUR 分别为 54 356 t 和 25 394 t,与真实值相比,误差均在 11.3% 以下,证明该模型性能良好.使用模型预测井 1 和井 2 的 EUR 分别为 58 758 t 和 26 034 t,与其邻井相比,预测的 EUR 数值更接近.

表1 井1和井2与邻井特征参数及预测结果

Table 1 Characteristic parameters and prediction results of well 1, well 2 and adjacent wells

井号	地层压力	泊松比/%	杨氏模量/GPa	最小水平主应力/MPa	$10^3 \times \text{平均渗透率}/\mu\text{m}^2$	水平段长/m
井1	61.88	0.28	24.55	64.17	0.81	1 300
井1邻井	63.39	0.23	41.20	61.70	1.88	1 175
井2	61.72	0.28	25.67	62.02	2.66	1 000
井2邻井	63.97	0.23	38.90	60.71	1.93	1 304

井号	钻遇率/%	平均孔隙度/%	$\bar{S}/\%$	I类油层厚度/m	裂缝间距/m	$V(\text{总液})/\text{m}^3$
井1	100.00	8.71	54.75	58.25	20	18 316
井1邻井	99.05	9.62	55.59	142.25	35	19 840
井2	82.73	11.23	55.54	188.25	30	11 009
井2邻井	100.00	9.86	53.96	212.75	63	12 803

井号	$V(\text{总砂})/\text{m}^3$	$\varphi(\text{前置液})/\%$	$\bar{\varphi}(\text{砂})/\%$	施工排量/ $(\text{m}^3 \cdot \text{min}^{-1})$	真实 EUR/t	预测 EUR/t
井1	1 150	24.80	16.00	13	—	58 758
井1邻井	1 180	54.00	19.30	12	46 130	51 356
井2	840	11.28	13.00	13	—	26 034
井2邻井	781	59.20	17.50	12	22 139	25 394

3.2 单井压裂设计优化

由2.1节分析主控因素得知,影响玛湖地区水平井的工程参数主要为裂缝间距、总液量以及总砂量.实际压裂过程中,多种参数之间存在着相互影响,需要在多参数相互影响下,利用正交试验的方法进行多参数影响下的优化研究.正交试验就是对多因素可能产生关联影响的探究,通过正交表进行因素的排布,在各因素不同水平值的排列组合的众多方案中,选取有代表性的水平值,形成最优方案^[22],井1及井2的正交表参数如表2和表3.

表2 井1正交参数表

Table 2 Well 1 orthogonal parameter table

方案	裂缝间距/m	$10^{-4} \times \text{总液量}/\text{m}^3$	$10^{-3} \times \text{总砂量}/\text{m}^3$
1	15	1.6	1.0
2	15	1.8	1.1
3	15	2.0	1.2
4	20	1.6	1.1
5	20	1.8	1.2
6	20	2.0	1.0
7	25	1.6	1.2
8	25	1.8	1.0
9	25	2.0	1.1

表3 井2正交参数

Table 3 Well 2 orthogonal parameter

方案	裂缝间距/m	$10^{-4} V(\text{总液})/\text{m}^3$	$10^{-3} V(\text{总砂})/\text{m}^3$
1	25	1.0	0.8
2	25	1.1	0.9
3	25	1.2	1.0
4	30	1.0	0.9
5	30	1.1	1.0
6	30	1.2	0.8
7	35	1.0	1.0
8	35	1.1	0.8
9	35	1.2	0.9

根据不同的设计方案,预测不用压裂规模下的油井产能情况,并测算相应方案下的压裂投资,综合考虑产能和经济成本,选取最优的压裂方案,得到井1与井2的EUR预测结果如图5和图6.

综合考虑产能与压裂投资,井1选择方案6(裂缝间距为20 m,总液量约为 $2 \times 10^4 \text{ m}^3$,总砂量约为 $0.1 \times 10^4 \text{ m}^3$)为最优施工参数组合,此时预测EUR为64 225 t,在此压裂规模下的压裂投资为1 213.5万元,符合投资限额要求.在此方案下的EUR相较于原58 758 t提高了9.3%.井2选择方

图 5 不同方案下井 1 预测产能与投资

Fig. 5 Productivity and investment of Well 1 under different plans

图 6 不同方案下井 2 预测产能与投资

Fig. 6 Productivity and investment of Well 2 under different plans

案 9 (裂缝间距为 35 m, 总液量约为 $1.2 \times 10^4 \text{ m}^3$, 总砂量约为 900 m^3) 为最优施工参数组合, 此时预测 EUR 为 34 264 t, 在此压裂规模下的压裂投资为 954.4 万元, 符合投资限额要求. 在此方案下的 EUR 相较于原 26 034 t 提高了 37.3%.

4 结 论

1) 随机森林算法确定了影响玛湖地区水平井 EUR 的主要储层参数依次为含油饱和度、孔隙度、钻遇率以及地层压力, 主要工程参数为裂缝间距、总液量和总砂量.

2) 基于各特征参数与最终可采储量之间的 BP 神经网络模型, 利用遗传算法对各层的权值及阈值进行优化, 该产能预测模型决定系数达到 0.78, 模型性能较好.

3) 应用建立的产能模型进行压裂参数优化, 井 1 与井 2 的最终可采储量分别提高了 9.3% 和 37.3%, 可为现场水平井压裂参数的优化设计提供参考.

4) 不同区块的储层特征、流体性质和井网特征均不相同, 水平井产能参数影响程度也会不同, 本研究所建立产能预测及压裂参数优化方法可推广至其他油藏, 为现场压裂设计提供参考.

基金项目: 新疆自治区科技厅天山青年计划资助项目 (2018Q030); 中石油战略科技资助项目 (ZLZX2020-02-07-03, ZLZX2020-02-07-05); 中国石油大学 (北京) 克校区基金资助项目 (RCYJ2017B-01-003); 新疆克拉玛依市科技重大专项基金资助项目 (2018ZD001B)

作者简介: 马俊修 (1987—), 中国石油新疆油田公司工程师. 研究方向: 非常规油气藏开发技术.
E-mail: majunxiu2017@petrochina.com.cn

引 文: 马俊修, 石胜男, 陈 进, 等. 基于机器学习的玛湖地区水平井压裂设计优化 [J]. 深圳大学学报理工版, 2021, 38(6): 621-627.

参考文献 / References:

- [1] 任宜伟, 楼宣庆, 段宝江, 等. 工程参数对 L 区煤层气直井产量影响的定量研究 [J]. 石油钻采工艺, 2016, 38(4): 487-493.
REN Yiwei, LOU Xuanqing, DUAN Baojiang, et al. Quantitative analysis on the effect of engineering parameters on production rate of CBM vertical well in Block L [J]. Oil Drilling & Production Technology, 2016, 38(4): 487-493. (in Chinese)
- [2] WANG Shuhua, CHEN Shengnan. Integrated well placement and fracture design optimization for multi-well pad development in tight oil reservoirs [J]. Computational Geosciences, 2019, 23(3): 471-493.
- [3] ANIFOWOSE F, LABADIN J, ABDULRAHEEM A. A hybrid of functional networks and support vector machine models for the prediction of petroleum reservoir properties [C]// The 11th International Conference on Hybrid Intelligent Systems (HIS). Melacca, Malaysia: IEEE, 2011: 85-90.
- [4] 叶俊华, 郭娇娇, 金伟东, 等. 灰色关联法在吉 006 断块压后主控因素分析中的应用 [J]. 中国石油石化, 2016(22): 111-113.
YE Junhua, GUO Jiaojiao, JIN Weidong, et al. Application of grey correlation method in the analysis of main control factors after Ji006 fault block compression [J]. China Petroleum and Petrochemical, 2016(22): 111-113. (in Chinese)
- [5] 张 杰. 数据挖掘技术在油井压裂施工决策中的应用 [D]. 北京: 中国科学院, 2007.
Zhang Jie. Application of data mining technology in oil well fracturing construction decision-making [D]. Beijing: Chinese Academy of Sciences, 2007. (in Chinese)
- [6] 陶 亮, 郭建春, 李凌铎, 等. 致密油藏体积压裂水平井产能评价新方法 [J]. 特种油气藏, 2019, 26

- (3): 89-93, 127.
- TAO Liang, GUO Jianchun, LI Lingduo, et al. A new productivity evaluation of horizontal well with volume-fracturing in tight oil reservoir [J]. *Special Oil & Gas Reservoirs*, 2019, 26(3): 89-93, 127. (in Chinese)
- [7] JIU Bo, HUANG Wenhui, LI Yuan, et al. Influence of clay minerals and cementation on pore throat of tight sandstone gas reservoir in the eastern Ordos Basin, China [J]. *Journal of Natural Gas Science and Engineering*, 2021, 87: 103762.
- [8] ORAK N H, REEDER M, PEKNEY N J. Identifying and quantifying source contributions of air quality contaminants during unconventional shale gas extraction [J]. *Atmospheric Chemistry and Physics*, 2021, 21(6): 4729-4739.
- [9] BALDONI E, PHILIPPIDIS G, SPEKREIJSE J, et al. Getting your hands dirty: a data digging exercise to unearth the EU's bio-based chemical sector [J]. *Renewable and Sustainable Energy Reviews*, 2021, 143: 110895.
- [10] 李政宏, 刘永福, 张立强, 等. 数据挖掘方法在测井岩性识别中的应用[J]. *断块油气田*, 2019, 26(6): 713-718.
- LI Zhenghong, LIU Yongfu, ZHANG Liqiang, et al. Application of data mining method in logging lithology identification [J]. *Fault-Block Oil and Gas Field*, 2019, 26(6): 713-718. (in Chinese)
- [11] 雷 蕾, 吴乃君, 刘 鹏, 等. 灵敏度分析: 分类器中的缺失数据[C]//第11届海峡两岸信息管理发展策略研讨会论文集. 武汉: 华中科技大学出版社, 2005: 153-157.
- Lei Lei, Wu Naijun, Liu Peng, et al. Sensitivity Analysis: Missing Data in Classifiers [C]//The 11th Cross Strait Conference of Information Management Development and Strategy. Wuhan: Huazhong University of Science and Technology Press. 2005: 153-157. (in Chinese)
- [12] TRONG D D, VIET T Q, KHOA V D, et al. A two-dimensional sideways problem with random discrete data [J]. *Computers and Mathematics with Applications*, 2021, 86: 16-32.
- [13] LEBANOV L, GHIASVAND A, PAULL B. Data handling and data analysis in metabolomic studies of essential oils using GC-MS [J]. *Journal of Chromatography A*, 2021, 1640: 461896.
- [14] 樊毅龙, 马先林, 连建文. 水力压裂缺失数据填补方法研究比较[J]. *石油化工应用*, 2020, 39(11): 48-55.
- FAN Yilong, MA Xianlin, LIAN Jianwen. Research and comparison of filling methods for missing data in hydraulic fracturing [J]. *Petrochemical Industry Application*, 2020, 39(11): 48-55. (in Chinese)
- [15] 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. (in Chinese)
- [16] 纪 磊, 李菊花, 肖佳林. 随机森林算法在页岩气田多段压裂改造中的应用[J]. *大庆石油地质与开发*, 2020, 39(6): 168-174.
- JI Lei, LI Juhua, XIAO Jialin. Productivity forecast based on support vector machine optimized by grey wolf optimizer [J]. *Petroleum Geology & Oilfield Development in Daqing*, 2020, 39(6): 168-174. (in Chinese)
- [17] 林 霞, 刘宗尚, 高 宇, 等. 基于机器学习的产油量主控因素分析[J]. *信息系统工程*, 2019(12): 94-97, 99.
- LIN Xia, LIU Zongshang, GAO Yu, et al. Application of random forest algorithm in the multistage fracturing stimulation of shale gas field [J]. *China CIO News*, 2019(12): 94-97, 99. (in Chinese)
- [18] ZHAO Ruidong, SHI Junfeng, ZHANG Xishun, et al. Research and application of the big data analysis platform of oil and gas production [C]//International Petroleum Technology Conference. Beijing: Society of Petroleum Engineers, 2019: 1-9.
- [19] YU Yinshan, SHAO Mingzhen, JIANG Lingjie, et al. Quantitative analysis of multiple components based on support vector machine (SVM) [J]. *Optik*, 2021, 237: 166759.
- [20] BU Fankang, HE Jun, LI Haorun, et al. Interval-valued intuitionistic fuzzy MADM method based on TOPSIS and grey correlation analysis [J]. *Mathematical Biosciences and Engineering*, 2020, 17(5): 5584-5603.
- [21] ALEXANDRINO P S L, GOMES G F, CUNHA Jr. S S. A robust optimization for damage detection using multiobjective genetic algorithm, neural network and fuzzy decision making [J]. *Inverse Problems in Science and Engineering*, 2020, 28(1): 21-46.
- [22] 孙兆亮. 吉木萨尔致密油藏水平井压裂裂缝优化设计研究[D]. 北京: 中国石油大学(北京), 2018.
- SUN Zhaoliang. Research on optimization design of fracturing fractures in horizontal wells in Jimsar tight oil reservoir [D]. Beijing: China University of Petroleum (Beijing), 2018. (in Chinese)