

引用格式:胡秋嘉,刘春春,张建国,等.基于机器学习的煤层气井产能预测与压裂参数优化[J].油气藏评价与开发,2025,15(2):266-273.

HU Qiuqia, LIU Chunchun, ZHANG Jianguo, et al. Machine learning-based coalbed methane well production prediction and fracturing parameter optimization[J]. Petroleum Reservoir Evaluation and Development, 2025, 15(2): 266-273.

DOI: 10.13809/j.cnki.cn32-1825/te.2025.02.011

基于机器学习的煤层气井产能预测与压裂参数优化

胡秋嘉¹, 刘春春¹, 张建国¹, 崔新瑞¹, 王千², 王琪¹, 李俊¹, 何珊¹

(1. 中国石油华北油田山西煤层气勘探开发分公司, 山西长治, 046000; 2. 中国矿业大学资源与地球科学学院, 江苏徐州 221116)

摘要:沁水盆地南部煤层气区块储层非均质性强,气井产能预测难度大,且压裂施工缺乏针对性设计,导致压裂后井间生产效果差异显著。为此,基于沁水盆地南部187口煤层气直井的地质、测井、压裂和生产数据,构建了基于多任务学习策略的随机森林算法的气井产能预测模型,并通过粒子群优化算法优化压裂参数。研究使用深度卷积自动编码-解码器处理测井曲线等非结构化数据,采用随机森林算法结合多任务学习策略,有效缓解了样本数据有限和泛化性能低的问题,使得模型在小样本数据下仍能保持较高的预测精度。分析结果表明:深度、施工液量和小粒径支撑剂用量是影响产能的主要因素;地质条件是决定气井长期产能的关键因素;压裂参数则主要影响气井的峰值产能。多任务学习的随机森林算法在小样本数据上表现出高预测精度,测试集中峰值30 d和5 a累产气量的决定系数(R^2)分别为0.883和0.887。对6口新井的5 a累产气量预测 R^2 达0.901,显示出模型在实际应用中的高准确性和稳定性。通过粒子群优化算法对压裂参数进行优化后的方案,能够显著提高气井的产能分类等级或提升气井的产能水平。优化后的预测单井产能比原实际方案提高了约153%至188%,显示出优化方案在实际应用中的显著效果。通过结合多任务学习和粒子群优化算法,成功解决了小样本数据下的产能预测及压裂参数优化问题。构建的产能预测模型和压裂参数优化算法为沁水盆地南部煤层气高效开发提供了理论支持和实践参考。

关键词:煤层气;随机森林算法;多任务学习;粒子群优化算法;产能预测;压裂参数优化

中图分类号:TE357.1

文献标识码:A

Machine learning-based coalbed methane well production prediction and fracturing parameter optimization

HU Qiuqia¹, LIU Chunchun¹, ZHANG Jianguo¹, CUI Xinrui¹, WANG Qian², WANG Qi¹, LI Jun¹, HE Shan¹

(1. Shanxi Coalbed Methane Exploration and Development Company, PetroChina Huabei Oilfield, Changzhi, Shanxi 046000, China; 2. College of Resources and Geosciences, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China)

Abstract: The coalbed methane (CBM) blocks in the southern Qinshui Basin exhibit strong reservoir heterogeneity, resulting in challenges for accurate productivity prediction of gas wells. Furthermore, the absence of tailored fracturing designs has caused substantial variations in post-fracturing production performance among adjacent wells. To address these issues, a predictive model for well production capacity was developed based on geological, well logging, fracturing, and production data from 187 vertical CBM wells in the southern Qinshui Basin. The model employs a random forest algorithm integrated with a multi-task learning strategy and utilizes a particle swarm optimization (PSO) algorithm to optimize fracturing parameters. A deep convolutional autoencoder-decoder was applied to unstructured data (e.g., well logs), and the integration of random forest with multi-task learning strategies effectively addressed limited sample sizes and poor generalization, ensuring high prediction accuracy under small-data conditions. The results indicate that well depth, fracturing fluid volume, and small-sized proppant dosage are the dominant factors affecting productivity. Geological conditions determine long-term productivity, whereas fracturing parameters predominantly affect peak production performance. The multi-task random forest algorithm achieved high accuracy on small datasets, with R^2 values of 0.883 for 30-day peak cumulative production and 0.887 for 5-year cumulative production in the test set. Furthermore, the R^2 for 5-year cumulative production predictions of six new wells reached 0.901, confirming the model's robustness and reliability in field applications. The PSO-optimized fracturing parameters significantly improved the productivity

收稿日期:2024-08-29。

第一作者简介:胡秋嘉(1982—),男,硕士,教授级高级工程师,从事煤层气勘探开发研究。地址:山西省长治市漳泽工业园区华北油田山西煤层气分公司,邮政编码:046011。E-mail:mcq_hqj@petrochina.com.cn

通信作者简介:张建国(1967—),男,本科,教授级高级工程师,从事煤层气勘探开发研究。地址:山西省长治市漳泽工业园区华北油田山西煤层气分公司,邮政编码:046011。E-mail:cz_zjg@petrochina.com.cn

基金项目:国家自然科学基金项目“深部煤系造穴激活储层地质约束机理与优化”(42272198)。

classification and overall productivity levels of the gas wells. The optimized parameters increased single-well productivity by 153–188% compared to original designs, demonstrating substantial practical efficacy. The combined multi-task learning and PSO framework successfully resolves productivity prediction and fracturing optimization challenges under small-data constraints. The proposed model and fracturing parameter optimization algorithm provide theoretical support and practical references for efficient CBM development in the southern Qinshui Basin.

Keywords: coalbed methane; random forest algorithm; multi-task learning; particle swarm optimization algorithm; production prediction; fracturing parameter optimization

沁水盆地位于山西省东南部,是煤层气勘探开发的重要区域之一^[1-2],资源分布广泛,开发潜力巨大^[3]。该区域储层非均质较强,压裂设计未充分考虑到地质差异的影响,导致压裂后井间生产效果差异较大^[4]。需分析该区域煤层气井产能的主要影响因素,并在此基础上对煤层气井产能进行预测,为压裂施工方案设计进行优化提供依据^[5]。

煤层气井产能受到地质、工程等方面因素共同影响^[6-7],传统的气井产能预测准确性较差,同时压裂参数的优化涉及多种因素的耦合作用,因此,针对性的压裂设计方案难度大^[8-9]。目前,机器学习方法在煤层气井产能预测和压裂参数优化方面得到了广泛的应用^[10-11]。随机森林算法是一种常用的集成学习方法,该算法基于煤层气井的地质、压裂和生产数据,通过构建多个决策树并进行投票或平均,更容易捕捉数据中的非线性关系,处理复杂的交互关系,从而提高产能预测的准确性,为压裂参数的调整提供科学的依据^[12]。

数据样本、数据质量、特征参数选择对随机森林算法性能的影响较大^[13]。研究区沁水盆地南部沁南、马必东、郑北3个区块生产井数相对较少,导致训练模型的可靠性下降。另外,地质数据中部分参数由测井数据解释而来,但其解释方法的准确性受多种因素干扰^[14],而预测过程中直接利用测井曲线数据能有效排除其解释结果的不确定性。

因此,基于上述问题,以多层感知机和卷积神经网络耦合的方法处理提取地质、测井、压裂和生产数据的特征值参数,基于随机森林算法,嵌入多任务学习策略,结合粒子群优化算法,训练187口直井小样本数据的多任务产能预测模型,分析3个区块的煤层气直井产能主要影响因素,对气井短期和长期产能水平进行预测和分析,并针对低产井开展了压裂参数优化设计和优化方案效果评价,为沁水盆地南部地区的煤层气高效开发提供支撑。

1 研究区煤层特征与煤层气生产概况

沁南、马必东、郑北区块位于沁水盆地南部,主要含煤地层为上石炭统太原组和下二叠统山西组^[15]。太原组

和山西组共含煤21层,煤厚介于6.75~16.50 m,含煤系数为9.53%。其中,3号和15号煤层全区发育,煤层稳定,是煤层气主力产层。3号煤层位于山西组下部,煤层厚度平均为5.71 m,属三角洲平原泥炭沼泽相沉积。15号煤层位于太原组下部,距3号煤底板80~87 m,属潟湖-海湾泥炭沼泽沉积^[16]。

沁南、马必东、郑北3个区块共202口直井,根据气井产量划分为3个类型:1类井稳产日产气量 $\geq 2\ 000\ \text{m}^3$,为高产井;2类井稳产日产气量 $[1\ 000, 2\ 000)\ \text{m}^3$,为中产井;3类井稳产日产气量 $< 1\ 000\ \text{m}^3$,为低产井。3个类型井占比分别为12%、25%、63%,总体开发效果未达预期,需针对区块煤层地质特征,对压裂施工进行优化设计以提高气井产能水平。

2 数据处理与模型构建

2.1 特征值选取与数据处理

将收集的所有直井的数据按照地质、测井、压裂和生产数据进行分级整理,共整理四大类15个小类特征参数(表1)。地质、压裂和生产参数为数值;测井参数为数据序列。利用测井曲线数据一般是将测井数据用理论或经验公式解释得到孔渗、含气量和煤体结构等对气井产能有重要影响的地质参数,然而,在解释过程中受到公式及参数选取等不确定因素的干扰^[17],其解释结果往往存在偏差,因此,在产能预测模型中直接利用测井曲线原始数据。另外,压裂方面表征压裂增产效果的压裂缝半长、高

表1 特征参数分级
Table 1 Categorization of feature parameters

一级参数	二级参数	一级参数	二级参数
地质参数	深度	生产参数	峰值30 d累产气量
	煤厚		5 a累产气量
	夹矸		
压裂参数	施工液量	测井参数	密度
	总砂量		自然伽马
	携砂液施工排量		声波时差
	12/20目、20/40目、 40/70目支撑剂用量		深侧向电阻率
	射孔长度		浅侧向电阻率

度和裂缝导流能力等参数都是影响气井产能的关键工程因素^[18]。上述参数在现场获取过程中均存在较大的不确定性,直接利用反映裂缝信息未经处理的原始压裂资料,减少人工干预进而提高训练模型的准确性。

2.1.1 数据统计与规范化

建模前对井数据进行全面统计与质量分析,针对缺失数据、异常值等情况,需要通过数据预处理进一步提高数据质量,以提高算法性能和准确性^[19]。以峰值30 d累产气量和5 a累产气量作为目标变量,删除测井曲线缺失的井,共筛选187个有效数据样本,对其进行统计分析,数据分布状况如表2所示。

187个数据样本中携砂液施工排量、施工液量、砂量等特征变量存在缺失情况,其缺失率低于5%,采用中位值填补方式。另外,采取孤立森林方法识别异常值,孤立森林方法中异常被定义为容易被孤立的离群点,可以将其理解为分布稀疏且离密度高的群体较远的点^[20]。最后,将原始数据集中的每个数值线性映射到0到1的区间内,使数据的尺度得到统一,增强训练模型的稳定性和加快收敛速度^[21],转换公式为:

$$x_s = \frac{x - \bar{x}}{\sigma} \quad (1)$$

式中: x 为原始数据; \bar{x} 为原始数据的均值; σ 为原始数据的标准差; x_s 为标准化后的数据。

2.1.2 特征值选取

选取地质、测井、压裂和生产数据等15个特征变量,在产能主要影响因素分析和产能预测前需要通过特征筛选,排除部分冗余特征,以减少过拟合、提高模型泛化能力并加快模型的训练速度^[22]。鉴于测井参数为序列数据(非结构数据),与地质、压裂和生产参数(结构数据)存在较大区别,采用深度卷积自动编码-解码器实现测井序列的非线性降维表征。随后,采用皮尔逊相关系数衡量参数相关性,进行特征参数的筛选^[23]。皮尔逊相关系数计算结果显示测井曲线浅侧向电阻率与深侧向电阻率相关系数为0.97,射孔段长度与煤层深度相关系数为0.89,其特征值相关性较强。过滤掉浅侧向电阻率和射孔段长度,剩余13个参数。

2.2 产能预测模型与压裂参数优化算法

2.2.1 随机森林算法模型

应用随机森林算法,以产能数据为目标函数,计算相关特征重要性并排序,将其作为产能预测模型的数据基础,对煤层气井短期及长期产能进行预测。训练过程利用训练数据和Adam随机优化算法优化神经网络参数,通过贝叶斯优化算法和验证数据集实现神经网络模型超参数优化,获得最优机器学习模型参数^[24]。整个训练过程如图1所示。

表2 特征参数数据分布情况

Table 2 Distribution of feature parameter data

特征	携砂液施工排量/(m ³ /min)	不同颗粒支撑剂用量/m ³			施工液量/m ³	总砂量/m ³
		12/20目	20/40目	40/70目		
数量	187	187	187	187	187	187
平均值	6.3	5.6	32.4	12.2	879.6	54.7
方差	1.5	9.2	15.0	18.9	368.3	25.2
最小值	0.6	0	0	0	287.3	8.1
25%	5.6	0	20.0	0	710.1	40.1
50%	6.2	0	37.5	0	811.8	50.1
75%	7.4	10.0	40.0	20.0	933.8	60.1
最大值	11.0	70.0	100.0	60.1	2 467.6	115.2

特征	煤厚/m	夹矸/m	深度/m	峰值30 d累产气量/m ³	5 a累产气量/m ³
数量	187	187	187	187	187
平均值	5.9	0.5	988.4	26 936.4	726 105.7
方差	1.2	0.4	412.2	23 909.1	775 065.3
最小值	0.7	0	657.1	0	0
25%	5.5	0.4	791.6	8 166.5	146 131.5
50%	6.0	0.6	965.7	21 218.0	498 499.0
75%	6.6	0.7	1 120.7	37 191.5	1 055 182.0
最大值	11.3	2.5	6 094.8	110 622.0	4 002 087.0

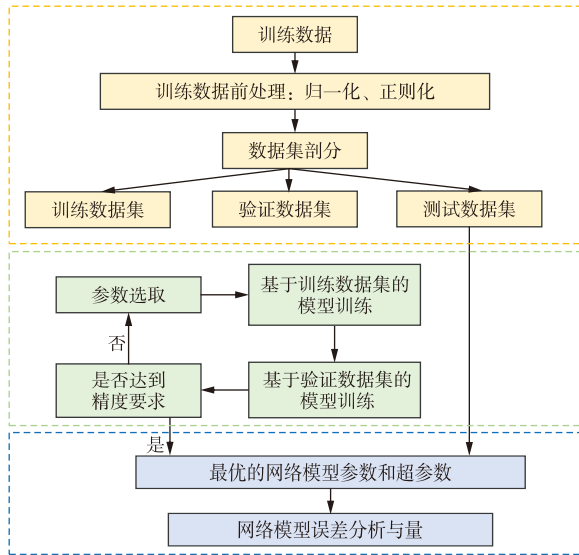


图1 机器学习模型训练及超参数优化示意图

Fig. 1 Schematic diagram of machine learning model training and hyperparameter optimization

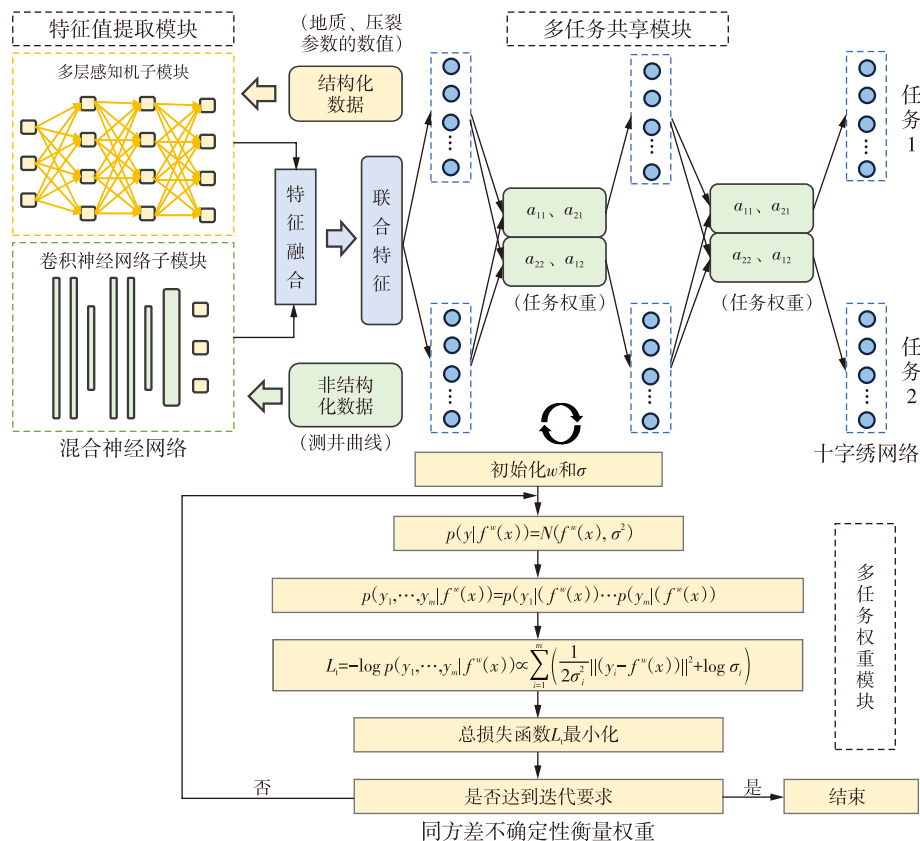
2.2.2 多任务学习策略的嵌入

此次收集的有效样本数量仅为187个,而煤层气井产能受储层条件和压裂施工等多种因素的影响,一般基于随机森林算法的单任务学习遇到有限样本时,很难理解数据集的真实分布,从而导致过拟合和弱泛化能力^[25]。

设置峰值30 d累产气量以及5 a累产气量2个预测任务目标,建立的单任务随机森林产能预测模型计算效率较低。因此,将多任务学习理念嵌入有限样本产能预测模型中,能充分利用多个相关任务训练信号中的共享表征,缓解样本数据有限、泛化性能低的问题,有效地提高机器学习的预测性能和运算效率^[26]。采用十字绣网络样式构建共享模块,使用多任务的共享表征和特定任务表征的线性组合来连接多个任务的神经网络,通过端到端的学习自动确定分割结构和共享程度^[27]。

从有限数据中充分提取有价值的特征,增强信息表达,对于有限样本多任务预测模型也同样重要。为此,在产能预测框架中嵌入了混合神经网络,借助多层感知机和卷积神经网络2个子模块,从多源、多类型和多维的输入中提取特征。多层感知机模块由多层全连接层构成,旨在处理地质、压裂和生产等数据的离散型特征。卷积神经网络子模块由卷积层、池化层、扁平层和全连接层等组成,旨在从作为非结构化数据的测井曲线提取有用表征。

多任务学习通过同时优化多个任务的损失来准确有效地进行预测,整个模型的性能高度依赖于每个损失的相对权重。基于任务的最优权重取决于任务噪声大小的认知^[28]。因此,引入同方差不确定性来衡量任务的权重,确定多任务损失权重,最大化高斯似然估计,总损失函数具体计算过程见图2中多任务权重模块。



注: σ 为标准差; w 为任务权重; $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$ 为任务权重; x 为输入数据; $f^*(x)$ 为任务权重 w 、输入数据 x 的神经网络的输出; y_1, y_2, \dots, y_m 为观测值; p 为似然函数; N 为正态分布; m 为任务数量; σ_i 为第 i 个数据的标准差; L_i 总损失函数。

图2 多任务产能预测模型示意图

Fig. 2 Schematic Diagram of the multi-task production prediction model

基于多任务学习的理念,针对有限样本产能预测,建立多任务产能预测模型,结构如图2所示。模型由特征提取、任务共享和损失函数3个要素构成^[29]。首先,多任务的多类型数据被输入到混合神经网络模块,以提取结构数据和非结构数据的有效特征;随后,多个任务的联合特征被传递到十字绣网络模块,通过端到端的学习自动确定多任务的共享结构和分享程度,进而提高多任务预测的鲁棒性和泛化性能;最后,使用同方差不确定性衡量损失权重,建立多任务损失函数,通过最小化损失训练模型。

在已建立的一般产量预测随机森林方法框架的基础上,搭建有限样本产量预测框架,有限样本多任务产能预测模型建模建立步骤如下:

收集地质参数、压裂施工参数、生产数据和煤层段的测井曲线等多类型数据,将不同尺度的数据缩放到0~1范围内,对地质、测井和压裂参数的特征值进行筛选;选取峰值30 d累产气量以及5 a累产气量作为预测任务;将统计的数据集随机划分为训练集和测试集,其中测试集占总数据集的15%;基于预处理后的多任务输入和输出,采用同方差不确定性建立的多任务损失函数训练多任务产能预测模型,利用超参数优化算法在5折交叉验证数据集上寻找最优模型超参数;最后采用平均绝对误差、均方根误差和决定系数(R^2)3个评价指标来评估模型的准确性。

2.2.3 粒子群优化算法压裂参数

采用粒子群优化算法基于地质、测井参数,以单井产能为目标函数,对压裂参数进行优化。粒子群优化算法是一种基于种群的随机方法^[30],用于求解连续和离散优化问题,在粒子群优化算法中,简单的软件代理粒子在优化问题的搜索空间中移动。粒子的位置表示当前优化问题的候选解决方案,每个粒子通过改变其速度来搜索空间中更好的位置。粒子群优化算法属于一类用于求解优化问题的群体智能技术,可用于寻求多项压裂参数中各个参数的群体最优值,而非其自身最优值。

3 结果与讨论

3.1 气井产能影响因素评价

分别将峰值30 d累产气量和5 a累产气量作为评估煤层气井早期最大产气能力以及长期产气的的能力,基于随机森林算法的影响因素重要性评价方法对地质、测井和压裂特征参数的重要性进行评价,评价结果如图3和图4所示。其中地质和测井参数均是表征煤层气井产量的地质条件因素,因此,两种参数作为一个整体进行评价。

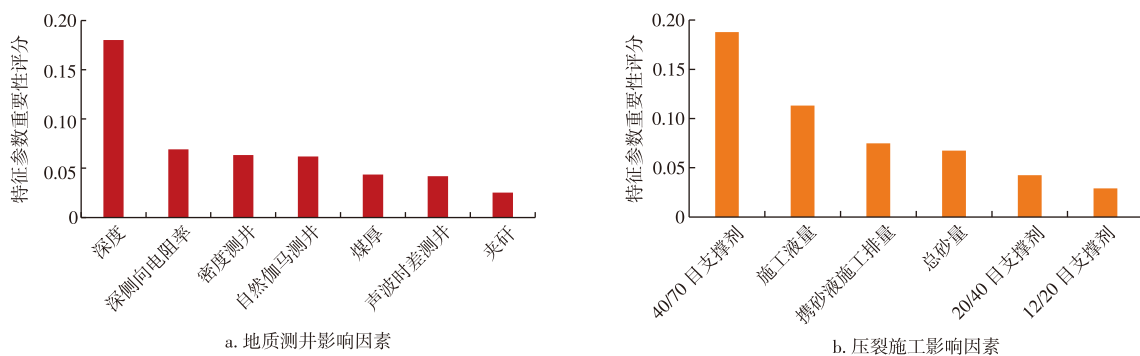


图3 以峰值30 d累产气量为评价指标的参数重要性评价

Fig. 3 Parameter importance evaluation based on peak 30-day cumulative gas production

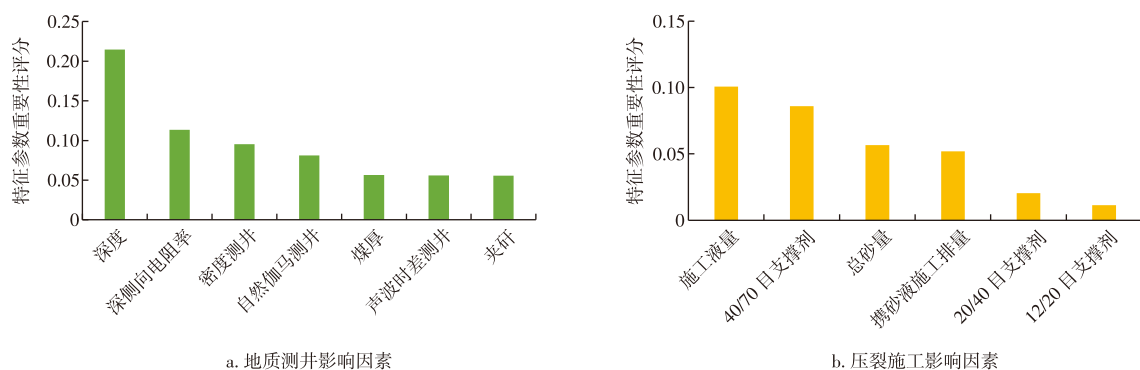


图4 以5 a累产气量为评价指标的参数重要性评价

Fig. 4 Parameter importance evaluation based on 5-year cumulative gas production

以峰值30 d累产气量作为产能评价指标,地质测井参数重要性总评分为0.48,压裂参数重要性评分为0.52,压裂参数的影响程度略高于地质测井参数。其中深度是对峰值30 d累产气量影响程度最高的地质参数,其原因在于深度与含气量以及应力机制关系较大。压裂参数中小粒径支撑剂的用量和施工液量的影响比重较大,而较大颗粒支撑剂的用量则影响最小。其原因可能在于大粒径的颗粒在压裂过程中容易沉降,其用量对裂缝导流能力支撑效果影响较小,而施工液量和小颗粒支撑剂能有效增加压裂裂缝体积的扩展和有效导流能力的维持。

以5 a累产气量作为产能评价指标,地质测井参数重要性总评分为0.67,压裂参数重要性评分为0.33,压裂参数的影响程度显著低于地质测井参数,且相较于峰值30 d累产气量,地质测井参数的重要性显著上升。该结果表明煤层气井短期的产能水平与压裂施工相关性更强,而地质参数是决定气井长期产能的主要因素。然而无论是对于峰值30 d累产气量和5 a累产气量,地质测井和压裂参数重要性排序基本一致,即对煤层气井产能影响最大的因素是深度、压裂液用量以及小颗粒支撑剂的用量。

3.2 煤层气井产能预测

以气井峰值30 d累产气量和5 a累产气量为预测目

标,基于嵌入多任务学习策略的随机森林算法训练产能预测模型结果如图5所示。

表3为训练集和测试集的3种回归评价指标,平均绝对误差和均方根误差数值均较低,说明其回归评价结果较好。气井5 a累产气量预测结果训练集和测试集的 R^2 分别为0.916和0.887,即预测精度分别为91.6%和88.7%,超过了一般方法80%左右的预测精度,说明基于多任务学习的产能预测结果可靠程度高。另外,由于数据样本来自3个不同的区块,为了评估该模型适应性差异,分别统计3个区块预测结果误差,如表4所示。可以看出该模型预测结果的误差相差较小,表明该模型在3个区块的适应性良好。其主要原因在于地质和测井特征参数的选取较全面地包含了影响气井产能的地质方面因素,而区块之间的地质差异也在模型中得到了体现。

在3个区块各收集2口(共6口)未进入过训练集及测试集的新井,利用模型进行产能预测。表5为新井产能预测值和真实值对比,6口新井的预测产能值与实际产能值均相差较小,新井5 a累产气量预测结果的 R^2 为0.901,预测精度为90.1%,进一步验证了基于嵌入多任务学习策略的随机森林算法训练产能预测模型的准确性和稳定性。

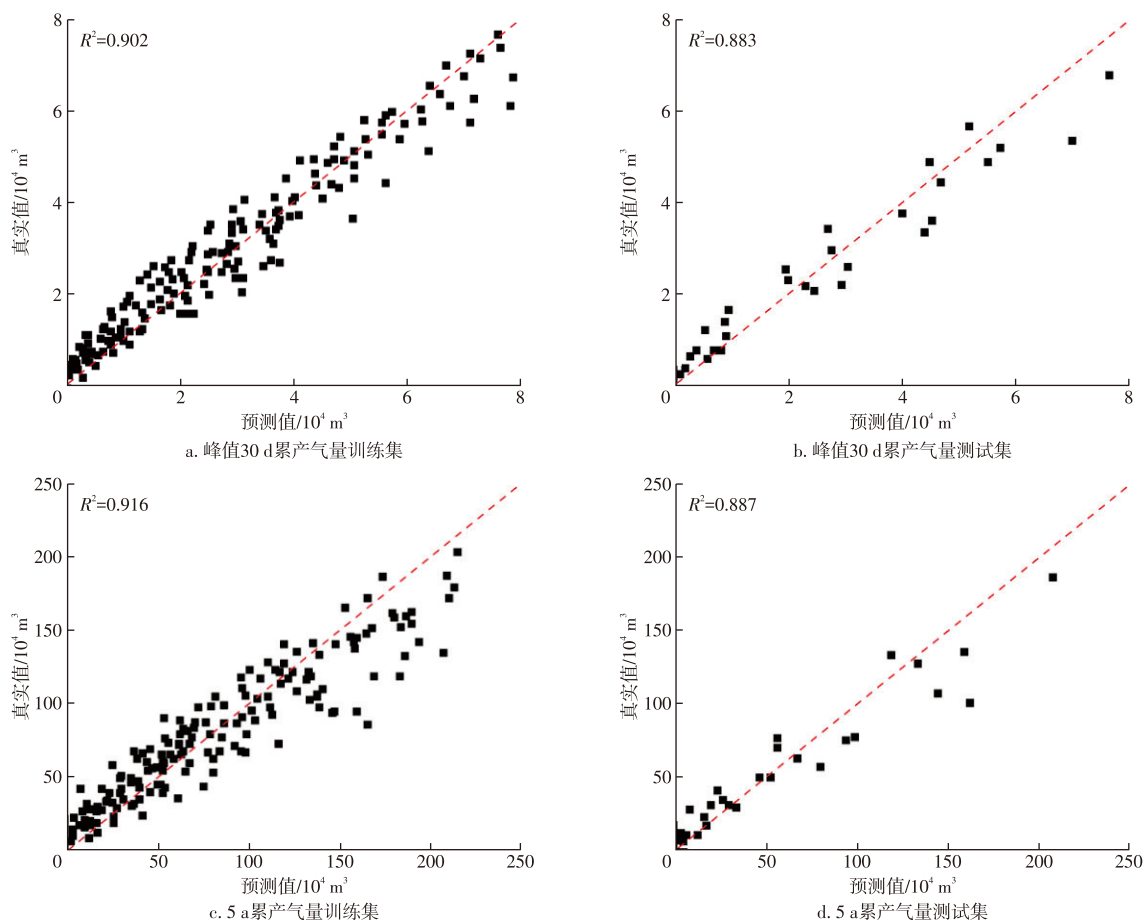


图5 模型预测结果与真实结果对比

Fig. 5 Comparison between model predictions and actual results

表3 回归评价指标计算结果

Table 3 Results of Regression Evaluation Metrics

预测指标	项目	平均绝对误差	均方根误差	决定系数 R^2
峰值30 d累产气量	训练集	0.199	0.281	0.902
	测试集	0.382	0.397	0.883
5 a累产气量	训练集	13.530	15.470	0.916
	测试集	9.851	14.080	0.887

表4 3个区块预测结果误差分析

Table 4 Error analysis of prediction results in three blocks

误差	预测指标	项目	沁南	马必东	郑北
决定系数	峰值30 d累产气量	训练集	0.879	0.913	0.909
		测试集	0.891	0.887	0.875
	5 a累产气量	训练集	0.921	0.912	0.917
		测试集	0.879	0.889	0.891

表5 新井产能真实值和预测值对比

Table 5 Comparison between actual and predicted production of new wells

单井产能	编号	真实值	预测值	决定系数
峰值30 d累产气量/ 10^4 m^3	M-1	2.29	1.96	0.896
	Z-1	1.45	1.56	
	X-1	1.44	1.27	
	M-2	1.64	1.73	
	Z-2	2.57	2.36	
	X-2	1.35	1.50	
5 a累产气量/ 10^4 m^3	M-1	54.44	47.06	0.901
	Z-1	24.45	26.57	
	X-1	49.85	45.87	
	M-2	36.17	32.53	
	Z-2	54.37	52.90	
	X-2	45.37	47.95	

3.3 压裂效果评价与参数优化

从3个区块另选取3口新井Z-3井、M-3井和X-3井,基于产能预测模型,通过粒子群优化算法,以提高其产能为目标,开展压裂施工方案优化,并对其优化方案的产量进行预测,如表6所示。Z-3井、M-3井为已压裂未投产井,对其实际压裂施工方案以及优化后的压裂设计方案进行产能预测;X-3井为一口未压裂未投产的新井,对其进行压裂参数优化设计及设计方案预测。

基于地质测井参数,按照实际的压裂施工方案产能,预测Z-3井、M-3井峰值30 d累产气量分别为 $1.37 \times 10^4 \text{ m}^3$ 、 $1.01 \times 10^4 \text{ m}^3$,对应日产气量为 456.6 m^3 、 327.8 m^3 ,均属于低产井。优化后压裂方案为峰值30 d累产气量分别为 $3.96 \times 10^4 \text{ m}^3$ 、 $2.56 \times 10^4 \text{ m}^3$,对应日产气量为 $1\,320.1 \text{ m}^3$ 、 852.4 m^3 ,Z-3井提升为中产井,虽然M-3井依然为低产

表6 Z-3井、M-3井和X-3井压裂参数优化及对应产能预测结果

Table 6 Fracturing parameter optimization and production prediction results for wells Z-3, M-3, and X-3

井名	Z-3		M-3		X-3
压裂施工方案	实际方案	优化方案	实际方案	优化方案	优化方案
携砂液施工排量/ (m^3/min)	5.3	5.9	8.2	7.6	6.1
12/20目支撑剂/ m^3	0	1.8	10.0	16.6	6.9
20/40目支撑剂/ m^3	10.0	21.2	20.0	30.3	25.7
40/70目支撑剂/ m^3	10.0	23.1	10.0	23.2	24.7
总砂量/ m^3	20.0	46.1	40.0	70.1	57.3
施工液量/ m^3	307.8	1\,204.9	812.2	1\,715.4	1\,195.6
峰值30 d累产气量/ 10^4 m^3	1.37	3.96	1.01	2.56	5.22
5 a累产气量/ 10^4 m^3	24.40	59.41	23.10	54.20	42.60

井,但其峰值30 d累产气量和5 a累产气量是实际压裂施工方案的2.5倍和2.3倍,产能改善效果显著。

对于未压裂未投产的新井X-3,压裂设计方案对应的峰值30 d累产气和5 a累产气分别为 $5.22 \times 10^4 \text{ m}^3$ 与 $42.60 \times 10^4 \text{ m}^3$,属于中产井,根据该区块煤层气井的产能状况,其产能水平将位于前20%,表明优化的压裂施工方案能有效发挥该井的产能潜力。

4 结论

基于沁南、郑北、马必东3个区块共187口井地质、测井、压裂以及生产数据,通过嵌入多任务学习策略的随机森林算法对煤层气井产能主要影响因素进行评估以及产能预测,利用粒子群优化算法对压裂施工参数进行优化,得到以下结论:

1) 压裂参数对煤层气井短期产能有较大影响,气井长期产能更受地质条件制约,其中关键影响因素包括深度、施工液量和小粒径支撑剂用量。

2) 多任务学习的随机森林算法在小样本数据上展现了较高的气井产能预测精度,新井的产能预测值与实际值相差较小,其 R^2 均大于0.875,且3个区块预测结果 R^2 差异较小,显示出模型在研究区的高预测精度和适应性。

3) 通过粒子群优化算法对压裂参数进行优化后的方案能提高气井产能分类等级或显著提高气井产能,优化后的预测单井产能比原实际方案提高约153%~188%。

参考文献

- [1] 张聪,李可心,贾慧敏,等.郑庄北中深部煤层气水平井产能影响因素及开发技术优化[J].煤田地质与勘探,2024,52(6):21-32.
ZHANG Cong, LI Kexin, JIA Huimin, et al. Factors influencing the productivity and technology optimization of horizontal wells for

- moderately deep coalbed methane in the northern Zhengzhuang block [J]. *Coal Geology & Exploration*, 2024, 52(6): 21–32.
- [2] 王琪,张聪,贾慧敏,等.煤层气筛管水平井低产原因分析和治理对策研究:以沁水盆地南部煤层气筛管水平井为例[J]. *中国煤层气*, 2024, 21(1): 21–24.
WANG Qi, ZHANG Cong, JIA Huimin, et al. Study on causes and treatment methods of low production of coalbed methane screen horizontal well: A case study in southern Qinshui Basin[J]. *China Coalbed Methane*, 2024, 21(1): 21–24.
- [3] 唐书恒,李洋,吕建伟.原位储层生物地球化学评价及其对煤层气开采的指示意义:以沁水盆地南部柿庄南区块为例[J]. *煤炭学报*, 2024, 49(1): 555–562.
TANG Shuheng, LI YANG, LYU Jianwei. In situ reservoir biogeochemical evaluation and its indicative significance for coalbed methane extraction: Taking the Shizhuangnan Block in the southern Qinshui Basin as an example[J]. *Journal of China Coal Society*, 2024, 49(1): 555–562.
- [4] 张聪,李梦溪,胡秋嘉,等.沁水盆地南部中深部煤层气储层特征及开发技术对策[J]. *煤田地质与勘探*, 2024, 52(2): 122–133.
ZHANG Cong, LI Mengxi, HU Qiuqia, et al. Moderately deep coalbed methane reservoirs in the southern Qinshui Basin: Characteristics and technical strategies for exploitation[J]. *Coal Geology & Exploration*, 2024, 52(2): 122–133.
- [5] 朱庆忠.沁水盆地高煤阶煤层气高效开发关键技术与实践[J]. *天然气工业*, 2022, 42(6): 87–96.
ZHU Qingzhong. Key technologies and practices for efficient development of high-rank CBM in the Qinshui Basin[J]. *Natural Gas Industry*, 2022, 42(6): 87–96.
- [6] LYU Y M, TANG D Z, XU H, et al. Production characteristics and the key factors in high-rank coalbed methane fields: A case study on the Fanzhuang Block, Southern Qinshui Basin, China[J]. *International Journal of Coal Geology*, 2012, 96: 93–108.
- [7] LI H Y, LAU H C, HUANG S. China's coalbed methane development: A review of the challenges and opportunities in subsurface and surface engineering[J]. *Journal of Petroleum Science and Engineering*, 2018, 166: 621–635.
- [8] DU S Y, WANG M Z, YANG J S, et al. An enhanced prediction framework for coalbed methane production incorporating deep learning and transfer learning[J]. *Energy*, 2023, 282: 128877.
- [9] 贾慧敏,胡秋嘉,张聪,等.煤层气双层合采直井产能预测及排采试验:以沁水盆地郑庄西南部为例[J]. *油气藏评价与开发*, 2022, 12(4): 657–665.
JIA Huimin, HU Qiuqia, ZHANG Cong, et al. Prediction of productivity and co-drainage trial of bilayer vertical coalbed methane wells: Cases study of the southwest of Zhengzhuang Block, Qinshui Basin[J]. *Petroleum Reservoir Evaluation and Development*, 2022, 12(4): 657–665.
- [10] 宋洪庆,都书一,杨焦生,等.基于机器学习的煤层气产能标定智能算法及影响因素分析[J]. *工程科学学报*, 2024, 46(4): 614–626.
SONG Hongqing, DU Shuyi, YANG Jiaosheng, et al. Forecasting and influencing factor analysis of coalbed methane productivity utilizing intelligent algorithms[J]. *Chinese Journal of Engineering*, 2024, 46(4): 614–626.
- [11] 王贝,门鹏,李腾,等.基于机器学习可解释性算法的石嘴山矿区煤层气井排采主控因素分析[J]. *中国煤层气*, 2023, 20(5): 3–7.
WANG Bei, MEN Peng, LI Teng, et al. Analysis of main control factors of coalbed methane well drainage in Shizuishan Mining Area based on machine learning interpretability algorithm[J]. *China Coalbed Methane*, 2023, 20(5): 3–7.
- [12] DU S Y, WANG J L, WANG M Z, et al. A systematic data-driven approach for production forecasting of coalbed methane incorporating deep learning and ensemble learning adapted to complex production patterns[J]. *Energy*, 2023, 263: 126121.
- [13] FU X H, QIN Y, WANG G G X, et al. Evaluation of gas content of coalbed methane reservoirs with the aid of geophysical logging technology[J]. *Fuel*, 2009, 88(11): 2269–2277.
- [14] GAN T, BALMAIN B, SIGBATULLIN A. Formation evaluation logoff results comparing new generation mining-style logging tools to conventional oil and gas logging tools for application in coalbed methane (CBM) field development[J]. *Journal of Natural Gas Science and Engineering*, 2016, 34: 1237–1250.
- [15] 胡秋嘉,张聪,贾慧敏,等.沁水盆地南部郑庄区块中北部煤层气直井增产新技术研究与应用[J]. *煤炭学报*, 2024, 49(3): 1518–1529.
HU Qiuqia, ZHANG Cong, JIA Huimin, et al. Research and application of a new stimulation technology for deep coalbed methane vertical wells in central and Northern Zhengzhuang block, southern Qinshui Basin[J]. *Journal of China Coal Society*, 2024, 49(3): 1518–1529.
- [16] 张聪,胡秋嘉,冯树仁,等.沁水盆地南部煤层气地质工程一体化关键技术[J]. *煤矿安全*, 2024, 55(2): 19–26.
ZHANG Cong, HU Qiuqia, FENG Shuren, et al. Key technologies for integration of coalbed methane geology and engineering in southern Qinshui Basin[J]. *Safety in Coal Mines*, 2024, 55(2): 19–26.
- [17] LI L J, LIU D M, CAI Y D, et al. Coal structure and its implications for coalbed methane exploitation: A review[J]. *Energy & Fuels*, 2021, 35(1): 86–110.
- [18] 胡秋嘉,李梦溪,乔茂坡,等.沁水盆地南部高阶煤煤层气井压裂效果关键地质因素分析[J]. *煤炭学报*, 2017, 42(6): 1506–1516.
HU Qiuqia, LI Mengxi, QIAO Maopo, et al. Analysis of key geologic factors of fracturing effect of CBM wells for high-rank coal in Southern Qinshui Basin[J]. *Journal of China Coal Society*, 2017, 42(6): 1506–1516.
- [19] 李敬松,王涛,王金伟,等.基于多层感知机模型的煤层气井压后生产动态反演研究[J]. *测井技术*, 2023, 47(5): 558–568.
LI Jingsong, WANG Tao, WANG Jinwei, et al. Production dynamic of coal-bed methane after well pressure based on multi-layer perceptron model inversion study[J]. *Well Logging Technology*, 2023, 47(5): 558–568.
- [20] MENG M, ZHONG R Z, WEI Z L. Prediction of methane adsorption in shale: Classical models and machine learning based models[J]. *Fuel*, 2020, 278: 118358.
- [21] MIN C, WEN G Q, GOU L J, et al. Interpretability and causal discovery of the machine learning models to predict the production of CBM wells after hydraulic fracturing[J]. *Energy*, 2023, 285: 129211.