

引文: 郭子熙, 张舒, 马翥, 等. 基于融合多模态特征的深层煤岩气产量预测[J]. 天然气工业, 2024, 44(10): 140-149.

GUO Zixi, ZHANG Shu, MA Biao, et al. Production prediction of deep coal-rock gas based on integrated multi-modal characteristics[J]. Natural Gas Industry, 2024, 44(10): 140-149.

基于融合多模态特征的深层煤岩气产量预测

郭子熙^{1,2} 张舒^{1,3,4,5} 马翥⁶ 康芸玮⁷ 李曙光⁸ 陈东⁹ 陈怡羽⁸ 陈博文¹⁰ 赵金洲¹

1. 油气藏地质及开发工程全国重点实验室·西南石油大学 2. 清华大学数学科学系 3. 西南石油大学计算机与软件学院

4. 四川省油气勘探开发智能化工程研究中心 5. 数据挖掘与知识管理南充市重点实验室·西南石油大学

6. 中国石油集团川庆钻探工程有限公司井下作业公司 7. 西南石油大学理学院 8. 中国石油煤层气有限责任公司

9. 中联煤层气国家工程研究中心 10. 中国石油西南油气田公司

摘要: 深层煤岩气生产在井型、压裂工艺、排采方式等方面显著区别于浅层煤层气, 针对浅层煤层气建立的产量预测模型无法处理多元异构的深层煤岩气数据, 且难以适应深层煤岩气的生产方式。为此, 基于卷积自编码结合空间金字塔池化提取深层煤岩气二维地质特征, 采用 PointNet++ 提取深层煤岩气缝网三维特征, 拼接得到的完整数据矩阵再经近邻传播法 (Affinity Propagation, AP) 聚类后划分训练集和测试集, 最后将注意力机制融入长短时记忆网络 (Long Short-Term Memory, LSTM) 建模, 建立了预测深层煤岩气产量的深度学习新模型并通过不同算法和消融实验对比分析了新模型的泛化能力。研究结果表明: ①提出的特征提取方法能够有效解决深层煤岩气数据多元异构、多尺度等问题。②对比深度神经网络 (Deep Neural Networks, DNN)、门控循环单元 (Gated Recurrent Unit, GRU)、LSTM、数值模拟、BP 神经网络, 融入三维特征提取及注意力机制的 AP—LSTM 预测产量的准确率最高。③融入三维特征提取, AP—LSTM 预测产量的数据波动得到有效抑制; 融入注意力机制, AP—LSTM 预测产量的趋势得到改善。④基于研究区块建立的模型迁移到新的区块预测产量, 仍具备较高的准确率。结论认为, 缝网特征是影响深层煤岩气产量的重要因素, 注意力机制有助于 AP—LSTM 区分自喷和抽吸, 融入三维特征提取及注意力机制的 AP—LSTM 适应了深层煤岩气的生产方式, 具有良好的普适性, 能够为深层煤岩气规模化开发提供技术支撑。

关键词: 深层煤岩气; 深度学习; 长短时记忆网络; 三维特征提取; 注意力机制; 产量预测

中图分类号: TE32²·8

文献标识码: A

DOI: 10.3787/j.issn.1000-0976.2024.10.011

Production prediction of deep coal-rock gas based on integrated multi-modal characteristics

GUO Zixi^{1,2}, ZHANG Shu^{1,3,4,5}, MA Biao⁶, KANG Yunwei⁷, LI Shuguang⁸, CHEN Dong⁹, CHEN Yiyu⁸, CHEN Bowen¹⁰, ZHAO Jinzhou¹

(1. National Key Laboratory of Oil & Gas Reservoir Geology and Exploitation//Southwest Petroleum University, Chengdu, Sichuan 610500, China; 2. Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China; 3. School of Computer Science and Software Engineering, Southwest Petroleum University, Chengdu, Sichuan 610500, China; 4. Engineering Research Center for Intelligent Oil & Gas Exploration and Development of Sichuan Province, Chengdu, Sichuan 610500, China; 5. Nanchong Key Laboratory of Data Mining and Knowledge Management//Southwest Petroleum University, Nanchong, Sichuan 610599, China; 6. Downhole Service Company, CNPC Chuangqing Drilling Engineering Company Limited <CCDC>, Chengdu, Sichuan 610052, China; 7. School of Sciences, Southwest Petroleum University, Chengdu, Sichuan 610500, China; 8. PetroChina Coalbed Methane Company Limited, Beijing 100028, China; 9. China United Coalbed Methane National Engineering Research Center, Beijing 100095, China; 10. PetroChina Southwest Oil & Gasfield Company, Chengdu, Sichuan 610051, China)

Natural Gas Industry, Vol.44, No.10, p.140-149, 10/25/2024. (ISSN 1000-0976; In Chinese)

Abstract: The production of deep coal-rock gas is significantly different from that of shallow coalbed methane in terms of well types, fracturing techniques and production modes. The production prediction model established for shallow coalbed methane cannot be used to process the deep coal-rock gas data of multivariate heterogeneity, and can hardly adapt to the production mode of deep coal-rock gas. To this end, this paper extracts the 2D geological features of deep coal-rock gas based on convolutional autoencoder and spatial pyramid pooling, and the 3D fracture network features of deep coal-rock gas by means of PointNet++ method. The complete data matrix obtained by means of mosaic is clustered through affinity propagation (AP) and then divided into training and testing sets. Finally, the attention mechanism is integrated into the long short-term memory network (LSTM). Thus, a new deep learning model for production prediction of deep coal-rock gas is established. In addition, its generalization ability is comparatively analyzed by virtue of different algorithms and ablation experiments. The following results are obtained. First, the proposed feature extraction methods can effectively handle the multivariate heterogeneity and multiple scales of deep coal-rock gas data. Second, compared with deep neural networks (DNN), the gated recurrent unit (GRU), LSTM, numerical simulation and the BP neural network, the AP—LSTM integrated with 3D feature extraction and attention mechanism has the highest accuracy in production prediction. Third, by incorporating 3D feature extraction, AP—LSTM effectively suppresses data fluctuations in production prediction; and by integrating attention mechanism, AP—LSTM improves the trend of production prediction. Fourth, the model established on the basis of the study area still has a high production prediction accuracy when it is applied to other blocks. In conclusion, the characteristics of the fracture network are important factors affecting the production of deep coal-rock gas. The attention mechanism is helpful for AP—LSTM to distinguish between flowing and pumping. The AP—LSTM which integrates 3D feature extraction and attention mechanism adapts to the production mode of deep coal-rock gas and has a good universality, providing technical support for the large-scale development of deep coal-rock gas.

Keywords: Deep coal-rock gas; Deep learning; Long short-term memory network; 3D feature extraction; Attention mechanism; Production prediction

基金项目: 国家自然科学基金重点项目“基于新一代信息技术的复杂油气储层地震勘探理论和方法”(编号: 42330801), 油气藏地质及开发工程全国重点实验室开放基金课题“深部煤层气缝网压裂缝网形成与延伸机理研究”(编号: PLN2024-29)、“基于迁移学习的气井产量预测方法研究”(编号: PLN2022-50)。

作者简介: 郭子熙, 1994年生, 助理研究员, 博士, 本刊青年编委; 主要从事油气藏增产改造技术及油气人工智能技术研究工作。地址: (100084) 北京市海淀区双清路30号。ORCID: 0000-0002-6624-7060。E-mail: guozixi@vip.163.com

通信作者: 张舒, 女, 1990年生, 副教授, 硕士研究生导师, 博士; 主要从事人工智能在油气勘探开发应用方面的研究工作。地址: (610500) 四川省成都市新都区新都大道8号。ORCID: 0000-0001-8589-1567。E-mail: zhangshu@swpu.edu.cn

0 引言

煤层气是清洁低碳能源，属于非常规天然气^[1]。煤层气开发利用不仅能提供清洁低碳能源，而且对促进煤矿安全生产和保护生态环境有积极的推动作用，对全球和国家“双碳”目标的实现具有重要意义^[2-3]。

煤层气产量预测能对煤层气井开发时的区块选择、经济效益评价等工作起到指导作用^[4-6]。煤层气储集主要是通过吸附作用且具有“吸附—解吸—扩散—渗流”的特点，使得煤层气的开采过程与煤层气产量之间构成复杂的非线性关系，因此像常规油气产量预测一样建立解析数学模型是极其困难的^[7-8]。针对煤层气产量预测，目前国内外研究集中在数值模拟和人工智能方法^[9]。影响煤层气产量的因素多而复杂，不仅受到储层地质因素影响，而且受到压裂工艺、排采制度等工程因素影响，在数值模拟中需要输入的渗透率、含气饱和度等参数既难获取，也是实时动态变化的（因煤层的压敏特性、煤岩易碎形成煤粉堵塞渗流通道等），因此基于数值模拟的煤层气产量预测模型计算结果精度往往很低^[10]。随着信息社会和现代科技的迅速发展，人工智能在各领域得到应用，也为煤层气产量预测提供了新的思路^[11]。如人工神经网络（ANN）、径向基函数（RBF）神经网络、反向传播（BP）神经网络等浅层神经网络已广泛用于煤层气产量预测，但预测结果容易出现过拟合现象^[12]。因此，浅层神经网络对煤层气产量的预测精度也较低。

2006 年，Hinton 等^[13-14]首次提出了深度学习的概念，该方法可以从大量复杂的数据中学习到处有效的特征。近年来石油工程领域已有许多学者将其应用到非常规油气产量预测中，如 Lee 等^[15]为解决传统方法应用于页岩气井出现的长瞬态流动、独特完井设计和高密度钻井等限制，充分考虑页岩气产量的时序关系，基于 LSTM 算法建立页岩气产量预测模型；Rahmanifard 等^[16]对基于 LSTM 和 GRU 的致密油产量模型的计算量和准确率进行了全面对比；笔者团队^[17]也基于深度学习理论建立了浅层煤层气产量预测模型并成功在中国韩城、潘河和保德等区块应用。

深层煤岩气是储存于埋深 1 500 m 以深煤层中的烃类气体，是重要的天然气勘探潜在领域^[18]。中国深层煤岩气资源丰富，仅鄂尔多斯盆地深层煤岩气资源量就初步评价为 $23 \times 10^{12} \text{ m}^3$ ，成为近年来非常规油气开发及国内外研究的新热点^[19]。笔者团队在将过去基于深度学习理论建立的煤层气产量预测模

型应用到深层煤岩气产量预测上后，发现预测精度显著变低。事实上深层煤岩气生产在井型、压裂工艺、排采方式等方面显著区别于浅层煤层气生产^[20]：浅层煤层气井采用直井压裂，而深层煤岩气井目前采用大规模水平井缝网压裂；浅层煤层气井以吸附气为主，在排水解吸后见气，受限于成本利用天然能量衰竭式开采，而深层煤岩气因为游离气比例随埋深增加而增高，开采即见气，高产带来的经济效益使其在衰竭式开采后能够及时转抽等。因此过去的预测模型存在以下问题：①由于水平井分段压裂涉及压裂时间、压裂段数、段点位置、每段加砂量、每段加液量等多个因素，导致出现人为刻画不全面，压裂段数不一致等难点问题；②深层煤岩气开采有自喷和抽吸两个阶段，过去的模型无法有效识别，导致预测产量的趋势与真实产量存在较大差异。为解决以上问题，不仅需要进一步明确深层煤岩气产量的重要影响因素，还需要对深度学习模型加以改进，以适应深层煤岩气不同的生产方式。

本文拟通过在数据特征提取模块加入 PointNet++ 提取缝网压裂后的三维缝网数据特征，准确刻画深层煤岩气水平井大规模缝网压裂后的缝网特征，在深度学习预测模型中融入注意力机制帮助区分自喷和抽吸阶段，建立基于深度学习的深层煤岩气产量预测方法，为深层煤岩气规模化开发提供技术支撑。

1 特征提取方法

深层煤岩气数据来源多，包括地质数据、压裂数据和排采数据，并且来源的格式可能是文本型、字符型和图像型，构成了多源异构的海量数据集^[21-22]，因此需要建立征提取方法将不同来源不同格式的数据特征提取并归一化到相同尺度。

1.1 二维数据特征提取方法

深层煤岩气地质数据的存储类型为二维矩阵，地质数据的尺寸大小不一存在多尺度的问题。深层煤岩气二维数据特征提取方法是基于卷积自编码结合空间金字塔池化。提取深层煤岩气数据中矩阵类型数据的全局特征，利用卷积自编码器经过多次非线性特征表征，从而达到降维的目的^[17]。空间金字塔池化通过将多尺度矩阵提取为相同维度以解决多尺度问题。

1.2 三维数据特征提取方法

由于深层煤岩气的缝网延伸模拟数据^[23]存储

为三维点云数据,且点云的数据规模较大,不适用于多视图^[24]与体素化方式^[25],点云神经网络(PointNet)局部提取特征的能力不足^[26],图神经网络效果不理想^[27],因此本文采用 PointNet 的进阶版 PointNet++^[28]提取三维点云数据特征,其网络结构图如图 1 所示。

PointNet++ 网络的方法是:通过两个抽象模块处

理后,点云数据送入 PointNet 网络提取特征,再利用全连接网络最终获得分类结果。该网络核心是抽象模块层,主要是由取样、分组和 PointNet 这三部分来组成。图中的 $N \times (d+C)$ 表示输入点云矩阵,其中 N 表示点云的个数, C 表示特征维数, d 表示点云维度(当坐标为 x 、 y 、 z 时, $d=3$), n_1 , n_2 , n' 表示采样后点数, C_1 , C_2 , C' 表示新的特征维数。

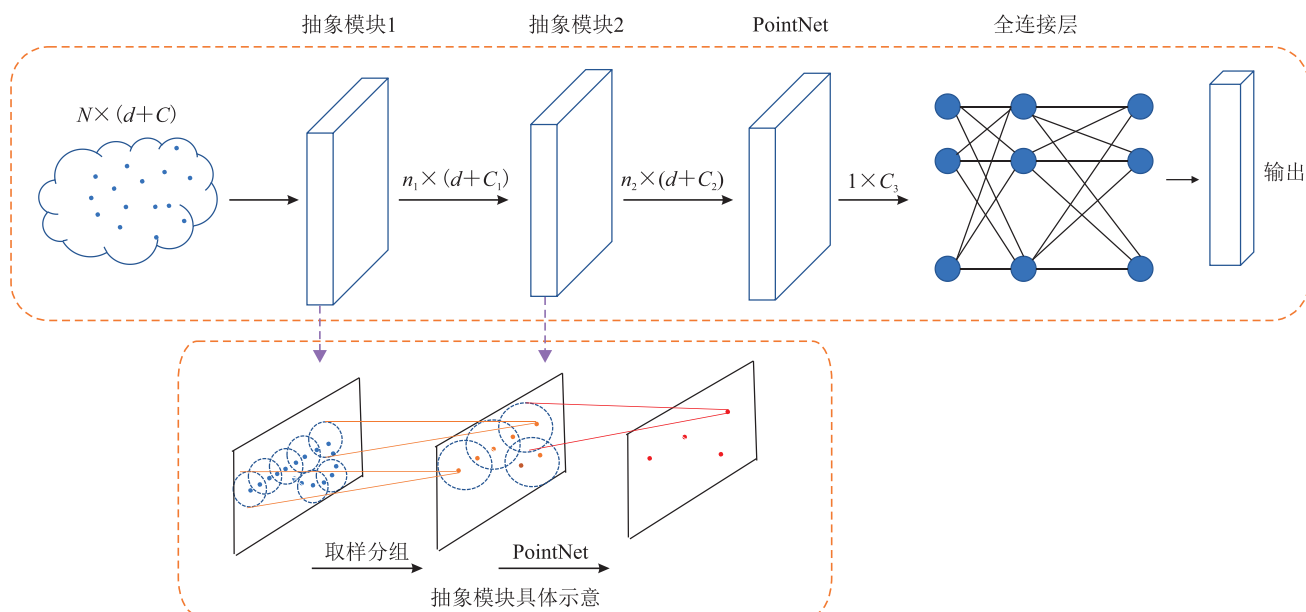


图 1 PointNet++ 网络结构图

深层煤岩气原始压裂数据为二维矩阵,直井可直接采用二维数据的特征提取。但是由于水平井压裂还包括段数、段长、簇间距等参数。这些压裂数据人为刻画不全面,压裂段数不一致,因此利用缝网模拟的三维数据代替原始压裂数据。要将二维的地质数据和排采数据,与缝网延伸的三维数据输入预测模型,两种数据的维度不一致,需要将二维数据与三维数据进行特征提取并降成与排采数据相同的维度,再进行特征拼接。因此,设计深层煤岩气三维数据特征提取的网络结构图见图 2。

具体算法如下:

- 1) 从包含 N 个点的点云数据集中随机选取点 P_0 作为初始点,构成采样点集 $S = \{P_0\}$ 。
- 2) 计算所有点到 P_0 的欧氏距离,构成 N 维数组 L ,选择距离最大值所对应的点 P_1 ,更新采样点集合 $S = \{P_0, P_1\}$ 。
- 3) 计算所有点到 P_1 的距离,若每一个点 P_i 到 P_1 的距离小于 $L[i]$,则更新 $L[i] = d[P_i, P_1]$ 。
- 4) 选取 L 中最大值所对应的点,将其作为 P_2 ,更新集合 $S = \{P_0, P_1, P_2\}$ 。

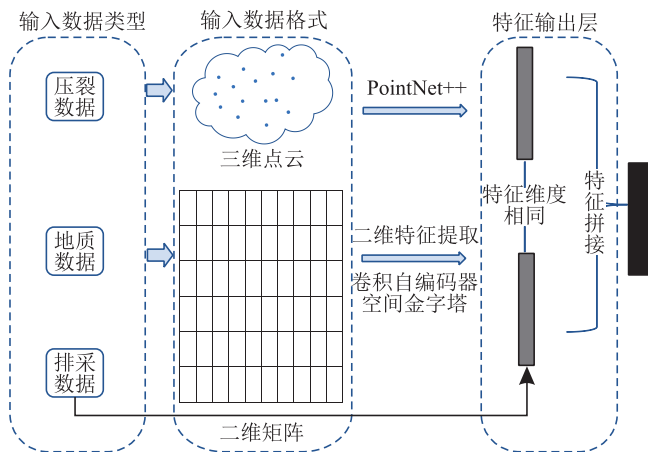


图 2 特征提取新方法的网络结构图

- 5) 重复步骤 2) ~ 4), 依次迭代直至满足采样数据点的数量要求为止。分组是选取采样中心点 n' , 以采样点为中心, 在小于半径参数的局部区域中选取 K 个点进行聚合, 聚合后的矩阵为 $n' \times K \times (d+C)$ 。PointNet 的作用是提取通过分组聚合后的矩阵局部特征, 提取特征后矩阵维度变为 $n' \times K \times (d+C)$ 。

上述方法通过点云神经网络的进阶版 PointNet++

解决了三维数据的特征提取问题，将缝网模拟的三维点云数据经 PointNet++ 提取后的特征与地质数据经过二维数据提取后的特征拼接，最后拼接后的特征与排采数据输入预测模型中。

2 预测模型

2.1 数据聚类

中国煤层气开发的显著特征是区块内和区块间的地质、岩石力学、地应力、压裂施工等参数差异性较大^[29]，也导致改造效果^[30-31]和最终的裂缝参数值差异大^[32-34]，因此对于新井产量预测即使是同一生产区块的井也需要数据聚类。

深层煤岩气数据的聚类具有不确定性，其聚类中心通常取决于研究目标。依据以上要求，本文沿用以前模型里聚类使用的近邻传播算法，并根据所研究的数据样本选出更合适、更高质量的聚类中心，进而进行聚类分析^[16]。AP 算法也称为仿射传播算法，是 Frey 等^[35]2007 年在《Science》上首次提出的一种聚类算法。与传统聚类算法相比，AP 算法进行了许多优化设计，例如不要求预先设定类的数量，无须距离矩阵的对称性，从而能够耗费更短的时间得到更好的聚类结果^[36]。

2.2 融入注意力机制的长短时记忆网络

对于一般常用于信息数据处理上的编码器—解码器模型，编码器的作用是将原序列信息数据转换为遵循一定规范长度的上下文向量，且权值相等。因为权值相同，当原序列数据信息较多时，预测结果下降，加入注意力机制可以有效地解决此类问题。注意力机制是用来提升编码—解码模型效果的机制，指在某些时刻，将注意力分配给需要关注的重点部分，将非重点部分注意力减少，以在重点区域获得更多的有用信息^[37]。其核心含义是放大所需信息及忽视无关信息，类似于人类脑力注意力分配功能。注意力机制对于模型输入序列的不同时刻的隐含层向量赋予相应的权重，将隐含层向量按照不同的重要程度合并为新的隐含层向量并输入到解码器中。

经过编码—解码结构处理的输入、输出为 Source、Target，Source 和 Target 分别由各自的元素序列构成：

$$\begin{cases} \text{Source} = \langle x_1, x_2, \dots, x_m \rangle \\ \text{Target} = \langle y_1, y_2, \dots, y_m \rangle \end{cases} \quad (1)$$

编码器对输入值 Source 进行编码，将输入值通过非线性变换转化为中间值 (C_i)，每个 C_i 对应着不同的源输入元素的注意力分配概率分布，再对构成元素加权求和。

解码器是根据输入值 Source 的中间值 (C) 和已生成的历史信息来生成 i 时刻的输出。

$$\begin{cases} y_1 = f_1(C_1) \\ y_2 = f_2(C_2, y_1) \\ y_3 = f_3(C_3, y_1, y_2) \end{cases} \quad (2)$$

式中 f_i 表示解码器的非线性变换函数，每个 y_i 依次产生，则系统根据输入值 Source 生成了目标值 Target。

长短时记忆神经网络融合注意力机制，可克服传统的 LSTM 在编码—解码过程在使用固定长度向量的缺陷，提高预测性能。融合了注意力机制的煤层气产量预测模型能判断各输入时刻的重要程度，通过凸显重要信息进而提高模型训练效率和模型预测结果的准确率，降低计算和存储开销。在传统的 LSTM 网络中，随机分配权重的方式往往会因数据信息过多造成关键信息的缺失的特点，运用概率分配权重的方法可改善这一缺点，融合注意力机制的 LSTM 预测模型结构见图 3。

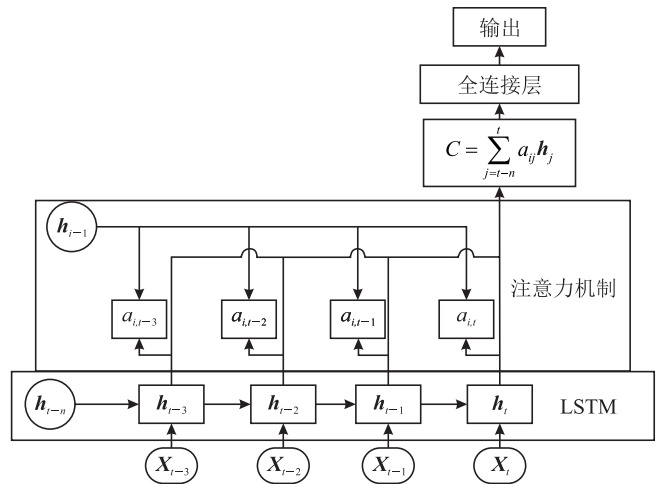


图 3 融合注意力机制的 LSTM 预测模型结构图

在图 3 中， $X_{t-n}, \dots, X_{t-3}, X_{t-2}, X_{t-1}, X_t$ 表示待预测时刻之前的 n 个输入的多维特征向量。输入的向量经过 LSTM 网络处理之后，可以得到若干个隐含层的向量 $h_j (k = t-n, \dots, t)$ 。由编码器中的隐含层 h_j 与解码器中另外一个长短时记忆神经网络前一时刻的隐含层 h'_{j-1} 计算得到注意力系数。注意力系数计算公式如下：

$$\begin{cases} e_{ij} = v \tanh(W \cdot h_j + U \cdot h'_{i-1} + b) \\ \alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=t-n}^i \exp(e_{ik})} \\ C = \sum_{j=t-n}^i a_{ij} h_j \end{cases} \quad (3)$$

式中 e_{ij} 表示通过计算求得的 h'_{i-1} 和 h_j 间的相关系数, 数值越高则关联度越大; α_{ij} 表示注意力系数; h_j 表示隐含层状态; C 表示注意力层的输出。

最后, 将注意力系数对应分配到每个不同的中间层状态 h_j , 并通过加权求和的方式求得输入到解码器的向量 C , 通过全连接层计算出下一时刻的产量预测结果。从而就得到融合注意力机制的 LSTM 网络, 它通过输入目标井多个参数的时间序列数据, 来获取煤层气产量的预测曲线。

2.3 预测步骤

1) 将煤层气测井数据以上文 1.1 中提出的二维特征提取方法进行特征提取, 获得地质数据特征 (X_A)。

2) 将煤层气缝网数值模拟结果以上文 1.2 中提出的三维数据特征提取方法进行特征提取, 获得缝网数据特征 (X_B)。

3) 将地质数据特征 (X_A) 与缝网数据特征 (X_B) 进行拼接, 为每个煤层气井建立特征 $x_i = (x_{i1}, x_{i2}, \dots, x_{id})^T$, 按照上文 2.1 中提出的近邻传播算法对数据样本进行聚类, 筛选出目标井的相似井, 得到建立煤层气产量预测模型的数据集。

4) 将包含地质数据特征、缝网数据特征和排采数据的数据集划分为训练集和测试集, 对多个时间序列数据集进行划分样本, 作为上文 2.2 提出的融合注意力机制的长短时记忆网络的样本数据。

5) 将训练样本输入融合注意力机制的长短时记忆网络进行网络模型的训练, 通过测试样本检测模型的准确率, 得到深层煤岩气产量预测模型。

针对相同的煤层气产量预测问题, 训练样本的时间序列个数增加并不会对煤层气井特征参数的个数和预测参数的个数产生影响, 目标是学习 T 个时刻的输入值 ($x_1, x_2, \dots, x_t$) 和 T 时刻的输出值 (y_t) 之间的非线性关系。多序列网络模型的构建关键在于处理样本数据的方法, 可根据时间图窗切割和划分多时序列, 然后批量训练。 B 个样本井的时间序列长度记为 $L_b (b = 1, 2, \dots, B)$ 。

据此, 数据样本的批次大小 $batch_size$ 为:

$$batch_size = \sum_{b=1}^B \left(\left\lceil \frac{L_b - T}{A} \right\rceil + 1 \right) \quad (4)$$

特别地, 当 $A=T$ 时数据样本的批次大小 $batch_size$ 为:

$$batch_size = \sum_{b=1}^B \frac{L_b}{T} \quad (5)$$

式中 B 表示与研究目标同类的煤层气井的个数; L_b 表示第 b 个煤层气井的时间序列长度; T 表示时间窗口大小。

找到批量大小 $batch_size$ 和输入尺寸 $inputs_size$ 来定义细胞状态, 并且初始化单元的状态。然后, 结合输出大小, 定义整个 AP-LSTM 的结构以执行网络的前向计算。通过 BPTT 算法的反向传播计算 AP-LSTM 的后传播。BPTT 算法不仅考虑了层级间的纵向传播还考虑了时间上的横向传播, 同时在两个方向上进行参数优化。以 T 个时刻中预测输出值与实际值的误差最小为目标, 通过梯度信息对模型参数进行修改, 学习到最优的权重矩阵及偏置。

2.4 评价指标

2.4.1 聚类模型评价指标

鉴于模型特点, 采用轮廓系数 (Silhouette) 指标进行评价。设有 n 个样本点的数据集划分为 k 个类别 $C_i (i=1, 2, \dots, k)$, $a(t)$ 是类别 C 中样本 t 与该类别中其他样本点的平均不相似度或距离, $d(t, C_i)$ 是一个类别中的样本 t 与另一个类别中的所有样本点的平均不相似度或距离, 则得到轮廓系数指标计算方式为:

$$Sil(t) = \frac{[b(t) - a(t)]}{\max\{a(t), b(t)\}} \quad (6)$$

$$b(t) = \min\{d(t, C_i)\} \quad (i=1, 2, \dots, k, i \neq j) \quad (7)$$

类别中所有样本的平均 Sil 值代表该类的紧密性和可分性, 取值范围 $[-1, 1]$, 平均 Sil 值越大, 表明聚类效果越好。

2.4.2 预测模型评价指标

在深层煤岩气预测模型中, 预测结果的优劣可以通过深层煤岩气产量预测值与真实值间的损失函数进行判断, 即平均绝对百分比误差 (MAPE), 表达为:

$$MAPE = \frac{100}{N} \sum_{i=1}^N \left| \frac{y_i - f_i}{y_i} \right| \quad (8)$$

式中 y_i 表示第 i 个样本的实际值; f_i 表示第 i 个样本的预测值; N 表示样本总数。

3 实例应用

利用 DJ 区块已有的并且数据量完整的 200 口深层煤岩气井（158 口直井与 42 口水平井）为研究对象。随机选取 120 口直井和 30 口水平井作为训练样本，其余 50 口井用于验证。其中，输入的二维地质特征数据集共包含 12 个参数，分别是：自然伽马、自然电位、声波时差、补偿中子、密度、泥质含量、

深侧向电阻率、浅侧向电阻率、孔隙度、温度、井斜角、井斜方位角。三维数据即各井段的缝网特征。在验证准确性后，对 DJ 区块的两口井在投产前进行了产量预测，并将预测结果与实际产量进行对比。

为了将预测结果与其他深度学习方法以及数值模拟和浅层神经网络方法进行比较，还使用 DNN^[13]、GRU^[15]、LSTM^[16]、数值模拟^[38]、BP 神经网络^[39] 方法用于验证集和测试集（图 4）。

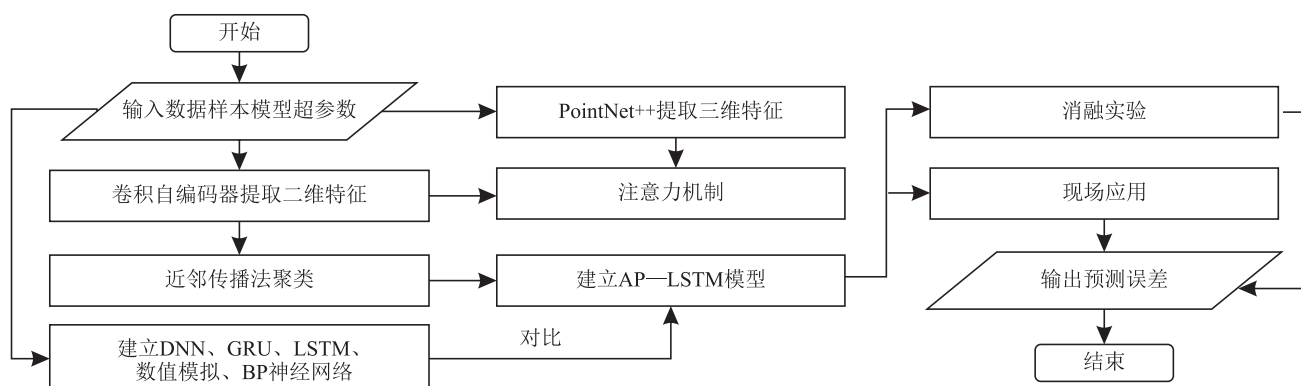


图 4 模型对比预测流程图

3.1 训练和验证

图 4 给出了模型对比及预测的整体框架流程，模型训练过程中具体的参数取值如下：

1) 特征提取。每口井的包含 12 个地质特征，因此特征提取网络输入层的宽度设置为 12，长度由井中测井数据的行数决定。在该网络结构中，卷积核大小设置为 5×5 ，dropout 比率设置为 0.25。在训练模型时，epoch 设置为 30，batch size 设置为 125。隐藏层数设置为 2。三层金字塔池将输入数据分成 16 块、4 块、1 块，输出 1×21 维的特征提取结果。

2) 样本聚类。采用 AP 算法对 150 个训练样本的特征提取结果进行聚类，初始值 P_0 取为相似度矩阵的中位数，阻尼因子 (λ) 取 0.5，最终聚类得到 A、B、C、D、E、F、G 共 7 个类别，聚类数如图 5 所示。

3) 建立 LSTM 预测模型。采用同类样本作为输入数据。在该 LSTM 网络结构中，时间窗 (T) 的初始值设为 25，滑动步长为 1。输入大小为 12，输出大小为 1。

3.2 应用和分析

在 DJ20 井开产前，将结合了三维特征提取及注意力机制的基础模型（即 AP-LSTM）应用于该井的产量预测，并结合深度学习 DNN、GRU、LSTM、数值模拟、BP 神经网络方法进行对比。其中，数值

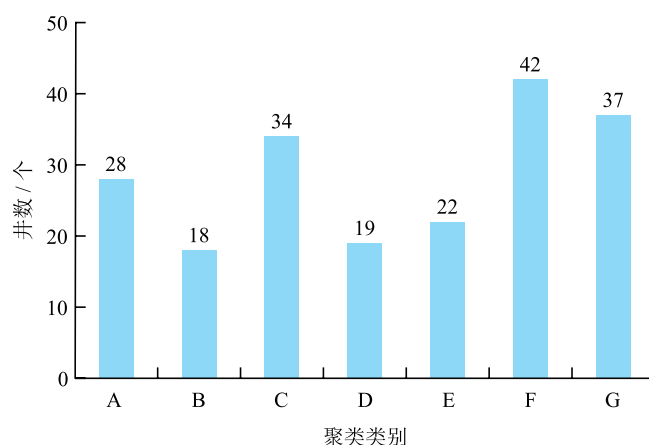


图 5 训练样本聚类结果图

模拟的参数设置如表 1 所示。该井的产气量进行了 110 d 的监测，并比较了实际值和预测值，BP 神经网络选择生产 15 ~ 20 d、45 ~ 50 d 和 105 ~ 110 d 进行计算。

此外，为验证三维特征提取和注意力机制模块对 AP-LSTM 的提升效果，通过消融实验的方式对不同模块下的深度学习模型预测结果进行了对比。该井的 110 d 的预测结果如图 6 所示，消融实验结果如图 7 所示。所有方法的预测误差汇总如表 2 所示。

从图 6 可以看出，文章建立的融入三维特征提取及注意力机制的 AP-LSTM 模型预测效果最好。

表 1 两口井数值模拟部分参数设置结果表

井号	深度 /m	孔隙度	渗透率 / mD	密度 / ($\text{g} \cdot \text{m}^{-3}$)	含气饱和度	排量 / ($\text{m}^3 \cdot \text{min}^{-1}$)	砂比	射孔密度 / (孔 $\cdot \text{m}^{-1}$)	射孔方式
DJ01	2 501 ~ 2 505	4.29%	0.03	1.55	94.02%	9.8	10.85%	16	相位 60°, 扇面向下
DJ20	2 446 ~ 2 450	4.24%	0.02	1.57	96.99%	17.6	15.27%	16	相位 60°, 扇面向下

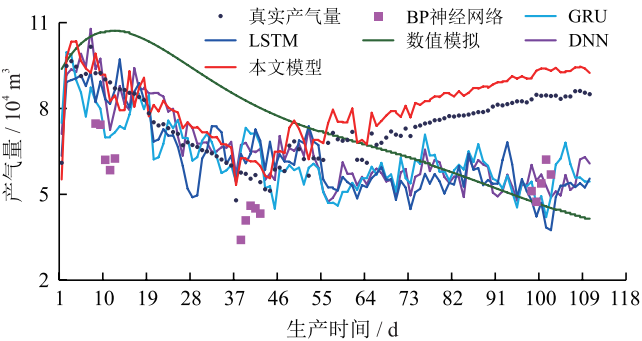


图 6 DJ20 井产气量预测结果图

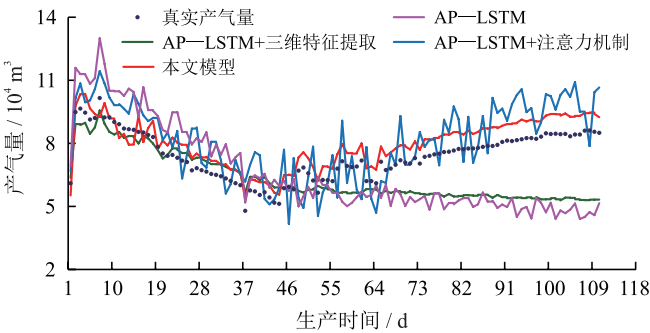


图 7 DJ20 井消融实验结果图

表 2 DJ20 井预测误差表

算法	添加模块	MAPE
AP—LSTM	三维特征提取 + 注意力机制	9.37%
	三维特征提取	17.23%
	注意力机制	15.18%
	无	24.35%
DNN	无	25.14%
GRU	无	25.72%
LSTM	无	26.43%
数值模拟	无	28.23%
BP 神经网络	无	29.16%

DNN、GRU、LSTM 预测产量的波动较大，且与真实产量的趋势存在较大差异。数值模拟和 BP 神经网络的预测效果较差。结合表 2 可以看出，本文模型的预测结果与真实产量的符合程度最高，预测的 MAPE

为 9.37%，明显低于 DNN、GRU、LSTM 等模型。

如图 7 所示，未添加任何模块的 AP—LSTM 在预测长期产量的过程中数据波动较大，前期 AP—LSTM 的预测值与真实产量的变化趋势基本一致，从第 45 d（该井实施转抽）开始二者差异越来越大。在 AP—LSTM 的基础上仅添加三维特征提取，结果显示预测产量的波动显著减小但趋势无明显改善，表明 AP—LSTM 在引入深层煤岩气缝网信息后预测效果有所提升但无法识别转抽等关键信息。在 AP—LSTM 的基础上仅添加注意力机制，结果显示预测产量存在较大波动但整体趋势与真实产量基本一致，表明注意力机制有助于 AP—LSTM 识别出转抽等关键信息但无法解决预测过程中丢失缝网信息后的数据波动问题。结合三维特征提取及注意力机制的深度学习新模型在预测长期产量的过程中数据波动较小，且预测产量与真实产量的趋势基本一致，预测效果最优。

结合表 2 的结果可以看出，数值模拟和 BP 神经网络在预测长期产量方面都不如深度学习。不添加任何模块的 AP—LSTM 预测误差为 24.35%，仅添加三维特征提取后预测误差下降到 17.23%，仅添加注意力机制后预测误差下降到 15.18%，表明三维特征提取和注意力机制都能够有效提升 AP—LSTM 的预测效果。融入三维特征提取和注意力机制的 AP—LSTM 预测误差仅为 9.37%，表明三维特征提取和注意力机制二者结合能够进一步提升模型的预测效果。

针对目标区块另一口生产时间较长的水平井（DJ01 井），收集了该井 800 d 的生产数据，将结合了三维特征提取及注意力机制的 AP—LSTM 应用于该井的产量预测，同时结合数值模拟进行对比，进一步验证本文模型的可靠性（图 8），不同天数下的预测误差如表 3 所示。

从图 8 可以看出，数值模拟方法在 DJ01 井生产初期的拟合效果较好，但由于无法有效识别出第 100 d 左右的转抽，导致之后的预测出现明显偏差。结合表 3 可以看出，随着生产天数的增加，数值模拟的预测误差基本呈显著递增的趋势，而文章建立的模型预测的产量无论在前中后期都与真实产量吻合较好，进一步验证了本文模型的可靠性。

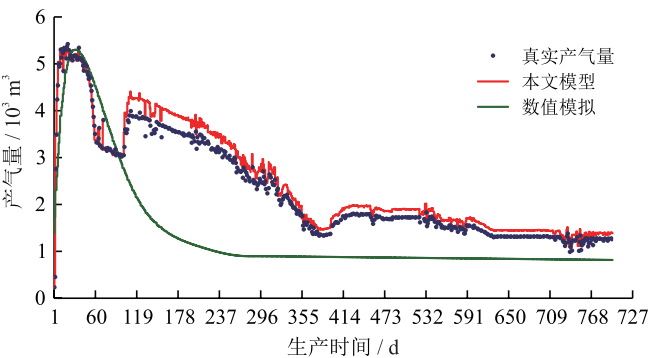


图 8 DJ01 井产气量预测结果图

基于 DJ 区块水平井和直井的生产数据建立了融入三维特征提取和注意力机制的 AP—LSTM 模型，为进一步分析该模型的普适性，收集整理了鄂尔多斯盆地临兴—神府区块、鄂尔多斯盆地神木—佳县区块、沁水盆地长治北区块的生产井数据进行验证。

表 3 DJ01 井不同天数产量预测误差表

算法	MAPE			
	50 d	200 d	500 d	800 d
本文模型	1.50%	5.27%	8.10%	8.64%
数值模拟	9.14%	27.35%	37.35%	35.26%

三大区块的主要地质和储层参数如表 4 所示。除了 DJ 区块 200 口井的迁移结果外，还对比了加入目标区块 10 口井之后和加入目标区块 20 口井之后的模型预测结果，如表 5 所示。

从表 5 可以看出，基于 DJ 区块建立的模型直接迁移到新的区块，产量预测的 MAPE 基本低于 20%，平均的 MAPE 为 18.84%，验证了文章模型的普适性。此外，随着新区块样本数量的增加，模型的预测准确率逐渐增大。

表 4 三大区块主要地质及储层参数表

区块	煤阶	深度 /m	储层厚度 /m	含气量 /(m ³ · t ⁻¹)	渗透率 /mD	压力系数
临兴—神府	中低	1 500 ~ 2 200	1.6 ~ 19.6	6.0 ~ 30.9	1.00 ~ 4.00	0.9 ~ 1.2
神木—佳县	高	1 600 ~ 2 800	4.0 ~ 18.0	17.7 ~ 22.8	0.01 ~ 0.12	0.9 ~ 1.3
长治北	中高	1 000 ~ 1 900	1.1 ~ 6.7	8.0 ~ 20.0	0.06 ~ 0.35	0.8 ~ 1.2

表 5 模型迁移结果表

区块	井号	MAPE		
		仅用 DJ 区块数据	加入目标区块 10 口井数据	加入目标区块 20 口井数据
临兴—神府	LS01	19.81%	16.04%	10.43%
	LS02	20.73%	17.31%	11.37%
	LS03	20.04%	16.38%	11.02%
神木—佳县	SJ01	17.79%	15.12%	9.28%
	SJ02	19.63%	16.25%	10.08%
	SJ03	17.38%	15.43%	9.54%
长治北	CZB01	18.52%	15.21%	9.91%
	CZB02	18.17%	16.07%	10.55%
	CZB03	17.45%	15.11%	11.03%
平均		18.84%	15.88%	10.36%

4 结论

1) 卷积自编码结合空间金字塔池化统一了二维地质数据的尺度，采用 PointNet++ 提取了缝网三维点云数据特征，有效解决了深层煤岩气数据多元异构、多尺度等问题。

2) 考虑缝网三维特征，将注意力机制融入长短时记忆网络，建立了预测深层煤岩气产量的深度学习新模型，有效提高了产量预测的准确率，预测的 MAPE 为 9.37%。

3) 缝网特征是影响深层煤岩气产量的重要因素，引入深层煤岩气缝网三维特征，AP—LSTM 模型预测的 MAPE 从 24.35% 下降到 17.23%，模型的泛化能力有效提升。

4) 融入三维特征提取和注意力机制的 AP—LSTM 模型预测出的产量与真实的生产曲线趋势基本一致，该模型能够有效区分自喷和抽吸，适应了深层煤岩气的生产方式，普适性较强，可为深层煤岩气规模化开发提供技术支撑。

参 考 文 献

[1] 门相勇, 娄钰, 王一兵, 等. 中国煤层气产业“十三五”以来发展成效与建议[J]. 天然气工业, 2022, 42(6): 173-178.
MEN Xiangyong, LOU Yu, WANG Yibing, et al. Development achievements of China's CBM industry since the 13th Five-Year Plan and suggestions[J]. Natural Gas Industry, 2022, 42(6): 173-178.
[2] 贾承造, 郑民, 张永峰. 中国非常规油气资源与勘探开发前景

- [J]. 石油勘探与开发, 2012, 39(2): 129-136.
- JIA Chengzao, ZHENG Min, ZHANG Yongfeng. Unconventional hydrocarbon resources in China and the prospect of exploration and development[J]. Petroleum Exploration and Development, 2012, 39(2): 129-136.
- [3] 门相勇, 韩征, 宫厚健, 等. 新形势下中国煤层气勘探开发面临的挑战与机遇[J]. 天然气工业, 2018, 38(9): 10-16.
- MEN Xiangyong, HAN Zheng, GONG Houjian, et al. Challenges and opportunities of CBM exploration and development in China under new situations[J]. Natural Gas Industry, 2018, 38(9): 10-16.
- [4] 徐凤银, 闫霞, 李曙光, 等. 鄂尔多斯盆地东缘深部(层)煤层气勘探开发理论技术难点与对策[J]. 煤田地质与勘探, 2023, 51(1): 115-130.
- XU Fengyin, YAN Xia, LI Shuguang, et al. Theoretical and technological difficulties and countermeasures of deep CBM exploration and development in the eastern edge of Ordos Basin[J]. Coal Geology & Exploration, 2023, 51(1): 115-130.
- [5] 闫霞, 徐凤银, 聂志宏, 等. 深部微构造特征及其对煤层气高产“甜点区”的控制——以鄂尔多斯盆地东缘大吉地区为例[J]. 煤炭学报, 2021, 46(8): 2426-2439.
- YAN Xia, XU Fengyin, NIE Zhihong, et al. Microstructure characteristics of Daji area in east Ordos Basin and its control over the high yield dessert of CBM[J]. Journal of China Coal Society, 2021, 46(8): 2426-2439.
- [6] YAO Shanshan, WANG Xiangzeng, YUAN Qingwang, et al. Production analysis of multifractured horizontal wells with composite models: Influence of complex heterogeneity[J]. Journal of Hydrology, 2020, 583: 124542.
- [7] 李曙光, 王成旺, 王红娜, 等. 大宁—吉县区块深层煤层气藏特征及有利区评价[J]. 煤田地质与勘探, 2022, 50(9): 59-67.
- LI Shuguang, WANG Chengwang, WANG Hongna, et al. Reservoir forming characteristics and favorable area evaluation of deep coalbed methane in Daning-Jixian Block[J]. Coal Geology & Exploration, 2022, 50(9): 59-67.
- [8] 周德华, 陈刚, 陈贞龙, 等. 中国深层煤层气勘探开发进展、关键评价参数与前景展望[J]. 天然气工业, 2022, 42(6): 43-51.
- ZHOU Dehua, CHEN Gang, CHEN Zhenlong, et al. Exploration and development progress, key evaluation parameters and prospect of deep CBM in China[J]. Natural Gas Industry, 2022, 42(6): 43-51.
- [9] YOUSEF A M, KAVOUSHI G P, ALNUAIMI M, 等. 预测数据分析法在中东某老油田提高采收率中的应用[J]. 石油勘探与开发, 2020, 47(2): 366-371.
- YOUSEF A M, KAVOUSHI G P, ALNUAIMI M, et al. Predictive data analytics application for enhanced oil recovery in a mature field in the Middle East[J]. Petroleum Exploration and Development, 2020, 47(2): 366-371.
- [10] 李曙光, 王红娜, 徐博瑞, 等. 大宁—吉县区块深层煤层气井酸化压裂产气效果影响因素分析[J]. 煤田地质与勘探, 2022, 50(3): 165-172.
- LI Shuguang, WANG Hongna, XU Borui, et al. Influencing factors on gas production effect of acid fractured CBM wells in deep coal seam of Daning-Jixian Block[J]. Coal Geology & Exploration, 2022, 50(3): 165-172.
- [11] ZHAO Zhibo, CHEN Yuhua, ZHANG Yi, et al. A deep learning model for predicting the production of coalbed methane considering time, space, and geological features[J]. Computers & Geosciences, 2023, 173: 105312.
- [12] 匡立春, 刘合, 任义丽, 等. 人工智能在石油勘探开发领域的应用现状与发展趋势[J]. 石油勘探与开发, 2021, 48(1): 1-11.
- KUANG Lichun, LIU He, REN Yili, et al. Application and development trend of artificial intelligence in petroleum exploration and development[J]. Petroleum Exploration and Development, 2021, 48(1): 1-11.
- [13] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [14] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [15] LEE K, LIM J, YOON D, et al. Prediction of shale-gas production at Duvernay Formation using deep-learning algorithm[J]. SPE Journal, 2019, 24(6): 2423-2437.
- [16] RAHMANIFARD H, GATES I, ASL A S. Comparison of machine learning and statistical predictive models for production time series forecasting in tight oil reservoirs[C]//SPE/AAPG/SEG Unconventional Resources Technology Conference. Houston: SPE, 2022: URTEC-3703284-MS.
- [17] GUO Zixi, ZHAO Jinzhou, YOU Zhenjiang, et al. Prediction of coalbed methane production based on deep learning[J]. Energy, 2021, 230: 120847.
- [18] 秦勇, 申建, 李小刚. 中国煤层气资源控制程度及可靠性分析[J]. 天然气工业, 2022, 42(6): 19-32.
- QIN Yong, SHEN Jian, LI Xiaogang. Control degree and reliability of CBM resources in China[J]. Natural Gas Industry, 2022, 42(6): 19-32.
- [19] 国土资源部油气资源战略研究中心. 煤层气资源动态评价[M]. 北京: 地质出版社, 2017.
- Oil and Gas Resources Strategic Research Center of the Ministry of Land and Resources. National coalbed methane resource assessment[M]. Beijing: Geological Publishing House, 2017.
- [20] GUO Zixi, CHEN Yiyu, YAO Shanshan, et al. Feasibility analysis and optimal design of acidizing of coalbed methane wells[J]. Journal of Energy Resources Technology, 2019, 141(8): 082906.
- [21] PENG Cheng, ZOU Changchun, ZHOU Tianning, et al. Factors affecting coalbed methane (CBM) well productivity in the Shizhuangnan block of southern Qinshui Basin, North China: Investigation by geophysical log, experiment and production data[J]. Fuel, 2017, 191: 427-441.
- [22] YONG Qin, MOORE T A, SHEN Jian, et al. Resources and geology of coalbed methane in China: A review[J]. International Geology Review, 2018, 60(5/6): 777-812.
- [23] GUO Zixi, ZHAO Jinzhou, SUN Xinhao, et al. A novel continuous fracture network model: Formation mechanism, numerical simulation, and field application[J]. Geofluids, 2022:

- 4026200.
- [24] 朱威, 绳荣金, 汤如, 等. 基于动态图卷积和空间金字塔池化的点云深度学习网络[J]. 计算机科学, 2020, 47(7): 192-198.
ZHU Wei, SHENG Rongjin, TANG Ru, et al. Point cloud deep learning network based on dynamic graph convolution and spatial pyramid pooling[J]. Computer Science, 2020, 47(7): 192-198.
- [25] MATURANA D, SCHERER S. VoxNet: A 3D Convolutional Neural Network for real-time object recognition[C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hamburg: IEEE, 2015: 922-928.
- [26] CHARLES R Q, SU Hao, MO Kaichun, et al. PointNet: Deep learning on point sets for 3D classification and segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 77-85.
- [27] GORI M, MONFARDINI G, SCARSELLI F. A new model for learning in graph domains[C]//Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005. Montreal: IEEE, 2005: 729-734.
- [28] QI C R, YI Li, SU Hao, et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017: 5105-5114.
- [29] 杨长鑫, 杨兆中, 李小刚, 等. 中国煤层气地面井开采储层改造技术现状与展望[J]. 天然气工业, 2022, 42(6): 154-162.
YANG Changxin, YANG Zhaozhong, LI Xiaogang, et al. Status and prospect of reservoir stimulation technologies for CBM surface well production in China[J]. Natural Gas Industry, 2022, 42(6): 154-162.
- [30] GUO Zixi, CHEN Yiyu, ZHOU Xiang, et al. Inversing fracture parameters using early-time production data for fractured wells[J]. Inverse Problems in Science and Engineering, 2020, 28(5): 674-694.
- [31] ZHAO Yunxiang, GUO Dali, ZHANG Chuanxin, et al. Calculation method of fracture conductivity based on performance of proppant material[J]. Science of Advanced Materials, 2021, 13(3): 417-426.
- [32] GUO Dali, ZHAO Yunxiang, GUO Zixi, et al. Theoretical and experimental determination of proppant crushing rate and fracture conductivity[J]. Journal of Energy Resources Technology, 2020, 142(10): 103005.
- [33] GUO Zixi, ZHAO Jinzhou, LI Yongming, et al. Theoretical and experimental determination of proppant-crushing ratio and fracture conductivity with different particle sizes[J]. Energy Science & Engineering, 2022, 10(1): 177-193.
- [34] ZHANG Qiushi, WANG Daohan, ZENG Fanhua, et al. Pressure transient behaviors of vertical fractured wells with asymmetric fracture patterns[J]. Journal of Energy Resources Technology, 2020, 142(4): 043001.
- [35] FREY B J, DUECK D. Clustering by passing messages between data points[J]. Science, 2007, 315(5814): 972-976.
- [36] LI Yang, GUO Chonghui, SUN Leilei. Fast clustering by affinity propagation based on density peaks[J]. IEEE Access, 2020, 8: 138884-138897.
- [37] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.
- [38] WANG Gongda, REN Ting, QI Qingxin, et al. Prediction of coalbed methane (CBM) production considering bidisperse diffusion: Model development, experimental test, and numerical simulation[J]. Energy & Fuels, 2017, 31(6): 5785-5797.
- [39] ZHOU Fengde, YAO Guangqing, TYSON S. Impact of geological modeling processes on spatial coalbed methane resource estimation[J]. International Journal of Coal Geology, 2015, 146: 14-27.

(修改回稿日期 2024-07-17 编辑 董 莎)



本文
互动