【环境与能源／Environment and Energy】

基于深度学习的稠油蒸汽驱汽窜时间预测方法

崔传智[1]，陆水青山[1]，吴忠维[1]，盖平原[2]，刘廷峰[2]

1）中国石油大学(华东)非常规油气开发教育部重点实验室，山东青岛 266580；
2）中石化胜利油田分公司石油工程技术研究院，山东东营 257099

摘　要：实时可靠的汽窜时间预测方法可以为及时采取防治措施提供支持．结合深度学习算法以及油井自身动态数据的约束，提出一种预测稠油井汽窜时间的新方法．针对单一指标波动频繁，噪声大，以及无法准确表征汽窜时间等问题，根据油田实际注采数据，以参数组合的方式构建表征蒸汽窜流通道形成时间的指标参数，并结合变异系数-G1混合交叉赋权法融合成汽窜综合判识曲线．基于标准互信息的相似性度量方法选择合适的时间序列数据作为输入特征，以相应的汽窜综合判识曲线作为输出时间序列构建学习样本．采用序列到序列深度学习框架建立汽窜时间的预测模型进行实际预测，并与传统的机器学习方法进行对比，验证模型的有效性和优越性．该方法通过数据驱动的方式模拟了注采时间序列特征与汽窜判识曲线之间的映射关系，可有效提高汽窜时间预测的效率和精度，对汽窜智能预警具有一定指导意义．

关键词：油藏工程；稠油油藏；汽窜判识；汽窜时间预测；自然语言处理；序列到序列深度学习框架
中图分类号：TE34　　**文献标志码：**A　　**doi：**10.3724/SP.J.1249.2023.05622

Deep-learning-based prediction method of steam channeling time in heavy oil reservoirs

CUI Chuanzhi[1], LU Shuiqingshan[1], WU Zhongwei[1], GAI Pingyuan[2], and LIU Tingfeng[2]

1) MOE Key Laboratory of Unconventional Oil & Gas Development, China University of Petroleum (East China), Qingdao 266580, Shandong Province, P. R. China
2) Petroleum Engineering Technology Research Institute, Shengli Oilfield Company, SINOPEC, Dongying 257099, Shandong Province, P. R. China

Abstract: A real-time and reliable prediction method for the onset of a steam breakthrough in heavy oil wells is proposed by integrating deep learning algorithms with dynamic data from the wells. To address the issues such as frequent fluctuations and high noise level in single indicator data, and inability to accurately characterize the steam breakthrough time, a set of indicator parameters is constructed based on the actual injection and production data of the oilfield, effectively representing the formation time of the steam channel. These parameters are combined with the variation coefficient-G1 hybrid cross-weighting method to fuse into a comprehensive breakthrough identification curve of steam channeling. On this basis, suitable time series data are selected as input features using the similarity measurement method based on standard mutual information, with the corresponding breakthrough identification curve as the output time series for constructing the learning samples. A sequence-to-sequence deep learning framework is used to establish a prediction model for steam breakthrough time, and the effectiveness and superiority

Received: 2022-07-13; **Revised:** 2023-01-08; **Accepted:** 2023-03-06; **Online** (CNKI)**:** 2023-04-20
Foundation: National Natural Science Foundation of China (51974343); Funding for Postdoctoral Applied Research Project in Qingdao (qdyy20200084); China University of Petroleum (East China) Independent Innovation Research Program (Science and Engineering) (20CX06089A); 2021 Graduate Innovation Project Funded by School of Petroleum Engineering, East China University of Petroleum (YCX2021047)
Corresponding author: Professor CUI Chuanzhi (ccz2008@126.com)
Citation: CUI Chuanzhi, LU Shuiqingshan, WU Zhongwei, et al. Deep-learning-based prediction method of steam channeling time in heavy oil reservoirs [J]. Journal of Shenzhen University Science and Engineering, 2023, 40(5): 622-630. (in Chinese)

http://journal.szu.edu.cn

of the model are verified by the actual predictions and comparison with the traditional machine learning methods. This method simulates the mapping relationship between the characteristics of injection production time series and the identification curve of steam channeling in a data-driven way, which can effectively improve the efficiency and accuracy of steam channeling time prediction, and has certain guiding significance for the intelligent early warning of steam channeling.

Key words: reservoir engineering; heavy oil reservoir; steam channeling identification; prediction of steam channeling time; natural language processing; sequence-to-sequence deep learning framework

蒸汽驱往往伴随着汽窜现象的发生，汽窜通道的形成严重影响稠油油藏的开发效果[1-2]．实时可靠的稠油井汽窜时间预测方法可以为及时采取防治措施提供支持，提高稠油油藏的开采效果．

国内外学者对汽窜时间开展了相关研究．示踪剂监测法和试井分析法等传统的汽窜识别方法所需时间长，成本高，且一定程度上会影响油田的正常生产[3-4]．动态分析法由于具备现场数据获取方便和成本低等优势，已成为蒸汽窜流通道判别的主流方法，并且生产动态数据的波动性能够很好地反映注采井间的窜流关系．传统数学方法建立的预测模型难以适用于目前大数据背景下的油田实际应用场景[5-6]．伴随大数据和人工智能等技术的兴起与发展，传统的油田开发问题可以依赖信息时代的新兴技术加以解决[7-8]．其中，朱璇等[9]针对压裂井网气驱效果预测，结合误差逆传播算法(back propagation algorithm，BP)、长短期记忆(long short-term memory，LSTM)神经网络和双向长短期记忆(bidirectional long short-term memory，Bi-LSTM)神经网络等深度学习算法，建立日产油、地层压力和采出程度预测代理模型，为致密低渗油藏压裂井网气驱模拟预测提供了新途径；SYED 等[10]针对非常规油藏提高采收率方案设计问题，基于深度神经网络建立了提高采收率效果评价模型，该模型可以很好地预测不同时间轴下的采油性能．

现阶段采用深度学习方法对汽窜时间预测的研究较少，本研究在深度学方法和油井自身注采数据的支撑和约束下，提出一种基于数据驱动方式预测稠油井汽窜时间的新方法．通过实际的预测验证，该方法可以有效提高汽窜时间预测的效率和精度，对汽窜智能预警具有一定指导意义．

1　基于变异系数-G1 混合交叉赋权法的汽窜时间表征方法

基于变异系数-G1 混合交叉赋权法是一种汽窜时间表征方法，通过组合动态数据，构建了表征汽窜时间的指标参数，引入变异系数-G1 混合交叉赋权方法对各表征指标进行赋权，从而得到综合的汽窜时间判别曲线．

1.1　基于生产动态时间序列的汽窜时间表征指标

温度是蒸汽突破过程中一个重要因素．当高温蒸汽前沿突破到生产井时，井底温度会突然上升，井底温度曲线的拐点则标志着蒸汽突破的开始．然而，在油田实际应用中，由于井口温度受多种因素的影响，噪声较大，没有明显的拐点．汽窜后生产井的产水量绝大部分来自邻近窜通的注汽井，产水量的数据波动与注汽量的数据波动具有很强的相关性[11]．因此，选取日产油量(q_o)、日产水量(q_w)、注汽井的井口温度、生产井的井口温度、累积注汽量以及累积产油量等具有代表性，且易于测量的特征参数作为基本表征参数，通过基本参数的组合，得到表征蒸汽时间的指标R_1(生产井日产水量与日产油量之比)、R_2(生产井井口温度和注入井井口温度之比)和R_3(注汽井的累积注汽量和生产井的累积产油量之比)分别为

$$R_1 = \frac{q_w}{q_o} \tag{1}$$

$$R_2 = \frac{t_p}{t_i} \tag{2}$$

$$R_3 = \frac{G_i}{N_p} \tag{3}$$

其中，t_p和t_i分别为生产井和注入井的井口温度．G_i和N_p分别为注汽井的累积注汽量和生产井的累积产油量．R_1从瞬时产量的角度反映了蒸汽驱过程中生产井随日产油量变化的产水程度，R_1值越大，说明生产井日产水量增加程度越大，形成窜流通道的可能性越大．生产井井口温度t_p越高，R_2值越大，说明高温蒸汽向生产井方向移动，蒸汽前缘在此方向移动较快，有可能发生突破．R_3值越大，说明累积注汽量在增加的情况下油却没有被有效的驱替出来，汽窜通道形成的可能性大．

http://journal.szu.edu.cn

1.2 基于变异系数-G1混合交叉赋权法的表征参数赋权

为综合考虑上述表征指标的影响，需要确定各指标时间序列的权重，从而构建汽窜时间的综合判别曲线[12].

变异系数法属于体现数据本身客观信息的赋权方法，变异系数越大说明指标的影响越大，重要程度越大[13]. 变异系数赋权法计算流程如下：

1）计算变异系数

$$\sigma_k = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_{ki} - \overline{X}_k\right)^2} \tag{5}$$

$$V_k = \frac{\sigma_k}{\overline{X}_k} \tag{6}$$

其中，σ_k为第k个指标的标准差；V_k为第k个指标的变异系数；X_{ki}为第i个目标中第k个指标的原始数据($i=1, 2, \cdots, n$;$k=1, 2, \cdots, m$)，$\overline{X}_k$为第k个指标的平均值.

2）计算权重值

$$\omega_k = \frac{V_k}{\sum_{k=1}^{m} V_k} \tag{7}$$

其中，ω_k为第k个指标的权重值.

G1赋权法是一种主观赋权法，该方法凭借决策者对指标的重要性排序，从而根据相邻指标之间的相对重要程度确定各指标的权重. G1赋权法计算流程如下：

1）决策者确定指标排序

$$X_1 > X_2 > \cdots > X_k \tag{8}$$

其中，X_k为第k个指标($k=1, 2, \cdots, m$).

2）决策者确定相邻指标重要性程度之比的理性赋值

$$r_k = \omega^{s}_{k-1}/\omega^{s}_{k} \tag{9}$$

其中，ω^{s}_{k}和ω^{s}_{k-1}分别为第k和k-1个指标的权重系数. 重要性程度之比的赋值是人为规定.

3）计算权重值

$$\omega^{s}_{m} = \left(1 + \sum_{k=2}^{m}\prod_{i=k}^{m} r_i\right)^{-1} \tag{10}$$

其中，ω^{s}_{m}为第m个评价指标的权重.

基于变异系数-G1混合交叉赋权法使用各指标的变异系数代替了G1法中人为规定的步骤，从而确定相邻指标的重要性程度之比. 该方法很好地契合了汽窜判识应用场景中综合考虑现场经验以及数据客观信息的需求. 具体的计算步骤如下：

1）决策者对评价因素进行重要程度排序；

2）利用变异系数法赋权原理，采用式(6)计算变异系数V_k；

3）利用变异系数数值确定相邻评价因素的重要程度的比值r_k；

$$r_k = \begin{cases} \dfrac{V_{k-1}}{V_k}, & V_{k-1} \geqslant V_k \\ 1, & V_{k-1} < V_k \end{cases} \tag{11}$$

4）根据式(11)计算所得的r_k代入式(9)和式(10)，即可得到基于变异系数-G1混合交叉赋权法确定的各指标权重.

1.3 窜时间综合判识曲线的构建

按照上述计算步骤，结合油田生产数据，可以求出各生产井的汽窜时间综合判识曲线，结果如图1. 其中，陡然上升的拐点即为蒸汽窜流通道的形成初始时刻. 以图1中4口生产井为例，J4井的汽窜时间为第678天和第782天，J23井的汽窜时间为第674、1 860、2 050和第2 359天，J30井的汽窜时间为第892天以及第2 001天，J40井的汽窜时间为第1 726天. 将根据汽窜时间综合判识曲线得到的汽窜时间与该井实际汽窜时间进行对比，绝对误差均小于10 d，验证了该判识曲线的准确性.

2 基于序列-序列的汽窜时间预测模型

2.1 样本生成

试验工区为S油田稠油蒸汽驱区块，共有8个蒸汽驱井组，其中，注汽井8口，生产井40口，区块原油黏度为98 500 mPa·s，平均孔隙度变化范围为32~34，平均渗透率变化范围为3.158 4~3.454 5 μm^2. 动态数据为注汽井与生产井从2016年1月—2019年12月的注采数据，约5.2万条记录. 以生产井为基本单位，对应注入井的注入参数作为影响因素特征构建样本集，每个样本记录了包括注入参数、静态数据和生产数据等数据项.

2.2 汽窜时间特征选择

在多维时间序列数据的分析过程中，当样本存在冗余特征时会大大增加分析的复杂难度. 为保证深度学习模型的精度，需要针对各样本的基础参数进行特征选择. 本研究使用了基于标准互信息(normalized mutual information)的相似性度量方法[14]，其表达式为

$$U(X,Y) = \frac{-2\sum_{y\in Y}\sum_{x\in X} p(x,y)\,\mathrm{lb}\left[\dfrac{p(x,y)}{p(x)p(y)}\right]}{\sum_{x\in X} p(x)\,\mathrm{lb}\,p(x) + \sum_{y\in Y} p(y)\,\mathrm{lb}\,p(y)} \tag{12}$$

http://journal.szu.edu.cn

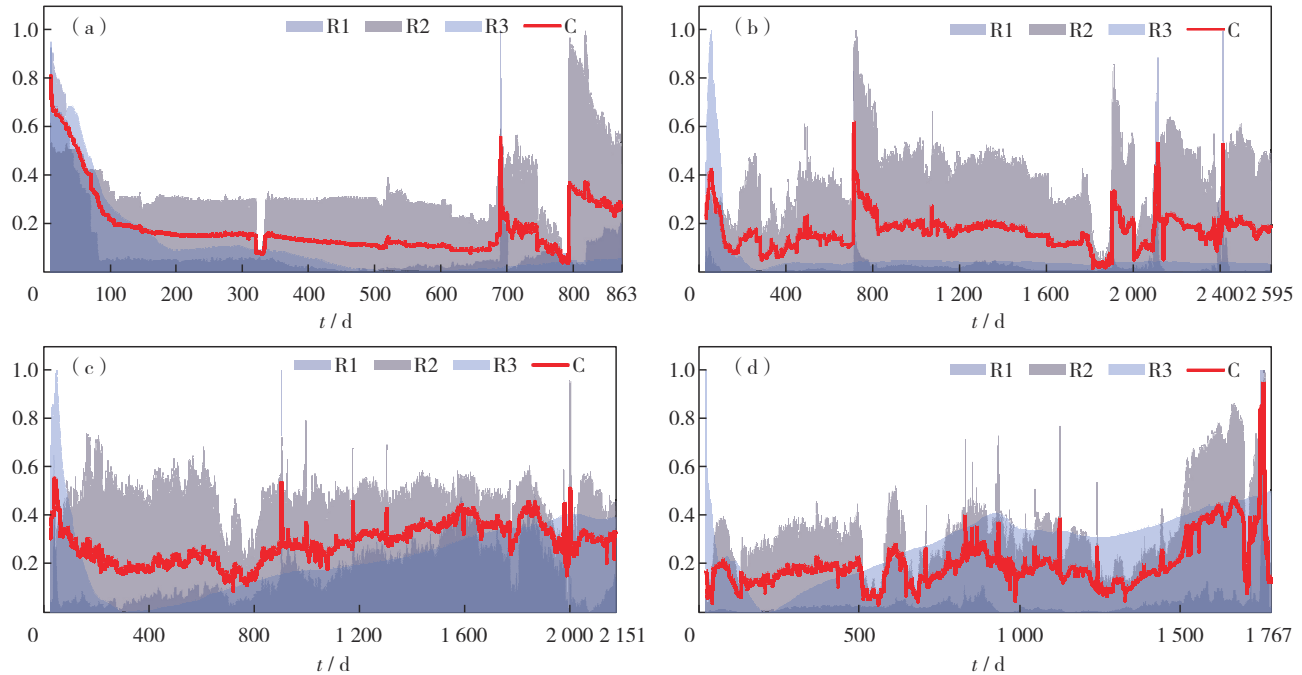


图 1　不同样本井的汽窜时间综合判识曲线　(a)J4 样本井；(b)J23 样本井；(c)J30 样本井；(d)J40 样本井

Fig. 1　Comprehensive identification curves of steam channeling time of (a) J4 sample well, (b) J23 sample well, (c) J30 sample well, and (d) J40 sample well.

其中，X 和 Y 分别为特征序列和目标序列；$p(x)$ 和 $p(y)$ 分别为特征值 x 和目标值 y 的边缘概率函数；$p(x,y)$ 为 x 和 y 的联合概率函数.

由式(12)定义可知，标准互信息的值越大，表明 Y 和 X 之间的相关性越强. 基于标准互信息的相似性度量方法，计算样本中每个注采参数与汽窜时间判识曲线之间的标准互信息值，并且进行排序，结果如图 2. 在该区块中，由于孔隙度、渗透率和有效厚度等储层参数呈现的变化幅度较小，相应的标准互信息值较小，表明其对汽窜时间判识曲线变化的影响较小. 从数据驱动的角度来看，在该实验区块下的储层参数对后续所建立的深度学习模型贡献不大，因此，选择标准互信息值在整体均值以上的 10 个注采参数作为预测汽窜时间的输入特征，即日产液量、日产油量、累积产油量、日产水量、累积产水量、生产井井口温度、注汽井井口温度、日注汽量、累积注汽量以及注采井距.

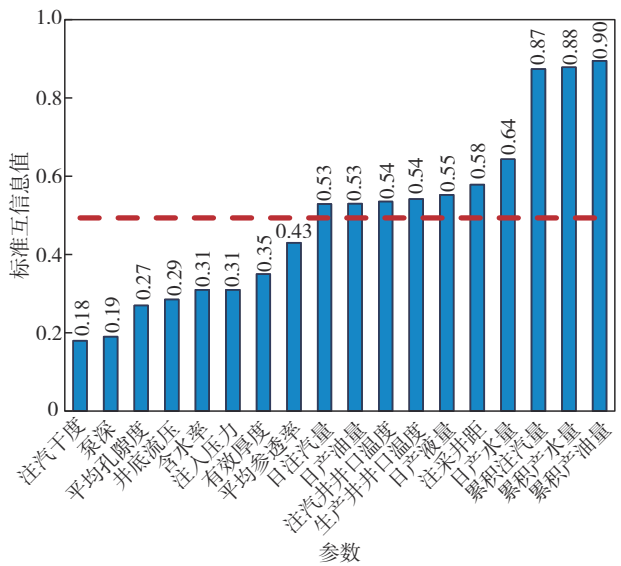


图 2　十个注采参数的标准互信息值

Fig. 2　Standard mutual information values of 10 injection - production parameters.

2.3　时间序列数据预处理

由于量纲不同，实际的注采动态数据在维度上通常表现出值域范围跨度大的特点，因此，需要对时间序列数据进行预处理，以提高后续模型的训练速度以及预测的准确率. 首先，对时间序列数据进行标准化处理，公式为

$$X_{\text{norm}} = \frac{X - \mu}{s_0} \tag{13}$$

其中，X_{norm} 为标准化处理后的数据；X 为原始数据；μ 为原始数据的均值；s_0 为原始数据的标准差.

然而，经过标准化后的数据在时间上还存在一定的噪声波动，因此，还需要运用高斯滤波器对时间序列进行降噪处理[15]，其表达式为

http://journal.szu.edu.cn

$$g(x)=\frac{1}{\sqrt{2\pi}\,s_G}e^{-x^2/(2s_G)} \tag{14}$$

其中，s_G为高斯分布的标准差．降噪处理只是去除了数据中不必要的峰值、趋势以及离群值，最大程度保留时间序列的结构信息，提高时间序列的辨识度．以J4井为例，经过标准化处理后的时间序列特征如图3(a)，运用高斯滤波器对这些时间序列数据进行降噪处理，结果如图3(b)．

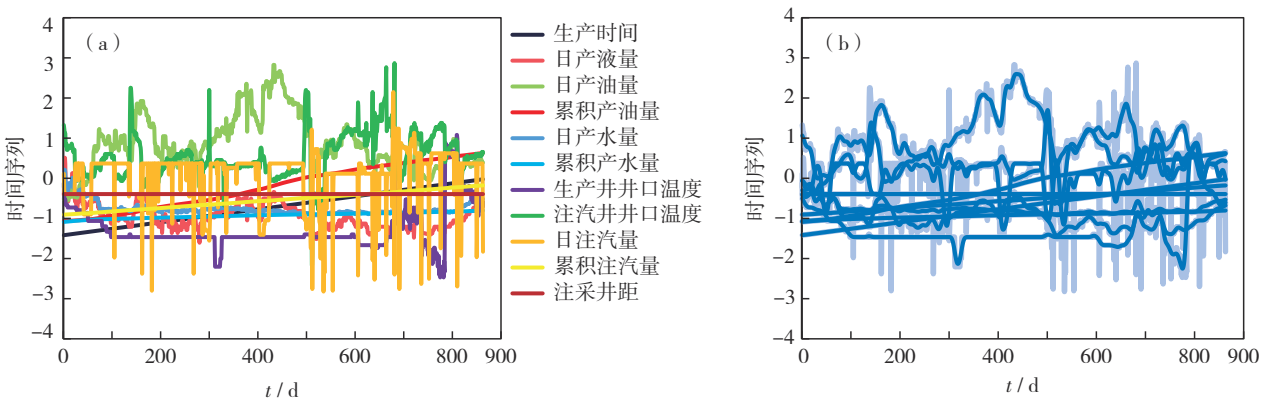


图3 J4样本井时间序列特征 (a)标准化；(b)降噪结果

Fig. 3 (a) Standardized and (b) smoothed time-series characteristic features of J4 sample well.

2.4 序列到序列深度学习框架

利用上述注采数据选择的特征作为输入，预测未来汽窜时间判识曲线的波动，以此确定生产井的汽窜时间．因此，选用序列到序列的深度学习框架(sequence-to-sequence，seq-to-seq)作为模型基础．序列到序列是一种通用的序列编码-解码深度学习框架，目前被广泛应用于自然语言处理(natural language processing，NLP)领域．与传统的循环神经网络(recurrent neural network，RNN)或LSTM只能处理单一维度且固定长度序列不同，序列到序列框架可以实现对多维度且任意长度序列之间的映射[16]．该框架的核心是由编码器和解码器组成的特殊网络架构，编码器将可变长度时间序列编码为一个固定长度的特征向量，然后由解码器将固定长度的特征向量进行解码作为预测输出，图4为序列到序列深度学习框架的基本结构．其中，S_N和h_N均为隐含层变量．由图4可见，输入的历史注采数据的时间序列以及对应的输出汽窜时间判识曲线时间序列的长度为n．此时构造的输入特征矩阵$\boldsymbol{I}$以及输出向量$\boldsymbol{O}$为

$$\boldsymbol{I}=\left[\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3,\cdots,\boldsymbol{X}_N\right] \tag{15}$$

$$\boldsymbol{O}=\left[\boldsymbol{Y}_1,\boldsymbol{Y}_2,\boldsymbol{Y}_3,\cdots,\boldsymbol{Y}_N\right] \tag{16}$$

$$\boldsymbol{X}_i=\begin{bmatrix}x_{11}^i & x_{12}^i & \cdots & x_{1n}^i\\ x_{21}^i & x_{22}^i & \cdots & x_{2n}^i\\ \vdots & \vdots & \ddots & \vdots\\ x_{m1}^i & x_{m2}^i & \cdots & x_{mn}^i\end{bmatrix} \tag{17}$$

$$\boldsymbol{Y}_i=\left[y_1^i,y_2^i,y_3^i,\cdots,y_n^i\right] \tag{18}$$

其中，N为样本总数($i=1,2,3,\cdots,N$)；X_i和Y_i分别为第i个样本的输入和输出；m为第i个样本中的特征维度；n为第i个样本中的时间序列长度．

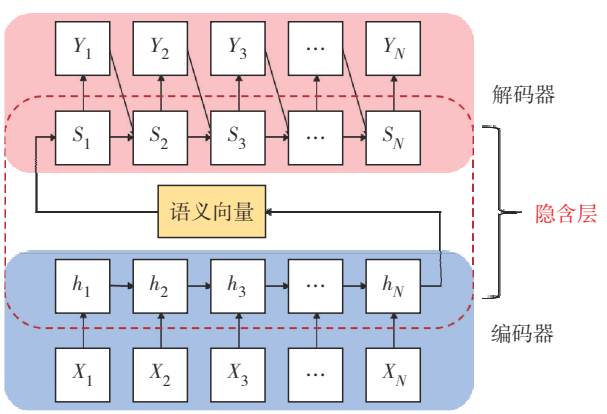


图4 序列到序列模型结构

Fig. 4 Model structure of seq-to-seq.

将构造好的输入特征矩阵和输出向量代入序列到序列结构中，随后将由编码器按照输入序列的时间顺序分步读入网络．编码器是由若干个LSTM门控制单元组成的，每个时刻的隐含状态h_t都由当前时刻的输入数据x_t与上一时刻的隐含状态h_{t-1}和单元状态c_{t-1}共同决定，其表达式为

$$h_t=f\left(x_t,h_{t-1},c_{t-1}\right) \tag{19}$$

其中，h_t和h_{t-1}分别为t和$t-1$时刻编码器门控制单元的隐含状态；x_t为t时刻的输入；c_{t-1}为$t-1$时刻编码器门控制单元的单元状态．

具体的门控制单元由候选门(g)、输入门(i)、

http://journal.szu.edu.cn

遗忘门(f)和输出门(o)4个功能模块组成．可学习的参数包括输入权重$\boldsymbol{W}$、循环权重$\boldsymbol{R}$以及偏置$\boldsymbol{b}$[17]．

2.5　模型设置

采用序列到序列深度学习结构建立汽窜时间预测模型．序列输入层支持输入不同长度的多维时间序列数据，每个样本都具有11个维度的时间序列特征，编码器和解码器隐含层则用于对输入序列进行训练学习，指定adam求解器来代替随机梯度下降作为网络的优化算法．最终，本研究以汽窜判识曲线拟合准确度为目标函数，使用优化算法得到最合适的超参数搭配，即200个隐含层，学习率为0.01，最小批分尺寸为16，迭代轮次为250次．

3　应用效果

3.1　效果评价指标

为评价模型的预测性能和精度，采用均方根误差(root mean square error，RMSE)和决定系数R^2两种指标定量评价模型的性能．

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}\left[y_{\text{true}}(i) - y_{\text{pred}}(i)\right]^2}{n}} \tag{20}$$

$$R^2 = 1 - \frac{S_{\text{res}}}{S_{\text{tot}}} \tag{21}$$

其中，$y_{\text{true}}(i)$为第i个样本的真实观测值；$y_{\text{pred}}(i)$为第i个样本的预测值；n为样本总数；S_{res}为样本真实观测值与预测值之间的残差平方和；S_{tot}为样本真实观测值与预测值之间的总离差平方和．

RMSE描述的是模型预测值和真实观测值之间的偏差，σ值越接近0，表示模型整体拟合效果越好．R^2描述的是预测值和真实观测值之间的变异程度，该值越大，说明预测值对真实观测值的解释程度越高，说明模型精度越高．

3.2　预测结果分析

将通过特征选择以及时间序列预处理的40个实验工区样本随机划分为80%的训练集和20%的测试集进行预测．测试样本数据中后10%时间步的汽窜综合判识曲线看作未知数据，运用前90%的历史注采时间序列数据输入模型进行预测，测试集整体的预测结果如图5．从图5可见，在该试验区快下运用序列到序列深度学习框架预测出的汽窜综合判识曲线与测试样本真实值之间的误差较小，并且随着时间序列长度的增加，预测误差没有显著的波动．说明序列到序列深度学习框架在此区块中具有一定的适用性，并且预测结果对测试样本的时间序列长度不敏感．

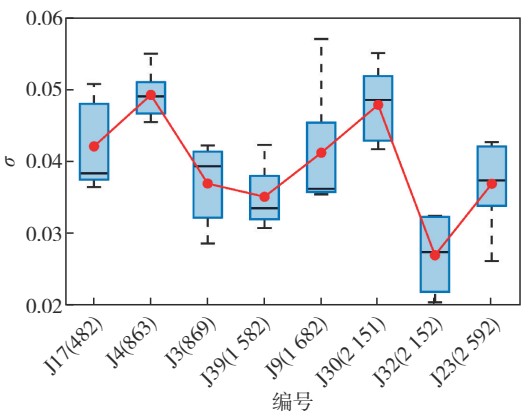


图5　测试集样本序列到序列深度学习框架预测结果(括号内数字为对应的序列长度)

Fig. 5　Prediction results of test set samples by seq-to-seq deep learning framework. (Values in parentheses are the length of the time series corresponding to the test samples.)

为验证所建立的汽窜时间预测模型的精度，以J4和J30样本为例，运用序列到序列深度学习框架预测的汽窜综合判识曲线结果分别为图6(绿色虚线左侧为训练集，右侧为测试集)，性能评价指标见表1．由表1可见，序列到序列深度学习框架在任意长度时间序列样本中均呈现了90%以上的R^2值，预测精度较高；其预测出的曲线幅度变化与实际综合判识曲线相同，提供了明显的汽窜时间拐点，并且与油田实际汽窜时间的绝对误差在10 d以内，符合油田现场汽窜预警的要求．

表1　性能评价指标对比

Table 1　Comparison of performance evaluation indexes

井号	模型	训练集		测试集		总数据集	
		σ	R^2	σ	R^2	σ	R^2
J4	序列到序列	0.014 6	0.988 8	0.034 8	0.866 7	0.017 7	0.982 6
	LSTM	0.014 3	0.982 2	0.098 2	0.625 3	0.034 1	0.931 2
J30	序列到序列	0.026 7	0.937 3	0.034 6	0.804 7	0.027 6	0.913 2
	LSTM	0.011 2	0.981 1	0.052 2	0.000 1	0.019 6	0.938 3

http://journal.szu.edu.cn

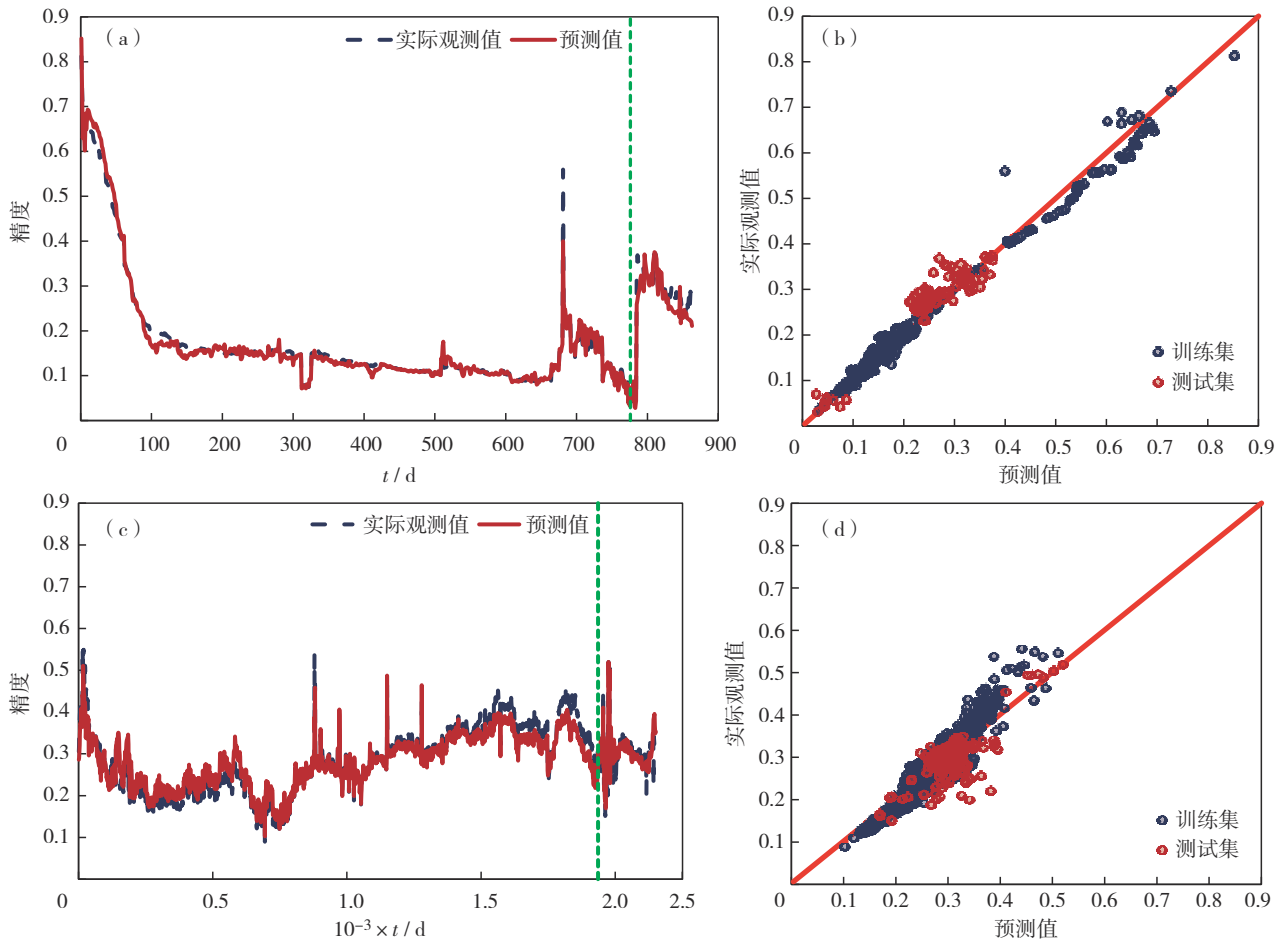


图6 J4和J30样本采用序列到序列深度学习框架的预测结果 (a)J4时间序列预测；(b)J4回归分析；(c)J30时间序列预测；(d)J30回归分析

Fig. 6 Prediction results of J4 and J30 samples by seq-to-seq deep learning framework. (a) Time series forecast plot of J4, (b) regression analysis plot of J4, (c) time series forecast plot of J30, and (d) regression analysis plot of J30. In the figure (a) and (c), the left side of the green dashed line represents the training set, and the right side represents the test set.

为说明所建立的汽窜时间预测模型在油田实际应用场景下的优越性，将所建立的模型与传统时间序列预测模型LSTM的预测结果进行了对比．运用传统的LSTM预测的结果分别为图7(绿色虚线左侧为训练集，右侧为测试集)．由图7可见，序列到序列深度学习结构在测试集中的精度高于传统的LSTM．这是因为传统的LSTM只能处理单一维度且固定长度的时间序列，在预测中不能准确反映实际汽窜时间判识曲线的真实波动．而序列到序列深度学习框架由于编码-解码的特殊结构，使其不受目标时间序列自身分布规律的限制，在预测过程中充分考虑了汽窜时间的多维影响因素，预测结果更符合油田实际情况．除此之外，在时间序列长度较长的样本J30中，传统的LSTM模型无法预测明确的拐点．这主要是因为传统的LSTM在处理长度较长的时间序列时存在梯度爆炸或者梯度消失的现象．而在序列到序列深度学习框架中，中间特征向量提取了整个输入序列的信息供解码器使用，使得该模型在处理长度较长的时间序列上也能保持很好的性能．

综合上述分析可见，本研究所建立的基于序列到序列深度学习框架的汽窜时间预测模型不仅可以准确的预测出汽窜综合判识曲线的拐点，而且在面对任意长度的时间序列数据依然能显示出较强的优越性，也符合油田实际现场应用的要求，为现场蒸汽驱汽窜预警提供了智能决策．

4 结 论

1）以生产动态时间序列数据为基础，提出一种汽窜时间表征方法．该方法通过组合动态数据构建了表征汽窜时间的指标参数，再引入变异系数-G1

http://journal.szu.edu.cn

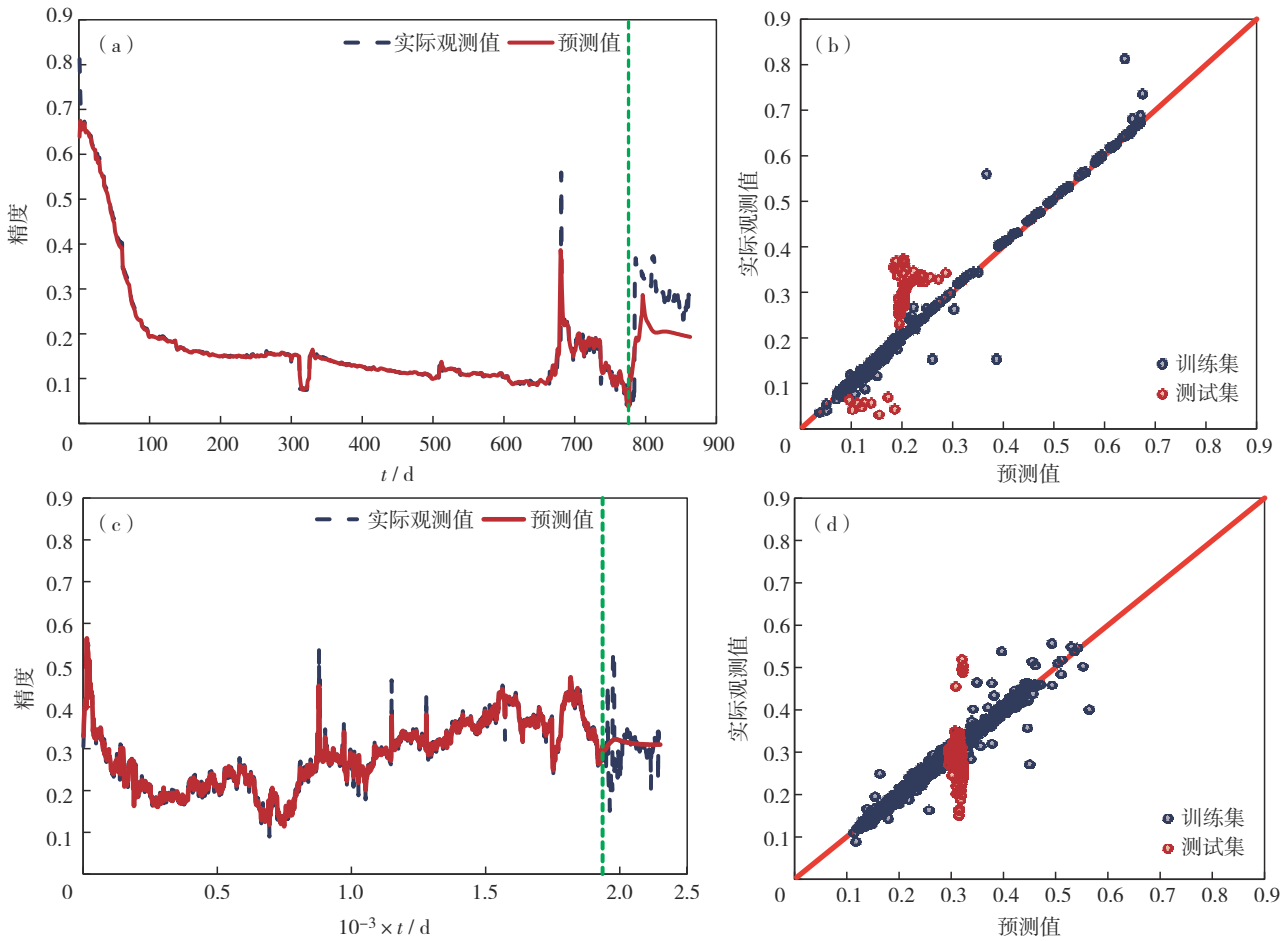


图 7　J4 和 J30 样本采用 LSTM 的预测结果　(a)J4 时间序列预测；(b)J4 回归分析；(c)J30 时间序列预测；(d)J30 回归分析

Fig. 7　Prediction results of J4 and J30 samples by LSTM. (a) Time series forecast plot of J4, (b) regression analysis plot of J4, (a) time series forecast plot of J30, and (b) regression analysis plot of J30. In the figure (a) and (c), the left side of the green dashed line represents the training set, and the right side represents the test set.

混合交叉赋权方法确定各表征指标的权重，从而得到综合的汽窜时间判别曲线．该曲线充分融合了温度变化、含水变化和汽窜突破程度等因素的影响，使其在时间序列上去除了多余的扰动，显现出更易于区分的突破拐点，这种陡然上升的拐点即为蒸汽窜流通道的形成初始时刻．

2）以构建的汽窜综合判识曲线作为输出目标时间序列，以标准互信息的相似性度量方法选择的注采参数作为预测汽窜时间的输入特征，采用序列到序列深度学习框架建立了基于数据驱动的汽窜时间预测模型，其预测结果与油田实际生产井汽窜时间相符合．通过与传统时间序列 LSTM 模型的对比，该模型在预测过程中不仅表现出更高的精度，而且其在任意长度的时间序列中都展现出了可观的性能．因此，针对同一区块相同地质条件的稠油井蒸汽驱汽窜时间预测场景来说，相比传统人工判识和数值模拟的方法，基于数据驱动的方式更适合现场应用，可为汽窜的智能预警工作提供新的科学依据．

基金项目：国家自然科学基金资助项目(51974343)；青岛市博士后应用研究资助项目(qdyy20200084)；中国石油大学(华东)自主创新科研计划资助项目(理工科)(20CX06089A)；中国石油大学(华东)石油工程学院 2021 年研究生创新工程资助项目(YCX2021047)

作者简介：崔传智(ccz2008@126.com)，中国石油大学(华东)教授、博士生导师．研究方向：油气渗流理论和油气田开发技术．

引　　文：崔传智，陆水青山，吴忠维，等．基于深度学习的稠油蒸汽驱汽窜时间预测方法[J]．深圳大学学报理工版，2023，40(5)：622-630.

参考文献 / References：

[1] 崔传智，郑文乾，祝仰文，等．蒸汽吞吐后转降黏化学驱加密井井位优化方法[J]．石油学报，2020，41(12)：1643-1648，1656.
CUI Chuanzhi, ZHENG Wenqian, ZHU Yangwen, et al. A method for optimizing the location of infill wells exploited by viscosity reduction chemical flooding after

http://journal.szu.edu.cn

steam huff and puff stimulation [J]. Acta Petrolei Sinica, 2020, 41(12): 1643-1648, 1656. (in Chinese)

[2] 王杰祥，刘晓，牛成成，等. 稠油油藏蒸汽吞吐防砂完井产能评价新方法[J]. 石油学报，2020，41(1)：80-87.
WANG Jiexiang, LIU Xiao, NIU Chengcheng, et al. A new productivity evaluation method of sand control completion for steam stimulation in heavy oil reservoirs [J]. Acta Petrolei Sinica, 2020, 41(1): 80-87. (in Chinese)

[3] WANG Peixi, ZHANG Jing. Application and design of fuzzy intelligent evaluation software for sand production and steam channeling prediction of steam injection well [J]. Procedia Engineering, 2011, 24: 546-550.

[4] WU Zhengbin, WANG Lei, XIE Congjiao, et al. Experimental investigation on improved heavy oil recovery by air assisted steam injection with 2D visualized models [J]. Fuel, 2019, 252: 109-115.

[5] 万慧清，张仲平，郝婷婷，等. 稠油油藏注蒸汽开发井间汽窜描述与影响因素[J]. 石油化工高等学校学报，2021，34(6)：57-63.
WAN Huiqing, ZHANG Zhongping, HAO Tingting, et al. The description and its influence factors of steam channeling between wells during steam injection in heavy oil reservoirs [J]. Journal of Petrochemical Universities, 2021, 34(6): 57-63. (in Chinese)

[6] 吴克柳，李相方，全一平，等. 稠油油藏蒸汽驱开发蒸汽突破预测模型[J]. 系统工程理论与实践，2014，34(2)：538-544.
WU Keliu, LI Xiangfang, QUAN Yiping, et al. Forecast model for steam breakthrough during steam drive exploration of inspissated pool [J]. Systems Engineering-Theory and Practice, 2014, 34(2): 538-544. (in Chinese)

[7] 梁成钢，李菊花，陈依伟，等. 基于朴素贝叶斯算法评估页岩油藏产能[J]. 深圳大学学报理工版，2023，40(1)：66-73.
LIANG Chenggang, LI Juhua, CHEN Yiwei, et al. Productivity evaluation based on naive Bayesian algorithm in shale reservoir [J]. Journal of Shenzhen University Science and Engineering, 2023, 40(1): 66-73. (in Chinese)

[8] 陆吉，林伯韬，史璨，等. 克拉玛依油田七区砂砾岩油藏智能岩性识别[J]. 深圳大学学报理工版，2023，40(3)：361-369.
LU Ji, LIN Botao, SHI Can, et al. Intelligent lithology identification of sandy conglomerate reservoir in District No. 7, Karamay oilfield [J]. Journal of Shenzhen University Science and Engineering, 2023, 40(3): 361-369. (in Chinese)

[9] 朱璇，袁彬，同元辉，等. 致密低渗油藏压裂井网气驱深度学习预测模型[J]. 深圳大学学报理工版，2022，39(5)：559-566.
ZHU Xuan, YUAN Bin, TONG Yuanhui, et al. Deep-learning-based proxy model for forecasting gas flooding performance of fractured well pattern in tight oil reservoirs [J]. Journal of Shenzhen University Science and Engineering, 2022, 39(5): 559-566. (in Chinese)

[10] SYED F I, MUTHER T, DAHAGHI A K, et al. CO_2 EOR performance evaluation in an unconventional reservoir through mechanistic constrained proxy modeling [J]. Fuel, 2022, 310, Part A: 122390.

[11] 郑强，刘慧卿，李芳，等. 蒸汽驱后汽窜通道定量描述[J]. 中国科学技术科学，2013，43(6)：684-688.
ZHENG Qiang, LIU Huiqing, LI Fang, et al. Quantitative description of steam channeling channel after steam flooding [J]. Scientia Sinica Technologica, 2013, 43(6): 684-688. (in Chinese)

[12] 李志军，向建军，盛涛，等. 基于G1-变异系数-KL改进TOPSIS雷达对抗干扰有效性评估[J]. 北京航空航天大学学报，2021，47(12)：2571-2578.
LI Zhijun, XIANG Jianjun, SHENG Tao, et al. G1-variation-coefficient-KL based TOPSIS radar jamming effectiveness evaluation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(12): 2571-2578. (in Chinese)

[13] 赵圆元，张行南，黄钰凯，等. 基于动态权重的水文时间序列相似性分析方法[J]. 水电能源科学，2021，39(2)：9-12，17.
ZHAO Yuanyuan, ZHANG Xingnan, HUANG Yukai, et al. Similarity analysis method of hydrological time series based on dynamic weight [J]. Water Resources and Power, 2021, 39(2): 9-12, 17. (in Chinese)

[14] 齐方鹏，王童，傅忠谦. 基于互信息的复杂网络链路预测[J]. 中国科学技术大学学报，2020，50(1)：57-63.
QI Fangpeng, WANG Tong, FU Zhongqian. Link prediction in complex networks based on mutual information [J]. Journal of University of Science and Technology of China, 2020, 50(1): 57-63. (in Chinese)

[15] MARTINO L, READ J. A joint introduction to gaussian processes and relevance vector machines with connections to Kalman filtering and other kernel smoothers [J]. Information Fusion, 2021, 74: 17-38.

[16] 俞佳佳，金赟，马勇，等. 基于Sinc-Transformer模型的原始语音情感识别[J]. 信号处理，2021，37(10)：1880-1888.
YU Jiajia, JIN Yun, MA Yong, et al. Emotion recognition from raw speech based on Sinc-Transformer model [J]. Journal of Signal Processing, 2021, 37(10): 1880-1888. (in Chinese)

[17] 刘艳，张婷，康爱卿，等. Seq2Seq模型的短期水位预测[J]. 水利水电科技进展，2022，42(3)：57-63.
LIU Yan, ZHANG Ting, KANG Aiqing, et al. Short-term water level prediction based on Seq2Seq model [J]. Advances in Science and Technology of Water Resources, 2022, 42(3): 57-63. (in Chinese)

【中文责编：晨兮；英文责编：天澜】

http://journal.szu.edu.cn