



DOI:10.12404/j.issn.1671-1815.2402520

引用格式:姚秀田,盖平原,张兆祥,等.稠油蒸汽驱见汽时间表征指标及预测方法[J].科学技术与工程,2025,25(8):3181-3189.

Yao Xiutian, Gai Pingyuan, Zhang Zhaoxiang, et al. Characterization indicators and prediction methods of steam breakthrough time in heavy oil steam flooding[J]. Science Technology and Engineering, 2025, 25(8): 3181-3189.

石油、天然气工业

## 稠油蒸汽驱见汽时间表征指标及预测方法

姚秀田<sup>1</sup>, 盖平原<sup>2,3</sup>, 张兆祥<sup>2,3</sup>, 郝婷婷<sup>2,3</sup>, 佟彤<sup>2,3</sup>, 张仲平<sup>2,3</sup>

(1. 胜利油田分公司孤岛采油厂, 东营 257000; 2. 胜利油田分公司石油工程技术研究院, 东营 257000;

3. 山东省稠油开采技术省级重点实验室, 东营 257000)

**摘要** 稠油作为全球能源结构中的关键非常规油气资源,其开发效率提升备受关注。蒸汽驱技术虽然提高了稠油开采效率,但汽窜现象对热效率和油藏开发产生不利影响。传统的预测方法在处理油田长时间序列时存在精度不足和响应迟缓的问题。基于胜利油田稠油区块13个蒸汽驱井组的数据,采用生产井与注汽井间瞬时温度比作为见汽时间的表征方法,并通过皮尔逊相关系数分析,精选出与汽窜时间相关的关键因素。在此基础上,构建基于Transformer架构的深度学习模型,成功实现了瞬时温度比的精确预测,并与实地观测数据高度吻合,展现了比传统长短期记忆网络(long short-term memory, LSTM) LSTM模型更高的预测精度与稳定性。研究成果不仅为稠油油藏汽窜时间的精确预测提供了新的视角,也进一步验证了深度学习技术在油田开发领域应用的广泛潜力,为油田智能化管理和决策支持系统的构建提供支撑。

**关键词** 稠油油藏; 蒸汽驱; 见汽时间预测; 自注意力机制; Transformer; 深度学习

中图分类号 TE34;

文献标志码 A

## Characterization Indicators and Prediction Methods of Steam Breakthrough Time in Heavy Oil Steam Flooding

YAO Xiu-tian<sup>1</sup>, GAI Ping-yuan<sup>2,3</sup>, ZHANG Zhao-xiang<sup>2,3</sup>, HAO Ting-ting<sup>2,3</sup>,

TONG Tong<sup>2,3</sup>, ZHANG Zhong-ping<sup>2,3</sup>

(1. Gudao Oil Production Plant of Shengli Oilfield, Dongying 257000, China; 2. Petroleum Engineering Technology

Research Institute of Shengli Oilfield, Dongying 257000, China; 3. Key Laboratory of Heavy Oil Recovery

Technology of Shandong Province, Dongying 257000, China)

**[Abstract]** As a critical unconventional oil and gas resource within the global energy framework, heavy oil has garnered significant attention for its development efficiency. Although steam flooding technology has improved the efficiency of heavy oil production, the phenomenon of steam breakthrough negatively impacts thermal efficiency and reservoir development. Traditional prediction methods have shown inadequate precision and delayed response when dealing with long-term oilfield time series data. Data from 13 steam flooding well groups in the Shengli oilfield heavy oil block were utilized. An innovative approach was adopted, using the instantaneous temperature ratio between production and injection wells as an indicator of steam breakthrough time. Pearson correlation coefficient analysis was employed to select key factors related to steam breakthrough time. Based on these factors, a deep learning model built on the Transformer architecture was developed, achieving accurate predictions of the instantaneous temperature ratio. The predictions closely aligned with oilfield observation data, demonstrating higher prediction accuracy and stability compared to traditional long short-term memory(LSTM) models. The research results not only provide a new perspective for the precise prediction of steam breakthrough time in heavy oil reservoirs but also further validate the extensive potential of deep learning technology in oilfield development applications, supporting the construction of intelligent oilfield management and decision support systems.

**[Keywords]** heavy oil reservoir; steam drive; steam emergence time prediction; self-attention mechanism; Transformer; deep learning

收稿日期: 2024-04-08; 修订日期: 2024-12-17

基金项目: 国家科技重大专项(2017ZX05072-006-004)

第一作者: 姚秀田(1978—),男,汉族,山东五莲人,博士,高级工程师。研究方向:油气田开发理论及技术。E-mail: yaoxiutian\_slyt@sinopec.com。

投稿网址: www.stae.com.cn

在全球能源结构中,稠油作为一种关键的非常规油气资源,对于满足不断增长的能源需求发挥着至关重要的作用。蒸汽驱技术,作为一种高效的稠油开采方法,虽然在提高油田开发效率方面发挥了显著作用,但开发过程中不可避免的汽窜现象,形成高渗透通道,会显著降低热效率并对油藏的开发效果产生负面影响<sup>[1-3]</sup>。因此,精确预测见汽时间,成为优化开发策略、提升稠油油藏开发效率的关键。然而,传统预测方法,如数值模拟,难以满足油田快速响应的需求。解析法虽计算速度较快,但由于简化了油藏的复杂条件,其在实际应用中的准确性和可靠性受到限制<sup>[4-6]</sup>。动态分析法因其能够便捷、低成本地获取现场数据并有效反映注采井间的窜流关系受到关注,但在处理蒸汽驱开发中的复杂因素时,仍显示出传统数学方法的局限,难以适应当前大数据环境下的复杂油田应用需求<sup>[7-8]</sup>。随着油田开发进程的加深,大量积累的数据资产逐渐转化为探索新规律的重要资源。智能油田领域因此紧密聚焦于此领域的研究。随着大数据和人工智能技术的演进,传统油田开发问题逐步依赖这些前沿技术寻找解决路径。

近年来,深度学习方法广泛应用于解决油田开发中的时间序列预测问题。在产量预测方面,刘合等<sup>[9]</sup>通过结合极限梯度提升算法(extreme gradient boosting, XGBoost)筛选关键特征,提出了一种解决油田高含水期产量预测的长短时记忆模型(long short-term memory, LSTM),有效提高了高含水期产量预测的准确性。Zhang等<sup>[10]</sup>通过建立基于卷积神经网络(convolutional neural networks, CNN)和LSTM的混合学习模型,成功实现了水驱油田中单井月度产量的精准预测。Kumar等<sup>[11]</sup>通过引入注意力机制提升了LSTM的预测性能,结果表明,带注意力机制的LSTM相对于传统的深度学习模型在石油产量预测方面表现更为出色。这些研究成果展示了深度学习在油田时间序列预测领域的显著优势,并为稠油蒸汽驱汽窜时间预测开辟了新的研究路径,突出了数据驱动预测方法的潜力。崔传智等<sup>[12]</sup>提出的基于序列到序列(seq-to-seq)的汽窜时间预警模型,通过实际预测案例并与传统机器学习方法的比较,证实了该模型的有效性和优越性。尽管该模型采用了序列到序列结构并集成了LSTM作为其核心模块,但在处理长期数据序列时,LSTM模块的固有缺陷依然存在。这提示着未来研究中需进一步探索改进深度学习模型以弥补这一不足。

综上所述,前人研究在汽窜时间预警模型的开发方面取得了进展,但对长期时间序列的处理却仍

显不足,指出深度学习模型在预测效率和精度上的待改进空间。鉴于稠油油藏开发过程中对汽窜现象精准预测的紧迫需求,现聚焦于胜利油田稠油区块的13个蒸汽驱井组,提出新路径来探究见汽时间的表征及预测技术。首次利用生产井与注汽井间瞬时温度比例,作为判定蒸汽驱汽窜时间的新颖指标,并通过皮尔逊相关系数分析,从多个潜在因素中筛选出与之紧密相关的关键变量,如日产量、累积产量、日注汽量和累积注汽量,以及重要的地质参数等。基于这些精选特征,构建针对深度学习的样本库,并进一步开发一种基于Transformer架构的预测模型。实证分析表明,该模型在预测精度与稳定性方面相比传统的LSTM序列到序列模型显示出明显的优势,与实地观测数据高度吻合。为稠油油藏中汽窜时间的精确预测提供新方法,也进一步拓宽深度学习技术在油田开发领域的应用范围。

## 1 基于Transformer的时间序列预测模型原理及方法

Transformer是一种以多头自注意力机制(multi-head self-attention)为基础的序列到序列模型,主要应用于机器翻译、语音识别、对话生成等领域<sup>[13]</sup>。相较于传统的循环神经网络,Transformer通过引入位置编码和自注意力机制,实现对所有位置信息的同时处理,更有效地捕捉全局依赖性<sup>[14]</sup>。其模型结构,如图1所示,包括编码器、解码器、多头自注意力机制、前馈神经网络以及位置编码。

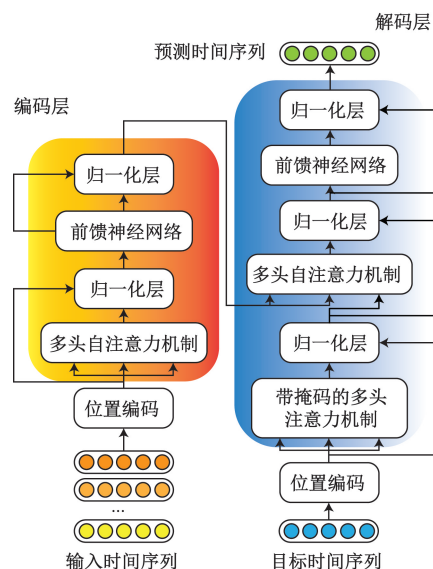


图1 Transformer模型结构

Fig. 1 Transformer model structure

### 1.1 编码器结构

Transformer模型的编码器通过正弦和余弦函数

对输入序列中的每个位置进行编码,确保模型能够识别位置信息<sup>[15]</sup>,如式(1)~式(3)所示。

假设一个输入的时间序列  $\mathbf{x}_{1:T} \in \mathbf{R}^{D \times T}$ ,其位置编码后的输入向量为

$$\mathbf{H}^{(0)} = [\mathbf{x}_1 + \mathbf{p}_1, \mathbf{x}_2 + \mathbf{p}_2, \dots, \mathbf{x}_T + \mathbf{p}_T] \quad (1)$$

$$p_{t,2i} = \sin\left(\frac{t}{10000^{\frac{2i}{D}}}\right) \quad (2)$$

$$p_{t,2i+1} = \cos\left(\frac{t}{10000^{\frac{2i}{D}}}\right) \quad (3)$$

式中: $\mathbf{H}^{(0)}$ 为经过位置编码后的初始隐含状态; $\mathbf{p}_t$ 为位置  $t$  的向量,即位置编码; $p_{t,2i}$ 为位置为  $t$  的第  $2i$  个维度; $D$ 为位置编码的维度。

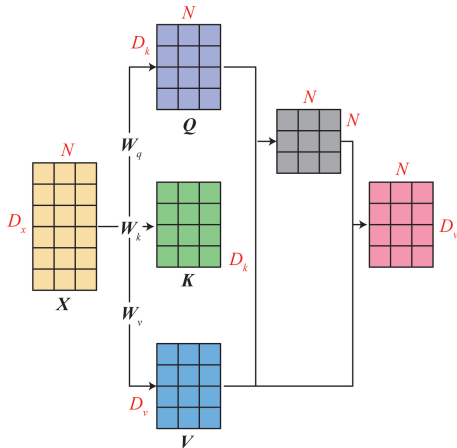
多头自注意力机制,作为编码器的核心部分,专注于捕获输入序列内各位置间的相互依赖<sup>[16]</sup>。自注意力的计算过程如图2所示。

$$\text{self-att}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathbf{V} \text{softmax}\left(\frac{\mathbf{K}^T \mathbf{Q}}{\sqrt{D_k}}\right) \quad (4)$$

$$\mathbf{Q} = \mathbf{W}_q \mathbf{H}, \mathbf{K} = \mathbf{W}_k \mathbf{H}, \mathbf{V} = \mathbf{W}_v \mathbf{H} \quad (5)$$

式中:slfe-att( $\cdot$ )为自注意力的计算过程,用于衡量输入序列中每个位置与其他位置之间的相关性,并通过权重分配整合关键信息; $D_k$ 为输入矩阵  $\mathbf{Q}$  和  $\mathbf{K}$  的维度; $\mathbf{H} \in \mathbf{R}^{N \times D}$ 为模型中当前层的隐含状态矩阵,其中, $N$ 为序列长度; $D$ 为特征维度; $\mathbf{W}_q \in \mathbf{R}^{D_k \times D_h}$ 、 $\mathbf{W}_k \in \mathbf{R}^{D_k \times D_h}$ 、 $\mathbf{W}_v \in \mathbf{R}^{D_v \times D_h}$ 为3个投影矩阵,即用于线性变换的权重矩阵,是模型需要学习的参数,其中, $D_h$ 为多头注意力机制的分头数; $D_v$ 为前馈神经网络隐藏层维度, $D_v = 4D_k$ ,设置  $D_k = 512$ ,8个多头注意力, $D_h = 512/8 = 64$ , $D_v = 4 \times 512 = 2048$ ;矩阵  $\mathbf{Q}$ 为当前信息的索引位;矩阵  $\mathbf{K}$ 为其他位置的输入信息;矩阵  $\mathbf{V}$ 为是当前位置的输入信息<sup>[17]</sup>。

自注意力机制的核心在于识别并加权输入信



$D_x$ 为位置编码后输入特征的维度; $N$ 为位置编码后输入序列的长度; $X$ 为在注意力机制里运算时输入的矩阵

图2 注意力机制

Fig. 2 Attention mechanism

息中各位置间的相关性,为当前位置汇总关键信息。多头自注意力机制进一步扩展此概念,在多个投影空间内并行计算注意力,以捕获信息的多维关系,丰富模型对输入数据的理解与处理<sup>[18]</sup>。

假设在  $M$  个投影空间中分别应用自注意力,其表达式为

$$\text{MultiHead}(\mathbf{H}) = \mathbf{W}_o [\text{head}_1; \text{head}_2; \dots; \text{head}_M] \quad (6)$$

式(6)中: $\mathbf{W}_o \in \mathbf{R}^{D_h \times M D_e}$ 为输出投影矩阵,该矩阵也为模型需要学习的权重矩阵; $m \in \{1, 2, \dots, M\}$ ;Multi-Head( $\cdot$ )为多头注意力机制; $\text{head}_m$ 为第  $m$  个头的注意力输出,是一个独立计算的注意力结果,定义为

$$\text{head}_m = \text{self-att}(\mathbf{Q}_m, \mathbf{K}_m, \mathbf{V}_m) \quad (7)$$

$$\mathbf{Q}_m = \mathbf{W}_q^m \mathbf{H}, \mathbf{K}_m = \mathbf{W}_k^m \mathbf{H}, \mathbf{V}_m = \mathbf{W}_v^m \mathbf{H} \quad (8)$$

式中: $\mathbf{Q}_m$ 、 $\mathbf{K}_m$ 、 $\mathbf{V}_m$ 分别为第  $m$  个头的索引位矩阵、其他位置的输入信息、当前位置的输入信息; $\mathbf{W}_q^m$ 为  $\mathbf{Q}_m$  的投影矩阵; $\mathbf{W}_k^m$ 为  $\mathbf{K}_m$  的投影矩阵; $\mathbf{W}_v^m$ 为  $\mathbf{V}_m$  的投影矩阵。

在信息传递过程中,假设  $\mathbf{H}^{(l-1)}$ 为第  $l-1$  层的隐含状态,则第  $l$  层隐含状态  $\mathbf{H}^{(l)}$  的计算公式为

$$\mathbf{Z}^{(l)} = \text{norm}\{\mathbf{H}^{(l-1)} + \text{Multihead}[\mathbf{H}^{(l-1)}]\} \quad (9)$$

$$\mathbf{H}^{(l)} = \text{norm}\{\mathbf{Z}^{(l)} + \text{FFN}[\mathbf{Z}^{(l)}]\} \quad (10)$$

式中:norm( $\cdot$ )表示归一化;FFN( $\cdot$ )为逐位置的前馈神经网络(position-wise feed-forward network); $\mathbf{Z}^{(l)}$ 为第  $l$  层所有位置的隐含状态集合,而对于输入序列中每个位置上向量  $\mathbf{z}$   $\mathbf{Z}^{(l)}$ ,则有

$$\text{FFN}(\mathbf{z}) = \mathbf{W}_2 \text{ReLU}(\mathbf{W}_1 \mathbf{z} + \mathbf{b}_1) + \mathbf{b}_2 \quad (11)$$

式(11)中: $\mathbf{W}_1$ 、 $\mathbf{W}_2$ 、 $\mathbf{b}_1$ 、 $\mathbf{b}_2$ 为FFN的网络参数<sup>[19-21]</sup>。

## 1.2 解码器结构

Transformer模型中的解码器与编码器结构相似,包含位置编码、多头自注意力机制、残差连接和层归一化等组件,以处理目标序列的编码。解码器增设了编码器-解码器注意力机制(encoder-decoder attention),以提供与当前目标序列生成步骤最相关的信息,从而在生成序列的每一步上,增强对输入序列关键部分的关注。此外,解码器还利用了掩码矩阵来优化序列生成过程<sup>[22]</sup>。掩码矩阵的设计确保在生成当前位置的标记时,解码器仅依赖已生成的标记,而无法访问未来位置的信息。这有助于确保生成的序列按顺序逐步完成,提高了生成序列的可靠性和合理性<sup>[23]</sup>。

## 2 基于Transformer的蒸汽驱见汽时间预测模型建立

### 2.1 实验数据选择

选择胜利油田稠油区块的实际资料作为研究

数据。本区块特征包括原始地层压力 11.5 MPa,平均渗透率  $2\ 786 \times 10^{-3} \mu\text{m}^2$ ,平均孔隙度 33%,初期含油饱和度 64%,油藏埋深 1 080 ~ 1 150 m,油层厚度范围为 6 ~ 24 m;而地面脱气原油黏度一般介于 3 390 ~ 420 000 mPa·s。由于井网密集度提高、蒸汽吞吐周期延长等因素,该区块出现了汽窜现象,随着开发进度的推进,汽窜情况日益严重,对稠油油藏的开发造成了显著的负面影响。选取该区块内 13 个蒸汽驱井组的日度时间序列数据作为对象,以注汽井 II 和生产井 P1 的数据为例,如表 1、表 2 所示。

2.2 汽窜特征选择以及数据样本集构建

在蒸汽驱开发过程中,汽窜现象的监测是一个关键问题。注入的高温气体直接流向生产井时,会引起生产井温度的显著上升,从而在温度曲线上形成拐点<sup>[24]</sup>。这一现象可以作为判断汽窜发生的一个重要依据。然而,由于地下环境的复杂性,油田实际应用中生产井的井口温度曲线受多种因素影响,波动较大,无法呈现明显的突变特征。因此,引入生产井与注汽井之间的瞬时温度之比来表征蒸汽驱汽窜的见汽时间,如式(12)所示。生产井的井口温度与注汽井的井口温度之比  $C$ ,通过无因次的

角度综合了生产井与注入井之间的温度变化, $C$  越大为注汽井与生产井之间的温差越大,蒸汽正在向生产井方向移动,发生汽窜的可能性越大。

$$C = \frac{T_p}{T_i} \tag{12}$$

式(12)中: $T_p$  为生产井井口的瞬时温度; $T_i$  为注汽井井口的瞬时温度。

图 3 展示了 P1、P2、P3 井瞬时温度之比  $C$  的变化曲线。以比值等于 1 作为基准,当  $C$  大于 1 时,表明注汽井到生产井的温度梯度发生逆转,生产井温度瞬时高于注入井,蒸汽直接到达生产井,未能有效加热油藏,发生汽窜。据此,通过这种无因次消除

表 1 注汽井基础数据  
Table 1 Injection well basic data

井号	日期	出口温度/ ℃	注汽干度	日注汽量/ ( $\text{m}^3 \cdot \text{d}^{-1}$ )
II	2013-07-18	308	75	126
II	2013-07-19	305	77	144
II	2013-07-20	304	76	144
II	2013-07-21	308	77	144
II	2013-07-22	309	76	144
II	...	...	...	...
II	2020-12-14	318	80	156

表 2 生产井基础数据  
Table 2 Production well basic data

井号	日期	日产液/ ( $\text{m}^3 \cdot \text{d}^{-1}$ )	日产油/ ( $\text{m}^3 \cdot \text{d}^{-1}$ )	温度/℃	泵深/m	渗透率/ ( $10^{-3} \mu\text{m}^2$ )	孔隙度	厚度/m	井距/m
P1	2013-02-01	55	3.8	68	947.39	3 000	0.3	10	150
P1	2013-02-02	51	3.4	62	947.39	3 000	0.3	10	150
P1	2013-02-03	51	3.1	68	947.39	3 000	0.3	10	150
P1	2013-02-04	50	3.0	68	947.39	3 000	0.3	10	150
P1	2013-02-05	48	2.7	66	947.39	3 000	0.3	10	150
P1	...	...	...	...	...	...	...	...	...
P1	2016-06-08	54	3.3	65	947.39	3 000	0.3	10	150

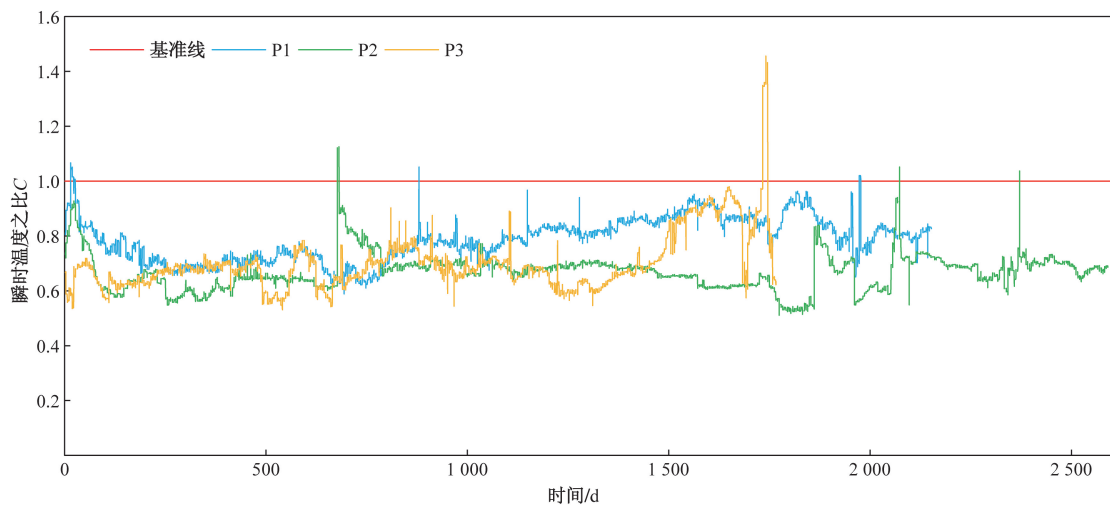


图 3 瞬时温度之比曲线

Fig. 3 Instantaneous temperature ratio curve

冗余的方法,可以清楚的区分出来 P1 井在第 21 天、第 897 天、第 1 986 天发生汽窜;P2 井在第 616 天、第 2 156 天、第 2 400 天出现汽窜;P3 井在第 1 604 天检测到汽窜。

在多维时间序列建模中,特征选择对于增强模型精确性与减少计算复杂度至关重要。通过筛选与目标高度相关、低冗余的特征,可以有效提升训练效率与模型的泛化能力。利用皮尔逊相关系数分析各影响因素与瞬时温度比值  $C$  的相关度,筛选出绝对值相关性强的因素,如图 4 所示。其中包括日产量数据、累积产量数据、日注入量数据、累积注入量数据以及静态地质参数,如油层厚度、渗透率,孔隙度以及注采井距,示例如表 3 所示。以筛选后的特征作为模型输入,结合瞬时温度比值变化曲线作为预测目标建立深度学习样本库,为进一步的时间序列建模奠定基础。

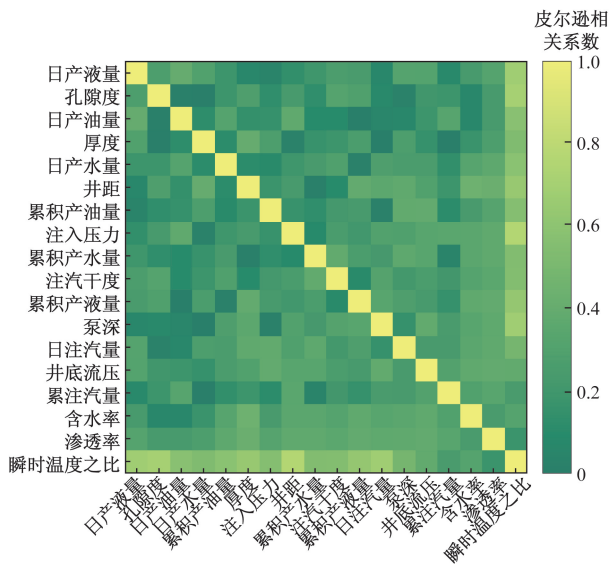


图 4 皮尔逊相关性系数热图

Fig. 4 Pearson correlation coefficient heatmap

### 2.3 数据标准化以及平滑

在建立样本库之后,由于实际动态数据在维度上因量纲不同而表现出较大的值域波动,对时间序列数据进行标准化至关重要。这不仅可以加快深度学习模型的训练速度,还能提升预测准确性。采用 Z-score 标准化方法,对数据进行标准化处理,该标准化方法可以保留原始数据以及后续更新样本库新数据的变化趋势<sup>[25]</sup>,其转化函数为

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

式(13)中: $X_{norm}$ 为标准化后的数据; $X$ 为原始数据; $\mu$ 为原始数据的均值; $\sigma$ 为原始数据的标准差。

标准化处理后,数据维度在同一分布特征下变化,保持了时间序列的结构特性,其结果如图 5 所示。尽管如此,标准化后的数据可能因为不均匀分布而表现出时间上的噪声波动。为此,采用小波降噪法进行数据平滑处理是有效的策略。该方法通过小波变换对时间序列进行分解,实现噪声与信号的分离,进而通过阈值处理提纯信号,并通过小波重构精炼最终信号<sup>[26]</sup>。此过程不仅消除了异常峰值,而且尽可能保留了时间序列的核心结构,增强了数据的清晰度,其结果如图 6 所示。

### 2.4 模型建立及超参数优化

在完成样本准备、数据标准化与平滑处理后,将处理后的样本井分为 80% 的训练集和 20% 的测试集,用于构建基于 Transformer 的蒸汽驱见汽时间预测模型。接下来,采用平均绝对百分比误差 (mean absolute percentage error, MAPE) 和均方根误差 (root mean square error, RMSE) 作为评估模型性能指标<sup>[27]</sup>,进行模型训练和超参数优化,以确定最优的超参数配置,确保模型达到最佳性能,评价指标计算公式为

表 3 样本输入特征

Table 3 Sample input feature data

时间/d	渗透率/ $10^{-3} \mu\text{m}^2$	孔隙度	厚度/ m	井距/ m	日产量/ $(\text{m}^3 \cdot \text{d}^{-1})$	日产油量/ $(\text{m}^3 \cdot \text{d}^{-1})$	日产水量/ $(\text{m}^3 \cdot \text{d}^{-1})$	累积产 油量/ $\text{m}^3$	累积产 水量/ $\text{m}^3$	日注汽量/ $(\text{m}^3 \cdot \text{d}^{-1})$	累积注 汽量/ $\text{m}^3$
1	2 451	0.4	8.5	170	39	15.0	24.0	15.0	24.0	126	126
2	2 451	0.4	8.5	170	38	14.8	23.2	29.8	47.2	144	270
3	2 451	0.4	8.5	170	38	14.8	23.2	44.6	70.4	144	414
4	2 451	0.4	8.5	170	38	14.8	23.2	59.4	93.6	144	558
5	2 451	0.4	8.5	170	40	15.4	24.6	74.8	118.2	144	702
6	2 451	0.4	8.5	170	40	15.4	24.6	90.2	142.8	144	846
7	2 451	0.4	8.5	170	40	15.4	24.6	105.6	167.4	144	990
8	2 451	0.4	8.5	170	40	15.7	24.3	121.3	191.7	144	1 134
9	2 451	0.4	8.5	170	40	15.7	24.3	137.0	216.0	144	1 278
10	2 451	0.4	8.5	170	40	15.7	24.3	152.7	240.3	144	1 422
...	...	...	...	...	...	...	...	...	...	...	...
987	2 451	0.4	8.5	170	29	9.4	19.6	14 869.8	27 424.0	156	141 819

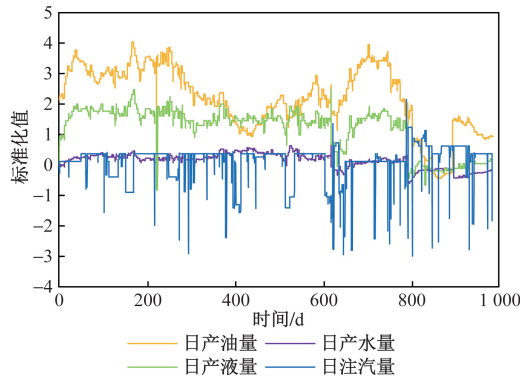


图5 时间序列数据标准化结果

Fig. 5 Standardization results of time series data

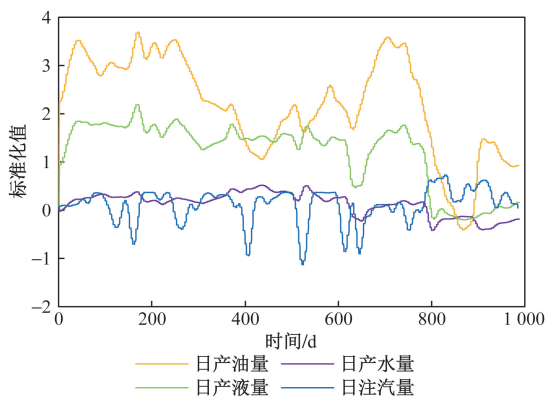


图6 时间序列数据平滑结果

Fig. 6 Time series data smoothing results

$$Y_{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_{true}(i) - Y_{pre}(i)}{Y_{true}(i)} \right| \times 100\% \quad (14)$$

$$Y_{RMSE} = \sqrt{\frac{\sum_{i=1}^n [Y_{true}(i) - Y_{pre}(i)]^2}{n}} \quad (15)$$

式中:  $Y_{true}(i)$  为  $i$  时刻的温度真实值;  $Y_{pre}(i)$  分别  $i$  时刻温度的预测值。

平均绝对百分比误差  $Y_{MAPE}$  衡量预测值与真实值之差相对于真实值的比例,  $MAPE$  越低, 模型预测性能越好。均方根误差  $Y_{RMSE}$  反映预测值与真实值之间的标准偏差,  $RMSE$  越小, 预测精度越高。

采用正交试验方法探究了关键超参数对预测性能的作用, 涉及隐含单元数、迭代次数、学习率、批处理大小 4 个维度, 每个维度设定 4 个水平。详细的参数配置如表 4 所示。

通过正交试验设计, 采用四因素四水平进行 16 组试验, 每种方案重复 5 次取其平均值以确保结果的准确性, 图 7 为 16 组超参数优化正交试验结果。结果表明, 实验编号 11 和 14 的模型在平均绝对百分比误差 ( $MAPE$ ) 和根均方误差 ( $RMSE$ ) 方面表现优异, 特别是编号 14 的模型不仅误差较小, 而且训

表 4 正交试验模型因素种类及水平

Table 4 Types and levels of factors in orthogonal experimental models

因素	隐含单元个数	迭代次数	学习率	批处理大小
水平 1	50	200	0.000 1	2
水平 2	100	400	0.001 0	4
水平 3	150	600	0.010 0	8
水平 4	200	800	0.020 0	16

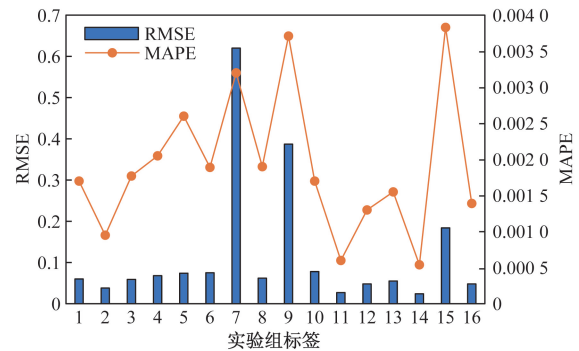


图7 正交实验结果

Fig. 7 Orthogonal experimental results

练时间仅为 883 s, 显著优于其他模型。因此, 选取实验编号 14 对应的参数组合作为模型的最优超参数, 即隐藏单元数为 150, 迭代次数为 800, 学习率为 0.001, 批处理大小 4。

### 3 基于 Transformer 的蒸汽驱见汽时间预测结果分析

#### 3.1 模型预测结果

针对训练好的蒸汽驱见汽时间预测模型, 在测试集上验证模型性能。将测试样本井时间序列后 10% 时间步的瞬时温度比值变化曲线看作未知数据, 运用前 90% 的历史时间序列数据输入模型进行预测, 测试集示例样本井的预测结果如图 8 所示。其中, 对于测试样本井 1, 其截至当前的生产时间为 876 d。在对最后 10% 时间步进行预测时, 第 790 天的瞬时温度比值曲线上升超过了预定基准, 表明发生了汽窜现象; 测试样本井 2 的截至当前的生产时间为 478 d, 在最后 10% 时间步的预测中, 第 436 天瞬时温度比值上升超过基准值, 说明该井出现汽窜; 测试样本井 3 的截至当前的生产时间为 889 d, 在最后 10% 时间步的预测中, 第 801 天瞬时温度比值变化上升超过基准值, 说明该井出现汽窜; 测试样本井 4 的截至当前的生产时间为 987 d, 在后 10% 时间步的预测中, 第 939 天瞬时温度比值上升超过基准值, 说明该井出现汽窜。这些预测结果与实际

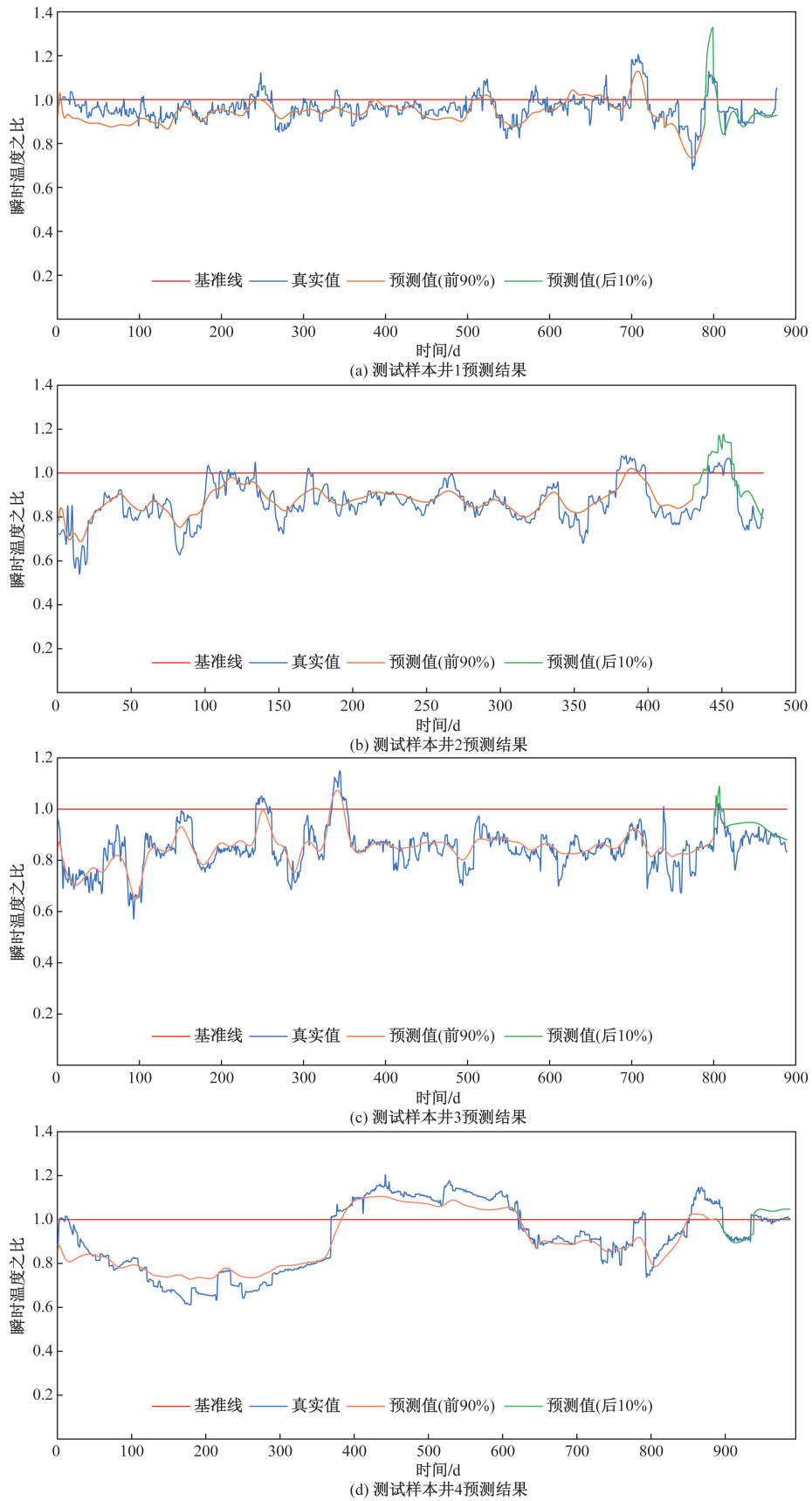


图8 测试样本井1~井4预测结果

Fig. 8 Test sample well 1 ~ well 4 prediction results

油田汽窜时间高度一致,证明了基于 Transformer 的蒸汽驱见汽时间预测模型有效识别了汽窜的关键时刻。

### 3.2 预测结果对比

为进一步展示所建立模型的优势,将基于 Transformer 的蒸汽驱见汽时间预测模型与 LSTM 模块的序列到序列结构进行了比较。采用时间序列较长的样本井 4 作为比较案例,分析两种模型在最后 10% 时间步上的预测结果。预测结果与根均方误差(RMSE)的对比分别如图 9 和图 10 所示。结果显示,基于 Transformer 模型在预测瞬时温度比值的曲线变化形态方面,相比 LSTM 模块展现出了更高的准确度。LSTM 模块虽能跟随当前曲线趋势,但未能明确指出汽窜的准确时刻。特别是在目标曲线波动较大时,LSTM 模型的预测误差相较于基于 Transformer 模型的更为显著。这种性能差异的核心原因在于 Transformer 模型的结构特性。其独有的多头自注意力机制为模型提供了强大的输入特征间关联分析能力,使其能在整个预测流程中自动优化特征之间的权重分配。这一机制使得 Transformer 模型不仅能够有效地识别和整合时间序列中的全局信息,而且还能够在预测过程中适应性地调整对不同时间点信息的重视程度,显著提升了预测的精度与鲁棒性。此外,Transformer 模型的这种高效信息处理能力还为处理复杂、长期的时间序列数据提供了新的可能性,为蒸汽驱见汽时间预测提供了一种更为先进和可靠的技术手段。

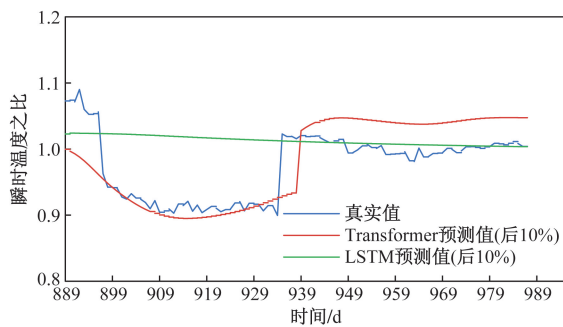


图 9 模型预测结果对比

Fig. 9 Comparison of model prediction results

综合上述分析,所建立的基于 Transformer 的蒸汽驱见汽时间预测模型不仅可以准确地预测出蒸汽驱见汽时间的突破点,而且在面对截至当前的生产时间为较长的油井依然能显示出较强的范性,更适合油田现场汽窜预警的要求。

## 4 结论

(1) 研究聚焦胜利油田稠油区块的 13 个蒸汽

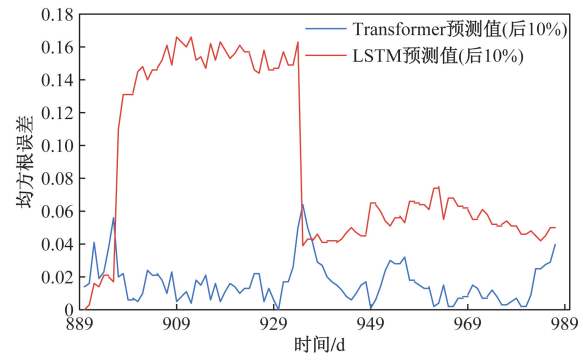


图 10 模型均方根误差对比

Fig. 10 Comparison of RMSE of models

驱井组,采用生产井与注汽井间瞬时温度比作为见汽时间的新表征方法。通过皮尔逊相关系数分析,筛选出与汽窜时间高度相关的主控因素,包括日产量、累积产量、日注汽量、累积注汽量及关键地质参数(油层厚度、渗透率、孔隙度、注采井距),并基于这些特征构建了瞬时温度比的深度学习样本库。

(2) 构建基于 Transformer 架构的蒸汽驱见汽时间预测模型,实现了瞬时温度比的精确预测,有效判断汽窜发生时刻。模型性能分析表明,预测结果与实际观测数据高度一致。与基于 LSTM 的序列到序列模型相比,Transformer 模型在预测精度、适应性和稳定性方面表现更佳。因此,数据驱动的预测策略在应用潜力和发展前景上显示出显著优势,为稠油油藏的有效开发提供了技术支持。

## 参考文献

- [1] 蒋琪,游红娟,潘竞军,等. 稠油开采技术现状与发展方向初步探讨[J]. 特种油气藏, 2020, 27(6): 30-39.  
Jiang Qi, You Hongjuan, Pan Jingjun, et al. A preliminary discussion on the current situation and development direction of thick oil extraction technology [J]. Specialized Oil and Gas Reservoirs, 2020, 27(6): 30-39.
- [2] 徐宏光,束青林,曹嫣媛,等. 稠油油藏化学降黏驱全耦合数值模拟方法[J]. 科学技术与工程, 2023, 23(33): 14201-14210.  
Xu Hongguang, Shu Qinglin, Cao Yanbin, et al. Full coupling numerical simulation of chemical viscosity reduction flooding in heavy oil reservoir [J]. Science Technology and Engineering, 2023, 23(33): 14201-14210.
- [3] 于连东. 世界稠油资源的分布及其开采技术的现状与展望[J]. 特征油气藏, 2001(2): 98-103, 110.  
Yu Liandong. Distribution of thick oil resources in the world and the current status and prospect of its extraction technology [J]. Characterization of oil and gas reservoirs, 2001(2): 98-103, 110.
- [4] 张豆娟,郭海敏,戴家才,等. 稠油油藏蒸汽驱阶段汽窜的研究[J]. 中国测试技术, 2004(3): 45-46, 50.  
Zhang Doujuan, Guo Haimin, Dai Jiakai, et al. Study on vapor runout in steam drive stage of thick oil reservoirs [J]. China Testing Technology, 2004(3): 45-46, 50.

- [5] 凌建军, 宋振宇, 王珏, 等. 蒸汽吞吐阶段的“汽窜”现象实质研究[J]. 江汉石油学院学报, 1996(1): 58-61.  
Ling Jianjun, Song Zhenyu, Wang Jue, et al. Substantial study on the phenomenon of “steam scrambling” in steam throughput stage [J]. Journal of Jianghan Petroleum Institute, 1996(1): 58-61.
- [6] 郑家朋, 东晓虎, 刘慧卿, 等. 稠油油藏注蒸汽开发汽窜特征研究[J]. 特征油气藏, 2012, 19(6): 72-75.  
Zheng Jiapeng, Dong Xiaohu, Liu Huiqing, et al. Characterization of vapor run in thick oil reservoirs developed by steam injection [J]. Characterization of Oil and Gas Reservoirs, 2012, 19(6): 72-75.
- [7] 王永鹏, 邱文昌. 基于 LSTM 神经网络的港口燃油价格预测 [J]. 中国水运, 2019(10): 30-31.  
Wang Yongpeng, Qiu Wenchang. Port fuel oil price prediction based on LSTM neural network [J]. China Water Transportation, 2019(10): 30-31.
- [8] Tatar A, Barati-Harooni A, Moradi S, et al. Prediction of heavy oil viscosity using a radial basis function neural network [J]. Petroleum Science and Technology, 2016, 34(21): 1742-1748.
- [9] 刘合, 李艳春, 杜庆龙, 等. 基于多变量时间序列模型的高含水期产量预测方法 [J]. 中国石油大学学报(自然科学版), 2023, 47(5): 103-114.  
Liu He, Li Yanchun, Du Qinglong, et al. Prediction of production during high water-cut period based on multivariate time series model [J]. Journal of China University of Petroleum (Edition of Natural Science), 2023, 47(5): 103-114.
- [10] Zhang L, Dou H G, Zhang K, et al. CNN-LSTM model optimized by Bayesian optimization for predicting single-well production in water flooding reservoir [J]. Geofluids, 2023, 2023: 1-16.
- [11] Kumar I, Tripathi B K, Singh A. Attention-based LSTM network-assisted time series forecasting models for petroleum production [J]. Engineering Applications of Artificial Intelligence, 2023, 123: 106440.
- [12] 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.  
Cui Chuazhi, 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.
- [13] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. arXiv Preprint, 2017; DOI:10.48550/arXiv.1706.03762.
- [14] Goel R, Susan S, Vashisht S, et al. Emotion-aware Transformer encoder for empathetic dialogue generation [J]. arXiv Preprint, 2022; DOI:10.48550/arXiv.2204.11320.
- [15] 卢婷婷, 刘博, 李纯柱. 基于注意力 Seq2Seq 模型的终端区航空器航迹预测 [J]. 科学技术与工程, 2024, 24(9): 3882-3895.  
Lu Tingting, Liu Bo, Li Chunzhu. Aircraft trajectory prediction within terminal area based on Seq2Seq-attention model [J]. Science Technology and Engineering, 2024, 24(9): 3882-3895.
- [16] Kalyan K S, Rajasekharan A, Sangeetha S. AMMU: a survey of transformer-based biomedical pretrained language models [J]. Journal of Biomedical Informatics, 2022, 126; DOI:10.48550/arXiv.2105.00827.
- [17] Guo M H, Liu Z N, Mu T J, et al. Beyond self-attention: external attention using two linear layers for visual tasks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(5): 5436-5447.
- [18] Liu L Q, Xu X L. Self-attention mechanism at the token level: gradient analysis and algorithm optimization [J]. Knowledge-Based Systems, 2023, 277: 110784.
- [19] Ezzameli K, Mahersia H. Emotion recognition from unimodal to multimodal analysis: a review [J]. Information Fusion, 2023, 99: 101847.
- [20] Peng S, Cao L, Zhou Y, et al. A survey on deep learning for textual emotion analysis in social networks [J]. Digital Communications and Networks, 2022, 8(5): 745-762.
- [21] Zhou M, Duan N, Liu S, et al. Progress in neural NLP: modeling, learning, and reasoning [J]. Engineering, 2020, 6(3): 275-290.
- [22] Lü H, Chen J, Pan T, et al. Attention mechanism in intelligent fault diagnosis of machinery: a review of technique and application [J]. Measurement, 2022, 199: 111594.
- [23] Reza S, Ferreira M C, Machado J J M, et al. A multi-head attention-based transformer model for traffic flow forecasting with a comparative analysis to recurrent neural networks [J]. Expert Systems with Applications, 2022, 202: 117275.
- [24] 付金刚, 杜殿发, 郑洋, 等. 超稠油油藏蒸汽驱动态预测模型 [J]. 石油与天然气地质, 2018, 39(1): 192-197.  
Fu Jingang, Du Dianfa, Zheng Yang, et al. Dynamic prediction model of steam flooding in extra heavy oil reservoirs [J]. Oil & Gas Geology, 2018, 39(1): 192-197.
- [25] 俞佳佳, 金赞, 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.
- [26] 牟振华, 李克鹏, 申栋夫. 基于小波降噪与贝叶斯神经网络联合模型的短时交通流量预测 [J]. 科学技术与工程, 2020, 20(33): 13881-13886.  
Mu Zhenhua, Li Kepeng, Shen Dongfu. Short-term traffic flow prediction based on wavelet denoising and Bayesian neural network model [J]. Science Technology and Engineering, 2020, 20(33): 13881-13886.
- [27] 刘擎龙, 张宏立, 王聪, 等. 基于序列到序列和注意力机制的超短期风速预测 [J]. 太阳能学报, 2021, 42(9): 286-294.  
Liu Bolong, Zhang Hongli, Wang Cong, et al. Ultra-short-term wind speed prediction based on sequence-to-sequence and attention mechanism [J]. Acta Energetica Solaris Sinica, 2021, 42(9): 286-294.