

基于 LSTM-BP 神经网络的地层孔隙压力计算方法

宋先知^{1,2*}, 姚学喆¹, 李根生^{1,2}, 肖立志^{1,2}, 祝兆鹏²

1 中国石油大学(北京)人工智能学院, 北京 102249

2 中国石油大学(北京)油气资源与探测国家重点实验室, 北京 102249

* 通信作者, songxz@cup.edu.cn

收稿日期: 2021-10-13

中国石油天然气集团有限公司-中国石油大学(北京)战略合作科技专项(ZLZX2020-03)、国家重点研发计划“变革性技术关键科学问题”项目(2019YFA0708300)资助

摘要 地层孔隙压力是油气井从设计到完钻过程中重要的基础数据, 准确计算地层孔隙压力是保障钻井安全、提高钻井效率的重要前提。为了克服传统地层孔隙压力计算方法精度不足、计算效率不高的问题, 本文考虑到钻井与地层沉积均为序列性和非线性过程, 提出了将长短期记忆神经网络(LSTM)和误差反向传播神经网络(BP)相结合计算地层孔隙压力的方法, 利用 LSTM 层提取钻-测-录多源数据中的序列性特征信息, 经过 BP 层构建特征信息与地层孔隙压力之间的非线性映射关系。通过对油田现场钻测录数据进行清洗并综合相关性分析和钻井经验知识优选了 18 种输入参数, 对 LSTM-BP 地层孔隙压力计算模型进行训练和测试, 并采用网格搜索法对 LSTM-BP 神经网络模型的 5 种模型超参数进行了优选, 效果最优的单井计算模型和邻井计算模型的平均绝对误差分别为 4.92 MPa 和 2.34 MPa, 均方根误差分别为 6.65 MPa 和 3.03 MPa, 平均相对误差分别为 4.36% 和 8.31%。最后与传统 BP 模型、LSTM 模型和支持向量机(SVM)模型的最优结果进行对比, 结果显示, 本文所建立的 LSTM-BP 神经网络模型精度均高于 BP 模型、LSTM 模型和 SVM 模型, 表明所提出的 LSTM-BP 地层孔隙压力计算模型具有较高的计算精度。

关键词 地层孔隙压力; LSTM-BP 神经网络; 深度学习; 多源数据

A novel method to calculate formation pressure based on the LSTM-BP neural network

SONG Xianzhi^{1,2}, YAO Xuezhe¹, LI Gensheng^{1,2}, XIAO Lizhi^{1,2}, ZHU Zhaopeng²

1 College of Artificial Intelligence, China University of Petroleum-Beijing, Beijing 102249, China

2 State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum-Beijing, Beijing 102249, China

Abstract The formation pore pressure is an important basic parameter in the process of oil and gas drilling from well design to well completion. It is an important basis for the rational design of drilling plans and analysis of wellbore stability, and accurate calculation of formation pore pressure is an important prerequisite for ensuring drilling safety and improving drilling efficiency. In order to overcome the problems of insufficient accuracy and low calculation efficiency of traditional formation pore pressure calculation methods, this paper takes into account that both the drilling process and the formation deposition process have a certain degree of sequentiality and complex nonlinearity, so this article proposes a method to calculate formation pore pressure

引用格式: 宋先知, 姚学喆, 李根生, 肖立志, 祝兆鹏. 基于 LSTM-BP 神经网络的地层孔隙压力计算方法. 石油科学通报, 2022, 01: 12-23

SONG Xianzhi, YAO Xuezhe, LI Gensheng, XIAO Lizhi, ZHU Zhaopeng. A novel method to calculate formation pressure based on the LSTM-BP neural network. Petroleum Science Bulletin, 2022, 01: 12-23. doi: 10.3969/j.issn.2096-1693.2022.01.002

by combining a Long Short-Term Memory (LSTM) neural network and an error Back Propagation (BP) neural network based on drilling-logging-recording data. The LSTM layer in the neural network is used to extract the serial feature information in the multi-source data of drilling, logging, and recording, and the BP layer in the neural network is used to construct a nonlinear mapping relationship between characteristic information and formation pore pressure. The field data of an oilfield was cleaned and processed, and 18 parameters such as drilling time, weight on bit, dc-exponent, sonic time difference, and density logging were optimized through comprehensive correlation analysis and drilling experience knowledge, and the LSTM-BP formation pore pressure calculation model was carried out with training, validation and testing, and using the grid search method to analyze and optimize the 5 hyperparameters of the LSTM-BP model, including the number of LSTM layers, the number of neural units in the LSTM unit gate, the number of BP layers, the number of neurons in the BP layer, and the activation function. The mean absolute error of the best single well calculation model and the best adjacent well calculation model were 4.92 MPa and 2.34 MPa, the root mean square error were 6.65 MPa and 3.03 MPa, and the mean relative error were 4.36% and 8.31%. Finally, the LSTM-BP model is compared with the optimized traditional BP neural network model, LSTM neural network model, and Support Vector Machine (SVM) model. The results show that the accuracy of the LSTM-BP neural network model established in this paper is higher than that of the BP neural network model, LSTM neural network model, and SVM model, which show that the LSTM-BP formation pore pressure calculation model proposed in this paper has a high calculation accuracy.

Keywords formation pore pressure; LSTM-BP neural network; deep learning; multi-resources data

doi: 10.3969/j.issn.2096-1693.2022.01.002

0 引言

地层孔隙压力是石油勘探开发过程中最重要的基础数据之一,是合理设计钻井方案、分析井壁稳定性的重要依据,在钻井过程中准确计算地层孔隙压力有利于提高机械钻速,保护储层,对提高油气采收率具有重要作用^[1]。

地层孔隙压力计算主要可分为钻前预测、随钻监测与钻后评估 3 个阶段^[2]。其中钻前预测常采用地质分析法和地震法;随钻监测常采用 dc 指数法^[3-4]、标准钻速法^[5]和岩石强度法^[6-7]等;钻后评估通过等效深度法^[8]、Eaton 法^[9-10]和有效应力法^[11]。等对地层压力进行计算评估。国内外学者对地层孔隙压力计算开展了大量研究,吴建忠^[12]等对比了 dc 指数和 sigma 指数法,认为 dc 指数法更适合计算川西地区地层压力;付明川^[13]针对 dc 指数值波动大,变化频繁的问题,对 dc 指数模型进行了修正,稳定性、准确性明显提升;Primasty^[14]等采用等效深度法分析了加拿大 Nova Scotia 地区的异常孔隙压力深度;Bektas^[15]等利用卡尔曼滤波过滤随钻测井噪声,基于 Eaton 法预测地层孔隙压力;Ziegler^[16]等利用 Bowers 等方法计算地层孔隙压力,并利用随钻数据模型进行校正;李中^[17]等考虑岩性和孔隙度对声波速度的影响,修正了传统 Bowers 法卸载方程;陶磊^[18]等提出一种多元回归模型的地层孔隙压力计算方法。但上述传统地层孔隙压力计算方法受岩性、水力参数等影响较大,未考虑数据的序列性,且需要建立正常压实趋势线方程,计算过程复杂,经验参数较多导致误差累加,计算精度有限。因此亟

需一种高效精准的地层孔隙压力实时计算方法。

近些年来人工智能算法快速发展,在解决复杂非线性问题上具有很大的优势,在石油领域中的地震反演^[19-20]、测井解释^[21-23]、产量预测^[24-26]和开发优化^[27-28]等方面的应用取得了显著效果。其中误差反向传播神经网络(error Back Propagation Neural Network, BP)具有高度的自学习自适应能力和较强的泛化能力,适用于解决地层孔隙压力计算等复杂非线性问题;地层沉积和钻井过程表明所产生的数据具有一定序列性,而长短期记忆循环神经网络(Long Short-Term Memory Neural Network, LSTM)因具有一定的记忆能力,能够较好的处理钻井、测井和录井等序列性数据。本文充分分析了钻井、测井和录井数据及其存在的序列性和非线性关系,创新地提出了利用 LSTM-BP 神经网络模型实时计算地层孔隙压力的方法,以进一步提高神经网络模型在此类问题上的表现。并基于钻井、测井和录井数据,对 LSTM-BP 神经网络模型进行了训练和测试,结果表明 LSTM-BP 神经网络模型比 BP 神经网络、LSTM 神经网络和支持向量机(Support Vector Machines, SVM)模型的精度更高,为地层孔隙压力的准确实时计算提供了一种新方法。

1 神经网络原理

1.1 神经元与 BP 神经网络

神经元是人工神经网络中最基本的单元,McCulloch 和 Pitts^[29]受生物学中神经系统的启发,将其抽

象为简单的数学模型,如图1所示。神经元接收其它神经元传递的输入信号 x_i ,通过权重 w_i 加权后计算总输入并与阈值 b 进行比较,最后通过激活函数 f 计算后产生神经元的总输出。其正向传播过程为:

$$y = f\left(\sum_{i=1}^n w_i x_i - b\right) \quad (1)$$

神经网络是神经元逐层组织以非循环方式连接而成的结构,如图2所示,一些神经元的输出可以是另外神经元的输入,通常层与层为全连接的形式,层内神经元不连接,神经元层按功能分为输入层、隐藏层和输出层。1984年Rumelhart^[30]等人提出了误差反向传播神经网络,其是按误差逆向传播算法训练的多层前馈网络,也是目前应用最广泛的神经网络模型之一。它通过在每轮迭代中采用广义感知机学习规则对权重和阈值参数进行更新估计,具有非常强的函数逼近能力,但不具备记忆能力。

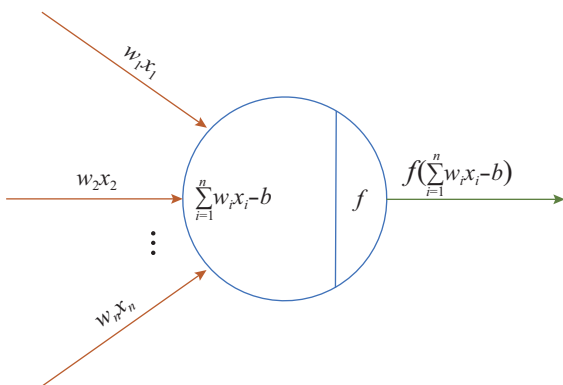


图1 M-P 神经元模型

Fig. 1 M-P neuron model

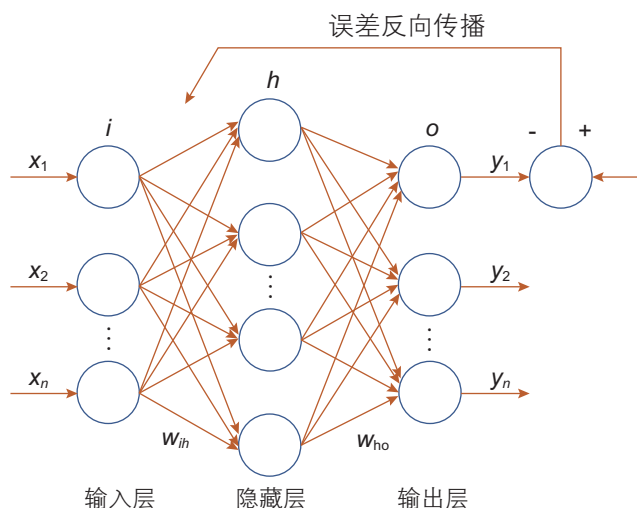


图2 一个隐藏层的BP神经网络

Fig. 2 A hidden layer of BP neural network

1.2 LSTM神经网络

循环神经网络(Recurrent Neural Network, RNN)主要用于处理序列数据,其最大的特点就是神经元在某时刻的输出可以作为输入再次输入到神经元,这种串联的网络结构非常适合于时间序列数据,可以保持数据中的依赖关系,且具有一定记忆能力^[31]。但大量实践表明,RNN很难实现信息的较长时间保存,Bengio^[32]等认为RNN在不断迭代过程中存在梯度消失和梯度爆炸的问题导致其应用效果不佳。

为解决长期依赖的问题,Hochreiter等^[33]在RNN的基础上提出了长短期记忆网络,之后Graves^[34]对其进行了优化。如图3所示,与M-P神经元相比,LSTM单元的结构更加复杂,其可通过遗忘门(f)、输入门(i)和输出门(o)三种门控使信息有选择地添加或减少,LSTM单元每个时刻的输入包含前一时刻的记忆单元(C)状态、隐状态和当前时刻输入。首先是当前时刻输入和前一时刻隐状态,依次通过遗忘门和输入门,完成记忆单元状态的更新,然后根据更新后的记忆单元状态和输出门,完成隐状态的更新,这使其拥有较长的短期记忆能力,并避免了梯度消失和爆炸的问题。

遗忘门可根据当前时刻输入和上一时刻隐状态决定要忘记的信息,

$$f_t = \sigma(W_f(h_{t-1}, x_t) + b_f) \quad (2)$$

输入门根据当前时刻输入和上一时刻隐状态决定要添加的新信息,

$$i_t = \sigma(W_i(h_{t-1}, x_t) + b_i) \quad (3)$$

$$\text{新信息: } \tilde{C}_t = \tanh(W_c(h_{t-1}, x_t) + b_c) \quad (4)$$

记忆单元根据上一时刻状态和产生的新信息对状态进行更新,

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (5)$$

输出门将当前时刻输入上一时刻隐状态进行合并计算,并利用当前时刻的记忆单元状态对隐状态进行更新,

$$O_t = \sigma(W_o(h_{t-1}, x_t) + b_o) \quad (6)$$

$$\text{当前隐状态输出: } h_t = O_t \tanh(C_t) \quad (7)$$

其中, x_t 为当前的输入; h_{t-1} 为上一时刻的输出; W_f 为LSTM遗忘门中神经元的权重矩阵; b_f 为LSTM遗忘门中神经元的偏置向量; f_t 为LSTM遗忘门中神经元的输出; W_i 为LSTM输入门中神经元的权重矩阵; b_i 为LSTM输入门中神经元的偏置向量; i_t

为 LSTM 输入门中神经元的输出； W_o 为 LSTM 输出门中神经元的权重矩阵； b_o 为 LSTM 输出门中神经元的偏置向量； O_t 为 LSTM 输出门中神经元的输出； σ 为 sigmoid 激活函数； \tilde{C}_t 为 LSTM 记忆单元中新的信息； W_c 为 LSTM 记忆单元中神经元的权重矩阵； b_c 为 LSTM 记忆单元中神经元的偏置向量； C_{t-1} 为上一时刻 LSTM 记忆单元的状态， \tanh 为激活函数。

2 数据处理

钻井、测井和录井原始数据中通常含有较多的异常值和缺失值，会对模型的训练效果和计算效率产生较大的影响。因此合理的数据处理对于人工智能模型来说是至关重要的，不仅能够提高模型的质量，还可以减少模型的计算时间。为了保证地层孔隙压力模型的计算精度，本文对原始钻井、测井和录井数据进行了清洗、补充、相关性分析和归一化等操作，建立了用于训练验证和测试模型的高质量数据集。

2.1 数据清洗

本文所使用的数据为国内某油田提供的现场数据，包括钻井工程参数、测斜参数、录井参数和测井参数等 50 多种参数，利用 python 相关第三方库对其进行了异常值删除、数据补充、钻测录数据拼接并分井保存等操作。基于声波时差、密度等测井数据利用 Eaton 法计算地层孔隙压力标签，并利用实测值对压力剖面进行校正作为地层孔隙压力实际值，与实测点的相对误差仅为 2% 左右，最终建立可用于训练的地层孔隙压力样本数据集。

2.2 相关性分析

相关性分析是量化不同因素间变动状况一致程度的重要指标。数据中往往存在大量的冗余或者不相关特征，且特征过多极易引发“维数灾难”，使得模型复杂度和计算时间大大增加。本文采用距离相关系数^[35]来计算各参数与地层孔隙压力之间的线性和非线性相关性，并结合钻井经验知识优选模型输入特征。

距离相关系数既能反映变量之间的线性关系，也可以在一定程度上反映各变量之间的非线性关系，且不需要任何的模型假设和参数条件^[36]。距离相关系数 $R_n(X, Y)$ 的计算公式如下，

$$R_n(X, Y) = \begin{cases} \frac{V_n^2(X, Y)}{\sqrt{V_n^2(X) V_n^2(Y)}}, & V_n^2(X) V_n^2(Y) > 0 \\ 0, & V_n^2(X) V_n^2(Y) = 0 \end{cases} \quad (8)$$

其中， $V_n^2(X, Y)$ 为 X, Y 的距离协方差； $V_n^2(X)$ 为 X 的距离方差； $V_n^2(Y)$ 为 Y 的距离方差； $R_n(X, Y)$ 为 X, Y 的距离相关系数。

距离相关系数仅从统计学角度在一定程度上分析了各参数与地层孔隙压力之间的线性和非线性相关性，但钻井、测井和录井参数与地层孔隙压力之间存在着复杂的非线性关系，因此需要结合钻井经验知识综合选取输入特征。由钻井经验知识可知，在正常地层沉积过程中，地层孔隙压力随井深的增加而逐渐增加；钻进时若遇到异常高压会导致钻速、电导率增加，钻时、钻压、dc 指数和钻井液密度降低；声波时差、自然伽马和密度测井则能反映地层孔隙度和岩性的变化情况。

本文计算了各参数与地层孔隙压力之间的距离

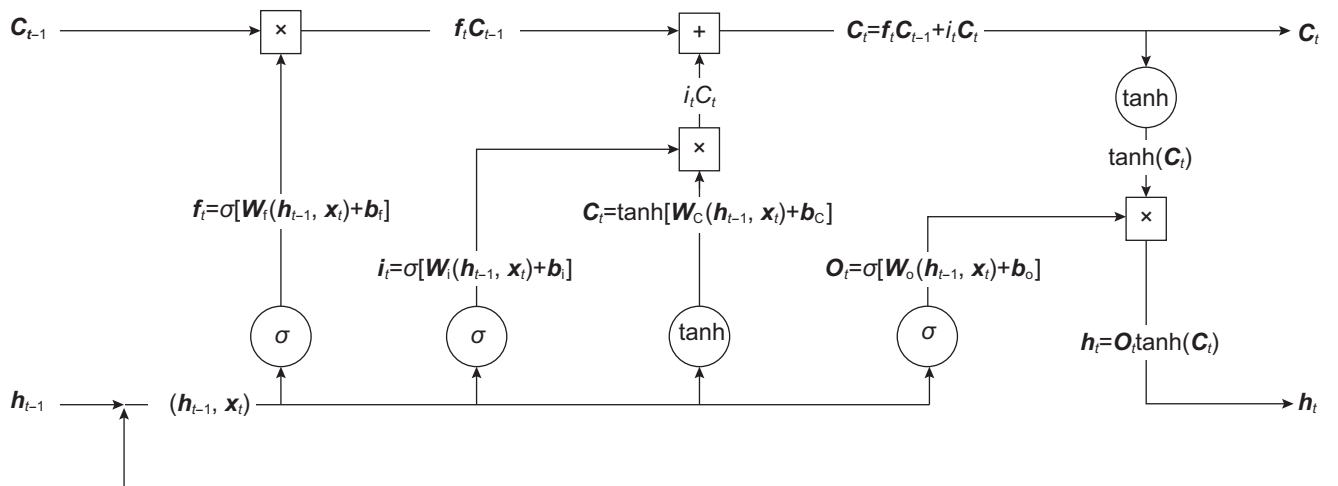


图 3 LSTM 神经网络结构

Fig. 3 Structure of LSTM neural network

相关系数,如图4所示,由统计学知识可知,相关系数为0.4~0.6为中等程度相关,0.6~0.8为强相关,0.8~1.0为极强相关,以相关性强度大于0.4为标准并结合经验知识确定了标准井深、钻时、dc指数、钻压、立管压力、出口流量、入口密度、当量密度、入口电导率、出口电导率、大钩负荷、泵冲速等12种钻录井参数和井径、声波时差、自然伽马、密度4种测井参数为输入。

2.3 数据规范化处理

数据规范化是一个常用的数据预处理操作,目的是处理不同规模和量纲的数据,使其缩放到相同的数据区间和范围,以减少规模、特征、分布差异等对模型的影响。常用的数据规范化方法有Z-Score和Max-Min等,考虑到钻井是一个随深度的变化的序列化过程,井深、钻压等参数并不符合正态分布,为消除量纲和减小各参数的不同数量级对模型性能的影响,本文采取Max-Min方法对测井、钻录井数据进行归一化处理。

$$\tilde{X}_i = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)} \quad (9)$$

其中, \tilde{X}_i 为第*i*个变量Max-Min归一化后的结果; X_i 为原始第*i*个变量; $\min(X_i)$ 为原始第*i*个变量的最小值; $\max(X_i)$ 为原始第*i*个变量的最大值。

3 模型搭建与参数优选

3.1 模型评价指标

对模型的表现性能进行评价是机器学习建模的一

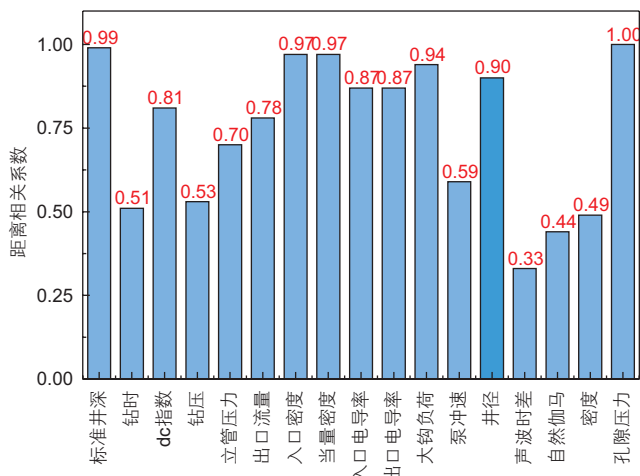


图4 距离相关系数

Fig. 4 Distance correlation coefficient

个重要环节,这不仅仅需要有效可行的估计方法,还需要有衡量模型泛化能力的评价标准。本文所研究的地层孔隙压力计算属于回归问题,常用的评价指标有平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Squared Error, RMSE)和平均相对误差(Mean Relative Error, MRE)。

平均绝对误差是真实值与预测值之间绝对误差的平均值,能更好地反映预测值误差的实际情况,其计算公式为:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_{\text{true}} - y_{\text{predict}}| \quad (10)$$

均方根误差可用来衡量预测值与真实值之间的偏差,计算公式为:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{\text{predict}} - y_{\text{true}})^2} \quad (11)$$

平均相对误差指真实值与预测值的相对误差的平均值,其计算公式为:

$$MRE = \frac{1}{N} \frac{|y_{\text{true}} - y_{\text{predict}}|}{y_{\text{true}}} \quad (12)$$

式中 y_{true} 表示样本真实值; y_{predict} 表示样本预测值; N 为样本数量。

3.2 模型搭建

考虑到钻井是一个与深度相关的序列性过程,且地层的沉积过程属于时序性过程,由于LSTM神经网络适合处理时序性问题,故采用LSTM神经网络算法来提取钻测录数据前后存在的序列性关系,进而利用BP神经网络进一步挖掘各变量与地层孔隙压力间的非线性关系,LSTM-BP串联神经网络模型结构如图5所示。

3.3 地层孔隙压力单井计算模型参数优选

本文首先采用某油田区块T-1井进行地层孔隙压力计算模型的训练验证和测试,其中共5980条数据,前75%数据用于训练和验证,后25%数据用于模型测试。模型采用Adam优化器进行训练,与传统梯度下降算法相比,Adam优化器能够在训练迭代过程中自适应调整学习率,显著提升模型学习速度^[37]。采用dropout方法随机使隐藏层神经元以一定概率失活,来提升模型泛化能力并防止模型过拟合。前期预训练实验结果表明,当输入序列长度为1,dropout参数为0.2,训练迭代轮数为300,批尺度为50时模型能达到较好的效果,故利用交叉验证方法对LSTM神经网络层数(2层、3层)、LSTM神经网络门控的神经单元

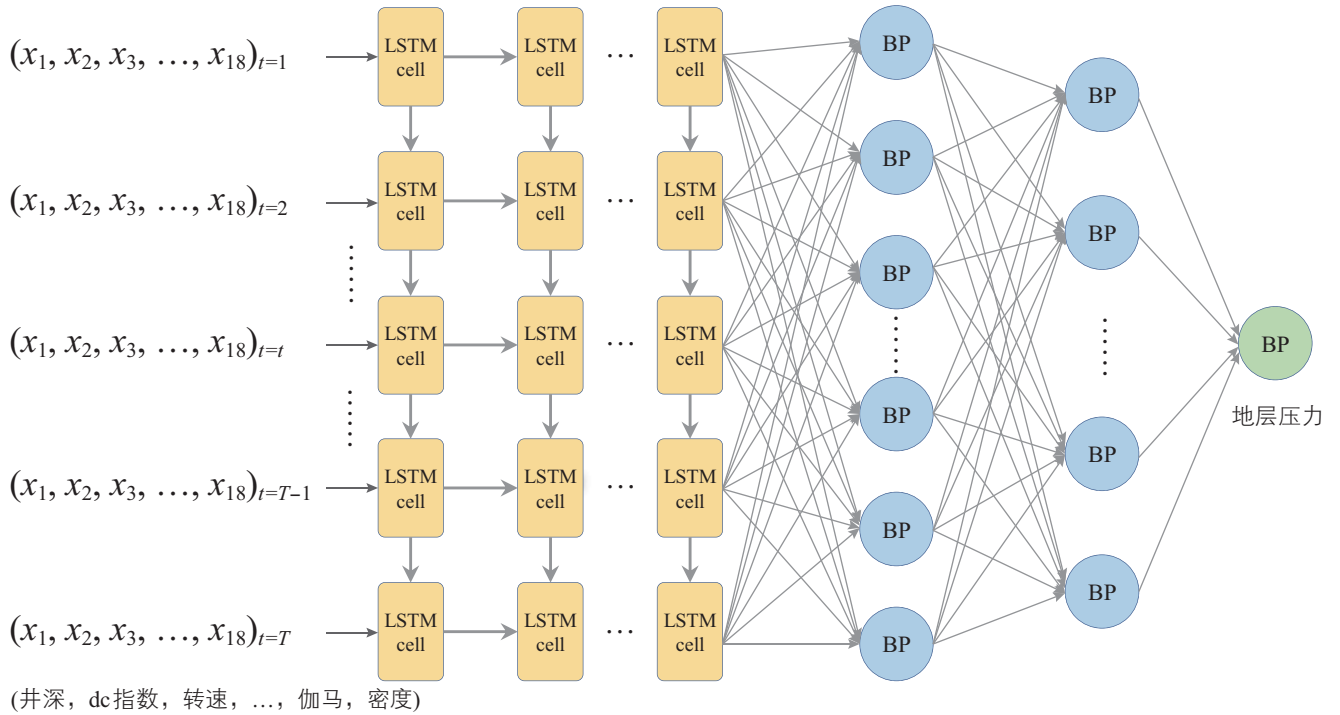


图 5 LSTM-BP 网络结构模型 (T 为输入步长, x 为不同的输入特征)

Fig. 5 LSTM-BP network structure model (T is the input step size, x is the different input features)

数(10 个、12 个)、BP 神经网络层数(2 层、3 层)、BP 神经网络每层单元数(20 个、24 个)和激活函数(reLU、tanh、sigmoid)共 48 种模型参数组合进行优选, 不同模型参数的部分测试结果如表 1 所示。

利用 MAE、RMSE、MRE 和计算时间 4 个评价指标对模型参数进行优选, 这 4 个指标越小表明模型的拟合能力越强。由表 1 和图 6 可以看出, 第 1 个参数组合的模型表现最佳, 其在测试集上的 MAE 为 4.9169 MPa, RMSE 为 6.6521 MPa, MRE 为 4.36%, 与实测

点的平均相对误差为 8.39%, 远低于其他模型, 虽然第二个模型的计算时间略低于第一个, 但是综合考虑所有参数考虑, 本文认为第一个模型为最优模型。

3.4 地层孔隙压力邻井计算模型参数优选

单井计算模型只是针对 1 口井进行训练测试, 往往只适用于利用一口井上部数据训练模型, 下部数据进行计算的情况。利用其来计算邻井地层孔隙压力效果并不理想, 为提高地层孔隙压力计算模型的在该油

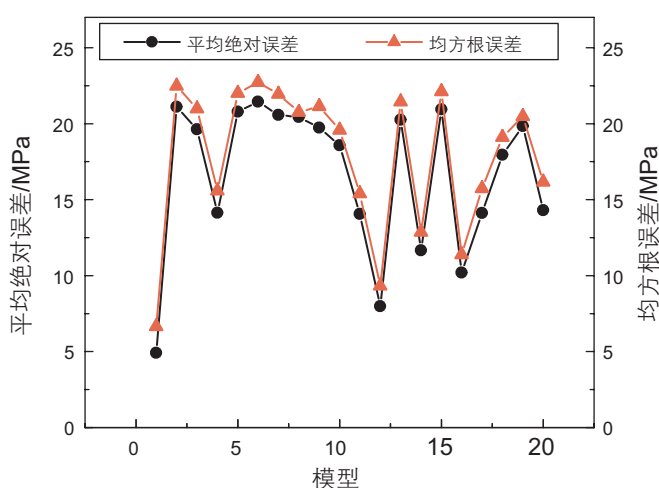
表 1 不同单井模型参数下的测试集效果 (前 20)

Table 1 Effect of test set under different single well model parameters (Top 20)

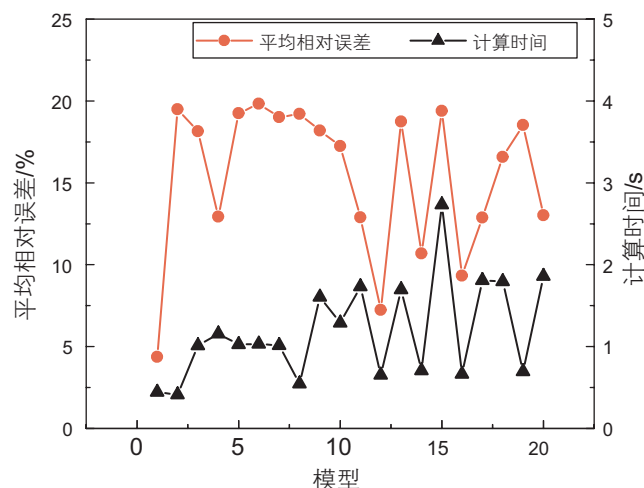
序号	LSTM 层参数		BP 层参数		激活函数	MAE/MPa	RMSE/MPa	MRE/%	计算时间/s
	层数	神经元个数	层数	神经元个数					
1	2	10	2	20	relu	4.9169	6.6521	4.36	0.4443
2	2	10	2	20	sigmoid	21.1150	22.4889	19.49	0.4122
3	2	10	2	20	tanh	19.6295	20.9872	18.14	1.0117
4	2	10	2	24	tanh	14.1326	15.5902	12.93	1.1563
5	2	10	3	20	tanh	20.7876	22.0132	19.25	1.0252
6	2	10	3	24	tanh	21.4395	22.7199	19.83	1.0315
7	2	12	2	24	tanh	20.5783	21.9487	19.01	1.0152
8	2	12	3	20	relu	20.4343	20.7589	19.21	0.5441
9	2	12	3	20	tanh	19.7303	21.1460	18.19	1.6046
10	2	12	3	24	tanh	18.5791	19.5880	17.24	1.2878
11	3	10	2	20	tanh	14.0569	15.3983	12.89	1.7327

(续表)

序号	LSTM层参数		BP层参数		激活函数	MAE/MPa	RMSE/MPa	MRE/%	计算时间/s
	层数	神经元个数	层数	神经元个数					
12	3	10	2	24	relu	7.9894	9.3284	7.24	0.6534
13	3	10	2	24	tanh	20.2545	21.4508	18.74	1.6970
14	3	10	3	20	relu	11.6561	12.8730	10.68	0.7053
15	3	10	3	24	tanh	20.9350	22.1088	19.39	2.7345
16	3	12	2	20	relu	10.2001	11.3827	9.33	0.6655
17	3	12	2	20	tanh	14.1200	15.7410	12.88	1.8102
18	3	12	3	20	tanh	17.9506	19.1068	16.58	1.7951
19	3	12	3	24	relu	19.8480	20.4802	18.53	0.6945
20	3	12	3	24	tanh	14.3033	16.1701	13.02	1.8601



(a) 平均绝对误差与均方根误差



(b) 平均相对误差与计算时间

图6 不同单井模型评价指标对比

Fig. 6 Comparison of evaluation indexes of different single well models

田区块的泛化能力,采用某油田区块T-2井、T-3井和T-4井进行训练和验证,利用T-1井进行模型的测试。选取模型的输入序列长度为1, dropout 参数为0.2,训练迭代轮数为400,批尺度为100,采用交叉验证方法LSTM神经网络层数(2层,3层)、LSTM神经网络门控的神经单元数(10个,12个,14个)、BP神经网络层数(2层,3层)、BP神经网络每层单元数(20个,24个,28个)和激活函数(relu, tanh, sigmoid)共5种模型参数进行优选。部分模型测试结果如表2所示。

由表2和图7可以看出,第4和第6个模型表现都较好,在测试集上的MAE分别为2.698 MPa、2.336 MPa, RMSE分别为3.708 MPa、3.030 MPa, MRE分别为8.76%、8.31%。与实测点的平均相对误差分别为2.04%和1.23%。由于均方根误差计算时将误差平方,其受异常值影响较大,当其于平均绝对误差的比值越

小,模型效果越好。结合实际预测效果,如图8所示,选取第6个模型为最优。

4 不同算法模型对比

为了证明LSTM-BP模型的效果,本文还将其与BP神经网络模型、LSTM神经网络模型和SVM模型进行了对比分析。

在地层孔隙压力单井计算模型方面,采用T-1井的数据对BP神经网络模型、LSTM神经网络模型和SVM模型进行训练验证和测试,前75%的数据进行训练和验证,后25%的数据进行测试;在地层孔隙压力邻井计算模型方面,采用T-2、T-3、T-4井数据进行训练验证,采用T-1井数据测试。利用交叉验证法对BP神经网络的层数(2层,3层,4层)、每层神经单元数(24个,28个,32个)、激活函数(tanh, relu,

表 2 不同邻井模型参数下的测试集效果 (前 20)

Table 2 Effect of test set under different adjacent well model parameters (Top 20)

序号	LSTM层参数		BP层参数		激活函数	MAE/MPa	RMSE/MPa	MRE/%	计算时间/s
	层数	神经元个数	层数	神经元个数					
1	2	10	2	20	tanh	4.6492	5.2439	13.90	1.2637
2	2	10	3	20	sigmoid	4.5182	5.3880	13.14	0.5259
3	2	12	2	24	tanh	3.9348	4.5261	13.37	1.1482
4	2	12	2	28	tanh	2.6977	3.7081	8.76	1.1445
5	2	12	3	24	tanh	3.5100	4.2924	11.74	1.1508
6	2	12	3	28	tanh	2.3357	3.0295	8.31	1.3043
7	2	14	2	28	sigmoid	4.1447	4.7892	13.15	0.5123
8	2	14	3	24	tanh	4.0730	4.6833	11.96	1.1568
9	2	14	3	28	tanh	3.8912	4.4348	13.74	1.1769
10	3	10	2	20	tanh	4.3950	5.0765	12.45	2.1015
11	3	10	3	20	sigmoid	4.7458	5.6665	14.20	0.8174
12	3	12	3	24	relu	3.8034	4.6933	8.35	0.7106
13	3	12	3	24	sigmoid	4.4092	5.2163	13.69	0.6980
14	3	12	3	28	relu	3.3993	4.4306	11.55	0.6996
15	3	12	3	28	sigmoid	3.7798	4.6109	13.58	0.7211
16	3	14	2	24	sigmoid	4.0927	5.0269	13.53	0.6550
17	3	14	2	28	tanh	4.5849	5.3907	11.98	1.6748
18	3	14	2	28	sigmoid	3.8769	4.4423	13.76	0.6598
19	3	14	3	24	tanh	4.2824	5.2434	13.22	1.6577
20	3	14	3	28	tanh	3.3308	4.1829	11.71	1.6362

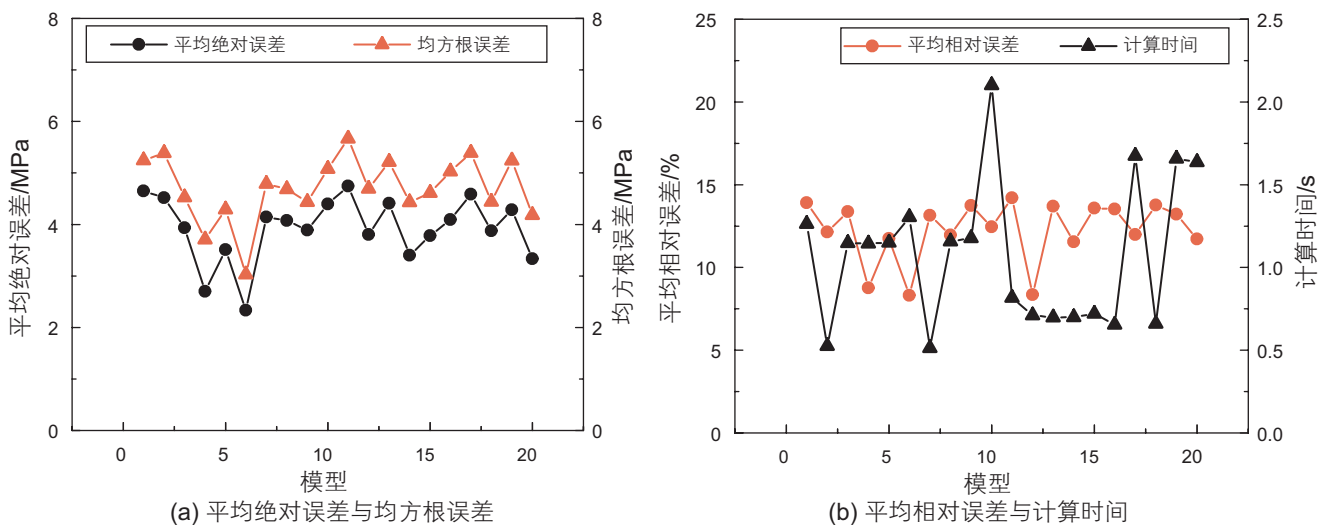


图 7 不同邻井模型评价指标对比

Fig. 7 Comparison of evaluation indexes of different adjacent well model

sigmoid)、训练迭代轮数(300, 400)和批尺度(20, 50)进行优选;对LSTM神经网络的层数(2层, 3层)、神经元个数(10个, 12个, 14个)、激活函数(tanh, relu, sigmoid)、训练迭代轮数(300, 400)和批尺度(20, 50)进行优选;对SVM模型的核函数、核函数系数、惩罚参数和epsilon参数进行优选。BP模型、LSTM模型与

SVM模型参数优选结果如表3所示。

在地层孔隙压力单井计算方面, LSTM-BP模型的MAE为4.9169 MPa、RMSE为6.6521 MPa、MRE为4.36%, 与实测点相对误差为8.39%。均低于传统的BP神经网络模型、LSTM模型和SVM模型。四种最优模型的计算值与实际值的对比如图9所示, 误差对

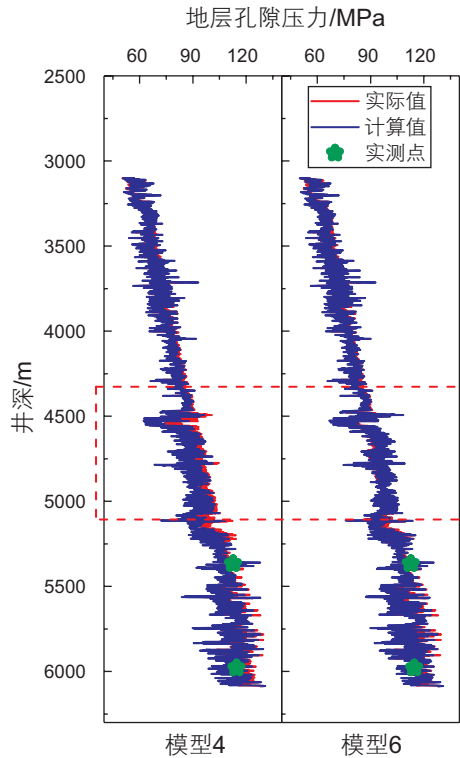


图 8 两种不同参数的邻井模型预测效果对比
 Fig. 8 Comparison of prediction effects of two adjacent well models with different parameters

比如图 10 所示。

在地层孔隙压力邻井计算方面，LSTM-BP模型的MAE为2.3357 MPa、RMSE为3.0295 MPa、MRE为8.31%，与实测点的平均相对误差为1.23%。也均小于传统的BP神经网络模型、LSTM模型和SVM模型，四种最优模型误差对比如图 11 所示，计算值与实际值

的对比如图 12 所示。

由此可以看出，本文所建立的LSTM-BP神经网络计算模型的表现要优于传统的BP神经网络模型、LSTM神经网络模型和SVM模型，这是可能由于增加了LSTM层后，神经网络模型能够较好的学习到测井

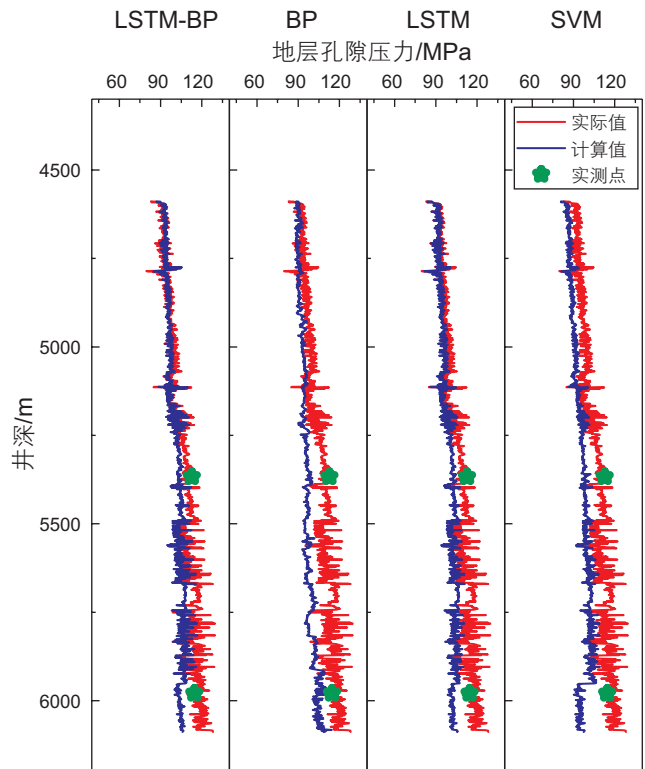


图 9 不同单井模型计算值与实际值对比 (T-1 井 4589 m~6088 m)
 Fig. 9 Comparison of calculated and actual values of different single well models (T-1 Well 4589 m~6088 m)

表 3 BP神经网络模型、LSTM神经网络模型和SVM模型参数优选结果

Table 3 Results of parameter preferences for BP neural network model, LSTM neural network model and SVM model

模型	最优参数组合					MAE /MPa	RMSE /MPa	MRE /%	计算时间 /s	实测点 误差/%	
	网络层数	神经元数	激活函数	迭代轮数	批尺度						
单井模型	BP	2	32	tanh	400	20	6.4084	8.1936	9.29	0.1456	11.85
	LSTM	3	12	relu	300	50	6.3811	8.0996	5.69	2.8477	10.12
	SVM	核函数	核函数系数	惩罚参数	<i>epsilon</i>						
		linear	auto	2	0.05		10.4017	11.8627	9.53	0.0006	15.87
邻井模型	BP	3	32	sigmoid	300	20	2.6547	3.2068	9.46	0.2406	2.29
	LSTM	网络层数	神经元数	激活函数	迭代轮数	批尺度					
		3	12	tanh	300	50	5.1425	5.8673	10.45	2.4757	2.61
	SVM	核函数	核函数系数	惩罚参数	<i>epsilon</i>						
		rbf	scale	3	0.05		2.6865	3.3112	10.60	0.0119	1.40

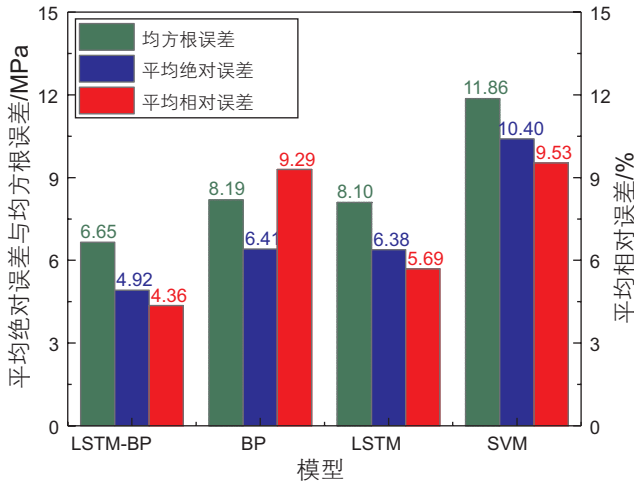


图 10 四种单井模型误差对比
Fig. 10 Comparison of errors in four single well models

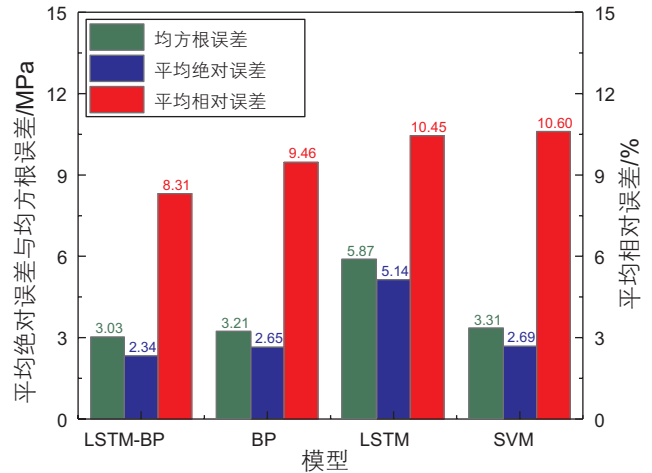


图 11 四种邻井模型误差对比
Fig. 11 Comparison of errors in four adjacent well models

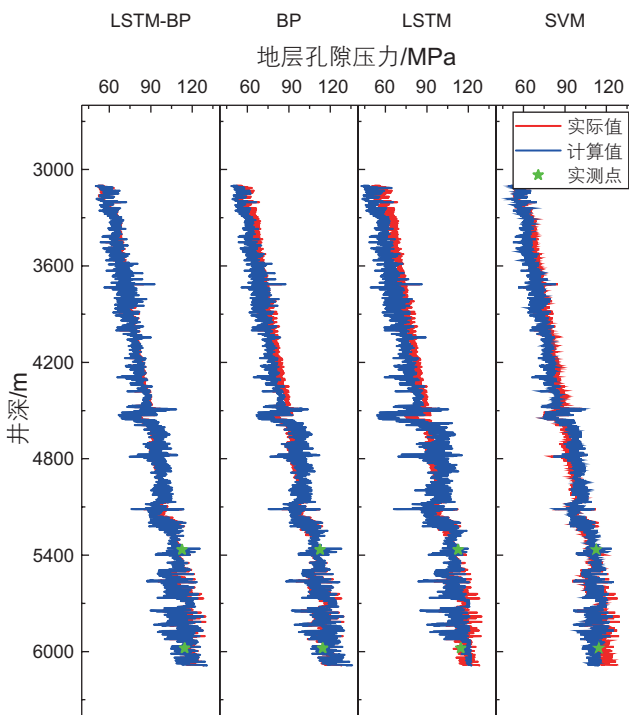


图 12 不同邻井模型计算值与实际对比 (T-1 井 3100 m~6088 m)
Fig. 12 Comparison of calculated and actual values from different adjacent well models (T-1 Well 3100 m~6088 m)

和钻井数据中存在的关键的序列性特征信息，这些信息经过 BP 神经网络层的进一步的融合，建立了钻-

测-录数据与地层孔隙压力之间复杂的非线性映射关系，使 LSTM-BP 神经网络模型在处理钻井工程和地层沉积等序列性问题上具有较大的优势。

5 结论

(1) 本文利用距离相关系数分析了地层孔隙压力与钻井、测井以及录井参数之间的相关性，结合钻井经验知识优选出了标准井深、钻时、dc 指数、立管压力、当量密度、出口电导率、泵冲速、井径、声波时差、自然伽马、密度测井等 16 种参数作为模型输入参数。

(2) 考虑了钻井过程与地层沉积过程的序列性，分析了钻井、测井和录井参数与地层孔隙压力之间具有较强的非线性关系，构建了 LSTM-BP 神经网络模型，采用 LSTM 层提取输入参数中关键的序列性特征信息，进而通过 BP 层来构建参数与地层孔隙压力之间的非线性映射关系。

(3) 将 LSTM-BP 模型与优化后的 BP 神经网络模型、LSTM 神经网络模型和 SVM 模型进行对比，结果表明 LSTM-BP 神经网络模型的精度最高，较传统方法有了较大的提升，为地层孔隙压力提供了一种新的智能计算方法。

参考文献

[1] 饶蕾, 崔杰. 用钻井参数实时监测地层压力的新方法[J]. 石油钻采工艺, 2004(02): 35-37. [RAO L, CUI J. New method of formation pressure real time monitoring using drilling parameters[J]. Oil Drilling & Production Technology, 2004(02): 35-37.]
 [2] 么春雨, 曹军, 苑仁国, 等. 渤海油田随钻地层压力监测 dc 指数法改进及应用[J]. 录井工程, 2018, 29(04): 29-31. [YAO C Y,

- CAO J, YUAN R G, et al. Improvement and application of dc index method for formation pressure monitoring while drilling in Bohai Oilfield[J]. *Mud Logging Engineering*, 2018, 29(04): 29–31.]
- [3] JORDEN J R, SHIRLEY O J. Application of drilling performance data to overpressure detection[J]. *Journal of Petroleum Technology*, 1966, 18(11): 1387–1394.
- [4] REHM B, MCCLENDON R. Measurement of formation pressure from drilling data[C]//Fall Meeting of the Society of Petroleum Engineers of AIME. *OnePetro*, 1971.
- [5] VIDRINE D J, BENIT E J. Field verification of the effect of differential pressure on drilling rate[J]. *Journal of Petroleum Technology*, 1968, 20(07): 676–682.
- [6] 杨进. 地层孔隙压力评估新方法研究及其应用[D]. 北京: 中国石油大学(北京), 1996. [YANG J. Research and application of a new method for evaluating formation pore pressure[D]. Beijing: China University of Petroleum (Beijing), 1996]
- [7] 杨进. 岩石抗钻强度与地层压力关系模型及其应用[J]. *石油大学学报(自然科学版)*, 2001(02): 1–2. [YANG J. Correlation model of rock strength and formation pore pressure and application of the model[J]. *Journal of China University of Petroleum (Edition of Natural Sciences)*, 2001(02): 1–2.]
- [8] HOTTMAN C E, JOHNSON R K. Estimation of formation pressures from log-derived shale properties[J]. *Journal of Petroleum Technology*, 1965, 17(06): 717–722.
- [9] EATON B A. The effect of overburden stress on geopressure prediction from well logs[J]. *Journal of Petroleum Technology*, 1972, 24(08): 929–934.
- [10] EATON B A. The equation for geopressure prediction from well logs[C]//Fall meeting of the Society of Petroleum Engineers of AIME. *OnePetro*, 1975, 5554: 1–5.
- [11] TERZAGHI K, PECK R B, MESRI G. *Soil mechanics in engineering practice*[M]. John Wiley & Sons, 1996.
- [12] 吴建忠, 王波, 李强, 等. 随钻地层压力监测技术在川西地区的应用[J]. *石油地质与工程*, 2009, 23(04): 82–84. [WU J Z, WANG B, LI Q, et al. Application of formation pressure monitoring technology while drilling in western Sichuan[J]. *Petroleum Geology and Engineering*, 2009, 23(04): 82–84.]
- [13] 付明川. dc指数校正研究[J]. *录井工程*, 2020, 31(03): 8–13. [FU M C. Dc index correction study[J]. *Mud Logging Engineering*, 2020, 31(03): 8–13.]
- [14] PRIMASTY A Q T, AKBAR K A. Detecting overpressure using the Eaton and Equivalent Depth methods in Offshore Nova Scotia, Canada[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2018, 132(1): 012016.
- [15] BEKTAS E, MISKA S Z, OZBAYOGLU E M, et al. Application of Kalman filter to predictions of pore pressure while drilling[C]//SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, 2015.
- [16] ZIEGLER F E, JONES J F. Pre-drill pore-pressure prediction and pore pressure and fluid loss monitoring during drilling: A case study for a deepwater subsalt Gulf of Mexico well and discussion on fracture gradient, fluid losses, and wellbore breathing[J]. *Interpretation*, 2014, 2(1): SB45–SB55.
- [17] 李中, 刘和兴, 李磊, 等. 基于改进的Bowers法预测南海陵水深水高温地层异常压力[J]. *西安石油大学学报(自然科学版)*, 2019, 34(06): 60–66. [LI Z, LIU H X, LI L, et al. Prediction of abnormal pressure in Lingshui Deep-water high temperature formation based on improved Bowers method[J]. *Journal of Xi'an Shiyou University(Natural Science Edition)*, 2019, 34(06): 60–66.]
- [18] 陶磊, 赵越哲, 何岩峰, 等. 一种碳酸盐岩气藏地层孔隙压力计算方法[J]. *常州大学学报(自然科学版)*, 2019, 31(03): 88–92. [TAO L, ZHAO Y Z, HE Y F, et al. A calculation method of pore pressure in carbonate gas reservoir[J]. *Journal of Changzhou University(Natural Science Edition)*, 2019, 31(03): 88–92.]
- [19] OGBAMIKHUMI A, EBENIRO J O. Reservoir properties estimation from 3D seismic data in the Alose field using artificial intelligence[J]. *Journal of Petroleum Exploration and Production*, 2021, 11(3): 1275–1287.
- [20] YAN X S, ZHANG M Z, WU Q H. Big-data-driven pre-stack seismic intelligent inversion[J]. *Information Sciences*, 2021, 549: 34–52.
- [21] 邹文波. 人工智能研究现状及其在测井领域的应用[J]. *测井技术*, 2020, 44(04): 323–328. [ZOU W B. Artificial intelligence research status and applications in well logging[J]. *Well Logging Technology*, 2020, 44(04): 323–328.]
- [22] 廖广志, 李远征, 肖立志, 等. 利用卷积神经网络模型预测致密储层微观孔隙结构[J]. *石油科学通报*, 2020, 5(01): 26–38. [LIAO G Z, LI Y Z, XIAO L Z, et al. Prediction of microscopic pore structure of tight reservoirs using convolutional neural network model[J]. *Petroleum Science Bulletin*, 2020, 5(01): 26–38.]
- [23] CAJA M Á, PEÑA A C, CAMPOS J R, et al. Image processing and machine learning applied to lithology identification, classification and quantification of thin section cutting samples[C]//SPE Annual Technical Conference and Exhibition. *OnePetro*, 2019.
- [24] NOSHI C I, EISSA M R, ABDALLA R M. An intelligent data driven approach for production prediction[C]//Offshore Technology Conference. *OnePetro*, 2019.
- [25] 薛亮, 吴雨娟, 刘倩君, 等. 裂缝性油气藏数值模拟与自动历史拟合研究进展[J]. *石油科学通报*, 2019, 4(04): 335–346. [XUE L, WU Y J, LIU Q J, et al. Advances in numerical simulation and automatic history matching of fractured reservoirs[J]. *Petroleum Science*

- Bulletin, 2019, 4(04):335-346.]
- [26] 张瑞, 贾虎. 基于多变量时间序列及向量自回归机器学习模型的水驱油藏产量预测方法[J]. 石油勘探与开发, 2021, 48(01): 175-184. [ZHANG R, JIA H. Production performance forecasting method based on multivariate time series and vector autoregressive machine learning model for waterflooding reservoirs[J]. Petroleum Exploration and Development, 2021, 48(01): 175-184.]
- [27] JIA D, GUO T, PEI X, et al. Intelligent waterflooding development of high-permeability reservoirs at the late development stage[C]// SPE Asia Pacific Oil and Gas Conference and Exhibition. OnePetro, 2018.
- [28] 张凯, 赵兴刚, 张黎明, 等. 智能油田开发中的大数据及智能优化理论和方法研究现状及展望[J]. 中国石油大学学报(自然科学版), 2020, 44(04): 28-38. [ZHANG K, ZHAO X G, ZHANG L M, et al. Current status and prospect for the research and application of big data and intelligent optimization methods in oilfield development[J]. Journal of China University of Petroleum (Edition of Natural Sciences), 2020, 44(04): 28-38.]
- [29] MCCULLOH W S, PITTS W. A logical calculus of the ideas immanent in neural nets[J]. Bull Math. Biophys, 1943, 5: 133-137.
- [30] DE RUMELHART, HINTON G E, WILLIAMS R J. Learning representations by back propagating errors[J]. Nature, 1986, 323(6088): 533-536.
- [31] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(S2): 1-6. [YANG L, WU Y X, WANG J L, et al. Research on recurrent neural network[J]. Journal of Computer Applications, 2018, 38(S2): 1-6.]
- [32] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE transactions on neural networks, 1994, 5(2): 157-166.
- [33] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [34] GRAVES A, SCHMIDHUBER J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. Neural Networks, 2005, 18(5-6): 602-610.
- [35] 宋亚斌, 邢元军, 江腾宇, 等. 基于距离相关系数和KNN回归模型的森林蓄积量估测研究[J]. 中南林业科技大学学报, 2020, 40(04): 22-27. [SONG Y W, XING Y J, JIANG T Y, et al. Forestry volume estimation based on distance coefficient and KNN regression model[J]. Journal of Central South University of Forestry & Technology, 2020, 40(04): 22-27.]
- [36] FINLEY A O, MCROBERTS R E. Efficient k-nearest neighbor searches for multi-source forest attribute mapping[J]. Remote Sensing of Environment, 2008, 112(5): 2203-2211.
- [37] 武中原, 张欣, 张春雷, 等. 基于 LSTM 循环神经网络的岩性识别方法[J]. 岩性油气藏, 2021, 33(03): 120-128. [WU Z Y, ZHANG X, ZHANG C L, et al. Lithology identification based on LSTM recurrent neural network[J]. Lithologic Reservoirs, 2021, 33(03): 120-128.]

(编辑 付娟娟)