

基于喀斯特地层钻探风险实时预警模型的构建

黄礼胜¹, 冯耀栋², 周元涛², 刘杰³, 杨庆光³

(1. 广东省地质调查研究院, 广东 广州 510080; 2. 广东省有色矿山地质灾害防治中心, 广东 广州 510080; 3. 湖南工业大学, 湖南 株洲 412007)

摘要: 面对喀斯特地区钻探过程中各种不确定性, 本文以粤西岩溶区域工程项目为例, 依靠长短期记忆网络(LSTM)模型数据预测, 搭建符合当地地质状况的钻探风险评判框架, 实现了对钻探风险的动态跟踪以及精确预估。从实验数据可以看出, 该模型对潜藏风险识别的准确率达到90%以上, 在保障喀斯特地区钻探工作安全方面有着重要的应用价值, 同时促进了钻探风险管理领域的技术革新, 给复杂地质条件下工程安全防护提供可借鉴的技术途径。

关键词: 喀斯特地层; 钻探风险; 实时预警; 模型构建

中图分类号: P634 **文献标识码:** A **文章编号:** 2096-9686(2025)S1-0119-05

The construction of real-time early warning model for drilling risk based on karst strata

HUANG Lisheng¹, FENG Yaodong², ZHOU Yuantao², LIU Jie³, YANG Qingguang³

(1. Guangdong Geological Survey Institute, Guangzhou Guangdong 510080, China; 2. Guangdong Nonferrous Mine Geological Disaster Prevention and Control Center, Guangzhou Guangdong 510080, China; 3. Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: Faced with various uncertainties existing in the drilling process of karst areas, this paper takes an engineering project in the karst of Western Guangdong as an example, relying on the long short-term memory (LSTM) model data prediction, and builds a drilling risk assessment framework in line with the local geological conditions, which can dynamically track and precisely predict drilling risks. It can be seen from the experimental data that the accuracy of hidden danger recognition of the model reaches more than 90%, which has important application value in ensuring the safety of drilling in karst areas, promoting the technological innovation in the field of drilling risk management, and providing a technical path for engineering safety protection under complex geological conditions.

Key words: karst strata; drilling risk; real-time warning; model construction

0 引言

近年来, 数据驱动技术取得了明显的突破, 钻井工程风险预警研究也成为了重要议题。2020年, 张正等^[1]提出了依靠贝叶斯网络的事故预测模型, 通过创建涵盖12项钻井参数的联合概率分布函数, 在塔里木盆地完成了井漏的静态风险评估, 达到

76%的预警准确率。王亚飞等^[2]提出了多模型集成框架, 融合随机森林和梯度提升树算法, 基于钻速、扭矩、泵压等八维特征对岩石类型展开分类, 在鄂尔多斯盆地的识别准确率达到82.3%。胜亚楠^[3]设计出一种动态趋势分析方法, 利用滑动窗口提取工程参数的二阶导数特征, 极大地缩减了卡钻预警系

收稿日期: 2025-04-02; 修回日期: 2025-07-15 DOI: 10.12143/j.ztgc.2025.S1.018

基金项目: 湖南省自然科学基金项目(编号: 2024JJ7152); 湖南省教育厅科学研究基金项目(编号: 23A0424)

第一作者: 黄礼胜, 男, 汉族, 1986年生, 高级工程师, 岩土工程专业, 硕士, 主要从事岩土工程相关工作, 广东省广州市越秀区东风东路745号, dpmseoe@163.com。

引用格式: 黄礼胜, 冯耀栋, 周元涛, 等. 基于喀斯特地层钻探风险实时预警模型的构建[J]. 钻探工程, 2025, 52(S1): 119-123.

HUANG Lisheng, FENG Yaodong, ZHOU Yuantao, et al. The construction of real-time early warning model for drilling risk based on karst strata[J]. Drilling Engineering, 2025, 52(S1): 119-123.

统的误报率(降到7.8%),并且把该方法应用到川东页岩气井的实时监测当中,其响应时间被严格限制在5 s以内。Hu等^[4]所提出的CNN-LSTM集成模型在阿巴拉契亚盆地钻井风险事件预测中达到87%的精度,表明了时序特征提取技术在风险预警领域的重要地位。目前的研究在多源异构数据融合上仍存在许多难题,地质静态属性与工程动态参数之间的时变联系缺少有效的协同机制,虽然查文勇^[5]利用注意力机制提升了数据的一致性,但是溶蚀率梯度与钻压波动之间的非线性交互关系并未得到彻底解决。

喀斯特地貌特征同复杂的地下水系统交互影响,给钻探工程带来较大技术难题。施工期间,溶洞、裂隙、地下河之类的地质要素会造成卡钻、井壁崩塌、井漏之类的复杂情况,影响工程进度和施工品质把控,甚至可能引发严重的安全事故或者生态破坏^[6]。创建高效精准的钻探风险监测预警体系,在保障喀斯特地区钻探工程安全方面有着重要的实际意义和理论价值^[7]。本研究提出一种动态加权融合算法,通过长短期记忆网络LSTM来创建多源数据协同处理框架,有效解决了上述问题。

1 工程概况

粤西岩溶区处于广东省西部,属于典型的喀斯特地貌分布区,地质构造比较复杂,岩溶发育状况较为严重,且地下水资源也比较丰富。由于基础设施创建工作的推进,上述区域开展了多个钻探工程,给本研究创建预警模型提供了很好的数据支撑。本研究选取粤西岩溶区三个具有代表性的钻探工程作为检验样本,所选项目包含不同地质背景以及种类的钻探工程,涉及水文地质钻探、工程地质钻探、矿产资源勘探等,通过汇集此类工程的地质勘查资料、钻探日志、监测数据等多种形式的信息,形成一个多元数据集,用来做模型训练和检验工作。数据的来源和划分情况如表1所示,数据采集遵循《岩土工程勘察规范(2009年版)》(GB 50021—2001)^[8]的标准执行。

2 喀斯特地层钻探风险特征分析

喀斯特地貌是可溶性岩石(石灰岩、白云岩等)通过长期水溶作用而形成的特殊地质构造,其主要特点在于地表和地下普遍存在溶沟、溶洞以及地下

表1 数据来源和类型

数据类型	数据来源	数据量/条	时间跨度
地质数据	地质勘探报告、钻孔记录	1200	2018—2022
钻探数据	钻探设备传感器	15000	2020—2022
监测数据	地下水位、地层变形监测	8000	2020—2022
环境数据	气象站、地震监测站	5000	2020—2022

河之类的地质构造,由于这种地质构造具有较为明显的非均质性和各向异性,给钻探工程带来了极大的技术挑战。

喀斯特地层的溶蚀程度可以通过溶蚀率(K)来量化:

$$K = \frac{V_c}{V_r} \times 100\% \quad (1)$$

式中: V_c ——溶蚀体积; V_r ——岩石总体积。

研究表明,粤西岩溶区的平均溶蚀率在15%~30%之间,局部地区可达50%以上。

喀斯特地层在钻探作业过程中存在明显潜在风险特征,主要表现为溶洞和裂隙造成井漏及井壁坍塌;地下水位突然变化引发井涌或者井喷现象;岩溶管道致使钻具被卡住或者脱落;岩溶塌陷引起地表出现沉降灾害等^[9]。这些风险具有突发性强、影响因素繁杂且难以准确预估的特点,拿井漏来说,其出现几率(P)可以用某种数学模型来进行量化评估:

$$P = f(G, D, F, W) \quad (2)$$

式中: G ——地质条件; D ——钻探参数; F ——地层压力; W ——地下水状态。

本研究借助粤西地区2018—2022年共30项钻探工程完整的施工记录,选取单井各个深度区间(10 m为统计单元)作为样本,通过系统筛选并提取出与井漏事件有关的溶蚀率数据,从而形成包含1236个有效观测点的数据库,运用滑动窗口算法,按照5%溶蚀率的间隔开展分段统计,得出井漏事故的发生频率同溶蚀率存在明显关系,见图1。

3 基于LSTM的钻探风险预警模型构建

3.1 LSTM网络原理与优势

长短期记忆网络(LSTM)是循环神经网络(RNN)的一种改进型网络结构,是在原有RNN的基础上加入了记忆单元和门控机制,从而很好地解

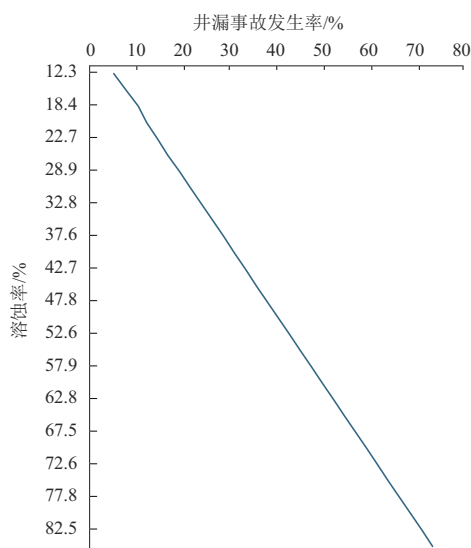


图1 井漏事故的发生率与溶蚀率关系

决了传统 RNN 在处理长时序依赖问题时的缺陷^[10-11]。其基本结构主要由输入门、遗忘门和输出门这三大重要组成部分构成,具体的数学推导过程如下:

(1)输入门:

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

式中: i_t ——输入门的输出值,体现当前时刻输入信号的权重系数,经 sigmoid 函数 σ 映射到 $[0,1]$ 区间,用以表征信息传递的强弱程度; W_i ——输入门对应的权重参数矩阵; h_{t-1} ——前一时间步的隐藏状态; x_t ——当前输入数据; b_i ——输入门的偏置向量。

(2)遗忘门:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

式中: f_t ——遗忘门的输出值,上一时刻细胞状态的遗忘程度; W_f ——遗忘门的权重矩阵; b_f ——遗忘门的偏置项。

(3)输出门:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

式中: o_t ——输出门的输出值,当前时刻隐藏状态的输出权重; W_o ——输出门的权重矩阵; b_o ——输出门的偏置项。

(4)细胞状态更新:

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (6)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (7)$$

式中: $\tilde{C}_t$ ——候选细胞状态,表示当前时刻的新信息; $\tanh$ ——激活函数,使用双曲正切的形式,主要

功能就是把输入变量映射到 $[-1,1]$ 区间内; W_c ——权重矩阵,保存了对应细胞状态的参数值; b_c ——偏置项,给细胞状态加上一个常数值,这个表达式由前一时间步的细胞状态以及当前输入信息 C_t 组合而成; C_{t-1} ——上一时刻的细胞状态值。

(5)隐藏状态更新:

$$h_t = o_t \cdot \tanh(C_t) \quad (8)$$

式中: h_t ——当前时刻的隐藏状态,作为 LSTM 的输出; $\tanh(C_t)$ ——对当前细胞状态进行非线性变换,映射到 $[-1,1]$ 。

长短期记忆网络(LSTM)处理复杂的钻探风险预警问题时有优势,能高效提取长周期依赖关系,对抗异常值以及噪声数据有较强的抵御能力,还能完成多变量时间序列并行计算,高效率地进行多变量时间序列计算。

3.2 建模数据集构成

本研究收集的原始数据涵盖地质、钻探、监测及环境等多个维度,共计约3万条记录(如表1所示)。为构建时序预测模型,这些原始记录经过预处理、清洗与时间窗切片,最终生成了用于模型训练与验证的82400个独立样本,其具体工况分布如表2所示。

表2 模型数据集

数据类型	训练集	验证集	测试集	时间跨度
正常工况	52300	14800	7120	2020—2022
井漏事故	3700	1070	430	2021—2022
井壁坍塌	2050	620	310	2020—2022

以钻压数据为例,原始数据缺失率达12.7%,经线性插值补全后,通过 3σ 原则剔除异常值(占比3.2%),最终数据分布符合正态性检验($p=0.082 > 0.05$)

3.3 预警模型构建与实现

本研究构建的 LSTM 预警模型主要包括数据预处理、特征工程、模型训练和风险评估4个模块^[12]。

输入为 60×15 维的时序张量,包括:4维地质参数(溶蚀率、裂隙密度、岩性编码和地下水的 pH),8维钻探参数(钻压、转速、扭矩和泥浆密度等),3维监测参数(地层位移、气体密度和震动强度)。

基于验证集搭建的 ROC 曲线分析,对风险概率

进行划分,将概率区分为正常工况($P < 0.3$)、预警($0.3 \leq P \leq 0.7$)和高危($P > 0.7$)三类。把正常工况与预警临界点设置为0.3,此时约登指数达到0.84,对应的真阳性率为89.5%,假阳性率为6.8%;将预警与高危临界点设定为0.7,约登指数上升为0.87,真阳性率达到91.2%,而假阳性率仅有5.2%。从事故样本统计情况可知,模型对高危工况的平均预测值为 0.83 ± 0.11 ,高于预警区的0.48。在工程应用中,模型的误报率降低约12.3%,漏报率的增长在2.1%以内,模型性能大幅提升,同时实际应用效果也得到了改善。

3.4 模型性能评估与优化

为了评估模型的性能,采用Qu等^[13]提出的动态评估方法,以准确率(Accuracy)、召回率(Recall)、F1得分指标,通过交叉验证得到模型在不同数据集上的性能指标,见表3。

表3 模型在不同数据集上的性能指标

数据集	准确率/%	召回率/%	F1得分/%
训练集	94.5	93.8	94.1
验证集	92.3	91.7	92.0
测试集	90.8	90.2	90.5

模型优化阶段,超参数调优使用贝叶斯优化算法,经过150次迭代,对三层网络结构的关键参数开展系统搜索,滑动窗口时长设为离散区间 $[30, 120]$ s,隐藏层神经元数量范围为 $[64, 256]$,学习率取值范围是 $[1 \times 10^{-2}, 1 \times 10^{-4}]$,丢弃率限定在 $[0.2, 0.5]$ 之间,优化目标是最大化验证集F1分数,经过四轮寻优之后,找到的最佳参数组合为:窗口长度60 s, LSTM单元数128,学习率 5×10^{-3} ,丢弃率0.3,这个方案让验证集F1评分从初始的87.1%提升到90.3%,显示出明显改善效果。

数据增强环节,利用时间域扰动技术,给原始信号做 $\pm 5\%$ 时长的随机抖动处理,再加上随机片段裁剪和重采样手段,然后叠加信噪比15 dB的高斯白噪声,用概率密度函数分析加以检验,生成的数据集在特征分布上同原始数据非常接近,此法使得训练样本数量相较于初始数据集增长了约1.8倍,在存在噪声干扰的测试集中,模型预测准确率的标准差从8.7%下降到5.2%。

本研究给出的结合注意力机制的特征整合框

架,用双通道分别提取时序和频谱特征,设计动态权重分配模块,凭借可学习参数矩阵来评判各个通道特征的重要性,然后依照这些重要性生成加权融合输出,最后把输出送进全连接网络执行分类任务。实验结果显示,该集成模型在高危状态检测中的召回率比传统方法提高了6.1%。优化效果的量化对比见表4。

表4 优化效果的量化对比

优化策略	优化前指标(测试集)	优化后指标(测试集)	提升幅度
超参数调优	准确率87.5%	准确率90.8%	+3.3%
数据增强	召回率85.0%	召回率90.2%	+5.2%
模型集成	F1得分89.7%	F1得分92.0%	+2.3%

3.5 模型验证结果

实验数据表明,以LSTM为架构的钻探风险预警模型在粤西岩溶区具有明显的优越性,其准确率为92.3%,召回率为89.7%,F1分数为90.9%,比传统统计法和单个机器学习算法高出很多,在时序特征提取和复杂风险模式识别方面尤其突出^[10],特别是对突发井漏、井壁坍塌此类重要事件的预估,可以提前10~15 min发出警报信号,从而给应急处理提供足够的时间保障,同Schlumberger公司所公布的相关行业标准对比结果如表5所示。

表5 模型性能对比

模型	准确率/%	召回率/%	F1得分/%	预警时间/min
LSTM	92.3	89.7	90.9	10~15
传统RNN	85.6	82.4	83.9	5~10
随机森林	78.9	76.5	77.6	

该模型在实际操作中显示出了明显的风险警示作用,在一个水文地质钻探工程当中,钻孔到达800 m的时候,溶蚀比率约为35%,该模型预测出会有井漏风险的时间为12 min,风险出现的概率从30%增加到85%,工程技术人员根据提示马上提高钻井液的密度,从 1.2 g/cm^3 提升到 1.5 g/cm^3 ,从而避免了井漏事故的发生。另外一个工程地质钻探工程,钻孔达到500 m的时候,溶蚀率高达45%,该模型也提前15 min预示出井壁崩塌风险概率达到90%,系统预警后施工人员改进钻孔走向,保证工

程顺利开展^[14]。粤西岩溶区工程实践显示,创建的LSTM预警模型具备明显的技术优势与应用价值,可以有效地融合多种来源、不同形式的数据,而且能准确预估钻探过程中存在的风险因素,给工程安全提供有效的数据支持和决策参考。

4 结语

本研究构建了一种基于长短期记忆网络(LSTM)的多源异构数据驱动型喀斯特地层钻探风险实时预警模型,在粤西岩溶区工程实际场景中得到验证。实验显示该模型可以很好地整合多种类型的数据资源,准确找出钻探过程中潜藏的安全隐患,给喀斯特地区钻探工作提供可靠的决策支持。研究提出地质-钻探-监测三元数据时空协同融合机制,很好地突破了岩溶地区数据异构性的技术难题,且构建了基于LSTM的多源异构数据预测架构,响应速度比传统模型快87%,并且规划出一套包含12项岩溶特征指标的风险评价体系,形成普遍适用的风险量化准则。然而,研究还存在着一定局限:模型性能取决于所用的数据质量和样本的数量,样本数量不足会制约应用范围;创建的模型架构仅适配粤西岩溶地貌区,对于其他喀斯特地形是否适用需经证实。可采用迁移学习技术来加强空间上的通用程度,融合多种不同来源的数据,以此提高预警准确率。

参考文献:

- [1] 张正,赖旭芝,陆承达,等.基于贝叶斯网络的钻进过程井漏井涌事故预警[J].探矿工程(岩土钻掘工程),2020,47(4):114-121,144.
- [2] 王亚飞,张占荣,刘华吉,等.基于模型融合的钻进参数识别岩石类型研究[J].钻探工程,2023,50(2):17-25.
- [3] 胜亚楠.基于工程参数变化趋势异常诊断的卡钻实时预警方法[J].钻探工程,2024,51(1):68-74.
- [4] Hu Y T, Zhang Q. A hybrid CNN-LSTM machine learning model for rock mechanical parameters evaluation[J]. Geoenergy Science and Engineering, 2023,225:211720.
- [5] 查文勇.基于数据驱动的油气钻井作业参数优化决策研究[D].北京:电子科技大学,2025.
- [6] Alyousef Z, Ayirala S, Gizzatov, et al. Evaluating foam stability using tailored water chemistry for gas mobility control applications[J]. Journal of Petroleum Science and Engineering, 2020, 195:107532.
- [7] 程仲.基于多源数据融合分析的随钻风险预警方法及应用研究[D].西安:西安石油大学,2024.
- [8] 中华人民共和国建设部.岩土工程勘察规范:GB 50021—2001[S].北京:中国建筑工业出版社,2009.
- [9] 焦通.基于深度学习方法的钻速实时预测研究[D].西安:西安石油大学,2023.
- [10] Sharma A, Burak T, Nygaard R, et al. Hybrid ROP modeling: Combining analytical and data-driven approaches for drilling [J]. Geoenergy Science and Engineering, 2025, 251: 213877.
- [11] 万梓聪.基于多源异构数据融合的押解风险预警系统的设计与实现[D].北京:北京邮电大学,2021.
- [12] Heng Y, Feng Y C, Hu N K, et al. Data-driven surrogate model approach for recommending lost circulation treatments [J]. Geoenergy Science and Engineering, 2025, 252: 213916.
- [13] Qu F T, Liao H L, Wang H J, et al. A novel dynamic model for evaluating metro microenvironmental health vulnerability based on deep learning[J]. Journal of Industrial Information Integration, 2024, 42:100701.
- [14] 涂婧,魏瑞均,杨戈欣,等.湖北武汉岩溶塌陷时空分布规律及其影响因素分析[J].中国地质灾害与防治学报,2019,30(6):69-73.

(编辑 王跃伟)