

“深部探测钻探技术装备专题”编者按:地球深部探测国家重大项目立项实施正在加快推进,深部探测技术和装备研发持续开展,致力形成我国万米空间入地能力。为此,山东省地矿局钻探工程技术研究中心与编辑部联合组织了“深部探测钻探技术装备专题”,由翟育峰高级工程师担任客座编辑。本专题遴选6篇论文发表,内容涵盖深部钻探技术与机具研究、深部钻探工程实践等内容,以期推动深部钻探技术的发展,服务地球深部探测工程。

基于区域多井数据优选与模型预训练的深部地质钻探过程钻速动态预测方法

甘超^{1,2,3,5}, 汪祥^{1,2,3}, 王鲁朝^{4,5}, 曹卫华^{*1,2,3}, 吴敏^{1,2,3}

(1. 中国地质大学(武汉)自动化学院,湖北武汉430074;

2. 复杂系统先进控制与智能自动化湖北省重点实验室,湖北武汉430074;

3. 地球探测智能化技术教育部工程研究中心,湖北武汉430074;

4. 山东省第三地质矿产勘查院,山东烟台264004; 5. 山东省地矿局钻探工程技术研究中心,山东烟台264004)

摘要:深部地质钻探过程钻速精准预测有助于提升钻探效率、降低钻探成本,可为安全高效的深部地质钻探施工提供关键技术支撑。本文提出了一种基于区域多井数据优选与模型预训练的深部地质钻探过程钻速动态预测方法。首先,选取岩性识别软件、钻进过程智能监控云平台、地质云系统等作为数据源,在此基础上设计深部地质钻探数据仓库。其次,运用区域多井数据优选技术在数据仓库中选择与目标井较匹配的数据,并开展钻速模型预训练。最后,结合深部地质钻探过程实钻数据,引入小波滤波、超限学习机、增量学习等技术,实现钻速预测模型动态更新。实验对比结果验证了所提方法具有很强的钻速预测性能与泛化能力。

关键词:深部地质钻探;区域多井数据优选;模型预训练;钻速动态预测;钻速模型更新

中图分类号:P634 文献标识码:A 文章编号:2096-9686(2023)04-0001-08

Dynamic prediction method of rate of penetration (ROP) in deep geological drilling process based on regional multi-well data optimization and model pre-training

GAN Chao^{1,2,3,5}, WANG Xiang^{1,2,3}, WANG Luzhao^{4,5}, CAO Weihua^{*1,2,3}, WU Min^{1,2,3}

(1. School of Automation, China University of Geosciences, Wuhan Hubei 430074, China;

2. Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan Hubei 430074, China;

3. Engineering Research Center of Intelligent Technology for Geo-Exploration, Ministry of Education, Wuhan Hubei 430074, China;

4. Shandong No.3 Exploration Institute of Geology and Mineral Resources, Yantai Shandong 264004, China;

收稿日期:2023-05-21; 修回日期:2023-06-24 DOI:10.12143/j.ztgc.2023.04.001

基金项目:国家自然科学基金青年项目“基于多源井震信息融合的地质钻进过程钻速智能优化”(编号:62003318);中央高校基本科研业务费专项资金科研项目“考虑复杂地质环境的钻进过程钻速优化”(编号:CUG2106350)

第一作者:甘超,男,汉族,1990年生,副教授,硕士生导师,主要从事复杂地质钻进过程建模与优化控制的研究工作,湖北省武汉市鲁磨路388号,ganchao@cug.edu.cn。

通信作者:曹卫华,男,汉族,1972年生,中国地质大学(武汉)自动化学院院长,教授,博士生导师,主要从事过程控制、智能系统和机器人技术的研究工作,湖北省武汉市鲁磨路388号,weihuacao@cug.edu.cn。

引用格式:甘超,汪祥,王鲁朝,等.基于区域多井数据优选与模型预训练的深部地质钻探过程钻速动态预测方法[J].钻探工程,2023,50(4):1-8.

GAN Chao, WANG Xiang, WANG Luzhao, et al. Dynamic prediction method of rate of penetration (ROP) in deep geological drilling process based on regional multi-well data optimization and model pre-training[J]. Drilling Engineering, 2023, 50(4): 1-8.

5. *Drilling Engineering Technology Research Center of Shandong Provincial Bureau of Geology & Mineral Resources, Yantai Shandong 264004, China*

Abstract: Accurate prediction of rate of penetration (ROP) in deep geological drilling process can help to improve drilling efficiency and reduce drilling costs, which can provide key technical support for safety and efficient deep geological drilling construction. In this paper, a dynamic prediction method of ROP in deep geological drilling process based on regional multi-well data optimization and model pre-training is proposed. First, the deep geological drilling data warehouse is designed by selecting lithology identification software, drilling process intelligent monitoring cloud platform, and geological cloud system as data sources. Secondly, the regional multi-well data optimization technique is used to select the matching data with the target well in the data warehouse, and the ROP model is pre-trained on this basis. Finally, the ROP prediction model is dynamically updated through combining the actual drilling data of deep geological drilling process, and introducing techniques such as the wavelet filtering, extreme learning machine, and incremental learning strategy. The experimental comparison results verify that the proposed method has strong ROP prediction performance and generalization capability.

Key words: deep geological drilling; regional multi-well data optimization; model pre-training; rate of penetration dynamic prediction; rate of penetration model update

0 引言

深部地质钻探过程通常采用多种钻探装备从地表钻至目标地层,以获取地下埋藏资源能源。整个过程会遭遇各种突变地层,导致钻探效率受到严重影响^[1-2]。钻速是评估钻探效率的核心因素,通过建立准确的钻速预测模型并在此基础上优化钻探参数,有助于缩减工程周期与钻探成本,为钻探过程智能优化控制奠定重要基础^[3]。钻速预测一直是地质钻探领域的重点研究问题之一,主要可以将其分为两类:第一类是基于钻速离线模型开展预测,以考虑历史钻探数据建模的静态预测为主;第二类是基于钻速在线模型开展预测,可以根据流式钻探数据动态更新模型。

第一类是利用机理分析或机器学习方法进行钻速静态建模并开展预测。其中机理分析是在钻头-岩石作用机制的基础上建立钻速预测模型^[4-6]。针对钻压、转速、岩石抗压强度等参数对钻速的共同影响,DENG Yong等^[7]根据单齿冲击压痕下的岩石破碎机理提出牙轮钻头钻速预测模型,钻探实验结果验证了该方法的有效性,模型平均相对误差低于15%。近年来,学者们将机器学习方法应用于钻速预测研究^[8-10],针对钻探过程中的强非线性、强耦合性、噪声干扰等特征,GAN Chao等^[11]提出基于混合支持向量回归的钻速预测方法:首先,利用小波滤波对钻探数据进行预处理,提升数据的整体质量。其次,运用相关性分析对不同钻探参数进行特征选择,降低建模过程的复杂性。最后,引入改进蝙蝠算法

优化的支持向量回归建立钻速模型,实验结果表明所提方法具有很高的预测精度。此类方法主要基于已知钻探数据进行离线钻速建模,在针对特定钻头或岩性变化不大的地层时具有优势,但在面对复杂地层环境时难以有效预测钻速。

第二类是根据实时钻探大数据对模型进行更新并开展预测^[12-14]。挪威斯塔万格大学SUI Dan等^[15]引入滚动时域估计方法建立了一种带有状态空间表达式的钻速动态预测模型,实验结果表明该方法可以很好的预测钻速变化趋势。考虑到钻探过程存在机理复杂、非线性强等特点,ZHOU Yang等^[16]采用非支配排序遗传算法与混合蝙蝠算法^[17]优化的支持向量回归建立钻速模型,并且引入在线学习策略更新模型,实验表明该方法可以取得较好的预测效果。另外,针对钻探过程中地层变化剧烈、参数耦合严重等问题,GAN Chao等^[18]提出一种新型的钻速动态预测方法:首先,针对8个钻探参数与钻速进行相关性分析,选取井深、钻压、转速作为模型输入。其次,利用重采样技术与小波滤波方法进行钻探数据预处理。最后,引入超限学习机与滑动窗口^[19]建立动态钻速模型。与多种钻速预测模型的对比结果验证了该方法的有效性,模型预测精度至少提升7%。但是上述方法大多只考虑单井信息进行钻速预测,未考虑区域多井信息,预测模型的泛化性能有待提高。

针对上述问题,本文在现有钻速在线模型的基础上开展相关钻速预测研究,提出了一种基于区域

多井数据优选与模型预训练的深部地质钻探过程钻速动态预测方法,具体内容如下所述。

1 深部地质钻探过程特点分析

在深部地质钻探过程中,现场工人可以在司钻房中实时监测钻探参数,并操作人机交互屏/按钮下发指令给可编辑逻辑控制器,所提方法输出的钻速预测值将为司钻员提供指导。现场工人需要在钻探平台上开展相关作业,例如接单根、下套管等任务。同时,由于深部地质钻探的复杂性与多变性,该过程存在参数来源不同、钻探数据价值密度低、多参数非线性强和地层岩性变化剧烈等特点。

1.1 参数来源不同

大多数井场会采集丰富的过程信息,按照不同来源可以分为地层参数、钻探参数和区域多完钻井信息。地层参数可以利用岩性识别软件进行辨识,能很好的描述地层变化情况,包括地层可钻性、单轴抗压强度、地层压力等参数;钻探参数是通过钻进过程智能监控云平台^[20]采集到的信息,可以直观反映当前钻探状态,主要包括扭矩、钻压、转速、钻速等参数;区域多完钻井信息是通过地质云系统^[21]获取到的信息,可以概述某一区域的井场情况与地质环境,主要包括井身柱状图、地层岩心、历史钻探数据等。

1.2 钻探数据价值密度低

深部地质钻探过程存在高温、高压、高地应力等复杂地质力学环境,钻探数据的价值密度与整体质量偏低,存在大量的异常样本点需要处理。除此之外,钻探过程时常发生粘滑振动、钻头跳钻、卡钻、断钻具等情况,导致钻探数据会受到不同频率的噪声干扰,数据的信噪比程度相对较低。

1.3 多参数非线性强

钻压、转速、钻速等钻探参数之间存在较强的非线性关系。在理想状态下,钻速会随钻压或转速的上升而增加,但受地层环境、钻头寿命等因素的影响和制约,导致钻压或转速在提升到一定程度后,钻速会不增反降。同时,在面对复杂多变的地质环境时,地层特征参数(尤其是地层可钻性)与上述钻探参数之间也存在较强的非线性关系^[22]。

1.4 地层岩性变化剧烈

实际钻探过程会穿越多种非平稳地层,例如软硬交替层、岩石破碎地带、孔隙环境、裂隙环境等。以中国东北地区某井为例,该井场属于特深孔取心

钻探,地层岩性变化大,先后钻遇到蚀变破碎带、矿区废旧坑道、卵砾石回填料等突变地层,导致现场工人难以有效克服复杂地层引发的钻速突变问题。同时,不同钻探现场在地层方面也会存在很大的差异性,现有钻速预测研究更适用于单一井场,难以将其应用在其它井场。

2 钻速在线预测模型框架

所提钻速在线预测模型框架如图1所示。整体框架包括核心数据源选取、深部地质钻探数据仓库设计、区域多井数据优选、钻速模型预训练、钻速模型动态更新等5个部分。

2.1 核心数据源选取与深部地质钻探数据仓库设计

由1.1节分析可以发现,岩性识别软件、钻进过程智能监控云平台、地质云系统可以采集到大量的地层信息、钻探信息和区域多完钻井信息,因而可以将上述软件或系统作为核心数据源。同时,在核心数据源的基础上设计一种综合性的深部地质钻探数据仓库,具体包括目标井样本库、完钻井样本库、数据选取样本库和模型参数库,使所提钻速预测模型可以快速提取相关信息。

2.2 区域多井数据优选

由1.1节与2.1节分析可知,深部地质钻探数据仓库包含区域多井的各类信息,涉及面广且信息丰富,本文引入钻探工程师的专家经验对不同井场进行全面分析,确定同目标井匹配度最高的完钻井。具体而言,从深部地质钻探数据仓库中提取模型参数、完钻井过程信息、完钻井地层岩性信息、完钻井钻探数据、目标井过程信息、目标井基础地质资料等多源异构信息,利用区域多井数据优选技术对不同完钻井与目标井进行特性匹配。通过分析不同井场在钻探区域、过程特点、地层环境等方面的相同点与不同点,确定区域完钻井数据优选结果(具体包括完钻井钻探、完钻井地层和模型参数优选结果),为后续钻速模型预训练提供关键信息。

2.3 钻速模型预训练

由1.2节分析可知,钻探过程存在传感器测量误差、传输检测异常、高低频噪声干扰等问题,导致钻探数据的价值密度较低,难以为后续研究提供有力支撑。针对上述问题,本文在确定区域完钻井数据优选结果后,运用两种预处理技术对原始钻探数

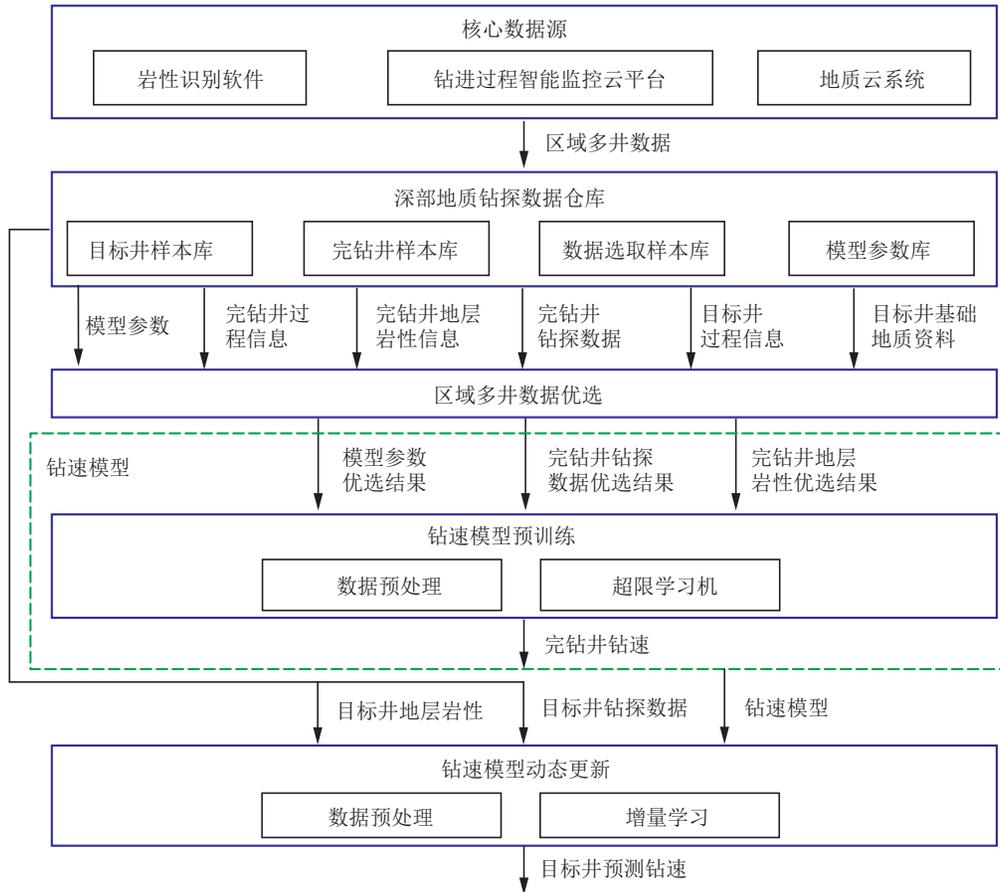


图1 所提钻速在线预测模型框架

Fig.1 Proposed model framework for ROP online prediction

据进行清洗。在第一种数据预处理技术中,根据专家经验确定钻探核心参数(扭矩、钻压、转速、钻速)的合理波动范围,并且将超出范围的异常样本点剔除。在第二种数据预处理技术中,根据钻探数据的分布特征设置小波滤波参数的阈值,针对筛选后的钻探数据进行降噪处理,提高钻探数据的价值密度。

由1.3节分析可知,钻探参数与地层参数之间呈现强烈的非线性关系,导致难以建立准确的钻速预测模型。针对上述问题,本文在预处理后的完钻井数据优选结果基础上,引入具有泛化性强、学习速度快的超限学习机算法进行钻速模型预训练,以提升所提方法的预测精度。同时,利用十折交叉验证方法对钻速预测模型中的关键参数做进一步优化,选取相对最优的模型结构及参数。

2.4 钻速模型动态更新

由1.4节分析可知,深部地质钻探过程地层复杂,岩性变化剧烈,较难感知和捕捉钻速的动态变化^[18]。此外,现有钻速在线预测研究以滑动窗口式

更新为主,在学习新钻探信息的同时会丢失一部分已知的钻探信息,可能会使区域多井数据优选阶段(图1)输出的完钻井数据优选结果没有得到有效利用。针对上述问题,将钻探时间间隔与地层岩性变化同时作为钻速预测模型的更新条件,引入增量式在线学习策略(递推最小二乘)对模型参数进行动态更新。该算法可以根据上一时刻的模型参数与这一时刻的过程信息不断更新模型参数,其计算公式为:

$$P_{k+1} = P_k - P_k H_{k+1}^T (I + H_{k+1} P_k H_{k+1}^T)^{-1} H_{k+1} P_k \quad (1)$$

式中: P_k ——中间参数矩阵, $P_k = (H_k^T H_k)^{-1}$; H_k —— k 时刻的输入参数矩阵; H_{k+1} —— $k+1$ 时刻的输入参数矩阵; H_{k+1}^T —— $k+1$ 时刻的输入参数矩阵的转置。

$$\beta_{k+1} = \beta_k + P_{k+1} H_{k+1}^T (S_{k+1} - H_{k+1} \beta_k) \quad (2)$$

式中: β_k —— k 时刻的模型参数; β_{k+1} —— $k+1$ 时刻的模型参数; S_{k+1} —— $k+1$ 时刻的输出参数。

通过递推最小二乘算法可以有效实现所提模型对多完钻井数据优选结果与目标井信息的整合提炼。

3 区域多井数据优选与预处理实验

从已建立的深部地质钻探数据仓库中选取河北保定地区某井、黑龙江安达地区某井和湖北襄阳地区某井作为完钻井,辽宁丹东地区某井作为目标井,

利用区域多井数据优选技术对不同完钻井与目标井进行适配性分析,输出完钻井数据优选结果。同时,针对完钻井数据优选结果进行预处理。

3.1 区域多井数据优选实验

从钻探类型、钻探区域、测量参数、过程特点、地层环境、钻探效率等方面对区域不同完钻井与目标井进行匹配优选,具体如表1所示。

表1 区域多完钻井与目标井的过程信息

Table 1 Process information of regional multi-completion wells and target well

过程信息	河北保定地区某井 (完钻井)	黑龙江安达地区某井 (完钻井)	湖北襄阳地区某井 (完钻井)	辽宁丹东地区某井 (目标井)
钻探类型	特深孔取心钻探	特深孔取心钻探	深孔全面钻探	特深孔取心钻探
钻探区域	中国华北地区	中国东北地区	中国华中地区	中国东北地区
测量参数	钻压、泵压、转速、钻速等	钻压、转速、钻速、扭矩等	井深、转速、泵量、钻速等	钻压、转速、扭矩、钻速等
过程特点	钻速非常低、钻压误差大、送钻精度差、深部取心艰难	钻进速度慢、钻头磨损严重、起下钻耗时、跳钻现象突出	钻速波动大、钻头更换频繁、测量误差大、取心工序繁杂	钻速较缓慢、钻头磨损严重、传感器检测异常、设备冲击性能弱
地层环境	地层岩性多变,多为硅质白云岩	地层坚硬复杂,岩石类型丰富	以灰岩、页岩和白云岩为主	岩石破碎带,大倾角地层
钻探效率	低	低	不高	低
钻探深度	4000 m	7000 m	2000 m	3000 m

由表1可以看到,湖北襄阳地区某井(处在中国华中地区)属于深孔全面钻探,存在钻速波动大、钻头更换频繁、测量误差大、取心工序繁杂等特点,在钻探类型、钻探区域、过程特点、地层岩性等方面与辽宁丹东地区某井的匹配度较低。同时,河北保定地区某井位于中国华北地区,相比于其它两口井(黑龙江安达地区某井、辽宁丹东地区某井)具有一定的区域差异性,存在钻速非常低、钻压误差大、送钻精度差等特征。另外,黑龙江安达地区某井与辽宁丹东地区某井都属于特深孔取心钻探且处在中国东北地区,存在钻速低、钻头磨损严重、地层复杂等特点,在钻探类型、测量参数、地层环境等方面高度相似。结合专家经验,综合考虑优选黑龙江安达地区某井作为最佳完钻井。

3.2 区域多井数据预处理实验

由于辽宁丹东地区某井(目标井)与黑龙江安达地区某井(完钻井)都可以实时采集到扭矩、钻压、转速、钻速4个钻探参数,并且扭矩、钻压、转速与钻速存在强烈的非线性关系,于是本文选取黑龙江安达

地区某井4244组钻探数据作为样本总集,数据分布如表2所示。可以发现,扭矩、钻压、钻速存在一些异常样本点,例如扭矩的最大值和标准差分别超过23000 N·m和3100 N·m、钻压的最小值低到-426.9 kN、钻速的最大值甚至高达13.1 cm/min,其主要原因可能是钻进现场的参数传感器出现失灵、检测异常、测量误差大等问题。结合2.3节中的异常值剔除思想,确定了黑龙江安达地区某井的钻探实测数据设定范围,具体如表3所示,最终选出500组真实有效的钻探数据。

表2 黑龙江安达地区某井钻探数据分布

Table 2 Distribution of drilling data from a well in Anda area, Heilongjiang

参数	扭矩/ (N·m)	钻压/ kN	转速/ (r·min ⁻¹)	钻速/ (cm·min ⁻¹)
最小值	0	-426.9	0	0
平均值	2326.1	179.3	19.2	0.3
最大值	23500	1072.1	80	13.1
标准差	3163.7	357.1	26.9	1.03

表3 黑龙江安达地区某井的钻探实测数据设定范围

参数	扭矩/(N·m)	钻压/ kN	转速/ (r·min ⁻¹)	钻速/ (cm·min ⁻¹)
范围	0<扭矩≤12000	>0	≥10	0<钻速≤8.3

在已筛选的500组钻探数据基础上运用小波滤波分析技术对其进行过滤,过滤后的黑龙江安达地区某井钻探数据分布如表4所示。

表4 过滤后的黑龙江安达地区某井钻探数据分布

参数	扭矩/ (N·m)	钻压/ kN	转速/ (r·min ⁻¹)	钻速/ (cm·min ⁻¹)
最小值	4521.2	6.3	29.2	0.3
平均值	6646.7	49.5	59.8	2.3
最大值	11826	149.6	80.8	8.1
标准差	1644.9	19.8	13.9	1.9

由表4可以看到,扭矩、钻压、转速、钻速4个钻探核心参数中的异常样本点已全部去除,并且在一定程度上将噪声干扰对钻探数据的影响降到最低,验证了异常值剔除与小波滤波分析技术的有效性和重要性。

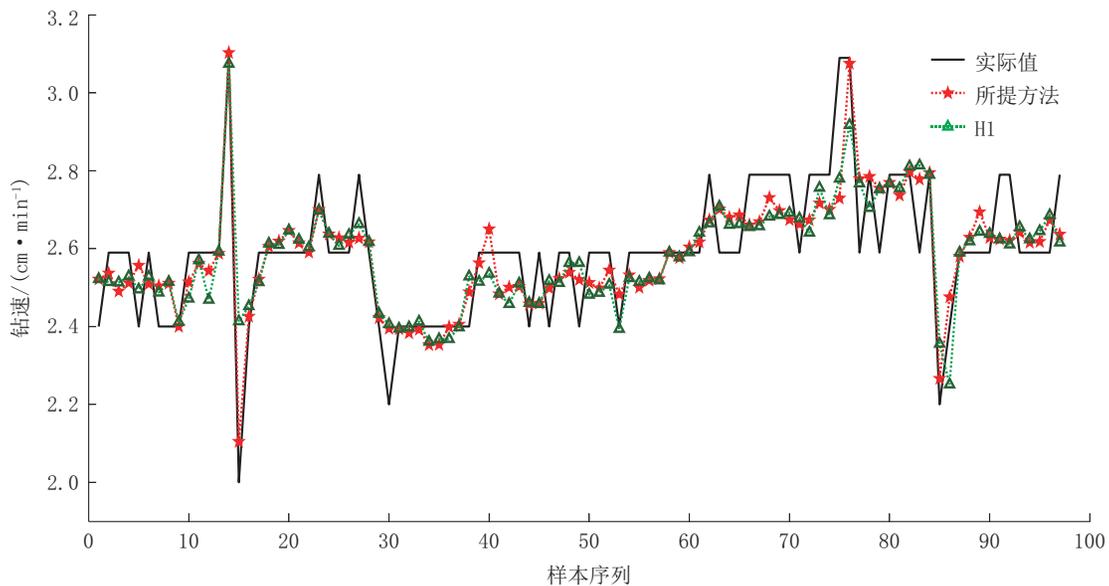


图2 所提方法与钻速在线对比方法的预测曲线

Fig.2 Prediction curves of the proposed method and ROP online comparison method

4 钻速预测对比验证实验

根据上面的实验结果,选取了辽宁丹东地区某井(目标井)、黑龙江安达地区某井(完钻井)进行钻速预测实验。引入2种模型性能评价指标,其计算公式定义为:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

$$NRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}}{\frac{1}{n} \sum_{i=1}^n y_i} \times 100\% \quad (4)$$

式中:RMSE——均方根误差;NRMSE——归一化均方根误差;n——样本数量;y_i——测量数据; \hat{y}_i ——预测数据。

均方根误差与归一化均方根误差越小,预测钻速与实际钻速的拟合效果越好,说明钻速模型的预测能力越强。同时,为测试所提模型的性能,在该模型的基础上设计了一种钻速在线预测对比方法(不包含黑龙江安达地区某井数据),简称为H1。

图2展示了所提方法与钻速在线对比方法的预测曲线。可以看到,所提方法的预测曲线可以很好地捕捉实际钻速变化趋势,相比于H1具有很好的实时感知和预测能力。同时,所提方法与钻速在线对比方法的预测误差如表5所示。

表5 所提方法与钻速在线对比方法的预测误差
Table 5 Prediction error of the proposed method and ROP online comparison method

方法	RMSE	NRMSE/%
所提方法	0.0917	3.54
H1	0.1059	4.09

由表5可以看到,所提方法的均方根误差和归一化均方根误差分别为0.0917和3.54%,相比于H1减少13.4%,说明所提模型在考虑黑龙江安达地区某井数据的情况下,其预测精度明显优于H1,验证了所提模型在区域多井数据优选阶段的有效性,通过区域多井数据优选技术可以很好地提升钻速模型的预测性能。同时,综合对比图2和表5可以发现,所提钻速预测模型还拥有丰富的泛化能力,可以在不同钻探现场为司钻工人提供操作指导,有助于提升实际钻探效率。

5 结论

本文针对深部地质钻探过程钻速预测精度不高、区域多井信息利用率低等问题,综合分析深部地质钻探过程中参数来源不同、钻探数据价值密度低、多参数非线性强、地层岩性变化剧烈等特点,提出了基于区域多井数据优选与模型预训练的深部地质钻探过程钻速动态预测方法。钻速预测精度较对比方法提升了13.4%,验证了所提方法的有效性。主要创新点与贡献如下:

(1)利用区域多井数据优选技术挖掘不同完钻井在地层环境、钻探区域等方面与目标井匹配度高的关键信息,并且通过预训练建立考虑区域多井信息的钻速预测模型。

(2)引入区域多井数据优选、小波滤波、超限学习机、增量学习等技术实现钻速预测模型的动态更新,提升钻速预测模型的普适性与鲁棒性。

(3)所提方法符合钻探工程需要,为深部钻探过程智能优化控制奠定了重要基础。

将来,基于区域多井数据优选与模型预训练的深部地质钻探过程钻速动态预测方法会不断进行优化和完善,为我国深部地质钻探工程安全高效施工提供关键技术支撑。

参考文献(References):

- [1] 甘超.复杂地层可钻性场智能建模与钻速优化[D].武汉:中国地质大学(武汉),2019.
GAN Chao. Intelligent modeling of formation drillability field and drilling rate of penetration optimization in complex conditions [D]. Wuhan: China University of Geosciences, 2019.
- [2] 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.
ZHANG Zheng, LAI Xuzhi, LU Chengda, et al. Lost circulation and kick accidents warning based on Bayesian network for the drilling process [J]. Exploration Engineering (Rock & Soil Drilling and Tunneling), 2020, 47(4): 114-121, 144.
- [3] 吴敏,曹卫华,陈鑫,等.复杂地质钻进过程智能控制[M].北京:科学出版社,2022.
WU Min, CAO Weihua, CHEN Xin, et al. Complex Geological Drilling Process Intelligent Control [M]. Beijing: Science Press, 2022.
- [4] VINICIUS Veloso De Melo, WOLFGANG Banzhaf. Automatic feature engineering for regression models with machine learning: An evolutionary computation and statistics hybrid [J]. Information Sciences, 2018, 430: 287-313.
- [5] HUANG Junping, YAN Xuefeng. Quality-driven principal component analysis combined with kernel least squares for multivariate statistical process monitoring [J]. IEEE Transactions on Control Systems Technology, 2019, 27(6): 2688-2695.
- [6] ZOU Lei, WANG Zidong, HU Jun, et al. Moving horizon estimation with unknown inputs under dynamic quantization effects [J]. IEEE Transactions on Automatic Control, 2020, 65(12): 5368-5375.
- [7] DENG Yong, CHEN Mian, JIN Yan, et al. Theoretical and experimental study on the penetration rate for roller cone bits based on the rock dynamic strength and drilling parameters [J]. Journal of Natural Gas Science and Engineering, 2016, 36: 117-123.
- [8] 李谦,曹彦伟,朱海燕.基于人工智能的钻速预测模型数据有效性下限分析[J].钻探工程,2021,48(3):21-30.
LI Qian, CAO Yanwei, ZHU Haiyan. Discussion on the lower limit of data validity for ROP prediction [J]. Drilling Engineering, 2021, 48(3): 21-30.
- [9] 熊虎林,李谦.基于地层成分和钻进参数的钻速预测模型[J].探矿工程(岩土钻掘工程),2018,45(10):195-201.
XIONG Hulin, LI Qian. ROP prediction model based on formation composition and drilling parameters [J]. Exploration Engineering (Rock & Soil Drilling and Tunneling), 2018, 45(10): 195-201.
- [10] 周长春,姜杰,李谦,等.基于融合特征选择算法的钻速预测模型研究[J].钻探工程,2022,49(4):31-40.
ZHOU Changchun, JIANG Jie, LI Qian, et al. Research on drilling rate prediction model based on fusion feature selection

- algorithm[J]. *Drilling Engineering*, 2022, 49(4):31-40.
- [11] GAN Chao, CAO Weihua, WU Min, et al. Prediction of drilling rate of penetration (ROP) using hybrid support vector regression: A case study on the Shennongjia area, Central China [J]. *Journal of Petroleum Science and Engineering*, 2019, 181: 106200.
- [12] ETESAMI Danial, ZHANG W J, HADIAN Mohsen. A formation-based approach for modeling of rate of penetration for an offshore gas field using artificial neural networks[J]. *Journal of Natural Gas Science and Engineering*, 2021, 95:104104.
- [13] YOUCEFI Mohamed Riad, HADJADJ Ahmed, BENTRIOU Abdelhak, et al. Rate of penetration modeling using hybridization extreme learning machine and whale optimization algorithm [J]. *Earth Science Informatics*, 2020, 13(4):1351-1368.
- [14] SOARES Cesar, GRAY Kenneth. Real-time predictive capabilities of analytical and machine learning rate of penetration (ROP) models[J]. *Journal of Petroleum Science and Engineering*, 2019, 172:934-959.
- [15] SUI Dan, AADNØY Bernt Sigve. Rate of penetration optimization using moving horizon estimation[J]. *Modeling, Identification and Control: A Norwegian Research Bulletin*, 2016, 37(3): 149-158.
- [16] ZHOU Yang, CHEN Xin, ZHAO Haibin, et al. A novel rate of penetration prediction model with identified condition for the complex geological drilling process[J]. *Journal of Process Control*, 2021, 100:30-40.
- [17] GAN Chao, CAO Weihua, LIU Kangzhi, et al. A new hybrid bat algorithm and its application to the ROP optimization in drilling processes [J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(12):7338-7348.
- [18] GAN Chao, CAO Weihua, LIU Kangzhi, et al. A novel dynamic model for the online prediction of rate of penetration and its industrial application to a drilling process[J]. *Journal of Process Control*, 2022, 109:83-92.
- [19] 甘超,曹卫华,王鲁朝,等. 深部地质钻探过程流式大数据分析 with 动态处理——以辽宁丹东 3000 m 科学钻探工程为例[J]. *钻探工程*, 2022, 49(4):1-7.
- GAN Chao, CAO Weihua, WANG Luchao, et al. Streaming big data analysis and dynamic pre-processing in deep geological drilling process: A case study on the 3000m scientific drilling project in Dandong, Liaoning province [J]. *Drilling Engineering*, 2022, 49(4):1-7.
- [20] FAN Haipeng, WU Min, CAO Weihua, et al. An operating performance assessment strategy with multiple modes based on least squares support vector machines for drilling process [J]. *Computers & Industrial Engineering*, 2021, 159:107492.
- [21] 中国地质调查局. 地质云 3.0 [DB/OL]. <https://geocloud.cgs.gov.cn/>, 2021.
- China Geological Survey. Geological cloud 3.0 [DB/OL]. <https://geocloud.cgs.gov.cn/>, 2021.
- [22] 刘添翼. 基于模糊神经网络的机械钻速预测方法 [D]. 大庆: 东北石油大学, 2022.
- LIU Tianyi. Prediction method of rate of penetration based on fuzzy neural network [D]. Daqing: Northeast Petroleum University, 2022.

(编辑 荐华)