临界尺度月壤颗粒辨识集成学习方法

陈 伟，赵德明，高兴文，胡 明
（浙江理工大学 机械工程学院，杭州 310018）

摘 要：针对月壤中临界尺度月壤颗粒对钻探过程中的影响，提出了一种基于集成学习算法的临界尺度月壤颗粒参数辨识模型。使用离散元法（Discrete Element Method，DEM）仿真软件进行建模分析，通过中央复合试验设计（Central Composite Design，CCD）构建多种工况，获得不同条件下月壤钻具的负载特性。以临界尺度月壤颗粒大小与偏置位置为核心变量，采集钻进过程中钻扭矩的信息，并提取关键特征值，在此基础上，首次引入集成学习算法用于临界尺度月壤颗粒参数辨识，构建了同时识别颗粒大小与偏置位置的双变量模型，试验结果表明：该模型在颗粒位置预测方面表现优异，其均方误差（Mean Squared Error，MSE）达到0，预测精度极高；而在颗粒尺寸识别方面也具有良好性能，MSE为1.61。该模型有效解决了颗粒尺寸与位置不可直接观测情况下的参数辨识难题，研究结果可为无人钻探取样的辨识模型技术建立提供参考。

关键词：离散元模型；月壤采样；模拟月壤；集成学习模型
中图分类号：V19 **文献标识码：**A **文章编号：**2096-9287(2025)06-0652-11
DOI:10.3724/j.issn.2096-9287.2025.20250006
引用格式：陈伟，赵德明，高兴文，等. 临界尺度月壤颗粒辨识集成学习方法[J]. 深空探测学报（中英文），2025，12（6）：652-662.
Reference format: CHEN W，ZHAO D M，GAO X W，et al. Integrated learning method for identifying critical-sized lunar soil particles[J]. Journal of Deep Space Exploration，2025，12（6）：652-662.

引 言

随着“嫦娥六号”成功完成月球背面采样返回任务，标志中国在探月工程取得了重要进展，在未来的探月任务中，保持月壤层理信息的完整性以及确保钻探取样顺利进行是非常重要的[1]。然而，月壤的复杂性和不可预测性给钻探取样带来了诸多挑战。前苏联“月球20号”（Luna 20）在钻进取样过程中因负载过大导致停钻。美国国家航空航天局（National Aeronautics and Space Administration，NASA）与Honey Robotics公司合作开发的火星探测钻探自动化（Drilling Automation for Mars Exploration，DAME）项目在极地考察中成功辨识6种典型工况，并且在钻进过程中排除了故障[2-3]。这些经验可以看出，无人自主钻进取样作业任务，最严峻的考验是无法提前预测月壤的结构形态以及物理化学性质，如遇到临界尺度月壤颗粒。由于月球环境的恶劣，加上钻进速度不当、月壤结构复杂等问题，在钻探取样过程中可能会发生故障，导致堵钻、卡钻，从而无法正常完成采样。为了确保月壤取样的正常进行，有必要完善和发展月壤探测技术。

临界尺度月壤辨识是指对月壤中具有特定粒径范围的颗粒进行识别、分析和研究。在月壤钻探取样中，各国学者大多数通过模拟月壤进行分析试验，如Takashi等[4]通过X射线衍射量化了月壤颗粒形态的分布；姜生元等[5]通过分析“嫦娥五号”任务的遥测数据，分析了月壤层序结构及钻进负载波动状态；Li等[6]建立了月岩钻进模型；孙淼等[7]通过月壤侵彻试验分析了颗粒的扰动；Cui等[8-9]通过离散元仿真分析钻具参数对取样效率的影响。然而上述研究均未考虑到临界尺度颗粒的局部干扰。刘天喜等[10]通过三维离散元仿真分析，分析了大颗粒岩块对采样过程的影响。

近年来，在月壤参数辨识中，李操[11]基于月壤–铲斗土力学理论、丁亮等[12]通过月球车轮地作用的力学积分模型分别提出了参数辨识方法，但是依赖于静态假设；周峻等[13]提出基于钻具进尺和回转运动模式的进尺力预测模型；唐钧跃等[14]提出基于可钻性在线辨识的钻进控制方法；在利用机械学习上，郑红燕等[15]利用神经网络算法区分月壤的密实度情况；Zou等[16]提出基于GA-BP神经网络识别月壤剪切参数；罗仁泽等[17]提出了基于集成学习岩性预测方法；Kodikara等[18]选择9种机械学习算法分析了月壤特征，该方法显著提升了

收稿日期：2025-01-22 修回日期：2025-03-09
基金项目：国家自然科学基金（51805488）

识别准确率。综上所述，以上学者尽管在月壤钻取过程层序和参数辨识方法上取得了重要进展，但是在临界尺度月壤辨识（如颗粒大小、颗粒位置的预测）研究方面仍缺乏有效手段，因此，开发一种临界尺度月壤颗粒辨识模型以预测工况，对于保障钻探取样的正常进行具有重要意义。

基于以上问题，本文针对临界尺度月壤颗粒对钻探过程的影响，提出了基于集成学习的临界尺度月壤颗粒参数辨识框架。首先，利用离散元法（Discrete Element Method，DEM）仿真软件建立模拟月壤颗粒模型，并通过试验设计得出具有代表性的工况进行钻取动态仿真，然后，根据钻进负载数据获取特征值，利用基于集成学习算法进行工况参数辨识。试验结果为后续钻具的钻进以及月壤参数辨识优化问题的研究提供了参考。

1　月壤颗粒离散元模型

1.1　EEPA接触模型

EEPA（Extended Elastic-Plastic Adhesion）接触模型是英国爱丁堡大学（The University of Edinburgh）John Paul Morrissey等开发创建的，该颗粒模型是DEM中一种重要的接触模型，用于模拟颗粒之间的弹性、塑性以及黏附特性[19]。它特别适用于研究微小颗粒或具有明显表面黏附作用的物质，例如粉末、颗粒流体和月壤[20]。

传统的颗粒接触模型有Hertz-Mindlin接触模型和JKR接触模型，这些接触模型不能对黏附颗粒的应力变化过程进行有效的描述，在仿真环境中，颗粒在压缩过程中会发生弹性和塑性变形，当压力卸载后颗粒间则会保留一定的重叠量，并且颗粒间接触面积直接影响到黏附强度。在本文中，为了模拟月壤的钻进过程，本文采用EEPA接触模型，该模型用于模拟颗粒间的弹性、塑性及黏附特性。将弹性、塑性以及黏附特性结合起来，能够更准确地描述颗粒间或颗粒与表面的复杂接触行为。EEPA接触模型中的法向黏附力与模型参数的非线性函数关系如图1所示，横坐标δ代表颗粒间重叠量；纵坐标f_n代表颗粒间弹力，正值代表颗粒间的排斥力，负值代表颗粒间的黏结力。

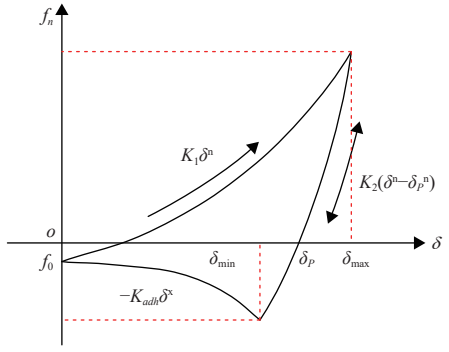


图 1　EEPA接触模型

Fig. 1　EEPA contact model

1.2　EEPA计算模型

当颗粒δ颗粒间重叠量为0时，颗粒间法向黏结力为f_0，随着颗粒δ颗粒间重叠量不断增加，从图1中可以看出力有3条计算路径，第1条计算路径是初始的力，也就是力的加载过程K_1，第2条路径是力的卸载以及再加载过程K_2，第3条路径是黏性力的变化过程K_{adh}。

各个路径的函数关系为

$$f_n=\begin{cases}K_1\delta^n, & K_2(\delta^n-\delta_P^n)\leqslant K_1\delta^n\\ K_2(\delta^n-\delta_P^n), & -K_{adh}\delta^x\leqslant K_2(\delta^n-\delta_P^n)<K_1\delta^n\\ -K_{adh}\delta^x, & K_2(\delta^n-\delta_P^n)^x<-K_{adh}\delta\end{cases}\tag{1}$$

其中：f_0为黏附强度；K_1为加载刚度；K_2为卸载/重新加载刚度；K_{adh}为黏附刚度；δ_p为颗粒间黏结力为 0 时的重叠量；n为刚度指数；x为黏附分支功率。

EEPA接触模型是对颗粒间的相互作用进行描述。考虑到月壤颗粒形状的多样性和不规则性，为简化计算，本文将复杂的非球形颗粒假设为球形颗粒进行仿真。在这种简化模型中，颗粒间的碰撞通过软球模型进行模拟。具体来说，EEPA接触模型中，当颗粒接触时，接触力随位移的增加表现为非线性，包括了弹性变形和塑性变形以及黏附特性，这种模型既保留了颗粒间相互作用的基本特性，又有效降低了计算复杂度。

2　仿真与试验设计

2.1　模拟月壤模型参数

为提高仿真环境的可信度与结果的实际参考价值，本研究采用物理力学特性与真实月壤高度相似的模拟月壤（CUG-1A）进行仿真试验。CUG-1A 的特征与真正月球土壤的特征非常接近[21]，颗粒尺为 0.01～1 mm，堆积密度大约为 2 500 kg/m³，与真实月壤的成分相似。CUG-1A与实际真实月壤的区别很小，因此CUG-1A月壤参数可以作为真实月壤模拟研究的替代物。仿真模型包含两类月壤颗粒：临界尺度月壤颗粒和月壤。真实月壤的颗粒尺寸远远小于1 mm，每个颗粒都有明显的不规则性。与球形颗粒相比，这些颗粒之间的相互作用更为复杂。但是，如果按照实际尺寸进行仿真，将需要大量的计算时间。因此仿真需要简化需求，提高计算效率，模拟月壤的优化参数详见表1。虽然仿真试验无法直接定量分析临界尺度月壤颗粒对

钻具负载的影响，但可以利用这些数据来研究钻具负载特性的趋势。

表 1　月壤颗粒DEM仿真模型参数

Table 1　DEM simulation model parameters for lunar soil particles

参数	数值
颗粒直径/mm	0.5～1
颗粒密度/（kg·m^{-3}）	2.6 × 10^3
颗粒剪切模量/Pa	1 × 10^7
颗粒泊松比	0.5
碰撞恢复系数	0.25
静摩擦系数	0.8
滚动摩擦系数	0.3
月壤桶尺寸/mm	半径30，高80
月壤桶泊松比	0.5
月壤桶密度/（kg·m^{-3}）	7.85 × 10^3
月壤桶剪切模量/Pa	1 × 10^7
月壤桶碰撞恢复系数	0.5
颗粒–月壤桶静摩擦系数	1.0
颗粒–月壤桶滚动摩擦系数	0.3

2.2　EEPA模型堆积试验

在离散元DEM软件中，EEPA颗粒模型参数设置主要是黏附强度f_0、接触黏附能γ、接触塑形比λ_P、刚度指数n、黏附分支功率x、切向刚度系数ζ_{tm}。利用DEM离散元软件进行漏斗法堆积角仿真试验来确定颗粒参数范围，模拟月壤堆积角在30°～40°是比较合理的，堆积角仿真模拟试验如图2所示。

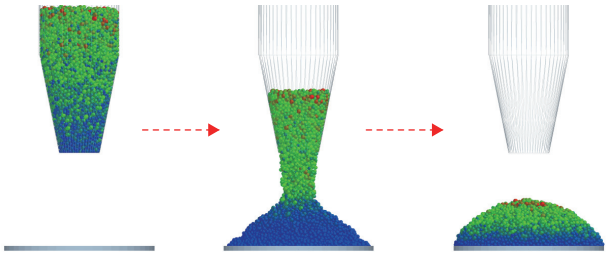


图 2　堆积角仿真模拟流程图试验

Fig. 2　Flowchart of simulation process for angle of repose experiment

初始参数范围设定，黏附强度f_0：–0.02～0 N（负值表示排斥，正值表示吸引），接触黏附能γ：0.5～1.0 J/m^2（参考CUG–月壤特性）。通过不断调整模型参数（见表2），使模拟出来的颗粒性质与真实情况一致，则认为该离散元模型参数是符合实际情况的。从表2中可以看出，当黏附力表面能小于0，黏附力表面能大于或者等于1，颗粒无法正常自然落下，又因月壤钻进为非线性，刚度指数默认为1.5。

由表2中数据可知，序号2～6之间的堆积角较为合理，为了便于试验的进行，本文仿真模拟月壤选择序号4的数据作为仿EEPA颗粒模型参数。

表 2　参数设置及堆积角试验

Table 2　Parameter settings and angle of repose experiment

序号	黏附强度 f_0/N	接触黏附能 γ/（J·m^{-2}）	接触塑形比 λ_P	刚度指数 n	黏附分支功率 x	切向刚度系数 ζ_{tm}	堆积角 θ/（°）
1	–0.01	0.5	0.7	1.5	5	0.28	88.96
2	0	0.5	0.7	1.5	5	0.28	32.95
3	0	0.6	0.7	1.5	5	0.28	34.14
4	0	0.7	0.7	1.5	5	0.28	36.56
5	0	0.8	0.7	1.5	5	0.28	37.74
6	0	0.9	0.7	1.5	5	0.28	39.54
7	0	1.0	0.7	1.5	5	0.28	89.72

综上所述，本文选用与真实月壤参数相近的模拟月壤CUG-1A，并结合堆积角仿真验证EEPA模型参数设定的合理性。通过调整接触模型参数，使颗粒堆积行为、摩擦特性更趋近于真实月壤状态，保证了仿真结果具备良好的代表性。

2.3　钻具参数设定

钻具参考螺旋取样钻具，其构型如图3所示，螺旋钻具主要由钻头、钻杆和取芯机构组成[22]。钻头与钻杆参数为：钻杆内径为12 mm，外径为20 mm，高为70 mm，螺距为7.5 mm，导程为1.5 mm，螺纹升角为16°。

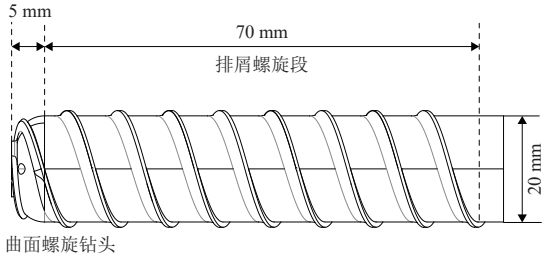


图 3　螺旋取样钻具

Fig. 3　Spiral sampling drill

在钻具材料方面，根据月壤钻具要求“轻量化、耐磨损、可靠性”的技术要求[10]，选用铝合金材质，具体参数如表3所示。

表 3　钻具仿真模型参数

Table 3　Parameters of spiral sampling drill

参数	数值
钻杆内径/mm	12
钻杆外径/mm	20
钻具密度	2.9 × 10^3
钻具剪切模量	1 × 10^7
泊松比	0.3
碰撞恢复系数	0.3
静摩擦系数	0.4
滚动摩擦系数	0.02
进给速度/(mm·s^{-1})	65
转速 n_r/(rad·min^{-1})	600
取样深度	80

2.4　仿真条件

为了揭示月壤钻进负载的产生机理，本文在DEM软件月壤钻取过程中进行动态仿真。仿真不仅能获得负载数据，还能获得钻具钻探取样的动态变化过程，进一步分析临界尺度月壤颗粒负载的机理。本文中对临界尺度月壤颗粒的取值范围为钻具的0.5～1倍，即临界尺度月壤颗粒直径为10～20 mm。基于第1章相关参数，月壤钻探过程中初始化仿真动态过程如图4所示。

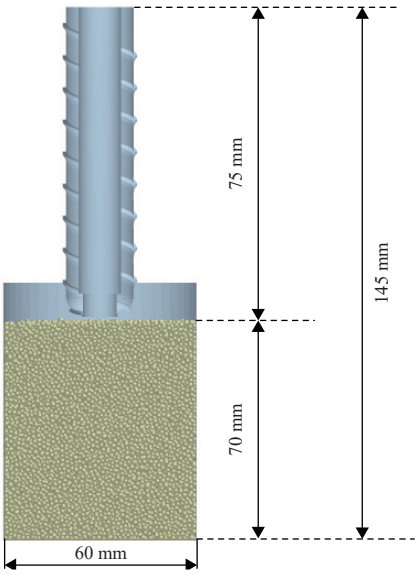


图 4　为钻杆在离散元仿真软件中的初始化仿真模型

Fig. 4　Initialization simulation model of drill rod in discrete element simulation software

2.5　试验设计

本文在定义仿真试验矩阵时，变量为颗粒大小和偏置位置，使用中央复合设计（Central Composite Design，CCD），得到有代表性的试验点。CCD试验设计在双因素非线性问题具有高效性，可减少仿真试验次数，同时覆盖尽可能多的关键变量和交互作用，并通过试验获得的数据，进而寻求最优参数。本文利用Design-Expert13.0软件设计，参数设置如表4所示，双因素试验表水平编码如表5所示，9组试验设计结果如表6所示。

表 4　试验设计软件参数设置

Table 4　DOE software settings

参数	设置值
因素与水平	颗粒大小（S）：–1.414 ～+1.414 偏置位置（*P*）：–1.414 ～+1.414
中心点重复次数	3次
星点距离	1.414
置信区间	$P<0.05$
试验顺序	随机

表 5　仿真试验因素与水平

Table 5　Factors and levels of simulation experiment

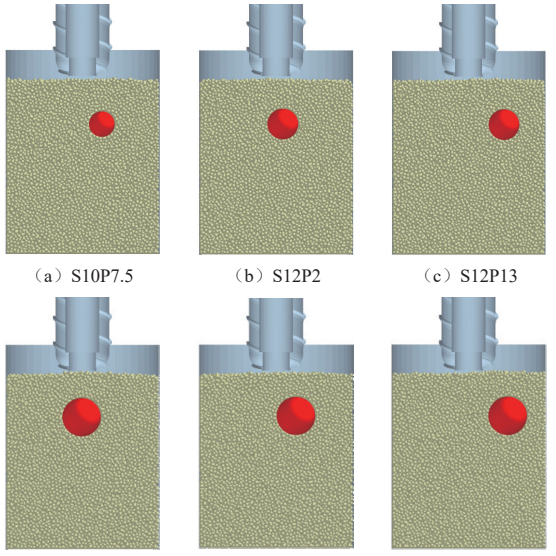
因素	水平				
	–1.414	–1	0	+1	+1.414
颗粒大小*S*/mm	10	12	15.0	18	20
偏置位置*P*/mm	0	2	7.5	13	15

表 6　工况仿真设计试验

Table 6　Experimental design for operational conditions simulation

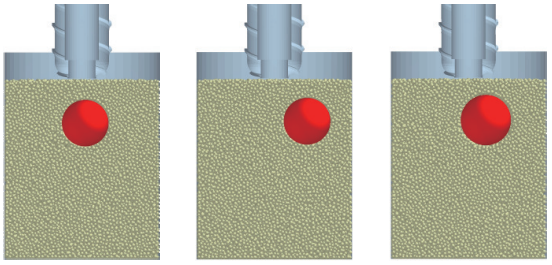
试验组	颗粒大小/mm	偏置位置/mm
1	10	7.5
2	12	2.0
3	12	13.0
4	15	0
5	15	7.5
6	15	15.0
7	18	2.0
8	18	13.0
9	20	7.5

因此根据CCD设计，共设计了9种工况，分别探究临界尺度月壤颗粒的大小、位置对钻进负载的影响，试验设计钻进量为65 mm，进给速度为65 mm/s，钻具转速为600 rad/min。颗粒指定位置为20 mm，月壤桶半径为30 mm，月壤桶高度为80 mm，月壤桶颗粒半径设置为0.5～1 mm，共有9组试验，为了便于区分，将颗粒大小标记为S（Size），偏置位置标记为P（Position），如S15P0，表示中置15 mm颗粒，S20P7.5表示颗粒大小为20 mm，偏离钻具中心位置为7.5 mm。等待月壤颗粒系统完全处于静止状态后开始钻进，钻取时间持续10 s，仿真时间步长设置为0.02 s，即每隔0.02 s记录一次数据，仿真初始效果图如图5所示，仿真完成后，临界尺度月壤颗粒运动轨迹图如图6所示。



（a）S10P7.5　（b）S12P2　（c）S12P13

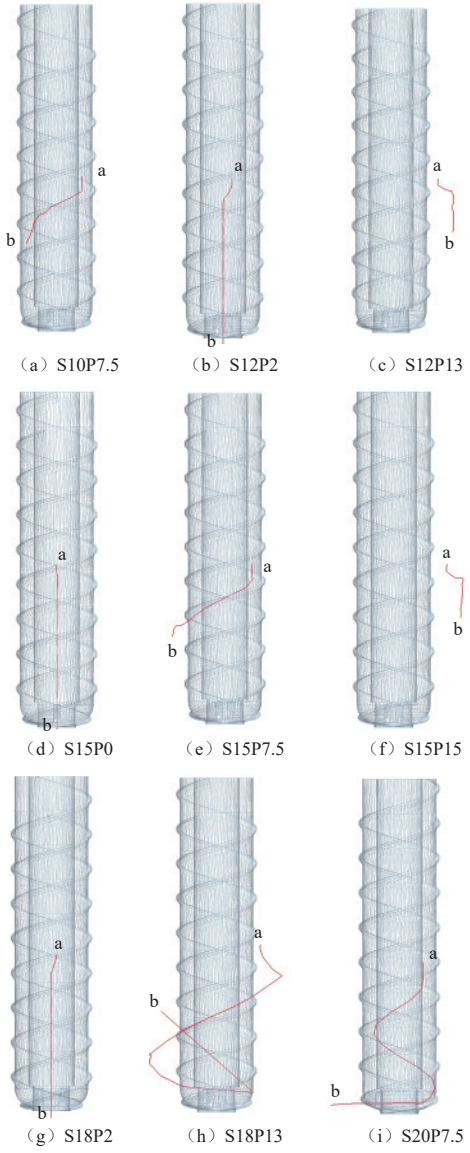
（d）S15P0　（e）S15P7.5　（f）S15P15



（g）S18P2 （h）S18P13 （i）S20P7.5

图 5 临界尺度月壤颗粒初始仿真状态

Fig. 5 Initial simulation state of critical-scale lunar soil particles



（a）S10P7.5 （b）S12P2 （c）S12P13

（d）S15P0 （e）S15P7.5 （f）S15P15

（g）S18P2 （h）S18P13 （i）S20P7.5

图 6 临界尺度月壤颗粒运动轨迹图

Fig. 6 Critical-scale lunar soil particle motion trajectory diagram

图6中展示了不同工况下临界尺度月壤颗粒的运动轨迹，在颗粒居中或距钻头较近时，如图S12P2〔（图6（b））、S15P0〔图6（d）〕、S18P2〔图6（g）〕，颗粒被钻头顶部持续推动，轨迹呈垂直直线运动；在颗粒偏置时，如图S15P7.5〔图6（e）〕、S18P13〔图6（h）〕、S20P7.5〔图6（i）〕，颗粒受离心力影响向桶壁偏移，轨迹呈螺旋扩散；在颗粒远离时，如图S10P7.5〔图6（a）〕、S12P13〔图6（c）〕、S15P15〔图6（f）〕，颗粒与钻头无直接接触，轨迹呈无规则偏移。

3 辨识模型构建

3.1 集成学习

集成学习是一种通过结合多个学习模型来提高整体预测性能的机器学习方法，广泛应用于分类、回归和检测问题。集成学习的核心思想是通过将多个弱学习器结合成一个强学习器，以提高整体模型的预测性能。其关键在于通过模型多样性，减少单一模型的误差。集成学习能弥补每个模型的不足，增强模型的泛化能力。本章先对数据进行预处理、提取特征值，然后进行参数优化、模型评估和试验验证。结果显示，经过调参后，模型预测误差变小，达到预测的目的。

3.2 数据获取

在仿真试验中，结合实际情况， 在钻取月壤时，主要通过钻杆的数据反馈当前的状态。钻进负载是钻取过程中的关键指标之一，仿真中获取的数据主要是钻的力和扭矩。本文中主要对扭矩进行分析，因为力的数据在测量中可能存在较大的瞬时波动，对颗粒特性表现出较大的随机性，不利于特征提取。扭矩是钻具在受载的过程中反映颗粒间相互作用特性的综合指标，颗粒大小和偏置位置直接影响扭矩的动态响应。扭矩曲线的波动性较强，能更直观地反映颗粒特性，干扰较少，信噪比相对较高，有利于后续信号处理和特征提取。

3.2.1 原始数据

仿真试验数据包含无临界尺度月壤颗粒和临界尺度月壤颗粒两类样本，其中，无临界尺度月壤颗粒进给力随着钻进深度的变化呈渐进上升趋势，如图7所示，数据图无明显的波动，为后续的数据去趋势分析提供了理想的输入条件。相较之下，临界尺度月壤颗粒的数据图呈显著波动特征，如图8左所示，说明不同大小和不同位置对钻进扭矩的影响较大。为更有效提取信号的关键变化特征，去趋势处理目的在于剔除随深度增长的趋势项，保留临界尺度月壤颗粒扰动相关的真实波动信息，如图8各分图右上所示，图8各分图右下为特征值的提取结果。

然而，当颗粒尺寸不断增大时，扭矩显著增加；其次是偏置位置，偏置位置接近钻杆竖直中心时，颗

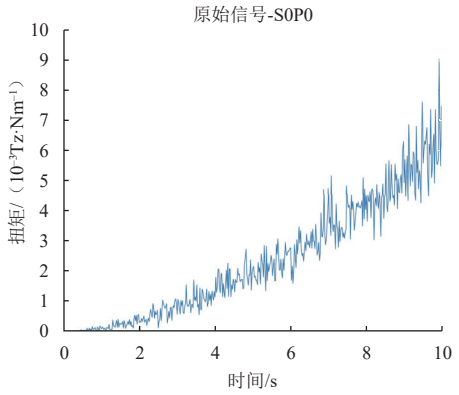


图 7　无临界尺度月壤颗粒扭矩图

Fig. 7　Torque diagram without critical-scale lunar soil particles

粒难以排开，被钻头顶着走，偏移量不大，当临界尺度月壤颗粒被顶着到桶的底部时候，钻头扭矩急剧上升，并发生严重堵钻，直至钻进停止，如图8（b）、（d）、（g）、（i）所示。当颗粒偏置位置与钻具较远，钻具碰不到临界尺度月壤颗粒，如图8（c）所示，扭矩负载缓慢增加，钻具抖动不是很明显，当颗粒偏置位置在7.5 mm时，在2～6 s之间钻具扭矩抖动明显，随着钻具向下钻进，颗粒被排开，扭矩缓慢增加，抖动不明显，如图8（a）、（e）所示。当颗粒大小在15～20 mm范围，偏置位置较远时，钻具抖动时而平缓，时而增加，如图8（f）、（h）所示。

综上，钻进过程中，较小颗粒的位置和形状变化对钻进负载的影响不明显；当颗粒位于中间位置时，负载显著增加，而在偏置情况下，尽管负载有所波动，但总体上仍保持稳定增加；较临界尺度月壤颗粒位于钻头直径范围内时，钻进过程中的负载波动幅度较大；尽管颗粒位于中间位置，钻头负载没有明显增加，但钻进过程中负载波动幅度显著，这是由于钻头推动临界尺度月壤颗粒直线运动，达到临界点时临界尺度月壤颗粒发生偏移和翻转，最终导致负载波动。以上说明颗粒的形状会影响钻进力载的大小以及趋势，也会造成钻具抖动。

3.2.2　数据处理及特征提取

为了便于特征提取，以及更能直观感受钻具抖动程度，在9组代表试验中，本文首先对信号进行去趋势处理，以更好地捕捉颗粒大小和偏置位置的相关特征，目的是去除信号中的长期变化趋势，从而关注信号中的短期波动和局部变化。去趋势后，选取信号中的 2～7 s 时间段进行特征提取。因为在钻进试验开始时（0～2 s）信号存在过渡过程，且钻进试验最后（7 s 之后）临界尺度月壤颗粒部分急剧增加，不利于特征提取。这一时间段内，颗粒的运动和试验设置对信号的影响较明显，能够捕捉到与颗粒大小和偏置位置密切相关的特征。

通过对去趋势后的信号进行特征提取，选择了6个特征来建立颗粒大小和偏置位置的预测模型。这些特征包括：均值、标准差、能量、最大值、最小值和均方根。这些特征从不同角度反映了信号的特性，能够有效捕捉颗粒行为的信息。均值反映了信号的中心趋势；标准差提供了信号的变化幅度；能量和均方根描述了信号的强度和有效幅度；最大值和最小值捕捉了信号中的极端波动。综上，这些特征为后续的建模和预测提供了重要依据。

3.3　辨识算法建模

3.3.1　数据标准化

集成学习算法的基本原理是通过组合多个基础学习器来提高整体预测性能。为了提升模型训练精度，本文先对特征数据进行标准化，标准化的数据能提高模型的收敛速度和准确性。本文使用Z-score标准化方法对数据进行了处理，标准化后的数据使得6个特征值分布一致，避免了不同量纲特征对模型训练产生的影响。标准化公式为

$$X' = \frac{X-\mu}{\delta} \tag{2}$$

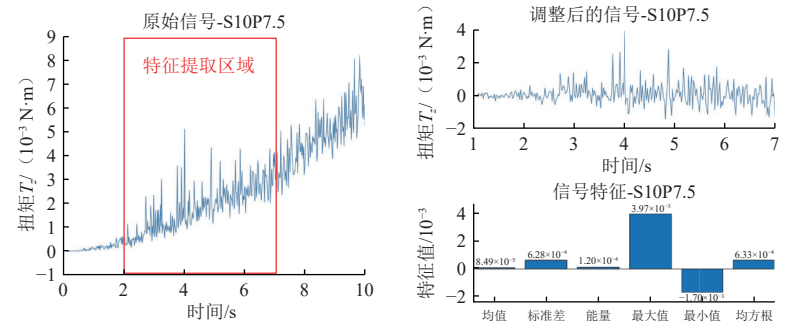
其中：X为原始数据；μ为特征数据的均值；δ为特征数据的标准差。

在完成数据标准化后，再利用集成学习算法模型对颗粒大小和偏置位置分别进行训练。因为有两个变量因素，所以建立了两个独立的集成学习算法模型，一个用于预测颗粒大小，另一个用于预测偏置位置。

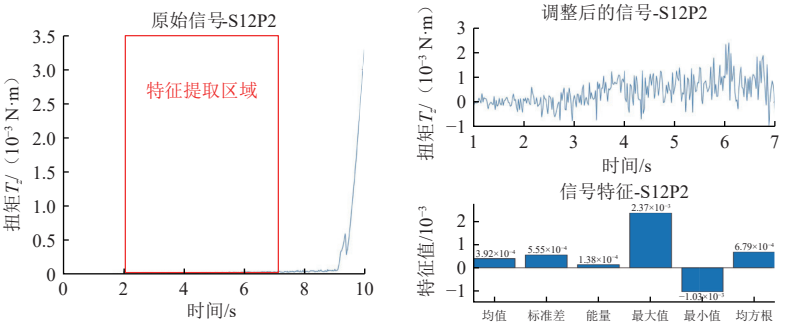
3.3.2　构建集成模型与评估

在本文中，构建集成学习模型时，选择了两个不同的决策树方法来提高预测性能，并且分别对决策树的超参数（最小叶节点大小、最大分裂数、分裂准则、特征采样方式）以及集成学习的学习循环次数进行了调优，以优化模型性能。

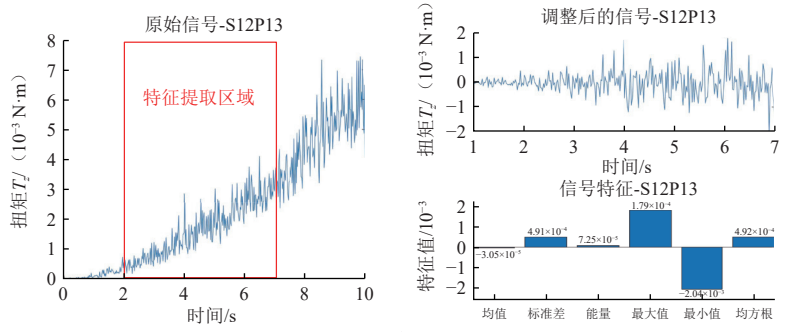
使用AdaBoost来预测颗粒的大小。这是因为对于颗粒大小的预测，模型可能会面临更多的复杂性和误差，尤其在处理不完全或有偏数据时，AdaBoost能通过不断强化对难以预测样本的关注，帮助提高模型的准确性。使用Bagging来预测颗粒的偏置位置是因为偏置位置的数据可能存在噪声和波动，此时Bagging能通过减小模型的方差，减少噪声对结果的影响，从而提供更加稳定的预测，其中算法交叉验证结果对比如表7所示。



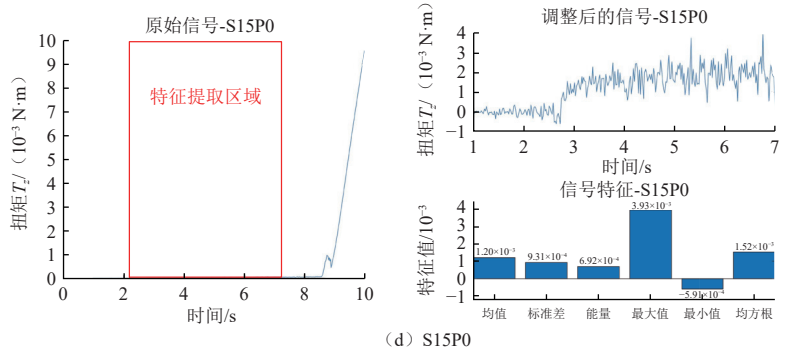
（a）S10P7.5



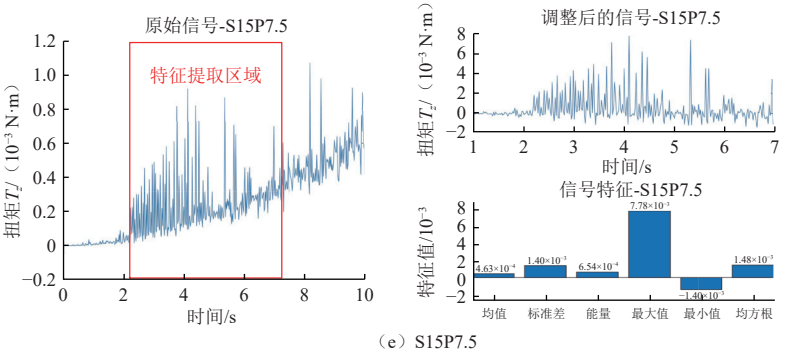
（b）S12P2



（c）S12P13



（d）S15P0



（e）S15P7.5

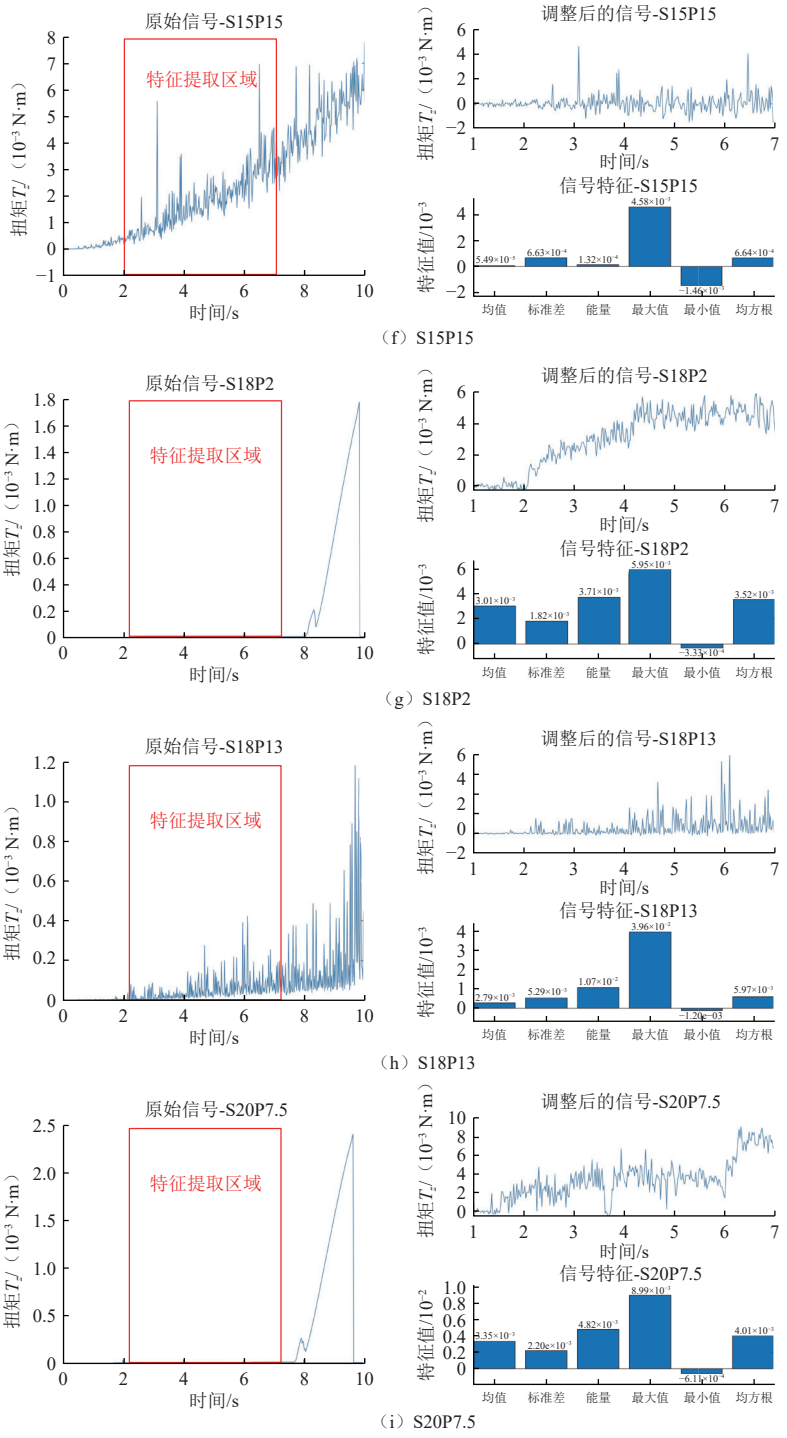


图 8　颗粒大小和偏置位置的原始图及取特征值图

Fig. 8　Original graphs of particle size and offset position and feature extraction graphs

表 7　交叉验证结果对比

Table 7　Comparison of cross-validation results

算法组合	大小MSE	偏置MSE
AdaBoost和Bagging	1.61	0
双AdaBoost	1.72	0.08
双Bagging	2.03	0.12
随机森林	1.85	0.15

本模型中总共有54个样本进行模型训练，在模型训练完成后，取其中18组测试集样本，如图9所示。为了避免过拟合并对模型进行合理评估，本文使用均方误差（Mean Squared Error，MSE）作为性能指标。均方误差是衡量预测值与实际值之间差异的一个常用指标，计算公式为

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_{\mathrm{pred},i}-y_{\mathrm{true},i})^2 \tag{3}$$

其中：$y_{\mathrm{pred},i}$为第i个样本的预测值；$y_{\mathrm{true},i}$为第i个样本的真实值；n为样本总数。

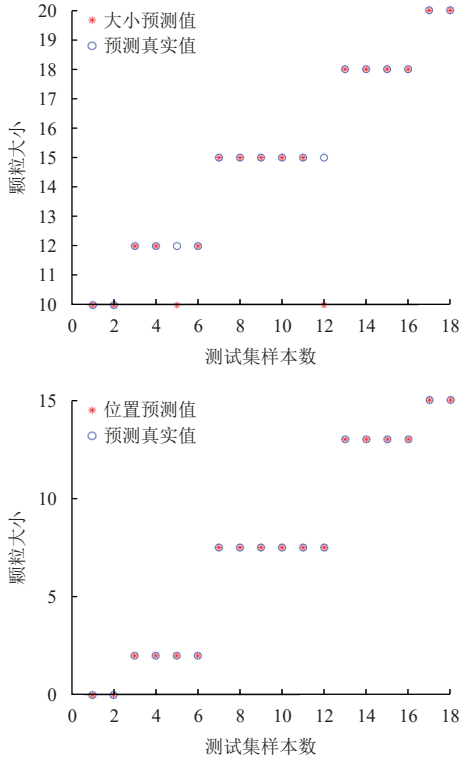


图 9　临界尺度月壤颗粒参数辨识结果

Fig. 9　Identification results of critical-scale lunar soil particle parameters.

在模型测试过程中，模型对测试集的预测结果分别是颗粒大小和偏置位置的预测值。通过计算预测值与实际值之间的MSE，可以量化模型的预测误差，评估模型的泛化能力。

在本文中，分别计算了颗粒大小和偏置位置的MSE值，以验证模型的预测精度。颗粒大小的MSE值为1.61，偏置位置的MSE值为0。较低的MSE值表明颗粒大小的预测结果误差相对较大，而偏置位置的预测准确。

通过对颗粒大小和偏置位置的预测结果进行分析，可以发现，模型在偏置位置的预测上具有较高的准确性，准确率为100%。而在颗粒大小的预测上表现稍弱，准确率为88.89%。临界尺度月壤参数辨识结果显示，模型能够较为准确地预测颗粒偏置位置，但对颗粒大小预测不准确，因为颗粒大小的预测可能受到一些非线性因素或数据噪声的影响，导致其预测精度稍有下降。

3.4　试验验证

为验证模型在试验中的效果，本小节以CUG-1A模拟月壤为对象，借助多动能钻取试验平台，进行了含临界尺度模拟月壤颗粒的钻进试验，如图10所示。钻具直径为60 mm，试验颗粒大小分别为16 mm和20 mm颗粒，放在同一位置约7.5 mm处进行预测，试验钻进规程为：进给速度5 mm/s，转速5 rad/min，采样频率50 Hz，这样仿真中微小的变化趋势就得以放大便于观察。最后再铺上模拟月壤土进行试验，试验结束后采集并分析了钻进过程的扭矩等信息，如图11所示。

图 10　临界尺度模拟月壤颗粒的钻进试验

Fig. 10　Drilling experiment of critical-scale lunar soil particles

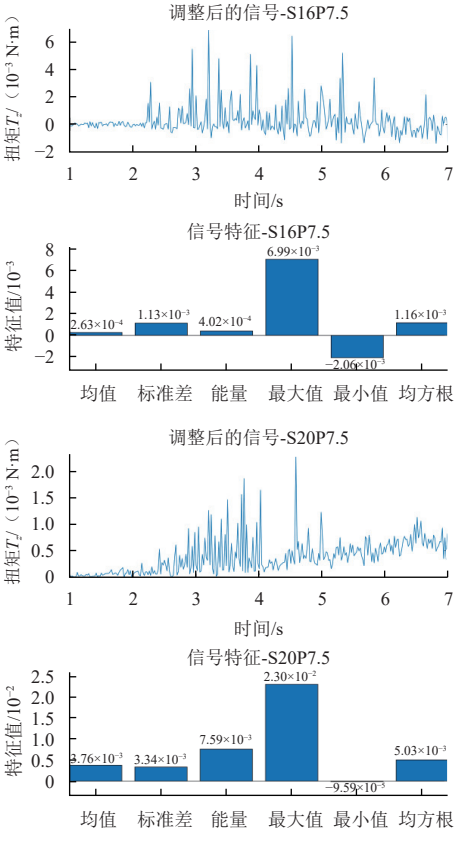


图 11　临界尺度模拟月壤颗粒试验数据及处理

Fig. 11　Critical-scale simulation of lunar soil particle experimental data and processing

最后采用使用训练好的集成学习模型，通过提取试验结果的特征值，以扭矩为输入，得出该钻进工况的测试结果。试验结果得出颗粒大小为16 mm的临界

尺度月壤颗粒的预测结果为S15P7.5，颗粒大小为20 mm的临界尺度月壤颗粒的预测结果为S20P7.5，说明试验结果预测与仿真预测模型结果相符。

4　结　论

1）EEPA（颗粒接触模型）在颗粒物理模拟中发挥了重要作用，尤其是在钻头与颗粒之间的相互作用模拟中。通过在离散元软件DEM中应用EEPA模型，能准确地描述颗粒间的接触力、黏附力及颗粒的动态响应。EEPA模型对颗粒接触的物理现象具有较高的模拟精度，能够有效预测不同颗粒大小、形态和物理属性下的颗粒行为。

2）本文采用CCD进行试验设计，并基于试验数据构建预测模型。通过合理的试验设计，研究了颗粒大小和偏置位置这两个因素对试验结果的影响，收集了不同条件下的仿真数据。试验设计的应用有效地减少了试验次数，优化了资源的使用。

3）集成学习模型被用于预测颗粒大小和偏置位置，模型训练基于标准化后的特征数据。该模型能够有效处理非线性问题，并在本研究中表现出良好的预测能力。预测误差结果显示，颗粒大小的预测误差较大，偏置位置的预测中具有较高的准确性。

参考文献

[1] 欧阳自远. 我国月球探测的总体科学目标与发展战略[J]. 地球科学进展，2004(3)：351-358.
OUYANG Z Y. Overall scientific objectives and development strategy of China's lunar exploration program[J]. Journal of Earth Sciences，2004(3)：351-358.

[2] CHEN W B，DIPIRRO M，MCKINLEY I. et al. Active cryocooling needs for NASA space instruments and future technology development[J/OL]. Cryogenics，2024，141：103877. https://www.sciencedirect.com/science/article/abs/pii/S0011227524000973.

[3] GLASS B J，DAVE A，AMCKAY C P，et al . Robotic and automation for icebreaker[J]. Journal of Filed Robotics，2014，31(1)：192-204.

[4] TAKASHI M，JUN K，KENTARO U. et al. 3D shape characterization. and image-based DEM simulation of the lunar soil simulant FIS-1[J]. Journal of Aerospace Engineering，2009，22(1)：15-23.

[5] 姜生元，梁杰能，赖小明，等. 嫦娥五号月壤剖面钻进取芯状态分析与解译[J]. 机械工程学报，2022，58(10)：348-360.
JIANG S Y，LIANG J N，LAI X M，et al. Analysis and interpretation of the drilling and core sampling status of Chang'E-5 lunar soil profile[J]. Journal of Mechanical Engineering，2022，58(10)：348-360.

[6] LI P，JIANG S Y，TANG D W，et al. Design and testing of coring bit on drilling lunar rock simulant[J]. Advances in Space Research，2017，59(4)：1057-1076.

[7] 孙淼，张鸿宇，迟润强，等. 侵彻式月壤探测地面模拟试验研究[J]. 深空探测学报(中英文)，2022，9(2)：141-149.
SUN M，ZHANG H Y，CHI R Q，et al. Ground simulation test of penetrator for lunar soil exploration[J]. Journal of Deep Space Exploration，2022，9(2)：141-149.

[8] CUI J S，HOU X Y，WEN G L，et al. DEM thermal simulation of bit and object in drilling of lunar soil simulant[J]. Advances in Space Research，2018，62(5)：967-975.

[9] ZHAO D M，TANG D W，HOU X Y，et al. Soil chip convey of lunar subsurface auger drill[J]. Advances in Space Research，2016，57(10)：2196-2203.

[10] 刘天喜，魏承，马亮，等. 大颗粒岩块对月壤钻取过程的影响分析[J]. 岩土工程学报，2014，36(11)：2118-2126.
LIU T X，WEI C，MA L，et al. Analysis of the impact of large particles on the lunar soil drilling process[J]. Journal of Geotechnical Engineering，2014，36(11)：2118-2126.

[11] 李操. 基于机械臂的月壤挖掘采样及力学参数辨识的研究[D]. 哈尔滨：哈尔滨工业大学，2014.
LI C. Research on lunar soil excavation，sampling，and mechanical parameter identification based on robotic arm[D]. Harbin: Harbin Institute of Technology，2014.

[12] 丁亮，高海波，邓宗全，等. 基于月球车轮地作用地面力学积分模型的月壤力学参数辨识方法[J]. 航空学报，2011，32(6)：1112-1123.
DING L，GAO H B，DENG Z Q，et al. Lunar soil mechanical parameter identification method based on a ground force mechanics integration model for lunar rover wheels[J]. Acta Aeronautica et Astronautica Sinica，2011，32(6)：1112-1123.

[13] 周峻，刘天喜，梁磊，等. 月壤钻采进尺力形成机理及进尺力预测建模[J]. 宇航学报，2023，44(1)：34-42.
ZHOU J，LIU T X，LIANG L，et al. Mechanism of lunar soil drilling and sampling penetration force and penetration force prediction modeling[J]. Journal of Astronautics，2023，44(1)：34-42.

[14] 唐钧跃，全齐全，姜生元，等. 一种基于可钻性在线辨识的月面钻进控制方法研究[J]. 深空探测学报(中英文)，2015，2(4)：325-332.
TANG J Y，QUAN Q Q，JIANG S Y，et al. Control method of lunar drilling based on online identification of drilling ability[J]. Journal of Deep Space Exploration，2015，2(4)：325-332.

[15] 郑燕红，邓湘金，庞勇，等. 月球风化层钻取采样过程密实度分类研究[J]. 航空学报，2020，41(4)：155-165.
ZHENG Y H，DENG X J，PANG Y，et al. Study on the compaction classification of the lunar regolith drilling and sampling process[J]. Acta Aeronautica et Astronautica Sinica，2020，41(4)：155-165.

[16] ZOU M，XUE L，GAI H J. et. al. Identification of the shear parameters for lunar regolith based on a GA-BP neural network[J]. Journal of Terramechanics，2020，89：21-29.

[17] 罗仁泽，庹娟娟，倪华玲，等. 基于改进集成学习的测井岩性识别方法研究[J]. 石油物探，2023，62(2)：212-224.
LUO R Z，TUO J J，NI H L，et al. Logging lithology identification method based on improved ensemble learning[J]. Geophysical Prospecting for Petroleum，2023，62(2)：212-224.

[18] KODIKARA G R，MCHENRY L J. Machine learning approaches for classifying lunar soils[J]. Icarus，2020，345：113-719.

[19] ZHOU L，LAN Y B，YU J Q. et al. Validation and calibration of soil parameters based on EEPA contact model[J]. Computational Particle Mechanics，2023，10：1295-1307.

[20] 万里鹏程，李永磊，苏辰，等. 基于EEPA接触模型的土壤耕作特性模拟及颗粒球型影响分析[J]. 中国农业大学学报，2021，26(12)：193-206.

WANLI P H，LI Y L，SU C，et al. Soil tillage characteristics simulation and particle spherical shape impact analysis based on the EEPA contact model[J]. Journal of China Agricultural University，2021，26(12)：193-206.

[21] 贺新星，肖龙，黄俊，等. 模拟月壤研究进展及CUG-1A模拟月壤[J]. 地质科技情报，2011，30(4)：137-142.
HE X X，XIAO L，HUANG J，et al. Research progress on lunar soil simulation and CUG-1A lunar soil simulant[J]. Geological Science and Technology Information，2011，30(4)：137-142.

[22] 庞勇，王国欣，汤滨，等. 月球浅层月壤螺旋钻具进芯机制设计[J]. 深空探测学报(中英文)，2021，8(3)：252-258.
PANG Y，WANG G X，TANG B，et al. Design of coring dynamics with Chang'E-5 anger drill[J]. Journal of Deep Space Exploration，2021，8(3)：252-258.

作者简介：
陈伟(1997–)，男，硕士研究生，主要研究方向：地外天体采样技术。
通信地址：杭州市钱塘区2号大街928号浙江理工大学(310018)
E-mail：cwtech2022@163.com
高兴文(1986–)，男，讲师，硕士生导师，主要研究方向：地外天体采样探测技术、颗粒物质力学及其仿真技术等。**本文通信作者**。
通信地址：杭州市钱塘区2号大街928号浙江理工大学(310018)
E-mail：xwgao@zstu.edu.cn

Integrated Learning Method for Identifying Critical-Sized Lunar Soil Particles

CHEN Wei，ZHAO Deming，GAO Xingwen，HU Ming
（School of Mechanical Engineering, Zhejiang Sci-Tech University，Hangzhou 310018，China）

Abstract： To address the impact of critical-scale lunar soil particles on the drilling process，a parameter identification model for critical-scale lunar soil particles based on ensemble learning algorithms was proposed. Simulation analysis was conducted using Discrete element method (DEM) simulation software，and a central composite design was used to conduct experiments under various operating conditions to obtain the load characteristics of lunar drilling tools under different conditions. Taking model took critical-scale lunar soil particle size and offset position as main variables，torque data from the lunar drilling process were collected. By extracting the features of the experimental data were extracted. Based on this，the ensemble learning algorithm was introduced for the first time for parameter identification of critical-scale lunar soil particles, and a bivariate model was developed to simultaneously identify particle size and offset position. Experimental results show that the ensemble learning algorithm achieved high accuracy in predicting particle positions，with a mean squared error (MSE) of 0，and it also demonstrated good performance in particle size identification，with an MSE of 1.61. This model effectively solves the parameter identification challenge under the condition that particle size and location cannot be directly observed. These findings can provide a reference for developing identification model technologies for autonomous drilling and sampling.

Keywords： DEM；lunar soil sampling；simulated lunar soil；ensemble learning model

Highlights：

- A novel ensemble learning-based model is developed for identifying critical-sized lunar regolith particle parameters during drilling.
- A bivariate model is constructed to simultaneously predict particle size and offset position using torque signal features from DEM simulations.
- The proposed method achieves high prediction accuracy with an MSE of 0 for particle position and 1.61 for particle size，addressing challenges in non-observable parameter identification.

[责任编辑：宋宏，英文审校：宋利辉]