

复杂地形条件下山区土石界面插值策略与预测方法研究

曹子君¹, 郑立宁², 赵士博³, 曾雪松², 王轩毫^{4,*}

1. 西南交通大学高速铁路线路工程教育部重点实验室, 智慧城市与交通学院, 四川 成都 611756;

2. 中国建筑西南勘察设计研究院有限公司, 四川 成都 610052;

3. 西南交通大学土木工程学院, 四川 成都 611756;

4. 武汉大学水资源工程与调度全国重点实验室, 湖北 武汉 430072;

摘要: 土石界面预测对山区机场建设具有重要意义。现有方法通常基于钻孔数据直接估计土石界面高程,但在山区地形起伏剧烈、界面高程差异显著的条件下,其结果准确性难以满足工程需求。

本文提出一种基于地形高程与界面深度的插值策略,土石界面高程可表示为已知地形高程减去土石界面深度。该策略只需预测土石界面深度,因此降低了地形起伏对预测结果的影响。本文系统对比了该策略在多种预测方法中的表现,包括反距离加权、径向基函数回归、高斯过程回归等核方法以及神经网络方法,如多层感知机和 Kolmogorov-Arnold 网络。山区机场工程案例表明,所提策略适用于不同预测方法,并显著提升土石界面预测准确性,为复杂地形条件下的机场选址、土石方工程量估算及施工方案优化提供技术支撑。

关键词: 山区; 地形; 土石界面; 插值策略; 预测方法

中图分类号: TV121.7

收稿日期: 2025-10-05

Interpolation strategy and prediction method of soil-rock interface in mountainous areas under complex topography

CAO Zijun¹, ZHENG Lining², ZHAO Shibo³, ZENG Xuesong², WANG Xuanhao^{4,*}

1. MOE Key Laboratory of High-Speed Railway Engineering, Institute of Smart City and Intelligent Transportation, Southwest Jiaotong University, Chengdu, Sichuan province 611756;

2. China Southwest Geotechnical Investigation & Design Institute Co., Chengdu, Sichuan province 610052;

3. Institute of Civil Engineering, Southwest Jiaotong University, Chengdu, Sichuan, province 611756;

4. State Key Laboratory of Water Resources Engineering and Management, Wuhan University, Wuhan, Hubei province 430072;

基金项目: 国家自然科学基金项目(52278368), 四川省自然科学基金(24NSFSC2017)

第一作者简介: 曹子君(1987—), 男, 教授, 研究方向为岩土工程数据分析、不确定性表征与可靠度设计. ORCID: 0000 - 0002 - 4712 - 1542. E-mail: zijuncao@swjtu.edu.cn.

*通讯作者: 王轩毫, E-mail: wangxuanhao@whu.edu.cn.

Abstract: The prediction of the soil–rock interface is crucial for airport construction in mountainous areas. Conventional methods typically rely on borehole data to directly estimate the elevation of the soil–rock interface. However, when topographic variability is pronounced and interface-elevation contrasts are large, the accuracy of such estimates is often insufficient to satisfy engineering requirements. This paper introduces an interpolation strategy that incorporates terrain elevation and interface depth. In this framework, the elevation of the soil–rock interface is expressed as the terrain elevation minus the depth of the interface. By shifting the prediction task to estimating interface depth alone, this approach reduces the influence of topographic variability on prediction accuracy. The performance of the proposed strategy is systematically evaluated using various prediction methods, including kernel methods such as Inverse Distance Weighting, Radial Basis Function Regression, and Gaussian Process Regression, as well as neural network approaches (e.g., Multilayer Perceptron and Kolmogorov–Arnold Networks). Case studies from airport projects in mountainous regions demonstrate that the strategy can be readily integrated with different prediction methods and substantially improves the accuracy of soil–rock interface predictions. The findings provide technical support for airport site selection, earthwork volume estimation, and construction planning in complex terrains.

Key words: mountainous areas; topography; soil-rock interface; interpolation strategy; prediction method

引言

近年来,我国越来越多的机场选址于地形复杂、地质条件多变的山区地带(谢春庆等,2015)。受地形起伏影响,这类山区机场施工通常需进行大规模土石方开挖与填筑(王巧等,2012;潘凯等,2017)。因此,确定土石界面对地基承载力计算(仇志龙等,2025)、场地安全(苏正阳等,2025)和优化工程量估算、施工成本控制等具有关键意义(Zhang等,2010)。

开展山区机场工程勘察时,钻探仍是获取地层信息的主要手段;钻孔资料可用于确定土层厚度、基岩埋深及土石分界点位置,为土石界面预测提供依据(谢春庆等,2013)。工程上通常采用直线连接钻孔揭示的土石分界点以重建土石界面。但是考虑到钻探成本和地形条件,实际工程中钻孔一般稀疏布置(Bai等,2024),进而使上述土石界面重建方法的预测准确性受限。因此,在有限钻孔条件下合理预测复杂山区土石界面是亟待解决的难题。

为此,文献中提出并应用了多种插值方法。Samui等(2015)比较了自适应神经模糊推理系统(ANFIS)、多元自适应样条回归(MARS)、普通克里金(OK)和反距离加权法(IDW)的岩层深度预测准确性;Li等(2016)基于条件随机场(CRF)方法预测岩层深度,可量化结果不确定性;Qi等(2020)开展了二维剖面内土石界线预测研究,对比了多项式回归、样条回归、样

条插值法以及条件随机场法；Qi 等（2021，2022）进一步开展了三维空间内土石界面的预测研究，对比了 MARS、CRF 与薄板样条插值法（TPSI）；Deng 等（2023）采用高斯过程回归（GPR）方法预测了岩层深度。表 1 汇总了前述文献的关键信息，包括钻孔数量、场地高程差、均方根误差（RMSE）。由表 1 可见，现有研究主要针对高程差较小（ ≤ 64.1 m）的场地。土石界面预测的准确性主要受钻孔数量（或密度）与场地高程差共同影响：RMSE 整体上随钻孔密度增加而降低、随高程差增大而升高。因此，在地形起伏剧烈且钻孔分布稀疏的山区机场场地，土石界面的准确预测尚未得到有效解决。

表 1 文献中土石界面预测方法汇总

文献来源	应用场景	预测方法	场地面积 (m ²)	钻孔 (个)	高程差 (m)	均方根误差 RMSE(m)
Samui (2015)	区域勘察	ANFIS	14000×14000	652	/	8.8
Samui (2015)	区域勘察	IDW	/	/	/	6.7
Li (2016)	工程勘察	CRF	55×59.3	49	9.4	1.0
Qi (2020)	线路勘察	MARS	3300×120	173	60	9.3
Qi (2021)	线路勘察	MARS	1200×400	154	25.6	4.4
Qi (2021)	工程勘察	MARS	550×350	135	54.4	9.2
Qi (2022)	工程勘察	CRF	550×350	135	54.4	7.1
Qi (2022)	线路勘察	TPSI	500×1000	47	64.1	12.1
Deng (2023)	工程勘察	GPR	160×55	87	/	2.3

虽然在山区机场场地勘察过程中钻孔数量有限，但遥感与地面测量可获取高精度数字高程模型 DEM，为研究区提供连续、致密的地表高程信息。已有研究表明岩土体界面高程与地表高程具有较强相关性（Cordonnier 等，2017）。除钻孔外，地表高程数据（如 DEM）可作为先验信息用于土石界面预测。据此，本文将 DEM 与稀疏钻孔联合用于土石界面预测，通过引入地形先验约束，缓解钻孔稀疏性并提升预测准确性。

本文提出了一种基于地形高程与界面深度插值策略：先预测土石界面深度，随后从 DEM 提取预测点高程，两者相减得到界面高程，从而在一定程度上降低地形起伏对预测的影响。为评估所提策略的预测效果，本文将其与传统基于界面高程的插值方法进行比较。在反距离加权（IDW）、径向基函数（RBF）回归、高斯过程回归（GPR）、多层感知机（MLP）与 Kolmogorov–Arnold 网络（KAN）5 种典型预测模型中分别应用两种策略，并基于山区机场工程数据对两种策略的预测表现进行对比分析。本文结构安排如下：先阐述所提土石界面插值策略，再介绍预测方法及对应

评价指标，继而开展工程案例研究，最后总结并提出应用建议。

1.土石界面插值策略

1.1 基于界面高程的插值策略

在工程实践中，土石分界的判定通常依托钻孔资料，并结合相关规范所规定的室内试验、原位测试结果及土石材料分级标准加以确定。例如，《民用机场勘测规范》(MH/T5025)则按照开挖难易程度和工程性质对材料进行分级分类（见表2）。如图1所示，传统的土石界面预测流程分为两步：首先基于土体分类规范确定钻孔位置的土石分界点高程；然后选择适当预测方法预测未知区域的土石分界点高程，直接生成密集土石分界点。本研究将其称为“基于界面高程的插值策略”。然而，在山区机场工程应用中，因钻孔稀疏且场地高程差显著，该策略的预测准确性难以保障。

表2 土、石按开挖难易程度分级规则（《民用机场勘测规范(GB50021)》）

土石等级	土石类别	代表性土、石名称	开挖难易程度
I	松土	植物土、中密或松散的砂土和粉土、软塑的粘性土。	用铁锹挖，脚蹬一下到底的松散土层
II	普通土	稍密或松散的碎石土（不包括块石或漂石）、密实的砂土和粉土、可塑的粘性土。	部分用镐刨松，再用锹挖，以脚蹬锹需连蹬数次才能挖动
III	硬土	中密的碎石土、硬塑粘性土、风化成土块的岩石。	必须用镐整个刨过才能用锹挖
IV	软石	块石或漂石碎石土、泥岩、泥质砂岩、弱胶结砾岩，中风化~强风化的坚硬岩或较硬岩。	部分用撬棍或十字镐及大锤开挖，部分用爆破法开挖
V	次坚石	砂岩、硅质页岩、微风化~中等风化的灰岩、玄武岩、花岗岩、正长岩。	用爆破法开挖
VI	坚石	未风化~微风化的玄武岩、石灰岩、白云岩、大理岩、石英岩、闪长岩、花岗岩、正长岩、硅质砾岩等。	用爆破法开挖

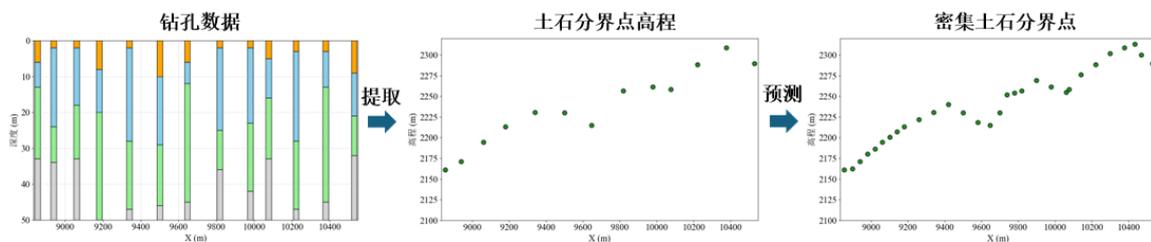


图1. 基于界面高程插值策略流程图

1.2 基于地形高程与界面深度的插值策略

为了克服基于界面高程插值策略的局限性，本文提出了一种基于地形高程与界面深度的插值策略（见图 2）。所提策略分为三步：先从钻孔数据提取土石界面深度，然后预测未知区域的土石界面深度，最后以预测点的地形高程（来自 DEM）减去预测深度，获得相应位置的土石界面高程，从而生成密集的界面点集。相比较于传统高程插值策略，所提策略创新之处在于：（1）土石界面深度预测过程不依赖界面高程数据，一定程度降低地形起伏对预测准确性的影响；（2）充分利用地表高程信息，在稀疏钻孔条件下提升预测准确性。

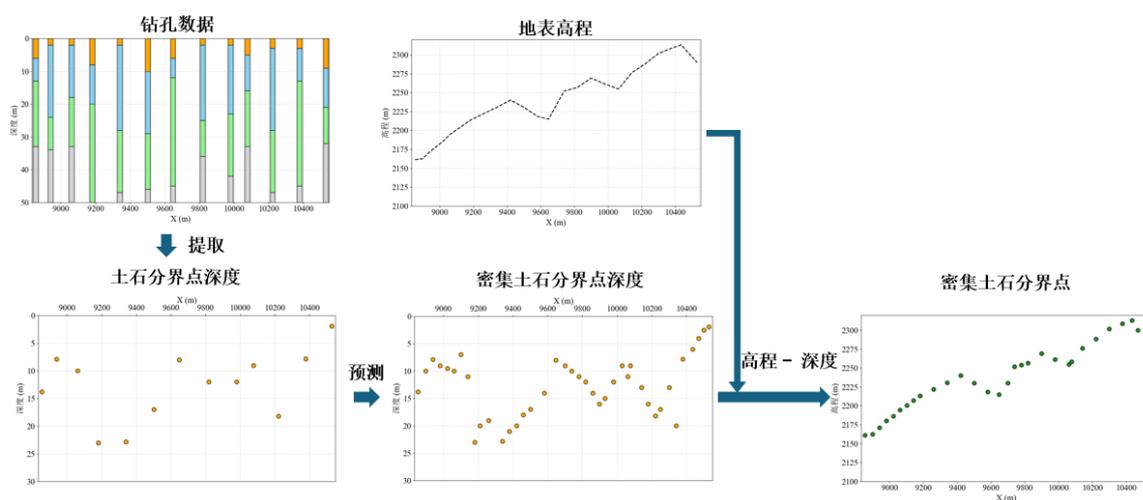


图 2 基于地形高程与界面深度插值策略流程图

2. 土石界面预测方法与评估指标

为系统评估所提插值策略，本研究选取了文献中多种典型土石界面预测方法，并采用三种评估指标，开展所提插值策略和基于界面高程插值策略的对比研究。下文依次介绍插值方法与评价指标。

2.1 预测方法

考虑不同预测方法的建模范式差异，本文选取了两类方法：（1）以核方法为代表的非参数方法；（2）以神经网络为代表的参数化回归方法（张翰等，2024）。核方法以核函数度量待预测位置与钻孔的相似性，据此计算权重，对钻孔处的土石界面高程（或深度）进行加权平均，以估计预测位置的土石界面高程（或深度）；常见实现包括反距离加权（IDW）、径向基函数（RBF）回归和高斯过程回归（GPR）。神经网络方法则基于可学习的参数化结构，通过迭代训练不断优化权重或函数模块，从而逼近输入特征与土石界面高程（或深度）之间的复杂非线性关系；典型方法包括多层感知机（MLP）和 Kolmogorov-Arnold 网络（KAN）。

2.1.1 核方法

核方法用于预测任务时，对应预测方程表达式如下：

$$\hat{y} = \sum_{i=1}^n \omega(\hat{\mathbf{x}}, \mathbf{x}_i) y_i \quad (1)$$

式中 $\hat{\mathbf{x}}$ 为预测点特征，本研究特指预测点的平面坐标； \mathbf{x}_i 为第 i 个观测点（钻孔）的特征，即其平面坐标； \hat{y} 为预测值，代表预测点的高程（或深度）； y_i 为第 i 个观测点的土石界面高程（或深度）； $\omega(\hat{\mathbf{x}}, \mathbf{x}_i)$ 为权重函数；本文将单次预测所用的观测点数记为 n ：在 IDW 中， $n=k$ （目标点的近邻观测点数）；在 RBF 回归和 GPR 中， $n=N$ （全部观测点数）。

反距离加权法（IDW）采用反距离幂次权重函数（Shepard, 1968; 李章林等, 2015; Chen 等, 2015），记作 $\omega_l(\hat{\mathbf{x}}, \mathbf{x}_i)$ ，其表达式如下：

$$\omega_l(\hat{\mathbf{x}}, \mathbf{x}_i) = \left(\frac{1}{\|\hat{\mathbf{x}} - \mathbf{x}_i\|^p} \right) / \sum_{i=1}^k \left(\frac{1}{\|\hat{\mathbf{x}} - \mathbf{x}_i\|^p} \right) \quad (2)$$

式中， $\|\hat{\mathbf{x}} - \mathbf{x}_i\|$ 为预测点和其近邻观测点之间的欧式距离（ $i=1, 2, \dots, N$ ）； p 为幂指数，可以是任意正实数或 0。

径向基函数（RBF）回归采用径向基函数核（Murphy, 2012、2022），其权重函数 $\omega_r(\hat{\mathbf{x}}, \mathbf{x}_i)$ 定为：

$$\omega_r(\hat{\mathbf{x}}, \mathbf{x}_i) = \left(\exp \left[\frac{-(\hat{\mathbf{x}} - \mathbf{x}_i)^2}{2\lambda^2} \right] \right) / \sum_{i=1}^N \left(\exp \left[\frac{-(\hat{\mathbf{x}} - \mathbf{x}_i)^2}{2\lambda^2} \right] \right) \quad (3)$$

式中， λ 为带宽，带宽越小，权重主要分配给距离预测点较近的观测点；其他符号的含义参考公式（2）。

高斯过程回归（GPR）作为一种贝叶斯非参数方法，不仅能够预测土石界面高程或深度，还能量化预测的不确定性（Wang, 2021）。根据高斯过程（GP）定义，假设预测点高程（或深度） y^* 和所有观测点高程（或深度）向量 \mathbf{y} 服从如下联合高斯分布：

$$\begin{pmatrix} \mathbf{y} \\ y^* \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu} \\ \mu^* \end{pmatrix}, \begin{pmatrix} \mathbf{K}_{\mathbf{X}, \mathbf{X}} & \mathbf{K}_{*, \mathbf{X}} \\ \mathbf{K}_{*, \mathbf{X}}^T & k_{**} \end{pmatrix} \right) \quad (4)$$

其中， $\mathbf{y} = [y_1, y_2, \dots, y_N]^T$ ； $\boldsymbol{\mu}$ 与 μ^* 分别为观测点的均值向量和预测点均值，一般假设为 0； $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ ， $\mathbf{K}_{\mathbf{X}, \mathbf{X}} = \mathcal{K}(\mathbf{X}, \mathbf{X})$ ， $\mathbf{K}_{*, \mathbf{X}} = \mathcal{K}(\hat{\mathbf{x}}, \mathbf{X})$ ， $k_{**} = \mathcal{K}(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ ， \mathcal{K} 代表正定核函数，用于计算两点间协方差。本文选取一种常用的 \mathcal{K} ，即高斯核函数

$$\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = C \exp \left[\frac{-(\mathbf{x}_i - \mathbf{x}_j)^2}{2l^2} \right] + \sigma_n^2 \quad (5)$$

其中, C 为核函数方差, 表示函数在整体上的变化幅度; σ_n^2 为噪声方差; l 为长度尺度, 控制 \mathcal{K} 的宽度。按高斯分布条件化的标准规则 (Murphy, 2012), y^* 的后验分布仍是高斯分布:

$$p(y^* | \hat{\mathbf{x}}, \mathbf{X}, \mathbf{y}) = \mathcal{N}(y^* | \mathbf{K}_{*,\hat{\mathbf{x}}}^T \mathbf{K}_{\mathbf{X},\mathbf{X}}^{-1} \mathbf{y}, k_{**} - \mathbf{K}_{\hat{\mathbf{x}},\hat{\mathbf{x}}}^{-1} \mathbf{K}_{\hat{\mathbf{x}},\mathbf{X}}) \quad (6)$$

因此, y^* 的后验最可能值 (\hat{y}) 等于其后验分布均值, 即

$$\hat{y} = \mathbf{K}_{*,\hat{\mathbf{x}}}^T \mathbf{K}_{\mathbf{X},\mathbf{X}}^{-1} \mathbf{y} = \sum_{i=1}^N (\mathbf{K}_{*,\hat{\mathbf{x}}}^T \mathbf{K}_{\mathbf{X},\mathbf{x}_i}^{-1}) y_i \quad (7)$$

其中, $\mathbf{K}_{\mathbf{X},\mathbf{x}_i} = \mathcal{K}(\mathbf{X}, \mathbf{x}_i)$ 。令 $\omega_G(\hat{\mathbf{x}}, \mathbf{x}_i)$ 代表 GPR 预测方程的权重函数, 则 $\omega_G(\hat{\mathbf{x}}, \mathbf{x}_i) = \mathbf{K}_{*,\hat{\mathbf{x}}}^T \mathbf{K}_{\mathbf{X},\mathbf{x}_i}^{-1}$ 。

2.1.2 神经网络方法

多层感知机 (MLP) 通过多层非线性映射和反向传播优化参数, 用于刻画复杂输入输出关系 (杜昕等, 2020)。该网络包含输入层、输出层以及多个隐含层, 以神经元为最小单位, 将输入的多个节点信息非线性映射到单一输出上。 $f(x)$ 代表激活函数, 可以为 Sigmoid、Tanh 或者 ReLU 函数中的一种。与第一层神经元的解析表达式:

$$z_j^1 = f \left[\sum_{t=1}^{N_x} (\omega_{t,j}^1 x_t + b_{t,j}^1) \right] \quad (8)$$

式中, z_j^1 为第 1 个隐含层上的第 j 个神经元, N_x 为输入层维度, x_t 为特征 \mathbf{x} 中第 t 个分量, $\omega_{t,j}^1$ 是输入层第 t 个分量与第 1 个隐含层上第 j 个神经元的连接权重, $b_{t,j}^1$ 为偏置常数。向后继续传播时, 以 z_j^1 作为新的输入, 对式 (8) 依次递归, 由式 (9) 进一步得到输出层的解析表达式:

$$\hat{y} = f \left[\sum_{t=1}^{N_{M-1}} (\omega_{t,j}^M z_t^M + b_{t,j}^M) \right] \quad (9)$$

式中, \hat{y} 代表输出层的预测值, M 为神经网络隐含层总数, N_{M-1} 代表第 $M-1$ 个隐含层包含的神经元总数。

与 MLP 不同, Kolmogorov-Arnold 网络 (KAN) 将用被参数化的函数代替权重参数 (Liu 等, 2024; Somvanshi 等, 2024)。KAN 中的每一条连接可表示为单变量连续函数加和的形式, 通过一组参数化的基函数 (样条函数、多项式函数等) 对这些单变量进行近似。KAN 网络输出层的解析表达式可表示为:

$$\hat{y} = f \left[\sum_{j=1}^{2N_x+1} \Phi_j \left(\sum_{t=1}^{N_x} \phi_{t,j}(x_t) \right) \right] \quad (10)$$

式中, $\phi_{t,j}$ 是对输入做非线性变换的单变量函数, Φ_j 表示在隐藏层做进一步变换的单变量函数。

2.2 评估指标

为全面、客观地评估不同预测方法的表现, 本文采用引入均方根误差 (RMSE)、归一化均

方根误差 (nRMSE) 与决定系数 (R^2) 三项评估指标。RMSE 表示预测值与真实值的平均绝对误差, 其计算公式为:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (11)$$

式中, N 为预测点个数, \hat{y} 为预测值, y 为已知值。由于研究山区地形起伏较大, RMSE 可能难以全面反映预测准确性的相对水平。为此, 本文提出归一化均方根误差 (nRMSE), 即

$$nRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} / \Delta h \quad (12)$$

式中, Δh 为最大地形高差。nRMSE 将 Δh 作为归一化因子对 RMSE 进行标准化, 从而更客观地比较高差较大场景下不同方法的误差水平。 R^2 表示模型对观测数据变异性的解释能力, 其计算公式为:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (13)$$

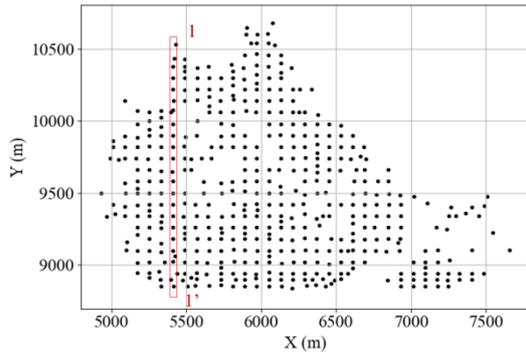
式中, \bar{y} 为观测数据的均值。 R^2 值越接近 1, 说明模型拟合效果越好。

综上所述, 本节简要介绍了典型土石界面预测方法与评价, 下文将通过工程案例验证所提基于地形高程与界面深度的插值策略。

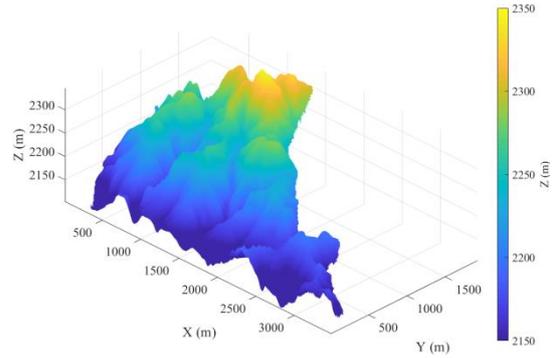
3. 工程案例——某山区机场工程

3.1 案例简介

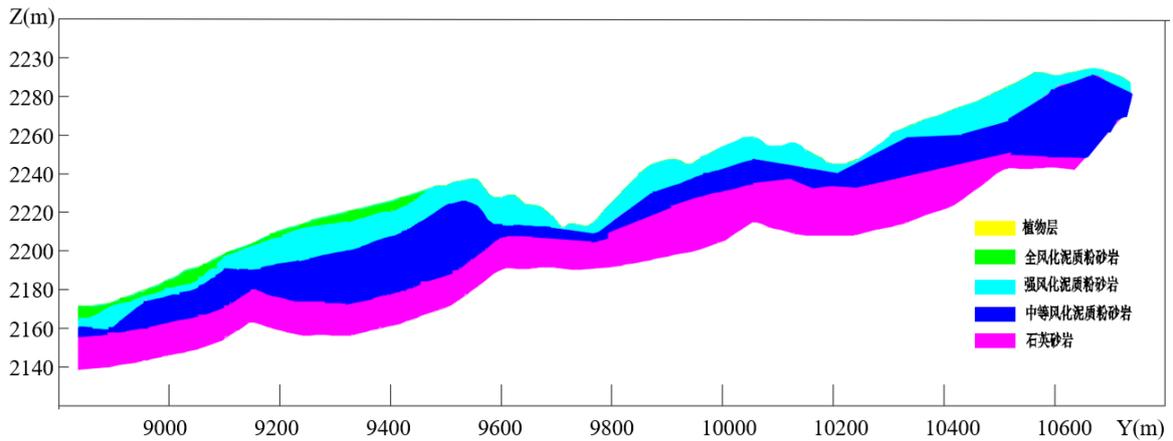
该机场位于我国西南山区, 场地范围约 2000 m × 3500 m。图 3 (a) 显示了研究区域的钻孔布置, 共计 473 个钻孔。由于钻孔分布稀疏, 仅基于钻孔资料难以准确重构土石界面。图 3 (b) 为该场地数字高程模型 (DEM), 其显示地表高程分布在 2100 m - 2350 m 之间, 最大高差约 250 m。综合钻孔和 DEM, 该场地土石界面高程范围为 2120 m-2322 m, 标准差 52 m; 对所有钻孔的地表高程与对应土石界面高程进行相关分析, 相关系数为 0.99, 表明二者相关性极强。因此, 本文在土石界面预测过程中引入地表高程信息, 以约束预测结果。图 3 (c) 为剖面 1-1' (见图 3a) 的工程地质剖面图。



(a) 钻孔分布图



(b) 数字高程模型



(c) 典型工程地质剖面图

图 3 研究区域的钻孔与地表高程信息及典型地质剖面

3.2 二维土石界线预测

本文首先在纵剖面 1-1' 上进行土石界线（一维）插值，用于结果展示与方法比较。由于地形条件限制，1-1' 剖面内钻孔的 X 坐标并不完全一致，但其最大偏差不超过 10m，而剖面沿 Y 方向的长度约为 1600 m。与剖面长度相比，该偏移量很小，因此在工程上可认为这些钻孔位于同一剖面。图 4 展示了剖面 1-1' 内 25 个真实钻孔数据的岩性分布与土石分界点深度分布。强风化泥质粉砂岩与中等风化泥质粉砂岩的分界线代表土石分界线，其深度范围为 1.9 - 25.0 m，标准差 5.9 m。基于图 4 所示土石分界点数据，本研究分别采用两种插值策略，并结合不同预测方法（即 IDW、RBF 回归、GPR、MLP 与 KAN）预测剖面 1-1' 内的土石分界线。

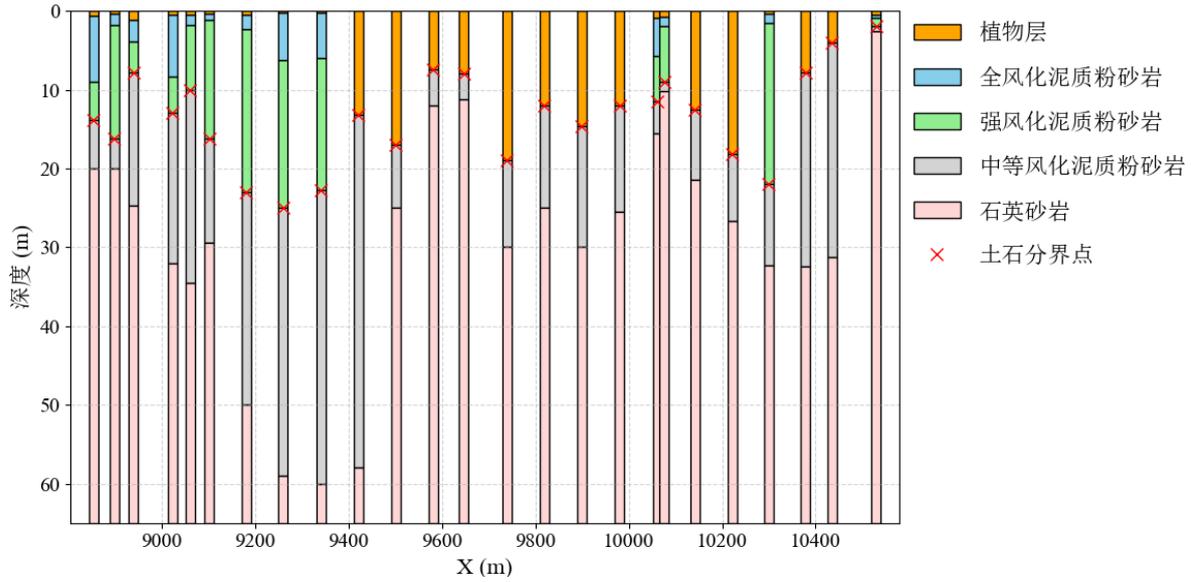


图 4 剖面 1-1'内各钻孔岩性与土石分界点

本文在预设参数网络上采用留一交叉验证 (LOOCV) 对各预测方法 (GPR 除外) 进行参数选择与性能评估: 以 LOOCV 的 RMSE 最小为准则确定最优参数配置, 并在该最优参数配置下计算 nRMSE 与 R^2 。GPR 为特例, 其最优参数采用最大化对数边缘似然直接估计, 随后在该最优参数配置下, 用 LOOCV 计算 RMSE、nRMSE 与 R^2 。表 3 给出了各方法的最优参数配置。其中, k 此外, GPR 噪声方差极小, 对结果无显著影响, 因此未单独列出。MLP 激活函数为 ReLU 函数; KAN 学习函数为 B-样条函数。神经网络方法的输入特征为平面 X 和 Y 坐标与地表高程, 输出为界面高程或深度。为了防止神经网络方法过拟合, 在所有全连接层加入了 L2 正则化项并采用早停机制。

表 3 二维土石界线插值各方法最优参数

	IDW	RBF 回归	GPR	MLP	KAN
预测高程	$k=7; p=2.5$	$\lambda=44$	$C=0.91^2; l=180$	$M=2; N_1=N_2=64$	$M=1; N_1=64$
预测深度	$k=5; p=3.5$	$\lambda=40$	$C=0.91^2; l=95.8$	$M=2; N_1=N_2=64$	$M=1; N_1=64$

注: k 为临近点个数, p 为幂指数; λ 为带宽; C 为核函数方差, l 为长度尺度; M 为隐藏层层数, N_i 为第 i 层神经元的个数。

图 5 分别展示了两种插值策略下五种预测方法的留一交叉验证结果, 即钻孔处土石界面高程预测值。所提插值策略下, 每种预测方法结果均接近对应真实值, 其中未发现预测值超出地表的情况。而基于界面高程插值策略下, 所有预测方法在 1-1' 剖面边界附近钻孔处土石界面高程预测值超出地表; 此外, 神经网络模型 (MLP 和 KAN 网络) 只能捕捉土石界面的整体趋势, 难以准确反映其局部起伏。图 5 结果说明了所提插值策略通过引入地形约束, 提升了预测的合理性。

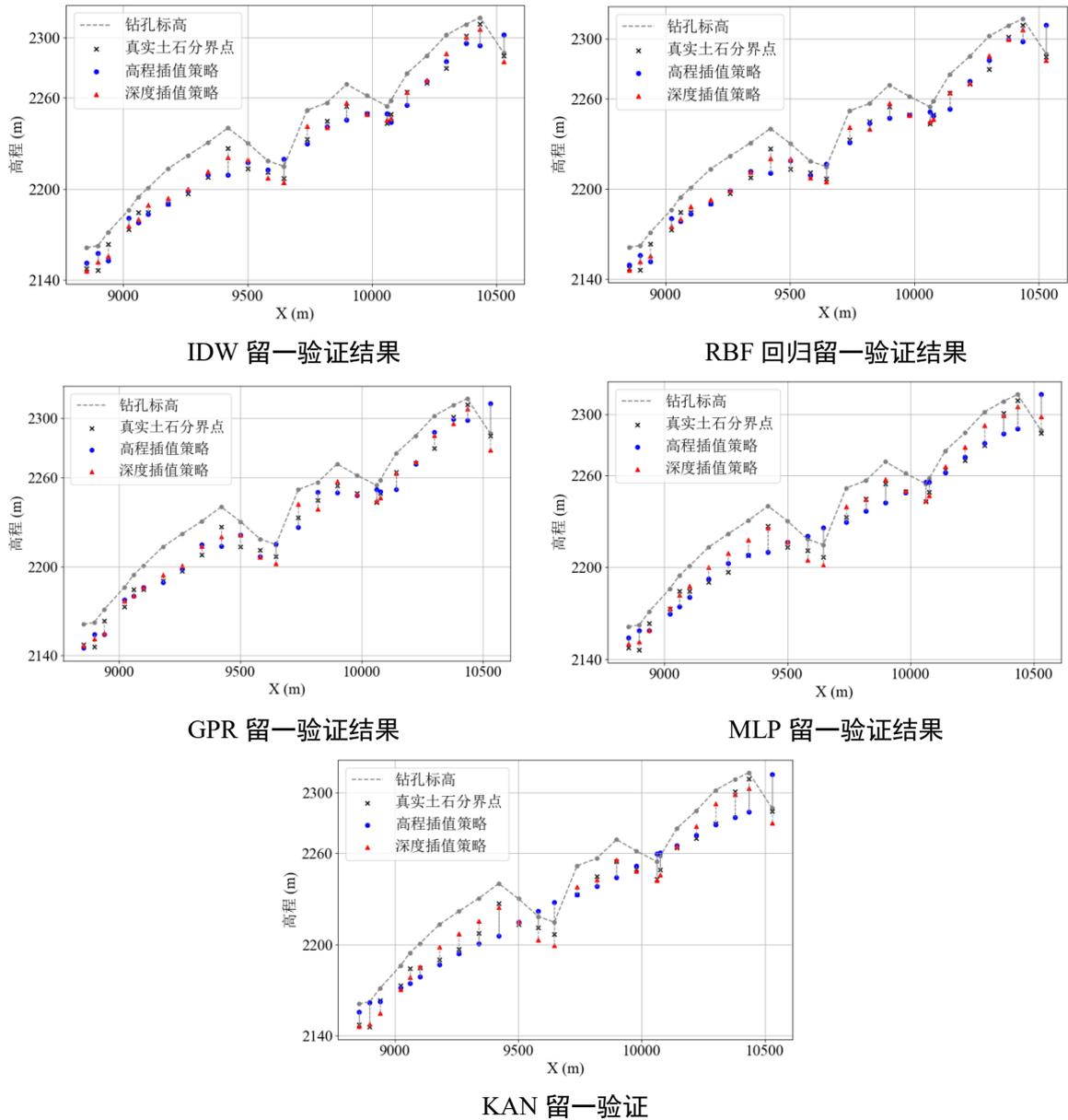


图 5 五种方法在剖面 1-1' 的留一验证结果

为了进一步定量评估预测方法，图 6 给出了两种插值策略下不同方法的留一验证 RMSE（或 nRMSE）对比结果。与传统基于界面高程的插值策略相比，采用所提策略后，各预测方法的 RMSE 均有所降低，降幅在 38% - 47% 之间（视具体预测方法而定）。nRMSE 的对比结果亦与 RMSE 保持一致，各预测方法均呈现相同的性能提升趋势。图 6 结果说明所提插值策略能够显著提升预测方法的准确性，其先预测土石界面深度，再由地形高程减去深度预测值估计土石界面高程，有助于降低地形起伏对预测的影响。表 4 给出了两种插值策略下不同方法的 R^2 指标。所提插值策略下各方法预测结果的 R^2 均大于基于界面高程的插值策略，表明所提插值策略能够提升各预测结果的拟合优度。

需要指出的是，不同评估指标在土石界面预测中侧重点各异。在本研究中 R^2 指标普遍较高

(>0.9)，说明各策略与各方法均能较好地反映整体趋势，但无法揭示复杂地形下的绝对误差。相比之下，RMSE 表示绝对误差，但其受地形起伏影响显著，在高差较大的山区场地往往数值较大，而在高差较小的场地即使预测准确度不高，RMSE 也可能较小，因此有其局限性。而相对误差指标 nRMSE 在基于最大地形高差对 RMSE 进行了归一化，因此适合用于不同场地的预测结果对比。由此可见，单一指标难以全面评价预测效果，需结合多种指标并结合工程实际背景进行综合判断。

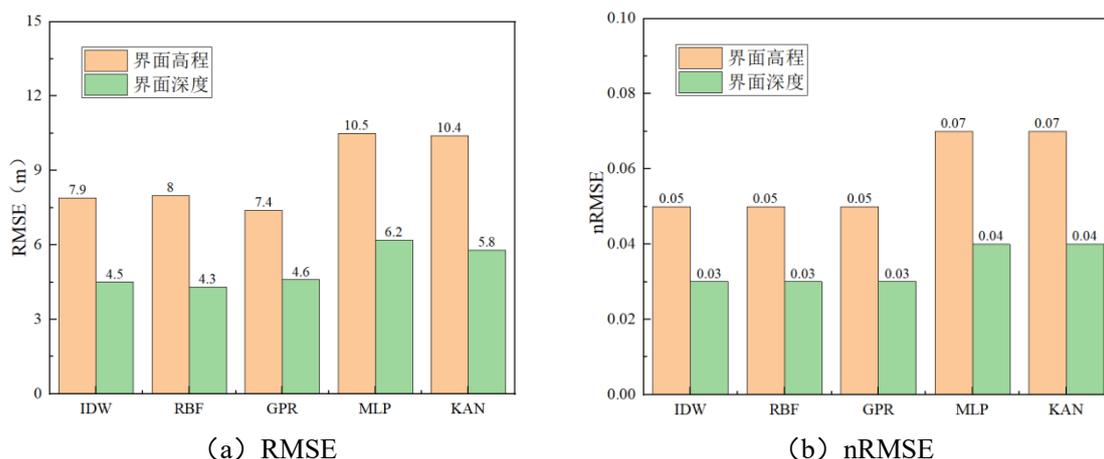


图 6 二维土石界线预测误差

表 4 二维土石界线预测结果 R^2

	IDW	RBF 回归	GPR	MLP	KAN
预测高程	0.97	0.97	0.97	0.95	0.93
预测深度	0.99	0.99	0.99	0.98	0.98

在剖面 1-1' 所在区域内，基于 DEM 选取 1734 个预测点（该处界面深度未知），在所提插值策略下，分别采用 IDW、RBF 回归、GPR、MLP、KAN 对这些点的土石界面深度进行估计。作为对比，给定基于界面高程插值策略，在相同位置直接预测土石界面高程。图 7 显示了剖面 1-1' 土石界线预测结果。对非参数方法（IDW、RBF 回归、GPR），无论采用哪种插值策略，预测结果均能较好刻画土石界线的空间起伏。相比之下，对参数化方法（MLP 和 KAN），在基于界面高程的策略下，结果整体呈平滑上升趋势，难以准确表征局部起伏，这可能与样本量较小 ($N=25$) 有关。所提插值策略将地形信息作为先验约束引入预测过程，在一定程度上缓解了数据稀缺导致的参数化方法对局部起伏欠拟合问题。

GPR 方法不仅可以给出平均预测结果，还可量化预测结果的不确定性。图 7 展示了 GPR 预测结果的 95% 贝叶斯可信区间 (95% Bayesian credible interval)，即不同色带，其宽度 $= 2 \times 1.96 \sigma_p$ 表示预测结果的不确定性大小； σ_p 代表预测标准差。结果显示所提插值策略的色带宽度明显更

窄，25个预测点的标准差平均值从5.9m减小为3.9m，由此说明所提插值策略有助于降低土石界面预测的不确定性，从而提高了预测结果的可靠性。

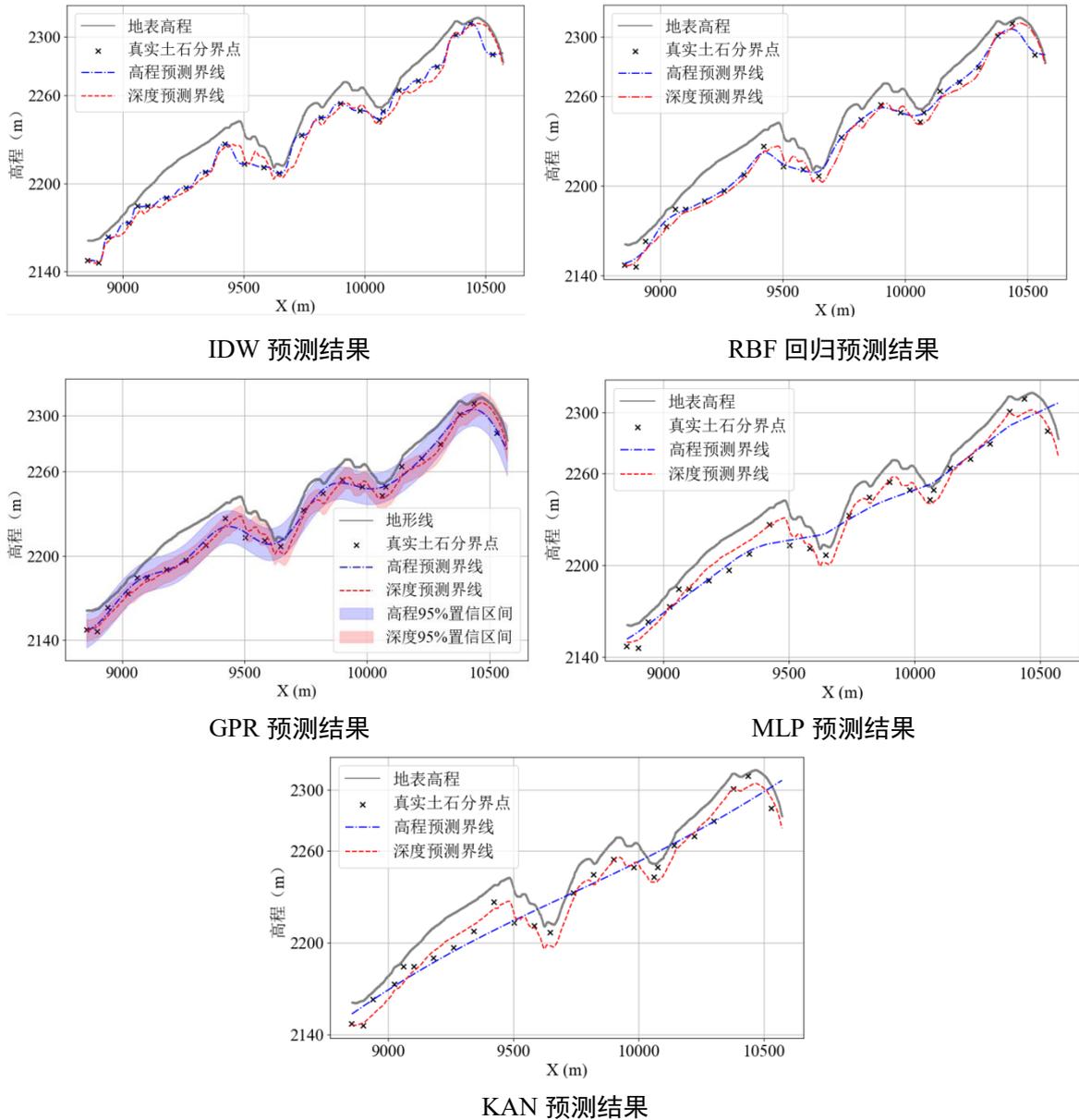


图7 五种方法在剖面 1-1' 的预测结果

3.3 三维土石界面预测

为评估所提插值策略在三维场景中的适用性，本文将研究区域内 473 个钻孔的土石界面点作为观测数据，进行土石界面预测。这些土石界面点深度范围约 0.6 - 42.0 m，标准差等于 6.5 m。本研究按照 8: 2 的比例划分了训练数据集与测试集，并在训练数据集内进行五折交叉验证。将训练数据集随机划分为 5 个大小相同的互斥子集。然后，每次用 4 个子集的并集作为训练集，余下的子集作为验证集。本研究的重点在于比较不同模型的预测性能，因此针对各模型均采用相同

的训练/验证/测试集划分，并进行一次完整的交叉验证，确定最优模型参数；将该最优模型参数对应的预测值作为模型的预测结果，以此计算评估指标 RMSE、nRMSE 和 R^2 。预测方法 IDW、RBF 回归、GPR、MLP、KAN 对应的参数寻优策略参考 3.2 节。表 5 展示了不同预测方法的最优参数配置，MLP 激活函数为 ReLU 函数，KAN 学习函数为 B-样条函数。

表 5 三维土石界面插值各方法最优参数

	IDW	RBF 回归	GPR	MLP	KAN
预测高程	$k=6; p=2.5$	$\lambda=50$	$C=0.71^2; l=86.4$	$M=2; N_1=128; N_2=64$	$M=1; N_1=128$
预测深度	$k=8; p=1.5$	$\lambda=70$	$C=0.99^2; l=45.2$	$M=2; N_1=64; N_2=32$	$M=1; N_1=64$

注： k 为临近点个数， p 为幂指数； λ 为带宽； C 为核函数方差， l 为长度尺度； M 为隐藏层数， N_i 为第 i 层神经元的个数。

图 8 展示了在两种插值策略下，各预测方法 RMSE（或 nRMSE）的对比图。表 6 则列出了每种预测方法的预测性能指标 R^2 。在所提插值策略下，每种预测方法的 RMSE 均在 6.5 m 左右；这些结果均小于对应基于界面高程的插值策略，并且 RMSE 降幅在 30%-57% 之间（视方法而定）。nRMSE 结果也呈现类似趋势。此外，所提插值策略下 R^2 计算结果整体大于基于界面高程的插值策略。上述结果再次表明，所提插值策略能够提升预测结果的准确性。

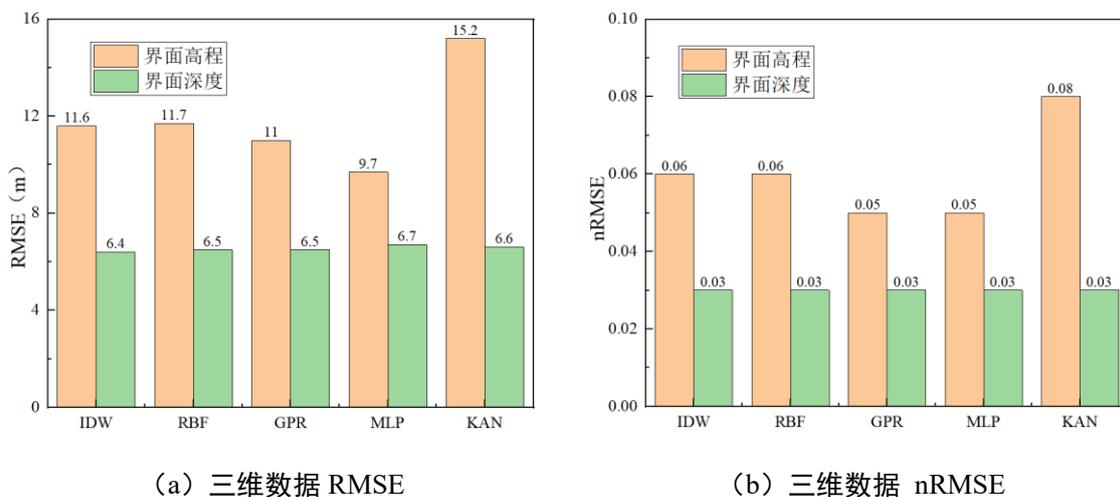


图 8 三维土石界面预测误差

表 6 三维土石界面预测性能指标 R^2

	IDW	RBF 回归	GPR	MLP	KAN
预测高程	0.95	0.95	0.96	0.97	0.92
预测深度	0.98	0.98	0.98	0.99	0.99

为全面评估所提插值策略，本文将第 3.2 节的二维案例与本节的三维案例进行了对比。结果

显示给定所提插值策略条件下，在三维案例中各预测方法的 RMSE 相较二维案例有所增大，但经由最大地形高差归一化后的 nRMSE 以及决定系数 R^2 基本保持稳定。这表明所提策略能够在不同地形起伏与数据变异性条件下维持一致的预测性能，具有良好的普适性。因此，单独使用 RMSE 不足以全面评价插值策略，需结合 nRMSE 与 R^2 等指标开展综合评估。作为对照，在基于界面高程的传统插值策略下，当研究区由二维拓展至三维时，各预测方法的结果更易出现 nRMSE 升高或 R^2 降低的趋势，表明传统插值策略对最大高程差与数据变异性的敏感性更强，从而进一步凸显了所提策略的普适性。

由于所提插值策略下，IDW、RBF 回归、GPR、MLP、KAN 方法的预测误差以及 R^2 总体一致，本文仅以 IDW 为例展示场地三维土石界面预测结果。在区域内的 DEM 上选取了 18000 预测点，并将 IDW 的参数设置为其最优参数配置（见表 6），IDW 在不同插值策略下的预测结果如图 9 所示。图中红色实心点标示了界面预测结果超过地表高程的异常位置。采用所提插值策略时，IDW 预测结果中未发现超出地表点，较好刻画了土石界面空间起伏。相反，采用基于界面高程的插值策略，在场地边界以及地形起伏较大区域出现众多超出地表的预测点。由此可见，所提插值策略在三维土石界面预测中提升了结果准确性以及合理性。

在该三维场景中，钻孔地表高程与对应土石界面高程的相关系数为 0.99；地表最大高差和标准差分别为 212.4 m 和 50.4 m，而界面深度的最大差值和标准差仅为 41.4 m 和 6.5 m。由此可见，界面深度的变化幅度明显小于界面高程。所提策略在工程中取得较高准确性主要源于两点：（1）将预测对象由界面高程转为界面深度，可有效降低预测难度，从而提升预测精度；（2）界面高程由“地表高程 - 预测深度”得到，地表高程天然作为土石界面高程的上限约束，能确保预测界面始终位于地表以下，使结果更加合理。基于上述原因，该策略能够在复杂地形条件下保持较高的预测准确性，可为机场选址、土石方工程量估算及施工方案优化等提供可靠的技术支持。

本研究中地表高程与界面高程高度相关，在该情况下所提策略优势尤为显著。然而，该策略的适用性在一定程度上依赖于地形对界面起伏的控制作用。当地表高程与界面高程相关性减弱时（如岩溶发育区、断裂破碎带），该策略有效性和稳定性有待验证。未来研究可从以下方面展开：进一步评估该策略在地表高程与界面高程弱相关或不相关条件下的有效性，考虑引入更多反映地下结构的先验信息（如物探数据等）以增强稳定性，也可以考虑基于地质构造进行分区预测，以提升策略在复杂地质条件下的适用性和可靠性。

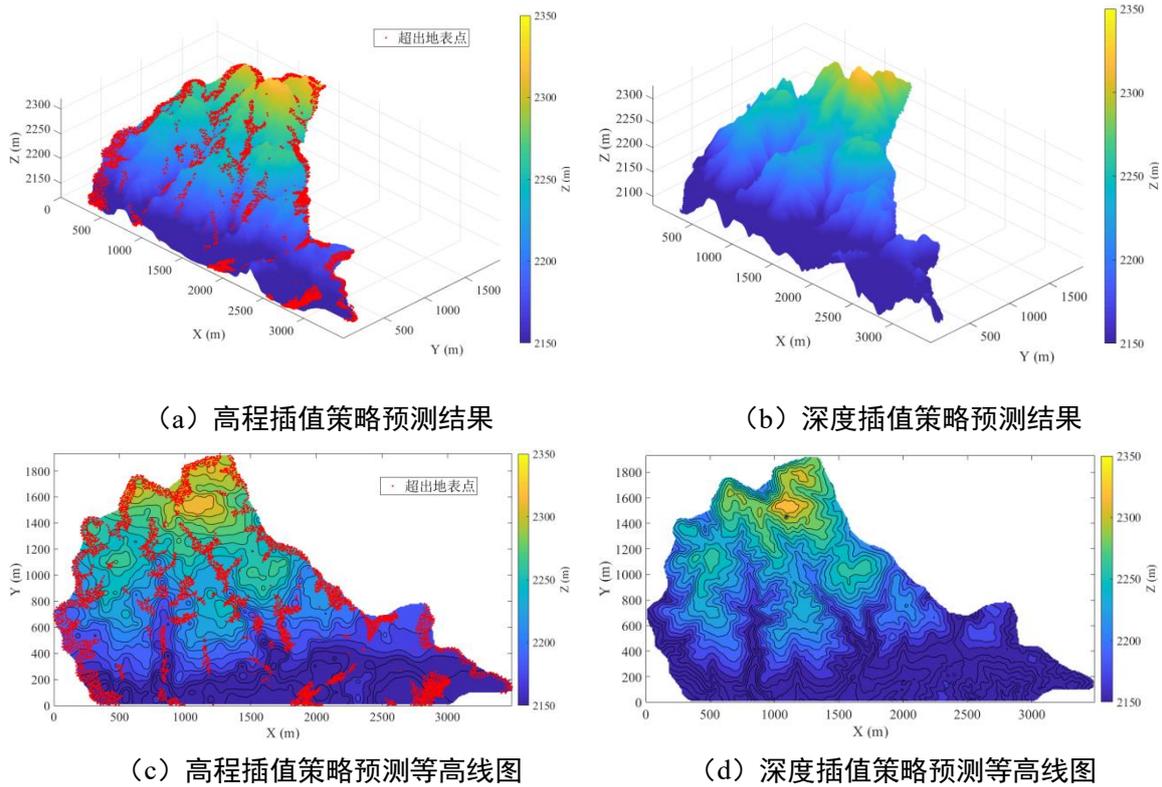


图9 三维土石界面预测结果

4.结论

本文针对山区复杂地形条件下的土石界面预测问题，提出了一种基于地形高程与界面深度的插值策略，并将该策略用于 IDW、RBF 回归、GPR、MLP 和 KAN 五种预测方法，通过 RMSE、nRMSE 和 R^2 三种评估指标系统评估了该策略的应用效果。采用山区机场工程案例验证了所提策略，主要结论如下：

(1) 与传统基于界面高程的插值策略相比，本文提出的基于地形高程与界面深度的插值策略先预测土石界面深度，再由地表高程减去预测深度估计土石界面高程，能够降低山区地形剧烈起伏对预测结果影响。工程案例结果表明，所提策略提升了预测准确性，预测结果更准确地刻画了土石界面地空间起伏特征，为机场选址、土石方工程量估算及施工方案优化提供了技术支撑。

(2) 本研究表明在小样本条件下，核方法（如 IDW、RBF 回归和 GPR）较神经网络方法（如 MLP 和 KAN）的表现更为稳定；随着样本数量增加，两类方法的误差均显著降低并趋于一致。所提插值策略通过引入地表高程作为先验约束，可在小样本条件下一定程度缓解神经网络方法的欠拟合问题，因此具有更广的适用性。

(3) 本文采用的 nRMSE 将 RMSE 相对于研究区最大地形高差进行归一化，提升了在不同地形起伏条件下误差指标的可比性，尤其适用于最大高差差异显著的场地。在对比不同场地（或研究区域）上的预测结果时，本研究建议联合使用多种评估指标（如 nRMSE、RMSE 和 R^2 ，以

确保评价结果的客观性。

5.作者贡献度说明

曹子君：负责核心数据分析与论文初稿撰写，以及基金支持；郑立宁：提供关键性技术指导；赵士博：负责理论框架优化、修改论文；曾雪松：提供研究资源及终稿审阅；王轩毫：提出科学问题，并提供技术指导。

References

- [1] Bai, J., Wang, S., Xu, Q., et al., 2024. Intelligent regional subsurface prediction based on limited borehole data and interpretability stacking technique of ensemble learning. *Bulletin of Engineering Geology and the Environment*, 83(7), 272.
- [2] Chen C, Zhao N, Yue T, et al., 2015. A generalization of inverse distance weighting method via kernel regression and its application to surface modeling. *Arabian Journal of Geosciences*, 8(9): 6623-6633.
- [3] Cordonnier G, Cani M P, Benes B, et al., 2017. Sculpting mountains: Interactive terrain modeling based on subsurface geology. *IEEE transactions on visualization and computer graphics*, 24(5): 1756-1769.
- [4] Du X, Fan T E, Dong J H, et al., 2020. Intelligent seismic facies classification based on an improved deep learning method(in Chinese with English abstract). *Oil Geophysical Prospecting*, 55(06):1178-1187+1159.
- [5] Deng Z P, Pan M, Niu J T, et al., 2023. Spatial prediction of rockhead profile using the Gaussian process regression method. *Canadian Geotechnical Journal*, 60(12): 1849-1860.
- [6] Li X Y, Zhang L M, Li J H., 2016. Using conditioned random field to characterize the variability of geologic profiles. *Journal of Geotechnical and Geoenvironmental Engineering*, 142(4): 04015096.
- [7] Li Z L, Wu C L, Zhang X L, et al., 2015. Uncertainty Assessment for IDW Ore Grade Estimates(in Chinese with English abstract). *Earth Science-Journal of China University of Geosciences*, 40(11):1796-1801.
- [8] Liu Z, Wang Y, Vaidya S, et al., 2024. Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*.
- [9] Specifications for Geotechnical Investigation and Surveying of Airports(in Chinese). MH /T5025—2011[S],2011.
- [10] Murphy K P., 2012. Machine learning: a probabilistic perspective. MIT press.
- [11] Murphy K P., 2022. Probabilistic machine learning: an introduction. MIT press.
- [12] Pan K, Xie C Q, Cheng R Y., et al., 2017. Engineering Example of Comprehensive Prospecting of Underground Karst in Mountain Airport Area(in Chinese with English abstract). *Site Investigation Science and Technology*, (01):28-32.
- [13] Qi X, Pan X, Chiam K, et al., 2020. Comparative spatial predictions of the locations of soil-rock interface. *Engineering Geology*, 272: 105651.
- [14] Qi X, Wang H, Chu J, et al., 2022. Two-dimensional prediction of the interface of geological formations: A

- comparative study. *Tunnelling and Underground Space Technology*, 121: 104329.
- [15] Qi X, Wang H, Pan X, et al., 2021. Prediction of interfaces of geological formations using the multivariate adaptive regression spline method. *Underground Space*, 6(3): 252-266.
- [16] Qiu Z L, Wu J D, Wan P, et al., 2025. Experimental Study on the Influence of Soil-Rock Ratio on the Dynamic Compaction Reinforcement Effect of High Fill Gravel Soil Subgrade. (in Chinese with English abstract) *Chinese Journal of Underground Space and Engineering*, 21(01):123-130.
- [17] Samui P, Kim D, Viswanathan R., 2015. Spatial variability of rock depth using adaptive neuro-fuzzy inference system (ANFIS) and multivariate adaptive regression spline (MARS). *Environmental Earth Sciences*, 73(8): 4265-4272.
- [18] Shepard D., 1968. A two-dimensional interpolation function for irregularly-spaced. *Proceedings of the 1968 23rd ACM national conference*. 517-524.
- [19] Somvanshi S, Javed S A, Islam M M, et al., 2024. A survey on Kolmogorov-Arnold network. *ACM Computing Surveys*.
- [20] Su Z Y, Li D Q, Wang S, et al., 2025. Large deformation analysis of 3D soil-rock mixture slopes using SPH-DEM method(in Chinese with English abstract). *Earth Science*,1-9.
- [21] Wang J, Robotics O., 2021. An intuitive tutorial to Gaussian processes regression. *Stat*, 1050: 2.
- [22] Wang Q, 2011. A Study on Quantitative Research Method about Soil-Rock Ratio in Red-Beds Airports' Excavation Area in Sichuan(in Chinese with English abstract) (Dissertation). Chengdu University of Technology.
- [23] Xie C Q, Li Z Z, Pan K, 2015. Study on Method Determining Soil-Rock Ratio in Large-area Excavation Works in Southwest Karst Mountains(in Chinese with English abstract). *Subgrade Engineering*, (02):1-5.
- [24] Xie C Q, Rong S J, Wang W, 2013. Study on Soil-Rock Ratio of the Projects in Red Bed Area(in Chinese with English abstract). *Subgrade Engineering*, (06):83-88.
- [25] Zhang L M, Dasaka S M., 2010. Uncertainties in geologic profiles versus variability in pile founding depth. *Journal of geotechnical and geoenvironmental engineering*, 136(11): 1475-1488.
- [26] Zhang H, Gui L, Wang T F, et al., 2024. Prediction of Quaternary Cover Thickness and 3D Geological Modeling Based on BP Neural Network(in Chinese with English abstract). *Earth Science*, (002):049.

附中文参考文献

- [27] 杜昕,范廷恩,董建华,等, 2020.基于多层感知机网络的薄储层预测.石油地球物理勘探, 55(06):1178-1187+1159.
- [28] 李章林,吴冲龙,张夏林,等, 2015.IDW 矿石品位估值结果的不确定性评价.地球科学(中国地质大学学报),40(11):1796-1801.
- [29] 民用机场勘测规范. MH /T5025—2011[S],2011.
- [30] 潘凯,谢春庆,程瑞驭,等, 2017.山区机场场地地下岩溶综合勘察工程实例.勘察科学技术, (01):28-32.

- [31] 仇志龙,吴嘉栋,万鹏,等, 2025.土石比对高填方碎石土地基强夯加固影响试验研究.地下空间与工程学报, 21(01):123-130.
- [32] 苏正洋,李典庆,王顺,等, 2025.三维土石混合体边坡大变形 SPH-DEM 模拟.地球科学,1-9.
- [33] 王巧,2011.四川红层机场挖方区土石比定量分析方法研究(硕士学位论文).成都理工大学.
- [34] 谢春庆,李正忠,潘凯,2015.西南某岩溶山区大面积挖方工程确定土石比方法研究.路基工程, (02):1-5.
- [35] 谢春庆,容树俭,王伟,2013.红层地区工程中土石比研究.路基工程, (06):83-88.
- [36] 张瀚,桂蕾,王腾飞,等,2024.基于 BP 神经网络的第四系覆盖物厚度预测及三维地质建模.地球科学, (002):049.