DOI：10.16285/j.rsm.2025.0099 CSTR:32223.14.j.rsm.2025.0099

融合先验物理知识和可解性分析的导热系数集成学习预测模型

王 琼[1,2]，田升奎[1]，刘观仕[3]，苏 薇[1]，刘宜春[1]，叶为民[1]

（1. 同济大学 岩土及地下工程教育部重点实验室，上海 200092；2. 同济大学 教育部城市环境与可持续发展联合研究中心，上海 200092；
3. 中国科学院武汉岩土力学研究所 岩土力学与工程安全全国重点实验室，湖北 武汉 430071）

摘 要：水热参数本构关系的精确获取是多场耦合研究的核心前提。然而，由于其多尺度的影响因素和高度非线性的响应模式，模型难以准确刻画因素间耦合效应、传热路径和传输机制，建立可靠的参数模型面临着诸多挑战。提出了一种融合先验物理知识与可解性分析的水热参数集成学习模型（physics-informed ensemble learning，简称 PIEL）。以土体导热系数λ为范例，综合评估 PIEL 模型输出结果的精度、鲁棒性和物理一致性，并结合沙普利加性解释（Shapley additive explanations，简称 SHAP）与部分依赖图（partial dependence plots，简称 PDPs）可视化 PIEL 模型决策过程的敏感性、响应模式和耦合效应。结果表明：集成学习模型精确捕捉了导热系数复杂的非线性耦合关系，预测精度较传统模型提升至 3～6 倍；融合先验传热知识的 PIEL 模型有效避免了纯数据驱动模型违背物理规律的输出，决策结果的物理一致性显著提升。采用麻雀搜索算法优化的融合知识的极端梯度提升模型（physics-informed extreme gradient boosting，简称 PXGBoost），展现出最佳的精度和鲁棒性；SHAP 和 PDPs 可视化的敏感性、响应模式和耦合效应与先验传热知识基本吻合，验证了 PIEL 决策过程的物理一致性。基于累计 SHAP 值识别的最优预测因子显著优于传统参数分析，构建的简化模型的计算效率和精度更高。融合先验物理知识和可解性分析的 PIEL 模型可为岩土参数预测提供一种有效途径，也为人工智能赋能的水热模拟提供支持。

关 键 词：导热系数；物理知识；可解释性分析；集成学习；物理一致性

中图分类号：TU 43 **文献标识码**：A **文章编号**：1000－7598 (2026) 02－674－17

Ensemble learning predictive model for thermal conductivity integrating priori physical knowledge and interpretive analysis

WANG Qiong[1,2], TIAN Sheng-kui[1], LIU Guan-shi[3], SU Wei[1], LIU Yi-chun[1], YE Wei-min[1]

(1. Key Laboratory of Geotechnical and Underground Engineering of Ministry of Education, Tongji University, Shanghai 200092, China;
2. United Research Center for Urban Environment and Sustainable Development, Ministry of Education, Tongji University, Shanghai 200092, China;
3. State Key Laboratory of Geomechanics and Geotechnical Engineering Safety, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan, Hubei 430071, China)

Abstract: Accurate determination of constitutive relationships for hydrothermal parameters is crucial for multi-field coupling studies. However, the presence of multi-scale influencing factors and highly nonlinear response behaviors makes it difficult for existing models to accurately capture coupling effects, heat transfer pathways, and transport mechanisms. Consequently, the development of reliable and robust parameter models remains a significant challenge. To address these issues, a physics-informed ensemble learning (PIEL) model was introduced, combining a priori physical knowledge with interpretive analysis. Using soil thermal conductivity (λ) as a case study, the PIEL model's accuracy, robustness, and physical consistency were comprehensively evaluated. Sensitivities, response patterns, and coupling effects in the decision-making process were visualized using Shapley additive explanations (SHAP) and partial dependence plots (PDPs). The results demonstrate that the proposed ensemble learning framework effectively captures the complex nonlinear coupling behavior governing thermal conductivity, improving prediction accuracy by a factor of 3 to 6 compared to traditional models. By incorporating a priori heat-transfer knowledge, the PIEL model effectively prevents physically implausible predictions—a common limitation of purely data-driven approaches—thereby substantially enhancing the physical consistency of model outputs. Among the evaluated methods, the physics-informed extreme gradient boosting (PXGBoost) model optimized via the

收稿日期：2025-01-26 录用日期：2025-07-02
基金项目：国家自然科学基金（No. 42172298, No. 42002289, No. 41907231, No. 42577180）；中央高校基本科研业务费专项资金（No. 22120250424）。
This work was supported by the National Natural Science Foundation of China (42172298, 42002289, 41907231, 42577180) and the Fundamental Research Funds for the Central Universities (22120250424).
第一作者简介：王琼，女，1982 年生，博士，教授，主要从事膨润土水热耦合的研究工作。E-mail: qiong.wang@tongji.edu.cn
通信作者：田升奎，男，1996 年生，博士研究生，主要从事膨润土水热耦合的研究工作。E-mail: kshengt@tongji.edu.cn

sparrow search algorithm exhibited the highest predictive accuracy and robustness. Sensitivity analyses and response patterns revealed by SHAP and PDPs are consistent with established heat-transfer theory, further validating the physical interpretability of the PIEL-based decision-making process. Furthermore, the optimal predictors identified through cumulative SHAP values significantly outperform traditional parameter analysis methods, enabling the development of more computationally efficient and accurate simplified models. The PIEL model, integrating physical knowledge, represents a powerful tool for geotechnical parameter prediction and paves the way for advancing AI-enabled hydrothermal simulations.

Keywords: thermal conductivity; physical knowledge; interpretive analysis; ensemble learning; physical consistency

1　引　言

为实现“双碳”目标和能源结构转型，能源岩土工程逐渐兴起，如地源热泵系统[1-2]、可燃冰开采[3]和高放核废物处置库建设[4-5]等，这些工程往往会遭受 100 ℃以上的高温[6-7]。以处置库为例，回填材料承担传热屏障，不合理的热参数设定，尤其是导热系数 λ，可能造成温度超出限制[8]，引发缓冲材料的微观裂隙、相变和宏观开裂，从而削弱处置库的屏障性能[9]。因此，精准确定土体导热系数，不仅有助于控制传热系统的温度分布，也对能源岩土工程的长期安全运营至关重要。

导热系数是多种因素协同作用的结果，大致分为 3 类：固有因素（如颗粒级配和矿物组成等）[1, 6, 10]、结构因素[11-13]（如含水率和干密度等）和环境因素[14-15]（如温度和盐分浓度等），且不同因素间存在高度的非线性耦合关系[16]，导致确定可靠的 λ 面临诸多挑战。当前，主流导热系数确定方法主要有测量法、理论模型、数值模拟和经验模型等[1, 17]。试验测量精度高，常作为其他方法的标准值，但繁琐费时的操作流程导致大尺寸现场调查的实践性差[18]。理论模型和数值模型大多以理想化的传热骨架为基础，其引入了部分物理意义不明确和难以获取的经验参数[9]，造成预测精度和代表性较差。经验模型通常基于有限的数据拟合，其预测精度依赖于数据库的适配度，泛化能力较差[19]。综上所述，传统模型均存在一定局限性，难以全面刻画因素间耦合效应、传热路径和传输机制，目前尚缺乏普适的 λ 预测模型。近年来，人工智能算法因其处理高维度非线性数据的能力，在岩土工程领域获得了广泛应用[3, 20-22]。李林等[20]研究了融入固结方程的神经网络框架，实现了原位固结系数的快速准确反演。张东明等[21]搭建了嵌入弹性地基模型的物理信息双驱动模型，实现了隧道力学性态智能诊断。Gong 等[22]开发了融合太沙基地基固结理论的物理信息神经网络（physics-informed neural network，简称 PINN），成功实现了分层土体固结系数的反演与长时间固结过程的预测，模型展现出良好的物理一致性与鲁棒性。由此可见，融合人工智能算法的 λ 预测模型具有巨大潜力。

当前，基于人工智能算法的导热系数预测模型已有初步探索，相关研究主要围绕纯数据驱动模型展开，重点研究数据库样本数量、算法类型、优化算法种类、输入变量维度对模型性能的影响[16-18]。如 Li 等[18]建立了 6 种机器学习的 λ 预测模型，研究表明机器学习模型的精度和鲁棒性显著优于传统模型。Kardani 等[17]对比了不同算法优化的深度神经网络 λ 预测模型的精度，结果表明优化算法的选择对其决策结果的可靠性至关重要。Zhao 等[16]对比了 7 种机器学习 λ 预测模型在不同类型土体中的适应性，结果表明模型的类型和土的种类均会影响预测精度。尽管上述人工智能模型相较于传统模型表现出更高的精度和鲁棒性，但现有模型大多未考虑影响因子与 λ 之间的响应模式和单调性，可能导致模型的预测过程违背非饱和土传热机制，影响决策结果的物理一致性（输出负值）[16-18]。建立融合先验传热知识的人工智能 λ 预测模型值得探索。此外，由于纯数据驱动算法的黑箱本质，模型的决策过程缺乏透明度，难以理解和信任模型产生的估计，这给其在实际工程中的大规模部署构成了重大挑战。Li[18]、Zhao[16]等运用相关性分析、重要性分析和灰度关联分析进行了影响因素的全局解释，但其结果存在悖论，更无法揭示特定实例中模型决策的原因。可解释性机器学习，如沙普利加性解释（Shapley additive explanations，简称 SHAP）与部分依赖图（partial dependence plots，简称 PDPs），作为一种新颖的模型决策可视化技术，已在部分领域获得应用[22-24]。然而，针对融合先验物理知识和可解性分析的 λ 集成学习预测模型尚未见诸文献报道。

本文提出了一种融合先验物理知识与可解性分析的水热参数本构关系集成学习框架（physics-informed ensemble learning，简称 PIEL）。以导热系数为范例，全面评估该框架的可靠性。首先通过广泛的文献检索汇编了一个具有代表性的 λ 数据库，基于此建立了融合非饱和土传热机制的 λ 集成学习模型。通过与传统模型的多维对比分析，评估模型决策结果的精度、鲁棒性和物理一致性。同时，引入 SHAP 和 PDPs 可解释性分析，可视化模型的决策过程，验证模型决策过程是否契合非饱和土传热

机制。最后，改进预测因子的识别方法，并构建简化模型。研究工作可为人工智能赋能的水热耦合研究提供潜在支持。

2 数据与方法

2.1 数据获取

本文共收集了 2 940 个数据，源自 28 篇文献[2, 4-8, 10-15, 25-40]中约 75 个地区的 118 种土[27]，涵盖了大部分类型的土样（见图 1）。与现有数据库不同的是，考虑到黏土复杂孔隙和水分形态对传热的影响，汇编了更多黏土。数据库收集了 λ 主要的影响因素，总计 11 个：固有因素包含黏粒含量 M_c、粉粒含量 M_{st}、砂粒含量 M_{sd}、石英含量 M_q、颗粒密度 ρ_s；结构因素有干密度 ρ_d、质量含水率 θ、孔隙率 n、体积含水率 w、饱和度 S_r；环境因素有温度 T。除 Nikolaev 等[33]的 λ 由稳态法获取，剩余数据均由瞬态法测定。鉴于两种方法测量差异较小，忽视测量方法对数据的影响。为便于定量分析，运用可靠的 WebPlot Digitizer 软件数字化测量结果。

M_q 未知时，部分学者假设 $M_q = M_{sd}$ [18]或 $M_q = 0.5\,M_{sd}$ [16]。除去数据库中石英砂样本，回归分析表明 $M_q = 0.32\,M_{sd}$，所以利用后者的假设确定王海波等[29]数据中的 M_q。为弥补现有数据库中温度区间窄和缺失的不足，收集了更广泛的温度数据。

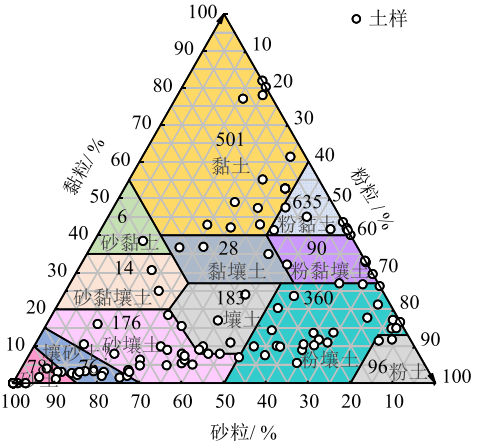


图 1 土样的粒径分布（数值代表土样的数量）

Fig.1 Particle size distribution of soil samples (the values represent the number of soil samples)

2.2 传统模型

将 4 个经典的传统模型视为基准，用来验证人工智能模型的性能。为了便于描述，4 个传统模型分别简称为 JO1975[41]、CK2005[42]、DV1963[16]、TF2009[9]模型。JO1975 模型是由 Johansen[41]首次提出的归一化经验模型（见式（1）～（5）），其计算结果由系数 K_e、干土导热系数 λ_{dry}、饱和土导热系数 λ_{sat} 决定。

$$K_e = \frac{\lambda - \lambda_{dry}}{\lambda_{sat} - \lambda_{dry}} \tag{1}$$

$$\lambda_{sat} = (\lambda_w)^n (\lambda_s)^{1-n} \tag{2}$$

$$\lambda_{dry} = (\lambda_a)^n (\lambda_s)^{1-n} \tag{3}$$

$$\lambda_s = (\lambda_q)^{\phi_q} (\lambda_o)^{1-\phi_q},\ \lambda_o = \begin{cases} 2, & \phi_q > 0.2 \\ 3, & \phi_q \leqslant 0.2 \end{cases} \tag{4}$$

$$K_e = \begin{cases} \eta_1 \lg S_r + \eta_2, & S_r > 10\%, \text{非冻土} \\ 0, & 0 \leqslant S_r \leqslant 10\%, \text{非冻土} \\ S_r, & \text{冻土} \end{cases} \tag{5}$$

式中：K_e 为一个 Kersten 函数；λ_s、λ_w 和 λ_a 分别为固体颗粒、液相和气相的导热系数；λ_q 为石英的导热系数；λ_o 为其他矿物的导热系数；ϕ_q 为石英的体积。

为了提升 JO1975 模型的精度和实用性，Côté 等[42]改进了 λ_{dry} 和 K_e，得到 CK2005 模型（见式（6）和式（7）），有效避免了 JO1975 模型中 K_e 的发散问题。

$$K_e = \frac{\eta_3 S_r}{1 + (\eta_3 - 1) S_r} \tag{6}$$

$$\lambda_{dry} = \eta_3 10^{-\eta_4 n} \tag{7}$$

式中：η_1、η_2、η_3 和 η_4 均为经验系数，与土的类型有关。考虑温度效应的液相和气相导热系数分别为

$$\left.\begin{aligned} \lambda_a &= 2.454\times10^{-2} + 7.27\times10^{-5} T \\ \lambda_w &= 0.569 + 1.884\times10^{-3} T - 0.772\times10^{-6} T^2 \end{aligned}\right\} \tag{8}$$

为克服经验模型的不足，一系列理论模型应运而生，DV1963 模型[16]是从 Maxwell 方程演化而来的。假设土体由自由漂浮在连续孔隙流体（空气或水）中的椭球体颗粒组成，考虑形状因子后，采用加权平均法计算导热系数：

$$\lambda = \frac{F_w \phi_w \lambda_w + F_a \phi_a \lambda_a + F_s \phi_s \lambda_s}{F_w \phi_w + F_a \phi_a + F_s \phi_s} \tag{9}$$

式中：ϕ_s、ϕ_w 和 ϕ_a 分别为固相、液相和气相的体积；F_s、F_w 和 F_a 分别为固相、液相和气相的形状因子，由下式计算：

$$F_i = \frac{2}{3\left[1 + (\lambda_i/\lambda_w - 1) g_i\right]} + \frac{1}{3\left[1 + (\lambda_i/\lambda_w - 1)(1 - 2g_i)\right]} \tag{10}$$

式中：$i = \mathrm{s,w,a}$；g_i 为椭球体的拟合参数，计算式为

$$g_i=\begin{cases}0.013+0.944\phi_{\mathrm{w}}, & i=\mathrm{w}\\ 0.333-(0.333-0.035)(1-S_{\mathrm{r}}), & i=\mathrm{s}\\ 0.125, & i=\mathrm{a}\end{cases} \tag{11}$$

TF2009 模型[9]以串并联理论的 Wiener 边界为基础，考虑饱和度变化，类比获得了三相混合物的导热系数上限 λ^{U} 和下限 λ^{L} 分别为

$$\lambda^{\mathrm{U}}=\eta_5(1-n)\lambda_{\mathrm{s}}+(1-n)(1-\eta_5)\lambda_{\mathrm{s}}+nS_{\mathrm{r}}\lambda_{\mathrm{w}}+n(1-S_{\mathrm{r}})\lambda_{\mathrm{a}} \tag{12}$$

$$\lambda^{\mathrm{L}}=\eta_5(1-n)\lambda_{\mathrm{s}}+\left[1-\eta_5(1-n)\right]^2\times\left[\frac{(1-n)(1-\eta_5)}{\lambda_{\mathrm{s}}}+\frac{nS_{\mathrm{r}}}{\lambda_{\mathrm{w}}}+\frac{n(1-S_{\mathrm{r}})^{-1}}{\lambda_{\mathrm{a}}}\right] \tag{13}$$

Tong 等[9]认为非饱和土的表观导热系数为串联和并联边界的加权组合：

$$\lambda=(1-\eta_6)\lambda^{\mathrm{L}}+\eta_6\lambda^{\mathrm{U}} \tag{14}$$

$$\lambda=\eta_5(1-n)\lambda_{\mathrm{s}}+(1-\eta_6)[1-\eta_5(1-n)]^2\times\left[\frac{(1-n)(1-\eta_5)}{\lambda_{\mathrm{s}}}+\frac{nS_{\mathrm{r}}}{\lambda_{\mathrm{w}}}+\frac{n(1-S_{\mathrm{r}})}{\lambda_{\mathrm{a}}}\right]^{-1}+\eta_6[(1-n)(1-\eta_5)\lambda_{\mathrm{s}}+nS_{\mathrm{r}}\lambda_{\mathrm{w}}]+\eta_6 n(1-S_{\mathrm{r}})\lambda_{\mathrm{a}} \tag{15}$$

$$\left.\begin{aligned}\eta_5&=0.069\ 2\,n^{-0.783\ 1}\\ \eta_6&=0.59S_{\mathrm{r}}^{1.487}n-0.040\ 4\end{aligned}\right\} \tag{16}$$

式中：η_5、η_6 均为与土孔隙有关的经验系数。

2.3　融合先验传热知识的 PIEL 模型框架

2.3.1 模型建立

融合先验物理知识与机器学习是当前岩土工程研究的热点之一，尤其是 PINN 模型，已在固结[20, 22]、渗流和变形[21]等领域获得了显著进展。但也存在部分缺点：①PINN 模型多以深度神经网络框架为基础，需要大量数据避免过拟合；②神经网络框架和超参数优化存在挑战，试错法（trial-and-error method）难以获得最优解；③神经网络常被视为“黑箱”，难以对模型决策路径做出可视化的解释。为此，选择 4 种可解释性更强的集成树学习模型，包括梯度提升决策树（gradient boosting decision tree，简称 GBDT）、极限梯度提升树（extreme gradient boosting，简称 XGBoost）、类别特征梯度提升树（categorical boosting，简称 CatBoost）和轻量级梯度提升机（light gradient boosting machine，简称 LightGBM），作为基准。为了区分纯数据驱动模型和融合先验物理信息模型，分别给模型简称增加前缀 D 和 P。尽管集成学习模型已在岩土工程高维度非线性问题中展现出优势，但现有模型多为纯数据驱动，融合先验物理信息的工作鲜有报道。纯数据驱动模型仅是数据的逼近，可能导致违背先验物理知识的决策行为，降低模型在风险预测和工程应用中的可靠性和稳健性[22]。为此，本文为了使集成学习模型的决策过程和结果均符合先验非饱和土传热机制，分别对基准集成学习模型进行两方面改进：①基于先验传热知识限定模型输入和输出的单调性，采用模型自带单调约束功能实现，确保模型输入和输出间的响应模式与先验传热机制吻合；②将先验传热知识融合到数据驱动模型损失函数中。

以纯数据驱动极限梯度提升树（data-driven extreme gradient boosting，简称 DXGBoost）模型为例，详细介绍融合先验物理信息模型的建立过程。D_{ij} 是样本数据库：

$$D_{ij}=\{(\boldsymbol{X}_1,Y_1);(\boldsymbol{X}_2,Y_2);\cdots;(\boldsymbol{X}_n,Y_n)\} \tag{17}$$

式中：$\boldsymbol{X}_i$ 和 Y_i 分别为模型输入（影响因素）和输出（导热系数）；i 为样本编号，$i=1, 2, \cdots, n$；j 为影响因素的个数，j=1, 2, …, 11。$\boldsymbol{X}_i$ 为 11 个影响因素组成的行向量：

$$\boldsymbol{X}_i=(x_1,x_2,\cdots x_j) \tag{18}$$

为消除数据量纲和范围差异，对汇编数据进行 Min-Max 归一化处理：

$$D_{ij}^{\mathrm{nor}}=\frac{D_{ij}-\mathrm{Min}_i(D_{ij})}{\mathrm{Max}_i(D_{ij})-\mathrm{Min}_i(D_{ij})} \tag{19}$$

通过训练模型和调整参数，寻找输入和输出之间最优的映射关系：

$$\hat{y}_i=\mathcal{F}(\boldsymbol{X}_i)=\sum_{k=1}^{K}a_k\mathcal{F}_k(\boldsymbol{X}_i) \tag{20}$$

式中：$\hat{y}_i$ 为模型的预测值；K 为回归树的个数；$\mathcal{F}$ 为最终的强学习器；$\mathcal{F}_k$ 为弱学习器；a_k 为学习速率。DXGBoost 模型的目标函数 $\mathcal{L}_{\mathrm{O}}(\cdot)$ 由两部分组成：

$$\mathcal{L}_{\mathrm{O}}(\mathcal{F})=\sum_{i=1}^{n}\mathcal{L}_{\mathrm{D}}(y_i,\hat{y}_i)+\sum_{k=1}^{K}\Omega(\mathcal{F}_k) \tag{21}$$

式中：$\mathcal{L}_{\mathrm{D}}(\cdot)$ 和 $\Omega(\cdot)$ 分别为标签值与预测值之间的损失函数和模型的复杂度（正则化项），分别限定模型的精度和复杂度，计算公式如下：

$$\left.\begin{aligned}\mathcal{L}_{\mathrm{D}}(y_i,\hat{y}_i)&=(y_i-\hat{y}_i)^2\\ \Omega(\mathcal{F}_k)&=\gamma T+\frac{1}{2}\chi\|W\|^2\end{aligned}\right. \tag{22}$$

式中：Γ 和 W 分别为叶子节点的总数和权重；γ 和 χ 均为惩罚因子；$\|\ \|$ 为二范数。

为了在 DXGBoost 模型中融合先验的传热理论，将 PXGBoost 模型的总损失 $\mathcal{L}_{\mathrm{T}}$ 概化成数据损失 $\mathcal{L}_{\mathrm{O}}$ 和物理损失 $\mathcal{L}_{\mathrm{P}}$ 之和：

$$\mathcal{L}_{\mathrm{T}} = w_{\mathrm{O}}\mathcal{L}_{\mathrm{O}} + w_{\mathrm{P}}(\mathcal{L}_{\mathrm{P}}^{1} + \mathcal{L}_{\mathrm{P}}^{2}) \tag{23}$$

式中：w_{O} 和 w_{P} 分别为数据损失和物理损失的权重；$\mathcal{L}_{\mathrm{P}}^{1}$ 和 $\mathcal{L}_{\mathrm{P}}^{2}$ 分别为 $\hat{\lambda}$ 与下限 $\lambda_{\mathrm{H\text{-}S}}^{\mathrm{L}}$ 和上限 $\lambda_{\mathrm{H\text{-}S}}^{\mathrm{U}}$ 的物理损失。通常设定 $w_{\mathrm{O}}=1$，通过试错法或自适应权重法确定最优权重组合。理论上，物理损失一般可以由微分方程、初值条件或边界条件等组成[20-21]。本文选择 Hashin-Shtrikman（H-S）边界约束作为物理损失。原因在于，Wiener 边界和 H-S 边界被广泛用于评价 λ 预测值的合理性，且后者的约束区间更严格[9, 19]。采用 H-S 边界引导模型的收敛方向，确保模型输出具有物理意义。将 $\mathcal{L}_{\mathrm{P}}^{1}$、$\mathcal{L}_{\mathrm{P}}^{2}$ 定义为预测值 $\hat{\lambda}$ 到 H-S 下边界和上边界距离的平方：

$$\mathcal{L}_{\mathrm{P}}^{1} = \left\| \hat{\lambda} - (\lambda_{\mathrm{H\text{-}S}}^{\mathrm{L}})_i \right\|^2 \tag{24}$$

$$\mathcal{L}_{\mathrm{P}}^{2} = \left\| \hat{\lambda} - (\lambda_{\mathrm{H\text{-}S}}^{\mathrm{U}})_i \right\|^2 \tag{25}$$

$$\left.\begin{aligned}
\lambda_{\mathrm{H\text{-}S}}^{\mathrm{L}} &= \lambda_{\mathrm{a}} + \frac{3\lambda_{\mathrm{a}}\left[\phi_{\mathrm{w}}/\left(1+f_{\mathrm{w\text{-}a}}\right)+\phi_{\mathrm{s}}/\left(1+f_{\mathrm{s\text{-}a}}\right)\right]}{\phi_{\mathrm{a}}+\phi_{\mathrm{w}}f_{\mathrm{w\text{-}a}}/\left(1+f_{\mathrm{w\text{-}a}}\right)+\phi_{\mathrm{s}}f_{\mathrm{s\text{-}a}}/\left(1+f_{\mathrm{s\text{-}a}}\right)} \\
\lambda_{\mathrm{H\text{-}S}}^{\mathrm{U}} &= \lambda_{\mathrm{s}} + \frac{3\lambda_{\mathrm{s}}\left[\phi_{\mathrm{w}}/\left(1+f_{\mathrm{w\text{-}s}}\right)+\phi_{\mathrm{a}}/\left(1+f_{\mathrm{a\text{-}s}}\right)\right]}{\phi_{\mathrm{s}}+\phi_{\mathrm{w}}f_{\mathrm{w\text{-}s}}/\left(1+f_{\mathrm{w\text{-}s}}\right)+\phi_{\mathrm{a}}f_{\mathrm{a\text{-}s}}/\left(1+f_{\mathrm{a\text{-}s}}\right)}
\end{aligned}\right\} \tag{26}$$

式中：$f_{\mathrm{a\text{-}s}} = 3\lambda_{\mathrm{s}}/(\lambda_{\mathrm{a}}-\lambda_{\mathrm{s}})$；$f_{\mathrm{w\text{-}s}} = 3\lambda_{\mathrm{s}}/(\lambda_{\mathrm{w}}-\lambda_{\mathrm{s}})$；$f_{\mathrm{w\text{-}a}} = 3\lambda_{\mathrm{a}}/(\lambda_{\mathrm{w}}-\lambda_{\mathrm{a}})$；$f_{\mathrm{s\text{-}a}} = 3\lambda_{\mathrm{a}}/(\lambda_{\mathrm{s}}-\lambda_{\mathrm{a}})$。最终，寻找最优学习器 $\mathcal{F}_t$ 使 PXGBoost 模型的 $\mathcal{L}_{\mathrm{T}}$ 最小。

$$\begin{aligned}
\mathcal{F}_t &= \arg\min_{\mathcal{F}_t\in\mathcal{F}} \mathcal{L}(\mathcal{F}_t) = \\
&\arg\min_{\mathcal{F}_t\in\mathcal{F}} \left[w_{\mathrm{O}}\sum_{i=1}^{n}\mathcal{L}_{\mathrm{D}} + w_{\mathrm{P}}\sum_{i=1}^{n}\mathcal{L}_{\mathrm{P}} + w_{\mathrm{O}}\sum_{k=1}^{t}\Omega(\mathcal{F}_k) \right]
\end{aligned} \tag{27}$$

式中：arg min 表示寻找总损失最小值时的模型参数；t 为当前迭代轮数。

按照 7∶3 的比例随机划分训练集（2 058 个样本）和测试集（882 个样本），采用 5 折交叉验证（five-fold cross-validation）训练模型，具体而言，将训练集随机分成 5 份，取其中 4 份训练，剩余 1 份用于验证模型，重复此过程 5 次，取其均值作为模型输出，以此提升模型的泛化能力和可信度[43]。最后，对模型输出做反归一化，再进行性能评价。

2.3.2 模型优化

集成树模型性能依赖超参数设定，传统随机搜索和网格搜索算法缺乏理论支持，且难以获得最优超参数组合。为此，部分学者提出了贝叶斯优化和元启发式算法[44]。前者利用有限的超参数组合与概率代理模型，识别验证集上性能曲线上的最小值作为最优超参数组合。后者通过模拟自然界生物种群或智能群体的社会行为，在众多高维、非线性优化问题中优势显著[44]。Zhang[44]、Gharehchopogh[45]等对比了多种元启发式算法，结果表明麻雀搜索算法（sparrow search algorithm，简称 SSA）具有搜索能力强、收敛速度快和稳定性高的特点，为此，本文采用此算法识别最优的超参数组合。麻雀搜索算法[44]模仿麻雀种群的觅食和反捕食者行为来解决复杂的优化问题。根据麻雀能量的高低，种群被分为 3 个角色：生产者、搜寻者和警告者，代表优化模型的超参数空间。为简洁起见，仅给出了生产者适应度的更新算法，细节见文献[44-45]。

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t}\exp\left(-\dfrac{i}{\alpha t_{\max}}\right), & R_2 < T_{\mathrm{S}} \\ X_{i,j}^{t} + Q\boldsymbol{L}, & R_2 \geqslant T_{\mathrm{S}} \end{cases} \tag{28}$$

式中：$X_{i,j}^{t+1}$ 为 t+1 次迭代中第 i 个个体第 j 维的适应度；α 和 Q 为随机数；$t_{\max}$ 为最大迭代次数；$\boldsymbol{L}$ 为 1×d 的矩阵，d 为优化变量的数量；R_2 和 T_{S} 分别代表报警和安全阈值。当 $R_2 < T_{\mathrm{S}}$ 时，种群处于安全环境，允许生产者进入广泛的搜索模式；反之，当 $R_2 \geqslant T_{\mathrm{S}}$ 时，发现捕食者存在，所有麻雀迅速迁移到安全区。

2.3.3 模型评价

为了验证模型决策结果的物理一致性，分别采用指标法[16]、排名分析[18]和精度热图[17]对模型的性能进行多维度评估。选择 10 个指标，部分指标见式（29）～（32），式（29）、（30）为误差指标，理想值均为 0，式（31）、（32）是趋势指标，理想值分别为 1 和 2。模型的指标越接近理想值，其性能越优。

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{m}(\lambda_{\mathrm{e}}-\lambda_{\mathrm{p}})^2 \Big/ m} \tag{29}$$

$$\mathrm{WMAPE} = \sum_{i=1}^{m}[(\lambda_{\mathrm{e}}-\lambda_{\mathrm{p}})/\lambda_{\mathrm{e}}]\times\lambda_{\mathrm{e}} \Big/ \sum_{i=1}^{m}\lambda_{\mathrm{e}} \tag{30}$$

$$R^2 = 1 - \sum_{i=1}^{m}(\lambda_{\mathrm{e}}-\lambda_{\mathrm{p}})^2 \Big/ \sum_{i=1}^{m}(\lambda_{\mathrm{e}}-\lambda_{\mathrm{em}})^2 \tag{31}$$

$$\mathrm{PI} = R_{\mathrm{Adj}}^2 + 0.01\mathrm{VAF} - \mathrm{RMSE} \tag{32}$$

式中：RMSE 为均方根误差；WMAPE 为加权平均绝对百分比误差；R^2 为决策系数；PI 为性能指标；R^2_{Adj} 为调整后的相关系数；VAF 为方差解释率；λ_{p} 为导热系数的预测值；m 为样本数量；λ_{e} 和 λ_{em} 分别为导热系数的测量值和测量值均值。

2.3.4 模型解释

鉴于现有可解释性技术在高维变量协同效应分析中的局限性，选择适用性广泛的 SHAP 与 PDPs 方法，仅对模型决策过程中主控因素的响应模式与耦合效应进行系统解析。

（1）沙普利加性解释（SHAP）

为了验证模型的决策过程是否与先验的非饱和土传热机制吻合，采用基于博弈论的 Tree-SHAP 技术增强对模型的理解和解释[23-24]。每个特征的贡献 φ_i 采用建模中使用的所有可行特征组合中的平均边际贡献衡量，可以实现全局和局部解释。SHAP 值计算如下：

$$\varphi_i=\sum_{S\in F\backslash\{i\}}\frac{|S|!(M-|S|-1)!}{M!}\left[f_x(S\cup\{i\})-f_x(S)\right]$$
$$f_x(S)=f(h_x^{-1}(z))=E[f(x)\,|\,x_{\mathrm{S}}] \tag{33}$$

式中：F 为所有特征的集合；M 为输入特征的个数；S 为不包含第 i 个特征的任意子集；$f_x(S\cup\{i\})$ 和 $f_x(S)$ 分别为模型包含和不包含第 i 个特征的输出；E 为期望。SHAP 通过函数 $h(x)$将 x 映射到简化的输入 z'，原始模型 $f(x)$ 便可以通过线性模型叠加来逼近：

$$\left.\begin{aligned}f(x)&=g(z')=\varphi_0+\sum_{i=1}^{M}\varphi_i z_i'\\ z'&=h(x)\end{aligned}\right\} \tag{34}$$

式中：$g(z')$ 为可解释模型；φ_0 为基值（$E(f(x))$）；简化输入 $z_i'=\{0,1\}^M$，当选取第 i 个特征时取 1，否则取 0。

（2）部分依赖图解释（PDPs）

PDPs 用于定量分析影响因素与模型输出之间的敏感性、响应模式和耦合效应。单因素 PDPs 具体地揭示目标和特定输入特征之间的响应模式，而双因素 PDPs 可视化两个特征之间协同效应对模型输出的影响。PDPs 计算如下[24]：

$$\hat{f}_{x_{\mathrm{s}}}(X_{\mathrm{s}})=E_{x_{\mathrm{c}}}[\hat{f}(x_{\mathrm{s}},x_{\mathrm{c}})]=\int\hat{f}(x_{\mathrm{s}},x_{\mathrm{c}})\mathrm{d}P(x_{\mathrm{c}}) \tag{35}$$

式中：$\hat{f}$ 为解释模型；X_{s} 为构建部分依赖关系的特征集合；$E_{x_{\mathrm{c}}}$ 为剩余特征的期望；x_{s} 为解释的特征；x_{c} 为训练模型的变量中除 x_{s} 的剩余特征；$P(x_{\mathrm{c}})$ 为 x_{c} 边缘概率分布函数。

图 2 系统总结了本文提出的融合先验物理信息和可解性分析的 λ 集成学习预测模型。提出的方法不局限于 λ，也适用不同岩土参数（如土-水特征曲线和非饱和渗透系数）；先验物理知识、集成学习模型和优化算法可依据实际问题进行调整。

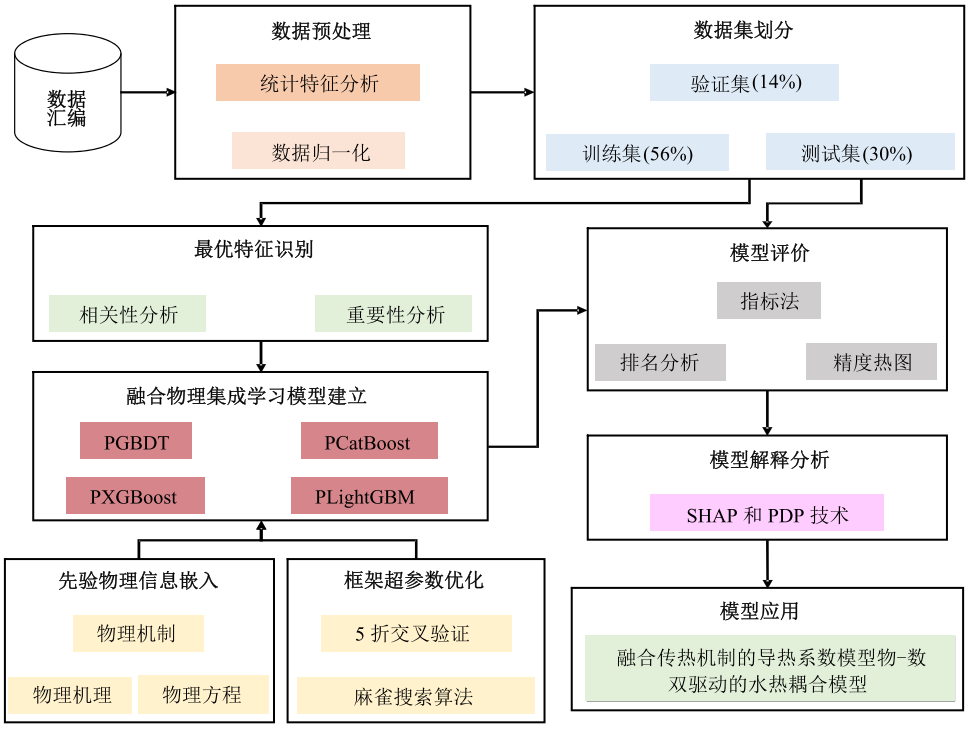


图 2　融合先验物理信息和可解性分析的λ预测 PIEL 模型流程图

Fig.2　Flowchart of PIEL models for predicting λ integrating a priori physical knowledge and interpretive analysis

3 结果与讨论

3.1 统计特征

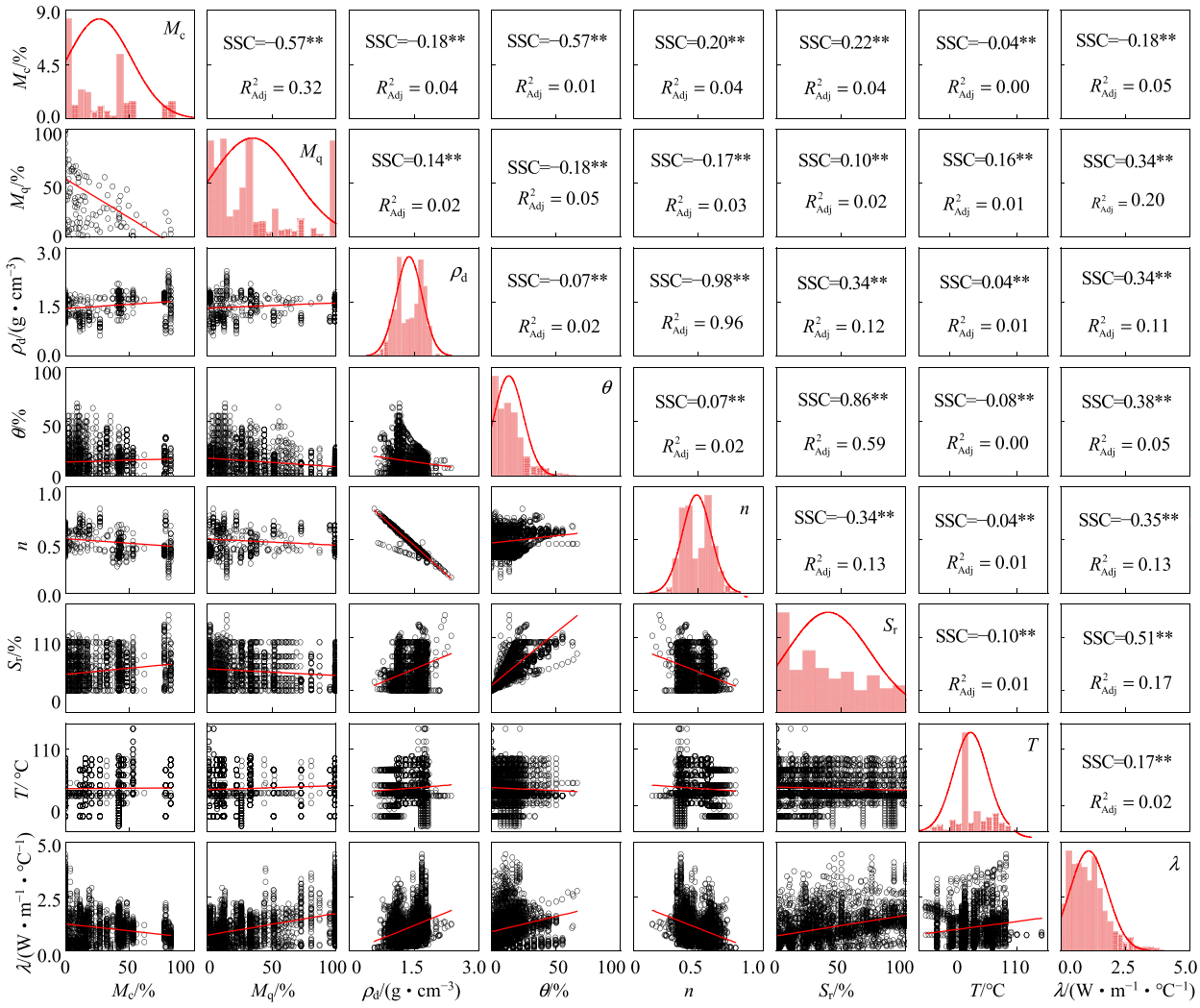
汇编 λ 和影响因子的基本统计特征见表 1。T 和 λ 的范围分别介于−38.50～150.10 ℃和 0.08～4.45 W/(m·℃)，变异系数分别为80.07%和67.41%，表明汇编的 λ 数据库有一定的代表性，可用于 PIEL 模型的训练和可解释性分析。偏度和峰度描述数据的分布形态，分别刻画数据的对称性和陡缓程度。各因素的峰度和偏度均不为 0，表明数据不满足正态分布，为此选择斯皮尔曼相关系数（Spearman's correlation coefficient，简称 SCC）定量分析关键因素与 λ 之间的相关性，结果如图 3 所示。M_q、θ、ρ_d、S_r、T 与 λ 之间存在单调递增关系，而 M_c、n 与 λ 之间存在单调递减关系。但石英作为土中导热系数最高的矿物，其主导地位在相关性分析中被低估，违背了先验传热规律。此外，数据分布较离散，表明 λ 与各因素之间存在高度非线性耦合关系。

表 1 汇编数据库的统计特征

Table 1 Statistical characteristics of the compilation database

因素	最小值	最大值	均值	标准差	变异系数/%	偏度	峰度
M_c	0.00	81.90	26.35	24.92	94.60	0.69	−0.49
M_{st}	0.00	83.85	32.63	24.84	76.11	0.12	−1.09
M_{sd}	0.00	100.00	41.02	39.92	97.31	0.51	−1.44
M_q	0.00	100.00	35.03	32.37	92.42	0.96	−0.38
ρ_s	2.54	2.89	2.70	0.05	1.72	−0.30	0.67
ρ_d	0.58	2.36	1.38	0.30	21.98	−0.12	−1.04
θ	0.00	0.93	0.14	0.11	84.06	1.25	2.37
n	0.15	0.79	0.49	0.11	22.82	0.12	−1.09
w	0.00	0.65	0.18	0.14	77.26	0.61	−0.21
S_r	0.00	155.97	40.06	32.17	80.30	0.51	−0.79
T	−38.50	150.10	34.34	27.49	80.07	0.40	0.47
λ[1]	0.08	4.45	1.07	0.72	67.41	1.11	1.55

注：M_c、M_{st}、M_{sd}、M_q 的单位为%；ρ_s、ρ_d 的单位为 g/cm³；S_r 的单位为%；T 的单位为℃；λ的单位为 W/(m·℃)。



注：*和**分别表示相关性达到显著性水平和极显著性水平。

图 3 汇编数据库的关键变量的相关系数矩阵

Fig.3 Correlation matrix of key variables of the compiled database

3.2　传统模型的精度评价

图 4 给出了 JO1975、CK2005、DV1963 和 TF2009 4 种传统模型的计算结果。模型整体精度较差，计算值偏小，大量误差超出 50%，尤其在 $\lambda>2$ W/(m·℃)时，这表明传统模型在不同类型土中的鲁棒性和泛化能力较差。CK2005 模型性能最优，其次是 JO1975 和 DV1963 模型，TF2009 模型最差。CK2005 模型是在 JO1975 模型基础上更深入考虑土体类型对 λ 的影响，故精度提升了 9.3%。DV1963 和 TF2009 模型进行平均化和理想化的简化处理过程中，一方面难以真实刻画土的孔隙结构和传热骨架，如颗粒与颗粒间有效接触的点数量和面积；另一方面，忽视了多种因素耦合控制的传热机制，如温度效应引起的对流传热和水汽潜热传输。针对上述不足，部分研究[2, 19, 36, 40]引入经验参数来完善模型框架，但是这些经验参数往往难以获取，实际适用性受限。由于土体矿物组成和近场环境的复杂性，导致传统模型准确地表征因素间耦合效应、传热路径和传输机制面临诸多挑战，当前仍缺乏一种预测 λ 的普适模型，基于人工智能算法的 λ 模型值得探索。

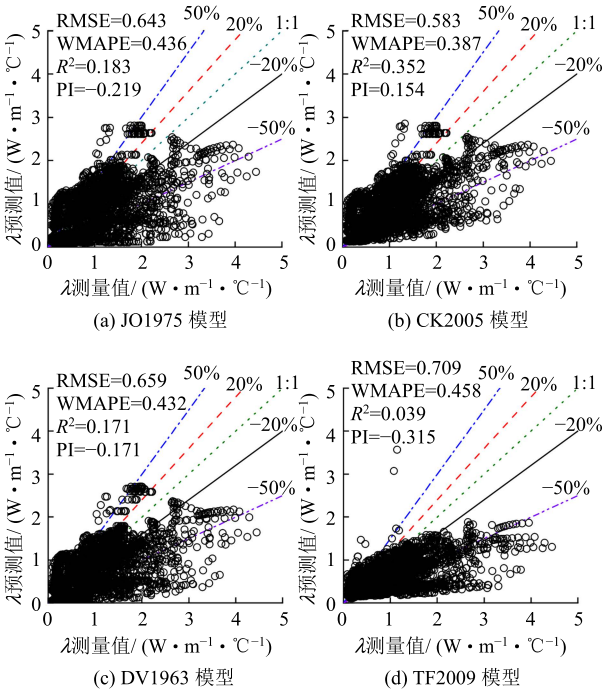


图 4　4 个传统模型的性能

Fig.4　Performance of four traditional models

3.3　纯数据驱动模型的精度评价

训练两组纯数据驱动的人工智能模型，第 1 组为支持向量机（support vector machine，简称 SVM）、深度神经网络（deep neural Network，简称 DNN）、高斯过程回归（Gaussian process regression，简称 GPR）和决策树（decision tree，简称 DT）组成的经典机器学习模型；另一组为 GBDT、XGBoost、CatBoost 和 LightGBM 组成的集成学习模型。每组排名前二的预测结果如图 5 所示。整体上，人工智能模型的性能普遍优于传统模型，集成学习模型优于经典机器学习模型。以 RMSE 指标为例，人工智能模型提升至最优传统 CK2005 模型的 3～5 倍，最优机器学习 DGPR 模型（RMSE=0.180，$R^2=0.938$）和最优集成学习 DXGBoost 模型（RMSE=0.136，$R^2=0.966$）展现了最高的精度和鲁棒性，且后者性能提升到前者的 24.4%，上述结果在可视化精度热图（见图 6）和排名分析（限于篇幅未提供）中更为突出，凸显了人工智能算法在捕捉 λ 非线性耦合关系上的显著优势。但两组纯数据驱动模型均给出少量违背先验物理知识的输出（负值），这一物理不一致性在传统模型中并未出现。尽管负值的偏离较小，但充分说明了数据驱动模型在应对噪声数据和极端情况时可能缺乏物理一致性，这将极大削弱人工智能模型在风险评估应用中的可靠性。

3.4　融合先验传热知识的 PIEL 模型评价

3.4.1　模型参数识别

图 7 给出了不同物理约束权重 w_p 下 XGBoost 模型的预测性能，w_p 对模型性能的影响显著，尤其是 $w_p>0.05$ 时。鉴于 w_p 对模型训练效率的影响较小，

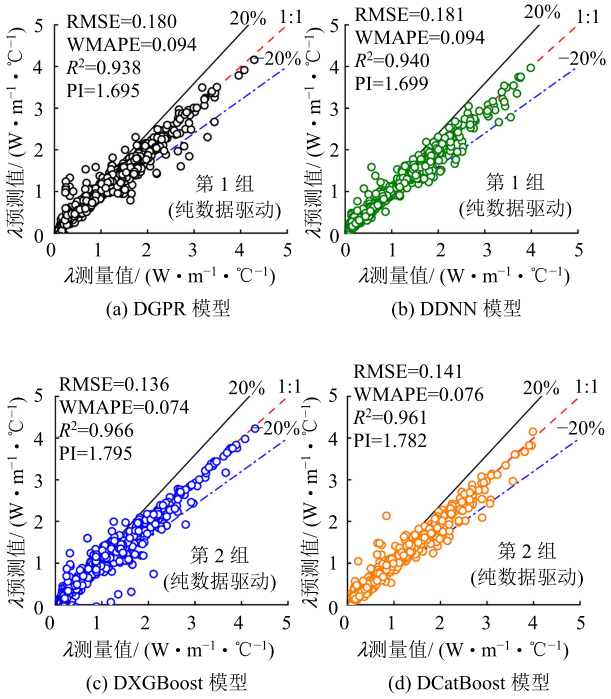


图 5　纯数据驱动机器学习模型预测性能

Fig.5　Predictive performance of pure data-driven models

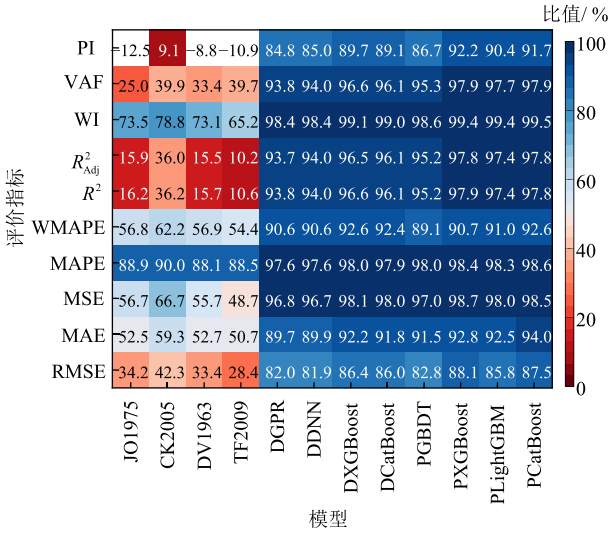


图 6 不同模型的精度热图

Fig.6 Heat map of accuracy for different models

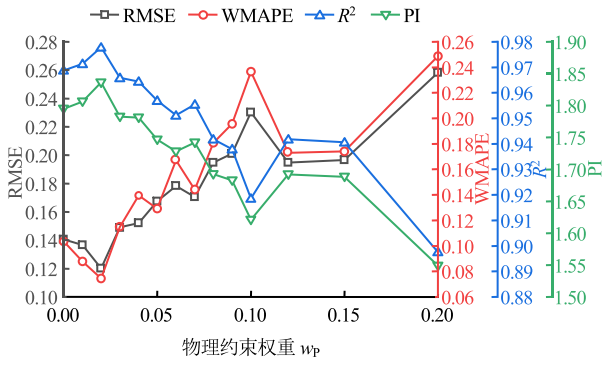


图 7 不同 w_P 设定 XGBoost 模型的性能

Fig.7 Performance of XGBoost model with different w_P settings

以最小 RMSE 为标准，识别最优 $w_p = 0.02$；PGBDT/PCatBoost 和 PLightGBM 模型的最优 w_p 分别为 0.01 和 0.02。参考 Zhang 等[44]研究，设定麻雀搜索算法的迭代次数为 60，种群规模为 30。搜索的 PXGBoost 模型最优超参数组合结果见表 2。

表 2 PXGBoost 模型超参数搜寻范围和最优值

Table 2 Search space and hyperparameter optimal values of PXGBoost model

超参数	取值范围	最优值
K	10～200	181
α	0.01～0.30	0.21
γ	0.0～0.5	0.1
最大深度	1～10	6
行采样比例	0.5～1.0	0.8
列采样比例	0.5～1.0	1.0
最小叶子节点样本数	1～10	6

为进一步验证麻雀搜索算法（SSA）的收敛性与稳定性，对比其与遗传算法（genetic algorithm，简称 GA）在相同超参数空间下的优化效果。SSA 在第 25 次迭代后趋于收敛，表现出更快的收敛速度。经 60 次迭代后，SSA 和 GA 所优化模型的均方误差 MSE 分别为 0.001 11 和 0.001 85，前者在预测精度上也更优。鉴于篇幅限制，部分最优适应度随迭代次数的变化关系见图 2。

3.4.2 模型性能评价

交叉验证结果表明，模型在各折中的表现稳定。以表现最差的 PGBDT 模型和表现最佳的 PXGBoost 模型为例，5 折验证中，前者的均方误差 MSE 和决定系数 R^2 分别介于 0.001 4～0.002 3 和 0.918～0.950 之间；后者的 MSE 和 R^2 分别限定在 0.001 1～0.001 4 和 0.947～0.968 之间。由此可见，模型整体性能波动较小，且具有良好的泛化能力。图 8 给出了融合先验传热知识的 PIEL 模型测试集的回归图。相较于传统模型，PIEL 模型精度提升显著，以 RMSE 指标为例，PXGBoost 模型的精度提升到了最优传统 CK2005 模型的 5 倍；相较于 DXGBoost 模型，PIEL 模型提升幅度较小，提升增幅最大的是 PLightGBM 模型（13.8%）和 PXGBoost 模型（12.5%）。PIEL 模型预测的 λ 分布范围更集中，且

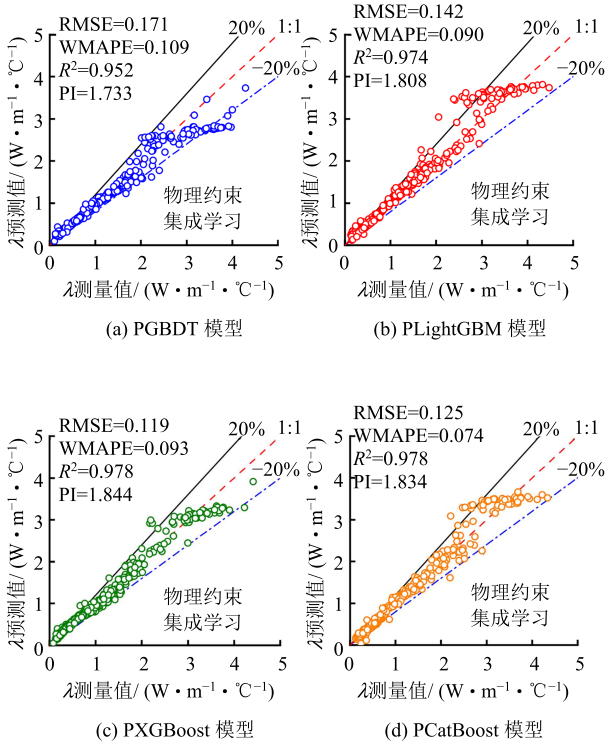


图 8 融合先验传热知识 PIEL 模型性能对比

Fig.8 Performance comparison of PIEL models incorporating a priori heat transfer knowledge

均为正值，几乎全部落在 20%以内，尤其是在 $\lambda<3$ W/(m·℃)时，而当 $\lambda>3$ W/(m·℃)时，误差略有增大。图 9 对比了不同模型的预测误差分布，PXGBoost 和 PCatBoost 模型的误差范围最小，分别介于−0.91～0.80 和−0.81～0.88，误差聚类在 0 附近，离群值较少，中间值与平均值相近，误差呈正态分布，表明融合先验传热知识的 PIEL 模型整体性能更稳定。

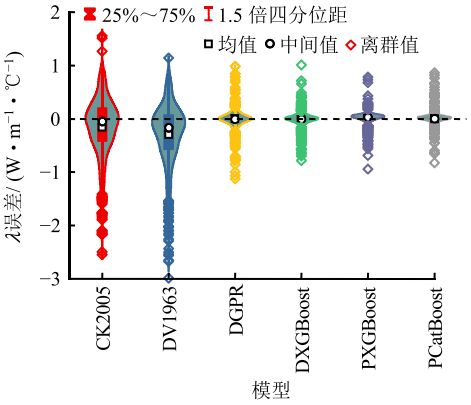
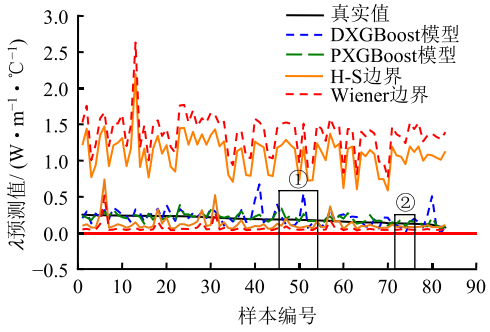


图 9　不同模型的误差分布

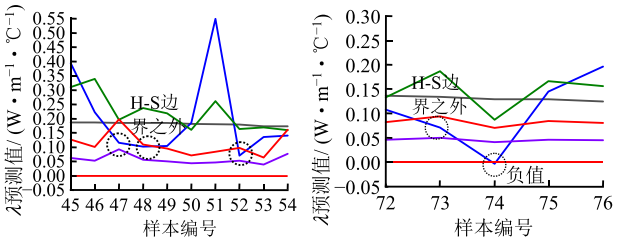
Fig.9　Error distribution of different models

为了进一步验证模型单个样本预测的合理性，对比了不同模型预测结果与 Wiener 边界和 H-S 边界之间的关系，结果如图 10 所示。PXGBoost 模型预测值收敛在真实值附近，仅有极少预测结果落在 H-S 边界之外，而 DXGBoost 模型的波动范围稍大，存在少量负值和 Wiener 边界之外的输出。此外，传感器噪声可能导致 λ 测量值超出界限，但这并不影响 PXGBoost 和 DXGBoost 模型的决策结果。

相较于纯数据驱动模型，融合先验传热知识的 PIEL 模型不仅展现出了较高的精度，且有效提升了模型决策的物理一致性。但 $\lambda>3$ W/(m·℃)时，融合物理信息模型的预测值相较于纯数据驱动模型更发散。原因在于，这些较高 λ 测量数据是由高温条件下对流传热和水汽潜热传输及测量噪声共同造成的[28]，而 H-S 边界中仅考虑了各组分的温度效应，忽视了对流和相变传热基质，导致 H-S 边界在高温条件下的约束趋于保守，进而影响模型训练的收敛范围。这一现象从侧面反映出物理约束提升了模型预测结果的鲁棒性和泛化能力，尤其在处理高噪声数据时，其决策过程更依赖于先验物理信息的引导。因此，进一步融合更全面的传热理论，提升模型对复杂热传输机制的适应能力，以增强 PIEL 框架在全域导热预测中的适用性与精确性。



(a) 不同模型的 λ 预测值



(b) 区间①放大图　　(c) 区间②放大图

图 10　不同模型 λ 预测值与边界之间的对比

Fig.10　Comparison between predicted λ using different models and boundaries

3.4.3 模型的泛化性能

为进一步验证 PIEL 模型的泛化能力，从 3 类模型中各选取排名前两位的共 6 个模型，选用 Zhao 等[46]采集的青藏高原 7 个干旱（80°3′46″E～79°37′50″E，33°27′18″N～32°25′22″N）、4 个半干旱（91°53′55″E～91°46′48″E，31°22′52″N～31°13′48″N）和 8 个半湿润区域（102°9′27″E～101°50′20″E，33°38′58″N～33°55′4″N）表土取样点进行导热系数测试，结果如图 11 所示。统计分析显示，融合物理约束的 PIEL 模型预测性能最佳，RMSE≤0.084，R^2 ≥0.962；其次为数据驱动的机器

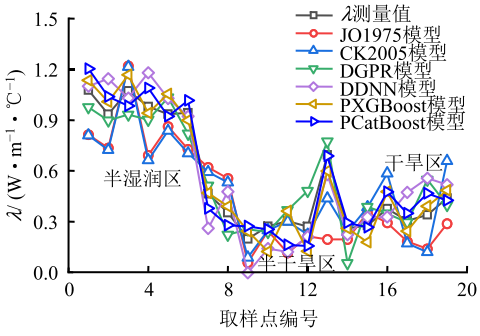


图 11　PIEL 模型反演的青藏高原表土导热系数空间分布图

Fig.11　Inverted spatial distribution map of topsoil thermal conductivity over the Qinghai-Xizang Plateau using PIEL models

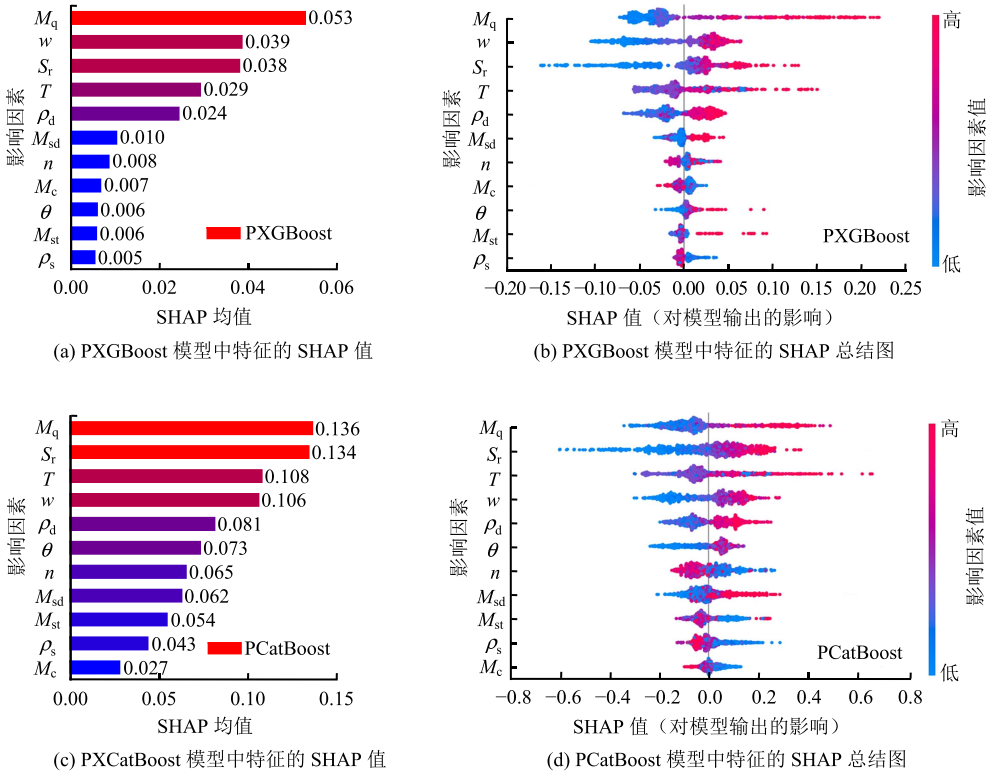
学习模型，RMSE≥0.113，R^2≤0.934；传统模型性能最差，RMSE≥0.178，R^2≤0.68。这表明 PIEL 模型在处理不同区域数据时仍具有良好的泛化能力。下一步将侧重于收集更多现场监测数据，以全面评估 PIEL 模型在不同土体类型和复杂耦合环境条件下的综合性能。同时，将训练完成的 PIEL 模型应用于大尺度网格化土体导热系数分布图的重建，旨在为降雨−蒸发和冻融循环等驱动下的区域水热行为动态模拟提供可靠的参数支持，进一步为多场耦合作用下边坡稳定性和寒区冻融灾害等风险的智能化监测预警提供潜在的技术支撑。

3.4.4 模型解释

为了验证融合先验传热知识的 PIEL 模型的决策过程是否符合非饱和土传热机制，采用 SHAP 和 PDPs 技术从全局和局部贡献率、响应模式和耦合效应 3 个角度揭示了模型的决策过程。鉴于现有可解释算法在多维变量解释的局限性[44]，仅选取关键变量 M_q、S_r、T 和 ρ_d 展开协同效应解释。

（1）因素的全局和局部解释（敏感性）

图 12 给出了排名前两位的模型中 11 个因素的全局和局部重要性。这些因素在不同模型中的累计 SHAP 值存在细微差异，其中 M_q、w、S_r、T 和 ρ_d 在 PXGBoost 模型中排前五位（见图 12(a)），而 M_q、S_r、T、w 和 ρ_d 在 PCatBoost 模型中排前五位（见图 12(b)），这与相关性分析排名（$S_r>\theta>n>\rho_d>M_q$）差异较大，SHAP 值排名更符合先验的传热知识。为了克服全局 SHAP 值提供的有限见解，图 12(c)、12(d)展示了每个因素在每个样本决策过程中的强度和影响趋势。分布图中的每个点代表一个样本，横轴的 SHAP 值表示相应特征对模型输出的贡献，颜色代表影响的正负（红色为正）。M_q 和 S_r 覆盖了最广泛的范围，且数值越大，SHAP 值也越大，表明 M_q、S_r 与 λ 存在正相关关系。其他因素对模型输出的影响较小，仅在某些极端情况占据重要贡献，局部 SHAP 值分析可视化了模型的决策过程，表明模型决策中因素的全局贡献符合先验传热规律。



(a) PXGBoost 模型中特征的 SHAP 值　(b) PXGBoost 模型中特征的 SHAP 总结图

(c) PXCatBoost 模型中特征的 SHAP 值　(d) PCatBoost 模型中特征的 SHAP 总结图

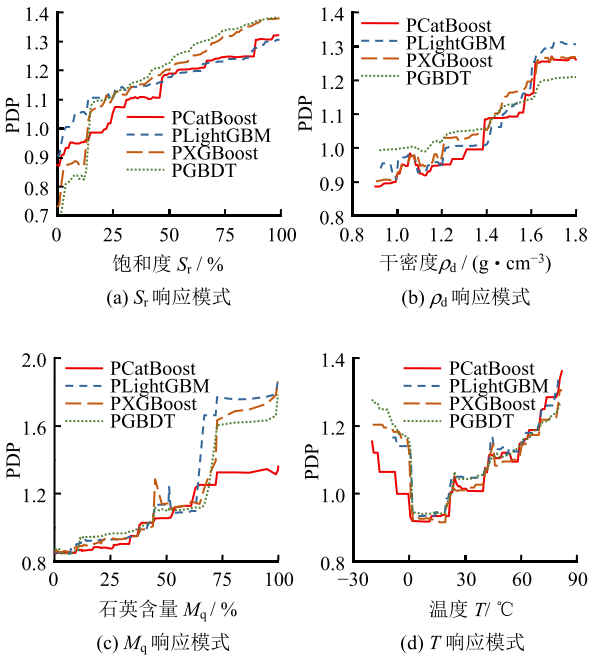
图 12　PXGBoost 和 PCatBoost 模型的全局和局部 SHAP 值

Fig.12　Global and local SHAP values for PXGBoost and PCatBoost models

（2）因素响应模式

图 13 可视化了融合先验传热知识的 PIEL 模型 λ 预测值与关键因素之间的响应模式。λ 随着 S_r、ρ_d、M_q 和 T 的增大而增加，敏感性分别为 0.56、0.33、0.98 和 0.32。部分依赖值 PDP 与 S_r 的关系近似 S 形，较好地捕捉了水形态对传热路径的影响[13, 39]；

PDP 与 ρ_d 的演化速率呈现先增大后减小的趋势，大致可分为缓慢增长段 I（0.9～1.2 g/cm^3）、快速增长段II（1.2～1.5 g/cm^3）、稳定段III（1.5～1.8 g/cm^3）3 个阶段，体现了固体颗粒接触点的数量和面积对传热效率的影响。T 的演化过程类似于 U 形，大致分为 4 个阶段：−20～−5、−5～0、0～40、40～100 ℃。当 T 低于冰点（约−5℃）时，液态水逐渐被冰代替，传热路径由固−液传热变成固−固传热[30]，且后者传热效率更高；反之，随着 T 的升高，尤其是大于 40 ℃时，液态水汽化的蒸汽增强了对流传热和水汽潜热传输，传热效率显著提升[2, 26]。这些潜在的非饱和土传热机制被融合先验传热理论的 PIEL 模型捕捉，表明了模型的决策过程与非饱和土的传热机制相吻合。



(a) S_r 响应模式　(b) ρ_d 响应模式

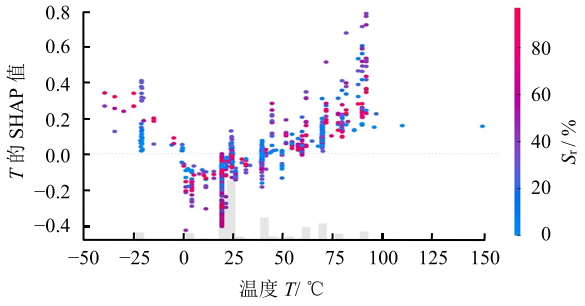
(c) M_q 响应模式　(d) T 响应模式

图 13　PIEL 模型中关键因素的响应模式

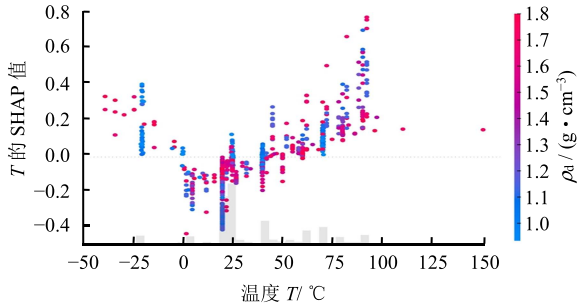
Fig.13　Response patterns of key factors in PIEL model

（3）因素间的耦合效应

图 14、15 给出了融合先验传热知识的 PIEL 模型 λ 预测值与关键因素间的协同作用。单因素响应模式与图 13 的结论完全一致，但不同因素之间会存在耦合效应。图 14(a)说明了 S_r、T 与 λ 的耦合效应，对 T 分段分析，当 $T<0$ 时，S_r 越大，越有助于形成更多的固−固传热路径[30, 41]，相较固−液传热效率显著提升；当 $T>40$ ℃时，随着 S_r 的增加，温度效应呈现出先增大后减小的趋势，在 $S_r=40\%$附近达到峰值。由菲克第一定律可知，蒸汽的传输不仅与蒸汽浓度有关，与传输空间也有关，尽管高温为蒸汽的产生提供了温度前提，但过高的 S_r 占据了水汽潜热的传输路径，温度效应被削弱，这一点在 ρ_d、T 与 λ 的耦合效应中也得以体现，正是因为 ρ_d 的增加导致蒸汽传输路径的减小，水汽潜热传输受限[15, 26]。



(a) S_r 与 T 的协同作用



(b) ρ_d 与 T 的协同作用

注：灰色柱状图表示数据分布。

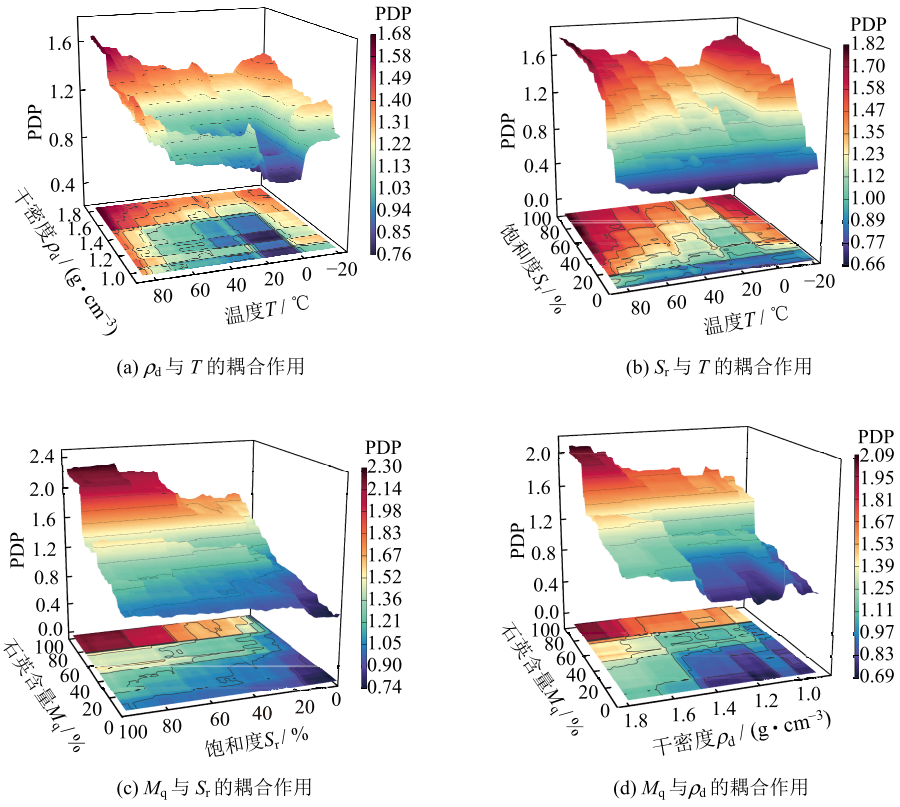
图 14　PXGBoost 模型中关键因素间的协同效应

Fig.14　Synergies between key factors in the PXGBoost model

为弥补 SHAP 定量分析的不足，采用 PDPs 生成二维等高线图，可视化的结果如图 15 所示。单因素响应模式与图 13、14 的规律一致，表明 PXGBoost 模型捕捉物理规律的一致性。ρ_d-T 的耦合响应模式主要受 ρ_d 控制，逐一分析 $T=20$、40、60、80 ℃时，不同 ρ_d 阶段的 PDP 分别为：0.79、0.92、1.03、1.36（I 阶段），0.92、1.03、1.12、1.40（II 阶段）以及 1.21、1.22、1.34、1.46（III 阶段），3 个阶段最大增幅分别为 0.33、0.28、0.12。可见，ρ_d 越小，孔隙空间越大，越有利于水汽潜热传输，反之亦然。但 ρ_d 太小，固体颗粒间的传热效率降低[4, 37]，反而造成 λ 降低。只有当 ρ_d 适中时，两者的耦合效应最显著。S_r-T 的耦合关系更复杂，受到水汽潜热传输、对流传热和水分形态的多重影响，由于土体孔隙较小，以及测量时间较短，忽视对流传热的影响。随着 T 升高，相同 S_r 下的 λ 差异显著，如 S_r 固定为 60%，当 $T=20$、40、60、90 ℃时，PDP 分别为 1.16、1.29、1.49、1.81，可见 T 越高，温度效应越显著[2, 28]。固

定 T 不变，当 S_r 增至某一临界值 S_{rc} 时，PDP 值会出现突变。当 $T=40$、60、80 ℃时，S_{rc} 分别为 48%、32%、13%，可见随着 T 升高，S_{rc} 逐渐减小[33]；当 $S_r > S_{rc}$ 时，水汽潜热传输的空间被水分占据，温度效应便趋于稳定[26, 31]。这一决策过程可解释性的定量分析，表明 PXGBoost 模型较好地捕捉了水汽潜热传热与固-液传热两种机制的博弈过程。ρ_d、S_r 与 M_q 的耦合效应不显著，主要受 M_q 控制。当 ρ_d 位于 I、II、III 阶段，M_q 分别大于 70%、60%、40%时，两者的耦合效应显著增加。当 S_r 介于 0～20%、20%～100%，M_q 分别大于 80%、65%时，协同作用显著提升。

综上所述可知，相较纯数据驱动模型，多维度的精度和泛化能力评价验证了融合先验传热知识的 PIEL 模型决策结果具有更高的物理一致性（见图 10、图 11）；联合 SHAP 和 PDPs 的可解释性分析表明其决策过程与先验的非饱和土传热机制完全吻合（见图 13～15），即使处理测量数据噪声的影响，局部预测结果和响应模式也未出现显著偏差。尽管取得良好表现，PIEL 模型框架仍存在进一步优化空间：①基准模型、整体框架及物理约束形式的选择仍需深入探讨，可考虑引入自适应权重算法或更全面的先验传热知识，以缓解物理约束可能引发的模型精度损失；②当前训练数据的代表性仍有限，未来将汇编更多土体类型以及盐碱环境等复杂工况数据，以提升模型在多场耦合条件下的泛化能力；③SHAP 与 PDPs 技术仅成功揭示了关键变量之间的耦合关系，但受限于现有算法，此类方法仅能揭示两个变量之间的交互效应，难以全面呈现多变量之间的高阶耦合关系。因此，未来可引入更多高维度的可解释性技术，深入拓展对模型内部机制的理解。总之，PIEL 模型较好地解释了主要因素之间的耦合效应，验证了模型决策过程的物理一致性。此框架有望为物理驱动人工智能的水热耦合模拟提供潜在的算法支持。



(a) ρ_d 与 T 的耦合作用　(b) S_r 与 T 的耦合作用

(c) M_q 与 S_r 的耦合作用　(d) M_q 与 ρ_d 的耦合作用

图 15　PXGBoost 模型中关键因素间的耦合作用

Fig. 15　Coupling effects between key factors in the PXGBoost model

3.5　简化 PIEL 模型

为了降低模型的复杂度，提升计算效率，往往需要识别模型的最优预测因子。常见方法有相关性分析、重要性分析和灰度关联分析等[16, 18]，但这些方法会忽视影响因素间的耦合效应和共线性，进而掩盖因素对模型输出的实际贡献。鉴于前文所述 SHAP 技术处理因素耦合效应的优势，采用累计 SHAP 值排名作为最优预测因子的筛选方法，识别

出 M_q、S_r、ρ_d 和 T 作为最优预测因子（见图 16）。这一结果与相关性分析（S_r、θ、n 和 ρ_d）、重要性分析（M_q、M_{sd}、S_r 和 w）和灰度关联分析（ρ_d、w、n 和 S_r）的结果表现出显著差异。显然，相关性分析和灰度关联分析低估了 M_q 的贡献，重要性分析忽视了 M_q 和 M_{sd} 的共线性。

分别以 4 种方法识别的最优预测因子为输入构建简化模型，结果如图 17 所示。对比误差分布图可知，以累计 SHAP 值识别的最优预测因子训练的简化模型精度约是其他模型的 2 倍，表明基于累计 SHAP 排名方法更适合非线性耦合因素的数据降维和特征提取。相较于 11 个因素训练的模型，简化模型计算效率提升了近 7 倍，预测精度虽略有降低，但仍为传统模型的 4 倍，介于工程应用的可接受范围。

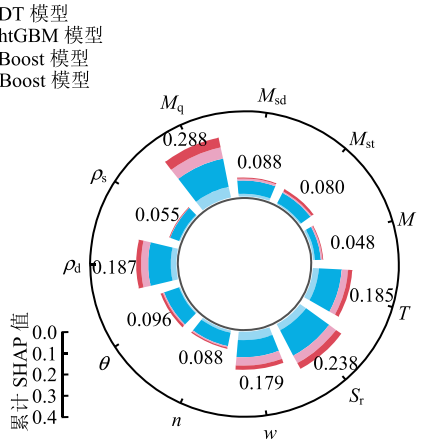
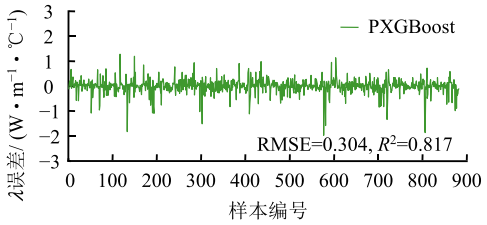
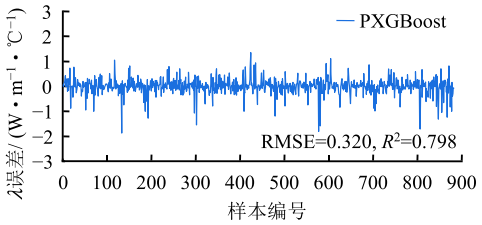


图 16　影响因素的累计 SHAP 值

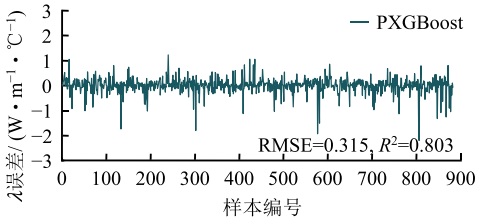
Fig.16　Cumulative SHAP values for the influencing factors



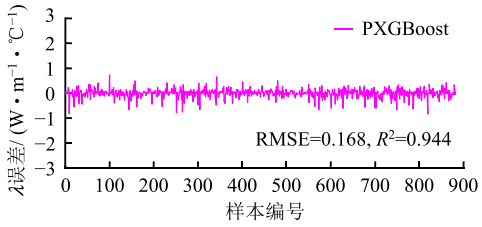
(a) 重要性排名前四位因素(M_q、M_{sd}、S_r 和 w)建立的模型



(b) 灰度关联性排名前四位因素(ρ_d、w、n 和 S_r)建立的模型



(c) 相关性排名前四位因素(S_r、θ、n 和ρ_d)建立的模型



(d) 累计 SHAP 值排名前四位因素(M_q、S_r、ρ_d 和 T)建立的模型

图 17　简化 PIEL 模型的预测误差

Fig.17　Prediction error of the simplified PIEL models

4　结　论

本文提出了一种融合先验物理信息和可解性分析的岩土水热参数 PIEL 框架，以导热系数为案例，系统验证了 PIEL 模型决策过程和结果的物理一致性。主要结论如下：

（1）导热系数是多种因素协同控制的非线性响应模式，传统模型难以系统刻画影响因素间耦合效应、传热路径和传输机制，当前缺乏普适的导热系数预测模型。

（2）纯数据驱动和物理驱动模型均能有效捕捉导热系数的非线性耦合响应模式，展现出了更好的鲁棒性和泛化能力，预测精度较传统模型提升 3～6 倍，而 PIEL 模型的性能最优（RMSE≤0.171，R^2≥0.952）。

（3）融合先验传热知识的 PIEL 模型有效规避了纯数据驱动模型违背物理规律的输出，决策结果基本限定在 H-S 边界内，物理一致性更高；麻雀搜索算法优化的 PXGBoost 模型具有最佳的精度和鲁棒性（RMSE＝0.119，R^2＝0.978）。然而，H-S 边界忽视了高温条件下对流和水汽潜热传输机制，导致 PIEL 模型在高导热区域（λ>3 W/(m・℃)）存在一定的精度损失。

（4）采用 SHAP 和 PDPs 技术可视化的 PIEL 敏感性、响应模式和耦合效应与先验传热知识基本吻合，有效捕捉了水分状态和温度效应引起的传热机

制差异，验证了其决策过程的物理一致性。

（5）采用累计 SHAP 值识别的最优预测因子（M_q、S_r、ρ_d 和 T）有效弥补了传统参数分析忽视的因素耦合效应，更适合非线性耦合因素的数据降维和特征提取。基于此简化的 PXGBoost 模型精度是传统模型的 4 倍，训练效率提高了 7 倍。

融合先验物理知识和可解性分析的 PIEL 模型框架，具有物理一致性、框架灵活、适用范围广等优点，可为人工智能赋能的岩土工程水热耦合研究提供新视野。受限于模型框架、物理约束和可解性技术，未来工作需进一步引入更全面的传热知识和更多可解释性算法，以增强 PIEL 框架在全域导热预测中的适用性与精确性。

参 考 文 献

[1] 张楠, 夏胜全, 侯新宇, 等. 土热传导系数及模型的研究现状和展望[J]. 岩土力学, 2016, 37(6): 1550-1562.
ZHANG Nan, XIA Sheng-quan, HOU Xin-yu, et al. Review on soil thermal conductivity and prediction model[J]. **Rock and Soil Mechanics**, 2016, 37(6): 1550-1562.

[2] 徐云山, 肖子龙, 孙德安, 等. 土体导热系数温度效应及其预测模型[J]. 岩土工程学报, 2023, 45(6): 1180-1189.
XU Yun-shan, XIAO Zi-long, SUN De-an, et al. Temperature effects and prediction model of thermal conductivity of soil[J]. **Chinese Journal of Geotechnical Engineering**, 2023, 45(6): 1180-1189.

[3] GJELSVIK E L, FOSSEN M, TØNDEL K. Current overview and way forward for the use of machine learning in the field of petroleum gas hydrates[J]. **Fuel**, 2023, 334: 126696.

[4] XU Y S, SUN D A, ZENG Z T, et al. Temperature dependence of apparent thermal conductivity of compacted bentonites as buffer material for high-level radioactive waste repository[J]. **Applied Clay Science**, 2019, 174: 10-14.

[5] 刘月妙, 蔡美峰, 王驹. 高放废物处置库缓冲材料导热性能研究[J]. 岩石力学与工程学报, 2007, 26(增刊 2): 3891-3896.
LIU Yue-miao, CAI Mei-feng, WANG Ju. Thermal properties of buffer material for high-level radioactive waste disposal[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2007, 26(Suppl.2): 3891-3896.

[6] TANG A M, CUI Y J, LE T T. A study on the thermal conductivity of compacted bentonites[J]. **Applied Clay Science**, 2008, 41(3-4): 181-189.

[7] YOON S, KIM M J, PARK S, et al. Thermal conductivity prediction model for compacted bentonites considering temperature variations[J]. **Nuclear Engineering and Technology**, 2021, 53(10): 3359-3366.

[8] YOON S, LEE G J, PARK T J, et al. Thermal conductivity evaluation for bentonite buffer materials under elevated temperature conditions[J]. **Case Studies in Thermal Engineering**, 2022, 30: 101792.

[9] TONG F G, JING L R, ZIMMERMAN R W. An effective thermal conductivity model of geological porous media for coupled thermo-hydro-mechanical systems with multiphase flow[J]. **International Journal of Rock Mechanics and Mining Sciences**, 2009, 46(8): 1358-1369.

[10] MCCOMBIE M L, TARNAWSKI V R, BOVESECCHI G, et al. Thermal conductivity of pyroclastic soil (Pozzolana) from the environs of Rome[J]. **International Journal of Thermophysics**, 2017, 38: 1-15.

[11] CHO W J, LEE J O, KWON S. An empirical model for the thermal conductivity of compacted bentonite and a bentonite–sand mixture[J]. **Heat and Mass Transfer**, 2011, 47: 1385-1393.

[12] 陈航, 张虎元, 郭永强, 等. 混合型缓冲回填材料导热性能测试与预测研究[J]. 岩石力学与工程学报, 2014, 33(增刊 2): 4312-4320.
CHEN Hang, ZHANG Hu-yuan, GUO Yong-qiang, et al. Measurement and prediction of thermal properties of bentonite-sand mixtures as buffer backfilling materials for high level radioactive waste[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2014, 33(Suppl.2): 4312-4320.

[13] TARNAWSKI V R, MCCOMBIE M L, MOMOSE T, et al. Thermal conductivity of standard sands. Part III. Full range of saturation[J]. **International Journal of Thermophysics**, 2013, 34: 1130-1147.

[14] YOON S, KIM M J, LEE S R, et al. Thermal conductivity estimation of compacted bentonite buffer materials for a high-level radioactive waste repository[J]. **Nuclear Technology**, 2018, 204(2): 213-226.

[15] 刘晨晖, 周东, 吴恒. 土壤热导率的温度效应试验和预测研究[J]. 岩土工程学报, 2011, 33(12): 1877-1886.
LIU Chen-hui, ZHOU Dong, WU Heng. Measurement and prediction of temperature effects of thermal conductivity of soils[J]. **Chinese Journal of Geotechnical Engineering**, 2011, 33(12): 1877-1886.

[16] ZHAO T Y, LIU S C, XU J, et al. Comparative analysis of seven machine learning algorithms and five empirical models to estimate soil thermal conductivity[J]. **Agricultural and Forest Meteorology**, 2022, 323: 109080.

[17] KARDANI N, BARDHAN A, SAMUI P, et al. Predicting the thermal conductivity of soils using integrated approach of ANN and PSO with adaptive and time-varying acceleration coefficients[J]. **International Journal of Thermal Sciences**, 2022, 173: 107427.

[18] LI K Q, LIU Y, Kang Q. Estimating the thermal conductivity of soils using six machine learning algorithms[J]. **International Communications in Heat and Mass Transfer**, 2022, 136: 106139.

[19] CHU Z X. Combination of the unifying model for the effective thermal conductivity of isotropic, porous and composite geomaterials[J]. **International Journal of Rock Mechanics and Mining Sciences**, 2023, 164: 105342.

[20] 李林, 左林龙, 胡涛涛, 等. 基于孔压静力触探试验测试数据的原位固结系数物理信息神经网络反演方法[J]. 岩土力学, 2024, 45(10): 2889-2899.
LI Lin, ZUO Lin-long, HU Tao-tao, et al. Physics-informed neural networks inversion method for in-situ consolidation coefficient based on piezocone penetration test pore pressure data[J]. **Rock and Soil Mechanics**, 2024, 45(10): 2889-2899.

[21] 张东明, 周烨璐, 黄宏伟, 等. 物理信息双驱动的长距离盾构隧道结构纵向力学性态智能诊断方法[J]. 岩土力学, 2023, 44(10): 2997-3010.
ZHANG Dong-ming, ZHOU Ye-lu, HUANG Hong-wei, et al. A physics and information dual-driven intelligent diagnosis method for longitudinal mechanical behavior of long-distance shield tunnels[J]. **Rock and Soil Mechanics**, 2023, 44(10): 2997-3010.

[22] GONG W B, ZUO L L, LI L, et al. Prediction of stratified ground consolidation via a physics-informed neural network utilizing short-term excess pore water pressure monitoring data[J]. **Computer-Aided Civil and Infrastructure Engineering**, 2025, 40(2): 147-165.

[23] 真嘉捷, 赖丰文, 黄明, 等. 基于 LightGBM-Informer 的盾构隧道管片上浮长时间序列预测模型[J]. 岩土力学, 2024, 45(12): 3791-3801.
ZHEN Jia-jie, LAI Feng-wen, HUANG Ming, et al. Long sequence time series model to predict uplift of segmental lining in shield tunnel based on LightGBM-Informer[J]. **Rock and Soil Mechanics**, 2024, 45(12): 3791-3801.

[24] KASHEM A, KARIM R, MALO S C, et al. Hybrid data-driven approaches to predicting the compressive strength of ultra-high-performance concrete using SHAP and PDP analyses[J]. **Case Studies in Construction Materials**, 2024, 20: e02991.

[25] MADSEN F T. Clay mineralogical investigations related to nuclear waste disposal[J]. **Clay Minerals**, 1998, 33(1): 109-129.

[26] XU Y S, SUN D A, ZENG Z T, et al. Effect of temperature on thermal conductivity of lateritic clays over a wide temperature range[J]. **International Journal of Heat and Mass Transfer**, 2019, 138: 562-570.

[27] TIAN Sheng-kui, WANG Qiong. Compiled data of soil thermal conductivity[EB/OL]. (2025-04-16)[2026-01-07]. https://www.researchgate.net/publication/390832390_Compiled_data_of_soil_thermal_conductivity.

[28] SMITS K M, SAKAKI T, HOWINGTON S E, et al. Temperature dependence of thermal properties of sands across a wide range of temperatures (30–70 ℃)[J]. **Vadose Zone Journal**, 2013, 12(1): vzj2012.0033.

[29] 王海波, 杨平, 何忠意. 孔隙率与饱和度对粉土导热特性的影响[J]. 南京林业大学学报（自然科学版), 2012, 55(2): 42-46.
WANG Hai-bo, YANG Ping, HE Zhong-yi. Influence of porosity and saturation on thermal properties of silty soils[J]. **Journal of Nanjing Forestry University (Natural Science Edition)**, 2012, 55(2): 42-46.

[30] ZHANG M Y, LU J G, LAI Y M, et al. Variation of the thermal conductivity of a silty clay during a freezing-thawing process[J]. **International Journal of Heat and Mass Transfer**, 2018, 124: 1059-1067.

[31] SEPASKHAH A R, BOERSMA L. Thermal conductivity

of soils as a function of temperature and water content[J]. **Soil Science Society of America Journal**, 1979, 43(3): 439-444.

[32] VILLAR M V. Thermo-hydro-mechanical characterisation of a bentonite from Cabo de Gata. A study applied to the use of bentonite as sealing material in high level radioactive waste repositories[D]. Madrid: Complutense University of Madrid, 2021.

[33] NIKOLAEV I V, LEONG W H, ROSEN M A. Experimental investigation of soil thermal conductivity over a wide temperature range[J]. **International Journal of Thermophysics**, 2013, 34: 1110-1129.

[34] KERSTEN M S. Thermal properties of soils[D]. Minneapolis: University of Minnesota Institute of Technology, 1949.

[35] TARNAWSKI V R, TSUCHIYA F, COPPA P, et al. Volcanic soils: inverse modeling of thermal conductivity data[J]. **International Journal of Thermophysics**, 2019, 40: 1-25.

[36] TARNAWSKI V R, MOMOSE T, MCCOMBIE M L, et al. Canadian field soils III. Thermal-conductivity data and modeling[J]. **International Journal of Thermophysics**, 2015, 36: 119-156.

[37] CHEN S X. Thermal conductivity of sands[J]. **Heat and Mass Transfer**, 2008, 44(10): 1241-1246.

[38] ZHANG N, YU X B, PRADHAN A, et al. Thermal conductivity of quartz sands by thermo-time domain reflectometry probe and model prediction[J]. **Journal of Materials in Civil Engineering**, 2015, 27(12): 04015059.

[39] TOKORO T, ISHIKAWA T, SHIRAI S, et al. Estimation methods for thermal conductivity of sandy soil with electrical characteristics[J]. **Soils and Foundations**, 2016, 56(5): 927-936.

[40] MISHRA P N, SURENDRAN S, GADI V K, et al. Generalized approach for determination of thermal conductivity of buffer materials[J]. **Journal of Hazardous, Toxic, and Radioactive Waste**, 2017, 21(4): 04017005.

[41] JOHANSEN O. Thermal conductivity of soils[D]. Trondheim: University of Trondheim, 1975.

[42] CÔTÉ J, KONRAD J M. A generalized thermal conductivity model for soils and construction materials[J]. **Canadian Geotechnical Journal**, 2005, 42(2): 443-458.

[43] 宋牧原, 杨明辉, 陈伟, 等. 基于自注意力−循环神经网络模型的盾构引发的土体沉降预测[J]. 岩土力学, 2025, 46(8): 2613-2625.
SONG Mu-yuan, YANG Ming-hui, CHEN Wei, et al. Prediction of shield tunneling-induced soil settlement based on self-attention recurrent neural network model[J]. **Rock and Soil Mechanics**, 2025, 46(8): 2613-2625.

[44] ZHANG B, XU C, DAI X G, et al. Research on mining land subsidence by intelligent hybrid model based on gradient boosting with categorical features support algorithm[J]. **Journal of Environmental Management**, 2024, 354: 120309.

[45] GHAREHCHOPOGH F S, NAMAZI M, EBRAHIMI L, et al. Advances in sparrow search algorithm: a comprehensive survey[J]. **Archives of Computational Methods in Engineering**, 2023, 30(1): 427-455.

[46] ZHAO H, ZENG Y J, LÜ S N, et al. Analysis of soil hydraulic and thermal properties for land surface modeling over the Tibetan Plateau[J]. **Earth System Science Data**, 2018, 10(2): 1031-1061.