

基于物理信息神经网络能量桩周土体 热 - 力耦合高效求解方法

王志良¹, 肖智桓¹, 申林方^{1*}, 刘海明¹, 李 泽¹, 李 淼²

(1. 昆明理工大学 建筑工程学院, 云南 昆明 650500; 2. 查尔斯特大学 计算机、数学与工程学院, 澳大利亚 巴瑟斯特 2795)

摘要: 为研究能量桩周土体的热 - 力耦合作用, 基于物理信息神经网络 (PINNs) 方法, 提出一种高效的无网格数值求解模型。该模型将热扩散方程与 Navier 应力平衡方程作为物理约束, 以控制方程残差的形式嵌入损失函数中, 从而确保训练过程严格遵循物理守恒定律。与有限元数值计算结果对比验证表明: 所构建的 PINNs 热 - 力耦合模型的最优网络配置为 4 层隐藏层, 每层 50 个神经元以及 Tanh 激活函数, 其温度场和位移场预测的最大相对误差分别为 0.53% 和 5.72%。在优化策略方面, Adam 与 L-BFGS 混合优化使总损失值降至 4.42×10^{-4} , 较单一使用 Adam 和 L-BFGS 优化分别提升了 51.52% 与 15.33%。此外, 拉丁超立方采样策略显著提升了模型精度, 其温度预测平均相对误差较均匀采样与随机采样分别降低了 77.03% 与 76.81%。引入迁移学习框架后, 该模型在新工况下仅需 9.35% 的计算成本即可达到与完整重新训练相当的精度, 整体计算效率较传统有限元方法提升了 16.55%。研究成果可为地下工程热 - 力耦合问题研究提供一种新的兼具物理严谨性和计算高效性的求解方案。

关键词: 桩基础; 物理信息神经网络; 能量桩; 热 - 力耦合作用; 深度学习; 迁移学习

中图分类号: TU 473.1

文献标识码: A

文章编号: 1000 - 6915(2026)03 - 0903 - 15

第一作者: 王志良 (1982 -), 现任教授, 主要从事隧道及地下建筑结构方面的研究工作。E-mail: wangzhiliangtj@126.com

***通信作者:** 申林方 (1982 -), 现任教授, 主要从事岩土工程多场耦合方面的研究工作。E-mail: linfangshen@126.com

An efficient thermo-mechanical coupling solution for soil surrounding energy pile based on physics-informed neural networks

WANG Zhiliang¹, XIAO Zhihuan¹, SHEN Linfang^{1*}, LIU Haiming¹, LI Ze¹, LI Miao²

(1. Faculty of Civil Engineering and Mechanics, Kunming University of Science and Technology, Kunming, Yunnan 650500, China;

2. School of Computing, Mathematics and Engineering, Charles Sturt University, Bathurst 2795, Australia)

Abstract: To investigate the thermo-mechanical coupling for soil surrounding energy pile, an efficient mesh-free numerical model based on the physics-informed neural networks (PINNs) method is proposed. The model incorporates the heat diffusion equation and Navier stress equilibrium equation as physical constraints by embedding the residuals of the governing equations into the loss function, thereby ensuring that the training process strictly adheres to physical conservation laws. Validation against finite element simulation results indicates that the optimal configuration of the proposed PINNs model consists of four hidden layers with 50 neurons per layer and the Tanh activation function, achieving maximum relative errors of only 0.53% and 5.72% for the temperature and displacement fields, respectively. In terms of optimization strategy, the hybrid Adam+L-BFGS optimizer reduces

收稿日期: 2025 - 07 - 31; **修回日期:** 2025 - 10 - 27

基金项目: 国家自然科学基金资助项目 (42167022, 11962008); 云南省基础研究计划重点项目 (202501AS070105)

Supported by the National Natural Science Foundation of China (Grant Nos. 42167022 and 11962008) and Key Project of Yunnan Fundamental Research Program (Grant No. 202501AS070105)

the total loss to 4.42×10^{-4} , improving performance by 51.52% and 15.33% compared with using Adam and L-BFGS individually. Moreover, the Latin hypercube sampling strategy significantly enhances model accuracy, reducing the average relative temperature prediction error by 77.03% and 76.81% compared with uniform and random sampling, respectively. With the introduction of a transfer learning framework, the model achieves comparable accuracy to full retraining under new working conditions at only 9.35% of the computational cost, improving overall computational efficiency by 16.55% compared with the traditional finite element method. The proposed approach provides a new solution for studying thermo-mechanical coupling problems in underground engineering, combining physical rigor with high computational efficiency.

Key words: pile foundations; physics-informed neural networks (PINNs); energy pile; thermo-mechanical coupling; deep learning; transfer learning

First author: WANG Zhiliang (1982 -), professor, is engaged in research on tunnel and underground structural engineering. E-mail: wangzhiliangtj@126.com

***Corresponding author:** SHEN Linfang (1982 -), professor, is engaged in research on multi-field coupling in geotechnical engineering. E-mail: linfangshen@126.com

1 引 言

在全球能源结构转型^[1]与低碳建筑技术^[2]发展的推动下,地源热泵系统^[3]因其高效节能特性,成为建筑供暖制冷领域的重要解决方案。作为该系统的关键组成部分,能量桩^[4]技术将传统桩基结构与地下热交换器结合,实现了建筑结构支撑与地热能开发的双重功能。然而,能量桩在运行过程中同时承受上部结构荷载与周期性温度荷载,桩体与周围土体之间产生复杂的热-力耦合作用,这种耦合效应不仅直接影响能量桩的承载性能,还会显著改变周围土体的热力学响应特征,目前已成为岩土工程领域的研究难点。

目前主要通过试验研究与数值计算 2 种手段研究能量桩-土体系统的热-力耦合作用机制。在试验研究方面,A. Moradshahi 等^[5]基于原位试验,研究了土体热物性参数(导热系数、热膨胀系数和弹性模量等)对能量桩的热-力响应机制;Y. Yin 等^[6]采用能量桩的室内模型试验,对比分析了能量桩单桩和群桩在夏季和冬季条件下的热-力响应差异,发现群桩效应会降低单桩的换热效率,并引起单桩轴力呈现“夏增冬减”的季节性变化特征。然而,现场试验虽能真实反映工程实际情况,但存在成本高、试验周期长等问题;而室内模型试验虽具有可控性优势,却难以精准模拟实际边界条件,且数据采集易受环境干扰,测量精度难以保证。

在数值计算方面,S. Cui 等^[7]考虑应力对热导率的影响,基于有限元法(FEM)建立了饱和/非饱和土

热导率的计算模型,并评估了能量桩和其他地质结构的传热效率;D. Salciarini 等^[8]建立了三维热-力耦合有限元模型,研究了地源热泵运行过程中能量桩与土体间的相互作用机制;T. Razmkhah 等^[9]基于 FLAC^{3D} 的有限差分方法,分析了能量桩的热-力耦合响应,并指出土体刚度增加会导致桩身应变减小而应力增大。尽管有限元法和有限差分法等传统数值方法在多物理场耦合分析中应用广泛,但在处理非线性热-力耦合问题时仍存在显著局限。一方面,这些方法计算效率强烈依赖于网格质量^[10],在模拟复杂几何或梯度剧烈变化时易引发收敛困难;另一方面,在模拟土体的温度与位移演变、桩-土接触面非稳态传热等复杂工况时,控制方程呈现强非线性,这会引发传统有限元法在求解收敛性和数值稳定性方面的固有局限^[11]。往往需要引入过多简化假设来达到数值稳定性的目的,导致计算结果与实际工程响应之间存在显著差异,这已成为提升预测精度亟待破解的核心难题。

近年来,基于深度学习技术发展起来的物理信息神经网络^[12](PINNs),为解决复杂多物理场耦合问题提供了一种新的解决方案。该方法通过将控制方程作为物理约束直接嵌入神经网络的训练过程,突破了传统数值方法依赖网格划分的局限,显著提升了高维非线性耦合问题的求解效率。目前,PINNs 已在固体力学以及传热学等领域得到了成功应用:R. Sharma 和 Y. B. Guo^[13]通过 PINNs 方法实现了激光金属沉积过程中的温度场和应力场的动态演变预测;M. You 等^[14]基于 PINNs 方法直接从温度场重构出多层复合材料结构的热致变形行为,并同时获得

了热诱导结构响应的稳态和瞬态解。上述研究充分证明 PINNs 方法在处理强非线性多物理场耦合问题方面具有独特优势。针对能量桩周土体的热 - 力耦合问题, PINNs 方法能够直接耦合热传导与应力平衡方程, 避免传统方法因网格离散化或本构简化带来的误差; 同时 PINNs 内嵌的物理机制与深度学习的灵活结构, 使其具备良好的迁移学习能力, 能够显著提升多工况情况下的计算效率, 为该问题的深入研究提供了新的研究路径。

为此, 将 PINNs 框架应用于能量桩周土体的热 - 力耦合问题, 通过将热传导方程与 Navier 平衡方程作为物理约束嵌入神经网络, 实现温度场与位移场的高精度同步预测, 并与传统有限元法的计算结果进行对比分析, 验证了该模型的计算精度。在此基础上, 系统探究了神经网络结构、采样策略等关键参数对预测模型性能的影响机制, 进一步通过迁移学习验证, 揭示了 PINNs 在计算效率方面的显著优势。

2 热 - 力耦合作用数学模型

2.1 问题描述

为了研究能量桩与周围土体间的热 - 力耦合作用机制, 将桩简化为刚性体, 并假设桩 - 土界面为理想连接, 建立了如图 1 所示的计算模型。在温度梯度驱动下, 桩体热量通过桩 - 土界面持续传递至周围土体, 引发以下物理演化过程: (1) 热传导主导的温度场演化: 热量以热传导方式由桩体传入土体, 形成随空间演变的温度场; (2) 物理约束条

件下的热 - 力转换: 土体受热膨胀产生的热应变, 在边界约束作用下转化为热应力, 引起应力场重分布; (3) 热应力 - 位移耦合响应: 在应力场与温度场的协同作用下, 土体产生与温度分布呈非线性耦合特征的位移场。这一耦合过程完整描述了能量桩系统从热量传递到力学响应的物理路径, 其耦合作用机制可以通过热扩散方程和 Navier 应力平衡方程进行定量表征。

2.2 控制方程

假设土体为各向同性的线弹性材料, 其稳态热力学条件严格遵循 Navier 方程^[15]:

$$\nabla \cdot \boldsymbol{\sigma} + \mathbf{f} = 0 \quad (1)$$

式中: $\mathbf{f}$ 为体力项; $\boldsymbol{\sigma}$ 为应力, 可表示为

$$\boldsymbol{\sigma} = \lambda \text{tr}(\boldsymbol{\varepsilon}) \mathbf{I} + 2\mu \boldsymbol{\varepsilon} - \beta(T - T_0) \mathbf{I} \quad (2)$$

式中: T 为土体温度; T_0 为参考温度; $\mathbf{I}$ 为单位二阶张量; μ , λ 为 Lamé 常数; β 为热膨胀系数; $\boldsymbol{\varepsilon}$ 为应变。 μ , λ , β 和 $\boldsymbol{\varepsilon}$ 可分别表示为

$$\mu = \frac{E}{2(1+\nu)} \quad (3)$$

$$\lambda = \frac{Ev}{(1+\nu)(1-2\nu)} \quad (4)$$

$$\beta = (3\lambda + 2\mu)\alpha \quad (5)$$

$$\boldsymbol{\varepsilon} = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^T) \quad (6)$$

式中: $\mathbf{u}$ 为位移矢量, α 为土体线性膨胀系数, E 为土体弹性模量, ν 为土体泊松比。

不同方向的应变可表示为

$$\left. \begin{aligned} \varepsilon_{xx} &= \frac{\partial u}{\partial x} \\ \varepsilon_{yy} &= \frac{\partial v}{\partial y} \\ \varepsilon_{xy} &= \varepsilon_{yx} = \frac{1}{2} \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \end{aligned} \right\} \quad (7)$$

热膨胀引起的应力 $\boldsymbol{\sigma}$ 可表示为

$$\left. \begin{aligned} \sigma_{xx} &= \lambda(\varepsilon_{xx} + \varepsilon_{yy}) + 2\mu\varepsilon_{xx} - \beta(T - T_0) \\ \sigma_{yy} &= \lambda(\varepsilon_{xx} + \varepsilon_{yy}) + 2\mu\varepsilon_{yy} - \beta(T - T_0) \\ \sigma_{xy} &= 2\mu\varepsilon_{xy} \end{aligned} \right\} \quad (8)$$

Navier 方程展开为二维稳态力平衡方程:

$$\left. \begin{aligned} \frac{\sigma_{xx}}{\partial x} + \frac{\sigma_{xy}}{\partial y} &= 0 \\ \frac{\sigma_{xy}}{\partial x} + \frac{\sigma_{yy}}{\partial y} &= 0 \end{aligned} \right\} \quad (9)$$

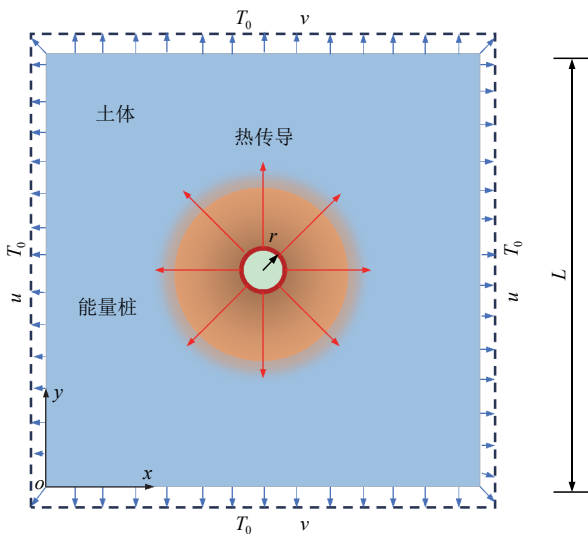


图 1 能量桩与土体间的热 - 力耦合作用模型

Fig.1 Thermo-mechanical coupling model of energy pile and surround soil

将式(8)代入式(9)可得含 u , v , T 的显式表达式:

$$\left. \begin{aligned} & \frac{\partial}{\partial x} \left[\lambda \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + 2\mu \frac{\partial u}{\partial x} - \beta(T - T_0) \right] + \\ & \frac{\partial}{\partial y} \left[\mu \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right] = 0 \\ & \frac{\partial}{\partial x} \left[\mu \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right] + \frac{\partial}{\partial y} \left[\lambda \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + \right. \\ & \left. 2\mu \frac{\partial v}{\partial y} - \beta(T - T_0) \right] = 0 \end{aligned} \right\} \quad (10)$$

基于傅里叶定律和能量守恒原理, 二维热传导的控制方程为

$$\rho c \frac{\partial T}{\partial t} - k \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) = Q \quad (11)$$

式中: k 为土体导热系数, ρ 为密度, c 为比热容, Q 为源项。

假设计算域和外界环境之间不存在热交换, 则控制方程可简化为拉普拉斯方程的稳态形式

$$\nabla(k \nabla T) = 0 \quad (12)$$

2.3 边界条件

计算模型如图 1 所示, 计算区域设置为: $0 \leq x \leq L$, $0 \leq y \leq L$, 能量桩处于计算域中心(x_c , y_c), 其半径为 r , 极角 $\varphi \in [0, 2\pi]$ 。热边界条件: 桩表面温度恒为 T_p ; 计算域外边界的温度设置为 T_0 。基于刚性桩假设, 桩-土界面处位移为 0。限制土体外边界的法向位移, 允许其切向自由变形。相应的边界条件可表示为

$$\left. \begin{aligned} & v(x, 0) = 0, \quad v(x, L) = 0 \\ & u(0, y) = 0, \quad u(L, y) = 0 \\ & T(x, 0) = T_0, \quad T(x, L) = T_0 \\ & T(0, y) = T_0, \quad T(L, y) = T_0 \\ & T(x_c + r \cos \varphi, y_c + r \sin \varphi) = T_p \\ & v(x_c + r \cos \varphi, y_c + r \sin \varphi) = 0 \\ & u(x_c + r \cos \varphi, y_c + r \sin \varphi) = 0 \end{aligned} \right\} \quad (13)$$

2.4 量纲一化处理

为了确保不同物理量的计算尺度保持一致, 并提高 PINNs 计算模型的训练效率和收敛速率, 针对热-力耦合问题所涉及的物理量, 进行量纲一化处理, 见表 1。

表 1 量纲一的变量定义

Table 1 Definition of dimensionless variables

物理量 (physical parameters)	量纲一的变换 (dimensionless transformation)
x, y	$x = Lx^*, y = Ly^*$
u, v	$u = Uu^*, v = Uv^*$
T	$T = T_0 + \Delta T T^*$

注: “*” 为无量纲变量; L 为特征长度, 取值为土体计算域长度; U 为特征位移, 取值为土体最大位移的估计量; ΔT 为桩身温度与边界温度之差, $\Delta T = T_p - T_0$ 。

将式(10)和(12)进行量纲一化处理, 得到以下无量纲方程组:

$$\left. \begin{aligned} & \frac{\partial^2 T^*}{\partial x^{*2}} + \frac{\partial^2 T^*}{\partial y^{*2}} = 0 \\ & \frac{\partial^2 u^*}{\partial x^{*2}} + \left(\frac{\lambda + \mu}{\lambda + 2\mu} \right) \frac{\partial^2 v^*}{\partial x \partial y} + \left(\frac{\mu}{\lambda + 2\mu} \right) \frac{\partial^2 u^*}{\partial y^{*2}} = K \frac{\partial T^*}{\partial x^*} \\ & \frac{\partial^2 v^*}{\partial y^{*2}} + \left(\frac{\lambda + \mu}{\lambda + 2\mu} \right) \frac{\partial^2 u^*}{\partial x \partial y} + \left(\frac{\mu}{\lambda + 2\mu} \right) \frac{\partial^2 v^*}{\partial x^{*2}} = K \frac{\partial T^*}{\partial y^*} \end{aligned} \right\} \quad (14)$$

式中: K 为温度梯度相对位移的驱动力强度, 可表示为:

$$K = \frac{\beta L \Delta T}{(\lambda + 2\mu)U}。$$

3 物理信息神经网络

基于物理约束的深度学习, 将 Navier 方程和热传导控制方程作为正则化项嵌入神经网络训练过程。具体而言, 首先构建全连接神经网络预测架构, 用于同时预测位移场(u , v)和温度场(T)的空间分布, 并利用自动微分技术精确计算各物理量的空间导数项。在此基础上, 建立包含控制方程残差和边界条件误差的复合损失函数, 并采用梯度下降算法对网络权重参数进行迭代优化求解。这种基于物理信息驱动的建模方法, 既继承了神经网络强大的非线性映射能力, 又保持了物理模型的严谨性。

3.1 前向传播

深度神经网络(DNN)通过前向传播计算实现输入数据与预测结果间的映射关系, 如图 2 所示。该网络架构采用分层设计, 包含数据输入层、多个隐藏层和输出层。其物理场的求解过程, 如图 3 所示, 首先对计算域进行离散采样, 然后基于多层非线性变换处理提取特征, 最终结合自动微分技术实现控制方程残差的解析计算和目标函数的优化求解。

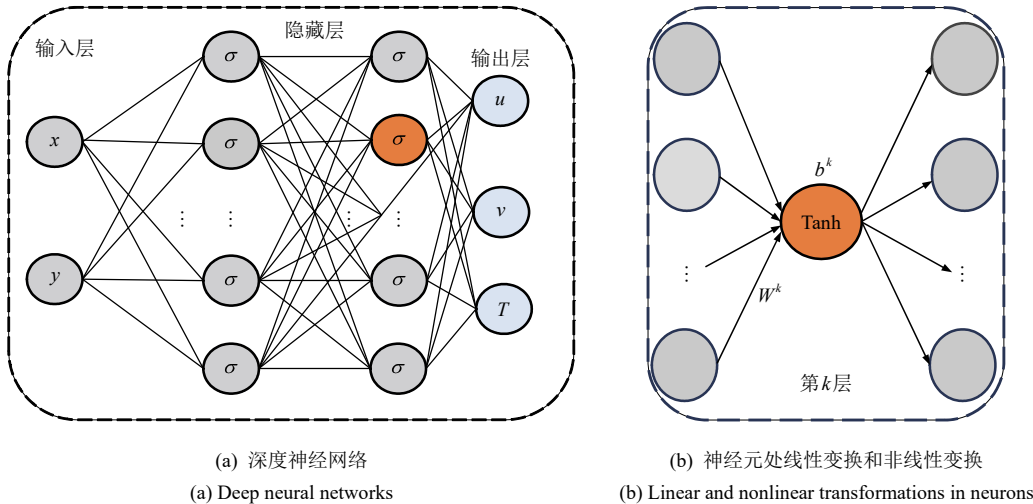


图 2 深度神经网络示意图

Fig.2 Schematic diagram of deep neural networks

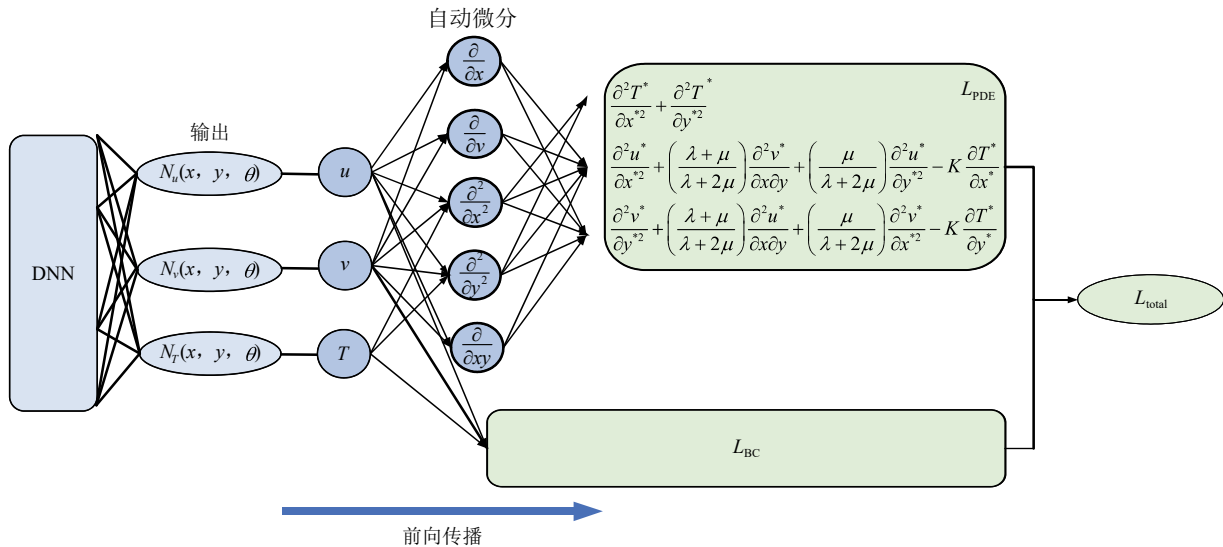


图 3 物理场前向传播示意图

Fig.3 Schematic diagram of physical field forward propagation

3.1.1 输入数据采样

在求解过程中，采用结构化空间离散方法对计算域进行预处理。针对二维热 - 力耦合问题，以空间坐标 (x, y) 作为神经网络输入变量，通过优化采样策略在计算域内构建训练样本集。如图 4 所示，根据物理场求解需求，将采样点分为两类：(1) 绿色标记的内部采样点，用于构建控制方程残差项，确保满足偏微分方程约束；(2) 黑色标记的边界采样点，严格施加 Dirichlet 边界条件。特别地，在能量桩 - 土体界面区域采用局部加密采样策略，以准确捕捉界面处的强梯度变化特征。精细化区域为边长 L_0 的正方形。这种融合物理约束的采样方法，有效平衡了计算精度与效率的需求。

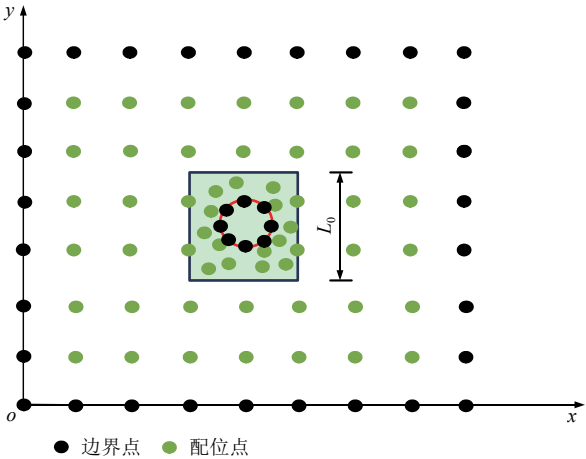


图 4 输入数据采样示意图

Fig.4 Schematic diagram of input data sampling

3.1.2 输入层到隐藏层的传播

将采样得到的输入变量 $\xi = (x, y)$ 传递到隐藏层, 并进行线性变换和非线性变换的交替操作:

$$z^{(1)} = W^{(1)}\xi + b^{(1)} \quad (15)$$

$$a^{(1)} = \sigma(z^{(1)}) \quad (16)$$

$\vdots$

$$z^{(k)} = W^{(k)}a^{(k-1)} + b^{(k)} \quad (17)$$

$$a^{(k)} = \sigma(z^{(k)}) \quad (18)$$

式中: $W^{(k)}$ 为第 k 层的权重矩阵; $b^{(k)}$ 为第 k 层的偏置; $z^{(k)}$ 为神经网络中第 k 层的线性变换的结果; $a^{(k)}$ 为神经网络中第 k 层的激活值; σ 为激活函数的通用占位符, 其作用是将线性模型转化为非线性模型。常用的激活函数主要有以下 3 种:

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (19)$$

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (20)$$

$$\text{ReLU}(x) = \max(0, x) \quad (21)$$

其中, Tanh 函数具有零中心、平滑有界的特性, 适用于需要精确计算高阶导数和生成平滑输出的场景; Sigmoid 函数能将任意输入值“压缩”到 0~1 范围, 适合表示概率, 传统上多用于二分类问题的输出层; ReLU 函数凭借计算高效的优点, 适用于计算资源敏感但对输出平滑性要求不高的场景。

每经过一层隐藏层, 依次进行线性变换和激活函数的非线性变换, 如图 2(b)所示, 直至到达最后一层隐藏层。

3.1.3 输出层的计算

在神经网络的前向传播过程中, 隐藏层末端的特征表示通过线性加权运算映射至输出层。输出层采用线性变换, 其数学表达式为

$$\hat{u} = W^{(l)}a^{(l-1)} + b^{(l)} \quad (22)$$

式中: $W^{(l)}$ 为最后一层的权重矩阵; $b^{(l)}$ 为最后一层的偏置; $\hat{u}$ 为神经网络中的输出, 其分量(N_u, N_v, N_T)为 x, y 方向位移与温度的待构造变量。

为了提高计算模型的准确性和收敛性, 将边界条件加入损失函数计算, 通过构造神经网络解的函数形式来满足大部分边界条件(除能量桩-土体界面温度和位移边界条件外), 将 u, v 和 T 设计如下

$$\left. \begin{aligned} u &= x(L-x)y(L-y)N_u(x, y, \theta) \\ v &= x(L-x)y(L-y)N_v(x, y, \theta) \\ T &= x(L-x)y(L-y)N_T(x, y, \theta) + T_0 \end{aligned} \right\} \quad (23)$$

3.1.4 自动微分及残差计算

u, v 和 T 的所有关于空间坐标 x, y 偏导数项, 通过调用 PyTorch 框架^[16]的自动微分引擎计算得到。具体而言, 在定义了神经网络的前向传播函数后, 将空间坐标(x, y)定义为需要追踪梯度的张量, 直接调用 `torch.autograd.grad(·)` 函数, 即可精确计算一阶导数; 在此基础上, 通过对此一阶导数结果再次调用该函数, 求得控制方程中所需的二阶偏导数。将网络预测的物理量及其导数项代入物理控制方程, 并通过计算方程左右两侧的数值偏差, 定量评估当前预测结果与理论方程的符合程度, 从而获得残差项。以式(14)的温度扩散方程为例, 单个采样点残差为

$$r_i = \frac{\partial^2 T^*}{\partial x_i^{*2}} + \frac{\partial^2 T^*}{\partial y_i^{*2}} \quad (24)$$

3.1.5 损失计算

通过残差项的计算, 进一步定义损失函数。首先, 物理方程的残差采用均方根计算得到物理损失 L_{PDE} ; 其次, 将边界条件预测值与真实值之间的偏差, 也采用均方根计算得到边界损失 L_{BC} ; 最后, 将两部分加权求和构成总损失函数:

$$L_{\text{total}} = \lambda_1 L_{\text{PDE}} + \lambda_2 L_{\text{BC}} \quad (25)$$

式中: λ_1, λ_2 分别为物理损失 L_{PDE} 和边界损失 L_{BC} 的权重系数, 用于平衡两者在总损失中的贡献。对于本研究问题, 通过试算发现当 $\lambda_1 = 1, \lambda_2 = 100$ 时, 两项损失能以相近的速率协同下降, 从而优化过程更为平稳。 $L_{\text{PDE}}, L_{\text{BC}}$ 可分别表示为

$$L_{\text{PDE}} = \frac{1}{N_p} \left\{ \sum_{i=1}^{N_p} \left| \frac{\partial^2 T^*}{\partial x^{*2}} + \frac{\partial^2 T^*}{\partial y^{*2}} \right|^2 + \sum_{i=1}^{N_p} \left| \frac{\partial^2 u^*}{\partial x^{*2}} + \left(\frac{\lambda + \mu}{\lambda + 2\mu} \right) \frac{\partial^2 v^*}{\partial x \partial y} + \left(\frac{\mu}{\lambda + 2\mu} \right) \frac{\partial^2 u^*}{\partial y^{*2}} - K \frac{\partial T^*}{\partial x^*} \right|^2 + \sum_{i=1}^{N_p} \left| \frac{\partial^2 v^*}{\partial y^{*2}} + \left(\frac{\lambda + \mu}{\lambda + 2\mu} \right) \frac{\partial^2 u^*}{\partial x \partial y} + \left(\frac{\mu}{\lambda + 2\mu} \right) \frac{\partial^2 v^*}{\partial x^{*2}} - K \frac{\partial T^*}{\partial y^*} \right|^2 \right\} \quad (26)$$

$$L_{\text{BC}} = \frac{1}{N_b} \left(\sum_{i=1}^{N_b} |u_p - u_{\text{BC}}|^2 + \sum_{i=1}^{N_b} |v_p - v_{\text{BC}}|^2 + \sum_{i=1}^{N_b} |T_p - T_{\text{BC}}|^2 \right) \quad (27)$$

式中: N_p, N_b 分别为物理损失和边界损失采样点数量; u_p, v_p, T_p 分别为 PINNs 模型预测的 x, y 方向位移和能量桩桩周温度; $u_{\text{BC}}, v_{\text{BC}}$ 和 T_{BC} 可分别为边界条件的 x, y 方向位移和能量桩桩周温度。相关取值可由式(13)确定。

3.2 反向传播

在完成前向传播计算后, 神经网络进入参数优化阶段。如图 5 所示, 反向传播通过优化损失函数来更新神经网络参数, 以提高模型预测精度。其计算过程主要包括损失函数梯度计算、参数更新和模型预测的重新计算。

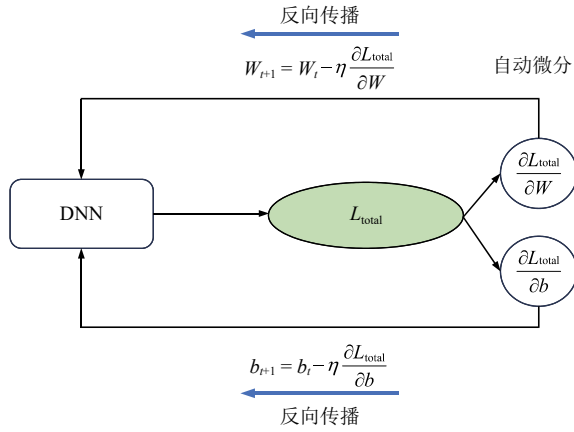


图 5 物理场反向传播示意图

Fig.5 Schematic diagram of physical field back propagation

3.2.1 梯度计算

通过自动微分工具计算总损失函数 L_{total} 对神经网络参数(权重 W 和偏置 b)的梯度:

$$\nabla_W L_{total} = \frac{\partial L_{total}}{\partial W} \quad (28)$$

$$\nabla_b L_{total} = \frac{\partial L_{total}}{\partial b} \quad (29)$$

误差反向传播过程通过计算图的拓扑结构实现参数优化。神经网络的分层架构决定了参数影响的传递特性: 各层的权重参数不仅决定当前层的特征表示, 还会通过层间连接影响后续所有层的输出。基于微分链式法则, 梯度信息从输出层开始, 沿着网络连接的反向路径逐层传递。在这个过程中, 损失函数 L_{total} 对神经网络参数的梯度通过连续乘法运

算实现误差信号的累积和传播, 使得网络能够精确量化每一个参数对整体损失的全局贡献, 从而实现从深层特征依赖关系的有效建模。

3.2.2 参数更新

基于损失函数 L_{total} 对模型权重 W 和偏置 b 的梯度 $\nabla_W L_{total}$ 和 $\nabla_b L_{total}$ 进行调整, 采用自适应动量估计算法^[17](Adam)或 L-BFGS 算法^[18]更新网络参数, 其更新公式为

$$W_{t+1} = W_t - \eta \nabla_W L_{total} \quad (30)$$

$$b_{t+1} = b_t - \eta \nabla_b L_{total} \quad (31)$$

式中: t 为迭代步数; W_t , W_{t+1} 分别为第 t 次和第 $t+1$ 次迭代时的权重矩阵; b_t , b_{t+1} 分别为第 t 次和第 $t+1$ 次迭代时的偏置矩阵; η 为学习率, 用来控制参数更新的步幅。

通过梯度下降法不断优化网络参数, 直至满足收敛条件($L_{total} \leq \tau$, 其中 τ 为损失函数的收敛判定阈值)或达到最大迭代次数。

3.2.3 更新后的模型预测

在参数更新完成后, 基于调整后的权重矩阵 W 和偏置向量 b 重新进行前向传播计算, 获得更新后的物理场预测值。这些新的预测结果随后再次经过自动微分处理, 并用于计算物理损失和边界损失, 进而根据新的误差分布优化损失函数, 完成一次完整的训练迭代。通过这一过程循环迭代, 直至预测模型达到收敛标准, 并获得最终计算结果。

3.3 计算流程体系

PINNs 通过损失函数设计架构(包括控制方程残差和边界条件误差), 建立了一个严格遵循物理规律的智能求解体系。图 6 为 PINNs 从多物理场坐标输入到计算结果预测的理论计算框架, 其详细的数值计算流程见图 7。其中, $Epochs$ 为训练周期, θ 为权重矩阵和偏置向量的组合。

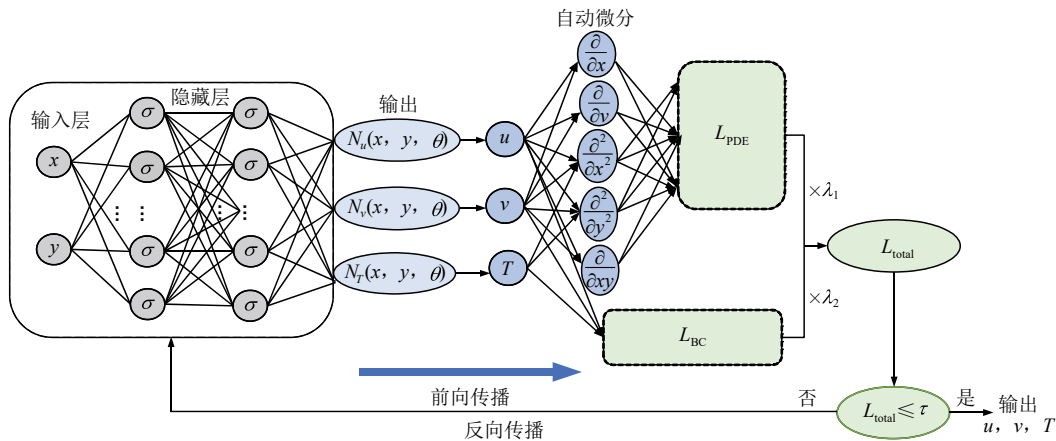


图 6 PINNs 的框架示意图

Fig.6 Schematic diagram of PINNs framework

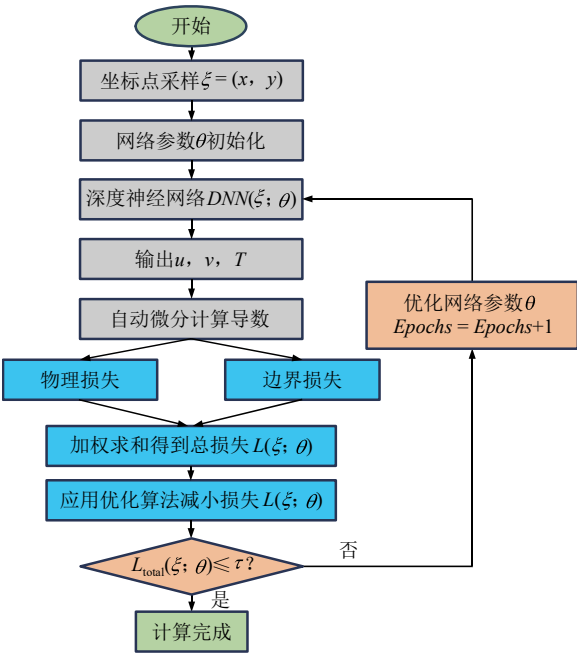


图 7 PINNs 数值计算流程图

Fig.7 Flow chart of PINNs numerical calculation

4 讨论分析

针对能量桩与周围土体的热 - 力耦合作用机制, 建立了图 1 所示的计算模型。模型的几何尺寸为 $L \times L = 5 \text{ m} \times 5 \text{ m}$, 能量桩半径 $r = 0.25 \text{ m}$, 精化区域边长 $L_0 = 1 \text{ m}$ 。桩周土体为膨胀土, 其热膨胀系数 $\alpha = 3 \times 10^{-5} / \text{K}$, 比热容 $c_p = 1\,800 \text{ J}/(\text{kg} \cdot \text{K})$, 导热率 $k = 0.8 \text{ W}/(\text{m} \cdot \text{K})$, 密度 $\rho = 1\,800 \text{ kg}/\text{m}^3$, 弹性模量为 $E = 15 \text{ MPa}$, 泊松比 $\nu = 0.35$ 。土体外边界温度 $T_0 = 10\text{ }^\circ\text{C}$, 能量桩表面温度 $T_1 = 30\text{ }^\circ\text{C}$ 。

4.1 最佳模型配置

为建立最优化的 PINNs 热 - 力耦合计算模型, 从神经网络架构、激活函数和优化算法 3 个方面开展参数敏感性分析。在神经网络架构方面, 重点考察了隐藏层层数(2~6 层)和神经元数量(20~80 个)

对预测模型计算精度的影响; 在激活函数选择上, 对比分析了 Tanh, Sigmoid 和 ReLU 三种函数的性能差异; 在优化算法方面, 考察了 Adam, L-BFGS 及其混合策略的收敛特性。采用均方误差 MSE 、均方根误差 $RMSE$ 、平均绝对误差 MAE 和平均相对误差 MRE 四项指标^[19]对 PINNs 模型的预测性能进行评估。

4.1.1 神经网络配置

为研究神经网络架构对 PINNs 热 - 力耦合计算模型预测精度的影响, 在保持其他参数恒定的条件下(内部配点数量 10 000 个, 精化区域配点 2 500 个, 激活函数采用 Tanh 函数), 对不同网络结构配置下的模型性能进行研究, 将预测结果与有限元数值解进行对比, 并计算相对误差来评价模型精度。

表 2 为不同网络架构下(隐藏层数及每层神经元)PINNs 模型的温度和位移预测结果与有限元计算结果的误差对比。由表 2 可知: 4 层 50 个神经元的网络配置在所有评估指标(MSE , $RMSE$, MAE , MRE)中均表现出最优性能。以最具代表性的 MRE 指标为例, 温度场预测精度随网络深度增加呈先增加后下降趋势: 当层数 2 层增至 4 层时, MRE 由 37.05%降至 0.51%, 而继续增至 6 层时则出现约 5% 的精度退化。类似地, 神经元数量在 50 个时达到最佳预测精度。值得注意的是, 温度预测对网络参数表现出更高的敏感性, 其 MRE 波动范围(37.05%~0.51%)远大于位移场(27.33%~3.92%)。同时, 温度(最优 $MRE = 0.51\%$)要显著优于位移预测精度(最优 $MRE = 3.92\%$), 这一差异主要源于温度场演化主要由热传导单一的物理过程主导; 而位移场则由温度场通过热膨胀驱动产生, 其求解涉及热 - 力耦合作用, 物理路径更为复杂。其他指标(MSE , $RMSE$, MAE)的变化趋势与 MRE 保持一致, 进一步验证了上述结论的可靠性。

表 2 网络结构对预测结果误差的影响

Table 2 Influence of neural network structure on the error of prediction outcomes

[隐藏层数, 神经元数] [hidden layers, neurons]	MSE		RMSE		MAE		MRE/%	
	温度	位移	温度	位移	温度	位移	温度	位移
	(temperature)/ $^\circ\text{C}$	(displacement)/ m^2	(temperature)/ $^\circ\text{C}$	(displacement)/ m	(temperature)/ $^\circ\text{C}$	(displacement)/ m	(temperature)	(displacement)
[2, 20]	2.436 2	1.63×10^{-9}	1.560 8	4.04×10^{-4}	1.468 1	3.76×10^{-5}	37.05	27.33
[2, 50]	1.683 5	9.49×10^{-10}	1.297 5	3.08×10^{-5}	1.182 7	2.88×10^{-5}	33.52	18.09
[2, 80]	1.937 6	4.56×10^{-10}	1.392 0	2.14×10^{-5}	1.235 7	1.99×10^{-5}	34.81	11.80
[4, 20]	0.276 3	2.79×10^{-10}	0.525 6	1.67×10^{-5}	0.404 7	1.53×10^{-5}	2.32	10.77
[4, 50]	0.017 1	2.45×10^{-11}	0.130 3	4.95×10^{-6}	0.088 9	4.08×10^{-6}	0.51	3.92
[4, 80]	0.330 6	2.38×10^{-10}	0.575 0	1.55×10^{-5}	0.462 0	1.40×10^{-5}	2.68	8.36
[6, 20]	0.332 3	6.81×10^{-10}	0.576 5	2.61×10^{-5}	0.452 0	2.17×10^{-5}	2.54	13.12
[6, 50]	0.263 6	2.92×10^{-10}	0.513 4	1.51×10^{-5}	0.402 9	1.38×10^{-5}	2.29	8.47
[6, 80]	0.374 1	2.98×10^{-10}	0.611 6	1.73×10^{-5}	0.486 2	1.59×10^{-5}	2.77	9.10

4.1.2 激活函数选择

在神经网络求解热-力耦合问题时, 激活函数的选择对预测模型性能具有决定性影响^[20]。为此, 基于前期优化网络架构(4 层隐藏层, 每层 50 个神经元), 在固定采样方案(全域 10 000 个内部配点及 2 500 个精化区域配点)下, 选择 Tanh, Sigmoid 和 ReLU 三种典型激活函数, 探究其数学特性对预测精度的影响机制。

图 8 和 9 分别为不同激活函数下温度场和位移场的 PINNs 预测结果与 FEM 数值计算结果对比。在温度场预测方面, 由于 ReLU 函数分段线性的数学特性与真实温度场应光滑连续的物理规律相冲突, 导致最大相对误差达 37.62%; Sigmoid 函数虽能生成连续的温度分布, 但由于其输出值恒为正, 这种固有的偏差会扭曲温度场的整体分布, 导致预测结果系统性偏离真实解。尤其在高梯度变化的桩周区域造成显著误差, 最大误差达到 8.60%, 并在该区域出现明显的误差集中。相比之下, Tanh 函数展现出最优预测性能, 其关于原点对称的特性使其能无偏差地学习, 如热胀冷缩这种方向相反的位移变化, 最大误差仅 0.53%, 即使在热梯度较大的桩-土界面区域仍保持较高的计算精度。在位移场预测方面, ReLU 函数产生明显的非物理振荡, 最大误差高达 47.14%, 桩-土界面处误差尤为显著; Sigmoid 函数虽改善了位移分布的连续性, 但最大误差反而升至 87.26%, 其非对称性导致结果出现系统性偏差; 而 Tanh 函数的预测结果与 FEM 的计算结果高度吻合, 最大误差仅为 5.72%, 能够精确捕捉热致位移场的分布特征。故 Tanh 函数在热-力耦合问题的温度场与位移场的预测方面具有显著优势。在模型配置方面, 确定的最优网络架构为 4 个隐藏层与 50 个神经元, 并选用 Tanh 作为激活函数。该配置不仅在本案例中表现出色, 同时也与 K. Q. Li 等^[21]在冻结管温度场研究中的结论一致, 共同印证了该类配置在求解岩土工程稳态场问题中具有良好的普适性与鲁棒性。

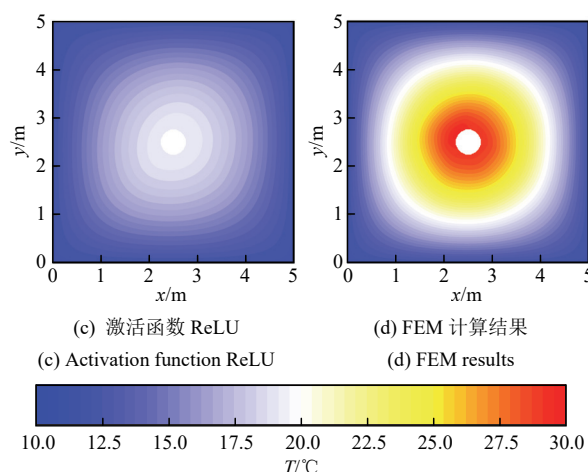
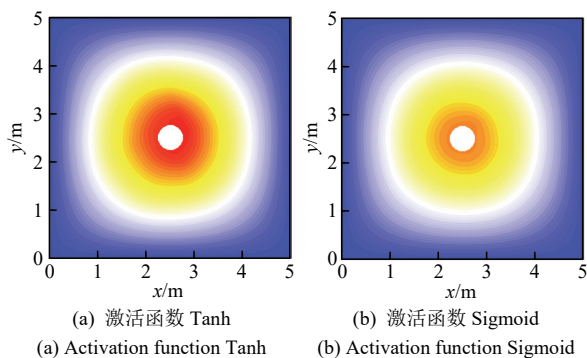


图 8 不同激活函数下温度场的 PINNs 预测结果与 FEM 结果对比

Fig.8 Comparison of PINNs predicted results and FEM results for temperature field under different activation functions

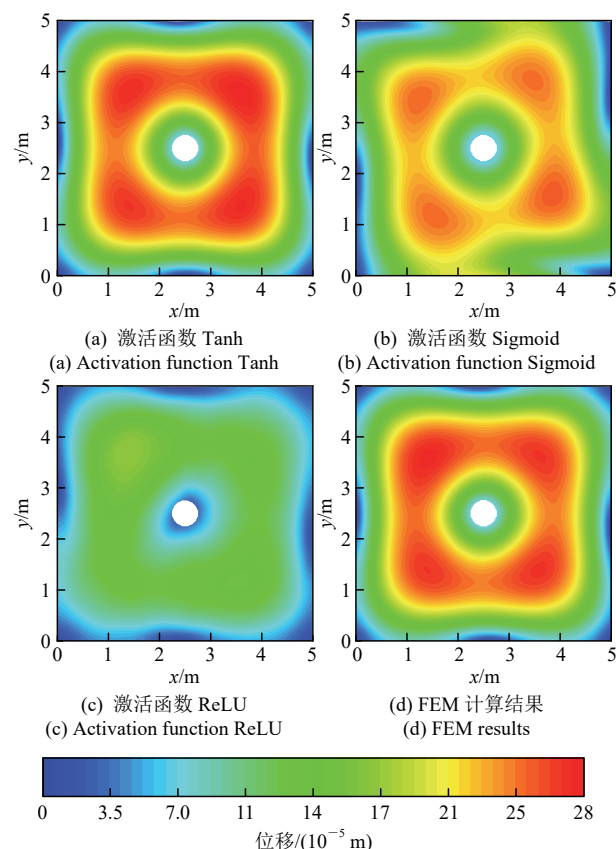


图 9 不同激活函数下位移场的 PINNs 预测结果与 FEM 结果对比

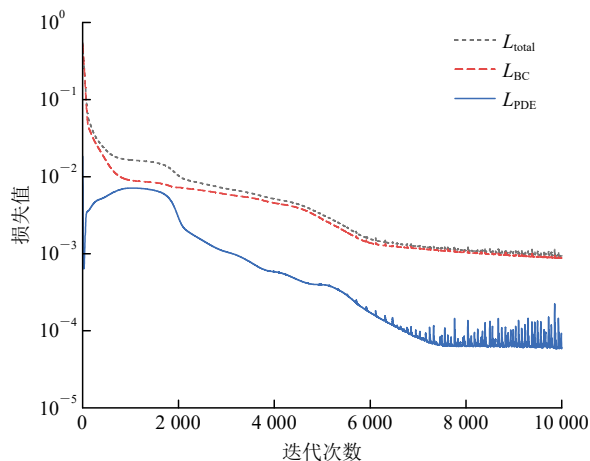
Fig.9 Comparison of PINNs predicted results and FEM results for displacement field under different activation functions

4.1.3 优化策略选择

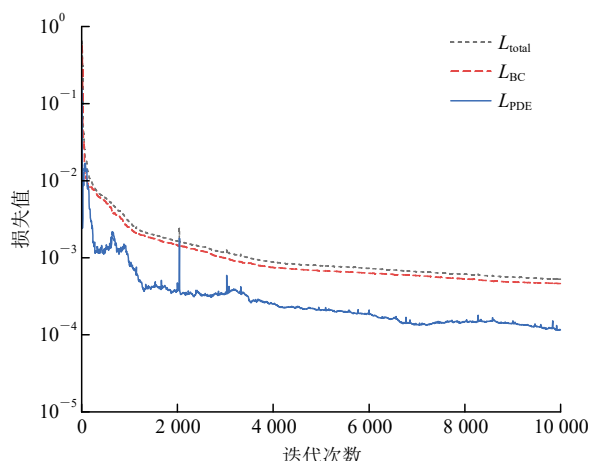
为揭示优化策略对 PINNs 模型收敛特性的调控作用, 基于最优网络参数(4 层隐藏层, 50 个神经元, Tanh 激活函数)和固定采样方案(全域 10 000 个内部配点及 2 500 个精化区域配点), 对比分析了 Adam, L-BFGS 及其混合优化策略在热-力耦合问题中的

性能表现。

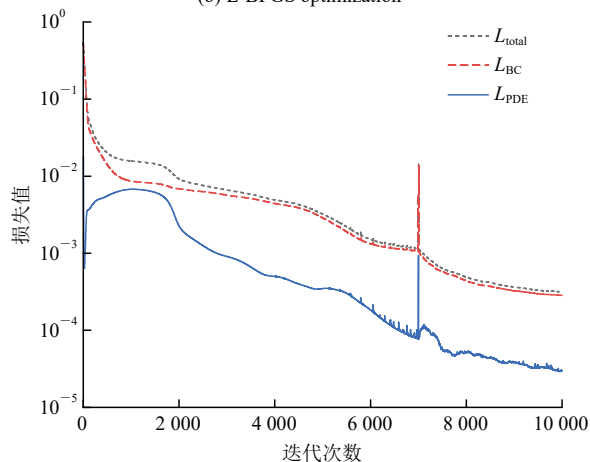
图 10 为不同优化策略下损失函数随迭代次数的演化过程。在 Adam 优化器的独立测试中, 经过 10 000 次迭代训练, 模型呈现显著的二阶段收敛特性, 总损失值从初始值 1.03 降至 9.31×10^{-4} , 其中



(a) Adam 优化
(a) Adam optimization



(b) L-BFGS 优化
(b) L-BFGS optimization



(c) Adam+L-BFGS 混合优化
(c) Adam+L-BFGS hybrid optimization

图 10 不同优化策略下损失函数随迭代次数的演化过程

Fig.10 Evolution process of loss function with iteration number under different optimization strategies

边界损失作为显式约束在前 1 000 步内即实现快速收敛。而物理损失由于涉及全域多物理场耦合计算, 其收敛速度相对较为滞后, 当迭代至 7 000 步时, 损失函数进入明显的收敛平台期, 这主要源于 2 个关键因素: (1) 固定学习率策略在优化后期难以自适应调整步长, 导致效率显著下降; (2) 一阶优化方法对控制方程强非线性特征的敏感性不足, 限制了其对复杂物理耦合机制的精确表征。

在 L-BFGS 优化器的独立测试中, 经过 10 000 步迭代训练, 模型表现出独特的收敛特性。总损失从初始值 1.03 下降至 5.22×10^{-4} , 且物理损失在优化后期未出现明显振荡, 表现出较 Adam 优化器更优的数值稳定性。这种平稳收敛特性源于 L-BFGS 算法的二阶优化特性, 其通过构建近似的逆 Hessian 矩阵能够更准确地捕捉损失函数的曲率信息。然而, 该方法对初始参数值具有较高敏感性, 在非理想初始化条件下易收敛至局部最优解; 此外, 随着网络参数规模的增加, 存储和更新 Hessian 矩阵近似所需的计算资源显著增长, 使其在处理大规模问题时受限。

在使用 Adam+L-BFGS 混合优化测试时, 通过分阶段优化机制有效整合了两种算法的优势。在全局探索阶段(0~7 000 步), 采用 Adam 优化器充分发挥其鲁棒的参数空间搜索能力, 为网络参数提供高质量的初始估计; 随后转入局部精修阶段(7 000~10 000 步), 切换至 L-BFGS 优化器(初始学习率 $\eta = 1$, 最大函数评估 500 次/步, 梯度容差 1×10^{-5})利用其二阶收敛特性进行精细调优。该混合策略兼具 Adam 的全局收敛性和 L-BFGS 的局部精确性, 最终实现三者中最优的总损失值(4.42×10^{-4}), 较纯 Adam 和纯 L-BFGS 分别提升 51.52%和 15.33%。

4.2 采样方案影响

为构建高精度的 PINNs 热-力耦合计算模型, 从采样策略和采样密度 2 个方面开展研究。在采样策略方面, 通过对比拉丁超立方采样、均匀采样和随机采样 3 种方法, 评估不同空间采样策略对模型性能的调控作用; 在采样数量方面, 探究了内部配点和精化区域配点数量差异对预测模型性能的影响规律。

4.2.1 采样策略影响

图 11 为 3 种采样策略在计算域内的点集空间分布特征的对比。均匀采样采用规则网格划分, 在计算域内生成等间距分布的采样点; 随机采样基于伪随机数生成器, 实现采样点完全随机的空间分布;

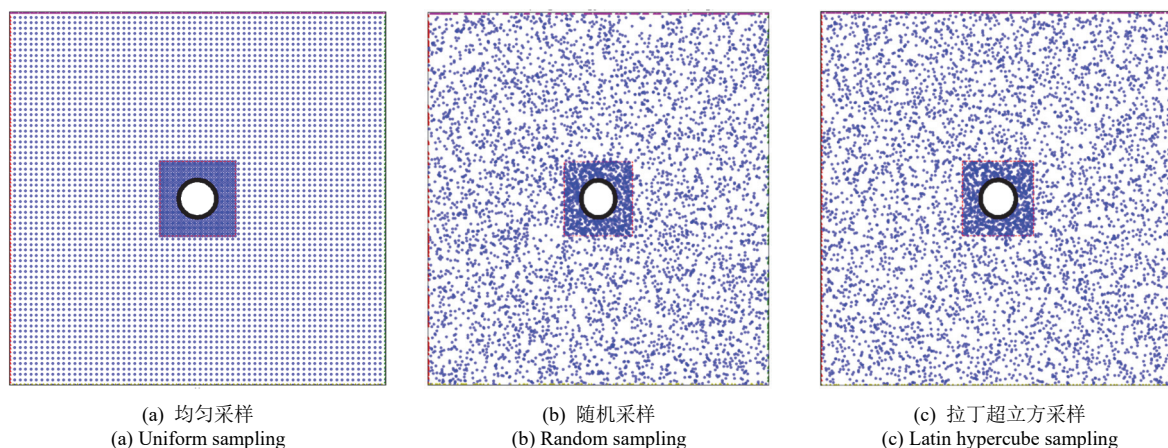
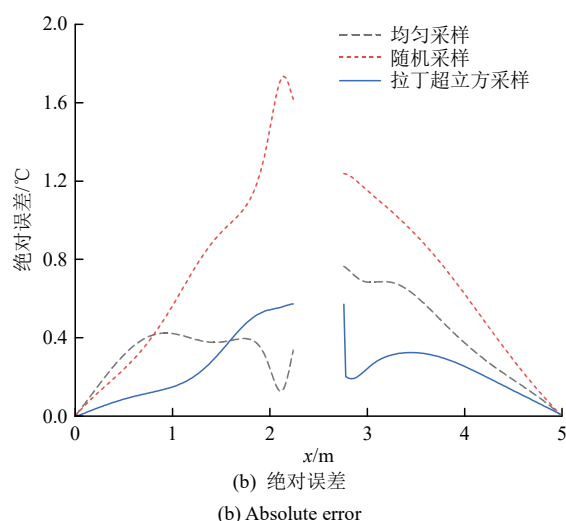
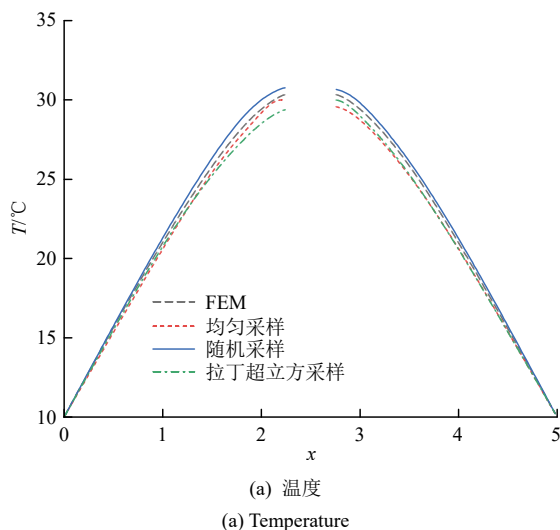


图 11 3 种不同采样策略下的采样点分布图

Fig.11 Sampling point distributions for three sampling strategies

而拉丁超立方采样^[22]作为随机采样改进方法,通过多维空间的分层抽样策略,在保持采样点随机性的同时,确保各维度投影的均匀分布,从而提升高维空间中的采样效率。

图 12 为 $y = 2.5 \text{ m}$ 截面处温度分布的 PINNs 预测结果与 FEM 数值解的对比。从图中可以发现:3 种采样策略(拉丁超立方采样、均匀采样和随机采样)的 PINNs 预测结果均与 FEM 的数值解表现出高度一致性,其中较大误差均集中在桩-土界面温度梯度变化较大的区域。误差定量分析表明:拉丁超立方采样策略表现出最优的预测精度,其最大绝对误差控制在 0.63°C 以内。相比之下,均匀采样策略和随机采样策略所产生的最大绝对误差分别为 0.78°C 和 1.68°C 。这一对比结果验证了拉丁超立方采样策略在关键区域(桩-土界面)的采样优势,其分层抽样特性通过优化采样点空间分布,能有效降低数值解的相位偏差和幅值误差,为热-力耦合问题的高精度求解提供了可靠的采样方案。

图 12 $y = 2.5 \text{ m}$ 处 3 种采样策略下温度的 PINNs 预测结果与 FEM 数值解对比Fig.12 Comparison of PINNs predicted results and FEM solutions for temperature at $y = 2.5 \text{ m}$ under three sampling strategies

针对 3 种采样策略下 PINNs 预测结果(温度和位移)与 FEM 数值解的误差进行了对比分析,见表 3 和 4。在温度场预测方面,拉丁超立方采样的全局平均相对误差 $MRE = 0.51\%$,较均匀采样和随机采样分别降低 77.03%和 76.81%;同时,其位移场预测表现出相似的趋势,拉丁超立方采样的 $MRE = 3.02\%$,较均匀采样和随机采样分别降低 68.31%和 66.14%。这主要源于拉丁超立方采样独特的空间分层特性:通过提升 40%的空间覆盖效率,在高梯度区域实现 2~3 倍的点密度增强,从而有效抑制数值振荡^[23]。而随机采样在桩土界面等关键区域的局部误差可达拉丁超立方采样的 2~6 倍;均匀采样虽较随机采样稳定,但仍存在明显的偏差。

表 3 采样策略对 PINNs 模型温度预测性能的影响

Table 3 Impact of sampling strategies on PINNs performance for temperature prediction

采样策略 (sampling strategies)	区域内温度误差 (temperature error within the target domain)				y = 2.5 m 处温度误差 (temperature error at y = 2.5 m)			
	MSE/℃ ²	RMSE/℃	MAE/℃	MRE/%	MSE/℃ ²	RMSE/℃	MAE/℃	MRE/%
均匀采样	0.256 7	0.506 7	0.393 9	2.22	0.556 9	0.746 3	0.635 7	3.27
随机采样	0.254 3	0.504 3	0.390 4	2.20	0.545 5	0.738 6	0.424 7	3.13
拉丁立方采样	0.017 0	0.130 3	0.088 9	0.51	0.035 4	0.188 1	0.127 6	0.62

表 4 采样策略对 PINNs 模型位移预测性能的影响

Table 4 Impact of sampling strategies on PINNs performance for displacement prediction

采样策略 (sampling strategies)	区域内位移误差 (displacement error within the target domain)				y = 2.5 m 处位移误差 (displacement error at y = 2.5 m)			
	MSE/m ²	RMSE/m	MAE/m	MRE/%	MSE/m ²	RMSE/m	MAE/m	MRE/%
均匀采样	2.316 9×10 ⁻¹⁰	1.522 1×10 ⁻⁵	1.403 3×10 ⁻⁵	9.53	2.245 1×10 ⁻¹⁰	1.498 4×10 ⁻⁵	1.375 6×10 ⁻⁵	8.63
随机采样	4.536 4×10 ⁻¹⁰	2.129 9×10 ⁻⁵	1.383 2×10 ⁻⁵	8.92	2.161 8×10 ⁻¹⁰	1.470 3×10 ⁻⁵	1.378 2×10 ⁻⁵	8.38
拉丁立方采样	2.454 7×10 ⁻¹¹	4.954 5×10 ⁻⁶	4.075 4×10 ⁻⁵	3.02	4.400 0×10 ⁻¹⁰	6.633 2×10 ⁻⁵	5.729 8×10 ⁻⁵	4.02

4.2.2 采样点数量影响

为研究采样密度对温度和位移预测精度的调控机制，基于最优网络参数和 Adam+L-BFGS 混合优化策略，提出了“基础+加密”的分区采样方案，并探究了采样点数量对预测模型性能的影响规律。分区采样方案设置为(2 000, 500), (4 000, 1 000), (6 000, 1 500), (8 000, 2 000), (10 000, 2 500)。以(2 000, 500)方案为例，2 000 个基础配点采用拉丁超立方采样策略分布于整个计算域；同时 500 个加密配点布置在高梯度区域(桩 - 土界面附近)。

表 5 和 6 给出了 5 种采样方案 PINNs 的温度和位移预测结果误差对比。在温度场预测方面，当采样点数量从(2 000, 500)增至(10 000, 2 500)的过程中，均方误差 MSE 由 11.932 1 ℃² 降至 1.937 7 ℃²，降幅达 83.76%。值得注意的是，误差下降过程表现出明显的阶段性特征：采样量从(2 000, 500)增至(4 000, 1 000)时，误差降幅最大，达到 67.34%。而

表 5 采样点数量对 PINNs 模型温度预测性能的影响

Table 5 Impact of sampling point quantity on PINNs performance for temperature prediction

采样数量 (sampling point quantity)		区域内温度误差 (temperature error within the target domain)			
计算域内 (within the computational domain)	精化区域 (within the refined region)	MSE/℃ ²	RMSE/℃	MAE/℃	MRE/%
2 000	500	11.932 1	3.454 3	2.707 1	39.39
4 000	1 000	3.896 0	1.973 8	1.620 4	9.07
6 000	1 500	2.309 4	1.519 7	1.262 1	2.05
8 000	2 000	1.993 5	1.411 9	1.130 1	1.94
10 000	2 500	1.937 7	1.392 0	1.111 8	1.89

当采样量继续增加时，误差改善幅度逐步趋缓，这反映了 PINNs 模型对采样密度存在收益递减效应。位移场预测虽然呈现类似的演化趋势，但其平均相对误差 MRE 始终高于温度场，这一差异本质上源于位移场受温度梯度和材料参数的双重影响。

表 6 采样点数量对 PINNs 模型位移预测性能的影响

Table 6 Impact of sampling point quantity on PINNs performance for displacement prediction

采样数量 (sampling point quantity)		区域内位移误差 (displacement error within the target domain)			
计算域内 (within the computational domain)	精化区域 (within the refined region)	MSE/m ²	RMSE/m	MAE/m	MRE/%
2 000	500	7.483 9×10 ⁻⁹	8.651 0×10 ⁻⁴	7.145 5×10 ⁻⁵	16.76
4 000	1 000	1.650 8×10 ⁻⁹	4.063 2×10 ⁻⁴	3.225 1×10 ⁻⁵	10.04
6 000	1 500	1.553 2×10 ⁻⁹	3.941 0×10 ⁻⁴	3.676 2×10 ⁻⁵	7.40
8 000	2 000	1.319 2×10 ⁻⁹	3.632 1×10 ⁻⁴	3.262 4×10 ⁻⁵	6.54
10 000	2 500	1.260 5×10 ⁻⁹	3.550 3×10 ⁻⁴	3.216 1×10 ⁻⁵	6.33

4.3 迁移学习能力

迁移学习是机器学习中一种高效的知识迁移技术，其核心是将已训练模型学到的特征或规律迁移到新问题中。在物理建模领域，迁移学习特别适用于参数化问题的快速求解，因其能够保留控制方程的普适性特征，仅需局部调整即可适应新参数或边界条件。

针对传统数值方法在几何或边界条件变化时需重复建模，导致计算效率低的问题，基于 PINNs，引入迁移学习技术，构建了高效求解框架。通过“预训练 - 微调”两阶段策略实现知识的迁移与复用：首先在基准工况($T_0=15\text{ }^{\circ}\text{C}$ ， $T_1=30\text{ }^{\circ}\text{C}$ ， $E=15\text{ MPa}$ ， $\nu=0.35$)下训练 10 000 步，使模型掌握热 - 力耦合问题的基本演化规律(如温度扩散特性、位移与应力的耦合关系等)。当处理新工况时，直接加载预训练模型的网络权重作为初始值，通过参数映射机制动态调整控制方程中的物理参数，在保留神经网络已学习的底层物理特征的同时，实现快速收敛。该方法既能保持模型预测精度，还可以显著提升计算效率。

基于上述所确立的最优模型配置，针对能量桩工程实际需求，设计以下工况开展数值验证：工况一： $T_0=15\text{ }^{\circ}\text{C}$ ， $T_1=30\text{ }^{\circ}\text{C}$ ， $E=15\text{ MPa}$ ， $\nu=0.35$ ；工况二： $T_0=10\text{ }^{\circ}\text{C}$ ， $T_1=40\text{ }^{\circ}\text{C}$ ， $E=15\text{ MPa}$ ， $\nu=0.35$ ；工况三： $T_0=10\text{ }^{\circ}\text{C}$ ， $T_1=30\text{ }^{\circ}\text{C}$ ， $E=25\text{ MPa}$ ， $\nu=0.35$ ；工况四： $T_0=10\text{ }^{\circ}\text{C}$ ， $T_1=30\text{ }^{\circ}\text{C}$ ， $E=15\text{ MPa}$ ， $\nu=0.30$ 。该工况设计分别考察了环境温度(工况一)、热荷载强度(工况二)、土体弹性参数(工况三和四)三类关键因素对迁移学习性能的影响。所有工况均保持其他参数与基准工况一致，确保迁移学习能力分析的可靠性。

表 7 和 8 分别为有/无迁移学习的情况下 PINNs 模型温度与位移预测的误差对比。在温度场预测中，误差演化呈先快速下降后稳定收敛的趋势：在初始 500 次迭代阶段，受参数重适应过程影响， MRE 维持在 2.35%~5.53% 的较高水平；随着迭代次数增至 1 000 次， MRE 降至 0.10%~0.33%，降幅达 88.51%~96.85%；继续增加至 1 500 次时， MRE 趋于平缓，其最大降低幅度仅为 21.21%。在位移场预测中，500 次迭代后 MRE 处于 9.23%~12.01% 范围；当迭代增至 1 000 次时， MRE 降至 1.11%~2.01%，最大降幅为 87.97%；而 1 500 次迭代后，最大 MRE 降幅仅为 19.76%，说明预测模型已进入稳定状态。由此可

表 7 有/无迁移学习的温度预测误差对比

Table 7 Temperature prediction error with and without transfer learning

工况编号 (case No.)	温度误差 MRE (temperature prediction error MRE)/%			
	无迁移学习 (without transfer learning)	迁移学习 (with transfer learning)		
		500 次	1 000 次	1 500 次
一	0.26	5.53	0.33	0.26
二	0.27	4.77	0.28	0.25
三	0.31	3.17	0.10	0.13
四	0.46	2.35	0.27	0.23

表 8 有/无迁移学习的位移预测误差对比

Table 8 Displacement prediction error with and without transfer learning

工况编号 (case No.)	位移误差 MRE (displacement prediction error MRE)/%			
	无迁移学习 (without transfer learning)	迁移学习 (with transfer learning)		
		500 次	1 000 次	1 500 次
一	8.90	12.01	2.01	1.83
二	1.30	9.23	1.11	1.45
三	2.16	11.33	1.57	1.26
四	3.10	9.71	1.33	1.35

见，1 000 次迁移学习迭代在温度与位移的预测中，能够达到 10 000 次无迁移学习的计算精度，该优化方案可在保证预测精度的同时大幅降低计算成本。这与 E. Haghighat 等^[24]在固体力学领域的相关研究结论相吻合，证明了 PINNs 通过迁移学习实现的“一次训练，多次快速复用”范式，在解决多工况问题时具有巨大的潜力。

4.4 计算环境与成本分析

为评估 PINNs 方法的计算效率，对 PINNs 模型的训练与测试耗时进行详细分析，并与传统有限元法进行对比。基于表 9 所列的计算环境，在最优网络配置下(4 层隐藏层，每层 50 个神经元，Tanh 激活函数)，完成一次 10 000 次迭代的训练过程至收敛，平均耗时为 497.84 s。作为对比，在相同硬件上，使用 COMSOL Multiphysics 有限元软件对同一工况进行求解，在包含 6 942 个域单元和 228 个边单元的网格下，单次计算耗时约为 46.53 s。这一对比清晰地揭示了 PINNs 方法的初始训练成本远高于传统 FEM 的单次求解成本。

表 9 计算环境与耗时对比

Table 9 Comparison of computational environment and runtime

类别 (category)	项目 (item)	详情 (details)
硬件环境	GPU	NVIDIA GeForce RTX 3060
硬件环境	CPU	AMD Ryzen 5 5600G
软件环境	语言与框架	Python 3.9, PyTorch 1.12.0
PINNs 配置	网络结构	4 层隐藏层, 每层 50 个神经元
PINNs 配置	激活函数	Tanh
计算耗时	PINNs 预训练	497.84 s
计算耗时	PINNs 边界微调	46.53 s
计算耗时	FEM 单次求解	55.76 s

然而, PINNs 的策略是以较高的初始训练投入, 换取一个参数化的代理模型。该模型具备瞬时预测与低成本微调(迁移学习)的能力。这使得在工况分析、参数化设计等需要多次求解的场景中, PINNs 的每次额外分析的成本极低, 以节 4.3 中工况一为例, 经 1 000 次迁移学习迭代, 在计算结果精度相近的前提下($MRE < 3\%$), 其计算效率相比 FEM 方法提升 16.55%。

5 结论与展望

5.1 结 论

基于 PINNs 将控制方程(热扩散方程和 Navier 应力平衡方程)的残差项作为正则化约束嵌入损失函数, 构建了一种能够严格满足物理守恒定律的能量桩周土体热 - 力耦合作用的高效求解方法。通过与有限元数值解对比, 探究了神经网络架构、激活函数选择、优化策略以及采样方案等因素对模型预测精度的影响, 并通过“预训练 - 微调”迁移学习策略评估了其计算效率, 主要结论如下:

(1) 经与有限元结果对比表明, 所提出的 PINNs 热 - 力耦合模型在采用 4 层隐藏层、50 个神经元以及 Tanh 激活函数时表现最优, 其温度场和位移场预测结果的最大相对误差仅分别为 0.53% 和 5.72%。

(2) Adam+L-BFGS 混合优化策略能够显著提升 PINNs 模型的收敛性能, 最终总损失值仅为 4.42×10^{-4} , 较单一采用 Adam 和纯 L-BFGS 策略分别提升 51.52% 和 15.33%。

(3) 拉丁超立方采样凭借其空间分层特性有效提升了模型预测精度。在温度预测中, 其 MRE 为

0.51%, 较均匀采样和随机采样分别降低 77.03% 和 76.81%; 在位移预测时, 拉丁超立方采样的 MRE 为 3.02%, 较二者分别降低 68.31% 和 66.14%。

(4) 基于“预训练 - 微调”的迁移学习策略能够显著提升能量桩周土体热 - 力耦合的计算效率。在新工况下, 仅需 1 000 次迁移学习迭代, 即可达到无迁移时 10 000 次迭代的预测精度, 计算成本仅为完整训练的 9.35%, 计算效率相较于 FEM 方法提升 16.55%。基于 PINNs 的迁移学习策略为求解多工况岩土工程问题提供了高效途径。

5.2 展 望

尽管当前研究取得了一些积极成果, 但仍存在若干局限性, 值得在未来工作中进一步深入探索:

(1) 在计算模型中将能量桩简化为刚性体, 且桩 - 土界面假设为理想连接。然而, 实际工程中桩体变形与界面的非线性接触行为可能会对计算结果产生重要影响。后续研究可考虑采用双网络架构分别表征桩与土的力学响应, 并引入更复杂的界面本构模型, 以更真实地模拟二者之间的相互作用。

(2) 当前研究聚焦于单一能量桩的二维简化模型。将其扩展至三维分析框架, 并进一步研究大规模群桩系统, 同时考虑土体的非均质性与材料非线性, 是推进其走向工程实际应用的必然路径。这一扩展也将对 PINNs 的训练效率与数值稳定性提出更高要求。

(3) 损失函数中权重系数 λ_1 和 λ_2 的合理设定对 PINNs 的训练效果具有关键影响。目前常采用试算法确定其取值, 未来可探索自适应权重调整策略, 以增强模型的鲁棒性, 并提高其自动化程度。

(4) 当前研究的验证工作依赖于有限元数值解。后续研究应结合现场实测数据或物理模型试验结果进行系统对比, 从而更全面地评估模型在实际工程场景中的预测精度与适用性。

参考文献(References):

[1] GIELEN D, BOSHELL F, SAYGIN D, et al. The role of renewable energy in the global energy transformation[J]. Energy Strategy Reviews, 2019, 24: 38 - 50.

[2] LAI X, LIU J, SHI Q, et al. Driving forces for low carbon technology innovation in the building industry: A critical review[J]. Renewable and Sustainable Energy Reviews, 2017, 74: 299 - 315.

[3] OMER A M. Ground-source heat pumps systems and applications[J].

- Renewable and Sustainable Energy Reviews, 2008, 12(2): 344 - 371.
- [4] SANI A K, SINGH R M, AMIS T, et al. A review on the performance of geothermal energy pile foundation, its design process and applications[J]. Renewable and Sustainable Energy Reviews, 2019, 106: 54 - 78.
- [5] MORADSHAHI A, FAIZAL M, BOUAZZA A, et al. Effect of nearby piles and soil properties on thermal behaviour of a field-scale energy pile[J]. Canadian Geotechnical Journal, 2021, 58(9): 1 351 - 1 364.
- [6] YIN Y, LI Q, QIAO L. Response of energy pile-soil structure and pile group effect: An indoor similarity simulation study[J]. Journal of Building Engineering, 2022, 51: 104247.
- [7] CUI S, ZHOU C, LIU J, et al. Stress effects on thermal conductivity of soils and heat transfer efficiency of energy piles in the saturated and unsaturated soils[J]. Computers and Geotechnics, 2023, 160: 105549.
- [8] SALCIARINI D, RONCHI F, CATTONI E, et al. Thermomechanical effects induced by energy piles operation in a small piled raft[J]. International Journal of Geomechanics, 2015, 15(2): 04014042.
- [9] RAZMKHAH T, JAFARZADEH F, GHASEMI-FARE O. Numerical investigation of the thermo-mechanical response of single energy pile[C]// International Foundations Congress and Equipment Expo. Orlando, USA: American Society of Civil Engineers, 2018: 664 - 674.
- [10] ZIENKIEWICZ O C, TAYLOR R L, ZHU J Z. The finite element method: its basis and fundamentals(Sixth edition)[M]. Oxford: Butterworth-Heinemann, 2005: 264 - 302.
- [11] BORST R D, CRISFIELD M A, REMMERS J J, et al. Non-linear finite element analysis of solids and structures[M]. 2nd ed. New York: John Wiley and Sons, 2012: 1 - 29.
- [12] KARNIADAKIS G E, KEVREKIDIS I G, LU L, et al. Physics-informed machine learning[J]. Nature Reviews Physics, 2021, 3(6): 422 - 440.
- [13] SHARMA R, GUO Y B. Physics-informed machine learning of thermal stress evolution in laser metal deposition[C]// Proceedings of the 154th Annual Meeting and Exhibition Supplemental. Las Vegas, USA: Springer, 2025: 550 - 559.
- [14] YOU M, LIU T, HUAI Y. Determining thermomechanical response of thin disk lasers with physics-informed neural networks[J]. Optics and Laser Technology, 2023, 165: 109590.
- [15] HUANG L, DING Z, SONG C. A novel three-dimensional analytical model for the thermomechanical responses of energy piles and the surrounding soil based on thermoelastic theory[J]. International Journal for Numerical and Analytical Methods in Geomechanics, 2025, 49(12): 2 749 - 2 762.
- [16] PASZKE A, GROSS S, CHINTALAS, et al. Automatic differentiation in PyTorch[C]// 31st Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates, Inc., 2017: 1 - 4.
- [17] MCCLENNY L D, BRAGA-NETO U M. Self-adaptive physics-informed neural networks[J]. Journal of Computational Physics, 2023, 474: 111722.
- [18] MORITZ P, NISHIHARA R, JORDAN M I. A linearly-convergent stochastic L-BFGS algorithm[C]// Proceedings of the 19th International Conference on Artificial Intelligence and Statistics. Cadiz, Spain: PMLR, 2016: 249 - 258.
- [19] LI K Q, YIN Z Y, ZHANG N, et al. A data-driven method to model stress-strain behaviour of frozen soil considering uncertainty[J]. Cold Regions Science and Technology, 2023, 213: 103906.
- [20] JAGTAP A D, KAWAGUCHI K, KARNIADAKIS G E. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks[J]. Journal of Computational Physics, 2020, 404: 109136.
- [21] LI K Q, YIN Z Y, ZHANG N, et al. Physics-informed neural networks for solving steady-state temperature field in artificial ground freezing[J]. Canadian Geotechnical Journal, 2025, 62: 1 - 17.
- [22] MCKAY M D, BECKMAN R J, CONOVER W J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code[J]. Technometrics, 2000, 42(1): 55 - 61.
- [23] SHIELDS M D, ZHANG J. The generalization of Latin hypercube sampling[J]. Reliability Engineering and System Safety, 2016, 148: 96 - 108.
- [24] HAGHIGHAT E, RAISSI M, MOURE A, et al. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics[J]. Computer Methods in Applied Mechanics and Engineering, 2021, 379: 113741.