用于非饱和土固结分析的滞后向后兼容物理信息神经网络

Dong Li[1], Shuai Huang[2], Yapeng Cao[3*], Yujun Cui[4], Xiaobin Wei[5], Hongtao Cao[6]

摘要： 本研究开发了一种滞后反向兼容物理信息神经网络（LBC-PINN），用于模拟和反演长期荷载作用下的一维非饱和土固结。为应对多尺度时间域内气水压力耦合消散所带来的挑战，该框架融合了对数时间分段、滞后兼容性损失约束和分段迁移学习。在正向分析中，采用推荐分段方案的 LBC-PINN 能准确预测孔隙气压和孔隙水压力的演变；与有限元法（FEM）结果对比验证表明，直至 10^{10} 秒的整个时间范围内，平均绝对误差均低于 10^{-2} 。基于特征气相消散时间的简化分段策略在保持预测精度的同时提高了计算效率。敏感性分析证实，该框架在气水渗透率比 k_a/k_w 从 10^{-3} 到 10^{3} 的范围内具有稳健性。

关键词： 非饱和土；物理信息神经网络（PINN）；一维固结；分段时间训练；滞后向后相容

[1]博士，乔治·梅森大学土木、环境与基础设施工程系，美国弗吉尼亚州费尔法克斯 22030 [2] 博士，应急管理部国家自然灾害防治研究院，中国北京 100085 [3*] 博士，中国科学院西北生态环境资源研究院冰冻圈科学与冻土工程国家重点实验室，中国兰州 730000；法国国立路桥学校纳维叶实验室，法国马恩河谷 Cedex 2，77455（通讯作者）邮箱：yapeng.cao@enpc.fr [4] 教授，法国国立路桥学校纳维叶实验室，法国马恩河谷 Cedex 2，77455 [5] 博士，河北工程大学土木工程学院，中国邯郸 056038 [6] 博士，浙江工业大学土木工程学院，中国杭州 310023

符号说明

符号 描述 单位 C_a 与气相相关的交互常数 – C_w 与水相相关的交互常数 – c_v^a 气相固结系数 m2/s c_v^w 水相固结系数 m2/s H 土层厚度 m k_a 空气渗透系数 m/s k_w 水渗透系数 m/s k_a/k_w 气水渗透率比 – $\mathrm{L_{BC}}$ 边界条件损失项 – $\mathrm{L_{IC}}$ 初始条件损失项 – L_R 偏微分方程残差（内部）损失项 – L_S 滞后兼容损失项 – $\mathrm{L_{STD}}$ 标准物理信息神经网络的总损失 – m_1^a 空气体积变化相对于净法向应力变化的系数 /kPa m_2^a 空气体积变化相对于净基质吸力变化的系数 /kPa m_1^w 水体积变化相对于净法向应力变化的系数 /kPa m_2^w 水体积变化相对于净基质吸力变化的系数 /kPa MAE 平均绝对误差 kPaMRE 平均相对误差 –n 土壤孔隙率 – R 通用气体常数 J/mol/KR^2 决定系数 –S_r 饱和度 % T 绝对温度 K $T_{\max}$ 最大模拟／训练时间范围 s t 时间 st_s 特征气相消散时间 s $u_a(z,t)$ 超孔隙气压 kPa $u_w(z,t)$ 超孔隙水压力 kPa u_a^0 初始孔隙气压 kPa u_w^0 初始孔隙水压力 kPa z 垂直坐标（深度，自地表起算） m γ_w 水的重度 kN/m3 ρ_a 空气密度 kg/m3 ρ_w 水的密度 kg/m3 σ 总竖向应力 kPaσ' 有效竖向应力 kPa

θ 可训练神经网络参数向量（权重和偏置）– ω_{BC} 边界条件损失项的损失权重参数 – ω_{IC} 初始条件损失项的损失权重参数 – ω_R 偏微分方程残差损失项的损失权重参数 – ω_s 滞后相容性损失项的损失权重参数 –

1. 引言

岩土工程中的固结是指土体在外部荷载作用下超孔隙水压力消散，从而产生时变沉降（Biot, 1941; Esrig, 1968; Mesri and Rokhsar, 1974; Cargill, 1984; Conte, 2004; Huang et al., 2010）。传统上，饱和土的时变沉降行为采用太沙基一维固结理论进行建模，该理论基于土体渗透性和压缩性恒定等理想化假设，将固结描述为持续荷载引起的超孔隙水压力消散（Terzaghi, 1943; Mesri and Rokhsar, 1974）。在许多实际情况下，例如压实填土、地下水位以上的天然沉积土，或受到波动地下水位影响的结构，土体在持续荷载和环境波动作用下仍保持部分非饱和状态（Barden, 1965; Lloret and Alonso, 1980; Ho 等人, 2014; Ho 和 Fatahi, 2016）。当气相存在并与水相相互作用时，固结的控制机制不同于饱和土中的情况（Conte, 2004）。关于非饱和土固结的早期研究包括 Scott（1963）对含封闭气泡土体的估算、Biot（1941）适用于此类条件的一般固结理论，以及 Barden（1965）对压实非饱和黏土进行的一维固结分析。

假设气相和水相均为连续相，Fredlund 和 Hasan（1979年）进一步发展了非饱和土一维固结理论，并基于质量守恒和达西定律建立了水相和气相的耦合控制方程。传统上，这些行为通过常微分方程和偏微分方程（ODE 和 PDE）进行描述和分析，它们构成了非饱和土固结分析中解析解和数值解的基础（Wang 等人，2020年；

Zhou 等人，2017年；Briaud，2023年）。针对单向（单面排水）和双向（双面排水）条件下的非饱和土固结，已有多种考虑气水耦合的解析解，包括 Qin 等人（2008年，2010年）、Zhou 等人（2014年）、Zhou 等人（2017年）、Shan 等人（2012年）、Ho 等人（2014年）以及 Ho 和 Fatahi（2016年，2018年）提出的解。Qin 等人（2008年，2010年）基于 Fredlund 和 Hasan（1979年）的控制方程，开发了一种基于拉普拉斯变换的方法，用于预测非饱和土的一维沉降。Shan 等人（2012年）通过变量分离法获得了非饱和土一维固结的精确解。Ho 等人（2014年）采用特征函数展开和拉普拉斯变换，推导出非饱和土层一维固结控制方程的精确解析解。然而，推导闭式解通常需要运用拉普拉斯变换及其逆变换、凯莱–哈密顿定理等高级数学方法，还需费力求解相关特征值问题。这些代数推导十分繁琐，是主要的计算瓶颈。因此，数值技术成为近似求解偏微分方程的首选方法，例如有限元（FE）方法（Wong 等人，1998年；Cheng 等人，2017年；Tang 等人，2018年；Liu 等人，2014年）和微分求积（DQ）法（Zhou，2013年；Zhou 和 Zhao，2014年）。然而，传统数值方法存在若干局限，包括大变形条件下易受网格畸变影响、强耦合分析中难以收敛、反演建模计算需求高，以及难以表示气相和水相的复杂边界条件（Reddy，1993年）。

科学机器学习的最新进展引入了物理信息神经网络（PINN），这是一种通过将控制方程、边界条件和初始条件嵌入神经网络的损失函数中，以无网格方式求解偏微分方程的方法（Cuomo 等人，2022年；Lawal 等人，2022年）。Raissi 等人（2019年）率先提出了这一方法，Bekele（2021年）则将其应用于太沙基固结。Bekele（2021年）证明，通过最小化合并物理损失和数据损失，PINN 能够准确复现一维固结解（正问题），并根据实测数据估计固结系数（反问题）。Wang 等人（2024年）进一步开发了一种

新型 PINN，用于利用有限的现场特定测量数据对二维土体固结进行正反演分析，并特别关注样本外预测性能。Lan 等（2024年）通过引入硬约束（PINNs-H），增强了物理信息神经网络（PINN），以模拟连续排水边界下非线性固结过程中超孔隙水压力的变化。与有限元法（FEM）相比，PINNs-H 使用更少的训练数据便达到了相当的准确度，并免去了时空离散的需要。Zhang 等（2022年）开发了一种物理信息数据驱动方法，可根据实测数据自动还原太沙基固结理论，并通过引入弱形式偏微分方程获得真实解，以增强抗噪能力；同时采用蒙特卡洛 dropout 进行不确定性估计。Zhou 等（2025年）提出了一种自适应 PINN（SA-PINN）方法，用于非线性大应变固结的正问题和反问题，并考虑了软黏土的蠕变特性。然而，尽管 PINN 在饱和固结问题中日益取得成功，却尚未应用于非饱和土固结。首先，非饱和土固结的控制方程涉及耦合偏微分方程，描述超孔隙气压力与超孔隙水压力的同步消散，且二者分别由不同的相互作用常数和固结系数控制。这种耦合会导致系统刚性，进而造成梯度传播不平衡和训练困难。其次，非饱和土固结的时间域通常跨越多个数量级（秒至数月或数年），从快速的气相消散到较慢的水相排水；由于传统 PINN 已知存在偏向拟合长期动力学的倾向，因此面临困难。因此，若不采用特殊策略，在如此长的时间尺度上进行训练往往会导致收敛不佳、后期预测分辨率不足以及损失传播失衡。本研究开发了一种滞后反向兼容物理信息神经网络（LBC-PINN），用于一维非饱和土固结的正反演分析。该框架结合了时空域归一化、滞后兼容损失和迁移学习，以提高长时间范围内的稳定性和准确度。正向模型预测长期荷载作用下耦合的孔隙气压和孔隙水压力场，并通过有限元模拟进行验证。此外，

本文对气水渗透率比（k_a/k_w）、时间分段方案、神经网络结构和数据噪声进行了详细的敏感性分析，证实了 LBC-PINN 方法的稳健性和泛化能力。

2. 问题描述

2.1 假设与控制方程

本研究采用 Fredlund 和 Hasan（1979年）提出的非饱和土一维固结模型，其示意图如图1所示。厚度为H的均质各向同性非饱和土层承受竖向表面荷载q，并沿水平方向无限延伸。假设水流、气流和沉降仅发生在竖直方向，并在分析中忽略空气溶解。水和空气的渗透系数（k_w 和 k_a）以及水的密度在固结过程中均视为常数。在这些假设下，固结过程涉及施加荷载 q 作用下超孔隙水压力 u_w 和超孔隙气压u_a 的耦合消散。各向同性非饱和土中u_a 和u_w 所对应的一维耦合控制方程可写成 Fredlund-Hasan （ 1979 年）形式：

$$\frac{\partial u_a}{\partial t} + C_a \frac{\partial u_w}{\partial t} + c_v^a \frac{\partial^2 u_a}{\partial z^2} = 0 \tag{1}$$

$$\frac{\partial u_w}{\partial t} + C_w \frac{\partial u_a}{\partial t} + c_v^w \frac{\partial^2 u_w}{\partial z^2} = 0 \tag{2}$$

其中 $C_a = \frac{1}{\left[\left(\frac{m_1^a}{m_2^a}-1\right) - \left(\frac{1-S_r\ n}{m_2^a\ u_a^0+u_{atm}}\right)()\right]}$为与气相相关的交互常数； $c_v^a = \frac{\frac{k_a RT}{gM}}{\left[m_2^a(u_a^0+u_{atm})\left(\frac{m_1^a}{m_2^a}-1\right)-(1-S_r)n\right]}$为气相固结系数（$m^2/s$）；$C_w = \left|\frac{m_1^w}{m_2^w} - 1\right|$为与水相相关的交互常数；$c_v^w = \frac{k_w}{m_2^w \gamma_w}$为水相固结系数（$m^2/s$），$k_a$ 和k_w 分别为气体和水的渗透系数（m/s）；m_1^a 和m_2^a 分别为空气体积变化相对于净法向应力变化的系数（$\sigma_z - u_a$）和相对于基质吸力变化的系数（$u_a - u_w$）（/kPa）；m_1^w 和m_2^w 为系数

水体积变化相对于净法向应力（$\sigma_z - u_a$）变化和基质吸力（$u_a - u_w$）变化的系数（/kPa）；u_a^0 和u_w^0 分别为初始孔隙气压力和孔隙水压力（kPa），u_{atm} 为大气压力（kPa）；R为通用气体常数（~8.3 J/mo/K）；T 为绝对温度（K）；M 为固结过程中空气饱和的分子质量；n为固结过程中的孔隙率；S_r 为固结过程中的饱和度（%）；γ_w 为水的重度（~ 9.8 kN/m^3）。在本研究中，假定这些材料参数沿深度保持恒定，且与不断演化的饱和度无关，从而使控制方程保持线性，并可直接与现有解析解和数值解进行比较。

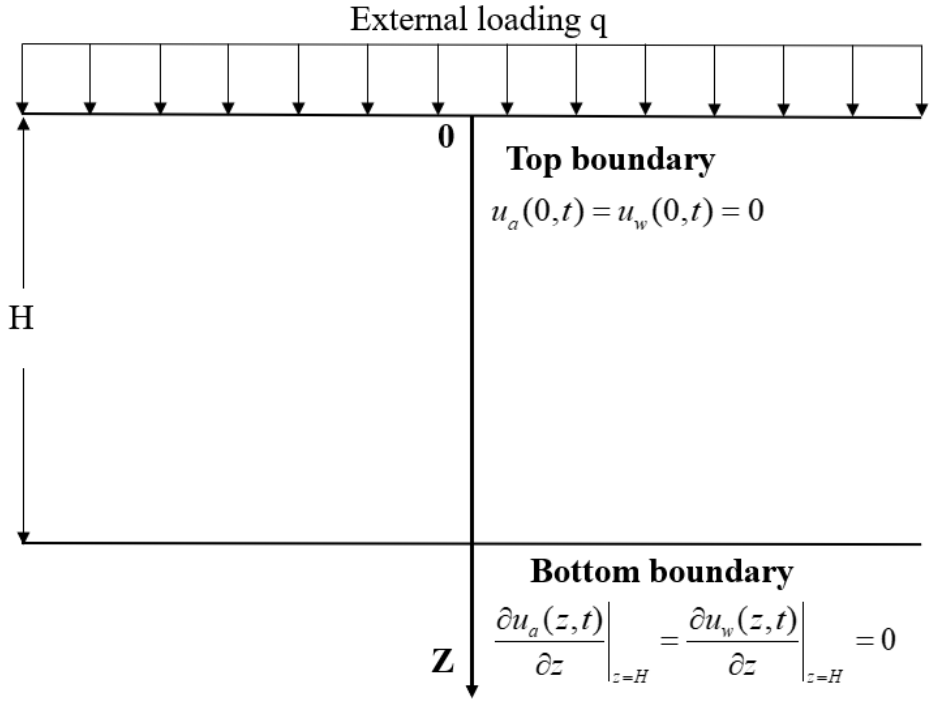


图1。简化　的单向排水非饱和土层一维固结模型　e

2.2 边界和初始条件

本研究评估单向排水条件，其特征是上边界可透水（排水），底部边界不透水（不排水），如图1所示。 气相和水相均采用单向（单面排水）条件。$z = 0$ 处的顶部边界可透水：

$$u_a(0,t) = u_w(0,t) = 0, \quad t > 0; \tag{3}$$

$z = H$ 处的底部边界不透水：

$$\frac{\partial u_a(z,t)}{\partial z}|_{z=H} = \frac{\partial u_w(z,t)}{\partial z}|_{z=H} = 0, \quad t > 0; \tag{4}$$

初始超孔隙气压和孔隙水压力取为：

$$u_a(z,0) = u_a^0, u_w(z,0) = u_w^0 \tag{5}$$

3. 非饱和土固结的物理信息神经网络

3.1 标准物理信息神经网络（STD-PINN）

物理信息神经网络（PINN）是一种将物理定律（通常为偏微分方程）嵌入其损失函数中的机器学习模型，即使没有标注数据，它也能学习满足控制方程的解（Raissi 等人，2019年；Tartakovsky 等，2020；He 和 Tartakovsky，2021；Lu 等，2021；Mao 等，2023）。标准物理信息神经网络（STD-PINN）通过训练神经网络，在任意时空坐标处计算近似解，并通过残差损失项惩罚控制方程、初始条件和边界条件的违背，从而为解场提供符合物理规律的无网格近似（Cai 等，2021；Toscano 等，2025）。图 2 展示了本研究采用的标准物理信息神经网络框架示意图。考虑一个定义在时空域 Ω × [0, T], 上的、包含 m 个偏微分方程的系统的一般形式，其中偏微分方程描述一个或多个因变量 u_i(z,t)（例如，u_a(z,t)、 u_w(z,t)）的演化，使得：

$$\mathcal{N}_i[u(z,t); \boldsymbol{\lambda}] = 0, z \in \Omega, t \in (0,T], i = 1,2,\dots,N \tag{6}$$

这里，$\mathcal{N}_i[\cdot]$ 表示微th 微分算子，其中可能包含空间和时间的混合偏导数； $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots]$ 是偏微分方程参数，其中 z 和 t分别为空间和时间坐标。偏微分方程系统对应的初始条件和边界条件可表示为：

$$u_i(z,0) = u_{i,0}(z), \qquad x \in \Omega \tag{7}$$

$$\mathcal{B}_i[u](z,t) = \varphi_i(z,t), \quad z \in \partial\Omega, \qquad t \in [0,T] \tag{8}$$

其中 $u_{i,\ 0}(z)$ 表示偏微分方程的初始条件，$\partial\Omega$ 表示偏微分方程的边界，且$\varphi_i(z,t)$ 表示相应的边界条件，可以是狄利克雷、诺伊曼或混合边界条件（Wang 等人，2023年）。

为近似解场 u_a 和 u_w，定义神经网络 $u\hat{}(\boldsymbol{z},t;\ \boldsymbol{\theta}) = u(\boldsymbol{z},t)$，其中 $\boldsymbol{\theta}$表示全连接前馈网络的可训练参数（权重和偏置）（Sharma et al., 2013; Xu et al., 2025）。该网络通常由两个输入层、多个带有非线性激活函数的隐藏层以及两个线性输出层组成。将第 kth 个隐藏层的输出记为$\boldsymbol{z}^{(k)}$，则网络架构可表示为：

$$\begin{cases} \boldsymbol{z}^0 = (z,t), \\ \boldsymbol{z}^m = \sigma(\boldsymbol{W}^m \boldsymbol{z}^{m-1} + \boldsymbol{b}^m), m = 1,2,\dots M \\ \boldsymbol{z}^{M+1} = \boldsymbol{W}^{M+1}\boldsymbol{z}^M + \boldsymbol{b}^{M+1} \end{cases} \tag{9}$$

其中，m（m=1,2,...M）表示第 mth 个隐藏层，$\boldsymbol{z}^m$ 是第 mth 个隐藏层中神经元的输出；W^m 和 b^m 表示第 mth 层的权重矩阵和偏置向量，$\mathbf{W}^{M+1}$ 和$\mathbf{b}^{M+1}$ 分别是输出层的权重矩阵和偏置向量；$\sigma(\cdot)$是非线性激活函数。本文将所有可训练模型参数，即权重和偏置，记为 $\boldsymbol{\theta}$ 。网络输出对时间和空间的所有所需偏导数（例如 $\partial_t u\hat{}_a$, $\partial_t u\hat{}_w$）均通过自动微分计算（Gunes et al., 2015）。在物理信息神经网络中，通过迭代更新 $\boldsymbol{\theta}$ ，将偏微分方程系统的求解转化为一个以最小化损失函数 L$_{std}$ 为目标的优化问题：

$$\mathcal{L}_{\text{std}}(\boldsymbol{\theta}) = \underbrace{\omega_{\text{IC}}\mathcal{L}_{\text{IC}}(z_k^i, 0) + \omega_{\text{BC}}\mathcal{L}_{\text{BC}}(z_k^b, t_k^b)}_{Physics\ Loss} + \underbrace{\omega_R\mathcal{L}_R(z_k^r, t_k^r)}_{Data\ Loss} \tag{10}$$

其中

$$\mathcal{L}_{\mathrm{IC}} = \frac{1}{N_{IC}} \sum_{k=1}^{N_{IC}} \left[\left(\hat{u}_a\left(z_k^i, 0\right) - u_{a,0}\left(z_k^i, 0\right) \right)^2 + \left(\hat{u}_w\left(z_k^i, 0\right) - u_{w,0}\left(z_k^i, 0\right) \right)^2 \right], \left(z_k^i, 0\right) \in \Omega \tag{11}$$

$$\mathcal{L}_{\mathrm{BC}} = \frac{1}{N_{BC}} \sum_{k=1}^{N_{BC}} \left[\left(\hat{u}_a\left(0, t_k^b\right) \right)^2 + \left(\hat{u}_w\left(0, t_k^b\right) \right)^2 + \left(\frac{\partial \hat{u}_a\left(H, t_k^b\right)}{\partial z} \right)^2 + \left(\frac{\partial \hat{u}_w\left(H, t_k^b\right)}{\partial z} \right)^2 \right], \quad \left(z_k^b, t_k^b\right) \in \partial\Omega \times [0, \mathrm{T}] \tag{12}$$

$$\mathcal{L}_R = \frac{1}{N_R} \sum_{k=1}^{N_R} \left[\left(\frac{\partial \hat{u}_a(z_k^r, t_k^r)}{\partial t} + C_a \frac{\partial \hat{u}_w(z_k^r, t_k^r)}{\partial t} + c_v^a \frac{\partial^2 \hat{u}_a(z_k^r, t_k^r)}{\partial z^2} \right)^2 + \left(\frac{\partial \hat{u}_w(z_k^r, t_k^r)}{\partial t} + C_w \frac{\partial \hat{u}_a(z_k^r, t_k^r)}{\partial t} + c_v^w \frac{\partial^2 \hat{u}_w(z_k^r, t_k^r)}{\partial z^2} \right)^2 \right], \left(z_k^r, t_k^r\right) \in \Omega \times [0, \mathrm{T}] \tag{13}$$

其中，L_{IC}、L_{BC}、L_R 分别为初始条件损失、边界条件损失和偏微分方程残差损失；ω_{IC}、ω_{BC}、ω_R 分别表示L_{IC}、L_{BC}、L_R的损失权重参数；N_{IC}、N_{BC}、N_R 分别为各损失项的采样点数量。物理损失（L_{IC}、L_{BC}）用于强制满足指定的初始条件和边界条件，确保符合问题约束。由偏微分方程残差定义的数据损失（L_R）促使网络在整个区域内满足潜在的物理定律（Raissi 等人，2019年）。

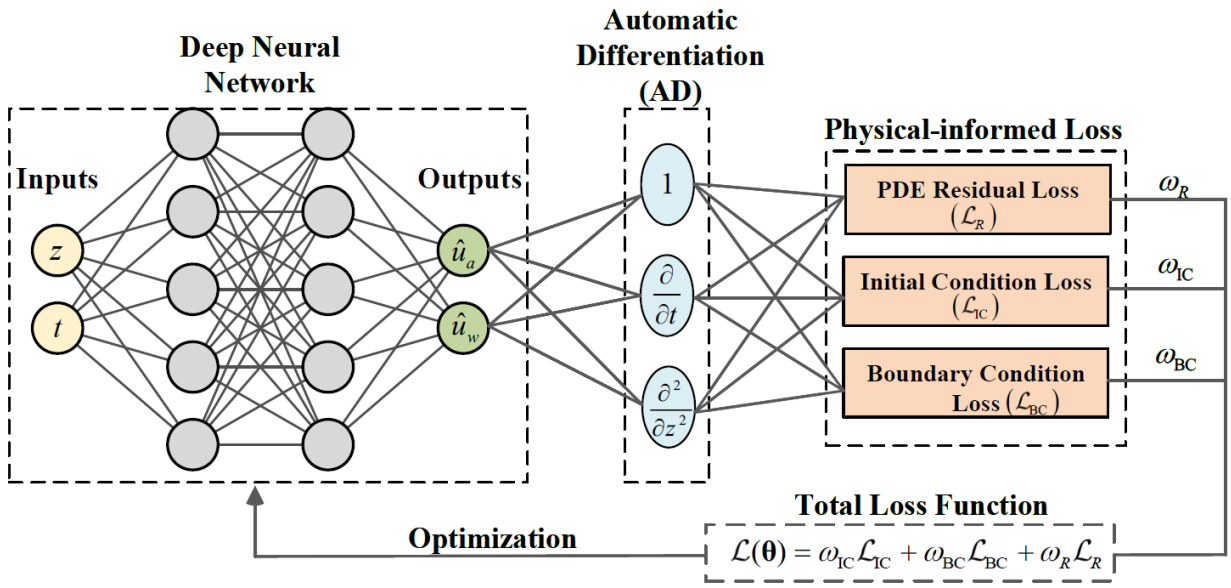


图 2. 标准物理信息神经网络（STD-PINN）示意图。

3.2 滞后向后兼容 物理信息神经网络（LBC-PINN）

物理信息神经网络（PINN）通过将物理定律直接嵌入神经网络的训练中，已成为求解偏微分方程（PDE）的强大范式。然而，在应用于涉及长时间尺度的问题时，例如非饱和土固结（其时间跨度可能跨越多个数量级，例如 10^{10} s），标准物理信息神经网络往往表现出收敛性差、准确度降低，或不同时间尺度间分辨率不均衡等问题（Chen 等，2024）。

为缓解这些问题，本文采用反向兼容物理信息神经网络（BC-PINN）框架，该框架针对长时固结进行了专门设计，并通过若干改进提升其实用性和稳定性（Mattey 和 Ghosh，2022年）。与 BC-PINN 一样，时间尺度被划分为若干线段，并采用热启动按顺序训练；不过，在推进到下一段时，不再累积此前所有线段的残差惩罚，而是引入滞后损失方案，在紧邻的前一段上评估目标函数。这种向后回顾的一致性检查结合显式的界面连续性约束，在减缓各线段间漂移的同时，降低了计算量和内存占用的增长。首先，将时间域划分为 n_{Max} 个连续线段（图3）：

$$[T_0, T_1], [T_1, T_2], \dots, [T_{n-1}, T_n], \dots, \left[T_{n_{Max}-1}, T_{n_{Max}} = T\right] \tag{14}$$

其中，第 n^{th} 个时间区间可表示为 $\Delta T_n = [T_{n-1}, T_n]$、n =1...、$n_{max}$。

每个线段共享相同的空间域 Ω，但各自关联独立的神经网络参数$\boldsymbol{\theta}_n$。对于第一个子域 $\Omega \times \Delta T_1$，采用标准物理信息神经网络的损失函数，该函数结合了初始条件（IC）、边界条件（BC）和偏微分方程残差项（如图3所示）：

$$\mathcal{L}_{\Delta T_1}(\boldsymbol{\theta_1}) = \underbrace{\omega_{\mathrm{IC}}\mathcal{L}_{IC}^{(1)}(z_k^i, 0) + \omega_{\mathrm{BC}}\mathcal{L}_{BC}^{(1)}(z_k^b, t_k^b)}_{Physics\ Loss} + \underbrace{\omega_R\mathcal{L}_R^{(1)}(z_k^r, t_k^r)}_{Data\ Loss},$$
$$(z_k^i, 0) \in \Omega,\ (z_k^b, t_k^b) \in \partial\Omega \times [T_0, T_1], (z_k^r, t_k^r) \in \Omega \times [T_0, T_1] \tag{15}$$

其中，$\mathrm{L}_{\mathrm{IC}}^{(1)}$是$t = T_0$处初始条件的均方误差（MSE），$\mathrm{L}_{\mathrm{BC}}^{(1)}$ 是$\partial\Omega \times \Delta T_1$上边界条件的均方误差，$\mathrm{L}_{\mathrm{R}}^{(1)}$ 是在$\Omega \times \Delta T_1$内部配点评估的控制偏微分方程残差的均方误差；ω_{IC}、 ω_{BC}和ω_{R} 为常数权重。

对于所有后续子域 Ω ×（n = 2,..., n_{Max}），其中 ΔT_n，提出了一种新的损失函数，其不同之处在于，该函数被要求满足前一线段训练所获得的解：

$$\mathcal{L}_{\Delta T_n}(\boldsymbol{\theta_n}) = \underbrace{\omega_{\mathrm{IC}}\mathcal{L}_{IC}^{(n)}(z_k^i, T_{n-1}) + \omega_{\mathrm{BC}}\mathcal{L}_{BC}^{(n)}(z_k^b, t_k^b)}_{Physics\ Loss} + \underbrace{\omega_R\mathcal{L}_R^{(n)}(z_k^r, t_k^r) + \omega_s\mathcal{L}_S^{(n)}(z_k^s, t_k^s)}_{Data\ Loss},$$
$$(z_k^i, T_{n-1}) \in \Omega \times \{T_{n-1}\},\ (z_k^b, t_k^b) \in \partial\Omega \times [T_{n-1}, T_n], (z_k^r, t_k^r) \in \Omega \times (T_{n-1}, T_n],\ (z_k^s, t_k^s)$$
$$\in \Omega \times [T_{n-2}, T_{n-1}] \tag{16}$$

此处，$\mathrm{L}_{\mathrm{IC}}^{(n)}$、$\mathrm{L}_{\mathrm{BC}}^{(n)}$ 和 $\mathrm{L}_{\mathrm{R}}^{(n)}$ 的定义与第一线段类似，但在 $\Omega \times \Delta T_n$ 上进行评估。 ω_s 控制向后兼容性约束的强度。

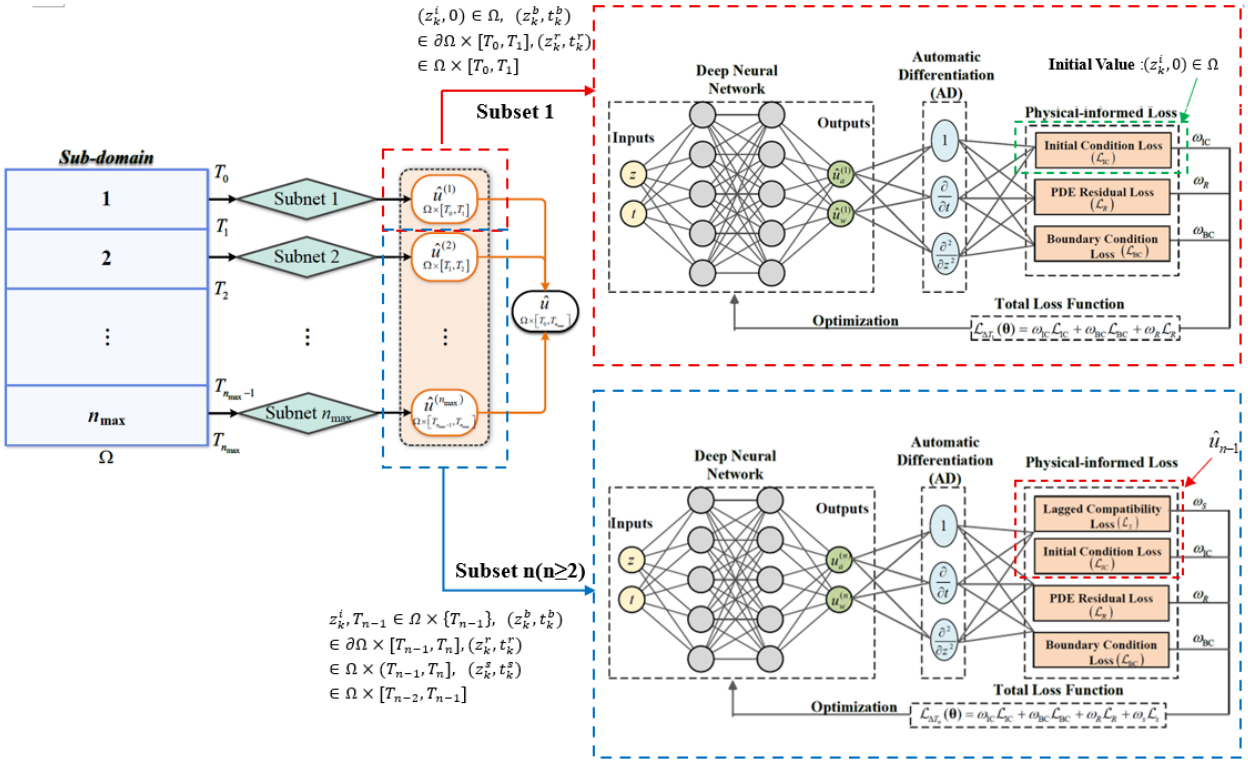


图3. 滞后反向兼容物理信息神经网络（LBC-PINN）示意图

如图 4 所示，附加的滞后相容性损失项 $L_S^{(n)}$ 强制与先前训练的线段 $\Delta T_{n-1} = [T_{n-2}, T_{n-1}]$ 保持反向兼容性。

$$\mathcal{L}_S^{(n)}(z_k^s, t_k^s) = \frac{1}{N_S}\sum_{k=1}^{N_S}\left[\left(\hat{u}_a^{(n)}(z_k^s, t_k^s) - \hat{u}_a^{(n-1)}(z_k^s, t_k^s)\right)^2 + \left(\hat{u}_w^{(n)}(z_k^s, t_k^s) - \hat{u}_w^{(n-1)}(z_k^s, t_k^s)\right)^2\right], (z_k^s, t_k^s) \in \Omega \times [T_{n-2}, T_{n-1}] \tag{17}$$

其中，$\hat{u}_a^{(n)}, \hat{u}_w^{(n)}$ 和 $\hat{u}_a^{(n-1)}$，$\hat{u}_w^{(n-1)}$ 分别表示当前及前一线段网络在前一时间步 $\Omega \times [T_{n-2}, T_{n-1}]$ 的孔隙空气压力和孔隙水压力预测。L_s 用于衡量当前预测解与先前训练所得解之间的偏差。引入该项可确保反向兼容性。

最后，通过在各自的时间区间内评估特定线段的模型，并拼接结果以覆盖整个时空区域，得到所报告的 LBC-PINN 解。与 STD-PINN 方法相比，所提出的 LBC-PINN 还采用了（i）时空域归一化和（ii）分段迁移学习。

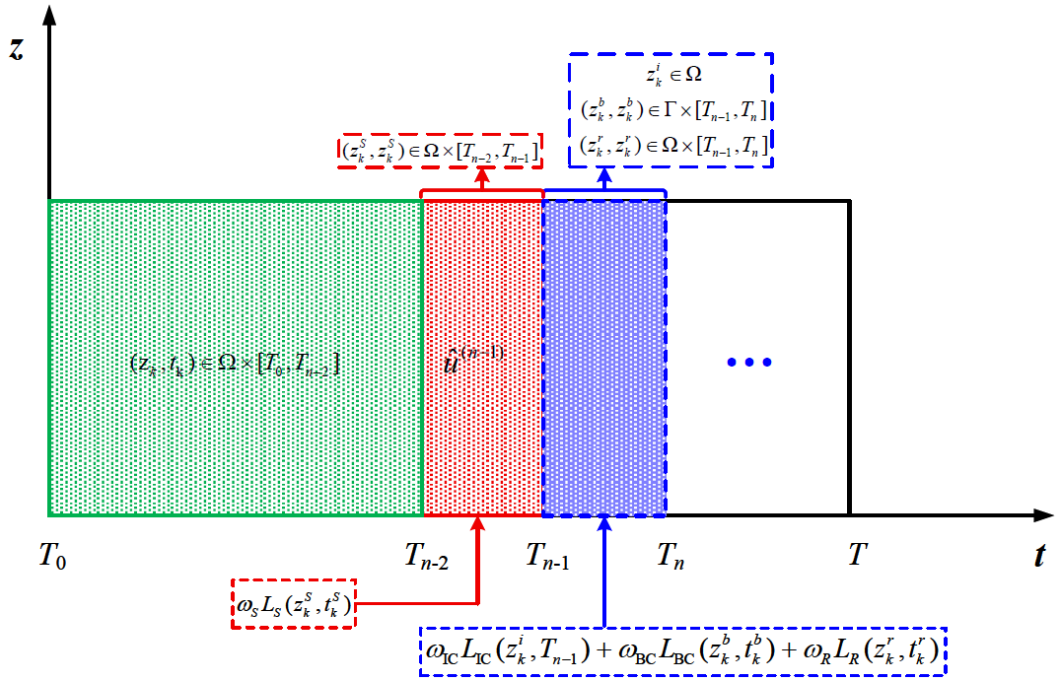


图3. 所提出的滞后向后兼容方案示意图。

3.2.1 时空域的归一化

为确保多个时间段内训练稳定，将时空区域归一化至单位计算域。这种归一化是数据驱动深度学习框架中的标准预处理技术，可避免原始坐标跨越多个数量级时产生的隐性重加权效应（Kumar 和 Premalatha，2021年；Chen 等，2024年；Xu 等，2025年）。在本研究中，深度 $z \in [0, H]$ 和物理时间 $t \in [0, T_{\max}]$ 首先以量纲形式定义。对于每个时间段 $[T_{n-1}, T_n]$，随后按如下方式生成内部配点。在层内均匀采样空间坐标，$z \in [0, H]$。为使对数时间轴上的分辨率均衡，在 $\log_{10} T_{n-1}$和$\log_{10} T_n$之间均匀采样物理时间的以 10 为底的对数，并

通过 $t = 10^{\log_{10} t}$映射回物理时间。最后，将所得数对（z， t）归一化至参考域 $\bar{z}$， $\bar{t} \in [0,1]$，具体如下：

$$\bar{z} = \frac{z - z_{min}}{z_{max} - z_{min}} \tag{18}$$

$$\bar{t} = \frac{t - T_{n-1}}{\Delta T_n} \tag{19}$$

其中，z_{min} 和z_{max} 分别是 z 的最小值和最大值。经过此变换后，控制方程及其相应的初始条件和边界条件均可用归一化坐标（$\bar{z}$， $\bar{t}$）表示，并施加于单位计算域上。

$$\frac{\partial u_a}{\partial \bar{t}} + C_a \frac{\partial u_w}{\partial \bar{t}} + \left(\frac{\Delta T_n}{H^2}\right) c_v^a \frac{\partial^2 u_a}{\partial \bar{z}^2} = 0 \tag{20}$$

$$\frac{\partial u_w}{\partial \bar{t}} + C_w \frac{\partial u_a}{\partial \bar{t}} + \left(\frac{\Delta T_n}{H^2}\right) c_v^w \frac{\partial^2 u_w}{\partial \bar{z}^2} = 0 \tag{21}$$

$$u_a(0, \bar{t}) = u_w(0, \bar{t}) = 0, \quad \bar{t} > 0; \tag{22}$$

$$\frac{\partial u_a(z, \bar{t})}{\partial \bar{z}}|_{\bar{z}=1} = \frac{\partial u_w(z, \bar{t})}{\partial z}|_{\bar{z}=1} = 0, \quad \bar{t} > 0; \tag{23}$$

在线段 k（$\bar{t} = 0$，即 $t = T_{n-1}$）开始时，初始条件可写为 ：

$$u_a(\bar{z}, 0) = u_a(\bar{z}, T_{n-1}), u_w(\bar{z}, 0) = u_w(\bar{z}, T_{n-1}) \tag{24}$$

3.2.2 迁移学习

迁移学习通过将从紧邻的前一线段中学到的权重和偏置，用于初始化每个 LBC-PINN 来实现。从线段 k-1 转至线段 k 时，使用线段 k-1 结束时获得的权重和偏置初始化线段 k 的网络。对于第一条线段，或切换到具有不同初始条件的数据时，则使用在该替代初始条件下训练的模型进行初始化。可在预热期内短暂冻结较早的层，然后解除冻结以进行完整微调（Goswami 等人，2020；Liu 等人，2023）。各线段的无量纲化和损失权重保持不变。此设置适用于在紧邻的前一线段上评估目标函数的滞后损失方案，

可改善接口处 u_a 和 u_w 的连续性，并减少优化器步数。

4 实现

对于正问题，神经网络（STD-PINN 和 LBC-PINN）的架构包含5个隐藏层，每层有50个神经元。STD-PINN 和 LBC-PINN 均选择双曲正切函数 作为激活函数。双曲正切函数激活函数的优点在于其连续（范围 $[-1,1]$）且可微（Dubey等，2022）。神经网络包含超过100,000个学习参数，并采用“Xavier初始化”技术进行初始化（Glorot和Bengio，2010）。使用 L-BFGS优化器（Kiyani等，2025）优化损失函数并更新学习参数（神经网络的权重和偏置）。每个线段的初始、边界和内部采样点数分别为2000、2000和10000。所有神经网络训练均使用 Python 3.12.3 和 PyTorch 2.5.0，在 NVIDIA V100（PCIe；32 GB HBM2、5,120个 CUDA核心、640个张量核心）上执行。我们在 COMSOL Multiphysics 中实现了一维有限元模型，以根据第2节所述的控制方程、边界条件和初始条件生成孔隙空气压力和孔隙水压力的参考解。除非另有说明，所有 STD-PINN 和 LBC-PINN 结果均与这些有限元解进行比较。为进一步验证，在选定深度和时间处计算 Ho和Fatahi（2014）的解析解，并将其与有限元解和 LBC-PINN 结果叠加比较。

PINN 训练完成后，测量 PINN 预测解与有限元解之间的相对误差和绝对误差。采用两个评价指标，即平均绝对误差（MAE）和平均相对误差（MRE），来评估 PINN 或 LBC-PINN 与 COMSOL 参考结果之间的误差，具体如下：

$$MAE = \frac{1}{N}\sum_{k=1}^{N}|u_i(z_k,t_k) - \tilde{u}_i(z_k,t_k)| \tag{25}$$

$$MRE = \frac{1/N \sum_{k=1}^{N} \left(u_i(z_k, t_k) - \tilde{u}_i(z_k, t_k)\right)}{1/N \sum_{k=1}^{N} \left(\tilde{u}_i(z_k, t_k)\right)} \tag{26}$$

其中，u_i 是 PINN 或 LBC-PINN 预测的模拟值（例如, u_a、 u_w ）；u_i 是 COMSOL 解的模拟值；n 为样本数。所有误差指标均相对于有限元解计算。

5. 结果与讨论

本节参照 Ho和Fatahi（2014）报道的单向排水条件下非饱和土一维固结解析解，展示相关结果。固结区域内土体的初始厚度 H 为 10 m。水（k_w）和空气（k_a）的渗透系数均为 10^{-10} m/s，固结过程中的孔隙率 (n) 为 $n = 0.5$，初始饱和度 (S_r) 为 80%。初始孔隙气压力和孔隙水压力分别为 20 kPa 和 40 kPa。其他相关土体参数列于表1。

表1 正问题中的土的物理性质参数

n	S_r (%)	k_a (m/s)	k_w (m/s)	C_a	C_w	c_v^w (m^2/s)	c_v^a (m^2/s)
0.50	80	10^{-10}	10^{-10}	0.0882	0.75	$6.3x10^{-8}$	$6.3x10^{-6}$

5.1. 正向分析：STD-PINN与有限元法结果

为评估 LBC-PINN 框架的有效性，首先将标准物理信息神经网络（STD-PINN）的预测性能与非饱和土一维固结的有限元法结果进行比较。STD-PINN 在整个时间域 $[0, T_{\max}]$ 内训练，未进行分段。所有物理损失和数据损失均直接在全局区域上计算。图 5 展示了在逐步扩展的时间范围 $[0, T_{\max}]$内训练得到的 STD-PINN 预测结果，其中 $T_{\max} \in [10^6, 10^7, 10^8, 10^9, 10^{10}]$s；同时将其与有限元法得到的归一化超孔隙气压 ($u_a/u_{a0}$) 和孔隙水压力($u_w/u_{w0}$)结果进行比较。

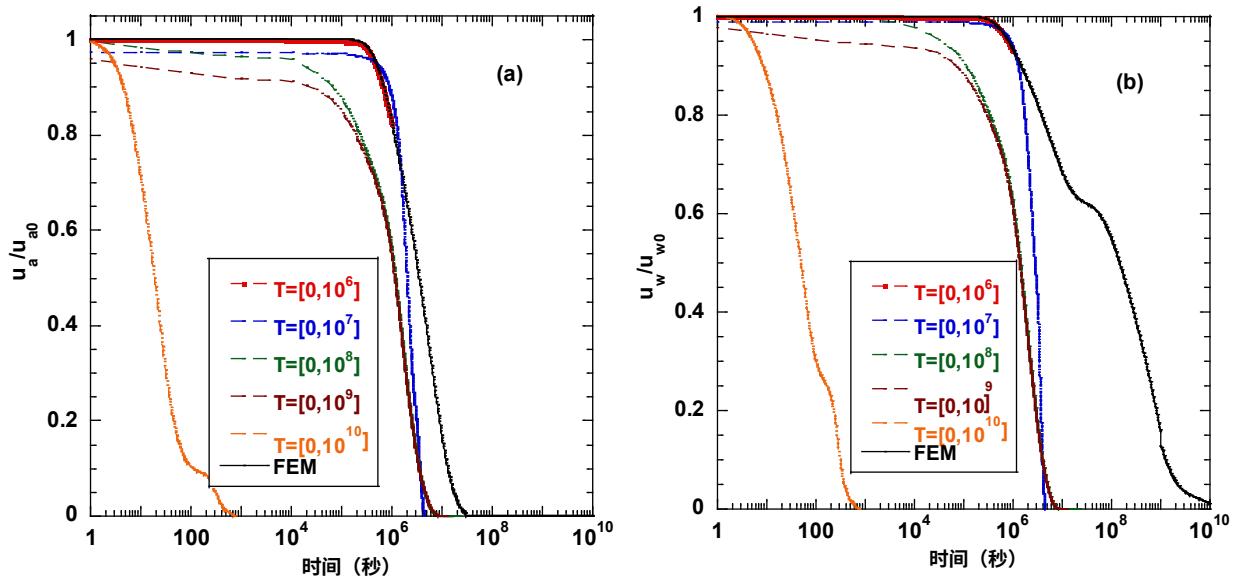


图 4. 所提出的 STD-PINN 与有限元法在不同时间范围 $[0, T_{\max}]$ 下的归一化 (a) 超孔隙气压 (u_a/u_{a0}) 和 (b) 超孔隙水压力 (u_w/u_{w0}) 的比较。

有限元法结果呈现出非饱和土固结中典型的两阶段消散过程。起初，超孔隙气压在较长时间内几乎保持不变，随后迅速消散，之后孔隙水压力缓慢消散。这种从空气主导流动向水主导流动的转变跨越了多个数量级的时间尺度，从而形成平滑且符合物理规律的压力衰减（Fredlund 和 Hasan，1979年）。相比之下，随着 $T_{\max}$ 增大，STD-PINN 与参考值之间的偏差逐渐增大。在最短训练时间范围 $[0, 10^6]$ s 内，STD-PINN 能够捕捉孔隙气压和孔隙水压力整体下降的趋势，并以合理的准确度再现早期消散过程。在 $[0, 10^7]$ s 内训练的 STD-PINN 模型，其孔隙气压和孔隙水压力消散相较有限元法结果均出现时间滞后。当扩展至 $[0, 10^8]$和 $[0, 10^9]$ s 时，该模型对孔隙水压力骤降的幅度和时机以及孔隙水压力完全消散的时机估计得越来越低，表明网络难以兼顾早期和后期的动态变化。采用完整时间范围 $[0, 10^{10}]$ s 后，STD-PINN 与参考值的偏差最大。随着 $T_{\max}$ 增大，预测精度不断下降，这凸显了采用单段标准PINN模拟非饱和土固结的局限性。

深度网络已被充分证实存在偏向低频/缓慢变化成分的频谱偏差，这使其无法充分解析早期时间的快速消散；当训练时间窗口跨度达数个数量级时，模型反而会产生滞后且过于陡峭的下降（Rahaman 等人，2019年）。此外，物理信息神经网络在不同损失项上的收敛速率和病态性存在显著差异，因此优化器会倾向于区域中“容易”处理的后期部分，而忽视早期瞬态；随着全局时间范围 $[0, T_{max}]$ 增大，这一问题会进一步加剧（Wang 等人，2022年）。

5.2. 正向分析：LBC-PINN 与有限元法结果

图 6 展示了 LBC-PINN 训练过程中四个代表性时间段（总时间段 N=5）内总损失、数据损失和物理损失的演化。各时间段中两个损失分量（例如数据损失、物理损失）均持续下降，证实了 LBC-PINN 框架在求解长期非饱和土固结问题时的稳定性和适应性。随着时间窗口从早期阶段转向后期阶段，数据损失和物理损失的相对贡献发生了显著变化。在最早的时间段（$0 \leq t \leq 10^2$ s；图 6a）中，物理损失占主导地位，并与总损失 closely tracks这一现象反映出，在早期阶段，孔隙空气和孔隙水压力的消散主要受初始条件和边界条件约束，因为此时孔隙空气、孔隙水压力几乎尚未发生消散，且偏微分方程残差仍然较小。随着时间推移（图 6b–6c），数据损失开始对总损失产生更显著的贡献。在最后一个时间段（$10^8 \leq t \leq 10^{10}$ s；图 6d）中，数据损失成为主导分量，并与总损失 closely aligns，而物理损失则保持较低且更加稳定。这一转变表明，随着系统进入水主导消散阶段，解越来越多地由偏微分方程所描述的内部动力学控制，而较少受边界条件和初始条件约束。

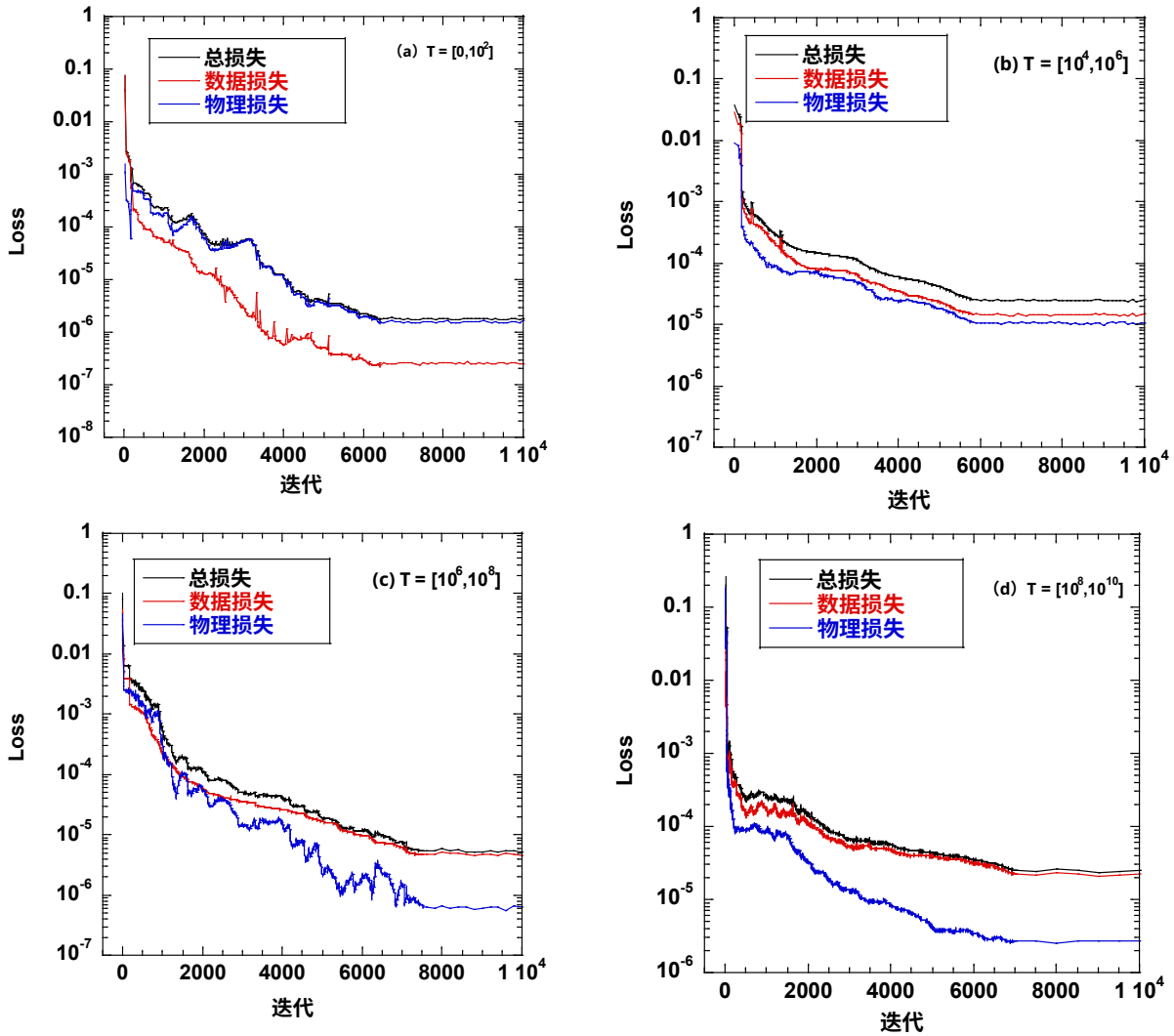


图 5. LBC-PINN 在以下条件下随迭代次数变化的损失：（a）$T = [0,10^2]$、（b）$T = [10^4,10^6]$、（c）$T = [10^6,10^8]$、（d）$T = [10^8,10^{10}]$。

图7 比较了深度 z = 2.5、5.0 和 7.5 m 处归一化超孔隙空气压力和孔隙水压力的 LBC-PINN 预测结果与有限元法结果。在所有深度处，LBC-PINN 在整个时间域内均 closely follows 有限元法基准结果。它准确捕捉了孔隙空气和孔隙水压力消散开始时间随深度变化而产生的延迟，并再现了u_w/u_{w0}的多阶段 S 形衰减，包括 $10^7 \sim 10^8$s 附近的典型肩部特征。从 Ho 和 Fatahi（2014）提取的方形标记点与三个深度处的有限元法结果和 LBC-PINN 结果均吻合良好。

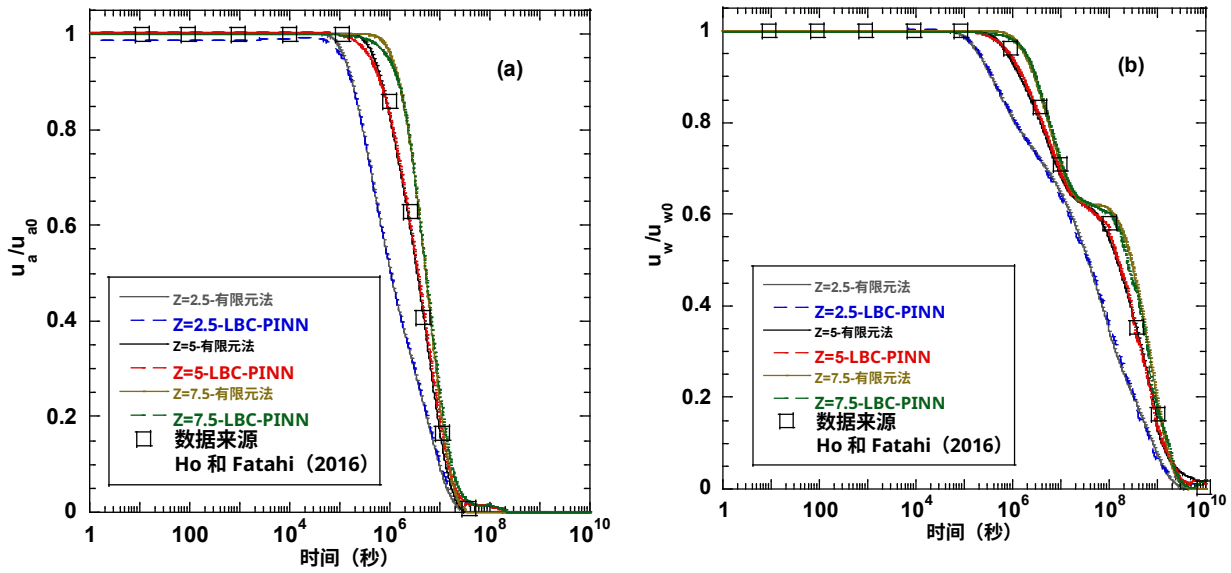


图 6。 LBC-PINN 与有限元法在选定深度处的 (a) (u_a/u_{a0})、(b) (u_w/u_{w0})结果对比。类似地，图 8 展示了在不同时间因子 T_v $(T_v = k_w t/m_s^1 \gamma_w H^2)$ 下，孔隙气压和孔隙水压力随深度的消散情况。LBC-PINN 的结果（虚线）在所有时间快照下都与有限元法（实线）的解高度吻合。图 9 对比了 LBC-PINN 和有限元法预测的归一化孔隙气压和孔隙水压力的时空场。LBC-PINN 的结果与整个区域内的有限元法结果吻合极佳。绝对误差图显示，误差较小，主要集中在早期的边界区域，且后期没有显著累积。对于孔隙气压，较大的误差出现在中间时刻和中等深度处；对于孔隙水压力，较明显的偏差则出现在后期的较深区域。峰值误差低于 0.05，证实 LBC-PINN 能够准确捕捉两个相在空间和时间上的变化。

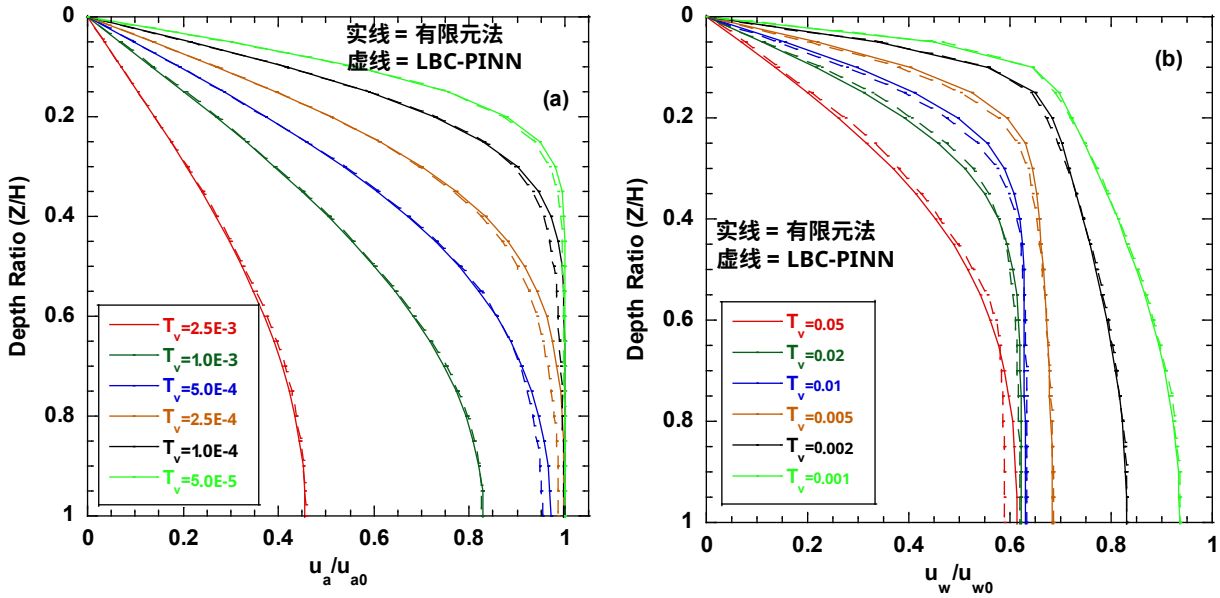
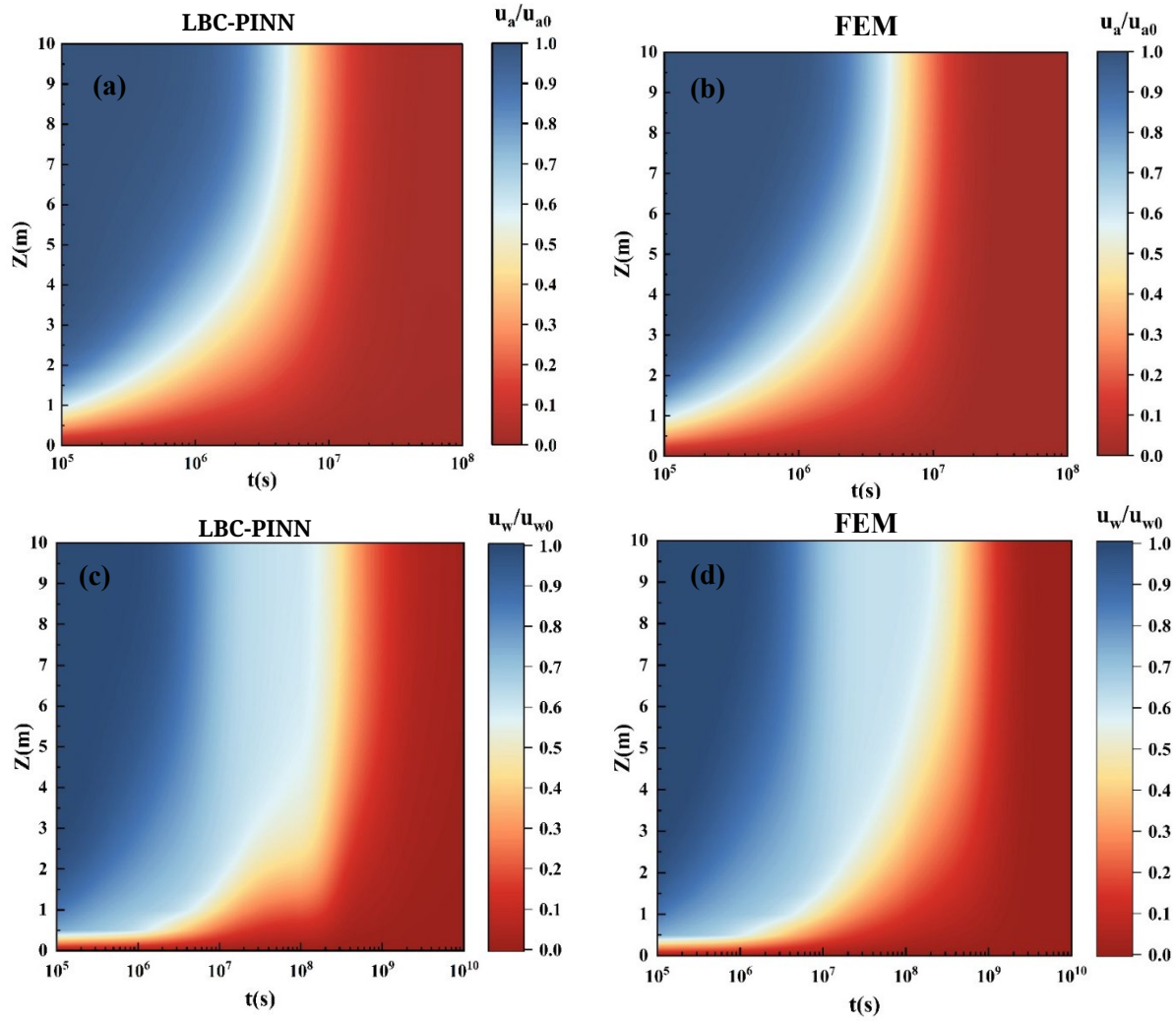


图7。 选定 T_v 下 LBC-PINN 与有限元法结果之间（a）（u_a/u_{a0}）和（b）（u_w/u_{w0}）的比较。



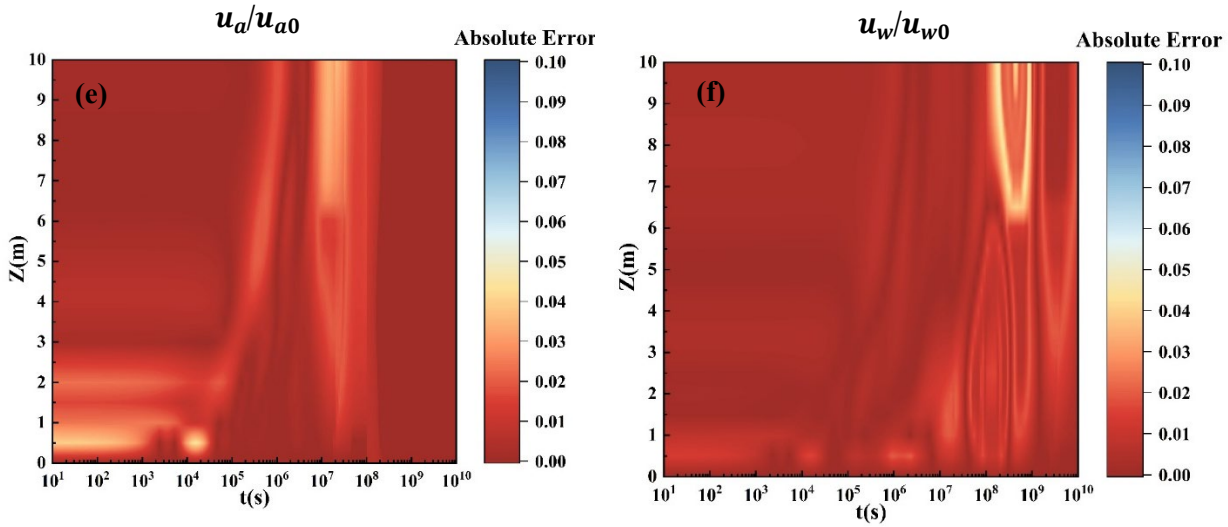


图 8。 (a) LBC-PINN 的 (u_a/u_{a0})、(b) 有限元法的 (u_a/u_{a0})、(c) LBC-PINN 的 (u_w/u_{w0})、(d) 有限元法的 (u_w/u_{w0}) 的时空分布图。LBC-PINN 与有限元解之间的绝对误差：(e) (u_a/u_{a0})、(f) (u_w/u_{w0})。

5.3 敏感性分析

5.3.1 时间分段的影响

为评估时间分段对模型精度的影响，在固定仿真时域 $[0, 10^{10}]$ 内，采用逐渐增加对数尺度时间分区数量（$N = 2$ 至 6）的方式训练分段式 LBC-PINN。如表 2 和图 10 所示，$N = 2$ 的情况表现最差，合并平均绝对误差（MAE）为 0.233，平均相对误差（MRE）为 0.321，峰值绝对误差超过 1.0。决定系数（$R^2 = 0.116$）表明，该模型未能捕捉非饱和土固结问题中固有的陡峭梯度和多尺度行为。增加到三个线段（N = 3）后，MAE 显著降至 0.028，且R^2有所提升（0.961）；但 MRE 仍然较高（0.123），这主要是由于在后期真实值较小的时段（如图 11 所示），孔隙气压和孔隙水压力的相对误差放大所致。一旦时间段数量达到四个（$N = 4$），模型在孔隙气压和孔隙水压力两种压力相下均实现了高准确度，合并 MAE 低于 0.011，MRE 约为 0.059，且两者的 R2 均超过 0.9995u_a 和u_w。结果$N = 5$和$N = 6$ 几乎相同，所有关键指标的变化都很小。这一平台期表明，五个时间段足以解析问题的时间复杂性。随时间和分段变化的相对误差见图 11。对于u_a，相对误差

主要集中在中间时间范围（10^6~10^8 s），此时快速耗散并与u耦合

w 引发陡峭梯度 · 相比之下，u_w 在 106 s 之后，相对误差逐渐上升。随着 N 增加，性能有所提升，这与已知的物理信息神经网络在长时间跨度的刚性偏微分方程上的局限性一致。神经网络往往更容易学习低频分量，这种现象称为频谱偏差（Rahaman 等人，2019年）；当网络在较宽的时间域上训练时，这会导致其无法充分拟合剧烈的早期时间动力学。此外，粗略的线段划分会导致梯度失衡：早期时间的物理损失被更平滑的长期行为掩盖，从而降低收敛性和准确度（Wang 等人，2021年）。将时间域划分为更多线段后，每个子问题所涵盖的时间范围会更窄，从而提高近似质量、优化器条件性和采样有效性。这种方法与域分解和序贯学习领域的既有研究一致，包括 XPINNs（Jagtap 等人，2020年）和 bc-PINNs（Mattey 和 Ghosh，2022年）；这些研究主张通过分阶段向后兼容训练来求解多尺度偏微分方程。总体而言，将时间域划分为五个或更多子区间，可以把刚性的长期学习任务转化为一系列适定问题，其中 N =5 在准确度与计算效率之间实现了最佳平衡。

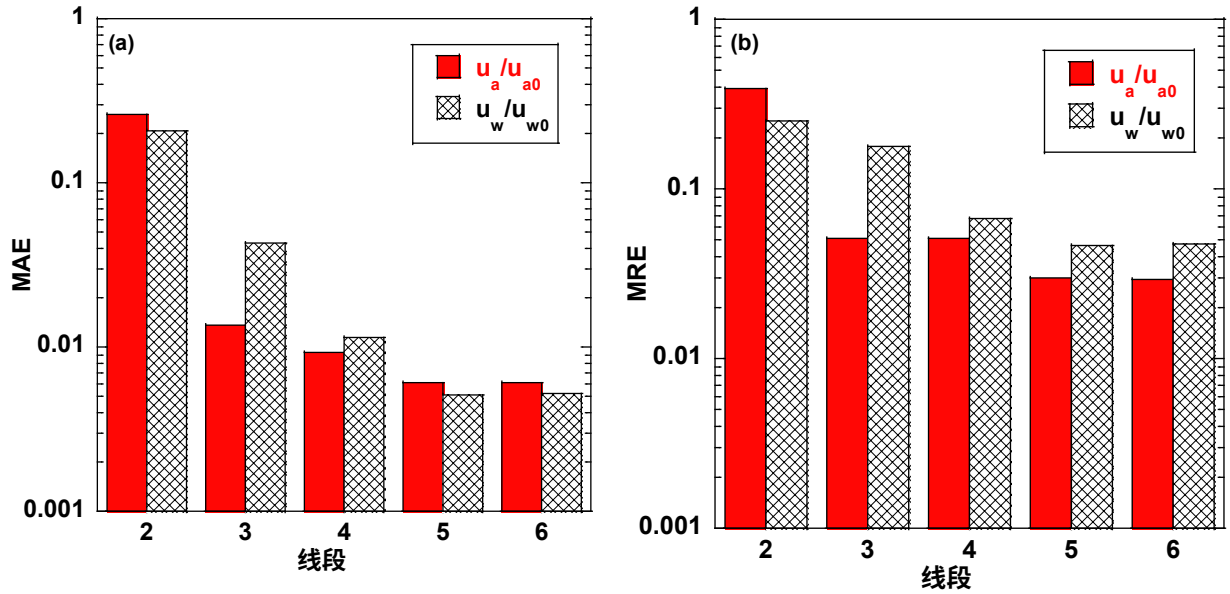


图 9. 准确度与时间分段级别（N）的关系：(u_a/u_{a0}) 和 (u_w/u_{w0}) 的 (a) 平均绝对误差与 (b) 平均相对误差。

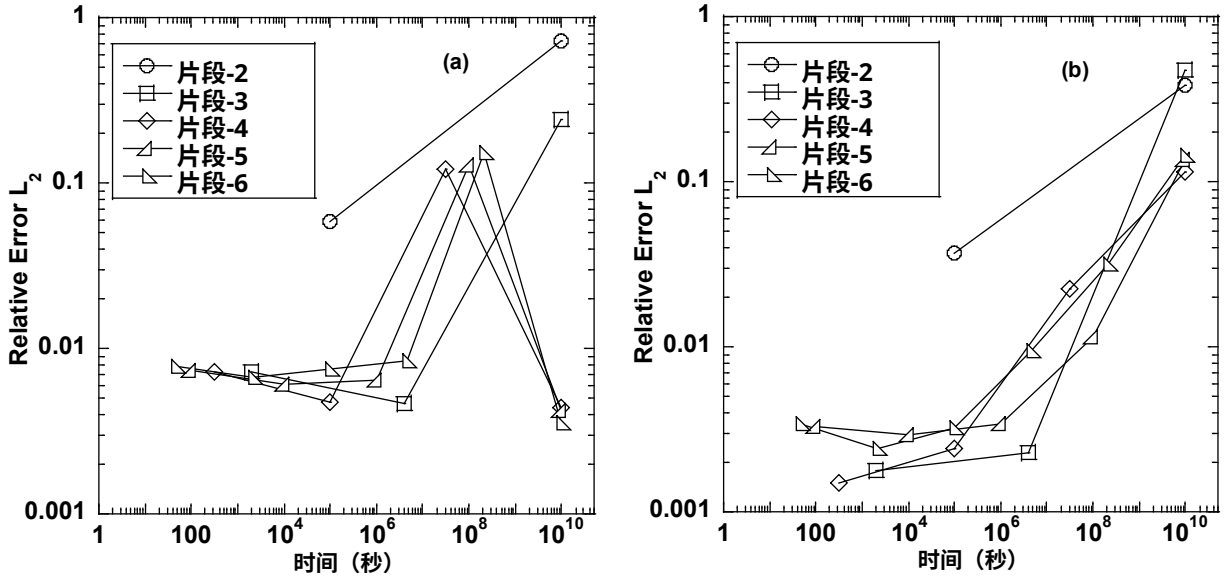


图10。不同训练分段随时间变化的相对误差：(a) (u_a/u_{a0})，(b) (u_w/u_{w0})。

表 2 有限元法与采用时间分段级别的 LBC-PINN 的比较

时间段 (N)	变量	MAE	MRE	最大绝对误差	R^2
2	u_a	0.2598	0.3880	>1.0	0.1271
	u_w	0.2071	0.2529	>1.0	0.0483
	合并 (u_a+u_w)	0.2334	0.3205	>1.0	0.1157
3	u_a	0.0135	0.0516	0.2796	0.9921
	u_w	0.0429	0.1779	0.5352	0.9098
	合并（u_a+u_w）	0.0282	0.1226	0.5352	0.9606
4	u_a	0.0093	0.0514	0.1041	0.9995
	u_w	0.0114	0.0668	0.0493	0.9995
	合并（u_a+u_w）	0.0114	0.0591	0.0767	0.9995
5	u_a	0.0061	0.0301	0.0421	0.9995
	u_w	0.0051	0.0469	0.0485	0.9995
	合并（u_a+u_w）	0.0056	0.0385	0.0453	0.9995
6	u_a	0.0061	0.0291	0.0432	0.9995
	u_w	0.0052	0.0477	0.0488	0.9996
	合并（u_a+u_w）	0.0054	0.0384	0.0469	0.9995

注： 平均绝对误差 = 平均绝对误差；平均相对误差 = 平均相对误差；决定系数 = 决定系数。合并 = 指标基于串联的 u_a 和 u_w序列计算。

5.3.2 气水渗透率比的影响

图12展示了所提出的 LBC-PINN 和有限元法在气水渗透率比为 k_a/k_w 时预测的孔隙气压和孔隙水压力。结果表明，不同渗透率比下的气压衰减曲线呈平行趋势，其中较高的 k_a/k_w 值会使消散更早发生，且趋势更陡。例如，当 k_a/k_w=10.0 时，孔隙气压迅速消散，并在约 5×10^6 秒时几乎完全释放；而当 k_a/k_w=1.0 时，孔隙气压要到 5×10^7 秒才完全释放。如图12b所示，单向排水条件下的孔隙水压力消散遵循两阶段过程。起初，孔隙水压力迅速下降，尤其是在较高的 k_a/k_w 条件下，这是由于气体同时流出，形成“上S形”曲线。一旦孔隙气压完全消散，所有孔隙水压力曲线便汇聚并逐渐下降，呈现出经典饱和土固结的特征。第二阶段持续至约 10^{10} 秒。LBC-PINN 在所有工况下均与数据高度吻合，准确捕捉了早期消散过程以及长达 10^{10} 秒时间尺度内的长期固结行为。在低渗透率工况（$k_a/k_w = 0.1$）下观察到轻微的相位滞后，但总体趋势和幅值均得到了良好保留。

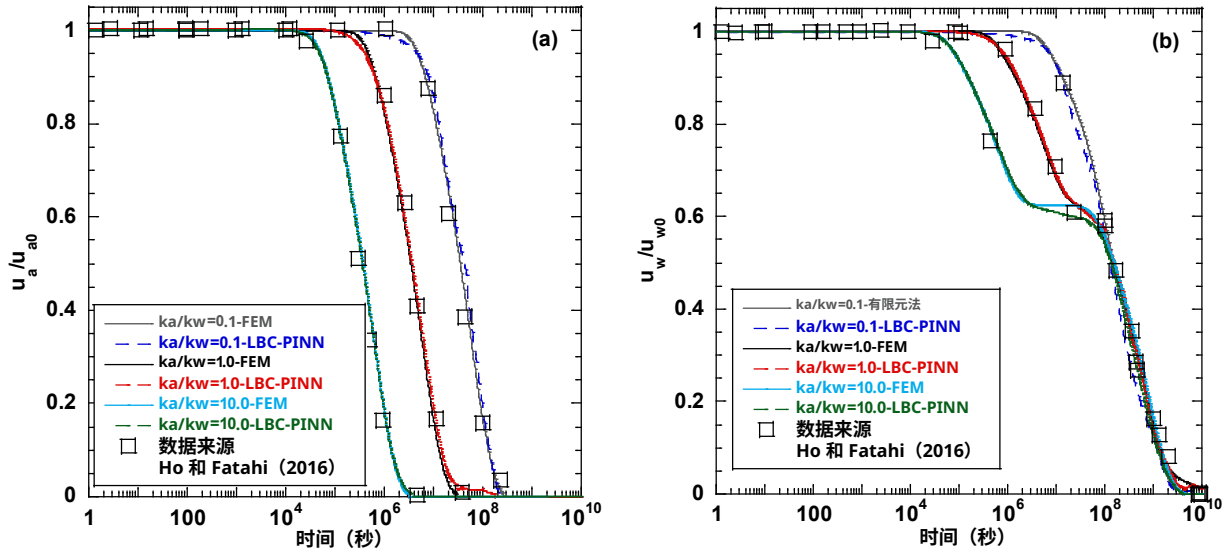


图11. 在气水渗透率比为 k_a/k_w 时，LBC-PINN 与有限元法结果之间的 (a) (u_a/u_{a0})、(b) (u_w/u_{w0}) 对比。

图13a和图13b展示了预测的孔隙气压和孔隙水压力的平均绝对误差（MAE）和平均相对误差（MRE）随 k_a/k_w 变化的情况，其范围为0.001至1000。 总体而言，LBC-PINN 在所有工况下的误差均较低，MAE 和 MRE 始终低于0.2。然而，当 $k_a/k_w < 1$ 时，误差明显增大，尤其是在 $k_a/k_w = 0.1$ 时，模型对 u_a和u_w的预测均出现最大的 MAE 和 MRE。 当 $k_a/k_w < 1$ 时，LBC-PINN 的准确度有所下降，尤其是在约0.1时，这是由于气压与水压力之间的耦合更强。空气渗透率较低会减缓气压消散，使 u_a和u_w 在较长时间内同时演变。由于在刚性条件下学习尖锐瞬态和相互作用的残差存在局限，物理信息神经网络对这类多尺度耦合动力学尤为敏感（Wang 等人，2021年）。相比之下，当 $k_a/k_w \geq 1$ 时，孔隙气压消散得更快，并在早期便与系统脱耦。因此，超孔隙水压力消散遵循更简单、平滑的固结路径，使网络能够实现更高的准确度和更快的收敛。

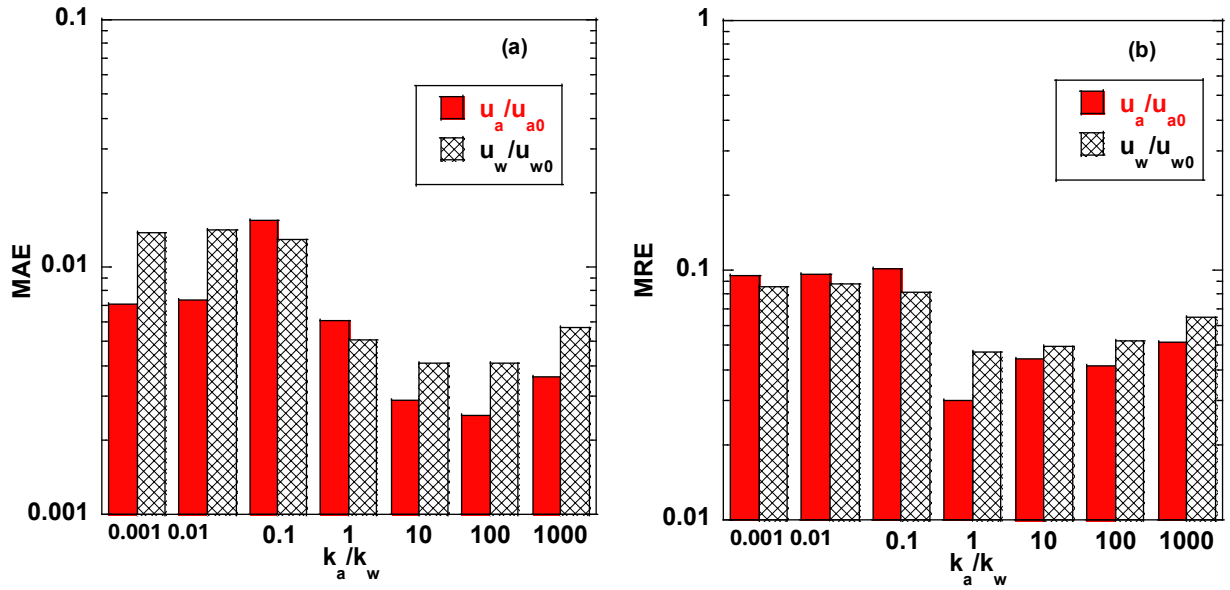


图12.准确度与 k_a/k_w的气水渗透率比：(a) (u_a/u_{a0})和(u_w/u_{w0})的平均绝对误差与(b)平均相对误差 .

图14展示了 LBC-PINN 和有限元解在单向排水条件下预测的归一化沉降（s^*）随时间的变化。结果表明，在所有情况下，LBC-PINN 与有限元解都高度吻合，说明该模型能够在较宽的时间范围内（最长达 10^{10}s）准确捕捉非饱和土的整体固结行为。不出所料，增大 k_a/k_w 会使沉降曲线向更早的时间移动，表明在此设置下固结响应更快。

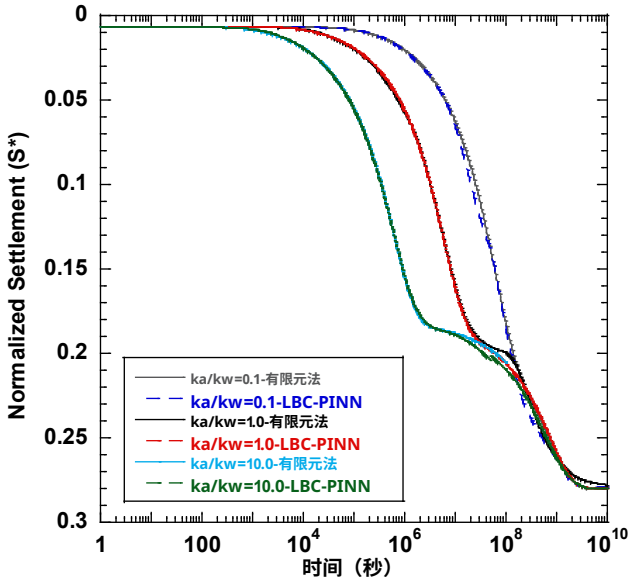


图13。 不同气水渗透率比k_a/k_w下，LBC-PINN与有限元法结果的归一化沉降（S*）对比。

表3 对比　　不同气水渗透率比k_a/k_w下有限元法与LBC-PINN之间的对比。

k_a/k_w。	变量	MAE	MRE	最大绝对误差	R^2
0.001	u_a	0.0071	0.0947	0.0541	0.996
	u_w	0.0138	0.0854	0.0523	0.996
	合并（u_a+u_w）	0.0104	0.0901	0.0532	0.996
0.01	u_a	0.0074	0.0957	0.0514	0.9960
	u_w	0.0142	0.0877	0.0502	0.9959
	合并（u_a+u_w）	0.0108	0.0917	0.0508	0.9960
0.1	u_a	0.0155	0.1019	0.0463	0.9972
	u_w	0.0129	0.0811	0.0484	0.9964
	合并（u_a+u_w）	0.0142	0.1019	0.0484	0.9969
1.0	u_a	0.0061	0.0301	0.0421	0.9995
	u_w	0.0051	0.0469	0.0485	0.9995

k_a/k_w。	变量	MAE	MRE	最大绝对误差	R^2
	合并（u_a+u_w）	0.0056	0.0385	0.0453	0.9995
10.0	u_a	0.0029	0.0441	0.0501	0.9997
	u_w	0.0041	0.0498	0.0787	0.9984
	合并（u_a+u_w）	0.0035	0.0469	0.0419	0.9987
100	u_a	0.0025	0.0414	0.0542	0.9950
	u_w	0.0041	0.0521	0.0414	0.9980
	合并（u_a+u_w）	0.0033	0.0467	0.0478	0.9965
1000	u_a	0.0036	0.0515	0.0485	0.9962
	u_w	0.0057	0.0644	0.0687	0.9941
	合并（u_a+u_w）	0.0046	0.0629	0.0486	0.9951

注： 平均绝对误差 = 平均绝对误差；平均相对误差 = 平均相对误差；决定系数 = 决定系数。合并 = 指标根据拼接后的 u_a 和 u_w 序列计算。

5.3.3 神经网络结构的影响

在 LBC-PINN 中，隐藏层数（深度）和每层神经元数（宽度）会显著影响预测精度。为评估模型对这些超参数的敏感性，我们测试了多种神经网络配置：隐藏层数设为 1、3、5、7 和 10，每层神经元数设为 10、30、50、70 和 100。如图 15a 所示，采用 5 个隐藏层、每层 50 个神经元的模型相对 L_2 误差最低，表现优于所有其他配置。总体而言，增大网络规模可以提高准确度；然而，当网络超过中等规模（例如，层数超过 5 层或神经元数超过 50 个）后，误差值便趋于平稳，并保持在同一数量级。仅有一个隐藏层的网络始终表现出较高误差，这凸显了在此应用中具备足够深度对于有效学习的必要性。

5.3.4 训练数据的影响

为探究配点密度对 LBC-PINN 框架准确度的影响，我们通过改变每条线段上的内部（残差）配点数和边界配点数，开展了敏感性研究。图 15b 展示了不同残差点数（1,000 至 20,000）和边界点数（50 至 2,000）组合下的相对 L_2 误差热力图。

结果表明，增加残差点数通常能提高准确度，尤其是将点数从 1,000 增至 10,000 时。值得注意的是，最低 L_2 误差（0.0385）出现在边界点数为 1,000、残差点数为 20,000 时。然而，边界点数超过 1,000 后，收益逐渐减小；在某些情况下（例如 2,000 个点），还会引发不稳定，这可能是过拟合所致。相反，无论边界点数多少，残差点数过少的模型（例如 1,000 个点）都会表现出明显更高的误差，这证实了在物理信息学习中充分施加内部约束至关重要。

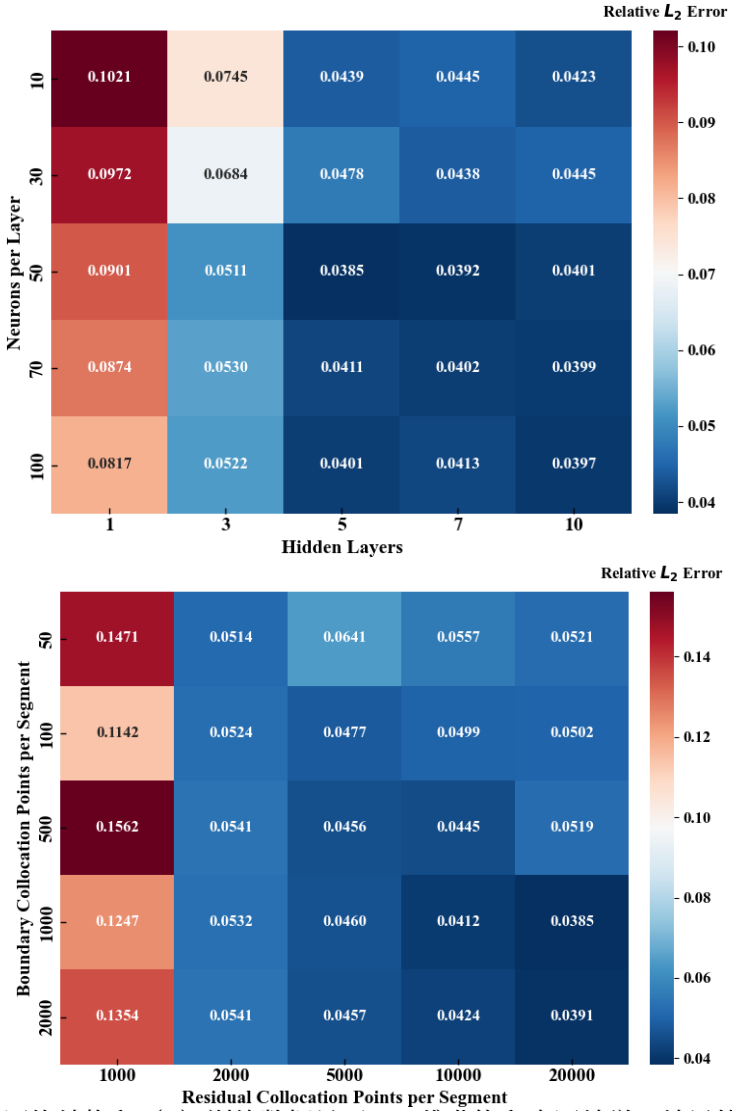


图14。不同（a）神经网络结构和（b）训练数据量下，一维非饱和土固结学习结果的平均相对误差。

5.4 简化时间分段策略

如图7所示，在早期，超孔隙气压几乎保持不变；随着消散开始，约在 10^6 s 时急剧下降。准确解析这一过程可能会产生较高的计算成本，尤其是在孔隙水压力消散之前发生这些动态变化时。为降低计算成本并提高训练效率，提出了一种简化策略，该策略基于孔隙气压完全消散的特征时间，表示为 $t_s = H^2/c_v^a$，其中 H 为排水路径长度，c_v^a 为气相固结系数。假定该时间（t_s）代表孔隙气压完全消散的时刻。为确保不同算例之间的一致性，将计算所得时间向下取整至最近的数量级，并将其作为时间分段的起点。随后，按照基线 LBC-PINN 所采用的相同方法对剩余时间域进行划分。

图 16 展示了采用简化时间分段策略预测的孔隙气压和孔隙水压力（u_a 和 u_w）的绝对误差分布。对于 $k_a/k_w = 0.1$，在时间范围 10^5 至 10^7s 内，底部边界附近出现较大的u_a 误差，这归因于孔隙气压消散滞后以及强烈的气水压力耦合。相应的u_w 误差也集中在消散前沿附近，尤其是在后期，表明解析耦合瞬态响应存在困难。在 $k_a/k_w = 1.0$ 时，u_a 和 u_w误差均处于中等水平且分布更均匀，反映了同步消散。当 k_a/k_w =10.0 时，气流加快使 u_a 误差迅速减小，而u_w 误差仍存在于底部和后期时段，此时孔隙水控制着主要的消散过程。

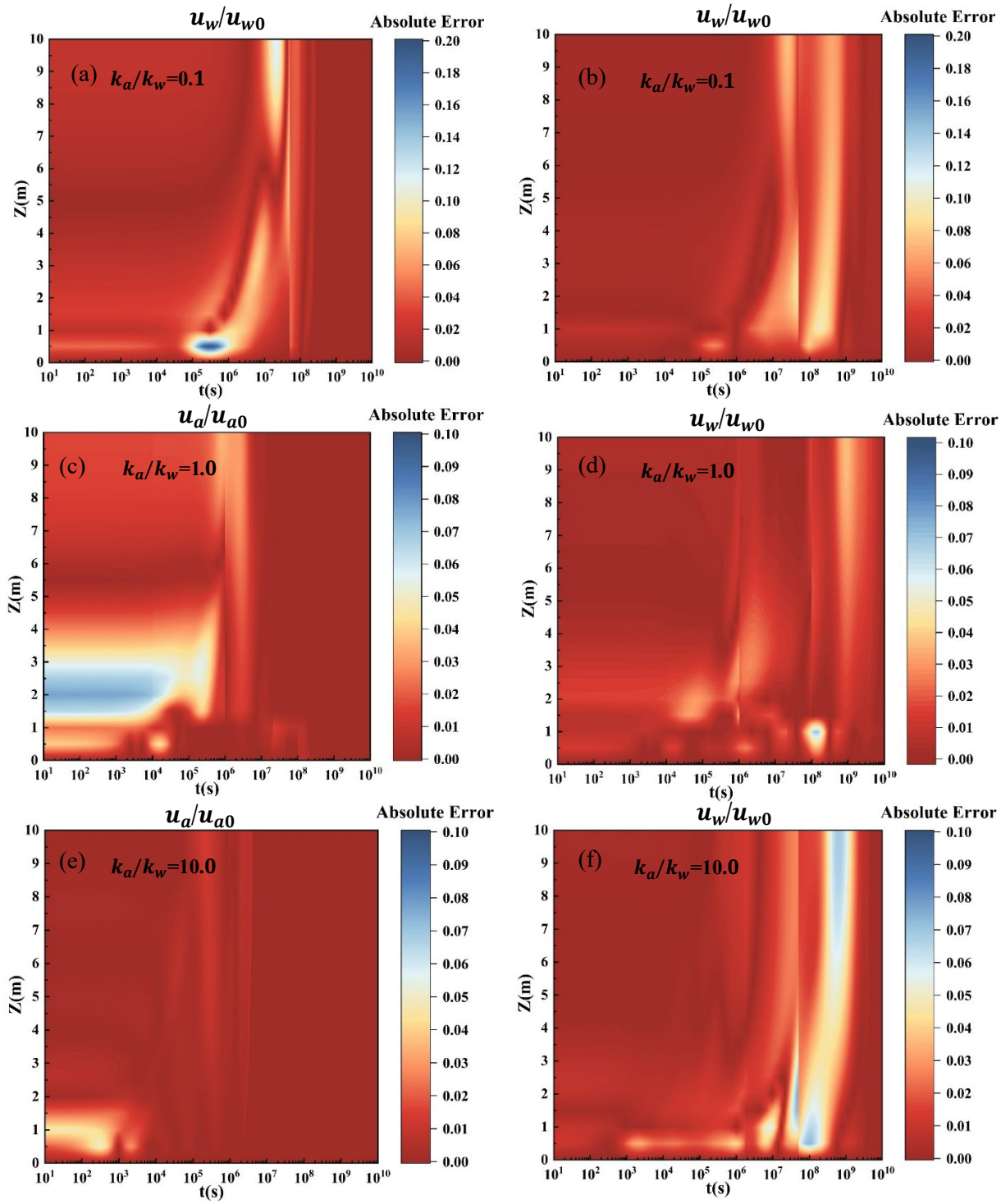


图15. 采用简化时间段划分策略时，LBC-PINN 与有限元解的绝对误差：(a) (u_a/u_{a0})，k_a/k_w= 0.1；(b) (u_w/u_{w0})，k_a/k_w= 0.1；(c) (u_a/u_{a0})，k_a/k_w= 1.0；(d) (u_w/u_{w0})，k_a/k_w= 1.0；(e) (u_a/u_{a0})，k_a/k_w= 10.0；(f) (u_w/u_{w0})，k_a/k_w= 10.0，

图17定量证实了绝对误差图中观察到的趋势。当$k_a/k_w \leq 0.1$时，u_a 和u_w 的合并平均绝对误差（MAE）和平均相对误差（MRE）明显较高，在$k_a/k_w = 0.01$时，MAE达到峰值0.0169，MRE达到峰值0.1764。这种较高的误差归因于该区域内强烈的气水压力耦合：两种压力同时变化并发生非线性交互，使解空间更加复杂，也更难以通过物理信息神经网络逼近。当k_a/k_w增大并超过1时，u_a 和两者的MAE与MRE均大幅下降。例如，u_a 的MAE从$k_a/k_w = 0.001$时的0.0174降至$k_a/k_w = 10$时的0.0032。u_w 的MRE则从$k_a/k_w = 0.01$时的0.1824降至$k_a/k_w = 10$时的0.0787。

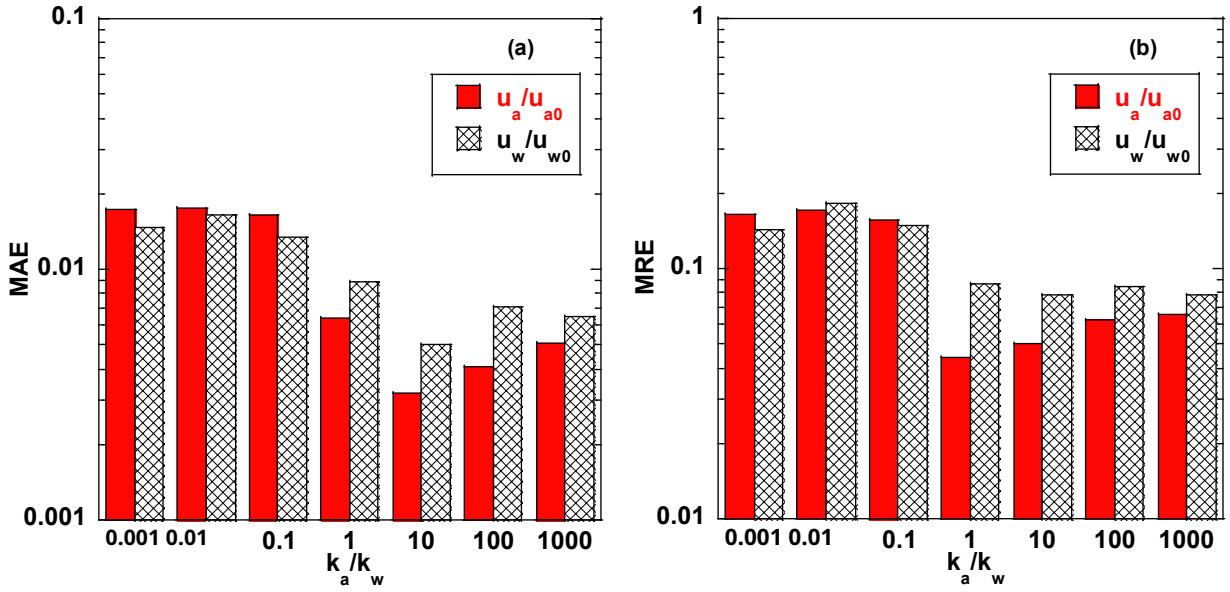


图 16. 采用简化时间段划分策略时，k_a/k_w 的准确度与气水渗透率比的关系：（a）(u_a/u_{a0})和（b）(u_w/u_{w0})的平均绝对误差和平均相对误差。

6. 总结与结论

本研究开发了一种滞后反向兼容物理信息神经网络（LBC-PINN），用于模拟一维非饱和土固结。该框架引入了时空域归一化、滞后兼容损失和迁移学习，以克服传统物理信息神经网络中的收敛问题和时间分辨率不平衡。通过有限元解对正向预测进行了基准验证，

并开展了一系列敏感性分析和反演分析，以评估其稳健性和参数可辨识性。得出以下结论和建议：

1) LBC-PINN 能够准确捕捉孔隙空气压力和孔隙水压力在长达 10^{10} s 的时间尺度内的耦合消散行为。对于推荐的分段设置（例如，$N \geq 5$），u_a 和u_w 的合并平均绝对误差约为 $10^{-3} \sim 10^{-2}$，与参考有限元解高度吻合。在时间段之间引入滞后兼容损失可确保平滑过渡并提高时间一致性，而迁移学习则加快了各时间段的收敛。

2) LBC-PINN 框架可处理广泛的渗透率比（k_a/k_w 从 0.001 到 1000），在气流或水流占主导的不同工况下均保持较低误差。在 $k_a/k_w = 0.1$ 处观察到准确度有所下降，这归因于气相和水相之间的强耦合。尽管如此，在整个测试范围内，预测结果始终符合物理规律。针对网络深度、宽度和配点密度进行的敏感性分析表明，增加残差点采样可提高预测精度，而过于密集的边界采样则会带来递减的收益。

3) 分段训练策略对于解析强烈的多尺度时间响应至关重要。线段数量较少时，误差仍然较大，且消散前沿的分辨率不足；而将时间域划分为四个或更多对数间隔窗口，则可获得稳定且高准确度的解。本文提出了一种基于特征气相消散时间（$t_s = H^2/c_{va}$）的简化策略，用于确定初始分段截断点。该简化方案的准确度与完整的对数分段相当，绝对误差仅集中在局部区域（例如底部附近或中间耦合阶段）。

当前模型仅适用于一维非饱和土固结。虽然它适用于许多实验室问题或理想化的现场问题，但实际岩土工程系统通常涉及二维或三维渗流与变形。尽管引入了一种基于物理机制的分段策略

该策略采用特征气相消散时间，但分段边界和网络设置仍需手动调整。未来可根据动态误差指标进行自适应分段，或自动优化训练时间窗。

致谢

作者声明，本研究由作者独立完成，未获得任何外部资助。

参考资料

Barden, L., 1965. 压实黏土和非饱和黏土的固结。《岩土工程》, 15(3), 267-286. Bekele, Y.W., 2021. 用于一维固结的物理信息深度学习。《岩石力学与岩土工程学报》,13(2),420-430. Biot, M.A., 1941. 三维固结的一般理论。《应用物理学杂志》, 12(2), 155- 164. Briaud, J.L., 2023. 岩土工程：非饱和土和饱和土。约翰·威立父子公司。Cargill, K.W., 1984. 极软土固结预测。《岩土工程学报》, 110(6),775-795. Cai, S., Mao, Z., Wang, Z., Yin, M. and Karniadakis, G.E., 2021. 流体力学中的物理信息神经网络（PINN）：综述。《力学学报》, 37(12), 1727-1738. Chen, Z., Lai, S.K. and Yang, Z., 2024. AT-PINN：用于结构振动分析的先进时间推进物理信息神经网络。《薄壁结构》, 196, p.111423. Cheng, Y., Zhang, L.L., Li, J.H., Zhang, L.M., Wang, J.H. and Wang, D.Y., 2017. 基于耦合流动–变形模拟的空间随机非饱和土固结。《国际岩土力学数值与分析方法杂志》, 41(5), 682-706. Conte, E., 2004. 非饱和土固结分析。《加拿大岩土工程杂志》, 41(4), 599-612. Cuomo, S., Di Cola, V.S., Giampaolo, F., Rozza, G., Raissi, M. and Piccialli, F., 2022. 基于物理信息神经网络的科学机器学习：现状与未来方向。《科学计算杂志》, 92(3), p.88. Dubey, S.R., Singh, S.K. and Chaudhuri, B.B., 2022. 深度学习中的激活函数：全面综述与基准测试。《神经计算》, *503*, 92-108. Esrig, M.I., 1968. 孔隙压力、固结与电动现象。《土力学与基础工程分部学报》, 94(4),899-921. Fredlund, D.G. and Hasan, J.U., 1979. 一维固结理论：非饱和土。《加拿大岩土工程杂志》, 16(3), 521-531. Glorot, X. and Bengio, Y., 2010, March. 理解深度前馈神经网络训练的难点。发表于第十三届国际人工智能与统计学会议论文集（pp. 249-256）。JMLR研讨会与会议论文集。Goswami, S., Anitescu, C., Chakraborty, S. and Rabczuk, T., 2020. 通过迁移学习增强物理信息神经网络，用于断裂相场建模。《理论与应用断裂力学》, 106, p.102447. Gunes Baydin, A., Pearlmutter, B.A., Andreyevich Radul, A. and Siskind, J.M., 2015. 机器学习中的自动微分：综述。arXiv电子预印本, pp.arXiv-1502. He, Q. and Tartakovsky, A.M., 2021. 用于正向和反向平流–弥散方程的物理信息神经网络方法。《水资源研究》,57(7), p.e2020WR029479.

Ho, L., Fatahi, B. and Khabbaz, H., 2014. 使用特征函数展开法求解非饱和土一维固结的解析解。《国际岩土力学数值与分析方法杂志》, 38(10), 1058-1077. Ho, L., Fatahi, B., 2016. 时变荷载作用下非饱和土的一维固结分析。《国际岩土力学杂志》, 16(2), 04015052. Ho, L. and Fatahi, B., 2018. 等应变条件下考虑涂抹效应的非饱和土层轴对称固结解析解。《国际岩土力学数值与分析方法杂志》, 42(15), 1890-1913. Huang, J., Griffiths, D.V. and Fenton, G.A., 2010. 土体耦合固结的概率分析。《岩土工程与环境岩土工程学报》, 136(3), 417-430. Jagtap, A.D. and Karniadakis, G.E., 2020. 扩展物理信息神经网络（XPINNs）：一种基于时空域分解的非线性偏微分方程通用深度学习框架。《计算物理通讯》, 28(5). Kiyani, E., Shukla, K., Urbán, J.F., Darbon, J. and Karniadakis, G.E., 2025. 哪种优化器最适用于物理信息神经网络和 Kolmogorov-Arnold 网络？*arXiv*预印本 *arXiv:2501.16371*。Kumar, G.S. and Premalatha, K., 2021. 通过数据扰动和统计变换结合三维剪切保护隐私信息。《应用软计算》, 112, p.107819. Lan, P., Su, J.J., Ma, X.Y. and Zhang, S., 2024. 改进的物理信息神经网络在连续排水边界条件下非线性固结问题中的应用。《岩土工程学报》, 19(1), 495-508. Lawal, Z.K., Yassin, H., Lai, D.T.C. and Che Idris, A., 2022. 物理信息神经网络（PINN）的演进及未来发展：系统文献综述与文献计量分析。《大数据与认知计算》, 6(4), p.140. Lloret, A. and Alonso, E.E., 1980. 考虑膨胀和塌陷行为的非饱和土固结。《岩土工程》, 30(4), 449-477. Liu, E., Yu, H.S., Deng, G., Zhang, J. and He, S., 2014. 非饱和土渗流–变形的数值分析。《岩土工程学报》, 9(6), 1045-1058. Liu, Y., Liu, W., Yan, X., Guo, S. and Zhang, C.A., 2023. 面向物理信息神经网络的自适应迁移学习。《计算物理学杂志》, 490, p.112291. Lu, L., Meng, X., Mao, Z. and Karniadakis, G.E., 2021. DeepXDE：用于求解微分方程的深度学习库。《SIAM评论》, 63(1), pp.208-228. Mao, Z. and Meng, X., 2023. 使用基于残差/梯度的自适应采样方法和物理信息神经网络求解具有尖锐解的偏微分方程。《应用数学与力学》, 44(7), 1069-1084. Mattey, R. and Ghosh, S., 2022. 一种用于训练物理信息神经网络求解 Allen-Cahn 方程和 Cahn-Hilliard 方程的新型序贯方法。《计算力学与工程中的方法》, 390, p.114474. Mesri, G. and Rokhsar, A., 1974. 黏土固结理论。《岩土工程分会杂志》, 100(8), 889-904. Qin, A.F., Chen, G.J., Tan, Y.W. and Sun, D.A., 2008. 非饱和土一维固结解析解。《应用数学与力学》, 29(10), 1329-1340. Qin, A., Sun, D.A. and Tan, Y., 2010. 随时间指数变化的荷载作用下非饱和土一维固结解析解。《计算机与岩土工程》, 37(1-2), 233-238. Reddy, J.N., 1993. 有限元法导论。纽约, 27(14). Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., Bengio, Y. Courville, A., 2019, May. 关于神经网络的频谱偏差。发表于国际机器学习会议（5301-5310）。PMLR. Raissi, M., Perdikaris, P. and Karniadakis, G.E., 2019. 物理信息神经网络：用于求解涉及非线性偏微分方程的正向与反问题的深度学习框架。《计算物理学杂志》, 378, 686-707. Scott, R.F., 1963.《土力学原理》，Addison-Wesley。马萨诸塞州雷丁, p.54.

Shan, Z., Ling, D. and Ding, H., 2012年. 单层非饱和土一维固结的精确解。《国际岩土力学数值与分析方法杂志》, 36(6), 708- 722. Sharma, P., Malik, N., Akhtar, N. and Rohilla, H., 2013年. 前馈神经网络：综述。《国际工程与应用科学高级研究杂志》(IJAREAS), 2(10), 25-34. Tartakovsky, A.M., Marrero, C.O., Perdikaris, P., Tartakovsky, G.D. and Barajas-Solano, D., 2020. 用于学习地下流问题中参数和本构关系的物理信息深度神经网络。《水资源研究》, 56(5), p.e2019WR026731. Tang, Y., Taiebat, H.A. and Russell, A.R., 2018年. 考虑水力滞后的非饱和土固结数值模拟。《国际岩土力学杂志》, 18(2), p.04017136. Terzaghi, K., 1943. 固结理论。《理论土力学》, 265-296. Toscano, J.D., Oommen, V., Varghese, A.J., Zou, Z., Ahmadi Daryakenari, N., Wu, C. and Karniadakis, G.E., 2025年. 从物理信息神经网络到物理信息科尔莫戈洛夫-阿诺德网络：物理信息机器学习的最新进展。《计算科学与工程机器学习》, *1*(1), 1-43. Xu, S., Dai, Y., Yan, C., Sun, Z., Huang, R., Guo, D. and Yang, G., 2025年. 论物理信息神经网络的预处理：如何更好地利用流体力学中的数据。《计算物理学杂志》, 528, p.113837. Wang, J., Mo, Y.L., Izzuddin, B. and Kim, C.W., 2023年. 用于固体力学的精确狄利克雷边界物理信息神经网络（EPINN）。《计算力学与工程中的方法》, 414, p.116184. Wang, L., Xu, Y., Xia, X., Li, L. and He, Y., 2020. 非饱和土一维固结的一系列半解析解。《国际岩土力学杂志》, 20(6), 06020005. Wang, S., Teng, Y. and Perdikaris, P., 2021 年 . 理解并缓解物理信息神经网络中的梯度流病态问题。《*SIAM*科学计算杂志》, *43*(5), pp.A3055-A3081. Wang, S., Yu, X. and Perdikaris, P., 2022年. 物理信息神经网络何时以及为何训练失败：神经切线核视角。《计算物理学杂志》, 449, 110768. Wang, Y., Shi, C., Shi, J. and Lu, H., 2024年. 利用物理信息神经网络对二维土体固结进行数据驱动的正反演分析。《岩土工程学报》,19 (12), 8051-8069. Wong, T.T., Fredlund, D.G. and Krahn, J., 1998年. 非饱和土耦合固结的数值研究。《加拿大岩土工程杂志》, 35(6), 926-937. Zhang, P., Yin, Z.Y. and Sheil, B., 2024年. 一种用于固结分析的物理信息数据驱动方法。《岩土工程》, 74(7), 620-631. Zhou, W.H., Zhao, L.S. and Li, X.B., 2014 年 . 非饱和土一维固结的简易解析解。《国际岩土力学数值与分析方法杂志》, 38(8), 794-810. Zhou, W.H., Zhao, L.S., Garg, A. and Yuen, K.V., 2017年. 部分透水边界条件下非饱和土固结的广义解析解。《国际岩土力学杂志》, 17(9), . 04017048. Zhou, W.H., 2013年. 采用微分求积法分析非饱和土的轴对称固结。《工程数学问题》, 2013(1), p.497161. Zhou, W.H. and Zhao, L.S., 2014年. 采用微分求积法分析非饱和土双层系统的固结。载于《2014年岩土工程大会：面向可持续发展的岩土表征与建模》（pp. 3994-4003）。Zhou, H., Wu, H., Sheil, B. and Wang, Z., 2025年. 一种用于大应变固结分析的自适应物理信息神经网络方法。《计算机与岩土工程》, 181, p.107131.