采用 IFENN 的物理信息融合 MIONet：耦合孔弹性问题的通用框架

Fouad M. Amin [a,b], **Rigoberto Moncada** [d], **PanosPantidis** [b], **Diab W.Abueidda** [b], **Fan Fei** [e], **Mostafa E. Mobasher*** [a,b]

[a]土木与城市工程系，坦登工程学院，纽约大学，美国纽约州布鲁克林，邮编 11201[b]土木与城市工程系，纽约大学阿布扎比分校，阿联酋阿布扎比，邮政信箱 129188[c]结构工程系，工程学院，曼苏拉大学，埃及曼苏拉 35516[d]劳伦斯利弗莫尔国家实验室，美国加利福尼亚州利弗莫尔市 East Avenue 7000 号，邮编 94550[e]国家超级计算应用中心，伊利诺伊大学厄巴纳-香槟分校，美国伊利诺伊州厄巴纳

arXiv:2611.31078v1 [cs.CE] 23 Nov 2026

摘要。 采用传统数值求解器进行复杂、高维、大规模孔隙弹性模拟，计算资源成本可能高得令人难以承受。我们提出一种针对孔弹性问题的创新框架，将物理信息机器学习（PI-ML）模型与高性能有限元求解器相结合。该框架通过集成有限元神经网络（IFENN）设置实现，其中有限元求解器负责力学平衡，PI-ML 模型负责流体流动预测，二者采用交错方案。我们提出一种新颖的物理信息融合 MIONet（PIF-MIONet）架构，其中融合网络对 MIONet 分支的输出进行编码。MIONet 对不同的输入函数进行编码，而融合网络可促进输入之间的早期交互，从而支持跨多个场景的泛化，同时遵循孔隙弹性问题的基本物理规律。我们研究了多个不同示例，涵盖二维和三维域，并考虑了可变材料属性和不同的时间离散。我们使用未见过的加载历程对模型进行测试，结果表明模型在不同场景下具有良好的准确度和泛化能力。对于复杂几何和边界条件，我们证明，在损失函数中同时考虑物理规律和数据时，PI-ML 模型仅需纯数据驱动训练所用训练增量的 0.1% 即可完成训练。此外，我们将 PIF-MIONet 扩展为纳入区域分解方法；事实证明，在复杂几何中对流动进行建模且涉及更多局部现象时，该方法尤为有效。IFENN 首次采用开源

用于有限元法部分的多物理场模拟器 GEOS，为该框架应用于工业规模模拟和复杂多物理场问题提供了途径。

关键词： IFENN、物理信息机器学习、孔隙弹性、DeepONet、MIONet。

***通讯作者：** Mostafa E. Mobasher，mostafa.mobasher@nyu.edu

1 引言

1.1 背景与文献综述

孔隙力学是研究力学与流体流动之间相互作用的多物理场领域，涵盖地质力学中的重要问题，例如边坡和钻孔稳定性分析、[1–3] 渗流、[4] 以及固结分析。[5] 这些问题涉及复杂的相互作用，给全面的实验研究带来了困难。数值模拟可通过有限差分法（FDM）、[6,7] 有限体积法（FVM）、[8,9] 和有限元法（FEM）[10,11] 等方法来研究这些问题。尽管如此，这些传统方法也面临自身的挑战，包括时空多尺度现象，这可能导致某些模拟的计算成本高得难以承受。[12–15] 研究人员已提出多种方法来应对这些数值挑战，例如自适应重新网格划分、并行计算、区域分解和交错求解器。[16–19] 然而，改进这些方法并提供更好的替代方案，仍是一个有诸多探索方向的开放研究领域。机器

学习（ML）是目前受到广泛探索的一种途径，旨在应对耦合多物理场模拟中的挑战。本文提出了集成有限元神经网络（IFENN）框架的一种新颖形式，利用创新的机器学习模型加速孔隙弹性模拟。

机器学习方法可分为数据驱动方法，其损失函数仅使用标注数据；以及物理信息方法，其损失函数包含支配物理定律的残差。数据驱动方法的泛化能力可能较差，而物理信息机器学习（PI-ML）能够生成符合物理规律的预测结果，并提升泛化能力。[20] PI-ML 能够融合标注数据与支配物理定律，从而提高数据效率并稳定预测结果。[20,21] 此外，PI-ML 在多孔介质的若干应用中展现出良好的前景，[22] 包括流体流动[23–25] 和孔隙弹性[26–30] 模拟。另一方面，PI-ML 也面临诸多挑战，例如收敛缓慢、对超参数敏感、损失景观复杂、难以扩展至多物理场问题、边界条件施加困难，以及神经网络架构设计方面的挑战。[31,32] 为应对这些挑战，文献中提出了多种技术，包括自适应权重、[33–37] 无量纲形式、[28,38,39] 形函数导数，[40,41] 以及区域分解。[42,43]

聚焦于孔隙弹性，已有多种数据驱动方法和物理信息方法被提出，用于求解反问题和正问题。在数据驱动方法中，Vasilyeva 和 Tyrylgin[44] 使用卷积神经网络（CNN）求解反问题，从局部微观尺度随机场估计宏观尺度性质。Kadeethum 等人[45] 采用条件生成对抗网络（cGAN）处理稳态响应的正问题和反问题。Choi 等人[46] 针对一维（1D）固结问题，采用数据驱动训练研究了不同的网络架构设计。物理信息神经网络（PINNs）已被应用于多个正问题。例如，Bekele[26,27] 使用全连接深度神经网络，预测二维 Barry-Mercer 注采问题和一维固结问题中的位移和流体压力。Haghighat 等人[28] 基于孔隙力学中的应力分裂方法提出了一种序贯训练技术，从而提高了训练收敛性，并将其应用于固结、注采和两相排水问题。Millevoi 等人[29] 探索了将数据纳入损失函数的方法，从而实现了更快的训练速度和更准确的结果。为处理非均质介质，Roy 等人[30] 提出了复合神经网络（CoNN），其中每个均质子域使用不同的网络。最近，Ren 等人[47] 通过处理观测到的系统响应，识别材料参数，从而解决了一个反问题。

尽管上述研究作出了重要贡献，数据驱动方法限制了训练后的模型的泛化能力[20]，并且需要大型数据集。[48] 另一方面，采用 PINNs 的研究[26–30] 存在两个主要局限。首先，这些研究没有探索三维空间域，也未研究地质材料孔隙弹性建模中的复杂边界条件。其次，所有提出的 PINNs 都是在固定加载情景下训练的。输入仅包含时空坐标，无法编码其他输入场，例如应变场、可变材料属性或随机边界条件。这些局限凸显了对先进机器学习架构和训练技术的需求，使其能够处理复杂几何形状、多种输入和随机加载历程，同时遵循孔弹性问题的基本物理规律。

随着机器学习的发展，基于算子的架构应运而生，在泛化能力和表达能力方面更具优势。[49–52] 深度算子网络（DeepONet）[52] 便是其中一种架构，在学习不同算子方面展现出良好的潜力。经典的（原始）DeepONet由

两个网络组成：一个分支用于编码输入函数，另一个主干用于编码输出点的坐标。分支和主干的输出通过点积运算进行组合，从而生成 DeepONet 的输出。DeepONet 架构是一种高层次的模块化设计，分支或主干有多种可选设计，包括多层感知机（MLP）、卷积神经网络（CNN）、循环神经网络（RNN）和 Kolmogorov-Arnold 网络（KAN）。[52,53] DeepONet 已衍生出许多变体，例如面向高维问题的可分离物理信息 DeepONet（Sep-PI-DeepONet）。[54] 另一种变体是多输入深度神经算子（MIONet），[55] 它通过多个分支处理多个输入函数。MIONet 架构已应用于多个领域，包括结构动力响应、[56] 相场水力压裂、[57] 多相流建模、[58] 以及孔隙弹性。[46]

本文介绍了一种物理信息融合 MIONet（PIF-MIONet），将融合概念[46,59–62] 融入 MIONet。PIF-MIONet 继承了 MIONet 对多个输入函数进行编码的能力，其中包括可变材料属性和随机注采通量历史。融合组件使输入变量能够及早交互。此外，主干和分支的输出采用与标准 DeepONet 相同的方式进行聚合，[52] 使 PIF-MIONet 成为许多现有模型的适宜扩展。PIF-MIONet 主要包含两个方面，即 Fusion-MIONet 架构和物理信息训练过程。为提高训练稳定性和效率，我们在 PIF-MIONet 中研究并整合了自适应权重、无量纲形式、形函数导数和区域分解。

尽管纯基于机器学习的方法具有较高的计算效率，但仍面临若干局限，包括可靠性[63,64]、泛化能力[65]、可解释性有限[66,67]、对超参数的敏感性[63,68]、训练稳定性[37,69] 以及网络收敛[70,71] 等方面的担忧。为弥合这一差距，一种颇具前景的研究趋势应运而生：针对不同问题类型，以不同方式融合机器学习与经典方法的混合方法。对于单物理场应用，已有若干方法被提出，包括用于动态载荷和碰撞仿真的 FEMIN[72] 和 PI-DeepONet-FEM[73]，以及 HINTS、[74] ANN-FEA、[75] 和 NN-PGD。[76] 对于耦合多物理场应用，混合方法包括 NN-Nested-FEM、[77] NeuroSEM、[78] PDNN、[79]以及集成有限元神经网络（IFENN）。[80]

IFENN 是一个通用框架，可为耦合多物理场问题提供高效、准确的仿真。IFENN 的核心思想是结合有限元法和预训练神经网络来求解控制偏微分方程。按照混合交错方案，有限元求解器通常负责力学场平衡，而神经网络则近似求解耦合场。两个求解器在每个荷载增量或时间增量处交换信息，从而在整个分析过程中实现紧密耦合。之所以采用这种设计（即用有限元法处理力学场、用神经网络处理耦合场），是因为力学响应更为复杂且包含多个分量，而另一耦合场通常是标量场。因此，鉴于机器学习模型的可靠性有限，一种审慎的做法是使用神经网络进行标量场预测，并使用有限元求解器处理复杂的力学响应。

IFENN 框架已成功应用于多种应用场景，包括损伤[41,80] 和相场断裂建模、[81,82] 热弹性、[48,83,84] 以及孔隙弹性。[48]这些研究各自存在一定局限性，通常与训练难题或泛化能力有关。对于孔隙弹性，IFENN 目前仅针对二维问题进行了研究，并采用大型数据集开展数据驱动的 DeepONet 训练。尽管 PI-ML 具有良好的应用潜力，

鉴于本综述中上述各种挑战，其在孔隙弹性地质力学问题中的应用仍存在研究空白。具体而言，广义PI模型在孔隙弹性中的可行性及其在二维和三维空间域中的实现，同时考虑时间和材料变化，仍未得到充分探索。本文提出 PIF-MIONet，解决其训练难题，并将其集成到 IFENN 中，以适用于二维和三维域，同时探索更强的泛化能力以及与现实问题的相关性。

1.2 范围与大纲

在本研究中，我们提出物理信息融合 MIONet（PIF-MIONet），作为流体质量平衡方程的代理求解器，并将其集成到 IFENN 中，用于求解耦合孔弹性问题。我们首次提出了这一创新集成方案，主要贡献如下：

- PIF-MIONet 训练采用自适应权重、无量纲形式和形函数导数。
- 我们将空间域分解融入 PIF-MIONet，在保持训练参数数量相近的同时，提高了其在复杂几何中的准确度和训练稳定性。
- 该框架可同时考虑多种时间增量大小、材料属性、通量历史和网格分辨率。
- IFENN 集成利用了开源有限元模拟器 GEOS，[85] 提升了混合有限元法与机器学习方法的可扩展性和可移植性，并在学术研究与工业规模模拟之间架起桥梁。

本文其余部分的结构如下：第2节概述孔弹性问题的数学公式，随后第3节详细介绍孔弹性问题的 IFENN 框架实现。第4 节和第5 节分别详细介绍所提出的 PIF-MIONet 设计与训练。第6节通过数值示例探讨训练性能和 IFENN 框架的能力。第7 节讨论并总结本文的主要结论。

2 孔弹性问题的数学公式

2.1 强形式

在孔弹性问题中，力学变形与流体在多孔介质中的流动相互作用并彼此影响。对于由固体颗粒和充满流体的孔隙组成的区域，耦合两个变量——位移（$\boldsymbol{u}$）和流体压力（p）——的控制方程可写为：[5, 86–89]

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}_t + \boldsymbol{b} = \boldsymbol{0} \tag{1}$$

$$\dot{\zeta} + \boldsymbol{\nabla} \cdot \boldsymbol{v}_f = Q_f \tag{2}$$

其中， ∇ 是梯度算子， $\nabla\cdot$ 是散度算子， $\boldsymbol{\sigma}_t$ 是由变形和流体压力引起的总柯西应力， $\boldsymbol{b}$ 是机械体力向量， ζ 是流体含量增量， $\boldsymbol{v}_f$ 是流体速度向量， Q_f 表示任意流体汇或源。

边界条件定义在以下边界上 $\Gamma = \Gamma_u \cup \Gamma_t = \Gamma_p \cup \Gamma_q$，如图1所示：

$$\begin{aligned} \boldsymbol{u} &= \overline{\boldsymbol{u}}, \quad x \in \Gamma_u, \\ \boldsymbol{t} &= \bar{\boldsymbol{t}}, \quad x \in \Gamma_t, \\ p &= \overline{p}, \quad x \in \Gamma_p, \\ q &= \overline{q}, \quad x \in \Gamma_q \end{aligned} \tag{3}$$

其中，$\overline{\boldsymbol{u}}$ 和 $\bar{\boldsymbol{t}}$ 分别为定义在边界 Γ_u和 Γ_t上的位移边界条件和牵引边界条件。 p 为边界 Γ_p上的流体压力。 q 为边界Γ_q上的流体通量，定义为 $q = \boldsymbol{v}_f \cdot \boldsymbol{n}$，其中 $\boldsymbol{n}$ 为边界 Γ_q处的单位法向量。

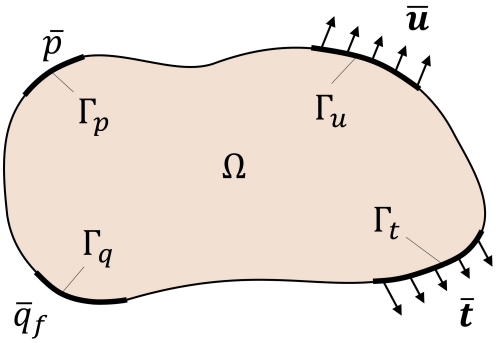


图 1 具有指定边界条件的孔隙弹性问题区域示意图

根据流体流动的达西定律，流体速度（$\boldsymbol{v}_f$）和流体含量增量（ζ）定义为：[5, 87–89]

$$\boldsymbol{v}_f = -K_H \nabla H_f \tag{4}$$

$$\zeta = \frac{p}{M} + \alpha \mathrm{tr}(\boldsymbol{\varepsilon}) \tag{5}$$

其中， K_H为各向同性水力传导系数，定义为 $K_H = \gamma_f K_p / \mu_f$； H_f为流体总水头，定义为 $H_f = p/\gamma_f + \boldsymbol{x} \cdot \boldsymbol{i}_g$。 K_p 为固有渗透率， γ_f 为流体单位重， μ_f 为流体动力黏度， $\boldsymbol{x}$ 为坐标向量， $\boldsymbol{i}_g$ 为平行于重力且方向与重力相反的单位向量。值得注意的是，根据式（3）和式（4）的定义，注入通量q为负值。Biot系数（α）定义为 $\alpha = 1 - K_b/K_s$，其中K_s 为固体颗粒体积模量， K_b 为固体骨架的体积模量。Biot模量（M）定义为 $1/M = \phi_b/K_f + (\alpha - \phi_b)/K_s$，其中 ϕ_b为介质的孔隙率， K_f 为流体体积模量。体积模量（K_b）可定义为 $K_b = \lambda + 2\mu/3$。

总应力 $\boldsymbol{\sigma}_t$ 可表示为固体有效应力 $\boldsymbol{\sigma}'$ 和流体压力 $\boldsymbol{\sigma}_f$的函数：

$$\begin{aligned} \boldsymbol{\sigma}_t &= \boldsymbol{\sigma}' + \boldsymbol{\sigma}_f \\ &= \boldsymbol{C} : \boldsymbol{\varepsilon} - \alpha p \boldsymbol{I} \end{aligned} \tag{6}$$

ˆ其中 $\boldsymbol{C}$ 是弹性张量，其定义为 Lam´e 常数（λ 和 μ）的函数，对于各向同性线性材料，有 $C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$。 $\boldsymbol{\varepsilon}$ 是应变，其定义为位移（$\boldsymbol{u}$）的函数，即 $\boldsymbol{\varepsilon} = \boldsymbol{\nabla}^s\boldsymbol{u} = \frac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{\nabla}\boldsymbol{u}^T)$，适用于小变形。将式（6）代入式（1），并将式（4）和式（5）中的定义代入式（2），可得时间离散化后的线性动量平衡方程和流体流动连续性方程：

$$\boldsymbol{\nabla}\cdot(\boldsymbol{C}:\boldsymbol{\varepsilon}_{n+1} - \alpha p_{n+1}\boldsymbol{I}) + \boldsymbol{b}_{n+1} = \boldsymbol{0} \tag{7}$$

$$\frac{p_{n+1} - p_n}{\Delta t M} + \alpha\,\mathrm{tr}\left(\frac{\boldsymbol{\varepsilon}_{n+1} - \boldsymbol{\varepsilon}_n}{\Delta t}\right) - \boldsymbol{\nabla}\cdot\left(\frac{K_p}{\mu_f}\cdot(\boldsymbol{\nabla}p_{n+1} + \gamma_f\boldsymbol{i}_g)\right) = Q_f \tag{8}$$

其中，下标 n 和 $n+1$ 表示两个连续的时间步，时间增量为Δt。该时间离散采用隐式欧拉格式，其中 $\dot{\boldsymbol{\varepsilon}} = (\boldsymbol{\varepsilon}_{n+1} - \boldsymbol{\varepsilon}_n)/\Delta t$，且$\dot{p} = (p_{n+1} - p_n)/\Delta t$。

2.2 弱形式与有限元求解器

给定线性动量平衡方程 Eq. (7) 和流体流动连续性方程 Eq. (8) 的最终强形式，相应的时间离散弱形式可分别写为：

$$\begin{aligned}\mathcal{D}_{n+1} = &\int_\Omega \boldsymbol{\nabla}^s\widehat{\boldsymbol{w}} : \hat{\boldsymbol{C}} : \boldsymbol{\varepsilon}_{n+1}\,d\Omega \;-\; \int_\Omega \boldsymbol{\nabla}^s\widehat{\boldsymbol{w}} : (\alpha p_{n+1}\boldsymbol{I})\,d\Omega \\ &- \int_{\Gamma_t}\widehat{\boldsymbol{w}}\cdot\bar{\boldsymbol{t}}_{n+1}\,d\Gamma \;-\; \int_\Omega \widehat{\boldsymbol{w}}\cdot\boldsymbol{b}_{n+1}\,d\Omega = \boldsymbol{0}, \quad \forall\,\widehat{\boldsymbol{w}} \in \mathcal{W}_u\end{aligned} \tag{9}$$

$$\begin{aligned}\mathcal{H}_{n+1} = &\int_\Omega \widehat{p}\,\frac{p_{n+1} - p_n}{\Delta t M}\,d\Omega \;+\; \int_\Omega \widehat{p}\alpha\,\mathrm{tr}\left(\frac{\boldsymbol{\varepsilon}_{n+1} - \boldsymbol{\varepsilon}_n}{\Delta t}\right)d\Omega \;+\; \int_\Omega \boldsymbol{\nabla}\widehat{p}\cdot\left(\frac{K_p\boldsymbol{I}}{\mu_f}\cdot\boldsymbol{\nabla}p_{n+1}\right)d\Omega \\ &+ \int_\Omega \boldsymbol{\nabla}\widehat{p}\cdot\left(\frac{K_p\boldsymbol{I}}{\mu_f}\cdot\gamma_f\boldsymbol{i}_g\right)d\Omega + \int_{\Gamma_q}\widehat{p}\,\bar{q}_{n+1}\,d\Gamma \;-\; \int_\Omega \widehat{p}\,Q_{f\,n+1}\,d\Omega = 0, \quad \forall\,\widehat{p} \in \mathcal{W}_p\end{aligned} \tag{10}$$

其中，$\widehat{\boldsymbol{w}}$ 和 $\widehat{p}$ 分别为位移和压力试函数，$\mathcal{W}_u$ 和 $\mathcal{W}_p$ 分别为位移和压力函数空间。

孔隙弹性模拟可能会出现压力振荡，[90] 可通过采用计算成本较高的双二次位移/双线性压力四边形单元来缓解（参见图 2中的 U2P1 单元）。另一种高效的方案是采用带稳定化弱形式的双线性位移/双线性压力四边形单元（参见图 2中的 U1P1 单元）；该弱形式由 White 和 Borja 提出，[90] 如下所示：

$$\mathcal{H}_{n+1} + \mathcal{H}^{stab}_{n+1} = 0 \tag{11}$$

$$\mathcal{H}^{stab}_{n+1} = \int_\Omega \frac{\tau}{2\mu\Delta t}\left(\widehat{p} - \Pi\widehat{p}\right)\left((p_{n+1} - p_n) - \Pi\left(p_{n+1} - p_n\right)\right)\,d\Omega \tag{12}$$

其中，$\mathcal{H}^{stab}_{n+1}$ 是添加到流体流动连续性方程中的时间离散化稳定化项，τ 是用于调节稳定化程度的常数乘子。投影算子 Π 计算 ρ^h 在单元域 Ω^e上的单元平均值，如$\Pi\rho^h|_{\Omega^e} = \frac{1}{V^e}\int_{\Omega_e}\rho^h\,d\Omega^e$。

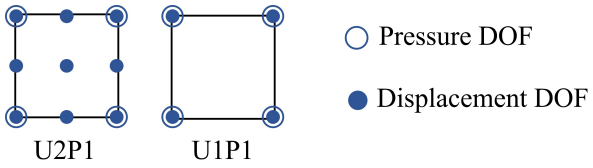


图 2 混合四边形单元示意图：二次位移/线性压力（U2P1）和线性位移/线性压力（U1P1）

对于方程（9）和方程（10所描述的孔隙弹性问题，本文采用 U2P1 单元并利用 deal.II[91] 有限元法库构建了一个全耦合有限元法整体求解器。该求解器用于生成数据，作为训练和测试期间误差评估的真实参考值。为研究效率，本文在第6.1.4节开发了两种采用 U2P1 和 U1P1 的交错求解器。值得注意的是，本文所有数值示例中均未检测到虚假压力振荡。

3 方法

3.1 IFENN 框架用于孔弹性问题

IFENN 框架已针对文献中的许多应用进行了介绍和讨论，[41, 80–84] 包括孔隙弹性。[48] 简而言之，IFENN 在概念上类似于交错混合格式，其中力学场采用传统有限元求解器进行求解，而另一个耦合场则使用预训练神经网络进行预测。图 3 展示了本研究针对孔弹性问题所采用的 IFENN 框架实现方式，该实现使用了预训练的 PIF-MIONet。在每个时间增量内，求解器会执行一组子循环。在每个子循环中，PIF-MIONet 与有限元求解器依次调用，以计算未知状态变量。关于子循环方案的更多详细信息见第 3.3节。

如式（7）和式（8）所示，位移 $\boldsymbol{u}$ 与压力 p 通过应变迹率 $\mathrm{tr}(\dot{\boldsymbol{\varepsilon}})$耦合。因此，预测压力会与有限元求解器共享，而应变迹（作为计算所得位移的后处理变量）则会与 PIF-MIONet 共享。有限元法代码通过 C++ 脚本运行，PIF-MIONet 则通过 Python 脚本运行。Python 层与 C++ 层之间的通信通过操作系统（OS）共享内存实现，如图 3所示。层间同步通过信号量实现；信号量充当协调机制，可阻止一个进程，直到另一个进程发出其已完成操作的信号。

3.2 IFENN 中有限元求解器的集成

在本研究中，我们展示了 IFENN 的适用性，并采用了两个在有限元法层面运行的不同平台：基于 deal.II 的求解器[91] 和 GEOS，[85] 如图 3所示。两个求解器均可部署于高性能计算环境，适用于大规模问题。基于 deal.II 的求解器具有更高的灵活性，便于针对广泛的目标应用进行开发和测试。另一方面，GEOS 是一款高性能模拟器，专用于地下流动和地质力学。此外，它还提供一套预构建的物理求解器，可轻松实现复杂的工业规模问题。展示 IFENN 与 GEOS 求解器结合使用的适用性是一项重要进展，因为这推动 IFENN 从纯粹面向研究的软件包迈向生产规模地质力学模拟的部署。这一进展有助于未来的扩展以及

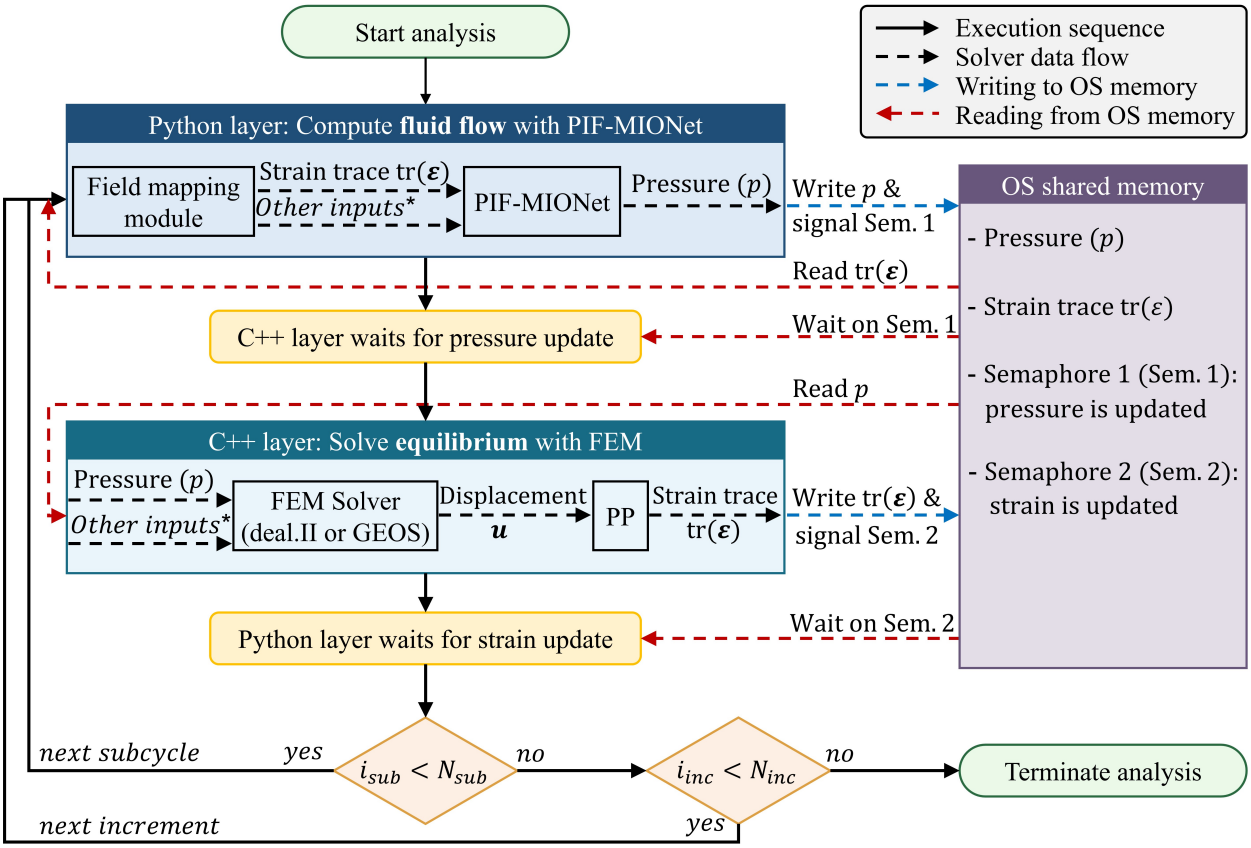


图3 IFENN 在孔弹性问题中的实现。Python 层与 C++ 层通过操作系统（OS）共享内存交换压力 p 和应变迹 tr(ε)。应变通过对位移进行后处理（PP）计算。两层之间的同步通过信号量实现，信号量用于指示压力和应变数据已更新、可供读取。索引 i_{sub} 是子循环计数器，子循环总数为 N_{sub}；i_{inc} 是时间增量计数器，时间增量总数为 N_{inc}。*其他输入 包括求解器所需的其余信息，例如材料属性、时间增量和载荷。场映射模块根据输入离散化情况有条件地启用。

集成其他机器学习模型。在第 6 节所示的数值示例中，根据求解器对问题的适用性或研究便利性，使用了这两种求解器。

对于 deal.II 集成，PIF-MIONet 在单元顶点处预测压力场，如图 4所示。压力值写入共享内存后，信号量会通知有限元求解器读取节点压力值，并使用双线性形函数将其插值到高斯积分点。随后，有限元求解器计算位移响应，并对其进行后处理，以计算单元节点处的应变。计算所得的应变值写入共享内存后，信号量会通知 PIF-MIONet 读取共享的应变数据并开始新的子循环。这种方法利用相邻单元之间共有的顶点，减少了需要与有限元求解器共享的压力评估次数。另一种方法是通过 PIF-MIONet 算子直接在高斯点处评估压力。但这需要在求解器之间共享更多数据，从而增加 IFENN 框架部署的内存需求。

PIF-MIONet 仅使用节点输入和输出进行训练。然而，现有的 GEOS

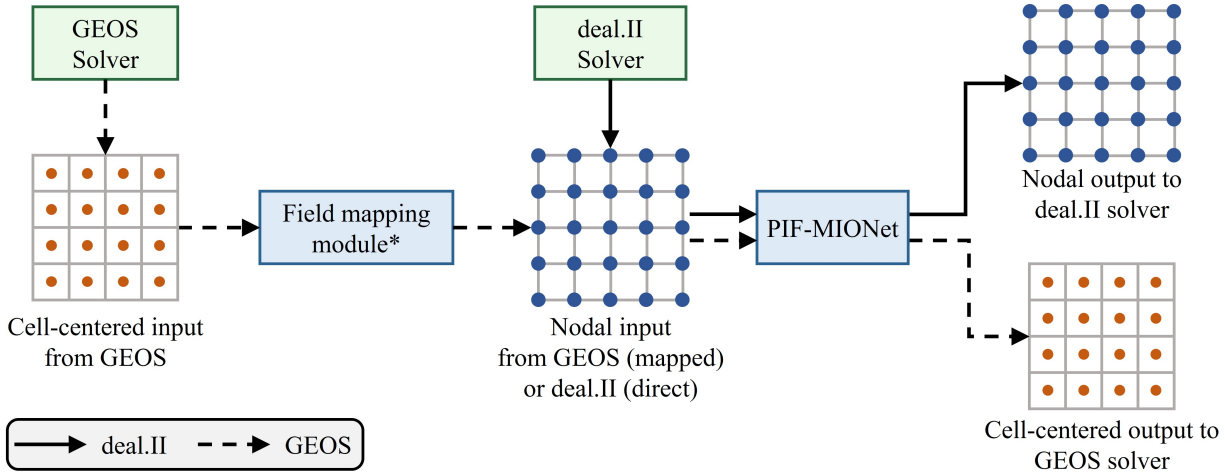


图 4 IFENN 框架的 Python 层中，PIF-MIONet 输入场和输出场空间离散的示意图。有限元求解器 deal.II 和 GEOS 分别以节点形式和单元中心形式交换数据。*场映射模块 仅在与 GEOS 集成时需要，此时输入场从单元中心传输到单元节点。PIF-MIONet 可直接在单元节点或单元中心提供输出。

力学求解器旨在与 GEOS 原生的基于有限体积法的流动求解器共享并接收单元中心值。为集成 IFENN，基于有限体积法的流动求解器被一个新定制的求解器所取代，该求解器作为接口与 Python 层通信，以交换压力和应变迹场。由于 PIF-MIONet 与 GEOS 使用不同的空间离散方式，因此在这两个层之间添加了一个场映射模块。该模块集成在 Python 层中，将 GEOS 计算得到的应变迹映射到 PIF-MIONet 所需的节点位置，如图 4 所示。对于内部节点，通过插值相邻单元中心值来计算其数值；对于边界节点，则直接将最近单元中心的值赋给该节点。为兼容 GEOS，压力直接在单元中心处求值，并与 GEOS 求解器共享。值得注意的是，在与 deal.II 和 GEOS 集成时，PIF-MIONet 分别用于预测节点和单元中心处的压力。因此，该实现为 IFENN 可移植性以及 PIF-MIONet 的灵活性提供了实用的测试。

3.3 预测器-多重校正器子循环方案

孔隙弹性本质上是一个耦合问题，在本文中，我们采用预测子-多校正子方案，具体如算法1所述。在每个新的增量（记为 $n+1$）中，p_{n+1}和 ε_{n+1} 均为未知量，这在两个变量之间产生了循环依赖（即，式（ p_{n+1}和 ε_{n+1} 7）以及式（8）中的）。因此，应对其中一个变量的初始值进行猜测，以打破这一循环，从而使交错方案得以推进。在我们的实现中，我们首先为应变率设定一个合理的初始猜测（预测阶段）。随后，首先调用 PIF-MIONet 计算压力 p_{n+1}，然后由有限元求解器计算 ε_{n+1}。接着执行若干个子循环（校正阶段）以校正这些预测结果，子循环次数为预设值。对于每次仿真，每个时间增量的循环均预先设定固定的子循环次数 N_{sub}，将其作为终止准则。或者，也可以在仿真过程中根据适当的收敛准则自动确定 N_{sub}。

算法 1：总计 N_{sub} 个子循环的预测-多重校正子循环方案

1 开始新的时间增量（从时间步 n 推进至 $n+1$） 2 设置子循环索引的初始值 $r=0$**3 while** $r<N_{sub}$ **do 4 if** $r=0$**then 5 预测子循环：设置应变率的初始猜测。6 if** $n=0$**then**7 $\dot{\varepsilon}=0$**8 else**9 $\dot{\varepsilon}=\left(\varepsilon_n-\varepsilon_{n-1}\right)/\Delta t$**10 end 11 else 12 校正子循环：设置应变率** $\dot{\varepsilon}=\left(\varepsilon_{n+1}^{r-1}-\varepsilon_n\right)/\Delta t$**13 end 14 使用 PIF-MIONet 预测压力** p_{n+1}^r**（该网络经过训练以遵循式（8）中的强形式） 15 使用有限元求解器求解式（9）以获得位移，然后计算** ε_{n+1}^r**16** $r=r+1$**17 end**

4 物理信息融合 MIONet（PIF-MIONet）的设计

PIF-MIONet 具有一组参数 ϕ，其训练目标是最小化损失函数 $\mathcal{L}(\phi)$，如下所示：

$$\phi^*=\arg\min_{\phi}\mathcal{L}(\phi) \tag{13}$$

其中，ϕ^* 表示使预定义损失函数 $\mathcal{L}(\phi)$ 达到最小值的优化后的模型参数集合。本节将详细介绍所采用的损失函数和 PIF-MIONet 架构。

4.1 损失函数

损失函数的一般形式可写为：

$$\mathcal{L}(\phi)=\sum_{i=1}^{N_{terms}}\lambda_i\mathcal{L}_i \tag{14}$$

其中，N_{terms} 为所考虑的损失项总数，λ_i 是分配给每个损失项（$\mathcal{L}_i$）的自适应权重，用于平衡各个损失项的贡献，并在训练过程中动态引导优化器的关注重点。未加权的损失项 $\mathcal{L}_i$ 可计算为：

$$\mathcal{L}_i=\mathcal{N}(S_i\boldsymbol{R}_i),\qquad i=1,\ldots,N_{terms}. \tag{15}$$

其中，$\boldsymbol{R}_i$ 是在预定义的评估点/配点集合上计算得到的损失项残差向量，S_i 是用于归一化残差幅值的标量缩放系数，而

$\mathcal{N}(\cdot)$ 是范数计算函数，例如二范数或均方值。公式（14）中的损失函数详细形式可通过将下标 i 替换为P、 D、 N和 d，分别表示偏微分方程、狄利克雷边界条件、诺伊曼边界条件和数据损失项，从而定义为：

$$\mathcal{L}(\phi) = \lambda_P \mathcal{N}(S_P \boldsymbol{R}_P) + \sum_{j=1}^{N_{\Gamma_p}} \lambda_D^j \mathcal{N}(S_D \boldsymbol{R}_D^j) + \sum_{j=1}^{N_{\Gamma_q}} \lambda_N^j \mathcal{N}(S_N \boldsymbol{R}_N^j) + \lambda_d \mathcal{N}(S_d \boldsymbol{R}_d) \tag{16}$$

其中 N_{Γ_p} 以及 N_{Γ_q} 分别表示狄利克雷边界条件和诺伊曼边界条件的数量。

值得注意的是，本研究中的所有示例均为纯物理信息型，不包含数据贡献（$\lambda_d = 0$），但有一个示例引入了数据项以进行正则化并增强训练。我们测试了两种自适应加权算法——GradNorm[34] 和 ReLoBRaLo，[33]用于训练 PIF-MIONet，具体实现细节见附录A.1。关于所提出的残差和缩放系数计算方法的更多细节，分别见第4.2 节和第4.3 节。

4.2 偏微分方程残差与边界条件的施加

各残差项定义如下：

$$\boldsymbol{R}_P = \frac{\widetilde{p}_{n+1} - p_n}{\Delta t M} + \alpha \mathrm{tr}\left(\frac{\boldsymbol{\varepsilon}_{n+1} - \boldsymbol{\varepsilon}_n}{\Delta t}\right) - \boldsymbol{\nabla} \cdot \left(\frac{K_p}{\mu_f} \cdot (\boldsymbol{\nabla}\widetilde{p}_{n+1} + \gamma_f \boldsymbol{i}_g)\right) - Q_f, \qquad \boldsymbol{x} \in \Omega \tag{17}$$

$$\boldsymbol{R}_D^j = \widetilde{p}_{n+1} - \overline{p}_{n+1}, \qquad \boldsymbol{x} \in \Gamma_p^j \tag{18}$$

$$\boldsymbol{R}_N^j = \left(-\frac{K_p}{\mu_f} \boldsymbol{\nabla}(\widetilde{p}_{n+1} + \gamma_f \boldsymbol{x} \cdot \boldsymbol{i}_g)\right) \cdot \boldsymbol{n} - \overline{q}_{n+1}, \qquad \boldsymbol{x} \in \Gamma_q^j \tag{19}$$

$$\boldsymbol{R}_d = \widetilde{p}_{n+1} - \check{p}_{n+1}, \qquad \boldsymbol{x} \in \Omega \tag{20}$$

其中 $\boldsymbol{R}_P$ 是通过域配点（$\boldsymbol{x}$）计算得到的偏微分方程残差向量， $\widetilde{p}_{n+1}$ 是关注坐标处的模型预测。 $\boldsymbol{R}_D^j$ 和 $\boldsymbol{R}_N^j$ 分别是边界 Γ_p^j 和 Γ_q^j上狄利克雷和诺伊曼边界条件软惩罚项的残差。 $\boldsymbol{R}_d$ 是标记数据误差，其中引入了来自预定义节点的参考数据，以提高模型精度。 $\check{}p_{n+1}$ 是通过耦合有限元求解器计算得到的真实参考响应。在此方法中， $\boldsymbol{R}_P$ 采用连续性方程的强形式进行公式化，见式（8）。文献中的其他方法基于变分法和能量方法，或结合强形式与弱形式。[84,92]

4.3 无量纲缩放

无量纲形式已应用于许多机器学习应用中，包括流体流动公式，[28,38,39] 主要有两个目的：(1) 归一化模型输入和输出；(2) 缩放残差方程以控制误差幅值。如果在训练前进行归一化过程，却未由训练软件进行跟踪，则应引入新的残差形式，以考虑归一化的影响。本文中，训练软件会跟踪所有归一化操作，并将其对梯度计算的影响纳入考虑。因此，无需对第4.2节所述的残差形式作进一步修改。因此，我们只关注这一过程的第二部分：缩放残差。

为此，按照文献中讨论的流程，[28,38,39] 我们将缩放系数确定为：

$$
\begin{aligned}
S_P &= \frac{\mu_f^* x^{*2}}{K_p^* p^*}, \\
S_D &= S_d = \frac{1}{p^*}, \\
S_N &= \frac{\mu_f^* x^*}{K_p^* p^*}
\end{aligned}
\tag{21}
$$

其中，S_P、S_D、S_d和 S_N分别是残差 $\boldsymbol{R}_P$、$\boldsymbol{R}_D$、$\boldsymbol{R}_d$和 $\boldsymbol{R}_N$的缩放因子。μ_f^*、x^*、K_p^*和 p^*分别是流体黏度、区域坐标、渗透率和压力的归一化系数。这些系数大致将对应变量归一化至 1。例如，x^* 设为区域中的最大坐标，p^* 设为训练增量中压力幅值的最大值。

4.4 Fusion-MIONet 架构

我们先简要说明所提架构的各个组成部分。最简单形式的 DeepONet（见图5a）由两个网络组成：（1）对输入函数信号进行编码的分支网络 F_I；（2）对目标位置进行编码的主干网络x_O，并在这些位置评估输出函数。[52,93] DeepONet 通过算子(F_I) 将离散化的输入函数(F_O) 映射为输出函数 $\mathcal{G}$，其采用点积过程，具体如下：

$$
\mathcal{G}(F_I)(x_O) = \sum_{k=1}^{O_{to}} b_k(F_I) t_k(x_O) + b_0 \tag{22}
$$

其中，$b(F_I)$ 和 $t(x_O)$ 分别是维数为 O_{to} 的分支和主干输出向量，$b_0 \in \mathbb{R}$是一个偏置。值得注意的是，在这一设置中，F_I 被定义为区域内固定位置处离散值构成的向量。输出查询坐标（x_O）可以任意选取，因此使用训练好的 DeepONet 时具有很强的泛化能力。

DeepONet 设置的一种扩展是多输入深度神经算子（MIONet），[55]它使用多个分支对多个输入函数进行编码，如图 5b 所示。在 PI-ML 中处理多个输入尤其具有挑战性，特别是当每个输入都具有各自的幅值范围、维数及时空分布时。使用 MIONet 可以提供一种可行的解决方案；然而，其原始设计限制了输入源之间的早期数据交换/交互。这一局限性可以通过融合来解决，即将多个源的数据整合为单一表示。[59,60] 近期研究已将融合概念引入基于 DeepONet 的数值模拟中，[61,62] 采用不同技术在分支网络与主干网络之间融合数据。

在本研究中，如图 5c 所示，我们在 MIONet 架构中提出了一种独特的融合方法，以实现输入场之间的早期交互。首先，每个分支将其输入函数 F_{Ii} 编码为输出向量 B_i。然后，将所有分支的输出向量拼接成一个向量 $U_I = \{B_1;\ B_2;\ .....\ B_n\}$，对应 n 个输入函数。最后，U_I 作为输入传入融合网络，由其将所有分支的信息编码为一个向量 (类似于 $b(F_I)$ 中的

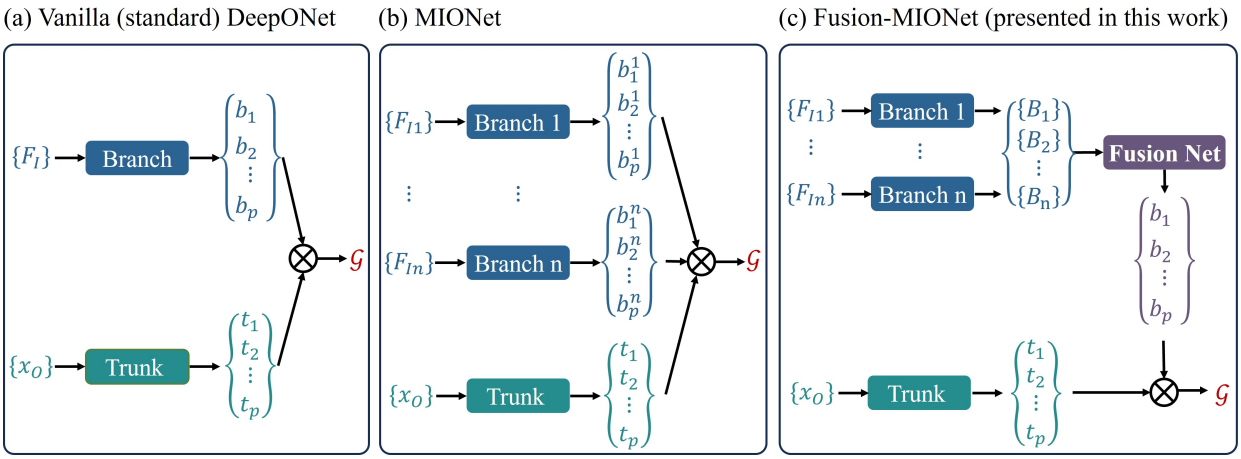


图 5 DeepONet 的不同变体：(a) 原始 DeepONet，(b) MIONet，(c) 我们提出的架构（Fusion-MIONet）

原始 DeepONet)。因此，Fusion-MIONet 的另一项优势是，最终输出可以采用式（22）中给出的标准 DeepONet 方法进行聚合。

所有分支、融合和主干网络均采用多层感知机（MLP）。对于输入向量为 $\mathbf{z}_i \in \mathbb{R}^{1\times N_i}$ 的多层感知机层，其输出向量 $\mathbf{z}_{i+1} \in \mathbb{R}^{1\times N_{i+1}}$ 可计算为[94]$\mathbf{z}_{i+1} = \sigma(\mathbf{z}_i\mathbf{W}_i + \mathbf{b}_i)$，其中 σ 是激活函数，$\mathbf{W}_i \in \mathbb{R}^{N_i\times N_{i+1}}$ 是权重参数，$\mathbf{b}_i \in \mathbb{R}^{1\times N_{i+1}}$ 是偏置向量。值得注意的是，Choi 等人近期的一项研究采用数据驱动方法，专门针对固结问题，考察了与 Fusion-MIONet 类似架构的有效性。

我们提出物理信息融合 MIONet（PIF-MIONet），将其作为流体质量平衡求解器，并纳入控制底层偏微分方程物理和数值行为的参数。图 6展示了我们提出的广义形式的 PIF-MIONet。它针对离散化的连续性方程（式（8））， 其中输出为压力(p_{n+1})，输入则包括通量 (q_{n+1})、渗透率(K_p)、时间增量 (Δt)、应变迹 tr 的变化量$(\Delta\varepsilon_{n+1})$ $=$ tr$(\varepsilon_{n+1} - \varepsilon_n)$、前一时间步的压力 (p_n)以及输出查询坐标 $(\boldsymbol{x})$。

如图 6所示，主干输入有两个维度，$[N_x, N_c]$，其中 N_x 是输出查询点的数量，N_c 是空间坐标分量的数量（$N_c = 2$ 或3）。主干输入被编码为大小为 $[N_x, O_{to}]$ 的输出张量。所有分支输入都有两个维度：第一个是 N_b，即采样时间步的批次大小；第二个是空间离散输入点的数量。例如，N_p 是压力场的空间离散输入点数量。各分支的输出被拼接为一个大小为$[N_b, N_{fi}]$的张量，再通过 MLP 融合网络进行编码，最终得到一个大小为 $[N_b, O_{to}]$ 的张量，其最后一个维度与主干输出的维度相同（即 O_{to}）。最后一个维度相同对于式（22）中的点积运算至关重要，该运算会生成大小为 $[N_b, N_x]$ 的最终压力输出张量。

在本研究的所有增量中保持不变的量，例如流体动力黏度 μ_f，可以从网络输入中省略，并直接并入偏微分方程残差中，如式（17）所示。在不同增量间变化但空间上均匀的量，例如通量、渗透率和时间步长，需要大小为 1的分支输入，因此 $N_q = N_K = N_t = 1$。

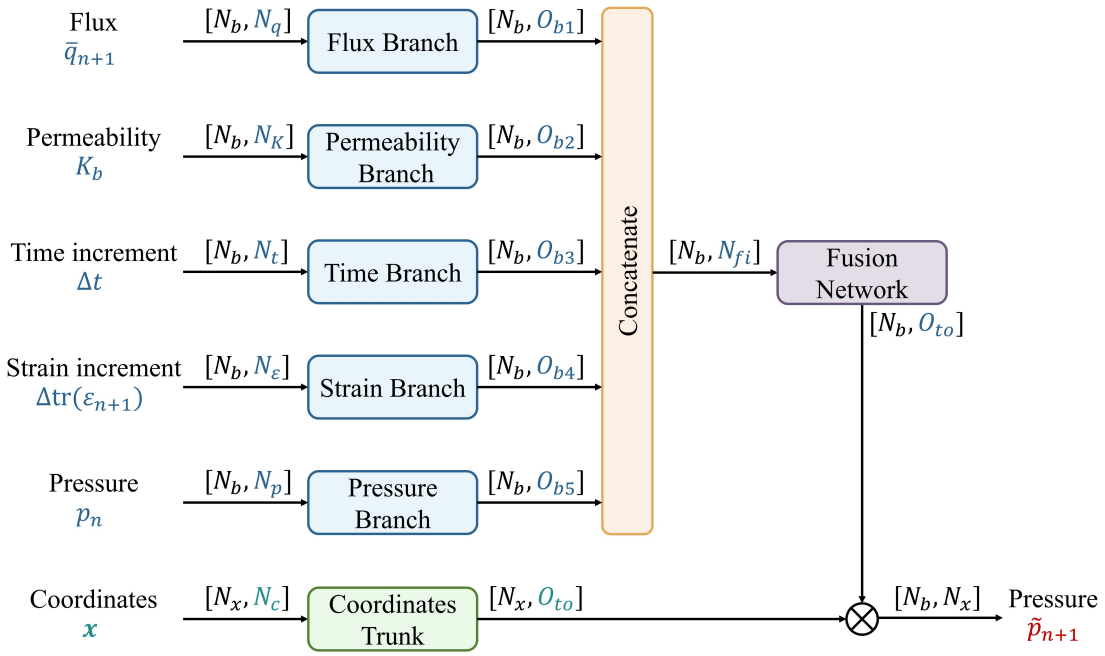


图 6 考虑可变渗透率 K_p 和时间增量 Δt的广义 PIF-MIONet。 N_b 为批次大小，N_x 为输出点数量。融合网络的输入尺寸等于 $N_{fi} = O_{b1} + O_{b2} + O_{b3} + O_{b4} + O_{b5}$。主干与融合网络的输出尺寸相同，为 O_{to}。

在残差评估期间，每个标量都会广播到所有配点。图 6所示的上述架构设计，是为第 6 节介绍的广义示例量身定制的。对于渗透率 K_p 和时间增量 Δt 保持不变的其他示例，我们引入了一个简化版 PIF-MIONet，如图7所示。

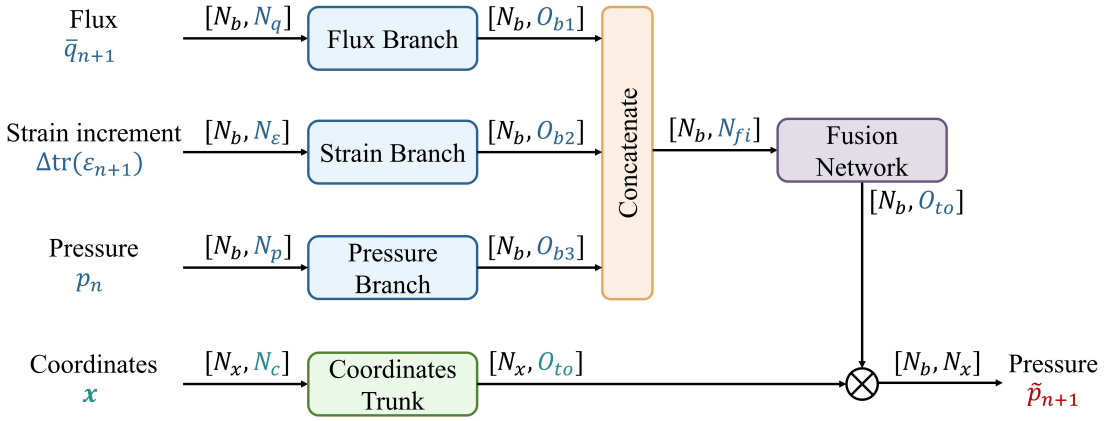


图 7 考虑恒定渗透率 K_p 和时间增量 Δt的简化 PIF-MIONet。 N_b 为批次大小，N_x 为输出点数量。融合网络的输入尺寸等于 $N_{fi} = O_{b1} + O_{b2} + O_{b3}$。主干与融合网络的输出尺寸相同，为 O_{to}。

对于空间变化的输入函数（例如应变或压力），输入函数由在节点（称为传感器）处采样的数值表示。[52] 每个分支对单个输入函数进行编码，并可使用其自身的采样或投影方法。[55] 输入

节点的数量和分布因此可能因输入函数而异。计算物理信息偏微分方程残差需要获取配点处的所有场值。当这些场值直接由分支输入函数获得时，每个空间变化函数的节点集必须包含所有配点。在此要求下，配点构成最小输入集，同时允许增加输入点以获得更高的空间分辨率。

在本研究中，我们在分支架构中加入了组归一化。我们此前的研究[48]表明，在采用 IFENN 的 DeepONets 中，对输入场随仿真逐步推进而出现误差累积的分支应用组归一化是有效的。有关此方法的更多细节，请参见附录A.2。

4.5 区域分解

若干方法采用区域分解来提高训练效率。[42,43,95] 如图8所示，我们通过引入多个主干网络，将区域分解融入 PIF-MIONet 架构。区域 Ω 被分解为一组较简单的子域 $\{\Omega_1, \ldots, \Omega_n\}$。对于每个子域 Ω_k，专用主干网络会对坐标进行编码，并据此预测该子域内的压力。受空间域分解连续性考量的启发，[43]我们通过对重叠区域的预测结果取平均，确保公共子域界面上的连续性。每个子域的输出均与权重张量逐元素相乘，然后将加权后的输出组合起来，形成总输出压力场 p_{n+1}。我们在组装全局压力场之后计算物理信息损失所需的空间导数，而不是针对每个子域分别计算。这样可以确保偏微分方程残差基于组装后的压力表示进行评估，其中包括子域界面。

权重张量能够实现高效的向量化计算和训练所需的梯度计算。每个子域都有一个专用的权重张量，其内部节点的权重为 1，1/2 接口处的权重为 1，$1/m_s$ 角点处的权重为 1，其中 m_s 表示重叠子域的数量。这种区域分解方法简化了每个主干的任务，使其能够专注于所分配的子域，而无需为整个区域设置单一且复杂的主干。第6.4节展示了该方法对复杂几何形状的有效性。

5 物理信息Fusion-MIONet（PIF-MIONet）的训练

本文所展示的所有示例均为纯物理信息示例，不包含数据损失项（即 $\lambda_d = 0$公式（16）中的 $\lambda_d = 0$6.4节的一个示例除外，该示例引入了数据损失以提高训练准确度和稳定性。值得注意的是，即使省略数据损失项，我们仍然利用全耦合有限元求解器的解来生成物理上保持一致的网络输入，例如压力场和应变场，我们将这些输入统称为训练数据。本节将介绍训练数据采样、所采用的误差指标以及空间导数的计算。

5.1 符合物理规律的输入场与稀疏增量采样

训练 PIF-MIONet 需要为压力分支和应变分支提供物理一致的输入场，如图 6和图 7所示。当训练涵盖完整的加载历史时，每个加载增量处预测的场都可用作后续增量的输入，从而在整个加载过程中进行顺序传播。然而，在本研究中，我们采用计算效率较高的稀疏采样策略，直接从全耦合有限元模拟中获取选定加载增量处的训练输入场。

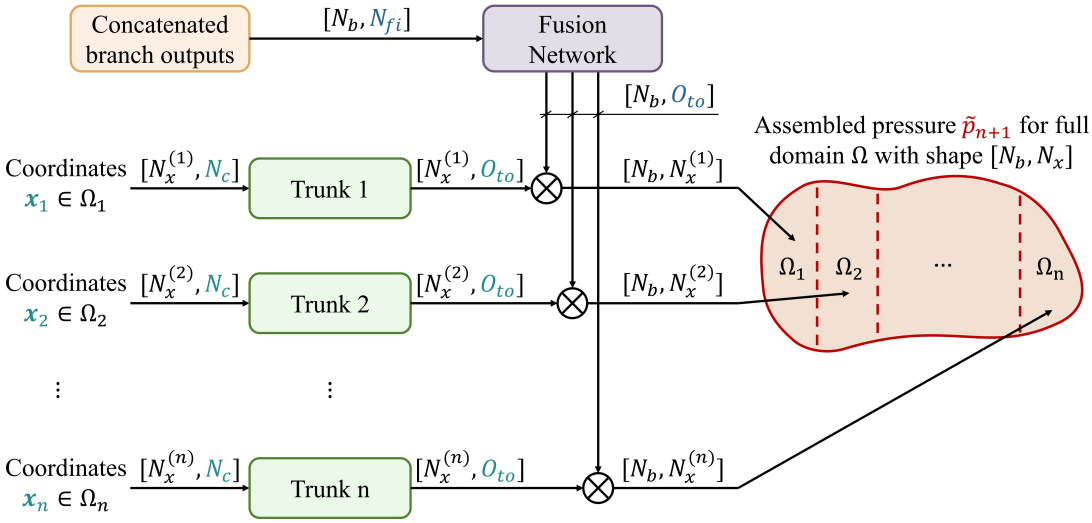


图 8 采用区域分解以高效利用主干的 PIF-MIONet。区域 Ω 被分解为不同的子域 $\{\Omega_1, \Omega_2, ...\Omega_n\}$，每个主干负责编码特定子域的坐标。利用各主干的输出计算相应子域的压力输出。

为此，对于特定问题，我们使用有限元求解器生成多个载荷历史的响应数据。生成的输入/输出历史被分为三组：训练、验证和测试，每个载荷历史仅分配给其中一组。对于每一组，通过半随机过程采样少量增量。具体而言，设定若干增量索引，通常在时间域内均匀分布。然后，对于每个索引，从多个随机选取的载荷历史中采样输入/输出数据。

在此采样技术中，我们使用不连续的稀疏增量进行训练，而不是使用完整的连续历史记录。该策略降低了训练成本，并能够更高效地覆盖更广泛的模拟范围。此外，这一特性凸显了 PIF-MIONet 的灵活性：它可以使用稀疏的时间增量而非完整历史记录进行训练。最后，对于 IFENN 的实现，我们使用来自测试组的完整历史记录对该框架进行测试，模型优化器在训练过程中未见过这些记录。

5.2 误差指标

在训练期间，针对训练增量和验证增量分别计算损失指标 $\mathcal{L}(\phi)$，以监控并检测过拟合。完成训练阶段后，按如下方式计算不同的测试误差：

$$\epsilon_{rel} = |y_{true} - y_{pred}|/(|y_{true}| + \epsilon_{tol}) \tag{23}$$

$$L_2^{inc} = \|y_{true} - y_{pred}\|_2 / \|y_{true}\|_2 \tag{24}$$

$$L_2^{all} = \left\|\{L_2^{inc=1}, L_2^{inc=2}, ..., L_2^{inc=N_{inc}}\}\right\|_2 / \sqrt{N_{inc}} \tag{25}$$

其中，y_{true} 是被视为真实参考值的全耦合有限元响应，y_{pred} 是预测响应，ϵ_{rel} 是用于在选定时间步进行可视化比较的逐点相对误差，$\epsilon_{tol} = 10^{-5}\max(|y_{true}|)$ 是数值稳定性的容差。L_2^{inc}用于衡量单个时间增量上的全域误差，而 L_2^{all}表示所有采样增量上的聚合误差。

5.3 使用有限元法形函数（*SF*）计算空间导数

在 PI-ML 中，通常使用自动微分（AD）计算损失项所需的空间导数，从而提供一种直接的实现方式，尤其适用于无网格方法。[20,96] 然而，自动微分可能导致无界误差、[97] 较高的计算成本，[96] 以及不稳定性，[40] 因此促使人们采用有限元法形函数（SF）导数，以实现更加一致、高效且准确的计算。[40,41] 基于形函数的导数需要建立网格，以对神经网络输出进行空间离散；然而，这不会给 IFENN 引入额外的网格划分要求，因为该框架已经包含了用于有限元求解器的离散网格。

式（17）和式（19）中的残差形式表明，需要计算一阶和二阶空间导数。为此，我们采用双二次四边形单元公式来计算导数。据作者所知，这是首次在 PI-ML 框架下，将九节点单元形函数用于算子学习中的空间导数计算。我们利用了空间离散具有结构化特征这一事实，所有预测坐标均匀分布。每个 3×3 二维空间中的节点集均被分组，以构成九节点宏单元的响应向量。节点按字典序排列，以便高效执行向量化张量积运算。有关所采用有限元公式的更多详情，请参见附录 A.3。值得注意的是，对于本文中的三维示例，计算空间导数时仅使用自动微分。

6 数值示例

本节将展示二维和三维示例，以体现所提框架的准确度和泛化能力。所有示例可分为两类主要问题：（1）具有注入通量的孔弹性柱，在二维和三维域中进行测试，并采用不同的材料属性；（2）具有复杂边界条件、受排水通量作用的二维开挖问题。注入问题完全基于物理信息，不包含任何数据损失项；排水问题则纳入数据损失项，以处理区域的几何复杂性。

对于每个示例，均按照第 5 节所述训练 PIF-MIONet，并将其集成到 IFENN 中进行测试。使用 deal.II 和 GEOS 作为力学场求解器的框架分别称为 IFENN-deal.II 和 IFENN-GEOS。IFENN 在线运行只需训练后的模型和目标荷载历史。全耦合有限元数据（y_{true}）仅用于监测运行时误差和评估测试指标（式（23）–式（25））。除非另有说明，IFENN-deal.II 的子循环次数为 $N_{sub} = 2$，IFENN-GEOS 的子循环次数为 $N_{sub} = 1$。除非另有说明，所有结果均采用国际单位制单位，位移单位为 m，压力单位为 N/m^2, 。以下各节中的所有模拟均在 Ubuntu 24.04 工作站上完成，该工作站配备 128 GB 内存、Intel Xeon W3-2435 处理器（8 个物理核心）以及具有 12 GB 专用显存的 NVIDIA RTX A2000 GPU。为确保性能比较一致，有限元法和 IFENN 模拟均使用 8 个进程运行。

6.1 恒定渗透率下的二维注入示例 K_p 和时间增量 Δ t

6.1.1 问题设置

第一个示例是一个受到注入通量作用的二维孔弹性柱体。如图9a所示，上边界（Γ^{top}）完全排水（$p = 0$），并承受大小为 $\sigma = 2000$ N/m^{2}的机械载荷。下边界（Γ^{bottom}）在z方向受到机械约束（$u_z = 0$），并受到

注入通量（$q\ (t)$）$^{-}$，该通量沿边界均匀分布，并随时间随机波动。左右边界（Γ^{left} 和 Γ^{right}）不透水，并在x方向受到机械约束（$u_x = 0$）。

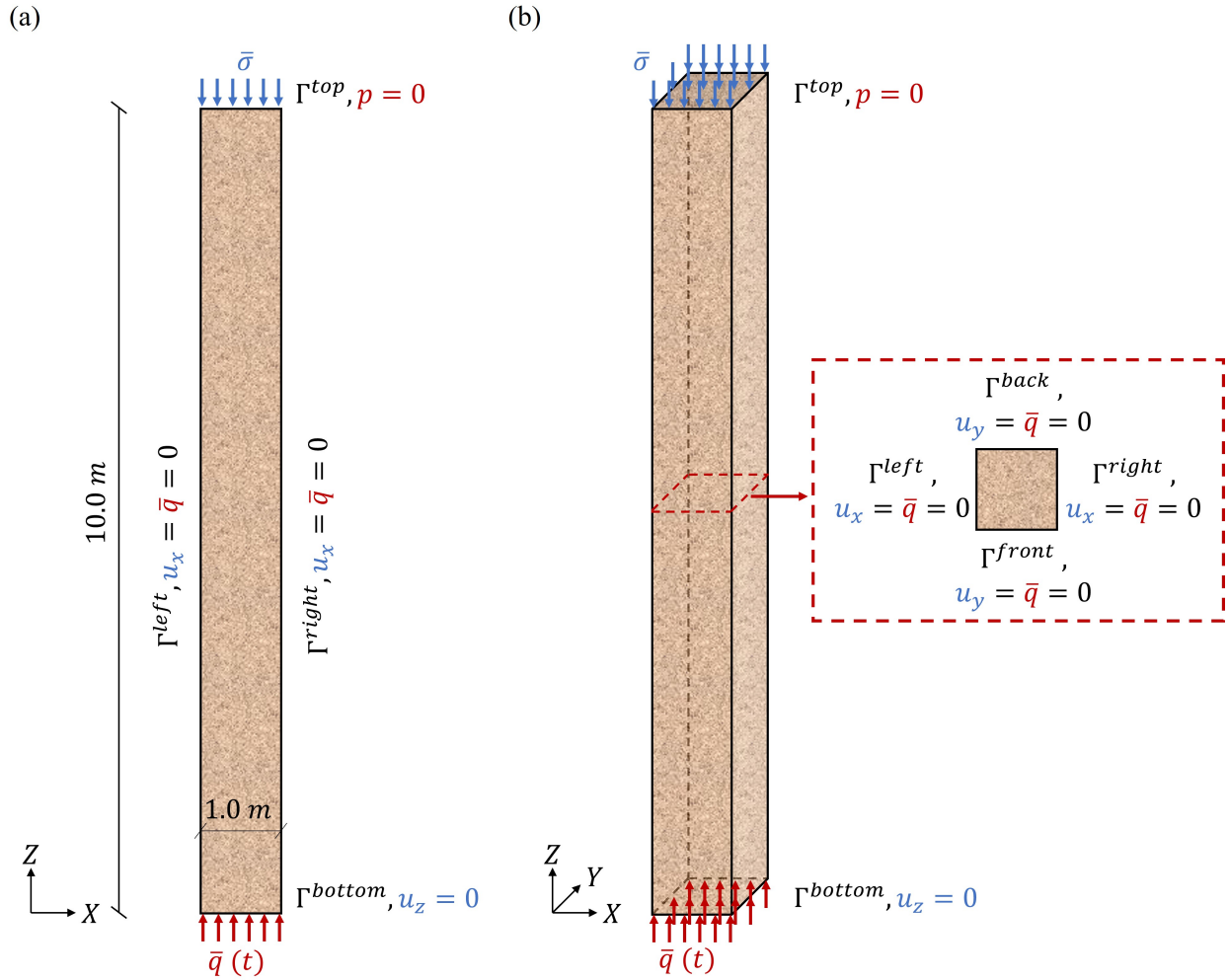


图 9 几何形状及边界条件示意图：(a) 二维注入问题；(b) 三维注入问题，横截面为宽度 1.0 米的正方形。

本例不考虑重力；因此，该问题的行为主要由机械载荷 σ 和注入通量 $q(t)$驱动。鉴于这两种驱动力，此问题设置适合用于研究水力-力学耦合作用，这些作用已被纳入多项研究，包括对多孔介质数值稳定性[98] 和损伤的研究。[99] 当注入通量为零时，该问题可简化为文献中所述的经典太沙基固结问题。[90]

受文献中不同示例的启发，[48,100] 孔隙弹性介质的属性设定为$K_s = 55556$ 兆帕、$\lambda = 8.375$ 兆帕、 $\mu = 5.58$ 兆帕、 $K_f = 2200$ 兆帕、 $\phi_b = 0.4$、 $\mu_f = 1.0 \times 10^{-3}$帕·秒，以及 $K_p = 1.0 \times 10^{-13}$ 米^2。波动注入通量的最大振幅设为 5.0×10^{-6} 米/秒，仿真时长设为 5.0×10^{7} 秒，并分为50个时间步，增量步长为 $\Delta t = 1.0 \times 10^{6}$ 秒。

6.1.2 PIF-MIONet 训练

如第5.1节所述，PIF-MIONet采用的训练/采样策略要求数据集包含物理一致的输入场。为此，将由100组高斯随机注入通量历程及其全耦合有限元响应组成的数据集划分为70 组训练、20 组验证和10组测试载荷工况（LC）。增量按照第5.1节所述方式采样。训练使用可用的3500个时间增量中的72个（$\approx 2\%$），验证则使用1000个中的24个。

训练期间，将采样增量处的输入场输入 PIF-MIONet，同时优化器最小化基于物理的损失函数。本问题采用纯物理信息训练，包含 5 个损失项，不包含数据损失项。第一个损失项，$\mathcal{L}_1^{pde}$，根据偏微分方程残差项 $\boldsymbol{R}_P$计算，该残差项定义于式（17）中。项 $\mathcal{L}_2^{top}$ 是边界 Γ^{top} （R_D 在式（18）中）上狄利克雷边界条件的残差。 $\mathcal{L}_3^{bottom}$通过在 q（t）上施加注入通量 作为诺伊曼边界条件来满足 Γ^{bottom} （R_N 在式（19）中）。最后，考虑 $\mathcal{L}_4^{left}$ 和 $\mathcal{L}_5^{right}$ 分别满足 Γ^{left} 和 Γ^{right}上的零诺伊曼边界条件。

为确定最优训练配置并找出当前问题中效果不佳的方法，我们开展了广泛的参数研究，考察了多种训练变体，详见附录 B.1.1。研究结果表明，采用形函数导数（SF）、GradNorm 自适应权重（GN）、无量纲缩放（DS），并使用 L2 范数 N（。） $= \|.\|_2$ 计算各损失项的范数时，效果最佳。这些配置称为“SF-GN-DS-L2”。

训练后的模型采用图 7 所示的架构。选取区域内等间距节点构成的结构化网格（$41 \times 5 = 205$ 个点），作为应变场和压力场的输入节点。表 1 列出了各多层感知机网络的输入/输出尺寸、隐藏层和组归一化的详细信息，总计包含 4592 个可训练参数。使用 Adam 优化器和自定义学习率调度器进行40000 轮训练，总训练时间为 35 分钟。关于自定义学习率的更多详情见附录 B.1.2。

表 1 二维注入示例训练的 PIF-MIONet 模型的超参数

网络	输入尺寸	隐藏层尺寸[a]	N_{ch}[b]	输出尺寸
分支 #1	N_q =1	[8,8]	-	8
分支 #2	N_ε =205	[8]	4	8
分支 #3	N_p =205	[8]	4	8
融合	N_{fi} =24	[16,16]	-	8
主干	N_c =2	[8,8]	-	8

[a] 尺寸数组的长度即为隐藏层的数量。[b] 在输入层的输出上使用 N_{ch} 个通道进行组归一化。“-”表示不进行组归一化。

图10 展示了未加权的训练损失项及其自适应权重随训练轮次的变化。图中显示了 GradNorm 如何自适应地更新权重，以确保所有损失项均衡下降。所有损失项均从初始值下降了三个数量级以上，突显了采用 SF-GN-DS-L2 的训练设置的效率。我们还

测试了另一种训练方案，即使用 ReLoBRaLo (RB) 代替 GradNorm（即训练组合 SF-RB-DS-L2）。结果显示训练不稳定性较高，且误差值更大。关于采用 SF-RB-DS-L2 进行训练的更多细节，见附录 B.1.3。

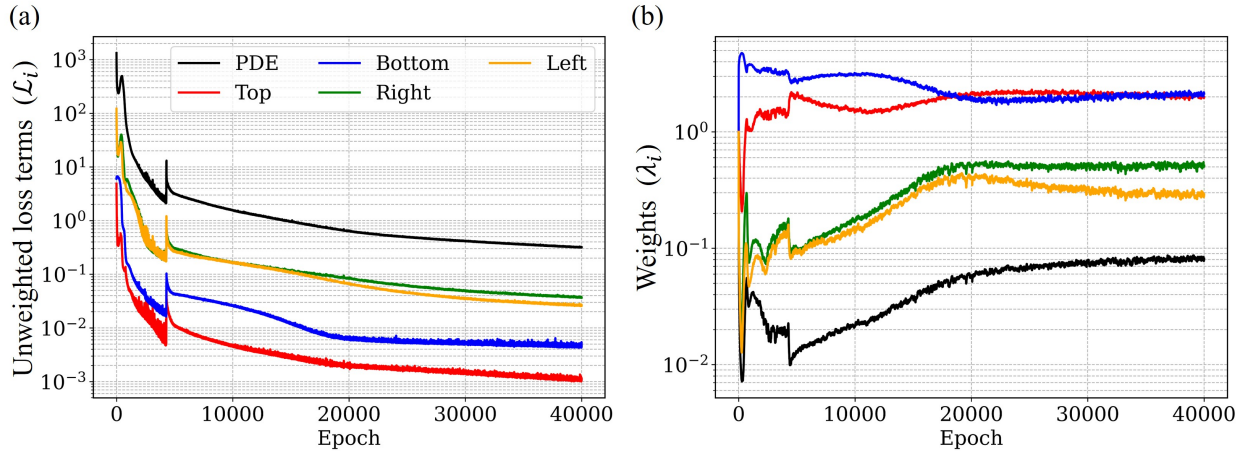


图 10 二维注入示例的训练指标：(a) 未加权损失项 $\mathcal{L}_i$，(b) 对应的自适应权重 λ_i。仅包括偏微分方程残差和边界条件损失；未使用数据损失。

6.1.3 IFENN 的准确度与可移植性

模型在稀疏随机采样的时间步上训练后，便集成到 IFENN- deal.II 框架中，并在 10 种训练优化器未见过的载荷工况（LC）上进行测试，如图 11所示。根据问题设置，响应实际上是一维的；因此，我们只关注压力场（p）和 z方向位移（u_z）。图12 展示了使用全耦合有限元求解器（$\|p\|_2$）和 IFENN-deal.II（$\|u_z\|_2$）获得的压力场（y_{true}）和位移场（y_{pred}）在每个时间增量上的范数。图中结果显示准确度稳定，所有测试工况下，IFENN-deal.II 的响应都与全耦合求解器的响应非常接近。此外，在所有时间增量和载荷工况下，压力和位移响应均随注入通量的变化而变化（见图 11）。

为进一步评估准确度和框架的可移植性，使用 IFENN-GEOS 重复进行了模拟。为此，采用了一个兼容 GEOS 的三维网格，该网格在面外方向仅包含单层单元，并施加适当的边界条件以实现平面应变条件。图 13显示了测试误差（L_2^{inc}定义见公式（24）），其中分别针对 p 和 u_z 场，计算了载荷工况 LC1-LC5 下 IFENN-deal.II 和 IFENN-GEOS 响应的测试误差。误差图显示误差幅值较低，范围从 1×10^{-2} 到1×10^{-4}，且两个场（p 和 u_z）中的误差相对接近。载荷工况 LC6-L10 的误差表现出一致的规律，如附录图 B.4所示。

误差度量的波动 L_2^{inc}可能与注入通量的波动有关$q(t)$。当注入通量的幅值处于最低水平时，真实压力范数 $\|y_{true}\|_2$ 会降至接近于零的数值，从而放大 L_2^{inc}的相对误差计算（见公式（24））。需要强调的是，这些结果来自一个仅基于 72 个非连续时间增量、完全采用物理信息方式训练的网络，损失函数中未使用任何数据，并且是通过此前未见过的连续时间序列进行评估的。

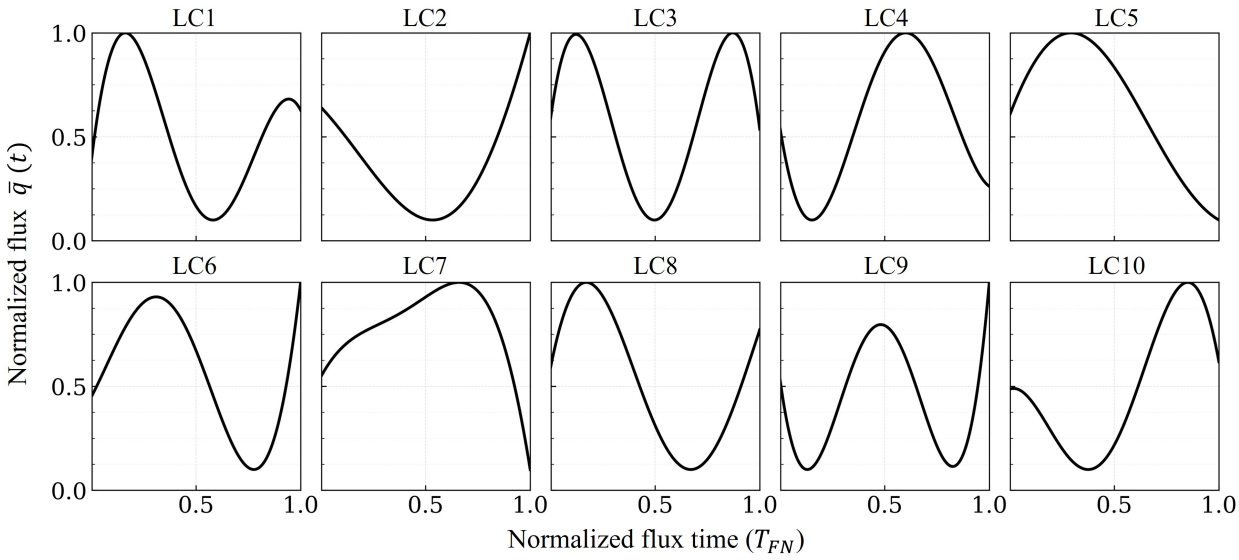


图 11 注入问题所有测试载荷工况（LC）的归一化通量幅值与归一化通量时间（T_{FN}）的关系。通量和时间分别以各自的最大值进行归一化。

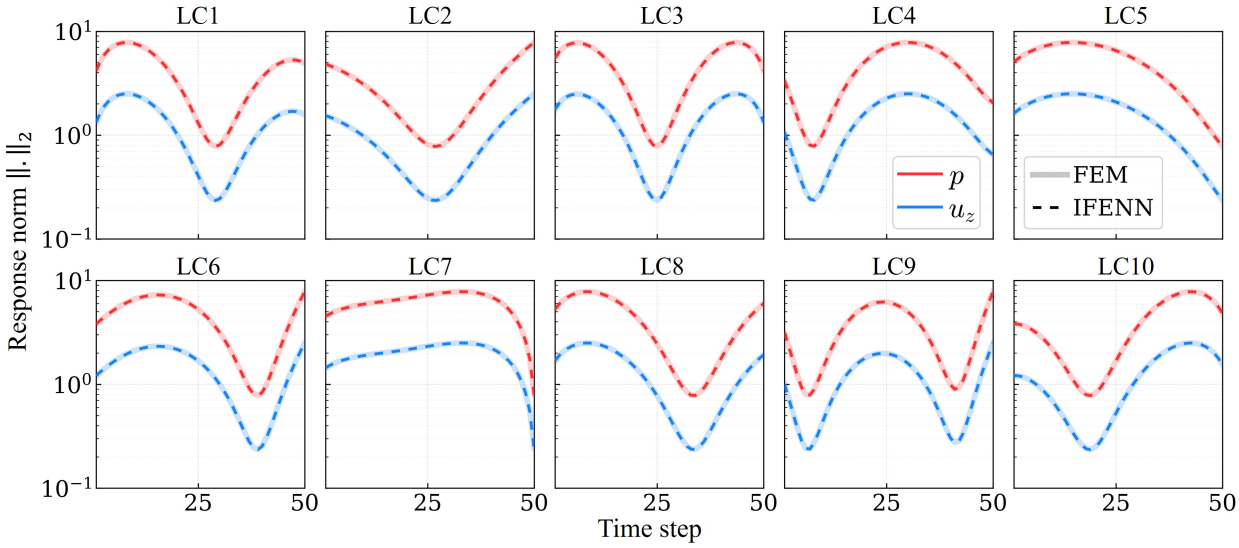


图 12 二维注入问题中，使用全耦合有限元求解器和 IFENN-deal.II 得到的压力范数 $\|p\|_2$ 和位移范数 $\|u_z\|_2$。压力和位移的响应范数单位分别为兆帕和米。

除了误差幅值较低之外，IFENN-deal.II 和 IFENN-GEOS 的误差也几乎完全相同。一个经过训练的 PIF-MIONet 与两个不同的求解器集成后，得到了几乎相同的响应，凸显了 IFENN 框架的可移植性。附录B.1.4中提供了更多结果，进一步展示 IFENN-deal.II 和 IFENN-GEOS 所达到的准确度与可移植性。

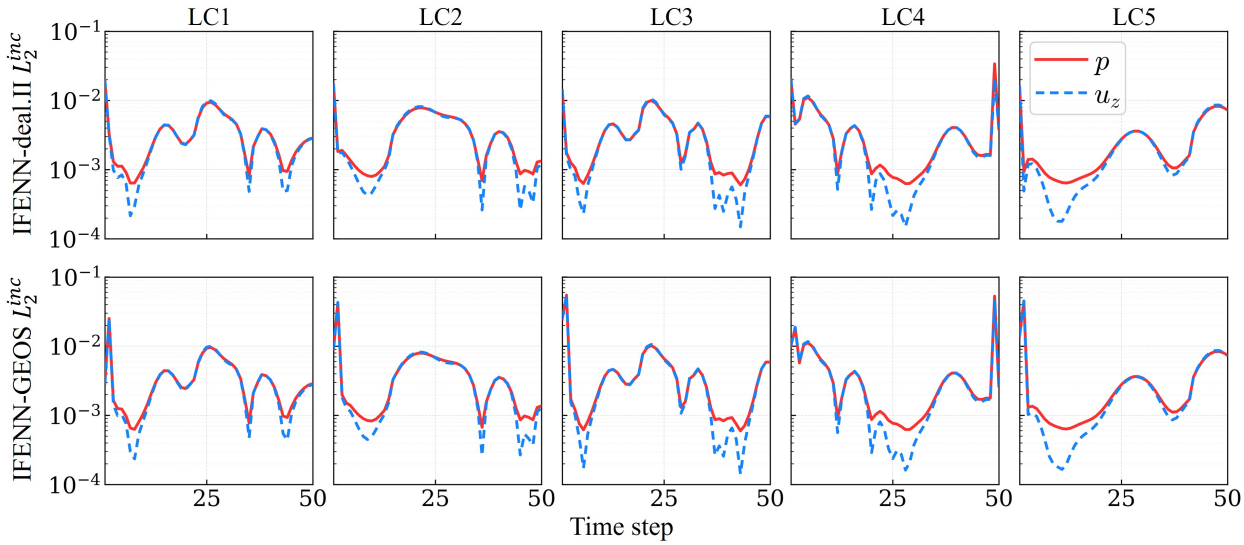


图 13 二维注入问题中，采用 IFENN-deal.II 和 IFENN-GEOS 时，测试载荷工况 LC1-LC5 的压力 (p) 和 z 方向位移 (u_z) 的测试误差 L_2^{inc}。

6.1.4 IFENN 效率

文献中已针对包括热弹性[48] 和相场断裂建模[81] 在内的不同问题研究过 IFENN 框架的效率。本文首次研究 IFENN 框架在孔弹性问题中的效率。由于孔弹性问题对数值有限元法稳定性有要求，本研究与以往工作有所不同。如第 2.2节所述，可使用双二次位移/双线性压力（U2P1）单元，或使用带稳定化算法的双线性位移/双线性压力（U1P1）单元进行适当的有限元分析。

在本研究中，我们实现了两个交错有限元求解器。具体而言，一个求解器针对未稳定化弱形式（式（9）和式（10））采用 U2P1 单元实现，另一个求解器针对稳定化弱形式（式（9）和式（11））采用 U1P1 单元实现。我们针对不同网格分辨率，将 U1P1-IFENN 的效率与这两种方案进行比较。结果见图 14，图中展示了归一化计算时间与区域尺寸的关系。归一化计算时间（T_{CN}）按总计算时间除以（$N_{inc} \times N_{DOF}$）计算，其中 N_{inc} 为时间增量数，N_{DOF} 为自由度数。

结果表明，对于较小的区域尺寸，IFENN 每个自由度的计算成本与有限元求解器相当。在较小区域尺寸下性能相近，可能是因为 IFENN 框架中用于计算运行时误差的调试功能带来了额外计算开销。然而，随着网格细化，IFENN 框架的性能优于有限元求解器。凭借这种持续提升的性能，与其他求解器相比，IFENN 展现出良好的可扩展性潜力，尤其适用于大规模工业问题。

值得注意的是，IFENN 的计算时间不包括 PIF-MIONet 的训练时间。模型只需训练一次，之后便可用于任意数量的模拟，并可采用不同的通量历史和网格细化设置。因此，随着所需模拟数量的增加，尤其是使用细网格时，训练时间相较于所需的

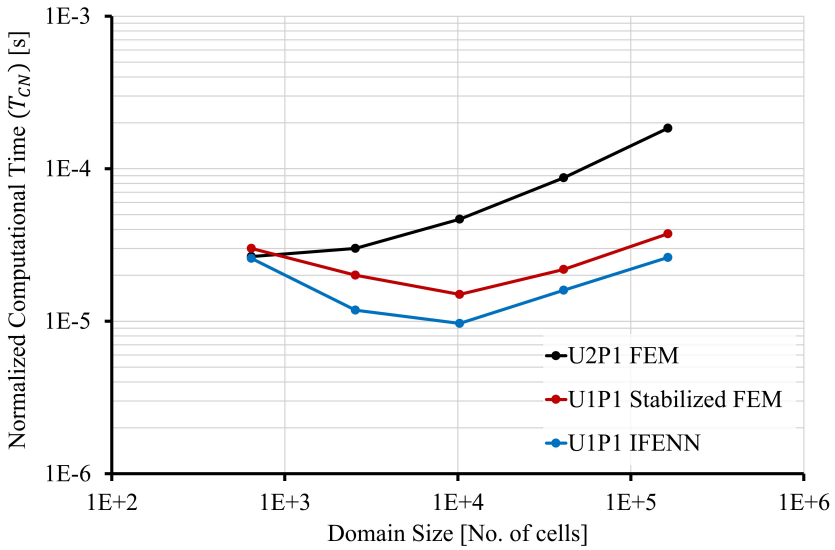


图14 效率分析：交错耦合有限元求解器与 IFENN-deal.II 的归一化计算时间（T_{CN}）。 T_{CN} 的计算方法为：总计算时间除以（时间增量数量 N_{inc} × 自由度数量 N_{DOF}）。

总计算时间而言会变得微不足道。

6.2 二维注入示例与可变渗透率 K_p 和时间增量 Δt

6.2.1 PIF-MIONet 训练

PIF-MIONet 在新的泛化设置下进行测试，在该设置中，模型对可变材料属性和时间离散进行编码。在此设置中，我们在可变渗透率（K_p）和可变时间步长（Δt）下训练模型。K_p 和 Δt 的变化被组织为三个渗透率组，每组包含三个对应于不同时间离散的子组，如表2所列。所有组均采用相同的注入历史；但是，注入速率（总时间）会根据渗透率 K_p 值进行调整。对于具有相同 K_p的子组，时间增量的数量被设置为使总时间保持一致。在相同注入幅值下如此广泛的渗透率值预计会导致超出线性孔弹性假设的变形。然而，为了确保全文一致，所有模拟均假设线弹性行为。

表2 不　同组属性（针对变量 K_p 和 Δ训练的广义 PIF-MIONet 模型）　t.

组编号	G1	G2	G3
渗透率（K_p）	1×10^{-14}	1×10^{-13}	1×10^{-12}
总时间	5.0×10^{8}	5.0×10^{7}	5.0×10^{6}
压力缩放因子（p^*）	5.0×10^{6}	5.0×10^{5}	5.0×10^{4}
应变缩放因子（ε^*）	2.5×10^{-1}	2.5×10^{-2}	2.5×10^{-3}
增量数（N_{inc}）	增量步长（Δt）		
500	1×10^{6}	1×10^{5}	1×10^{4}
50	1×10^{7}	1×10^{6}	1×10^{5}
5	1×10^{8}	1×10^{7}	1×10^{6}

对于这一广义问题，渗透率的两个分支 K_p 和时间增量大小 Δt 被纳入 PIF-MIONet 架构中，如图6所示。本示例中的训练完全采用物理信息方法，数据损失项未包含在损失函数中。然而，第5.1节所述的稀疏增量训练策略需要物理一致的输入场数据集。为此，训练、验证和测试数据集从表2中列出的所有组中进行采样。因此，压力场和应变场的量级预计会相差多个数量级。对所有采样增量进行全局归一化可能会损害网络编码某些输入的能力。例如，将跨越 1.0×10^4 至 1.0×10^6 的增量的压力输入归一化 $[0, 1]$ 到范围 $[0, 1]$ 会压缩小量级压力场中的变化。因此，我们采用逐样本归一化策略，其中针对特定增量的每个输入/输出都使用预定义的缩放因子进行缩放。

逐样本归一化策略仅应用于每个增量的压力场和应变场。其他参数，如通量、渗透率和时间增量大小，则进行全局归一化，使网络能够将这些输入视为连续场并接受任意值。压力（p^*）和应变（ε^*）的缩放因子见表 2。值得注意的是，压力

缩放因子（p^*）与残差形式的无量纲缩放中使用的因子相同（参见第 4.3节）。为简洁起见，训练过程的更多详情见附录 B.2。

6.2.2 IFENN-deal.II 在可变渗透率 K_p 和时间增量 Δt下的结果

将训练后的模型集成到 IFENN-deal.II 框架中，并针对表 2所列的三组（九个子组）进行测试。第 9 个测试荷载工况出现较大波动，并在最后一个时间增量达到最大振幅（见图 11），因此选取该工况进行评估，如图 15所示。

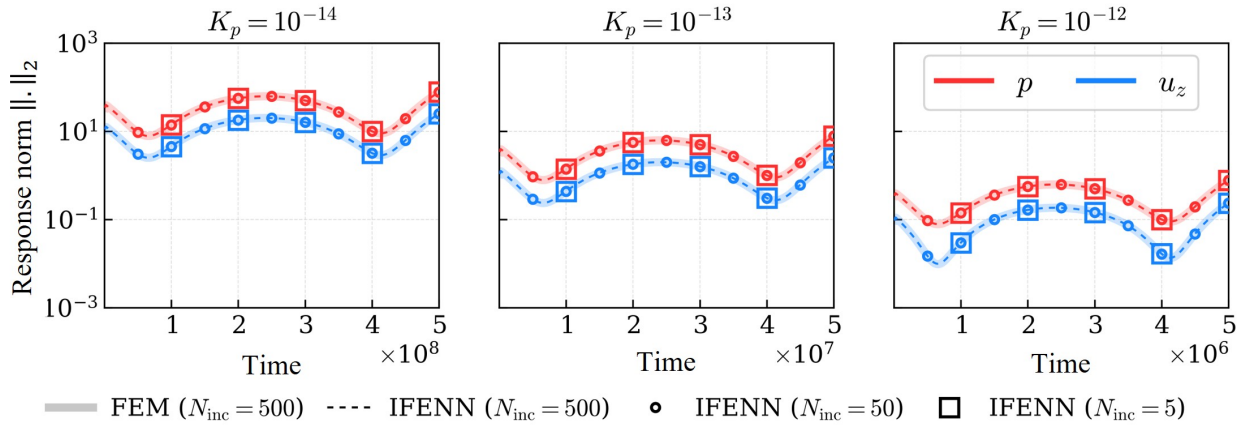


图 15 在二维注入问题的第 9 个测试荷载工况下，使用全耦合有限元求解器和 IFENN-deal.II 对 9 种不同设置（3 个渗透率值 × 3 和时间增量值）得到的压力范数 $\|p\|_2$ 和位移范数 $\|u_z\|_2$。压力和位移的响应范数单位分别为兆帕和米。

在图 15的每组中，对于采用不同时间离散的子组，压力范数 $\|p\|_2$ 和位移范数 $\|u_z\|_2$ 保持一致。具体而言，将增量总数 N_{inc} 从 500 减至5 仍能保持仿真终止时的精度。对 G1–G3 的比较进一步表明，同一个训练后的 PIF-MIONet 能够准确预测跨越两个至三个数量级的压力场。基于物理的缩放因子（表 p^* 中的 ε^* 2）处理这一压力范围（及相应的应变），而 PIF-MIONet 则捕捉渗透率 K_p 和时间增量 Δt带来的其余变化。附录图 B.9中第 7 个测试荷载工况也有类似观察结果。

6.3 三维注入示例

6.3.1 问题设置

前文展示了 IFENN 在二维注入问题中的性能。在此，我们将第 6.1 节中的相同问题扩展至三维区域，以测试所提框架能否处理体积输入、三维压力导数以及额外的边界条件项。尽管预计所得压力响应将呈现一维行为，但该问题完全在三维空间中进行建模和求解。图 9b 展示了该问题的三维区域及其边界条件。该三维区域采用八节点六面体单元的结构化网格进行离散，在三个空间方向上每个单元的边长均为 1/8 米。

6.3.2 PIF-MIONet 训练

对于这一三维设置，采用与三维兼容的 PIF-MIONet 进行训练，输入为定义在区域上的压力场和应变迹场，主干输入包含三个坐标分量。该三维问题保留了二维示例的主要加载和流动构型，同时在 y 方向上增加了前后边界 Γ^{front} 和 Γ^{back}。这些边界的处理方式与 Γ^{left} 和 Γ^{right}类似。因此，对于流体流动问题，前后表面均为不透水边界，并在 y 方向上受到机械约束。由于三维网格在 x 和 y 方向上采用相同的尺寸和分辨率，因此预计左/右边界与前/后边界的响应将表现出相似性。

在三维训练研究中，我们使用原始的二维 deal.II 模拟构建了一个受控的伪三维数据集。二维坐标网格沿 y 方向延拓，压力和应变迹历史数据则沿新增的深度方向复制。该伪三维数据集为分支输入提供了物理上相容的压力和应变迹历史数据。值得注意的是，该数据集仅用于在训练期间提供网络输入，不会计入任何数据驱动损失项。

三维物理信息损失在二维损失定义的基础上，增加了前后表面的两个边界项。这两个附加项，$\mathcal{L}_6^{front}$ 和 $\mathcal{L}_7^{back}$，用于施加 y 方向上的零通量条件。偏微分方程残差中的压力拉普拉斯算子（$\boldsymbol{R}_P$ 见式（17））也扩展到了 y 方向。采用自动微分计算空间导数。为施加诺伊曼边界条件，沿离散区域在 x、 y和 z 方向计算相应边界面上的压力梯度。关于训练过程和所采用的自定义学习率的更多细节见附录B.3。图16展示了监测的三维 PIF-MIONet 损失项及其自适应权重。图中可见，前后边界项的变化与左右边界项相似，这符合 x 方向和 y 方向之间的对称性。

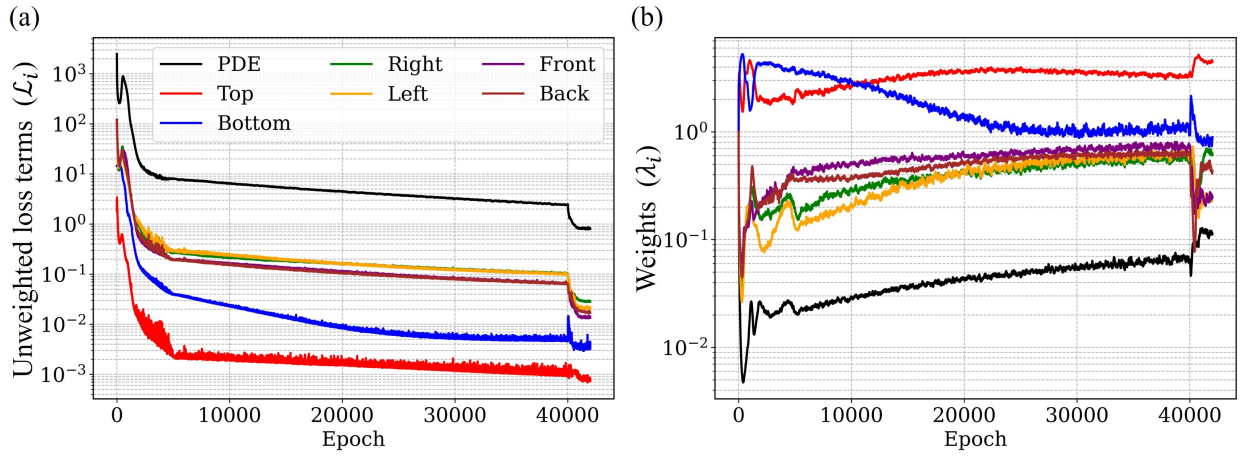


图 16 三维注入示例的训练指标：(a) 未加权损失项 $\mathcal{L}_i$，(b) 对应的自适应权重 λ_i。仅包含偏微分方程残差和边界条件损失；不使用数据损失。

6.3.3 三维注入示例的 *IFENN-GEOS* 结果

图17 比较了 IFENN-GEOS 与全耦合有限元求解器在三维问题中得到的响应范数。两者高度吻合，证明了该框架在三维情况下的准确度，与图12中的二维结果一致。图18展示了一个示例载荷工况，其中给出了压力场和位移场及其相对误差。压力分布在 x方向和 y方向上均匀，这与前、后、左、右边界上的零通量条件一致。最大相对误差出现在场值接近零的区域，而其他区域的误差均远低于5% 。

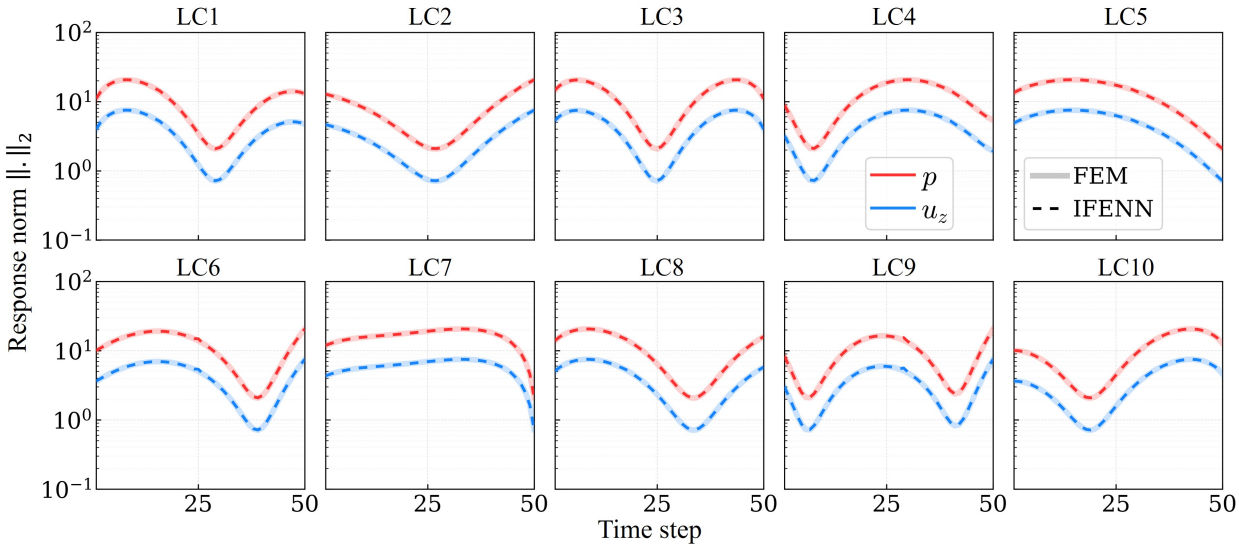


图 17 使用全耦合有限元求解器和 IFENN-GEOS 获得的三维注入问题的压力范数 $\|p\|_2$ 和位移范数 $\|u_z\|_2$ 。压力和位移的响应范数单位分别为兆帕和米。

总体而言，结果表明，在引入体积场和额外的 y方向边界项后，无论是离线训练还是在线实现，该框架仍能保持稳定。此示例表明，IFENN 可以从二维扩展到三维域，并与 GEOS 耦合，用于体积孔隙力学模拟。附录 B.3.2 中提供了展示不同载荷工况下相对误差和预测场的补充结果。

6.4 开挖问题

6.4.1 问题设置

最后一个示例是一个受排水通量作用的开挖问题。该示例受到文献中岩土工程问题的启发[101,102] ，作者此前曾使用数据驱动训练方法对其进行研究。[48] 在本节中，我们采用基于物理的方法重新研究该问题，讨论复杂边界条件带来的挑战，并强调使用 PI-ML 处理此类问题的优势。

问题几何如图 19所示，其中展示了排水通量 $q(t)$ 以及不同的边界条件。材料属性设为 $K_s = 55556$ 兆帕， $\lambda = 8.375$ 兆帕，$\mu = 5.58$ 兆帕， $K_f = 2200$ 兆帕， $\phi_b = 0.4$， $\mu_f = 1.0 \times 10^{-3}$ 帕·秒，以及 $K_p = 1.0 \times 10^{-11}$ 米^2。模拟时间设为 3×10^6 秒，总增量数为 $N_{inc} = 60$。

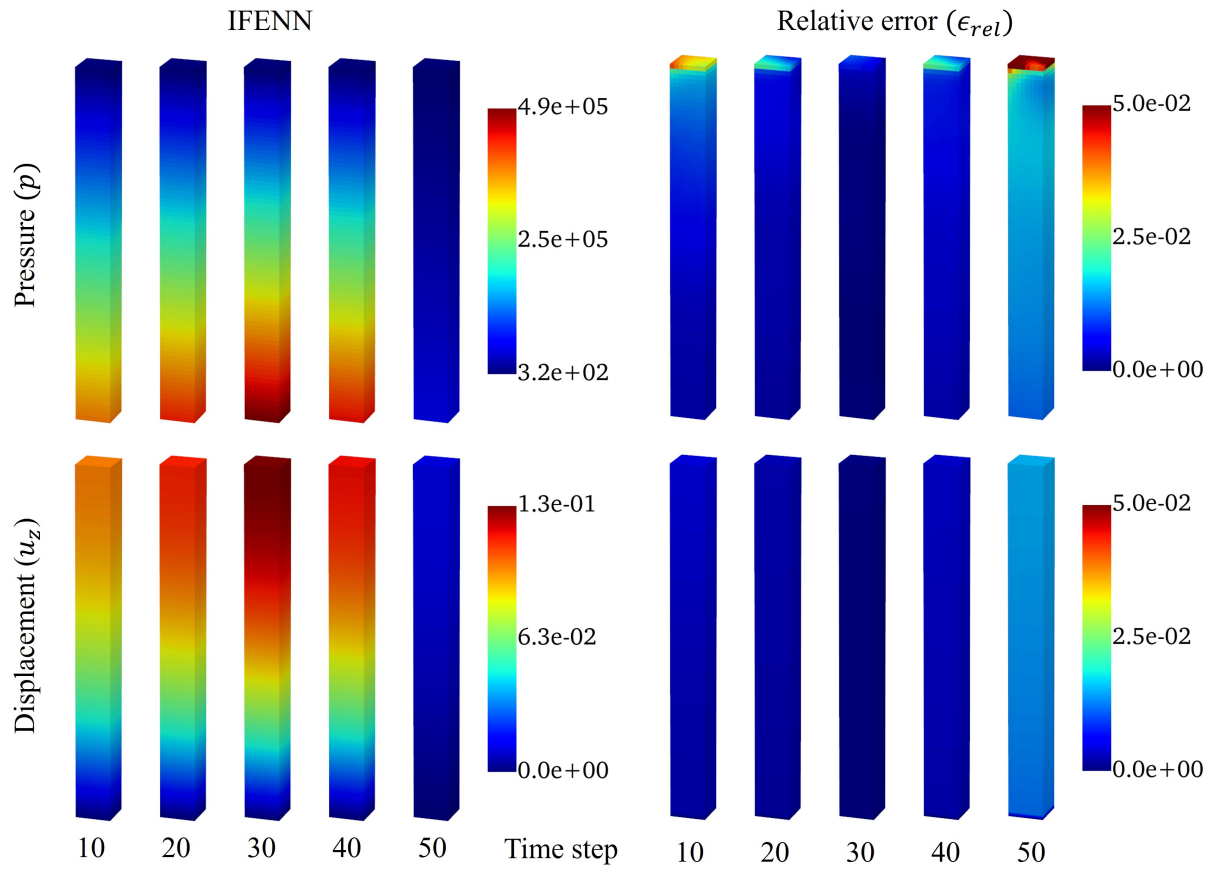


图 18 第 7 个测试载荷工况在不同时间步下，三维注入示例使用 IFENN-GEOS 得到的压力 p 和位移 u_z，以及相应的相对误差（ϵ_{rel}）。

6.4.2 PIF-MIONet 设计与训练选项

生成了 100 组排水通量 $q(t)$ 的高斯随机历程，并用于构建相应的输入/输出对。针对这 100 次模拟生成的数据集被划分为训练、验证和测试子集，数量分别为 70、20 和 10。从全部训练模拟（4200 个时间增量）中抽取 33 个增量，并从验证模拟（1200 个时间增量）中抽取 33 个增量。

给定模型需要考虑 7 个边界条件，如图19所示。初步训练试验表明，排除三个边界部分 Γ^{right}、Γ^{wall2}和 Γ^{wall1} 可以提高训练准确度。因此，我们将边界条件划分为两个子集：$\Gamma^{base} =\{\Gamma^{top1}, \Gamma^{top2}, \Gamma^{left}, \Gamma^{bottom}\}$ 和 $\Gamma^{comp} = \{\Gamma^{wall1}, \Gamma^{wall2}, \Gamma^{right}\}$。为求完整，我们考察排除和纳入 Γ^{comp}这两种情况。为便于后续讨论，分别将与边界子集相关的损失项加权和 Γ^{base} 和 Γ^{comp} 记为 $\mathcal{L}^{base} =\sum_{\Gamma^i \in \Gamma^{base}} \lambda_i \mathcal{L}_i$ 和 $\mathcal{L}^{comp} = \sum_{\Gamma^i \in \Gamma^{comp}} \lambda_i \mathcal{L}_i$。与 Γ^{comp} 处边界条件相关的自适应权重汇总为向量 $\boldsymbol{\lambda}_{comp} = \{\lambda_{wall1}, \lambda_{wall2}, \lambda_{right}\}$。是否纳入这些边界条件由其对应的权重控制；权重为零表示相应的损失项不纳入训练。

对于这种复杂几何结构，使用自动微分计算空间导数会导致训练不稳定，且无法收敛（表B.6中的模型M1-AD）；造成这种现象的一个可能原因是

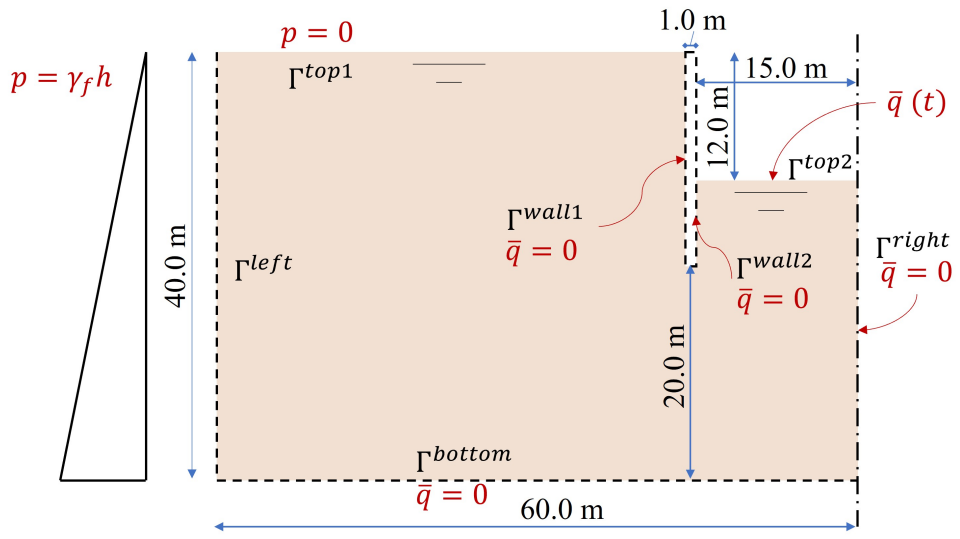


图 19 开挖问题的几何形状和边界条件示意图。出于训练方面的考虑，将边界分为两个子集：$\Gamma^{base} = \{\Gamma^{top1}, \Gamma^{top2}, \Gamma^{left}, \Gamma^{bottom}\}$ 和 $\Gamma^{comp} = \{\Gamma^{wall1}, \Gamma^{wall2}, \Gamma^{right}\}$。

附录B.4.1讨论了这一现象。根据这一说明以及先前示例的发现，我们采用形函数导数（SF）、GradNorm 自适应加权（GN）、无量纲缩放（DS）和 L_2 范数，$N(\cdot) = \|\cdot\|_2$ 来评估各个损失项。默认情况下，后续所有架构和训练配置均采用这一组合，记为“SF-GN-DS-L2”。

我们开展了多次训练试验，以评估采用单干网络架构和多干网络架构的 PIF-MIONet，分别如图 7 和 8所示。对于多干网络架构，区域被划分为三个子域（图 20a），每个子域均配备一个专用主干，用于编码其空间坐标。表3 列出了本问题所采用的三主干网络架构的详细信息。单干网络架构的详细信息见附录表 B.5。

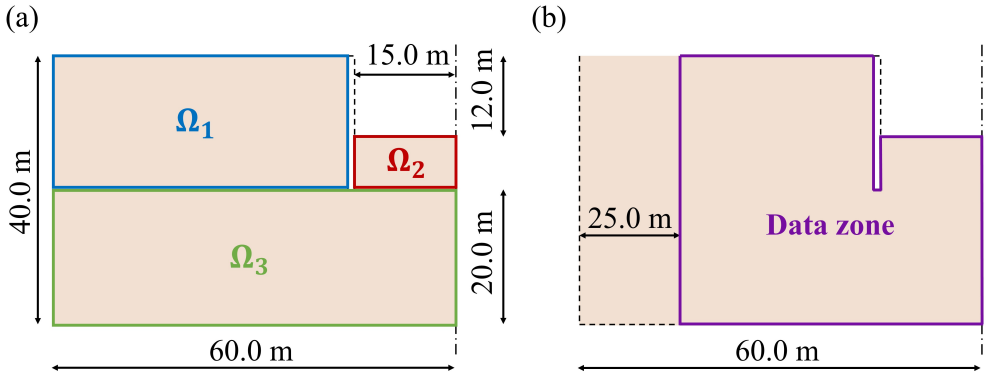


图 20 以下内容的示意图：（a）将区域分解为 3 个子域，以利用区域分解进行训练；（b）数据区域的边界，用于对数据损失项的节点进行采样。

训练试验中测试的另一方面是引入数据损失项（$\mathcal{L}^{data}$），其对应的自适应权重记为（$\lambda_{data} \equiv \lambda_d$），如式（16）所示。

表 3 针对二维开挖示例训练的三主干PIF-MIONet模型的超参数

网络	输入尺寸	隐藏层尺寸[a]	N_{ch}[b]	输出尺寸
分支 #1	N_q =1	[64,64]	-	64
分支 #2	N_ε =499	[64,64]	-	64
分支 #3	N_p =499	[64,64]	-	64
融合	N_{fi} =192	[192]	-	256
主干 #1	N_c =2	[128,128,128]	-	256
主干 #2	N_c =2	[128,128,128]	-	256
主干 #3	N_c =2	[128,128,128]	-	256

[a] 尺寸数组的长度即为隐藏层的数量。[b] 在输入层的输出上，采用 N_{ch} 个通道进行组归一化。“-”表示不考虑组归一化。

值得注意的是，这是唯一一个在物理信息损失函数中加入数据损失项的示例。用于施加数据项的配点采样自图20b 所示的数据区域。

附录 B.4.1 给出了所有损失项的详细参数化结果和说明。表现最佳的配置为采用三个分支的 PIF-MIONet，并将其作为参考模型。其损失函数结合了偏微分方程和数据损失，以及关于 Γ^{base} 的边界条件项，同时排除了关于 Γ^{comp} 的边界条件项。训练采用 Adam 优化器，训练 30000 轮次，学习率保持恒定为 1.0×10^{-5}。

图 21比较了参考模型与三种其他架构和训练配置的未加权损失项。图 22展示了 IFENN-deal.II 对第 9 个测试荷载工况的压力预测和相对误差，测试荷载历史见图 B.15。以下各节将考察不同选择对训练和预测性能的影响。

纳入所有边界条件的影响： 纳入 $\mathcal{L}^{comp}$ 以施加所有边界条件，会导致训练损失出现振荡（图21b），比参考模型更为明显（图21a）。此配置还使附加边界附近的局部误差更高，Γ^{wall1} 以及 Γ^{wall2} （图22b）。值得注意的是，排除与 Γ^{comp} 相关的边界条件，目的是提高优化稳定性，而非改变底层物理问题。

主干数量的影响： 另一个模型采用单主干进行训练，该主干包含三个隐藏层，每层有256个神经元，详情见附录表B.5。三主干模型和单主干模型分别有374848 个和373696 个可训练参数，因此模型规模相当。它们最终的验证误差分别为$L_2^{all} = 3.7 \times 10^{-3}$ 和 $L_2^{all} = 1.2 \times 10^{-2}$。

与参考模型相比，单主干模型在偏微分方程和边界条件损失方面也表现出更大的振荡（图 21c）。这一现象表明，使用一组基函数来表示复杂区域中的解更为困难。单主干模型的压力预测在壁面和顶部边界附近也表现出更高的误差（图22c）。

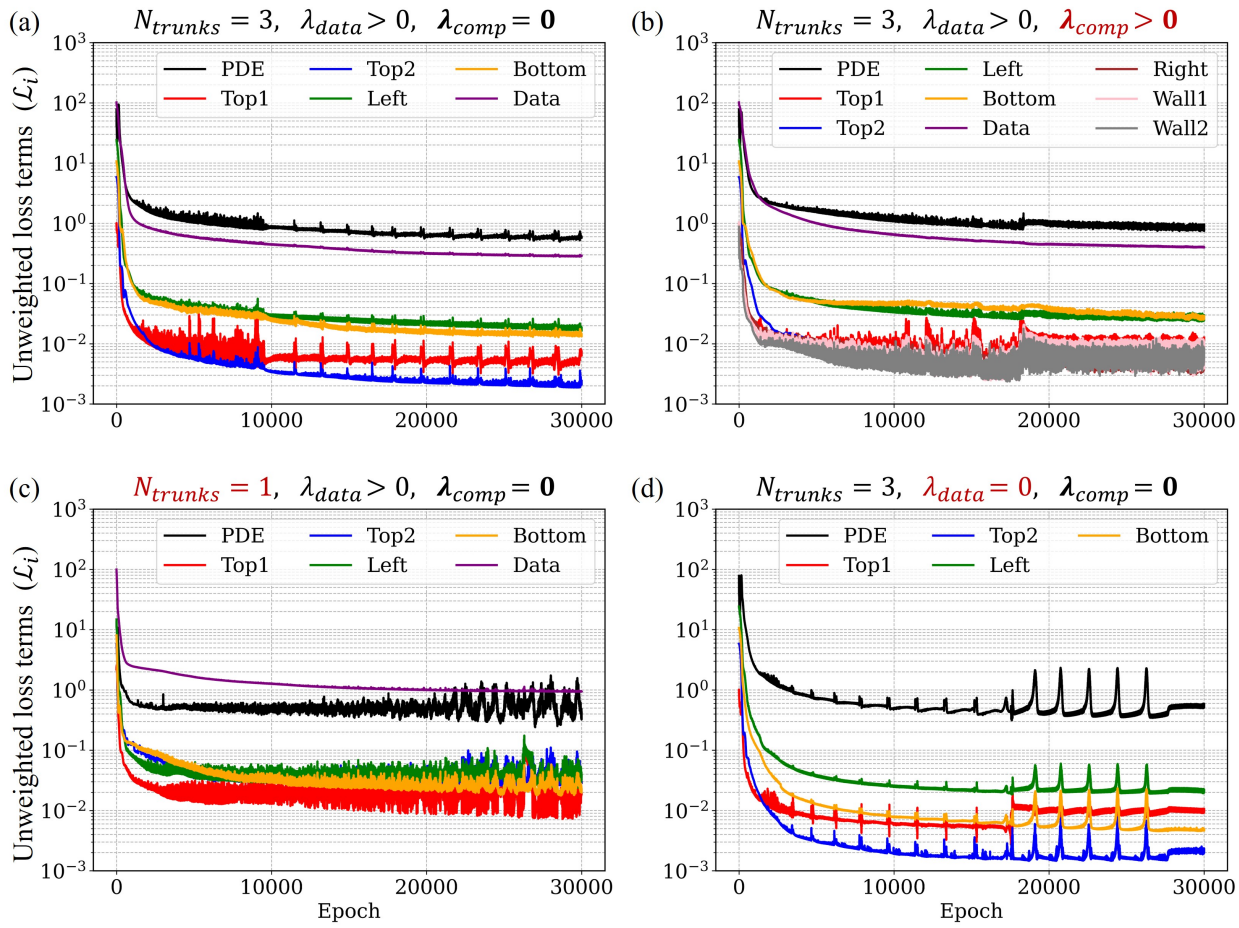


图21 四种训练配置下开挖问题的未加权损失项训练：(a) 具有 $\mathcal{L}^{data}$ ($\lambda_{data} > 0$) 且不具有 $\mathcal{L}^{comp}$ ($\boldsymbol{\lambda}_{comp} = \mathbf{0}$) 的三个主干的参考模型，(b) 同时具有 $\mathcal{L}^{data}$($\lambda_{data} > 0$) 和 $\mathcal{L}^{comp}$ ($\boldsymbol{\lambda}_{comp} > \mathbf{0}$) 的三个主干，(c) 具有 $\mathcal{L}^{data}$ ($\lambda_{data} > 0$) 且不具有 $\mathcal{L}^{comp}$ ($\boldsymbol{\lambda}_{comp} = \mathbf{0}$) 的一个主干，以及 (d) 两个项均不具备的三个主干（$\lambda_{data} = 0$ 和 $\boldsymbol{\lambda}_{comp} = \mathbf{0}$）。

这些结果凸显了所提出的采用区域分解的 PIF-MIONet 架构的优势：在总参数量相当的情况下，该架构能够达到更高的准确度。

数据损失对物理信息训练的影响： 未使用数据损失项进行的初始训练实验表现出振荡收敛现象（图 21d）以及较大的预测误差（图22d）。这些困难可能与不可渗透壁尖端处尖角附近的奇异行为有关，在该处可观察到局部应变集中（附录，图 B.16）。此类几何奇异性给物理信息神经网络带来了挑战，因此促使人们采用响应增强[103] 和自适应配置点采样等方法。[104,105]

本研究通过在损失函数中加入补充数据项（$\mathcal{L}^{data}$），解决复杂几何形状下的这些训练难题。值得注意的是，这是本研究中唯一引入补充数据损失项以增强训练的例子。因此，我们探讨三个关键问题：(1) 数据项如何影响训练稳定性和预测精度；(2) 它是否会以牺牲物理信息约束为代价主导优化；以及 (3) 与纯数据驱动模型相比，该方法在

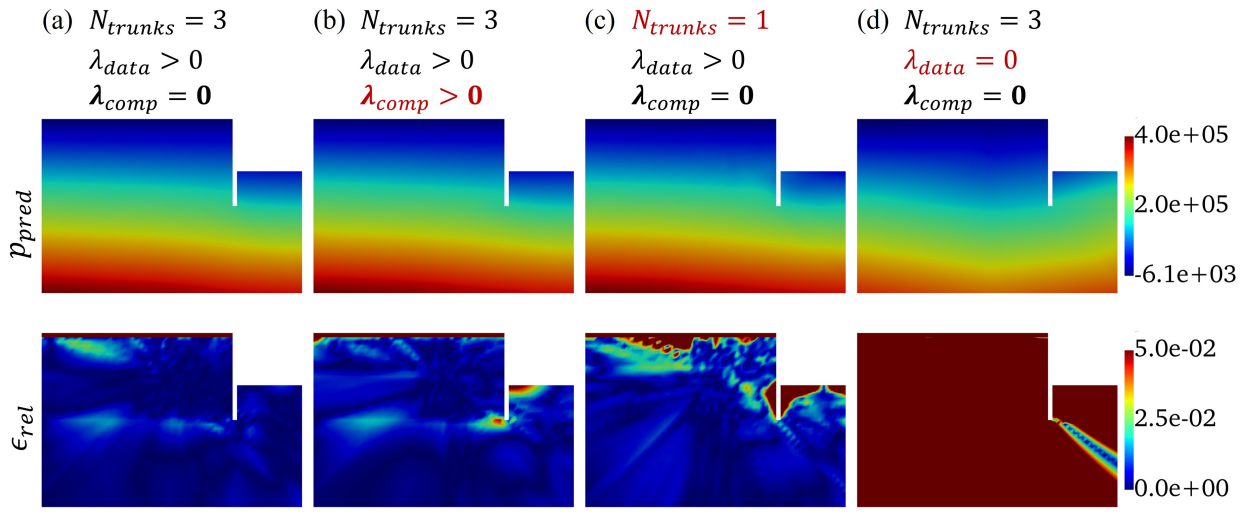


图 22 开挖示例中第 9 个测试载荷工况的第30个增量对应的预测压力场 p_{pred} （上排）和相对误差 ϵ_{rel} （下排）。从左至右的模型：(a) 包含 $\mathcal{L}^{data}$ ($\lambda_{data} > 0$)、不包含 $\mathcal{L}^{comp}$ ($\boldsymbol{\lambda}_{comp} = \mathbf{0}$) 的三分支参考模型；(b) 同时包含 $\mathcal{L}^{data}$ ($\lambda_{data} > 0$) 和 $\mathcal{L}^{comp}$($\boldsymbol{\lambda}_{comp} > \mathbf{0}$) 的三分支模型；(c) 包含 $\mathcal{L}^{data}$ ($\lambda_{data} > 0$)、不包含 $\mathcal{L}^{comp}$ ($\boldsymbol{\lambda}_{comp} = \mathbf{0}$) 的单分支模型；(d) 两项均不包含的三分支模型（$\lambda_{data} = 0$ 和 $\boldsymbol{\lambda}_{comp} = \mathbf{0}$）。

数据需求。对比图21a 和21d，以及图22a 和22d，可以看出，加入数据项能够稳定训练并显著降低预测误差，从而回答第一个问题。为回答第二个问题，我们将参考 PIF-MIONet 的偏微分方程损失项和数据损失项分离进行比较，如图23所示。

图 23a 显示两种未加权损失整体均有所下降，表明偏微分方程项和数据项同时得到改善。尽管未加权数据损失小于偏微分方程损失，但加权后两者的大小关系发生逆转（图 23b），这反映出 GradNorm 为数据项分配了更大的自适应权重（图 23c）。然而，对标量损失贡献更大并不一定意味着梯度贡献更大。事实上，在训练的大部分时间里，加权偏微分方程梯度范数约比数据项的梯度范数大一到两个数量级（图 23d）。因此，尽管数据项的加权损失更大，它在训练反向传播过程中并未在梯度幅值上压过偏微分方程项，偏微分方程残差仍持续下降。

为回答第三个问题，我们比较所提出方法与纯数据驱动模型的数据需求。[48] 所提出的物理信息、数据辅助方法使用了 33 个稀疏采样、非连续的时间增量。相比之下，数据驱动模型使用连续历史记录中的 28,800 个时间增量进行训练。因此，所提出方法使用的训练增量约为数据驱动模型的 0.1%。这凸显了物理信息训练大幅降低数据需求的潜力。未来研究可着重优化数据点的数量及空间分布。附录中还展示了探索其他训练配置的参数化研究B.4.1。

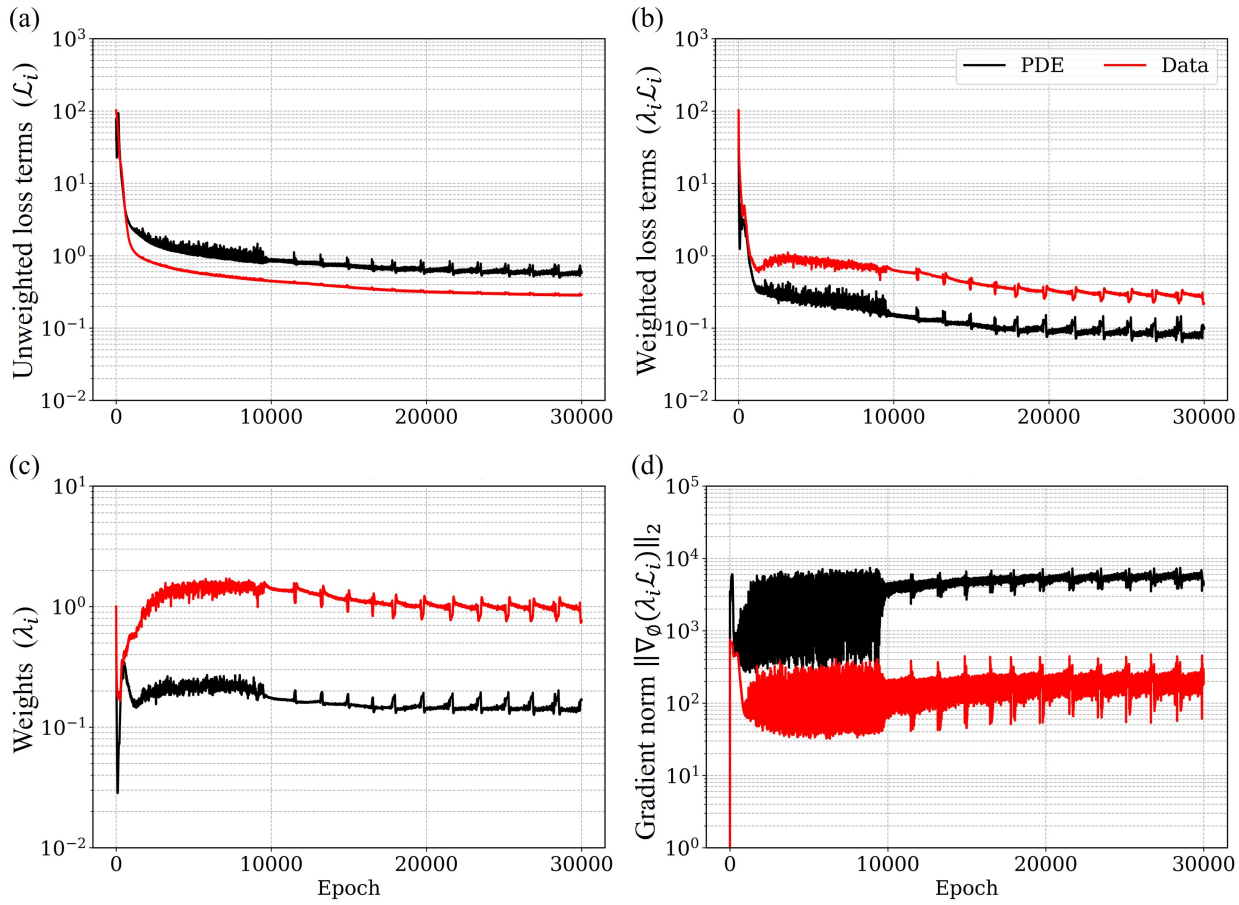


图23 开挖示例中采用三个分支的参考 PIF-MIONet 的训练指标：(a) 未加权损失项 $\mathcal{L}_i$，(b) 加权损失项 $\lambda_i\mathcal{L}_i$，(c) 自适应权重 λ_i，以及 (d) 加权梯度范数 $\|\nabla_\phi(\lambda_i\mathcal{L}_i)\|_2$。为清晰起见，此图仅展示偏微分方程和数据项。

6.4.3 IFENN-deal.II 使用多个子循环

使用训练后的三分支参考模型得到的初步 IFENN-deal.II 结果显示，采用预测器-多校正器方法并设置总子循环数 $N_{sub} = 2$ 会导致较大的误差，尤其是在初始时间步。进一步研究表明，增加子循环数（N_{sub}）可以大幅降低这些误差。图 24 展示了测试载荷工况（见图 B.15）的压力L_2^{inc}误差，采用了不同的子循环数（$N_{sub} = 2$、3、4）。位移误差（u_z，u_x）见附录 B.4.3，其表现一致。

在预测-多重校正方案中，预测子循环使用前序增量中的已知应变率值，从而改进初始应变率估计。然而，对于第一个时间增量，由于缺乏前序增量提供的先验信息，应变率迹的初始估计设为零。这一点在图 24所示结果中体现得十分清楚。增加子循环数（N_{sub}）可以改善预测结果，并在最初约五个增量内将误差 L_2^{inc}降低一个数量级。对于后续增量，两个子循环（$N_{sub} = 2$）即可提供令人满意的准确度；增加子循环数并不会带来改进，因为预测器-多校正器算法在开始时已利用前序

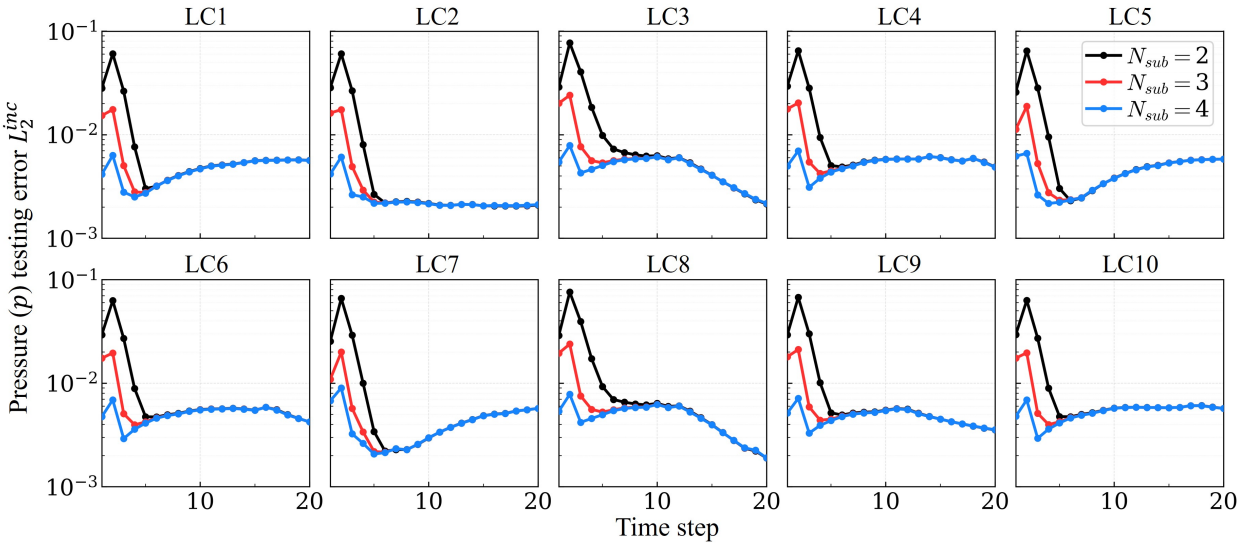


图 24 使用 IFENN-deal.II 求解开挖问题时，所有测试载荷工况（LC）前 20 个增量的压力归一化测试误差 L_2^{inc}（p），其预测-多重校正子循环数分别为 $N_{sub} = 2$、3、4。

增量。

使用四个子循环（$N_{sub} = 4$）时，如图 25 所示，所有载荷工况和时间增量下，预测压力范数和垂直位移范数均与全耦合有限元求解器的结果高度吻合。与注入示例不同，开挖问题中的压力范数和垂直位移范数呈现相反的变化趋势。随着排水通量增加，压力范数降低，而位移范数增加。预测结果与有限元结果高度吻合，表明 IFENN 能够捕捉这一独特响应，体现了其应用于不同孔弹性问题的能力。附录B.4 中提供了更多展示预测精度的结果。

7 讨论与结论

7.1 讨论

使用 IFENN-deal.II 和 IFENN-GEOS 对 PIF-MIONet 模型进行测试后发现，对于二维和三维注入问题，模型在所有载荷工况下均能给出准确且一致的预测（图 12、13 和 17）。这些结果证明，单个训练好的网络可在 IFENN 中的不同有限元求解器之间移植。研究采用创新技术，成功地首次引入了 IFENN-GEOS。进一步开发后，任何训练好的物理信息模型都可以作为求解器接入 IFENN-GEOS，而无需更改集成工作流。这样便可在复杂的工业规模应用中直接使用 GEOS 功能。

广义 PIF-MIONet 针对可变渗透率 K_p 和时间增量Δt 进行了训练，并使用预定义的缩放因子对每个样本的压力场和应变场进行归一化。当前实现仍受限于需要为这些场预定义缩放因子。不过，可以通过一种能够自动分类输入场并选择合适缩放因子的算法来克服这一限制。或者，未来的工作可以探索针对跨越多个数量级的输入的编码方法，以提升 PIF-MIONet 的泛化能力。

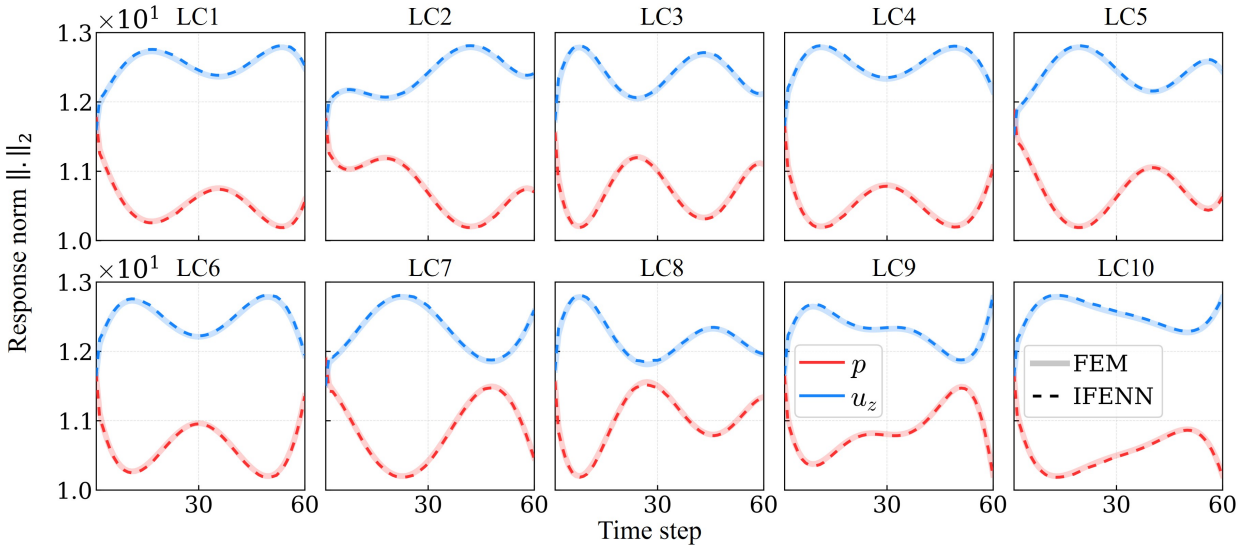


图 25 使用全耦合有限元求解器和 IFENN- deal.II 通过四个预测-多重校正子循环（$N_{sub} = 4$）获得的所有开挖问题测试载荷工况（LCs）的压力范数 $\|p\|_2$ 和位移范数 $\|u_z\|_2$。压力和位移的响应范数单位分别为兆帕和米。

注入问题完全采用物理信息方法，无需数据损失项；而开挖问题则需要数据损失项，以提升复杂边界条件下的训练准确度。该问题仅使用了0。1%数据驱动训练所用的带标签数据增量，[48]凸显了物理信息训练对数据需求较低。结果还表明，数据项能够对训练起到正则化作用，同时不会主导优化过程。未来的工作可以进一步优化所纳入数据的规模，或探索如何消除对数据的需求。

7.2 结论

本文提出了用于耦合多物理场问题物理信息学习的 PIF-MIONet，并通过二维和三维孔隙弹性示例展示了其与 IFENN 的集成。主要贡献和发现如下：

- 融合设计可耦合任意数量的输入场，并支持扩展至其他相互作用的物理场。

- 通过区域分解，专用子域主干网络能够在复杂几何中提高准确度，同时保持可比的可训练参数数量。

- 在来自不同载荷工况的稀疏、不连续增量上进行训练，可以预测未见历史；在非平凡区域内，相对误差低于 5%。

- 注入示例在训练时不使用数据损失项。对于存在几何奇异性的开挖示例，加入补充数据损失可提高训练稳定性和准确度，同时所需的监督训练增量仅约为此前报告的数据驱动方法所用数量的 0.1%。

• 单个训练完成的 PIF-MIONet 可在 IFENN 中的不同有限元求解器之间移植，并能生成一致的预测结果。

• 首次IFENN-GEOS集成将IFENN扩展至工业仿真环境，其预测结果通过与deal.II有限元结果进行独立对比得到验证。

• IFENN将其计算效率优势扩展至孔隙弹性问题；随着区域尺寸增大，相较于全耦合有限元法，其表现出良好的可扩展性。

这些发现支持进一步发展IFENN，以应对涉及数百万自由度的复杂地下和多物理场应用。

致谢

本研究得到沙灾害、韧性、能源与可持续发展机遇中心（SHORES Center）的部分支持；该中心由Tamkeen通过NYUAD研究院CG013奖项资助。作者还感谢NYUAD研究计算中心提供的资源、服务和人员专业支持，这些支持为本研究作出了贡献。本研究的部分工作由劳伦斯利弗莫尔国家实验室代表美国能源部，在合同DE-AC52-07NA27344框架下完成。

数据可用性

本研究结果所依据的数据和源代码将在文章发表后公开。

附录A：PIF-MIONet 训练的更多详情

A.1 自适应权重

当高梯度项主导训练过程时，物理信息训练可能会出现数值刚性，从而对收敛产生不利影响。[28,37] 为克服这种数值刚性，研究人员采用自适应加权算法，在训练过程中动态平衡各损失项。[33–37] 本文测试了两种采用根本不同策略的自适应权重算法。第一种自适应权重算法 GradNorm 由 Chen等人提出。 [34] 该算法提高了准确度并减少了过拟合。此外，文献中也已将其用于孔弹性问题。[28] 在介绍算法细节之前，我们先列出一组基本定义如下：

$$G_i^{\phi} = \|\nabla_{\phi}(\lambda_i \mathcal{L}_i)\|_2 \tag{A.1}$$

$$\overline{G}^{\phi} = E_{\forall i}[G_i^{\phi}] \tag{A.2}$$

$$\tilde{\mathcal{L}}_i = \mathcal{L}_i / \mathcal{L}_i^0 \tag{A.3}$$

$$r_i = \tilde{\mathcal{L}}_i / E_{\forall i}[\tilde{\mathcal{L}}_i] \tag{A.4}$$

其中， G_i^{ϕ} 表示加权损失分量 $\lambda_i \mathcal{L}_i$相对于模型参数 ϕ的梯度范数。 G^{ϕ} 表示所有 G_i^{ϕ} 值的平均范数， $\mathcal{L}_i^0$ 表示在˜第一个训练轮次时的初始损失值 $\mathcal{L}_i$ ， $\mathcal{L}_i$ 表示项 i的逆训练速率，而 r_i 表示相对逆训练

速率。为节省计算成本，梯度范数(G_i^ϕ) 可以针对模型权重的一个子集进行计算。[34] 然而，由于PIF-MIONet架构复杂，我们计算G_i^ϕ 所有可训练参数上的该值。

在第一个训练轮次中，所有权重均设为 1。随后，对于每个训练轮次 $\mathcal{E}$，新的权重 $\lambda_i^\mathcal{E}$可按如下方式更新：[28,34]

$$\hat{\lambda}_i^\mathcal{E} = \lambda_i^{\mathcal{E}-1}\frac{\overline{G}^\phi}{G_i^\phi}r_i^{\alpha_g} \tag{A.5}$$

$$\lambda_i^\mathcal{E} = (1-\beta_g)\lambda_i^{\mathcal{E}-1} + \beta_g\hat{\lambda}_i^\mathcal{E} \tag{A.6}$$

其中 α_g 和 β_g 是可调节算法行为的 GradNorm 调优参数。α_g 控制学习率较低的项所受惩罚项的强度。 β_g 是平滑因子 $(0 < \beta_g \leq 1)$，用于防止更新后的权重出现较大波动。最后需要注意的是，使用公式（A.6）计算新权重后，会对权重进行归一化，使其均值为 1，从而将自适应权重算法与优化器的全局学习率解耦。[34]

本文考虑的第二种自适应权重算法是随机回溯相对损失平衡（ReLoBRaLo），由 Bischof 和 Kraus 提出。[33] 与 GradNorm 不同，ReLoBRaLo 不使用梯度统计量，而仅跟踪损失统计量来计算自适应权重。该算法依赖于两个可定义为如下形式的函数：

$$\lambda_i^{bal}(\mathcal{E},\mathcal{E}') = N_{terms}\frac{\exp\left(\frac{\mathcal{L}_i^\mathcal{E}}{\theta_r\mathcal{L}_i^{\mathcal{E}'}}\right)}{\sum_{j=1}^{N_{terms}}\exp\left(\frac{\mathcal{L}_j^\mathcal{E}}{\theta_r\mathcal{L}_j^{\mathcal{E}'}}\right)} \tag{A.7}$$

$$\lambda_i^{hist}(\mathcal{E}) = \rho_r\lambda_i^{\mathcal{E}-1} + (1-\rho_r)\lambda_i^{bal}(\mathcal{E},0) \tag{A.8}$$

其中， $\mathcal{L}_i^\mathcal{E}$是项 i 在训练轮次 $\mathcal{E}$时的损失， θ_r 是温度参数，用于控制缩放因子的大小； ρ_r 是一个伯努利随机变量（$\rho_r \in \{0,1\}$），其期望值 $E[\rho_r]$接近于 1。利用前述函数，可以计算出新的 ReLoBRaLo 权重 $\lambda_i^\mathcal{E}$：

$$\lambda_i^\mathcal{E} = \alpha_r\lambda_i^{hist}(\mathcal{E}) + (1-\alpha_r)\lambda_i^{bal}(\mathcal{E},\mathcal{E}-1) \tag{A.9}$$

其中， α_r 是指数衰减率，用于控制先前权重与新平衡权重的相对贡献。

A.2 模型稳定化

如图所示。 A.1，在两层（L_l 和 L_{l+1}）之间进行组归一化：将大小为 N_{l+1} 的输出向量拆分为 N_{ch} 个通道，对每个通道分别进行组归一化，然后将输出重新整形为形状为 N_{l+1} 的统一向量。

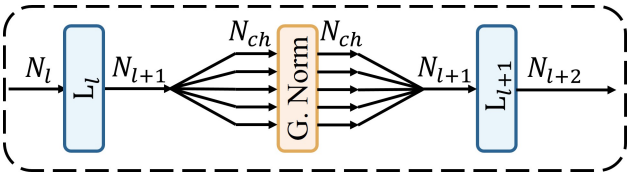


图 A.1 两个线性前馈层之间所采用的稳定化方法示意图。某层 L_l 的输入尺寸为 N_l，输出尺寸为 $O_l = N_{l+1}$。

表 A.1 形函数及其自然坐标导数。

Node	N_n	$\partial N_n/\partial\xi$	$\partial N_n/\partial\eta$	$\partial^2 N_n/\partial\xi^2$	$\partial^2 N_n/\partial\eta^2$	$\partial^2 N_n/\partial\xi\partial\eta$
1	$\xi\eta(\xi-1)(\eta-1)/4$	$\eta(2\xi-1)(\eta-1)/4$	$\xi(\xi-1)(2\eta-1)/4$	$\eta(\eta-1)/2$	$\xi(\xi-1)/2$	$(2\xi-1)(2\eta-1)/4$
2	$-\xi(\xi-1)(\eta^2-1)/2$	$-(2\xi-1)(\eta^2-1)/2$	$-\xi\eta(\xi-1)$	$1-\eta^2$	$-\xi(\xi-1)$	$-\eta(2\xi-1)$
3	$\xi\eta(\xi-1)(\eta+1)/4$	$\eta(2\xi-1)(\eta+1)/4$	$\xi(\xi-1)(2\eta+1)/4$	$\eta(\eta+1)/2$	$\xi(\xi-1)/2$	$(2\xi-1)(2\eta+1)/4$
4	$-\eta(\xi^2-1)(\eta-1)/2$	$-\xi\eta(\eta-1)$	$-(\xi^2-1)(2\eta-1)/2$	$-\eta(\eta-1)$	$1-\xi^2$	$-\xi(2\eta-1)$
5	$(\xi^2-1)(\eta^2-1)$	$2\xi(\eta^2-1)$	$2\eta(\xi^2-1)$	$2(\eta^2-1)$	$2(\xi^2-1)$	$4\xi\eta$
6	$-\eta(\xi^2-1)(\eta+1)/2$	$-\xi\eta(\eta+1)$	$-(\xi^2-1)(2\eta+1)/2$	$-\eta(\eta+1)$	$1-\xi^2$	$-\xi(2\eta+1)$
7	$\xi\eta(\xi+1)(\eta-1)/4$	$\eta(2\xi+1)(\eta-1)/4$	$\xi(\xi+1)(2\eta-1)/4$	$\eta(\eta-1)/2$	$\xi(\xi+1)/2$	$(2\xi+1)(2\eta-1)/4$
8	$-\xi(\xi+1)(\eta^2-1)/2$	$-(2\xi+1)(\eta^2-1)/2$	$-\xi\eta(\xi+1)$	$1-\eta^2$	$-\xi(\xi+1)$	$-\eta(2\xi+1)$
9	$\xi\eta(\xi+1)(\eta+1)/4$	$\eta(2\xi+1)(\eta+1)/4$	$\xi(\xi+1)(2\eta+1)/4$	$\eta(\eta+1)/2$	$\xi(\xi+1)/2$	$(2\xi+1)(2\eta+1)/4$

A.3 有限元法双二次单元的形函数导数

对于具有 9 个节点的双二次四边形，其自然坐标系中的形函数及其导数列于表 A.1中。为实现高效的向量化张量积运算，节点采用字典序排列，而非优先排列边界节点或角节点。

假设我们有一个以自然坐标 $f(\xi,\ \eta)$ 定义的函数。变量 f 在等参自然坐标 ξ 和 η 处的导数可计算为：

$$\frac{\partial f(\xi,\eta)}{\partial\xi} = \sum_{n=1}^{9} \frac{\partial N_n(\xi,\eta)}{\partial\xi} f_n \tag{A.10}$$

$$\frac{\partial^2 f(\xi,\eta)}{\partial\xi\partial\eta} = \sum_{n=1}^{9} \frac{\partial^2 N_n(\xi,\eta)}{\partial\xi\partial\eta} f_n \tag{A.11}$$

其中 N_n 和 f_n 分别是形函数和变量 f 在等参节点处的值 n。 $\partial f(\xi,\ \eta)/\partial\eta$ 可采用与式（A.10）相同的方法计算。同理，$\partial^2 f(\xi,\ \eta)/\partial\xi^2$ 和 $\partial^2 f(\xi,\ \eta)/\partial\eta^2$ 可采用与式（A.11）相同的方法计算。

现在，我们推导物理坐标 x 和 y中的导数。对$\partial f(\xi,\ \eta)/\partial\xi$和 $\partial f(\xi,\ \eta)/\partial\eta$应用链式法则，可以写出以下表达式：

$$\frac{\partial f}{\partial\xi} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial\xi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial\xi} \tag{A.12}$$

$$\frac{\partial f}{\partial\eta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial\eta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial\eta} \tag{A.13}$$

这两个表达式（式（A.12）和式（A.13））可以合并为如下矩阵形式：

$$\begin{Bmatrix} \frac{\partial f}{\partial \xi} \\ \frac{\partial f}{\partial \eta} \end{Bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} \begin{Bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{Bmatrix} \tag{A.14}$$

对于物理坐标 x 和 y下的二阶导数，我们对式（A.12）关于 ξ求导，得到：

$$\begin{aligned} \frac{\partial^2 f}{\partial \xi^2} &= \frac{\partial}{\partial \xi}\left(\frac{\partial f}{\partial \xi}\right) = \frac{\partial}{\partial \xi}\left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial \xi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \xi}\right) \\ &= \frac{\partial}{\partial \xi}\left(\frac{\partial f}{\partial x}\right)\frac{\partial x}{\partial \xi} + \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial \xi^2} + \frac{\partial}{\partial \xi}\left(\frac{\partial f}{\partial y}\right)\frac{\partial y}{\partial \xi} + \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial \xi^2} \end{aligned} \tag{A.15}$$

式（$\frac{\partial}{\partial \xi}\left(\frac{\partial f}{\partial x}\right)$和 $\frac{\partial}{\partial \xi}\left(\frac{\partial f}{\partial y}\right)$）中的两个导数可以利用链式法则计算如下：

$$\frac{\partial}{\partial \xi}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial \xi} + \frac{\partial^2 f}{\partial x \partial y}\frac{\partial y}{\partial \xi}, \tag{A.16}$$

$$\frac{\partial}{\partial \xi}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}\frac{\partial x}{\partial \xi} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial \xi} \tag{A.17}$$

将式（A.16）和式（A.17）代入式（A.15），二阶导数项可表示为：

$$\begin{aligned} \frac{\partial^2 f}{\partial \xi^2} = &\frac{\partial^2 f}{\partial x^2}\left(\frac{\partial x}{\partial \xi}\right)^2 + \frac{\partial^2 f}{\partial y^2}\left(\frac{\partial y}{\partial \xi}\right)^2 \\ &+ 2\frac{\partial^2 f}{\partial x \partial y}\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \xi} + \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial \xi^2} + \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial \xi^2}. \end{aligned} \tag{A.18}$$

分离出 f 关于物理坐标（x和 y）的二阶导数，得到：

$$\frac{\partial^2 f}{\partial x^2}\left(\frac{\partial x}{\partial \xi}\right)^2 + \frac{\partial^2 f}{\partial y^2}\left(\frac{\partial y}{\partial \xi}\right)^2 + \frac{\partial^2 f}{\partial x \partial y}\left(2\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \xi}\right) = \frac{\partial^2 f}{\partial \xi^2} - \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial \xi^2} - \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial \xi^2} \tag{A.19}$$

按照处理 $\frac{\partial^2 f}{\partial \eta^2}$和 $\frac{\partial^2 f}{\partial \xi \partial \eta}$的相同方法，关于物理坐标的孤立二阶项可表示如下

$$\frac{\partial^2 f}{\partial x^2}\left(\frac{\partial x}{\partial \eta}\right)^2 + \frac{\partial^2 f}{\partial y^2}\left(\frac{\partial y}{\partial \eta}\right)^2 + \frac{\partial^2 f}{\partial x \partial y}\left(2\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \eta}\right) = \frac{\partial^2 f}{\partial \eta^2} - \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial \eta^2} - \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial \eta^2} \tag{A.20}$$

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial \xi}\frac{\partial x}{\partial \eta} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial \xi}\frac{\partial y}{\partial \eta} + \frac{\partial^2 f}{\partial x \partial y}\left(\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta} + \frac{\partial y}{\partial \xi}\frac{\partial x}{\partial \eta}\right) = \frac{\partial^2 f}{\partial \xi \partial \eta} - \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial \xi \partial \eta} - \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial \xi \partial \eta} \tag{A.21}$$

三个方程 Eq. (A.19)、Eq. (A.20) 和 Eq. (A.21) 可组装成如下矩阵形式：

$$\begin{bmatrix} \left(\frac{\partial x}{\partial \xi}\right)^2 & \left(\frac{\partial y}{\partial \xi}\right)^2 & 2\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \xi} \\ \left(\frac{\partial x}{\partial \eta}\right)^2 & \left(\frac{\partial y}{\partial \eta}\right)^2 & 2\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \eta} \\ \frac{\partial x}{\partial \xi}\frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \xi}\frac{\partial y}{\partial \eta} & \frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta}+\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \xi} \end{bmatrix} \begin{Bmatrix} \frac{\partial^2 f}{\partial x^2} \\ \frac{\partial^2 f}{\partial y^2} \\ \frac{\partial^2 f}{\partial x \partial y} \end{Bmatrix} = \begin{Bmatrix} \frac{\partial^2 f}{\partial \xi^2} - \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial \xi^2} - \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial \xi^2} \\ \frac{\partial^2 f}{\partial \eta^2} - \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial \eta^2} - \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial \eta^2} \\ \frac{\partial^2 f}{\partial \xi \partial \eta} - \frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial \xi \partial \eta} - \frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial \xi \partial \eta} \end{Bmatrix} \tag{A.22}$$

为提高计算效率，方程（ $\frac{\partial f}{\partial x}$ 右端的 $\frac{\partial f}{\partial y}$ 一阶导数可从方程（A.22）的右端分离如下：

$$\begin{bmatrix} \left(\frac{\partial x}{\partial \xi}\right)^2 & \left(\frac{\partial y}{\partial \xi}\right)^2 & 2\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \xi} \\ \left(\frac{\partial x}{\partial \eta}\right)^2 & \left(\frac{\partial y}{\partial \eta}\right)^2 & 2\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \eta} \\ \frac{\partial x}{\partial \xi}\frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \xi}\frac{\partial y}{\partial \eta} & \frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta}+\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \xi} \end{bmatrix} \begin{Bmatrix} \frac{\partial^2 f}{\partial x^2} \\ \frac{\partial^2 f}{\partial y^2} \\ \frac{\partial^2 f}{\partial x \partial y} \end{Bmatrix} = \begin{Bmatrix} \frac{\partial^2 f}{\partial \xi^2} \\ \frac{\partial^2 f}{\partial \eta^2} \\ \frac{\partial^2 f}{\partial \xi \partial \eta} \end{Bmatrix} - \begin{bmatrix} \frac{\partial^2 x}{\partial \xi^2} & \frac{\partial^2 y}{\partial \xi^2} \\ \frac{\partial^2 x}{\partial \eta^2} & \frac{\partial^2 y}{\partial \eta^2} \\ \frac{\partial^2 x}{\partial \xi \partial \eta} & \frac{\partial^2 y}{\partial \xi \partial \eta} \end{bmatrix} \begin{Bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{Bmatrix} \tag{A.23}$$

根据前面的定义，直接实现时，可以先计算 f_n 单元节点处的值，然后利用表 A.1、方程（A.10）和方程（A.11）中的定义，在自然坐标下计算函数导数。在计算出相对于自然坐标的导数后，可以利用方程（A.14）和方程（A.23）中系统矩阵的逆，分别按照方程（A.24）和方程（A.25）计算相对于物理坐标的导数。

$$\begin{Bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{Bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix}^{-1} \begin{Bmatrix} \frac{\partial f}{\partial \xi} \\ \frac{\partial f}{\partial \eta} \end{Bmatrix} \tag{A.24}$$

$$\begin{Bmatrix} \frac{\partial^2 f}{\partial x^2} \\ \frac{\partial^2 f}{\partial y^2} \\ \frac{\partial^2 f}{\partial x \partial y} \end{Bmatrix} = \begin{bmatrix} \left(\frac{\partial x}{\partial \xi}\right)^2 & \left(\frac{\partial y}{\partial \xi}\right)^2 & 2\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \xi} \\ \left(\frac{\partial x}{\partial \eta}\right)^2 & \left(\frac{\partial y}{\partial \eta}\right)^2 & 2\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \eta} \\ \frac{\partial x}{\partial \xi}\frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \xi}\frac{\partial y}{\partial \eta} & \frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta}+\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \xi} \end{bmatrix}^{-1} \left\{ \begin{Bmatrix} \frac{\partial^2 f}{\partial \xi^2} \\ \frac{\partial^2 f}{\partial \eta^2} \\ \frac{\partial^2 f}{\partial \xi \partial \eta} \end{Bmatrix} - \begin{bmatrix} \frac{\partial^2 x}{\partial \xi^2} & \frac{\partial^2 y}{\partial \xi^2} \\ \frac{\partial^2 x}{\partial \eta^2} & \frac{\partial^2 y}{\partial \eta^2} \\ \frac{\partial^2 x}{\partial \xi \partial \eta} & \frac{\partial^2 y}{\partial \xi \partial \eta} \end{bmatrix} \begin{Bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{Bmatrix} \right\} \tag{A.25}$$

附录B：其他结果和参数化研究

B.1 二维恒定渗透率注入示例 K_p 和时间增量 Δ t

B.1.1 使用自适应权重、无量纲形式及其他超参数的影响

我们首先通过列出本研究中探索的不同技术及其选项来开展参数研究，如表B.1所示。例如，如果某种方法被称为“SF-GN-DS-L2”，则表示该方法使用形函数导数计算空间梯度，使用 GradNorm 设置自适应权重，采用无量纲缩放，并使用 L2 范数计算残差项。我们测试了 GradNorm 和 ReLoBRaLo 的不同调参变化。在此测试中，仅使用 3 个采样增量

进行训练，并在训练结束后计算这些增量的 L_2^{all}误差度量，如表B.2 和表B.3 所列，分别对应“SF-GN-DS-L2”和“SF-RB-DS-L2”。需要注意的是，所测试的调优参数值（α_g、 β_g、 α_r、 θ_r和 ρ_r）基于文献中的推荐值，[33,34] 以及一些探索性试验。

表 B.1 PI ML 中探讨的不同技术及其选项和缩写

机器学习技术	选项	缩写
空间梯度计算		
	自动微分	AD
	形函数导数	SF
更新损失项权重		
	GradNorm	GN
	ReLoBRaLo	RB
	均匀权重 [a]	UW
损失残差缩放		
	无量纲缩放	DS
	自定义缩放 [b]	CS
	均匀缩放 [a]	US
残差范数计算		
	L2 范数	L2
	均方值范数	MS

[a] 固定值为 1。[b] 调整尺度，使初始损失项几乎为 1，而不受其物理单位的影响。

表 B.2 使用 SF-GN- DS-L2 训练 3 个采样增量时，不同 GradNorm 设置对应的 L_2^{all}误差度量

试验名称	α_g	β_g	L_2^{all}
GN1	1.0	0.01	6.1×10^{-3}
GN2	0.1	0.01	4.1×10^{-3}
GN3	1.5	0.01	7.8×10^{-3}
GN4	1.0	0.10	2.3×10^{-2}
GN5	1.0	0.05	2.1×10^{-2}

根据表 B.2 和表 B.3中的结果，我们选取了 GN1 和 RB2 试验的调优参数，以对表 B.1中列出的多种方法变体进行更详细的参数研究。GN2 的误差相对低于 GN1；然而，其训练不稳定性更明显，因此未被选作进一步分析的对象。图B.1 展示了参数研究结果的一些见解。值得注意的是，所有采用单位均匀尺度（US）的组合在训练过程中均产生了“nan”值，或在训练结束后误差值很高；因此，图中排除了所有采用均匀缩放的组合。

表 B.3 使用 SF-RB- DS-L2 训练 3 个采样增量时，不同 ReLoBRaLo 设置的 L_2^{all}误差度量。“–”表示该方法产生了“NaN”值，无法完成训练。

试验名称	α_r	θ_r	$E[\rho_r]$	L_2^{all}
RB1	0.99	1.0×10^{-5}	0.99	–
RB2	0.99	5.0×10^{-1}	0.9	7.3×10^{-3}
RB3	0.999	1.0×10^{-1}	0.9999	1.0
RB4	0.999	1.0×10^{-2}	0.9999	–
RB5	0.99	1.0×10^{-2}	0.99	–
RB6	0.999	1.0	0.99	1.0

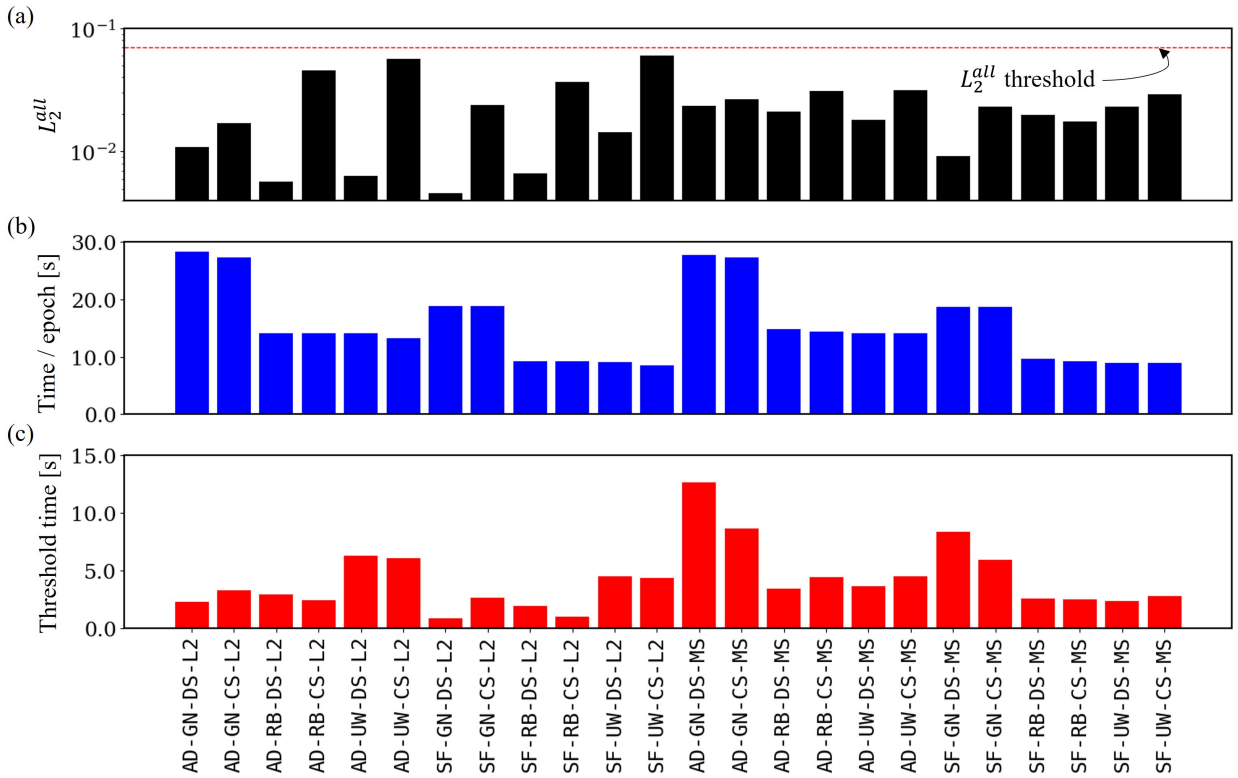


图 B.1 参数研究摘要，考察了多种训练变体：a) 训练 40000 个训练轮次后的最终 L_2^{all}，b) 每个训练轮次的平均计算时间，c) 达到 7×10^{-2}阈值精度所需的计算时间

就给定训练轮次数后的准确度而言，AD-RB-DS-L2、AD-UW-DS-L2、SF-GN-DS-L2 和 SF-RB-DS-L2 的误差值最低，如图 B.1a 所示。此外，结果（见图 B.1b）表明，自动微分（AD）与 GradNorm（GN）的组合在每训练轮次耗时方面成本最高。另一方面，形函数导数（SF）与 ReLoBRaLo（RB）的组合成本最低。

有些组合的计算成本相对较高；不过，它们的效率很高，能够在较少的训练轮次内达到较低的误差值。通过比较所有组合达到特定准确度所需的时间，可以推断其效率，如图 B.1c 所示。阈值的设定较为合理，

设定的值已满足所有被比较的组合。SF-GN-DS-L2、SF-RB-DS-L2 和 SF-RB-CS-L2 这几种组合效率最高，因为它们以最短时间达到了阈值误差。综合考虑不同组合的最终准确度和效率，我们选择 SF-GN-DS-L2 和 SF-RB-DS-L2，对注入示例进行更详细的训练/验证过程，具体说明见第 6.1.2节。

B.1.2 使用 *SF-GN-DS-L2* 训练的补充图

采用的自定义学习率见图 B.2a。该组合的训练损失和验证损失（$\mathcal{L}(\phi)$）如图 B.2b 所示，两条曲线均呈现出常见的误差下降趋势，没有过拟合迹象。

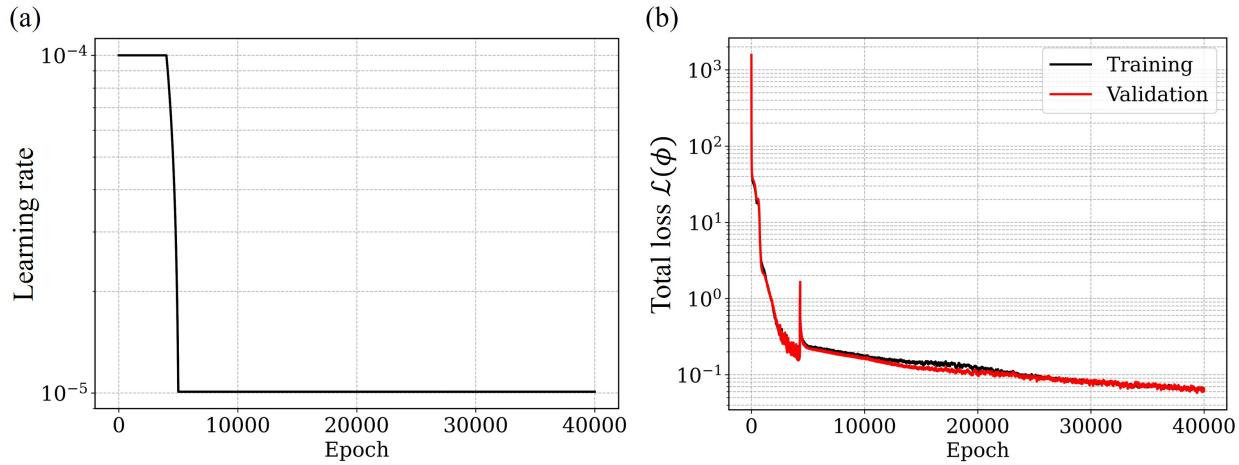


图 B.2 使用 SF-GN-DS-L2 时二维注入示例的训练指标：（a）自定义学习率，（b）训练损失和验证损失 $\mathcal{L}(\phi)$。

B.1.3 使用 *SF-RB-DS-L2* 的训练结果

图 B.3 展示了使用 SF-RB-DS-L2 对第6.1.1 节和第 6.1.2节中讨论的二维注入问题进行训练时的损失项及其自适应权重。除损失项 $\mathcal{L}_3^{bottom}$外，所有误差项在训练过程中均成功降低；该损失项对应于注入通量边界条件。自适应权重算法试图通过增加分配给该项的权重来解决这一问题；然而，这并未影响该损失项。结合这些结果以及附录 B.1.1中的结果，可以得出一个总体结论：ReLoBRaLo 算法可能非常高效；然而，每当训练过程发生变化时，都需要进行微调。

B.1.4 其他 *IFENN-deal.II* 和 *IFENN-GEOS* 结果。

为进一步研究结果，我们从图中选取一个样本增量。12以绘制相对误差指标 ϵ_{rel}，其定义见式（23）。第七个测试载荷工况（LC7）在最后一个时间增量处的 L_2^{inc}误差值较高，因此选取该增量绘制逐点相对误差（ϵ_{rel}）。图B.5 和图B.6 分别展示了 LC7 在第 50 个时间增量处 IFENN-deal.II 和 IFENN-GEOS 的响应（y_{pred}）。IFENN 结果与耦合有限元法结果高度吻合，在5% 的误差范围内覆盖了

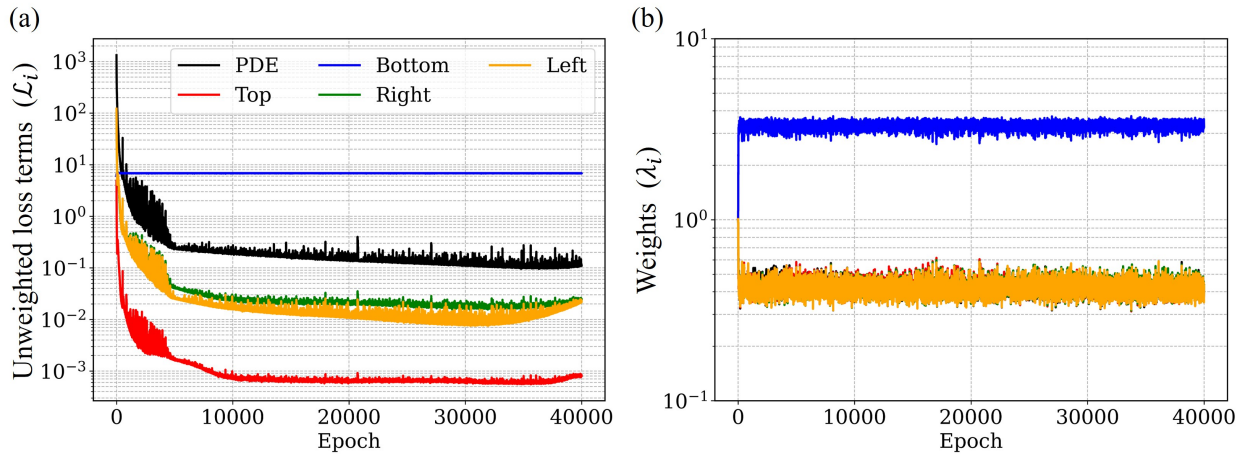


图 B.3 使用 SF-RB-DS-L2 的二维注入示例训练指标：(a) 未加权损失项 $\mathcal{L}_i$，(b) 相应的自适应权重 λ_i。仅包含偏微分方程残差和边界条件损失；未使用数据损失。

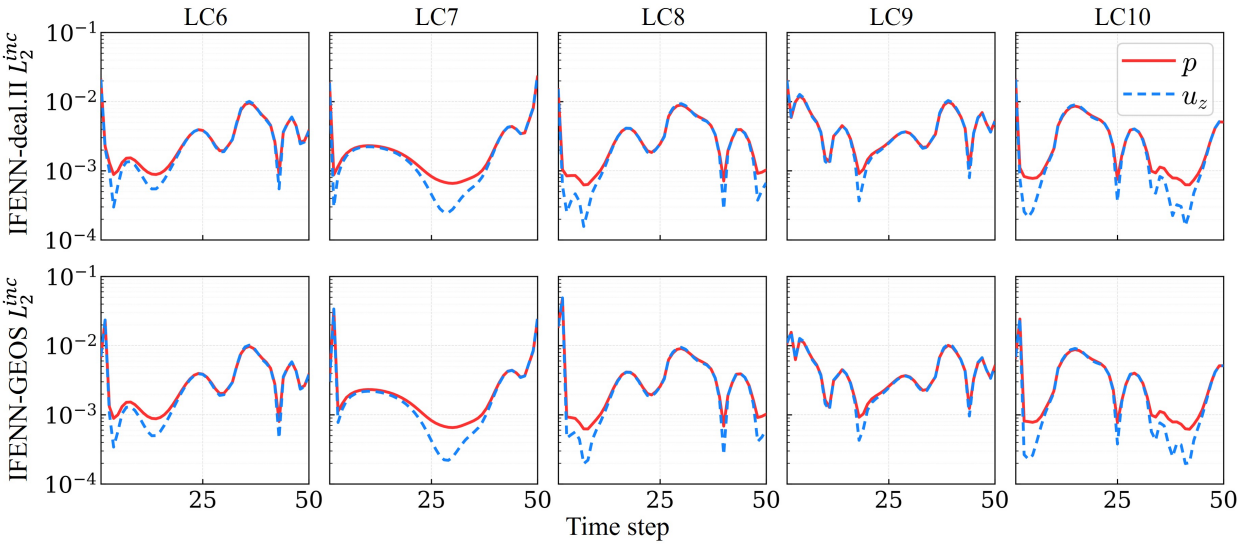


图 B.4 使用 IFENN-deal.II 和 IFENN-GEOS 对二维注入问题测试载荷工况 LC6-LC10 的压力 (p) 和 z方向位移 (u_z) 的测试误差 L_2^{inc}。

大部分区域。相对误差 ϵ_{rel} 通过将误差除以真实值 y_{true}计算得出。因此，当真实值接近零时，相对误差值会被放大。

B.2 渗透率可变 K_p 且时间增量为 Δt 的二维注入示例

模型参数列于表 B.4。 K_p-Δt 广义 PIF-MIONet 使用 Adam 优化器和自定义学习率进行 80,000 个训练轮次，如图 B.7a 所示。图 B.7b 中的训练和验证损失曲线呈现一致的下降趋势，且在训练后期振荡较少，这与学习率逐渐降低的情况一致。未加权损失项及其自适应权重见图 B.8。

图 B.9 给出了第 7 个测试荷载工况的其他测试结果，显示出一致的

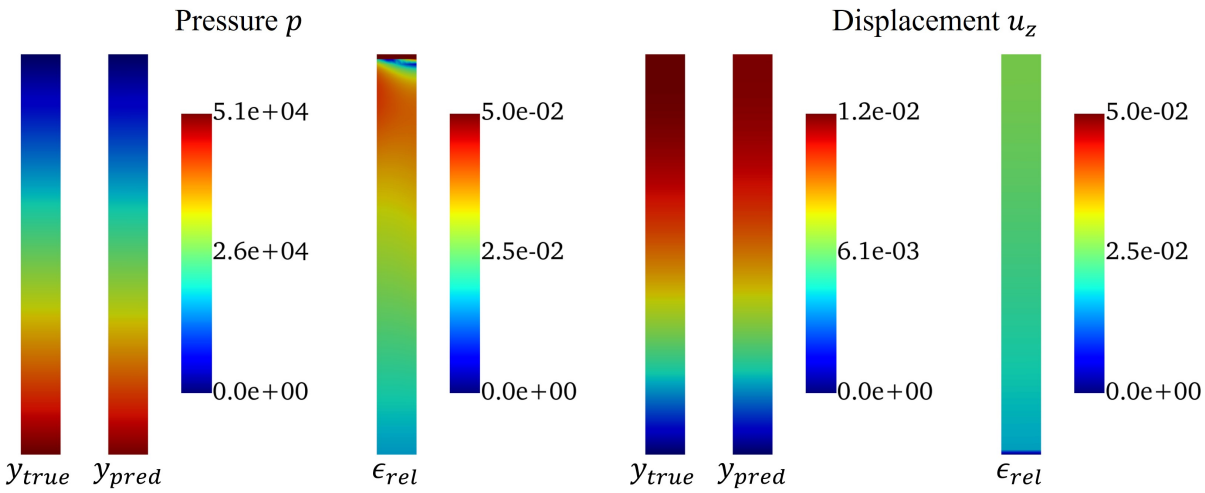


图 B.5 二维注入示例中第 7 个测试荷载工况在第 50 个时间步时，使用全耦合有限元求解器(y_{true})和 IFENN-deal.II (y_{pred})得到的解及对应的相对误差(ϵ_{rel})。

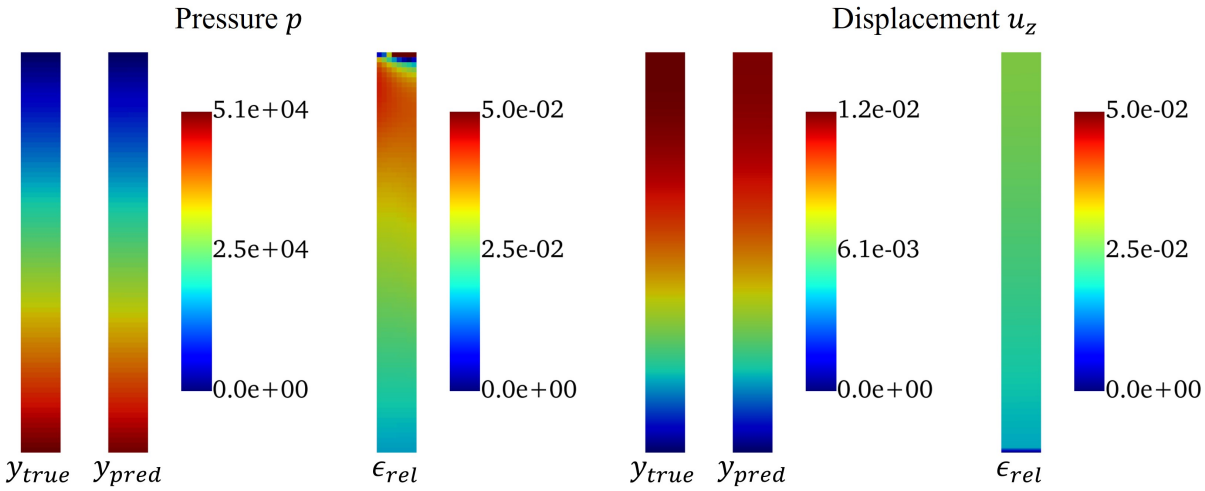


图 B.6 二维注入示例中第 7 个测试荷载工况在第 50 个时间步时，使用全耦合有限元求解器(y_{true})和 IFENN-GEOS (y_{pred})得到的解及对应的相对误差(ϵ_{rel})。

在不同渗透率和时间离散化配置下，IFENN-deal.II 与全耦合有限元求解器之间的一致性，进一步支持了图 15中所述的观察结果。

B.3 三维注入示例

B.3.1 训练

二维案例中使用的优化器设置保持不变，随后进行2000轮次的 L-BFGS 优化，学习率为 1.0×10^{-2}，如图 B.10a 所示。图 B.10b 中的训练和验证历史表明，在 Adam 和 L-BFGS 阶段之后，损失均持续下降。

Table B.4 针对变量 K_p 和 Δt训练的广义 PIF-MIONet 模型的超参数

网络	输入尺寸	隐藏层尺寸[a]	N_{ch}[b]	输出尺寸
分支 #1	N_q =1	[8,16]	-	16
分支 #2	N_K =1	[8,8]	-	8
分支 #3	N_t =1	[8,8]	-	8
分支 #4	N_ε =205	[32]	16	8
分支 #5	N_p =205	[32]	16	8
融合	N_{fi} =48	[48,48]	-	48
主干	N_c =2	[48,48]	-	48

[a] 尺寸数组的长度就是隐藏层的数量。[b] 在输入层的输出上实施组归一化，使用 N_{ch} 个通道。“-”表示不考虑组归一化。

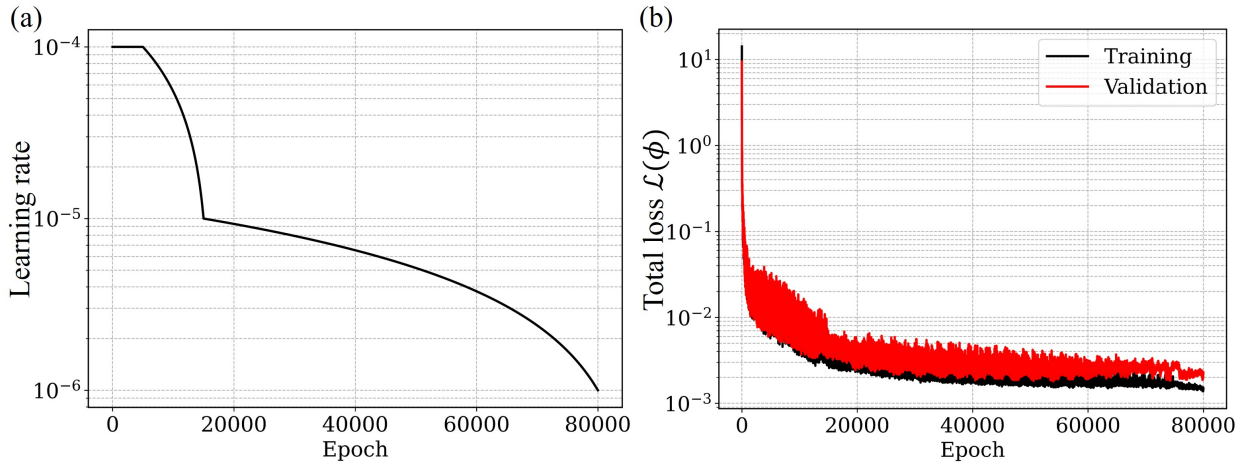


图 B.7 考虑可变渗透率 K_p 和时间增量 Δt时，二维注入示例的训练指标：(a) 自定义学习率；(b) 训练损失和验证损失 $\mathcal{L}(\phi)$。

B.3.2 补充结果

图 B.11和图 B.12展示了使用 IFENN-GEOS 的三维注入示例的补充结果。图 B.11 中的测试载荷工况误差总体较低，而图 B.12 中的预测结果表明，最大的逐点相对误差出现在响应幅值接近零的区域。

B.4 开挖问题

B.4.1 关于训练的补充讨论

表 B.6 列出了不同训练试验及其验证损失指标的汇总（L_2^{all}），其计算方式如式（25）所示。结果表明，尽管可训练参数的数量几乎相同，三主干模型的准确度仍高于单主干模型。此外，加入数据损失项对所有情况都有积极影响。最后，在复杂几何中使用自动微分计算空间导数会导致不稳定收敛和训练失败。

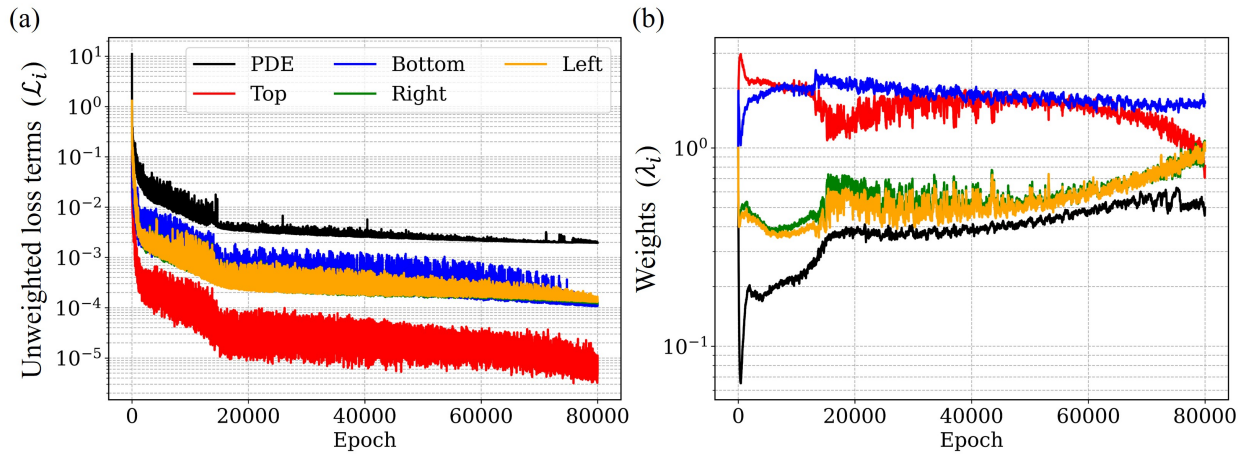


图 B.8 考虑可变渗透率 K_p 和时间增量 Δt的二维注入示例的训练指标：(a) 未加权损失项 $\mathcal{L}_i$；(b) 对应的自适应权重 λ_i。仅包含偏微分方程残差和边界条件损失；未使用数据损失。

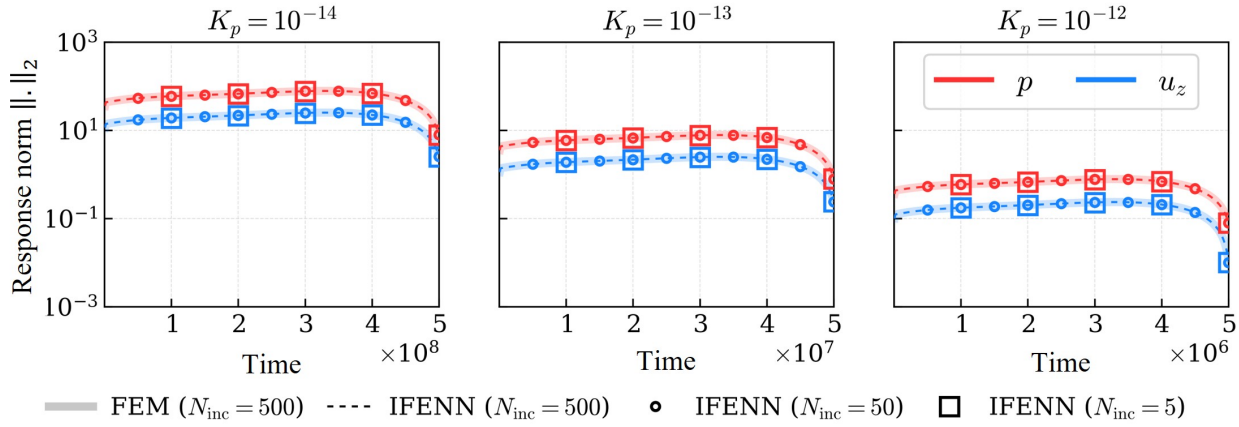
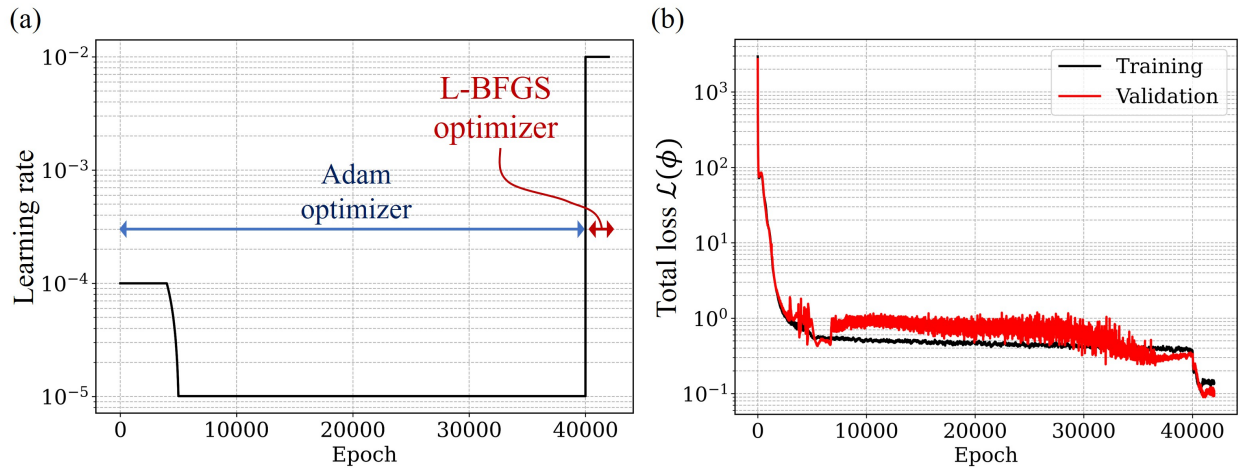


图 B.9 在二维注入问题中，针对第 7 个测试载荷工况，使用全耦合有限元求解器和 IFENN- deal.II 求得的压力范数 $\|p\|_2$ 和位移范数 $\|u_z\|_2$，对应 9 种不同设置（3 个渗透率值 $\times$ 3 和时间增量值）。压力和位移的响应范数单位分别为兆帕和米。

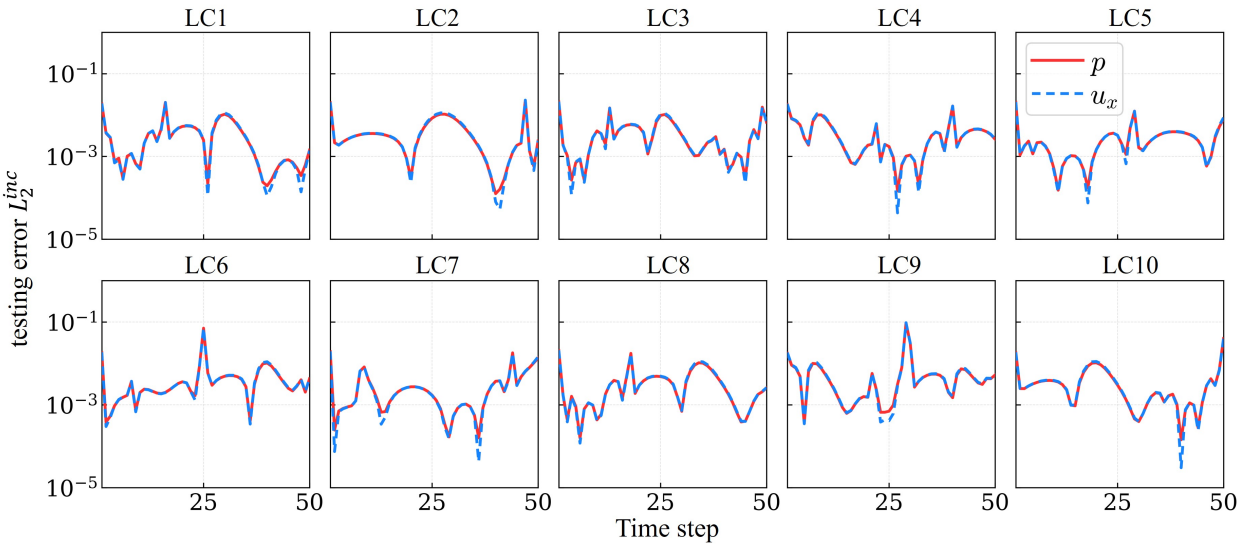
表 B.5 针对二维开挖示例训练的单主干 PIF-MIONet 模型的超参数

网络	输入尺寸	隐藏层尺寸[a]	N_{ch}[b]	输出尺寸
分支 #1	N_q =1	[64,64]	-	64
分支 #2	N_ε =499	[64,64]	-	64
分支 #3	N_p =499	[64,64]	-	64
融合	N_{fi} =192	[192]	-	256
主干 #1	N_c =2	[256,256,256]	-	256

[a] 尺寸数组的长度就是隐藏层的数量。[b] 在输入层的输出上采用包含 N_{ch} 个通道的组归一化。“-”表示不考虑组归一化。



图B.10 使用AD-GN-DS-L2求解三维注入示例的训练指标：(a) 自定义学习率，(b) 训练损失和验证损失 $\mathcal{L}(\phi)$



图B.11 使用IFENN-GEOS求解三维注入问题时，所有测试载荷工况（LC）的压力 L_2^{inc}和z方向位移（p）的归一化测试误差u_z。

图中进一步展示了采用自动微分时的训练行为。B.13，其中损失项持续振荡，且收敛不稳定。一个可能的原因是，如图B.16 所示并在第6.4.2节讨论的，不透水壁尖端附近存在奇异行为。采用自动微分时，偏微分方程残差直接根据网络输出计算，导致壁尖端角点附近的配点处出现较大的残差。相比之下，SF导数通过节点预测值的九节点宏单元插值（第5.3节）计算，提供了类似于有限元离散的基于网格的正则化。这种方法对于奇异解仍能保持稳定。这可能解释了模型M1-AD（$L_2^{all} = 4.6 \times 10^{1}$）为何发散，而其他方面完全相同、采用SF导数的模型M1（$L_2^{all} = 3.7 \times 10^{-3}$）却不会。对于注入算例（附录B.1.1 和第6.3节），基于自动微分的训练效果良好；这些算例中的解是平滑的，因此支持了这一解释。这一观察也与通过有限元插值来避免逐点计算二阶导数的方法相一致

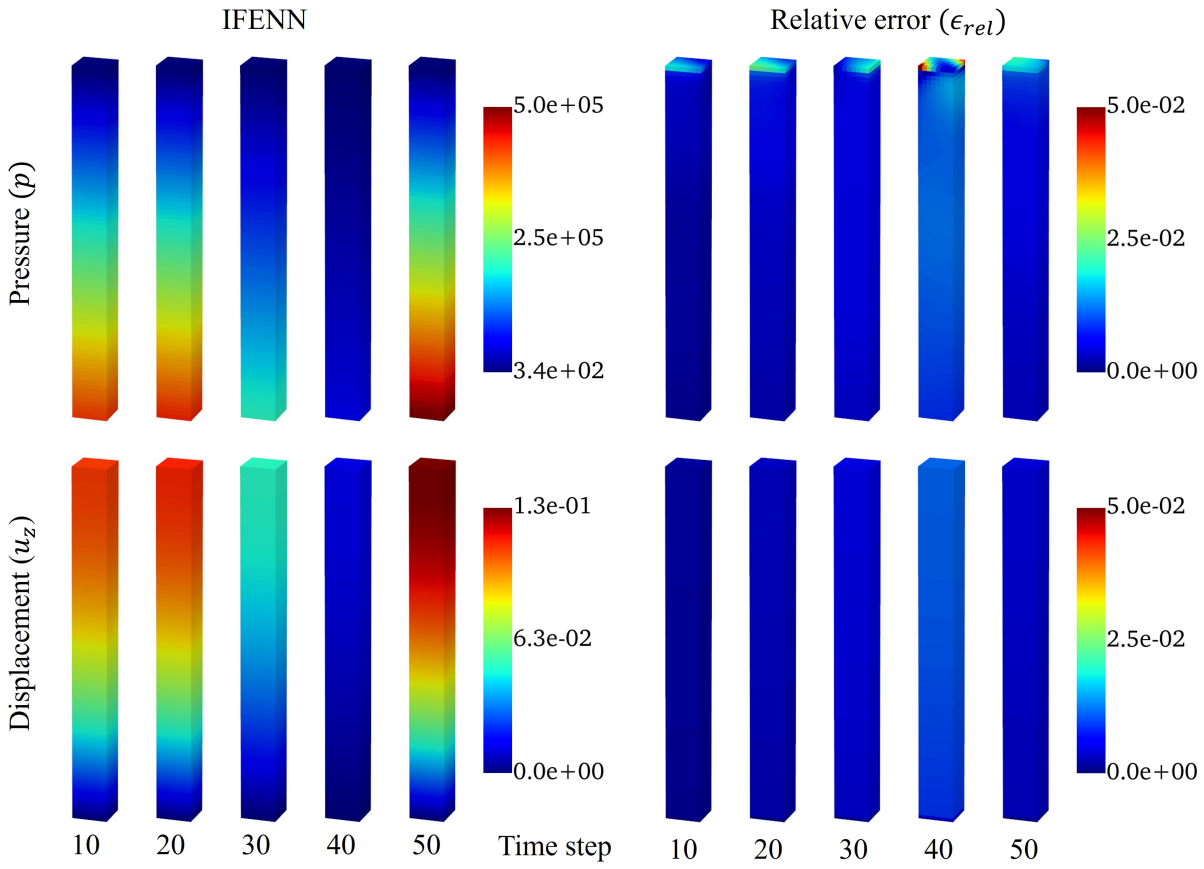


图 B.12 在第 7 个测试荷载工况的不同时间步下，使用 IFENN-GEOS 获得的压力 p 和位移 u_z，以及相应的相对误差（ϵ_{rel}）（三维注入示例）。

，如.[106]

图B.14比较了不同训练配置下的自适应权重，突出了损失项和网络架构的变化如何影响加权过程。这些观察结果补充了表B.6 中的准确度对比，并支持对参考训练配置的选择。

B.4.2 测试载荷工况及其他 *IFENN* 结果

开挖示例的测试载荷工况如图 B.15所示。图 B.16 展示了开挖壁尖端附近的局部应变集中，突出了该问题的几何复杂性。使用四个子循环（$N_{sub} = 4$）时压力 (p) 以及 z 和 x方向位移的测试误差 (u_z, u_x)如图 B.17所示，而图 B.18 则给出了第 9 个测试荷载工况下 IFENN-deal.II 与全耦合有限元解之间的空间场对比。这些结果进一步证明了所提框架对于开挖问题的准确度。

B.4.3 使用多个子循环的影响

图 B.19 和图 B.20 分别展示了预测-多重校正子循环数量对垂直和水平位移误差的影响。与压力结果一致，

表 B.6 开挖问题中不同训练后的模型及其最终验证损失值（L_2^{all} ）.

模型 code	空间梯度*	数量主干的	$\mathcal{L}^{data}$	$\mathcal{L}^{right}$, $\mathcal{L}^{wall1}$, 以及 $\mathcal{L}^{wall2}$	验证损失 (L_2^{all})
M1	SF	3	✓	×	3.7×10^{-3}
M2	SF	3	✓	✓	4.5×10^{-3}
M3	SF	3	×	×	1.0×10^{-1}
M4	SF	3	×	✓	1.2×10^{-1}
M5	SF	1	✓	×	1.2×10^{-2}
M6	SF	1	✓	✓	1.8×10^{-2}
M7	SF	1	×	×	1.3×10^{-1}
M8	SF	1	×	✓	9.6×10^{-2}
M1-AD	AD	3	✓	×	4.6×10^{1}

* 空间梯度计算方法：AD 表示自动微分，SF 表示形函数导数。✓ 表示包含的项。× 表示排除的项。

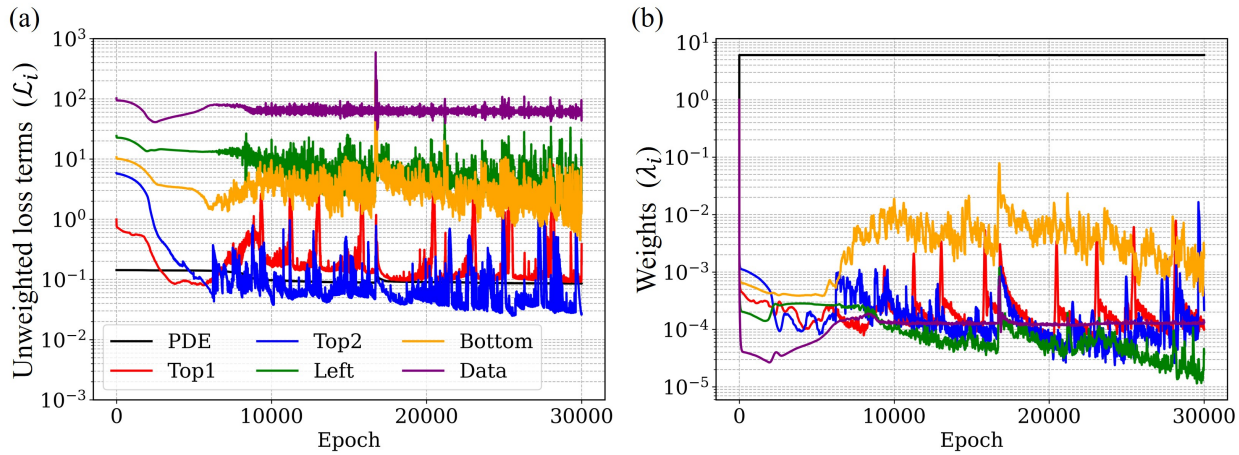


图 B.13 采用自动微分（AD）的开挖示例训练指标：(a) 未加权损失项$\mathcal{L}_i$，(b) 对应的自适应权重 λ_i。

图 24中，增加子循环数量主要能提高初始时间增量阶段的准确度，而在后续时间增量中，差异则不那么明显。

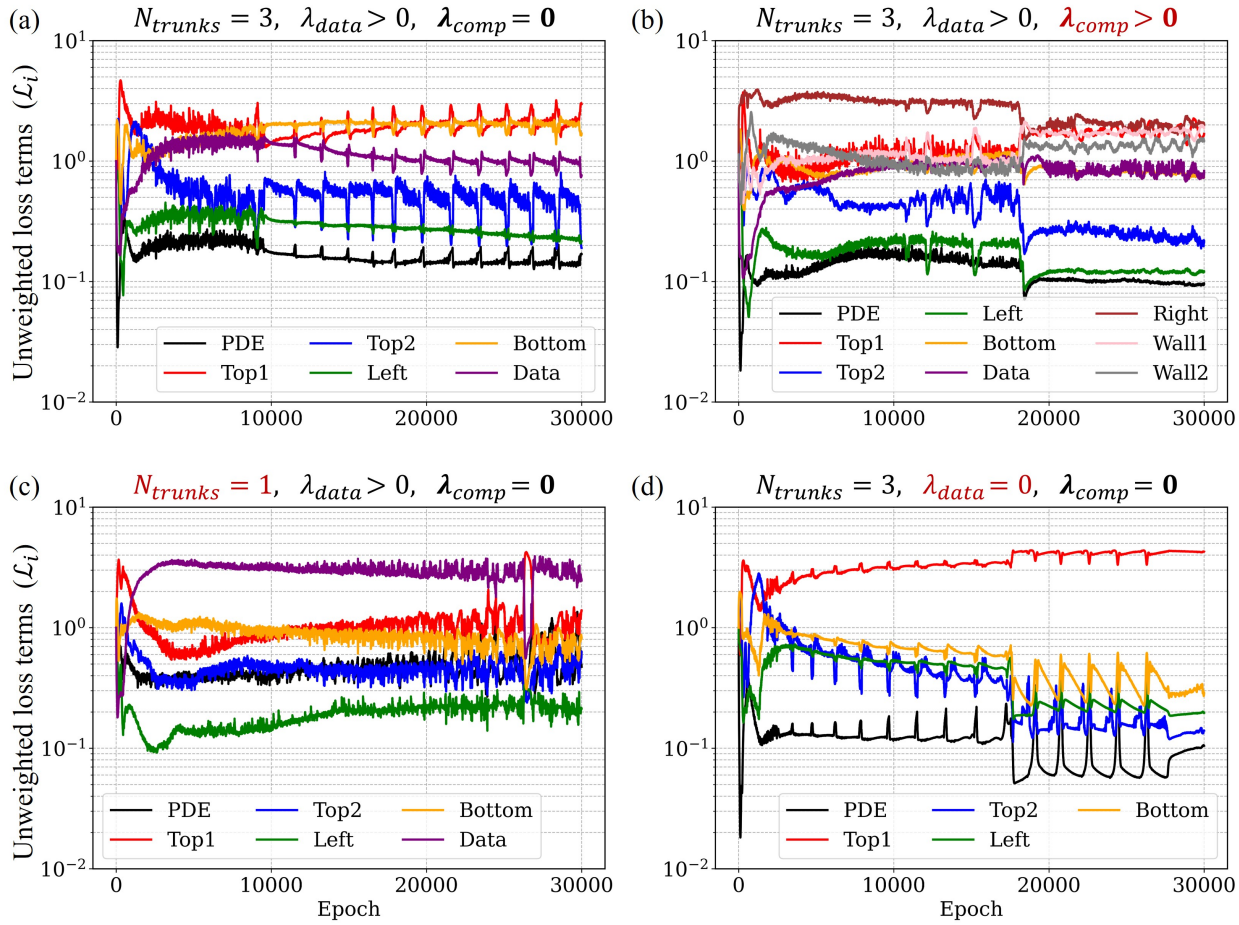


图 B.14 四种训练配置下开挖问题的自适应训练权重：(a) 包含 $\mathcal{L}^{data}$ ($\lambda_{data} > 0$) 和不包含 $\mathcal{L}^{comp}$ ($\boldsymbol{\lambda}_{comp} = \mathbf{0}$) 的三个主干参考模型，(b) 同时包含 $\mathcal{L}^{data}$($\lambda_{data} > 0$) 和 $\mathcal{L}^{comp}$ ($\boldsymbol{\lambda}_{comp} > \mathbf{0}$) 的三个主干，(c) 包含 $\mathcal{L}^{data}$ ($\lambda_{data} > 0$) 和不包含 $\mathcal{L}^{comp}$ ($\boldsymbol{\lambda}_{comp} = \mathbf{0}$) 的一个主干，以及 (d) 两项均不包含的三个主干（$\lambda_{data} = 0$ 和 $\boldsymbol{\lambda}_{comp} = \mathbf{0}$）。

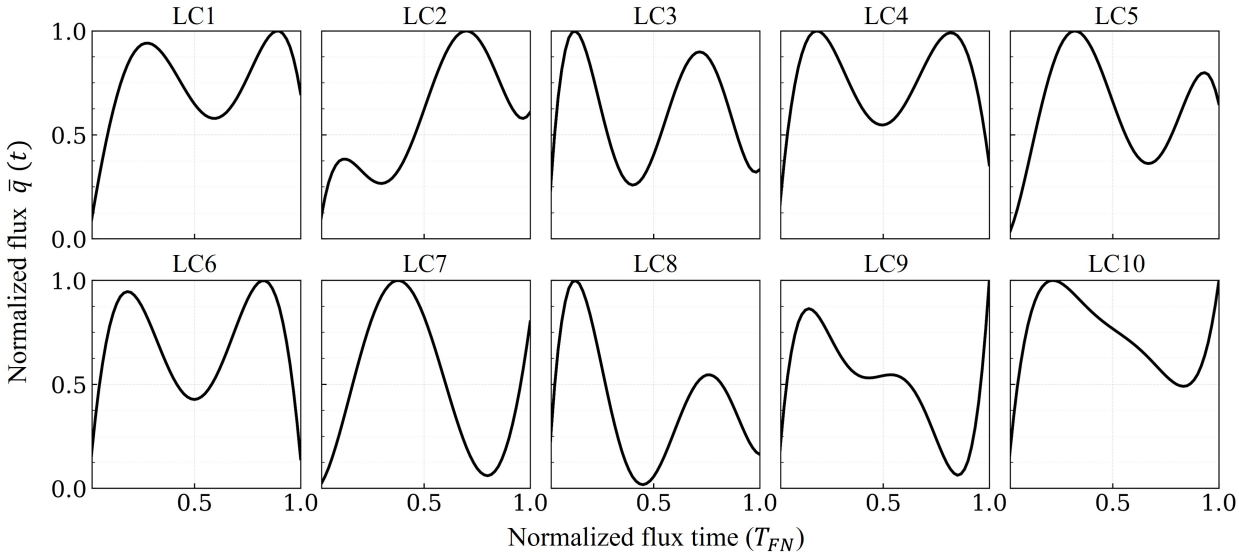


图 B.15 所有测试载荷工况（LC）的开挖问题中，归一化通量幅值与归一化通量时间（T_{FN}）的关系。通量和时间分别使用各自的最大值进行归一化。

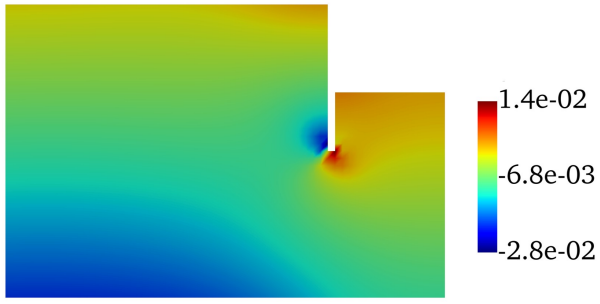


图 B.16 使用耦合有限元求解器计算的开挖示例第 9 个测试载荷工况第30个增量的应变迹值 tr(ε)。

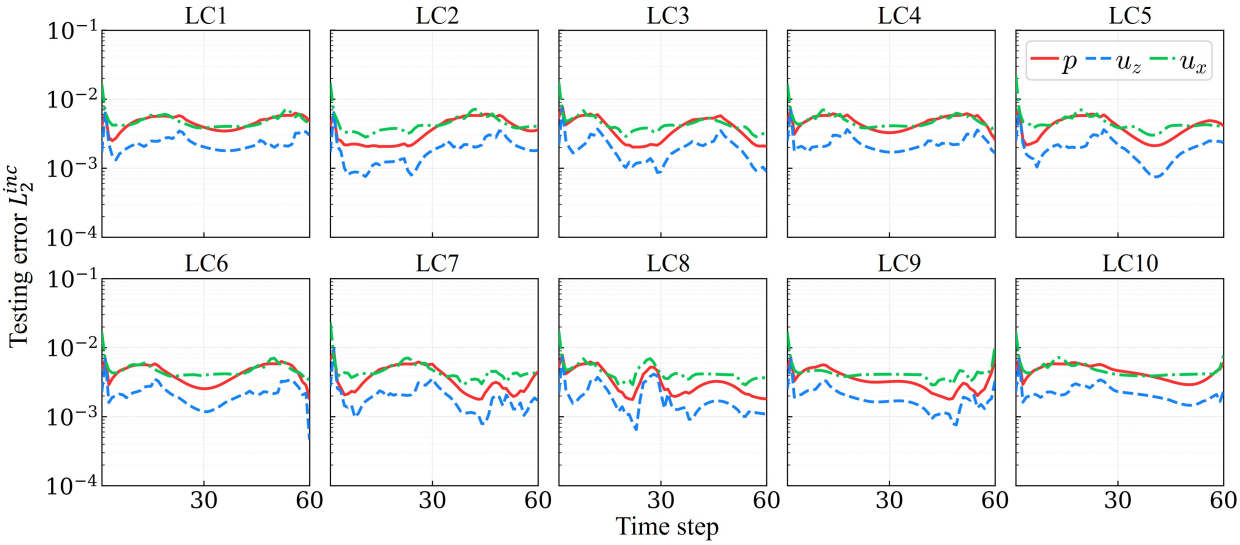


图 B.17 使用 IFENN-deal.II 和四个预测-多重校正子循环 ($N_{sub} = 4$) 得到的开挖问题所有测试载荷工况 (LCs) 的压力 (p) 以及 z 和 x 方向位移 (u_z, u_x) 的归一化测试误差 L_2^{inc}。

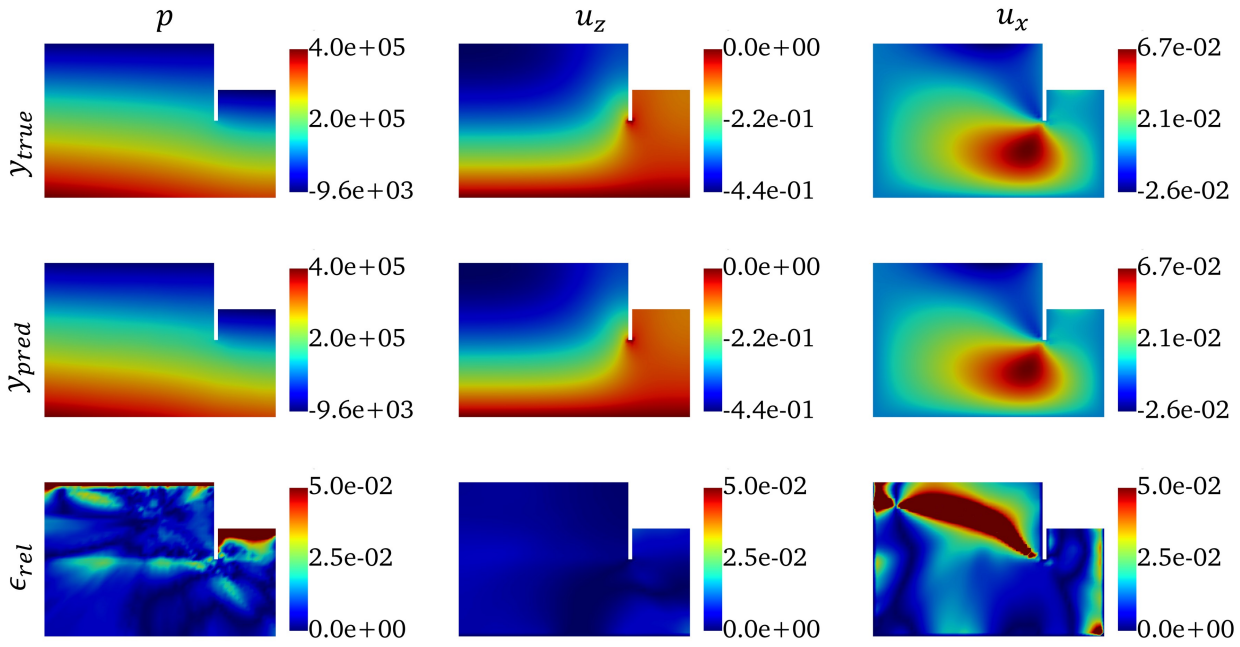


图 B.18 开挖示例中第 9 个测试荷载工况在第 60 个时间步得到的全耦合有限元求解器 (y_{true}) 和 IFENN-deal.II (y_{pred}) 的解，以及相应的相对误差 (ϵ_{rel})。

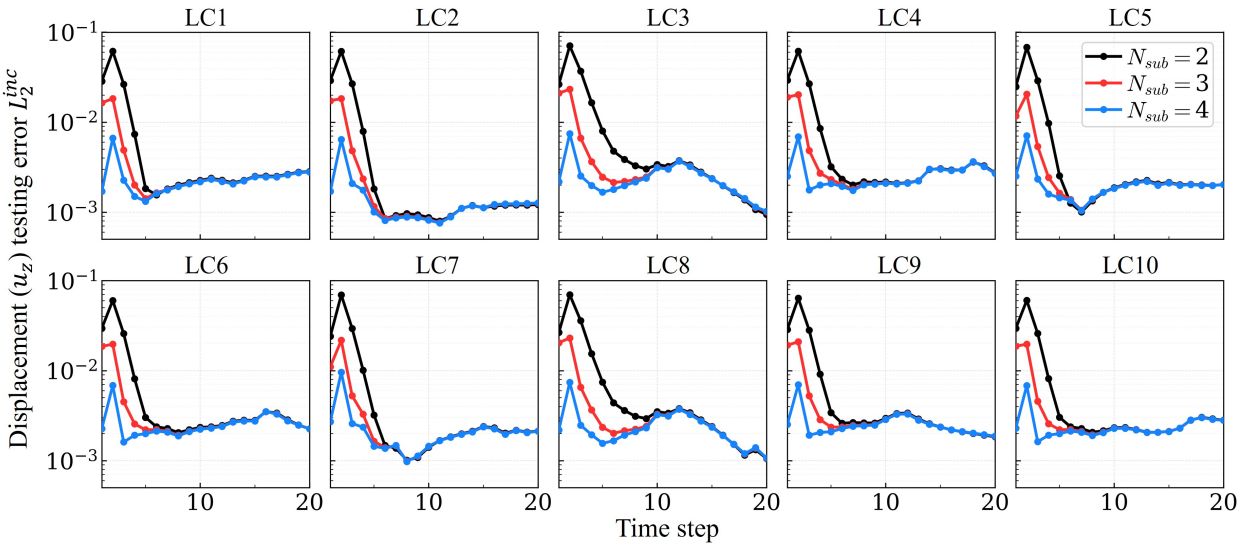


图 B.19 采用不同预测-多重校正子循环次数（$N_{sub} = 2$、3、4）的 IFENN-deal.II 对开挖问题所有测试载荷工况（LCs）中 z方向位移（u_z）的归一化测试误差 L_2^{inc}。

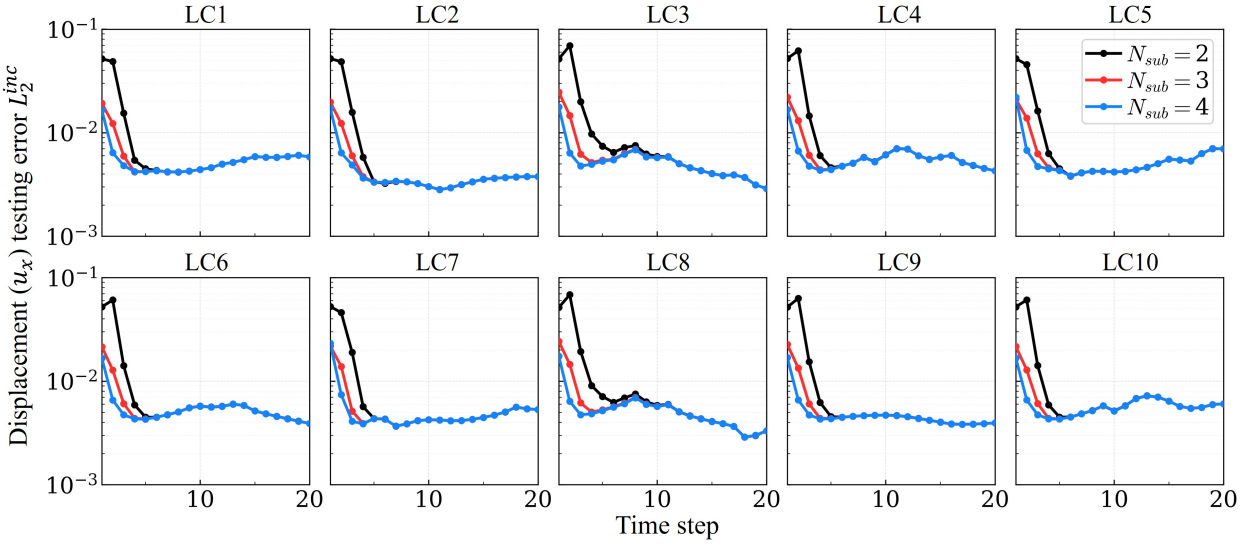


图 B.20 采用不同预测-多重校正子循环次数（$N_{sub} = 2$、3、4）的 IFENN-deal.II 对开挖问题所有测试载荷工况（LCs）中 x方向位移（u_x）的归一化测试误差 L_2^{inc}。

参考文献

1 S. Moradi, J. A. Huisman, H. Vereecken, H. Class, 《水力-力学模型中用于边坡稳定性评估的不同耦合与建模策略比较》, Water 16 (2) (2024) 312。doi:10.3390/w16020312。2 P. Maneepong, C. Leelasukseree, T. Pipatpongsa, P. Sangkhaphan, T. Maneewat, A. Chai- wan, 《泰国 Mae Moh 矿中继斜坡控制破坏机制的岩质边坡稳定性与水力-力学耦合评估》, International Journal of Rock Mechanics and Mining Sciences 191 (2025) 106132。doi:10.1016/j.ijrmms. 2025.106132。3 L. Ding, Z. Wang, B. Liu, J. Lv, Y. Wang, 《钻孔稳定性分析：基于孔隙弹性理论、考虑层状地层各向异性渗透率影响的新模型》, Journal of Natural Gas Science and Engineering 69 (2019) 102932。doi: 10.1016/j.jngse.2019.102932。4 L. Guo, M. Fahs, H. Hoteit, R. Gao, Q. Shao, 《裂隙多孔介质中渗流诱发固结的不确定性分析》, Computer Modeling in Engineering & Sci- ences 129 (1) (2021) 279–297。doi:10.32604/cmes.2021.016619。5 D. M. Potts, L. Zdravkovi´c, 《岩土工程中的有限元分析：理论》, Thomas Telford，伦敦，1999年。6 K. Shwetank, D. Deb, R. Pramanik, 《用于模拟流体在可变形多孔介质中流动的无网格（光滑粒子流体动力学）与基于网格（有限差分法）耦合流程》, International Journal of Rock Mechanics and Mining Sciences 170 (2023) 105494。doi:10.1016/j.ijrmms. 2023.105494。7 Y. Liu, X. Rao, H. Zhao, W. Zhan, Y. Xu, Y. Liu, 《基于广义有限差分法的无网格分析：耦合两相多孔介质流与地质力学》, Engineering Anal- ysis with Boundary Elements 146 (2023) 184–203。doi:10.1016/j.enganabound. 2022.10.011。8 K. M. Terekhov, Y. V. Vassilevski, 《耦合地下流动问题的有限体积法，II：孔隙弹性》, 《计算物理学杂志》 462 (2022) 111225。doi: 10.1016/j.jcp.2022.111225。9 W. M. Boon, S. E. Gasda, T. H. Sandve, S. Tveit, 《通过将 OPM Flow 与两点应力近似有限体积法耦合求解 Biot 孔隙弹性》, Computational Geosciences 30 (3) (2026) 23。doi:10.1007/s10596-026-10419-4。10 L. He, L. Jing, M. Feng, 《用于低渗透率双场孔隙弹性的全新稳定化混合有限元法》, Applied Mathematics and Computation 494 (2025) 129285。doi:10.1016/j.amc.2025.129285。11 Z. Ge, S. Pei, Y. Yuan, 《用于非线性孔隙弹性模型的两层多物理场有限元法》, Computers & Mathematics with Applications 124 (2022) 63–73。doi:10.1016/j.camwa.2022.08.021。12 Z. Moradinia, H. Vandierendonck, A. Murphy, 《探索多物理场模拟中的速度–精度权衡》, Meccanica doi:10.1007/s11012-024-01867-2。13 S. Meinecke, M. Selig, F. K¨oster, A. Knorr, K. L¨udge, 《多物理场模拟的数据驱动加速》, Machine Learning: Science and Technology 5 (4) (2024) 045011。doi:10.1088/2632-2153/ad7572。

14 X. Lyu, W. Wang, D. Voskov, P. Liu, L. Chen, 《地下能源储存中多相流与反应性质量传输的多尺度建模：综述》, Advances in Geo-Energy Research 15 (3) (2025) 245–260。doi:10.46690/ager.2025.03.07。15 Z. P. Du, Q. N. Fan, B. Wu, H. H. Zhang, 《利用嵌套网格应对多物理场模拟中的空间多尺度挑战》, IEEE Antennas and Wireless Propagation Letters 25 (6) (2026) 2491–2495。doi:10.1109/LAWP.2026.3684236。16 H. S. Tang, R. D. Haynes, G. Houzeaux, 《流体流动仿真中的区域分解方法综述：概念、算法与应用》, Archives of Computational Methods in Engineering 28 (3) (2021) 841–873。doi:10.1007/ s11831-019-09394-0。17 C. J. Permann, D. R. Gaston, D. Andrˇs, R. W. Carlsen, F. Kong, A. D. Lindsay, J. M. Miller, J. W. Peterson, A. E. Slaughter, R. H. Stogner, R. C. Martineau, 《MOOSE：实现大规模并行多物理场模拟》, SoftwareX 11 (2020) 100430。doi: 10.1016/j.softx.2020.100430。18 G. Castellazzi, A. D'Altri, S. De Miranda, H. Emami, L. Molari, F. Ubertini, 《用于多孔建筑材料盐结晶诱发损伤的交错式多物理场框架》, Construction and Building Materials 304 (2021) 124486。doi:10.1016/j. conbuildmat.2021.124486。19 A. J. Stershic, C. R. D'Elia, L. L. Beghini, M. R. Hill, B. Clausen, D. K. Balch, M. Maguire, C. W. San Marchi, J. W. Foulk, A. A. Hanson, K. L. Manktelow, 《电阻锻焊的自适应重网格多物理场建模：残余应力场与测量过程的实验验证》, International Journal of Solids and Structures 306 (2025) 113112。doi:10.1016/j.ijsolstr.2024.113112。20 G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, 《物理信息机器学习》, Nature Reviews Physics 3 (6) (2021) 422–440。doi:10. 1038/s42254-021-00314-5。21 P. Sharma, W. T. Chung, B. Akoush, M. Ihme, 《流体力学中物理信息机器学习综述》, Energies 16 (5) (2023) 2343。doi:10.3390/ en16052343。22 M. Delpisheh, B. Ebrahimpour, A. Fattahi, M. Siavashi, H. Mir, H. Mashhadimoslem, M. A. Abdol, M. Ghorbani, J. Shokri, D. Niblett, K. Khosravi, S. Rahimi, S. M. Alirahmi, H. Yu, A. Elkamel, V. Niasar, M. Mamlouk, 《机器学习在多孔介质中的应用》, Journal of Materials Chemistry A 12 (32) (2024) 20717–20782。doi:10.1039/D4TA00251B。23 O. Fuks, H. A. Tchelepi, 《物理信息机器学习在多孔介质非线性两相传输中的局限性》, Journal of Machine Learning for Modeling and Computing 1 (1) (2020) 19–37。doi:10.1615/ JMachLearnModelComput.2020033905。24 C. G. Fraces, H. Tchelepi, 《多孔介质流动与输运的物理信息深度学习》, 载于：SPE Reservoir Simulation Conference，SPE，按需，2021年，第 D011S006R002 页。doi:10.2118/203934-MS。25 L. Shan, C. Liu, Y. Liu, Y. Tu, L. Deng, X. Hei, 《用于求解多孔介质偏微分方程的物理信息机器学习》, Advances in Geo-Energy Research 8 (1) (2023) 37–44。doi:10.46690/ager.2023.04.04。

26 Y. W. Bekele，孔隙弹性介质中流动与变形的物理信息深度学习，arXiv:2010.15426 [cs.CE]（2020年10月）。doi:10.48550/arXiv.2010.15426。27 Y. W. Bekele，物理信息深度学习在一维固结中的应用，《岩石力学与岩土工程学报》13（2）（2021年）420–430。doi:10.1016/j.jrmge.2020.09.005。28 E. Haghighat、D. Amini、R. Juanes，使用应力分裂序贯训练的多相孔隙弹性物理信息神经网络仿真，《计算机方法：应用力学与工程》397（2022年）115141。doi:10.1016/j.cma.2022.115141。29 C. Millevoi、N. Spiezia、M. Ferronato，耦合水力-孔隙力学问题的物理信息神经网络训练，《计算物理学杂志》516（2024年）113299。doi:10.1016/j.jcp.2024.113299。30 S. Roy、C. Annavarapu、P. Roy、D. M. Valiveti，异质孔隙弹性介质的物理信息神经网络，《International Journal for Computational Methods in Engineering Science and Mechanics》26（2）（2025年）187–207。doi:10.1080/15502287.2024.2440420。31 S. Cuomo、V. S. Di Cola、F. Giampaolo、G. Rozza、M. Raissi、F. Piccialli，通过物理信息神经网络开展科学机器学习：现状与未来，《科学计算杂志》92（3）（2022年）88。doi:10.1007/s10915-022-01939-z。32 K. Luo、J. Zhao、Y. Wang、J. Li、J. Wen、J. Liang、H. Soekmadji、S. Liao，偏微分方程问题的物理信息神经网络：综述，《人工智能评论》58（10）（2025年）323。doi:10.1007/s10462-025-11322-7。33 R. Bischof、M. A. Kraus，物理信息深度学习的多目标损失平衡，《计算机方法：应用力学与工程》439（2025年）117914。doi:10.1016/j.cma.2025.117914。34 Z. Chen、V. Badrinarayanan、C.-Y. Lee、A. Rabinovich，GradNorm：深度多任务网络的梯度归一化与自适应损失平衡（2017年）。doi:10.48550/ARXIV.1711.02257。35 A. A. Heydari、C. A. Thompson、A. Mehmood，SoftAdapt：具有多部分损失函数的神经网络自适应损失加权技术（2019年）。doi:10.48550/ARXIV.1912.12355。36 K. Nath、A. Pandey、B. T. Susi、H. Babaee、G. E. Karniadakis，AMORE：用于刚性化学动力学的自适应多输出算子网络，《计算物理学杂志》562（2026年）115030。doi:10.1016/j.jcp.2026.115030。37 S. Wang、Y. Teng、P. Perdikaris，理解并缓解物理信息神经网络中的梯度流病态，《SIAM Journal on Scientific Computing》43（5）（2021年）A3055–A3081。38 D. Amini、E. Haghighat、R. Juanes，多孔介质中热-水-力过程的物理信息神经网络求解，《Journal of Engineering Mechanics》148（11）（2022年）04022070。doi:10.1061/(ASCE)EM.1943-7889.0002156。39 D. Amini、E. Haghighat、R. Juanes，使用物理信息神经网络对非等温多相孔隙力学进行反演建模，《计算物理学杂志》490（2023年）112323。doi:10.1016/j.jcp.2023.112323。

40 J. He、D. Abueidda、S. Koric、I. Jasiuk，图神经网络与基于形函数的梯度计算在深度能量法中的应用，《International Journal for Numerical Methods in Engineering》124（4）（2023年）864–879。doi:10.1002/nme.7146。41 P. Pantidis、H. Eldababy、D. Abueidda、M. E. Mobasher，采用时间卷积网络的I-FENN：加速非局部梯度损伤传播的加载历史分析，《计算机方法：应用力学与工程》425（2024年）116940。doi:10.1016/j.cma.2024.116940。42 Z. Wu、M. Li、C. He、B. Zhang、J. Ren、H. Yu、Q. Chen，使用解耦–耦合训练框架对化学反应器系统进行物理信息学习，《AIChE期刊》70（7）（2024年）e18436。doi:10.1002/aic.18436。43 R. Sharma、V. Shankar，面向算子学习的集成与混合专家DeepONet（2024年）。doi:10.48550/ARXIV.2405.11907。44 M. Vasilyeva、A. Tyrylgin，机器学习加速随机介质中孔隙弹性宏观参数预测，《Computers & Mathematics with Applications》84（2021年）185–202。doi:10.1016/j.camwa.2020.09.024。45 T. Kadeethum、D. O’Malley、J. N. Fuhg、Y. Choi、J. Lee、H. S. Viswanathan、N. Bouklas，使用条件生成对抗网络进行偏微分方程数据驱动求解与参数估计的框架，《自然·计算科学》1（12）（2021年）819–829。doi:10.1038/s43588-021-00171-3。46 Y. Choi、C. Liu、J. Macedo，面向固结的算子学习：DeepONet变体的架构比较，《计算机与岩土工程》194（2026年）108017。doi:10.1016/j.compgeo.2026.108017。47 Y. Ren、Z. L. Liu、M. Liu，In-poro：孔隙弹性的可微逆分析框架，《Journal of Geotechnical and Geoenvironmental Engineering》152（6）（2026年）04026030。doi:10.1061/JGGEFK.GTENG-14580。48 F. M. Amin、D. W. Abueidda、P. Pantidis、M. E. Mobasher，采用DeepONet的I-FENN：加速耦合多物理场问题中的仿真，《计算机方法：应用力学与工程》451（2026年）118645。doi:10.1016/j.cma.2025.118645。49 B. Raoni´c、R. Molinaro、T. De Ryck、T. Rohner、F. Bartolucci、R. Alaifari、S. Mishra、E. de B´ezenac，用于稳健且精确学习偏微分方程的卷积神经算子（2023年）。doi:10.48550/ARXIV.2302.01178。50 S. Sarkar、S. Chakraborty，用于求解不规则域和非结构网格上计算力学问题的时空谱图神经算子，《计算机方法：应用力学与工程》435（2025年）117659。doi:10.1016/j.cma.2024.117659。51 Z. Li、N. Kovachki、K. Azizzadenesheli、B. Liu、K. Bhattacharya、A. Stuart、A. Anandkumar，参数化偏微分方程的傅里叶神经算子（2020年）。doi:10.48550/ARXIV.2010.08895。52 L. Lu、P. Jin、G. Pang、Z. Zhang、G. E. Karniadakis，基于算子通用逼近定理的DeepONet学习非线性算子，《Nature Machine Intelligence》3（3）（2021年）218–229。doi:10.1038/s42256-021-00302-5。53 D. W. Abueidda、P. Pantidis、M. E. Mobasher，DeepOKAN：基于Kolmogorov-Arnold网络的深度算子网络，用于力学问题，《计算机方法：应用力学与工程》436（2025年）117699。doi:10.1016/j.cma.2024.117699。

54 L.Mandl, S. Goswami, L. Lambers, T. Ricken, 可分离物理信息 DeepONet：打破物理信息机器学习中的维数灾难，《计算机方法：应用力学与工程》434 (2025) 117586。doi:10.1016/j.cma.2024.117586。55 P. Jin, S. Meng, L. Lu, MIONet：通过张量积学习多输入算子，《SIAM Journal on Scientific Computing》44 (6) (2022) A3490–A3514。doi:10.1137/22M1477751。56 B. Ahmed, Y. Qiu, D. W. Abueidda, W. El-Sekelly, T. Abdoun, M. E. Mobasher, 用于高效预测结构动态响应的物理信息多输入算子，《人工智能工程应用》162 (2025) 112703。doi:10.1016/j.engappai.2025.112703。57 X. Wang, P. Li, K. Jia, S. Zhang, C. Li, B. Wu, Y. Dong, D. Lu, 用于相场水力压裂代理建模的 SPI-MIONet，《计算机方法：应用力学与工程》427 (2024) 117054。doi:10.1016/j.cma.2024.117054。58 Z. Jiang, M. Zhu, L. Lu, Fourier-MIONet：用于地质碳封存多相建模的傅里叶增强型多输入神经算子，《可靠性工程与系统安全》251 (2024) 110392。doi:10.1016/j.ress.2024.110392。59 C. Zhang, Z. Yang, X. He, L. Deng, 多模态智能：表征学习、信息融合与应用，《IEEE信号处理选题杂志》14 (3) (2020) 478–493。doi:10.1109/JSTSP.2020.2987728。60 R. S. Kiziltepe, J. Q. Gan, J. J. Escobar, 将特征融合与决策融合结合深度学习架构用于视频分类，《IEEE Access》12 (2024) 19432–19446。doi:10.1109/ACCESS.2024.3360929。61 D. Liu, X. Guo, Z. Liu, J. Tan, 一种图像生成器增强型深度算子网络，用于预测随机粗糙表面接触问题中的几何变形，《工程计算与信息科学杂志》25 (8) (2025) 081004。doi:10.1115/1.4068456。62 A. Peyvan, V. Kumar, G. E. Karniadakis, Fusion-DeepONet：一种用于几何相关高超声速与超声速流动的数据高效神经算子，《计算物理学杂志》544 (2026) 114432。doi:10.1016/j.jcp.2025.114432。63 P. Pantidis, H. Eldababy, C. M. Tagle, M. E. Mobasher，物理信息神经网络的误差收敛与工程指导的超参数搜索：迈向 I-FENN 性能优化，《计算机方法：应用力学与工程》414 (2023) 116160。64 A. Daw, J. Bu, S. Wang, P. Perdikaris, A. Karpatne，重新思考采样在物理信息神经网络中的重要性，arXiv 预印本 arXiv:2207.02338。65 J. N. Fuhg, N. Bouklas，基于概率机器学习和空间填充采样的物理信息数据驱动各向同性与各向异性本构模型，《计算机方法：应用力学与工程》394 (2022) 114915。66 X. Sun, B. Bahmani, N. N. Vlassis, W. Sun, Y. Xu，利用不确定性传播对深度学习材料定律的可解释因果关系进行数据驱动发现，《Granular Matter》24 (2022) 1–32。

67 M. Manfren, P. A. James, L. Tronchin，数据驱动的建筑能耗建模——通过可解释机器学习分析泛化潜力，《可再生能源与可持续能源评论》167 (2022) 112686。doi:10.1016/j.rser.2022. 112686。68 S. Markidis，新与旧：物理信息深度学习能否取代传统线性求解器？《大数据前沿》4 (2021) 669097。69 A. Krishnapriyan, A. Gholami, S. Zhe, R. Kirby, M. W. Mahoney，刻画物理信息神经网络中可能的失效模式，《神经信息处理系统进展》34 (2021) 26548–26560。70 Y. Shin, J. Darbon, G. E. Karniadakis，关于物理信息神经网络求解线性二阶椭圆型和抛物型偏微分方程的收敛性，《Communications in Computational Physics》28 (5) (2020) 2042–2074。doi:10.4208/cicp.OA-2020-0193。71 S. Mishra, R. Molinaro，物理信息神经网络近似一类偏微分方程反问题的泛化误差估计，《IMA Journal of Numerical Analysis》42 (2) (2022) 981–1022。doi:10.1093/imanum/drab032。72 S. Thel, L. Greve, B. Van De Weg, P. Van Der Smagt，介绍有限元方法集成网络（FEMIN），《计算机方法：应用力学与工程》427 (2024) 117073。doi:10.1016/j.cma.2024.117073。73 W. Wang, M. Hakimzadeh, H. Ruan, S. Goswami，时间推进神经算子–有限元耦合：人工智能加速物理建模，《计算机方法：应用力学与工程》446 (2025) 118319。doi:10.1016/j.cma.2025.118319。74 E. Zhang, A. Kahana, A. Kopaniˇc´akov´a, E. Turkel, R. Ranade, J. Pathak, G. E. Karniadakis，在偏微分方程数值求解器中融合神经算子与松弛方法，《Nature Machine Intelligence》6 (11) (2024) 1303–1313。doi:10.1038/s42256-024-00910-x。75 R. Quan, F. Liu, S. Cantero Chinchilla，将机器学习与有限元仿真相结合，实现对3D打印骨支架力学的可解释预测，《国际生物打印杂志》11 (6) (2025) 327。doi:10.36922/IJB025270257。ˇ 76 A. Daby-Seesaram, K. Skardov´a, M. Genet，有限元神经网络插值：第二部分——与适当广义分解相结合，用于非线性代理建模，《计算力学》77 (2) (2026) 521–546。doi:10.1007/ s00466-025-02676-4。77 R. He, M. Xia, P. Polak, B. Saleem, X. Yu, J. Pan，嵌入多物理场有限元分析的机器学习：在闪速烧结中的应用，《Journal of Intelligent Manufacturing》doi:10.1007/s10845-025-02725-y。78 K. Shukla, Z. Zou, C. H. Chan, A. Pandey, Z. Wang, G. E. Karniadakis，NeuroSEM：通过耦合物理信息神经网络与谱元来仿真多物理场问题的混合框架，《计算机方法：应用力学与工程》433 (2025) 117498。doi:10.1016/j.cma.2024.117498。79 W. Zhang, J. Li, PDNN：用于耦合热方程的纳维–斯托克斯方程的并行深度神经网络，《流体数值方法国际期刊》95 (4) (2023) 666–681。doi:10.1002/fld.5164。80 P. Pantidis, M. E. Mobasher，用于非局部连续体损伤力学的集成有限元神经网络（I-FENN），《计算机方法：应用力学与工程》404 (2023) 115766。doi:10.1016/j.cma.2022.115766。

81 P. Pantidis、L. Svolos、D. Abueidda、M. E. Mobasher，采用最少输入和广义几何-载荷处理的相场断裂集成有限元神经网络（IFENN），《计算机方法：应用力学与工程》448（2026年）118485。doi:10.1016/j.cma.2025.118485。 82 P. Pantidis、F. Amin、D. Abueidda、M. E. Mobasher，用于相场断裂起始与扩展可泛化建模的混合IFENN求解器（2026年6月）。doi: 10.48550/ARXIV.2606.27177。 83 D. W. Abueidda、M. E. Mobasher，基于物理信息时序卷积网络（PI-TCN）的热弹性I-FENN，《计算力学》74（6）（2024年）1229–1259。doi:10.1007/s00466-024-02475-3。 84 D. W. Abueidda、M. E. Mobasher，用于I-FENN热弹性的变分时序卷积网络，《计算机方法：应用力学与工程》429（2024年）117122。doi:10.1016/j.cma.2024.117122。 85 R. R. Settgast、R. M. Aronson、J. R. Besset、A. Borio、Q. M. Bui、T. J. Byer、N. Castel- letto、A. Citrain、B. C. Corbett、J. Corbett、P. Cordier、M. A. Cremon、C. M. Crook、M. Cusini、F. Fei、S. Frambati、J. Franc、A. Franceschini、M. Frigo、P. Fu、T. Gazzola、H. Gross、F. Hamon、B. M. Han、Y. Hao、R. Hasanzade、M. Homel、J. Huang、T. Jin、I. Ju、D. Kachuma、M. Karimi-Fard、T. Kim、S. Klevtsov、A. Lapene、V. A. P. Magri、A. Mazuyer、M. N’diaye、D. Osei-Kuffuor、S. Povolny、G. Ren、S. J. Semnani、C. S. Sher- man、M. Rey、H. A. Tchelepi、W. R. Tobin、P. Tomin、L. Untereiner、A. Vargas、S. Waziri、X. Wen、J. A. White、H. Wu，GEOS：面向地下应用的高性能可移植多物理场仿真框架，《开放源代码软件期刊》9（102）（2024年）6973。doi:10.21105/joss.06973。 86 O. Coussy，《孔隙力学》，Wiley，奇切斯特，2004年。 87 A. H.-D. Cheng，《孔隙弹性》，《多孔介质传输理论与应用》第27卷，施普林格国际出版社，查姆，2016年。doi:10.1007/ 978-3-319-25202-5。 88 L.-P. Yi、H. Waisman、Z.-Z. Yang、X.-G. Li，孔弹性介质中水力裂缝扩展的一致相场模型，《计算机方法：应用力学与工程》372（2020年）113396。doi:10.1016/j.cma.2020.113396。 89 Y. Chen、M. E. Mobasher、C. Gu、D. Zheng、H. Waisman，流体驱动断裂的热力学一致非局部损伤形式：适用于孔隙黏弹性介质，《岩土工程学报》17（11）（2022年）5321–5350。doi:10.1007/s11440-022-01557-x。 90 J. A. White、R. I. Borja，流动与地质力学耦合的稳定化低阶有限元及其在断层带瞬态过程中的应用，《计算机方法：应用力学与工程》197（49-50）（2008年）4353–4366。doi:10.1016/ j.cma.2008.05.015。 91 P. C. Africa、D. Arndt、W. Bangerth、B. Blais、M. Fehling、R. Gassm¨oller、T. Heister、L. Heltai、S. Kinnewig、M. Kronbichler、M. Maier、P. Munch、M. Schreter-Fleischhacker、J. P. Thiele、B. Turcksin、D. Wells、V. Yushutin，deal.II 程序库，9.6版，《数值数学期刊》32（4）（2024年）369–380。doi:10.1515/jnma-2024-0137。

92 A. Harandi、A. Moeineddin、M. Kaliske、S. Reese、S. Rezaei，适用于热-力耦合系统和非均质区域的物理信息神经网络混合形式，《International Journal for Numerical Methods in Engineering》125（4）（2024年）e7388。doi:10.1002/nme.7388。 93 L. Lu、X. Meng、S. Cai、Z. Mao、S. Goswami、Z. Zhang、G. E. Karniadakis，基于FAIR数据对两种神经算子（含实用扩展）进行全面且公平的比较，《计算机方法：应用力学与工程》393（2022年）114778。doi:10.1016/j.cma.2022.114778。 94 W. Xiang、H.-D. Tran、X. Yang、T. T. Johnson，神经网络控制系统的可达集估计：一种仿真引导方法，《Neural Networks and Learning Systems》32（5）（2021年）1821–1830。doi:10.1109/TNNLS.2020.2991090。 95 X. Yu、S. Hooten、Z. Liu、Y. Zhao、M. Fiorentino、T. Van Vaerenbergh、Z. Zhang，可分离算子网络（2024年）。doi:10.48550/ARXIV.2407.11253。 96 S. Rezaei、R. N. Asl、S. Faroughi、M. Asgharzadeh、A. Harandi、R. N. Koopas、G. Laschet、S. Reese、M. Apel，将微结构的弹性性质映射到其力学变形的有限算子学习技术，《International Journal for Numerical Methods in Engineering》126（1）（2025年）e7637。doi:10.1002/nme.7637。 97 D. Johnson、T. Maxfield、Y. Jin、R. Fedkiw，基于软件的自动微分存在缺陷，《计算物理学杂志》542（2025年）114319。doi:10.1016/j.jcp. 2025.114319。 98 J. Kim、H. A. Tchelepi、R. Juanes，流动与地质力学耦合顺序方法的稳定性、准确度和效率，载于：SPE油藏模拟研讨会，SPE，伍德兰兹，得克萨斯州，2009年，第SPE–119084–MS页。doi:10.2118/119084-MS。 99 M. E. Mobasher、L. Berger-Vergiat、H. Waisman，多孔介质输运与损伤的非局部形式，《计算机方法：应用力学与工程》324（2017年）654–688。doi:10.1016/j.cma.2017.06.016。 100 W. Li、C. Wei，强耦合孔隙力学问题的稳定化低阶有限元，《International Journal for Numerical Methods in Engineering》115（5）（2018年）531–548。doi:10.1002/nme.5815。 101 D. M. Potts、L. Zdravkovi´c，《岩土工程中的有限元分析。2：应用》，Telford，伦敦，2001年。 102 C. Li，适用于大变形动力分析的孔隙弹性有限元形式（2004年）。 103 T. Hu、B. Jin、Z. Zhou，采用奇异性增强物理信息神经网络求解多边形域中的泊松问题，《SIAM Journal on Scientific Computing》46（4）（2024年）C369–C398。doi:10.1137/23M1601195。 104 Z. Mao、X. Meng，采用基于残差/梯度的自适应采样方法的物理信息神经网络求解具有尖锐解的偏微分方程，《Applied Mathematics and Mechanics》44（7）（2023年）1069–1084。doi:10.1007/ s10483-023-2994-7。 105 Z. Gao 、 L. Yan 、 T. Zhou，面向失效信息的 PINNs 自适应采样，《 SIAM Journal on Scientific Computing》45（4）（2023年）A1971–A1994。doi:10.1137/22M1527763。

106 G. Grekas、C. G. Makridakis、T. Pryer，PINN-DG：采用有限元训练的残差神经网络方法，arXiv:2507.03521 [数值分析] (2025年7月)。doi:10.48550/arXiv.2507.03521。