arXiv:2303.03402v1 [cs.CE] 6 Mar 2023

用于模拟非弹性的不同神经网络架构比较研究

预印本

Max Rosenkranz德累斯顿工业大学固体力学研究所，01062 德累斯顿，德国

Karl A. Kalina德累斯顿工业大学固体力学研究所，01062 德累斯顿，德国

Jörg Brummund德累斯顿工业大学固体力学研究所，01062 德累斯顿，德国

Markus Kästner*德累斯顿工业大学固体力学研究所，01062 德累斯顿，德国

2023年3月8日

摘要

用于描述材料路径相关（即非弹性）行为的本构模型，其数学表述是一项具有挑战性的任务，数十年来一直是力学研究的重点。近年来，借助数据驱动技术来简化或自动化这项任务的努力日益增多，这尤其受到神经网络（NN）在计算力学领域复兴的推动。然而，在基于神经网络的方法中，完全不考虑过去几十年本构建模领域的基础性研究成果，似乎并不可取。本文提出了一项比较研究，探讨不同前馈和循环神经网络架构对非弹性行为的建模能力。在本研究中，我们将模型分为三类：黑箱神经网络、以弱形式施加物理约束的神经网络，以及以强形式施加物理约束的神经网络。第一类网络能够从数据中学习本构关系，却完全忽略底层物理规律；后两类则经过专门构造，能够考虑基本物理规律，本文尤其关注热力学第二定律。采用传统的线性和非线性黏弹性模型以及弹塑性模型生成训练数据，并在之后将其作为参考。使用包含应力、应变以及部分模型的内变量信息的随机游走时间序列进行训练后，我们将基于神经网络的模型与参考解进行比较，并考察插值和外推情况。除了应力预测的质量外，我们还分析相关的自由能和耗散率，以评估这些模型。总体而言，本研究清晰地呈现了不同神经网络架构在非弹性建模方面的优缺点，并为如何训练和应用这些模型提供了指导。

关键词 神经网络 · 循环神经网络 · 强制遵循物理规律 · 本构建模 · 热力学一致性 · 黏弹性 · 塑性

1 引言

数十年来，通过本构模型准确描述材料在机械载荷作用下的行为一直是力学研究的重点。然而，本构模型的构建与参数化仍然是一项艰巨的任务，尤其是对于表现出路径相关（即非弹性）行为的材料。作为传统模型的替代方案，基于数据的 或数据驱动技术前景广阔，有望改进或取代传统模型。近年来，这些技术在计算力学领域日益流行

*通讯 作者，电子邮箱：markus.kaestner@tu-dresden.de。

，其中神经网络（NNs）的应用可能是最常见的技术。下文将简要概述神经网络在本构建模中的应用。

在 20 世纪 90 年代初，Ghaboussi 等人[3] 最早提出了将神经网络用于本构建模的概念。然而，在这一早期阶段，通常采用纯黑箱技术，即不考虑任何物理原理的网络，因此只能准确复现训练数据（此处由应力-应变对构成），但在外推时表现不佳。为解决这一问题，一种在基于神经网络的本构建模以及一般科学机器学习（ML）中相对较新的方法，是以强形式或弱形式融入关键的底层物理。这些方法称为物理信息 4,5、力学信息[6,7],物理增强 8,9、物理约束 10或基于热力学 11的方法，能够提升外推能力，并允许使用稀疏训练数据。

最容易建模的材料行为是弹性，因为在这种情况下，合适的模型只需预测特定变形状态下的应力。在机器学习领域，相关研究[12,13] 旨在使用以三个变形类型不变量为输入的前馈神经网络 （FNN）近似弹性势。因此，若干本构要求可通过构造得到满足，例如热力学一致性、客观性或材料对称性。然而，这些模型的训练直接需要弹性势。与此同时，以不变量为输入、以超弹性势为输出的前馈神经网络是一种十分成熟的方法[8–10,14–18]。此外，采用了改进的训练方法，使网络能够直接根据应力和应变元组进行标定；也就是说，将能量对变形的导数纳入损失函数，这种方法也称为索博列夫训练[19,20]。另一种方法是使用预先训练来预测应力系数的网络构建伪势，从而事后确保基于神经网络的弹性模型满足热力学一致性[21]。

与弹性相比，基于神经网络的方法对具有路径依赖性的本构行为（即非弹性行为）进行建模更加复杂。一些早期方案[22–26] 已经能够实现相当准确的预测，例如将前几个时间步的应力和应变状态加入前馈神经网络的输入层[23]。这样，网络便能间接学习某种演化方程。该模型[22] 利用材料的内变量，可靠地再现黏塑性材料的应力—应变曲线。另一种方法是不借助内变量，而是使用所谓的循环神经网络 （RNN）来表征载荷历史相关行为。研究表明，循环神经网络具有普适性和准确性，尤其是采用更复杂的循环单元时，例如 Hochreiter 和 Schmidhuber 提出的方法[27]。循环神经网络，尤其是长短期记忆 （LSTM）单元，已被广泛用于非弹性建模，例如相关研究[28–31]，其预测质量也十分可观。最近，脉冲长短期记忆网络在各向同性硬化塑性建模中得到应用[32]；与传统神经网络相比，这类网络可大幅降低内存占用和能耗。此外，一种称为线性化最小状态单元的新型循环神经网络[33] ，能够避免其响应依赖于路径采样，因此在弹塑性建模中具有优势。该方法已用于二维和三维数据集，其结果与通过晶体塑性模拟确定的铝合金真实力学行为相符[34]。另一种有前景的各向异性弹塑性表征方法，将李代数与循环神经网络相结合[35]。最后，尽管完全采用单调数据进行训练，一种结合数据驱动编码器和基于物理的解码器的混合模型[36] 仍能够准确预测弹塑性卸载/再加载路径。
尽管基于机器学习的本构建模已取得巨大进展，但迄今提及的用于描述非弹性的神经网络方法有一个共同点：它们不了解热力学第二定律。然而，按照 Masi 等人[11], 的研究，纳入此类基本物理原理具有显著优势：训练目标更加明确，因而训练速度更快，所需数据量也更少，同时外推能力也大幅提升。文献[37–40]介绍了一种名为深度材料网络 的数据驱动框架。该框架通过一个由多个相互连接的机理构建模块组成的网络来再现代表性体积单元的响应；这些模块具有解析均匀化解，因此能够描述复杂的有效响应，同时不损失关键物理特性。文献[41–45] 提出以神经网络替代经典模型中的部分内容。比如，使用前馈神经网络描述屈服函数或演化方程，而非采用特定模型。文献[46]展示了将神经网络与所谓的微球方法相结合的方案。文献[11,47]提出了一种更灵活的率无关非弹性建模方法，其网络架构经过调整，由两个前馈神经网络构成。第一个网络用于学习内变量的演化，第二个网络用于近似自由能；训练过程需要内部状态变量。为确保热力学一致性，即耗散率 $\mathcal{D}$ 始终大于或等于零，将这一由自由能推导出的项加入损失函数。文献[48]展示了一种类似的、专门用于非弹性建模的方法。与前一种模型不同，该方法从循环神经网络的隐藏状态中自动推断出能够捕捉路径依赖性的内部状态变量。因此，这种方法的优势在于训练时只需要应力和应变。尽管如此，前述两种模型[11,48] 的不足在于，对于任意载荷工况，它们并非通过设计就能满足要求 $\mathcal{D} \geq 0$ ，而只是通过向损失函数中添加惩罚项来强制满足。在文献[7,49], 中，网络架构经过设计，可确保热力学一致性。这是通过将广义标准材料[50] 的概念与输入凸神经网络 （ICNN）[51]相结合实现的。上述文献展示了该方法在黏弹性建模中的应用。采用类似方法，通过将亥姆霍兹自由能函数和耗散势替换为先验满足热力学第二定律的数据驱动函数，利用神经常微分方程（NODE）对有限黏弹性进行

通过将亥姆霍兹自由能函数和耗散势替换为数据驱动函数来建模，这些函数先验地满足热力学第二定律，并采用神经常微分方程（NODEs）[52]。

综上所述，目前有多种基于神经网络的方法可用于非弹性建模，其所融入的物理原理程度各不相同。大多数方法仅用于描述某一特定材料类别，即弹塑性或黏弹性。因此，本研究旨在将不同方法纳入统一框架，并将其应用于一维情况下的弹塑性和黏弹性数据，以进行比较。为此，特别关注热力学第二定律的满足情况。鉴于此，将模型分为三类：黑箱神经网络、以弱形式施加物理约束的神经网络，以及以强形式施加物理约束的神经网络。第一类网络从数据中学习本构关系，而完全忽略底层物理；后两类网络的构造方式则使其能够考虑基本物理原理。不过，以弱形式施加物理约束的神经网络不一定能在任意载荷工况下满足热力学第二定律，这是因为 $D \geq 0$ 仅通过惩罚项纳入损失函数。相比之下，以强形式施加物理约束的神经网络，其网络架构经过专门设计，能够在所有情况下满足这一条件，即通过构造来保证。本文采用传统的线性和非线性黏弹性模型以及弹塑性模型生成训练数据，并在后续用作参考。使用包含应力、应变信息，以及部分模型的内变量信息的随机游走时间序列进行训练后，将基于神经网络的模型与参考解进行比较。除应力预测的质量外，还分析预测的自由能和耗散率，以评估模型的插值和外推能力。除所提供的比较外，本文还对部分基于神经网络的模型进行了多处扩展和/或修改，尤其是属于以强形式施加物理约束的神经网络的方法。

本文的组织结构如下：第 2 节和第 3 节分别介绍连续介质固体力学中本构建模的基础知识以及人工神经网络。随后，第 4 节介绍所研究的基于神经网络的本构模型。第 5 节说明训练数据库的生成过程。接着，第 6 节研究各种基于神经网络的模型的预测质量。在讨论结果之后，本文在第 7 节以总结性意见以及对未来必要工作的展望作结。

2 经典本构模型

描述材料的行为需要本构方程。本节将介绍用于表述不同类型材料行为的框架。根据 Haupt[53], 对材料行为的分类，材料可分为四类：(i) 弹性、(ii) 弹塑性、(iii) 黏弹性和 (iv)黏弹塑性。第 (ii) 至 (iv) 类称为非弹性 , 并表现出耗散行为。为了确保此类过程的不可逆性，在制定材料定律时必须考虑热力学一致性 。

2.1 一般框架

2.1.1 耗散不等式

利用熵平衡、能量平衡和热力学第二定律，可以推导出一个称为克劳修斯-迪昂不等式的表达式。假设过程为等温一维过程，该不等式可写为

$$\sigma\dot{\varepsilon} - \dot{\psi} \geq 0 \quad , \tag{1}$$

其中， σ 表示应力， ε 表示应变， ψ 表示亥姆霍兹自由能密度，为简洁起见，下文简称自由能。从式（1）出发，根据ψ的具体选择，可以推导出不同的、严格满足热力学第二定律的本构模型。一般而言，自由能 $\psi := \psi\left(\varepsilon, \xi^1, \xi^2, \ldots, \xi^N\right)$ 是 ε 和内变量 ξ^α， $\alpha \in \{1, \ldots, N\}$的函数。这组内变量用于描述材料点依赖于载荷历史的内部状态，不一定代表可测量的物理量。为简化记号，下文在适当之处将内变量汇总为广义向量 $\boldsymbol{\xi} \in \mathbb{R}^N$。根据等置原理[53],，假设应力 $\sigma := \sigma(\varepsilon, \boldsymbol{\xi})$ 是同一组变量的函数。对式（1）求值可得

$$\mathcal{D} := \left(\sigma - \frac{\partial\psi}{\partial\varepsilon}\right)\dot{\varepsilon} - \underbrace{\frac{\partial\psi}{\partial\boldsymbol{\xi}}}_{=:-\boldsymbol{\tau}} \cdot\dot{\boldsymbol{\xi}} \geq 0 \quad , \tag{2}$$

其中， $\boldsymbol{a} \cdot \boldsymbol{b}$是两个向量 $\boldsymbol{a}$、 $\boldsymbol{b} \in \mathbb{R}^N$的标量积。量 $\mathcal{D}$和 $\boldsymbol{\tau} \in \mathbb{R}^N$分别表示耗散率和热力学共轭力向量，后者也称为内力[50],，对应于内变量 $\boldsymbol{\xi}$。

为满足不等式 (2)，必要且充分条件

$$\mathcal{D} \geq 0 \quad \Longleftrightarrow \quad \sigma = \frac{\partial \psi}{\partial \varepsilon} \quad \wedge \quad \boldsymbol{\tau} \cdot \dot{\boldsymbol{\xi}} = \mathcal{D} \geq 0 \tag{3}$$

出现。因此，演化 方程的内变量仍有待定义，以确保 $\mathcal{D} \geq 0$ 始终成立 s.

2.1.2 广义标准材料

构建必要演化方程的一种常见方式，是采用广义标准材- 料[50,54,55] 的概念，下面将对此进行简要说明。在这一概念中，除了 ψ之外，还引入了一个耗散势 $\phi := \phi(\dot{\boldsymbol{\xi}}, \boldsymbol{\xi}, \varepsilon)$，其定义满足以下条件：(i) 关于其第一个自变量 $\dot{\boldsymbol{\xi}}$是凸函数，并且关于 $\dot{\boldsymbol{\xi}}$进行归一化，即满足条件(ii) $\phi(\mathbf{0}, \boldsymbol{\xi}, \varepsilon) = 0$，以及(iii) $\phi(\mathbf{0}, \boldsymbol{\xi}, \varepsilon) \geq 0$ 在 $\dot{\boldsymbol{\xi}}$中取最小值。对于率无关材料，即弹塑性材料，耗散势可能是不光滑的。现在可以根据 ϕ，按照 $\boldsymbol{\tau} \in \partial_{\dot{\boldsymbol{\xi}}} \phi$确定内力，其中算子 $\partial_{\dot{\boldsymbol{\xi}}}(\cdot)$ 表示非光滑凸函数的次微分。另一方面，如果 ϕ 是光滑的，则关系变为 $\boldsymbol{\tau} = \partial_{\dot{\boldsymbol{\xi}}} \phi$，此时引入的算子表示标准偏导数。因此，利用公式(2)，可得到描述内变量演化的比奥方程，其形式为

$$\mathbf{0} \in \partial_{\boldsymbol{\xi}} \psi + \partial_{\dot{\boldsymbol{\xi}}} \phi \quad \text{or} \quad \mathbf{0} = \partial_{\boldsymbol{\xi}} \psi + \partial_{\dot{\boldsymbol{\xi}}} \phi \quad \text{with} \quad \boldsymbol{\xi}(t=0) = \boldsymbol{\xi}_0 \tag{4}$$

分别适用于率无关和率相关本构行为。请注意，式（3）中的不等式，即$\mathcal{D} \geq 0$，由于对 ϕ提出了凸性和归一化要求，依据式（4）可自动满足。通过勒让德-芬切尔变换可得到另一种表述，即对偶耗散势 ϕ^*

$$\phi^*(\boldsymbol{\tau}, \boldsymbol{\xi}, \varepsilon) := \sup_{\dot{\boldsymbol{\xi}}} \left[\boldsymbol{\tau} \cdot \dot{\boldsymbol{\xi}} - \phi(\dot{\boldsymbol{\xi}}, \boldsymbol{\xi}, \varepsilon) \right] . \tag{5}$$

使用对偶耗散势可得到

$$\dot{\boldsymbol{\xi}} \in \partial_{\boldsymbol{\tau}} \phi^* \quad \text{or} \quad \dot{\boldsymbol{\xi}} = \partial_{\boldsymbol{\tau}} \phi^* \quad \text{with} \quad \boldsymbol{\xi}(t=0) = \boldsymbol{\xi}_0 \quad , \tag{6}$$

分别取代式（4）。

注1 需要指出的是，式（4）和式（6）在上述形式下仅在内变量的变化率或内力彼此独立时成立。如果各个量之间存在约束，则必须在计算式（4）和式（6）时明确考虑这些约束。否则，例如，耗散势可能表示为内变量变化率的一个约化集合的函数，而相对于被省略变化率的偏导数则为零。对于表1所列的弹塑性模型，这种情况即是如此。

在率无关本构行为这一特殊情况下，现在通过最大耗散概念[50,55]求得耗散函数 ϕ。因此，它由以下约束最大化问题定义：

$$\phi(\dot{\boldsymbol{\xi}}, \boldsymbol{\xi}, \varepsilon) := \sup_{\boldsymbol{\tau} \in \mathcal{E}} \left(\boldsymbol{\tau} \cdot \dot{\boldsymbol{\xi}} \right) \quad \text{with} \quad \mathcal{E} := \left\{ \boldsymbol{\tau} \in \mathbb{R}^N \mid f(\boldsymbol{\tau}, \boldsymbol{\xi}) \leq 0 \right\} . \tag{7}$$

在上式中，$\mathcal{E}$表示内力可容许域，其屈服函数为 $f(\boldsymbol{\tau}, \boldsymbol{\xi})$。假设该函数关于 $\boldsymbol{\tau}$是凸的、经过归一化且为一次齐次函数。式（7）的解给出了内变量的演化方程，以及 Karush-Kuhn-Tucker 条件：

$$\dot{\boldsymbol{\xi}} = \lambda \partial_{\boldsymbol{\tau}} f \ \wedge \ \lambda \geq 0 \ \wedge \ f \leq 0 \ \wedge \ \lambda f = 0 . \tag{8}$$

其中，标量 $\lambda \in \mathbb{R}_+$ 表示塑性乘子。

2.2 特定本构模型

所概述的框架 （续）见第 2.1 节，可用于描述黏弹性或弹塑性等行为，其中 h
相应的模型 简要总结如下。

2.2.1 黏弹性

固体的黏弹性行为表现为应变率相关的应力响应，并具有弹性平衡曲线，见图1c。这些特性可由广义 *Maxwell* 模型[53] 描述，如图1a所示。其中，E 和E_α 是流变模型中弹簧的杨氏模量，η_α 是相应的黏度，$\varepsilon_\alpha^{\mathrm{el}}$ 和 $\varepsilon_\alpha^{\mathrm{vi}}$分别是弹性应变和黏性应变，$\sigma_\alpha^{\mathrm{ov}}$是非平衡应力，亦称过应力。黏度 η_α 可能是过应力 $\sigma_\alpha^{\mathrm{ov}}$的函数，见图1b。下文将黏度取决于过应力的模型称为非线性模型，否则称为线性模型。此类广义 Maxwell 模型的行为由表1中的控制方程描述。

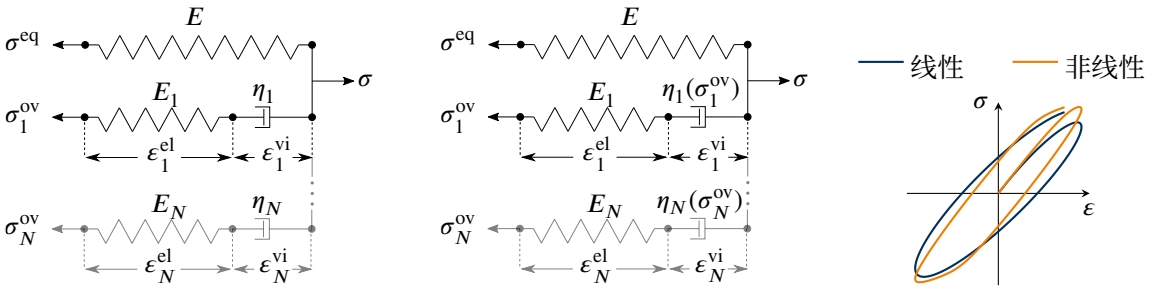


图1：黏弹性本构模型：**(a)** 线性广义麦克斯韦模型和**(b)** 非线性广义麦克斯韦模型的流变模型，以及**(c)**(a)和(b)的典型滞回曲线。

表1：所考虑的黏弹性和弹塑性模型的控制方程。

	黏弹性	**弹塑性**
自由能:	$\psi := \frac{1}{2}E\varepsilon^2 + \sum_{\alpha=1}^{N}\frac{1}{2}E_\alpha\left(\varepsilon-\varepsilon_\alpha^{\mathrm{vi}}\right)^2$	$\psi := \frac{1}{2}E\left(\varepsilon-\varepsilon^{\mathrm{pl}}\right)^2 + \frac{1}{2}H\alpha^2 + \frac{1}{2}\hat{H}\hat{\alpha}^2$
耗散势:	$\phi = \sum_{\alpha=1}^{N}\frac{\eta_\alpha}{2}\left(\dot{\varepsilon}_\alpha^{\mathrm{vi}}\right)^2$	$\phi = \sigma_{\mathrm{y}_0}\lvert\dot{\varepsilon}^{\mathrm{pl}}\rvert$ 其中 $\dot{\alpha} = \dot{\varepsilon}^{\mathrm{pl}}$ 和 $\hat{\dot{\alpha}} = \lvert\dot{\varepsilon}^{\mathrm{pl}}\rvert$
对偶耗散势:	$\phi^* = \sum_{\alpha=1}^{N}\frac{1}{2\eta_\alpha}\left(\sigma_\alpha^{\mathrm{ov}}\right)^2$	$\phi^* = \begin{cases} 0, & \text{如果 } (\sigma^{\mathrm{pl}}, p, \hat{p}) \in \mathcal{E} \\ \infty, & \text{else} \end{cases}$
应力:	$\sigma = E\varepsilon + \sum_{\alpha=1}^{N}E_\alpha\left(\varepsilon-\varepsilon_\alpha^{\mathrm{vi}}\right)$	$\sigma = \partial_\varepsilon\psi = E\left(\varepsilon-\varepsilon^{\mathrm{pl}}\right)$
内力:	$\sigma_\alpha^{\mathrm{ov}} = -\partial_{\varepsilon_\alpha^{\mathrm{vi}}}\psi = E_\alpha\left(\varepsilon-\varepsilon_\alpha^{\mathrm{vi}}\right)$	$\sigma^{\mathrm{pl}} = -\partial_{\varepsilon^{\mathrm{pl}}}\psi = \sigma\,,\quad p = -\partial_\alpha\psi = -H\alpha\,,$ $p = -\partial_{\hat{\alpha}}\psi = -\hat{H}\hat{\alpha}$
演化方程:	$\dot{\varepsilon}_\alpha^{\mathrm{vi}} := \frac{1}{\eta_\alpha}\sigma_\alpha^{\mathrm{ov}}$	$\dot{\varepsilon}^{\mathrm{pl}} = \dot{\alpha} = \lambda\,\mathrm{sign}\,(\sigma+p) \quad,\quad \hat{\dot{\alpha}} = \lambda$
耗散率:	$\mathcal{D} = \sum_{\alpha=1}^{N}\sigma_\alpha^{\mathrm{ov}}\dot{\varepsilon}_\alpha^{\mathrm{vi}} \geq 0$	$\mathcal{D} = \sigma\dot{\varepsilon}^{\mathrm{pl}} + p\dot{\alpha} + \hat{p}\hat{\dot{\alpha}} \geq 0$
黏度:	$\eta_\alpha\left(\sigma_\alpha^{\mathrm{ov}}\right) := \hat{\eta}_\alpha\exp\left(a_\alpha\lvert\sigma_\alpha^{\mathrm{ov}}\rvert^{b_\alpha}\right)$	—
屈服函数:	—	$f := \lvert\sigma+p\rvert - \sigma_{\mathrm{y}}(\hat{p}) \quad,\quad \sigma_{\mathrm{y}}(\hat{p}) = \sigma_{\mathrm{y}_0} - \hat{p}$
Karush-Kuhn-Tucker 条件:	—	$\lambda \geq 0 \;\wedge\; f \leq 0 \;\wedge\; \lambda f = 0$

2.2.2 弹塑性

与黏弹性不同，弹塑性材料行为的特点是应变率无关，但存在平衡滞回[53]。在初始区域 $\sigma \in \left(-\sigma_{\mathrm{y}_0}, \sigma_{\mathrm{y}_0}\right)$内，材料表现为纯弹性。若超出该初始区域的边界，即 $|\sigma| > \sigma_{\mathrm{y}_0}$，则会发生弹塑性变形，弹性区域既可以沿相应方向移动（运动硬化），也可以扩大（各向同性硬化）[56]。为了模拟这些特性，需要定义一个所谓的屈服函数 $f := f(\boldsymbol{\tau}, \boldsymbol{\xi}) \leq 0$。它确定弹性区域的边界：当 $f < 0$ 时，变形为纯弹性；当 $f = 0$时，则发生弹塑性变形。 $f > 0$ 这种情况不允许出现。演化方程可由最大耗散原理导出，见第2.1节。请注意，内变量的变化率彼此并非独立，因此在计算式(4)时必须予以考虑，详见注释1。由此可推导出具有随动硬化与各向同性硬化混合特性的弹塑性行为控制方程，参见表1。相应的流变模型如图2a所示。其中， E 为杨氏模量， H 和 $\hat{H}$ 分别为运动硬化模量和各向同性硬化模量，σ_{y_0} 为初始屈服应力。此外， p 表示背应力， $\varepsilon^{\mathrm{el}}$ 和 $\varepsilon^{\mathrm{pl}}$ 分别表示弹性应变和塑性应变。

将硬化模块 H 或 $\hat{H}$ 设为零，会得到理想塑性（$H = \hat{H} = 0$）、仅包含运动硬化（$H = 0$）或仅包含各向同性硬化（$\hat{H} = 0$）这些特殊情况，见图2b–2e。此时，设为零的硬化模块所对应的内变量仍可能取非零值，但不再对 σ、 ψ或 $\mathcal{D}$ 产生任何影响，因此可以忽略。例如，在理想塑性 $H = \hat{H} = 0$的情况下，内变量集可简化为塑性应变 $\varepsilon^{\mathrm{pl}}$。

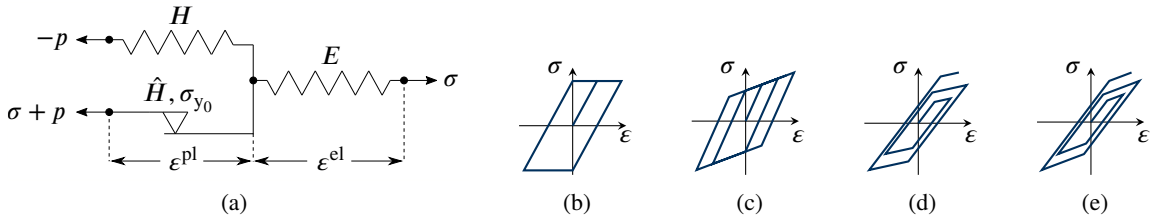


图 2：弹塑性本构模型：（**a**）具有各向同性硬化和运动硬化的流变模型，以及典型的ˆ ˆ滞回曲线：（**b**）理想塑性（$H = H = 0$）、（**c**）运动硬化（$H = 0$）、（**d**）各向同性硬化（$H = 0$）和（**e**）运动硬化与各向同性硬化。

3 人工神经网络基础

神经网络可分为 分为不同类别，分别适用于不同类型的任务[57,58]。对于本构建模，
使用前馈神经网络和循环神经网络
尤其合适。

3.1 前馈神经网络

前馈神经网络由排列在多个层中的人工神经元组成，如图 3a 所示。输入变量（特征） i_n的信息汇总于输入 $\boldsymbol{I} := (i_1, \dots, i_{N_1})$中，并传递给后续第 2 层的每个神经元，用于计算该层的输出。第 2 层的全部输出又作为下一层每个神经元的输入，以此类推。从输入层到最终输出$\boldsymbol{O} := (o_1, \dots, o_{N_L})$（第 L 层）的信息正向流动，且不反馈中间结果，这种方式称为前馈。图 3b 展示了第 l 层的第 n 个神经元。该神经元的输出 $\mathcal{N}_n^l$ 由下式确定：

$$\mathcal{N}_n^l = \mathcal{A}_n^l\left(x_n^l\right) = \mathcal{A}_n^l\left(\sum_{m=0}^{N_{l-1}} \mathcal{N}_m^{l-1} w_{nm}^l\right) \quad . \tag{9}$$

其中， $\mathcal{A}_n^l$ 表示该神经元的激活函数， N_{l-1} 表示第 $l-1$ 层中的神经元数量， w_{nm}^l 表示输出 $\mathcal{N}_m^{l-1}$在激活函数 $\mathcal{A}_n^l$ 的自变量中的权重。虚拟神经元 $\mathcal{N}_0^{l-1} \equiv 1$ 的输出权重 w_{n0}^l 称为该神经元的偏置。常见的激活函数包括双曲正切 $\tanh(x)$、整流线性单元 $\mathrm{ReLU}(x) := \max(0, x)$ 或 Softplus 激活函数 $\mathrm{SP}(x) := \ln(1 + \exp(x))$[11,57,58]。

特别地，输出层 L 的输出 $\mathcal{N}_n^L$也可以用式（9）表示。递归地，可以逐层替换 $\mathcal{N}_n^l$，直到到达第 1 层及其输入 $\mathcal{N}_n^1 = i_n$。这表明，此类网络可以任意嵌套；但对于固定的 w_{nm}^l，它是输入 $\boldsymbol{I}$的良定义函数。从这个意义上说，网络作为一个函数

$$\mathcal{F} : \mathbb{R}^{N_1} \to \mathbb{R}^{N_L} \quad , \quad \boldsymbol{I} \mapsto \mathcal{F}(\boldsymbol{I}) = \boldsymbol{O} \tag{10}$$

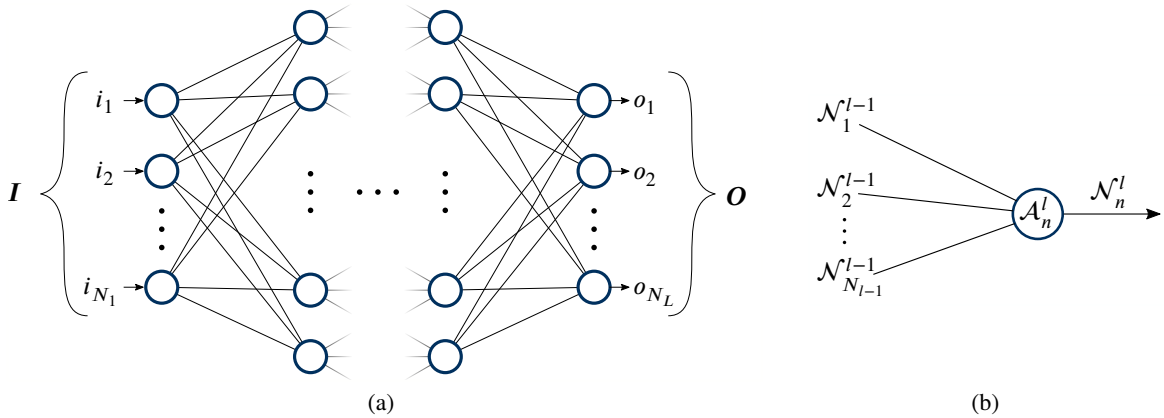


图3：前馈神经网络的结构：**(a)** 具有任意数量层和神经元的完整网络的一般表示，以及**(b)** 第 l层第 n个神经元的功能。

将输入 $\boldsymbol{I}$ 映射为输出 $\boldsymbol{O}$。为了调整权重 w^l_{nm}，使神经网络的预测与期望值相匹配，需要一个由 $N^{\mathrm{ds}} \in \mathbb{N}$ 个数据元组 $\mathcal{T}_\alpha$ 组成的训练数据集 $\mathcal{S} := \{\mathcal{T}_1, \dots, \mathcal{T}_{N^{\mathrm{ds}}}\}$。在每个数据元组 $\mathcal{T}_\alpha := (\boldsymbol{I}_\alpha, \boldsymbol{O}_\alpha)$中，都会为一个输入 $\boldsymbol{I}_\alpha$ 指定其期望输出 $\boldsymbol{O}_\alpha$。期望输出通常包含最后一层神经元的目标值，但也可能包含其他信息，例如相对于某个特定输入的目标导数。神经网络预测相对于这组数据元组的误差由损失函数 $\mathcal{L}(\boldsymbol{w}, \mathcal{S})$汇总，其中 $\boldsymbol{w}$ 表示神经网络所有权重 w^l_{nm} 构成的向量。对于此处展示的应用，损失函数主要由平均绝对误差项组成；对于前馈神经网络，其形式为

$$\mathrm{MAE}(q) := \frac{1}{N^{\mathrm{ds}}} \sum_{i=1}^{N^{\mathrm{ds}}} \left| q(\boldsymbol{I}_i, \boldsymbol{w}) - \bar{q}_i \right| \quad . \tag{11}$$

其中， q 是待评估其预测结果的量。相对于权重 $\boldsymbol{w}$对该误差函数进行最小化

$$\boldsymbol{w} = \arg\min_{\boldsymbol{\kappa}} \mathcal{L}(\boldsymbol{\kappa}, \mathcal{S}) \tag{12}$$

被称为神经网络的训练过程。可以使用多种方法来解决这一最小化问题，例如随机梯度下降 （SGD）或 *Adam*，这两者是最常用的优化器中的代表。在本研究中，采用序列最小二乘规划 （SLSQP）进行优化。

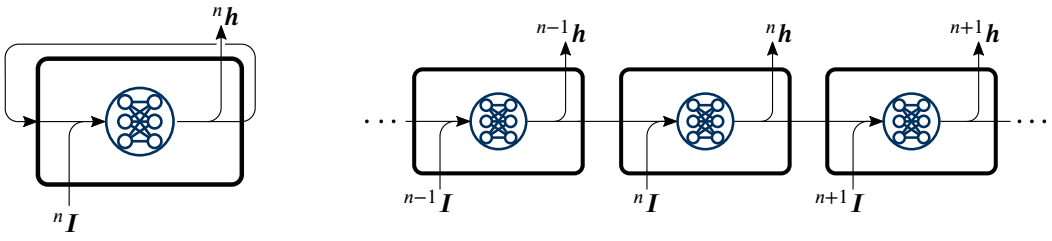
ICNN 是一类特殊的前馈神经网络，最初由 Amos 等人提出。[51], 即，这类网络具有关于其输入参数为凸的性质。通过使用凸且非递减的激活函数以及非负权重，可以实现这一点。

3.2 循环神经网络

与前馈神经网络不同，循环神经网络允许将中间结果反馈回来，见图 4a。这一特性使网络不仅能够考虑一个时间步，还能考虑多个时间步的输出，从而学会在输入变量的（伪）时间上下文中对其进行解释。例如，这对于文本翻译、语音识别或物理量演化的预测尤其有用。因此，在循环神经网络中，会对连续数据的完整序列进行评估；在本研究所展示的应用中，这些序列是由多个时间步组成的集合。因此，循环神经网络的运行方式与前馈神经网络有很大不同。不过，其内部结构同样由一个或多个前馈神经网络组成。

这些前馈神经网络接收输入向量 $\boldsymbol{I}$ 以及所谓的隐藏状态 ${}^n\boldsymbol{h}$。在循环单元之外，新的隐藏状态随后会被进一步处理，以获得输出 ${}^n\boldsymbol{O}$。图 4b 展示了一个简单的循环神经网络单元在时间轴上展开后的这一过程。

图示单元仅由一个前馈神经网络组成，无法将许多前置时间步的信息纳入当前隐藏状态的输出中。因此，为了使单元能够检测长期依赖关系，研究人员开发了采用多个前馈神经网络的更复杂内部结构。本文介绍的循环神经网络使用所谓的长短期记忆 (LSTM)[27] 单元来解决这一问题。除了隐藏状态之外，这种长短期记忆单元还会在各个时间步之间传递另一组信息，即所谓的单元状态 ${}^n\boldsymbol{c}$。 N^{c} 中条目的数量 ${}^n\boldsymbol{c}$ 可解释为该单元的记忆容量，是影响网络性能的关键参数。与前馈神经网络不同，训练数据集 $\mathcal{S} := \{\mathcal{T}_1, \dots, \mathcal{T}_{N^{\mathrm{seq}}}\}$ 包含 $N^{\mathrm{seq}} \in \mathbb{N}$ 个由输入序列 $\mathcal{T}\mathrm{I}_s$ 和对应预期输出序列 $\bar{\mathcal{T}}\mathrm{O}_s$ 组成的元组 $\mathcal{T}_s := (\mathcal{T}\mathrm{I}_s, \bar{\mathcal{T}}\mathrm{O}_s)$。每个输入序列 $\mathcal{T}\mathrm{I}_s := ({}^1\boldsymbol{I}_s, \dots, {}^{N^{\mathrm{ts}}_s}\boldsymbol{I}_s)$ 都是由 $N^{\mathrm{ts}}_s \in \mathbb{N}$ 个输入向量按顺序排列而成，每个向量代表该序列中的一个 N^{ts}_s时间步。预期输出序列 $\bar{\mathcal{T}}\mathrm{O}_s := ({}^1\bar{\boldsymbol{O}}_s, \dots, {}^{N^{\mathrm{ts}}_s}\bar{\boldsymbol{O}}_s)$



(a) (b)

图 4：标准循环神经网络单元的功能和内部结构的两种等价示意图：输出（**a**）反馈到循环神经网络单元中，（**b**）则从一个时间步传递到下一个时间步。也就是说，（b）可以理解为（a）在时间轴上的展开。

包含相同数量的预期输出向量。

循环神经网络预测的误差通过损失函数 $\mathcal{L}(\boldsymbol{w}, S)$ 衡量，其中 $\boldsymbol{w}$ 现在包含循环单元内部和外部所有前馈神经网络的权重。对于循环神经网络，量 q 的平均绝对误差形式为

$$\mathrm{MAE}(q) := \frac{1}{N^{\mathrm{seq}}} \sum_{s=1}^{N^{\mathrm{seq}}} \left(\frac{1}{N_s^{\mathrm{ts}}} \sum_{n=1}^{N_s^{\mathrm{ts}}} \left| q(\mathcal{T}_s^{\mathrm{I}}, \boldsymbol{w}) - {}^{n}\bar{q}_s \right| \right) \quad . \tag{13}$$

与前馈神经网络类似，训练过程通过最小化损失函数来调整网络中的所有权重。

4 基于神经网络的本构模型

尽管第 2 节介绍的模型能够以符合物理规律的方式再现真实材料行为的基本特征，但用于求得一组控制方程的方法存在一些缺点。一方面，这组方程必须手动构建；对于三维空间中更复杂的材料，这一过程会变得愈发困难。另一方面，该方法只能再现建模时考虑到的现象。本节将介绍一些基于神经网络的方法，旨在自动寻找任意关联，即无需额外的人工建模工作。这些方法分为三类，分别对应于所融入物理学的程度：（i）不包含物理信息的黑箱模型；（ii）以弱形式施加物理约束的神经网络，其训练目标是满足第二定律；（iii）以强形式施加物理约束的基于神经网络的模型，其中第二定律先验地得到满足。每种方法都需要一个训练数据集，其生成过程将在第 5 节中概述。模型使用 Python 和 TensorFlow 库实现，结果将在第 6 节中呈现。

4.1 黑箱模型

黑箱模型的特点是未融入物理知识，即它们接收一组输入值（取决于具体模型），并输出所需的量（应力），而不遵循连续介质力学中已知的关系。本文考察了两种简单架构：一种使用前馈网络，另一种使用循环单元。

4.1.1 应力作为输出的前馈架构（FNN$^{\sigma}$）

该架构是本文介绍的六种架构中最简单的一种，仅由一个前馈神经网络组成，该网络直接输出应力。它接收新的应变 ${}^{n+1}\varepsilon$，以及任意数量的前置时间步 $N^{\mathrm{pt}} \geq 1$ 对应的应变 ${}^{m}\varepsilon$、应力 ${}^{m}\sigma$ 和时间增量 ${}^{m}\Delta t := {}^{m+1}t - {}^{m}t$，并输出新的应力，即

$$\boldsymbol{I} := ({}^{n+1}\varepsilon, {}^{n}\varepsilon, {}^{n}\sigma, {}^{n}\Delta t, \ldots, {}^{n+1-N^{\mathrm{pt}}}\varepsilon, {}^{n+1-N^{\mathrm{pt}}}\sigma, {}^{n+1-N^{\mathrm{pt}}}\Delta t) \mapsto O := {}^{n+1}\sigma \quad . \tag{14}$$

输入中的时间步 ${}^{m}\Delta t$ 仅对率相关行为（即黏弹性）是必需的。损失函数仅包含应力预测，

$$\mathcal{L} := \mathcal{L}^{\sigma} := \mathrm{MAE}(\sigma) \quad . \tag{15}$$

此前已有类似方法的描述，例如用于本构建模的嵌套自适应神经网络[23,24,26]。然而，正如第 6 节的结果所示，该方法只能应用于特定类别的材料行为。基于循环单元的改进架构可用于克服这一限制。

4.1.2 以应力为输出的循环架构（RNN$^{\sigma}$）

第二种黑箱架构由一个循环单元 和其后的前馈神经网络（FNN）组成，如图 6 所示。此类模型已用于相关研究[29]。对于每个序列中的每个时间步，循环神经网络单元以新的应变以及必要时的时间增量作为输入，并输出新的隐藏状态；随后将该隐藏状态输入前馈神经网络，以获得新的应力，

$${}^{n}\boldsymbol{I}^{\mathrm{RNN}} := ({}^{n}\varepsilon, {}^{n-1}\Delta t) \mapsto {}^{n}\boldsymbol{h} =: {}^{n}\boldsymbol{I}^{\mathrm{FNN}} \mapsto {}^{n}O^{\mathrm{FNN}} := {}^{n}\sigma \quad . \tag{16}$$

第一个时间步之前的隐藏状态和单元状态初始值分别设为 ${}^{0}\boldsymbol{h} = \boldsymbol{0}$ 和 ${}^{0}\boldsymbol{c} = \boldsymbol{0}$。损失函数

$$\mathcal{L} := \mathcal{L}^{\sigma} := \mathrm{MAE}(\sigma) \tag{17}$$

计算所有时间步和序列上的应力预测误差的平均值。

这种架构能够对种类繁多的材料进行建模。不过，网络既没有获得任何物理知识，也无法获得应力响应之外的其他物理信息。因此，下面介绍两种融入物理信息的模型。

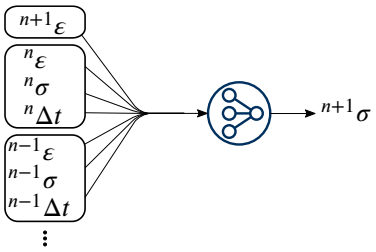


图 5：前馈神经网络$^{\sigma}$ 架构的功能：单个前馈神经网络将多个时间步的信息映射为新的应力。

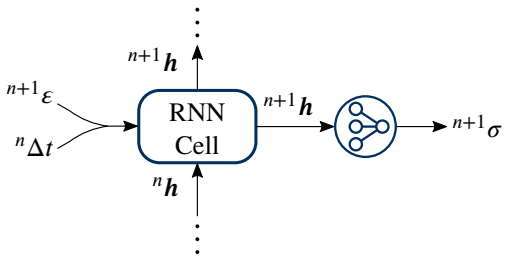


图 6：循环神经网络$^{\sigma}$ 架构的功能：循环神经网络单元根据新的应变和时间增量输出新的隐藏状态，再将其输入前馈神经网络以预测新的应力。

4.2 以弱形式施加物理约束的神经网络

与黑箱模型不同，以弱形式施加物理约束的基于神经网络的模型会考虑已知关系，即式（3）$_2$ 和式（3）$_3$。这样做旨在获得真正符合物理规律的预测，并获取额外信息，尤其是自由能和耗散率。网络经过训练以遵循这些关系，但架构本身无法先验地保证 $\mathcal{D} \geq 0$ 这一点。为将这些关系纳入训练过程，必须对网络关于其输入进行求导。对网络求导会限制激活函数的选择：由于优化时需要损失函数的梯度，因此激活函数需要进行二阶求导。因此，激活函数必须二阶可微，才能进行优化[10,11,47]。因此，以下相应网络采用双曲正切和 Softplus 激活函数。

4.2.1 以前馈架构处理内变量并以自由能为输出（前馈神经网络$^{\xi+\psi}$ ）

下文所述架构基于 Masi 等人[11]，并在其研究中进行了详细探讨。如图 7 所示，该架构使用两个前馈神经网络，分别执行不同任务。第一个子网络 sFNN$^{\xi}$ 接收新的和旧的应变、旧的应力，以及前一时间步的时间增量和内变量，并通过训练来预测新的内变量:

$$\boldsymbol{I}^{\mathrm{sFNN}^{\xi}} := (\,{}^{n+1}\boldsymbol{\varepsilon}, {}^{n}\boldsymbol{\varepsilon}, {}^{n}\boldsymbol{\sigma}, {}^{n}\Delta t, {}^{n}\boldsymbol{\xi}^{1}, \dots, {}^{n}\boldsymbol{\xi}^{N}\,) \mapsto \boldsymbol{O}^{\mathrm{sFNN}^{\xi}} := {}^{n+1}\boldsymbol{\xi} \quad . \tag{18}$$

下面的第二个子网络 sFNN$^{\psi}$ 利用第一个网络的输出以及新的应变来确定自由能，即

$$\boldsymbol{I}^{\mathrm{sFNN}^{\psi}} := (\,{}^{n+1}\boldsymbol{\varepsilon}, {}^{n+1}\boldsymbol{\xi}^{1}, \dots, {}^{n+1}\boldsymbol{\xi}^{N}\,) \mapsto O^{\mathrm{sFNN}^{\psi}} := {}^{n+1}\psi \quad . \tag{19}$$

由于 sFNN$^{\xi}$ 以 ${}^{n}\boldsymbol{\xi}$ 作为输入，因此训练数据集中必须包含内变量。自由能并非必需，但能够提升网络性能。在本研究中，目标是尽可能少地使用信息，因此将自由能视为训练期间不可用的信息。以下所有架构同样适用。

与原始工作[11]不同，训练过程并非同时考虑两个网络并作为一次优化执行，而是分为两个较小的步骤，从而加快优化过程：首先，独立训练 sFNN$^{\xi}$，即将其与 sFNN$^{\psi}$分离，使用

$$\mathcal{L}^{\mathrm{sFNN}^{\xi}} := \mathcal{L}^{\xi} \quad \text{with} \quad \mathcal{L}^{\xi} := \frac{1}{N}\sum_{\alpha=1}^{N} \mathrm{MAE}\,(\xi^{\alpha}) \tag{20}$$

来学习预测内变量。随后，将两个网络结合使用，通过损失函数训练第二个网络

$$\mathcal{L}^{\mathrm{sFNN}^{\psi}} := w^{\sigma}\mathcal{L}^{\sigma} + w^{\psi}\mathcal{L}^{\psi} + w^{\mathcal{D}\geq 0}\mathcal{L}^{\mathcal{D}\geq 0} \quad \text{with} \quad \mathcal{L}^{\sigma} := \mathrm{MAE}\,(\sigma) \quad , \tag{21}$$

$$\mathcal{L}^{\psi} := \mathrm{MAE}\,(\psi) \quad \text{and} \quad \mathcal{L}^{\mathcal{D}\geq 0} := \sum_{k=1}^{N^{\mathrm{ds}}} \mathrm{ReLU}\left(-{}^{n+1}\mathcal{D}_{k}\right) \quad \text{where} \quad {}^{n+1}\mathcal{D} = -\sum_{\alpha=1}^{N} \frac{\partial\, {}^{n+1}\psi}{\partial\, {}^{n+1}\xi^{\alpha}} \frac{{}^{n+1}\xi^{\alpha} - {}^{n}\xi^{\alpha}}{{}^{n}\Delta t} \quad . \tag{22}$$

不可训练参数 w^{σ}、w^{ψ} 和 $w^{\mathcal{D}\geq 0}$ 调节各项对损失函数值的影响。在此优化过程中，损失函数仅针对 sFNN$^{\psi}$的权重进行求导，因此只有这些权重

会进行调整，而 sFNN$^{\xi}$ 的权重则保留其在第一步训练中获得的值。然而，为更复杂的材料行为提供训练数据所需的内变量并非易事，Masi 和 Stefanou[47]对此进行了探讨。为了避免提供内变量，一种类似的架构似乎是合适的：用循环单元替换 sFNN$^{\xi}$。

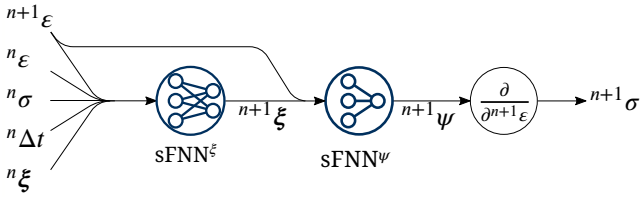


图 7：前馈神经网络$^{\xi+\psi}$ 架构的功能：基于输入值，第一个子网络 sFNN$^{\xi}$ 预测新的一组内部变量；这些变量与新应变一起传递给第二个子网络 sFNN$^{\psi}$，以预测新的自由能。对新应变求导即可得到新应力。

4.2.2 以内部变量和自由能为输出的循环架构（RNN$^{\xi+\psi}$ ）

由于循环神经网络能够存储很久以前的时间步中的信息，因此利用这一特性来模拟内变量似乎是合理的。图 8 展示了这一思想的一种可能实现方式，与 He 和 Chen[48]的方法相似。在每个时间步，循环神经网络单元以新应变以及时间增量作为输入（对于黏性行为而言），并生成一个新的向量值输出 $^{n}\boldsymbol{h}$，该输出携带历史信息，并以 $^{0}\boldsymbol{h} = {}^{0}\boldsymbol{c} = \boldsymbol{0}$ 初始化。随后，这些值被输入前馈神经网络，并被缩减为数量可自由选择的内变量 N^{ξ}：

$$^{n}\boldsymbol{I}^{\mathrm{RNN}} := (\,^{n}\boldsymbol{\varepsilon}, {}^{n-1}\Delta t\,) \mapsto {}^{n}\boldsymbol{h} =: {}^{n}\boldsymbol{I}^{\mathrm{FF}^{\xi}} \mapsto {}^{n}\boldsymbol{O}^{\mathrm{FF}^{\xi}} := {}^{n}\boldsymbol{\xi} \in \mathbb{R}^{N^{\xi}} \quad . \tag{23}$$

Wit在另一个前馈神经网络中，这些值与新的应变一起被映射为自由能，即 .,

$$\boldsymbol{I}^{\mathrm{FF}^{\psi}} := (\,^{n}\boldsymbol{\varepsilon}, {}^{n}\xi^{1}, \ldots, {}^{n}\xi^{N^{\xi}}\,) \mapsto O^{\mathrm{FF}^{\psi}} := {}^{n}\psi \quad . \tag{24}$$

损失函数

$$\mathcal{L}^{\mathrm{FF}^{\psi}} := w^{\sigma}\mathcal{L}^{\sigma} + w^{\xi}\mathcal{L}^{\xi} + w^{\psi}\mathcal{L}^{\psi} + w^{\mathcal{D}\geq 0}\mathcal{L}^{\mathcal{D}\geq 0} \quad \text{with} \tag{25}$$

$$\mathcal{L}^{\sigma} := \mathrm{MAE}\,(\sigma) \quad , \quad \mathcal{L}^{\xi} := \frac{1}{N_{\mathrm{av}}} \sum_{\alpha=1}^{N_{\mathrm{av}}} \mathrm{MAE}\left(\xi^{\alpha}_{\mathrm{av}}\right) \quad , \quad \mathcal{L}^{\psi} := \mathrm{MAE}\,(\psi) \quad \text{and} \tag{26}$$

$$\mathcal{L}^{\mathcal{D}\geq 0} := \frac{1}{N^{\mathrm{seq}}} \sum_{s=1}^{N^{\mathrm{seq}}} \left(\frac{1}{N_s^{\mathrm{ts}}} \sum_{n=1}^{N_s^{\mathrm{ts}}} \mathrm{ReLU}\left(-{}^{n}\mathcal{D}_s\right) \right) \quad \text{where} \quad {}^{n}\mathcal{D} = -\sum_{\alpha=1}^{N} \frac{\partial\,{}^{n}\psi}{\partial\,{}^{n}\xi^{\alpha}} \frac{{}^{n}\xi^{\alpha} - {}^{n-1}\xi^{\alpha}}{{}^{n}\Delta t} \quad , \tag{27}$$

由四项组成，分别用于控制应力预测、内变量和自由能的预测，以及对热力学一致性的遵循，即 $\mathcal{D} \geq 0$。在内变量的误差项 $\mathcal{L}^{\xi}$ 中，只考虑那些

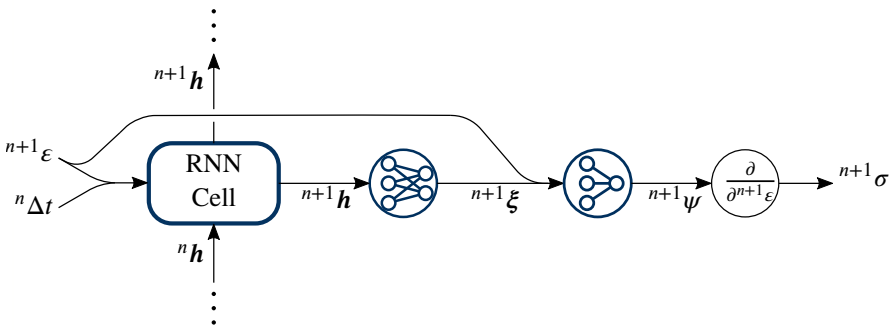


图 8：循环神经网络$^{\xi+\psi}$ 架构的功能：对于序列中的每个时间步，新的应变和时间增量都会输入循环神经网络单元，其输出随后传递给前馈网络 sFNN$^{\xi}$，以获得新的一组内部变量。这些变量与新的应变一起，作为最终前馈网络 sFNN$^{\psi}$ 的输入，用于预测新的自由能。对 sFNN$^{\psi}$ 求导即可得到新的应力。

在训练数据集中实际可用的内变量，在式 (26) 中记为 $\boldsymbol{\xi}_{\text{av}}$，表示 N_{av} 可用内变量。例如，这可能是可测量的塑性应变。因此，如果没有已知的内变量，则省略此项，网络必须自行找到合理的内变量表示。自由能项 $w^{\psi}\mathcal{L}^{\psi}$ 也是可选的，如果自由能未知，则可以省略。

4.3 以强形式施加物理约束的神经网络

此处将最后一类模型称为以强形式施加物理约束的神经网络。与前一模型类别不同，该类模型通过模型构造先验地满足 $\mathcal{D} \geq 0$。为此，将第 2.1.2 节概述的广义标准材料概念引入数据驱动范式，并采用输入凸神经网络[51]。Huang 等人[49] 以及 As'ad 和 Farhat[7]已将这一方法用于黏弹性。下文将详细介绍三个模型：前馈神经网络$\psi{+}\phi$、前馈神经网络$\psi{+}\phi^{*}$和前馈神经网络$\psi{+}\phi{+}\xi$。

注 2 应当注意，将所介绍的方法前馈神经网络$\psi{+}\phi$ 和前馈神经网络$\psi{+}\phi^{*}$ 应用于属于率无关材料的数据时，会自动导致正则化[59],，即用率相关模型近似这些数据。这是因为所选激活函数无法表示塑性中典型的非光滑耗散势，参见表1。

4.3.1 以自由能和耗散势为输出的前馈神经网络架构 (FNN$^{\psi+\phi}$)

此类别中的第一种架构 FNN$^{\psi+\phi}$由两个前馈神经网络组成，分别对自由能 ψ和耗散势 ϕ进行建模。根据第 2.1 节对 ϕ的要求，即 (i) 在 $\dot{\boldsymbol{\xi}}$中具有凸性，(ii) $\phi(\dot{\boldsymbol{\xi}}=\mathbf{0},\boldsymbol{\xi},\varepsilon)=0$，以及 (iii) $\phi(\dot{\boldsymbol{\xi}}=\mathbf{0},\boldsymbol{\xi},\varepsilon)$ 在 $\dot{\boldsymbol{\xi}}$中达到最小值，该模型可用于生成先验热力学一致的预测。通过乘法分解，耗散势仅相对于 $\dot{\boldsymbol{\xi}}$（而非 $\boldsymbol{\xi}$和 ε）的凸性得以实现

$$\phi^{\text{NN}}(\dot{\boldsymbol{\xi}},\boldsymbol{\xi},\varepsilon) := \phi^{\text{con}}(\dot{\boldsymbol{\xi}})\,\phi^{+}(\varepsilon,\boldsymbol{\xi}) \tag{28}$$

分解为凸部分 $\phi^{\text{con}}(\dot{\boldsymbol{\xi}})$（仅取决于内变量的变化率）和正部分 $\phi^{+}(\varepsilon,\boldsymbol{\xi})$（取决于应变和内变量本身）。每个部分均由单个前馈神经网络建模。要使 ϕ^{con} 具有凸性，需要使用凸且非递减的激活函数，此处采用 Softplus 激活函数 $\text{SP}(x) := \ln(1+\exp(x))$，并要求整个网络中的权重均为非负值[8,16,51]。而要使 ϕ^{+} 为正，只需保证输出层的权重和偏置为非负值，并使用正激活函数，此处同样采用 Softplus 激活函数。构建网络的另一种更通用的方法是使用部分输入凸神经网络 (PICNN)，如 Amos 等人[51] 所述；这种网络仅对部分输入具有凸性。为满足归一化条件 (ii) 和 (iii)，需通过以下方式修改式 (28) 中组合网络的输出 ϕ^{NN}

$$\phi(\dot{\boldsymbol{\xi}},\boldsymbol{\xi},\varepsilon) := \phi^{\text{NN}}(\dot{\boldsymbol{\xi}},\boldsymbol{\xi},\varepsilon) - \phi^{\text{NN}}(\dot{\boldsymbol{\xi}}=\mathbf{0},\boldsymbol{\xi},\varepsilon) - \frac{\partial\phi^{\text{NN}}(\dot{\boldsymbol{\xi}}=\mathbf{0},\boldsymbol{\xi},\varepsilon)}{\partial\dot{\boldsymbol{\xi}}}\cdot\dot{\boldsymbol{\xi}} \geq 0 \;\forall\dot{\boldsymbol{\xi}},\boldsymbol{\xi},\varepsilon \quad . \tag{29}$$

同样，试探形式

$$\psi(\varepsilon,\boldsymbol{\xi}) := \psi^{\text{NN}}(\varepsilon,\boldsymbol{\xi}) - \psi^{\text{NN}}(\varepsilon=0,\boldsymbol{\xi}=\mathbf{0}) - \frac{\partial\psi^{\text{NN}}(\varepsilon=0,\boldsymbol{\xi}=\mathbf{0})}{\partial\varepsilon}\varepsilon - \frac{\partial\psi^{\text{NN}}(\varepsilon=0,\boldsymbol{\xi}=\mathbf{0})}{\partial\boldsymbol{\xi}}\cdot\boldsymbol{\xi} \tag{30}$$

可确保自由能、应力和内力在初始无载状态下均为零[49]。

接下来，在给出上述模型表述后，将说明训练算法以及如何利用经过校准的前馈神经网络$\psi{+}\phi$ 架构预测新的时间步。为此，训练 ϕ 和 ψ 需要应力 σ、应变 ε和内变量 $\boldsymbol{\xi}$ 组成的元组，并使用方程（29）和（30）。[2]

三个网络 $\phi^{\text{con}}(\dot{\boldsymbol{\xi}})$、$\phi^{+}(\varepsilon,\boldsymbol{\xi})$ 和 $\psi^{\text{NN}}(\varepsilon,\boldsymbol{\xi})$ 的训练过程如下：首先，根据公式 (30) 预测 ${}^{n}\psi$，并通过求导得到 ${}^{n}\sigma$ 和 ${}^{n}\tau^{\alpha}$。根据公式 (4)，内力应等于耗散势对相应内变量变化率的导数。该变化率通过 ${}^{n}\dot{\xi}^{\alpha} := ({}^{n}\xi^{\alpha} - {}^{n-1}\xi^{\alpha})/{}^{n-1}\Delta t$ 近似。预测的内力 ${}^{n}\tau^{\alpha} = -\frac{\partial\psi}{\partial{}^{n}\xi^{\alpha}}$ 与

[2]为了仅使用应力和应变进行训练，也可以在每个训练轮次中对 $\boldsymbol{\xi}$ 的完整路径进行积分。然而，由于这需要对所有时间步进行反向传播，训练将耗费更多时间[52]。另一种方法是按照 As'ad 和 Farhat[7], 的做法，在训练期间使用另一个网络来近似内变量 $\boldsymbol{\xi}$。这样也可以得到仅包含 σ 和 ε 元组的训练数据集，与第一种方法相比，可缩短训练时间。还可以在预处理步骤中预先确定内变量。Ladevèze 等人[60] 或 Gerbaud 等人[61]介绍了具体做法。此处采用第二种方法，见第 4.3.3 节。

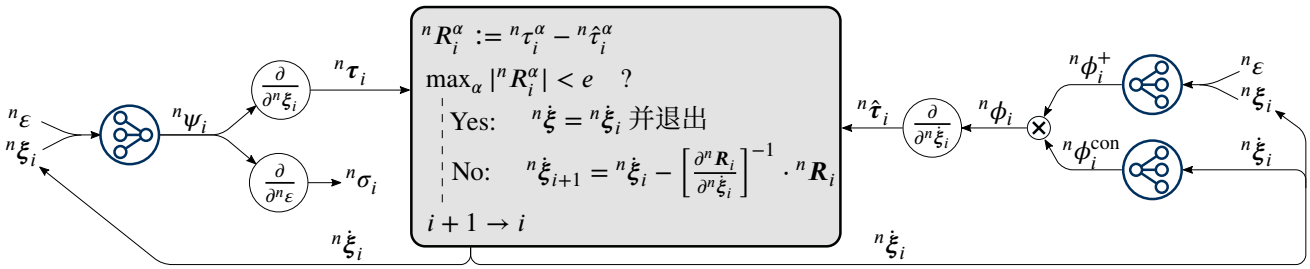


图 9：前馈神经网络$^{\psi+\phi}$ 架构的功能：利用当前迭代的内变量变化率，通过自由能网络（$\boldsymbol{\tau}$，左）和耗散势网络（$\hat{\boldsymbol{\tau}}$，右）计算内力。随后对内变量变化率进行迭代调整，使得 $\boldsymbol{\tau} \approx \hat{\boldsymbol{\tau}}$。

${}^n\hat{\tau}^\alpha := \frac{\partial \phi}{\partial {}^n\dot{\xi}^\alpha}$ 以及应力预测均纳入损失函数

$$\mathcal{L} := w^\sigma \mathcal{L}^\sigma + w^{\text{Biot}} \mathcal{L}^{\text{Biot}} \quad \text{with} \quad \mathcal{L}^\sigma := \text{MAE}(\sigma) \quad \text{and} \quad \mathcal{L}^{\text{Biot}} := \frac{1}{N}\sum_{\alpha=1}^{N} \text{MAE}(\tau^\alpha) \quad , \tag{31}$$

这是过程中需要最小化的量。

训练完成并找到 ψ 和 ϕ 的良好表示后，这些势可用于预测给定应变路径下的材料响应。因此，在每个时间步 n中，使用牛顿-拉夫森方法对内变量的变化率 ${}^n\dot{\xi}^\alpha$ 进行迭代调整，使得 $\max_\alpha |{}^n\tau^\alpha - {}^n\hat{\tau}^\alpha| < e$，并满足给定容差 e；其中，新的内变量通过 ${}^n\xi^\alpha = {}^{n-1}\xi^\alpha + {}^{n-1}\Delta t\, {}^n\dot{\xi}^\alpha$获得，详见图 9 所示的方法。

4.3.2 以自由能和对偶耗散势作为输出的前馈神经网络架构（FNN$^{\psi+\phi^*}$）

这里介绍的第二种架构是前馈神经网络$^{\psi+\phi^*}$ 模型，该模型根据公式（5）使用对偶耗散势$\phi^*(\boldsymbol{\tau}, \boldsymbol{\xi}, \varepsilon)$。因此，模型方程与公式（28）–（30）类似，区别在于需要将 ϕ 替换为 ϕ^*。对偶耗散势 ϕ^* 必须保证：（i）关于 $\boldsymbol{\tau}$ 的凸性；（ii） $\phi^*(\boldsymbol{\tau} = \mathbf{0}, \boldsymbol{\xi}, \varepsilon) = 0$；以及（iii） $\phi^*(\boldsymbol{\tau} = \mathbf{0}, \boldsymbol{\xi}, \varepsilon)$ 在 $\boldsymbol{\tau}$处为最小值，从而预先获得一个热力学一致模型。同样，通过类似于公式（28）的乘法分解，可以实现 ϕ^*仅关于 $\boldsymbol{\tau}$ 的凸性。

除了模型方程不同之外，FNN$^{\psi+\phi^*}$ 架构在训练算法和预测新时间步的方法上也有所不同。在 FNN$^{\psi+\phi^*}$中，使用由 σ、 ε 和 $\boldsymbol{\xi}$ 组成的数据集进行训练，过程如下：首先计算自由能 ${}^n\psi$，并通过求导得到 ${}^n\sigma$ 和 ${}^n\tau^\alpha$。随后，将内力 ${}^n\tau^\alpha$ 输入由对偶耗散势组成的组合网络，并得到 ${}^n\phi^*$。对 ${}^n\phi^*$ 关于内力求导，可得内变量的变化率 ${}^n\dot{\xi}^\alpha$。损失函数

$$\mathcal{L} := w^\sigma \mathcal{L}^\sigma + w^{\dot{\xi}} \mathcal{L}^{\dot{\xi}} \quad \text{with} \quad \mathcal{L}^\sigma := \text{MAE}(\sigma) \quad \text{and} \quad \mathcal{L}^{\dot{\xi}} := \frac{1}{N}\sum_{\alpha=1}^{N} \text{MAE}\left(\dot{\xi}^\alpha\right) \tag{32}$$

现在分别将应力和内变量变化率的预测值与其期望值 ${}^n\bar{\sigma}$ 和 ${}^n\bar{\dot{\xi}}^\alpha$进行比较，其中 ${}^n\bar{\dot{\xi}}^\alpha = ({}^n\xi^\alpha - {}^{n-1}\xi^\alpha)/{}^{n-1}\Delta t$。三个网络训练完成后，可按照图10中的模式获得某一时间步的材料状态。由于 ${}^n\xi^\alpha = {}^{n-1}\xi^\alpha + {}^{n-1}\Delta t\, {}^n\dot{\xi}^\alpha$是隐式描述的，因此内变量的演化通过牛顿-拉夫森方法迭代确定，使得在给定容差 e的情况下，
$\max_\alpha |{}^n\dot{\xi}^\alpha - {}^n\hat{\dot{\xi}}^\alpha| < e$。

4.3.3 以自由能和耗散势为输出的前馈神经网络架构的改进训练（FNN$^{\psi+\phi+\xi}$）

前两种架构，即 FNN$^{\psi+\phi}$ 和 FNN$^{\psi+\phi^*}$ 的训练，需要训练数据集中包含内变量。然而，由于这些量通常是未知的，As'ad 和 Farhat[7] 提出了一种基于对偶耗散势模型的高效训练方法，该方法不再需要内变量。这里将类似地采用这种方法来训练架构 FNN$^{\psi+\phi}$。

为避免直接指定内变量，同时保持训练速度，本文仅在训练过程中集成另一个前馈神经网络，该网络以时间 t 为输入，并输出内变量 $\boldsymbol{\xi}^{\text{NN}}(t)$。

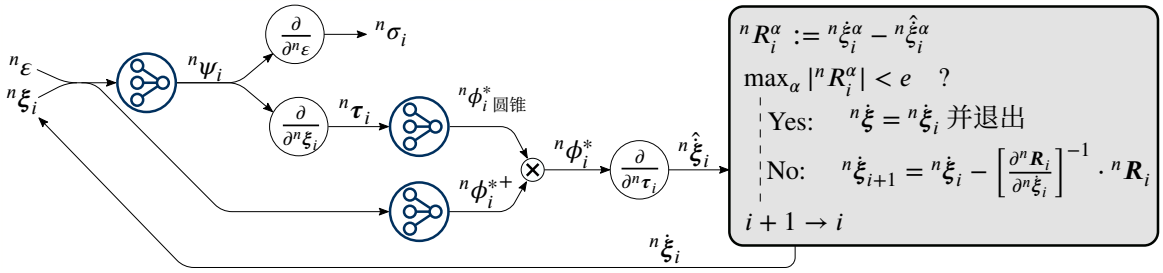


图10：前馈神经网络$\psi{+}\phi^*$架构的功能：使用当前迭代的内变量变化率，将内变量与应变一起输入自由能网络。通过求导得到应力和内力。内力传递给凸网络以计算 $\phi^{*\mathrm{con}}$；应变和内变量则输入正值网络以计算 ϕ^{*+}。将 $\phi^{*\mathrm{con}}$ 与 ϕ^{*+} 相乘得到对偶耗散势，再对内力求导以获得 $\hat{\dot{\xi}}$。随后迭代调整内变量变化率，使得 $\dot{\xi} \approx \hat{\dot{\xi}}$。

自由能网络和耗散势网络所需的内变量值及其变化率取自该辅助网络，其中变化率通过 ${}^n\dot{\xi}^{\mathrm{NN}} = \frac{1}{\Delta t}\left(\xi^{\mathrm{NN}}({}^nt) - \xi^{\mathrm{NN}}({}^{n-1}t)\right)$进行一致近似，以匹配预测过程。通过从实际输出中减去 $\xi^{\mathrm{NN}}(0)$，强制满足条件 $\xi^{\mathrm{NN}}(0) = \mathbf{0}$。在训练过程中，调整该网络的权重，使得预测的 $\xi^{\mathrm{NN}}(t)$ 时间历程能够令应力（$\mathcal{L}^\sigma$）和比奥项（$\mathcal{L}^{\mathrm{Biot}}$）的误差尽可能小。这样，无需明确指定内变量的值，也能合理地表示内变量。训练完成后，辅助网络便不再需要，预测过程可按图 9 所述的相同方式进行。为了让优化器更容易完成训练，辅助网络需要表示的 $\xi^{\mathrm{NN}}(t)$ 历程应当平滑且相对简单。为此，按照第 5.1 节所述，选择平滑应变路径。As'ad 和 Farhat[7] 还提出将自由能加性分解为平衡部分和耗散部分。为保持与前馈神经网络$\psi{+}\phi$的可比性和一致性，本文未采用这种分解。不过需要指出的是，这一额外假设能显著简化训练。

除了损失函数方程（31）外，还添加了另一项 $w^{|\xi|}\mathcal{L}^{|\xi|}$。此项旨在使内变量保持较小。否则，所选内变量会变得无限大。为说明这一点，考虑一个自由能函数为 $\psi = \psi(\varepsilon, \xi) = \frac{1}{2}E(\varepsilon - \xi)^2$ 的本构模型，其对应的应力和内力为$\sigma = \tau = E(\varepsilon - \xi)$。使用另一个内变量 $\tilde{\xi}$ 的等价表述式为 $\tilde{\psi}(\varepsilon, \tilde{\xi}) = \frac{1}{2}E(\varepsilon - \frac{1}{A}\tilde{\xi})$，其中 $\tilde{\xi} = A\xi$、$\tilde{\sigma} = E(\varepsilon - \frac{1}{A}\tilde{\xi}) = \sigma$ 且 $\tilde{\tau} = \frac{1}{A}E(\varepsilon - \frac{1}{A}\tilde{\xi})$。损失项 $\mathcal{L}^{\mathrm{Biot}}$ 倾向于使内力的绝对值较小，~即采用较大的缩放因子 $|A|$ 进行表示，从而得到 $|\xi| \gg |\xi|$。~这会给之后的预测过程带来困难。因此，新的损失函数为

$$\mathcal{L} := w^\sigma \mathcal{L}^\sigma + w^{\mathrm{Biot}}\mathcal{L}^{\mathrm{Biot}} + w^{|\xi|}\mathcal{L}^{|\xi|} \quad \text{with} \quad \mathcal{L}^{|\xi|} := \frac{1}{N^{\mathrm{ds}}}\sum_{n=1}^{N^{\mathrm{ds}}}\sum_{\alpha=1}^{N}|\xi_\alpha^{\mathrm{NN}}({}^nt)| \quad , \tag{33}$$

其中，$\mathcal{L}^{|\xi^{\mathrm{NN}}|}$ 会惩罚所选内变量的较大绝对值。请注意，这并不意味着 $\mathrm{sign}(\xi) = \mathrm{sign}(\xi)$，因为 $A < 0$ 仍然是有效选项。在本研究范围内，仅预测一个内变量。也可以使用针对另一个内变量的另一个网络，或具有两个输出的网络，但本文不予展示。

5 用于训练和验证的数据库生成

在研究基于神经网络的材料模型的可用性之前，必须先生成一个数据库，用于训练过程，并在之后作为验证的参考。因此，采用第2.2.1节和第2.2.2节所述的四种不同模型，生成对应于指定应变路径的材料状态：一个含有一个麦克斯韦元件的非线性黏弹性模型（V1）、一个含有两个麦克斯韦元件的线性黏弹性模型（V2）、一个采用运动硬化的弹塑性模型（P1），以及一个采用运动-各向同性混合硬化的弹塑性模型（P2）。通过应用隐式欧拉格式，求解表1中给出的模型控制方程。这些参考本构模型所选用的参数见表2。

5.1 训练数据的生成

采用两种不同的方法生成训练数据。第一种方法适用于除前馈神经网络$^{\psi+\phi+\xi}$之外的所有架构，并使用随机游走序列。前馈神经网络$^{\psi+\phi+\xi}$需要一条平滑应变路径，该路径由一组三次样条节点拟合而成。下面将详细介绍这两种方法。

5.1.1 随机游走应变路径

针对每种待研究材料 V1、V2、P1 和 P2，分别生成一个数据库。由于循环架构的训练数据需要包含多个相互独立的序列，因此数据点以多个序列的形式生成。对于前馈架构，则将这些序列重新拆分为独立的时间步，以获得所需的输入-输出对或数据元组。每个序列都从初始材料状态出发，围绕应变路径进行随机游走，初始状态为 ${}^0\sigma = {}^0\varepsilon = {}^0\xi^\alpha = 0$。在时间增量 ${}^0\Delta t$ 内，对初始状态施加应变增量 ${}^0\Delta\varepsilon$；该增量从均值为 0、标准差为 $s^{\Delta\varepsilon}$ 的正态分布中采样，而时间增量 ${}^0\Delta t$ 则从均匀分布 ${}^0\Delta t \in (\Delta t_{\min}, \Delta t_{\max})$ 中采样。在此区间内，应变率保持恒定。随后，在另一个时间增量 ${}^1\Delta t$ 内施加另一个应变增量 ${}^1\Delta\varepsilon$，以此类推，直到序列包含 N^{ts} 个时间步。研究表明，对于循环架构而言，长度为 $N^{\text{ts}} = 100$ 的序列表现良好，且与所研究的材料无关。为将应变限制在合理范围内，其绝对值不得超过 $|\varepsilon|_{\max}$。也就是说，如果 $|{}^n\varepsilon + {}^n\Delta\varepsilon| > |\varepsilon|_{\max}$，则重新采样应变增量 ${}^n\Delta\varepsilon$，直至$|{}^n\varepsilon + {}^n\Delta\varepsilon| \leq |\varepsilon|_{\max}$。随机游走的参数选取为 $s^{\Delta\varepsilon} = 0.25\,\%$、$|\varepsilon|_{\max} = 2\,\%$、$\Delta t_{\min} = 0.02\,\text{s}$ 和 $\Delta t_{\max} = 0.1\,\text{s}$。作为示例，图11a展示了 V1 数据库中第一个序列的应变路径及相应的应力响应。

实际训练数据集，即训练过程中实际使用的数据，包含从该数据库中提取的若干 N^{seq} 序列或 N^{ds} 数据元组。研究发现，对于循环架构，$N^{\text{seq}} = 100$个序列已足够；对于前馈架构，$N^{\text{ds}} = 1000$个数据元组已足够，且不会降低预测质量，与所要建模的材料行为无关。从每个数据库中选取数据时，先从第一个序列开始；若选取数据元组，则从第一个序列的第一个时间步开始，然后按时间顺序继续选取。这意味着，所有前馈架构的训练数据集都包含来自前 10 个序列中相同材料状态的信息，而循环架构则使用相同的 100 个序列。

5.1.2 平滑应变路径

为了借助辅助网络实现内变量的神经网络表示，FNN$^{\psi+\phi+\xi}$ 架构要求训练数据由更简单、更平滑的应变路径生成。该平滑应变路径由一条三次样条描述，它连接了在 ε-t 空间[7]中人工选定的一组节点。定义样条后，将该应变路径应用于解析模型，并以均匀分布的时间增量$\Delta t \in (0.01\,\text{s}, 0.02\,\text{s})$ 生成数据点，直至数据集包含 900 个时间步。所有材料 V1、V2、P1 和 P2 均使用相同的应变路径；图 11b 展示了该路径以及 V1 对应的应力响应。节点的选择使得沿此路径生成的数据与随机游走序列中的数据具有可比性。

5.2 验证数据的生成

为评估和比较网络对未见数据的性能，网络需要预测图 12a 所示应变路径中一系列连续时间步的本构响应，其中采用恒定的时间增量 $0.05\,\text{s}$。选择该路径时，应变、应变率和时间增量均不超出训练范围。

表 2：用于生成训练数据和验证数据的四种测试材料模型所选参数。

描述	标签	参数
带有.....的非线性黏弹性模型 一个麦克斯韦元件	V1	$E = 1\,\text{GPa}$, $E_1 = 10\,\text{GPa}$, $\hat{\eta}_1 = 200\,\text{GPa s}$, $a_1 = -0.7$, $b_1 = 0.1$
含两个元件的线性黏弹性模型 麦克斯韦元件	V2	$E = 1\,\text{GPa}$, $E_1 = 10\,\text{GPa}$, $\eta_1 = 10\,\text{GPa s}$, $E_2 = 20\,\text{GPa}$, $\eta_2 = 5\,\text{GPa s}$
含运动学 硬化	P1	$E = 20\,\text{GPa}$, $\sigma_y = 100\,\text{MPa}$, $H = 10\,\text{GPa}$
混合运动-各向同性硬化弹塑性模型 运动-各向同性混合硬化	P2	$E = 20\,\text{GPa}$, $\sigma_y = 100\,\text{MPa}$, $H = 3\,\text{GPa}$, $\hat{H} = 3\,\text{GPa}$

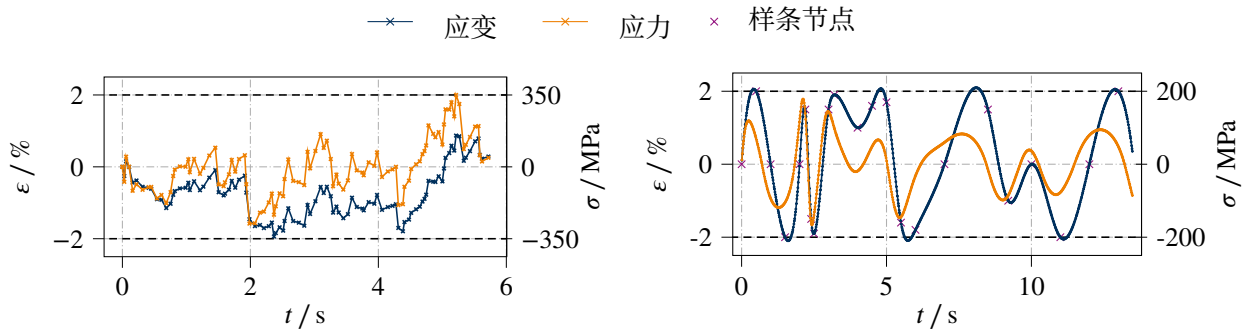


图11：非线性黏弹性材料（V1）训练数据中的应力和应变随时间的变化：**(a)**取自非光滑随机游走数据库的第一个序列，以及**(b)** 使用三次样条生成、供前馈神经网络$^{\psi+\phi+\xi}$使用的平滑路径。

因此，在此基准测试中，网络无需进行外推。与训练数据类似，参考数据点采用隐式欧拉格式获得。

其次，生成一条用于进一步验证的路径，其应变、应变率和时间增量均超出训练范围，见图12b。稍后将用它来评估不同模型的外推能力。该路径包括一段超出训练应变界限（$2\,\mathrm{s} < t \leq 4\,\mathrm{s}$）的小滞回，以及一个由16个时间步组成的短区间（$5\,\mathrm{s} < t \leq 5.08\,\mathrm{s}$），其间 $\Delta t = 0.005\,\mathrm{s}$ 和 $\dot{\varepsilon} = 62.5\,\%\,\mathrm{s}^{-1}$低于 Δt 的界限，并在另一段滞回中超过 $\dot{\varepsilon}$ 的界限。随后，在应变保持恒定为 $\varepsilon = 0\,\%$ 的情况下进行短暂松弛，直至$t = 6\,\mathrm{s}$。接着，施加超出训练范围的大时间增量（分别在$6\,\mathrm{s} < t \leq 7\,\mathrm{s}$和 $7\,\mathrm{s} < t \leq 8\,\mathrm{s}$中为$\Delta t = 0.125\,\mathrm{s}$和 $\Delta t = 0.2\,\mathrm{s}$）。其余所有时间步的时间增量均为 $\Delta t = 0.05\,\mathrm{s}$。最后，应变线性增加，直至 $t = 10\,\mathrm{s}$ 和 $\varepsilon = 6\,\%$。

5.3 归一化

大多数常见激活函数的一阶和二阶导数在自变量绝对值较大时会趋于零，导致无法进行训练。为确保梯度不为零，必须将训练数据集缩放至量级为 1 的数值范围内，例如将其中数量级为若干 MPa 的应力缩小。由于训练数据集中的相关特征通常并非彼此独立，因此必须恰当选择缩放因子，以下将对此进行说明。

˜一般而言，独立量 f 可通过以下方式归一化至范围 $(-1, 1) \ni \tilde{f}$

$$\tilde{f} = \frac{f - m_f}{s_f} \quad \Longleftrightarrow \quad f = s_f \tilde{f} + m_f \quad \text{with} \quad m_f = \frac{1}{2}(f_{\max} + f_{\min}) \quad \text{and} \quad s_f = \frac{1}{2}(f_{\max} - f_{\min}) \quad , \tag{34}$$

˜其中，f 是 f 的缩放值，而 $f_{\max}$ 和 $f_{\min}$ 分别表示 f 在整个数据集中的最大值和最小值[21]。

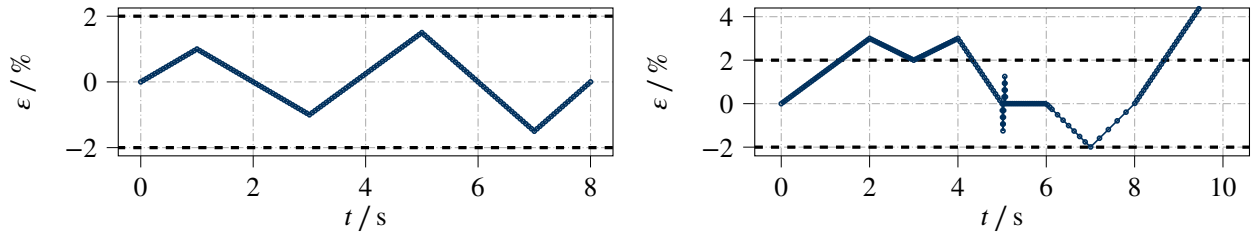


图12：用于评估网络插值 **（a）**和外推 **（b）**性能的验证路径。（a）中的路径经过设计，使应变和时间增量均不超出训练范围。（b）中的路径则分别超出这些界限，并包含应变率较大的区段。黑色虚线表示训练数据界限。

对于黑箱模型前馈神经网络（FNN）$^{\sigma}$ 和循环神经网络（RNN）$^{\sigma}$，必要量 ε、 σ 和 Δt 可视为相互独立，因此可使用公式（34）求得缩放因子。对于所有其他架构，并非每个特征都与其他特征相互独立，这是由微分算子的存在所致。此外，合理的做法是选择适当的缩放因子，使未缩放量之间成立的所有关系在其缩放值之间也成立，从而避免在训练过程中进行反变换。这可以简化并加快损失函数的计算。为简单起见，以下将所有m_f 设为零。对于其余所有架构，相关量为ε、 σ、 Δt、 ξ、 τ、 ψ、 ϕ 和 ϕ^*。将 ε、 Δt 和 ψ 选为相互独立的量后，可使用公式（34）求得其缩放因子$_4$。其余缩放因子则通过在所述条件下求解相关方程得出，结果如下：

$$s_\sigma = \frac{s_\psi}{s_\varepsilon} \quad , \quad s_\xi = s_\varepsilon \quad , \quad s_\tau = \frac{s_\psi}{s_\varepsilon} \quad , \quad s_\phi = \frac{s_\psi}{s_{\Delta t}} \quad \text{and} \quad s_{\phi^*} = \frac{s_\psi}{s_{\Delta t}} \quad . \tag{35}$$

对相应数据集中的所有特征进行缩放，训练仅使用这些缩放后的值。因此，网络本身仅预测缩放后的值。通过反变换方程（34）$_2$ 可得到熟悉的物理单位下的数值。现在可以使用训练好的网络，根据给定的应变路径预测材料响应。

6 应用

本节基于第 5 节所述方法生成的数据，对第 4 节介绍的七种架构，即 FNN$^{\sigma}$、RNN$^{\sigma}$、FNN$^{\xi+\psi}$、RNN$^{\xi+\psi}$、FNN$^{\psi+\phi}$、FNN$^{\psi+\phi^*}$和 FNN$^{\psi+\phi+\xi}$，分别进行测试，以评估其在训练后重现材料 V1、V2、P1 和 P2 本构响应的能力。评估依据是基于神经网络的模型对图 12a 所示未知应变路径预测的应力响应；该路径经过选择，使应变、应变率和时间增量均不超出训练范围。此外，为分析外推能力，还考察了模型对第二条验证路径的预测响应；该路径的应变、应变率和时间增量均超出训练范围，参见图 12b。为将本研究的范围控制在合理水平，第二条验证路径仅针对 V1 进行考察。所有使用的前馈神经网络均由输入层、单个隐藏层和输出层组成。隐藏层中的神经元数量记为 N_2。

6.1 无外推的训练结果与验证路径

6.1.1 黑箱模型

前馈神经网络$^{\sigma}$ 对于每种测试材料模型，均按照第 4.1.1 节所述，使用表 3 中给出的超参数和训练结果创建一个前馈神经网络$^{\sigma}$。模型预测结果见图 13。

黏弹性方面的结果表明，使用 FNN$^{\sigma}$ 对材料 V1 进行应力行为预测并无困难。相比之下，对于 V2，只有在输入中包含两个前置时间步（即 $N^{\mathrm{pt}} = 2$）时，才能进行精确预测。架构的这些局限性可归因于给定特定输入时，预期输出存在歧义。也就是说，给定 V1 的材料参数，仅凭输入中给出的 ${}^{n}\varepsilon$ 和 ${}^{n}\sigma$ 的值，就能确定过应力 σ_1^{ov}。因此，材料状态可以被明确确定。另一方面，对于材料 V2，依据这些量只能确定过应力之和 $\sigma_1^{\mathrm{ov}} + \sigma_2^{\mathrm{ov}}$，而无法确定各个过应力的具体值。因此，内部状态无法得到清晰描述，应力预测也无法做到精确，如图 13b 所示。为此，网络还需要获得额外一个时间步的信息，另见图 13b。

表 3：前馈神经网络$^{\sigma}$ 架构：网络超参数，包括前置时间步数 N^{pt}、隐藏层中的神经元数 N_2和激活函数 $\mathcal{A}_2^l$；以及训练的特征值，包括训练数据数量 N^{ds}和经过给定迭代次数后的损失值 $\mathcal{L}$。对应验证载荷工况的图号列于“验证”栏。

材料	N^{pt}	N_2	$\mathcal{A}_2^l$	N^{ds}	迭代次数	$\mathcal{L}$	Val.
V1	1	15	tanh	1000	3094	$2 \cdot 10^{-4}$	图 13a
V2	1	25	tanh	1000	911	$2 \cdot 10^{-3}$	图 13b
V2	2	25	tanh	1000	2903	$1 \cdot 10^{-4}$	图 13b
P1	1	15	ReLU	1000	356	$3 \cdot 10^{-6}$	图 13c
P2	1	15	ReLU	1000	1240	$8 \cdot 10^{-3}$	图 13d

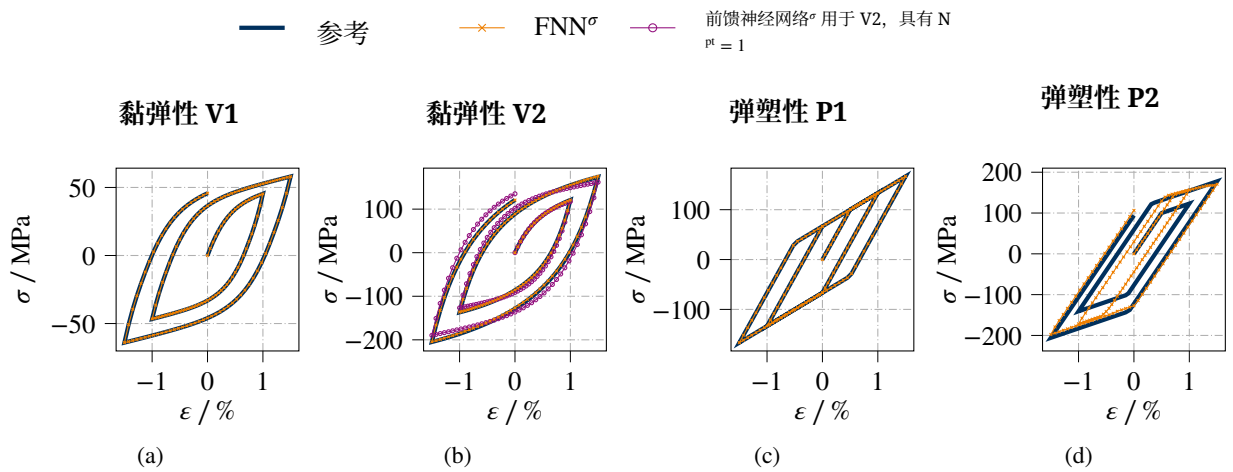


图13：前馈神经网络$^{\sigma}$ 对四种基础测试材料的应力预测。所选超参数见表 3。

与 V1 类似，具有运动硬化的弹塑性材料 P1 的状态可由 ${}^{n}\varepsilon$ 和 ${}^{n}\sigma$ 充分描述。然而，与 V2 不同，P2 中的运动-各向同性混合硬化弹塑性无法通过在输入中增加一个时间步来描述。为此，需要提供有关弹性区域当前位置和范围的信息，而固定数量的前置时间步无法提供这些信息。有趣的是，该模型会尽可能准确地近似具有运动硬化的材料行为。
总而言之，前馈神经网络$^{\sigma}$ 因此仅适用于某些类型的材料行为。此外，除应力预测外，无法获得更多信息。另一方面，训练过程计算开销较低，数据集仅需包含应力-应变对和时间步。

循环神经网络$^{\sigma}$ 现在使用由 V1、V2、P1 和 P2 生成的相同参考数据，研究 Wu 等人[29] 提出的循环神经网络$^{\sigma}$ 模型的能力，所使用的网络超参数见表 4。由此，在验证测试中获得了图14所示的结果。循环神经网络$^{\sigma}$ 模型因此能够对所有四种测试材料的应力响应进行近乎完美的预测。由于权重数量较多，且循环网络单元的梯度计算开销较大，这种广泛的适用性需要付出更长训练时间的代价。与前馈神经网络$^{\sigma}$类似，训练数据集只需包含应力和应变，但无法获得额外的物理信息。不过，前馈神经网络$^{\sigma}$ 和循环神经网络$^{\sigma}$这两种模型都无法说明各自描述的过程是否嵌入了有意义的热力学框架。

表 4：循环神经网络$^{\sigma}$ 架构：网络超参数包括单元状态中的值的数量 N^{c}、前馈网络隐藏层中的神经元数量 N_2^{FF} 以及激活函数 $\mathcal{A}_2^{l}\text{FF}$；还包括训练特征值，如序列数量 N^{seq}、序列长度 N^{ts}，以及按指定迭代次数训练后的损失值 $\mathcal{L}$。

材料	N^{c}	N_2^{FF}	$\mathcal{A}_2^{l}\text{FF}$	N^{seq}	N^{ts}	迭代次数	$\mathcal{L}$	Val.
V1	6	10	tanh	100	100	3465	$9 \cdot 10^{-4}$	图 14a
V2	10	10	tanh	100	100	2219	$8 \cdot 10^{-5}$	图 14b
P1	10	10	ReLU	100	100	5352	$7 \cdot 10^{-3}$	图 14c
P2	12	20	ReLU	100	100	5281	$1 \cdot 10^{-3}$	图 14d

6.1.2 在弱形式下强制符合物理学的神经网络

前馈神经网络$^{\xi+\psi}$ 在考察了上述纯黑箱模型之后，本文分析了第一种属于弱形式下强制符合物理学的神经网络类别的方法，即 Masi 等人提出的前馈神经网络$^{\xi+\psi}$ 模型[11]。这里同样使用由 V1、V2、P1 和 P2 生成的参考数据。

采用表 5 中给出的超参数，构建了四个前馈神经网络$^{\xi+\psi}$。图 15 所示验证加载工况的结果表明，前馈神经网络$^{\xi+\psi}$ 能够准确预测应力和内变量，即

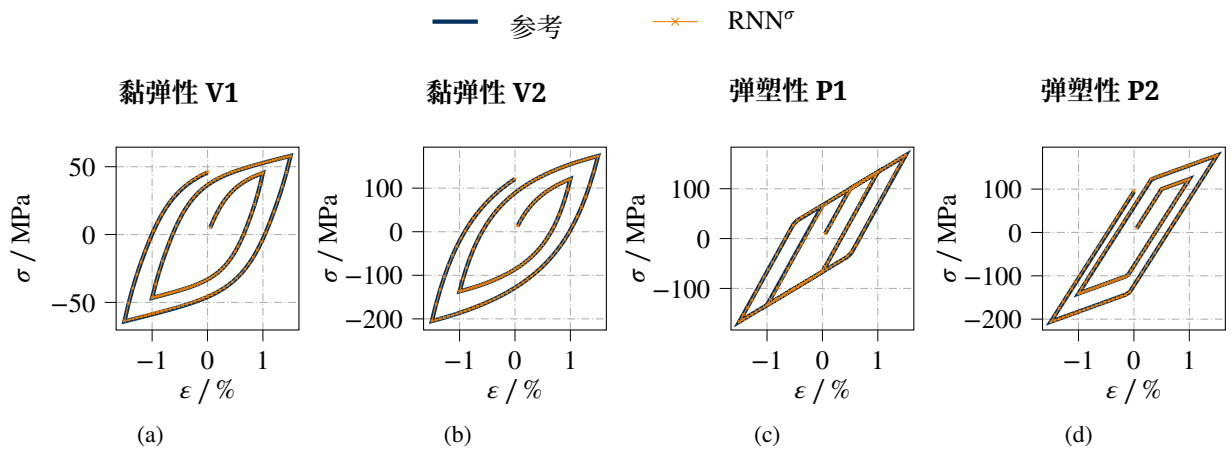


图14：RNN$^\sigma$ 对四种基础测试材料的应力预测。所选超参数见表 4。

表5：FNN$^{\xi+\psi}$ 架构：前馈网络$N_2^{\mathrm{FF}^\xi}$和 $N_2^{\mathrm{FF}^\psi}$的隐藏层神经元网络超参数，其激活函数分别为 $\mathcal{A}_2^l\mathrm{FF}^\xi$和 $\mathcal{A}_2^l$ FF$^\psi$，以及损失函数 w^σ、 $w^{\mathcal{D}\geq 0}$的权重因子、训练数据数量 N^{ds}和经过给定迭代次数后的损失值 $\mathcal{L}^\sigma$、$\mathcal{L}^{\mathcal{D}\geq 0}$。验证结果见验证部分。

材料	$N_2^{\mathrm{FF}^\xi}$	$N_2^{\mathrm{FF}^\psi}$	$\mathcal{A}^l_{\mathrm{FF}^\xi_2}$	$\mathcal{A}^l_{\mathrm{FF}^\psi_2}$	w^σ / $w^{\mathcal{D}\geq 0}$	N^{ds}	迭代次数	$\mathcal{L}^\xi$	$\mathcal{L}^\sigma$ / $\mathcal{L}^{\mathcal{D}\geq 0}$	Val.
V1	15	15	tanh	tanh	1/1	1000	5414/504	$1\cdot10^{-4}$	$1\cdot10^{-3}/4\cdot10^{-6}$	图 15a
V2	15	15	tanh	tanh	1/1	1000	1410/700	$5\cdot10^{-5}$	$2\cdot10^{-4}/4\cdot10^{-8}$	图 15b
P1	15	15	ReLU	tanh	1/1	1000	973/522	$5\cdot10^{-6}$	$2\cdot10^{-4}/4\cdot10^{-4}$	图 15c
P2	15	15	ReLU	tanh	1/1	1000	1114/608	$1\cdot10^{-7}$	$4\cdot10^{-5}/4\cdot10^{-5}$	图 15d

可用的量。除了应力-应变图之外，还添加了自由能 ψ、内变量 ξ^α和耗散率 $\mathcal{D}$ 的图示。可以看出， ψ 和 $\mathcal{D}$ 与预期结果有所偏差，尽管其数量级和总体变化趋势大多一致。条件 $\mathcal{D}\geq 0$ 得到了可靠满足。因此，与黑箱模型不同，前馈神经网络$^{\xi+\psi}$ 将预测置于更符合物理规律的框架中，这对网络的泛化能力有积极影响。有关该架构的详细研究，请参见 Masi 等人[11]。然而，前馈神经网络$^{\xi+\psi}$ 需要提供训练数据中的内变量信息，这并非易事。Masi 和 Stefanou[47] 讨论了多尺度模拟中的这一问题。

循环神经网络$^{\xi+\psi}$ 第二种所分析的、以弱形式施加物理学约束的基于神经网络的方法，是类似于 He 和 Chen[48] 所提出方法的循环神经网络$^{\xi+\psi}$ 方法。网络超参数及训练信息见表 6。基于此，针对所考虑的验证路径，可得到图 16 所示的结果。

可以看出，对于所有四种材料，应力预测都非常精确。此外，除个别时间步外，也都满足克劳修斯-迪昂不等式。然而，内变量与预期结果相差很大。由于训练数据集中没有提供内变量，也没有对其变化过程施加进一步限制，网络可以自由选择任何一组内变量，以实现准确的应力预测并满足 $\mathcal{D}\geq 0$。因此，网络找到了内变量集合和自由能函数的另一种表示形式。

注意，对于运动硬化材料 P1，网络仅预测一个内变量（$N^\xi=1$），尽管 $N=2$。这是可行的，因为塑性应变 $\varepsilon^{\mathrm{pl}}$ 与运动硬化变量 α 重合，并且材料响应可以仅用塑性应变表示。对于材料 P2，同样可以进行这种简化，其中 $N^\xi=2$ 尽管 $N=3$，但选择了 $N^\xi=3$。

与 FNN$^{\xi+\psi}$相比，要获得精确且符合物理规律的应力预测，并不明确需要提供内变量。这一优势的代价是计算开销较大的循环单元。对于 FNN$^{\xi+\psi}$ 和 RNN$^{\xi+\psi}$，尽管 $\mathcal{D}\geq 0$ 通过损失函数中的惩罚项得到强制约束，并且在该验证路径中也满足这一条件且不存在外推，但对于这类模型而言，仍无法完全保证这一点，详见第 6.2 节。

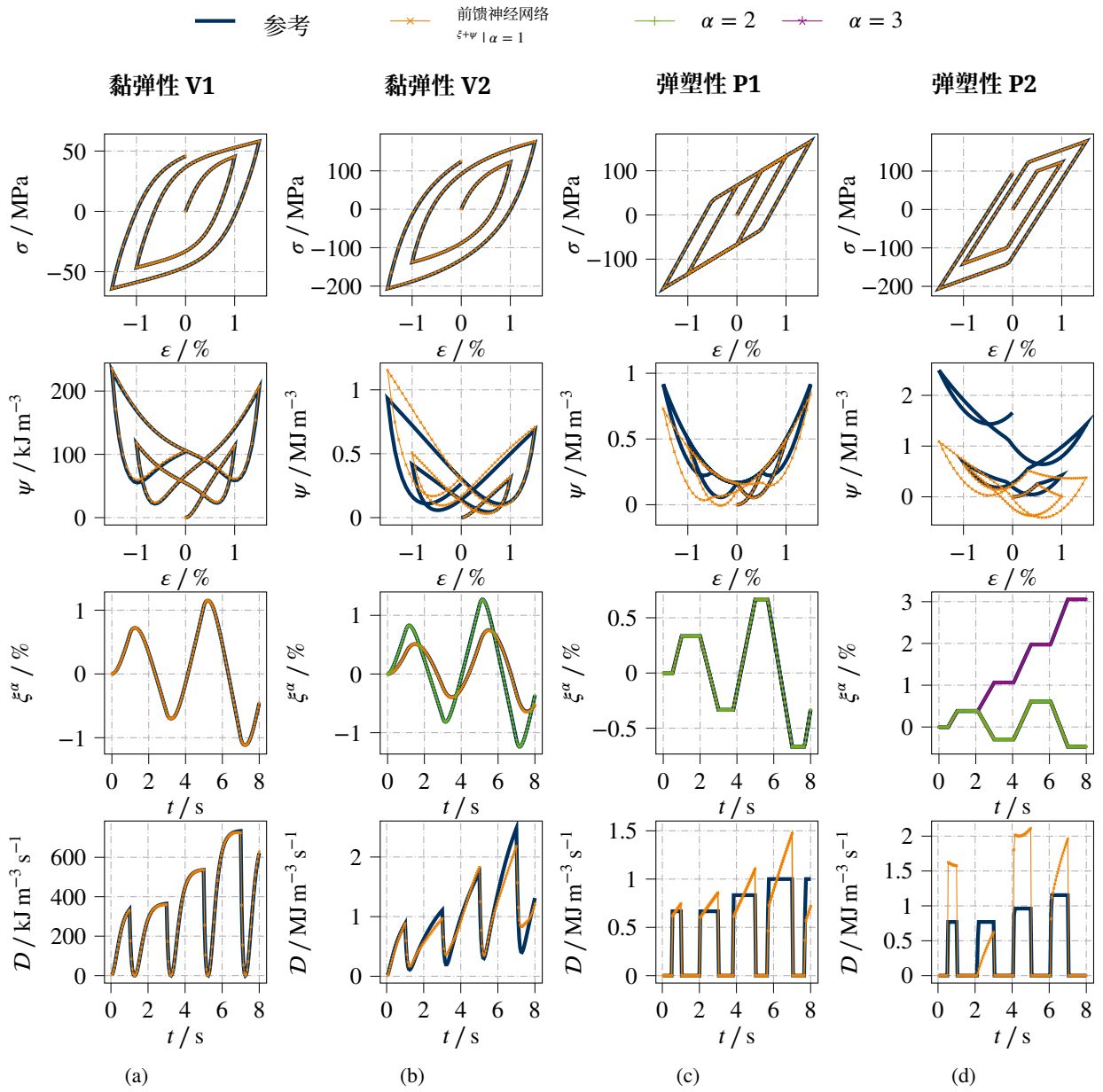


图15：使用前馈神经网络$^{\xi+\psi}$对四种测试材料的应力、自由能、内变量和耗散率进行预测。

表6：RNN$^{\xi+\psi}$ 架构：网络超参数，包括单元状态中的值数量 N^{c}、预测的内变量数量 N^{ξ}、前馈网络隐藏层中的神经元数量 $N_2^{\mathrm{FF}^{\xi}}$、 $N_2^{\mathrm{FF}^{\psi}}$，以及带有激活函数$\mathcal{A}_2^l\mathrm{FF}^{\xi}$、 $\mathcal{A}_2^l\mathrm{FF}^{\psi}$ 时损失函数的权重因子 w^{σ}、 $w^{D\geq 0}$；训练序列数量 N^{seq}、序列长度N^{ts}，以及在给定迭代次数下训练后的损失值 $\mathcal{L}^{\sigma}$、 $\mathcal{L}^{D\geq 0}$。验证载荷工况见验证部分中的相关图示。

Mat.	N^{c}	N^{ξ}	$N_2^{\mathrm{FF}^{\xi}}$	$N_2^{\mathrm{FF}^{\psi}}$	$\mathcal{A}_2^l\mathrm{FF}^{\xi}$	$\mathcal{A}_2^l\mathrm{FF}^{\psi}$	w^{σ} / $w^{D\geq 0}$	N^{seq}	N^{ts}	迭代次数	$\mathcal{L}^{\sigma}$ / $\mathcal{L}^{D\geq 0}$	Val.
V1	6	1	10	10	tanh	tanh	1/5	100	100	5079	$2\cdot10^{-3}/4\cdot10^{-7}$	图16a
V2	6	2	10	10	tanh	tanh	1/5	100	100	5244	$8\cdot10^{-3}/3\cdot10^{-5}$	图16b
P1	10	1	10	10	ReLU	tanh	1/5	100	100	5329	$7\cdot10^{-3}/1\cdot10^{-4}$	图 16c
P2	12	3	15	20	ReLU	tanh	1/5	100	100	5140	$1\cdot10^{-3}/5\cdot10^{-6}$	图 16d

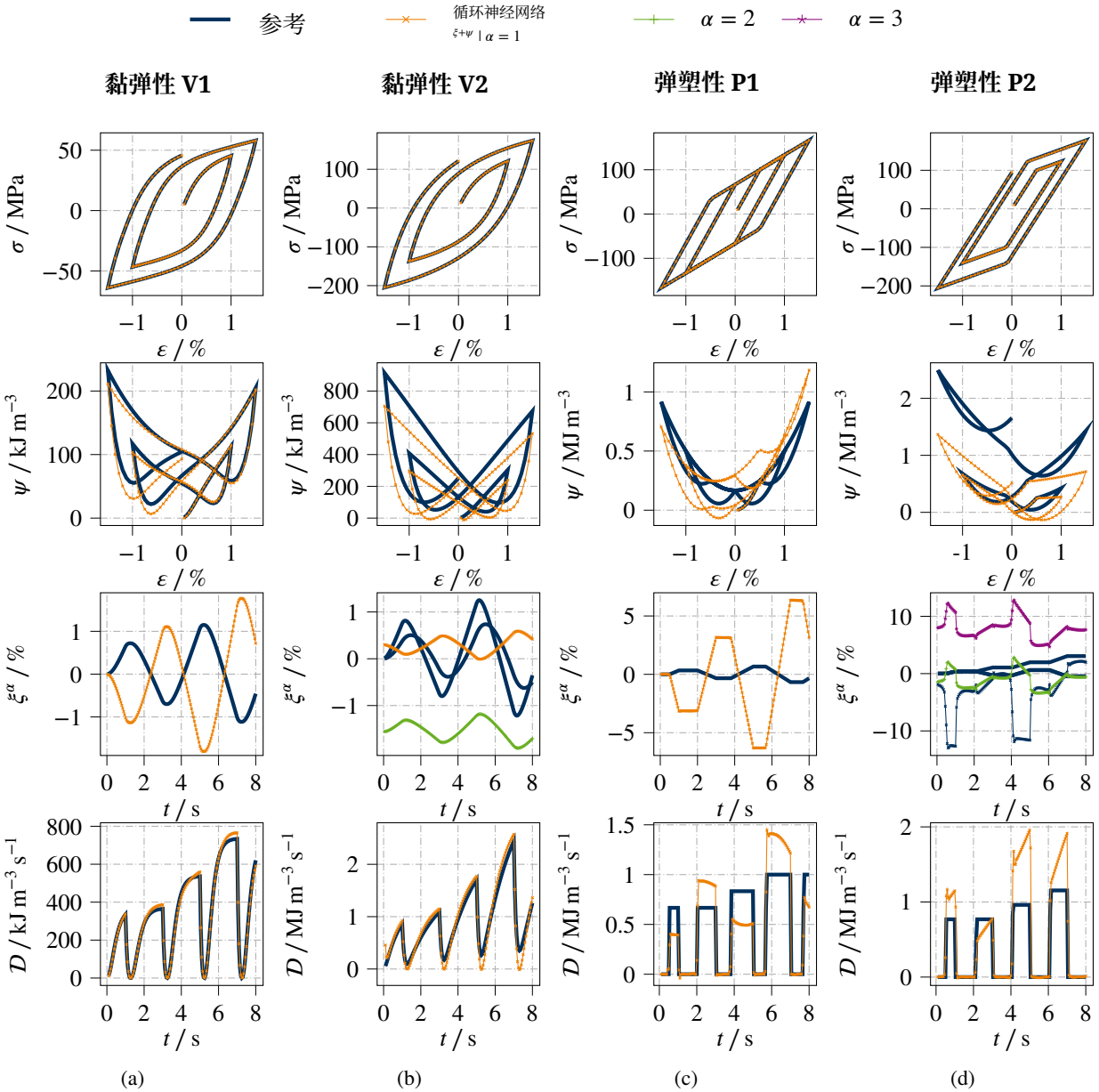


图 16：使用循环神经网络$^{\xi+\psi}$对四种测试材料的应力、自由能、内变量和耗散率进行预测。

6.1.3 以强形式施加物理约束的神经网络

前馈神经网络$^{\psi+\phi}$ 最后，使用由 V1、V2、P1 和 P2 生成的参考数据，对属于以强形式施加物理约束的神经网络 类别的方法进行分析。第一个模型是根据 Huang 等人[49]提出的 FNN$^{\psi+\phi}$ 模型。网络超参数和训练信息见表 7，所考虑验证路径所取得的结果见图 17。

可以看出，对于两种黏弹性材料 V1 和 V2，应力预测都非常精确。除了应力-应变图外，还添加了显示内变量 ξ^α、自由能 ψ、内力 τ^α、耗散率 $\mathcal{D}$和耗散势 ϕ 的图表。由此可以看出，所有这些量都几乎与参考模型一致，尽管这些值并未显式出现在损失函数中，参见公式（31）。只有对于 V1，耗散势存在明显差异。此外，与前两个模型不同，通过网络架构的构造，现在可以确保重要条件 $\mathcal{D} \geq 0$ 对所有容许加载路径均成立。然而，在未采用自适应训练时，FNN$^{\psi+\phi}$ 需要提供训练数据中的内变量信息[7]。

对于率无关材料，FNN$^{\psi+\phi}$ 模型能够重建弹塑性材料 P1 的响应。然而，有趣的是，内力 τ^α 与参考结果并不一致。这是因为 $\xi^1 = \varepsilon^{\mathrm{pl}}$ 和 $\xi^2 = \alpha$ 重合，即 $\alpha = \varepsilon^{\mathrm{pl}}$，参见表1。因此，内力存在任意多种划分方式，而这些方式都会得到相同的良好结果，即使网络中未纳入内变量之间的关系，参见备注1。此外，关于 $\mathcal{D}$ 和 ϕ的图，可以看到，零线处原本不连续的曲线形状被平滑曲线所逼近。这一过程通常称为率无关模型的正则化，即用率相关模型[59]进行逼近。由于采用了 Softplus 激活函数，这一过程会在训练过程中自动实现。对于 P2 的结果，一旦进入塑性阶段，图17中就会明显出现较差的预测结果。与 FNN$^{\sigma}$类似，应力响应仅采用运动硬化进行逼近，而不是同时采用运动硬化和各向同性硬化。根据作者的说法，这是因为不存在函数 ϕ，使得无需明确考虑塑性应变 $\varepsilon^{\mathrm{pl}}$ 与各向同性硬化变量 $\hat{\alpha}$的变化率之间的关系即可应用公式（4），参见备注1。如果将这一关系，即 $\dot{\hat{\alpha}} = |\dot{\varepsilon}^{\mathrm{pl}}|$，纳入模型架构，模型便能够做出准确的预测。作者已经对此进行了验证。

表 7：前馈神经网络$^{\psi+\phi}$ 架构：前馈网络$N_2^{\phi^{\mathrm{con}}}$，$N_2^{\phi^+}$，$N_2^{\psi}$ 隐藏层中的神经元数量等网络超参数、激活函数 $\mathcal{A}_2^{l\,\phi_{\mathrm{con}}}$、$\mathcal{A}_2^{l\,\phi^+}$、$\mathcal{A}_2^{l\,\psi}$，以及损失函数的权重因子 w^σ、w^{Biot}、训练数据数量 N^{ds} 和迭代次数，所得最终损失值为 $\mathcal{L}^\sigma$、$\mathcal{L}^{\mathrm{Biot}}$。验证栏中的数值表示验证案例。

Mat.	$N_2^{\phi^{\mathrm{con}}}$	$N_2^{\phi^+}$	N_2^{ψ}	$\mathcal{A}^{l\,\phi_{\mathrm{con}}}_2$	$\mathcal{A}_2^{l\,\phi^+}$	$\mathcal{A}_2^{l\,\psi}$	w^σ / w^{Biot}	N^{ds}	Val.	$\mathcal{L}^\sigma$ / $\mathcal{L}^{\mathrm{Biot}}$	Fig.
V1	20	20	15	SP	SP	tanh	1/0.3	1000	1251	$6\cdot10^{-4}/1\cdot10^{-2}$	图17a
V2	20	20	15	SP	SP	tanh	1/0.3	1000	995	$6\cdot10^{-5}/2\cdot10^{-4}$	图17b
P1	20	20	15	SP	SP	tanh	1/0.3	1000	1047	$2\cdot10^{-4}/2\cdot10^{-2}$	图17c
P2	20	20	15	SP	SP	tanh	1/0.3	1000	832	$4\cdot10^{-4}/6\cdot10^{-2}$	图17d

前馈神经网络$^{\psi+\phi^*}$ 该模型前馈神经网络$^{\psi+\phi^*}$ 并未采用耗散势 ϕ，而是根据 As'ad 和 Farhat[6] 采用对偶耗散势 ϕ^*。此处，网络超参数及训练信息见表8。由此，可针对所考虑的验证路径获得图18所示的结果。

与前馈神经网络$^{\psi+\phi*}$类似，对于黏弹性材料 V1 和 V2，应力、自由能、耗散率等的预测都非常精确。同样，只有 V1 的对偶耗散势存在明显差异。在未采用自适应训练时，前馈神经网络$^{\psi+\phi^*}$需要提供有关训练数据中的内变量的信息。

关于 FNN$^{\psi+\phi^*}$ 对弹塑性材料 P1 和 P2 的预测，即使对于运动硬化，图18中也显示出相当差的结果。然而，这并不令人意外，因为参考模型 P1 和 P2 的对偶耗散势 ϕ^* 具有一种无法通过所选激活函数合理表示的形状（参见表1），并且网络权重还受到限制。

前馈神经网络$^{\psi+\phi+\xi}$ 该架构对自由能和耗散势进行建模，训练时无需了解内变量。该架构仅限于单个内变量，即使用一个输出为标量的前馈神经网络来表示内变量。使用表 9 中的网络参数，可以获得图 19 所示的结果。这些结果表明，该架构能够学习 V1、V2 和 P1 势的合理表示，但无法学习 P2 势。所有预测的精度均低于前馈神经网络$^{\psi+\phi}$。与循环神经网络$^{\xi+\psi}$类似，内变量

表 8：前馈神经网络$\psi{+}\phi^*$架构：前馈网络$N_2^{\phi^{*\mathrm{con}}}$、$N_2^{\phi^{*+}}$、$N_2^{\psi}$隐藏层中神经元数量对应的网络超参数及激活函数$\mathcal{A}_2^{l\,\phi^*{}_{\mathrm{con}}}$、$\mathcal{A}_2^{l\,\phi^{*+}}$、$\mathcal{A}_2^{l\,\psi}$，以及损失函数的权重因子$w^\sigma$、$w^{\dot{\xi}}$、训练数据数量$N^{\mathrm{ds}}$、迭代次数和最终损失值$\mathcal{L}^\sigma$、$\mathcal{L}^{\dot{\xi}}$。图中展示验证载荷工况的部分链接见 Val.

Mat.	$N_2^{\phi^{*\mathrm{con}}}$	$N_2^{\phi^{*+}}$	N_2^{ψ}	$\mathcal{A}_2^{l\,\phi^*{}_{\mathrm{con}}}$	$\mathcal{A}_2^{l\,\phi^{*+}}$	$\mathcal{A}_2^{l\,\psi}$	w^σ / $w^{\dot{\xi}}$	N^{ds}	迭代次数	$\mathcal{L}^\sigma$ / $\mathcal{L}^{\dot{\xi}}$	Val.
V1	20	20	15	SP	SP	tanh	1/1	1000	2429	$1\cdot10^{-4}/1\cdot10^{-4}$	图18a
V2	20	20	15	SP	SP	tanh	1/1	1000	859	$1\cdot10^{-4}/1\cdot10^{-5}$	图 18b
P1	20	20	15	SP	SP	tanh	1/1	1000	525	$8\cdot10^{-5}/8\cdot10^{-3}$	图 18c
P2	20	20	15	SP	SP	tanh	1/1	1000	673	$1\cdot10^{-5}/9\cdot10^{-3}$	图 18d

与参考模型中的结果并不完全一致。然而，归一化后可以看出，两者的变化趋势相似。由于当$\xi(t)$与参考模型不同时，学习到的是其他势，因此会得到等效的结果。

值得注意的是，该架构为 V2 找到了一组势，尽管仅使用一个内变量，却能实现出乎意料地准确的应力预测。P1 的预测表现出相当强的速率依赖性。这可归因于第二种正则化机制：除了耗散势的正则化之外，离开弹性区域时$\dot{\varepsilon}^{\mathrm{pl}}$的骤增也无法由前馈神经网络$\dot{\xi}^{\mathrm{NN}}(t)$准确建模。反过来，P2 完全无法建模，因为描述路径依赖性至少需要两个内变量；更重要的是，原因在于上文讨论的前馈神经网络$\psi{+}\phi$的问题。因此，架构前馈神经网络$\psi{+}\phi{+}\xi$更适用于黏弹性材料。

表 9：前馈神经网络$\psi{+}\phi{+}\xi$架构：前馈网络$N_2^{\phi^{\mathrm{con}}}$、$N_2^{\phi^{+}}$、$N_2^{\psi}$、$N_2^{\xi}$的网络超参数，包括隐藏层中的神经元数量；激活函数为$\mathcal{A}_2^{l\,\phi}{}_{\mathrm{con}}$、$\mathcal{A}_2^{l\,\phi^{+}}$、$\mathcal{A}_2^{l\,\psi}$、$\mathcal{A}_2^{l\,\xi}$；以及损失函数的权重因子$w^\sigma$、$w^{\mathrm{Biot}}$、$w^{|\dot{\xi}|}$、训练数据数量$N^{\mathrm{ds}}$、迭代次数和最终损失值$\mathcal{L}^\sigma$、$\mathcal{L}^{\mathrm{Biot}}$。验证载荷工况见验证部分。

Mat.	$N_2^{\phi^{\mathrm{con}}}$	$N_2^{\phi^{+}}$	N_2^{ψ}	N_2^{ξ}	$\mathcal{A}_2^{l\,\phi_{\mathrm{与}}}$	$\mathcal{A}_2^{l\,\phi^{+}}$	$\mathcal{A}_2^{l\,\psi}$	$\mathcal{A}_2^{l\,\xi}$	$w^\sigma/w^{\mathrm{Biot}}/w^{|\dot{\xi}|}$	N^{ds}	迭代次数	$\mathcal{L}^\sigma$ / $\mathcal{L}^{\dot{\xi}}$	Val.
V1	20	20	15	40	SP	SP	tanh	tanh	10/1/1	900	920	$5\cdot10^{-2}/2\cdot10^{-1}$	19a
V2	20	20	15	40	SP	SP	tanh	tanh	10/1/1	900	1161	$4\cdot10^{-2}/3\cdot10^{-1}$	19b
P1	20	20	15	75	SP	SP	tanh	tanh	10/1/1	900	1588	$4\cdot10^{-2}/1\cdot10^{-1}$	19c
P2	20	20	15	75	SP	SP	tanh	tanh	10/1/1	900	1163	$3\cdot10^{-2}/2\cdot10^{-1}$	19d

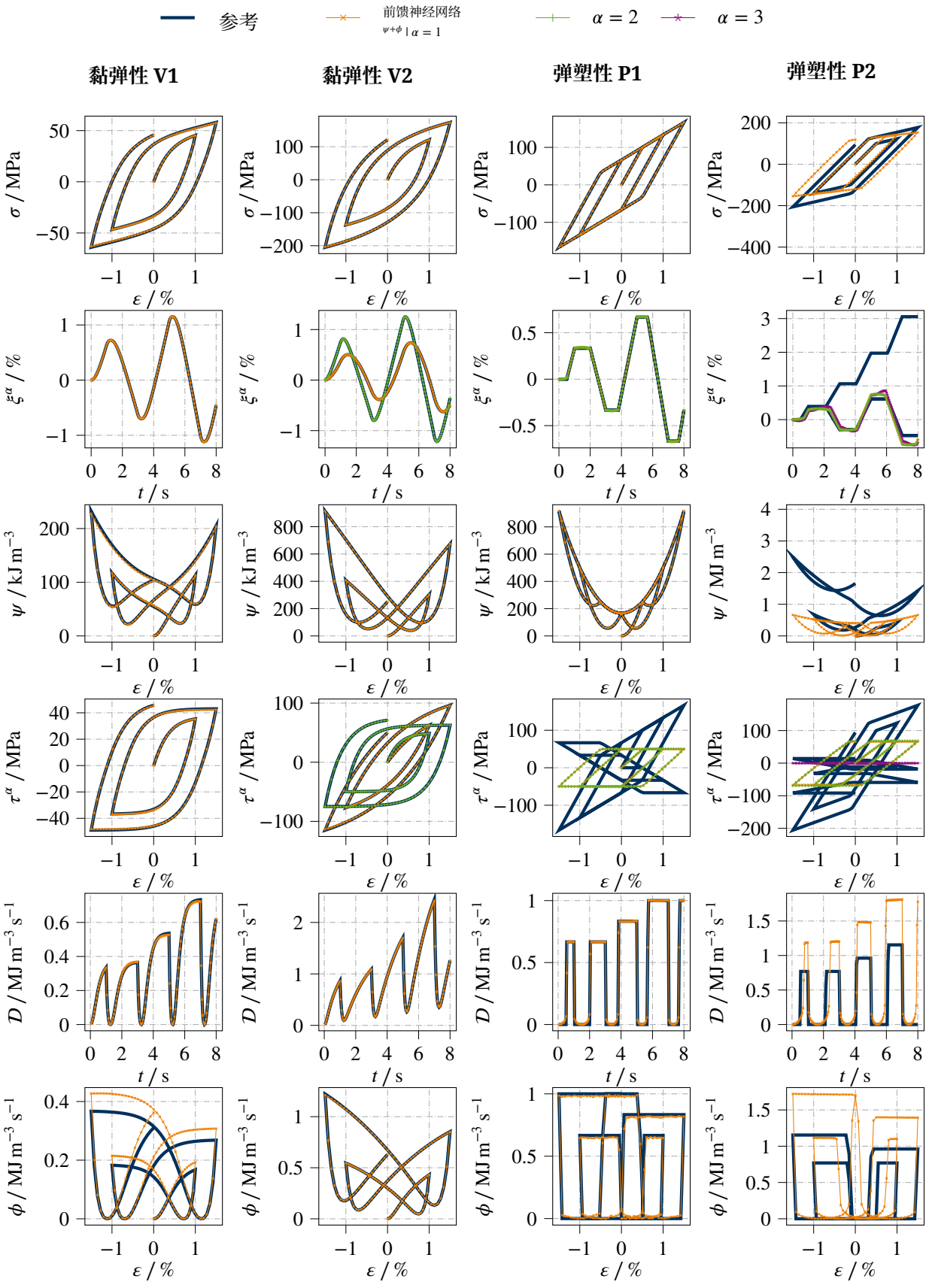


图17：使用前馈神经网络$\psi+\phi$对四种测试材料的应力预测、内变量、自由能、内力、耗散率和耗散势进行预测。

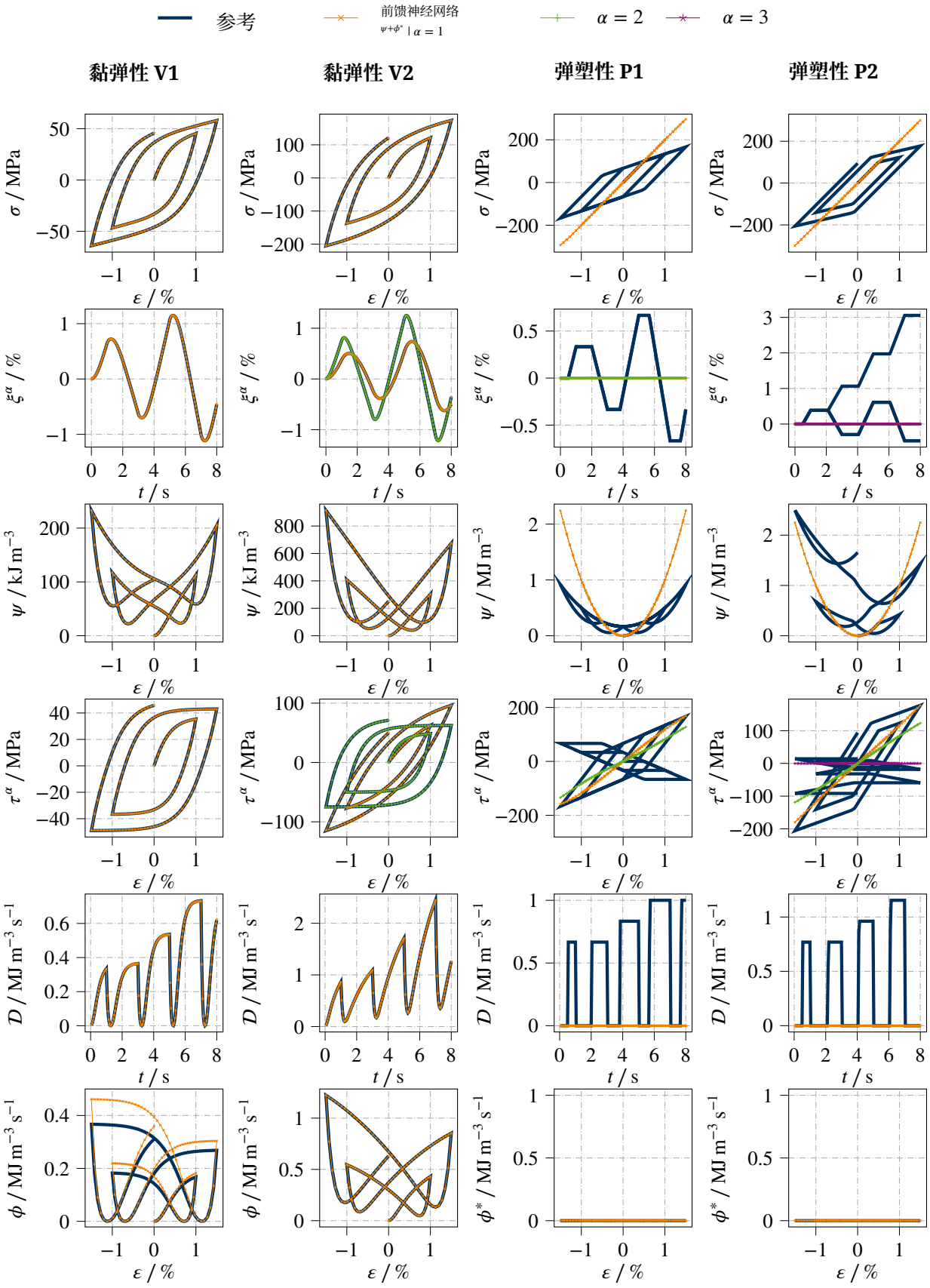


图18：使用前馈神经网络$\psi+\phi^*$对四种测试材料的应力、内变量、自由能、内力、耗散率和对偶耗散势进行预测。

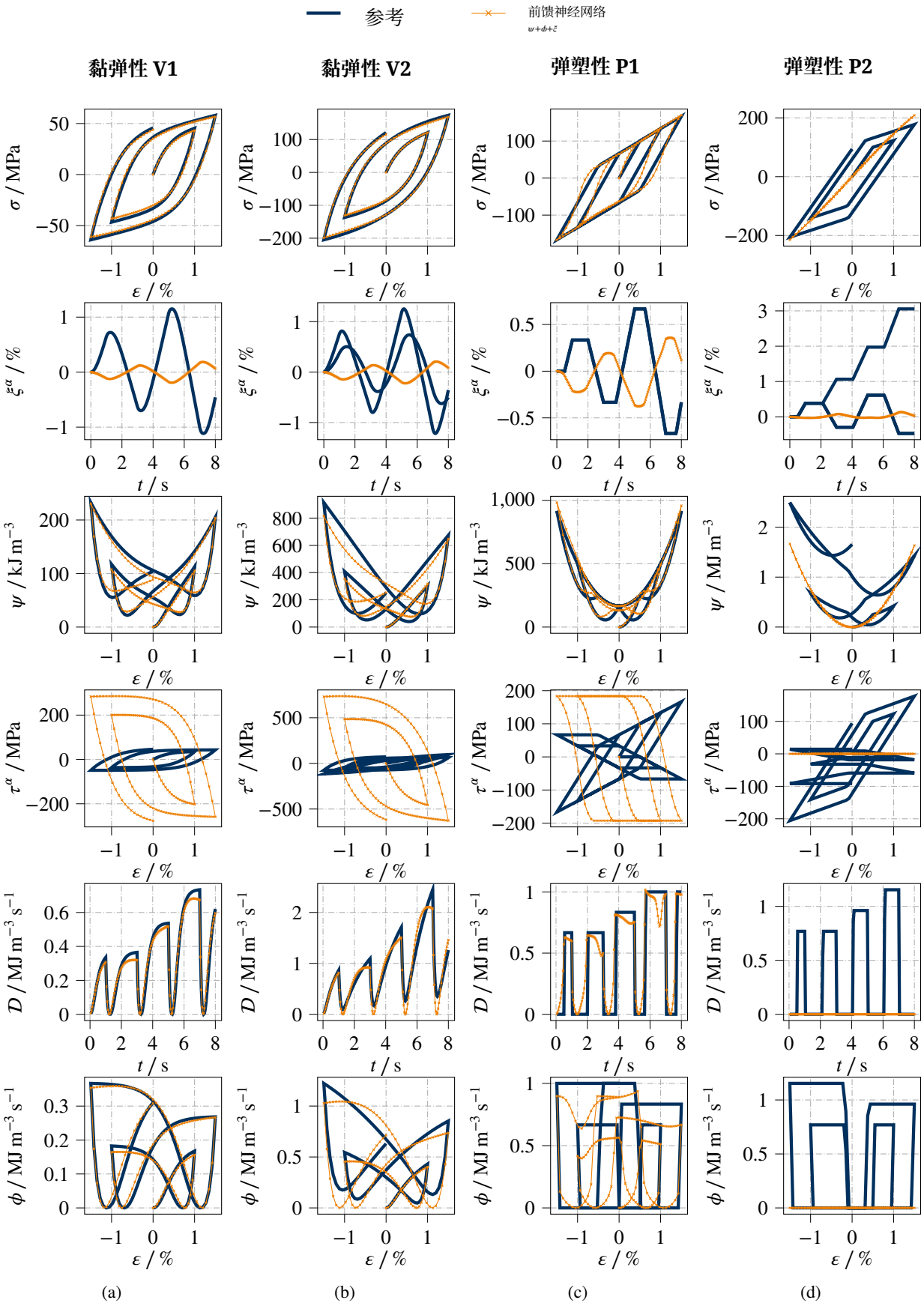


图 19：预测 of 应力、内变量、自由能、内力、耗散率和耗散势的国产自拍四种测试材料的 a使用高级训练方法的前馈神经网络$\psi+\phi+\xi$。25

6.2 带外推的验证路径

在研究了哪些基于神经网络的方法能够良好地再现哪些材料行为（即 V1、V2、P1 和 P2）之后，现在对这些方法的外推能力进行比较。比较采用图12b所示的路径。由于所有基于神经网络的方法都表明能够良好地再现黏弹性材料 V1，以下仅考虑这一材料。现在使用与无外推应变路径验证时完全相同的模型开展外推研究。这项最终研究的结果见图20。

6.2.1 黑箱神经网络

首先，评估第一个外推黑箱模型前馈神经网络$^{\sigma}$的能力。可以看出，该模型能够轻松外推至应变 $\varepsilon = 3\,\%$，超出了训练范围 $\varepsilon^{\text{train}} \in [-2, 2]\,\%$。模型重新进入训练范围后，虽未出现明显误差，但无法外推至应变率更高的范围，并产生较大误差。不过，经过短暂松弛后，该模型能够准确预测训练范围之外的应变增量，甚至能够对高达 4 %的应力给出合理数值。因此，除应变率增加 $\dot{\varepsilon}$的情况外，前馈神经网络$^{\sigma}$ 模型的外推能力出乎意料地好。不过需要指出的是，无法对该模型的热力学性质作出判断。

第二种黑箱模型——循环神经网络$^{\sigma}$——起初同样表现出良好的预测质量，但在应变率较大时也无法预测应力，并且在时间增量为 $\Delta t = 0.2\,\text{s}$时出现偏差。需要注意的是，循环神经网络的外推行为往往会因训练配置不同而有显著差异。在某些情况下，甚至可以观察到强烈的振荡。因此，与前馈神经网络$^{\sigma}$ 模型相比，循环神经网络$^{\sigma}$ 的外推能力较差。不过，与前馈神经网络$^{\sigma}$相比，它可用于更广泛的材料行为类别，见图13和图14。它仍然无法对热力学做出判断。

6.2.2 在弱形式中强制满足物理学的神经网络

现在考察外推行为在弱形式中强制满足物理学的神经网络。如图20所示，模型前馈神经网络$^{\xi+\psi}$ 在最大应变达到 2.5 %之前与参考值V1吻合得非常好，但超过这一数值后会略有偏差。重新进入训练范围后，模型再次做出了准确预测。提高应变率也会导致误差，但与之前的模型相比，误差显著减小。该模型能够应对增大的时间增量，并再次准确预测应力，直至 2.5 %。自由能 ψ 和耗散率也同样如此；对于前者， $\mathcal{D} \geq 0$ 在此之前均成立。因此，在加载路径到达这一点之前，该模型不会违反热力学第二定律。然而，当应变进一步增大到 $\varepsilon > 3\,\%$时， σ 中会出现显著误差。此外，可以看到模型预测出了负值 $\mathcal{D}$。因此，在加载路径的大部分范围内，前馈神经网络$^{\xi+\psi}$ 的外推表现非常好。但当应变达到一定水平后，会出现不符合物理规律的预测。这是因为热力学第二定律仅通过损失函数中的惩罚项来强制满足，并非先验满足。

循环神经网络$^{\xi+\psi}$ 模型同样表现良好，但在应变不超过 3 %时预测更为精确。与前馈神经网络$^{\xi+\psi}$一样，当应变大于 3 %时，$\mathcal{D} < 0$ 会出现完全不符合物理规律的预测。因此，循环神经网络$^{\xi+\psi}$ 在外推至 $\varepsilon = 3\,\%$时能够给出可接受的结果。不过，在此处的训练中，无需使用内部变量。最后，与前馈神经网络$^{\xi+\psi}$的情况一样，也可能出现不符合物理规律的预测。

6.2.3 以强形式施加物理约束的神经网络

最后，分析三种以强形式施加物理约束的神经网络外推性能：前馈神经网络$^{\psi+\phi}$、前馈神经网络$^{\psi+\phi^*}$和前馈神经网络$^{\psi+\phi}$。如图20所示，使用前馈神经网络$^{\psi+\phi}$ 和$^{\psi+\phi^*}$，在 σ、 ψ以及 $\mathcal{D}$的情况下，均可获得极为准确的预测。这适用于整个加载路径，但前馈神经网络$^{\psi+\phi^*}$ 在 $\varepsilon > 4\,\%$时的最后一段除外。前馈神经网络$^{\psi+\phi}$ 甚至能够针对 $\varepsilon = 6\,\%$的应变给出非常精确的结果。使用前馈神经网络$^{\psi+\phi+\xi}$ 进行外推的准确度不及前两种架构，但考虑到其插值结果平平，这一表现仍然不错。总而言之，应变或应变率大幅提高，以及时间增量超出训练范围，都不会导致材料响应显著偏离预期。尤其值得注意的是，在任何情况下都能确保 $\mathcal{D} \geq 0$ 。因此，预测始终符合 CDI (1) 给出的第二定律。总的来说，这一模型类别最适合用于外推，这得益于其中融入的扎实物理基础。

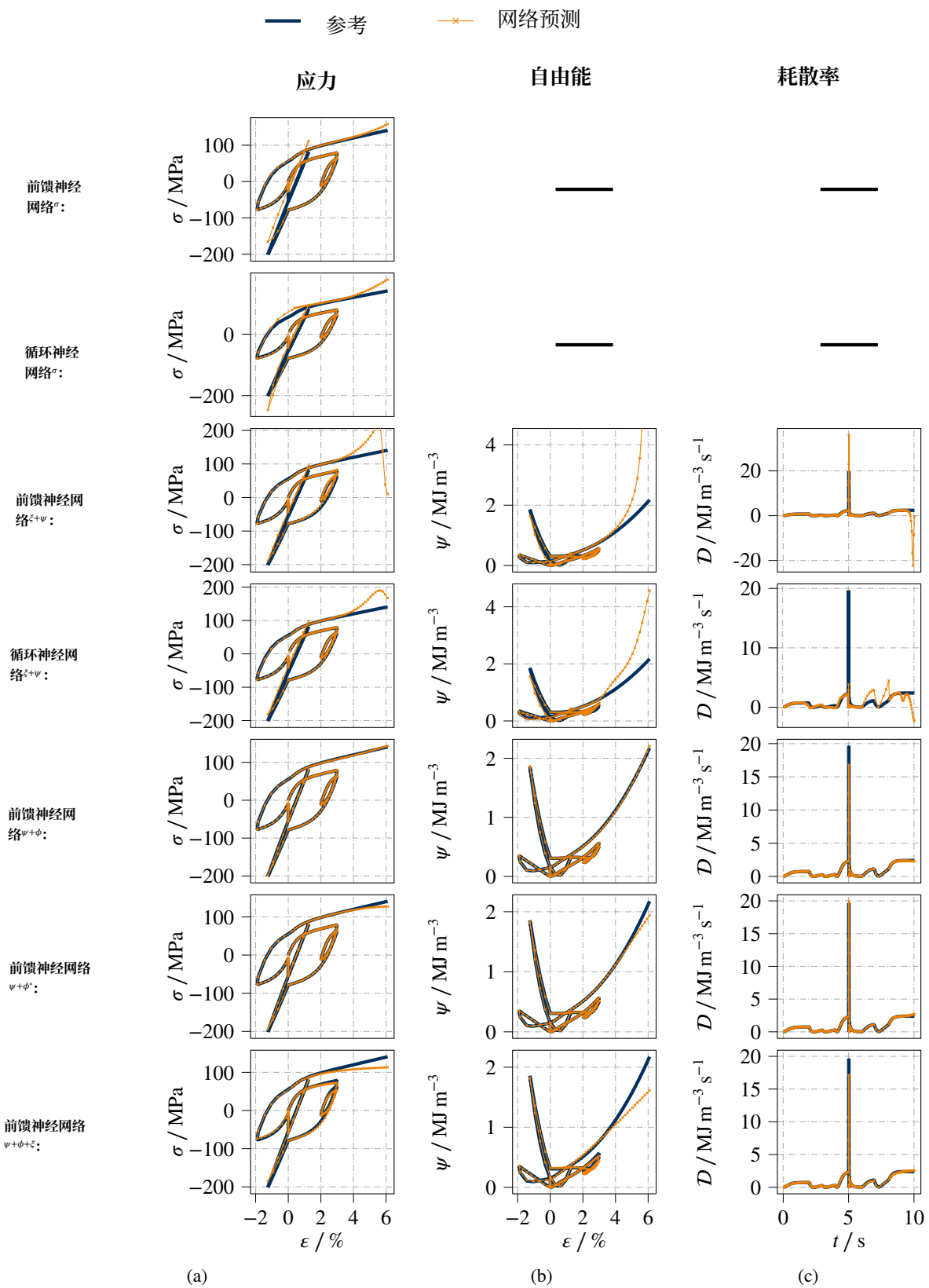


图20：对未包含在训练数据集中的外推路径 ε、 $\dot{\varepsilon}$和 Δt 进行测试的六种神经网络的预测结果：（**a**）应力 σ、（**b**）自由能 ψ和（**c**）耗散率 $\mathcal{D}$。黏弹性模型 V1 作为参考，并用于生成训练数据。

7 结论

本文对用于模拟非弹性本构行为的各种基于神经网络的方法进行了分类，特别关注热力学框架，并对这些方法进行了统一表述。为此，将基于神经网络的方法分为黑箱神经网络、以弱形式施加物理约束的神经网络和以强形式施加物理约束的神经网络，并将其应用于一维弹塑性数据和黏弹性数据。

在简要回顾相关文献后，本文概述了基于连续介质的本构建模，包括标准黏弹性模型和弹塑性模型，并简要回顾了前馈神经网络和循环神经网络的基础知识。在此基础上，介绍了七种基于神经网络的方法，并详细说明其训练和应用过程。本文阐述了各模型以何种方式考虑热力学第二定律，以及训练所需的数据。随后，展示了训练数据的生成过程，以及七种基于神经网络的方法在这些数据上的应用。结果表明，所有考察的模型都能够表示黏弹性，但并非所有方法都能表示弹塑性。此外，本文还分析了各模型的外推能力。

总之，本研究的结果表明，基于神经网络的模型在描述复杂的非弹性行为方面前景广阔，且具有很强的灵活性。它们有望逐步取代耗时的传统本构模型构建与标定工作，并实现自动化工作流。黑箱模型然而，无法据此判断相应模型所描述的过程是否嵌入了合理的热力学框架，因此不建议使用。此外，当应用于训练域之外的未知加载路径时，神经网络较差的外推能力可能会导致应力预测出现较大误差。以弱形式施加物理约束的神经网络则在模型中纳入了更多物理学内容。只有以强形式施加物理约束的神经网络 这一类模型，才能真正确保不违反热力学定律。此外，它们也展现出在所考察模型中最佳的外推行为。

关于本文研究的后续工作，未来计划开展若干扩展。例如，需要将研究扩展至有限应变下的一般三维情形。因此，还需将多种其他物理原理和条件纳入比较，例如客观性或材料对称性等原理。

CRediT 作者贡献声明

Max Rosenkranz: 概念构思、形式分析、调查研究、方法学、可视化、软件、验证、撰写初稿、审阅与编辑。**Karl A. Kalina:** 概念构思、形式分析、方法学、撰写初稿、审阅与编辑。**Jörg Brummund:** 形式分析、方法学、审阅与编辑。**Markus Kästner:** 经费获取、资源提供、审阅与编辑。

参考

[1] Frederic E. Bock、Roland C. Aydin、Christian J. Cyron、Norbert Huber、Surya R. Kalidindi 和 Benja min Klusemann。机器学习与数据挖掘方法在连续体材料力学中的应用综述。*Frontiers in Materials*，6:110，2019年5月。国际标准连续出版物号 2296-8016。数字对象标识符:10.3389/fmats.2019.00110。[2] Francisco J. Montáns、Francisco Chinesta、Rafael Gómez-Bombarelli 和 J. Nathan Kutz。科学与工程中的数据驱动建模与学习。*Comptes Rendus Mécanique*，347(11):845–855，2019年11月。国际标准连续出版物号 16310721。数字对象标识符:10.1016/j.crme.2019.11.009。[3] J. Ghaboussi、J. H. Garrett 和 X. Wu。利用神经网络对材料行为进行基于知识的建模。*Journalof Engineering Mechanics*，117(1):132–153，1991年1月。国际标准连续出版物号 0733-9399，1943-7889。数字对象标识符:10.1061/(ASCE)0733-9399(1991)117:1(132)。[4] M. Raissi、P. Perdikaris 和 G.E. Karniadakis。物理信息神经网络：用于求解涉及非线性偏微分方程的正问题和逆问题的深度学习框架。*Journal of ComputationalPhysics*，378:686–707，2019年2月。国际标准连续出版物号 00219991。数字对象标识符:10.1016/j.jcp.2018.10.045。[5] Alexander Henkes、Henning Wessels 和 Rolf Mahnken。用于连续体微观力学的物理信息神经网络。*Computer Methods in Applied Mechanics andEngineering*，393:114790，2022年4月。国际标准连续出版物号 0045-7825。数字对象标识符:10.1016/j.cma.2022.114790。[6]Faisal As'ad、Philip Avery 和 Charbel Farhat。数据驱动本构建模中的一种力学信息人工神经网络方法。*InternationalJournal for NumericalMethods inEngineering*，123(12):2738–2759，2022年。国际标准连续出版物号 1097-0207。数字对象标识符:10.1002/nme.6957。 _电子预印本: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.6957。

[7] Faisal As'ad 和 Charbel Farhat。用于数据驱动非线性黏弹性的力学信息神经网络框架。2022年。数字对象标识符:10.13140/RG.2.2.21694.36168。[8] Dominik K. Klein、Rogelio Ortigosa、Jesús Martínez-Frutos 和 Oliver Weeger。物理增强神经网络下的有限电弹性。*Computer Methods in Applied Mechanicsand Engineering*，400:115501，2022年10月。国际标准连续出版物号 0045-7825。数字对象标识符:10.1016/j.cma.2022.115501。[9]Lennart Linden、Dominik K. Klein、Karl A. Kalina、Jörg Brummund、Oliver Weeger 和 Markus Kästner。神经网络与超弹性相遇：施加物理约束指南。2023年2月。数字对象标识符:https://doi.org/10.48550/arXiv.2302.02403。[10] Karl A. Kalina、Lennart Linden、Jörg Brummund 和 Markus Kästner。FE^{ANN}：一种基于物理约束神经网络和自动化数据挖掘的高效数据驱动多尺度方法。*Computational Mechanics*，2023年2月。国际标准连续出版物号 1432-0924。数字对象标识符:10.1007/s00466-022-02260-0。[11] Filippo Masi、Ioannis Stefanou、Paolo Vannucci 和 Victor Maffi-Berthier。用于本构建模的基于热力学的人工神经网络。*Journal of the Mechanics and Physics of Solids*，147:104277，2021年2月。国际标准连续出版物号 0022-5096。数字对象标识符:10.1016/j.jmps.2020.104277。[12] Y. Shen、K. Chandrashekhara、W. F. Breig 和 L. R. Oliver。用于橡胶材料的基于神经网络的本构模型。*RubberChemistry and Technology*，77(2):257–277，2004年5月。国际标准连续出版物号 1943-4804，0035-9475。数字对象标识符:10.5254/1.3547822。[13] G. Liang 和 K. Chandrashekhara。弹性体泡沫的基于神经网络的本构模型。*EngineeringStructures*，30(7):2002–2011，2008年7月。国际标准连续出版物号 01410296。数字对象标识符:10.1016/j.engstruct.2007.12.021。[14] Kevin Linka、Markus Hillgärtner、Kian P. Abdolazizi、Roland C. Aydin、Mikhail Itskov 和 Christian J. Cy-ron。本构人工神经网络：一种通过深度学习进行预测性数据驱动本构建模的快速通用方法。*Journal of ComputationalPhysics*，429:110010，2021年3月。国际标准连续出版物号 00219991。数字对象标识符:10.1016/j.jcp.2020.110010。[15] Lennart Linden、Karl A. Kalina、Jörg Brummund、Philipp Metsch 和 Markus Kästner。基于人工神经网络的各向同性超弹性热力学一致本构建模。*PAMM*，21(1)，2021年12月。国际标准连续出版物号 1617-7061，1617-7061。数字对象标识符:10.1002/pamm.202100144。[16] Dominik K. Klein、Mauricio Fernández、Robert J. Martin、Patrizio Neff 和 Oliver Weeger。利用神经网络实现多凸各向异性超弹性。*Journalof the Mechanics and Physicsof Solids*，第104703页，2021年11月。国际标准连续出版物号 00225096。数字对象标识符:10.1016/j.jmps.2021.104703。[17] Jan N. Fuhg、Nikolaos Bouklas 和 Reese E. Jones。通过张量基神经网络从数据中学习超弹性各向异性。*Journal of theMechanics and Physics of Solids*，168:105022，2022年11月。国际标准连续出版物号 00225096。数字对象标识符:10.1016/j.jmps.2022.105022。[18] Vahidullah Tac、Francisco Sahli Costabal 和 Adrian B. Tepole。采用多凸神经常微分方程的数据驱动组织力学。*Computer Methods inApplied Mechanics and Engineering*，398:115248，2022年8月。国际标准连续出版物号 0045-7825。数字对象标识符:10.1016/j.cma.2022.115248。[19] Nikolaos N. Vlassis、Ran Ma 和 WaiChing Sun。计算力学中的几何深度学习 第一部分：各向异性超弹性。*Computer Methods in Applied Mechanics andEngineering*，371:113299，2020年11月。国际标准连续出版物号 0045-7825。数字对象标识符:10.1016/j.cma.2020.113299。[20] Nikolaos N. Vlassis、Puhan Zhao、Ran Ma、Tommy Sewell 和 WaiChing Sun。基于分子动力学推断的迁移学习模型，用于单斜晶体的有限应变超弹性：索博列夫训练及物理约束验证。*International Journal for Numerical Methods in Engineering*，123(17):3922–3949，2022年。国际标准连续出版物号 1097-0207。数字对象标识符:10.1002/nme.6992。 _电子预印本: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.6992。[21] Karl A. Kalina、Lennart Linden、Jörg Brummund、Philipp Metsch 和 Markus Kästner。基于人工神经网络的各向同性超弹性自动化本构建模。*Computational Mechanics*，2021年10月。国际标准连续出版物号 0178-7675，1432-0924。数字对象标识符:10.1007/s00466-021-02090-6。[22]Tomonari Furukawa 和 Genki Yagawa。利用神经网络进行黏塑性的隐式本构建模。*International Journal for NumericalMethodsin Engineering*，43(2): 195–219，1998年。国际标准连续出版物号 1097-0207。数字对象标识符 :10.1002/(SICI)1097-0207(19980930)43:2<195::AID-NME418>3.0.CO;2-6。 _ 电子预印本: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291097-0207%2819980930%2943%3A2%3C195%3A%3AAID-NME418%3E3.0.CO%3B2-6。[23] J. Ghaboussi 和 D.E. Sidarta。用于本构建模的新型嵌套自适应神经网络（NANN）。《计算机与岩土工程》

，22:(1)29–52 ，一月 y 1998年。国际标准连续出版物号 0266352X。数字对象标识符: 10.1016/S0266-352X (97)00034-7.

[24] Y. M. A. Hashash、S. Jung 和 J. Ghaboussi。基于神经网络的材料模型在有限元分析中的数值实现：基于神经网络的材料模型。*InternationalJournalfor Numerical Methods in Engineering*，59(7):989–1005，2004年2月。国际标准连续出版物号 00295981。数字对象标识符:10.1002/nme.905.[25] M.S. Al-Haik、M.Y. Hussaini 和 H. Garmestani。利用人工神经网络预测聚合物基复合材料的非线性黏弹行为。*International Journal ofPlasticity*，22(7):1367–1392，2006年7月。国际标准连续出版物号 07496419。数字对象标识符:10.1016/j.ijplas.2005.09.002.[26] Sungmoon Jung 和 Jamshid Ghaboussi。率相关材料的神经网络本构模型。*Com-puters&Structures*，84(15-16):955–963，2006年6月。国际标准连续出版物号 00457949。数字对象标识符:10.1016/j.compstruc.2006.02.015.[27] Sepp Hochreiter 和 Jürgen Schmidhuber。长短期记忆。*Neural Computation*，9(8):1735–1780，1997年11月。国际标准连续出版物号 0899-7667。数字对象标识符:10.1162/neco.1997.9.8.1735。会议名称：Neural Computation。[28] F. Ghavamian 和 A. Simone。利用循环神经网络加速历史相关材料的多尺度有限元模拟。*Computer Methods in Applied Mechanicsand Engineering*，357:112594，2019年12月。国际标准连续出版物号 00457825。数字对象标识符:10.1016/j.cma.2019.112594.[29] Ling Wu、Van Dung Nguyen、Nanda Gopala Kilingar 和 Ludovic Noels。一种由循环神经网络加速的多尺度模型，用于随机循环和非比例加载路径下的弹塑性非均质材料。*Computer Methods in AppliedMechanics andEngineering*，369:113234，2020年9月。国际标准连续出版物号 00457825。数字对象标识符:10.1016/j.cma.2020.113234.[30] Bin Li 和 Xiaoying Zhuang。基于前馈神经网络和循环神经网络的多尺度计算。*Frontiers ofStructural and CivilEngineering*，14(6):1285–1298，2020年12月。国际标准连续出版物号 2095-2430, 2095-2449。数字对象标识符:10.1007/s11709-020-0691-7.[31] Alexander Fuchs、Yousef Heider、Kun Wang、WaiChing Sun 和 Michael Kaliske。DNN2：用于自设计基于神经网络的弹塑性本构描述的超参数强化学习博弈。*Computers & Structures*，249:106505，2021年6月。国际标准连续出版物号 00457949。数字对象标识符:10.1016/j.compstruc.2021.106505.[32] Alexander Henkes、Jason K. Eshraghian 和 Henning Wessels。用于非线性回归的脉冲神经网络。2022年10月。数字对象标识符:10.48550/arXiv.2210.03515.[33] Colin Bonatti 和 Dirk Mohr。自洽性对于表征弹塑性固体的循环神经网络模型的重要性。*Journal of the Mechanics and Physicsof Solids*，158:104697，2022年1月。国际标准连续出版物号 00225096。数字对象标识符:10.1016/j.jmps.2021.104697.[34] Colin Bonatti、Bekim Berisha 和 Dirk Mohr。从 CP-FFT 到 CP-RNN：晶体塑性的循环神经网络代理模型。*International Journal of Plasticity*，158:103430，2022年11月。国际标准连续出版物号 0749-6419。数字对象标识符:10.1016/j.ijplas.2022.103430.[35] Yousef Heider、Kun Wang 和 WaiChing Sun。基于信息图的深度神经网络对各向异性弹塑性材料的 SO(3)不变性。*Computer Methods inApplied Mechanicsand Engineering*，363:112875，2020年5月。国际标准连续出版物号 00457825。数字对象标识符:10.1016/j.cma.2020.112875.[36] I. B. C. M. Rocha、P. Kerfriden 和 F. P. van der Meer。面向多尺度固体力学的演化的基于物理的材料模型的机器学习。2023年1月。数字对象标识符:10.48550/arXiv.2301.13547.[37] Zeliang Liu、C.T. Wu 和 M. Koishi。用于非均质材料多尺度拓扑学习和加速非线性建模的深度材料网络。*Computer Methodsin Applied Mechanics and Engineering*，345: 1138–1168，2019年3月。国际标准连续出版物号 00457825。数字对象标识符:10.1016/j.cma.2018.09.020.[38] Zeliang Liu 和 C.T. Wu。探索数据驱动多尺度力学中深度材料网络的三维架构。*Journalof theMechanics and Physicsof Solids*，127:20–46，2019年6月。国际标准连续出版物号 00225096。数字对象标识符:10.1016/j.jmps.2019.03.004.[39] Sebastian Gajek、Matti Schneider 和 Thomas Böhlke。深度材料网络的细观力学。*Journal of the Mechanics and Physics of Solids*，142:103984，2020年9月。国际标准连续出版物号 00225096。数字对象标识符:10.1016/j.jmps.2020.103984.[40] Sebastian Gajek、Matti Schneider 和 Thomas Böhlke。一种用于热力学复合材料多尺度分析的 FE-DMN方法。*Computational Mechanics*，2022年2月。国际标准连续出版物号 1432-0924。数字对象标识符:10.1007/s00466- 021-02131-0.[41] Christoph Settgast、Geralf Hütter、Meinhard Kuna 和 Martin Abendroth。一种利用神经网络模拟泡沫均质化不可逆弹塑性变形和损伤的混合方法。*InternationalJournalofPlasticity*，126:102624，2020年3月。国际标准连续出版物号 07496419。数字对象标识符:10.1016/j.ijplas.2019.11.003。

[42] Alexander Malik、Martin Abendroth、Geralf Hütter 和 Bjoern Kiefer。采用神经网络模拟三维泡沫结构弹塑性变形行为的混合方法。*Advanced Engineering Materials*，n/a(n/a):2100641，2021年。国际标准连续出版物号 1527-2648。数字对象标识符:10.1002/adem.202100641。_电子版:https://onlinelibrary.wiley.com/doi/pdf/10.1002/adem.202100641。[43] Nikolaos N. Vlassis 和 WaiChing Sun。为具有水平集硬化的可解释弹塑性模型训练热力学信息神经网络的索博列夫训练。*Computer Methods in Applied Mechanics andEngineering*，377:113695，2021年4月。国际标准连续出版物号 00457825。数字对象标识符:10.1016/j.cma.2021.113695。[44] Nikolaos Napoleon Vlassis 和 Waiching Sun。用于发现率相关、压力敏感水平集塑性模型的基于组件的机器学习范式。*Journal of Applied Mechanics*，第1–13页，2021年10月。国际标准连续出版物号 0021-8936, 1528-9036。数字对象标识符:10.1115/1.4052684。[45] Jan N. Fuhg、Craig M. Hamel、Kyle Johnson、Reese Jones 和 Nikolaos Bouklas。模块化机器学习弹塑性：有限数据情境下的泛化。2022年10月。数字对象标识符:10.48550/arXiv.2210.08343。[46] C. Zopf 和 M. Kaliske。采用基于神经网络的方法对未硫化弹性体进行数值表征。*Computers&Structures*，182:504–525，2017年4月。国际标准连续出版物号 00457949。数字对象标识符:10.1016/j.compstruc.2016.12.012。[47] Filippo Masi 和 Ioannis Stefanou。利用基于热力学的人工神经网络（TANN）对非弹性材料进行多尺度建模。*Computer Methods in Applied Mechanics and Engineering*，398:115190，2022年8月。国际标准连续出版物号 0045-7825。数字对象标识符:10.1016/j.cma.2022.115190。[48] Xiaolong He 和 Jiun-Shyan Chen。用于路径相关材料数据驱动建模的热力学一致机器学习内状态变量方法。*ComputerMethods inApplied Mechanics andEngineering*，第115348页，2022年7月。国际标准连续出版物号 00457825。数字对象标识符:10.1016/j.cma.2022.115348。[49] Shenglin Huang、Zequn He、Bryan Chem 和 Celia Reina。变分昂萨格神经网络（VONNs）：用于非平衡偏微分方程的基于热力学的变分学习策略。*Journal of the Mechanics andPhysics of Solids*，163:104856，2022年6月。国际标准连续出版物号 0022-5096。数字对象标识符:10.1016/j.jmps.2022.104856。[50] Christian Miehe。基于增量变分表述的非弹性微观结构与复合材料应变驱动均匀化。*International Journalfor Numerical Methods in Engineering*，55(11):1285–1322，2002年12月。国际标准连续出版物号 0029-5981, 1097-0207。数字对象标识符:10.1002/nme.515。[51] Brandon Amos、Lei Xu 和 J. Zico Kolter。输入凸神经网络。载于第*34*届国际机器学习会议论文集，第146–155页。PMLR，2017年7月。[52] Vahidullah Tac、Manuel K. Rausch、Francisco Sahli Costabal 和 Adrian Buganza Tepole。利用神经常微分方程进行数据驱动的各向异性有限黏弹性建模。*SSRN*电子期刊，2023年。国际标准连续出版物号 1556-5068。数字对象标识符:10.2139/ssrn.4332501。[53] Peter Haupt。连续介质力学与材料理论。Springer Berlin Heidelberg，德国柏林、海德堡，2000年。国际标准书号 978-3-662-04109-3。OCLC：851363702。[54] C Miehe、J Schotte 和 M Lambrecht。基于增量极小化原理的有限应变下非弹性固体材料均匀化：在多晶体织构分析中的应用。*J. Mech. Phys.Solids*，第45页，2002年。[55] Christian Miehe、Björn Kiefer 和 Daniele Rosato。宏观连续体层级耗散磁致伸缩的增量变分表述。《国际固体与结构杂志》，48(13):1846–1866，2011年6月。国际标准连续出版物号 00207683。数字对象标识符:10.1016/j.ijsolstr.2011.02.011。[56] Juan Carlos Simo 和 Thomas J. R. Hughes。计算非弹性力学。跨学科应用数学：力学与材料丛书第7卷。Springer，纽约，纽约州，2000年更正后的第2版。国际标准书号 978-0-387-97520-7 978-1-4757-7169-5。OCLC：254534973。[57] Rudolf Kruse、Christian Borgelt、Christian Braune、Sanaz Mostaghim 和 Matthias Steinbrecher。计算智能。计算机科学丛书。Springer London，伦敦，2016年。国际标准书号 978-1-4471-7294-9 978-1-4471- 7296-3。数字对象标识符:10.1007/978-1-4471-7296-3。[58] Stefan Kollmannsberger、Davide D'Angella、Moritz Jokeit 和 Leon Herrmann。计算力学中的深度学习：入门课程，计算智能研究第977卷。Springer International Publishing，Cham，2021年。国际标准书号 978-3-030-76586-6 978-3-030-76587-3。数字对象标识符:10.1007/978-3-030-76587-3。[59] Michaela Nagler、Astrid Pechstein 和 Alexander Humer。弹塑性的混合有限元表述。*International Journal for Numerical MethodsinEngineering*，123(21):5346–5368，2022年。国际标准连续出版物号 1097-0207。数字对象标识符:10.1002/nme.7070。_电子版:https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.7070。[60]Pierre Ladevèze、David Néron 和 Paul-William Gerbaud。历史相关材料的数据驱动计算。*Comptes Rendus Mécanique*，347(11):831–844，2019年11月。国际标准连续出版物号 16310721。数字对象标识符:10.1016/j.crme.2019.11.008。

[61] Paul-William Gerbaud、David Néron 和 Pierre Ladevèze。涉及隐状态变量的数据驱动弹（黏）塑性。*Computer Methodsin Applied Mechanics andEngineering*，402:115394，2022年12月。国际标准连续出版物号 00457825。数字对象标识符:10.1016/j.cma.2022.115394。