

华南理工大学硕士学位论文

LaTeX 模板使用说明

redfu

指导教师：周嘉嘉 教授

华南理工大学

2025 年 4 月 4 日

摘要

流变学是研究物质变形和流动的科学，流变学本构方程作为其核心内容之一，是描述材料在外力作用下变形和流动行为的数学模型。传统本构方程的获取和研究方法主要依赖实验测定、理论推导和数值模拟，这些方法存在成本高、耗时长、难以泛化等问题，难以满足现代材料科学对高效、精确建模的需求。

近年来，基于数据驱动的机器学习方法开始应用于流变学本构方程的构建研究中，为解决传统方法的局限性提供了新思路。然而，当前的机器学习本构建模方法仍存在物理约束不足、未充分考虑流变学本构方程的历史依赖性、特征稀疏以及难以反推制备参数等问题。

针对上述问题，本文从时间域流变数据建模和频域数据建模两个方向开展研究。在时间域应力应变数据方面，本文创新性地采用物理信息门控循环单元（PI-GRU）进行时变本构数据的预测建模，利用 GRU 的门控机制有效捕捉流变学中的记忆效应和应变历史依赖性，同时引入物理残差作为损失函数的一部分，引入物理约束。相比传统深度前馈神经网络，PI-GRU 在经典本构方程的数值模拟数据和真实流变学实验数据上均表现出更好的预测效果和泛化能力。

在频域数据方面，本文采用物理信息神经网络（PINN）对材料的制备参数与储存模量、损耗模量、损耗角正切等频域数据关系进行正向建模，通过在损失函数中引入经典本构方程残差，为模型提供物理约束。同时，针对实验中多个制备参数作为特征时出现的稀疏性和分散性问题，本文使用注意力特征融合方法进行特征降维，进一步优化了模型效果。此外，本文使用条件变分自编码器（CVAE）对实验数据进行反向建模，实现了从特定流变学性质数据生成实验制备参数的逆向设计，与 PINN 形成正逆向联合建模框架，为高分子材料的设计提供了有效的辅助支持工具。

关键词：流变学；本构方程；门控循环单元；物理信息神经网络；条件变分自编码器

Abstract

Rheology is the science that studies the deformation and flow of matter, and rheological constitutive equations, as one of its core components, are mathematical models that describe the deformation and flow behaviors of materials under external forces. Traditional methods for obtaining and studying constitutive equations primarily rely on experimental measurements, theoretical derivation, and numerical simulation; however, these methods are often costly, time-consuming, and lack generalizability, which makes it challenging to meet the demands of modern materials science for efficient and accurate modeling.

In recent years, data-driven machine learning approaches have been applied to the construction of rheological constitutive equations, offering new avenues to overcome the limitations of traditional methods. Nonetheless, current machine learning-based constitutive modeling methods still suffer from insufficient physical constraints, inadequate consideration of the historical dependency inherent in rheological constitutive equations, feature sparsity, and challenges in reverse engineering the processing parameters.

To address these issues, this study investigates two modeling directions: time-domain rheological data modeling and frequency-domain data modeling. For time-domain stress – strain data, an innovative approach employing a physics-informed gated recurrent unit (PI-GRU) is proposed for predicting time-varying constitutive data. The gated mechanism of the GRU effectively captures the memory effects and strain-history dependency in rheology, while the inclusion of a physical residual term in the loss function imposes physical constraints. Compared with traditional deep feedforward neural networks, the PI-GRU demonstrates superior predictive performance and generalization ability on both numerical simulation data derived from classical constitutive equations and experimental rheological data.

For frequency-domain data, a physics-informed neural network (PINN) is employed to model the relationship between the material processing parameters and frequency-domain properties, such as storage modulus, loss modulus, and loss tangent. By incorporating the residual of the classical constitutive equation into the loss function, physical constraints are imposed on the model. Furthermore, to address the sparsity and dispersion issues encountered when multiple processing parameters are used as features in experiments, an attention-based feature fusion

method is applied for dimensionality reduction, which further optimizes the model performance. In addition, a conditional variational autoencoder (CVAE) is used for inverse modeling of the experimental data, enabling the reverse design of processing parameters from specific rheological property data. Together with the PINN, this forms a forward–inverse joint modeling framework that serves as an effective tool for the design of polymer materials.

Keywords: Rheology; Constitutive Equations; Gated Recurrent Unit; Physics-Informed Neural Networks; Conditional Variational Autoencoder

目 录

摘 要	I
Abstract	II
第一章 绪论	1
1.1 流变学介绍	1
1.1.1 流变学的核心研究内容	1
1.1.2 流变学研究及应用方向	2
1.2 本构方程	2
1.2.1 线性本构方程	2
1.2.2 非线性本构方程	5
1.2.3 传统的本构方程的构建方法	7
1.3 数据驱动方法流变学本构建模	9
1.3.1 机器学习方法介绍	9
1.3.2 机器学习流变学应用研究现状	10
1.4 引入物理约束的神经网络研究	13
1.4.1 引言	13
1.4.2 物理信息神经网络研究现状	14
1.4.3 其他物理信息深度学习模型	17
1.5 本课题研究介绍	17
1.5.1 研究内容	18
1.5.2 研究创新点	19
1.5.3 研究意义	19
第二章 算法介绍	21
2.1 引言	21
2.2 算法介绍	21
2.2.1 时间序列数据介绍	21
2.2.2 循环神经网络	21
2.2.3 其他时间序列算法	24

2.3 物理信息神经网络	25
2.3.1 理论基础	25
2.3.2 损失函数构建	26
2.3.3 未来技术优化	27
2.4 特征融合方法	28
2.4.1 简单特征融合	28
2.4.2 注意力特征融合	28
2.5 生成式模型	29
2.5.1 变分自编码器	29
2.5.2 条件变分自编码器	30
2.5.3 其他生成式模型	32
2.6 本章小结	33
第三章 物理信息门控循环单元流变学本构方程建模研究	34
3.1 引言	34
3.2 深度学习模型设计	35
3.3 实验设计	36
3.3.1 数据获取	36
3.3.2 Giesekus 模型模拟数据	38
3.3.3 模型训练	39
3.3.4 模型测试	39
3.4 结果与讨论	40
3.4.1 Herschel-Bulkley 模型建模	40
3.4.2 Maxwell 模型建模	41
3.4.3 Doi-Edwards 模型建模	45
3.4.4 Giesekus 模型建模	50
3.4.5 真实实验数据建模	56
3.5 本章小结	63
第四章 物理信息-生成式正逆向流变学本构建模研究	65
4.1 引言	65

4.2 实验设计	66
4.2.1 PINN-CVAE 正逆向建模架构	66
4.2.2 PINN 数据预处理	66
4.2.3 PINN 模型训练	68
4.2.4 PINN 模型测试	70
4.2.5 CVAE 反向建模训练	71
4.2.6 CVAE 反向建模测试	71
4.2.7 正逆向联合建模	71
4.3 结果与讨论	72
4.3.1 低保真数据拟合	72
4.3.2 单 PBA 流体本构建模	74
4.3.3 单组分 PBA 注入的 PFGs 本构建模	77
4.3.4 多组分 PBA 注入的 PFGs 本构建模	79
4.3.5 CVAE 组分预测的多模态解分析	81
4.3.6 正逆向联合建模分析	85
4.4 本章小结	87
总结与展望	89
参考文献	92
攻读博士/硕士学位期间取得的研究成果	104
致 谢	105

第一章 绪论

1.1 流变学介绍

1.1.1 流变学的核心研究内容

流变学是研究物质在外力作用下变形和流动的科学，其研究对象涵盖了流体、软固体以及在特定条件下可以流动的固体^[1]。流变学的核心在于揭示材料的应力、应变和时间之间的内在关系，并通过本构方程对这些关系进行定量描述。流变学的研究不仅深化了对材料力学行为的理解，还为工程应用和科学研究提供了重要的理论基础。流变学的核心研究内容主要包括以下几个方面^[2-3]：

- (1) 材料的流动与变形行为：材料的流动与变形行为是流变学研究的核心内容之一。通过实验和理论模型，流变学揭示了材料在外力作用下的复杂力学行为。例如，蠕变现象（即在恒定应力下，材料的变形随时间逐渐增加）和应力松弛现象（即在恒定应变下，材料的应力随时间逐渐减小）是流变学中重要的研究对象^[4-5]。这些现象不仅反映了材料的时间依赖性行为，还为材料的长期性能评估提供了理论依据。此外，流变学还研究了材料的非线性力学行为，如屈服、塑性变形和断裂等，这些研究对于理解材料的宏观力学性能具有重要意义^[6-7]。
- (2) 本构方程的构建：本构方程是流变学中用于描述材料力学行为的数学工具，其核心在于建立应力、应变和时间之间的定量关系。对于牛顿流体，其本构方程基于牛顿黏性定律，即应力与应变率成正比^[2]。然而，对于非牛顿流体和软固体，其本构方程则更为复杂，通常需要考虑材料的非线性、黏弹性以及时间依赖性等特性^[8]。通过构建合理的本构方程，流变学能够对各种物理现象进行精确的数学描述，从而为工程设计和材料开发提供理论支持。
- (3) 实验与模拟方法：流变学实验是研究材料流变性能的重要手段，常见的实验方法包括蠕变实验、应力松弛实验和动力试验等^[3]。这些实验能够直接测量材料在不同条件下的力学响应，为理论模型的验证和优化提供实验数据。近年来，随着计算模拟技术的发展，流变学研究逐渐从唯象模型向定量科学转变。微观实验技术（如 X 射线散射、中子散射）与计算模拟的结合，使得研究者能够在微观尺度上揭示材料的流变机制，从而推动流变学向更高精度和更深层次发展^[9-10]。

1.1.2 流变学研究及应用方向

流变学的研究方向广泛，涵盖了多个学科和领域，例如高分子流变学研究高分子材料的分子结构与其流变行为的关系，包括聚合物熔体和溶液的拉伸流变行为^[11]。生物流变学研究生物材料（如血液、肌肉）的流变特性，揭示生理和病理过程中的力学机制^[12-13]。地质流变学研究岩石、土壤等地质材料的流变行为，应用于地震预测、矿产资源开发等领域^[14]。在材料加工^[5]、食品工业^[15]、化妆品和医药制造等领域^[16-17]，流变学用于优化工艺和产品性能^[18]。非牛顿流体力学研究不符合牛顿黏性定律的流体（如油漆、泥浆、血液）的流动特性^[8,19-20]。

1.2 本构方程

1.2.1 线性本构方程

凝聚相物质分为固体或液体，固体和液体之间的一个区别特征是它们对施加的力的响应。固体在变形时储存能量，如果变形很小，则在消除力后会恢复到原来的形状。相比之下，液体则会通过耗散能量和调整其形状来抵抗外力^[21-22]。这种区别可以通过两种经典的力学模型来描述：胡克固体（Hookean Solid）和牛顿流体（Newtonian Fluid）。胡克定律（Hooke's Law）可以来描述小变形下的弹性固体行为。胡克定律表明，固体的应力 σ 与应变 γ 成正比，如公式(1-1)所示。其中， G 为弹性模量，用于描述材料在弹性变形范围内抵抗外力的能力。它反映了材料的刚度，即材料在受力时发生变形的难易程度。弹性模量越大，材料越难变形；弹性模量越小，材料越容易变形。胡克固体是理想化的弹性固体模型，适用于描述金属、陶瓷等材料在小变形条件下的力学行为^[8]。

$$\sigma = G\gamma \quad (1-1)$$

而对于液体，其对外力的响应则完全不同。液体无法储存能量以恢复形状，而是通过内部的粘性阻力来耗散能量，并持续流动以适应外力。这种行为可以用牛顿流体的本构方程来描述，即剪切应力与剪切速率成正比，如公式(1-2)所示。其中， η 为粘性系数，用于描述液体在运动过程中耗散能量的能力。粘性系数越大，液体越容易耗散能量，反之亦然。牛顿流体是理想化的粘性流体模型，适用于描述水、空气等液体在运动过程中耗散能量的行为^[21]。

$$\sigma = \eta \dot{\gamma} \quad (1-2)$$

然而，这些特征都是理想化的，代表了特定条件的行为。许多凝聚相材料不容易归入这些经典类别，因为它们的机械性能取决于变形的大小、速率、变形历史，加载过程等等。例如，考虑牙膏，它像液体一样流动，可以将其从管中挤出，但一旦放在牙刷上，它就会像固体一样保持其形状。这类物质同时具有黏性和弹性，被认为是黏弹性材料，被称为软物质或者复杂流体^[23]。

线性黏弹性理论认为在小变形范围（线性范围内）应力-应变关系是线性的，即应力与应变成正比。同时线性黏弹性区间内材料具有时间依赖性，材料的力学响应不仅取决于当前的应力或应变，还依赖于时间或加载历史。图1-1概述了两种不同类型的粘弹

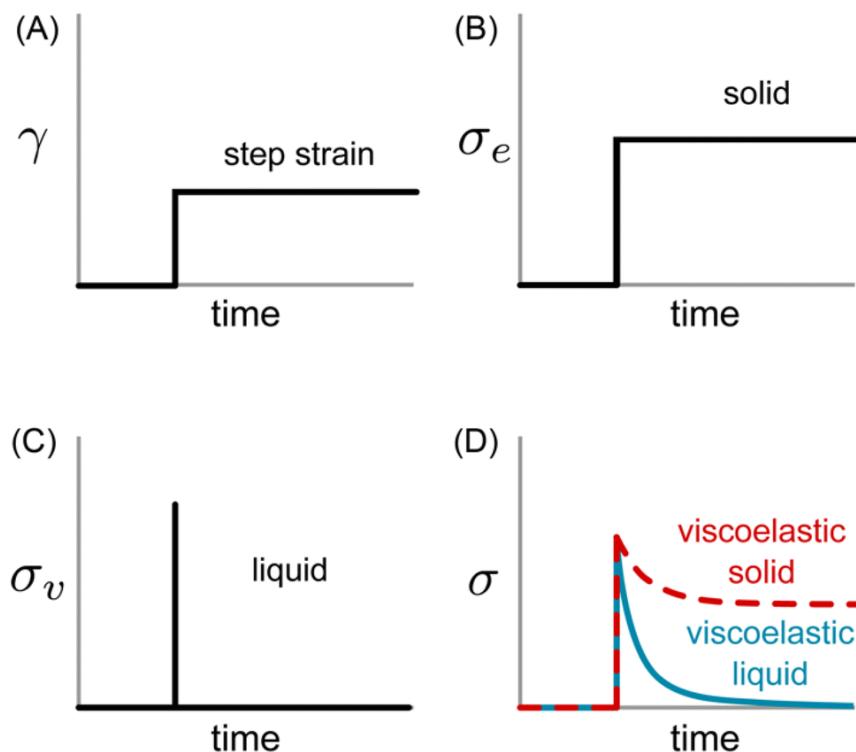


图 1-1 (A) 施加的应变曲线以及对应的剪切应力 (B) 理想弹性固体 (C) 理想粘性液体 (D) 粘弹性样品^[21]

Figure 1-1 (A) Applied strain profile and resulting shear stress (B) Ideal elastic solid (C) Ideal viscous liquid (D) Viscoelastic samples^[21]

性材料的简单剪切行为。对于粘弹性固体和液体，阶跃应变会引起瞬时弹性响应，从而产生 σ 峰值。然而，应力不是保持不变或立即降至零，而是逐渐降低。它在很长一段时

间内接近粘弹性固体的有限平台值，而粘弹性液体则完全衰减到零^[21]。

线性本构理论中最经典的是 Maxwell 模型^[24]，如图1-2所示，Maxwell 模型将材料的弹性行为和粘性行为结合起来，它用一个弹簧（弹性元件）和一个粘壶（粘性元件）串联表示黏弹性关系。Maxwell 模型的微分形式如公式(1-3)所示，其中 τ 表示松弛时间，等于 η/G 。将两边积分得到 Maxwell 模型的积分形式如公式(1-4)所示。

$$\frac{d\sigma}{dt} + \frac{\sigma}{\tau} = G \frac{d\gamma}{dt} \quad (1-3)$$

$$\sigma(t) = \int_{-\infty}^t G e^{-\frac{t-t'}{\tau}} \frac{d\gamma(t')}{dt'} dt' \quad (1-4)$$

积分形式的方程显示了任何时刻的应力是松弛模量乘以应变速率的积分。由于被积函数中的衰减指数，模型具有衰落的记忆，因此最近的应变历史比过去的应变历史更重要。如果将多个 Maxwell 模型并联，便可以得到广义 Maxwell 模型方程，如公式(1-5)所示。

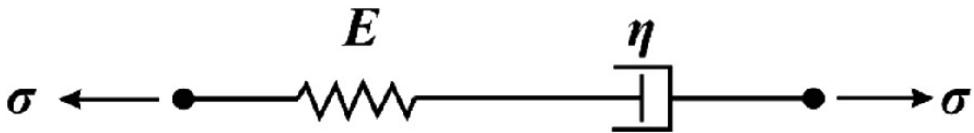


图 1-2 Maxwell 模型示意图

Figure 1-2 Schematic illustration of Maxwell model

$$\sigma(t) = \int_{-\infty}^t G(t-t') \frac{d\gamma(t')}{dt'} dt' \quad (1-5)$$

其中松弛模量 $G(t)$ 定义为公式(1-6)。

$$G(t) = \sum_{i=1}^n G_i e^{-\frac{t}{\tau_i}} \quad (1-6)$$

Maxwell 模型通过将黏弹性抽象为黏性元件和弹性元件串联来得到本构方程。如果将弹簧和粘壶进行并联，则得到 Kelvin-Voigt 模型的本构方程，即公式(1-7)^[25]。

$$\sigma(t) = G\gamma(t) + \eta \frac{d\gamma(t)}{dt} \quad (1-7)$$

将多个 Kelvin-Voigt 模型的元件进行串联，便可以得到广义 Kelvin-Voigt 模型，如公式(1-8)。

$$\sigma(t) = \sum_{i=1}^n \left(G_i \gamma_i(t) + \eta_i \frac{d\gamma_i(t)}{dt} \right) \quad (1-8)$$

原则上，任何广义 Voigt 模型都可以在数值上映射到等效的广义 Maxwell 模型，这是线性本构方程构建的基本研究方法的不同角度^[21]。

如果将传统的整数阶导数模型改为分数阶导数，能够更加准确描述材料的记忆效应。例如 Bagley 和 Torvik 等在 20 世纪 80 年代提出的分数阶 Maxwell 模型，是对传统 Maxwell 模型的推广^[26]。无论是什么形式的线性本构方程，均满足一个基本假设：材料的响应是线性的，即多个应变历史的叠加效应等于各自效应的线性相加，当公式(1-5)中的松弛模量 G 表示为公式(1-6)时是 Maxwell 模型的形式，事实上当 G 为一个抽象的松弛函数时，公式(1-5)抽象为更一般的线性本构方程，被称为玻尔兹曼叠加原理（BSP），玻尔兹曼叠加原理广泛应用于描述线性黏弹性材料的行为，例如应力松弛、蠕变、动态力学响应等^[27]。

1.2.2 非线性本构方程

线性本构方程只能用于描述只能描述简单的材料行为，例如弹性变形、小应变下的黏弹性行为^[11,21,28]。适用于材料在小变形范围内的线性响应，而非线性本构方程可以用来描述更为复杂的行为例如塑性变形、硬化或软化、各向异性、大变形、剪切稀化、剪切增稠等行为。绝大部分高分子材料具有较复杂的非线性关系。Binham 认为当剪应力低于屈服应力时，屈服性流体表现为刚性固体；当剪应力超过屈服应力时，流体开始流动，且流动行为类似于牛顿流体^[29]。Herschel-Bulkley 模型在此基础之上引入了剪切稀化和剪切增稠的表示，如公式(1-9)，Herschel-Bulkley 模型是描述非线性本构方程的一套公式^[30]。

$$\sigma = \sigma_0 + K\dot{\gamma}^n \quad (1-9)$$

改进的 Herschel-Bulkley 类模型在 Herschel-Bulkley 模型的基础上，通过引入额外的参数或修正项，显著提升了对流体行为的描述精度。例如，在 Herschel-Bulkley 模型中引入高阶项（如剪切速率的二阶项），可以更准确地刻画非线性流变行为^[31]。此外，通过引入 Papanastasiou 正则化方法，可以解决原始模型在低剪切速率下的数值不稳定性问题，这一改进模型被称为 Herschel-Bulkley-Papanastasiou (HBP) 模型。HBP 模型在磁流变液等复杂流体的流变特性描述中得到了广泛应用^[32]。

Herschel-Bulkley 类模型不涉及弹性流体，主要用于解决屈服应力流体的本构问题。在非线性本构方程中，松弛模量函数不仅是时间的函数，也是应变的函数，能够表征大变形下的非线性响应。对于黏弹性流体的非线性本构方程而言，可以写出一般的通式，

如公式(1-10), 非线性黏弹性流体的本构方程主要通过积分和微分形式分别进行描述^[3]。

$$\sigma(t) = \int_{-\infty}^t G(t-t', \gamma) \frac{d\gamma(t')}{dt'} dt' \quad (1-10)$$

微分形式的 Oldroyd-B 模型, 如公式(1-11)所示是在 Maxwell 模型的基础上作了修正, 增加了延迟时间项 λ_2 , 从而能够描述更复杂的流变行为, 同时这个模型引入了上随体导数来代替普通导数, 在物理上更加符合真实世界的材料行为^[33]。Oldroyd-B 模型可以在 Weissenberg 数 ($W_i = \lambda_1 \dot{\gamma}_i$) 较小的情况下描述线性黏弹性, 其中的上随体导数在一定程度上包含了部分非线性效应, 因此该模型也可以一定程度上描述非线性关系。Oldroyd-B 模型是非线性本构方程的经典基础模型, 研究者在此基础上为了更准确地描述非线性黏弹性行为, 提出来多种基于此模型的本构模型。

$$\boldsymbol{\sigma} + \lambda_1 \frac{D\boldsymbol{\sigma}}{Dt} = \eta \left(\dot{\gamma} + \lambda_2 \frac{D\dot{\gamma}}{Dt} \right) \quad (1-11)$$

Giesekus 模型 (公式(1-12)) 在 Oldroyd-B 模型基础上, 增加了一个非线性项, 通过非线性系数 α 来表示非线性行为^[34]。Giesekus 模型常用于描述高聚物溶液、熔体以及其他粘弹性流体的流变行为。这些流体通常表现出剪切稀化和弹性效应。尤其是在中等至高剪切速率范围内。其模型参数 (如迁移因子 α) 可以调节剪切稀化的强度和拐点形状, 具有较高的灵活性^[35-36]。Giesekus 模型在高 Weissenberg 数 (Wi) 条件下仍能保持数值稳定性, 适用于强弹性效应的流动场景。通过引入对数构象重构等方法, 可以进一步提高其在高 Wi 条件下的计算稳定性^[37]。

$$\boldsymbol{\sigma} + \lambda_1 \frac{D\boldsymbol{\sigma}}{Dt} + \alpha \frac{\lambda_1}{\eta} \boldsymbol{\sigma} \cdot \boldsymbol{\sigma} = \eta \left(\dot{\gamma} + \lambda_2 \frac{D\dot{\gamma}}{Dt} \right) \quad (1-12)$$

PTT 模型 (公式(1-13)) 通过引入一个非线性应力函数 f 扩展了 Oldroyd-B 模型, 该函数通常取指数形式^[38]。

$$f(\text{tr}(\boldsymbol{\sigma})) \boldsymbol{\sigma} + \lambda_1 \frac{D\boldsymbol{\sigma}}{Dt} = \eta \left(\dot{\gamma} + \lambda_2 \frac{D\dot{\gamma}}{Dt} \right) \quad (1-13)$$

FENE-P 模型 (公式(1-14)) 在 Oldroyd-B 模型的基础上引入了有限拉伸效应, 通过项 $\frac{\lambda_1}{\eta} \frac{\boldsymbol{\sigma}}{1 - \text{tr}(\boldsymbol{\sigma})/b}$ 描述聚合物链的有限拉伸行为^[39]。参数 b 表示聚合物链的最大拉伸比。该模型适用于描述聚合物溶液在强流动条件下的非线性行为。

$$\boldsymbol{\sigma} + \lambda_1 \frac{D\boldsymbol{\sigma}}{Dt} = \eta \left(\dot{\gamma} + \lambda_2 \frac{D\dot{\gamma}}{Dt} \right) - \frac{\lambda_1}{\eta} \frac{\boldsymbol{\sigma}}{1 - \text{tr}(\boldsymbol{\sigma})/b} \quad (1-14)$$

通过将 Oldroyd-B 模型与分数阶导数结合, 可以得到分数阶 Oldroyd-B 模型, 如公

式(1-15)。分数阶 Oldroyd-B 模型通过引入分数阶导数描述非局部记忆效应^[40]。参数 α 和 β 是分数阶导数的阶数，该模型能够捕捉更复杂的流变行为，适用于具有非局部记忆效应的复杂流体。

$$\boldsymbol{\sigma} + \lambda_1^\alpha \frac{D^\alpha \boldsymbol{\sigma}}{Dt^\alpha} = \eta \left(\dot{\gamma} + \lambda_2^\beta \frac{D^\beta \dot{\gamma}}{Dt^\beta} \right) \quad (1-15)$$

另一类本构模型如 K-BKZ 模型源于积分型 Maxwell 模型（公式(1-4)^[41-42]）。K-BKZ 模型的基本形式如公式(1-16)所示。在 K-BKZ 模型中， h 称为阻尼函数，它是形变张量的第一不变量 I_1 和第二不变量 I_2 的函数。 $\mathbf{C}^{-1}(t, t')$ 是 Finger 形变张量的逆，用于描述从时间 t' 到 t 的形变历史。 $m(t - t')$ 是瞬态函数或记忆函数，用于表征材料对历史形变的记忆效应。K-BKZ 模型广泛应用于聚合物加工（如挤出、注塑、热成型等）、生物流体力学（如血液、蛋白质悬浮液等复杂流体的流变行为研究）以及涂料和润滑剂的流动行为和流变特性分析^[43]。

$$\boldsymbol{\sigma}(t) = \int_{-\infty}^t m(t - t') h(I_1, I_2) \mathbf{C}^{-1}(t, t') dt' \quad (1-16)$$

Doi 和 Edwards 尝试从分子角度构建本构方程，在管子模型的基础上提出了 Doi-Edwards 模型（公式(1-17)）。Doi-Edwards 模型的核心思想是将高分子链的缠结效应简化为一条光滑管道对链的限制作用，链在管道中的运动通过松弛和扩散来描述^[44-47]。其中 G_0 表示松弛时间，而 Q 表示为公式(1-18)，反映了高分子链在形变历史下的方向分布变化，即链的取向如何随形变而变化。

$$\boldsymbol{\sigma}(t) = G_0 \int_{-\infty}^t \frac{\partial Q(\mathbf{F}(t, t'))}{\partial t'} dt' \quad (1-17)$$

$$Q(\mathbf{F}(t, t')) = \frac{5}{2} \left\langle \frac{\mathbf{F}(t, t') \cdot \mathbf{u} \mathbf{u}}{|\mathbf{F}(t, t') \cdot \mathbf{u}|^2} \right\rangle_{\mathbf{u}} \quad (1-18)$$

传统的 Doi-Edwards 模型基于单链平均场近似，将缠结效应简化为一条光滑的“管子”对高分子链的限制作用。然而，这种简化忽略了链-链间的直接相互作用，难以解释快速大形变条件下的非线性流变现象（如应力过冲、缠结点破损和重组等）。近年来，研究者通过引入多链相互作用，提出了修正的管子模型，在此基础之上能够更好地描述缠结高分子流体的非线性行为^[48-50]。

1.2.3 传统的本构方程的构建方法

传统本构方程尤其是 Oldroyd-B、Doi-Edwards 等复杂的非线性模型具有较多的待定参数，通过方程来描述材料的应力-应变关系及其对时间、温度和形变历史的依赖性。

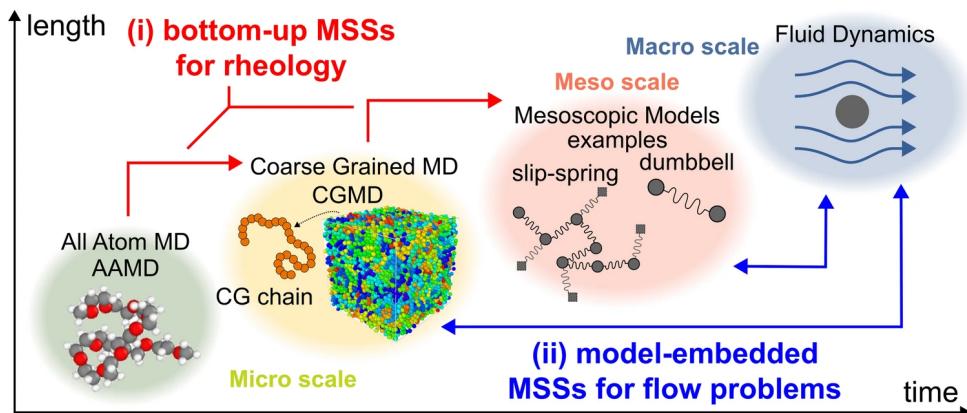
其构建方法是实验观测、理论推导和数值模拟的结合，具体可以分为以下几个步骤。

首先，实验观测是构建本构方程的基础。通过流变学实验（如剪切流变、拉伸流变等），研究者可以获取材料在不同形变条件下的应力响应数据。这些实验数据为理论模型的构建提供了关键依据。例如，通过动态力学分析（DMA）测量材料的储存模量和损耗模量，可以确定其粘弹性特性；通过应力松弛和蠕变实验，可以推断材料的记忆效应和时间依赖性。实验数据的精确性和全面性直接决定了本构方程的适用性和预测能力^[51-52]。

其次，对于非线性材料，则需要引入更复杂的本构关系，这里主要还是基于线性黏弹性方程，通过各类引入非线性关系的方法来进行非线性关系推导，理论推导的关键在于如何将微观结构信息（如高分子链的缠结、颗粒的相互作用等）与宏观力学行为联系起来。而对于理论推导的结果，每一个非线性项或参数应当从数学角度进行证明^[53]。近年来，数学研究者对 Doi-Edwards 模型的适定性进行了深入研究。Chupin 等人通过 Schauder 不动点定理和 Galerkin 近似方法，证明 Doi-Edwards 模型在二维情况下的全局解存在性和唯一性^[49]。这一结果为模型的数学基础提供了严格的理论支持。

数值模拟在本构方程的构建和验证中发挥着关键作用。有限元分析（FEA）、有限差分法（FDM）和有限体积法（FVM）等数值方法能够将本构方程应用于复杂几何和边界条件下的力学问题，从而模拟材料的应力分布、形变行为和流动特性^[51]。分子动力学模拟（MD）是另一种重要的数值工具，能够从分子尺度模拟材料的力学行为，为宏观本构方程的构建提供微观依据。例如，通过粗粒度分子动力学模拟（CGMD），研究者可以研究高分子链的缠结力学，验证管子模型的假设^[54-55]。由于聚合物液体具有单体、中观和宏观流尺度的多尺度特征，因此关联不同层次的多尺度模拟（MSS）能够更准确地再现其流动特性^[56]。近年来，基于分子动力学的多尺度模拟方法显著推进了复杂流体本构关系的构建。Webb 等人提出基于图论的系统粗粒化方法，通过拓扑约简实现分子结构特征的有效保留，为建立跨尺度本构方程提供了数学基础^[57]。在此框架下，Behbahani 开发了原子模拟-粗粒化-滑移弹簧的分层耦合算法，成功预测了聚乙烯熔体的非线性粘弹性响应^[58]。Morii 提出的拉格朗日多尺度框架实现了从分子涨落到连续介质流动的全域耦合，在高弹性数工况下展现出优于传统欧拉方法的数值稳定性^[59]。

传统本构方程的研究方法主要依赖于物理实验和理论推导，通过建立数学方程来描述材料的力学行为。这种方法虽然具有明确的物理意义，但在处理复杂材料或非线性行为时，往往面临模型精度不足、参数识别困难等问题^[60]。随着数据驱动技术的快速发展

图 1-3 多尺度模拟示意图^[56]Figure 1-3 Schematic illustration of Multiscale Simulation^[56]

展，机器学习为材料本构关系的研究提供了新的思路。通过利用大量实验或仿真数据，机器学习能够自动挖掘材料行为中的潜在规律，构建高精度的预测模型，从而弥补传统方法的不足，并为材料科学的研究开辟了更加智能化的路径。

1.3 数据驱动方法流变学本构建模

1.3.1 机器学习方法介绍

机器学习是人工智能的一个重要分支，其核心是通过算法从数据中自动学习规律，并利用这些规律进行预测或决策。与传统编程不同，机器学习不依赖于明确的规则，而是通过训练数据优化模型参数，从而实现对复杂问题的建模和解决。它在图像识别、自然语言处理、推荐系统等领域取得了显著成果^[61]。

监督学习是最常见的机器学习类型，适用于有标签的数据集。常见的算法包括线性回归和逻辑回归，分别用于连续值的预测和二分类问题^[62-63]。决策树（DT）通过树状结构进行决策，适用于分类和回归任务^[64]。随机森林（RF）是多个决策树的集成，通过投票或平均提高预测准确性^[65]。支持向量机（SVM）通过寻找最优超平面进行分类，适用于高维数据^[66]。K 近邻算法（KNN）基于距离度量进行分类或回归，简单但计算量大^[67]。

随着数据规模和计算能力的提升，深度学习作为机器学习的一个子领域迅速崛起。深度学习通过构建多层的神经网络结构，能够自动提取数据中的多层次特征，从而在处理高维、非线性问题（如图像、语音和文本）时表现出更强的能力，成为推动人工智能发展的核心技术之一^[61]。

近年来，深度学习领域在理论创新与应用拓展方面取得了显著进展，极大地推动了机器学习技术的前沿发展。在生成模型领域，扩散模型通过模拟数据从噪声分布到目标分布的逆扩散过程，逐步生成高保真度的图像样本，在图像生成、风格迁移和图像修复等任务中展现出卓越的性能^[68]。在多模态学习方面，研究者通过联合建模文本、图像、音频和视频等多种模态数据，实现了更全面的语义理解与跨模态生成。以模型为代表的多模态预训练框架，通过对比学习策略学习图像-文本对的联合表示空间，在零样本分类和跨模态检索任务中取得了突破性进展^[69]。在自监督学习领域，研究者通过设计预训练任务从数据本身生成监督信号，显著降低了对人工标注数据的依赖^[70]。对比学习作为自监督学习的核心范式之一，通过最大化正样本对的表示一致性并最小化负样本对的相似性，有效提升了模型的特征提取能力^[71]。这些前沿进展不仅深化了深度学习理论体系，也为解决实际问题提供了新的方法论支持。未来，随着计算能力的提升和数据规模的扩大，深度学习技术有望在医疗影像分析、自动驾驶、智能内容创作等领域展现出更广泛的应用潜力。同时，模型的可解释性、鲁棒性和能效优化等方向仍面临重要挑战，需要跨学科合作以推动该领域的持续发展^[61]。

1.3.2 机器学习流变学应用研究现状

1.3.2.1 引言

传统机器学习或者是目前最前沿的深度学习研究，最初都集中于计算机科学、人工智能领域，但是近年来在物理学领域，机器学习与物理学问题的研究结合也日益密切^[72]。在量子物理中，机器学习被用于量子态重构、量子电路优化和量子相变识别，显著提高了量子系统的分析和计算效率^[73]。在凝聚态物理领域，机器学习通过预测材料性质和分类相图，加速了新材料的发现和设计^[72]。在天体物理中，机器学习被用于引力波探测和宇宙微波背景辐射分析，揭示了宇宙的起源和结构^[74]。在流体力学中，机器学习通过数据驱动的方法直接从流体数据中学习湍流动力学，显著提高了湍流模型的精度和计算效率^[75]。此外，机器学习中的符号回归技术能够从实验数据中自动发现物理定律，为探索未知物理规律提供了新工具^[76]，在多尺度物理系统中，机器学习被用于气候模拟和生物信息学建模，预测复杂动态和处理多重空间和时间尺度的混沌系统。

在流变学领域，机器学习也被广泛应用。传统的流变学研究方法如上一章所述，首先通过实验获取数据，再通过数学物理方程解释数据。流变学家总是期待于对每个流变学体系建立一个可表示的数学方程，尽管这个方程可能完全无法求出解析解也极难获取

稳定精确的数值解。这本质上是一个由因推果的过程，而数据驱动的机器学习方法则从真实的结果出发，构建一个完全匹配的数学方程，这个方程的形式与参数都不确定，但是可以泛化本构现象，即通过含有大量参数的非方程化的计算机模型代替具体形式的本构方程来解决流变学本构问题^[77-79]。

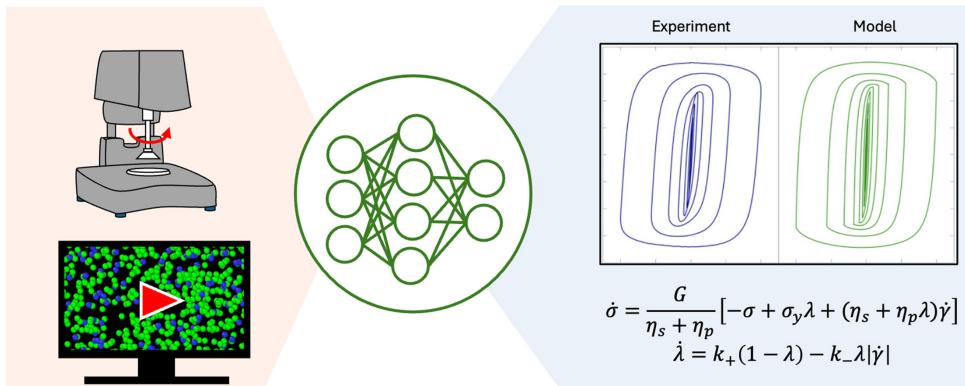


图 1-4 数据驱动方法应用于流变学研究示意图^[79]

Figure 1-4 Schematic illustration of data-driven methods applied to rheology research^[79]

1.3.2.2 流变特性预测

机器学习应用于流变学本构关系研究的其中一个应用是流变特性的预测。人工神经网络（ANN）已被用于自动监测合成油基泥浆的各种流变特性，使用泥浆密度和沼泽漏斗作为输入，预测值与实际测量值非常接近，平均绝对误差小于 9.66%^[80]。混合机器学习模型结合了 ANN 和 SVM 准确预测了纳米基水基钻井液的流变性和过滤特性。四种不同的模型——多元线性回归（MLR）、SVM、回归决策树（CART）和 ANN 被一起使用，以根据溶液特性预测聚合物溶液的粘度，有助于提高石油采收率^[81]。此外，机器学习模型可以与微流体或其他设备集成，用于复杂流体的原位粘度测量。例如，Mustafa 等设计了一种微流体传感设备，利用流固耦合和 ML 算法来测量复杂流体的粘度^[82]。他们采用 SVM 和 KNN 算法，分别实现了 89.7% 和 98.9% 的平均准确率。Ponick 等利用卷积神经网络（CNN）通过立体相机图像预测 Bingham 流体的流变特性^[83]。Chen 等人通过将群贡献（GC）方法与人工神经网络（ANN）相结合，以预测离子液体（IL）-水混合物的粘度^[84]，取得比单一 ANN 结构更好的预测效果。Zhang 等人开发了一种深度半监督的基于即时学习的高斯过程回归（DSSJITGPR）用于门尼粘度估计^[85]。它将即时学习、半监督学习和深度学习集成到一个统一的建模框架中。DSSJITGPR 的有效性和优越性已通过工业橡胶混炼过程的 Mooney 粘度预测结果得到验证。

1.3.2.3 材料表征和工艺设计优化

机器学习在流变学本构关系研究中的应用不仅限于材料表征和工艺设计/优化，还通过直接引入材料制备和表征中的各种参数作为特征，显著简化了传统本构方程中针对特定材料体系拟合不同参数的复杂性。与传统统计方法相比，机器学习算法能够更高效地对这些复杂关系进行建模，其中主成分分析和深度神经网络等技术在管理和减少材料表征中常见的高维数据方面发挥了重要作用，从而实现了更高效的分析和解释。这些方法已广泛应用于 3D 打印混凝土、纤维增强混凝土、聚合物纳米复合材料、生物墨水、食品材料和高分子科学等领域。Zhang 和 Shao 的研究进一步探讨了基于图像的机器学习技术在材料科学中的应用，以应对各种挑战和任务，由于这些方法的通用性和可转移性，它们同样适用于流变学表征^[86]。机器学习模型还彻底改变了高通量表征，为数据分析、预测和优化提供了强大的工具。例如，Zhang 等人开发了一种自主高通量系统，采用支持向量机（SVM）、随机森林（RF）和极端梯度提升（XGB）分类器等监督式机器学习算法，快速表征水凝胶的流变特性，所有模型在水凝胶相分类中均表现出色^[87]。Verheyen 等人则利用基于树的集成算法（如随机森林和梯度提升）将数据驱动建模与颗粒水凝胶基质系统的实验优化相结合，这些技术因其在处理分类或回归任务、非线性关系、高维数据和混合数据类型方面的灵活性而被选用^[88]。Martineau 等人开发了一种基于高斯过程（GP）建模的管道，用于识别细菌嵌入的丝水凝胶的凝胶化状态，通过比较人类专家和机器学习引导算法的性能，发现两者结合最能加速发现过程^[89]。此外，随机森林和多元线性回归（MLR）等技术已被用作回归工具，将从大振幅振荡剪切（LAOS）实验中获得的流变特性与人类感官感知（如涂抹性）相关联，从而改进化妆品行业的配方设计。Lee 等人的研究表明，由于流变测量和铺展性之间的非线性关系，随机森林模型的性能优于多元线性回归模型^[90]。

1.3.2.4 流变学计算模拟

机器学习为流变学的计算模拟提供助力，计算平台已经成为流变学研究的重要支柱，广泛应用于从聚合物动力学到颗粒系统的详细模拟，为复杂流体在不同条件下的行为提供了深刻的物理见解。然而，这些技术的主要瓶颈在于模拟与实验条件相关的时间和长度尺度的能力。由于大规模模拟中的大部分计算资源消耗在重复的数学运算上，而这些运算可以通过数据驱动技术高效学习，因此将机器学习与模拟相结合能够显著提升效率和精度。例如，Lu 等人基于卷积神经网络的仿真模型在颗粒流模拟中实现了显著

的加速，同时保持了较高的精度^[91]。高斯过程回归则被 Seryo 等人用于从微观聚合物模拟中推断本构关系，并将这些关系应用于宏观流动仿真，从而调整流体中的应力分布以反映微观动力学^[92]。类似的方法还被应用于缠结良好的聚合物熔体，通过机器学习的本构关系优化模拟效果。此外，Bai 等人介绍了一种数据驱动的平滑粒子流体动力学方法，该方法利用实验数据提高了牛顿流体和非牛顿流体的建模精度，即使在数据集较小的情况下也能准确预测幂律流体的速度曲线，同时显著优化计算时间^[93]。多尺度建模通过在不同尺度之间交换信息来捕捉材料的复杂行为，例如 Li 等人基于生成对抗网络的方法实现了从粗粒度到原子级别的结构回溯映射^[94]。主动学习作为一种高效的机器学习范式，通过选择性地标记信息量最大的数据点，以最少的数据优化学习过程，例如在多尺度建模中，Zhao 等人基于高斯过程回归的主动学习策略将所需的模拟次数大幅减少，显著提高了研究复杂材料行为的效率^[95]。这些技术的结合不仅提升了模拟的精度和效率，还为流变学研究提供了更强大的工具。

1.3.2.5 小结

无论其架构或算法如何设计，以数据为中心的机器学习模型严重依赖丰富的训练数据来确保预测的可靠性和准确性。此外，材料特性通常对成分和加工条件的细微变化表现出高度敏感性，这可能导致材料行为的显著变化。因此，数据收集过程必须足够全面和细致，以捕捉这种复杂性，从而确保以数据为中心的模型能够做出准确的预测。例如，对于具有整体 Herschel-Bulkley 类型行为的屈服应力流体，如果算法仅暴露在大变形率或小变形率下的数据中，它将无法对观察域之外的一般行为做出可靠预测^[79,96-97]。更重要的是，这些模型完全依赖于数据相关性和统计规律，而忽略了基础科学理论的支持，例如黏弹性材料的时间依赖性和应变依赖性等关键物理特性。纯数据驱动方法往往忽视了这些物理细节，导致模型在实际应用中可能失效。因此，在开发算法时，我们应当引入物理约束，将黏弹性的时间依赖性等基本物理规律纳入模型框架，以弥补数据驱动方法的不足，从而提升模型的预测能力和泛化性能。

1.4 引入物理约束的神经网络研究

1.4.1 引言

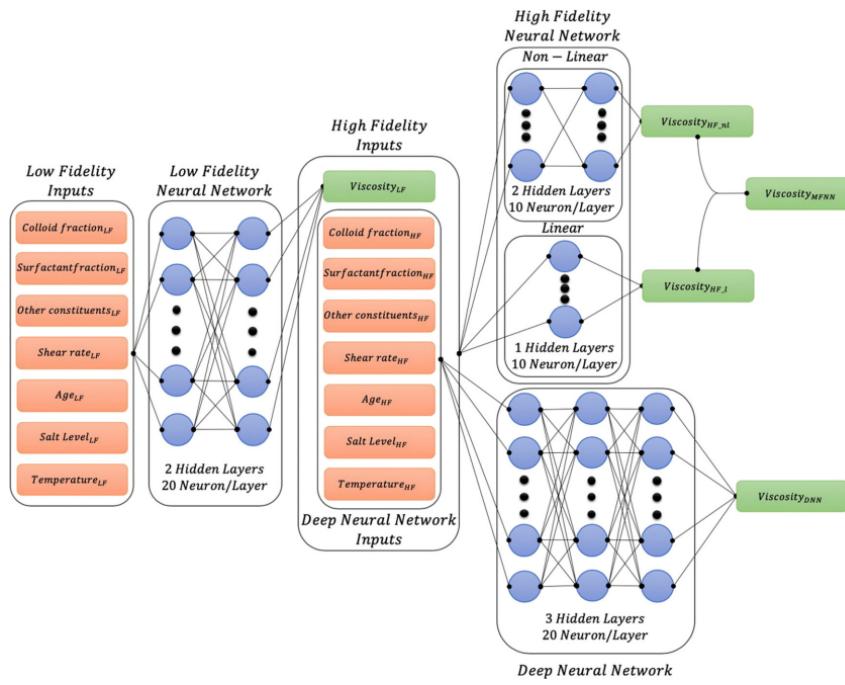
近年来，多种方法被提出并广泛应用于将基本物理规则和领域知识整合到机器学习框架中，这些方法在众多领域展现了深远的影响。这类模型的总体框架通常被称为物理

信息机器学习（PIML）。PIML 的一个开创性范例是“物理信息神经网络”（PINN），它代表了科学计算和应用数学领域的重大突破^[98-99]。PINN 将机器学习算法与从系统观察、经验以及物理或数学理解中获得的先验知识无缝结合，通过将这些先验知识直接嵌入模型架构中，显著减少了对大规模训练数据集的依赖，从而能够利用更少的观测数据解决复杂的物理问题。

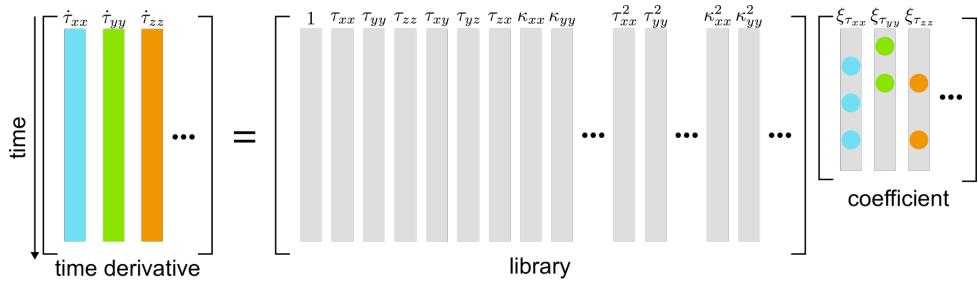
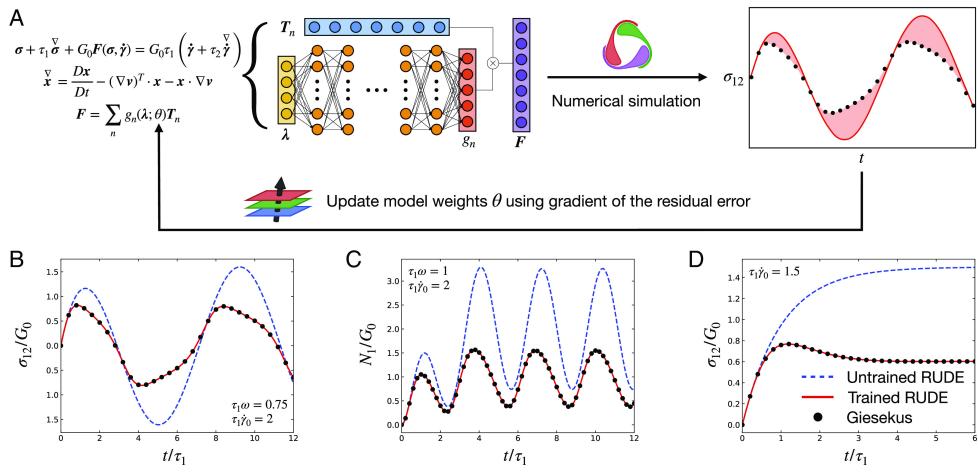
1.4.2 物理信息神经网络研究现状

近年来，物理信息神经网络（PINN）在流变学建模领域取得了重要进展，有效克服了传统方法依赖简化假设和难以处理复杂材料特性的局限。Mahmoudabadbozchelou 等人提出的流变学信息神经网络（RhINN）框架，能够处理触变性弹性粘塑性（TEVP）模型等多种本构模型^[100]。该框架以时间 t 和剪切速率为输入，剪应力 σ 和结构参数 λ 为输出，通过最小化包含方程残差和初始条件差异的损失函数来训练网络参数。在逆问题求解中，RhINN 可同时学习网络参数和流变参数，其预测结果与真实数据高度吻合，展现了良好的鲁棒性。Dabiri 等人进一步将 RhINN 扩展至分数阶微分方程，提升了其在复杂本构模型参数识别中的应用价值^[101]。在相关研究中，Zhang 等人开发的 RheologyNet 成功应用于胶凝材料触变特性评估，其预测结果与有限元分析高度一致^[102]。Nagrani 等人则利用 PINN 研究了导热硅脂的流变行为，通过实验数据确定了流变参数^[103]。基于物理的机器学习方法还可从间接观测中学习未知流变模型，例如通过三维流场数据推断广义牛顿流体的稳态剪切粘度，这一方法在聚合物熔体和颗粒悬浮液研究中取得了显著成果。此外，PINN 在稀薄气体流动预测中也展现出应用潜力^[104]。Howard 等人将悬浮平衡模型与神经网络结合，成功预测了单分散悬浮液的颗粒应力，但在双分散系统中的应用仍存在局限。这些研究充分展示了物理信息神经网络在流变学领域的广泛应用前景^[105]。

多保真度建模是一种强大的技术，通过整合来自不同保真度和计算成本的数据来增强模型预测能力。这种方法在物理定律不精确或高保真数据生成成本较高的情况下尤为有用。通过高效结合高保真和低保真数据，多保真度建模为复杂系统的精确建模提供了一种更高效且更具成本效益的解决方案。例如，Mahmoudabadbozchelou 等人利用多保真度方法构建了一个流变学元模型，用于预测复杂多组分系统的流变响应，同时考虑了老化、混合物的盐度和温度依赖性等输入参数^[106]。在该框架中，低保真数据通过多种本构方程生成，而高保真数据则来自实验测量。与传统的高保真神经网络相比，这种组合方法显著提高了预测精度，充分展示了多保真度建模的优势。

图 1-5 多保真神经网络 (MFNN) 结构示意图^[106]Figure 1-5 Schematic illustration of Multifidelity Neural Network (MFNN) structure^[106]

类似的方法也被应用于预测纤维悬浮液的流变特性，其中基于法向载荷相关摩擦系数模型的数值模拟用于生成高保真数据，而低保真数据则通过不同的本构模型生成。纤维的物理特性，包括纵横比、纤维刚度、表面粗糙度以及体积分数，均被用作输入参数。这种将基础物理特性融入机器学习算法的模式并不局限于特定方法，因此多保真度建模也可以与直接或反向流变学信息神经网络（RhINN）相结合，以利用不同级别的数据准确性来优化模型预测。这种混合方法能够从有限的实验观察中做出高度准确的流变学预测，例如 Mahmoudabadbozchelou 等人对气相硅胶流变学的成功预测^[107]。虽然多保真度算法通常依赖于不同准确性数据的可用性，但在上述应用中，这些数据是通过描述目标系统的物理或现象学模型直接生成的，因此这些方法被归类为物理信息建模。然而，如果不同保真度级别的数据源与基础物理学无关，则该方法更倾向于以数据为中心。例如，分层机器学习模型通过结合墨水和打印机参数之间的已知物理关系，能够在小数据集上实现精确建模。这种灵活性使得多保真度建模在流变学和其他复杂系统的研究中具有广泛的应用潜力。机器学习技术的另一个重要应用在于发现闭式本构关系，这些关系从数学上描述了应力和变形之间的关联。这一过程通常通过分析实验数据并结合非线性动力学稀疏识别（SINDy）和符号回归等技术来实现^[108]。在这些方法中，通常采用多步骤流程：首先从广泛的潜在函数列表中筛选出最重要的候选函数，然后精确恢复

图 1-6 Rheo-SINDy 的示意图^[108]Figure 1-6 Schematic illustration of Rheo-SINDy^[108]图 1-7 (A) RUDE 训练循环示意图 (B-D) 基于 Giesekus 模型合成 LAOS 数据训练的 RUDE 评估结果, 黑色圆圈: 测试数据; 红色线条: 训练后的 RUDE; 蓝色线条: 未训练的 RUDE (B) 中间频率下的剪切应力响应 (C) 训练频率下的法向应力 (D) 稳态剪切流启动时的剪切应力响应^[109]Figure 1-7 (A) Schematic depiction of a RUDE within the training loop (B-D) Evaluation of a RUDE trained on synthetic LAOS data for the Giesekus model, Black circles: test data; red lines: trained RUDE; blue lines: untrained RUDE (B) Shear stress response at an intermediate frequency (C) Normal stress at a training frequency (D) Shear stress during startup of steady shear flow^[109]

这些函数的参数, 以最简洁的形式描述动力学系统。例如, SINDy 方法被直接扩展到流变学应用中, 开发了一种称为 Rheo-SINDy 的稀疏识别方法, 用于从已知方程库生成的流变学数据中恢复本构方程。然而, 现有方法在处理不完整、稀缺或噪声数据时面临挑战, 特别是需要数值微分来计算导数以构建控制方程^[110]。这一问题可以通过结合自动微分功能的物理信息神经网络 (PINN) 来解决, 从而避免数值微分的需求。这种组合

方法被称为 PINN-SR（具有稀疏回归的 PINN），已成功应用于从模型弹性粘塑性流体的实验数据中提取精确的本构关系。

尽管目前大多数机器学习技术（无论是数据驱动还是物理信息）都集中在单个标量分量（如剪切应力、微观结构参数或粘度）作为流变学特征，并与粘度流动和体材料函数预测相关，但真正可推广的流变学机器学习技术必须具备张量性质，并避免客观性问题。Lennon 等人提出的 RUDE 框架优雅地解决了这些基本约束，构建了包含物理信息的可学习模型，同时对特定实验协议或流动运动学的细节保持不可知^[109]。在该框架中，广义粘弹性模型中的未知非线性项被表示为张量基函数的线性组合，这些函数强制执行对称性和框架不变性等物理约束，而张量基函数的系数则通过神经网络输出建模。RUDE 框架的一个显著优势在于，其预测不仅能够推广到观测域之外，还可以扩展到应力张量的其他分量，从而实现完全解析的流动预测。这种能力使得 RUDE 在流变学应用中展现出强大的泛化能力和实用性。

1.4.3 其他物理信息深度学习模型

近年来，深度神经算子作为一种新的机器学习模型被提出，用于隐式学习物理现象的解算子^[111]。与使用有限维向量空间的标准神经网络不同，神经算子学习函数空间之间的映射。典型的架构包括 DeepONet 和傅里叶神经算子（FNO）及其变体。DeepONet 由两个子网络组成：分支网络和中继网络，分别提取输入函数和输入坐标的潜在表示，并通过点积合并输出。FNO 通过在傅里叶空间中参数化积分核来实现高效的架构。神经算子可以作为隐藏控制方程的隐式解算子，使其成为理解复杂物理系统的强大工具。它们可以与物理场和其他域约束结合，以获得高保真解和良好的泛化能力。

神经算子在简洁且准确地学习复杂动力学方面表现出色，已在流体力学应用中得到验证，如天气预报和碳捕获的储层工程。然而，在复杂流体领域的应用仍有限。Rashid 等人采用 FNO 架构预测了数字复合材料中的应力和应变场^[112]。FNO 的扩展，如隐式傅里叶神经算子（IFNO），已被证明可以预测不可见载荷条件下的材料响应，并且同样适用于流变学应用^[113]。在 IFNO 中，层之间的增量由积分算子建模，因此所得架构可以解释为未知控制定律的定点方法。与传统的本构模型相比，这种方法将预测误差减少了十倍。Howard 等人还应用了模型算子回归（MOR-physics）框架，从单分散和双分散悬浮液中的体积分数和速度测量中学习粒子应力^[105]。

1.5 本课题研究介绍

1.5.1 研究内容

本文旨在通过多方法融合的系统性研究，探索深度学习在模拟数据和实验数据下本构方程的建模与评估，文章主要结构如下，研究路线如图1-8所示。

第一章首先综述了流变学研究的基本问题和应用方向。其次讨论了本构方程理论的发展，经典本构方程的发展历史，数学形式，适用范围等。然后综述了本构方程的研究方法，包括传统的数值方法和近年来被广泛研究的数据驱动机器学习方法。最后综述了将物理约束引入机器学习方法的研究现状，并简述了本文后续的课题设计方案。

第二章主要介绍了本文课题中主要使用的各类算法，从算法理论到选型依据做了综述讨论。第三章介绍了本文的第一部分工作，使用门控循环单元（GRU）对经典本构方

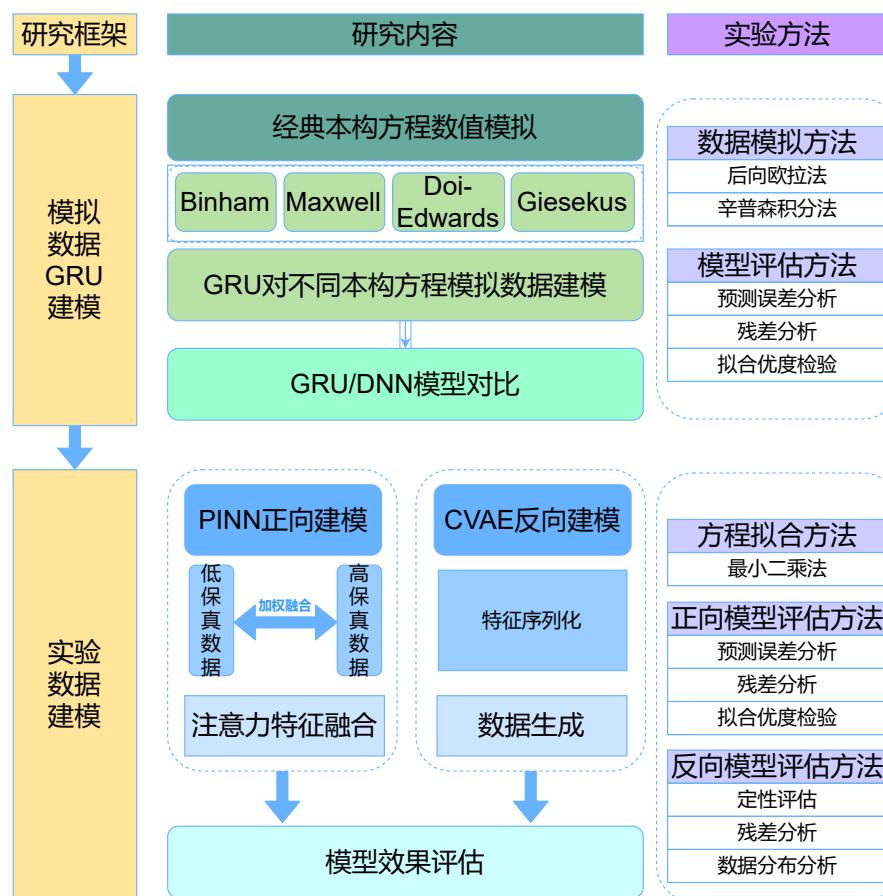


图 1-8 研究路线图

Figure 1-8 Research roadmap

程的模拟数据进行深度学习建模，并使用各种指标对模型性能进行评估。具体来说，首先基于经典本构方程（Binham 模型、Maxwell 模型、Doi-Edwards 模型、Giesekus 模型），

采用后向欧拉法和辛普森积分法等数值模拟方法生成模拟数据，并划分训练集、验证集、测试集。随后使用 GRU 分别对不同的模型模拟数据进行建模，训练后使用测试集进行模型的评估，采用决定系数 (R^2)、平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE) 等指标，对模型性能进行定量评估。同时在相同的数据集上使用普通深度神经网络 (DNN) 进行训练，将两套模型进行比较，以验证 GRU 模型在处理复杂非线性时间依赖性流变学本构方程时的优势。

第四章介绍了本文的第二部分工作，针对黏弹性凝胶材料的真实实验流变数据进行物理信息神经网络 (PINN) 的建模，在 Mahmoudabadbozchelou 等人的模型基础之上，进一步优化 PINN 的模型，采用可学习的损失权重进行损失函数的优化，探究 PINN 对特定参数下凝胶材料模量及损耗因子的预测能力。随后，采用条件变分自编码器 (CVAE) 反向建立生成式模型，通过这个模型生成特定流变学数据下的制备参数，并进行生成效果评估。

1.5.2 研究创新点

- (1) 本文采用门控循环单元 (GRU)，通过其门控机制捕捉时间序列数据中的时间依赖性，使模型更贴近复杂流体的物理特性。
- (2) 优化 Mahmoudabadbozchelou 提出的物理信息神经网络 (PINN)，引入可学习的权重损失以缓解多损失函数的梯度消失问题，并结合注意力特征融合解决实验数据稀疏问题。
- (3) 使用 CVAE 反向建模，以流变学特性反预测制备参数，辅助实验设计。

1.5.3 研究意义

传统机器学习方法在流变学建模中往往局限于单点预测，难以捕捉复杂流体的时间依赖性，而本文采用门控循环单元 (GRU) 进行建模，充分利用其时间序列处理能力，使模型更贴近黏弹性材料的流变学特性。这一研究不仅拓展了深度学习在流变学领域的应用范围，还为复杂流体建模提供了新的理论支持。随着计算机科学的发展，自注意力机制（如 Transformer）在自然语言处理等领域展现了强大的长程依赖捕捉能力。尽管受限于流变学数据的稀缺性，本文未直接采用 Transformer 架构，但通过探索能够解决长期序列依赖的模型，为未来本构方程建模的研究指明了方向。这一思路有望进一步推动流变学本构建模的发展，为解决复杂流体建模问题提供更强大的工具。

本文通过优化物理信息神经网络 (PINN)，引入可学习权重和注意力特征融合，为

小样本学习提供了新的解决方案，提升了模型在数据稀缺情况下的性能。同时，利用条件变分自编码器（CVAE）进行反向建模，以流变学特性反推制备参数，为实验设计提供了理论支持。这些研究不仅推动了流变学建模方法的进步，还为材料科学和工程领域的实验优化提供了新的思路，具有一定的理论意义和实践价值。

第二章 算法介绍

2.1 引言

本文基本的研究思路是使用数值模拟算法和深度学习算法来研究流变学本构方程建模的问题。本文第一部分基于动态时序建模理论，构建门控循环单元网络（GRU）的预测模型，通过深度学习算法处理数值模拟生成的时变本构关系数据，重点解析循环神经网络在时间序列建模中的特征提取机制，主要涉及的算法有 GRU 算法，本章同时综述了时间序列数据的其他深度学习算法。第二部分引入物理约束的智能建模策略，针对实际流变实验数据构建融合物理信息的混合神经网络模型，重点整合了物理信息神经网络（PINN）的微分方程嵌入技术、基于注意力机制驱动的多源特征融合方法，以及结合条件变分自编码框架（CVAE）的生成式建模策略，实现物理规律约束下的数据-模型双驱动建模。

本章将在此基础上，对研究中所涉及的关键算法进行理论概述，并阐述算法选型的依据，以系统性地综述相关领域的研究进展，为后续研究奠定理论基础。

2.2 算法介绍

2.2.1 时间序列数据介绍

时间序列数据是指在一系列时间点上按时间顺序排列的数据点集合。这些数据点通常是连续或定期记录的，反映了某个变量随时间的变化情况。在流变学研究中，时间域的应力应变数据在小应变时满足玻尔兹曼叠加原理，即当前应力状态为之前所有应变历史的线性和^[27]。大应变时应力应变数据虽然不满足玻尔兹曼的线性叠加，但是依旧可以表示为过去应变历史的非线性关系叠加。

2.2.2 循环神经网络

2.2.2.1 简单 RNN

循环神经网络（Recurrent Neural Network, RNN）是一种专门处理序列数据的神经网络结构，其核心在于利用循环结构捕捉时间或顺序上的依赖关系^[114]。RNN 的基本结构由输入层、隐藏层和输出层组成（图2-1 (a)）。隐藏层中的神经元不仅接收来自输入层的信号，还接收来自前一时刻隐藏层的信号。这种循环连接使得 RNN 能够捕获序列

中的时间依赖关系。RNN 的数学模型可以通过公式(2-1)描述:

$$\mathbf{h}_t = \sigma(\mathbf{W}_{xh}\mathbf{x}_t + \mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{b}_h) \quad (2-1)$$

$$\mathbf{o}_t = \mathbf{W}_{ho}\mathbf{h}_t + \mathbf{b}_o \quad (2-2)$$

其中 \mathbf{h}_t 是时刻 t 的隐藏状态, \mathbf{x}_t 是时刻 t 的输入, \mathbf{o}_t 是时刻 t 的输出, \mathbf{W}_{xh} 是输入到隐藏层的权重矩阵, \mathbf{W}_{hh} 是隐藏层到隐藏层的权重矩阵, \mathbf{W}_{ho} 是隐藏层到输出层的权重矩阵, \mathbf{b}_h 和 \mathbf{b}_o 是偏置。 σ 是激活函数, 通常使用 tanh 或 ReLU。RNN 的训练过程通常使用反向传播算法。通过计算梯度来更新权重矩阵, 从而最小化损失函数。

在简单的 RNN 中, 存在梯度消失问题, 这主要源于反向传播过程中梯度的连乘效应^[115]。损失函数 L 对 \mathbf{W}_{hh} 的梯度可以表示为公式(2-3):

$$\frac{\partial L}{\partial \mathbf{W}_{hh}} = \sum_{t=1}^T \frac{\partial L}{\partial \mathbf{h}_t} \cdot \frac{\partial \mathbf{h}_t}{\partial \mathbf{W}_{hh}} \quad (2-3)$$

展开后, 隐藏状态关于 \mathbf{W}_{hh} 的导数为公式(2-4):

$$\frac{\partial \mathbf{h}_t}{\partial \mathbf{W}_{hh}} = \sigma'(\mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h) \cdot \mathbf{h}_{t-1}^T \quad (2-4)$$

注意到, $\frac{\partial L}{\partial \mathbf{h}_t}$ 依赖于前一个时间步的梯度, 如公式(2-5):

$$\frac{\partial L}{\partial \mathbf{h}_t} = \frac{\partial L}{\partial \mathbf{h}_{t+1}} \cdot \frac{\partial \mathbf{h}_{t+1}}{\partial \mathbf{h}_t} = \frac{\partial L}{\partial \mathbf{h}_{t+1}} \cdot \mathbf{W}_{hh} \cdot \sigma'(\mathbf{h}_t) \quad (2-5)$$

这表明梯度在通过时间反向传播时会乘以 \mathbf{W}_{hh} 。从时间 t 到时间 T 的梯度可以表示为公式(2-6):

$$\frac{\partial L}{\partial \mathbf{h}_t} = \left(\prod_{k=t+1}^T \mathbf{W}_{hh} \cdot \sigma'(\mathbf{h}_{k-1}) \right) \cdot \frac{\partial L}{\partial \mathbf{h}_T} \quad (2-6)$$

其中乘积 $\prod_{k=t+1}^T \mathbf{W}_{hh} \cdot \sigma'(\mathbf{h}_{k-1})$ 在某些情况下可能会非常小, 例如如果 \mathbf{W}_{hh} 的特征值小于 1, 那么随着 $(T-t)$ 的增加, \mathbf{W}_{hh}^{T-t} 将呈指数级减小并趋近于零。如果反向传播过程中每一项都小于 1, 那么整个乘积将随着 $(T-t)$ 的增加呈指数级减小。这意味着对于远离输出端的时间步 t , 梯度 $\frac{\partial L}{\partial \mathbf{h}_t}$ 将非常小, 导致权重更新几乎停止, 这就是 RNN 的梯度消失问题。由于 RNN 为了捕捉时间依赖性, 我们不可避免地设置长时间步, 这导致 RNN 梯度消失问题几乎难以避免。

2.2.2.2 GRU

为了解决简单 RNN 的梯度消失问题, Schmidhuber 等提出了长短期记忆网络 (Long Short-Term Memory, LSTM)^[115]。LSTM 通过引入门控机制来控制信息的流动, 从而有效地缓解梯度消失问题。LSTM 的核心结构包括输入门、遗忘门和输出门, 这些门控机制能够有效地控制信息的流动, 从而捕获长期依赖关系。

LSTM 虽然通过引入门控机制有效地解决了简单 RNN 中的梯度消失问题, 但在实际应用中仍然存在一些问题。LSTM 包含三个门 (输入门、遗忘门、输出门), 导致其参数数量较多, 增加了模型的复杂度和计算成本。由于 LSTM 的复杂结构, 其训练和推理速度相对较慢, 尤其是在处理大规模数据时。为了解决 LSTM 的上述问题, Cho 等提出了 GRU 作为 LSTM 的简化版本^[116] (图2-1 (b))。GRU 的设计初衷是为了简化 LSTM 的结构, 同时保持其捕获长期依赖关系的能力。GRU 只有重置门 (Reset Gate) 和更新门 (Update Gate) 两种门控单元, 与 LSTM 相比, GRU 的参数数量较少, 训练和推理速度更快。重置门的计算公式为公式(2-7)。

$$\mathbf{r}_t = \sigma(\mathbf{W}_r \mathbf{x}_t + \mathbf{U}_r \mathbf{h}_{t-1} + \mathbf{b}_r) \quad (2-7)$$

更新门的计算公式为公式(2-8), 重置门决定了前一时间步的隐藏状态在多大程度上被忽略。当重置门的输出接近 0 时, 网络倾向于“忘记”前一时间步的信息, 仅依赖于当前输入; 而当输出接近 1 时, 前一时间步的信息将被更多地保留。更新门决定了多少过去的信息将被保留, 而新信息将占据多少比例。当更新门的输出接近 1 时, 新的隐藏状态几乎等同于旧的隐藏状态, 从而实现了长期依赖的捕捉; 当更新门的输出接近 0 时, 新的隐藏状态将主要由当前输入决定。

$$\mathbf{z}_t = \sigma(\mathbf{W}_z \mathbf{x}_t + \mathbf{U}_z \mathbf{h}_{t-1} + \mathbf{b}_z) \quad (2-8)$$

候选隐藏状态决定了当前时间步的新信息, 如公式(2-9)所示。

$$\tilde{\mathbf{h}}_t = \tanh(\mathbf{W}_h \mathbf{x}_t + \mathbf{U}_h (\mathbf{r}_t \odot \mathbf{h}_{t-1}) + \mathbf{b}_h) \quad (2-9)$$

当前隐藏状态如公式(2-10)所示, 这里的当前隐藏状态与上文 LSTM 的隐藏状态含义是一样的。

$$\mathbf{h}_t = (1 - \mathbf{z}_t) \odot \mathbf{h}_{t-1} + \mathbf{z}_t \odot \tilde{\mathbf{h}}_t \quad (2-10)$$

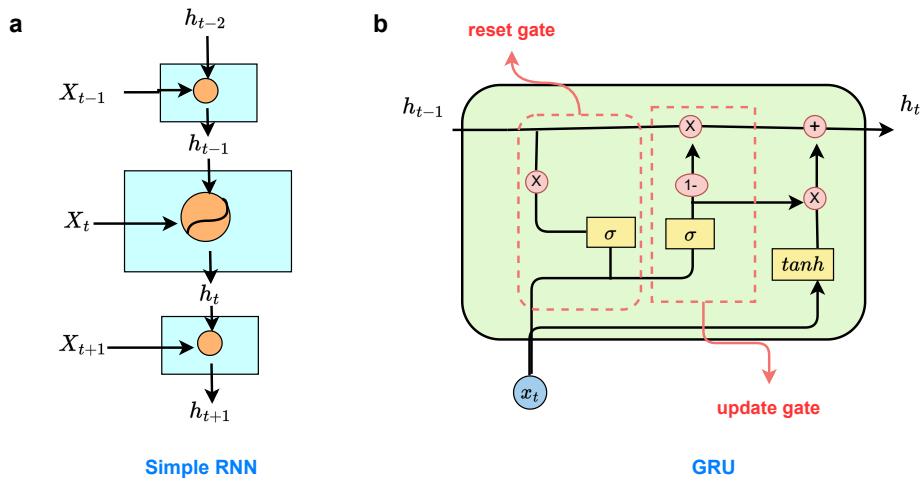


图 2-1 (a) RNN 示意图; (b) GRU 示意图

Figure 2-1 Schematic illustration of (a) RNN and (b) GRU

GRU 移除了单元状态的概念，直接从门控单元来计算隐藏状态，GRU 通过重置门和更新门的协同工作，实现了对信息流动的有效控制，使得网络能够在需要时记住长期依赖关系，同时减少梯度消失/爆炸的问题。GRU 的结构比 LSTM 更简单，计算效率更高，因此在实践应用中广受欢迎。本章使用 GRU 模型进行流变学本构方程的预测建模。

2.2.3 其他时间序列算法

在深度学习领域，针对时间序列数据的处理方法已呈现出多样化的趋势。除了传统的循环神经网络（RNN）及其变体，近年来研究人员提出多种新型网络架构，包括卷积神经网络（CNN）、Transformer 以及 Mamba 网络等，这些方法在时间序列分析中展现出显著的优势。

卷积神经网络（CNN）最初是为图像处理任务设计的，但其在时间序列分析中的应用也取得了显著成效^[117]。通过引入一维卷积操作，CNN 能够有效地从时间序列数据中提取局部特征。由于卷积操作具有权值共享的特性，CNN 在处理长序列数据时表现出较高的计算效率，并且能够有效规避传统 RNN 中常见的梯度消失或梯度爆炸问题。通过多层卷积结构的堆叠，CNN 能够捕获更为复杂的时间依赖性，进而在语音识别、金融预测等实际应用中表现出优异的性能^[118]。

Transformer 模型是近年来在自然语言处理（NLP）领域取得突破性进展的架构，其核心在于自注意力机制（Self-Attention）^[119]。与传统的 RNN 类模型不同，Transformer

摒弃了序列顺序处理的限制，转而通过全局上下文信息建模序列中各元素之间的依赖关系。这种机制使得 Transformer 能够并行处理长序列数据，并捕获全局时序特征。在时间序列分析任务中，Transformer 通过自注意力机制能够聚焦于序列中的关键时间点，从而有效建模时序依赖性和非线性关系。这一特性使其在机器翻译、语音识别等任务中超越了传统模型的表现。

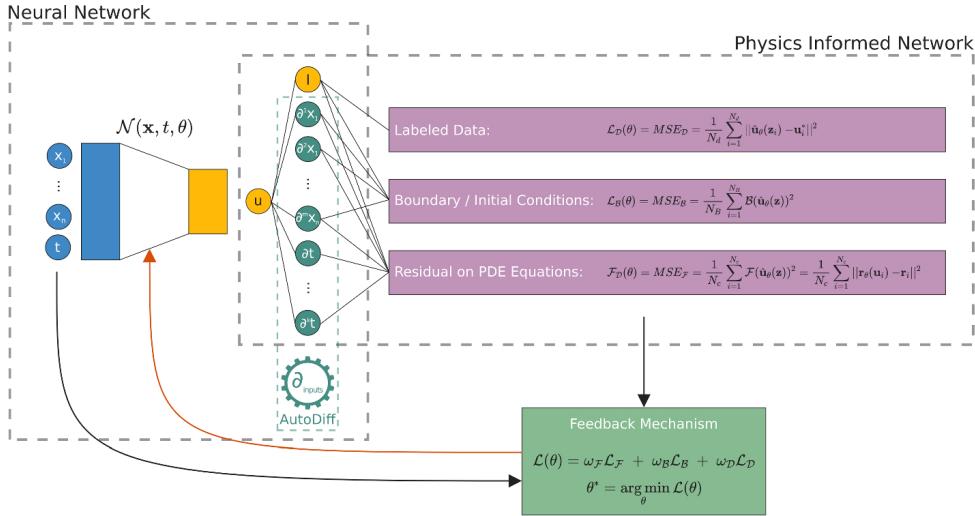
Mamba 网络是近年来提出的一种新型时间序列处理架构，专门针对复杂时间序列数据的高效建模而设计^[120]。与传统的 RNN 或 CNN 模型相比，Mamba 网络通过引入多尺度时序建模策略，能够同时捕获时间序列中不同时间尺度的特征。这种多尺度建模方法不仅增强了模型的表达能力，还提升了其在处理多变且复杂时序数据时的鲁棒性。Mamba 网络的独特之处在于其能够整合多层次时序信息进行联合建模，从而为不同时间粒度的变化提供更为精确的预测。这一特性使其在金融市场分析、气象预测等领域展现出广阔的应用前景。

这些方法不仅拓展了时间序列建模的理论边界，也为实际应用提供了更为强大的工具。本文的研究工作是选择一种可以处理时间序列的算法来对数值模拟的流变学数据进行建模预测，样本量在万级，属于中小规模数据集，综合考虑算法时空间复杂度和训练成本，选择 RNN 中的 GRU 作为算法模型。未来针对更多流变学数据和复杂场景，可以进一步研究其他前沿模型的应用。

2.3 物理信息神经网络

2.3.1 理论基础

物理信息神经网络（Physics-Informed Neural Networks, PINN）是一种结合深度学习和物理知识的神经网络模型，旨在通过数据驱动的方式求解偏微分方程（PDE）^[121]。与传统数据驱动的神经网络模型相比，PINN 在训练过程中引入了物理约束，使得模型不仅能够学习数据中的模式，还能够满足物理定律，从而提高模型的泛化能力和预测精度。PINN 的基本思想是将物理定律（如偏微分方程）作为先验知识嵌入到神经网络中，通过最小化数据损失和物理损失来训练模型。这样，PINN 不仅能够学习数据中的模式，还能够满足物理定律，从而提高模型的泛化能力和预测精度。在实际应用中，PINN 可以通过损失函数强制模型满足物理约束，从而提高模型的物理合理性和可解释性。PINN 的优势在于其能够通过物理方程进行训练，极大减少了对大量标注数据的需求。此外，PINN 能够确保模型的预测符合实际的物理规律，从而避免了数据驱动模型可能出现的

图 2-2 PINN 示意图^[122]Figure 2-2 Schematic illustration of PINN^[122]

不合理结果。尽管 PINN 在处理复杂问题时可能会面临训练难度和求解精度的挑战。

2.3.2 损失函数构建

PINN 的核心原理在于其损失函数的构建，普通深度神经网络（DNN）的损失函数为 $\mathcal{L} = \mathcal{L}_{data}$ ，即预测数据与真实数据之间的误差损失。而 PINN 的损失函数如公式(2-11)所示，由两部分组成。其中， \mathcal{L}_{data} 为数据损失， $\mathcal{L}_{physics}$ 为物理损失。

$$\mathcal{L} = \mathcal{L}_{data} + \lambda \mathcal{L}_{physics} \quad (2-11)$$

数据损失用于衡量神经网络预测值与实际观测数据之间的差异。常用的度量方法是均方误差（Mean Squared Error，MSE），如公式(2-12)所示。其中， N_d 是观测数据点的数量， $u(x_i, t_i)$ 是神经网络在点 (x_i, t_i) 处的预测值， u_i 是对应的真实观测值。

$$\mathcal{L}_{data} = \frac{1}{N_d} \sum_{i=1}^{N_d} (u(x_i, t_i) - u_i)^2 \quad (2-12)$$

物理损失用于确保神经网络的预测结果满足物理定律，通常通过计算物理方程的残差来实现。以纳维-斯托克斯方程为例，其物理损失可以表示为公式(2-13)。可以看到物理损失本质是神经网络的预测值在物理方程中的残差。

$$\mathcal{L}_{physics} = \frac{1}{N_p} \sum_{j=1}^{N_p} \left(\frac{\partial u}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} - \nu \nabla^2 \mathbf{u} - \nabla p + \mathbf{f} \right)^2 \Big|_{(x_j, t_j)} \quad (2-13)$$

2.3.3 未来技术优化

公式(2-11)中的 λ 参数用于控制物理损失对数据损失之间的权重关系。当 λ 较大时，物理损失对数据损失有较大的影响，反之亦然，最开始的 PINN 研究不考虑权重关系，这使得最终的损失函数受数据量的影响较大。如果采用同一份实验数据，分别计算物理损失和数据损失，那么影响还不会太大，但是如果按照多保重神经网络的思路，使用高保真数据来计算数据损失，低保真数据作为物理残差损失，那么在低保真数据远远大于高保真的情况下，权重参数需要谨慎设置，以防止梯度消失问题^[123]。对于 λ 的取值，我们可以针对实际应用进行调参，以获得更好的结果，但是这样的调参缺乏科学解释，且耗费时间，所以近年来，一些学者开始研究如何自动化确定 λ 的取值，以获得更好的结果。Farmer 等人提出了一种经验性的损失权重优化方法，用于 PINN 模型在求解激光生物效应中的 1D 热方程^[124]。该方法通过自动归一化损失函数的各个部分，确保不同损失项之间的平衡。Xiang 等人提出了一种自适应损失平衡方法（lbPINN），通过高斯概率模型定义自适应损失函数^[125]。该方法在每个训练周期中自动更新损失项的权重，基于最大似然估计来平衡不同损失项的影响。实验结果表明，lbPINN 在求解多个方程时，均表现出比传统 PINN 更好的性能。Song 等人提出了一种基于损失注意力的 PINN 架构（LA-PINN），为每个损失项配备独立的损失注意力网络（LAN）^[126]。该方法通过将每个训练点的平方误差（SE）输入到 LAN 中，动态地为不同点的 SE 分配不同的权重。实验结果表明，LA-PINN 在求解多个基准 PDE 时，表现出比传统 PINN 更高的预测精度和更快的收敛速度。本文的研究工作也借鉴了 Song 的方案，做了简化处理，实现了可学习权重的 PINN 框架进行训练。

在面对更高维度和更为复杂的偏微分方程（PDE）时，PINN 计算成本依然是一个待解决的挑战。为了有效应对大规模问题，未来的研究需要着重提升计算效率^[127]。此外，PINN 在处理小数据集时的表现仍有待提高，这要求我们进一步开展研究以降低其对大量训练数据的依赖。自适应采样和数据增强方法有望在这一领域发挥关键作用。将复杂的物理约束有效地整合到 PINN 中也是当前面临的一大挑战。未来的研究需要开发更加灵活且强大的方法，以应对复杂的物理方程和边界条件。在多物理场耦合问题中，PINN 的应用范围仍然较为有限^[128-129]。因此，未来的研究需要探索如何将 PINN 扩展至多物理场耦合问题，从而解决更为复杂的实际问题。

2.4 特征融合方法

2.4.1 简单特征融合

特征融合（Feature Fusion）是深度学习中一种重要的技术，旨在将来自不同来源或不同层次的特征进行组合，以创建一个捕获集体信息的统一表示。这种技术通过利用来自不同特征集的互补信息，能够显著增强模型的性能。

简单的特征融合方法包括 Concat、Add 和 Hadamard Product 等。其中，Concat 是将两个特征向量拼接在一起，如公式(2-14)，适用于需要保留两个特征向量的所有信息的场景。例如，在流体力学中，当需要同时考虑不同物理场（如速度场和压力场）或不同尺度的特征时，Concat 是一种合适的方法。Add 是将两个特征向量逐元素相加，如公式(2-15)，适用于两个特征向量维度相同，且需要强调某些共同特征时。例如，需要将速度场的不同分量相加，可以突出重要的流动特征的时候，可以使用 Add 方法。Hadamard Product 是将两个特征向量逐元素相乘，如公式(2-16)。当需要突出共同出现的特征并减弱不重要的特征时，Hadamard Product 是一种合适的方法。例如，在流体力学中，如果需要将速度场和压力场的特征相乘，希望突出共同的流动特征时，在材料科学中，如果需要将材料编码和组分含量相乘，突出加权的含量特征时，都可以使用 Hadamard Product。

$$\mathbf{v} = [\mathbf{v}_1; \mathbf{v}_2] \quad (2-14)$$

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 \quad (2-15)$$

$$\mathbf{v} = \mathbf{v}_1 \odot \mathbf{v}_2 \quad (2-16)$$

2.4.2 注意力特征融合

深度学习中，注意力特征融合（AFF）是一种重要的技术，它通过引入注意力机制来动态调整特征的权重，从而更好地融合特征^[130]。注意力特征融合的核心思想是利用注意力机制来捕获特征之间的关系，从而增强重要特征并减弱不重要的特征。注意力特征融合的数学公式如公式(2-17)所示，其中 \mathbf{v}_i 为特征向量， α_i 为注意力权重， \mathbf{q} 为查询向量，用于计算注意力权重。

$$\begin{aligned} \mathbf{v} &= \sum_{i=1}^n \alpha_i \mathbf{v}_i \\ \alpha_i &= \frac{\exp(\mathbf{q}^T \mathbf{v}_i)}{\sum_{j=1}^n \exp(\mathbf{q}^T \mathbf{v}_j)} \end{aligned} \quad (2-17)$$

在多尺度特征融合中，注意力机制的应用极大地提升了模型对不同尺度特征的动态权重调整能力，从而更有效地捕获全局和局部信息。这种方法在目标检测和图像分割等任务中表现出色，因为它能够更好地处理不同尺度的目标。例如，CM-UNet 模型通过多尺度注意力聚合模块（MSAA），在遥感图像语义分割任务中高效捕捉局部和全局信息，提升了特征表达能力^[131]。

在实现注意力特征融合时，通常会使用注意力模块，如自注意力（Self-Attention）或交叉注意力（Cross-Attention）。自注意力机制允许模型在同一个序列内部捕获特征之间的关系，而交叉注意力机制则允许模型在两个不同的序列之间捕获特征之间的关系。例如，交叉注意力可以用于将图像特征与文本特征进行融合，从而实现更准确的图像描述生成。在多模态学习中，交叉注意力机制通过在不同模块之间引入注意力机制，让信息交流更高效，也让模型在处理复杂任务时表现得更出色^[132]。

此外，注意力特征融合还可以与其他特征融合方法（如 Concat 和 Add）结合使用。例如，在某些模型中，可以先使用 Concat 将不同来源的特征拼接在一起，然后使用注意力机制对拼接后的特征进行加权，从而进一步增强特征的表示能力。这种方法在多模态学习中特别有用，因为它可以有效地融合来自不同模态的特征。例如，多模态融合网络使用多头自注意力机制来最小化不同模态之间的噪声干扰，并利用局部区域特征表示之间的相关性来提取互补信息^[133]。

总结注意力机制在多尺度特征融合和多模态学习中具有广泛的应用前景，能够显著提升模型的性能和泛化能力。通过合理设计和应用注意力模块，可以更好地捕获和融合不同尺度和模态的特征，从而在各种任务中取得更好的效果。而针对本文流变学的本构建模任务，由于制备参数（分子量、组分比例等）特征存在隐藏联系，但是特征本身非常稀疏，所以采取注意力特征融合的方法，期待显著提高模型的性能和泛化能力。

2.5 生成式模型

2.5.1 变分自编码器

生成式模型是一类能够学习数据生成过程的统计模型。它们通过建模数据的联合概率分布，能够生成与数据相似的新样本。与别判别式模型（关注于建模条件概率分布）不同，生成式模型关注数据的整体结构和分配。

变分自编码器是一种生成模型，结合了自编码器架构和变分推断方法。其核心思想是通过编码器将输入数据映射到潜在变量的分布参数（通常是均值和方差），然后通过

解码器从这个分布中采样并重构输入数据变分自编码器（Variational Autoencoder, VAE）的概念最早由 Diederik 提出。VAE 的数学表达式如公式(2-18)所示。其中编码器将输入数据映射到潜在变量空间，解码器将潜在变量映射回原始输入空间。VAE 的目标是最小化重构损失和变分下界的差异^[134]。

$$\mathcal{L}(\theta_E, \theta_D) = \mathbb{E}_{q_{\theta_E}(z|x)} [\log p_{\theta_D}(x|z)] - D_{KL}(q_{\theta_E}(z|x)\|p(z)) \quad (2-18)$$

重构损失 $L_{rec} = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|x_i - \hat{x}_i\|^2$ 用于衡量模型输出的重构精度，通常采用均方误差（MSE）或交叉熵等损失函数，变分下界 $L_{vae} = \mathbb{E}_{q_{\theta_E}(z|x)} [\log p_{\theta_D}(x|z)] - D_{KL}(q_{\theta_E}(z|x)\|p(z))$ 用于最大化潜在变量分布的重构似然。为了使 VAE 的训练过程可微分，引入了重参数化技巧。具体来说，假设潜在变量 z 服从均值为 μ 、方差为 σ 的高斯分布，可以通过公式 $z = \mu + \sigma \cdot \epsilon$ 从分布中采样， ϵ 是标准正态分布的随机变量。

变分自编码器（VAE）能够生成与训练数据类似的样本，例如人脸、文字等。在标签数据稀缺的情况下，VAE 可以利用无标签数据学习数据的潜在结构，从而增强模型的泛化能力。凭借这一特性，VAE 在图像生成、数据增强、无监督学习等多个领域展现出广泛的应用潜力，已成为生成模型领域的重要研究方向。

2.5.2 条件变分自编码器

条件变分自编码器（Conditional Variational Autoencoder, CVAE）的概念最早由 Sohn 等人提出^[135]。CVAE 是 VAE 的扩展，它在 VAE 的基础之上引入了条件变量，使得生成的样本可以根据特定条件进行控制。CVAE 的核心思想是将条件变量 y 作为输入的一部分，与输入数据 x 一起编码到潜在变量 z 中，然后通过解码器生成样本。CVAE 的目标是最大化条件似然 $p(x|y) = \int p(x|z, y)p(z|y)dz$ 。

$$\mathcal{L}(\theta_E, \theta_D) = \mathbb{E}_{q_{\theta_E}(z|x,y)} [\log p_{\theta_D}(x|z,y)] - D_{KL}(q_{\theta_E}(z|x,y)\|p(z|y)) \quad (2-19)$$

其中，第一项是重构损失，用于衡量模型输出与真实数据的差异；第二项是 KL 散度，用于衡量后验分布与先验分布之间的差异。通过最小化这一损失函数，CVAE 能够学习到数据的生成分布，并生成与训练数据类似的样本。CVAE 引入条件变量 y 使得模型能够根据特定条件生成样本，这对于许多实际应用非常重要。例如，在图像生成任务中，条件变量可以控制生成图像的类别或属性。通过指定不同的条件变量，CVAE 可以生成不同类别的图像，如不同类型的动物、不同的场景等。这使得 CVAE 在图像生成领域具有广泛的应用前景。

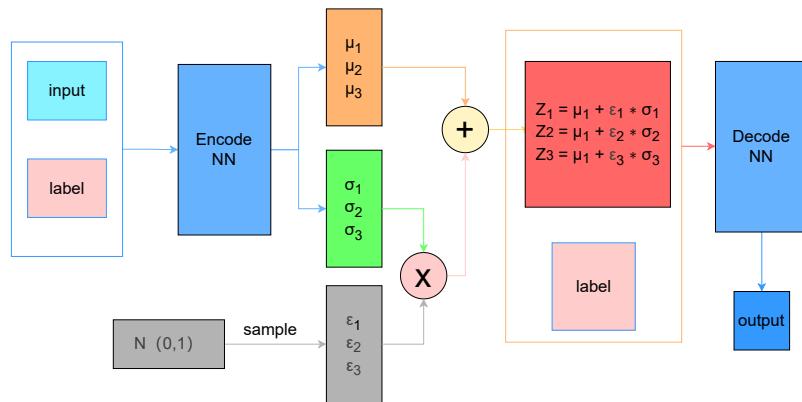


图 2-3 CVAE 示意图

Figure 2-3 Schematic illustration of CVAE

在自然语言处理任务中，条件变量可以控制生成文本的主题或风格。通过指定不同的条件变量，CVAE 可以生成不同主题的文本，如新闻报道、故事、诗歌等。此外，条件变量还可以控制文本的风格，如正式、幽默、悲伤等。这使得 CVAE 在自然语言生成领域具有重要的应用价值。

此外，CVAE 还可以在标签数据稀缺的情况下，利用无标签数据学习数据的潜在结构，从而增强模型的泛化能力。在实际应用中，标签数据往往非常稀缺，而无标签数据相对丰富。CVAE 可以利用无标签数据学习数据的潜在结构，从而提高模型的性能。这使得 CVAE 在半监督学习和无监督学习中具有重要的应用价值。

在本文的工作中，我们希望通过已有的流变学性质参数，如特定频率下的储存模量 (G')、损耗模量 (G'') 和损耗角正切 ($\tan\delta$) 等，来预测特定的制备参数。CVAE 能够将输入数据映射到高维潜在空间，并通过将流变学参数作为条件 (y) 输入，生成特定的制备参数。

与传统的深度神经网络 (DNN)，尤其是回归或分类网络不同，DNN 通常直接进行输入到输出的映射，缺乏对潜在空间的建模，因此可能无法捕捉到数据背后复杂的分布。相比之下，CVAE 作为生成模型，能够学习输入数据的潜在分布。它不仅依赖于输入数据的特征，还引入了条件信息，使得它能够在潜在空间中生成符合条件的参数，而不仅仅是进行简单的映射。

此外，CVAE 通过变分推断捕获数据中的不确定性，能生成多个样本。对于制备参数，可能存在多个合理的组合或生成路径，CVAE 通过潜在变量的不同采样来生成这些

多样化的样本，从而增强了生成结果的多样性。而传统的 DNN 模型通常是确定性的，即给定相同的输入总是生成相同的输出，无法有效地表示这种不确定性。

最后，CVAE 通过引入变分推断，平衡了重建误差和潜在变量的 KL 散度，从而在优化过程中不仅考虑了生成结果的质量，还保证了潜在空间的结构化。这使得 CVAE 能够生成接近原始数据的样本，并通过潜在空间的结构化生成有意义的样本。相比之下，DNN 训练时通常只关注误差最小化，缺少对潜在空间结构的显式建模。

2.5.3 其他生成式模型

生成式模型（Generative Models）是机器学习领域的重要研究方向，旨在学习数据的分布特征，从而生成与原始数据相似的新样本。除了变分自编码器（VAE）和条件变分自编码器（CVAE）之外，还有多种生成式模型在不同领域取得了显著成果。

自回归模型基于条件概率链式分解对数据进行建模，通过逐步生成数据单元完成整体构建^[136]。这类模型在密度估计方面展现出卓越性能，典型代表如 PixelCNN 和 PixelSNAIL 已成功应用于图像生成、语音合成等领域。然而其固有特性也带来若干限制：采样过程必须严格遵循序列生成路径，导致高维数据生成效率显著降低；同时模型强制要求将输入数据线性化为固定顺序，这在文本和音频等具有自然时序结构的模态中尚可适用，但对于图像等空间数据而言，最优排列方式的确定缺乏明确依据，且不同的顺序选择可能通过神经网络架构的归纳偏置对最终效果产生潜在影响^[137]。

生成对抗网络（Generative Adversarial Networks, GAN）是一种基于博弈论框架的生成模型，其核心由生成器（Generator）与判别器（Discriminator）构成动态博弈系统^[138]。生成器通过参数化映射函数将潜在空间向量转化为合成数据，旨在捕捉真实数据分布的统计特性；判别器则作为二元分类器，通过迭代优化提升对真实数据与生成数据的鉴别能力。二者在对抗性训练过程中形成“最小-最大”博弈关系：生成器试图生成以假乱真的样本来欺骗判别器，而判别器则持续升级其辨别能力以识别生成样本的统计缺陷，最终推动系统向纳什均衡收敛。相较于传统生成模型，GAN 的突出优势在于其能通过对抗机制隐式学习复杂数据分布，生成具有高度视觉保真度的样本（如图像、视频）^[139]。

扩散模型是新一代生成式人工智能的核心范式，其创新性地将数据生成过程建模为物理学启发的渐进式去噪机制^[140]。该模型通过构建马尔可夫链，系统性地模拟两个互逆过程：前向扩散阶段将原始数据通过逐步添加高斯噪声退化为随机噪声，反向生成阶段则通过参数化的神经网络学习逆向扩散轨迹，从纯噪声出发通过连续的去噪操作重建出目标数据分布。这种基于随机微分方程（SDE）或概率流常微分方程（ODE）的数学

框架，使得模型能够通过变分推断精确优化对数似然下界^[141]。相较于 GAN，其生成过程具有可解释的物理意义，通过调节去噪步长可实现生成质量与速度的灵活平衡；无需对抗训练避免了模式崩溃风险，确保生成样本的多样性，且理论框架的严密性支持精确的概率密度估计^[142]。

2.6 本章小结

本章围绕深度学习与物理建模的融合框架展开系统性论述，为流变学本构方程的数据驱动建模奠定了坚实的理论基础与技术体系。

在时序建模层面，系统解析了循环神经网络及其衍生架构的时序建模机理，重点探讨了门控循环单元（GRU）在流变学本构时序响应建模中的适用性与局限性，阐明其在处理材料非线性记忆效应（如应力松弛、触变恢复）中的优势及梯度稳定性问题。针对物理约束嵌入方法，深入剖析了物理信息神经网络（PINN）的数学基础与实现范式，为后续构建具有物理一致性的流变学代理模型提供理论支撑。在特征融合方面，通过对比回归多尺度数据整合策略，揭示了注意力机制在动态协调流变实验多源异构数据中的核心价值，为解决复杂流变行为的特征解耦问题开辟新途径。针对生成式建模技术，综述了变分自编码器的概率框架与条件扩展形式。

本研究通过算法体系的多维度解析，为突破传统唯象模型的表征瓶颈、后续构建发展“数据-物理”双驱动的智能本构建模范式提供了理论架构与方法论指导。

第三章 物理信息门控循环单元流变学本构方程建模研究

3.1 引言

近年来，深度学习在流变学本构建模研究中得到广泛应用。例如 Lennon、Mahmoudabadbozchelou 等人的研究工作，虽然具体方法各有不同，但在模型选择上基本都采用了多层感知机模型（MLP）^[106,109]。MLP 是一种经典的前馈人工神经网络，由输入层、多个隐藏层及输出层的全连接层堆叠而成，通过非线性激活函数实现复杂函数逼近。当 MLP 的隐藏层数达到一定值时，它被视为深度神经网络（DNN）。传统的前馈性质 DNN 模型虽具备一定的非线性行为捕捉能力，但在处理具有时间依赖性的数据时，其性能往往受到限制。Lennon 和 Mahmoudabadbozchelou 的工作使用 DNN 模型，难以有效捕捉黏弹性材料中的长程应变历史依赖性。本章研究在前人基础上，采用物理信息门控循环单元（PI-GRU）进行本构建模，旨在解决 DNN 模型在处理时间序列数据时性能受限的问题。PI-GRU 模型创新性地结合了物理约束和 GRU 的循环神经网络架构，通过在损失函数中引入物理方程残差作为惩罚项，同时利用 GRU 的循环结构处理时序数据。该模型包含输入层感知机、GRU 循环层和输出层感知机三部分组成，输入为应变张量分量、应变速率张量分量和时间等参数，输出为预测的应力张量。

本章研究首先通过数值模拟方法构建了多种经典本构方程的应力应变数据，包括 Herschel-Bulkley 模型、Maxwell 模型、Doi-Edwards 模型。这些模型在流变学领域具有重要的理论和应用价值，能够描述不同类型的流变行为。随后，本章详细设计了 GRU 模型的结构、参数初始化及训练优化策略，并使用数值模拟生成的应力应变数据进行训练和参数优化，使模型能够准确拟合训练数据。为了验证 GRU 模型在探索未知流变学本构方程方面的能力，本章将训练好的 GRU 模型应用于未经训练的测试集数据，并通过与传统 DNN 等基准模型的对比，验证了 GRU 模型在处理时间序列数据方面的优势。在此基础上，本章进一步研究了更复杂的非线性本构关系，通过数值模拟方法生成 Giesekus 模型数据，并引入物理信息约束，构建 PI-GRU 模型进行建模。通过与其他基准模型的对比，验证了 PI-GRU 模型在复杂非线性本构关系建模中的有效性。最后，本章使用一类黏弹性聚合物的真实流变学实验数据进行 PI-GRU 训练和测试，进一步验证了该模型在真实工况下的适用性和优越性。

3.2 深度学习模型设计

本章针对 Herschel-Bulkley 模型、Maxwell 模型、Doi-Edwards 模型使用 GRU 模型进行建模，其基本架构在第二章已经详细介绍。本章针对 Giesekus 模型和真实实验数据使用的算法模型为 PI-GRU 模型，PI-GRU 与基本的 GRU 神经网络模型架构类似，但不同的是，PI-GRU 循环神经网络模型在 GRU 的基础上，加入了物理约束损失，用于约束模型的预测结果符合基本的本构物理规律，增强模型的泛化能力。其架构图如图3-1所示。

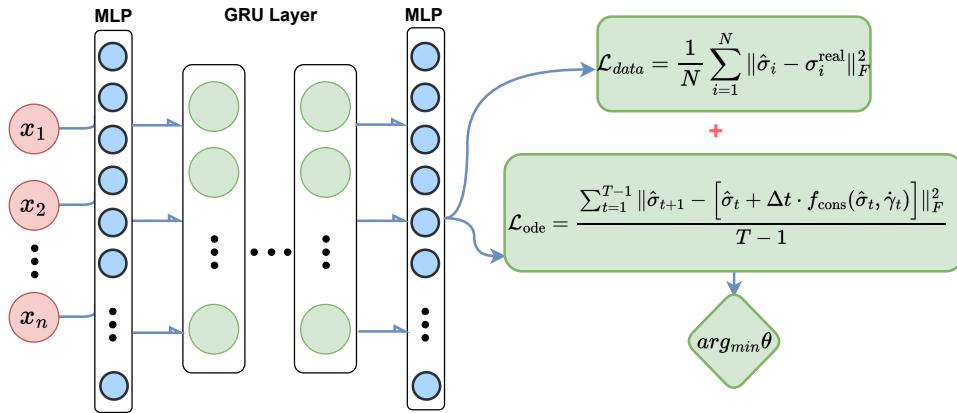


图 3-1 本工作使用的 PI-GRU 架构示意图

Figure 3-1 Schematic of the PI-GRU model used in this work

示，模型的输入为应变张量分量、应变速率张量分量、时间等，首先经过输入层的感知机进行特征提取，然后经过 GRU 层进行时序信息处理，最后经过输出层的感知机输出预测的应力张量。PI-GRU 的损失函数主体由数据损失和物理约束损失共同组成，具体公式如3-1所示。

$$\mathcal{L}_{\text{total}} = \lambda_{\text{data}} \mathcal{L}_{\text{data}} + \lambda_{\text{ode}} \mathcal{L}_{\text{ode}} + \lambda_{\text{dae}} \mathcal{L}_{\text{dae}} + \mathcal{L}_{L_2} \quad (3-1)$$

其中数据损失 $\mathcal{L}_{\text{data}}$ 是预测值与真实值的均方误差（公式3-2），目的是确保模型的预测结果与数值模拟数据尽可能接近，这是监督学习的基础，直接拟合数据，也是最重要和最基本的损失项。

$$\mathcal{L}_{\text{data}} = \frac{1}{N} \sum_{i=1}^N \|\hat{\boldsymbol{\sigma}}_i - \boldsymbol{\sigma}_i^{\text{real}}\|_F^2 \quad (3-2)$$

$$\mathcal{L}_{\text{ode}} = \frac{1}{T-1} \sum_{t=1}^{T-1} \|\hat{\boldsymbol{\sigma}}_{t+1} - [\hat{\boldsymbol{\sigma}}_t + \Delta t \cdot f_{\text{constitutive}}(\hat{\boldsymbol{\sigma}}_t, \dot{\gamma}_t)]\|_F^2 \quad (3-3)$$

物理损失分为 ODE 损失 \mathcal{L}_{ode} 和 DAE 损失 \mathcal{L}_{dae} 。 \mathcal{L}_{ode} 确保模型的预测遵循本构方程描述的物理规律，如公式3-3所示。该公式计算的是时间序列中相邻时间步之间的平均损失，其中 $\hat{\boldsymbol{\sigma}}_{t+1}$ 是模型在 $t+1$ 时刻的预测应力，而方括号内的表达式 $\hat{\boldsymbol{\sigma}}_t + \Delta t \cdot f_{\text{constitutive}}(\hat{\boldsymbol{\sigma}}_t, \dot{\gamma}_t)$ 是根据本构方程从 t 时刻推导出的 $t+1$ 时刻的理论应力值。这种损失函数通过比较模型预测值与基于物理定律推导值之间的差异，确保模型学习结果符合我们想要的本构关系。这里的本构方程在 Giesekus 模型模拟数值训练时为上对流 Maxwell 模型（UCM），在真实实验数据时为最小二乘法拟合的各类本构方程。

\mathcal{L}_{dae} 损失则是为了保证预测的应力张量满足基本的对称性要求。物理损失相当于一个独特的惩罚机制，迫使模型的预测结果约束在基本的框架内，一定程度上防止模型在纯数据驱动下出现过拟合现象。

$$\mathcal{L}_{\text{dae}} = \frac{1}{T} \sum_{t=1}^T \left(\underbrace{\|\hat{\boldsymbol{\sigma}}_t - \hat{\boldsymbol{\sigma}}_t^\top\|_F^2}_{\text{symmetry constraint}} \right) \quad (3-4)$$

$$\mathcal{L}_{L_2} = \lambda \sum_i w_i^2 \quad (3-5)$$

最后在损失函数中加入正则化项 \mathcal{L}_{L_2} ，具体公式如公式3-5。 \mathcal{L}_{L_2} 通过类似于归纳推理的机制，使得模型倾向于学习更通用的模式，参数更小更简单，对未知数据的预测更可靠。

3.3 实验设计

3.3.1 数据获取

3.3.1.1 Herschel-Bulkley 模型模拟数据

Herschel-Bulkley 模型的本构方程如公式(1-9)所示，其中剪切应力 σ 与剪切率 $\dot{\gamma}$ 之间存在函数关系。该模型包含流变参数 K 、流动指数 n 及屈服应力 σ_0 。在模拟过程中，本章设置 σ_0 为 1.0 Pa， K 为 1，而 n 则取值为 0.2、0.6、1.0、1.4 及 1.8。剪切率 $\dot{\gamma}$ 范围设定为 [0,100]，并以 0.01 的时间步长进行离散采样。

3.3.1.2 Maxwell 模型模拟数据

Maxwell 模型的微分本构方程如公式(1-3)所示。该模型包含松弛时间 $\tau = \eta/G$, 其中 η 表示黏性系数, G 为剪切模量。本章设置 η 为 $0.1 \text{ Pa} \cdot \text{s}$, G 为 1.0 Pa 。采用后向欧拉法来离散化微分方程, 具体推导如下:

首先将微分离散化, 设 $d\sigma = \sigma_i - \sigma_{i-1}$, $dt = \Delta t$, $d\gamma = \gamma_i - \gamma_{i-1}$, 则原方程可以化简为公式(3-6),

$$\frac{\sigma_i - \sigma_{i-1}}{\Delta t} + \frac{\sigma_i}{\tau} = G \frac{\gamma_i - \gamma_{i-1}}{\Delta t} \quad (3-6)$$

移项并化简可得公式(3-7)。

$$\sigma_i = \frac{\sigma_{i-1} + G(\gamma_i - \gamma_{i-1})}{1 + \frac{\Delta t}{\tau}} \quad (3-7)$$

根据公式(3-7), 本章首先使用 NumPy 库生成不同应变变化协议的应变数据, 时间步为 0.01 s , 每个协议模拟 2000 个数据点, 并存为 NumPy 数组, 随后通过迭代法计算单个应变数据对应的应力数据, 并存为 NumPy 数组。

3.3.1.3 Doi-Edwards 模型模拟数据

Doi-Edwards 模型的本构方程如公式(1-17)所示。该方程为积分形式的本构方程, 本章首先对该方程进行处理, 将取向张量函数 $\mathbf{Q}(t', t)$ 写为球坐标形式, 即公式(3-8)。

$$\mathbf{Q}(t', t) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi 5 \left(\frac{\mathbf{u}' \cdot \mathbf{F}^{-1} \mathbf{u}' \cdot \mathbf{F}^{-1}}{|\mathbf{u}' \cdot \mathbf{F}^{-1}|^2} \right) \sin \theta d\theta d\phi \quad (3-8)$$

$$G(t, t') = \frac{G_0}{\lambda_i} \exp \left(\frac{t' - t}{\lambda_i} \right) \quad (3-9)$$

$$\boldsymbol{\sigma}(t) = \int_{t_0}^t G(t, t') \cdot \mathbf{Q}(\gamma(t')) dt' \quad (3-10)$$

本章模拟的为简单剪切流动, 只在 xy 方向存在应变, 因此可以将逆变形梯度张量 \mathbf{F}^{-1} 写为公式(3-11)的矩阵形式, \mathbf{u}' 为球坐标系下的单位向量。

$$\mathbf{F}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \gamma & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3-11)$$

公式(3-9)为松弛模量函数, 本章使用 NumPy 库生成不同的应变协议的 NumPy 数组, 单个协议时间区间为 $[0, 4\pi]$, 总数据点为 2000 个, 生成形式为 3×3 的张量矩阵, 代入公

式(3-8)计算 \mathbf{Q} 值数组。设置 G_0 为 1.0 Pa, λ 为 1.0 s, $i = 1$, 根据公式(3-10)计算应变张量, 使用的积分工具为 Python 的 `scipy.integrate` 库, 生成形式为 3×3 的张量矩阵, 如公式(3-12)所示。本章提取 σ_{12} 、 σ_{11} 、 σ_{22} 分量作为模拟实验数据, 与对应的应变分量数据一起通过 Pandas 库存入 Excel 文件, 便于后续的模型训练。

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \quad (3-12)$$

3.3.2 Giesekus 模型模拟数据

Giesekus 模型的本构方程如公式(1-12)所示, 迁移因子 α 用于引入剪切稀化的强度。本节中 Giesekus 模型模拟的为简单剪切流动, 只在 xy 方向存在应变, 应变张量 γ 仅在 γ_{12} 分量上存在值。本章首先使用 NumPy 库生成不同速度场协议的简单剪切流动速度数据, 再使用公式(3-13)对应生成应变速率张量数据, 单个协议的模拟数据点为 2000, 生成形式为 3×3 应变速率张量矩阵, 之后根据 $\gamma_t = \dot{\gamma}_{t-1} \Delta t + \gamma_{t-1}$ 迭代计算应变张量。

$$\begin{aligned} \dot{\gamma}_{ii} &= 2v_{ii}(t) \\ \dot{\gamma}_{ij} &= v_{ij}(t) + v_{ji}(t) \end{aligned} \quad (3-13)$$

随后设置迁移因子 α 为 0.8, 其余松弛时间参数 (λ_1 、 λ_2) 为 1.0, 使用 Python 的 `scipy.integrate.solve_ivp` 函数对 Giesekus 模型的微分方程组进行求解计算, 生成 3×3 应力张量矩阵, 如公式(3-12)所示。本章提取 σ_{12} 、 σ_{11} 、 σ_{22} 分量作为模拟实验数据。

分别按照上述方法模拟小振幅振荡剪切 (SAOS) 数据和大振幅振荡剪切 (LAOS) 数据。

3.3.2.1 真实实验数据

首先通过原子转移自由基聚合 (ATRP) 和开环异位聚合 (ROMP) 设计合成了由聚甲基丙烯酸 (亲水单元) 和聚甲基丙烯酸叔丁酯 (疏水单元) 组成的两亲性刷状聚合物 (ABPs)。随后将 ABPs 与甲氧基丙烯酸酯按特定比例混合, 在紫外光照射下聚合并经过浸泡、干燥处理, 最终得到均匀的两亲性刷掺杂的弹性体 (ABE 弹性体)。

通过安东帕型流变仪在 25°C 下测量具有 0.8mm 厚度的材料的振荡剪切试验。将样品置于 8mm 直径的平行板下, 在 1% 到 300% 的应变范围内以 10rad/s 的固定频率对样

品进行应变扫描，获取 ABE 弹性体的 SAOS 和 LAOS 数据。

3.3.3 模型训练

3.3.3.1 数据集划分

首先本章对模拟生成的数据进行数据集划分，将数据集分为训练集、验证集和测试集。其中每种本构模型数据集和真实数据数据集分别选取一组待预测数据作为测试集，其余数据按照 9:1 比例划分为训练集和验证集。

3.3.3.2 训练细节

本研究使用 PyTorch 框架进行模型训练，使用 Adam 优化器进行优化。超参数的设置采用随机搜索算法，该算法通过在预定义的超参数空间中进行随机采样来寻找最优组合。具体而言，对于学习率 η ，在对数空间 $[10^{-5}, 10^{-2}]$ 内随机采样；batch size 在 16, 32, 64, 128 中随机选择；隐藏层神经元数量在 10, 20, 30, 40 范围内随机采样。随机搜索的优化目标函数可表示为：

$$\theta^* = \arg \min_{\theta \in \Theta} \mathcal{L}_{\text{val}}(f_\theta(X_{\text{val}}), y_{\text{val}}) \quad (3-14)$$

其中 Θ 表示超参数空间， \mathcal{L}_{val} 为验证集上的损失函数。

为了进一步提高训练效率，本文还采用了学习率自适应调整策略。具体使用带有预热 (warmup) 的余弦退火调度，学习率 η 随训练轮次 t 的变化可表示为：

$$\eta(t) = \begin{cases} \eta_{\max} \cdot \frac{t}{T_{\text{warmup}}} & \text{if } t \leq T_{\text{warmup}} \\ \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min})(1 + \cos(\frac{t-T_{\text{warmup}}}{T_{\text{total}}-T_{\text{warmup}}} \pi)) & \text{if } t > T_{\text{warmup}} \end{cases} \quad (3-15)$$

其中 T_{warmup} 为预热期轮次数， T_{total} 为总训练轮次， η_{\max} 和 η_{\min} 分别为最大和最小学习率。这种策略既能避免训练初期不稳定，又可以在后期实现更精细的参数调整。

3.3.4 模型测试

模型验证首先通过可视化方法直观比较预测结果与真实值。主要采用预测值-真实值曲线对比和残差分析图两种可视化方式。预测值-真实值曲线对比将模型预测的应力-应变曲线与真实数据在同一坐标系中绘制，直观展示模型预测的准确性。残差分析图则计算每个时间点的预测值与真实值之间的差异（残差），并绘制残差随时间或应变

的变化图，能够揭示模型预测中的系统性偏差。理想情况下，残差应随机分布在零线附近，不呈现明显的趋势或模式。

本章的模型训练均为回归问题，因此采用的测试指标包括决定系数（Coefficient of Determination, R²），如公式(3-16)所示，平均绝对误差（Mean Absolute Error, MAE），如公式(3-17)所示，以及平均百分比误差（Mean Absolute Percentage Error, MAPE），如公式(3-18)所示。

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3-16)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3-17)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3-18)$$

此外，加上训练时间作为训练成本指标，全面评估模型性能。

3.4 结果与讨论

3.4.1 Herschel-Bulkley 模型建模

本节分别使用 GRU 和 DNN 两种算法模型对 Herschel-Bulkley 模型模拟数据进行深度学习建模，之后使用预测模型在测试集上进行验证，测试结果如图3-2所示。图3-2 (a) 为两种算法测试的真实值-预测值曲线，从曲线可以定性看出两种算法的预测值曲线与真实值曲线都非常接近。图3-2 (b) 为两种算法测试结果的残差图，可以看出两种算法的残差点离散程度接近，均没有明显趋向性，均呈现无序分布，说明两种算法均可以比较好地捕捉到所有的输入特征。图3-2 (c-f) 分别比较了两种算法预测结果的 R², MAE, MAPE 指标，从结果中可以看出，两种算法的预测效果都很好。GRU 算法的平均训练时间为 378 s，高于 DNN 的 155 s，这是由于 GRU 网络的参数量更大，且由于其循环神经网络的特点，只能顺序运算，限制了 GPU 的并行计算能力，导致训练时间较长。

综合看来 GRU 和 DNN 两种算法在 Herschel-Bulkley 模型模拟数据上的预测表现比较接近，从预测指标的绝对数值看，GRU 略优于 DNN。这个结果符合预期，因为 Herschel-Bulkley 模型本质上是模拟了剪切稀化过程，不涉及黏弹性材料的时间依赖性，并且我们模拟的过程中，对于某个特定时间的应力状态也仅仅是当前应变状态的函数。

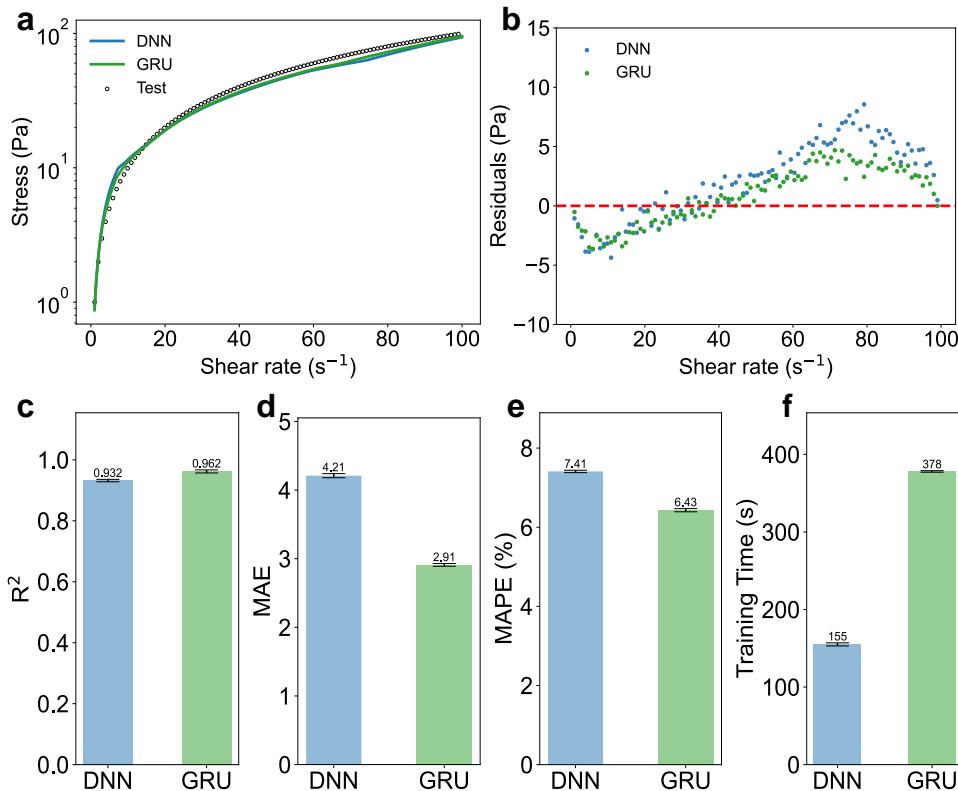


图 3-2 GRU 和 DNN 在 Herschel-Bulkley 模型测试集上的预测效果对比示意图: (a) GRU 和 DNN 在测试集上的预测值与真实值的应力-应变率曲线; (b) GRU 和 DNN 在测试集上的预测值残差图; (c-f) GRU 和 DNN 在测试集上的预测指标对比: (c) R^2 ; (d) MAE; (e) MAPE; (f) 训练时间

Figure 3-2 Comparison of prediction performance between GRU and DNN on Herschel-Bulkley model test set: (a) Stress-strain rate curves of predicted vs. true values; (b) Residual plots of predicted values; (c-f) Comparison of prediction metrics: (c) R^2 ; (d) MAE; (e) MAPE; (f) Training time

3.4.2 Maxwell 模型建模

3.4.2.1 振荡剪切预测振荡剪切效果验证

为了验证 GRU 模型在时间序列本构方程数据中的预测效果, 本节使用振荡剪切协议生成的数据作为训练集, 振荡剪切协议生成的数据作为测试集。分别使用了 GRU 和 DNN 进行训练, 并在测试集上进行验证, 测试结果如图3-3所示。

图3-3 (a-b) 展示了两种算法模型在测试集上的预测效果的定性分析曲线。从 Lissajous 曲线可见, GRU 预测值与真实值高度吻合, 而 DNN 预测值在大应变区域存在

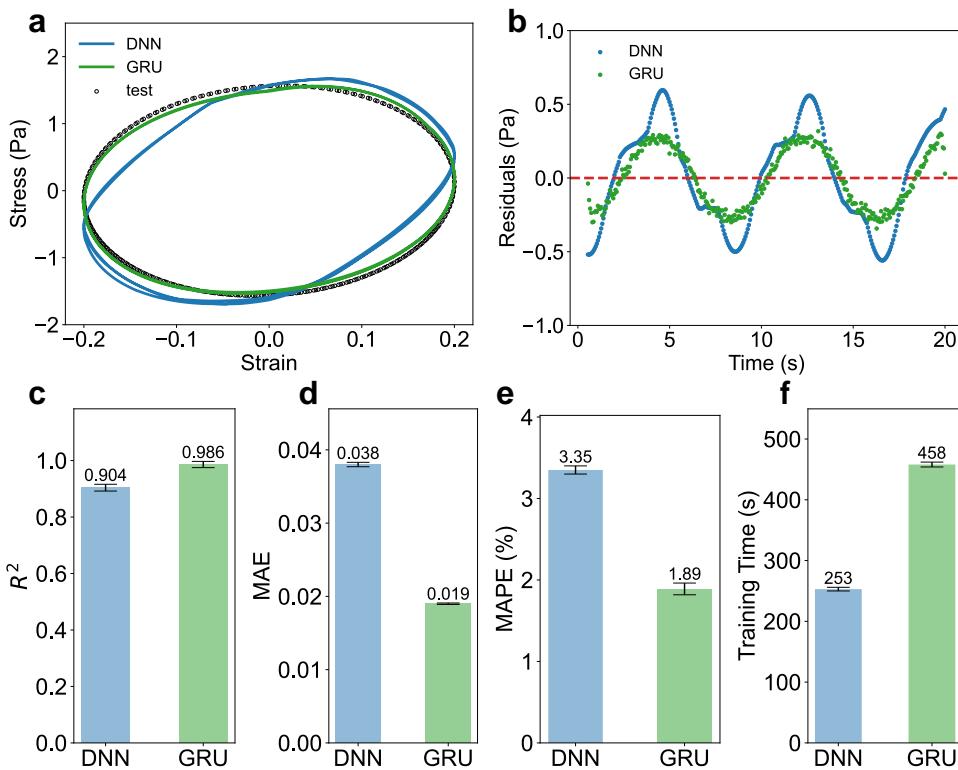


图 3-3 GRU 和 DNN 在 Maxwell 模型振荡协议测试集上的预测效果对比示意图: (a) GRU 和 DNN 在测试集上的预测值与真实值的应力-应变曲线 (Lissajous 曲线); (b) GRU 和 DNN 在测试集上的预测值残差图; (c-f) GRU 和 DNN 在测试集上的预测指标对比: (c) R^2 值; (d) MAE 值; (e) MAPE 值; (f) 训练时间

Figure 3-3 Comparison of prediction performance between GRU and DNN on Maxwell model oscillatory protocol test set: (a) Stress-strain curves (Lissajous curves) of predicted vs. true values; (b) Residual plots of predicted values; (c-f) Comparison of prediction metrics: (c) R^2 value; (d) MAE value; (e) MAPE value; (f) Training time

明显周期性偏差。从残差图可以看到，尽管 GRU 和 DNN 的残差图都存在周期性，但是 GRU 的残差图的周期性更小，且残差图的离散程度更小，说明 GRU 的预测效果更好。

图3-3 (c-f) 的定量指标进一步证实了 GRU 的优势，GRU 的 R^2 值更高，MAE 值和 MAPE 值更小，说明 GRU 的预测效果更好。当然，由于 GRU 的循环神经网络结构，GRU 的训练时间较 DNN 更长。

综合分析表明，对于同类型应变变化过程，GRU 算法能更好地学习 Maxwell 模型的内在特征，包括周期性响应、黏弹性和时间依赖性。在计算资源充足的情况下，GRU

算法在此类任务上明显优于 DNN。

3.4.2.2 振荡剪切预测稳态剪切效果验证

为了验证 GRU 算法模型在不同形式的应变历史下的泛化预测效果，本节使用振荡剪切应变协议生成的数据作为训练集，稳态剪切应变协议生成的数据作为测试集。分别使用了 GRU 和 DNN 进行训练，并在测试集上进行验证，测试结果如图3-4所示。

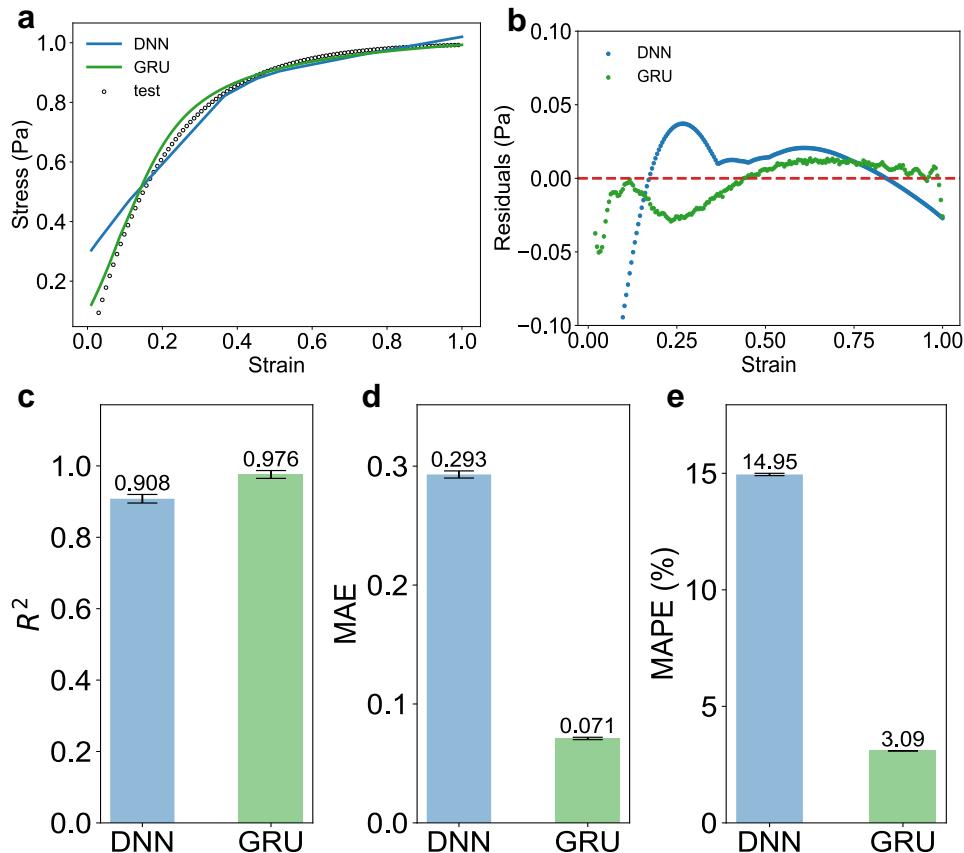


图 3-4 GRU 和 DNN 在 Maxwell 模型稳态剪切测试集上的预测效果对比示意图：(a) GRU 和 DNN 在测试集上的预测值与真实值的应力-应变曲线 (Lissajous 曲线); (b) GRU 和 DNN 在测试集上的预测值残差图; (c-e) GRU 和 DNN 在测试集上的预测指标对比：(c) R^2 值; (d) MAE 值; (e) MAPE 值

Figure 3-4 Comparison of prediction performance between GRU and DNN on Maxwell model steady-state shear test set: (a) Stress-strain curves (Lissajous curves) of predicted vs. true values; (b) Residual plots of predicted values; (c-e) Comparison of prediction metrics: (c) R^2 value; (d) MAE value; (e) MAPE value

图3-4 (a) 展示了 GRU 和 DNN 在 Maxwell 模型稳态剪切测试集上的 Lissajous 曲线预测效果。可以看出 GRU 的预测曲线与真实值更为接近，仅在小应变区域存在轻微偏差，这可能与门控单元对初始状态的敏感性有关。相比之下，DNN 的预测曲线偏差更大，尤其在小应变区域表现不佳。图3-4 (b) 的残差分析进一步证实了这一点，GRU 的残差点分布更加均匀地围绕零线，而 DNN 的残差在初始阶段明显偏离零线，整体分布效果不如 GRU。

图3-4 (c-e) 的定量指标也体现出同样的结论。GRU 的 R^2 值高于 DNN，而 MAE 值和 MAPE 值则远低于 DNN，说明 GRU 的预测效果更好。这些数据清晰地表明 GRU 在预测精度上具有显著优势，尽管其训练时间较长。

综合分析表明，当训练数据与测试数据为不同类型应变变化过程时，GRU 算法表现出明显优势。这主要归功于 GRU 能够有效捕捉 Maxwell 模型中材料响应的时间依赖性特征，即使仅通过学习振荡剪切应变数据，也能较为准确预测稳态剪切应变下的材料行为。

3.4.2.3 不同时间步的预测效果对比

对于 GRU 这类循环神经网络来说，网络架构的时间步数（序列长度）是非常重要的超参数。较大的时间步可以使模型在每个时间步上处理更多的信息，有助于捕捉长期依赖关系，但是可能导致模型学习到数据中的噪声，而不是其潜在的结构，从而导致过拟合。较小的时间步可以更细致地捕捉序列中的短期变化和细节信息，有助于模型更好地理解数据的短期动态特征，但是可能导致模型过于关注噪声，而忽略重要的长期依赖关系，从而导致欠拟合。本节研究了不同时间步下训练的模型在测试集上的预测效果。如图3-5所示。由图可见，预测模型 R^2 值随着时间步的增加而增加，但是幅度较小。MAE 值和 MAPE 值随时间步的增加而减少，说明随着时间步的增加，模型的预测效果变的更好，时间步小于 10 时，这种误差减小的趋势较明显，时间步大于 20 时，MAE 值下降幅度变小，时间步大于 10，MAPE 值下降幅度变小。这说明模型的优化指标上升存在阈值，这是因为 GRU 算法在较长序列中容易遗忘丢失信息，对于时间序列的处理长度有一定限制。而随着时间步的增加，当时间步大于 20 后，模型的训练时间急剧增加，训练成本陡增。

本节分析针对此项任务，最佳时间步在 10-20 之间，MAE 值和 MAPE 值开始下降到阈值，且训练时间还未显著增加。

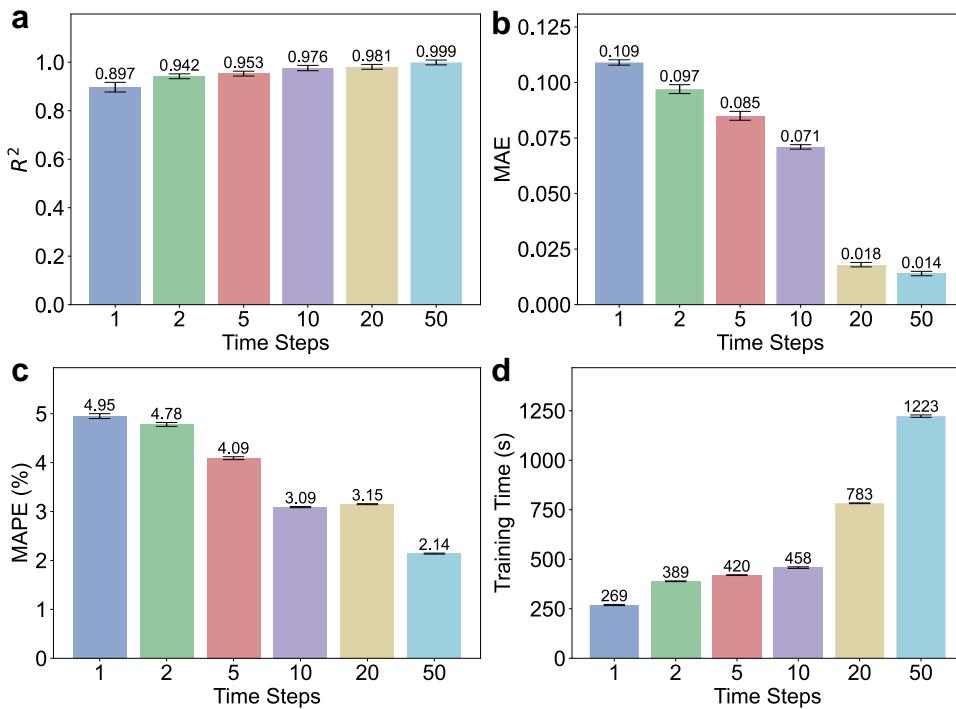


图 3-5 GRU 算法在 Maxwell 模型不同时间步下的预测效果对比图: (a) R^2 值; (b) MAE 值; (c) MAPE 值; (d) 训练时间

Figure 3-5 Comparison of prediction performance of the GRU algorithm on the Maxwell model with different time steps: (a) R^2 value; (b) MAE value; (c) MAPE value; (d) Training time

3.4.3 Doi-Edwards 模型建模

3.4.3.1 振荡剪切预测振荡剪切

上一节验证了 GRU 在简单 Maxwell 模型本构方程数据中的预测效果, 为了进一步探究 GRU 的循环神经网络结构在复杂本构模型中的预测效果, 本节使用更为复杂的 Doi-Edwards 模型进行验证。Doi-Edwards 模型是描述黏弹性材料的一种经典模型, 不仅可以描述剪切应力-应变关系, 还可以描述第一法向应力差 N_1 , 更加符合实际的材料行为。

为了验证 GRU 在 Doi-Edwards 模型本构方程数据中的预测效果, 本节使用振荡剪切协议生成的数据作为训练集, 振荡剪切协议生成的数据作为测试集。分别使用了 GRU 和 DNN 进行训练, 并在测试集上进行验证, 测试结果如图3-6、图3-7所示。

图3-6 (a-c) 为两种不同算法构建的预测模型预测的剪切应力 σ_{12} 的 Lissajous 曲线、

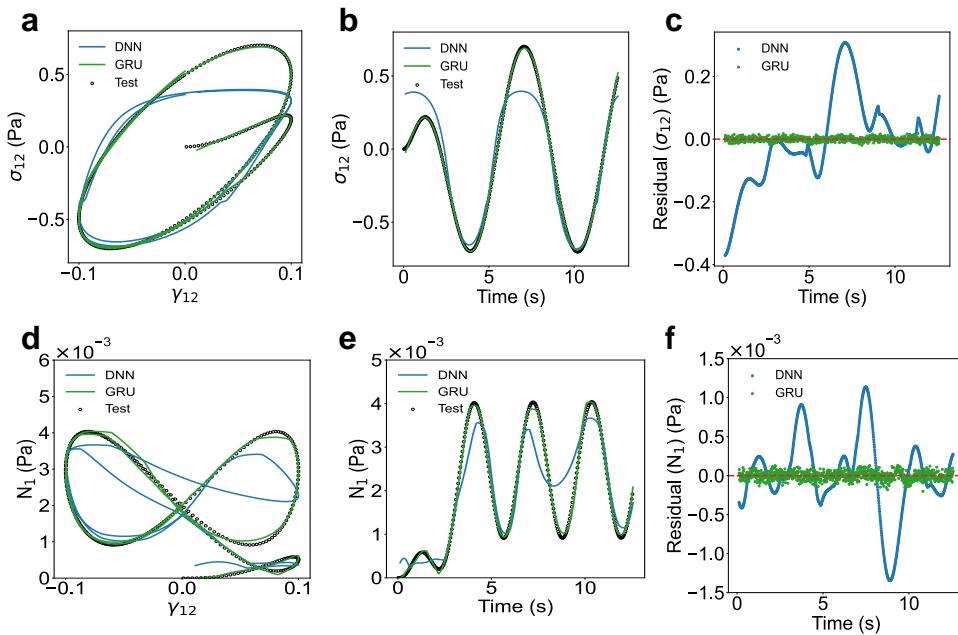


图 3-6 GRU 和 DNN 在 Doi-Edwards 模型振荡剪切测试集上的预测效果对比示意图: (a) GRU 和 DNN 在测试集上的预测值与真实值的剪切应力-应变曲线 (Lissajous 曲线); (b) GRU 和 DNN 在测试集上的预测值与真实值的时间-应力曲线; (c) GRU 和 DNN 在测试集上剪切应力的预测值残差图; (d) GRU 和 DNN 在测试集上的预测值与真实值的第一法向应力差 (N_1) 的 Lissajous 曲线; (e) GRU 和 DNN 在测试集上的预测值与真实值的时间-第一法向应力差曲线; (f) GRU 和 DNN 在测试集上的第一法向应力差预测值残差图

Figure 3-6 Comparison of prediction performance between GRU and DNN on Doi-Edwards model oscillatory shear test set: (a) Shear stress-strain curves (Lissajous curves) of predicted vs. true values; (b) Time-stress curves of predicted vs. true values; (c) Residual plots of shear stress predicted values; (d) First normal stress difference (N_1) Lissajous curves of predicted vs. true values; (e) Time-first normal stress difference curves of predicted vs. true values; (f) Residual plots of first normal stress difference predicted values

时间-应力曲线和残差图, (d-f) 为预测的第一法向应力差 (N_1) 的 Lissajous 曲线、时间-应力曲线、残差图。从图 a 和图 b 可以看出 GRU 预测的 σ_{12} 值与真实的 σ_{12} 接近, 曲线拟合较好, 而 DNN 算法的预测值在某些区域存在较大误差。图 c 的残差图显示, GRU 算法预测值与真实值的残差紧贴零刻度线, 呈现无序离散的分布状态, 而 DNN 算法预测值和真实值的残差呈现明显的曲线规律, 且与零刻度线性距离偏差较大, 残差分

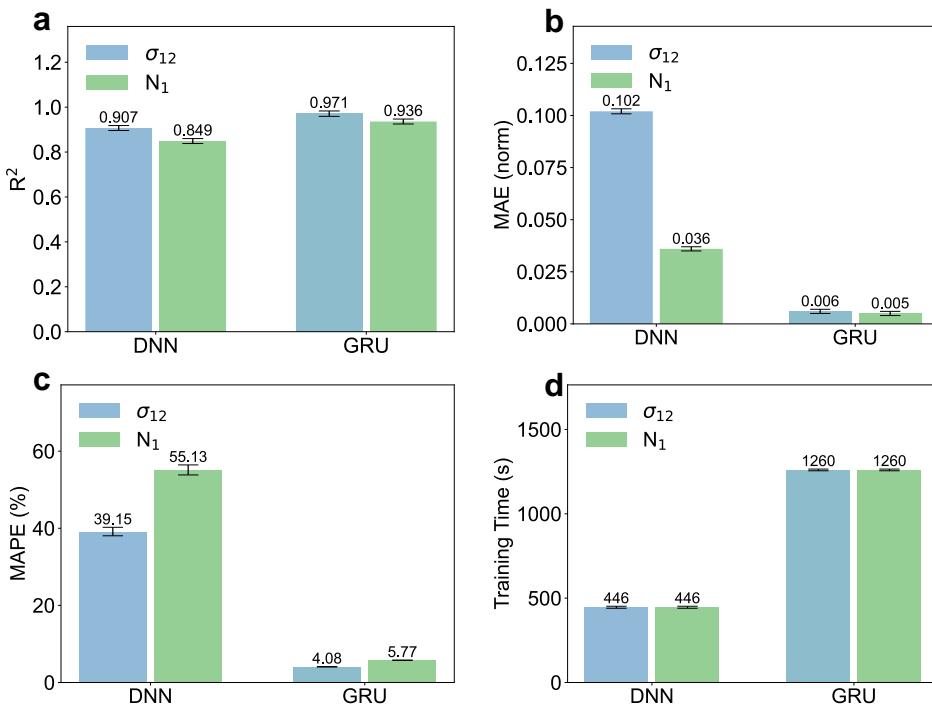


图 3-7 GRU 和 DNN 在 Doi-Edwards 模型振荡剪切测试集上的预测指标对比图：(a) R^2 值；(b) MAE 值；(c) MAPE 值；(d) 训练时间

Figure 3-7 Comparison of prediction metrics between GRU and DNN on Doi-Edwards model oscillatory shear test set: (a) R^2 value; (b) MAE value; (c) MAPE value; (d) Training time

布区间远远大于 GRU 部分。这表明 GRU 成功地学习到了训练数据中的各项特征和复杂的非线性关系，不存在明显周期性，总体残差较小，而 DNN 存在特征未能完全学习的问题，拟合效果较差。从图 d、图 e 和图 f 的 N_1 的预测效果来看，结果与 σ_{12} 类似，GRU 的预测效果明显优于 DNN。

对比同是 GRU 预测的 N_1 值和 σ_{12} 值， σ_{12} 值的预测效果更好，这一点在图3-7的定量分析中也得到了验证。从 R^2 指标看，GRU 在 σ_{12} 和 N_1 的预测指标均优于 DNN，而同种算法下 σ_{12} 的 R^2 值高于 N_1 。MAE 和 MAPE 指标显示，DNN 的 σ_{12} 和 N_1 的误差值均远高于 GRU，证实了 GRU 的预测误差显著小于 DNN。值得注意的是，同种算法下 σ_{12} 的 MAE 值高于 N_1 ，但 MAPE 值却相反，这是因为 MAE 受到数据尺度影响，而 MAPE 消除了这种影响，因此应以 MAPE 值判断，结果表明 σ_{12} 值的预测效果确实优于 N_1 。这可能是因为 σ_{12} 与输入的剪切速率特征 γ_{12} 之间的函数关系更为简单，且更加符合时间叠加原理，使 GRU 能更好地捕捉其时间依赖性。

从图3-7 (d) 可以看出，GRU 在此项任务上的训练时间约为 DNN 的 3 倍，这是更

复杂模型结构的必然代价。综合所有分析结果，当训练数据和测试数据为同类型应变变化过程（都为振荡剪切）时，GRU 算法能更好地学习 Doi-Edwards 模型数据的内在特征，包括周期性响应、黏弹性和时间依赖性响应。GRU 算法的预测泛化效果明显优于 DNN，在计算资源充足的情况下是更佳选择，而在同种算法下， σ_{12} 的预测效果优于 N_1 。

3.4.3.2 振荡剪切预测稳态剪切效果验证

为了验证 GRU 算法在不同形式的应变历史下对于 Doi-Edwards 模型泛化预测效果，本节使用振荡剪切协议生成的数据作为训练集，稳态剪切协议生成的数据作为测试集。分别使用了 GRU 和 DNN 进行训练，并在测试集上进行验证，测试结果如图3-8所示。

图3-8 (a) 展示了两种算法预测模型在测试集上的结果对比。可以看到 GRU 的预测曲线与真实值曲线更为接近，无明显偏差区间，而 DNN 的预测曲线与真实曲线存在明显偏差。在应变加载初始和结束阶段，DNN 预测值偏大，而在中间阶段（2 ~ 6 s）则偏小，整体拟合效果较差。

图3-8 (b) 的残差图显示，DNN 的残差分布不均匀，呈现出明显的曲线趋势，表明其未能充分捕捉数据特征。相比之下，GRU 的残差呈现无序分布，均匀分布在零刻度线两侧，且残差值普遍小于 DNN。这表明 GRU 能更好地捕捉训练数据特征，预测偏差更小。

图3-8 (c-e) 展示了两种算法模型的预测性能指标。GRU 的 R^2 值达到 0.979，而 DNN 仅为 0.882；GRU 的 MAE 值为 0.006，远低于 DNN 的 0.044；GRU 的 MAPE 值为 1.03%，而 DNN 高达 95.09%。这些数据表明 GRU 在该预测任务中的泛化能力显著优于 DNN。

从整体分析数据来看，当训练数据与测试数据涉及不同类型的应变变化过程时，GRU 在学习 Doi-Edwards 模型数据的内在特性方面有显著优势。这主要归因于 GRU 能够高效地从振荡剪切数据中提取关键的时间依赖特征，并将其成功应用于稳态应变的预测之中，这里与上一节中 Maxwell 模型的结论一致。

3.4.3.3 不同时间步的预测效果对比

为了探究在 Doi-Edwards 模型上 GRU 的最佳时间步，本节研究了不同时间步下训练的模型在测试集上的预测效果，如图3-9所示。由图可见，预测模型 R^2 值在时间步小于 40 时随着时间步的增加而增加，在时间步大于 40 后反而略有下降。MAE 值和 MAPE 值在时间步小于 40 时随时间步的增加而减少，时间步为 40 的值只有不到时间步为 20

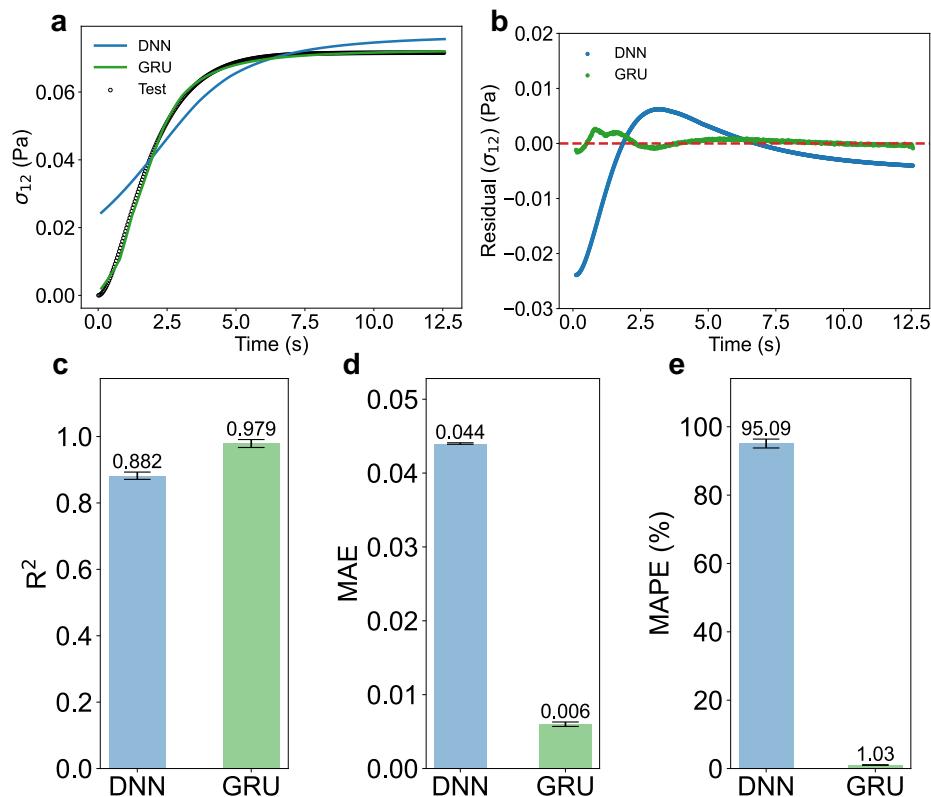


图 3-8 GRU 和 DNN 在 Doi-Edwards 模型稳态剪切测试集上的预测效果对比示意图: (a) GRU 和 DNN 在测试集上的预测值与真实值的剪切应力-应变曲线 (Lissajous 曲线); (b) GRU 和 DNN 在测试集上的预测值残差图; (c-e) GRU 和 DNN 在测试集上的预测指标对比: (c) R^2 值; (d) MAE 值; (e) MAPE 值

Figure 3-8 Comparison of prediction performance between GRU and DNN on Doi-Edwards model steady-state shear test set: (a) Shear stress-strain curves (Lissajous curves) of predicted vs. true values; (b) Residual plots of predicted values; (c-e) Comparison of prediction metrics: (c) R^2 value; (d) MAE value; (e) MAPE value

的值的十分之一，显著下降，而时间步大于 40 后，MAE 值和 MAPE 值则趋于稳定。这说明对于本项任务，时间步为 40 左右时，恰好可以开始获得不错的预测效果，而从训练成本来看，训练时间是随着时间步设置增加而单调增的，所以综合看来，时间步为 40 左右时，性价比最高。

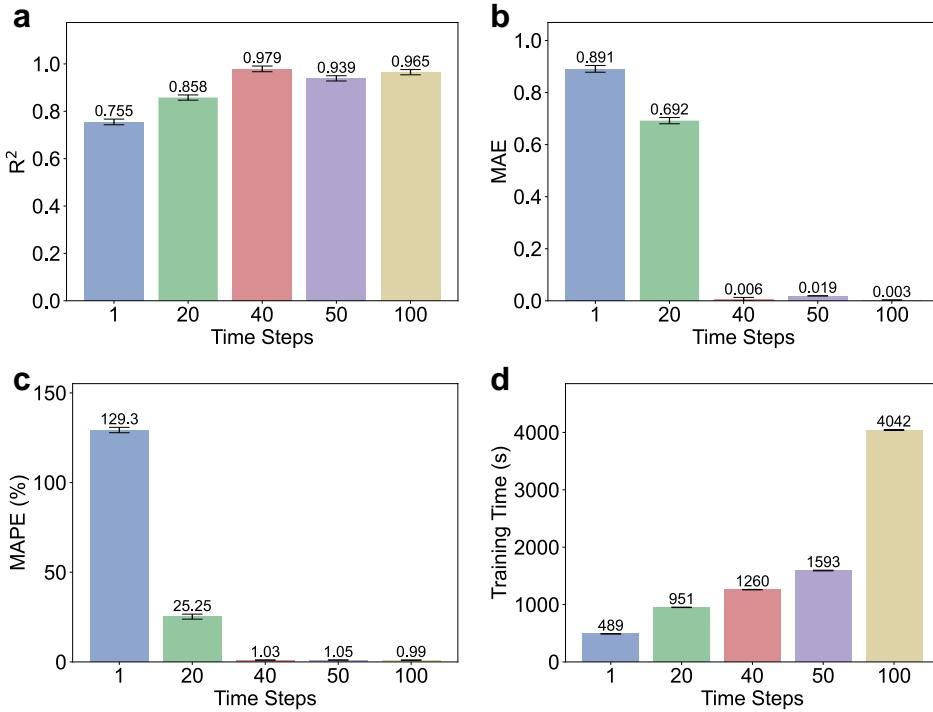


图 3-9 GRU 和 DNN 在 Doi-Edwards 模型不同时间步长下的预测指标对比图：(a) R^2 值；(b) MAE 值；(c) MAPE 值；(d) 训练时间

Figure 3-9 Comparison of prediction metrics between GRU and DNN under different time steps of the Doi-Edwards model: (a) R^2 value; (b) MAE value; (c) MAPE value; (d) Training time

3.4.4 Giesekus 模型建模

3.4.4.1 振荡剪切预测振荡剪切

前一部分的工作证明 GRU 的门控机制相比普通 DNN 的 MLP 架构在简单的时间域流变学本构方程的泛化预测中表现良好，本节在 GRU 的基础上，引入物理约束，构建了 PI-GRU 模型，并验证其在复杂非线性流变学本构方程（Giesekus 模型）的泛化预测效果。

本节通过数值模拟不同应变幅值和频率的 Giesekus 模型振荡剪切数据，将其输入 PI-GRU 模型进行训练，并评估模型对未知振荡剪切数据的预测能力。图3-10展示了 PI-GRU、MLP 和 GRU 模型（不含物理约束损失）在测试集上的预测结果与真实数据对比，图3-11则展示了相应的评价指标。

本节中 PI-GRU 模型采用上对流 Maxwell 模型（UCM）作为物理约束，其表达式如

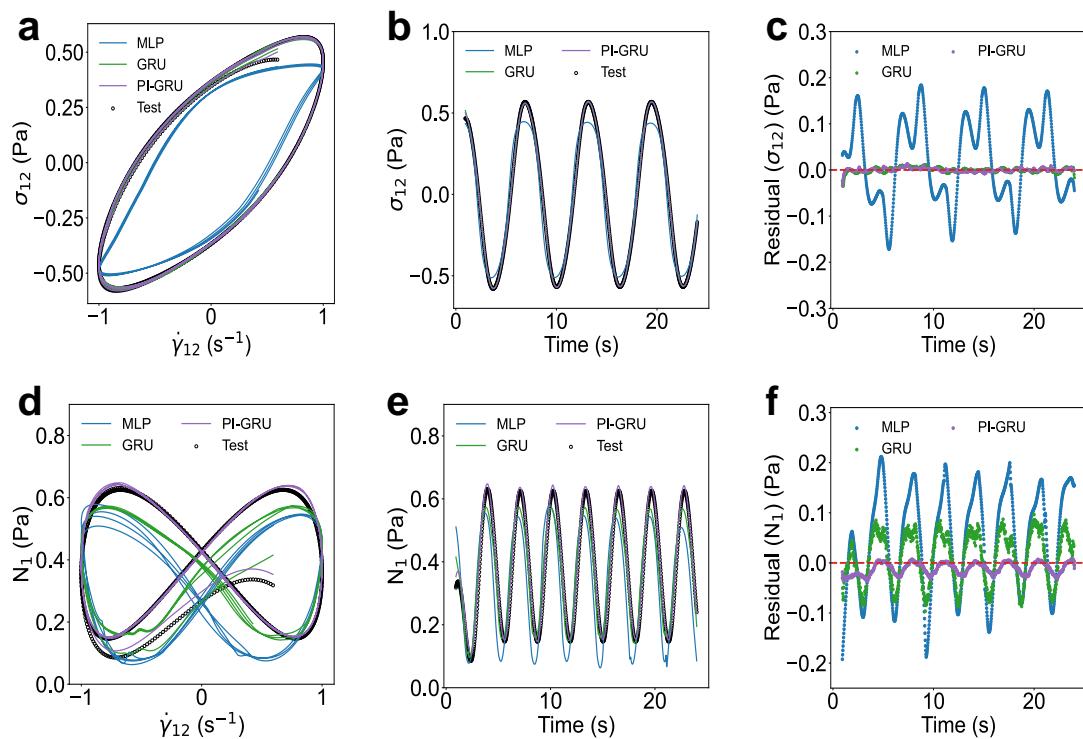


图 3-10 Giesekus 数值模拟振荡剪切数据测试集上 PI-GRU、MLP 和 GRU 模型的预测结果与真实数据对比，这里训练数据与测试数据均为大振幅振荡剪切数据：(a) 剪切应力 σ_{12} 真实值与模型预测值 Lissajous 曲线；(b) 剪切应力 σ_{12} 真实值与模型预测值的时间-应力曲线；(c) 剪切应力 σ_{12} 真实值与模型预测值残差图；(d) 第一法向应力差 N_1 真实值与模型预测值 Lissajous 曲线；(e) 第一法向应力差 N_1 真实值与模型预测值的时间-应力曲线；(f) 第一法向应力差 N_1 真实值与模型预测值残差图

Figure 3-10 Comparison schematic of the prediction results of the PI-GRU, MLP, and GRU models on the Giesekus numerical simulation oscillatory shear data test set: (a) Lissajous curve of the true value and predicted value of the shear stress σ_{12} for the GRU model; (b) time-stress curve of the true value and predicted value of the shear stress σ_{12} for the GRU model; (c) residual plot of the true value and predicted value of the shear stress σ_{12} for the GRU model; (d) Lissajous curve of the true value and predicted value of the first normal stress difference N_1 for the GRU model; (e) time-stress curve of the true value and predicted value of the first normal stress difference N_1 for the GRU model; (f) residual plot of the true value and predicted value of the first normal stress difference N_1 for the GRU model

公式3-19所示。UCM 可视为 Giesekus 模型的简化形式，省略了流动指数 α 项，因此理论上 UCM 只能部分捕捉 Giesekus 模型的特征。

$$\boldsymbol{\sigma} + \lambda \overset{\nabla}{\boldsymbol{\sigma}} = \eta \dot{\boldsymbol{\gamma}} \quad (3-19)$$

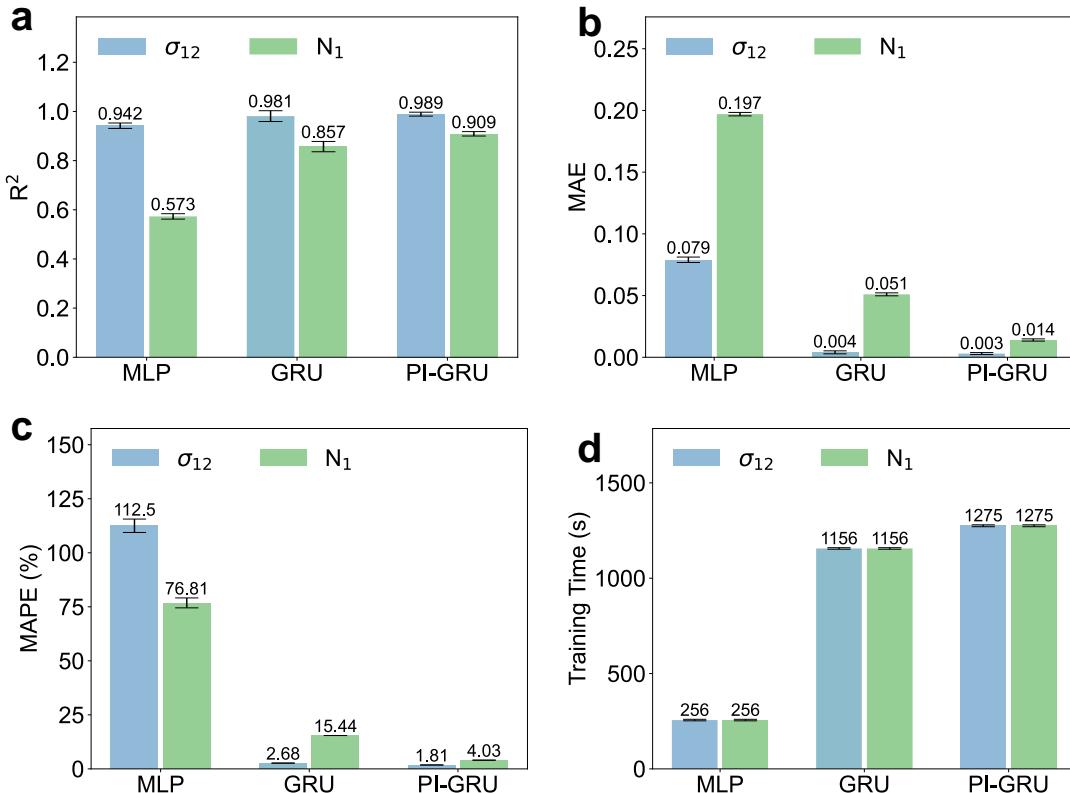


图 3-11 Giesekus 数值模拟振荡剪切数据测试集上 PI-GRU、MLP 和 GRU 模型的评价指标：(a) R^2 值；(b) MAE 值；(c) MAPE 值；(d) 训练时间

Figure 3-11 Comparison schematic of the evaluation metrics of the PI-GRU, MLP, and GRU models on the Giesekus numerical simulation oscillatory shear data test set: (a) R^2 value; (b) MAE value; (c) MAPE value; (d) training time

图3-10和图3-11的结果表明，作为前馈神经网络的 MLP 模型在预测剪切应力 σ_{12} 和第一法向应力差 N_1 时均表现最差，这符合前馈神经网络难以有效捕捉长序列时间依赖关系的预期。GRU 模型在处理时间依赖关系方面表现出色，中其预测曲线与真实值高度吻合，明显优于 MLP 模型。而 PI-GRU 模型在 σ_{12} 和 N_1 的预测中均表现最佳，优于 GRU 和 MLP。值得指出的是，在剪切应力预测任务中，引入 UCM 物理约束的 PI-GRU 模型相比 GRU 模型提升不显著。然而，在第一法向应力差 N_1 的预测中，PI-GRU 模型

明显优于 GRU 和 MLP。例如在 N_1 预测任务的 MAPE 指标上, PI-GRU 模型为 4.03%, GRU 模型为 15.44%, MLP 模型为 76.81%, 提升明显。

第一法向应力差预测难度高于剪切应力的原因在于 N_1 是应力张量对角元素的差值 ($\sigma_{11} - \sigma_{22}$), 其计算涉及多个应力分量, 误差容易累积。其次, N_1 对材料非线性响应更敏感, 在大应变下变化更为剧烈, 需要模型具备更强的非线性特征提取能力。UCM 模型虽然是 Giesekus 模型的简化形式, 但仍包含了应力张量的上对流导数项, 能够部分捕捉材料在剪切流动中的非线性行为, 特别是法向应力差的产生机制。这种物理约束引导模型学习符合流变学基本原理的特征表示, 从而在预测具有挑战性的 N_1 时表现出更强的泛化能力。

PI-GRU 模型由于其循环神经网络的内在特性, 计算复杂度较高, 导致训练时间显著长于 MLP 结构, 这一现象在本研究的所有实验任务中均有体现。尽管如此, 本任务中 1275 s 的训练时间仍处于可接受范围内。PI-GRU 模型为复杂流变学本构关系的高效建模提供了新思路, 原本需要大量不同振荡条件的实验数据才能描述的本构关系, 或许可以使用有限的数据建立模型, 再去泛化到未知实验条件的数据, 从而减少实验成本并加速材料开发过程。

3.4.4.2 振荡剪切预测稳态剪切

当训练集与测试集都为正弦振荡剪切时, 这种同分布测试无法有效验证模型是否真正建立了 Giesekus 模型的本构关系映射。Giesekus 模型可以描述非线性应力响应特性, 包括剪切稀化等特性。因此, 本节做了振荡剪切预测稳态剪切的实验来探究 PI-GRU 模型是否能够正确预测稳态剪切下的应力应变关系, 真正重现 Giesekus 本构方程。图3-12展示了 PI-GRU 模型在测试集上的预测结果与真实数据对比。

从图3-12可以看到面对不同应变协议的 Giesekus 测试数据, GRU 和 PI-GRU 的预测效果明显优于 MLP。MLP 完全无法从振荡剪切数据泛化稳态剪切数据, MLP 模型的 R^2 值甚至出现了负数, 这说明纯数据驱动的 MLP 模型无法理解 Giesekus 模型中的非线性特征。将神经网络结构改为 GRU 后, 预测效果得到明显改善。

进一步引入物理约束后, PI-GRU 模型则表现出最佳的预测泛化效果, R^2 值为 0.993(>0.99), MAE 指标再降低至 0.0085, MAPE 指标仅为 2.07%, 这表明 PI-GRU 模型能够很好学习到 Giesekus 模型中的非线性特征, 并能够从振荡剪切数据泛化到稳态剪切数据。图3-12a 显示 Giesekus 模型在稳态剪切下首先表现为弹性主导的应力增长响应,

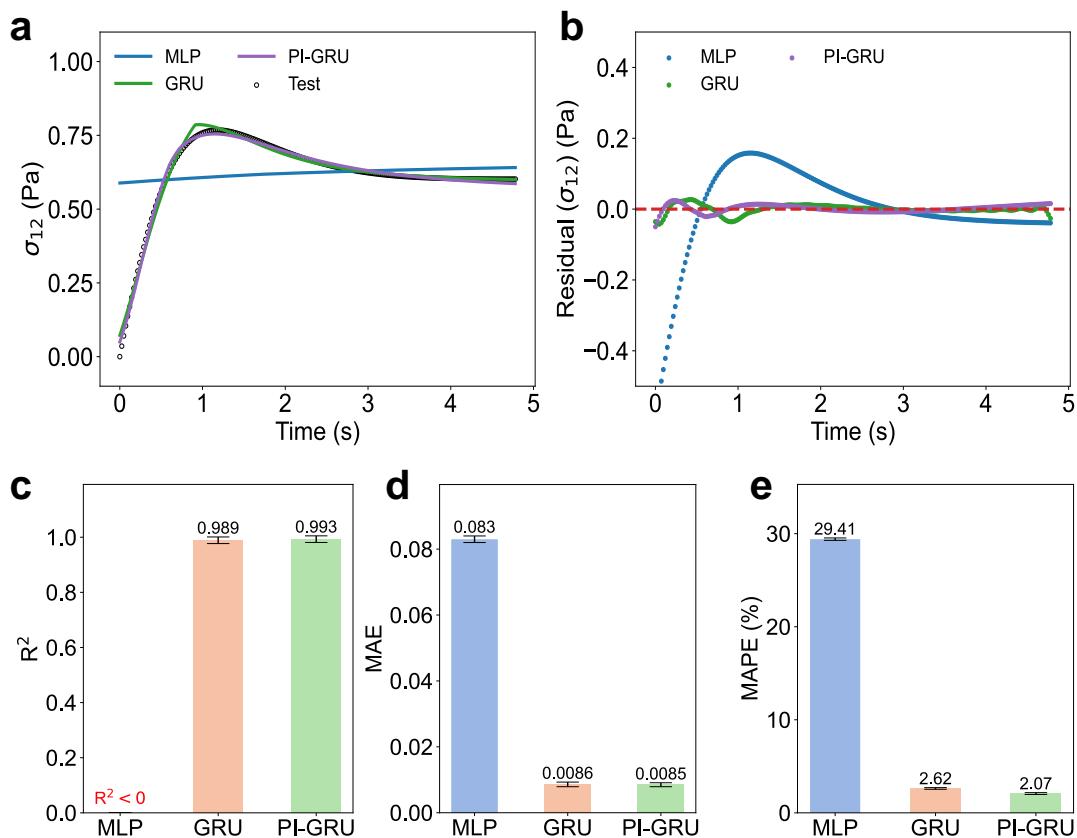


图 3-12 Giesekus 数值模拟稳态剪切数据测试集上 PI-GRU、MLP 和 GRU 模型的预测结果与真实数据对比，这里训练数据为大振幅振荡剪切（LAOS）数据，测试数据为稳态剪切数据：(a) 剪切应力 σ_{12} 真实值与模型预测值的时间-应力曲线；(b) 剪切应力 σ_{12} 真实值与模型预测值残差图；(c) R^2 值；(d) MAE 值；(e) MAPE 值

Figure 3-12 Comparison schematic of the prediction results of the PI-GRU, MLP, and GRU models on the Giesekus numerical simulation steady-state shear data test set: (a) time-stress curve of the true value and predicted value of the shear stress σ_{12} for the GRU model; (b) residual plot of the true value and predicted value of the shear stress σ_{12} for the GRU model; (c) R^2 value; (d) MAE value; (e) MAPE value

之后应力衰减对应剪切稀化现象，最后黏性主导下呈现平台，PI-GRU 模型能够从振荡剪切数据中正确识别出这些 Giesekus 模型的特征，将相关的时间依赖性特征与非线性特征嵌入到模型的参数中，最后在稳态剪切下预测出与真实数据高度吻合的应力应变曲线。

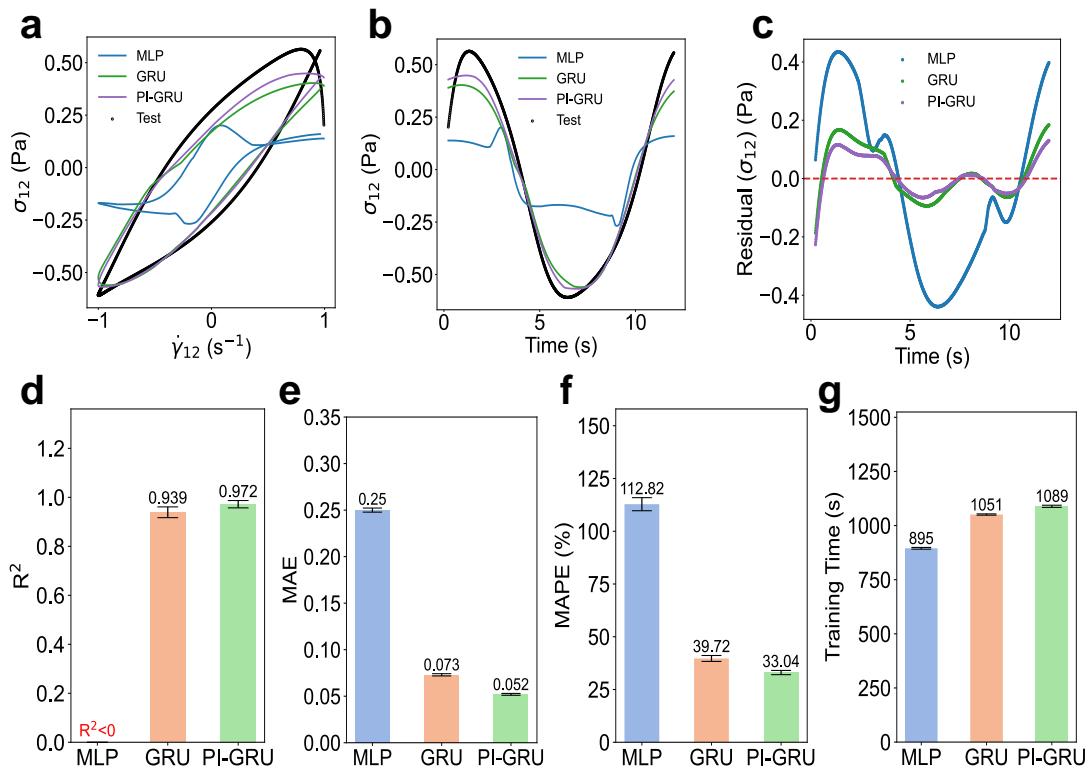


图 3-13 Giesekus 数值模拟线性 SAOS 数据预测非线性 LAOS 数据测试集上 PI-GRU、MLP 和 GRU 模型的预测结果与真实数据对比，这里训练数据为 SAOS 数据，测试数据为 LAOS 数据：(a) 剪切应力 σ_{12} 真实值与模型预测值的 Lissajous 曲线；(b) 剪切应力 σ_{12} 真实值与模型预测值的时间-应力曲线；(c) 剪切应力 σ_{12} 真实值与模型预测值残差图；(d) R^2 值；(e)MAE 值；(f)MAPE 值；(g) 训练时间

Figure 3-13 Comparison schematic of the prediction results of the PI-GRU, MLP, and GRU models on the Giesekus numerical simulation linear SAOS data prediction nonlinear LAOS data test set: (a) Lissajous curve of the true value and predicted value of the shear stress σ_{12} for the GRU model; (b) time-stress curve of the true value and predicted value of the shear stress σ_{12} for the GRU model; (c) residual plot of the true value and predicted value of the shear stress σ_{12} for the GRU model; (d) R^2 value; (e) MAE value; (f) MAPE value; (g) Training time

3.4.4.3 线性区间预测非线性区间

在流变学本构方程的研究中，非线性本构方程的数学建模和参数拟合难度往往远远大于线性本构方程，在实际的流变学测试中，能够获取非线性数据的 LAOS 实验成本也远远高于 SAOS。因此，如果能够从线性的 SAOS 数据出发，探究出非线性区间的全部

或部分本构关系，是非常有意义的。本节通过构建 PI-GRU 模型，做了这样的尝试：训练数据全部采用 SAOS 数据，测试数据全部采用 LAOS 数据，探究 PI-GRU 模型是否能够从线性区间的 SAOS 数据泛化到非线性区间的 LAOS 数据。

这里还是选取了 MLP，GRU 和 PI-GRU 模型进行对比，所使用的训练集全部都是 SAOS 数据，最终图3-13绘制的是 LAOS 测试集的对比图，可以发现 MLP 在预测 LAOS 数据时，预测曲线与真实曲线偏差较大，各指标表现较差，GRU 和 PI-GRU 模型在预测 LAOS 数据时，预测效果相对 MLP 得到一定提升，图3-13（a-c）可以定性看出不同模型之间的预测效果是：PI-GRU>GRU>MLP。在定量指标方面（图3-13（d-f）），也可以得到同样的模型效果比较顺序，但是 PI-GRU 相比 GRU 的效果提升远不如 GRU 相比 MLP 的效果提升，这说明物理约束对模型在此项任务上的提升有限。这里的結果可能是因為，训练数据中数据都是线性区间数据，非线性特征较少，深度学习方法很难从几乎没有的特征中学到未知领域的知识，引入物理约束后，这种物理约束的先验信息一定程度上拓展了模型的想象能力，但受限于训练数据中确实缺少非线性特征，物理约束的先验信息无法有效指导模型学习到非线性特征，从而导致预测效果提升有限。三个模型的训练时间也是 PI-GRU>GRU>MLP，这一点符合预期，但是 PI-GRU 的训练时间为 1089 s，相比 MLP 的 895 s，耗时增加仅仅 21.7%，说明 PI-GRU 模型的训练成本可以接受。

综合来看，PI-GRU 模型在物理约束和 GRU 结构的加持下，可以在一定程度上通过线性 SAOS 数据预测非线性的 LAOS 数据，尽管预测的误差没有上文中的振荡剪切预测稳态剪切任务中那么小，但仍然表明 PI-GRU 模型的应用潜力。由于线性区间的 SAOS 数据是较容易获取的，因此未来也许可以考虑先通过大量的 SAOS 数据预训练一个基础模型，再通过有限的 LAOS 数据进行迁移学习，从而提升模型的泛化能力，不失为一个降本增效的策略。

3.4.5 真实实验数据建模

上一节中，我们验证了 PI-GRU 模型在模拟数据上的泛化预测效果，然而模拟的数据往往噪声较小，且受到人为的精确控制，而真实实验数据往往噪声较大，且受到实验条件的影响，真实的黏弹性材料中往往存在更为复杂的非线性特征，例如多松弛谱等现象，因此本节将探究 PI-GRU 模型在真实实验数据上的泛化能力，即 PI-GRU 模型是否能够在真实的材料科研中起到作用。本研究选取的实验材料为一类具有黏弹性的 ABE 弹性体阻尼材料，具体材料情况在上一节中已经介绍。

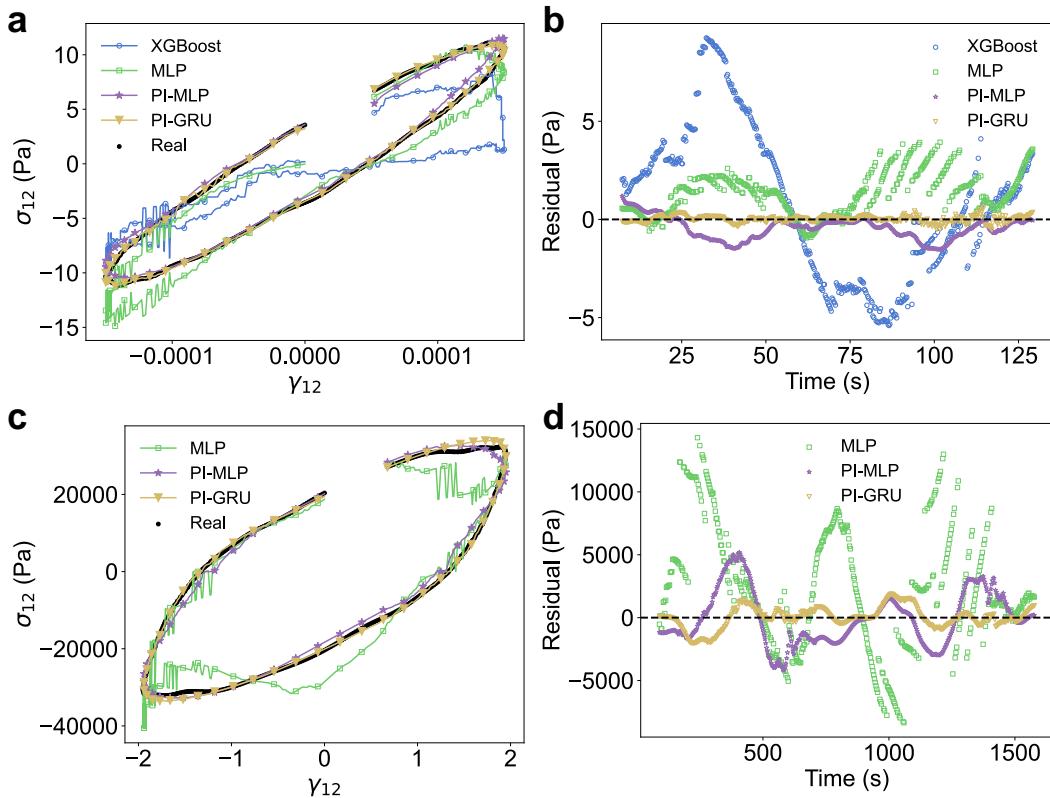


图 3-14 ABE 弹性体数据测试集上 XGBoost、MLP、PI-MLP 和 PI-GRU 模型的预测结果与真实数据对比 (a) SAOS 测试集剪切应力 σ_{12} 真实值与模型预测值的 Lissajous 曲线; (b) SAOS 测试集剪切应力 σ_{12} 真实值与模型预测值的残差图; (c) LAOS 测试集剪切应力 σ_{12} 真实值与模型预测值的 Lissajous 曲线; (d) LAOS 测试集剪切应力 σ_{12} 真实值与模型预测值的残差图

Figure 3-14 Comparison schematic of the prediction results of the XGBoost, MLP, PI-MLP, and PI-GRU models on the ABE elastomer data test set: (a-b) SAOS test set, both training and test data are SAOS data, and (c-d) LAOS test set, training data are SAOS data and LAOS data, test data are LAOS data: (a) Lissajous curve of the true value and predicted value of the shear stress σ_{12} for the SAOS test set; (b) residual plot of the true value and predicted value of the shear stress σ_{12} for the SAOS test set; (c) Lissajous curve of the true value and predicted value of the shear stress σ_{12} for the LAOS test set; (d) residual plot of the true value and predicted value of the shear stress σ_{12} for the LAOS test set

3.4.5.1 SAOS 数据预测

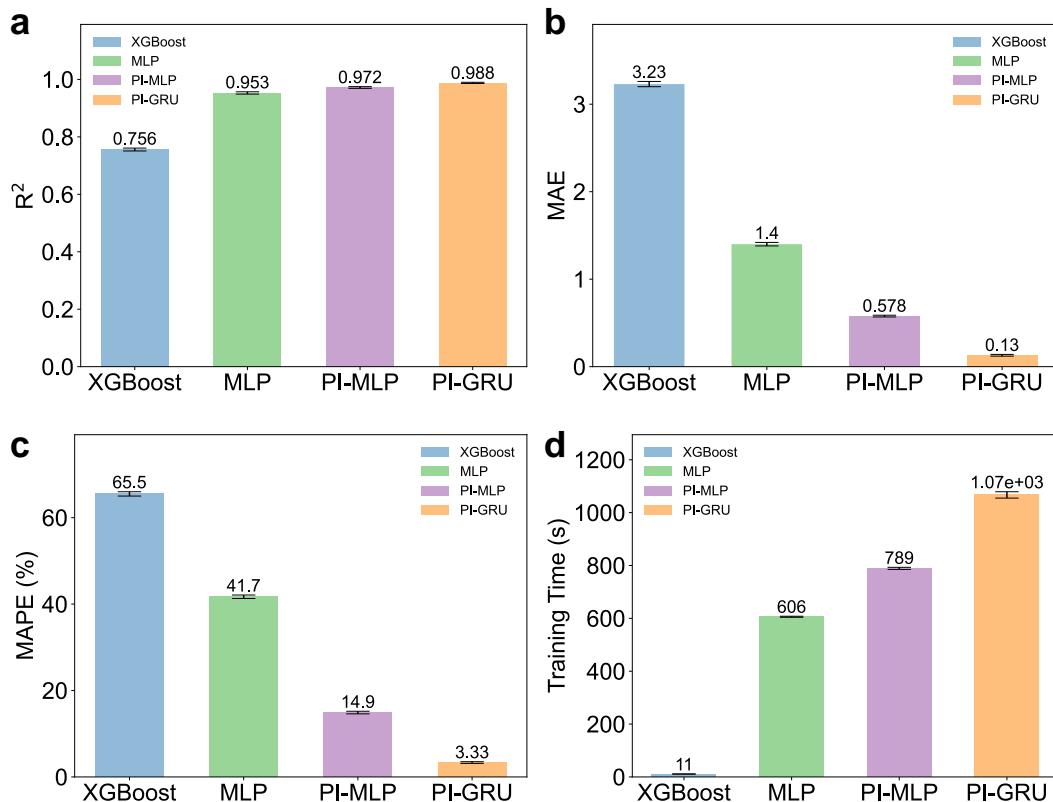


图 3-15 ABE 弹性体数据 SAOS 测试集上 XGBoost、MLP、PI-MLP 和 PI-GRU 模型的预测指标对比: (a) R^2 值; (b) MAE 值; (c) MAPE 值; (d) 训练时间

Figure 3-15 Comparison schematic of the prediction metrics of the XGBoost, MLP, PI-MLP, and PI-GRU models on the ABE elastomer data SAOS test set: (a) R^2 value; (b) MAE value; (c) MAPE value; (d) Training time

图3-14 (a, b) 和图3-15展示了 PI-GRU 模型在真实数据集的 SAOS 数据上的预测结果, 这里对比了 XGBoost 模型, MLP 模型, PI-MLP 模型(增加物理约束, 但使用 MLP 结构)和 PI-GRU 模型。这里使用的物理约束为 UCM 模型, 其表达式如公式3-19所示。

从真实-预测值曲线(图3-14a)和残差图(图3-14b)可以定性看出不同模型之间的预测效果是: PI-GRU>PI-MLP>MLP>XGBoost。从定量指标来看, 纯数据驱动的 XGBoost 模型和 MLP 模型在预测真实数据时表现较差, MAPE 指标均高于 40%, 在物理约束的加持下, PI-MLP 模型和 PI-GRU 模型在预测真实数据时表现较好, MAPE 指标均低于 20%, 其中 PI-GRU 模型在预测真实数据时表现最佳为 3.33%。而训练时间上,

PI-GRU 模型由于其循环神经网络结构特性，训练时间最高为 1070 s。

本节的研究表明，PI-GRU 模型在真实的 SAOS 数据上可以取得较好的预测效果，通过学习不同振幅的 SAOS 数据，模型可以学到材料线性黏弹性特征，并泛化到未知应变协议的数据，这意味着只需有限的的流变学实验，便可以得到一个较为准确的某种材料的线性黏弹本构方程。

3.4.5.2 LAOS 数据预测

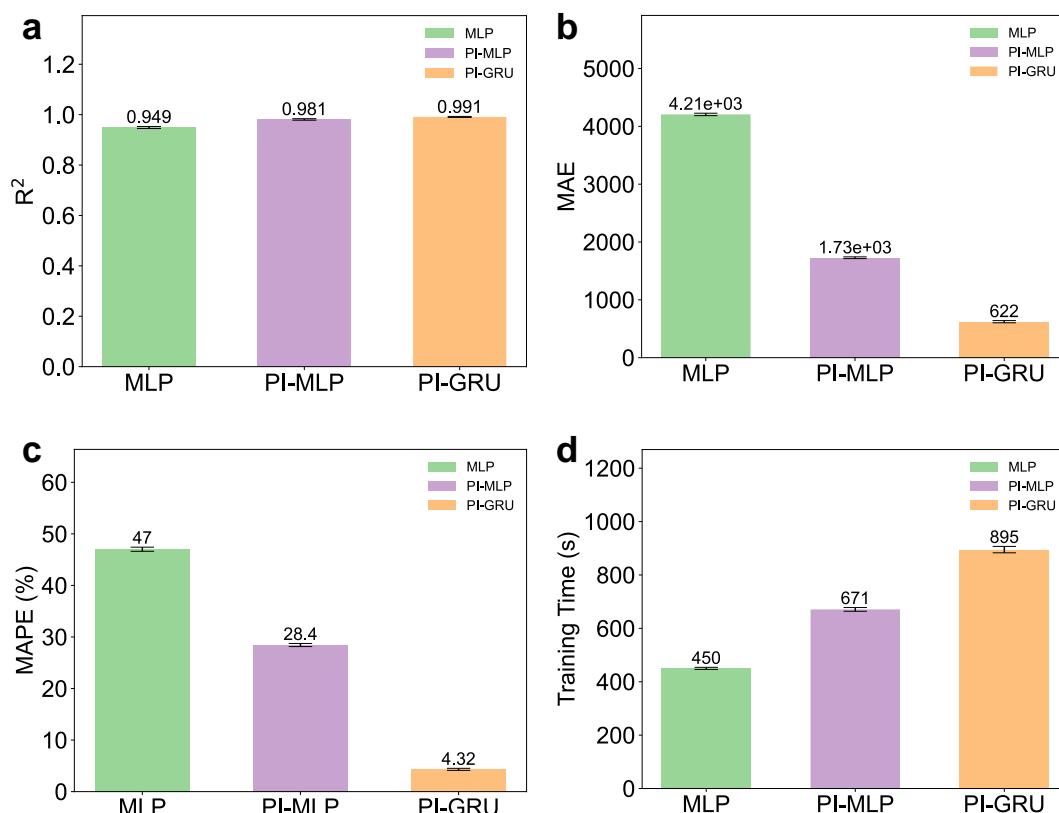


图 3-16 ABE 弹性体数据 LAOS 测试集上 XGBoost、MLP、PI-MLP 和 PI-GRU 模型的预测指标对比：(a) R^2 值；(b) MAE 值；(c) MAPE 值；(d) 训练时间

Figure 3-16 Comparison schematic of the prediction metrics of the XGBoost, MLP, PI-MLP, and PI-GRU models on the ABE elastomer data LAOS test set: (a) R^2 value; (b) MAE value; (c) MAPE value; (d) training time

在进一步的研究中，图3-14 (c, d) 和图3-16展示了 PI-GRU 模型在真实数据集的 LAOS 数据上的预测结果，由于 LAOS 数据的测量难度和时间都高于 SAOS 且需要的样品量更多，这里的训练数据同时包含 SAOS 和 LAOS 数据，这里对比了 MLP 模

型, PI-MLP 模型和 PI-GRU 模型, 其中 PI-GRU 模型预测的曲线与真实曲线最为吻合, 图3-16的预测相关指标显示 PI-GRU 在这项任务上的 R^2 为 0.991, MAE 值为 622, MAPE 值为 4.32%, 均领先其他算法模型。

PI-GRU 模型的优越性能源于其门控机制与黏弹性材料应力响应机制高度契合, 可选择性保留或遗忘历史信息, 而物理约束的引入使其遵循流变学原理, 避免纯数据驱动方法的物理不一致性。通过 SAOS 与 LAOS 数据协同训练, 模型可学习材料从线性到非线性区域的连续转变特性, 构建完整响应图景, 实现不同应变幅度和频率下的精准预测。这种基于“多尺度”学习的方法, 在 SAOS 数据易获取、LAOS 数据难获取的实际条件下, 能高效建立黏弹性材料非线性本构关系, 显著降低高分子加工领域的实验成本与时间, 为材料开发和工艺优化提供高效数据支持。

3.4.5.3 不同本构约束

本研究使用的 PI-GRU 模型一个很重要的组件就是物理方程的约束, 而 PI-GRU 作为物理信息机器学习的一种, 存在一个研究者都颇为关心的问题, 即如何在一个具体的科学问题中从浩繁的物理方程中选择一个最适配的物理方程。为了了解 PI-GRU 模型对于本构方程的依赖性, 本节探究了不同本构方程作为物理约束下的建模预测效果。图3-17和图3-18展示了不同本构方程作为物理约束的 PI-GRU 的预测效果, 其中纯数据驱动 (DD) 组的效果最差, 这里本质上使用的 GRU 模型, 与上文的结果不同, 在真实的数据集上, 单纯的 GRU 模型预测非线性区间的 LAOS 数据表现不佳, MAPE 值高达 36.4%, 这里考虑是因为真实场景下材料流变性质的复杂性超过有限参数的数值模拟数据。而引入物理约束后, 模型的预测效果则得到较大提升, 本节的工作中, M 代表的是 Maxwell 模型, SLS 为标准线性固体模型, 方程如所示, FSLS 为分数阶线性固体模型, 方程如所示。

$$\sigma(t) + \lambda_1 \frac{d\sigma(t)}{dt} = G \left(\gamma(t) + \lambda_2 \frac{d\gamma(t)}{dt} \right) (SLS) \quad (3-20)$$

$$\sigma(t) + \lambda_1 \frac{d^\alpha \sigma(t)}{dt^\alpha} = G \left(\gamma(t) + \lambda_2 \frac{d^\beta \gamma(t)}{dt^\beta} \right) (FSLS) \quad (3-21)$$

研究中使用的 ABE 弹性体是一类黏弹性固体材料, Maxwell 模型在实践中多用于高分子流体的线性区间拟合, SLS 模型则多用于黏弹性固体的线性区间拟合, FSLS 模型通过分数阶导数的记忆性和幂律特性, 突破了整数阶模型的指数衰减限制, 兼容材料的长期记忆、宽弛豫时间分布和幂律响应等非线性特征^[3,21]。这几个本构模型中理论上

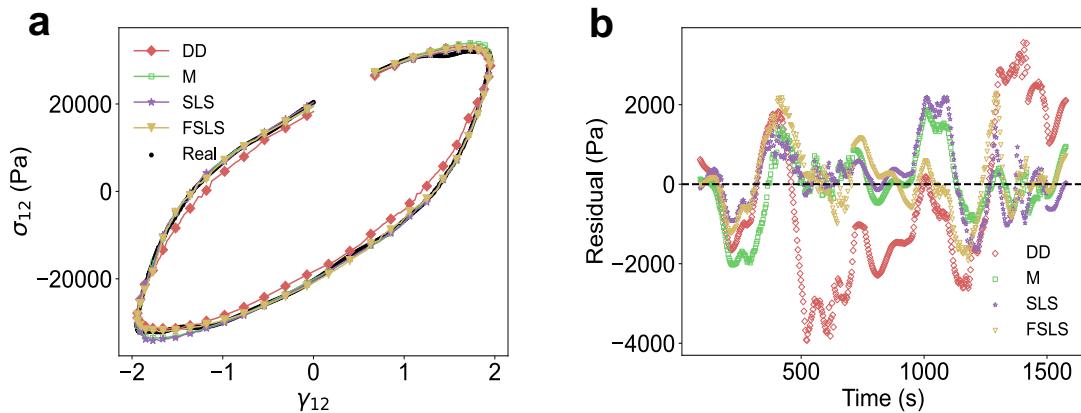


图 3-17 ABE 弹性体数据 LAOS 测试集使用不同本构方程作为物理约束的 PI-GRU 模型的预测效果对比，这里 DD 指的是没有物理约束的纯数据驱动模型，M 代表的是 Maxwell 模型，SLS 为标准线性固体模型，方程如3-20所示，FSLS 为分数阶线性固体模型，方程如3-21所示：(a) 剪切应力 σ_{12} 真实值与模型预测值的 Lissajous 曲线；(b) 剪切应力 σ_{12} 真实值与模型预测值的残差图

Figure 3-17 Comparison schematic of the prediction results of the PI-GRU model with different constitutive equations as physical constraints on the ABE elastomer data LAOS test set. Here DD refers to a pure data-driven model without physical constraints, M represents the Maxwell model, SLS is the standard linear solid model as shown in equation 3-20, and FSLS is the fractional-order linear solid model as shown in equation 3-21: (a) Lissajous curve of the true value and predicted value of the shear stress σ_{12} ; (b) residual plot of the true value and predicted value of the shear stress σ_{12}

FSLS 模型最能描述 ABE 弹性体的本构关系，而应用到 PI-GRU 的模型中，结果表明三种不同适配场景的本构方程作为物理约束情况下，最终的泛化预测效果非常接近，且没有呈现预期的梯度变化趋势。具体而言，在 R^2 指标上 Maxwell 模型约束的模型表现最优，在 MAE 指标上 SLS 模型略胜一筹，而在 MAPE 指标上 Maxwell 模型又占优势，但这些差异在统计学意义上并不显著。这一现象表明，PI-GRU 模型在 ABE 弹性体流变数据建模过程中表现出一定程度的本构方程无关性，即使是理论上更适合流体而非固体的 Maxwell 方程作为物理约束，模型也能取得令人满意的预测效果。

PI-GRU 的本构无关性可能源于物理信息与 GRU 网络结构之间的协同增强作用。图3-14的结果清晰地表明 PI-GRU 相比 PI-MLP 具有显著的预测优势，而图3-17的结果则

证实 PI-GRU 相比纯数据驱动的 GRU 模型有着本质的提升，这两组对比共同证明了物理约束和门控循环单元结构在提升预测性能方面均发挥着关键作用。物理约束为模型提供了一个基础的理论框架，使得模型不必严格符合特定材料的全部特性，只需满足黏弹性材料的基本数学假设；而 GRU 结构则通过其记忆机制有效捕捉了材料响应中的时间依赖性和非线性特征。这种设计理念与 Mahmoudabadbozchelou 和 Jamali 提出的 MFNN 或 PINN 模型有相似之处^[106]，但 PI-GRU 模型通过其特殊的神经网络架构，有效地捕捉了时间域流变数据中的历史依赖性特征，在模型设计上更为贴近真实的黏弹性材料。

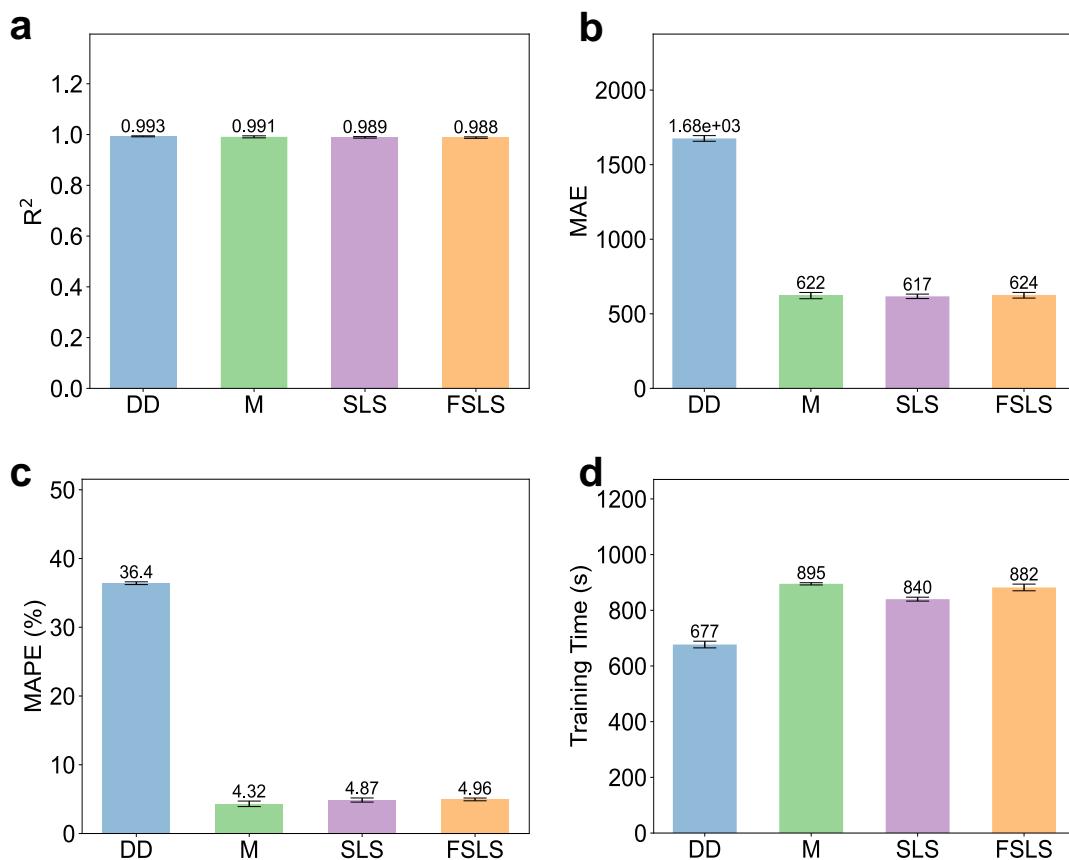


图 3-18 ABE 弹性体数据 LAOS 测试集使用不同本构方程作为物理约束的 PI-GRU 模型的预测指标对比：(a) R^2 值；(b) MAE 值；(c) MAPE 值；(d) 训练时间

Figure 3-18 Comparison schematic of the prediction metrics of the PI-GRU model with different constitutive equations as physical constraints on the ABE elastomer data LAOS test set: (a) R^2 value; (b) MAE value; (c) MAPE value; (d) Training time

Lennon 等人提出的 RUDE 框架采用了 Oldroyd 形式的微分方程作为基础框架，Oldroyd 形式的微分方程事实上已经包含时间依赖性，仅通过神经网络训练框架无法拟

合的非线性部分，因此 RUDE 不需要事先确定精确的本构方程^[109]。相比之下，我们的 PI-GRU 模型将物理约束作为一个大致约束，依靠 GRU 的门控单元结构来学习材料的时间依赖性，本质上是将 RUDE 的设计以非解析方程的形式融入到模型中。

这一发现对于实际应用具有重要意义，它表明在使用 PI-GRU 模型进行黏弹性材料建模时，研究者可以选择相对简单的物理约束方程，而不必过分关注物理方程与材料特性的精确匹配度，这大大简化了模型的设计和应用过程。

3.5 本章小结

本章首先探索了 GRU 算法在流变学建模中的应用潜力，通过对 Herschel-Bulkley 模型、Maxwell 模型及 Doi-Edwards 模型生成的模拟数据进行建模分析。研究表明，GRU 凭借其门控单元机制能有效捕捉黏弹性材料的时间依赖特性，即使在非时间序列的 Herschel-Bulkley 模型数据上也略优于 DNN 模型。对于时间序列特性明显的 Maxwell 模型和复杂的 Doi-Edwards 模型，GRU 的预测效果更是全面超越传统方法。这些初步探索为后续研究奠定了基础。

在此基础上，本章重点开发了物理信息门控循环单元（PI-GRU）模型，通过将物理约束引入 GRU 的损失函数，实现了黏弹性材料本构关系的高精度建模。PI-GRU 模型在数值模拟数据和真实实验数据上均展现出卓越性能。在 Giesekus 模型生成的数值模拟数据上，PI-GRU 不仅能准确预测振荡剪切数据，还能从振荡剪切训练数据泛化到稳态剪切工况，预测百分比误差低至 2.07%，证明其对材料应变历史依赖性和剪切变稀等复杂行为的深刻理解。更值得注意的是，PI-GRU 模型在一定程度上能够仅通过线性黏弹性数据预测非线性黏弹性行为，这一能力对于降低实验成本具有重要意义。PI-GRU 模型在振荡剪切数据上训练后能够准确预测稳态剪切下的应力应变关系，表明其具有构建完备统一本构关系的能力。这一特性使得相对易获取的振荡剪切数据可以低成本地用于构建各种复杂工况下的本构关系，如注塑成型、挤出成型等大剪切形变工况，有望显著缩短实验周期，降低工业试错成本。

在真实 ABE 弹性体数据测试中，PI-GRU 模型在 SAOS 和 LAOS 数据集上的 MAPE 指标分别为 3.33% 和 4.32%，显著优于 XGBoost、MLP 和 PI-MLP 等对比模型。这一结果验证了 PI-GRU 模型在处理真实流变学数据时的实用价值，能够同时适应线性和非线性黏弹区的材料响应特性。

本章的研究还发现了 PI-GRU 模型具有一定程度的本构方程无关性。即使使用理论

上不太匹配的物理约束方程（如用 Maxwell 模型约束非线性黏弹性固体），模型仍能取得令人满意的预测效果。这一特性源于物理约束与 GRU 网络结构之间的协同增强作用：物理约束提供基础理论框架，而 GRU 结构有效捕捉时间依赖性和非线性特征。这意味着在实际应用中，只需物理约束本构方程符合基本物理假设，而不必严格匹配每种材料的具体特性，大大简化了模型的设计和应用过程。

综上所述，本章通过系统研究证明了 PI-GRU 模型在流变学本构建模中的优越性能。该模型不仅能够准确捕捉黏弹性材料的时间依赖性和非线性特征，还展现出卓越的泛化能力和预测稳定性。这一研究为深度学习在流变学本构建模领域的应用开辟了新思路，即从模型结构升级的角度优化本构建模过程，为高分子材料的设计与加工提供了高效的数据驱动方法。

第四章 物理信息-生成式正逆向流变学本构建模研究

4.1 引言

流变学本构建模长期面临实验数据与理论模型间的鸿沟。传统方法常基于理想化假设构建本构方程，虽能获得形式优美的数学模型，却难以准确表征真实材料在复杂工况下的非线性响应。基于前章通过 PI-GRU 网络对流变数据的时序建模基础，本章则提出物理信息驱动与生成式建模融合的新范式，通过融合频率域的真实实验数据与物理守恒定律的强约束机制，构建兼具高预测精度与可定制性的智能建模框架。

当前流变学本构建模领域存在两大核心问题，第一个问题是传统物理模型在描述真实材料复杂流变行为时，常因过度简化导致预测偏差累积，而纯数据驱动的黑箱模型虽能实现高精度拟合，却丧失了物理可解释性这一流变学研究的本质诉求，同时由于流变学数据往往依赖于各类流变仪的实验测量，所以优质的流变学数据是稀有的，本文前一章使用数值模拟方法可以生成大量模拟数据，但是这在真实实验上不可以复刻。Mahmoudabadbozchelou 提出的一类 PINN 方法，通过对实验数据进行本构方程的参数拟合，然后借用拟合后的方程进行数值模拟生成大量低保真数据，之后使用高保真数据（原本的实验数据）和低保真数据（模拟数据）联合进行深度学习训练，在一定程度上缓解了数据不足问题，同时由于低保真数据本身是符合经典本构方程的，具有一定物理约束意义。本章参考了 Mahmoudabadbozchelou 的方法，来对一类黏弹性聚合物凝胶的流变学数据：储存模量 (G')、损耗模量 (G'') 和损耗角正切 ($\tan\delta$) 进行 PINN 建模，尝试构建材料制备参数、频率到流变学性质的模型映射。在 Mahmoudabadbozchelou 的基础模型中，本文引入了几个优化，首先是通过引入可学习的损失函数权重，使得模型在训练过程中，能够根据训练数据中不同参数的分布情况，动态调整损失函数的权重，从而提高模型泛化能力，其次在特征工程方面，引入注意力特征融合的方法，进一步解决实验数据特征稀疏的问题。

第二个问题是材料逆向设计过程中，基于试错法的实验优化模式耗费大量资源，而现有生成模型在流变学参数空间的可控生成方面缺乏物理约束，导致生成结果常偏离热力学可行域。近年来，许多研究者聚焦于材料制备参数到流变学性质的建模，但是实际应用中，根据已知的期望的材料性质，通过反向建模来确定这种材料应该如何制备也是非常重要。考虑到多组分材料体系的流变指纹具有高维参数空间中的低维流形特性，变分自编码器 (VAE) 的潜空间建模应用于流变学的反向建模是可行的，本章使用了 VAE

类模型中的条件变分自编码器（CVAE）来进行反向建模，探究定制制备参数到流变学特性的可控映射。

本章创新性地将物理信息神经网络（PINN）与条件变分自编码器（CVAE）进行耦合，建立双向建模通道：在正向建模路径中，通过将已知的本构方程编码为 PINN 的软约束条件，有效解决了小样本实验数据下的过拟合问题；在逆向设计路径中，利用 CVAE 的潜空间探索能力，结合流变响应的物理可行性验证模块，实现了从目标流变特性到材料制备参数的可控映射。这种混合建模策略不仅突破了传统方法在数据-物理融合层面的技术壁垒，更重要的是构建了闭环的材料设计-验证工作流，为智能流变学提供一定方法论基础。

4.2 实验设计

4.2.1 PINN-CVAE 正逆向建模架构

4.2.2 PINN 数据预处理

4.2.2.1 实验数据来源

本章所有训练和测试的实验数据是基于 Huang 等人的工作^[143]。Huang 等人介绍了一种将粘性聚合物流体（PBA）注入弹性网络中，形成聚合物流体凝胶（PFGs），以实现宽频带可控超高能量耗散的策略，如图4-1所示。在实验过程中，Huang 等人对不同分子量的 PBA 流体，注入不同分子量 PBA 制备的 PFGs，分别进行了流变学实验，得到了对应材料特定频率下的储存模量、损耗模量和损耗角正切等数据。本章采用他们的实验数据进行具体的深度学习建模。

4.2.2.2 实验数据分类

首先本节对真实实验数据进行分类，第一类为单 PBA 流体数据，为不同分子量的 PBA 流体的流变学数据，特征包括：聚合度（DP）、数均分子量（M_n）、分散指数（PDI）、频率（ ω ），标签包括：储存模量（G'）、损耗模量（G''）和损耗角正切（tan δ ），第二类为单分子量 PBA 注入制备的 PFGs 数据，特征和标签和第一类单 PBA 流体数据一致，第三类为多分子量 PBA 注入制备的 PFGs 数据，特征包括：不同分子量的 PBA 特征和组分信息（M_{n*i*}、PDI_{*i*}、DP_{*i*}、 ϕ_i , $i = 1, 2, 3$ ）、频率（ ω ），标签包括：储存模量（G'）、损耗模量（G''）和损耗角正切（tan δ ）。第三类数据的组分信息见表4-1所示。

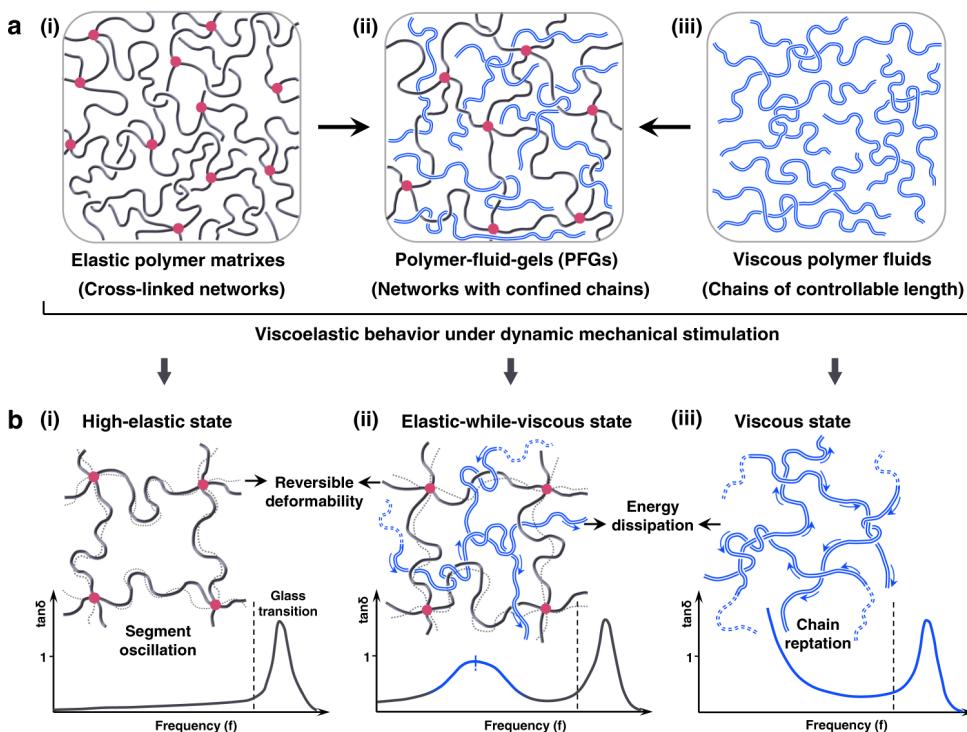


图 4-1 Huang 等人制备的 PFGs 示意图

Figure 4-1 Schematic illustration of the PFGs prepared by Huang et al.

表 4-1 多分子量 PBA 流体制备 PFGs 的配比参数表

Table 4-1 Formulation Parameters of Multi-MW PBA Fluids for PFGs

实验分组	$\phi_{PBA}(\%)$
	Mn(20k,35k,52k,78k,102k,152k)
PFG-b1	(0,20,0,0,30,10)
PFG-b2	(0,30,0,30,0,0)
PFG-b3	(20,0,0,40,0,0)
PFG-b4	(10,0,20,30,0,0)
PFG-b5	(0,0,30,0,0,30)
PFG-b6	(0,0,20,0,0,40)

4.2.2.3 低保真数据生成

分类后，本节使用 Doi-Edwards 模型来进行低保真数据拟合。首先假设实验数据可以使用 Doi-Edwards 模型描述。Doi-Edwards 模型的频率域公式可以通过时间域公式进

行傅里叶变换得到公式(4-1)和公式(4-2)

$$G'(\omega) = G_0 \frac{8}{\pi^2} \sum_{p=1,3,5,\dots}^{\infty} \frac{1}{p^2} \frac{1}{1 + (\omega\tau_d/p^2)^2} \quad (4-1)$$

$$G''(\omega) = G_0 \frac{8}{\pi^2} \sum_{p=1,3,5,\dots}^{\infty} \frac{1}{p^2} \frac{\omega\tau_d/p^2}{1 + (\omega\tau_d/p^2)^2} \quad (4-2)$$

本节使用 Python 的 `scipy.optimize` 库对真实数据进行最小二乘法的拟合，为了简化运算设置公式中的 p 为 1，得到 Doi-Edwards 模型的拟合参数。之后根据拟合后的方程，通过 Numpy 库生成频率-模量数据，用于后续的 PINN 模型训练，这一部分数据被称为低保真数据（LF-data）。

4.2.2.4 数据集划分

本节将数据集划分为训练集、验证集和测试集，其中训练集和验证集用于模型训练，测试集用于模型验证。对于 PBA 流体数据取 DP 值为 162 的数据为测试集，其余数据为训练集和验证集，训练集和验证集按照 9:1 的比例划分。对于单分子量 PBA 注入制备的 PFGs 数据，取 DP 值为 162 的数据为测试集，其余数据为训练集和验证集，训练集和验证集按照 9:1 的比例划分。对于多分子量 PBA 注入制备的 PFGs 数据，取 PFG-b2 组（分组见表4-1）的数据为测试集，其余数据为训练集和验证集，训练集和验证集按照 9:1 的比例划分。

4.2.3 PINN 模型训练

4.2.3.1 损失函数构建

本节使用 PINN 模型对低保真数据（LF-data）和高保真数据（HF-data）进行训练。模型架构包含两个多层感知机（MLP）网络：低保真网络 \mathbf{M}_{low} 和高保真网络 \mathbf{M}_{high} 。训练流程如下：

首先低保真数据 X_{low} 首先通过低保真网络得到预测值 \hat{y}_{low} ，与真实低保真标签 y_{low} 计算物理损失。之后高保真数据 X_{high} 通过低保真网络得到中间特征 f_{mid} ，将 f_{mid} 与 X_{high} 拼接后输入高保真网络，得到最终预测值 \hat{y}_{high} 。 \hat{y}_{high} 与真实高保真标签 y_{high} 计算数据损失，并与物理损失拼接得到总损失。

具体计算过程可表示为公式(4-3)和公式(4-4)：

$$\hat{y}_{low} = \mathbf{M}_{low}(X_{low}) \quad (4-3)$$

$$\hat{y}_{high} = \mathbf{M}_{high}(X_{high}, \mathbf{M}_{low}(X_{high})) \quad (4-4)$$

本节使用 Python 语言和 PyTorch 深度学习框架实现模型训练。在优化策略方面，采用 Adam 优化器进行参数更新，并引入基于性能指标的学习率调度机制实现自适应学习率调整。为获得最优的模型结构，本文结合网格搜索和随机搜索两种方法对关键超参数进行系统调优，包括网络层数、每层神经元数量、初始学习率、正则化强度、训练轮次以及批次大小等。

在损失函数选择上，考虑到实验数据中可能存在的异常值和噪声，本文采用鲁棒性较好的 Huber 损失函数，其定义如公式(4-5)所示：

$$L_\delta(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{if } |y - \hat{y}| \leq \delta \\ \delta|y - \hat{y}| - \frac{1}{2}\delta^2 & \text{otherwise} \end{cases} \quad (4-5)$$

Huber 损失函数通过参数 δ 实现对异常值敏感度的动态调节：当 δ 趋近于 0 时，其行为接近平均绝对误差 (MAE)，表现出对异常值的强鲁棒性；当 δ 较大时，其特性近似于均方误差 (MSE)，保持了较高的训练效率。这种自适应特性使得模型能够在保持训练稳定性的同时，有效应对实验数据中的噪声干扰。

本节在基本的 PINN 损失函数构建公式中添加可学习的权重 α ，如公式(4-6)，通过权重 α 平衡损失强度，适应不同的训练任务。

$$L_{PINN} = \underbrace{L_\delta(y_{low}, \hat{y}_{low})}_{\text{低保真物理损失}} + \alpha \cdot \underbrace{L_\delta(y_{high}, \hat{y}_{high})}_{\text{高保真数据损失}} \quad (4-6)$$

4.2.3.2 特征融合细节

本章中 4.2.1 中提到的第一类和第二类数据的 PINN 训练均直接按照 PINN 方案进行训练，第三类数据涉及不同分子量的分子量和组分数据，在训练时分别采用哈达玛积特征融合和注意力特征融合的方法进行特征融合，再进行训练。哈达积融合方法如公式(4-7)，

$$\mathbf{Mn} \circ \mathbf{w} = \begin{bmatrix} M_{n_1} \cdot w_1 \\ M_{n_2} \cdot w_2 \\ \vdots \\ M_{n_k} \cdot w_k \end{bmatrix} \quad (4-7)$$

即将不同的分子量与其对应的组分进行哈达玛积融合，这有助于模型更好地理解特征间的关系，优化训练效果。该操作的本质是构建流变学特征基元，其中高分子链的松弛行为同时受 M_n (链长) 和 w_i (浓度) 调控。

公式(4-8)表示经过哈达玛积融合后的结果，根据 \mathbf{H} 按照公式(4-9)到(4-11)计算 \mathbf{Q} 、 \mathbf{K} 和 \mathbf{V} 。 $\mathbf{Q}(\text{Query})$ 是编码目标流变性能的特征查询， $\mathbf{K}(\text{Key})$ 表征各组分分子量分布的“响应指纹”， $\mathbf{V}(\text{Value})$ 携带原始流变特征的实际物理量级信息。

$$\mathbf{H} = \begin{bmatrix} M_{n_1} \cdot w_1 \\ M_{n_2} \cdot w_2 \\ M_{n_3} \cdot w_3 \end{bmatrix} \quad (4-8)$$

$$\mathbf{Q} = \mathbf{W}_q \mathbf{H}, \quad \mathbf{W}_q \in \mathbb{R}^{d \times 3} \quad (4-9)$$

$$\mathbf{K} = \mathbf{W}_k \mathbf{H}, \quad \mathbf{W}_k \in \mathbb{R}^{d \times 3} \quad (4-10)$$

$$\mathbf{V} = \mathbf{W}_v \mathbf{H}, \quad \mathbf{W}_v \in \mathbb{R}^{d \times 3} \quad (4-11)$$

之后使用公式(4-12)计算注意力分数，注意力分数量化了组分间的流变学相互作用，该机制可自动识别关键组分，例如高分子量组分 ($M_{n_1} > M_{n_2}$) 对模量等的主导作用。

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax} \left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d}} \right) \mathbf{V} \quad (4-12)$$

最后使用公式(4-13)计算最终的注意力输出。

$$\mathbf{Z} = \text{LayerNorm}(\mathbf{H} + \text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V})) \quad (4-13)$$

本节分别使用原始特征的 PINN、哈达玛积特征融合后的 PINN 和注意力特征融合后的 PINN 训练多分子量 PBA 注入制备的 PFGs 数据，得到不同的训练模型。

4.2.4 PINN 模型测试

本节对训练模型进行测试，并分析训练结果，对于第一类数据（单 PBA 流体数据）和第二类数据（单分子量 PBA 注入制备的 PFGs 数据），使用 DNN 和 PINN 两种模型分别训练。保存模型后，使用测试数据进行预测，并分析预测结果。绘制真实值-预测值曲线，残差曲线进行定性分析。计算指标：决定系数 R^2 、平均绝对误差 MAE、平均百分比误差 MAPE 和训练时间 Training Time，并绘制指标对比图进行定量分析。

对于第三类数据（多分子量 PBA 注入制备的 PFGs 数据），本节分别使用 DNN、PINN、哈达玛积特征融合后的 PINN 和注意力特征融合后的 PINN 进行训练，对不同的训练模型进行测试，并分析训练结果，具体分析方法同上。

4.2.5 CVAE 反向建模训练

本节使用条件变分自编码器（CVAE）进行反向建模，将 PFG-b4 的数据作为测试集，其余数据为训练集和验证集，训练集和验证集按照 9：1 划分，输入特征改为 $[\omega, G', G'', \tan\delta]$ 序列，输出标签为制备参数： M_{n_i} 和 w_i 。

本节使用 Python 的 Pytorch 库编写 CVAE 代码，并使用训练数据进行训练，得到模型参数。CVAE 的损失函数公式如公式(4-14)所示，其中 $x = [w, G', G'', \tan\delta]$, $p = M_{n_i}, w_i'$ 。

$$\mathcal{L}(\theta_E, \theta_D) = \mathbb{E}_{q(\theta_E(z|\mathbf{x}, p))} [\log p_{\theta_D}(\mathbf{x}|z, p)] - D_{KL}(q_{\theta_E}(z|\mathbf{x}, p)||p(z|p)) \quad (4-14)$$

CVAE 的训练模型采用 Adam 优化器来进行优化，采用余弦退火算法自动调整学习率，采用网格搜索算法和随机搜索算法对超参数进行调优，其中超参数包括：隐藏层数、隐藏层节点数、学习率、迭代次数、批次大小，隐藏层激活函数采用 ReLU6 激活函数。

4.2.6 CVAE 反向建模测试

本节通过 CVAE 训练频率-模量序列到组分信息的映射，将训练好后的模型保存。生成测试数据时，首先加载保存的模型参数，并输入目标制备条件 $p = M_{n_i}, w_i'$ 作为生成约束；随后从条件先验分布 $p(z|p)$ 中采样潜在变量 z （通过重参数化技巧 $z = \mu_p + \epsilon\sigma_p$ 实现可导性，其中 ϵ 服从标准正态分布），将其与条件变量 p 拼接后输入解码器网络 $p_{\theta_D}(\mathbf{x}|z, p)$ ，生成对应的 $x = [w, G', G'', \tan\delta]$ 。为增强生成多样性，通过潜在空间插值或对条件参数 p 施加微小扰动生成 100 个多模态解。

针对生成的多模态解，本节通过多模态分布分析揭示其内在结构特征，采用核密度估计与箱线图结合的小提琴图可视化方法，呈现不同模态下解的密度分布、峰值位置及离散程度。使用残差分析对预测值与理论解的偏离模式进行系统性检验。使用真实数据误差分析，采用平均绝对百分比误差（MAPE）量化多模态解的整体预测精度。

4.2.7 正逆向联合建模

本节将训练好的 PINN 模型和 CVAE 模型进行正逆向联合建模，将 CVAE 生成的组分信息作为 PINN 模型的输入的特征，验证 CVAE 生成的数据对最终的流变学性质参数的预测效果，并分析 CVAE 生成的数据对 PINN 模型预测效果的提升作用，绘制真实组

分特征与 CVAE 生成的数据输入 PINN 模型后预测的流变学性质参数的对比图，并计算 MAPE 指标，进行定量分析。

4.3 结果与讨论

4.3.1 低保真数据拟合

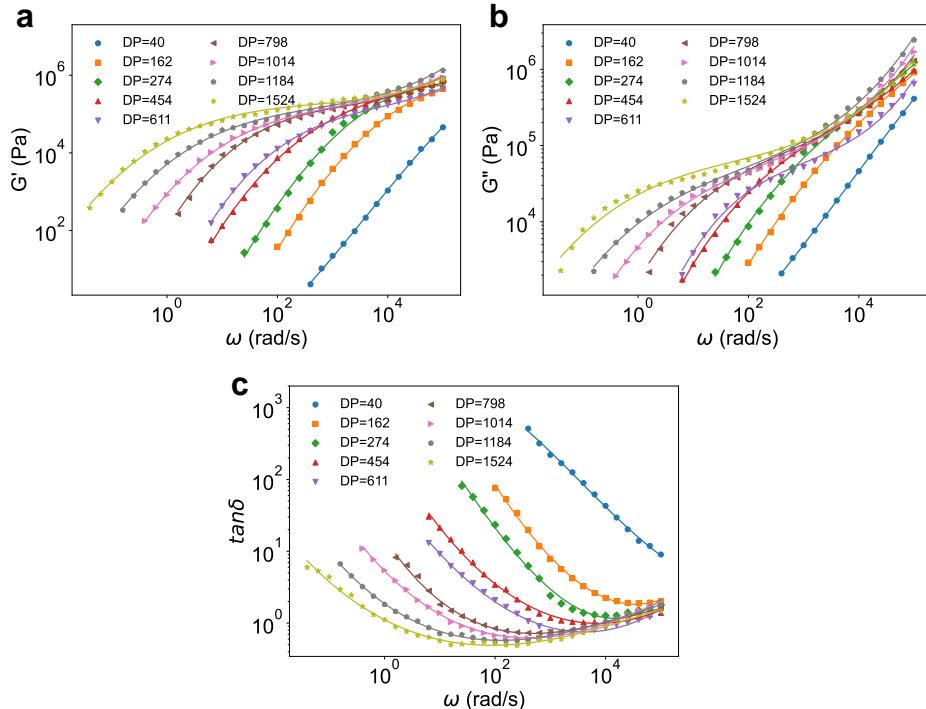


图 4-2 不同 PBA 流体的频率-流变学性质参数数据低保真拟合结果：(a) 不同 DP 值的 PBA 流体的频率-储存模量 (ω - G') 数据拟合结果；(b) 不同 DP 值的 PBA 流体的频率-损耗模量 (ω - G'') 数据拟合结果；(c) 不同 DP 值的 PBA 流体的频率-损耗角正切 (ω - $\tan\delta$) 数据拟合结果

Figure 4-2 Low-fidelity fitting results of frequency-rheological property parameter data of different PBA fluids: (a) Fitting results of frequency-storage modulus (ω - G') data of PBA fluids with different DP values; (b) Fitting results of frequency-loss modulus (ω - G'') data of PBA fluids with different DP values; (c) Fitting results of frequency-loss tangent (ω - $\tan\delta$) data of PBA fluids with different DP values

本节使用 Python 的 `scipy.optimize` 库对真实数据进行最小二乘法拟合，既定函数为公式(4-1)和(4-2)。首先对 PBA 流体进行拟合，训练数据为不同聚合度 (DP) 的 PBA

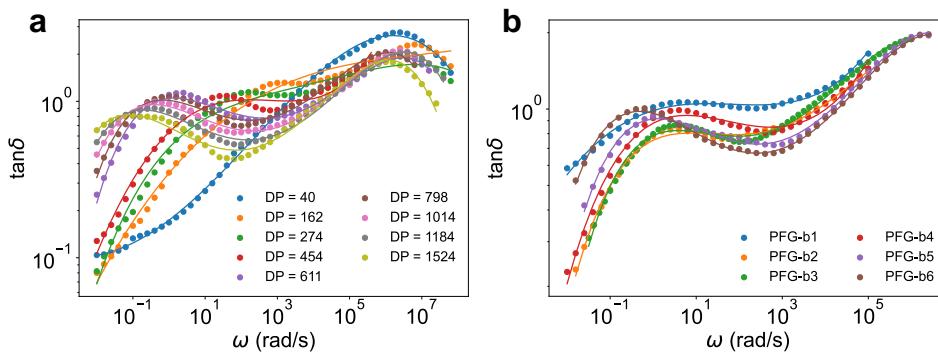


图 4-3

Figure 4-3 不同 PBA 流体组合制备的 PFGs 的频率-流变学性质参数数据低保真拟合结果: (a) 不同 DP 值的 PBA 流体制备的 PFGs 的频率-储存模量 ($\omega\text{-}G'$) 数据拟合结果; (b) 不同 PBA 流体组合制备的 PFGs 的频率-损耗模量 ($\omega\text{-}G''$) 数据拟合结果; (c) 不同 PBA 流体组合制备的 PFGs 的频率-损耗角正切 ($\omega\text{-}\tan\delta$) 数据拟合结果
Low-fidelity fitting results of frequency-rheological property parameter data of PFGs prepared by different PBA fluid combinations: (a) Fitting results of frequency-storage modulus ($\omega\text{-}G'$) data of PFGs prepared by PBA fluids with different DP values; (b) Fitting results of frequency-loss modulus ($\omega\text{-}G''$) data of PFGs prepared by different PBA fluid combinations;
(c) Fitting results of frequency-loss tangent ($\omega\text{-}\tan\delta$) data of PFGs prepared by different PBA fluid combinations

流体的频率-储存模量 ($\omega\text{-}G'$) 数据、频率-损耗模量数据 ($\omega\text{-}G''$) 和频率-损耗角正切 ($\omega\text{-}\tan\delta$) 数据，拟合结果如图4-2所示。从拟合曲线来看，不同 PBA 流体的拟合质量较优，拟合曲线和数据点贴合。表4-2展示了低保真数据的回归拟合指标，其中 g_1 、 g_2 、 g_3 分别对应不同 PBA 流体的 $\omega\text{-}G'$ 、 $\omega\text{-}G''$ 和 $\omega\text{-}\tan\delta$ ，可以看到这三组拟合指标的 R^2 指标都大于 0.95，拟合效果优秀。 g_1 组的 MAPE 值为 19.72%， g_2 组的 MAPE 值为 15.89%， g_3 组的 MAPE 值为 12.45%，根据统计学标准，MAPE 值小于 20% 被认为是可接受误差，这三组均在可接受范围，可以用于后续实验。

之后，针对 PBA 流体注入制备的 PFGs 数据，使用最小二乘法进行拟合。图4-3 (a) 展示了不同 DP 值的注入制备的 PFGs 数据的拟合曲线，拟合曲线和数据点贴合，拟合效果较优。图 (a) 对应的定量指标为表4-2的 g_4 ， R^2 值为 0.985，MAPE 值为 19.91%，均在可接受范围，可以用于后续实验。

图4-3 (b) 展示了多分子量 PBA 注入制备的 PFGs 数据的拟合曲线，其中 b1-b6 是

表 4-2 低保真数据拟合回归指标表

Table 4-2 Regression Metrics of Low-Fidelity Data Fitting

实验分组	R ²	MAE	MAPE(%)
g_1	0.977	456.78	19.72
g_2	0.982	2456.23	15.89
g_3	0.964	5.17	12.45
g_4	0.985	0.25	19.91
g_5	0.962	0.11	9.89

不同的 PBA 流体组合，具体组合如表4-1所示。拟合曲线显示拟合效果优秀，定量指标如表4-2的 g_5 所示，R² 值为 0.962，MAPE 值为 9.89%，误差范围在可接受范围，可以用于后续实验。

4.3.2 单 PBA 流体本构建模

本节首先对不同分子量的单 PBA 流体进行深度学习本构建模，分别建模预测储存模量 (G')、损耗模量 (G'') 和损耗角正切 (tanδ)。图4-4为 G' 的预测结果，图4-4 (a) 为真实-预测值曲线，从曲线图可以看到 PINN 建模的预测曲线相比 DNN 建模的预测曲线更为贴近真实数据点，拟合效果更佳。图4-4 (b) 为残差图，PINN 预测值与真实值的残差总体更为贴近 0 刻度线，极值范围相比 DNN 更小。根据图4-4，可以定性分析出 PINN 相比 DNN 在此项任务上具有更好的泛化预测效果。

图4-5为 G'' 的预测结果，图4-5 (a) 为真实-预测值曲线。从真实-预测值曲线 (图4-5 (a)) 可观察到 PINN 建模的预测值序列与真实值分布趋势呈现更高程度的吻合，而 DNN 的预测曲线在低频率段明显偏高，拟合效果欠佳。进一步通过残差分布图 (图4-5 (b)) 进行评估，发现 PINN 模型的残差绝对值显著低于 DNN 模型的残差绝对值，且 PINN 模型的残差点分布更为集中，更为邻近 0 刻度线。综上所述，可以定性认为在此项任务中，PINN 模型的泛化预测效果更佳。

图4-6为 tanδ 的预测结果，图4-6 (a) 为真实-预测值曲线。从真实-预测值曲线 (图4-6(a)) 可观察到 PINN 建模的预测值序列与真实值分布趋势呈现更高程度的吻合，相较于传统深度神经网络 (DNN) 建模产生的预测曲线展现出更优的轨迹跟踪特性。进一步通过残差分布图 (图4-6(b)) 进行评估，两种模型的预测残差虽均呈现以负偏差为

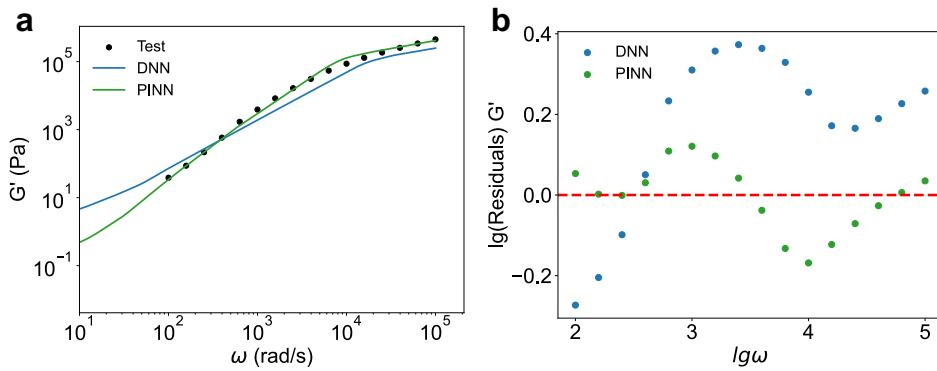


图 4-4 不同分子量的 PBA 流体 PINN 和 DNN 建模的 G' 预测结果: (a) 真实-预测值曲线; (b) 残差图

Figure 4-4 Prediction results of G' of PBA fluids with different DP values using PINN and DNN: (a) Real-predicted value curve; (b) Residual diagram

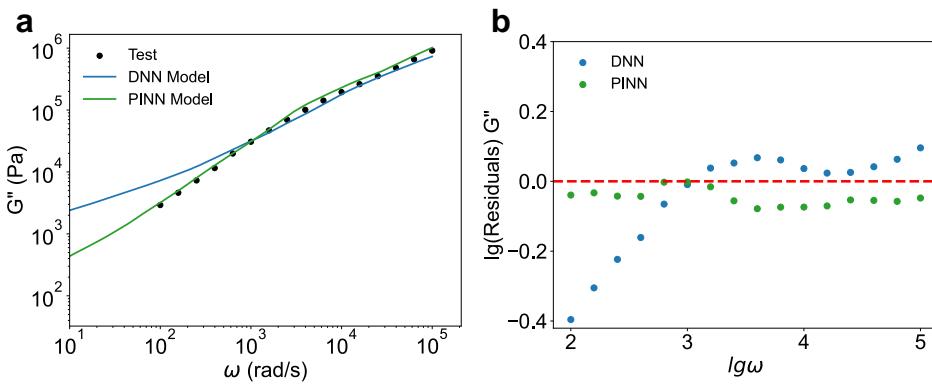


图 4-5 不同分子量的 PBA 流体 PINN 和 DNN 建模的 G'' 预测结果: (a) 真实-预测值曲线; (b) 残差图

Figure 4-5 Prediction results of G'' of PBA fluids with different DP values using PINN and DNN: (a) Real-predicted value curve; (b) Residual diagram

主的分布特征，但 PINN 模型的残差绝对值显著低于 DNN 模型的残差绝对值，极值绝对值较小。为了进一步定量分析 PBA 流体的模量数据 PINN 建模效果，本节计算测试集的 R^2 、MAE、MAPE、Training Time。图4-7为两个模型在 G' 预测、 G'' 预测和 $\tan\delta$ 预测三种不同任务上的指标对比图。从图 4-7 可以看出，PINN 模型在 G' 、 G'' 和 $\tan\delta$ 三个任务上的 R^2 均略高于 DNN 模型。此外，PINN 模型的 MAE 值在所有任务中均小于 DNN 模型，而 MAPE 值均小于 10%，属于较小误差范围。相比之下，DNN 模型的 MAPE 值均大于 10%，表明其误差较大。这些结果表明，PINN 模型在 G' 、 G'' 和 $\tan\delta$

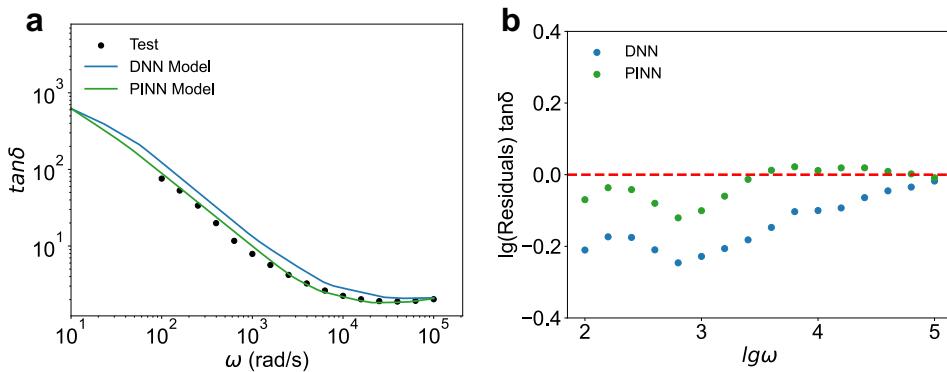


图 4-6 不同分子量的 PBA 流体 PINN 和 DNN 建模的 $\tan\delta$ 预测结果: (a) 真实-预测值曲线; (b) 残差图

Figure 4-6 Prediction results of $\tan\delta$ of PBA fluids with different DP values using PINN and DNN: (a) Real-predicted value curve; (b) Residual diagram

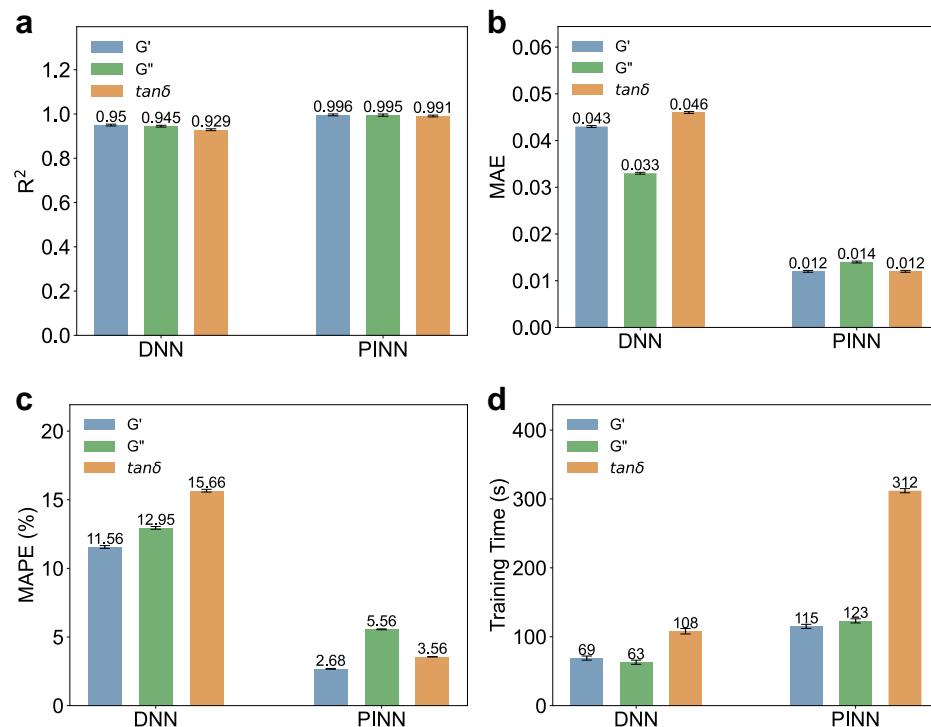


图 4-7 不同分子量的 PBA 流体 PINN 和 DNN 建模的指标对比图: (a) R^2 对比; (b) MAE 对比; (c) MAPE 对比; (d) 训练时间对比

Figure 4-7 Metrics comparison of PINN and DNN modeling for PBA fluids with different molecular weights: (a) R^2 comparison; (b) MAE comparison; (c) MAPE comparison; (d) Training time comparison

三个任务上均具有更低的误差。然而，PINN 模型的训练时间显著更长，约为 DNN 模型的 2-3 倍。

这种差异表明 PINN 通过嵌入物理守恒方程作为正则化约束，有效抑制了 DNN 模型因纯数据驱动导致的过拟合现象，其残差分布的紧致性和对称性改善印证了物理先验知识对模型泛化能力的提升作用。综合可视化分析与统计指标可知，PINN 框架在本研究涉及的偏微分方程反演任务中，通过融合物理机理与数据特征的双重驱动，实现了比传统数据驱动范式更优的泛化预测性能。

4.3.3 单组分 PBA 注入的 PFGs 本构建模

本节对第二类数据，即单一分子量 PBA 注入制备的 PFGs 进行深度学习本构建模，探究在不同分子量 PBA 注入的 PFGs 上，PINN 模型与 DNN 模型的性能对比。图4-8为两种算法模型对这类数据的预测性能对比。图4-8（a）为真实值-预测值曲线，这里对比的是 $\tan\delta$ 的预测结果。从图中可以看出，PINN 模型的预测曲线与真实值曲线更为贴近，拟合效果更佳。图4-8（b）为残差图，PINN 模型的残差绝对值显著低于 DNN 模型的残差绝对值，PINN 的残差更为邻近 0 刻度线。综上所述，可以定性认为在此项任务中，PINN 模型的泛化预测效果更佳。为了进一步定量分析单组分 PBA 注入的 PFGs 数据的

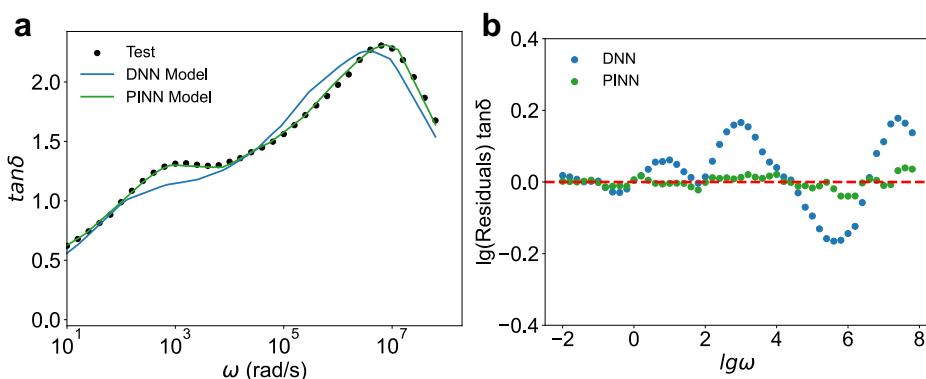


图 4-8 单分子量 PBA 注入的 PFGs 数据 PINN 和 DNN 建模的 $\tan\delta$ 预测结果：(a) PINN 和 DNN 在测试集上的预测值与真实值对比曲线；(b) PINN 和 DNN 在测试集上的预测值残差图

Figure 4-8 Prediction results of $\tan\delta$ of PFGs prepared by single molecular weight PBA injection using PINN and DNN modeling: (a) Comparison curves of predicted vs. true values for PINN and DNN on test set; (b) Residual plots of predicted values for PINN and DNN on test set

PINN 建模效果，本节计算测试集的 R^2 、MAE、MAPE、Training Time。图4-9为两个模型 $\tan\delta$ 预测的指标对比图。从图中可以看出，PINN 模型在该任务上的 R^2 略低于 DNN 模型。在 MAE 指标上，PINN 为 0.012，DNN 为 0.071，PINN 的 MAE 值相比 DNN 降低了 83.38%。MAPE 指标方面，PINN 为 0.03%，DNN 为 8.59%，相差近 3 个数量级别。这里 PINN 模型的 R^2 值意外地低可能是因为 PINN 引入的物理约束可能降低了模型对训练数据的过拟合程度，但模型的泛化能力和实际预测精度（MAPE）更好这些结果表明，PINN 模型在预测单组分 PBA 注入的 PFGs 流变性能时具有更高的准确性和稳定性。然而，与前文类似，PINN 模型的训练时间显著长于 DNN 模型，大约是后者的 5 倍，这是由于物理约束的引入增加了计算复杂度。综上所述，PINN 模型在单组分 PBA 注入的

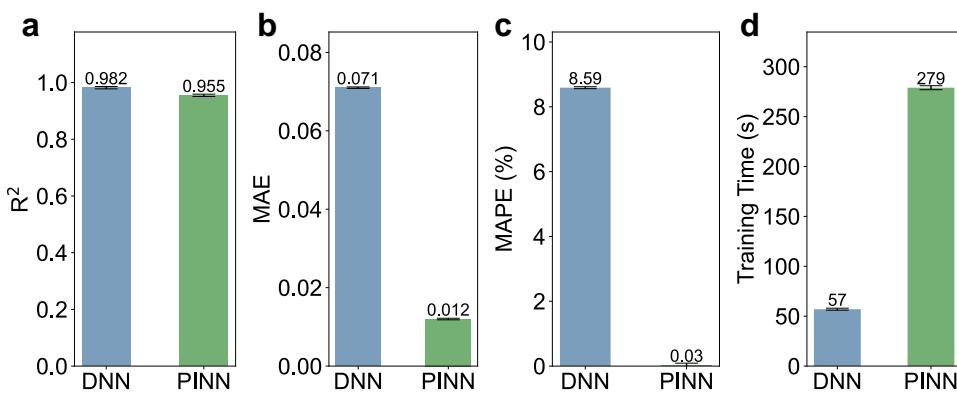


图 4-9 单分子量 PBA 注入的 PFGs 数据 PINN 和 DNN 建模的 $\tan\delta$ 预测指标对比图：(a) R^2 对比；(b) MAE 对比；(c) MAPE 对比；(d) 训练时间对比

Figure 4-9 Comparison of R^2 , MAE, MAPE, and Training Time metrics of PINN and DNN for $\tan\delta$ prediction of PFGs prepared by single molecular weight PBA injection: (a) R^2 comparison; (b) MAE comparison; (c) MAPE comparison; (d) Training Time comparison

PFGs 数据的建模预测中展现出显著优势。从定性分析角度来看，PINN 模型生成的预测曲线与实验数据点更为贴合，残差分布更加集中且接近零值。从定量分析的角度来看，尽管 PINN 的 R^2 指标略低于 DNN 模型，但在 MAE 和 MAPE 这两个更能反映实际预测精度的指标上，PINN 模型都取得了显著优势。特别是在 MAPE 指标上，PINN 与 DNN 指标差近 3 个数量级。这表明 PINN 通过引入物理约束，有效抑制了过拟合现象，提高了模型的泛化能力和预测准确性。然而需要注意的是，由于引入物理约束增加了计算复杂度，PINN 模型的训练时间约为 DNN 模型的 5 倍，这是可以预见的成本增加。

4.3.4 多组分 PBA 注入的 PFGs 本构建模

上一节研究了单组分 PBA 注入的 PFGs 数据建模，验证了 PINN 模型在处理简单制备参数特征（单一分子量）时，通过引入物理约束能够有效提升模型的泛化能力。本节将进一步探讨 PINN 模型在多组分 PBA 注入的 PFGs 数据上的表现，研究其在处理复杂制备参数特征（多分子量组合）时的泛化性能，以期为实际应用提供更有价值的参考。图4-10为多组分 PBA 注入的 PFGs 数据建模结果，图4-10 (a) 为真实值-预测值曲

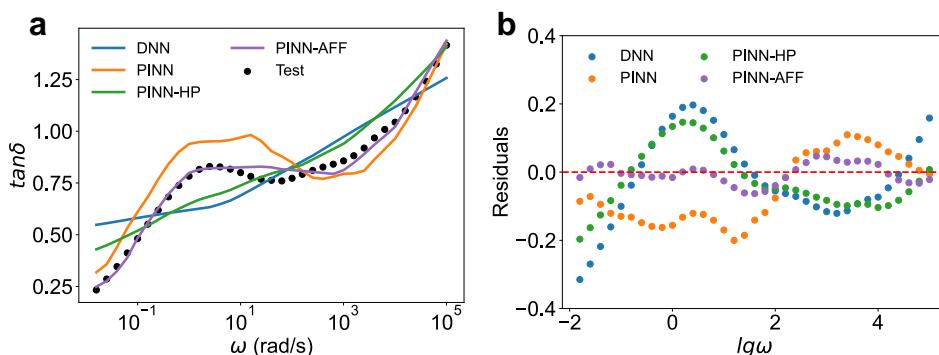


图 4-10 多组分 PBA 注入制备的 PFGs 数据不同算法建模的 $\tan\delta$ 预测结果：(a) 不同算法在测试集上的预测值与真实值对比曲线；(b) 不同算法在测试集上的预测值残差图

Figure 4-10 Prediction results of $\tan\delta$ of PFGs prepared by multiple molecular weight PBA injection using different algorithms: (a) Comparison curves of predicted vs. true values on test set; (b) Residual plots of predicted values on test set

线，图4-10 (b) 为残差图。本项预测建模任务一共使用 4 组不同的算法建模，分别是普通 DNN、普通 PINN、采用哈达玛积特征融合的 PINN (PINN-HP) 和采用注意力特征融合的 PINN (PINN-AFF)。图4-10显示 PINN-AFF 的预测效果最佳，预测曲线与真实值曲线最为贴近，残差分布最为集中且接近 0 刻度线。而其他算法模型的预测存在不同类型的偏差。

为了进一步定量分析多组分 PBA 注入的 PFGs 数据的 PINN 建模效果，本节计算了测试集的 R^2 、MAE、MAPE 和训练时间等评价指标。图4-11展示了 4 种不同算法模型在 $\tan\delta$ 预测任务上的性能对比。从图中可以看出，在所有评价指标上，注意力特征融合的 PINN (PINN-AFF) 模型均表现最佳，其次是哈达玛积特征融合的 PINN (PINN-HP) 模型，再次是普通 PINN 模型，而传统 DNN 模型的表现最差。这一结果表明，引入特征融合机制能够有效提升 PINN 模型的预测性能，其中基于注意力机制的特征融合方

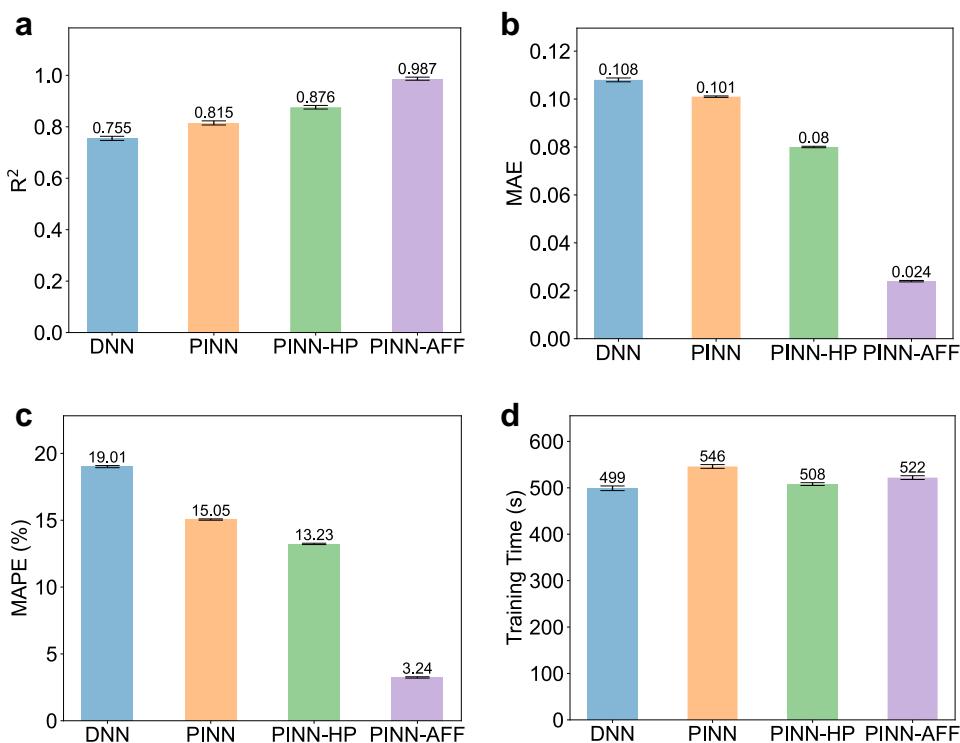


图 4-11 多组分 PBA 注入制备的 PFGs 数据不同算法建模的 $\tan\delta$ 预测指标对比图：(a) R^2 对比；(b) MAE 对比；(c) MAPE 对比；(d) 训练时间对比

Figure 4-11 Comparison of R^2 , MAE, MAPE, and Training Time metrics of PINN and DNN for $\tan\delta$ prediction of PFGs prepared by multiple molecular weight PBA injection: (a) R^2 comparison; (b) MAE comparison; (c) MAPE comparison; (d) Training Time comparison

案尤为有效。从时间指标上看，PINN 的训练时间略大于 DNN，这与之前的实验一致，PINN-HP 的训练时间略小于 PINN，这是因为 PINN-HP 通过将分子量和组分含量特征融合减少了总的特征数量，减少了训练参数的数量。PINN-AFF 的训练时间又略大于 PINN-HP，这是因为 PINN-AFF 引入了注意力机制，增加了计算复杂度。

综合图4-10和图4-11的分析结果可以看出，在处理特征数量较多且特征之间存在复杂物理关系的情况下，普通 PINN 模型的学习泛化能力相对有限。针对这一问题，本研究提出了两种改进方案。首先，哈达玛积特征融合的 PINN (PINN-HP) 模型通过将分子量和组分含量特征进行哈达玛积运算进行特征融合，这种方式相当于为模型预设了特征之间的物理关联，从而提升了模型的预测效果。在此基础上，注意力特征融合的 PINN (PINN-AFF) 模型进一步引入了注意力机制，对不同分子量与组分含量融合后的联合特征进行注意力加权。这种机制能够有效学习不同高分子链段之间的相互作用关系，通过

注意力分数定量描述组分间的流变学相互作用强度，并能自动识别关键组分的影响程度，例如能够捕捉到高分子量组分对整体流变学性能的主导作用。实验结果表明，这两种改进方案都显著提升了模型性能，其中 PINN-AFF 模型取得了最优的预测效果。

4.3.5 CVAE 组分预测的多模态解分析

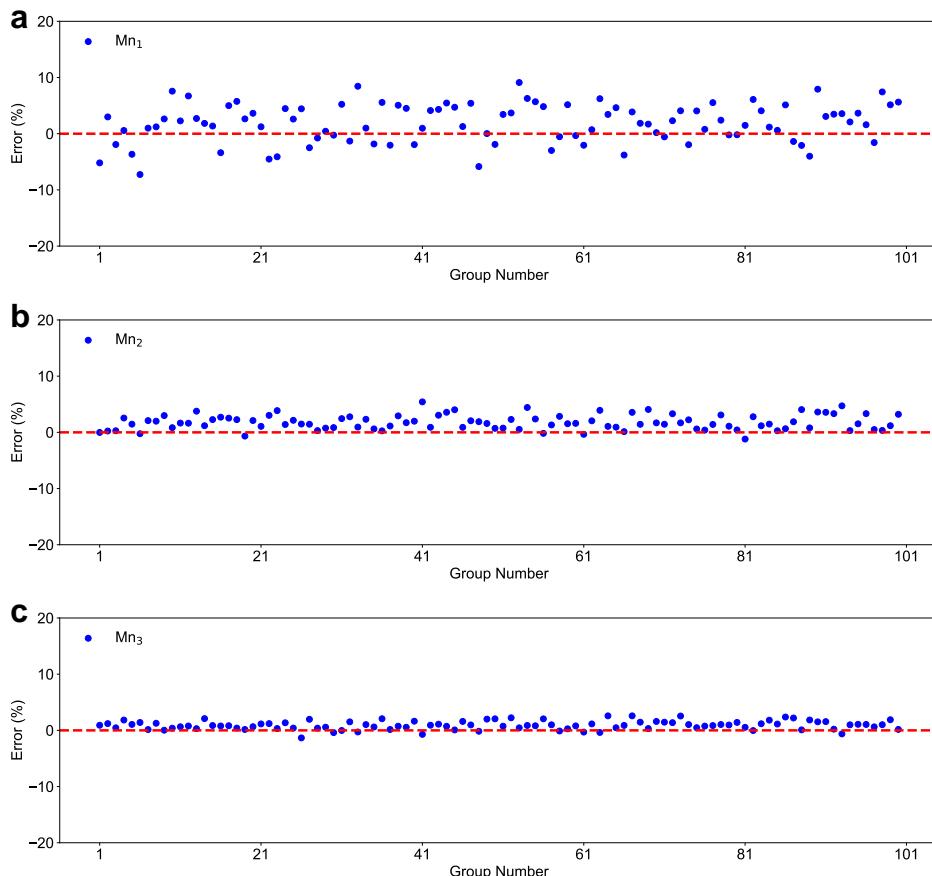


图 4-12 CVAE 生成的 100 组多模态解中不同分子量组分 Mn_1 、 Mn_2 、 Mn_3 的残差分布：(a) Mn_1 残差分布；(b) Mn_2 残差分布；(c) Mn_3 残差分布

Figure 4-12 Residual distributions of 100 multimodal solutions generated by CVAE for different molecular weight components Mn_1 , Mn_2 , Mn_3 : (a) Residual distribution of Mn_1 ; (b) Residual distribution of Mn_2 ; (c) Residual distribution of Mn_3

本节探索了使用 CVAE 模型从材料的流变学性质反向预测其组分配比的可行性。基于训练完成的生成模型，我们输入目标流变学参数 $[\omega, G', G'', \tan\delta]$ ，生成 100 组多模态解，并对生成结果进行系统的残差分析。

图4-12展示了分子量参数 Mn 的残差分布。结果表明，三种分子量组分 Mn_1 、 Mn_2 、 Mn_3 的残差均呈现正态分布特征，最大偏差约为 10%，处于可接受范围内。残差大小依

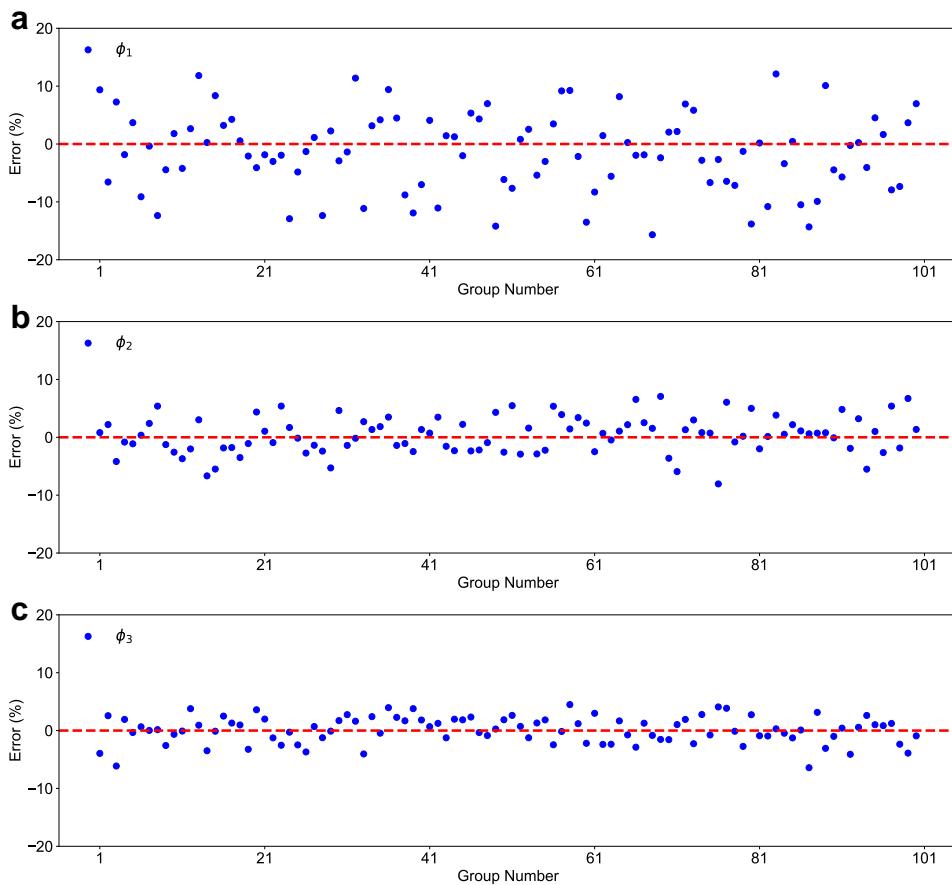


图 4-13 CVAE 生成的 100 组多模态解中不同组分含量 ϕ_1 、 ϕ_2 、 ϕ_3 的残差分布: (a) ϕ_1 残差分布; (b) ϕ_2 残差分布; (c) ϕ_3 残差分布

Figure 4-13 Residual distributions of 100 multimodal solutions generated by CVAE for different component contents ϕ_1 , ϕ_2 , ϕ_3 : (a) Residual distribution of ϕ_1 ; (b) Residual distribution of ϕ_2 ; (c) Residual distribution of ϕ_3

次为 $Mn_1 > Mn_2 > Mn_3$ 。图4-13则显示了组分含量参数 ϕ 的残差分布。 ϕ_1 、 ϕ_2 、 ϕ_3 同样呈现正态分布特征，最大偏差约为 20%，残差大小顺序为 $\phi_1 > \phi_2 > \phi_3$ 。

值得注意的是，在测试集中 $Mn_1 < Mn_2 < Mn_3$ ， $\phi_1 < \phi_2 < \phi_3$ ，由此观察到残差大小与参数实际值呈现负相关关系 - 即参数绝对值越小，其预测误差反而越大。这一现象揭示了模型在预测较小数值参数时的精度局限性，实际应用时，可以通过调整量纲一定程度缓解这一问题。

图4-14为生成数据的 violin 图。(a-f) 分别展示了分子量参数 Mn_1 、 Mn_2 、 Mn_3 和组分含量参数 ϕ_1 、 ϕ_2 、 ϕ_3 的多模态解分布情况。由图可见，真实分子量数据 (Mn) 在生成分子量数据的下四分位数附近，总体生成偏大，大部分数据分布在中位数附近，真实

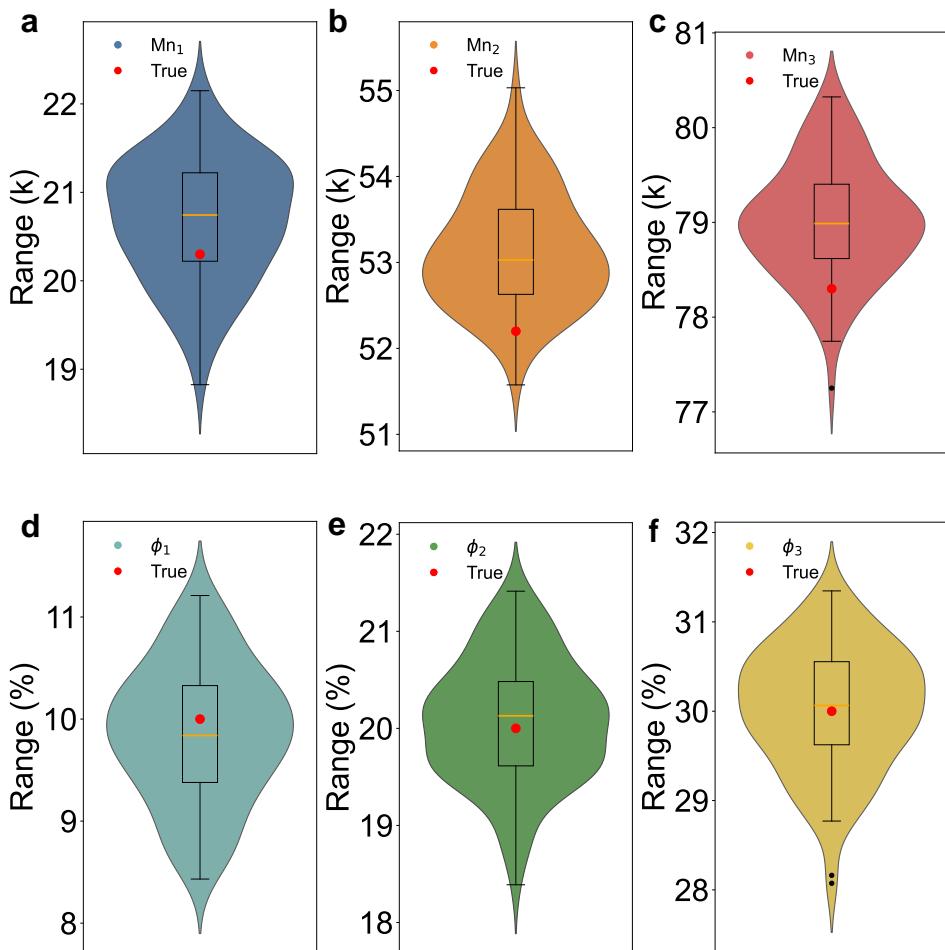


图 4-14 CVAE 生成的 100 组多模态解的 violin 图: (a) 分子量参数 Mn_1 的多模态解分布; (b) 分子量参数 Mn_2 的多模态解分布; (c) 分子量参数 Mn_3 的多模态解分布; (d) 组分含量参数 ϕ_1 的多模态解分布; (e) 组分含量参数 ϕ_2 的多模态解分布; (f) 组分含量参数 ϕ_3 的多模态解分布

Figure 4-14 Violin plots of 100 multimodal solutions generated by CVAE: (a) Distribution of multimodal solutions for molecular weight parameter Mn_1 ; (b) Distribution of multimodal solutions for molecular weight parameter Mn_2 ; (c) Distribution of multimodal solutions for molecular weight parameter Mn_3 ; (d) Distribution of multimodal solutions for component content parameter ϕ_1 ; (e) Distribution of multimodal solutions for component content parameter ϕ_2 ; (f) Distribution of multimodal solutions for component content parameter ϕ_3

组分含量数据 (ϕ) 在生成组分含量数据中位数附近，总体生成效果良好。从核密度曲线的形态来看，所有参数的分布均呈现出中间宽、两端窄的典型小提琴形状，表明数据分布较为集中。整体数据分布服从正态分布，无明显多峰值特征。具体分析各参数的核密度曲线最宽处（即数据最集中处）与中位数线的相对位置： Mn_1 的核密度曲线最宽处位于中位数线上方，表明生成数据相对真实值略有高估； Mn_2 的核密度曲线最宽处位于中位数线下方，显示生成数据略低于中位数水平；而其余各参数 (Mn_3 、 ϕ_1 、 ϕ_2 、 ϕ_3) 的核密度曲线最宽处与中位数线基本重合，说明这些参数的生成分布更为准确。异常点分析结果显示，所有参数的异常点数量均较少，表明生成数据整体质量较高。

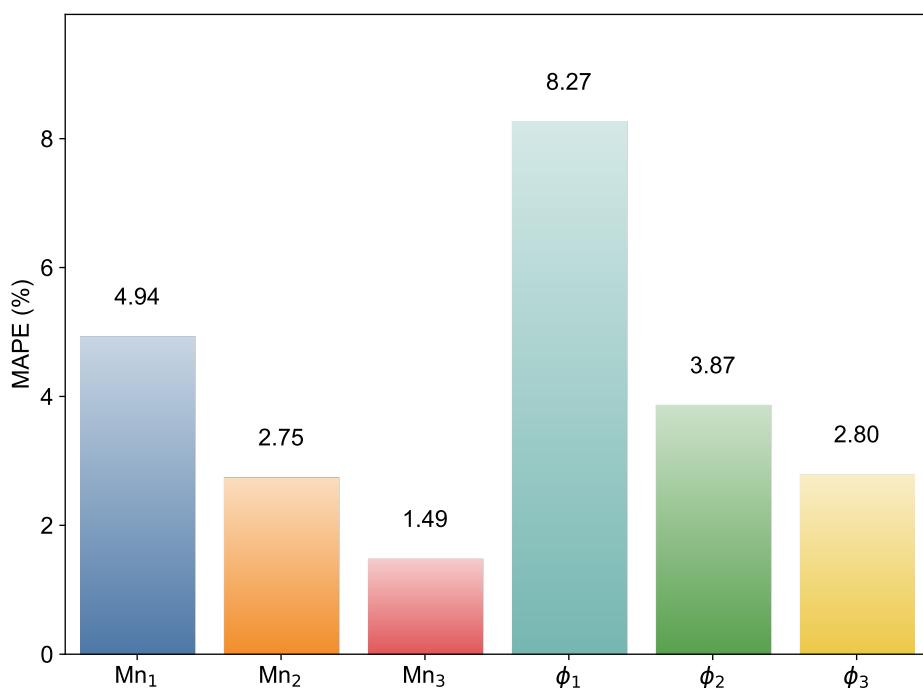


图 4-15 CVAE 生成的 100 组多模态解中分子量参数 Mn_1 、 Mn_2 、 Mn_3 和组分含量参数 ϕ_1 、 ϕ_2 、 ϕ_3 的 MAPE 图

Figure 4-15 MAPE chart of molecular weight parameters Mn_1 , Mn_2 , Mn_3 and component content parameters ϕ_1 , ϕ_2 , ϕ_3 of 100 multimodal solutions generated by CVAE

图4-15为MAPE的条形图。从图中可以看出，分子量参数 Mn_1 、 Mn_2 、 Mn_3 的 MAPE 值分别为 4.94%、2.75%、1.49%，组分含量参数 ϕ_1 、 ϕ_2 、 ϕ_3 的 MAPE 值分别为 8.27%、3.87%、2.80%。这些数据揭示了几个重要的趋势：首先，分子量参数的预测总体优于组分含量参数，这表明模型在预测分子量特征时具有更好的准确性；其次，无论是分子量还是组分含量参数，都呈现出随着数值增大（即 $Mn_1 < Mn_2 < Mn_3$ 和 $\phi_1 < \phi_2 < \phi_3$ ）预测

误差逐渐减小的规律，这与之前残差分析的结果相一致。特别值得注意的是，即使是预测误差最大的 ϕ_1 ，其 MAPE 值也仅为 8.27%，而其他参数的 MAPE 值都控制在 5% 以内，这说明模型整体的预测精度达到了较高水平，具有良好的实用价值。

综合以上分析结果，CVAE 模型在从流变学性质反向预测组分配比的任务中展现出了良好的性能。通过残差分析、violin 图分析和 MAPE 定量分析，我们可以得出以下主要结论：首先，模型生成的组分配比数据整体呈现正态分布特征，预测结果的分布集中且稳定，异常值较少；其次，预测误差随着参数数值的增大而减小，这一特征在分子量和组分含量两类参数中均有体现。这种误差分布特征可能源于在数据标准化过程中，较小数值参数更容易受到数值舍入和归一化误差的影响。第三，从 MAPE 指标来看，分子量参数的预测精度普遍优于组分含量参数，但两类参数的预测误差均控制在可接受范围内（最大 MAPE 为 8.27%）。这些结果表明，CVAE 模型能够有效地实现从材料性质到材料制备参数的反向设计，为高分子材料配方开发提供了一种可靠的数据驱动方法。

4.3.6 正逆向联合建模分析

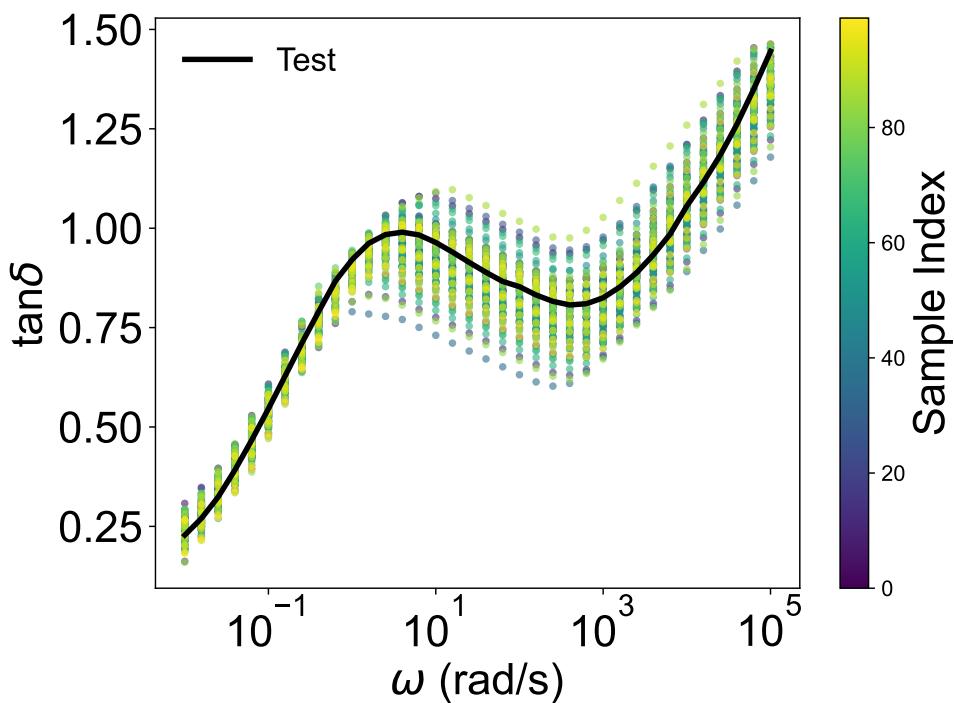


图 4-16 100 组 CVAE 生成的多模态解输入 PINN 模型后的预测值曲线

Figure 4-16 Prediction curves of 100 multimodal solutions generated by CVAE after input into PINN model

上一节通过分析 CVAE 生成的 100 组生成组分数据与真实组分数据的误差来确定

CVAE 模型的生成效果，但是这种分析方法存在一定的局限性。首先，组分数据与频率-损耗角正切曲线之间存在多对一的映射关系，即理论上可能存在多种不同的组分配比，制备出来的材料具有相同或近似的频率-损耗角正切曲线。这种多解性是高分子材料设计中的一个普遍现象，源于不同分子量组分之间复杂的相互作用。其次，组分配比的微小变化可能导致材料性能的显著差异，这种非线性关系使得仅通过比较组分数据的误差来评估模型效果是不够全面的。

从实验验证的角度来看，最准确的分析方法应该是按照生成的组分配比分别制备材料样品，通过 DMA 频率扫描实验测试其损耗角正切，并与目标曲线进行对比。然而，考虑到实验成本和时间效率，这种方法在实际操作中并不现实。为了在保证验证可靠性的同时提高效率，本节采用了一种替代方案：将 CVAE 模型生成的 100 组组分特征输入到已训练好的 PINN 模型中，通过 PINN 的预测结果来评估 CVAE 生成结果的质量。这种方法虽然不能完全替代实验验证，但可以在一定程度上反映生成组分的合理性。

图4-16展示了这 100 组特征输入 PINN 后得到的频率-损耗角正切曲线，其中 Test 线代表原始真实特征对应的预测曲线。从图中可以观察到，这 100 组预测特征绘制的曲线形成了一个连续的曲线带，且具有以下特点：首先，曲线带的整体趋势与 Test 线高度一致，表明生成的组分特征能够较好地捕捉材料的主要流变学特性；其次，Test 线位于曲线带的中心区域，说明生成结果的分布是合理的；最后，曲线带的宽度随频率变化呈现出不均匀的特征，这反映了在不同频率区间，组分配比对材料性能的影响程度是不同的。这种曲线带的分布特征也从另一个角度验证了 CVAE 模型在材料组分反演任务中的有效性。

通过建立 PINN-CVAE 的联合建模，可以建立一个自适应增强的流变学机器学习系统。该系统的工作流程如下：首先，PINN 模型通过物理约束和特征融合机制对流变学数据进行正向建模，建立组分特征到流变学性质的映射关系。在实际应用过程中，通过真实实验获得的新数据可以不断补充到训练集中，进一步提升 PINN 模型的预测精度。与此同时，CVAE 模型负责从目标流变学性质反向生成可能的组分配比方案，为材料设计提供多样化的参考。这些由 CVAE 生成的组分方案可以输入到 PINN 模型中进行快速评估和筛选，通过比较预测曲线与目标曲线的吻合度，筛选出最具潜力的组分配比进行实验验证，从而大大减轻实验负担。最后，实验验证结果又可以作为新的训练数据反馈给两个模型，形成一个不断优化的闭环系统。这种联合建模方法充分发挥了两种模型的优势，既保证了预测的物理合理性，又提供了材料设计的多样化方案，同时通过数据驱

动和实验验证的结合，实现了系统性能的持续提升。

4.4 本章小结

本章主要围绕 PINN 与 CVAE 两类深度学习模型在高分子流变学数据建模及组分反演中的应用进行了研究。本章首先对比了传统 DNN 模型与基于物理先验约束的 PINN 模型在处理高维数据与复杂特征时的性能差异。在 PINN 建模中使用自适应权重机制，首先对 PBA 流体和单 PBA 注入制备的 PFGs 流体进行建模，结果表明 PINN 相比 DNN 具有更好的泛化预测效果。实验结果显示，普通 PINN 模型在一定程度上能够提高预测精度，但其泛化能力在特征数量较多且分布稀疏场景时表现一般，具体为在预测多组分 PBA 注入制备的 PFGs 流体时，模型预测效果较差。为此，本章进一步提出了两种特征融合改进方案——哈达玛积特征融合（PINN-HP）和注意力特征融合（PINN-AFF），通过预设和学习输入特征之间的物理关联，有效地改善了模型对各类复杂非线性关系的捕捉能力，从而解决特征稀疏的问题。从 R^2 、MAE、MAPE 多项评价指标来看，PINN-AFF 模型在预测准确性和稳定性上均明显优于传统 DNN 和普通 PINN 模型，尽管其训练时间略长，但整体优势十分明显，这为流变学数据的高精度预测提供了新思路。

在 PINN 模型取得显著进展的同时，本章还深入探讨了利用 CVAE 模型进行材料组分反演的可行性。该部分工作主要通过将目标流变学参数（如频率 ω 、 G' 、 G'' 和 $\tan\delta$ ）作为输入，生成多模态解，并对生成数据的残差分布、核密度及异常点进行了详细分析。结果表明，无论是分子量参数 (M_n) 还是组分含量参数 (ϕ)，生成数据均基本呈现正态分布特征，且随着参数数值的增大，预测误差逐步减小，MAPE 指标均控制在较低水平，验证了模型在反向预测材料配方比例方面的高精度和鲁棒性。这一研究内容解决了传统回归方法单一输出的问题，而且为基于数据驱动的材料设计提供了具有多解性的参考方案，在真实的实验应用中可以使用这些多模态解来辅助调整实验设计，从而提高实验效率。

此外，本章还对各模型在训练时间和资源消耗上的表现进行了比较和分析，指出尽管理论上 PINN 类模型由于引入了物理约束及复杂特征融合模块而导致训练时间较长，但其在捕捉复杂物理特性和有效泛化方面的优势足以弥补这一不足。同时，通过对比分析，可以发现深度学习模型在处理小数值参数时仍存在归一化误差，未来可通过调整量纲转换和数据预处理进一步优化模型效果。

最后，本章设计了 PINN-CVAE 联合建模方法，通过将 PINN 的正向预测能力与

CVAE 的反向生成能力相结合，构建了一个自适应增强的流变学机器学习系统。该系统利用 PINN 模型进行正向建模，通过物理约束和特征融合机制建立组分特征到流变学性质的映射关系；同时利用 CVAE 模型进行反向生成，从目标流变学性质生成多种可能的组分配比方案。两个模型相互配合，形成闭环优化系统：CVAE 生成的组分方案可通过 PINN 快速评估筛选，实验验证结果又可作为新数据反馈给模型进行训练。这种联合建模方法既保证了预测的物理合理性，又提供了材料设计的多样化方案，为高分子材料的智能设计提供了新思路。

综合来看，本章对 PINN 系列模型和 CVAE 模型进行了初步构建、改进与验证，展示了深度学习在流变学数据预测和材料组分反演中的应用潜力。虽然目前的结果仍存在不足，但实验数据为后续高分子材料设计和流变特性模拟提供了一定的理论依据和参考。总体而言，本章工作初步验证了结合物理信息与深度学习方法的思路，并为未来进一步改进材料配方设计提出了一些启示。

总结与展望

本文以流变学本构建模为研究对象，系统探索了深度学习方法在流变学本构建模中的应用前景。研究内容主要包含两个方面：第一个方面聚焦于深度学习模型结构对流变学本构建模的适用性研究。考虑到黏弹性材料流变特性中的应变历史依赖性，本文创新性地引入了天然适合处理时序数据的循环神经网络，特别是采用了门控循环单元（GRU），并与传统深度神经网络（DNN）进行了系统对比分析。第二个方面是在真实实验数据上，使用物理信息神经网络（PINN）和条件变分自编码器（CVAE）两种方法对流变学本构模型分别进行正逆向训练，通过引入注意力特征融合机制，探讨了在实验数据有限，特征信息稀疏的情况下使用 PINN 建模以预测材料流变学性质的可行性。而 CVAE 的引入则探讨了另一个材料科研中的重要问题，即如何通过期望的流变学性质反推出材料的制备参数来辅助材料设计。

在第一部分的研究工作中，本文首先采用了数值模拟与深度学习相结合的方式。首先通过数值模拟生成了包括 Herschel-Bulkley 模型、Maxwell 模型、Doi-Edwards 模型和 Giesekus 模型在内的多种典型流变学模型的数据。这些模型涵盖了从简单到复杂、从线性到非线性的不同类型本构关系，为深度学习方法的验证提供了全面的数据基础。在深度学习方面，本文重点比较了传统深度神经网络（DNN）和门控循环单元（GRU）两种算法的建模效果。研究发现，GRU 算法凭借其独特的门控机制，在处理具有时间依赖性的流变学数据时展现出显著优势。在模型泛化能力方面，本文通过交变应变预测和线性应变预测两类任务进行了验证。之后本文对一类阻尼黏弹性材料的应力应变数据进行本构建模，在 GRU 模型中引入物理本构残差，构建物理信息门控循环单元（PI-GRU）模型，通过物理约束来增强模型的泛化能力，同时保持对材料动态行为的准确描述。

结果显示，PI-GRU 模型表现出更强的泛化能力，能够准确预测不同应变条件下的材料响应。这一发现对于实际应用具有重要意义，说明深度学习模型不仅能够进行简单的数据拟合，还能够捕捉材料本构关系的本质特征。然而，研究也发现 GRU 模型的优势是以更高的计算成本为代价的，其训练时间和资源消耗都显著高于 DNN。

第二部分的研究工作中，本文以高分子流变学数据建模及组分反演为研究对象，分别使用 PINN 和 CVAE 两类深度学习模型进行正逆向建模。在正向建模方面，研究发现普通 PINN 模型虽然相比传统 DNN 具有更好的泛化预测效果，但在特征数量较多且分布稀疏的场景下表现一般。为此，本文提出了哈达玛积特征融合（PINN-HP）和注

意力特征融合（PINN-AFF）两种改进方案，通过预设和学习输入特征之间的物理关联，有效改善了模型对复杂非线性关系的捕捉能力。从 R^2 、MAE、MAPE 等评价指标来看，PINN-AFF 模型在预测准确性和稳定性上均明显优于传统 DNN 和普通 PINN 模型。在逆向建模方面，本文利用 CVAE 模型从目标流变学参数反向生成材料组分配比，结果表明生成的分子量参数和组分含量参数均呈现正态分布特征，最大误差控制在 10% 以内。最后，本文设计了 PINN-CVAE 联合建模方法，通过将 PINN 的正向预测能力与 CVAE 的反向生成能力相结合，构建了一个自适应增强的流变学机器学习系统，为高分子材料的智能设计提供了新思路。

总的来说，本文的研究工作为深度学习方法在流变学本构建模中的应用提供了创新性的探索。在时间域流变学数据（应力、应变率-应力）的研究中，本文创新性地引入了更擅长处理时序数据的 GRU 模型，在频率域流变学数据（频率-储存模量、损耗模量、损耗角正切）的研究中，本文在传统 PINN 的基础上提出了注意力特征融合的优化方案，并首次尝试了基于 CVAE 的逆向训练来辅助材料设计。基于本文的研究成果，未来还可以在以下几个方向进行深入研究：

- (1) 针对黏弹性流体的非线性本构关系，本文采用的 GRU 模型在处理长时间尺度的非线性本构关系时存在长期依赖问题。在实验条件允许且具备充足高质量流变学数据的情况下，可以考虑引入基于自注意力机制的 Transformer 架构。该架构通过多头注意力机制和位置编码，能够更好地处理长序列数据，有望捕捉材料在更长时间尺度下的流变学特性，尤其是在应力松弛和蠕变等长期行为的建模中具有潜在优势。此外，还可以探索将 Transformer 与物理约束相结合的方法，通过在注意力层中引入物理先验知识来增强模型对材料本构关系的理解。
- (2) 可以考虑引入多尺度建模方法，将分子动力学模拟与宏观流变学行为建模相结合。通过分子动力学模拟获取材料微观结构演化信息，结合宏观流变实验数据，构建跨尺度的深度学习模型。这种多模态数据训练方式不仅可以从理论和实验两个角度提高模型的泛化能力，还能帮助揭示材料微观结构与宏观流变性质之间的关联机制。同时，可以探索使用图神经网络等新型架构来处理分子结构数据，进一步提升模型对材料结构-性能关系的表征能力。
- (3) 在 PINN-CVAE 联合建模框架的基础上，可以进一步探索引入强化学习方法，构建端到端的材料配方优化系统。通过将 PINN 的物理约束预测能力、CVAE 的多模态生成能力与强化学习的决策优化能力相结合，可以实现从目标性能出

发，自动搜索和优化材料配方的智能设计流程。这种方法有望大幅提高材料开发效率，减少试错成本。同时，还可以考虑在优化过程中引入多目标约束，平衡材料性能、成本和工艺等多个维度的需求。

参考文献

- [1] Dealy J M, Wissbrun K F. Introduction to Rheology[M]//Melt Rheology and Its Role in Plastics Processing: Theory and Applications. Boston, MA: Springer US, 1990: 1-41.
- [2] Ellero M, Housiadas K D, Phan-Thien N. Tanner: 90 Years of Rheology[J]. Physics Of Fluids, 2024, 36(11): 110401.
- [3] Ewoldt R H, Saengow C. Designing Complex Fluids[J]. Annual Review of Fluid Mechanics, 2022, 54(1): 413-441.
- [4] Barnes H A. Thixotropy—a Review[J]. Journal of Non-Newtonian Fluid Mechanics, 1997, 70(1): 1-33.
- [5] Banerjee R, Ray S S. Role of Rheology in Morphology Development and Advanced Processing of Thermoplastic Polymer Materials: A Review[J]. ACS Omega, 2023, 8(31): 27969-28001.
- [6] Zener C M, Siegel S. Elasticity and Anelasticity of Metals.[J]. The Journal of Physical and Colloid Chemistry, 1949, 53(9): 1468-1468.
- [7] Hajikarimi P, Sadat Hosseini A. Viscoelasticity Theoretical Background[M]//Constructional Viscoelastic Composite Materials: Theory and Application. Singapore: Springer Nature Singapore, 2023: 9-41.
- [8] Sun H, Jiang Y, Zhang Y, et al. A Review of Constitutive Models for Non-Newtonian Fluids[J]. Fractional Calculus and Applied Analysis, 2024, 27(4): 1483-1526.
- [9] Kuschel S, Ho P J, Al Haddad A, et al. Non-Linear Enhancement of Ultrafast X-ray Diffraction through Transient Resonances[J]. Nature Communications, 2025, 16(1): 847.
- [10] Sun R, Yang J, Patil S, et al. Relaxation dynamics of deformed polymer nanocomposites as revealed by small-angle scattering and rheology[J]. Soft Matter, 2022, 18(46): 8867-8884.
- [11] Ling S, Wu Z, Mei J. Comparison and Review of Classical and Machine Learning-Based Constitutive Models for Polymers Used in Aeronautical Thermoplastic Composites[J]. Reviews On Advanced Materials Science 2023, 62(1): 20230107.
- [12] Martín-Roca J, Bianco V, Alarcón F, et al. Rheology of *Pseudomonas fluorescens* biofilms: From experiments to predictive DPD mesoscopic modeling[J]. The Journal of

- Chemical Physics, 2023, 158(7): 074902.
- [13] Jeon E, Kim H, Kim G, et al. A review of bacterial biofilm formation and growth: rheological characterization, techniques, and applications[J]. Korea-Australia Rheology Journal, 2023, 35(4): 267-278.
- [14] Campbell G A, Zak M E, Wetzel M D. Newtonian, Power Law, and Infinite Shear Flow Characteristics of Concentrated Slurries Using Percolation Theory Concepts[J]. Rheologica Acta, 2018, 57(3): 197-216.
- [15] Schreuders F K, Sagis L M, Bodnár I, et al. Non-linear rheology reveals the importance of elasticity in meat and meat analogues[J]. Scientific reports, 2022, 12(1): 1334.
- [16] Kim J, Jeong E H, Baik J H, et al. The Role of Rheology in Cosmetics Research: A Review [J]. Korea-Australia Rheology Journal, 2024, 36(4): 271-282.
- [17] Murch W L, Spiridigliozi J, Heller A, et al. Non-invasive, continuous oral delivery of solid levodopa-carbidopa for management of Parkinson's disease[J]. Scientific Reports, 2024, 14(1): 26826.
- [18] Zhang H, Li A, Su Y, et al. Modification Technologies and Constitutive Models of Viscoelastic Damping Materials: Progress and Future Trends[J]. Construction and Building Materials, 2024, 441: 137406.
- [19] Wang L, Martínez J A I, Ulliac G, et al. Non-reciprocal and non-Newtonian mechanical metamaterials[J]. Nature Communications, 2023, 14(1): 4778.
- [20] Lowe G D, Forbes C. Rheology of cardiovascular disease[J]. Clinical blood rheology, 2019: 113-140.
- [21] Ricarte R G, Shanbhag S. A Tutorial Review of Linear Rheology for Polymer Chemists: Basics and Best Practices for Covalent Adaptable Networks[J]. Polymer Chemistry, 2024, 15(9): 815-846.
- [22] Yao D, Zatloukal M. Inelastic Fluid Models with an Objective Stretch Rate Parameter[J]. Journal of Non-Newtonian Fluid Mechanics, 2024, 334: 105320.
- [23] Song J K, Holten-Andersen N, McKinley G H. Non-Maxwellian Viscoelastic Stress Relaxations in Soft Matter[J]. Soft Matter, 2023, 19(41): 7885-7906.
- [24] Maxwell J C. IV. On the dynamical theory of gases[J]. Philosophical transactions of the Royal Society of London, 1867(157): 49-88.

- [25] Voigt W. Ueber Innere Reibung Fester Körper, Insbesondere Der Metalle[J]. Annalen der Physik, 1892, 283(12): 671-693.
- [26] Bagley R L, Torvik P J. On the fractional calculus model of viscoelastic behavior[J]. Journal of Rheology, 1986, 30(1): 133-155.
- [27] Boltzmann L. Zur Theorie Der Elastischen Nachwirkung[J]. Annalen der Physik, 1878, 241(11): 430-432.
- [28] Fedorowicz K, Prosser R. The Elastic Perfectly Plastic Constitutive Equation for Yield Stress Fluids[J]. Journal Of Non-Newtonian Fluid Mechanics, 2024, 325: 105201.
- [29] Bingham E C. An investigation of the laws of plastic flow[J]. Bulletin of the Bureau of Standards, 1916, 13: 309-353.
- [30] Herschel W H, Bulkley R. Konsistenzmessungen von gummi-benzollösungen[J]. Kolloid-Zeitschrift, 1926, 39: 291-300.
- [31] Magnon E, Cayeux E. Precise method to estimate the herschel-bulkley parameters from pipe rheometer measurements[J]. Fluids, 2021, 6(4): 157.
- [32] Papanastasiou T C. Flows of materials with yield[J]. Journal of rheology, 1987, 31(5): 385-404.
- [33] Oldroyd J. Non-Newtonian effects in steady motion of some idealized elastico-viscous liquids[J]. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 1958, 245(1241): 278-297.
- [34] Giesekus H. A simple constitutive equation for polymer fluids based on the concept of deformation-dependent tensorial mobility[J]. Journal of Non-Newtonian Fluid Mechanics, 1982, 11(1-2): 69-109.
- [35] Peng S, Li J y, Xiong Y l, et al. Numerical simulation of two-dimensional unsteady Giesekus flow over a circular cylinder[J]. Journal of Non-Newtonian Fluid Mechanics, 2021, 294: 104571.
- [36] Kim S K. Viscosity model based on Giesekus equation[J]. Applied Rheology, 2024, 34(1): 20240004.
- [37] Fattal R, Kupferman R. Constitutive laws for the matrix-logarithm of the conformation tensor[J]. Journal of Non-Newtonian Fluid Mechanics, 2004, 123(2-3): 281-285.
- [38] Thien N P, Tanner R I. A new constitutive equation derived from network theory[J].

- Journal of Non-Newtonian Fluid Mechanics, 1977, 2(4): 353-365.
- [39] Bird R B, Dotson P J, Johnson N. Polymer solution rheology based on a finitely extensible bead—spring chain model[J]. Journal of Non-Newtonian Fluid Mechanics, 1980, 7(2-3): 213-235.
- [40] Qi H, Xu M. Stokes' first problem for a viscoelastic fluid with the generalized Oldroyd-B model[J]. Acta Mechanica Sinica, 2007, 23(5): 463-469.
- [41] Kaye A. Non-Newtonian flow in incompressible fluids[J]. College of Aeronautics Note 134 & 149, 1962.
- [42] Bernstein B, Kearsley E, Zapas L. A study of stress relaxation with finite strain[J]. Transactions of the Society of Rheology, 1963, 7(1): 391-410.
- [43] Mitsoulis E, Hatzikiriakos S G. 60 Years of the Kaye-Bernstein, Kearsley, Zapas Rheological Constitutive Law for Polymers[J]. Physics Of Flow, 2023, 35(10): 103111.
- [44] Doi M, Edwards S. Dynamics of concentrated polymer systems. Part 1.—Brownian motion in the equilibrium state[J]. Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics, 1978, 74: 1789-1801.
- [45] Doi M, Edwards S. Dynamics of concentrated polymer systems. Part 2.—Molecular motion under flow[J]. Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics, 1978, 74: 1802-1817.
- [46] Doi M, Edwards S. Dynamics of concentrated polymer systems. Part 3.—The constitutive equation[J]. Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics, 1978, 74: 1818-1832.
- [47] Doi M, Edwards S. Dynamics of concentrated polymer systems. Part 4.—Rheological properties[J]. Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics, 1979, 75: 38-54.
- [48] O'Connor N P T, Ball R C. Confirmation of the Doi-Edwards model[J]. Macromolecules, 1992, 25: 5677-5682.
- [49] Hassager O, Hansen R. Constitutive equations for the Doi-Edwards model without independent alignment[J]. Rheologica acta, 2010, 49: 555-562.
- [50] Chupin L. Mathematical Existence Results for the Doi-Edwards Polymer Model[J]. Archive for Rational Mechanics and Analysis, 2017, 223: 1-55.

- [51] Alves M, Oliveira P, Pinho F. Numerical Methods for Viscoelastic Fluid Flows[J]. Annual Review of Fluid Mechanics, 2021, 53(1): 509-541.
- [52] Stadler F J. What Are Typical Sources of Error in Rotational Rheometry of Polymer Melts?[J]. Korea-Australia Rheology Journal, 2014, 26(3): 277-291.
- [53] Zhai X, Chen Z M. Global Well-Posedness to the n-Dimensional Compressible Oldroyd-B Model Without Damping Mechanism[J]. Journal of Dynamics and Differential Equations, 2024, 36(2): 1405-1433.
- [54] Li J, Si Z, Shang K, et al. Coupling effect of LDPE molecular chain structure and additives on the rheological behaviors of cable insulating materials[J]. Polymers, 2023, 15(8): 1883.
- [55] Sgouros A, Megariotis G, Theodorou D. Slip-spring model for the linear and nonlinear viscoelastic properties of molten polyethylene derived from atomistic simulations[J]. Macromolecules, 2017, 50(11): 4524-4541.
- [56] Sato T, Yoshimoto K. Recent Developments on Multiscale Simulations for Rheology and Complex Flow of Polymers[J]. Korea-Australia Rheology Journal, 2024, 36(4): 253-269.
- [57] Webb M A, Delannoy J Y, De Pablo J J. Graph-based approach to systematic molecular coarse-graining[J]. Journal of chemical theory and computation, 2018, 15(2): 1199-1208.
- [58] Behbahani A F, Schneider L, Rissanou A, et al. Dynamics and rheology of polymer melts via hierarchical atomistic, coarse-grained, and slip-spring simulations[J]. Macromolecules, 2021, 54(6): 2740-2762.
- [59] Morii Y, Kawakatsu T. Lagrangian multiscale simulation of complex flows[J]. Physics of Fluids, 2021, 33(9).
- [60] Amamoto Y. Data-Driven Approaches for Structure-Property Relationships in Polymer Science for Prediction and Understanding[J]. Polymer Journal, 2022, 54(8): 957-967.
- [61] Wang H, Fu T, Du Y, et al. Scientific discovery in the age of artificial intelligence[J]. Nature, 2023, 620(7972): 47-60.
- [62] Uesaka Y, Aizawa T, Ebara T, et al. A theory of learnability[J]. Kybernetik, 1973, 13: 123-131.
- [63] Liu X, Zhang F, Hou Z, et al. Self-supervised learning: Generative or contrastive[J]. IEEE transactions on knowledge and data engineering, 2021, 35(1): 857-876.
- [64] Quinlan J R. Induction of decision trees[J]. Machine learning, 1986, 1: 81-106.

- [65] Breiman L. Random forests[J]. Machine learning, 2001, 45: 5-32.
- [66] Cortes C. Support-Vector Networks[J]. Machine Learning, 1995, 20: 273-297.
- [67] Cover T, Hart P. Nearest neighbor pattern classification[J]. IEEE transactions on information theory, 1967, 13(1): 21-27.
- [68] Yang L, Zhang Z, Song Y, et al. Diffusion models: A comprehensive survey of methods and applications[J]. ACM Computing Surveys, 2023, 56(4): 1-39.
- [69] Xu P, Zhu X, Clifton D A. Multimodal learning with transformers: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(10): 12113-12132.
- [70] Xie Y, Xu Z, Zhang J, et al. Self-Supervised Learning of Graph Neural Networks: A Unified Review[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(2): 2412-2429.
- [71] Zhu Y, Wu Y, Sebe N, et al. Vision + X: A Survey on Multimodal Learning in the Light of Data[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12): 9102-9122.
- [72] Choudhary K, DeCost B, Chen C, et al. Recent advances and applications of deep learning methods in materials science[J]. npj Computational Materials, 2022, 8(1): 59.
- [73] Biamonte J, Wittek P, Pancotti N, et al. Quantum machine learning[J]. Nature, 2017, 549(7671): 195-202.
- [74] Bufano F, Riggi S, Sciacca E, et al. Machine Learning for Astrophysics: Proceedings of the ML4Astro International Conference 30 May-1 Jun 2022[M]. Springer Nature, 2023.
- [75] Brunton S L, Noack B R, Koumoutsakos P. Machine Learning for Fluid Mechanics[J]. Annual Review of Fluid Mechanics, 2020, 52(1): 477-508.
- [76] Udrescu S M, Tegmark M. AI Feynman: A Physics-Inspired Method for Symbolic Regression[J]. Science Advances, 2020, 6(16): eaay2631.
- [77] Colen J, Han M, Zhang R, et al. Machine Learning Active-Nematic Hydrodynamics[J]. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118(10): e2016708118.
- [78] Bahiuddin I, Mazlan S A, Imaduddin F, et al. Review of Modeling Schemes and Machine Learning Algorithms for Fluid Rheological Behavior Analysis[J]. Journal of the Mechanical Behavior of Materials, 2024, 33(1): 20220309.

- [79] Mangal D, Jha A, Dabiri D, et al. Data-Driven Techniques in Rheology: Developments, Challenges and Perspective[J]. *Current Opinion in Colloid & Interface Science*, 2025, 75: 101873.
- [80] Alsabaa A, Gamal H, Elkataatny S, et al. Machine Learning Model for Monitoring Rheological Properties of Synthetic Oil-Based Mud[J]. *ACS Omega*, 2022, 7(18): 15603-15614.
- [81] Shakeel M, Pourafshary P, Hashmet M R, et al. Application of machine learning techniques to predict viscosity of polymer solutions for enhanced oil recovery[J]. *Energy Systems*, 2023: 1-24.
- [82] Mustafa A, Haider D, Barua A, et al. Machine Learning Based Microfluidic Sensing Device for Viscosity Measurements[J]. *Sensors & Diagnostics*, 2023, 2(6): 1509-1520.
- [83] Ponick A, Langer A, Beyer D, et al. Image-Based Deep Learning for Rheology Determination of Bingham Fluids[J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2022, 43: 711-720.
- [84] Chen Y, Peng B, Kontogeorgis G M, et al. Machine learning for the prediction of viscosity of ionic liquid–water mixtures[J]. *Journal of Molecular Liquids*, 2022, 350: 118546.
- [85] Zhang Y, Jin H, Liu H, et al. Deep Semi-Supervised Just-in-Time Learning Based Soft Sensor for Mooney Viscosity Estimation in Industrial Rubber Mixing Process[J]. *Polymers*, 2022, 14(5): 1018.
- [86] Zhang L, Shao S. Image-based machine learning for materials science[J]. *Journal of Applied Physics*, 2022, 132(10): 100701.
- [87] Zhang J, Liu Y, Chandra Sekhar.P D, et al. Rapid, Autonomous High-Throughput Characterization of Hydrogel Rheological Properties via Automated Sensing and Physics-Guided Machine Learning[J]. *Applied Materials Today*, 2023, 30: 101720.
- [88] Verheyen C A, Uzel S G, Kurum A, et al. Integrated Data-Driven Modeling and Experimental Optimization of Granular Hydrogel Matrices[J]. *Matter*, 2023, 6(3): 1015-1036.
- [89] Martineau R L, Bayles A V, Hung C S, et al. Engineering Gelation Kinetics in Living Silk Hydrogels by Differential Dynamic Microscopy Microrheology and Machine Learning [J]. *Advanced Biology*, 2022, 6(1): 2101070.
- [90] Lee S, Kim S R, Lee H J, et al. Predictive model for the spreadability of cosmetic for-

- mulations based on large amplitude oscillatory shear (LAOS) and machine learning[J]. Physics of Fluids, 2022, 34(10): 103109.
- [91] Lu L, Gao X, Dietiker J F, et al. Machine learning accelerated discrete element modeling of granular flows[J]. Chemical Engineering Science, 2021, 245: 116832.
- [92] Seryo N, Sato T, Molina J J, et al. Learning the Constitutive Relation of Polymeric Flows with Memory[J]. Physical Review Research, 2020, 2(3): 033107.
- [93] Bai J, Zhou Y, Rathnayaka C M, et al. A Data-Driven Smoothed Particle Hydrodynamics Method for Fluids[J]. Engineering Analysis with Boundary Elements, 2021, 132: 12-32.
- [94] Li W, Burkhart C, Polinska P, et al. Backmapping Coarse-Grained Macromolecules: An Efficient and Versatile Machine Learning Approach[J]. The Journal Of Chemical Physics, 2020, 153(4): 041101.
- [95] Zhao L, Li Z, Caswell B, et al. Active Learning of Constitutive Relation from Mesoscopic Dynamics for Macroscopic Modeling of Non-Newtonian Flows[J]. Journal of Computational Physics, 2018, 363: 116-127.
- [96] Saadat M, Mangal D, Jamali S. A Rheologist's Guideline to Data-Driven Recovery of Complex Fluids' Parameters from Constitutive Models[J]. Digital Discovery, 2023, 2(4): 915-928.
- [97] Reyes B, Howard A, Perdikaris P, et al. Learning Unknown Physics of Non-Newtonian Fluids[J]. Physical Review Fluids, 2021, 6(7): 073301.
- [98] Karniadakis G E, Kevrekidis I G, Lu L, et al. Physics-Informed Machine Learning[J]. Nature Reviews Physics, 2021, 3(6): 422-440.
- [99] Raissi M, Perdikaris P, Karniadakis G. Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations[J]. Journal of Computational Physics, 2019, 378: 686-707.
- [100] Mahmoudabadbozchelou M, Jamali S. Rheology-Informed Neural Networks (RhINNs) for Forward and Inverse Metamodelling of Complex Fluids[J]. Scientific Reports, 2021, 11(1): 12015.
- [101] Dabiri D, Saadat M, Mangal D, et al. Fractional Rheology-Informed Neural Networks for Data-Driven Identification of Viscoelastic Constitutive Models[J]. Rheologica Acta, 2023, 62(10): 557-568.

- [102] Zhang T, Wang D, Lu Y. RheologyNet: A Physics-Informed Neural Network Solution to Evaluate the Thixotropic Properties of Cementitious Materials[J]. *Cement and Concrete Research*, 2023, 168: 107157.
- [103] Nagrani P P, Kulkarni R V, Kelkar P U, et al. Data-Driven Rheological Characterization of Stress Buildup and Relaxation in Thermal Greases[J]. *Journal of Rheology*, 2023, 67(6): 1129-1140.
- [104] Tucny J M, Durve M, Montessori A, et al. Learning of Viscosity Functions in Rarefied Gas Flows with Physics-Informed Neural Networks[J]. *Computers and Fluids*, 2024, 269: 106114.
- [105] Howard A, Dong J, Patel R, et al. Machine Learning Methods for Particle Stress Development in Suspension Poiseuille Flows[J]. *Rheologica Acta*, 2023, 62(10): 507-534.
- [106] Mahmoudabadbozchelou M, Caggioni M, Shahsavari S, et al. Data-Driven Physics-Informed Constitutive Metamodeling of Complex Fluids: A Multifidelity Neural Network (MFNN) Framework[J]. *Journal of Rheology*, 2021, 65(2): 179-198.
- [107] Mahmoudabadbozchelou M, Kamani K M, Rogers S A, et al. Digital Rheometer Twins: Learning the Hidden Rheology of Complex Fluids through Rheology-Informed Graph Neural Networks[J]. *Proceedings of the National Academy of Sciences*, 2022, 119(20): e2202234119.
- [108] Sato T, Miyamoto S, Kato S. Rheo-SINDy: Finding a Constitutive Model from Rheological Data for Complex Fluids Using Sparse Identification for Nonlinear Dynamics [J]. *Journal of Rheology*, 2025, 69(1): 15-34.
- [109] Lennon K R, McKinley G H, Swan J W. Scientific Machine Learning for Modeling and Simulating Complex Fluids[J]. *Proceedings of the National Academy of Sciences*, 2023, 120(27): e2304669120.
- [110] Mahmoudabadbozchelou M, Kamani K M, Rogers S A, et al. Unbiased Construction of Constitutive Relations for Soft Materials from Experiments via Rheology-Informed Neural Networks[J]. *Proceedings of the National Academy of Sciences*, 2024, 121(2): e2313658121.
- [111] Lu L, Jin P, Pang G, et al. Learning Nonlinear Operators via DeepONet Based on the Universal Approximation Theorem of Operators[J]. *Nature Machine Intelligence*, 2021,

- 3(3): 218-229.
- [112] Rashid M M, Pittie T, Chakraborty S, et al. Learning the stress-strain fields in digital composites using Fourier neural operator[J]. *Iscience*, 2022, 25(11): 105452.
- [113] You H, Zhang Q, Ross C J, et al. Learning deep implicit Fourier neural operators (IFNOs) with applications to heterogeneous material modeling[J]. *Computer Methods in Applied Mechanics and Engineering*, 2022, 398: 115296.
- [114] Elman J L. Finding Structure in Time[J]. *Cognitive Science*, 1990, 14(2): 179-211.
- [115] Schmidhuber J, Hochreiter S, et al. Long short-term memory[J]. *Neural Comput*, 1997, 9(8): 1735-1780.
- [116] Cho K, Van Merriënboer B, Gulcehre C, et al. Learning Phrase Representations Using RNN Encoder – Decoder for Statistical Machine Translation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar: Association for Computational Linguistics, 2014: 1724-1734.
- [117] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [118] Li Z, Liu F, Yang W, et al. A survey of convolutional neural networks: analysis, applications, and prospects[J]. *IEEE transactions on neural networks and learning systems*, 2021, 33(12): 6999-7019.
- [119] Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need[C]//Advances in Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000-6010.
- [120] Gu A, Dao T. Mamba: Linear-Time Sequence Modeling with Selective State Spaces[C] //First Conference on Language Modeling. Philadelphia, Pennsylvania, United States: OpenReview, 2024: 1-32.
- [121] Raissi M, Perdikaris P, Karniadakis G. Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations[J]. *Journal of Computational Physics*, 2019, 378: 686-707.
- [122] Cuomo S, Di Cola V S, Giampaolo F, et al. Scientific Machine Learning Through Physics-Informed Neural Networks: Where We Are and What's Next[J]. *Journal of Scientific Computing*, 2022, 92(3): 88.

- [123] Lu L, Meng X, Mao Z, et al. DeepXDE: A Deep Learning Library for Solving Differential Equations[J]. SIAM Review, 2021, 63(1): 208-228.
- [124] Farmer J, Oian C A, Bowman B A, et al. Empirical Loss Weight Optimization for PINN Modeling Laser Bio-Effects on Human Skin for the 1D Heat Equation[J]. Machine Learning with Applications, 2024, 16: 100563.
- [125] Xiang Z, Peng W, Liu X, et al. Self-Adaptive Loss Balanced Physics-informed Neural Networks[J]. Neurocomputing, 2022, 496: 11-34.
- [126] Song Y, Wang H, Yang H, et al. Loss-Attentional Physics-Informed Neural Networks [J]. Journal of Computational Physics, 2024, 501: 112781.
- [127] Shah J, Lineswala R, Chopra A. Benchmarking Quantum-Assisted PINN (QA-PINN) for Computational Fluid Dynamics[C]//2024 IEEE International Conference on Quantum Computing and Engineering (QCE). Montreal, QC, Canada: IEEE, 2024: 1707-1712.
- [128] Hillebrecht B, Unger B. Certified machine learning: A posteriori error estimation for physics-informed neural networks[C]//2022 International Joint Conference on Neural Networks (IJCNN). Padua, Italy: IEEE, 2022: 1-8.
- [129] Haitsiukovich K, Ilin A. Improved Training of Physics-Informed Neural Networks with Model Ensembles[C]//2023 International Joint Conference on Neural Networks (IJCNN). Gold Coast, Australia: IEEE, 2023: 1-8.
- [130] Dai Y, Gieseke F, Oehmcke S, et al. Attentional feature fusion[C]//Proceedings of the IEEE/CVF winter conference on applications of computer vision. Waikoloa, HI, USA: IEEE, 2021: 3560-3569.
- [131] Cui M, Li K, Chen J, et al. CM-Unet: A Novel Remote Sensing Image Segmentation Method Based on Improved U-Net[J]. IEEE Access, 2023, 11: 56994-57005.
- [132] Rong Y, Wei X, Lin T, et al. DynStatF: an efficient feature fusion strategy for LiDAR 3D object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, BC, Canada: IEEE, 2023: 3237-3246.
- [133] Nagrani A, Yang S, Arnab A, et al. Attention bottlenecks for multimodal fusion[J]. Advances in neural information processing systems, 2021, 34: 14200-14213.
- [134] Kingma D P, Welling M. Auto-Encoding Variational Bayes[C]//International Confer-

- ence on Learning Representations. Banff, Canada: OpenReview, 2014: 1-14.
- [135] Sohn K, Yan X, Lee H. Learning Structured Output Representation Using Deep Conditional Generative Models[C]//Advances in neural information processing systems. Cambridge, MA, USA: MIT Press, 2015: 3483-3491.
- [136] Bengio Y, Senécal J S. Adaptive importance sampling to accelerate training of a neural probabilistic language model[J]. Journal of Machine Learning Research, 2003, 3: 1137-1155.
- [137] Bond-Taylor S, Leach A, Long Y, et al. Deep Generative Modelling: A Comparative Review of VAEs, GANs, Normalizing Flows, Energy-Based and Autoregressive Models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(11): 7327-7347.
- [138] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets[C]// Advances in Neural Information Processing Systems. Montreal, Canada: Curran Associates, Inc., 2014: 2672-2680.
- [139] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [140] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. Advances in neural information processing systems, 2020, 33: 6840-6851.
- [141] Song Y, Sohl-Dickstein J, Kingma D P, et al. Score-Based Generative Modeling through Stochastic Differential Equations[C]//International Conference on Learning Representations. Vienna, Austria: OpenReview, 2021: 1-36.
- [142] Cao H, Tan C, Gao Z, et al. A Survey on Generative Diffusion Models[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(7): 2814-2830.
- [143] Huang J, Xu Y, Qi S, et al. Ultrahigh Energy-Dissipation Elastomers by Precisely Tailoring the Relaxation of Confined Polymer Fluids[J]. Nature Communications, 2021, 12(1): 3610.

攻读博士/硕士学位期间取得的研究成果

一、已发表（包括已接受待发表）的论文，以及已投稿、或已成文打算投稿、或拟成文投稿的论文情况（只填写与学位论文内容相关的一部分）：

序号	作者（全体作者，按顺序排列）	题 目	发表或投稿刊物名称、级别	发表的卷期、年月、页码	与学位论文哪一部分（章、节）相关	被索引收录情况
1	redfu、Mahmoudabad-bozchelou、Lennon	Dissipative particle dynamics simulations of self-assembly in solution of bottlebrush copolymers with different topologies	ACM Transactions on Modeling and Computer Simulation (TOMACS)	2024.04 100(10)	第三章	
2						

注：在“发表的卷期、年月、页码”栏：

1. 如果论文已发表，请填写发表的卷期、年月、页码；
2. 如果论文已被接受，填写将要发表的卷期、年月；
3. 以上都不是，请据实填写“已投稿”，“拟投稿”。

不够请另加页。

二、与学位内容相关的其它成果（包括专利、著作、获奖项目等）

致 谢