arXiv:2602.07184v1 [cs.CE] 6 Feb 2026

不确定性条件下基于物理信息机器学习的批式结晶建模

Dingqi Nai[a]、Huayu Li[b]、Martha Grover[a]、Andrew J. Medford[a]*

[a]佐治亚理工学院化学与生物分子工程学院 亚特兰大，GA 30332[b]材料与分析科学部 勃林格殷格翰制药公司 里奇菲尔德，CT 06877

2026年2月10日

摘要

由于各种不确定性来源导致真实数据中存在非理想性，开发稳健且可靠的结晶过程建模方法通常颇具挑战。本研究考察了物理信息循环神经网络（PIRNN）的有效性。该网络将机理群体平衡模型与循环神经网络相结合，并应用于存在系统不确定性和模型不确定性的情形。研究通过包含受控噪声、溶解度偏移和有限采样的合成数据来表征这些不确定性。结果表明，尽管训练数据中存在显著的随机变化，PIRNN仍能实现较强的泛化能力和物理一致性，保持稳定的学习行为，并准确恢复动力学参数。在溶解度模型存在系统误差的情况下，与纯数据驱动模型相比，加入物理正则化使测试性能提升了一个数量级以上；然而，过度提高物理项的权重会因模型失配而增加误差。结果还表明，即使在采样分辨率极低的情况下，PIRNN也能够恢复模型参数并复现结晶动力学。这些发现验证了物理信息机器学习在处理数据缺陷和不完整领域知识方面的稳健性，为实现可靠且实用的结晶动力学混合建模以及工业过程监测与控制提供了潜在路径。

关键词 物理信息机器学习；批次结晶；循环神经网络；不确定性；参数估计

1 引言

结晶是多个行业中分离、纯化和生产固体材料的基础工艺，这些行业包括制药、精细化学品和食品加工 [1–3]。尤其是在制药行业，超过 80% 的固体产品至少涉及一个结晶步骤，而生产的活性药物成分中超过 90% 为晶体 [4–7]。结晶对纯度、粒度分布和多晶型等产品特性的影响，使得理解并精确控制这一过程成为持续生产符合质量规范产品的必要条件 [8]。因此，大量研究致力于深入、定量地认识结晶过程动力学 [9–12]。

近几十年来，过程分析技术（PAT）的发展彻底改变了人们对结晶过程开发的理解，使动态系统能够得到实时监测 [13, 14]。过程分析技术（PAT）工具，如聚焦光束反射测量（FBRM）、衰减全反射傅里叶变换红外（ATR-FTIR）

*通讯作者。电子邮箱：ajm@gatech.edu

光谱学、拉曼光谱和原位成像，能够深入揭示结晶动力学，包括过饱和度和晶体粒度分布（CSD） [9, 15–18]。然而，从过程分析技术（PAT）中获取的信息通常具有不确定性且并不完善；如果没有能够解读测量结果并将其与潜在系统状态联系起来的适当过程模型，要有效利用这些数据进行过程控制仍然具有挑战性 [19–21]。

基于物理的结晶建模传统上一直是结晶研究与开发的基础。基于群体平衡模型（PBM），并且通常与质量衡算和能量衡算相耦合，基于物理的模型能够对成核、晶体生长、溶解、聚集和破碎现象进行严格的数学描述 [9, 22–24]。其理论基础和外推能力赋予了基于物理的模型独特优势。然而，由于方程具有高度非线性的结构，且需要大量实验数据，这些模型的参数估计可能具有挑战性 [25, 26]。尽管存在这些挑战，基于物理的模型对于过程放大和设计，以及深入理解结晶现象的基本机理，仍然至关重要。

机器学习算法的快速发展以及大量实验数据集的可用性，也推动了纯数据驱动建模方法的出现 [10, 27]。人工神经网络（ANN）、支持向量机、随机森林和其他深度学习框架等技术，已成功用于捕捉结晶动力学与输入过程变量之间的“黑箱”经验关系[6, 28–37]。近期研究表明，循环神经网络和时间序列Transformer在预测结晶动力学以及实施基于神经网络的控制策略方面具有较高效率 [6, 38–42]。与基于物理的模型相比，数据驱动模型能够处理大量数据，而无需明确的机理知识。然而，在处理新颖系统或超出训练域的操作条件时，它们通常面临物理可解释性差和外推能力有限等固有局限 [43]。

混合建模将基于物理学的知识与数据驱动技术相结合，已成为一种很有前景的框架，能够融合这两种范式的优势 [44, 45], ；物理信息神经网络（PINNs）等架构正日益应用于各种化学工程系统，涵盖反应动力学、流体动力学和过程控制等领域 [46–60]。这种“灰箱”方法保留了物理模型的可解释性和外推能力，同时利用灵活的机器学习模型处理大量数据以及简化模型与真实系统之间的差异，从而在非理想条件或稀疏数据情形下提升预测性能 [30, 61–65]。近期的结晶研究验证了混合建模的优势，表明与纯数据驱动方法相比，它有望实现更高的准确性和稳健性 [10, 65]。

然而，将这些混合建模方法应用于实际结晶过程时，仍面临诸多挑战。测量噪声、过程扰动以及其他固有的系统变异性会带来两种基本的不确定性：偶然不确定性和认知不确定性。偶然不确定性指观测中不可消除的变化，其来源包括成核事件的随机性、测量设备限制以及随机的操作波动。认知不确定性的范畴更广，源于对系统和模型认知的不完备，例如简化的模型假设、模型与数据之间的不一致，或实验数据覆盖有限 [66–71]。以往研究采用蒙特卡罗方法将输入不确定性传播至群体平衡模型，揭示了不确定性对预测的晶体粒度分布和过程结果的显著影响 [21]。近期研究还尝试通过贝叶斯方法量化这些不确定性，使得在结晶过程建模中理解并考虑这些不确定性受到更多关注 [72]。

本文旨在通过研究数据非理想性对结晶建模的影响，拓展现有研究。这项研究采用混合方法，基于Wu等人近期开发的物理信息循环神经网络（PIRNN） [65]。我们聚焦于一个相对简单但具有代表性的案例：对乙酰氨基酚间歇冷却结晶。先前研究已建立了一个物理PBM模型，我们将其视为“真实值” [73] ，并探究合成“观测数据”中的随机不确定性和几种形式的认知不确定性如何影响结晶建模的性能与可靠性。通过系统地引入不同程度的测量噪声、模型差异和采样限制，我们揭示了在数据不完善且领域知识不完整的情况下，所提方法的稳健性与可靠性。本研究结果表明，在存在随机测量噪声和认知采样不确定性的情况下，混合PIRNN模型为PBM模型参数化提供了一种稳健的选择。此外，我们发现，经过适当优化后，这些混合模型能够灵活地处理溶解度校准曲线和采样频率中的认知不确定性，且性能优于纯数据驱动模型和纯物理模型。这些发现表明，PIRNN方法有望成为一种构建稳健结晶模型的策略，将实际过程分析技术数据与已知的物理化学主导机制相结合。

2 方法学

2.1 群体平衡模型与矩量法

在本研究中，我们聚焦于晶种间歇冷却场景。遵循Ramkrishna提出的框架[74] 并采用Li等人 [73], 的参数化方法，我们使用一维PBM描述该系统 [75, 76]。PBM描述晶体数密度 n 随时间 t 和特征晶体尺寸 L 的演变，其形式如下:

$$\frac{\partial n}{\partial t} + G\frac{\partial n}{\partial L} = 0 \tag{1.1}$$

其中， G 表示晶体生长速率。矩量法 [77, 78] 用于求解该偏微分方程，方法是将其拆分为一组常微分方程（ODE），如Kim等人 [9]所实施。在矩量法中，第 i阶矩 μ_i 定义为

$$\mu_i = \int_0^\infty L^i n(L) dL, \qquad i = 0, 1, 2, 3, ... \tag{1.2}$$

从零阶矩到三阶矩， μ_0、 μ_1、 μ_2、 μ_3分别与晶体的总数、长度、表面积和体积成正比。假设生长速率 G [73, 79]与尺寸无关，则可根据式（1.1）和（1.2）推导出各阶矩的时间导数:

$$\frac{\mathrm{d}\mu_i}{\mathrm{d}t} = iG\mu_{i-1}, \qquad i = 1, 2, 3, ... \tag{1.3}$$

对于表示晶体总数的零阶矩，其时间导数可定义为成核速率，即

$$\frac{\mathrm{d}\mu_0}{\mathrm{d}t} = B_1 + B_2 \tag{1.4}$$

其中， B_1 和 B_2 分别表示初级和二次成核速率。初级成核是指晶体直接从澄清溶液中形成，而二次成核则是指在已有晶体的促进下形成新晶体 [76, 80]。如先前研究所示，在晶种批次结晶中，添加晶种会促进二次成核速率，因此在此情况下忽略初级成核。由此，零阶矩的变化速率可仅通过二次成核速率的经验模型计算 [22]

$$\begin{aligned} \frac{\mathrm{d}\mu_0}{\mathrm{d}t} &= B_2 \\ &= k_{b2}(S-1)^\alpha m_s^\beta \end{aligned} \tag{1.5}$$

其中， k_{b2} 为二次成核速率常数 [（g 晶体/kg 溶剂）$^{-\beta}$/（min·g 溶剂）]， m_s 为单位质量溶液中的晶体质量 [g 晶体/kg 溶剂]， α 和 β 为模型指数参数， S 为过饱和度。过饱和度 S 可根据浓度 C [g 溶质/g 溶剂] 和随温度变化的饱和浓度 C_S [g 溶质/g 溶剂] 计算，公式如下:

$$S = C/C_S \tag{1.6}$$

$$C_S = -16.17 + 1.765 \times 10^{-1} T - 6.439 \times 10^{-4} T^2 + 7.915 \times 10^{-7} T^3 \tag{1.7}$$

其中， T 表示以开尔文为单位的过程温度，饱和浓度多项式参数取自Li等人先前的研究 [81]。

鉴于第三阶矩 μ_3 对应晶体的体积，批次过程中溶质浓度 C 可根据质量衡算计算为

$$C(t) = C_0 - k_\nu \rho \mu_3(t) \tag{1.8}$$

其时间导数为

$$\frac{\mathrm{d}C}{\mathrm{d}t} = -k_\nu \rho \frac{\mathrm{d}\mu_3}{\mathrm{d}t} = -3k_\nu \rho G \mu_2 \tag{1.9}$$

其中， C_0 为初始浓度 [g 溶质/g 溶剂]， k_ν为晶体的体积形状因子， ρ 为晶体的固体密度 [g/cm^3]。有意将晶种质量设得较小，因此可在质量衡算中忽略。如上所述，在假设生长与尺寸无关的情况下， G 可表示为以绝对过饱和度为驱动力 $C - C_S$ 的阿伦尼乌斯型方程:

$$G = k_g \exp\left(-\frac{E_{a_g}}{RT}\right)(C - C_S)^{\gamma_g} \tag{1.10}$$

其中， k_g 是晶体生长的指前因子 [（μ米/分钟）（克溶质/克溶剂）$^{-\gamma_g}$]， E_{a_g} 是活化能 [焦耳/摩尔]， γ_g 是驱动力的指数参数， R 是通用气体常数 [焦耳/（摩尔·开尔文）]。

2.2 数据生成与准备

合成数据集使用第2.1节中描述的群体平衡模型及相关矩方程生成。仿真设置旨在模拟动态温度曲线下的间歇冷却结晶实验。按照Li等人 [73],描述的加晶种间歇分段冷却曲线，每个实验由三个阶段组成：初始等温平台、线性冷却阶段和最终等温保持阶段。温度变化使用以下函数构建:

$$T(t) = \begin{cases} T_{\text{plat}}, & 0 \le t < t_{\text{plat}}, \\ T_{\text{plat}} + r(t - t_{\text{plat}}), & t_{\text{plat}} \le t < t_{\text{cool}}, \\ T_{\text{final}}, & t_{\text{cool}} \le t < t_{\text{end}}, \end{cases} \tag{2.1}$$

其中， T_{plat} 为初始平台温度 [℃]， r 为冷却速率 [℃/分钟]， t_{plat} 为平台持续时间 [分钟]，t_{cool} 为冷却时间 [分钟]， T_{final} 为最终保持温度（ 0℃）， t_{end} 为实验结束时间，此处设为500分钟。参数 T_{plat}、 r和 t_{plat} 在预定义范围内从均匀分布中随机采样，以生成多种实验条件； t_{cool} 则根据 T_{final} 和 T_{plat}之间的差值计算。初始溶质浓度 C_0 也从均匀分布中随机取值，以引入过饱和度动态变化的差异。合成温度曲线以及 T_{plat}、 r、 t_{plat} 和 C_0 的分布见补充信息中的图S1至S4。

对于每个模拟实验，基于生成的温度曲线和Li等人报告的参数值 [73]，使用SciPy中实现的LSODA方法 [82] 求解第2.1节中的常微分方程组。正向求解包含五个状态变量，从 μ_0 到 μ_3 和 C；求解常微分方程组时需指定这五个状态变量的初始条件。为确保模拟系统处于加晶种条件，特意将 μ_0 的初始条件设为最小值，且该值至少比 μ_0 的数值小3个数量级，因此在质量衡算中可忽略。模拟详情见补充信息中的S1.2节。每次模拟都会生成浓度以及零阶至三阶矩的时间序列输出。需要注意的是，实验中通常无法直接测量矩。以往研究已将FBRM信号与这些矩关联起来，Kim提出了一种浅层神经网络框架，可将FBRM测量值准确转换为矩数据 [6, 83, 84]。此外，也可以使用原位相机估算矩。本研究假设使用“理想传感器”，可直接获取矩和浓度；在基准情形下，我们还假设溶解度已得到完美校准。

尽管无噪声模拟为比较提供了理想基线，但真实实验数据包含测量噪声。为以受控方式复现这种非理想性，在第3.1节中向矩和浓度轨迹均引入高斯噪声。噪声幅值设定为每条状态轨迹上所有值的标准差的10%、30%和100%（即从 $t = 0$ 到 $t = 500$分钟内所有浓度/矩的标准差），以分别代表“低噪声”、“中噪声”和“高噪声”情形。生成的数据剖面见补充信息中的图 S6至图 S9。

除引入受控噪声水平以反映偶然不确定性外，还对群体平衡模型模拟中使用的溶解度表达式进行调整，以复现第3.2节中与实际测量相关的非理想性。溶解度值可能因物料纯度和分析方法而有所不同。结晶领域中常报告偏移，其中预测与测得的平衡溶解度之差可达百分之几，具体取决于数据集质量及所用拟合方法 [9, 79, 85–87]。引入这种系统偏差反映了“真实”热力学溶解度与通常用于关联过程模拟和实验的工作关联式之间的不匹配。为此，在恒定溶解度偏移情形中，利用式(1.7)作为温度的多项式函数计算得到的饱和浓度 C_S在整个温度范围内被均匀偏移10%。该偏置是认知模型不确定性的一个来源，源于对系统和模型的不完全认识。

我们还在第 3.3 节中将采样频率视为认知不确定性的另一来源。为模拟实验约束，在这种情况下，训练集中仅包含 10 次运行；我们有策略地评估每次运行分别采用 2、3、5 和 9 个点的情形，对应的训练点总数分别为 20、30、50 和 90。我们采用非均匀且近似对数的采样方案。在 2 点情形下，仅能获得状态变量的初始值和最终值；在 3 点情形下，由于早期阶段的动力学更快、梯度更陡，因此在 t_{plat} 之前增加一个数据点；在 5 点情形下，分别在三个冷却阶段增加三个点，如式 (2.1) 所示；在 9 点情形下，则采用近似对数间隔，使早期采样更密、后期采样更疏。详细的采样点见补充信息中的第 S4 节。需要注意的是，尽管状态变量的采样频率从每分钟一次大幅降低，我们仍假设可以获得完整的、以分钟为分辨率的温度曲线，因为在该结晶系统中，温度是被控变量。将这些误差来源纳入合成数据集，进一步提高了其现实性，并

可用于评估在关键热力学知识并非完全已知的条件下，不同建模策略的稳健性。

为对模型进行无偏评估，将生成的数据按 6:2:2 的比例划分为训练集、验证集和测试集。在本研究中，共生成 100 次合成运行，因此训练集、验证集和测试集分别包含 60、20 和 20 次运行。训练数据用于优化机器学习模型，包括网络权重、偏置、PBM 参数及其他可学习参数。验证数据不用于优化或调整模型配置超参数（例如网络宽度/深度、学习率等），而仅用于通过早停选择“性能最佳”的模型。早停是一种正则化技术，可提升模型对未见数据的泛化能力 [88, 89]。这一流程能够以相对公平的方式选出单个最佳模型来评估模型性能，并防止模型可能对训练数据过拟合。随后，使用测试数据对模型性能进行最终的无偏评估。在本文中，除非另有说明，验证误差和测试误差均指基于验证集通过早停选出的最佳模型所对应的误差。训练前，使用训练集中各状态变量对应的最大值对其进行归一化，以确保数值稳定性，并平衡各变量对损失函数的贡献。

2.3 物理信息循环神经网络（PIRNNs）

遵循 Raissi等人提出的物理信息神经网络（PINNs）原理 [90], , 并直接基于 Wu等人近期为结晶过程改进的 PIRNN 架构 [65], , 我们采用一种混合框架，将长短期记忆网络的序列建模能力与基于物理的约束相结合，以模拟结晶动力学 [91], , 如图 1 所示。PIRNN 模型由一个 LSTM编码器组成，用于处理拼接后的状态变量 $\mathbf{x}(t)$ 和控制输入 $T(t)$，其中

$$\mathbf{x}(t) = [\mu_0, \mu_1, \mu_2, \mu_3, C]^\intercal \tag{3.1}$$

随后接入批归一化和 dropout层，以防止协变量偏移和过拟合并提升泛化能力；再接入一个全连接线性解码器，在每个前向步输出预测状态变量 [92–95]。为保持架构简洁，此处使用了两个 LSTM 层。输出采用 Softplus激活函数，以确保预测值非负。详细的设计和超参数（即节点数/层数、丢弃率等）见 SI 中的表 S1。训练期间，PIRNNs 接收初始状态 $\mathbf{x}_0 \in \mathbb{R}^5$ 和温度控制序列 $T \in \mathbb{R}^{N\times 1}$，针对 N 个时间戳生成五个状态变量的预测轨迹 $\hat{\mathbf{x}} \in \mathbb{R}^{N\times 5}$。

总训练损失 $\mathcal{L}$ 是数据损失项和物理损失项的加权和：

$$\mathcal{L} = \mathcal{L}_{data} + \lambda \mathcal{L}_{physics} \tag{3.2}$$

其中，λ 是一个权重超参数，用于调节数据拟合与物理匹配的相对重要性。数据损失$\mathcal{L}_{data}$ 由三个主要部分组成：预测状态与观测训练状态之间的均方误差（MSE）、Huber损失和平滑度损失，其中

$$\mathcal{L}_{data} = \omega \text{MSE}(\hat{\mathbf{x}}, \mathbf{x}_{\text{obs}}) + (1-\omega)\text{Huber}_\delta(\hat{\mathbf{x}}, \mathbf{x}_{\text{obs}}) + \mathcal{L}_{smooth} \tag{3.3}$$

$$\omega = \sigma(-\eta_{noise}) \tag{3.4}$$

$$\text{MSE}(\hat{\mathbf{x}}, \mathbf{x}_{\text{obs}}) = \frac{1}{N}\sum_{k=1}^{N} \|\hat{\mathbf{x}}_k - \mathbf{x}_{k,\text{obs}}\|_2^2 \tag{3.5}$$

$$\mathcal{L}_{smooth} = \frac{1}{N-2}\sum_{k=2}^{N-1} \|\hat{\mathbf{x}}_{k+1} - 2\hat{\mathbf{x}}_k + \hat{\mathbf{x}}_{k-1}\|^2 \tag{3.6}$$

这里，ω 是关于可学习噪声尺度参数 η_{noise} 的Sigmoid函数，可实现MSE与Huber损失的自适应混合。这种方法受异方差似然损失的启发，该损失允许模型对被认为噪声较大的数据点赋予较低权重 [96]。Huber损失是一种稳健回归损失，它通过转换阈值 δ，结合了MSE对小残差的敏感性与平均绝对误差（MAE）对大残差的稳健性，具体而言，

$$\text{Huber}_\delta(\hat{\mathbf{x}}, \mathbf{x}_{\text{obs}}) = \begin{cases} 0.5(\hat{\mathbf{x}} - \mathbf{x}_{\text{obs}})^2, & \text{if} |\hat{\mathbf{x}} - \mathbf{x}_{\text{obs}}| < \delta \\ \delta(|\hat{\mathbf{x}} - \mathbf{x}_{\text{obs}}| - 0.5\delta) & \text{otherwise} \end{cases} \tag{3.7}$$

在这里，如果误差小于阈值，损失函数呈二次形式，与均方误差相同，从而确保梯度平滑并施加强惩罚；另一方面，对于较大的误差，损失函数随残差线性增长，类似于平均绝对误差，这有助于减小离群值和噪声尖峰的影响，防止它们主导

训练过程。这种组合还确保梯度处处平滑，而平均绝对误差在残差为零时不可微，因此 Huber损失是此处基于梯度的优化的理想选择 [97]。公式 3.6 中定义的平滑度损失 $\mathcal{L}_{smooth}$ 也用于获得更平滑的时间行为，并进一步防止噪声过拟合。它借鉴了信号处理和最优控制理论中的正则化技术，通过二阶有限差分惩罚预测状态中的快速振荡。这些损失项的相对权重被任意设为相等。

遵循典型的物理信息神经网络设计，物理损失 $\mathcal{L}_{physics}$ 被纳入式 3.2，以确保学习得到的轨迹与底层物理学 [90]保持一致。它惩罚基于预测状态计算的时间导数有限差分近似与

$$\dot{\mathbf{x}}_{\mathrm{FD}}(t_k) \approx \frac{\hat{\mathbf{x}}_{k+1} - \hat{\mathbf{x}}_{k-1}}{2\Delta t}, \quad k = 2, ..., N-1 \tag{3.8}$$

以及采用均方误差准则、在PBM矩常微分方程中定义的基本变化率，如下

$$\mathcal{L}_{physics} = \mathrm{MSE}\left(\dot{\mathbf{x}}_{\mathrm{FD}}, \mathrm{PBM}(\hat{\mathbf{x}}, T_k; \boldsymbol{\theta})\right) \tag{3.9}$$

其中 $\boldsymbol{\theta}$ 包含物理信息循环神经网络的所有可学习参数，包括权重和偏置，以及群体平衡模型的动力学参数 [98]。在本文中，所有已开展案例研究中的全部PBM参数均初始化为1，以避免已知知识对我们的模型产生偏置。随后可将物理信息循环神经网络学习得到的PBM参数与表1中给出的参考值进行比较。

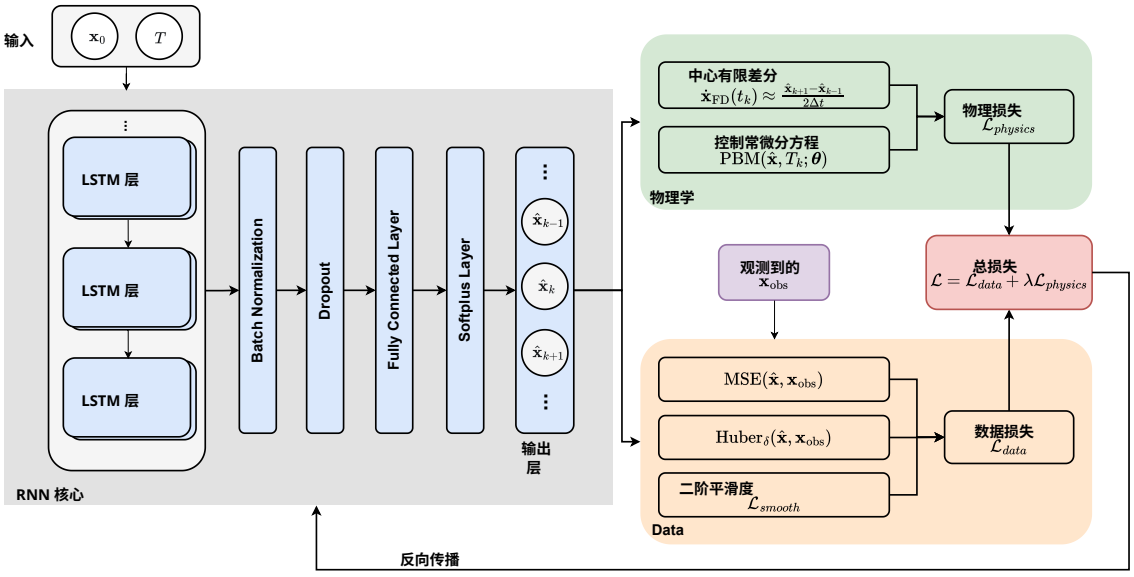


图1：物理信息循环神经网络架构与损失项。

在本文中，模型性能主要通过两项指标进行评估：PIRNN预测和重新评估的常微分方程的预测。PIRNN预测是指利用网络的权重和偏置进行的直接前向预测，而重新评估的常微分方程是指使用由PIRNN优化估计得到的PBM参数求解第2.1节中的常微分方程组。

3 结果与讨论

在不同测量噪声水平和训练数据可用性条件下评估了PIRNN的性能。以受控方式引入了偶然信号噪声，以及以溶解度校准偏差和降低采样频率形式出现的认知不确定性，并分别进行了独立研究。

3.1 信号噪声

图2展示了不同噪声水平下 PIRNN 的测试损失和物理一致性。正如预期，增加训练数据规模通常会降低测试均方误差，表明随着纳入更多

数据，模型的泛化能力得到提升。在无噪声条件下，两项评估指标的误差均低于 10^{-4}。然而，PIRNN 的直接前向预测对训练数据规模的依赖性更强，而使用 PIRNN 估计参数重新评估的常微分方程曲线，在训练运行次数超过 10 次时表现稳定。这表明，在理想条件下，PIRNN 模型仅凭少量数据就能够提取可靠的动力学参数。当数据量较大（60 次训练运行）时，灵活的 PIRNN 结构比僵化的 ODE 积分更精确地插值状态动力学；后者仍对估计参数的微小偏差敏感，这表明使用有限差分可能存在局限。改用更精确的导数计算方法或许可以解决这一问题，但对微分方法的详细评估超出了本文的讨论范围。这一结果表明，PIRNN 模型可能具有双重用途：既可用于直接预测状态，也可用于提取参数。

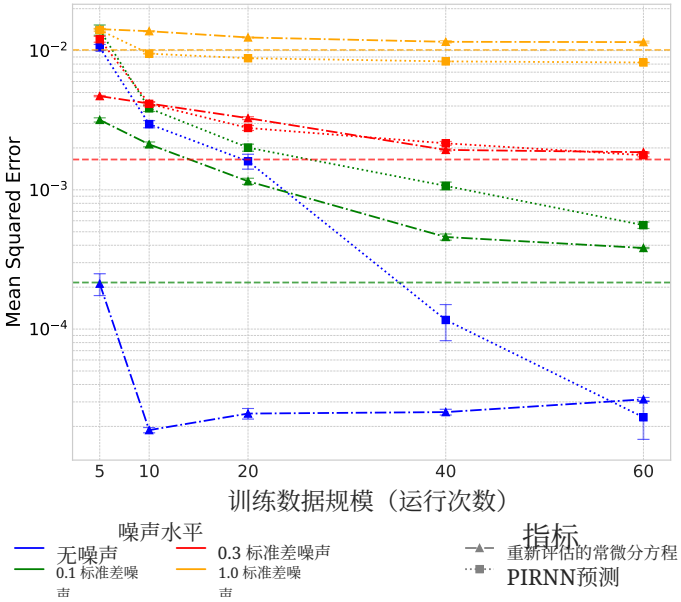


图2：在两种不同指标下，测试数据预测均方误差随训练数据量的变化情况，并比较了不同噪声水平：PIRNN前向预测和使用从PIRNN中学习到的动力学参数计算的重新评估常微分方程。图中展示了随着训练运行次数增加，无噪声、0.1标准差、0.3标准差和1.0标准差噪声条件下的均方误差。水平虚线表示含噪数据与无噪声真实数据之间的参考误差。

引入噪声后，PIRNN预测和重新评估的常微分方程误差都会在较高数值附近趋于平稳，这一数值接近含噪数据与无噪声真实数据之间的参考误差，该误差在图2中以水平虚线表示。这些误差在水平线附近趋于平稳，说明误差升高并非源于模型本身的性能下降，而是由数据中固有的、不可消除的偶然不确定性所致，这是添加噪声造成的。在低噪声水平（0.1标准差）下，重新评估的常微分方程误差通常低于直接的PIRNN预测误差，表明低水平误差不会显著影响提取出的物理规律的质量。然而，在高噪声水平（1.0标准差）下，重新评估的常微分方程误差高于PIRNN前向预测误差，表明恢复出的物理规律有所退化。值得注意的是，在高噪声水平（1.0标准差）下，PIRNN预测误差低于参考水平，这表明模型可能对含噪数据过拟合。调整不同损失项的相对权重或许可以解决这一问题，但为了进行受控比较，在不同噪声水平的案例研究中，不同损失项的权重保持不变。总体而言，与重新评估的常微分方程保持稳定的理想情况不同，这里的误差显然随输入噪声和训练数据规模而变化，这表明恢复出的物理规律的准确性最终受限于观测数据的质量。

图3展示了基于无噪声数据而非含噪数据计算的测试误差。使用无噪声数据计算的测试误差降至约 10^{-3}，且在不同噪声水平下表现出相似的趋势。这进一步证实，图2中误差升高是由于添加噪声导致内在信息损失，而非模型性能下降；噪声的存在对模型的泛化能力和预测能力影响甚微。即使训练规模较大（60次运行），高噪声情形下重新评估的常微分方程误差仍比低噪声情形高出约一个数量级。这表明，尽管物理约束能够缓解过拟合，但高幅度的偶然不确定性仍不可避免地会在参数估计中引入残余偏差，

从而阻碍物理状态收敛至真实参考状态。随着噪声水平升高，这两个评估指标也表现出切换行为。在低噪声情形下，重新评估的常微分方程误差低于PIRNN预测；在中等噪声情形下，两项指标收敛至相近的误差水平；而在高噪声情形下，PIRNN预测明显优于重新评估的常微分方程。这一现象表明，应根据可用数据的质量谨慎决定PIRNN的双重用途。

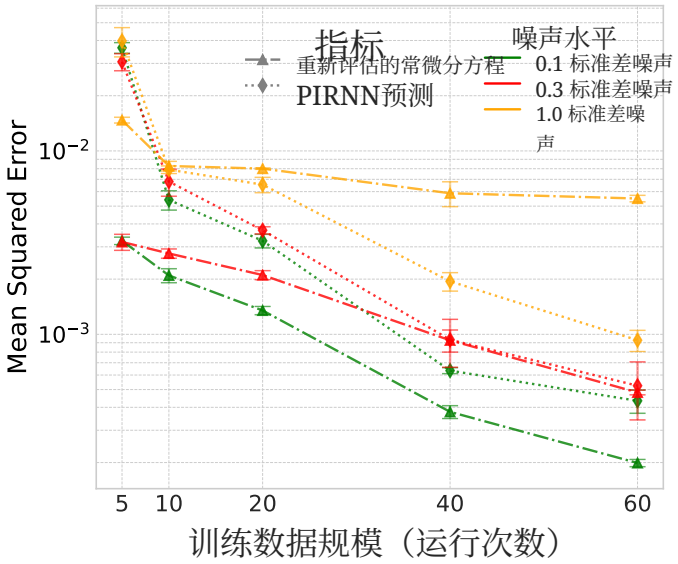


图3：不同噪声水平下，PIRNN预测状态和学习得到的PBM参数与真实无噪声参考之间的测试均方误差。

图4展示了一个随机选取的测试浓度轨迹，以及使用不同数量的训练运行次数、在不同噪声水平下训练的模型所给出的预测结果。在所有情况下，PIRNN预测和重新评估的常微分方程曲线均与真实值参考相符。对于无噪声至中噪声情形（无噪声、0.1 标准差噪声和 0.3 标准差噪声），预测的浓度分布与参考值完全重合，表明模型能够精确、准确地重建结晶动力学。随着噪声水平从 0.3 标准差升至 1.0 标准差，PIRNN预测开始出现轻微的噪声过拟合，但整体预测仍然平滑，波动极小，体现了PIRNN结构的高噪声容忍度 [56]。与此同时，在 1.0 标准差噪声下，重新评估的常微分方程曲线与其他所有情形一致，说明即使噪声量相当大，物理信息循环神经网络仍具有鲁棒性和物理一致性。根据这些结果，物理信息循环神经网络架构能够有效约束模型，使其遵循物理上合理的轨迹。即使存在较高的偶然不确定性，模型仍能捕捉重要动态特征的幅度和时序，例如过饱和度衰减和最终稳态。通过对学习得到的动力学参数进行积分得到的重新评估的常微分方程曲线，进一步证实了学习到的动力学与控制PBM之间的一致性。

评估PIRNN性能的另一种方法是分析物理模型所反演出的参数值。表1将PIRNN估计的动力学参数与Li等人此前报道的参考值进行比较， [73], 这些参考值用于生成合成数据。在无噪声情况下，所有学习到的参数都迅速收敛至文献中的参考值；当运行次数达到10次或更多时，所有量的偏差均低于10%。标准差通常较小，并且随着参与训练的运行次数增加而减小。例外情况是 k_{b2}和 β，这是因为这两个参数高度相关，并且出现在模型的同一项中。如第2.1节所示，PBM参数并非完全相互独立，尤其是在可用数据量有限时。因此， k_{b2} 的降低会由 β的增加来补偿，因为可用数据量不足以分离这两个参数的影响。

随着偶然不确定性水平升高，估计得到的PBM参数精度自然会下降，这体现在标准差增大以及与参考值的偏差增加上。然而，将这些参数差异与图4中的重新评估的常微分方程曲线进行比较后可以清楚地看出，参数精度的下降并不会导致动态预测出现显著且成比例的失效。这种脱钩表明，在高偶然不确定性存在的情况下，逆问题在实践中无法唯一确定解。此时，优化景观会变得平坦，使高度相关的参数能够相互补偿。因此，PIRNN开始提取一组在动力学上等效的参数，从而再现相似的状态轨迹。这种行为

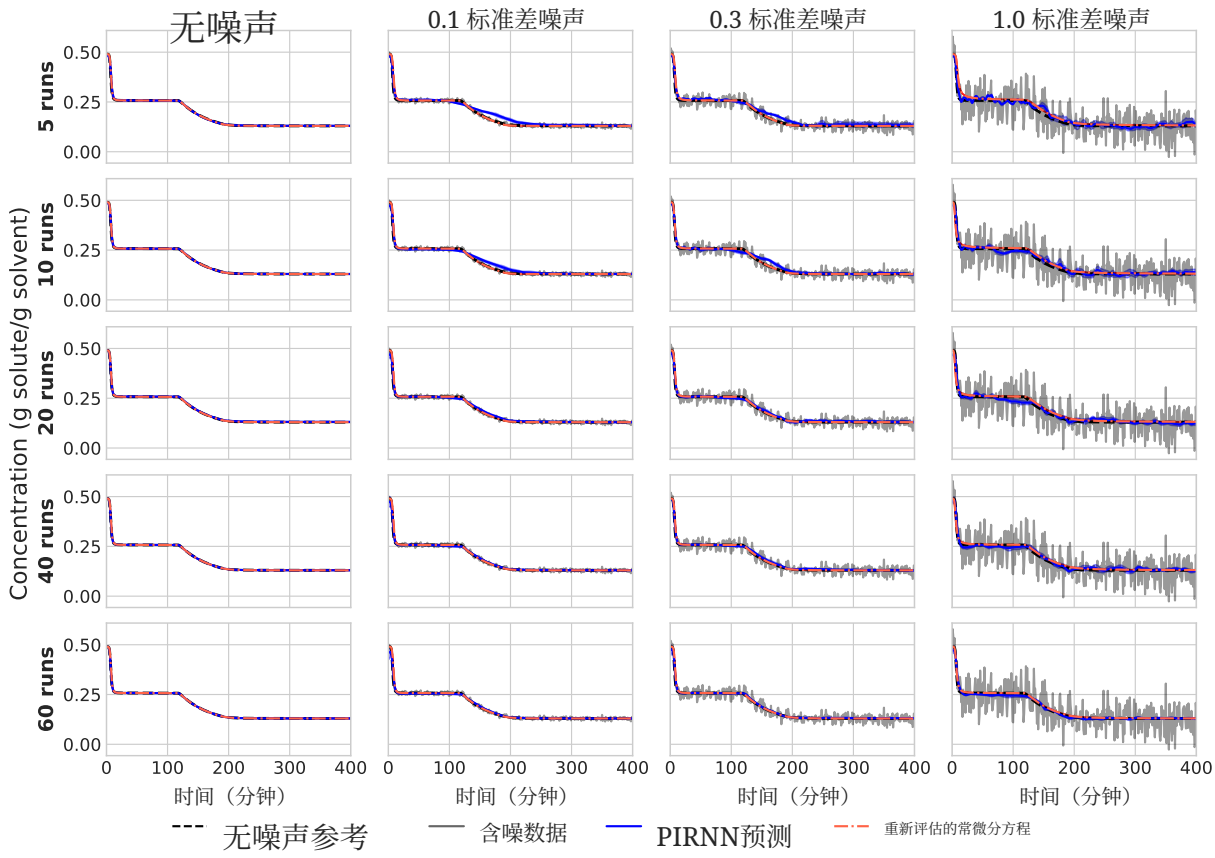


图4：使用 10、20、40 和 60 次训练运行对一组测试数据进行浓度预测。图中给出了 95% 置信区间。每种情况下均显示原始含噪数据和无噪声真实参考值。其他状态变量的预测结果见补充信息中的图 S10 至图 S13。

进一步表明了物理信息循环神经网络双重用途的稳健性。尽管较高的测量噪声限制了内在动力学参数的唯一性，该框架仍能有效地充当稳健的状态估计器，准确重建过程动力学。

3.2 溶解度偏移

实际结晶数据中常见的一种情况是溶解度模型的校准存在不确定性，并且可能在不同运行之间发生漂移。如第2.2节所述，我们在此对溶解度关系引入恒定的 10% 偏移，并评估 PIRNN 直接从测量数据中学习这一偏差的能力。为简化分析，我们省略了信号噪声。

为应对这种模型校准偏差，PIRNN 将物理正则化权重 λ 作为可调参数，用于控制经验灵活性与机理理解之间的权衡。当 λ 较低时，模型的行为类似于纯循环神经网络，优先将数据残差降至最低。尽管这使网络能够在溶解度偏移的情况下保持对观测到的动力学的拟合，却牺牲了群体平衡模型中其他物理定律所提供的有用信息，例如成核和生长速率。相反，当 λ较高时，群体平衡模型方程几乎成为硬约束。这会强制模型严格遵循物理一致性，但也迫使预测遵循存在偏置的溶解度曲线，忽略实验测量中包含的校正信息。因此，确定一个合理的正则化区间至关重要：在此区间内，物理学能够引导优化，使其趋向物理上合理的解，同时又不忽略观测数据，从而有助于应对认知不确定性的影响。

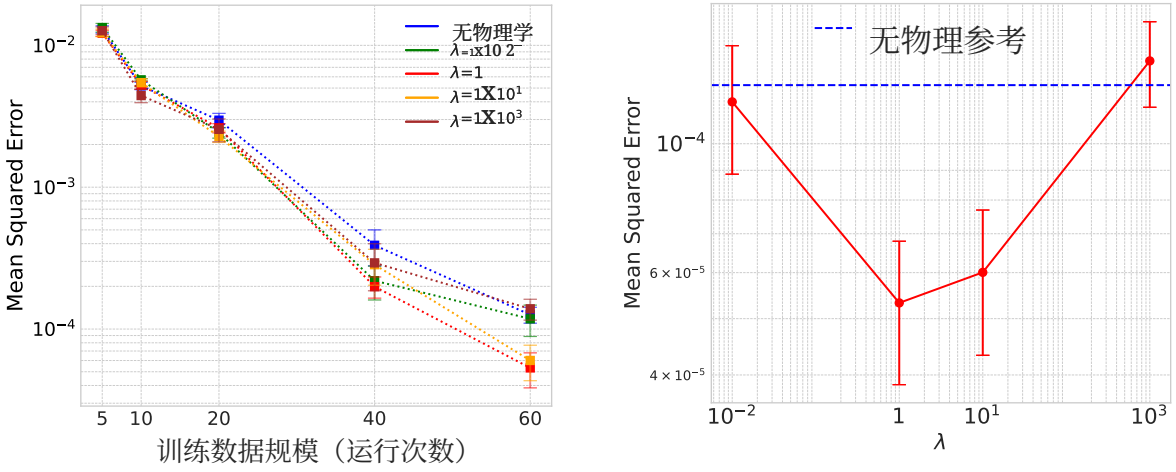
图5展示了在此模型失配条件下，改变物理权重系数对测试均方误差的影响。可以看出，与纯数据驱动基线或弱正则化（$\lambda = 10^{-2}$）的对照案例相比，采用基于物理的正则化显著提升了模型的泛化能力。当物理权重相对于数据损失较低时，训练主要受数据驱动，损失曲线与纯数据驱动基线高度吻合

PBM 参数	k_{b2}	α	β	k_g	E_{ag}	γ_g
参考值	6.000×10^3	2.080	0.713	2.730×10^5	4.130×10^4	1.240
无噪声						
5 次运行	$5.455(43) \times 10^3$	2.041(8)	0.725(1)	$8。626(7044) \times 10^4$	$3.792(161) \times 10^4$	1.239(6)
10 次运行	$5.467(40) \times 10^3$	2.052(4)	0.726(1)	$2.036(266) \times 10^5$	$4.066(34) \times 10^4$	1.228(1)
20次运行	$5.400(30) \times 10^3$	2.054(2)	0.729(1)	$2.213(316) \times 10^5$	$4.088(40) \times 10^4$	1.227(1)
40次运行	$5.380(10) \times 10^3$	2.057(1)	0.729(1)	$2.465(108) \times 10^5$	$4.120(12) \times 10^4$	1.225(1)
60次运行	$5.376(9) \times 10^3$	2.058(0)	0.730(0)	$2.563(65) \times 10^5$	$4.130(7) \times 10^4$	1.225(1)
0.1 标准差噪声						
5 次运行	$5.184(49) \times 10^3$	1.878(5)	0.683(3)	$1.259(984) \times 10^2$	$2.116(191) \times 10^4$	1.279(9)
10 次运行	$4.839(61) \times 10^3$	1.944(5)	0.712(4)	$2.083(1072) \times 10^3$	$2.875(150) \times 10^4$	1.242(7)
20次运行	$5.220(55) \times 10^3$	1.961(8)	0.697(3)	$1。169(352) \times 10^4$	$3.345(81) \times 10^4$	1.232(9)
40次运行	$4.865(78) \times 10^3$	1.975(3)	0.714(4)	$7。437(1771) \times 10^4$	$3.839(71) \times 10^4$	1.204(4)
60次运行	$4.791(42) \times 10^3$	1.974(3)	0.718(2)	$1。257(179) \times 10^5$	$3.982(41) \times 10^4$	1.198(1)
0.3 标准差噪声						
5 次运行	$4.570(286) \times 10^3$	2.065(56)	0.717(20)	$1.450(731) \times 10^1$	$1。527(130) \times 10^4$	1.434(13)
10 次运行	$4.884(177) \times 10^3$	2.104(36)	0.716(9)	$2.200(862) \times 10^2$	$2.249(113) \times 10^4$	1.379(12)
20次运行	$6。272(300) \times 10^3$	2.165(50)	0.669(10)	$3.171(1737) \times 10^3$	$2.914(53) \times 10^4$	1.368(11)
40次运行	$5。259(274) \times 10^3$	2.069(33)	0.704(15)	$3.517(1130) \times 10^4$	$3.595(107) \times 10^4$	1.288(11)
60次运行	$4.984(228) \times 10^3$	2.059(30)	0.716(11)	$6.538(1755) \times 10^4$	$3.770(72) \times 10^4$	1.270(8)
1.0 标准差噪声						
5 次运行	$2.245(454) \times 10^3$	2.322(185)	0.779(52)	$2.475(3385) \times 10^1$	$1.524(305) \times 10^4$	1.629(42)
10 次运行	$3.851(644) \times 10^3$	2.495(235)	0.688(40)	$1.517(1381) \times 10^2$	$2.022(308) \times 10^4$	1.615(69)
20次运行	$5.527(894) \times 10^3$	2.381(89)	0.622(44)	$8.688(7578) \times 10^3$	$2.521(211) \times 10^4$	1.573(39)
40次运行	$5.834(927) \times 10^3$	2.601(296)	0.620(41)	$1.059(874) \times 10^5$	$3.774(240) \times 10^4$	1.496(66)
60次运行	$5.228(578) \times 10^3$	2.540(268)	0.648(29)	$1.474(931) \times 10^5$	$3.915(204) \times 10^4$	1.446(81)

表1：在不同训练数据规模下，PIRNN在不同噪声水平下估计的PBM参数。参考值来自Li等人此前的工作 [73]。所有参数均在对数尺度上进行优化，以保证其为正值。标准差采用括号表示，其中括号内的数字是对应末尾位数的数值不确定性。例如，1.974(3) = 1.974 ± 0.003。关于这种括号表示法的详细信息，请参阅参考文献 [99]。

。对于小型数据集（≤10 次运行），所有模型的误差水平相近，这表明可用数据量是主要限制因素。随着训练规模增大，适度的物理权重（$\lambda = 1 \times 10^{[0,1]}$）可加快收敛，并使测试性能提升数个数量级；在60次运行时，测试均方误差降至 10^{-4} 以下，尽管嵌入的物理规律与训练数据并非完全匹配。在此，加入作为软约束的物理规律，将假设空间缩小到近似遵循“真实”物理模型的函数范围内。尽管数据与模型之间存在失配，但模型的结构为各量之间的平滑性和相关性提供了有用信息。如果失配程度适中，加入物理正则化仍可降低模型方差。同时，由于物理定律被纳入为软约束，网络仍保有足够的灵活性，能够补偿溶解度误差，而不会对观测到的动态趋势造成欠拟合。这种平衡符合机器学习领域广泛采用的经典偏差-方差权衡：引入物理规律所带来的偏置小于其消除的方差 [88, 100–103]。近似的群体平衡模型（PBM）充当归纳偏置，在正则化解的同时允许数据驱动的调整，以应对认知不确定性，从而提高模型的泛化能力并使训练过程更加稳定。

有趣的是，过度增加物理权重会降低模型性能：当训练规模超过40次运行，或正则化强度超过 10^1时，误差会明显增大。这是因为网络会偏向基于错误溶解度模型的物理模型，从而限制其适应经验数据空间的能力。图6展示了这一点：PIRNN预测和重新评估的常微分方程都开始



(a) (b)

图5：不同强度的物理正则化下，偏移溶解度模型的测试误差比较。(a) 绘制了测试数据上的均方误差随训练数据规模变化的曲线，对比了从 10^{-2} 到 10^{3} 的不同物理权重（方程 3.2 中的 λ）。(b) 更详细地展示了 60 次训练运行情况下均方误差随 λ 的变化。

偏离测试数据，尽管它们仍然都优于使用参考参数值进行直接 ODE 积分的结果。这表明模型过于僵化地遵循了可能存在误设的约束。与物理权重适中的情况相比，此时物理约束起到归纳偏置的作用，应与数据保真度达到最佳平衡；数据损失与模型损失开始相互竞争，导致模型性能如预期般下降。因此，在存在认知不确定性的情况下，直接使用 PIRNN 预测显然优于使用参考参数值进行 ODE 积分；通过物理正则化，重新评估的常微分方程也更接近测试数据，尽管由于模型失配，它不可能达到完美拟合。

总体而言，适当加权的物理学信息能够提升模型的鲁棒性，使物理信息循环神经网络在保持较低泛化误差的同时，弥补认知不确定性的影响。重要的是，这些结果凸显了混合架构的一项独特优势：与使用参考参数或学习得到的参数进行常微分方程评估相比，直接使用 PIRNN 进行前向预测能够获得更高的准确性，如图6所示。由于底层物理学存在结构性溶解度偏差，严格积分控制方程会将这一误差传递至状态轨迹。相比之下，物理信息循环神经网络能够灵活地进行修正：利用物理学信息进行正则化，同时学习与经验观测相符的轨迹。该结果表明，混合模型有望弥合不完善的领域知识与现实世界观测之间的差距。

3.3 采样频率的局限性

实际结晶场景中的另一项常见挑战是采样频率受限。为评估 PIRNN 在采样不确定性（即测量极为稀疏的条件）下的稳健性，我们首先考虑了一种极端情况：训练时仅使用每个状态变量轨迹的初始（$t = 0$ 分钟）和最终（$t = 500$ 分钟）时间点。这种设置模拟了这样一种情形：在批次运行的大部分时间内无法获取 PAT 测量值，只有运行结束时才能获得筛分晶体质量。图7展示了在不同物理约束强度下，μ_0、μ_1、μ_2和 C 的归一化预测轨迹。在没有物理约束的情况下（$\lambda = 0$），纯数据驱动模型无法捕捉瞬态动力学，只能在起点和终点之间进行基本上任意的平滑插值，因而无法正确反映冷却诱导成核和生长事件。相反，在另一个极端，即物理正则化过强时（$\lambda = 10^{10}$），模型几乎完全由“物理学驱动”，但它并不具备准确的物理参数先验知识（也就是说，它满足微分方程，但微分方程的参数与数据不一致）。由于所有 PBM 参数的初始值均为 1，远离参考值，优化器可以通过选择使 PBM 计算出的导数几乎衰减至零的物理参数，收敛到平凡解。同时，由于数据损失相对于物理损失的权重可以忽略不计，优化器可以偏离已知边界点以减小有限差分导数，从而将物理损失项降至 $\sim$0。在中间区间可以看到，逐步引入物理约束（随着 λ 从 1 增至 10^2 再增至 10^4,），能够提升模型恢复未观测动力学的能力。在最优点（$\lambda = 10^4$），物理损失项与数据损失项可以达到平衡，

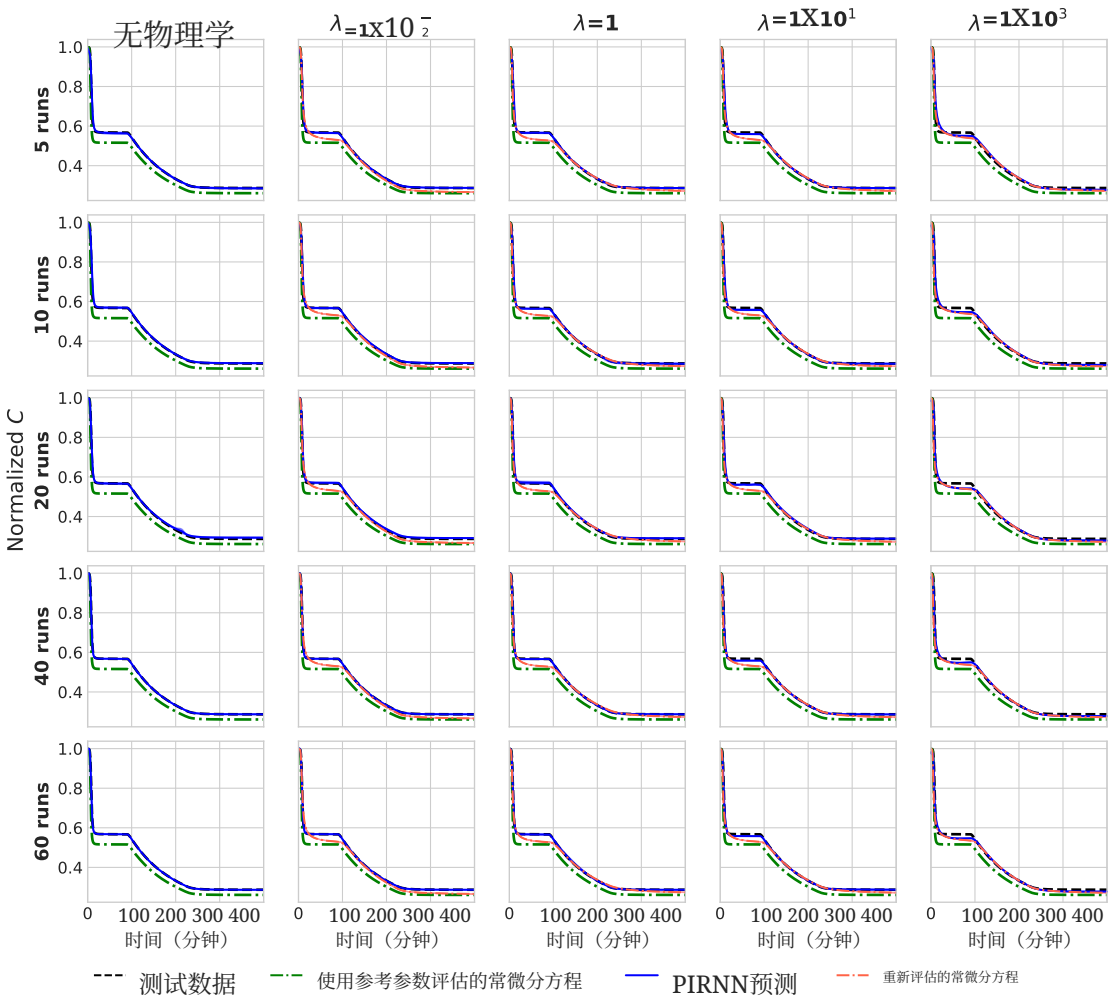


图6：在施加 10% 恒定溶解度偏移时，使用不同 λ 下的 5、10、20、40 和 60 次训练运行，对一组测试数据进行归一化 C 预测。由于未将 PBM 参数纳入优化，因此未展示无物理学情形下重新评估的常微分方程曲线。常微分方程曲线使用两组不同的参数进行评估：参考值（绿色）和 PIRNN 学习得到的数值（红色）。其他状态变量的预测结果见补充信息中的图S14至S17。

从而使边界点的信息能够影响学习得到的 PBM 参数，生成具有物理意义的轨迹。需要注意的是，尽管状态测量值仅包含边界值，模型仍可获取完整的温度轨迹，因为实验中的受控变量包括温度。这也使 PIRNN 能够捕捉与温度变化高度耦合的动力学，即使无法获得完整的状态测量值。

为确定能够平衡数据保真度与物理一致性的最佳物理约束强度，我们分析了在不同 λ量级下数据损失与加权物理损失（$\lambda\mathcal{L}_{physics}$）之间的权衡。如图8所示，在低权重区间（$\lambda \leq 10^2$），加权物理损失保持较小，表明物理约束对优化景观的影响微乎其微；缺乏物理正则化，导致了此前图7中观察到的物理不一致插值。随着 λ 增大并趋近最佳值（$\sim 10^4$），加权物理损失开始上升，并与数据损失相交。这表明物理约束已开始发挥作用，并开始塑造潜在轨迹。在此区间内，数据损失保持稳定且平坦，说明模型成功满足了控制PBM，同时未违反边界数据点。因此，在这种情况下，λ 约为 10^4 时最优，两个损失项在 10^4处达到平衡。超过这一临界阈值（$\lambda > 10^6$）后，随着优化过程被物理项主导、拟合观测边界的能力随之降低，数据损失迅速增大。当 $\lambda = 10^{10}$ 时，加权物理损失随后下降，这表明网络偏向于一种平凡解：通过不符合物理规律地将导数推至零来满足常微分方程。

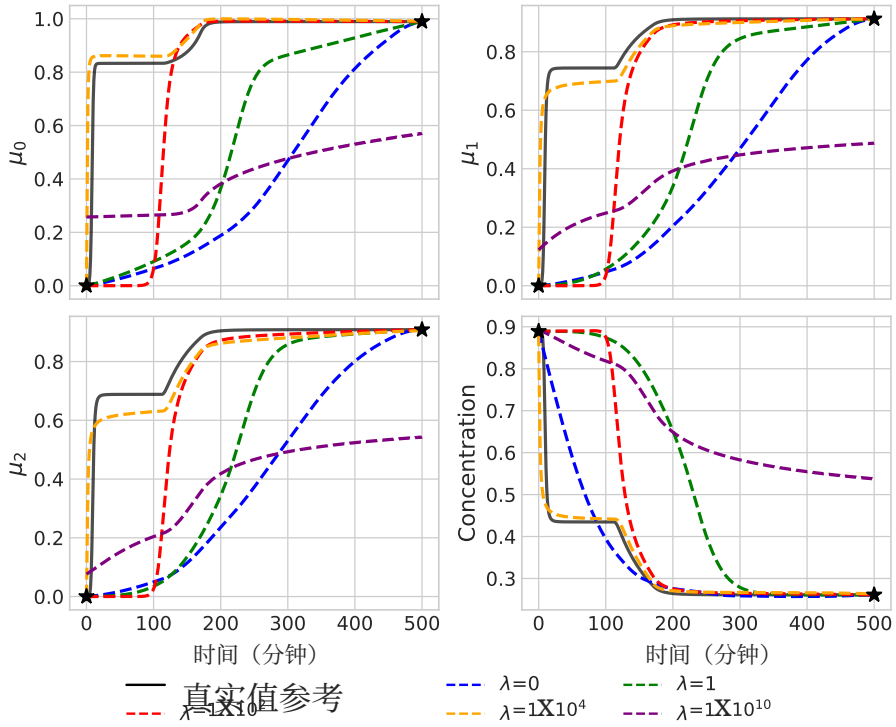


图7：仅使用初始点和终止点进行训练时，PIRNN对归一化的 μ_0、 μ_1、 μ_2和 C 的预测轨迹（图中的黑色星形标记）。真实参考值以黑色实线表示。由于 μ_3 可通过使用 C进行质量衡算得出，因此此处未显示。

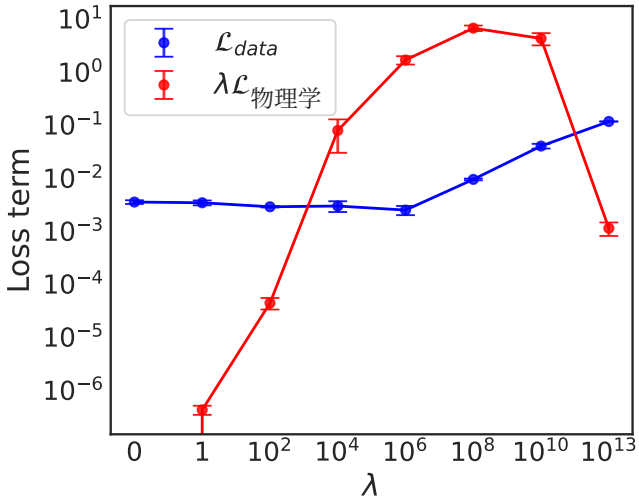


图8：仅使用初始点和终止点进行训练时，不同 λ 值对应的数据和PBM模型训练损失。数据点和误差棒由5次独立训练的集成结果获得。

尽管PIRNN仅利用边界点就成功重建了未观测到的状态轨迹，但底层动力学参数的准确性揭示了这种极端降采样的局限性，表 2 对此进行了进一步量化。表 2 展示了PIRNN在不同采样密度下估计的PBM参数，范围从仅包含边界点的情况到每次运行包含9个点的稀疏网格。在2点情形下，尽管图7所示的预测轨迹准确，估计参数却与参考值存在显著偏差。这种偏差表明，虽然物理约束将解空间缩小到一条物理上可行的路径，但该问题

在参数可辨识性方面仍然病态，使网络能够找到满足边界状态和预测轨迹的其他参数组合。然而，引入单个中间数据点可显著改善参数恢复，使除 k_g 之外的所有参数都恢复到正确的数量级。随着采样密度增加至每次训练运行 9 个点，估计参数会非常接近参考值，对数尺度上的均方误差为 ∼3.6%。这一发现对实际应用场景颇具前景，因为测量 < 10 个中间点比进行高分辨率采样实用得多。

PBM 参数	k_{b2}	α	β	k_g	E_{ag}	γ_g
参考值	6.000×10^3	2.080	0.713	2.730×10^5	4.130×10^4	1.240
每次运行 2 个点	1。288(271) $\times 10^2$	3.901(46)	1.954(76)	0.307(19)	2.608(183) $\times 10^1$	2.273(8)
3 每次运行的点数	3.344(53) $\times 10^3$	2.078(40)	0.737(4)	2.632(1687) $\times 10^4$	3.602(229) $\times 10^4$	1.211(36)
每次运行的点数：5	3.289(25) $\times 10^3$	1.983(21)	0.726(2)	5.397(2254) $\times 10^4$	3.861(137) $\times 10^4$	1.147(21)
每次运行的点数 9	5.250(19) $\times 10^3$	1.975(1)	0.720(1)	1。759(170) $\times 10^5$	4.054(29) $\times 10^4$	1.189(1)

表 2：不同下采样频率下 PIRNN 估计的 PBM 参数。所有参数均在对数尺度上进行优化，以确保其为正值。λ 设置为 10^4，如正文所述。标准差以括号形式给出。参考值取自 Li等人 的先前研究 [73]。所有模型均成功收敛，收敛情况见补充信息中的图 S18。收敛情况下重新评估的常微分方程曲线和 PIRNN预测也见补充信息中的图S19至S21。

这项对采样不确定性的研究展示了 PIRNN框架 兼具动态插值器和参数估计器的双重用途。PIRNN 能够弥合稀疏测量之间未观测到的间隙，即使测量有限，也能实现符合物理规律的状态重构。同时，尽管仅使用起点和终点也能正确恢复状态趋势，却无法准确求得动力学参数；要解决逆问题固有的非唯一性，至少还需要一个额外的中间观测。这些结果表明，尽管 PIRNN框架 在存在采样不确定性时表现出很强的稳健性和有效性，但时间分辨率对于可靠的参数估计仍至关重要。

4 结论

本文系统评估了 PIRNN 在数据非理想性（包括偶然不确定性和认知不确定性）影响下对间歇结晶进行建模的稳健性和有效性。我们使用测量噪声、模型失配和数据稀疏性水平受控的合成数据集，评估了物理信息循环神经网络架构的泛化能力、物理一致性和参数可解释性。结果表明，PIRNN 为在非理想条件下对结晶动力学建模提供了一种稳健可靠的方法。

引入基于物理的约束能够有效减轻偶然不确定性的影响。尽管提高数据噪声水平会形成不可消除的误差下限，但模型对潜在无噪声动力学的预测准确度仍然很高且稳定。此外，该框架在噪声条件下始终能够恢复具有物理意义、且与参考值高度吻合的参数，这凸显了其即使在训练数据存在随机变化的情况下，仍能学习潜在支配性物理定律的能力。不过仍须指出，测量数据质量对于准确学习动力学至关重要。在高噪声环境中，PIRNN 更像是一种稳健的状态估计器：它能够找到满足动力学约束的动力学等效解，即便数据质量限制了对一组唯一内在参数的识别。

当溶解度模型失配引入认知不确定性时，将不完善的物理模型作为软约束纳入其中，便会显现出明显的偏差-方差权衡。权重设置恰当的物理约束能够正则化模型训练，同时不忽略修正结构模型误差所需的观测证据。在这种情况下，与严格积分控制常微分方程相比，物理信息循环神经网络直接前向预测具有更高的预测准确度，这凸显了混合方法弥合不完善的领域知识与实验观测之间差距的能力。

对采样不确定性的分析表明，该框架能够利用有限信息重构复杂的不可观测瞬态动力学。尽管仅凭边界点不足以提取可靠的物理学，

但加入少量中间数据点即可显著提高所学习动力学的准确性。

本研究验证了物理信息机器学习作为一种强大工具的有效性，可利用不完美数据为结晶过程开发稳健的预测模型。PIRNN框架能够处理随机不确定性和系统不确定性，为实现更可靠的过程监测、优化和控制提供了潜在途径。未来的工作将扩展该框架，使其能够直接纳入原始过程分析技术数据，并探索自适应加权方案，以动态平衡数据拟合度与物理一致性。

致谢

作者感谢 Thomas A.Fanning公平导向工程讲席和佐治亚理工学院工程学院 ADVANCE教授职位提供的经费支持。作者还感谢 Fernando Arrais R.D. Lima 博士就群体平衡模型和矩量法的实现进行讨论。

参考文献

[1] Byeongho Ahn、Luca Bosetti 和 Marco Mazzotti。粒子间能量引发的二次成核。II. 动力学。《晶体生长与设计》，22(1):74–86，2022年1月。ISSN 1528-7483。doi: 10.1021/acs.cgd.1c00928.[2] Christopher L. Burcham、Michael F. Doherty、Baron G. Peters、Sarah L. Price、Matteo Salvalaglio、Susan M. Reutzel-Edens、Louise S. Price、Ravi Kumar Reddy Addula、Nicholas Francia、Vikram Khanna 和 Yong- sheng Zhao。药物数字化设计：从化学结构到晶体多晶型，再到概念性结晶过程。《晶体生长与设计》，24(13):5417–5438，2024年7月。ISSN 1528-7483。doi: 10.1021/acs.cgd.3c01390.[3] Narayan S. Tavare。《工业结晶》。Springer US，波士顿，马萨诸塞州，1995年。ISBN 978-1-4899-0235-1。doi: 10.1007/978-1-4899-0233-7.[4] Mohd R. Abu Bakar、Zoltan K. Nagy、Ali N. Saleemi 和 Chris D. Rielly。直接成核控制对药物结晶过程中晶体粒度分布的影响。《晶体生长与设计》，9(3):1378–1384，2009年3月。ISSN 1528-7483。doi: 10.1021/cg800595v.[5] Jie Chen、Bipul Sarma、James M. B. Evans 和 Allan S. Myerson。药物结晶。《晶体生长与设计》，11(4):887–895，2011年4月。ISSN 1528-7483。doi: 10.1021/cg101556s.[6] Fernando Arrais Romero Dias Lima、Marcellus Guedes Fernandes de Moraes、Martha A. Grover、Amaro Gomes Barreto Junior、Argimiro Resende Secchi 和 Maurício B. de Souza。用于对乙酰氨基酚无晶种间歇冷却结晶的神经网络逆模型控制器。《工业与工程化学研究》，63，2024年11月。ISSN 15205045。doi: 10.1021/ACS.IECR.4C02060.[7] Weiwei Tang、Yufeng Quan、Junbo Gong、Jingkang Wang、Qiuxiang Yin 和 Tonglei Li。伴生多晶型的晶型选择：基于结晶动力学建模的案例研究。《*AIChE*期刊》，67(4)，2021年4月。ISSN 0001-1541。doi: 10.1002/aic.17129.[8] Stephanie J. Urwin、Guillaume Levilain、Ivan Marziano、Jeremy M. Merritt、Ian Houson 和 Joop H. Ter Horst。应对工业结晶开发过程中杂质问题的系统方法。*Organic ProcessResearch & Development*，24(8):1443–1456，2020年8月。ISSN 1083-6160。doi: 10.1021/acs.oprd.0c00166.[9] Youngjo Kim、Yoshiaki Kawajiri、Ronald W. Rousseau 和 Martha A. Grover。采用温度循环策略进行无晶种间歇冷却结晶时，对乙酰氨基酚在乙醇溶液中的成核、生长与溶解建模。《工业与工程化学研究》，62(6):2866–2881，2023年2月。ISSN 15205045。doi: 10.1021/ACS.IECR.2C03465.[10] Fernando Arrais R.D. Lima、Marcellus G.F. de Moraes、Amaro G. Barreto、Argimiro R. Secchi、Martha A. Grover 和 Maurício B. de Souza。机器学习在结晶过程建模与先进控制中的应用：进展与展望。《数字化学工程》，14:100208，2025年3月。ISSN 2772-5081。doi: 10.1016/J.DCHE.2024.100208.[11] J. Ulrich 和 P. Frohberg。工业结晶的问题、潜力与未来。《化学科学与工程前沿》，7(1):1–8，2013年3月。ISSN 2095-0179。doi: 10.1007/s11705-013-1304-y.[12] Z.Q. Yu、J.W. Chew、P.S. Chow 和 R.B.H. Tan。结晶控制的最新进展。《化学工程研究与设计》，85(7):893–905，2007年1月。ISSN 02638762。doi: 10.1205/cherd06234。

[13] Huiquan Wu、Zedong Dong、Haitao Li 和 Mansoor Khan。用于理解药物结晶过程、确保产品质量与安全性的集成式过程分析技术（PAT）方法：FDA科学家的观点。*Organic Process Research &Development*，19(1):89–101，2015年1月。ISSN 1083-6160。doi: 10.1021/op500056a.[14] Lawrence X Yu、Robert A Lionberger、Andre S Raw、Rosario D'Costa、Huiquan Wu 和 Ajaz S Hussain。过程分析技术在结晶过程中的应用。*Advanced DrugDelivery Reviews*，56(3):349–369，2004年。ISSN 0169-409X。doi: 10.1016/j.addr.2003.10.012.[15] Huayu Li、Martha A. Grover、Yoshiaki Kawajiri 和 Ronald W. Rousseau。建立关联晶体粒度分布与FBRM测量值的经验方法。*ChemicalEngineering Science*，89:142–151，2013年2月。ISSN 00092509。doi: 10.1016/j.ces.2012.10.031.[16] Matthew A. McDonald、Hossein Salami、Patrick R. Harris、Colton E. Lagerman、Xiaochuan Yang、Andreas S. Bommarius、Martha A. Grover 和 Ronald W. Rousseau。反应结晶：综述。*Reaction Chemistry&Engineering*，6(3):364–400，2021年。ISSN 2058-9883。doi: 10.1039/D0RE00272K.[17] Saikat Roy、Brianna Chamberlin 和 Adam J. Matzger。利用低波数拉曼光谱鉴别多晶型物。*OrganicProcess Research & Development*，17(7):976–980，2013年7月。ISSN 1083-6160。doi: 10.1021/op400102e.[18] Hossein Salami、Matthew A. McDonald、Andreas S. Bommarius、Ronald W. Rousseau 和 Martha A. Grover。原位成像与深度学习相结合的结晶过程监测：在头孢氨苄生产中的应用。*Organic Process Research & Development*，25(7):1670–1679，2021年7月。ISSN 1083-6160。doi: 10.1021/acs.oprd.1c00136.[19] Faten Farouk、Rania M. Hathout 和 Ehab F. Elkady。利用多变量分析方法解决制药过程监测中的分析难题：在过程理解、控制与改进中的应用。*Spectroscopy(Santa Monica)*，38(3)，2023年3月。ISSN 08876703。doi: 10.56530/SPECTROSCOPY.OP4571N3.[20] Zoltan K. Nagy 和 Richard D. Braatz。结晶控制的进展与新方向。*Annualreview of chemical andbiomolecular engineering*，3:55–75，2012年7月。ISSN 1947-5438。doi: 10.1146/ANNUREV-CHEMBIOENG-062011-081043.[21] Noor Asma Fazli Abdul Samad、Gürkan Sin、Krist V. Gernaey 和 Rafiqul Gani。在结晶过程的过程分析技术（PAT）系统设计中引入成核与晶体生长模型的不确定性分析。*EuropeanJournal of Pharmaceutics and Biopharmaceutics*，85(3):911–929，2013年11月。ISSN 09396411。doi: 10.1016/j.ejpb.2013.05.016.[22] Hecham M. Omar 和 Sohrab Rohani。晶体群体衡算的表述与求解方法：综述。《晶体生长与设计》，17(7):4028–4041，2017年7月。ISSN 1528-7483。doi: 10.1021/acs.cgd.7b00645.[23] Åke C Rasmuson。基于群体平衡建模的结晶过程分析。*Handbook of IndustrialCrystallization*，第172–196页，2019年。doi: 10.1017/9781139026949.006.[24] Botond Szilagyi、Wei-Lee Wu、Ayse Eren、Jaron Mackey、Shivani Kshirsagar、Emoke Szilagyi、Iben Ostergaard、Haiyan Qu、Kushal Sinha、Laurie Mlinar、Daniel Pohlman、Jie Chen、Nandkishor Nere、Moussa Boukerche、Akshay Korde、Ryan Ellis、Michael Lovette、Simon Black、Anna Jawor-Baczynska、Huayu Li、Bing-Shiou Yang、Erwin Irdam、Daniel Patience、Rahn McKeown、Megan Ketchum、Daniel A. Green、Venkata Ramana Reddy、Christopher Polster、Christopher Burcham、Daniel Jarmer、Jeremy Miles Merritt、Aaron Cote、Eric Sirota、Lorenzo Codan、Jochen Schoell、Yiqing Claire Liu、Kevin P. Girard、Samir A. Kulkarni、Yihui Yang、Justin L. Quon 和 Zoltan K. Nagy。跨制药企业合作开发用于药物结晶过程基于模型的数字化设计的仿真工具（CrySiV）。《晶体生长与设计》，21(11):6448–6464，2021年11月。ISSN 1528-7483。doi: 10.1021/acs.cgd.1c00904.[25] Atul H. Bari 和 Aniruddha B. Pandit。用于确定成核、生长、破碎和团聚动力学的序贯结晶参数估计方法。《工业与工程化学研究》，57(5):1370–1379，2018年2月。ISSN 0888-5885。doi: 10.1021/acs.iecr.7b03995.[26] Maximilian. O. Besenhard、Anwesha Chaudhury、Thomas Vetter、Rohit Ramachandran 和 J ohannes G. Khinast。对采用群体衡算方程建模的结晶过程参数估计方法的评估。《化学工程研究与设计》，94:275–289，2015年2月。ISSN 02638762。doi: 10.1016/j. cherd.2014.08.004.[27] Meijin Lu、Silin Rao、Hong Yue、Junjie Han 和 Jingtao Wang。机器学习在晶体行为与结晶过程控制中的应用最新进展。《晶体生长与设计》，24(12):5374–5396，2024年6月。ISSN 1528-7483。doi: 10.1021/acs.cgd.3c01251。

[28] Nahyeon An、Hyukwon Kwon、Hyungtae Cho 和 Junghwan Kim。连续结晶过程中岩浆密度的数据驱动建模。计算机辅助化学工程，49:1813–1818，2022年1月。ISSN 1570-7946。doi: 10.1016/B978-0-323-85159-6.50302-X.[29] Wachira Daosud、Jedsada Thampasato 和 Paisan Kittisupakorn。用于间歇加热/冷却蒸发结晶过程的基于神经网络的建模与控制。《工程学报》，21(1):127–144，2017年1月。ISSN 01258281。doi: 10.4186/ej.2017.21.1.127.[30] Paisan Kittisupakorn、Preeya Somsong、Mohd Azlan Hussain 和 Wachira Daosud。基于神经网络混合模型改进精对苯二甲酸间歇结晶器的晶体粒度分布控制。《工程学报》，21(7):319–331，2017年12月。ISSN 01258281。doi: 10.4186/ej.2017.21.7.319.[31] Qingwei Li 和 Hongjun Duan。中药糖沉淀过程的模型预测控制。载于第*11*届世界智能控制与自动化大会论文集，第4850–4853页。IEEE，2014年6月。ISBN 978-1-4799-5825-2。doi: 10.1109/WCICA.2014.7053535.[32] Fernando Arrais R.D. Lima、Marcellus G.F. de Moraes、Amaro G. Barreto、Argimiro R. Secchi 和 Maurício B. de Souza。用于控制过饱和区与未饱和区结晶过程的神经网络非线性模型预测控制。计算机辅助化学工程，52:1835–1840，2023年1月。ISSN 1570-7946。doi: 10.1016/ B978-0-443-15274-0.50291-2.[33] Siyang Ma、Chao Li、Jie Gao、He Yang、Weiwei Tang、Junbo Gong、Fu Zhou 和 Zhenguo Gao。人工神经网络预测碳酸锂反应结晶的亚稳区宽度。《工业与工程化学研究》，59(16):7765–7776，2020年4月。ISSN 0888-5885。doi: 10.1021/acs.iecr.9b06074.[34] V. Manee、W. Zhu 和 J. A. Romagnoli。用于晶体粒度分布实时表征的深度学习图像传感器。《工业与工程化学研究》，58(51):23175–23186，2019年12月。ISSN 0888-5885。doi: 10.1021/acs.iecr.9b02450.[35] Baggie W. Nyande、Zoltan K. Nagy 和 Richard Lakerveld。结晶动力学的数据驱动辨识。《*AIChE*期刊》，70(5)，2024年5月。ISSN 0001-1541。doi: 10.1002/aic.18333.[36] K. Vasanth Kumar、P. Martins 和 F. Rocha。使用人工神经网络对间歇蔗糖结晶动力学进行建模：与传统回归分析的比较。《工业与工程化学研究》，47(14):4917–4923，2008年7月。ISSN 0888-5885。doi: 10.1021/ie701706v.[37] Matthew R. Wilkinson、Uriel Martinez-Hernandez、Laura K. Huggon、Chick C. Wilson 和 Bernardo Cas- tro Dominguez。利用人工智能预测药物晶体形貌。*CrystEngComm*，24 (43):7545–7553，2022年。ISSN 1466-8033。doi: 10.1039/D2CE00992G.[38] Fernando Arrais R.D. Lima、Marcellus G.F. de Moraes、Argimiro R. Secchi 和 Maurício B. de Souza Jr.。开发用于控制结晶过程浓度的循环神经网络非线性模型预测控制器。《数字化学工程》，5:100052，2022年12月。ISSN 27725081。doi: 10.1016/j.dche.2022.100052.[39] Niranjan Sitapure 和 Joseph Sang-Il Kwon。探索时间序列Transformer在化学系统过程建模与控制中的潜力：不可避免的范式转变？《化学工程研究与设计》，194:461–477，2023年6月。ISSN 02638762。doi: 10.1016/j.cherd.2023.04.028.[40] Niranjan Sitapure 和 Joseph Sang-Il Kwon。机器学习与过程控制相结合：揭示LSTMc的潜力。《*AIChE*期刊》，70(7)，2024年7月。ISSN 0001-1541。doi: 10.1002/aic.18356.[41] Yingzhe Zheng、Xiaonan Wang 和 Zhe Wu。批次结晶过程的机器学习建模与预测控制。《工业与工程化学研究》，61(16):5578–5592，2022年4月。ISSN 0888-5885。doi: 10.1021/acs.iecr.2c00026.[42] Yingzhe Zheng、Tianyi Zhao、Xiaonan Wang 和 Zhe Wu。批次间参数漂移条件下结晶过程的基于在线学习的预测控制。《*AIChE*期刊》，68(11)，2022年11月。ISSN 0001- 1541。doi: 10.1002/aic.17815.[43] Daniel J. Griffin、Martha A. Grover、Yoshiaki Kawajiri 和 Ronald W. Rousseau。将数据驱动建模与动态规划应用于间歇冷却结晶。《工业与工程化学研究》，55(5):1361–1372，2016年2月。ISSN 0888-5885。doi: 10.1021/acs.iecr.5b03635.[44] William Bradley、Jinhyeun Kim、Zachary Kilwein、Logan Blakely、Michael Eydenberg、Jordan Jalvin、Carl Laird 和 Fani Boukouvala。关于第一性原理建模与数据驱动建模融合的观点。《计算机与化学工程》，166:107898，2022年10月。ISSN 0098-1354。doi: 10.1016/J.COMPCHEMENG. 2022.107898.[45] Artur M. Schweidtmann、Dongda Zhang 和 Moritz von Stosch。混合建模方法综述与展望。《数字化学工程》，10:100136，2024年3月。ISSN 27725081。doi: 10.1016/j.dche.2023. 100136.

[46] Mohammed S. Alhajeri、Fahim Abdullah、Zhe Wu 和 Panagiotis D. Christofides。利用含噪数据进行预测控制的物理信息机器学习建模。《化学工程研究与设计》，186: 34–49，2022年10月。ISSN 0263-8762。doi: 10.1016/J.CHERD.2022.07.035.[47] Tuse Asrav 和 Erdal Aydin。用于动态过程系统的物理信息循环神经网络与超参数优化。《计算机与化学工程》，173:108195，2023年5月。ISSN 0098-1354。doi: 10.1016/J.COMPCHEMENG.2023.108195.[48] Federico Atzori、Federico Zedda、Giacomo Cao 和 Alessandro Concas。利用多线性回归程序求解物理信息神经网络的逆问题：在化学工程三个案例研究中的应用。《工业与工程化学研究》，64(34):16916–16936，2025年8月。ISSN 15205045。doi: 10.1021/ACS.IECR.5C01752.[49] C. P. Batuwatta-Gamage、C. M. Rathnayaka、H. C.P. Karunasena、W. D.C.C. Wijerathne、H. Jeong、Z. G. Welsh、M. A. Karim 和 Y. T. Gu。基于物理信息神经网络的代理框架，用于预测干燥过程中植物细胞的水分浓度和收缩率。《食品工程杂志》，332:111137，2022年11月。ISSN 0260-8774。doi: 10.1016/J.JFOODENG.2022.111137.[50] Shengze Cai、Zhiping Mao、Zhicheng Wang、Minglang Yin 和 George Em Karniadakis。流体力学中的物理信息神经网络（PINNs）：综述。《力学学报》，37(12):1727– 1738，2021年12月。ISSN 16143116。doi: 10.1007/S10409-021-01148-1.[51] Solji Choi、Ikhwan Jung、Haeun Kim、Jonggeol Na 和 Jong Min Lee。用于数据驱动计算流体动力学求解的物理信息深度学习。《韩国化学工程学报》，39(3):515–528，2022年3月。ISSN 19757220。doi: 10.1007/S11814-021-0979-X.[52] Gabriel S. Gusmão、Adhika P. Retnanto、Shashwati C. da Cunha 和 Andrew J. Medford。动力学信息神经网络。《今日催化》，2022年4月。ISSN 09205861。doi: 10.1016/j.cattod.2022.04.002.[53] Weiqi Ji、Weilun Qiu、Zhiyu Shi、Shaowu Pan 和 Sili Deng。Stiff-PINN：用于刚性化学动力学的物理信息神经网络。《物理化学杂志*A*》，125(36):8098–8106，2021年9月。ISSN 15205215。doi: 10.1021/ACS.JPCA.1C05102.[54] Ece Serenat Koksal、Tuse Asrav、Elif Ecem Esenboga、Ahmet Cosgun、Gizem Kusoglu 和 Erdal Aydin。采用实际验证的工业废水处理厂物理信息与数据驱动建模。《计算机与化学工程》，189:108801，2024年10月。ISSN 0098-1354。doi: 10.1016/J.COMPCHEMENG.2024. 108801.[55] Farshad Moayedi、Aswin Chandrasekar、Sarah Rasmussen、Samardeep Sarna、Brandon Corbett 和 Prashant Mhaskar。面向过程系统的物理信息神经网络：处理工厂与模型失配。《工业与工程化学研究》，63(31):13650–13659，2024年8月。ISSN 15205045。doi: 10.1021/ACS.IECR.4C00690.[56] Dingqi Nai、Gabriel S. Gusmão、Zachary A. Kilwein、Fani Boukouvala 和 Andrew J. Medford。利用动力学信息神经网络对产品时间分析数据进行微观动力学建模。《数字发现》，3(11):2327–2340，2024年11月。ISSN 2635098X。doi: 10.1039/D4DD00163J.[57] Kar Ming Shaw、Phaik Eong Poh、Yong Kuen Ho、Zhi Yuan Chen 和 Irene Mei Leng Chew。通过物理信息深度学习对棕榈油厂废水的厌氧消化进行建模。《化学工程杂志》，485:149826，2024年4月。ISSN 1385-8947。doi: 10.1016/J.CEJ.2024.149826.[58] Zhiyong Wu、Huan Wang、Chang He、Bingjian Zhang、Tao Xu 和 Qinglin Chen。物理信息机器学习在化学工程多物理场建模中的应用。《工业与工程化学研究》，62(44):18178–18204，2023年11月。ISSN 15205045。doi: 10.1021/ACS.IECR.3C02383.[59] Ming Xiao 和 Zhe Wu。利用物理信息迁移学习对化学过程网络进行建模与控制。《工业与工程化学研究》，62(42):17216–17227，2023年10月。ISSN 15205045。doi: 10.1021/ACS.IECR.3C01435.[60] Baoyu Zhu、Shaojun Ren、Qihang Weng 和 Fengqi Si。考虑单调关系的物理信息神经网络，用于预测燃煤锅炉的NOx排放。《燃料》，364:131026，2024年5月。ISSN 0016-2361。doi: 10.1016/J.FUEL.2024.131026.[61] V. Galvanauskas、P. Georgieva 和 S.F. de Azevedo。基于混合（机理+人工神经网络）模型的工业制糖结晶过程动态优化。载于*2006*年*IEEE*国际神经网络联合会议论文集，第2728–2735页。IEEE，2006年。ISBN 0-7803-9490-9。doi: 10.1109/IJCNN.2006. 247177.

[62] P. Georgieva、M.J. Meireles 和 S. Feyo de Azevedo。考虑成核、生长和团聚现象的间歇结晶知识型混合建模。*Chemical Engineering Science*，58 (16):3699–3713，2003年8月。ISSN 00092509。doi: 10.1016/S0009-2509(03)00260-4。[63] Fernando Arrais R.D. Lima、Carine M. Rebello、Erbet A. Costa、Vinícius V. Santana、Marcellus G.F. de Moares、Amaro G. Barreto、Argimiro R. Secchi、Maurício B. de Souza 和 Idelfonso B.R. Nogueira。利用通用微分方程改进结晶过程建模。《化学工程研究与设计》，200：538–549，2023年12月。ISSN 02638762。doi: 10.1016/j.cherd.2023.11.032。[64] Niranjan Sitapure 和 Joseph Sang-Il Kwon。将时间序列变换器引入混合建模：间歇结晶中串联与并联方法的比较研究。《工业与工程化学研究》，62(49):21278–21291，2023年12月。ISSN 0888-5885。doi: 10.1021/acs.iecr.3c02624。[65] Guoquan Wu、Wallace Tan Gian Yion、Khoa Le Nguyen Quang Dang 和 Zhe Wu。用于模型预测控制的物理信息机器学习：在间歇结晶过程中的应用。《化学工程研究与设计》，192:556–569，2023年4月。ISSN 0263-8762。doi: 10.1016/J.CHERD.2023.02.048。[66] Matthew A. Chan、Maria J. Molina 和 Christopher A. Metzler。使用单一模型估计认知不确定性和偶然不确定性。2024年11月。doi: 10.48550/arXiv.2402.03478。[67] Eyke Hüllermeier 和 Willem Waegeman。机器学习中的偶然不确定性和认知不确定性：概念与方法简介。*Machine Learning*，110(3):457–506，2021年3月。ISSN 0885-6125。doi: 10.1007/s10994-021-05946-3。[68] Armen Der Kiureghian 和 Ove Ditlevsen。偶然不确定性还是认知不确定性？这重要吗？《结构安全》，31(2):105–112，2009年3月。ISSN 01674730。doi: 10.1016/j.strusafe.2008.06.020。[69] Hermann G. Matthies。量化不确定性：概率的现代计算表示及其应用。载于《结构动力学中的极端人为与自然灾害》，第105–135页。施普林格·荷兰，多德雷赫特。doi: 10.1007/978-1-4020-5656-7{_}4。[70] Daniel J. Segalman、Matthew R. Brake、Lawrence A. Bergman、Alexander F. Vakakis 和 Kai Willner。建模中的认知不确定性和偶然不确定性。载于第*8*卷：第*22*届可靠性、应力分析与失效预防会议；第*25*届机械振动与噪声会议。美国机械工程师学会，2013年8月。ISBN 978-0-7918-5599-7。doi: 10.1115/DETC2013-13234。[71] Freddie Bickford Smith、Jannik Kossen、Eleanor Trollope、Mark van der Wilk、Adam Foster 和 Tom Rainforth。重新思考偶然不确定性与认知不确定性。2025年6月。doi: 10.48550/arXiv.2412.20892。[72] Daniele Pessina、Jorge Calderon De Anda、Claire Heffernan、Jerry Y. Y. Heng 和 Maria M. Papathanasiou。用于蛋白质结晶模型的体外/计算机模拟一体化不确定性量化方法。《工业与工程化学研究》，64(24):12025–12035，2025年6月。ISSN 0888-5885。doi: 10.1021/acs.iecr.4c04517。[73] Huayu Li、Yoshiaki Kawajiri、Martha A. Grover 和 Ronald W. Rousseau。无晶种间歇冷却结晶的成核与生长动力学建模。《工业与工程化学研究》，56(14):4060–4073，2017年4月。ISSN 15205045。doi: 10.1021/acs.iecr.6b04914。[74] Doraiswami Ramkrishna。《群体衡算：工程颗粒系统中的理论与应用》。爱思唯尔，2000年。ISBN 0080539246。doi: 10.1016/B978-0-12-576970-9.X5000-0。[75] Doraiswami Ramkrishna 和 Meenesh R. Singh。群体衡算建模：现状与未来展望。*Annual Review of Chemical and Biomolecular Engineering*，5（第5卷，2014年）：123–146，2014年6月。ISSN 19475438。doi: 10.1146/ANNUREV-CHEMBIOENG-060713-040241。[76] Alan Randolph。《颗粒过程理论：连续结晶的分析与技术》。爱思唯尔，2012年。ISBN 978-0-12-579650-7。doi: 10.1016/B978-0-12-579650-7.X5001-5。[77] K. O. Bowman 和 L. R. Shenton。估计：矩量法。载于《统计科学百科全书》。威立，2005年12月。doi: 10.1002/0471667196.ess1618.pub2。[78] Karl Pearson。矩量法与最大似然法。《生物统计学》，28(1/2):34，1936年6月。ISSN 00063444。doi: 10.2307/2334123。[79] Jörg Worlitschek 和 Marco Mazzotti。对乙酰氨基酚间歇冷却结晶中粒度分布的基于模型优化。《晶体生长与设计》，4(5):891–903，2004年9月。ISSN 1528-7483。doi: 10.1021/cg034179b。[80] Allan Myerson。*HandbookofIndustrial Crystallization*。巴特沃思-海涅曼，2002年。doi: 10.1016/ B978-0-7506-7012-8.X5000-9。

[81] Huayu Li、Yoshiaki Kawajiri、Martha A. Grover 和 Ronald W. Rousseau。将经验 FBRM 模型应用于间歇结晶中的晶体粒度分布估计。《晶体生长与设计》，14(2):607–616，2014年2月。ISSN 15287483。doi: 10.1021/cg401484d。[82] Pauli Virtanen、Ralf Gommers、Travis E. Oliphant、Matt Haberland、Tyler Reddy、David Cournapeau、Evgeni Burovski、Pearu Peterson、Warren Weckesser、Jonathan Bright、Stéfan J. van der Walt、Matthew Brett、Joshua Wilson、K. Jarrod Millman、Nikolay Mayorov、Andrew R. J. Nelson、Eric Jones、Robert Kern、Eric Larson、C J Carey、 ̇Ilhan Polat、Yu Feng、Eric W. Moore、Jake VanderPlas、Denis Laxalde、Josef Perktold、Robert Cimrman、Ian Henriksen、E. A. Quintero、Charles R. Harris、Anne M. Archibald、Antônio H. Ribeiro、Fabian Pedregosa、Paul van Mulbregt、Aditya Vijaykumar、Alessandro Pietro Bardelli、Alex Rothberg、Andreas Hilboll、Andreas Kloeckner、Anthony Scopatz、Antony Lee、Ariel Rokem、C. Nathan Woods、Chad Fulton、Charles Masson、Christian Häggström、Clark Fitzgerald、David A. Nicholson、David R. Hagen、Dmitrii V. Pasechnik、Emanuele Olivetti、Eric Martin、Eric Wieser、Fabrice Silva、Felix Lenders、Florian Wilhelm、G. Young、Gavin A. Price、Gert-Ludwig Ingold、Gregory E. Allen、Gregory R. Lee、Hervé Audren、Irvin Probst、Jörg P. Dietrich、Jacob Silterra、James T Webber、Janko Slavi ̆c、Joel Nothman、Johannes Buchner、Johannes Kulick、Johannes L. Schönberger、José Vinícius de Miranda Cardoso、Joscha Reimer、Joseph Harrington、Juan Luis Cano Rodríguez、Juan Nunez-Iglesias、Justin Kuczynski、Kevin Tritz、Martin Thoma、Matthew Newville、Matthias Kümmerer、Maximilian Bolingbroke、Michael Tartre、Mikhail Pak、Nathaniel J. Smith、Nikolai Nowaczyk、Nikolay Shebanov、Oleksandr Pavlyk、Per A. Brodtkorb、Perry Lee、Robert T. McGibbon、Roman Feldbauer、Sam Lewis、Sam Tygier、Scott Sievert、Sebastiano Vigna、Stefan Peterson、Surhud More、Tadeusz Pudlik、Takuya Oshima、Thomas J. Pingel、Thomas P. Robitaille、Thomas Spura、Thouis R. Jones、Tim Cera、Tim Leslie、Tiziano Zito、Tom Krauss、Utkarsh Upadhyay、Yaroslav O. Halchenko 和 Yoshiki Vázquez-Baeza。SciPy 1.0：Python 科学计算的基础算法。《自然·方法》，17(3):261–272，2020年3月。ISSN 1548-7091。doi: 10.1038/s41592-019-0686-2。[83] Youngjo Kim。无晶种间歇冷却结晶最优控制方法的开发：第一性原理与机器学习方法的结合。博士论文，博士论文，佐治亚理工学院，2021年。[84] Clifford T. Ó'Ciardhá、Kieran W. Hutton、Niall A. Mitchell 和 Patrick J. Frawley。有晶种反溶剂结晶的参数估计与优化同步进行。《晶体生长与设计》，12(11):5247–5261，2012年11月。ISSN 1528-7483。doi: 10.1021/cg3006822。[85] Roger A Granberg 和 Åke C Rasmuson。对乙酰氨基酚在纯溶剂中的溶解度。《化学与工程数据杂志》，44(6):1391–1395，1999年。doi: 10.1021/je990124v。[86] Inga Hahnenkamp、Gitte Graubner 和 Jürgen Gmehling。活性药物成分溶解度的测量与预测。《国际药剂学杂志》，388(1-2):73–81，2010年3月。ISSN 03785173。doi: 10.1016/j.ijpharm.2009.12.036。[87] D.J. Widenski、A. Abbas 和 J.A. Romagnoli。用于冷却结晶建模的不同溶解度方程比较。《化学工程与加工：过程强化》，49(12):1284–1297，2010年12月。ISSN 02552701。doi: 10.1016/j.cep.2010.09.018。[88] Federico Girosi、Michael Jones 和 Tomaso Poggio。正则化理论与神经网络架构。《神经计算》，7(2):219–269，1995年3月。ISSN 0899-7667。doi: 10.1162/NECO.1995.7.2.219。[89] Yuan Yao、Lorenzo Rosasco 和 Andrea Caponnetto。关于梯度下降学习中的早停。《构造逼近》，26(2):289–315，2007年8月。ISSN 01764276。doi: 10.1007/S00365-006-0663-2/METRICS。URL https://link.springer.com/article/10.1007/s00365-006-0663-2。[90] M. Raissi、P. Perdikaris 和 G.E. Karniadakis。物理信息神经网络：用于求解涉及非线性偏微分方程的正问题与逆问题的深度学习框架。《计算物理学杂志》，378:686–707，2019年2月。ISSN 00219991。doi: 10.1016/j.jcp.2018.10.045。[91] Sepp Hochreiter 和 Jürgen Schmidhuber。长短期记忆。《神经计算》，9(8):1735–1780，1997年11月。ISSN 0899-7667。doi: 10.1162/neco.1997.9.8.1735。[92] Geoffrey E. Hinton、Nitish Srivastava、Alex Krizhevsky、Ilya Sutskever 和 Ruslan R. Salakhutdinov。通过防止特征检测器的协适应来改进神经网络。2012年7月。doi: 10.48550/arXiv.1207.0580。[93] Sergey Ioffe 和 Christian Szegedy。批归一化：通过减少内部协变量偏移加速深度网络训练。2015年3月。doi: 10.48550/arXiv.1502.03167。[94] Li Wan、Matthew Zeiler、Sixin Zhang、Yann Le Cun 和 Rob Fergus。使用 DropConnect 对神经网络进行正则化。载于国际机器学习会议，第1058–1066页，2013年。doi: 10.5555/3042817。3043055。

[95] Nitish Srivastava、Geoffrey Hinton、Alex Krizhevsky、Ilya Sutskever 和 Ruslan Salakhutdinov。Dropout：一种防止神经网络过拟合的简单方法。机器学习研究期刊，15(1): 1929–1958，2014。doi: 10.5555/2627435.2670313。[96] Alex Kendall 和 Yarin Gal。计算机视觉的贝叶斯深度学习需要哪些不确定性？2017 年 10 月。doi: 10.48550/arXiv.1703.04977。[97] Peter J. Huber。位置参数的稳健估计。数学统计年刊，35(1):73–101，1964 年 3 月。ISSN 0003-4851。doi: 10.1214/aoms/1177703732。[98] Kart Leong Lim、Rahul Dutta 和 Mihai Rotaru。使用有限差分法的物理信息神经网络。会议论文集——*IEEE*国际系统、人类与控制论会议，2022-October: 1828–1833，2022。ISSN 1062922X。doi: 10.1109/SMC53654.2022.9945171。[99] Peter J. Mohr、David B. Newell、Barry N. Taylor 和 Eite Tiesinga。CODATA 推荐的基本物理常数值：2022。《现代物理评论》，97(2):025002，2025 年 4 月。ISSN 0034-6861。doi: 10.1103/RevModPhys.97.025002。[100] Mikhail Belkin、Daniel Hsu、Siyuan Ma 和 Soumik Mandal。调和现代机器学习实践与经典偏差-方差权衡。《美国国家科学院院刊》，116(32):15849–15854，2019 年 8 月。ISSN 10916490。doi: 10.1073/PNAS.1903070116。[101] Stuart Geman、Elie Bienenstock 和 René Doursat。神经网络与偏差-方差困境。《神经计算》，4(1):1–58，1992 年 1 月。ISSN 0899-7667。doi: 10.1162/NECO.1992.4.1.1。[102] Brady Neal、Sarthak Mittal、Aristide Baratin、Vinayak Tantia、Matthew Scicluna、Simon Lacoste-Julien 和 Ioannis Mitliagkas。神经网络中偏差-方差权衡的新视角。2018 年 10 月。doi: 10.48550/ arXiv.1810.08591。[103] Tomaso Poggio、Vincent Torre 和 Christof Koch。计算视觉与正则化理论。《计算机视觉读本》，第 638–643 页，1987 年 1 月。doi: 10.1016/B978-0-08-051581-6.50061-1。

补充信息

利用物理信息机器学习对不确定性下的间歇结晶进行建模

Dingqi Nai[a]、Huayu Li[b]、Martha Grover[a]、Andrew J. Medford[a*]佐治亚理工学院化学与生物分子工程学院，亚特兰大，GA 30332[b]勃林格殷格翰制药公司材料与分析科学部，里奇菲尔德，CT 06877

S1 数据准备

S1.1 合成温度曲线

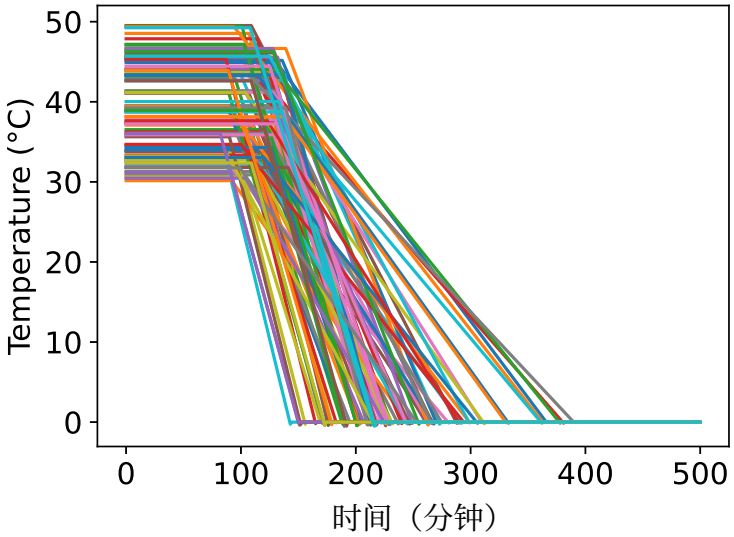


图 S1：100 次合成间歇冷却运行的温度曲线。初始温度、冷却速率和平台持续时间均从均匀分布中随机抽样。

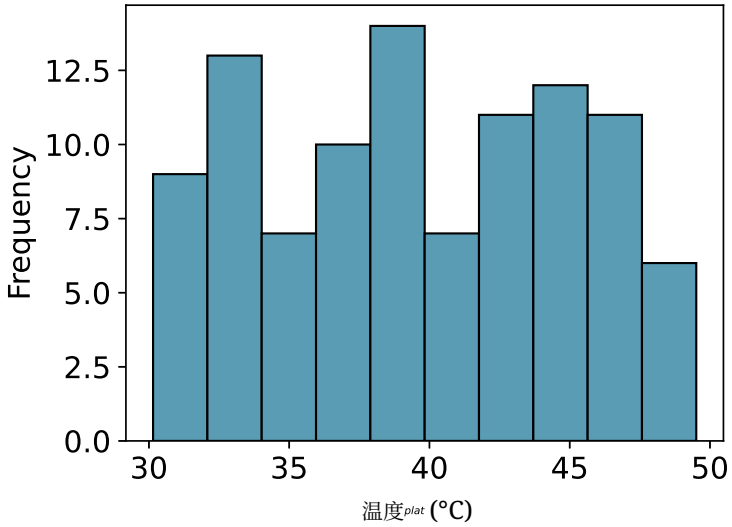


图 S2：平台温度的分布。从 30 到 50 的均匀分布中抽取 100 个点
°C.

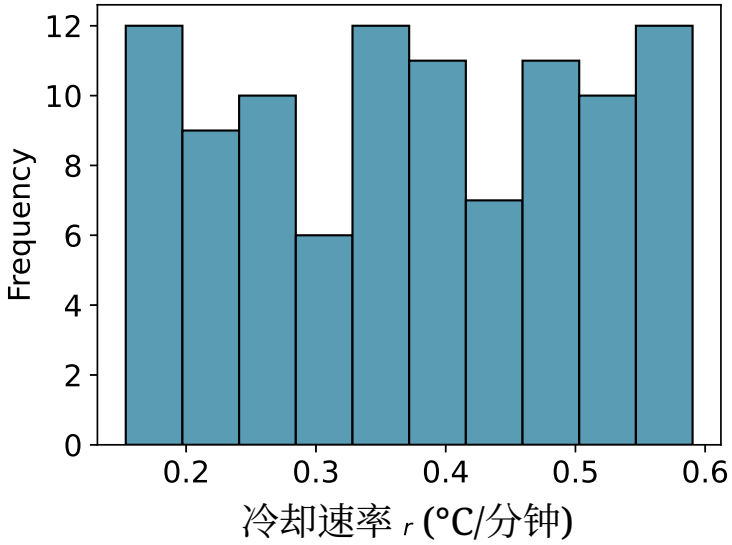


图 S3：冷却速率的分布。从 0.15 到 0.60 的均匀分布中抽取 100 个点
°C/分钟。

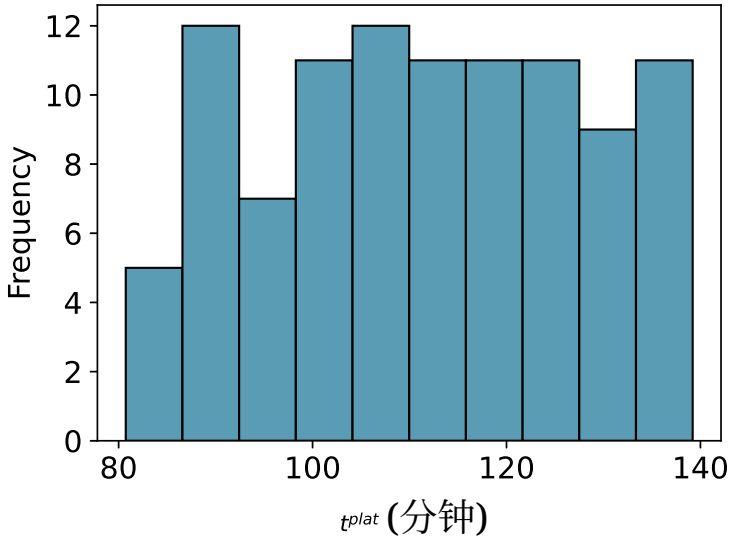


图 S4：平台持续时间的分布。从 80 到 140 分钟的均匀分布中抽取了 100 个点。

S1.2 合成数据

通过使用矩量法对群体平衡模型进行正向求解来合成数据

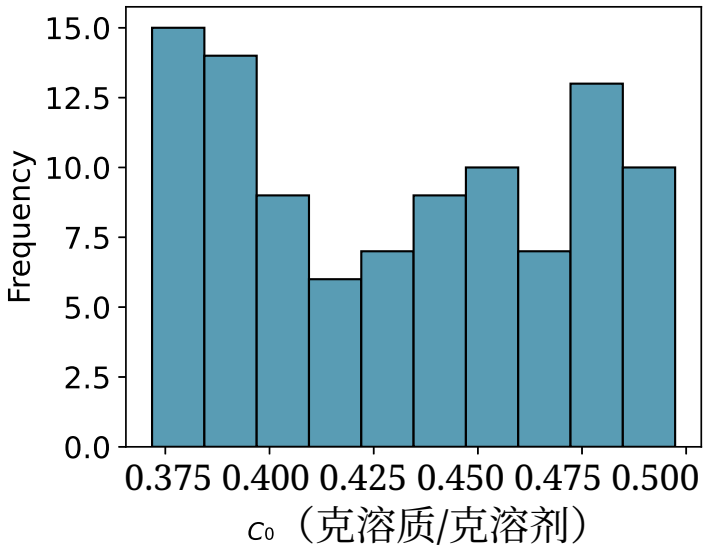


图 S5：用于生成合成数据的初始浓度分布。从 0.37 到 0.50 克溶质/克溶剂的均匀分布中采样 100 个点。

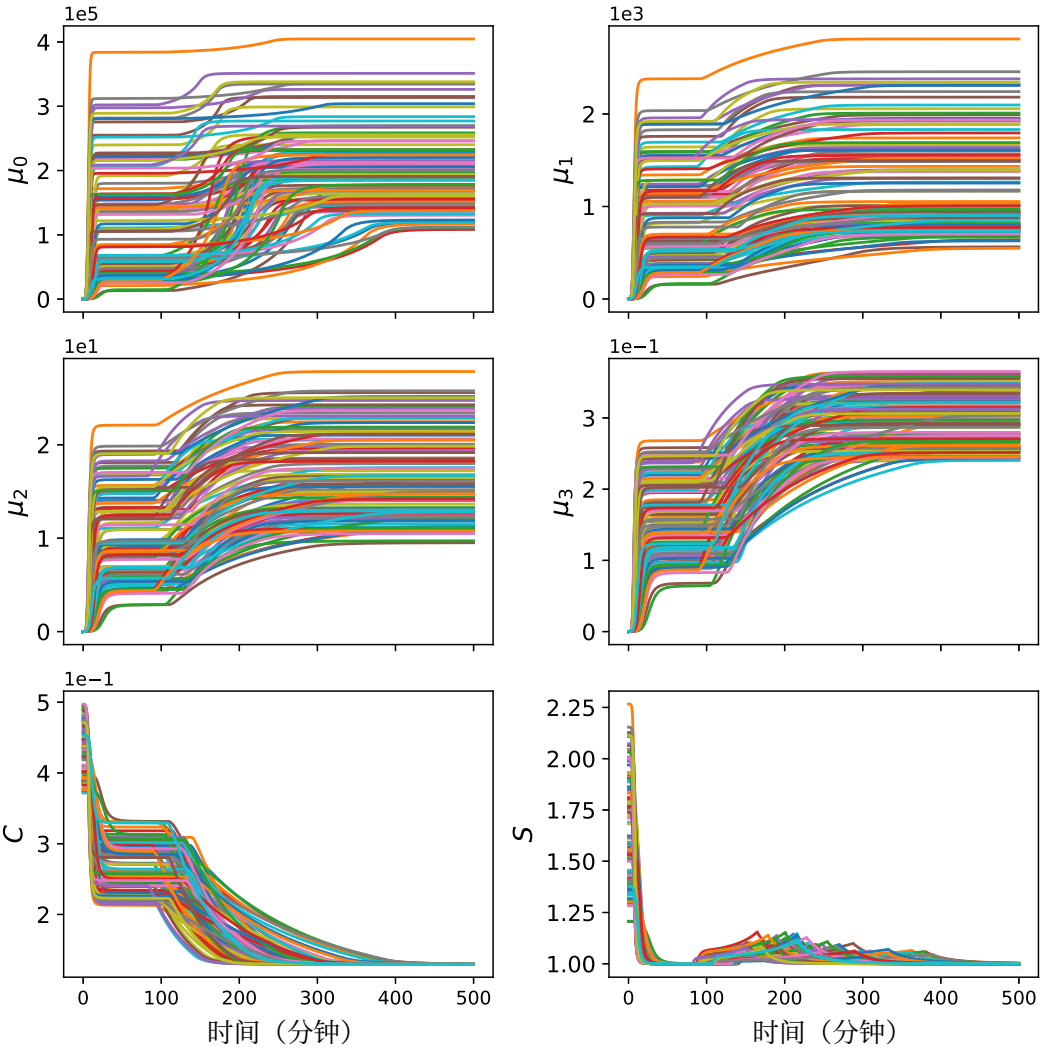


图 S6：合成无噪声数据剖面。同时也包含过饱和度作为参考值，但未将其纳入训练。

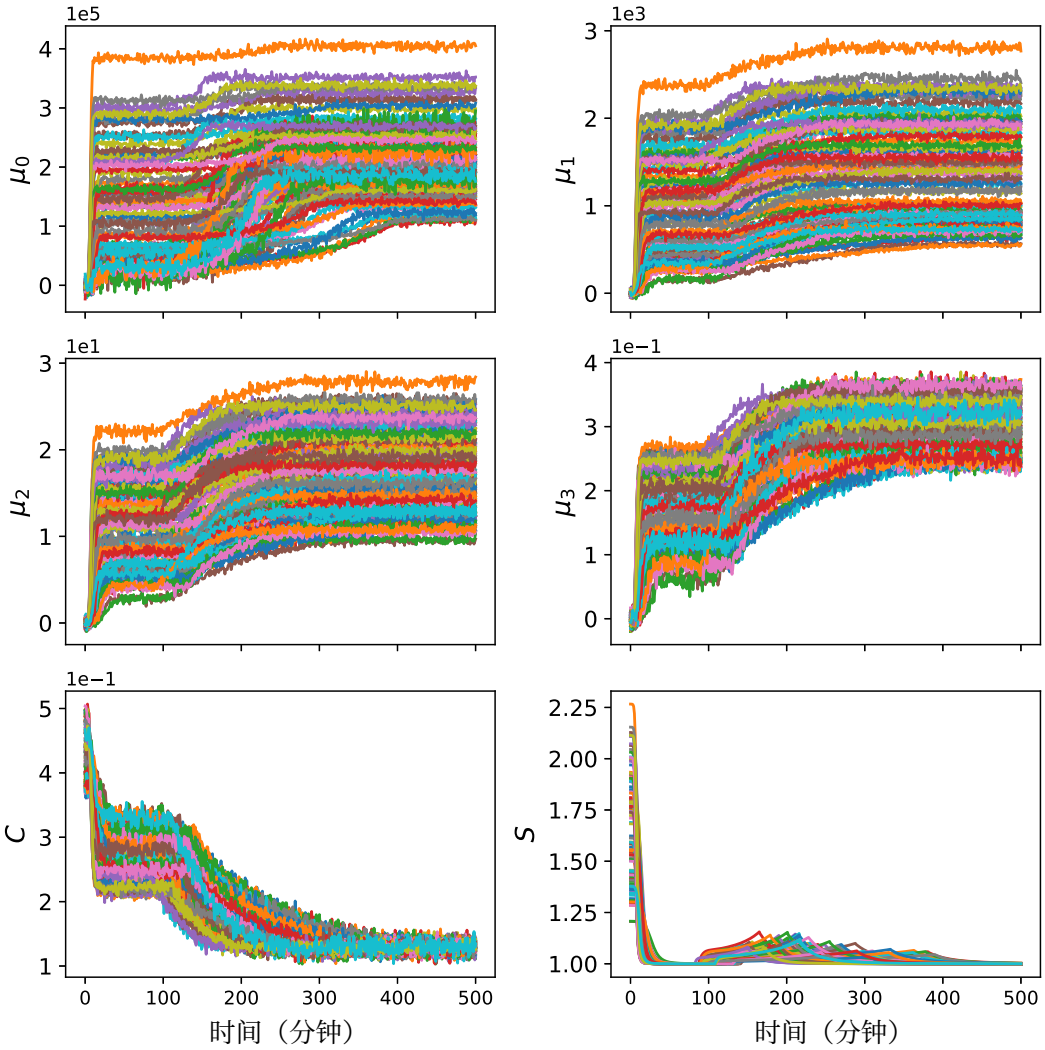


图 S7：合成的0.1标准差噪声水平数据剖面。过饱和度也作为参考值列出，但未纳入训练。

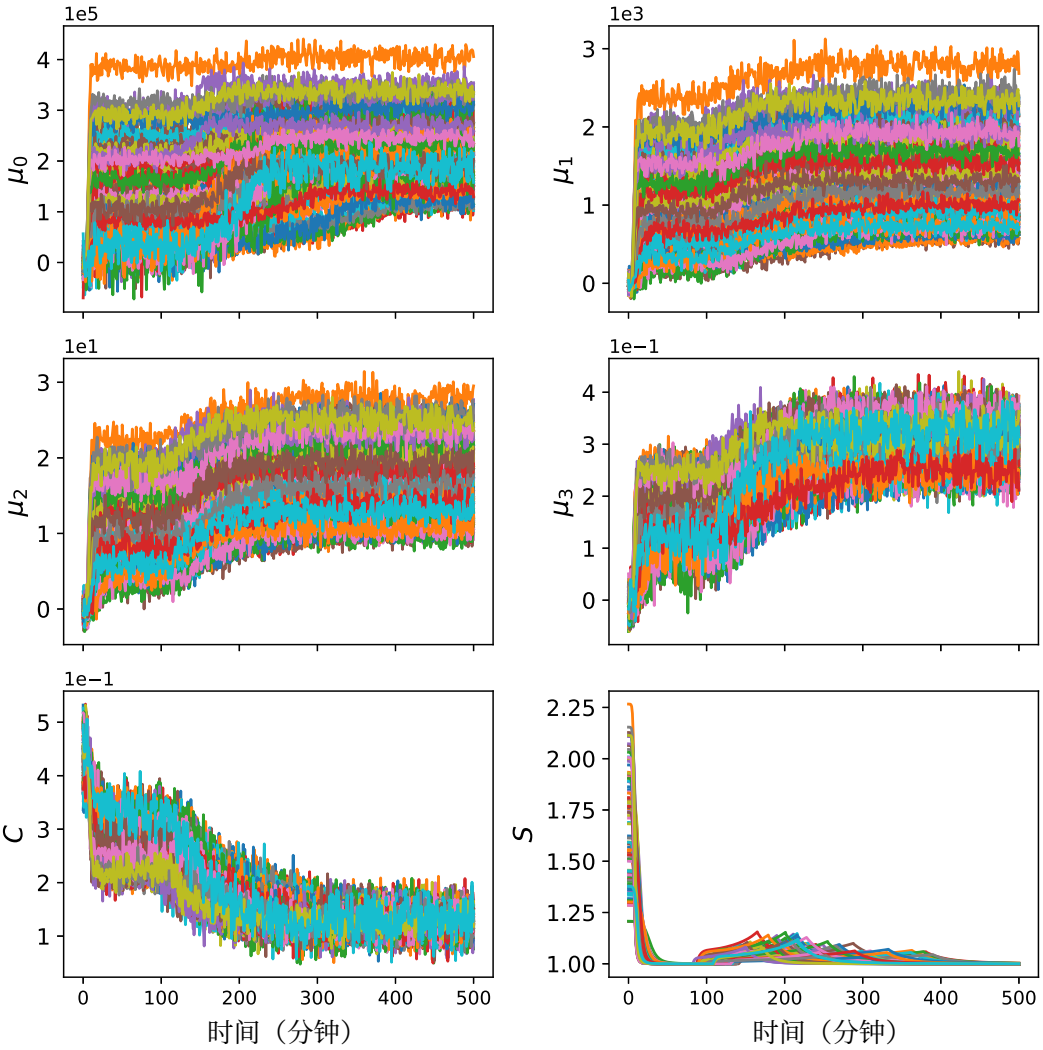


图 S8：合成的 0.3 标准差噪声水平数据剖面。过饱和度也列出供参考，但未纳入训练。

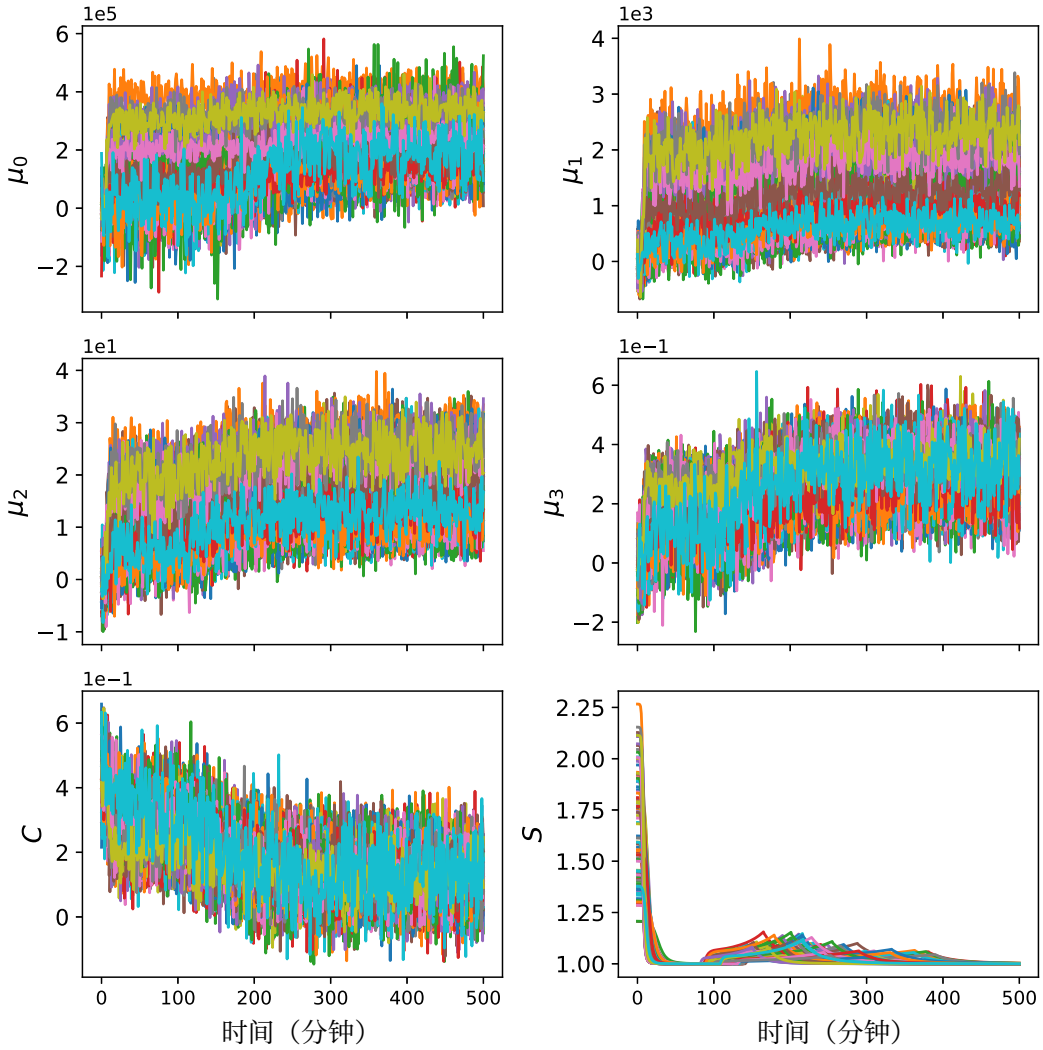


图 S9：合成的 1.0 标准差噪声水平数据剖面。过饱和度也作为参考值一并展示，但未纳入训练。

S1.3 神经网络架构

表 S1 总结了网络架构和训练配置所采用的具体设计与超参数。输入和输出维度根据状态变量和控制变量的数量确定。我们选用 64 个隐藏单元和 2 个 LSTM 层，以确保具备足够的表示能力，同时避免模型过于复杂；选择 64 个隐藏单元也有利于提高硬件效率，使其适配二进制逻辑，而且这也是常用的经验起始值。丢弃率设为 0.2，处于常用范围内，如 Pytorch 文档中的示例所示。由于已知 RNN/LSTM 对 Dropout 比卷积神经网络更敏感，因此采用较低的丢弃率。选择 Softplus 激活函数是因为它平滑且能保证输出非负。由于所有状态变量都归一化至 0 到 1 的范围内，因此初始噪声尺度和转换阈值均设为 0.1。在这种情况下，0.1（10%）可视为相当大的偏差。Adam 是一种广泛使用的优化器，而 1×10^{-3} 是 Pytorch Adam 优化器的默认学习率。

这些设置在所有案例研究中均保持不变。由于本文旨在展示 PIRNN框架在受控不确定性下的总体稳健性和有效性，因此我们没有进行细致的超参数搜索。系统的超参数优化或许可以改善结果，但不应影响结论。

超参数	**值**	**描述**
输入维度	6	状态变量（$\mathbf{x}$） + 控制输入（T）
输出维度	5	预测状态变量
隐藏状态大小	64	LSTM 隐藏状态中的特征数量
层数	2	堆叠的 LSTM 层数
丢弃率	0.2	应用于第一个长短期记忆网络层之后的丢弃概率
激活函数	Softplus	应用于最终输出以确保非负性
归一化	一维批归一化	应用于解码器之前的长短期记忆网络输出
初始噪声尺度 η	0.1	可学习噪声参数的初始值
转换阈值 δ	0.1	Huber损失的转换阈值
优化器	Adam	针对模型权重和 PBM 参数进行优化
基础学习率	1×10^{-3}	Adam 优化器的初始学习率
学习率调度器	余弦退火	在总训练轮次内将学习率衰减至 1×10^{-7}

表 S1：网络超参数和训练配置。

S2 信号噪声

随着训练数据规模的增加，PIRNN预测与无噪声参考值更加吻合，表明网络本身的泛化能力更强。模型在 μ_2、 μ_3 和 C 中的表现通常优于在 μ_0 和μ_1 中的表现，原因在于 PIRNN 在优化过程中并未“平等地”对待每个状态变量。尽管所有状态变量都归一化到 0 至 1 的范围内，但它们的导数仍相差若干数量级。因此，由于各状态变量的拟合难易程度和时间导数的尺度不同，有限的泛化模型可能会优先拟合相关性较强且相对容易拟合的状态变量；在本例中，这些变量是 μ_2、 μ_3和 C。除高噪声情形外，所有情况下重新评估的常微分方程曲线都一致且准确，这表明 PIRNN 估计的动力学参数具有很高的可靠性。在高噪声情形下，重新评估的 μ_0 与真实值参考之间出现偏差，表明提取出的动力学参数质量有所下降。偏差从 μ_0到 μ_1 再到 μ_2逐渐减小，表明常微分方程积分中可能存在误差累积。在高噪声情形下，尽管重新评估的常微分方程预测存在明显偏差，PIRNN预测仍然准确，凸显了混合建模可能具备双重用途这一优势。

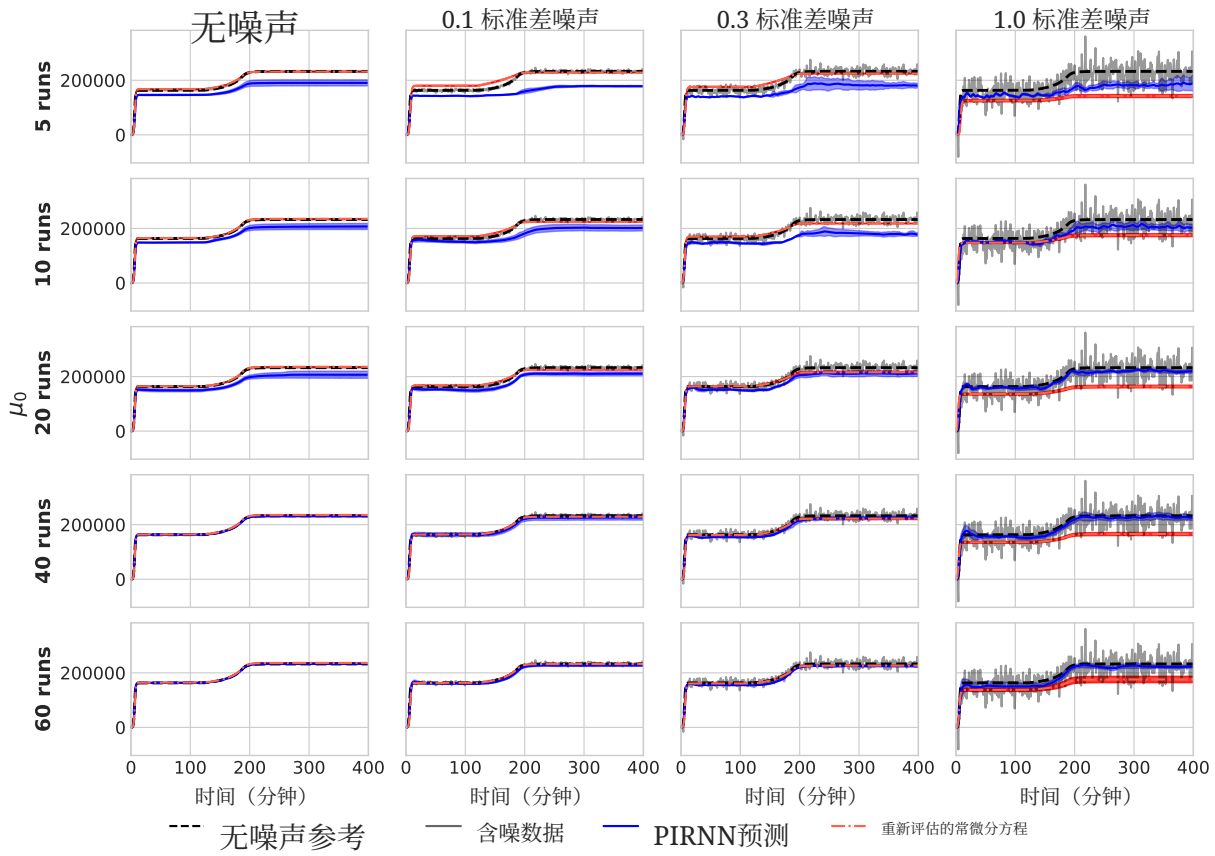


图 S10： μ_0 使用 5、10、20、40 和 60 次训练运行对一组测试数据的预测。图中给出了 95% 置信区间。每种情况下均显示原始含噪数据和无噪声真实参考。

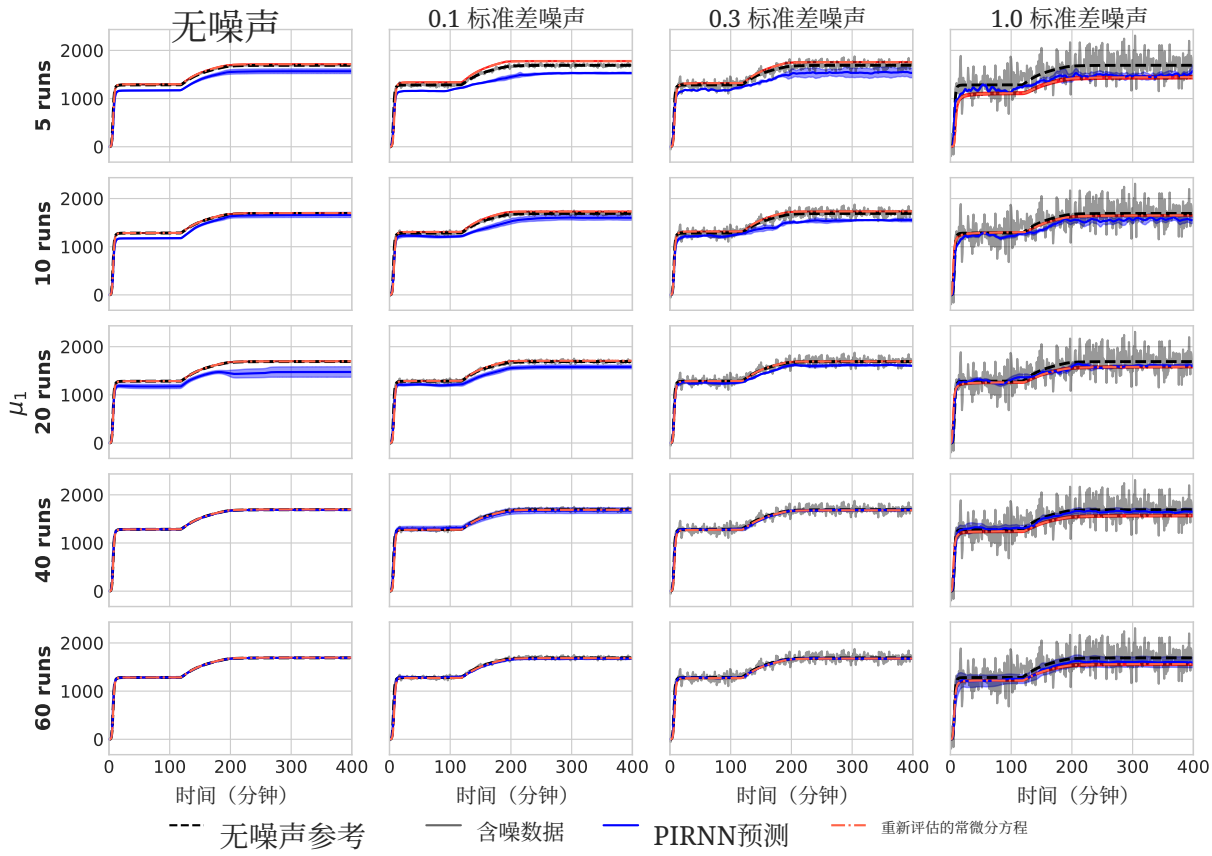


图 S11： μ_1 使用 5、10、20、40 和 60 次训练运行对一组测试数据的预测。图中给出了 95% 置信区间。每种情况下均显示原始含噪数据和无噪声真实参考。

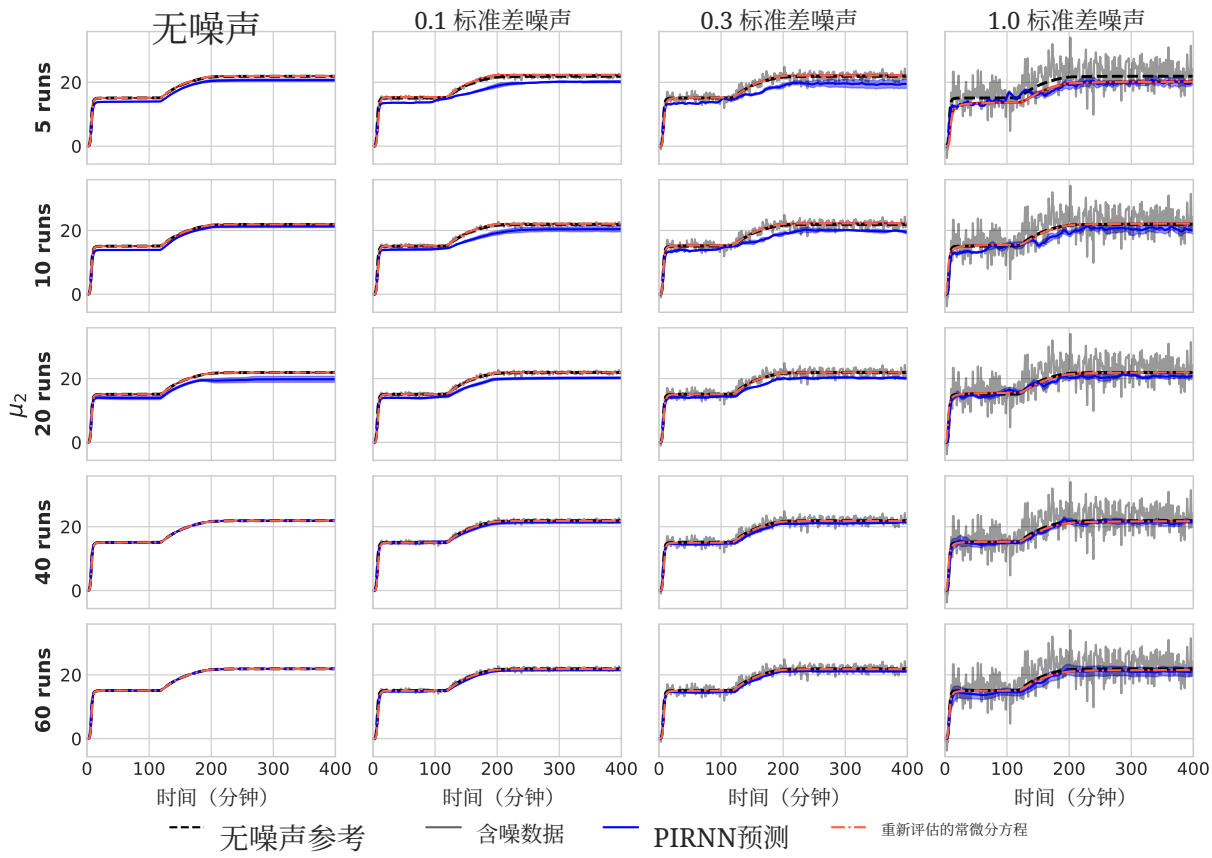


图 S12：使用 5、10、20、40 和 60 次训练运行次数对一组测试数据进行 μ_2 预测。图中给出了 95% 置信区间。每种情况下均展示了原始含噪数据和无噪声真实参考。

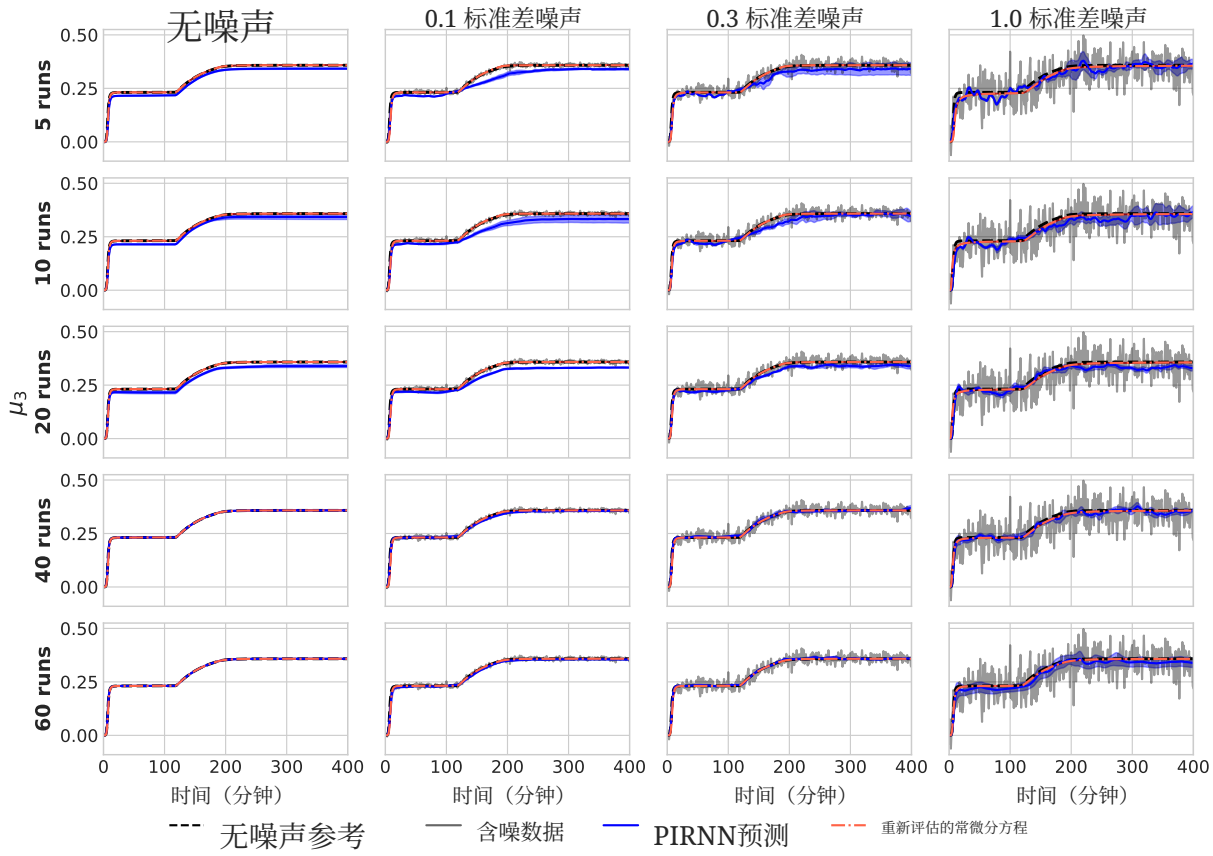
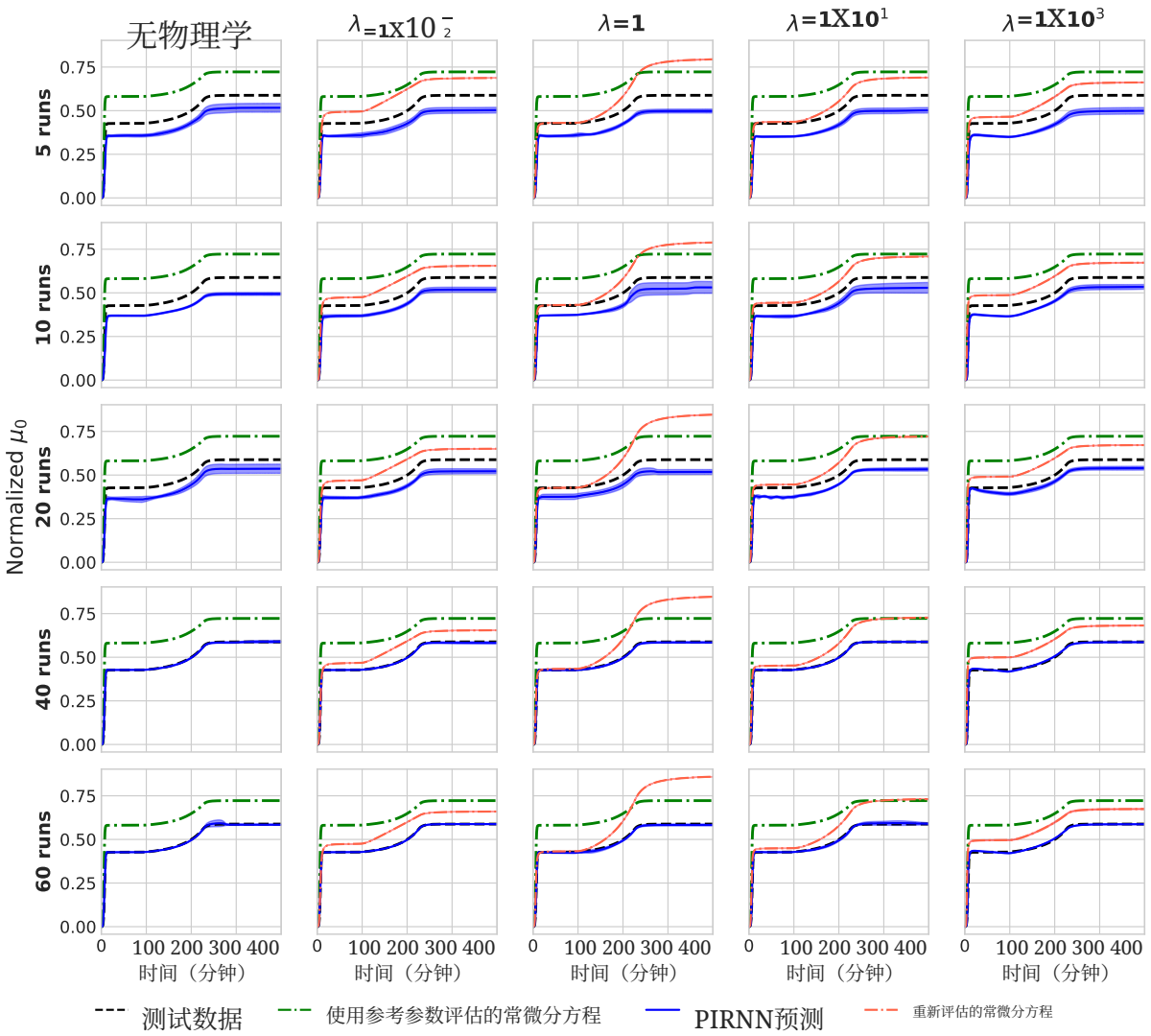


图 S13： μ_3 使用 5、10、20、40 和 60 次训练运行对一组测试数据的预测。提供了 95% 置信区间。每种情况下均显示原始含噪数据和无噪声真实参考。

S3溶解度偏移



图S14：在施加 10% 恒定溶解度偏移时，使用不同的 λ 条件下，通过 5、10、20、40 和 60 次训练运行对一组测试数据进行归一化 μ_0 预测。对于无物理学情形，由于未将 PBM 参数纳入优化（未包含物理损失），因此不显示重新评估的常微分方程曲线。常微分方程曲线使用两组不同的参数进行评估：参考值（绿色）和 PIRNN 优化值（红色）。图中显示了 95% 置信区间。

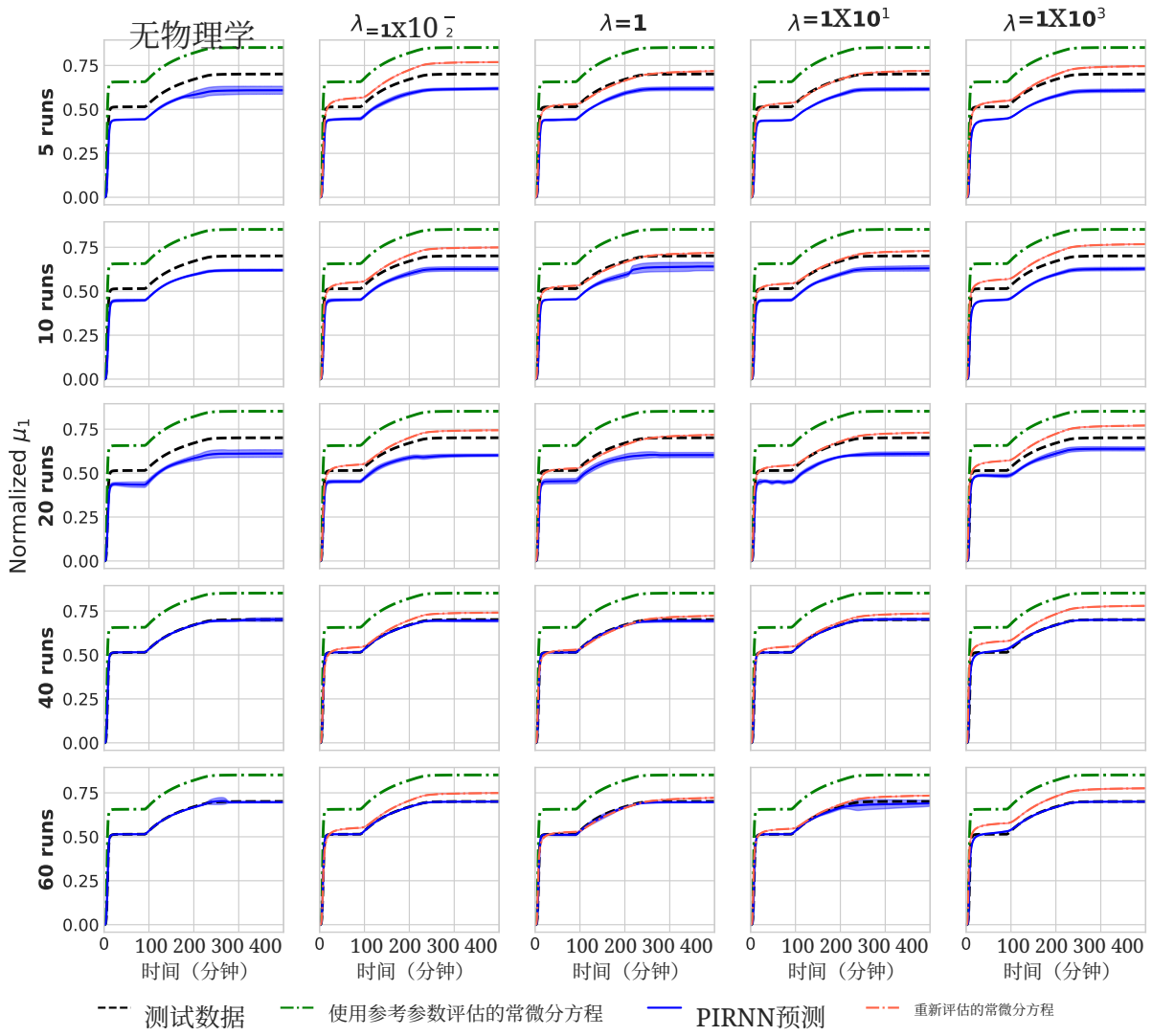


图 S15：施加 10% 恒定溶解度偏移时，在不同 λ 下，使用 5、10、20、40 和 60 次训练运行得到的测试数据集的归一化 μ_1 预测。对于无物理学情形，由于未将 PBM 参数纳入优化（未包含物理损失），因此不显示重新评估的常微分方程曲线。使用两组不同的参数评估常微分方程曲线：参考值（绿色）和 PIRNN 优化值（红色）。图中显示了 95% 置信区间。

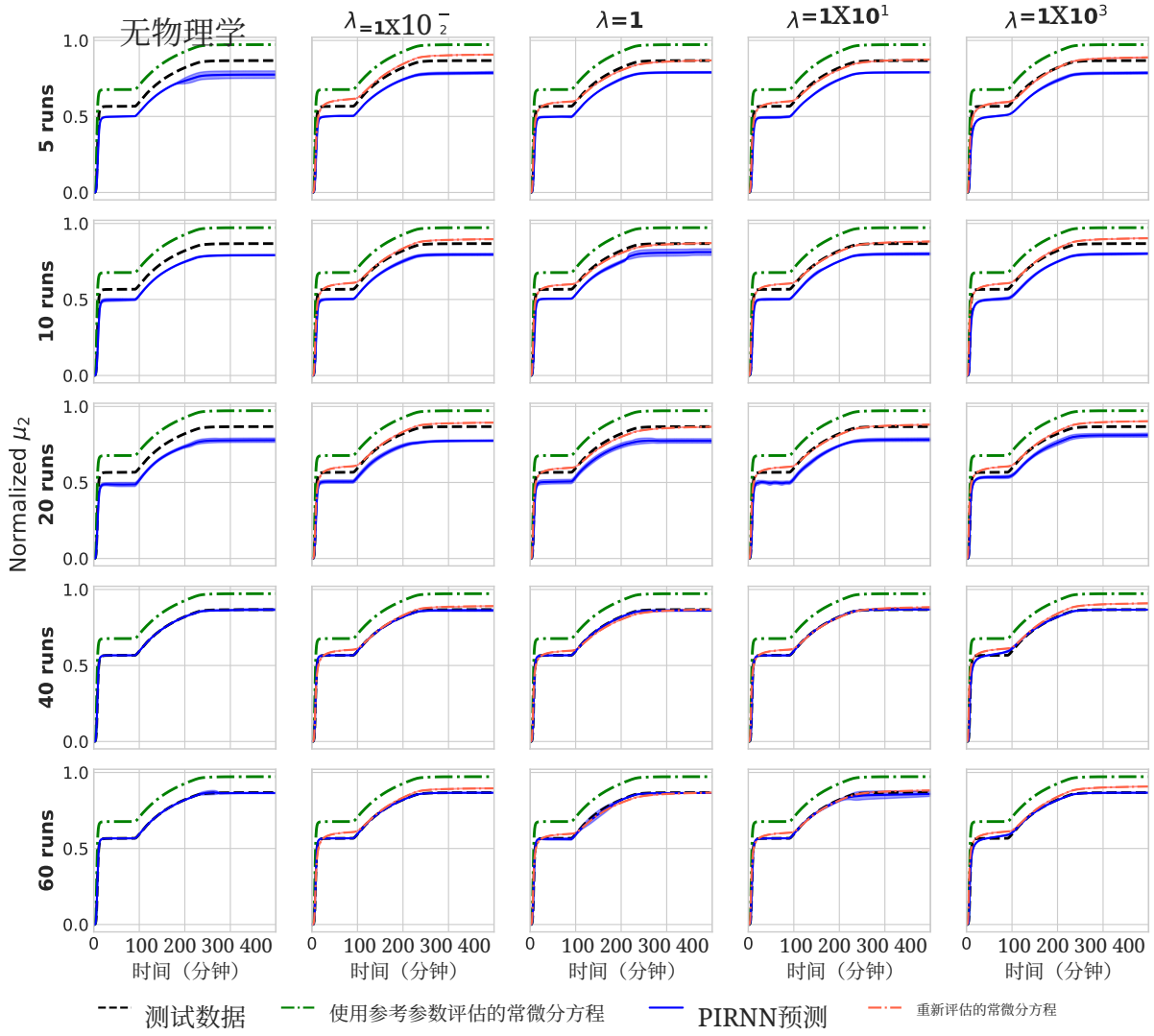


图 S16：在应用 10% 恒定溶解度偏移时，使用不同 λ 下的 5、10、20、40 和 60 次训练运行，对一组测试数据进行归一化 μ_2 预测。对于无物理学情形，由于未将 PBM 参数纳入优化（未包含物理损失），因此不显示重新评估的常微分方程曲线。使用两组不同的参数评估常微分方程曲线：参考值（绿色）和 PIRNN 优化值（红色）。图中显示了 95% 置信区间。

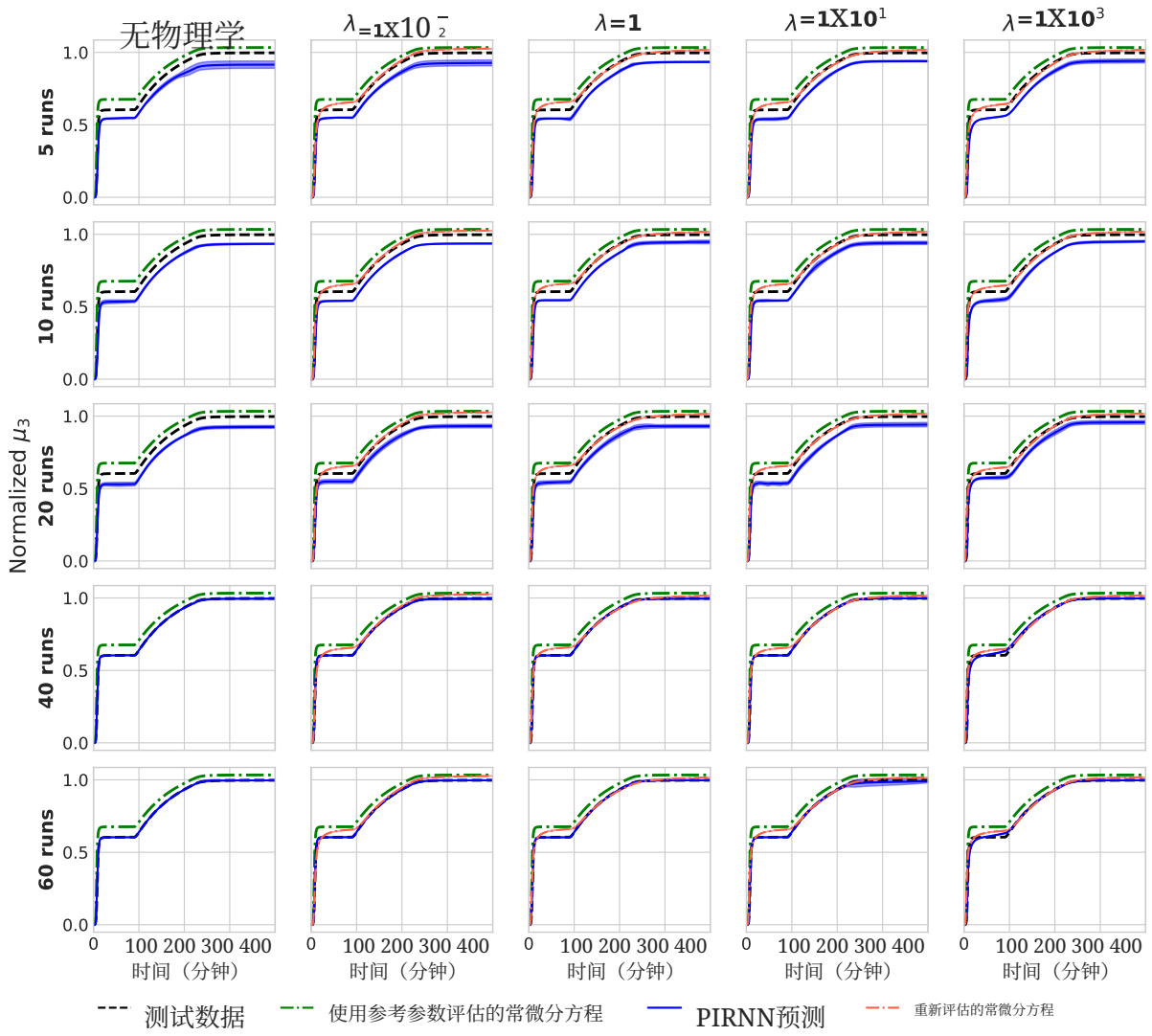


图 S17：应用 10% 恒定溶解度偏移时，在不同 λ 下使用 5、10、20、40 和 60 次训练运行对一组测试数据进行的归一化 μ_3 预测。无物理学情形未显示重新评估的常微分方程曲线，因为未将 PBM 参数纳入优化（未包含物理损失）。常微分方程曲线使用两组不同的参数进行评估：参考值（绿色）和 PIRNN 优化值（红色）。图中显示了 95% 置信区间。

S4 采样频率限制

有限差分估算的时间导数与使用学习得到的PBM参数计算的时间导数高度一致，这表明在此情况下模型损失已成功收敛。因此，学习得到的参数值与参考值之间的差异，可能是由于PBM参数之间高度相互相关性所导致的动力学等效参数集。

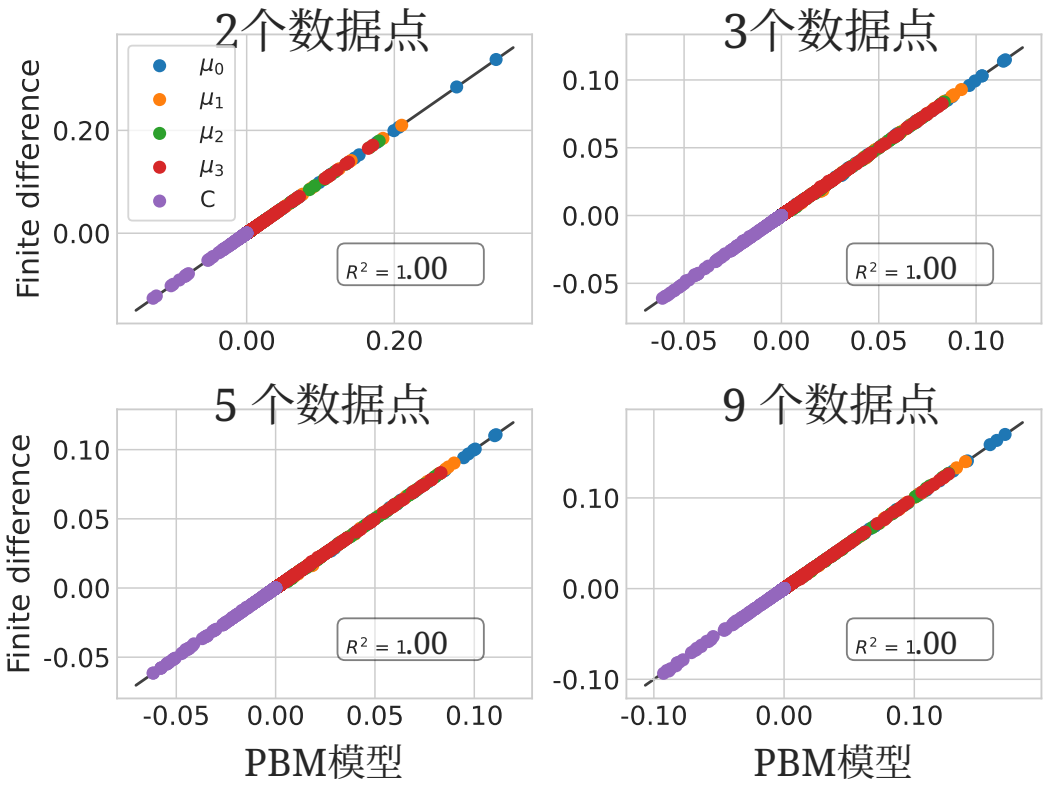
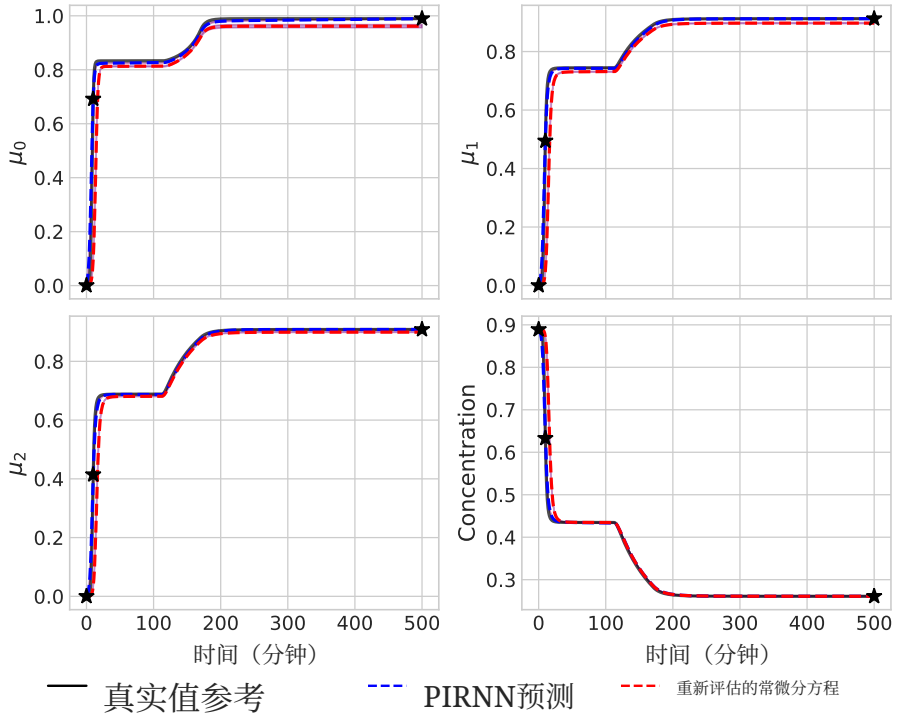
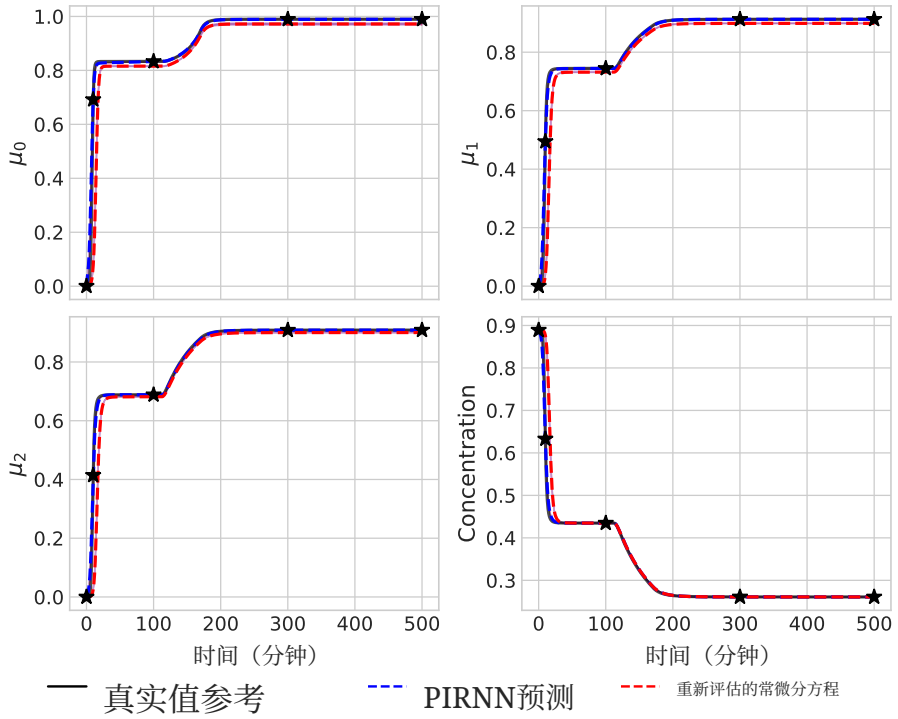


图 S18：有限差分估算的时间导数与使用表 2 中 PIRNN 反演得到的 PBM模型参数计算的时间导数之间的模型收敛情况。



图S19：使用 PIRNN 估计的 PBM 参数，针对归一化的 μ_0、μ_1、μ_2 和 C，在每次运行使用 3 个数据点时，PIRNN 预测的轨迹和评估得到的常微分方程轨迹。黑色实线表示真实参考值。



图S20：使用 PIRNN 估算的 PBM 参数时，PIRNN 对归一化 μ_0、μ_1、μ_2 和 C 的预测轨迹，以及根据这些参数评估得到的常微分方程轨迹；每次运行使用 5 个数据点。真实参考值以黑色实线表示。

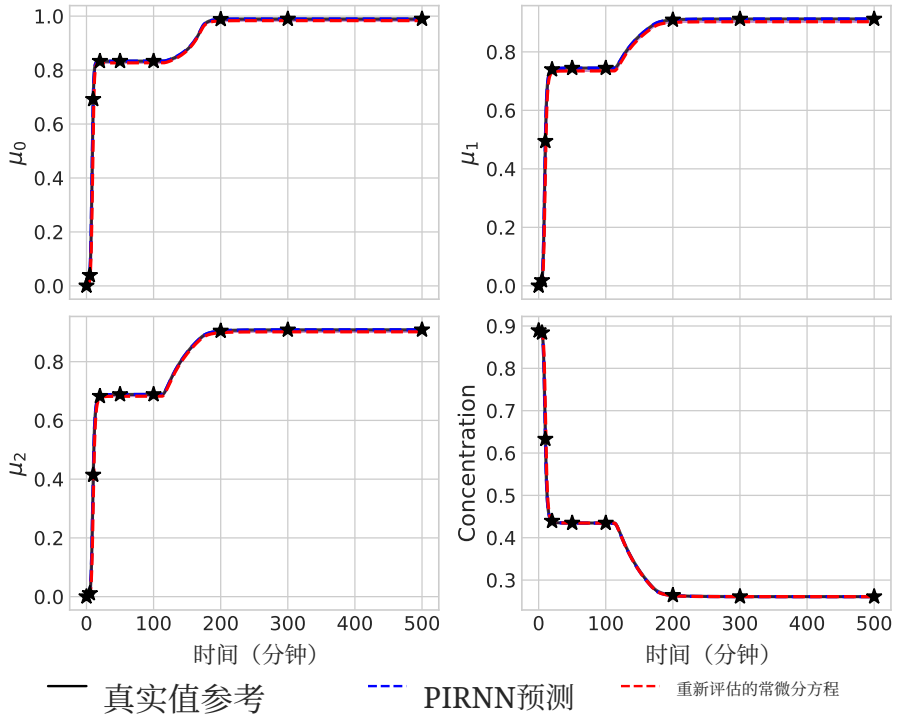


图 S21：使用 PIRNN 估计的 PBM 参数得到的 PIRNN 预测轨迹和计算得到的 ODE 轨迹，其中归一化的 μ_0、μ_1、μ_2 和 C 使用每次运行 9 个数据点。真实参考值以黑色实线表示。

S5 代码可用性