# 基于 BPINNs 的圣维南方程求解与不确定性量化

景象

2025 年 11 月 29 日

# PINN 基本原理

**核心思想**: 将物理定律（偏微分方程 PDE）作为正则化项加入神经网络的损失函数中，实现数据与物理双驱动。

- **数据驱动**: 拟合稀疏观测数据（边界、初始条件）。
- **物理驱动**: 最小化 PDE 残差，无需网格。
- **技术关键**: 自动微分 (Automatic Differentiation) 精确计算导数。

**通用近似**

$$\hat{u}(x,t) \approx \mathcal{NN}(x,t;\theta)$$

# 数学表述

**损失函数** $\mathcal{L}$ 由数据误差 $\mathcal{L}_{data}$ 和物理残差 $\mathcal{L}_{PDE}$ 组成:

$$\mathcal{L}(\theta) = w_{data}\mathcal{L}_{data} + w_{PDE}\mathcal{L}_{PDE}$$

- **数据损失** (观测值匹配):

$$\mathcal{L}_{data} = \frac{1}{N_d}\sum_{i=1}^{N_d} \|\hat{u}(x_i, t_i) - u^i_{obs}\|^2$$

- **物理残差** (控制方程 $f(u, u_t, u_x, \dots) = 0$):

$$\mathcal{L}_{PDE} = \frac{1}{N_f}\sum_{j=1}^{N_f} \|f(\hat{u}(x_j, t_j); \frac{\partial \hat{u}}{\partial t}, \nabla \hat{u})\|^2$$

# 面临的主要问题

1. **优化困难**
   - 损失函数不同损失项 ($\mathcal{L}_{data}$ vs $\mathcal{L}_{PDE}$) 梯度量级差异大，难以平衡。需要仔细调参才可以较好训练。
2. **谱偏置**
   - 深度神经网络倾向于优先学习低频函数 (F-Principle)。
   - 难以捕捉高频振荡、剧烈变化的波前或多尺度特征。
3. **因果性缺失**
   - 时间演化问题中，标准 PINN 缺乏时间因果性，导致解在整个时空域同时收敛，而非按时间推进。

# 常用改进方法

针对上述问题，学术界提出了多种改进方案：

| 改进方向 | 代表性方法 |
|---|---|
| 特征工程 | **傅里叶特征嵌入 (Fourier Features)** (解决谱偏置)<br>自适应激活函数 (Adaptive Activation Functions) |
| 损失加权 | 自适应权重 (Gradient Normalization, SoftAdapt)<br>神经正切核 (Neural Tangent Kernel, NTK) 加权 |
| 训练策略 | 课程学习 (Curriculum Learning)<br>因果训练 (Causal Training), 序列到序列学习 |
| 模型架构 | 区域分解 (XPINNs), 算子学习 (DeepONet, FNO) |

# 项目基本情况

1. **目标**: 发展基于 BPINNs 的圣维南方程求解与不确定性量化方法，实现利用有限观测数据对中线工程全线水情进行模拟
2. **直接采用监测数据训练难度大**: 利用水动力模拟结果开展训练
3. **闸门处 loss 较大**: 仅对渠道进行训练（事后证明渠道模拟只是 rmse 小，其实渠道才是真正谬误所在）
4. **本次汇报内容**: 最简化圣维南方程，仅拟合矩形河道单一工况加噪声的模拟结果，最后较为成功

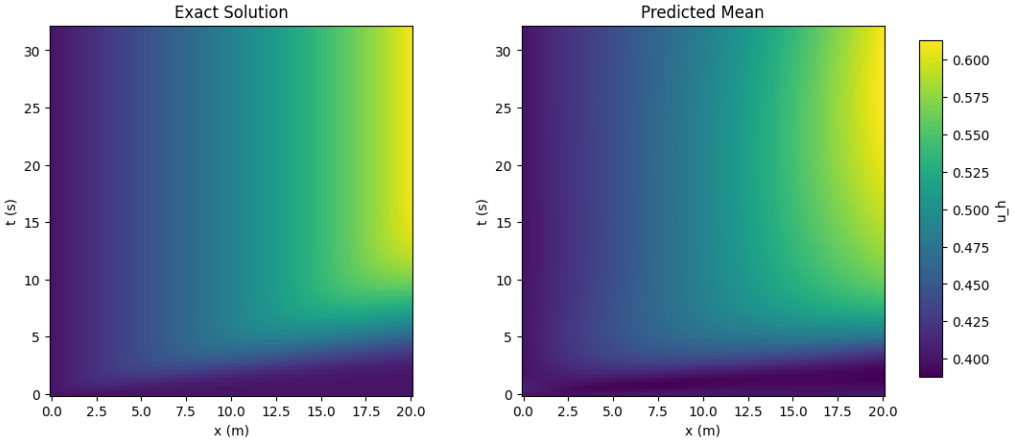
# 项目演进路线图

1. **阶段一：原型验证** - 小规模问题，可行性验证。
2. **阶段二：理想基准** - 大规模问题，无噪声验证。
3. **阶段三：噪声挑战** - 引入噪声，暴露双曲型方程弱点。
4. **阶段四：频谱陷阱** - 稳定边界数据后，揭示低频偏差问题。
5. **阶段五：最终方案** - 人工粘性 + 频域特征。

实际实施的顺序是：$3 \to 4 \to 1 \to 5 \to 2$。

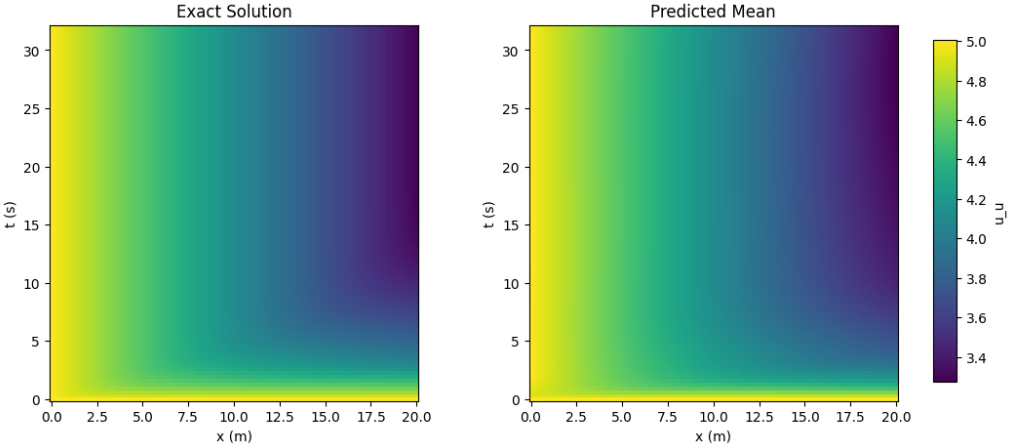
# 阶段一：Demo 预测结果 - 水深 (h)

**小规模验证**：矩形渠道平底，长 20m，模拟 32s，边界条件不变
FDM(左) vs PINN (右)



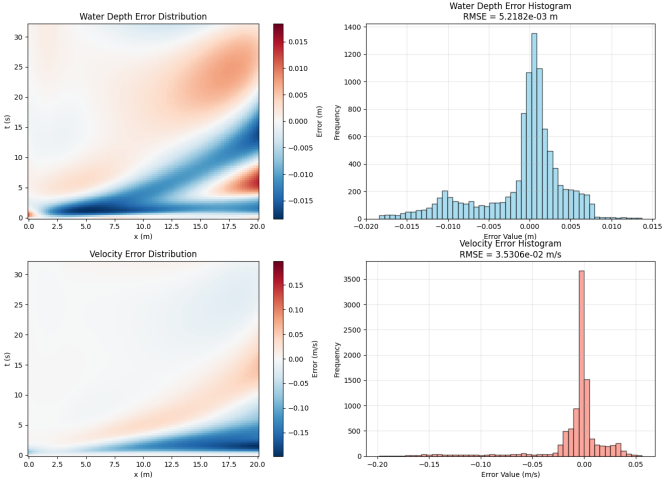
# 阶段一：Demo 预测结果 - 流速 (u)

**小规模验证**：流速场捕捉准确。



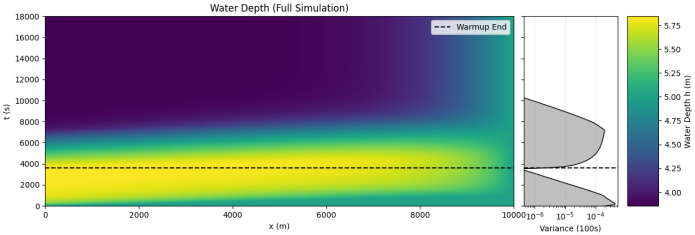
# 阶段一：Demo 误差分析

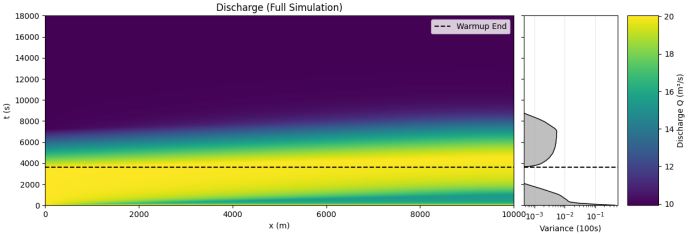
误差分布均匀且量级极低，水深 $h$ 误差约为 5.21e-3 m，流速 $u$ 误差约为 3.53e-2 m/s。



# 大规模工况定义

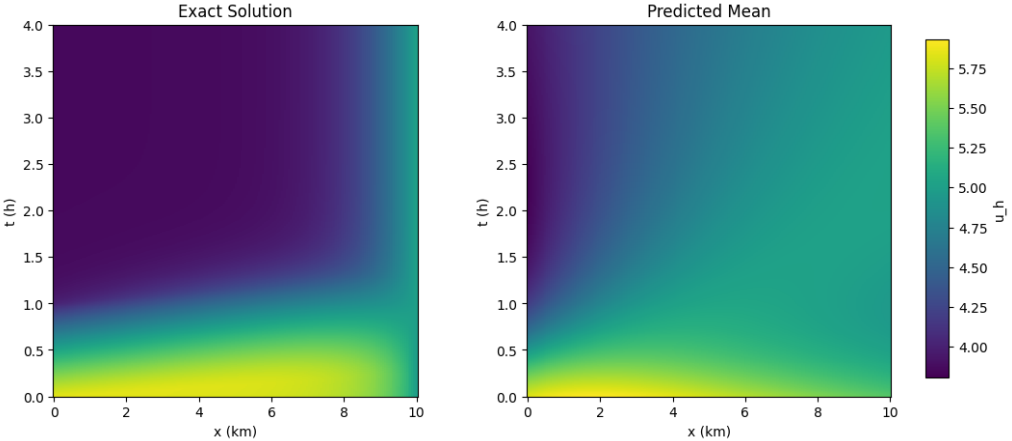
**设置**: 长度 10km，时长 4h，采用上游流量、下游水位边界。正式模拟前预热 1h。模拟过程中前一小时流量 $Q_{in}$ 线性减半。



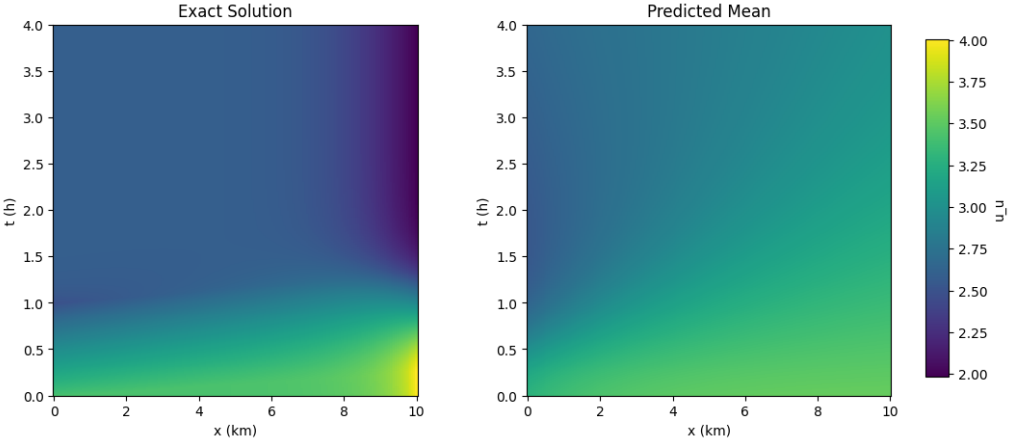
# 阶段二：无噪声基准 - 水深 (h)

**目的**: 验证网络容量。预测值 (右) 与真值 (左) 完美重合。



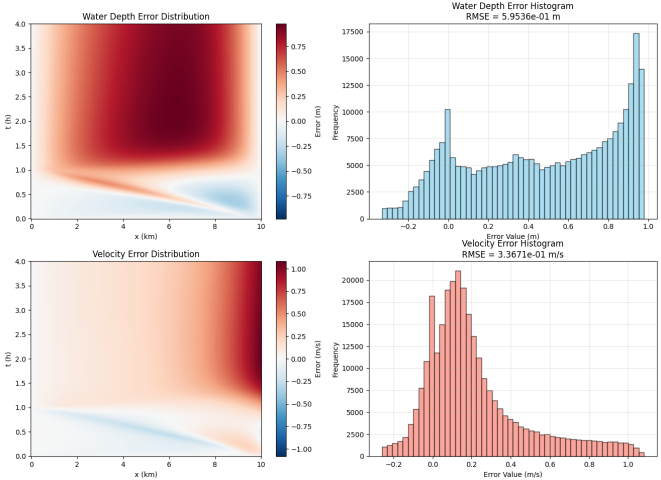
# 阶段二：无噪声基准 - 流速 (u)

流速预测同样精准，无伪影。



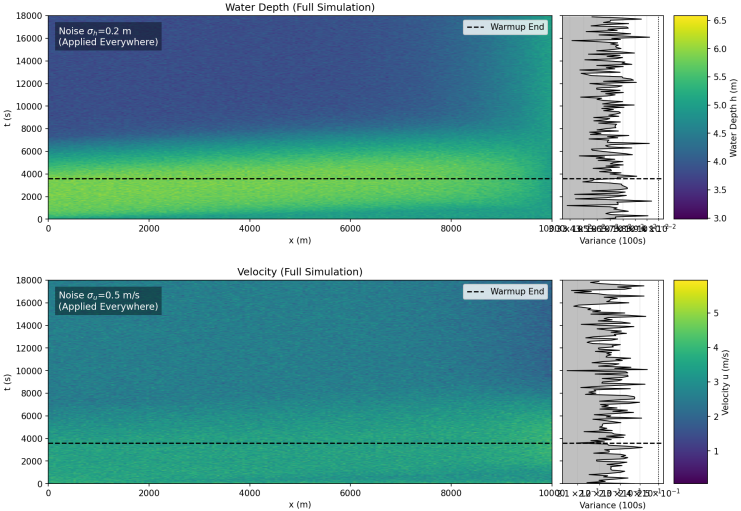
# 阶段二：Demo 误差分析

上游误差沿特征线向下游传播，上游 1h 后的稳定边界特征完全丢失



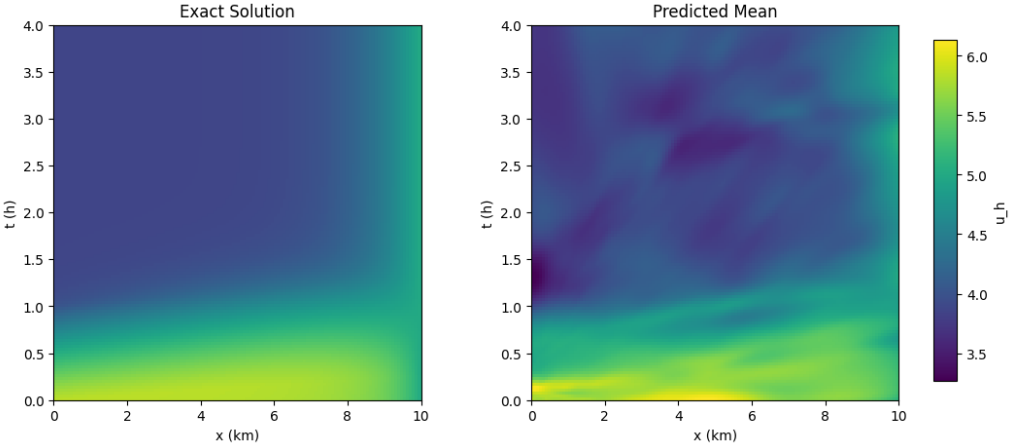
# 阶段三：工况设置

**设置**：在阶段二的基础上，向观测数据添加噪声，模拟实际测量误差对模型的影响。



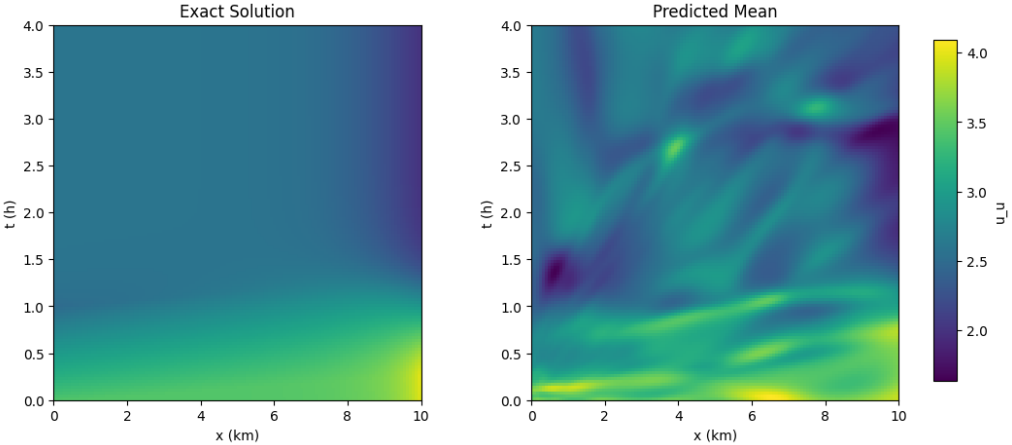
# 阶段三：全域水深模拟结果 (h)

加入 5% 高斯噪声后，水深预测出现严重伪影。



# 阶段三：全域流速模拟结果 (u)

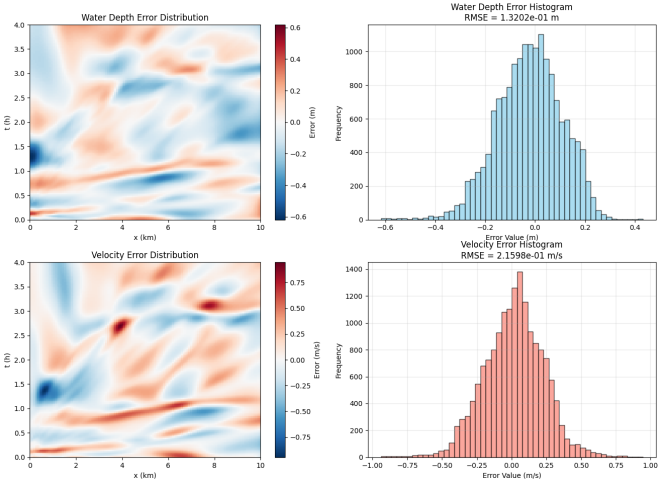
流速场也出现剧烈震荡，无法使用。



# 阶段三：误差传播机制分析

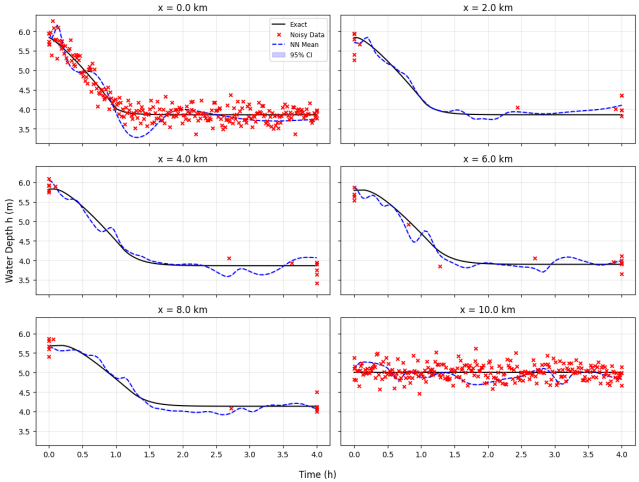
误差连片聚集分布，沿特征线扩散。

1. sol(内部误差) = 0.2
2. par(先验误差) = 0.0
3. bnd(边界误差) = 2.0
4. pde(物理误差) = 0.005
5. noise_h_std_phys (水深噪声标准差) = 0.2
6. noise_Q_std_phys (流量噪声标准差) = 0.5
7. weight_std (权重标准差) = 0.5



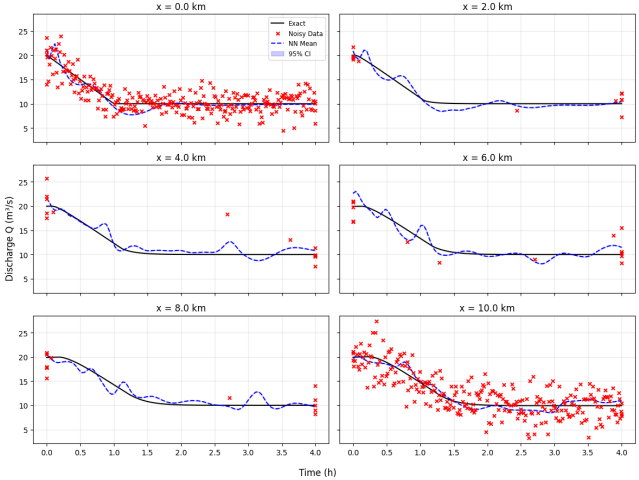
# 阶段三：时序分析 (水深 h)

预测值剧烈震荡，物理守恒性丧失。
黑色：FEM 解。红色：带误差的模型训练输入。蓝色：BPINN 预测值。



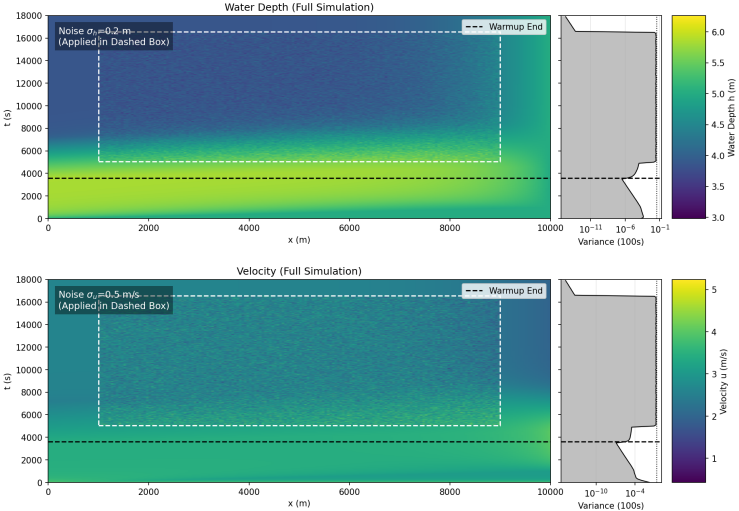
# 阶段三：时序分析 (流量 Q)

预测值剧烈震荡，物理守恒性丧失。



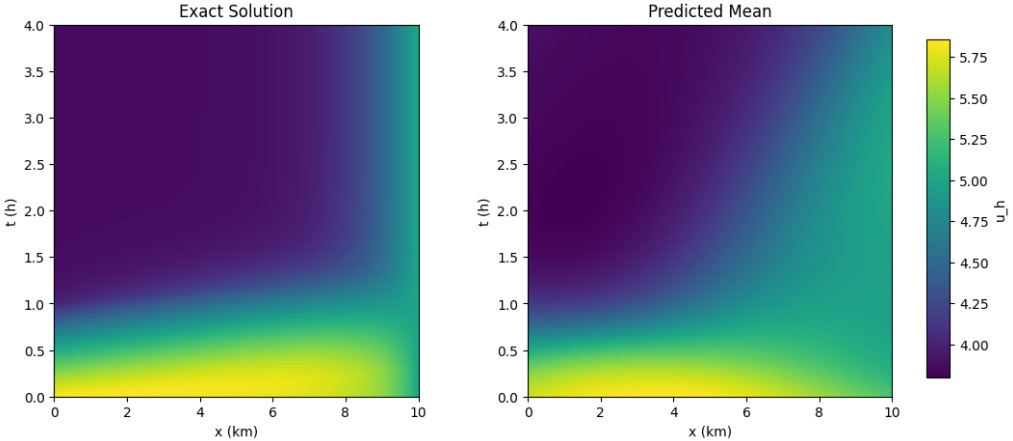
# 阶段四：工况设置

**设置**: 仅在内部观测数据添加噪声，模拟模型理论边界 + 实测误差对模型的影响。



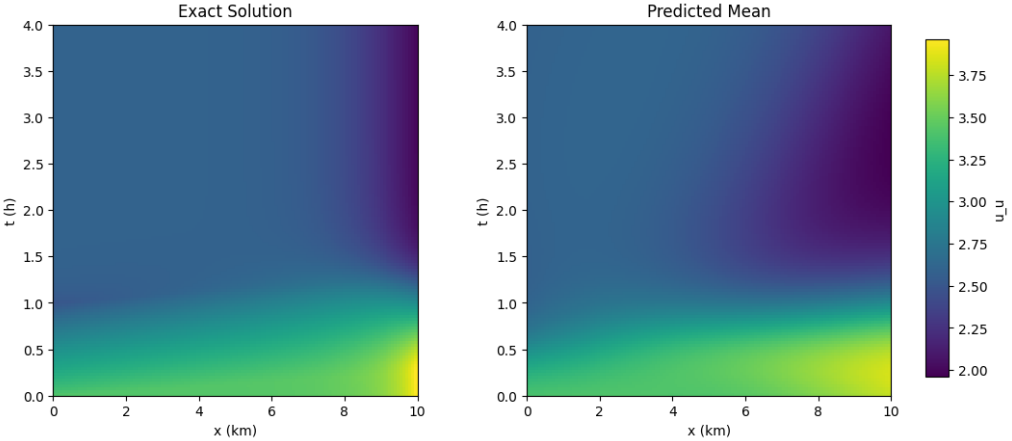
# 阶段四：频谱陷阱 - 水深 (h)

结果过于平滑，丢失波峰。上游流量边界在 (x=0m,t=3600s) 时流量的一阶导数不连续，网络未能捕捉到该变化。



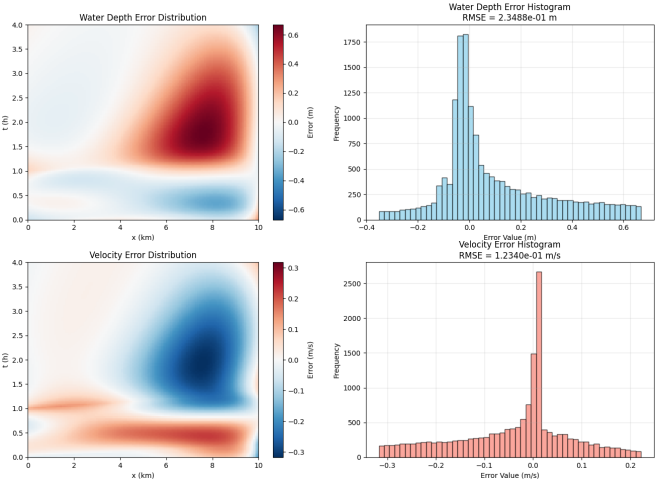
# 阶段四：频谱陷阱 - 流速 (u)

流速同样平滑化，未能捕捉锐利变化。



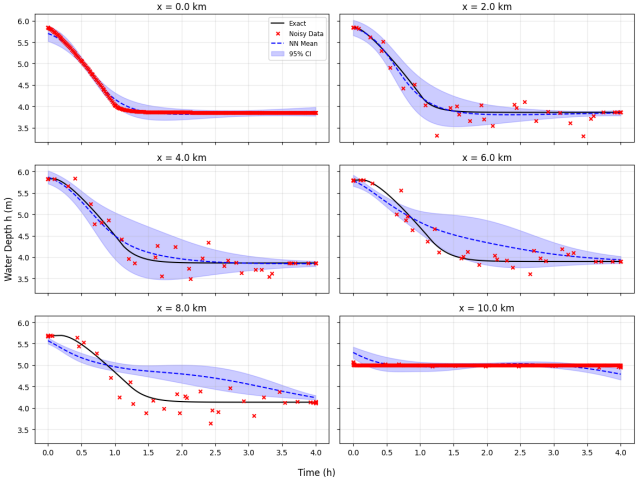
# 阶段四：系统性偏差分析

低频偏差 (Spectral Bias) 导致结构性误差。



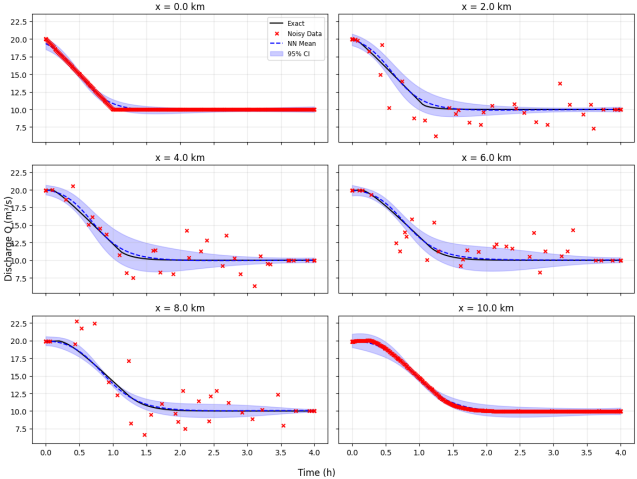
# 阶段四：时序分析 (水深 h)

预测均值 (蓝虚线) 明显偏离真值。



# 阶段四：时序分析 (流量 Q)

波前捕捉失败。



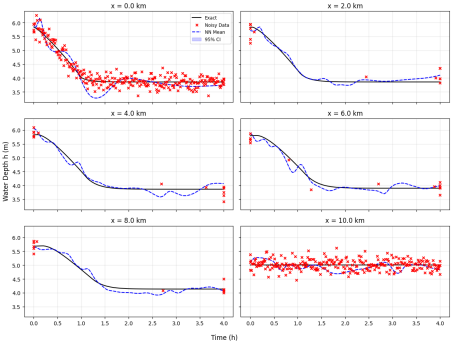
# 改进一：人工粘性 (Artificial Viscosity)

**控制方程修正**

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = \mathbf{S} + \nu \frac{\partial^2 \mathbf{U}}{\partial x^2}$$

为何引入？

- **物理层**：模拟流体微团内摩擦，**抹平激波间断。**
- **优化层**：平滑 Loss Landscape，防止 **HMC 采样发散。**




图：目标：消除此类数值震荡

# 改进二：傅里叶特征嵌入 (Fourier Features)

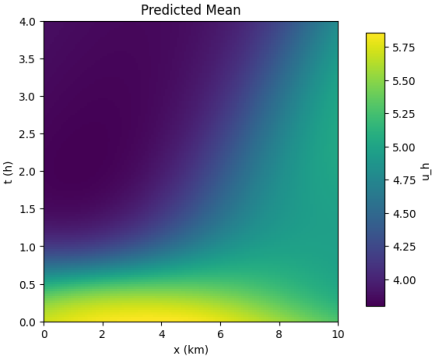
**坐标映射 (放大变化)**

$$\gamma(\mathbf{x}) = \begin{bmatrix} \cos(2\pi\mathbf{B}\mathbf{x}) \\ \sin(2\pi\mathbf{B}\mathbf{x}) \end{bmatrix}$$

$$\mathbf{B} \sim \mathcal{N}(0, \sigma^2)$$

**核心机制**

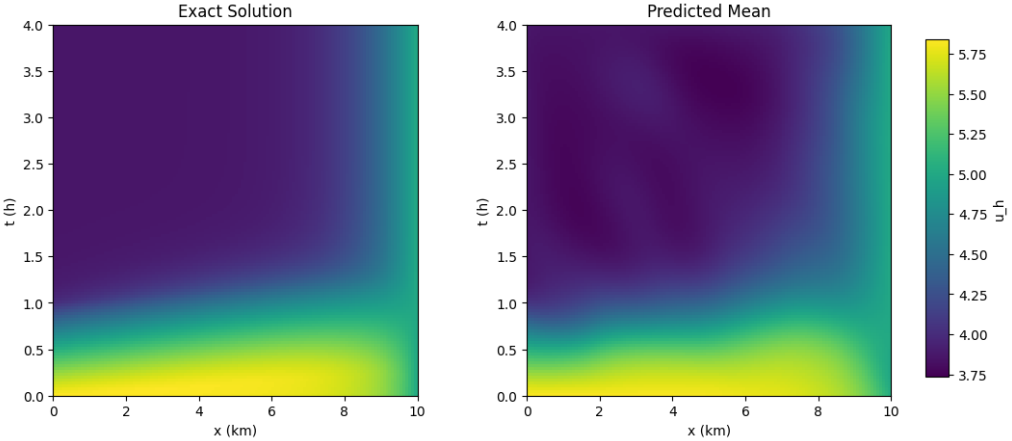
- **问题**：神经网络 tanh 具有低通滤波器的特性，导致神经网络出现了**谱偏置**问题，偏爱低频信息。
- **对策**：映射至高维频率空间，强迫网络关注高频细节。




图：目标：修正过度平滑 (丢失波前)

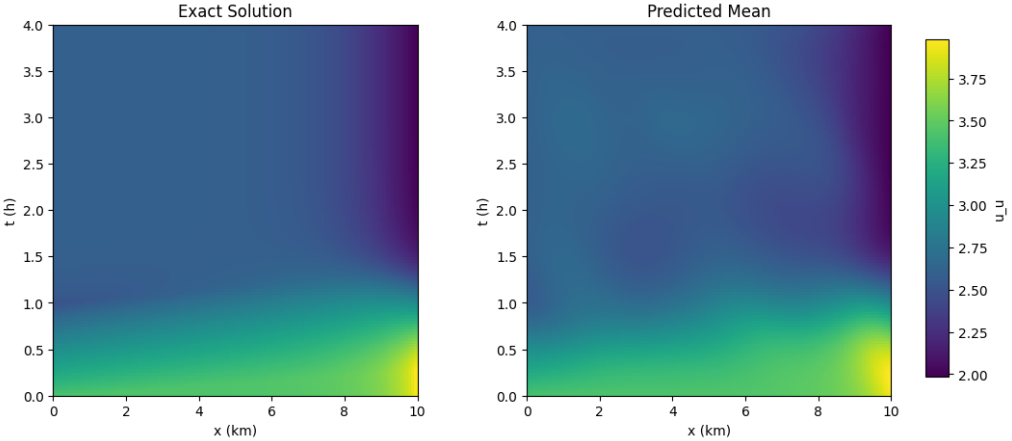
# 阶段五结果：精准预测 - 水深 (h)

完美重建流场，伪影消失，强噪声鲁棒性。



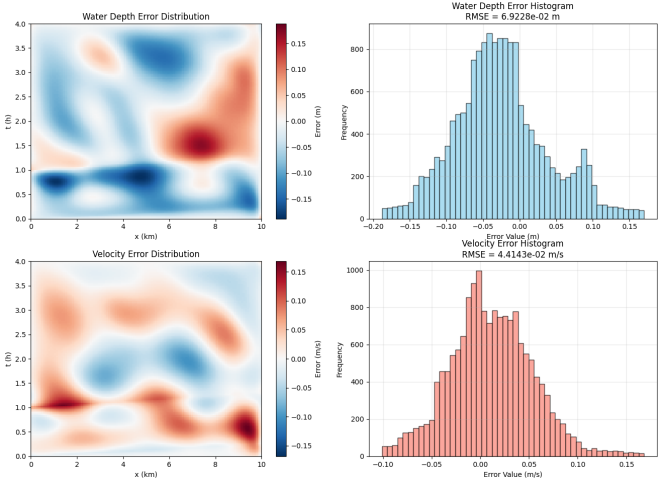
# 阶段五结果：精准预测 - 流速 (u)

流速场恢复清晰，细节保留完整。



# 阶段五结果：全域误差分析

全域误差显著降低，无明显结构性偏差。

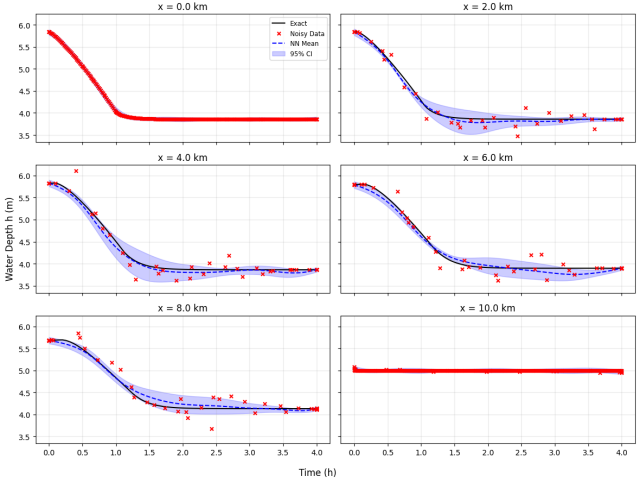


# 阶段五结果：训练收敛性 (Loss)

Loss 曲线平稳下降，无过拟合或震荡迹象。

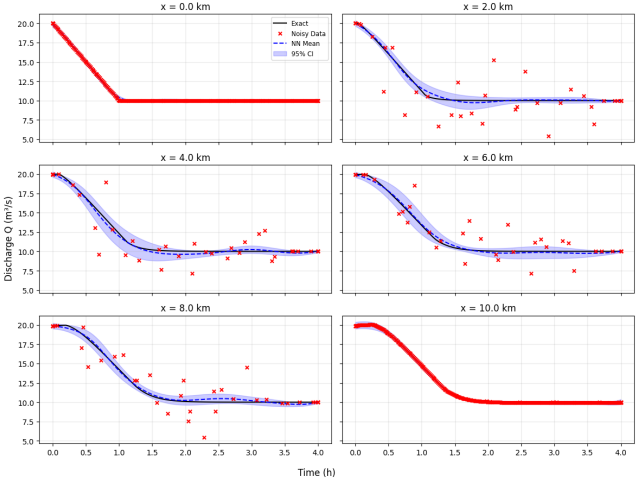
# 阶段五：不确定性量化 (水深 h)

**95% CI**: 蓝色阴影覆盖良好，不确定性随时间合理扩散。



# 阶段五：不确定性量化 (流量 Q)

准确捕捉流量减半动态，物理守恒性得到保持。



# 总结

- **当前结论**：BPINN 可以通过人工粘性、傅里叶特征嵌入等技术，有效应对噪声和频谱偏置问题，实现对含噪声简单流场的预测。
- **存在问题**：对于复杂流场（如渐变段、多段渠道联合），当前方法的可行性未经验证。
- **存在问题**：目前的实现还是对流场的复现，也即 $f(x,t)=[u,h]$，但我们事实上希望的是一个预测模型，即 $f(x,t|BC,IC)=[u,h]$，这需要进一步的研究工作。

谢谢聆听！