

 <https://doi.org/10.71573/4p6bcf78>

© Authors. This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

# HR-PIGNN A High Resolution Prediction Method for Urban Drainage Network: Combining Graph Neural Networks and Discrete Form Physics Informed Neural Networks

Yizhou Qian <sup>1</sup>  <https://orcid.org/0009-0004-6783-1703>, Xin Dong <sup>1,\*</sup>

<sup>1</sup> School of Environment, Tsinghua University, 100084, Beijing, China

\*Corresponding author email: [donqixin@tsinghua.edu.cn](mailto:donqixin@tsinghua.edu.cn)

## Abstract

This article presents a novel hybrid model combining data and mechanisms for high-resolution water level prediction in pipeline networks. The model utilizes graph convolutional neural networks to integrate network topology information for precise predictions and incorporates de Saint-Venant system equations through physics-driven neural networks. Compared to traditional data-driven, mechanism-driven, and hybrid methods, this model achieves 5-minute, 1-centimeter resolution predictions while maintaining computational efficiency and high accuracy. In an experimental drainage system in Suzhou, China, the model's RMSE for predicting water levels and pipeline flow rates is 0.014 and 0.012, respectively, with NSE values of 0.802 and 0.883. The model's computation time for 24 hours of data at 5-minute intervals is 0.981 seconds.

## Highlights

- Developed hybrid model (GNN+PINN) for accurate urban drainage network simulations.
- Incorporating de Saint-Venant equations via difference method enhances model performance.

## Introduction

Recent advancements in water level simulation have led to the exploration of data-driven methods for predicting urban drainage systems. However, these methods often suffer from over-reliance on samples, limiting their interpretability and generalization. With the introduction of physics-informed neural networks (PINNs) (Raissi et al. 2017), researchers have begun integrating urban water system equations into traditional data-driven approaches to enhance model accuracy, interpretability, and generalization (Ye et al., 2022). Nevertheless, current hybrid models frequently focus on incorporating mass conservation constraints or using graph neurophysics to strengthen topological coupling (Zhang et al., 2024). While PINNs using complete hydraulic equations have been applied mainly in river and water conservancy simulations (Li et al., 2024), this article proposes a discretized PINN integrated with graph convolutional networks for drainage networks, incorporating the full de Saint-Venant system as a physical constraint for reliable and high-resolution water level predictions.

## Methodology

### Model formulation

The hybrid model proposed in this study consists of three main components: a graph neural network coupled with topological relationships, a convolutional neural network generating high-resolution

results, and a discretized differential loss calculator used to calculate the complete de Saint-Venant system of equations.

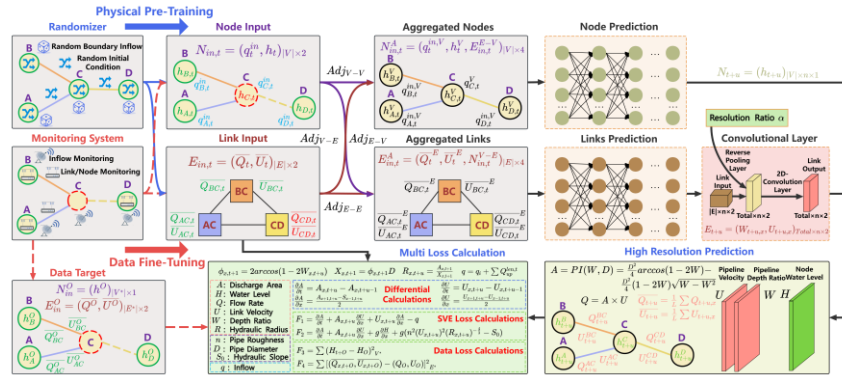


Figure 1. High resolution hybrid driven model HD-PIGNN model framework

In the hybrid model (Figure 1), the simulation begins by importing boundary conditions into the graph neural network (GNN). The GNN aggregates relevant node water levels, pipeline flow velocities, and boundary flow data based on the network’s topological structure to make preliminary predictions. Concurrently, a convolutional neural network (CNN) is used to upsample the pipeline flow and velocity predictions, generating high-resolution results with a 1 cm precision. Next, the high-resolution matrices for pipeline velocity, flow, and node water levels are fed into the discretized differential loss calculator. The de Saint-Venant equation system loss is discretized by evaluating the filling degree based on flow rate, basic information, and spatiotemporal variations in flow rate, filling degree, and water level. The monitoring data loss is balanced with the physical constraint loss, and the final model loss is computed for training and iterative updating of model parameters.

### Model performance test

Mathematical experiments were conducted to evaluate the performance of the hybrid model. In these experiments, the HD-PIGNN model with complete physical constraints and high-resolution predictions, the GNN model without physical constraints, the PINN model incorporating mass conservation constraints, and the traditional SWMM model were independently optimized for hyperparameters using the Bayesian Optimization framework on the same computing platform. The models simulated the UDS under identical inflow conditions. Performance was assessed based on simulation time (T), root mean square error (RMSE) of water levels and flow rates, and Nash-Sutcliffe efficiency (NSE), with T indicating efficiency and RMSE/NSE reflecting accuracy.

## Case study

### Case study area

The case study is located in Suzhou, Jiangsu Province, China, focusing on the urban water system simulation station in Kunshan. The underground drainage system consists of 19 pipes, totaling 280 meters in length. Three parallel variable frequency pumps introduce municipal sewage from the research station’s periphery into inspection wells, simulating varying inflow conditions. A regulating tank is installed at the end to balance flow. Figure 2 provides an overview.

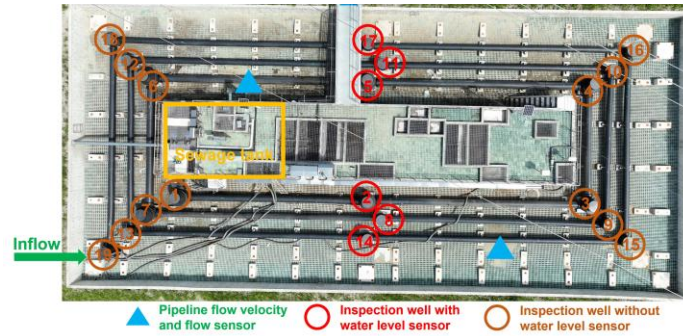


Figure 2. Overview of study area.

### Datasets for calibration, training and testing

The experimental platform for the urban drainage pipe network connects nodes 19 through 1 sequentially. Water level sensors are installed at inspection wells 2, 5, 8, 11, 14, and 17, while flow and velocity sensors are located in pipelines 6-5 and 15-14. Sewage is pumped into node 19 at hourly intervals using three variable frequency pumps. Sensor data, recorded every 5 minutes, represents the actual conditions of the urban drainage system (UDS). This data forms a dataset for model calibration and training.

The dataset consists of 4124 data points collected at 5-minute intervals from 00:30 on May 25, 2024, to 08:10 on June 7, 2024. Of these, 3163 points from May 25 to June 9 are used for model training, while the remainder serves as the testing set for performance evaluation. Both training and testing sets are utilized to perform Bayesian optimization for PIGNN, GNN, and PINN models using the HyperOpt framework, and to calibrate the physical parameters of SWMM models through a genetic algorithm (GA).

## Results and discussion

### The optimal hyperparameters

Using Bayesian optimization algorithm to iterate 100 cycles to determine the optimal hyperparameters for each neural network model, as shown in Table 1 below.

**Table 1.** The optimal hyperparameters for each model (where Edge-hop and Node-hop refer to the k nearest neighbour pipelines and the number of nodes connected to the model, Node-Edge and Edge-Node refer to the number of hidden neurons in each layer of the model's aggregated k nearest neighbour nodes and pipelines, and Node and Edge refer to the number of hidden neurons in each layer of the model's predicted node and pipeline results).

	Edge-hop	Node-hop	Node-Edge	Edge-Node	Node	Edge	Learning-rate	Weight-decay
PIGNN	3	3	[32,64]	[128,64]	[8,16,32]	[32]	1e-4	1e-5
GNN	2	4	[32,64]	[128,64,32]	[8,32]	[64,128]	1e-3	1e-7
PINN	-	-	-	-	[64,64,32]	-	1e-3	1e-5

### Node water level and pipeline flow prediction results

Table 2 presents the monitoring data from water level sensors in 5 inspection wells and flow sensors in 2 pipelines for the inflow process nodes between June 9th and 12th, along with the predicted Nash Sutcliffe efficiency (NSE), root mean square error (RMSE), and mean absolute error (MAE) for various models. The simulation error and NSE are shown in Figure 3, while the predicted water level at inspection well 5 and the flow rate in pipeline 6-5 are shown in Figure 4.

For the water level prediction in the test set, the high-resolution hybrid model PIGNN achieved an average NSE of 0.802. This represents improvements of 0.124, 0.108, and 0.133 over individual data driven model GNN, mechanism model SWMM, and traditional PINN models, respectively. Additionally, the average RMSE for water level prediction in PIGNN was 0.014, which is a reduction of 9.09%,

12.50%, and 21.84% compared to the individual data models, mechanism models, and traditional PINN models, respectively.

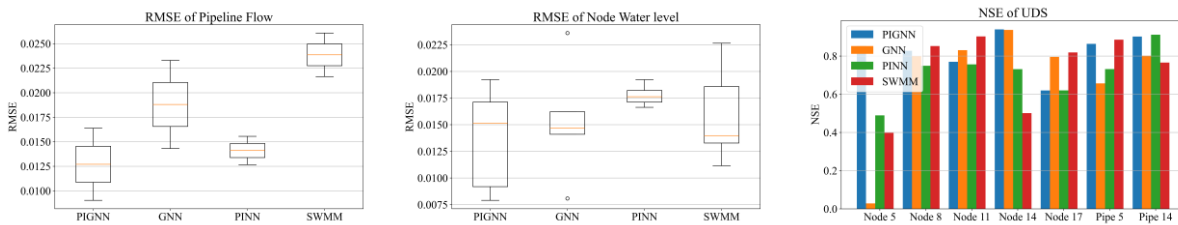
In flow prediction, PIGNN demonstrated even greater improvements over the traditional PINN model. The traditional PINN, relying solely on node mass conservation, cannot capture the physical dynamics within pipelines, leading to less accurate flow predictions. By incorporating the complete de Saint-Venant system of equations constraint through the discretization difference method, PIGNN integrates the physical laws governing the pipeline system. The mean NSE for flow prediction was 0.883, approaching the performance of the SWMM mechanism model. Furthermore, the average RMSE for flow prediction in PIGNN was 0.012, which is a reduction of 35.14%, 50.00%, and 17.24% compared to individual data models, mechanism models, and traditional PINN models, respectively.

Overall, the lowest NSE and highest RMSE for the PIGNN model were 0.620 and 0.017, respectively. When compared to the standalone data driven model GNN, mechanism model SWMM, and traditional PINN model, the lowest NSE increased by 0.592, 0.222, and 0.131, respectively, and the highest RMSE decreased by 29.17%, 34.62%, and 10.53%, respectively. These results further demonstrate the strong generalization ability of the high-resolution hybrid model PIGNN.

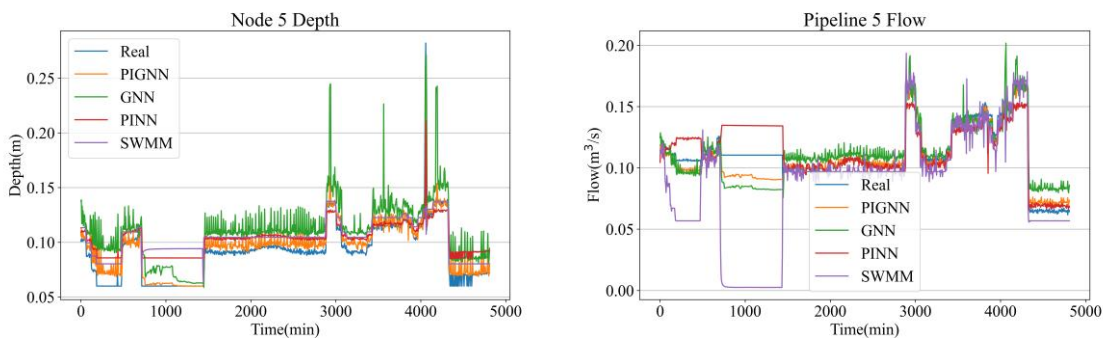
The overall computation time for the PIGNN model on the test set is 0.981 seconds, which is 95.43% shorter than the traditional SWMM mechanism model, which requires 21.469 seconds. This demonstrates that the high-resolution PIGNN model, driven by hybrid technology, has significant potential to enhance the simulation efficiency of urban drainage systems.

**Table 2.** The average NSE, RMSE, and MAE of four types of models at five inspection wells and two pipelines (in a single cell, the first value is NSE, the second value is RMSE).

	Node 5	Node 8	Node 11	Node 14	Node 17	Pipe 6-5	Pipe 15-14
PIGNN	0.853, 0.009	0.826, 0.015	0.770, 0.017	0.939, 0.008	0.620, 0.017	0.863, 0.009	0.902, 0.016
GNN	0.028, 0.024	0.801, 0.016	0.831, 0.015	0.934, 0.008	0.795, 0.014	0.657, 0.014	0.802, 0.023
PINN	0.489, 0.017	0.749, 0.018	0.756, 0.018	0.731, 0.017	0.620, 0.019	0.731, 0.013	0.912, 0.016
SWMM	0.398, 0.019	0.852, 0.014	0.902, 0.011	0.501, 0.023	0.819, 0.013	0.886, 0.022	0.765, 0.026



**Figure 3.** Four types of model simulation errors RMSE and Nash Sutcliffe efficiency in the test set.



**Figure 4.** Monitoring of water level at Node 5 and flow rate in Pipeline 6-5, along with four types of model prediction results.

## Conclusions and future work

In this study, we developed and applied a high-resolution hybrid model that combines the de Saint-Venant system of equations, graph neural network, and Discrete Physics Informed Neural Networks to improve the accuracy and generalization ability of water level and flow simulation in UDS. Looking ahead to the future, it is recommended to apply more multi-source data fusion and numerical calculation methods to further improve and validate the efficiency of UDS models while ensuring a certain level of simulation accuracy. This approach may lead to more efficient and reliable UDS models, thereby supporting the decision-making process of urban water resource management.

## References

- Raissi, M., Perdikaris, P., & Karniadakis, G.E. (2017). Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations. ArXiv, abs/1711.10561.
- Ye, J., Do, N.C., Zeng, W., & Lambert, M.F. (2022). Physics-informed neural networks for hydraulic transient analysis in pipeline systems. *Water research*, 221, 118828 .
- Zhang, Z., Tian, W., Lu, C., Liao, Z., & Yuan, Z. (2024). Graph neural network-based surrogate modelling for real-time hydraulic prediction of urban drainage networks. *Water Research*, 263, Article 122142.
- Li, H., & Shatarah, M. (2024). Operator learning for urban water clarification hydrodynamics and particulate matter transport with physics-informed neural networks. *Water research*, 251, 121123 .