Towards foundational (meta)models of Water Distribution Networks with Graph Neural Networks

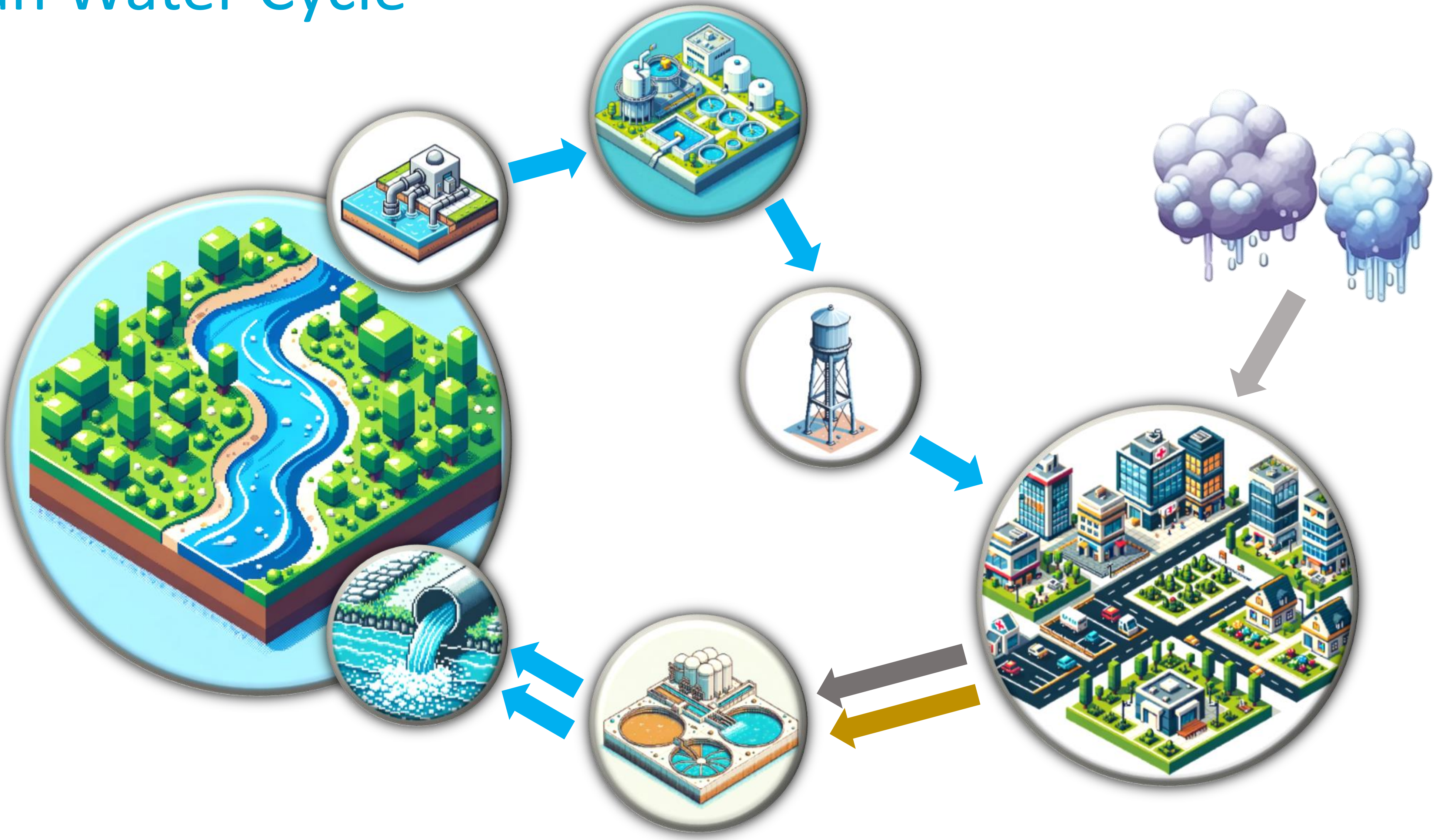
Dr Riccardo Taormina, r.taormina@tudelft.nl
Bulat Kerimov, bulat.kerimov@ntnu.no

Outline of Today’s presentations

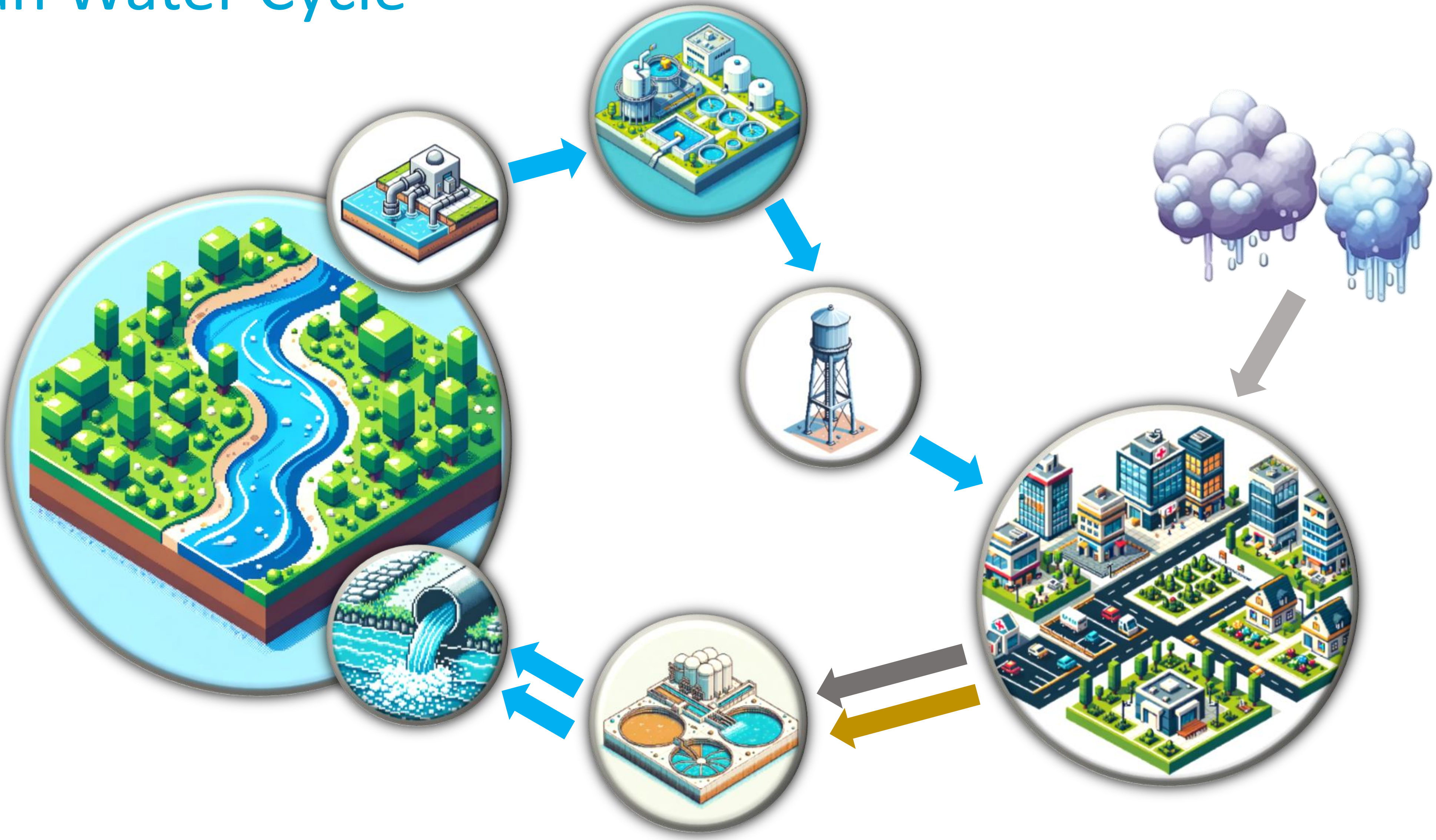
- *Towards foundational (meta)models of Water Distribution Networks with Graph Neural Networks* — Dr R. Taormina
- *Faster and Transferable Urban Drainage Simulations with Graph Neural Networks* — A. Garzón
- *Relating complex network theory metrics with discoloration activity in Water Distribution Systems* — Dr G. Kyritsakas

PART 1
Urban Water Networks and Graph Machine Learning

Urban Water Cycle



Urban Water Cycle



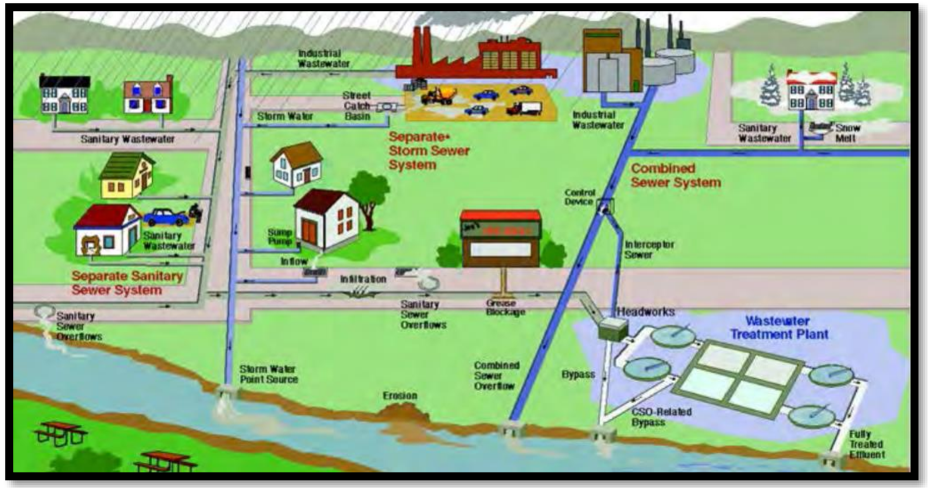
Urban Water Networks

Water Distribution Networks

Deliver treated, **potable water,** with sufficient pressure from source facilities to residential, commercial, and industrial consumers.

Urban Drainage Networks

Collect and transport **storm water** and **wastewater** away from populated areas to prevent flooding and grant sanitation.



Physical vs. Non-Physical Networks

Node/Edge Features:

- Physical: Attributes have physical meaning (e.g., pipe diameters and flowrate).
- Non-Physical: Attributes are more abstract or symbolic (e.g., number of connections in a social network or ratings in a recommender system).

Laws & Constraints:

- Physical: Governed by physical laws (e.g., fluid dynamics in a water network).
- Non-Physical: Governed by non-physical patterns and principles (e.g., user behavior or social norms).

Spatial Relationship:

- Physical: Spatial relation is crucial (e.g., physical distance affects energy loss).
- Non-Physical: Spatial relation is typically not relevant or abstracted away (e.g., a social network connection can span large distances).

Other considerations for *temporal dynamics*, *noise/uncertainty*, *scalability*, ...

Water Distribution Networks

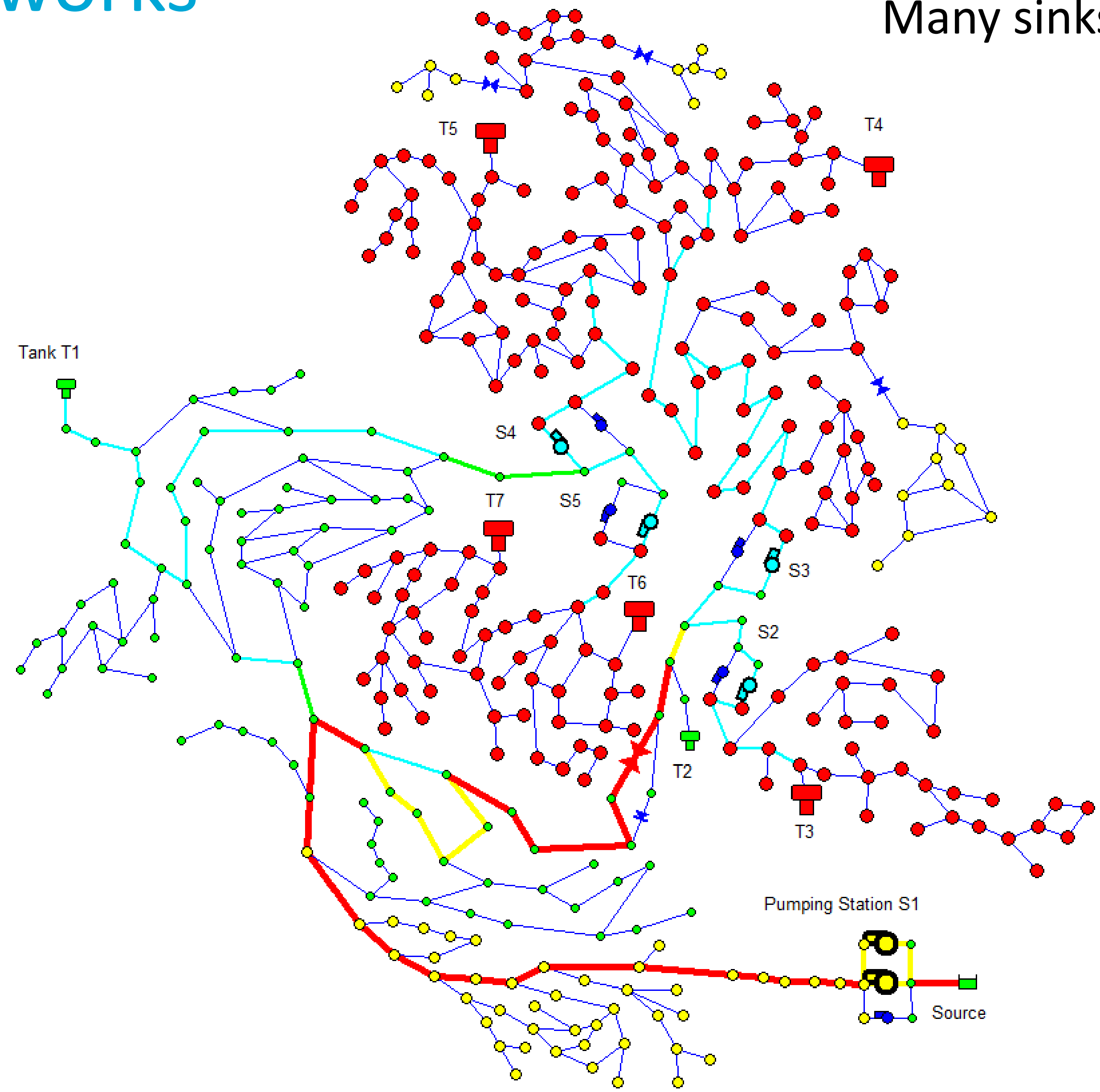
Few sources,
Many sinks

Node Features

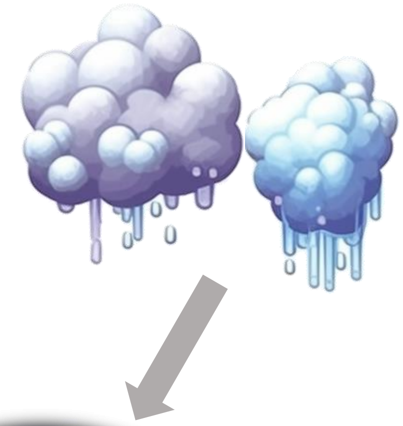
Static	Dynamic
Type: junction, tank, reservoir	Pressure, Head(=pressure + node elevation)
Node elevation	Water demand
Tank volume	Water quality parameters
…	…

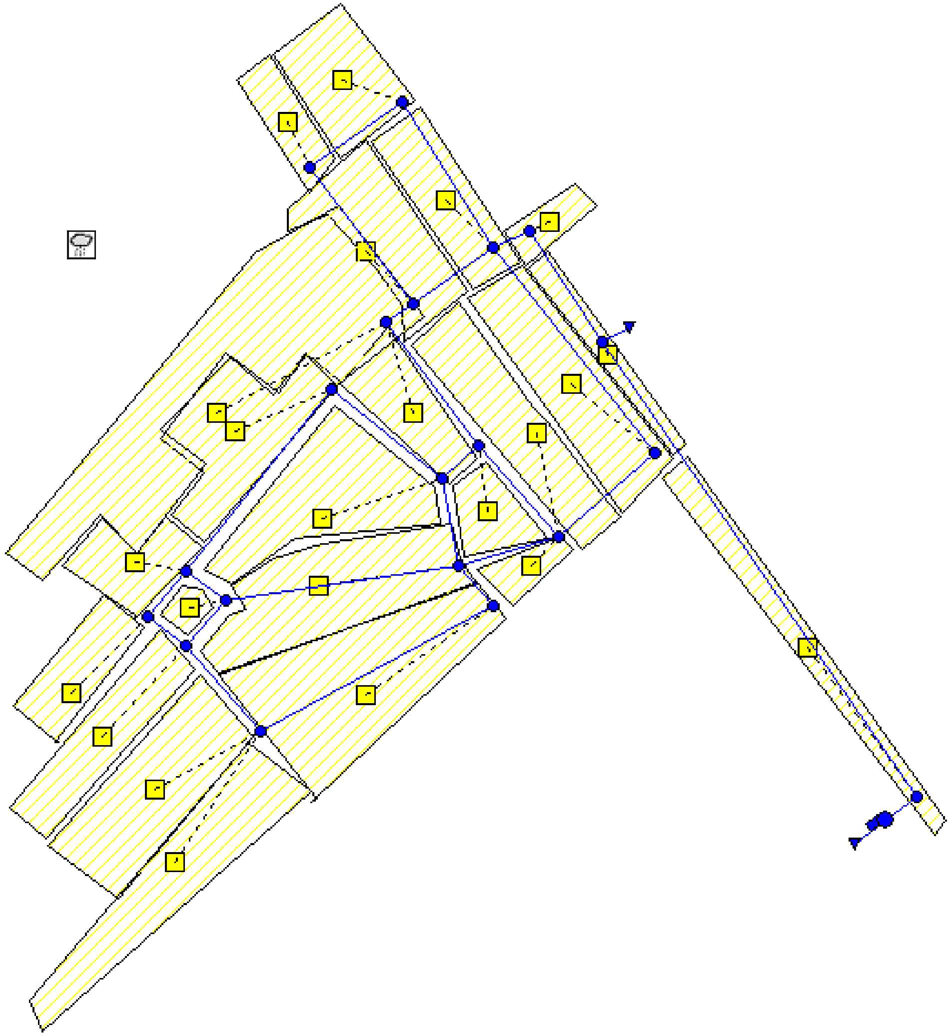
Edge Features

Static	Dynamic
Type: pipes, valves, pumps	Flowrate, velocity
Pipe geometry: diameter, length, area.	Headloss
Roughness coefficient	Water quality parameters
Pump curve	Link status (e.g., on/off, closed/open)



Urban Drainage Systems: Stormwater Sewers

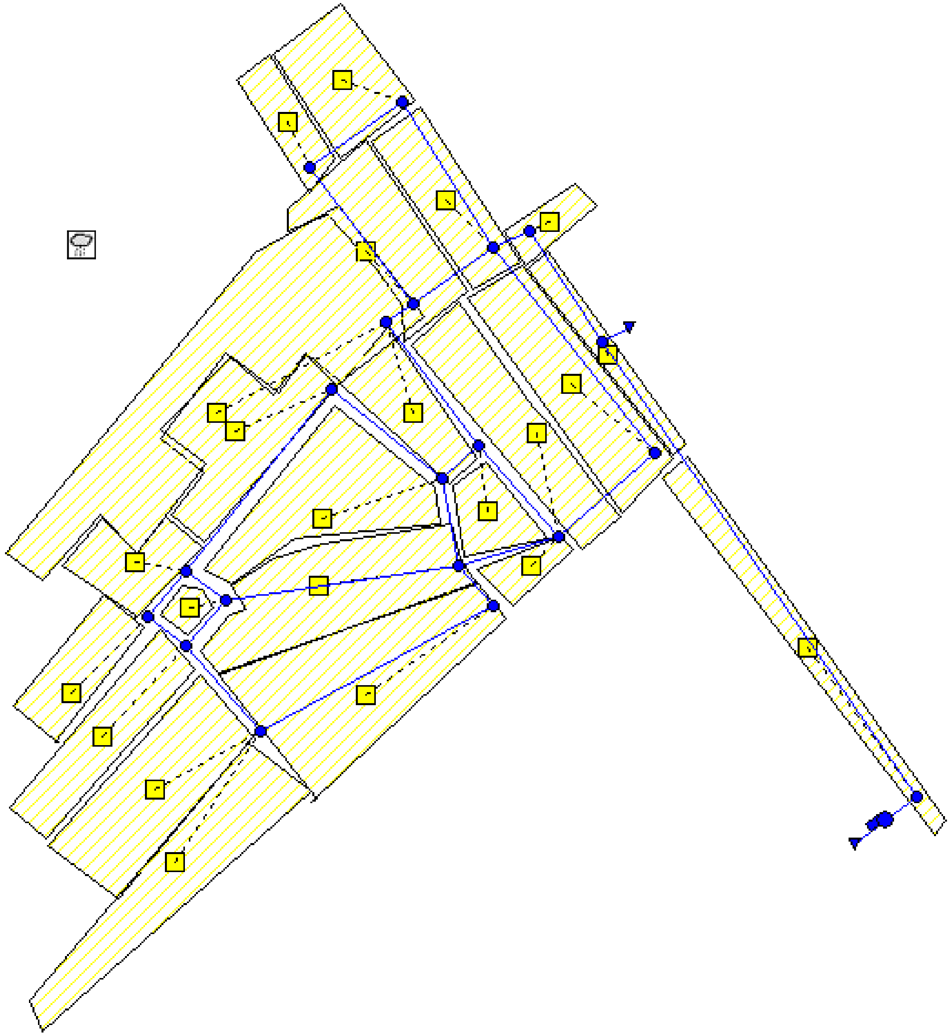




Collect and transport **storm water** from populated areas to prevent flooding.

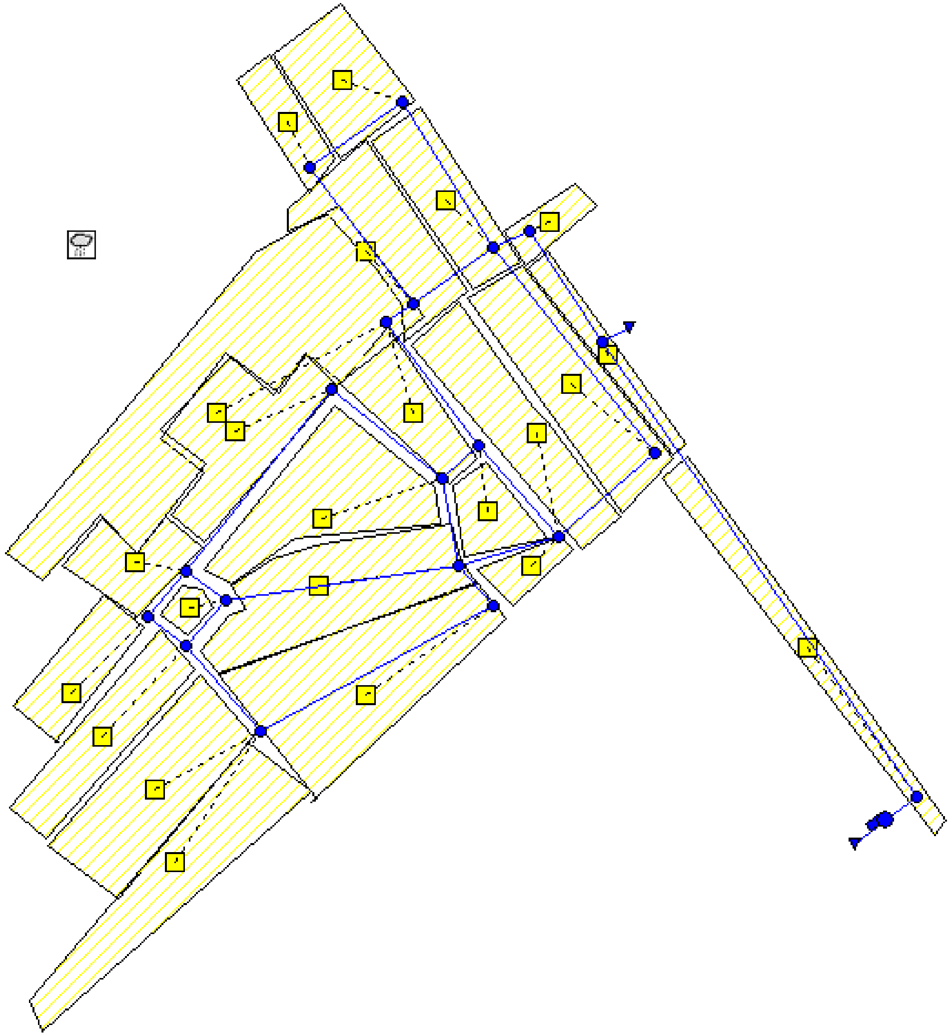
Urban Drainage Systems: Wastewater Sewers

Also known as **sanitary sewers**: collect and transport **wastewater** from populated areas to provide sanitation.



Urban Drainage Systems: Combined Sewers

Perform **both functions together**, mainly to save space in densely populated areas.



Urban Drainage Systems

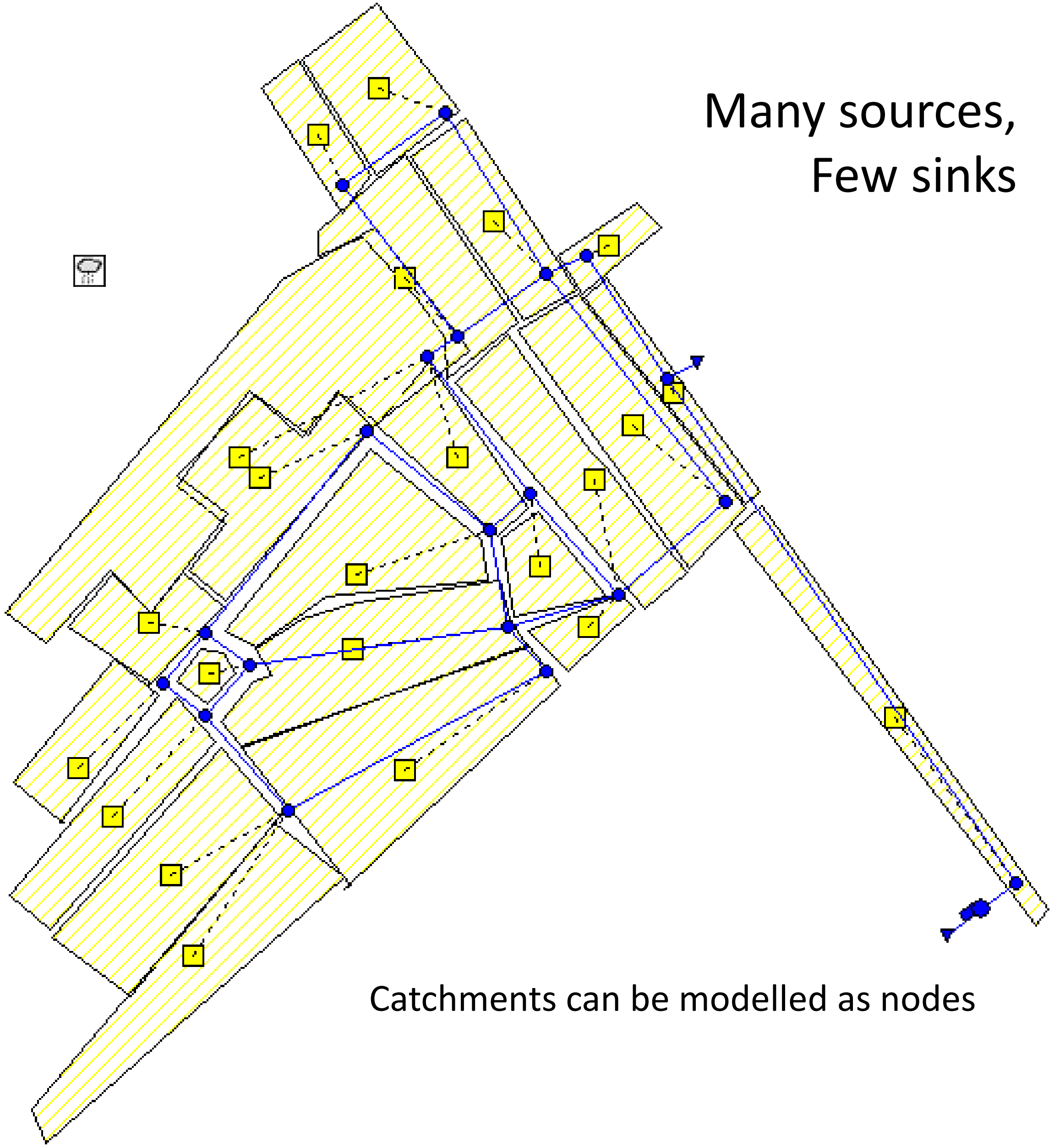
Node Features

Static	Dynamic
Node type: junction, storage, outlets, catchments, ...	Pressure, Water Depths
Node elevation	Head (=pressure + elevation)
Storage volume	Runoff, Wastewater inflow
Catchment characteristics	Water quality parameters

Edge Features

Static	Dynamic
Link type: pipes, valves, pumps, weirs, ...	Flowrate
Pipe geometry: diameter, length, ...	Headloss
Roughness coefficient	Water quality parameters
Pump curve	Link status (e.g., on/off, closed/open)

Many sources,
Few sinks



Catchments can be modelled as nodes

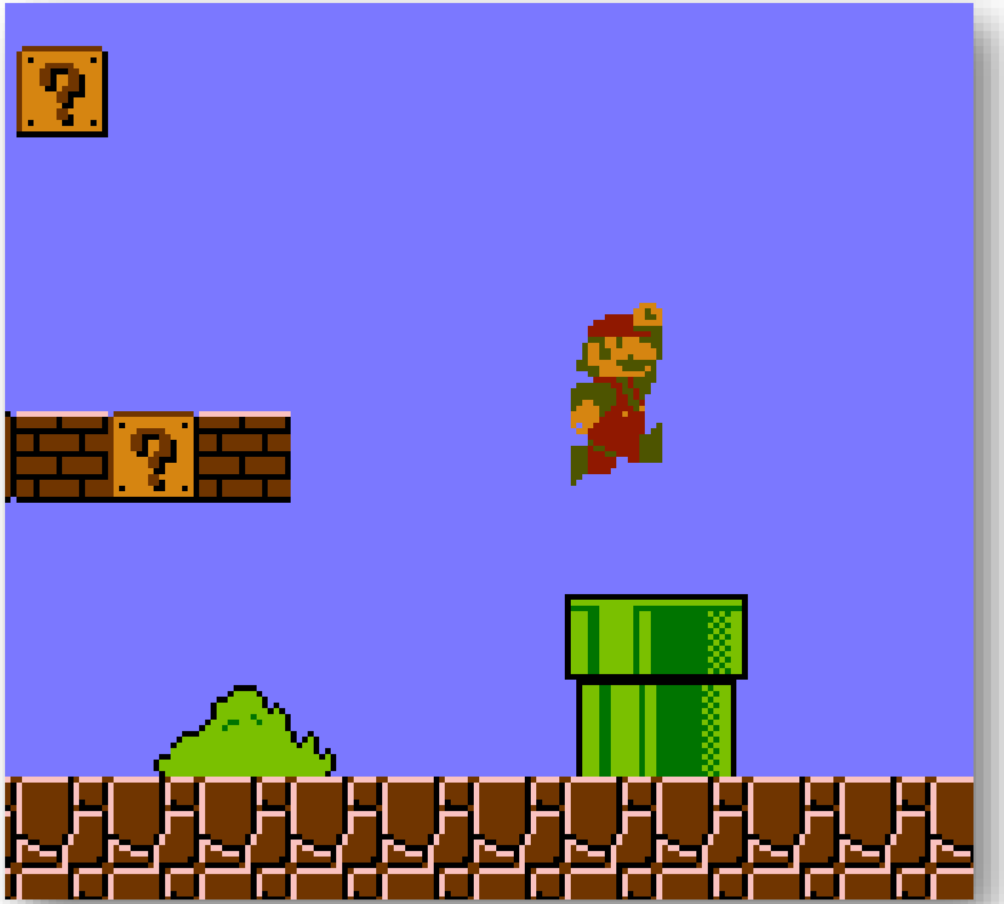
What pipes does Mario use?

A) Water Distribution Network

B) Stormwater Sewer

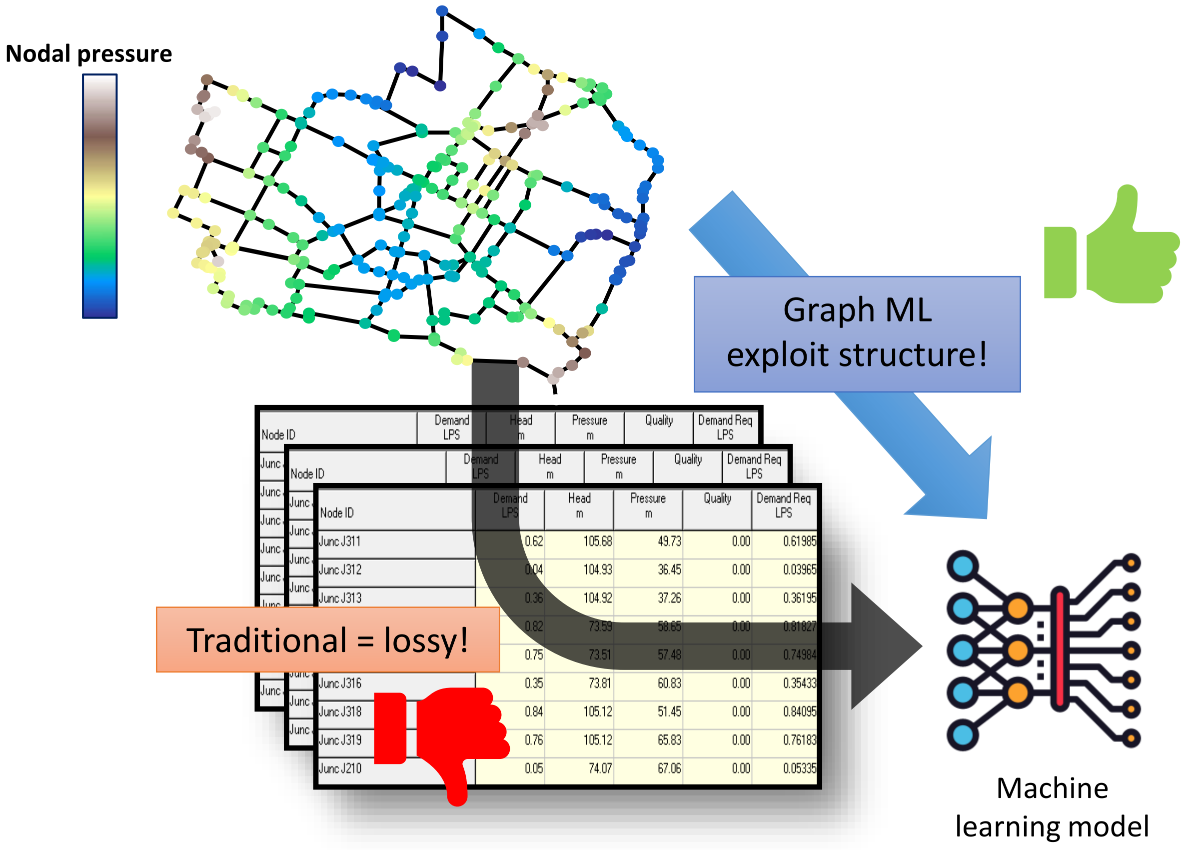
C) Sanitary Sewer

D) Combined Sewer



Graph Machine Learning for Water Networks

Inductive bias: data ***is*** or ***comes*** from a **network**



Graph ML tasks in Water Networks (1/2)

- **State Estimation**: infer the state of the networks (e.g., head, flows) from a few sensors (e.g., pressure gauge, flow meter)
  - Node/edge regression
- **Leak Detection**: Identifying potential leaks in the water network by detecting anomalies in water pressure or flow data
  - Node/edge classification/regression
- **Water quality** monitoring/forecasting
  - Node/edge regression (e.g., depends where water quality is sampled)
- **Asset Maintenance**: what components to repair or substitute?
  - Node/edge ranking

Graph ML tasks in Water Networks (2/2)

- **Blockage identification**: detect blockages in UDS (e.g., from accumulation of debris, collapses, roots, ...) or unreported closed valves
  - Edge classification
- **Sewer overflow**: predict the release of untreated sewage into the environment due to reached system capacity or blockages
  - Node regression (at the outlet node)
- **Estimation of Network Resilience**: predict the ability of the system to withstand and recover from disruption (e.g., due to redundancy)
  - Graph regression
- **Metamodelling**: reproduce and generalize physics-based simulation with high accuracy and considerable speedups
  - Node/edge regression

Graph ML tasks in Water Networks (2/2)

- **Blockage identification**: detect blockages in UDS (e.g., from accumulation of debris, collapses, roots, ...) or unreported closed valves
  - Edge classification
- **Sewer overflow**: predict the release of untreated sewage into the environment due to reached system capacity or blockages
  - Node regression (at the outlet node)
- **Estimation of Network Resilience**: predict the ability of the system to withstand and recover from disruption (e.g., due to redundancy)
  - Graph regression
- **Metamodelling**: reproduce and generalize physics-based simulation with high accuracy and considerable speedups
  - Node/edge regression

PART 2
Transferable Metamodels for Water Distribution Networks

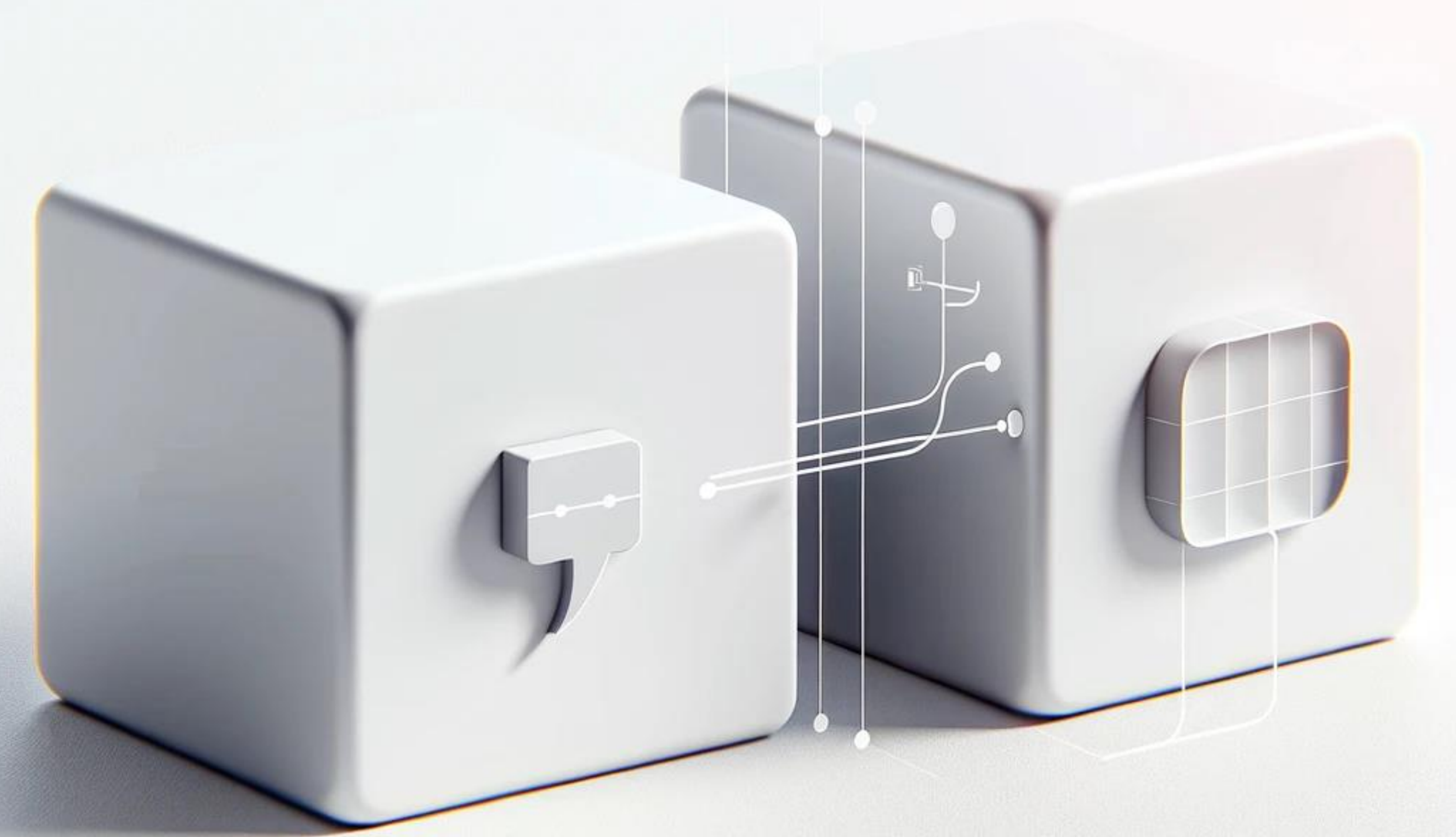
AI & Digital Twins

… but where is the data?

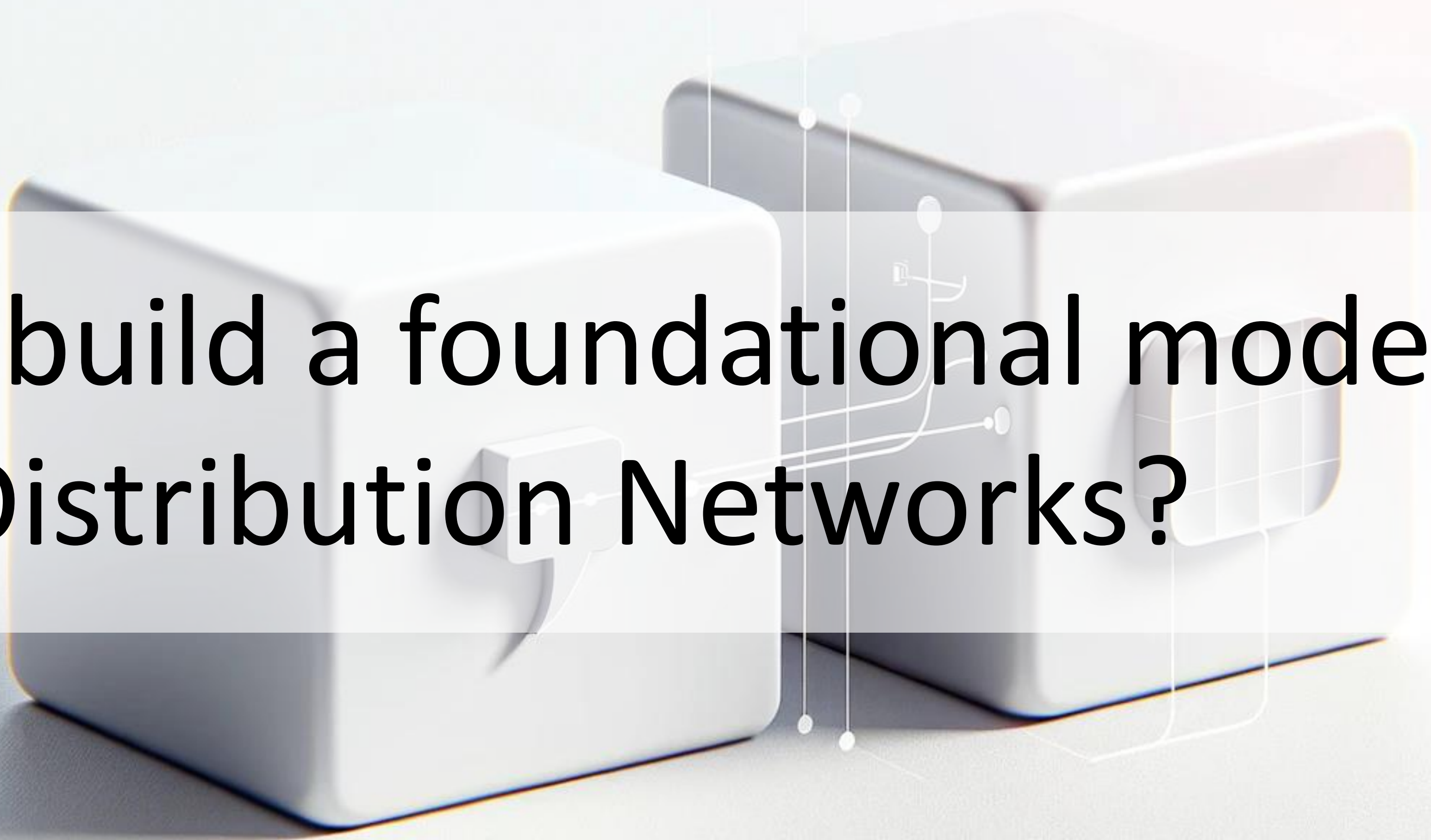
How it should be....

How it is...

Transfer Learning



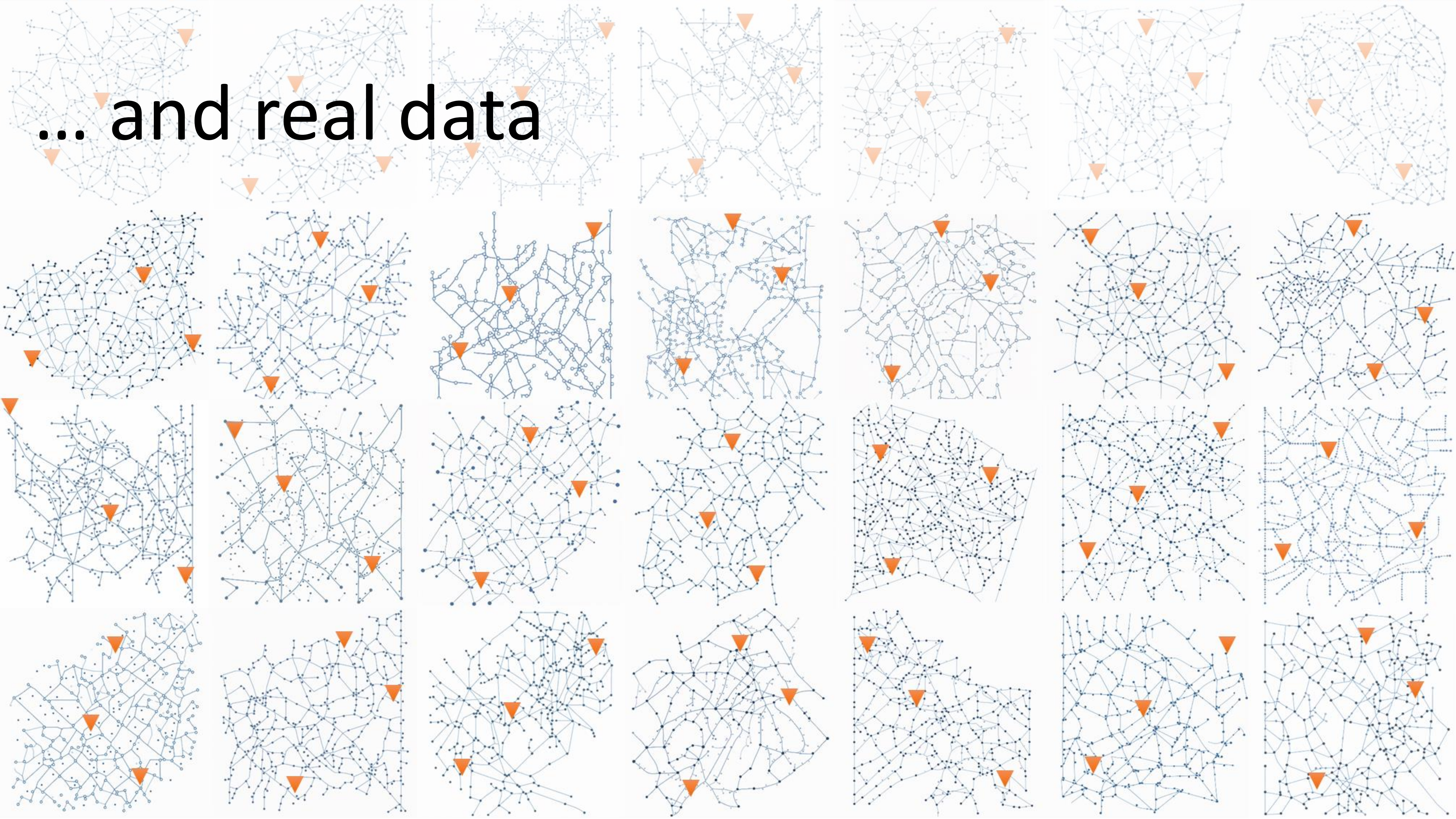
Transfer Learning



Can we build a foundational model for Water Distribution Networks?

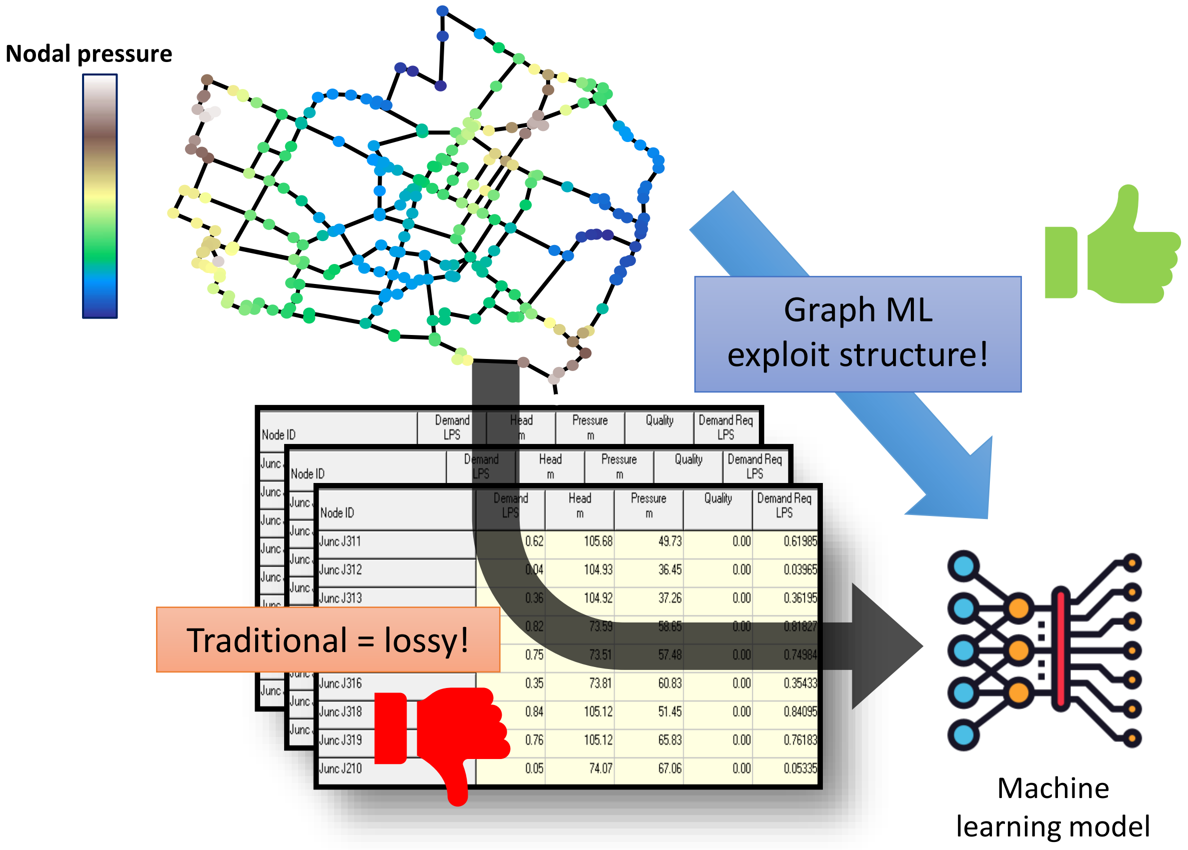
Yes? From lots of simulations….

… and real data

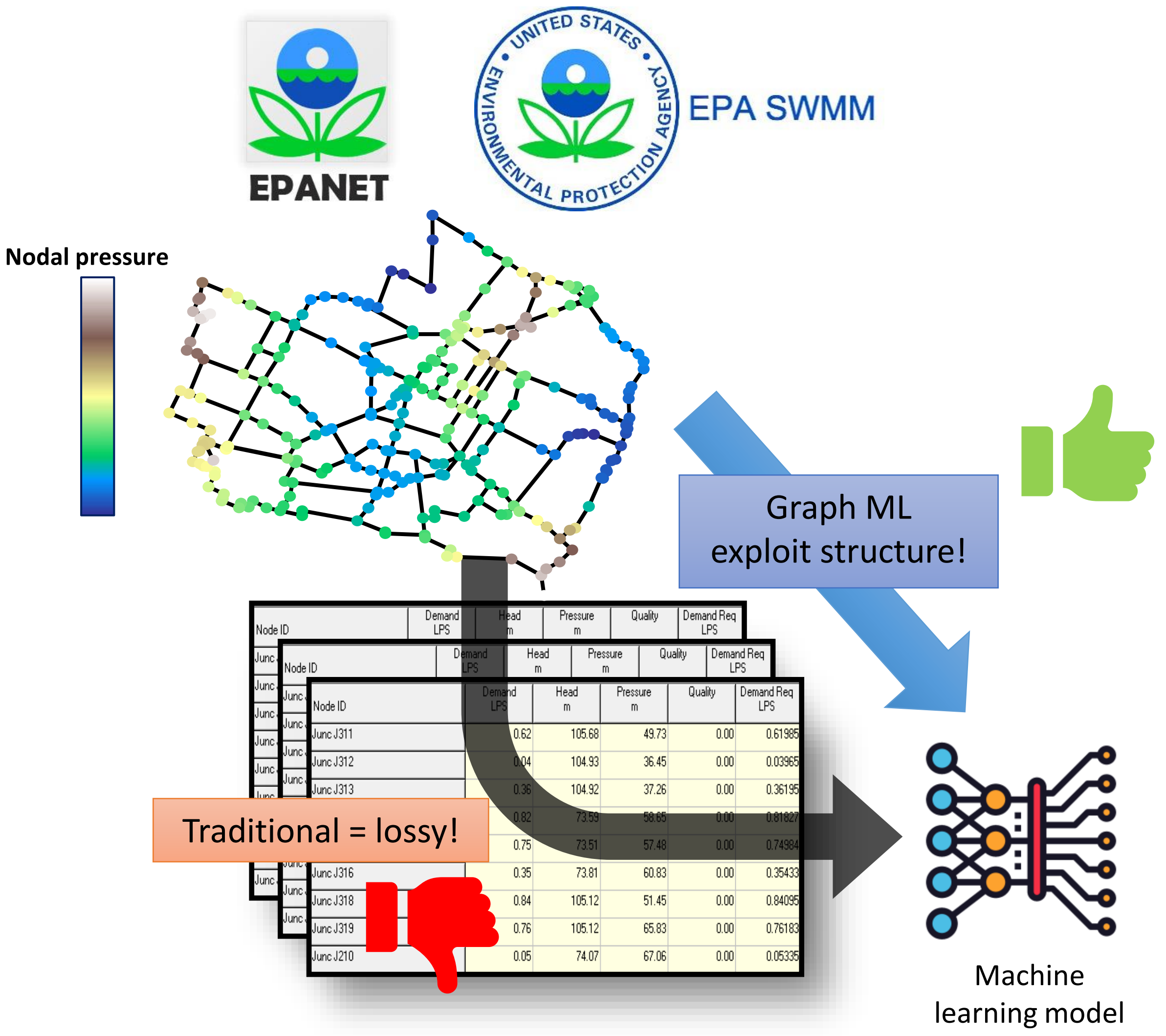


We need Graph Neural Networks…

Learn transferable representation across multiple graphs

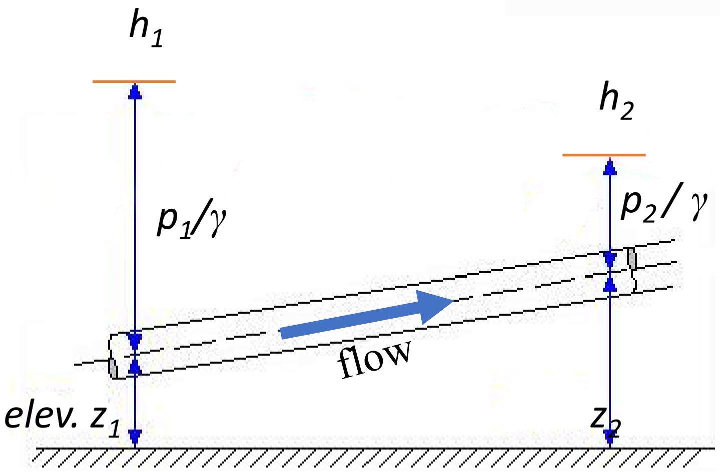


Learning from simulations: metamodelling



Basic Physics in Water Distribution System

Conservation of Energy	$h_i - h_j = h_{Lij}(q_{ij})$
Headloss formula	$h_{Lij} = r\, q_{ij}\, \lvert q_{ij} \rvert^{n-1} + m\, q_{ij}\, \lvert q_{ij} \rvert$
Conservation of Mass	$\sum_j q_{ij} - D_i = 0$



γ =specific weight of water

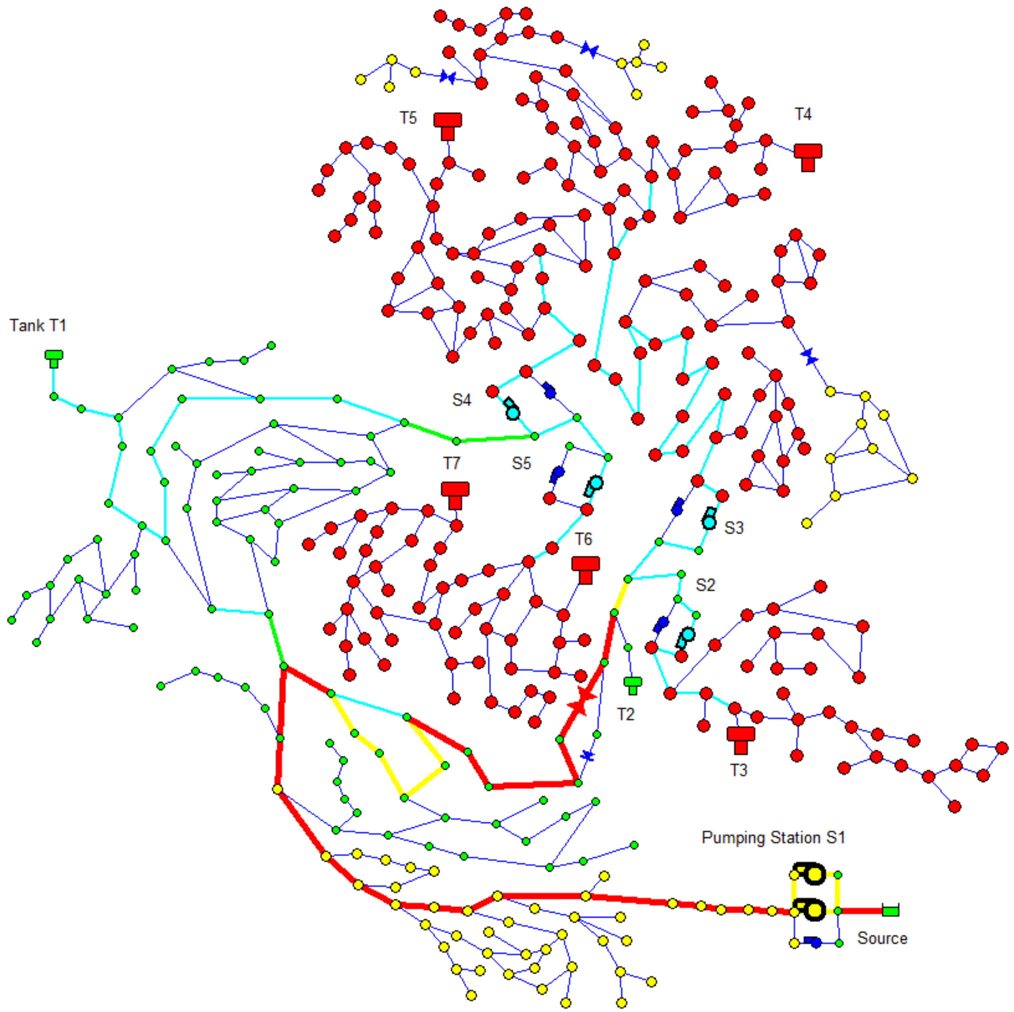
WDS flow is always pressurized

h_i	head in node *i*
h_{Lij}	headloss in pipe *ij*
q_{ij}	flow in pipe *ij*
D_i	water demand of node *i*
h_i	head in node *i*
r	resistance coefficient
n	flow exponent
m	minor losses

Source https://epanet22.readthedocs.io/

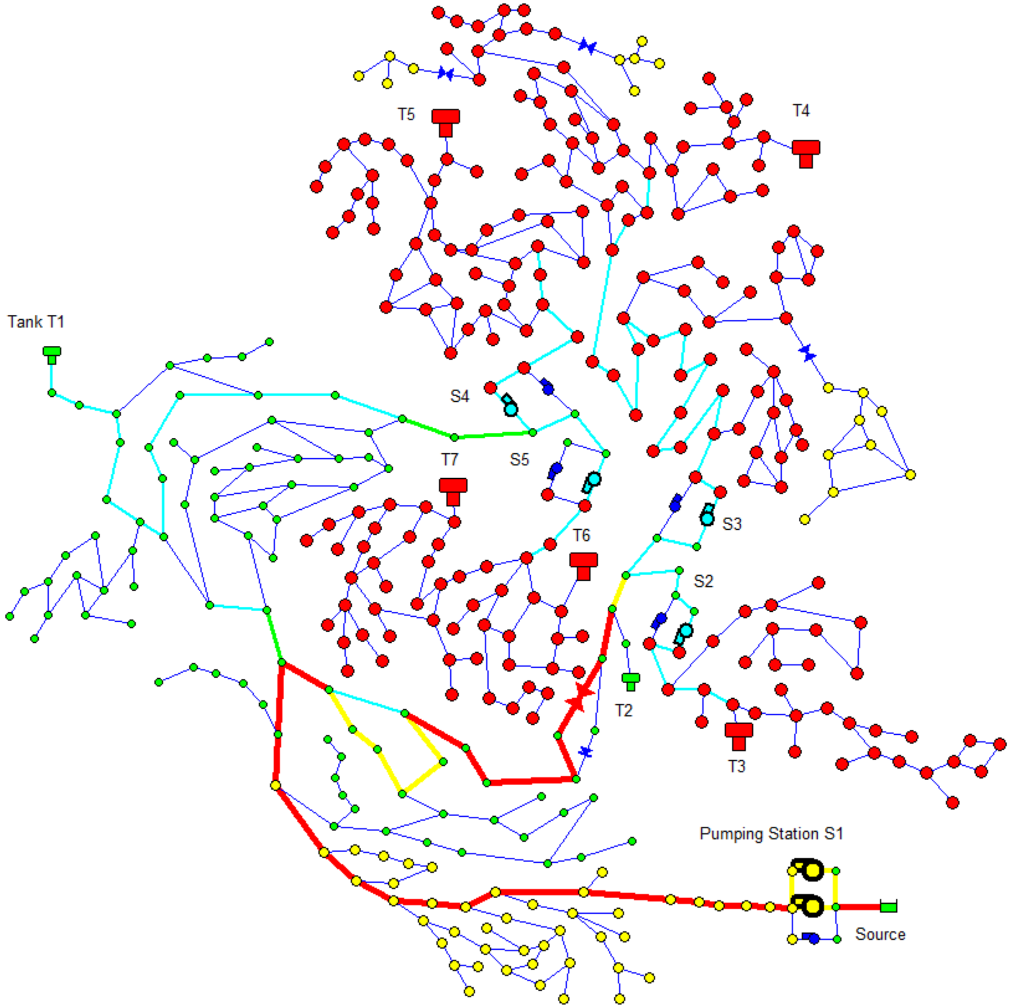
Problem Definition: steady-state simulations

- Given inputs
  - Water demand requested by all nodes of the system
  - Head (energy) of water source (reservoir)
  - Network geometry/characteristics
- Determine
  - Pressure at all junctions (nodes)
  - Water flow in all pipes (edges)
- Simplifications
  - Steady-state conditions
  - No valves/pumps/storage tanks

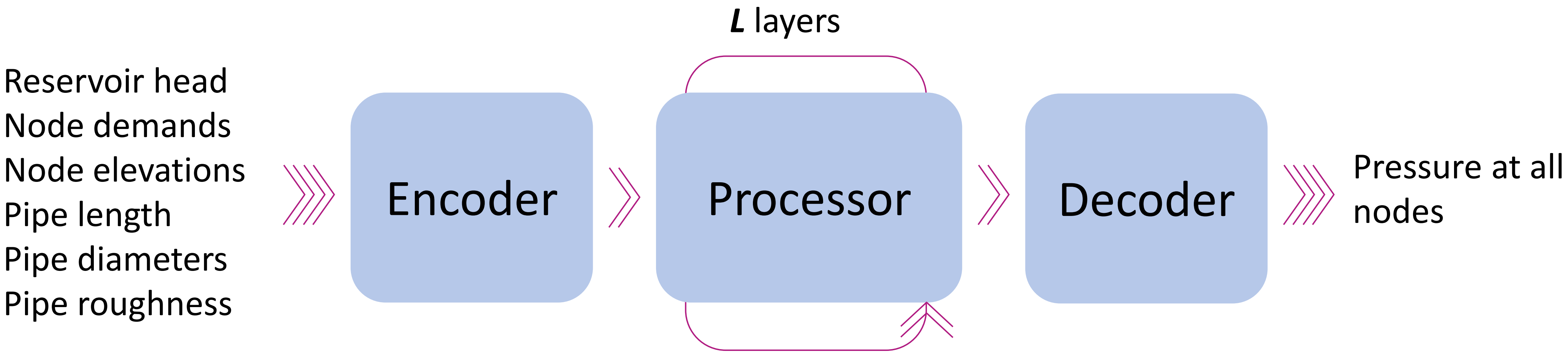


Problem Definition: transferable metamodeling

- We run several simulations in EPANET on different case studies
- We train a GNN to reproduce the simulations
- We check whether the GNN can learn shared representations across case studies.

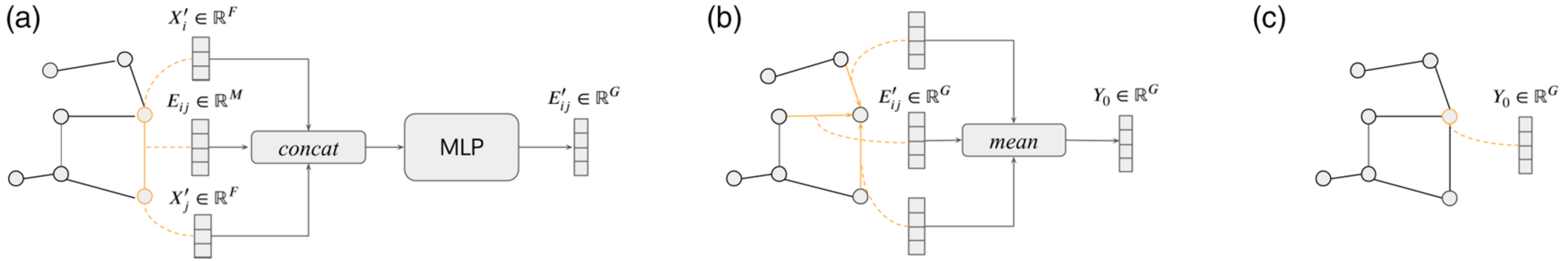


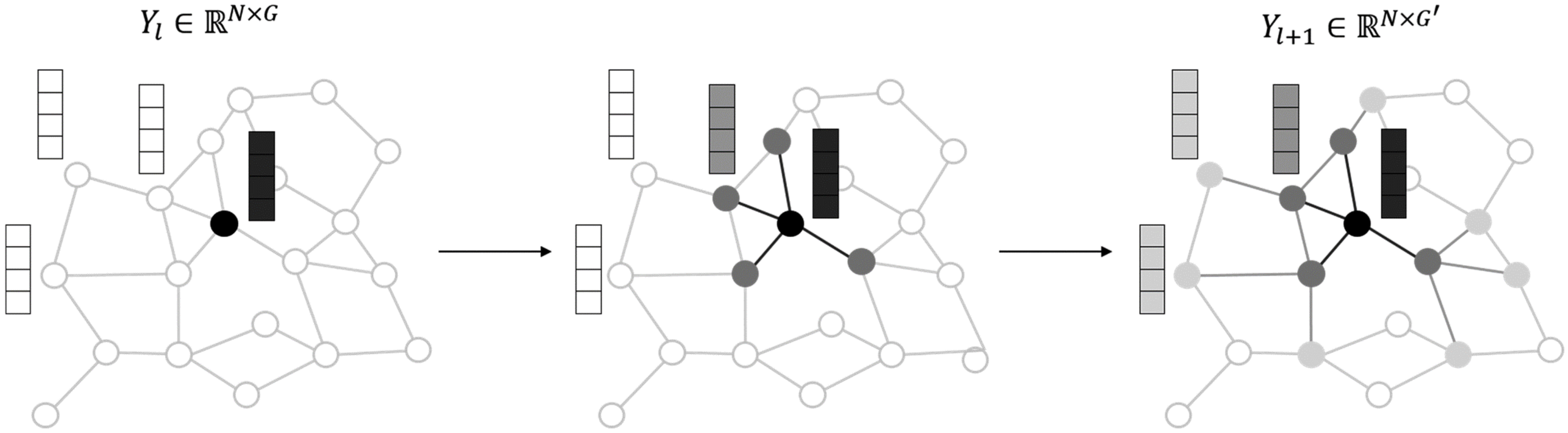
Node-based GNN for metamodeling



Kerimov, Bulat, et al. "Assessing the performances and transferability of graph neural network metamodels for water distribution systems." *Journal of Hydroinformatics* 25.6 (2023): 2223-2234.

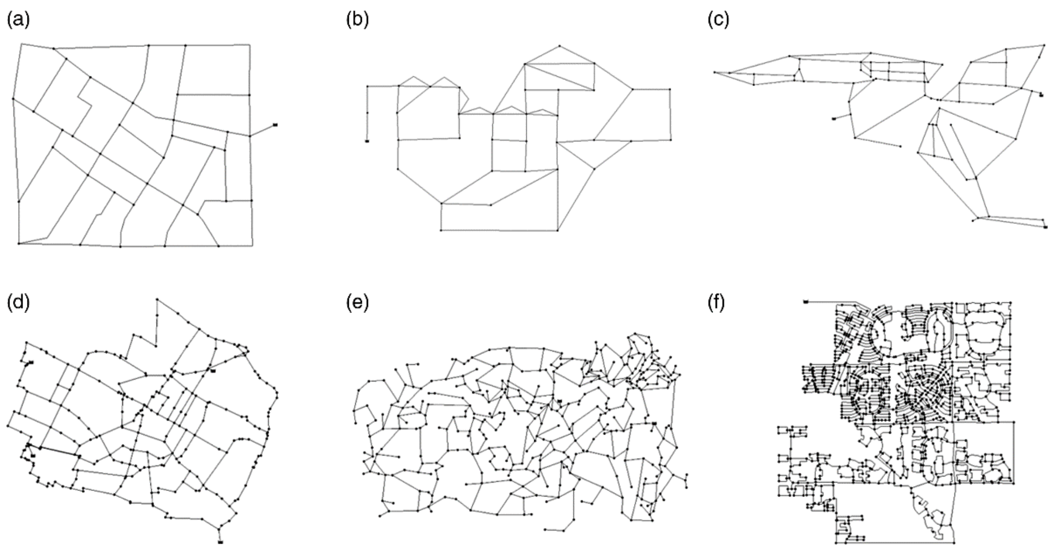
Node-based GNN for metamodeling





Kerimov, Bulat, et al. "Assessing the performances and transferability of graph neural network metamodels for water distribution systems." *Journal of Hydroinformatics* 25.6 (2023): 2223-2234.

Results: single case study



	MLP			GNN		
	R_squared	*Speedup*	*# params*	*R_squared*	*Speedup*	*# params*
(a) FOS	0.379	879	200k	**0.815**	71	60k
(b) BAK	**0.993**	1393	50k	**0.993**	56	60k
(c) PES	**0.561**	1241	200k	0.445	43	200k
(d) MOD	**0.868**	2223	300k	0.763	24	200k
(e) RUR	**0.929**	2029	500k	0.906	27	200k
(f) KL	**0.482**	4001	300k	0.468	22	200k

Kerimov, Bulat, et al. "Assessing the performances and transferability of graph neural network metamodels for water distribution systems." *Journal of Hydroinformatics* 25.6 (2023): 2223-2234.

Results: learning shared representations

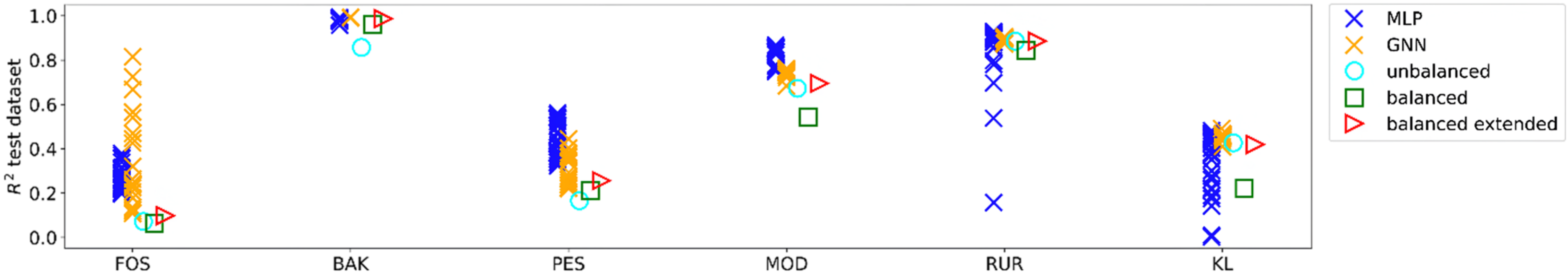


Table 2 | Training datasets used for the study on transferability

Name	FOS	BAK	PES	MOD	RUR	KL	Total
Unbalanced	1,024	1,024	1,024	1,024	1,024	1,024	6,144
Balanced	2,048	2,048	1,067	279	199	81	5,722
Balanced extended	8,000	8,000	4,169	1,088	777	316	22,350

Kerimov, Bulat, et al. "Assessing the performances and transferability of graph neural network metamodels for water distribution systems." *Journal of Hydroinformatics* 25.6 (2023): 2223-2234.

Conclusions for node-based GNN

- Overall worse performance than MLPs
- Prone to over-smoothing
- Limited transferability

Kerimov, Bulat, et al. "Assessing the performances and transferability of graph neural network metamodels for water distribution systems." *Journal of Hydroinformatics* 25.6 (2023): 2223-2234.

Edge-based GNNs

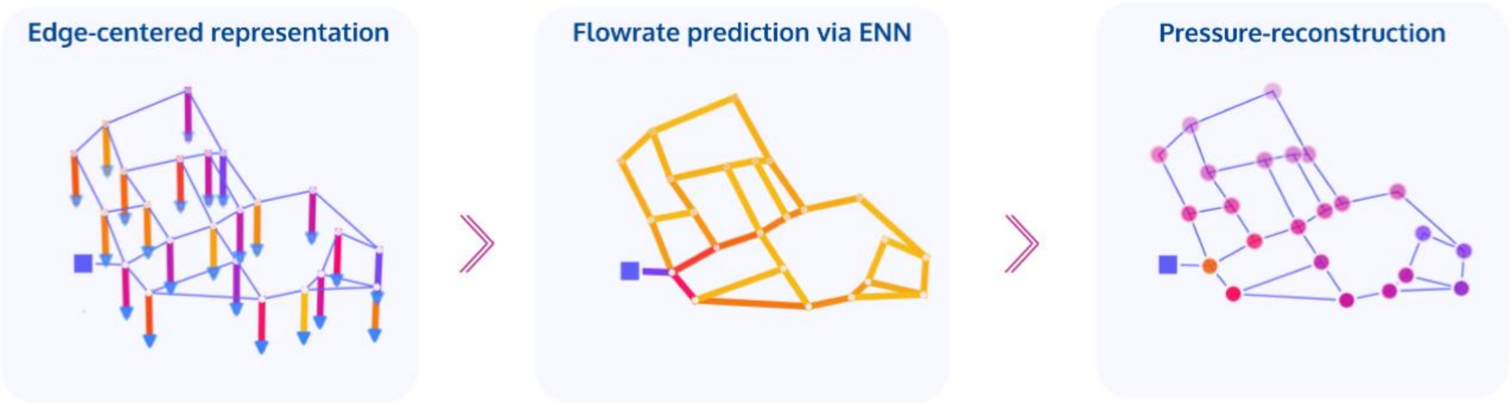


Figure 2: Overview of the model predicting steps. The second step (center) reconstructs flowrates based on the augmented representation of the network with virtual sinks. Next, nodal pressures are calculated based on Hazen-Williams and conservation laws.

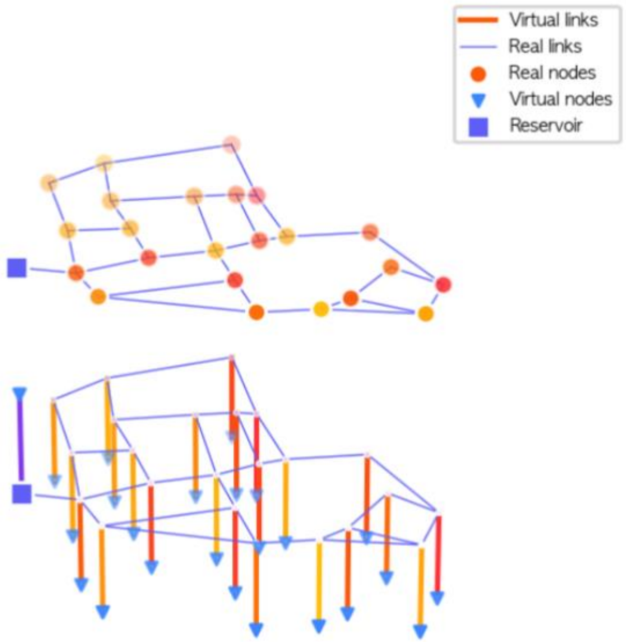


Figure 3: Visualization of virtual sinks. Each node is augmented with a virtual sink that emulates the flow out of the system based on the consumption volume. In the edge level representation the virtual sinks act as flowmeter sensors.

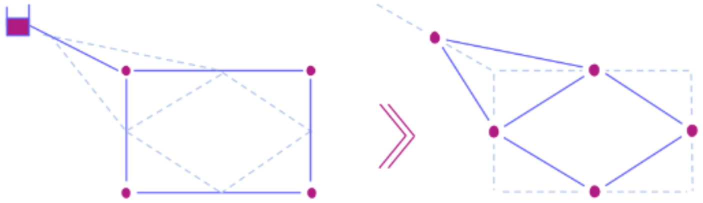


Figure 4: Lifting of the representation to edge level with corresponding connectivity matrices.

Basic Physics in Water Distribution System

Conservation of Energy

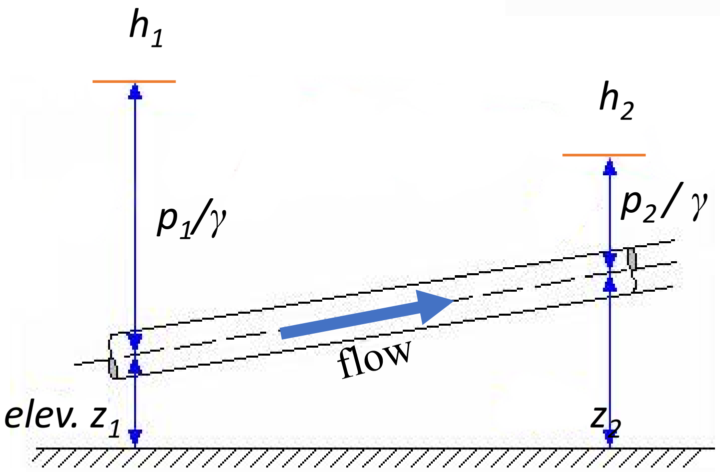
$$h_i - h_j = h_{Lij}(q_{ij})$$

Headloss formula

$$h_{Lij} = r\, q_{ij}\, |q_{ij}|^{n-1} + m\, q_{ij}\, |q_{ij}|$$

Conservation of Mass

$$\sum_j q_{ij} - D_i = 0$$



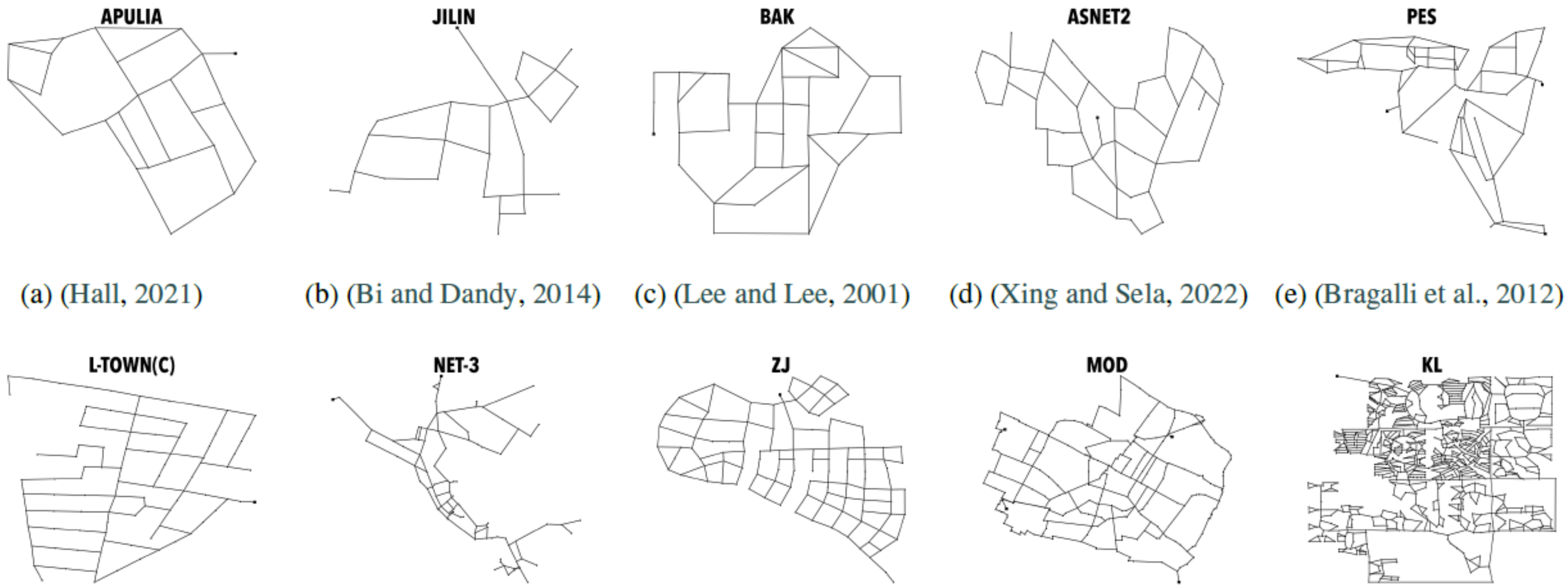
γ =specific weight of water

WDS flow is always pressurized

h_i	head in node *i*
h_{Lij}	headloss in pipe *ij*
q_{ij}	flow in pipe *ij*
D_i	water demand of node *i*
h_i	head in node *i*
r	resistance coefficient
n	flow exponent
m	minor losses

Source https://epanet22.readthedocs.io/

Case Studies

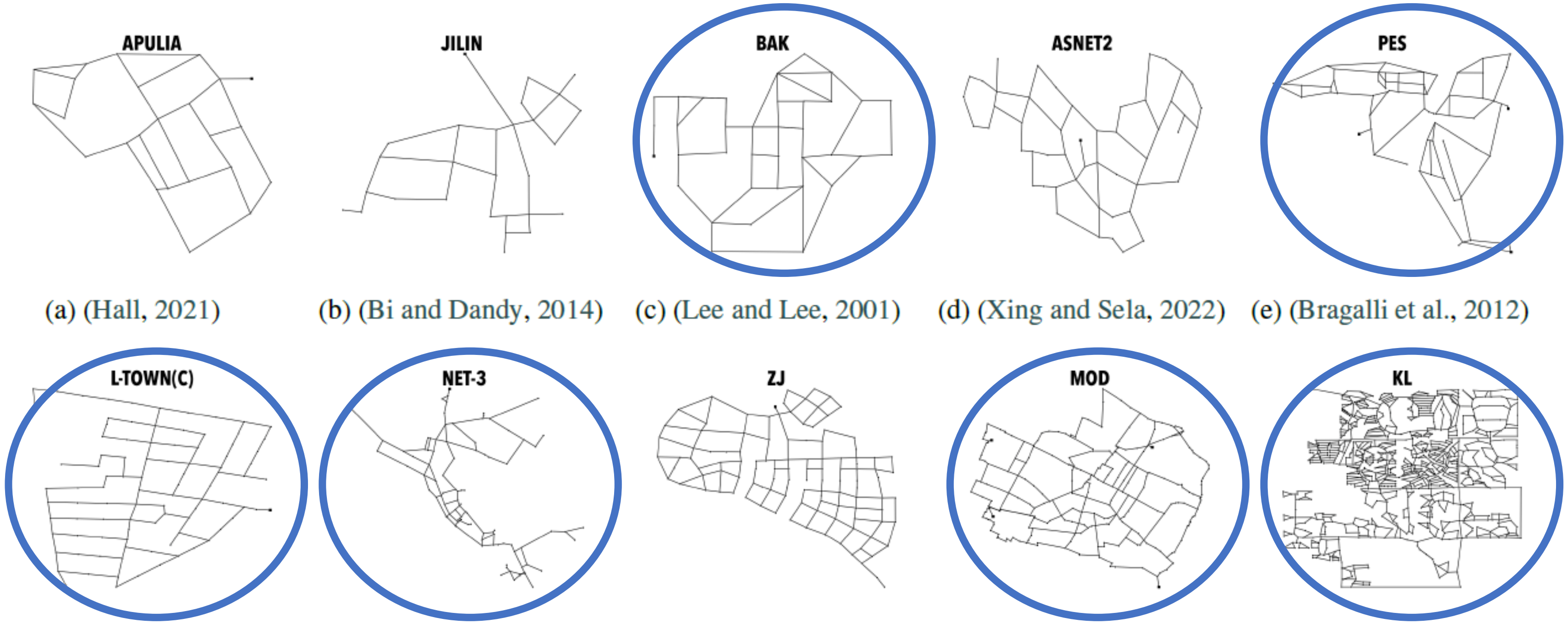


(a) (Hall, 2021) (b) (Bi and Dandy, 2014) (c) (Lee and Lee, 2001) (d) (Xing and Sela, 2022) (e) (Bragalli et al., 2012)

(f) (Vrachimis et al., 2022) (g) (Rossman, 2016) (h) (Dandy, 2016) (i) (Bragalli et al., 2012) (j) (Kang and Lansey, 2012)

Kerimov, Bulat, et al. "Towards transferable metamodels for water distribution systems with edge-based graph neural networks." *in Review*

Experiment 1: in-the-domain generalization



(a) (Hall, 2021) (b) (Bi and Dandy, 2014) (c) (Lee and Lee, 2001) (d) (Xing and Sela, 2022) (e) (Bragalli et al., 2012)

(f) (Vrachimis et al., 2022) (g) (Rossman, 2016) (h) (Dandy, 2016) (i) (Bragalli et al., 2012) (j) (Kang and Lansey, 2012)

Kerimov, Bulat, et al. "Towards transferable metamodels for water distribution systems with edge-based graph neural networks." *in Review*

Results: in-the-domain generalization

Towards transferable metamodels of water distribution systems with edge-based graph neural network

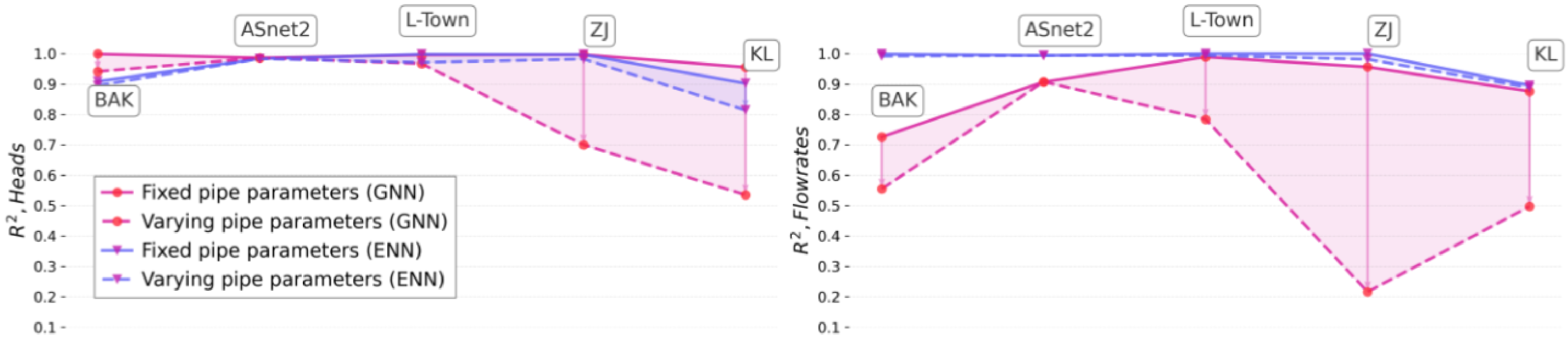
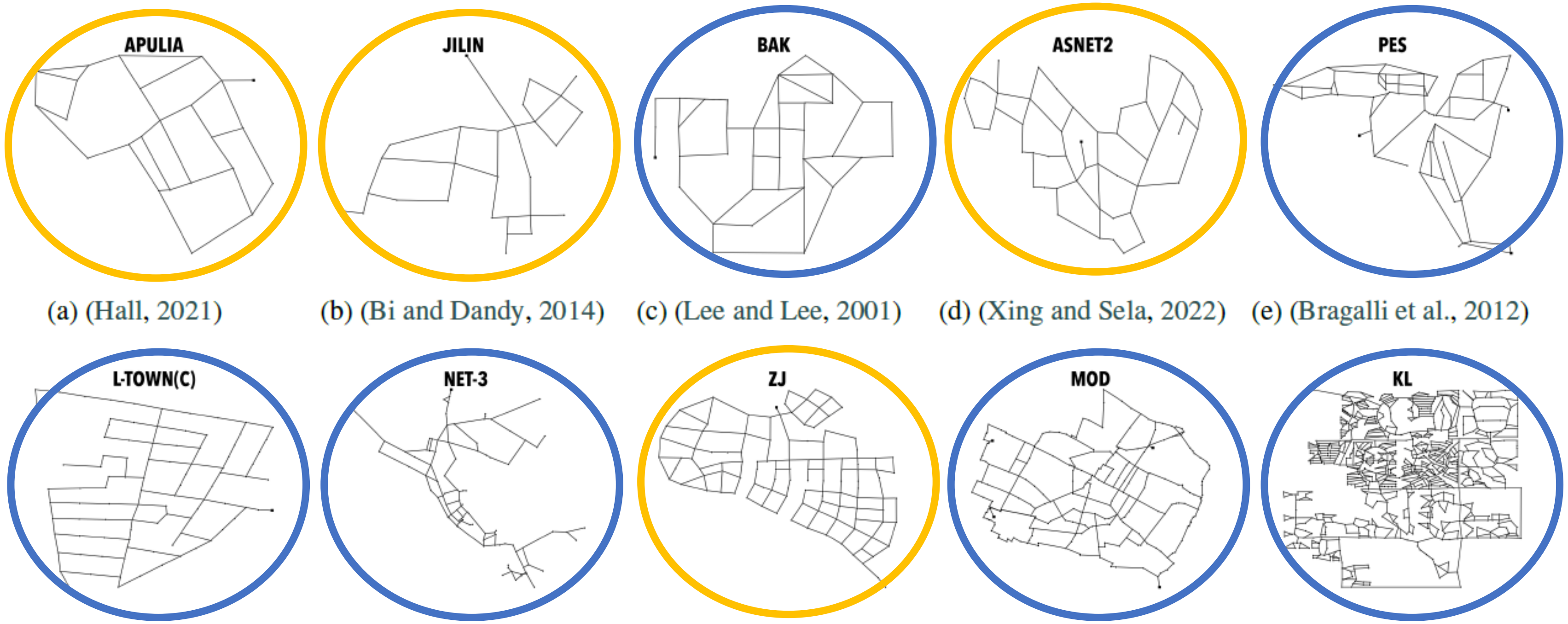


Figure 7: In-the-domain comparison of accuracy in terms of R^2 of predicted heads (left) and flowrates (right) between GNN (pink) and ENN (purple). Dashed line shows the performance of the model trained on the subset with varying pipe parameters, while a solid line indicates the performance of a model trained on the subset with fixed and known pipe parameters.

Kerimov, Bulat, et al. "Towards transferable metamodels for water distribution systems with edge-based graph neural networks." *in Review*

Experiment 2: transferability



(a) (Hall, 2021) (b) (Bi and Dandy, 2014) (c) (Lee and Lee, 2001) (d) (Xing and Sela, 2022) (e) (Bragalli et al., 2012)

(f) (Vrachimis et al., 2022) (g) (Rossman, 2016) (h) (Dandy, 2016) (i) (Bragalli et al., 2012) (j) (Kang and Lansey, 2012)

Kerimov, Bulat, et al. "Towards transferable metamodels for water distribution systems with edge-based graph neural networks." *in Review*

Results: transferability

Results of evaluation of out-of-domain water networks

Case study	Average demand $\bar{\mathbf{q}}$, L/s	Maximum κ_f	Heads, R^2	Flowrates, R^2
ASnet2	5	$4.5 \cdot 10^1$	0.832	0.793
ZJ	5	$3.5 \cdot 10^2$	0.858	0.848
Jilin	5	$1.5 \cdot 10^4$	0.950	0.983
Apulia	5	$7.4 \cdot 10^5$	0.883	0.982

Kerimov, Bulat, et al. "Towards transferable metamodels for water distribution systems with edge-based graph neural networks." *in Review*

Conclusions for edge-based GNNs

- Model based on edge convolutions are more accurate
- They show much better transferability
- Speedups from 350 to 10 times with respect to EPANET simulations, depending on size of the network
- Reduction of speedups mainly due to pressure reconstruction from flows.
- Simultaneous prediction of pressures and flowrates can provide better speedups.

Kerimov, Bulat, et al. "Towards transferable metamodels for water distribution systems with edge-based graph neural networks." *in Review*

Future work

- Representation of valves, tanks and pumps for more realism
- From single steady-state simulation to extensive simulations (i.e., over 1 day, 1 week, …)
- Training on a much larger set of networks (e.g., synthetic)
- Fine-tuning with real data

Kerimov, Bulat, et al. "Towards transferable metamodels for water distribution systems with edge-based graph neural networks." *in Review*

Thanks for listening! Questions?

Dr Riccardo Taormina, r.taormina@tudelft.nl
Bulat Kerimov, bulat.kerimov@ntnu.no

APPENDIX

Physics of Urban Drainage Systems

- More complex than in water distribution systems
- Flow regime changes depending on volumes
  - Gravity flow when the pipes are partially full
  - Pressurized flow when the pipes are full (like in WDS)

Results: single case study

Table 5 | Hyperparameters and R^2scores for the best metamodels

Model		FOS	PES	PES	MOD	RUR	KL
MLP	# hidden units	256	128	256	256	256	64
	# hidden layers	4	3	3	2	2	2
	Dropout	0	0.25	0.25	0	0	0.25
	R^2 validation	0.364	0.991	0.570	0.859	0.944	0.472
	R^2 test	0.360	0.993	0.561	0.868	0.929	0.482
GNN	Embedding dimension	32	32	64	32	32	64
	# conv. layers	3	2	3	3	3	3
	# hidden units	64	128	128	128	128	128
	K-hop neigh.	6	3	6	6	6	6
	R^2 validation	0.748	0.991	0.496	0.759	0.924	0.463
	R^2 test	0.815	0.993	0.445	0.763	0.906	0.468

Kerimov, Bulat, et al. "Assessing the performances and transferability of graph neural network metamodels for water distribution systems." *Journal of Hydroinformatics* 25.6 (2023): 2223-2234.