Digital Twins of Urban Drainage Systems: innovative data assimilation algorithm for continuous state update

Miloš Milašinović [1], Željko Vasilić [1], Damjan Ivetić[1]*

[1]*University of Belgrade, Department of Hydraulic and Environmental Engineering, Digital Water Engineering Lab, Belgrade, Serbia*

**Corresponding author email: mmilasinovic@grf.bg.ac.rs*

Abstract
Urban drainage systems (UDSs) are facing increasing challenges due to aging infrastructure and external factors, requiring innovative decision support solutions. Digital Twins (DT) of the real-world systems offer a promising decision-support tool for UDS management. These digital replicas, updated in real-time through sensor data integration, can assist with energy efficiency, resource allocation, and scenario analysis for contingency planning. DTs require up-to-date simulation model, and integration of the sensor data into the model is perceived as a critical component. This research introduces an innovative data assimilation method, utilizing Proportional-Integrative-Derivative (PID) controllers to update UDS model states based on sensor data. Proposed approach, tested with PySWMM on a synthetic dataset, demonstrates the potential for improving UDS performance and reducing uncertainty through continuous updates.

Highlights
- Digital Twins of UDSs represent viable decision support tool for urban water utilities
- Simulation model, along with sensor data, represent the essential component of Digital Twin
- PID controllers-based DA algorithm is proposed for continuous model state update

Introduction
Urban drainage infrastructure experience difficulties to keep desired performance levels due to aging systems, exacerbated by both natural and anthropogenic factors. To address these challenges and enhance the sustainability, safety, and resilience of urban drainage systems (UDSs), innovative management solutions and strategies are a must. Digital Twins (i.e. up-to-date virtual replicas) could be a viable decision support tool for water utilities. Creating a dynamically updated digital replica of the real-world system can help water utilities improve energy and resources allocation efficiency through real-time monitoring and modelling of the UDS. Furthermore, by creating a safe virtual environment, DTs allow assessing various "what-if" scenarios and inform contingency planning and operation decisions.
Creating DT of an UDS requires integrating sensor data from the system into the simulation model for better insights into the system dynamics and performance (*Bartos and Kerkez 2021; Kim et al. 2025; Pedersen et al. 2021*). To create a DT of an UDS, sensor network distributed across the system key components is necessary. Along with these sensors, appropriate algorithms for processing measured raw data (e.g. detecting and replacing data anomalies) and integrating them into the simulation model are the essence of a DT. Simulation model represents a mathematical approximation of the real-world process and is often affected by numerous uncertainty sources (e.g. uncertainty in model forcings, model parameters, model structure). To reduce the model uncertainty and improve the DT by continuous update of model's states (heads/flows) and possibly model parameters are required. This necessitates the application of data assimilation (DA) methods.
This paper presents innovative data assimilation method for UDS models, based on the simplified model states updating concept. Here, the integration of sensor data and model states update is conducted by inserting correction flows into the observation nodes. Proportional-Integrative-

Derivative (PID) controllers are used to estimate the correction flows based on the head measurements (*Milašinović et al. 2020*). These innovative data assimilation algorithm is coupled with PySWMM (*McDonnell et al. 2020*) on a synthetic dataset as a proof of concept.

Methodology

Urban Drainage Systems Digital Twins – role of data assimilation

Simulation models have a significant role in the Digital Twins concept. These models are affected by numerous uncertainty sources. For example, rainfall (used as model input/forcing) is often collected from distant rain gauges without considering its spatial and temporal variation. This can result in over- or underestimating hydraulic variables (heads/flows) representing system hydraulic state. When sensor data is collected across the system, its assimilation into the model can produce better estimate of the hydraulic variables in the system in (near) real-time. Eventually, this can provide better assessment of system performance and determine effective control and maintenance measures. This paper presents an algorithm capable for reducing model uncertainty caused by input data uncertainty. Here, the algorithm assumes that model uncertainty is dominantly caused by the unreliable inflow data (e.g. precipitation, inflows, or water usage by inhabitants). Consequently, the discrepancy between the simulated and measured water levels is treated as either a lack or an excess of water in the model. Hence, the problem is solved by adding or subtracting correction flows at the location where the water level measurements are available. This process is controlled by PID controllers (Figure 1).

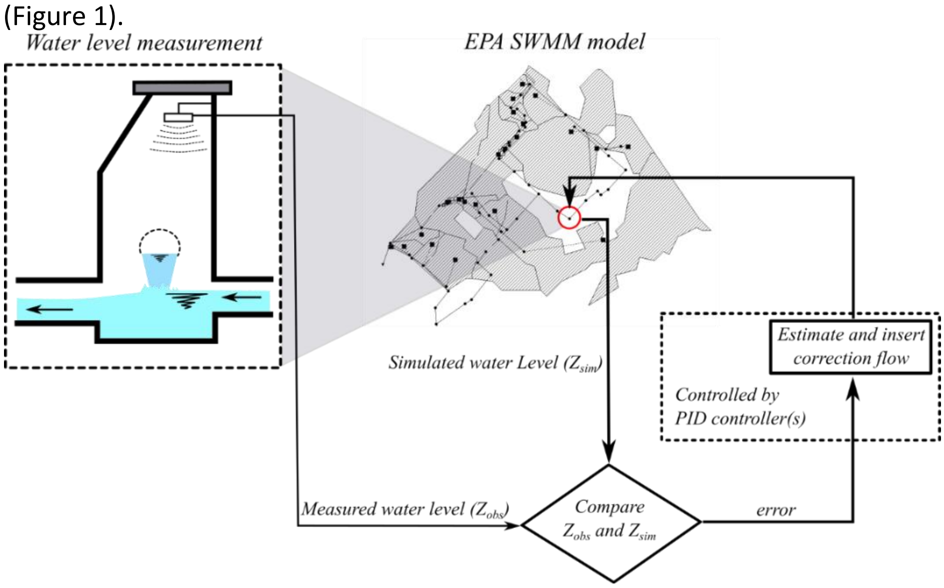

Figure 1. PID controller-based data assimilation algorithm for UDS Digital Twins - overview.

Continuous state estimation algorithm – PID controllers as data assimilation tool

PID controller is utilized to estimate the amount of the water that should be added or subtracted from the model at each simulation time step. It uses the process error *e(t)* as the input. This variable represents the difference between the setpoint which PID controller tends to reach (observed head - Z_{obs}) and model output (simulated head - Z_{sim}):

$$e(t) = Z_{obs}(t) - Z_{sim}(t) \quad (1)$$

Error variable is used to estimate the correction flow Q_{corr} using the PID controller's theory:

$$Q_{corr} = K_p \cdot e(t) + K_i \cdot \int_0^t e(t)dt + K_d \cdot \frac{de}{dt} \quad (2)$$

where K_p, K_i and K_d represent proportional, integrative, and derivative gain factors, respectively. This correction flow is inserted into the EPA SWMM model as the additional flow using the PySWMM software package.

Case study

Described methodology is tested on the part of New Belgrade UDS system, covering the highly urban area of 0.5 km^2 (Figure 2a). A synthetic test case is used to demonstrate the applicability of the proposed algorithm. First, true states (heads and flows) are generated using the observed rain event as the input data (Figure 2b). After the simulation, generated data (heads and flow) at observation locations are collected and considered as true ("observed") data (Figure 2c). A new simulation is then conducted with modified rain event (rainfall intensity scaled down to 50% of the original rain event and lagged for 30 minutes), in which data assimilation algorithm is used to compare the simulated and "observed" data and to insert/subtract correction flows at assimilation nodes (red dots on Figure 2). The effect of inserting the correction flows at assimilation nodes is evaluated at head validation nodes (yellow stars on Figure 2) and flow validation node (green dot).

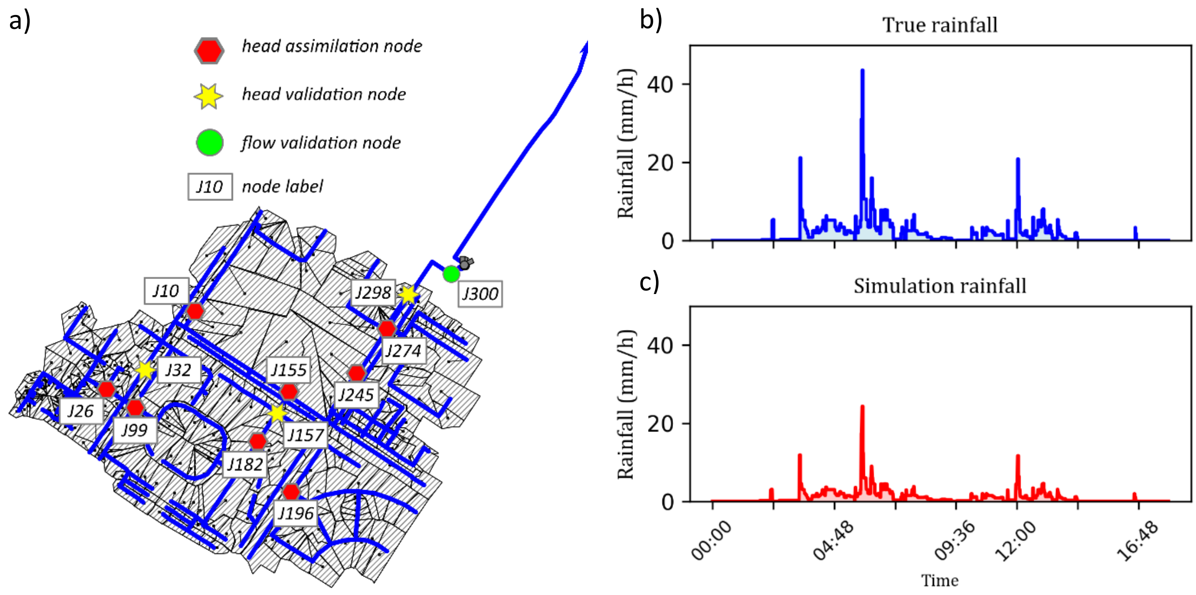

Figure 2. a) Case study: New Belgrade stormwater system, **b)** true rainfall event, and **c)** modified rainfall event

Results and discussion

True head timeseries (used to mimic the observed data) differ from those obtained by model free run (simulation without assimilation, both at head assimilation and validation locations (Figure 3a, b). This leads to discrepancy between the true and simulated flow timeseries at flow validation location (Figure 3c). To minimize this discrepancy and provide a better estimate of system's state, the proposed data assimilation (DA) algorithm is employed. PID controllers (hereinafter referred to as assimilators) are assigned to head assimilation locations to reduce the difference between the true and simulated data, represented as the process error (eq. 1) by inserting correction flows (eq. 2). Assimilators are manually tuned (K_p = 100, K_i = 100, K_d = 0) and applied to estimate correction flows at head assimilation locations. Tuning the assimilators should be carefully conducted, and special attention should be given to data assimilation sensitivity analysis to assimilators' parameters. This approach in data assimilation reduced the discrepancy between true and simulated heads at assimilation locations (Figure 3a) which led to improved estimate of hydraulic states across the system (head and flow validation locations, Figure 3b, c). These preliminary results demonstrate the high potential of the proposed DA algorithm for state estimation in UDS digital twins, even when the model forcings contain significant uncertainty.

This study analyses the idealistic scenario where the true values (i.e. measured values) are considered without uncertainty and the same temporal resolution is used for model simulation and measuring data. Real - world applications require further development and refinement of the proposed DA algorithm (e.g. considering measurement uncertainty, different timescales of the measurements and simulation, etc).

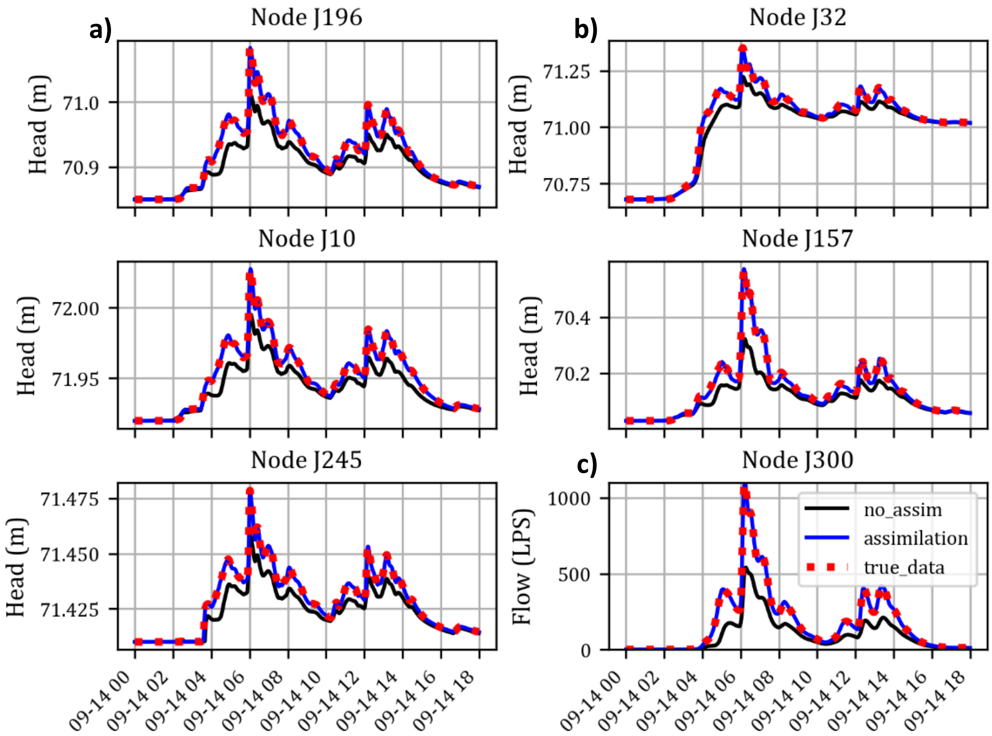

Figure 3. a) Head timeseries for some of the head assimilation nodes, b) head timeseries at head validation nodes, c) flow timeseries at flow validation node

Conclusions and future work

This research presents the DA algorithm customized for UDS models used for continuous model state update. The method uses PID controllers as assimilators to control the model state according to the measured data at specific locations. This provides valuable tool for digital twins of UDSs, where keeping the virtual replica of the real-world system up to date is one of the crucial requirements. Preliminary results show the promising potential of the proposed DA algorithm, but also emphasize the necessity for further investigations. To create a reliable digital twin, based on continuously updated simulation model, addressing other model uncertainties is needed. Hence, appropriate algorithms for regular update of model parameters and model structure should be developed and applied. Additionally, future sensor network is planned to be installed on the New Belgrade location and the proposed algorithm will be tested using the real-world data.

Acknowledgement

The research presented herein is part of the DIGIDRAIN project (funding from the Science Fund of the Republic of Serbia, research program DIASPORA 2023: joint research, under grant agreement number: 17823 – City-scale digital twins for urban drainage systems: bringing “smart” to water infrastructure).

References

Bartos, Matthew and Branko Kerkez. 2021. “Pipedream: An Interactive Digital Twin Model for Natural and Urban Drainage Systems.” *Environmental Modelling and Software* 144.

Kim, Yeji, Jeil Oh, and Matthew Bartos. 2025. “Stormwater Digital Twin with Online Quality Control Detects Urban Flood Hazards under Uncertainty.” *Sustainable Cities and Society* 118.

McDonnell, Bryant, Katherine Ratliff, Michael Tryby, Jennifer Wu, and Abhiram Mullapudi. 2020. “PySWMM: The Python Interface to Stormwater Management Model (SWMM).” *Journal of Open Source Software* 5(52):2292.

Milašinović, Miloš, Dušan Prodanović, Budo Zindović, Nikola Rosić, and Nikola Milivojević. 2020. “Fast Data Assimilation for Open Channel Hydrodynamic Models Using Control Theory Approach.” *Journal of Hydrology* 584(May 2020):124661.

Pedersen, Agnethe N., Morten Borup, Annette Brink-Kjær, Lasse E. Christiansen, and Peter S. Mikkelsen. 2021. “Living and Prototyping Digital Twins for Urban Water Systems: Towards Multi-Purpose Value Creation Using Models and Sensors.” *Water (Switzerland)* 13(5).