



HAL
open science

A digital twin testbed for real-time stormwater modeling with online quality control

Yeji Kim, Jeil Oh, Matthew Bartos

► **To cite this version:**

Yeji Kim, Jeil Oh, Matthew Bartos. A digital twin testbed for real-time stormwater modeling with online quality control. Novatech 2023, Graie, Jul 2023, Lyon, France. hal-04186329

HAL Id: hal-04186329

<https://hal.science/hal-04186329v1>

Submitted on 18 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A digital twin testbed for real-time stormwater modeling with online quality control

Un banc d'essai numérique jumeau pour la modélisation en temps réel des eaux pluviales avec contrôle qualité en ligne

Yeji Kim, Jeil Oh, and Matthew Bartos

The University of Texas at Austin (yejikim@utexas.edu; jeoh@utexas.edu; mdbartos@utexas.edu)

RÉSUMÉ

La surveillance et le contrôle en temps réel des systèmes de drainage urbain promettent de permettre une meilleure gestion des risques liés aux eaux pluviales comme les inondations et les débordements d'égouts unitaires. Pour atteindre ces objectifs, les gestionnaires de l'eau exigent de plus en plus des modèles jumeaux numériques basés sur les données qui utilisent des données de capteurs en continu et des modèles en ligne pour surveiller la dynamique des eaux pluviales en temps réel. Cependant, les modèles de jumeaux numériques pilotés par les données sont susceptibles d'erreurs causées par l'assimilation de données de capteurs bruitées. Les capteurs à faible coût, en particulier, souffrent d'un grand nombre de valeurs aberrantes qui peuvent provoquer de fausses alarmes dans les systèmes d'alerte précoce aux inondations. Pour combler ce manque de connaissances, cette étude se concentre sur le développement d'un modèle de jumeau numérique avec contrôle de la qualité des données en ligne pour la modélisation en temps réel des eaux pluviales à l'échelle du bassin versant. Tout d'abord, un réseau de capteurs sans fil est déployé dans le bassin versant de Waller Creek à Austin, au Texas, pour collecter des données en temps réel sur la profondeur de l'eau. Ensuite, un modèle de jumeau numérique de Waller Creek est créé à l'aide du solveur hydraulique PipeDream et est simulé en continu à l'aide des prévisions de précipitations. Enfin, un nouveau schéma de contrôle de la qualité des données est développé qui combine des données de capteur de profondeur, des estimations basées sur un modèle de la profondeur de l'eau et un nouvel algorithme de détection d'anomalies avec le filtrage de Kalman pour détecter et supprimer les défauts des capteurs. Ensemble, nous montrons comment ce système de modèle de jumeau numérique combiné peut être utilisé pour suivre avec précision l'état du système de drainage urbain en temps réel.

ABSTRACT

Real-time monitoring and control of urban drainage systems promises to enable improved management of stormwater hazards like floods and combined sewer overflows. To meet these goals, water managers are increasingly demanding data-driven digital twin models that use streaming sensor data and online models to monitor stormwater dynamics in real-time. However, data-driven digital twin models are susceptible to errors caused by assimilation of noisy sensor data. Low-cost sensors in particular suffer from large numbers of outliers which can cause false alarms in flood early warning systems. To address this knowledge gap, this study focuses on developing a digital twin model with online data quality control for real-time stormwater modeling at the watershed scale. First, a wireless sensor network is deployed in the Waller Creek watershed in Austin, TX to collect real-time water depth data. Next, a digital twin model of Waller Creek is created using the PipeDream hydraulic solver and is continuously simulated using precipitation forecasts. Finally, a novel data quality control scheme is developed that combines depth sensor data, model-based estimates of water depth, and a novel anomaly detection algorithm with Kalman Filtering to detect and remove sensor faults. Together, we show how this combined digital twin model system can be used to accurately track the state of the urban drainage system in real-time.

KEYWORDS

Anomaly detection, Digital Twin Model, Online quality control, Real-Time, Stormwater Modeling

1 INTRODUCTION

Stormwater modeling is widely used to understand, estimate, and manage water quantity and quality at the watershed scale (Wang et al. 1992). Classically, stormwater models have been used for infrastructure planning purposes—for instance, determining the sizes of pipes and reservoirs needed to mitigate flooding and reduce pollutant loads in natural water bodies. In recent years, however, *digital twin* models enabled by real-time sensor data have emerged as an important new tool for stormwater management. Briefly, digital twin models for urban watersheds are online virtual representations that simulate rainfall-runoff processes in real-time. Unlike traditional offline models used for infrastructure design, these models allow for improved real-time management of hazards—for instance, enabling online detection of flooding at low water crossings and facilitating more effective early emergency response.

Despite the transformative potential of digital twin models for stormwater management, important knowledge gaps remain surrounding their reliability for safety-critical applications. To facilitate smart stormwater management, many cities have installed large numbers of low-cost ultrasonic depth sensors and pressure transducers to monitor water levels in real-time (Zhang et al. 2019, Panagopoulos et al. 2021). However, these low-cost sensors suffer from large numbers of outliers caused by obstructing vegetation and water surface turbulence (Bae et al. 2019). These sensor faults cause problems for digital twin models that employ online data assimilation, leading to false alarms in flood early warning systems. To address these issues, this study explores how real-time quality control can be used to enable reliable digital twins of urban drainage systems. Towards this goal, we first deploy a pilot digital twin testbed consisting of a real-world sensor network deployment and an online hydrologic-hydraulic model. We then investigate the extent to which online quality control successfully detects and rejects sensor faults in streaming sensor data using measurements from a real-world storm event.

2 METHODS

This study describes the design and implementation of a real-world stormwater digital twin system, consisting of a wireless sensor network deployed within a real-world urban watershed and an online hydraulic model that simulates stormwater dynamics and assimilates quality-controlled sensor data in real-time. In this section, we describe (i) the deployment of the wireless sensor network, (ii) the implementation of the real-time hydrologic-hydraulic model, and (iii) an online quality control scheme used for data assimilation.

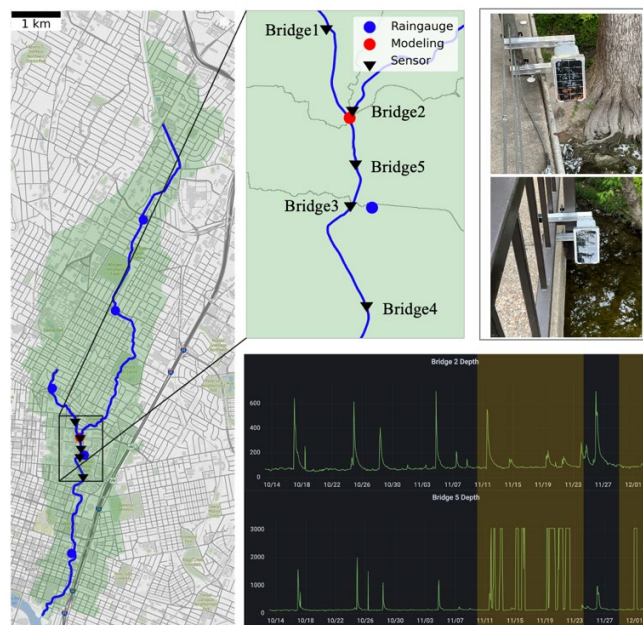


Figure 1. Map and data of wireless sensor network in Waller Creek watershed

2.1 Wireless sensor network

A wireless sensor network is deployed in the Waller Creek urban drainage system in Austin, TX to measure water levels and detect flooding. For the case study described in this paper, four ultrasonic depth sensors are concentrated near a flood-prone confluence point where high water-level peaks occur during rainfall. Figure 1

shows a map of the sensor locations in the drainage network (left), along with site photographs of the sensors (top-right), and a sample data stream from two sites (bottom-right). The architecture of the wireless sensor network is based on the open storm wireless sensing platform (Bartos et al., 2018), which consists of a low-power wireless sensor node and a cloud architecture for storing and processing real-time sensor data.

2.2 Watershed digital twin model

A digital twin model is developed to simulate stormwater dynamics in real-time and track the current state of the stormwater system. The digital twin model consists of (i) a hydrologic model that uses precipitation forecasts along with an infiltration model to compute a runoff hydrograph, and (ii) a hydraulic model based on the Saint-Venant equations that routes surface runoff through the channel network to estimate flood depths.

2.2.1 Hydrologic model

Modeling of infiltration and runoff processes is accomplished using the Soil Conservation Service (SCS) Curve Number (CN) method (NRCS et al. 2004), which estimates precipitation excess from storm events:

$$P_e = \frac{(P - I_a)^2}{(P - I_a) + S}, \quad \text{For } P > I_a$$

Where P_e is the excess precipitation (in), P is the cumulative precipitation (in), I_a are initial abstractions, and S is the total available infiltration capacity (in). The curve number (CN) is used to characterize the infiltration capacity of a particular soil and is calculated based on pervious CN (natural surface) and the percentage of impervious area (Rawls et al. 1981). Finally, the direct runoff hydrograph is derived through discrete convolution of the precipitation excess with subcatchment-specific unit hydrographs.

2.2.2 Hydraulic model

The hydraulic model of the urban drainage system is implemented using the PipeDream solver, previously developed by Bartos et al. (2021). This software solves the Saint-Venant equations for unsteady flow in channel networks using a stable implicit discretization scheme that is suitable for online simulations.

2.2.3 Online quality control

An online quality control system is developed to accurately track the state of the urban drainage system by combining raw sensor data with simulated water depths predicted by the digital twin model. As shown in Fig. 1 (bottom-right), raw sensor data is noisy, and exhibits frequent sensor spikes. In a real-time context, these spikes are often indistinguishable from peaks caused by the arrival of a flood wave, which may cause false alarms unless these two events are properly classified. Thus, for effective emergency management of flood events, it is essential that spurious sensor faults are promptly detected and flagged.

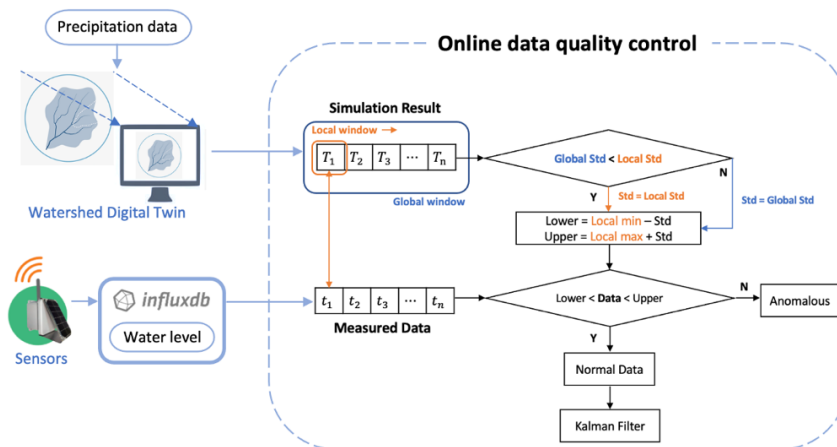


Figure 2. Online quality control architecture diagram

The new online quality control system detects faults by comparing the modeled and measured estimates of the water level together using a rolling filter to isolate sensor data points that deviate significantly from the modeled results. Details of the quality control procedure are described in Fig. 2. First, the digital twin is

simulated using rainfall forecasts to produce modeled estimates of water depths in the channel network. Next, a rolling window is passed over the modeled data to determine local and global outlier thresholds based on the sampled standard deviation. Streaming sensor data is then compared against these moving thresholds and outlier points are rejected. Finally, refined inlier data are fused into simulation results via a Kalman Filter to keep the modeled and measured states mutually consistent.

3 RESULTS

The digital twin model with online quality control successfully distinguishes sensor faults from real hydrograph peaks in a real-world test case. Here, the performance of the digital twin model with online quality control is evaluated against a single storm event on October 25th, 2022. Fig. 3 shows the results of the online quality control procedure without data assimilation. Here, the quality control procedure detects all outliers in the sensor data, but also flags several inlier points near the beginning of the storm event.

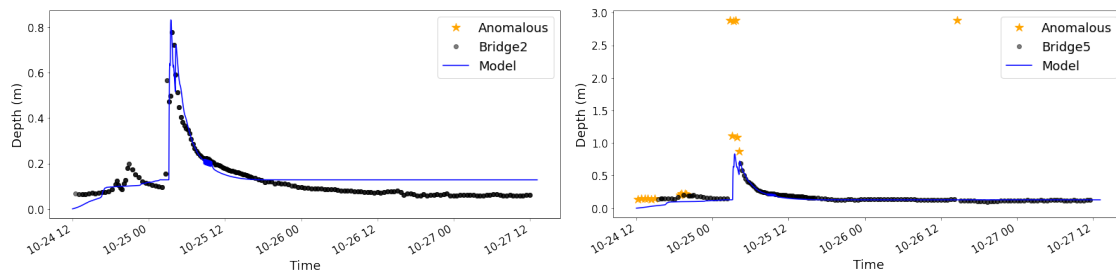


Figure 3. Anomaly detection with semi-supervised algorithm. (Left): Results of quality control after removing outliers. (Right): Results of quality control with outliers shown.

Fusing sensor data into the model via a Kalman Filter helps to address false negatives by keeping the modeled and measured states consistent. Fig. 4 (left) shows the results of fusing sensor data into the model after quality control using a Kalman Filter. Here, the filtered results (red) show nearly perfect agreement with the non-outlier sensor data. On the other hand, without anomaly detection, Kalman filtering can fuse sensor faults to raw model result, leading to incorrect predictions (Fig. 4, right). Thus, for real-time digital twin models of urban drainage networks, it is important to combine data assimilation with preliminary quality control routines to ensure accurate tracking of hydraulic states like depth and discharge and enable effective response to flood events.

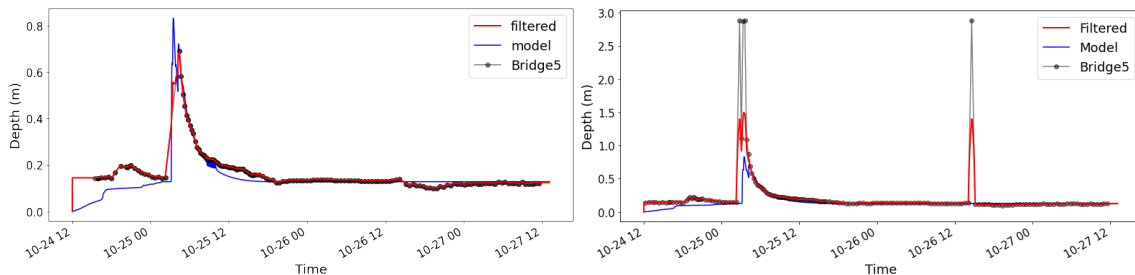


Figure 4. Kalman filtered results with (left) and without (right) anomaly detection.

LIST OF REFERENCES

- Bae, I. and Ji, U., 2019. Outlier detection and smoothing process for water level data measured by ultrasonic sensor in stream flows. *Water*, 11(5), p.951.
- Bartos, M. and Kerkez, B., 2021. Pipedream: An interactive digital twin model for natural and urban drainage systems. *Environmental Modelling & Software*, 144, p.105120.
- Bartos, M., Wong, B., and Kerkez, B., 2018. Open Storm: a complete framework for sensing and control of urban watersheds. *Environmental Science: Water Research & Technology*, 4(3), 346-358.
- Panagopoulos, Y., Papadopoulos, A., Poulis, G., Nikiforakis, E. and Dimitriou, E., 2021. Assessment of an Ultrasonic Water Stage Monitoring Sensor Operating in an Urban Stream. *Sensors*, 21(14), p.4689.
- Wang, G.T., Singh, V.P. and Yu, F.X., 1992. A rainfall-runoff model for small watersheds. *Journal of Hydrology*, 138(1), 97-117.
- Zhang, D., Heery, B., O'Neil, M., Little, S., O'Connor, N.E. and Regan, F., 2019. A low-cost smart sensor network for catchment monitoring. *Sensors*, 19(10), p.2278.