Integrated Real-Time Monitoring of Sewer Systems via SPC and Graph-Based Validation

Jun Li[1] and Finia Buitkamp[2]

[1]Geodetic Institute and Chair of Computing in Civil Engineering & GIS (gia), Aachen, Germany

[2]Wasser-Abwasser Systemtechnik GmbH (W.A.S.)

E-mail(s): jun.li@gia.rwth-aachen.de, buitkamp@was-bs.de

Abstract: Urban sewer systems face growing challenges due to aging infrastructure, climate-induced stress, and sensor unreliability. This paper presents a hybrid framework that combines Statistical Process Control (SPC) with graph-based spatial validation to monitor sewer networks in real time. Temporal anomalies are identified through SPC techniques - including Western Electric rules, CUSUM, EWMA - applied to streaming sensor data. To enhance detection performance under varying hydrological conditions, the framework incorporates weather-based classification prior to SPC computation. Simultaneously, flow consistency across the network is validated using a directed graph model, detecting structural inconsistencies and sensor faults. Experimental results using real-world data demonstrate that SPC methods capture local anomalies effectively, while spatial validation uncovers topological violations that temporal methods may miss. The combination of these methods enables robust, real-time detection of flow disturbances and supports infrastructure monitoring in smart sewer systems.

Keywords: real-time monitoring, sewer networks, anomaly detection, SPC, spatial validation

DOI: 10.18154/RWTH-CONV-254896. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,

1 Introduction

Urban drainage systems are increasingly stressed by aging infrastructure and climate-induced extremes. In Germany, annual precipitation has risen by approximately 7% since 1881, and the frequency of localized heavy rainfall events has significantly increased [1], [2]. Combined sewer systems (CSS), still prevalent in many cities, are often overwhelmed by stormwater during intense rainfall, resulting in untreated combined sewer overflows (CSOs) into surface waters [3].

Beyond hydraulic capacity issues, modern urban sewer systems rely on a dense sensor network to monitor flow rates, water levels, and other hydrodynamic parameters. However, these sensors are prone to noise, drift, and measurement bias [4], [5]. The resulting data quality challenges complicate early-warning detection tasks. Traditional anomaly detection in such systems still involves labor-intensive visual inspections or heuristic thresholds, limiting the scalability of real-time monitoring [6], [7].

To address these challenges, data-driven methods such as Statistical Process Control (SPC) and graph-based validation frameworks are gaining attention for their ability to support automated, scalable, and interpretable anomaly detection [5], [8].

The research project "Kanalanalyse", a collaboration between xxx and xxx, addresses this gap by integrating real-time sensor data with historical and simulated records. This paper presents two key modules of the system:

(1) A temporal anomaly detection component using Statistical Process Control (SPC) methods, structured in two phases - baseline modeling and real-time monitoring - applied to sewer flow and level data;

(2) A spatial consistency check based on directed graph topology, which validates flow monotonicity to detect implausible measurements and sensor faults.

Recent studies emphasize the potential of data-driven methods to enhance operational awareness in sewer networks, particularly under dynamic weather conditions [5]. To support this, rainfall events are further characterized based on their temporal structure and intensity, enabling the differentiation between convective and stratiform patterns. These contextual insights guide the interpretation of flow anomalies, allowing the system to distinguish between hydraulic responses to actual events and deviations caused by sensor faults.

Together, these methods enable a scalable, real-time monitoring framework that supports early warning and improves data quality across urban drainage networks.

2 Spatial Validation of Flow Consistency

Urban sewer systems can be broadly categorized into separate and combined systems. In combined systems, both sanitary sewage and stormwater runoff are conveyed in the same pipes, leading to potential overflows during heavy rain events.

Validating flow consistency requires understanding flow direction and upstream-downstream relations. We model the sewer network as a directed graph and check spatial consistency via flow monotonicity constraints.

2.1 Graph-Based Modeling and Flow Monotonicity Validation

Sewer networks with junctions, pipes, and manholes map naturally to directed graphs $G = (V, E)$, where V are junctions or sensor locations and E are conduits aligned with flow direction. This structure supports efficient modeling of hydraulic dependencies and topology-aware analysis.

Directed graphs have been widely used in sewer system design and monitoring. For optimization, they support layout simplification and hydraulic simulations [9], while in monitoring applications, they capture flow dependencies for spatio-temporal analysis [4], [10].

Building on this structure, we introduce a physically motivated validation method based on flow monotonicity: under normal conditions, the cumulative flow from upstream nodes should not exceed that at the downstream node, aside from minor deviations due to measurement noise or unmodeled inflows. Formally, for a downstream node v and its set of upstream neighbors $\mathcal{U}(v)$, we expect:

$$\sum_{u \in \mathcal{U}(v)} Q_u(t) \leq Q_v(t) + \epsilon, \tag{1}$$

where $Q_u(t)$ and $Q_v(t)$ are flow measurements and ϵ accounts for sensor uncertainty and spatial delays. Systematic violations of this inequality may indicate sensor faults, data desynchronization, undocumented inflows, or structural inconsistencies. We apply this constraint across all connected node pairs to validate the spatial coherence of sensor data in real-time.

2.2 Algorithm: Upstream–Downstream Consistency Check

Based on the monotonicity constraint introduced in Section 2, we implement a spatial validation procedure that systematically checks for flow inconsistencies across upstream-downstream sensor pairs in the sewer network.

Let $G = (V, E)$ be the directed graph representing the sewer network, where each node $v \in V$ is associated with a time series $Q_v(t)$ denoting the measured flow at time t. For each node v, let $\mathcal{U}(v)$ denote the set of its direct upstream neighbors (i.e., $(u, v) \in E$). For each time step t, the algorithm performs the following:

1. For each node $v \in V$ with incoming edges:
   (a) Retrieve $Q_v(t)$ and $\{Q_u(t) \mid u \in \mathcal{U}(v)\}$.
   (b) Compute the difference $\Delta_v(t) = Q_v(t) - \sum_{u \in \mathcal{U}(v)} Q_u(t)$.
   (c) If $\Delta_v(t) < -\epsilon$, flag this as a monotonicity violation.
2. Store all violations $(t, u, v, \Delta_v(t))$ for reporting or visualization.

The tolerance term ϵ is set as a fixed violation rate of 10% of the downstream flow at time t, i.e., $\epsilon = 0.1 \cdot Q_v(t)$. This accounts for measurement uncertainty, short-term flow fluctuations, and minor unrecorded inflows.

To facilitate practical application, the consistency check algorithm has been integrated into a stream-processing pipeline and deployed as part of a real-time monitoring system developed in the research project "Kanalanalyse". Detected violations are streamed to a web-based dashboard, enabling operators to visually inspect the spatial distribution and temporal evolution of flow anomalies in the sewer network.

3 Statistical Process Control for Sewer Flow Monitoring

Sewer systems are increasingly equipped with flow and level sensors that enable continuous monitoring. However, interpreting this time series data remains challenging due to its susceptibility to seasonal variations, rainfall events, sensor drift, and infrastructure anomalies. Statistical Process Control (SPC) offers a framework originally developed for industrial quality control that has proven effective in environmental and infrastructure monitoring [8], [11].

SPC distinguishes between normal process variability and deviations that may indicate underlying issues. In the context of wastewater networks, SPC has been applied for detecting anomalies such as blockages, overflows, and sensor failures [6]. It has also been used for burst detection in water distribution systems [11], flood risk index construction based on control limits [12], and overflow behavior analysis under different rainfall intensities [13]. Its application to sewer monitoring is motivated by its interpretability, computational efficiency, and suitability for real-time deployment.

3.1 SPC Methods

SPC methods rely on control charts to distinguish random variability from significant process deviations. Each chart consists of a center line representing the process average and upper/lower control limits that define acceptable fluctuation ranges. As illustrated in 1, When observations fall outside the control limits, they are flagged as potential anomalies. The example shows a Shewhart control chart with center line, upper control limit (UCL), and lower control limit (LCL), highlighting out-of-control points.

It is been proven that 6 sigma is a very effective framework for data quality control [14] This range captures approximately 99.73% of all data points under normal operation, thus providing a statistically justified threshold for anomaly detection [8].

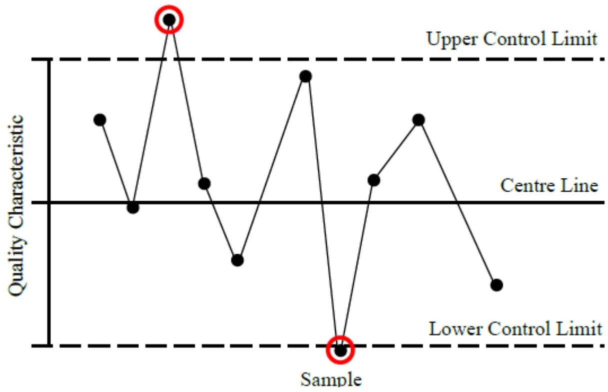

Figure 1: Example Shewhart Control Chart showing centre line, UCL, and LCL. Out-of-control points are highlighted.

Among SPC techniques, **the Western Electric Company (WEC) rules** represent a foundational approach. The first and most basic rule states that a process is considered out of control if any single data point falls outside the three-sigma range $[\mu - 3\sigma, \mu + 3\sigma]$. In this study, we adopt this rule by directly setting the control chart bounds to this interval. This provides a balance between sensitivity and false alarm rate in high-frequency environmental monitoring scenarios, and serves as a baseline against which more advanced methods (e.g., CUSUM and EWMA) are later compared.

EWMA (Exponentially Weighted Moving Average): This method assigns more weight to recent observations, making it effective for detecting gradual trends or small drifts in the data [8]. The control statistic is recursively updated as $z_t = \lambda x_t + (1-\lambda)z_{t-1}$, where $\lambda \in (0,1]$ is the smoothing factor. Control limits are calculated as $\mu \pm L \cdot \sigma\sqrt{\frac{\lambda}{2-\lambda} \cdot (1-(1-\lambda)^{2t})}$, based on the segment-specific mean μ and standard deviation σ.

CUSUM (Cumulative Sum Control Chart): A sequential analysis technique designed for identifying small, sustained shifts in the process mean [11]. It accumulates deviations from the target mean over time and flags anomalies when the cumulative statistic crosses a decision interval. In our implementation, the CUSUM statistic is computed as the cumulative sum of deviations from the segment-specific mean μ.

3.2 SPC Phases and Implementation

Following standard SPC methodology [8], our framework comprises two phases:

Phase I — Baseline modeling: Historical data under stable conditions is used to compute process statistics - mean μ, standard deviation σ. Control limits (e.g., $\mu \pm 3\sigma$) are defined per method. Periods with anomalies or strong seasonality are excluded.

Phase II — Online monitoring: Incoming sensor data is checked against these limits. Violations - such as boundary crossings or sensitizing patterns (e.g., WEC rules) - trigger anomaly flags.

We implemented this two-phase approach in a real-time stream processing setup. Sensor data is transmitted using MQTT (Message Queuing Telemetry Transport), a lightweight publish-subscribe protocol optimized for reliable delivery in low-bandwidth environments.

The processing pipeline includes:

- **Preprocessing:** hourly resampling, gap filling, and outlier filtering.
- **Dynamic SPC:** incremental updates of CUSUM and EWMA per new data point.
- **Alerting:** a 10% violation threshold within a time window triggers alerts, filtering transient noise.

This architecture enables scalable integration of SPC into live sewer monitoring systems.

3.3 Experimental Setup

This section investigates whether SPC methods can reliably detect process anomalies in wastewater level data and how weather classification affects their interpretability. We evaluate three univariate SPC methods: WEC rules, EWMA, and CUSUM. For EWMA, we adopt Montgomery's recommended setting of a smoothing factor $\lambda = 0.2$ and control width $L = 3$, which balances sensitivity and false alarms [14]. [14] notes that $\lambda \in [0.05, 0.25]$ is commonly used, with $L = 3$ working well, especially for larger λ. For CUSUM, we use the standard decision interval $H = 5\sigma$, which reliably detects moderate persistent shifts with low false-alarm rates [14]. These parameter choices follow established practice and ensure fair comparability across methods.

Since rainfall strongly affects sewer behavior, especially in runoff-sensitive lines, we follow prior studies [5], [6], [13] in applying weather classification before SPC. Rainfall introduces transient variability that violates the stationarity assumptions of SPC. To address this, we divide the monitoring period into three meteorologically homogeneous segments based on a practical rainfall threshold of 0.3 mm per 15-minute interval: a dry weather period, a rain-affected period, and a post-event recovery phase.

To illustrate this need, we first applied SPC methods to the unsegmented series. When SPC methods were applied directly to the full time series, both WEC and EWMA exhibited strong responsiveness to rainfall dynamics (Figure 2 b)-d)). Each alarmed segment corresponded precisely to precipitation peaks (Figure 2 a)), suggesting that these methods are well-suited for early warning of potential overflow risks under storm conditions. In contrast, CUSUM produced ambiguous results. Since the full monitoring window spans dry, rainy, and post-event conditions, the use of a single global mean and

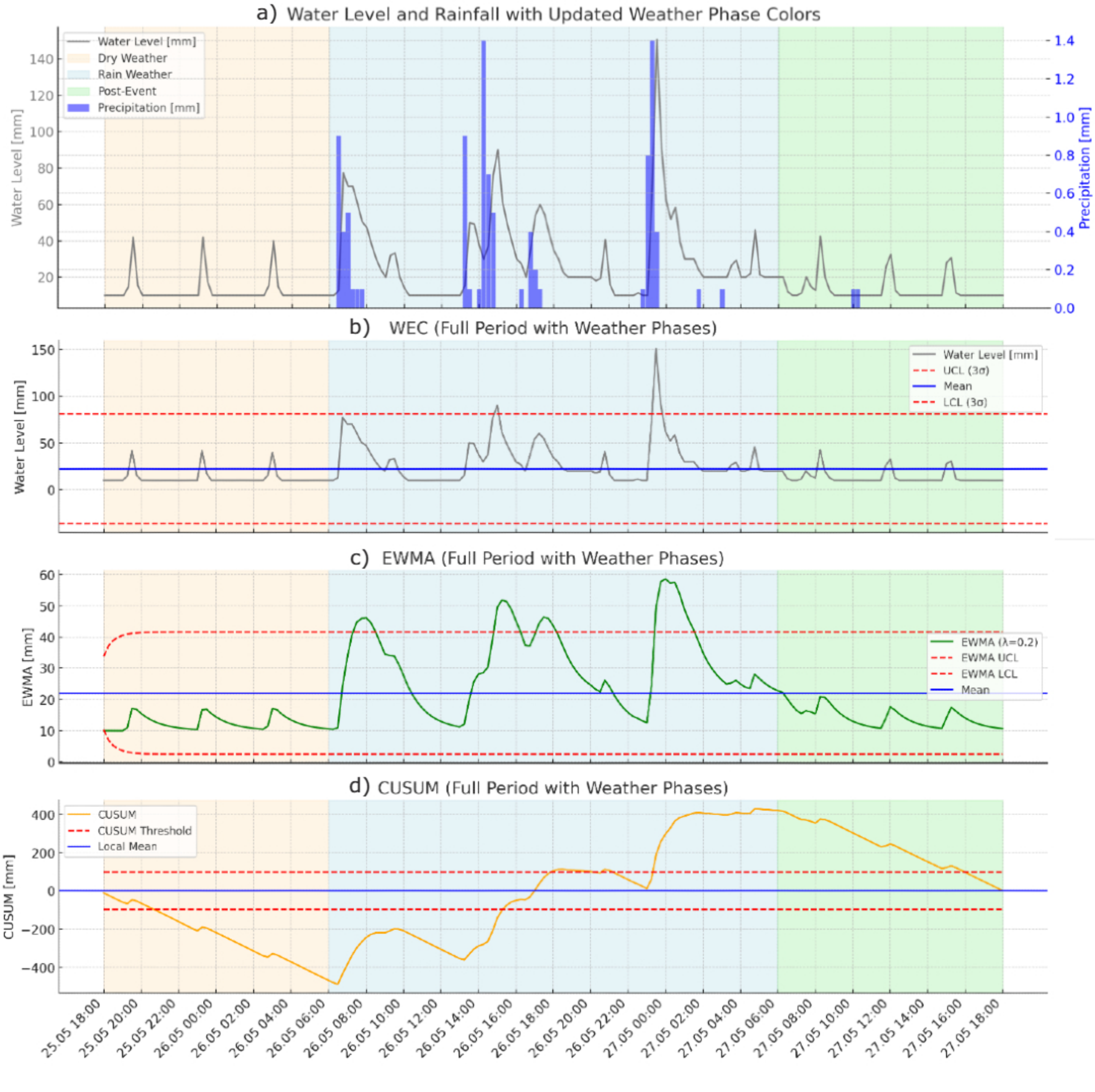

Figure 2: Rainfall with annotated weather phases and SPC output without weather classification.

standard deviation distorted its anomaly detection capability. CUSUM failed to distinguish between genuine system anomalies and statistical shifts caused by changing weather states. This highlights that, in order to accurately detect localized system anomalies, it is essential to first perform weather classification to eliminate weather-induced variability.

We then reapplied the SPC methods to the segmented data, distinguishing between dry weather and rain-affected periods to represent baseline and disturbed system states. The results confirmed statistical stability for all SPC methods, with no control limit violations detected during dry weather days. In contrast, significant deviations associated with precipitation events were identified in the rain affected segments. Notably, the CUSUM method, as illustrated in Figure 3, effectively addressed the issue of false cumulative signals caused by precipitation, which previously led to misinterpretation. It also delivered performance comparable to the other two methods. These findings demonstrate that applying weather classification prior to SPC significantly enhances interpretability and detection accuracy in real world sewer monitoring scenarios. Although our evaluation focuses on a single measurement site, the proposed stratified SPC approach is readily transferable to other locations within urban drainage networks where flow behavior is sensitive to weather conditions.

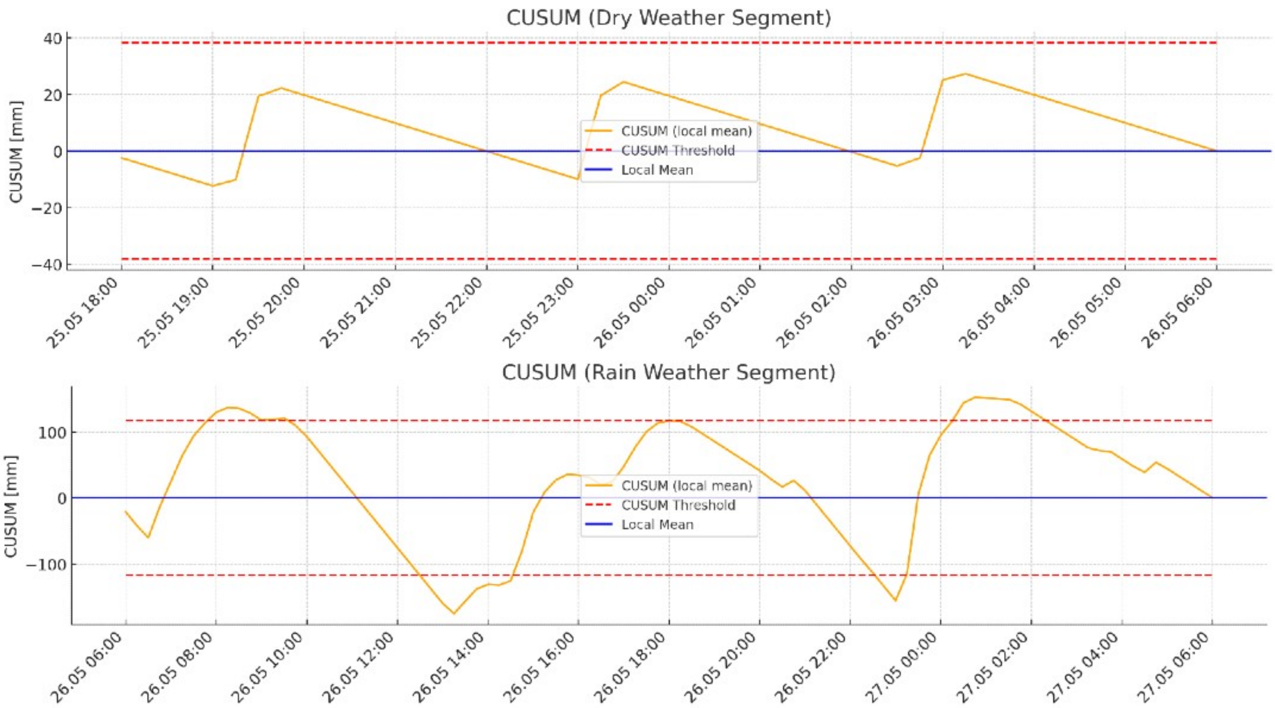

Figure 3: For the dry weather segment, the CUSUM method indicates statistical control; for the rain affected segment, it detects significant deviations associated with precipitation events.

4 Summary and Conclusions

We evaluated three SPC methods for temporal anomaly detection in sewer monitoring and found that applying them to unsegmented time series often produced misleading results, especially for CUSUM. In contrast, WEC and EWMA remained reliable under global conditions. When rainfall-based segmentation was introduced, all methods, including CUSUM, showed improved interpretability and localization of anomalies. Complementing temporal analysis, we introduced a spatial validation approach based on flow-directed graph topology to detect violations of flow monotonicity. These violations can indicate backwater effects, undocumented inflows, or potential sensor inconsistencies - capturing anomalies that temporal methods alone might miss. Together, our findings highlight the necessity of weather classification and spatial validation as preprocessing steps for robust sewer monitoring. SPC should serve as a fast, interpretable pre-alarm system, while further diagnosis should involve physical modeling. Future work will explore adaptive time windows and integrate SPC with hydraulic simulations in a multi-stage monitoring architecture.

References

[1] IPCC, *Climate change 2022: Impacts, adaptation and vulnerability. working group ii contribution to the sixth assessment report*, Online, 2022.

[2] Umweltbundesamt, *Klimawandel und urbane sturzfluten: Risiken für städte*, Online, 2022.

[3] U.S. Environmental Protection Agency, "Report to congress on combined sewer overflows to the great lakes basin", U.S. EPA, Washington, DC, Tech. Rep., 2016.

[4] L. He et al., "Spatiotemporal graph convolutional network using sparse monitoring data for accurate water-level reconstruction in urban drainage systems", *Journal of Hydrology*, vol. 652, p. 132 681, 2025.

[5] S. C. Troutman, N. Schambach, N. G. Love, and B. Kerkez, "An automated toolchain for the data-driven and dynamical modeling of combined sewer systems", *Water Research*, vol. 126, pp. 88–100, 2017.

[6] T. Rosin, *Data Analytics for Automated Near Real Time Detection of Blockages in Smart Wastewater Systems*. University of Exeter (United Kingdom), 2020.

[7] V. Edmondson, M. Cerny, M. Lim, B. Gledson, S. Lockley, and J. Woodward, "A smart sewer asset information model to enable an 'internet of things' for operational wastewater management", *Automation in Construction*, vol. 91, pp. 193–205, 2018.

[8] G. J. da Silva and A. C. Borges, "Statistical process control in the environmental monitoring of water quality and wastewaters: A review", *Water*, vol. 17, no. 9, p. 1281, 2025.

[9] N. Duque, D. Duque, A. Aguilar, and J. Saldarriaga, "Sewer network layout selection and hydraulic design using a mathematical optimization framework", *Water*, vol. 12, no. 12, p. 3337, 2020.

[10] D. Meijer, M. Van Bijnen, J. Langeveld, H. Korving, J. Post, and F. Clemens, "Identifying critical elements in sewer networks using graph-theory", *Water*, vol. 10, no. 2, p. 136, 2018.

[11] D. Jung, D. Kang, J. Liu, and K. Lansey, "Improving the rapidity of responses to pipe burst in water distribution systems: A comparison of statistical process control methods", *Journal of Hydroinformatics*, vol. 17, no. 2, pp. 307–328, 2015.

[12] A. S. M. Saudi, H. Juahir, A. Azid, and F. Azaman, "Flood risk index assessment in johor river basin", *Malaysian Journal of Analytical Sciences*, vol. 19, no. 5, pp. 991–1000, 2015.

[13] R. Ibrahim, B. Saâdia, and R. Mohamed, “Comparison between different intensities of rainfall to identify overflow points in a combined sewer system using storm water management model.”, *AIMS Environmental Science*, vol. 9, no. 5, 2022.

[14] D. C. Montgomery, *Introduction to statistical quality control*. John wiley & sons, 2020.