SUPPORTING BETTER PUBLIC RESPONSE BY ENHANCING WASTEWATER MONITORING WITH MACHINE LEARNING

Nasrine Tomasi (Mott MacDonald), Chris Park (Mott MacDonald), Nathan Donald (Watercare)

ABSTRACT

Wastewater overflows can have a significant impact on the environment. During these overflows, sewage is released into nearby water bodies which can lead to several environmental problems and can affect human health. For years, water utilities have invested in improving the management of wastewater overflows by establishing long term monitoring plans of their constructed sewer overflow structures. These monitoring programmes alert operators of overflows occurring with the goal of better prioritising resources on the ground.

This paper outlines a recent trial undertaken to enhance wastewater monitoring datasets with the latest technologies in artificial intelligence (AI) and support operators with the early identification of potential blockages, forecasted overflows and trends in inflow/infiltration at a catchment level. Here are some ways AI and machine learning (ML) have been used for these purposes:

- Wastewater blockage detection: AI can be used to analyse data from sensors that measure sewer depth and velocity in wastewater pipes. By analysing changes in patterns in this data, AI can identify potential signs of blockages and alert operators when an inconsistency in the data is identified. This can help prevent minor or partial blockages from escalating to overflows.

- Wet weather overflow forecast: AI can be used to analyse weather data in conjunction with sewer depth to create a model that predicts when and where wet weather events are likely to occur based on available rain forecasts. By combining this with the historical performance of the system during wet weather events, AI can forecast when and where overflows are likely to occur. This can help operators prepare for and respond to overflows more effectively.
- Inflow infiltration characterisation: Statistical models can be used to compare dry weather data with wet weather profiles to identify sources of inflow and infiltration (such as leaky pipes or illegal connections) and estimate their magnitude. Undertaking this analysis in near-real-time can help operators identify degrading performance and condition of the asset to support the prioritisation of repairs, renewals and maintenance to reduce the impact of inflow and infiltration on the system.

Overall, AI, ML and statistical models can be powerful tools for optimising the operation and maintenance of wastewater systems and reducing the impact of wastewater on the environment. It can enhance traditional wastewater monitoring to proactively manage network issues before seeing their impact on the receiving environment and communities.

KEYWORDS

Wastewater, overflows, artificial intelligence, machine learning, operation

PRESENTER PROFILE

Nasrine Tomasi is Mott MacDonald's Smart Water Lead. She drives innovation through water projects and has been instrumental in the deployment of award-winning digital twins. Nasrine strongly believes in the power of data to deliver better insights for decision makers in the water sector.

Chris Park leads Mott MacDonald's Auckland based Network Insights team. Chris' career is founded on experience in wastewater flow monitoring, telemetered monitoring technology, data analysis, and digital solutions to enable data driven infrastructure planning, and levels of service improvement.

INTRODUCTION

Wastewater overflows represent a critical challenge in modern urban infrastructure and environmental management. These incidents occur when the capacity of wastewater systems is exceeded, resulting in the release of sewage into the environment. Sewer overflows can be related to multiple factors:

- Inflow and infiltration: Inflow and infiltration are generally seen in older, urbanised and densely populated areas where unwanted stormwater enters the system. Inflow and infiltration mostly affect combined sewer infrastructure but are also observed in aging or deteriorated separate systems. This means that during dry periods and insignificant wet weather, the systems generally route all sewer flow to the treatment facilities, yet, when wet weather increases flows, the collection and conveyance systems are surcharging or overflowing.
- Temporary blockages: Blockages can also cause capacity issues and overload sewer systems resulting in dry weather overflows and the discharge of untreated and non-diluted sewer effluent into the receiving environment.

Wastewater overflows can have far-reaching consequences, affecting public health, water quality, and the integrity of ecosystems. As urbanisation continues to accelerate and the impact of climate change results in more frequent and more intense rainfall events, the risk and recurrence of sewer overflows are set to increase.

For these reasons, and encouraged by more public awareness around this issue, water utilities in New Zealand and around the globe have been investing in the deployment of more IoT sensors capable of detecting overflows and/or measuring sewer depth and velocity. These sensors are generally linked to an alarming system set to inform operations teams of an ongoing overflow event, the data collected is also used for reporting purposes and increasingly being shared with the public (through platforms such as www.safeswim.org.nz in New Zealand or www.thameswater.co.uk/edm-map in the United Kingdom).

While the process of monitoring critical overflow locations is not new, there has been a clear increase in the number of sensors being deployed in recent years. The volume of data collected and recent advancement in data analysis technologies mean that enhancing wastewater monitoring with machine learning has the potential to significantly improve public response and overall environmental management. This approach can lead to more effective and timely interventions. This paper details the outcome of a recent trial undertaken to enhance wastewater monitoring datasets focusing on wastewater blockage detection, wet weather overflow forecast and inflow and infiltration characterisation.

TRADITIONAL APPROACH

PREDICTION OF WET WEATHER OVERFLOWS

Traditionally, hydraulic models have been used to predict wet weather overflows. A hydraulic model is a physical or mathematical representation of a hydraulic system. These models consider multiple aspects of the drainage network including asset characteristics, along with their structure, topology, and other determining factors such as flows from other parts of the network and rainfall.

The interaction between such elements is then consolidated into public or proprietary software packages, which are then used to simulate flow under different conditions (Hedges, 1994; Carstensen et al., 1998; Duchesne et al., 2001; Politano et al., 2007; Grum et al., 2011; Garofalo et al., 2017; Morales et al., 2017).

Hydraulic models are considered closed-form solutions, which are transparent and interpretable (Thorndahl et al., 2008; Szeląg et al., 2018). In addition, with the mechanics clearly defined from the beginning, they do not require abundant data to calibrate. However, by relying on a static framework, these models become less flexible when the context or the set of variables changes (e.g., population growth, GIS updates, new infrastructure, etc). Finally, these models are computationally intensive, and thus, hard to deploy and maintain (Osiadacz, 1996; Kochevsky & Nenya, 2004). Their long run time also makes real-time predictions of overflow difficult, which explains why these models are mostly used for planning and optioneering purposes.

DETECTION OF BLOCKAGES IN SEWER SYSTEMS

Sewer blockages can occur for various reasons, and often happen unexpectedly. Some common causes include the buildup of foreign objects in the sewer pipelines such as grease and fat, sanitary products or sediments, tree roots growth, or debris related by collapsed pipes.

Water utilities can take several proactive measures to identify sewer blockages and ensure the smooth operation of their wastewater systems. In particular, CCTV inspections play an important role in regularly assessing the condition of sewer lines and associated infrastructure. Monitoring systems are also implemented to help identify blockages at critical sites and track sewer system performance. Real-time monitoring is traditionally used to send alarms to sewer operators when sewer depth exceeds a specific threshold, indicating increased flows that are often linked to network constraints during dry conditions.

INFLOW AND INFILTRATION CHARACTERISATION

Accurately quantifying and characterising inflow and infiltration (I&I) is crucial for the proper management and maintenance of wastewater systems and directly impacts the functionality, efficiency, and cost-effectiveness of sewer networks by limiting wastewater flow, especially during wet weather events.

I&I studies typically consist of an in-depth analysis of monitoring data. Data collected from flow monitoring is used to identify an average dry weather flow (ADWF) pattern. I&I is estimated by subtracting this dry weather flow hydrograph from the recorded wet weather flow hydrograph (Water New Zealand infiltration & Inflow Control Manual 2nd Edition, March 2015).

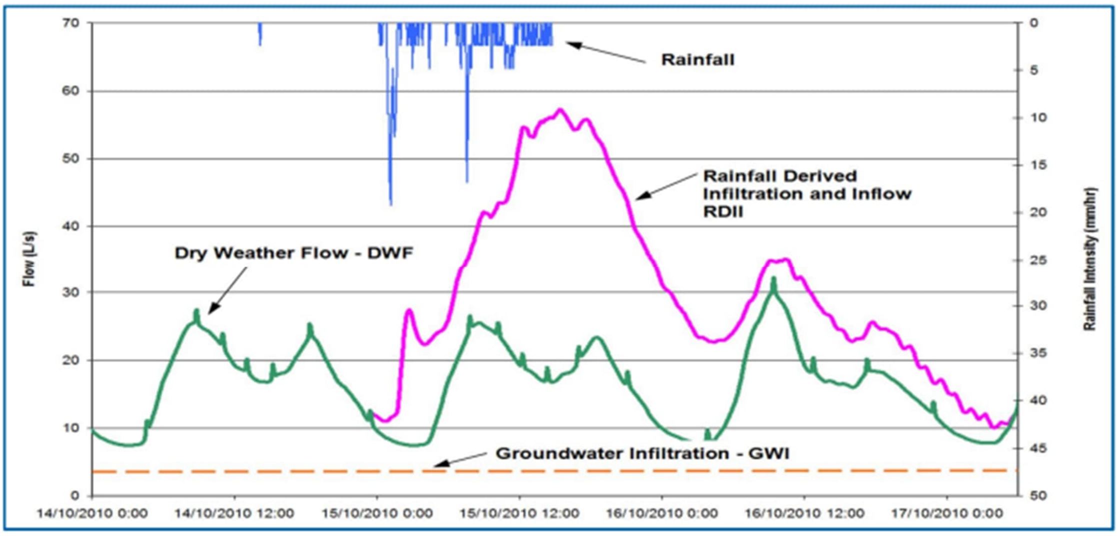

Figure 1: Wet-Weather Flow Components (Water New Zealand infiltration & Inflow Control Manual 2nd Edition, March 2015)

High ratios of wet to dry weather, are indicative of high I&I. Correlating flow data with rainfall data is also essential to characterise I&I. Rainfall data from local weather stations or rain gauges is used to determine the timing and intensity of rain events, and these data can be compared with flow data to identify periods of excessive flow.

I&I studies are computation heavy and are usually carried out on an annual basis to prioritise catchments and further investigations.

A NEW APPROACH

In the last few years, the water industry across the globe has put a real focus on reducing wastewater overflows in order to improve community and environmental outcomes. This can be seen with the development of a new Plan for Water published in April 2023 by the UK Department for Environment, Food & Rural Affairs which sets the following targets for water companies (Drainage and

Wastewater Management Plan Alignment to the Storm Overflows Discharge Reduction Plan, May 2023, SouthWest Water):

Companies are only permitted to discharge from a storm overflow where it can be demonstrated that there is no local adverse ecological impact, profiled such that:

- 75% of storm overflows discharging into or close to high priority sites are addressed by 2035.
- 100% of storm overflows discharging into or close to high priority sites are addressed by 2045.
- 100% of all storm overflows are addressed by 2050.

By 2050, companies will not be allowed to discharge from storm overflows for more than 10 rainfall events per year.

These targets have resulted in a significant increase in the number of monitors being installed in the UK, with Anglian water investing in approximately 20,000 sensors for its entire network in the next two years or Thames Water increasing monitors to up to 18,500 by the end of 2025. The increase in monitoring devices needs to be combined with the latest technology in data analysis to turn this dataset into actionable insights and achieve a significant reduction in overflows.

We have combined state-of-the-art technologies in sewer monitoring, hydraulic engineering, statistics, and computer science to develop solutions maximising insights from available IoT data to address the above-mentioned issues more proactively.

The result is an ML and statistics-based methodology that takes minutes to calibrate and seconds to predict once calibrated, thus allowing for scalable deployment to 1000's of sites. In addition, ML and statistical models have been set up to learn from past data and continually self-calibrate. The speed of the models allows us to run analysis in near real-time which offers major improvements to wastewater management by enabling a preventive approach. We have currently developed and tested our solution across multiple sites in Auckland, New Zealand where at least 1 year of flow and rainfall data was readily available. The outcome of the analysis was designed to be integrated with any in-house or external platform and enable real-time predictions, alerts and isualization across any area of interest.

METHODOLOGY

Our system to detect wastewater blockages and forecast wet weather overflow employs a dual methodology approach, consisting of a forecasting model and a statistical anomaly detection model that work together in tandem.

WET WEATHER OVERFLOW FORECAST

At the foundation of our system of predictive models is an ML model – trained via XGBoost, a decision tree-based algorithm – to forecast sewer depth, given inputs of historical sewer depth and forecasted rain. We trained our models on 5 years of historical sewer and rain data across 27 sites and use the same model at every site to predict sewer depth 24 hours ahead. During the training process, we performed feature transformations appropriate for time series input – we considered various weighted means and rolling sums of recent rain data as a proxy for accumulated rain at the moment of prediction, as well as looking at rain forecasts, and previous values of the sewer depth data, to aid in predicting future depth values.

With the ML model deployed to the cloud, these predicted forecasts for sewer depth can be made available in real time. The live forecasts can then be fed into specialised statistical models to identify anomalous events in the live data.

WASTEWATER BLOCKAGE DETECTION

We opted for a statistical model given the unlabelled nature of historical blockage events in the dataset, ruling out other supervised ML models for anomaly detection. In tuning a statistical model, we are better able to build up our understanding of the model performance and edge cases.

By utilising sewer depth forecasts as an indication of expected depth levels, we picked up any deviations from this expected depth that are characteristic of wastewater blockages. We focused on two different profiles of blockages, each with different detection criteria:

Significant Dry Blockage

- Rolling 12-hour median is significantly larger than the 1-week baseline.
- “Expected Data” predictions are significantly larger than the depth for the last 6 hours.
- The gradient of the series is not decreasing.
- There has been less than 15mm of rain in the last 24 hours.

Partial Dry Blockage

- The daytime median (7:00am – 9:00pm) is larger than the model predicted values at this time.
- The night-time median (9:00pm – 7:00am) is smaller than the model predicted value at this time.
- The gradient of the series is not decreasing.
- There has been less than 15mm of rain in the last 24 hours.

This produces a binary labelled series whether that data is indicative of a blockage.

INFLOW AND INFILTRATION CHARACTERISATION

We created an automated dry day selection application to select dry days automatically, rather than manually for each gauge. The average dry weather flow for each day of the week is calculated and updated hourly as data becomes available. The automated process performs the following tasks:

- Data cleansing: This step includes interpolating data to obtain regular timesteps, removing outliers from flow data, excluding public holidays, school holidays and days where missing data exceeds two hours.

- Dry day selection: Selecting dry days based on antecedent wetness conditions described below:

$$\sum_{Day=0}^{Day=14} 0.7^{day} \times Rain\ (mm)_{day} < 3.5mm$$

- Generate ADWF Hydrograph: Calculates the average dry day for each day of the week, if enough dry days are available or for weekdays and weekends.

Data is processed hourly and ADWF are updated as new data comes in, looking back to a maximum period of 1 year.

Once an ADWF hydrograph has been generated an automated storm event selector uses rain gauge data to identify the start of an event by verifying a set of rules related to rainfall depth in a specified period of time (e.g., 20mm in 24hours). The storm event starts when rainfall is greater than 0.5mm and stops when the observed flow returns within 10% of the ADWF.

Once an event is selected and the ADWF is computed, the RDII for the event can be calculated. The RDII volume is simply the difference between computed average dry weather flow and measured flow for a specific time period.

OUTCOMES

WET WEATHER OVERFLOW FORECAST

Our wet weather overflow forecasting model is connected to rainfall predictions to provide 24 hour forecasted levels and identify the extent of overflows related to an upcoming rain event.

Our model achieves a mean error of 0.188, or 3.6% of the depth range. Practically, this means that (after accounting for anomalies in the dataset) our model is on average about 3.6% away from the true value. This isn't uniformly distributed as the model accuracy depends on the behaviour of the gauge. We have trained the model such that it maintains performance across these metrics when we filter for overflow events. Ultimately, we have found that the model can predict when overflow events will happen to within the hour ~91% of the time, alerting users 24-hours in advance when this is predicted.

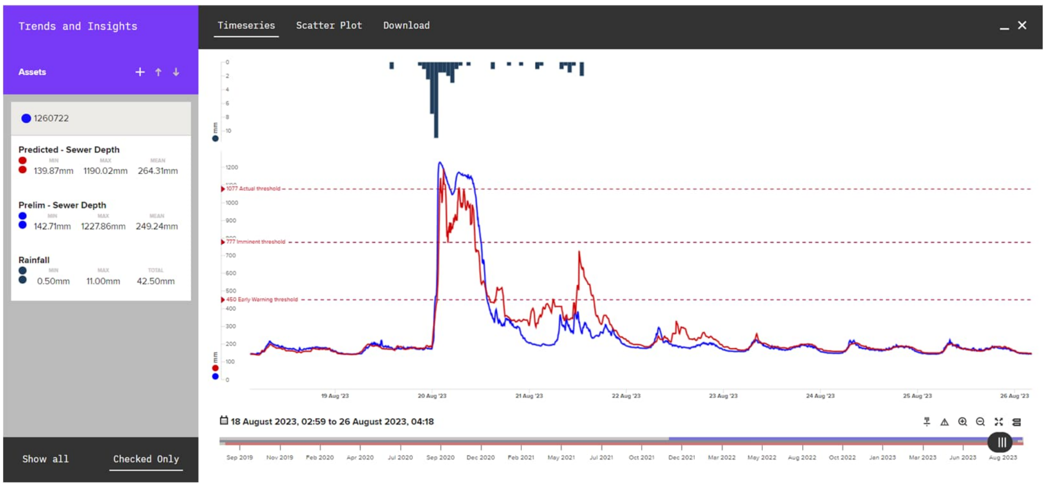

Figure 2: Wet weather sewer level prediction

The blockage identification trial returned the following results:

- 74.0% precision on all data (2096 out of 2831 model labelled points were during an event). Example shown in Figures 3 and 4 below.
- 78.7% precision on all data when removing gauge with many negative values (1925 out of 2445 model labelled points were during an event)
- 44.6% recall on event data (at least one anomalous value raised in 50 out of 112 labelled events)

These results show that the model was effectively trained to reduce the number of False Positive events. The recall is acceptable for a sample this small but will be the focus for improving the model. With access to labelled data, more complex models could be trained to further improve the performance of the model.

Figure 3: Positive Detection of Significant Blockage Event. Detection is shown in red. Successfully detected due to increase in actual depth (compared to predicted depth in red) outside of a rain event.

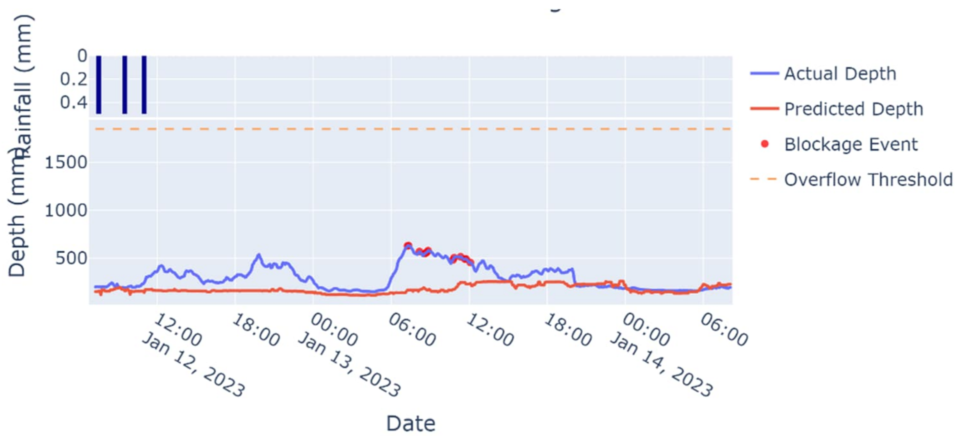

Figure 4: False Positive Detection Example. Detection is shown in red. Unsuccessful detection caused by difference between the median depth and predicted depth.

INFLOW AND INFILTRATION CHARACTERISATION

The selection of dry weather days and automated calculation of ADWF was tested for over 50 temporary gauges in the Auckland area and returned satisfactory results. The ADWF was checked by wastewater modellers as part of Watercare's Network Performance Monitoring and Modelling who used automated dry weather flows as part of their projects. This process has increased efficiency and consistency in the calculation of ADWF.

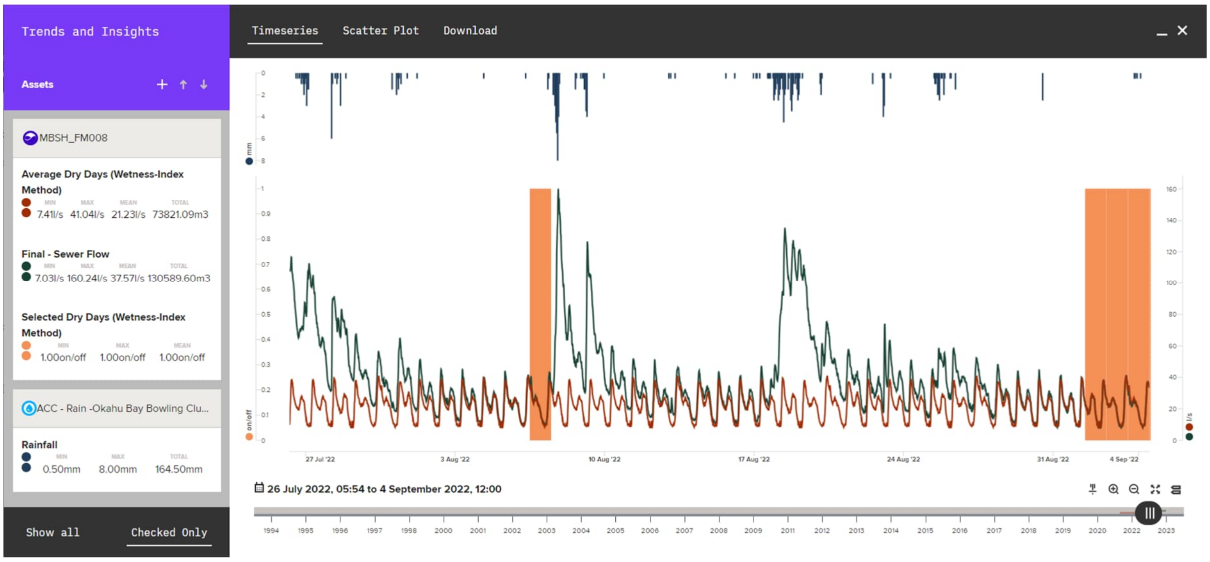

Figure 5: Automated dry weather day selection and ADWF calculation.

The RDII volumes automatically generated can already be used to prioritise high II catchments. We are currently improving our storm event detection to better define the end of events for long term events. These improvements include:

- discarding events which do not get back within 10% of ADWF before the next rain event occurs,
- ending an event when flow gets back within 10% of the closest dry day (instead of ADWF) to consider dry weather seasonality.

Another step for improvement will be to consolidate all identified RDII events and compile them in a tabular format and or geospatial layers to return the following key indicators:

- Groundwater Infiltration (GWI) Volume
- Dry Weather Flow Production
- Peaking Factor
- Percentage Ingres (II Volume / (Catchment Area * Total Rainfall)
- Leakage Severity (II Volume / Length of Pipe / Total Rainfall)

CONCLUSIONS

The water industry has recently intensified its efforts to mitigate wastewater overflows in both wet and dry weather conditions, aiming to protect public health and the environment. This heightened focus, accompanied by more stringent targets, has driven a substantial investment in IoT devices for monitoring sewer depth, flow, and overflow incidents at constructed sewer overflows.

This surge in the number of monitoring devices deployed must now be harnessed in conjunction with recent technological advancements, including machine learning, statistical analysis, and cloud computing, to elevate the effectiveness of wastewater monitoring and optimise return on investment. The trial carried out in Auckland has demonstrated that technology can be leveraged to automate early blockage detection, to provide predictive overflow analysis, and to support data-driven decision-making and prioritise inflow and infiltration remediation efforts.

The next phase in this journey involves refining ML and statistical models as more training data is captured and feedback is received from operational teams. This step is critical to enhance system performance and the accuracy of the proposed solutions, thereby increasing their reliability and effectiveness, ultimately maximising the benefits of wastewater monitoring.

While technology will play a pivotal role in reducing overflows in the foreseeable future, it is important to acknowledge the inherent complexity of this challenge. Effective solutions will necessitate a collaborative approach, drawing upon multidisciplinary resources such as planning teams to expand network capacity, asset management teams to renew and maintain aging infrastructure, and operations teams to address unforeseen blockages. In this context, the prioritisation of issues and the real-time sharing of critical insights among relevant stakeholders will also be paramount to successfully reducing overflows and mitigating their impact.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Watercare for their support in using their data for this trial.

Mott MacDonald Digital Ventures for funding the research.

Mott MacDonald technical team who developed these solutions and contributed to this paper: David Burton, George Doughty, Evan Harwin, Matthew Fredericks, Yeunjin Kim, Sreedath Ikarath.

REFERENCES

Hedges, P. D. (1994) ‘The prediction of combined sewer overflow performance from models - A case study’, Water Science and Technology.

Carstensen, J., Nielsen, M. K. and Strandbæk, H. (1998) ‘Prediction of hydraulic load for urban storm control of a municipal WWT plant’, Water Science and Technology.

Duchesne, S. et al. (2001) ‘Mathematical modeling of sewers under surcharge for real time control of combined sewer overflows’, Urban Water.

Politano, M., Odgaard, A. J. and Klecan, W. (2007) ‘Case study: Numerical evaluation of hydraulic transients in a combined sewer overflow tunnel system’, Journal of Hydraulic Engineering.

Garofalo, G. et al. (2017) ‘A distributed real-time approach for mitigating CSO and flooding in urban drainage systems’, Article in Journal of Network and Computer Applications.

Grum, M. et al. (2011) ‘Full-Scale Real Time Control Demonstration Project in Copenhagen’s Largest Urban Drainage Catchments’, in Proceedings of the 12th international conference on urban drainage, Porto Alegre.

Morales, V. M., Mier, J. M. and Garcia, M. H. (2017) ‘Innovative modeling framework for combined sewer overflows prediction’, Urban Water Journal.

Szeląg, B., Chmielowski, K. and Dacewicz, E. (2018) ‘Simulation of a storm overflow with probabilistic and hydrodynamic models’, Urban Water Journal.

Thorndahl, S., Schaarup-Jensen, K. and Jensen, J. B. (2008) ‘Probabilistic modelling of combined sewer overflow using the First Order Reliability Method’, Water Science and Technology.

Osiadacz, A. J. (1996) ‘Different Transient Flow Models - Limitations, Advantages, And Disadvantages’, Pipeline Simulation Interest Group.

Kochevsky, A. N. and Nenya, V. G. (2004) ‘Contemporary Approach for Simulation and Computation of Fluid Flows in Centrifugal Hydromachines’.

Vu, Q. & Joseph, T (2020) ‘Flow and level prediction with machine learning – digital twin solution’, Water New Zealand Stormwater Conference.

Joseph, T & Stephens P (2015) ‘Real time inflow and infiltration analysis – a new tool to focus your renewal spend, Water New Zealand Conference.

Water New Zealand (2015) ‘Infiltration & Inflow Control Manual 2nd Edition’ (March 2015).

UK Department for Environment, Food & Rural Affairs, (2023) ‘Plan for Water: our integrated plan for delivering clean and plentiful water’.

SouthWest Water (2023) ‘Drainage and Wastewater Management Plan - Alignment to the Storm Overflows Discharge Reduction Plan’.