A Strategic Digital Transformation for the Water Industry

Edited by

Oliver Grievson
Timothy Holloway
Bruce Johnson

Published by

IWA Publishing
Republic - Export Building, 1st Floor
2 Clove Crescent, E14 2BE
London, United Kingdom

T: +44 (0)20 7654 5500
F: +44 (0)20 7654 5555
E: publications@iwap.co.uk
W: www.iwapublishing.com

First published 2022

ISBN: (Paperback) 9781789063394
ISBN: (eBook) 9781789063400
DOI: 10.2166/9781789063400

DISCLAIMER

The information provided and the opinions given in this publication are not necessarily those of IWA and should not be acted upon without independent consideration and professional advice. IWA and the Editors and Authors will not accept responsibility for any loss or damage suffered by any person acting or refraining from acting upon any material contained in this publication.

This book is a collection of the 9 white papers published in the IWA Digital Water Programme's White Paper Series.

The core content and message of the white papers remain the same. However, additional information has been added to keep the topics up to date.

COPYRIGHT NOTICE CHAPTER 2

Contributors

Chapter 1

Randolf Waters *Senior Director of Global Partnerships & Branding, Xylem Inc.*
Will Sarni *CEO, Water Foundry*
Cassidy White *Consultant, Water Foundry*
Katharine Cross *Strategy and Partnerships Lead, Australian Water Partnership*
Raül Glotzbach *Scientific Researcher, KWR Water Research Institute*

Chapter 2

Oliver Grievson *Technical Lead, Z-Tech Control Systems Ltd.*
Daniel Aguado *Associate Professor, Institut Universitari d'Investigació d'Enginyeria de l'Aigua i Medi Ambient (IIAMA), Universitat Politècnica de València, València*
Frank Blumensaat *Senior Research Fellow, ETH Zurich/Eawag, Institute of Environmental Engineering, Chair of Urban Water Management Systems*
Juan Antonio Baeza *Full Professor, Department of Chemical, Biological and Environmental Engineering, Universitat Autònoma de Barcelona, Cerdanyola del Vallès*
Kris Villez *Sr. R&D Staff Member, Oak Ridge National Laboratory*
Maria Victoria Ruano *Associate Professor, Chemical Engineering Department, Universitat de València*
Oscar Samuelsson *Researcher, IVL Swedish Environmental Research Institute*
Queralt Plana *R&D Engineer, modelEAU, Université Laval*
Dragan Savic *Chief Executive Officer, KWR Water Research Institute*
Peter van Thienen *Senior Researcher & Chief Information Officer, KWR Water Research Institute*

Chapter 3

Bruce Johnson *Technology Fellow at Jacobs*
Borja Valverde- Pérez *Assistant Professor at Department of Environmental Engineering, Technical University of Denmark*
Christoffer Wärff *Industrial PhD student at RISE Research Institutes of Sweden / Lund University*
Douglas Lumley *DHI Sweden Ab*
Elena Torfs *Assistant Professor at Department of Data analysis and Modelling, Ghent University*
Ingmar Nopens *Full Professor, BIOMATH, Ghent University*
Lloyd Townley *Chief Scientist at Nanjing University Yixing Environmental Research Institute and Founder of Nanjing Smart Technology Development Co. Ltd*
Zoran Kapelan *Delft University of Technology, Department of Water Management*
Emma Weisbord *Digital Water Partnerships, Royal HaskoningDHV Digital*
Vladan Babovic *National University Singapore, Department of Civil and Environmental Engineering*
Timothy Holloway *Senior Scientific Officer at University of Portsmouth School of Civil Engineering and Surveying*

John Barry Williams *Professor of Environmental Technology at University of Portsmouth School of Civil Engineering and Surveying*
Djamila Ouelhadj *Professor of Operational Research at the University of Portsmouth Department of Mathematics*
Gong Yang *Process Capacity Manager at Southern Water Services Ltd*
Ben Martin *Lead Research Scientist at Thames Water*
Matthew Wade *Science Lead: Wastewater Research and Development, Environmental Monitoring for Health Protection, Data Analytics and Surveillance at the UK Health Security Agency*

Chapter 4

Alex Knezovich *Toilet Board Coalition*
Brian Moloney *Co-Founder and Managing Director of StormHarvester*
Jonathan Lavercombe *Chief Operating Officer at StormHarvester*
Jody Knight *Asset Technology Manager at Wessex Water*
Jacobo Illueca *Wastewater Specialist at Idrica*
Raimundo Seguí *Epidemiologist at Global Omnium*

Chapter 5

Hagimar von Ditfurth *Digital Ambassador for GIZ Uganda*
Emma Weisbord *Digital Water Partnerships, Royal HaskoningDHV Digital*
Thor Danielsen *Sector Expert – Water at the Embassy of Denmark in Poland*
Lloyd Fisher-Jeffes *Digital Water Partnerships, Royal HaskoningDHV Digital*
Ana Claudia Hafemann *Regulatory Analyst at AGIR*
Juljana Hima *Organization for Security and Cooperation in Europe*
Temple Chukwuemeka Oraeki *State Focal Person, Consultant at British Council*
Oliver Grievson *Technical Lead, Z-Tech Control Systems Ltd.*

Chapter 3

Artificial intelligence, digital twins and dynamic resilience

Introduction

Global digitalisation and the accelerated development of digital technologies and analytical tools have ushered rapid changes in how humans interact with the different sectors of our society. This has had a significant impact on the international water industry, as described in this book. The amount of data we have at our fingertips is typically much more than any person or group can effectively use, so help is needed to improve our decision-making based on these large data sets. These vast datasets are generated intensively from instrumentation mounted in water/wastewater assets (see Chapter 2) leading to vast data silos of individually formatted and difficult to access data. For example, data logged by water companies is typically done at 15 min intervals, equating to 87,000 data points per annum. When multiple parameters are measured and logged for water and wastewater systems, it can result in millions of data points per annum. Processing this data can present a significant challenge for standard spreadsheet-based packages; therefore, new methods are required to convert this information into valuable insights.

The previous chapter discussed the importance of data and how to collect and transform it into information along with the challenges associated with these processes. When the data/information has been generated and stored, it should be used to generate insights that are both communicable and informed. These insights should extract otherwise unidentifiable characteristics, such as hidden dimensions or trends from pre-existing data or modelled outputs.

This chapter will present case studies where analytical tools (artificial intelligence, digital twins and dynamic resilience) were used to collect and better understand this available data. These analytical tools can help us understand data in distinct ways.

- Artificial intelligence (AI) helps us find and make use of obscured hidden patterns in large data sets, whether it be in numerical or visual patterns. This valuable tool enables computer systems to rapidly analyse water-related data in ways that human operators may not.
- Digital twins (DT) are essentially tools that force the integration of all the available data into a digital simulation of a system. This enables users to make more time-relevant and actionable decisions about operations by providing the bigger picture about how all the data streams are working together and are related to each other.
- Dynamic resilience (DR) uses machine learning to extract system operating conditions from actual WRRF data. Data-driven simulations are then performed before generated data is transformed into a dynamic resilience heat map of system stresses to represent how a water resource recovery facility (WRRF) system reacts to extreme events or stressors.

While the three tools are distinct from each other, they can be combined in some very powerful ways, as the case studies in this chapter show.

3.1 Artificial intelligence solutions for the water sector

Artificial intelligence (AI) is a term used to describe the theory and development of computer systems that are able to perform tasks normally requiring human intelligence. Examples of AI type systems include various expert systems, speech recognition, and systems used for financial trading. Two important concepts that are frequently mentioned in the context of AI are Machine Learning and Computer Vision.

Machine Learning is the study of computer methods and algorithms that improve automatically through experience. Machine Learning typically involves building some computer “black-box” model from data with the ability to make predictions without the need for additional software programming. One of the best-known machine learning methods is the Artificial Neural Network (ANN) that works by mimicking biological neural networks that exist in a human brain. ANNs learn from training data which is used to capture the functional relationships among the data, even if the underlying relationships are not known or the physical meaning is difficult to explain. This enables the ANNs to discover patterns in data that are often unknown, even to the best experts in the field. Computer Vision denotes a set of AI-type methods that are used to train computers to interpret and understand digital images and videos. Examples of computer vision applications include systems for facial recognition, medical diagnostics, and driverless cars.

Many AI methods exist and it is not our intention to describe these in detail. Instead, this section presents examples of AI-based solutions that make use of these methods.

AI-based solutions

REAL-TIME DETECTION OF PIPE BURSTS IN WATER DISTRIBUTION NETWORKS

Leakage is a major issue in water distribution systems worldwide. This example presents an AI-based system that detects pipe bursts/leaks but also equipment and other failures in these systems. The detection system works by automatically processing pressure and flow sensor signals in near real time to forecast the signal values in the near future (using ANN). These are then compared with incoming observations to collect different forms of evidence about the failure event taking place. The evidence collected this way is processed using Bayesian Networks to estimate the likelihood of the event occurrence and raise corresponding alarms (Romano et al., 2014). The system effectively learns from historical burst and other events to predict future ones. Elements of the detection system, developed initially as part of a research project, were built into a commercial Event Detection System (EDS).

Figure 3.1 EDS Screenshot with Real-Time Analysis of a Specific Alarm

An EDS has been in use by a large UK water company since 2015. It processes data from over 7,000 pressure and flow sensors every 15 minutes (see Figure 3.1). This enables EDS to detect pipe bursts and related leaks in a timely and reliable manner, i.e., shortly after their occurrence and with high true and low false alarm rates. In addition to detection, EDS can proactively prevent burst events by detecting equipment failures that often precede these events (e.g., pressure reducing valve failures). An EDS does not make use of a hydraulic or any other simulation/mechanistic model of the analysed water distribution network, i.e., it works solely by extracting useful information from sensor signals where bursts and other events leave their imprints (i.e., deviations from normal pressure and flows signals). This makes the EDS robust and scalable as it enables data to be processed in near real time (i.e., within the 15 minute time window). The use of EDS has resulted in major operational cost savings to date, significantly reduced customer supply minutes lost, reduced leakage and several other benefits to the water company (full details not mentioned for commercial reasons). All this has led to a change in company business culture and improved service to over 7 million customers.

AUTOMATED ASSET CONDITION ASSESSMENT USING AI AND COMPUTER VISION

The inspection of urban drainage (i.e., sewer) systems’ pipes is important as undetected structural and other faults (e.g., displaced joints, cracks, etc.) may result in severe pollution and/or flooding incidents. This inspection is done usually by recording CCTV videos and then analysing these manually. This process is time consuming (i.e., costly) and subjective/inconsistent in nature hence not necessarily always reliable.

The AI-based solution automates the process of analysing CCTV videos and detection of faults in pipes. It does this by using computer vision and machine learning methods (Myrans et al., 2018). Image processing is conducted first to process and convert the CCTV images into suitable data. This data is used then to detect faults with a help of a Random Forest machine learning method. This method is trained before it is used on a number of pre-labelled CCTV images. The automated detection works similarly to the human face recognition system although the task of fault detection is more complex in sewers due too many different types of faults that exist and that can manifest themselves in very different ways in CCTV images.

The above solution was successfully evaluated and validated on unseen real CCTV data from several water companies in the UK, Finland, and Australia. It has a high true detection rate accompanied with low false alarms rate. This technology is currently being commercialised by a variety of companies around the world.

PREDICTIVE WASTEWATER TREATMENT PLANT CONTROL

Royal Haskoning DHV's Aquasuite® software was deployed at PUB Singapore's Integrated Validation Plant at Ulu Pandan Water Reclamation Plant in March 2019. It was introduced to provide operators and managers with predictive insights while improving plant performance. Real-time data is collected on the plant's flows and qualitative measurements, including those for ammonia, nitrates, oxygen, phosphates, and dry solids, and builds a historical database. The software then makes use of advanced analytics and Machine Learning algorithms to predict the plant's wastewater flows and loads, oxygen needs, chemical dosing needs, and other requirements (see Figure 3.2). The system controls key treatment processes, automatically optimising them in real-time based on its predictions and the plant's historical performance.

It further detects anomalies in the plant's processes through quantile regression techniques based on multiple measurements. Prediction accuracy of the influent flow increases over time as the software is learning, reaching a prediction accuracy of 88% after just one month.

Connecting to the plant's Supervisory Control and Data Acquisition (SCADA) system to gather data and control key processes, data is sent to the cloud component of this solution. With the data collected in near real time in the cloud, the software tracks the actual performance of the on-premises solution through a digital twin, a digital replica of a physical system. More advanced analytics are made available to operators through the cloud, while the on-premises part keeps optimising real-time performance. Machine learning is used to understand the efficiency of each process, and the learnt relationship is then used with the prediction of the influent load to decide the most optimal efficient set points for treatment. Results show that this digital twin learns and predicts operations several days ahead and it can function as an autopilot, able

Figure 3.2 Dashboard view of Aquasuite Pure

Figure 3.3 Dashboard view of flow levels and predicted outcomes

to perform unattended operation. Preliminary results show a reduced aeration flow of up to 15% with predictive control, resulting in corresponding energy savings.

SMART ALARMS FOR PROACTIVE WASTEWATER NETWORK MANAGMENT

The flow prediction component of this digital twin is an AI-powered predictive analytics tool for wastewater networks that enables prediction and early warning of critically high-water levels in sewers, potential pollution events, and detection of anomalous levels that could indicate blockages in the network (see Figure 3.3).

Water utilities are most at risk for blockages during rainfall events, in part because of sewer flooding, capacity limitations, and foreign objects.

AI is used to detect when high level flows in sewers exceed critical levels or are not consistent with the expected flows. The system identifies these as anomalies through a customisable

1
2
3
4
5

smart alarm notification system. Using AI in this context drastically reduces false alarms and prioritises the remaining alarms to manage blockages accurately and efficiently through early detection and real-time response, preventing further problems like fatbergs. Monitors located strategically around the catchment collect flow and level data. That data is transmitted to a secure, reliable cloud platform that displays the information in easy-to-use dashboards and reports for analysis. Using smart algorithms, real-time data, historical data sets, and information from other sources such as rainfall, data are translated into actionable insights for operators.

The flow prediction AI model is built on two multi-layer perceptron ANNs that work together to support operators. A regressive ANN predicts flow levels for the next 48 hours and a heuristic ANN identifies event anomalies like blockages. Integrating machine learning allows the anomaly detection to continually improve over time. Human data technicians review the data alongside the predictive analytics software to verify the alarm protocols. The additional human intervention helps detect any existing or temporary anomalies, seasonal variations, and external influences. Prediction of a high sewer level enables water utilities to proactively inform their customers of potential sewer flooding, and, in critical environmental regions, to actively prevent spills. Anomaly detection enables the utilities to detect and clear blockages before they cause sewer overflows.

REAL-TIME FORECASTING OF SEA CURRENTS

Over the past few decades ANNs have evolved in a popular approximation and forecasting tool, frequently used in a range of problems and application areas (ASCE Task Force, 2000). In this solution, a recurrent ANN is used to create a real-time hybrid data assimilation system resulting in extremely accurate forecasts of sea surface currents. This solution was developed to support the construction of Øresund Link connecting Danish capital Copenhagen and City of Malmø in Sweden (Babovic et al., 2001). The combined roadway and rail line bridge run nearly 8 km where it then transitions into an underwater tunnel for the remaining 3.5 km. Due to the material of the seafloor, a tunnel was not possible. Instead, engineers chose to sink and connect 20 prefabricated reinforced concrete segments – the largest in the world at 55,000 tonnes each – and interconnect then in a trench dug at the seabed. The elements were prefabricated in a special purpose-built facility north of Copenhagen, sealed shut and using a specially designed barge along with seven tugboats, lowered into place at required accuracy of alignment of 2.5 cm. The towing operation for each element could be conducted within a “window of opportunity” of 36 hours during which sea surface currents had to be guaranteed to be less than 0.75 m/s. Despite extremely challenging conditions, all 20 elements of the Øresund link’s tunnel were successfully placed at their positions between 11 August 1997 and 6 January 1999 – 20 towing operations in 17 months! It is suggested that the accurate ANN-based forecasting of seal surface currents was one of the key factors in this achievement. The Øresund Strait Link was opened to the public on 1 July 2000.

BAYESIAN NETWORKS FOR PRO-ACTIVE ASSET MANAGEMENT

The economic and social costs associated with pipe bursts and leakage in modern water supply systems are rising to unacceptably high levels. The challenge for the decision-maker is to determine what pipes in the network to rehabilitate, by which rehabilitation method, and at what time within the planning horizon. Advanced AI, machine learning, and statistical methods are used in order to establish risks of pipe bursts (Economou et al., 2014). For example, analysis of the database of already occurred burst events can be used to train a risk model as a function of associated characteristics of bursting pipe (its age, diameter, material of which it is built, etc.), soil type in which a pipe is laid, climatological factors (such as temperature), traffic loading, etc. In this context, a machine learning model based on a Bayesian network is used to predict which pipes are most vulnerable to failure including a metric for failure probability (Babovic et al., 2002). Bayesian networks are a probabilistic graphical model that use Bayesian inference for probability computations. The approach models conditional dependence and, therefore, causation. Through these relationships, one can efficiently conduct inference on the variables in the graph exemplifying pipe failure mechanism. Pilot projects using the approach have been conducted in Sweden, Singapore, UK, and Denmark.

COMPUTER VISION FOR OPPORTUNISTIC RAINFALL MONITORING

The quantity and quality of precipitation data are crucial in meteorological and water resource management applications. Rain gauges are a classic approach to measuring rainfall. However, as we enter the age of the Internet of Things (IoT) in which “anything may become data” so-called opportunistic sensing using unconventional data sources offer a promise to enhance the spatiotemporal representation of existing observation networks. One particular area attracting attention is the estimation of quantitative and analytical rainfall intensity from video feeds acquired by smart phones or CCTV surveillance cameras. Technological advances in image processing and computer vision enable extraction of diverse features, including identification of rain streaks enabling estimation of the instantaneous rainfall intensity (Allamano et al., 2015). Recent AI and machine learning approaches rely on the use of autoencoders, deep learning, and convolutional neural networks to address the problems. Companies such as WaterView (Italy), Hydroinformatics Institute (Singapore), as well as universities (e.g., Southern University of Science and Technology China, Shenzhen) have proposed and implemented practical approaches to weather hazards in energy, automotive and smart cities application domains (Jiang et al., 2019).

3.2 Operational digital twins in the urban water sector

Due to advances in instrumentation and the increasing availability of online data and computing capacity for utilities (e.g., via cloud computing), the development of digital twins has recently attracted large interest in the urban water sector. This interest is primarily driven by two trends in our industry:

1) The increasing amount of data available at our facilities; and

2) The business drivers around both improved capital and operational efficiency targets.

The amount of data available at many facilities now exceed the amount that most operational staff can utilise in their day-to-day operations. Digital twins can take those data and present them in terms that staff can use effectively on a day-to-day basis (Garrido-Baserba et al., 2020). The second driver towards efficiency relates to the ability of a digital twin to go beyond conventional PID (i.e., proportional–integral–derivative) control, using model-based control and data-driven optimisation. This improvement has implications for capital expenditure since more reliable control can result in reductions in conservatism, and thus lower capital expenditure (Stentoft, 2020).

In a broad sense, "digital twin" refers to a digital replica of physical assets. Within the water sector, digital twins are combinations of models that provide a digital representation of a specific part of the water system (e.g., water resource recovery facilities, sewers, etc.) and utilise automated real-time data from multiple sources to, e.g., simulate expected, desired, or critical behaviour of the physical system (Pedersen et al., 2021).

While there is agreement in a broad sense about the definition, there is a lack of consensus on which elements characterise a digital twin. Some experts consider traditional mechanistic models with frequent recalibration using sensor data enough to qualify as a digital twin, whereas others consider the interaction with real-time data as a key element in an operational digital

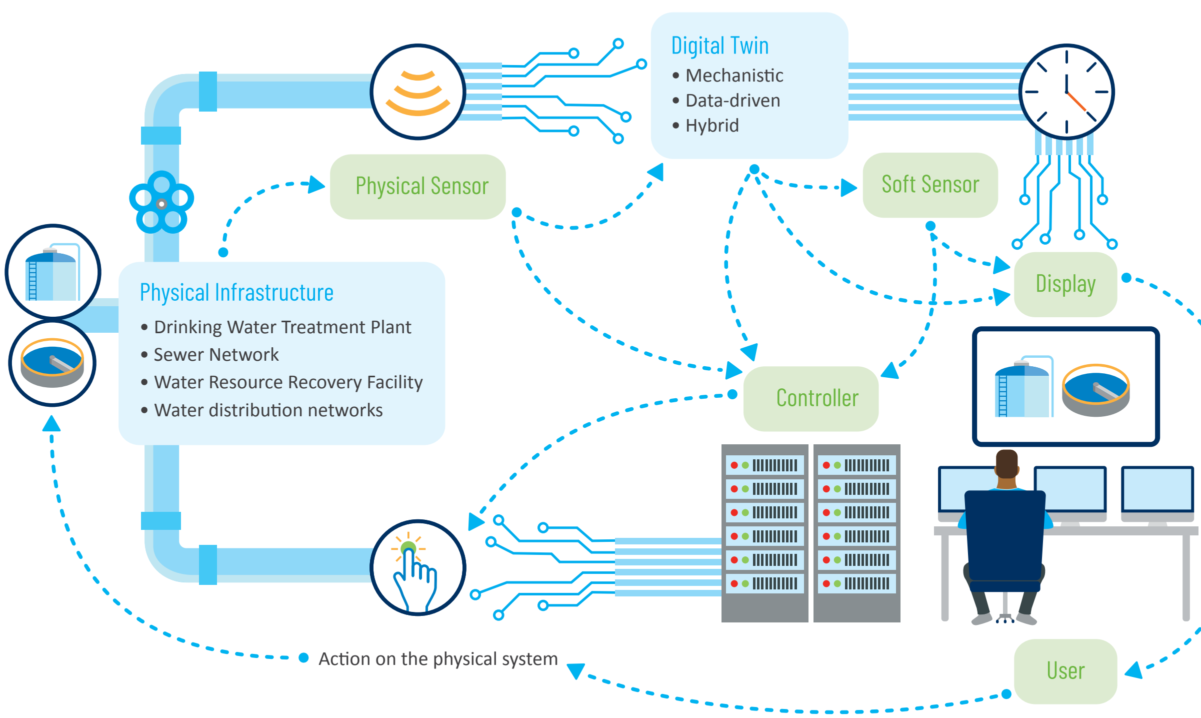


Figure 3.4 Basic structure of a digital twin application. Elements in blue are basic elements needed for implementation of a digital twin, while elements in green are complementary to the digital twin, ultimately to enable an action on the physical system.

twin. In this discussion, it is important to distinguish between digital twins for operational use (relying on near- real-time data from the physical system) and digital twins for planning, design, construction, or investment purposes (relying on historical data from similar physical systems or expected demands given the operational environment of the physical system under development). In some cases, digital twins are considered as soft or virtual sensors of the physical system. Thus, aspects like model complexity (mechanistic, data-driven, or a hybrid combination of both), handling of uncertainty and data requirements, among others, need proper assessment to ensure successful alignment between digital twins and their intended application. Furthermore, data assimilation from different data sources remains a key challenge to be addressed. Figure 3.4 illustrates the most basic structure of a digital twin and its interaction with the real system and users.

Digital twin applications include (but are not limited to):

(1) Investment planning via future scenario analysis, traditionally performed with models calibrated with historical data;

(2) Real-time decision support for selection of different operational strategies;

(3) Operator training;

(4) Online optimisation (e.g. model predictive control) for energy or resource savings or compliance management (e.g., to minimise carbon footprint);

(5) Asset management and interaction with different stakeholders; and

(6) Sewage epidemiology (Poch et al., 2020).

Within the water sector, digital twins can support the transition towards more proactive management of water infrastructure, whereby different processes and systems can be operated to mitigate disturbances before they have adverse impacts on performance (Karmous-Edwards et al., 2019). Thus, there is large potential for economic savings (e.g., online energy optimisation), more effective protection of the environment (e.g., model predictive control for effective nutrient removal) and increased benefits for society (e.g., improved storm water management to minimise risk of flooding in urban areas).

In this section, two examples of online digital twins for operational decision support are presented: one focuses on improved control of combined sewer overflows (CSOs) in a sewer system and the second focuses on proactive maintenance and process optimisation of a WRRF intended for water reclamation. Both are considered hybrid digital twins which uses a combination of mechanistic modelling and machine learning computational methods.

Digital twin for sewer networks

This case study describes the digital twin approach used for sewer networks in the Swedish cities of Gothenburg and Helsingborg within the project Future City Flow. The approach is similar in both cities, but this case study will focus on the digital twin developed in Gothenburg (WRRF: 900 000 PE; catchment: 240 km^2 with 20 km^2 of impervious surface; 40% combined sewers).

The utilities regularly experience high flows in the sewer collection systems leading to spills from CSOs that lead to 3 billion litres per year (2.2% of total flow) of untreated wastewater being discharged to the environment. The CSOs mostly occur due to large flow variations caused by heavy rainfall events. To manage issues related to CSOs and to reduce storm weather impacts on the WRRF, an operational digital twin approach was envisioned as a decision support system with online flow prediction and suggestions for control strategies (the final decision is currently made by the operator). Part of the last stage of the project (2019- 2021) focuses on implementation of full model predictive control (MPC).

COMPONENTS

The structure and different components of the digital twin are shown in Figure 3.5. The main part of the digital twin consists of a model of the sewer system built in MIKE URBAN using several different modules, including:

(1) A dynamic model for the hydraulics of the pipes and tunnels;

(2) Conceptual hydraulic models to describe sub-catchments;

(3) Optimisation modules for real-time control; and

(4) Modules for handling of precipitation forecasts and rain gauge data quality control.

The model has been calibrated manually for many of the sub-catchments in the network where rain gauges and flow measurements are available, as well as for the hydraulic model of the main tunnel system. The catchment model is re-calibrated manually to achieve better flow prediction as more sensors are added to the system over time, thus providing greater spatial resolution in the system for model calibration.

The physical infrastructure connected to the digital twin consists mainly of flow and water level sensors in the central part of the sewer system and at the WRRF that provide online, real-time updates of these measurements, as well as status from actuators (pumps, valves, and gates) in the system.

The operator visualises simulation results on a dedicated website, separate from the SCADA system, showing flows in

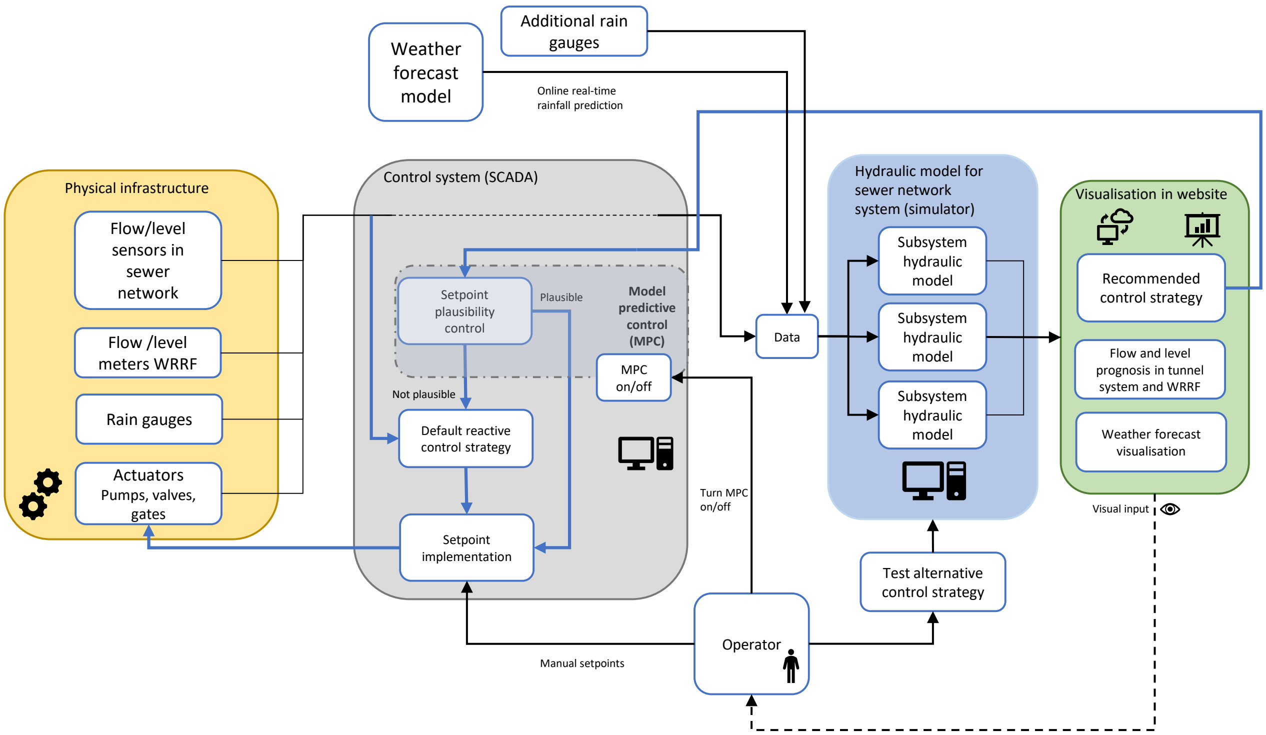


Figure 3.5 Sewer network digital twin structure and components in the current implementation as well as with model predictive control (work is now underway with the goal of implementing full real-time model predictive control).

the catchment, influent flow to the WRRF, and a comparison of predicted flows with both the default reactive control strategy (based on an empirical level-flow relationship) and a recommended control strategy (provided by the simulator) (Figure 3.6).

Both sets of predictions are for the next 2.5 days, based on the latest weather forecast. The recommended control strategy is updated every hour, resulting in new recommended set points. The operator then has the option to implement the recommended control strategy or to evaluate their own strategy in the simulator and evaluate those results. The future MPC will include plausibility control of the recommended control strategy before final implementation. If the plausibility check fails, the controller will revert to the default reactive control strategy.

To determine the recommended control strategy, the controller uses an algorithm to optimise a pumping scheme to achieve inflow to the WRRF that is as constant as possible during the next 12 hours, while considering boundary conditions, such as allowable water levels in the sewer (to avoid CSOs), pumping capacity, pumping conditions etc. To avoid discrepancies between current measured values of flows and water levels in the sewer network and pumping stations and the corresponding values in the model, corrections are required. Data assimilation techniques are therefore used to initialise the model using real-time data from the sewer network before each simulation and to adjust the forecast according to identified patterns of deviations between measured and simulated values.

CHALLENGES

Predicting the future in such a varying system is always a challenge. Since accurate flow and level predictions depend on accurate rain forecasting, this has presented a major challenge due to the uncertainty of weather forecasting models.

Another challenge is to represent variations in extraneous water from different sources (such as direct inflow, infiltration water, drainage water, etc.) for different hydrological events both in the short term and the longer term. This requires a modelling approach that allows an accurate representation of the urban hydrology and of extraneous water impacts on the sewer system, not only for single events but continuously in time, using both hindcast and forecast data. Figure 3.6 illustrates the effect of the prediction time horizon on predicted flows by showing past predictions of WRRF inflow with two prediction horizons (0—1 h and 3—6 h in advance). The curves display the mean predicted flow values for the specified interval after the time of forecast which are saved at the time in the middle of the interval (i.e., the curve tagged “prediction horizon 0—1 h” shows the prediction made 0.5 h earlier for the mean flow during the next hour after

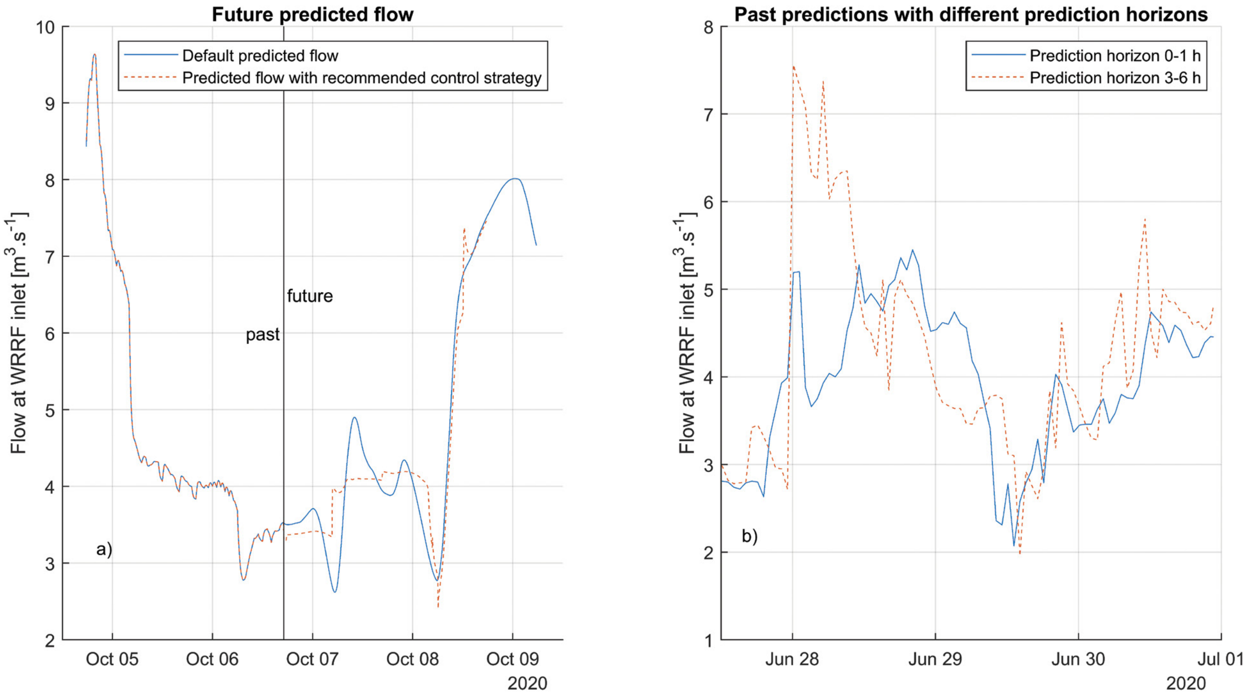


Figure 3.6 (a) Example of model predictions for influent flow to the WRRF using the default control strategy (solid line) and the recommended control strategy suggested by the model (dashed line); (b) Example of past model predictions with different prediction horizons (0 — 1 hours and 3 — 6 hours).

the time of forecast, while the curve tagged "prediction horizon 3—6 h" shows the prediction made 4.5 h earlier for the mean flow in the period 3—6 h after the time of forecast).

A large rainfall event was initially predicted to occur on 28 June, as suggested in the 3—6 h curve, but the rainfall volume and WRRF inflow were predicted to be much smaller when predicting 0—1 h in advance.

Data management has also been a challenge, mainly to create a secure and robust data transfer. Specific issues include time stamps when combining data from different systems (e.g., because of daylight savings time); ensuring data quality for precipitation data; managing the size of files for weather data which was initially too large for efficient transfer; and integration of the SCADA system, which normally only handles historical data, with local databases and interfaces for storage and visualisation of future predicted scenarios, respectively.

A critical part of the digital twin is the connection between the user and the models. An important aspect has been to develop confidence in the digital twin among the operators in the control room. A model that is sufficiently accurate and whose limitations are well understood is therefore important. A related issue has been the question of how to display uncertainty in predictions to the user (since weather predictions can include many different scenarios and many pumping strategies can be tested).

RESULTS AND BENEFITS

The digital twin allows for increased confidence in decision-making as the effects of a change in control strategy can be visualised quickly. Simulations of the real-time control strategies indicate that CSO spill events can (ideally) be reduced by 50% in the Gothenburg case. In Helsingborg, CSO spill events have already been reduced by 32% before implementation of MPC, by thoroughly studying the sewer system and exploiting opportunities that emerged. Benefits at the WRRF include:

(1) A more constant influent load which leads to more stable treatment processes;

(2) Lower risk of reaching critical load situations (e.g., by pumping water from the tunnel system before large flows are expected, so that more volume is available for attenuation, thus avoiding large flow peaks that can be hard to handle); and

(3) An increased margin for handling issues with pumps or accumulation of material in the coarse screens at the inlet of the WRRF. A major benefit with MPC is that the control system takes into account both current conditions and predicted future conditions.

Digital twin for a water resource recovery facility

Jacobs partnered with PUB, Singapore's National Water Agency in an R&D project to create a digital twin of the Changi Water Reclamation Plant (CWRP). This digital twin provides new insights into the ongoing operations and maintenance of the facility, supporting increased productivity and enhancing operational resilience.

The digital twin is envisioned as an advisory tool without direct control capabilities, which can grow into control functions as staff gain confidence in the tool. It has automated data inputs directly from both the SCADA system and the laboratory information management system (LIMS), as well as auto-calibration and soft sensor capabilities. This model is expected to assist PUB in simulated scenarios to test and calibrate strategies to enhance the plant's water quality as well as optimise its energy and chemical consumption. The use of the model is also in line with PUB's goal of tapping smart technologies to increase productivity and improve resilience in operations. A process flow diagram of both CWRP and the digital twin simulation is shown in Figure 3.7.

COMPONENTS

The digital twin includes dynamic semi-mechanistic models of the CWRP whole plant hydraulics, controls, and processes. Hydraulics and controls are simulated within Jacobs' Replica™ simulator. Processes are modelled using Dynamita's Sumo© whole plant simulator with direct communication between Replica and Sumo. The inputs to the digital twin are only those currently available at the facility. Data-driven (machine learning) influent predictions are used as part of the digital twin functionality for predicting plant performance up to five days into the future.

The influent soft sensor workflow of the digital twin is illustrated in Figure 3.8. Input data to the model are first conditioned by removal of "bad" data, defined as negative or zero data that are not consistent with other related data. A historical test is then made to remove all data that are not within normal variations (± 2.5*log normal interquartile range, IQR). The digital twin inputs from SCADA are the various flows measured in the facility, the online primary effluent ammonia values, air rates to the various bioreactor zones, and other relevant operational set points. From the data, in combination with the LIMS data, a dynamic raw sewage influent file is generated, thus creating a soft sensor of the actual influent to the facility. Air rates and operations set points read from SCADA are directly input into the digital twin.

The initial calibration of the digital twin was done manually; control and hydraulics calibrations were based on actual measurements. The process calibration was first done in a steady state, then dynamic calibration was done based upon the first 6 months of a full data set. The digital twin also included limited auto-calibration while running. This was accomplished as measurements became available where, for example, primary effluent laboratory total suspended solids (TSS) and chemical oxygen demand (COD) data were used to calibrate both the primary clarifier TSS removal and the soluble COD/COD fraction in the raw sewage based on the most recent performance data.

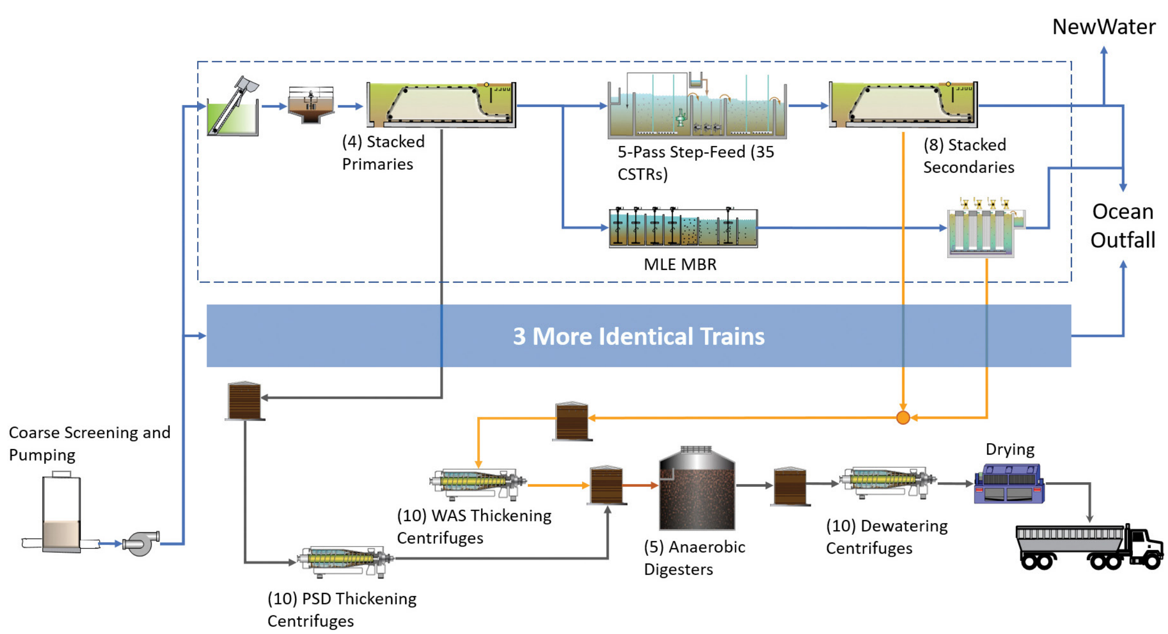


Figure 3.7 Process flow diagram of Changi WRP and scope of digital twin simulation (inside dashed line). The digital twin includes the four water treatment trains.

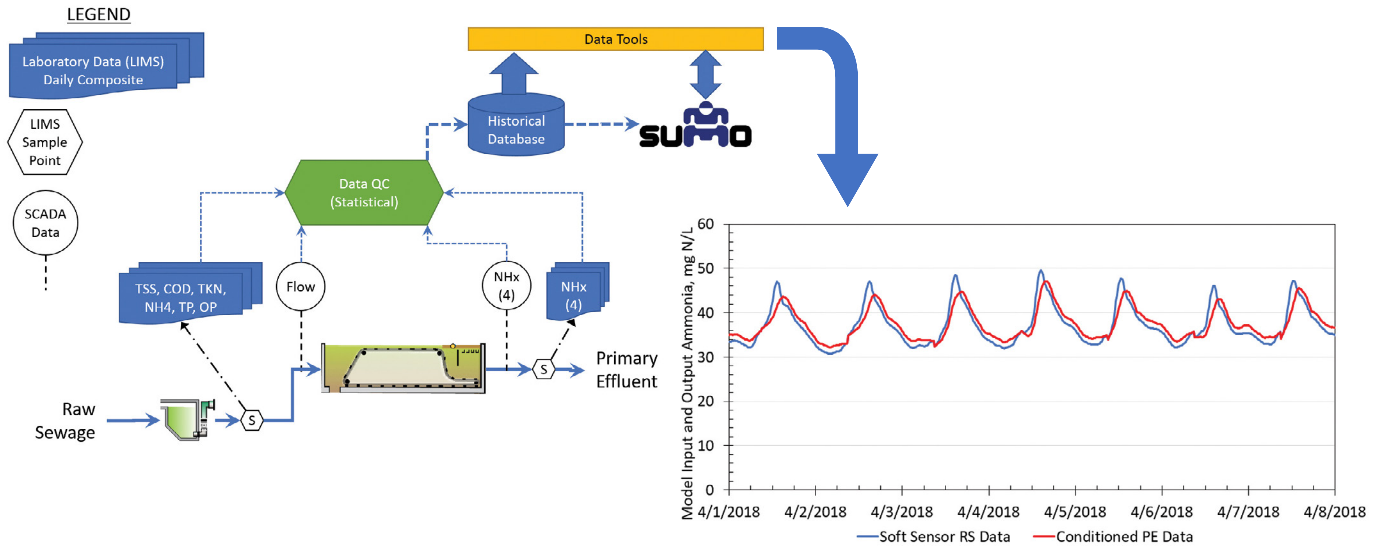


Figure 3.8 Data flow schematic within digital twin for influent load development for the influent "soft sensor"

CHALLENGES

The largest challenge in this project is speed of computation, as results are needed for both current operation/model comparison functions, scenario evaluation, and future predictions using Monte Carlo techniques. This must be accomplished within 24 hours of receiving the relevant data. For example, the process (Sumo) model is over 40 MB in size and the Replica hydraulics and control model is over 170MB in size, thus giving an indication of the model complexity.

RESULTS AND BENEFITS

There are three primary functionalities of the digital twin. First, comparison between model predictions and measured data can be used to highlight areas which require particular attention by operators and maintenance staff, thus minimizing their efforts. For example, a warning will be issued if one of the primary effluent online ammonia probes differ significantly from the others and the model or if the laboratory primary sludge total solids does not match up with the model results.

For the latter, a check of the primary sludge flow meter or confirmation of the laboratory results would be required. The second use of the calibrated model includes evaluation of various operational scenarios, both operator-defined, and fixed. Lastly, the auto-calibrated model is being used to predict the likelihood of future events at the wastewater facility up to five days in the future, a "wastewater weather forecast" that operations can use to help proactively operate the facility.

Conclusions

The case studies presented in this work demonstrate the water industry's eagerness to move to digital solutions such as digital twins to improve the performance of infrastructure. This section demonstrates that large potential savings can be derived from better process automation, online optimisation, fault detection, maintenance, and more proactive operation. However, there is no "one recipe fits all" for digital twin applications, and this may hinder the rate at which the full potential of this tool can be realised. Good practice protocols need to be established to support digital twin developers. The protocols should provide answers to the following questions:

- **Purpose of the digital twin:** what are the objectives and the expected gains? Can the objectives be attained with the existing infrastructure and available data?
- **Data:** what is the impact of data quality and quantity on the applicability of a digital twin? How to feed data to it? How to calibrate it?
- **Model:** how to identify the right modelling approach for a given objective? What is the right level of detail (granularity) in a model? How to be confident that model predictions are acceptable for a digital twin? How to manage model uncertainty?
- **Robustness of the digital twin:** how does the digital twin continue to evolve? How to perform model validation using available online data?

3.3 Digital dynamic resilience for wastewater treatment processes:

Exploiting real data for long-term resilience

With societal (e.g., COVID-19) and climatic factors coinciding with increasing regulatory pressures, the resilience of wastewater networks and infrastructure is reducing globally. Historically, water companies have relied on reserve capacity, but are now being forced to manage extreme dynamic responses as wastewater assets react new stressors. An example of this is rainfall intensity, which has already increased from 12 to 24 % (Fischer et al., 2014) and has been commonly followed by prolonged dry periods driven by climate change (NASA, 2016). These dramatic variations generate complex dynamic changes and can drastically reduce the resilience of networks and, in turn, water resource recovery facilities (WRRF). Without unified quantification methods, it is impossible to compare the resilience of different wastewater processes or systems, when exposed to climate and societal stressors. Also, the complexity of dynamic changes makes it virtually impossible to simulate the numerous dynamic factors that combined cause these reductions in resilience.

To avoid possibly cumbersome modelling and simulation of possible scenarios, dynamic resilience uses actual WRRF data to visualise zones of process stress and resilience as a heat map. The methods presented in this section separate stressors present in water company data as the 'cause' of an event, and the 'effect', whether a WRRF experiences process stress or resilience as a result. This separation of stressors and process stress is key to isolating the cause of an event then its manifestation as the effect. This separation of the stressor and process stress requires data feedback from WRRF process and systems. Data generated by water companies is ideal for computation of dynamic resilience in response to extreme events, where the cause (stressor) and effect (process stress) can be seperated. This data is generated in vast quantities daily (typically < 1 h intervals) and is used to make operational decisions, but often remains in silos, or as described by Aguado et al. (2021), 'data graveyards'. Another challenge is that WRRF data can be difficult to interpret when instruments are poorly maintained or installed incorrectly (Grievson, 2020), but meaningful observations can still be made to interpret resilience metrics. Therefore, this data must be exploited to understand the dynamic resilience. Without data (real-time or from silos) connecting a WRRF or process to resilience, evaluations may remain theoretical and iterative, which is computationally intensive.

This section starts by describing the historical context of resilience, before moving onto the dynamic resilience of wastewater processes and WRRF. Real-world examples of societal and process related resilience from industrial and academic experts are provided, which discuss the challenge of generating data under uncertainty of ageing infrastructure. A case study is then presented on dynamic resilience using actual WRRF data. The case study shows how actual water company instrument data could be used to evaluate stressors and process stresses independently. The outcomes of dynamic resilience case studies are then presented as a series of contoured heat maps.

This section covers the fundamental and historical context of resilience and its evolution within the water industry globally

THE FUNDAMENTALS OF RESILIENCE FOR WASTEWATER TREATMENT

The word resilience dates back to the 1620s, described as the 'act of rebounding', i.e., the ability to recover from an external event (Harper, 2019). By definition, the term rebound indicates that an external event causes a system to deviate from its initial reference position, and must be reconciled for an event to be considered complete. Classical resilience theory, originated from social sciences, focusing on ecological resilience in predator-prey systems, where a system absorbs change until moving to a completely new state as described by Canadian ecologist Crawford Holling (Holling, 1973). This description was later expanded to include engineering resilience, which focuses on maintaining stability close to an equilibrium stready state. An example of engineering resilience is a set point, where the system is controlled to maintain predictable operation close to user defined optimum.

Wastewater systems can be a combination of the two types of resilience (ecological and engineered). Extraneous events shift the wastewater collection systems to a new operating state and, as result, the system performance is controlled, in order to manage undesirable process upsets or disturbances. Using the example of an activated sludge (AS) system, the microbial ecology is controlled by an engineered system of mechanical/ electrical components (aeration and pumps). Therefore, a more accurate classification of a biological treatment process would be engineered ecology, where process engineers apply engineering principals to control biological ecology. Wastewater catchments can also behave ecologically, where systems can move to a completely different operational state as populations vary, or extreme events occur. Many of these changes in the catchment have a direct effect on the WRRF engineered ecology due to dramatic operational state changes which reduce resilience by

1
2
3
4
5

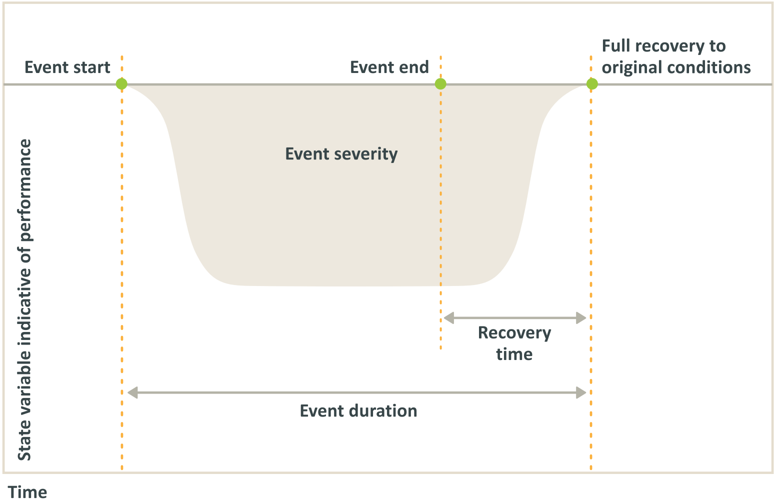


Figure 3.9 Resilience to a stressor presented by Juan-García et al. (2017) adapted from Mugume et al. (2015)

eroding operational safety factors (reserve capacity). Therefore, to evaluate the dynamics of resilience, we must understand that we are actually managing engineered ecology (i.e., engineered systems that aims to control microbial ecology). This is where, dynamic resilience extends existing theories: it accepts that dynamic changes in resilience are both, positive (resilience) and negative (stress).

Another difficulty for resilience has been selecting an appropriate definition. A commonly accepted definition was proposed by Walker et al. (2004) as:

'Resilience is the capacity of a system to absorb disturbance and reorganise while undergoing change so as to retain essentially the same pre-disturbance process, form, identity, and feedbacks'

This definition suggests that a system undergoing change adapts to an event before returning to its original condition (Figure 3.9). However, the definition does not account for the complexity associated with biological wastewater systems, which have numerous states (ecological and engineered), some temporary and some permanent. Resilience is also reduced further, by additional novel extreme states occurring outside of diurnal and seasonal patterns, which related to climate change and modifications to human behaviour. When normal operation is combined with novel stressors, simulation of their response is cumbersome, vast numbers of theoretical iteratons. Therefore, if WRRF data has sufficient resolution (adequate range and quantity), it could hold clues to these novel states, where many parameters are affected by stressors, and the process stress/resilience response (dynamic resilience).

There are many interrelated factors must be considered to evaluate the dynamics of resilience. Therefore, it is crucial to not only consider wastewater infrastructure and WRRF, but also factors that lead to stressors under the following headings:

1. **P**olitical resilience
2. **E**conomic resilience
3. **S**ocial resilience
4. **T**echnological resilience
5. **E**nvironmental resilience
6. **L**egal resilience

PESTEL factors and their interactions are responsible for many of the stressors exerted on ageing assets and infrastructure. However, in times of societal adaptaion and change, numerous PESTEL factors can become interrelated, generating significant complexity and uncertainty. Nevertheless, pre-existing clues of dynamic stressors and process stresses are apparent in water company transactional data and surveillance data used by government agencies to monitor COVID-19. This section focuses on stressors occuring from a wastewater catchment that have a significant effect on the stresses generated within a specific WRRF. As shown in Figure 3.10, these direct external stressors have the greatest potential to have significant influence on wastewater volumes and concentrations.

THE EVOLUTION OF RESILIENCE AS A CONCEPT IN THE WATER SECTOR

Since the original research of Holling (1973) and Walker et al. (2004), interest in resilience has grown in all subject areas. The investigation of resilience as a determinant for water and wastewater systems has been well documented, see Butler et al. (2014). A reasonable level of success has been achieved in water supply resilience, although this is often not suitable for the complex multi-variate mechanisms, and often competing processes, associated with wastewater delivery and treatment. When wastewater is generated numerous factors are involved, e.g., urban creep or groundwater infiltration to sewers can dramatically increase the flow to a receiving WRRF. The randomness and unpredictability of changes leads to significant reductions in resilience over a short period of time (i.e., hrs). Attempts have been made to circumvent this with research focussing more toward general resilience (Sweetapple et al., 2022a), which utilises a more systems of systems approach that can be highly complex. From an industry perspective, there has been growing interest in resilience as a concept by the UK water companies; however, the principles have not been fully embedded. This was exemplified by the Ofwat price review 2019 (PR19), where only two water companies provided evidence of securing the long-term resilience of their assets (Ofwat, 2019). Therefore indicating that the challenge is not the generation of greater knowledge; but applying resilience theory and embedding its principles into daily operation.

Figure 3.10 Direct (external) and indirect (internal) stressors adapted from Butler et al. (2014).

For this to happen, dynamic changes in operation that influence resilience should consider the use of existing WRRF data to build a knowledge base of past stressors and process stresses/ resilience (dynamic resilience) in preparation to future events. For instance, to avoid potentially damaging pollution incidents, it is crucial to understand how stressors manifest leading up to such dramatic changes in wastewater volume and concentration. To understand these dynamic events, the interaction of stressors and process stresses/resilience should be considered as dynamic resilience (Figure 3.11). The principle of dynamic resilience in Figure 3.11 shows the stressor as having a bell-shaped peak and the process stresses generated as a well-defined peak concentration. Differences between stressors (cause) and process stresses (effect) occur because the WRRF behaves differently to the generated stressor due to recirculations and sludge extractions. Understanding the magnitude and duration of events has benefits, gives rise to the possibility of classifying stressor influence at the WRRF or for separate processes. It also allows for a reaction time between an event occurring (stressor) and its effect on the process (stress).

The definition of dynamic resilience used in this section, as shown in Figure 3.10 and also presented in the International Water Association (IWA) Modelling and Integrated Assessment (MIA) Specialist Group (SG) webinar (Holloway et al. 2021) is detailed below as:

"The dynamic, temporal variation of stressors and process stresses (and resilience) in response to events outside of standard operating conditions"

Aims and objectives

This section starts with expert perceptions of resilience, then approaches the challenge of evaluating the dynamic resilience in the context of actual WRRF data. It does this by demonstrating some common pitfalls of using actual WRRF data, then how stressors can be extracted for the evaluation of process stress/ resilience effects. Dynamic resilience is demonstrated through a case study using 10 years of data to visualise a specific event. The broader impacts of dynamic resilience are also discussed, along with how methods could be incorporated to provide a holistic, resilience based approach which is applicable to water companies (Holloway et al., 2021).

Expert's perspectives on resilience

The following two sections provide an overview of the resilience based challenges faced by water companies and those interpreting data from wastewater-based epidemiology.

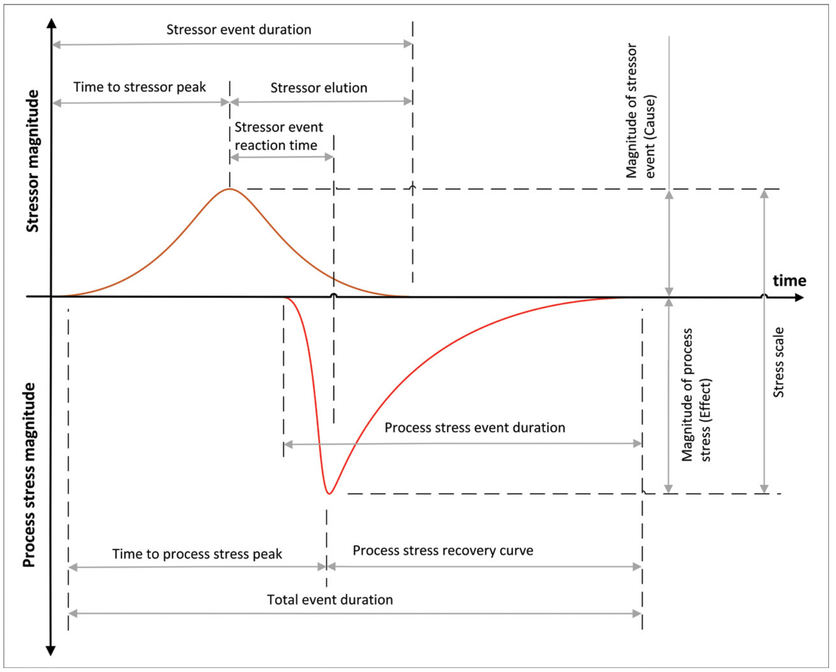


Figure 3.11 Dynamic resilience diagram showing the separation of stressors and process stresses.

RESILIENCE FROM A WATER COMPANY PERSPECTIVE

Wastewater flow data is central to understanding the dynamic resilience of wastewater infrastructure and WRRF and to understand the pressure exerted by wastewater catchments (stressors). Water companies in the UK are making great strides in understanding how their assets react to climate change while incorporating meteorological and demographic data. In this section, Dr Ben Martin speaks of the challenges associated with asset and infrastructure resilience and how Thames Water are embedding dynamic digital systems.

Improving asset and infrastructure resilience is a significant challenge for the water industry as operational disruptions become more common and difficult to predict. The most significant of these disruptions are extreme weather events, which have recently delivered a month's worth of rain in 1 hour. Most sewers and wastewater treatment plants are many decades old and were simply not designed to cope with the loads and temperature fluctuations recorded in recent years. Such weather events have been coupled with the COVID pandemic, which resulted in spatial disruptions to normal wastewater loads as the mobility of the population reduced during lockdowns.

Thames Water is currently working on a host of data-driven projects to increase the resilience of ageing asset and infrastructure base. These digital initiatives will output control systems that can manage disruptive loads more effectively and efficiently. Additionally, a digital twin is being developed for Beckton sewage treatment works, the largest wastewater treatment plant in the UK. This brings together a number of hydraulic, pneumatic, and biological models to digitally represent its physical assets. This is expected to deliver a 10—20% reduction in workforce planning, a 30—50% reduction in predictive maintenance, and a 20—40% reduction in reactive maintenance.

A broad rollout of connected digital monitoring instruments throughout assets and infrastructure is planned. These include prediction systems so that assets can be maintained and replaced long before failure, machine learning to enable autonomous waste catchments, and open data frameworks for sharing across water companies. Linkages will also be established with meteorological and demographic datasets. These will inform operational control strategies that can respond in real time, or even ahead of time, to enable rapid recovery from events. These dynamic digital systems will ensure that as disruptions become the norm, wastewater can still be treated to exceed environmental targets.

RESILIENCE OF WATER MANAGEMENT SYSTEMS FOR PUBLIC HEALTH PROTECTION

Social resilience has a key role in the protection of public health. In this section, Dr Matthew Wade writes of the challenges associated with asset and infrastructure resilience and its impact on monitoring wastewater-based epidemiology (WBE).

Developments in the water sector to embrace Industry 4.0 and digital transformation philosophy have demonstrated an evolution in the function of wastewater. While the focus of treatment and transport infrastructure has been generally on protecting the environment from harmful pollutants and, more recently, resource and energy recovery from wastewater (Daigger, 2009; Guest et al., 2009; Kehrein et al., 2020), the COVID-19 pandemic has given greater visibility to a broader function of wastewater systems, public health protection.

Wastewater monitoring as a tool for public health intervention dates to the mid-19th Century, when physician John Snow mapped data of cholera incidence in Soho, London to determine the source of the outbreak (a water pump contaminated with sewage) (Tulchinsky, 2018). The detection of sewage borne indicators of public health, commonly known as WBE, has been used to monitor a range of targets from poliovirus to illicit drug use in urban centres (Larsen et al., 2021). From the onset of the COVID-19 pandemic, the SARS-CoV-2 viral RNA was shown to both be detectable and quantifiable in sewer samples collected throughout the sewer network, typically at the inlet of treatment works, within-network, or at near-source (e.g., building scale) (Ahmed et al., 2020; Sweetapple et al., 2022b).

Given the evolution of the virus, subsequent work has also demonstrated the ability to detect its variants (Crits-Christoph et al., 2021). Once evidenced, the challenge for those working with WBE for COVID-19 was to determine the value of these datasets for public health policy and decision-making. The rapid uplift of COVID-19 science in wastewater and the fragmented nature of its utilisation across the globe means that the true value proposition of WBE as tool to complement existing measures of public health remains unproven and uncertain. Factors influencing its resilience include a lack of empirical data to understand and mitigate for wide range of uncertainties associated with the data from WBE (Wade et al., 2022), a robust understanding of the relationship between the target marker(s) of public health (Mao et al., 2020), and the general lack of standardisation and protocols to enable WBE to be implemented as a function of public health policy (Wu et al. 2021). Nevertheless, WBE has great potential as a tool for public health protection beyond COVID and across a broad range of targets (e.g., lifestyle chemicals, pathogens, metabolites of health), settings (e.g., community-wide, critical infrastructure), environments (e.g., urban centres, low-income settings), and functions (e.g., rapid response, long-term surveillance of health trends, targeted monitoring).

For WBE to be useful in the long-term, its research, development, and use needs to be considered together with the wider efforts to ensure infrastructural and data resilience. This should be viewed as a resilience of the systems where information is collected, as a failing sewer network will inevitably lead to greater measurement uncertainty. Additionally, it is crucial to consider its ability to ensure public health resilience as the information acquired by WBE must be reliable and usable by the stakeholders receiving it. This effort must be global. Water and disease know no borders and ensuring the resilience of water systems in highly resourced regions (e.g., those with the means to embrace digital tools to manage the increasingly voluminous and valuable data streams), must be matched by initiatives to maximise the value and potential for WBE in low-income settings (e.g., development of low-cost but smart technologies (Gwenzi, 2022)).

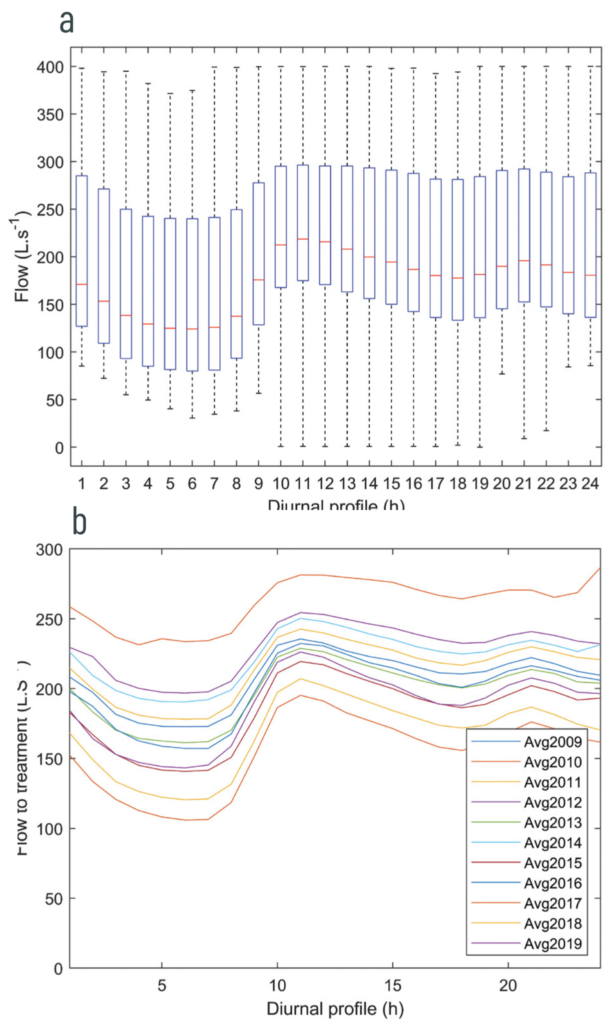


Figure 3.12 10 years of hourly flow data over a 24 h period aggregated into boxplots (a) and average flows by year (b).

The importance of data for the evaluation of dynamic resilience

This section evaluates the challenge of using actual stochastic data, presents possible methods for generating valuable insights, and finally evaluates dynamic resilience. It finishes with the future perspectives of dynamic resilience and how the best possible environmental outcomes can be achieved.

DATA AND THE IMPORTANCE OF SCALE: CHANGING FROM MICRO TO MACRO ANALYSIS

Before examples of dynamic resilience are presented, it is crucial to note some of the pitfalls of using actual WRRF data. The use of large datasets captured over multiple years can be overwhelming (e.g., 174,720 data points when logged at 15 min intervals over 5 years). Conventional concentration-time plots are not sufficient, providing poor resolution due to the variation frequency of actual stochastic WRRF data. An example of this can be seen in Figure 3.12a, where aggregated boxplots of hourly WRRF influent flow over 10 years show significant variability. It also indicates that this data is not normal or lognormal, so when averages are taken (Figure 3.12b), the central tendency is skewed providing inaccurate flow predictions. Therefore, significant uncertainty can be created when making statistical assumptions, particularly when data is being used as a timeseries for data-driven models or other statistical evaluations (Newhart et al., 2019). To avoid introducing significant uncertainty the entire data set must be considered (macro analysis). This also avoids the temptation to be miopic, analysing each discrete variation (micro analysis) which can also lead to erroneous outputs when performing simulations.

TOWARDS DIGITAL DYNAMIC RESILIENCE OF WRRF

The dynamics of digital systems have long been the interest of those working on solutions where real-world interfacing (sensors) is crucial for process control. Examples of this occur throughout robotic manufacturing and high-value fluids such as oil and gas. Unfortunately, being termed as a 'waste' rather than a 'product' places less importance on the value of the end product, resulting in less instrumentation (less concequential loss). Fortunately, for water companies, they generate data intensively from WRRF and processes. In decentralised rural locations, instruments are commonly used for compliance monitoring and are less likely to be used for real-time process control. However, larger centralised WRRF often combine monitoring of complicane with signals/data generated for process control. These larger WRRF processes have vast data silos spanning years or even decades. Much of this data generated for larger centralised WRRF holds clues to how dynamic stressors emerge and wheter they generate process stress/resilience. Therefore, to avoid exhaustive iterations for modelling and simulation of scenarios, this data could be used as a timeseries to evaluate the dynamic resilience and generate knowledge of past and present events through empirical and mechanistic modelling. Modelling improves the context of events, and the data permits evaluating them relative to actual conditions. Over time it may be possible to make predictions on how future stressors may influence specific WRRF processes. A data-driven dynamic resilience philosophy therefore uses actual WRRF data and established modelling practices to compute the impact of a stressor, then process related stresses/resilience. This takes the focus from the intensive simulation of event-based scenarios (stressors), to evaluating actual scenarios in the context WRRF instrument data.

Case study: dynamic resilience using 10 years of data from a WRRF in the south of the UK

In this section, 10 years of data has been used to (1) identify significant events through the evaluation of prominence and dominance, (2) extract a standard operating condition for an existing WRRF and (3) provide examples of dynamic resilience visualisations.

EVALUATING THE DOMINANCE AND PROMINENCE OF EVENTS UNDER DYNAMIC CONDITIONS

Events or stressors are typically characterised by their magnitude (the prominance of the variation). The most significant event magnitudes can be isolated then scaled for direct comparison. This can be done for both the stressor (cause) and process stresses/resilience (effect) which can be evaluated independently to isolate only the most severe, as shown in Figure 3.13. When a significant event has been isolated using prominence, the event dominance can be estimated as the difference between the stressor exerted to the WRRF and the resultant process stresses/resilience. The difference in time between the stressor and process stress peaks allows estimation of the reaction time for applying interventions (maximum time to react). It is then possible to classify events as stressors and process stresses to estimate a reaction time for future events (i.e., the time between the stressor peak and the process stress occurring). Without event insight, it may not possible to learn from interventions applied to WRRF that reduce process-related stresses.

Examples of prominence are shown in Figure 3.13a as time-based examples of stressor prominence (brown line) and process stresses prominence (red line). Significant events can be seen in Figure 3.13a, when stressors are close to 1 and the resulting process stresses approach -1. The event dominance is shown in Figure 13b (difference between the stressor and process stresses), allowing for isolation of events that have the most significant effect on process performance. Therefore, using prominence and dominance based analysis, stressor events can be isolated, then evaluated based on the effect the event has on the WRRF (process stress/resilience).

FAILURE AS A CONSEQUENCE OF NORMAL OPERATION

As much as we don't like to admit it, failure is a consequence of standard operation, originating from specific and operational changes or stressor influences. Therefore, as provided in the definition of resilience, failure also has a magnitude, with the extent of failure present over a scale. An example of this is failure to remove sludge from a system, which over time accumulates solids until a compliance breach occurs. This is far less significant, than a toxicity event that essentially kills off microbes in a biological system. Therefore, the extent of failure is crucial when evaluating resilience, meaning dynamic metrics then become scalable from a standard operating condition (where the process normally operates). Therefore, when we consider WRRF process failure, we must first consider the failure extent, and what constitutes a failure (*a failure of what and to what extent?*). For example, operational staff may consider the the color of the process fluids or visual tests to evaluate process conditions. These are empirical observations of indirect/internal stressors that prevent the process from functioning and are common. However, process scientists and engineers take more of a theoretical method of diagnostics, with failure defined as a compliance breach. Therefore, when considering events, it is essential to appreciate both empirical and theretical thought processes.

WRRF processes are subject to diurnal variation, which is also dynamic, but typical of standard operation for a specific wastewater catchment. Therefore, it differs from original WRRF process engineering design information, particularly when urban creep and populations in the wastewater catchment increase. The difference with dynamic resilience is that it takes the standard operation from actual operational data, rather than historial design information. This extracted condition is called the standard operating condition (SOC) relating specifically to the nuances of a catchment or WRRF. This can be done by using clustering methods to extract three flow conditions similar to that of Borzooei et al. (2020). The clustered outputs for a particular WRRF are shown in Figure 3.14a, and the extracted SOC in Figure 3.14b alongside actual time-based data. This SOC is important when considering WRRF that have been retrofitted with supplementary processes to increase WRRF capacity, which is common practice internationally. The extent of process failure is classified as the process stress index (PSI) and is anything outside a WRRF specific SOC.

The outputs of the clustering shown in Figure 3.14 take an engineered resilience stance (Holling, 1996) and must be further elaborated to include a safety factor or degree of freedom to exclude normal variation (Figure 3.14b). Also, outputs based on failure should consider the balance of compliance risk to operating costs, with evaluations based on the WRRF meeting permitted limits at the lowest operating cost. However, caution should be applied to a minimum operational cost approach so the process does not become unstable or vulnerable to changes resulting from external stressors. Therefore, evaluating the magnitude of variation (stressor) as the extent of failure allows

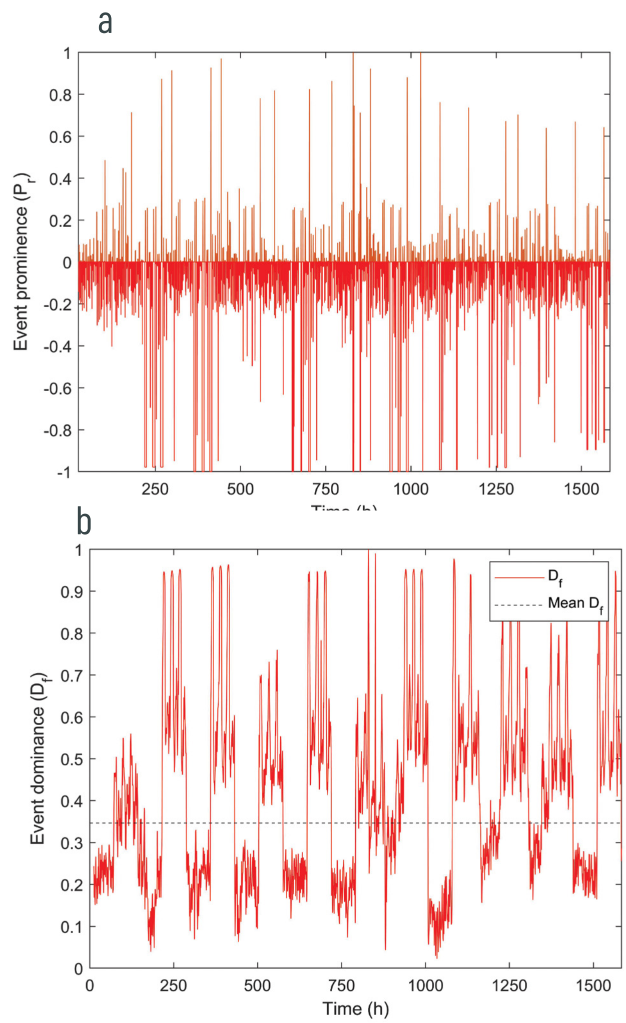


Figure 3.13 Prominence (a) and dominance (b) of stressors and process stresses adapted from Holloway et al., 2021.

failure prediction and numerous event classification possibilities based on process stress or resilience (i.e., dynamic resilience).

COMMUNICATION OF DYNAMIC RESILIENCE

Visual communication is often overlooked, which is possibly why Corominas et al. (2018) found that only 16% of publications on transforming data into knowledge reference a commercially available product. Therefore, it is possible that there is a distinct gap between the evaluation of water resilience and how it is communicated to those operating and maintaining wastewater processes. The communication of resilience data in the run up to a significant event can be challenging, particularly if there is no knowledge of past events and the associated interventions. Unfortunately, inadequate operational communication often comes as an incident investigation, such as the Longford gas plant explosion (Conlin and O'Meara, 2006). This highlights the

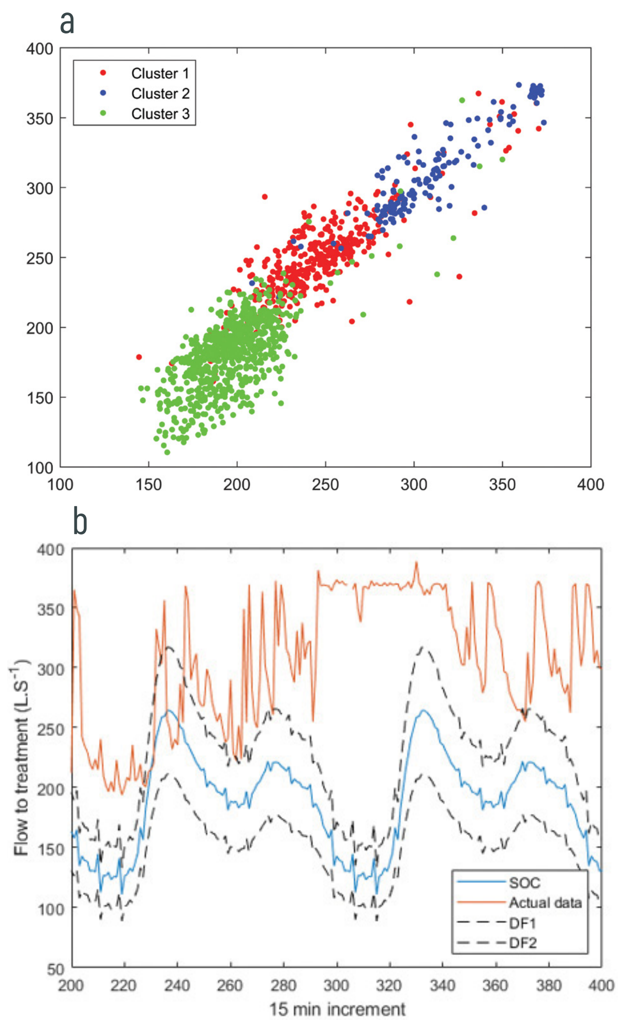


Figure 3.14 Examples of clustering methods for the generation (a) and outputted SOC including degrees for freedom (b).

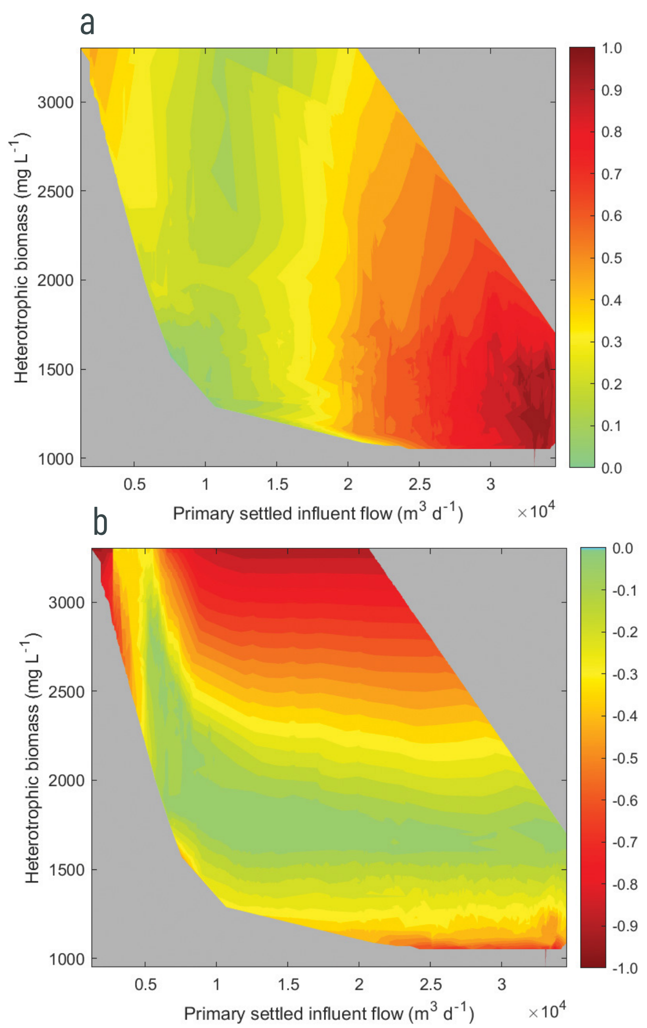


Figure 3.15 Influence of a stressor and generated process stresses using the SOW approach for activated sludge.

importance of linking resilience with complex modelled outputs in a communicable form for operational interpretation. It should also reflect the time-based dynamics of actual WRRF operating resilience.

To address this, the concept of dynamic resilience aims to incorporate time-based evaluations to visually communicate event severity to operational staff. self ordering windows (SOW) are used as a visualisation method for stressor and process stress/resilience observations. These SOW use a 48 hrs event window and, through transformation, plot the PSI as a contoured heat map. The SOW then becomes a significant event window based on its duration, prominence and dominance. In most cases, SOW represent a unique dynamic resilience fingerprint extracted from actual WRRF data, while keeping complex modelling practices out of sight (IWA ASM model series). The SOW principles also avoid having numerous number of iterations required for methods such as Monte-Carlo and other iterative simulations.

Examples of dynamic resilience are shown in Figure 3.15 to demonstrate the impact of flow on heterotrophic biomass concentrations, where Figure 3.15a shows the stressor and Figure 3.15b the process stresses resulting from the stressor. The stressor in Figure 3.15a shows concentrated zones at the highest flows and lowest concentrations and process stress at the highest concentration and middle of the flow range. The grey areas in each plot indicates zones of no data. Therefore, the concept of data richness is also crucial to SOW success, where inputs must reflect the time-based range and specific variation of the WRRF.

The main challenge of dynamic resilience is the duration of the time-based window selected for evaluations. It can also be extremely difficult to predict when an event starts and ends, with characteristics varying depending on specific or multiple stressors. This will be crucial for the expansion of dynamic resilience to incorporate real-time control (RTC) for live management of stressors and process stresses/resilience generated at the WRRF. An extension of that could be a traffic light system, as shown in Figure 3.16, mapping the process stresses, but also communicating them for the application automated interventions.

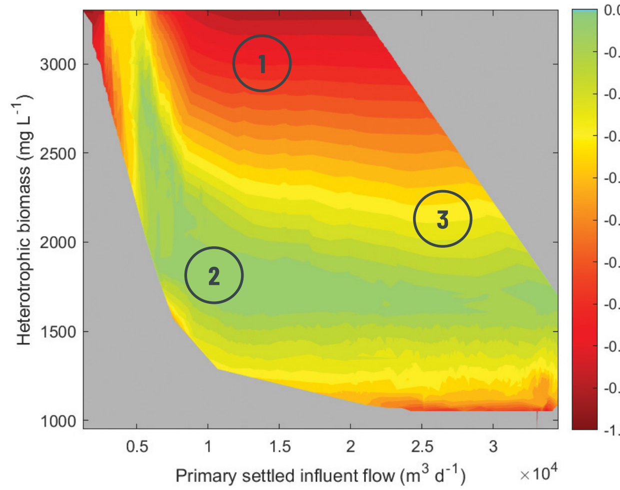
Summary of dynamic resilience methods

The dynamic resilience approach proposed in this section has demonstrated the possibility of using actual WRRF data to understand and communicate dynamic resilience. The methods presented used prominence and dominance to isolate significant events, then a macro data analysis approach to extract a dynamic SOC. Actual data points were used to scale system failure magnitude and compute the dynamic resilience of a specific WRRF as a SOW, which reflects specific process nuances in response to significant events. This includes the possibility of isolating novel events generated by climate or societal change. However, the main challenge for dynamic resilience is selecting a suitable time over which dynamic resilience is monitored, as this can dramatically influence SOW output and affect the classification of events.

Overall, the dynamic resilience methodology has provided a possible link between resilience, data-driven modelling and visualisations through time based contoured heat plots (SOW). It is hoped that these methods could eventually close gap between evaluating and modelling resilience and its communication to wastewater operators within water companies globally. However, the methods presented are limited to countries that (1) have the instrumentation installed in WRRF or networks and (2) have the capacity and knowledge to maintain these instruments.

Reflection on dynamic resilience for an uncertain future

The future of the planet is reliant on how resilient the human race can be to changes in the climate and rapidly emerging stressors. As we face increasing uncertainty from political, social, and environmental factors, water companies and government agencies are forced to manage the dynamics of resilience resulting from changes outside of their control. Although many theoretical methodologies of resilience have been proposed, a unified approach has not yet been developed to (1) satisfy the dynamics of resilience that occurs from an actual WRRF and (2) communicate outputs to operational and maintenance staff.



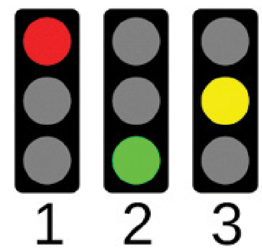


Figure 3.16 Process stress analysis using the SOW principles, adapted from Holloway et al., 2021.

At the time of writing, the Russian invasion of Ukraine is causing significant political and economic instability. This, again, will likely generate novel stressors that result in process stresses in wastewater assets and infrastructure as transient populations exit Ukraine to neighbouring countries. These countries will now be subject to increased demand for clean water and increased capacity to treat additional wastewaters generated by those that have exited Ukraine. It is extremely likely (Pörtner et al., 2022) that we will continue to see the emergence of novel, rapidly emerging stressors, and if we continue on the same path, there is high confidence that their occurrences will increase.

It is also important to consider the factors that have contributed to stressors in recent history. Reflecting on the past 3 years, the following factors/stressors have emerged globally:

1. **February 2022**: IPCC WGII sixth assessment report predicts with high confidence global increases in inland flooding, flood and storm damage in coastal areas and damages to infrastructure (Pörtner et al., 2022).

2. **February 2022**: the Russian invasion of Ukraine caused migration into neighbouring European countries.

3. **March 2020 to present**: the COVID-19 pandemic escalates, causing unprecedented damage to human health, global economies and freedom of movement (Ramos and Hynes, 2020).

1 2 3 4 5

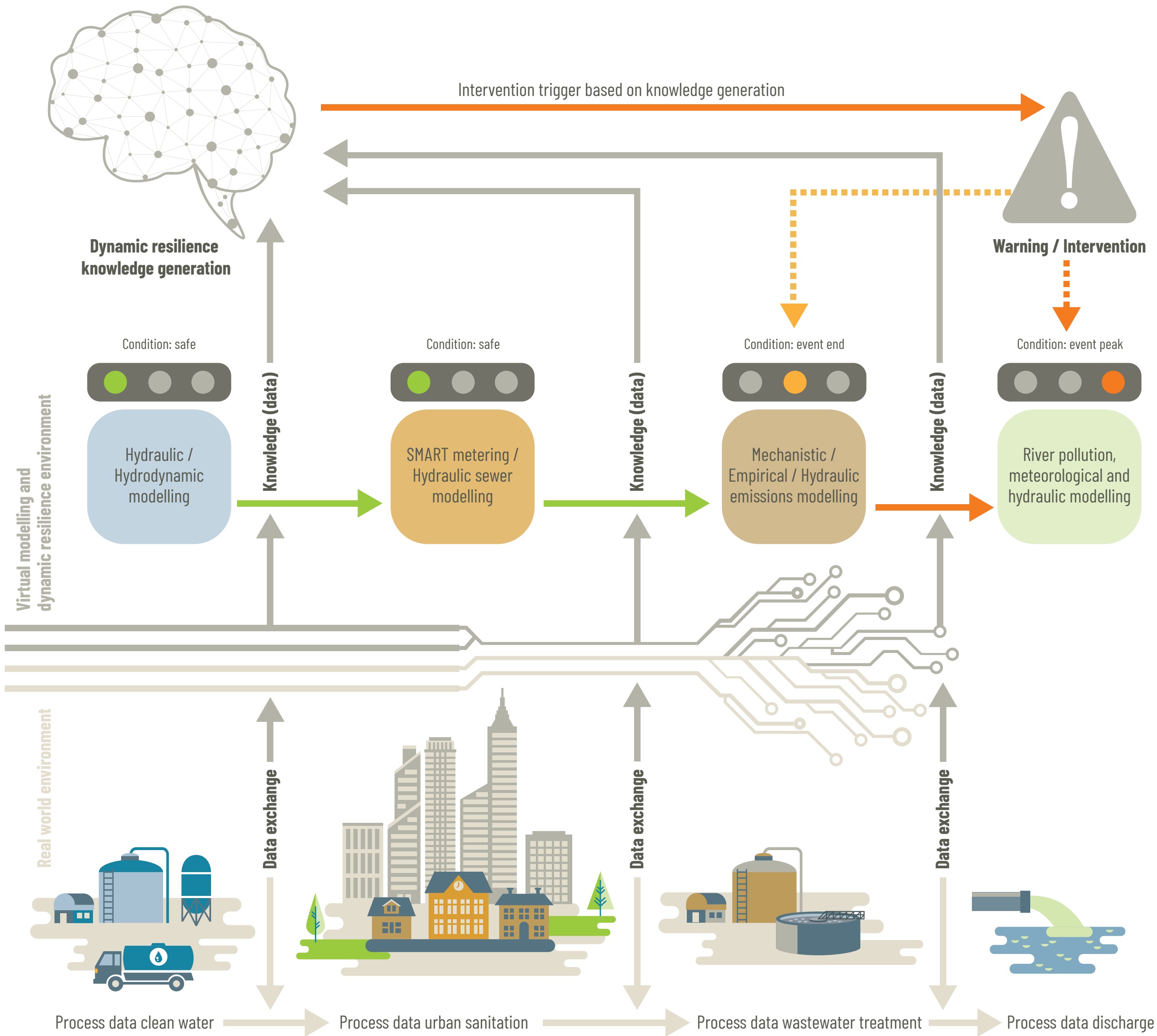


Figure 3.17 Dynamic resilience for autonomy of the urban water cycle under the precept of Industry 4.0.

The above emphasises the absolute need to understand the dynamic resilience of not just in assets and infrastructure but also due to societal change. Dynamic resilience has the premise of providing the means to understand the real-time adaptive capacity by monitoring stressors and process stresses/resilience directly from WRRF and associated processes (instruments). We must consider how existing instruments and data from these instruments can drive us toward Industry 4.0. Instrumentation combined with an understanding of dynamic resilience could allow us to understand and to adapt resilience in response to novel global events, while understanding where improvements can be made.

We already have many connected networks for intensive data exchange, along with large data silos (historical knowledge). Using these tools, it may be possible to access the real-time dynamic resilience of not just WRRF, but any interconnected infrastructure network. Therefore, the future of dynamic resilience and dynamic processes should embrace the possibility of improving resilience through digitally connected assets and infrastructure as the beating heart of the modern world (Figure 3.17).

Conclusion

Transforming information to insight

As shown by the case studies in this chapter, AI, DT and DR are three powerful tools that can be used by the international water industry to take advantages of 'big data'. These are also tools that water professionals now have at their fingertips. Most utilities and system designers/engineers have more data available than they can readily integrate and analyse efficiently. So any analytical tool than can rapidly extract valuable insights from existing data has the potential to reduce both operating and capital costs and also understand the resilience to external stressors such as climate change.

Artificial intelligence is now being used in the global water industry to find hidden patterns in large data streams and recommend the best approach to meet specific performance targets. As discussed above, the pattern recognition capabilities of AI offer the rapid identification of adverse conditions in water/wastewater assets, such as pipes and pumps. Additionally, AI has been used to recommend control actions that achieve objectives such as lower energy usage, predicting future events, and the identification of water/wastewater system data patterns. The ability of AI to identify patterns in large data sets with many variables allows additional value to be extracted from generated data. Instead of simply investigating trended data, many dimensions and variables can be investigated to understand any hidden characteristics within the data.

Many global industry challenges are driving the emergence of digital twins in the water/wastewater industry. These challenges relate to: (1) staffing, (2) the need to improve efficiency in water/wastewater related systems, and (3) the need to make more efficient decisions in operations. A digital twin of the system answers most of these drivers. Digital twins force the reconciliation of many data streams into a coherent platform that enables operations and maintenance staff to see the operational conditions more clearly to make more effective and efficient decisions.

Resilience is rapidly becoming one of the most critical water/wastewater asset-related challenges of this century and is aligned with how we '*do more with less*' where assets are commonly retrofitted rather than replaced. The emergence of climate change along with extreme modifications to human behaviour caused by the COVID-19 pandemic, which changed the volume and composition of wastewaters, has further compounded these asset-related challenges. These stressors have reduced the resilience of our water/wastewater assets and infrastructure increasing their vulnerability and the potential for pollution events. To address the the impact of these stressors, it is crucial to understand that resilience is dynamic and may vary diurnally, monthly, seasonally, and even in response to the external events and stressors. The dynamic resilience methodology presented in this chapter uses actual WRRF data, machine learning, data-driven modelling, and heat map visualisations to capture the apparent stresses. These visualisations can be used to evaluate the stress response of a process to an extreme event or investigate its prominence/dominance and the possibility of the process failing (pollution event or reduction in process performance). This could lead to the prediction and anticipation of extreme events and the rapid deployment of event-based mitigations.

The combination of AI, DT, and DR could entirely change how the water/wastewater industry operates, maintains, and understands its facilities. The pattern recognition capabilities of AI can be used to augment DT to free up operations staff for critical time relevant decision-making and not get stuck in the mire of digging through large data sets (micro analysis). Also, the relevance of understanding and evaluating dynamic resilience from actual data can support water/wastewater operational decisions by relating them to the resilience or reserve capacity of a WRRF system. This combination of analytical tools also opens up possibilities for making more complex (and more efficient) systems designs (biological, chemical, and physical) for a broader range of utilities that may not be able to recruit and retain the expert workforce typically required for complex operational systems.