2 | 2021

#HydrolinkMagazine

hydrolink

ARTIFICIAL INTELLIGENCE

IAHR AIRH
International Association for Hydro-Environment Engineering and Research
Hosted by
Spain Water and IWHR, China

IAHR.org

Dr Angelos N. Findikakis
Hydrolink Editor

Prof. Dragan Savić
Guest Editor

EDITORIAL

Last year Hydrolink published four articles on applications of Artificial Intelligence (AI) in water management and hydro-environment problems. The present issue includes four more articles that illustrate the great potential of AI methods and techniques to improve the design, performance, and maintenance of water systems.

During the last thirty years we have seen not only significant developments in the methods and technologies that fall under the general umbrella of AI, but also a change in the attitudes of the hydro-environment community towards the use of these methods and tools. What was initial scepticism about the value of AI in solving water management and other hydraulic problems, evolved in few years into full embracement of these methods and vigorous exploration of their potential by both water managers, academics and others. The article by Babovic *et al.* chronicles the use of different AI methods and tools by the IAHR community over the years and provides some thoughts on the way forward.

Among the tools used in recent years to improve the design, operation and maintenance of large and complex water systems are different nature-inspired optimization techniques combined with other AI methods. These methods have been used to optimize the layout, configuration and pipe size of drinking water distribution networks and wastewater collection systems, as well as to determine the optimal location of water quality and pressure sensors. One category of such techniques is that of the evolutionary algorithms, inspired by the concept of Darwinian evolution. The article of van Thienen *et al.* discusses briefly these methods and presents examples of their application to water systems, including the optimization of the phased expansion of the water distribution system of the Ottawa in Canada and optimizing the rehabilitation plan for a water distribution network in the United Kingdom.

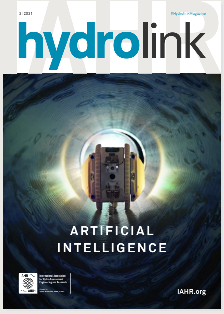
A major problem faced by water utilities is the repair and maintenance of underground pipe systems. The condition of these systems is often unknown until there is a problem that calls for repair action. Manually operated video cameras and acoustic loggers are used in some cases to assess pipe conditions. The data collected from such devices are interpreted by specially trained individuals which can be time consuming. Recent technological developments make it easier to switch from reactive to proactive maintenance of these systems. Autonomous robots can significantly facilitate the inspection of pipe systems and collect data to assess their condition and operational performance and determine maintenance needs. Ongoing research explores the use of swarms of micro-robots designed to work in buried pipe networks autonomously and cooperatively. The article by Mounce *et al.* describes the development of different autonomous inspection platforms and new AI algorithms to extract useful information from the large volume of data collected by such mobile systems.

Some utilities have developed digital twins of their water supply and distribution networks, which can be used to analyze system behavior, detect anomalies, test new ideas and potential changes to improve performance. AI and advanced analytics methods can be used to develop valuable information from the observation data from many real situations recorded by the digital twin of the system. For example, Artificial Neural Networks (ANN) algorithms can be used to characterize and classify demand patterns and detect anomalies in system performance. The article by Alzamora *et al.* discusses the development of digital twins of water systems and describes two examples of such systems, for the water distribution networks of Valencia Metropolitan Area (Spain) and the city of Eindhoven (Netherlands).

AI methods can also be used to develop empirical understanding of the performance of complex pipe networks by learning from the detailed analysis of similar such systems. An article by Telci describes the use of an ANN to provide a preliminary estimate of the pipe characteristics of different parts of loading systems of liquified natural gas facilities, based on learning from past detailed hydraulic transient analysis studies of similar systems. Because of space limitations this article will be included in the print edition of a future issue of Hydrolink, but it is available online.

The articles published in this issue of Hydrolink clearly show the great potential of using AI methods to improve the design, performance, and maintenance of water systems. Although the water sector is already benefiting from advances in AI and other Hydroinformatics tools, the adoption of these techniques is still in an early stage and the case studies in this issue of Hydrolink are examples from only a small part of the areas where AI and Hydroinformatics can add value to water management. Extended reality, serious games, cloud computing, remote sensing, to name but a few of these rapidly evolving digital technologies, provide almost unlimited opportunities for water specialists. However, knowledge and expertise of water systems and understanding of the capabilities of digital technologies will be required for successful applications in the water sector, thus always requiring a human in the loop.

IAHR HYDROLINK 2 | 2021



IAHR
International Association for Hydro-Environment Engineering and Research

IAHR Global Secretariat
Madrid Office
Paseo Bajo Virgen del Puerto, 3
28005 Madrid, SPAIN
T +34 91 335 7908 | F +34 91 335 7935

Beijing Office
A-1 Fuxing Road, Haidian District
100038 Beijing, CHINA
T +86 10 6878 1128 | F +86 10 6878 1890

iahr@iahr.org
www.iahr.org

Cover picture: Sprintbot first prototype in-pipe testing

ISSN: 1388-3445

GUEST EDITOR

Dragan Savić

Dragan Savić is Chief Executive Officer at KWR Water Research Institute based in Netherlands. He is also Professor of Hydroinformatics at the University of Exeter in the UK and Distinguished International Professor at the National University of Malaysia. Professor Savić is an international expert in smart water systems with over 35 years of experience working in engineering, academia and research consultancy. His work has resulted in patentable innovation and spinout companies. In addition to innovation and leadership skills, he is known for believing in bridging science to practice in the wider water sector and utilities in general. He is a long-term member and supporter of IAHR, having served as the Editor-in-Chief of the IAHR/IWA Journal of Hydroinformatics and the Chair of the IAHR Technical Committee on Hydroinformatics.

Artificial Intelligence within IAHR: Past, present and future

By Vladan Babovic, Dawei Han and Yiheng Chen

Artificial Intelligence (AI) is the area of computer science that focuses on creating intelligent machines that can perceive their environment and make decisions to optimise against a goal. Two disciplines frequently mentioned in the context of AI are Machine Learning and Computer Vision. Machine Learning (ML) aims to provide computers with the ability to learn iteratively, improve predictive models and find insights from data without being explicitly programmed. Computer Vision relies on a set of AI methods to train computers to interpret digital images and videos. Examples of computer vision applications include systems for facial recognition and medical diagnostics.

Traditionally fields of hydraulics, hydrology and hydro-environment in general are concerned with natural phenomena described through deterministic equations derived from our best understanding of conservation laws and other underlying physics, chemistry and biology. There was a limited interest in the scientific and communities of practice in early 1990s, even a degree of scepticism about AI. At the time, AI was going through its own formative phase. On the heel of successes of artificial intelligence over last two decades, we are experiencing a growing interest and accelerating adoption by the water community of state-of-the-art AI techniques originated in computer science.

The aim of this article is to review the evolution of AI in IAHR, as well as attempt to outline the present state-of-affairs and future directions. The paper focuses on tangible solutions that were applied to address specific challenges in hydro-environment systems.

Introduction

The debate about the implications of technology on work and jobs is as old as the industrial era. In early nineteenth century, English textile workers called the Luddites protested against the introduction of spinning frames and power looms, fearing that the machines would leave them without jobs. Since then, new technological advancements brought with themselves another wave of concern about a possible displacement of labour. Seen from such a historical perspective it should not be too surprising that the IAHR community was initially quite sceptical about Artificial Intelligence.

Irrespective of the early distrust, the opportunities that AI brings to the hydro-environment community remain very significant. Applications of AI are merely in the first 30 years of a shift in helping humans advance discoveries, leveraging scientific progress made over the past century via improved computing power and enhanced datasets. As a result, we are witnesses of the early successes in practical applications and barely starting to gain a glimpse into the potential of the technology to match human intelligence. This is the time to fully embrace these opportunities. The opportunity is not just about enhancing capabilities but opening a completely new chapters and avenues.

Several white and other papers have been written recently on the potential and use of AI in the water sector[1,2]. The aim of this article is to profile the evolution of AI and ML within IAHR, capture the current state-of-affairs in the hydro-environment community and speculate on future developments.

Earliest AI efforts in hydraulics, hydrology and hydroinformatics

With its origins in computational hydraulics, hydroinformatics[3] was established as a technology born as an integration of numerical modelling with data collection and processing. The field is broadly defined as the application of communication and information technologies to solve water-related problems. Hydroinformatics provided an early platform for (at the time) younger researchers to introduce AI topics in the context of hydraulics and hydrology. The first conference on Hydroinformatics held in Delft in 1994 offered early glimpses of applications of neural networks and evolutionary computation. As a matter of that fact, at that stage the field was not ready to refer to the work as being AI at all. Instead, the presented research was referred to as "adaptive computing methods"[4].

It took few additional years before more substantial work started to appear in peer reviewed literature. A first influential paper on the application of neural networks in hydrology appeared in 1996[5], followed a by two papers on evolutionary computation[6,7] that appeared in the Journal of Hydraulic Research.

There were also some notable applications of AI outside IAHR-related circles. Simpson *et al.*[8], Savić *et al.*[9], Savić and Walters[10] among others reported applications of Genetic Algorithms to water supply systems. Hsu *et al.*[11] and Maier *et al.*[12] described applications of artificial neural networks on rainfall-runoff and water quality problems respectively, whereas Duan and colleagues[13] focused their attention on automatic calibration problems.

Expert systems never really took off in IAHR. Assessment of the rise and fall of expert systems in the water sector[14] lead authors to the conclusion that although some expert systems have been developed for a few specific cases, in general, these expert systems have proved to be too simplistic or unwieldy, and a more open, unstructured access to information and knowledge is preferred.

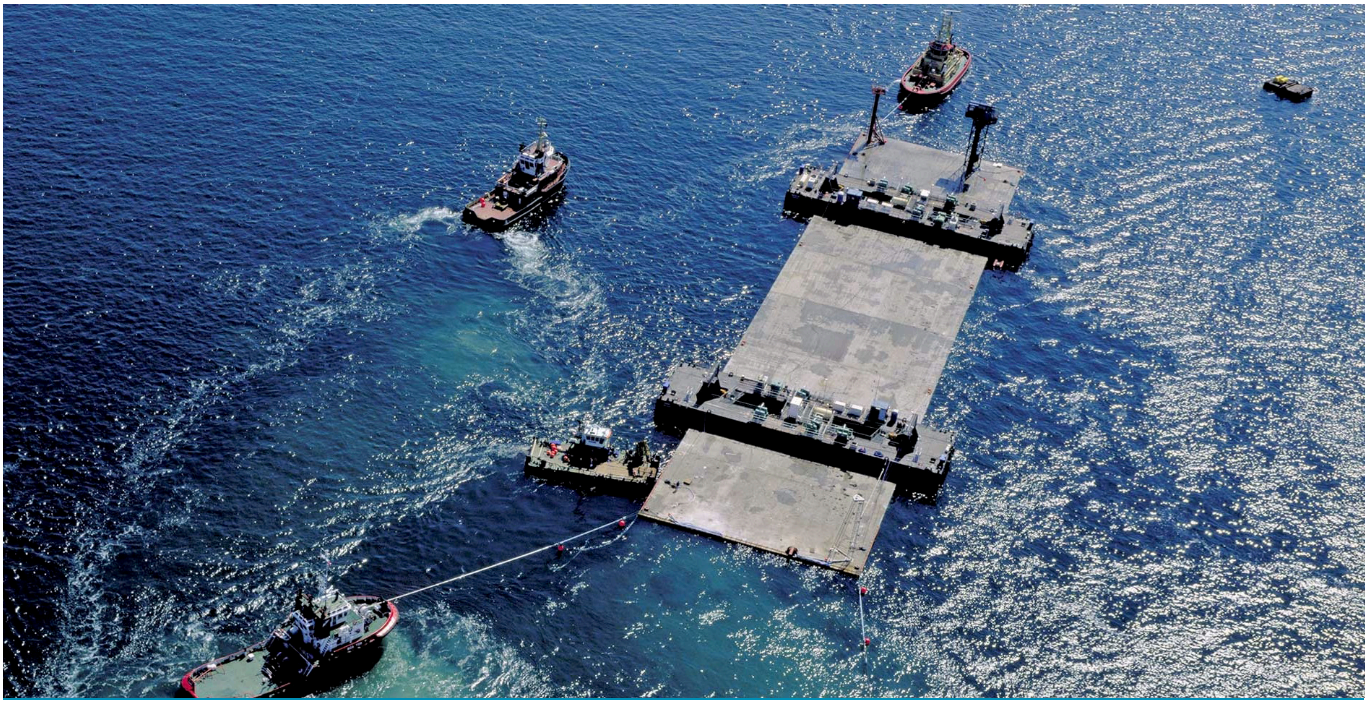


Figure 1 | Barge towing tunnel segments of the Øresund Link under favorable sea conditions based on accurate ANN-based forecasting of surface currents.

Picking-up momentum

The 3rd International Conference on Hydroinformatics held in Copenhagen in 1998 was a super-spreader event for a wider adoption of AI in IAHR circles. This was due to three reasons.

Firstly, one of the keynote speakers at the conference was David Goldberg, a global authority of Genetic Algorithms. Goldberg, himself a hydraulic engineer with a PhD degree on pipeline operations delivered a passionate and engaging address which eloquently articulated a need for use of more intelligent algorithms in hydraulics. His own research was largely shaped under mentorship of his PhD supervisors, hydraulician Benjamin Wylie and computer scientist John Holland. During the conference, Goldberg also offered a short course on evolutionary computing. Secondly, the Journal of Hydroinformatics was launched at the conference hence providing a publication and communication platform for less traditional research and applications, including AI related work. Finally, at the meeting held during the conference, the IAHR Hydroinformatics Section decided to broaden its base and it became a join section together with International Water Association (IWA) and International Association for Hydrological Sciences (IAHS). By doing so, Hydroinformatics and IAHR more closely engaged many more researchers and instigated a momentum and an increase in the critical mass needed to advance AI applications.

As a consequence, the years that followed witnessed a growing interested in AI. Significantly a larger group of authors started to apply data driven techniques to a growing range or problems. Typical examples include runoff prediction and downscaling of climate models. Also, interest in the algorithmic aspects of AI broadened. Authors started exploring and reporting on the performance of a wider range of machine learning algorithms: Support Vector Machines [15], Fuzzy Logic [16], model tree induction [17], Chaos Theory [18] are but few examples.

First real-world applications – real time forecasting of sea currents

While the popularity of AI increased among the research community and a number of academic contributions grew, it was quite significant to demonstrate the value of AI to real world applications. The early years of AI in hydraulics were dominated by Artificial Neural Networks (ANNs) which have evolved to a popular approximation and forecasting tool used in a range of problems and application areas. It should be little surprise, therefore, that the first real-world application was based on a recurrent ANN to create a real-time hybrid data assimilation system resulting in extremely accurate forecasts of sea surface currents.

In the late 1990s, the Danish Hydraulic Institute developed a solution to support the construction of the Øresund Link connecting the Danish capital Copenhagen with the City of Malmø in Sweden [19]. 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 them in a trench dug at the seabed. The elements were prefabricated in a special-purpose build facility North of Copenhagen, sealed shut and using a specially designed barge along with 7 tugboats, were lowered into place at required accuracy of alignment of 2.5 cm. The towing operation (Figure 1) 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 in 17 months. It is alleged that the accurate ANN-based forecasting of sea surface currents was one of the key factors in this achievement [2].

Formative years

Perhaps the key obstacle to earlier and broader acceptance of AI within the larger IAHR community was related to opposition to the use of models perceived to be "black boxes". It was often claimed that such models do not add to scientific knowledge or improved understanding to the field of hydro-environment. While it might be true that early applications of AI were primarily focused on enhancing forecasting abilities and non-linear approximation of input-output relationships, in the years that followed there was an increasing trend towards opening up the black boxes and trying to understand how these models work and, more importantly, how we can relate them to process knowledge emerged.

Dibike and collaborators [20] were among early proponents of the idea, and explored the encapsulation of numerical-hydraulic models in neural networks. Babovic and Keijzer [21] dedicated considerable attention to exploration of methods for incorporation of domain knowledge into machine learning. Among others Giustolisi [22] and Elshorbagy *et al.* [23] reported on a range of machine techniques generating interpretable equations thus having the potential to contribute to knowledge discovery in different hydro-environment disciplines.

In the review of so-called human-competitive results produced by genetic programming Koza [24] highlights the work of Babovic [25] and Baptist *et al.* [26] as examples of machine learning outcomes that are matching or better than the results produced by human experts.

Growing Volumes of Data and Accelerating Computing Power

The capabilities of digital devices continue to increase, and Internet of Things (IoT) sensors provide greater amounts of information than ever, at lower cost and with greater reliability than previously possible. The confluence of these two trends only increases the relevance of AI to IAHR community.

On one side we witness an increase in so-called opportunistic sensing, lately facilitated by crowdsourcing, social media and citizen science that enables the general public to observe local conditions [27]. Couple these with Earth Observation (EO) developments which gather information about planet's physical, chemical and biological systems and you get the perfect ingredients to assess the status of, and changes in, the natural and man-made environment. In recent years, EO has become more and more sophisticated with the advancement of remote-sensing satellites and increasingly high-tech "in-situ" instruments. Today's Earth observation instruments include floating buoys for monitoring ocean currents, temperature and salinity; rainfall trends and similar. Some recent AI-enabled systems based on a very large EO data sets coupled with crowd sensed data are described below.

Water Surface Changes

EO data sets are expected to generate petabytes of data. As such, they developed a system to monitor changes in surface water worldwide in order to understand how our planet is changing as a result of human activities or climate change—and all this relying on millions of megapixels and intelligent algorithms to process the data. The analysis of EO Landsat images made it possible to assess medium (last 15-30 years) and short-term (last 1-5 years) large scale changes, which mainly include erosion and accretion of river banks, large sandbars and islands. Such automated applications allow fluvial geomorphology users, to monitor and detect land and water changes over several decades.

Water Quality Sensing

EO data has been utilised to retrieve water quality information since 1970s. While most of the remote-sensing-based water quality monitoring methods focus on optically active parameters such as turbidity, chlorophylla, Suspended Particulate Matter (SPM) and Coloured Dissolved Organic Matter (CDOM) recent AI-enabled method are able to achieve satisfactory estimates of non-optically active parameters such as Total Phosphorous (TP), Total Nitrogen (TN), and chemical oxygen demand [28].

Researchers from Nankai University, China, created an AI system to predict water quality using the EO data in Shenzhen Bay. The system consists of an AI model and a mobile app that predicts and visualizes the distribution of chlorophyll, turbidity, dissolved oxygen and total dissolved solids in an area of 35 km^2. In this case, AI models learn from the ground measure-ments of water quality and the multispectral remote sensing data to build regression models. The model follows a two-step approach: (i) first it corrects the spectrum of the remote sensing data with the spectrum of the water surface, and (ii) it predicts the water quality parameters from the corrected spectrum. This approach provides a low-cost solution to retrieve the areal distribution of water quality over a large area, which is hardly possible by using the common water sampling methods.

This technology has also been applied to Lake Ontario, Lake Huron and Lake Simcoe in Canada, Lake Bolong, Shenzhen Bay and an anonymous reservoir in China. These cases demonstrate the synergy of AI and remote sensing on managing the water environment for lakes, reservoirs, and coastal areas.

Computer Vision for Opportunistic Rainfall Monitoring

The quantity and quality of precipitation data are crucial in meteorological and water resource management applications. Using rain gauges is the classic approach to measuring rainfall. However, as we enter the age of the Internet of Things in which "anything may become data" so-called opportunistic sensing using unconventional data sources offers the promise to enhance the spatiotemporal representation of existing observation networks [2]. 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 the estimation of the instantaneous rainfall intensity [29]. 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), the Hydroinformatics Institute (Singapore), as well as universities (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 [30].

Where do we go from here?

It is quite obvious from this brief and incomplete review of AI-related developments in IAHR that significant progress has been made over the past 25-30 years, and that the IAHR community has embraced and significantly benefitted from data science.

What are the boundaries and obstacles that need to be addressed to enable even more significant progress?

Freedom to the data!

An area that needs significant attention is the democratisation of observations and data. Issues associated with data privacy on one hand, and the need for openness and data exchange on the other are essential–particularly within the hydro-environment community. As highlighted in a recent White Paper [31] IAHR was founded by 66 hydraulic laboratories, making the science of measuring and data acquisition entrenched in the origins of the association. Observations and measurements are fundamental for scientific progress in general and remain essential for advancing our insights and knowledge. Induction of relationships and conversion of data to a better understanding of the processes that generated or produced those data has always been at the very heart of hydraulics.

These developments open avenues for taking advantage of big data, an area that is gaining attention in the hydroinformatics community [32]. Big data is enabled by the extremely large datasets that cannot be processed within a tolerable time using traditional data processing methods. IAHR can benefit from the big data technology and analytical tools to handle large datasets, from which creative ideas and new insights could be mined.

Sharing the data as a community would result in not only more affordable access to quality-assured data but would also accelerate scientific and technological advances within our community at large. Perhaps this is a role in which IAHR as an association should take a pro-active and strong leadership role. With increasing data availability deep learning is gaining popularity. Deep learning is a class of machine learning algorithms that uses multiple layers to progressively extract higher-level features from raw inputs. For example, in image processing, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits or letters or faces. It has been used to replace the conventional hydrodynamic models (very slow to run due to complex numerical computations) in simulating flood inundations and vulnerabilities [33].

Hydro-environment-informed Machine Learning?

As we enter the true digital information era, one of the greatest challenges facing organisations and individuals is related to turning the rapidly expanding data stores into accessible, and actionable knowledge. Without such developments, we risk missing most of what the data have to offer. The traditional approach of a human analyst, intimately familiar with a data set, serving as a conduit between raw data and synthesised knowledge by producing useful analyses and reports, is breaking down.

What is to be done with all these data? Ignoring whatever we cannot analyse would be wasteful and unwise. This is partcularly pronounced in scientific endeavours, where data represent carefully collected observations about particular phenomena that are under study. Data science models, although successful in a number of commercial domains, have had limited applicability in scientific problems involving complex physical phenomena. Theory-guided data science is an emerging paradigm that aims to leverage the wealth of scientific knowledge for improving the effectiveness of data science models in enabling scientific discovery.

Theory Guided Machine Learning (TGML) recognises that applying the AI alone is not the entire story. At least not in scientific domains, such as hydro-environment! Scientific theories encourage the acquisition of new data and this data in turn leads to the generation of new theories. Traditionally, the emphasis is on a theory, which demands that appropriate data be obtained through observation or experiment. In such an approach, the process is what we may refer to as theory-driven. Especially when a theory is expressed in mathematical form, theory-driven discovery may make extensive use of strong methods associated with mathematics or with the subject matter of the theory itself. The converse view takes a body of data as its starting point and searches for a set of generalisations, or a theory, to describe the data parsimoniously or even to explain it. Usually such a theory takes the form of a precise mathematical statement of the relations existing among the data. This would be the AI (and ML in particular) driven discovery process. The new data driven, ML-discovered models, combined with the understanding of the physical processes – the theory – can result in an improved understanding and novel formulations of physical laws and an improved predictive capability. Tech giant Google had launched a flood prediction service using machine learning to identify areas of land prone to flooding and alert users before the waters arrive in 2018 for India's Patna region with some success https://blog.google/technology/ai/ flood-forecasts-india-bangladesh/. Realising the limits of machine learning tools, Google has built a forum to connect computer data scientists with hydrologists/ hydraulicians to merge ML and process knowledge to achieve the best results (see Google Flood Forecasting Meets Machine Learning Workshop, 2019, Tel Aviv https://ai.googleblog.com/2019/03/a-summary-of-google-flood-forecasting.html.

Some early activities along those lines are starting to take shape [34, 35], but there is a still a long road ahead to realise the full potential of combining the two approaches: theory-driven, understanding-rich with state-of-the-art AI- algorithms to accelerate knowledge discovery in hydro-environment.

Nevertheless, the future is very bright!

Vladan Babovic
Vladan Babovic is professor at the National University of Singapore (NUS). He is a leading scientist in the field of hydroinformatics where he has been spearheading research in artificial intelligence, machine learning and computer modelling of hydraulics and hydrological phenomena from 1990s. In more recent years, his work on real options pertaining to decision-making under deep uncertainties in water- and climate-related domains is gaining wider recognition. In addition to being a leading researcher and educator, Babovic is a scientist entrepreneur who secured research and venture capital-funding for several applied and fundamental research organisations.

Dawei Han
Dawei Han is professor of hydroinformatics from the Department of Civil Engineering, University of Bristol, UK. His research interests include hydrological modelling, Real-time Flood Forecasting, Flood Risk Assessment and Management, Climate Change impact, Remote Sensing and Geographic Information System and natural hazards.
Further information can be found here: https://research-information.bris.ac.uk/en/persons/dawei-han.

Yiheng Chen
Dr Yiheng Chen is a postdoctoral research associate at Nankai University, China, a staff of Sino-Canada Joint R&D Center on Water and Environmental Safety (SCWE), and secretary general of UK China Association of Resources and Environment (UK-CARE). He received his master's degree and PhD degree in Civil Engineering at University of Bristol, UK. Dr Chen's research interests focus on hydroinformatics, particularly on big data and water quality. He worked on a number of research projects on developing automatic real-time water quality monitoring devices. He is also experienced in flood forecasting and precipitation nowcasting.

References

1 | D. Savić, 2019, Artificial Intelligence-How can water planning and management benefit from?, *IAHR White Paper.*

2 | Z. Kapelan, E. Weisbord, and V. Babovic, *Digital Water.* Artificial Intelligence Solutions for the Water Sector, 2020.

3 | M. B. Abbott, 1999, Introducing Hydroinformatics, *Journal of Hydroinformatics,* vol. 1, no. 1, pp. 3-19.

4 | V. Babovic and A. W. Minns, 1994, Use of computational adaptive methodologies in hydroinformatics, in P*roceedings of the First International Conference on Hydroinformatics,* pp. 201-210.

5 | A. W. Minns and M. J. Hall, 1996, Artificial neural networks as rainfall-runoff models, *Hydrological Sciences Journal,* vol. 41, no. 3, pp. 399-417.

6 | V. Babovic and M. B. Abbott, 1997, The evolution of equations from hydraulic data Part I: Theory, *Journal of Hydraulic Research,* vol. 35, no. 3, pp. 397-410.

7 | V. Babovic and M. B. Abbott, 1997, The evolution of equations from hydraulic data Part II: Applications, *Journal of Hydraulic Research,* vol. 35, no. 3, pp. 411-430.

8 | A. R. Simpson, G. C. Dandy, and L. J. Murphy, 1994, Genetic Algorithms Compared to Other Techniques for Pipe Optimization, *Journal of Water Resources Planning and Management,* vol. 120, no. 4, pp. 423-443.

9 | D. A. Savić, G. A. Walters, and J. Knezevic, 1995, Optimal opportunistic maintenance policy using genetic algorithms, 1: formulation, *Journal of Quality in Maintenance Engineering,* vol. 1, no. 2, pp. 34-49.

10 | D. A. Savić and G. A. Walters, 1995, An evolution program for optimal pressure regulation in water distribution networks, *Engineering Optimization,* vol. 24, no. 3, pp. 197-219.

11 | K. Hsu, H. V. Gupta, and S. Sorooshian, 1995, Artificial Neural Network Modeling of the Rainfall-Runoff Process, *Water Resources Research,* vol. 31, no. 10, p. 2517.

12 | H. R. Maier and G. C. Dandy, 1996, The Use of Artificial Neural Networks for the Prediction of Water Quality Parameters, *Water Resources Research,* vol. 32, no. 4, pp. 1013-1022.

13 | Q. Y. Duan, V. K. Gupta, and S. Sorooshian, 1993, Shuffled complex evolution approach for effective and efficient global minimization, *Journal of Optimization Theory and Applications,* vol. 76, no. 3, pp. 501-521.

14 | R. K. Price and D. P. Solomatine, A brief guide to hydroinformatics, UNESCO-IHE Institute for water education, Delft, Delft, The Netherlands, 2009.

15 | Y. B. Dibike, S. Velickov, D. Solomatine, and M. B. Abbott, 2001, Model Induction with Support Vector Machines: Introduction and Applications, *Journal of Computing in Civil Engineering,* vol. 15, no. 3, pp. 208-216.

16 | A. J. Abebe, D. P. Solomatine, and R. G. W. Venneker, 2000, Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events, Hydrological Sciences Journal, vol. 45, no. 3, pp. 425-436.

17 | D. P. Solomatine and Y. Xue, 2004, M5 Model Trees and Neural Networks: Application to Flood Forecasting in the Upper Reach of the Huai River in China, *Journal of Hydrologic Engineering,* vol. 9, no. 6, pp. 491-501.

18 | A. Elshorbagy, S. P. Simonovic, and U. S. Panu, 2002, Estimation of missing streamflow data using principles of chaos theory, *Journal of Hydrology,* vol. 255, no. 14, pp. 123-133.

19 | V. Babovic, R. Cañizares, H. R. Jensen, and A. Klinting, 2001, Neural Networks as Routine for Error Updating of Numerical Models, *Journal of Hydraulic Engineering,* vol. 127, no. 3, pp. 181-193.

20 | Y. B. Dibike, A. W. Minns, and M. B. Abbott, 1999, Applications of artificial neural networks to the generation of wave equations from hydraulic data, *Journal of Hydraulic Research,* vol. 37, no. 1, pp. 81-97.

21 | V. Babovic and M. Keijzer, 2002, Rainfall runoff modelling based on genetic programming, *Hydrology Research,* vol. 33, no. 5, pp. 331-346.

22 | O. Giustolisi, 2004, Using genetic programming to determine Chèzy resistance coefficient in corrugated channels, *Journal of Hydroinformatics,* vol. 6, no. 3, pp. 157-173.

23 | A. Elshorbagy and S. L. Barbour, 2007, Probabilistic Approach for Design and Hydrologic Performance Assessment of Reconstructed Watersheds, Journal of Geotechnical and Geoenvironmental Engineering, vol. 133, no. 9, pp. 1110-1118.

24 | J. R. Koza, 2010, Human-competitive results produced by genetic programming, *Genetic Programming and Evolvable Machines* 2010 11:3, vol. 11, no. 3, pp. 251-284.

25 | V. Babovic, 2000, Data Mining and Knowledge Discovery in Sediment Transport, *Computer-Aided Civil and Infrastructure Engineering,* vol. 15, no. 5, pp. 383-389.

26 | M. J. Baptist *et al.,* 2007, On inducing equations for vegetation resistance, *Journal of Hydraulic Research,* vol. 45, no. 4, pp. 435-450.

27 | J. Seibert, B. Strobl, S. Etter, P. Hummer, and H. J. (Ilja) van Meerveld, 2019, Virtual Staff Gauges for Crowd-Based Stream Level Observations, *Frontiers in Earth Science,* vol. 7, p. 70.

28 | H. Guo, J. J. Huang, B. Chen, X. Guo, and V. P. Singh, 2021, A machine learning-based strategy for estimating non-optically active water quality parameters using Sentinel-2 imagery, *International Journal of Remote Sensing,* vol. 42, no. 5, pp. 1841-1866.

29 | P. Allamano, A. Croci, and F. Laio, 2015, Toward the camera rain gauge, *Water Resources Research,* vol. 51, no. 3, pp. 1744-1757.

30 | S. Jiang, V. Babovic, Y. Zheng, and J. Xiong, 2019, Advancing Opportunistic Sensing in Hydrology: A Novel Approach to Measuring Rainfall With Ordinary Surveillance Cameras, *Water Resources Research,* vol. 55, no. 4, pp. 3004-3027.

31 | D. Valero *et al.,* 2021, Pathways towards democratization of hydro-environment observations and data, IAHR White Paper, no. 1, pp. 1-8.

32 | Y. Chen and D. Han, 2016, Big data and hydroinformatics, *Journal of Hydroinformatics,* vol. 18, no. 4, pp. 599-614.

33 | D. Tien Bui *et al.,* 2020, A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area, *Science of The Total Environment,* vol. 701, p. 134-413.

34 | H. M. V. V. Herath, J. Chadalawada, and V. Babovic, 2020, Hydrologically Informed Machine Learning for Rainfall-Runoff Modelling: Towards Distributed Modelling, *Hydrology and Earth System Sciences Discussions,* pp. 1-42.

35 | J. Chadalawada, H. M. V. V. Herath, and V. Babovic, 2020, Hydrologically Informed Machine Learning for Rainfall-Runoff Modeling: A Genetic Programming-Based Toolkit for Automatic Model Induction, *Water Resources Research,* vol. 56, no. 4, p. e2019WR026933.

Nature-based optimization techniques and their application in the water industry

By Peter van Thienen, Edward Keedwell, Raziyeh Farmani and Ina Vertommen

Even though formal optimization techniques have been applied to several types of problems in the water industry, the vast majority of examples from the literature is based on (meta)heuristics, and in particular those inspired on nature. Their application has been shown quite successful and has allowed significant boosts in performance while reducing costs. This article provides an overview of nature-inspired optimization techniques, and briefly discusses a number of case studies in which they have been applied.

There is already a long history of numerical optimization in the water industry[1]. Many advances have been made regarding the solution methods from linear and non-linear programming[2] to nature-based optimization techniques, like genetic algorithms, simulated annealing, particle swarm[3], ant colony and many others. Besides the optimization methods, the design problems have evolved from single to multiobjective, and from deterministic to stochastic[2] and robust approaches. For a thorough overview of the optimal design of water distribution networks and the applied evolutionary algorithms and metaheuristics, the reader is referred to[4,5]. The application of the methods has been shown quite successful and has allowed significant boosts in performance of many aspects of systems while reducing costs. In this article, we present an overview of nature-inspired optimization techniques, and briefly discuss a number of case studies where they have been applied.

Nature-inspired optimization techniques

Nature-based or nature-inspired optimization techniques have been in existence since the work of Rechenberg and Schwefel in the early 60s on evolutionary strategies[6]. Since then, thousands of approaches have been developed based on a wide variety of natural systems. A complete taxonomy of these approaches would be much too large for this article, but generally speaking these methods fall into the following categories:

- Evolutionary based approaches:[4] where the underlying iterative process includes perturbation by crossover and/or random mutation and variously selection and replacement of individuals. Methods here include evolutionary strategies, genetic algorithms, multi and many objective evolutionary algorithms, genetic programming and differential evolution.
- Swarm intelligence based approaches:[3] where the underlying iterative process is based on the movement and interaction of individual agents working as part of a collective (e.g. a herd, a flock or a swarm). Methods here include particle swarm optimization, ant colony optimization, artificial bee colony and many other approaches based on a diverse array of organisms including fireflies, wolves, herons, fish schools and whales.

- Other approaches:[7] other sets of metaheuristic optimization methods exist that are perhaps more distantly related to their natural inspiration but are nonetheless nature-inspired. Methods such as simulated annealing (inspired by the cooling of metal), chemical reaction optimization and gravity search are examples of these. Other methods that have a related mode of operation but are not naturally inspired in the conventional sense include methods such as tabu search and many variants of local search.

Although each of the above approaches is inspired by a different natural system, it is an open question as to whether they all represent distinctly separate methods in the exploration of the search space. However, each method has to balance the trade-off between exploration and exploitation of the space of possible solutions, which for most problems in the water industry is very large indeed. Methods that solely exploit (e.g. hillclimbing) will find mediocre solutions quickly, whereas those that solely explore (e.g. random search) will eventually find good solutions, but over huge timescales. Nature-inspired algorithms typically embed exploration through the use of a random or probabilistic operation (e.g. mutation in evolutionary algorithms, probabilistic path selection in ant colony optimization) and exploitation by conferring some preference on solutions that perform well at the optimization task. A further feature that characterizes these nature-inspired methods is that they are general-purpose optimization algorithms that can be applied to many different problems. For most algorithms application to new problems requires the specification of three elements:

- Representation/encoding: this is the mapping between the features of the problem being solved and the numeric decision variables that will be optimized by the algorithm. Although some representations will be straightforward, there are often choices to be made that will determine the efficacy of the search.
- Fitness or objective function(s): provides an assessment of solution quality in terms of one or more objectives of the problem being solved. Objectives in water industry network problems usually include calculations of CAPEX, OPEX, hydraulic constraints such as pressures, velocities and tank penalties, the coverage or detection likelihood for sensor networks, etc.
- Parameter settings: most algorithms have parameter settings that can affect the efficacy of the algorithm and can vary on different problems. Common parameter settings include population sizes, number of iterations, perturbation operator selections and application rates (evolutionary approaches) and various momentum, velocity and pheromone evaporation terms (swarm intelligence). These are usually set by rule of thumb or through prior experimentation, although adaptive methods that set these parameters automatically through the search are becoming increasingly popular.

The choice of algorithm is often dictated by the number and complexity of the decision variables and constraints, the computational complexity of the objective function and the number of objectives to be optimized. Often the computational complexity can be the largest factor and if the problem is particularly time consuming to solve, on the optimization run may require access to high performance or cloud computing resources, although many problems can be optimized with modern desktop equipment.

Nature-inspired optimization algorithms are among the best-known approaches for discovering good solutions to highly complex large-scale problems in reasonable time and have the potential to transform the design and operation of the complex assets and systems that characterize the water industry. This is particularly the case when these methods are combined with other popular AI methods such as machine learning, where the predictive power of these methods can be coupled with optimization to yield asset and operations upgrade programmes designed for future system demands. AI is revolutionizing many sectors and as such, it offers great potential for the water industry as well.

Applications in the drinking water industry

A wide range of problems exists in the water industry for which nature-based optimization algorithms can provide valuable solutions. We give a generic overview here and discuss a number of case studies for different fields of application and/or geographies in the following paragraphs. For water resource management, areas of application include model calibration, choosing sampling locations for monitoring, and risk-based water supply portfolio planning, the optimization of reservoir operation, and regulation of the abstraction from different sources for scarcity management. Water pipe networks, both drinking water distribution and wastewater collection networks have also been subjected to numerical optimization in numerous cases. Not only their layout and sizing, but also their subdivision into functional sections and the optimal placement of different types of sensors, including water quality sensors for contami-nation detection and pressure sensors for leak detection have been explored.

Network design with multiple planning horizons

Climate change, population mobility and urban development in cities necessitates the planning of major distribution network upgrades and requires a phased approach where changes to population and demand increase over time. Work in 2007[8] brought together researchers, consultants and city planners to develop the water distribution system master plan for the City of Ottawa. This phased expansion reflected expected population and demand increases over a 25-year planning horizon, optimized by using an evolutionary algorithm. When the work was carried out, the population of Ottawa was about 800,000 which has risen to almost one million at the time of writing, highlighting the extent of the planned urban growth and underlining the accuracy of the projections which anticipated 1.07 million by 2021. The city's demand is fed by two water

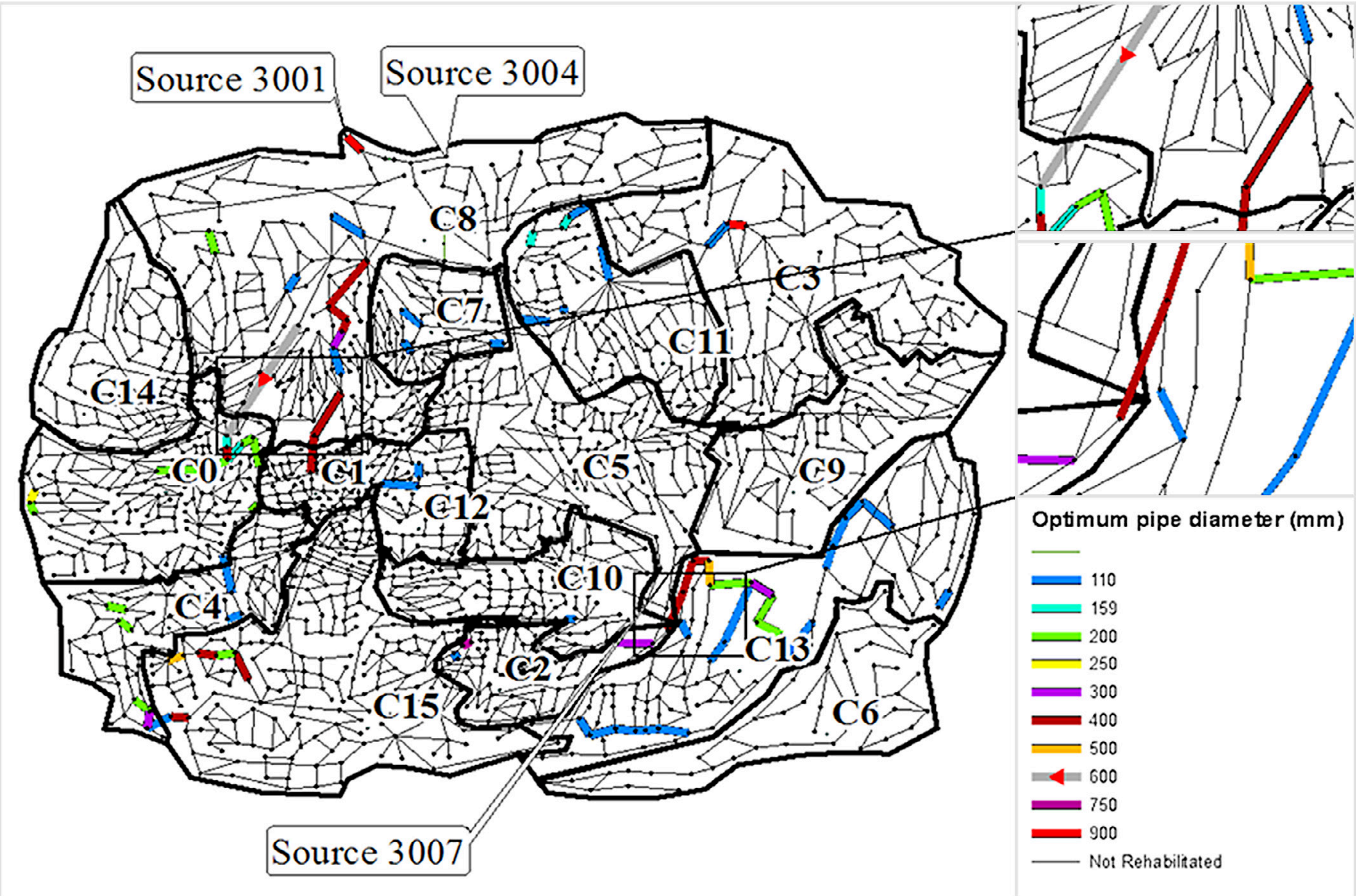


Figure 1 | Layout of proposed pipes for rehabilitation in different clusters for solution with no deficiency.

purification plants, Britannia and Lemieux, located along the Ottawa River, with a combined capacity of 640 Millions of Liters per Day (MLD) and the East and West Urban Communities are fed by a 1,200 mm transmission feed. A 1,220 mm main also feeds the west portion of the South Urban Community and a 762 mm main feeds a small area on the east side of the river. There are two major storage reservoirs, one located in the center of the city and the other in the East Urban Community with storage capacity of 108 ML and 82 ML respectively. There are also two smaller reservoirs in the West and South with storage capacity of 34 ML and 18 ML respectively, and additional elevated storage in the communities.

As with many evolutionary optimization applications, the majority of the work was involved in the development of representation and objective (fitness) function formulations to enable the evolutionary algorithm to effectively solve the problem. The representation establishing the link between the algorithm and problem, provided the options to introduce new infrastructure and upgrade existing assets. A key element here was the introduction of single variables that combined logical sets of infrastructure upgrades. An example of this is in the introduction of a new tank; this must be accompanied by the pipework necessary to connect the tank to the network and so was established as a single 'decision' for the algorithm to take. The introduction of these variables had the dual effect of increasing the engineering feasibility of the developed solutions and reducing the search space for the evolutionary algorithm. A single objective function minimized costs (CAPEX and OPEX) and hydraulic penalties under demand scenarios in 2011, 2021 and the final planning horizon in 2031. The single objective function required a coefficient to balance the cost and individual hydraulic components, which allowed the optimization to be tailored towards end-user requirements, although it would also suit a multi-objective approach. Extensive optimization runs were conducted, and a final 2031 solution was developed at an estimated CAD 402M, including CAD 205M for plant expansions, CAD 110M for new water mains, CAD 45M for pumping stations and CAD 24M and CAD 17M for reservoir expansions and elevated tanks respectively. However, the optimization was able to show only CAD 79M was required to satisfy 2011 demands and a further CAD 152M was required between 2011 and 2021, demonstrating the benefit of using multiple planning horizons within the project.

Optimum rehabilitation schemes

The water sector is under growing pressure to deliver service that satisfies customer expectations and regulatory requirements. Urbanization and growing water demand are putting great stress on ageing or inadequate infrastructure in many countries. This example[9] demonstrates how existing scientific and engineering knowledge benefited from advances in soft computing analytics to address deficiencies in a water distribution networks. This network is part of a water distribution network of a city in the UK. The network has grown over years from a small network to a system that serves 400,000 customers.

The network has two reservoirs, 5 connections that import water from adjacent systems, 1,891 nodes and 2,462 pipes. The pipes have small to moderate diameter sizes with no major transmission mains due to the way the system evolved in the past. The existing network is unable to satisfy the recent growth and projected future demands with adequate pressure. The problem was set as a multi-objective optimization problem in order to generate a set of optimum rehabilitation schemes that trade-off between capital investment and system performance. A two-stage methodology was proposed. In the first stage, using network connectivity and topology, the system was divided into a number of clusters with stronger internal than external connectivity. In the second stage, three different problem setting strategies, for optimal rehabilitation, were considered including: 1 | rehabilitation of pipes within clusters 2 | rehabilitation of feed pipelines, pipes that connect the clusters with deficiency to other clusters or to the sources, and 3 | rehabilitation of pipes within clusters and feed pipelines.

Using an undirected graph algorithm of the Gephi tool[10], 16 clusters with different degrees of pressure deficiency were identified for this network. The pipes in the clusters that have no performance issues and do not participate in water transmission to other areas of the network will have no contribution towards reducing deficiency in the system. Therefore, they were not considered as candidate pipes for rehabilitation of the system. A total of 248, 149 and 349 pipes were considered for rehabilitation (as decision variables) for strategies 1, 2, and 3 respectively. The non-dominated sorting genetic algorithm II (NSGA-II) was used to generate optimum Pareto-front between total cost and number of nodes with pressure deficiency for different strategies.

The generated results (**Figure 1**) based on strategy 3 dominated the results generated by both strategies 1 and 2. The results of strategy 2 were also compared with those generated based on considering i) all the pipes as design variables and ii) a subset of pipes (567 pipes) based on the engineering judgment (water company). The optimum Pareto front generated by strategy 3, again dominated the results generated based on these two problem settings. The optimum solution with no pressure deficiency generated by strategy 3, has a total cost of GBP 3.05 million. A solution, with total cost of GBP 4.15 million with 195 nodes with pressure deficiency, was generated independently by the water company by trial and error. A solution, with a similar number of deficient nodes, on Pareto-front of strategy 3 has a cost of GBP 1.5 million which is 65% cheaper than the solution generated manually.

Network design and transition optimization

Different challenges arise when applying optimization techniques to larger real-world networks: the computational effort involved in applying numerical optimization techniques to such large networks and being able to translate practical challenges and constraints to formal problem formulations with clear objectives, constraints and decision variables. To tackle the first challenge, one might think of high-performance computing and problem-specific variators. Regarding the second one, we have learned that these types of problems are best solved in an iterative process between researchers and practitioners, wherein each result is assessed, and the optimization problem is adjusted accordingly to the gained insights. This approach leads to results that are a perfect fit for what water utilities are looking for and has the added bonus of providing them with new insights into their own water supply systems.

An optimization tool has been applied to the rehabilitation of real-life networks in the Netherlands. It uses (modified) genetic algorithms and NSGAII as optimization methods. Network rehabilitation is approached as a two-phased problem: (1) the optimal design of the network (so called blueprint or master plan) and (2) the optimal transition between the currently existing network and the blueprint, i.e., the rehabilitation timeline. The design of the network blueprint considers the minimization of costs (a function of the diameter and length of the new pipes), constrained by minimum pressure requirements and commercially available pipe diameters and materials. For the rehabilitation timeline both hydraulic (improvement of current pressure deficiencies) and risk based (reduction of pipe failures, which are a function of pipe diameter, material and age) objectives have been considered, in combination with a practical aspect regarding the number of construction sites in each rehabilitation step. A construction site is a cluster of valve sections where old pipes are replaced by new ones. Water utilities prefer to concentrate rehabilitation works in a few sites, instead of working in a very disperse manner.

This approach was applied to the water distribution network serving the area of Helmond-Mierlo, with 105,000 inhabitants in the Netherlands[11]. The network model has about 12,000 pipes. Adding to that 32 commercially available pipe diameters, it is clear how large the solution space for this problem is. By starting the optimization problem from the current pipe diameter values and using problem specific variators in the GA, it was possible to effectively explore the solution space.

Figure 2 illustrates the obtained results. The costs for rebuilding entirely the network currently in the ground would be EUR 41.1M. At the peak demand conditions (maximum demand in the past 10 years) the 30 m pressure requirement is not met at several nodes of the network. The costs of the optimized blueprint are significantly lower, at EUR 26.4M. At the same time, the hydraulic performance is significantly improved: the total pressure deficiency in the network (sum of all pressures below the required 30 m) is reduced by 97%. Regarding the rehabilitation timeline two Pareto fronts were obtained: (1) trade-off between maximization of hydraulic performance and number of construction sites, and (2) trade-off between the minimization of pipe failures and the number of construction sites. **Figure 2** (b) illustrates one of the identified solutions.

Network rehabilitation is an opportunity for re-designing, an often organically grown network. The achieved results prove that numerical optimization techniques can be used in this context. Moreover, the achieved amount of savings allows the

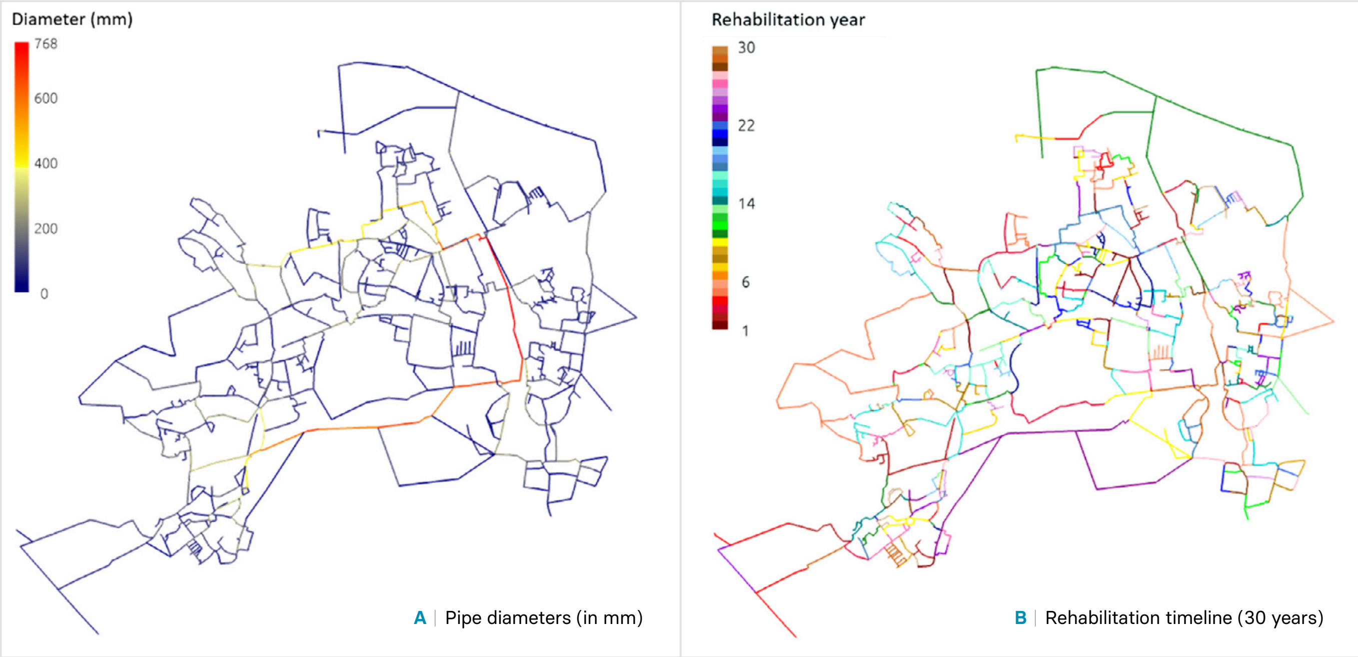


Figure 2 | Optimized solution for (A) pipe diameters to minimize costs while guaranteeing adequate network performance given by different colors, and (B) rehabilitation timeline that maximizes the reduction of pipe failures with a maximum of 10 rehabilitation sites per year (the colors and numbers indicate the year in which the pipes should be rehabilitated, pipes with the same color are rehabilitated in the same year).

water utility to rehabilitate their networks at a higher rate. Proactively replacing old and fragile pipes with new ones reduces the risk of pipe failure and thus, water losses due to leakage.

Moreover, having the initial optimization problem defined and all relevant data organized, it makes it easy to accommodate different objectives, constraints, and scenarios. In this way, the optimization problem can be re-run when new information, such as changes in urban development or water demand, becomes available, making it a very flexible approach.

Optimal water quality sensor placement

Both societal events and technological advances have pushed the development of techniques for online water quality monitoring in drinking water distribution systems since the beginning of this century. Their purpose is generally to protect customers from incidental and/or intentional drinking water contamination. The number of online monitoring sensors that can be placed in any system is always constrained by budgetary limitations. Therefore, methods have been developed to determine optimal sensor placement [12,13] within a drinking water distribution network. Optimality is, however, a matter of definitions and requirements. The objectives that have been presented in the literature can be classified roughly into three categories[14], aimed at obtaining information, facilitating utility response, and mitigating the effects of contamination (Table 1).

Of the three classes of sensor placement optimization objectives, those that are information-oriented are the simplest to compute[12], requiring only a network model (hydraulics and material transport). The more complex effect-oriented approach has been implemented in the Threat Ensemble Vulnerability Assessment and Sensor Placement Optimization Tool (TEVA-SPOT)[13]; here we present some results of its application to a network model of part of the network of Vitens, the largest water utility in the Netherlands. Many simulations were performed

Objective class	Orientation of optimization strategy towards...		
	Information	Utility response	Effect mitigation
Examples	Detection likelihood, time to first detection, network/customer coverage	Redundant detection, identifiability of contamination source	Population affected, ingested volume, numbers of people above does threshold
(Dis) advantages	Simple, but several steps from information to actual customer protection	Close to operational practice	Objective matches final objective of utility, but the latter is complex to compute and the results show a strong dependence of utility response (time)

Table 1 | Rate of increase of potential energy as a function of the lake trophic state with Po = 395,343 W

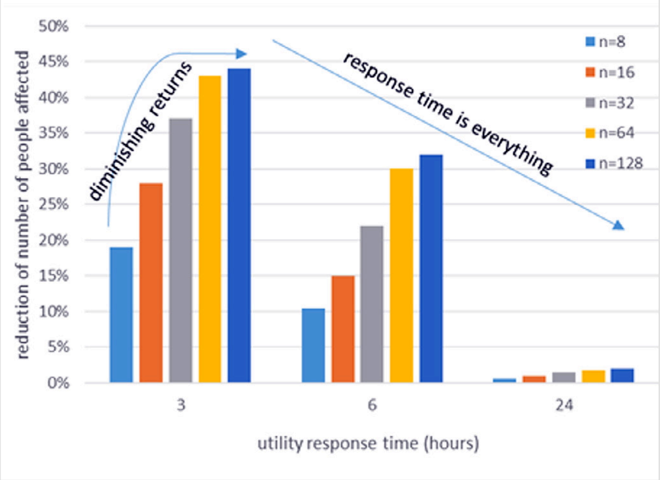


Figure 3 | Performance of optimized water quality sensor networks as a function of number of sensors (n) and utility response times. Results from[14].

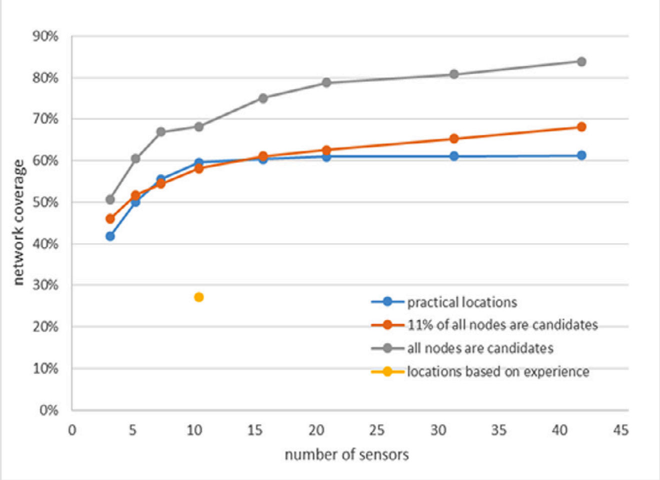


Figure 4 | Water quality sensor network performance for ideal and practically feasible locations. Results from[14].

to study the relationship between the assumed and actual response times of the utility (after which water consumption is assumed to cease), on the one hand, and the performance of the sensor network (reduction in the number of people affected) on the other. In all cases, the sensor locations were optimized using a genetic algorithm. Some results are shown in Figure 3. Two important observations can be made: Sensors are useless for event detection if the utility's response time is too long and Every additional sensor contributes less to the objective than its predecessors (the law of diminishing returns).

When the water utility elects to consider only practically suitable locations for sensor placement, this may have a significant effect on the network's performance as shown in a different study presented in Figure 4 which presents a comparison between the performance of optimizations (again using a genetic algorithm) for different sets of uniformly distributed nodes used as potential locations (300, and all 2,700, respectively). In some cases, the network based on the practical set of potential locations performs better than one based on a larger number of uniformly distributed nodes. The practical set may include, by chance, suitable locations that are absent from the uniformly distributed sets. The best performance is seen when all network nodes are considered as candidate locations (grey curve in Figure 4). But the main conclusion must be that even though optimizing a sensor network configuration based on practically available locations results in some performance loss compared to networks in which sensor placement is not restricted, performing an optimization is still worthwhile.

Conclusions

Several decades of development of ideas, methods and applications have resulted in a myriad of cases which demonstrate the added value of applying numerical optimization techniques to water industry problems. Nature-based metaheuristic methods have been and continue to be particularly popular and successful because of their ability to deal with the scale and complexity that are typical in this field of application. Nevertheless, the vast majority of system design projects in practice continues to rely on human designs and expert judgment. This is not to say that the human factor should be taken out of the equation, rather the opposite: the application of numerical optimization techniques taking into account the deep domain knowledge of the water industry's experts holds the potential for performance increase and cost reduction (both monetary and in terms of environmental impact) in all these projects. The primary gains for the industry from numerical optimization will come from taking the step to actually more or less universally applying these methods.

Ongoing development in this area is focused on the development of new and faster formulations of algorithms, often through the combination of one or more techniques and in the development of methods that can take advantage of modern CPU and GPU (graphical processing unit architectures). Other areas of development are learning optimization (hyperheuristic) methods[15] and multi-method[16] search which combine machine learning and optimization components to create methods that can adapt to new search space domains on the fly. Real-world applications are being addressed through the development of many-objective[17] and human-in-the-loop[18] algorithms that aim to consider the large number of objectives that characterize real world problems and leverage the domain expertise of experienced staff. In this way, the people that have always been responsible for the design of systems and their operation continue to be so, but with a new and very powerful tool in their toolbox.

We are becoming more aware of the uncertainties that exist in the models and data that we are applying optimization to. In addition to this we observe that the world is changing at an ever quicker pace, and progressing climate change can be expected to have more changes and con-sequences in store for us for the rest of the century, both in terms of water availability and demand. Considering these uncertainties in the present state and future conditions and requirements (discussed in[19]), it becomes urgent to start taking these uncertainties into account in the formulation of our optimization problems. Academic work on robust and resilient optimization has been presented in the past decades[2,20] this should become the standard in real world applications as well.

Peter van Thienen

Peter van Thienen is a senior researcher and chief information officer of KWR. He holds a PhD in geophysics and has over 10 years of experience in the water industry. At KWR, Peter works with a number of colleagues on research and development questions with respect to the drinking water distribution network, from a quantitative and modelling point of view. Examples include the analysis of flow data for the understanding of the network and the detection of leaks (including the development of the CFPD method); the numerical optimization of the design of sensor and pipe networks (development of the Contamination Source Toolkit and the Gondwana platform); and the development of an inspection robot for drinking water pipes.

Ed Keedwell

Ed Keedwell is a Professor of Artificial Intelligence at the University of Exeter and has 20 years of experience in the research and development of novel optimization techniques for application in the water industry. He is currently Director of Research for Computer Science and has research interests in optimization (e.g. genetic algorithms, swarm intelligence, hyperheuristics) machine learning and AI-based simulation and their application to a variety of difficult problems in engineering and bioinformatics that has led to over 150 journal and conference publications. Particular areas of current interest are the optimization of transportation systems, the development of sequence-based hyperheuristics and human-in-the-loop optimization methods for applications in engineering.

Raziyeh Farmani

Raziyeh Farmani is an associate professor of Water Engineering and industrial fellow of Royal Academy of Engineering at Centre for Water Systems, University of Exeter, UK. She is the Chair of IWA's Intermittent Water Supply Specialist Group and associate editor of Journal of Hydroinformatics. She specialises in urban water systems modelling, asset management, water resources management, many-objective optimization, uncertainty and risk assessment, and decision aid. Her research interests cover interdisciplinary field of Hydroinformatics including Artificial Intelligence, data mining and optimization techniques and their application for real-time control for smart water systems.

Ina Vertommen

Ina Vertommen is a scientific researcher in the Water Infrastructure team at KWR. She works with her colleagues on the development of the Gondwana optimization platform and translates water-practice problems, such as the design of network masterplans and sensor networks (for instance for sectorization or leak detection), into mathematical optimization problems. Ina also researches the impact of the weather, holidays and vacation periods on water consumption, and has experienced in the detection of leakages and changes in consumption patterns based on the CFPD method. Moreover, Ina contributes to the monitoring and investigation of trends in lead in drinking water in the Netherlands.

References

1 | Jared L. Cohon David H. Marks (1975) A review and evaluation of multiobjective programing techniques. *Water Resources Research,* Volume 11, Issue 2, https://doi.org/10.1029/WR011i002p00208

2 | Lansey, K. E., Duan, N., Mays, L. W., and Tung, Y. K. (1989). Water distribution system design under demand uncertainty. *Journal of Water Resources Planning and Management,* 115(5). doi:10.1061/(ASCE)0733-9496(1989)115:5(630)

3 | Suribabu, C. R., and Neelakantan, T. R. (2006). Design of water distribution networks using particle swarm optimization. *Urban Water Journal,* 3(2), 111-120. doi:10.1080/15730620600855928

4 | Maier, H. R., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L. S., Cunha, M. C., Reed, P. M. (2014). Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions. *Environmental Modelling & Software,* 62, 271-299. doi: https://doi.org/10.1016/j.envsoft.2014.09.013

5 | Mala-Jetmarova, H., Sultanova, N., and Savić, D. (2018). Lost in Optimisation of Water Distribution Systems? A Literature Review of System Design. *Water,* 10, 307. doi:10.3390/w10030307

6 | Rechenberg I., Küppers U., Scheel A., Mattheck C., Harzheim L. (1998) Evolution und Optimierung. *Bionik.* Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-06114-5_14

7 | Kirkpatrick, S., C.D. Gelatt Jr. and M.P. Vecchi (1983) Optimization by Simulated Annealing. *Science,* 13 May 1983, 671-680.

8 | Rogers *et al.* (2007) Application of optimization technology to the development of a water distribution system master plan for the City of Ottawa, American Water Works Association - AWWA Annual Conference and Exposition, ACE 2007

9 | K. Muhammed, R. Farmani, K. Behzadian, K. Diao, D. Butler (2017) Optimal rehabilitation of water distribution systems using a cluster-based technique, *Journal of Water Resources Planning and Management,* 143 (7)

10 | Gephi (2014). Gephi, an open source graph visualization and manipulation software. https://gephi.org (Jun. 22, 2014)

11 | Vertommen, I., Laarhoven, K. v., Thienen, P. v., Agudelo-Vera, C., Haaijer, T., & Diemel, R. (2018). Optimal Design of and Transition towards Water Distribution Network Blueprints. *Proceedings,* 2(11), 584.

12 | Ostfeld A. *et al.* 2008 The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms, *Journal of Water Resources Planning and Management* 134(6), 556-568.

13 | Berry J., Boman E., Riesen L.A., Hart W.E., Phillips C.A., Watson J.P and Murray R. 2010 User s Manual TEVA-SPOT Toolkit Version 2.4. *Office of Research and Development,* USEPA 600/R-08/041B.

14 | P. van Thienen; B. de Graaf; J. Hoogterp; J. van Summeren; A. Vogelaar (2018) Bounds on water quality sensor network performance from design choices and practical considerations. *Water Practice and Technology* (2018) 13 (2): 328 334. https://doi.org/10.2166/wpt.2018.044

15 | Yates, W. B., & Keedwell, E. C. (2020). Offline Learning with a Selection Hyper-heuristic: An Application to Water Distribution Network Optimisation. *Evolutionary computation,* 1-24.

16 | Raad, D. N., Sinske, A., & Van Vuuren, J. H. (2011). Water distribution systems design optimisation using metaheuristics and hyperheuristics. *ORiON,* 27(1).

17 | Matrosov, E. S., Huskova, I., Kasprzyk, J. R., Harou, J. J., Lambert, C., and Reed, P. M. (2015). Many-objective optimization and visual analytics reveal key trade-offs for London s water supply. *Journal of Hydrology,* 531, 1040-1053.

18 | Johns, M. B., Mahmoud, H. A., Walker, D. J., Ross, N. D., Keedwell, E. C., and Savić, D. A. (2019, July). Augmented evolutionary intelligence: combining human and evolutionary design for water distribution network optimisation. In *Proceedings of the Genetic and Evolutionary Computation Conference* (pp. 1214-1222).

19 | Peter van Thienen and Dragan Savić (2020) The importance of knowing what we do not know - Uncertainty in planning, designing and modelling of urban water infrastructure. *IWA Digital Water whitepaper,* 2020

20 | Cunha, M., & Sousa, J. (2010). Robust design of water distribution networks for a proactive risk management. *Journal of Water Resources Planning and Management,* 136(2). doi:10.1061/(ASCE)WR.1943-5452.0000029

Digital Twins - A new paradigm for water supply and distribution networks

By Fernando Martínez Alzamora, Pilar Conejos, Mario Castro-Gama and Ina Vertommen

A digital twin (DT) is a virtual copy (a digital model) of a real system continuously fed with data to mimic the systems' past, present and future behaviour. This makes it possible to detect anomalies, test new ideas and changes in the virtual system and assess how it reacts, minimizing the risks to the real system. In this sense, the DT can be seen as a playground to explore the effects of different scenarios and to practice how to best react and operate the physical system under these circumstances. The concept of DT has been used traditionally in the industry field[1] but it can also be developed and exploited in a city management context, and in particular in Water Supply and Distribution Networks (WSDN), where it can be applied to all aspects of the system[2].

How DTs help for better management of WSDN

A DT can help to make short and long-term informed decisions in order to improve water distribution systems management. In the system design phase, it can be applied to:

- Develop masterplans by simulating the system behaviour under long-term demand projections and new scenarios. This allows for new infrastructure to be designed considering different needs for water, the most appropriate components to be added or replaced, and test the system resilience as a whole.
- Planning reengineering projects aimed at saving energy, integrating new water sources or improving the resilience of the network.
- Design the future operation of the system and determine the new infrastructure commissioning stages.
- Develop a sectorization plan for anomaly detection and gain insight into the performance of the system.
- Determine the best places where to locate the isolation, washout and purge valves for maintenance of the network with minimal disturbance to users.
- Plan the progressive implementation of Automatic Meter Reading (AMR).

For operation and maintenance, a DT can be applied to:

- Achieve a better understanding of the performance of the whole system.
- Train the operators by familiarizing them with the response of the system under different failure scenarios.
- Help operators make the best decisions in real time by simulating the effects of any operation before taking the action in the real system.
- Optimize the operation of the system, minimizing energy consumption and maximizing the quality of the service.
- Plan flushing operations to guarantee good water quality.
- Predict the behaviour of the system under short term demand forecasting.
- Detect anomalies in the system by comparing the observed values with those expected and simulated by the DT, e.g., leaks, valve failures or malfunctioning of other elements.
- Develop emergency response plans, simulating the behaviour of the system under emergency conditions.
- Develop an early warning system against possible attacks or contamination into the network.
- Improve predictive maintenance (i.e., maintenance of components before they fail) taking into account the stresses each component is submitted to and its role in ensuring service.

The components of a DT

Hydraulic model

A detailed and accurate representation of the WSDN, in the form of a hydraulic model, is the basis of a DT. The model should include all elements of the system, from pipes, junctions, demand nodes, reservoirs, pumps, valves and other minor components, to current water demands. Manually building and keeping such a hydraulic model up to date is a laborious task.

The availability and maturity of these models vary between water utilities and around the world. Nowadays, some utilities have detailed models and in general, these are updated annually, regarding water demand (average and peak demand) and new elements in the network. Models are often not updated during maintenance or repair works. This means that, for instance, valve statuses in the real system and the model differ.

The purpose of a DT imposes different requirements on the hydraulic model. For instance, for operation and anomaly detection (water quality or quantity) the hydraulic models need to be continuously updated and paired with the physical systems. In other words, a DT has to include, at every moment, changes made during maintenance/repair, variations in demands and control rules in the operation. To reproduce an isolated segment during a repair all valves must be included in the model, and if the emptying time is to be calculated then the discharge and air release valves must be also included.

Water demand

In addition to the physical elements, to accurately model the flows in the network it is important to correctly assign the demands to the nodes of the model, as well as their evolution over time. Most utilities measure water consumption for the purpose of billing. It can be measured with different levels of aggregation and frequency. In some utilities, for instance, consumption is measured directly at service connections, while in others it is measured through domestic meters. The frequency can also vary from daily, or monthly to once a year. The DT must incorporate the greatest amount of information available in this regard (at least daily, preferably hourly).

Currently, there are WSDN that have digital water meters to read daily or hourly users' demands, like in the city of Valencia (Spain). Incorporating this information into the DT makes it possible to have a more reliable hydraulic model since demands are assigned at the house connection level. In addition, demand patterns can be established depending on the type and number of users supplied, which is of great help in, for instance, locating leaks, regulating the system and managing demand in situations of scarcity.

When information from digital water meters is not available (which is currently expected to be the case for most water utilities around the world), it is necessary to find an alternative way of feeding current demands to a DT, in an accurate and dynamic way. There are several water demand models available in the literature. However, the DT requires more than a model for the average consumption of a typical user, but the actual water demand in a given area at a given moment in time. Besides understanding how consumers use water, it is necessary to know where they are at different moments. External information to grasp people movement throughout the day, like, for instance, data from traffic, use of public transport, energy consumption, and mobile phone data can be used to this end. As an example, KWR Water Research Institute in the Netherlands followed an approach wherein mobile phone data is used to capture population dynamics and couples this information to the water demand model SIMDEUM[3,4]. SIMDEUM is an end-use model that simulates stochastic residential and non-residential water demand patterns, based on statistical data on water appliances and users. In this way, the water demand is dynamically estimated over time based on the actual number of users present at each node of the network. This approach offers an additional advantage: SIMDEUM is able to estimate water demand on very short time steps (up to one second), while smart meters often provide information only at an hourly or daily basis due to battery restrictions. For some applications, such as water quality modelling, one-hour time steps are too coarse. Hence, even in cases where smart meters are available, it could be beneficial to combine both methods by using live measurements to calibrate a SIMDEUM model and then proceed to use such patterns to model demand at shorter time steps.

Real-time data

One of the most important characteristics of DTs is their continuous use of field data to reproduce the real state of the system. We refer here to those variables that change continuously and are registered by Supervisory Control And Data Acquisition (SCADA) systems and sensors in general, such as tank levels, flow meters, pressures, etc. It must be taken into account that these signals can be registered and sent at different times. Connecting them with a DT is therefore not straightforward. In addition, a data management system is necessary to filter and replace incorrect information, which can be a challenge[11].

Computerized maintenance management system (CMMS) services

One of the most outstanding features of DTs is their ability to manage the maintenance of an industrial product or an installation, by continuously monitoring its behaviour and evolution through the measurement of the most relevant variables and subsequent analysis. Unlike classic predictive maintenance systems, which are based solely on statistical data analysis, a DT provides the additional ability to reproduce past, present and future dynamic behaviour of the system as it is a virtual replica of the real system continuously updated and calibrated from a reduced number of measurements. For that, the hydraulic model must incorporate all the maintenance operations carried out since they affect the state of its elements. This is possible if the hydraulic model is linked to the CMMS. In this way, the maintenance management can be improved with the capability of incorporating predictive maintenance, based not only on the expected use of the different components, but its real behaviour as part of the system.

Additional information sources

A DT has to incorporate also complementary information that affects its behaviour or decision making, such as topography, availability and quality of water sources, type of dwellings and local facilities, types of consumers, electricity tariffs, weather forecast, and social behaviour, amongst others.

Calibrating a DT

A DT has to behave like the real system, so the calibration of the hydraulic model is crucial to achieving a reliable DT. There are different techniques and methodologies for calibrating a hydraulic model. It is useful to develop an initial pre-calibration stage, reviewing and correcting all possible errors in the information, and only afterwards calibrate the model parameters. Fortunately, with the DT many scenarios are continuously available, which allows for frequent calibration of the hydraulic model, instead of using only for single situations or days, as it has traditionally been done.

One of the aspects which is commonly challenging for model calibration is the demand allocation. In cases where smart water meters are available, such as in Valencia (Spain), these data can be used to calibrate the model. Internal pipe diameters and roughness coefficients are the most relevant parameters bound to uncertainty. In other countries, such as in the Netherlands, the calibration of water distribution network models is an iterative process of updating of assets, demands and operation criteria. For most utilities, these updates are scheduled at an annual or bi-annual basis. Almost all systems in the Netherlands are operated as a single zone, for that reason most data are known at the booster stations, and only a few flow meters and pressure loggers are located within the network. While industrial and 'large' users, such as sports facilities, carwashes and hospitals, are monitored using digital water meters, this is not the norm at the household level, introducing a serious challenge to model calibration.

The existence of a large number of valves in the distribution system poses an additional challenge for modelling. Although most valve manipulations are registered, it is estimated that at least 2-3% of the valves are not properly displayed or their status is changed (i.e., partially opened). This requires a large amount of effort as these mis-registrations are not easy to detect until additional operations in the surroundings are performed. Likewise, the presence of numerous regulation elements can significantly complicate the calibration of the model[5].

DTs, Decision Support Systems and Artificial Intelligence

One of the most important reasons that justify the development and maintenance of a DT is to use a replica of reality as Decision Support System (DSS). The concept of DSS has usually been linked to optimization techniques, aimed at minimizing one or more objectives, whether technical or economical, by modifying the values of the decision variables subject to certain restrictions; sometimes some of these restrictions can be relaxed by being incorporated as additional objectives. Among the most important applications of DSS in the field of the WSDN that can be cited are those used to determine the adjustment parameters in a calibration process, for optimal sizing of pipes and control elements, for optimal location of valves and other accessories to facilitate the network maintenance, for optimal identification of Demand Metered Areas (DMA's) in a sectorization plan, for optimal sensor location to identify leaks or for the early detection of contaminant intrusion, for optimal operation of the system to reduce energy consumption or the associated cost, to design optimal strategies to renew water in stagnant areas or to reduce the retention time in tanks, to plan preventive maintenance operations, to plan investments in asset management, etc. Methods used by optimizers to achieve these goals range from classical Linear Programing (LP), Mixed-Integer Linear Programming (MILP) and Non-linear programming (NLP) to the most advanced Evolutionary Algorithms (EA), depending on the nature and complexity of the problem, and the type of variables involved. For some of these applications a simplified model of the network may be sufficient, but in others cases it is very important to take into account each and every one of the elements that make up the real network as per a DT, for example in asset management, maintenance issues, network sectorization, etc. When simplified models are enough, they can also be derived from a DT.

Another field in which DTs can play a relevant role is to exploit the capabilities offered by modern advanced analytics techniques and Artificial Intelligence (AI)[6]. Optimization should not be confused with them. They are fundamentally based on the data observation of a large number of real situations and on learning about the system behavior from these data, which usually come from field sensors but not limited to them. In its application to WSDN management, the signals will come from the SCADA system, from the maintenance management system or from the remote readings of consumption. But if we have a well-calibrated DT, the training variables can also be synthesized from the results provided by the DT under certain scenarios, with the advantage of its low acquisition cost and the high quality of the data, in contrast to the data taken from the actual operation of the system. AI uses Machine Learning (ML) algorithms to achieve its purposes. In a first instance, they can be arranged in supervised, unsupervised, mixed, or reinforced, being the latter the most promising for future.

In supervised learning, sets of paired values for the input and output variables are given. The algorithm must be able to reproduce the outputs from the inputs with the minimum error. Actually, the classical regression techniques would fall in this group, but in the last decades other much more powerful methods have been developed to tackle more complex problems having a high number of input/output variables with strongly non-linear relationships, such as k-Nearest Neighbors, Logistic Regression, Support Vector Machines (SVM), Decision Trees (DT), Random Forest (RF) and particularly Artificial Neural Networks (ANN), initially developed around the concept of the Multilayer Perceptron (MLP). In recent years ANN have been developed greatly with the introduction of new architectures under the concept of Deep Learning, like the Convolutional Neural Networks (CNN) and the Recurrent Neural Networks (RNN), with results as astonishing as facial or speech recognition.

Unsupervised learning instead tries either to group the set of data (observed or synthetic) into differentiated classes using cluster analysis techniques, to reduce the size of the problem, or to discover behavior laws among the data set, in an attempt in all cases to abstract the information and synthesize it, which constitutes one of the pillars of the development of human intelligence. Compared with the classical statistical techniques used for this purpose such as k-Means, Hierarchical Cluster Analysis (HCA), Expectation Minimization (EM) or Principal Component Analysis (PCA), the ANNs, and in particular the architectures associated with Deep Learning such as the Autoencoders and the Reinforced Learning (RL), seem very promising.

To finish this brief description of the state of the art of AI, it should be noted that in any application it is necessary to differentiate whether the variables managed are continuous or discrete, if the data set is static or dynamic (real-time systems), and in the latter case, if the goal of the algorithm is to properly reproduce the recent past or to forecast the future.

All these techniques have applications to WSDN management, although most of them are still incipient compared to other areas such as image recognition, marketing or business. The first applications have been aimed at characterizing and classifying demand patterns, by differentiating the type of consumers or the effect of exogenous factors such as the day of the week, the season or the temperature. Past data and unsupervised methods are used for this purpose. These patterns could be used later to detect deviations from the expected values, due to the occurrence of a leak for example, and to issue the corresponding warnings.

Other applications try to directly predict the demand in the coming hours, either at consumer or sector level, based on past and recent data. In contrast to the classic Box-Jenkins techniques, which assume a linear behavior of the time series, Recurrent Neural Networks (RNN), and in particular Long Short Term Memory (LSTM) networks characterized by progressively reducing the weight of the oldest readings, have provided so far the best results. The RNN is fed in this case with continuous and dynamic data to carry out a supervised training.

AI techniques can also be used to detect sudden anomalies, such as a pipe break, a sensor failure, or a contaminant intrusion into the network. Supervised methods fed by synthetic data provided by DT can be used to train convolutional ANNs for this purpose. However, when the nature of the anomaly is not anticipated, unsupervised methods would be more appropriate.

Regarding predictive maintenance, the use of AI techniques can lead also to significant advances to improve WSDN management. Supervised training techniques such as Decision Trees (DT) or Random Forest (RF) have been mainly applied for this purpose, but using the Gradient Boosting, a variant of RF more suitable when the number of leaves on the tree is reduced, or the LSTM already discussed above, are more promising in the future. The source of data in this case must be real data because it is very difficult to physically model a fault. In WSDN is common to have a lack of recorded data concerning faults and maintenance operations, so a greater sensorization is needed in the future to take full advantage of these techniques. One of the most important applications along this line would be the capability to anticipate new leaks.

By considering the power of AI, new applications for improving WSDN management are constantly arising. For example, AI can be used to fast respond in emergency situations, to reduce the daily energy consumption, to manage the pressure in DMAs in order to reduce leaks or to control demand, to manage DMAs in case of unforeseen incidents, and to detect incipient leaks by observing the drift of certain signals in a zone. All these applications require huge data for training the AI algorithms, but fortunately DTs working upon well calibrated models can produce such data automatically at low cost, by subjecting them to multiple randomly generated scenarios. A training data set can be built just with the results of the simulations or with the outcomes of a subsequent optimization process looking for the best solution for each scenario, thus combining optimization with AI techniques. In the future it is possible that, thanks to the power of Reinforced Learning (RL) algorithms, ANNs can reach by themselves the optimal solution to each situation posed thanks to a previous self-training process aid by DTs, just as AlphaGO Zero learned to play GO on his own, defeating the world champion in 2017, without the need for a prior supervised training.

Viewers and User Interface

For a DT to be used by water utilities, it is necessary to build a user-friendly and intuitive graphical user interface. The interface has to be interactive and fine-tuned to its use (daily operation or long-term design for instance).

As a DT manages a significant amount of information of different nature and origin it can be useful to use a combination of different products and interfaces, such as Application Programming Interfaces (APIs), web services, map-based interfaces, GIS integration, dashboards, and web interfaces.

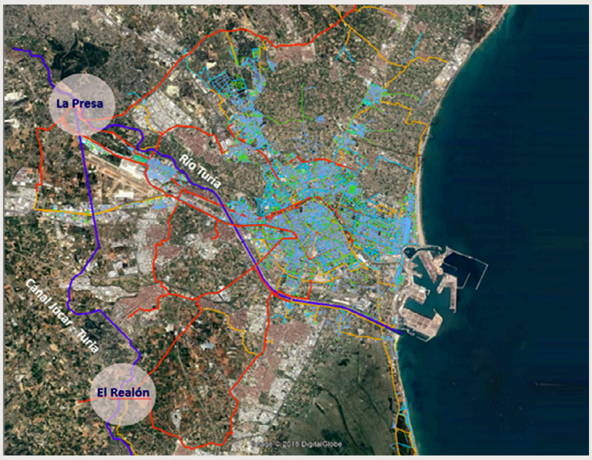
Applications

In this section, two application cases (at different maturity levels) are presented to illustrate the possibilities, benefits and challenges of DTs applied to WSDN. These cases refer to the DTs of Valencia, Spain and Eindhoven, The Netherlands (**Figure 1**).

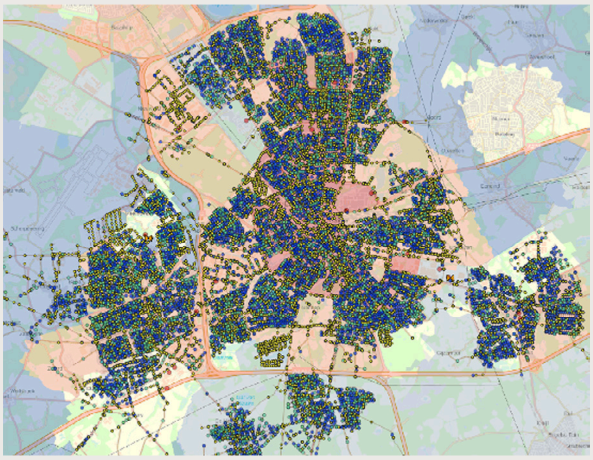
The DT of Valencia

Today Global Omnium (GO) operates a DT for the water distribution network of Valencia Metropolitan Area. The DT works upon a hydraulic model connected with the main sources of the information provided by the physical system. The addition of advanced analytics like AI starts to exploit the potential of the DT, particularly to identify demand patterns, to forecast demands and to detect anomalies in the hydraulic variables. In a near future much more applications are envisaged.

The first strategic model of the city of Valencia was created in 1993 in collaboration between GO and the Universitat Politècnica de València (UPV) and since then, significant progress has been made. In 2007 the hydraulic model was connected to SCADA for the first time[7] in order to run live simulations and help the operators make decisions in the Network Control Center. The AMR implementation in Valencia opened up new opportunities, so in 2016 GO and UPV began the ambitious project of building a full DT for the system by connecting the hydraulic model with all information sources: SCADA, sensors, GIS, CMMS, AMR, etc. The DT had to be interoperable with new IT platforms and be scalable to any size of the supply system. The result was the Digital Twin developed with GoAigua, a smart water platform by Idrica, a Spanish company that provides technological services. It is now fully operational and in use in the Control Room of the water supply system of Valencia and its metropolitan area. The model is connected in real time with 600 sensors and replicates the real behaviour of the network with a 95% accuracy for flows and 98% for pressures[5]. It is now a vital tool in support of decision-making for both daily operations and planning tasks[8].



Valencia
1,700,000 inhabitants | 101,300 junctions | 113,267 pipes | 46,801 manual valves



Eindhoven
223,300 inhabitants[1] | 29,444 junctions | 57,194 pipes | 13,828 manual valves

Figure 1 | Characteristics of the WSDN of Valencia and Eindhoven.

How the DT is built and maintained: detailed and strategic models

The Valencia DT uses the GoAigua platform to integrate information from various sources. From there a set of algorithms configuring the application GO2HydNet, builds automatically from scratch and by querying different sources for the required information, an EPANET-based detailed model for the whole network or a selected area, which reproduces with accuracy its behaviour for a certain time period. This detailed model includes all the pipes, operating elements and auxiliary elements that affect water flows, and is connected with the SCADA information to make live simulations. Hence, it can be used as an assistant to test and make real-time decisions. Building the detailed 24-hour model of Valencia, including service connections with their corresponding consumption pattern when available, to reach a complete model of 325,000 nodes takes about 1 minute of computation time on a standard PC i7-3.2 GHz (the time required for data pre-processing is not included). Thus, as data sources are updated, the model is also updated. However, depending on the use, a strategic model containing only the main elements is more useful to have a general view of the systems' behaviour. For this reason, the Control Center works with a 10,000 nodes strategic model—a simplified model obtained from the detailed model, and always connected to it. This strategy makes it possible to perform every operation with either the detailed or the strategic model, in such a way that, both models together constitute the DT of Global Omnium (**Figure 2**).

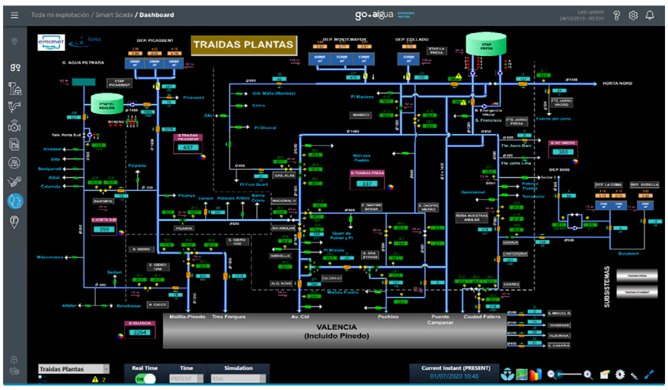


Figure 2 | The DT of Valencia WSDN connected in real time with field data. Real-time data are compared with simulated ones next to each box. The proposed actions can be simulated before carrying them out in the real system.

Use cases

GoAigua's DT is used in GO for planning, design and management of the daily operations in the Valencia Metropolitan Area since it provides a complete overview of the network in real time, along with informative and actionable dashboards 24/7. Valencia's DT provides operational teams with:

- Simulation of past, present and future scenarios under all kinds of operating conditions.
- On-the-fly analysis of what-if situations for both present and the future, facilitating support decision-making on the best time for network maintenance and other operations.
- Anomaly detection: the DT calculates in real time pressures and flows at all nodes and pipes of the strategic model, providing a great understanding of the performance of the system and allowing the fast detection of incidents (**Figure 3**).
- Forecast of the network behaviour in the next 24 hours, which facilitates the prediction of potential events.
- A playground for training new staff in network operations.

Valencia's DT is also used for planning tasks such as:

- Development of contingency plans for emergencies.
- Designing the new infrastructure required according to the network needs.
- Defining, in advance, the operation of the new infrastructure and determine the network commissioning stages.
- Planning long-term actions, including investments to optimize Capex and risk levels.

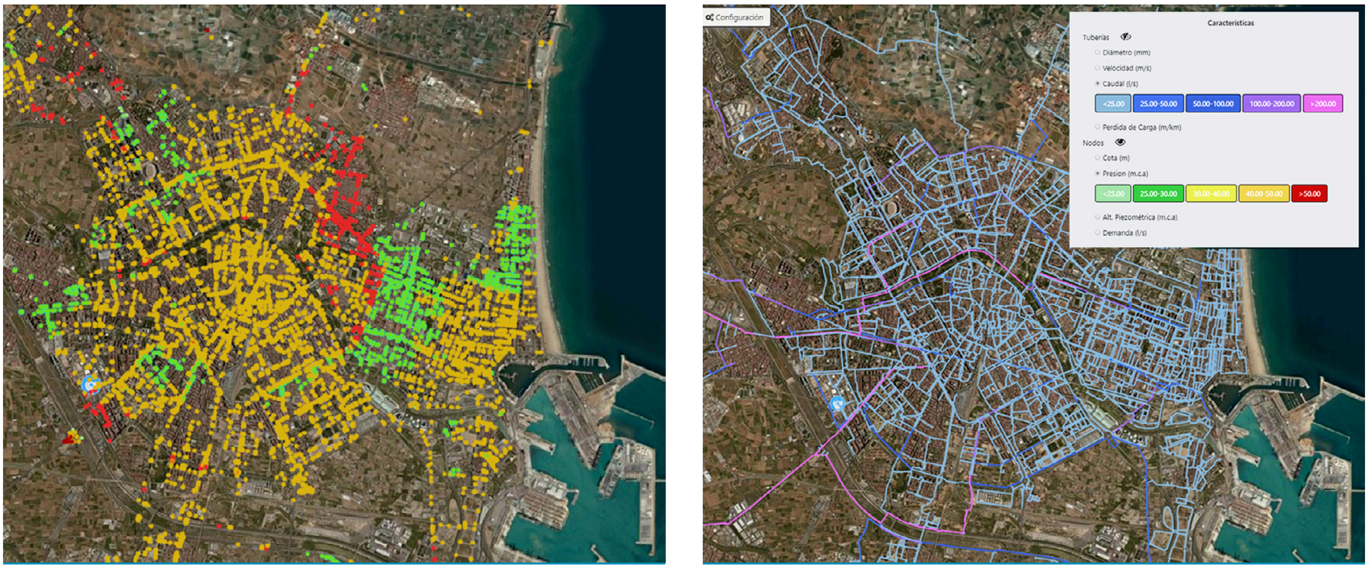


Figure 3 | With 600 measurements (pressures, flows and levels), it is possible to know in real time all flows (right) and pressures (left) of a 10,000 nodes strategic model.

The DT of Eindhoven

Recently KWR took the first steps in building a DT of the WSDN serving the city of Eindhoven with about 223,300 inhabitants[9]. The WSDN, operated by Brabant Water, is fed by a total of five pumping stations. One of the aspects to consider is that the topography of the city is reasonably flat. For that reason, this network is operated as a single pressure management zone. This means that no sectors, like district or pressure metered areas, are implemented, and the available flow measurements regard the total area. The assets registration is in general of excellent quality and continuously updated by the Brabant Water operators. The network model is coupled to different data sources, in particular, data from mobile phones (i.e. how many mobile phones present in a given area at a given moment in time) are used to capture population dynamics, and together with weather, land use and population data (ranging from numbers of inhabitants to household size and composition), which feed the water demand model SIMDEUM. Maintenance and repair activities, as well as unplanned events such as leakages, offer additional relevant information. By linking all the aforementioned data, the DT was used to model water demand, pressure and flow in the WSDN of Eindhoven, at three different times in the year: a regular week, a week with warm temperatures and a week during the vacation period. From the obtained results it is clear there is a correlation between the number of users in an area (estimated with the mobile phones) and the water consumption measured in the same period.

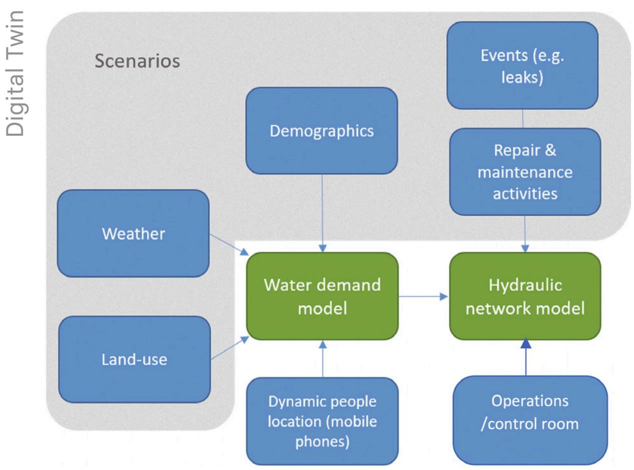


Figure 4 | Conceptual design of DT using different data sources in the Netherlands. Demographics and land use are based on large governmental databases (Kadaster, CBS). Weather data is available from the meteorological organization (KNMI). Dynamic population information is obtained using mobile phone data (third-party vendor). The water demand model is fed with this information using a stochastic demand simulator (SIMDEUM®), while the hydraulic network model is obtained from the water utility (Brabant Water). Several scenarios can be simulated based on this information.

Moreover, it was possible to model the effects of leakages and wrongly registered valve statuses on the network performance, for both regular, warm and vacation periods, identifying for instance areas with lower pressures. This information is valuable for water utilities to help them identify sensitive areas and to anticipate how to best operate the network when facing particular conditions.

The conceptual design of the DT (**Figure 4**) is suitable for both (near) real-time modelling and scenario analysis. The use of mobile data as input for consumer demands suggests that aspects of population dynamics can be integrated into a DT. However, at the time being, mobile phone data are not (yet) available on a real-time basis in the Netherlands, making the proposed approach more suitable for scenario studies. Once the data of mobile phones can be fetched in a timely manner, the DT could be used for real-time modelling as well.

The conducted research shows the potential of DTs for the Dutch drinking water industry. In the Dutch context, DTs can be developed in the short-term, as water utilities have good network models, and multiple data sources are already available. In the upcoming months, the Eindhoven DT will be further developed to include additional data, such as traffic information, and to better model non-household water demand. Dutch water utilities want to use DTs to understand the effects of the lockdown and other governmental measures imposed in control the spread of Covid 19 in combination with the drought.

1 | For privacy reasons the smallest area for which data is available is the 4-number postcode used in the Netherlands (more detailed postcodes include 4 numbers and 2 letters). If less than 10 mobile phones are available in the area, the data is not shared. Only the number/amount of mobile phones is shared, data about the mobile phones such as the number/owner is not. Data treatment takes 2 weeks' time.

Conclusions

The introduction of DTs in the reality of water utilities requires a paradigm shift in the way of managing WSDN. The current capabilities of computers to simulate the behaviour of networks in real-time is beyond debate, even using a standard PC. The main challenges lie in sensor deployment of the networks and the collection and treatment of a large amount of data, ranging from SCADA signals to consumption data and maintenance operation. The potential results of this complex effort are considerable in improving the service provided to customers, as has been shown in the case of the DT of Valencia. While sensor data is not widely available (the shift towards smart water metering, for instance, can take years), one can think of alternative data sources and modelling approaches, such as those used in the Netherlands taking advantage of mobile phone data in combination with SIMDEUM to better model water demand. The cost of mobile phone data is lower than that for large-scale implementation of digital water metering. However, a complete analysis of the relationship between mobile phone data and water consumption during different seasons must be performed in order to reduce the uncertainty of their use within DTs.

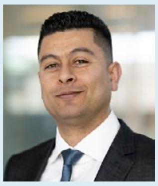
The current coronavirus pandemic imposes new short-term challenges to water utilities as consumers change their habits. In combination with other factors such as drought and ageing infrastructure, DTs are a powerful tool to provide insight into network behaviour under new circumstances and into how to best operate it in the long and short term. For many operators, a WSDN is a black box. DTs are a key step in unravelling it.

Fernando Martínez Alzamora

Fernando Martínez Alzamora is a PhD industrial engineer and full professor of Hydraulic Engineering since 1995 at the Universitat Politècnica de València, Spain, where he teaches on hydraulic machinery, hydropower plants and renewable energy systems. He has been director of the Institute of Water and Environmental Engineering at the same University. His main research is on water distribution modelling and the efficient use of water and energy in water supply and irrigation systems, working very closely to the industry sector. He has led the development of several software packages oriented to modelling the behaviour of the hydraulic systems, using jointly GIS info and real time data provided by SCADA systems. He has been leader of many research projects, funded by the Spanish research councils, European Commission or Spanish industry, and has published more than 120 technical papers on these topics in indexed journals, magazines and conference proceedings.

Pilar Conejos

Pilar Conejos has a PhD in industrial engineering and is a part-time assistant professor at the Universitat Politècnica de València, Spain. She has over twenty years of experience in hydraulic engineering and water management. She has been for a long time Head of Water Network Control and Regulation for Valencia and its metropolitan area at Global Omnium. She is the co-author of several technical papers and is an international guest speaker.



Mario Castro-Gama

Prior scientific researcher at KWR Water Research Institute, now specialist infrastructure for Vitens B.V. Netherlands. He is part of Vitens' Water Expertise Center (WEC). Currently, he is in charge of different projects related to digital water and water distribution and quality monitoring and asset management. Born in Bogota, Colombia. Civil Engineer from the Universidad Nacional de Colombia and MSc in hydroinformatics from UN-IHE Delft. He has more than seventeen years of experience, and he is author of several articles and proceedings in subjects ranging from hydraulics to hydrology.

Ina Vertommen

Ina Vertommen is the team leader of Hydroinformatics at KWR and a scientific researcher in the fields of water demand modelling, robust design of water distribution networks and optimization problems. She likes to explore the added value and application possibilities of different new techniques and methods for the water sector. She works with her colleagues on the development of tools, such as the optimization platform Gondwana. Ina gets energy from collaborating with researchers and water management professionals and strives to achieve innovative and valuable research results and solutions together.

References

1 | Tao, F., and Q. Qi. (2019). "Make More Digital Twins." *Nature* 573 (7775): 490–491. doi: 10.1038/d41586-019-02849-1.

2 | Conejos, P.; Martínez Alzamora, F.; Hervás, M.; Alonso, J.C. (2020): "Building and exploiting a Digital Twin for the management of drinking water distribution networks", *Urban Water Journal,* doi: 10.1080/1573062X.2020.1771382

3 | Blokker, E.J.M.; Vreeburg, J.; Pieterse-Quirins, E.J.; van Dijk, J. (2010) "Simulating Residential Water Demand with a Stochastic End-Use Model", *Journal of Water Resources Plannning and Management,* 136(1) 19-26, 2010.

4 | Blokker, E.J.M.; Agudelo-Vera, C.M.; Moerman, A.; Thienen, P. van; Pieterse-Quirijns, E.J. (2017) "Review of applications for SIMDEUM, a stochastic drinking water demand model with a small temporal and spatial scale", *Drinking Water Engineering and Science 10* (2017) 1-12. doi: https://doi.org/10.5194/dwes-10-1-2017.

5 | Martínez Alzamora, F.; Conejos, P.; Urbán, P. (2017). Calibración de un Sistema de Distribución de Agua Complejo a partir de variables SCADA. *V Jorn. Ingeniería del Agua,* Octubre 2017, A Coruña.

6 | Savić, D. (2019). What is Artificial Intelligence and how can water planning and management benefit from it? *IAHR White Papers.*

7 | Bou, V.; Martínez Alzamora, F.; Conejos, P. (2006). SCARed, a general purpose SCADA application for taking decisions in real time with the aid of a hydraulic model. *8th Annual Water Distribution Systems Analysis Symposium. WDSA.* Cincinnati (Ohio).

8 | Conejos, P.; Martínez Alzamora, F.; Gamón, A., Urbán, P., Alonso, J.C. (2017). La simulación hidráulica como herramienta de apoyo a la toma de decisiones en Global Omnium–Grupo Aguas de Valencia. *XXXIV Jorn. AEAS,* May 2017, Tarragona.

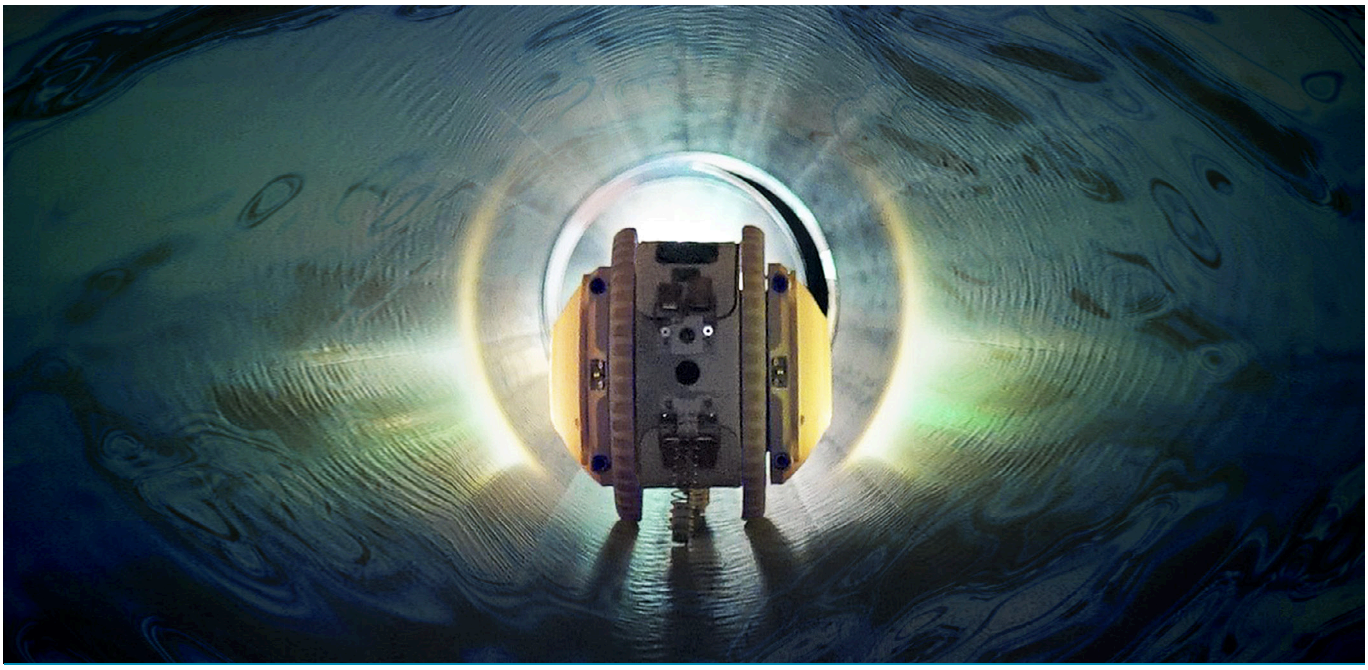
9 | Castro-Gama, M. *et al.* (2020) "Using mobile phone data in digital twins for drinking water distribution networks". Report Number: BTO 2020.029 pp 43. *KWR Water Research Institute.* Nieuwegein, Netherlands.

10 | CBS (2018). Dashboard: de inwoners van Eindhoven 2017. Last viewed online: 2020-09-01. https://www.cbs.nl/nl-nl/maatwerk/2019/45/dashboard-de-inwoners-van-eindhoven-2017

11 | Castro-Gama, M.; Agudelo-Vera, C.; Bouziotas, D. (2020) A bird's eye view of data validation in the drinking water industry of The Netherlands. In: *The Handbook of Environmental Chemistry.* Springer, Berlin, Heidelberg. https://doi.org/10.1007/698_2020_609

Autonomous robotics for water and sewer networks

By Stephen R. Mounce, Will J. Shepherd, Joby B. Boxall, Kirill V. Horoshenkov and Jordan H. Boyle



Smart water networks are at the forefront of investment plans for water companies in the developed world as part of a progression to a circular economy. Technological advancements allow water companies to gather more information about their networks and assets than ever before and to connect the sector to the Internet of Things (IoT). Edge computing will help make IoT rollouts more integral and core to the way businesses work in coming years using new sensing approaches including using in-pipe robotics. Pervasive Robotic Autonomous Systems (RAS) will facilitate a move from reactive to truly proactive practice, enabling ongoing and repeat assessment of pipe condition and operational performance. It is foreseen that robotic inspection and data collection will add increasing amounts of data to more traditional data sources. The more intelligence that is captured, the more that can be learned, understood and predicted about the network. Extra data provides new opportunities for asset condition monitoring, performance assessment, maintenance and event analytics. This article provides background on cutting edge research which aims to revolutionise buried pipe infrastructure management with the development of swarms of microrobots designed to work in underground pipe networks autonomously and cooperatively. New Artificial Intelligence algorithms are being developed that uniquely incorporate Lagrangian (mobile sensing) rather than traditional Eulerian (fixed sensing) based coordinate systems. The resulting big data can be used for pipe condition assessment and to inform simulation of hydrodynamic performance of pipe networks, for example identifying pinch points or spare capacity.

Water utilities operate large, complex pipe networks with often limited information on system connectivity and asset condition. The unknown condition, performance, and often even location, of such buried assets is a significant problem for the companies that manage large networks of pipes. In the EU, buried water and wastewater networks have a combined length of approximately 6.1 million km; the replacement value is an estimated EUR 3.5 trillion[1]. While pipe inspection technologies used by these industries have progressed, the lack of comprehensive knowledge about the condition of buried pipes results in sporadic, unforeseen failures. For example, there are 1.5 million road excavations per year in the UK causing full or partial road closures and a cost to the UK of at least GBP 5.5 billion per year (GBP 7 billion including dig costs)[2]. The repairs are conducted in a reactive fashion, with well-developed and efficient industry protocols. However, with current industry replacement rates of less than 1% annually[3] inferred asset lives are 100-800 years. Without the transformative step-change in pervasive sensing proposed herein, this situation will worsen exponentially as infrastructure ages.

Most utilities manage their pipe networks using several data sources: (i) historical records such as pipe location; (ii) age and material; (iii) system failure notifications (including surface leaks, no customer supply, sewage spills, etc.); (iv) and, in the case of sewers, limited contemporary inspections. Action is typically only taken reactively once failure occurs and performance is compromised. Such asset failure is undesirable because it can cause service disruption or economic loss to the customer, damage to other infrastructure such as roads plus the potential polluting environmental impact of spills and added congestion when unplanned roadworks are required to repair or replace the asset.

There are no 'business as usual' inspection systems which measure pipe condition accurately over time at a high spatial granularity (thus enabling models of degradation). As an example, currently wastewater networks are generally inspected using CCTV, requiring a manually operated camera to be passed through the network. The collected footage is analysed by a trained engineer. Manual techniques such as listening sticks or acoustic noise loggers are similarly utilised in drinking water distribution systems. All these reactive approaches require human intervention, cause disruption at the surface, particularly on roads, and are difficult to apply in complex networks.

Internet of Things (IoT) objects and sensors with IP addresses can be connected via the cloud giving rise to the concept of 'smartness' and the development of 'Smart cities' and 'Smart Water Networks' (SWaN). Smart water means using technologies for optimising water resources and waste treatment, monitoring and controlling water, and providing real-time information to help water companies and households manage their water better [4]. SWaNs are currently being rolled out on scale[5] and autonomous monitoring systems hold the key to transform our awareness of inaccessible buried pipe infrastructure. SWaNs have been described as a layered architecture, beginning with the sensing-and-control layer through continuous and pervasive data collection, proactive data management, and ending with the data fusion-and-analytics layer[6]. It seems clear that in the future the whole water sector is going to be completely penetrated by Information and Communications Technology (ICT) and IoT-like technologies. In a decade, tens or even hundreds of petabytes of data may be routinely available. As these technological capabilities advance, so does the ability to collect information from remote devices and correlate that information across diverse systems. An infrastructure that can connect the monitoring and control systems to an IoT platform allows effective use of the operational information that the systems hold to help achieve near-real time situational awareness. Demands for solutions and tools will become more urgent to meet the aspiration for intelligent water networks, proactively managed through access to timely information. While a step change, the spot sensors of a SWaN can only be installed where there is access to the pipe network, e.g. manholes and fire hydrants. While frequent in the networks such features are still only a tiny fraction of the total systems, and commonly only the performance at these points can be measured, not the condition of the infrastructure between these access points. Autonomous pervasive robotics offers the potential to radically transform this situation by going from spot sensing (at fixed locations) to pervasive, Lagrangian (mobile) sensing.

State of the art robotic devices for buried water pipes

A comprehensive review[7] of robots for pipeline inspection revealed that robots currently available are mainly laboratory prototypes designed for large diameter pipes, human controlled, heavy (tens or hundreds of kg) individual devices suitable for a single short duration intervention. Locomotion is often limited to wheeled or tracked approaches. Hardly any of these devices are autonomous.

Autonomous robots appear to have great potential for inspecting difficult to access water pipe networks[8]. A report[9] on Robotic Autonomous Systems (RAS) by TWENTY65 (www.twenty65.ac.uk), a UK collaborative initiative between academic research institutions and the water industry, sets out the opportunities for the use of RAS in the Water Industry, specifically for use in underground infrastructure and more generally in all operational activities in water. A key opportunity was identified as "mapping, condition assessment and rehabilitation within underground pipe assets". The report confirms that Inspection Robots are usually multi-sensor platforms that carry a variety of condition assessment tools inside the pipeline in a single deployment that also provides live video (CCTV) that can aid in detecting anomalies within the pipe. These also tend to be tethered tools which provide condition assessment with limited spatial and temporal resolution, and require human intervention and service disruption.

Tethered robotic crawlers suitable for water and wastewater are available with multiple sensors for condition assessment including laser profiling. There are technologies that can assess a variety of pipe materials to identify structural deterioration that could lead to pipe failure. Examples are the PipeDiver® and Sahara® which are free swimming and tethered devices, respectively.

An alternative to crawlers is the so called 'soft' robot, such as Lighthouse (https://www.digitaltrends.com/cool-tech/ leaky-pipe-detecting-robot-james-dyson/) that is a low-cost unit designed to travel through water pipes hunting for leaks before they turn into major problems. Lighthouse is inserted into a water pipe by way of an existing hydrant. It then passively flows through the pipe, traversing around pipe elbows, discovering leaks by measuring the suction associated with escaping water. The device can then be retrieved when it is flushed out of the pipes through a hydrant, and wirelessly downloads a map of leaks. Smartball (https://puretechltd.com/technology/smartball-leak-detection) is a similar passive untethered approach which is a 'dumb' ball following the current (flow) through water, wastewater, and oil and gas pipelines that can complete long inspections in a single deployment. It should be stressed that such devices are non-autonomous and driven by network flow, designed for single release and inspection. Positional and condition data is of poor quality and stored on-board, and since they are used alone, data coverage is sparse because these devices must follow the flow and are not able to deviate from this path. Recovery of such passive devices can be challenging in complex and uncertain networks.

In the last decade, a number of interesting projects and initiatives have explored the feasibility of RAS for the water sector and several of the more promising are now outlined.

Ariel. KWR Watercycle Research Institute and Wetsus, alongside

Dutch water utilities, developed an initial prototype for an Autonomous Inspection Robot: Ariel (https://www.kwrwater.nl/en/actueel/autonomous-inspection-robots-game-changer-for-asset-management/). Van Thienen *et al.*[10] provided details on a tethered prototype and its testing, in particular its modular design as a segmented 'snake like' train with modules for propulsion/vision, centering and battery/electronics. Van Thienen *et al.*[11] provided further progress with prototypes and the design of a large pilot scale network for testing and presented a comprehensive business case study by way of costs and benefits. The robot's further development has resulted in autonomously operating robots equipped with various sensors which determine the condition of the pipe with exact positioning (x, y, z coordinates). Base stations provide locations for up/downloading of route/inspection data and recharging of batteries. A data ecosystem framework facilitates the analysis of large volumes of sensor data. In practice, testing is ongoing, with further development under the banner of SubMerge b.v.

EU TRACT project. In collaboration with SINTEF and Spanish and Italian research partners, the project (https://www.sintef.no/en/latest-news/2014/robot-water-pipe-inspectors/) has developed a long, torpedo-like and propeller-driven robot equipped with 64 large ultrasound transducers. This is designed to operate in branched water and district heating pipe systems in pipe diameters from 0.1 m, with a range of 150 m. It collects data which enable the calculation of the thickness of, and levels of corrosion in, the pipes.

TISCA. In Netherlands, the Technology Foundation STW, together with Stichting RIONED, STOWA and Kennis Programma Urban Drainage (KPUD), have been cooperating since 2016 on the programme Technology Innovation for Sewer Condition Assessment (TISCA) (https://www.nwo.nl/en/researchprogrammes/joint-programme-technology-innovation-sewer-condition-assessment-tisca). Five projects are currently in progress and of particular interest is FOULC (Fast Over-all scanning of Underground and Linear Constructions). An aquatic drone is being developed as a sensor platform and data-acquisition system for sewer systems. Use of a laser scanner, IR camera and turbidity/velocity profiler were investigated, with preliminary laboratory results reported in [12].

PUB robotics. Singapore's National Water Agency and NTU, with co-funding from the National Research Foundation, developed a mobile robotic platform that can travel in trunk sewers, which is equipped with CCTV, profiling sonar and laser scanners for monitoring the sewers[13]. The initial objective was to design and develop a sewage inspection robot to inspect concrete sewage tunnels with internal diameter of 3 m or larger and for incursions of up to 400 m. An Unmanned Aerial Vehicle (UAV) system[14] equipped with cameras and sensors has also been developed and deployed to inspect the Deep Tunnel Sewerage System. The system is capable of autonomous operation in a signal-denied environment.

Despite these interesting studies, it is reasonable to question why products have not in general reached the market to date. One of the barriers is the difficulty in testing prototypes in realistic networks (note that[11] tackles this by means of a full scale network above ground with various network elements). Other challenges limiting deployment relate to:

- Developing a timely and affordable capability to inspect and quantify performance of individual assets in large pipe networks.
- Synthesizing the inspection data to enable planned intervention at an asset level and prevent unforeseen failure and unplanned repair.
- Ensuring end user requirements are strongly embedded so that pervasive data can be transformed into knowledge that is actionable to prevent failures.
- Implementing maintenance based on such additional information derived from data i.e. providing capacity to act in a timely manner.

A future RAS highway for water infrastructure

Significant problems must be addressed and solved to make buried water pipe infrastructure a robot-friendly RAS highway. Autonomous robots can cover the whole infrastructure moving freely whenever required to detect objects, obstacles and contraventions to the norm, supplying data continuously on an unprecedented scale and integrating safety into their decision making including motion control through dynamic motion planning. High level communications, map generation and adaptive planning through optimisation of space and time usage, analysing motion and analysing power usage will all be essential components to enable adaptive strategies in complex pipe networks.

Robotic autonomous systems are differentiated from other machines by their ability to perform physical tasks with little or no human intervention. They have the potential to enact a wide range of individual tasks without direct human supervision. Their work is a combination of the following three sub-tasks: (i) manipulation and processing; (ii) data gathering and monitoring; (iii) data sorting and storage. Artificial Intelligence (AI) methods and tools are essential for success with these tasks because of the massive volume of data and complexity of the problem associated with the inspection of buried pipes. AI methods have already been embraced by many water utilities which use them to support the planning, operation and maintenance of their distribution and sewerage networks, improve customer service and predict demand[15]. Robots are widely used in other industrial sectors and the significant development of AI and machine learning will result in a rapid growth of RAS having a major impact on nearly all market sectors within the next decade. This economic impact is not just related to an expansion in the market for robotics technology but also to the deep impact robotics technology will have on competitiveness and service provision across all economic sectors. The early signs of this impact are visible

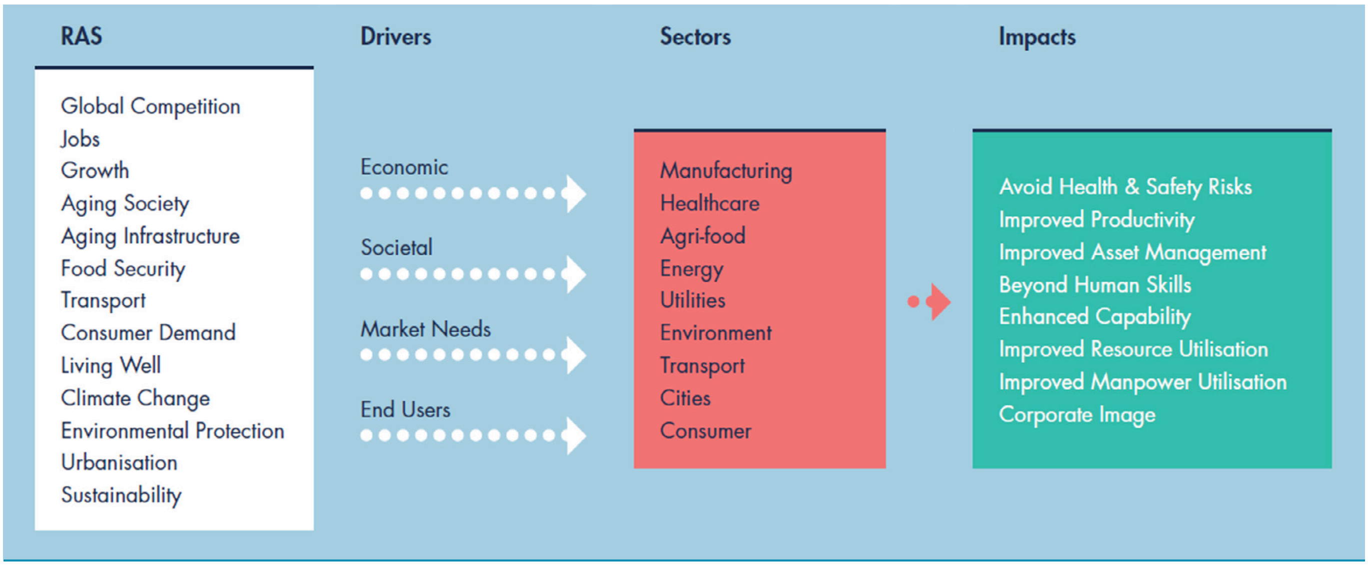


Figure 1 | Global drivers, sector applications, and potential impacts of RAS[9].

in manufacturing, utilities, agriculture, transport, logistics, energy supply and healthcare. In these sectors robotics and autonomous systems are already deployed in niche applications. There are numerous drivers for change which are common to numerous market sectors and these have significant impacts as illustrated in Figure 1.

As RAS developments continue to progress it is clear that commercial robots will no longer be largely confined to use within manufacturing and consumer applications but expand to environmental and utility applications. In the first instance this is likely to be centred on mapping and condition/performance assessment in underground infrastructure. Such application will ultimately transition to a full "find and fix" solution in the future which integrates with other city transport and utility systems.

It is reasonable to expect that most water companies in the developed world will be using the impact of AI and big data analytics in the current decade. It should however be noted that this is unlikely to be the case in many other less economically developed countries, with a more gradual trickle down of technology transfer occurring over time. RAS inspection, collection and condition monitoring will add increasing amounts of data. It is the data frequency from sensors and geo-distribution of data points that provide the granularity required to produce actionable information and knowledge. Further, the transition from Eulerian (fixed) sensors to Lagrangian (mobile) sensors opens up both the prospect of repeat sensing for condition monitoring as well as converting performance data to actionable information (such as identifying pinch points and spare capacity). The reliability and information content of low-resolution monitoring has been such that its use is typically confined to reactively demonstrating compliance to regulators and/or to calibrate idealised single snapshot deterministic network models derived from generic understanding of processes. High resolution monitoring is now sufficiently reliable that it should be integral to derivation of information from data through the building and running of site-specific, continually updated predictive models that can be used proactively to make management proactive, more cost-efficient and effective. These deployments will enable the shift to much richer detailed water network models such as digital twins based on real-time pervasive sensing. In relation to such digital twins, autonomous robots promise a step change in the data driven construction, calibration and utilisation of such models.

Vision for autonomous pipe robots

In 2019 the UK government invested in research to develop pervasive sensing for buried pipes which will be based on autonomous robotic systems[16]. The vision for the UK Engineering and Physical Sciences Research Council (EPSRC) Pipebots Grant (2019-2024) is of intelligent, robust and resilient buried pipe systems with the development of autonomous and pervasive micro-robots which are smart and (almost) failure free[17]. Such systems reduce the service disruption to society by avoiding unnecessary and unplanned road excavation. Key challenges that have been identified in the industry are Asset Mapping, Leakage, Condition Monitoring, Cost-Benefit, Blockages and No Disruption. Ideas of timescales have been developed for some of these applications (such as asset mapping) expected to be feasible at a small scale by 2025, compared to full implementation of swarms of robots by 2030. The experimental validation and demonstration proposed is taking place both in the new UK Collaboratorium for Research on Infrastructure and Cities (UKCRIC) facilities at Sheffield (Figure 2, https://icair.ac.uk/), and on carefully selected field sites with support from industry partners to guarantee the safety of the robotics technology platform.

Pipebots prototypes

Sprintbot is Pipebot's first autonomous sewerage inspection platform which is a result of a 4-week long hardware sprint exercise conducted in March 2020. Sprintbot is designed like a ball (Figure 3) allowing it to easily move and turn around inside a pipe[18]. As a first prototype, the Sprintbot is designed

Figure 2 | UKCRIC facilities at Sheffield.

to operate in relatively dry pipes, but future prototypes will be designed to operate in live sewers. The physical platform is custom-designed and largely 3D printed. The electronic package is built around a Raspberry Pi 4 as the primary controller, interfaced with an Arduino Nano as a secondary low-level controller. The sensor payload consists of a camera (Arducam MIPI), Inertial Measurement Unit (IMU, Arduino Nano 33 BLE Sense), laser range finders (STMicroelectronics VL53L1X), ultrasonic transducers (Murata MA40S4R), speaker (Pimoroni 4Ω COM1601) and microphones (Adafruit I2S MEMS Microphone Breakout). These provide data for localisation, autonomous control and blockage sensing.

Sprintbot has been tested and videoed at the Integrated Civil and Infrastructure Research Centre (ICAIR) facility at the University of Sheffield. The Sprintbot is relatively large needing a minimum pipe diameter of 300 mm to safely operate. However, this size was a function of using off the shelf electronics for the short development period. The design also highlighted the need for the platform to be stable to allow camera images to be processed. Using experience from the development of Sprintbot, a new pipe inspection robot has been designed and is being assembled for testing. The new robot is significantly smaller than the original with a maximum dimension of 60 mm (Figure 4). The team is experimenting with the use of whegs (a hybrid of wheels and legs) for motive traction. The reduced size of robots means that the internal electronics can rely much less on off the shelf modules, so custom electronics boards are being designed and built. Initially twenty of these new robots will be constructed to allow a variety of tests to be carried out, including swarm applications, involving the use of a larger 'Marsupial' robot for deploying the swarm. An iterative prototyping approach to design will continue to be employed throughout the project lifetime. Robots for deployment in pressurised water supply pipes face a different set of challenges and design of these will commence soon.

Software integration and AI control

A full software architecture has been produced following the development of the Sprintbot, by holding a further sprint. This three week event was run using agile methodologies, specifically Scrum (https://www.scrum.org/), and brought the Pipebots themes[16] together daily to ensure smooth communication to establish intermodule dependencies. The aim was to agree a flexible software architecture that would allow reconfiguration of a Pipebot across different variants with different sensing capabilities. This software architecture (see Figure 5) is now available to the team on the Pipebots Github repository (https://github.com/pipebots), providing a software skeleton for any Pipebot and allowing each theme to populate black-box elements with code developed through their research. This will enable the rapid development of future use-cases, as interoperability between modules has been captured within the model. This software approach helps in the design of control algorithms that can easily adapt to the robotic and sensor designs that have yet to be created. An example of such control is that required for self-assembly Pipebots, in which robots would link together and cooperate to perform certain tasks such as collaboratively moving against a strong flow, ascending steep pipes or climbing steps and obstacles. Ideas for algorithms have been proposed and simulation models are currently being built for evaluation.

Autonomous navigation

Effective interventions in buried pipes rely on accurate knowledge of the pipe network itself and on the location of the robot sensors

Figure 3 | Sprintbot first prototype testing in mock pipe network.

Figure 4 | Swarm Pipebot prototype concept.

within the network. The former is for the operators to have accurate and up to date information on their assets. The latter is to ensure that the robot swarm efficiently inspects the whole network in a timely fashion and to pinpoint faults for interventions. It is also critical to the success of the distributed (swarm) robotic sensing that the sensor node positions are known accurately. Novel algorithms are being developed by researchers for simultaneous localisation and mapping (SLAM) and subsequent navigation in feature-sparse pipe networks[19]. The first major challenge is to generate accurate 3D maps of pipe networks using 1D and pseudo-2D movements in feature sparse environments in the pipes below the ground. The second challenge is to incorporate prior knowledge (e.g. from geographical information systems) to enhance SLAM initialisation and performance. The third major challenge is to combine SLAM information from swarm robots to produce a real-time fused pipe network map. Researchers have been determining the feasibility of tuning the parameters of visual odometry methods to recover the camera position along the pipe without the use of a tether[19]. Simulations with the water distribution network model EPANET were used to show that a swarm of autonomous robots could operate without a centralized controller and benefit from having some degree of in-pipe communication[20]. Results indicate that 10-20 robots with simple 'ant colony' style intelligence could be used to autonomously inspect an (approximately) 30 km water distribution network with a regularity of at least one inspection/month.

Applications

Real world applications, and integration, of the various technologies are being investigated by means of a number of case studies. Four example challenges are now provided.

Asset mapping. Asset databases for buried pipe networks are regularly incomplete and uncertain, both in terms of network coverage and specific details, such as material and diameter. This is due to the age of the network, changes in ownership over time, changes in database technology (e.g. from paper records to computer), and repairs and replacement of the original pipes. Overcoming such challenges relies on the development of pipe network SLAM described above. Research using pose-graph optimization for localisation of a robot in an underground water pipe has been demonstrated[19]. As an alternative to visual localization methods, four methods of incorporating information from the measurement of an acoustic spatial field were developed and designed to be applicable to any spatially varying property along the robot's trajectory, such as magnetic or electric fields. Experimental results in[19] showed that the use of acoustic information in pose-graph optimization reduces errors by 39% compared to the use of typical pose-graph optimization using landmark features only.

Blockage. Blockage of sewers, specifically small diameter laterals and pipes downstream of combined sewer overflows can result in flooding and spills from the network. Blockages of smaller pipes can accumulate rapidly hence the robot swarms would likely need to be based in a local area in order to visit the small pipes close to properties that are most prone to blockage. Existing and emerging technologies monitor water levels in Combined Sewer Overflows (CSOs) and are widely used in the UK to monitor spill durations. Robots could react to investigate alerts from automated analysis of the water level data.

Condition monitoring. The condition of buried pipes is very difficult to assess, but understanding the condition accurately is important to maintain or improve service performance and extend the life of assets at an affordable cost. This is a major challenge due to the many potential failure modes between water distribution and drainage, different pipe materials, different ground conditions, etc. Condition monitoring robots could carry

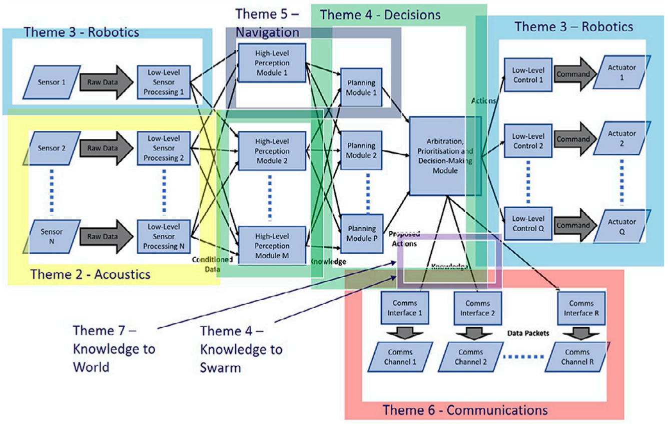


Figure 5 | Pipebots software integration diagram.

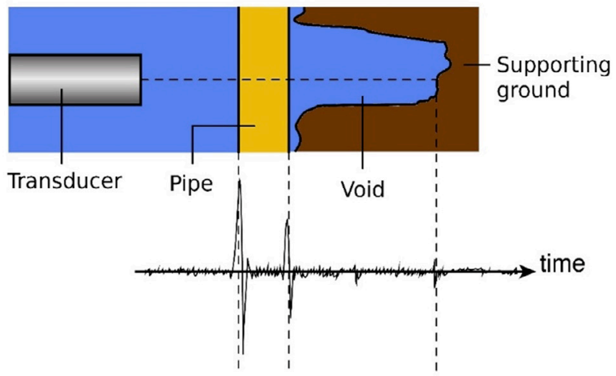


Figure 6 | Ultrasonic detection of voids in water pipes[30].

out repeat surveys (with frequency based on the known condition of individual pipes) in order that changes in condition can be recorded and to inform accurately the predictive condition modelling. Robots do not need to communicate condition back to the cloud regularly, the frequency is likely to be determined by the available data storage, or the frequency with which it passes a hub that allows communication back to the water utility. Work has been conducted on the ultrasonic detection of voids (and water content in soil) as an early indicator of the onset of failure in plastic water pipes[21]. The ultrasound technique is shown to be capable of detecting water filled voids and assessing the soil support, both of which are critical early indicators of failure (Figure 6). Such solutions are ideal to be deployed on Pipebots working inside pipes. Work has explored using acoustic sensing for blockage detection in sewer pipes to characterise the blockage shapes and sizes[22].

Leakage. The leakage from piped water distribution networks is a key (and enduring) challenge. While water utilities are able to locate larger leaks the process is time consuming. Locating smaller leaks, especially in plastic pipes, remains challenging. Leakage detection robots would continuously trawl the network, listening and searching for new (or changed) leaks and intrusion. A main advantage of autonomous robotic technology is that it will be possible to deploy leak sensors sufficiently close to the position of each leak to pinpoint it much more accurately and over a shorter period of time than is currently possible with Eulerian based leak detectors. Initially these robots could be deployed for a short period at a local level targeting areas with high leakage. A challenge for Pipebots is to detect and locate smaller leaks with reasonable precision (e.g. within centimetres) and to do this in a timely manner.

Conclusions

Water distribution and wastewater pipe infrastructures are ageing, resulting in regular failures requiring costly, disruptive, reactive maintenance. Mobile robots could be used for autonomous, persistent monitoring of a buried pipe network, locating faults and reporting information enabling proactive rather than reactive interventions. This article assesses the state-of-the-art in the development of autonomous robots for monitoring of buried pipe networks and describes technologies being developed by Pipebots[16]. This collaboration between four UK universities and industry aims to revolutionise buried pipe inspection with the development of autonomous micro-robots designed to work in complex pipe networks (clean and waste water) to generate potentially massive amounts of real-time data. Swarms of miniaturised autonomous robots equipped with novel sensors will be deployed in buried pipe networks. New algorithms uniquely incorporating Lagrangian (mobile) rather than traditional Eulerian (fixed) based coordinate systems are being developed to process the autonomously collected data to inform condition assessment and system performance. The outputs from these algorithms could be directly mapped to the hydrodynamic performance of single pipes or fed into pipe deterioration models that can predict, with AI and machine learning support, the remaining service life of a pipe. Robotic autonomous systems will enable maximising the capacity of existing infrastructure, detecting deterioration proactively, increasing safety and reducing downtime of city infrastructure and generating data to drive better maintenance and investment models. However, significant problems must be solved to result in pipes becoming a RAS highway. Technologies are required for effective robot deployment and recovery, navigating in often unmapped and uncertain pipe environments and communicating in order to contextualise the condition and allow rehabilitation work to be planned. With sufficient technological and governance progress utilities would be empowered to run a fully automated inspection, repair and maintenance system (using find and fix swarming robots). This would radically reduce the risk of service failure and along with new repair technologies significantly reduce the cost of individual repairs – releasing funds to rehabilitate assets over the longer term. The Pipebots team intends to have a full Pipebots system demonstrated in a realistic (initially sewerage) network before 2024. Once that is successful, a thorough certification and compliance process will be required to ensure that pervasive autonomous Pipebots will be safe to adopt and deploy in live water and sewer networks for mapping, sensing and communicating.

Stephen R. Mounce
Stephen R. Mounce is a visiting research fellow in Hydroinformatics at the University of Sheffield and the director of Mounce HydroSmart Ltd. His PhD is in Computer Science and he has over twenty years of experience on research projects with over one hundred academic publications. His research combines Artificial Intelligence and Water Engineering for such applications as leakage (including smart meter data mining), CSO analytics, water quality and burst event detection systems and fuzzy RTC.

Will Shepherd
Will Shepherd is a research associate in the Department of Civil and Structural Engineering at the University of Sheffield in the UK. He has over twenty years research experience. His main interests lie in the performance of urban drainage systems, including asset performance modelling; hydraulic modelling; laboratory and field testing; and application of artificial intelligence in urban drainage.

Kirill Horoshenkov
Kirill Horoshenkov is professor of Acoustics at the University of Sheffield. His research interests are in acoustic sensing, robotics and porous media. He leads the EPSRC UK Acoustics Network with over 1200 members (www.acoustics.ac.uk). He also leads the EPSRC Programme Grant (www.pipebots.ac.uk, Pervasive Sensing of Buried Pipes). He is Fellow of the Royal Academy of Engineering, UK Institute of Acoustics and Acoustical Society of America (ASA). He is also a coordinating editor for the Journal of the Acoustical Society of America. He is a founder of two successful spin-off companies which commercialised his research into acoustics sensing for sewer pipes (SewerBatt device) and sustainable noise absorbing materials (now marketed as ArmaSound). He is a chartered engineer.

Joby Boxall
Joby Boxall is professor of Water Infrastructure Engineering at the University of Sheffield, and head of the department of Civil and Structural Engineering. He is a chartered engineer and environmentalist and fellow of the Chartered Institution of Water and Environmental Management. Joby's research interests are concerned with understanding and modelling hydraulic, water quality and infrastructure performance and interactions, usually with a focus on water supply. He is focused on research addressing the grand challenges facing water, including leading the EPSRC grand challenge consortium on sustainable clean water for all, TWENTY65.

Jordan Boyle
Jordan Boyle is associate professor in the School of Mechanical Engineering at the University of Leeds. His primary research area is miniature mobile robotics for infrastructure inspection. His PhD research in the School of Computing at Leeds involved computational modelling of the neural control of animal locomotion, and he maintains a particular interest in biologically-inspired robotics.

Acknowledgements

Pervasive Sensing for Buried Pipes (Pipebots) is a programme grant funded by the Engineering and Physical Sciences Research Council (EP/S016813/1) and consisting of the Universities of Sheffield, Bristol, Birmingham and Leeds in addition to commercial partners.

References

1 | IKT (2017) Länge und Wiederbeschaffungswert der Unterirdischen Infrastruktur in Deutschland und in der Europäischen Union. www.ikt.de/wpcontent/uploads/2017/03/ikt-unterirdische-infrastrukturen-netzlaengen-wiederbschaffungswerte-deutschland-eu.pdf

2 | UKWIR (2006) Minimising Street Works Disruption: The Real Costs of Street Works to the Utility Industry and Society. 05/WM/12/8. https://ukwir.org/view/$J_t9zGm!/streetworks%20costs

3 | EurEau (2017). Europe's Water in Figures www.eureau.org/resources/publications/1460-eureau-data-report-2017-1/file

4 | Mounce, S. R. (2021). Chapter 1: Data science trends and opportunities for smart water utilities, ICT for smart water systems: Measurements and data science. In: *The Handbook of Environmental Chemistry.* Springer, Berlin, Heidelberg. ISBN 978-3-030-61972-5, Eds: Scozzari, Mounce, Han, Soldovieri and Solomatine.

5 | Romano, M., Boatwright, S., Mounce, S. R., Nikoloudi, E. and Kapelan, Z. (2020). AI- based event management at United Utilities. *Hydrolink,* IAHR. No. 4/2020, pp. 104-108.

6 | Cahn, A. (2014). "An overview of smart water networks". *Journal AWWA (American Water Works Association),* 106 (7), pp. 68-74.

7 | Mills, G. H., Jackson, A. E. and Richardson, R. C. (2017). Advances in the Inspection of Unpiggable Pipelines. *Robotics* 2017, 6, 36; doi: 10.3390/robotics6040036.

8 | Mirats Tur, J. M. and Garthwaite, W. (2010). Robotic devices for water main in-pipe inspection: A survey, *Journal of Field Robotics,* vol. 27, no. 4, pp. 491-508.

9 | Caffoor, I. (2019) Robotics and Autonomous Systems in the Water Industry. TWENTY65 Report, 18/2/19.

10 | Van Thienen, P., Maks, M., Yntema, D., Janssens, J. P., Bergmans, B., Diemel, R., Helgers, M., Horst, P. and Trietsch, E. (2017). Continuous robotic inspection of pipes for data rich asset management. *Proceedings of Computing and Control for the Water Industry (CCWI2017),* Sheffield, UK, 5th to 7th September 2017.

11 | Van Thienen, P., Bergmans, B., Diemel, R., Helgers, M., Holzhaus, P., Horst, P., Trietsch, E., Vogelaar, H., Summeren, J. V., de Waal, L. and Hillebrand, B. (2018). Advances in Development and Testing of a System of Autonomous Inspection Robots for Drinking Water Distribution Systems. *Proceedings of 1st International WDSA/CCWI 2018 Joint Conference,* Kingston, Ontario, Canada–July 23-25, 2018.

12 | Lepot, M., and Clemens, F. (2017). Towards a new technology for sewer inspection. New Zealand Hydrological Society Annual Conference, at Napier, New-Zealand.

13 | Seeta, G., Yeoa, S. H., Lawa, W.C., Burhana, C.Y., Wong, A., Saparib, S., Liau, K. K. (2018). Design of Tunnel Inspection Robot for Large Diameter Sewers. *Procedia Computer Science* vol. 133, pp. 984–990.

14 | Tan, C.H., Ng, M., Shaiful, D.S.B., Win, S.K.H., Ang, W.J., Yeung, S.K., Lim, H.B., Do, M.N., Foong, S., (2018). A smart unmanned aerial vehicle (UAV) based imaging system for inspection of deep hazardous tunnels. *Water Practice and Technology* 13, 991-1000.

15 | Savić, D. and Findikakis, A. (2020). Artificial intelligence in water management and hydro-environment problems. *Hydrolink,* Artificial Intelligence, IAHR. No. 4, pp. 98-99.

16 | EPSRC Pipebots grant, https://pipebots.ac.uk/

17 | Doychinov V, (2021) An army of sewer robots could keep our pipes clean, but they'll need to learn to communicate, The Conversation, 26th January 2021

18 | Sprint 2 summary, https://www.youtube.com/watch?v=OrFOY6kFfQM

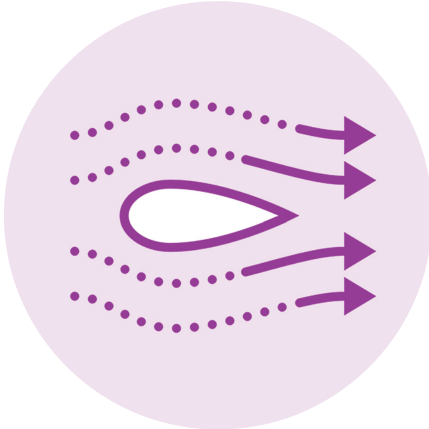
19 | Worley R, Ma K, Sailor G, Schirru M.M, Dwyer-Joyce R, Boxall J, Dodd T, Collins R, Anderson S, (2020). Robot Localization in Water Pipes Using Acoustic Signals and Pose Graph Optimization. *Sensors* 2020, 20, 5584, https://www.mdpi.com/1424-8220/20/19/5584

20 | Parrott C, Dodd T.J, Boxall J, Horoshenkov K, (2020). Simulation of the behavior of biologically-inspired swarm robots for the autonomous inspection of buried pipes. *Tunnelling and Underground Space Technology,* vol. 101, 2020, 103356, ISSN 0886-7798, https://doi.org/10.1016/j.tust.2020.103356

21 | Zhu J, Boxall JB, Hills AF, Dwyer-Joyce RS, Anderson SR and Collins RP, (2021) Assessing Ground Support of Plastic Pipes Using Ultrasound. Infrastructures 2021, 6 (30). https://doi.org/10.3390/infrastructures

22 | Yu, Y., Krynkin, A., Li, Z., Horoshenkov, K.V., (2021) Analytical and empirical models for the acoustic dispersion relations in partially filled water pipes. *Applied Acoustics* 179, 108076. https://doi.org/10.1016/j.apacoust.2021.108076

FAMOUS WOMEN IN FLUID MECHANICS



Irmgard Flügge-Lotz
1903, Germany – 1974, USA

Irmgard Flügge-Lotz entered Technical University of Hannover in 1923 and graduated in engineering thereby specializing in aeronautics. She received the PhD title in 1929 and derived two years later the Lotz-Method for calculating the distribution of pressure in aircraft wings of different shapes. Later, she took interest in automatic flight control of aircraft, notably of the discontinuous or the 'on-off' type. By 1928, Flügge-Lotz headed the Department of Theoretical Aerodynamics, of Aerodynamische Versuchsanstalt AVA, at Göttingen University. She was also a consultant for Deutsche Versuchsanstalt für Luftfahrt in Berlin-Adlershorst from 1938-1945. However, she and her husband Flügge felt themselves increasingly discriminated by the Nazi regime. In 1948 they emigrated to the USA where Flügge was offered a professorship in engineering, whereas his wife got the first woman full professorship in the Engineering Department at Stanford University. She extended her work to automatic flight control and to the guidance of rockets and missiles, earning herself the description 'a female Wernher von Braun'.

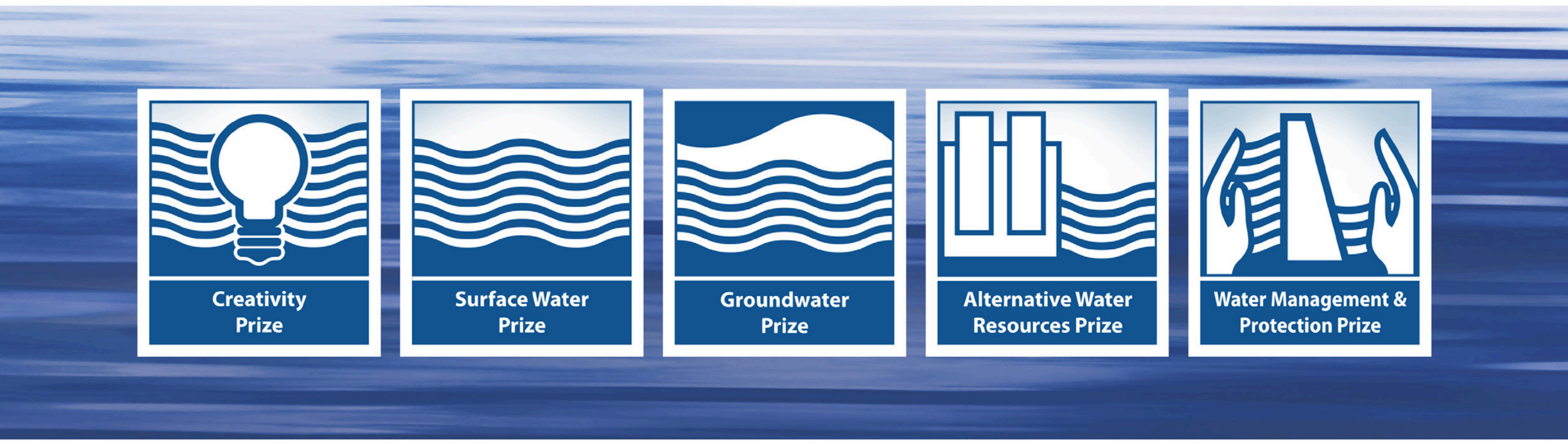
Flügge-Lotz was awarded the Women Engineers Achievements Award in 1970, and she was a Fellow of the Institution of Aero-nautics and Astronautics. She delivered the von Karman Lecture on Trends in the field of automatic control in the last two decades in 1972 during the AIAA Annual Meeting in Washington DC. A Wilhelm Flügge and Irmgard Flügge-Lotz Memorial Award was installed by the Applied Mechanical Division of Stanford University.

Anonymous (1964). Irmgard Flügge-Lotz. Who's who in engineering: *A biographical dictionary of contemporaries* 9: 594. Lewis: New York.
Anonymous (1972). von Karman Lecture by Flügge-Lotz well received. *Astronautics and Aeronautics* 10(6): 74. P
Anonymous (1974). Irmgard Flügge-Lotz. Astronautics and Aeronautics 12(7/8): 77.
Anonymous (1996). Flügge-Lotz, Irmgard. *Biographical dictionary of the history of technology:* 263, L. Day, I. McNeil, eds. Routledge: London.
Flügge-Lotz, I. (1968). *Discontinuous and optimal control.* McGraw Hill: New York.
Lotz, I. (1936). Beeinflussung der Auftriebsverteilung durch den Schraubenstrahl. *Jahrbuch Lilienthal-Gesellschaft für Luftfahrtforschung:* 57-74.
Poggendorff, J.C. (2002). Flügge-Lotz, Irmgard Wilhelmine Dorette. *Biographisch-Literarisches Handwörterbuch* 7b: 1431-1432; 8: 1225-1227. Akademie-Verlag. Berlin.

Biography extracted from the book *Hydraulicians in Europe* by Willi H. Hager. Volume 2. IAHR, 2009, p. 1059. ISBN: 978-90-78046-06-6

Prince Sultan Bin Abdulaziz International Prize for Water

Recognizing Innovation



Invitation for Nominations

10th Award (2022)

Nominations open online until 31 December 2021

www.psipw.org
e-mail: info@psipw.org