water

MDPI

Article

Computational Tools for Supporting the Operation and Management of Water Distribution Systems towards Digital Transformation

Nelson Carriço [1,*], Bruno Ferreira [1], André Antunes [2], João Caetano [1] and Dídia Covas [3]

[1] INCITE, Escola Superior de Tecnologia do Barreiro, Instituto Politécnico de Setúbal, 2839-001 Lavradio, Portugal
[2] Sustain.RD, Escola Superior de Tecnologia de Setúbal, Instituto Politécnico de Setúbal, 2914-508 Setúbal, Portugal
[3] CERIS, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal
* Correspondence: nelson.carrico@estbarreiro.ips.pt; Tel.: +351-212-064-660

Abstract: This paper presents a set of computational tools specially developed for supporting the operation and management of water distribution systems towards digital transformation of water services. These tools were developed in the scope of two R&D projects carried out in Portugal, DECIdE and WISDom, during 2018–2022. The DECIdE project focused on the development of tools for importing cadastral and operational data, as well as on the three operational tools for supporting the performance assessment: the first allows the calculation of different key performance indicators, both at a global and sectorial level, which is an annual requirement of the water regulator, and the other two allow the calculation of the water and the energy balances and a set of complementary indices. The WISDom project aimed at the implementation of applications that directly address specific water utility needs, namely, the flow rate data processing, the optimal location of pressure sensors, the identification of critical areas in the distribution network for pipe burst location, and the prioritization of pipes for rehabilitation. Implemented tools are useful to support water utilities in the daily operation and management of their systems, being a step forward towards digital transformation of the water sector.

Keywords: data processing; digitalization; digital twins; energy efficiency; leakage operation; management; water distribution networks

Citation: Carriço, N.; Ferreira, B.; Antunes, A.; Caetano, J.; Covas, D. Computational Tools for Supporting the Operation and Management of Water Distribution Systems towards Digital Transformation. *Water* **2023**, *15*, 553. https://doi.org/10.3390/w15030553

Academic Editor: Enrico Creaco

Received: 30 December 2022
Revised: 21 January 2023
Accepted: 29 January 2023
Published: 31 January 2023

1. Introduction

Acknowledging that the core business of water utilities is to provide drinking water and wastewater services to residential, commercial, and industrial sectors of the economy, information is the key element for operating and managing these services [1,2]. Water services bring together a wide variety of collected data (e.g., flow rate, pressure, energy consumption, end-user consumption) [3–5]. Depending on the water utility maturity level, different data can be collected with different acquisition frequencies, manually in situ by an operator (e.g., domestic consumption) or automatically by smart meters or by sensors installed in the systems (e.g., flow rate, pressure) [6]. Independently of the type of metering used, data are periodically collected and should provide good knowledge about water distribution systems (WDS). The fact is that big and disorganized data are generated daily and stored in dispersed information systems (IS), which turns the operation and management of WDS into a challenging and backbreaking task [7]. Many IS used by the utilities were reactively acquired to answer specific needs. Therefore, the efficiency of data management processes is often poor, especially in small- and medium-sized utilities, such as municipalities whose management is structured according to local government model conventions, where insufficient planning leads to a lack of integration that involves the information flows and their use [4].

An efficient and effective operation and management of WDS requires that the gathered and stored data can be combined into a single platform to help the utility managers and operators to address their operational and infrastructural problems, namely, related to assets' ageing, non-revenue water, water stress and scarcity, and variability of hydrologic events from climate change, among others [8–11]. This process is typically called data integration. Additionally, the use of smart Internet of Things (IoT) devices to collect time series of different parameters (e.g., water quality, hydrometeorological data, flood data, water consumption) significantly increases the amount of data that is collected, processed, and stored [12,13]. Furthermore, data typically have several types of anomalies that ought to be automatically processed to avoid providing erroneous information and leading to inadequate decisions [12]. Therefore, these big data require the use of artificial intelligence (AI) techniques, such as data mining and machine learning, to extract relevant information and present them in a more user-friendly format that can assist utility experts in the decision-making processes. As a result, new tools and processes are needed to collect, gather, manage, analyze, and extract knowledge from existing data [5,14,15] and to assist decision-making processes [16].

For these tools and processes to effectively support the operation and management of WDS, it is also of the utmost importance that utility managers are open-minded and assume a proactive attitude in the use of these tools, otherwise, no knowledge or added value can be generated. For instance, flow rate or pressure data in specific locations are used solely for triggering automatic alarms of abnormal behaviors, whereas such data could also be used to locate and size the source of such abnormal events with the help of more advanced tools.

Water utilities need to climb a ladder with several steps to become a fully smart utility, each being step a small advance towards the digital transformation. Full water digitalization is envisioned as having a reliable virtual copy of the real system, capable of describing its behavior based on real-time monitoring and serving as a basis for experimentation and simulations [16–18], also called a Digital Twin of the WDS (DT-WDS). For example, the first step consists of just having the physical assets managed without any digital tool, the second step of having control and measuring devices and sensors to collect data from the physical systems, and the third step may be having communication technologies connected to the sensors and automation to control the system, and so on. This means that water utilities are at quite different digital maturity levels, notwithstanding most having to address similar daily problems and sharing the same challenges.

Despite this being one of the main research areas of the most mature utilities and much research having already been developed in this domain, the implementation of DT-WDS still faces major challenges related to the interoperability and integration of different IS, the definition of a common ontology between utilities, and the management of different digitization maturity levels [3,14,16].

Some of these challenges have been addressed by the authors in two projects, DECIdE and WISDom, funded by the Portuguese National Funding Agency for Science and Technology (FCT), carried out between 2018 and 2022. The first project, DECIdE—Multi-Criteria DECision Support Platform for Urban Water InfrastructurEs (accessible at https://decide.ips.pt/, accessed on 20 December 2022)—developed a platform for integrating data from different IS and calculating water and energy balances to support the annual reporting of key performance indicators (KPI) required by the Portuguese Water and Wastewater Authority (ERSAR). The second project, WISDom—Water Intelligence System Data (accessible at https://wisdom.ips.pt/, accessed on 22 December 2022)—developed several artificial intelligence techniques to extract knowledge from the data and to address several WDS problems, such as data processing and analysis [12], pressure sensor location [19], pipe burst detection [20] and location [21], pipe condition assessment [22], and water age performance assessment [23].

This paper aims to present a set of computational tools developed in the scope of referred projects that are made available for the scientific community or utilities in the projects' websites, upon request. These tools were implemented as prototypes to test developed techniques and were specifically tailored for the five participating utilities that represent the reality of most small- and medium-sized European water utilities. The five utilities encompass urban and rural territorial topologies, with populations ranging from 3000 to 80,000 inhabitants and covering areas from 7 to 1232 km^2. More detailed information about the participating water utilities can be found in the work of Carriço et al. [4].

The computational tools developed under the scope of the DECIdE and WISDom R&D projects are listed in Table 1, and these tools will be further detailed in the following sections.

Table 1. Computational tools developed under the scope of the DECIdE and WISDom R&D projects.

R&D Project	Computational Tool
DECIdE	Data integration and analytics platform
WISDom	Flow rate time series processing
	Optimal number and location of pressure sensors
	Identification of critical areas for pipe burst location
	Prioritization of rehabilitation units for interventions

2. Data Integration and Analytics Platform

Data integration depends on the technology used to collect and store data in the different IS. For example, data can be linked to a third-party application using interoperability web services (this technology may be called middleware data integration) or by an Extract, Transform, Load (ETL) process, which extracts data from a source system and loads the data after transformation into a target destination.

A computational platform was developed for the integration and analysis of data from the different IS used in five small- and medium-sized water utilities, with different digital maturity levels and information management processes, representative of the national and European water utilities. The platform allows importing data exported from geographic information systems (GIS), in the shapefile format, and measurements in spreadsheet format or text files. A dashboard was specifically developed to this end, in which the user may import the files and map the file content attributes with the data model elements. Once infrastructural data and measurements have been imported, the platform allows the performance assessment using one of three modules, namely, calculation of a set of performance indicators [24] (selected by the participating water utilities and necessary for the annual reporting to the ERSAR) and calculation of the water and energy balances. These balances allow the assessment of the water efficiency [25], and the energy efficiency associated with the water transport and distribution in the WDS [26]. The assessments can be carried out at both the system and sectoral levels (i.e., sector, subsystem, or district metering area). The three assessment modules follow a similar basic functional procedure:

- The user selects a given area of analysis (e.g., the complete system or a particular subsystem) and defines the period of analysis (e.g., 12 months).
- The platform prepares the required data for analysis by querying the database and performs necessary calculations for performance assessment (e.g., performance indicators, water and energy balance). For water and energy balances, the user has the possibility to validate and change the input values before the final calculation.
- The results are displayed in a user-friendly dashboard.

The input data for the water balance follow the IWA recommendations [25] and include the total system input volume of water, billed metered and unmetered consumption that is automatically fetched and summed by the platform, as well as additional values that should be introduced by the user (for instance, metering error percentage or estimation of non-authorized consumption). The same input data are used for the calculation of the energy balance, in addition to energy consumption and the hydraulic head in the pumping station, storage tank elevation and level, and consumers' elevation data.

The dashboards include different types of graphical representations (e.g., bar graph, line, circular) and user-customizable filters (i.e., by analysis area, dataset and indicators, and time) for each of the performance indicators and balance analysis.

Figure 1 shows a screenshot of the window of the data integration module in the DECIdE platform, specifically the measurement importation module, whilst Figure 2 shows a screenshot of a result window of the water balance. More detailed information about this platform can be found in Carriço et al.'s work [7]. Note that this platform is developed in Portuguese, thus translated versions are presented herein.

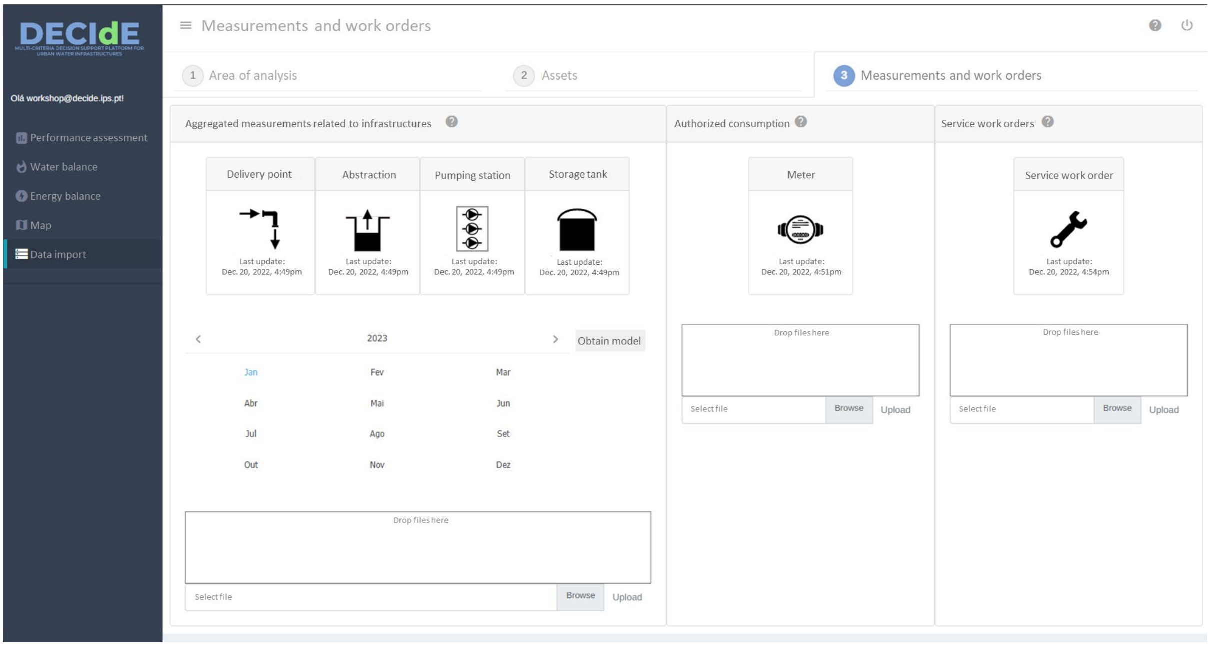

Figure 1. Screenshot of the window of data integration in the DECIdE platform.

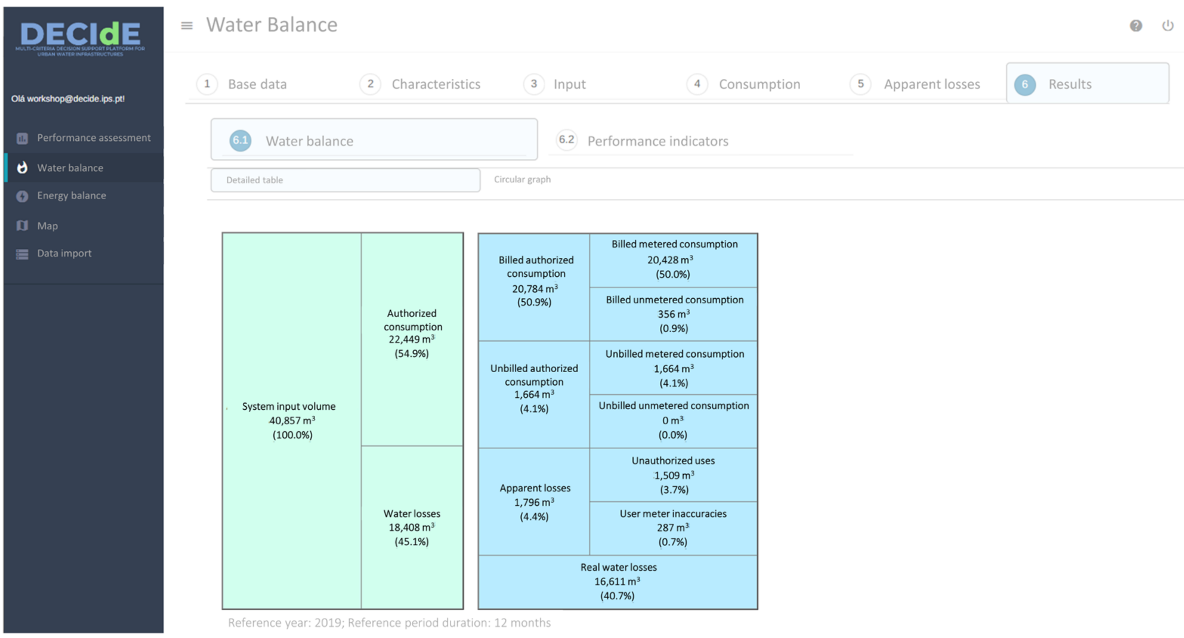

Figure 2. Screenshot of the window of the water balance in the DECIdE platform.

The developed solution facilitates the data integration and the KPI calculation required of end-users (water utilities) to support the decision-making process and to annually report to the Water Authority. This platform may also be useful for any regional or national utilities needing data integration for combined analysis, such as sanitation services or irrigation projects.

3. Flow Rate Time Series Processing

The use of telemetry creates unprecedented opportunities for WDS operation and management, water demand management, leak detection and localization, and the recognition of anomalous events. The collected time series by the installed pressure sensors and flow meters result in big data with detailed and useful information on the system's behavior and condition. However, these time series frequently present errors, namely missing, duplicate, or anomalous values. Processing the collected (raw) flow rate time series, before any use by engineering tools, is essential to ensure reliable data that do not compromise the success of the subsequent applications. Data processing usually requires identifying and removing outliers, normalizing the time step, and filling data gaps with reliably estimated values. Often, this time series processing is manually carried out, ultimately limiting the amount of data that can be simultaneously treated [27].

The implemented data processing tool is based on the methodology proposed by Ferreira et al. [6] that allows the automatic processing of flow rate time series with different characteristics (e.g., consumption pattern, data acquisition system, transmission settings), both for normal operating conditions and during the occurrence of abnormal events (e.g., pipe bursts). The methodology consists of a four-step procedure: (i) anomaly identification and removal, (ii) short-duration gap reconstruction, (iii) time step normalization, and (iv) long-duration gap reconstruction. In the first step, specifically developed tests are used for the automatic identification of the most common anomalies in flow rate time series due to acquisition and transmission problems (e.g., abnormally high and low values). Short-duration gaps are directly reconstructed in the second step using simple reconstruction techniques (e.g., linear interpolation). The time step normalization is carried out in the third step by a numerical procedure before the fourth step of the reconstruction process of long-duration gaps. This reconstruction uses a pattern model coupled with regression techniques (i.e., autoregressive integrated moving average, exponential smoothing).

The flow rate time series processing methodology was implemented in a computational application both as a desktop for Windows and web-based versions. The main functionality is the possibility to import a raw flow rate time series (in CSV or TXT format) to be processed using the developed methodology. Once the time series have been imported, a set of parameters necessary for the method is automatically suggested, by default, to the user, who has the possibility to accept or to change the value of each parameter. These parameters include the time step after normalization, and parameters related to thresholds for anomalous values' identification (see parameter box on the right side of Figure 3). More details regarding the parameters of the processing methodology can be found in Ferreira et al.'s work [6]. Already processed flow rate time series of the same sensor can also be imported to be used in the reconstruction step.

Figure 3 presents the screenshot of the window for the tool for flow rate time series processing, specifically the web-based application available to use at: https://wisdom.ips.pt (accessed on 22 December 2022). The desktop version can be found in the GitHub repository at: https://github.com/Ferreira-B/Flowrate-time-series-processing (accessed on 23 December 2022).

The implemented data processing tool is very useful for preparing the time series to be used in different engineering applications, namely, hydraulic simulation, model calibration, or online burst detection, both in the daily operation of the water services and in the planning and management of future works and activities.

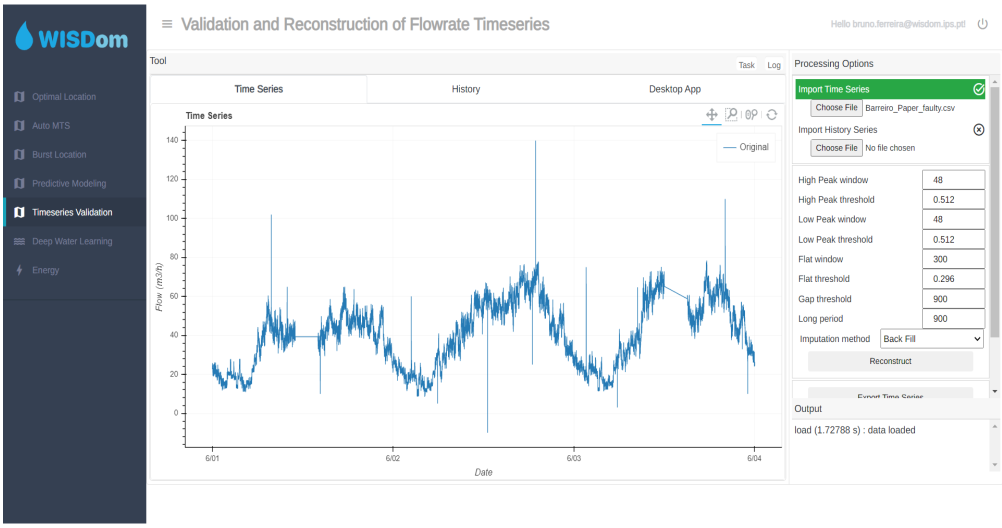

Figure 3. Screenshot of tool for the flow rate time series processing.

4. Optimal Number and Location of Pressure Sensors

The implemented tool, based on the methodology proposed by Ferreira et al. [19,28], allows the determination of the optimal number and the best location of pressure sensors for the simultaneous calibration of the hydraulic network model and the detection of pipe bursts. The methodology was implemented in a computational tool that uses the hydraulic simulation model and solves the optimization problem for different numbers of pressure sensors, evaluates the technical benefit of installing each set of pressure sensors using the hypervolume indicator, and determines the optimal number of sensors by the variation of the hypervolume. A set of optimal solutions is returned. The decision on the optimal solution results from the compromise of those two objectives and the maximization of the benefit of installing these costly devices.

Figure 4 shows a screenshot of the tool for the optimal number and location of the pressure sensors, available at: https://wisdom.ips.pt (accessed on 22 December 2022).

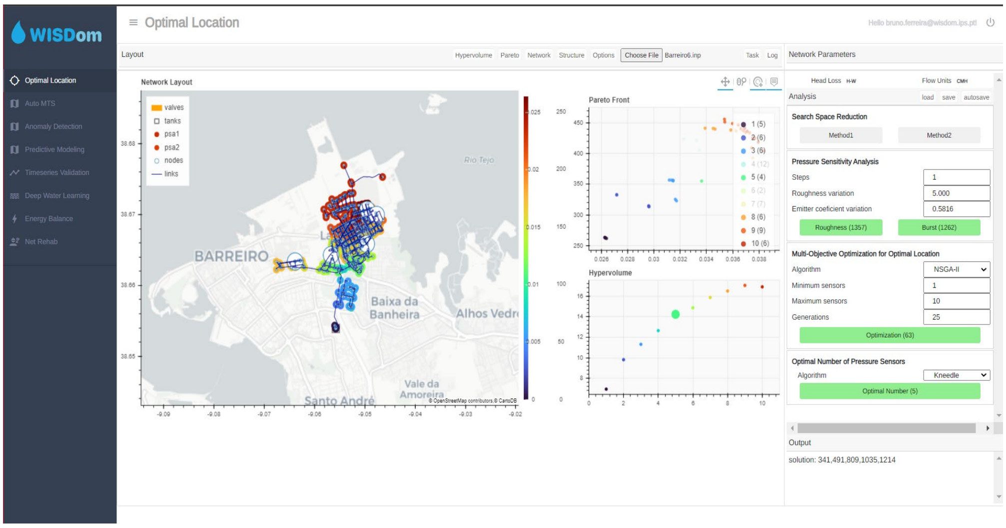

Figure 4. Screenshot of tool for the optimal number and location of the pressure sensors.

The developed tool directly helps water utilities to address a real-life problem that is often too complex to be solely solved by engineering judgment. The use of such a tool allows utilities to explore a higher number of solutions and, ultimately, increase the benefit of installing pressure sensors.

5. Identification of Critical Areas for Pipe Burst Location

The detection and location of pipe bursts in the shortest time possible is essential to reduce the water volume lost and can be achieved by using numerical methods and network-monitoring data. The ability to correctly locate the burst is directly related to the number and location of pressure sensors. As a result, bursts are more difficult to locate in certain zones of the network due to the topology of the system or the lack of nearby sensors. These zones, which are not sufficiently covered by pressure sensors, and in which burst location cannot be accurately identified, are herein referred to as critical areas.

The identification of such critical areas can be carried out by simulating a pipe burst event (i.e., using the hydraulic simulation model) in distinct network locations and by attempting to locate such simulated bursts in the network using a given advanced algorithm, for instance, inverse analysis, sensitivity-based methods, or classifier-based methods [29]. If the simulated event is correctly located, it means that a real event occurring in such a location can also be potentially located. On the other hand, if the simulated event is not correctly located, it means that a real event occurring in such a location cannot be located by using that algorithm with the available pressure sensor data. Repeating this procedure for every possible location of a burst in the network allows for identifying critical areas of the WDN in which pipe burst events would not be located. The aforementioned process could also be repeated by considering that the burst starts at distinct periods of the day, since the demand patterns may considerably vary, and which may affect the performance of the burst location method.

A computational tool was developed to carry out the aforementioned process. This tool uses a hydraulic simulation model of the network and the existing pressure sensors' location. Currently, the tool only allows the simulation of one pipe burst event at a given location and with a given size. The pressure measurements generated by such a simulated event are then used by the advanced algorithm to locate such an event. The implemented advanced algorithm is based on a pattern recognition classifier (further details are presented in [29]), but distinct methods could be used (e.g., inverse analysis). The assessment of the correct burst location is given by the distance between the identified location and the correct location, measured along the network pipes (the shortest distance) or given by the Euclidean distance. When this distance is lower than a given threshold, the burst is considered correctly located. Critical areas can be identified by systematically simulating bursts in all potential locations. The implementation of this systematic procedure is currently under development.

Such areas can be subjected to further hydraulic studies, namely, to identify the location of pressure sensors or network sectorization to improve pressure management. The tool, shown in Figure 5, is available at: https://wisdom.ips.pt (accessed on 22 December 2022).

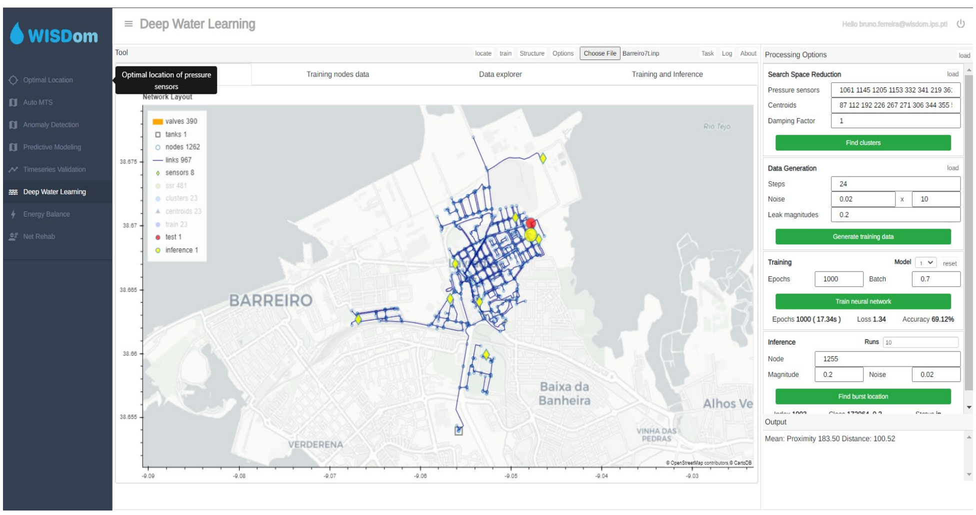

Figure 5. Screenshot of tool for the identification of the critical areas for pipe burst location.

6. Prioritization of Rehabilitation Units for Interventions

The establishment of effective medium- and long-term rehabilitation strategies is essential to counter the continuous deterioration process of WDS as well as to ensure the infrastructure sustainability, the continuity of the supply, and the improvement of the quality of the provided service. The existing decision support methodologies in the literature commonly become impractical as they use single pipes as basic rehabilitation units. Water utilities plan and carry out rehabilitation activities in sets of contiguous pipes to meet the pre-determined available budget. Even if single pipes are identified as priorities, they generally do not cover the annual available budget and it is unrealistic to set up a work front to rehabilitate small sections of pipes scattered throughout the network. On the other hand, in the case of a burst in a single pipe, the entire segment (i.e., pipes connected between isolation valves) to which the pipe belongs will be taken out of service for repair or replacement. In this sense, it is appropriate to rehabilitate the segment, restoring the structural condition of all pipes.

A novel approach to support pipe rehabilitation has been developed to group network segments instead of grouping individual pipes and scheduling rehabilitation in a medium- and long-term perspective. This approach is divided into two main phases: the first is the definition of pipe grouping according to an available budget and contiguity (referred to as rehabilitation units), and the second is the interventions' scheduling overtime according to predefined criteria, defined by Caetano et al. [30].

The methodology generically described above was implemented in a computational tool. This tool is useful for engineers and managers, aiding in the process of defining rehabilitation strategies, and being fully adaptable to the specifications of each utility (e.g., financial capabilities, O&M practices). Figure 6 shows the graphical interface of the tool which is available at: https://wisdom.ips.pt (accessed on 22 December 2022).

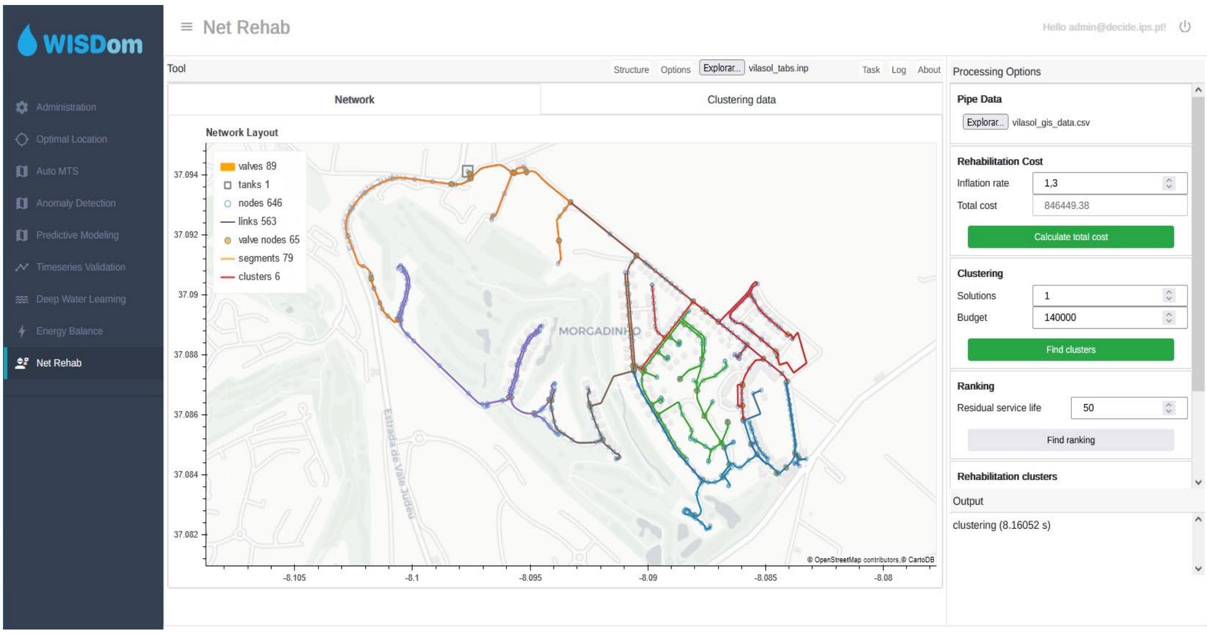

Figure 6. Screenshot of tool for the prioritization of rehabilitation units for interventions.

7. Discussion

7.1. Practicality of the Developed Tools

The presented computational tools correspond to a user-friendly implementation of advanced techniques and methodologies, developed by academia together with industry, in the scope of two R&D projects. In this regard, the developed tools are prototypes for testing the developed techniques, specifically tailored for the five participating utilities that represent the European reality in terms of small- and medium-sized services. By no means are these tools aimed to be commercial solutions. In this context, there are some aspects to be improved in the future concerning the presented tools, namely:

- The developed tools are usually very specific and require training to be used by utilities. This training can only be provided by the R&D team.
- As the project funding has ended, part of the R&D team is unavoidably dispersed (in particular, researchers on computer science and mathematicians), and thus, it is not easy to continuously carry out updates on these tools. However, the core R&D team, mainly composed of the authors of this paper, will pursue further developments in the scope of their doctoral projects and will continue to make efforts to raise more funding.
- Software and hardware limitations may affect the tools made available as, for instance, the university servers may not be capable of handling the expected traffic of free-to-use tools (especially when the provided service involves complex optimization or AI analyses).
- Problems with the standardization of input files may occur since the tools were developed to meet the requirements of specific utilities. In fact, certain details, data formats, or data specificities not considered in the five participating utilities may exist in other water utilities.

Nonetheless, several advantages can be pointed out from the development of these tools. First, there is the direct benefit not only to the five participating utilities that have tools specifically tailored to their problems but also to many other utilities that share similar engineering realities and scarce human, technological, and financial resources. Secondly, the smart tools developed under the national R&D projects' funding are a starting point for software companies and digitally advanced water utilities to further enhance that software, addressing their concerns and challenges. In this regard, the collaboration between the industry and academia is essential to define the priorities for the industry and the state-of-the-art solutions by academia.

As referred, these tools are now freely available for the direct use of water utilities (e.g., in a computer application with a graphical user interface), and a special training course on Advanced Analysis of Water Distribution Systems will be offered in 2023, by the core R&D team of this research, for capacitating the water utilities and interested consultancy engineer staff to use these tools in the daily operation and management of their services or for making investment plans.

7.2. Envisioned Journey towards Digital Water Transformation

Nowadays, the operation and management of WDS have become a challenging task due to a huge amount of buried assets having been built decades ago and having bad operating conditions nowadays, the existing limitations of available budgets do not allow attending to all rehabilitation and maintenance needs, and a huge amount of data that are continuously generated, aiming to provide useful insights on system daily operation. Utilities currently struggle with several monitoring and information systems (IS) and with multiple computational tools (such as those presented herein) attempting to effectively and efficiently operate and manage their systems.

In the authors' opinion, the DT-WDS will be the future for the operation, management, and planning of water utilities. Nevertheless, several challenges need to be overcome, such as data assimilation from different data sources, data quality and uncertainty, the need to forecast systems' malfunctions based on accurate and real-time data, and computation time for some simulation scenarios. These challenges need further research from academia to help the industry to advance in knowledge and to improve the service provided to the users and save the scarce natural resource that is water, in particular, in countries where climate change leads to more frequent droughts and water scarcity.

To transform data into useful insights, the water utilities need to verify the quality of their collected data, aiming to avoid misinformation and always bearing in mind that "garbage in, garbage out". Ultimately, effective digitalization requires the cultivation of good metadata management practices amongst water utility personnel [14].

Hydraulic models are often erroneously considered "simplified" versions of Digital Twins of WDS; nonetheless, currently, these are not fed with real-time data nor connected to existing IS, usually being used for running offline analyses [18,31]. Ideally, a hydraulic model should be automatically updated with processed real-time data and continuously calibrated to accurately describe the WDS behavior, being essential to efficiently anticipate, detect, and manage anomalous behaviors and to assist in real-time operation. Although distinct works can be found related to real-time hydraulic simulation [13,32–34], the main challenges still lie in the development of tools capable of discarding unrealistic calibration parameter values that may be obtained from measurements [32] and which will lead to misleading results. To the authors' knowledge, no hydraulic model was used nor calibrated using real-time data, being one of the main challenges to developing tools capable of discarding unrealistic calibration parameter values that may be obtained from measurements [32] and that will lead to misleading results.

Most existing tools process stored offline data time series and are used when specific analyses are necessary (e.g., for demand forecasting). Utilities with telemetry systems use SCADA to monitor their networks, setting up alarms for when something is out of the expected range, often creating false alarms. The fact is that, despite the big historical data available, seldom utilities use these data for real-time decision-making and system operation considering the data management and analysis limitations of existing IS. The lack of suitable methods, algorithms, or tools has been a barrier to the usefulness of real-time data.

Real-time measurements generally present errors (e.g., duplicated values, negative values, high or low values), called outliers, caused by equipment failure, network coverage, or data corruption. These errors should ideally be automatically identified, removed, and corrected before the data are used in engineering applications (e.g., hydraulic modelling, anomalous event detection). The tools used should not ignore the time correlation. Regardless of the acquisition system, measurements may not be simultaneously transmitted from all sensors, but rather by batches of sensors (e.g., a few hundred sensors at a time step).

Missing data are ubiquitous in real-time monitoring systems, causing bias in any analysis based on those data. To deal with this, there are two basic tactics: removing and imputing. The former is rarely used, as it reduces the dataset's size, causing information loss and irregular time series. The latter is the most effective option. Imputation algorithms should produce values matching the true data distribution, handle different types of data, and scale to larger datasets. Mean imputation is the most naive method, but it presents the worst performance since it ignores the link between close values. For real-time data, advanced data imputation algorithms should be used, such as Gaussian copula, temporal and spatial kNN, Hot Deck, multivariate imputation by chained equations, and recurrent neural networks.

An essential application of a DT-WDS is its short-term predictive capabilities. Considering that both real-time data and offline data from sensors are available, the Digital Twin can correctly predict the real network operation, assuming that no anomalous event is affecting the network. When the Digital Twin shows a different behavior (e.g., pressure at a node) compared to the real network, an anomaly may exist and affect the real network condition (e.g., unknown leak) or operation (e.g., wrong valve setting). Methods to detect anomalous events using offline time series can be adapted to be used in a real-time mode. Such techniques may include prediction and classification methods, statistical approaches, and unsupervised clustering techniques. Since the DT-WDS must reflect the current network behavior through real-time measurements, data quality is critical as raw data can be noisy, incomplete, and include various types of faulty measurements.

A DT-WDS can be used to run "what-if" scenarios of operations, to locate anomalous events, and to predict and prevent service failures (e.g., pump shutoff). Advanced data analytical tools (e.g., machine learning, optimization) may help in the daily operation of the WDS.

Many anomalous detection and location methods have also been developed, based on artificial intelligence (AI), inverse analysis, and data mining [35,36]. An anomalous event is usually characterized by an increase in water consumption which, in turn, leads to a decrease in the pressure-head in the surrounding area. This pressure drop is captured by installed sensors that can be used to infer the approximate location of the anomalous event in the network. To this end, clustering analysis of the network topology may be used to group the possible network locations (usually thousands of possible pipes) into a smaller number of sectors. Different techniques need to be studied, adapted, and compared, namely, graph-based proximity measures (e.g., proximity and adjacency matrices).

Water utilities have been using hydraulic models for several engineering applications and run simulations in batch mode with input parameters based on historical data. These models need to be updated using the data processed as explained previously, and real-time operational data from the real twin need to be periodically calibrated to accurately reproduce the network's real behavior.

A Digital Twin of a WDS may also measure the performance state of a system by calculating KPI and, based on the overall information of the system, simulations of operation scenarios (e.g., valve settings) can be run to support operational decisions. The DT-WDS will automatically assess the current state of the WDS by determining KPI for a specified and regular time step with the available data. Examples of these KPIs are the percentage of real losses, normalized energy consumption, and the ratio of energy in excess. Some of these KPIs require the calculation of water–energy balances running the hydraulic model. For that, several tools should be coupled to enhance this capability, such as water and energy balances' computation, optimization algorithms, and AI.

Notwithstanding the fast digital solution advances in recent years, the success of digital transformation in a water utility fundamentally relies upon the human element. Therefore, employees should be involved in the process and motivated to embrace this hard journey. Though enthusiasm is essential for digitalization, so is having knowledge and capabilities. In many water utilities, employees simply make spreadsheets and compile data for reports [37,38]. A water utility will be much more efficient, if its employees can explore the available data by using more advanced tools, for instance, machine learning or forecasting models. However, many water utilities face several difficulties in achieving their digital transformation, namely, ageing assets, ageing employees, and a lack of financial resources to invest.

8. Conclusions

This paper presented several computational tools developed under two FCT-funded R&D projects, which showed that there is a long road ahead in small- to medium-sized utilities towards full water digitalization. The tools developed and presented in this paper are a direct contribution to the digital transformation of utilities and should be able to be used in the daily management of their systems. These tools can have a direct impact on how WDS are managed, namely, by facilitating the data integration and KPI calculation required to support the decision-making process, by preparing time series to be used in different engineering applications, by increasing the benefit of installing pressure sensors, or by contribution of the planning process in defining rehabilitation strategies. However, for these tools to be properly used by the experts of the utilities, namely those of smaller size and with fewer financial resources, it is necessary to train them. Additionally, water utilities with fewer financial resources often have employees overloaded, which means that they lack the time to use the tools. In any case, it is considered that once tools are available and with the proper training, experts will use them very soon, even if only because of pressure from society because of the inefficiencies of these utilities in the use of resources.

The two projects have reached an end, but some R&D gaps of knowledge are still missing and should be addressed in the future, namely: (i) a lack of tools for the real-time processing of a massive amount of sensor data acquired with different measurement frequencies and synchronization times, (ii) the need for more accurate data-driven techniques to detect and locate anomalous events in real-time, (iii) a lack of knowledge on how to dynamically update and calibrate a hydraulic model for near real-time engineering applications, (iv) a lack of tools for the simulation of operation scenarios using near-real-time operational data, and (v) an absence of guidelines for creating tools based on experience-based knowledge.

Author Contributions: Conceptualization, N.C. and D.C.; methodology, B.F., A.A. and J.C.; software, A.A.; validation, B.F. and J.C.; investigation, N.C., B.F., A.A., J.C. and D.C.; data curation, A.A.; writing—original draft preparation, N.C.; writing—review and editing, B.F., A.A., J.C. and D.C.; visualization, A.A., B.F. and J.C.; supervision, N.C. and D.C.; project administration, N.C.; funding acquisition, N.C. and D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was jointly funded by Fundação para a Ciência e a Tecnologia, Programa Operacional Regional Lisboa 2020, and Programa Operacional Regional Alentejo 2020, through the DECidE project (grant number SAICT-POL/24135/2016 and LISBOA-01-0145-FEDER-024135 | ALT20-03-0145-FEDER-024135), and solely by Fundação para a Ciência e a Tecnologia through the WISDom project (grant number DSAIPA/DS/0089/2018), as well as the Ph.D. Research Studentship of Bruno Ferreira (grant number SFRH/BD/149392/2019).

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to being obtained from third parties.

Acknowledgments: The authors acknowledge Fundação para a Ciência e a Tecnologia, Programa Operacional Regional Lisboa 2020, and Programa Operacional Regional Alentejo 2020 for funding this research. The authors also acknowledge the experts from the participating water utilities for their relevant and constructive contributions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cardoso, M.A.; Brito, R.; Ribeiro, R.; Alegre, H. Infrastructure Asset Management—Maturity Assessment of Water Utilities Based on International Standards ISO 55000. 2017. Available online: http://repositorio.lnec.pt:8080/xmlui/handle/123456789/1010252 (accessed on 21 December 2022).
2. Cardoso, M.A.; Silva, M.S.; Coelho, S.T.; Almeida, M.C.; Covas, D.I.C. Urban Water Infrastructure Asset Management—A Structured Approach in Four Water Utilities. *Water Sci. Technol.* **2012**, *66*, 2702–2711. [CrossRef] [PubMed]
3. Okwori, E.; Pericault, Y.; Ugarelli, R.; Viklander, M.; Hedström, A. Data-Driven Asset Management in Urban Water Pipe Networks: A Proposed Conceptual Framework. *J. Hydroinform.* **2021**, *23*, 1014–1029. [CrossRef]
4. Carriço, N.; Ferreira, B.; Barreira, R.; Antunes, A.; Grueau, C.; Mendes, A.; Covas, D.; Monteiro, L.; Santos, J.; Brito, I.S. Data Integration for Infrastructure Asset Management in Small to Medium-Sized Water Utilities. *Water Sci. Technol.* **2020**, *82*, 2737–2744. [CrossRef] [PubMed]
5. Carriço, N.; Ferreira, B. Data and Information Systems Management for the Urban Water Infrastructure Condition Assessment. *Front. Water* **2021**, *3*, 670550. [CrossRef]
6. Ferreira, B.; Carriço, N.; Barreira, R.; Dias, T.; Covas, D. Flowrate Time Series Processing in Engineering Tools for Water Distribution Networks. *Water Resour. Res.* **2022**, *58*, e2022WR032393. [CrossRef]
7. Carriço, N.; Ferreira, B.; Antunes, A.; Grueau, C.I.C.; Barreira, R.; Mendes, A.; Covas, D.I.C.; Monteiro, L.; Santos, J.F.; Brito, I.S. An Information Systems for Infrastructure Asset Management Tailored to Portuguese Water Utilities: Platform Conceptualization and A Prototype Demonstration 2022. Available online: https://www.preprints.org/manuscript/202211.0295/v1 (accessed on 20 December 2022).
8. Alegre, H.; Coelho, S.T.; Covas, D.I.C.; do Céu Almeida, M.; Cardoso, A. A Utility-Tailored Methodology for Integrated Asset Management of Urban Water Infrastructure. *Water Sci. Technol. Water Supply* **2013**, *13*, 1444–1451. [CrossRef]
9. Cardoso, M.A.; Poças, A.; Silva, M.S.; Ribeiro, R.; Almeida, M.C.; Brito, R.S.; Coelho, S.T.; Alegre, H. Innovation Results of IAM Planning in Urban Water Services. *Water Sci. Technol.* **2016**, *74*, 1518–1526. [CrossRef] [PubMed]
10. Güngör, M.; Yarar, U.; Cantürk, Ü.; Fırat, M. Increasing Performance of Water Distribution Network by Using Pressure Management and Database Integration. *J. Pipeline Syst. Eng. Pract.* **2019**, *10*, 04019003. [CrossRef]
11. Kanakoudis, V.K. A Troubleshooting Manual for Handling Operational Problems in Water Pipe Networks. *J. Water Supply Res. Technol. Aqua* **2004**, *53*, 109–124. [CrossRef]
12. Barrela, R.; Amado, C.; Loureiro, D.; Mamade, A. Data Reconstruction of Flow Time Series in Water Distribution Systems—A New Method That Accommodates Multiple Seasonality. *J. Hydroinform.* **2017**, *19*, 238–250. [CrossRef]

13. Shafiee, M.E.; Rasekh, A.; Sela, L.; Preis, A. Streaming Smart Meter Data Integration to Enable Dynamic Demand Assignment for Real-Time Hydraulic Simulation. *J. Water Resour. Plan. Manag.* **2020**, *146*, 06020008. [CrossRef]
14. Grievson, O.; Holloway, T.; Johnson, B. (Eds.) *A Strategic Digital Transformation for the Water Industry*; IWA Publishing: London, UK, 2022; ISBN 978-1-78906-340-0.
15. von Ditfurth, H.; Weisbord, E.; Danielsen, T.; Zutari, L.F.-J.; Hafemann, A.C.; Hima, J.; Oraeki, T.C. Digital Water: An Overview of the Future of Digital Water from a YWP Perspective. Available online: https://iwa-network.org/publications/digital-water-an-overview-of-the-future-of-digital-water-from-a-ywp-perspective/ (accessed on 13 September 2022).
16. Karmous-Edwards, G.; Tomić, S.; Cooper, J.P. Developing a Unified Definition of Digital Twins. *J. AWWA* **2022**, *114*, 76–78. [CrossRef]
17. Bonilla, C.A.; Zanfei, A.; Brentan, B.; Montalvo, I.; Izquierdo, J. A Digital Twin of a Water Distribution System by Using Graph Convolutional Networks for Pump Speed-Based State Estimation. *Water* **2022**, *14*, 514. [CrossRef]
18. Conejos Fuertes, P.; Martínez Alzamora, F.; Hervás Carot, M.; Alonso Campos, J.C. Building and Exploiting a Digital Twin for the Management of Drinking Water Distribution Networks. *Urban Water J.* **2020**, *17*, 704–713. [CrossRef]
19. Ferreira, B.; Carriço, N.; Covas, D. Optimal Number of Pressure Sensors for Real-Time Monitoring of Distribution Networks by Using the Hypervolume Indicator. *Water* **2021**, *13*, 2235. [CrossRef]
20. Gomes, S.C.; Vinga, S.; Henriques, R. Spatiotemporal Correlation Feature Spaces to Support Anomaly Detection in Water Distribution Networks. *Water* **2021**, *13*, 2551. [CrossRef]
21. Capelo, M.; Brentan, B.; Monteiro, L.; Covas, D. Near–Real Time Burst Location and Sizing in Water Distribution Systems Using Artificial Neural Networks. *Water* **2021**, *13*, 1841. [CrossRef]
22. Cabral, M.; Loureiro, D.; Flores-Colen, I.; Covas, D. A Distress-Based Condition Assessment Approach of Urban Water Assets Using Novel Deterioration Indices. *Water Resour. Manag.* **2022**, *36*, 1075–1092. [CrossRef]
23. Monteiro, L.; Algarvio, R.; Covas, D. Enhanced Water Age Performance Assessment in Distribution Networks. *Water* **2021**, *13*, 2574. [CrossRef]
24. ERSAR. *Water and Waste Services Quality Assessment Guide: 2nd Generation of the Assessment System*; ERSAR: Lisbon, Portugal, 2017; ISBN 978-989-8360-11-3.
25. Alegre, H.; Baptista, J.M.; Cabrera, E.; Cubillo, F.; Duarte, P.; Hirner, W.; Merkel, W.; Parena, R. *Performance Indicators for Water Supply Services*, 3rd ed.; Alegre, H., Baptista, J.M., Cabrera, E., Jr., Cubillo, F., Duarte, P., Hirner, W., Merkel, W., Parena, R., Eds.; IWA Publishing: London, UK, 2017; ISBN 978-1-84339-051-0.
26. Mamade, A.; Loureiro, D.; Alegre, H.; Covas, D. A Comprehensive and Well Tested Energy Balance for Water Supply Systems. *Urban Water J.* **2017**, *14*, 853–861. [CrossRef]
27. Quevedo, J.; Puig, V.; Cembrano, G.; Blanch, J.; Aguilar, J.; Saporta, D.; Benito, G.; Hedo, M.; Molina, A. Validation and Reconstruction of Flow Meter Data in the Barcelona Water Distribution Network. *Control Eng. Pract.* **2010**, *18*, 640–651. [CrossRef]
28. Ferreira, B.; Antunes, A.; Carriço, N.; Covas, D. Multi-Objective Optimization of Pressure Sensor Location for Burst Detection and Network Calibration. *Comput. Chem. Eng.* **2022**, *162*, 107826. [CrossRef]
29. Ferreira, B.; Antunes, A.; Carriço, N.; Covas, D. Comparison of Model-Based Techniques for Pipe Burst Location in Water Distribution Networks. *IOP Conf. Ser. Earth Environ. Sci.* **2023**, *1136*, 012039. [CrossRef]
30. Caetano, J.; Carriço, N.; Covas, D. Lessons Learnt from the Application of MCDA Sorting Methods to Pipe Network Rehabilitation Prioritization. *Water* **2022**, *14*, 736. [CrossRef]
31. Pedersen, A.N.; Borup, M.; Brink-Kjær, A.; Christiansen, L.E.; Mikkelsen, P.S. Living and Prototyping Digital Twins for Urban Water Systems: Towards Multi-Purpose Value Creation Using Models and Sensors. *Water* **2021**, *13*, 592. [CrossRef]
32. Shao, Y.; Chu, S.; Zhang, T.; Yang, Y.J.; Yu, T. Real-Time Water Distribution System Hydraulic Modeling Using Prior Demand Information by Formal Bayesian Approach. *J Water Resour. Plan. Manag.* **2019**, *145*, 04019059. [CrossRef] [PubMed]
33. Abu-Mahfouz, A.M.; Hamam, Y.; Page, P.R.; Adedeji, K.B.; Anele, A.O.; Todini, E. Real-Time Dynamic Hydraulic Model of Water Distribution Networks. *Water* **2019**, *11*, 470. [CrossRef]
34. Zhou, X.; Xu, W.; Xin, K.; Yan, H.; Tao, T. Self-Adaptive Calibration of Real-Time Demand and Roughness of Water Distribution Systems. *Water Resour. Res.* **2018**, *54*, 5536–5550. [CrossRef]
35. Hu, Z.; Tan, D.; Chen, B.; Chen, W.; Shen, D. Review of Model-Based and Data-Driven Approaches for Leak Detection and Location in Water Distribution Systems. *Water Supply* **2021**, *21*, 3282–3306. [CrossRef]
36. Wu, Y.; Liu, S. A Review of Data-Driven Approaches for Burst Detection in Water Distribution Systems. *Urban Water J.* **2017**, *14*, 972–983. [CrossRef]
37. Ng, J.H.; Seah, H.; Pang, C.M. *Digitalising Water—Sharing Singapore's Experience*; IWA Publishing: London, UK, 2020.
38. Pesantez, J.E.; Alghamdi, F.; Sabu, S.; Mahinthakumar, G.; Berglund, E.Z. Using a Digital Twin to Explore Water Infrastructure Impacts during the COVID-19 Pandemic. *Sustain. Cities Soc.* **2022**, *77*, 103520. [CrossRef] [PubMed]