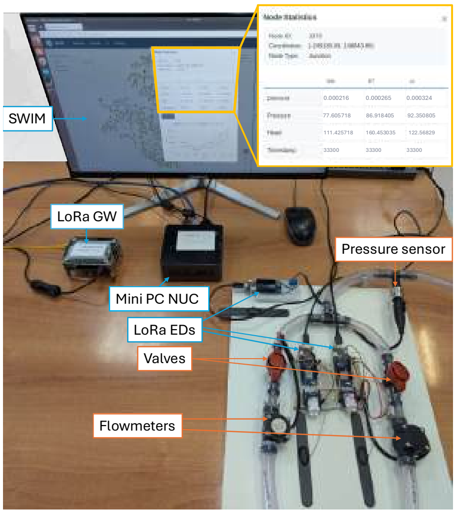

Fig. 5: Real-world demonstration of SWIM, showcasing integration with IoT devices and Deep Neural Network predictions [21].

- **SIM:** Data derived from classical hydraulic simulations performed with WNTR, based on EPANET-compatible ".inp" files.
- **AI:** Predictions from Deep Neural Network, estimating hydraulic parameters based on initial conditions.

SWIM includes an IoT device management interface that displays device properties such as ID, topic, IP address, network port, assigned node, and status, which become accessible upon selecting the "Devices" tab in Fig. 5. Users can manage devices via action buttons, enabling seamless interactions. SWIM's modularity and flexibility enable the integration of automatic logic which can easily implemented through custom code (e.g., Python) to enable DECISION/REPORT routines embedded within the framework. Lastly, SWIM popup menu showing detailed node information and hydraulic data from SIM, RT, and AI sources.

4.2 Graph representation of WDS

In this section, we present a concise view of the fundamental notions of graph theory, collected in a single reference point for the basic terminology, following standard definitions in [27]. We then outline how graph-theoretic models can be employed to represent WDS.

A graph $\mathcal{G} = (V, E)$ consists of a finite set of vertices V and a finite set of edges $E \subseteq [V]^2$, where each edge is an unordered pair of vertices. The elements of V are referred to as vertices (or nodes), while the elements of E are edges (or links). A vertex is said to be incident to an edge if it belongs to that edge, which in turn connects its two end vertices. An edge joining vertices x and y is denoted by xy. Two vertices are adjacent if they are connected by an edge. The neighborhood of a vertex v, denoted by $\mathcal{N}(v)$, is the set of vertices adjacent to v, and its degree is the cardinality of this set. Two distinct edges are called adjacent if they share a common vertex.

Let $|V| = N$ and $|E| = N_e$. The adjacency matrix $\mathbf{A} \in \mathbb{R}^{N\times N}$ provides a matrix representation of the graph connectivity, where the entry a_{ij} is zero when no edge exists between nodes i and j. Depending on the application, $\mathbf{A}$ may be binary or weighted. In weighted graphs, edge weights can encode quantities such as statistical dependencies between node signals or physical attributes of WDS components, for example pipe diameters. The degree matrix $\mathbf{D}$ is a diagonal matrix whose entries are given by $d_{ii} = \sum_{j=1}^{N} a_{ij}$. The graph Laplacian is then defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \tag{7}$$

and plays a central role in spectral graph analysis by capturing key structural characteristics of the network.

Within the context of GNN-based modeling of WDSs, it is necessary to formalize graph-structured data representations tailored to these systems. Nodes correspond to physical elements of the network, such as junctions, reservoirs, or pumps, while edges represent hydraulic connections, typically pipes. A graph signal associates a scalar or vector quantity to each node or edge, describing system variables such as pressure, flow, or water quality over time. Node attributes are collected in a feature matrix $\mathbf{X} \in \mathbb{R}^{N\times d}$, where each row $x_v \in \mathbb{R}^d$ contains the features of node v. Similarly, edge attributes are stored in $\mathbf{X}^e \in \mathbb{R}^{N_e\times c}$, with each vector $x^e_{u,v} \in \mathbb{R}^c$ encoding properties of a pipe, such as length, diameter, or roughness. Graph-level features may also be defined to summarize global properties or labels of the entire network. These representations constitute the inputs to GNN architectures, which often rely on normalized adjacency or Laplacian operators to enable stable training and effective information propagation across the graph.

A common and intuitive way to represent a WDS using graph theory relies on a direct abstraction of its physical structure, where junctions are modeled as nodes and pipes as edges. Under this representation, water transfer between two points in the network is captured by an edge linking the corresponding nodes. Specifically, an edge is included whenever a pipe physically connects the junctions associated with nodes. This graph formulation allows the inclusion of additional information through node and edge attributes. Node features may describe properties such as the elevation of a junction above sea level or its water demand, typically expressed in meters and liters per second, respectively. Edge weights can encode physical or hydraulic characteristics of the pipes, including their length, diameter, or head loss, as adopted for instance in [28]. An illustrative example of this graph construction is shown in Fig. 6. Owing to its simplicity and physical interpretability, this representation is widely adopted in the literature, particularly for tasks related to leak detection and localization.

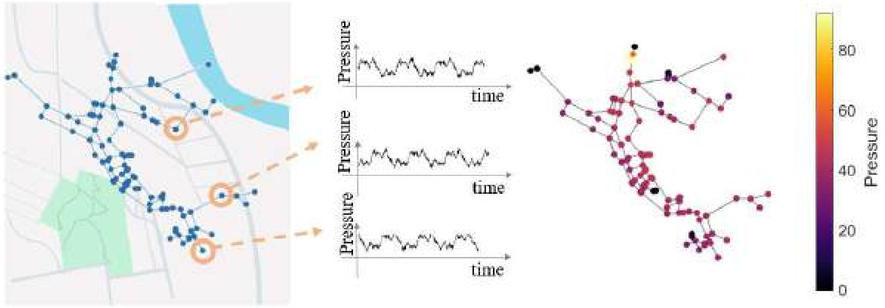

Fig. 6: Graph construction from a WDS by abstracting junctions and pipes. The left panel shows the Net3 network overlaid on the geographic map, while the right panel presents the corresponding graph representation, where nodes represent junctions and edges represent pipes. Node attributes correspond to time-varying pressure values for the highlighted junctions, with colors indicating pressure magnitude [7].

In real-world deployments, it is often infeasible to instrument every junction of a WDS with sensors due to cost limitations or physical constraints. As a result, sensing infrastructure is typically installed only at selected locations in the network [29]. Under this constraint, the graph representation includes only the junctions for which measurements are available, and the resulting edges do not necessarily coincide with physical pipes. Instead, they represent abstract relationships among sensed locations, leading to a reduced or incomplete graph in which connectivity reflects data availability rather than physical topology.

Sensors installed in WDS can monitor different variables, such as hydraulic pressure [30] or water quality, either continuously or at discrete intervals. The resulting observations form time series (signals) associated with each sensing location. These time series can be analyzed to assess similarity or dependency across different points in the network. When two signals exhibit strong similarity, the corresponding nodes are connected by an edge in the graph.

Similarity between time series can be quantified using a variety of metrics, including pointwise distances such as the Euclidean distance [31] or error-based measures like the mean absolute percentage error, as well as statistical indicators such as Pearson's correlation coefficient [32]. More advanced techniques, including dynamic time warping [33] and compression-based dissimilarity [34], provide increased robustness to temporal misalignment.

Beyond physical and correlation-based abstractions, alternative graph representations can provide complementary perspectives for graph-based analysis of WDSs. One option is to exploit the operational partitioning commonly adopted by water utilities. WDSs are often divided into District Metered Areas (DMAs), each monitored by flow and pressure sensors at key entry points and selected internal locations. A graph model reflecting this segmentation can naturally encode operational practices for leak detection and pressure control while limiting sensor deployment costs [29].

Another representation, proposed in [8], is based on the line graph $\mathcal{G}_L = (V_L, E_L)$ associated with the original graph $\mathcal{G}$. In this construction, each node $v_l \in V_L$ corresponds to an edge of $\mathcal{G}$, and two nodes in $\mathcal{G}_L$ are connected if the corresponding edges in $\mathcal{G}$ share a common endpoint [35]. The adjacency matrix $\mathbf{A}_L \in \mathbb{R}^{N_e \times N_e}$ is given by

$$\mathbf{A}_L = \mathbf{B}\mathbf{B}^T - 2\mathbf{I}, \tag{8}$$

where $\mathbf{B} \in \mathbb{R}^{N \times N_e}$ is the incidence matrix of $\mathcal{G}$.

This transformation enables edge-based quantities, such as hydraulic flows, to be treated as node signals in $\mathcal{G}_L$, allowing the application of graph signal processing techniques [36]. This framework supports the reconstruction of flow values from sparse measurements, as demonstrated in [8]. Moreover, a node-ranking strategy on the line graph can identify which flows are more predictable from neighboring information, guiding sensor placement toward less interconnected edges. These typically correspond to inter-cluster links in the original graph, which are known to strongly influence network spectral properties [37]. Motivated by this insight, node centrality measures [38] are adopted to quantify the structural relevance of network components.

5 Machine learning for water demand and network traffic analysis

The progressive digitalization of WDSs, enabled by the diffusion of Low-Power Wide-Area Network communications such as LoRaWAN and advanced metering infrastructures, has opened the way for the systematic application of AI and data-driven models. The continuous acquisition of high-resolution observations on pressure, flow and consumption patterns provides an informational layer that complements traditional hydraulic modeling, allowing water utilities to evolve from reactive to predictive and proactive management strategies.

The scope of AI in the modern WDS spans three complementary domains. First, it targets the analytics of hydraulic variables, such as pressures, flows, and most distinctively, water demand, by learning temporal patterns and producing insights and short-term or long-term forecasts. Second, AI can operate at the communication and infrastructural layer of the smart metering system itself, where the behavior of LoRaWAN devices and GWs can be monitored to ensure data reliability and continuity of service, or even more, to inspect the physical layer of the network looking for anomalies or other unusual patterns. Finally, AI methods enhance the overall interaction with the WDS ecosystem by powering operator-centric interfaces, LLM-based assistants and simulation tools that improve situational awareness and decision-making.

Within this framework, machine learning techniques make it possible to extract knowledge from raw time series, forecast the near-future behavior of the system and identify anomalous or unexpected operational conditions. Over the last years, a wide range of state-of-the-art models has emerged to address these tasks, supported by the growing scientific and industrial interest in data-driven water management. In this scenario, our work introduces new approaches tailored to the characteristics of modern WDSs. These methods demonstrate significant advantages in their ability to address task-specific challenges in the water domain.

5.1 Demand characterization

High-resolution smart meter data enable the analysis of user behavior. Clustering methods can group residents according to their typical weekly or seasonal profiles, revealing temporal patterns that can support tariff design, demand-response programs, or targeted infrastructure planning. One of the main contributions of the project in the area of demand characterization occurs in soft clustering. The k-means algorithm is the most prominent clustering method. It separates the data into k clusters for a given k, but enforces a strict cluster assignment that works well only if there is a clear distinction between groups. However, the reality may be more heterogeneous due to some user characteristics that span multiple clusters. The assignments by hard clustering could be too restrictive, as opposed to soft clustering, which provides membership scores.

Fuzzy C-means (FCM) is the soft relaxation of k-means as it eases the strict constraint on hard clustering by providing each data sample with a degree of membership to all clusters. In this context, to further empower FCM, we have implemented a deep learning architecture that obtains cluster-friendly latent representations for soft clustering. Specifically, we employ the Deep Normalized Fuzzy Compactness and Separation (DNFCS [39]) neural network, which is trained by minimizing the reconstruction error and the within-cluster distance while maximizing the between-cluster distance. This unsupervised method yields the desired cluster memberships.

The demand characterization is performed in a real district metered area comprising around 1000 LoRaWAN smart meters handled by Unidata S.p.A. where users are clustered by the DNFCS based on their weekly profile averaged over one year of hourly observations from July 2022 to July 2023. The algorithm delivers the most interesting results using $k = 4$ classes. Their median profiles are plotted in Fig. 7: cluster 1 in blue shows a strong evening peak, cluster 3 in green a strong morning peak, while clusters 2 and 4 display different properties. Examples in Fig. 8 and Fig. 9 provide the likelihood of some users' profiles belonging to each class, where the highest probability corresponds to the most likely class. A visual inspection suggests that the first user belongs to cluster 1 due to a distinct evening peak in water usage. In contrast, the second user reveals characteristics of both clusters 1 and 3, with high demand observed in both morning and evening. The cluster memberships reflect strong confidence in the first assignment (83.3% to cluster 1), but high uncertainty for the second user (50% for cluster 1 and 47.6% for cluster 3). These scores provided by soft clustering can be more informative for water utilities compared to hard clustering as they provide additional insights into users' behavior.

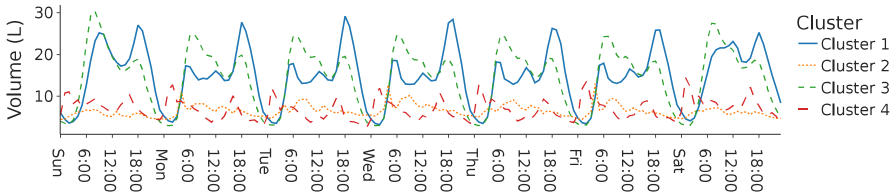

Fig. 7: Clusters average weekly consumption profiles by DNFCS [40]. Cluster 1: evening peak; Cluster 3: morning peak; Cluster 2, Cluster 4: other.

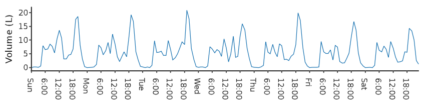

Fig. 8: Cluster memberships: [83.3% 2.7% 12.2% 1.8%] [40].

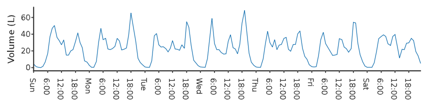

Fig. 9: Cluster memberships: [50% 1.1% 47.6% 1.3%] [40].

5.2 Demand forecasting

Predicting water demand is a crucial task for operational planning of water distribution. An important aspect emerging from studies in the field is the need to move from classical deterministic forecasting to probabilistic forecasting, since this approach enables water utilities to manage pump scheduling and reservoir levels more efficiently. Single-point predictions overlook the fact that various factors, such as weather, human activity, and unexpected events, can cause significant variations in water demand throughout the day and week. If the single-point forecast misses a peak, releasing a reduced amount of water can lead to low pressures at network nodes and tank depletion, potentially incurring contractual penalties when supply requirements are not met; on the other hand, releasing a large volume can cause high pressures, resulting in significant water waste and economic losses. Single-point forecasts neglect the risks associated with demand volatility, forcing operators to be very conservative in resource allocation.

In this context, we are the first to explore denoising diffusion probabilistic models (DDPMs [42]) for water demand forecasting. These generative methods estimate the target distribution (i.e., the future demand) by iteratively denoising stochastic trajectories conditioned on input data (i.e., the past observations). Rather than single-point forecasts, the proposed diffusion model, which we have named OSMOSIS, produces multi-point forecasts that form an empirical distribution of future water demand. The prediction intervals (such as 50% and 95%) are then derived by computing the upper and lower quantiles, thus showing the uncertainty level at each time step, as illustrated in Fig. 10. In this example, a solid deterministic deep learning model like the Transformer fails to match the

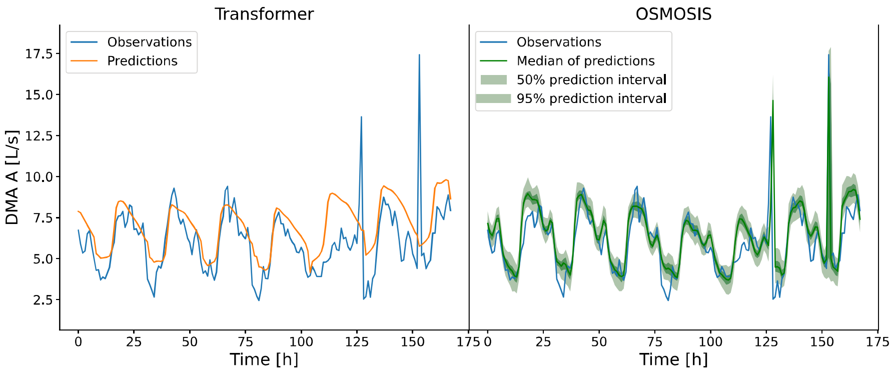

Fig. 10: Plots of net flows over 7 days predicted by Transformer and OSMOSIS for a district of the BWDF dataset [41]. The blue line represents the actual observations, the orange line the deterministic forecasts, and the prediction intervals are colored in green.

amplitude of consumption and anticipate peak demands for 7-day hourly water demand forecasting of a district metered area in the BWDF dataset [43]. In contrast, OSMOSIS has shown the ability to better capture spikes and amplitudes while quantifying uncertainty through prediction intervals. This probabilistic perspective is particularly relevant for daily operational routines since they are influenced by the volatility of water demand.

5.3 Demand disaggregation

The availability of sub-minute smart meter data enables in-depth analysis of water consumption habits in households. In particular, water disaggregation refers to estimating the consumption of multiple items like shower, toilet, washing machine, or dishwasher based only on the aggregate water consumption. Users can track ongoing activities at home and receive feedback to reduce water consumption. Several studies have reported that providing users with detailed real-time feedback can significantly encourage water conservation practice. Water disaggregation is also known as non-intrusive water monitoring because it separates analytically the total consumption, hence without installing a smart meter for each device. In contrast, intrusive monitoring is expensive and usually faces user privacy concerns. Sub-minute monitoring makes disentangling the individual signals very challenging. For instance, washbasins and toilets are used sporadically during the day, while washing machines and dishwashers follow a predefined and long program. Therefore, their combination results in an intricate aggregate signal.

We have contributed to water disaggregation with the first application of a UNet architecture that extracts fine-grained and coarse temporal features from the aggregate signal using a series of CNN-based encoding and decoding blocks

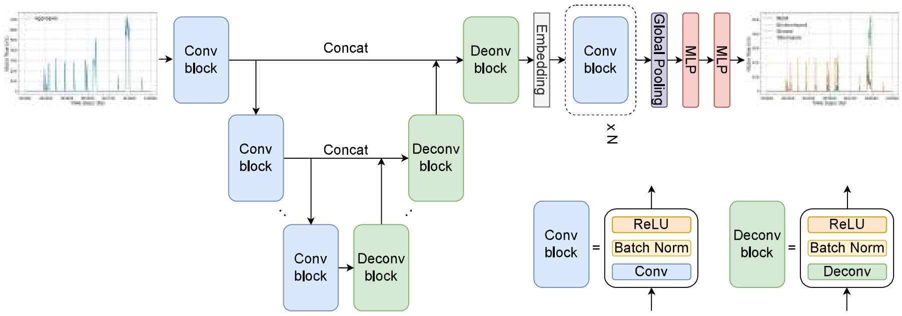

Fig. 11: The UNet model for non-intrusive monitoring of water fixtures. [44].

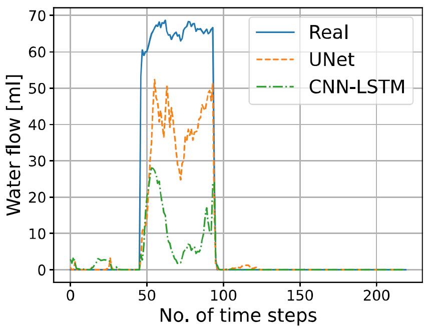

Fig. 12: Shower [44].

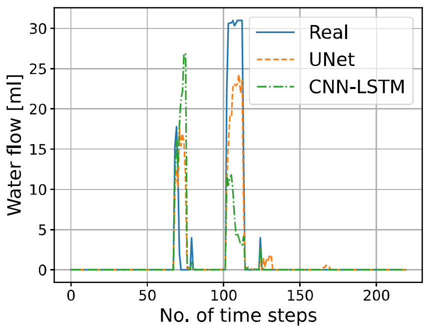

Fig. 13: Washbasin [44].

[44], as shown in Fig. 11. The proposed method is evaluated against a CNN-LSTM on the WEUSEDTO dataset [45] with temporal resolution of 6 seconds. A visual comparison of shower water flow disaggregation in Fig. 12 demonstrates the UNet's superior performance. Its predictions (orange line) align more closely with the real flow data (blue line), depicting a higher peak usage. Similarly, for washbasin disaggregation in Fig. 13, the UNet again outperforms the CNN-LSTM. The UNet tracks the real water flow more closely during periods of activity, whereas the CNN-LSTM shows significant discrepancies, initially over-estimating and then underestimating actual water usage. In general, further research must prioritize accurate non-intrusive water monitoring, which remains an underexplored field, as a crucial step for achieving sustainable consumption goals.

5.4 Network monitoring

Monitoring the health of large-scale IoT deployments is a fundamental requirement for ensuring reliable data acquisition in modern WDSs. In LoRaWAN-based smart metering networks, devices transmit asynchronously, with mostly unconfirmed uplinks and highly variable radio conditions [46]. In real deployments, a single district may include tens of thousands of meters, each producing irregular traffic influenced by propagation conditions, gateway availability, firmware behavior and environmental dynamics [47]. Under circumstances, manual inspection of communication logs becomes impractical: packet, unusual frame-counter patterns, empty payloads or unstable data rates often appear intertwined, and distinguishing a genuine malfunction from a temporary network fluctuation rarely possible by visual inspection alone [48]. A missing transmission, for example, may be caused by interference or shadowing rather than by a faulty device, while regular uplink sequences may still conceal subtle firmware or protocol inconsistencies [49].

This ambiguity makes device diagnostics in LoRaWAN fundamentally more complex than in traditional network monitoring. The goal is to detect anomalies and more importantly to understand why they occur and whether they originate from the device, the radio environment or higher-level network conditions. This distinction is essential in water utilities, where maintenance actions are costly and typically planned over geographical clusters: operators need to identify coherent groups of devices exhibiting similar behavioral drifts rather than reacting to isolated alerts [50].

To address this challenge, we designed a data-driven diagnostic architecture that leverages machine learning to analyse the time-series statistics extracted from the communication history of each device. Instead of relying on fixed thresholds or manual rules, the proposed pipeline learns to identify recurrent behavioral patterns and infer the most likely health state of each device. The process begins with the extraction of descriptive indicators from raw uplinks—such as radio-quality fluctuations, protocol-level inconsistencies and transmission continuity—which jointly capture how the device interacts with the network. These heterogeneous signals are then represented in a unified feature space and fed to an unsupervised learning stage that groups devices exhibiting comparable operating conditions. The resulting clusters serve as behavioral archetypes, enabling the model to recognise degradation trajectories that extend across large subsets of the fleet.

The full diagnostic workflow, combines these unsupervised embeddings with a lightweight neural module that refines the inference by integrating evidence from all feature domains. The outcome is a probabilistic assessment of device health that is both interpretable and actionable. This probability summarizes the reliability of each meter, while the spatial distribution of unhealthy devices highlights areas where correlated issues are emerging.

In operational deployments exceeding 200,000 devices and tens of millions of transmissions, this architecture has proved able to identify coherent degradation clusters and anticipate device malfunction well before data loss becomes critical (a reference real application scenario is reported in Fig. 14).

By transforming raw LoRaWAN traffic into structured health indicators, the approach provides utilities with continuous visibility over their metering infrastructure, reduces the risk of large-scale data gaps and supports proactive field

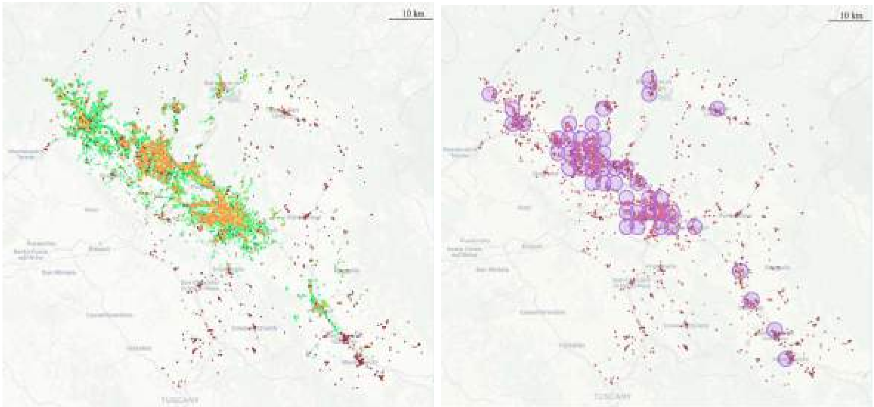

Fig. 14: Spatial inference of device health across a selected deployment. **Left:** all devices color-coded by operational probability (gradient from green: $\approx 100\%$ working confidence, to red: $\approx 10\%$ working confidence). **Right:** only faulty devices, with purple circles marking areas exceeding 200 faulty nodes per km^2, an example of optimal sites for the operator's targeted maintenance.

interventions. Most importantly, it replaces manual inspection and threshold-based rules with a principled, scalable and data-driven method, forming a robust digital foundation upon which demand analysis, hydraulic modeling and real-time network analysis can reliably operate.

6 Holistic Tools for Water Operators

The progressive digitalization of WDSs has opened up new opportunities for operators to observe, analyse and manage their networks with a level of detail that was previously inaccessible. At the same time, the increasing complexity of the data ecosystem, comprising IoT telemetry, consumption time series, hydraulic simulations and AI models, poses a significant challenge for utilities. Operational workflows require tools that are not only scientifically robust, but also practical, interpretable and aligned with real-world decision-making processes.

In this context, holistic platforms play a central role by bridging the gap between research-driven methods and the concrete needs of water utilities. Instead of addressing isolated tasks, these platforms provide a unified environment where advanced analytics, simulation engines and smart-meter data are seamlessly integrated. Their goal is to offer a comprehensive operational view, enabling operators to explore both the physical and digital layers of the network, detect anomalies, validate sensor behavior, anticipate demand and support long-term strategic planning.

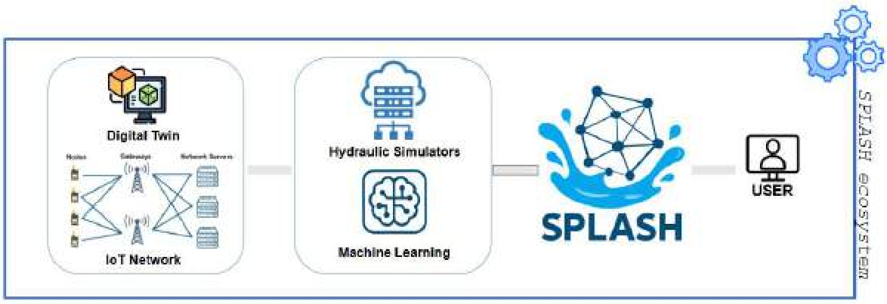

Fig. 15: High-level overview of the SPLASH ecosystem. IoT telemetry and the Digital Twin feed hydraulic simulators and machine learning modules, which in turn provide an operator-facing interface for monitoring and decision support.

6.1 SPLASH - Smart Platform for Large Scale Holistic Water Management

The SPLASH platform represents a concrete example of how digital-twin technologies, IoT networks, hydraulic models and machine learning can be combined into a single, operator-oriented ecosystem. It is particularly designed to transform heterogeneous data sources into actionable insights for analysts, engineers and field personnel. Its architecture integrates the physical water infrastructure with its digital counterpart, enabling a continuous exchange of information between devices, simulators and decision-support modules.

At its core, SPLASH is built around three main pillars, illustrated in Fig. 15:

- **IoT and Digital Twin layer**: a real LoRaWAN-based metering infrastructure in which smart meters, GWs and NSs are mirrored through Eclipse Ditto digital twins, providing real-time device state abstraction;
- **Hydraulic Simulation layer**: simulation engines such as EPANET and DYN-WNTR reconstruct pressures, flows and system behavior at scale, creating a virtual replica of the water network for validation, analysis and predictive studies;
- **Machine Learning layer**: advanced models support tasks including demand forecasting, anomaly detection, device diagnostics and pattern-based analytics, enriching the simulation outputs with data-driven insights.

These components converge within an integrated web interface specifically designed for water operators. SPLASH allows users to interact with the network through multiple coordinated visualizations:

Mapping: gateway-centric and device-centric radio-signal maps allow operators to inspect network connectivity, detect coverage gaps and identify areas where communication anomalies may originate, as shown in Fig. 16.

Time-series: aggregated and per-device consumption curves reveal abnormal patterns, peak demands, and irregular behaviors across the metering fleet, as illustrated in Fig.17.