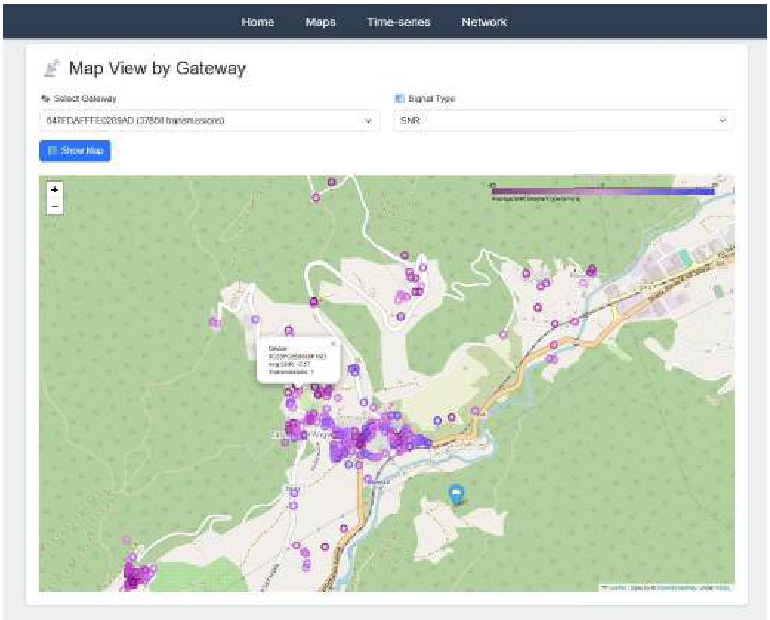
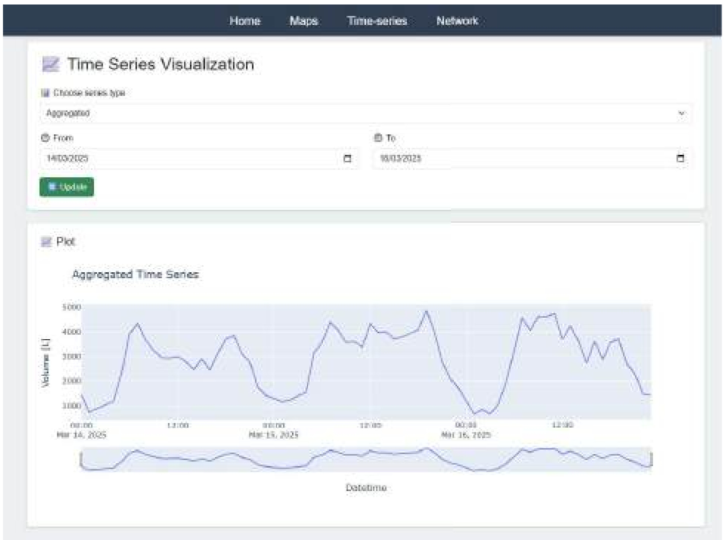


Map View by Gateway

Select Gateway: 647FDAFFFE0289AD (37850 transmissions)

Signal Type: SNR

Fig. 16: Example of SPLASH geospatial mapping for device and gateway inspection.



Time Series Visualization

Choose series type: Aggregated

From: 14/03/2025

To: 16/03/2025

Fig. 17: Consumption time-series visualization within the SPLASH platform.

Fig. 18: SPLASH hydraulic simulation view, enabling comparison between real and modelled network states.

Hydraulic simulation: real and simulated states, such as pressures, expected demand, or valve conditions, are compared at both network and device scales, providing a practical way to validate sensor readings and detect system-wide discrepancies (see Fig. 18).

6.2 Agent

This section presents the design and implementation of the intelligent virtual agent within our framework. We discuss the rationale behind the chosen architecture, the integration of the language model, and the strategies employed to enable effective interaction with domain-specific data. The section also provides an evaluation of the implemented system, highlighting its effectiveness and potential areas for improvement.

In recent years, Large Language Models (LLMs) have represented a breakthrough in artificial intelligence, offering advanced capabilities in natural language processing. Models such as GPT (Generative Pretrained Transformer) and Claude demonstrate the ability to understand, generate, and manipulate text with consistency and relevance. Furthermore, LLMs can serve as the cognitive core of intelligent virtual agents, enabling natural language interaction, context understanding, data access, API usage, and decision-making. Especially in technical and industrial fields, these agents simplify user interaction with complex systems, reducing the need to engage directly with the underlying technologies.

The use of LLMs in critical infrastructure contexts, such as WDS, is an emerging area of research with significant potential for improving human-machine interaction, accessibility to technical data, and operational efficiency. In WDS, LLMs are finding innovative applications, particularly as natural language interfaces to complex technical models and hydraulic simulators.

A prime example of this trend is the LLM-EPANET framework [51], which connects an LLM to the EPANET hydraulic simulator. In this system, users can interact with sophisticated water supply models simply through text requests, without necessarily having to know the technical details of the domain. The goal of these solutions is twofold: on the one hand, to allow even non-specialist operators to perform hydraulic simulations, configure scenarios, and obtain results in an intuitive way; on the other hand, to provide clear and contextualized explanations of the data and analyzes produced by the simulator, contributing to a broader and more accessible understanding of network dynamics.

Design of an Agent for WDS In the implemented framework, we chose not to fine-tune the language model. Instead, we adopted a pre-trained LLM, integrated into a LangChain-based software environment. The main contribution concerns interaction engineering, through: i) the formulation of structured prompts; ii) the design of the agent and its decision-making tools; iii) the dynamic integration of domain data. This choice made it possible to reduce the development effort and explore the potential of a generalist LLM.

The design of the model involves the orchestration of two main pipelines. The first is for extracting the network topology from a Neo4j model, and the second is for retrieving historical and real-time sensory data from the time-series database. Indeed, as already presented, the WDS can be effectively represented as a directed graph composed of nodes and edges.

In particular, in the dataset used in this implementation, each node is associated with various pieces of information, called features, which describe its hydraulic and structural status over time, such as consumption (demand), pressure, piezometric head, geographical coordinates (position), and any leaks present. Similarly, each arc has specific operational attributes such as flow rate, flow velocity, and opening status.

The model was designed to simultaneously learn topological relationships (i.e., the connection structure between nodes and pipes) and hydraulic dynamics (how sensory data evolve over time). To this end, temporal snapshots were constructed, representing the overall state of the network at regular time intervals, integrating both the graph topology and information from the sensors. The ultimate goal is to create a predictive system capable of combining structural and operational information in a single processing flow.

Orchestrator agent as the central component of the architecture The implementation of the orchestrator agent is the core of the proposed architecture. This component has the function of putting the human operator in direct communication with the intelligent support models, in particular, the leak detection system based on GNN and the dynamic data visualization module.

The goal is to offer a natural, intelligent, and contextually aware conversational interface capable of interpreting user requests, analyzing their historical context, and autonomously activating the most appropriate computational tools to provide relevant and useful responses.

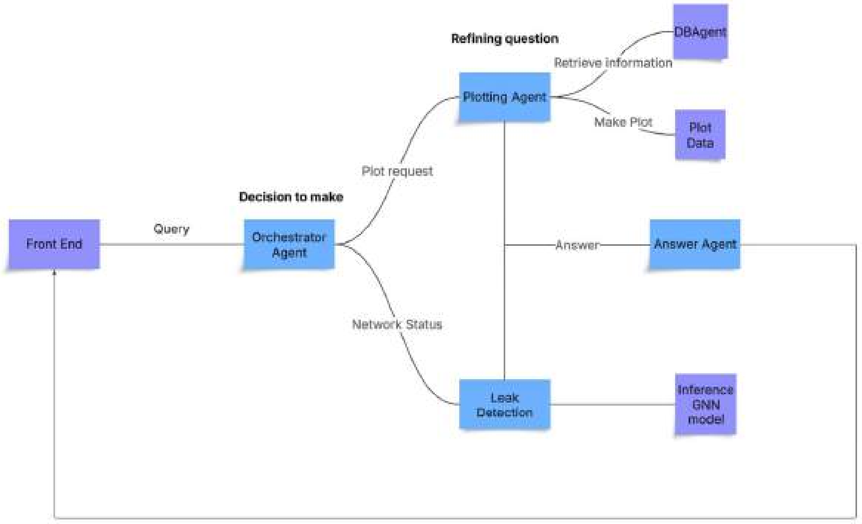


Fig. 19: Diagram of the logical architecture of the conversational agent for water network management.

As outlined in Fig. 19, the process is initiated via a user interface, the Front End, which transmits a `Query` to the central controller. The Orchestrator Agent plays a critical role in this architecture, acting as the central decision-maker that interprets operator queries and routes them to the most appropriate module. This hierarchical orchestration allows the system to distinguish between different types of requests. For instance, queries related to network status and anomalies are routed to the Leak Detection module, while requests for data retrieval and visualization are managed by the Plotting Agent.

The Leak Detection module exemplifies the benefits of specialized processing within the framework. By employing a GNN, the system captures the inherent graph structure of the WDS as nodes—and models the complex hydraulic and spatial-temporal relationships within the network. Similarly, the Data Analysis and Visualization pathway demonstrates how the framework supports operator insights through historical data review and predictive plotting. The Plotting Agent, in coordination with the DBAgent, retrieves and refines data from the WDS repository, then generates visualizations such as pressure profiles, demand trends, or pump efficiency charts. Its main function is to translate the user's natural language requests, interpreted by the main agent, into correctly formatted SQL queries optimized for access to the PostgreSQL database containing the water network's sensor data.

In this way, the LLM agent plays a crucial role in abstracting the complexity of the data layer, allowing the orchestrator to obtain relevant hydraulic information (e.g., pressure, demand, local anomalies) without direct interaction with the database. This approach promotes modularity, security, and reusability, configuring the agent as a specialized functional module for mediating between natural language and data querying in complex operating environments.

These activities ensured that the agent could bridge operators with intelligent models through a natural, context-aware conversational interface capable of interpreting user requests, analyzing historical context, and activating computational tools autonomously to provide relevant and actionable outputs. The agent was developed using the LangChain framework, following the ReAct (Reasoning + Acting) paradigm. This approach combines step-by-step reasoning with concrete actions executed via specialized tools. The cognitive core is based on the Claude 3.5 Sonnet language model, accessed through the Anthropic API and configured with parameters optimized to ensure reliability, coherence, and low randomness. Finally, the system outputs are seamlessly communicated back to the operator through the Front End.

6.3 Integration of the agent in the SPLASH platform

This section introduces the operational prototype of the proposed system. Using annotated screenshots, it illustrates how an operator engages with the virtual agent to analyze network data, diagnose anomalies, and access results in a clear and interpretable manner.

Current Conversation

Your question:
I am at node 880171 ari and I notice an abnormal pressure value—what is happening?

Agent's response:
At node 880171, a critical leak has been detected with a very high probability (99.98%). The situation requires immediate intervention since:

1. The system classifies the risk as CRITICAL.
2. The pressure anomaly you noticed is confirmed by the analysis.
3. The leak appears to be localized specifically at this node, as nearby nodes do not show significant anomalies (the average probability of leakage in neighboring nodes is only 2%).

I recommend immediately activating emergency maintenance procedures for this specific node.

Fig. 20: System response with notification of a critical leak diagnosis in node 880171.

Anomaly detection. Fig. 20 illustrates an initial example of interaction with the AI Agent system. At the center of the interface, a text box enables the operator to submit natural-language requests. In this case, the operator queries node 880171, highlighting an abnormal pressure reading and asking the system to diagnose the underlying cause.

At the bottom of the interface, the operational status of the virtual agent is displayed; in the figure, the agent is active and currently handling the request. This information is conveyed to the user through a descriptive message positioned beneath the question input field. This stage marks the initiation of the system's operational cycle: the natural language query is processed by the orchestrator agent, which interprets its semantics and determines the appropriate handling strategy. Specifically, the orchestrator selects the most relevant computational modules—such as the Graph Neural Network–based leak detection model or the temporal data querying component—to generate an appropriate and prompt response. The workflow is structured to enable a natural and transparent interaction between the user and the decision-support system.

The figure illustrates the system's response to the operator's query about node 880171. After processing sensor data and running the leak detection module, the agent indicates a critical leak at the specified node with a 99.98% probability.

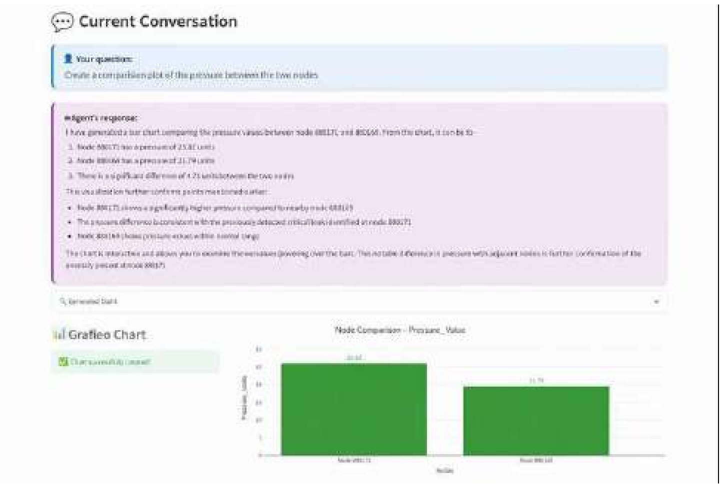
The response provides a detailed explanation:

- The system classifies the risk as CRITICAL, suggesting immediate intervention.
- The pressure anomaly reported by the operator has been confirmed by the automatic analysis.
- The leak is located with high precision in the requested node, as the adjacent nodes do not show any significant anomalies (average probability of leakage in nearby nodes is 2%).

Data visualization module. As part of the data visualization module, Fig. 21 presents a bar chart comparing pressure values at two selected nodes in the water network. Here, the operator requested a comparison between node 880171 (previously classified as critical) and node 880169 (used as a stable reference). The agent then generated a comparative pressure plot with the following outcomes:

- Node 880169: pressure of 26.52 meters. Node 880171: pressure of 21.79 meters.
- The pressure difference between the two nodes is significant, equal to 4.73 meters, which is an anomalous value in the context of a well-balanced water network.

The agent's output highlights that the observed pressure fluctuation aligns with the critical leak previously identified at node 880171, while node 880169 maintains normal values, supporting the hypothesis that the anomaly is localized. The graph at the bottom of the interface provides an immediate visual



Create a comparison plot of the pressure between the two nodes

Fig. 21: Comparison between node 880171 and node 880169.

representation: the bars correspond to the two nodes and report their exact pressure readings, making the information accessible even to staff without expertise in databases or predictive modeling.

7 Conclusions and Next Steps

This chapter has systematically analysed the role of digital transformation in making more resilient, sustainable and intelligent the water distribution systems. The synergistic integration of large-scale IoT technologies, LPWAN communications, advanced hydraulic models, Digital Twins and machine learning algorithms emerges as a key factor in overcoming the limitations of traditional, predominantly reactive approaches and enabling predictive and proactive management strategies.

From a communications infrastructure perspective, the results show that LPWAN technologies now account for approximately 55% of the solutions adopted in WDS monitoring, with LoRaWAN alone covering approximately 37% of the cases analysed. Coexistence analyses conducted in high-density scenarios (up to 1000 devices/km^2, equivalent to over 3000 smart meters in an area with a 1 km radius) show that CSS (LoRaWAN) and UNB (Sigfox) systems can coexist reliably, with interference probabilities of less than 3% for RSS levels above -90 dBm. Furthermore, with the same duty cycle, LoRaWAN achieves significantly higher throughput, up to over 8000 packets/day with SF7, compared to approximately 100 packets/day for UNB, highlighting a clear trade-off between robustness and transmission capacity.

In terms of digital representation, the adoption of advanced Digital Twins, such as the SWIM platform, allows us to go beyond simple passive monitoring.

The integration of real-time IoT data with EPANET/WNTR-based hydraulic simulations and machine learning models enables automatic "what-if" analyses and supports informed operational decisions. Demonstration cases show, for example, the ability to identify localised leaks with a probability greater than 99.9% by associating pressure anomalies of the order of 4-5 m between adjacent nodes, values clearly indicative of critical operating conditions.

The contributions of machine learning highlight further quantitative benefits. In demand characterization, analysis of approximately 1,000 LoRaWAN smart meters with hourly data for an entire year identified four main consumption clusters, distinguishing profiles with morning and evening peaks and providing more informative probabilistic membership degrees than rigid clustering. In the field of forecasting, the introduction of probabilistic models based on diffusion models allows confidence intervals (e.g., 50% and 95%) of future demand to be estimated, improving the management of operational uncertainty compared to traditional deterministic models. Similarly, consumption disaggregation techniques on data with a resolution of less than one minute show a marked improvement in the reconstruction of domestic usage profiles, with UNet models able to track peaks and dynamics more accurately than CNN-LSTM architectures.

Looking ahead, several next steps emerge from this work. First, future research should focus on scaling Digital Twin frameworks to city-wide and multi-DMA deployments, integrating heterogeneous data sources while maintaining real-time responsiveness. Second, tighter coupling between probabilistic demand forecasts, hydraulic simulations, and automated control strategies could further reduce water losses and energy consumption in pumping operations. Third, the integration of intelligent virtual agents and natural language interfaces represents a promising direction to enhance operator interaction, reduce cognitive load, and democratize access to advanced analytics. Finally, long-term field validations are needed to quantitatively assess the cumulative impact of these digital solutions in terms of water savings, energy efficiency, and resilience under extreme events.

By combining scalable IoT infrastructures, high-fidelity Digital Twins, and data-driven intelligence, we believe that future WDSs can evolve toward more sustainable, resilient, and human-centric systems, fully aligned with the principles of Industry 5.0 and the objectives of modern water utilities.

Acknowledgements This work was supported by the Italian Ministry of University and Research (MUR) under the RESTART program.

References

1. Junguo Liu, Hong Yang, Simon N Gosling, Matti Kummu, Martina Flörke, Stephan Pfister, Naota Hanasaki, Yoshihide Wada, Xinxin Zhang, Chunmiao Zheng, et al. Water scarcity assessments in the past, present, and future. *Earth's future*, 5(6):545–559, 2017.
2. UNESCO World Water Assessment Programme. WWAP: World Water Assessment Programme, The United Nations World Water Development Report 2020: Water and Climate Change, 2020. Accessed: 2024-09-23.

3. Francesco Gino Ciliberti, Luigi Berardi, Daniele Biagio Laucelli, and Orazio Giustolisi. Digital transformation paradigm for asset management in water distribution networks. In *2021 10th International conference on energy and environment (CIEM)*, pages 1–5. IEEE, 2021.
4. Antonino Pagano, Domenico Garlisi, Ilenia Tinnirello, Fabrizio Giuliano, Giovanni Garbo, Mariana Falco, and Francesca Cuomo. A survey on massive iot for water distribution systems: Challenges, simulation tools, and guidelines for large-scale deployment. *Ad Hoc Networks*, 168:103714, 2025.
5. Helena M. Ramos, Maria Cristina Morani, Armando Carravetta, Oreste Fecarrotta, Kemi Adeyeye, P. Amparo López-Jiménez, and Modesto Pérez-Sánchez. New Challenges towards Smart Systems' Efficiency by Digital Twin in Water Distribution Networks. *Water*, 14(8):1304, April 2022.
6. Salvatore Cavalieri and Salvatore Gambadoro. Digital Twin of a Water Supply System Using the Asset Administration Shell. *Sensors*, 24(5):1360, February 2024.
7. Giacomo Vittori, Yelizaveta Falkouskaya, Daniel M Jimenez-Gutierrez, Tiziana Cattai, and Ioannis Chatzigiannakis. Graph neural networks to model and optimize the operation of water distribution networks: A review. *Journal of Industrial Information Integration*, page 100880, 2025.
8. Tiziana Cattai, Stefania Colonnese, Domenico Garlisi, Antonino Pagano, and Francesca Cuomo. Graphsmart: a method for green and accurate iot water monitoring. *ACM Transactions on Sensor Networks*, 20(6):1–32, 2024.
9. Lu Xing and Lina Sela. Graph neural networks for state estimation in water distribution systems: Application of supervised and semisupervised learning. *Journal of Water Resources Planning and Management*, 148(5):04022018, 2022.
10. Larry W Mays. Water distribution systems handbook. 2000.
11. Thomas M Walski, Donald V Chase, Dragan A Savic, Walter Grayman, Stephen Beckwith, and Edmundo Koelle. Advanced water distribution modeling and management. 2003.
12. Lewis A Rossman et al. Epanet 2: users manual. 2000.
13. Carol Boyle, Greg Ryan, Pratik Bhandari, Kris MY Law, Jinzhe Gong, and Douglas Creighton. Digital transformation in water organizations. *Journal of Water Resources Planning and Management*, 148(7):03122001, 2022.
14. Antonino Pagano, Domenico Garlisi, Fabrizio Giuliano, Tiziana Cattai, and Francesca Cuomo. Application-aware lorawan gateway placement in massive iot water distribution networks. In *2024 IEEE 10th World Forum on Internet of Things (WF-IoT)*, pages 475–480. IEEE, 2024.
15. Domenico Garlisi, Gabriele Restuccia, Ilenia Tinnirello, Francesca Cuomo, and Ioannis Chatzigiannakis. Leakage detection via edge processing in lorawan-based smart water distribution networks. In *2022 18th International Conference on Mobility, Sensing and Networking (MSN)*, pages 223–230, 2022.
16. Domenico Garlisi, Antonino Pagano, Fabrizio Giuliano, Daniele Croce, and Ilenia Tinnirello. Interference analysis of lorawan and sigfox in large-scale urban iot networks. *IEEE Access*, 2025.
17. Domenico Garlisi, Antonino Pagano, Fabrizio Giuliano, Daniele Croce, and Ilenia Tinnirello. A coexistence study of low-power wide-area networks based on lorawan and sigfox. In *2023 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–7. IEEE, 2023.
18. Nagib Matni, Jean Moraes, Helder Oliveira, Denis Rosário, and Eduardo Cerqueira. Lorawan gateway placement model for dynamic internet of things scenarios. *Sensors*, 20(15):4336, 2020.

19. Julia Robles, Cristian Martín, and Manuel Díaz. OpenTwins: An open-source framework for the development of next-gen compositional digital twins. *Computers in Industry*, 152:104007, November 2023.
20. Zhiwei Shen, Felipe Arrano-Vargas, and Georgios Konstantinou. What-if scenario testing through power system digital twins. In *2024 IEEE 22nd International Conference on Industrial Informatics (INDIN)*, pages 1–6, 2024.
21. Gabriele Restuccia, Domenico Garlisi, Fabrizio Giuliano, Antonino Pagano, and Ilenia Tinnirello. The synergy of digital twins and machine learning in sustainable water management: Swim. *Computer*, 58(12):46–54, 2025.
22. Eclipse Foundation. Eclipse ditto. https://eclipse.dev/ditto/. Accessed: 2024-10-05.
23. Katherine A Klise, Regan Murray, and Terra Haxton. An overview of the water network tool for resilience (wntr). 2018.
24. European Commission. Industry 5.0 - towards a sustainable, human-centric, and resilient european industry, 2023. Accessed: 2024-10-10.
25. Mohammed H Alsharif, Abu Jahid, Anabi Hilary Kelechi, and Raju Kannadasan. Green iot: A review and future research directions. *Symmetry*, 15(3):757, 2023.
26. Carlos A. Bonilla, Ariele Zanfei, Bruno Brentan, Idel Montalvo, and Joaquín Izquierdo. A Digital Twin of a Water Distribution System by Using Graph Convolutional Networks for Pump Speed-Based State Estimation. *Water*, 14(4):514, February 2022.
27. B. Bollobas. *Modern Graph Theory*. Springer-Verlag, 1998.
28. Garðar Örn Garðarsson, Francesca Boem, and Laura Toni. Graph-based learning for leak detection and localisation in water distribution networks. *IFAC-PapersOnLine*, 55(6):661–666, 2022.
29. Adrià Soldevila, Joaquim Blesa, Sebastian Tornil-Sin, Rosa M. Fernandez-Canti, and Vicenç Puig. Sensor placement for classifier-based leak localization in water distribution networks using hybrid feature selection. 108:152–162.
30. Sen Peng, Jing Cheng, Xingqi Wu, Xu Fang, and Qing Wu. Pressure sensor placement in water supply network based on graph neural network clustering method. *Water*, 14(2):150, 2022.
31. Gustavo Batista, Eamonn Keogh, Oben Tataw, and Vinícius Alves de Souza. Cid: An efficient complexity-invariant distance for time series. *Data Mining and Knowledge Discovery*, 28, 04 2013.
32. Michael R Berthold and Frank Höppner. On clustering time series using euclidean distance and pearson correlation. *arXiv preprint arXiv:1601.02213*, 2016.
33. Pavel Senin. Dynamic time warping algorithm review. *Information and Computer Science Department University of Hawaii at Manoa Honolulu, USA*, 855(1-23):40, 2008.
34. Eamonn Keogh, Stefano Lonardi, Chotirat Ratanamahatana, Li Wei, Sang-Hee Lee, and John Handley. Compression-based data mining of sequential data. *Data Mining and Knowledge Discovery*, 14:99–129, 02 2007.
35. Lowell W. Beineke. Characterizations of derived graphs. *Journal of Combinatorial Theory*, 9(2):129–135, 1970.
36. Antonio Ortega, Pascal Frossard, Jelena Kovačević, José MF Moura, and Pierre Vandergheynst. Graph signal processing: Overview, challenges, and applications. *Proceedings of the IEEE*, 106(5):808–828, 2018.
37. Elena Ceci and Sergio Barbarossa. Graph signal processing in the presence of topology uncertainties. *IEEE Transactions on Signal Processing*, 68:1558–1573, 2020.

38. Francisco Aparecido Rodrigues. Network centrality: an introduction, 2019.
39. Qiying Feng, Long Chen, CL Philip Chen, and Li Guo. Deep fuzzy clustering—a representation learning approach. *IEEE Transactions on Fuzzy Systems*, 28(7):1420–1433, 2020.
40. Redemptor Jr Laceda Taloma, Francesca Cuomo, Danilo Comminiello, and Patrizio Pisani. Machine learning for smart water distribution systems: exploring applications, challenges and future perspectives. *Artificial Intelligence Review*, 58(4):120, 2025.
41. Redemptor Jr Laceda Taloma. Deep temporal modeling of high-resolution time series for water demand analysis and management. 2025.
42. Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.
43. Stefano Alvisi, Marco Franchini, Valentina Marsili, Filippo Mazzoni, Elad Salomons, Mashor Housh, Ahmed Abokifa, Kristina Arsova, Faten Ayyash, Hyansu Bae, et al. Battle of water demand forecasting. *Journal of Water Resources Planning and Management*, 151(10):04025049, 2025.
44. Redemptor Jr Laceda Taloma, Danilo Comminiello, Patrizio Pisani, and Francesca Cuomo. Unet-WD: Deep learning for multi-appliance water disaggregation. In *2024 IFIP Networking Conference (IFIP Networking)*, 2024.
45. Anna Di Mauro, Andrea Cominola, Andrea Castelletti, and Armando Di Nardo. Urban water consumption at multiple spatial and temporal scales. a review of existing datasets. *Water*, 13(1):36, 2021.
46. Domenico Garlisi, Alessio Martino, Jad Zouwayhed, Reza Pourrahim, and Francesca Cuomo. Exploratory approach for network behavior clustering in lorawan. *Journal of Ambient Intelligence and Humanized Computing*, 14(12):15745–15759, Dec 2023.
47. Siqi et al. Wang. Multiview deep anomaly detection: A systematic exploration. *IEEE Transactions on Neural Networks and Learning Systems*, 35(2):1651–1665, 2024.
48. Ferran Adelantado, Xavier Vilajosana, Pere Tuset-Peiro, Borja Martinez, Joan Melia-Segui, and Thomas Watteyne. Understanding the limits of lorawan. *IEEE Communications magazine*, 55(9):34–40, 2017.
49. Pietro Spadaccino, Francesco Giuseppe Crinó, and Francesca Cuomo. Lorawan behaviour analysis through dataset traffic investigation. *Sensors*, 22(7), 2022.
50. Giovanni Valtorta, Dario Bonino, and Mario Alessi. A clustering approach for profiling lorawan iot devices. In *2021 IEEE International Conference on Smart Internet of Things (SmartIoT)*, pages 125–132. IEEE, 2021.
51. Y. Goldshtein and altri. Llm-epanet: A natural language interface for hydraulic simulation. *Water Resources Research*, 2025.