ASCE

Battle of Water Demand Forecasting

S. Alvisi, Ph.D.[1]; M. Franchini[2]; V. Marsili, Ph.D.[3]; F. Mazzoni, Ph.D.[4]; E. Salomons, Ph.D.[5]; M. Housh, Ph.D.[6]; A. Abokifa, Ph.D.[7]; K. Arsova[8]; F. Ayyash[9]; H. Bae, Ph.D.[10]; R. Barreira, Ph.D.[11]; L. Basto[12]; S. Bayer, Ph.D.[13]; E. Z. Berglund, Ph.D.[14]; D. Biondi, Ph.D.[15]; F. Boloukasli Ahmadgourabi[16]; B. Brentan, Ph.D.[17]; J. Caetano[18]; F. Campos[19]; H. Cao, Ph.D.[20]; S. Cardona[21]; E. P. Carreño Alvarado, Ph.D.[22]; N. Carriço, Ph.D.[23]; G.-A. Chatzistefanou[24]; Y. Coy[25]; E. Creaco, Ph.D.[26]; S. Cuomo, Ph.D.[27]; A. de Klerk[28]; A. Di Nardo, Ph.D.[29]; M. DiCarlo, Ph.D.[30]; U. Dittmer, Ph.D.[31]; R. Dziedzic, Ph.D.[32]; A. Ebrahim Bakhshipour, Ph.D.[33]; D. Eliades, Ph.D.[34]; R. Farmani, Ph.D.[35]; B. Ferreira, Ph.D.[36]; A. Gabriele, Ph.D.[37]; M. M. Gamboa-Medina, Ph.D.[38]; F. Gao[39]; J. Gao, Ph.D.[40]; R. Gargano, Ph.D.[41]; M. Geranmehr, Ph.D.[42]; C. Giudicianni, Ph.D.[43]; K. Glynis[44]; S. Gómez[45]; L. González[46]; M. Groß[47]; H. Guo[48]; M. N. Habibi[49]; A. Haghighi, Ph.D.[50]; B. Hammer, Ph.D.[51]; L. Hans[52]; M. Hayslep[53]; Y. He[54]; L. Hermes[55]; M. Herrera, Ph.D.[56]; F. Hinder[57]; B. Hou, Ph.D.[58]; A. Iglesias-Rey[59]; P. L. Iglesias-Rey, Ph.D.[60]; I.-S. Jang, Ph.D.[61]; J. Izquierdo, Ph.D.[62]; M. S. Jahangir, Ph.D.[63]; C. Jara-Arriagada[64]; B. Jenks[65]; G. Johnen[66]; M. Kalami Heris, Ph.D.[67]; M. Kalumba, Ph.D.[68]; M.-S. Kang, Ph.D.[69]; M. Khashei Varnamkhasti[70]; K.-J. Kim, Ph.D.[71]; J. Kley-Holsteg[72]; T. Ko[73]; A. Koochali, Ph.D.[74]; P. Kossieris, Ph.D.[75]; P. Koundouri, Ph.D.[76]; C. Kühnert, Ph.D.[77]; A. Kulaczkowski[78]; J. Lee, Ph.D., P.E.[79]; K. Li[80]; Y. Li[81]; H. Liu, Ph.D.[82]; Y. Liu[83]; C. A. López-Hojas[84]; A. Maier, Ph.D.[85]; C. Makropoulos, Ph.D.[86]; F. J. Martínez-Solano, Ph.D.[87]; N. H. Marzouny[88]; A. Menapace, Ph.D.[89]; C. Michalopoulos[90]; G. Moraitis[91]; H. Mousa[92]; H. Namdari[93]; D. Nikolopoulos, Ph.D.[94]; M. Oberascher, Ph.D.[95]; A. Ostfeld, Ph.D.[96]; M. Pagano[97]; F. Pasha, Ph.D.[98]; J. Perafán[99]; G. Perelman[100]; J. Pesantez, Ph.D.[101]; M. Polycarpou, Ph.D.[102]; M. G. Quarta[103]; Q. Que[104]; J. Quilty, Ph.D.[105]; C. Quintiliani, Ph.D.[106]; A. Ramachandran[107]; G. Reynoso Meza, Ph.D.[108]; V. Rodriguez[109]; Y. Romano, Ph.D.[110]; J. Saldarriaga, Ph.D.[111]; A. K. Salem[112]; P. Samartzis, Ph.D.[113]; G. F. Santonastaso, Ph.D.[114]; D. Savic, Ph.D.[115]; V. Schiano Di Cola, Ph.D.[116]; D. Schol, Ph.D.[117]; A. G. Seyoum, Ph.D.[118]; R. Shen[119]; K. Simukonda, Ph.D.[120]; A. Sinske, Ph.D.[121]; R. Sitzenfrei, Ph.D.[122]; B. Sonnenschein[123]; I. Stoianov, Ph.D.[124]; A. Tabares, Ph.D.[125]; E. Todini, Ph.D.[126]; L. Tsiami[127]; I. Tsoukalas, Ph.D.[128]; A.-J. Ulusoy, Ph.D.[129]; L. Vamvakeridou-Lyroudia, Ph.D.[130]; A. van Heerden[131]; J. Vaquet[132]; V. Vaquet[133]; S. Wallner[134]; M. Walraad[135]; D. Wang[136]; S. Wu, Ph.D.[137]; W. Wu, Ph.D.[138]; A. Wunsch, Ph.D.[139]; Y. Yao[140]; J. Yu[141]; A. Zanfei, Ph.D.[142]; D. Zanutto[143]; H. Zhang[144]; M. Ziebarth, Ph.D.[145]; F. Ziel, Ph.D.[146]; and J. Zou[147]

Abstract: As part of the Battle of Water Networks competition series, the Battle of Water Demand Forecasting (BWDF) was organized in the context of the 3rd Water Distribution Systems Analysis and Computing and Control in the Water Industry (WDSA-CCWI) joint conference held in Ferrara (Italy) in 2024. In line with the previous editions of the Battle of Water Networks—the main objective of which was to address a specific problem related to the design and operation of water distribution networks—the BWDF aims to compare the effectiveness of methods for the short-term forecast of urban water demand in a set of real district metered areas. During the conference, 31 teams across the world participated in the BWDF and presented their approaches. The results obtained demonstrate the importance of (1) considering integrated approaches for short-term water demand forecasting; and (2) evaluating their performance in relation to more than one metric, case study, and period. **DOI: 10.1061/JWRMD5.WRENG-6887.**

Author keywords: Water demand; Forecasting; Water distribution network; Field data; Time series analysis; Artificial intelligence.

[1]Associate Professor, Univ. of Ferrara.
[2]Professor, Univ. of Ferrara.
[3]Postdoctoral Researcher, Univ. of Ferrara.
[4]Postdoctoral Researcher, Univ. of Ferrara (corresponding author). Email: filippo.mazzoni@unife.it
[5]Water Resources Consultant, Optiwater.
[6]Associate Professor, Univ. of Haifa.
[7]Assistant Professor, Univ. of Illinois Chicago.
[8]Data Analyst, Brabant Water.
[9]Ph.D. Candidate, Univ. of Exeter.
[10]Researcher, Electronics and Telecommunications Research Institute.
[11]Associate Professor, Instituto Politécnico de Setúbal.

Note. This manuscript was submitted on October 18, 2024; approved on April 1, 2025; published online on July 18, 2025. Discussion period open until December 18, 2025; separate discussions must be submitted for individual papers. This paper is part of the *Journal of Water Resources Planning and Management*, © ASCE, ISSN 0733-9496.

Introduction

The Battle of the Water Demand Forecasting (BWDF) is part of the Battle of Water Networks, a series of competitions related to the design and operation of water distribution networks (WDNs) that dates back to the 1980s (Walski et al. 1987). Since the beginning, the competition's goal has been to attract groups with different backgrounds, such as academic researchers, engineers and technicians from water utilities, and consultants, to propose strategies for solving complex problems concerning WDNs under various conditions.

[12]Researcher, Univ. of the Andes.
[13]Researcher, Friedrich Alexander Univ.
[14]Professor, North Carolina State Univ.
[15]Associate Professor, Univ. of Calabria.
[16]Ph.D. Candidate, Concordia Univ.
[17]Associate Professor, Federal Univ. of Minas Gerais.
[18]Ph.D. Candidate, Instituto Politécnico de Setúbal.
[19]Master's Candidate, Univ. of São Paulo.
[20]Professor, Harbin Institute of Technology.
[21]Researcher, Univ. of the Andes.
[22]Postdoctoral Researcher, Federal Univ. of Paraná.
[23]Assistant Professor, Instituto Politécnico de Setúbal.
[24]Ph.D. Candidate, Univ. of Exeter.
[25]Researcher, Univ. of the Andes.
[26]Professor, Univ. of Pavia.
[27]Professor, Univ. of Naples Federico II.
[28]Senior Engineer, GLS Consulting.
[29]Associate Professor, Univ. of Campania Luigi Vanvitelli.
[30]American Association for the Advancement of Science Fellowship Programs.
[31]Professor, Univ. of Kaiserslautern-Landau.
[32]Assistant Professor, Concordia Univ.
[33]Postdoctoral Researcher, Univ. of Kaiserslautern-Landau.
[34]Assistant Professor, Univ. of Cyprus.
[35]Professor, Univ. of Exeter.
[36]Visiting Assistant Professor, Instituto Politécnico de Setúbal.
[37]Researcher, Univ. of Cassino and Southern Lazio.
[38]Assistant Professor, Univ. of São Paulo.
[39]Researcher, Dalian Univ. of Technology.
[40]Associate Professor, Harbin Institute of Technology.
[41]Associate Professor, Univ. of Cassino and Southern Lazio.
[42]Research Associate, Univ. of Sheffield.
[43]Assistant Professor, Univ. of Pavia.
[44]Ph.D. Candidate, KWR Water Research Institute.
[45]Researcher, Univ. of the Andes.
[46]Ph.D. Candidate, Univ. of the Andes.
[47]Ph.D. Candidate, Cologne Univ. of Applied Sciences.
[48]Researcher, Dalian Univ. of Technology.
[49]Research Assistant, Concordia Univ.
[50]Professor, Shahid Chamran Univ. of Ahvaz.
[51]Professor, Bielefeld Univ.
[52]Ph.D. Candidate, Cologne Univ. of Applied Sciences.
[53]Research Fellow, Univ. of Exeter.
[54]Ph.D. Candidate, Harbin Institute of Technology.
[55]Ph.D. Candidate, Bielefeld Univ.
[56]Lecturer, Newcastle Univ.
[57]Ph.D. Candidate, Bielefeld Univ.
[58]Associate Professor, Beijing Univ. of Technology.
[59]Customer Advisor, SAS Institute.
[60]Professor, Universitat Politècnica de València.
[61]Researcher, Electronics and Telecommunications Research Institute.
[62]Professor, Universitat Politècnica de València.
[63]Postdoctoral Researcher, Univ. of Waterloo.
[64]Ph.D. Candidate, Imperial College London.
[65]Ph.D. Candidate, Imperial College London.
[66]Research Assistant, Univ. of Duisburg-Essen.
[67]Assistant Professor, Sheffield Hallam Univ.
[68]Lecturer, Univ. of Zambia.
[69]Researcher, Electronics and Telecommunications Research Institute.
[70]Ph.D. Candidate, Concordia Univ.
[71]Researcher, Electronics and Telecommunications Research Institute.
[72]Ph.D. Candidate, Univ. of Duisburg-Essen.
[73]Ph.D. Candidate, Univ. of Exeter.
[74]Research Associate, German Research Center for Artificial Intelligence.
[75]Postdoctoral Researcher, National Technical Univ. of Athens.
[76]Professor, Athens Univ. of Economics and Business.
[77]Senior Scientist, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation.
[78]Environmental Engineer, Manhattan College.
[79]Professor, Manhattan College.
[80]Ph.D. Candidate, Harbin Institute of Technology.
[81]Master's Candidate, Beijing Univ. of Technology.
[82]Professor, Dalian Univ. of Technology.
[83]Ph.D. Candidate, Imperial College London.
[84]Master's Candidate, Universitat Politècnica de València.
[85]Professor, Friedrich Alexander Univ.
[86]Professor, National Technical Univ. of Athens.
[87]Associate Professor, Universitat Politècnica de València.
[88]Ph.D. Candidate, Concordia Univ.
[89]Researcher, Free Univ. of Bozen-Bolzano.
[90]Ph.D. Candidate, National Technical Univ. of Athens.
[91]Ph.D. Candidate, National Technical Univ. of Athens.
[92]Researcher, Friedrich Alexander Univ.
[93]Ph.D. Candidate, Shahid Chamran Univ. of Ahvaz.
[94]Postdoctoral Researcher, National Technical Univ. of Athens.
[95]Postdoctoral Researcher, Univ. of Innsbruck.
[96]Professor, Technion–Israel Institute of Technology.
[97]Research Fellow, Univ. of Campania Luigi Vanvitelli.
[98]Professor, California State Univ. Fresno.
[99]Researcher, Univ. of the Andes.
[100]Ph.D. Candidate, Technion–Israel Institute of Technology.
[101]Assistant Professor, California State Univ. Fresno.
[102]Professor, Univ. of Cyprus.
[103]Ph.D. Candidate, Univ. of Salento.
[104]Master's Candidate, Harbin Institute of Technology.
[105]Assistant Professor, Univ. of Waterloo.
[106]Data Analyst, Brabant Water.
[107]Ph.D. Candidate, Friedrich Alexander Univ.
[108]Associate Professor, Pontifical Catholic Univ. of Paraná.
[109]Researcher, Univ. of the Andes.
[110]Assistant Professor, Technion–Israel Institute of Technology.
[111]Professor, Univ. of the Andes.
[112]Ph.D. Candidate, Univ. of Illinois Chicago.
[113]Assistant Professor, Univ. of Macedonia.
[114]Associate Professor, Univ. of Campania Luigi Vanvitelli.
[115]Professor, KWR Water Research Institute.
[116]Data Analyst, Cyberneid.
[117]Data Analyst, Brabant Water.
[118]Research Associate, Univ. of Sheffield.
[119]Master's Candidate, Harbin Institute of Technology.
[120]Lecturer, Univ. of Zambia.
[121]Head of Technology, GLS Consulting.
[122]Professor, Univ. of Innsbruck.
[123]Research Assistant, Univ. of Duisburg-Essen.
[124]Reader, Imperial College London.
[125]Assistant Professor, Univ. of the Andes.
[126]Honorary President, Italian Hydrological Society.
[127]Ph.D. Candidate, KWR Water Research Institute.
[128]Assistant Professor, Democritus Univ. of Thrace.
[129]Research Fellow, Imperial College London.
[130]Senior Researcher, Univ. of Exeter.
[131]Senior Engineer, GLS Consulting.
[132]Master's Candidate, Bielefeld Univ.
[133]Ph.D. Candidate, Bielefeld Univ.
[134]Scientist, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation.
[135]Data Analyst, Brabant Water.
[136]Master's Candidate, Beijing Univ. of Technology.
[137]Associate Professor, Beijing Univ. of Technology.
[138]Professor, Birmingham City Univ.
[139]Scientist, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation.
[140]Ph.D. Candidate, Dalian Univ. of Technology.
[141]Researcher, Electronics and Telecommunications Research Institute.
[142]Environmental Engineer, AIAQUA.
[143]Ph.D. Candidate, Politecnico di Milano.
[144]Master's Candidate, Harbin Institute of Technology.
[145]Scientist, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation.
[146]Professor, Univ. of Duisburg-Essen.
[147]Bachelor's Candidate, Dalian Univ. of Technology.

The BWDF, organized in the context of the 3rd International Water Distribution Systems Analysis and Computing and Control in the Water Industry (WDSA-CCWI) joint conference held in Ferrara (Italy) in 2024, is aimed at comparing the effectiveness of methods for the short-term forecast of urban water demand considering a set of district metered areas (DMAs) in a real WDN. This challenge attracted 31 teams from different backgrounds worldwide who participated in the competition. The approaches applied, and the results obtained, by each team were presented during the plenary session of the WDSA-CCWI 2024 conference dedicated to the BWDF. In this plenary session, the winning team was announced, along with second and third place.

After introducing the BWDF problem and summarizing the approaches adopted by the 31 participating teams, this paper compares the results submitted by each team and provides considerations and insights for future research.

Background

Population growth, urbanization, and climate change have been raising people's awareness about the impact of human activities on the environment and the available natural sources, such as water resources (Daniel et al. 2018; Jeandron et al. 2019; Attallah et al. 2023). In this context, sustainable management of water systems is crucial to avoid water shortage or the depletion of available sources. The operational and strategic decisions made by drinking water utilities and authorities can benefit from reliable and accurate forecasts of water demand—i.e., the sum of all the types of water consumption within WDNs, spacing from residential to nonresidential activities (Mazzoni et al. 2024)—which is the main driver of these systems (Hussien et al. 2016; Grespan et al. 2022; Li and Song 2023). Water demand and, strictly related, the total net inflow—i.e., the total amount of water entering a WDN, including all types of water consumption and leakages—can vary significantly due to factors affecting user behavior, spacing from weather to socio-economic conditions or the type of day (Zounemat-Kermani et al. 2020), making its forecast a challenging task.

Water demand forecasting can be performed at different time scales (Bakker et al. 2013; Zounemat-Kermani et al. 2020; Hao et al. 2024), i.e., with different time resolutions (the time steps at which water demand forecasts are generated) and forecast horizons (the forecasted time intervals). In general, water demand forecasting models can be classified as *long-term* and *short-term* models, according to the levels of planning in relation to which they are used (Donkor et al. 2014; Ghalehkhondabi et al. 2017; Pacchin et al. 2019). On the one hand, long-term models are typically used by water utilities in decision-making processes for urban water management in the distant future (e.g., upcoming years or decades), such as those related to design, water pricing, resource allocation, or water use restrictions (Herrera et al. 2010; Babel and Shinde 2011; Donkor et al. 2014). Long-term models generally provide demand forecasts on a monthly or yearly basis with a time horizon ranging from 1 to 10 years, and they are used for offline simulations to evaluate different planning options. On the other hand, short-term models allow drinking water utilities to estimate water demand and its pattern over a limited time horizon. Short-term models are generally used for management or real-time control purposes (Odan and Ribeiro Reis 2012), e.g., to appropriately regulate pumps, valves, and other network elements (Kley-Holsteg and Ziel 2020). Short-term models typically provide water demand forecasts over time horizons ranging from 1 day to 1 month with a time step ranging from daily to subhourly (Arandia et al. 2016; Bakker et al. 2013; Msiza et al. 2008; Shabani et al. 2018).

The BWDF focuses on short-term forecasting. With specific reference to daily and weekly forecast horizons, various methods have been proposed and developed in the scientific literature in recent years (Tian et al. 2016), and promising results have been obtained (Guo et al. 2018). However, available methods have been generally tested on specific case studies, and comparing their performance is not straightforward (Ghalehkhondabi et al. 2017).

Forecasting Methods

Approaches for the generation of water demand forecasts can be classified into *soft computing* methods and *time series analysis* methods (Pacchin et al. 2019), whereas approaches including the combination of two or many water demand forecasting techniques are defined as *hybrid* methods (Niknam et al. 2022).

Soft computing methods for water demand forecasting are based on the use of emerging data-driven tools that produce approximate solutions to high-complexity problems (de Souza Groppo et al. 2019). In greater detail, these data-driven approaches can forecast water demand based on historical demand data, possibly coupled with data related to *exogenous variables*, i.e., the potential drivers of water demand (spacing from weather to socio-economic factors, or related to the calendar). Clearly, selecting the most relevant exogenous variables can contribute to improving model accuracy while limiting its dimensionality (Hao et al. 2024).

Among soft computing methods, artificial neural networks (ANNs) are extensively used for forecasting purposes in various domains and in many occasions are shown to have excellent predictive ability (Ghalehkhondabi et al. 2017). The ability of ANNs to model nonlinear dynamics is a key advantage over other methods (Hao et al. 2024). By mimicking the function of biological neurons, ANNs can provide an accurate prediction in the face of little or no prior knowledge of the problem. This performance can be significantly enhanced after adapting the model to the specific problem. Among ANNs—and, specifically, deep learning methods—approaches relying on long-short-term memory networks have emerged as promising tools in the field of urban water demand (Hao et al. 2024), better suited for time series forecasts compared to traditional ANNs due to their ability to preserve previous information in the learning process (Mu et al. 2020). Despite their advantages, ANN-based models may face issues related to their lack of explainability, limitations in extrapolating data, and the need for long time series of data for training and testing (Zounemat-Kermani et al. 2020; Niknam et al. 2022).

Other soft computing (OSC) methods for short-term water demand forecasting are available beyond ANNs. For example, among the several machine learning models, models based on decision trees—such as support vector machines and random forests—emerge as the most promising in the literature on water demand forecasting (Niknam et al. 2022). In addition, fuzzy logic reproduces the human way of thinking in computational form and is generally applied to overcome data uncertainty and the statistical assumptions of linearity and time invariance (Palomero et al. 2022).

Time series analysis (TS) methods are based on the analysis of historical data, which is conducted by decomposing their main statistical elements, such as level, trend, seasonality, and noise (Alvisi et al. 2007). Although TS models struggle to capture nonlinear relationships amongst data, a clear advantage of TS for water demand forecasting lies in their great explainability, which cannot be overlooked in the decision-making process (Niknam et al. 2022). Two main categories can be identified among TS methods: univariate TS and TS with exogenous variables (Anele et al. 2017). On the one hand, univariate TS methods forecast future water demand based on past observations and associated error terms.

Examples of univariate TS forecasting include exponential smoothing and the autoregressive integrated moving average (ARIMA) class of models. These methods may not be the most accurate alternative when variations in weather conditions are likely to influence the underlying determinants of water demand (Anele et al. 2017). On the other hand, TS forecasting models with exogenous variables can generate forecasts based on the relationship between water demand and its drivers by including these variables (e.g., weather factors, socio-economic factors, or information about the type of day) as inputs. Notable examples of TS regression models are multiple linear and nonlinear regression and ARIMA with exogenous variables (Anele et al. 2017). Overall, the quality and reliability of the exogenous variable data to input in the TS methods for water demand forecasting are crucial as far as prediction accuracy is concerned (Sardinha-Lourenço et al. 2018).

Finally, hybrid methods combine two or more different models to take advantage of the strengths of each technique, with the aim of outperforming the different methods when used separately (Niknam et al. 2022). In the case of the BWDF, every possible combination of forecasting methods (even if belonging to the same category, e.g., two ANN models) is considered as a hybrid method.

Open Questions and Aim of the Challenge

Based on the literature, a large set of tools for water demand forecasting is currently available. However, the selection of the most suitable technique to adopt can be hard (Herrera et al. 2010), since it generally depends on the focus of a given water demand forecasting problem, e.g., prediction of the average or peak demand (Niknam et al. 2022). Therefore, the results may differ based on the methods used, making it difficult to identify a single method as the overall best (Donkor et al. 2014; Ghalehkhondabi et al. 2017). In general, companies and water utilities prefer interpretable (i.e., transparent) methods instead of complex ones (e.g., soft computing methods) when similar predictive capabilities are obtained by both methods. This is because interpretable methods are more straightforward to include in the decision-making process for a range of WDN operations (Niknam et al. 2022). However, the increasing number of soft computing methods tested in the field of water demand forecasting—including ANN-based models (e.g., Mu et al. 2020; Menapace et al. 2021; Zanfei et al. 2022b), but also the application of OSC techniques (e.g., Liu et al. 2023)—is opening a promising research direction (Guo et al. 2018; Xenochristou and Kapelan 2020). Furthermore, available methods have been mostly trained and validated on specific case studies and considering a variety of different metrics and indicators, often resulting in outputs hard to cross-compare (Zounemat-Kermani et al. 2020). As a result, water demand forecasting remains a research problem, leaving room for improvement (de Souza Groppo et al. 2019).

Considering the above, the BWDF aims to compare the effectiveness of different methods for the short-term urban water demand forecasting using a benchmark data set from the northeast of Italy. This data set is based on supervisory control and data acquisition (SCADA) measurements and mass-balance calculations for a set of real DMAs, i.e., delimited zones within a WDN where boundaries are defined and the quantities of water entering and leaving the area are metered (Pesantez et al. 2020; Sharma et al. 2022). This paper summarizes the main approaches and solutions proposed in the context of the BWDF, along with the key outcomes, and addresses future research directions in the context of short-term water demand forecasting. The rest of the paper is structured as follows: (1) the BWDF problem, data, and assessment criteria are described in the Problem Formulation section; (2) an overview of the forecasting methods presented by participating teams is provided in the Competing Methods section; (3) the main results performance of different methods is outlined in the Results and Discussion section; and (4) key findings and future research directions are discussed in the Conclusions section.

Problem Formulation

Scope and Materials

The goal of the BWDF is to develop new models—or to apply one or many existing models—to perform water demand forecasting on a weekly time horizon. More precisely, the BWDF is focused on forecasting hourly water demand of a real WDN located in the northeast of Italy. The WDN includes 10 DMAs that have different sizes, features, and average water demands. Participants of the BWDF were required to forecast the net inflow for all 10 DMAs, here assumed to represent water demand.

The main features of each DMA are summarized in Table 1, which includes area characteristics, the number of users supplied, and the average net inflow.

The water utility managing the DMAs provided hourly net inflow time series Q_{net} (L/s) for each DMA from January 1, 2021, to March 31, 2023. Flow data were acquired by means of the water utility SCADA systems. Net inflow time series include water consumption and leakage and are obtained through a water balance as shown in Eq. (1):

$$Q_{\text{net}} = \sum_{i=1}^{n_{\text{in}}} Q_{\text{in},i} - \sum_{j=1}^{n_{\text{out}}} Q_{\text{out},j} \tag{1}$$

Table 1. DMA characteristics

DMA	Area characteristics	Number of users supplied (approx.)	Average net inflow (L/s)[a]
1	Hospital district	200	8.2
2	Residential district in the countryside	500	9.6
3	Residential district in the countryside	600	4.2
4	Suburban residential/commercial district	2,100	32.9
5	Residential/commercial district close to the city center	800	78.7
6	Suburban district including sport facilities and office buildings	1,100	8.4
7	Residential district close to the city center	3,200	25.2
8	City center district	2,900	21.2
9	Commercial/industrial district close to the port	400	21.1
10	Commercial/industrial district close to the port	800	26.1

[a]Assessed with reference to the period between January 1, 2021, and March 31, 2023.

J. Water Resour. Plann. Manage., 2025, 151(10): 04025049

in which $Q_{\mathrm{in},i}$ = flow rate entering the DMA through the ith inlet point ($i = 1, 2, \ldots n_{\mathrm{in}}$, n_{in} being the total number of inflow points), whereas $Q_{\mathrm{out},j}$ is the flow rate outgoing from the DMA through the jth outlet point ($j = 1, 2, \ldots n_{\mathrm{out}}$, n_{out} being the total number of outflow points). DMAs with storage facilities were not considered in the BWDF. Moreover, net inflow data were not processed. Therefore, they may include gaps related to SCADA system malfunctioning and other data collection/transmission issues. In particular, missing data constitute about 5% of the data. The small amount of missing data was believed to be acceptable for the purposes of the BWDF.

In addition to the historical net inflow data, weather data such as air temperature, rainfall depth, air humidity, and wind speed observed at a weather station located within the case-study WDN were made available from January 1, 2021, to March 31, 2023. Participants were allowed to use weather data as exogenous variables (although it was not a requirement). Finally, calendar information was provided as water demand patterns may sometimes deviate from those typically observed during working days (Mazzoni et al. 2024). Calendar information provided to BWDF participants included Sundays, holidays, and local-event days along with all time changes from Central European Time to Central European Summer Time and vice-versa. All historical and calendar data are available in the Supplemental Materials.

In the context of the BWDF, teams were asked to deterministically forecast the hourly net inflow of the 10 DMAs over four distinct weeks of the period from January 1, 2021, and March 31, 2023 (hereinafter referred to as *evaluation weeks*, details of which are shown in Table 2). Forecasts were asked, in turn, for each evaluation week, providing: (1) only historical data (i.e., demand and weather data from week January 1, 2021, to the day preceding the evaluation week considered); and (2) weather forecasts for the evaluation week considered (i.e., weather data observed during each evaluation week, excluding demand data for prior evaluation weeks, i.e., problem solutions), so that participants could use this information in their forecasting models. As far as weather data are concerned, these were provided under the assumption that they represent a *perfect* forecast of future weather conditions. In a real-world scenario, these *perfect* weather forecasts would be replaced by other methods, such as numerical weather predictions (e.g., Tian et al. 2016). Overall, the above step-by-step data release was conducted with the aim of reproducing a real time process of water demand forecasting, in which only historical data and the weather forecasts for the period concerned are available. Specifically,

- Teams were first provided with historical demand and weather data from week 1/2021 to week 29/2022 to set up the forecasting models. In addition, weather forecasts for week 30/2022 were made available, and teams were asked to forecast the hourly net inflow time series of the 10 DMAs for week 30/2022 (*evaluation week W*1) and to submit their solution by the first deadline (see Fig. S1a and Table 2).
- After the first deadline, historical demand and weather data from week 31/2022 to week 43/2022 were made available along with weather forecasts for week 44/2022, and teams were asked to forecast the hourly net inflow time series of the 10 DMAs for week 44/2022 (*evaluation week W*2) and to submit their solution by the second deadline (see Fig. S1b and Table 2).

Table 2. Evaluation weeks

Evaluation week	Week number	Initial–final day
W1	30/2022	July 25–31, 2022
W2	44/2022	October 31–November 6, 2022
W3	3/2023	January 16–22, 2023
W4	10/2023	March 6–12, 2023

- After the second deadline, historical demand and weather data for the period between week 45/2022 and week 2/2023 were then made available along with weather forecasts for week 3/2023, and teams were asked to forecast the hourly net inflow time series of the 10 DMAs for week 3/2023 (evaluation week *W*3) and to submit their solution by the third deadline (see Fig. S1c and Table 2).
- After the third deadline, historical demand and weather data for the period between week 4/2023 and week 9/2023 were finally made available along with weather forecasts for week 10/2023, and teams were asked to forecast the hourly net inflow time series of the 10 DMAs for week 10/2023 (evaluation week *W*4) and to submit their solution by the last deadline (see Fig. S1d and Table 2).

Assessment Criteria

The performance of each water demand forecasting model is evaluated by considering its accuracy in forecasting the hourly water demand time series in relation to (1) the first day of each evaluation week (i.e., the first 24 h of the weekly time window), given that WDN operations—such as pumping control strategies—are typically programmed on a daily basis over the next 24 h (Alvisi et al. 2007); and (2) the subsequent period of each evaluation week, since the different behaviors observable on a weekly scale—e.g., between weekdays and holidays—may still affect WDN operational controls on a longer term (Bakker et al. 2013; Sardinha-Lourenço et al. 2018).

From an operational standpoint, model performance is assessed by considering the forecasted net inflow time series of each DMA d ($d = 1, \ldots, D$, where $D = 10$) and for each evaluation week w ($w = 1, \ldots, W$ where $W = 4$). Three performance indicators (PIs) were used in the evaluation:

1. the mean absolute error (MAE) for the 24 h making up to the initial day of each evaluation week w [i.e., $PI1_w^d$, as shown in Eq. (2)];
2. the MAE for the period between the second and the final day of each evaluation week [i.e., $PI2_w^d$, as shown in Eq. (3)]; and
3. the maximum absolute error for the 24 h making up to the initial day of each evaluation week w [i.e., $PI3_w^d$, as shown in Eq. (4)]

$$\mathrm{PI1}_w^d = \frac{1}{24}\sum_{h=1}^{24}|O_{d,w,h} - F_{d,w,h}| \tag{2}$$

$$\mathrm{PI2}_w^d = \frac{1}{144}\sum_{h=25}^{168}|O_{d,w,h} - F_{d,w,h}| \tag{3}$$

$$\mathrm{PI3}_w^d = \max\{|O_{d,w,1} - F_{d,w,1}|, |O_{d,w,2} - F_{d,w,2}|, \ldots, |O_{d,w,24} - F_{d,w,24}|\} \tag{4}$$

In Eqs. (2)–(4), $O_{d,w,h}$ = observed net inflow in DMA d at hour h of evaluation week w ($h = 1, \ldots, 168$); and $F_{d,w,h}$ = respective forecasted value, i.e., forecasted net inflow. It is worth noting that performance indicators PI1, PI2, and PI3 are expressed in the same units as the target variable (i.e., L/s).

J. Water Resour. Plann. Manage., 2025, 151(10): 04025049

For each DMA d, evaluation week w, and performance indicator $\mathrm{PI}_{j_w}^d$ ($j = 1, \ldots, J$, being $J = 3$), the solutions provided by nt water demand forecasting models were ranked by (1) sorting water demand forecasting models in ascending order based on the value of performance indicator $PI_{j_w}^d$; and (2) assigning a value $r_{j_w}^d$ ($r_{j_w}^d = 1, \ldots, nt$) based on the position of each model in the sorted list (so that $r_{j_w}^d = 1$ in the case of the model providing the lowest $PI_{j_w}^d$ value and $r_{j_w}^d = nt$ in the case of the model providing the highest $PI_{j_w}^d$ value). The overall performance of each forecasting model was then assessed by calculating the average rank value R, namely, the sum of the rank values $r_{j_w}^d$ obtained for each DMA, week, and performance indicator divided by the product of the total number of DMAs (D), weeks (W), and performance indicators (J), as shown in Eq. (5):

$$R = \frac{\sum_{d=1}^{D} \sum_{w=1}^{W} \sum_{j=1}^{J} r_{j_w}^d}{D \cdot W \cdot J} \tag{5}$$

Based on Eqs. (2)–(5), it emerges that, the lower R, the lower $PI_{j_w}^d$, and, therefore, the higher the performance of a given forecasting model.

Considering the above, the most effective water demand forecasting model (i.e., the winning solution of the challenge) was the model with the lowest average rank value R. The average rank value, R, is not necessarily an integer number and can range between a minimum possible value, $R_{\min} = 1$, in in the ideal case of a team placing in the first position for all PIs, DMAs, and evaluation weeks, and a maximum potential value, $R_{\max} = 31$, in the case of a team consistently placing in the last position.

The missed submission deadline of a generic solution, associated with one week and with one DMA, resulted in the highest rank being assigned to the team for that week and for that DMA.

Competing Methods

A total of $nt = 31$ teams were involved in the BWDF, each proposing a water demand forecasting method (WDFM) to address the BWDF challenge. Details about each of the nt methods (from now on denoted as $\{M1, M2, M3, \ldots, M30, M31\}$) are available in the single contributions of the participating teams, which are gathered in a special volume dedicated to the conference (Alvisi et al. 2024). The references to the individual contributions are available in Table 3, where the top-10 teams are sorted based on ranking position, whereas teams placed from the 11th to the 31st position are sorted in alphabetical order. Table 3 also summarizes WDFMs' main features, based on which a clustering analysis is conducted to outline the main characteristics of the approaches—highlighting similarities and dissimilarities among different methods—and making an initial and qualitative discrimination before focusing on the quantitative results brought by the methods. More specifically, WDFM clustering is conducted based on two major features: (1) forecasting-method type; and (2) input data exploited.

As far as the forecasting-method type is concerned, the vast majority of WDFM proposed (i.e., 30 methods out of 31, 97%) use soft computing approaches to perform water demand forecasting, optimize forecasting parameters, or select the best solution from different tools. Among soft computing-based methods, ANNs emerge as the most adopted technique, a component of nearly 70% of methods (i.e., 21). In greater detail, among ANNs, deep learning methods such as long-short-term memory networks are mainly used (M1, M8, M10, M13, M14, M15, M17, M18, M19, M22, M24, M26, and M29), confirming the increase in the application of these soft computing approaches in the field of urban water demand forecasting (Mu et al. 2020; Hao et al. 2024). Conversely, a more limited number of methods (48%, i.e., 15 methods) make use of OSC techniques, such as gradient boosting (M2, M3, M4, M8, M11, and M22), regression trees (M7, M28), support vector regression (M8, M31) or random forests (M6, M20, M21, M22, M24, and M27). Besides relying on soft computing-based models, about 26% of WDFMs (i.e., eight methods) also include TS-analysis methods to increase model robustness and provide a more affordable water demand forecast. In some cases (e.g., M1, M8), the adopted method based on TS analysis is the naive method (i.e., a method producing water demand forecasts for a given time of a given day and using the data observed at the same time of the same day in a given number of previous weeks). In other cases, authors considered autoregressive models (M8, M22, and M28) or pattern-based methods (M6). TS analysis techniques and, particularly, generalized autoregressive moving average models, are used as the only forecast strategy in the single case of WDFM developed without including AI tools (M16). Overall, hybrid methods including a combination of different models cover about 42% of the proposed approaches (i. e., 13/31 WDFMs). More specifically, some hybrid methods include the application of different approaches for water demand forecast in relation to different DMAs and evaluation weeks, whereas other methods are based on the application of different approaches to the same DMA, thus providing an ensemble of forecast, which in some way could be useful to take into account uncertainty related to the forecast itself, even though at the very end a deterministic forecast is provided given the very nature of the indicators proposed within the framework of the Battle. In contrast with hybrid approaches, methods based on a single technique represent the remaining 58%. In particular, approaches based on a single ANN model represent 36% of the total (i.e., 11/31 WDFMs), those based on a single OSC technique represent 19% of the total (i.e., 6/31 WDFMs), and the residual 3% of the models is exclusively based on TS analysis (i.e., 1/31 WDFMs).

As far as input-data type is concerned, WDFM clustering reveals that all methods perform water demand forecasting by making use of historical net inflow data coupled with other types of data, such as weather data or information about the type of each day of the forecasting horizon (i.e., calendar data). In about 74% of cases (i.e., 23 methods), water demand forecasting is conducted by exploiting all these kinds of data, whereas 16% of WDFM (i.e., five methods) do not rely on weather data and only 10% (i.e., three methods) do not consider calendar data. This analysis reveals that (1) in the majority of cases, water demand forecasting is conducted by considering not only the historical water demand data but also weather forecasts (i.e., weather data) and information about day type (i.e., calendar data), if available; and (2) calendar data are exploited in more than 90% of cases, suggesting the general need of considering information on day type when dealing with water demand forecasting (due to considerably different behaviors in terms of water demand throughout the week).

Finally, to complete the analysis, the following additional aspects are highlighted:

- Almost all the teams dealt with the presence of net inflow data gaps before the application of the forecast models. As mentioned in the previous section, net inflow data were not post-processed and showed gaps related to other data collection or transmission issues, as usually happens with real-world data (Zanfei et al. 2022a). Teams applied different approaches to perform data imputation, such as interpolation, historical

J. Water Resour. Plann. Manage., 2025, 151(10): 04025049

Table 3. Main features of the water demand forecasting methods (WFDM) presented

WDFM	Ranking position	Reference	Model type	Input data			Data preprocessing	Method calibration	Probabilistic approach	Code language
				HID	CD	WD				
M1	1	Kley-Holsteg et al. (2024)	Hybrid (ANN, TS)	Yes	Yes	T, R, H	Yes	Local	—	Hybrid (Python, R)
M2	2	Zanutto et al. (2024)	Hybrid (ANN, OSC, TS)	Yes	Yes	T, R, H, W	Yes	Multiple	—	Python
M3	3	Bakhshipour et al. (2024)	Hybrid (ANN, OSC, TS)	Yes	Yes	—	Yes	Local	—	Python
M4	4	Groß and Hans (2024)	OSC	Yes	Yes	T, R, H, W	Yes	Multiple	—	Python
M5	5	Gabriele et al. (2024)	Hybrid (ANN)	Yes	Yes	T, R, H, W	Yes	Local	Yes	Hybrid (MATLAB, Fortran)
M6	6	Creaco et al. (2024)	Hybrid (OSC, TS)	Yes	Yes	T, R, H, W	Yes	Local	—	Hybrid (Python, MATLAB)
M7	7	Pagano et al. (2024)	OSC	Yes	Yes	—	Yes	Local	—	Python
M8	8	Ferreira et al. (2024)	Hybrid (ANN, OSC, TS)	Yes	Yes	T, R, H, W	Yes	Local	—	Python
M9	9	Ramachandran et al. (2024)	ANN	Yes	Yes	T	Yes	Local	—	Python
M10	10	Wunsch et al. (2024)	ANN	Yes	Yes	T, R	Yes	Local	Yes	Python
M11	>10	Arsova et al. (2024)	OSC	Yes	Yes	—	Yes	Local	—	Python
M12	>10	Ayyash et al. (2024)	ANN	Yes	Yes	T, R, H, W	Yes	Local	—	Python
M13	>10	Boloukasli Ahmadgourabi et al. (2024)	ANN	Yes	Yes	T, R, H, W	Yes	Local	—	Python
M14	>10	Brentan et al. (2024)	ANN	Yes	Yes	T, H	Yes	Local	—	Hybrid (Python, MATLAB)
M15	>10	Coy et al. (2024)	ANN	Yes	Yes	T, H, W	Yes	Local	—	Python
M16	>10	Gamboa-Medina and Campos (2024)	TS	Yes	Yes	—	—	Local	Yes	R
M17	>10	Geranmehr et al. (2024)	ANN	Yes	Yes	T, R, H, W	Yes	Global	—	Python
M18	>10	Iglesias-Rey et al. (2024)	Hybrid (ANN)	Yes	—	T, R, H, W	Yes	Local	—	Other (SAS Viya)
M19	>10	Jahangir and Quilty (2024)	Hybrid (ANN)	Yes	—	T, R, H, W	Yes	Multiple	—	Python
M20	>10	Kossieris et al. (2024) (adapted from)	OSC	Yes	Yes	T, R	Yes	Global	—	R
M21	>10	Kulaczkowski and Lee (2024)	OSC	Yes	Yes	T, R, H, W	Yes	Local	—	R
M22	>10	Perelman et al. (2024)	Hybrid (ANN, OSC, TS)	Yes	Yes	T, R, H, W	Yes	Local	—	Python
M23	>10	Pesantez et al. (2024)	ANN	Yes	Yes	T, R, H, W	Yes	Global	—	MATLAB
M24	>10	Que et al. (2024)	Hybrid (ANN, OSC)	Yes	Yes	T, R, H ,W	Yes	Global	—	Python
M25	>10	Reynoso-Meza and Carreño-Alvarado (2024)	Hybrid (ANN, OSC)	Yes	Yes	T, R, H, W	Yes	Local	—	MATLAB
M26	>10	Salem and Abokifa (2024)	ANN	Yes	Yes	T, R, H, W	Yes	Local	—	Python
M27	>10	Sinske et al. (2024)	Hybrid (OSC)	Yes	Yes	—	Yes	Local	—	Hybrid (Python, VBA)
M28	>10	Ulusoy et al. (2024)	Hybrid (OSC, TS)	Yes	Yes	T, R, H, W	Yes	Local	—	Other (Julia)
M29	>10	Wang et al. (2024)	ANN	Yes	Yes	T, R	Yes	Local	—	Hybrid (Python, MATLAB)
M30	>10	Yao et al. (2024)	ANN	Yes	—	T, R	Yes	Local	—	Python
M31	>10	Yu et al. (2024)	OSC	Yes	Yes	T	Yes	Local	—	MATLAB

Note: ANN = artificial neural networks; OSC = other soft computing; TS = time series analysis; HID = historical inflow data; CD = calendar data; WD = weather data; T = air temperature; R = rainfall depth; H = air humidity; and W = wind speed.

statistical values, moving average, K-nearest neighbors, neural networks, or random forest.
- In most cases (nearly 80%, i.e., 24/31 WDFMs), models were locally calibrated; i.e., water demand forecasting parameters were individually obtained for each DMA. Conversely, a global calibration was performed considering all DMAs grouped together in about 13% of cases (i.e., 4/31 WDFMs). Only models M2, M4, and M19 were parametrized at multiple spatial levels, and the solution that led to the highest forecast accuracy was finally selected.
- Despite the deterministic nature of the BWDF, a limited number of teams (i.e., 3) included a probabilistic component in their method. In greater detail, model conditional processors, probability functions, and copulas were considered to account for the uncertain nature of water demand.
- The majority of the WDFMs (55%) were developed by using the Python programming language, followed by MATLAB (10%), R (10%), a combination of those (19%), or other languages (6%). This finding aligns with Python's popularity in scientific computing and data analysis due to its flexibility and transferability (Ryzhkov et al. 2024).

Results and Discussion

The forecasts provided by the 31 teams are compared to the observed net inflow data and the results analyzed at three different levels, namely, (1) a general comparison between observed and forecasted time series; (2) the analysis of PIs values; and (3) the obtainment of challenge ranking and the discussion of the performance of different WDFMs. Overall, there was a high level of participation amongst teams. Only a single team (i.e., M25) did not submit their solution for only a single evaluation week (i.e., W2). Based on the BWDF rules, team M25 was assigned to the last position for evaluation week W2.

In addition to the 31 WDFMs, the naive method was considered as a benchmark. The naive method—defined in the literature as the "mean" model (Gelazanskas and Gamage 2015; Gagliardi et al. 2017)—performs weekly water demand forecasting based on the mean values $\mu = [\mu_1, \mu_2, \ldots \mu_{168}]$ associated with each of the 168 weekly hours calculated based on the 4 weeks preceding each evaluation week. Specifically, the water demand forecast for a generic hour j of the day is assumed to be equal to the corresponding mean demand μ_j.

Analysis of Observed and Forecasted Net Inflow Time Series

The analysis of the observed and forecasted net inflow time series is conducted to compare the quality of the forecasts provided by the 31 teams across the different DMAs and evaluation weeks. The comparisons of the observed inflows and forecasted ones submitted by the teams are introduced in Fig. 1 for each of the 10 DMAs in the first evaluation week (W1). The results of the other evaluation weeks (W2, W3, and W4) are in Figs. S2–S4, along with a Microsoft Excel spreadsheet including all water demand time series forecasted by each team.

Concerning Fig. 1, some interesting aspects can be emphasized. Considering DMA 1 [Fig. 1(a)], the observed trend is generally well reproduced by the forecasting models, but only some of them correctly reproduce the increase in water demand occurring on Monday and Tuesday in the very early morning. In greater detail, these peaks in demand are not related to problems in flow data acquisition or exceptional uses but are traceable back to the nature of DMA 1. The first DMA is, in fact, a hospital district, and these periodic withdrawals of water are related to flushing purposes, as confirmed by the water utility manager. The same considerations apply to the other evaluation weeks (Figs. S2a, S3a, and S4a).

An overall poor performance of the forecasting models for W1 is observed in DMA 2 and DMA 3 [Figs. 1(b and c)]. In this regard, the net inflow of these two DMAs—both residential districts in the countryside and showing similar features in terms of number of users supplied and average net inflow—is susceptible to weather conditions and, especially, rainfall during summer. In particular, a significant decrease in water demand is observed on summer rainy days for both DMAs, and this is due to the presence of several houses featuring large gardens requiring irrigation and pools. This behavior is emphasized in summer (to which W1 refers), whereas it tends to have less impact during the rest of the year (i.e., W2–W4, Figs. S2–S4). Overall, the majority of the methods, which overestimated water demand for both the DMAs over week W1, did not correctly consider the correlation between water demand and weather. In addition, some issues with the forecasts emerge with a few methods providing an almost constant value (around the average water demand) over the whole evaluation week.

The forecast of water demand over the four evaluation weeks for DMA 4 [Figs. 1(d), S2(d), S3(d), and S4(d)] and DMA 5 [Figs. 1(e), S2(e), S3(e), and S4(e)] was a success for almost all the methods proposed, and the majority of teams provided forecasted time series in line with the observed one. This can be related to the nature of the districts (residential/commercial) and the high number of users supplied, thus creating a high value of the average net inflow entering the areas. This latter aspect is formally analyzed in the subsequent phase of analysis.

The water demand forecast of DMA 6 [Figs. 1(f), S2(f), S3(f), and S4(f)], DMA 7 [Figs. 1(g), S2(g), S3(g), and S4(g)], and DMA 8 [Figs. 1(h), S2(h), S3(h), and S4(h)] is generally good over the four evaluation weeks, but some problems with the forecast models emerge. Some problems include the underestimation of the DMA night consumption (especially for DMA 6), the obtainment of almost constant forecasted time series, or the temporal shifting of the forecasted series with respect to the observed one.

DMA 9 [Figs. 1(i), S2(i), S3(i), and S4(i)] and DMA 10 [Figs. 1(j), S2(j), S3(j), and S4(j)] are commercial/industrial districts—close to the port—featuring similarities in the type of users supplied and values of net inflow. A different weekly pattern characterizes the water demand time series of these DMAs as opposed to the patterns that are generally associated with residential DMAs. Indeed, the DMA inflow trend highlights an evident difference between weekdays and weekends, with a clear water-consumption drop occurring over the latter as a consequence of the decrease or halting of most industrial activities (Mazzoni et al. 2024). Despite the similarity between DMA 9 and DMA 10, the forecasting methods show a different performance for the two DMAs and the same evaluation week. For example, for W1, the forecast for DMA 10 [Fig. 1(j)] is generally good, whereas the forecast for DMA 9 [Fig. 1(i)] tends to be more affected by inaccuracies related to very smooth, almost constant forecasted time series, or the overestimation of water demand mainly for the second day of the week.

To evaluate the performance of each water-forecasting model, the three PIs are considered, and the results of this quantification are discussed as the second level of analysis.

Performance Indicator Analysis

The second level of analysis focuses on the three PIs. On the one hand, PI1 and PI2 refer to the MAE for a shorter term period

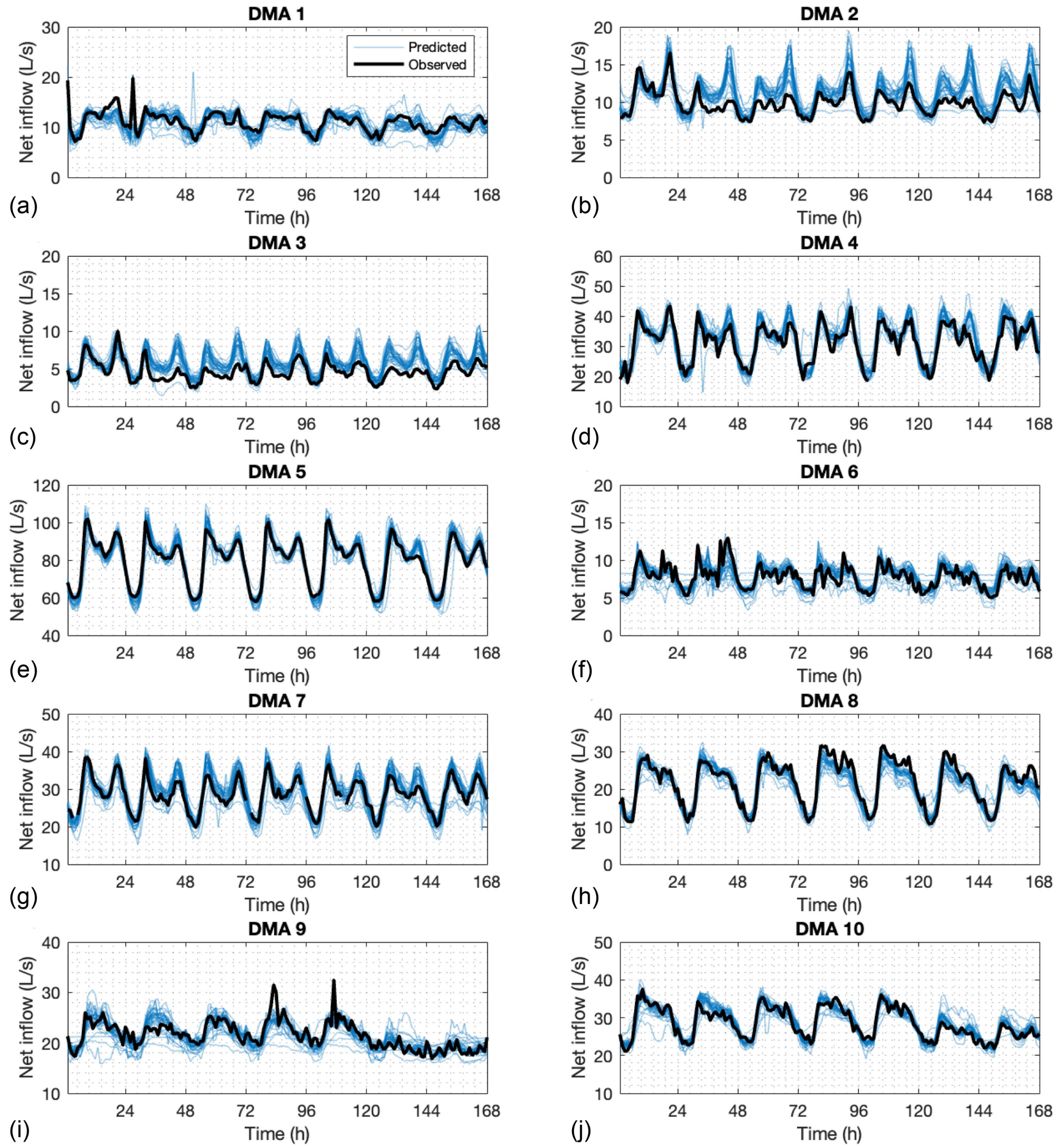

Fig. 1. Observed and forecasted inflow in the 10 DMAs during the first evaluation week (W1).

(i.e., the first 24 h of each evaluation week) and to a longer term period (i.e., the period between the second and the final day of each evaluation week), respectively. On the other hand, PI3 is evaluated as the maximum absolute error for the same shorter term period as for PI1.

The values of the single PIs for all teams and referring to all evaluation weeks (i.e., W1, W2, W3, and W4) are considered for each DMA. The boxplots of these values are reported in Fig. 2, where the y axis is normalized with respect to the average net inflow of the single DMA (to facilitate the comparison among the different districts), and the dots point out the mean value of the considered PI. In detail, Fig. 2 highlights that the average MAEs for the shorter term period (i.e., PI1) and the longer term period (i.e., PI2) are generally very similar, with values lower than 10% of the average net inflow (except for few cases) and rather limited dispersion of the single values of PI1 and PI2. Otherwise, the average maximum absolute error (i.e., PI3) exceeds 20% of the average net inflow for almost all DMAs with single values of PI3 generally quite spread (up to values of more than 100%).

For most of the DMAs (i.e., DMAs 1–5, 7, and 8), the average value of PI1 remains slightly lower than the average value of PI2. In other words, the forecast in the shorter term tends to be marginally better than the forecast in the longer term. However, this is not the case for a few DMAs (i.e., DMAs 6, 9, and 10), for which the accuracy of the forecast methods is slightly higher in the longer term period.

From the analysis of PIs, a second aspect related to the values of the first two PIs can be emphasized: PI1 and PI2 tend to decrease with the increase in the average net inflow of the DMAs, meaning that the higher the net inflow of the DMA, the better the performance of the forecast methods. This aspect is highlighted by the graphical results reported in Fig. 3 and seems generally valid, regardless of the characteristics of the DMA and the users supplied. In detail, on the one hand, higher values of PI1 are associated with DMA 1, DMA 3, and DMA 6, which are the three districts with the lowest values of average net inflow (i.e., 8.2, 4.2, and 8.4 L/s, respectively). On the other hand, the lowest value of PI1 is associated with DMA 5, which is the most demanding in terms of water delivered to the area (i.e., 78.7 L/s). Similar considerations are valid if PI2 is considered. These results confirm that larger DMAs featuring a high number of users tend to result in a more

J. Water Resour. Plann. Manage., 2025, 151(10): 04025049

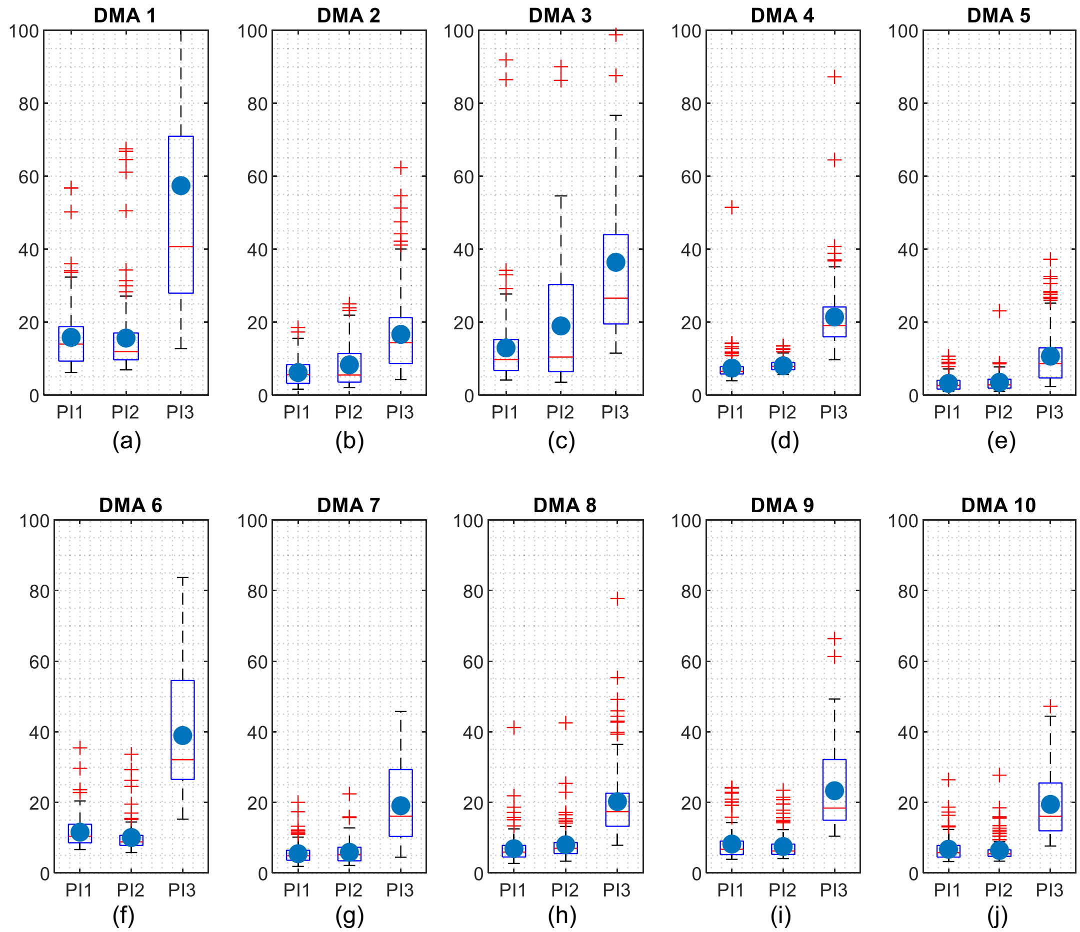

Fig. 2. Boxplot of the values of PI1, PI2 and PI3 for each of the (a–j) 10 DMAs and evaluation weeks, considering all the 31 WDFMs. The *y* axis is normalized with respect to the average net inflow of the DMA, whereas dots point out the mean value assumed by the single PIs.

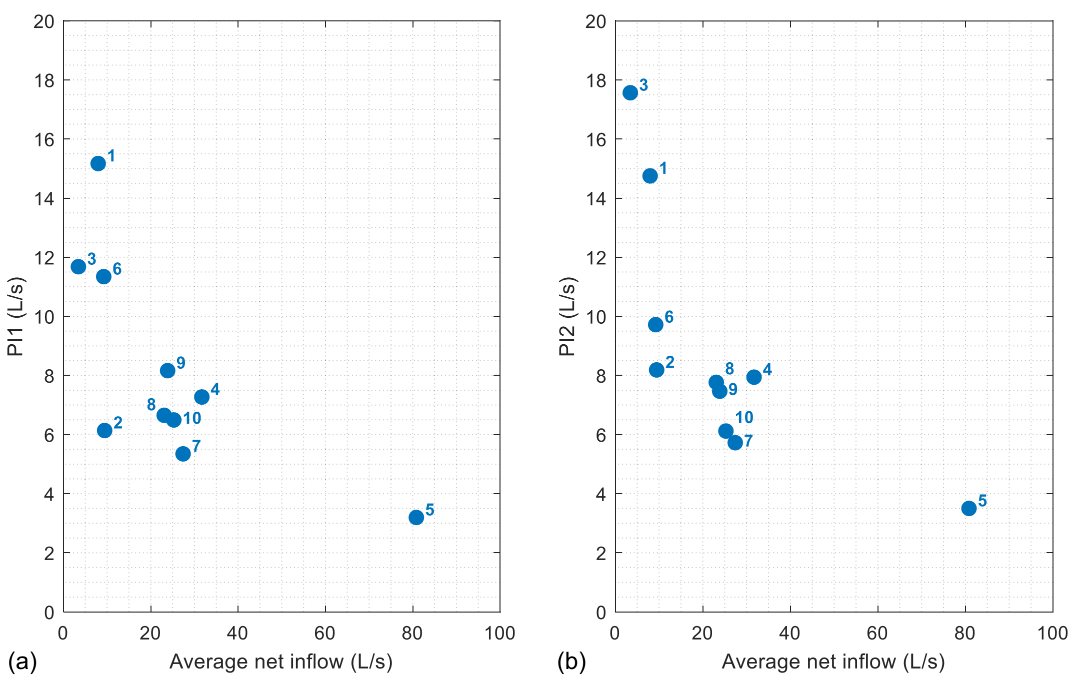

Fig. 3. Average values of (a) PI1; and (b) PI2 as a function of the DMA average net inflow. The number next to each marker refers to the DMA considered in the competition.

stable trend of water demand over time and, thus, are easier to forecast.

Ranking and WDFM Performances

The third level of analysis is based on the ranking of the BWDF and discusses the key aspects that contribute to the definition of the final placement of the participating teams, as well as the features of the WDFMs applied.

The ranking is obtained by evaluating the average rank value R defined in Eq. (5) for each one of the 31 teams starting from the single rank values, $r^d_{j_w}$, obtained considering each DMA d ($d = 1, \ldots, D$ being $D = 10$), evaluation week w ($w = 1, \ldots, W$ being $W = 4$), and performance indicator PI ($j = 1, 2, 3$). The anonymized ranking of the BWDF is reported in Fig. 4, where for each team the single rank values, $r^d_{j_w}$, are considered to build the boxplots, whereas the average rank value, R, is shown with a dot. In relation to Fig. 4, teams are ordered on the x axis based on their ranking and, consequently, based on the associated average rank value, R. The winning team corresponds to the x coordinate equal to 1 (i.e., first position), whereas the worst performing team corresponds to the x coordinate equal to 31 (i.e., last position). The ranking presented in the figure is thus not related to the order in which WDFMs are introduced in the Competing Methods section. However, the WDFMs related to the highest ranking—i.e., placed in the first 10 positions of the BWDF—are indicated in Table 3.

Fig. 4 reveals that the first-placed team obtained a value of R of almost 7, which is higher than the minimum possible average rank value, $R_{\min} = 1$, that could have been obtained only if the team overall placed in the first position would have been placed in the first position in each of the $D \cdot W \cdot J = 120$ single rank values, $r^d_{j_w}$. Conversely, the team placed in the last position is associated with an R value of almost 27, which is lower than the maximum possible average rank value, $R_{\max}$. This means that no teams achieved the highest or the lowest performance for all PIs, DMAs, and evaluation weeks.

More specifically, the teams placed in the first four positions achieved average rank values R rather distant from each other (i.e., R equal to 6.6, 7.7, 9.2, and 10.3, respectively), whereas the teams covering the positions from the fifth to the ninth have rather similar R values (i.e., R overall ranging between 11.2 and 12.1). A sort of plateau with a slightly increasing trend is observable between the 10th and the 24th positions (R between 13.5 and 18.2). By contrast, the last seven teams are associated with considerably higher values of the average rank value R (with values from 21.4 to 26.7, the latter referring to the team in the last position). Beyond the chart, additional aspects can be observed from Fig. 4. For example, with a focus on outliers, it emerges that also the winning team was placed in the lowest positions at least once (i.e., in the case of at least one combination of DMA, evaluation week, and performance indicator). At the same time, teams overall placed in the last several positions performed better than the majority of the teams for a given PI, week, and DMA at least once. This confirms that, overall, no method always outperformed or underperformed the others for all PIs, evaluation weeks, and DMAs.

To further validate the above observations, the boxplots of the single rank values are produced for individual PIs (Fig. S5), evaluation weeks (Fig. S6, Supplemental Materials), and DMAs (Fig. S7 for DMAs 1–4, Fig. S8 for DMAs 5–8, and Fig. S9 for DMAs 9 and 10) where the spread of single rank values can be observed. When a single PI, evaluation week, or DMA is considered, the tendency of some forecast models to perform better or worse is emphasized. By way of example, if evaluation week W1 was only considered, the team positioned 29th would be placed in the middle of the ranking (Fig. S6a). Similarly, if only DMA 8 was considered (Fig. S8d), the team placing seventh would be the best performing team. In light of the above considerations, it can be stated that assessment criteria considering more DMAs, more evaluation periods, and more evaluation metrics emerge as essential requirements when it comes to effectively evaluating the performance of a water demand forecasting model.

If ranking is coupled with the available information regarding the type of methods proposed by the teams, further considerations can be set forth. A graphical representation of the link between chart and method characteristics is reported in Fig. 5, where teams are marked with different colors and symbols based on the forecasting model type resulting from clustering. In addition, methods that included a time series analysis component in the forecasting method are further highlighted.

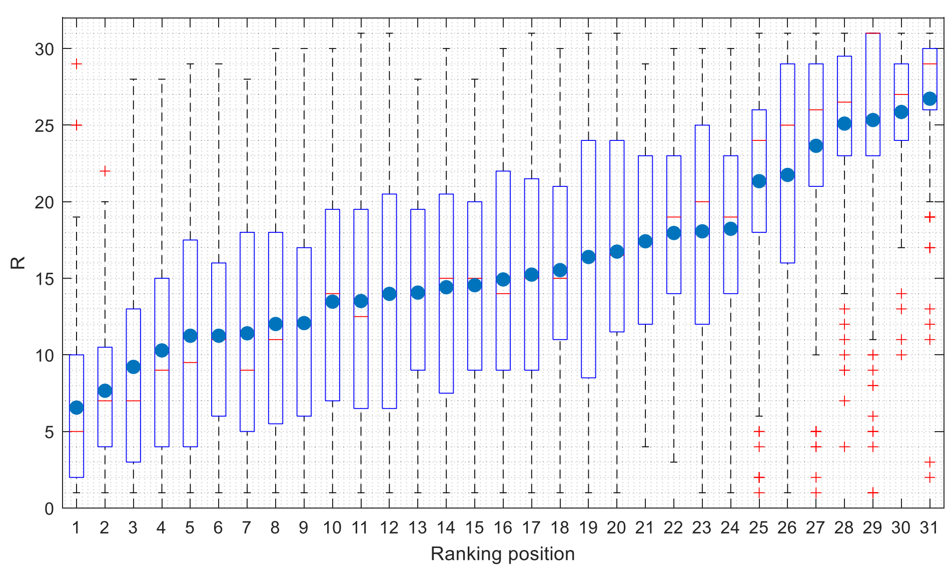

Fig. 4. Ranking of the 31 participating teams. Boxplot of the single rank values considering the three PIs, the four evaluation weeks, and the 10 DMAs. Correspondent average rank values R are also shown (dots).

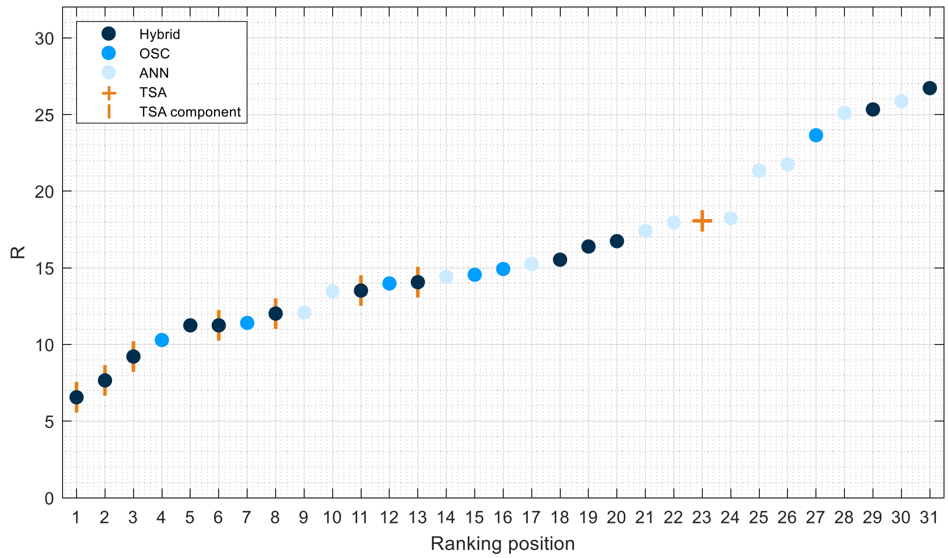

Fig. 5. Forecast methods positioning based on the average rank value *R* with thematic map based on forecasting model type.

Hybrid methods tend to have better performance than single-model approaches, whereas the use of a single method generally results in lower accuracy forecasts. Overall, including TS analysis as a component of the forecast method leads to better results. It is worth noting that the teams placed in the first three positions—that is, Kley-Holsteg et al. (2024), Zanutto et al. (2024), and Bakhshipour et al. (2024) (i.e., M1, M2, and M3 in Table 3)—used hybrid models including TS analysis as a component of the forecasting method. Specifically, (1) M1 (Kley-Holsteg et al. 2024) includes a combination of more than 50 individual WDFMs—spacing from ANN-based models (e.g., long-short-term memory networks and multilayer perceptron-based techniques) to TS models—that are individually applied in relation to each DMA and evaluation week, while the overall forecast is provided by mean of an automated aggregation method that dynamically combines individual WDFM forecasts by adjusting weights based on past performance; (2) M2 (Zanutto et al. 2024) integrates an ensemble of different WDFMs of different natures—including ANNs (convolutional neural networks) and OSC techniques (gradient boosting)—to provide a deterministic forecast based on the selection and the weighted combination of the top-performing WDFMs (compared against the results of TS approaches, e.g., rolling average); and (3) M3 (Bakshipour et al. 2024) relies on the integration of three soft computing methods—including ANN-based models and OSC techniques (convolutional neural networks, multilayer perceptron-based approach, and gradient boosting)—the weights of which are fine-tuned to provide an overall, optimized forecast.

In this regard, the naive method, which generates water demand forecasts based on the mean values belonging to weeks preceding each evaluation week, was considered as a benchmark. In particular, in light of the ranking of the challenge (Fig. 5), the naive method would place at the 12th position, thus performing better than two-thirds of the methods proposed by the participating teams. Although exclusively based on historical net inflow data and the related mean values, the naive method allows seizing the periodicities that characterize water demand time series at different temporal scales, from daily to weekly and seasonal. Therefore, it can be considered as an effective tool for water demand forecasting, allowing the obtainment of a raw—but robust—solution to use as a benchmark without facing the limitations related to the application of data-driven methodologies, e.g., large amounts of time series data needed. In this regard, it is worthy of note that the majority of the forecast methods outperforming the naive method include a simple TS approach (in line with the naive method) in the forecasting process, or consider it as a benchmark. This allowed to automatically or manually check for the absence of gross errors in the forecasted time series, while verifying that the water demand periodical trend is correctly forecasted. Hence, this choice proved to be successful when it comes to increasing method robustness and performance.

Conclusions

Organized in the context of the 3rd International WDSA-CCWI Joint Conference held in Ferrara (Italy) in 2024, the BWDF provided an excellent opportunity for researchers and practitioners to solve a complex problem related to real-world WDNs. In particular, the challenge aimed to compare the effectiveness of methods for the short-term forecast of water demand considering 10 DMAs in a real WDN located in the northeast of Italy. Beyond defining the challenge and the assessment criteria, this paper summarizes the approaches adopted by the 31 teams participating from different contexts worldwide. This work also offers a comparison of the results submitted by each team, outlining the major insights on the state-of-the-art of short-term water demand forecasting. The outcomes and the key findings that emerged from the study are reported in the following:

- In general, the participating teams provided forecasted time series in line with the observed ones over the four evaluation weeks defined in the context of the challenge. In a few cases, peculiar issues emerged with some models, such as the obtainment of the forecasted net inflow series characterized by almost constant values over the whole evaluation week, along with time series affected by a temporal shifting with respect to the observed one.
- The assessment of the three PIs for all the teams concerning the four evaluation weeks and the 10 DMAs allows for making additional considerations. On the one hand, the average MAEs within the first 24 h (i.e., PI1) and in the period between the second and the final day of each evaluation week (i.e., PI2) are generally very similar, with values lower than 10% of the

average net inflow (except for a few cases), and single values of the indicators show somewhat limited dispersion. On the other hand, the average maximum absolute error (i.e., PI3) exceeds 20% of the average net inflow for almost all DMAs with single values of PI3 generally quite spread (up to values of more than 100%). Overall, the forecast in the shorter term tends to be slightly better than the forecast in the longer term.
- No method always outperformed or underperformed the others for all PIs, evaluation weeks, and DMAs. However, hybrid methods tend to perform better than single-model approaches.
- Surprisingly, the naive method would have outperformed two-thirds of the methods proposed by the participating teams, since it can seize the periodicities in water demand and provide an approximate but robust solution exclusively based on historical inflow data. It is worthy of note that the majority of the forecast methods that performed better than the naive (including the teams placed in the first three positions) consider, among the others, a simple TS approach in line with the naive in the process.

Overall, the results obtained in the context of the BWDF reveal that a combination of methods of different natures can effectively improve the accuracy of a forecast model, as demonstrated by the top-performing WDFMs. Despite this, water demand forecast is still a complex problem, and the selection of the most appropriate water demand forecasting model is not always straightforward. To this regard, the following general recommendations for future WDFM applications are provided: (1) assessment criteria that consider more than one metric, more than one DMA, and more than one evaluating week emerge as an essential requirement to evaluate the robustness of a water demand forecasting method and understand the circumstances under which the selected method is more effective (e.g., 24-h versus weekly or peak demand forecast, large versus small DMA, residential versus nonresidential water demand, winter versus summer forecast, etc.); (2) automated or human-based postprocessing checks emerge as a critical component of the forecasting process to exclude the presence of gross errors in the forecasted water demand and check the periodicity of the related time series; and (3) TS analysis represents a relevant component for direct forecast benchmarking; i.e., the inclusion of TS techniques as a benchmark component of the forecast can lead to better and more consistent results.

The complex problem of water demand forecasting still remains an open topic with room for improvement, as also shown in the context of this challenge. For example, even though water demand forecasting is characterized by a certain degree of uncertainty due to the variability of water consumption (Gagliardi et al. 2017), applications of probabilistic approaches for short-term water demand forecasting have not been widely investigated in the scientific literature. Within the context of the BWDF, most of the teams made use of deterministic methods. This may be related to the nature of this competition, being a deterministic solution requested, and characterization of the forecasting uncertainty disregarded in the PIs used for evaluating the performance of the methods proposed. In fact, in general, the application of probabilistic approaches is dependent on the framework in which the forecast has to be used and on whether this framework can benefit from the application of stochastic outcomes. However, probabilistic approaches for short-term water demand forecasting can represent a field to be further explored and tested, particularly whenever the estimation of the uncertainty related to the forecasted demands can be effectively exploited, for example, in the decision-making process.

In conclusion, the BWDF can be considered an important milestone in the area of short-term water demand forecasting, providing a freely accessible database that can become a benchmark for future studies.

Data Availability Statement

All data, models, and code generated or used during the study appear in the published article. Specifically, observed and forecasted water demand data are available in the Supplemental Materials, along with weather and calendar data. Details on the forecast models used are available in the individual contributions gathered in a special volume dedicated to the conference (Alvisi et al. 2024) or by request from each participating team.

Author Contributions

S. Alvisi: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Supervision; Validation; Visualization; Writing – original draft; Writing – review and editing. M. Franchini: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Supervision; Validation; Visualization; Writing – original draft; Writing – review and editing. V. Marsili: Data curation; Formal analysis; Investigation; Methodology; Resources; Software; Validation; Visualization; Writing – original draft; Writing – review and editing. F. Mazzoni: Data curation; Formal analysis; Investigation; Methodology; Resources; Software; Validation; Visualization; Writing – original draft; Writing – review and editing. E. Salomons: Data curation; Formal analysis; Investigation; Methodology; Resources; Software; Validation; Visualization; Writing – original draft; Writing – review and editing. M. Housh: Data curation; Formal analysis; Investigation; Methodology; Resources; Software; Validation; Visualization. A. Abokifa: Formal analysis; Investigation; Methodology; Software. K. Arsova: Formal analysis; Investigation; Methodology; Software. F. Ayyash: Formal analysis; Investigation; Methodology; Software. H. Bae: Formal analysis; Investigation; Methodology; Software. R. Barreira: Formal analysis; Investigation; Methodology; Resources. L. Basto: Formal analysis; Investigation; Methodology; Software. S. Bayer: Formal analysis; Investigation; Methodology; Software. E. Z. Berglund: Formal analysis; Investigation; Methodology; Software. D. Biondi: Formal analysis; Investigation; Methodology; Software. F. Boloukasli Ahmadgourabi: Formal analysis; Investigation; Methodology; Software. B. Brentan: Formal analysis; Investigation; Methodology; Software. J. Caetano: Formal analysis; Investigation; Methodology; Software. F. Campos: Formal analysis; Investigation; Methodology; Software. H. Cao: Formal analysis; Investigation; Methodology; Software. S. Cardona: Formal analysis; Investigation; Methodology; Software. E. P. Carreño Alvarado: Formal analysis; Investigation; Methodology; Software. N. Carriço: Formal analysis; Investigation; Methodology; Software. G.-A. Chatzistefanou: Formal analysis; Investigation; Methodology; Software. Y. Coy: Formal analysis; Investigation; Methodology; Software. E. Creaco: Formal analysis; Investigation; Methodology; Software. S. Cuomo: Formal analysis; Investigation; Methodology; Software. A. de Klerk: Formal analysis; Investigation; Methodology; Software. A. Di Nardo: Formal analysis; Investigation; Methodology; Software. M. DiCarlo: Formal analysis; Investigation; Methodology; Software. U. Dittmer: Formal analysis; Investigation; Methodology; Software. R. Dziedzic: Formal analysis; Investigation; Methodology; Software. A. Ebrahim Bakhshipour: Formal analysis; Investigation; Methodology; Software. D. Eliades:

Formal analysis; Investigation; Methodology; Software. R. Farmani: Formal analysis; Investigation; Methodology; Software. B. Ferreira: Formal analysis; Investigation; Methodology; Software. A. Gabriele: Formal analysis; Investigation; Methodology; Software. F. Gao: Formal analysis; Investigation; Methodology; Software. M. M. Gamboa-Medina: Formal analysis; Investigation; Methodology; Software. J. Gao: Formal analysis; Investigation; Methodology; Software. R. Gargano: Formal analysis; Investigation; Methodology; Software. M. Geranmehr: Formal analysis; Investigation; Methodology; Software. C. Giudicianni: Formal analysis; Investigation; Methodology; Software. K. Glynis: Formal analysis; Investigation; Methodology; Software. S. Gómez: Formal analysis; Investigation; Methodology; Software. L. González: Formal analysis; Investigation; Methodology; Software. M. Groß: Formal analysis; Investigation; Methodology; Software. H. Guo: Formal analysis; Investigation; Methodology; Software. M. N. Habibi: Formal analysis; Investigation; Methodology; Software. A. Haghighi: Formal analysis; Investigation; Methodology; Software. B. Hammer: Formal analysis; Investigation; Methodology; Software. L. Hans: Formal analysis; Investigation; Methodology; Software. M. Hayslep: Formal analysis; Investigation; Methodology; Software. Y. He: Formal analysis; Investigation; Methodology; Software. L. Hermes: Formal analysis; Investigation; Methodology; Software. M. Herrera: Formal analysis; Investigation; Methodology; Software. F. Hinder: Formal analysis; Investigation; Methodology; Software. B. Hou: Formal analysis; Investigation; Methodology; Software. A. Iglesias-Rey: Formal analysis; Investigation; Methodology; Software. P. L. Iglesias-Rey: Formal analysis; Investigation; Methodology; Software. I.-S. Jang: Formal analysis; Investigation; Methodology; Software. J. Izquierdo: Formal analysis; Investigation; Methodology; Software. M. S. Jahangir: Formal analysis; Investigation; Methodology; Software. C. Jara-Arriagada: Formal analysis; Investigation; Methodology; Software. B. Jenks: Formal analysis; Investigation; Methodology; Software. G. Johnen: Formal analysis; Investigation; Methodology; Software. M. Kalami Heris: Formal analysis; Investigation; Methodology; Software. M. Kalumba: Formal analysis; Investigation; Methodology; Software. M.-S. Kang: Formal analysis; Investigation; Methodology; Software. M. Khashei Varnamkhasti: Formal analysis; Investigation; Methodology; Software. K.-J. Kim: Formal analysis; Investigation; Methodology; Software. J. Kley-Holsteg: Formal analysis; Investigation; Methodology; Software. T. Ko: Formal analysis; Investigation; Methodology; Software. A. Koochali: Formal analysis; Investigation; Methodology; Software. P. Kossieris: Formal analysis; Investigation; Methodology; Software. P. Koundouri: Formal analysis; Investigation; Methodology; Software. C. Kühnert: Formal analysis; Investigation; Methodology; Software. A. Kulaczkowski: Formal analysis; Investigation; Methodology; Software. J. Lee: Formal analysis; Investigation; Methodology; Software. K. Li: Formal analysis; Investigation; Methodology; Software. Y. Li: Formal analysis; Investigation; Methodology; Software. H. Liu: Formal analysis; Investigation; Methodology; Software. Y. Liu: Formal analysis; Investigation; Methodology; Software. C. A. López-Hojas: Formal analysis; Investigation; Methodology; Software. A. Maier: Formal analysis; Investigation; Methodology; Software. C. Makropoulos: Formal analysis; Investigation; Methodology; Software. F. J. Martínez-Solano: Formal analysis; Investigation; Methodology; Software. N. H. Marzouny: Formal analysis; Investigation; Methodology; Software. A. Menapace: Formal analysis; Investigation; Methodology; Software. C. Michalopoulos: Formal analysis; Investigation; Methodology; Software. G. Moraitis: Formal analysis; Investigation; Methodology; Software. H. Mousa: Formal analysis; Investigation; Methodology; Software. H. Namdari: Formal analysis; Investigation; Methodology; Software. D. Nikolopoulos: Formal analysis; Investigation; Methodology; Software. M. Oberascher: Formal analysis; Investigation; Methodology; Software. A. Ostfeld: Formal analysis; Investigation; Methodology; Software. M. Pagano: Formal analysis; Investigation; Methodology; Software. F. Pasha: Formal analysis; Investigation; Methodology; Software. J. Perafán: Formal analysis; Investigation; Methodology; Software. G. Perelman: Formal analysis; Investigation; Methodology; Software. J. Pesantez: Formal analysis; Investigation; Methodology; Software. M. Polycarpou: Formal analysis; Investigation; Methodology; Software. M. G. Quarta: Formal analysis; Investigation; Methodology; Software. Q. Que: Formal analysis; Investigation; Methodology; Software. J. Quilty: Formal analysis; Investigation; Methodology; Software. C. Quintiliani: Formal analysis; Investigation; Methodology; Software. A. Ramachandran: Formal analysis; Investigation; Methodology; Software. G. Reynoso Meza: Formal analysis; Investigation; Methodology; Software. V. Rodriguez: Formal analysis; Investigation; Methodology; Software. Y. Romano: Formal analysis; Investigation; Methodology; Software. J. Saldarriaga: Formal analysis; Investigation; Methodology; Software. A. K. Salem: Formal analysis; Investigation; Methodology; Software. P. Samartzis: Formal analysis; Investigation; Methodology; Software. G. F. Santonastaso: Formal analysis; Investigation; Methodology; Software. D. Savic: Formal analysis; Investigation; Methodology; Software. V. Schiano Di Cola: Formal analysis; Investigation; Methodology; Software. D. Schol: Formal analysis; Investigation; Methodology; Software. A. G. Seyoum: Formal analysis; Investigation; Methodology; Software. R. Shen: Formal analysis; Investigation; Methodology; Software. K. Simukonda: Formal analysis; Investigation; Methodology; Software. A. Sinske: Formal analysis; Investigation; Methodology; Software. R. Sitzenfrei: Formal analysis; Investigation; Methodology; Software. B. Sonnenschein: Formal analysis; Investigation; Methodology; Software. I. Stoianov: Formal analysis; Investigation; Methodology; Software. A. Tabares: Formal analysis; Investigation; Methodology; Software. E. Todini: Formal analysis; Investigation; Methodology; Software. L. Tsiami: Formal analysis; Investigation; Methodology; Software. I. Tsoukalas: Formal analysis; Investigation; Methodology; Software. A.-J. Ulusoy: Formal analysis; Investigation; Methodology; Software. L. Vamvakeridou-Lyroudia: Formal analysis; Investigation; Methodology; Software. A. van Heerden: Formal analysis; Investigation; Methodology; Software. J. Vaquet: Formal analysis; Investigation; Methodology; Software. V. Vaquet: Formal analysis; Investigation; Methodology; Software. S. Wallner: Formal analysis; Investigation; Methodology; Software. M. Walraad: Formal analysis; Investigation; Methodology; Software. D. Wang: Formal analysis; Investigation; Methodology; Software. S. Wu: Formal analysis; Investigation; Methodology; Software. W. Wu: Formal analysis; Investigation; Methodology; Software. A. Wunsch: Formal analysis; Investigation; Methodology; Software. Y. Yao: Formal analysis; Investigation; Methodology; Software. J. Yu: Formal analysis; Investigation; Methodology; Software. A. Zanfei: Formal analysis; Investigation; Methodology; Software. D. Zanutto: Formal analysis; Investigation; Methodology; Software. H. Zhang: Formal analysis; Investigation; Methodology; Software. M. Ziebarth: Formal analysis; Investigation; Methodology; Software. F. Ziel: Formal analysis; Investigation; Methodology; Software. J. Zou: Formal analysis; Investigation; Methodology; Software.

Supplemental Materials

Observed water demand data, weather data, calendar data, forecasted water demand data, and Figs. S1–S9 are available online in the ASCE Library (www.ascelibrary.org).

References

Alvisi, S., M. Franchini, and A. Marinelli. 2007. "A short-term, pattern-based model for water-demand forecasting." *J. Hydroinf.* 9 (1): 39–50. https://doi.org/10.2166/hydro.2006.016.

Alvisi, S., M. Franchini, V. Marsili, and F. Mazzoni. 2024. "Preface of the 3rd international joint conference on water distribution systems analysis and computing and control for the water industry (WDSA/CCWI 2024)." *Eng. Proc.* 69 (1): 1. https://doi.org/10.3390/engproc2024069001.

Anele, A. O., Y. Hamam, A. M. Abu-Mahfouz, and E. Todini. 2017. "Overview, comparative assessment and recommendations of forecasting models for short-term water demand prediction." *Water* 9 (11): 887. https://doi.org/10.3390/w9110887.

Arandia, E., A. Ba, B. Eck, and S. McKenna. 2016. "Tailoring seasonal time series models to forecast short-term water demand." *J. Water Resour. Plann. Manage.* 142 (3): 04015067. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000591.

Arsova, K., C. Quintiliani, D. Schol, and M. Walraad. 2024. "Predicting net inflow for 10 DMAs in North-East Italy." *Eng. Proc.* 69 (1): 178. https://doi.org/10.3390/engproc2024069178.

Attallah, N., J. Horsburgh, and C. Bastidas Pacheco. 2023. "An open-source, semisupervised water end-use disaggregation and classification tool." *J. Water Resour. Plann. Manage.* 149 (7): 04023024. https://doi.org/10.1061/JWRMD5.WRENG-5444.

Ayyash, F., M. Hayslep, T. Ko, M. Kalumba, K. Simukonda, and R. Farmani. 2024. "Application of a neural network model to short-term water demand forecasting." *Eng. Proc.* 69 (1): 123. https://doi.org/10.3390/engproc2024069123.

Babel, M. S., and V. R. Shinde. 2011. "Identifying prominent explanatory variables for water demand prediction using artificial neural networks: A case study of Bangkok." *Water Resour. Manage.* 25 (6): 1653–1676. https://doi.org/10.1007/s11269-010-9766-x.

Bakhshipour, A. E., H. Namdari, A. Koochali, U. Dittmer, and A. Haghighi. 2024. "An ensemble data-driven approach for enhanced short-term water demand forecasting in urban areas." *Eng. Proc.* 69 (1): 69. https://doi.org/10.3390/engproc2024069069.

Bakker, M., J. H. G. Vreeburg, K. M. van Schagen, and L. C. Rietveld. 2013. "A fully adaptive forecasting model for short-term drinking water demand." *Environ. Modell. Software* 48 (Apr): 141–151. https://doi.org/10.1016/j.envsoft.2013.06.012.

Boloukasli Ahmadgourabi, F., M. Khashei Varnamkhasti, M. Nosrati Habibi, N. Hedaiaty Marzouny, and R. Dziedzic. 2024. "Enhancing water demand forecasting: Leveraging LSTM networks for accurate predictions." *Eng. Proc.* 69 (1): 120. https://doi.org/10.3390/engproc2024069120.

Brentan, B. M., A. Zanfei, M. Oberascher, R. Sitzenfrei, J. Izquierdo, and A. Menapace. 2024. "Cascade machine learning approach applied to short-term medium horizon demand forecasting." *Eng. Proc.* 69 (1): 42. https://doi.org/10.3390/engproc2024069042.

Coy, Y., L. González, L. Basto, V. Rodríguez, S. Gómez, J. Perafán, S. Cardona, A. Tabares, and J. Saldarriaga. 2024. "A methodology for forecasting demands in a water distribution network based on the classical and neural networks approach." *Eng. Proc.* 69 (1): 29. https://doi.org/10.3390/engproc2024069029.

Creaco, E., C. Giudicianni, and M. Herrera. 2024. "Multi-model demand forecasting in water distribution network districts." *Eng. Proc.* 69 (1): 188. https://doi.org/10.3390/engproc2024069188.

Daniel, D., S. J. Marks, S. Pande, and L. Rietveld. 2018. "Socio-environmental drivers of sustainable adoption of household water treatment in developing countries." *npj Clean Water* 1 (1): 12. https://doi.org/10.1038/s41545-018-0012-z.

de Souza Groppo, G., M. Azevedo Costa, and M. Libânio. 2019. "Predicting water demand: A review of the methods employed and future possibilities." *Water Supply* 19 (8): 2179–2198. https://doi.org/10.2166/ws.2019.122.

Donkor, E. A., T. A. Mazzucchi, R. Soyer, and J. A. Robertson. 2014. "Urban water demand forecasting: Review of methods and models." *J. Water Resour. Plann. Manage.* 140 (2): 146–159. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000314.

Ferreira, B., R. Barreira, J. Caetano, M. Quarta, and N. Carriço. 2024. "Optimizing short-term water demand forecasting: A comparative approach to the battle of water demand forecasting." *Eng. Proc.* 69 (1): 48. https://doi.org/10.3390/engproc2024069048.

Gabriele, A., D. Biondi, R. Gargano, and E. Todini. 2024. "From deterministic to probabilistic forecasts of water demand." *Eng. Proc.* 69 (1): 122. https://doi.org/10.3390/engproc2024069122.

Gagliardi, F., S. Alvisi, Z. Kapelan, and M. Franchini. 2017. "A probabilistic short-term water demand forecasting model based on the Markov chain." *Water* 9 (7): 507. https://doi.org/10.3390/w9070507.

Gamboa-Medina, M. M., and F. S. Campos. 2024. "Water demand forecast using generalized autoregressive moving average models." *Eng. Proc.* 69 (1): 125. https://doi.org/10.3390/engproc2024069125.

Gelazanskas, L., and K. Gamage. 2015. "Forecasting hot water consumption in residential houses." *Energies* 8 (11): 12702–12717. https://doi.org/10.3390/en81112336.

Geranmehr, M., A. Seyoum, and M. Heris. 2024. "Battle of water demand forecasting: An optimized deep learning model." *Eng. Proc.* 69 (1): 56. https://doi.org/10.3390/engproc2024069056.

Ghalehkhondabi, I., E. Ardjmand, W. A. Young II, and G. R. Weckman. 2017. "Water demand forecasting: Review of soft computing methods." *Environ. Monit. Assess.* 189 (Jul): 313. https://doi.org/10.1007/s10661-017-6030-3.

Grespan, A., J. Garcia, M. P. Brikalski, E. Henning, and A. Kalbusch. 2022. "Assessment of water consumption in households using statistical analysis and regression trees." *Sustainable Cities Soc.* 87 (Dec): 104186. https://doi.org/10.1016/j.scs.2022.104186.

Groß, M., and L. Hans. 2024. "Leveraging potentials of local and global models for water demand forecasting." *Eng. Proc.* 69 (1): 129. https://doi.org/10.3390/engproc2024069129.

Guo, G., S. Liu, Y. Wu, J. Li, R. Zhou, and X. Zhu. 2018. "Short-term water demand forecast based on deep learning method." *J. Water Resour. Plann. Manage.* 144 (12): 04018076. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000992.

Hao, W., A. Cominola, and A. Castelletti. 2024. "Combining wavelet-enhanced feature selection and deep learning techniques for multi-step forecasting of urban water demand." *Environ. Res.: Infrastruct. Sustainability* 4 (3): 035005. https://doi.org/10.1088/2634-4505/ad5e1d.

Herrera, M., L. Torgo, J. Izquierdo, and R. Pérez-García. 2010. "Predictive models for forecasting hourly urban water demand." *J. Hydrol.* 387 (1–2): 141–150. https://doi.org/10.1016/j.jhydrol.2010.04.005.

Hussien, W. A., F. A. Memon, and D. Savic. 2016. "Assessing and modelling the influence of household characteristics on per capita water consumption." *Water Resour. Manage.* 30 (Jul): 2931–2955. https://doi.org/10.1007/s11269-016-1314-x.

Iglesias-Rey, A., C. A. López-Rojas, F. J. Martínez-Solano, and P. L. Iglesias-Rey. 2024. "An approach based on the use of commercial codes and engineering judgement for the battle of water demand forecasting." *Eng. Proc.* 69 (1): 176. https://doi.org/10.3390/engproc2024069176.

Jahangir, M. S., and J. Quilty. 2024. "Sequence-to-sequence deep learning for urban water demand forecasting." *Eng. Proc.* 69 (1): 41. https://doi.org/10.3390/engproc2024069041.

Jeandron, A., O. Cumming, L. Kapepula, and S. Cousens. 2019. "Predicting quality and quantity of water used by urban households based on tap water service." *npj Clean Water* 2 (1): 23. https://doi.org/10.1038/s41545-019-0047-9.

Kley-Holsteg, J., B. Sonnenschein, G. Johnen, and F. Ziel. 2024. "Water demand forecasting based on online aggregation for district meter areas-specific adaption." *Eng. Proc.* 69 (1): 15. https://doi.org/10.3390/engproc2024069015.

Kley-Holsteg, J., and F. Ziel. 2020. "Probabilistic multi-step-ahead short-term water demand forecasting with lasso." *J. Water Resour. Plann. Manage.* 146 (10): 04020077. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001268.

Kossieris, P., I. Tsoukalas, D. Nikolopoulos, G. Moraitis, and C. Makropoulos. 2024. "Probabilistic forecasting of hourly water demand." *Eng. Proc.* 69 (1): 100. https://doi.org/10.3390/engproc2024069100.

Kulaczkowski, A., and J. Lee. 2024. "Harnessing the power of random forest for precise short-term water demand forecasting in Italian water districts." *Eng. Proc.* 69 (1): 81. https://doi.org/10.3390/engproc2024069081.

Li, J., and S. Song. 2023. "Urban water consumption prediction based on CPMBNIP." *Water Resour. Manage.* 37 (13): 5189–5213. https://doi.org/10.1007/s11269-023-03601-1.

Liu, G., D. Savic, and G. Fu. 2023. "Short-term water demand forecasting using data-centric machine learning approaches." *J. Hydroinform.* 25 (3): 895–911. https://doi.org/10.2166/hydro.2023.163.

Mazzoni, F., V. Marsili, S. Alvisi, and M. Franchini. 2024. "Detection and pre-localization of anomalous consumption events in water distribution networks through automated, pressure-based methodology." *Water Resour. Ind.* 31 (Jun): 100255. https://doi.org/10.1016/j.wri.2024.100255.

Menapace, A., A. Zanfei, and M. Righetti. 2021. "Tuning ANN hyperparameters for forecasting drinking water demand." *Appl. Sci.* 11 (9): 4290. https://doi.org/10.3390/app11094290.

Msiza, I., F. Nelwamondo, and T. Marwala. 2008. "Water demand prediction using artificial neural networks and support vector regression." *J. Comput.* 3 (11): 1–8. https://doi.org/10.4304/jcp.3.11.1-8.

Mu, L., F. Zheng, R. Tao, Q. Zhang, and Z. Kapelan. 2020. "Hourly and daily urban water demand predictions using a long short-term memory based model." *J. Water Resour. Plann. Manage.* 146 (9): 05020017. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001276.

Niknam, A., H. Khademi Zare, H. Hosseininasab, A. Mostafaeipour, and M. Herrera. 2022. "A critical review of short-term water demand forecasting tools—What method should I use?" *Sustainability* 14 (9): 5412. https://doi.org/10.3390/su14095412.

Odan, F. K., and L. F. Ribeiro Reis. 2012. "Hybrid water demand forecasting model associating artificial neural network with Fourier series." *J. Water Resour. Plann. Manage.* 138 (3): 245–256. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000177.

Pacchin, E., F. Gagliardi, S. Alvisi, and M. Franchini. 2019. "A comparison of short-term water demand forecasting models." *Water Resour. Manage.* 33 (Mar): 1481–1497. https://doi.org/10.1007/s11269-019-02213-y.

Pagano, M., G. Santonastaso, A. Di Nardo, S. Cuomo, and V. Schiano Di Cola. 2024. "Machine learning model for battle of water demand forecasting." *Eng. Proc.* 69 (1): 37. https://doi.org/10.3390/engproc2024069037.

Palomero, L., V. García, and J. S. Sánchez. 2022. "Fuzzy-based time series forecasting and modelling: A bibliometric analysis." *Appl. Sci.* 12 (14): 6894. https://doi.org/10.3390/app12146894.

Perelman, G., Y. Romano, and A. Ostfeld. 2024. "Optimizing time series models for water demand forecasting." *Eng. Proc.* 69 (1): 9. https://doi.org/10.3390/engproc2024069009.

Pesantez, J. E., E. Z. Berglund, and G. Mahinthakumar. 2020. "Geospatial and hydraulic simulation to design district metered areas for large water distribution networks." *J. Water Resour. Plann. Manage.* 146 (7): 06020010. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001243.

Pesantez, J. E., M. DiCarlo, F. Pasha, and E. Z. Berglund. 2024. "Predictive model for short-term water demand forecasting and feature analysis in urban networks." *Eng. Proc.* 69 (1): 155. https://doi.org/10.3390/engproc2024069155.

Que, Q., J. Gao, W. Wu, H. Cao, K. Li, H. Zhang, Y. He, and R. Shen. 2024. "Short-term water demand forecasting using machine learning approaches in a case study of a water distribution network located in Italy." *Eng. Proc.* 69 (1): 177. https://doi.org/10.3390/engproc2024069177.

Ramachandran, A., H. Mousa, A. Maier, and S. Bayer. 2024. "Week-ahead water demand forecasting using convolutional neural network on multi-channel wavelet scalogram." *Eng. Proc.* 69 (1): 179. https://doi.org/10.3390/engproc2024069179.

Reynoso-Meza, G., and E. P. Carreño-Alvarado. 2024. "Water demand forecasting with multi-objective computational intelligence." *Eng. Proc.* 69 (1): 79. https://doi.org/10.3390/engproc2024069079.

Ryzhkov, F. V., Y. E. Ryzhkova, and M. N. Elinson. 2024. "Python tools for structural tasks in chemistry." *Mol. Diversity* https://doi.org/10.1007/s11030-024-10889-7.

Salem, A. K., and A. A. Abokifa. 2024. "A multivariate LSTM model for short-term water demand forecasting." *Eng. Proc.* 69 (1): 167. https://doi.org/10.3390/engproc2024069167.

Sardinha-Lourenço, A., A. Andrade-Campos, A. Antunes, and M. S. Oliveira. 2018. "Increased performance in the short-term water demand forecasting through the use of a parallel adaptive weighting strategy." *J. Hydrol.* 558 (Mar): 392–404. https://doi.org/10.1016/j.jhydrol.2018.01.047.

Shabani, S., A. Candelieri, F. Archetti, and G. Naser. 2018. "Gene expression programming coupled with unsupervised learning: A two-stage learning process in multi-scale, short-term water demand forecasts." *Water* 10 (2): 142. https://doi.org/10.3390/w10020142.

Sharma, A. N., S. R. Dongre, R. Gupta, and L. Ormsbee. 2022. "Multiphase procedure for identifying district metered areas in water distribution networks using community detection, NSGA-III optimization, and multiple attribute decision making." *J. Water Resour. Plann. Manage.* 148 (8): 04022040. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001586.

Sinske, A., A. de Klerk, and A. van Heerden. 2024. "Battle of water demand forecasting: Integrating machine learning with a heuristic post process for short term prediction of urban water demand." *Eng. Proc.* 69 (1): 203. https://doi.org/10.3390/engproc2024069203.

Tian, D., C. J. Martinez, and T. Asefa. 2016. "Improving short-term urban water demand forecasts with reforecast analog ensembles." *J. Water Resour. Plann. Manage.* 142 (6): 04016008. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000632.

Ulusoy, A.-J., C. Jara-Arriagada, Y. Liu, B. Jenks, and I. Stoianov. 2024. "Interpretable AI for short-term water demand forecasting." *Eng. Proc.* 69 (1): 101. https://doi.org/10.3390/engproc2024069101.

Walski, T. M., et al. 1987. "Battle of the network models: Epilogue." *J. Water Resour. Plann. Manage.* 113 (2): 191–203. https://doi.org/10.1061/(ASCE)0733-9496(1987)113:2(191).

Wang, D., Y. Li, B. Hou, and S. Wu. 2024. "Short-term water demand forecasting based on LSTM using multi-input data." *Eng. Proc.* 69 (1): 103. https://doi.org/10.3390/engproc2024069103.

Wunsch, A., C. Kühnert, S. Wallner, and M. Ziebarth. 2024. "Urban water demand forecasting using DeepAR-models as part of the battle of water demand forecasting (BWDF)." *Eng. Proc.* 69 (1): 25. https://doi.org/10.3390/engproc2024069025.

Xenochristou, M., and Z. Kapelan. 2020. "An ensemble stacked model with bias correction for improved water demand forecasting." *Water Resour. Manage.* 17 (3): 212–223. https://doi.org/10.1080/1573062X.2020.1758164.

Yao, Y., H. Liu, F. Gao, H. Guo, and J. Zou. 2024. "Short-term urban water demand forecasting using an improved NeuralProphet model." *Eng. Proc.* 69 (1): 175. https://doi.org/10.3390/engproc2024069175.

Yu, J., H. Bae, M.-S. Kang, K.-J. Kim, and I.-S. Jang. 2024. "A study on short-term water-demand forecasting using statistical techniques." *Eng. Proc.* 69 (1): 154. https://doi.org/10.3390/engproc2024069154.

Zanfei, A., A. Menapace, B. M. Brentan, and M. Righetti. 2022a. "How does missing data imputation affect the forecasting of urban water demand?" *J. Water Resour. Plann. Manage.* 148 (11): 04022060. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001624.

Zanfei, A., A. Menapace, F. Granata, R. Gargano, M. Frisinghelli, and M. Righetti. 2022b. "An ensemble neural network model to forecast drinking water consumption." *J. Water Resour. Plann. Manage.* 148 (5): 04022014. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001540.

Zanutto, D., et al. 2024. "A water futures approach on water demand forecasting with online ensemble learning." *Eng. Proc.* 69 (1): 60. https://doi.org/10.3390/engproc2024069060.

Zounemat-Kermani, M., E. Matta, A. Cominola, X. Xia, Q. Zhang, Q. Liang, and R. Hinkelmann. 2020. "Neurocomputing in surface water hydrology and hydraulics: A review of two decades retrospective, current status and future prospects." *J. Hydrol.* 588 (Sep): 125085. https://doi.org/10.1016/j.jhydrol.2020.125085.

J. Water Resour. Plann. Manage., 2025, 151(10): 04025049