Utilities Policy 82 (2023) 101559

Contents lists available at ScienceDirect

Utilities Policy

journal homepage: www.elsevier.com/locate/jup

Full-length article

Performance analytics for regulation in retail water utilities: Guiding asset management by identifying peers and targets

Hermilio Vilarinho[a,*], Giovanna D'Inverno[b], Henriqueta Nóvoa[a], Ana S. Camanho[a]

[a] Faculty of Engineering, University of Porto, Porto, Portugal
[b] Department of Economics and Management, University of Pisa, Pisa, Italy

ARTICLE INFO

Keywords:
Data envelopment analysis
Asset management
Benchmarking
Water regulation
Benefit-of-the-doubt
Composite indicator

ABSTRACT

This research evaluates the performance of water supply utilities operating at the retail level in Portugal concerning asset management practices. The study's main innovative feature is identifying peers and targets to guide improvements in the sector. Reliable data collected by the regulatory authority for water and waste services in Portugal (ERSAR) are employed to design two composite indicators reflecting different dimensions of asset management: operational conditions and management systems. Based on the Data Envelopment Analysis technique, the Benefit-of-the-Doubt model is employed in robust and conditional formulations. The role of the context on utilities' performance is also investigated. The results show that the direct management model is unfavourable concerning developing structured management systems, whilst urban environments favour managerial advancement. Rural and semi-urban environments favour "good" operational results in infrastructures. The pool of peers obtained for each utility and the quantification of targets based on the observed achievements by those peers facilitates the search for industry best practices and promotes continuous improvement. Given the high heterogeneity in asset management performance within the sector, the utility-specific target-setting approach illustrated in this paper can support a regulatory policy review for determining more realistic goals.

1. Introduction

Providing access to clean water is of utmost importance for the health and well-being of all individuals. Water is essential for human life and increasingly needed to produce energy, generate food, manufacture products, and provide services. According to the latest Global Water Security Report, the percentage of the world's population using safely managed drinking water increased from 70% to 74% between 2015 and 2020. However, this rate is insufficient to meet the target set by the United Nations Sustainable Development Goals of achieving universal access to water by 2030. Two billion people were still lacking access to water services in 2020 (United Nations, 2022). This problem is further exacerbated by climate change, which brings more frequent and severe droughts, floods, and increasing sea levels. Additional challenges include population growth, urbanisation, and lack of infrastructure maintenance and management. Consequently, it is necessary for the water sector to become more resilient so that it can withstand shocks and stresses (Lombana Cordoba et al., 2022).

These challenges require that water infrastructures are maintained in reliable conditions. Unfortunately, the majority of water systems are in severe disrepair due to the short-term and narrow focus strategies used in the sector, which have resulted in the deferment of necessary investments. According to a recent report by EurEau - the European Federation of National Associations of Water Services representing 30 European countries, the rate of water loss in Europe was found to be 25.1% (EurEau, 2021). The report also revealed substantial discrepancies in the rates among member countries, with the Netherlands reporting the lowest rate at 5%, while Bulgaria had the highest rate at 61%. According to the American Association of Civil Engineers (ASCE, 2021), water mains break every two minutes in the United States, resulting in the loss of 6 billion gallons of treated water daily. This volume is equivalent to filling more than 9,000 swimming pools, and equates to US$ 7.6 billion lost in 2019.

Urban water infrastructures are capital-intensive, expensive, long-lasting, and exclusive assets, which cannot be shared by multiple

* Corresponding author.
E-mail addresses: up202001529@edu.fe.up.pt (H. Vilarinho), giovanna.dinverno@unipi.it (G. D'Inverno), hnovoa@fe.up.pt (H. Nóvoa), acamanho@fe.up.pt (A.S. Camanho).

https://doi.org/10.1016/j.jup.2023.101559
Received 29 January 2023; Received in revised form 16 April 2023; Accepted 16 April 2023
Available online 27 April 2023

service providers, and represent a significant portion of municipal public assets' value. Asset management, a modern expression for a centuries-old practice that focuses on managing infrastructure assets, has emerged as a more comprehensive and well-devised strategic approach over the past few decades (Amaral et al., 2017). It represents a potential solution to deal with water infrastructures, ensuring the economic health and welfare of modern communities (Alegre, 2010). Formalised since 2014 in a series of international standards, the ISO 55000 (ISO, 2014), asset management involves an extensive system that requires organisations to balance cost, risk, performance and life cycle and extract the maximum value from their physical assets. Almeida et al. (2021) discuss that asset management techniques and principles allow an organisation to structure a governance model to achieve sustainable levels of service and performance. To ensure integration and successful implementation, asset management should be grounded in plan-do-check-act (PDCA) principles and divided into three levels of planning: strategic, tactical, and operational. At each level, defined objectives, assessment criteria and targets, diagnosis, action plan development, and implementation are key activities that should be undertaken for effective asset management (Alegre, 2010).

Benchmarking is a common practice in organisational management used to evaluate processes against best practices of peer entities in an industry or sector. When the best-in-class entities are identified, the managers are able to set targets that enable them to learn from others, measure their performance and guide improvements. This research is aimed at performing a benchmarking exercise with a set of Portuguese water utilities operating at the retail level, by focusing on their asset management practices.

Asset management applications were introduced in the Portuguese water sector in the beginning of century XXI, and even though some improvement effort has been carried out, the results are not uniform among all operators (Luís and Almeida, 2021). Different ownership and management structures coexist, resulting in significant heterogeneity in governance mechanisms and asymmetrical access to funding that are necessary to cope with necessary investments. More than 200 utilities operate in the retail market and most of them are directly controlled by municipalities. Those service providers are responsible for the storage and distribution of treated water received upstream by larger-scale utilities operating the bulk market. The bulk or wholesale utilities extract, treat and distribute water to the retail market in Portugal.

In Portugal, the regulatory agency ERSAR (acronym in Portuguese for Water and Waste Services Regulation Authority) annually gathers a vast collection of metrics that are suitable for benchmarking purposes. Several of those metrics reflect asset management factors that can be employed to assess the performance of the retail water operators. These indicators are used as input data for this research.

The Benefit-of-the-Doubt (BoD) technique based on Data Envelopment Analysis (DEA) is applied to construct composite indicators, in order to measure the performance of the various utilities in two dimensions: assets' condition and managerial features. Additionally, the role of the context in which those utilities operate was also taken into consideration in the benchmarking study. Finally, suitable peers and targets for operators are determined.

The relevance of the study relies on the critical conditions presented by the Portuguese water sector in terms of asset management and the urgent needs to foster improvements in this area. The use of ERSAR metrics to perform benchmarking studies in the retail water market in Portugal has been explored by several studies (Marques, 2006; De Witte and Marques, 2010; Henriques et al., 2022; Mergoni et al., 2022; Pinto et al., 2017a,b; Amaral et al., 2022, 2023). The study developed by Vilarinho et al. (2023) focused exclusively on asset management practices, but covered only the bulk market. The development of a methodology to identify the most appropriate benchmark counterparts and targets for a set of water retailers concerning asset management practices has yet to be covered in the literature, representing the main contribution of this paper.

The remainder of the paper is divided into the following sections: Section 2 provides a short literature review, Section 3 explains the methodology, Section 4 gives details about the case study, Section 5 displays the results and discusses the findings, and, finally Section 6 presents the conclusions.

2. Literature review

The literature review includes a discussion about benchmarking practices using DEA applications in retail water utilities (subsection 2.1), the characterisation of the Portuguese water sector (subsection 2.2), and a discussion of asset management features of the water sector in Portugal (subsection 2.3).

2.1. Benchmarking practices in retail water utilities

Benchmarking is a common practice used to compare performance with standards, aiming to identify areas for improvement. Metric benchmarking employs indicators to measure an entity's performance over time and compare it to its peers. Entities can thereby evaluate how they measure up against industry standards or observed practices of peers, track progress toward goals, explore best practices, and optimise operations and resource utilisation. Developing a reliable benchmarking method enables spotlighting the better and worse performing service providers, setting incentives on organisation performance and offering visibility on the processes and mechanisms that work and those that do not work (Mumssen et al., 2018).

Marques and De Witte (2010) detail the vital role of benchmarking practices in the performance the water sector. Regulators often evaluate the efficiency of utilities using a set of practices known as "yardstick competition" that is utilised when direct competition between public service providers is not possible. The main idea of yardstick competition is to compare the performance of service providers in the same sector creating an artificial competition between them. The key advantages of yardstick competition include incentives to boost information sharing and openness, as well as efficiency, innovation, and quality of service. As a result of yardstick competition, the knowledge acquired from other utilities is used to redirect the incentive of the utility under examination to enhance its efficiency. Yardstick competition is performed in the water sector using two approaches. The first one, known as price yardstick competition, employs benchmarking practices to set tariffs. The second approach to yardstick competitions is often softer and involves mandatory benchmarking mixed with open disclosure of performance data with no relation to price setting. This lighter mode of yardstick competition is known as sunshine regulation and has been implemented in many countries, such as Australia, Argentina, Holland, Denmark and Portugal. Sunshine regulation became popular in the water sector, and is sometimes employed as an initial step toward more demanding and tighter regulation processes (Marques, 2006).

In a benchmarking context, a systematic process can be helpful to estimate efficiencies and obtain by-products of the measurement exercise corresponding to targets for inputs and outputs and peers that serve as benchmarks for each utility. Data Envelopment Analysis (DEA) models can address these requirements. Among the techniques available, Marques (2006) championed using DEA as the most consensual and widespread approach for evaluating water systems. Non-parametric approaches, including DEA, differ from those using engineering standards or production functions with conceptually stipulated functional forms. DEA, established initially by Charnes et al. (1978), is a data-driven, non-parametric approach that evaluates performance compared to best practices identified across a group of units known as DMUs (Decision Making Units). The use of DEA identifies an efficient best-practice frontier, and inefficient units are rated based on their distance from that frontier. The calculations are performed using linear programming models to identify the optimal weights applied to inputs and outputs, from which the efficiency scores are obtained.

The selection of an appropriate reference set is of critical importance when conducting benchmarking activities. It requires organisations to identify a peer group in their sector or industry that presents proper performance measures from which to learn. DEA-based applications facilitate the establishment of best practices and benchmarks, as ultimately DEA analyses offer information on both target setting and peer identification. Methodologically, in DEA, for a particular DMU under assessment, only a section of the DEA efficiency frontier should be considered the common best-practice frontier. This common best-practice frontier will be the facet of the DEA efficiency frontier spanned by a set of technically efficient DMUs, which can be seen as a common reference group. This reference set represents the DMU's peers. Targets will then result from projections of this DMU toward the common best-practice frontier. Selecting the peer set that provides the closest targets ensures the identification of the globally most similar best practices. Therefore, the DMU can identify the easiest way to improvement (Ruiz and Sirvent, 2016, 2022). Thanassoulis et al. (2008) explains that targets represent the levels of inputs and outputs that render a DMU efficient. These authors highlight that, by focusing on observed operating practices, DEA tends to be very useful in providing a starting point for setting performance targets. Simplified and interactive procedures incorporating user preferences have been applied to target identification. Their relevance lies on improving nontechnical users' comprehension of the evaluation process which supports the organisational learning (Pereira et al., 2021).

DEA applications can be used to aggregate several metrics in the form of a *composite indicator (CI)*, which is known as the "Benefit-of-the-Doubt" (BoD) approach. This strategy was proposed to evaluate macroeconomic performance by Melyn and Moesen (1991) and popularised by Cherchye et al. (2007). Zanella et al. (2015) explain that there are only output measures to be aggregated in a BoD, so all DMUs are assumed to be similar regarding the inputs. Thus, a unitary input is considered in the BoD as opposed to a standard DEA linear programming model that presents inputs and outputs. Being based on DEA, the BoD method is data-driven and avoids the need for consultation with stakeholders to determine the aggregation weights for the individual metrics. Additionally, since the weights generated as outcomes of a BoD model can handle the conversion of units, the metrics can be employed using their own units of measurement, avoiding the need for normalisation. BoD models are also used to identify peers (Lavigne et al., 2019; Zanella et al., 2013; Morais and Camanho, 2011) and targets (Wüst and Rogge, 2021; Pereira et al., 2021).

The efficiency of water and wastewater operators has been vastly explored in the literature with studies performed worldwide, including Australia (Byrnes et al., 2010), Brazil (Tourinho et al., 2022a,b), Canada (Wang et al., 2018), China (Dong et al., 2018), Italy (Romano and Guerrini, 2011; Lo Storto, 2018; D'Inverno et al., 2021), Japan (Marques et al., 2014), Palestina (Alsharif et al., 2008), Peru (Berg and Lin, 2008), United Kingdom (Walker et al., 2019; Thanassoulis, 2000a,b) among others. Several works using DEA have investigated the Portuguese water sector (Marques, 2006; De Witte and Marques, 2010; Henriques et al., 2022; Mergoni et al., 2022; Amaral et al., 2022). Literature reviews covering benchmarking practices in water systems can be found in Berg and Marques (2011), that analysed 190 benchmarking studies using quantitative methods and Goh and See (2021), that reviewed 142 articles published between 2000 and 2019 on that subject.

The DEA approaches available in the literature include the computation of robust efficiency scores to minimise the effect of outliers and robust conditional efficient scores, that allow statistical inference and adjust the scores produced according to the environment. The robust and robust conditional approaches have been widely employed to evaluate water systems (e.g. De Witte and Marques 2010, Mbuvi et al. 2012, Marques et al. 2014, D'Inverno et al. 2021, Mergoni et al. 2022). The effect of the context is characterised by a separate set of data from variables that do not enter directly in the computation of the scores but are used to guide the sampling process of the DMUs under evaluation. Carvalho and Marques (2011), in their study of overall performance measurement in 66 Portuguese water utilities, explain that the influence of the operational environment on efficiency must be taken into account. The comparison between water utilities operating under highly diverse contexts should be avoided. Therefore studies that do not adjust the efficiency measurement to the context in which the utilities operate can lead to unrealistic scores.

In a fragmented and heterogeneous market that is typical from the retail water sector, benchmarking exercises based on DEA/BoD models present the ideal fit for sunshine regulation practices, given their capacity to consider the environment in which utilities operate, identify genuine reference peers, and suitable performance targets. Those features support the selection of those tools for this study. Another benefit that reduces the chance of complaints in case of undesirable results is the flexibility that DEA-based techniques offer to determine the most favourable weights for each DMU.

2.2. *The water market in Portugal*

The Portuguese water sector experienced a complete structural reconfiguration by implementing new public policies for water and waste services initiated in 1993. Since then, Portugal has suffered a substantial transformation in social well-being, with relevant impacts on the environment and public health. Baptista (2014) describes that the public water systems served only 81% of the homes on Portugal's mainland in 1993. Regarding water quality, just 50% of the population was supplied with safe water according to national and European legislation. The service currently covers 96% of residential units with a quality above 99% (ERSAR, 2021a,b). Despite the notable geographical discrepancies between urban and rural regions, as roughly 99% of urban residences have access to public water supply services, compared to less than 90% in rural areas, the water sector reforms in Portugal represent an outstanding achievement (Baptista, 2014). Paul Reiter, former Executive Director of the International Water Association (IWA), has referred to this success as the 'Portuguese miracle'. The progress can be attributed to establishing a coherent public policy, implementing major reforms in the legal and institutional frameworks, and practising sound strategic planning (Alegre et al., 2020). This implementation involved an overall perspective integrating various components, such as strategic planning, legislation, institutional framework, governance systems, introducing competition,[1] access targets and quality of service goals, tariff and tax policy, labour force qualification, information publishing, promotion of research and development, and construction of the infrastructure (Baptista, 2014).

The rising financial inflows that supported the structural transformations of the Portuguese water sector were motivated by the entry to the European Union in the 1990s. The European integration clearly accelerated the reversal of the state's authority over the financial sector, stimulating the creation of a private banking system. The wave of privatisations across the Portuguese economy was led by the development of capital markets in the new robust banking sector, and significant investments were made in the water sector with the help of external financing (Teles, 2015). During the time of the Strategic Plan for Water Supply and Wastewater Services (PEAASAR I) from 2000 to 2006, Portugal invested between 5 and 6 billion euros in construction, expansion, or rehabilitation of infrastructure for water supply and wastewater treatment (Alegre, 2010).

Until 1993, the local municipalities were exclusively responsible for the provision of water. The only exception was the state-owned

[1] Given the natural monopoly characteristics of the water industry, the concept of competition needs to be clarified. Baptista (2014) refers to it as "virtual competition". The benchmarking among utilities as well as the introduction of different models of governance have enabled competition to increase and ultimately the efficiency and quality of the services to improve.

utility EPAL (Empresa Pública de Águas de Lisboa), which supplied Lisbon. The Decree-Law no. 372/93 instituted the participation of private capital in the sector through concessions. The sector's property remained with the State, but, in many cases, the management was given to the private sector, which was supposed to bring more investments, mainly from European funds (Pato, 2011). Water supply and wastewater management were divided into bulk and retail services as part of the sector's corporatization process. Bulk or wholesale companies are capital-intensive and multi-municipal. They include water abstraction, treatment, lifting, and abduction, while retail services include storage and final distribution to end-consumers. The retail utilities are also in charge of tariff setting and collection.

Regarding wastewater services, bulk companies are responsible for wastewater elevation, transport, treatment, and disposal (ERSAR, 2021a). The municipalities remained as minor shareholders of the multi-municipal bulk companies, but their actual control was then limited to the retail sector. The central state concentrated investment efforts in the bulk sector, so several municipal concessions were created to enable the entry of private capital to support the needs of retail systems. Under those concessions, celebrated as public–private partnerships, the municipalities leveraged the investment capacity without jeopardising their control. The rural population in Portugal's most remote locations was the socioeconomic category that, in relative terms, gained the most from the water sector's investments. Moreover, the introduction of private capital to enable investments in retail utilities benefited large construction companies that received generous contracts for infrastructure projects and, in many cases, managed to acquire the retail concessions (Teles, 2015).

There are now three main options available for managing Portuguese water utilities: direct management, delegation and concession. Municipalities and associations of municipalities control and run the water services under the direct management model, often without the involvement of private businesses. The delegation model is applicable to parishes, user organisations, municipal companies or companies created in collaboration with the State (municipal or state utilities). Without a concession agreement, the State (central, local or both) owns and controls the utility directly under the delegation system. In this case, a contract of management must be signed, defining goals and tariff policies for the operator. In the concession model, a public–private partnership with municipalities and other private operators is created under a long-term contract, often ranging from 30 to 50 years. Private capital may participate primarily through the delegation and concession models, and subsequently through direct management in cases of partnerships with the government or local governments (Marques and Berg, 2011; Pérez et al., 2019; ERSAR, 2021a).

Carvalho and Marques (2016) explain that the water and wastewater sectors in Portugal present a clearly unique market structure. Only a few nations, like Belgium, The Netherlands, and Romania, have separate wholesale and retail marketplaces. The retail water sector in Portugal is highly fragmented, with a large number of utilities, which is partially explained by the fact that the municipalities handle the majority of the services. The direct management of water provision is currently adopted by 158 municipalities (68% of the total), but these utilities cover only 26% of the population, being more frequent in rural areas with lower population densities. Another type of direct management occurs when a self-managed utility is created under the ownership of one or more municipalities. This model covers 22% of the population in Portugal. Although the utilities using the direct management model still prevail in the retail water sector, there has been a trend toward corporatisation of the sector in the last two decades. At the beginning of the 2000s, the concession and delegation management models counted only for 20% of the population, while today, they account for around half, more than doubling their share in the sector (ERSAR, 2021a).

Various research projects have looked at the market structure of the Portuguese water sector. Marques (2008), Correia and Marques (2011) and Marques and Simões (2020) studied Portugal's general efficiency of public and private utilities. In all those studies, the results favoured private utilities compared to public ones. The possibility of scale economies in eventual mergers and scope economies by integrating water and wastewater systems have been also examined. Correia and Marques (2011) found increasing returns of scale and decreasing economies of scope, suggesting that there are no advantages in the joint production of water and wastewater activities. Marques and De Witte (2011) concluded that the number of retail water utilities in Portugal should be reduced from more than 200 to around 60 to operate at the optimal scale. As a result, each utility should serve an average population of between 160,000 and 180,000 people.

Regarding scope, this study did not recommend joint activities of water and wastewater by the same utility. Pinto et al. (2017b) identified 40,000 customers as the optimal scale for water utilities. Carvalho and Marques (2016) and Marques and Carvalho (2014) also pointed out opportunities for economies of scale. Moreover, these studies identified some opportunities for merging bulk and retail operators and water and wastewater activities. Carvalho and Marques (2014) concluded that there are economies of vertical integration between wholesale and retail activities and economies of scale in water utilities. However, diseconomies of scope were found, suggesting that the utilities should choose only one specialisation between water and wastewater activities. Marques and Berg (2011) investigated how regulatory contracts for infrastructure deal with risk. They concluded that risk is a major concern when the public and private sectors collaborate and must be addressed in regulatory contracts. Tariff structure (Pinto and Marques, 2015; Marques and Berg, 2011; Martins et al., 2020, 2013; Gonçalves et al., 2014; Silvestre and Gomes, 2017), quality of service (Pinto et al., 2017a; Duarte et al., 2009) and sustainability (Pérez et al., 2019; Mergoni et al., 2022) have also been relevant themes of study in the water market in Portugal.

The establishment of a regulatory entity for the sector made mandatory the use of market-oriented management practices. The Water and Waste Services Regulation Authority (ERSAR), the regulating body for the whole water and waste industry, was created in 2009 after being founded in 1995 as the Supervisory Commission for Concessions. ERSAR's role rests on the idea that a natural monopoly should be controlled to guarantee proper protection to costumers, but keeping the market efficiency (Santos et al., 2018). Sunshine regulation has been adopted by ERSAR as an incentive for the utilities to improve their performance and has been addressed by several studies (e.g. Gonçalves et al., 2014; Marques, 2006; Marques and Pinto, 2018; De Witte and Marques, 2010; Cardoso et al., 2012). Following the sunshine regulation model, a set of comprehensive performance measures is established and collected by ERSAR from utilities operating in the sector, and their outcomes are made available to the public. Since the regulator is not actively involved in the pricing formulation process, ERSAR's authority is not coercive (Gonçalves et al., 2014). In Portugal, sunshine regulation can be a particularly suitable approach due to the high inefficiency levels and the fragmented structure of the Portuguese market. Besides that, this practice can help minimise the existing political interference in the sector and increase transparency. Portugal faces the challenge of improving the performance of its utilities, which cannot be accomplished solely by publicising performance indicators (Marques, 2006). Thus, a structured methodology is needed to tackle this challenge, which supports the relevance of this study.

2.3. Asset management practices in the Portuguese water sector

The sustainable management of water infrastructure in Portugal has become a prominent issue in recent years and has resulted in various measures. The Decree-Law 194/2009, effective in 2013, required the existence of an asset management system in all water supply services and urban wastewater management services serving 30,000 people and above. In response to this law, ERSAR, jointly with LNEC (*Laboratório Nacional de Engenharia Civil* - National Civil Engineering Laboratory) and the Technical University of Lisbon, released technical guidelines that described a framework for integrated asset management. Several relevant research and collaboration projects have been conducted at the national and international levels as a result of LNEC's active leadership in asset management research, development, training and awareness efforts. Additionally, many utilities were used to test and design a decision support software for asset management. Portugal has hosted several conferences, seminars, courses and meetings on this subject. There has also been an intense activity on academic training, as shown by the recent development of multiple master and doctorate dissertations on this topic at Portuguese universities. For more detail on this process, see Matos and Baptista (1999), Alegre (2010), Leitão et al. (2016) and Amaral et al. (2017).

Alegre et al. (2020) highlighted that the primary goal of the reform process, started in 1993, was the creation of new infrastructures to improve the availability and quality of services. However, in recent years the focus has been shifted toward the value maximisation of existing infrastructures in a long-term perspective to ensure sustainable service delivery. The massive investment of 13 billion Euros from 1993 to 1999 was mainly applied to bulk systems, which is noteworthy given the country's population of approximately ten million. The significant asset portfolio generated by this spending has a high value, although some assets are too old, complex and demanding in management. Therefore, effective asset management is a priority to ensure that the value of these assets is maintained and sustainable water services are provided.

In 2015 a new strategic plan for the water sector, the PENSAAR 2020 (*Plano Estratégico de Abastecimento de Água e Saneamento de Águas Residuais* - Strategic Plan for Water Supply and Wastewater Sanitation 2020), was launched bringing the management of the sectors' assets to the centre of the discussion. As explicitly stated in the Plan: "The strategy should be less centred in new infrastructures to increase the served population and focuses more on the management of the sector assets, its operations and the quality of the provided services with an overall sustainability" (Frade et al., 2015). The plan determines five strategy axis, being Axis 3 dedicated to the optimisation and efficient use of the existing resources. This axis establishes six operational objectives, as follows: (i) optimisation of the installed capacity use and increase of service adhesion; (ii) reduction of physical water losses; (iii) control of rainwater to foul sewerage; (iv) efficient management of assets and rehabilitation increase; (v) upgrade resources and sub products; (vi) allocation and efficient use of the water resources.

The control of water losses, one of the objectives of Axis 3 of PENSAAR 2020, is commonly researched in this field. According to the study conducted by EurEau (2021), Portugal experiences a high rate of water losses, with an estimated 30% of the total water supply being lost. Marques and Monteiro (2001) indicated a critical low level of asset rehabilitation and non-existence of preventive maintenance as the responsible for the considerable volume of water losses. Those authors suggest a set of indicators to monitor and control water losses. Marques and Monteiro (2003) also reinforce that the high volume of water losses in Portugal is associated with the focus on building new assets instead of giving more attention to the existing systems' operation and maintenance. This study also recommends the application of performance indicators to control losses. The minimisation of water losses is discussed by Machado et al. (2009) that reports a case study in a bulk water system. The use of energy resources is the focus of the research conducted by Loureiro et al. (2020) that proposed a comprehensive framework assessment for energy efficiency and concluded that energy inefficiencies are related to water losses or network layout, not to pumping inefficiencies.

The rehabilitation of water assets is considered vital in increasing the efficiency of water utilities. Ferreira and Carriço (2019) analysed practical applications of asset management approaches by comparing alternatives for rehabilitation strategies employing performance indicators. It was found that decreasing proactive management spending may result in future problems and unanticipated costs. A case study describing rehabilitation of infrastructures in a utility in the Algarve region is described by Cabral et al. (2019), and the results indicate that the assets' economic valuation accuracy is essential to determine a rehabilitation strategy. The application of a performance assessment framework for water systems tested in two Portuguese retail water utilities by Santos et al. (2022) identified vulnerable areas to flooding and the need for rehabilitation investments. Carriço et al. (2012) developed a methodology to prioritise rehabilitation interventions, using the technique ELECTRE III.

The current situation underlined by PENSAAR 2020 displays an inadequate rehabilitation rate, lack of asset knowledge and difficulties in ensuring cost recovery. For the current rehabilitation rates to be sustainable, pipes would need to last, on average, 100 and 200 years for water and wastewater networks, respectively. There are also serious problems of economic and financial sustainability. Over 3.5 million people, or 33% of the country's population, are served by utilities that do not ensure cost recovery. A large number of utilities are not able to quantify the actual cost of their services. The strategic plan also addresses new tariff regulations and utility mergers (Amaral et al., 2017).

The development of a strategy to implement effective asset management systems is also at the core of PENSAAR 2020's Axis 3. Those methodologies should complement and support the approach initiated by ERSAR and LNEC. In that sense, the structured procedure developed by Cardoso et al. (2012) include elements of strategic, tactical and operational planning. It was tested in four operators with different characteristics, focusing on the diversity of the utilities to ensure flexibility.

Leitão et al. (2016) presented the results of a collaborative project led by LNEC comprising asset management system implementations in 19 retail water utilities, covering different sizes, management models and scope (water, wastewater, storm water). The utilities took advantage of the simultaneous implementation process by sharing difficulties and solutions, and at the end, they could successfully develop their own strategic and tactical plans. This process proved to be successfully suited for the water industry scenario in Portugal and many of the strategic and tactical plans developed were actively applied to the systems.

The water systems in Portugal show highly diverse results in terms of asset management performance. The results of a survey conducted by the *Specialised Commission for Asset Management* from the *Portuguese Association for Water Distribution and Drainage (Associação Portuguesa de Distribuição e Drenagem de Águas - APDA)* in 2019, using data from bulk and retail utilities, indicate that asset management practices are not used by 54% of those utilities. Asset management goals are not established by 41% that declare to have implemented an asset management system. Besides that, 57% of those utilities do not dedicate personnel exclusively to asset-related activities. Many of those utilities do not perform preventive maintenance, do not analyse their assets' condition and record their data on paper and spreadsheet records (APDA, 2019). Amaral et al. (2017) mention the highly fragmented market structure, the politicised nature of municipal water utility management and the existing accounting procedures as some of the main barriers to spreading asset management best practices. When discussing the applicability of asset management to small and medium utilities, Alegre (2010) reinforces the option to establish realistic targets and network connections with relevant peers for sharing problems and solutions.

Benchmarking studies have been undertaken in Portugal employing asset management elements. Santos et al. (2022) performed a comparison assessment of two Portuguese retail utilities, using a multidimensional performance framework, where infrastructural sustainability is one of the examined dimensions. The results show the potential of an assessment framework to support planning and monitoring of activities and investments.

Targets for asset management were also proposed by Ferreira and Carriço (2019), that performed a case-study in a water supply system in Lisbon. This study evaluated the operator's performance in fulfilling the proposed tactical objectives by the use of thirteen metrics.

The comprehensive set of metrics that ERSAR annually requests from the sector's retail operators under the sunshine regulation strategy enables a multifaceted assessment of utilities' performance. Those metrics were already employed to undertake benchmarking studies in the literature. Pinto et al. (2017a,b) used those metrics to evaluate *quality of service*; (Henriques et al., 2020) assessed the *general performance of wastewater operators*, Mergoni et al. (2022) evaluated *environment achievements* and Amaral et al. (2022) addressed the *techno-economic efficiency* of wastewater utilities using ERSAR's metrics. The study developed by Vilarinho et al. (2023) selected metrics related to asset management practices to construct composite indicators following the BoD approach. The role of the environment was also examined including contextual variables. However, that study focused on wholesale utilities, a different market, and emphasised the progress of utilities along a five-year period. This study aims to extend the developments of that research by focusing on the retail water market and the role of the context is also explored. More importantly, the use of ERSAR's metrics to identify peers and targets for the retail water operators in asset management performance represents the main innovative contribution of this work. The relevance of this study relies on the need for immediate actions due to the unsatisfactory water infrastructure conditions, both in Portugal and worldwide, from which considerable room for improvement can be noticed.

3. Methodology

The proposed methodology includes three stages covering the development of the BoD methods employed in the study. First, the standard computation of the CIs using a deterministic approach is presented. The techniques employed to determine the peers and targets are explained in the second stage. Finally, the robust and conditional methods are discussed.

3.1. Calculation of the standard deterministic composite indicator

This subsection explains how the composite indicator can be computed using the standard deterministic BoD model. The standard deterministic CI, which is the baseline method used to calculate the composite indicators, is described in this subsection. BoD linear programming models and the metrics aggregated as outputs are employed to generate the CIs.

The BoD Model (1) is used when only desirable metrics are aggregated. Desirable metrics are the ones that are targeted to increase, so better performance results correspond to higher values. On the contrary, lower values are preferable for undesirable metrics.

$$\begin{aligned}
\text{maximise} \quad & \beta_{j_0} \\
\text{subject to} \quad & \textstyle\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{rj_0} + g_y\beta_{j_0} \qquad r = 1,\dots,s \\
& \textstyle\sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0, \qquad j = 1,\dots,n \\
& \beta_{j_0} \in \mathbb{R}
\end{aligned} \tag{1}$$

The BoD Model (2) based on the Directional Distance Function (DDF) and introduced by Zanella et al. (2015) is employed to handle both desirable and undesirable metrics.

$$\begin{aligned}
\text{maximise} \quad & \beta_{j_0} \\
\text{subject to} \quad & \textstyle\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{rj_0} + g_y\beta_{j_0} \qquad r = 1,\dots,s \\
& \textstyle\sum_{j=1}^{n} b_{kj}\lambda_j \leq b_{kj_0} - g_b\beta_{j_0} \qquad k = 1,\dots,l \\
& \textstyle\sum_{j=1}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0, \qquad j = 1,\dots,n \\
& \beta_{j_0} \in \mathbb{R}
\end{aligned} \tag{2}$$

BoD Models (1) and (2) are presented in their envelopment formulation, often employed in peer identification for benchmarking purposes. In the BoD models, y_{rj} represents the desirable metrics, whereas b_{kj} represents the undesirable ones. r is an index for desirable metrics, ranging from 1 to the total number of desirable metrics s, while k represents each undesirable metric, ranging from 1 to the total number of undesirable metrics l. The parameters y_{rj_0} and b_{kj_0} are the values of desirable and undesirable metrics observed for the DMU j_0 under assessment.

The BoD model must be solved n times, where n represents the number of assessed DMUs. For each DMU under evaluation denoted as j_0, the values of the decision variables λ_j and β_{j_0} are obtained as the solution of the BoD model. The variable β_{j_0} represents the factor by which the desirable metrics should proportionally increase and the undesirable metric should proportionally decrease toward the best-practice frontier. Note that the model's objective function aims to maximise β_{j_0}, by finding the optimal results for the DMU under assessment. As discussed by Lavigne et al. (2019), the values of λ_j identify how relevant other DMUs are for representing the benchmark against the DMU under assessment. Therefore, λ_j different from zero identify the peers; the higher their values the more relevant the peer is.

The direction of expansion of the desirable metrics and reduction of the undesirable ones is indicated by the Directional Distance Vector defined as $(g_y, -g_b)$. The direction vector used in DEA and BoD models is a crucial factor that can impact the calculated scores. To address this issue, various solutions have been proposed in the literature, depending on the research objectives. Fried et al. (2008) and Rogge et al. (2017) have discussed different options for selecting the direction vectors in DEA and BoD models. In this study, the values of $(g_y, -g_b)$ were used as $(y_{rj_0}, -b_{kj_0})$, following Zanella et al. (2015) and Rogge et al. (2017), so that each DMU may guide its improvement using the values of its own performance metrics. This results in a proportional interpretation of the composite indicator value.

Since the maximum feasible level of β_{j_0} is obtained by optimisation, DMU j_0 under assessment is given the best possible results. The CI for j_0 is calculated as $1/(1+\beta_{j_0})$. The best-performing DMUs are located in the best-practice frontier, meaning that for those DMUs neither the reduction of undesired metrics nor the expansion of desirable metrics is required. For those instances, the obtained score for β_{j_0} equals zero, and for CI_{j_0} is equal to 1. For all the other cases in the deterministic approach, β_{j_0} is a positive number, meaning that CI_{j_0} ranges from 0 to 1.

3.2. Determination of peers and targets for benchmarking

This subsection explains how the peers and targets are obtained in the standard deterministic BoD model. The first set of constraints, one for each s desirable metric, in Models (1) and (2) are shown in expression (3).

$$\sum_{j=1}^{n} y_{rj}\lambda_j \geq \underbrace{y_{rj_0} + g_y\beta_{j_0}}_{Target} \qquad r = 1,\dots,s \tag{3}$$

For each r desirable metric, the right-hand side term of expression (3) is the sum of the observed desirable metric of the DMU under assessment y_{rj_0} and its expansion toward the best-practice frontier $g_y\beta_{j_0}$. Therefore we can say that $y_{rj_0} + g_y\beta_{j_0}$ defines the target for each desirable metric that DMU j_0 should have to reach the best-practice frontier.

Following the same rationale, the targets for the undesirable metrics are displayed in the set of l constraints in (4) taken from Model (2). The values of each undesirable metric b_{kj_0} are subtracted by $g_b\beta_{j_0}$, representing each indicator's contraction toward the best-practice frontier.

$$\sum_{j=1}^{n} b_{kj}\lambda_j \leq \underbrace{b_{kj_0} - g_b\beta_{j_0}}_{Target} \qquad k = 1, \ldots, l \tag{4}$$

If there are no undesirable metrics, such as displayed in (1), the calculation of the targets is conducted only using expression (3).

The linear programming model generates a vector of values of λ_j ($j = 1, \ldots, n$) for each DMU under evaluation. The peers of DMU j_0 are the DMUs that present λ_j different from zero, and their obtained intensity values highlight their role in the benchmarking exercise.

3.3. Use of robust and conditional approaches for composite indicators

This section explains how CIs are generated, and peers and targets are determined using the robust and robust conditional (or simply conditional) approaches.

The robust approach for computing composite indicators was developed to overcome the high sensitivity that the deterministic technique displays in presence of outliers and atypical observations in the sample. The conditional approach is employed to provide adjustments to the CIs by accounting for the influence of external contextual variables. Those techniques have been developed initially by Cazals et al. (2002) and Daraio and Simar (2005, 2007), and have been employed and extended by numerous research such as De Witte and Kortelainen (2013), Rogge et al. (2017), De Witte and Schiltz (2018), Lavigne et al. (2019), D'Inverno and De Witte (2020), Fusco et al. (2020) and Mergoni et al. (2022).

The robust method for estimating CIs involves computing a BoD model many times using randomly selected sub-samples from the collection of DMUs instead of doing so only once as the deterministic approach. This sampling procedure, known as bootstrapping, is performed with replacement, meaning that each unit can be drawn many times in the same sample. The number of sub-samples, denoted as B, is often a very high number, large enough to minimise the effect of outliers in calculating averages. The arithmetic average of the several CIs ($CI_{j_0}^{b,m}$) produced for each sub-sample yields the final robust CI for a given DMU. The effect of extreme values will be mitigated in the computation of the average CI, because they will be not present in all the sub-samples. The resulting robust CI, referred as $CI_{j_0}^{m}$, is expressed by (5).

$$CI_{j_0}^{m} = \frac{1}{B}\sum_{b=1}^{B} CI_{j_0}^{b,m} \tag{5}$$

It is possible that the DMU being evaluated (j_0) is not included in the sub-sample used for BoD calculation, and that this DMU is better-performing than all the DMUs in the sub-sample. In this case, $\beta_{j_0}^{b,m}$ displays a negative value and the value of $CI_{j_0}^{b,m}$, obtained from the expression $1/(1+\beta_{j_0}^{b,m})$, does not express the proportional performance improvement expected for a better-performing DMU. Besides that, if $\beta_{j_0}^{b,m}$ is lower than -1, $1/(1+\beta_{j_0}^{b,m})$ can assume negative values, and if β_{j_0} equals -1, CI cannot be obtained. This situation does not reflect the "super-performing" nature of those DMUs. Following Mergoni et al. (2022), we employ an alternative way to compute $CI_{j_0}^{b,m}$, when $\beta_{j_0}^{b,m}$ is negative, in order to solve this problem. This alternative solution is shown in (6).

$$CI_{j_0}^{b,m} = \begin{cases} \frac{1}{1+\beta_{j_0}^{b,m}}, & \text{if} \quad \beta_{j_0}^{b,m} \geq 0; \\ \log_{10}(1-\beta_{j_0}^{b,m}) + 1, & \text{if} \quad \beta_{j_0}^{b,m} < 0 \end{cases} \tag{6}$$

The conditional approach accounts for the contextual variables in the computation process, allowing the CIs to be adjusted by comparing the DMUs with more similar units. In that sense, fairer evaluations can be performed. As in the robust approach, B sub-samples of size m are collected, but not randomly. The sub-sample collection is performed according to a similarity function. A kernel function developed according to the contextual factors is employed to estimate the similarity between the DMU under evaluation and the other DMUs. The context can be characterised using continuous or categorical variables that can be included in the same model (Li and Racine, 2003). The BoD model is solved B times for each DMU j_0 and the CIs for each sub-sample b, designated as $CI_{j_0}^{b,m,z}$, are computed according to the expressions shown in (6). The average of $CI_{j_0}^{b,m,z}$ for a total of B sub-samples represents the conditional CI, as indicated in (7).

$$CI_{j_0}^{m,z} = \frac{1}{B}\sum_{b=1}^{B} CI_{j_0}^{b,m,z} \tag{7}$$

The influence of the contextual variables can be assessed by looking at the score ratio between the robust CI and the conditional CI ($CI_{j_0}^{m}/CI_{j_0}^{m,z}$). Using non-parametric regression between the score ratios and the contextual variables, partial plots with bias-corrected bootstrapped non-parametric confidence intervals can be obtained. Confidence intervals that do not overlap reveal a statistically significant relationship between the contextual variable and the utilities' performance (see also D'Inverno et al., 2021).

A significantly higher score ratio for a given level of the contextual variable indicates that the context is more favourable for better performance at this level. This happens when the conditional and the robust scores are similar, that is, irrespective of whether the unit under evaluation is compared against more similar units or not.

In the case of the robust and conditional BoD approaches, a set of λ_j values is generated for each computation. In each of the B computations, a number of DMUs for which λ_j is different from zero can be identified as a peer for the DMU under assessment. This makes the number of peers in those approaches to increase significantly compared to the standard BoD approach. As previously discussed, the relevance of a peer for benchmarking purposes increases as the intensity value λ_j increases. Lavigne et al. (2019) explain that the most relevant peers in the case of robust and conditional approaches are given by the higher average values of λ_j in B samples collected.

4. Case study

This section presents the data used in the study in three parts. The first one (subsection 4.1) presents the metrics used for the construction of two different and complementary composite indicators (CIs). The second part (subsection 4.2) details the dataset used to build the composite indicators (CIs). Finally, the third part (subsection 4.3) presents the data about the exogenous variables employed to characterise the context.

4.1. Metrics employed for the composite indicators (CIs)

The metrics utilised to construct the CIs in this study are described in this subsection.

Two distinct composite indicators are created by combining those metrics. The strategy for developing two different indicators is justified by the fact that improvements in managerial aspects of asset management usually take some time to generate operational benefits in a utility's performance (Luís and Almeida, 2021). Therefore, one of the CIs reflects the business's observable operational achievement: the *Resource and Infrastructure Sustainability Index (RISI)*. On the other hand, the evolution of the management system maturity is assessed by the *Asset Management Maturity Index (AMMI)*. These two composite indicators have been introduced by Vilarinho et al. (2023), but used to assess the Portuguese wholesale utilities and to provide a different empirical analysis.

The *Resource and Infrastructure Sustainability Index (RISI)* is composed by the metrics: pipeline rehabilitation (AA09b), occurrence of pipeline failure (AA10b), actual water losses (AA12b) and energy efficiency in pumping stations (AA13b). All the data reported by the water operators to ERSAR have been analysed in order to choose the metrics. In line with the literature review, which emphasises the importance of water losses, mains failure, mains rehabilitation and energy usage for managing infrastructures in water systems, those metrics have been selected to compose the RISI. They comprise the information on the operational performance of the utilities' assets reported to ERSAR and reflect the tangible results of asset management. The definition of the four metrics and their units of measurement are displayed in Table 1. The letter "b" presented in all ERSAR's metric codes indicates that the metrics come from retail utilities ("baixa" in Portuguese) to distinguish from the metrics collected from wholesale utilities that present the letter "a" ("alta" in Portuguese).

The annual report issued by ERSAR (ERSAR, 2021a) presents the results of the main performance metrics and their general reference values. ERSAR determines the reference values for the metrics that compose the RISI in three levels: good, medium and unsatisfactory. This study considers the "good" or desirable level as the ERSAR's goal for the utility.

The metric *Pipeline Rehabilitation (AA09b)* is the average yearly percentage of pipelines with an age greater than ten years undergoing rehabilitation during the previous five years. This metric aims to determine whether there is a continuous practice of pipeline restoration to guarantee their continuous renewal and an acceptable average age of the network. ERSAR defines this metric as higher than 1%, with a good result between 1% and 4%. Values above 4% are considered medium. However, the pipeline networks in Portugal are, on average, far from being at a good level, with the average result for the retail water utilities being at most 0.6% in all years from 2016 to 2020.

The second metric in RISI is the *Occurrence of Pipeline Failure (AA10b)*, which is intended to evaluate the occurrence of pipeline faults that can cause water losses and potential supply interruptions. It measures the number of failures per 100 km of pipelines per year. ERSAR considers a positive ("good") outcome for this metric to be less than 30 occurrences per 100 kilometres per year. Retail operators had an average of 38 to 42 between 2016 and 2020, regarded as a medium level.

The metric *Actual Water Losses (AA12b)* assesses the water losses in leakages and overflows, defined as the daily volume of real losses divided by the extension of the utilities' pipelines. ERSAR collects this information using two different units of measurement. This metric is expressed in litre per branch per day for denser pipeline networks, with more or equal to 20 connection branches per kilometre. If the pipeline network density is inferior to 20 branches per kilometre, the water losses are measured in cubic metres per kilometre per day. According to ERSAR, the good result for water losses is lower than 100 litres per branch per day for the denser-network utilities and inferior to 3 cubic metres per kilometre per day for the remaining utilities. According to these limits, the actual losses in Portugal are at the medium level for the less dense networks ranging from 125 to 137 litres per branch per day from 2016 to 2020. For high-density utilities, the average performance in water losses is better. This average was 2.6 cubic metres per kilometre per day in 2020. In this study, the variable density of branches per kilometre of a pipeline, collected by ERSAR with the code PiAA01b, was employed to transform the units of measurement, enabling all the water loss data to use the same unit. Because most utilities present a density superior to 20 branches per kilometre, all the data were converted to litres per branch per day.

The fourth metric that composes RISI is the energy efficiency in pumping stations (AA13b). This metric aims to assess the use of energy resources by the management entities. It is defined as the average normalised energy consumption of the pumping facilities. The performance may be judged as medium up to a value of 0.54 kilowatts per year per 100 metres elevation; however, the good performance result may be at most 0.4.

The second CI developed for asset management measurement is the *Asset Management Maturity Index (AMMI)*. The AMMI is composed by two metrics: the *Infrastructure Knowledge Index (PAA31b)* and the *Infrastructure Asset Management Index (PAA32b)*. Those are the only metrics that reflect managerial elements directly related to asset management in the dataset collected by ERSAR. They were selected because they are the two critical facets of managing infrastructures in water systems, (1) the knowledge about the assets and (2) the organisational systems that were implemented.

The *Infrastructure Knowledge Index* expresses the level of knowledge that the utilities hold about their assets. It is measured as a score taken from a questionnaire issued by ERSAR, using a scale from 0 to 200. This metric deals with the existence of engineering drawings and other records, as well as detailed information about asset conservation and the interventions performed. This information is crucial for water supply operators' business, considering that part of water systems' assets is buried and constructed to last for many years.

The second metric that composes the AMMI is the *Infrastructure asset management index*, which reflects the features of the management systems that the water utilities have implemented. The *Infrastructure asset management index* is also measured using the scores taken from a questionnaire issued by ERSAR on a scale from 0 to 200. The questionnaire used to generate the *Infrastructure asset management index* deals with the utilities' management systems, assessing aspects such as general asset management framework, strategic, tactical and operational planning, documentation and communication. The elements included in the questionnaire used for the computation of the *Infrastructure asset management index* are inspired by the international standard for asset management, the ISO 55001 (ERSAR and LNEC, 2017). Table 1 displays the metrics that compose AMMI with their definition and the codes employed by ERSAR.

4.2. Data used for building the composite indicators

This subsection details the data employed to build the two proposed composite indicators: the RISI (Resource and Infrastructure Sustainability Index) and the AMMI (Asset Management Maturity Index).

Pipeline Rehabilitation (AA09b) is a desirable metric for the metrics employed for the RISI. On the other hand, *Occurrence of Pipeline Failure (AA10b)*, *Actual water losses (AA12b)* and *Energy Efficiency in Pumping Stations (AA13b)* represent undesirable metrics. For the AMMI, the *Infrastructure Knowledge Index (PAA31b)*, and the *Infrastructure Asset Management Index (PAA32b)* are desirable metrics.

The research includes indicators acquired by ERSAR and widely publicised on the regulator's website in line with the sunshine regulation policy. ERSAR has regularly reviewed its assessment system and the indicators that make it up to ensure they are consistent with its strategic goals. This study looked at the third generation of indicators, which covered the years 2016 through 2020, and selected the data from 2020 to perform the benchmarking assessment.[2]

A list of 233 water utilities at the retail level may be found in the ERSAR dataset for 2020. The dataset is incomplete since many utilities

[2] The data is available in ERSAR's website: https://www.ersar.pt/pt/site-publicacoes/Paginas/edicoes-anuais-do-RASARP.aspx.

Table 1
Metrics for constructing the Composite Indicators.

CI	Metric code	Metric description	Metric definition	ERSAR's goals	N	Average	St. Dev.	Min.	Max.
RISI	AA09b	Pipeline Rehabilitation (%/year)	Average annual percentage of pipelines with life higher than ten years rehabilitated in the last five years.	≥ 1	223	0.58	0.86	0.01	5.40
	AA10b	Occurrence of Pipeline Failure (n^o/100 km year)	Number of failures in pipelines per 100 km in a year.	≤ 30	223	53.12	70.30	0.01	350.00
	AA12b	Actual water losses (l/branch day)	Actual water losses due to leakages and overflows per unit of pipeline length.	≤ 100	223	173.74	169.05	2.00	706.30
	AA13b	Energy efficiency in pumping stations (kWh/m^3.100 m)	Average normalised energy consumption of pumping stations.	≤ 0.4	223	1.71	1.27	0.35	3.24
AMMI	PAA31b	Infrastructure Knowledge Index (Score 0–200)	Evaluation score of the knowledge of the several infrastructures in different classes ranging from 0 to 200.	200	223	132.20	41.93	29.00	200
	PAA32b	Infrastructure Asset Management Index (Score 0–200)	Evaluation score in a questionnaire about asset management practices ranging from 0 to 200.	200	223	40.17	67.94	0.01	200

have not reported their results, and multiple missing data are present. The operators included in the sample studied were only those who have provided data for at least two metrics used in the RISI and presented no missing data in the metrics of AMMI. This approach guaranteed the consistency and practical relevance of the obtained results. We removed ten utilities from the original sample, resulting in a final number of 223 water operators for evaluation.[3] Even after removing these utilities, the remaining sample still accounts for 95.7% of the total number, representing a significant proportion of the original dataset. The remaining missing data instances were treated following the procedure employed by Kuosmanen et al. (2002), Morais and Camanho (2011) and Henriques et al. (2020). For the desirable metrics, a small value equal to the minimum value of each metric replaced the missing data. In the case of undesirable metrics, the missing instances were changed to a large number equivalent to the maximum value of each metric. This procedure ensures that the DMU cannot benefit from the lack of data for its performance evaluation. Several scores of the *Infrastructure Asset Management Index*, one of the AMMI's components, present a value of zero in 2020. This fact reflects the low level of maturity in many retail water utilities concerning the development of management systems. The same situation occurs for the metric *Pipeline Rehabilitation*, one of the RISI components, meaning that those utilities could not recover their pipelines as expected.

However, a few utilities presented zero *occurrences of failures in pipelines*, which is another component of the RISI and represents the best result for this undesirable metric. In general, DEA formulations require that the inputs and outputs are positive. Even though, this "positivity property" can be relaxed, as detailed by Charnes et al. (1991), we chose to replace the zero values with 0.01, a small positive number as recommended by Bowlin (1998) and discussed by Sarkis (2007). Since the BoD model emphasises the indicators for which the DMU performs best, an indicator with a minimal value would not be expected to contribute to any bias in the efficiency assessment. Table A.1 in Appendix displays the list of the evaluated utilities (DMUs) with the identification codes used in the study ranging from B1 to B223.

The descriptive statistics for the data related to the metrics that compose both CIs are presented in Table 1. Looking at the average results of the metrics in 2020 (Table 1) and comparing them with ERSAR's goals, the asset-management-related metrics perform worse than the ideal levels. In terms of the operational metrics, the average AA09b is less than 1%, the average AA10b is greater than 30, the average AA12b is much higher than 100 and the AA13b significantly surpasses 0.4. The managerial metrics also indicate poor average outcomes. PAA31b and PAA32b are far below the ideal score of 200. The benchmarking exercise performed in this study indicates realistic targets for the operators to pursue, in comparison with the market best-performs. Given this scenario, such targets may not always reach the expected goals set by ERSAR.

4.3. Data used as exogenous variables

This subsection presents the data employed to characterise the environment in which the utilities under evaluation operate.

Contextual factors were selected among the data reported by the retail water utilities to ERSAR to characterise the environment in which the utilities operate. Four contextual variables or exogenous variables were chosen: the management system, the typology of intervention area, the geographic location and the volume of activity.

The management system indicates the kind of utility ownership, according to the models available for the water sector in Portugal. Municipalities own and operate most retail water utilities directly, 74.4% in 2020. Direct management is thus the management system of 166 retail water utilities. The remaining 25.6% are divided in Concession (12.6%) and Delegation (13.0%).

The typology of intervention area is mainly related with the population density. According to the kind of intervention area, the water operators can be classified in urban, semi-urban or rural. Based on this criterion, most of the utilities (147) are rural, representing 65.9% in 2020. Other 55 utilities (24.7%) are considered semi-urban and only 21 (9.4%) utilities are urban.

The geographic location is based on the region of Portugal where the utility primarily operates. This classification is based on the European Union's Nomenclature of Territorial Units for Statistics (NUTS) standard. At its second level, known as NUTS 2, the locations presented in ERSAR's reports for mainland Portugal are Algarve, Alentejo, Centre, Lisbon and North.

The volume of activity expressed as the metric *PAA50b* represents the amount of water (in m^3) supplied by the operator in a year. This metric can be used as a proxy for the utility size and is available in ERSAR's reports. However, following Mergoni et al. (2022), we chose to characterise the context by classifying the utilities as small, medium and large. We use the approach of the *Drinking Water Directive*, Council Directive 98/83/EC (European Commission, 1998). This directive defines the limit between small and large utilities as 1,000 m^3/day of average supplied water volume or 5,000 persons in the population served. Small utilities were defined in this study as those that provide less than 1,000 m^3/day. Only 63 utilities fall under this threshold; thus, the remaining 160 operators were split into groups of 80 units each, including medium and large utilities.

Table 2 presents the statistics for the exogenous variables in 2020.

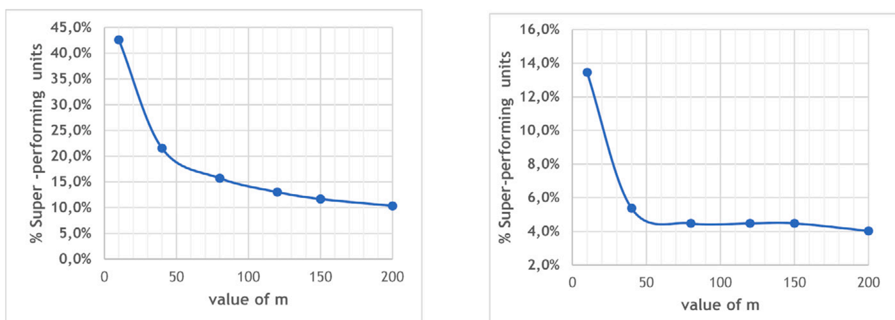
5. Results and discussion

This section presents and discusses the study's findings in three parts. The computation of composite indicators using the deterministic, robust and robust conditional approaches is presented in subsection 5.1. Subsection 5.2 describes the identification of peers and targets for benchmarking practices. The role of the environment on the performance of the water utilities is discussed in subsection 5.3.

[3] The utilities removed from the sample for presenting missing data are: APIN, CM de Cabeceiras de Basto, CM de Caminha, CM de Idanha-a-Nova, CM de Marco de Canaveses, CM de Monchique, CM de Paredes, CM de Santo Tirso, CM de Vila Nova de Paiva, and CM de Vila Viçosa.

Table 2
Categorical exogenous variables.

ERSAR Code	Variable description	Categories	Obs.	Number of utilities and percentage per category
PAA02b	Management System	Concession, Delegation or Direct Management	223	Concession - 28 (12.6%) Delegation- 29 (13.0%) Dir. Manag. - 166 (74.4%)
PAA14b	Typology of Intervention Area	Rural areas, Semi-Urban areas or Urban areas	223	Rural - 147 (65.9%) Semi-urban - 55 (24.7%) Urban - 21 (9.4%)
NUTS2	Geographic Location	Alentejo, Algarve, Centre, Lisbon or North	223	Alentejo - 54 (24.2%) Algarve - 18 (8.1%) Centre - 68 (30.5%) Lisbon - 16 (7.2%) Norte - 67 (30.0%)
PAA50b	Volume of Activity	Small, Medium or Large	223	Small - 63 (28.3%) Medium - 80 (35.9%) Large - 80 (35.9%)



(a) % Super-performing DMUs for RISI relative to the value of m

(b) % Super-performing DMUs for AMMI relative to the value of m

Fig. 1. Sensitivity analysis for m selection.

5.1. Composite indicator results

The results from the calculation of the CIs are presented and discussed in this subsection.

The methods explained in subsections 3.1 and 3.3 are employed to compute the deterministic, robust unconditional and robust conditional CIs. BoD Model (2), which can handle both desirable and undesirable metrics, calculates the RISI. BoD Model (1) calculates the AMMI since this indicator only includes desirable metrics. The R program solved the BOD models using the R program's packages *Rglpk* (Theussl and Hornik, 2019) and *lpSolve* (Berkelaar et al., 2023). An additional R package, the *np* package (Hayfield and Racine, 2008), was used to handle the collection of sub-samples according to the similarity level of DMUs in the robust conditional approach. This R package was also used to compute the bias-corrected bootstrapped non-parametric confidence intervals of the utilities' performance concerning the environment.

For the robust and robust conditional CIs, the values of the parameters m and B must be determined. B is often a high number, and for this study, the value of 2,000 was employed for B. According to Daraio and Simar (2007), there are no formal guidelines for choosing m, but for smaller values of m, the presence of numerous "super-performing units" might be problematic. Therefore a sensitivity analysis is recommended to select a value of m. Fig. 1 shows two graphs that present the resulting percentage of "super-performing" DMUs in each of the robust CIs' computations for several values of m. These findings led to the choice of m = 80 for both CIs since, at this value, the proportion of "super-performing" units reduces, whereas it remains relatively stable at higher values.

The CI scores obtained with the deterministic, robust and robust conditional techniques are shown in Table A.1 in Appendix for all DMUs. Table 3 presents the descriptive statistics for both composite indicators. A close look at the average of both indicators reveals that the performance of the retail water operators in asset management may be significantly improved.

A combined visualisation model shown in Fig. 2, based on the BCG (Boston Consulting Group) matrix, is displayed, following Vilarinho et al. (2023), to enable the joint analysis of water operators in both indicators (RISI and AMMI). Considering that the robust conditional approach provides the most accurate and fair comparison of the utilities, this version of the composite indicators was used to represent the performance of the utilities in the following analyses. The 2×2 matrix in Fig. 2 divides the utilities according to the median of their robust conditional indicators. Fig. 2 classifies the utilities' performance into four categories, as listed below, to show how they operate compared to their competitors:

(i) Stars. These utilities present better operational results and better management systems than their peers. In this category, both RISI and AMMI are higher than the median values.
(ii) Soldiers. This group takes care of the assets, keeping their operational conditions, but in comparison to their counterparts, the management procedures are not effectively established. For the Soldiers, the RISI is higher than the median while the AMMI is lower or equal.
(iii) Infants. This category gives the initial moves in the organisation for asset management, and they show worse tangible results than

Table 3
Descriptive statistics for RISI and AMMI Results.

CI	CI formulation	Average	St. Dev.	Min	Q1	Median	Q3	Max
RISI	Deterministic RISI CI (CI_{j_0})	0.697	0.161	0.529	0.535	0.656	0.818	1.000
	Robust Unconditional RISI CI ($CI^{m}_{j_0}$)	0.766	0.214	0.530	0.585	0.731	0.879	2.082
	Robust Conditional RISI CI ($CI^{m,z}_{j_0}$)	0.824	0.161	0.533	0.667	0.856	0.985	1.076
AMMI	Deterministic AMMI CI (CI_{j_0})	0.669	0.213	0.145	0.510	0.675	0.840	1.000
	Robust Unconditional AMMI CI ($CI^{m}_{j_0}$)	0.670	0.213	0.145	0.511	0.676	0.841	1.005
	Robust Conditional AMMI CI ($CI^{m,z}_{j_0}$)	0.719	0.218	0.148	0.562	0.741	0.925	1.001

Table 4
Summary of utilities' categories in asset management performance.

Class	Count	%
Infant	71	31.8%
Learner	41	18.4%
Soldier	41	18.4%
Star	70	34.4%

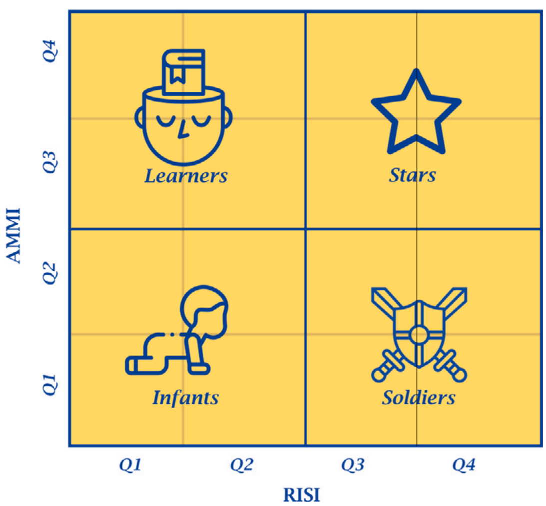


Fig. 2. Visualisation model - RISI and AMMI.

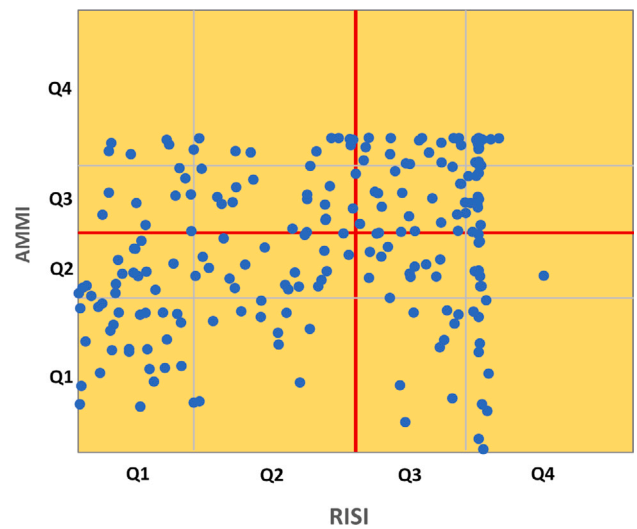


Fig. 3. Results for RISI and AMMI in 2020. Conditional Approach - All Retail Utilities.

their peers. The Infants present both RISI and AMMI below or equal to the medians of all utilities.

(iv) Learners. Although these utilities have been working on effective management systems, they have performed poorly than most of their counterparts regarding operational results in asset management. This group's AMMI is above the median, while its RISI is equal to or lower than the median.

Table A.1 in Appendix outlines all utilities' classifications in 2020. A summary of the classification is shown in Table 4 and the distribution of the 223 utilities is presented in Fig. 3.

A correlation test was conducted to verify if improved performance on the AMMI dimension is associated with better performance on the RISI dimension. The results for Pearson correlation show that even though the correlation is significant (p-value = 0), the correlation coefficient (ρ) is only 0.325, indicating that the correlation is not strong between the two CIs. These results suggest that the maturity in asset management systems is not necessarily associated with good operational performance in the short term. As previously discussed, it takes time for management efforts to generate operational results.

It is possible to examine the results obtained by some utilities and compare them with past data collected from the literature to illustrate the potential impact of the conversion of managerial actions in asset management into operational results. A group of 19 water retailers simultaneously started to implement asset management practices in 2012–2013 as described by Leitão et al. (2016). At the end of this collaborative project, each utility issued strategic and tactical plans aiming to develop asset management practices. These plans in most cases were effectively implemented. One of those utilities, *SMAS de Almada*, was the only utility in Portugal operating exclusively in the retail market to hold the international certification in asset management, ISO 55001. The results from the computation of RISI and AMMI in 2020 for those 19 utilities are encouraging, as 14 of them have achieved the status of Star. Three of the remaining five utilities had been merged into other systems, making their results incomparable. This trend is an indication of the positive operational results that can be achieved with the implementation of managerial practices.

The ten best performers and the ten worst performers in both dimensions are presented in Table 5. In the case of AMMI, 11 top-performance utilities are presented because the value of AMMI for the 10th and 11th utility is the same. Their position in the 2 × 2 Matrix is also shown in Fig. 4. DMU B45 is identified in black colour in Fig. 4 because it is classified both as *Bottom 10 RISI* and as *Bottom 10 AMMI*. DMUs B115, B123 and B165 also identified in black colour because they are classified both as *Top 10 RISI* and as *Bottom 10 AMMI*.

A closer look at the AMMI outcomes from Table 5 reveals that the majority of best performers are managed by concession or delegation (8 utilities), located in urban or semi-urban areas (8 utilities) and are large (9 utilities). However, all of the worst performers are small and managed directly by the municipalities, most of which are located in rural areas (9 utilities). In the case of RISI, no patterns can be observed

Table 5
Top 10 and bottom 10 performers in each composite indicator.

RISI top 10 performers		RISI bottom 10 performers		AMMI top 10 performers		AMMI bottom 10 performers	
Code	Utility	Code	Utility	Code	Utility	Code	Utility
B163	CM de Sousel	B68	CM de Castelo de Paiva	B44	CM de Alfândega da Fé	B133	CM de Penedono
B168	CM de Vale de Cambra	B120	CM de Moura	B4	Águas da Figueira	B90	CM de Gouveia
B119	CM de Mora	B101	CM de Manteigas	B13	Águas de Gondomar	B57	CM de Arronches
B164	CM de Tábua	B70	CM de Castro Daire	B61	CM de Barreiro	B115	CM de Mondim de Basto
B123	CM de Nisa	B116	CM de Monforte	B66	CM de Bragança	B45	CM de Alijó
B60	CM de Barrancos	B109	CM de Miranda do Douro	B191	Indaqua Fafe	B112	CM de Moimenta da Beira
B165	CM de Tabuaço	B179	CM de Vila Nova de Foz Coa	B192	Indaqua Feira	B123	CM de Nisa
B63	CM de Bombarral	B110	CM de Mirandela	B193	Indaqua Matosinhos	B117	CM de Montalegre
B115	CM de Mondim de Basto	B45	CM de Alijó	B195	Indaqua Santo Tirso/Trofa	B98	CM de Lousada
B75	CM de Condeixa-a-Nova	B180	CM de Vila Pouca de Aguiar	B196	Indaqua Vila do Conde	B165	CM de Tabuaço
				B201	INOVA		

Table 6
Descriptive statistics for the Composite Indicators' Targets.

CI	Target	ERSAR's goals	Average performance	N	Robust Cond. DEA/BoD Targets			
					Average	St Dev	Min	Max
RISI	TG_AA09b - Pipeline Rehabilitation (%/year)	≥ 1	0.58	223	0.66	0.91	0.01	5.40
	TG_AA10b - Occurrence of Pipeline Failure (n°/100 km.year)	≤ 30	53.12	223	31.82	41.22	0.01	350.00
	TG_AA12b - Actual Water Losses (l/branch day)	≤ 100	173.74	223	111.63	121.54	2.00	706.30
	TG_AA13b - Energy Efficiency in Pumping Stations (kWh/m^3.100 m)	≤ 0.4	1.71	223	1.00	0.78	0.35	3.24
AMMI	TG_PAA31b - Infrastructure Knowledge Index (Score 0–200)	200	132.2	223	184.70	17.55	71.67	200.0
	TG_PAA32b - Infrastructure Asset Management Index (Score 0–200)	200	40.17	223	43.91	72.16	0.01	200.0

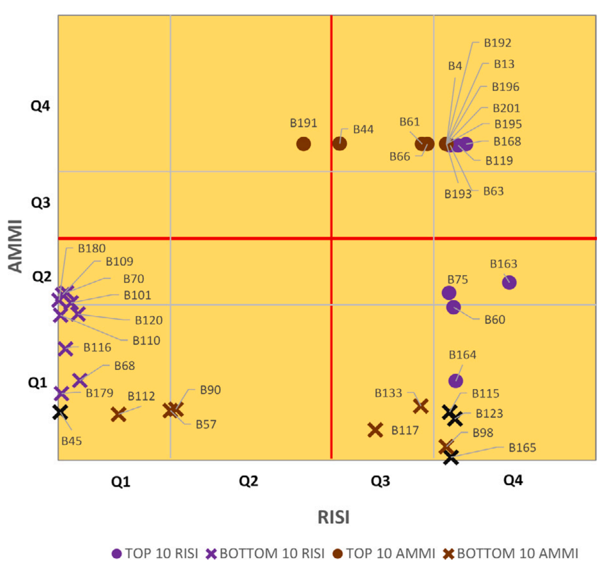


Fig. 4. Top 10 performers and Bottom 10 performers in both dimensions.

between the top and bottom performers and the context in which they operate. The trends observed in AMMI results are analysed in more detail using the CI conditional techniques in subsection 5.3.

5.2. Identification of peers and targets

This subsection addresses the designation of utility peers for best-practice identification and selecting the most suitable benchmarking targets for individual utilities.

The targets for the desirable metrics assigned for each utility are obtained from the expressions (3) present in Models (1) and (2). The targets for the undesirable metrics in Model (2) are obtained from expression (4). These targets represent the projection of the performance metrics toward the efficient best-practice frontier, meaning that if the utility under evaluation can leverage its performance to reach those objectives, it will reach the benchmarking level compared to the other utilities. After the BoD models are solved, using the robust conditional approach, the descriptive statistics for the targets in RISI and AMMI are presented in Table 6.

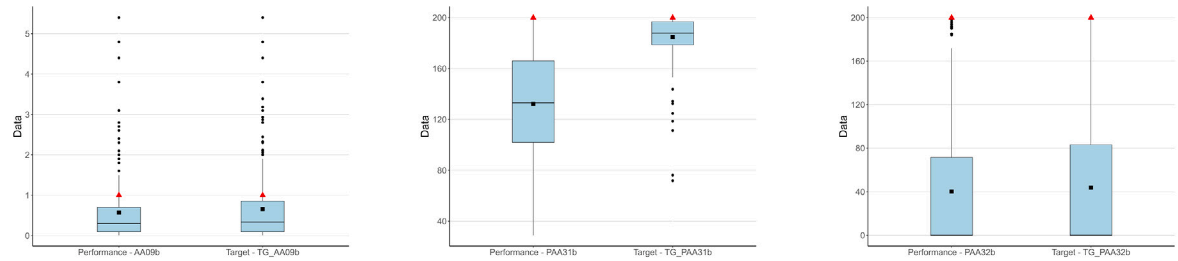
From the results in Table 6, one can notice that the average target values do not reach the goals established by ERSAR, meaning that in several cases utilities may achieve levels comparable to their peers without complying with the regulator's objectives. Given the poor performance of the retail operators in these metrics, this result indicates that a revision of ERSAR policy in setting up the goals for asset management targets may be necessary.

Figs. 5 and 6 display the distribution of the desirable and undesirable metrics' targets, respectively, alongside the distribution of the actual metrics (observed performance). The black squares in the box plots of Figs. 5 and 6 indicate the averages and the red triangle is the goal determined by ERSAR. Looking at the box plots and the values of standard deviations in Table 6, it is noticeable that the variation among the targets is considerable, reinforcing the sector's heterogeneity regarding asset management practices.

This heterogeneity is also seen when the results are analysed in each category assigned in the research (Table 7). Even for the Stars, considered the top-performing category, it is noticeable that given their current performance, the goals set by ERSAR look unrealistic in many cases.

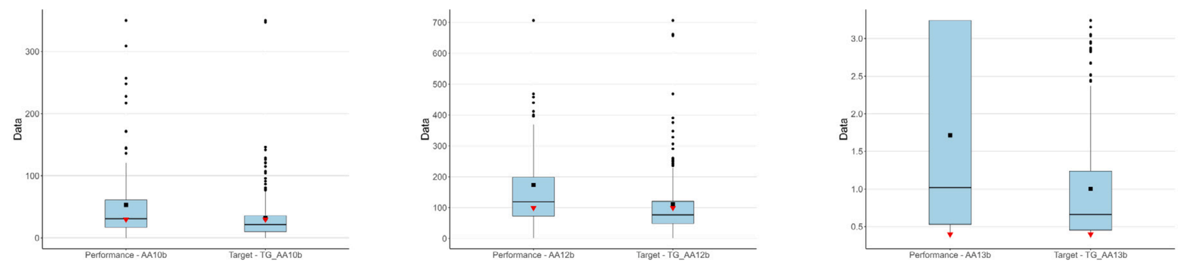
The peers of the utilities are identified as the ones that present the intensity variable λ_j different from zero as an output of the BoD models. The peer set for a given DMU represents its closest anchor on the best-practice frontier, meaning a potentially suitable choice to guide improvements. In robust and robust conditional approaches, this peer set is more extensive due to the high number of efficient frontiers ($B = 2000$ in this study). The average of λ_j for the sub-sample of B interactions in which the peer was actually selected to compose the sub-sample gives the relevance of the peers.

Following Lavigne et al. (2019), we built an intensity matrix for each CI presenting the average values of λ_j that identify the peers for the robust conditional case. The intensity matrix is a n by n matrix (223 × 223 in this study), where each row represents a vector of intensities for the evaluated DMUs. The vectors of intensities include the average values of λ_j for each peer identified. One part of the intensity matrix generated in the computation of RISI is displayed in Table 8 to illustrate this process. In Table 8, it is possible to notice



(a) Pipeline Rehabilitation (b) Infrastructure Knowledge Index (c) Infrastructure Asset Management Index

Fig. 5. Comparison between the distributions of actual performances and targets for the desirable metrics.



(a) Occurrence of Pipeline Failure (b) Actual Water Losses (c) Energy Eff. in Pumping Stations

Fig. 6. Comparison between the distributions of actual performances and targets for the undesirable metrics.

Table 7
Average Metrics and Average Targets in Each Category.

Class	Number of units	CI Average RISI	AMMI	AA09b Average performance	AA09b Average target	AA10b Average performance	AA10b Average target	AA12b Average performance	AA12b Average target	AA13b Average performance	AA13b Average target	PAA31b Average performance	PAA31b Average target	PAA32b Average performance	PAA32b Average Target
INFANT	71	0.600	0.539	0.27	0.40	77.61	29.15	210.82	77.20	2.63	1.04	99.32	185.41	4.37	7.97
LEARNER	41	0.639	0.883	0.37	0.51	60.07	32.32	179.06	91.44	2.05	1.07	154.61	175.32	50.88	54.22
SOLDIER	41	0.961	0.530	1.08	1.11	36.44	34.61	227.90	217.77	1.42	1.34	99.98	188.40	4.08	6.18
STAR	70	0.930	0.918	0.71	0.74	33.99	32.59	101.28	96.23	0.76	0.72	171.29	187.30	91.35	96.43

Table 8
Partial intensity matrix for RISI: Average λ_j values in the robust conditional approach.

Utility	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
B1 - AGERE	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B2 - Águas da Azambuja	0.00	0.06	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.00
B3 - Águas da Covilhã	0.00	0.00	0.14	0.00	0.02	0.01	0.00	0.00	0.00	0.00
B4 - Águas da Figueira	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
B5 - Águas da Região de Aveiro	0.00	0.00	0.00	0.17	0.20	0.00	0.00	0.00	0.00	0.00
B6 - Águas da Teja	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.00
B7 - Águas de Alenquer	0.00	0.00	0.00	0.87	0.00	0.00	0.00	0.00	0.00	0.00
B8 - Águas de Barcelos	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B9 - Águas de Carrazeda	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.93	0.00
B10 - Águas de Cascais	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07

that all utilities in this part of the intensity matrix except B7 and B8 are peers of themselves, meaning that they were identified as efficient in some or all *B* sub-samples. The utility *B4 - Águas da Figueira* presents a λ_j equal to 1 when compared to itself, which means that it was the only utility included in the efficient frontier in all the *B* computations performed for its robust conditional BoD assessment. A different situation is noticed for the utility *B7 - Águas de Alenquer*. In this case, *B4 - Águas da Figueira* is found as being a relevant peer for *B7* with λ_j equal to 0.87. The complete intensity matrices are too large to be displayed in the text, but are available upon request to the authors.

In the DEA/BoD conditional approach, the selection of samples is not random; utilities that are more alike, and operating in similar environments, have a higher probability of being included in the *B* subsamples considered. Therefore, as highlighted by Lavigne et al. (2019), higher values of λ_j indicate that the peer is more relevant because it is comparable in performance and operating environment.

Table 9 illustrates the results of the benchmarking assessment using one utility, *SM de Alcobaça - B206*, as an example, employing the robust conditional BoD approach. The table displays the value of each metric alongside the computed targets and ERSAR's goals. *SM de Alcobaça* is classified as Infant, and the calculated targets for most of the metrics are generally realistic and do not reach ERSAR's expectations. For the only two metrics for which this utility's performance is acceptable according to ERSAR's goals, pipeline failures and water losses, the

Table 9
Example of target and peer determination: DMU B206 - utility SM de Alcobaça.

CI	Metric component of CI	Unit	ERSAR's goals	Actual performance	DEA/BoD targets	Peer performance
RISI	AA09b - Pipeline Rehabilitation	%/year	≥ 1	0.5	0.6	0.4 (B100); 0.6 (B66); 2.1 (B201).
	AA10b - Occurrence of Pipeline Failure	no./(100 km year)	≤ 30	27	19	19 (B100); 20 (B66); 7 (B201).
	AA12b - Actual Water Losses	l/(branch day)	≤ 100	71	50	6 (B100); 35 (B66); 108 (B201).
	AA13b - Energy Effic. in Pumping Stations	kWh/(m^3.100 m)	≤ 0.4	0.86	0.61	0.95 (B100); 0.47 (B66); 0.35 (B201).
AMMI	PAA31b - Infrastructure Knowledge Index	score	200	85	169	147 (B28); 186 (B66); 190 (B201).
	PAA32b - Infrastructure Asset Manag. Index	score	200	80	159	184 (B28); 200 (B66); 200 (B201).

assigned targets are more challenging than the ones set by the regulator. In that sense, this example reveals the BoD technique's ability to provide more suitable targets for each particular DMU.

In the calculation for RISI, 64 peers were identified for *SM de Alcobaça - B206*, while in AMMI model, 39 peers were determined. The three most relevant peers of *SM de Alcobaça* presenting the highest values of λ_j, according to the performance in RISI are *CM de Mangualde - B100* ($\lambda_{100} = 0.21$), *CM de Bragança - B66* ($\lambda_{66} = 0.18$) and *INOVA - B201* ($\lambda_{201} = 0.11$). Regarding AMMI, the three most relevant peers are *INOVA - B201* ($\lambda_{201} = 0.39$), *CM de Bragança - B66* ($\lambda_{66} = 0.16$) and *Águas do Planalto - B28* ($\lambda_{28} = 0.10$). The performance metrics of the peers of *SM de Alcobaça* are presented in Table 9 as well. *SM de Alcobaça* should look at their performance and learn from their practices. This exercise is facilitated because they share similar environments as the peer selection generated from the BoD conditional approach.

The entire list of targets and peers is available upon request from the authors.

5.3. Role of context on the utilities' performance

This subsection presents and discusses the role of contextual factors on the utilities' performance in asset management.

As explained in subsection 3.3, the partial plots with bias-corrected bootstrapped non-parametric confidence intervals of the score ratios (between the robust CI and conditional CI) can be used to assess the relationship between the context and the utilities' performance. The partial plots are obtained using the *np* package in R (Hayfield and Racine, 2008).

Considering the variable Management System (PAA002b), partial plots with confidence intervals are shown in Fig. 7.

Regarding RISI, no differences are noticed for the different management systems, as shown in Fig. 7(a). Regarding AMMI, the partial plots in Fig. 7(b) reveal that the performance of the utilities directly managed by the municipalities is significantly different from the ones managed by concession and delegation. Specifically, the direct management system displays score ratios significantly lower and an unfavourable role on the utilities' AMMI performance. A direct management system implies that the water utility is managed and controlled exclusively by the public sector (municipalities). The other two management systems assume the responsibility of a designated utility for the services, either through a concession contract or a delegation from the public sector. From the results in Fig. 7(b), concession and delegation models have been more successful in implementing structured practices to manage their infrastructures. The concession and delegation utilities often specialise in the water supply sector and may present more proficient administration. The fact that those utilities present higher maturity in management systems can leverage their operational results in the future.

The effect of the typology of intervention area in asset management performance is displayed in Fig. 8.

Regarding RISI, the urban environment is less favourable for operational results, as the graph in Fig. 8(a) reveals. The younger infrastructure in Portugal's rural and semi-urban areas may be why urban regions have inferior operational performance. Several extension projects have been carried out in Portugal to expand pipeline networks into rural areas in recent decades. More recent water systems are more likely to be free of leakages or failures, which comprise the RISI metrics.

On the other hand, regarding AMMI, urban settings are more favourable to maturity in management systems, as indicated by the AMMI score ratio in Fig. 8(b). Better knowledge about their assets (such as accurate engineering drawings and records) may explain the better performance in management systems by urban utilities. Notice that the *Infrastructure knowledge index* is one of the components of AMMI.

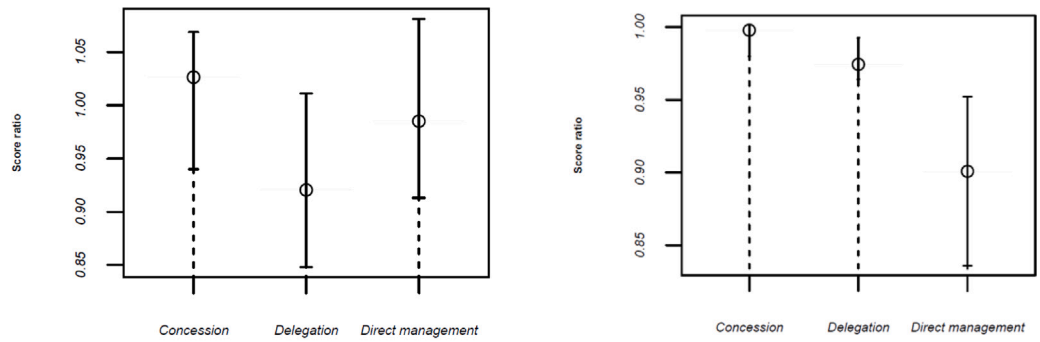
Regarding the geographic location, Fig. 9 displays the partial plots for both indicators' performance scores. Results from RISI (Fig. 9(a)) and AMMI (Fig. 9(b)) show that the geographic regions present similar performance in both dimensions of asset management performance measurement.

Concerning the volume of activity, the partial plots displayed in Fig. 10 confirm that the volume of activity makes no statistically significant difference in the performance of both RISI and AMMI.

6. Conclusion

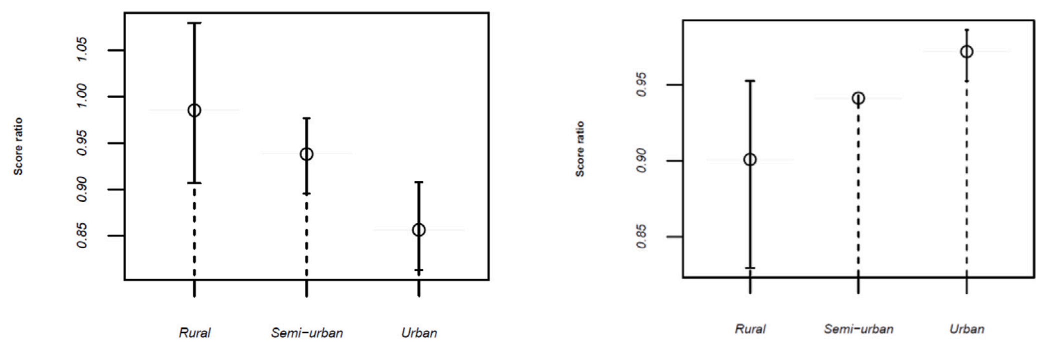
This research contributes to the literature by providing a novel method to identify peers and targets for benchmarking asset management practices in the retail water sector. The benchmarking exercise is carried out using the Portuguese water sector. This unique market is fragmented, displaying hundreds of operators and, is heterogeneous in many facets, such as governance, utility size and service scope. The country's regulatory authority policies actively focus on benchmarking, making this topic relevant. Rather than increasing investments in new assets, the national strategic policy for the industry encourages strengthening current infrastructure management (Frade et al., 2015).

Moreover, the current state of infrastructure preservation is inadequate, making this subject even more critical. The study provides new and adaptable tools for the regulators and utilities to strengthen sunshine regulation practices. Official metrics issued by the regulatory authority (ERSAR) are employed to construct two Composite



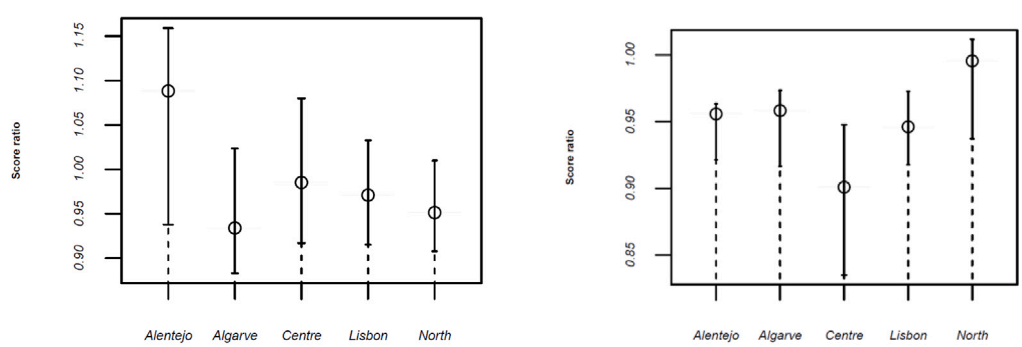
(a) RISI (b) AMMI

Fig. 7. Effect of exogenous variable - management system.



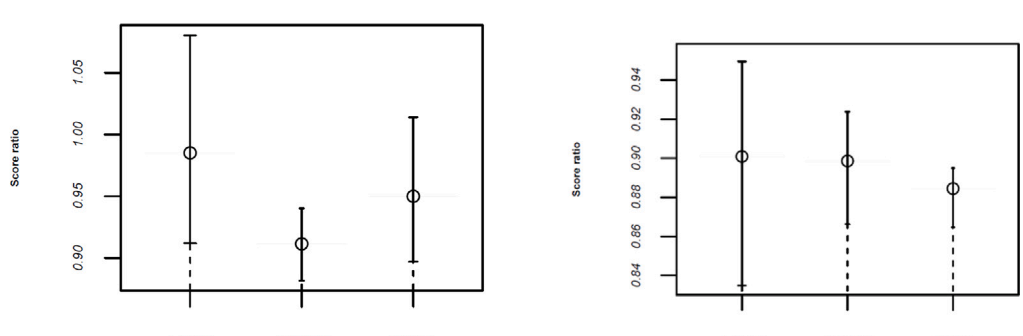
(a) RISI (b) AMMI

Fig. 8. Effect of exogenous variable - typology of intervention area.



(a) RISI (b) AMMI

Fig. 9. Effect of exogenous variable - geographic location.



(a) RISI (b) AMMI

Fig. 10. Effect of exogenous variable - volume of activity.

Indicators (CIs) reflecting the managerial practices and operational performance of asset management in 223 retail operators. Benefit-of-the-Doubt (BoD) directional distance models enable the computation of the indicators. They are employed to identify the most suitable peers and targets for the benchmarking exercise, representing this study's innovative contribution. A visualisation model for the combined assessment of the two CIs is also provided.

To facilitate statistical inference and investigate the relationship between contextual elements and utility performance, robust and conditional techniques are utilised in addition to the deterministic strategy for building CIs. The most recent models of concession or delegation favour a better performance in managerial practices but do not influence the operational results. The operation in urban areas is favourable for managerial practices but unfavourable for operational results. The performance for tangible results and management features is not sensitive to the volume of water supplied and the geographical location in mainland Portugal.

The targets generated are specific for each operator and reflect the most suitable way to pursue and conduct improvements. The fact that those targets, in most cases, fall short of ERSAR's ideals means that the utilities may need to follow specific and more realistic pathways for their performance. In that sense, the regulatory authority can take advantage of the procedure detailed in this study, setting individual and feasible targets for the sector's operators. By identifying a group of peers for benchmarking asset management practices, the study offers guidance on where to look for recommendations in this highly fragmented and diverse scenario of retail water businesses. The methodology outlined in this study has the potential to be replicated in other developed countries facing similar challenges related to the maintenance and renewal of water supply and sanitation infrastructures, as long as the necessary data is available.

The main limitations of the study rely on the data set available. Even though the data are reliable and provided by an official source, the information gathered needs to be enlarged in several aspects, such as investment details and preservation of vertical assets, such as storage tanks. The results of studies like this can reinforce the practical importance of this information and stimulate the regulatory authority and the operators to expand the collected data set. Furthermore, the study did not involve stakeholders such as regulatory authorities, and other relevant groups in the benchmarking exercise. Including these groups might have provided valuable insights from different perspectives, and incorporating their preferences into the models could have added value to the research and enhanced its practical applicability. Future research can look at the progression of the asset management practices over time, determine each utility's strengths and weaknesses, and explore other perspectives of ERSAR's metrics, such as environmental issues or quality of service.

Further studies could also examine how different ownership and governance structures, such as private sector participation and public–private partnerships, affect asset management practices. Comparing the asset management practices and operational performance of retail water operators in Portugal with those in other countries facing similar challenges could provide valuable insights. Furthermore, examining the effectiveness of different benchmarking methods and tools for improving asset management and investigating the role of regulatory policies on asset management practices in the retail water sector could be worthwhile research avenues to pursue.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Acknowledgements

The first author acknowledges the financial support provided by the Portuguese Foundation for Science and Technology (FCT, from the Portuguese abbreviation of *Fundação para a Ciência e a Tecnologia*) through PhD research grant 2021.05244.BD and INESC TEC – Institute for Systems and Computer Engineering.

Appendix. Composite indicators (CIs) for all utilities in 2020

See Table A.1.

Table A.1
CIs for all utilities in 2020.

Utility ID	Utility	Category	Deterministic CI		Robust CI		Robust Conditional CI	
			RISI	AMMI	RISI	AMMI	RISI	AMMI
B1	AGERE	STAR	0.961	0.895	1.004	0.896	0.997	0.897
B2	Águas da Azambuja	LEARNER	0.665	0.705	0.694	0.709	0.849	0.997
B3	Águas da Covilhã	STAR	0.735	0.715	0.771	0.716	0.909	0.743
B4	Águas da Figueira	STAR	0.938	1.000	1.018	1.002	1.000	1.000
B5	Águas da Região de Aveiro	STAR	0.706	0.930	0.761	0.931	0.866	0.939
B6	Águas da Teja	STAR	0.767	0.730	0.922	0.731	0.998	0.817
B7	Águas de Alenquer	LEARNER	0.792	0.980	0.831	0.985	0.836	1.000
B8	Águas de Barcelos	LEARNER	0.607	0.980	0.819	0.982	0.850	0.980
B9	Águas de Carrazeda	STAR	0.825	0.675	0.846	0.675	0.993	0.743
B10	Águas de Cascais	LEARNER	0.747	0.990	0.772	0.994	0.854	0.995
B11	Águas de Coimbra	STAR	0.824	0.955	1.013	0.956	1.000	0.970
B12	Águas de Gaia	LEARNER	0.730	0.885	0.741	0.886	0.737	0.887
B13	Águas de Gondomar	STAR	1.000	1.000	1.051	1.003	1.000	1.000
B14	Águas de Ourém	STAR	0.839	0.850	0.876	0.851	0.911	0.850
B15	Águas de Paços de Ferreira	STAR	0.751	0.825	0.948	0.833	0.999	0.839
B16	Águas de Paredes	STAR	1.000	0.975	1.276	0.976	1.000	0.975
B17	Águas de S. João	INFANT	0.750	0.580	0.768	0.580	0.790	0.593
B18	Águas de Santarém	STAR	0.797	0.955	0.833	0.960	0.976	0.998
B19	Águas de Santo André	STAR	1.000	0.820	1.012	0.821	1.000	0.904
B20	Águas de Valongo	STAR	0.747	0.975	0.818	0.976	0.868	0.975
B21	Águas de Vila Real de Santo António	LEARNER	0.629	0.865	0.670	0.866	0.717	0.865
B22	Águas do Alto Minho	STAR	0.590	0.715	0.661	0.716	0.976	0.746
B23	Águas do Baixo Mondego e Gândara	LEARNER	0.539	0.720	0.563	0.721	0.664	0.745
B24	Águas do Interior - Norte	SOLDIER	0.699	0.660	0.751	0.661	0.873	0.688

(*continued on next page*)

Table A.1 (*continued*).

Utility ID	Utility	Category	Deterministic CI		Robust CI		Robust Conditional CI	
			RISI	AMMI	RISI	AMMI	RISI	AMMI
B25	Águas do Lena	STAR	0.785	0.845	0.872	0.846	0.997	0.995
B26	Águas do Marco	INFANT	0.536	0.735	0.591	0.736	0.797	0.735
B27	Águas do Norte	LEARNER	0.537	0.770	0.569	0.771	0.713	0.825
B28	Águas do Planalto	STAR	1.000	0.920	1.057	0.925	1.000	0.927
B29	Águas do Porto	STAR	1.000	0.985	1.038	0.986	1.000	0.987
B30	Águas do Ribatejo	STAR	0.818	0.740	0.844	0.741	0.956	0.768
B31	Águas do Sado	STAR	0.797	0.950	0.820	0.956	0.896	0.956
B32	AMBIOLHÃO	INFANT	0.638	0.665	0.646	0.665	0.678	0.675
B33	Aquaelvas	SOLDIER	0.768	0.470	0.782	0.470	0.976	0.517
B34	Aquafundalia	SOLDIER	0.642	0.440	0.731	0.440	0.972	0.492
B35	Aquamaior	STAR	0.767	0.815	0.811	0.816	0.988	0.896
B36	Aquanena	STAR	0.814	0.870	0.879	0.871	0.969	0.922
B37	CARTÁGUA	INFANT	0.532	0.555	0.546	0.555	0.715	0.589
B38	CM de Aguiar da Beira	INFANT	0.534	0.375	0.550	0.375	0.613	0.422
B39	CM de Alandroal	LEARNER	0.573	0.840	0.640	0.841	0.811	0.963
B40	CM de Albufeira	LEARNER	0.533	0.805	0.569	0.810	0.734	0.961
B41	CM de Alcácer do Sal	INFANT	0.547	0.425	0.590	0.425	0.636	0.448
B42	CM de Alcochete	INFANT	0.699	0.480	0.708	0.480	0.746	0.555
B43	CM de Alcoutim	STAR	0.637	0.765	0.786	0.766	0.984	0.824
B44	CM de Alfândega da Fé	STAR	0.620	0.990	0.652	0.991	0.872	1.001
B45	CM de Alijó	INFANT	0.532	0.265	0.559	0.265	0.534	0.271
B46	CM de Aljezur	INFANT	0.532	0.445	0.575	0.445	0.652	0.495
B47	CM de Aljustrel	LEARNER	0.547	0.780	0.564	0.781	0.600	0.822
B48	CM de Almeida	INFANT	0.535	0.465	0.672	0.465	0.631	0.522
B49	CM de Almodôvar	SOLDIER	0.632	0.595	0.725	0.595	0.955	0.668
B50	CM de Alter do Chão	INFANT	0.554	0.295	0.613	0.295	0.791	0.331
B51	CM de Alvito	LEARNER	0.539	0.730	0.586	0.731	0.700	0.820
B52	CM de Amares	SOLDIER	0.924	0.625	1.024	0.626	0.995	0.636
B53	CM de Anadia	INFANT	0.533	0.510	0.539	0.510	0.560	0.547
B54	CM de Arganil	LEARNER	0.787	0.715	0.799	0.716	0.826	0.868
B55	CM de Armamar	INFANT	0.532	0.510	0.555	0.510	0.648	0.518
B56	CM de Arraiolos	INFANT	0.532	0.455	0.582	0.455	0.746	0.510
B57	CM de Arronches	INFANT	0.532	0.245	0.552	0.245	0.667	0.275
B58	CM de Arruda dos Vinhos	LEARNER	0.638	0.840	0.644	0.841	0.676	0.916
B59	CM de Avis	SOLDIER	0.539	0.465	0.643	0.465	0.924	0.522
B60	CM de Barrancos	SOLDIER	1.000	0.495	1.114	0.495	1.009	0.556
B61	CM de Barreiro	STAR	0.787	1.000	0.839	1.000	0.971	1.000
B62	CM de Belmonte	INFANT	0.532	0.535	0.537	0.535	0.576	0.601
B63	CM de Bombarral	STAR	0.878	0.930	0.996	0.931	1.005	0.996
B64	CM de Borba	INFANT	0.535	0.455	0.551	0.455	0.576	0.575
B65	CM de Boticas	SOLDIER	1.000	0.585	1.064	0.586	1.002	0.594
B66	CM de Bragança	STAR	0.849	1.000	0.959	1.002	0.977	1.000
B67	CM de Cadaval	LEARNER	0.533	0.790	0.558	0.791	0.636	0.995
B68	CM de Castelo de Paiva	INFANT	0.533	0.350	0.548	0.350	0.558	0.357
B69	CM de Castelo de Vide	INFANT	0.548	0.330	0.587	0.330	0.634	0.371
B70	CM de Castro Daire	INFANT	0.529	0.365	0.532	0.367	0.542	0.595
B71	CM de Castro Marim	INFANT	0.557	0.575	0.711	0.576	0.822	0.635
B72	CM de Castro Verde	INFANT	0.534	0.495	0.580	0.495	0.580	0.522
B73	CM de Celorico da Beira	INFANT	0.532	0.465	0.575	0.465	0.611	0.521
B74	CM de Chaves	LEARNER	0.575	0.855	0.586	0.856	0.651	0.918
B75	CM de Condeixa-a-Nova	SOLDIER	1.000	0.555	1.036	0.556	1.003	0.595
B76	CM de Constância	LEARNER	0.532	0.680	0.569	0.681	0.611	0.762
B77	CM de Crato	INFANT	0.532	0.460	0.552	0.460	0.604	0.516
B78	CM de Cuba	SOLDIER	0.539	0.550	0.637	0.550	0.872	0.617
B79	CM de Entroncamento	LEARNER	0.532	0.710	0.552	0.711	0.664	0.846
B80	CM de Espinho	LEARNER	0.684	0.775	0.692	0.777	0.856	0.902
B81	CM de Estremoz	SOLDIER	0.727	0.700	0.735	0.701	0.881	0.738
B82	CM de Évora	INFANT	0.601	0.360	0.721	0.360	0.653	0.376
B83	CM de Felgueiras	INFANT	0.700	0.595	0.725	0.595	0.774	0.598
B84	CM de Ferreira do Alentejo	STAR	0.722	0.885	0.777	0.886	0.957	0.934
B85	CM de Figueira de Castelo Rodrigo	INFANT	0.622	0.660	0.648	0.661	0.842	0.739
B86	CM de Fornos de Algodres	LEARNER	0.539	0.695	0.617	0.696	0.667	0.968
B87	CM de Fronteira	INFANT	0.532	0.560	0.552	0.560	0.584	0.628
B88	CM de Gavião	STAR	0.533	0.700	0.592	0.701	0.919	0.786
B89	CM de Golegã	INFANT	0.688	0.425	0.703	0.425	0.803	0.477
B90	CM de Gouveia	INFANT	0.532	0.260	0.564	0.260	0.674	0.279
B91	CM de Grândola	STAR	0.604	0.930	0.874	0.931	0.980	0.981
B92	CM de Guarda	INFANT	0.532	0.395	0.552	0.395	0.572	0.419
B93	CM de Lagoa	INFANT	0.533	0.620	0.550	0.621	0.644	0.656
B94	CM de Lagos	SOLDIER	0.818	0.595	0.839	0.596	0.919	0.629
B95	CM de Lamego	INFANT	0.532	0.405	0.569	0.405	0.766	0.434
B96	CM de Loulé	INFANT	0.531	0.445	0.540	0.445	0.570	0.472
B97	CM de Lourinhã	LEARNER	0.532	0.785	0.582	0.786	0.646	0.842

(*continued on next page*)

Table A.1 (*continued*).

Utility ID	Utility	Category	Deterministic CI		Robust CI		Robust Conditional CI	
			RISI	AMMI	RISI	AMMI	RISI	AMMI
B98	CM de Lousada	SOLDIER	1.000	0.165	1.079	0.165	1.000	0.177
B99	CM de Macedo de Cavaleiros	LEARNER	0.740	0.820	0.748	0.821	0.800	0.833
B100	CM de Mangualde	STAR	0.713	0.755	1.131	0.756	0.999	0.810
B101	CM de Manteigas	INFANT	0.530	0.505	0.534	0.505	0.547	0.568
B102	CM de Marinha Grande	STAR	0.761	0.880	0.825	0.881	0.996	0.934
B103	CM de Marvão	LEARNER	0.530	0.375	0.535	0.377	0.571	0.986
B104	CM de Mealhada	SOLDIER	0.616	0.595	0.688	0.596	0.894	0.702
B105	CM de Mêda	INFANT	0.603	0.415	0.672	0.415	0.765	0.467
B106	CM de Melgaço	STAR	0.927	0.820	1.039	0.821	1.000	0.833
B107	CM de Mértola	LEARNER	0.656	0.655	0.821	0.655	0.821	0.779
B108	CM de Miranda do Corvo	SOLDIER	0.754	0.490	0.861	0.490	0.994	0.525
B109	CM de Miranda do Douro	INFANT	0.532	0.580	0.556	0.580	0.537	0.589
B110	CM de Mirandela	INFANT	0.532	0.525	0.549	0.525	0.535	0.534
B111	CM de Mogadouro	SOLDIER	1.000	0.430	1.060	0.430	1.001	0.438
B112	CM de Moimenta da Beira	INFANT	0.590	0.260	0.593	0.260	0.605	0.265
B113	CM de Moita	INFANT	0.614	0.680	0.630	0.681	0.848	0.680
B114	CM de Monção	INFANT	0.723	0.575	0.734	0.575	0.778	0.586
B115	CM de Mondim de Basto	SOLDIER	0.855	0.265	1.093	0.265	1.004	0.271
B116	CM de Monforte	INFANT	0.529	0.395	0.530	0.395	0.541	0.443
B117	CM de Montalegre	SOLDIER	0.615	0.215	0.661	0.215	0.914	0.222
B118	CM de Montemor-o-Novo	STAR	0.763	0.750	0.796	0.751	0.975	0.790
B119	CM de Mora	STAR	1.000	0.815	1.090	0.816	1.014	0.996
B120	CM de Moura	INFANT	0.531	0.510	0.537	0.510	0.556	0.537
B121	CM de Mourão	INFANT	0.532	0.635	0.548	0.636	0.606	0.719
B122	CM de Nelas	LEARNER	0.566	0.830	0.587	0.831	0.657	0.890
B123	CM de Nisa	SOLDIER	0.926	0.225	1.029	0.225	1.010	0.252
B124	CM de Óbidos	INFANT	0.532	0.650	0.558	0.650	0.599	0.697
B125	CM de Odemira	LEARNER	0.590	0.945	0.614	0.946	0.674	0.999
B126	CM de Oleiros	LEARNER	0.590	0.755	0.675	0.756	0.800	0.846
B127	CM de Oliveira de Frades	SOLDIER	0.845	0.640	0.868	0.641	1.001	0.717
B128	CM de Oliveira do Hospital	SOLDIER	0.535	0.630	0.596	0.631	0.886	0.676
B129	CM de Ourique	LEARNER	0.576	0.860	0.610	0.861	0.716	0.964
B130	CM de Palmela	LEARNER	0.769	0.635	0.792	0.636	0.821	0.776
B131	CM de Penalva do Castelo	INFANT	0.532	0.370	0.568	0.370	0.592	0.415
B132	CM de Penamacor	INFANT	0.532	0.555	0.554	0.555	0.602	0.622
B133	CM de Penedono	SOLDIER	0.776	0.280	0.791	0.280	0.969	0.287
B134	CM de Pinhel	INFANT	0.553	0.465	0.590	0.465	0.690	0.499
B135	CM de Pombal	STAR	0.867	0.895	0.891	0.896	0.942	0.960
B136	CM de Ponte da Barca	SOLDIER	0.658	0.730	0.748	0.731	0.883	0.741
B137	CM de Ponte de Sor	SOLDIER	1.000	0.590	1.084	0.590	1.001	0.622
B138	CM de Portel	INFANT	0.532	0.555	0.559	0.555	0.669	0.623
B139	CM de Porto de Mós	INFANT	0.683	0.570	0.738	0.570	0.816	0.611
B140	CM de Póvoa de Lanhoso	LEARNER	0.629	0.910	0.951	0.911	0.803	0.924
B141	CM de Póvoa de Varzim	INFANT	0.535	0.525	0.592	0.525	0.723	0.525
B142	CM de Proença-a-Nova	SOLDIER	0.578	0.415	0.830	0.415	0.960	0.447
B143	CM de Redondo	LEARNER	0.547	0.875	0.587	0.876	0.638	0.982
B144	CM de Reguengos de Monsaraz	STAR	0.666	0.780	0.994	0.781	0.991	0.822
B145	CM de Resende	INFANT	0.532	0.325	0.549	0.325	0.621	0.333
B146	CM de Ribeira de Pena	SOLDIER	0.723	0.610	0.831	0.611	0.950	0.620
B147	CM de Rio Maior	INFANT	0.532	0.630	0.556	0.631	0.579	0.666
B148	CM de Sabugal	INFANT	0.531	0.335	0.538	0.335	0.592	0.422
B149	CM de Santiago do Cacém	SOLDIER	1.000	0.610	1.192	0.611	1.000	0.637
B150	CM de São Brás de Alportel	STAR	0.740	0.785	0.749	0.786	0.879	0.853
B151	CM de São João da Pesqueira	INFANT	0.537	0.585	0.597	0.585	0.813	0.594
B152	CM de São Pedro do Sul	INFANT	0.669	0.590	0.679	0.591	0.785	0.632
B153	CM de Sátão	STAR	0.735	0.880	0.811	0.881	0.999	0.985
B154	CM de Seia	INFANT	0.653	0.645	0.756	0.645	0.820	0.691
B155	CM de Seixal	INFANT	0.540	0.615	0.563	0.616	0.709	0.615
B156	CM de Sernancelhe	SOLDIER	0.650	0.315	0.672	0.315	0.908	0.323
B157	CM de Serpa	LEARNER	0.534	0.840	0.585	0.841	0.594	0.956
B158	CM de Sertã	INFANT	0.532	0.650	0.558	0.651	0.598	0.697
B159	CM de Sesimbra	SOLDIER	0.888	0.665	0.969	0.666	1.000	0.737
B160	CM de Silves	INFANT	0.655	0.675	0.666	0.676	0.750	0.700
B161	CM de Sines	STAR	0.872	0.765	0.902	0.766	0.999	0.836
B162	CM de Sobral de Monte Agraço	INFANT	0.532	0.545	0.553	0.545	0.623	0.584
B163	CM de Sousel	SOLDIER	1.000	0.555	1.088	0.555	1.076	0.623
B164	CM de Tábua	SOLDIER	1.000	0.315	1.464	0.315	1.012	0.355
B165	CM de Tabuaço	SOLDIER	1.000	0.145	1.211	0.145	1.005	0.148
B166	CM de Tarouca	LEARNER	0.529	0.705	0.534	0.709	0.568	0.964
B167	CM de Terras de Bouro	INFANT	0.538	0.360	0.561	0.360	0.615	0.367
B168	CM de Vale de Cambra	STAR	1.000	0.985	2.082	0.986	1.024	0.999
B169	CM de Valpaços	STAR	1.000	0.600	1.279	0.601	1.003	0.925
B170	CM de Vendas Novas	LEARNER	0.532	0.715	0.556	0.716	0.568	0.851

(*continued on next page*)

Table A.1 (*continued*).

Utility ID	Utility	Category	Deterministic CI		Robust CI		Robust Conditional CI	
			RISI	AMMI	RISI	AMMI	RISI	AMMI
B171	CM de Viana do Alentejo	STAR	0.853	0.830	0.982	0.831	0.915	0.931
B172	CM de Vidigueira	LEARNER	0.532	0.670	0.618	0.671	0.782	0.752
B173	CM de Vieira do Minho	INFANT	0.533	0.715	0.606	0.716	0.702	0.725
B174	CM de Vila de Rei	STAR	1.000	0.680	1.316	0.681	1.002	0.762
B175	CM de Vila do Bispo	INFANT	0.531	0.435	0.541	0.435	0.573	0.489
B176	CM de Vila Flor	INFANT	0.539	0.625	0.547	0.625	0.612	0.634
B177	CM de Vila Nova de Cerveira	SOLDIER	1.000	0.705	1.114	0.706	1.000	0.714
B178	CM de Vila Nova de Famalicão	INFANT	0.682	0.650	0.691	0.651	0.727	0.654
B179	CM de Vila Nova de Foz Coa	INFANT	0.532	0.315	0.562	0.315	0.536	0.321
B180	CM de Vila Pouca de Aguiar	INFANT	0.530	0.565	0.533	0.565	0.533	0.575
B181	CM de Vila Velha de Ródão	INFANT	0.560	0.575	0.603	0.575	0.685	0.644
B182	CM de Vila Verde	STAR	0.763	0.765	0.784	0.769	0.861	0.765
B183	CM de Vimioso	SOLDIER	0.740	0.645	0.754	0.646	0.939	0.654
B184	CM de Vinhais	SOLDIER	1.000	0.405	1.082	0.405	1.000	0.414
B185	CM de Vouzela	LEARNER	0.534	0.705	0.564	0.706	0.561	0.790
B186	EMAR de Portimão	STAR	0.743	0.800	0.775	0.801	0.887	0.812
B187	EMAS de Beja	STAR	0.805	0.985	0.879	0.986	0.934	1.000
B188	EPAL	STAR	0.973	0.990	0.997	0.999	1.000	1.000
B189	Esposende Ambiente	STAR	0.539	0.905	0.683	0.906	0.903	0.910
B190	FAGAR - Faro	STAR	0.821	0.925	0.844	0.930	0.930	0.993
B191	Indaqua Fafe	LEARNER	0.535	1.000	0.589	1.001	0.828	1.000
B192	Indaqua Feira	STAR	0.832	1.000	1.008	1.001	1.000	1.000
B193	Indaqua Matosinhos	STAR	1.000	1.000	1.027	1.005	1.000	1.000
B194	Indaqua Oliveira de Azeméis	STAR	0.728	0.970	0.784	0.979	0.956	0.987
B195	Indaqua Santo Tirso/Trofa	STAR	0.645	1.000	0.984	1.001	1.000	1.000
B196	Indaqua Vila do Conde	STAR	0.835	1.000	0.996	1.001	1.000	1.000
B197	INFRALOBO	SOLDIER	0.758	0.635	0.878	0.636	0.925	0.648
B198	INFRAMOURA	STAR	0.872	0.980	0.911	0.981	0.999	0.995
B199	INFRAQUINTA	STAR	1.000	0.985	1.106	0.986	1.000	1.000
B200	INFRATRÓIA	SOLDIER	0.778	0.515	0.796	0.515	0.963	0.528
B201	INOVA	STAR	1.000	1.000	1.103	1.003	1.000	1.000
B202	Penafiel Verde	STAR	0.855	0.870	0.896	0.871	0.979	0.875
B203	SIMAR de Loures e Odivelas	STAR	0.787	0.835	0.822	0.836	0.946	0.835
B204	SIMAS de Oeiras e Amadora	STAR	0.818	0.945	0.936	0.947	1.000	0.979
B205	SM de Abrantes	STAR	0.907	0.915	1.004	0.916	1.000	0.972
B206	SM de Alcobaça	INFANT	0.646	0.425	0.690	0.426	0.776	0.521
B207	SM de Castelo Branco	STAR	0.943	0.910	1.074	0.911	1.000	0.971
B208	SM de Nazaré	LEARNER	0.645	0.770	0.654	0.771	0.695	0.839
B209	SMAS de Almada	STAR	0.787	0.885	0.798	0.887	0.920	0.929
B210	SMAS de Caldas da Rainha	STAR	0.710	0.800	0.757	0.801	0.882	0.849
B211	SMAS de Leiria	SOLDIER	0.820	0.585	0.839	0.585	0.920	0.620
B212	SMAS de Mafra	STAR	0.941	0.835	0.990	0.836	1.000	0.936
B213	SMAS de Montijo	SOLDIER	0.927	0.460	0.978	0.460	1.000	0.510
B214	SMAS de Peniche	LEARNER	0.639	0.770	0.716	0.771	0.821	0.818
B215	SMAS de Sintra	STAR	0.743	0.795	0.772	0.796	0.985	0.795
B216	SMAS de Torres Vedras	LEARNER	0.746	0.700	0.768	0.701	0.799	0.743
B217	SMAS de Vila Franca de Xira	LEARNER	0.787	0.730	0.820	0.731	0.853	0.807
B218	SMAS de Viseu	SOLDIER	0.818	0.530	0.886	0.530	0.896	0.562
B219	SMAT de Portalegre	SOLDIER	0.604	0.395	0.680	0.395	0.954	0.427
B220	SMEAS de Maia	STAR	0.571	0.745	0.628	0.746	0.926	0.746
B221	Taviraverde	STAR	0.675	0.975	0.725	0.979	0.897	1.000
B222	Tejo Ambiente	INFANT	0.532	0.610	0.547	0.610	0.597	0.631
B223	VIMÁGUA	STAR	0.874	0.845	0.904	0.849	0.979	0.874

References

Alegre, H., 2010. Is strategic asset management applicable to small and medium utilities? Water Sci. Technol. 62 (9), 2051–2058.

Alegre, H., Amaral, R., Brito, R., Baptista, J., 2020. Public policies as strategic asset management enablers: The case of Portugal. H2Open J. 3 (1), 428–436.

Almeida, N.M.d., Vieira, J., Silva, J.G., et al., 2021. The impact of asset management development programs in infrastructure organizations. In: Sustainability and Automation in Smart Constructions. Springer, pp. 247–258.

Alsharif, K., Feroz, E.H., Klemer, A., Raab, R., 2008. Governance of water supply systems in the Palestinian Territories : A data envelopment analysis approach to the management of water resources. J. Environ. Manag. 87 (1), 80–94.

Amaral, R., Alegre, H., Matos, J., 2017. Highlights of key international water infrastructure asset management initiatives, and trends, challenges and developments in Portugal. Water Policy 19 (1), 128–146.

Amaral, A.L., Martins, R., Dias, L.C., 2022. Efficiency benchmarking of wastewater service providers: An analysis based on the Portuguese case. J. Environ. Manag. 321, 115914.

Amaral, A.L., Martins, R., Dias, L.C., 2023. Drivers of water utilities' operational performance–An analysis from the Portuguese case. J. Clean. Prod. 136004.

APDA, 2019. Principais resultados do inquérito elaborado pela Comissão Especializada de Gestão de Ativos (CEGA). Available on https://www.apda.pt/site/ficheiros_eventos/202104121030-inquerito_cega_2019.pdf.

ASCE, 2021. 2021 Report card for America's infrastructure: Drinking water. American Society of Civil Engineers. Available on https://infrastructurereportcard.org/cat-item/drinking-water-infrastructure/.

Baptista, J.M., 2014. The Regulation of Water and Waste Services. IWA Publishing.

Berg, S., Lin, C., 2008. Consistency in performance rankings: The Peru water sector. Appl. Econ. 40 (6), 793–805.

Berg, S., Marques, R., 2011. Quantitative studies of water and sanitation utilities: a benchmarking literature survey. Water Policy 13 (5), 591–606.

Berkelaar, M., et al., 2023. lpSolve: Interface to 'Lp_solve' v. 5.5 to solve linear/integer programs. R package version 5.6.18.

Bowlin, W.F., 1998. Measuring performance: An introduction to data envelopment analysis (DEA). J. Cost Anal. 15 (2), 3–27.

Byrnes, J., Crase, L., Dollery, B., Villano, R., 2010. The relative economic efficiency of urban water utilities in regional New South Wales and Victoria. Resour. Energy Econ. 32 (3), 439–455.

Cabral, M., Loureiro, D., Covas, D., 2019. Using economic asset valuation to meet rehabilitation priority needs in the water sector. Urban Water J. 16, 205–214.

Cardoso, M.A., Silva, M.S., Coelho, S.T., Almeida, M.C., Covas, D.I., 2012. Urban water infrastructure asset management - A structured approach in four water utilities. Water Sci. Technol. 66, 2702–2711.
Carriço, N., Covas, D.I., Almeida, M.C., Leitão, J.P., Alegre, H., 2012. Prioritization of rehabilitation interventions for urban water assets using multiple criteria decision-aid methods. Water Sci. Technol. 66 (5), 1007–1014.
Carvalho, P., Marques, R.C., 2011. The influence of the operational environment on the efficiency of water utilities. J. Environ. Manag. 92 (10), 2698–2707.
Carvalho, P., Marques, R.C., 2014. Computing economies of vertical integration, economies of scope and economies of scale using partial frontier nonparametric methods. European J. Oper. Res. 234 (1), 292–307.
Carvalho, P., Marques, R.C., 2016. Estimating size and scope economies in the Portuguese water sector using the Bayesian stochastic frontier analysis. Sci. Total Environ. 544, 574–586.
Cazals, C., Florens, J.-P., Simar, L., 2002. Nonparametric frontier estimation: a robust approach. J. Econometrics 106, 1–25.
Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European J. Oper. Res. 2 (6), 429–444.
Charnes, A., Cooper, W., Thrall, R.M., 1991. A structure for classifying and characterizing efficiency and inefficiency in data envelopment analysis. J. Prod. Anal. 2 (3), 197–237.
Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., 2007. An introduction to 'benefit of the doubt' composite indicators. Soc. Indic. Res. 82, 111–145.
Correia, T., Marques, R., 2011. Performance of Portuguese water utilities: How do ownership, size, diversification and vertical integration relate to efficiency? Water Policy 13 (3), 343–361.
Daraio, C., Simar, L., 2005. Introducing environmental variables in nonparametric frontier models: a probabilistic approach. J. Prod. Anal. 24, 93–121.
Daraio, C., Simar, L., 2007. Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications. Springer Science & Business Media.
De Witte, K., Kortelainen, M., 2013. What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. Appl. Econ. 45 (17), 2401–2412.
De Witte, K., Marques, R.C., 2010. Designing performance incentives, an international benchmark study in the water sector. CEJOR Cent. Eur. J. Oper. Res. 18 (2), 189–220.
De Witte, K., Schiltz, F., 2018. Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional Benefit-of-the-Doubt approach. European J. Oper. Res. 267 (3), 1172–1181.
D'Inverno, G., Carosi, L., Romano, G., 2021. Environmental sustainability and service quality beyond economic and financial indicators: A performance evaluation of Italian water utilities. Soc.-Econ. Plan. Sci. 75.
D'Inverno, G., De Witte, K., 2020. Service level provision in municipalities: A flexible directional distance composite indicator. European J. Oper. Res. 286, 1129–1141.
Dong, X., Du, X., Li, K., Zeng, S., Bledsoe, B., 2018. Benchmarking sustainability of urban water infrastructure systems in China. J. Clean. Prod. 170, 330–338.
Duarte, A., Rodrigues, G., Ramos, R., 2009. A global service quality index to evaluate the performance and sustainability in water supply utilities. WSEAS Trans. Environ. Dev. 5 (12), 759–769.
ERSAR, 2021a. Relatório Anual dos Serviços de Águas e Resíduos em Portugal (2021): Volume 1 - Caracterização do Setor de Águas e Resíduos 1. Issued by ERSAR - Entidade Reguladora dos Serviços de Águas e Resíduos and LNEC - Laboratório Nacional de Engenharia Civil. Available on https://www.ersar.pt/pt/site-publicacoes/Paginas/edicoes-anuais-do-RASARP.aspx.
ERSAR, 2021b. Relatório Anual dos Serviços de Águas e Resíduos em Portugal (2021): Volume 2 – Controlo da qualidade da água para consumo humano 2. Issued by ERSAR - Entidade Reguladora dos Serviços de Águas e Resíduos and LNEC - Laboratório Nacional de Engenharia Civil. Available on https://www.ersar.pt/pt/site-publicacoes/Paginas/edicoes-anuais-do-RASARP.aspx.
ERSAR, LNEC, 2017. Desenvolvimento e implementação de processos de gestão patrimonial de infraestruturas - guia técnico 21. Sér. Guias Técnicos 21, 1–165, Issued by ERSAR - Entidade Reguladora dos Serviços de Águas e Resíduos and LNEC - Laboratório Nacional de Engenharia Civil. Available on http://www.ersar.pt/pt/publicacoes/publicacoes-tecnicas/guias.
EurEau, 2021. Europe's Water in Figures—An Overview of the European Drinking Water and Waste Water Sectors. The European Federation of National Water Associations Brussels, Belgium.
European Commission, 1998. Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. Available on https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31998L0083.
Ferreira, B., Carriço, N., 2019. Urban water infrastructure asset management plan: Case study. Open Eng. 9 (1), 459–467.
Frade, et al., 2015. PENSAAR 2020: Uma Estratégia a Serviço da População: Serviços de Qualidade a um Preço Sustentável. Volume 1. Available on https://apambiente.pt/agua/plano-estrategico-de-abastecimento-de-agua-e-saneamento-de-aguas-residuais-2020.
Fried, H.O., Lovell, C.K., Schmidt, S.S., Schmidt, S.S., et al., 2008. The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press.
Fusco, E., Vidoli, F., Rogge, N., 2020. Spatial directional robust Benefit of the Doubt approach in presence of undesirable output: An application to Italian waste sector. Omega (United Kingdom) 94.
Goh, K.H., See, K.F., 2021. Twenty years of water utility benchmarking: A bibliometric analysis of emerging interest in water research and collaboration. J. Clean. Prod. 284, 124711.
Gonçalves, I., Alves, D., Robalo, G., 2014. Social tariffs for water and waste services in mainland Portugal: An impact analysis. Water Sci. Technol. Water Supply 14 (4), 513–521.
Hayfield, T., Racine, J.S., 2008. Nonparametric econometrics: The np package. J. Stat. Softw. 27 (5), 1–32.
Henriques, A.A., Camanho, A.S., Amorim, P., Silva, J.G., 2020. Performance benchmarking using composite indicators to support regulation of the Portuguese wastewater sector. Util. Policy 66, 101082.
Henriques, A.A., Fontes, M., Camanho, A.S., D'Inverno, G., Amorim, P., Silva, J.G., 2022. Performance evaluation of problematic samples: a robust nonparametric approach for wastewater treatment plants. Ann. Oper. Res. 315 (1), 193–220.
ISO, 2014. ISO 55000:2014 . Asset Management — Overview, Principles and Terminology. Technical Report, International Organization for Standardization.
Kuosmanen, T., et al., 2002. Modeling blank data entries in data envelopment analysis. In: Econ-WPA Working Paper at WUSTL, Vol. 210001.
Lavigne, C., Jaeger, S.D., Rogge, N., 2019. Identifying the most relevant peers for benchmarking waste management performance: A conditional directional distance Benefit-of-the-Doubt approach. Waste Manage. 89, 418–429.
Leitão, J., Coelho, S., Alegre, H., Cardoso, M., Silva, M., Ramalho, P., Ribeiro, R., Covas, D., Poças, A., Vitorino, D., Almeida, M., Carriço, N., 2016. Moving urban water infrastructure asset management from science into practice. Urban Water Journal 13 (2), 133–141.
Li, Q., Racine, J., 2003. Nonparametric estimation of distributions with categorical and continuous data. J. Multivariate Anal. 86 (2), 266–292.
Lo Storto, C., 2018. Efficiency, conflicting goals and trade-offs: A nonparametric analysis of the water and wastewater service industry in Italy. Sustainability (Switzerland) 10 (4).
Lombana Cordoba, C., Saltiel, G., Penalosa, F.P., 2022. Utility of the Future, Taking Water and Sanitation Utilities Beyond the Next Level. A Methodology to Ignite Transformation in Water and Sanitation Utilities. World Bank, Washington DC.
Loureiro, D., Silva, C., Cardoso, M., Mamade, A., Alegre, H., Rosa, M., 2020. The development of a framework for assessing the energy efficiency in urban water systems and its demonstration in the Portuguese water sector. Water (Switzerland) 12 (1).
Luís, A.M., Almeida, N.M., 2021. Gestão de ativos no setor da água em Portugal: passado, presente e futuro. Rev. APDA Assoc. Port. Distrib. Drenagem Águas 24–33, Available on https://www.apda.pt/pt/pagina/87/revista-apda/.
Machado, B., Carvalho, T., Cupido, C., Almeida, M., Alegre, H., 2009. Minimization of losses in water supply systems: Strategy definition in a Portuguese case study. Desalin. Water Treat. 2 (1–3), 24–29.
Marques, R.C., 2006. A yardstick competition model for Portuguese water and sewerage services regulation. Util. Policy 14 (3), 175–184.
Marques, R.C., 2008. Comparing private and public performance of Portuguese water services. Water Policy 10 (1), 25–42.
Marques, R.C., Berg, S., 2011. Risks, contracts, and private-sector participation in infrastructure. J. Constr. Eng. Manag. 137 (11), 925–932.
Marques, R.C., Berg, S., Yane, S., 2014. Nonparametric benchmarking of Japanese water utilities: Institutional and environmental factors affecting efficiency. J. Water Resour. Plan. Manag. 140 (5), 562–571.
Marques, R., Carvalho, P., 2014. Optimising the water sector market structure in Portugal. J. Water Supply Res. Technol.—AQUA 63 (4), 303–310.
Marques, R.C., De Witte, K., 2010. Towards a benchmarking paradigm in European water utilities. Public Money Manag. 30 (1), 42–48.
Marques, R., De Witte, K., 2011. Is big better? On scale and scope economies in the Portuguese water sector. Econ. Model. 28 (3), 1009–1016.
Marques, R., Monteiro, A., 2001. Application of performance indicators in water utilities management - A case-study in Portugal. Water Sci. Technol. 44 (2–3), 95–102.
Marques, R.C., Monteiro, A., 2003. Application of performance indicators to control losses - Results from the Portuguese water sector. Water Sci. Technol. Water Supply 3 (1–2), 127–133.
Marques, R.C., Pinto, F.S., 2018. How to watch the watchmen? The role and measurement of regulatory governance. Util. Policy 51, 73–81.
Marques, R.C., Simões, P., 2020. Revisiting the comparison of public and private water service provision: an empirical study in Portugal. Water 12 (5), 1477.
Martins, R., Antunes, M., Fortunato, A., 2020. Regulatory changes to Portugal's social tariffs: Carrying water in a sieve? Util. Policy 64, 101038.
Martins, R., Cruz, L., Barata, E., 2013. Water price regulation: A review of Portuguese tariff recommendations. Public Organ. Rev. 13 (2), 197–205.
Matos, M., Baptista, J., 1999. Sustainable development strategies for the water supply and wastewater sectors. Aqua 48 (5), 186–190.
Mbuvi, D., De Witte, K., Perelman, S., 2012. Urban water sector performance in Africa: A step-wise bias-corrected efficiency and effectiveness analysis. Util. Policy 22, 31–40.
Melyn, W., Moesen, W., 1991. Towards a synthetic indicator of macroeconomic performance: Unequal weighting when limited information is available. In: Public Economics Research Paper 17. Center for Economic Studies, KU Leuven.

Mergoni, A., D'Inverno, G., Carosi, L., 2022. A composite indicator for measuring the environmental performance of water, wastewater, and solid waste utilities. Util. Policy 74, 101285.

Morais, P., Camanho, A.S., 2011. Evaluation of performance of European cities with the aim to promote quality of life improvements. Omega 39 (4), 398–409.

Mumssen, Y., Saltiel, G., Kingdom, B., 2018. Aligning Institutions and Incentives for Sustainable Water Supply and Sanitation Services. World Bank, Washington, DC.

Pato, J.H., 2011. História das políticas públicas de abastecimento e saneamento de águas em Portugal. ERSAR/ICS.

Pereira, M.A., Camanho, A.S., Figueira, J.R., Marques, R.C., 2021. Incorporating preference information in a range directional composite indicator: the case of Portuguese public hospitals. European J. Oper. Res. 294 (2), 633–650.

Pérez, F., Delgado-Antequera, L., Gómez, T., 2019. A two-phase method to assess the sustainability of water companies. Energies 12 (13).

Pinto, F.S., Costa, A.S., Figueira, J.R., Marques, R.C., 2017a. The quality of service: An overall performance assessment for water utilities. Omega 69, 115–125.

Pinto, F., Marques, R., 2015. Tariff recommendations: A panacea for the Portuguese water sector? Util. Policy 34, 36–44.

Pinto, F.S., Simões, P., Marques, R.C., 2017b. Water services performance: Do operational environment and quality factors count? Urban Water J. 14, 773–781.

Rogge, N., Jaeger, S.D., Lavigne, C., 2017. Waste performance of NUTS 2-regions in the EU: A conditional directional distance Benefit-of-the-Doubt model. Ecol. Econom. 139, 19–32.

Romano, G., Guerrini, A., 2011. Measuring and comparing the efficiency of water utility companies: A data envelopment analysis approach. Util. Policy 19, 202–209.

Ruiz, J.L., Sirvent, I., 2016. Common benchmarking and ranking of units with DEA. Omega (United Kingdom) 65, 1–9.

Ruiz, J.L., Sirvent, I., 2022. Benchmarking within a DEA framework: setting the closest targets and identifying peer groups with the most similar performances. Int. Trans. Oper. Res. 29, 554–573, 1.

Santos, L., Cardoso, M., Galvão, A., 2022. Storm water systems' performance: assessment framework application to Portuguese water utilities. Int. J. Water Resour. Dev.

Santos, A.C., Rodrigues, J., Teles, N., 2018. Semi-peripheral financialisation and social reproduction: the case of Portugal. New Political Econ. 23 (4), 475–494.

Sarkis, J., 2007. Preparing your data for DEA. In: Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis. Springer, pp. 305–320.

Silvestre, H., Gomes, R., 2017. A resource-based view of utilities: The key-determinant factors for customer prices and organizational costs in the Portuguese water industry. Water Resour. Econ. 19, 41–50.

Teles, N., 2015. Financialisation and Neoliberalism: The Case of Water Provision in Portugal. FESSUD Working Paper Series No. 102, FESSUD, pp. 1–45, Available on http://hdl.handle.net/10316/40969.

Thanassoulis, E., 2000a. DEA and its use in the regulation of water companies. European J. Oper. Res. 127 (1), 1–13.

Thanassoulis, E., 2000b. The use of data envelopment analysis in the regulation of UK water utilities: water distribution. European J. Oper. Res. 126 (2), 436–453.

Thanassoulis, E., Portela, M.C., Despic, O., 2008. Data envelopment analysis: the mathematical programming approach to efficiency analysis. In: The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press New York, pp. 251–420.

Theussl, S., Hornik, K., 2019. Rglpk: R/GNU linear programming kit interface. R package version 0.6-4.

Tourinho, M., Barbosa, F., Santos, P.R., Pinto, F.T., Camanho, A.S., 2022a. Productivity change in Brazilian water services: A benchmarking study of national and regional trends. Soc.-Econ. Plan. Sci.

Tourinho, M., Santos, P.R., Pinto, F.T., Camanho, A.S., 2022b. Performance assessment of water services in Brazilian municipalities: An integrated view of efficiency and access. Soc.-Econ. Plan. Sci. 79, 101139.

United Nations, 2022. The Sustainable Development Goals Report: 2022. Available on https://unstats.un.org/sdgs/report/2022/.

Vilarinho, H., D'Inverno, G., Nóvoa, H., Camanho, A.S., 2023. The measurement of asset management performance of water companies. Soc.-Econ. Plan. Sci. 101545.

Walker, N., Norton, A., Harris, I., Williams, A., Styles, D., 2019. Economic and environmental efficiency of UK and Ireland water companies: Influence of exogenous factors and rurality. J. Environ. Manag. 241, 363–373.

Wang, S., Qiu, S., Ge, S., Liu, J., Peng, Z., 2018. Benchmarking Toronto wastewater treatment plants using DEA window and Tobit regression analysis with a dynamic efficiency perspective. Environ. Sci. Pollut. Res. 25 (32), 32649–32659.

Wüst, C., Rogge, N., 2021. How is the European Union progressing towards its Europe 2020 targets? A benefit-of-the-doubt window analysis. Empirica.

Zanella, A., Camanho, A.S., Dias, T.G., 2013. Benchmarking countries' environmental performance. J. Oper. Res. Soc. 64 (3), 426–438.

Zanella, A., Camanho, A.S., Dias, T.G., 2015. Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. European J. Oper. Res. 245 (2), 517–530.