Socio-Economic Planning Sciences 87 (2023) 101545

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

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

The measurement of asset management performance of water companies

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

[a] Faculty of Engineering, University of Porto, Porto, Portugal
[b] INESC TEC, Porto, Portugal
[c] 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 study explores asset management performance of Portuguese water supply companies operating in the bulk market. The focus of the analysis are the managerial practices and the condition of infrastructures. This assessment is based on the information conveyed by the indicators collected by the Portuguese water and waste services' regulator authority (ERSAR) between 2016 and 2020. The main contribution of this research is to propose innovative methods to enhance the knowledge on asset management practices in the water sector. Two Benefit-of-the-Doubt (BoD) Composite Indicators are developed to highlight different aspects of asset management approaches. The first reflects organisations' performance in maintaining their infrastructures at acceptable operational levels, and the other reveals their maturity in asset management practices. Robust and conditional approaches for estimating the BoD indicators are applied, allowing to obtain results that account for the effect of contextual variables on companies' performance. Additionally, the performance of the companies is analysed over a 5-year period. The results show that there is significant room for improvement given the indicators' values estimated in the benchmarking analysis. The type of management systems and areas of intervention (urban, semi-urban or rural) are factors that present significant impact in asset management performance. The analysis of trends in the evolution of performance over time revealed improvements both in the companies' managerial practices and operational results.

1. Introduction

Sustainability can be defined as development that "meets the needs of the present without compromising the ability of future generations to meet their own needs" [1]. This notion entails balancing three key interconnected factors: economic growth, social inclusion, and environmental preservation. Water availability and its associated services substantially impact all of these elements, making water essential for supporting economic activity, improving societal well-being, and protecting the environment [2]. Goal 6 of the United Nations' Sustainable Development Goals (SDGs) is to "ensure the availability and sustainable management of water and sanitation for all". This goal is closely interlinked with the other SDGs, which demands a structured strategy for managing the resources required to meet the intended targets associated with SDG 6 [3].

The infrastructures designed to provide water services demand special attention due to the serious consequences in case of failures or leakages. A simple water main break can lead to damages or failures to adjacent infrastructures, such as roads, oil or gas distribution systems, besides the direct effects of water supply shortages [4]. Under the framework of asset management, organisations can employ an integrated strategy to ensure that assets will fulfil the intended goals. According to the United Nations Technical Committee for Asset Management Systems (TC-251), asset management represents a key enabler contributing to the achievement of SDGs by organisations. Consequently, there is a natural alignment between asset management and the desires represented in the SDGs [5], which is specially important in what concerns public service utilities, such as water, gas and electricity companies.

Asset management is defined by ISO 55000 as a "coordinated activity of an organisation to realise value from assets". By covering strategy, safety, environment, cost, risk and life cycle, this approach represents more than an extension of maintenance. Value realisation entails balancing costs, risks, opportunities, and performance rewards. However, the concept of value will vary depending on the demands of each organisation and its stakeholders [6]. ISO 55001:2014 [7] is the

* 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.seps.2023.101545
Received 27 July 2022; Received in revised form 2 January 2023; Accepted 19 February 2023
Available online 21 February 2023

international standard that specifies the requirements for an organisation to develop an asset management system including a comprehensive set of tools, rules, processes, and information systems to ensure that the management objectives are satisfied.

According to Luís and Almeida [8], the adoption of asset management strategies in the water sector was triggered by different reasons around the world. The regulation was the primary motivator in the United Kingdom, whilst in Australia and New Zealand the first issue was maintenance optimisation. In the United States and Canada, the critical issue was concerned with asset ageing and deterioration, while in the Netherlands, the emphasis was on establishing and ensuring service levels.

In the first decade of the twenty-first century, the concept of asset management began to be internalised in the water sector in Portugal. Despite the significant involvement of the leading agents, the results of the dissemination and application of asset management are not yet visible in a uniform manner in the national panorama [8]. The water, wastewater and solid waste services are overseen in Portugal by ERSAR, *Entidade Reguladora dos Serviços de Águas e Resíduos*, the sector's regulatory authority. ERSAR's monitoring process relies on the comparison of operators based on performance indicators that are made public. This practice, known as *sunshine regulation*, has successfully encouraged performance improvement in the sector, praising good practices alongside exposing companies to "embarrassment" for bad performance [9].

The extensive set of indicators requested annually by ERSAR to the sector's operators allows the analysis of companies' performance in multiple facets. This study explores asset management practices by selecting and analysing the metrics collected by ERSAR and which are clearly related to that field to perform benchmarking. More specifically, this study is focused on Portuguese water supply companies operating at the bulk level and evaluates performance trends from 2016 to 2020.

The information provided by ERSAR is used to construct composite indicators (CIs) that aggregate the selected metrics to obtain a summary measure that incorporates multiple dimensions. The technique utilised to build the CIs is the Benefit-of-the-Doubt approach (BoD), popularised by Cherchye et al. [10] based on Data Envelopment Analysis (DEA) models. The BoD method was chosen for its ability to estimate the most favourable weights for the unit under consideration when compared to the peers in the sample, so no water company can object to those weights, making this strategy appropriate for sunshine regulation purposes. The robust and conditional formulations in the BoD approach are employed to overcome the effect of outliers and to assess the influence of the environment on companies' performance. A novel visualisation framework for the assessment of companies' performance is also presented.

Benchmarking studies using data collected by regulatory bodies are common in the literature. However, to the best of our knowledge, there are no benchmarking studies using those data with a specific focus on water system's asset management. In summary, this work aims to fill this literature gap by developing two complementary composite indicators focused on asset management performance (namely, the *Resource and Infrastructure Sustainability Index* and the *Asset Management Maturity Index*). The practical relevance of the proposed approach is demonstrated using the information collected by ERSAR to compare the bulk water supply companies operating in Portugal in the period from 2016 to 2020. This period corresponds to the most recent ERSAR's framework described as "third generation of indicators". The composite indicators developed in this work are used to compare the performance of different companies in a given year (a cross sectional approach), as well as to reveal performance trends over a five-year period. This novel evaluation method using ERSAR's data with focus on asset management represents the innovative contribution of this study.

The relevance of this study is justified by the urgent need to establish improvement processes in the management of assets at water systems. Sustainability is a major driver for enhancing water distribution system management. Water is a crucial resource for human life and according to Vieira et al. [11], 30% to 40% of treated water is lost worldwide due to degradation of water system infrastructures. Water and wastewater systems are deteriorating all around the world. Furthermore, because the water sector is capital intensive, and infrastructure expenditures are intended to last for a long time, physical asset management must be particularly efficient. According to Marlow and Burn [12], efficient asset management requires the appropriate selection of metrics related to the inputs and outputs. Monitoring key performance indicators must be an effective practice, capable of providing feedback on the implementation of strategies, in order to guide decision-making and promote improvements in the sector.

The remaining parts of this paper are structured as follows: Section 2 presents a brief literature review, an overview of the water sector in Portugal is discussed in Section 3, the proposed methodology is explained in Section 4, the case study is detailed in Section 5, Section 6 discusses the results, and the conclusions are presented in Section 7.

2. Literature review

The literature review presents the approaches available for benchmarking in the water sector in subsection 2.1, and subsection 2.2 discusses the use of performance measurement techniques in the field of asset management.

2.1. Benchmarking approaches using performance indicators in the water sector

The concept of benchmarking or relative performance assessment implies a systematic comparison between similar entities, known as Decision Making Units (DMUs). Examples of entities that can be considered as DMUs include groups of companies, organisations, countries, projects, among others. The main objective of a benchmarking approach is to foster performance improvement, which can happen implicitly by drawing the attention of DMUs to the issues highlighted in comparative studies. However, more explicitly, benchmarking may lead to incentives or rewards for the DMUs under evaluation, in the form of salary plans, tariff regulations or budget rules. This practice is especially useful when applied in monopoly regulations [13]. Benchmarking practices can also help in conflict resolution by redirecting the focus of stakeholders to performance improvements [14].

Marques and De Witte [15] describe the benefits of benchmarking public service activities. These authors highlight two different perspectives applied in benchmarking approaches: (i) metric benchmarking, which allows organisations to evaluate performance and compare it with competitors, and (ii) process benchmarking, where the companies map their internal processes and look for best practices in the industry to enable superior performance. In that sense, metric benchmarking identifies *what* to improve, whereas process benchmarking emphasises *how* to improve. However, in many cases, because different companies do not typically share information about their performance among themselves, benchmarking studies can only be conducted with the involvement of regulatory entities, who receive data from companies that operate in natural monopoly contexts. To this extent, the aim of the regulators is to stimulate, support and monitor benchmarking processes among organisations, set rules and standards of comparison, collect and publish results, and find out *where* to improve. This process can contribute to identify best practices and guide the design of strategies for improvement.

Regulators seek to create a pseudo-competitive environment, which stimulates companies to raise their efficiency levels and reduce their prices [16]. Independent regulators have legislative, executive, and judicial authority to monitor several operators by enforcing the required regulations. The governance system of regulators is discussed by Marques and Pinto [9], emphasising their independence and responsibility, the interaction with policymakers, operators, and customers, as well as their internal processes including judgement criteria and transparency. Those authors conclude that deficient governance systems may lead to excessive governmental influence in regulator's activities impacting their transparency and accountability.

The governance model that became prominent in the recent decades, known as yardstick competition, reinforces the comparison of the regulated firm's performance with that of other firms in the same sector [17]. Benchmarking instruments applied by the regulator are always included in the various types of yardstick competition. The incentive for the operators to improve their efficiency comes from the information received from other firms so that the regulatory process becomes an artificial competition process among them. Marques [17] explains that there are two main approaches in yardstick competition. The first strategy, which often has more authoritarian characteristics, uses benchmarking to set pricing and decide the operators' tariffs. It is known as *price yardstick competition*. In contrast, the second one, known as *sunshine regulation*, represents a lighter variant of yardstick competition and includes a comparison and public debate on the operators' performance.

In water utilities, regulators perform a macro or top-down benchmarking to get information about the operators' level of performance and set policy targets for the sector. At the same time, the companies themselves use bottom-up benchmarking, looking at their performance to perform a diagnosis and identify areas or activities to improve [15].

Performance indicators can be used to perform benchmark analysis at different levels, and global measures of efficiency are commonly employed by regulators to get information about the operator's performance. They are usually employed as decision support tools to prioritise improvement actions and analyse the effect of previous measures [18]. The use of indicators for performance benchmarking has become a crucial strategy to promote improvements within the water sector [19]. For an comprehensive discussion about the choice of indicators in the water sector, see [18] and [20].

Models based on Data Envelopment Analysis (DEA), originally developed by Charnes et al. [21], represent very useful tools to support benchmarking processes. Non-parametric techniques such as DEA differ from methods that employ production functions with theoretical imposed functional forms or engineering standards. DEA is a data-driven non-parametric method that assesses performance against the best practices observed in a set of DMUs [13].

In a literature review covering 190 studies on water services performance published between 1969 and 2008, Berg and Marques [22] report that 34% of the studies reviewed use non-parametric methods and, among them, 72% apply DEA. In a more recent review, Goh and See [23] confirm the interest in DEA methods in water sector research, since the term "DEA" represents 33.80% of author's keywords used among the studies reviewed that were published from 2000 to 2019. Several benchmarking works applying DEA have dealt with the efficiency of water systems worldwide: Thanassoulis [24,25] and Walker et al. [26] in United Kingdom, Byrnes et al. [27] in Australia, Wang et al. [28] in Canada, Berg and Lin [29] in Peru, Alsharif et al. [30] in Palestina, Dong et al. [31] in China, Marques et al. [32] in Japan, Lo Storto [33], D'Inverno et al. [34] and Romano and Guerrini [35] in Italy, among others. Bogetoft [36,37] developed techniques based on DEA to deal with the regulatory agencies' incentive mechanisms. Those incentive schemes were also addressed in a cross-country study performed by De Witte and Marques [38] that compared the water sector from the Netherlands, England and Wales, Australia, Portugal and Belgium. The results suggest that the incentives have positively impacted the sector's efficiency. In a specific study that examines the adoption of the sunshine regulation in the Netherlands, De Witte and Saal [39] describe the effectiveness of this approach examining data from different periods before and after the employment of sunshine regulation, by using DEA methods. Those authors conclude that the adoption of sunshine regulation beneficially resulted in higher productivity, that was transferred to customers as price reductions.

Techniques based on DEA may also be employed for the construction of composite indicators (CIs). CIs entail the combined analysis of a set of performance indicators to compare multiple-dimensional activities. According to Vilanova et al. [18], even though the collection of data and generation of multiple indicators represent a complicated process, the aggregation of those indicators into an overall measure of performance may be even more challenging involving creativity and experienced judgement. The use of a method based on DEA presents the advantage of being data-driven, avoiding the extensive interaction with stakeholders to decide the relative importance of indicators. This strategy known as the "Benefit-of-the-Doubt" (BoD) approach overcomes the concerns about the need for normalisation and identification of "right" weights, allowing an easy and intuitive interpretation of results [10,40]. BoD models were initially proposed for macroeconomic performance assessment [41] and have been extensively applied in many areas such as transportation [42], competitiveness [43,44], human development [45,46], quality of life [47], social inclusion [48], public health [49], environmental performance [50], and active ageing of population [51].

The standard DEA models, including BoD, present the inconvenience of being too sensitive to outliers and not allowing statistical inference. Several approaches have been proposed in the literature to tackle these issues. For example, detection outlier procedures or *robust* approaches have been introduced to mitigate the impact of outlying observations (see all the discussion in [52]). In addition, *one-stage* or *two-stage* approaches have been suggested to investigate the influence of external conditions on the efficiency estimates (see for example [53,54]). In this vein, the works of Henriques et al. [19], Molinos-Senante et al. [55], Dong et al. [56] and Romano and Guerrini [35] applied DEA methods to evaluate water systems.

An alternative method was developed by Daraio and Simar [57,58] to compute conditional scores while accounting for the influence of external factors directly in the efficiency score estimation (*conditional* approach). Then, the influence of exogenous variables on the performance is estimated using a smoothed non parametric regression between the ratio of conditional and unconditional efficiencies. However, those studies allowed the appraisal of the context factors using only continuous variables. De Witte and Kortelainen [59] introduced the use of both continuous and discrete variables as external factors. Since then, many studies have adopted that approach for evaluating water systems, such as De Witte and Marques [38], Marques et al. [32], D'Inverno et al. [34] and Mergoni et al. [60].

In the literature review issued by Berg and Marques [22], about 35% of the non parametric studies analysed the context of water utilities using explanatory exogenous factors, and more than twenty different variables were identified as being used in those studies. Those exogenous variables include customer density, proportion of non-residential customers, peak factor, and water losses. Tourinho et al. [61] presented an overview of the contextual variables used in studies that deal with performance of water supply systems. According to those authors, the most frequent contextual variables used in the literature are: ownership, regional differences, scope of services, customer density, population density, water source, water losses and peak factor (ratio between the highest and the average water consumption within a month).

2.2. Performance measurement in asset management

Indicators are frequently used to measure performance and make decisions in asset management. At the strategic level, an asset management system emphasises key performance indicators in conformity with higher-level objectives. By doing that, the alignment between asset management objectives and business objectives can be pursued [62]. Galar et al. [63] and Cecconi et al. [64] discuss the popularity of indicators as decision-making tools for asset management. The selection of the most suitable performance indicators for asset maintenance was addressed by Gonçalves et al. [65], and Dutuit and Rauzy [66] analysed 'importance measures' applied to complex components. Attwater et al. [67] investigated the state of play of performance measurement for asset management systems. Their findings revealed that it is still an unsettled issue how to measure the performance of asset management systems. Further research is needed to understand the linkage between organisation performance, asset performance and asset management performance.

Galar et al. [63] compares the use of individual indicators versus aggregated metrics in the form of composite indicators to measure asset management performance. At the core of this discussion, there is the possible loss of information that arises when aggregating many indicators and the resulting misconception or misunderstanding of the actual phenomenon. This idea is counterbalanced by the fact that an aggregate indicator can be more intuitive and simpler to communicate for managers. For this reason, in asset management, weight summations using aggregating weights provided by specialists are the most frequently used, even though direct ratios between pairs of indicators are also frequently employed (e.g., the maintenance cost divided by the asset replacement value). Statistical techniques can also be used to perform aggregated metrics. Galar et al. [63] also highlight the use of some statistical techniques such as principal component analysis (PCA) in setting suitable weights. Other types of aggregation strategies found in the literature are the fuzzy logic [68,69] and Analytical Hierarchical Process (AHP) [70]. To our knowledge, asset management performance at the corporate level has not been measured using composite indicators based on BoD and it is therefore the object of this work. Due to the complexity of this subject and the many dimensions involved, a vast unexplored area of research exists [63].

3. The water sector in Portugal

In recent decades, Portugal has gone through substantial changes related to water supply services, mainly concerning service access, quality of service and structure of the market. Before 1993, the public sector had full ownership of water services. That was modified by Executive Law No. 372, which promoted the water sector restructuring, allowing the private capital to participate in the sector and establishing a regulatory authority to deal with water services. Since this period the service coverage increased from 81% to 96% and the acceptable level of water quality raised from 50% to 99% [71].

In the Portuguese water sector, the national regulatory agency, ERSAR, specifies a set of key performance indicators and collects data for each operator. Following the sunshine regulation approach, the results are publicly disclosed. The powers of ERSAR are not coercive, and the regulator does not actively engage in the pricing formulation process [72].

Another consequence of the updated water sector organisation after 1993 is the separation between bulk or wholesale systems and retail systems, which occurred both in the water and wastewater businesses [73]. The water supply wholesale companies are responsible for water abstraction, treatment and storage before distributing the water to the retail companies that supply water to end-users.

Portuguese water companies can currently be managed according to three different models, namely direct management, delegation and concession. In the direct management model, municipalities, municipal services and associations of municipalities own and operate the water services, usually without participation of private companies. The delegation model works with a municipal company or a company established in partnership with the State (municipal or state company), parishes, or user associations. In the delegation system, the company is owned and controlled exclusively by the State (central, municipal or both), without a contract of concession. However a contract of management must be celebrated, defining goals and tariff policies for the operator. In the concession, a municipal concessionaire or public–private partnership with municipalities and other private operators is established under a long term contract, usually from 30 to 50 years. The participation of private capital is allowed mainly in the delegation and concession models, and eventually in the direct management in case of partnership with State or municipalities [74–76].

According to the annual report issued by ERSAR in 2021 [76,77], in Portugal, there are ten companies operating in the wholesale water supply market. Those companies and their identification codes used in the study are: Águas de Santo André (A1), Águas do Algarve (A2), Águas do Douro e Paiva (A3), Águas do Centro Litoral (A4), Águas do Norte (A5), Águas do Vale do Tejo (A6), Águas do Vouga (A7), Águas Públicas do Alentejo (A8), EPAL (A9) and ICOVI (A10). The wholesale companies are predominantly managed by concession (seven companies). The other three wholesale companies are managed by delegation. The retail water sector includes 233 companies, and most of them are managed directly by municipalities.

The indicator system used by ERSAR for benchmarking practices is detailed in Technical Guide 22 [78]. The volume of information annually acquired from the operators is vast, comprising water, wastewater and solid waste services. In the case of water supply companies, the performance indicator system of ERSAR presents 14 main metrics, grouped in three different dimensions: (i) *Adequacy of the Interaction with the User*, (ii) *Service Management Sustainability* and (iii) *Environmental Sustainability*.

The ERSAR indicators directly related to asset management are included in the subgroup *Infrastructure Sustainability*, in the dimension of *Service Management Sustainability*. The other subgroups in this dimension are *Economic Sustainability* and *Physical Productivity of Human Resources*. The *Infrastructure Sustainability* subgroup contains two main indicators: *pipeline rehabilitation* (%/year) and *occurrence of pipeline failures* (number of failures/100 km/year).

The dimension *Environmental Sustainability* in subgroup *Efficiency of Utilisation of Environmental Resources* includes also two indicators that are related to asset management and its effect on the use of resources: *actual water losses* (m^3/year) and *energy efficiency of pumping stations* (kWh/(m^3.100 m)).

Additional metrics regarding asset management status are also collected by ERSAR, including the Infrastructure Knowledge Index, the Infrastructure Asset Management Index, the Infrastructure Current Value and the Infrastructure Replacement Cost. All this information has been informed annually by wholesale and retail companies.

The Portuguese water sector has been explored by several works that employed benchmarking techniques using DEA, such as Marques [17], De Witte and Marques [38] and Henriques et al. [52]. ERSAR's indicators in a BoD composite-indicator approach are utilised by Henriques et al. [19] to identify best practices and foster continuous improvement in wastewater operators. Mergoni et al. [60] employs also ERSAR's indicators in a BoD approach to evaluate the environmental performance of Portuguese utilities. The quality of water supply service is evaluated by Pinto et al. [79,80] using ERSAR's metrics. Those benchmarking studies take advantage of using indicators developed under the procedures of ERSAR system, such as submission of data, validation and processing of results [79]. None of these studies applied ERSAR metrics to assess the performance of water companies with a focus in asset management, which reinforces the innovative nature of this research.

In terms of asset management performance, the water systems in Portugal present quite heterogeneous results. In a survey conducted by the *Specialised Commission for Asset Management* from the *Portuguese Association for Water Distribution and Drainage (APDA - Associação Portuguesa de Distribuição e Drenagem de Águas)* in 2019, the results, including both retail and wholesale companies, show that in 54% of the companies do not follow asset management practices. From the companies that claim to have an asset management system, 41% do not set objectives for asset management and 57% work on asset issues using staff that is not dedicated only to that task. Only 4% of the water supply companies present a certification in ISO 55001. A considerable number of companies do not undertake asset condition analysis, and when they do, visual inspections prevail. Many businesses still do not do preventative maintenance. There is a significant reliance on paper and spreadsheet-based records. These results are worse in retail companies compared to bulk systems [81]. Based on such findings, there is a significant space for enhancement.

4. Methodology

The methodology we propose consists of three steps. The first one consists of identifying the metrics that should be considered in the construction of the composite indicators (CIs). The second one deals with the development of a deterministic approach to compute the CIs. Finally, the third step describes the calculation and evaluation of the robust and conditional CIs, accounting for contextual factors.

4.1. Construction of composite indicators (CIs)

This subsection presents the method of selecting the measures used for building the CIs in this study. For the construction of the CIs, we selected metrics among the data collected by ERSAR that reflect asset management practices in two distinct perspectives. Those metrics are aggregated to generate two different composite indicators.

Luís and Almeida [8] explain that the practical results of adopting an asset management philosophy do not become apparent immediately after the start of asset management development programmes in organisations. These programmes typically require several years to effectively implement an asset management culture before the full material benefits become visible. As a result, managerial practices may be implemented, but the tangible results may not instantly reflect their impact on company performance. This fact supports the approach adopted in this study to develop one indicator that indicates tangible operational achievements (*Resource and Infrastructure Sustainability Index — RISI*) and another that represents the maturity stage in management systems (*Asset Management Maturity Index — AMMI*).

4.1.1. The Resource and Infrastructure Sustainability Index - RISI

The first CI is related to the companies' performance for the activities that aim to keep their infrastructures at suitable and sustainable operational levels. In that sense, the companies' tangible results in asset management can be expressed by this indicator. We named this indicator as *Resource and Infrastructure Sustainability Index (RISI)*. The RISI is made up of the following ERSAR metrics: pipeline rehabilitation (AA09a), occurrence of pipeline failures (AA10a), actual water losses (AA12a) and energy efficiency in pumping stations (AA13a).

The choice of these metrics has been driven by the available data collected by ERSAR and supported by previous studies, as those metrics are considered critical to monitor the performance of assets in water systems. The rate of pipeline rehabilitation and failures in water mains in Portugal and the importance of monitoring those indicators is discussed by Marques and Monteiro [82], Ferreira and Carriço [83], Cabral et al. [84] and Santos et al. [85]. The use of water losses as one key indicator for the sector is detailed by Marques and Monteiro [82, 86] and Machado et al. [87]. Moreover, Loureiro et al. [88] studied the energy efficiencies in water systems and concluded that inefficiencies are more related to the conditions of infrastructure and network layouts than to pumping issues.

All those metrics are included in the set of the aforementioned 14 main indicators required by ERSAR's system. Pipeline rehabilitation (AA09a) and occurrence of pipeline failures (AA10a) are included in the dimension *Infrastructure Sustainability*, and are directly related to assets' performance. The other two metrics, actual water losses (AA12a) and energy efficiency in pumping stations (AA13a), are included in the dimension *Efficiency in the utilisation of environmental resources*, but they reflect the impact of assets' performance on the use of the available resources.

According to the Technical Guide 22 issued by ERSAR and LNEC [78], the pipeline rehabilitation metric (AA09a) is defined as the annual average percentage of supply and distribution pipelines older than ten years that were rehabilitated in the last five years. This metric is designed to assess the level of sustainability of service management, reflecting a continuous practice of pipeline repair to ensure their progressive renewal and appropriate average age of the network. The occurrence of pipeline failure (AA10a) is calculated as the number of pipeline faults per hundred kilometres. The actual water losses (AA12a) is the average daily volume of losses per unit of pipeline length in a year, expressed in cubic meters per pipeline kilometres in a day (m^3/km day). This metric reflects the level of sustainability in the water supply service when utilising water as an environmental resource. Berg and Marques [22] explain that the water-loss variable can be used as a proxy for inadequate maintenance costs, and recommend that it is modelled as an undesirable output. Finally, the energy efficiency in water pumping stations (AA13a) is defined as the normalised average energy usage for water pumping, indicating the sustainability of the assets in terms of using energy. It is expressed in kilowatt-hours by cubic meters per hundred meters of elevation.

Three of the metrics employed to compose RISI are undesirable, meaning that lower values are expected to denote better performance: AA10a, AA12a and AA13a. Only the metric AA09a that measures the pipeline rehabilitation is desirable, meaning that higher values indicate that the performance is better.

4.1.2. The Asset Management Maturity Index - AMMI

The second CI designed from ERSAR metrics expresses the focus of the companies in managerial practices regarding their physical assets. ERSAR highlights the importance of those aspects and requests water operators information about the knowledge of the their assets (*Infrastructure knowledge index — PAA31a*) and the features of the management systems they have implemented (*Infrastructure asset management index — PAA32a*). Those two facets of the companies' managerial practices represent crucial aspects of water systems' management. They used to be expressed by only one metric in the earlier versions of ERSAR's indicator system. However, ERSAR decided to specify these two indicators to obtain more detailed information, such as data about non-buried assets and a greater focus on data records in geographical information systems rather than on paper [89]. We propose to integrate the two indicators in the form of the *Asset Management Maturity Index (AMMI)*.

The *Infrastructure knowledge index (PAA31a)* aims to assess the company's knowledge about the infrastructure of the water supply service in its area of intervention [78]. The accuracy of asset information is crucial for successful asset management, and it depends on the quality of the data stored and the way the information is managed. The selection and specification of the data to be collected, and the quality of the strategic information systems where the information is stored and made available to users, are essential aspects of asset information management. Furthermore, the effectiveness of linking the various information systems is also important, so that data from different information systems can be cross-referenced.

It is essential to evaluate the data quality regarding its accuracy, the scale used, consistency and reliability as well as ensuring a proper georeferencing of data to manage infrastrutures. In addition, data storage must be reliable, and the flow of information must be ensured at all stages of the data system processes, including acquisition, evaluation, recording, updating, archiving and use. All those aspects are reflected in the *Infrastructure knowledge index* [89].

The *Infrastructure knowledge index* is calculated by adding the scores taken from the company's answers to a questionnaire. The total score results from the sum of the question scores and may vary between 0 and 200. The questionnaire is divided into classes covering different topics, as follows:

(a) class A — Existence of infrastructure engineering drawings and layout,
(b) class B — Information recorded on pipelines and connection branches,
(c) class C — Information recorded on other infrastructure,
(d) class D — Information recorded on measuring equipment,
(e) class E — Information recorded on the state of conservation of infrastructures,
(f) class F — Information recorded on interventions in the public network,
(g) class G — Interconnection between the Geographic Information and other company's information systems and recording of risk factors.

According to ERSAR and LNEC [78], the *Infrastructure asset management index (PAA32a)* is also determined by adding the score attributed to a set of questions related to the assessment of the company's asset management system concerning:

(a) general asset management framework,
(b) documentation and communication,
(c) strategic planning
(d) tactical planning
(e) operational planning

This index may vary between 0 and 200. ERSAR takes advantage of the existence of international asset management reference standards, ISO 55000 and ISO 55001 [6,7], and includes many of the principles and requirements present in the standard into the organisational aspects indicated in the *Infrastructure asset management index* [89]. Therefore, following ISO 55001, the companies are encouraged to deal with relevant internal and external features, major stakeholders, appropriate planning, leadership and commitment, responsibility and authority definitions, proper procedures and documentation, process controls, continuous improvement actions and other managerial aspects.

Both the *Infrastructure knowledge index — PAA31a* and the *Infrastructure asset management index — PAA32a* are only provided in their aggregate form, with no information about the partial scores that give rise to them. If detailed information about these partial scores were available, the composite indicator might be constructed including the specific scores of each question.

4.2. Deterministic approach for CI calculation

In this subsection, we describe the first approach applied to the calculation of the composite indicators which is the standard deterministic CI. The CI is computed from BoD linear programming models. BoD models are DEA models that handle multiple outputs, corresponding to several metrics to be aggregated, and a dummy input with a unitary value for all DMUs. The outputs, in this case, are the selected metrics collected from ERSAR. We employed the BoD model based on a Directional Distance Function (DDF), as formulated by Zanella et al. [90]. This model can deal with desirable and undesirable outputs, without needing to adjust the scales of measurement. The weights formulation of the Directional Distance Function BoD CI model is presented in (1).

$$\begin{aligned}
\text{minimise} \quad & \beta_{j_0} = -\sum_{r=1}^{s} y_{rj_0} u_r + \sum_{k=1}^{l} b_{kj_0} p_k + v \\
\text{subject to} \quad & \sum_{r=1}^{s} g_y u_r + \sum_{k=1}^{l} g_b p_k = 1 \\
& -\sum_{r=1}^{s} y_{rj} u_r + \sum_{k=1}^{l} b_{kj} p_k + v \geq 0 \qquad j = 1, \ldots, n \\
& u_r \geq 0, \qquad r = 1, \ldots, s \\
& p_k \geq 0, \qquad k = 1, \ldots, l \\
& v \in \mathbb{R}
\end{aligned} \tag{1}$$

In formulation (1), y_{rj} and b_{kj} are, respectively, the desirable and undesirable indicators for DMUs j ($j = 1, \ldots, n$) and the values of y_{rj_0} and b_{kj_0} represent the indicators of the DMU j_0 under assessment. The index r stands for the set of desirable outputs ($r = 1, \ldots, s$) and the index k stands for the set of undesirable outputs ($k = 1, \ldots, l$). The model's decision variables are the weights, where v is associated with the dummy input, u_r is associated to the desirable outcomes r, and p_k with the undesirable outcomes k. The total number of DMUs is n, the total number of desirable outputs is s and the total number of undesirable outputs is l.

The directional distance vector is specified as $(g_y, -g_b)$, indicating the direction of expansion of desired outputs and contraction of undesired ones. The decision about the direction vector used in the models is critical since it can influence the computed scores. Several solutions have been presented in the literature depending on the study's objective. Fried et al. [91] address different options for applying direction vectors to guide the improvement of inputs and outputs in DEA models. Those authors discuss suggestions for the vectors' selection and advocate that this decision should be made according to the research purpose. Rogge et al. [92] also explores alternatives for the vector to set the directions of improvement for desirable and undesirable outputs in BoD models. In this work, following Zanella et al. [90] and Rogge et al. [92] we choose the values of $(g_y, -g_b)$ as being equal to $(y_{rj_0}, -b_{kj_0})$. In this case, each DMU can improve by following the path indicated by its specific output metrics, allowing for a proportional interpretation of the resulting composite indicator value.

The factor β_{j_0} in (1) expresses the inefficiency level of DMU j_0, representing the maximum expansion of desirable outputs and contraction of undesirable outputs that is feasible to satisfy the model's restrictions. The minimum feasible level of β_{j_0} is determined by optimisation, such that the DMU j_0 under assessment can select the weights that show it in the best possible light. The value of CI associated with j_0, can be obtained as $1/(1+\beta_{j_0})$. Consequently, the CI score ranges from 0 to 1, where 1 represents the best performance level. The deterministic CI is referred in this work as CI_{j_0}. If $CI_{j_0} < 1$, there is a linear combination of other DMUs that dominates in terms of overall performance. If $CI_{j_0} = 1$, the DMU j_0 is located in the best-practice frontier, meaning that it is not outperformed by any of the others DMUs included in the assessment.

Weight restrictions must also be included in the model to prevent assessments that could disregard certain indicators by assigning them weights equal to zero. A more detailed discussion about the several kinds of weight restrictions for DEA models is available in Wong and Beasley [93], Allen et al. [94], Sarrico and Dyson [95], among others. Zanella et al. [90] proposes a formulation for AR-I restrictions in BoD models, using virtual weights restricted in terms of the proportional importance of the variables. These restrictions consider a hypothetical DMU whose outputs are equal to the average of all values observed in the DMUs in the sample, represented by $(\bar{y}_r, \bar{b}_k)$. The virtual weights of the "average DMU" are then constrained by percentage-based restrictions. The use of those AR-I restrictions presents the advantage of being identical for all DMUs, and according to Zanella et al. [90], they represent the best choice to construct composite indicators and ranks. The AR-I restrictions are the most used weight restrictions in BoD models. In this study, only lower bounds expressed as percentages are used (ϕ_r and ϕ_k, respectively for desirable and undesirable indicators). Following Zanella et al. [90], the weight restrictions are added to the BoD model and formulated as shown in (2). By avoiding zero weights, all indicators are given some degree of importance when computing the composite indicators.

AR-I weight restrictions

$$\begin{aligned}
& \frac{u_r \bar{y}_r}{\sum_{r=1}^{s} u_r \bar{y}_r + \sum_{k=1}^{l} p_k \bar{b}_k} \geq \phi_r, \quad r = 1, \ldots, s \\
& \frac{p_k \bar{b}_k}{\sum_{r=1}^{s} u_r \bar{y}_r + \sum_{k=1}^{l} p_k \bar{b}_k} \geq \phi_k, \quad k = 1, \ldots, l
\end{aligned} \tag{2}$$

A detailed explanation of the BoD model formulation and the use of weight restrictions AR-I is available in Zanella et al. [90],D'Inverno and De Witte [96],Van Puyenbroeck et al. [97].

In case there are no undesirable indicators among the components to be aggregated in the CI, the BoD model and the AR-I weight restrictions can be simplified as shown in (3).

$$\begin{aligned}
&\text{minimise} && \beta_{j_0} = -\sum_{r=1}^{s} y_{rj_0} u_r + v \\
&\text{subject to} && \sum_{r=1}^{s} g_y u_r = 1 \\
& && -\sum_{r=1}^{s} y_{rj_0} u_r + v \geq 0 \qquad j = 1, \ldots, n \\
& && u_r \geq 0, \qquad r = 1, \ldots, s \\
& && v \in \mathbb{R} \\
& && \text{AR-I weight restrictions} \\
& && \frac{u_r \bar{y}_r}{\sum_{r=1}^{s} u_r \bar{y}_r} \geq \phi_r, \quad r = 1, \ldots, s
\end{aligned} \tag{3}$$

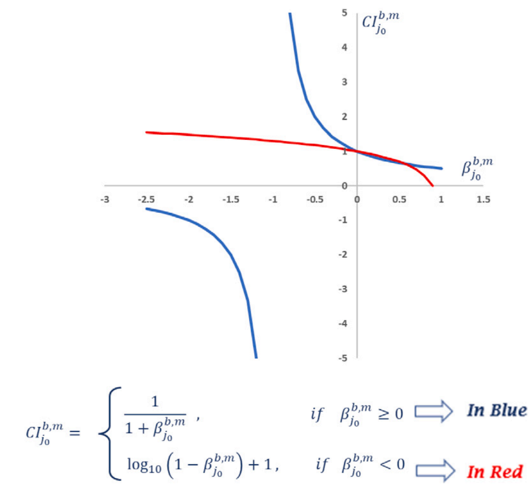


Fig. 1. Comparison between the expressions used to calculate CI.

4.3. Robust and conditional approaches for CI calculation

This subsection describes the generation of CIs following the robust and conditional approaches.

Some limitations on the use of the deterministic CI have been discussed in the literature, namely its great sensitivity to outliers in the sample and the difficulty in performing statistical inference. These limitations can be overcome by the use of the robust CI approach. The conditional approach allows accounting for the effect of exogenous contextual variables in a single stage when constructing CIs. Since its initial conceptualisation by Cazals et al. [98] and Daraio and Simar [57,58], these techniques have been applied, revised and enhanced by an extensive number of studies: De Witte and Kortelainen [59], Rogge et al. [92], De Witte and Schiltz [99], Lavigne et al. [100], D'Inverno and De Witte [96], Fusco et al. [101] and Mergoni et al. [60], among others.

In line with this stream of the literature, the computation of the robust CI is performed by drawing (for a very large number of times) at random with replacement units from the original set of DMUs and computing the CI estimates for each sample through the resolution of the BoD model. If a sample of size m is considered, the resulting CI will reflect the comparison with the best-practice frontier composed only by DMUs included in the sample of size m. If this sampling and calculation process is performed B times, where B is typically a high number, the effect of the outliers on the average efficiencies will be lessened since they will not appear in all the collected samples. The resulting robust CI, referred as $CI_{j_0}^{m}$ in this study, is the average of the CIs generated for all B samples of size m, as shown in (4), where $CI_{j_0}^{b,m}$ is the CI of DMU j_0 calculated using sample b.

$$CI_{j_0}^{m} = \frac{1}{B}\sum_{b=1}^{B} CI_{j_0}^{b,m} \tag{4}$$

When the results are calculated, it may happen that, for a given sample, the DMU under assessment (j_0) is not included in that sample, such that it may be more efficient than all the DMUs in the sample. In this case, the DMUs would be classified as "super-performing" and its score, $\beta_{j_0}^{b,m}$, would have a negative value. The more negative $\beta_{j_0}^{b,m}$ is, the higher the performance of the DMU, so $CI_{j_0}^{b,m}$ should increase as $\beta_{j_0}^{b,m}$ decreases. However, this cannot happen if $CI_{j_0}^{b,m}$ is calculated as $1/(1+\beta_{j_0})$. A solution to this problem is suggested by Mergoni et al. [60] by modifying the calculation of $CI_{j_0}^{b,m}$ to adapt for the case of negative values of $\beta_{j_0}^{b,m}$ as detailed in (5).

$$CI_{j_0}^{b,m} = \begin{cases} \frac{1}{1+\beta_{j_0}^{b,m}}, & \text{if} \quad \beta_{j_0}^{b,m} \geq 0; \\ \log_{10}(1-\beta_{j_0}^{b,m})+1, & \text{if} \quad \beta_{j_0}^{b,m} < 0 \end{cases} \tag{5}$$

Fig. 1 displays both functions applied to calculate $CI_{j_0}^{b,m}$. The original formulation (in blue), besides of being discontinued in $\beta_{j_0}^{b,m} = -1$, yields negative values for $CI_{j_0}^{b,m}$ if $\beta_{j_0}^{b,m} < -1$. This curve does not reflect the proportional increments in performance expected for the CI when the units are "super-performing". On the other hand, the proposed formulation when $\beta_{j_0}^{b,m} < 0$ (in red) follows a similar trend as the original formulation for positive values of $\beta_{j_0}^{b,m}$ allowing the value of $CI_{j_0}^{b,m}$ to increase as the performance of the DMUs improves. Therefore, following Mergoni et al. [60], the robust CIs proposed in this study are computed using the expressions in (5). This applies also for the computation of the robust conditional CIs as presented hereinafter.

In order to account for the influence of the contextual variables, the robust conditional approach or, simply conditional approach, needs to be employed. This strategy is used to adjust the CIs to allow fairer comparisons by forcing the DMUs assessment to be performed with more similar DMUs according to exogenous characteristics. The procedure is analogous to the robust CI strategy, using B samples of size m and computing the CI as the average of all samples. The difference between this strategy and the robust approach is that instead of performing random sampling from a uniform distribution, the sampling is conducted using a similarity function. The similarity is measured using a kernel function estimated using the contextual variables. There are currently computing models to deal with both continuous and categorical context or exogenous variables [102]. The conditional CI, referred to as $CI_{j_0}^{m,z}$, is computed as the average of the conditional $CI_{j_0}^{b,m,z}$ for B samples as shown in (6).

$$CI_{j_0}^{m,z} = \frac{1}{B}\sum_{b=1}^{B} CI_{j_0}^{b,m,z} \tag{6}$$

After computing $CI_{j_0}^{m,z}$, the significance and the direction of influence of the contextual variables can be evaluated. The score ratio between the robust CI and the robust conditional CI ($CI_{j_0}^{m}/CI_{j_0}^{m,z}$) is non-parametrically regressed against the exogenous variables. Partial plots showing the variables' confidence intervals for different levels of the exogenous variables can be generated, and non-overlap intervals indicate that the effect of the context is significant.

If the ratio is decreasing as the environmental variable increases (that is, the regression plot displays a negative slope), it means that the conditional score is larger than the unconditional one just because compared among units more similar in terms of the contextual variables. In that sense, the environment plays an unfavourable role when it comes to the performance evaluation. On the contrary, if the regression plot displays a positive slope, the environment plays a favourable role. For a more detailed explanation, see Rogge et al. [92] and D'Inverno et al. [34]. For an analogous approach using DEA efficiency scores, see Walker et al. [26].

5. Case study

This section details the case study analysed by the research. In subsection 5.1, the primary data set containing the metrics collected by ERSAR that form the composite indicators is presented. Subsection 5.2 displays the variables utilised to characterise the environment in which the companies operate.

5.1. Data collected

The study employs the metrics collected by ERSAR during five years, from 2016 to 2020, to construct the CIs RISI (Resource and Infrastructure Sustainability Index) and AMMI (Asset Management Maturity Index) as detailed in Section 4.1[1].

The methods developed for comparative performance, such as those based on DEA, provide better results if the number of DMUs is large. Thanassoulis [24] explains that one way to increase the number of DMUs is to treat each unit as a separate comparative entity in distinct units of time, through the use of a panel data. By doing that, the basic assumption to consider is that the technology remains stable over time to enable meaningful comparisons of performance. Given the observed time span and the nature of the water industry, this assumption is verified. The infrastructure cannot be changed rapidly as the investments in assets are primarily underground and deemed to last several decades. In that sense, the DMUs in this study are formed by the combination of company and year. For example, DMU A1-2016 means that the data of company *Águas de Santo André (A1)* for 2016 is being assessed. By choosing this strategy, the companies can be compared not only with other companies but also with themselves in different years, allowing the evaluation of their performance over time. Since *Águas do Douro e Paiva (A3)* was created in 2017, only four years of data are available for this company. Therefore the number of DMUs employed in the study is 49 instead the expected number of 50, considering that there are ten companies for five years of evaluation.

An examination in the data set indicates that two data instances are missing: the values for the metric AA13a for DMUs A5-2016 and A8-2020. The procedure recommended by Kuosmanen et al. [103], Morais and Camanho [47] and Henriques et al. [19] for treatment of missing data in DEA was used in this case. Since the metric is undesirable, a large value corresponding to the maximum value of metric AA13a in the sample was assigned to both DMUs. This implies that the absence of data cannot favour the DMU in the performance assessment.

The descriptive statistics for the data related to the metrics that compose both CIs are presented in Tables 1 and 2.

The correlation among the various metrics employed to build the CIs was investigated. The estimated Pearson correlation coefficients do not reflect a significant association between the pairs of metrics used in each CI, as the resulting absolute values of the coefficients are not close to one, as shown in Table 3. In this scenario, the low correlation supports incorporating all variables into the models.

[1] ERSAR reports are available online in https://www.ersar.pt/pt/site-publicacoes/Paginas/edicoes-anuais-do-RASARP.aspx.

5.2. Exogenous contextual variables

Four characteristics covering various contexts in which the organisations operate were chosen to analyse the influence of the background on their performance. The four factors are expressed also by variables collected and publicised by ERSAR on the annual report. Two of those variables, the management system and the typology of intervention area are categorical, and the other two, the volume of activity and the pipeline network length are continuous.

Variable PAA02a identifies the company's management system, and reflects the market structure of the water sector in Portugal. The companies *Águas Públicas do Alentejo (A8)*, *EPAL (A9)* and *ICOVI (A10)*, are operated by delegation, whereas all the other wholesale companies are operated by concession. This status remains for the whole period from 2016 to 2020.

Variable PAA14a reflects the typology of intervention area, in which the companies are classified as operating in rural, urban, or semi-urban settings. This criterion is mostly determined by population density. The urban companies are *Águas do Douro e Paiva (A3)* and *EPAL (A9)*, the rural companies are *Águas de Santo André (A1)*, *Águas Públicas do Alentejo (A8)* and *ICOVI (A10)*. The remaining five companies operate in semi-urban environment. The companies' status also does not change during the assessment period.

Table 4 presents the statistics for categorical exogenous variables between 2016 and 2020.

The two continuous exogenous factors are represented by variables PAA50a and dAA15a. Variable PAA50a indicates a company's volume of activity, meaning the total billed volume of water supplied by the company per year. Variable dAA15a expresses the pipeline network length of the company in kilometres.

According to Haider et al. [104], water supply systems include vertical components and linear components. Examples of vertical components are treatment plants, pumping stations and storage tanks, and the linear components include the water mains and pipeline networks. The linear components are usually much more expensive representing from 60% to 80% of the total cost of the water system. Therefore, the variable dAA15a was chosen to reflect the amount of assets that the company manages. Both continuous exogenous variables included are proxies of the company's size, but they are not strongly correlated with each other. The Pearson correlation coefficient is 0.377 and the p-value is 0.008.

Table 5 displays the descriptive statistics of the continuous exogenous variables. The effects of problematic and small samples have been already discussed by Henriques et al. [52]. Following those authors, and in order to maintain consistency in the study, we choose to include the two continuous variables as discrete variables (a similar approach can be found also in D'Inverno et al. 34). Therefore those variables are split in two classes: above and below the median.

6. Results and discussion

In this section, the study's results are presented and discussed in three stages. The first part discusses the results of the calculation of the deterministic CIs and robust non-conditional (or simply robust) CIs. The second stage presents the estimation of the robust conditional (or simply conditional) CIs and the findings regarding the effect of contextual conditions on asset management performance. The last part presents a visualisation tool conceived to enable the combined analysis of both AMMI and RISI.

6.1. Deterministic and robust composite indicators' results

This subsection presents the findings from the calculation of the deterministic and robust CIs.

The deterministic CIs calculation follows the procedure detailed in Section 4.2. Since RISI presents undesirable outputs, they are computed

Table 1
Metrics that compose RISI.

ERSAR Code	Metric description	Metric definition	No. Obs.	Mean	St. Dev.	Min.	Max.
AA09a	Pipeline rehabilitation (%/year)	Average annual percentage of pipelines with life higher than ten years rehabilitated in the last five years.	49	0.19	0.31	0	1.3
AA10a	Occurrence of pipeline failure (n^o/100 km year)	Number of failures in pipelines per 100 km in a year.	49	7.92	8.78	1	40
AA12a	Actual water losses (m^3/km day)	Actual water losses due to leakages and overflows per unit of pipeline length.	49	6.46	8.13	0.1	31.4
AA13a	Energy efficiency in pumping stations (kWh/m^3.100 m)	Average normalised energy consumption of pumping stations.	49	0.47	0.12	0.36	0.73

Table 2
Metrics that compose AMMI.

ERSAR Code	Metric description	Metric definition	No. Obs.	Mean	St. Dev.	Min.	Max.
PAA31a	Infrastructure knowledge Index (Score 0–200)	Score of evaluation of the knowledge of the several infrastructures in different classes ranging from 0 to 200.	49	170.37	20.08	111	197
PAA32a	Infrastructure asset management Index (Score 0–200)	Score of evaluation in a questionnaire about asset management practices ranging from 0 to 200.	49	109.06	83.9	0	200

Table 3
Pearson correlation coefficients — RISI and AMMI metrics.

CI	Pair of metrics	Pearson Correl. Coefficient
RISI	AA09a–AA10a	0.334
	AA09a–AA12a	0.047
	AA09a–AA13a	0.362
	AA10a–AA12a	−0.210
	AA10a–AA13a	0.253
	AA12a–AA13a	−0.284
AMMI	PAA31a–PAA32a	0.357

through the resolution of BoD model (1). For the AMMI calculation, BoD model (3) is employed, where only desirable outputs are considered. For the weight restrictions shown in (2) and (3), the values of parameters ϕ_r and ϕ_k were set to be equal to 0.05. Different values of ϕ_r and ϕ_k from 0.02 to 0.10 were tested, and the results remained very stable. After running this sensitivity analysis, an intermediate value of 0.05 was chosen.

For the robust CIs calculation, a sensitivity analysis was performed to decide the value of bootstrapping sample size m. Daraio and Simar [58] explain that there are no fixed rules neither automatic procedures to select the value of m. This value is typically an integer number smaller than n. These authors recommend to perform a sensitivity analysis, choosing several levels of m and evaluating the number of super-performing units. This number should decrease as m increases. In small dimension samples, the choice of m as being equal to the number of DMUs is recommended by Henriques et al. [52]. Following those authors, we choose for both indicators to use $m = n = 49$. Furthermore, at this level there is already a substantial decrease on the number of super-performing units and on the average CI values. The number of bootstrapping replications is chosen as $B = 2000$. The results are obtained using packages *Rglpk* [105] and *lpSolve* [106] in R programme.

Table 6 shows the descriptive statistics for RISI results and Table 7 displays the similar information for AMMI results. Looking at the averages of both indicators, a significant room for improvement can be noticed. Note that lower average scores signal a larger degree of heterogeneity among firms, taking the best-observed practices of the sample in a five-year period as reference. The CIs for all DMUs are reported in Table A.1 in Appendix.

A close look at the results for AMMI reveals that half of the DMUs present a CI above 0.866 in the deterministic case and above 0.873 in the robust conditional analysis. Overall, the CIs allows the identification of the poorly performing companies and the highly performing ones, to guide improvements of the former by looking at the good practices of the latter.

Both indicators suggest that there is potential room for improvement among the companies. As the BoD model assigns the weights to each metric in the most favourable way, the underperforming companies cannot complain about the fairness of the evaluation. The highest ratings given to the top performers may not always indicate that there is no potential for further improvement in absolute terms. It simply means that, based on the data available, these companies represent the best observed performance in the period under consideration. The evolution of productivity levels over time is captured by the movement of the best-practice frontier, whereas cross-sectional assessments of efficiency only evaluate the distance to the frontier at a given moment in time. This benchmarking exercise can be beneficial to the determination of policies both for the regulatory entity and the companies themselves.

6.2. Effect of exogenous contextual variables

This subsection presents the calculation of robust conditional CIs, which reveals companies' performance taking into account the operating context. Besides the use of R packages *Rglpk* [105] and *lpSolve* [106], in this analysis the *np* package was employed in R software to perform sampling according to the similarity level of DMUs and also to execute the non parametric tests of significance [107]. The *np* package focuses on kernel approaches that are suitable for the combination of continuous and categorical data.

As previously discussed, the continuous variables need to be converted to discrete ones in this analysis due to the small sample size. For the same reason, the effect of exogenous factors cannot be addressed using one model combining all variables as discussed in [52]. The computation of the conditional CIs and significance tests has to be performed individually for each exogenous variable in a first stage. In a second phase, only the significant variables are included in the final calculation to generate the conditional CIs. The potential for omitted variable bias in this case may be a concern, yet this method was deemed

Table 4
Categorical exogenous variables.

ERSAR code	Description	Definition	Obs.	Number of companies and percentage per category
PAA02a	Management system	Concession or Delegation.	49	Concession - 7 (69.4%) Delegation- 3 (30.6%)
PAA14a	Typology of intervention area	Rural areas, semi-urban areas or urban areas.	49	Rural - 3 (30.6%) Semi-urban - 5 (51.0%) Urban - 2 (18.4%)

Table 5
Continuous exogenous variables.

ERSAR code	Description	Definition	Obs.	Mean	St. Dev.	Median	Min.	Max.
PAA50a	Volume of activity	Volume of water supplied (m^3/year)	49	60,680,458	62,916,421	30,448,818	978,630	215,392,064
dAA15a	Pipeline network length	Total length of pipelines (km)	49	1009.0	1148.4	497	26.8	3578.8

Table 6
Descriptive statistics for RISI results in deterministic and robust unconditional approaches.

	Average	St. Dev.	Min	Q1	Median	Q3	Max
Deterministic RISI CI (CI_{j_0})	0.799	0.079	0.637	0.771	0.801	0.840	1.000
Robust unconditional RISI CI ($CI^m_{j_0}$)	0.835	0.092	0.646	0.796	0.834	0.877	1.138

Table 7
Descriptive statistics for AMMI results in deterministic and robust unconditional approaches.

	Average	St. Dev.	Min	Q1	Median	Q3	Max
Deterministic AMMI CI (CI_{j_0})	0.861	0.111	0.602	0.762	0.866	0.979	1.000
Robust unconditional AMMI CI ($CI^m_{j_0}$)	0.865	0.110	0.606	0.768	0.873	0.979	1.004

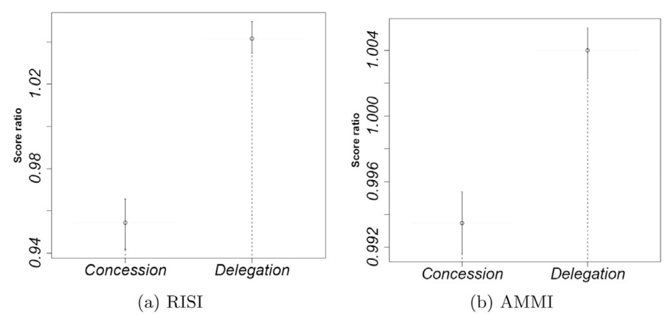


Fig. 2. Effect of each exogenous variables — Management system.

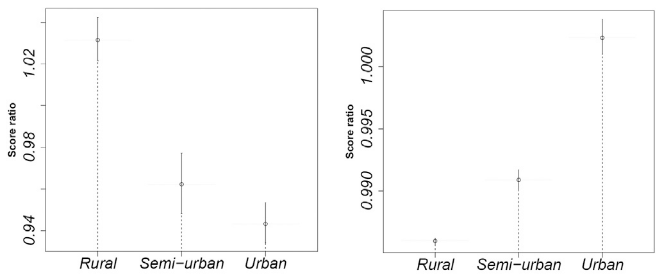
valid as the aim of the analysis was to pursue evidence of correlation, not necessarily causal relationships. The option to include one variable at a time furthers a practical approach that enables the identification of the contextual factors influencing the outcomes.

Fig. 2 reports the results obtained for the Conditional BoD model considering the variable PAA02a (Management System). The confidence intervals shown in Fig. 2 do not overlap, and the p-value of the hypothesis test used to compare the groups (Kernel regression significance test) is smaller than 2.22×10^{-16} in both cases, which means that the difference between concession and delegation management systems is significant regarding the performance measured by both indicators. The score ratio between robust and robust conditional CIs for the delegation management system in both indicators is higher, indicating that the delegation environment is more favourable for the performance in both perspectives of RISI and AMMI. We hypothesise that this fact is concerned with the more experience the delegation companies have already got in asset management practices. Historically, the emphasis on asset management began in Portugal with delegation-managed firms. Since 2006, EPAL, one of the largest delegation firms, has been the first wholesale water provider to focus on asset management procedures. It was also the first wholesale company in the country to acquire ISO 55001 certification [8].

The ownership and management approaches of water systems have been extensively discussed in the literature, with controversial results. In the review conducted by Berg and Marques [22], out of 47 studies focusing on that issue, 18 found that private water companies perform more efficiently than the public ones, 12 concluded that public water utilities are more efficient than the privates, and 17 reach inconclusive results. In general, the private sector tends to improve labour productivity but often increases capital expenses, and the opposite holds for the public sector. We highlight that those studies do not emphasise only asset management practices, but efficiency in general. Furthermore, in the management systems for the bulk companies in Portugal, the public control is more direct in the delegation system than in the concession.

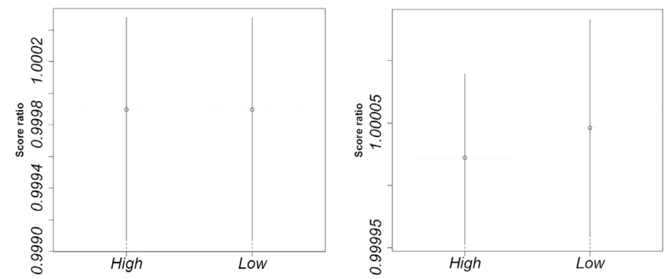
The results for the conditional CI approach using variable PAA14a (Typology of Intervention Area) can be seen in Fig. 3. The results for the kernel regression significance test in this case indicate that this contextual factor is also significant for the companies' performance in asset management. The p-values are less than 2.22×10^{-16}, both for RISI and AMMI.

In the case of RISI, rural environment is more favourable, achieving higher values in the score ratios between the robust and conditional CIs. A possible reason for that may be related to the fact that the rural water networks are younger in Portugal, mainly due to the expansion in investments towards the rural areas in recent decades. Younger



(a) RISI (b) AMMI

Fig. 3. Effect of each exogenous variables — Typology of intervention area.



(a) RISI (b) AMMI

Fig. 4. Effect of each exogenous variables — Volume of activity.

water assets have reduced chances of deterioration and leakage, which may explain why rural settings operate more efficiently. Rurality has been already studied as an exogenous factor in the context of global efficiency by Walker et al. [26], that concluded that higher population densities in urban setting are more favourable to increase the efficiency due to scale economies.

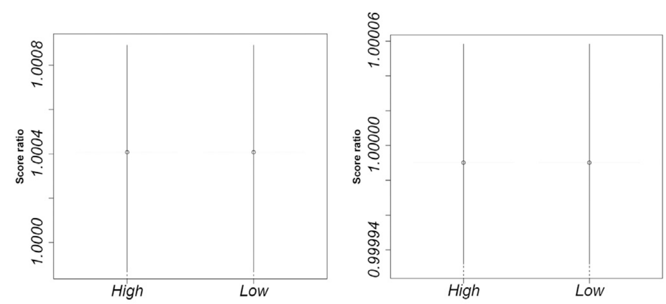
The urban environment score ratio, on the other hand, is much higher than the other settings for the AMMI, suggesting that urban enterprises have superior asset management systems. This phenomenon might be connected to urban companies' knowledge of their assets. Since the *Infrastructure knowledge index*, required by ERSAR, is a component of AMMI, the information the companies retain about their assets affects the result of that indicator. In urban settings, the assets' inventories and records are more accurate, which may explain this finding.

The results obtained from the analysis of the conditional CIs employing the variable PAA50a (Volume of activity) are displayed in Fig. 4. As previously stated, the level "High" in the graph includes the DMUs that present the volume of activity higher than the median of all DMUs, while the DMUs labelled "Low" have a lower volume of activity than the median. In this case, the differences between the groups were found to be non-significant for both indicators. The p-values are 0.399 for RISI and 0.231 for AMMI, revealing that the volume of activity expressed by the amount of water supplied by the wholesale companies does not affect their performance in asset management. Water systems are considered large by the European Commission, if they supply more than 1000 m^3 of water per day or serve more than 5000 people [108]. Looking at the data in Table 5, we can see that all the companies included in this analysis provide a larger volume of water than 1000 m^3/day, and at this scale no difference can be noticed among the analysed companies in asset management regarding the volume of activity.

Similar results are seen for variable dAA15a (Pipeline network length). The graphs in Fig. 5 suggest that there are no significant differences regarding the two levels of pipeline network length considered in the analysis. As previously noted, the two categories are "Low" for less than the median of all DMUs and "High" for larger than the median.

After the effect of all variables is evaluated individually, the robust conditional CIs are computed utilising the two factors considered significant: *Management system* and *Typology of intervention area*. The resulting CIs are also shown in Table A.1 in Appendix and the descriptive statistics are displayed in Table 8.

In the conditional assessment, the companies are predominantly compared to more similar units. Also in this case, it is possible to identify potential room for improvement. Once more, the companies found poorly performing are granted the fairness of the assessment and cannot blame the evaluation system.



(a) RISI (b) AMMI

Fig. 5. Effect of each exogenous variables — Pipeline length.

Table 8
Descriptive statistics for RISI and AMMI in robust conditional approach.

	Average	St. Dev.	Min	Q1	Median	Q3	Max
Robust conditional RISI CI ($CI_{j_0}^{m,z}$)	0.891	0.088	0.667	0.829	0.913	0.953	1.057
Robust conditional AMMI CI ($CI_{j_0}^{m,z}$)	0.906	0.084	0.615	0.865	0.898	0.998	1.000

6.3. Visualisation framework for the combined analysis of asset management dimensions

A visualisation model inspired by the BCG (Boston Consulting Group) matrix [109] was created to enable for the combined analysis of companies in both indicators (RISI and AMMI) in an integrated manner. In this framework, the companies under assessment are classified according to the value of the CIs compared to the median of the entire sample. Fig. 6 distinguishes four categories to illustrate the companies' performance compared to peers, as follows:

(a) Stars — when both RISI and AMMI are higher than their median values. In this case, the companies provide tangible results and demonstrate consistent asset management techniques, compared to peers.
(b) Soldiers — when RISI is higher than median and AMMI is lower or equal than median. In this category, the companies take good care of the assets, meaning that the assets are maintained in suitable operational conditions compared to peers, but management strategies are not properly implemented.

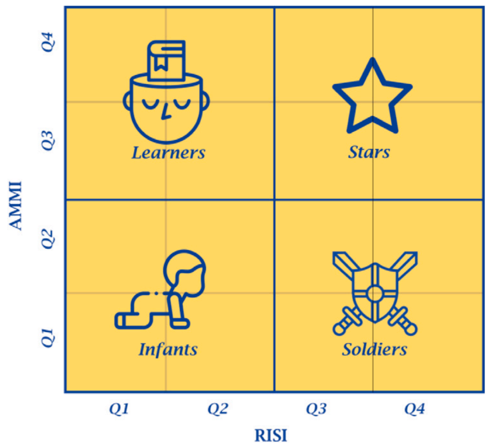


Fig. 6. Visualisation model — RISI and AMMI.

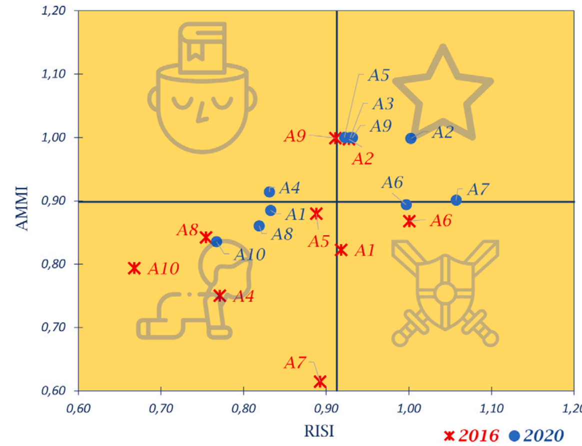


Fig. 7. Visualisation model — Comparison between the first and the last year of assessment.

(c) Infants — In this class, both RISI and AMMI are lower or equal than the medians. The Infants give the first steps in the organisation for asset management and their operational performance is worse compared to peers.
(d) Learners — The Learners present AMMI higher than the median and RISI lower or equal than the median. They have been working on robust management systems but their achievements in asset management are worse than most of their peers.

The complete classification for all the companies is presented in Table A.1 in Appendix.

Fig. 7 displays the positions of the companies for the first and last years considered in the robust conditional assessment to highlight the changes in both CIs across time. When the data for 2016 and 2020 are compared, a tendency towards increasing both indicators can be noticed for most companies, suggesting an improvement in the sector's asset management practices.

Looking at the AMMI results, all the companies present better results for their management practices, between the first and the last years of evaluation. The same comparison for the RISI results reveals only two exceptions to this trend: the companies *Águas de Santo André (A1)* and *Águas do Vale do Tejo (A6)*. *Águas de Santo André (A1)* displays significantly worse results for water losses (AA12a), which raised from 0.5 m^3/km day in 2016 to a range between 1.8 and 2.6 m^3/km day in the following years. The unfavourable trend is also repeated for the energy efficiency in pumping stations (AA13a) which was 0.49 kWh/m^3.100 m in 2016 and jumped to values superior to 0.62 kWh/m^3.100 m from 2017 . This company is also disfavoured by the lack of investment in network rehabilitation, as metric AA09a is null for the whole period, and by the significant number of pipeline failures (AA10a), which are higher than the sector average for all years. As a result of this poor performance, Águas de Santo André dropped from the category Soldier to Infant in 2017, and remained in the same category since then. The case of *Águas do Vale do Tejo (A6)* is different since the worsening in RISI is minimal. This company presents relatively stable results over time, but when the metrics between 2016 and 2020 are compared, the number of failures in pipelines (AA10a) increased from 6 to 7 failures per 100 km.year.

In all the other cases, improvements are noticed in both indicators. This information may be utilised as a motivator to continue with asset management practices in the future.

The use of combinations of company and year as units of assessment allow for the comparison of a firm's performance with itself across time. This procedure is known as internal benchmarking (see also [110], for further details on this topic). The visualisation framework may be used to depict the progression of the companies' performance throughout the period under assessment. Fig. 8 displays examples for three companies.

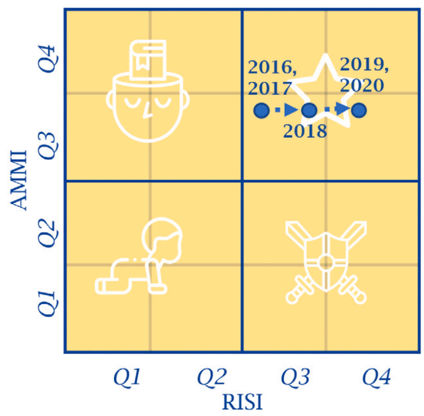
The progression of company *Águas do Algarve (A2)* is presented in Fig. 8(a). This organisation has consistently learned from implementing managerial approaches over the years. It started in 2016 already as a Star, and by keeping the performance in managerial features at a high level, it keeps improving its operational results, remaining always in the same category. *Águas do Norte (A5)* has also been improving its management system compared to peers over the years, being certified in ISO 55001 in 2019. The tangible results have also been improving as shown in Fig. 8(b), even though an unsustainable major progress from 2016 to 2017 led to a decline in 2018. Since 2019, the company performs as a Star. The results of *Águas do Vouga (A7)* depicted in Fig. 8(c) indicate that the company has also learned from the implementation of managerial approaches over the years. Its operational results have improved, and finally, in 2020, it performs as a Star.

By analysing each company's evolution individually through the 2 × 2 matrix, one can identify in which period the company adopted best practices and better understand what actions are required to support improvements. The fact that performance can be evaluated in two dimensions, using the joint visualisation of managerial elements (AMMI) and tangible results (RISI), may support the companies' overall internal analyses.

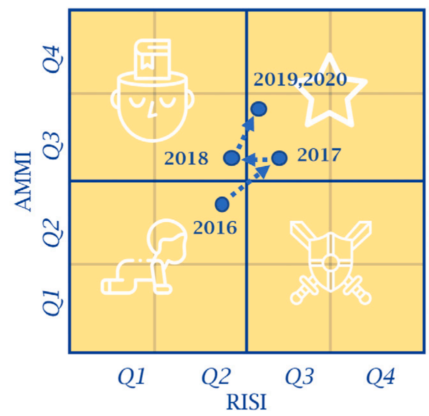
A first policy recommendation for a given company should be to analyse its evolution over time through an internal benchmarking process. If there is change between categories, or even if there is a variation in performance within the same category, the company can use the periods when its performance was superior and try to determine which factors led to that success. Next, the company should analyse the performance of its peers, especially those that are subject to the same context, and try to set targets based on the results of these peers that may help the company improve its performance.

7. Conclusion

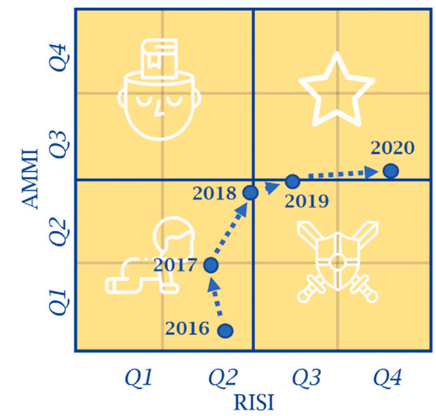
Among the main findings of this work, a novel approach to benchmark wholesale water supply companies regarding asset management practices is developed using Benefit-of-the-Doubt (BoD) directional distance models to construct composite indicators (CI). The BoD models provide an innovative way of applying the metrics collected annually by the Portuguese regulatory authority, ERSAR. This strategy benefits from reliability of ERSAR's data and well-established procedures for monitoring companies and acquiring information. This study is the first to use these data to evaluate the success of organisations in terms



(a) A2 - Águas do Algarve



(b) A5 - Águas do Norte



(c) A7 - Águas do Vouga

Fig. 8. Examples of three companies' evolution from 2016 to 2020.

of asset management methods, which fills an important gap in the literature.

In addition to the traditional deterministic strategy for generating CIs, robust and conditional approaches are used to allow statistical inference and examine the influence of contextual factors on firms' performance. The findings suggest that companies with a management system based on a delegation model show better asset management performance. Furthermore, a rural setting appears to be more favourable for achieving good operational results in assets, whereas better management systems are expected to benefit from urban environments.

The findings of this study enable water businesses to understand better where they stand in terms of asset management performance compared to other firms and themselves over time. These findings are highly relevant because they may help organisations make better decisions about where to focus on promoting continuous improvement efforts related to asset management techniques, which represent a vital issue in the water industry. Furthermore, the insights uncovered by this research may be used by the regulator to set policy targets for the water sector following the objectives of the sunshine regulation method and promoting the overall efficiency of the sector. A visualisation model for the combined evaluation of the two CIs is also presented, and examples of companies' evolution across the assessment period are discussed.

The presented study gathers in an innovative way a number of relevant aspects. First, it makes use of reliable and accurate data collected from the regulatory entity in the Portuguese water industry. Second, the BoD technique is suitable to reduce potential conflicts in the evaluation assessment, since companies cannot complain about the fairness of the aggregating scheme (being by design the most favourable one). Third, the conditional analysis favours the comparison of units under a more similar context, allowing for an even fairer analysis. Finally, the strategy combining internal and external benchmarking allows the assessment of a company over time and the visualisation model enables the combined evaluation of operational results and managerial enablers.

Some limitations of the study derive from the small sample size used. As a result, the contextual variables cannot be all included in a single model and the interaction between all variables cannot be investigated. The reduced sample size also prevents the use of continuous exogenous variables in their original form. The investigation can be extended to the 234 Portuguese retail companies in future developments. A larger sample size may allow a more detailed analysis of contextual factors. Furthermore, the envelopment formulation of BoD models may be used to determine the best peers and targets regarding asset management practices for this broader sample of companies.

The proposed tools and the overall analysis should encourage the companies and the regulator to collaborate for a richer collection of indicators associated with the assets and asset management practices. For example, in most of the cases, companies do not report relevant information such as the *Infrastructure Value Index*, the *Infrastructure Current Value* and the *Infrastructure Current Value*. A richer dataset can help the regulator and the companies better identify the best practices, enhance internal management, and design policies to foster a continuous improvement of asset management activities.

CRediT authorship contribution statement

Hermilio Vilarinho: Conceptualisation, Methodology, Software, Formal analysis, Investigation, Writing – original paper. **Giovanna D'Inverno:** Methodology, Software, validation, Formal analysis, Writing – review and editing. **Henriqueta Nóvoa:** Methodology, Validation, Formal analysis, Writing – review and editing, Supervising. **Ana S. Camanho:** Methodology, Validation, Formal analysis, Writing – review and editing, Supervising.

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 "FCT- Fundação para a Ciência e Tecnologia" through PhD research grant 2021.05244.BD and INESC TEC – Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência.

Appendix. Composite indicators (CIs) and categories for all companies in each year

See Table A.1.

Table A.1
CIs and categories for all companies in each year.

Year	Company ID	Company	Category	Deterministic CI		Robust CI		Robust Conditional CI	
				RISI	AMMI	RISI	AMMI	RISI	AMMI
2016	A1	Águas de Santo André	Soldier	0,713	0.764	0,732	0,770	0,918	0,823
	A2	Águas do Algarve	Star	0,854	0.978	0,883	0,979	0,928	0,998
	A4	Águas do Centro Litoral	Infant	0,724	0.735	0,750	0,740	0,771	0,750
	A5	Águas do Norte	Infant	0,823	0.841	0,846	0,848	0,888	0,880
	A6	Águas do Vale do Tejo	Soldier	0,803	0.851	0,836	0,852	1,001	0,868
	A7	Águas do Vouga	Infant	0,823	0.602	0,853	0,606	0,893	0,615
	A8	Águas Públicas do Alentejo	Infant	0,754	0.709	0,818	0,714	0,754	0,842
	A9	EPAL	Learner	0,764	0.993	0,788	0,996	0,911	0,999
	A10	ICOVI	Infant	0,665	0.670	0,686	0,676	0,667	0,794
2017	A1	Águas de Santo André	Infant	0,653	0.764	0,663	0,770	0,765	0,822
	A2	Águas do Algarve	Star	0,858	0.978	0,887	0,979	0,933	0,998
	A3	Águas do Douro e Paiva	Star	0,811	0.891	0,840	0,898	0,974	0,917
	A4	Águas do Centro Litoral	Learner	0,775	0.918	0,794	0,925	0,826	0,937
	A5	Águas do Norte	Star	0,874	0.935	0,908	0,942	0,987	0,954
	A6	Águas do Vale do Tejo	Soldier	0,801	0.859	0,838	0,860	1,009	0,876
	A7	Águas do Vouga	Infant	0,821	0.708	0,855	0,708	0,885	0,722
	A8	Águas Públicas do Alentejo	Infant	0,772	0.734	0,831	0,739	0,773	0,872
	A9	EPAL	Star	0,787	0.993	0,824	0,996	0,956	0,999
	A10	ICOVI	Infant	0,877	0.703	0,936	0,709	0,877	0,835
2018	A1	Águas de Santo André	Infant	0,659	0.764	0,671	0,770	0,790	0,822
	A2	Águas do Algarve	Star	0,865	0.978	0,904	0,979	0,950	0,998
	A3	Águas do Douro e Paiva	Star	0,793	1.000	0,820	1,000	0,938	1,000
	A4	Águas do Centro Litoral	Infant	0,769	0.881	0,788	0,887	0,819	0,898
	A5	Águas do Norte	Learner	0,816	0.935	0,843	0,942	0,908	0,954
	A6	Águas do Vale do Tejo	Infant	0,776	0.860	0,799	0,860	0,855	0,884
	A7	Águas do Vouga	Infant	0,861	0.866	0,928	0,873	0,913	0,887
	A8	Águas Públicas do Alentejo	Soldier	0,939	0.738	0,988	0,744	0,940	0,876
	A9	EPAL	Star	0,791	0.993	0,827	0,996	0,960	0,999
	A10	ICOVI	Learner	0,695	0.760	0,752	0,765	0,734	0,902
2019	A1	Águas de Santo André	Infant	0,637	0.802	0,646	0,808	0,733	0,863
	A2	Águas do Algarve	Star	0,880	0.979	0,914	0,980	0,995	0,999
	A3	Águas do Douro e Paiva	Star	0,809	1.000	0,832	1,000	0,942	1,000
	A4	Águas do Centro Litoral	Learner	0,788	0.904	0,810	0,911	0,842	0,923
	A5	Águas do Norte	Star	0,826	0.981	0,855	0,988	0,939	1,000
	A6	Águas do Vale do Tejo	Infant	0,778	0.862	0,799	0,863	0,863	0,894
	A7	Águas do Vouga	Soldier	0,880	0.867	0,947	0,875	0,922	0,894
	A8	Águas Públicas do Alentejo	Soldier	1,000	0.738	1,046	0,744	1,001	0,879
	A9	EPAL	Star	0,799	0.993	0,834	0,996	0,963	0,999
	A10	ICOVI	Learner	0,786	0.801	0,871	0,808	0,833	0,949
2020	A1	Águas de Santo André	Infant	0,660	0.825	0,672	0,832	0,833	0,885
	A2	Águas do Algarve	Star	0,898	0.979	0,932	0,980	1,002	0,999
	A3	Águas do Douro e Paiva	Star	0,801	1.000	0,823	1,000	0,929	1,000
	A4	Águas do Centro Litoral	Learner	0,778	0.896	0,799	0,902	0,831	0,914
	A5	Águas do Norte	Star	0,811	0.981	0,839	0,988	0,922	1,000
	A6	Águas do Vale do Tejo	Soldier	0,801	0.862	0,836	0,863	0,997	0,894
	A7	Águas do Vouga	Star	1,000	0.869	1,138	0,876	1,057	0,901
	A8	Águas Públicas do Alentejo	Infant	0,818	0.724	0,850	0,730	0,819	0,861
	A9	EPAL	Star	0,775	1.000	0,807	1,004	0,932	1,000
	A10	ICOVI	Infant	0,710	0.703	0,776	0,709	0,767	0,836

References

[1] Brundtland GH. Our common future—Call for action. Environ Conserv 1987;14(4):291–4.

[2] United Nations. The united nations world water development report 2015: Water for a sustainable world. Paris: UNESCO; 2015, p. 122.

[3] Hall N, Richards R, Barrington D, Ross H, Reid S, Head B, Jagals P, Dean A, Hussey K, Abal E, Ali S, Boully L, Willis J. Achieving the UN sustainable development goals for water and beyond. Brisbane, Australia: Global Change Institute, The University of Queensland; 2016, URL: http://www.gci.uq.edu.au/un-sustainable-development-goals-water-sanitation-and-hygiene.

[4] Mazumder RK, Salman AM, Li Y, Yu X. Performance evaluation of water distribution systems and asset management. J Infrastruct Syst 2018;24:24.

[5] ISO/Technical Committee 251. Asset management - Achieving the UN sustainable development goals asset management, First edition. 2018.

[6] ISO. ISO 55000:2014 . Asset management — Overview, principles and terminology. Technical report, International Organization for Standardization; 2014.

[7] ISO. ISO 55001:2014. Asset management — Management systems — Requirements. Technical report, International Organization for Standardization; 2014.

[8] Luís AM, Almeida NM. Gestão de ativos no setor da água em Portugal: passado, presente e futuro. In: Revista APDA -Associação Portuguesa de distribuição e drenagem de águas. 2021, p. 24–33, https://www.apda.pt/pt/pagina/87/revista-apda/.

[9] Marques RC, Pinto FS. How to watch the watchmen? The role and measurement of regulatory governance. Util Policy 2018;51:73–81.

[10] Cherchye L, Moesen W, Rogge N, Van Puyenbroeck T. An introduction to 'benefit of the doubt' composite indicators. Soc Indic Res 2007;82:111–45.

[11] Vieira J, Cabral M, Almeida N, Silva JG, Covas D. Novel methodology for efficiency-based long-term investment planning in water infrastructures. Struct Infrastruct Eng 2020;16:1654–68.

[12] Marlow D, Burn S. Effective use of condition assessment within asset management. Journal-American Water Works Association 2008;100(1):54–63.

[13] Afsharian M, Ahn H, Kamali S. Performance analytics in incentive regulation: A literature review of DEA publications. Decis Anal J 2022;4:100079.

[14] Berg SV. Conflict resolution: benchmarking water utility performance. Public Adm Dev: Int J Manag Res Pract 2007;27(1):1–11.

[15] Marques RC, De Witte K. Towards a benchmarking paradigm in European water utilities. Public Money Manag 2010;30:42–8.

[16] Heesche E, Asmild M. Controlling for environmental conditions in regulatory benchmarking. Util Policy 2022;77:101390.

[17] Marques RC. A yardstick competition model for Portuguese water and sewerage services regulation. Util Policy 2006;14(3):175–84.
[18] Vilanova MRN, Filho PM, Balestieri JAP. Performance measurement and indicators for water supply management: Review and international cases. Renew Sustain Energy Rev 2014;43:1–12.
[19] Henriques AA, Camanho AS, Amorim P, Silva JG. Performance benchmarking using composite indicators to support regulation of the Portuguese wastewater sector. Util Policy 2020;66:101082.
[20] Alegre H, Baptista JM, Jr. HC, Cubillo F, Duarte P, Hirner W, Merkel W, Parena R. Performance indicators for water supply services. 3rd ed.. London: IWA Publishing; 2017, p. 1–404.
[21] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European J Oper Res 1978;2(6):429–44.
[22] Berg S, Marques R. Quantitative studies of water and sanitation utilities: a literature survey. Water Policy 2011;591–606.
[23] Goh KH, See KF. Twenty years of water utility benchmarking: A bibliometric analysis of emerging interest in water research and collaboration. J Clean Prod 2021;284:124711.
[24] Thanassoulis E. DEA and its use in the regulation of water companies. European J Oper Res 2000;127(1):1–13.
[25] Thanassoulis E. The use of data envelopment analysis in the regulation of UK water utilities: Water distribution. European J Oper Res 2000;126(2):436–53.
[26] Walker N, Norton A, Harris I, Williams A, Styles D. Economic and environmental efficiency of UK and Ireland water companies: Influence of exogenous factors and rurality. J Environ Manag 2019;241:363–73.
[27] Byrnes J, Crase L, Dollery B, Villano R. The relative economic efficiency of urban water utilities in regional New South Wales and Victoria. Resour Energy Econ 2010;32(3):439–55.
[28] Wang S, Qiu S, Ge S, Liu J, Peng Z. Benchmarking Toronto wastewater treatment plants using DEA window and Tobit regression analysis with a dynamic efficiency perspective. Environ Sci Pollut Res 2018;25(32):32649–59.
[29] Berg S, Lin C. Consistency in performance rankings: The Peru water sector. Appl Econ 2008;40:793–805.
[30] Alsharif K, Feroz EH, Klemer A, Raab R. Governance of water supply systems in the Palestinian Territories : A data envelopment analysis approach to the management of water resources. J Environ Manag 2008;87:80–94.
[31] Dong X, Du X, Li K, Zeng S, Bledsoe B. Benchmarking sustainability of urban water infrastructure systems in China. J Clean Prod 2018;170:330–8.
[32] Marques RC, Berg S, Yane S. Nonparametric benchmarking of Japanese water utilities: Institutional and environmental factors affecting efficiency. J Water Resour Plan Manag 2014;140:562–71.
[33] Lo Storto C. Efficiency, conflicting goals and trade-offs: A nonparametric analysis of the water and wastewater service industry in Italy. Sustainability (Switzerland) 2018;10(4).
[34] D'Inverno G, Carosi L, Romano G. Environmental sustainability and service quality beyond economic and financial indicators: A performance evaluation of Italian water utilities. Socio-Econ Plan Sci 2021;75.
[35] Romano G, Guerrini A. Measuring and comparing the efficiency of water utility companies: A data envelopment analysis approach. Util Policy 2011;19:202–9.
[36] Bogetoft P. Incentives and productivity measurements. Int J Prod Econ 1995;39(1):67–77, Efficiency measurement.
[37] Bogetoft P. Incentive efficient production frontiers: An agency perspective on DEA. Manage Sci 1994;40(8):959–68.
[38] De Witte K, Marques RC. Designing performance incentives, an international benchmark study in the water sector. CEJOR Cent Eur J Oper Res 2010;18:189–220.
[39] De Witte K, Saal DS. Is a little sunshine all we need? On the impact of sunshine regulation on profits, productivity and prices in the Dutch drinking water sector. J Regul Econ 2010;37(3):219–42.
[40] Nardo M, Saltelli A, Tarantola S, Hoffman A, Giovannini E. Handbook on constructing composite indicators: methodology and user guide. OECD and European Union and Joint Research Centre-European Commission; 2008, p. 162.
[41] Melyn W, Moesen W. Towards a synthetic indicator of macroeconomic performance: Unequal weighting when limited information is available, public economics research paper 17. Center for Economic Studies, KU Leuven; 1991.
[42] Gruetzmacher SB, Vaz CB, Ferreira ÂP. Performance assessment of the transport sustainability in the European union. Commun Comput Inf Sci 2021;1359:106–19.
[43] Bowen HP, Moesen W. Composite competitiveness indicators with endogenous versus predetermined weights: An application to the World Economic Forum's global competitiveness index. Compet Rev: Int Bus J 2011.
[44] Lafuente E, Araya M, Leiva JC. Assessment of local competitiveness: A composite indicator analysis of Costa Rican counties using the 'Benefit of the Doubt' model. Socio-Econ Plan Sci 2020;100864.
[45] Rogge N. Composite indicators as generalized benefit-of-the-doubt weighted averages. European J Oper Res 2018;267:381–92.
[46] Van Puyenbroeck T, Rogge N. Comparing regional human development using global frontier difference indices. Socio-Econ Plan Sci 2020;70:1–8.
[47] Morais P, Camanho AS. Evaluation of performance of European cities with the aim to promote quality of life improvements. Omega 2011;39:398–409.
[48] Verbunt P, Rogge N. Geometric composite indicators with compromise Benefit-of-the-Doubt weights. European J Oper Res 2018;264:388–401.
[49] Pereira MA, Camanho AS, Figueira JR, Marques RC. Incorporating preference information in a range directional composite indicator: The case of Portuguese public hospitals. European J Oper Res 2021;294:633–50.
[50] Zanella A, Camanho AS, Dias TG. Benchmarking countries' environmental performance. J Oper Res Soc 2013;64:426–38.
[51] Amado CA, José JMS, Santos SP. Measuring active ageing: A Data Envelopment Analysis approach. European J Oper Res 2016;255:207–23.
[52] Henriques AA, Fontes M, Camanho AS, D'Inverno G, Amorim P, Silva JG. Performance evaluation of problematic samples: a robust nonparametric approach for wastewater treatment plants. Ann Oper Res 2022.
[53] De Witte K, Marques RC. Incorporating heterogeneity in non-parametric models: a methodological comparison. Int J Oper Res 2010;9(2):188–204.
[54] Bădin L, Daraio C, Simar L. Explaining inefficiency in nonparametric production models: the state of the art. Ann Oper Res 2014;214(1):5–30.
[55] Molinos-Senante M, Sala-Garrido R, Lafuente M. The role of environmental variables on the efficiency of water and sewerage companies: a case study of Chile. Environ Sci Pollut Res 2015;22(13):10242–53.
[56] Dong X, Zhang X, Zeng S. Measuring and explaining eco-efficiencies of wastewater treatment plants in China: An uncertainty analysis perspective. Water Res 2017;112:195–207.
[57] Daraio C, Simar L. Introducing environmental variables in nonparametric frontier models: a probabilistic approach. J Product Anal 2005;24:93–121.
[58] Daraio C, Simar L. Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications. Springer Science & Business Media; 2007.
[59] De Witte K, Kortelainen M. What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. Appl Econ 2013;45(17):2401–12.
[60] Mergoni A, D'Inverno G, Carosi L. A composite indicator for measuring the environmental performance of water, wastewater, and solid waste utilities. Util Policy 2022;74:101285.
[61] Tourinho M, Santos PR, Pinto FT, Camanho AS. Performance assessment of water services in Brazilian municipalities: An integrated view of efficiency and access. Socio-Econ Plan Sci 2022;79:101139.
[62] Gavrikova E, Volkova I, Burda Y. Strategic aspects of asset management: An overview of current research. Sustainability (Switzerland) 2020;12:9–11.
[63] Galar D, Berges L, Sandborn P, Kumar U. The need for aggregated indicators in performance asset management. Eksploat Niezawodn 2014;16:120–7.
[64] Cecconi FR, Moretti N, Dejaco MC. Measuring the performance of assets: A review of the facility condition index. Int J Strateg Prop Manag 2019;23:187–96.
[65] Gonçalves CDF, Dias JAM, Machado VAC. Multi-criteria decision methodology for selecting maintenance key performance indicators. Int J Manag Sci Eng Manag 2015;10:215–23.
[66] Dutuit Y, Rauzy A. On the extension of Importance Measures to complex components. Reliab Eng Syst Saf 2015;142:161–8.
[67] Attwater A, Wang JQ, Parlikad A, Russell P. Measuring the performance of asset management systems. In: IET conference publications, Vol. 2014. 2014, p. 1–6.
[68] Jasiulewicz-Kaczmarek M, Żywica P. The concept of maintenance sustainability performance assessment by integrating balanced scorecard with non-additive fuzzy integral. Eksploat Niezawodn 2018;20(4).
[69] Famurewa SM, Stenström C, Asplund M, Galar D, Kumar U. Composite indicator for railway infrastructure management. J Mod Transp 2014;22(4):214–24.
[70] Hassan J, Khan F. Risk-based asset integrity indicators. J Loss Prev Process Ind 2012;25(3):544–54.
[71] Marques RC, Simões P. Revisiting the comparison of public and private water service provision: an empirical study in Portugal. Water 2020;12(5):1477.
[72] Gonçalves I, Alves D, Robalo G. Social tariffs for water and waste services in mainland Portugal: An impact analysis. Water Sci Technol: Water Supply 2014;14(4):513–21.
[73] Marques RC. Comparing private and public performance of Portuguese water services. Water Policy 2008;10:25–42.
[74] Marques RC, Berg S. Risks, contracts, and private-sector participation in infrastructure. J Constr Eng Manag 2011;137:925–32.
[75] Pérez F, Delgado-Antequera L, Gómez T. A two-phase method to assess the sustainability of water companies. Energies 2019;12.
[76] ERSAR. Relatório anual dos serviços de águas e resíduos em Portugal (2021): Volume 1 - Caracterização do setor de águas e resíduos. 2021, https://www.ersar.pt/pt/site-publicacoes/Paginas/edicoes-anuais-do-RASARP.aspx.
[77] ERSAR. Relatório anual dos serviços de águas e resíduos em Portugal (2021): Volume 2 – Controlo da qualidade da água para consumo humano. 2021, https://www.ersar.pt/pt/site-publicacoes/Paginas/edicoes-anuais-do-RASARP.aspx.
[78] ERSAR and LNEC. Guia de avaliação da qualidade dos serviços de águas e resíduos prestados aos utilizadores – 3.ª geração do sistema de avaliação. - guia técnico 22. 2021, Issued by ERSAR - Entidade Reguladora dos Serviços de Águas e Resíduos and LNEC - Laboratório Nacional de Engenharia Civil. http://www.ersar.pt/pt/publicacoes/publicacoes-tecnicas/guias.

[79] Pinto FS, Costa AS, Figueira JR, Marques RC. The quality of service: An overall performance assessment for water utilities. Omega (United Kingdom) 2017;69:115–25.
[80] Pinto FS, Simões P, Marques RC. Water services performance: Do operational environment and quality factors count? Urban Water J 2017;14:773–81.
[81] APDA. Principais resultados do inquérito elaborado pela Comissão Especializada de Gestão de Ativos (CEGA). 2019, https://www.apda.pt/site/ficheiros_eventos/202104121030-inquerito_cega_2019.pdf.
[82] Marques R, Monteiro A. Application of performance indicators in water utilities management - A case-study in Portugal. Water Sci Technol 2001;44(2–3):95–102.
[83] Ferreira B, Carriço N. Urban water infrastructure asset management plan: Case study. Open Eng 2019;9(1):459–67, cited By 4.
[84] Cabral M, Loureiro D, Covas D. Using economic asset valuation to meet rehabilitation priority needs in the water sector. Urban Water J 2019;16:205–14.
[85] Santos L, Cardoso M, Galvão A. Storm water systems' performance: assessment framework application to Portuguese water utilities. Int J Water Resour Dev 2022. cited By 2.
[86] Marques RC, Monteiro A. Application of performance indicators to control losses - Results from the Portuguese water sector. Water Sci Technol: Water Supply 2003;3(1–2):127–33, cited By 8.
[87] Machado B, Carvalho T, Cupido C, Almeida M, Alegre H. Minimization of losses in water supply systems: Strategy definition in a Portuguese case study. Desalin Water Treat 2009;2(1–3):24–9, cited By 1.
[88] Loureiro D, Silva C, Cardoso M, Mamade A, Alegre H, Rosa M. The development of a framework for assessing the energy efficiency in urban water systems and its demonstration in the Portuguese water sector. Water (Switzerland) 2020;12(1). cited By 6.
[89] ERSAR and LNEC. Desenvolvimento e implementação de processos de gestão patrimonial de infraestruturas - Guia Técnico 21. 2017, Issued by ERSAR - Entidade Reguladora dos Serviços de Águas e Resíduos and LNEC - Laboratório Nacional de Engenharia Civil. http://www.ersar.pt/pt/publicacoes/publicacoes-tecnicas/guias.
[90] Zanella A, Camanho AS, Dias TG. Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. European J Oper Res 2015;245:517–30.
[91] Fried HO, Lovell CK, Schmidt SS, Schmidt SS, et al. The measurement of productive efficiency and productivity growth. Oxford University Press; 2008.
[92] Rogge N, Jaeger SD, Lavigne C. Waste performance of NUTS 2-regions in the EU: A conditional directional distance benefit-of-the-doubt model. Ecol Econom 2017;139:19–32.
[93] Wong Y-HB, Beasley JE. Restricting weight flexibility in data envelopment analysis. J Oper Res Soc 1990;41:829–35.
[94] Allen R, Athanassopoulos A, Dyson RG, Thanassoulis E. Weights restrictions and value judgements in Data Envelopment Analysis: Evolution, development and future directions. Ann Oper Res 1997;73:13–34.
[95] Sarrico CS, Dyson RG. Restricting virtual weights in data envelopment analysis. European J Oper Res 2004;159(1):17–34.
[96] D'Inverno G, De Witte K. Service level provision in municipalities: A flexible directional distance composite indicator. European J Oper Res 2020;286:1129–41.
[97] Van Puyenbroeck T, Montalto V, Saisana M. Benchmarking culture in Europe: A data envelopment analysis approach to identify city-specific strengths. European J Oper Res 2021;288:584–97.
[98] Cazals C, Florens J-P, Simar L. Nonparametric frontier estimation: a robust approach. J Econometrics 2002;106:1–25.
[99] De Witte K, Schiltz F. Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional Benefit-of-the-Doubt approach. European J Oper Res 2018;267(3):1172–81.
[100] Lavigne C, Jaeger SD, Rogge N. Identifying the most relevant peers for benchmarking waste management performance: A conditional directional distance Benefit-of-the-Doubt approach. Waste Manage 2019;89:418–29.
[101] Fusco E, Vidoli F, Rogge N. Spatial directional robust Benefit of the Doubt approach in presence of undesirable output: An application to Italian waste sector. Omega (United Kingdom) 2020;94.
[102] Li Q, Racine J. Nonparametric estimation of distributions with categorical and continuous data. J Multivariate Anal 2003;86(2):266–92.
[103] Kuosmanen T, et al. Modeling blank data entries in data envelopment analysis. In: Econ-WPA working paper at WUSTL, Vol. 210001. 2002.
[104] Haider H, Sadiq R, Tesfamariam S. Performance indicators for small-and medium-sized water supply systems: A review. Environ Rev 2014;22:1–40.
[105] Hornik K, Theussl S, Theussl MS. Package 'Rglpk'. 2019.
[106] Berkelaar M, et al. Package 'lpSolve'. 2015.
[107] Hayfield T, Racine JS. Nonparametric econometrics: The np package. J Stat Softw 2008;27:1–32.
[108] European Commission. Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. 1998, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31998L0083.
[109] Hax AC, Majluf NS. The use of the growth-share matrix in strategic planning. Interfaces 1983;13(1):46–60.
[110] Piran FS, Lacerda DP, Camanho AS, Silva MC. Internal benchmarking to assess the cost efficiency of a broiler production system combining data envelopment analysis and throughput accounting. Int J Prod Econ 2021;238:108173.

Hermilio Vilarinho is a Ph.D. student in Industrial Engineering and Management at the Faculty of Engineering (FEUP) of the University of Porto (Portugal). He has a degree in Mechanical Engineering from the Federal University of Bahia, Brazil and in Business Administration from the Catholic University of Salvador, Brazil. He has a master's degree in management and a specialisation in Environmental Management from the Federal University of Bahia. He has worked as an Engineering professional in leading positions in Brazil, USA, Angola and Japan, with emphasis on Quality Management and Project Management. He has also worked as a university professor.

Giovanna D'Inverno is a tenure-track assistant professor (RTDb) in the Quantitative Methods Area in the Department of Economics and Management at the University of Pisa (Italy). She received a joint Ph.D. degree in Economics at IMT School for Advanced Studies Lucca (Italy) and in Business Economics at KU Leuven (Belgium) in 2018. Until March 2022 she was a postdoctoral researcher at the Faculty of Economics and Business, KU Leuven.

Henriqueta Nóvoa is an Assistant Professor at the Faculty of Engineering (FEUP) of the University of Porto(Portugal). She holds a Ph.D. in Engineering Sciences from FEUP (2000) and a first degree in Electrical Engineering from FEUP. She has worked as a production test engineer at Texas Instruments for 6 years. She has been teaching courses in Quality Management, Statistics and Strategic Planning of Information Systems since 1990. She has been involved in research projects in diverse areas, such as wine industry, construction industry and packaging. She published more than 20 refereed papers in international conferences and scientific journals.

Ana S. Camanho is an Associate Professor at the Faculty of Engineering (FEUP) of the University of Porto (Portugal), and currently Vice-Rector of the University of Porto. She has a degree in Industrial Engineering and Management from FEUP and a Ph.D. in Industrial and Business Studies by Warwick Business School, UK. She is author of more than 80 articles in peer-reviewed journals in the areas of Operations Research and Data Science, with emphasis on the development of models for evaluation of efficiency and productivity using Data Envelopment Analysis. Her articles have more than 2800 citations (h-index 26) in Scopus database.