



Benchmarking the Performance of Productive Units Using Cross-Efficiency Techniques: An Empirical Approach for Water Companies

Ramon Sala-Garrido¹ · Manuel Mocholi-Arce¹ · Maria Molinos-Senante^{2,3} · Michail Smyrnakis⁴ · Alexandros Maziotis³

Received: 6 October 2021 / Accepted: 12 September 2023 / Published online: 30 September 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

Water companies provide essential services to society, such as drinking water and sanitation. Improving the management of these companies is fundamental for evaluating their performance. This study employs cross-efficiency data envelopment analysis techniques to assess a sample of water companies during the years 2010–2018. The assessment focuses on three main topics: i) the impact of service quality on water companies' performance; ii) the influence of environmental variables on efficiency and eco-efficiency in water companies and; iii) the effect of ownership on water company performance. The results reveal that service quality significantly influences water companies' performance. Additionally, it demonstrates that customer density and ownership of water companies impact economic and environmental efficiency, with public water companies showing the best performance. However, the quality of service of public water companies had deteriorated over time.

Keywords Cross-efficiency · Data envelopment analysis · Cluster analysis · Water utilities · Quality of service

1 Introduction

The production process involves transforming inputs into outputs. To assess the outcome of this process, it is essential to introduce the concept of efficiency. Efficiency refers to a situation in which all inputs could be jointly contracted when output levels do

✉ Maria Molinos-Senante
mmolinos@uc.cl

¹ Department of Mathematics for Economics, University of Valencia, Avd. Tarongers S/N, Valencia, Spain

² Centro de Derecho y Gestión del Agua UC, Avda. Libertador Bernardo O'Higgins 340, Santiago, Chile

³ Departamento de Ingeniería Hidráulica y Ambiental, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna, 4860 Santiago, Chile

⁴ Hartree Centre, Science and Technology Facility Council (STFC), Daresbury Warrington WA4 4AD, UK

not rise, or all outputs could be expanded when input levels do not rise (Ben Amor and Mellah 2023). The assessment of efficiency has been widely applied to various industries worldwide over the years (Emrouznejad and Yang 2018) and plays a crucial role in the water industry. This is because it can serve as a fundamental factor in setting water tariffs and can also help evaluate the impact of ownership on the performance of water companies (Suárez-Varela et al. 2017; Nuru et al. 2023).

Frontier methods offer two primary techniques for evaluating the efficiency of water companies: parametric (econometric) techniques, such as Stochastic Frontier Analysis (SFA), and non-parametric techniques, like Data Envelopment Analysis (DEA) (Goh and See 2021; Moosavian et al. 2022). The main advantage of DEA techniques over parametric ones is that they do not assume a specific functional form for the underlying technology (Marques and Simões 2020). However, the traditional DEA model, as developed by Charnes et al. (1978), computes the efficiency of Decision Making Units (DMUs) with total weight flexibility. In this approach, each DMU selects the weights that maximize its efficiency score. Consequently, several DMUs may be identified as DEA efficient, making further discrimination challenging (Wu et al. 2016; Liu et al. 2017). In other words, the traditional DEA model estimates efficiency based on a self-evaluation framework, potentially leading to overestimated efficiency scores (Ding et al. 2015).

To address this limitation, this study introduces the cross-efficiency technique for the first time to evaluate the efficiency of a sample of water companies. Unlike traditional methods, this approach requires each Decision Making Unit (DMU) not only to evaluate itself but also to be evaluated by its peers. It's important to note that optimal weights for efficient DMUs and cross-efficiency scores are not always unique (Tavana et al. 2021). To tackle this issue, different secondary models, such as the aggressive and benevolent models, were introduced by Sexton et al. (1986) and Doyle and Green (1994). The aggressive cross-efficiency DEA model involves minimizing the efficiency of other DMUs while maximizing the efficiency of each DMU, treating each as a competitor. In contrast, the benevolent cross-efficiency DEA model treats each DMU as a collaborator, aiming to maximize the efficiency of all units when evaluating one (Cui and Li 2020).

The objectives of this study are threefold. The first objective is to evaluate the impact of quality of service in the economic performance of water companies. In doing so, economic efficiency and eco-efficiency scores are computed using cross-efficiency techniques. The second objective is clustering water companies based on its economic and environmental performance. The third objective is to analyze the influence of exogenous or environmental variables on the performance of water companies.

To demonstrate the utility and advantages of this methodological approach, an empirical application focused on the Chilean water industry was conducted. The Chilean water industry underwent privatization between 1998 and 2004, resulting in two types of private companies: full private and concessionary (Sala-Garrido et al. 2019). Additionally, a small proportion of customers are served by a public company (SISS 2021). The unique structure of the Chilean water industry and the evaluation of both private and public companies' efficiency are of great interest to researchers and policymakers. Furthermore, it's worth noting that the Chilean water regulator currently does not provide financial incentives to companies to address undesirable outputs, such as water leakage and unplanned interruptions. Therefore, studying the environmental performance of water companies, specifically their quality of service, becomes crucial. A literature review conducted by Cetrulo et al. (2019) revealed that most studies conducted in developing countries, including Chile, did not include quality variables due to methodological limitations and data availability.

This research contributes to the literature as follows. In contrast to conventional DEA methods applied by previous studies evaluating the performance of water companies, this study employs a unique approach that in addition to self-evaluation considers peer-evaluation. By incorporating peer-evaluation, this method offers a more robust and realistic efficiency evaluation, providing valuable insights for policy-making and performance improvement in the water sector (Sala-Garrido et al. 2023). Additionally, this study takes into account an often overlooked aspect of service quality within the current regulatory framework, incorporating factors such as water leakage and unplanned supply interruptions as undesirable outputs in efficiency assessment. This analytical process is conducted at the ownership level, shedding light on the historical efficiency levels of both public and private water companies in Chile. Lastly, the analysis conducted extends beyond cost and quality of service, examining the impact of factors like customer density and the source of water resources on water companies efficiency.

2 Methodology

The methodology employed in this study comprises three primary stages. Firstly, efficiency scores of water companies were assessed by initially considering only desirable outputs (Eqs. 1–4) and subsequently incorporating quality of service variables as undesirable outputs (Eqs. 5–8). In the second stage, a clustering analysis of water companies based on the efficiency scores was conducted. The final stage focused on identifying potential external variables that may influence the performance of water companies (Eq. 9).

2.1 Economic and Environmental (Eco-Efficiency) Assessment

This section describes the methodology used to estimate the efficiency of the water companies with and without the inclusion of undesirable outputs using cross-efficiency DEA techniques.

Let's assume that there are m DMUs (i.e., water companies) that use a set of n inputs, $x_{ij}(i = 1 \dots n)$ to generate a set of p desirable outputs, $y_{rj}(r = 1 \dots p)$. The relative efficiency of any DMU l , θ_{ll} when desirable outputs are considered only in the analysis can be derived by solving the following classical DEA model:

$$\begin{aligned}
 & \max \theta_{ll} = \sum_{r=1}^p u_{rl} y_{rl} && l = 1 \dots m \\
 & \text{subject to :} \\
 & \sum_{i=1}^n v_{il} x_{il} = 1 \\
 & \sum_{r=1}^p u_{rl} y_{rj} - \sum_{i=1}^n v_{il} x_{ij} \leq 0 \quad j = 1 \dots m \\
 & u_{rl} \geq 0 && r = 1 \dots p \\
 & v_{il} \geq 0 && i = 1 \dots n
 \end{aligned}
 \tag{1}$$

where u_{rl} and v_{il} denote the weight for each desirable output and input for each DMU l , respectively. The self-evaluated efficiency of any DMU l can be calculated from model (1) as follows:

$$\theta_{ll}^* = \sum_{r=1}^p u_{rl}^* y_{rl}
 \tag{2}$$

where * represents optimal values (Bevilacqua et al. 2015). The cross-efficiency which is its peer evaluation to DMU l can be calculated as follows:

$$\theta_{jl} = \frac{\sum_{r=1}^p u_{rl}^* y_{rl}}{\sum_{i=1}^n v_{il}^* x_{il}} \quad j = 1 \dots m; j \neq l \tag{3}$$

Since cross-efficiency scores calculated from Eq. (3) are not unique (Liang et al. 2008; Moeini et al. 2015), previous studies have proposed secondary goals to optimize input and output weights while keeping the efficiency scores of Eq. (1) unchanged (Aparicio et al. 2020). In line with prior research (Wu et al. 2009; Falagario et al. 2012), and given the preference for relative dominance among DMUs, it was adopted the aggressive formulation of cross-efficiency DEA. This approach aims to minimize efficiency scores (Falagario et al. 2012). To achieve this, the following linear programming is solved:

$$\begin{aligned} & \min \sum_{r=1}^p u_{rl} \left(\sum_{j=1, j \neq l}^m y_{rj} \right) \\ & \text{subject to :} \\ & \sum_{i=1}^n v_{il} \left(\sum_{j=1, j \neq l}^m x_{ij} \right) = 1 \\ & \sum_{r=1}^p u_{rl} y_{rl} - \theta_{ll}^* \sum_{i=1}^n v_{il} x_{il} = 0 \\ & \sum_{r=1}^p u_{rl} y_{rj} - \sum_{i=1}^n v_{il} x_{ij} \leq 0 \quad j = 1 \dots m; j \neq m \\ & u_{rl} \geq 0 \quad r = 1 \dots p \\ & v_{il} \geq 0 \quad i = 1 \dots n \end{aligned} \tag{4}$$

As a result, the cross-efficiency value matrix gathers for each DMU, both on efficiency value ($\theta_{ll}, l = 1 \dots m$) and the $(m - 1)$ cross-efficiency values. The cross-efficiency score is the average of the self-evaluation and peers evaluations efficiencies (Bevilacqua et al. 2015).

To estimate efficiency of water companies from an environmental and economic (eco-efficiency) perspective, a set of undesirable outputs $q, y_{bj}(b = 1 \dots q)$ was included in the above models. Thus, the DEA model in Eq. (1) with undesirable outputs can be written as follows (Liu et al. 2017):

$$\begin{aligned} & \max \theta_{ll} = \sum_{r=1}^p u_{rl} y_{rl} + \sum_{b=1}^q z_{bl} y_{bl} \quad l = 1 \dots m \\ & \text{Subject to :} \\ & \sum_{i=1}^n v_{il} x_{il} = 1 \\ & \sum_{r=1}^p u_{rl} y_{rj} + \sum_{b=1}^q z_{bl} y_{bj} - \sum_{i=1}^n v_{il} x_{ij} \leq 0 \quad j = 1 \dots m \\ & u_{rl} \geq 0 \quad r = 1 \dots p \\ & z_{bl} \geq 0 \quad b = 1 \dots q \\ & v_{il} \geq 0 \quad i = 1 \dots n \end{aligned} \tag{5}$$

where z_{bl} denotes the weight of undesirable output q for any DMU l . The self-evaluated efficiency and cross eco-efficiency of any DMU l in the presence of undesirable outputs can be calculated as follows, respectively:

$$\theta_{ll}^* = \sum_{r=1}^p u_{rl}^* y_{rl} + \sum_{b=1}^q z_{bl}^* y_{bl} \tag{6}$$

and

$$\theta_{jl}^* = \frac{\sum_{r=1}^p u_{rl}^* y_{rl} + \sum_{b=1}^q z_{bl}^* y_{bl}}{\sum_{i=1}^n v_{il}^* x_{il}} \quad j = 1 \dots m; j \neq l \tag{7}$$

The secondary optimization model in Eq. (5) with undesirable outputs is as follows:

$$\begin{aligned}
 & \min \sum_{r=1}^p u_{rl} \left(\sum_{j=1, j \neq l}^m y_{rl} \right) + \sum_{b=1}^q z_{bl} \left(\sum_{j=1, j \neq l}^q y_{bl} \right) \\
 & \text{Subject to :} \\
 & \sum_{i=1}^n v_{il} \left(\sum_{j=1, j \neq l}^n x_{ij} \right) = 1 \\
 & \sum_{r=1}^p u_{rl} y_{rl} + \sum_{b=1}^q z_{bl} y_{bj} - \theta_{il}^{UD*} \sum_{i=1}^n v_{il} x_{il} = 0 \\
 & \sum_{r=1}^p u_{rl} y_{rj} + \sum_{b=1}^q z_{bl} y_{bj} - \sum_{i=1}^n v_{il} x_{ij} \leq 0 \qquad j = 1 \dots m; j \neq m \\
 & u_{rl} \geq 0 \qquad r = 1 \dots p \\
 & z_{bl} \geq 0 \qquad b = 1 \dots q \\
 & v_{il} \geq 0 \qquad i = 1 \dots n
 \end{aligned} \tag{8}$$

The cross-efficiency scores with undesirable outputs is the average of self-evaluations and peer evaluation efficiencies.

2.2 Clustering of Water Companies

To gain deeper insights into the performance of water companies, a cluster analysis was conducted to group them based on their efficiency scores. This categorization allows us to identify which water companies belong to the best-performing and worst-performing groups. Consequently, policymakers can introduce more tailored and group-specific policies to enhance the performance of water companies.

Cluster analysis is a statistical technique that groups research objects, such as companies, persons, or countries, based on shared characteristics (Xie et al. 2019). Several previous studies have employed clustering methods like hierarchical cluster analysis, k-means, and two-step cluster analysis (for a comprehensive review, please refer to Wu et al. 2010; Samoilenko and Osei-Bryson 2010; Jiang et al. 2019) to categorize entities based on their efficiency scores. Studies by Sikka et al. (2009), Omrani et al. (2018), Cinaroglu (2020), and others have highlighted the advantages of combining DEA and k-means clustering techniques to identify Decision Making Units (DMUs) with similar efficiencies and characteristics. Additionally, the k-means clustering technique is known for its speed, ease of use, and ability to handle large datasets, making it a widely preferred choice for pattern analysis (Swenson et al. 2016; Cinaroglu 2020).

The K-means algorithm involves an iterative process comprising three main steps: initialization, assignment, and update. In the initialization step, it was defined the number of clusters, 'k,' and select initial centroids, typically chosen randomly (Tang et al. 2018). Next, in the assignment step, each object is assigned to the nearest centroid (Park and Jun 2009). In the update step, new centroids are calculated for each cluster by averaging the values of the objects within that cluster (Bayaraa et al. 2019). The assignment and update steps are repeated iteratively until the algorithm converges, meaning there are no further changes in the assignments of objects to clusters.

A crucial parameter in the K-means algorithm is the number of clusters, 'k.' To determine the optimal 'k,' it was utilized the silhouette value (Cinaroglu 2020) which measures the similarity among objects within the same cluster and ranges from zero to one, with one indicating perfect similarity among cluster members.

2.3 Analysis of the Influence of Environmental Variables on Efficiency

In the final step of the performance assessment, it was investigated the impact of various environmental (exogenous) factors on the efficiency of water companies.

Cross-efficiency scores were regressed against a set of environmental variables. The dependent variable, namely the efficiency score, is non-negative and truncated, with values falling between zero and 1 and therefore, to ensure robust estimates in this context, it was employed a Tobit regression (Wang et al. 2024):

$$\theta_{ll,t} = \beta_0 + \beta_i \eta'_{l,t} + s_l + \varepsilon_{l,t} \quad (9)$$

where θ_{ll} presents the cross-efficiency of each DMU l with desirable or undesirable outputs, β_0 is the constant term, $\eta'_{l,t}$ is the set of operating characteristics (explanatory variables) of any DMU l , s_l presents DMU unobserved heterogeneity (e.g., managerial ability) which is uncorrelated with the explanatory variables and follows the standard normal distribution, $\varepsilon_{l,t}$ is error (noise) term that follows a standard normal distribution and t denotes time.

3 Data and Sample Selection

The sample evaluated in this study comprises 21 water companies providing water and sanitation services in Chile from 2010 to 2018. This group includes 10 full private, 10 concessionary, and 1 public water company. The data used is sourced from the national water regulator, the Superintendencia de Servicios Sanitarios (SISS), which monitors water company performance and sets tariffs (Molinos-Senante et al. 2018a).

The selection of inputs, desirable and undesirable outputs was informed by prior studies in the water industry (e.g., Carvalho et al. 2012; Pinto et al. 2017; Molinos-Senante et al. 2018a; Cetrulo et al. 2019; Goh and See 2021). The first input was the operating expenditure for water and sanitation services, measured in thousands of Chilean pesos per year (CLP/year) (Molinos-Senante and Sala-Garrido 2015; Ferro and Mercadier 2016). The second input was the network length, measured in kilometers (km) (Mellah and Ben Amor 2016). Desirable outputs selected were the volume of drinking water delivered, measured in thousands of cubic meters per year (Brea-Solis et al. 2017; Maziotis et al. 2020) and the annual number of customers receiving wastewater treatment (Molinos-Senante et al. 2019). To incorporate the quality of service provided by water companies into the assessment, it was multiplied the water and wastewater outputs by the drinking water quality and wastewater treatment quality indexes, respectively (Saal et al. 2007; Sala-Garrido et al. 2019). These quality indexes, ranging from zero to one, are developed by the water regulator (SISS 2021). Additionally, the two undesirable outputs considered were the volume of water leakage, measured in thousands of cubic meters per year (Molinos-Senante et al. 2019) and the number of unplanned water supply interruptions, measured in hours per year (Sala-Garrido et al. 2019; Maziotis et al. 2020).

Previous studies on the water industry have highlighted the potential influence of various environmental variables on companies' efficiency, making their consideration crucial in the analysis (e.g., Pinto et al. 2017; López-Ruiz et al. 2023). Therefore, the following environmental variables were considered: i) Customer Density: Calculated as the ratio of the number of customers to the network length; ii) Type of Water Resource: A categorical variable capturing whether the water resource is surface, groundwater, or a mix of both and; iii) Type of Ownership: Categorized to capture whether the company is public, fully private, or concessionary. Table 1 presents a summary of the variables used the study.

Table 1 Descriptive statistics of the variables used

Variables	Unit of measurement	Mean	Std. Dev	Minimum	Maximum
Volume of water delivered	000s m ³ /year	52328	94251	1064	473846
Customers receiving wastewater treatment	nr	762173	1334811	5835	6497126
Operating expenditure	000s CLP/year	34328761	44140687	823816	210144159
Network length	km	3439	4968	88	21859
Drinking water quality indicator	Index	0.974	0.052	0.700	1.000
Wastewater treatment quality indicator	Index	0.967	0.053	0.667	1.000
Volume of water leakage	000s m ³ /year	17049	29,530	104	142922
Water supply unplanned interruptions	hours/year	4345	7031	3	34051
Customer density	nr/km	57.890	14.618	19.407	89.237

Observations: 189

Operating expenditure is expressed in 2018 prices

4 Results and Discussion

4.1 Efficiency Assessment

GAMS (General Algebraic Modeling System) with the CPLEX solver was employed to run DEA models, allowing us to obtain cross-efficiency scores. The DEA model that considers only desirable outputs (excluding undesirable outputs) is denoted as Model DO, while the model that includes both desirable and undesirable outputs is denoted as Model DUO. The results, as depicted in Fig. 1, reveal that, on average, the Chilean water industry performed poorly when the quality of service was not included in the analysis. Specifically, it was found that, on average, water companies needed to reduce their operating costs and

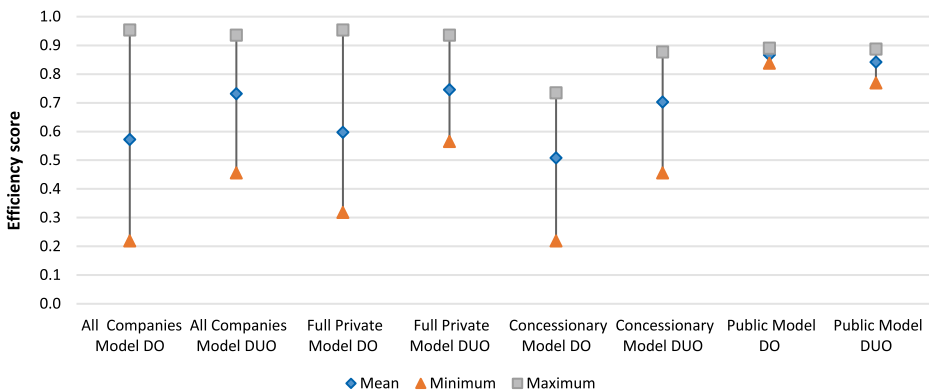


Fig. 1 Mean, minimum and maximum efficiency scores for all, full private, concessionary and public Chilean water companies based on models excluding and integrating undesirable outputs (Model DO and Model DUO, respectively)

capital inputs by 43% to generate the same level of output. The least efficient company had to make substantial improvements, requiring a contraction of inputs by 78.1% to maintain the same output level. However, when quality of service variables were included in the assessment, water companies' efficiency considerably improved, reaching an average efficiency level of 0.732. This suggests that, on average, water companies needed to reduce costs, water leakage, and unplanned interruptions by 26.8% to provide the same level of outputs to their customers.

From a management perspective, the observed low mean efficiency in the Chilean water industry signals a need for substantial improvements in daily operations to enhance overall efficiency. These findings align with previous studies by Molinos-Senante et al. (2018a) and Sala-Garrido et al. (2019), which also reported limited productivity improvements over time for the average Chilean water company. It is worth noting that some earlier studies in the Chilean water industry reported slightly higher efficiency levels (Ferro and Mercadier 2016; Molinos-Senante et al. 2018b). However, it's essential to consider the methodologies used in these studies. For instance, Ferro and Mercadier (2016) employed parametric techniques, which assume a priori a specific functional form. In contrast, Molinos-Senante et al. (2018b) used traditional DEA techniques, where each DMU self-evaluated relative to others without incorporating peer evaluation.

Figure 1 illustrates that, in both models, the public water company exhibited higher efficiency compared to private companies. However, it appeared to have not improved its quality of service, as its efficiency decreased. On average, the public company needed to further reduce water leakage, unplanned interruptions, and costs by 15.8% to achieve the same output level. Moreover, full private water companies demonstrated greater efficiency improvements in quality of service compared to concessionary companies. On average, a 25.4% reduction in inputs and undesirable outputs was required to enhance their efficiency. Finally, concessionary companies had an efficiency score of 0.703 in both models, indicating a need for a 30% reduction in inputs and undesirable outputs to improve efficiency. Additionally, there was significant room for efficiency improvement within the concessionary category, with a notable difference between the least and most efficient companies. Furthermore, a statistically significant difference in average efficiency scores between the two models highlights the importance of considering quality of service when assessing water company efficiency (see Table 5 in the Appendix).

Table 2 presents the trends in efficiency scores for different water company types. When undesirable outputs were not considered in the analysis (Model DO), full private companies exhibited a relatively stable trend with a gradual improvement over time. Their efficiency increased at an average rate of 0.22% per year, particularly showing an upward trend during the years 2015–2017. This slight increase in efficiency could be attributed to small savings in operating costs. In contrast, when undesirable outputs were included in the analysis (Model DUO), efficiency became more volatile but maintained an overall upward trend. This improvement was primarily due to a reduction in the frequency of unplanned water supply interruptions. Full private companies' efficiency increased from 0.735 in 2010 to 0.765 in 2018, marking an average increase of 4.14%. However, despite these gains, the overall performance of full private companies remained unsatisfactory. Enhancing efficiency would require more effective asset management and improvements in quality of service, such as reducing water leakage levels.

Concessionary water companies consistently reported the lowest efficiency scores among the various types of companies. When quality of service was not factored into the analysis (Model DO), they exhibited significant inefficiency over time, primarily stemming from their inability to reduce operating costs. Efficiency declined from 0.539 in 2010 to 0.499 in 2018,

Table 2 Evolution over time of efficiency scores by water company type

Year	Full private water companies		Concessionary water companies		Public water company	
	Model DO	Model DUO	Model DO	Model DUO	Model DO	Model DUO
2010	0.587	0.735	0.539	0.706	0.885	0.861
2011	0.598	0.748	0.534	0.713	0.884	0.852
2012	0.594	0.734	0.505	0.687	0.89	0.853
2013	0.59	0.736	0.503	0.684	0.844	0.821
2014	0.588	0.734	0.496	0.688	0.885	0.836
2015	0.598	0.746	0.49	0.685	0.86	0.855
2016	0.603	0.759	0.511	0.726	0.853	0.887
2017	0.615	0.761	0.498	0.716	0.867	0.839
2018	0.597	0.765	0.499	0.725	0.838	0.769
Average 2010–2018	0.597	0.746	0.508	0.703	0.867	0.842

marking an average annual decrease of 0.93%. On average, concessionary companies needed to reduce costs by 50% to generate the same output level. However, the situation improved when quality of service was included in the analysis (Model DUO). Efficiency showed an average annual increase of 0.32%, following a trend similar to that of full private companies. Efficiency declined during the years 2010–2012, largely due to an increase in water leakage and unplanned interruptions. Nevertheless, in subsequent years, improvements in reducing the frequency of unplanned interruptions offset any increases in water leakage. Despite this improvement, concessionary companies still required a 30% reduction in costs and undesirable outputs to provide the same level of water services in 2018 as in 2010.

Similar to concessionary companies, the efficiency of the public water company showed a downward trend when quality of service was disregarded (Model DO). Over time, its efficiency declined at a rate of 0.6% per year, resulting in a decrease of 5.37% from an efficiency score of 0.885 in 2010 to 0.838 in 2018. However, unlike private companies, the public company reported lower efficiency scores when quality of service was included in the analysis (Model DUO), indicating a deterioration in the service quality it provided. Despite reporting the lowest levels of unplanned interruptions, this company had the highest levels of water leakage in the industry.

In summary, these results underscore the need for water companies to enhance their efficiency by implementing improved management practices. Improving the quality of service to customers, possibly through the adoption of new technologies for better prediction and prevention of network leakage, is of paramount importance.

4.2 Clustering of Water Companies According to its Performance

For the entire study period (2010–2018), the optimal number of clusters was determined by selecting the highest silhouette measure. This analysis led to the formation of two distinct clusters: one comprising the worst-performing water companies (Cluster I) and the other consisting of the top-performing water companies (Cluster II).

In Model DO (excluding quality of service) over the entire study period, the worst-performing group (Cluster I) comprised 12 water companies with an average efficiency score of 0.470. This indicated that these companies needed to reduce their costs by an average of 53% to achieve the same output level. The efficiency range within this group varied between 0.467 and 0.478 on average, emphasizing the need for substantial cost reduction—up to 53.2% on average—to improve managerial practices. Conversely, the top-performing group (Cluster II) included the same 9 water companies throughout the study period, with an average efficiency score of 0.707. This group needed to enhance daily operations by an average of 29.3% to achieve the same output level. Remarkably, these same companies also constituted the top-performing group in Model DUO (including quality of service). These findings suggest that both groups struggled to manage their assets efficiently, with over half of the companies falling into the worst-performing group. The potential for improvement lies in the adoption of best management practices from the top-performing group (Table 3).

When quality of service was factored into the analysis (Model DUO), fewer water companies (9 companies) fell into the worst-performing cluster over the entire study period. Notably, several companies (WaSC5, WaSC6, and WaSC13) moved from the worst-performing cluster to the best-performing one. Cluster I, which included the worst-performing companies, had an average eco-efficiency score of 0.658. This suggests that, on average, these water companies needed to reduce costs and undesirable outputs such as water leakage and unplanned supply interruptions by 35.2% to improve efficiency. In contrast, the top-performing group, comprising 12 companies across all years, reported an average eco-efficiency score of 0.788. Their efficiency ranged from 0.763 to 0.816, indicating room for substantial cost reduction and enhancement of undesirable output management to improve efficiency further. Although the findings indicated improvements in service quality, there remains potential for additional savings and performance enhancement, ranging from 18.4% to 23.4% on average.

4.3 Influence of Environmental Variables on the Performance of Water Companies

Two Tobit regression models were conducted, one using cross-efficiency without the inclusion of service quality variables (Model DO) and the other with the inclusion of service quality (Model DUO). The results are presented in Table 4. In both models, we found that customer density and type of ownership had a statistically significant impact on water company efficiency. This positive and statistically significant relationship indicates the presence of economies of customer density in the Chilean water industry, aligning with the findings of Molinos-Senante et al. (2018a). It suggests that densely populated areas may require lower operational costs for water and wastewater services than less densely populated areas. Additionally, improving service quality in densely populated areas might be more cost-effective. Conversely, the negative and statistically significant parameter related to the type of ownership suggests that public and full private water companies exhibited higher efficiency than concessionary companies. However, the impact of ownership type on efficiency was less pronounced when quality of service was included in the analysis (Model DUO).

From a policy perspective, some valuable insights for policy makers are as follows. Firstly, it is introduced a methodology that not only allows water companies to be assessed individually but also enables comparisons among them and their peers. This enhances the accuracy and robustness of efficiency scores. It was quantified the extent to which less efficient companies must reduce costs to catch up with the industry's most efficient firms. Furthermore, it enables an examination of the impact of service quality inclusion in the efficiency analysis, shedding light on whether service quality matters in the efficiency assessment of water

Table 3 Efficiency scores of cluster analysis by different models

	Mean	Minimum	Maximum	Water Companies
<i>Model DO</i>				
Cluster I (Worst performance)	0.470	0.464	0.478	WaSC4; WaSC5; WaSC6; WaSC7; WaSC12; WaSC13; WaSC15; WaSC17; WaSC18; WaSC19; WaSC20; WaSC21
Cluster II (Best performance)	0.707	0.691	0.725	WaSC1; WaSC2; WaSC3; WaSC8; WaSC9; WaSC10; WaSC11; WaSC14; WaSC16
<i>Model DUO</i>				
Cluster I (Worst performance)	0.658	0.648	0.670	WaSC4; WaSC7; WaSC12; WaSC15; WaSC17; WaSC18; WaSC19; WaSC20; WaSC21
Cluster II (Best performance)	0.788	0.763	0.816	WaSC1; WaSC2; WaSC3; WaSC5; WaSC6; WaSC8; WaSC9; WaSC10; WaSC11; WaSC13; WaSC14; WaSC16

Table 4 Estimates of Tobit regression for analyzing the influence of environmental variables on efficiency

Variables	Model DO				Model DUO			
	Coef	Std. Err	Z- stat	P-value	Coef	Std. Err	Z- stat	P-value
Constant	0.316	0.069	4.550	0.000	0.641	0.056	11.530	0.000
Customer density	0.007	0.001	10.080	0.000	0.003	0.001	4.970	0.000
Type of ownership	-0.089	0.030	-2.910	0.004	-0.037	0.022	-1.660	0.096
Type of water resource	0.001	0.024	0.030	0.978	-0.027	0.018	-1.530	0.127
Year								
2011	0.001	0.009	0.140	0.891	0.009	0.010	0.890	0.373
2012	-0.012	0.009	-1.340	0.179	-0.010	0.010	-0.990	0.323
2013	-0.022	0.009	-2.450	0.014	-0.014	0.010	-1.430	0.154
2014	-0.033	0.009	-3.540	0.000	-0.016	0.010	-1.640	0.100
2015	-0.036	0.009	-3.820	0.000	-0.012	0.010	-1.250	0.213
2016	-0.046	0.010	-4.650	0.000	0.004	0.010	0.360	0.716
2017	-0.040	0.010	-4.120	0.000	0.000	0.010	0.020	0.985
2018	-0.048	0.010	-4.970	0.000	0.004	0.010	0.440	0.662
Log-likelihood	354.04				350.016			
X ² (11)	131.34				61.870			
Prob > X ² (11)	0.000				0.000			

Bold coefficients are statistically significant from zero at the 5% level

Bold italic coefficients are statistically significant from zero at the 10% level

companies. Additionally, the conducted analysis empowers managers to gain deeper insights into the factors driving efficiency within their companies, such as customer density, water resource type, and ownership. Lastly, the applied methodology facilitates a comparison of efficiency scores across different types of ownership, providing further evidence on whether private water companies outperform or underperform public ones.

5 Conclusions

Assessing the efficiency of regulated companies is a crucial tool for the water regulator to safeguard the interests of citizens. It aids in establishing appropriate water tariffs and encouraging innovation within the sector. Operational cost efficiency alone does not necessarily equate to overall efficiency, as a water company could excel in cost control but lag in delivering quality service. Therefore, incorporating variables like water leakage volume and unplanned supply interruptions into performance assessments offers a more comprehensive view of water company performance. These variables directly influence the provision of water services and provide valuable insights into how effectively water companies operate.

Traditional parametric techniques like DEA often evaluate the efficiency of DMUs based solely on self-evaluation, leading to potentially unrealistic weights and inaccurate efficiency scores. This approach is unacceptable in a regulatory context. To address this limitation, the cross-efficiency DEA technique was employed for efficiency assessment. This method was applied to assess the economic and environmental efficiency (eco-efficiency) of various private and public water companies in Chile spanning the years

2010–2018. The key findings of the assessment are as follows. Firstly, when quality of service was not considered in the analysis, the average efficiency of Chilean water companies was low, with a score of 0.572. However, when quality of service was included in the analysis, the average efficiency significantly improved to 0.732. This indicates the positive impact of accounting for service quality on overall efficiency scores. Secondly, the public water company outperformed private ones on average, with full private companies performing better than concessionary ones in terms of efficiency. It is important to note that the public water company's performance in terms of quality of service deteriorated when undesirable outputs were included in the analysis, as reflected by a lower efficiency score. Thirdly, both full private and concessionary companies reported eco-efficiency scores of 0.746 and 0.703, respectively. Nevertheless, regardless of whether service quality was considered or not, the worst-performing companies needed substantial efficiency improvements, with approximately 53% improvement without service quality variables and 34.2% improvement with their inclusion, to catch up with the best-performing group. Finally, customer density and ownership type were found to have a statistically significant impact on companies' efficiency. More densely populated areas tend to require lower costs for delivering water and wastewater services than less densely populated areas.

The methodology proposed and the results obtained in this study hold significant value for researchers and policy makers for several compelling reasons: i) It offers a non-parametric methodology that connects the efficiency of each water company with the efficiency of all water companies, employing cross-efficiency. This technique mitigates the issue of over-estimation of efficiency scores, which is a common concern with traditional non-parametric methods; ii) It provides valuable insights into how water companies in Chile have performed over time, particularly in terms of reducing operational costs and enhancing the quality of service. It highlights the improvements made in reducing unplanned supply interruptions but also underscores the need for more substantial efforts to decrease daily operational costs and water leakage; iii) these findings can serve as a basis for regulatory and policy decisions. Regulators and policy makers can use this information to design policies that involve financial incentives or penalties to encourage water companies to become more innovative and efficient, thereby enhancing their economic and environmental performance and iv) water companies can benefit from this study by identifying areas where they can allocate resources more efficiently. Additionally, investing in technologies that enable accurate prediction and prevention of network leakages can be a strategic move to improve operational efficiency.

Appendix

Figures 2, 3, 4 and 5

Table 5 Correlations and difference comparison between efficiency measurements

	Model DO vs Model DUO
<i>Correlations</i>	
Pearson	0.860
Spearman	0.804
<i>Difference comparisons</i>	
Mean difference	-0.161
T test	-23.240

All estimates are statistically significant from zero at the 1% level

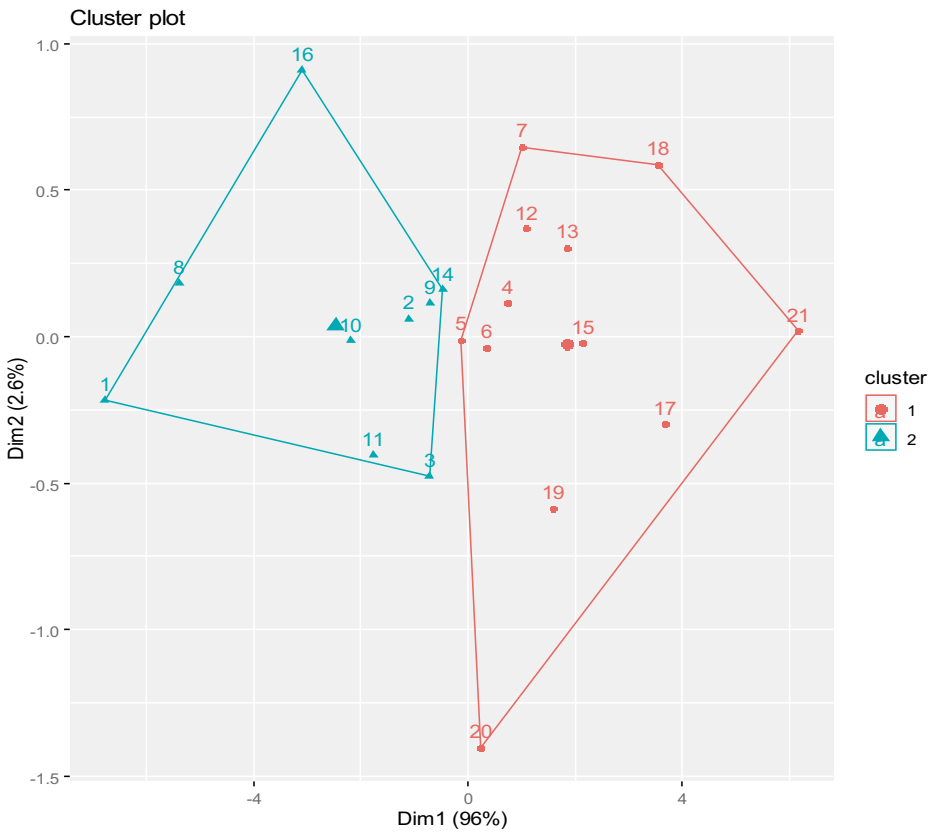


Fig. 2 Cluster analysis – Model DO

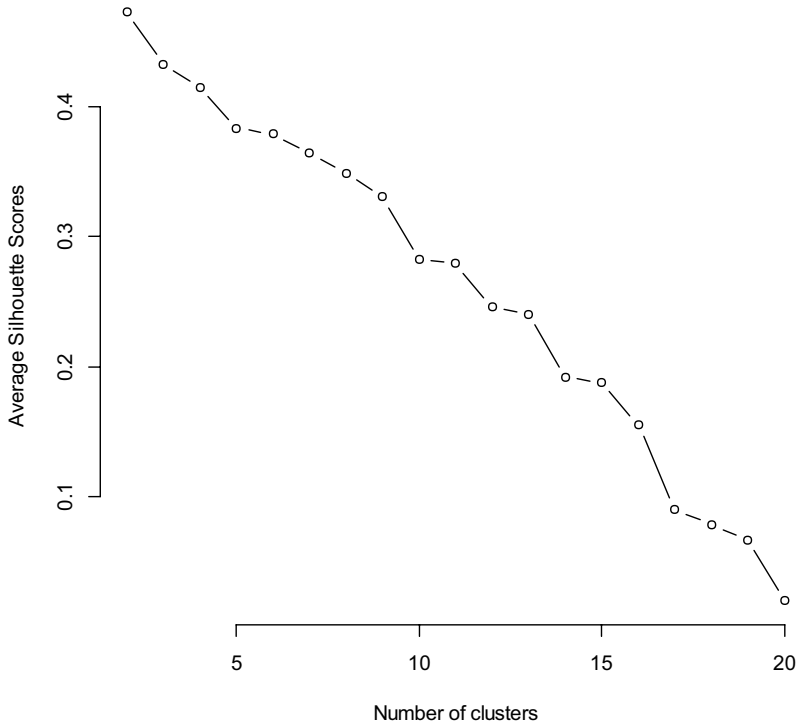
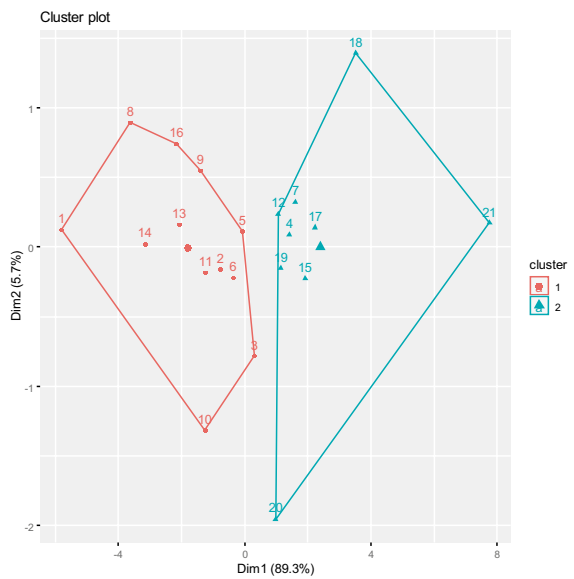


Fig. 3 Silhouette score – Model DO

Fig. 4 Cluster analysis – Model DUO



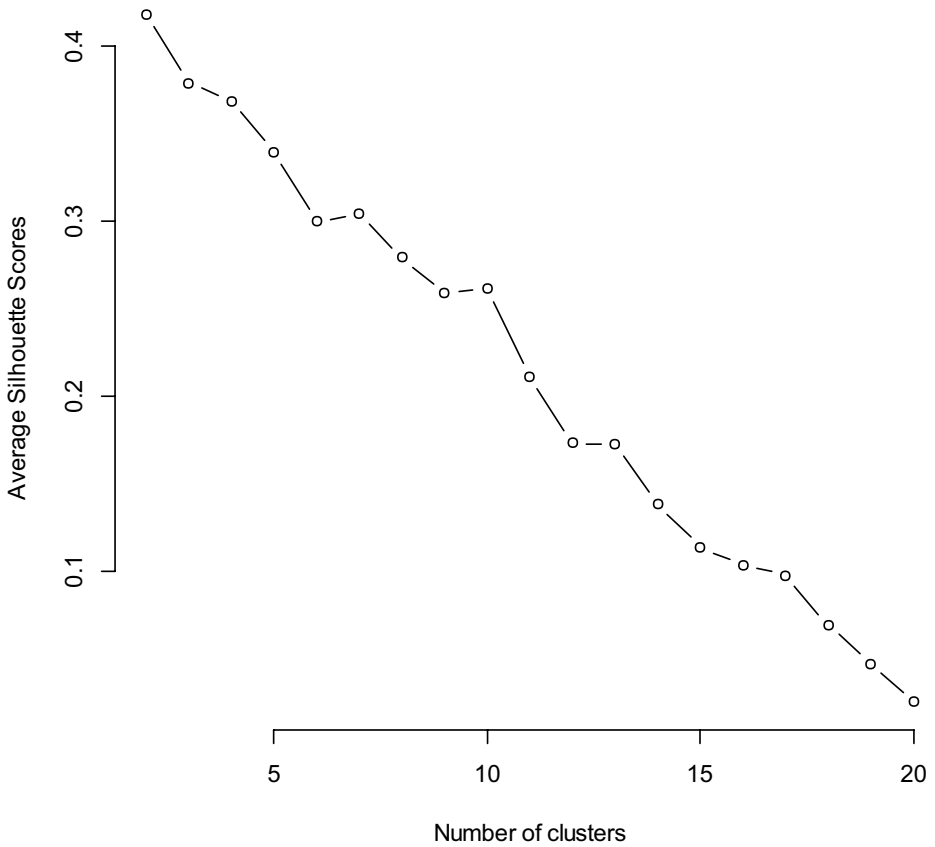


Fig. 5 Silhouette score – Model DUO

Authors Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Ramón Sal-Garrido, Manuel Mocholi-Arce and Alexandros Maziotis. The first draft of the manuscript was written by Alexandros Maziotis and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability Data will be available upon a reasonable request.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

References

Aparicio J, Ortiz L, Pastor JT, Zabala-Iturriagoitia JM (2020) Introducing cross-productivity: A new approach for ranking productive units over time in Data Envelopment Analysis. *Comput Ind Eng* 144:106456

- Bayarara B, Tarnoczi T, Fenyves V (2019) Measuring performance by integrating k-medoids with DEA: Mongolian case. *J Bus Econ Manag* 20(6):1238–1257
- Ben Amor T, Mellah T (2023) Cost efficiency of Tunisian water utility districts: Does heterogeneity matter? *Utilities Policy* 84:101616
- Bevilacqua M, Ciarapica FE, Mazzuto G, Paciarotti C (2015) Efficiency assessment of blanching and deep-freezing systems through data envelopment analysis. *Eng Agric Environ Food* 1–6
- Brea-Solis H, Perelman S, Saal DS (2017) Regulatory incentives to water losses reduction: the case of England and Wales. *J Prod Anal* 47(3):259–276
- Carvalho P, Marques RC, Berg S (2012) A meta-regression analysis of benchmarking studies on water utilities market structure. *Utilities Policy* 21:40–49
- Cetrulo TB, Marques RC, Malheiros TF (2019) An analytical review of the efficiency of water and sanitation utilities in developing countries. *Water Res* 161:372–380
- Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *Eur J Oper Res* 2:429–444
- Cinaroglou S (2020) Integrated k-means clustering with data envelopment analysis of public hospital efficiency. *Health Care Manag Sci* 23:325–338
- Cui Q, Li Y (2020) A cross efficiency distinguishing method to explore the cooperation degree in dynamic airline environmental efficiency. *Transp Policy* 99:31–43
- Ding ZY, Jo GS, Wang Y, Yeo GT (2015) The relative efficiency of container terminals in small and medium-sized ports in China. *Asian J Ship Log* 31(2):231–251
- Doyle J, Green RH (1994) Efficiency and cross-efficiency in DEA: Derivations, meanings and uses. *J Oper Res Soc* 45(5):567–578
- Emrouznejad A, Yang G (2018) A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socioecon Plann Sci* 61(1):1–5
- Falagario M, Sciancalepore F, Costantino N, Pietroforte R (2012) Using a DEA-cross efficiency approach in public procurement tenders. *Eur J Oper Res* 218:523–529
- Ferro G, Mercadier AC (2016) Technical efficiency in Chile's water and sanitation provides. *Utilities Policy* 43:97–106
- Goh KH, See KF (2021) Twenty years of water utility benchmarking: A bibliometric analysis of emerging interest in water research and collaboration. *J Clean Prod* 284:124711
- Jiang Z, Ding Z, Zhang H, Caic W, Liu Y (2019) Data-driven ecological performance evaluation for remanufacturing process. *Energy Convers Manag* 198:111844
- Liang L, Wu J, Cook WD, Zhu J (2008) Alternative secondary goals in DEA cross-efficiency evaluation. *J Clean Prod* 113(2):1025–1030
- Liu X, Chu J, Yin P, Sun J (2017) DEA cross-efficiency evaluation considering undesirable output and ranking priority: a case study of eco-efficiency analysis of coal-fired power plants. *J Clean Prod* 142:877–885
- López-Ruiz S, Ibáñez-Rueda N, Guardiola J, González-Gómez F (2023) Does the ownership of water utilities influence water-saving advice provided to service users? An Analysis of the Spanish water sector. *Water Resour Manag* 37(8):3299–3318
- Marques RC, Simões P (2020) Revisiting the comparison of public and private water service provision: An empirical study in Portugal. *Water (Switzerland)* 12(5):1477
- Maziotis A, Villegas A, Molinos-Senante M, Sala-Garrido R (2020) Impact of external costs of unplanned interruptions on water company efficiency: Evidence from Chile. *Util Policy* 66:101087
- Mellah T, Ben Amor T (2016) Performance of the Tunisian Water Utility: An input distance function approach. *Util Policy* 38:18–32
- Moeini M, Karimi B, Khorram E (2015) A cross-efficiency approach for evaluating decision making units in presence of undesirable outputs. In H.A. Le Thi et al. (eds.), *Model Comput Optim Inf Syst Manag Sci Adv Intell Syst Comput* 360
- Molinos-Senante M, Donoso G, Sala-Garrido R, Villegas A (2018a) Benchmarking the efficiency of the Chilean water and sewerage companies: a double-bootstrap approach. *Environ Sci Pollut Res* 25:8432–8440
- Molinos-Senante M, Porcher S, Maziotis A (2018b) Productivity change and its drivers for the Chilean water companies: A comparison of full private and concessionary companies. *J Clean Prod* 183:908–916
- Molinos-Senante M, Sala-Garrido R (2015) The impact of privatization approaches on the productivity growth of the water industry: A case study of Chile. *Environ Sci Policy* 50:166–179
- Molinos-Senante M, Villegas A, Maziotis A (2019) Are water tariffs sufficient incentives to reduce water leakages? An empirical approach for Chile. *Util Policy* 61:100971
- Moosavian SF, Borzuei D, Ahmadi A (2022) Cost analysis of water quality assessment using multi-criteria decision-making approach. *Water Resour Manag* 36(12):4843–4862

- Nuru H, Raji P, Manivasagam VS, Sudha A, Raj C (2023) Global sustainable water management: a systematic qualitative review. *Water Resour Manag*, In Press
- Omrani H, Shafaat K, Emrouznejad A (2018) An integrated fuzzy clustering cooperative game data envelopment analysis model with application in hospital efficiency. *Expert Syst Appl* 114:615–628
- Park HS, Jun CH (2009) A simple and fast algorithm for K-medoids clustering. *Expert Syst Appl* 36(2):3336–3341
- Pinto FS, Simoes P, Marques RC (2017) Water services performance: do operational environmental and quality factors account? *Urban Water J* 14(8):773–781
- Saal DS, Parker D, Weyman-Jones T (2007) Determining the contribution of technical efficiency, and scale change to productivity growth in the privatized English and Welsh water and sewerage industry: 1985–2000. *J Prod Anal* 28(1):127–139
- Sala-Garrido R, Maziotis A, Mocholi-Arce M, Molinos-Senante M (2023) Assessing eco-efficiency of wastewater treatment plants: A cross-evaluation strategy. *Sci Total Environ* 900:165839
- Sala-Garrido R, Molinos-Senante M, Mocholi-Arce M (2019) Comparing changes in productivity among private water companies integrating quality of service: A metafrontier approach. *J Clean Prod* 216:597–606
- Samoilenko S, Osei-Bryson K-M (2010) Determining sources of relative inefficiency in heterogeneous samples: Methodology using Cluster Analysis, DEA and Neural Networks. *Eur J Oper Res Soc* 206:479–487
- Sexton TR, Silkman RH, Hogan AJ (1986) Data envelopment analysis: Critique and extensions. *New Dir Prog Eval* 32:73–105
- Sikka V, Luke RD, Ozcan YA (2009) The efficiency of hospital based clusters: evaluating system performance using data envelopment analysis. *Health Care Manag Rev* 34(3):251–261
- SISS (Superintendencia de Servicios Sanitarios) (2021) Management reports about Chilean water companies. Available at: <https://www.siss.gob.cl/586/w3-propertyvalue-6415.html>
- Suárez-Varela M, de los Angeles García-Valiñas M, González-Gómez F, Picazo-Tadeo AJ (2017) Ownership and performance in water services revisited: does private management really outperform public? *Water Resour Manag* 31(8):2355–2373
- Swenson ER, Bastian ND, Nembhard HB (2016) Data analytics in health promotion: health market segmentation and classification of total joint replacement surgery patients. *Expert Syst Appl* 60:118–129
- Tang T, Chen S, Zhao M, Huang W, Luo J (2018) Very large-scale data classification based on K-means clustering and multi-kernel SVM. *Soft Comput* 23:3793–3801
- Tavana M, Toloo M, Aghayi N, Arabmaldar A (2021) A robust cross-efficiency data envelopment analysis model with undesirable outputs. *Expert Syst Appl* 167:114117
- Wang X, Lu Y, Chen C, Yi X, Cui H (2024) Total-factor energy efficiency of ten major global energy-consuming countries. *J Environ Sci (china)* 137:41–52
- Wu J, Chu J, Sun J, Zhu Q, Liang L (2016) Extended secondary goal models for weights selection in DEA cross-efficiency evaluation. *Comput Ind Eng* 93:143–151
- Wu J, Liang L, Song M (2010) Performance based clustering for benchmarking of container ports: An application of dea and cluster analysis technique. *Int J Comput Intell Syst* 3(6):709–722
- Wu J, Liang L, Yang F (2009) Achievement and benchmarking of countries at the Summer Olympics using cross efficiency evaluation method. *Eur J Oper Res* 197:722–730
- Xie L, Chen C, Yu Y (2019) Dynamic assessment of environmental efficiency in Chinese industry: A multiple DEA model with a gini criterion approach. *Sustainability* 11:2294

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.