Efficiency benchmarking of wastewater service providers: an analysis based on the Portuguese case

A. Luís Amaral[a,b,c,d*], Rita Martins[c], Luís C. Dias[c]

[a] Instituto Politécnico de Coimbra, ISEC, Rua Pedro Nunes, 3030-199 Coimbra, Portugal
[b] CEB - Centre of Biological Engineering, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[c] Universidade de Coimbra, CEBER – Centre for Business and Economics Research, Faculdade de Economia, Av. Dias da Silva, 165, 3004-512 Coimbra, Portugal
[d] Instituto de Investigação Aplicada, Laboratório SiSus, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal

*Corresponding author. Email address: lpamaral@isec.pt (A.L. Amaral)

Abstract

Sustainability is increasingly important in the wastewater treatment sector, and benchmarking studies are needed to identify the best practices in using resources efficiently. This work addresses this need by characterizing and benchmarking this sector in Portugal. Data from around 120 service providers, from 2015 to 2019, regarding inputs consumption, desirable and undesirable outputs, are used in a data envelopment analysis to benchmark these operators and to determine improvement potentials. A productivity change analysis revealed an increase from 2015 to 2019, essentially due to an increased technical efficiency. On the other hand, no real technological change seems to have occurred. A positive effect of the service providers dimension and certifications (quality, environmental and energy) on efficiency was found. Service providers with concessionary governance models and urban typology presented the highest efficiencies, whereas treatment plants underuse contributed negatively to efficiency. These results can inform the definition of strategic policies for the water sector, namely concerning the aggregation of small providers. The proposed methodology, based on the use of a comprehensive new set of studied variables and explanatory factors, allowed to determine the best characteristics regarding the techno-economic efficiency of Portuguese service providers, and can potentially be applied to identify role models in other countries.

Keywords: Wastewater treatment service providers; Techno-economic efficiency benchmarking; Data envelopment analysis; Productivity analysis; Explanatory factors

1. Introduction

Sustainability is currently of extreme importance for policy makers, regulators, managers, and the general public. In this respect, environmental protection, as well as technical, economic and social development, are fundamental pillars (UN, 2005; Davidson et al., 2007). Wastewater (WW) treatment systems are no exception and, therefore, must also pursue sustainability goals (Balkema et al., 2002; Molinos-Senante et al., 2016). A growing number of studies focus on WW treatment sustainability regarding techno-economic efficiency, mostly aiming to identify the best practices through benchmarking techniques. However, most of them assess facilities (treatment plants - TP) or treatment systems, rather than service providers (SP). On the other hand, assessing, and comparing SP allows looking for potential efficiency gains within a regional or national scope. Therefore, this study focuses on the WW treatment sector characterization and benchmarking of SP at a country scale. The Portuguese case is analyzed as an example.

One of the most important pillars of the WW treatment sector sustainability relates to its social and financial impact, since costs are paid by consumers through tariffs. Personnel and energy consumption expenditures are the main operational costs (Molinos-Senante et al., 2014; Castellet and Molinos-Senante, 2016), counterbalanced by WW revenues (RevWW). These economic factors are therefore studied in detail in the present work. The energy produced within a TP, using the produced biogas or renewable energy generation systems, is also considered in this study for its potential to reduce the procurement of energy to operate such systems. In addition, concerns with the rational use of clean water sources are growing, due to the increasing seasonal demand (related to more extreme and frequent climatic events) and supply scarcity (associated with the reduction of clean water sources in nature). Thus, the possibility of reusing (treated) WW, either internally or externally (public networks, crop irrigation, etc.), becomes increasingly important (Gonzalez-Serrano et al., 2005; Rodriguez-Garcia et al., 2011). As such, both the treated WW (TWW) and the reused WW (RWW) volumes were incorporated in this study.

Sludge generation is another issue WW treatment systems must deal with (Molinos-Senante et al., 2014; Castellet and Molinos-Senante, 2016). Although some sludge could be used for energetic valorization, e.g. in anaerobic biodigesters or by dehydrated sludge incineration (Tchobanoglous et al., 2003), solely a fraction enters that pathway in Portugal. Therefore, the produced sludge leaving a TP is considered as an undesirable output, carrying costs to the SP that need to pay external entities for sludge treatment and/or deposition.

Identifying the best systems, practices or units using their inputs (resources) efficiently for the generation of desirable outputs is increasingly relevant in sustainability studies (Molinos-Senante et al., 2014; Castellet and Molinos-Senante, 2016; Molinos-Senante et al., 2016; Dong et al., 2017). The most efficient SP, identified through a data envelopment analysis (DEA), can become a reference for the remaining ones. The Slacks-Based Measure (SBM) DEA model (Tone, 2001) is used in this work to compute the efficiency of each SP, as well as the improvement potential in the country as a whole. The analysis considers three inputs (allocated personnel, energy consumption and expenditures), four desirable outputs (RevWW, TWW, RWW and produced energy) and one undesirable output (produced sludge). The volume of collected WW (CWW) is used as a non-discretionary input variable, since SP cannot control how much WW they are required to treat.

WW treatment is a dynamic process and, thus, both the position of a given SP within a benchmark, as well as the efficiency frontier, may change over time (Sala-Garrido et al., 2012; Pointon et al., 2016; Sala-Garrido et al., 2019). Accordingly, a Luenberger-Malmquist (LM) productivity change analysis between 2015 and 2019 (encompassing LM total and technical efficiencies and technological changes) is also performed.

Finally, in order to assess other explanatory factors that may affect the SP efficiency, cross-correlation (CC), ordinary least squares (OLS) and Kruskal-Wallis (KW) analyses were performed, relating the obtained efficiencies with the number of TP managed by each SP, septic tanks and lift stations, WW collection grid length and rehabilitation, energy efficiency (pumping systems), served households, population and TP dimensioning, governance model, TP typology, costs recovery, satisfaction and quality parameters compliance, and quality, environmental and energy certifications.

2. Literature survey on service providers benchmarking

Studies made at the SP level, rather than at the plant scale, tend to focus on drinking water supply, although some studies in the water sector (including WW treatment) can already be found and are next described. Marques et al. (2011) analyzed the implementation of DEA and other benchmarking techniques in more than 50 countries, concluding that over 70% of the regulators used benchmarking tools for assessing the quality of service or to assist in the economic regulation of the water sector.

Concerning WW treatment services, Byrnes et al. (2009) are among the first to study SP performance. Using DEA, they studied the economic efficiency of 56 non-metropolitan SP in the New South Wales and Victoria regions to provide policy recommendations. Nafi et al. (2014) analyzed the cost functions of three SP in the Lyon region (France) to obtain performance indicators for benchmarking urban WW treatment. Pointon et al. (2016) studied the English and Welsh water and WW industry dynamic efficiency, for 10 SP in the 1997-2011 period, using a DEA approach, concluding that the inadequate intertemporal allocation of the inputs is the largest contributor to economic inefficiency.

More recently, Sala-Garrido et al. (2019) have studied the differences between 22 Chilean fully-privatized and concessionary water SP productivity, assessed by the LM directional distance function approach, between 2010 and 2016. Both the drinking water supply and the WW treatment services were studied and a decrease on the fully-privatized productivity was observed, mainly attributed to technical gap regression. A holistic evaluation was undertaken by Walker et al. (2021) to benchmark the sustainability of 12 water sector SP in the United Kingdom over a 6-year period by the Hicks-Moorsteen productivity index. These authors endorsed customer satisfaction and water quality measures as preferable outputs to reflect sustainability and identified economies of scale and scope as drivers for productivity change. Greenhouse gas (GHG) emissions were considered by Molinos-Senante and Maziotis (2021) to compare productivity change of English and Welsh water sector SP, between 2011–2019, using LM. These authors found that the inclusion of GHG emissions decreased productivity, which is important when evaluating water companies for regulatory purposes. Furthermore, water source and treatment level, type of WW treatment facility, density and pumping capacities were found to influence productivity. Pereira and Marques (2022) proposed a two-stage DEA model to determine technical and scale efficiencies of 2160 Brazilian municipal water and sanitation services, observing significant scale inefficiencies. Water source, ownership and geography were found to be significant explanatory factors. These authors further suggested regulation and policy-making debureaucratization, stricter expenditure control, and investment in facilities expansion. Mocholi-Arce et al (2022) applied a two-stage network DEA model to evaluate the cost and operational performance of 189

Chilean water sector SP, between 2010–2018, considering their service quality, noting it was possible to reduce costs and improve operational performance. They also reported public SP outperformed fully private and concessionary providers.

Considering the Portuguese reality, Carvalho and Marques (2014) applied DEA and partial frontier nonparametric methodologies to benchmark 74 SP operating on the water supply and/or WW treatment sectors in Portugal. Their results evidenced economies of vertical integration and scale in the drinking water supply and WW treatment sectors. In subsequent studies, using parametric techniques (Carvalho and Marques, 2015, 2016), economies of scope between water supply and WW treatment services were observed. Caldas et al. (2019) studied, using DEA, the scale economies of 308 Portuguese municipalities, establishing optimal services size. Scale economies were found, regardless of municipality size, with the authors suggesting the formulation of local government policies on a case-by-case basis. Marques and Simões (2020) analyzed the performance of Portuguese private and public water sector SP, in 2011-2015, based on investment effectiveness, quality of service, tariffs and prices. Results showed that the performance of private exceeded public SP. Benefit-of-the-doubt composite indicators formulated with a directional distance function were used by Henriques et al (2020) to benchmark 169 Portuguese WW SP. These authors found that larger scale, investment subsidies and energy production, alongside a concession model and urban typology, improve the quality of service for retail SP. Expanding the previous work, Mergoni et al (2022) benchmarked the environmental performance of 149 Portuguese utilities jointly active in the water supply, WW collection and solid waste management sectors by an analog methodology. Their results highlighted existing improvement potential especially for small and very large units, mostly operating in urban areas.

To the authors' best knowledge, the present work is one of the first benchmark studies, assessing the WW treatment sector from a techno-economic perspective for a long time period, aiming at identifying the potential for inputs (expenditure, personnel and energy) and undesirable outputs (sludge) reduction, and desirable outputs (RevWW, TWW, RWW and produced energy) increase. Indeed, the use of energy consumption and production, reused and revenue WW and sludge production have seldom been used in such studies (Sala-Garrido et al, 2019; Henriques et al, 2020; Molinos-Senante and Maziotis, 2021; Walker et al, 2021), and never simultaneously. Another contribution of this study is the assessment of a comprehensive set of explanatory variables (dimension, techno-economic, environmental, governance, typology and quality). Furthermore, this study directly addresses a country scale perspective, rather than a TP facility, aiming to identify the best practices in the Portuguese reality. Moreover, besides being the first study with such characteristics, the proposed methodology can be used with data from other countries in order to identify their role model SP.

3. Materials and Methods

The methodology used in the current study is presented next. Upon a prior validation step, the collected data is further organized and analyzed in terms of its descriptive statistics. Next, a DEA (SBM model) analysis is carried out to study the SP efficiencies and improvement potentials. A LM productivity analysis is further employed to study the evolution of the productivity factors from 2015 to 2019. Finally, an explanatory factors analysis is performed employing CC, OLS and KW analyses.

3.1 Data collection and validation

The data used in this work was collected from the *Annual Report of Water and Waste Services in Portugal* (RASARP) from 2015 to 2019 made available by *The Water and Waste Services Regulation Authority* (ERSAR, 2016, 2017, 2018, 2019, 2020). SP that did not operate TP or did not present energy or economic data for the WW treatment were excluded. The identification and removal of potential outliers was performed by iteratively removing the SP presenting ratio values for *CWW volume per personnel*, *CWW volume per consumed energy*, *CWW volume per expenditure*, *RevWW volume per CWW volume* and *TWW volume per CWW volume* larger than twice the second highest value, in order to avoid biasing the results. Indeed, such outliers would artificially expand the efficiency frontier, creating unrealistic efficiency targets for the remaining SP. The 117 (2015) to 124 (2019) SP remaining after outliers' removal accounted for more than 2000 TP each year, serving roughly 90% of the population and representing 90.4% of the WW treated in TP.

The expenditure data, regarding WW treatment, was lacking from the RASARP source for a few SP (5 in 2016, 6 in 2017, 5 in 2018 and 4 in 2019), and was therefore estimated from their annual financial reports. For this purpose, total expenditures were obtained by considering the following costs: inventory variation, supplies and external services, personnel, amortizations, provisions, investments, operational costs, financial costs, taxes and other costs. All expenditures were normalized to the fiscal year of 2018, considering the Portuguese inflation rates.

The selected inputs, non-discretionary input, desirable outputs and undesirable output are presented in Table 1. Their total, minimum, maximum, average and standard deviation values are also presented in Table S.1 of the Supplementary material.

Table 1 – Inputs, non-discretionary input, desirable and undesirable outputs.

	Acronym	Significance	Units
Inputs	**Per**	Allocated personnel	(p.e. in one year)
	EnC	Energy consumption	(kWh $year^{-1}$)
	Exp	Total expenditure	(€ $year^{-1}$)
Non-discretionary input	**CWW**	Collected WW	m^3 $year^{-1}$
Desirable outputs	**TWW**	Treated WW	m^3 $year^{-1}$
	RevWW	Revenue WW	m^3 $year^{-1}$
	RWW	Reused WW	m^3 $year^{-1}$
	EnP	Energy production	(kWh $year^{-1}$)
Undesirable output	**Slu**	Sludge generation	(ton $year^{-1}$)

3.2 Explanatory factors

A set of factors can affect the SP efficiency and, therefore, should be considered upon discussing the efficiency results. The parameters presented in Table 2 (dimension, techno-economic, environmental, governance, typology and quality), found in the RASARP database or directly determined, were addressed as possible explanatory factors. Hence, a study was conducted to determine the influence of each of these factors on the efficiency coefficients for energy consumption, allocated personnel, total expenditure, RevWW, TWW, RWW, produced energy and produced sludge.

Table 2 – Studied explanatory factors regarding the efficiency coefficients.

	Acronym	Significance	Units
Discrete	**Gover.**	Governance model (direct management, delegation or concession)	--
	Typol.	TP typology (rural, mostly urban or urban)	--
	Qual. Cert.	Quality certification	--
	Env. Cert.	Environmental certification	--
	Ener. Cert.	Energy certification	--
SP dimension	**CWW**	Volume of collected WW	m^3 $year^{-1}$
	TP	Number of treatment plants	#
	ST	Number of septic tanks	#
	LS	Number of lift stations	#
	Cap.	TP capacity	m^3 $year^{-1}$
	HH	Number of served households	#
	Pop.	Effectively served population	p.e. $year^{-1}$

	GL	WW collection grid length	Km
Non SP dimension	**Recov.**	Costs recovery in percentage	%
	Over.	TP overuse (total treated volume over the TP capacity)	%
	Under.	TP underuse (total treated volume under the TP capacity)	%
	EnE	Energy efficiency of the pumping systems	kWh m^{-3}
	Rehab.	Length of rehabilitated collector pipes per grid length	m Km^{-1}
	Adeq.	TP adequateness (effectively served to dimensioned population)	%
	Comp.	Effluent quality parameters compliance	%
	Satisf.	Satisfactory WW treatment (percentage of population with satisfactory treatment)	%

3.3 Data envelopment analysis (DEA)

The model proposed by Farrell (1957) establishes a best practices frontier composed by the efficient productive units (decision-making units – DMU) in a sample. Such DMUs obtain the maximum possible output from a given set of inputs or use the minimum inputs possible to reach a given output. In the latter case, the (technical) efficiency of a DMU can be determined by calculating the maximum possible proportional reduction in input utilization in order to maintain the same level of production. Based on these ideas, DEA is a non-parametric technique to assess efficiency, which is able to consider a multiplicity of inputs and outputs and does not require assuming a specific production function. In addition, both inputs and outputs can be measured in different units (Cooper et al., 2006).

To determine the efficiency benchmark, consider $k = 1, 2, ..., K$ DMUs, each one using a vector $x^k = (x_1^k, x_2^k, ..., x_N^k)$ of N inputs, to produce a vector $y^k = (y_1^k, y_2^k, ..., y_M^k)$ of M outputs. Let k' be a specific DMU within the sample set. One can define an efficiency score $E\ (x^{k'}, y^{k'})$, assessing the ability of the k' DMU to achieve a pre-established production using the minimum inputs. For a given output vector, defined for each DMU, the objective is to determine to what extent its inputs vector can be minimized. An efficient DMU cannot further reduce inputs consumption, contrary to an inefficient DMU. The efficiency score E for each DMU k' can be calculated by solving the following linear program (Charnes et al., 1978; Cooper et al., 2007):

$$\min(\lambda)\,, s.t.:$$

$$y_m^{k'} \leq \sum_{k=1}^{K} z_k\, y_m^k \qquad m = 1, \dots, M \quad (Eq.1)$$

$$\lambda\, x_n^{k'} \geq \sum_{k=1}^{K} z_k\, x_n^k \quad n = 1, \dots, N \quad (Eq.2)$$

Here, the efficiency benchmark $E\ (x^{k'}, y^{k'}) = \lambda$ (considering the optimal value) indicates whether the k' DMU is efficient (if $\lambda = 1$) or inefficient (if $0 \leq \lambda < 1$).

Furthermore, the difference between λ and 1 can be considered as the potential inputs reduction that can be achieved without changing production.

DEA models such as the Charnes, Cooper and Rhodes – CCR (Charnes et al., 1978), assume constant returns to scale (CRS), with the constraint present in Eq. 3, whereas models as Banker, Charnes and Cooper – BCC (Cooper et al., 2007) consider variable returns to scale by assuming additionally the constraint present in Eq. 4.

$$z_k \geq 0, \qquad k = 1, \dots, K \qquad (Eq. 3)$$

$$\sum_{k=1}^{K} z_k = 1,\ k = 1, \dots, K \quad (Eq. 4)$$

Radial models, as CCR and BCC, determine the overall efficiency of a DMU by assessing its maximum allowed proportional inputs reduction, or outputs increase, considering the efficiency frontier defined by its peers. On the contrary, a non-radial DEA model allows minimizing inputs in different proportions (Molinos-Senante et al., 2014). In non-radial DEA models, λ may have a different value (λ_n) for each input (Färe and Lovell, 1978), modifying the optimization problem function to $\frac{1}{N}\min(\sum_{n=1}^{N} \lambda_n)$ and adding the following restriction to the model:

$$1 \geq \lambda_n \geq 0, \qquad n = 1, \dots, N \quad (Eq. 5)$$

One of the most used non-radial DEA models for determining efficiency benchmarks in TP (Lorenzo-Toja et al., 2015; Castellet and Molinos-Senante, 2016) is the slack-based model (SBM) (Tone, 2001), due to its elasticity in the calculation of inefficiencies for the different DMUs (Thrall, 1996; Cooper et al., 2007). The SBM model considers the non-radial characteristics of the inputs and outputs, which makes it more appropriate to monitor inputs with vague interconnections (Cooper et al., 2007). The SBM model can be employed by solving the following minimization problem:

$$\min\left(\frac{1 - \frac{1}{N}\sum_{n=1}^{N} \frac{s_n^-}{x_n^k}}{1 + \frac{1}{M}\sum_{m=1}^{M} \frac{s_m^+}{y_m^k}}\right), s.t.:$$

$$y_m^k = \sum_{k=1}^{K} z_k\, y_m^k - s_m^+, \quad m = 1, \dots, M \quad (Eq. 6)$$

$$x_n^k = \sum_{k=1}^{K} z_k\, x_n^k + s_n^-, \qquad n = 1, \dots, N \quad (Eq. 7)$$

$$z_k \geq 0;\ s_n^- \geq 0;\ s_m^+ \geq 0, \qquad (Eq. 8, 9, 10)$$

where s_n^- are the input slacks (excess) and s_m^+ the output slacks (shortfall).

3.4 Productivity analysis

The determination of the productivity evolution was based primarily on the calculation of the Luenberger-Malmquist (*LM*) total efficiency change (Chambers et al., 1996). This index represents the productivity of the production point (x_{t+1}, y_{t+1}) relatively to the production point (x_t, y_t). A positive value indicates a productivity increase from period *t* to period *t*+1, whereas a negative value indicates a productivity decrease.

$$LM\ (y_{t+1}, x_{t+1}, y_t, x_t) =$$
$$\frac{1}{2}(\theta^{t+1}(x_t, y_t) - \theta^{t+1}(x_{t+1}, y_{t+1}) + \theta^t(x_t, y_t) - \theta^t(x_{t+1}, y_{t+1})) \quad (Eq. 11)$$

where $\theta_k = \frac{\sum_{i=1}^{M} w_i\, x_{ik}}{\sum_{j=1}^{N} w_j y_{jk}}$ represents the distance function (or technical efficiency) of DMU *k* to the efficiency frontier (Färe et al., 1989, Chambers et al., 1996). Namely, $\theta^t(x_t, y_t)$ is the DMU efficiency at time *t* comparing to its efficiency frontier, $\theta^t(x_{t+1}, y_{t+1})$ is the DMU efficiency at time *t*+1 comparing to the time *t* efficiency frontier, $\theta^{t+1}(x_t, y_t)$ is the DMU efficiency at time *t* comparing to the time *t*+1 efficiency frontier, and $\theta^{t+1}(x_{t+1}, y_{t+1})$ is the DMU efficiency at time *t*+1 comparing to the time *t*+1 efficiency frontier.

The LM index can be decomposed as the sum of the technical efficiency change (*EC*) and technological change (*TC*) indices (Chambers *et al.*, 1996):

$$EC = [\theta^t(x_t, y_t) - \theta^{t+1}(x_{t+1}, y_{t+1})] \quad (Eq. 12)$$

$$TC = \frac{1}{2}[\theta^{t+1}(x_{t+1}, y_{t+1}) - \theta^t(x_{t+1}, y_{t+1}) + \theta^{t+1}(x_t, y_t) - \theta^t(x_t, y_t)] \quad (Eq. 13)$$

The *EC* index evaluates the evolution of the DMU's technical efficiency (increase if above 0 and decrease if below), indicating if the DMU moved closer to (if positive), or away from (if negative), the convex efficiency envelope. The *TC* index evaluates the evolution of technological changes: if positive it denotes technological progress leading to an expansion of the efficiency frontier; if negative it denotes a regression, i.e., a contraction of the efficiency frontier.

Furthermore, the *TC* index can be decomposed in three parts: output biased technological change (OBTC), the bias of technological change caused by output change; input biased technological change (IBTC), the bias of technological change caused by input change; and magnitude of technological change (MATC).

$$OBTC = \frac{1}{2}\left[\left(\theta^{t+1}(x_t, y_t) - \theta^t(x_t, y_t)\right) - (\theta^{t+1}(x_t, y_{t+1}) - \theta^t(x_t, y_{t+1}))\right] \quad (Eq. 14)$$

$$IBTC = \frac{1}{2}\left[\left(\theta^{t+1}(x_t, y_{t+1}) - \theta^t(x_t, y_{t+1})\right) - \left(\theta^{t+1}(x_{t+1}, y_{t+1}) - \theta^t(x_{t+1}, y_{t+1})\right)\right] \quad (Eq. 15)$$

$$MATC = \theta^{t+1}(x_{t+1}, y_{t+1}) - \theta^{t}(x_{t+1}, y_{t+1}) \quad (Eq. 16)$$

The use of TC and EC indices is already established in the water and WW sector literature (Pointon and Matthews, 2016; Molinos-Senante and Maziotis, 2021; Sala-Garrido et al, 2019), and a few works further detail input and/or output based indices (Walker et al, 2021). In the current work, the SBM model was used, with variable returns to scale, for the benchmark analysis. Three different DEA models were used: i) an inputs' oriented SBM (SBM-in) to minimize inputs consumption and undesirable outputs generation; ii) an outputs' oriented SBM (SBM-out) to maximize desirable outputs production; and iii) a non-oriented SBM (SBM-none) simultaneously pursuing both goals. LM productivity indices, based on the SBM model with undesirable outputs, were employed for the productivity analysis, using *MaxDEA 8* software for computations.

3.5 Multivariate statistics

Multivariate statistics were employed in the study of the selected explanatory factors influence on the efficiency coefficients for energy consumption, allocated personnel, total expenditure, RevWW, TWW, RWW, produced energy and produced sludge. Regarding the continuous explanatory factors (CWW, TP, ST, LS, GL, Rehab., Cap., Over., Under., Recov., EnE, HH, Pop., Adeq., Comp. and Satisf.), the p-values of an OLS analysis with the studied efficiency values were determined. Considering the different scale ranges of some explanatory factors, this study was conducted with raw as well as normalized values (both by log-norm and per CWW) in those cases (CWW, TP, ST, LS, GL, Cap., HH and Pop.). With respect to the discrete explanatory factors (Gover., Typol, Qual. Cert., Env. Cert. and Ener. Cert.), a KW analysis was performed between each of these factors group allocation and the studied efficiency coefficients. In both cases a CC analysis was performed among all variables to determine their inner relationships.

The non-parametric KW test was performed to determine the influence of the discrete explanatory factors on the efficiency coefficients of the SP inputs and outputs. This test allows surveying for significant differences in the mean values of a given input or output efficiency with respect to each studied explanatory factor groups. Samples were considered not significantly different from each other for a level of statistical significance (p-value) greater than 0.05, and significantly different otherwise. Matlab 8.3.0 was employed for the KW analysis.

A schematic representation of the benchmarking method is presented in Figure 1.

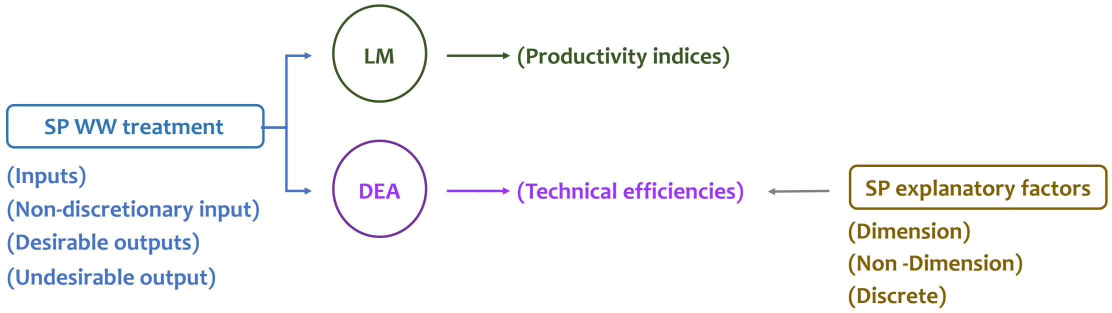

Figure 1 – Schematic representation of the benchmarking methodology.

4. Results and Discussion

4.1 Data characterization and analysis

Figure 2 shows the evolution of the absolute and normalized (per CWW) values for the DEA factors (recall Table 2) concerning the SP analyzed (117 in 2015, 121 in 2016, 119 in 2017, 120 in 2018 and 124 in 2019).

During the analyzed period (2015 to 2019), CWW averaged 761 million m^3 $year^{-1}$, increasing 104 million m^3 $year^{-1}$, and allocated personnel affected to WW treatment, both directly and indirectly, averaged 4280 p.e. $year^{-1}$, with an overall increase in absolute terms (580 p.e. $year^{-1}$), though somewhat stationary when normalized per CWW volume (from 5.85×10^{-6} to 5.81×10^{-6} p.e. m^{-3}). Total expenditure averaged 611 M € $year^{-1}$, with an estimated increase of 17.6 M €, although slightly decreasing in normalized terms from 0.91 to 0.81€ m^{-3}. Finally, total energy consumption averaged 367 GWh $year^{-1}$, increasing 50.2 GWh $year^{-1}$ in absolute terms, while remaining stationary (from 0.51 to 0.50 GWh m^{-3}) in normalized terms. Therefore, although inputs (resources) consumption increased in absolute terms, when normalized per CWW inputs consumption was relatively stable throughout this period.

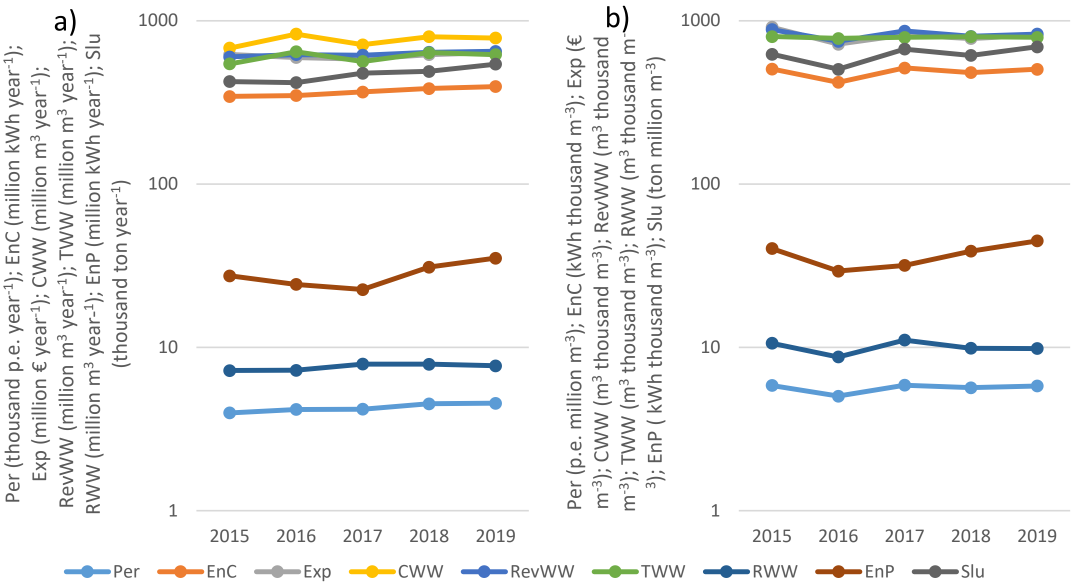

Figure 2 – Evolution of the absolute (a) and normalized (per CWW) (b) values for the inputs, desirable and undesirable outputs, from 2015 to 2019.

From 2015 to 2019, RevWW volume averaged 625 million m^3 $year^{-1}$ and TWW volume averaged 602 million m^3 $year^{-1}$. RevWW and TWW increased 49.0 million m^3 $year^{-1}$ and 75.5 million m^3 $year^{-1}$, respectively. Upon normalization per CWW, RevWW and TWW slightly decreased in that period, respectively from 8.82×10^{-1} m^3 m^{-3} to 8.27×10^{-1} m^3 m^{-3} and from 8.00×10^{-2} m^3 m^{-3} to 7.90×10^{-2} m^3 m^{-3}. These amounts dwarf the RWW volumes averaging 7.59 million m^3 $year^{-1}$, representing 0.99% of the CWW volume (1.26% of the TWW volume). Indeed, solely 20 to 27 SP reused WW from 2015 to 2019. Although the absolute RWW volume has slightly increased 511 thousand m^3, the normalized values (per CWW) slightly decreased from 1.06×10^{-2} m^3 m^{-3} to 9.84×10^{-3} m^3 m^{-3}. Considering energy production, only 14 to 16 SP have reported producing their own energy. The produced energy averaged 28.1 GWh $year^{-1}$, representing around 7.7% of the total energy consumed, increasing 7.80 GWh $year^{-1}$ in absolute terms, and from 4.03×10^{-2} to 4.49×10^{-2} GWh m^{-3} in normalized terms.

Concerning sludge generation, the absolute values averaged 470 thousand tons $year^{-1}$, increasing 11.9 tons $year^{-1}$ in absolute terms, and from 6.23×10^{-4} to 6.91×10^{-4} tons m^{-3} in normalized terms, from 2015 to 2019. Considering this undesirable output, both the absolute and the normalized values slightly increased throughout this period.

4.2 Efficiency analysis according to SBM orientation

The resulting average efficiencies (not weighted by SP size) obtained for the three DEA models are presented in Figure 3.

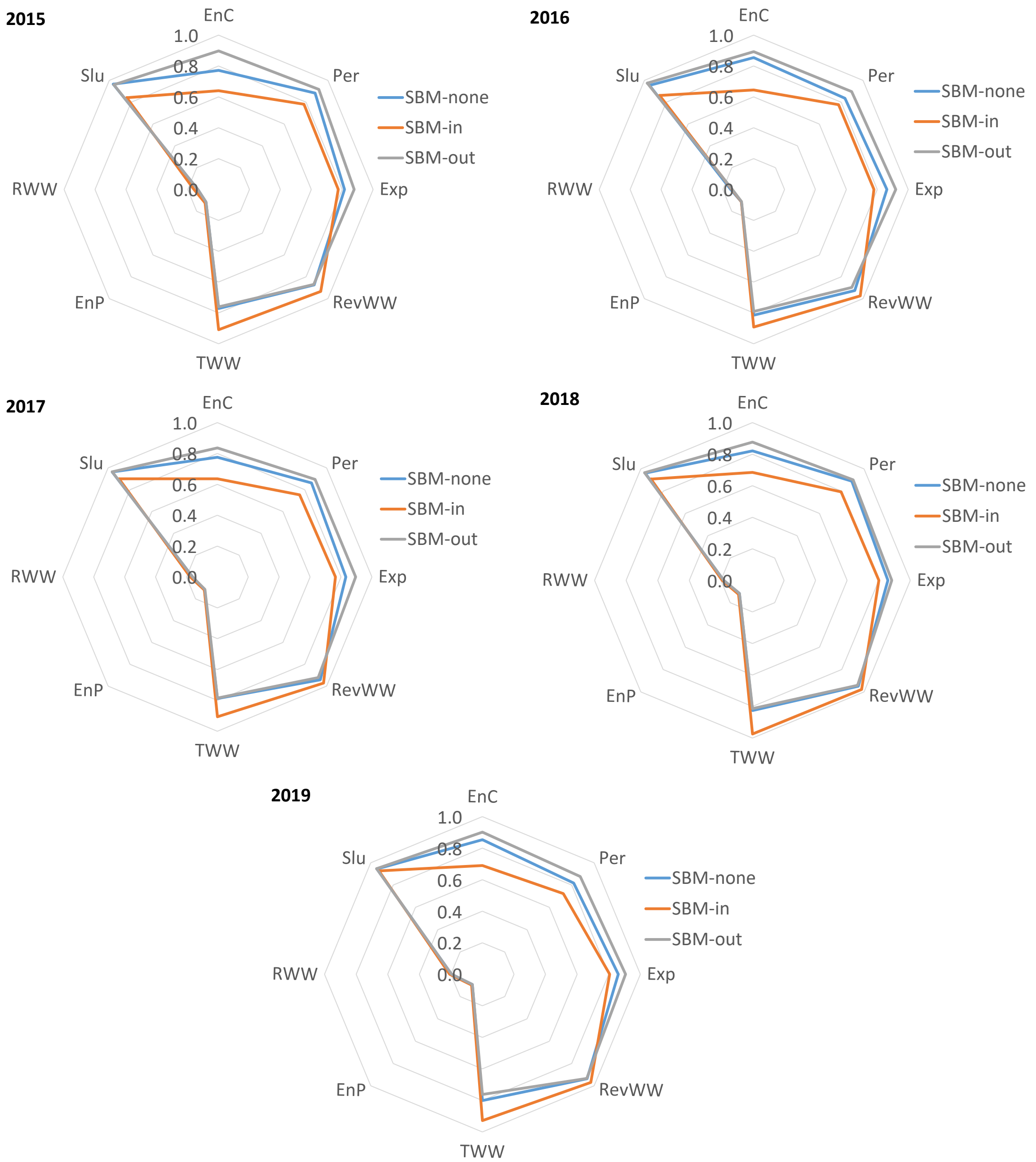

Figure 3 – Non-weighted average values of the SBM technical efficiencies for the inputs, desirable and undesirable outputs, in 2015-2019.

Considering simultaneous minimization of inputs consumption and undesirable outputs production, and maximization of desirable outputs (SBM-none), the obtained non-weighted average efficiencies in 2015-2019 were 0.517 overall, 0.815 for energy consumption, 0.857 for allocated personnel and 0.845 for expenditures. Regarding inputs consumption, efficiencies were relatively similar, although slightly higher for allocated personnel and lower for energy consumption. On the other hand, quite lower efficiencies were obtained for two of the desirable outputs, namely 0.111 for produced energy (the least efficient) and 0.164 for RWW. On the contrary, higher efficiency values were attained for RevWW averaging 0.926, TWW averaging 0.800 and produced sludge averaging 0.957.

When the primary goal is minimizing inputs consumption, the SBM-in results should be analyzed instead. In such case, the non-weighted average efficiencies were 0.659 for energy consumption, 0.766 for allocated personnel and 0.785 for expenditures. Comparing these results with the SBM-none results, it is apparent a decrease in the obtained efficiencies (5-year average), with energy consumption presenting the lowest and expenditures the highest efficiencies.

On the other hand, if the primary goal is the production of desirable outputs, the SBM-out results should be analyzed. In this case, the non-weighted average efficiencies were 0.913 for RevWW, 0.782 for TWW, 0.110 for produced energy and 0.162 for RWW. This represents a decrease (vs. SBM-none) in the efficiencies (5-year average), though still presenting higher efficiencies for RevWW than for TWW, and for RWW than for energy generation. Finally, regarding sludge generation, SBM-in results (minimization of an undesirable output) reveal a 5-year average efficiency decrease (regarding SBM-none), though still presenting, by far, the highest average efficiency (0.962) of the studied variables.

The use of non-weighted averages allows for the characterization of individual SP efficiencies. However, given the quite diverse CWW volumes that each provider accounts for, such averages do not represent the global WW treatment efficiency. For that purpose, a weighted (by CWW) average should be used (Figure 4).

The 2015-2019 weighted average efficiencies obtained for the SBM-none model were 0.840 overall, 0.912 for energy consumption, 0.909 for allocated personnel and 0.895 for expenditures. Considering that larger (in terms of CWW) SP obtained higher efficiencies than smaller (in most cases), this resulted in a clear increase of the weighted average efficiencies, except for generated sludge. Furthermore, a shift could also be observed in the inputs' efficiency, with energy consumption now presenting the highest efficiency and expenditures the lowest. Another noticeable change is the increased (vs. non-weighted) efficiency in generating desirable outputs, averaging 0.976 for RevWW, 0.984 for TWW, 0.793 for produced energy and 0.917 for RWW. The large difference in the latter two can be explained by the large number of small SP, not reusing TWW or generating energy, with their null efficiency lowering the non-weighted efficiency values. And, although the produced sludge still attained large efficiencies, averaging 0.954, it was now surpassed by both RevWW and TWW.

Figure 4 – Weighted average values of the SBM technical efficiencies for the inputs, desirable and undesirable outputs, in 2015-2019.

The SBM-in weighted average efficiencies in 2015-2019 were 0.900 for energy consumption, 0.856 for allocated personnel and 0.874 for expenditures. This represents a small decrease (regarding SBM-none) in the obtained efficiencies (5-year average), presenting slightly higher efficiencies for energy consumption and lower for allocated personnel. On the other hand, the SBM-out results averaged 0.971 for RevWW, 0.981 for TWW, 0.771 for produced energy and 0.911 for RWW, representing a small decrease (vs. SBM-none) in the obtained efficiencies (5-year average), though maintaining the same relative ranking among outputs efficiency. Finally, regarding sludge generation, the analysis of the SBM-in results reveals no significant efficiency change regarding SBM-none (negligible decrease in 5-year average), with the average efficiency (0.951) lagging solely to RevWW and TWW.

4.3 Improvement potential analysis

The potential improvement values for the inputs and undesirable outputs minimization and desirable outputs maximization, i.e. the distance from each SP to the efficiency frontier in a given year (representing the lost potential in that year), determined from 2015 to 2019, are shown in Figure 5 (and Table S.2 of the Supplementary material).

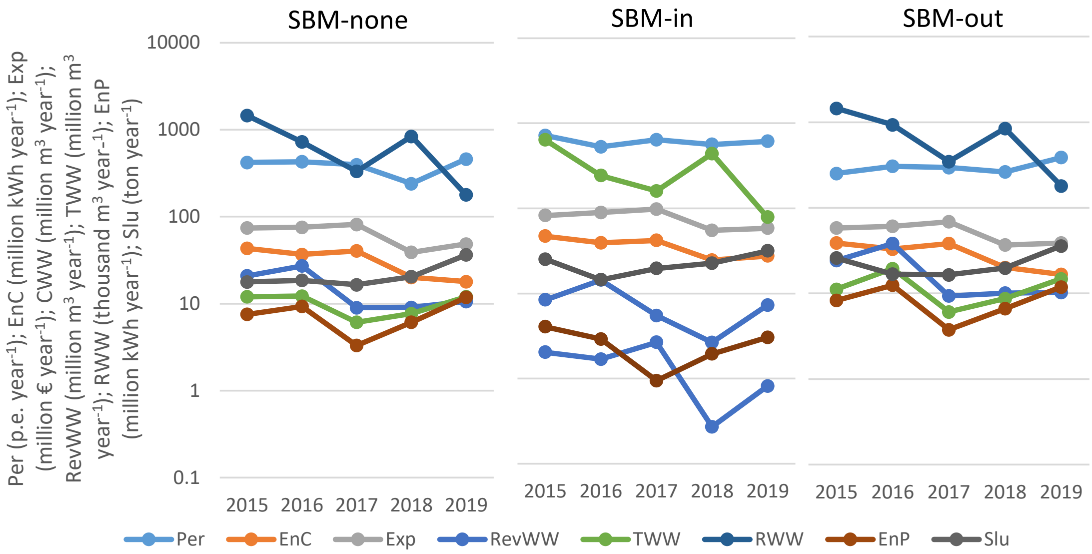

Figure 5 – Improvement potential for the inputs and undesirable outputs minimization and desirable outputs maximization, according to the efficiency frontier, from 2015 to 2019.

Considering model SBM-none, the improvement potential averaged 31.7 GWh year^{-1} for energy consumption, 388.5 p.e. year^{-1} for allocated personnel and 63.8 M € year^{-1} for total expenditure. Considering the SBM-in model instead, the improvement potential for the inputs slightly increases, averaging 36.2 GWh year^{-1} for energy consumption, 612.6 p.e. year^{-1} for allocated personnel and 76.7 M € year^{-1} for total expenditure. Regarding the undesirable produced sludge, the improvement potential averaged 21.9 thousand tons year^{-1}, according to SBM-none. This value amounts to 22.8 thousand tons year^{-1} if one seeks to minimize the generation of undesirable outputs. And, although the produced sludge is inherently dependent on the TWW, other factors are also important, such as the treatment system, use of coagulation/flocculation agents, the dehydration process or even its use for energy generation.

Concerning the generated outputs, and the SBM-none model, the improvement potential averaged 1.53 million m^3 for RevWW, 1.00 million m^3 for TWW, 7.60 GWh year^{-1} for produced energy and 703 thousand m^3 for RWW. On the other hand, if one seeks to maximize the production of desirable outputs (SBM-out), the improvement potential increases to an average 1.85 million m^3 for RevWW, 1.21 million m^3 for TWW, 8.63 GWh year^{-1} for produced energy and 751 thousand m^3 for RWW.

Furthermore, one should note that these potential improvements result from the efficiency frontier set by the surveyed SP and do not consider efficiency gains that the SP in the efficiency frontier could also achieve (thus changing the efficiency frontier and the improvement potentials). Indeed, within the surveyed SP, solely 1.27% of the TWW was reused, with the produced energy accounting for solely 7.62% of the energy consumption (and only in 6 and 27, out of 601 occasions, the SP reused more than 5% of the TWW or produced more than 10% of their energy consumption, respectively). These results are in sharp contrast with the 30% self-produced energy (by 2030) target defined by the Portuguese Government for SP in Portugal (Portuguese Secretary of State for the Environment, 2022). Considering the average energy production in 2015-2019, this would demand an increase in the produced energy of roughly 110.6 GWh year^{-1}. Analogously, a target of reusing 30% of the WW in 2030 would represent the need to increment RWW by roughly 157.8 million m^3 year^{-1}.

4.4 Evolution of the productivity indices from 2015 to 2019

The SP efficiency evolution was determined by the use of LM productivity indices, based on the SBM model with undesirable outputs. Considering the simultaneous minimization of the inputs' consumption and undesirable outputs generation, and maximization of the desirable outputs, the SBM-none approach was used. The resulting total efficiency (LM), technical efficiency (EC) and technological (TC) change indices are presented in Figure 6.

Considering the total (LM) efficiency change, 50.6% of the SP increased their total efficiency (comparing 2019 to 2015) against 49.4% that decreased. Furthermore, in all adjacent years, the SP increasing total efficiency were above 50%. Regarding the technical (EC) efficiency change, again the number of SP that increased (approaching the efficiency frontier) its technical efficiency (37.1%,) was larger than the ones that decreased (32.6%). On the other hand, a larger number of SP (53.9%) saw its technological efficiency change (TC) decrease, against 46.1% that saw an increase. However, this trend was not homogenous in all adjacent years, with opposite results appearing from 2015 to 2016 and from 2018 to 2019. Considering all the above, it seems that the efficiency frontier did not expand from 2015 to 2019 but rather slightly contracted (representing a possible technological regression) with a larger number of SP closing their gap to the efficiency frontier as well as gaining net (total) efficiency.

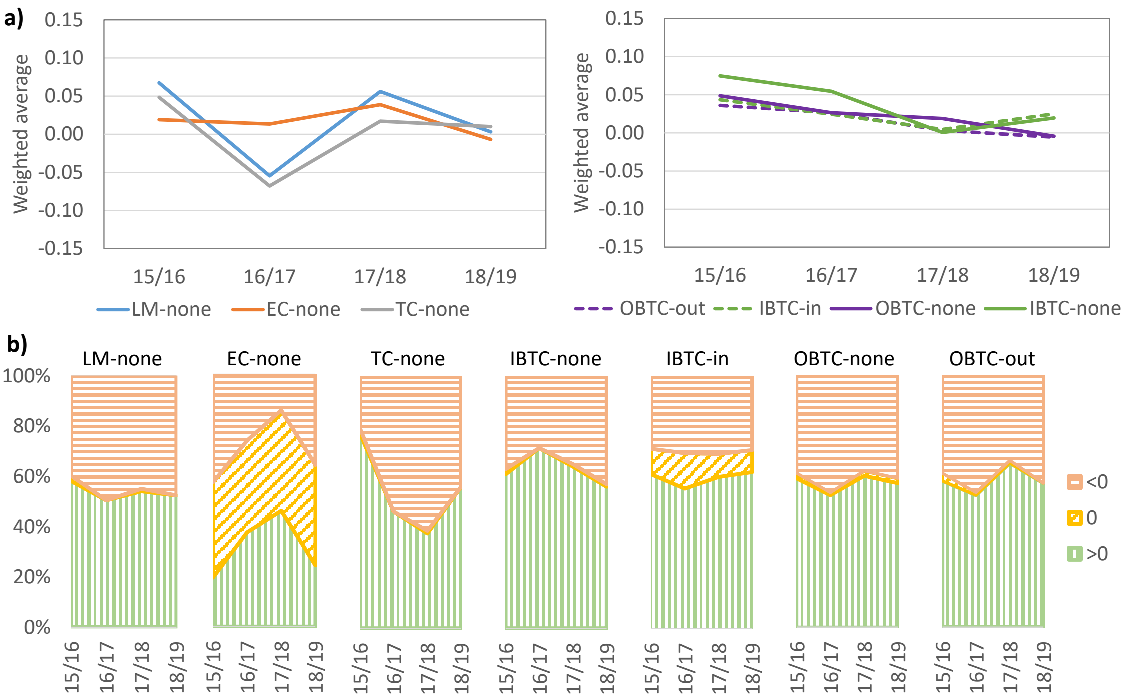

Figure 6 – Weighted average (a) and percentage of SP with values larger than, equal to, and smaller than 0 (b), for LM-none, EC-none, TC-none, OBTC-none, IBTC-none, OBTC-out and IBTC-in, from 2015 to 2019.

Analyzing the weighted average of the efficiency change indices, an overall positive evolution was apparent in the SP total efficiency as well as towards the efficiency frontier (0.05 for LM and 0.06 for EC, comparing 2019 to 2015), confirming the prior analysis. However, no significant evolution of the efficiency frontier was observed (TC of -0.01), pointing towards no real technological progress or regress, given that the larger SP (in terms of TWW) tendentially presented higher TC values than the smaller SP.

Addressing the inputs minimization efficiency change, and considering the SBM-none model, 59.6% of the SP saw an increase in IBTC-none. Furthermore, in all adjacent years, over 50% of the SP increased IBTC efficiency. In accordance, when the weighted average was analyzed, a (slight) positive IBTC-none value (0.02) was observed from 2015 to 2019. On the other hand, if the primary goal is minimizing inputs consumption, the SBM-in approach should be used. In this case, a larger number (47.2%) of SP saw an increase in the IBTC-in values in all adjacent years. Again, the weighted average of the ITBC-in was slightly positive (0.02), representing a minimal increase in inputs efficiency in this period.

Regarding the outputs' maximization efficiency change, and considering SBM-none, 61.8% of the SP saw an increase in OTBC-none. Furthermore, in all adjacent years, the SP increasing OBTC efficiency were above 50%. However, a marginally negative (-0.01) OTBC-none weighted average was obtained, since larger SP tendentially presented lower OTBC-none values than smaller SP. On the other hand, if the primary goal is the production of desirable outputs, the SBM-out approach should be employed. In this case, and in accordance with the SBM-none approach, the number of SP that saw an increase (57.3%) in OTBC-out from 2015 to 2019 surpassed the SP that saw a decrease (41.6%), which also happened in all adjacent years. On the other hand, the OTBC-out weighted average was slightly negative (-0.02), again due to larger SP tendentially presenting lower OTBC-out values than smaller SP. These results lead to a minimal decrease of the desirable outputs efficiency from 2015 to 2019.

4.5 Analysis of explanatory factors

A set of factors may affect the SP efficiency. Table 3 presents the results of the R and p-values of an OLS analysis between the continuous explanatory factors introduced in Section 3.2 and the efficiencies obtained in the DEA, namely the SBM-none efficiency coefficients obtained for energy consumption, allocated personnel, total expenditure, RevWW. TWW, RWW, produced energy and produced sludge.

The CC performed within the set of explanatory factors (recall Table 2) configuring the SP dimension (CWW, TP, ST, LS, Cap. HH, Pop. and GL) provided for high correlation coefficients (R), as expected, between CWW and the remaining factors ($R>0.8$ for the LS and GL (in log-norm terms), and $R>0.9$ for the TP, Cap., HH, and Pop.), except for ST. As expected, the importance of the set of these explanatory factors, regarding the studied efficiencies, was, for the most, comparable. The lowest correlation coefficient was obtained for the ST ($R=0.029$), given that lower volumes of generated WW can be more easily dealt with by the sole use of ST, rather than requiring the use of TP.

Table 3 – Obtained R and p-values for the continuous explanatory factors regarding the efficiency coefficients.

	EnC	Per.	Exp.	RevWW	TWW	EnP	RWW	Slu.
CWW	**0.181** ^{b***}	**0.194** ^{b***}	**0.176** ^{b***}	**0.113** ^{b**}	**0.277** ^{b***}	**0.516** ^{b***}	**0.635** ^{b***}	0.034
TP	**–0.137** ^{c***}	**–0.109** ^{c**}	**–0.124** ^{c**}	–0.078	**0.224** ^{b***}	**0.357** ^{a***}	**0.336** ^{a***}	–0.048
ST	**0.086** ^{a*}	**–0.125** ^{c**}	0.044	–0.056	–0.040	**–0.088** ^{c*}	**–0.176** ^{c***}	0.049
LS	**–0.314** ^{c***}	**–0.123** ^{c**}	**–0.236** ^{c***}	0.072	**–0.173** ^{c***}	**0.341** ^{a***}	**0.460** ^{b***}	–0.069
Cap.	**0.253** ^{b***}	**0.157** ^{b*}	**–0.236** ^{c**}	0.100	**0.539** ^{b***}	**0.463** ^{b***}	**0.524** ^{b***}	0.115
HH	**0.165** ^{b***}	**–0.167** ^{c***}	**0.165** ^{b***}	**0.187** ^{b***}	**0.305** ^{b***}	**0.541** ^{b***}	**0.648** ^{b***}	0.032
Pop.	**0.111** ^{a*}	**0.175** ^{b***}	**0.101** ^{a*}	0.085	**0.376** ^{b***}	**0.444** ^{a***}	**0.512** ^{b***}	**–0.117** ^{b**}
GL	**–0.202** ^{c***}	**–0.160** ^{c***}	**–0.233** ^{c***}	**0.091** ^{b*}	**0.260** ^{b***}	**0.341** ^{a***}	**0.501** ^{b***}	0.037
Recov.	**0.122** **	**0.129** **	**0.222** ***	**0.172** ***	**0.201** ***	**0.237** ***	**0.324** ***	–0.022
Over.	**0.215** **	**0.192** *	**0.209** **	–0.046	**0.171** *	–0.099	0.025	0.069
Under.	**–0.240** **	**–0.268** ***	**–0.235** **	0.068	–0.057	0.027	0.019	–0.079
EnE	0.022	0.100	0.054	0.066	–0.104	**–0.224** ***	**–0.292** ***	**0.160** *
Rehab.	–0.038	–0.012	0.046	–0.056	–0.052	–0.078	–0.075	0.004
Adeq.	–0.035	**0.117** **	0.026	–0.067	**0.098** *	**0.096** *	**0.083** *	0.010
Comp.	0.075	0.046	0.085	**0.114** *	**0.246** ***	**0.271** ***	**0.382** ***	**–0.121** **
Satisf.	0.077	0.027	0.084	**0.148** **	**0.252** ***	**0.255** ***	**0.352** ***	**–0.121** **

^{a} Raw values; ^{b} Log-norm values; ^{c} Values normalized by CWW; * p-value<0.05: ** p-value<0.01; *** p-value<0.001

Overall, the log-norm CWW was found to be an explanatory factor (p-value < 0.05) for all studied efficiency coefficients except for the undesirable output – sludge (similarly to the GL and HH). Indeed, as the SP CWW increased, higher inputs minimization and outputs maximization efficiencies were obtained. The primal importance of the normalized (per CWW) LS and GL for the inputs efficiency (and mainly for the EnC and Exp.) is also noteworthy. Indeed, an increase in the number of lift stations, and in the collectors' grid length, was found to negatively influence the energy consumption and expenditures SP efficiency. Alongside the log-norm CWW, also the normalized HH and GL (both per CWW) and log-norm Pop. could be established as explanatory factors for the personnel efficiency. An increase in the CWW volume and log normalized served population seems to positively affect the efficiency, but the opposite occurs for the normalized (per CWW) collectors grid length and number of households.

Regarding RevWW and TWW, the log-norm HH is among the most decisive explanatory factors, with larger number of served households positively influencing these outputs efficiencies. In fact, the number of households was the main explanatory factor regarding the RevWW and TWW (in the later alongside the served population and TP capacity). As for the above inputs and desirable outputs efficiencies (with the exception of RevWW), the energy production and RWW efficiencies were found to have multiple explanatory factors. In both cases, however, the CWW and the number of served households (both in log-norm terms) seem to be among the most important. Care should be taken, however, in establishing definitive conclusions given the relative low number of SP that produce energy and/or reuse WW. Finally, regarding the Slu. efficiency, within this set of explanatory factors, solely the log-norm Pop. presented a p-value under 0.05, with the sludge efficiency increasing with the decrease of the served population, which can be explained by larger populations producing higher WW volumes requiring treatment in TP instead of ST.

Our results align with the literature, confirming the importance of SP dimension explanatory factors to its efficiency, emphasizing the treated and/or collected WW. Indeed, Byrnes et al (2009), Walker et al (2021) and Pereira and Marques (2022) had already found economies of scale for SP up to an 'optimal' size, and (dis)economies of scale thereafter, which also applies to the Portuguese reality (Carvalho and Marques, 2014, 2015, 2016; Caldas et al, 2019; Henriques et al, 2020; Mergoni et al, 2022). Byrnes et al. (2009) and Molinos-Senante and Maziotis (2021) also highlighted the positive relevance of the collectors' grid and connections to efficiency. Also, economies of vertical integration, output density and scope have been found by Carvalho and Marques (2014, 2015, 2016).

Regarding the continuous explanatory factors not directly related to the SP dimension, a strong inner correlation (R=0.939) occurred solely between the Comp. and Satisf., which could be expected given that both represent quality parameters (of the TWW and of the service as perceived by the population).

Similarly to the explanatory factors configuring the SP dimension, a p-value under 0.05 could be found regarding the Recov. for all inputs minimization and desirable outputs maximization efficiencies. Indeed, these efficiencies rose alongside the SP costs recovery. High inputs and desirable outputs efficiencies can both affect and be affected by costs recovery. Lower demands for overall expenditures (encompassing energy and personnel expenditures), and higher revenue WW and produced energy, enable better costs recovery. On the other hand, higher costs recovery may lead the SP to invest on improved treatment technology and equipment and mains construction and rehabilitation. No significant differences could be found regarding the SP costs recovery for the sludge minimization efficiency.

Neither a decreased WW quality, nor a less than satisfactory treatment (highly correlated as reported above), were found to be trade-offs for inputs minimization and desirable outputs maximization efficiencies. Indeed, the desirable outputs efficiencies rose alongside the increase of the quality compliance and satisfactory treatment, whilst no significant p-values could be found regarding the inputs' efficiency. A synergistic effect could be observed between the Comp. and Satisf, on one hand, and EnP and RWW on the other. Indeed, as the amount of removed pollution increases, so generally occurs to the produced biomass (excess sludge), which can be used for energy production (via anaerobic digestion, waste energy valorization, etc.). On the other hand, higher treated WW quality fosters its reuse, which could be behind the positive influence found for the RWW. Regarding the sludge minimization efficiency, the efficiency increase is accompanied by a decrease of the quality parameters. Again, this could be due to the increase of produced biomass.

Both TP over and underuse could be inferred as explanatory factors for the inputs minimization efficiency. However, contrary to the expected negative influence of TP underuse in the inputs efficiencies, TP overuse seemed to have a positive effect. As the CC analysis found no strong negative correlation between the TP overuse and the quality parameters ($R=0.215$ for Comp. and $R=0.198$ for Satisf.), the TP may present an increased ability to sustain WW loads higher than their stated capacities. On the other hand, the TP over and underuse do not appear as explanatory factors for both the desirable outputs maximization and undesirable outputs minimization efficiencies (except for the Over., relating to the TWW). Furthermore, the SP served to dimensioned population adequateness, although being slightly positively correlated with the Over. ($R=0.295$) and negatively correlated with the Under. ($R=-0.266$) does not appear to be an explanatory factor for the inputs efficiencies (as well as for the undesirable output), except for the personnel to some extent. However, it was found to be so regarding the desirable outputs efficiencies (except for RevWW).

The length of rehabilitated collector pipes per grid length (Rehab.) seems to not significantly affect any of the efficiencies, whereas the energy efficiency of the pumping systems seems to affect the EnP, RWW and sludge efficiencies. However, when considering the p-values obtained for the energy production, RWW and sludge efficiencies of other explanatory factors, the EnE importance seems dimmed.

Table 4 presents the results of the KW analysis between the discrete explanatory factors introduced in Section 3.2 and the efficiencies obtained for the factors considered in the DEA. The studied governance groups (Gover.) were direct management (1), delegation (2) and concession (3). TP typology groups (Typol.) were rural (1), mostly urban (2) and urban (3). Regarding the quality, environmental and energy certification, two groups (absence – 0 and presence – 1) were considered for each certification.

Table 4 – Obtained p-values for the discrete explanatory factors regarding the efficiency coefficients.

	Group	EnC	Per.	Exp.	RevWW	TWW	EnP	RWW	Slu.
Gover.	1 *vs.* 2	NS	NS	NS	NS	***	***	***	NS
	1 *vs.* 3	**	**	**	NS	***	***	***	*
	2 *vs.* 3	NS	NS	NS	NS	NS	***	***	NS
Typol.	1 *vs.* 2	NS	NS	NS	NS	***	***	***	***
	1 *vs.* 3	**	***	**	NS	***	***	***	NS
	2 *vs.* 3	**	***	*	NS	NS	***	***	NS
Qual. Cert.	0 *vs.* 1	***	*	*	NS	***	***	***	NS
Env. Cert.	0 *vs.* 1	***	**	**	NS	***	***	***	NS
Ener. Cert.	0 *vs.* 1	**	**	***	**	**	***	***	NS

NS – Not significant; * p-value<0.05: ** p-value<0.01; *** p-value<0.001

The SP quality, environmental and energy certifications influence in the studied efficiency coefficients was also assessed. In all cases, and for all inputs and desirable outputs, except for the quality and environmental certifications for the RevWW efficiency, the above certifications can be considered as possible explanatory factors (p-values below 0.05), with the SP presenting such certifications attaining higher efficiencies. Furthermore, regarding the energy production and RWW, the obtained p-values indicate these parameters are among the most influential ones, alongside the SP dimension parameters. On the other hand, with respect to the sludge efficiency, these certifications seem not to be explanatory factors. Again, it is also important to note that the presence of such certifications is positively correlated with the SP dimension. Furthermore, and with respect to the energy certification, R values of 0.634 and 0.462 were found with the LS and TP, respectively, stressing the importance of adequate energy practices for these energy intensive installations.

The governance model was found to be related with the inputs minimization and outputs maximization efficiencies. Regarding inputs minimization, the SP within the concession model tend to obtain higher efficiency values, and the direct management the lowest. On the other hand, the governance model was not found to relate with the RevWW efficiency, contrary to the TWW where the direct management resulted in lower efficiency values. With respect to both EnP and RWW efficiencies, a significant distinction could be found between the three models, again with the direct management presenting the lowest values and the concession model the highest. Finally, the governance model was also found to relate with the Slu. efficiency, with the SP within the concession model presenting the lowest efficiency values, and within the direct management the highest.

Published literature presents mixed results concerning the relevance of the governance model for the SP efficiency, depending on the studied country, period and efficiency parameter. Whereas Byrnes et al (2009), Sala-Garrido et al (2019), Henriques et al (2020) and Marques and Simões (2020) determined a positive influence of private and/or concessionary SP, Mocholi-Arce et al (2022) and Pereira and Marques (2022) noted higher performances for public SP and/or under direct management.

The typology of the TP was also studied, with the SP within the urban TP group being found significantly different from the remaining ones in terms of inputs minimization efficiencies and presenting the highest efficiencies for all inputs (EnC, Per. and Exp.). On the other hand, the typology did not influence the RevWW efficiency, contrary to the TWW where the rural TP resulted in lower efficiencies. With respect to both the produced energy and RWW efficiencies, a significant distinction could be found between the three typologies, again with the rural presenting the lower values and the urban the highest. Finally, the TP typology was related to the sludge efficiency, with the SP having mostly urban TP presenting the lowest efficiency values. Furthermore, the SP within the direct management model were overwhelmingly of rural typology, unlike the delegation and the concession models mainly of the mostly urban typology (R=0.511 between TP typology and governance model).

The obtained TP typology results are in accordance with the ones obtained by Henriques et al (2020) for the Portuguese WW SP in 2018. On the other hand, the results found by Mergoni et al (2022), although seemingly contradicting the current results (positive relationship with environmental pressure for rural areas), are circumscribed to a much narrow set of studied variables (water losses, structural collapses, gas emission and recycled waste) and not directly related to WW, SP range (solely SP incorporating water, WW and solid waste treatment and a single year (2018).

Possible explanations for the above results may reside on the fact that:

i) the governance model and TP typology correlated with the SP dimension parameters (R=0.737 for both the log-norm of CWW and HH, with the typology, and R=0.590 between the log-norm CWW and the governance model). The SP dimension increased from rural, to mostly urban and to urban typologies, and from direct management to delegation and to concession. Indeed, the SP dimension could be a driving force behind the governance model and TP typology relationships with the studied efficiencies.

ii) the quality parameters values (explanatory factors for sludge efficiency) increased from direct management, to delegation and to concession model, as well as from rural, to mostly urban and to urban typology (R=0.401 and R=0.384 for the governance model, R=0.363 and R=0.354 for the TP typology, respectively for quality compliance and satisfactory treatment). Again, the quality parameters values could be an explanation for the relationships of the governance model and TP typology with the efficiencies.

iii) the SP within the direct management model, and of rural typology, seldom present environmental or energy certifications, whereas the presence of quality certifications (all explanatory factors) increased from direct management, to delegation and to concession model, as well as from rural, to mostly urban and to urban typology (R=0.570, R=0.462 and R=0.321 for the governance model, R=0.492, R=0.351 and R=0.233, for the typology, respectively for the environmental, quality and energy certifications).

5. Conclusions

In this study the collected, revenue, treated and reused WW volumes, sludge generation, allocated personnel, expenditure and energy consumption and production data, from 2015 to 2019, were used to benchmark the techno-economic efficiency of the Portuguese WW sector.

Simultaneous minimization of inputs used and sludge production, and maximization of desirable outputs, led to an overall weighted average efficiency of 0.840. Among inputs, energy consumption obtained the highest efficiency, closely followed by allocated personnel and expenditures. Regarding the desirable outputs, the treated WW presented the highest efficiency, closely followed by revenue WW, produced sludge (undesirable output) and reused WW with the produced energy falling short of the above. Strict inputs minimization and desirable outputs maximization were analyzed and accounted for other improvement potentials scenarios. These results do not consider possible efficiency gains that SP in the efficiency frontier could also undertake. Indeed, within the surveyed SP, solely 1.27% of the treated WW was reused with the produced energy attaining solely 7.62% of the energy consumption (against the 30% threshold for these OI, expected to be set for Portuguese goals for 2030).

Most SP increased their efficiency from 2015 to 2019, essentially by approaching the efficiency frontier, leading to a 5% weighted overall efficiency increase and 6% in the technical efficiency. On the other hand, most SP saw their technological efficiency decrease, though the corresponding weighted average pointed towards no real technological change. Furthermore, a minimal increase in the inputs' consumption efficiency could also be inferred in this period, as well as a minimal decrease on the efficiency of desirable outputs generation.

The most relevant explanatory factors affecting efficiency were related to SP dimension. Collected WW volume, TP capacity, served population and number of served households had a positive impact on most inputs and desirable outputs efficiencies, whereas collectors grid length and number of TP and lift stations exerted a negative influence on the inputs and a positive on the desirable outputs' efficiencies. Regarding the particular case of the revenue WW, only the number of served households, collected WW volume and collectors grid length (to some extent) were influential. Finally, regarding sludge efficiency, solely the served population was found to exert a (negative) influence.

Costs recovery was found to be positively associated to the inputs and desirable outputs efficiencies. A synergistic effect could be observed between the WW quality, satisfactory WW, energy production and reused WW, although a possible trade-off may exist between the quality parameters and sludge efficiency. TP underuse and overuse analysis suggests that the studied TP may present an increased ability to sustain higher WW loads than their stated capacities. The SP quality, environmental and energy certifications exerted a positive influence over the inputs and desirable outputs efficiencies. The importance of adequate energy practices for these energy intensive installations was also stressed. SP within the concession model led to higher efficiency values of inputs and desirable outputs efficiencies, and direct management to lower efficiencies. Likewise, SP with urban TP presented the highest efficiencies for the inputs and most desirable outputs efficiencies, and rural TP the lowest. Yet, SP dimension, WW quality and certification could be driving forces behind the governance model and TP typology relationships with the studied efficiencies.

Considering the above, the SP techno-economic efficiency seems to benefit from large SP, policy of compliance to environmental, quality and energy certification, and a strict control of the number of lift stations, collectors grid length and TP underuse. Given that governance models and typology were found to correlate with each other and with SP dimension, quality parameter values and certifications, the concession model and urban typology resulted in higher efficiency values, opposite to the direct management model and rural typology. That being the case, benefits will arise also regarding costs recovery, WW quality compliance and satisfactory WW treatment which, far from being a trade-off of improved efficiencies, are rather a positive consequence in most cases. Indeed, the sole exception found was a possible trade-off between sludge efficiency and WW quality compliance.

These results reinforce, and complement, the recommendations of the Portuguese Strategic Plan for Water Supply and Wastewater and Stormwater Management – PENSAARP 2030 (Portuguese Secretary of State for the Environment, 2022), regarding the aggregation of smaller SP to form larger entities, pursue of WW quality compliance and certification practices, increased control of expenditures (or costs recovery), energy and allocated personnel, and environmental (energy, water) sustainability. Indeed, this study allowed identifying the main factors that should be taken into consideration for each of the above objectives. Furthermore, considering the governance model, this study may also be useful regarding possible reorganization of the Portuguese WW sector, stressing the strong and weak points of the current SP models, pointing towards the adoption of more suitable management practices by direct management SP (mostly municipalities). This conclusion is also aligned with PENSAARP 2030, which states the need for the adoption of autonomous management accounting and the creation of supporting mechanisms for smaller SP aggregation and business management restructuring.

Credit author statement

A. Luís Amaral, Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing.

Rita Martins, Conceptualization, Writing – review & editing, Supervision.

Luís C. Dias, Conceptualization, Writing – review & editing, Supervision.

Declaration of competing interest

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

Acknowledgements

This work has been funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/05037/2020.

References

Balkema, A.J., Preisig, H.A., Otterpohl, R., Lambert, F.J., 2002. Indicators for the sustainability assessment of wastewater treatment systems. *Urban Water*, 4(2): 153–161. https://doi.org/10.1016/S1462-0758(02)00014-6

Byrnes, J., Crase, L., Dollery, B., Villano, R., 2009. An analysis of the relative efficiency of wastewater utilities in nonmetropolitan New South Wales and Victoria. *Australasian Journal of Regional Studies*, 15(2): 153–169.

Caldas, P., Ferreira, D., Dollery, B., Marques R. 2019. Are there scale economies in urban waste and wastewater municipal services? A non-radial input-oriented model applied to the Portuguese local government. *Journal of Cleaner Production*, 219: 531–539. https://doi.org/10.1016/j.jclepro.2019.02.076

Carvalho, P., Marques, R.C., 2014. Computing economies of vertical integration, economies of scope and economies of scale using partial frontier nonparametric methods. *European Journal of Operational Research*, 234: 292–307. https://doi.org/10.1016/j.ejor.2013.09.022

Carvalho, P., Marques, R.C., 2015. Estimating size and scope economies in the Portuguese water sector using the most appropriate functional form. *The Engineering Economist*, 60:109–137. https://doi.org/10.1080/0013791X.2013.873507

Carvalho, P., Marques, R.C., 2016. Estimating size and scope economies in the Portuguese water sector using the Bayesian stochastic frontier analysis. *Science of the Total Environment*, 544: 574–586. https://doi.org/10.1016/j.scitotenv.2015.11.169

Castellet, L., Molinos-Senante, M., 2016. Efficiency assessment of wastewater treatment plants: A data envelopment analysis approach integrating technical, economic, and environmental issues. *Journal of Environmental Management*, 167: 160–166. https://doi.org/10.1016/j.jenvman.2015.11.037

Chambers, R.G., Färe, R., Grosskopf, S., 1996. Productivity growth in APEC countries. *Pacific Economic Review*, 1(3): 181-190. https://doi.org/10.1111/j.1468-0106.1996.tb00184.x

Chambers, R.G., Chung, Y., Färe, R., 1998. Profit, directional distance functions, and nerlovian efficiency. *Journal of optimization theory and applications*, 98(2): 351–364. https://doi.org/10.1023/A:1022637501082

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6): 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Cooper, W.W., Seiford, L.M., Tone, K., 2006. *Introduction to Data Envelopment Analysis and Its Uses: With DEA-Solver Software and References.* New York: Springer. https://doi.org/10.1007/0-387-29122-9

Cooper, W.W., Seiford, L.M., Tone, K., 2007. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. New York: Springer. http://dx.doi.org/10.1057/palgrave.jors.2601257

Davidson, C.I., Matthews, H.S., Hendrickson, C.T., Bridges, M.W., Allenby, B.R., Crittenden, J.C., Chen, Y., Williams, E., Allen, D.T., Murphy, C.F., Austin, S., 2007. Adding sustainability to the engineer's toolbox: A challenge for engineering educators. *Environmental Science and Technology*, 41(14): 4847–4850. https://doi.org/10.1021/es072578f

Dong, X., Zhang, X., Zeng, S., 2017. Measuring and explaining eco-efficiencies of wastewater treatment plants in China: An uncertainty analysis perspective. *Water Research*, 112: 195–207. https://doi.org/10.1016/j.watres.2017.01.026

ERSAR – Entidade Reguladora dos Serviços de Águas e Resíduos (Regulatory Body for Water and Waste Services), 2016. *Caracterização do setor das águas residuais: Dados e indicadores do ciclo de avaliação da qualidade do serviço prestado aos utilizadores relativos a 2015*. http://www.ersar.pt/pt/setor/factos-e-numeros/dados-de-base

ERSAR – Entidade Reguladora dos Serviços de Águas e Resíduos, 2017. *Caracterização do setor das águas residuais: Dados e indicadores do ciclo de avaliação da qualidade do serviço prestado aos utilizadores relativos a 2016*. http://www.ersar.pt/pt/setor/factos-e-numeros/dados-de-base

ERSAR – Entidade Reguladora dos Serviços de Águas e Resíduos, 2018. *Caracterização do setor das águas residuais: Dados e indicadores do ciclo de avaliação da qualidade do serviço prestado aos utilizadores relativos a 2017*. http://www.ersar.pt/pt/setor/factos-e-numeros/dados-de-base

ERSAR – Entidade Reguladora dos Serviços de Águas e Resíduos, 2019. *Caracterização do setor das águas residuais: Dados e indicadores do ciclo de avaliação da qualidade do serviço prestado aos utilizadores relativos a 2018*. http://www.ersar.pt/pt/setor/factos-e-numeros/dados-de-base

ERSAR – Entidade Reguladora dos Serviços de Águas e Resíduos, 2020. *Caracterização do setor das águas residuais: Dados e indicadores do ciclo de avaliação da qualidade do serviço* *prestado aos utilizadores relativos a 2019*. http://www.ersar.pt/pt/setor/factos-e-numeros/dados-de-base

Färe, R., Lovell, C.A.K., 1978. Measuring the technical efficiency of production. *Journal of Economic Theory*, 19(1): 150–162. https://doi.org/10.1016/0022-0531(78)90060-1

Färe, R., Grosskopf, S., Weber, W.L., 1989. Measuring school district performance. *Public Finance Quarterly*, 17(4): 409–428. https://doi.org/10.1177%2F109114218901700404

Farrell, M.J., 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3): 253–290. https://doi.org/10.2307/2343100

Gonzalez-Serrano, E., Rodriguez-Mirasol, J., Cordero, T., Koussis, A.D., Rodriguez, J.J., 2005. Cost of reclaimed municipal wastewater for applications in seasonally stressed semi-arid regions. *Journal of Water Supply: Research and Technology - AQUA*, 54(6): 355–369. http://dx.doi.org/10.2166/aqua.2005.0034

Henriques, A.A., Camanho, A.S., Amorim, P., Silva, J.G. 2020. Performance benchmarking using composite indicators to support regulation of the Portuguese wastewater sector. *Utilities Policy*, 66 : 101082. https://doi.org/10.1016/j.jup.2020.101082

Lorenzo-Toja, Y.. Vázquez-Rowe, I., Chenel, S., Marín-Navarro, D., Moreira, M. T., Feijoo, G. 2015. Eco-efficiency analysis of Spanish WWTPs using the LCA+DEA method. *Water Research*, 68: 637–650. https://doi.org/10.1016/j.watres.2014.10.040

Marques, R.C., Simões, P., Pires, J.S., 2011. Performance benchmarking in utility regulation: the worldwide experience. *Polish Journal of Environmental Studies*, 20:1, 125-132.

Marques R.C., Simões, P., 2020. Revisiting the Comparison of Public and Private Water Service Provision: An Empirical Study in Portugal. *Water*, 12: 1477. https://doi.org/10.3390/w12051477

Mergoni, A., D'Inverno, G., Carosi, L. 2022. A composite indicator for measuring the environmental performance of water, wastewater, and solid waste utilities. *Utilities Policy*, 74: 101285. https://doi.org/10.1016/j.jup.2021.101285

Mocholi-Arce, M., Sala-Garrido, R., Molinos-Senante, M., Maziotis, A. 2022. Performance assessment of the Chilean water sector: A network data envelopment analysis approach. Utilities Policy, 75: 101350. https://doi.org/10.1016/j.jup.2022.101350

Molinos-Senante, M., Hernandez-Sancho, F., Sala-Garrido, R., 2014. Benchmarking in wastewater treatment plants: A tool to save operational costs. *Clean Technologies and Environmental Policy*, 16(1): 149–161. https://doi.org/10.1007/s10098-013-0612-8

Molinos-Senante, M., Sala-Garrido, R., Hernandez-Sancho, F., 2016. Development and application of the Hicks-Moorsteen productivity index for the total factor productivity assessment of wastewater treatment plants. *Journal of Cleaner Production*, 112: 3116–3123. https://doi.org/10.1016/j.jclepro.2015.10.114

Molinos-Senante, M., Maziotis, A. 2021. The impact of greenhouse gas emissions on the performance of water companies: a dynamic assessment. *Environmental Science and Pollution Research*, 28:48284–48297. https://doi.org/10.1007/s11356-021-13879-6

Nafi, A., Bentarzia, Y., Granger, D., Cherqui, F., 2014. Eco-EAR: A method for the economic analysis of urban water systems providing services. *Urban Water Journal*, 11(6): 467–481. https://doi.org/10.1080/1573062X.2013.868497

Pereira, M.A., Marques, R.C. 2022. Technical and Scale Efficiency of the Brazilian Municipalities' Water and Sanitation Services: A Two-Stage Data Envelopment Analysis. *Sustainability*, 14: 199. https://doi.org/10.3390/su14010199

Pointon, C., Matthews, K., 2016. Dynamic efficiency in the English and Welsh water and sewerage industry. *Omega*, 58: 86–96. https://doi.org/10.1016/j.omega.2015.04.001

Portuguese Secretary of State for the Environment (2022). Portuguese Strategic Plan for Water Supply and Wastewater and Stormwater Management – PENSAARP 2030 (https://participa.pt/pt/consulta/projeto-de-resolucao-do-conselho-de-ministros-que-aprova-o-pensaarp-2030)

Rodriguez-Garcia, G., Molinos-Senante, M., Hospido, A., Hernández-Sancho, F., Moreira, M.T., Feijoo, G., 2011 Environmental and economic profile of six typologies of wastewater treatment plants. *Water Research*, 45(18): 5997–6010. https://doi.org/10.1016/j.watres.2011.08.053

Sala-Garrido, R., Hernández-Sancho, F., Molinos-Senante, M., 2012. Assessing the efficiency of wastewater treatment plants in an uncertain context: A DEA with tolerances approach. *Environmental Science and Policy*, 18: 34-44. https://doi.org/10.1016/j.envsci.2011.12.012

Sala-Garrido, R. Molinos-Senante, M., Mocholí-Arce, M., 2019. Comparing changes in productivity among private water service providers integrating quality of service: A metafrontier approach. *Journal of Cleaner Production*, 216: 597–606. https://doi.org/10.1016/j.jclepro.2018.12.034

Tchobanoglous, G., Burton, F.L., Stensel, H.D., 2003. Wastewater Engineering: Treatment and Reuse. Columbus: McGraw Hill education.

Tone, K.A. 2001. "Slacks-Based Measure of Efficiency in Data Envelopment Analysis". European Journal of Operational Research, 130, 498-509. https://doi.org/10.1016/S0377-2217(99)00407-5

Walker. N.L., Styles, D., Gallagher, J., Williams, A.P. 2021. Aligning efficiency benchmarking with sustainable outcomes in the United Kingdom water sector. *Journal of Environmental Management*, 287: 112317. https://doi.org/10.1016/j.jenvman.2021.112317