

<https://doi.org/10.1038/s41545-024-00307-8>

Understanding energy performance in drinking water treatment plants using the efficiency analysis tree approach

Check for updates

Alexandros Maziotis¹ & Maria Molinos-Senante^{2,3} ✉

Water treatment processes are known to consume substantial amounts of energy, making it crucial to understand their efficiency, drivers, and potential energy savings. In this study, we apply Efficiency Analysis Tree (EAT), which combines machine learning and linear programming techniques to assess the energy performance of 146 Chilean drinking water treatment plants (DWTPs) for 2020. Additionally, we utilize bootstrap regression techniques to examine the influence of operating characteristics on energy efficiency. The results indicate that the evaluated DWTPs exhibited poor energy performance, with an average energy efficiency score of 0.197. The estimated potential energy savings were found to be 0.005 kWh/m³. Several factors, such as the age of the facility, source of raw water, and treatment technology, were identified as significant drivers of energy efficiency in DWTPs. The insights gained from our study can be valuable for policymakers in making informed decisions regarding the adoption of practices that promote efficient and sustainable energy use within the water cycle.

Energy plays a crucial role in various aspects of managing water resources in urban settings, including abstraction, treatment, and distribution¹. In particular, the production of drinking water is an energy intensive activity because raw water needs to be cleaned by removing high levels of pollutants². The Sustainable Development Goals established by the United Nations, specifically Goals 6 and 7, impose an obligation on governments to ensure universal access to clean water and enhance energy efficiency³. Consequently, it is crucial for water resources to meet rigorous quality standards before being deemed suitable for drinking purposes. Simultaneously, the challenges posed by climate change and population growth necessitate the sustainable and efficient utilization of water resources^{4,5}. Moreover, drinking water needs to be provided to people at an affordable price⁶.

To address the aforementioned challenges, it is crucial to gain a deeper understanding of the energy efficiency of the water treatment process, including the factors that influence energy performance and the potential for energy savings. By examining the energy efficiency of the water treatment process and exploring opportunities for energy conservation, we can work towards achieving sustainable and efficient use of energy in the water sector. Previous studies, such as those conducted by Loubet et al.⁷, Chini et al.⁸, and Lam et al.⁹, have examined the relationship between energy intensity and the water treatment process. These studies have highlighted the growing energy demands in water treatment resulting from factors like

climate change, population growth, and urbanization. However, it is worth noting that these studies did not specifically explore the link between energy efficiency and the water treatment process.

Energy intensity and energy efficiency are different concepts¹⁰. Energy intensity is the energy consumed [kWh] per unit volume [m³] of drinking water produced and therefore, it does not consider how the quality of raw water and drinking water affected the energy consumed by drinking water treatment plants^{11,12}. On the other hand, energy efficiency is a synthetic indicator which in addition to the volume of drinking water produce integrates the quantity of pollutants removed from raw water¹³.

Only a limited number of previous studies have specifically focused on evaluating the energy efficiency of drinking water treatment plants (DWTPs). These studies, conducted by Molinos-Senante and Guzman¹⁴, Molinos-Senante and Sala-Garrido^{10,15}, Ananda¹⁶, Sala-Garrido and Molinos-Senante¹⁷, and Maziotis et al.¹³, utilized the data envelopment analysis (DEA) method. DEA is a non-parametric technique that employs linear programming to measure energy efficiency, enabling the integration of multiple inputs and outputs for each DWTP¹⁸. One advantage of DEA is that it does not require a priori definition of the functional form of the production frontier, which represents the relationship between inputs and outputs¹⁹. While DEA has its positive characteristics in evaluating energy efficiency of DWTPs, it is important to acknowledge its limitations. One

¹Departamento de Ingeniería Hidráulica y Ambiental, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna, 4860 Santiago, Chile. ²Institute of Sustainable Processes, University of Valladolid, C/Dr. Mergelina, S/N, Valladolid, Spain. ³Centro de Desarrollo Urbano Sustentable ANID/FONDAP/15110020, Av. Vicuña Mackenna, 4860 Santiago, Chile. ✉ e-mail: maria.molinos@uva.es

such limitation is its deterministic nature, which makes it sensitive to outliers in the data²⁰. Consequently, the presence of extreme values can significantly impact the efficiency scores derived from DEA analysis.

DEA method needs implicitly to determine the boundary of the underlying technology, which constitutes the reference benchmark²¹. Its estimation allows the calculation of the corresponding inefficiency score for each unit (DWTP in our case study) as the deviation of each activity or production plan from the frontier of the production possibility set. For this reason, by definition, DEA approach suffers from an overfitting problem^{21–23}. Overfitting occurs when the model becomes too closely tailored to the specific dataset used for analysis, potentially resulting in less robust efficiency scores. Therefore, caution should be exercised when interpreting and relying solely on efficiency scores derived from DEA analysis.

In order to address the limitations of DEA and enhance the accuracy and robustness of efficiency scores, Esteve et al.²¹ introduced a novel method called Efficiency Analysis Trees (EAT). The EAT method combines machine learning and linear programming techniques to measure efficiency. Specifically, it utilizes regression (or decision) trees to estimate the value of the response variable based on thresholds of predictor variables. The EAT approach assumes free disposability to estimate a step function production frontier and calculate efficiency scores. Esteve et al.²¹ demonstrated from a mathematical point of view how the EAT method overcomes overfitting improving the accuracy of the efficiency results. In particular, they demonstrated that the EAT method outperforms other non-parametric techniques such as DEA because the estimated values are not overfitted, ensuring more reliable efficiency measurements²⁴.

In light of the aforementioned context, the main objective of this study is to assess the energy efficiency of the drinking water treatment process using the newly developed method EAT. By employing the EAT approach, this study aims to quantify the optimal level of energy consumption at various thresholds of volume pollutants removed. Additionally, the EAT method allows for the estimation of potential energy savings that could be achieved with efficient water treatment practices. Moreover, in order to gain a deeper understanding of the factors influencing energy efficiency in the water treatment process, this study utilizes bootstrap regression techniques. Specifically, it examines the impact of operational characteristics such as the age of the facility and the type of treatment technology on energy performance. Through these analyses, this study seeks to provide valuable insights into improving energy efficiency in drinking water treatment operations.

Our study makes several significant contributions to the existing literature. Firstly, it stands as the pioneering research that applies machine learning and linear programming techniques to assess the energy

performance of the drinking water treatment process. By employing the EAT approach, we are able to generate robust efficiency scores that are not overfitted, in contrast to other non-parametric methods like DEA. Furthermore, our study estimates the potential energy savings achievable in the drinking water treatment process. This allows us to gain insights into the optimal energy utilization required for different pollutant removal volumes. This information is invaluable for managers and decision-makers as it sheds light on the factors influencing energy intensity and aids in the decision-making process. Importantly, our innovative research was implemented and applied to multiple DWTPs in Chile, enhancing its applicability and relevance to real-world scenarios. This empirical application further strengthens the validity and reliability of our findings.

Results and discussion

Optimal energy use in drinking water treatment

By applying the EAT algorithm, we can determine the optimal level of energy utilization in DWTPs based on the volume of drinking water produced and efficiency in pollutants removal. The findings from our analysis, as depicted in Fig. 1, highlight the significant impact that removing arsenic and sulfates from raw water has on the energy consumption of DWTPs. By contrast, the removal of the other pollutants considered in this study, i.e., turbidity and total dissolved solids, does not significantly explain energy use in DWTPs. This finding aligns with Molinos-Senante and Sala-Garrido²⁵ conclusions, which assessed how various pollutants and the volume of treated water affect the energy intensity in a range of water treatment facilities. They evidenced that total dissolved solids only affect the energy usage in DWTPs employing coagulation-flocculation and pressure filtration techniques. Conversely, the reduction of turbidity was found to influence energy consumption only in DWTPs that utilize pressure filtration. However, for facilities relying on rapid gravity filtration, the energy usage is not affected by the removal of turbidity and total dissolved solids. For those facilities producing more than 2,111,834 m³ per year of drinking water adjusted by arsenic removal efficiency the maximum energy use is 1,054,754 kWh per year, i.e., 0.499 kWh m⁻³. In the case of DWTPs that produce less than 2,111,834 m³ per year of drinking water adjusted by arsenic removal and of more than 428,440 m³ per year adjusted by sulfates removal, the maximum use of energy utilization could reach the level of 539,412 kWh per year. Hence, the optimal energy usage ranges from 0.255 kWh m⁻³ to 1.259 kWh m⁻³ depending on whether the assessment takes into account the efficiency of arsenic or sulfates removal, respectively. Finally, DWTPs producing less than 428,440 m³ per year adjusted by sulfates efficiency removal, the maximum energy use required could be lower than 136,201 kWh/year, i.e., 0.318 kWh m⁻³. It is

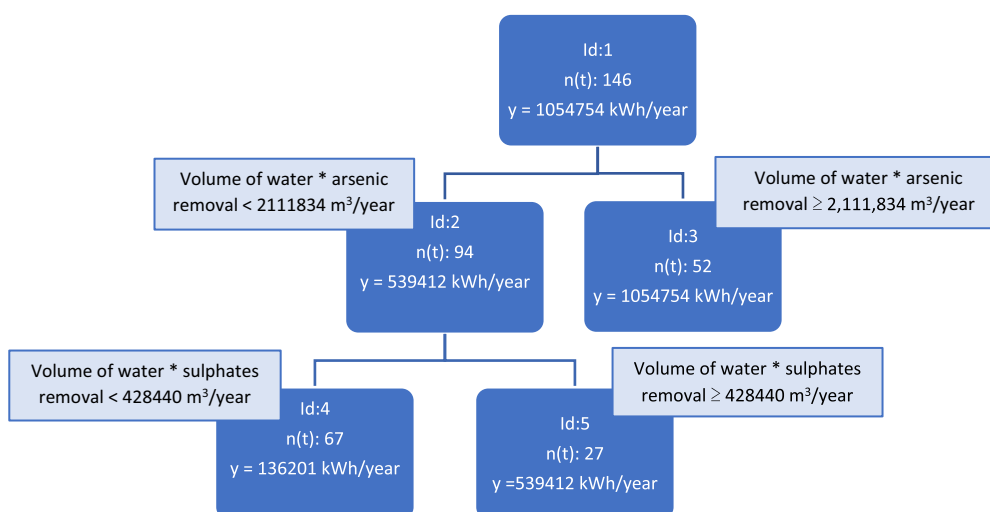


Fig. 1 | Efficiency Analysis Tree (EAT) for estimating optimal energy consumption in DWTPs. The volume of drinking water produced and the quantities of arsenic and sulfates removed from raw water influences of the energy use of DWTPs.

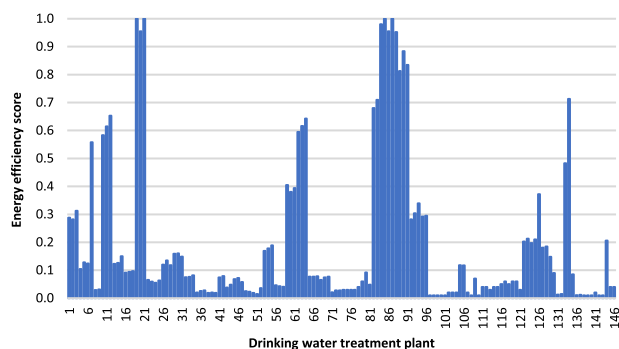


Fig. 2 | Energy efficiency scores of drinking water treatment plants evaluated. Relevant differences on the energy efficiency among drinking water treatment plants are reported.

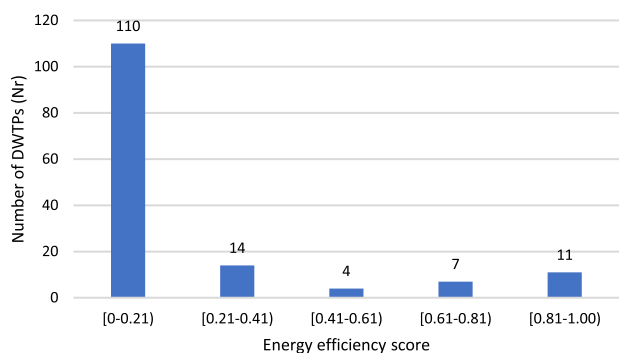


Fig. 3 | Energy efficiency scores across DWTPs. Most of the facilities evaluated present a very poor energy efficiency with a score lower than 0.21.

illustrated the large range of optimal energy use depending on the volume of drinking water produced and the quantity of arsenic and sulfates removed. Considering that the average volume of water produced adjusted by arsenic removal is $1,998,544 \text{ m}^3$ per year, results on Fig. 1, illustrate that smaller facilities can have an energy intensity as low (or lower) than larger ones as even the quantity of sulfates to be removed is large ($>428,440 \text{ m}^3$ per year).

Indeed, the results of our study indicate that optimal energy utilization in DWTPs can vary based on different thresholds of water treated to remove specific pollutants such as sulfates and arsenic. This finding underscores the importance of tailoring energy consumption to the specific requirements of pollutant removal in order to achieve optimal energy efficiency. By considering different thresholds or levels of pollutant removal, DWTPs can determine the appropriate energy usage for their particular circumstances. The results suggest that the energy required to remove pollutants like sulfates and arsenic can have a significant impact on overall energy consumption in DWTPs. Therefore, by optimizing energy usage based on these thresholds, water treatment facilities can enhance their energy efficiency and reduce unnecessary energy consumption.

Energy efficiency of drinking water treatment plants

The findings of our study indicate that the water treatment processes examined exhibit high levels of energy inefficiency, with an average energy efficiency score of 0.197 (Fig. 2). This implies that, on average, there is significant room for improvement in terms of energy consumption, with a potential reduction of almost 80% in energy usage. The distribution of energy efficiency scores among DWTPs is shown in Supplementary Fig. 1.

Moreover, it is observed that a small proportion of the evaluated facilities demonstrated full energy efficiency. Specifically, out of the 146 DWTPs analyzed, only four facilities, representing ~2.7% of the total,

achieved a full efficiency score of 1.00. This indicates that these particular facilities have effectively optimized their energy usage and are operating at the highest level of energy efficiency within the context of our study. The four energy-efficient facilities utilize pressure filters (PF) for drinking water production. However, they differ in their primary raw water sources, with two of them using groundwater and the other two relying on surface water.

Findings from this study slightly differ from those of Molinos-Senante and Sala-Garrido¹⁵, who reported an average energy efficiency score of 0.28 and classified 6 out of 146 DWTPs as energy-efficient. Notably, two facilities deemed efficient in their study were considered inefficient in ours. Furthermore, there are more pronounced discrepancies when compared to Molinos-Senante and Maziotis²⁶, who reported an average energy efficiency score of 0.462, with none of the evaluated facilities being fully energy-efficient. These variations can be attributed to the different methodological approaches used in assessing energy efficiency. Molinos-Senante and Sala-Garrido¹⁵ utilized a double-bootstrap DEA method, which reduces data uncertainty but does not address overfitting issues. In contrast, Molinos-Senante and Maziotis²⁶ used stochastic non-parametric envelopment of data (StoNED), a technique that incorporates both inefficiency and noise in the assessment but still has overfitting limitations. Our study, however, employed the EAT approach, a method that is not prone to overfitting issues. This methodological improvement strengthens the validity of our findings and underscores the need for targeted efforts to improve energy efficiency in water treatment plants.

Figure 3 provides valuable insights into the distribution of energy efficiency scores across the evaluated DWTPs. The majority of the facilities reported energy efficiency scores below 0.21, indicating a significant level of energy inefficiency. On average, these plants would need to reduce their energy consumption by nearly 80% to achieve optimal energy efficiency. On the other hand, there are 18 treatment plants that reported relatively higher energy efficiency scores. However, their scores still fall within the range of 0.21 and 0.61, indicating room for improvement. These facilities have the potential to achieve substantial energy savings, ranging from 40% to 80% on average, and bridge the gap with the most energy-efficient plants in the sector. No common characteristics were observed in terms of source of raw water, treatment train, and ownership for this group of DWTPs. The treatment trains employed are coagulation-flocculation with rapid gravity filters (CF-RGF) by 8 facilities, coagulation-flocculation with pressure filters (CF-PF) by 5, and PF alone by another 5. Similarly, the distribution of the main source of raw water varies, with 8 DWTPs treating mixed raw water, 5 treating groundwater, and 5 treating surface water. Ownership-wise, 14 of the 18 DWTPs are fully privately owned, while the remaining 4 are operated by concessioned companies.

Furthermore, the analysis reveals a group of 11 DWTPs that can be considered as best performers. These facilities attained energy efficiency scores ranging between 0.81 and 1.00, indicating a high level of energy efficiency. These best performers serve as examples of successful energy management practices and provide insights into the potential for achieving optimal energy efficiency in the water treatment sector. The shared characteristic among this group of DWTPs is their use of PF as the primary treatment method for producing drinking water. Additional information of the influence of treatment train on the energy efficiency of DWTPs is shown in Fig. 5. As with the previous group, diverse features are observed in terms of ownership and the main source of raw water. Of these DWTPs, 8 out of 11 are owned by full private water companies, while the remaining three are operated by concessioned water companies. Regarding the source of water, the distribution is as follows: 5 out of 11 DWTPs treat mixed water, 4 out of 11 treat groundwater, and 2 out of 11 treat surface water.

In our study, we investigated the energy-saving potential of the energy-inefficient DWTPs. By applying Eq. (5) and considering the current energy use of the 146 assessed DWTPs, we estimated the potential energy savings for these facilities, as depicted in Fig. (4). The results revealed that the assessed DWTPs have a combined potential energy-saving of 13,344,093 kWh/year. This indicates the substantial opportunity for reducing energy consumption in these facilities while maintaining the same

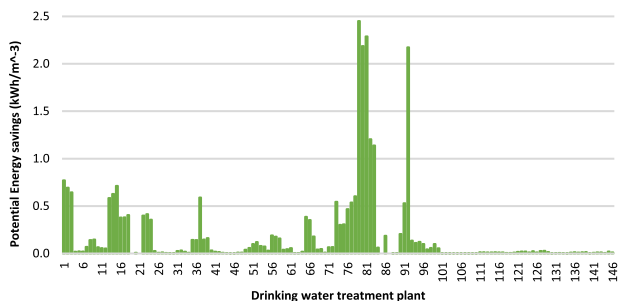


Fig. 4 | Potential energy savings whether DWTPs are energy efficient. DWTPs with the lowest energy performance can save up to 2.45 kWh per cubic meter of water.

Table 1 | Influence of operational characteristics on energy efficiency. Estimates of the bootstrap truncated regression model

Variables	Coeff.	St. Err.	Z-stat	p-value
Constant	-1.131	0.0631	-17.924	0.000
Age of plant	-0.140	0.013	-10.769	0.000
Source of water	-0.231	0.050	-4.620	0.000
Type of ownership	0.103	0.100	1.030	0.300
Type of technology	-0.312	0.043	-7.256	0.000
Sigma	0.175	0.042	4.167	0.000
$\chi^2(5)$	64.120			
p-value	0.000			

Energy efficiency score is the dependent variable

Bold indicates that coefficients are statistically significant at 5% significance level

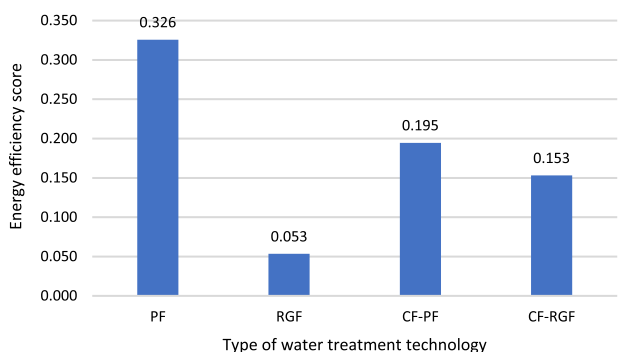


Fig. 5 | Energy efficiency scores across water treatment technologies (PF: pressure filters; RGF: rapid gravity filters; CF-PF: coagulation-flocculation and pressure filters; CF-RGF: coagulation-flocculation and rapid gravity filters). DWTPs using PF are those with the best energetic performance.

volume of drinking water production and pollutant removal from raw water. The mean potential energy savings for the 146 assessed DWTPs were estimated to be 0.005 kWh m⁻³. Additionally, we analyzed the distribution of potential energy savings across the assessed DWTPs. The 25th percentile indicates that 25% of the facilities could achieve energy savings of 0.008 kWh per cubic meter water, while the 75th percentile suggests that 25% of the facilities could achieve energy savings of 0.146 kWh m⁻³. Current energy intensity of DWTPs ranges between 0.002 kWh m⁻³ and 0.215 kWh m⁻³ with an average value of 0.007 kWh m⁻³. The distribution of potential energy savings among DWTPs is shown in Supplementary Fig. 2.

Factors influencing energy efficiency of DWTPs

To gain a deeper understanding of the factors impacting the energy performance DWTPs, it is essential to assess how their operational characteristics influence the previously estimated energy efficiency scores. The findings of this analysis are presented in Table 1. We check the existence of multicollinearity among the explanatory variables in the regression using the variance inflation factor (VIF) test. The estimated value of VIF was 2.31 indicating that there is no multicollinearity in the regression model. The results indicate that the age of the treatment plant, the source of raw water, and the type of treatment technology have a negative effect on energy efficiency. By contrast, the ownership of the facility does not statistically influence on its energy performance which is consistent with past research^{15,26,27}.

Specifically, the study reveals that older DWTPs tend to have lower energy performance. This can be attributed to the lack of updates and improvements in energy-efficient equipment within these aging plants. Molinos-Senante and Sala-Garrido¹⁵ found that newer facilities tend to exhibit better energy performance, suggesting a positive correlation between a facility’s age and its energy efficiency. Conversely, the study by Molinos-Senante and Maziotis²⁶ did not provide a definitive conclusion on how the age of a facility influences its energy efficiency, indicating that the relationship between these factors remains unclear and further research is needed. Furthermore, the analysis demonstrates that DWTPs relying on mixed water resources, such as both surface and groundwater, experience a decrease in energy efficiency. This implies that treating water from multiple sources necessitates extensive treatment processes, potentially leading to higher energy consumption. Consequently, this combination of surface and groundwater treatment may have a detrimental impact on overall energy performance. This finding is consistent with those reported by Molinos-Senante and Maziotis²⁶ whereas Molinos-Senante and Sala-Garrido¹⁵ did not identify the source of raw water as an explanatory factor of energy efficiency of DWTPs.

When focusing on the primary technology used for treating raw water, there are statistically significant differences in energy efficiency scores. Figure 5 provides an illustration of these differences, highlighting that the most energy-efficient technology is PF. On the other hand, treatment plants utilize rapid gravity filters (RGF) technology to remove pollutants from raw water are the least energy-efficient. Treatment plants that utilize a combination of coagulation and flocculation (CF) and PF or RGF to purify water demonstrated slightly higher energy efficiency scores. However, there is still considerable room for improvement in their energy performance to catch up with plants utilizing more energy-efficient technologies. These findings present a partial divergence from the conclusions of Molinos-Senante and Sala-Garrido¹⁰, who conducted a metafrontier DEA assessment on DWTPs. Their study concluded that DWTPs utilizing a combination of CF and rapid gravity filtering (CF + RGF) were the most energy-efficient. They also found evidence that facilities employing RGF as their treatment process were the least energy-efficient. It is important to note that differences in methodologies, data sources, and specific contexts may contribute to variations in the results between studies.

Results from this study provide evidence that the evaluated DWTPs exhibit inadequate energy performance, highlighting significant opportunities to reduce energy consumption. Such reductions can lead to cost savings and help mitigate greenhouse gas emissions, particularly if the energy sources are non-renewable. Water managers and regulators can implement various actions and policies to enhance energy efficiency in water treatment processes (Fig. 6). Potential strategies could be categorized as follows:

DWTPs can improve energy efficiency by both reducing energy use and by removing more pollutants from raw water. Focusing on the first alternative, as it has reported by Sowby et al.²⁸, some practices for managing energy in DWTPs are: (i) Implementing optimized operational procedures can help minimize energy waste. This includes strategies such as optimizing flow rates, ensuring appropriate maintenance of equipment, and adopting efficient operating schedules; (ii) Energy recovery systems, such as energy-

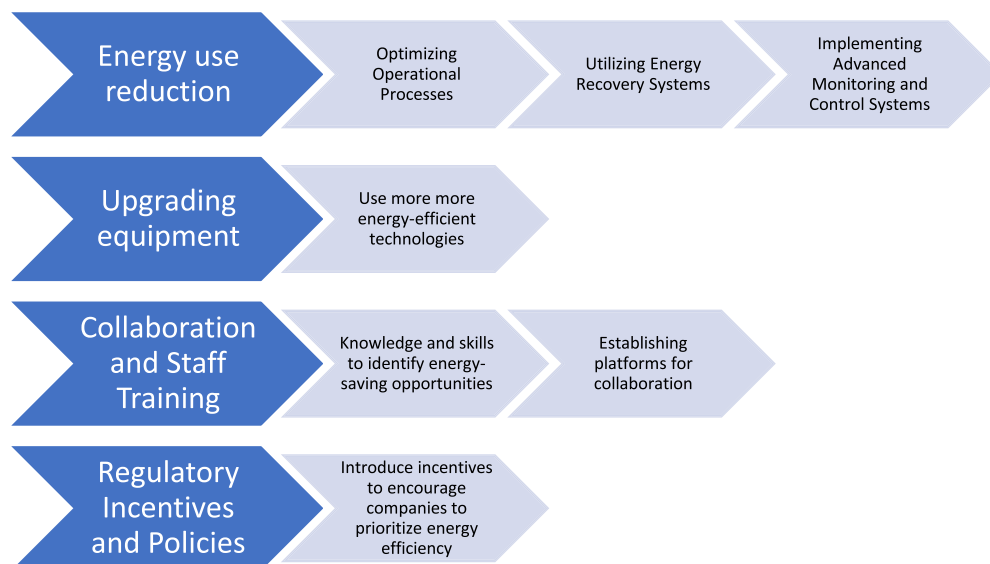


Fig. 6 | Some strategic actions to promote energy efficiency in drinking water treatment plants. Energy efficiency improvement can be achieved by applying a diverse range of approaches.

efficient pumps or turbines, can be installed to capture and utilize energy that would otherwise be wasted during the treatment process and (iii) Deploying advanced monitoring and control systems can enable real-time monitoring of energy consumption and process optimization. This allows operators to identify areas of energy inefficiency promptly and take corrective actions. Results from this study evidenced that older DWTPs tend to have lower energy performance. Thus, water managers can invest in modernizing treatment plant equipment to utilize more energy-efficient technologies. This may involve replacing outdated machinery with newer models that offer improved energy performance. Generation and transfer of specific knowledge is also a valuable tool for improving energy efficiency. Thus, training programs and awareness campaigns can educate DWTP staff about the importance of energy efficiency and provide them with the necessary knowledge and skills to identify energy-saving opportunities in their daily operations. Moreover, establishing platforms for collaboration and knowledge sharing among water managers, researchers, and industry experts can facilitate the exchange of best practices and innovative approaches to energy efficiency in water treatment processes. Finally, given the large room of Chilean DWTPs to improve energy efficiency, the Chilean water regulators can introduce incentives and policies to encourage water companies to prioritize energy efficiency. These can include offering financial incentives for adopting energy-efficient technologies or setting energy efficiency targets that must be met to ensure tariff adjustments or other benefits.

The policy implications of your study’s findings are indeed significant and can provide valuable guidance to stakeholders involved in water treatment processes. By employing a novel approach that combines machine learning and linear programming techniques, this study offers a visually intuitive way for water regulators to understand the maximum energy requirements for different pollutant removal scenarios. This can aid decision-making processes by providing clear insights into the energy implications of water treatment operations. The new method used overcomes overfitting issues often encountered in other efficiency techniques. As a result, the energy efficiency scores derived from this approach are more robust and reliable. This increased reliability can contribute to more informed decision-making, as water regulators can have greater confidence in the efficiency assessments provided. The analysis conducted identifies key factors influencing energy performance in water treatment processes. This knowledge enables water regulators to gain insights into the specific aspects that impact efficiency. For example, recognizing that newer treatment plants tend to be more energy-efficient can inform decisions regarding facility

upgrades or replacements. Similarly, understanding the energy intensity associated with different water sources and treatment technologies can guide choices in resource allocation and process optimization. In this context, Sowby²⁹ empirically proved that those water utilities with energy management policies or plans use less energy. This correlation was attributed to the organization’s culture and operation and also to the identification of energy use as a relevant topic within the organization.

Methods
Energy performance assessment based on efficiency analysis tree

According to Esteve et al.²¹, let’s make the assumption that there is a vector of predictors variables, i.e., factors influencing energy use in DWTPs, defined as x_1, \dots, x_m with $x_i \in \mathbb{R}^m$. Let’s also assume that this set of variables are employed to predict a vector of response variables, i.e., energy used, defined as y, \dots, y_n with $y_i \in \mathbb{R}^n$. The EAT approach selects a predictor variable j and a threshold $s_j \in \mathcal{S}_j$ where \mathcal{S}_j consists of the vector of potential thresholds for the variable j to split the dataset into the right and left node, t_R and t_L , respectively²². The mean squared error (MSE) is used to define the threshold and consequently, the right and left node. This is shown mathematically as follows:

$$R(t_L) + R(t_R) = \frac{1}{n} \sum_{(x_i, y_i) \in t_L} (y_i - y(t_L))^2 + \frac{1}{n} \sum_{(x_i, y_i) \in t_R} (y_i - y(t_R))^2 \tag{1}$$

where t shows the node of the regression tree, $R(t)$ presents the MSE of each node t , n is the size of the sample, and $y(t_L)$ and $y(t_R)$ are the estimated values of the response variable y . Note that the nodes t_L and t_R present the left and right nodes of the tree, respectively. The value of the response variable is calculated using the data that goes to nodes, t_L and t_R .

The estimated values of the response variable for each node of the regression tree are calculated as follows²³:

$$y(t_L) = \max \left\{ \max \{ y_i : (x_i, y_i) \in t_L \}, y \left(I_{T(k|t^* \rightarrow t_L, t_R)}(t_L) \right) \right\}$$

$$y(t_R) = \max \left\{ \max \{ y_i : (x_i, y_i) \in t_R \}, y \left(I_{T(k|t^* \rightarrow t_L, t_R)}(t_R) \right) \right\} \tag{2}$$

where T is the sub-tree that is derived using the EAT method, k is the number of splits, $y(I_{T(k|t^* \rightarrow t_L, t_R)}(t_L))$ and $y(I_{T(k|t^* \rightarrow t_L, t_R)}(t_R))$ is the set of leaf nodes of the regression tree generated after performing the k -th break that Pareto dominates node t_L and t_R ^{21,23}.

To avoid any overfitting problems, the EAT approach employs cross validation techniques to select the best regression tree²¹. Therefore, the production technology that is estimated takes the following form:

$$PT_{T_k} = \left\{ (x, y) \in R^{m+1} : y \leq d_{T_k}(x) \right\} \quad (3)$$

where $d_{T_k}(x)$ is the predictor estimator regarding the sub-tree T_k .

The energy efficiency score, which is a synthetic index embracing energy use and pollutants removed from raw water, for each analyzed DWTP is derived after solving the following linear programming model:

$$\begin{aligned} \theta(x_k, y_k) = \min \theta & \quad (4) \\ \text{s.t.} & \\ \sum_{t \in T^*} \lambda_t a_j^t \leq \theta x_{jk}, j = 1, \dots, m & \\ \sum_{t \in T^*} \lambda_t d_{T^*}^t(a^t) \geq y_{jk}, r = 1, \dots, p & \\ \sum_{t \in T^*} \lambda_t = 1 & \\ \lambda_t \in \{0, 1\}, i = 1, \dots, n & \end{aligned}$$

where θ is the energy efficiency score which is ranged between 0 and 1. We note that when energy efficiency score equals to one, then the unit is energy-efficient. A value lower than one indicates energy inefficiency. $a^t, d_{T^*}(a^t)$ are locations in the input-output space for all $t \in T^*$ where $*$ presents the final sub-tree, and λ are intensity variables that are part of the process to construct the efficient frontier³⁰.

Based on the energy efficiency scores estimated by using Eq. (4), potential energy savings if a DWTP was efficient can be estimated using the following equation:

$$Energy_s = Energy_c * (1 - \theta) \quad (5)$$

where $Energy_s$ is the potential saving in energy and $Energy_c$ is the actual level of energy consumption of the evaluated DWTP.

Factors influencing energy efficiency of DWTPs

In the second step of our analysis, we examine the relationship between energy efficiency scores of DWTPs and their operating characteristics. To accomplish this, we utilize bootstrap truncated regression techniques, as proposed by Simar and Wilson³¹. The choice of employing a truncated regression approach is motivated by the fact that energy efficiency scores are bounded between zero and one. This approach allows us to account for this bounded nature and ensure that the estimated relationship is valid within this range.

By using bootstrap regression techniques, we can mitigate any potential issues related to serial correlation among efficiency scores, error terms, and the operating characteristics. This is an improvement over the standard Tobit regression approach, which may encounter difficulties when dealing with such correlations³¹. The bootstrap method provides a robust framework for analyzing the relationship between energy efficiency scores and the various operating characteristics of DWTPs in our study.

Mathematically, the regression model takes the following form:

$$\theta_i = \mu_0 + \mu_i \xi'_i + time + v_i \quad (6)$$

where θ_i is the EAT energy efficiency score, μ_0 is the constant term, ξ'_i is the set of operating characteristics of any DWTP i , and μ_i are parameters that are estimated. Finally, v_i is the error term which is distributed following the standard normal distribution²⁴.

Case study and variables used

The case study conducted in Chile focuses on assessing the energy performance of 146 DWTPs in the country. The study focused on assessing the energetic performance of water treatment facilities excluding energy use for raw water abstraction. It is important to note that the water industry in Chile operates under a system of private ownership, which was established during the privatization process between 1998 and 2004. Two types of water companies emerged from this process: full private water companies, responsible for the long-term operation and maintenance of the water network, and concessionary water companies, tasked with supplying water for a specific period, typically around 30 years³². Due to the monopolistic nature of the water sector, a national regulator called the Superintendencia de Servicios Sanitarios (SISS) was established. This regulatory body is responsible for setting water tariffs for customers, using an efficient company standard as a benchmark³³. Additionally, the national regulator, SISS, monitors the environmental performance of the water sector. The Ministry of Health establishes quality standards that must be met by treated water before it is distributed to end-users for consumption. These quality standards are based on guidelines set by the World Health Organization¹⁰.

The selection of predictor and response variables is based on past literature on this topic and data availability^{34–39}. The response variable is captured by the energy consumption and is measured in kWh per year. In order to account for the removal of pollutants during the water treatment process and consider its impact on energy efficiency, our study incorporates four quality adjusted predictor variables. Following past practice^{10,40,41}, they are estimated as follows:

$$Quality\ adjusted\ y_s = Volume_w \cdot \left(\frac{Pollutant_{sin} - Pollutant_{sef}}{Pollutant_{sin}} \right) \quad (7)$$

where $Volume_w$ denotes the volume of drinking water produced and is measured in m^3 per year; $Pollutant_{sin}$ is the concentration of the pollutant s in the influent and $Pollutant_{sef}$ is the concentration of the pollutant s in the effluent. Pollutant concentrations are measured in g/m^3 . This study employs four quality adjusted predictors because four pollutants are removed during the water treatment process. The four pollutants considered are sulfates, turbidity, arsenic and total dissolved solids. The four pollutants considered are sulfates, turbidity, arsenic and total dissolved solids. This selection is based on their significant impact on the energy consumption of Chilean DWTPs^{10,25}.

Regarding operational characteristics influencing the energetic performance of DWTPs, the following variables are considered: (i) age of the DWTP measured in years; (ii) source of the raw water treated (surface water; groundwater or mixed water resources, which involves groundwater and surface water blending before its treatment); (iii) ownership of the DWTP which is captured through the use of a dummy variable, i.e., whether the treatment plant owned by a full private or concessionary water company and; (iv) the type of treatment technology used in the DWTPs, i.e., PF ($n = 66$), RGF ($n = 36$), CF-PF ($n = 18$) and CF-RGF ($n = 26$). The pre-treatment of all facilities assessed is a simple screening process and all use chlorine for water disinfection. The descriptive statistics of the variables used in our analysis are reported in Table 2. Data of the variables (predictor, response variables, and operational characteristics of DWTPs) was provided by the Chilean Urban Water Regulator (SISS) requested under the right to public information in Chile and correspond to 2020.

Table 2 | Descriptive statistics of the variables used in the analysis

Variables	Unit of measurement	Mean	St. Dev.	Min	Max	Median
Energy consumed	kWh/year	148,049	195,323	1152	1,054,754	85,792
Turbidity_adjusted	m ³ /year	2,294,674	3,027,762	3126	11,538,685	549,211
Total dissolved solids_adjusted	m ³ /year	1,741,844	2,471,374	4145	10,983,351	370,076
Sulphates_adjusted	m ³ /year	1,370,566	1,913,461	5604	8,113,707	431,100
Arsenic_adjusted	m ³ /year	1,998,544	2,633,425	4669	10,695,107	548,882
Age of plant	Years	28	16	11	72	21

Observations: 146.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Code availability

The codes generated and/or used during the current study are available from the corresponding author upon reasonable request.

Received: 18 July 2023; Accepted: 9 February 2024;

Published online: 29 February 2024

References

- Negi, R. & Chandel, M. K. Assessment on embodied energy and greenhouse gas emissions in urban water system from life cycle perspective: A typical case of India. *Sustain. Cities Soc.* **86**, 104152 (2022).
- Rodríguez-Merchan, V., Ulloa-Tesser, C., Baeza, C. & Casas-Ledón, Y. Evaluation of the water–energy nexus in the treatment of urban drinking water in Chile through exergy and environmental indicators. *J. Clean. Prod.* **317**, 128494 (2021).
- UN (United Nations). Sustainable development goals. Available at: <https://sustainabledevelopment.un.org/?menu=1300> (2015).
- Rabaey, K., Vandekerckhove, T., de Walle, A. V. & Sedlak, D. L. The third route: using extreme decentralization to create resilient urban water systems. *Water Res.* **185**, 116276 (2020).
- Garrido-Baserba, M. et al. The third route: a techno-economic evaluation of extreme water and wastewater decentralization. *Water Res.* **218**, 118408 (2022).
- Pinto, F. S. & Marques, R. C. Tariff recommendations: a panacea for the Portuguese water sector? *Uti. Policy* **34**, 36–44 (2015).
- Loubet, P., Roux, P., Loiseau, E. & Bellon-Maurel, V. Life cycle assessments of urban water systems: a comparative analysis of selected peer-reviewed literature. *Water Res.* **67**, 187–202 (2014).
- Chini, C. M., Schreiber, K. L., Barker, Z. A. & Stillwell, A. S. Quantifying energy and water savings in the U.S. residential sector. *Environ. Sci. Technol.* **50**, 9003–9012 (2016).
- Lam, K. L. & Van Der Hoek, J. P. Low-carbon urban water systems: opportunities beyond water and wastewater utilities? *Environ. Sci. Technol.* **54**, 14854–14861 (2020).
- Molinos-Senante, M. & Sala-Garrido, R. Evaluation of energy performance of drinking water treatment plants: Use of energy intensity and energy efficiency metrics. *Appl. Energy* **229**, 1095–1102 (2018).
- Mo, W., Wang, H. & Jacobs, J. M. Understanding the influence of climate change on the embodied energy of water supply. *Water Res.* **95**, 220–229 (2016).
- Santana, M. V. E., Zhang, Q. & Mihelcic, J. R. Influence of water quality on the embodied energy of drinking water treatment. *Environ. Sci. Technol.* **48**(5), 3084–3091 (2017).
- Maziotis, A., Mocholi-Arce, M., Sala-Garrido, R. & Molinos-Senante, M. Energy efficiency of drinking water treatment plants: a methodological approach for its ranking. *Sci. Total Environ.* **862**, 160840 (2023).
- Molinos-Senante, M. & Guzmán, C. Benchmarking energy efficiency in drinking water treatment plants: quantification of potential savings. *J. Clean. Prod.* **176**, 417–425 (2018).
- Molinos-Senante, M. & Sala-Garrido, R. Assessment of energy efficiency and its determinants for drinking water treatment plants using a double-bootstrap approach. *Energies* **12**, 765 (2019).
- Ananda, J. Explaining the environmental efficiency of drinking water and wastewater utilities. *Sustain. Prod. Consum.* **17**, 188–195 (2019).
- Sala-Garrido, R. & Molinos-Senante, M. Benchmarking energy efficiency of water treatment plants: effects of data variability. *Sci. Total Environ.* **701**, 134960 (2020).
- Mamghaderi, M., Mamkhezri, J. & Khezri, M. Assessing the environmental efficiency of OECD countries through the lens of ecological footprint indices. *J. Environ. Manag.* **338**, 117796 (2023).
- Xian, C., Gong, C., Lu, F., Wu, H. & Ouyang, Z. The evaluation of greenhouse gas emissions from sewage treatment with urbanization: Understanding the opportunities and challenges for climate change mitigation in China's low-carbon pilot city, Shenzhen. *Sci. Total Environ.* **855**, 158629 (2023).
- Costa, M. A., Lopes-Ahn, A. L., Kilger, A. D. C. & Micas, A. F. Limitations of weight restrictions in data envelopment analysis for benchmarking Brazilian electricity distribution system operators. *Uti. Policy* **82**, 101540 (2023).
- Esteve, M., Aparicio, J., Rabasa, A. & Rodríguez-Sala, J. J. Efficiency analysis trees: a new methodology for estimating production frontiers through decision trees. *Expert Syst. Appl.* **162**, 113783 (2020).
- Esteve, M., Rodríguez-Sala, J. J., Lopez-Espin, J. J. & Aparicio, J. Heuristic and backtracking algorithms for improving the performance of efficiency analysis trees. *IEEE Access* **9**, 17421–17428 (2021).
- Esteve, M., Aparicio, J., Rodríguez-Sala, J. J. & Zhu, J. Random forests and the measurement of super efficiency in the context of Free Disposal Hull. *Eur. J. Oper. Res.* **304**(2), 729–744 (2022).
- Molinos-Senante, M., Maziotis, A., Sala-Garrido, R. & Mocholi-Arce, M. Assessing the influence of environmental variables on the performance of water companies: an efficiency analysis tree approach. *Expert Syst. Appl.* **212**, 118444 (2023).
- Molinos-Senante, M. & Sala-Garrido, R. Energy intensity of treating drinking water: understanding the influence of factors. *Appl. Energy* **202**, 275–281 (2017).
- Molinos-Senante, M. & Maziotis, A. Influence of environmental variables on the energy efficiency of drinking water treatment plants. *Sci. Total Environ.* **833**, 155246 (2022).
- Sowby, R. B. Comparison of operational energy requirements in publicly and privately owned U.S. water utilities. *Uti. Policy* **54**, 92–95 (2018a).
- Sowby, R. B., Morehead, N. & Burdette, S. Review of energy management guidance for water and wastewater utilities. *Energy Nexus* **11**, 100235 (2023).

29. Sowby, R. B. Correlation of energy management policies with lower energy use in public water systems. *J. Water Resour. Plan. Manag.* **144**(11), 06018007 (2018b).
30. Zuluaga-Ortiz, R., Camelo-Guarín, A. & Delahoz-Domínguez, E. Efficiency analysis trees as a tool to analyze the quality of university education. *Int. J. Electr. Comput. Eng.* **13**(4), 4412–4421 (2023).
31. Simar, L. & Wilson, P. W. Estimation and inference in two-stage, semiparametric models of production processes. *J. Econ.* **136**(1), 31–64 (2007).
32. Ferro, G. & Mercadier, A. C. Technical efficiency in Chile’s water and sanitation provides. *Ut. Policy* **43**, 97–106 (2016).
33. Molinos-Senante, M., Porcher, S. & Maziotis, A. Productivity change and its drivers for the Chilean water companies: a comparison of full private and concessionary companies. *J. Clean. Prod.* **183**, 908–916 (2018).
34. Berg, S. & Marques, R. Quantitative studies of water and sanitation utilities: a benchmarking literature survey. *Water Policy* **13**(5), 591–606 (2011).
35. Marques, R. C., Berg, S. & Yane, S. Nonparametric benchmarking of Japanese water utilities: institutional and environmental factors affecting efficiency. *J. Water Resour. Plan. Manag.*, **140**(5), 562–571 (2014).
36. Pinto, F. S., Simoes, P. & Marques, R. C. Water services performance: do operational environment and quality factors count? *Urban Water J.* **14**(8), 773–781 (2017).
37. Cetrulo, T. B., Ferreira, D. F. C., Marques, R. C. & Malheiros, T. F. Water utilities performance analysis in developing countries: on an adequate model for universal access. *J. Environ. Manag.* **268**, 110662 (2020).
38. Cetrulo, T. B., Marques, R. C. & Malheiros, T. F. An analytical review of the efficiency of water and sanitation utilities in developing countries. *Water Res.* **161**, 372–380 (2019).
39. Goh, K. H. & See, K. F. Twenty years of water utility benchmarking: a bibliometric analysis of emerging interest in water research and collaboration. *J. Clean. Prod.* **284**, 124711 (2021).
40. Saal, D. S., Parker, D. & Weyman-Jones, T. Determining the contribution of technical change, efficiency change and scale change to productivity growth in the privatized English and Welsh water and sewerage industry: 1985–2000. *J. Product. Anal.* **28**(1–2), 127–139 (2007).
41. Stang, S., Wang, H., Gardner, K. H. & Mo, W. Influences of water quality and climate on the water-energy nexus: A spatial comparison of two water systems. *J. Environ. Manag.* **218**, 613–621 (2018).

Acknowledgements

This work has been supported by project CL-El-2021-07 funded by the Regional Government of Castilla y León and the EU-FEDER and projects TED-130807A-100 and CNS2022-135573 funded by MCIN/AEI/10.13039/501100011033 and by the “European Union NextGenerationEU/PRTR”.

Author contributions

Alexandros Maziotis: Methodology; Software; Writing—Original Draft. María Molinos-Senante: Conceptualization; Visualization; Writing—Review & Editing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41545-024-00307-8>.

Correspondence and requests for materials should be addressed to Maria Molinos-Senante.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024