Journal of Water Process Engineering 64 (2024) 105596

Contents lists available at ScienceDirect

Journal of Water Process Engineering

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

Assessing energy efficiency of water services and its drivers: A case study from water companies in England and Wales

Maria Molinos-Senante [a,*], Alexandros Maziotis [b,c]

[a] Institute of Sustainable Processes, Universidad de Valladolid, C/ Mergelina S/N, Valladolid, Spain
[b] Departamento de Ingeniería Hidráulica y Ambiental, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna, 4860 Santiago, Chile
[c] Department of Business, New York College, Leof. Vasilisis Amalias 38, Athina 105 58, Greece

ARTICLE INFO

Editor: Guangming Jiang

Keywords:
Energy efficiency
Artificial neural networks
Data envelopment analysis
Operating characteristics
Water services
Water-energy nexus

ABSTRACT

Understanding how energy efficient the water services are and what drives inefficiency can greatly assist water utilities in delivering sustainable services. This study employs a neural network (NN) approach to measure the energy efficiency of water services in relation to the volume of drinking water supplied and the number of connected properties. Unlike other non-parametric approaches, NN allows capturing the complex relationships and dependencies between various factors influencing energy efficiency of water companies. An empirical application for English and Welsh water utilities embracing water only companies (WoCs) and water and sewerage companies (WaSCs) over 2008–2020 was conducted. The average energy efficiency score was found to be 0.411, indicating that water utilities could potentially save 0.54 kWh per cubic meter of drinking water supplied. Notably, WaSCs exhibited better energy performance compared to WoCs, with energy efficiency scores of 0.559 and 0.239, respectively. Nevertheless, based on the volume of water delivered, WaSCs could save 0.65 kWh/m^3 whereas WoCs potential energy savings are 0.24 kWh/m^3. Energy efficiency remained relatively stable across the years, with average values of 0.440, 0.388 and 0.454 for the periods 2008–2010, 2011–2015, and 2016–2020, respectively. The analysis conducted using decision tree methods highlighted the relevance of water treatment quality and the source of raw water as key variables influencing the energy efficiency of water utilities. These findings can be valuable for policymakers, enabling them to gain a deeper understanding of the driving factors behind energy efficiency in water service provision.

1. Introduction

Energy plays a crucial role in the provision of water services, encompassing the entire water supply chain, from abstraction to treatment and distribution [1]. The utilization of energy within the water sector is influenced by a range of variables such as the source and quality of raw water and the distance to the final destination [2–4]. It could also be affected by the level of water lost in the network because it would require more water to be abstracted and thus, more energy ([95,5]). It is worth noting that water and energy resources are inherently interconnected, and their interdependence is expected to grow in the coming years due to factors such as population growth, economic development, and climate change [6]. The management of these resources should consider their mutual relationship to ensure sustainable and efficient water and energy systems.

The sustainable utilization of energy and ensuring universal access to clean water at an affordable cost aligns with the Sustainable Development Goals established by the United Nations in 2015 [7]. Achieving efficient and sustainable energy use in the provision of water services can yield significant economic and environmental benefits [8]. It enables water utilities to realize cost savings and deliver clean water to all individuals at affordable rates. Furthermore, it helps in mitigating the excessive exploitation of groundwater and surface water resources [9], as well as reducing the emission of greenhouse gases and air pollutants [10,11]. Hence, there is a growing emphasis on understanding the energy use of water services and identifying the factors driving inefficiencies within the water-energy nexus among policymakers [12].

Several research studies have highlighted the benefits of reducing energy demand and water consumption within the context of climate change and population growth (e.g., [13–17]). Other studies have focused on quantifying energy savings at water supply system level [18,19] or specific components of the water supply process such as

* Corresponding author.
E-mail address: maria.molinos@uva.es (M. Molinos-Senante).

https://doi.org/10.1016/j.jwpe.2024.105596
Received 11 April 2024; Received in revised form 21 May 2024; Accepted 3 June 2024
Available online 7 June 2024

drinking water treatment [6,20]. However, a key limitation of these studies is the lack of measurement of energy efficiency in water services. Energy efficiency in this context refers to reducing the amount of energy consumed by water utilities while maintaining the same level of outputs (i.e., volume of drinking water) [21,22].

Assessing the energy efficiency of water services is crucial for understanding the effectiveness of energy management practices and identifying areas for improvement. Measuring energy efficiency enables water utilities to evaluate their energy performance and identify opportunities to optimize energy use without compromising the quality or quantity of water supplied [23,24]. Addressing this limitation in previous studies by evaluating the energy efficiency of water services can provide valuable insights for policymakers, water utilities, and researchers. It facilitates a more comprehensive understanding of the energy-water nexus and supports the development of strategies and initiatives to enhance energy efficiency, reduce environmental impacts, and promote sustainable water management practices [25].

Within the framework of water utilities, stochastic frontier analysis (SFA) and data envelopment analysis (DEA) are indeed widely employed methodologies for assessing efficiency, although they primarily focus on overall efficiency rather than specifically targeting energy efficiency [26,27]. SFA is a parametric method that constructs an efficient frontier by estimating the best line that fits the data. It requires the specification of a functional form for the production technology and results are sensitive to the different assumption of the distribution of inefficiency (e.g. half-normal, exponential) [28,29]. On the other hand, DEA is a non-parametric method that uses linear programming techniques to build a piecewise frontier based on the observed data. Therefore, no statistical estimation of the frontier is performed [30,31].

While SFA and DEA are valuable tools for evaluating overall efficiency in the context of water utilities, artificial neural networks (ANNs or NNs) have gained significant attention in recent years as a non-parametric method for estimating and predicting complex relationships in various fields [32]. ANNs are a type of machine learning technique inspired by the functioning of the human brain [33]. ANNs have proven to be effective in capturing non-linear and complex patterns in data, making them suitable for situations where the relationships between inputs and outputs are unknown or non-linear [34,35]. They consist of interconnected layers of artificial neurons that process and transform input data to produce output predictions. One of the advantages of ANNs is their ability to learn from data and adapt to complex relationships without making strong assumptions about the underlying functional form. This makes them particularly useful when dealing with non-linear relationships and capturing intricate patterns that may exist in the data [36].

Previous research has demonstrated that NNs offer a viable alternative approach for assessing efficiency [37]. Within the context of water utilities, prior investigations utilizing NNs have primarily concentrated on predicting water leaks [38,39], evaluating customer satisfaction [40,41], forecasting water quality [42–45], material design and performance assessment for wastewater treatment [46,47], among others. Regarding efficiency assessment, only Nafi and Brans [48] and Molinos-Senante and Maziotis [49] employed NNs to evaluate the performance of water companies. Both studies represent notable contributions in assessing the economic performance of water utilities but do not specifically address energy efficiency. This suggests that there is a research gap in utilizing NNs for evaluating the energy efficiency of water services provided by water utilities.

Based on the aforementioned literature, the following hypotheses and research questions are proposed:

Hypothesis 1. Artificial neural networks serve as a reliable methodology to estimate and predict energy efficiency in water utilities, capturing the complex relationships and dependencies between various factors influencing energy efficiency.

Hypothesis 2. English and Welsh water companies are energy inefficient and therefore, present potential energy savings.

Hypothesis 3. Operational characteristics of water companies have a statistically significant influence on their energy efficiency.

Research Question 1: *What are the energy efficiency and potential energy savings of English and Welsh water companies across different years?*

Research Question 2: *Which operational variables, and to what extent, influence the energy efficiency of water companies in England and Wales?*

Against this background, this study has three main goals. The primary aim is to utilize NNs to estimate and predict the energy efficiency of water services provided by various water utilities. To validate the accuracy and dependability of the energy efficiency scores obtained through NNs, a comparison is made with scores obtained from DEA. The second objective is to quantify the potential energy savings that could be achieved if water utilities were to operate with high energy efficiency in delivering their services. Lastly, the study aims to analyze how various operational characteristics of water companies influence their energy efficiency. This is done through the use of decision trees methods such as regression trees and random forest. Our empirical study focuses on the water services that are provided by several water utilities in England and Wales.

Our contribution to the existing literature can be summarized in the following ways. Firstly, we assess the energy efficiency of water utilities by employing NNs which allows to identify hidden patterns in water-energy nexus that may not be evident through traditional analysis methods. This allows for a more accurate and comprehensive understanding of energy efficiency factors. Water utilities often involve intricate nonlinear relationships between energy consumption and various operational parameters. NNs are capable of capturing and modeling such nonlinearities effectively. Finally, NNs can be utilized to forecast energy efficiency outcomes based on historical data and current operating conditions. Hence, this proactive approach will allow water companies to optimize their energy management strategies and make informed decisions. Secondly, we utilize decision tree methods to uncover concealed interactions within the data, enabling us to comprehend how energy efficiency might vary based on various operational characteristics.

Following this brief introduction, the structure of the article is organized as follows: Section 2 details the methodology used to estimate energy efficiency and the influence of environmental variables on efficiency. Section 3 describes the case study, including the presentation of the variables employed. Section 4 presents and discusses the main findings. The final section outlines the main conclusions drawn from the study

2. Methodology

2.1. Energy efficiency estimation

The method employed in this study to assess the energy efficiency of water supply processes and identify its driving factors involves the use of NNs. Unlike traditional methods, i.e., DEA and SFA, NNs do not rely on a predetermined functional form to describe the relationship between inputs and outputs, making them non-linear and non-parametric models [34].

The specific type of NN utilized in this study is the multilayer perceptron (MLP) network, which consists of three layers of interconnected nodes or neurons: the input layer, hidden layers, and output layer [50]. The input layer receives the relevant input variables, while the hidden layers perform intermediate computations, and the output layer produces the desired output, which in this case is the predicted energy consumption. Based on this desired output we can generate energy efficiency scores [51].

Fig. 1 provides a visual representation of the general structure of a MLP model, showcasing the interconnected nodes and layers. Through iterations and adjustments of the weights (coefficients) between the

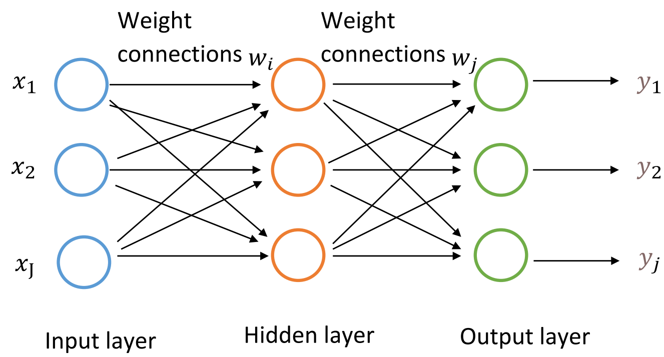

Fig. 1. Architecture of a neural network based on layer perceptron.

neurons, the MLP learns and improves its predictive capabilities, allowing it to eventually estimate the energy efficiency scores based on the provided inputs [52].

The input layer includes the data sample denoted as $x_1, \ldots, x_j$ while the output layer includes the desired outcome (target) category and is denoted as $y_1, \ldots, y_j$. In our study, the objective is to predict energy efficiency, which serves as the output or target variable, based on a set of inputs, specifically water connected properties. The MLP model consists of hidden layers that act as intermediate layers, where the inputs are connected to each other through weights or coefficients. These weights are updated during the learning process, following methodologies proposed by Emrouznejad and Shale [53], Azadeh and Javanmardi [54], Ciampi and Gordini [55], and Kwon [56]. The updated weights are then utilized as inputs to generate the desired output, as described by Emrouznejad and Shale [53]. A mathematical presentation of a MLP for any neuron j is provided below:

$$y_j = f(u_j + b_j) \tag{1}$$

$$u_j = \sum_{i=1}^{N} w_{ji} x_i \tag{2}$$

where in Eq. (2) $x_1, \ldots, x_n$ presents inputs, $w_{j1}, \ldots, w_{jn}$ are the weights that connect inputs and u_j is the weighted outcome of inputs. In Eq. (1) b captures the constant (bias) term, y is the desired output and $f(.)$ is the activation function [57]. The activation functions play a crucial role in defining the output format of a neural network. These functions are mathematical forms that are continuous, bounded, differentiable, and monotonically increasing [34]. The choice of activation function determines the range of values the output can take, such as being restricted to positive values only or binary values between zero and one. In the output layer, the activation function is typically a linear function of the inputs, including the weights and the constant term [36,56]. Various functional forms can be employed as activation functions in the hidden layers of the neural network.

The back-propagation (BP) method is the most common method to make predictions and therefore our study adopts this method [58]. The BP technique is a repeated process where in each iteration the predicted output derived from running the MLP is compared with the observed (actual) output. The difference between predicted and actual output gives the error. The error is then fed back to the model to update weights and predictions. The final predicted output is the one with the minimum error [56].

In MLP the dataset is separated into two datasets, the training and the testing dataset. The training dataset is used to fit (train) the model and produce in-sample results [59,60]. The testing dataset is used for predictions, i.e., to produce the out-sample result. Based on these predictions, efficiency scores can be estimated as well [61].

Determining the optimal architecture of the MLP model involves making decisions regarding several key aspects, including the number of hidden layers, the number of neurons within each hidden layer, the choice of activation functions, and the learning rate. Studies by Kwon et al. [62] and Zhu et al. [60] discuss the significance of these decisions in the context of model fitting. Typically, MLP has one or, at most, three hidden layers, with a recommendation of one hidden layer [53]. The determination of the number of neurons in the hidden layer was defined based on the minimum root-mean-squared Error (RMSE) [63,64]:

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}(z_i - d_i)^2\right)^{1/2} \tag{3}$$

where n is the number of data in each set, z_i is the network predicted value, and d_i is the desired output that the network is trying to reach.

Different forms of activation functions can be utilized in the hidden layers, such as the logistic function or rectified linear unit (ReLU), among others. The selection of an appropriate activation function depends on the specific requirements and characteristics of the problem under consideration. The learning rate captures how slowly or quickly the MLP model learns the problem [65]. The best MLP model is used for training. Once the best MLP model has been determined and fitted, the predicted output is obtained using the testing dataset. Efficiency score (energy efficiency in our study) is then derived as follows [36]:

$$Efficiency = \frac{y^{actual}}{\hat{y} + max(\varepsilon)} \tag{4}$$

where ε is the error, the difference between the predicted and actual output, $\hat{y}$ is the predicted output and y^{actual} is the observed (actual) output. Efficiency takes a value between zero and one. If efficiency is one, then the unit is on the efficient frontier, whereas a value lower than one indicates that the unit is inefficient and therefore, presents room for improvement.

To assess the robustness of the efficiency scores obtained from the MLP model, we compare them with the results obtained using the non-parametric DEA method. DEA is a deterministic method that evaluates the efficiency of units by employing linear programming techniques to construct a piecewise frontier based on observed data [21]. The DEA approach involves constructing an efficient frontier based on the observed data and any deviations from this frontier indicate inefficiency. Hence, it provides a deterministic assessment of efficiency. The generic form of a DEA model is as follows:

$$max \sum_{s=1}^{q} \omega_s y_{so} \tag{5}$$

$$\sum_{s=1}^{q} \omega_s y_{sj} - \sum_{i=1}^{n} \mu_i x_{ij} \le 0 \; j = 1, \ldots, J$$

$$\sum_{i=1}^{n} \mu_i x_{i0} = 1$$

$$\omega_s, \mu_i \ge 0, i = 1, \ldots, n; s = 1, \ldots, q$$

where o presents the unit under evaluation, J is the total number of units, q is the total number of outputs y, n is the total number of inputs, and ω_s and μ_i are the corresponding weights for outputs and inputs, respectively. The solution of the linear programming model produces an efficiency score for each unit which, as in the case of the MLP model, ranges between zero and one. A value of one means that the unit is fully efficient, whereas a value lower than one implies inefficiency.

2.2. Influence of operational characteristics on energy efficiency

To evaluate the impact of operating characteristics (environmental variables) on the energy performance of water services, we employ a regression tree analysis approach. This approach allows us to visualize the interactions between different elements of water services and study the role of operating characteristics in influencing energy performance [66]. In our analysis, we consider the energy efficiency score of each unit as the dependent variable, while the set of operating characteristics serves as the independent variables. The regression tree algorithm divides the dataset into multiple subsets based on thresholds determined for each operating characteristic. Each subset represents a distinct combination of operating characteristics and generates a predicted average value of the efficiency score based on the observations within that subset. This process helps uncover patterns and relationships between the operating characteristics and the efficiency score [67]. One of the key advantages of using regression trees is that they provide visual representations of the importance of each independent variable, i.e., operating characteristic, on the dependent variable, i.e., efficiency score. The higher the value of an independent variable, the greater its impact on the efficiency score.

To assess the robustness of the regression tree model, we employed another decision tree method known as random forest (RF). The RF method, initially developed by Ho [68] and Breiman [69], offers certain advantages such as reduced sensitivity to outliers and the ability to generate reliable predictions [70]. In the RF approach, multiple regression trees are estimated using bootstrap methods [71]. The predicted variable at the subset is calculated as an average of the predictions of all decision trees [72–74].

To ensure that decision trees in RF regression are not highly correlated with each other, the following steps are followed. In the first step, a bootstrap sample is randomly selected from the original training dataset. This process involves randomly selecting data points with replacement, creating a new training dataset for each decision tree. In the second step, each time a division is performed, the algorithm randomly selects a subset of m independent variables from the full set of p independent variables. By default, $m = p/3$ [66,75]. Once all decision trees are constructed, the predicted (outcome or target) variable is estimated by taking the mean prediction from all the decision trees [74].

Similar to the regression tree approach, the RF regression also allows for the visualization of the importance of each independent variable on the dependent variable. Variables with higher values indicate a greater influence on the dependent variable, while variables with lower values have relatively less impact [66,70,74].

3. Case study description

Our empirical study focuses on the drinking water services provided by several water utilities in England and Wales over the years 2008–2020. Drinking water is supplied by Water and Sewerage companies (WaSCs) and Water only companies (WoCs) who are under private ownership and their performance is monitored by the economic water regulator, the Water Services Regulation Authority (Ofwat). Every five years Ofwat determines the future revenue that water utilities are allowed to recover from their customers by challenging their business plans (price review).

The selection of variables in this study was based on previous research assessing the performance of water companies (e.g., [26,27,76–79]), as well as the availability of statistical data for all water companies for all assessed years. Considering that the main objective of this study is assessing the energy efficiency of water companies, the output variable in the MLP model was the energy consumed for the provision of water services. It was measured in MWh per year [80–84]. Past research [8,22,85–88] demonstrated that energetic performance of water companies presents economies of scale. Hence, the two input variables considered were the volume of water delivered, measured in megalitres per year, and the annual number of water connected properties, measured in thousands per year.

In examining the environmental variables that potentially impact the energy efficiency of water companies, according to specific literature for English and Welsh water companies, the following factors were analyzed. The first two operating characteristics focused on the source of raw water and were expressed as the percentage of water abstracted from reservoirs and boreholes, drawing insights from studies by Saal et al. [89] and Villegas et al. [90]. The subsequent three environmental variables pertained to the treatment quality of the water process. These variables were defined as the number of treatment works necessary to purify water sourced from surface and groundwater resources. Additionally, a variable indicating the percentage of water undergoing extensive treatment was included, utilizing information from sources such as Ofwat [91] and Walker et al. [84]. The final environmental variable considered was density, which was measured as the population-to-water mains length ratio. This variable aimed to capture the relationship between population density and water infrastructure [88].

The statistical information for the variables used in the study was collected from Ofwat and the annual reports of water companies. Table 1 presents the summary statistics of these variables.

4. Results and discussion

4.1. Basics of the estimated neural network

In line with previous research conducted by Wang et al. [92] and Liao et al. [36], the data used in the MLP was normalized using the min-max process. This normalization technique ensures that the data falls within a specific range for improved model performance:

$$y = \frac{(x - min)}{(max - min)} \tag{6}$$

The dataset was divided into two separate sets: 70 % for training the MLP and 30 % for testing its performance. This division of data into training and testing sets allows for the evaluation of the model's generalization ability on unseen data, as suggested by Kwon [56]. For the MLP architecture, the input layer consisted of 2 neurons, reflecting the number of input variables used in the model. The output layer had one neuron, corresponding to the predicted energy efficiency value. A single hidden layer was employed in the MLP, as it has been established that a single hidden layer is capable of modeling any function adequately, as noted in the works of Azadeh et al. [35], Kwon et al. [62], and Kwon [56].

The RMSE, used to determine the optimal number of neurons in the hidden layer, was found to be 0.0608, as indicated in Table 2. We tested a range of neurons (5–12) in the hidden layer as suggested by Nabavi et al. [93]. Thus, the optimal number of neurons was determined to be five. In the appendix, additional performance metrics for the training and testing processes are presented in Table 1 - Appendix. These metrics further support the suitability of the MLP model in predicting energy efficiency. The activation function used in the hidden layer was Rectified Linear Unit (ReLU), a commonly employed activation function known for its effectiveness in deep learning architectures:

$$f(x) = x^{+} = max(0, x) = \frac{x + |x|}{2} = \begin{cases} x \text{ if } x > 0 \\ 0 \text{ otherwise} \end{cases} \tag{7}$$

where x is the input to a neuron [96].

For the output layer, a linear activation function was applied. The backpropagation algorithm was employed with a maximum of 1000 epochs (iterations). The learning rate, which determines the step size during the training process, was 0.01 following the approach described by Azadeh et al. [34]. The momentum rate was 0.02 following the approach described in Nabavi et al. [93] and Nabavi et al. [94].

4.2. Energy efficiency of water companies in England and Wales

Based on the estimations made by the MLP, the average energy efficiency of water companies in England and Wales was found to be 0.411, as shown in Fig. 2. This indicates that, on average, there is a

Table 1
Descriptive statistics of variables to assess energy efficiency of English and Welsh water companies.

Variables	Unit of measurement	Mean	Std. Dev.	Min.	Max.
Energy use	Mwh/year	200,188	150,545	16,317	561,564
Water connected properties	000s/year	1439	1132	120	4047
Volumes of water delivered	Ml/year	695	558	56	2169
Water taken from rivers	%	29	25	0	86
Water taken from boreholes	%	39	30	0	92
Number of surface water treatment works	nr	16	15	1	54
Number of groundwater treatment works	nr	49	40	2	127
Water receiving high levels of treatment	%	0.59	0.23	0.22	0.99
Population density	000s/km	0.47	0.28	0.14	1.26

Observations: 228.

Table 2
Estimated neural network parameters to assess energy efficiency of water companies in England and Wales.

Network architecture	2–5–1
Activation function	ReLU/linear
Algorithm	Back propagation
Epochs	1000
Leaning rate	0.010
Root-mean-squared error	0.0608

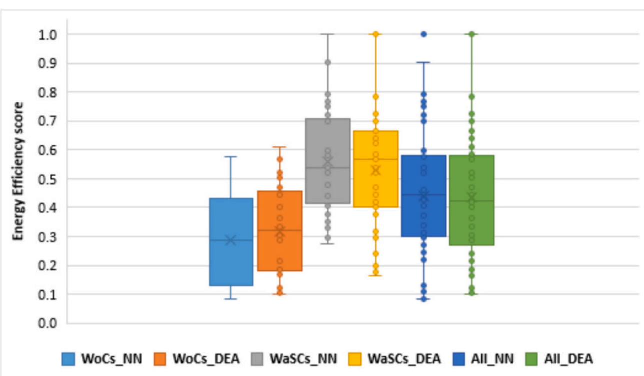

Fig. 2. Average energy efficiency scores for water only companies (WoCs), water and sewerage companies (WaSCs) and all water companies (all) based on neural network (NN) and data envelopment analysis (DEA) estimations.

potential for utilities to achieve energy savings of up to 55.9 % compared to their current energy consumption levels. The study revealed significant disparities in energy efficiency among the evaluated water companies. The utility with the lowest energy efficiency score, at 0.084, would need to reduce its energy use by >91 % in order to provide water services at a more efficient level. On the other hand, the most energy efficient utility achieved a perfect efficiency score of 1, representing 100 % efficiency. This demonstrates a substantial divergence in energy efficiency levels across the water companies analyzed. Consequently, the findings suggest that the English and Welsh water companies require different efforts to enhance their energy efficiency. It highlights the need for tailored strategies and interventions to improve energy efficiency based on the specific circumstances and performance of each utility.

Upon comparing the performance of both types of water companies, the study found that, on average WaSCs exhibited higher energy efficiency levels compared to WoCs, as depicted in Fig. 2. WoCs, on average, performed poorly from an energy perspective, with a mean energy efficiency score of 0.289. This indicates that there is a potential for energy savings of up to 71.1 % among WoCs. The most energy efficient WoC achieved an efficiency score of 0.575, indicating that there is room for improvement in the energy consumption of an average WoC. On the other hand, WaSCs performed relatively better with an average energy efficiency score of 0.559. These findings highlight the need for targeted energy efficiency measures and improvements for both types of water companies. While WoCs require substantial improvements in their energy consumption, WaSCs also have room for enhancement to achieve even higher energy efficiency levels.

In Fig. 2, the energy efficiency scores obtained through DEA techniques are presented for comparison purposes. The results obtained from DEA are similar to those obtained from the MLP model. The potential savings in energy use among water utilities were estimated to be around 56.4 % on average. WoCs were found to be significantly less energy efficient than WaSCs, with mean energy efficiency scores of 0.317 and 0.527, respectively. This suggests that both types of water companies need to make substantial improvements to enhance their energy performance.

The Pearson correlation coefficient, as presented in Table 3, supports the finding that the energy efficiency scores obtained from the MLP model and DEA model are strongly correlated. This indicates that the predicted energy efficiency scores from the MLP model can serve as a reliable proxy for the DEA energy efficiency scores. Consequently, the energy efficiency scores derived from the MLP modeling approach can be considered as reliable, robust, and suitable for efficiency analysis purposes.

Fig. 3 illustrates the distribution of energy efficiency scores, estimated using the MLP, across water companies. The majority of observations related to WoCs indicate high levels of energy inefficiency, with none of them reporting an energy efficiency score exceeding 0.60. Specifically, 11 out of 30 observations (36.7 %) performed extremely poorly, with a mean energy efficiency score below 0.20. In contrast, several WaSCs demonstrated satisfactory energy performance. Among the 39 observations related to WaSCs, 12 (30.8 %) exhibited an energy efficiency score >0.61.

To determine the statistical significance of the energy efficiency differences between WaSCs and WoCs, a non-parametric Mann-Whitney test was conducted. The resulting p-value was lower than 0.005, indicating that the energy efficiency disparities between the two types of water companies are statistically significant. These findings highlight the notable differences in energy efficiency performance between WaSCs and WoCs, with WaSCs generally outperforming WoCs.

Fig. 4 presents the evolution of energy efficiency scores for each type of water utility and the overall sector, divided into three sub-periods that correspond to the regulatory cycles. The first sub-period, from 2008 to 2010, corresponds to the 2004 price review, during which the regulator implemented various incentive schemes to encourage performance improvement in the water industry. As part of these schemes, water

Table 3
Pearson correlation coefficient between energy efficiency scores estimated using MLP and DEA methods.

All water companies	0.916
WoCs	0.821
WaSCs	0.935

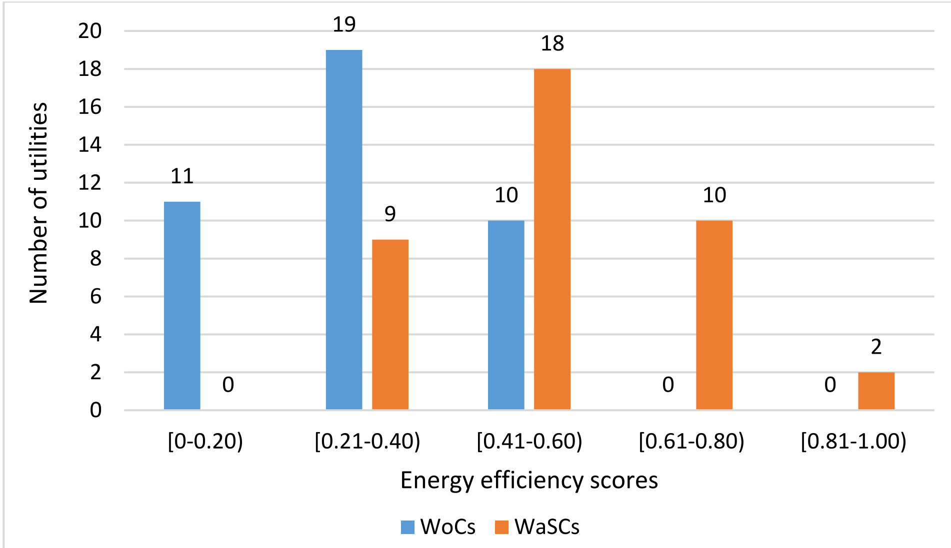

Fig. 3. Histogram of energy efficiency scores for English and Welsh water only companies (WoCs) and water and sewerage companies (WasCs).

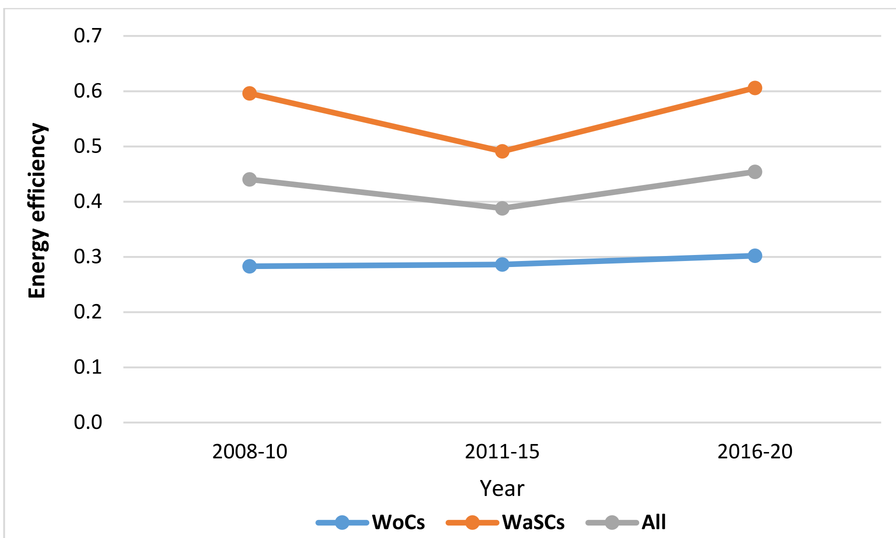

Fig. 4. Evolution of energy efficiency of English and Welsh water companies.

utilities were allowed to retain any savings in operational costs, including energy costs, regardless of the year in which these savings were achieved [90]. The results indicate that the energy efficiency of the water sector was relatively low during this period, with an average efficiency score of 0.440. This suggests that there was room for improvement in energy efficiency across the sector at that time.

During the sub-period from 2011 to 2015, the regulator continued to implement the rolling incentive mechanism for operational expenditure. Additionally, any savings in infrastructure maintenance were shared between utilities and customers. However, the findings indicate a deterioration in energy efficiency for an average WaSC, while it remained unchanged for an average WoC. This suggests that there were no significant gains in energy efficiency during this period. The water industry as a whole still had a considerable way to go in terms of improving energy efficiency, as it would have required reducing energy use by >60 % to reach an efficient level.

The period from 2016 to 2020 corresponds to the 2014 price review, during which the regulator implemented a set of common performance indicators to monitor the economic and environmental sustainability of all water utilities. Additionally, each water utility had the opportunity to report customized performance indicators after consulting with their customers. These indicators included metrics such as water leakage and pollution incidents, which were linked to financial rewards or penalties based on performance. Other indicators, such as greenhouse gas emissions, had a reputational impact on the utilities. The results of this period demonstrated an increase in energy efficiency for both WoCs and WaSCs. This indicates that water utilities made efforts to control their production costs and improve their overall performance. However, there is still room for further improvements in energy efficiency. Potential savings in energy could reach the level of 40 % for WaSCs and 70 % for WoCs, respectively. These findings suggest that there are significant opportunities for water utilities, particularly WoCs, to enhance their energy performance and reduce their energy consumption.

Overall, the results of the study highlight the high levels of energy inefficiency within the water sector in England and Wales. The less energy efficient utilities in the industry have a significant gap to bridge in order to improve their energy performance and reach the level of the most efficient utilities. This finding emphasizes the need for both regulators and regulated utilities to take action in addressing and eliminating energy inefficiencies. By implementing measures and strategies aimed at improving energy efficiency, regulators can set guidelines and incentives to encourage utilities to prioritize energy conservation and sustainable practices. At the same time, utilities themselves need to invest in technologies, processes, and infrastructure that promote energy efficiency and reduce energy consumption. Addressing energy inefficiencies in the water sector not only leads to cost savings for the utilities but also contributes to environmental sustainability and resource conservation. It is crucial for both regulators and utilities to collaborate and work towards eliminating energy inefficiencies, ensuring a more efficient and sustainable water industry in England and Wales.

4.3. Potential energy use savings in the provision of drinking water

Based on the energy efficiency scores ranging between 0 and 1, the potential savings in energy use can be estimated by comparing the current energy use of water companies with the energy use that would be expected if the companies were fully efficient.[1] The total potential energy savings for the 228 evaluated observations were 5,688,192 MWh based on DEA estimations and 5,417,825 MWh based on MLP estimations. These estimations suggest that there is significant potential for energy savings in the water sector. Although WaSCs performed better in terms of energy efficiency compared to WoCs, the higher savings in energy could still be achieved by WaSCs on average. This is attributed to the larger size of WaSCs compared to WoCs. On average, WoCs could potentially reduce their energy use by 53,261 MWh per year to align with the most energy-efficient companies in the industry. On the other hand, the average potential energy savings for WaSCs were estimated to be 97,949 MWh per year, indicating a greater scope for energy reduction.

Based on the volume of drinking water supplied by each water company and their energy efficiency, potential energy savings per cubic meter of water were estimated (Fig. 5). On average, the English and Welsh water companies could potentially save around 0.54 kWh/m^3 of water. However, significant variations exist between different water companies, with the minimum estimated value being 0.0 kWh/m^3 for the most energy-efficient company, and the maximum value being 2.05 kWh/m^3. The observed differences in potential energy savings per cubic meter of water are attributed to both the current energy use and the energy efficiency of water companies. The variability in potential energy savings is evident for both WaSCs and WoCs. According to the energy efficiency estimations from the MLP model, WoCs could potentially save an average of 0.24 kWh/m^3, while WaSCs have a higher potential for energy savings, with an average of 0.65 kWh/m^3.

These findings highlight the importance of considering energy efficiency in the water sector, as it has a direct impact on the energy consumption per unit of water supplied. By implementing energy-saving measures and improving energy efficiency, water companies can significantly reduce their energy use per cubic meter of water, leading to cost savings and environmental benefits.

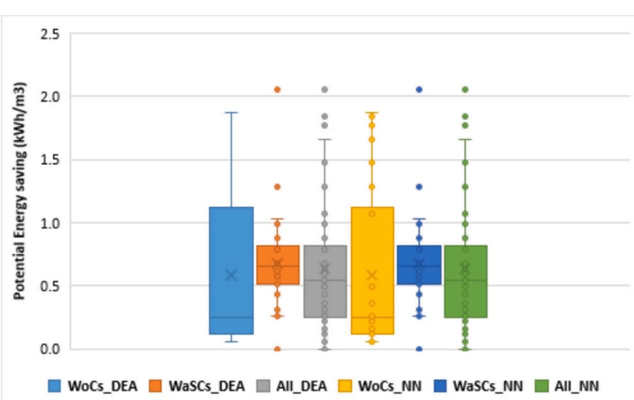

Fig. 5. Potential energy savings for water only companies (WoCs), water and sewerage companies (WaSCs) and all water companies (all) based on data envelopment analysis (DEA) and neural network (NN) and estimations.

4.4. Influence of environmental variables on energy efficiency of water companies

The influence of various operating characteristics on energy efficiency of water companies was analyzed using a regression tree and the results are presented in Fig. 6. In the regression tree, each branch represents a specific condition, and the mean predicted energy efficiency score is displayed at the bottom of each branch. Based on the results, it was found that three operating characteristics had the highest impact on influencing energy efficiency: treatment works when water is taken from groundwater (WGW), water treatment quality (wtq), and water taken from rivers (wriver). These factors were identified as significant predictors of energy efficiency in the water sector. To gain a more detailed understanding of the relationships between these operating characteristics and energy efficiency, please refer to the Appendix - Fig. 2 for a comprehensive visualization.[1]

According to the analysis (Fig. 6), the number of groundwater treatment works plays a significant role in energy efficiency. When the number of treatment works is below 12 on average, energy efficiency is relatively low at 0.14. This suggests that a higher energy input is required at the initial stages of the treatment process to ensure basic water treatment from groundwater sources. As the number of treatment works increases, energy can be used more efficiently. The complexity of the treatment process is also important. If >12 treatment works are needed to clean water from groundwater resources and <74 % of water receives high levels of treatment, the average energy efficiency score can reach 0.62. This indicates that utilities can benefit from economies of

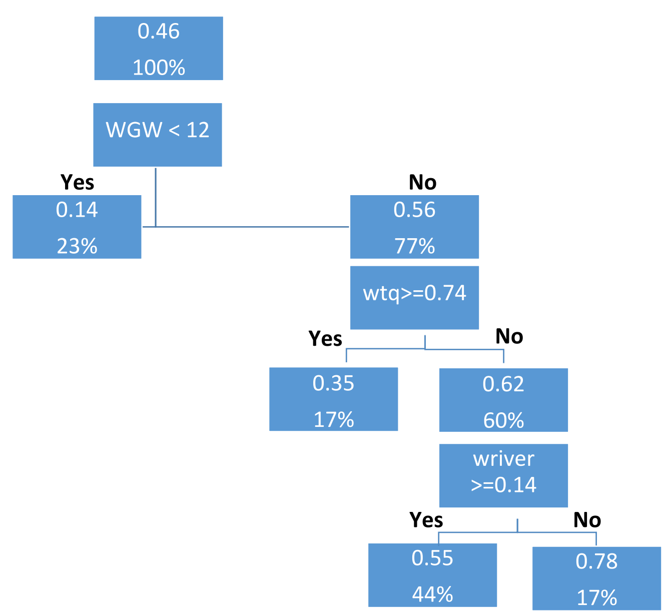

Fig. 6. Regression tree estimations to assess the influence of environmental variables on energy efficiency of water companies in England and Wales.

[1] Potential Energy Savings = (1 − Energy Efficiency Score) * Current Energy Use.

scale in high-level water treatment processes. Furthermore, the treatment of water from rivers can impact energy use. When 14 % of water is abstracted from rivers, energy efficiency reaches 0.55. However, if >75 % of water undergoes high-level treatment, energy requirements may be higher, resulting in a drop in energy efficiency to 0.35 on average.

The robustness of the results from the regression tee are checked with the ones obtained from the random forest regression. By applying the random forest algorithm, a more comprehensive understanding of the relationship between the explanatory variables and energy efficiency in water companies is obtained. We had six explanatory variables so the optimal value of *m* was set to 2. The number of trees was set to 3000 so that we have a stable RF regression [74]. The optimal number of trees, determined by reaching the minimum prediction error, was found to be 100. This suggests that after 100 trees, the random forest regression model's performance doesn't significantly improve. It's important to note that the choice of the number of trees can depend on the specific dataset and research context (additional details are provided at Appendix – Fig. 2).

According to the analysis (Fig. 7), treatment works for cleaning water from groundwater and surface resources, water treatment quality, and water taken from rivers were identified as the most important factors influencing energy efficiency. This suggests that the complexity and extent of treatment processes have a significant impact on energy consumption in water companies. Higher levels of water treatment are associated with higher energy use, indicating the need for optimizing treatment processes to improve energy efficiency. Additionally, the analysis highlights that water taken from boreholes and density (population per km of water main) also have an impact on energy performance, albeit to a lesser degree. Taking water from boreholes appears to have a positive influence on energy efficiency, possibly due to differences in the quality or accessibility of groundwater resources. Moreover, delivering water to densely populated areas seems to enhance energy efficiency, implying that it may be easier to manage energy consumption in urban areas compared to rural areas.

Indeed, the conclusions drawn from the random forest regression analysis align with the findings from the regression tree analysis, providing further support and robustness to the results. Both analyses highlight the significance of factors such as treatment works for different water sources (groundwater and surface resources), water treatment quality, water taken from rivers and boreholes, and population density in influencing energy efficiency in water companies. The consistency between the two analyses reinforces the importance of these factors in determining energy performance and provides a more comprehensive understanding of their impact. The findings emphasize the need for optimizing energy use during water abstraction from rivers and boreholes, as well as the importance of efficient treatment processes to minimize energy consumption. Moreover, the identification of population density as a contributing factor to energy efficiency suggests that tailored approaches might be required for different types of areas (urban vs. rural) to optimize energy usage in water services. This highlights the need for targeted strategies in managing energy resources based on the characteristics of the service area.

Water utilities and regulators have a crucial role to play to reduce energy inefficiencies. Some policies and actions they could carried out are as follows: i) invest in modernizing water infrastructure systems to minimize energy losses during water abstraction, treatment and, distribution; ii) implement water conservation and demand management programs to reduce the overall energy requirements for water provision; iii) conduct energy audits of water treatment and distribution systems to identify areas of high energy consumption and inefficiency. This helps in pinpointing opportunities for improvement and implementing energy-saving measures, such as upgrading equipment, optimizing operational processes, and reducing system losses; iv) provide financial incentives to encourage water utilities and consumers to adopt energy-efficient practices and technologies and; v) establish and enforce regulations that promote energy efficiency in the water sector. Set energy performance standards for water infrastructure, mandate energy audits, and require utilities to report on their energy consumption and efficiency measures.

5. Conclusions

Assessing the energy efficiency of water services holds significant significance for both society and the environment.

In contrast to previous studies, this research employs neural networks to evaluate the energy efficiency of water companies. The empirical application conducted for English and Welsh water companies evidenced their poor energy efficiency. The mean energy efficiency score was 0.441 which means that the potential energy savings could reach the level of 55.9 %. It is found that on average WaSCs performed better than WoCs since their mean energy efficiency scores were 0.559 and 0.239, respectively. Energy inefficiency in the provision of drinking water involves that on average water companies could save 0.54 kWh/m^3. It is evidenced that water treatment quality and topography had the major impact on energy efficiency. From a policy perspective, the findings of the MLP approach are proven to be a reliable alternative to traditional techniques used to estimate efficiency scores in the water sector.

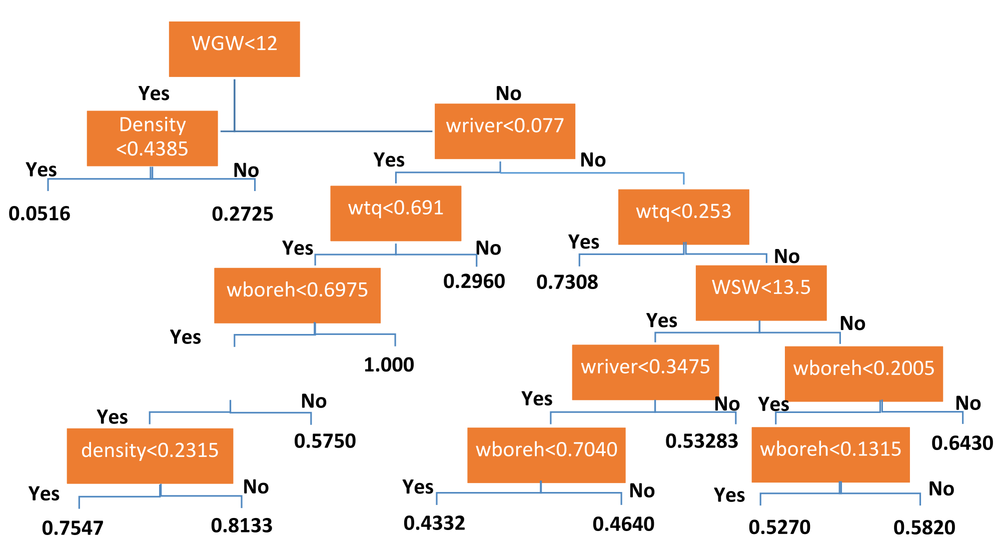

Fig. 7. Random forest regression tree.

CRediT authorship contribution statement

Maria Molinos-Senante: Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Alexandros Maziotis:** Writing – original draft, Software, Methodology, Conceptualization.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgement

This work has been supported by project CL-EI-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".

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jwpe.2024.105596.

References

[1] R.B. Sowby, Making waves: research to support water and wastewater utilities in the transition to a clean-energy future, Water Res. 233 (2023) 119739.
[2] A. Majid, M. Mortazavi-Naeini, J.W. Hall, Efficient pathways to zero-carbon energy use by water supply utilities: an example from London, UK, Environ. Res. Lett. 16 (2021) 105010.
[3] R.B. Sowby, S.J. Burian, Survey of energy requirements for public water supply in the United States, J. Am. Water Works Assoc. 109 (2017) E320–E330.
[4] E.S. Spang, A.J. Holguin, F.J. Loge, The estimated impact of California's urban water conservation mandate on electricity consumption and greenhouse gas emissions, Environ. Res. Lett. 13 (2018) 014016.
[5] S.J. Kenway, K.L. Lam, J. Stokes-Draut, K.T. Sanders, A.N. Binks, J. Bors, B. Head, G. Olsson, J.E. McMahon, Defining water-related energy for global comparison, clearer communication, and sharper policy, J. Clean. Prod. 236 (2019) 117502.
[6] K.L. Lam, S.J. Kenway, P.A. Lant, Energy use for water provision in cities, J. Clean. Prod. 143 (2017) 699–709.
[7] UN (United Nations), Sustainable development goals, Available at: https://sustainabledevelopment.un.org/?menu=1300, 2015.
[8] N.L. Walker, D. Styles, J. Gallagher, A.P. Williams, Aligning efficiency benchmarking with sustainable outcomes in the United Kingdom water sector, J. Environ. Manage. 287 (2021) 112317.
[9] Y. Wada, T. Gleeson, L. Esnault, Wedge approach to water stress, Nat. Geosci. 7 (2014) 615–617.
[10] A. Facchini, C. Kennedy, I. Stewart, R. Mele, The energy metabolism of megacities, Appl. Energy 186 (2017) 86–95.
[11] S.G. Rothausen, D. Conway, Greenhouse-gas emissions from energy use in the water sector, Nat. Clim. Chang. 1 (2011) 210.
[12] W. Wu, H.R. Maier, G.C. Dandy, M. Arora, A. Castelleti, The changing nature of the water–energy nexus in urban water supply systems: a critical review of changes and responses, J. Water Clim. Change 11 (4) (2020) 1095–1122.
[13] S. Ahmad, H. Jia, Z. Chen, Q. Li, C. Xu, Water-energy nexus and energy efficiency: a systematic analysis of urban water systems, Renew. Sustain. Energy Rev. 134 (2020) 110381.
[14] A.V. Mercedes Garcia, P.A. López-Jiménez, F.-J. Sánchez-Romero, M. Pérez-Sánchez, Objectives, keys and results in the water networks to reach the sustainable development goals, Water 13 (2021) 1268.
[15] K. Smith, S. Liu, Y. Liu, T. Chang, X. Wu, Impact of urban water supply on energy use in China: a provincial and national comparison, Mitig. Adapt. Strat. Glob. Chang. 21 (8) (2016) 1213–1233.
[16] M. Wakeel, B. Chen, T. Hayat, A. Alsaedi, B. Ahmad, Energy consumption for water use cycles in different countries: a review, Appl. Energy 178 (2016) 868–885.
[17] D. Zaman, M.K. Tiwari, A.K. Gupta, D. Sen, Performance indicators-based energy sustainability in urban water distribution networks: a state-of-art review and conceptual framework, Sustain. Cities Soc. 72 (2021) 103036.
[18] L. Liu, E. Lopez, L. Dueρas-Osorio, L. Stadler, Y. Xie, P.J.J. Alvarez, Q. Li, The importance of system configuration for distributed direct potable water reuse, Nat. Sustain. (2020) 1–8.
[19] Y. Liu, M.S. Mauter, Marginal energy intensity of water supply, Energy Environ. Sci. 14 (2021) 4533–4540.
[20] V. Rodríguez-Merchan, C. Ulloa-Tesser, C. Baeza, Y. Casas-Ledon, Evaluation of the Water–Energy nexus in the treatment of urban drinking water in Chile through exergy and environmental indicators, J. Clean. Prod. 317 (2021) 128494.
[21] J. Ananda, Productivity implications of the water-energy-emissions nexus: an empirical analysis of the drinking water and wastewater sector, J. Clean. Prod. 196 (2018) 1097–1105.
[22] M. Molinos-Senante, A. Maziotis, R. Sala-Garrido, M. Mocholi-Arce, Understanding water-energy nexus in drinking water provision: an eco-efficiency assessment of water companies, Water Res. 225 (2022) 119133.
[23] B.J. Cardoso, Á. Gomes, A.R. Gaspar, Barriers and drivers to energy efficiency in the Portuguese water sector: survey analysis, Appl. Energy 333 (2023) 120630.
[24] A. Maziotis, R. Sala-Garrido, M. Mocholi-Arce, M. Molinos-Senante, A comprehensive assessment of energy efficiency of wastewater treatment plants: an efficiency analysis tree approach, Sci. Total Environ. 885 (2023) 163539.
[25] R. Sala-Garrido, M. Mocholí-Arce, A. Maziotis, M. Molinos-Senante, Benchmarking the performance of water companies for regulatory purposes to improve its sustainability, npj Clean Water 6 (1) (2023) 1.
[26] T.B. Cetrulo, R.C. Marques, T.F. Malheiros, An analytical review of the efficiency of water and sanitation utilities in developing countries, Water Res. 161 (2019) 372–380.
[27] K.H. Goh, K.F. See, Twenty years of water utility benchmarking: a bibliometric analysis of emerging interest in water research and collaboration, J. Clean. Prod. 284 (2021) 124711.
[28] A. Maziotis, M. Molinos-Senante, The impact of model specification and environmental variables on measuring the overall technical efficiency of water and sewerage services: evidence from Chile, Struct. Chang. Econ. Dyn. 61 (2022) 191–198.
[29] S.A. Nyathikala, T. Jamasb, M. Llorca, M. Kulshrestha, Utility governance, incentives, and performance: evidence from India's urban water sector, Util. Policy 82 (2023) 101534.
[30] A. Robles-Velasco, M. Rodríguez-Palero, J. Muñuzuri, L. Onieva, Sustainable development and efficiency analysis of the major urban water utilities in Spain, Water (Switzerland) 14 (9) (2022) 1519.
[31] K.F. See, Exploring and analysing sources of technical efficiency in water supply services: some evidence from Southeast Asian public water utilities, Water Resour. Econ. 9 (2015) 23–44.
[32] A. Patel, A. Kethavath, N.L. Kushwaha, K.R. Sheetal, P.S. Renjith, Review of artificial intelligence and internet of things technologies in land and water management research during 1991–2021: a bibliometric analysis, Eng. Appl. Artif. Intel. 123 (2023) 106335.
[33] D. Santín, F.J. Delgado, A. Valiño, The measurement of technical efficiency: a neural network approach, Appl. Econ. 36 (6) (2004) 627–635.
[34] A. Azadeh, S.F. Ghaderi, S. Tarverdian, M. Saberi, Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption, Appl. Math Comput. 186 (2007) 1731–1741.
[35] A. Azadeh, M. Saberi, M. Anvari, An integrated artificial neural network algorithm for performance assessment and optimization of decision making units, Expert Syst. Appl. 37 (2010) 5688–5697.
[36] H. Liao, B. Wang, T. Weyman-Jones, Neural network based models for efficiency frontier analysis: an application to east Asian economies' growth decomposition, Glob. Econ. Rev. Perspect. East Asian Econ. Ind. 36 (4) (2007) 361–384.
[37] I.E. Tsolas, V. Charles, T. Gherman, Supporting better practice benchmarking: a DEA-ANN approach to bank branch performance assessment, Expert Syst. Appl. 160 (2020) 113599.
[38] L. Basnet, D. Brill, R. Ranjithan, K. Mahinthakumar, Supervised machine learning approaches for leak localization in water distribution systems: impact of complexities of leak characteristics, J. Water Resour. Plan. Manag. 149 (8) (2023) 04023032.
[39] A. Momeni, K.R. Piratla, A. Anderson, K. Chalil Madathil, D. Li, Stochastic model-based leakage prediction in water mains considering pipe condition uncertainties, Tunn. Undergr. Space Technol. 137 (2023) 105130.
[40] A.K.S. Ong, Y.T. Prasetyo, M.C.C. Sacro, R. Nadlifatin, K.P.E. Robas, Determination of factors affecting customer satisfaction towards "maynilad" water utility company: a structural equation modeling-deep learning neural network hybrid approach, Heliyon 9 (3) (2023) e13798.
[41] X. Tian, I. Vertommen, L. Tsiami, P. van Thienen, S. Paraskevopoulos, Automated customer complaint processing for water utilities based on natural language processing—case study of a Dutch water utility, Water (Switzerland) 14 (4) (2022) 674.
[42] Y. Deng, X. Zhou, J. Shen, G. Xiao, H. Hong, H. Lin, F. Wu, B.Q. Liao, New methods based on back propagation (BP) and radial basis function (RBF) artificial neural networks (ANNs) for predicting the occurrence of haloketones in tap water, Sci. Total Environ. 772 (2021) 145534.
[43] W.A. Fillah, D. Purwitasari, Prediction of water quality index using deep learning in mining company, in: Proceeding - 6th International Conference on Information Technology, Information Systems and Electrical Engineering: Applying Data Sciences and Artificial Intelligence Technologies for Environmental Sustainability, ICITISEE 2022, 2022, pp. 574–578.
[44] H. Hong, Z. Zhang, A. Guo, L. Shen, H. Sun, Y. Liang, F. Wu, H. Lin, Radial basis function artificial neural network (RBF ANN) as well as the hybrid method of RBF

ANN and grey relational analysis able to well predict trihalomethanes levels in tap water, J. Hydrol. 591 (2020) 125574.
[45] H. Lin, Q. Dai, L. Zheng, H. Hong, W. Deng, F. Wu, Radial basis function artificial neural network able to accurately predict disinfection by-product levels in tap water: taking haloacetic acids as a case study, Chemosphere 248 (2020) 125999.
[46] B. Li, L. Shen, Y. Zhao, W. Yu, H. Lin, C. Chen, Y. Li, Q. Zeng, Quantification of interfacial interaction related with adhesive membrane fouling by genetic algorithm back propagation (GABP) neural network, J. Colloid Interface Sci. 640 (2023) 110–120.
[47] B. Li, R. Yue, L. Shen, C. Gheng, L. Renjie, X. Yanchao, Z. Meijia, H. Hong, H. Lin, A novel method integrating response surface method with artificial neural network to optimize membrane fabrication for wastewater treatment, J. Clean. Prod. 376 (2022) 134236.
[48] A. Nafi, J. Brans, Prediction of water utility performance: the case of the water efficiency rate, Water (Switzerland) 10 (10) (2018) 1443.
[49] M. Molinos-Senante, A. Maziotis, Prediction of the efficiency in the water industry: an artificial neural network approach, Process Saf. Environ. Prot. 160 (2022) 41–48.
[50] L. Yi, H. Thomas, A decision support system for the environmental impact of ICT and ebusiness, Int. J. Inf. Technol. Decis. Mak. 8 (2) (2009) 361–377.
[51] H.-B. Kwon, J. Lee, Two-stage production modeling of large U.S. banks: a DEA-neural network approach, Expert Syst. Appl. 42 (2015) 6758–6766.
[52] M.R. Elkharbotly, M. Seddik, A. Khalifa, Toward sustainable water: prediction of non-revenue water via artificial neural network and multiple linear regression modelling approach in Egypt, Ain Shams Eng. J. 13 (5) (2022) 101673.
[53] A. Emrouznejad, E.A. Shale, A combined neural network and DEA for measuring efficiency of large scale data sets, Comput. Ind. Eng. 56 (2009) 249–254.
[54] A. Azadeh, A. Javanmardi, The impact of decision-making units features on efficiency by integration of data envelopment analysis, artificial neural network, fuzzy C-means and analysis of variance, Int. J. Operational Res. 7 (3) (2010) 387–411.
[55] F. Ciampi, N. Gordini, Small enterprise default prediction modeling through artificial neural networks: an empirical analysis of Italian small enterprises, J. Small Bus. Manag. 51 (1) (2013) 23–45.
[56] H.-B. Kwon, Exploring the predictive potential of artificial neural networks in conjunction with DEA in railroad performance modeling, Int. J. Prod. Econ. 183 (2017) 159–170.
[57] S. Jomthanachai, W.-P. Wong, C.-P. Lim, An application of data envelopment analysis and machine learning approach to risk management, IEEE Access 9 (2021) 85978–85994.
[58] O. Tosun, Using data envelopment analysis–neural network model to evaluate hospital efficiency, Int. J. Product. Quality Manag. 9 (2) (2012) 245–257.
[59] P. Hanafizadeh, H.R. Khedmatgozar, A. Emrouznejad, M. Derakhshan, Neural network DEA for measuring the efficiency of mutual funds, Int. J. Appl. Decision Sci. 7 (3) (2014) 255–269.
[60] N. Zhu, C. Zhu, A. Emrouznejad, A combined machine learning algorithms and DEA method for measuring and predicting the efficiency of Chinese manufacturing listed companies, J. Manag. Sci. Eng. 6 (4) (2020) 435–448.
[61] A. Athnassopoulos, S. Curram, A comparison of data envelopment analysis and artificial neural networks as tools for assessing the efficiency of decision-making units, J. Operational Res. Soc. 47 (1996) 1000–1016.
[62] H.-B. Kwon, J. Lee, J.J. Roh, Best performance modeling using complementary DEA-ANN approach: application to Japanese electronics manufacturing firms, BIJ 23 (3) (2016) 704–721.
[63] D. Celebi, D. Bayraktar, An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information, Expert Syst. Appl. 35 (4) (2008) 1698–1710.
[64] S.R. Nabavi, M. Abbasi, Black box modeling and multiobjective optimization of electrochemical ozone production process, Neural Comput. & Applic. 31 (2019) 957–968.
[65] F. Delgado, Measuring efficiency with neural networks. An application to public sector, Econ. Bull. 3 (15) (2005) 1–10.
[66] S. Rebai, F.B. Yahia, H. Essid, A graphically based machine learning approach to predict secondary schools performance in Tunisia, Socioecon. Plann. Sci. 70 (2019) 100724.
[67] G. James, D. Witten, R. Tibshirani, T. Hastie, An Introduction to Statistical Learning With Applications in R, Springer, New York, 2013.
[68] T.K. Ho, Random decision forests, in: Proceedings of the International Conference on Document Analysis and Recognition 1, ICDAR, 1995, pp. 278–282.
[69] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[70] M. Hallet, J. Fan, X. Su, R. Levine, M. Nunn, Random forest and variable importance for correlated survival data, with applications to tooth loss, Stat. Model. Int. J. 14 (6) (2014) 523–547.
[71] A. Nandy, P.K. Singh, Application of fuzzy DEA and machine learning algorithms in efficiency estimation of paddy producers of rural Eastern India, BIJ 28 (1) (2021) 229–248.
[72] R. Genuer, J.M. Poggi, C. Tueleau-Malot, N. Villa-Vialaneix, Random forests for big data, Big Data Res. 9 (2017) 28–46.
[73] E. Scornet, G. Biau, J.P. Vert, Consistency of random forests, Ann. Stat. 43 (4) (2015) 1716–1741.
[74] K. Thaker, V. Charles, A. Pant, T. Gherman, A DEA and random forest regression approach to studying bank efficiency and corporate governance, J. Oper. Res. Soc. 73 (6) (2021) 1258–1277.
[75] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning Data Mining, Inference, and Prediction, 2nd edition, Springer, 2009.
[76] S. Berg, R. Marques, Quantitative studies of water and sanitation utilities: a benchmarking literature survey, Water Policy 13 (5) (2011) 591–606.
[77] P. Carvalho, R.C. Marques, The influence of the operational environment on the efficiency of water utilities, J. Environ. Manage. 92 (10) (2011) 2698–2707.
[78] P. Carvalho, R.C. Marques, S. Berg, A meta-regression analysis of benchmarking studies on water utilities market structure, Util. Policy 21 (1) (2012) 40–49.
[79] F.S. Pinto, P. Simoes, R.C. Marques, Water services performance: do operational environment and quality factors count? Urban Water J. 14 (8) (2017) 773–781.
[80] A.P.P. da Silveira, H. Mata-Lima, Assessing energy efficiency in water utilities using long-term data analysis, Water Resour. Manag. 35 (9) (2021) 2763–2779.
[81] X. Dong, X. Zhang, S. Zeng, Measuring and explaining eco-efficiencies of wastewater treatment plants in China: an uncertainty analysis perspective, Water Res. 112 (2017) 195–207.
[82] B. Lin, K. Du, Measuring energy efficiency under heterogeneous technologies using a latent class stochastic frontier approach: an application to Chinese energy economy, Energy 76 (2014) 884–890.
[83] M. Molinos-Senante, R. Sala-Garrido, Evaluation of energy performance of drinking water treatment plants: use of energy intensity and energy efficiency metrics, Appl. Energy 229 (2018) 1095–1102.
[84] N.L. Walker, A.P. Williams, D. Styles, Key performance indicators to explain energy & economic efficiency across water utilities, and identifying suitable proxies, J. Environ. Manage. 269 (2020) 110810.
[85] H. Brea-Solis, S. Perelman, D.S. Saal, Regulatory incentives to water losses reduction: the case of England and Wales, J. Product. Anal. 47 (3) (2017) 259–276.
[86] R. Sala-Garrido, M. Mocholi-Arce, M. Molinos-Senante, A. Maziotis, Marginal abatement cost of carbon dioxide emissions in the provision of urban drinking water, Sustain. Prod. Consumption 25 (2021) 439–449.
[87] R. Sala-Garrido, M. Mocholi-Arce, M. Molinos-Senante, M. Smyrnakis, A. Maziotis, Eco-efficiency of the English and Welsh water companies: a cross performance assessment, Int. J. Environ. Res. Public Health 18 (2021) 2831.
[88] N.L. Walker, A. Norton, I. Harris, A.P. Williams, D. Styles, Economic and environmental efficiency of UK and Ireland water companies: influence of exogenous factors and rurality, J. Environ. Manage. 241 (2019) 363–373.
[89] D.S. Saal, D. Parker, T. Weyman-Jones, Determining the contribution of technical efficiency and scale change to productivity growth in the privatized English and Welsh water and sewerage industry: 1985–2000, J. Prod. Anal. 28 (1) (2007) 127–139.
[90] A. Villegas, M. Molinos-Senante, A. Maziotis, Impact of environmental variables on the efficiency of water companies in England and Wales: a double-bootstrap approach, Environ. Sci. Pollut. Res. 26 (2019) 31014–31025.
[91] Ofwat, PR19 Final Determinations: Securing Cost Efficiency Technical Appendix, The Water Services Regulation Authority, Birmingham, UK, 2019.
[92] Q. Wang, X. Sun, B.L. Golden, Using artificial neural networks to solve generalized orienteering problems, Intelligent Eng. Syst. Artif. Neural Netw. 6 (1996) 1063–1067.
[93] R. Nabavi, D. Salari, A. Niaei, M.T. Vakil-Baghmisheh, A neural network approach for prediction of main product yields in methanol to olefins process, Int. J. Chem. React. Eng. 7 (1) (2009) 1542–6580.
[94] S.R. Nabavi, M.J. Jafari, Z. Wang, Deep learning aided multi-objective optimization and multi-criteria decision making in thermal cracking process for olefines production, J. Taiwan Inst. Chem. Eng. 152 (2023) 105179.
[95] Y. Liu, M. Hejazi, P. Kyle, Y. He, D. Niyogi, Global and regional evaluation of energy for water, Environ. Sci. Technol. 50 (17)) (2016) 9736–9745.
[96] R. Hu, Q. Huang, H. Wang, J. He, S. Chang, Monitor-Based Spiking Recurrent Network for the Representation of Complex Dynamic Patterns, Int. J. Neural Syst. 29 (8) (2019) 1950006.