




ARTICLE



<https://doi.org/10.1057/s41599-023-02524-x>

OPEN

# Determinants of efficient water use and conservation in the Colombian manufacturing industry using machine learning

Carolina Henao <sup>1✉</sup>, Jenny Paola Lis-Gutiérrez<sup>1</sup>, Melissa Lis-Gutiérrez<sup>2</sup> & Janitza Ariza-Salazar<sup>1</sup>

Water is a fundamental aspect of achieving sustainable industries, and it is essential for firms to make efficient use of it. However, not all companies have sustainable water management practices in place. In this study, we aim to answer the research question: What factors influenced the efficient use and saving of water in the Colombian manufacturing industry in 2020? We used data from the Annual Manufacturing Survey and the Industrial Environmental Survey conducted by DANE to investigate this question. Our study grouped the variables into four categories: Environmental commitment, Risk mitigation, Innovation, and Location. We used two supervised learning machine-learning algorithms—decision tree and logit regression with LASSO regularization—to analyze the data. The results underscore those investments in wastewater treatment, total water consumption, and commitments to air and climate protection emerge as pivotal drivers for embracing sustainable practices. Moreover, the geographical location of companies significantly influences the likelihood of implementing water conservation initiatives. To enhance efficient water use and conservation in the Colombian manufacturing sector, several targeted public policies are recommended. These include offering fiscal and financial incentives to encourage investments in advanced wastewater treatment and reuse technologies. Additionally, advocating for specialized training and environmental education programs tailored for corporate personnel is crucial. Moreover, establishing initiatives for environmental certification and eco-labeling within the corporate sector can significantly promote sustainable water practices. Lastly, enforcing more stringent environmental regulations and standards, especially for industries like Textile, Metallurgy, and Manufacturing of rubber and plastic products, is essential to ensure responsible water management across the sector.

<sup>1</sup>Fundación Universitaria Konrad Lorenz, Bogotá, Colombia. <sup>2</sup>Universidad Nacional de Colombia, Bogotá, Colombia. ✉email: [lindac.henaor@konradlorenz.edu.co](mailto:lindac.henaor@konradlorenz.edu.co)

## Introduction

The linear economic model has resulted in the excessive depletion of natural resources, with products being purchased, used, and discarded, leading to environmental pollution from excessive waste (Akhtar et al., 2021). Efforts to counteract this environmental damage include advocating for production systems with reduced waste, such as the circular economy. This approach has the potential to drive social and economic development within a country (Penna & Fadel, 2021). Sustainable development hinges on prioritizing environmental, social, and economic elements. Industrial activity plays a pivotal role in conserving resources (Nodoushan et al., 2022), necessitating firms to develop models that ensure sustainability. A sustainable developmental trajectory, given limited resources, ensures the long-term viability of process industries (Jia et al., 2017).

Increasing environmental awareness has prompted governments to implement policies (Calise et al., 2019), while there is a global demand for investments from all quarters to enhance governance and management capabilities required to combat climate change, urbanization, and environmental degradation (Jiang, 2022).

Effectively managing water resources stands as one of the critical challenges in achieving sustainable development, entailing the delicate balance between economic interests and sustainable water stewardship (Chu et al., 2022; Ganesan & Rajiv, 2022; Song et al., 2022). Water scarcity has propelled countries to develop innovative technologies for their management (Gebre et al., 2021). With the world's population on the rise, substantial volumes of wastewater are generated by domestic, industrial, and agricultural activities. These waters can be treated and repurposed for other uses, aligning with the circular economy ethos (Issaoui et al., 2022). Rahim et al. (2020) identified four factors contributing to water scarcity: (1) uneven geographic distribution of water sources; (2) rapid urbanization leading to population and economic growth; (3) inadequate management of water resources; and (4) prolonged drought.

The burgeoning industrial sector necessitates ongoing evaluation of both the quantity and quality of water (Wang, 2022). Often, the industry's focus leans towards maximizing production rather than prioritizing water efficiency and conservation (Navarro-Ramírez et al., 2020), given the limited availability of water and the resulting wastewater from production processes (Hung et al., 2020; Akhtar et al., 2021). While conserving water resources can drive sustainable industrial advancement (Egbueri, 2022), it presents challenges in the production phases today (Ene et al., 2013). Achieving a green industrial transformation in certain areas calls for a systematic evaluation system and a comprehensive theoretical framework. Furthermore, understanding how various actors (users and management agencies) in a water resources system make decisions regarding resource utilization is paramount (Meempatta et al., 2023).

Empirical evidence underscores the correlation between water management and the industrial development levels of regions. Kosolapova et al. (2021) discovered a positive relationship between business innovation and effective practices in water resource utilization. Wang et al. (2019) scrutinized the ecological transformation of Chongqing Gepai Wire and Cable Co., Ltd., revealing that interdisciplinary methods are imperative in selecting and dynamically evaluating indicators. Cai et al. (2021) conducted a quantitative spatial analysis across 232 countries and territories to assess Sustainable Development Goal 6 (clean water and sanitation), finding that global water governance confronts challenges in resource management and international collaboration. Nations must implement appropriate political measures to attain universal and equitable access to potable water, sanitation services, and proper hygiene.

In Colombia, the water quality index highlights the industrial sector as a major contributor to the organic load in water bodies. This load is estimated at 51% in terms of biological oxygen demand and 62% in terms of chemical oxygen demand, in relation to the total load contributed by agriculture, industry, and domestic sources. Between 11.8 and 19 million Colombians are at high risk due to poor water quality (IDEAM, 2019).

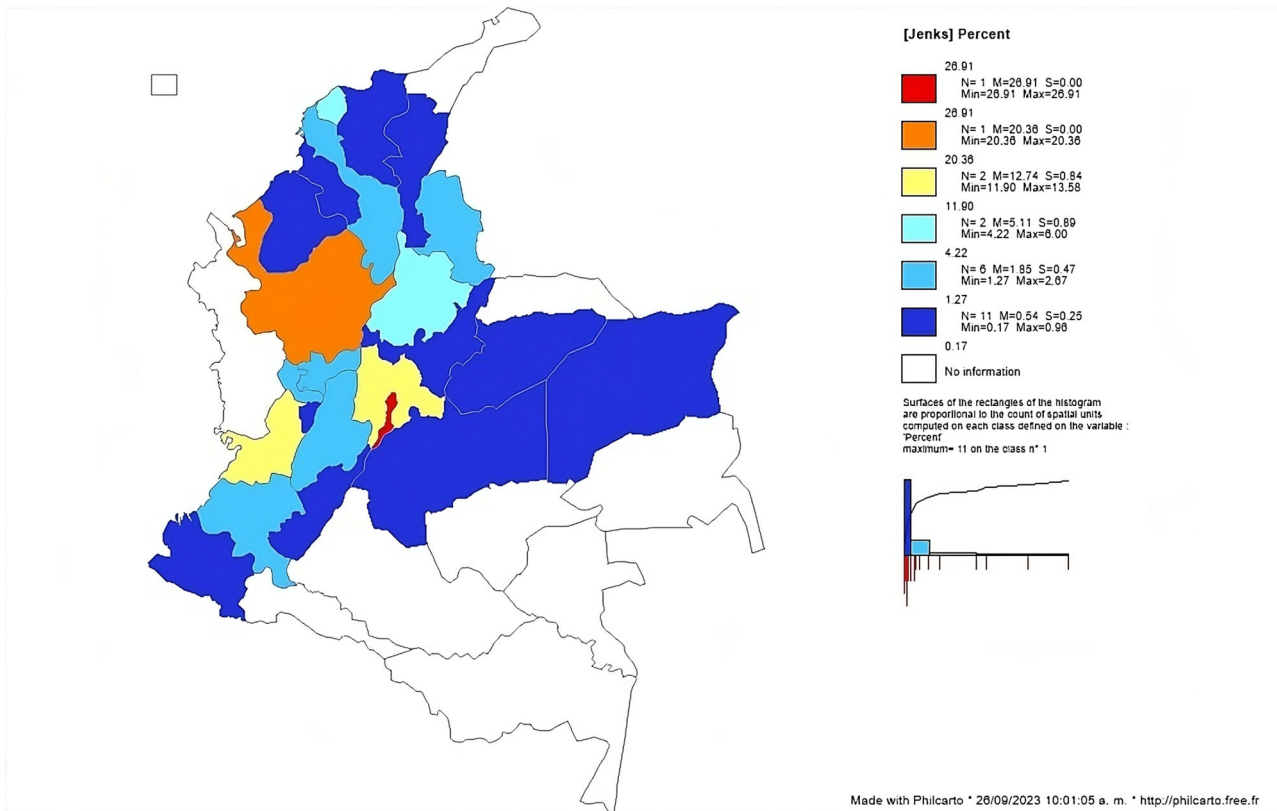
Despite empirical studies, Rahim et al. (2020) identified a research gap. Most studies rely on total water consumption data, with specific analyses on disaggregated water consumption behavior being scarce. This is even less prevalent in the case of studies encompassing the industrial sector (Burri et al., 2019).

The literature gap lies in the lack of comprehension regarding the connection between business performance and water conservation in the Colombian manufacturing industry. In this context, we aim to address the following question: What factors influenced the efficient use and conservation of water in the Colombian manufacturing industry in 2020? To address this query, we analyze data from the Annual Manufacturing Survey (Departamento Administrativo Nacional de Estadística, 2022a) and the Industrial Environmental Survey (Departamento Administrativo Nacional de Estadística, 2022b). We employ two supervised learning machine-learning algorithms: decision trees and logit regression with the Least Absolute Shrinkage and Selection Operator regularization algorithm.

## Methods

**Data.** In this research, we used data from the Annual Manufacturing Survey (Departamento Administrativo Nacional de Estadística, 2022a) and the Industrial Environmental Survey (Departamento Administrativo Nacional de Estadística, 2022b)<sup>1</sup>. These two surveys meet the requirements of OECD statistical operations. They are publicly accessible, and the data is anonymized. According to the documentation of the "Industry Environmental Survey" (Departamento Administrativo Nacional de Estadística, 2022b) the necessary elements to describe the sampling are:

- Unit of observation. Industrial establishments located in the national territory engaged in manufacturing activities that meet the parameters established in the target population.
- Reporting unit. The person responsible for providing the required information; typically, this is the professional in charge of process engineering or environmental engineering within the industrial establishment.
- Unit of analysis. Industrial establishments located in the national territory engaged in manufacturing activities that meet the parameters established in the target population.
- Sampling Unit. Industrial establishments located in the national territory engaged in manufacturing activities that meet the parameters established in the target population.
- Sampling method. The design of this sample is probabilistic with triple stratification: probabilistic because each establishment in the sample frame has a probability greater than zero of being selected, and triple stratification because a partition of the universe is done, first by regions, second by industrial divisions, and lastly, through the Hidiroglou algorithm, the population is divided into a stratum of forced inclusion and another of probabilistic inclusion. In some cases, according to statistical and thematic criteria, the population is classified only as forced inclusion. It is stratified according to the following variables:



**Fig. 1 Spatial distribution of the sample.** Note. Fig. 1 displays the spatial distribution of the sample. The observational unit was industrial establishments located in Colombia (7743), and the sample consisted of 2,917 establishments across 9 productive sectors. Own elaboration used Philcarto (Waniez, 2023).

- Region of the country to which the establishment belongs (Caribbean, Eastern, Central, Pacific, Bogota, and Amazonia-Orinoquia).
- Industrial activity carried out by the company<sup>2</sup>.
- Size of the establishment, in terms of employed personnel and industrial production.

The observational unit was industrial establishments located in Colombia (7743), and the sample consisted of 2917 establishments across 9 productive sectors. The sample design explained earlier allowed for estimates with an estimated coefficient of variation (sampling error) of less than 5% at the national level. Figure 1 displays the spatial distribution of the sample. It should be noted that there are no industrial activities in the departments to the east of Colombia.

**Variables.** Table 1 summarizes the information for each variable, with Prowater as the target variable. The variables were selected based on two criteria: (i) availability of information; and (ii) relationship with efficient and sustainable water use. The variables can be categorized into five groups: (i) Environmental commitment; (ii) Risk mitigation; (iii) Innovation; (iv) Location; and (v) Entrepreneurial performance. Sendawula et al. (2021) explored the relationship between environmental commitment and good water use practices in small and medium-sized companies. Similarly, Kosolapova et al. (2021) found a positive relationship between business innovation and sustainable water resource management. Chen et al. (2022) analyzed the effect of environmental liability insurance (ELI) on firms' environmental performance in developing economies.

**Ethical statements.** The data collection process employed in this study involved self-completion of an electronic form through a DANE website. Data transmission was conducted in real-time via the electronic form accessed through the website. It is important to note that access to the anonymized microdata for public use is provided free of charge and will be made available on the DANE website. This approach ensures transparency and accessibility while upholding ethical standards in data handling and dissemination.

**Algorithms.** Supervised learning algorithms have already been used to analyze water consumption and usage. For example, Pesantez et al. (2018) used eight variables to predict water demand 6 hours in advance and employed a three-layer neural model with three neurons in each layer. Similarly, Platsko & van Beek (2018) fitted models based on sets of regression trees using the LSBoost algorithm to predict outdoor residential water consumption. Deep learning techniques and algorithms have also been applied to urban water system planning and management problems (Fu et al., 2022).

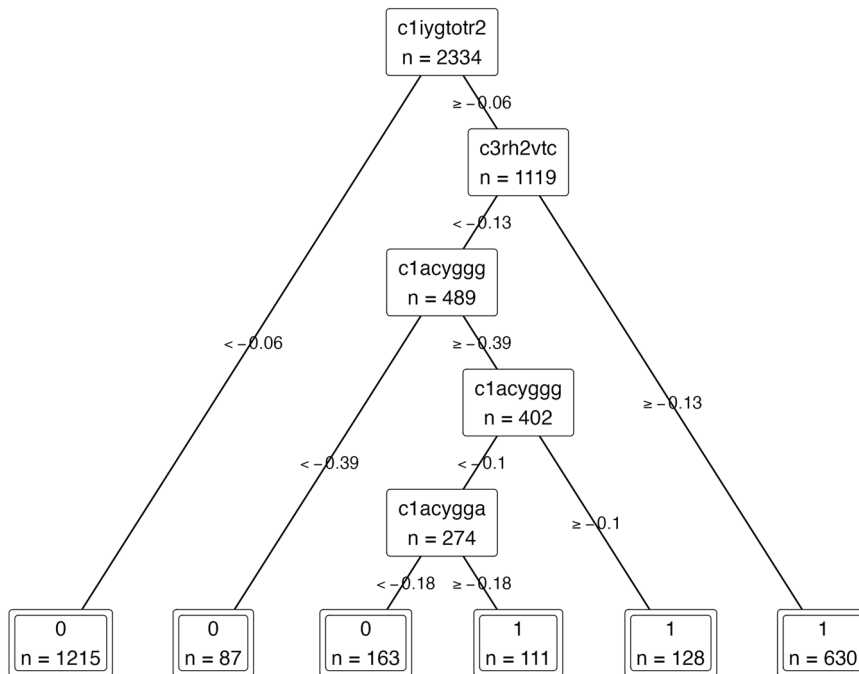
We applied two supervised learning algorithms: Logistic regression with LASSO regularization and decision tree. All variables were normalized because they had different units of measurement. As a result, standardization provides numerical stability, with a Z-score standardization of mean 0 and standard deviation 1.

We divided the sample into two parts. The algorithms used 80% of the sample for training and the remaining 20% for validation.

The training parameters for the classification decision tree were:

**Table 1 Variables.**

Name	Category	Description	Unit of measure
Prowater	Target	Dichotomous variable that assesses whether the companies: (i) had a program for water conservation and efficient use of this resource, (ii) treated the water discharged by the establishment, and (iii) reused water.	1. If the company complies with the three conditions. 0. If the company fails to meet at least one condition.
c4ga3ip1	Environmental commitment	Dichotomous variable that evaluates whether the companies have an environmental diagnosis as a planning tool	1. If the companies have an environmental diagnosis as a planning tool. 2. If the companies do not have an environmental diagnosis as a planning tool
activfi	Entrepreneurial performance	Fixed Assets. Numerical variable expressed in national currency units	Thousands of Colombian pesos
valagri	Entrepreneurial performance	Value added. Numerical variable expressed in national currency units	Thousands of Colombian pesos
valorven	Entrepreneurial performance	Sales. Numerical variable expressed in national currency units	Thousands of Colombian pesos
ppolizamb	Risk mitigation	Payments for Environmental Policies (environmental insurance). Numerical variable expressed in national currency units	Thousands of Colombian pesos
c1iygotr2	Risk mitigation	Investment for the prevention of wastewater generation. Numerical variable expressed in national currency units	Thousands of Colombian pesos
ingaprincon	Risk mitigation	Total investment and expenses for the prevention of contaminant infiltration. Numerical variable expressed in national currency units	Thousands of Colombian pesos
c1acygga	Environmental commitment	Payment of environmental licenses, permits, fees, and fines. Numerical variable expressed in national currency units	Thousands of Colombian pesos
c1acyggg	Environmental commitment	Expenditure on staff engaged in environmental protection activities. Numerical variable expressed in national currency units	Thousands of Colombian pesos
c1acygge	Innovation	Research and development (basic, experimental, or applied). Numerical variable expressed in national currency units	Thousands of Colombian pesos
c3rh2vtc	Environmental commitment	The total volume of water used by the property in m <sup>3</sup> /year	m <sup>3</sup> /year
c1inygrmye	Environmental commitment	Total investment and expenses for the Management of mineral and energy resources. Numerical variable expressed in national currency units	Thousands of Colombian pesos
c1iygotR1	Environmental commitment	Total investment and expenses for air and climate protection. Numerical variable expressed in national currency units	Thousands of Colombian pesos
c1iygotr3r2	Environmental commitment	Total investment and expenditure for the treatment and disposal of hazardous waste. Numerical variable expressed in national currency units	Thousands of Colombian pesos
codigo_actividad	Location	Nominal variable corresponding to the company's sector.	One-digit code (i) Food, beverages, and tobacco. (ii) Coking, manufacture of petroleum refining products and nuclear fuel. (iii) Manufacture of rubber and plastic products. (iv) Manufacture of chemical products and substances. (v) Manufacture of wood and cork products, manufacture of paper and paper products and publishing and printing activities. (vi) Manufacture of non-metallic mineral products. (vii) Metallurgy and manufacture of metal products. (viii) Textile, apparel, footwear and leather products. (ix) Other industrial divisions.
dept	Location	Nominal variable that corresponds to the name of the place where the company is located.	Two-digit code
filtcertecocapv	Environmental commitment	Dichotomous variable. Does the establishment have any environmental certification or eco-labeling, or is it implementing it?	1. Yes 0. No



**Fig. 2 Decision tree.** Note: Fig. 2 illustrates the estimated decision tree generated by the algorithm, defining three terminal nodes. own elaboration figure made in Japs 0.16.2.

- (i) the minimum number of observations per division was 20;
- (ii) the minimum number of observations in one or terminal was seven;
- (iii) the maximum depth of any node in the final tree was 30;
- (iv) the complexity of the parameters was 0.01 (any division not decreasing the overall unfit in a factor of this parameter is not attempted);
- (v) the seed was randomized; and
- (vi) all processes were randomized.

The classification decision tree assumptions are:

- (i) the target variable is a nominal or ordinal variable; and
- (ii) predictors are scalar, nominal, or ordinal variables.

Chernozhukov and Wei (2016) most that Lasso implemented one plugin estimator:

$$\lambda \log it = \frac{c}{2n^{1/2}} \Phi^{-1} \left[ 1 - \frac{1.1}{2 \max[n, p \ln n]} \right]$$

*n* is the sample size, and *p* is the number of covariates in the model.

**Results**

**Decision tree.** Figure 2 illustrates the estimated decision tree generated by the algorithm, defining three terminal nodes<sup>3</sup>. The algorithm demonstrates robust adjustment indicators, with a correct predictions-to-total predictions ratio of 0.81. The harmonic mean of accuracy and recovery scores also stands at 0.81, further underlining the algorithm’s performance. Additionally, the area under the curve (AUC) reaches a value of 0.79.

The most influential predictors for ranking establishments’ programs for efficient water use and conservation, ranked in order of importance, are as follows:

- Total investment and expenses for wastewater treatment.
- Total volume of water used by the facility.

Table 2 Relative Importance of Variables.	
Variable	Relative importance
c1iygotr2	27.539
c3rh2vtc	17.703
c1acyggg	15.509
c1acygga	12.558
activfi	12.433
c1iygotr1	12.401
valorven	1.711
valagri	0.080
c1inygrmye	0.036
c1iygotr3r2	0.031

Note: own elaboration made in Japs 0.16.2. (Intel).

- Personnel expenses dedicated to environmental protection activities.
- Payment of environmental licenses, permits, fees, and fines.
- Fixed assets.
- Total investment and expenses for air and climate protection (Table 2).

**Logistic regression with LASSO regularization algorithm.** The Logistic regression model with LASSO regularization yielded a ROC (receiver operating characteristic) score of 0.83 for the training sample and 0.81 for the validation sample, indicating strong predictive performance<sup>4</sup>. This algorithm identified twelve significant predictors (Table 3).

The results of the logit regression analysis reveal predictors that either decreased or increased the probability of establishments having a program for efficient water use and saving.

Predictors decreasing probability:

**Table 3 Logistic regression with LASSO regularization algorithm.**

Predictor	Coefficient	Elasticity
0.filtcertecocapv	-0.4205039	-0.6567158
1. Food, beverages, and tobacco	0.4491343	1.56695509
11. Bogotá	-0.0181282	-0.9820351
13. Bolívar	-0.0393538	-0.9614105
15. Boyacá	-0.0604993	-0.9412944
19. Cauca	0.0604259	1.06228888
2. Coking, manufacture of petroleum refining products and nuclear fuel	0.0723728	1.07505605
20. Cesar	0.0733447	1.07610141
25. Cundinamarca	0.0421088	1.04300795
3. Manufacture of rubber and plastic products.	-0.1663736	-0.8467298
4. Manufacture of chemical products and substances.	0.2031624	1.22527144
41. Huila	0.028787	1.02920535
47. Magdalena	-0.020114	-0.9800869
50. Meta	0.0686674	1.07107991
52. Nariño	-0.0073785	-0.9926487
6. Manufacture of non-metallic mineral products	0.0935644	1.09808132
66. Risaralda	0.0204832	1.02069442
68. Santander	-0.0197699	-0.9804242
7. Metallurgy and manufacture of metal products.	-0.0249021	-0.9754054
70. Sucre	0.0185345	1.01870733
73. Tolima	-0.0294656	-0.9709643
76. Valle del Cauca	0.1978713	1.21880552
8. Textile, apparel, footwear and leather products.	-0.0392153	-0.9615437
8. Atlántico	0.0340927	1.03468052
85. Casanare	0.0807537	1.08410385
9. Other industrial divisions	-0.0640447	-0.9379631
99. Vichada	-0.0099147	-0.9901343
c1acygga	0.1058259	1.11162833
c1acygge	0.0329876	1.03353772
c1acyggg	0.8522824	2.34499296
c1inygrmye	0.0165909	1.01672929
c1iygtotr1	0.4206342	1.52292709
c1iygtotr3r2	-0.1787424	-0.8363213
c3rh2vtc	0.8445018	2.3268183
Ppolizamb	0.1588814	1.17219892
Valagri	0.2556796	1.29133892
0. filtcertecocapv	-0.4205039	-0.6567158

Note: own elaboration figure made in Stata (17).

- Absence of environmental certification or eco-labeling or being in the process of implementing it.
- Total investment and expenditures for the treatment and disposal of hazardous waste.
- Establishments in the manufacture of rubber and plastic products and textile, apparel, footwear, and leather products.
- Establishments located in Boyacá and Bolívar (see Table 3).

Predictors increasing probability:

- Expenditures for personnel dedicated to environmental protection activities.
- Total volume of water used by the establishment.
- Establishments in the food, beverages, and tobacco, and manufacture of chemical products and substances.
- Total investment and expenditures for air and climate protection.

- Value added.
- Payments for environmental policies.
- Establishments located in Valle del Cauca (see Table 3).

Table 4 provides the calculated elasticity, representing the percentage change in the dependent variable for every 1% change in the predictor variable. This facilitates the interpretation of the coefficients.

The results of logistic regression and decision tree analyses revealed the following factors to be the most influential:

- Investment in wastewater treatment: Companies that invest more in wastewater treatment are more likely to have water use efficiency programs.
- Volume of water used: Companies that use more water are more likely to implement water use efficiency programs.
- Investment in air and climate protection: Companies that invest in air and climate protection are more likely to implement water use efficiency programs.
- Geographical location: Companies located in certain regions, such as Valle del Cauca, are more likely to implement water use efficiency programs.
- Investment in personnel dedicated to environmental protection: Companies that invest in personnel dedicated to environmental protection are more likely to implement water use efficiency programs.

In Fig. 3, the departments with a direct relationship to efficient water use and conservation (in red) are: Sucre, Risaralda, Huila, Atlántico, Cundinamarca, Cauca, Meta, Cesar, Casanare, and Valle del Cauca. On the other hand, those with a negative relationship (in blue) are: Nariño, Vichada, Bogotá, Santander, Magdalena, Tolima, Bolívar, and Boyacá.

These findings suggest that companies that are more committed to environmental protection and sustainability are more likely to implement water use efficiency programs. Additionally, the findings suggest that geographical location and industry type may play a role in the adoption of these programs.

Now, to foster efficient water use and conservation in the Colombian manufacturing industry, the following public policies can be enacted:

- Offer fiscal and financial incentives to encourage investments in wastewater treatment and reuse technologies.
- Advocate for training and environmental education programs tailored for corporate staff.
- Establish initiatives for environmental certification and eco-labeling within the corporate sector.
- Enforce more stringent environmental regulations and standards for businesses, for example: (i) Textile, (ii) Metallurgy and manufacture of metal products, and (iii) Manufacture of rubber and plastic products.

**Discussion**

Colombia faces a considerable challenge in achieving sustainable development, as 40% of surveyed firms lack programs focused on efficient water use. This is particularly concerning given that the industrial sector is the primary contributor to water pollution in the country. The findings of this study shed light on the intricate relationship between water use efficiency and the manufacturing industry in Colombia. The decision tree and logistic regression model, coupled with LASSO regularization, have provided invaluable insights into the influential factors affecting water use efficiency within the Colombian manufacturing sector. The decision tree illuminated that variables such as total investment in wastewater treatment and the overall volume of water utilized by

**Table 4 Interpretation of elasticities.**

Predictor	Elasticity	Elasticity Interpretation
0.filtcertecocapv	-0.6567	For every 1% increase in the probability of not having environmental certification, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.6567%.
1. Food, beverages, and tobacco	1.567	For every 1% increase in the probability of belonging to the category of food, beverages, and tobacco, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.567%.
11. Bogotá	-0.982	For every 1% increase in the probability of being in Bogotá, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.982%.
13. Bolívar	-0.961	For every 1% increase in the probability of being in Bolívar, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.961%.
15. Boyacá	-0.941	For every 1% increase in the probability of being in Boyacá, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.941%.
19. Cauca	1.062	For every 1% increase in the probability of being in Cauca, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.062%.
2. Coking, manufacture of petroleum refining...	1.075	For every 1% increase in the probability of belonging to the category of coking and manufacturing of petroleum refining products and nuclear fuels, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.075%.
20. Cesar	1.076	For every 1% increase in the probability of being in Cesar, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.076%.
25.Cundinamarca	1.043	For every 1% increase in the probability of being in Cundinamarca, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.043%.
3. Manufacture of rubber and plastic products.	-0.847	For every 1% increase in the probability of belonging to the category of manufacture of rubber and plastic products, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.847%.
4. Manufacture of chemical products and substan...	1.225	For every 1% increase in the probability of belonging to the category of manufacture of chemical products and substances, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.225%.
41. Huila	1.029	For every 1% increase in the probability of being in Huila, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.029%.
47. Magdalena	-0.98	For every 1% increase in the probability of being in Magdalena, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.98%.
50. Meta	1.071	For every 1% increase in the probability of being in Meta, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.071%.
52. Nariño	-0.993	For every 1% increase in the probability of being in Nariño, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.993%.
6. Manufacture of non-metallic mineral products	1.098	For every 1% increase in the probability of belonging to the category of manufacture of non-metallic mineral products, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.098%.
66. Risaralda	1.021	For every 1% increase in the probability of being in Risaralda, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.021%.
68. Santander	-0.98	For every 1% increase in the probability of being in Santander, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.98%.
7. Metallurgy and manufacture of metal products.	-0.975	For every 1% increase in the probability of belonging to the category of metallurgy and manufacture of metal products, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.975%.
70. Sucre	1.019	For every 1% increase in the probability of being in Sucre, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.019%.
73. Tolima	-0.971	For every 1% increase in the probability of being in Tolima, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.971%.
76. Valle del Cauca	1.219	For every 1% increase in the probability of being in Valle del Cauca, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.219%.
8. Textile, apparel, footwear and leather products.	-0.962	For every 1% increase in the probability of belonging to the category of textile, apparel, footwear, and leather products, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.962%.
8. Atlántico	1.035	For every 1% increase in the probability of being in Atlántico, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.035%.
85. Casanare	1.084	For every 1% increase in the probability of being in Casanare, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.084%.
9. Other industrial divisions	-0.938	For every 1% increase in the probability of belonging to other industrial divisions, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.938%.
99. Vichada	-0.99	For every 1% increase in the probability of being in Vichada, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.99%.
clacygga	1.112	For every 1% increase in the probability of making payments for environmental licenses, permits, fees, and fines, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.112%.

**Table 4 (continued)**

Predictor	Elasticity	Elasticity Interpretation
c1acygge	1.034	For every 1% rise in the probability of expenditure on research and development (including basic, experimental, or applied, expressed in national currency units), the likelihood of the company meeting all three criteria for the target variable (Prowater) increases by 1.034%.
c1acyggg	2.345	For every 1% increase in the probability of spending on staff engaged in environmental protection activities, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 2.345%.
clinygrmye	1.017	For every 1% increase in the probability of investing in the management of mineral and energy resources, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.017%.
cliygtotr1	1.523	For every 1% increase in the probability of investing in air and climate protection, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.523%.
cliygtotr3r2	-0.836	For every 1% increase in the probability of investing and spending on the treatment and disposal of hazardous waste, the probability that the company meets all three requirements of the Target variable (Prowater) decreases by 0.836%.
c3rh2vtc	2.327	For every 1% increase in the total volume of water used by the property in cubic meters per year, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 2.327%.
Ppolizamb	1.172	For every 1% increase in payments for environmental policies (environmental insurance), the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.172%.
Valagri	1.291	For every 1% increase in value added expressed in units of national currency, the probability that the company meets all three requirements of the Target variable (Prowater) increases by 1.291%.

Note: own elaboration.

facilities emerged as the most potent predictors for classifying companies' programs regarding efficient water use and conservation. This discovery aligns seamlessly with existing literature emphasizing the significance of investments in water treatment technologies and the adoption of sound water resource management practices (Chu et al., 2022; Gebre et al., 2021).

Additionally, the logistic regression model identified specific predictors that either heightened or lessened the likelihood of companies implementing water-efficient programs. Notably, substantial investment in air and climate protection, along with the total volume of water consumed, surfaced as pivotal drivers for embracing sustainable practices. This underscores the critical importance of considering both economic and environmental dimensions in business decision-making pertaining to water management (Wang et al., 2019).

Furthermore, we observed that the geographical location of companies significantly influences the adoption of water conservation practices. Companies situated in specific regions, such as: Sucre, Risaralda, Huila, Atlántico, Cundinamarca, Cauca, Meta, Cesar, Casanare, and Valle del Cauca, demonstrated a higher propensity for implementing water-efficient programs, indicating the substantial influence of local and regional factors in business decision-making processes. Departments with a positive relationship to efficient water use were not necessarily the places with the highest industrial development in Colombia, contrasting with findings by Kosolapova et al. (2021).

It is imperative to highlight that this study also unearthed a notable gap in the existing literature. Despite the growing awareness of water management and its pivotal role in the industry, there is a conspicuous dearth of studies delving into disaggregated water consumption behavior, particularly within the context of the Colombian manufacturing industry (Burri et al., 2019). This underscores the need for further research to gain a comprehensive understanding of the intricate dynamics and underlying factors influencing water management within this sector.

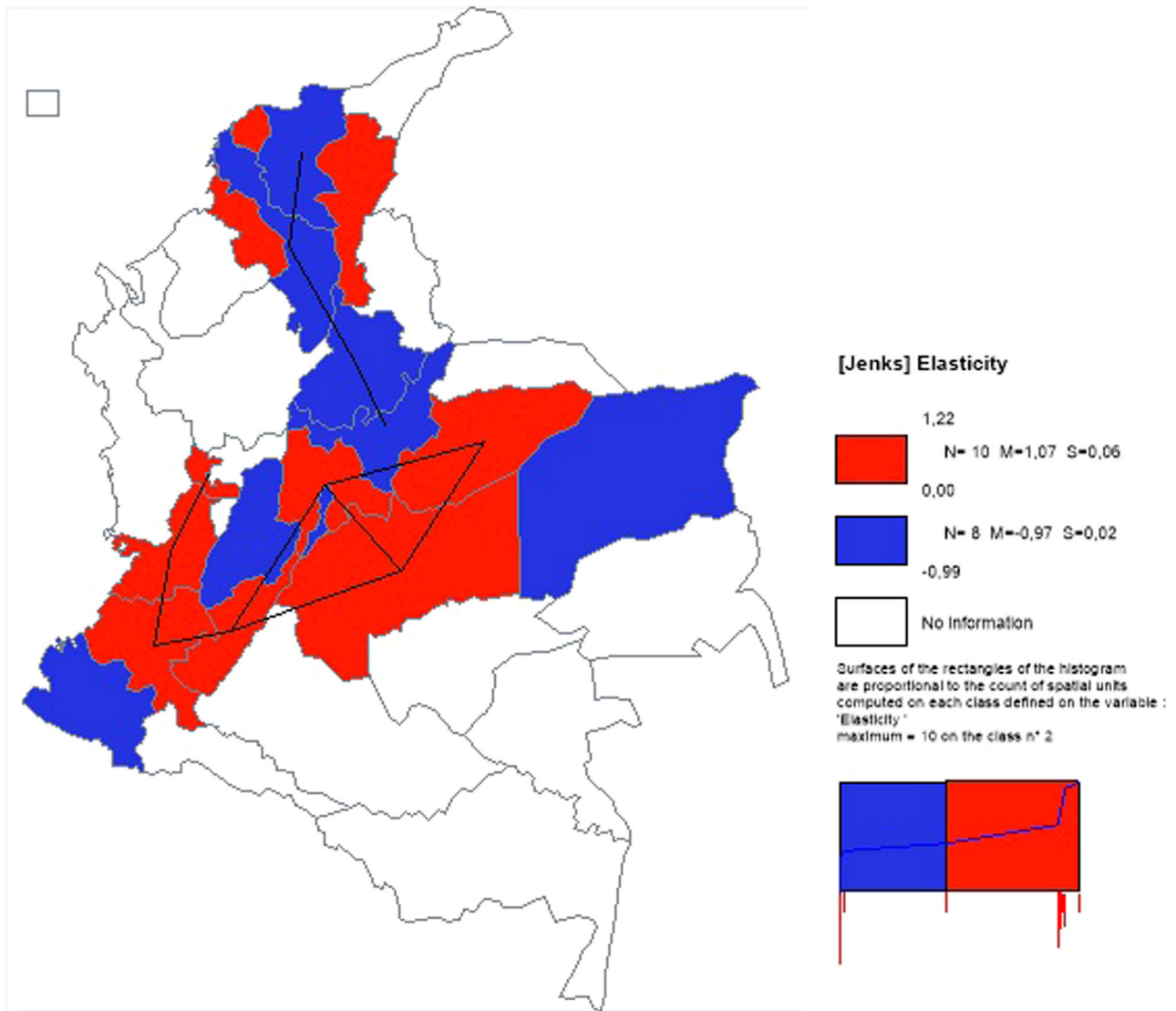
Assegid et al. (2022) underscore that the substantial water consumption and wastewater production within the textile sector aligns with our findings, indicating that companies in the Colombian textile, apparel, footwear, and leather products sector exhibit less efficient water resource management practices. Similarly, the global food, beverages, and tobacco sector, accounting for 70% of worldwide water consumption, demonstrates a greater likelihood of implementing programs for water efficiency and conservation among Colombian firms. It is noteworthy that firms adopting water management programs often experience short-term financial gains.

The coefficient and its positive sign for c3rh2vtc corroborate the observations of Chen & Cheung (2021), who noted that companies with high water usage benefit from water management programs, particularly in cost reduction through water recirculation. In this instance, there is a direct correlation between high water usage and its efficient management.

Conversely, the economic investment in technology for industrial water treatment and reuse (c1acygge) offers a longer-term return on investment. This may explain why companies choose to make such investments, albeit with a relatively modest effect (Navarro-Ramírez et al., 2020). The willingness of companies to invest in personnel dedicated to environmental protection activities (c1acyggg) emerges as the most influential factor in driving efficient water use within facilities. Finally, there exists a notable and positive correlation between conscientious water use and (i) payments investing in climate energy savings (cliygtotr1) and (ii) payments for environmental policies (ppolizamb).

It's important to note that not all surveyed companies have programs in place to ensure the efficient use of water resources, as investments in these programs often yield returns in the long-term. This lack of commitment to environmental certifications and best practices may be attributed to the abundance of water resources in Colombia, particularly in cities where scarcity is not a prevailing concern. Consequently, firms may lack incentives to implement efficient water management plans.





**Fig. 3 Elasticity by geographical location.** Note: Fig. 3 illustrates departments directly associated with effective water use and conservation in red, while those displaying a negative relationship are highlighted in blue. The diagram was created using Philcarto (Waniez, 2023).

The findings of this study underscore the environmental challenges confronting the Colombian government and underscore the need for sector-specific policies to address these issues. However, the economic interests of firms can sometimes impede investments in responsible water resource management programs. This challenge has also been emphasized in prior studies by Chu et al. (2022), Ganesan & Rajiv (2022), and Sound et al. (2022).

**Conclusion**

This paper substantiates that efficient water utilization in Colombia is intricately tied to companies’ willingness to invest, affirming Jiang’s (2022) assertions. Committing resources to water resource management not only strengthens a firm’s ability to mitigate environmental degradation risks, particularly those stemming from irresponsible water consumption in industrial operations but also bolsters its overall environmental stewardship.

The study has successfully identified sectors that are most likely to have robust resource management plans and at least two commendable practices for wastewater treatment and water reuse. These sectors encompass (i) Food, beverages, and tobacco, (ii) Manufacture of chemical products and substances, (iii) Manufacture of

non-metallic mineral products, and (iv) Coking, manufacture of petroleum refining products, and nuclear fuel. Conversely, sectors that demonstrate a divergence from these practices include (i) Manufacture of rubber and plastic products, (ii) Textile, apparel, footwear, and leather products, (iii) Metallurgy and manufacture of metal products, and (iv) Other industrial divisions.

In predicting efficient water usage, the study found a significant predictor in the inclusion of an environmental diagnosis by Colombian industrial firms as a planning instrument. These findings align with the discoveries of Wang et al. (2019). Furthermore, the study’s algorithms reinforce the results of Navarro et al. (2020), illustrating the interdependence between firms’ readiness to optimize water usage and their financial capacity.

Moreover, Colombian manufacturing enterprises should strongly contemplate the adoption of water use efficiency programs. These initiatives not only aid in cost reduction but also elevate their environmental performance, contributing significantly to the sustainability of Colombia’s water resources.

Based on these findings, the study advocates for the incorporation of the following points into public policies aimed at safeguarding water resources:

- Encourage all firms to establish a program for efficient water use.
- Develop targeted plans for sectors with the highest pollution contributions.
- Evaluate a firm's water footprint and offer tax incentives to those effectively mitigating their water impact and achieving neutrality.
- There is a need for government-led initiatives to enhance awareness among businesses and the public about the critical importance of water conservation.
- Encourage the widespread adoption of circular economy principles within the industrial sector.

The study's limitations include reliance on data from a single country and an exclusive focus on quantitative aspects. Future research endeavors could significantly augment this study by incorporating a qualitative analysis phase with a specific focus on economic sectors.

### Data availability

The data is publicly accessible and can be consulted at the following links <https://microdatos.dane.gov.co/index.php/catalog/724> and <https://microdatos.dane.gov.co/index.php/catalog/746> (Departamento Administrativo Nacional de Estadística (2022a); Departamento Administrativo Nacional de Estadística (2022b)).

Received: 12 April 2023; Accepted: 12 December 2023;

Published online: 02 January 2024

### Notes

- 1 The data is publicly accessible and can be consulted at the following links <https://microdatos.dane.gov.co/index.php/catalog/724> and <https://microdatos.dane.gov.co/index.php/catalog/746> (DANE, 2022a, 2022b).
- 2 The nine sectors are: (i) Food, beverages, and tobacco; (ii) Coking, manufacture of petroleum refining products and nuclear fuel; (iii) Manufacture of rubber and plastic products; (iv) Manufacture of chemical products and substances; (v) Manufacture of wood and cork products, manufacture of paper and paper products and publishing and printing activities; (vi) Manufacture of non-metallic mineral products; (vii) Metallurgy and manufacture of metal products; (viii) Textile, apparel, footwear and leather products; (ix) Other industrial divisions.
- 3 A decision tree is a non-parametric supervised learning algorithm. This can be used for both regression and classification tasks. A decision tree has a hierarchical structure (has a root node and branches). The predictor variables (features) are used, until arriving at the possible results, with a specific probability of occurrence and looking for homogeneous subsets. This algorithm has among its advantages: ease of interpretation, flexibility, and the possibility of application to different types of data.
- 4 A logit model with the LASSO regularization algorithm is a regression with a binary dependent variable; this model has a procedure of shrinking some parameters to zero, thus selecting the variables and minimizing a cross-validation parameter. It is possible to determine the variables that contribute to the prediction of the target variable. The sign of the coefficients shows whether the relationship of the predictor with the target variable is direct or inverse.

### References

- Akhtar N, Syakir Ishak MI, Bhawani SA, Umar K (2021) Various natural and anthropogenic factors responsible for water quality degradation: a review. *Water* 13(19):2660. <https://doi.org/10.3390/w13192660>
- Burri NM, Weatherl R, Moeck C, Schirmer M (2019) A review of threats to groundwater quality in the Anthropocene. *Sci Total Environ* 684:136–154. <https://doi.org/10.1016/j.scitotenv.2019.05.236>
- Cai J, Zhao D, Varis O (2021) Match words with deeds: Curbing water risk with the Sustainable Development Goal 6 index. *J Clean Prod* 318:128509. <https://doi.org/10.1016/j.jclepro.2021.128509>
- Calise F, Vicidomini M, Costa M, Wang Q, Østergaard PA, Duić N (2019) Toward an efficient and sustainable use of energy in industries and cities. *Energies* 12(16):3150. <https://doi.org/10.3390/en12163150>
- Chen Z, Cheung KCK (2021) Deal or no deal on water-saving program? Exploring the optimal interval for government subsidies and internal incentives. *Water Environ Res* 93(2):274–305. <https://doi.org/10.1002/wer.1411>
- Chen S, Ding X, Lou P, Song H (2022) New evidence of moral hazard: Environmental liability insurance and firms' environmental performance. *J Risk Insur* 89(3):581–613
- Chu Y, Wang Y, Zhang Z, Dai S (2022) Decoupling of economic growth and industrial water use in Hubei province: From an ecological–economic interaction perspective. *Sustainability* 14(20):13338. <https://doi.org/10.3390/su142013338>
- Departamento Administrativo Nacional de Estadística (2022a) Encuesta Ambiental Industrial-EAI–2020. <https://microdatos.dane.gov.co/index.php/catalog/746>
- Departamento Administrativo Nacional de Estadística (2022b) Encuesta Anual Manufacturera–EAM–2020. <https://microdatos.dane.gov.co/index.php/catalog/724>
- Egbueri JC (2022) Predicting and analysing the quality of water resources for industrial purposes using integrated data-intelligent algorithms. *Groundwater Sustain Dev* 18:100794. <https://doi.org/10.1016/j.gsd.2022.100794>
- Ene SA, Teodosiu C, Robu B, Volf I (2013) Water footprint assessment in the winemaking industry: a case study for a Romanian medium size production plant. *J Clean Prod* 43:122–135. <https://doi.org/10.1016/j.jclepro.2012.11.051>
- Ganesan S, Rajiv A (2022) Role of science in environmental conservation leading to sustainable development. *Phys Sci Rev.* <https://doi.org/10.1515/psr-2022-0146>
- Gebre SL, Cattrysse D, Van Orshoven J (2021) Multi-criteria decision-making methods to address water allocation problems: a systematic review. *Water* 13(2):125. <https://doi.org/10.3390/w13020125>
- Fu G, Jin Y, Sun S, Yuan Z, Butler D (2022) The role of deep learning in urban water management: a critical review. *Water Res* 223:118973. <https://doi.org/10.1016/j.watres.2022.118973>
- Hung Y, Aziz HA, Ramli SF, Paul HH, Huhnke CR, Adesanmi BM (2020) Chemical waste and allied products. *Water Environ Res* 92(10):1504–1509. <https://doi.org/10.1002/wer.1399>
- IDEAM (2019) Estudio Nacional del Agua 2018. IDEAM, Bogotá. 452 pp
- Issaoui M, Jellali S, Zorpas AA, Dutournie P (2022) Membrane technology for sustainable water resources management: challenges and future projections. *Sustain Chem Pharm* 25:100590. <https://doi.org/10.1016/j.scp.2021.100590>
- Jia X, Foo DCY, Tan RR, Li Z (2017) Sustainable development paths for resource-constrained process industries. *Resour Conserv Recy* 119:1–3. <https://doi.org/10.1016/j.resconrec.2016.11.004>
- Jiang Y (2022) Financing water investment for global sustainable development: challenges, innovation, and governance strategies. *Sustain Dev* sd.2412. <https://doi.org/10.1002/sd.2412>
- Kosolapova NA, Matveeva LG, Nikitaeva AY, Molapisi L (2021) The rational use of water resources in the strategy of industry 4. 0. *Water Resour Manag* 35(9):3023–3041. <https://doi.org/10.1007/s11269-021-02889-1>
- Meempatta L, Webb JA, Keogh LA, Horne AC, Stewardson MJ (2023) Exploring the role and decision-making behaviour of irrigation water supply authorities in Australia. *Int J Water Resour Dev* 39(2):314–336. <https://doi.org/10.1080/07900627.2021.1982680>
- Navarro-Ramírez V, Ramírez-Hernández J, Gil-Samaniego M, Eliana Rodríguez-Burguño J (2020) Methodological frameworks to assess sustainable water resources management in industry: a review. *Ecol Indic* 119:106819. <https://doi.org/10.1016/j.ecolind.2020.106819>
- Nodoushan HJ, Padash A, nadoushan AJ (2022) Modeling of industrial strategy based on sustainable development approaches. *Int J Environ Res* 16(3):32. <https://doi.org/10.1007/s41742-022-00406-9>
- Penna LB, Fadel ALC (2021) Circular economy as an alternative to better use of natural/economic resources and to contribute to sustainable development goals: a case study in pig farming. *Int J Hum Sci Res* 1(1):2–14. <https://doi.org/10.22533/at.ed.5582121104>
- Pesantez JEB, Berglund EZ, Kaza N (2018) Modeling and forecasting short-term water demand reported by smart meters. In: Proceedings of the 1st International WDSA/CCWI 2018 Joint Conference, Kingston, ON, Canada, 23–25 July 2018
- Platsko V, van Beek P (2018) Identification, prediction, and explanation of outdoor residential water consumption using smart meter data:(078). In: WDSA/CCWI Joint Conference Proceedings. vol. 1. <https://ojs.library.queensu.ca/index.php/wdsa-ccw/article/view/12162>
- Rahim MS, Nguyen KA, Stewart RA, Giurco D, Blumenstein M (2020) Machine learning and data analytic techniques in digital water metering: a review. *Water* 12(1):294. <https://doi.org/10.3390/w12010294>
- Sendawula K, Bagire V, Mbidde CI, Turyakira P (2021) Environmental commitment and environmental sustainability practices of manufacturing small and medium enterprises in Uganda. *J Enterp Commun People Place Glob Econ* 15(4):588–607
- Song M, Tao W, Shang Y, Zhao X (2022) Spatiotemporal characteristics and influencing factors of China's urban water resource utilization efficiency from

- the perspective of sustainable development. *J Clean Prod* 338:130649. <https://doi.org/10.1016/j.jclepro.2022.130649>
- Waniez P (2023) Philcarto [software]. Available at <http://philcarto.free.fr>
- Wang M, Lin X, Yu L (2019) Comprehensive evaluation of green transportation in Chongqing main urban area based on sustainable development theory. *Syst Sci Control Eng* 7(1):369–378. <https://doi.org/10.1080/21642583.2019.1681032>
- Wang X (2022) Design and development of an integrated environmental waste management system with a sustainable solution. *Water Supply* 22(8):6516–6531. <https://doi.org/10.2166/ws.2022.241>

### Author contributions

CH contributed to the conceptualization and analysis of data for the paper. JPL-G contributed to the interpretation of the data for the paper. ML-G critically reviewed it for important intellectual content. JA-S contributed to the writing of the first draft of the paper. All authors who contributed to the manuscript read and approved the submitted version.

### Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Ethical approval

Ethical approval was not required as the study did not involve human participants.

### Informed consent

Informed consent was not required as the study did not involve human participants.

### Additional information

**Correspondence** and requests for materials should be addressed to Carolina Henao.

**Reprints and permission 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024