ALFA
Revista de Investigación en Ciencias Agronómicas y Veterinarias

ALFA. Revista de Investigación en Ciencias Agronómicas y Veterinarias
enero - abril 2026 / Volumen 10, Número 28
ISSN: 2664-0902 / ISSN-L: 2664-0902
https://revistaalfa.org
pp. 68 - 89

Clasificación de la calidad del agua subterránea utilizada en la agricultura mediante aprendizaje automático con datos desbalanceados

Classification of the quality of groundwater used in agriculture, using machine learning in unbalanced data

Classificação da qualidade da água subterrânea utilizada na agricultura, utilizando machine learning em dados desequilibrados

Basthean Pino-Cabezas[1]
bpinoc@unjbg.edu.pe

Edwin Pino-Vargas[2]
epinov@unjbg.edu.pe

Edilberto Mamani-López[1]
emamanil@unjbg.edu.pe

Edgar Taya-Acosta[3]
etayaa@unjbg.edu.pe

Fredy Cabrera-Olivera[4]
fcabrerao@unjbg.edu.pe

ARTÍCULO ORIGINAL

[1]**Faculty of Engineering, Academic Department of Chemistry and Chemical Engineering, Universidad Nacional Jorge Basadre Grohmann. Tacna, Perú**
[2]**Department of Civil Engineering, Jorge Basadre Grohmann National University. Tacna, Perú**
[3]**Department of Computer Engineering and Systems, Jorge Basadre Grohmann National University, Tacna, Perú**
[4]**Department of Geological Engineering-Geotechnics, Jorge Basadre National University. Tacna, Perú**

Escanea en tu dispositivo móvil o revisa este artículo en:
https://doi.org/10.33996/revistaalfa.v10i28.421

Artículo recibido: 16 de octubre 2025 / **Arbitrado:** 24 de noviembre 2025 / **Publicado:** 7 de enero 2026

RESUMEN

En zonas costeras áridas donde la sobreexplotación y la intrusión marina amenazan tanto la calidad como la disponibilidad del agua subterránea, esta es fundamental para la agricultura. Se utilizaron índices hidroquímicos y modelos de aprendizaje automático (AA) para determinar si el agua subterránea del acuífero Caplina (sur de Perú) era apta para el riego. El Índice de Calidad del Agua de Riego (ICAR) se calculó utilizando siete iones principales: Na^+, Ca^{2+}, Mg^{2+}, K^+, Cl^-, SO_4^{2-} y HCO_3^-. Este índice también se empleó para entrenar algoritmos de clasificación supervisada en un esquema binario (zonas críticas y no críticas). El clasificador XGBoost resultó ser el mejor modelo de los evaluados, con una puntuación F1 de 0,897, un área bajo la curva ROC (AUC-ROC) de 0,968 y una precisión de 0,927 mediante validación cruzada dejando uno fuera (n = 41). El análisis de sensibilidad reveló que los predictores más efectivos fueron Na^+, Ca^{2+}, Mg^{2+} y K^+. Esto significa que el intercambio iónico y la congelación del agua ocurrieron simultáneamente. La precisión del modelo se mantuvo estable, ya que el número de predicciones disminuyó de siete a cuatro iones, mientras que el costo del monitoreo se redujo hasta en un 43 %. Esto demuestra que la red hidrológica puede mejorarse. Una combinación de modelado basado en datos y análisis geoquímico reveló indicios tempranos de salinización del agua en la zona costera del estuario. Este resultado demuestra la eficacia de los enfoques basados en aprendizaje automático (ML) como sistemas de alerta temprana para la exploración de aguas subterráneas en áreas con recursos limitados.

Palabras clave: Calidad del agua subterránea; Agua de riego; Aprendizaje automático; Acuífero Caplina; Agricultura

ABSTRACT

In dry coastal areas where overuse and marine intrusion threaten both quality and availability, groundwater is very important for farming. We used hydrochemical indices and machine learning (ML) models to find out if the groundwater in the Caplina aquifer (southern Peru) was good for irrigation. The Irrigation Water Quality Index (IWQI) was calculated using seven major ions: Na^+, Ca^{2+}, Mg^{2+}, K^+, Cl^-, SO_4^{2-}, and HCO_3^-. It was also used to train supervised classification algorithms in a binary scheme (critical and non-critical zones). The XGBoost classifier was the best overall model tested, with an F1 score of 0.897, a ROC-AUC score of 0.968, and an accuracy score of 0.927 via leave-one-out cross-validation (n = 41). Sensitivity analysis revealed that the most effective predictors were Na^+, Ca^{2+}, Mg^{2+}, and K^+. This means that ion exchange and water freezing occurred simultaneously. The model accuracy remained stable, as the number of predictions decreased from seven to four ions, while the monitoring cost was reduced by up to 43%. This shows that the hydrological network can be improved. A combination of data-driven modeling and geochemical analysis revealed early signs of water salinization in the coastal zone of the estuary. This result demonstrates the effectiveness of machine learning (ML)-based approaches as early warning systems for groundwater exploitation in resource-limited areas.

Key words: Groundwater Quality; Irrigation Water; Machine learning; Aquifer Caplina, Agriculture

RESUMO

Nas zonas costeiras secas, onde o uso excessivo e a intrusão marinha ameaçam tanto a qualidade como a disponibilidade da água, as águas subterrâneas são de importância primordial para a agricultura. Utilizámos índices hidroquímicos e modelos de aprendizagem automática (ML) para determinar se a água subterrânea do aquífero Caplina (sul do Peru) era adequada para a irrigação. O Índice de Qualidade da Água para Rega (IWQI) foi calculado utilizando sete iões principais: Na^+, Ca^{2+}, Mg^{2+}, K^+, Cl^-, SO_4^{2-} e HCO_3^-. Foi também utilizado para treinar algoritmos de classificação supervisionada num esquema binário (zonas críticas e não críticas). O classificador XGBoost apresentou o melhor desempenho global entre os modelos testados, com uma pontuação F1 de 0,897, uma pontuação ROC-AUC de 0,968 e uma precisão de 0,927 através de validação cruzada leave-one-out (n=41). A análise de sensibilidade revelou que os preditores mais eficazes foram o Na^+, Ca^{2+}, Mg^{2+} e K^+. Isto significa que a permuta iónica e o congelamento da água ocorreram simultaneamente. A precisão do modelo manteve-se estável, dado que o número de previsões diminuiu de sete para quatro iões, enquanto o custo da monitorização foi reduzido até 43%. Isto demonstra que a rede hidrológica pode ser melhorada. Uma combinação de modelação baseada em dados e análise geoquímica revelou sinais precoces de salinização da água na zona costeira do estuário. Este resultado demonstra a eficácia das abordagens baseadas em machine learning (ML) como sistemas de alerta precoce para a exploração de águas subterrâneas em áreas com recursos limitados.

Palavras-chave: Qualidade da água subterrânea; Água para rega; Aprendizagem de máquina; Aquífero Caplina; Agricultura

INTRODUCTION

Water resources are crucial for economic and social development worldwide, particularly in arid and semi-arid regions. In these areas, groundwater is often the primary source of water for domestic and industrial purposes, particularly agriculture, which is the world's largest water consumer (1). Unfortunately, over-extraction and human activities in coastal aquifers, such as seawater intrusion and salt deposition, are leading to ongoing problems that pose serious threats to the sustainability of our groundwater resources (2).

Furthermore, factors such as land-use change, climate variability, and agricultural expansion are exacerbating the situation, making water less suitable for irrigation and deteriorating water quality. These challenges not only threaten agricultural productivity but also the integrity of the ecosystem as a whole (3,4). Prolonged reliance on weak groundwater can damage soil structure and fertility, reduce crop yields, and exacerbate food security issues. This emphasizes the need for rigorous water quality monitoring, assessment, and forecasting. To ensure that this important resource is managed sustainably (5), to achieve this sustainability, multiple chemical assessments are required, including sodium absorption rate (SAR), residual sodium carbonate (RSC), sodium percentage (Na%), permeability index (PI), magnesium risk (MR), and Kelly index (KI), to evaluate the efficiency of irrigation water use (6).

One commonly used way to understand whether water is good for irrigation is to use the Irrigation Water Quality Index (IWQI), which combines multiple chemical parameters into a simple value. This method allows water managers to make sense of complex data sets and use them more easily for decision-making (7). In the past decade, machine learning (ML) techniques have been very useful in water chemistry research, as they increase the accuracy and adaptability of predictions by dealing with non-linear relationships and limited or ambiguous data sets (8,9). In a comparison of the two previously mentioned methods, the tree ensemble- based models demonstrated superior performance in predicting and classifying water quality, with special emphasis on sparse or irregular databases (10,11).

The Caplina aquifer, located in the extreme south of Peru, on the border with Chile and within the arid environment of the Atacama Desert, is the main source of water supply for human consumption and agriculture in the Tacna region. The expansion of the agricultural frontier —primarily olive cultivation —has increased water extraction, exacerbating the problems of marine intrusion and salinity observed in previous studies (12-14). All of these conditions reinforcement the need to develop predictive tools that integrate hydrochemical indicators and machine learning algorithms to support sustainable management and decision-making regarding groundwater use.

Therefore , the objectives of this study are: (i) to assess the hydrochemical composition and suitability of groundwater for irrigation using indices and graphical interpretations, (ii) to analyze the associations between water quality parameters to identify processes related to salinization and overexploitation, and (iii) to design machine learning-based classification models capable of predicting irrigation water suitability in a binary classification : critical zone and non- critical zone.

In addition, hydrochemical monitoring es being optimized, Evaluating reduced subsets of parameters (from two to seven) that maintain predictive capacity and reduce analysis costs. To our knowledge, this es the first study to combine hydrochemical approaches and artificial intelligence in the Caplina here, providing a scientific basis for more efficient and sustainable water resource coastal management arid areas.

METHODOLOGY

In this sense, the field campaign consisted of the sampling of 41 wells Figure 1, from the Caplina aquifer in September 2022. Before sampling, five volumes of water were extracted from each well. The following in situ measurements were made on water parameters: pH, dissolved oxygen (DO), electrical conductivity (EC), temperature, salinity, and total dissolved solids using a pre-calibrated YSI ProQuatro multiparameter. Samples were collected in pre-washed and rinsed 4-liter polyethylene bottles, filling them completely without leaving headspace to avoid bubbles. After that, samples were transported in ice-filled coolers to the laboratory at Jorge Basadre Grohmann University in Tacna.

Volumen 10, Nro. 28, enero - abril 2026
ISSN: 2664-0902 / ISSN: 2664-0902, www.revistaalfa.org

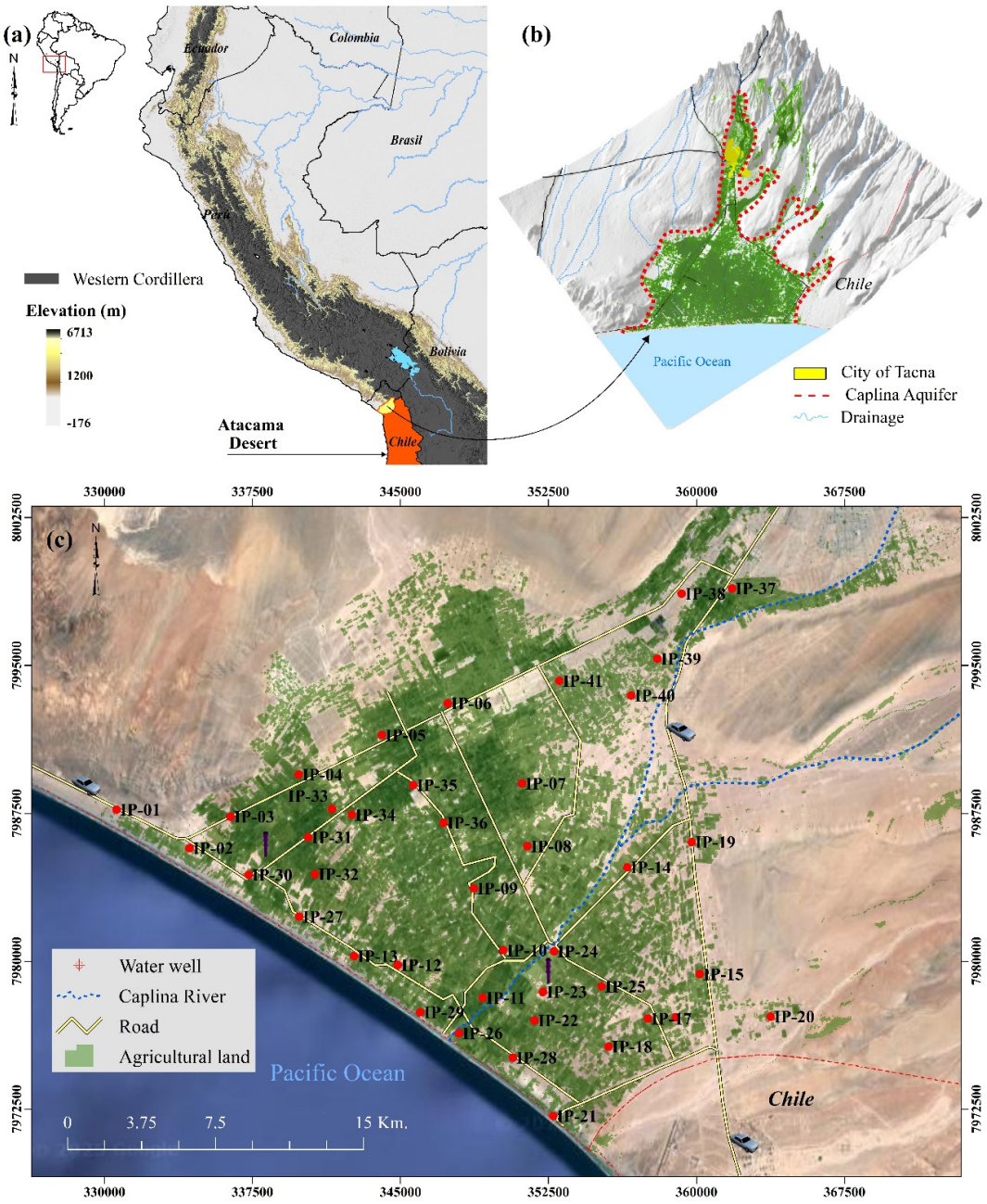

***Figura* 1**. Map showing the location of the study area and distribution of wells in he aquifer Caplina, Tacna, Peru. Note. Panel (a) shows the location of southern Peru and the Atacama Desert. Panel (b) represents the area of the Caplina aquifer in a 3D view. Panel (c) presents the location of 41 sampled wells in agricultural areas. Prepared by the authors.

The Caplina aquifer is located in southern Peru, primarily in the La Yarada district of Tacna province, and extends into northern Chile. This aquifer is an important source of groundwater for agricultural use and is located in a semi-arid region facing significant challenges related to water resource management due to overexploitation and marine intrusion (15).

The region has an arid coastal climate, with average annual rainfall of barely 6 mm and temperatures ranging between 18 °C and 20 °C (12, 15). Recharge originates mainly from rainfall and river infiltration, particularly from the Caplina and Uchusuma Rivers, whose headwaters lie in the Andean highlands where precipitation may reach up to 350 mm yr^{-1} (14).

In the laboratory, water samples were filtered using 0.45 µm pore-size membranes on the same day of collection. Bicarbonates were measured in unfiltered samples by titration to a pH of 4.3. All filtered samples were divided into different pre-washed and pre-rinsed polyethylene bottles: a 200 mL aliquot for the determination of major anions and another 200 mL aliquot preserved with ultrapure nitric acid (pH < 2) for the analysis of major and trace elements (16).

Major ions and trace elements were analyzed at the University of Waterloo Environmental Isotope Laboratory (UW-EIL), Canada. Major cations (Na^+, K^+, Ca^{2+}, and Mg^{2+}, B, Si) and trace elements (Li^+, Rb^+, Cs^+, Mo, Sr^{2+}, Ba^{2+}, Zr, Sb, As, and U) were measured by inductively coupled plasma mass spectrometry (ICP-MS) using a Perkin Elmer Sciex Elan 9000. Major anions (F^-, Cl^-, Br^-, NO_3^-, and SO_4^{2-}) were analyzed by ion chromatography using a Dionex ICS 1600 (16). The charge imbalance ranged between 0.1% and 6%. The saturation index (SI) of the mineral phases in the groundwater was determined using PHREEQC software (17). To verify the accuracy of the ICP-MS analyses, the international standard IV-STOCK-1643 (trace metals in water) was used. The relative differences between the standard measurements and the certified values of the elements considered in this study ranged between 0% and 11.8%.

Approach for evaluating groundwater quality in arid environments

The unweighted arithmetic water quality index (WQI UA) is used to reduce a large amount of data into numerical values that represent overall water quality, with which we can assess the suitability of groundwater for irrigation purposes (18).

$$WQI_{UA} = \frac{1}{n}\sum_{i=1}^{n} qi \quad (1)$$

For the case study, 6 parameters are used. With these, it will be possible to evaluate the suitability of the groundwater found in the aquifer (19).

Volumen 10, Nro. 28, enero - abril 2026
ISSN: 2664-0902 / ISSN: 2664-0902, www.revistaalfa.org

Sodium Adsorption Ratio:

$$SAR = \frac{Na^{+}}{\sqrt{\frac{(Ca^{2+} + Mg^{2+})}{2}}} \quad (2)$$

Evaluate the risk of sodification in the soil (20, 21).

Percent Sodium

$$\%Na = \left(\frac{Na^{+} + K^{+}}{Na^{+} + K^{+} + Ca^{2+} Mg^{2+}} \right) \times 100 \quad (3)$$

Indicates the percentage of sodium in relation to other cations (20, 21).

Permeability Index

$$PI = \left(\frac{Na^{+} + \sqrt{HCO_3^{-}}}{Ca^{2+} + Mg^{2+} + Na^{+}} \right) \times 100 \quad (4)$$

Measures soil permeability based on water composition (22).

Magnesium Hazard

$$MH = \left(\frac{Mg^{2+}}{Ca^{2+} + Mg^{2+}} \right) \times 100 \quad (5)$$

Evaluate the risk of magnesium toxicity in the soil (23).

Potential Salinity

$$PS = Cl^{-} + \frac{SO_4^{2-}}{2} \quad (6)$$

Measures the salinity potential in water (22).

Kelly Ratio:

$$KR = \frac{Na^{+}}{Ca^{2+} + Mg^{2+}} \quad (7)$$

Water alkalinity indicator (24).

Analysis Hydrochemical

The hydrochemical analysis was performed by characterizing the main cations and anions, expressed in meq/L. Descriptive statistics (mean, median, minimum, maximum, and standard deviation) were calculated to describe the variability in ionic composition. Additionally, the Pearson correlation between each ion and the Water Quality Index (WQI) was evaluated, which allowed us to identify the parameters with the greatest influence on groundwater quality and justify the subsequent selection of variables in the prediction models (25,26-30).

The Piper diagram will then be constructed with an ionic balance within a range of (< ±5%)

ensuring reliability. (25). This tool is widely used to classify facies and represent in a single graph the relative composition of the main cations and anions, facilitating the identification of the hydrochemical characteristics of an aquifer and the understanding of processes that influence its evolution (26).

In this study, the carbonate ion CO_3^{2-} was obtained from the bicarbonate HCO^{3-} and pH values by using the temperature field-adjusted dissociation constant pK_2, which ensured a more accurate representation of the anionic composition in the diagram (26).

$$\frac{\left[CO_3^{2-}\right]}{\left[HCO_3^{2-}\right]} = 10^{(pH-pK_2)} \qquad (8)$$

The CO_3^{2-} concentration es calculated by multiplying the above fraction by the total measured baking soda concentration. This procedure allows us to obtain the carbonate ion needed to complete the speciation of inorganic carbon and, therefore, to create the hydrochemical diagrams.

Machine Learning Modeling

Predictive modeling was developed with the purpose of classifying groundwater quality in a binary classification: critical zone and non- critical zone, according to the values of the irrigation water quality IWQI index. Supervised learning algorithms were applied, selected for their ability to handle non-linear relationships, small data sets and possible imbalances between classes (8, 32). The models used were: Random Forest (RF), Extreme Gradient Boosting (XGBoost), Gradient Boosting (GB), Support Vector Regression (SVR) and K- Nearest Neighbors (KNN) (28, 30).

The XGBoost and Gradient Boosting models are models that focus on the principle of sequentially assembling decision trees, revising previous predictions to iteratively correct errors. This process allows complex relationships between hydrochemical variables and IWQI to be found with high accuracy (28). The algorithms mentioned combine regularization mechanisms to avoid overfitting, demonstrating superior performance in the classification of environmental data with high collinearity and reduced sample size (27, 31). The SVR model, in contrast, was employed as both a linear and non-linear alternative, utilizing kernel functions, to assess its generalization capability in comparison to ensemble models (32). Finally, the KNN was used as a distance-based reference, taking into account the weighted average of the closest neighbors based on the best number of k, which was found through experimentation (29).

Because the sample size was small, leave-one-out (LOO) cross-validation was used this means that the model was trained with n–1 observations

Volumen 10, Nro. 28, enero - abril 2026
ISSN: 2664-0902 / ISSN: 2664-0902, www.revistaalfa.org

and evaluations, repeat the process until all samples are included by adding the item with the remaining observations and this method lessens bias and lets the model's predictive power be tested in a strong way (11). A standard scaling pipeline was used in each validation iteration for the SVR and KNN algorithms to avoid data leakage. The tree-based models, On the other hand, this approach is not scalable because it does not reduce the size of the variables, as non-critical categories have more samples than critical categories. Automatically assigning class weights (class weights = "balanced") and adjusting the decision criteria increase the model's sensitivity to critical areas (minority categories). This approach facilitates identification of the most critical scenarios for appropriate resource management (10).

To improve water chemistry monitoring, Pearson correlation analysis and variable importance analysis are used to analyze groups of 2-7 variables. This approach allows us to understand how the number of parameters used affects the model's ability to accurately predict water quality.

Evaluation Metrics in Water Quality Classification Models

We used classification parameters to evaluate the performance of machine learning models for groundwater quality detection. Due to class heterogeneity in the dataset, minority population-specific measures are preferred over general accuracy (28, 33). These include precision, sensitivity or recall, F1 score, ROC-AUC, and PR-AUC. These parameters were adopted because, using these, model performance could be evaluated even when a special component is unmatched, (34, 35).

Accuracy is the proportion of samples correctly classified as a class greater than all predicted values for that class. Clearly, it is a positive diagnosis. Sensitivity/memory represents the proportion of correctly identified nodes by the model out of the total number of nodes for that class. This becomes important in water resources management because the failure to identify critical areas (which is false negatives) will lead to improper irrigation decisions (35, 31). The F1 score represents a simultaneous average of precision and sensitivity; this quantifies both measures and thus makes the model suitable for evaluation under non-equilibrium conditions (33). Also, ROC-AUC and PR-AUC values were computed based on the classification ability. While the former computes the competency of the model in distinguishing between groups in different decisions, the latter provides valuable information in cases of intra-class inconsistency, reflecting the accuracy gained with increasing representation of the dominant

class (10). Then we use the confusion matrix as a representation of true positives and false positives and negatives.

Model Validation

Model performance was used to evaluate the model and also to make sure that there is no overfitting and the results obtained are reproducible. The dataset was small, and hydrochemical parameters varied; hence, this process was repeated until all observations were included, resulting in more accurate and less biased estimates of the total error (32, 35). LOO validation is ideal for environmental studies with limited datasets because it allows one to make use of all the available information while still separating the data into training and test sets (17, 28). The scale-sensitive models, including SVR and ANN, have been implemented at the pre-processing step.

This pre-processing used a predefined scale at each validation iteration to avoid data leakage and ensure that only the training data were used to derive normalization statistics. We did not prescale tree models such as XGBoost and Gradient Boosting since both are insensitive to the variability of the variables (8, 27, 28). We evaluated the models using bootstrap uncertainty analysis with 1000 random iterations. This allowed us to determine the confidence intervals of the most important metrics: accuracy, detection, F1 score, and ROC-AUC. Also, 1000 random prediction iterations were performed in determining the confidence interval of the critical metrics: accuracy, detection, F1 score, and ROC-AUC. This method allowed us to assess the robustness of each model to variations in the data and to statistically compare its performance across algorithms (10, 33). We derived the final performance metric from the mean of the validation test along with its standard deviation and confidence interval (27). This procedure allowed the identification of the best predictive model and the determination of the degree of its change with changes in training data.

Importance of Variables and Mapping

Mapping and Importance of Variables we used explainability analysis and feature importance analysis to figure out which hydrochemical variables had the biggest impact on separating the Caplina aquifer into critical and non-critical zones (29). To find out which hydrochemical factors had the biggest impact on separating critical and non-critical areas in the Caplina aquifer. The two methods for determining feature importance employed are permutation importance, which looks at the effect of each variable on the model's accuracy, and gain-based importance derived from ensemble techniques (31, 32). In tree-

based models (XGBoost and Gradient Boosting), the importance was estimated by quantifying either the performance penalty associated with permuting each variable or by measuring gain contribution at the node, a widely used approach in current water quality studies (11, 31). The results point out that the most informative ions are Na^+, Ca^{2+}, Mg^{2+}, and K^+, although Cl^- is relegated to a lower position because of its high collinearity with Na^+ ($r > 0.95$). Sodium is the dominant explanatory variable within the framework of salinization processes (11, 31, 33).

The relevance of Ca^{2+} and Mg^{2+} is consistent with ion exchange and carbonate dissolution processes, while K^+ serves as a complementary marker linked to agricultural practices and anthropogenic enrichment (11, 12).

To improve the individual interpretability of the forecasts, the SHAP (Shapley Additive Planations) method was applied. This method allows us to visualize the marginal effect (direction and magnitude) of each ion on the probability of entering the critical class, thus supporting the tractability of the model for management (33, 35). Finally, binary classification maps were developed in QGIS to represent the spatial distribution of the influencing variables and the delimitation of critical and non- critical zones, facilitating the prioritization of vulnerable areas and the optimization of monitoring with a reduced set of parameters without substantial loss of performance (30).

DEVELOPMENT AND DISCUSSION

Water Quality Assessment using $IWQI_{UA}$

In the Table 1, summarizes the distribution of wells by IWQI-UA category and their operational aggregation for the binary model. It is observed that “Excellent” = 16 wells and “Good” = 11 wells, which together make up the NON-CRITICAL ZONE (n = 27), while “Bad” = 8, “Very bad” = 3 and “Not suitable” = 3 make up the CRITICAL ZONE (n = 14), configuring a moderate imbalance that justifies prioritizing metrics sensitive to the minority class in the evaluation of the classifier, such as recall, F1 and PR-AUC. This binary partition adopts an IWQI threshold > 50 to define critical condition (Poor, Very Poor and Unsuitable), consistent with the practice reported in irrigation water assessments in arid and coastal environments (6,7).

Table 1. Distribution of wells by IWQI-UA category and their operational aggregation for the binary model.

WQI UA	Number of Wells	
Excellent	16	NON-CRITICAL ZONE (69.8%)
Good	11	
Bad	8	CRITICAL ZONE (30.2%)
Very Bad	3	
Not Suitable	3	

Note. The binary classification representing 69.8% and 30.2% of the total wells, respectively.

Hydrochemical Facies

The Piper diagram Figure 2, shows a system in evolution hydrochemistry (15). In the Table 2, group uncritical concentrated towards carbonate facies (Ca–Mg–HCO_3), indicative of recharge waters, and a gradient progressive towards saline facies in minor proportion (Na–Cl) where they are located samples reviews (“Bad”, “Very bad”, “Not suitable”), consistent with marine intrusion and salinization in the coastal sector (39, 40).

Table 2. Group uncritical concentrated towards carbonate facies.

Hydrochemical facies	Percentage	Type	Interpretation
Ca–Mg–HCO_3	68%	Dominant	Recharge waters, low TDS, dominate in the middle- upper zone of the aquifer.
Ca–Na–HCO_3 –Cl (mixed)	22%	Transition	They indicate exchange ionic progressive (Na^+ replacing Ca^{2+}), typical in stages salinization initials.
Na–Cl	10%	Minority	Represents wells further close to the coast, affected by marine mixing and overexploitation.

Note. The predominance of Ca–Mg–HCO_3 indicates recharge waters of low salinity, while the Na–Cl facies near the coast reflect progressive salinization and seawater intrusion processes.

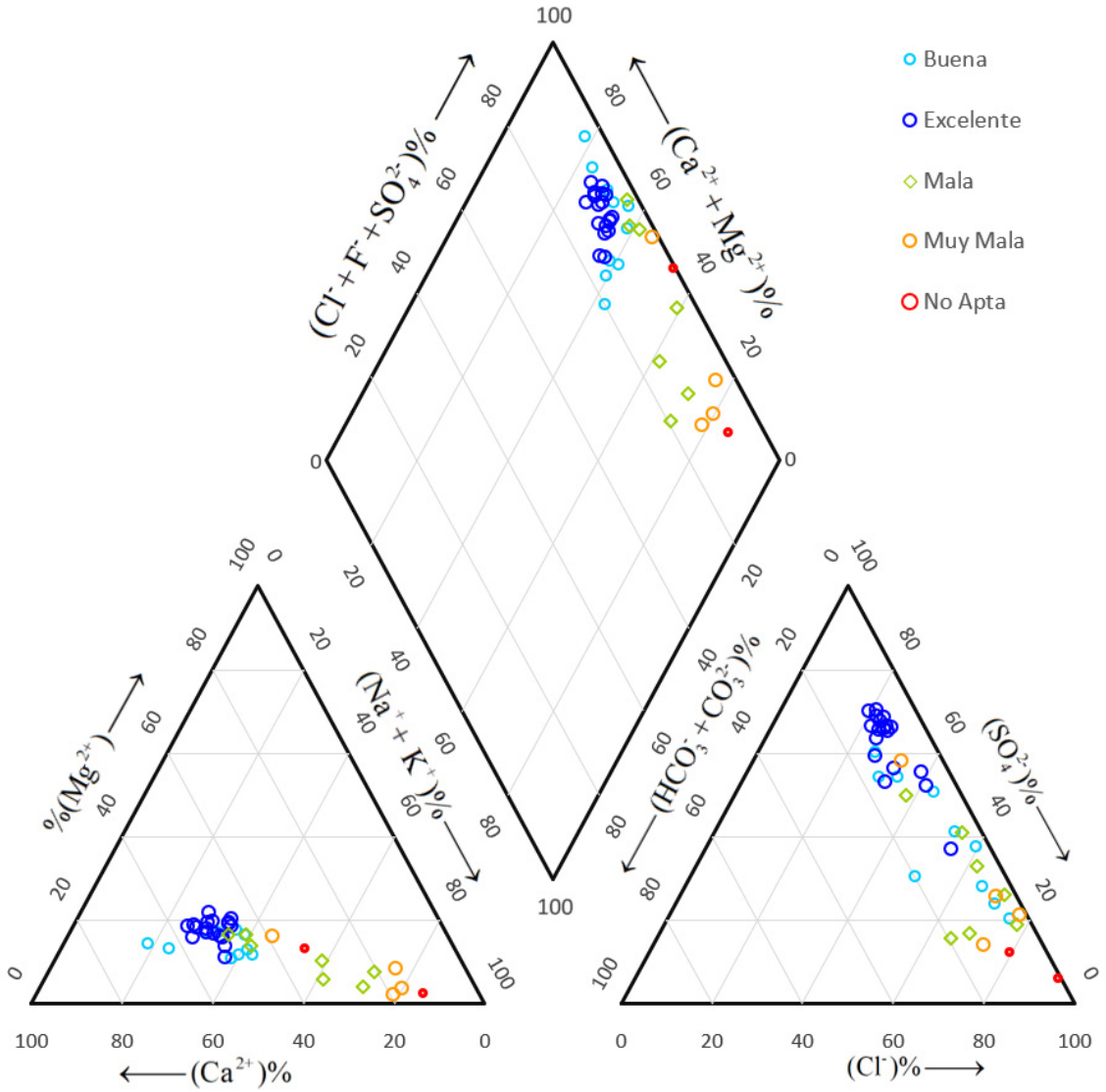

***Figura* 2**. Shows a system in evolution hydrochemistry. Note. The facies were classified based on major ion dominance using Piper diagram interpretation. Preparation Own using EXCEL.

The cationic triangle show a dominance of (Na^+, K^+) and a relative depletion of Ca^{2+}, a pattern compatible with ion exchange (replacement of Ca^{2+} by Na^+ in the exchange matrix), while in the anionic triangle an increase in Cl^- and a decrease in HCO_3^- + CO_3^{2-} are observed, reinforcing the marine/evaporite mixing signal (36, 38). Estimating CO_3^{2-} from HCO_3^- and pH using $pK_2(T)$ allowed for adequate representation of the carbonate vortex and improved sample projection in the diamond field, which is critical in waters with high pH (25).

Classification binary and performance of learning models automatic

Machine learning algorithms evaluated in the Figure 3, This research incorporated RF, SVM, KNN and XGBoost (30, 32). The best models had F1 scores between 0.88 and 0.90 and ROC-AUC scores above 0.96. This means they were very good at telling the difference between important and unimportant areas where XGBoost model stood out in particular with an F1 score of 0.897 and a ROC-AUC of 0.968 so this means that it is both accurate and sensitive, and it can also handle class imbalance (8, 33). Sensitivity analysis Figure 4, Shows that the Random Forest and XGBoost models correctly detected more than 90% of the important samples. The SVM model did the same, but with less variance (Recall = 0.86.

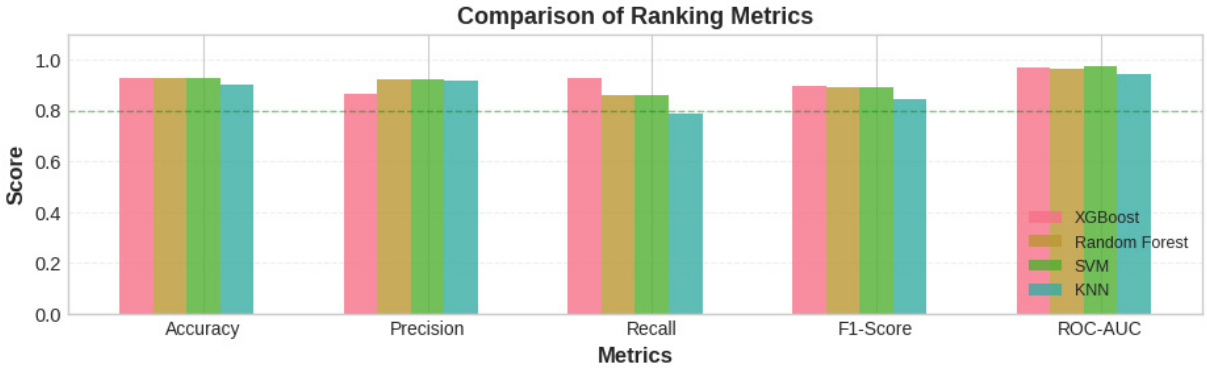

***Figura* 3**. Machine learning algorithms evaluated. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

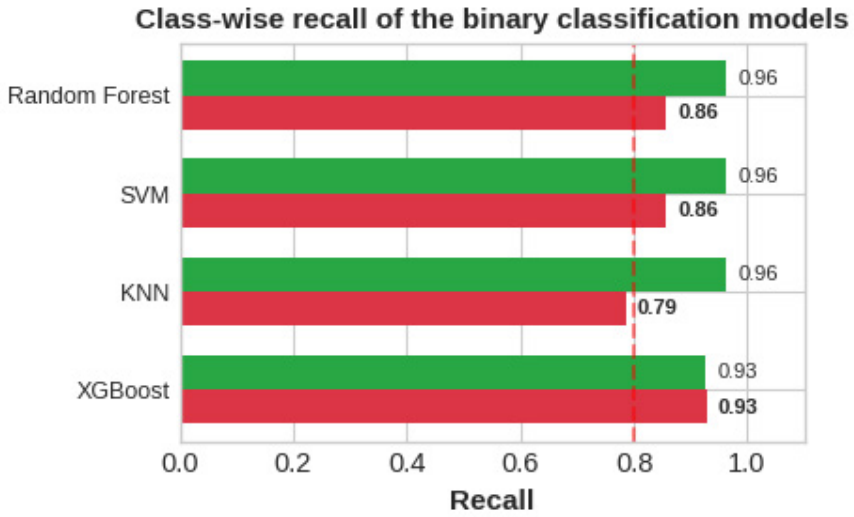

***Figura* 4**. Sensitivity analysis. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

The ROC-AUC comparison Figure 5, indicates that ensemble-based models (Random Forest and XGBoost) maintained the highest discrimination areas (AUC = 0.966 and 0.968, respectively), followed by SVM (AUC = 0.974) and KNN (AUC = 0.943), confirming excellent separability between classes (30, 34).

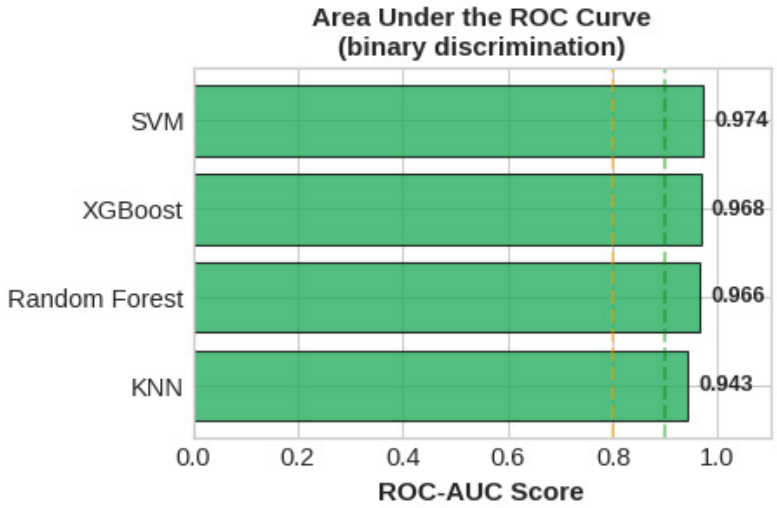

***Figura* 5**. The ROC-AUC comparison. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

The confusion matrices Figure 6, confirm that the best-performing classifiers produced very few type II errors (false negatives). For XGBoost, 26 out of 27 non-critical wells and 13 out of 14 critical wells were correctly classified this corresponds to an overall accuracy of over 94%. The performance demonstrates that the model generalizes well even in a moderately unbalanced sample without overfitting.

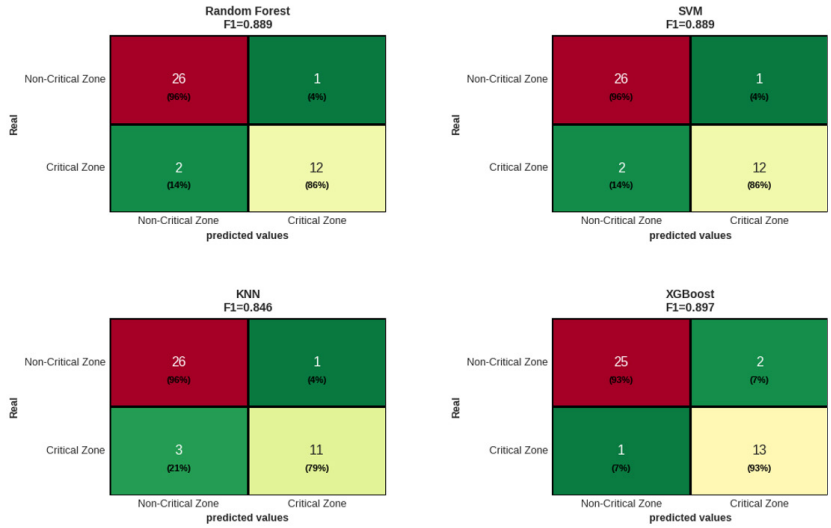

***Figura* 6**. The confusion matrices . Note. Elaboration own from the results obtained in Python (Google Colab , 2025).

Analysis detailed model XGBoost

The confusion matrices Figure 7, confirm that the best-performing classifiers produced very few type II errors (false negatives). For XGBoost, 26 out of 27 non-critical wells and 13 out of 14 critical wells were correctly classified; this corresponds to an overall accuracy of over 94%. The performance demonstrates that the model generalizes well even in a moderately unbalanced sample without overfitting (34).

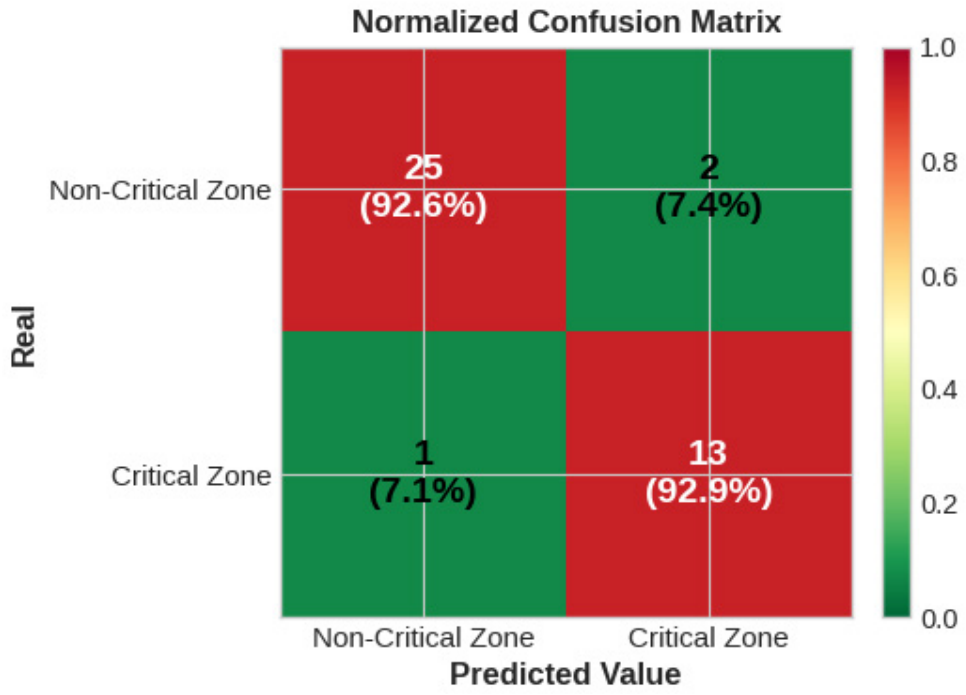

***Figura* 7**. Analysis detailed model XGBoost. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

The probability distribution in Figure 8, makes it easy to tell the difference between classes: most non-critical samples are below 0.3, and most critical samples are above 0.7. The best balance between recall and precision was found by setting the decision-bound threshold to 0.5. This value provided the best balance between recall and precision. This means that the model can be used as an early warning system to detect deterioration in groundwater quality. The sensitivity-accuracy curve Figure 9, confirms this result and shows an average precision (AP) of 0.931, indicating excellent diagnostic performance despite the lack of agreement between the datasets and the mean (33, 8). These results indicate the high sensitivity of the model in spotting key conditions (8) with great values: Non-Critical Zone: precision = 0.962, recall = 0.926. Critical Zone: precision = 0.867, recall = 0.929.

Volumen 10, Nro. 28, enero - abril 2026
ISSN: 2664-0902 / ISSN: 2664-0902, www.revistaalfa.org

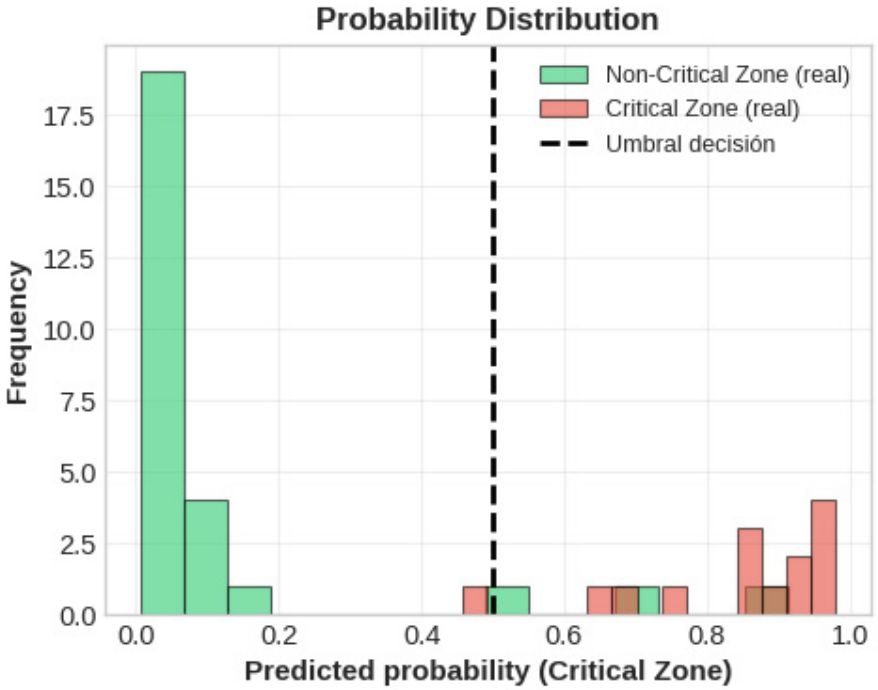

***Figura* 8**. The probability distribution. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

Comparative Evaluation

Evaluating Comparisons Figure 9, shows the performance heatmap, which gives an overview of the results of all classifiers. XGBoost, Random Forest, and SVM showed the most consistent performance across all metrics, with values above 0.85 for precision, recall, F1-score, and ROC-AUC. KNN, on the other hand, had a lower recall (0.786), which demonstrated their higher likelihood of overfitting on small datasets (33). The plot Figure 10, shows that the top three models are very similar as their accuracy, precision, and ROC-AUC are almost the same. XGBoost, on the other hand, had a slightly higher recall (0.93), which is useful for finding important areas with few false negatives (34). Overall, these results confirm that ensemble-based analysis algorithms are robust because they balance interpretability and predictive power better than distance (8, 30).

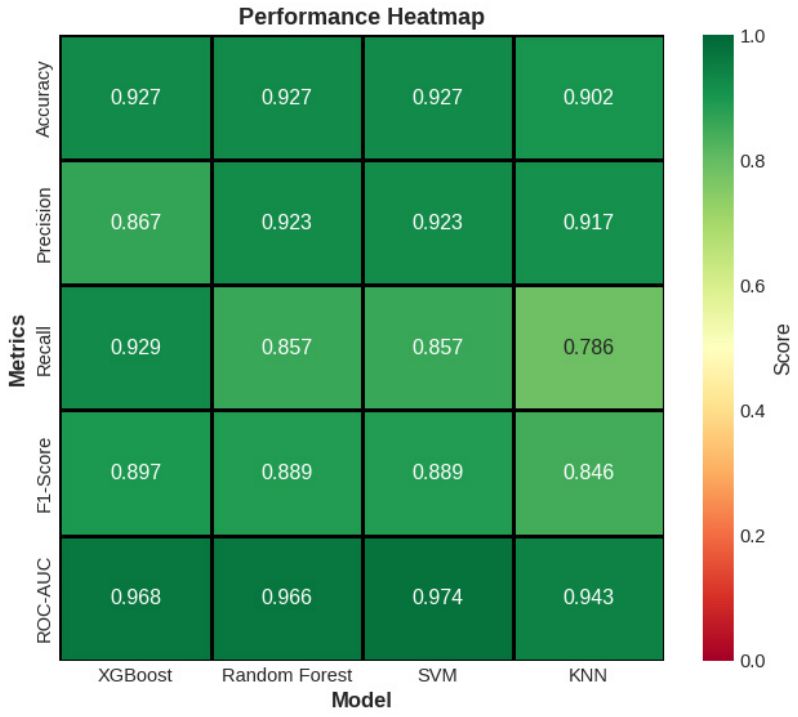

***Figura* 9**. Comparative Evaluation. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

***Figura* 10**. Radar comparision top models. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

Variable Reduction and Tracking Improvement

The incremental ROC analysis Figure 11, shows that the Random Forest model maintains AUC values above 0.94 even when only two features are used. This demonstrates the robustness of the model. The performance-cost ratio Figure 12, shows that the best configuration is the one with Na^+, Ca^{2+}, HCO_3^- and K^+, which reduces monitoring costs by 43% without losing accuracy. The performance curve shows that the predictive capacity stabilizes when four hydrochemical variables are added, with AUC = 0.960. This combination captures both the signature carbonate recharge (Ca–HCO_3) and marine salinization trend (Na–Cl), strengthening the predictions and chemical consistency of the model (9, 2).

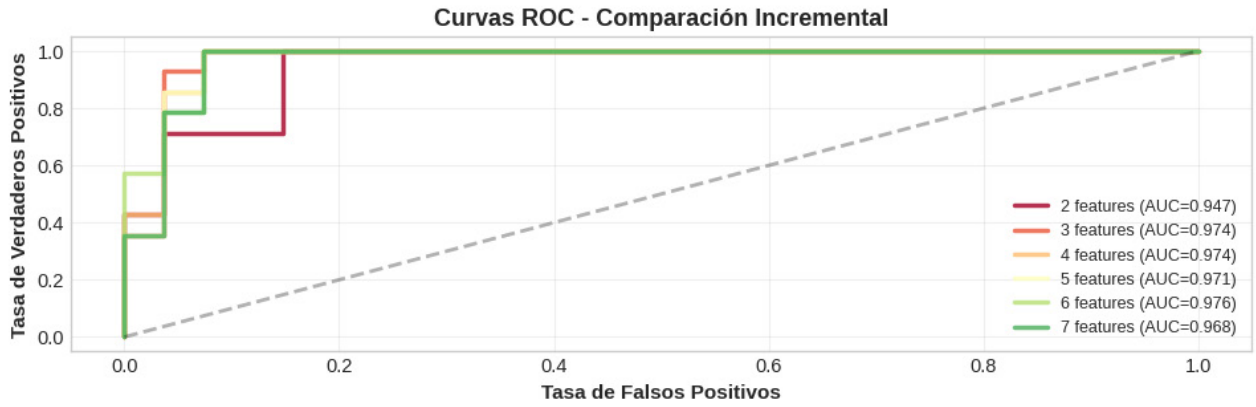

***Figura* 11**. The incremental ROC analysis. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

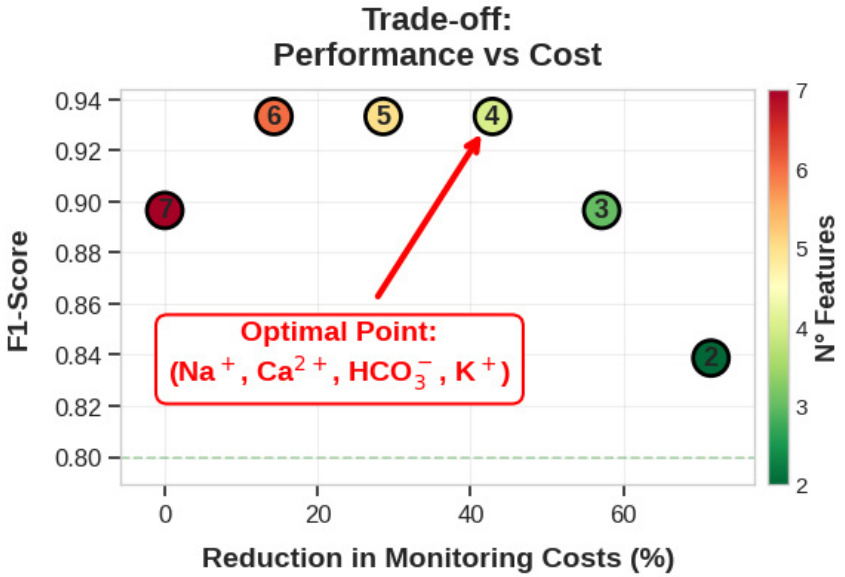

***Figura* 12**. The performance-cost ratio. Note. Elaboration own from the results obtained in Python (Google Colab, 2025).

Finally, the spatial map Figure 13, shows the results of the binary classification. It shows that nearshore wells are located in Critical Zones, which is consistent with what was observed in the field and what was learned from hydrogeochemical indicators of marine intrusion. The feature importance ranking supports these results by showing that Na^{+} is the most important predictor, followed by (Ca^{2+}, Mg^{2+}, K^{+}).

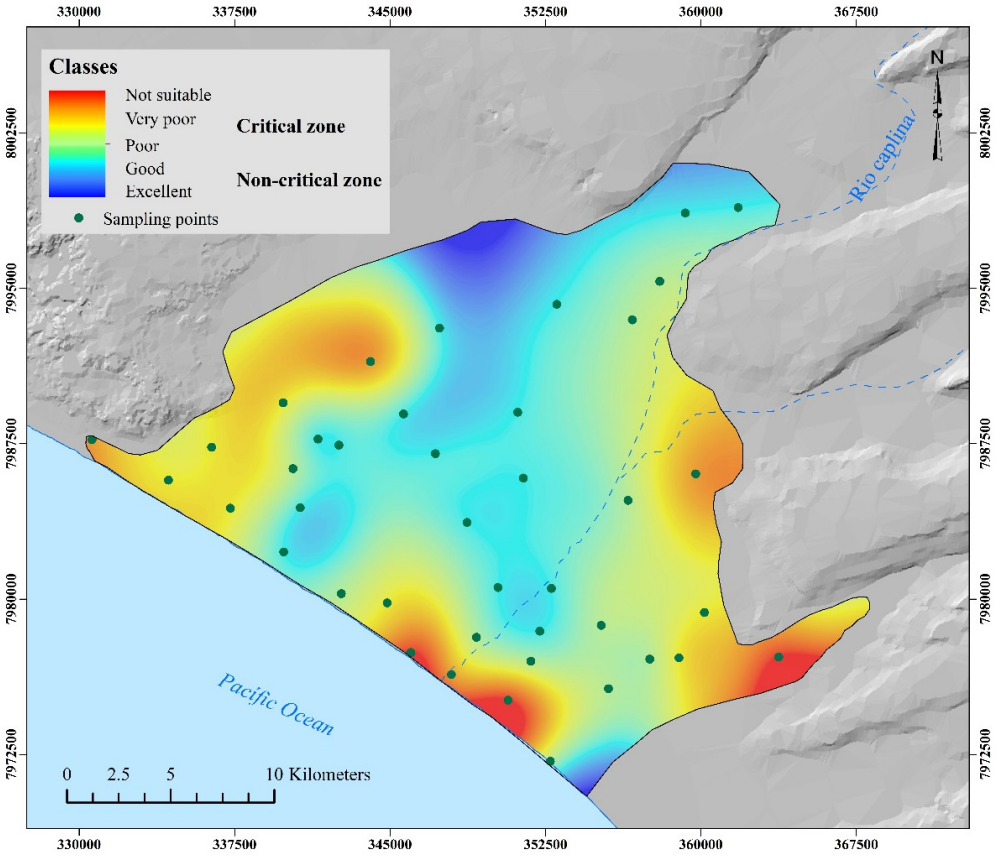

***Figura* 13**. Classification map binary. Nota. Spatial distribution map of groundwater quality index (WQI-UA) in the Caplina aquifer, southern Peru.The map shows Critical areas (red–orange tones) are concentrated near the coast, influenced by seawater mixing and over-extraction, whereas non-critical areas (blue–cyan tones) correspond to recharge zones of better quality.

CONCLUSION

Integrating machine-learning models with hydrochemical indicators proved effective for assessing and predicting groundwater quality in a dry coastal aquifer such as Caplina. Ensemble methods, particularly XGBoost and Random Forest, achieved the strongest class separation under class imbalance, with F1 scores > 0.89 and ROC-AUC > 0.96, in line with recent evidence on the advantages of boosting/bagging for non-linear groundwater problems (31, 30, 8). From the hydrochemical standpoint, Na^{+}, Ca^{2+}, Mg^{2+} and

K^+ emerged as the most informative predictors, reflecting progressive salinization and Na^+–Ca^{2+} ion-exchange processes that typify early stages of marine intrusion (2, 12).

Model parsimony analyses showed that reducing inputs from seven to four ions (Na^+, Ca^{2+}, HCO_3^-, K^+) preserved high performance (F1 = 0.897; AUC = 0.960) while lowering monitoring costs by 43%, supporting the design of lean observation networks for hydrochemical surveillance in resource-constrained settings (7). The Piper facies distribution (Ca–Mg–HCO_3: 68%; Na–Cl: 10%) is consistent with early salinization during the 2022 survey, indicating a transition from carbonate-recharge signatures toward coastal saline influence as abstraction pressures increase. Together, these results provide an operational basis to prioritize critical zones and to deploy cost-effective monitoring guided by a minimal, yet chemically coherent, set of ions.

ACKNOWLEDGEMENTS. The authors gratefully acknowledge the support of the research project "Determination of freshwater reserves in the Caplina aquifer and assessment of salinization risks for the sustainable management of groundwater, Tacna, Peru", approved under Rectoral Resolution No. 15241-2025-UNJBG. This study was conducted within the framework of said project, whose support was essential for the development of the research.

CONFLICT OF INTERESTS. The authors declare that they have no conflict of interest.

REFERENCES

1. Siebert S, Burke J, Faures J, Frenken K, Hoogeveen J, Döll P, Portmann F. Groundwater use for irrigation – a global inventory. Hydrology and Earth System Sciences. 2010; 14(10): 1863–1880. https://doi.org/10.5194/hess-14-1863-2010

2. Ayari J, Ouelhazi H, Charef A, Barhoumi A. Delineation of seawater intrusion and groundwater quality assessment in coastal aquifers: The Korba coastal aquifer (Northeastern Tunisia). Marine Pollution Bulletin. 2023; 188: 114643. https://doi.org/10.1016/j.marpolbul.2023.114643

3. Dao P, et al. The impacts of climate change on groundwater quality: A review. Science of the Total Environment. 2024; 912: 169241. https://doi.org/10.1016/j.scitotenv.2023.169241

4. Davamani V, John J, Poornachandhra C, Gopalakrishnan B, Arulmani S, Parameswari E, Naidu R. A critical review of climate change impacts on groundwater resources: Current status, future possibilities, and role of simulation models. Atmosphere. 2024; 15(1): 122. https://doi.org/10.3390/atmos15010122

5. Anyango W, Bhowmick D, Bhattacharya N. A critical review of irrigation water quality index and water quality management practices in micro-irrigation for efficient policy making. Desalination and Water Treatment. 2024; 318: 100304. https://doi.org/10.1016/j.dwt.2024.100304

6. Abadi H, Alemayehu T, Berhe A. Assessing the suitability of water for irrigation purposes using irrigation water quality indices in the Irob catchment, Tigray, Northern Ethiopia. Water Quality Research Journal. 2025; 60(1): 177–195. https://doi.org/10.2166/WQRJ.2024.055

7. Masoud M, El Osta M, Alqarawy A, Elsayed S, Gad M. Evaluation of groundwater quality for agriculture under different conditions using water quality indices, partial least squares regression models, and GIS approaches. Applied Water Science. 2022; 12(10): 1–22. https://doi.org/10.1007/s13201-022-01770-9

8. Ibrahim H, Yaseen M, Scholz M, Ali M, Ga M, Elsayed S, Khadr M, Hussein H, Ibrahim H, Eid H, Kovács A, Péter S, Khalifa M. Evaluation and prediction of groundwater quality for irrigation using integrated water quality indices, machine learning models and GIS approaches: A representative case study. Water. 2023; 15(4): 694. https://doi.org/10.3390/w15040694

9. Abu S, Ismael S, El-Sabri A, Abdo S, Farhat I. Integrated machine learning–based model and WQI for groundwater quality assessment: ML, geospatial, and hydro-index approaches. Environmental Science and Pollution Research. 2023; 30(18): 53862. https://doi.org/10.1007/s11356-023-25938-1

10. Basharat U, Zhang W, Han C, Khan S, Abbasi A, Mahroof S, Li S. Optimizing machine learning methods for groundwater quality prediction: Case study in District Bagh, Azad Kashmir, Pakistan. Ecotoxicology and Environmental Safety. 2025; 302: 118610. https://doi.org/10.1016/j.ecoenv.2025.118610

11. Rudrani A, Madhnure P, Yadav A, Kumari B, Roy A, Kumar V, Dauji S, Tirumalesh K. Machine learning approaches for predicting water quality towards climate-resilient groundwater management in southern India. Hydrology Research. 2025; 56(8): 754–773. https://doi.org/10.2166/nh.2025.042

12. Chucuya S, Vera A, Pino-Vargas E, Steenken A, Mahlknecht J, Montalván I. Hydrogeochemical characterization and identification of factors influencing groundwater quality in coastal aquifers, case: La Yarada, Tacna, Peru. International Journal of Environmental Research and Public Health. 2022; 19(5): 2815. https://doi.org/10.3390/ijerph19052815

13. Narvaez-Montoya C, Torres-Martínez J, Pino-Vargas E, Cabrera-Olivera F, Loge F, Mahlknecht J. Predicting adverse scenarios for a transboundary coastal aquifer system in the Atacama Desert (Peru/Chile). Science of the Total Environment. 2022; 806: 150386. https://doi.org/10.1016/j.scitotenv.2021.150386

14. Pino-Vargas E, Espinoza-Molina J, Chávarri-Velarde E, Quille-Mamani J, Ingol-Blanco E. Impacts of groundwater management policies in the Caplina aquifer, Atacama Desert. Water. 2023; 15(14): 2610. https://doi.org/10.3390/w15142610

15. Vera A, Pino-Vargas E, Verma P, Chucuya S, Chávarri E, Canales M, Torres-Martínez A, Mora A, Mahlknecht J. Hydrodynamics, Hydrochemistry, and Stable Isotope Geochemistry to Assess Temporal Behavior of Seawater Intrusion in the La Yarada Aquifer in the Vicinity of Atacama Desert, Tacna, Peru. Water 2021. 2021; 13(22): 3161. https://doi.org/10.3390/W13223161

16. González-Domínguez, J., Mora, A., Chucuya, S., Pino-Vargas, E., Torres-Martínez, J. A., Dueñas-Moreno, J., Ramos-Fernández, L., Kumar, M., & Mahlknecht, J. (2024). Hydraulic recharge and element dynamics during salinization in an overexploited coastal aquifer of the world's driest zone: Atacama Desert. Science of the Total Environment, 954, 176204. https://doi.org/10.1016/j.scitotenv.2024.176204

17. Parkhurst, D. L., & Appelo, C. A. J. (2013). Description of input and examples for PHREEQC version 3. U.S. Geological Survey Techniques and Methods, book 6, chap. A43.

18. El Tahlawi, M. R., Abo-El Kassem, M., Baghdadi, G. Y., & Saleem, H. A. (2016). Estimating and plotting of groundwater quality using WQIUA and GIS in Assiut Governorate, Egypt. World Journal of Engineering and Technology, 4(1), 59–70. https://doi.org/10.4236/wjet.2016.41007

19. Wilcox, L. V. (1955). Classification and use of irrigation waters. U.S. Department of Agriculture, Circular No. 969.

20. Ayers, R. S., & Westcot, D. W. (1985). Water quality for agriculture. FAO Irrigation and Drainage Paper 29. Food and Agriculture Organization, Rome.

21. Richards, L. A. (1954). Diagnosis and improvement of saline and alkali soils (Vol. 60). U.S. Department of Agriculture, Agriculture Handbook No. 60.

22. Doneen, L. D. (1964). Notes on water quality in agriculture. Department of Water Science and Engineering, University of California, Davis.

23. Raghunath, H. M. (1987). Ground Water (2nd ed.). Wiley Eastern Ltd., New Delhi.

24. Kelly, W. P. (1951). Alkali soils: Their formation, properties and reclamation. Reinhold Publishing Corporation, New York.

25. Rice, E. W., Baird, R. B., & Eaton, A. D. (Eds.). (2017). Standard methods for the examination of water and wastewater (23rd ed.). APHA; AWWA; WEF.

26. Piper, A. M. (1944). A graphic procedure in the geochemical interpretation of water analyses. Eos, Transactions American Geophysical Union, 25(6), 914–928. https://doi.org/10.1029/TR025i006p00914

27. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

28. Chen, W., Xu, D., Pan, B., Zhao, Y., & Song, Y. (2024). Machine learning-based water quality classification assessment. Water, 16(20), 2951. https://doi.org/10.3390/w16202951

29. Güler, C., Thyne, G. D., McCray, J. E., & Turner, A. K. (2002). Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeology Journal, 10(4), 455–474. https://doi.org/10.1007/s10040-002-0196-6

30. Hussein, E. E., Derdour, A., Zerouali, B., Almaliki, A., Wong, Y. J., Ballesta-de los Santos, M., Minh Ngoc, P., Hashim, M. A., & Elbeltagi, A. (2024). Groundwater quality assessment and irrigation water quality index prediction using machine learning algorithms. Water, 16(2), 264. https://doi.org/10.3390/w16020264

31. Gad, M., Saleh, A. H., Hussein, H., Elsayed, S., & Farouk, M. (2023). Water quality evaluation and prediction using irrigation indices, artificial neural networks, and partial least square regression models for the Nile River, Egypt. Water, 15(12), 2244. https://doi.org/10.3390/w15122244

32. Karthick, K., Krishnan, S., & Manikandan, R. (2024). Water quality prediction: A data-driven approach exploiting advanced machine learning algorithms with data augmentation. Journal of Water and Climate Change, 15(2), 431–452. https://doi.org/10.2166/wcc.2023.403

33. Nasaruddin, N., Masseran, N., Idris, W. M. R., & Ul-Saufie, A. Z. (2025). A SMOTE PCA HDBSCAN approach for enhancing water quality classification in imbalanced datasets. Scientific Reports, 15(1), 1–12. https://doi.org/10.1038/s41598-025-97248-0

34. Saito, T., & Rehmsmeier, M. (2015). The precision–recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE, 10(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

35. Irwan, D., Ibrahim, S. L., Latif, S. D., Winston, C. A., Ahmed, A. N., Sherif, M., El-Shafie, A. H., & El-Shafie, A. (2025). River water quality monitoring using machine learning with multiple possible in-situ scenarios. Environmental and Sustainability Indicators, 26, 100620. https://doi.org/10.1016/j.indic.2025.100620

36. Chadha, D. K. (1999). A proposed new diagram for geochemical classification of natural waters and interpretation of chemical data. Hydrogeology Journal, 7(5), 431–439. https://doi.org/10.1007/s100400050216

37. Langelier, W. F., & Ludwig, H. F. (1942). Graphical methods for indicating the mineral character of natural waters. Journal AWWA, 34(3), 335–352. https://doi.org/10.1002/j.1551-8833.1942.tb19682.x

38. Teng, W. C., Fong, K. L., Shenkar, D., Wilson, J. A., & Foo, D. C. Y. (2016). Piper diagram—A novel visualization tool for process design. Chemical Engineering Research and Design, 112, 132–145. https://doi.org/10.1016/j.cherd.2016.06.002

39. Vasilache, N., Vasile, G. G., Diacu, E., Modrogan, C., Paun, I. C., & Pirvu, F. (2020). Groundwater quality assessment for drinking and irrigation purpose using GIS, Piper diagram, and water quality index. Romanian Journal of Ecology & Environmental Chemistry, 2(2), 109–117. https://doi.org/10.21698/RJEEC.2020.214

40. Moreno Merino, L., Aguilera, H., González-Jiménez, M., & Díaz-Losada, E. (2021). D-Piper, a modified Piper diagram to represent big sets of hydrochemical analyses. Environmental Modelling & Software, 138, 104979. https://doi.org/10.1016/j.envsoft.2021.104979