agronomy

MDPI

Article

Artificial Neural Networks to Predict Electrical Conductivity of Groundwater for Irrigation Management: Case of Campo de Cartagena (Murcia, Spain)

Luis F. Mateo [1], M. Isabel Más-López pez [1], Eva M. García-del-Toro [2,*], Sara García-Salgado [2] and M. Ángeles Quijano [2]

[1] Departamento de Matemática e Informática Aplicadas a las Ingenierías Civil y Naval, ETSI Caminos, Canales y Puertos-Edificio Retiro, Universidad Politécnica de Madrid, Alfonso XII, 3, 28014 Madrid, Spain; luis.f.mateo@upm.es (L.F.M.); mariaisabel.mas@upm.es (M.I.M.-L.)

[2] Departamento de Ingeniería Civil: Hidráulica, Energía y Medio Ambiente, ETSI Caminos, Canales y Puertos-Edificio Retiro, Universidad Politécnica de Madrid, Alfonso XII, 3, 28014 Madrid, Spain; sara.garcia@upm.es (S.G.-S.); marian.quijano@upm.es (M.Á.Q.)

* Correspondence: evamaria.garcia@upm.es

Abstract: Groundwater is a crucial water resource, particularly in regions with intensive agriculture and a semi-arid climate, such as Campo de Cartagena (Murcia, Spain). Groundwater salinity in the area can be attributed to hydrogeological characteristics, irrigation return water, or even marine intrusion and communication between aquifers. The management of these waters is essential to maintain sustainable agriculture in the area. Therefore, two groundwater salinity prediction models were developed, a backpropagation artificial neural network (ANN) model and a multiple linear regression (MLR) model, based on EC (electrical conductivity) data obtained from official information sources. The data used were the bicarbonate, calcium, chloride, magnesium, nitrate, potassium, sodium, and sulphate concentrations, as well as EC, pH, and temperature, of 495 water samples from 38 sampling stations between 2000 and 2023. Variables with the least influence on the model were discarded in a previous statistical analysis. Based on seven evaluation metrics (RMSE, MAE, R^2, MPE, MBE, SSE, and AARD), the ANN model showed a slightly better accuracy in predicting EC compared to the MLR model. As a result, the ANN model, together with crop tolerance to EC, may be an effective tool for groundwater irrigation management in these areas.

Keywords: groundwater; salinity; electrical conductivity; sustainable agriculture; artificial neural networks; irrigation management; Campo de Cartagena (Murcia, Spain)

Citation: Mateo, L.F.; Más-López, M.I.; García-del-Toro, E.M.; García-Salgado, S.; Quijano, M.Á. Artificial Neural Networks to Predict Electrical Conductivity of Groundwater for Irrigation Management: Case of Campo de Cartagena (Murcia, Spain). *Agronomy* **2024**, *14*, 524. https://doi.org/10.3390/agronomy14030524

Academic Editor: Paul A. Ty Ferré

Received: 29 January 2024
Revised: 28 February 2024
Accepted: 1 March 2024
Published: 3 March 2024

1. Introduction

Groundwater is the most important freshwater resource when surface water supply is insufficient, and accounts for one-third of total water withdrawals worldwide [1]. This resource is relevant to various aspects of human development and ecosystem sustainability [2]. However, groundwater is sometimes excessively extracted, causing the depletion of the resource, since the recharge of aquifers is very slow and sometimes non-existent. Pollution from excess nutrients and pesticides, along with salinization problems, poses major challenges for sustainable agricultural management [3]. Among these problems, salinization is the most widespread, which causes a reduction in the quality of water. The salinity of groundwater is controlled by many factors such as the composition of aquifer material, climate, the infiltration of salt-laden irrigation water from agricultural treatments, and in coastal areas by marine intrusion.

Electric conductivity (EC) is the transmission capacity of electric currents in water and it depends on the ions present in the water and therefrom the water salinity, expressed through the concentration of total dissolved solids (TDS) through the expression

TDS (mg L^{-1}) = K × EC (µS cm^{-1}), where K takes values between 0.55 and 0.85, depending on the salinity level. In general, a value close to the average is usually taken, specifically 0.64. However, in hypersaline waters, the relationship loses linearity and reflects an exponential trend [4]. EC can be measured using an EC meter, and must be measured in situ, and it depends on the temperature. EC measurement in situ is labor-intensive, time-consuming, costly, and can be difficult, especially in the case of groundwater, where access depends on the characteristics of the terrain. In addition, other factors interfere with the flow of electrical currents and affect the precision and accuracy of the measurement, such as the presence of impurities in the water, which is not easy to control in groundwater.

EC prediction is a useful tool in groundwater management for determining groundwater quality and usability. In the case of use for irrigation, the EC will determine the type of crops and the amount of fertilizer to be applied [5]. These predictions are especially important in areas with intensive agriculture and semi-arid climates, which are affected by salinity problems [6]. The Campo de Cartagena region (SE of Spain) presents an important and developed area of intensive agriculture that supplies Spain and a large part of Europe and has an associated important agro-food industry that is the basis of the region's economy [7]. Most of the region's crop fields, about 40,000 hectares, are irrigated with water from the Tajo-Segura Water Transfer (TSWT) and the multilayer coastal aquifer of Campo de Cartagena. The average demand for water for irrigation is about 220 hm^3 yr^{-1}. The maximum annual allocation from the TSWT is around 120 hm^3 yr^{-1}. The rest comes from other sources, mainly groundwater and, to a lesser extent, from the Segura basin, treated water, and desalinated water [8]. The amount of groundwater withdrawn is a function of the TSWT's water availability [9]. In general, according to official sources, the overall inflow–outflow balance for the aquifers in the area is slightly negative, around -8 hm^3 yr^{-1} [8]. The high saline concentration of groundwater is due to hydrogeology and from the irrigation return waters containing salts from fertilizers. It is also due to irregular discharges of brine from small desalination plants, scattered throughout the region, either directly into the ground or injected into abandoned wells. In addition, there are some marine intrusion problems due to the communication of aquifers in some areas with the sea and between each other [10,11].

Both the climatic and hydrogeological characteristics and the fact that intensive agriculture grows in the area make the management of irrigation groundwater of great interest. The assessment of the water quality involves large-scale sampling, analysis, and data processing, and therefore the need for equipment, chemicals, and human resources, which is time-consuming and expensive. Therefore, the development of prediction models for water quality parameters based on artificial intelligence (AI) is of great interest [12].

Some researchers have shown that AI techniques can be an effective and reliable tool for predicting groundwater salinity. Barzegar and Asghari Moghaddam [13] proposed a committee neural network (CNN) developed by the combination of the salinity results from three individual ANN models (multi-layer perceptron (MLP), radial basis function (RBFNN), and generalized regression (GRNN)) for predicting the groundwater EC of the Tabriz plain confined aquifer. Ekemn Keskin et al. [14] forecasted EC of groundwater, from some areas in Turkey, using an artificial neural network (ANN) trained with backpropagation and a statistical model (MLR (multiple linear regression)) and reported higher performance for the ANN model. Kheradpisheh et al. [15] developed a feed-forward backpropagation neural network (FNN-BP) model to estimate EC together with Cl^- and SO_4^{2-} in groundwater samples from Bahabad plain in Iran. Among the input parameters were the concentration of major ions in water (mainly Ca^{2+}, Mg^{2+}, Na^+, SO_4^{2-}, and Cl^-), as well as heavy metals in mining areas (Fe, Mn, and Al), and even other parameters such as pH, evaporation, water elevation, and discharge. The salinity of groundwater has also been estimated by ANN models through TDS, sodium adsorption ratio (SAR), residual sodium carbonate (RSC), permeability index (PI), Kelly ratio (KR), and/or Na percent (%Na) or groundwater level [12,16,17]. Al-Waeli et at. [18] applied a backpropagation neural network model to predict groundwater salinity in the plateau of Najaf–Kerbala (Iraq). The authors

reported accurate results in predicting TDS, SAR, and %Na using the concentrations of major cations and pH as input parameters. Singh et al. [19] evaluated ANN and MLR models to predict groundwater quality for irrigation purposes in an intensively cultivated area in Punjab, India. The authors reported differences in the performance of both models depending on the predicted parameter and the season of the year. During the pre-monsoon season, MLR had the best performance in predicting RSC and SAR, while the ANN model was best in predicting Na%, PI, and KR. However, during the post-monsoon season, the ANN model was better than MLR in predicting all parameters except RSC. Kouadri et al. [20] reached similar results for prediction irrigation groundwater quality in the Akot basin area (India).

The quality of groundwater for irrigation depends largely on its salinity since it determines the type and yield of crops. A very useful parameter to determine the degree of salinity of water is the EC, but its experimental measurement can be complex, especially in groundwater, due to environmental conditions and terrain characteristics, and can be affected by possible random errors. This study aims to predict this parameter in groundwater in Campo de Cartagena (Murcia) by using theoretical models. As far as we know, there is no work on this topic in this area. Therefore, the main objective of this work is to provide tools for irrigation water management in semi-arid regions with low rainfall and intensive agriculture, such as the Campo de Cartagena region (Murcia, Spain). To this end, two theoretical models have been developed to predict EC as an indicator of groundwater salinity, comparing their accuracy. These consist of a traditional MLR-based model and an AI model using ANN. The studies also analyze the tolerance of the predominant crops in the area to water salinity and the possible need for desalination.

2. Materials and Methods

2.1. Studied Zone and Origin of Data

The Campo de Cartagena is located in the SE of the Iberian Peninsula in the region of Murcia (Spain). A description of the location and climate characteristics of the area can be found in previous work [21].

The hydrogeological system has been widely characterized [10,11,22–24]. Briefly, several layers constitute it with aquifers made up of detrital sediments, sandstone, and limestone, separated by aquitards generally composed of marl and clay. Two aquifers stand out for their extension and agricultural use: the intermediate Pliocene aquifer (with an area of 817 km^2), and the deeper Andalusian (upper Miocene) aquifer (570 km^2 thickest in the north). The Quaternary unconfined aquifer is the most superficial and wide in extension (1135 km^2) and communicates with the Mar Menor. There is also a fourth deeper confined aquifer from the Tortonian age, which is the smallest in extension and least used, present in the NO area of the region. The substrate of the Campo de Cartagena mass is the Betic substrate, formed by materials of dolomites, marbles, schists, and phyllites that include hypersaline formations of a confined thermal nature and fossils. The recharge of the groundwater mass comes mainly from the infiltration of rainwater and irrigation returns, which mainly affect the Quaternary aquifer. In addition, there is a small lateral supply from the Sierra de Cartagena. The discharge is carried out by pumping from the confined aquifers, as well as by lateral outlets towards the Mar Menor and the Mediterranean Sea through the unconfined aquifer. On the other hand, there exists an internal interconnection between different aquifers, both for natural conditions and for abandoned and poorly constructed boreholes. The vertical hydraulic communication through boreholes has been reported to be about 50 $hm^3\ yr^{-1}$ [11].

The physicochemical parameters of groundwater selected for this study were EC, pH, and water temperature, all measured in situ, and the concentration of major ions, bicarbonate (HCO_3^-), calcium (Ca^{2+}), chloride (Cl^-), magnesium (Mg^{2+}), nitrate (NO_3^-), potassium (K^+), sodium (Na^+), and sulphate (SO_4^{2-}), corresponding to 495 groundwater samples taken from 38 sampling stations between May 2000 and August 2023. The data come from the Groundwater Quality Monitoring Network from the Water Authority and

are available in open access on the web [25]. Table S1 indicates the geographical coordinates of the 38 sampling stations studied in the Campo de Cartagena (Confederación Hidrográfica del Segura (CHS) designations) and the sampling depth, differentiating between those located in the south, center, and north. The website provides information on the analytical results of groundwater for a water mass (Campo de Cartagena—070.052 in this study) and the desired date range. It also has a viewer that shows the location of the sampling stations in the region.

Table 1 shows the minimum, maximum, standard deviation, and mean values for each of the parameters considered in this study.

Table 1. Physicochemical parameters of groundwater in the studied region, between May 2020 and August 2023 (N = 495).

Parameter	Unit	Min	Max	Mean	Standard Deviation
Bicarbonate (HCO_3^-)	mg L^{-1}	60.8	1504	4×10^2	2×10^2
Calcium (Ca^{2+})	mg L^{-1}	13.7	686	3×10^2	2×10^2
Chloride (Cl^-)	mg L^{-1}	135	3797	12×10^2	6×10^2
Electric conductivity (EC)	mS cm^{-1}	0.437	12.460	6	2
Magnesium (Mg^{2+})	mg L^{-1}	23.0	564	2×10^2	1×10^2
Nitrate (NO_3^-)	mg L^{-1}	n.d.	703	2×10^2	1×10^2
pH	pH unit	6.30	9.68	7.5	0.4
Potassium (K^+)	mg L^{-1}	0.8	309	2×10^1	4×10^1
Sodium (Na^+)	mg L^{-1}	146	2637	8×10^2	3×10^2
Sulphate (SO_4^{2-})	mg L^{-1}	n.d.	4482	11×10^2	7×10^2
Temperature (T)	°C	7.05	36.9	22	4

n.d.: not detected.

2.2. Development of Mathematical Models

The first step in defining the MLR model was the selection of the most influential independent variables on EC (dependent variable) from those studied (HCO_3^-, Ca^{2+}, Cl^-, Mg^{2+}, NO_3^-, K^+, Na^+, SO_4^{2-}, pH, and temperature). To choose them, the Pearson linear correlation coefficient (r) between EC and the rest of the parameters, the statistical significance of each of the variables, and the adjusted coefficient of determination (R^2) were considered. Statistical calculations were performed with SPSS software Version 28.0 for Windows (IBM Corp. Armonk, NY, USA). We used the totality of the data in the period studied (495 groundwater samples).

The assumptions of linearity, independence, homoscedasticity, and normality of the residuals were checked to probe if the MLR model hypotheses fit the data collected.

The architecture of the ANN-based EC prediction model was developed by studying different configurations in terms of the number of neurons and hidden layers. Specifically, processes were tested with 2, 3, and 4 hidden layers and 100, 50, and 30 neurons per layer, respectively, and 100 cycles for all. The number of neurons in the input layer was set from the previously selected variables, while the output layer corresponded to a single neuron (predicted CE).

The ANN was created in two phases: training and validation of the model. The network weights were initialized as random values. A standardization process was also applied to avoid scale variations between the variables used in the network to transform the values into a natural system that varied from 0 to 1. The data were divided into two subsets. The first subset was used to generate a model for the training or learning phase of the network, and the second subset was used for the validation phase. The training/validation ratios tested ranged from 90–10% up to 50–50%, decreasing and increasing, respectively by 5% of data in each test.

The network was designed as a regression network with numerical results, so no activation functions were used. This study used the backpropagation learning algorithm, which consists of learning from a predetermined set of inputs and outputs through a

propagation cycle. The design of the ANN architectures for EC prediction was carried out with the KNIME application version 4.7.7. (KNIME Analytics Platform, open source). This software application is open-source and highly flexible. It is powerful and can be connected to other tools. This application showed a good results in previous studies [21,26].

To summarize, a flowchart is included. Figure 1 corresponds to the general process carried out in this study.

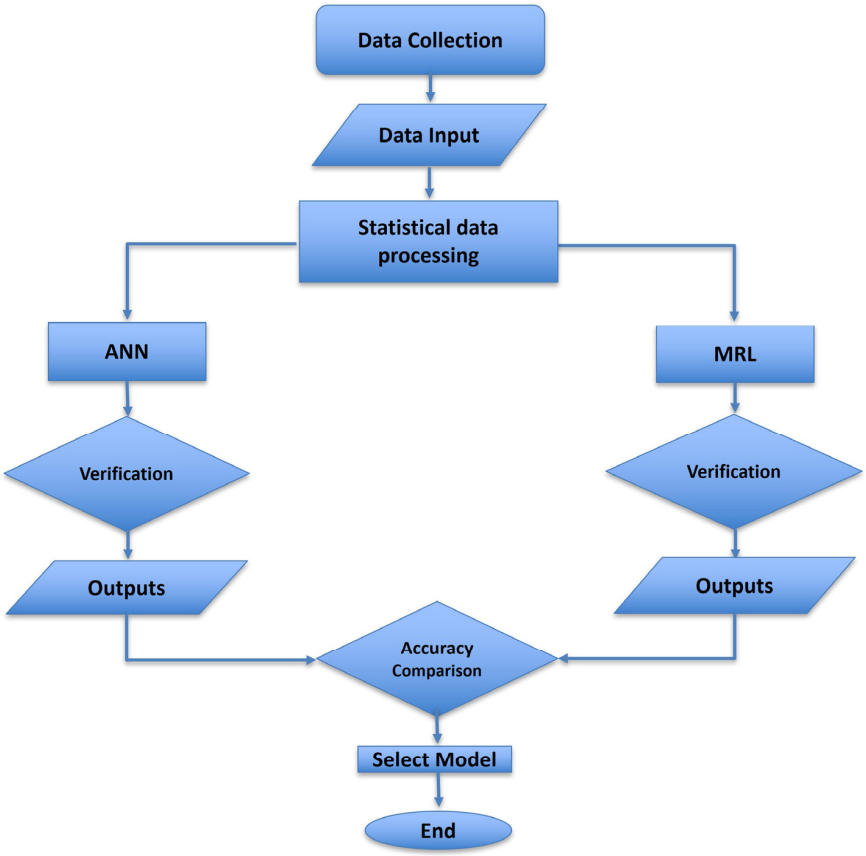


Figure 1. General flowchart of the modeling process.

3. Results and Discussion

3.1. Variables Selection for Development of Theoretical Models

Table 2 shows the linear Pearson correlation between the ten physicochemical groundwater variables considered (major ion concentrations, pH, and T) and the EC, for N = 495. From the result obtained, five variables (calcium, chloride, magnesium, sodium, and sulfate concentrations) presented a significant correlation with EC (p-value < 0.01). Therefore, these parameters were selected for their greatest influence in both mathematical models. The remaining variables (bicarbonate, nitrate, and potassium concentrations, as well as pH and T) were discarded from the model due to the low r values obtained for the groundwater under study.

The concentration of major ions was expected to have a strong influence on the EC of water as it contributes to the TDS content. However, the same was not observed for temperature, which is also known to affect EC. This may be due to several factors: hydrogeological differences between aquifers, possible marine intrusion in some areas, and the consequences of intensive agriculture in the area. Thus, the composition of groundwater can be affected through irrigation returns, irregular discharges of brines, and even the type of crops [10,11,21]. This heterogeneity is reflected in the wide range of groundwater EC values shown in Table 1. Our results agree with those reported by other authors who did not consider temperature as an influential parameter in the development of their groundwater EC prediction models [13,27].

Table 2. Pearson correlation coefficients (r) and *p*-values obtained between EC and the different variables (N = 495).

Variable	Electrical Conductivity (EC)	
	r	*p*-Value
Bicarbonate (HCO_3^-)	0.057	0.208
Calcium (Ca^{2+})	0.680	0.000
Chloride (Cl^-)	0.812	0.000
Magnesium (Mg^{2+})	0.842	0.000
Nitrate (NO_3^-)	0.026	0.567
pH	−0.233	0.018
Potassium (K^+)	0.336	0.040
Sodium (Na^+)	0.828	0.000
Sulphate (SO_4^{2-})	0.680	0.000
Temperature (T)	−0.077	0.098

3.2. MLR Model

For the development of the MLR model for EC prediction, Ca^{2+}, Cl^-, Mg^{2+}, Na^+, and SO_4^{2-} concentrations were used as independent variables. The best-fitting model obtained is shown in Equation (1).

$$EC = 965.203 + 2.374Mg^{2+} + 2.276Na^+ + 0.240SO_4^{2-} + 1.065Cl^- + 1.919Ca^{2+} \quad (1)$$

where EC is the electric conductivity of groundwater in mS cm^{-1} and Mg^{2+}, Na^+, SO_4^{2-}, Cl^-, and Ca^{2+} are the corresponding ion concentrations in mg L^{-1}.

Table 3 shows the characteristics of the MLR model. R represents the multiple linear correlation coefficient between the dependent variables and the independent variable. A value obtained very close to 1 indicates that the selected dependent variables maintain a strong relationship with the independent variable. R^2 is a measure of the overall accuracy of the model and indicates that 83.4% of the EC is explained by the five independent variables selected.

Table 3. Characteristics of the MLR model developed.

R	R^2	Ajusted R^2	Standar Error	Durbin-Watson
0.913	0.834	0.883	693.223	1.913

To verify the validity of the proposed model, the assumptions of linearity, independence, homoscedasticity, and normality of the residuals were checked.

Table 4 shows the results of the ANOVA analysis of the proposed MLR model. As can be seen, the linearity assumption is met, because the *p*-value is significant as it is lower than the level of significance ($\alpha < 0.01$).

Table 4. ANOVA analysis results of the MLR proposed model.

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1,768,292,671.458	5	294,715,445.243	613.278	0.000
Residuals	231,628,896.472	482	480,557.876		
Total	1,999,921,567.930	488			

The value obtained for the independence test of the Durbin–Watson residuals is equal to 1.913 (Table 3), which is within the range of 1.5–2.5. Therefore, it meets the assumption of independece.

The correlation between the absolute value of the residuals and their estimated values was calculated to test the homoscedasticity of the model. The results (Table 5) show that

the *p*-value is 0.185, so it is above the significance level (α = 0.01). The null hypothesis H_0 was accepted, and it can be affirmed that there is no correlation between the variables, demonstrating the model's homoscedasticity.

Table 5. Correlation between absolute value of residuals and their estimated values (N = 495).

		ABS Residuals	Predicted Value
ABS Residuals	r	1	0.060
	p-value	-	0.185
Predicted Value	r	0.060	1
	p-value	0.185	-

From the results obtained in the Kolmogorov–Smirnov test (Table 6), the *p*-value was 0.016, and therefore higher than 0.01. It confirms that the variable is distributed according to a normal probability function.

Table 6. Kolmogorov–Smirnov test results.

		Unstandardized Residual
N		495
Normal parameters	Mean	0.0038336
	Deviation	0.98888571
Most extreme differences	Absolute	0.046
	Positive	0.046
	Negative	−0.045
Kolmogorov–Smirnov Z		0.046
p-value		0.016

In conclusion, the MLR model proposed can be a valid prediction model for EC in the groundwater studied, using the concentrations of calcium, chloride, magnesium, sodium, and sulphate, since it does meet the hypotheses assumed by the model.

3.3. ANN Model

The proposed ANN architecture for the prediction model of the EC present in the groundwater of the Campo de Cartagena region is shown in Figure 2.

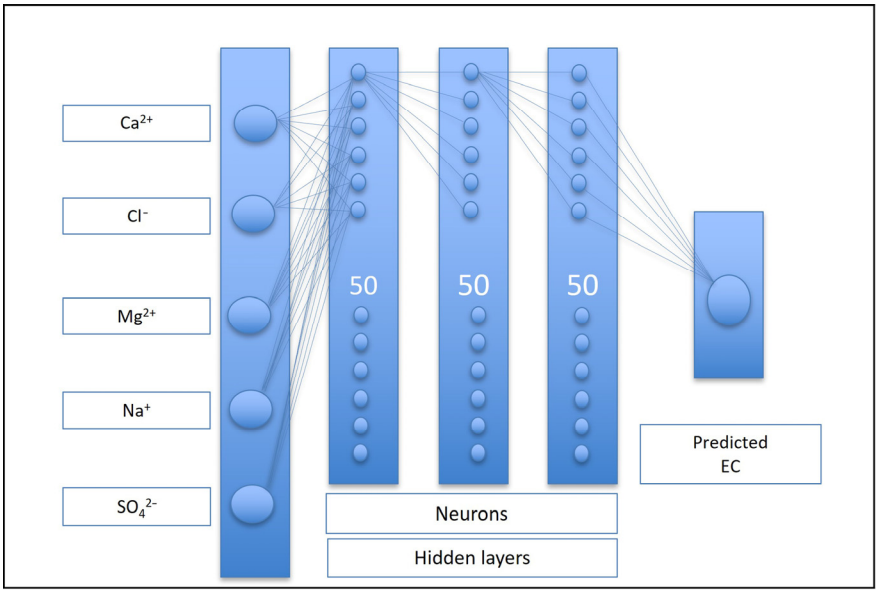


Figure 2. Schematic representation of the ANN architecture proposed for the prediction model of the EC in the groundwater of Campo de Cartagena region (Murcia, Spain).

The developed architecture consisted of a first input layer with five neurons, corresponding to the selected variables (Ca^{2+}, Cl^{-}, Mg^{2+}, Na^{+}, and $SO_4{}^{2-}$ groundwater concentrations) and an output layer with the EC as the dependent variable. Among those studied, the configuration with 3 hidden layers and 50 neurons was selected as it produced the best response time. The corresponding errors were a root mean square error (RMSE) of 0.7474, a mean absolute error (MAE) of 0.5684, and an R^2 value of 0.8774. The hidden layers/neurons configurations of 2/100 and 4/30 resulted in higher errors and worse coefficients of determination. The RMSE values were 0.7516 and 0.7625, the MAE values were 0.5724 and 0.5822, and the R^2 values were 0.8322 and 0.8042 for the configurations 2/100 and 4/30, respectively. The optimal data distribution for the training phase was 70% and for the validation tests was 30%. This relationship has also been shown to be the most appropriate in previous studies on the development of neural networks [26] using the backpropagation learning algorithm and 100 cycles.

Figure 3 shows the process used to design the KNIME network. The first phase is data entry, followed by the selection of the independent variables generated by the five input neurons (Ca^{2+}, Cl^{-}, Mg^{2+}, Na^{+}, and $SO_4{}^{2-}$ concentrations). Next, data cleaning is carried out, where possible anomalous data are eliminated, followed by normalization of the values (to natural numbers between 0 and 1) to avoid scale differences. Data partitioning for training and validation is then carried out. The system starts the backpropagation learning algorithm with 100 cicles. Finally, the model proceeds to the validation and output of results.

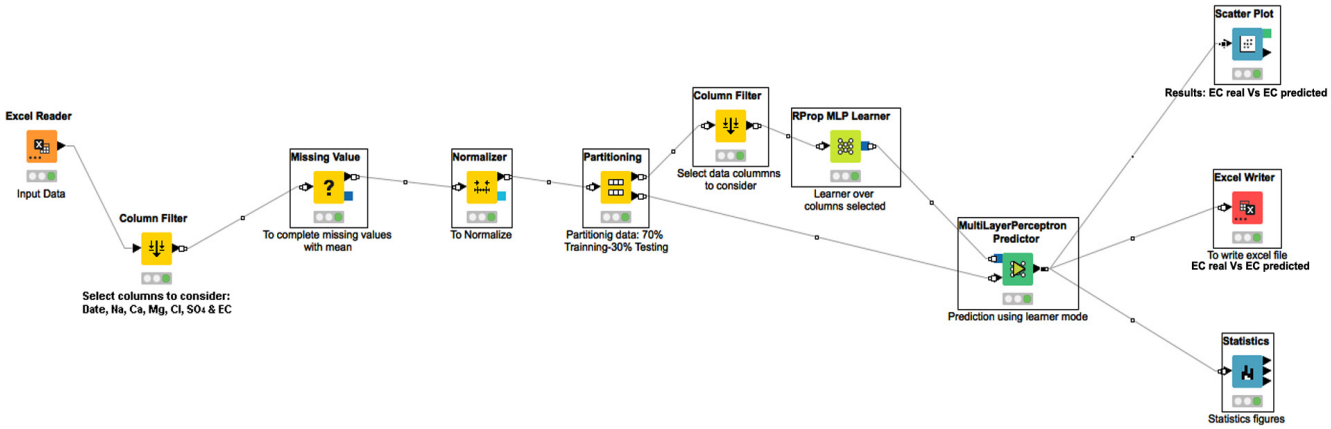


Figure 3. KNIME tool output diagram for the ANN model.

3.4. Comparison between Two Developed Models

The theoretical models were compared using various evaluation metrics: the root mean square error (RMSE), the mean absolute error (MAE), the coefficient of determination (R^2), the mean percentage error (MPE), the mean bias error (MBE), the total squared error (SSE), and the average absolute relative deviation (AARD). Equations from (2)–(8) show the mathematical expressions from these evaluation metrics:

$$\mathrm{RMSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left(O_i - E_i\right)^2} \tag{2}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|O_i - E_i\right| \tag{3}$$

$$R^2 = \frac{\left[\sum_{i-1}^{n}\left(O_i - \overline{O}\right) \times \left(E_i - \overline{E}\right)\right]^2}{\left[\left[\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2\right]^{\frac{1}{2}} \times \left[\sum_{i=1}^{n}\left(E_i - \overline{E}\right)^2\right]^{\frac{1}{2}}\right]^2} \tag{4}$$

$$\mathrm{MPE} = \frac{1}{n}\sum_{i=1}^{n} 100\left(\frac{O_i - E_i}{O_i}\right) \tag{5}$$

$$\mathrm{MBE} = \frac{1}{n}\sum_{i=1}^{n}(E_i - O_i) \quad (6)$$

$$\mathrm{SSE} = \frac{1}{n}\sum_{i=1}^{n}(E_i - \overline{E})^2 \quad (7)$$

$$\mathrm{AARD} = \frac{1}{100}\sum_{i=1}^{n}\left|\frac{O_i - E_i}{O_i}\right| \quad (8)$$

where O_i is the ith observed data, E_i is the ith estimated data, and n is the total number of observations in the validation phase.

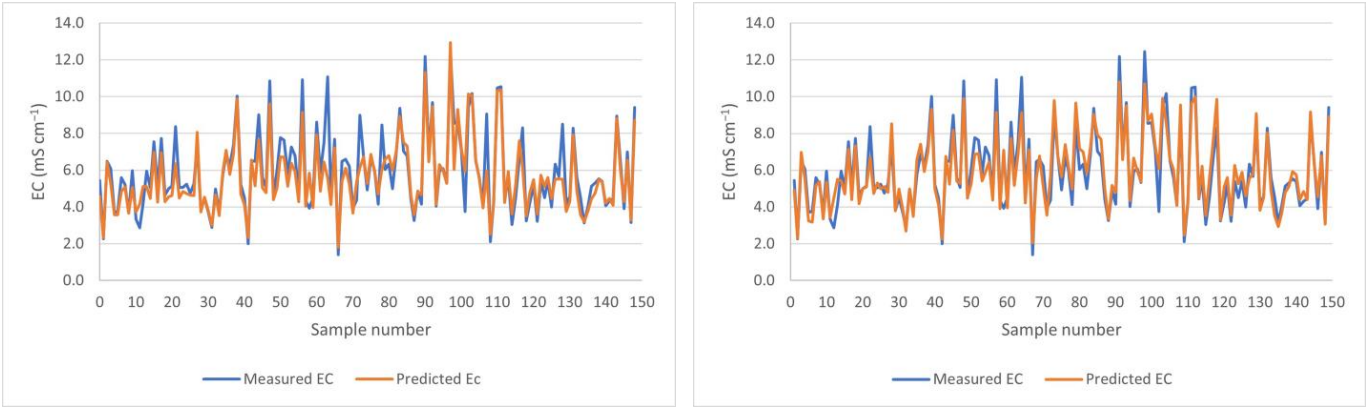
The best method will be the one that presents the lowest errors values, and an R^2 value closest to 1. The results obtained (Table 7) show that both models allow an accurate prediction of the EC of the groundwater studied; however, the ANN model provided a slightly better fit, due to the highest value of R^2 and the lowest values of RMSE, MAE, MPE, MBE, SSE, and AADR.

Table 7. Performance of the theoretical models developed.

Model	RMSE	MAE	R^2	MPE	MBE	SSE	AADR
MLR	0.9392	0.6128	0.8340	−2.4679	−0.281	132.043	33.52%
ANN	0.7474	0.5684	0.8774	1.9338	0.007	83.225	33.38%

In addition, an evaluation was carried out to verify that the predictions made by both models did not present significant differences with the real values of EC measured in situ. For this purpose, a chi-squared test (χ^2) was performed. The values obtained were 12.932 and 22.974 for the ANN and MRL models, respectively. The χ^2 was 536.58 for 592 degrees of freedom and a *p*-value of 0.05. Therefore, the null hypothesis was fulfilled in both cases, indicating that the predictions do not significantly differ from the experimental data used to design the models.

Figure 4 shows the profiles obtained by applying the two theoretical models developed in the validation phase (predicted EC), in comparison with the data obtained through the experimental measurement of the EC of groundwater.



(**a**) (**b**)

Figure 4. Data profile of measured and predicted EC by MLR (**a**) and ANN (**b**) models in validation phase. The groundwater samples used in validation phase have been numbered in chronologic order.

Figure 4 also shows that both models can predict the EC in the study groundwater with high accuracy. Of the 149 groundwater samples used in the validation phase of the ANN model, 91 had a relative error lower than 10%, 52 had errors between 10 and 25%, and only 6 showed an error greater than 25%. Regarding the MLR predictions, the results are summarized in that 89, 49, and 11 samples presented relative errors in the prediction

of CE lower than 10%, between 10% and 25%, and greater than 25%, respectively. This is consistent with the error values calculated in the table presented above.

Therefore, it can be concluded that the ANN model obtained a better estimation accuracy than the MLR model. This can also be observed in the curve profiles shown in Figure 4a (MLR model) and Figure 4b (ANN model). Furthermore, as the number of samples used in the estimation increases, the accuracy achieved with ANN models will be higher than that achieved with classical models such as MLR. This neural network makes accurate predictions even when designed with limited data.

Finally, an additional check was made using independent data taken from the Groundwater Quality Monitoring Network [25]. The data corresponded to 35 groundwater samples from September to October 2023. The results were very similar to those obtained in the model validation phases. For the ANN model, 22 samples presented a relative error of less than 10% in the EC prediction, 12 between 10 and 25%, and 1 with an error greater than 25%. For the MLR model, 17, 14, and 4 groundwater samples showed relative prediction errors of less than 10%, between 10 and 25%, and greater than 25%, respectively.

Our results agree with those reported by other works that found ANN and MLR models suitable for the prediction of parameters related to salinity and groundwater quality [14,16,19,20]. Ekemen Keskin et al. [14] also concluded that the ANN-based model was slightly more accurate than the MLR for irrigation groundwater management in a region of Turkey. Singh et al. [19] also found the ANN model more accurate than MLR for predicting various salinity parameters in the groundwater of Punjab (India), at least during the post-monsoon season. According to Kouadri et al. [20], the ANN technique may have a higher accuracy in predicting groundwater quality compared to MLR due to its ability to model complex interactions. The author found similar results in the prediction of irrigation groundwater quality in the Akot basin area (India).

3.5. Suitability of Agricultural Use of Groundwater

Agriculture is a key sector for the development of the Campo de Cartagena region and the whole of the Region of Murcia due to the high quality of land and its mild climatic conditions. In this century, the area of the region dedicated to dryland agriculture in Campo de Cartagena has been notably reduced (from 53,000 ha in 2000 to 9000 ha in 2022), while it has remained practically constant for irrigated agriculture (from 36,000 to 30,000 ha) [28]. Regarding the type of crop, the distribution is heterogeneous, with herbaceous crops above 20,000 ha in 2021 (vegetables, root crops, grain cereals, forage crops, etc.) and woody crops above 15,000 ha (citrus, non-citrus fruit, olive, and vineyards).

The FAO guidelines for the interpretation of water quality for irrigation classify water based on its salinity, measured as EC at 25 °C. There is no restriction on the use of water for irrigation when the EC is less than 0.7 mS cm^{-1}, and there are severe restrictions at values above 3.0 mS cm^{-1}. Between 0.7 and 3.0 mS cm^{-1} EC, water is characterized as light to moderate in terms salinity and its use will depend on the type of crop [28].

From the results shown in Table 1, the EC of groundwater in the Campo de Cartagena region varied from 0.4 to 12 mS cm^{-1}. The average value of 6 ± 2 mS cm^{-1} does not exceed the threshold value (6.3 mS cm^{-1}), established by Spanish regulation as indicator of groundwater salinization in the region [21]. Approximately 32% of the groundwater samples exceeded this threshold value. However, the hydrogeology of the area causes a high natural salinity in the aquifers (the background level of EC is 5.8 mS cm^{-1}), so 95% of the groundwater samples had an EC above the limit considered by the FAO to be used in irrigation crops (3.0 mS cm^{-1}).

Table 8 shows some crops of Campo de Cartagena for which the FAO has compiled tolerance and yield potentials based on the salinity of irrigation water [28]. Information on the area occupied by these crops is also included, according to the statistical portal of the region (data for 2021) [29].

Table 8. Tolerance and yield potential for some crops of Campo de Cartagena to the salinity of irrigation water in terms of EC (mS cm^{-1}) [28,29].

Crops		Area (ha)	Tolerance to Salinity	EC (mS cm^{-1})/Crop Yield			
				100%	75%	50%	0%
Cereals for grain	Barley	403	Tolerant	5.3	8.7	12	19
	Wheat	626	Moderately tolerant	4.0	6.3	8.7	13
Forage crops	Alfalfa	64	Moderately sensitive	1.3	3.6	5.9	10
Vegetable crops	Broccoli	2056	Moderately sensitive	1.9	3.7	5.5	9.1
	Bean	184	Sensitive	0.7	1.5	2.4	4.2
	Lettuce	3502	Moderately sensitive	0.9	2.1	3.4	6.0
	Tomato	103	Moderately sensitive	1.7	3.4	5.0	8.4
	Watermelon	3504	Moderately sensitive	3.1	4.9	6.7	10
Root crops	Potato	3007	Moderately sensitive	1.1	2.5	3.9	6.7
Woody crops	Almond	5093	Sensitive	1.0	1.9	2.8	4.5
	Apricot	4	Sensitive	1.1	1.8	2.5	3.8
	Circle	3	Sensitive	1.0	1.9	2.9	4.7
	Grape	57	Moderately sensitive	1.0	2.7	4.5	7.9
	Citrus	8399	Sensitive	1.1	2.2	3.2	5.3
	Peach	54	Sensitive	1.1	1.9	2.7	4.3

As can be seen from the information provided in Table 8, most of the dominant crops in the region, such as almond trees and citrus trees, are very sensitive to the salinity of irrigation water. More than 90% of groundwater considered in this study would cause significant reductions in the yield of these crops, to less than 50%, if used for direct irrigation. Other crops of great importance in the area, such as some vegetables and tubers, would see their yields considerably affected. This is the case for broccoli, lettuce, watermelon, and potatoes, which are moderately sensitive to EC irrigation water. Only the crops most tolerant to water salinity (barley) could be irrigated with approximately 50% of the area's groundwater without seeing their yield diminished.

With the purpose of reducing the salinity of the aquifers, there are more than 1000 small private desalination plants in the area, the oldest of which were built in the late 1980s [30]. Irrigation is carried out with a mixture of desalinated groundwater and water from some aquifers, up to the required salinity concentration, or, depending on availability, with TSWT or water from seawater desalination plants (Torrevieja and El Mojón) [31]. In 2022, annual desalinated volumes were in the order of 33 hm^3 [32]. The brines from desalination are managed by local authorities and ultimately discharged into coastal waters. However, sometimes the existing infrastructure cannot receive all the rejected brine generated in the area, causing uncontrolled leaks in certain places, which affect the soil and groundwater.

The desalination of groundwater can contribute to the agricultural development of Campo de Cartagena region, but it also has associated economic and environmental costs due to the need to process the waste appropriately. Therefore, EC prediction using ANN can help water management, through quick and easy-to-use methods, to assess the need to carry out desalination actions, depending on the specific characteristics of the crop and the water in the selected area.

4. Conclusions

The sustainability and survival of agricultural areas are very important in Spain, as well as for vegetable and fruit supply to Europe. Areas with intensive agriculture and low rainfall depend largely on the management of irrigation water from groundwater, particularly during times of drought. Groundwater with high salinity levels, such as those found in Campo de Cartagena, cannot be used for the irrigation of predominant crops in the area without prior desalination to obtain high yields.

Therefore, groundwater quality assessment is crucial for environmental and agricultural purposes. The advantages of using ANN models in this field are becoming increasingly evident. Accurate prediction models of groundwater characteristics are essential to economize on groundwater assessment and monitoring.

This work examines groundwater salinity from a theoretical perspective, specifically in terms of EC, using both traditional MLR and ANN models. The input variables considered include the concentrations of the main anions (bicarbonate, chloride, nitrate, and sulphate) and cations (calcium, magnesium, potassium, and sodium) in the water, as well as pH and temperature. Preliminary statistical analyses revealed that the concentration of calcium, chloride, magnesium, sodium, and sulphate had the most significant impact. Both models showed high accuracy in predicting the EC of the waters under study, although is worth noting that the accuracy of the ANN model was slightly better than that of the MLR model.

This study has demonstrated that ANNs are a valuable tool for predicting groundwater salinity, in terms of EC, without having to make measurements in situ, which is of great interest in the case of deep groundwater, where in situ measurement can be difficult and susceptible to possible measurement errors. This is particularly relevant in semi-arid and coastal areas, such as the Campo de Cartagena region (Murcia, Spain). In these areas, crop management is hindered by the scarcity of surface irrigation water and the high salinity levels of irrigation groundwater, and crop tolerance to salinity may also be considered. Therefore, theoretical prediction models can be a useful tool for irrigation water management in semi-arid environments where aquifers are subject to intensive exploitation due to irrigation needs.

To enhance the results, future research should focus on implementing studies for a more accurate assessment of water quality. It may be of interest to develop advanced AI models to evaluate and predict other irrigation water quality indicators, such as SAR, PI, KR, and %Na, mentioned in this work, and also, other contaminants such as excess nutrients (especially nitrates) or the presence of heavy metals. Furthermore, it is essential to investigate how climate change affects the quantity and quality of groundwater in Campo de Cartagena to adopt the required management strategies.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/agronomy14030524/s1: Table S1: Geographic coordinates of the 38 sampling stations studied in Campo de Cartagena (CHS designations) and sampling depth, differentiating between those located to the south, center, and north.

Author Contributions: Conceptualization, L.F.M., M.I.M.-L., E.M.G.-d.-T., M.Á.Q. and S.G.-S.; methodology, M.I.M.-L., E.M.G.-d.-T., M.Á.Q. and S.G.-S.; software, L.F.M.; validation, M.Á.Q., S.G.-S. and M.I.M.-L.; formal analysis, L.F.M. and E.M.G.-d.-T.; investigation, L.F.M., M.I.M.-L., E.M.G.-d.-T., M.Á.Q. and S.G.-S.; writing—original draft preparation, M.I.M.-L., E.M.G.-d.-T. and M.Á.Q.; writing—review and editing, M.Á.Q., M.I.M.-L., S.G.-S. and E.M.G.-d.-T.; supervision, M.Á.Q. and M.I.M.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wada, Y.; Van Beek, L.P.H.; Van Kempen, C.M.; Reckman, J.W.T.M.; Vasak, S.; Bierkens, M.F.P. Global Depletion of Groundwater Resources. *Geophys. Res. Lett.* **2010**, *37*, 1–5. [CrossRef]
2. Velis, M.; Conti, K.I.; Biermann, F. Groundwater and Human Development: Synergies and Trade-Offs within the Context of the Sustainable Development Goals. *Sustain. Sci.* **2017**, *12*, 1007–1017. [CrossRef] [PubMed]
3. Qian, H.; Chen, J.; Howard, K.W.F. Assessing Groundwater Pollution and Potential Remediation Processes in a Multi-Layer Aquifer System. *Environ. Pollut.* **2020**, *263*, 114669. [CrossRef] [PubMed]
4. Taylor, M.; Elliott, H.A.; Navitsky, L.O. Relationship between Total Dissolved Solids and Electrical Conductivity in Marcellus Hydraulic Fracturing Fluids. *Water Sci. Technol.* **2018**, *77*, 1998–2004. [CrossRef] [PubMed]

5. Lam, V.P.; Kim, S.J.; Park, J.S. Optimizing the Electrical Conductivity of a Nutrient Solution for Plant Growth and Bioactive Compounds of Agastache Rugosa in a Plant Factory. *Agronomy* **2020**, *10*, 76. [CrossRef]
6. Foster, S.; Pulido-Bosch, A.; Vallejos, Á.; Molina, L.; Llop, A.; MacDonald, A.M. Impact of Irrigated Agriculture on Groundwater-Recharge Salinity: A Major Sustainability Concern in Semi-Arid Regions. *Hydrogeol. J.* **2018**, *26*, 2781–2791. [CrossRef]
7. Ministerio para la Transición Ecológica y el Reto Demográfico. *Análisis de Soluciones para el Vertido Cero al Mar Menor Proveniente del Campo de Cartagena. Estudio del Impacto Ambiental después de la Información Pública*; APÉNDICE 1 Diagnóstico: Madrid, Spain, 2019; pp. 1–23.
8. Confederación Hidrográfica del Segura (CHS). Anejo 02 Inventario de Recursos. Plan Hidrológico de la Demarcación del Segura 395 2022/2027. Available online: https://www.chsegura.es/es/cuenca/planificacion/planificacion-2022-2027/plan-hidrologico-2022-2027/ (accessed on 12 December 2023).
9. Martinez Mechón, M.; Senent Alonso, M. El Agua En El Campo de Cartagena. *Rev. Murc. Antropol.* **2007**, *14*, 47–62.
10. Jiménez-Martínez, J.; Aravena, R.; Candela, L. The Role of Leaky Boreholes in the Contamination of a Regional Confined Aquifer. A Case Study: The Campo de Cartagena Region, Spain. *Water. Air. Soil Pollut.* **2011**, *215*, 311–327. [CrossRef]
11. Domingo-Pinillos, J.C.; Senent-Aparicio, J.; García-Aróstegui, J.L.; Baudron, P. Long Term Hydrodynamic Effects in a Semi-Arid Mediterranean Multilayer Aquifer: Campo de Cartagena in South-Eastern Spain. *Water* **2018**, *10*, 1320. [CrossRef]
12. Rajaee, T.; Ebrahimi, H.; Nourani, V. A Review of the Artificial Intelligence Methods in Groundwater Level Modeling. *J. Hydrol.* **2019**, *572*, 336–351. [CrossRef]
13. Barzegar, R.; Asghari Moghaddam, A. Combining the Advantages of Neural Networks Using the Concept of Committee Machine in the Groundwater Salinity Prediction. *Model. Earth Syst. Environ.* **2016**, *2*, 26. [CrossRef]
14. Ekemen Keskin, T.; Özler, E.; Şander, E.; Düğenci, M.; Ahmed, M.Y. Prediction of Electrical Conductivity Using ANN and MLR: A Case Study from Turkey. *Acta Geophys.* **2020**, *68*, 811–820. [CrossRef]
15. Kheradpisheh, Z.; Talebi, A.; Rafati, L.; Ghaneian, M.T.; Ehrampoush, M.H. Groundwater Quality Assessment Using Artificial Neural Network: A Case Study of Bahabad Plain, Yazd, Iran. *Desert* **2015**, *20*, 65–71. [CrossRef]
16. Haggerty, R.; Sun, J.; Yu, H.; Li, Y. Application of Machine Learning in Groundwater Quality Modeling—A Comprehensive Review. *Water Res.* **2023**, *233*, 119745. [CrossRef]
17. Tao, H.; Hameed, M.M.; Marhoon, H.A.; Zounemat-Kermani, M.; Heddam, S.; Kim, S.; Sulaiman, S.O.; Tan, M.L.; Sa'adi, Z.; Mehr, A.D.; et al. Groundwater Level Prediction Using Machine Learning Models: A Comprehensive Review. *Neurocomputing* **2022**, *489*, 271–308. [CrossRef]
18. Al-Waeli, L.K.; Sahib, J.H.; Abbas, H.A. ANN-Based Model to Predict Groundwater Salinity: A Case Study of West Najaf-Kerbala Region. *Open Eng.* **2022**, *12*, 120–128. [CrossRef]
19. Singh, G.; Singh, J.; Wani, O.A.; Egbueri, J.C.; Agbasi, J.C. Assessment of Groundwater Suitability for Sustainable Irrigation: A Comprehensive Study Using Indexical, Statistical, and Machine Learning Approaches. *Groundw. Sustain. Dev.* **2024**, *24*, 101059. [CrossRef]
20. Kouadri, S.; Pande, C.B.; Panneerselvam, B.; Moharir, K.N.; Elbeltagi, A. Prediction of Irrigation Groundwater Quality Parameters Using ANN, LSTM, and MLR Models. *Environ. Sci. Pollut. Res.* **2022**, *29*, 21067–21091. [CrossRef] [PubMed]
21. García-del-Toro, E.M.; García-Salgado, S.; Mateo, L.F.; Quijano, M.Á.; Más-López, M.I. Machine Learning as a Diagnosis Tool of Groundwater Quality in Zones with High Agricultural Activity (Region of Campo de Cartagena, Murcia, Spain). *Agronomy* **2022**, *12*, 3076. [CrossRef]
22. Alcolea, A.; Contreras, S.; Hunink, J.E.; García-Aróstegui, J.L.; Jiménez-Martínez, J. Hydrogeological Modelling for the Watershed Management of the Mar Menor Coastal Lagoon (Spain). *Sci. Total Environ.* **2019**, *663*, 901–914. [CrossRef] [PubMed]
23. Instituto Tecnológico Geominero de España (ITGE). *Estudio Hidrogeológico del Campo de Cartagena (2a FASE)*; Geological Survey of Spain: Madrid, Spain, 1991; p. 145.
24. Instituto Tecnológico Geominero de España (ITGE). *Las Aguas Subterráneas del Campo de Cartagena (Murcia)*; Geological Survey of Spain: Madrid, Spain, 1993; p. 61. Available online: https://www.igme.es/igme/publica/libros1_HR/libro84/Lib84.htm (accessed on 17 December 2023).
25. Confederación Hidrográfica del Segura (CHS). Redes de Control de las Masas de Agua Subterráneas. Available online: https://www.chsegura.es/es/cuenca/redes-de-control/calidad-en-aguas-subterraneas/acceso-a-los-datos/ (accessed on 26 February 2024).
26. García Del Toro, E.M.; Mateo, L.F.; García-Salgado, S.; Más-López, M.I.; Quijano, M.Á. Use of Artificial Neural Networks as a Predictive Tool of Dissolved Oxygen Present in Surface Water Discharged in the Coastal Lagoon of the Mar Menor (Murcia, Spain). *Int. J. Environ. Res. Public Health* **2022**, *19*, 4531. [CrossRef] [PubMed]
27. Nazari, H.; Taghavi, B.; Hajizadeh, F. Groundwater Salinity Prediction Using Adaptive Neuro-Fuzzy Inference System Methods: A Case Study in Azarshahr, Ajabshir and Maragheh Plains, Iran. *Environ. Earth Sci.* **2021**, *80*, 152. [CrossRef]
28. Ayers, R.S.; Westcot, D.W. *Water Quality for Agriculture*; FAO Irrigation and Drainage Paper; Food and Agriculture Organization of the United Nations: Rome, Italy, 1985; Volume 29, pp. 7–10, 31–35.
29. Centro Regional de Estadística de Murcia (CREM). Anuario Estadístico de la Región de Murcia 2021. Available online: https://econet.carm.es/inicio/-/crem/anuario/actual/anuario_tomoII_4.html (accessed on 3 December 2023).
30. Aparicio, J.; Candela, L.; Alfranca, O.; García-Aróstegui, J.L. Economic Evaluation of Small Desalination Plants from Brackish Aquifers. Application to Campo de Cartagena (SE Spain). *Desalination* **2017**, *411*, 38–44. [CrossRef]

31. Confederación Hidrográfica del Segura (CHS). Desalinizacion. Available online: https://www.chsegura.es/es/cuenca/caracterizacion/recursos-hidricos/desalinizacion/ (accessed on 25 March 2023).
32. Comunidad de Regantes del Campo de Cartagena. Memoria de Actividades Año 2022/2023. 2023. Available online: https://www.crcc.es/wp-content/uploads/2023/07/Memoria-2022.pdf (accessed on 12 December 2023).