

## Article

# Water Quality Assessment and Forecasting Along the Liffey and Andarax Rivers by Artificial Neural Network Techniques Toward Sustainable Water Resources Management

Eyad Abushandi 

Department of Geology, Faculty of Science, Technology, Engineering and Mathematics, Trinity College Dublin, The University of Dublin, D02 PN40 Dublin, Ireland; e.abushandi@tcd.ie

**Abstract:** This research evaluates water quality in two contrasting hydro-climatic regions: the River Liffey in Ireland and the Andarax River in Spain. It utilizes an Artificial Neural Network (ANN) to simulate potential changes in key water-quality parameters based on field measurements. The ANN models showed strong predictive efficiency and performance, achieving  $R^2$  values of 0.89 for dissolved oxygen (DO), 0.98 for electrical conductivity (EC), 0.87 for pH, 0.95 for total dissolved solids (TDS), and 0.96 for turbidity. The root mean-square-error (RMSE) values for important parameters were DO (1.25 mg/L), EC (48.06  $\mu\text{S}/\text{cm}$ ), and turbidity (8.9 FNU). The models were able to capture complex nonlinear relationships under different environmental conditions. The results showed that DO levels in the Liffey will decline by up to 20% over the next decade due to rising nutrient pollution, while TDS levels in the Andarax River are expected to rise by approximately 15% during the same period as a result of ongoing agricultural runoff. The study also simulated potential future hypothetical scenarios by applying the model to four different “what-if” situations. Overall, the research underscores the significance of machine learning in understanding intricate water-quality dynamics.

**Keywords:** artificial intelligence; hydrological modeling; climate impact; water resource sustainability; ANN applications



Academic Editors: Christos S. Akratos and Anastasia Nikolaou

Received: 21 December 2024

Revised: 25 January 2025

Accepted: 1 February 2025

Published: 6 February 2025

**Citation:** Abushandi, E. Water Quality Assessment and Forecasting Along the Liffey and Andarax Rivers by Artificial Neural Network Techniques Toward Sustainable Water Resources Management. *Water* **2025**, *17*, 453. <https://doi.org/10.3390/w17030453>

**Copyright:** © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the context of sustainability, one of the greatest challenges facing the regions of the modern world is the management of freshwater resources, which is primarily due to population growth, climatic changes, and human activities. Rivers are susceptible to these pressures, which can substantially impact their water quality. Understanding and predicting changes in river water quality is essential for the development of effective water management policies. However, this task requires sophisticated modeling techniques, as these are interconnected through the complex dynamics of water systems, which in turn are driven by both natural and human-induced factors. The River Liffey in Ireland represents a temperate maritime climate, while the River Andarax in Spain represents the semi-arid part of the country. Therefore, both rivers can be used as an ideal case study for a comparative assessment. These rivers have different climatic regimes and hydrological parameters, and they differ in terms of anthropogenic pressures. The River Liffey is subject to urban runoff, sewage emissions, and agricultural functions, leading to nutrient enrichment with fluctuating dissolved oxygen (DO) levels and nutrient pollution. However, the challenges faced by the Rivers Liffey and Andarax mirror situations many rivers worldwide face. A study by [1] has recorded urbanization’s strong impact on the River Liffey’s water quality,

particularly high nitrate levels and low oxygen concentrations in parts near Dublin. The Andarax River is facing a lack of water quality and quantity. This is due to low rainfall, high evaporation rates, and return flows from irrigation leading to high levels of total dissolved solids (TDS), electrical conductivity (EC) [2], and fertilizers such as nitrogen, phosphorus, and potassium (NPK).

In recent years, Artificial Neural Networks (ANNs) have gained wide recognition due to their ability to model such complexities in environmental science with high accuracy in both the classification and prediction of water-quality parameters (e.g., [3,4]). The ANN models have been applied in various studies to predict water-quality parameters such as DO, TDS, pH, and EC, often with better results than traditional statistical methods. Ref. [5] achieved a notable performance of the ANN models in predicting water-quality parameters in urban rivers by incorporating natural and human-induced factors. The results showed that the proposed model showed better prediction accuracy with an  $R^2$  as high as 0.93 in some cases and has been proven robust to capture the complex nonlinear relationships of water-quality variables. A study conducted by [6] investigated the application of ANN in modeling DO in the Mathura River. The predicted DO values show a significant correlation with the measured values. The predicted values of DO show prominent accuracy by producing high correlations ( $R^2$  up to 0.9) between measured and predicted values.

Similarly, Ref. [7] assessed the performance of machine-learning and regression techniques in predicting TDS and EC in river water. They found that gene-expression programming (GEP) performed better among the models than the regression techniques. They emphasized the potential contribution of machine-learning models, such as ANN, to the management of sustainable water resources through efficient water-quality predictions. In [8], five artificial intelligence methods were applied to predict daily discharge for sustainable management of water resources. In general, the five methods showed satisfactory prediction performance with few outliers. Detection of the impact of agricultural activities on river water quality using ANN was presented by [9]. The prediction of the water quality of the Nile River emphasizes the integration of physio-chemical parameters to achieve accurate predictions. The assessment of key physio-chemical parameters at 51 surface water sites shows that 98% of the water samples have no limitations for irrigation.

The primary objectives of this study are to assess water quality across two contrasting climatic environments, the River Liffey in Ireland and the Andarax River in Spain, and simulate potential changes using Artificial Neural Network (ANN) techniques under potential hypothetical scenarios. In addition, the ANN model was also used to test the water-quality dynamics against four hypothetical “what-if” scenarios that comprised increased nitrogen and phosphorus loads from industrial and agricultural effluents, plus extreme hydrological alterations associated with flash flooding. Such scenarios offer ample knowledge of the potential effects of anthropogenic and environmental stressors on water quality within diversified climatic conditions.

The study hypothesizes that ANN models can efficiently simulate key water-quality parameters under contrasting climatic and hydrological conditions for proactive water management strategies. However, it is important to examine the relationship between water-quality degradation due to nutrients in urban rivers and draw analogies with the River Liffey, whose water quality has deteriorated due to nutrient loading from urban growth. The Andarax River has undergone an intensive study of the impact of human activity and arid conditions on water quality and sediment composition. This highlights the problem of how to control water quality in an expanding urbanized area such as Dublin or Almeria.

## 2. Materials and Methods

### 2.1. Study Areas Description

The catchment area of the River Liffey is characterized by a temperate climate and flows in the eastern part of Ireland (Figure 1: left). The length of the river is about 125 km from its source in the Wicklow Mountains to Dublin Bay, where it then flows into the Irish Sea. The catchment area is approximately 1200 km<sup>2</sup>. The river flows through several towns before reaching the estuary, providing a diverse urban and rural population. Rainfall is evenly distributed throughout the year, with an average of about 2000 mm/year in the mountains of Wicklow and about 750 mm/year in the catchment downstream. The perennial river has seasonal fluctuations, with a higher flow in winter and a lower, more stable flow in summer. The river also supports the Pollaphuca Reservoir, which controls the flow and supplies Dublin with fresh water. The main source of pollution along the river is agricultural and urban runoff. Nutrient pollution from nitrogen and phosphorus is the cause of eutrophication in these sections. Municipal wastewater, particularly from the middle and downstream sections of the river, introduces a certain number of pollutants into the river, although treatment is conducted before discharge into the river.

The Andarax River, within a semiarid climate, is an ephemeral stream that flows in the southeast part of Spain with a length of 67 km starting from the Sierra Nevada Mountains to the Mediterranean Sea, approximately to the east of the city of Almería (Figure 1: right). The catchment area is relatively bigger than the River Liffey catchment area, with about 2000 km<sup>2</sup>. The Andarax River is characterized by flow that depends on seasonal rainfall and snowmelt from the surrounding mountains. This geographical entity has been characterized by long, dry summers with very low annual rainfall. The average annual rainfall is estimated at 250 mm a year. Severe, spasmodic rainfall events causing flash floods are usually separated by longer periods of dry season.

The River Liffey was selected for this study because it has been considered an important source of drinking water for the major metropolitan area of Dublin, as well as agricultural and industrial activities. The continued flow and resulting steady monitoring data present a helpful baseline in which to test the application of ANN models for water-quality forecasting in temperate conditions. The Andarax River represents a complicated hydrologic regime associated with relevant water management problems related to quantity and quality. The flow is very variable, with highly fluctuated rainfall, which makes this case relevant for testing ANN model performance under extreme environmental conditions.

The mean monthly flow values of the River Liffey and the Andarax River show significant differences between the two regimes (Figure 2). The River Liffey has its highest flow in winter months, from December to February, gradually falling to the lowest value in summer, namely June, July, and August. This reflects the regular annual pattern that would be expected in a temperate climate, with constant rainfall and lower evaporation in cooler months. In contrast, the River Andarax has much lower flows, reflecting its semi-arid environment. It has a minor increase in late winter–early spring from March to April, which can be attributed to intermittent rain or snow melt upstream. However, during summer, the flow is practically zero. The collected data months (March and April) capture the transition from high flow in winter due to rainfall and reduced evapotranspiration to moderate flow levels in spring.

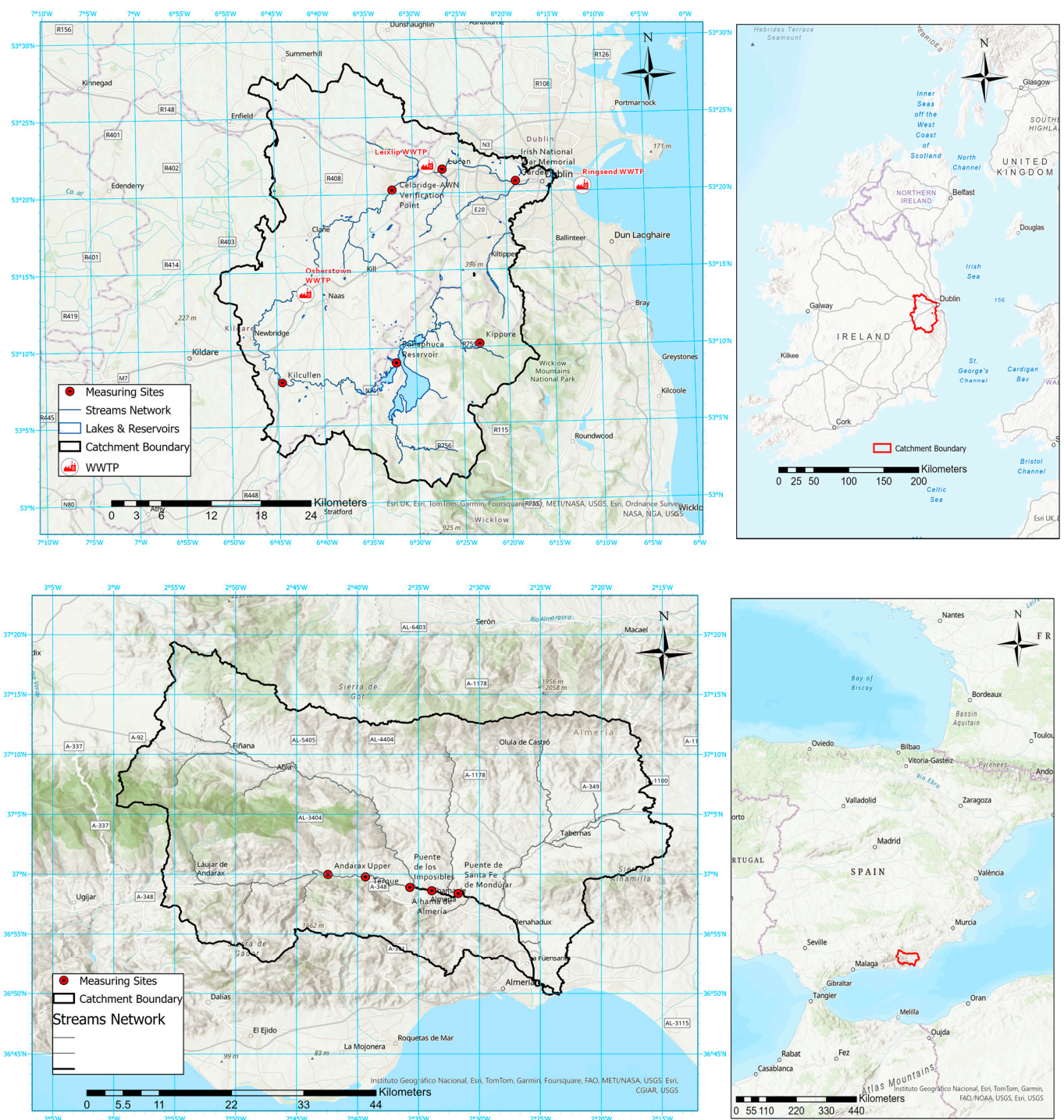


**Figure 1.** Catchment and stream network map, including the sites of measurements for the River Liffey catchment (**top**) Ireland, and the Andarax River catchment (**bottom**) Spain. (Scale 1:320,000).

The catchments of the rivers Liffey and Andarax represent distinctly different landscape types with different land covers that determine their hydrological characteristics, water quality, and ecosystem dynamics (Figure 3). Irrigation of agricultural land is widespread in the wider area. Land cover in the Liffey catchment is dominated by agricultural land, woodland, and semi-natural areas, with numerous wetlands and peatlands also present. Agricultural land use is widespread in most central and eastern parts of the catchment, while woodland and semi-natural areas dominate at higher altitudes or in regionally restricted areas such as the Wicklow Mountains. Artificial surface cover is quite widespread

near urban centers, particularly around Dublin, where land development for agriculture and housing has occurred, resulting in urbanized landscapes.

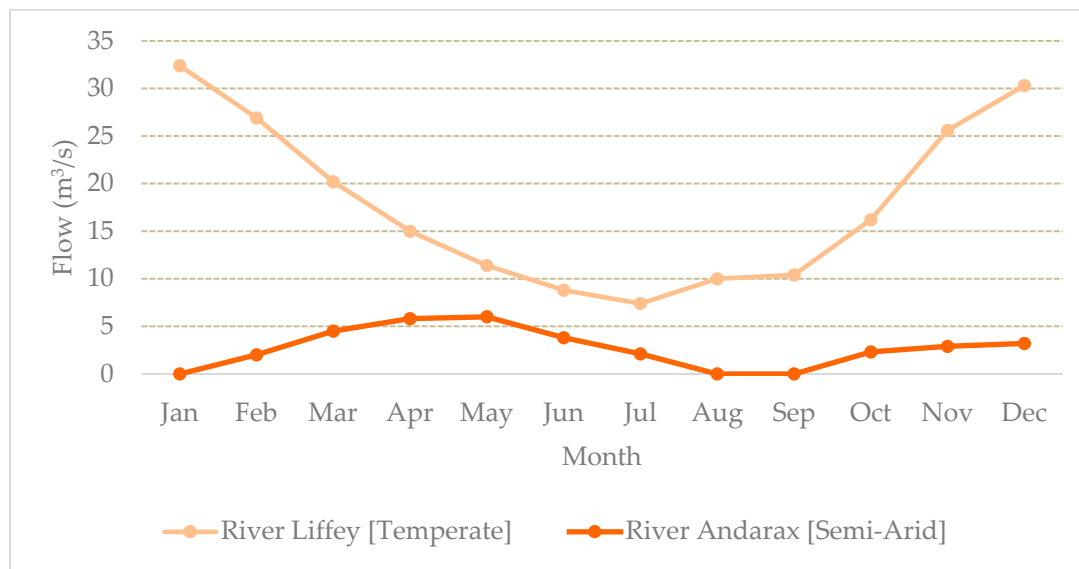


**Figure 2.** Downstream mean monthly flow for the River Liffey (Leixlip Station) from 1978 to 2021, and for the River Andarax (Terque Station) from 1997 to 2024 (Raw Data from Environmental Protection Agency (Ireland) and Automatic Hydrological Information Systems (SAIH) (Spain)).

In contrast, the catchment area of the Andarax extends over a much drier and steeper landscape in Southern Spain, where the semi-arid climate means that, in addition to bare rock or soil, there is mainly natural grassland. The agricultural land is limited to valleys, plains, and flat areas where irrigation is possible, while the man-made areas are found in the urban areas near Almería. Only small parts of the catchment area are covered with forests and semi-natural areas, most of them at higher altitudes. The sparse vegetation and the intensive cover of bare soil led to higher erosion rates and greater variability in water flow.

## 2.2. Data Sources, Variables, and Preparation

Water-quality data from the River Liffey (Ireland) and the River Andarax (Spain) were used to model and predict water-quality trends. The main water-quality parameters monitored in different sections of the two rivers include DO, pH, turbidity, TDS, EC, and NPK levels. These parameters were selected because they are directly related to water-quality classification and understanding the overall state of the aquatic ecosystem.

Data were collected in both rivers using a combination of in situ water-quality meters and analysis. Two monitoring campaigns were conducted in 2023 and 2024 for each river across 11 monitoring stops; 6 stops along the River Liffey during March from each year, and 5 points along the River Andarax during April from each year, resulting in a total of 44 samples and a range of two samples per each point to ensure accuracy. This temporal window was chosen to capture typical hydrological conditions for both rivers during periods of seasonal variation. The collected key physicochemical parameters were then integrated into artificial neural network models.

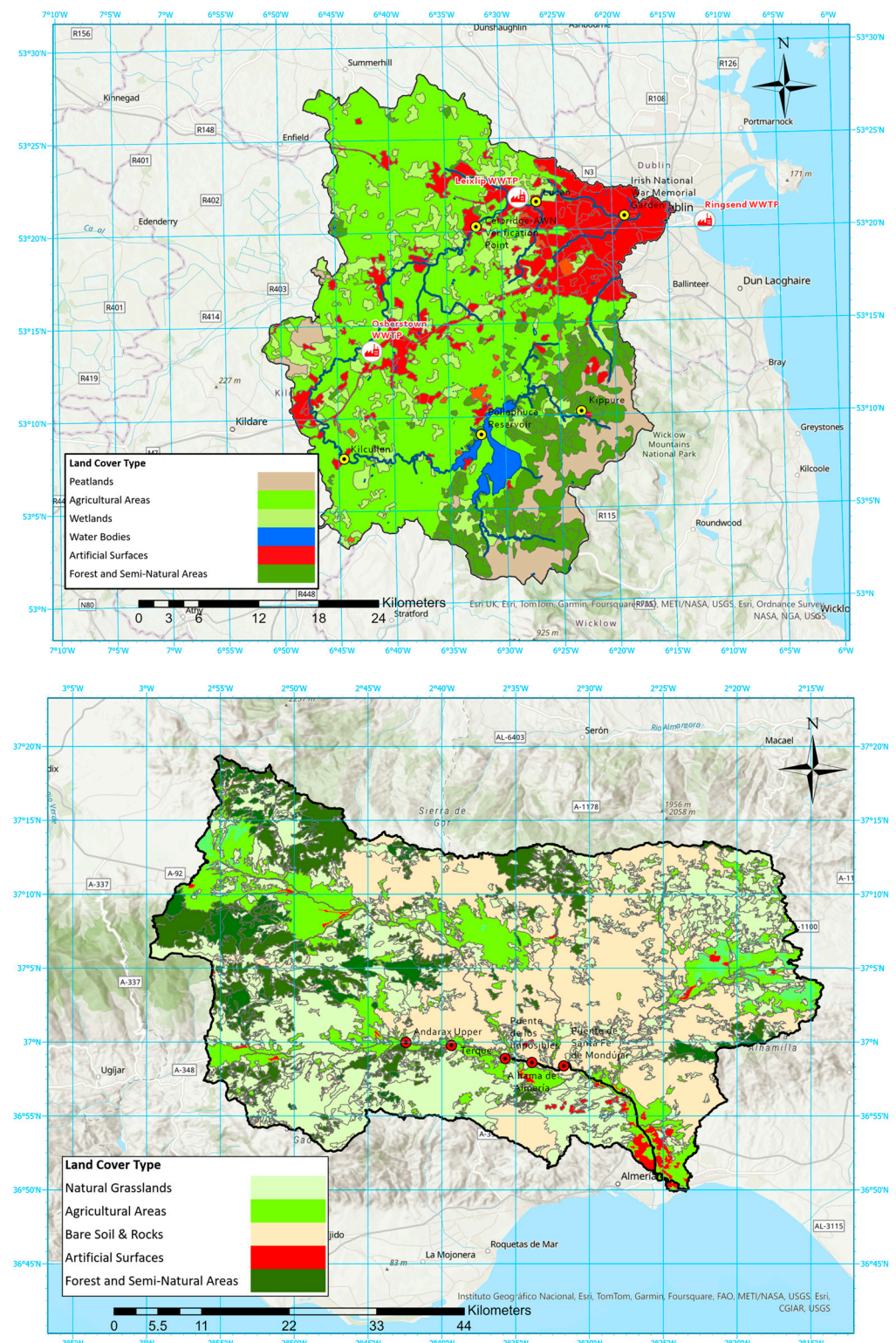


**Figure 3.** Land-cover types for the River Liffey catchment, Ireland (**top**), and the Andarax River catchment, Southern Spain (**bottom**).

### 2.3. Pearson Correlation Analysis

Pearson correlation analysis was performed to understand the direction of relationships that exist among the water-quality parameters and how strong relationships are between pairs of variables. A correlation coefficient ranging from  $-1$  to  $+1$  indicates the strength and direction of a relationship, with an approximate value of  $+1$  representing a perfect positive correlation. Conversely, if the  $r$  value is close to  $-1$ , that shows a high de-

gree of negative correlation. When  $r$  is around 0, then there is no serious linear relationship. The equation for Pearson's correction is as follows:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

where:

$r$ : Pearson correlation coefficient

$x_i$ : Individual data points for variable  $x$

$y_i$ : Individual data points for variable  $y$

$\bar{x}$ : Mean of variable  $x$

$\bar{y}$ : Mean of variable  $y$ .

$n$ : Number of data points.

Correlation analysis was conducted for parameters such as dissolved oxygen, turbidity, electrical conductivity, and total dissolved solids to show interdependence and identify patterns in the data.

#### 2.4. Artificial Neural Network (ANN) Model Description

A forward-looking ANN (Multilayer Perceptron) has been developed to predict water-quality parameters at various locations on the Rivers Liffey and Andarax based on measured water-quality parameters. The greatest strength of the ANN is its ability to represent complex non-linear relationships between the input variables DO, EC, TDS, pH, and turbidity. This involves the use of 12 neurons in the input layer to represent the water-quality parameters, followed by two hidden layers representing 64 and 32, respectively. The hidden neurons use the ReLU (Rectified Linear Unit) as the activation function and nonlinearity. The output layer consisted of one neuron used for the prediction of each water-quality parameter. The selection was informed by performance metrics such as the root-mean-squared error and the coefficient of determination during cross-validation to make sure that the architecture was optimal with regard to dataset size and structure. This architecture allowed the ANN to learn complex patterns in the data while limiting the risk of overfitting.

The architecture of the Artificial Neural Network (ANN) was designed to effectively capture the relationships between water-quality parameters and their variations under different scenarios. The input layer consists of 12 neurons, corresponding to the 12 measured water-quality parameters, ensuring that all relevant variables are represented in the model. Two hidden layers were included to capture the complex, non-linear interactions between the input parameters. The first hidden layer contains 64 neurons, while the second has 32 neurons. This configuration was determined through iterative experimentation and hyperparameter tuning, balancing the trade-off between model complexity and computational efficiency. The selection was guided by performance metrics such as the root-mean-squared error (RMSE) and the coefficient of determination ( $R^2$ ) during cross-validation, ensuring that the architecture was optimal for the dataset size and structure. This design allowed the ANN to learn intricate patterns in the data while minimizing the risk of overfitting.

Synthetic data points were generated by the Gaussian process to overcome limitations due to the small dataset size. This augmentation process conserved the variability and structure of the observed data (44 points) while effectively increasing the size of the training set, thus enabling better model generalization. In addition, cross-validation and dropout layers were further employed to enhance model performance. Each of the 44 original samples was augmented by 4 synthetic samples, adding 176 new data points and bringing the total number of samples to 220. This was done in such a way that the statistical properties and physical relevance were preserved in the augmented data.

This model used data split into training and test data: 70% for training, 15% for validation, and 15% for testing. The backpropagation training algorithm was used with Adam as the optimization method and a stack size of 32 (Figure 4).

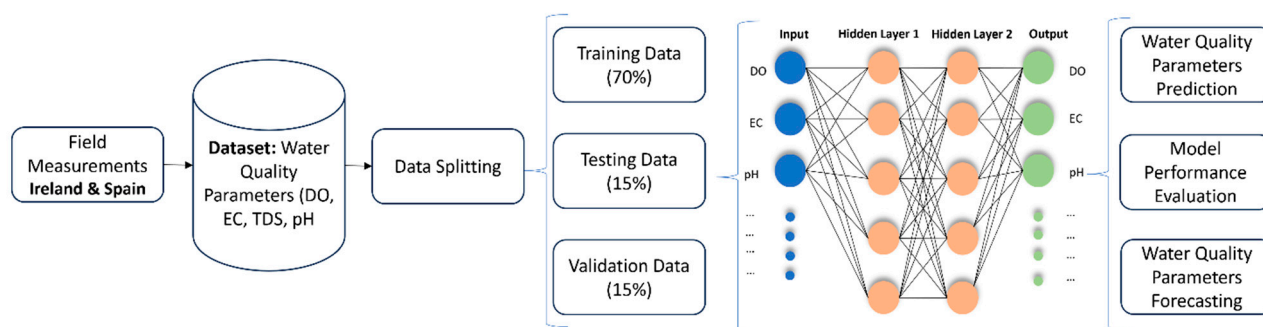


**Figure 4.** Methodology and ANN model architectures used for this study.

All input parameters were normalized using min–max scaling to provide consistency across all variables with different units and scales. This preprocessing step was very important to enhance ANN model performance by keeping the magnitude of input variables standardized, which cannot allow large-scale variables to influence the behavior of the model disproportionately. The normalization process ensures the representation of all the parameters and helps the model converge much faster with good numerical stability.

Furthermore, various strategies were employed during the ANN model's training to prevent overfitting. An early stopping step with monitoring based on validation loss was employed to ensure that training stops when it fails to generalize well on some unseen dropout layers at a rate of 0.2.

L2 regularization of the weights, which penalizes large parameter values, encourages more generalized models. These techniques greatly combine to reduce the possibility of overfitting, hence letting the ANN work well on a comparatively small dataset.

Consequently, the performance of the ANN model was assessed using the test set. The effectiveness of the model was quantified using the root-mean-square error (RMSE) of the model fit and the  $R^2$ , as these two indicators show the overall accuracy of the model.

Table 1 shows the potential of implementing the ANN model for different environmental conditions. It compares the applications under different climatic, hydrological, and data-related conditions that can support sustainable water resources management in different regions. The table provides a comparative framework for the study areas and identifies several broader challenges that need to be dealt with in these two contrasting environments.

Later, “what-if” scenarios of pollution or hydrological conditions, e.g., increased nitrogen or phosphorus due to increasing agricultural activities, are developed, and how the ANN predicts water quality under these conditions is observed to inform decisions for sustainable management. Accordingly, based on the measured nitrogen and phosphorus concentrations and high-flow events reported in the literature, four simulation scenarios of nitrogen, phosphorus, and flooding were developed. Potassium was not used in the model because of the poor correlation with other parameters. Besides that, the flooding impact was simulated by changing related parameters, such as turbidity values. The model developed different pollution scenarios:

Scenario 1: Higher nitrogen levels due to agricultural runoff

Scenario 2: Higher phosphorus levels due to industrial or agricultural runoff

Scenario 3: Higher nitrogen and phosphorus

Scenario 4: Severe hydrological changes (flash flooding)

**Table 1.** Evaluation of ANN model adaptability for varied climatic and hydrologic regimes.

| Aspect                       | Andarax River<br>(Semi-Arid Environment)                                                                                                 | Liffey River (Temperate<br>Environment)                                                                                 | Comparative Insight for<br>ANN Model                                                   |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|
| Climate                      | Arid, characterized by high evaporation, low precipitation, and episodic flows<br>Rainfall: 250 to 350 mm<br>Temperature: 10 °C to 30 °C | Temperate, with regular rainfall and consistent flow patterns<br>Rainfall: 750 to 2000 mm<br>Temperature: 5 °C to 15 °C | Tests the ANN model's ability to handle extreme variability and stable conditions      |
| Hydrological Dynamics        | Irregular and ephemeral flow, often leading to flash floods and dry periods<br>Flow: 6 m <sup>3</sup> /s (During wet season)             | Steady and perennial flow with seasonal variations, fewer extremes<br>Flow: 18 m <sup>3</sup> /s                        | Assesses the model's adaptability to varying hydrological regimes                      |
| Pollution Sources            | Predominantly agricultural runoff, salinity issues due to low flow                                                                       | Urban, industrial, and agricultural runoff, more regular contamination                                                  | Evaluate how well the model handles different types of pollutants and sources          |
| Water-Quality Fluctuations   | Large fluctuations in parameters like salinity, turbidity, and temperature due to inconsistent flow                                      | More stable water quality with periodic variations in DO, nutrients, and pH                                             | Demonstrates the ANN's capacity to predict both highly variable and stable conditions  |
| Key Water-Quality Challenges | Water scarcity, high salinity, agricultural contamination                                                                                | Urban runoff, industrial pollution, nutrient loading                                                                    | Highlights the model's flexibility in addressing diverse water-quality issues          |
| Sustainability Concerns      | Water management focuses on scarcity, allocation, and salinity control                                                                   | Pollution management, flood risk mitigation, and ecosystem health                                                       | Provides insights into region-specific water management strategies and ANN application |

The scenarios were designed with parameter ranges inside or near the domain of the training data in order to keep the model valid. This way, ANN predictions remain credible and interpretable. The nitrogen-, phosphorus-, and flooding-simulated scenarios were built on plausible variations derived from historical trends and literature values (e.g., [10,11]).

### 3. Results

The water quality of the Liffey appears to be largely stable, while the Andarax shows higher variability. In both rivers, DO is higher in the upper catchment, with around +50% higher value than downstream, indicating that the water flows through more urbanized areas. In general, the Liffey catchment shows relatively stable DO values compared to the Andarax River, with an average of 9.1 mg/L, while the Andarax DO values are 11 mg/L. Due to lower temperatures, the evaporation rates and chemical reactions are much lower in the River Liffey. Therefore, EC values in the River Liffey ranged from 74 µS/cm (Kippure) to 454 µS/cm. The Andarax catchment has much higher EC values, between 776 and 1400 µS/cm, and TDS between 199 and 666 ppm, suggesting a more mineral-rich water, although it may be exposed to drier conditions with a higher proportion of agricultural runoff. However, both catchments have slightly alkaline water (average pH is around 7.78), indicating good buffering capacity. Turbidity in the Andarax catchment is significantly higher with an average of 44.2 FNU, indicating that erosion and sediment transport are a

serious problem in this catchment, probably exacerbated by the dry climate and seasonal flash floods. In contrast, turbidity levels tend to be lower in the Liffey catchment, although sediment load increases downstream due to urban runoff and bank erosion (Table 2).

**Table 2.** Average field measurements of water quality for the River Liffey catchment sites 1–6 and the Andarax River catchment sites 7–11.

| ID | Name                                       | DO (mg/L)   | Temp (°C)   | EC (µS/cm)    | TDS (ppm)     | pH         | Turbidity (FNU) |
|----|--------------------------------------------|-------------|-------------|---------------|---------------|------------|-----------------|
| 1  | Kippure                                    | 12          | 7.5         | 74            | 43            | 7.8        | 2               |
| 2  | Poulaphouca Reservoir                      | 10.5        | 7.5         | 155           | 85            | 7.5        | 5               |
| 3  | Kilcullen                                  | 6           | 9           | 454           | 239           | 8.1        | 8               |
| 4  | Celbridge                                  | 10          | 8           | 454           | 239           | 8          | 4.5             |
| 5  | Lucan                                      | 6           | 8.5         | 399           | 199           | 7.9        | 6.5             |
| 6  | Irish National War Memorial Garden (INWMG) | 10          | 7.5         | 399           | 199           | 7.7        | 3               |
|    | <b>Standard Deviation for the Liffey</b>   | <b>2.50</b> | <b>0.63</b> | <b>164.98</b> | <b>83.08</b>  | <b>0.2</b> | <b>2.21</b>     |
| 7  | Canjáyar                                   | 17          | 16          | 776           | 404           | 7.2        | 12.45           |
| 8  | Terque                                     | 14          | 18          | 1000          | 500           | 7.5        | 10.5            |
| 9  | Puente de los Imposibles (PDLI)            | 12          | 20          | 1200          | 600           | 7.9        | 15              |
| 10 | Alhama de Almería                          | 7           | 22          | 1240          | 610           | 7.9        | 15.2            |
| 11 | Puente de Santa Fe de Mondújar             | 5           | 22.5        | 1400          | 666           | 8          | 168             |
|    | <b>Standard Deviation for the Andarax</b>  | <b>4.95</b> | <b>2.73</b> | <b>240.77</b> | <b>103.91</b> | <b>0.3</b> | <b>69.22</b>    |

Generally, the soil nutrient levels are higher in the Andarax catchment and may be attributed to agricultural runoff, lower water volumes, and higher evaporation rates that concentrate nutrients. The Liffey displays a lot lower nutrient concentrations, which is rather consistent with its temperate climate and higher water flows that tend to dilute pollutants more effectively.

Some key relationships and interdependencies driving the hydrological and chemical dynamics of the catchments were established using Pearson correlation. However, the DO has a significant negative correlation, with a pH of  $-0.813$  and a moderate negative correlation with a turbidity of  $-0.425$ . The temperature is highly correlated to EC with a value of  $0.972$  and to TDS with a value of  $0.971$ . Electrical conductivity (EC) and TDS are highly positively correlated to  $0.99$ . Furthermore, turbidity positively correlates with the EC of  $0.588$  and TDS of  $0.558$ . These findings, therefore, provided a background understanding of environmental and hydrological dynamics in the two contrasting catchments.

The ANN models considered the differences between the two catchments; for example, the impact of agricultural runoff in the Andarax River and urban wastewater in the River Liffey for proactive management to maintain water quality within acceptable limits. Therefore, the input data include the land-cover type and contribution between the measuring points. The model's performance generally showed strong and optimistic results, and the difference between the observed and modeled values was small (Figure 4).

The general trend of dissolved oxygen was captured by the model very well with an  $R^2$  of  $0.89$  and a reasonable RMSE of  $1.25$  mg/L. Uncommonly, the Andarax River's modeled values show better alignment of modeled DO values than the River Liffey catchment, with the  $R^2$  value higher at  $0.94$  vs.  $0.68$ , respectively. This minor disagreement at site levels was found due to some specific localized environmental conditions, such as vegetation-cover types, wastewater discharge, or aeration process differences, which were not fully captured by the model due to dataset limitations. These findings are consistent with previous

research that has focused on the need to include localized environmental variables to increase model precision (e.g., [12,13]).

The ANN model has excellent performance for simulating EC values, bringing about a very high  $R^2$  of 0.98 and a low RMSE of 48.06  $\mu\text{S}/\text{cm}$ . Similarly, the observed and modeled pH generally show minor deviations but stay within a narrow range of about 7.0 to 8.2. The  $R^2$  is high, at about 0.87, and the RMSE shows a small deviation, about 0.095. The pH variation is well captured by the model, but with minor discrepancies. These discrepancies may be because pH is dependent on a local geochemical or biological activity that the model does not well represent due to dataset limitations. The model performed well for the TDS, as indicated by an  $R^2$  of 0.95 and an RMSE of 44.84 ppm, capturing the main trend in variations in the TDS across sites. A high performance of this parameter reflects its stability even under a fluctuating environment [14,15]. The overall turbidity trend of the data was quite well captured, with a sharp increase in sediments at the last measuring site of the Andarax River, with an  $R^2$  of 0.96 and a RMSE of 8.90. The increase of urban land-cover types downstream increases nutrient loading and sediment transport [16].

Figure 5 shows, in detail, the results, underlining how the model is able to catch the general patterns of variation and the differences in water-quality parameters of both catchments while pinpointing aspects that need further refinement. The figure represents the results of the ANN model tested on the validation dataset, independent of the data used during training. This will ensure that the performance metrics are unbiased and representative of the model's predictive capability.

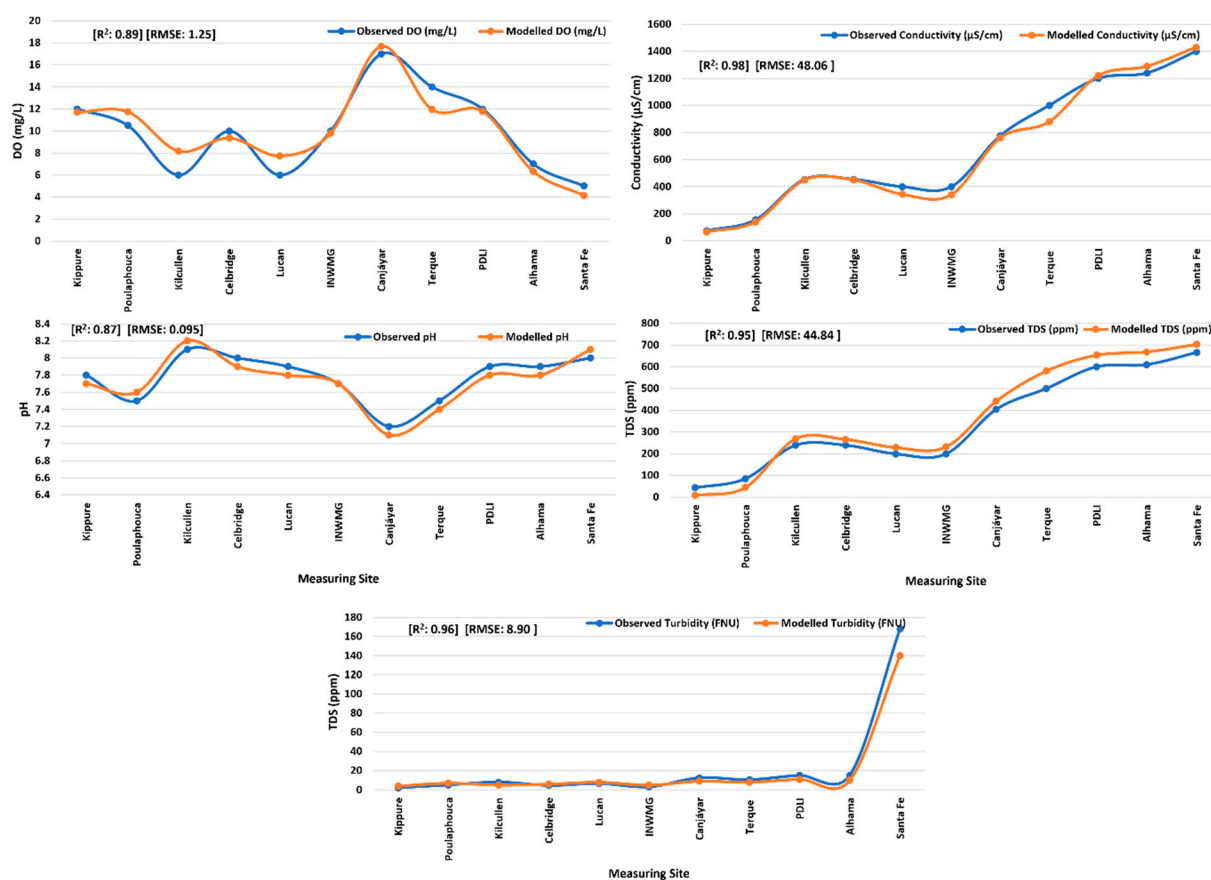


**Figure 5.** Observed versus modeled water-quality parameters at different measuring sites in the Liffey and Andarax catchments.

## 4. Discussion

The DO concentrations in the River Liffey catchment range between 6 mg/L and 12 mg/L, while modeled values range between 4.0 and 6.0% of the observed values, reflecting a relatively close fit. However, in some places, such as Kilcullen, the observed value was 6 mg/L but was modeled at 8.18 mg/L, which somewhat shows that the model is more sensitive at lower DO values. For the Andarax River, higher DO variability is observed with a good fit for the model in upstream measuring sites such as Canjáyar, while it tends to underestimate at points of low DO, such as at Santa Fe (downstream).

The River Liffey observed pH values fall within a range of near-neutral to slightly alkaline, measuring between 7.5 and 8.1, while the modeled values were around the observed and changed by approximately 0.1 to 0.2 units. However, the lower pH at the highest elevation in the Andarax River catchment could indicate a localized input of more acidic water from geological formations or rainwater–landscape interactions. The observed pH values for the Andarax River catchment are slightly underestimated.

The range in conductivity is from 74  $\mu\text{S}/\text{cm}$  at Kippure to 454  $\mu\text{S}/\text{cm}$  at Kilcullen and Celbridge. Modeled values are generally in good agreement with observed values, with most differences in the range of 15–20  $\mu\text{S}/\text{cm}$ , except for Lucan and Irish National War Memorial Garden (INWMG), which are marginally underestimated by the model (340–343  $\mu\text{S}/\text{cm}$ ). This would suggest that changes in conductivity within the range of 50–100  $\mu\text{S}/\text{cm}$  have little effect on the performance of this model for the catchment. Indeed, higher values of conductivity are recorded for the Andarax River catchment, especially in the case of Santa Fe (downstream), where observed conductivity is 1400  $\mu\text{S}/\text{cm}$  and modeled conductivity is 1430  $\mu\text{S}/\text{cm}$ . The model slightly overestimates conductivity for higher values. This could be a sign of sensitivity to higher ranges and might affect the prediction of extreme quantiles for high-conductivity events.

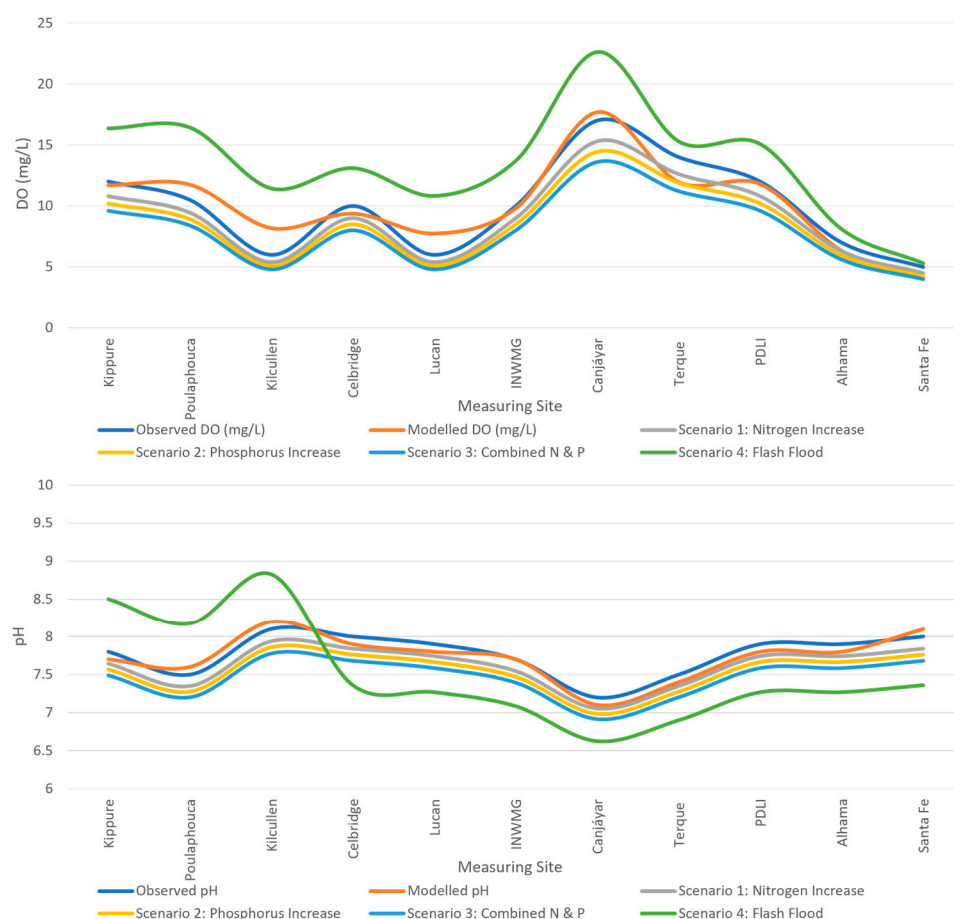
Generally, turbidity levels are low over most sites in the Liffey catchment, while the highest values were recorded for Lucan and the Irish National War Memorial Garden (INWMG). The turbidity is considerably raised for the Andarax River in most cases, especially at the downstream measuring site (Santa Fe) due to high sediment input and changes in the channel geometry during the measurement time. The model was able to capture the observed values through slight underestimations or overestimations at highly variable sites like Alhama and Santa Fe. Indeed, this points out that the model is sensitive to big shifts in turbidity but generally can predict the trends pretty well.

The TDS values range with moderate variation in the River Liffey catchment area, starting from 43 ppm at Kippure to about 239 ppm at both locations of Kilcullen and Celbridge. Much higher values for TDS characterize the Andarax River catchment, especially at Santa Fe, with observations of 666 ppm and modeled at 700 ppm. While the model reproduces the increasing trend from site to site, slight overestimations have been apparent at measuring sites with very high TDS, such as Alhama and Santa Fe. That reveals slight sensitivity regarding TDS in zones of extreme values; the model is still fitted to the observed data. These findings are consistent with previous research that has focused on the need to include localized environmental variables to increase model precision (e.g., [12,13]).

Once the model performance was high, four scenarios were generated to evaluate each pollution scenario on the river's water-quality parameters by the increase of agricultural or industrial runoff (Scenarios 1–3) or flash flood scenario (Scenario 4) (Table 3 and Figure 6).

**Table 3.** Percentage increase in each case over nitrogen and phosphorus concentrations.

| Scenario                                                               | Average Percentage Increase (%) | Influence                                                                                                                                                                                        |
|------------------------------------------------------------------------|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Scenario 1: Nitrogen Increase                                          | 223.5                           | <ul style="list-style-type: none"> <li>The increased amount of nitrogen leads to a 10% drop in DO (eutrophication would deplete DO),</li> <li>The pH slightly decreases.</li> </ul>              |
| Scenario 2: Phosphorus Increase                                        | 442.6                           | <ul style="list-style-type: none"> <li>A doubled phosphorus input leads to a DO decline by 15% and also to an acidification decrease in pH.</li> <li>Eutrophication would deplete DO.</li> </ul> |
| Scenario 3: Combined Nitrogen and Phosphorus                           | 223.5<br>442.6                  | <ul style="list-style-type: none"> <li>Combined nitrogen and phosphorus yield a 20% DO sag with a larger pH drop.</li> <li>Eutrophication would deplete DO.</li> </ul>                           |
| Scenario 4: Flash Flood (Leach High Levels of Nitrogen and Phosphorus) | 365.9<br>737.7                  | <ul style="list-style-type: none"> <li>Flash flooding produces a sharp increase in TDS and conductivity of 50% or more.</li> <li>Eutrophication would deplete DO.</li> </ul>                     |

**Figure 6.** Key water-quality parameters (observed, modeled, and scenarios); for example, dissolved oxygen (DO) and pH values under different water-quality scenarios for the River Liffey and Andarax catchments.

Land-cover-type composition in river catchments influences various water-quality parameters. The major land-cover types in the Liffey and Andarax catchments are summarized below in Table 4, with associated impacts on the various water-quality metrics

expected, as received from the existing literature and an attempt to build into the context on how land use bears directly and indirectly on water quality in these river systems.

Accordingly, these water-quality parameters reflect a combination of natural and anthropogenic influences. Parameters like electrical conductivity and pH, for instance, are largely controlled by natural conditions, specifically by the chemical composition of the riverbed and the input of groundwater. Nevertheless, human contributions remain very relevant in the lower courses of the rivers. High values of nitrogen and phosphorus were measured in the Andarax catchment, indicating agricultural runoff that would explain the eutrophication process, while urban wastewater was associated with localized reductions in dissolved oxygen within the Liffey catchment. Parameters provided priority in this study, therefore, include DO, TDS, turbidity, and EC, which are most relevant in the understanding of the relation between natural and artificial influences. While this, therefore, does not represent the full Water-Quality Index setup, it allows for critical drivers of changes in water quality in these two contrasting catchments. Further research will be done by adding more parameters like biochemical oxygen demand or nitrate concentrations to move closer to anthropogenic impacts on an aquatic ecosystem.

**Table 4.** Land-cover types for the two catchments and their impact on water-quality parameters.

| Land-Cover Type               | Expected Impact on Water Parameters                                                                                                                                                                                                                                                                                                        | Supporting Literature |
|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Forest and Semi-Natural Areas | <ul style="list-style-type: none"> <li>- Slightly lower pH due to organic matter breakdown</li> <li>- Increases DO from oxygenation due to tree roots</li> <li>- Decrease TDS from natural filtration</li> </ul>                                                                                                                           | [17–19]               |
| Agricultural Areas            | <ul style="list-style-type: none"> <li>- Leads to nutrient runoff, lowering the pH</li> <li>- Decreased DO from fertilizer runoff</li> <li>- Higher TDS from agricultural chemicals and sediments</li> </ul>                                                                                                                               | [20–22]               |
| Artificial Surfaces           | <ul style="list-style-type: none"> <li>- Tends to increase runoff, lowering pH</li> <li>- Decreased DO due to urban pollutants Increased TDS from contaminants and sediment</li> </ul>                                                                                                                                                     | [16,23,24]            |
| Wetlands                      | <ul style="list-style-type: none"> <li>- Stabilizes pH levels</li> <li>- Increases DO from vegetation</li> <li>- Lowers TDS due to natural filtration</li> </ul>                                                                                                                                                                           | [25–27]               |
| Peatlands                     | <ul style="list-style-type: none"> <li>- Lowers pH due to organic acids</li> <li>- Can lower DO if water is stagnant</li> <li>- May increase TDS from dissolved organic matter</li> </ul>                                                                                                                                                  | [28–30]               |
| Water Bodies                  | <ul style="list-style-type: none"> <li>- Neutral effect on pH</li> <li>- DO levels depend on water circulation</li> <li>- TDS levels vary based on upstream inputs</li> </ul>                                                                                                                                                              | [31,32]               |
| Natural Grasslands            | <ul style="list-style-type: none"> <li>- Slightly acidic pH due to soil organic matter</li> <li>- Moderate DO levels, with slight variations depending on surrounding vegetation</li> <li>- TDS levels are generally low but can increase with erosion and sediment transport</li> </ul>                                                   | [20–22]               |
| Bare Rocks or Soil            | <ul style="list-style-type: none"> <li>- High TDS potential from sediment and mineral dissolution, especially during rainfall and runoff events</li> <li>- Limited impact on DO, as sparse vegetation limits organic matter breakdown</li> <li>- Neutral to slightly alkaline pH due to minimal vegetation and organic material</li> </ul> | [20–22,33]            |

## 5. Conclusions

This present study illustrates the efficiency of Artificial Neural Network (ANN) models for assessing and simulating water-quality parameters across two hydrologically and

climatically diverse river catchments. The models were successfully able to capture the complex nonlinear relationships and extreme values of observed and future scenario figures with a high degree of accuracy. This is evidenced by the high  $R^2$  values (e.g.,  $R^2 = 0.98$  for EC), hence proving to be robust enough for both temperate and semi-arid settings.

Further analysis indicates how varied types of land cover, such as agricultural areas, artificial surfaces, and even forested zones, might impinge upon the water-quality parameters of each catchment. For example, the Andarax catchment gains a higher TDS and conductivity due to nutrient runoff and salinity issues. These results further stress that the land-cover data should be integrated within the water-quality models for better predictive capabilities and a more detailed understanding of the localized impacts. These simulation results will be very useful for the development of proactive management strategies to avoid further pollution and maintain sustainable levels of water quality. The ANN-based modeling presented in the paper may be applied to other river systems presenting different environmental challenges. It should, therefore, be a useful tool for a wide range of regional and global initiatives in water resource management. Future studies should try to incorporate the use of additional land use and climate change data to increase predictability and, thus, help the dynamic, long-term strategy of water management where climate variability and anthropogenic pressures increasingly threaten freshwater ecosystems.

However, the dataset size remains limited and, hence, it is a critical constraint with respect to the robustness and predictive power of the model. In this framework, ongoing activities are dedicated to enlarging the dataset through further monitoring campaigns and by integrating long-term historical data when available. In any case, due to generally limited funding, research often cannot afford extensive field measurement campaigns, which is a strong constraint for data acquisition in order to improve models.

**Funding:** This research received no external funding.

**Data Availability Statement:** Data used in this study can be provided on request.

**Acknowledgments:** The author would like to thank Anthea Cameron and Frances Leogue from Trinity College Dublin for providing the logistics necessary to complete field measurements in Ireland and Spain. Special thanks are extended to the Urrea Field Centre Team in Sorbas, Spain, for their great support and logistics during the fieldwork in Spain. The author would like to acknowledge the Environmental Protection Agency (EPA) Team in Ireland and the Instituto Geológico y Minero de España (IGME) Team for providing the respective essential data publicly, adding much value to the success of this research.

**Conflicts of Interest:** The author declares no conflicts of interest.

## References

1. Peñalver, R.; Jacobs, M.R.; Hegarty, S.; Regan, F. Assessment of anthropogenic pollution by monitoring occurrence and distribution of chemicals in the river Liffey in Dublin. *Environ. Sci. Pollut. Res.* **2021**, *28*, 53754–53766. [[CrossRef](#)] [[PubMed](#)]
2. Sanchez Martos, F.; Bosch, A.P.; Calaforra, J.M. Hydrogeochemical processes in an arid region of Europe (Almeria, SE Spain). *Appl. Geochem.* **1999**, *14*, 735–745. [[CrossRef](#)]
3. Zhang, J.; Zhong, Y.; Huai, W. Transverse distribution of streamwise velocity in open-channel flow with artificial emergent vegetation. *Ecol. Eng.* **2018**, *110*, 78–86. [[CrossRef](#)]
4. Isik, H.; Akkan, T. Water Quality Assessment with Artificial Neural Network Models: Performance Comparison Between SMN, MLP and PS-ANN Methodologies. *Arab. J. Sci. Eng.* **2024**, *50*, 369–387. [[CrossRef](#)]
5. Zhang, Y.; Gao, X.; Smith, K.; Inial, G.; Liu, S.; Conil, L.B.; Pan, B. Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network. *Water Res.* **2019**, *164*, 114888. [[CrossRef](#)]

6. Sarkar, A.; Pandey, P. River Water Quality Modelling Using Artificial Neural Network Technique. *Aquat. Procedia* **2015**, *4*, 1070–1077. [\[CrossRef\]](#)
7. Alqahtani, A.; Shah, M.I.; Aldrees, A.; Javed, M.F. Comparative Assessment of Individual and Ensemble Machine Learning Models for Efficient Analysis of River Water Quality. *Sustainability* **2022**, *14*, 1183. [\[CrossRef\]](#)
8. Niu, W.; Feng, Z. Evaluating the performances of several artificial intelligence methods in forecasting daily streamflow time series for sustainable water resources management. *Sustain. Cities Soc.* **2021**, *64*, 102562. [\[CrossRef\]](#)
9. Gad, M.; Saleh, A.H.; Hussein, H.; Elsayed, S.; Farouk, M. Water Quality Evaluation and Prediction Using Irrigation Indices, Artificial Neural Networks, and Partial Least Square Regression Models for the Nile River, Egypt. *Water* **2023**, *15*, 2244. [\[CrossRef\]](#)
10. Sánchez-García, C.; Schulte, L. Analysis of the Correlation Between Climate and Historical Extreme Flood Events in the Southeastern Iberian Peninsula: The Case of the Andarax River. In Proceedings of the 6th PAGES Open Sciences Meeting (OSM), Agadir, Morocco, 16–20 May 2022.
11. Environmental Protection Agency Water Quality Variations and Management Challenges in the River Liffey. 2013. Available online: <https://www.google.com.hk/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.epa.gov/&ved=2ahUKEwjutpjit66LAXes1YBHezwLcEQFnoECAGQAQ&usq=AOvVaw0FxrXzt-ArcktUfRBFLD5f> (accessed on 20 January, 2025).
12. Smith, D.I. Flood damage estimation-A review of urban stage-damage curves and loss functions. *Water* **1994**, *20*, 231–238.
13. Howarth, R.W.; Sharples, A.; Walker, D. Sources of nutrient pollution to coastal waters in the United States: Implications for achieving coastal water quality goals. *Estuaries* **2002**, *25*, 656–676. [\[CrossRef\]](#)
14. Covino, T. Hydrologic connectivity as a framework for understanding biogeochemical flux through watersheds and along fluvial networks. *Geomorphology* **2017**, *277*, 133–144. [\[CrossRef\]](#)
15. McGrane, S.J. Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: A review. *Hydrol. Sci. J.* **2016**, *61*, 2295–2311. [\[CrossRef\]](#)
16. Brezonik, P.L.; Stadelmann, T.H. Analysis and predictive models of stormwater runoff volumes, loads, and pollutant concentrations from watersheds in the Twin Cities metropolitan area, Minnesota, USA. *Water Res.* **2002**, *36*, 1743–1757. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Likens, G.E.; Bormann, F.H. Linkages between terrestrial and aquatic ecosystems. *Bioscience* **1974**, *24*, 447–456. [\[CrossRef\]](#)
18. Hornung, M.; Reynolds, B.; Stevens, P.A.; Hughes, S. Water Quality Changes from Input to Stream. In *Acid Waters in Wales*; Springer: Berlin/Heidelberg, Germany, 1990; pp. 223–240.
19. Neff, J.C.; Chapin, F.S., III; Vitousek, P.M. Breaks in the cycle: Dissolved organic nitrogen in terrestrial ecosystems. *Front. Ecol. Environ.* **2003**, *1*, 205–211. [\[CrossRef\]](#)
20. Carpenter, S.R.; Caraco, N.F.; Correll, D.L.; Howarth, R.W.; Sharples, A.N.; Smith, V.H. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Appl.* **1998**, *8*, 559–568. [\[CrossRef\]](#)
21. Cunha, D.G.F.; Bottino, F.; Calijuri, M.d.C. Land use influence on eutrophication-related water variables: Case study of tropical rivers with different degrees of anthropogenic interference. *Acta Limnol. Bras.* **2010**, *22*, 35–45. [\[CrossRef\]](#)
22. Lal, R. Soil erosion impact on agronomic productivity and environment quality. *CRC. Crit. Rev. Plant Sci.* **1998**, *17*, 319–464. [\[CrossRef\]](#)
23. Paul, M.J.; Meyer, J.L. Streams in the urban landscape. *Annu. Rev. Ecol. Syst.* **2001**, *32*, 333–365. [\[CrossRef\]](#)
24. Kaushal, S.S.; Groffman, P.M.; Likens, G.E.; Belt, K.T.; Stack, W.P.; Kelly, V.R.; Band, L.E.; Fisher, G.T. Increased salinization of fresh water in the northeastern United States. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13517–13520. [\[CrossRef\]](#)
25. Mitsch, W.J.; Gosselink, J.G. *Wetlands*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
26. An, S.; Verhoeven, J.T.A. *Wetland Functions and Ecosystem Services: Implications for Wetland Restoration and Wise Use*; Springer: Berlin/Heidelberg, Germany, 2019; ISBN 3030148602.
27. Kadlec, R.H.; Wallace, S. *Treatment Wetlands*; CRC Press: Boca Raton, FL, USA, 2008.
28. Armstrong, A.; Holden, J.; Kay, P.; Francis, B.; Foulger, M.; Gledhill, S.; McDonald, A.T.; Walker, A. The impact of peatland drain-blocking on dissolved organic carbon loss and discolouration of water; results from a national survey. *J. Hydrol.* **2010**, *381*, 112–120. [\[CrossRef\]](#)
29. Clymo, R.S. The limits to peat bog growth. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1984**, *303*, 605–654.
30. Prijac, A.; Gandois, L.; Taillardat, P.; Bourgault, M.-A.; Riahi, K.; Ponçot, A.; Tremblay, A.; Garneau, M. Hydrological connectivity controls dissolved organic carbon exports in a peatland-dominated boreal catchment stream. *Hydrol. Earth Syst. Sci.* **2023**, *27*, 3935–3955. [\[CrossRef\]](#)
31. Wetzel, R.G. *Limnology: Lake and River Ecosystems*; Academic Press: Cambridge, MA, USA, 2001; Volume 1.

32. Hem, J.D. *Study and Interpretation of the Chemical Characteristics of Natural Water*; Department of the Interior, US Geological Survey: Reston, VA, USA, 1985; Volume 2254.
33. Sridhar, A.; Thakur, B.; Basavaiah, N.; Seth, P.; Tiwari, P.; Chamyal, L.S. Lacustrine record of high magnitude flood events and climate variability during mid to late Holocene in the semiarid alluvial plains, western India. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2020**, *542*, 109581. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.