Article

Leveraging Machine Learning and Remote Sensing for Water Quality Analysis in Lake Ranco, Southern Chile

Lien Rodríguez-López [1,*], Lisandra Bravo Alvarez [2], Iongel Duran-Llacer [3,4], David E. Ruíz-Guirola [5], Samuel Montejo-Sánchez [6], Rebeca Martínez-Retureta [7,8], Ernesto López-Morales [1], Luc Bourrel [9], Frédéric Frappart [10] and Roberto Urrutia [11]

1 Facultad de Ingeniería, Arquitectura y Diseño, Universidad San Sebastián, Lientur 1457, Concepción 4030000, Chile; ernesto.lopez@uss.cl

2 Department of Electrical Engineering, Universidad de Concepción, Edmundo Larenas 219, Concepción 4030000, Chile; lisanbravo@udec.cl

3 Escuela de Ingeniería en Medio Ambiente y Sustentabilidad, Escuela de Ingeniería Forestal, Facultad de Ciencias, Ingeniería y Tecnología, Universidad Mayor, Camino La Pirámide 5750, Santiago 8580745, Chile; iongel.duran@umayor.cl

4 Hémera Centro de Observación de la Tierra, Facultad de Ciencias, Ingeniería y Tecnología, Universidad Mayor, Camino La Pirámide 5750, Santiago 8580745, Chile

5 Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland; david.ruizguirola@oulu.fi

6 Instituto Universitario de Investigación y Desarrollo Tecnológico, Universidad Tecnológica Metropolitana, Santiago 8940577, Chile; smontejo@utem.cl

7 Departamento de Ingeniería de Obras Civiles, Facultad de Ingeniería y Ciencias, Universidad de La Frontera, Francisco Salazar 1145, Temuco 4811186, Chile; rebeca.martinez@ufrontera.cl

8 Departamento de Ciencias Ambientales, Facultad de Recursos Naturales, Universidad Católica de Temuco, Rudecindo Ortega 02950, Temuco 4780000, Chile

9 Géosciences Environnement Toulouse, UMR 5563, Université de Toulouse, CNRS-IRD-OMP-CNES, 31400 Toulouse, France; luc.bourrel@ird.fr

10 ISPA, UMR 1391 INRAE, Bordeaux Sciences Agro, UMR 1391, 33140 Villenave-d'Ornon, France; frederic.frappart@inrae.fr

11 Facultad de Ciencias Ambientales, Universidad de Concepción, Concepción 4030000, Chile; rurrutia@udec.cl

* Correspondence: lien.rodriguez@uss.cl

Abstract: This study examines the dynamics of limnological parameters of a South American lake located in southern Chile with the objective of predicting chlorophyll-a levels, which are a key indicator of algal biomass and water quality, by integrating combined remote sensing and machine learning techniques. Employing four advanced machine learning models (recurrent neural network (RNNs), long short-term memory (LSTM), recurrent gate unit (GRU), and temporal convolutional network (TCNs)), the research focuses on the estimation of chlorophyll-a concentrations at three sampling stations within Lake Ranco. The data span from 1987 to 2020 and are used in three different cases: using only in situ data (Case 1), using in situ and meteorological data (Case 2), using in situ, and meteorological and satellite data from Landsat and Sentinel missions (Case 3). In all cases, each machine learning model shows robust performance, with promising results in predicting chlorophyll-a concentrations. Among these models, LSTM stands out as the most effective, with the best metrics in the estimation, the best performance was Case 1, with $R^2 = 0.89$, an RSME of 0.32 µg/L, an MAE 1.25 µg/L and an MSE 0.25 $(µg/L)^2$, consistently outperforming the others according to the static metrics used for validation. This finding underscores the effectiveness of LSTM in capturing the complex temporal relationships inherent in the dataset. However, increasing the dataset in Case 3 shows a better performance of TCNs ($R^2 = 0.96$; MSE = 0.33 $(µg/L)^2$; RMSE = 0.13 µg/L; and MAE = 0.06 µg/L). The successful application of machine learning algorithms emphasizes their potential to elucidate the dynamics of algal biomass in Lake Ranco, located in the southern region of Chile. These results not only contribute to a deeper understanding of the lake ecosystem but also highlight the utility of advanced computational techniques in environmental research and management.

Keywords: water quality; lake; remote sensing; machine learning; southern Chile

1. Introduction

In a climatic scenario of water deficit, such as the one we have been experiencing in recent decades, freshwater bodies become essential resources to meet human, agricultural, and industrial needs [1,2]. These bodies of water are not only vital for direct human consumption but also for food production and hydropower generation [3,4]. They also play a crucial role in the conservation of aquatic biodiversity and in the regulation of local and regional climates [5,6]. Therefore, the sustainable management and conservation of freshwater resources become imperative in a context where climate change and human activity are putting increasing pressure on these vital ecosystems [7].

Lakes provide a wide range of ecosystem services vital to human well-being and environmental health [8]. Lakes play a critical role in regulating local climate by influencing precipitation and temperature patterns in surrounding regions [1,9,10]. These aquatic ecosystems are also critical habitats for diverse plant and animal species, providing shelter and food and enabling reproduction for numerous life forms [11,12]. In addition, lakes can have important recreational and aesthetic value, attracting tourists and providing opportunities for leisure activities, such as fishing, boating, and bird watching [13]. Therefore, lakes are key components of natural landscapes, providing essential services that support human life and ecological balance [14].

The use of satellite technology to monitor water quality in aquatic bodies has been increasing worldwide, including in Chile, in response to the growing concern for the health of our water resources [15–17]. Advanced satellites, such as Landsat, Sentinel, and other systems, provide increasingly wide spatial and temporal coverage, allowing for the obtainment of detailed and updated information on the characteristics of water bodies [18,19]. These images, generated by sophisticated sensors onboard satellites, allow for the identification of changes in water quality parameters, such as Chl-a in [20,21], the distribution of pollutants studied by [22,23], and other important indicators for the management and conservation of aquatic ecosystems [24–26]. This innovative approach not only provides a global view of the health of water bodies but also facilitates early detection of problems and informed decision making for their protection and sustainable management [27].

In addition, the combination of traditional water sampling techniques with machine learning algorithms has proven to be complementary, thus improving monitoring processes and decision making on the state of water quality in lakes [20,21,28,29]. Therefore, the objective of this study is to estimate chlorophyll-a levels in Lake Ranco by combining machine learning and remote sensing techniques. For this purpose, the behavior of physicochemical, biological, and meteorological variables, as well as relevant satellite bands, will be described. Four machine learning algorithms will be used to estimate chlorophyll-a levels in the period from 1987 to 2020, validating their effectiveness. In addition, the ecosystem services associated with Lake Ranco will be analyzed, providing a more complete understanding of their importance and functioning in the local and regional ecosystem.

2. Materials and Methods

2.1. Lake Ranco: Study Area

Lake Ranco, located in the extreme south of Chile, in the Los Ríos Region, at coordinates 40°22′57″S–72°30′41″W, has an area of approximately 442 km^2, which places it among the largest lakes in Chile (see Figure 1). The local climate presents traits of a humid temperate climate with Mediterranean influences [30]. The average annual precipitation is around 2000 mm, while the temperature ranges between 6 and 9 °C, with maximum temperatures in January (20 °C) and minimum temperatures in July (2 °C), according to data from the Chilean Meteorological Directorate (DMC, http://www.meteochile.cl/, accessed on 1 January 2024). Nestled in a mountainous terrain, flanked by the Cordillera de la Costa and the Andes Mountains, the picturesque setting of Lake Ranco owes much to its natural environment. Characterized by numerous islands and peninsulas, the lake's landscape adds to its charm, offering unique habitats for local flora and fauna, thus increasing the region's biodiversity [29].
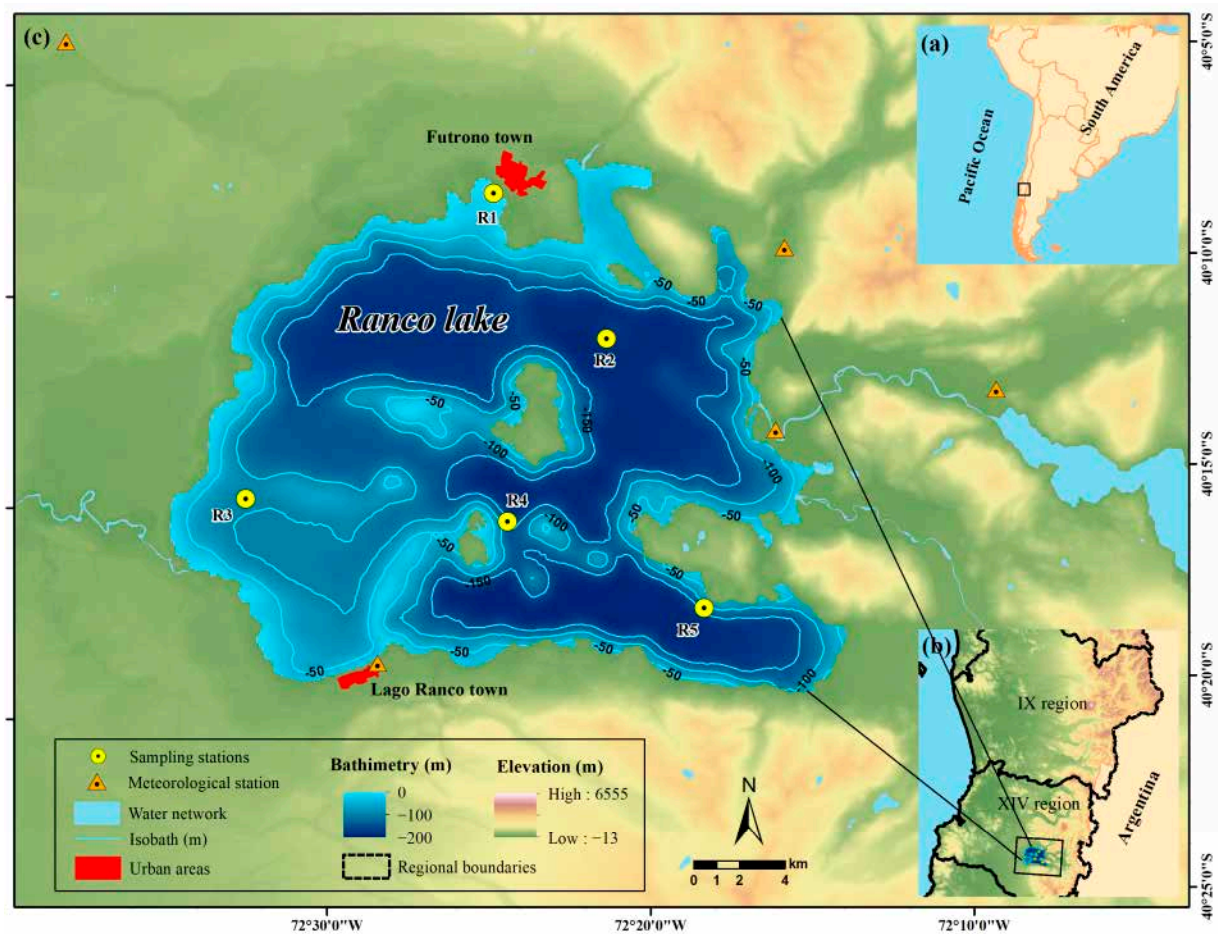
Figure 1. (**a**) Location of Chile in South America; (**b**) XIV Región de los Ríos; (**c**) Lake Ranco. Monitoring stations shown in yellow circles; meteorological stations shown in orange triangles and bathymetry.

2.2. Analysis of Ecosystem Services

An analysis of the ecosystem services provided by Lake Ranco was conducted within a 10 km buffer zone using preliminary information gathered from various studies, with a primary focus on biodiversity, natural elements, and anthropogenic factors.

To characterize the area, land use coverage data from the year 2014 were utilized, obtained from a raster file with a spatial resolution of 30 m, sourced from www.gep.uchile. cl/Landcover_CHILE.html, accessed on 10 January 2024. Additionally, to assess the level of human influence, data from the Ministry of Public Works of Chile were consulted for information regarding the road network (https://www.mop.gob.cl/, accessed on 12 January 2024), while data on surrounding populated areas within the study zone were obtained from the National Statistics Institute of Chile (https://www.ine.gob.cl/herramientas/portal-de-mapas/geodatos-abiertos, accessed on 20 February 2024).

2.2.1. Ecosystem Services of Lake Ranco

Lake Ranco plays a crucial role in providing freshwater to the region, serving as a vital source for domestic and agricultural supply (Figure 2). Additionally, the lake acts as a regulator of the local climate, moderating temperatures, creating favorable microclimates for agriculture, and helping mitigate the risk of floods by absorbing and retaining excess water during periods of heavy rainfall.
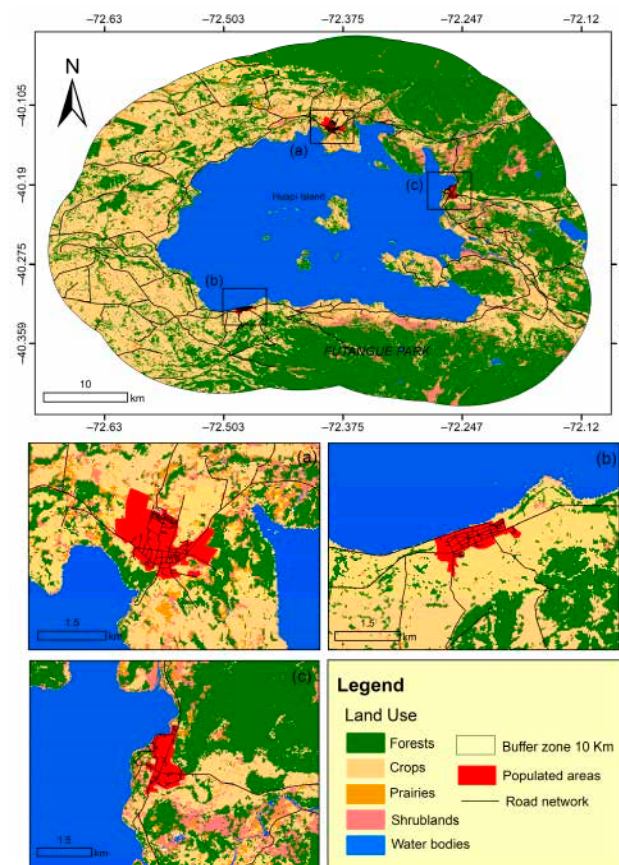
Figure 2. Land use, road network, and populated areas within a 10 km buffer zone around Lake Ranco. (**a**) Town of Futrono; (**b**) town of Lake Ranco; and (**c**) town of Llifén.

However, this environment faces considerable challenges due to the intensification of agricultural activity in the surrounding region. Despite Lake Ranco being naturally oligotrophic, agricultural activity introduces additional nutrients, such as phosphorus, nitrogen, and organic matter, which can result in lake eutrophication and threaten its ecosystem health [31] (Figure 2).

Additionally, pollution from the lake's tributaries, especially due to fish farming activity, adds another layer of concern about the overall health of the ecosystem. These concerns underscore the urgent need to comprehensively address the environmental impacts of human activities in the region to preserve the biodiversity and crucial ecosystem services offered by this lacustrine ecosystem [31].

In terms of biodiversity, Lake Ranco harbors various endemic and endangered species, highlighting its importance for regional conservation.

2.2.2. Human–Nature Interaction

Recreational tourism around the lake is a significant economic activity, generating income and employment for local communities. Cultural values rooted around the lake have been identified, which are fundamental to local identity and community cohesion [32,33].

To the southeast of Lake Ranco lies Parque Futangue, a privately protected area covering a vast expanse of land from the Andes Mountains to the vicinity of the Lake Ranco shore, offering a variety of outdoor activities and accommodations [34] (Figure 2).

The landscape on the edge of Lake Ranco is distinguished by a unique interaction between natural and human elements, with two main urban areas standing out: the towns of Ranco and Futrono, and a rural area, the locality of Llifén, along with extensive agricultural lands (Figure 2a–c). In the center of the lake lies Huapi Island, whose inhabitants depend

entirely on the resources provided by the lake for their sustenance. Although the lake is vital for these communities, its development does not entirely define the landscape.

Regarding accessibility, although pathways have been established along the entire lakeshore (Figure 2), there are issues associated with the appropriation of certain lakeside sites by private landowners, mainly second homes occupying beach areas [35].

The shoreline of Lake Ranco is primarily characterized by a natural–anthropogenic-rural interface, consisting mostly of agricultural use zones. These areas have intervened in the natural state of the shoreline, leading to an increasing threat to the biodiversity and ecosystem services that this ecosystem provides [35].

Although urban areas do not encompass the entirety of the shoreline surface, there are two significant urban centers that have played a decisive role in the ecological deterioration of the area.

2.3. In Situ and Meteorological Data

In this study, a time series spanning 33 years (1987–2020) was analyzed, incorporating data gathered from multiple in situ monitoring campaigns, sourced from the Dirección General de Aguas (DGA) database (accessed on 20 December 2023). Three sampling stations were strategically selected, as depicted in Figure 1. The selection criteria for these stations aimed to encompass various zones within the lake, ensuring comprehensive coverage across its expanse (see Figure 1). Additionally, five meteorological stations with monthly precipitation information near the study area were considered, with information from the Dirección Meteorológica de Chile DMC (accessed on January 2024).

2.4. Pre-Processing Satellite Images

From 2014 to 2022, during different seasons, satellite images from Landsat 8 (L8) and Sentinel 2 (S2) were obtained with a low percentage of cloud cover. A total of 10 images (3-L8, 7-S2) were downloaded to cover the dates of in situ sampling. The details of each image are detailed in Table 1. L8 (collection 2, level 1) was obtained from the United States Geological Survey (USGS) at Earth-Explorer https://earthexplorer.usgs.gov/ (accessed on 20 February 2024). S2 (Level-1C) was obtained from https://dataspace.copernicus.eu/ (accessed on 20 February 2024) of the European Commission, European Space Agency (ESA), and the Copernicus program.

Table 1. Satellite image characteristics.

Id	Path/Row	In Situ-Date	Year	Image Date	Day Differences
LC08_L1TP_233088_20140226_20200911_02_T1	233/88	18 February 2014	2014	26 February 2014	8
LC08_L1TP_232088_20150206_20200909_02_T1	232/88	11 February 2015	2015	6 February 2015	5
LC08_L1TP_233088_20151011_20200908_02_T1	233/88	6 October 2015	2015	11 October 2015	5
S2A_MSIL1C_20160122T142942_N0201_R053_T18GYA_20160122T144141	T18GYA	19 January 2016	2016	22 January 2016	3
S2B_MSIL1C_20200126T142649_N0208_R053_T18GYA_20200126T174938	T18GYA	27 January 2020	2020	26 January 2020	1
S2A_MSIL1C_20201126T142731_N0209_R053_T18GYA_20201126T180345	T18GYA	23 November 2020	2020	26 November 2020	3
S2A_MSIL1C_20210306T142731_N0209_R053_T18GYA_20210306T180339	T18GYA	4/5 March 2021	2021	6 March 2021	2, 1
S2B_MSIL1C_20210811T143729_N0500_R096_T18GYA_20230211T201839	T18GYA	4/5 August 2021	2021	8 August 2021	4, 3
S2B_MSIL1C_20211116T142729_N0301_R053_T18GYA_20211116T175548	T18GYA	7/8 November 2021	2021	9 November 2021	2, 1
S2A_MSIL1C_20220311T142741_N0400_R053_T18GYA_20220311T175316	T18GYA	9 March 2022	2022	11 March 2022	2

Five bands of the multispectral Operational Land Imager (OLI) sensor onboard L8 were used from blue (B) to shortwave infrared (SWIR), with a spatial resolution of 30 m. The same bands of the MultiSpectral Instrument (MSI) onboard S2 were used. MSI, onboard each of the two identical satellites (A/B) of the Sentinel-2 mission, has a spatial resolution of 10 m (Table 2). Subsequently, all images were atmospherically corrected in ACOLITE software (version 20231023.0, https://github.com/acolite, accessed on 20 February 2024). ACOLITE is a generic processor that was specifically developed for marine, coastal, and inland water and supports the free processing of both L8 and S2 [28,36]. The default atmospheric correction using the "Dark Spectrum Fitting" (DSF algorithm) and the older

"exponential extrapolation" or EXP approach were selected and used in the ACOLITE processor [37,38]. The ACOLITE products obtained correspond to surface-level reflectance (Rrs) for L8 and S2. The corrected bands were carried to the same spatial resolution, and different vegetation indices and combinations of bands were calculated from the same bands (Table 2).

The Quality Assessment bands of each satellite were used to mask clouds, cirrus, and shadows, and only cloud-free coverage data were used for data extraction [17]. The lake limits (Roi) were established using geospatial information from the DGA and only the water body was considered for analysis. QGIS software (version 3.32.0) was used to extract the data. The methodology used for the extraction was a 3 × 3-pixel sampling window. This step is important in terms of reducing the possible errors during the geometric correction of images and the dynamics of water bodies [28,39]. The bands, band combinations, and indices calculated are widely used to detect aquatic vegetation in lake systems [28,40]; these are shown in Table 2. For more details about each index, the cited references in the table can be consulted.

Table 2. Bands and indices characteristics.

Bandas/Indices Names	Sensor	Resolution	Equation	References
Blue	L8-OLI/S2-A/B	30 m/10 m	B (/490 nm)	[41,42]
Green	L8-OLI/S2-A/B	30 m/10 m	G (/560 nm)	[43,44]
Red	L8-OLI/S2-A/B	30 m/10 m	R (/665 nm)	[17,45]
Near Infrared	L8-OLI/S2-A/B	30 m/10 m	NIR (/842 nm)	[17,43]
Shortwave Infrared	L8-OLI/S2-A/B	30 m/20 m	SWIR	[45,46]
Red/Infrared	L8-OLI/S2-A/B	30 m/10 m	R/NIR	[17,44]
Infrared/Red	L8-OLI/S2-A/B	30 m/10 m	NIR/R	[17,47]
Normalized Difference Vegetation Index	L8-OLI/S2-A/B	30 m/10 m	$NDVI = (NIR - R)/(NIR + R)$	[48,49]
Floating Algae Index	L8-OLI/S2-A/B	30 m/10–20 m	$FAI = NIR - NIR$ with $NIR = R + (SWIR - R) \times (\lambda_{NIR} - \lambda_R)/(\lambda_{SWIR} - \lambda_R)$	[15,50]
Surface Algal Bloom Index	L8-OLI/S2-A/B	30 m/10 m	$SABI = (NIR - R)/(B + G)$	[51,52]
Green Normalized Difference Vegetation Index	L8-OLI/S2-A/B	30 m/10 m	$GNDVI = (NIR - G)/(NIR + G)$	[15,42]
Chlorophyll Index—Green	L8-OLI/S2-A/B	30 m/10 m	$CI\text{-}G = (NIR/G) - 1$	[53,54]
Green Blue (Two-Band) Ratio	L8-OLI/S2-A/B	30 m/10 m	$CHL\_OC2 = B/G$	[55,56]

2.5. Machine Learning Algorithms

2.5.1. Recurrent Neural Networks (RNNs)

RNNs represent a crucial class of neural networks adept at capturing long-term dependencies within sequential data. By processing one element of the sequence at a time and integrating it with the previous temporal information, RNNs discern underlying patterns. However, RNNs face the challenge of gradient fading or exploding during training, mainly due to recurrent operations. To mitigate this obstacle, several improved variants of RNNs have emerged, most notably long short-term memory (LSTM) and gate recurrent units (GRUs) [57]. In Figure 3, we illustrate a typical RNN architecture with N inputs, two layers of hidden neurons, and one output. In this context, the inputs encompass the current data from Cases 1–3 along with historical Chl-a data. The hidden layers serve to discover correlations between the input data, while the output layer provides the predicted Chl-a level, which is continuously refined with each iteration.
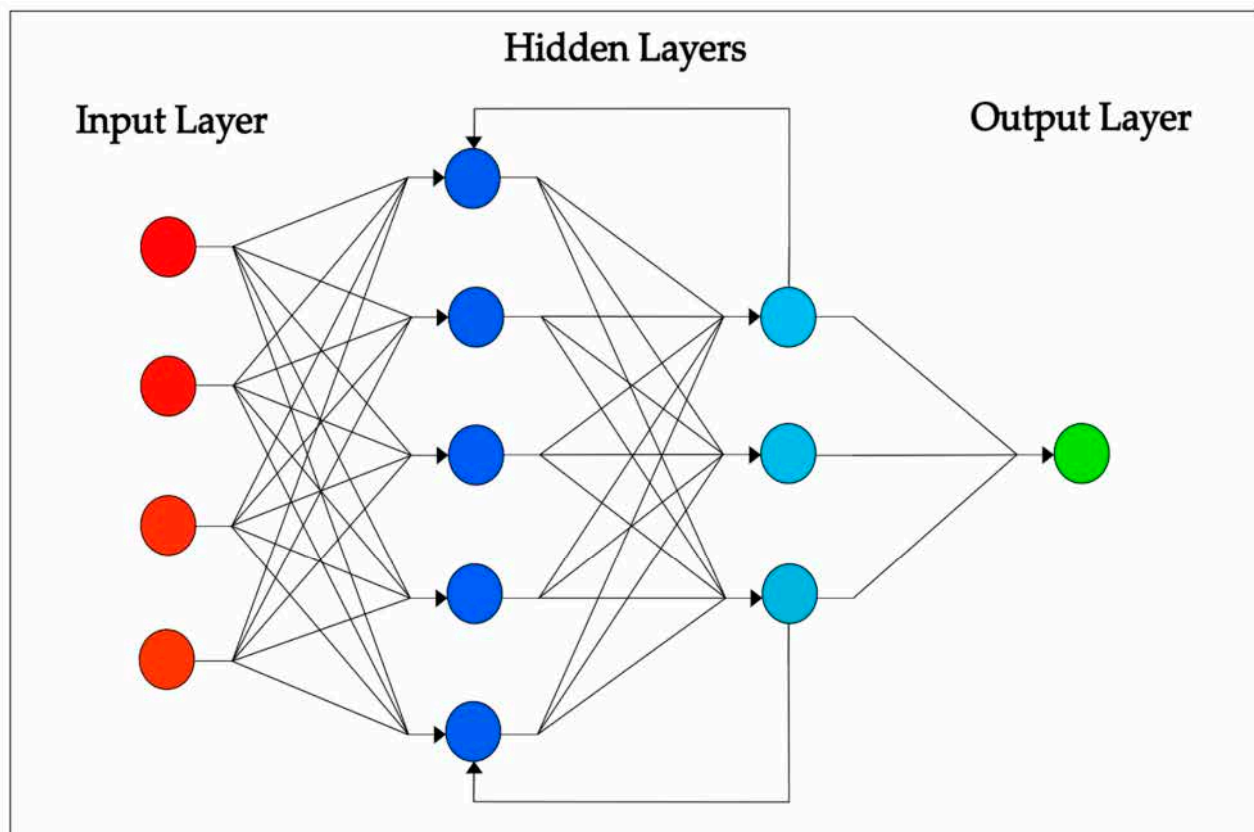
Figure 3. Recurrent neural network architecture. Adapted from [58].

2.5.2. Long Short-Term Memory (LSTM)

Long short-term memory models (LSTMs) represent a specialized type of RNN designed to effectively manage the complexities of long-term dependencies in sequential data. Conventional RNNs often have difficulty maintaining relevant information over extended periods of time, a phenomenon known as leaky gradients [59]. This occurs when the gradients in the backpropagation algorithm decrease significantly as they traverse numerous time steps, impeding the network's ability to adjust its weights effectively. To combat this problem, LSTM models incorporate memory cells and gating mechanisms, allowing the model to retain or discard information as needed over time [60].

Unlike conventional recurrent neural networks, which often struggle to hold information for long periods of time, LSTM networks are specifically designed to excel in this regard. They take advantage of a gating mechanism that allows them to precisely control the retention and discarding of information at each time step. In an LSTM network, each neural unit incorporates three fundamental gates: the input gate, the forgetting gate, and the output gate. These gates allow the network to discern which information is worth retaining and that which should be discarded at each time step. The input gate regulates the incorporation of new information into the network's short-term memory. The forgetting gate regulates the retention or deletion of information in the long-term memory. Finally, the output gate determines which information is used to make predictions or decisions at any given time (see Figure 4).
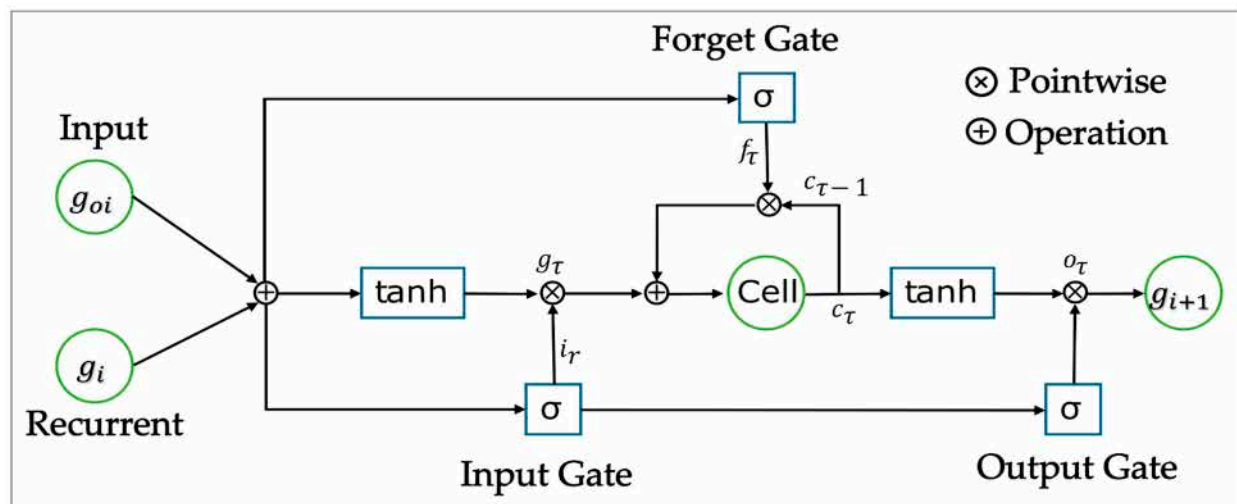
Figure 4. Long short-term memory architecture, adapted from [61].

2.5.3. GRU

A GRU presents an alternative to the LSTM in the field of recurrent neural networks, simplifying the control mechanism to only two crucial signals: the update gate and the reset gate. Unlike LSTM, a GRU simplifies the architecture by reducing the number of gating signals, thus saving computational resources. This reduction is achieved by merging the forget and input gates into a single update gate and eliminating the output gate. Consequently, a GRU consists of only two gates: the update gate and the reset gate. Figure 5 illustrates the architecture of a GRU, showing its simplified design. Although a GRU shares similarities with LSTM in its ability to capture temporal dependencies, it is distinguished by its parsimonious use of gate signals and associated parameters. This efficiency not only improves computational performance but also streamlines the training process, as demonstrated by [62].
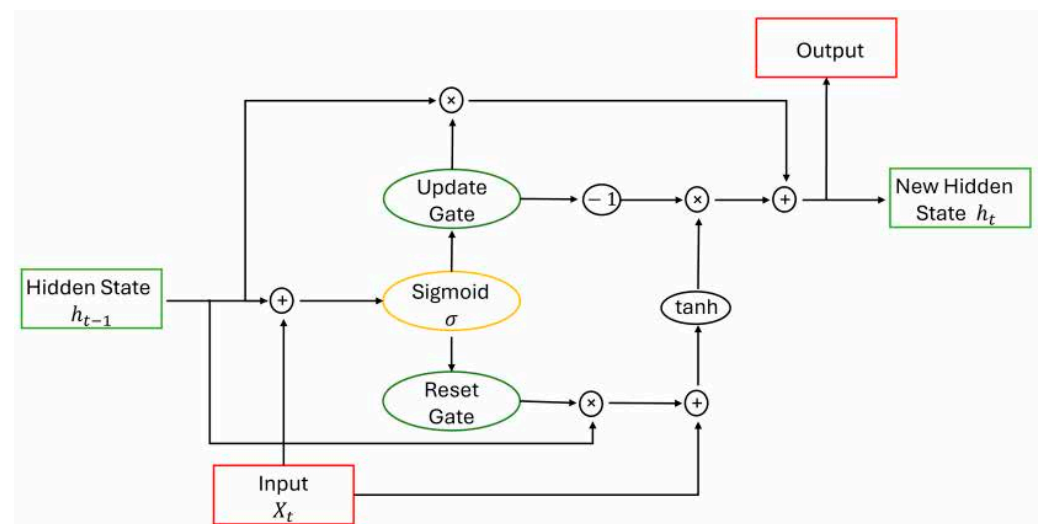
Figure 5. Gated recurrent unit architecture adapted from [63].

2.5.4. Temporal Convolutional Network (TCN)

A TCN represents a neural network architecture that is distinguished by the use of a causal convolution layer, which is strategically implemented to avoid leakage of future information into the past. This design feature ensures that predictions at any time step depend exclusively on the inputs at that time instant or earlier [57]. However, although causal convolution effectively maintains the temporal order, its dilation mechanism can sometimes

prevent the extraction of local connections between adjacent time steps, particularly in higher layers. In essence, a TCN generates its output through a temporal convolution of the input data. In particular, Figure 6 shows a TCN configuration comprising three dilated convolution layers, each of which contributes to the network's ability to capture the temporal dependencies across various time scales.

Causal convolution guarantees that the output at a given time depends only on the inputs up to that time, from the beginning of the sequence. For a sequence input x_0, x_1, ..., x_N, this implies that the output at time t is influenced only by x_t, x_{t-1}, ..., x_{t-r+1} for a given r. In contrast, dilated causal convolution allows one to obtain a large receptive field with minimal layer depth. This is achieved by summing the product of a filter $f(i)$ and the input sequence $x_{t-\lambda i}$ for i between 0 and $z-1$, where the dilation factor λ grows exponentially between layers. However, in higher layers, the dilated causal convolution may not fully capture the local connection between adjacent time steps [64]. Comparatively, temporal convolutional networks (TCNs) suffer from a drawback in inference due to their large memory requirement, as the complete sequence must be computed at the subsequent time step, unlike recurrent neural networks (RNNs).
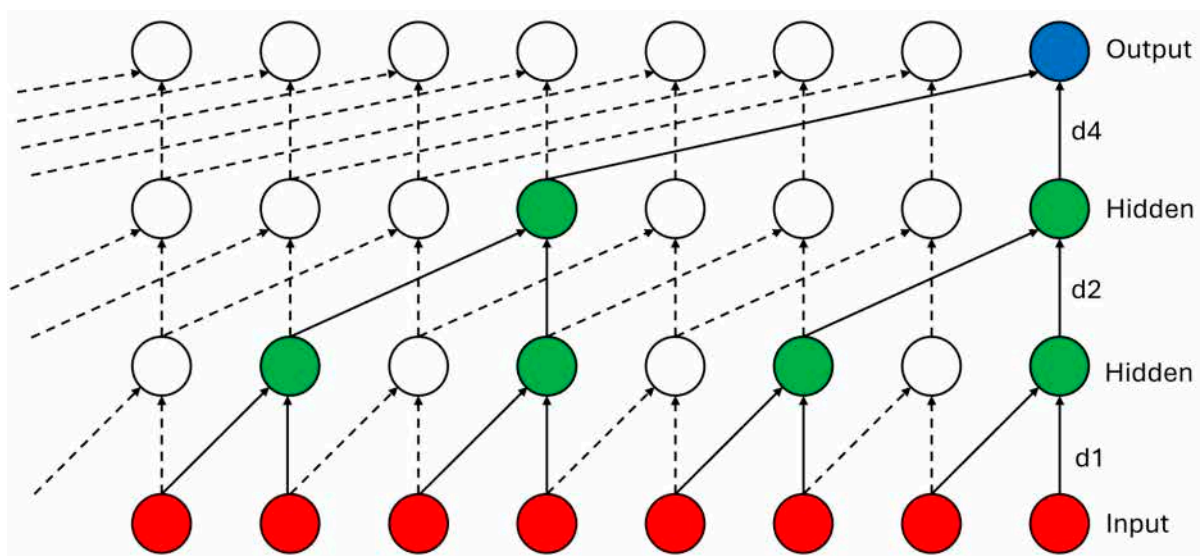
Figure 6. TCN architecture adapted from [65].

2.6. Algorithm Processsing

Recurrent neural network (RNN), long short-term memory (LSTM), and gate recurrent unit (GRU) architectures were constructed, with a hidden layer in each of them. The number of neurons (units) ranged from 5 to 250, employing an initial learning rate of 0.0001, a root mean square error (RMSE) loss function, a maximum of 100 epochs, and the Adam optimizer [66].

In contrast, a temporal convolutional network (TCN) architecture features three convolutional layers, incorporating dropout regularization, with a dropout factor of 0.005. The optimization process adjusts the number of neurons, ranging from 8 to 256 in the layer that connects to the same region in the input. In addition, it uses a filter size of 4, an initial learning rate of 0.0001, and the Adam optimizer. A TCN excels in handling long-term dependencies within input sequences, facilitated by the causal convolution layer, which guarantees a larger receptive field compared to traditional convolutional neural networks (CNNs). However, the inference process of a TCN occupies more memory than that of RNNs, as it requires processing the entire sequence to compute the next time step.

These architectures find application in a variety of tasks, such as predictive modeling, anomaly detection or classification. The selection of an appropriate architecture depends on the specific requirements of the problem, the characteristics of the available data and the

available computing resources. Through a nuanced understanding of the strengths and limitations of different architectures, we can adapt neural networks more effectively and efficiently to various applications.

To evaluate the performance of the four models mentioned above, we define cases 1, 2 and 3 as follows:

The Case 1 in situ data are as follows: The parameters included in this case are only in situ data (1—water temperature (T); 2—transparency (SD); 3—dissolved oxygen (DO); 4—pH; 5—chlorophyll (Chl-a); 6—total nitrogen (TN); 7—total phosphorus (TP); 8—turbidity (NTU); 9—conductivity).

The Case 2 in situ data and meteorological data are as follows: The parameters included in Case 2 are in situ data (Case 1) and meteorological data (1—air temperature; 2—relative humidity; 3—precipitation; 4—wind speed).

The Case 3 in situ data, meteorological data, and satellite data are as follows. The parameters included in Case 3 are the in situ data, meteorological data, (Case 2) and satellite data: 1—blue band; 2—green band; 3—red band; 4—infrared band; 5—near-infrared band; 6—R/N; 7—NDVI; 8—FAI; 9—SABI; 10—GNDVI; 11—GCI)

2.7. Statistics Validation

To evaluate the performance of the models described in Section 2.5, several metrics were employed, including the mean squared error (MSE) described in [17], the root mean squared error (RMSE) described in [15], the mean absolute error (MAE), described in [40], and R^2 described in [67]. This approach mirrors the methodology used in [29] and helps to understand the accuracy, precision, and potential limitations associated with chlorophyll-a estimation.

The number of samples for training and testing is given in Table S1. Sequential splitting, following a 70/30% rule, was adopted to calculate the different error metrics. In this procedure, the time series are split in such a way that the first part is used for training, while the latter is used for multi-station validation (Figure 1). This method has demonstrated commendable performance in the evaluation of time series competence in deep learning models.

3. Results

3.1. Water Quality Parameters

In Table 3, we can observe the statistical metrics of the limnological parameters used in this study. Transparency measured using a Secchi disk showed values ranging between 5.5 m at station R4 during fall, probably due to mixing, and 23 m at station R2 during summer, attributed to more favorable weather conditions: low wind, low rainfall, and low cloud cover. Water temperature values ranged from a minimum of 9.2 °C in the winter season to a maximum of 21.5 °C in summer. It should be noted that similar water temperature values in other lakes of the Araucano chain, such as Lake Villarrica, have triggered algal bloom events [15,30]. The water column of Lake Ranco is hydrodynamically mixed most of the year (autumn–winter–spring) and stratified during summer according to reports from the Chilean Water Authority (Direccion General de Aguas de Chile). In addition, thermal stratification in the summer period (summer) in Lake Ranco would lead to hydrodynamic isolation and an eventual decrease in oxygen concentration, which produces the desertification of the system by the abandonment of N2 (g) to the atmosphere, thus generating the limitation of nitrogen for biological assemblages during the summer. Chlorophyll-a reaches the highest values during the summer, up to 3.2 µg/L, coinciding with the highest values of water temperature, which is optimal for the aquatic vegetation to perform the photosynthetic process. Nitrogen and total phosphorus values are slightly higher than those reported in other Araucanian lakes, with maximum values of 310 mg/L and 13 mg/L, 11 mg/L and 12 mg/L in R2, R3 and R4 station, respectively; however, this lake still shows an oligotrophic condition.

Table 3. Water quality statistical parameters.

		Summer			Autumn			Winter			Spring		
		R2	R3	R4	R2	R3	R4	R2	R3	R4	R2	R3	R4
SD (m)	Avg	15.7	15.8	15.0	13.5	16.4	13.1	12.3	12.3	12.7	13.6	13.6	13.4
	Max	23	21.5	19.5	18.5	20.5	18	19	17.7	17.5	18.6	18	18.5
	Min	10.5	10	8.5	10	11	5.5	7	8	8	8.5	8	9
	SD	3.9	3.9	3.58	3.0	2.9	3.2	2.9	2.7	2.8	3.1	2.4	2.9
	CV (%)	0.2	0.2	0.23	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.17	0.2
	N	14	14	13	13	13	12	16	15	15	15	15	15
T (°C)	Avg	19.2	19.0	19.4	14.1	14.1	13.9	10.1	10.7	10.7	13.5	13.0	13.7
	Max	21.5	20.2	20.6	16.6	17	16.7	15.7	15.4	15.2	16.2	16.2	16.5
	Min	17.8	17.8	17.9	12.9	12.3	11.7	9.2	9.2	10.0	11.5	10.9	10.4
	SD	1.0	0.8	0.8	1.3	1.3	1.4	1.4	1.3	1.2	1.2	1.6	1.8
	CV (%)	0.5	0.3	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.11	0.1
	N	15	14	14	15	15	15	18	17	17	16	16	16
pH (unit)	Avg	7.7	7.6	7.7	7.4	7.4	7.4	7.5	7.4	7.4	7.5	7.6	7.5
	Max	8.1	8.1	8.1	7.73	7.8	7.8	7.9	7.8	7.8	8.3	8.3	8.3
	Min	6.9	6.8	6.9	6.7	6.7	6.2	6.8	6.9	6.9	6.2	7.1	6.7
	SD	0.4	0.4	0.4	0.4	0.3	0.4	0.3	27	0.3	0.5	0.3	0.4
	CV (%)	0.1	0.1	0.1	0.4	0.4	0.1	0.3	0.0	0.3	0.0	0.4	0.5
	N	15	14	14	14	15	14	18	17	17	16	16	16
DO (mg/L)	Avg	9.1	9.1	8.9	9.4	9.6	9.8	10.6	10.7	10.5	10.2	10.2	10.2
	Max	10.7	14	9.7	10.3	10.3	10.5	12.1	12.1	11.9	11.1	11.2	11.6
	Min	8.1	8.3	8.3	7.9	8.6	9.2	9.8	9.7	9.8	8.61	8.9	8.3
	SD	0.7	0.5	0.3	0.7	0.4	0.4	0.6	0.6	0.52	0.6	0.6	0.8
	CV (%)	0.1	0.1	0.3	0.1	0.1	0.3	0.1	0.1	0.1	0.1	0.1	0.1
	N	15	14	14	15	15	15	18	16	17	16	16	16
(NTU)	Avg	4.3	3.2	3.2	2.4	2.8	2.7	1.5	1.2	2.2	2.1	3.9	1.1
	Max	10.3	14	14	6.4	7.9	15	18	3.7	5.8	2.5	6.9	2.6
	Min	0.8	0.7	1.1	0.3	0.3	14	0.4	0.2	0.4	1.6	2.3	0.3
	SD	4.5	2.9	2.4	2.1	2.9	2.5	1.7	1.3	2.5	0.6	2.6	1.3
	CV (%)	1.0	0.9	0.8	0.9	1.0	0.9	0.8	1.1	1.1	0.3	0.7	1.1
	N	4	5	4	6	5	5	5	7	4	2	3	3
Chl-a (µg/L)	Avg	0.7	0.8	0.7	1.1	1.1	1.3	1.1	1.4	1.8	0.8	0.8	0.8
	Max	1.9	1.9	1.9	1.9	2.0	2.2	3.1	3.3	3.7	1.9	1.9	1.9
	Min	0.2	0.1	0.1	0.5	0.5	0.5	0.3	0.3	0.3	0.1	0.1	0.2
	SD	0.5	0.6	0.6	0.4	0.4	0.6	0.6	0.8	0.8	0.5	0.5	0.5
	CV (%)	0.8	0.7	0.8	0.3	0.4	0.5	0.6	0.7	0.7	0.7	0.7	0.6
	N	13	11	12	12	12	12	17	16	15	13	13	13
NT (mg/L)	Avg	156	158	156	157	157	157	90.0	92.1	89.9	94.4	85.2	94.3
	Max	310	310	310	166	166	166	177	177	177	166	166	166
	Min	15.6	15.6	15.6	148	148	148	30.7	30.7	30.7	40.4	35.6	40.4
	SD	111	108	111	9.0	8.9	8.9	56.6	56.4	56.6	56.1	50.8	56.1
	CV (%)	0.7	0.7	0.7	0.1	0.1	0.1	0.6	0.6	0.6	0.6	0.6	0.6
	N	6	9	6	3	3	3	6	9	6	5	7	5
PT (mg/L)	Avg	127	71.4	48.1	6.0	6.9	7.3	6.4	7.1	9.8	8.4	9.1	12.7
	Max	13	11	12	8.6	9.3	11.0	11.3	13.1	17.6	12.6	18.6	31.2
	Min	2.9	3.2	3.4	3	4.5	4.4	2.6	3.1	2.8	3.3	3.3	3.4
	SD	275	128	85.3	1.8	2.1	2.9	2.8	3.4	5.8	4.2	6.2	11.2
	CV (%)	2.2	1.8	1.8	0.3	0.3	0.4	0.4	0.5	0.6	0.5	0.8	0.9
	N	5	5	5	5	5	5	8	7	7	5	5	5

3.2. Meteorological Conditions

Observations made at Lake Ranco between 1989 and 2022 indicate that the mean monthly value of daily air temperature was 14.00 °C, also varying between a summer maximum of 19.63 °C (February) and a winter minimum of 10.18 °C (July). In general, cloudiness increases during the winter months, with a consequent decrease in temperature, and vice versa during the summer months. Average cloudiness in the area was 68.33%, with a winter maximum of 92% (August) and a summer minimum of 27% (February). According to the hydrometric statistics, total precipitation during the entire study period was 2789.66 mm, ranging from a maximum value of 346.80 mm (July) to a minimum value of 104.24 mm (February). There is no dry season in this region, but there is a significant decrease in precipitation during the summer. The surface wind regime during the winter was mainly from the northwest (see Figure 7), with a maximum speed of 2.86 m/s, and during the summer, with a minimum speed of 1.03 m/s. The relative humidity was high throughout the study period, excluding summer (>61.76%). For more details, see Table S2.
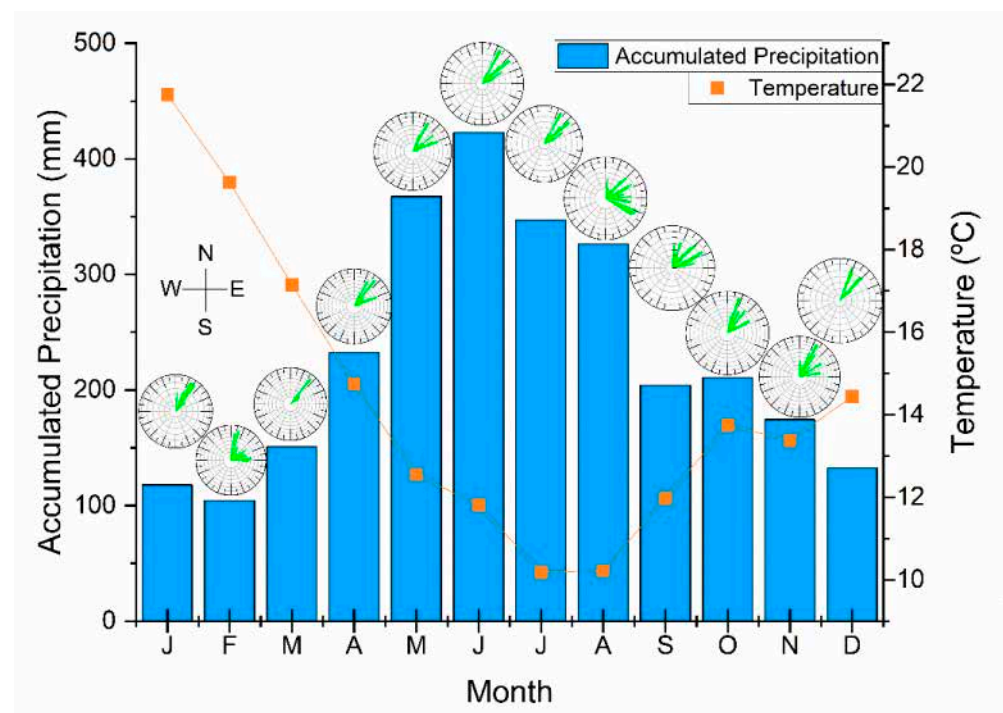

Figure 7. Meteorological conditions in Lake Ranco over 1989–2022.

3.3. Models Estimation

3.3.1. Case 1

Figure 8 shows the behavior of the estimated chlorophyll-a at the three-sampling station and station 6 with all values included for variables used of the Case 1. The estimation models used for each case were RNN, LSTM, GRU, and TCN.

From the results, we can observe in Figure 8 that for most of the stations, the models offer good recovery in terms of temporal variations. In addition, the LSTM and GRU models present more accurate values of performance metrics for LSTM (MSE = 0.10; RMSE = 0.32 µg/L; and MAE = 1.25 µg/L and R^2 = 0.89), while the GRU presents MSE = 0.11 (µg/L)^2, RMSE = 0.34 µg/L, MAE = 1.65 µg/L and R^2 = 0.88 compared to the other models (RNN and TCN) (see detail in Table 4). However, the modeling for station 6, in which all total values are used, also showed good results for TCN and RNN.
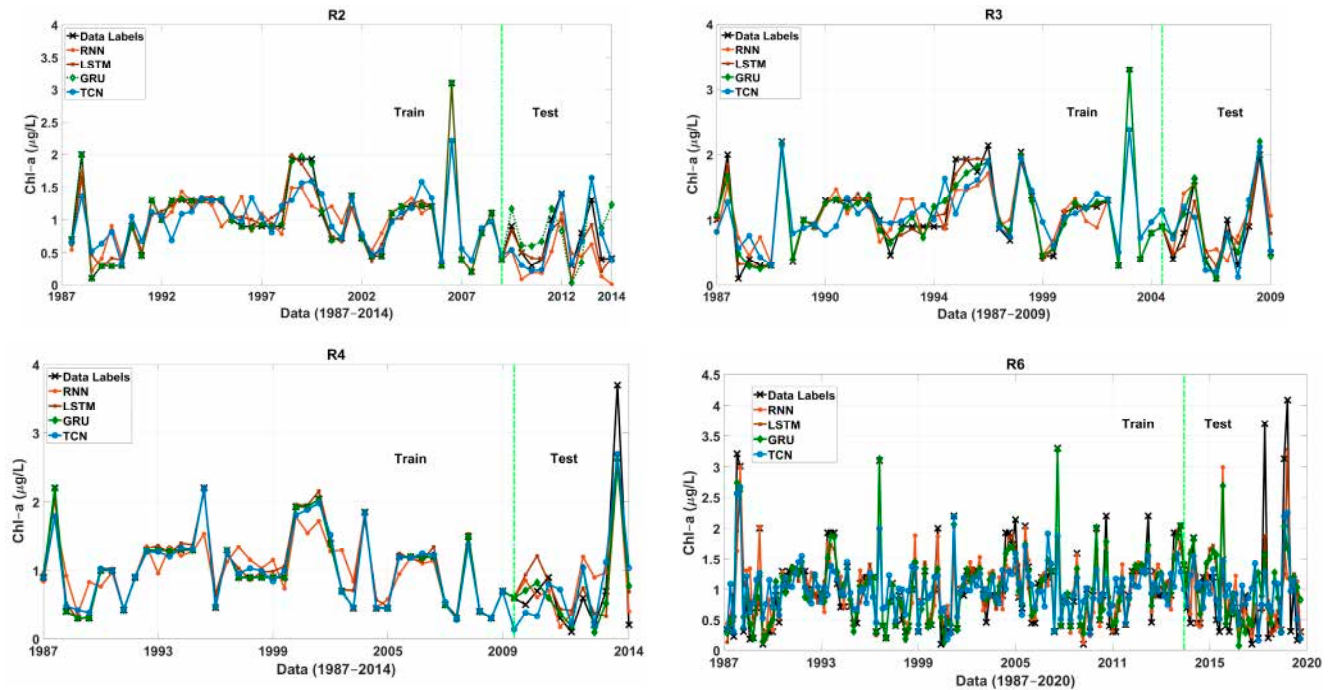
Figure 8. Chl-a estimation values Case 1 in each lake station.

Table 4. Model performance metrics for chlorophyll-a prediction based on RNN, LSTM, GRU, and TCN.

Case	Model	R^2	RMSE (µg/L)	MAE (µg/L)	MSE (µg/L)^2
1	RNN	0.72	0.27	1.97	0.25
	LSTM	0.89	0.32	1.25	0.10
	GRU	0.88	0.34	1.65	0.11
	TCN	0.73	0.49	2.40	2.40
2	RNN	0.39	0.66	1.36	0.43
	LSTM	0.86	0.37	0.28	0.13
	GRU	0.85	0.38	1.08	0.14
	TCN	0.66	0.54	1.73	0.29
3	RNN	0.66	0.71	4.16	0.50
	LSTM	0.69	0.42	2.00	0.64
	GRU	0.82	0.30	2.86	0.52
	TCN	0.96	0.13	0.06	0.33

3.3.2. Case 2

Figure 9 shows the behavior of the four estimation models of Chl-a at each lake sampling station in Case 2, where meteorological variables were used.

Based on the results shown in Figure 9, we can observe a similar behavior in the performance of the models as that seen in Case 1. The LSTM and GRU models prevail as the two with the best metrics in estimating Chl-a: LSTM with R^2 up to 0.89 and a minimum of 0.86, an RSME of 0.37 µg/L, an MAE of 0.28 µg/L, and an MSE of 0.13 (µg/L)^2; GRU resulted in R^2 = 0.85, RMSE = 0.38 µg/L, MAE = 1.08 µg/L, and MSE = 0.14 (µg/L)^2.
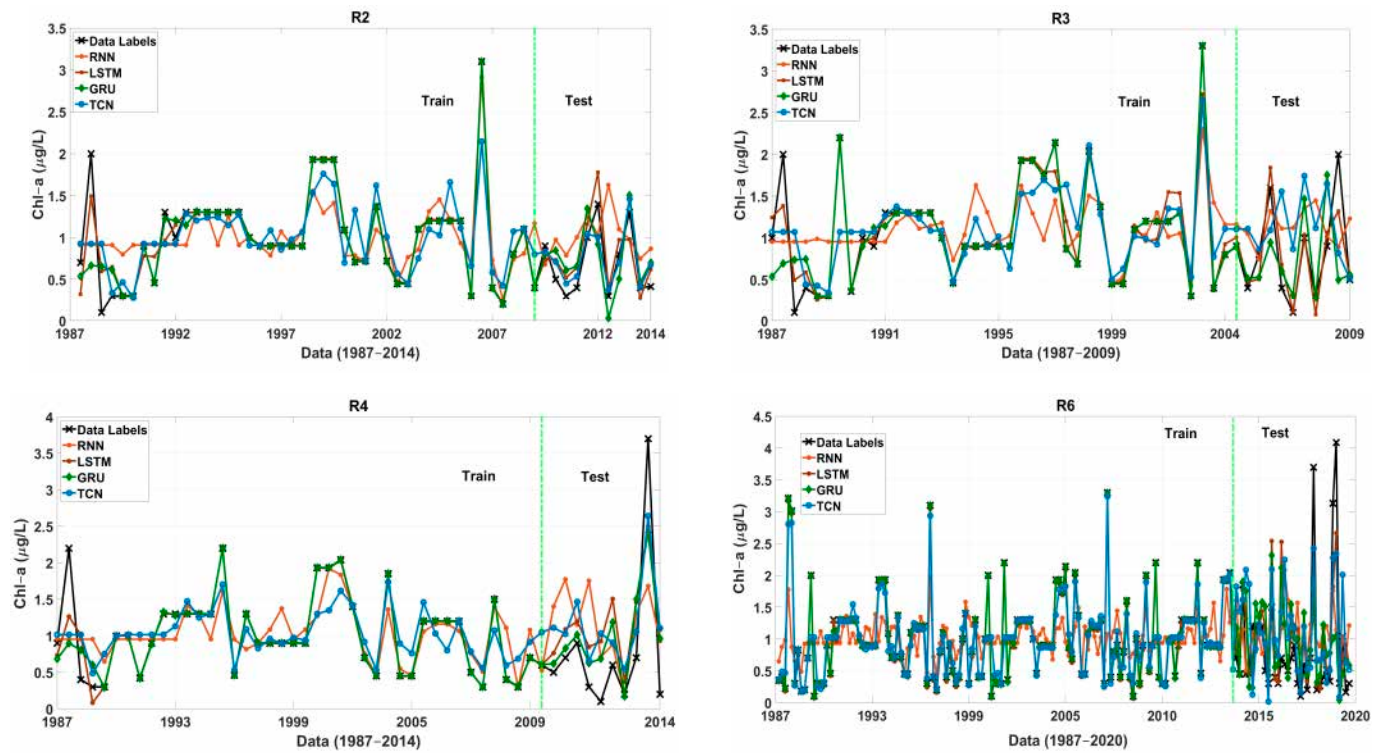
Figure 9. Chl-a estimation values Case 2 in each lake station.

3.3.3. Case 3

Figure 10 shows the behavior of the train/test for the four estimation models of Chl-a in each lake sampling station in Case 3 when we used the satellite spectral bands and indices described in Table 2.
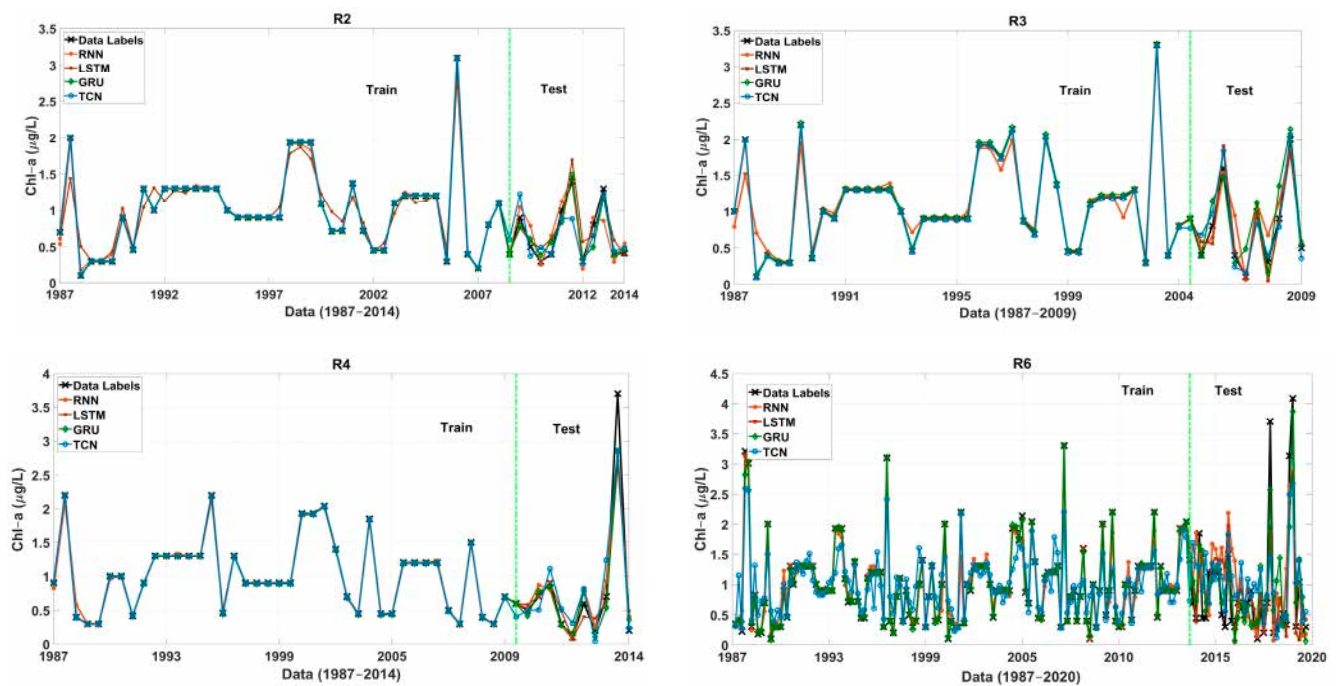

Figure 10. Chl-a estimation values for Case 3 at each lake station.

When comparing the results between Case 2 and Case 3, a marked improvement in the performance of all models in Case 3 is observed. This improvement is manifested in

lower MSE ($0.01 < (\mu g/L)^2$), RMSE ($0.09 < \mu g/L$), and MAE ($0.05 < \mu g/L$) values, along with higher R^2 (0.96) values. However, it is noteworthy that the TCN, with an increasing number of data, performed excellently ($R^2 = 0.96$; MSE = 0.33 $(\mu g/L)^2$; RMSE = 0.13 $\mu g/L$; and MAE = 0.06 $\mu g/L$), as illustrated in Table 4. This suggests that the integration of a broader set of variables related to chlorophyll-a, both directly and indirectly, improves the predictive ability of the other algorithms. As in the previous cases, the GRU model performed well with metrics of $R^2 = 0.82$, RMSE = 0.30 $\mu g/L$, MSE = 0.52 $(\mu g/L)^2$, and MAE = 2.86 $\mu g/L$.

3.4. Statistical Analysis

The results show very good performance, with a coefficient of determination until 0.96, a mean square error between 0.10 $(\mu g/L)^2$ and 2.40 $(\mu g/L)^2$, a root mean square error between 0.13 and 0.54 $\mu g/L$, and a mean absolute error between 0.06 and 4.16 (Table 4). The worst performance is observed when using the RNN. In the meantime, the best-performing models are LSTM and TCN.

4. Discussion

Lakes are lentic ecosystems that are of profound importance on a global scale due to the diverse range of ecosystem services they provide [68,69]. In an era marked by increasing water stress and scarcity, the preservation of freshwater reservoirs becomes paramount to maintain their pristine conditions and sustain their invaluable contributions [70]. These bodies of water not only serve as vital habitats for numerous species but also play a critical role in regulating local and regional climates, supporting agriculture, providing recreational opportunities, and supplying drinking water to communities [71]. Ensuring the resilience and vitality of lakes is, therefore, essential to the well-being of both ecosystems and human societies. Efforts to safeguard these aquatic environments require multifaceted approaches that encompass effective management strategies, conservation initiatives, and proactive measures to mitigate threats, such as pollution, habitat degradation, and the effects of climate change. Through collaboration and continued dedication, we can aspire to a future in which lakes thrive as resilient and biodiverse ecosystems, enriching our planet and sustaining life for generations to come.

This study investigates the effectiveness of employing four different machine learning models to measure the presence of algae in a South American lake located in the picturesque Lake District of southern Chile. The focus is on estimating the concentration of chlorophyll-a (Chl-a), a key pigment indicative of the presence of algae. For the initial estimation approach, the study uses a large dataset spanning from 1989 to 2023, covering various water quality parameters. These data were collected during monitoring campaigns conducted by prestigious institutions such as the Chilean General Water Directorate and the EULA Environmental Studies Center. In a second case, meteorological parameters from the Chilean Meteorological Directorate are integrated into the analysis to improve the predictive capability of the models. In addition, the study in Case 3 incorporates spectral bands and indices derived from multispectral satellite images obtained from reputable sources such as Landsat and Sentinel satellites.

The results of the models show a very good performance, with a coefficient of determination until 0.96, a mean square error between 0.28 and 2.40 $(\mu g/L)^2$, a root mean square error between 0.13 and 0.54 $\mu g/L$, and a mean absolute error between 0.06 and 4.16 $\mu g/L$. The worst performance is observed for the RNN, probably because the number of layers chosen is not sufficiently complex to establish adequate to establish adequate relationships between inputs and output. In the meantime, the best performing models are LSTM, with $R^2 = 0.89$, an RSME of 0.32 $\mu g/L$, an MAE 1.25 $\mu g/L$, and an MSE 0.25 $(\mu g/L)^2$) and TCN ($R^2 = 0.96$; MSE = 0.33 $(\mu g/L)^2$; RMSE = 0.13 $\mu g/L$; and MAE = 0.06 $\mu g/L$). These findings are consistent with previous research in other Chilean lakes, which have also reported increased chlorophyll-a concentrations and algal biomass. For example, a cyanobacterial bloom with elevated chlorophyll-a levels were observed in Lake Laja in 2020 through

the use of various vegetation indices using a combination of satellite spectral bands [16] and, since the 1990s, Lake Villarrica has experienced recurrent summer blooms, which have been detected through Landsat satellite images [15]. In addition, a persistent bloom of cyanobacteria has been observed in Laguna Grande de San Pedro throughout 2023 and 2024, and spectral models have been created through the use of satellite images [72]. Furthermore, studies have been conducted to estimate chlorophyll-a levels in lakes such as Llanquihue and Maihue using combined machine learning and remote sensing techniques [28,29]. Chile's lakes are increasingly affected by the combined effects of human activities and climate change, leading to more frequent algal blooms, deteriorating water quality, and accelerated eutrophication. Therefore, it is essential to closely monitor Lake Ranco to prevent similar episodes from occurring in other lakes in the country. Chl-a estimates help to better understand the behavior of the lake's water quality in the absence of in situ measurements (high cost and several monitoring campaigns), and thus, contribute to assessing the behavior of the ecosystem services provided by the lake, whether drinking water for different uses, such as tourism and recreation.

Lake Ranco in South America is a vital provider of ecosystem services that a re crucial to the livelihoods of the surrounding communities and biodiversity. These services encompass a wide range of benefits, including water supply for domestic, agricultural, and industrial purposes. In addition, the lake supports recreational activities, tourism and cultural practices, fostering economic growth and social well-being in the region. Its rich biodiversity also contributes to regulating water quality, nutrient cycling, and carbon sequestration, thus maintaining ecological balance and resilience.

New methodologies and ecosystem services mapping outputs are essential pillars for reaching targets set by environmental policy instruments at the national level [73]. However, the conservation status of this pristine ecosystem is of great concern. Rapid urbanization, agricultural expansion, and industrialization of the surrounding areas pose significant threats to the ecological integrity of the lake. Pollution from agricultural runoff, industrial discharges, and untreated sewage can lead to eutrophication, algal blooms, and habitat degradation, events that have occurred in lakes in the same chain, such as Lake Villarrica. In addition, deforestation and land use changes exacerbate soil erosion and sedimentation, impacting water quality and aquatic habitats. To safeguard the ecological health and long-term viability of this invaluable natural resource, effective conservation strategies and sustainable management practices are imperative. To mitigate these anthropogenic pressures and ensure the lake's continued provision of ecosystem services, collaboration between government agencies, local communities, and stakeholders is essential.

5. Conclusions

The use of machine learning-based estimation models and the incorporation of multiple data types, such as water quality parameters, meteorological data, and data from satellite sources, has proven to be an effective and accurate method. The models used in this study have demonstrated good performance in most cases, as evidenced by the statistical metrics provided by each model. In Case 3, the TCN model stands out with excellent chlorophyll-a estimation metrics with an R^2 of 0.96, MSE = 0.33 $(\mu g/L)^2$, RMSE = 0.13 $\mu g/L$, and MAE = 0.06 $\mu g/L$. This suggests that integrating a broader set of chlorophyll-a-related variables, both directly and indirectly, improves the predictive ability of the algorithms. In addition, we evaluated the meteorological conditions as they may influence the variability of the studied lake. Finally, we described the ecosystem services associated with the multiple uses provided by this lake and its surrounding watershed. The methodology developed by us will be used in other lake bodies in Chile and will serve as a basic tool for water resource managers.

The shoreline of Lake Ranco exhibits an interaction between natural, human, and rural elements, with agricultural areas predominating. This has disrupted the natural balance, threatening biodiversity, and ecosystem services. Although urban areas do not encompass the entire perimeter, two key urban centers have contributed to ecological deterioration. It

is essential to implement conservation measures and sustainable management practices to mitigate these impacts and preserve the biodiversity and ecosystem services essential for the community and the natural environment along the shoreline of Lake Ranco.

References

1. Chen, B.; Zhang, M.; Yang, R.; Tang, W. Spatiotemporal Variations in the Carbon Sequestration Capacity of Plateau Lake Wetlands Regulated by Land Use Control under Policy Guidance. *Land* **2023**, *12*, 1695. [CrossRef]
2. Shen, W.; Liu, Z. Urbanization Shifts Freshwater Service Flows in the Highly Urbanized Watersheds of Dongjiang River, China. *Appl. Geogr.* **2023**, *161*, 103140. [CrossRef]
3. Mishra, R.K. The Effect of Eutrophication on Drinking Water. *Br. J. Multidiscip. Adv. Stud.* **2023**, *4*, 7–20. [CrossRef]
4. Ramanaiah, S.V.; Chandrasekhar, K.; Cordas, C.M.; Potoroko, I. Bioelectrochemical Systems (BESs) for Agro-Food Waste and Wastewater Treatment, and Sustainable Bioenergy-A Review. *Environ. Pollut.* **2023**, *325*, 121432. [CrossRef] [PubMed]
5. Cantonati, M.; Poikane, S.; Pringle, C.M.; Stevens, L.E.; Turak, E.; Heino, J.; Richardson, J.S.; Bolpagni, R.; Borrini, A.; Cid, N.; et al. Characteristics, Main Impacts, and Stewardship of Natural and Artificial Freshwater Environments: Consequences for Biodiversity Conservation. *Water* **2020**, *12*, 260. [CrossRef]
6. Heino, J.; Alahuhta, J.; Bini, L.M.; Cai, Y.; Heiskanen, A.S.; Hellsten, S.; Kortelainen, P.; Kotamäki, N.; Tolonen, K.T.; Vihervaara, P.; et al. Lakes in the Era of Global Change: Moving beyond Single-Lake Thinking in Maintaining Biodiversity and Ecosystem Services. *Biol. Rev.* **2021**, *96*, 89–106. [CrossRef]
7. Bănăduc, D.; Simić, V.; Cianfaglione, K.; Barinova, S.; Afanasyev, S.; Öktener, A.; McCall, G.; Simić, S.; Curtean-Bănăduc, A. Freshwater as a Sustainable Resource and Generator of Secondary Resources in the 21st Century: Stressors, Threats, Risks, Management and Protection Strategies, and Conservation Approaches. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16570. [CrossRef]
8. Gong, D.; Huang, M.; Lin, H. Construction of an Ecological Security Pattern in Rapidly Urbanizing Areas Based on Ecosystem Sustainability, Stability, and Integrity. *Remote Sens.* **2023**, *15*, 5728. [CrossRef]
9. Zhang, G.; Yao, T.; Xie, H.; Yang, K.; Zhu, L.; Shum, C.K.; Bolch, T.; Yi, S.; Allen, S.; Jiang, L.; et al. Response of Tibetan Plateau Lakes to Climate Change: Trends, Patterns, and Mechanisms. *Earth Sci. Rev.* **2020**, *208*, 103269. [CrossRef]
10. Zhao, L.; Li, T.; Przybysz, A.; Liu, H.; Zhang, B.; An, W.; Zhu, C. Effects of Urban Lakes and Neighbouring Green Spaces on Air Temperature and Humidity and Seasonal Variabilities. *Sustain. Cities Soc.* **2023**, *91*, 104438. [CrossRef]
11. Ferreira, V.; Albariño, R.; Larrañaga, A.; LeRoy, C.J.; Masese, F.O.; Moretti, M.S. Ecosystem Services Provided by Small Streams: An Overview. *Hydrobiologia* **2023**, *850*, 2501–2535. [CrossRef]
12. Fernández-Martínez, M.; Barquín, J.; Bonada, N.; Cantonati, M.; Churro, C.; Corbera, J.; Delgado, C.; Dulsat-Masvidal, M.; Garcia, G.; Margalef, O.; et al. Mediterranean Springs: Keystone Ecosystems and Biodiversity Refugia Threatened by Global Change. *Glob. Change Biol.* **2024**, *30*, e16997. [CrossRef] [PubMed]
13. Guo, R.Z.; Lin, L.; Xu, J.F.; Dai, W.H.; Song, Y.B.; Dong, M. Spatio-Temporal Characteristics of Cultural Ecosystem Services and Their Relations to Landscape Factors in Hangzhou Xixi National Wetland Park, China. *Ecol. Indic.* **2023**, *154*, 110910. [CrossRef]

14. Ghasemi, M.; González-García, A.; Charrahy, Z.; Serrao-Neumann, S. Utilizing Supply-Demand Bundles in Nature-Based Recreation Offers Insights into Specific Strategies for Sustainable Tourism Management. *Sci. Total Environ.* **2024**, *922*, 171185. [CrossRef] [PubMed]

15. Rodríguez-López, L.; Duran-Llacer, I.; Bravo Alvarez, L.; Lami, A.; Urrutia, R. Recovery of Water Quality and Detection of Algal Blooms in Lake Villarrica through Landsat Satellite Images and Monitoring Data. *Remote Sens.* **2023**, *15*, 1929. [CrossRef]

16. Rodríguez-López, L.; Duran-Llacer, I.; González-Rodríguez, L.; Abarca-del-Rio, R.; Cárdenas, R.; Parra, O.; Martínez-Retureta, R.; Urrutia, R. Spectral Analysis Using LANDSAT Images to Monitor the Chlorophyll-a Concentration in Lake Laja in Chile. *Ecol. Inform.* **2020**, *60*, 101183. [CrossRef]

17. Rodríguez-López, L.; González-Rodríguez, L.; Duran-Llacer, I.; Cardenas, R.; Urrutia, R. Spatio-Temporal Analysis of Chlorophyll in Six Araucanian Lakes of Central-South Chile from Landsat Imagery. *Ecol. Inform.* **2021**, *65*, 101431. [CrossRef]

18. Adjovu, G.E.; Stephen, H.; James, D.; Ahmad, S. Overview of the Application of Remote Sensing in Effective Monitoring of Water Quality Parameters. *Remote Sens.* **2023**, *15*, 1938. [CrossRef]

19. Radeloff, V.C.; Roy, D.P.; Wulder, M.A.; Anderson, M.; Cook, B.; Crawford, C.J.; Friedl, M.; Gao, F.; Gorelick, N.; Hansen, M.; et al. Need and Vision for Global Medium-Resolution Landsat and Sentinel-2 Data Products. *Remote Sens. Environ.* **2024**, *300*, 113918. [CrossRef]

20. Amieva, J.F.; Oxoli, D.; Brovelli, M.A. Machine and Deep Learning Regression of Chlorophyll-a Concentrations in Lakes Using PRISMA Satellite Hyperspectral Imagery. *Remote Sens.* **2023**, *15*, 5385. [CrossRef]

21. Shahvaran, A.R.; Kheyrollah Pour, H.; Van Cappellen, P. Comparative Evaluation of Semi-Empirical Approaches to Retrieve Satellite-Derived Chlorophyll-a Concentrations from Nearshore and Offshore Waters of a Large Lake (Lake Ontario). *Remote Sens.* **2024**, *16*, 1595. [CrossRef]

22. de Liz Arcari, A.; Tavora, J.; van der Wal, D.; Salama, M.S. The Wastewater Contamination Index: A Methodology to Assess the Risk of Wastewater Contamination from Satellite-Derived Water Quality Indicators. *Front. Environ. Sci.* **2023**, *11*, 1130655. [CrossRef]

23. D'Ugo, E.; Mukherjee, A.; Giuseppetti, R.; Tucci, M.; Bucci, P.; Aulenta, F.; Laneve, G.; Magurano, F. Integration of Satellite Surveillance and Metagenomics for the Monitoring and Protection of Water Basins from Oil Spills. *Environ. Adv.* **2024**, *15*, 100498. [CrossRef]

24. da Silva, L.D.; Mahmoud, M.; González-Rodríguez, L.; Mohammed, S.; Rodríguez-López, L.; Arias, M.I.A. Assessment of the IMERG Early-Run Precipitation Estimates over South American Country of Chile. *Remote Sens.* **2023**, *15*, 573. [CrossRef]

25. Mukonza, S.S.; Chiang, J.L. Meta-Analysis of Satellite Observations for United Nations Sustainable Development Goals: Exploring the Potential of Machine Learning for Water Quality Monitoring. *Environments* **2023**, *10*, 170. [CrossRef]

26. Sudriani, Y.; Sebestyen, V.; Abonyi, J. Surface Water Monitoring Systems—The Importance of Integrating Information Sources for Sustainable Watershed Management. *IEEE Access* **2023**, *11*, 36421–36451. [CrossRef]

27. Kamyab, H.; Khademi, T.; Chelliapan, S.; SaberiKamarposhti, M.; Rezania, S.; Yusuf, M.; Farajnezhad, M.; Abbas, M.; Hun Jeon, B.; Ahn, Y. The Latest Innovative Avenues for the Utilization of Artificial Intelligence and Big Data Analytics in Water Resource Management. *Results Eng.* **2023**, *20*, 101566. [CrossRef]

28. Rodríguez-López, L.; Usta, D.B.; Duran-Llacer, I.; Alvarez, L.B.; Yépez, S.; Bourrel, L.; Frappart, F.; Urrutia, R. Estimation of Water Quality Parameters through a Combination of Deep Learning and Remote Sensing Techniques in a Lake in Southern Chile. *Remote Sens.* **2023**, *15*, 4157. [CrossRef]

29. Rodríguez-López, L.; Alvarez, D.; Bustos Usta, D.; Duran-Llacer, I.; Bravo Alvarez, L.; Fagel, N.; Bourrel, L.; Frappart, F.; Urrutia, R. Chlorophyll-a Detection Algorithms at Different Depths Using In Situ, Meteorological, and Remote Sensing Data in a Chilean Lake. *Remote Sens.* **2024**, *16*, 647. [CrossRef]

30. Aranda, A.C.; Rivera-Ruiz, D.; Rodríguez-López, L.; Pedreros, P.; Arumí-Ribera, J.L.; Morales-Salinas, L.; Fuentes-Jaque, G.; Urrutia, R. Evidence of Climate Change Based on Lake Surface Temperature Trends in South Central Chile. *Remote Sens.* **2021**, *13*, 4535. [CrossRef]

31. Municipalidad Lago Ranco. Ilustre Municipalidad de Lago Ranco "Modificación Plan Regulador Comunal de Lago Ranco"; Ranco. 2013. Available online: https://eae.mma.gob.cl/storage/documents/02_3er_IA_PRC_Lago_Ranco.pdf.pdf (accessed on 1 August 2024).

32. Hermosilla, K.; Gutiérrez, M.; Escalona, M. Caracterizacion de la oferta turistica y zonificacion en la cuenca del lago. *Estud. Perspect. Tur.* **2011**, *20*, 943–959.

33. Figueroa, E.S. *El Turismo Sustentable Como Base el Desarrollo Económico Local: El Caso de la Comuna de Lago Ranco*; Mérida, Venezuela, 2018.

34. Salinas, O. El Secreto Verde del Lago Ranco. Available online: https://parquefutangue.com/blog/secreto-lago-ranco/ (accessed on 1 August 2024).

35. Saravia, G. Nuevas Configuraciones del Paisaje del Agua En Los Lagos Nor-Patagónicos de Chile. *Rev. Bordes Lacustres* **2021**, *1*, 31–79.

36. Vanhellemont, Q. Sensitivity Analysis of the Dark Spectrum Fitting Atmospheric Correction for Metre- and Decametre-Scale Satellite Imagery Using Autonomous Hyperspectral Radiometry. *Opt. Express* **2020**, *28*, 29948. [CrossRef] [PubMed]

37. Vanhellemont, Q.; Ruddick, K. Atmospheric Correction of Metre-Scale Optical Satellite Data for Inland and Coastal Water Applications. *Remote Sens. Environ.* **2018**, *216*, 586–597. [CrossRef]

38. Vanhellemont, Q. Adaptation of the Dark Spectrum Fitting Atmospheric Correction for Aquatic Applications of the Landsat and Sentinel-2 Archives. *Remote Sens. Environ.* **2019**, *225*, 175–192. [CrossRef]

39. Abirhire, O.; Davies, J.M.; Guo, X.; Hudson, J. Understanding the Factors Associated with Long-Term Reconstructed Turbidity in Lake Diefenbaker from Landsat-Imagery. *Sci. Total Environ.* **2020**, *724*, 138222. [CrossRef]

40. Rodríguez-López, L.; Bustos Usta, D.; Bravo Alvarez, L.; Duran-Llacer, I.; Lami, A.; Martínez-Retureta, R.; Urrutia, R. Machine Learning Algorithms for the Estimation of Water Quality Parameters in Lake Llanquihue in Southern Chile. *Water* **2023**, *15*, 1994. [CrossRef]

41. Ma, J.; Jin, S.; Li, J.; He, Y.; Shang, W. Spatio-Temporal Variations and Driving Forces of Harmful Algal Blooms in Chaohu Lake: A Multi-Source Remote Sensing Approach. *Remote Sens.* **2021**, *13*, 427. [CrossRef]

42. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N.; Blaustein, J. Use of a Green Channel in Remote Sensing of Global Vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]

43. Kravitz, J.; Matthews, M.; Lain, L.; Fawcett, S.; Bernard, S. Potential for High Fidelity Global Mapping of Common Inland Water Quality Products at High Spatial and Temporal Resolutions Based on a Synthetic Data and Machine Learning Approach. *Front. Environ. Sci.* **2021**, *9*, 587660. [CrossRef]

44. Rodríguez-López, L.; Duran-Llacer, I.; González-Rodríguez, L.; Cardenas, R.; Urrutia, R. Retrieving Water Turbidity in Araucanian Lakes (South-Central Chile) Based on Multispectral Landsat Imagery. *Remote Sens.* **2021**, *13*, 3133. [CrossRef]

45. Pamula, A.S.P.; Gholizadeh, H.; Krzmarzick, M.J.; Mausbach, W.E.; Lampert, D.J. A Remote Sensing Tool for near Real-Time Monitoring of Harmful Algal Blooms and Turbidity in Reservoirs. *J. Am. Water Resour. Assoc.* **2023**, *59*, 929–949. [CrossRef]

46. Cao, H.; Han, L.; Li, L. A Deep Learning Method for Cyanobacterial Harmful Algae Blooms Prediction in Taihu Lake, China. *Harmful Algae* **2022**, *113*, 102189. [CrossRef] [PubMed]

47. Duan, H.; Zhang, Y.; Zhang, B.; Song, K.; Wang, Z. Assessment of Chlorophyll-a Concentration and Trophic State for Lake Chagan Using Landsat TM and Field Spectral Data. *Environ. Monit. Assess.* **2007**, *129*, 295–308. [CrossRef] [PubMed]

48. Raynolds, M.K.; Comiso, J.C.; Walker, D.A.; Verbyla, D. Relationship between Satellite-Derived Land Surface Temperatures, Arctic Vegetation Types, and NDVI. *Remote Sens. Environ.* **2008**, *112*, 1884–1894. [CrossRef]

49. Markogianni, V.; Kalvas, D.; Petropoulos, G.P.; Dimitriou, E. An Appraisal of the Potential of Landsat 8 in Estimating Chlorophyll-a, Ammonium Concentrations and Other Water Quality Indicators. *Remote Sens.* **2018**, *10*, 1018. [CrossRef]

50. Hu, C. A Novel Ocean Color Index to Detect Floating Algae in the Global Oceans. *Remote Sens. Environ.* **2009**, *113*, 2118–2129. [CrossRef]

51. Boucher, J.; Weathers, K.C.; Norouzi, H.; Steele, B. Assessing the Effectiveness of Landsat 8 Chlorophyll a Retrieval Algorithms for Regional Freshwater Monitoring. *Ecol. Appl.* **2018**, *28*, 1044–1054. [CrossRef]

52. Alawadi, F. Detection of Surface Algal Blooms Using the Newly Developed Algorithm Surface Algal Bloom Index (SABI). In *Remote Sensing of the Ocean, Sea Ice, and Large Water Regions 2010*; SPIE: Bellingham, WA, USA, 2010; Volume 7825, p. 782506. [CrossRef]

53. Gitelson, A.A.; Viña, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote Estimation of Canopy Chlorophyll Content in Crops. *Geophys. Res. Lett.* **2005**, *32*, L08403. [CrossRef]

54. Hassan, M.A.; Yang, M.; Rasheed, A.; Jin, X.; Xia, X.; Xiao, Y.; He, Z. Time-Series Multispectral Indices from Unmanned Aerial Vehicle Imagery Reveal Senescence Rate in Bread Wheat. *Remote Sens.* **2018**, *10*, 809. [CrossRef]

55. Institute of Natural Sciences. *Acolite Manual*; Institute of Natural Sciences: Brussels, Belgium, 2022.

56. Piech, K.R.; Schott, J.R.; Stewart, K.M. The Blue-to-Green Reflectance Ratio and Lake Water Quality Monitoring from Aircraft or Space Considerably Reduces the Amount of in Situ Sampling Required. *Photogramm. Eng. Remote Sens.* **1978**, *44*, 1303–1310.

57. Ruiz-Guirola, D.E.; Lopez, O.L.A.; Montejo-Sanchez, S.; Souza, R.D.; Bennis, M. Performance Analysis of ML-Based MTC Traffic Pattern Predictors. *IEEE Wirel. Commun. Lett.* **2023**, *12*, 1144–1148. [CrossRef]

58. Gers, F.A.; Urgen Schmidhuber, J.; Cummins, F. Learning to Forget: Continual Prediction with LSTM. *Neural Comput.* **2020**, *12*, 2451–2471. [CrossRef] [PubMed]

59. Choi, J.; Won, J.; Jang, S.; Kim, S. Learning Enhancement Method of Long Short-Term Memory Network and Its Applicability in Hydrological Time Series Prediction. *Water* **2022**, *14*, 2910. [CrossRef]

60. de la Fuente, A.; Meruane, V.; Meruane, C. Hydrological Early Warning System Based on a Deep Learning Runoff Model Coupled with a Meteorological Forecast. *Water* **2019**, *11*, 1808. [CrossRef]

61. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A Review of Recurrent Neural Networks: Lstm Cells and Network Architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [CrossRef]

62. Huang, G.; Li, X.; Zhang, B.; Ren, J. PM2.5 Concentration Forecasting at Surface Monitoring Sites Using GRU Neural Network Based on Empirical Mode Decomposition. *Sci. Total Environ.* **2021**, *768*, 144516. [CrossRef]

63. Nosouhian, S.; Nosouhian, F.; Khoshouei, A.K. A Review of Recurrent Neural Network Architecture for Sequence Learning: Comparison between LSTM and GRU. *Preprints* **2021**, 2021070252. [CrossRef]

64. Duc, T.N.; Minh, C.T.; Xuan, T.P.; Kamioka, E. Convolutional Neural Networks for Continuous QoE Prediction in Video Streaming Services. *IEEE Access* **2020**, *8*, 116268–116278. [CrossRef]

65. Liu, S.; Xu, T.; Du, X.; Zhang, Y.; Wu, J. A Hybrid Deep Learning Model Based on Parallel Architecture TCN-LSTM with Savitzky-Golay Filter for Wind Power Prediction. *Energy Convers. Manag.* **2024**, *302*, 118122. [CrossRef]

66. Sharma, M.; Pachori, R.B.; Rajendra Acharya, U. A New Approach to Characterize Epileptic Seizures Using Analytic Time-Frequency Flexible Wavelet Transform and Fractal Dimension. *Pattern Recognit. Lett.* **2017**, *94*, 172–179. [CrossRef]

67. Rodríguez-López, L.; González-Rodríguez, L.; Duran-Llacer, I.; García, W.; Cardenas, R.; Urrutia, R. Assessment of the Diffuse Attenuation Coefficient of Photosynthetically Active Radiation in a Chilean Lake. *Remote Sens.* **2022**, *14*, 4568. [CrossRef]

68. Adrian, R.; O'Reilly, C.M.; Zagarese, H.; Baines, S.B.; Hessen, D.O.; Keller, W.; Livingstone, D.M.; Sommaruga, R.; Straile, D.; Van Donk, E.; et al. Lakes as Sentinels of Climate Change. *Limnol. Oceanogr.* **2009**, *54*, 2283–2297. [CrossRef] [PubMed]

69. Williamson, C.E.; Saros, J.E.; Vincent, W.F.; Smol, J.P. Lakes and Reservoirs as Sentinels, Integrators, and Regulators of Climate Change. *Limnol. Oceanogr.* **2009**, *54*, 2273–2282. [CrossRef]

70. Moss, B. Cogs in the Endless Machine: Lakes, Climate Change and Nutrient Cycles: A Review. *Sci. Total Environ.* **2012**, *434*, 130–142. [CrossRef]

71. Woolway, R.I.; Kraemer, B.M.; Lenters, J.D.; Merchant, C.J.; O'Reilly, C.M.; Sharma, S. Global Lake Responses to Climate Change. *Nat. Rev. Earth Environ.* **2020**, *1*, 388–403. [CrossRef]

72. Yépez, S.; Velásquez, G.; Torres, D.; Saavedra-Passache, R.; Pincheira, M.; Cid, H.; Rodríguez-López, L.; Contreras, A.; Frappart, F.; Cristóbal, J.; et al. Spatiotemporal Variations in Biophysical Water Quality Parameters: An Integrated In Situ and Remote Sensing Analysis of an Urban Lake in Chile. *Remote Sens.* **2024**, *16*, 427. [CrossRef]

73. Inácio, M.; Das, M.; Barceló, D.; Pereira, P. Frameworks for Mapping Lake Ecosystem Services. An Example from Lithuania. *MethodsX* **2023**, *10*, 102015. [CrossRef]