

Article

Deep Learning-Based Time Series Forecasting Models Evaluation for the Forecast of Chlorophyll a and Dissolved Oxygen in the Mar Menor

Francisco Javier López-Andreu ^{*} , Juan Antonio López-Morales , Zaida Hernández-Guillen ,
Juan Antonio Carrero-Rodrigo , Marta Sánchez-Alcaraz , Joaquín Francisco Atenza-Juárez 
and Manuel Erena 

Institute of Agricultural and Environment Research and Development of Murcia—IMIDA, Mayor Street,
La Alberca, 30150 Murcia, Spain

* Correspondence: franciscoj.lopez26@carm.es; Tel.: +34-968-39-4554

Abstract: The Mar Menor is a coastal lagoon of great socio-ecological and environmental value; in recent years, different localized episodes of hypoxia and eutrophication have modified the quality of its waters. The episodes are due to a drop in dissolved oxygen levels below 4 mg/L in some parts of the lagoon and a rise in chlorophyll a to over 1.8 mg/L. Considering that monitoring the Mar Menor and its watershed is essential to understand the environmental dynamics that cause these dramatic episodes, in recent years, efforts have focused on carrying out periodic measurements of different biophysical parameters of the water. Taking advantage of the data collected and the versatility offered by neural networks, this paper evaluates the performance of a dozen advanced neural networks oriented to time series forecasted for the estimation of dissolved oxygen and chlorophyll a parameters. The data used are obtained in the water body by means of sensors carried by a multiparameter oceanographic probe and two agro-climatic stations located near the Mar Menor. For the dissolved oxygen forecast, the models based on the Time2Vec architecture, accompanied by BiLSTM and Transformer, offer an R2 greater than 0.95. In the case of chlorophyll a, three models offer an R2 above 0.92. These metrics are corroborated by forecasting these two parameters for the first time step out of the data set used. Given the satisfactory results obtained, this work is integrated as a new biophysical parameter forecast component in the monitoring platform of the Mar Menor Observatory developed by IMIDA. The results demonstrate that it is feasible to forecast the concentration of chlorophyll a and dissolved oxygen using neural networks specialized in time series forecasts.

Keywords: coastal; monitoring; environment; water quality; hypoxia; eutrophication; deep learning; machine learning; time series; forecasting



Citation: López-Andreu, F.J.; López-Morales, J.A.; Hernández-Guillen, Z.; Carrero-Rodrigo, J.A.; Sánchez-Alcaraz, M.; Atenza-Juárez, J.F.; Erena, M. Deep Learning-Based Time Series Forecasting Models Evaluation for the Forecast of Chlorophyll a and Dissolved Oxygen in the Mar Menor. *J. Mar. Sci. Eng.* **2023**, *11*, 1473. <https://doi.org/10.3390/jmse11071473>

Academic Editors: Feng Zhou and Pengbin Wang

Received: 28 June 2023
Revised: 14 July 2023
Accepted: 22 July 2023
Published: 24 July 2023



Copyright: © 2023 by the authors. 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

The Mar Menor is a Mediterranean coastal saline lagoon located in the Region of Murcia (Spain). It is an ecosystem of high ecological fragility and great socio-economic value, under great pressure from the different human activities in its environment. In recent years, several episodes of eutrophication and hypoxia have caused the water to become turbid and the fauna to die. To determine the level of eutrophication, the chlorophyll a content of algae in the water column is usually measured, combined with other parameters such as phosphorous and nitrogen content and transparency. On the other hand, the depletion of dissolved oxygen until hypoxia or anoxia can lead to deterioration of the water quality or mortality of fauna. For this reason, it is necessary to constantly monitor the environmental and ecological parameters [1] of the Mar Menor. Monitoring the environmental and ecological parameters of the Mar Menor is essential to detect and investigate negative impacts on the environment and to take measures for protection and conservation.

The Autonomous Community of the Region of Murcia approved the Integrated Coastal Zone Management Strategy of the Socio-Ecological System of the Mar Menor and its surroundings through Decree 42/31 March 2021. This strategy establishes the creation of the Observatory of the Mar Menor (OMM), which is responsible for coordinating the control and monitoring of the ecological state of the Mar Menor for monitoring, data collection, and the establishment of a system that offers the greatest amount of information to those who have to make decisions about the lagoon.

Geospatial technologies are used to carry out this monitoring, as they provide accurate, real-time data on the conditions of the sea and its watershed. These technologies include various Copernicus satellites [2], drones, oceanographic buoys, and multi-parametric oceanographic probes equipped with sensors (CTDs) [3] that measure multiple parameters such as temperature, salinity, turbidity, chlorophyll a concentration, dissolved oxygen and organic matter content, sea level oscillations, and surface runoff inputs from the catchment. These data are collected and analyzed using data science to obtain information on trends and patterns in the Mar Menor ecosystem [4]. These elements allow for more accurate and efficient monitoring, leading to better decision-making and sustainable ecosystem management.

The explosive emergence of Machine Learning (ML) and Deep Learning (DL) allows their application in various fields. These techniques are often used for forecasting, monitoring, and as computational tools for model evaluation. Compared to traditional numerical models, they have become more efficient and less computationally intensive tools. If the models, their hyperparameters, and their input parameters are finely tuned, they can accurately predict the target variable. Moreover, as new data are added to the base dataset, they can be re-trained to increase the accuracy of the models [5].

This work aims to develop a mechanism to forecast the dissolved oxygen and chlorophyll a values for each of the 12 measurement points based on the weekly samples in the Mar Menor water body. In this way, hypoxia and eutrophication episodes can be detected in advance. To do this, we use the data generated by a CTD and those provided by the agro-climatic stations established in the watershed of the Mar Menor hypersaline lagoon. Once homogenized at the temporal frequency level, the data are provided to advanced neural networks specialized in time series forecasts: firstly, to obtain the most reliable forecast of future dissolved oxygen and chlorophyll a values, and secondly, to understand the strengths and weaknesses of each of these neural networks to improve them. Finally, this mechanism is transformed into a process integrated into the ML Model Generator module of the OMM's Mar Menor lagoon monitoring system [6].

2. Study Area

The Mar Menor is a hypersaline coastal lagoon with a surface area of 135 km², the largest in the Iberian Peninsula and one of the largest coastal lagoons in Europe. Its total water volume is 653 hm³, and its maximum depth is 7 m. As shown in Figure 1, it is located on the southeastern coast of Spain, specifically in the so-called Campo de Cartagena, which occupies an area of 1609 km², draining into the Mar Menor.

The weather in the area is hot and dry and has a Mediterranean climate characterized by an arid or semi-arid climate. The average year-round temperature in the area varies between 18 and 23 °C. In recent years, the average annual rainfall has been around 250 mm, with much of it occurring during autumn storms.

The Mar Menor, like all coastal lagoons, is a particular ecosystem on the border between the mainland and the sea [7]. It is one of the most important ecosystems in the Mediterranean. We highlight the existence of seagrass meadows (*Cymodocea nodosa* and *Ruppia cirrhosa*), fish species of special interest such as seahorses (*Hippocampus ramulosus*), high densities of nacre (*Pinna nobilis*), as well as important communities of aquatic birds, some of enormous biological importance such as Audouin's gull (*Larus audouinii*) and common terns (*Sternula albifrons*). It also has two lagoon systems converting into salt flats, five interchange zones with the Mediterranean Sea, five islands of volcanic origin, and three wetlands.

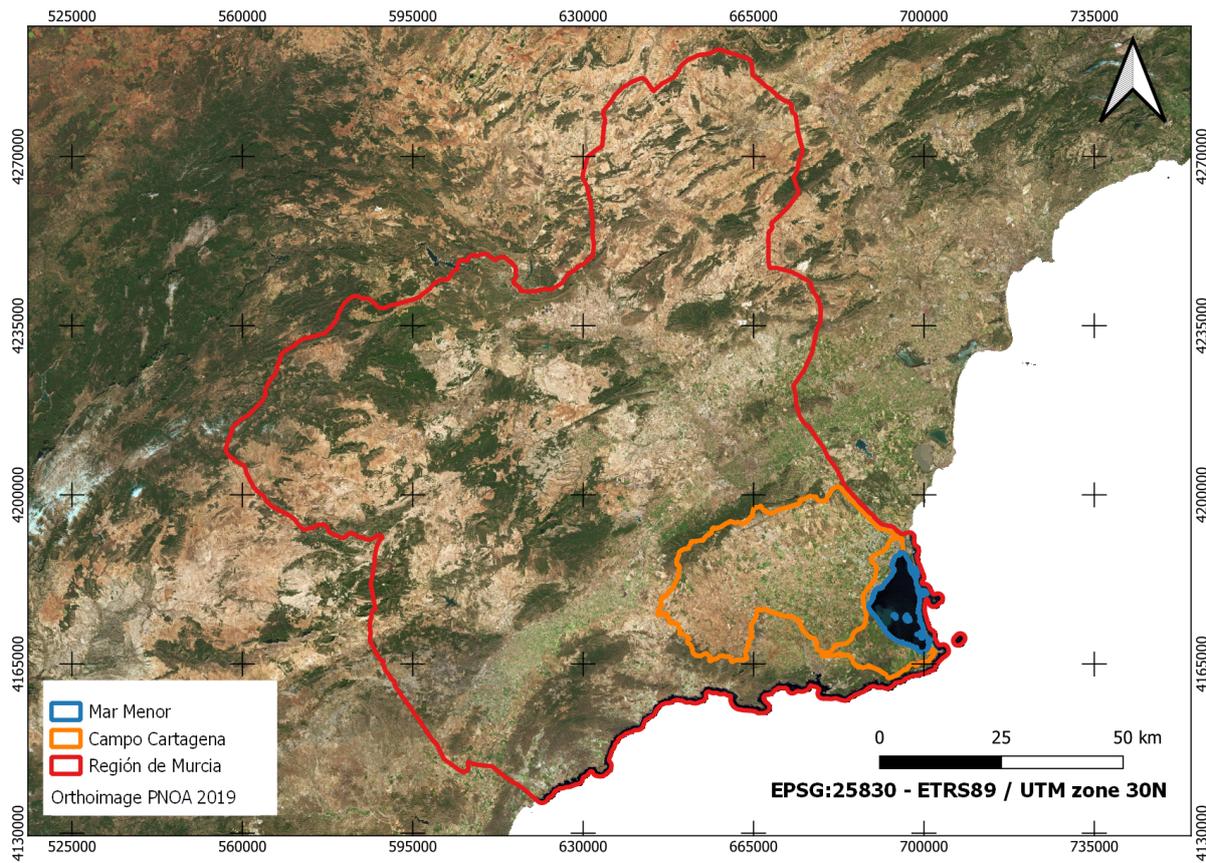


Figure 1. Work study area.

In addition to its ecological importance, it has great socioeconomic importance and relevance due to tourist, recreational, and fishing use and exploitation, as well as the agricultural activity that develops around the lagoon. Therefore, the Mar Menor can be considered a key ecosystem from an ecological perspective and a socioeconomic and cultural point of view [8].

3. Materials

The data used in this work come from two different sources: those collected by a CTD are used to measure the biophysical variables of the water body and the data provided by two agroclimatic stations located in its watershed area a few kilometers from the Mar Menor water body. Figure 2 shows the location of the sources of information.

The processes implemented for this work have been developed with the Python programming language. The most important libraries used are keras [9] and tensorflow [10] for the construction of neural networks, the pandas library [11] is used for data manipulation, and scikit-learn [12] is used for the comparison of metrics and data scaling. The data relating to the CTD are stored in a PostgreSQL database, and those generated by the agroclimatic stations are in an Oracle database. As for the execution environment, a computer with a Windows Server 2019 operating system is used, with an AMD Ryzen Threadripper 2950X processor, 64GB of RAM, and an Nvidia RTX2080Ti graphics card.

3.1. Lagoon Biophysical Data

On a weekly frequency, different biophysical parameters of the Mar Menor are monitored using a CTD [13]. The SeaBird model SBE 19plus is equipped with seven sensors on board. These seven sensors provide temperature, conductivity, dissolved oxygen, salinity, chlorophyll a, turbidity, and depth data. The measurement in each of the points is carried out manually. In the first step, the CTD is left suspended in the water so that it is tempered,

and the different sensors and conduits are filled with water. The CTD is then slowly lowered to the bottom of the sea. The CTD used performs approximately four readings per second. The Figure 3 shows the CTD outside and during the in-water measurement process.

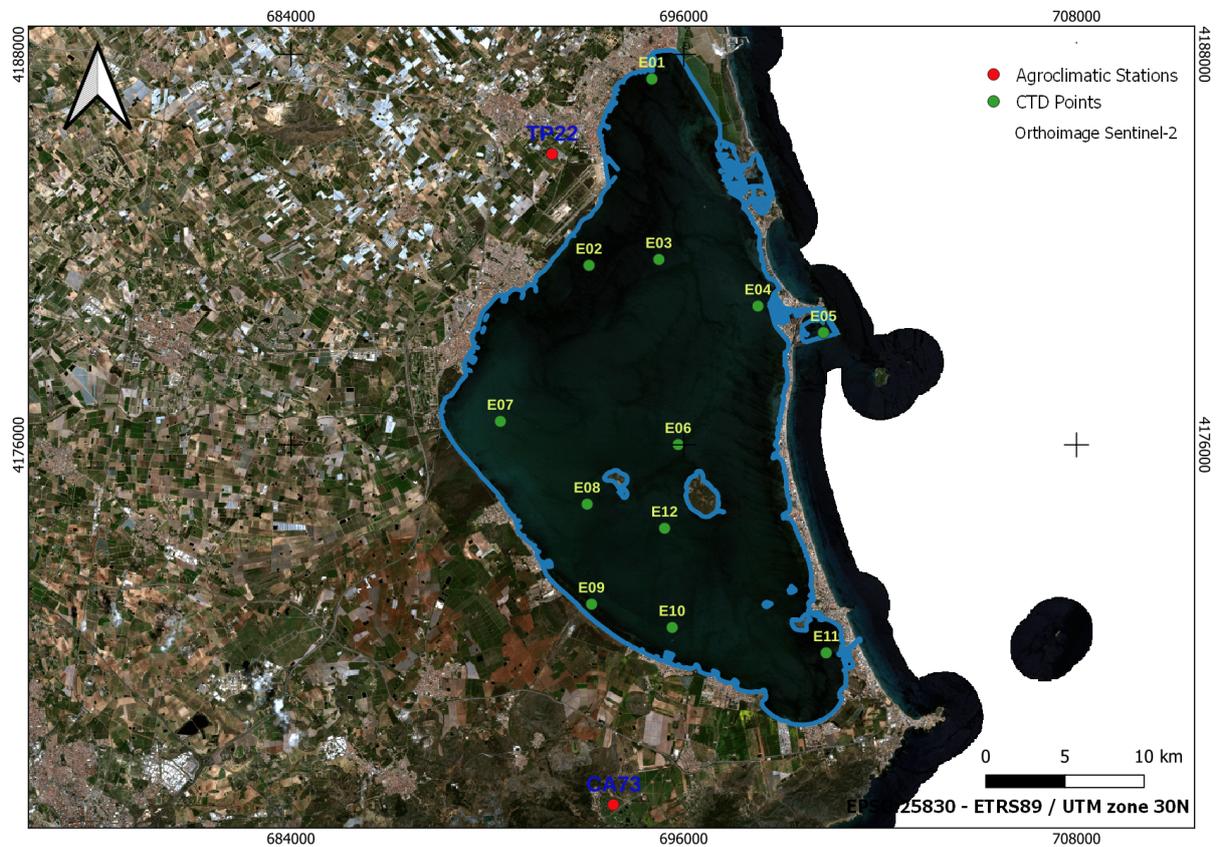


Figure 2. Locations of CTD measuring points and agroclimatic stations.

Once the measurements have been taken, the CTD device generates a file for each measurement point in binary format. These files are converted to plain text format through the software provided by the manufacturer. Each file is processed through a proprietary implementation process, and data curation is performed. This data curation removes readings from the CTD tempering process and readings with incongruent values. Incongruent values are those with negative values of any parameter or values totally out of range, caused, for example, by the CTD hitting the seabed.

The measurements are carried out by different agencies, dependent on the Regional Ministry of Water, Agriculture, Livestock, and Fisheries of the Autonomous Community of the Region of Murcia [14]. The General Directorate of Livestock, Fisheries, and Aquaculture and the Institute of Agricultural and Environment Research and Development of Murcia (IMIDA) [15] are these agencies. These organizations carry out the measurements in 12 sites of the Mar Menor. These measurements have been carried out from April 2017 to the present on a weekly frequency. Regardless of the organization that performs the measurements, they are made with the same equipment and are calibrated with the same procedure.

Although the CTD performs measurements for the entire water column, all the collected values are grouped by date and point through averaging for this work. From April 2017 to the end of April 2023, there were 317 weeks. The dataset used, coming from the CTD, has data for 298 weeks. This means there are 19 weeks for which CTD data are unavailable, usually due to adverse weather conditions.

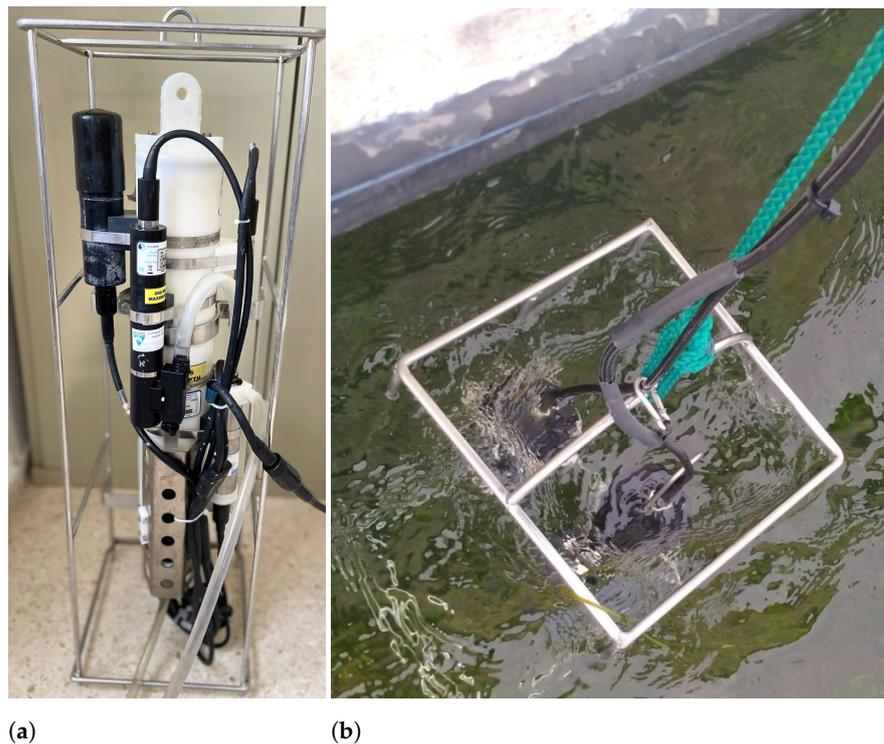


Figure 3. External and at-sea images of the CTD employed. (a) CTD external view. (b) View of CTD taking data in water.

The Figure 4 shows the changes in concentrations and trend of dissolved oxygen in the Mar Menor for the complete dataset. A four-week rolling window is used for the trend. The seasonal characteristics of dissolved oxygen are detected through this image, which experiences a decrease when temperatures rise during spring and summer and a rise in values in autumn and spring when temperatures are lower.

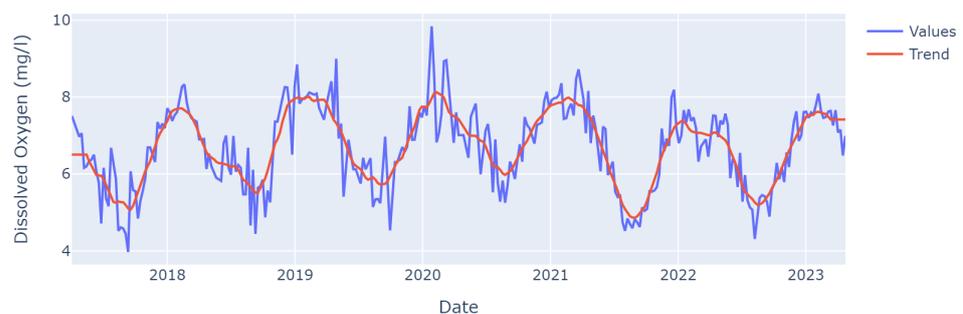


Figure 4. Dissolved Oxygen changes in concentrations and trend in the Mar Menor with a 4-week rolling window.

In the dataset used in this study, the range of dissolved oxygen values ranges from 2.93 and 9.88 mg/L. For chlorophyll a, this range is from 0.11 and 25.15 mg/m³. Chlorophyll a values are generally below 2 mg/m³, and in extraordinary events, such as an Isolated High-Level Depression (DANA) [16] or a prolonged rise in ambient temperature, they rise above 2 mg/m³ to 25 mg/m³.

The Figure 5 illustrates the evolution of chlorophyll a and its trend, combined with two exogenous parameters, such as ambient temperature and precipitation. Looking only at Figure 5a, showing its trend using a four-week rolling window, no seasonality is detected [17]. If we add the accumulated precipitation, as illustrated in Figure 5b, we notice that, after episodes of abnormal rainfall, in subsequent weeks, there is an increase in chlorophyll a, probably due to nutrients [18] and sediments carried by runoff from the

watershed. Similarly, confronting the evolution of chlorophyll a together with temperature, as shown in Figure 5c, a certain correlation is observed between high temperatures and an increase in chlorophyll a levels [19], in the whole time series, except in 2018.

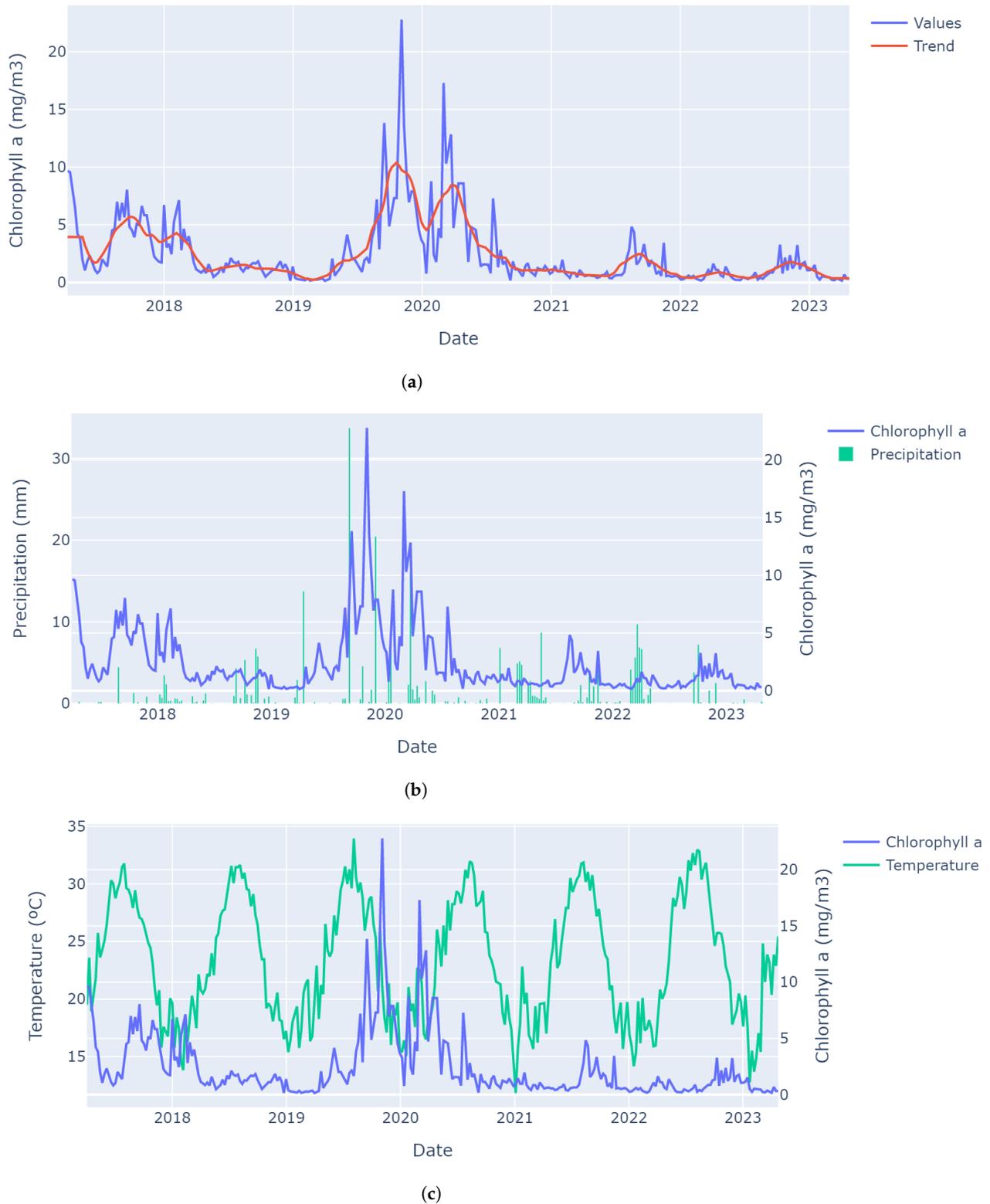


Figure 5. Chlorophyll a evolution with its trend and the ambient temperature and precipitation. (a) Chlorophyll a and trend with 4-week rolling window. (b) Chlorophyll a compared with precipitation. (c) Chlorophyll a compared with ambient temperature.

3.2. Agroclimatic Data

To complete the series of biophysical data, atmospheric data provided by the Murcian Institute for Agricultural and Environmental Research and Development (IMIDA) through its Agricultural Information System of the Region of Murcia (SIAM) [20] have been used. The network consists of 54 automatic agro-meteorological stations, as shown in Figure 6, located in the irrigable areas of the Region of Murcia, from which the stations closest to or with the greatest influence on the Mar Menor lagoon have been selected. Observations from the following variables have been processed: temperature ($^{\circ}\text{C}$), humidity (%), precipitation (mm), wind speed (m/s^2), wind direction ($^{\circ}$), and radiation (w/m^2), obtained every ten minutes from 1 January 2000 to 30 April 2023.

The data are analyzed to verify that the variables are consistent within optimal thresholds. These filters verify whether an observation is within a predetermined range, which can be fixed or dynamic. For data validation, six levels of validation are defined, according to the UNE 500540:2004 standard, "Networks of automatic meteorological stations: Guidelines for the validation of meteorological records from networks of automatic stations. Validation in real time", indicating which level of validation each record reaches.



Figure 6. Type of automatic station used to collect agro-climatic information.

4. Methods

This work aims to obtain a forecast, as close as possible, of the biophysical parameters of chlorophyll a and dissolved oxygen in the Mar Menor lagoon. The forecasts are issued for the 12 different measurement points. The CTD makes four measurements per second as it descends through the water column. Thus, hundreds of records are provided for each measurement point, which varies depending on the depth of each record. In this work, the predictions are formulated for the entire water column, i.e., issuing a single prediction per point. For this purpose, 12 advanced neural networks specially designed for time series forecasting are evaluated. The input data are provided on a weekly frequency, and the forecast horizon is one week.

The methodology used in this work is based on the typical workflow for time series forecasts using neural networks. The main difference between this typical workflow and the one proposed in this work is the introduction of two iterative processes interlaced. The first one iterates among all the defined neural networks, and the second one, for each of the neural networks to be trained, iterates to find the best set of features for each model.

Since we have 12 different measurement points and the same techniques are applied in the measurements, we can consider that we have 12 different time series. Traditionally, each time series is considered independent, so each neural network would be trained as often

as we have time series. This approach is called Local Forecasting Models (LFM). Taking advantage of the fact that the time series are similar, in this work, we have considered these 12 series as a single series, an approach called Global Forecasting Models (GFM) [21,22]. With this approach, we expect that the patterns learned by the model at a particular measurement point can be applied to the rest of the points, which is called cross-learning [23]. In addition, GFMs can be considered multi-task learning paradigms [24], in which a single model is trained to learn multiple tasks, with the individual task being understood as the forecasting of each time series. Among the advantages of multitask models is the ability to understand useful features by observing those that have been useful for other tasks, the ability to learn difficult features by using the datasets of the other time series, or the type of regularization that multitask learning introduces, whereby it forces the model to find a performance that works correctly in all tasks, thus reducing the risk of overfitting. As for the negative aspects of GFM models, the datasets are complex, which implies a higher processing capacity.

Figure 7 shows, in a graphical way, the defined workflow. Each of its main parts is depicted below.

The methodology used in this work is finally transformed into a process integrated into the ML Model Generator module of the Mar Menor Observatory Monitoring System [6]. Each time new CTD data are uploaded, normally on a weekly frequency, the process is run to generate a series of models to obtain a forecast of the state of dissolved oxygen and chlorophyll a in the Mar Menor.

4.1. Base Dataset Creator

As mentioned above, two datasets are available, the first relating to biophysical parameters of the water of the Mar Menor provided by the CTD and the second one containing atmospheric variables provided by the network of agroclimatic stations of the SIAM. The first dataset provides depth, dissolved oxygen, chlorophyll a, salinity, conductivity, and turbidity data. The second dataset provides data on temperature, radiation, evapotranspiration, wind speed and direction, precipitation, and radiation. The CTD time series used in this work is between April 2017 and April 2022. The time series of agroclimatic variables is much longer, exceeding 20 years, so we adjusted the dataset to the CTD data's time-space.

It has been repeatedly mentioned that CTD measurements are carried out at 12 locations in the Mar Menor water body. As for agroclimatic data, these are provided by two stations located in the watershed of the study area. To relate the agroclimatic data with the CTD data, the distance in a straight line of the nearest climatological station concerning each CTD measurement point has been incorporated. This distance is incorporated into the dataset in different units, such as kilometers, land, or nautical miles.

Although CTD data are available at the depth level, as shown in Figure 8, this work focuses on issuing a point forecast for the entire water column. In the dataset, data are grouped by measurement point and date through averaging. This is applied to each of the biophysical parameters captured by the CTD.

The temporal frequency of the CTD dataset is weekly, and that of the agroclimatic dataset is daily. In order to merge both datasets, we resampled to the highest temporal frequency, i.e., weekly frequency, with Monday as the first day. After resampling, the dataset has 3804 records, 317 for each of the 12 points. Of the total number of records, there are 228 with no data, which represents 19 per point. To fill these 228 records, the average of the immediately preceding and following records is used as a method.

Once the complete time series and all the records with values for all the variables are available, the dataset is divided to obtain training and test datasets. When using a GFM model approach, date ranges to carry out splitting. The training dataset has data between 1 April 2017 and 30 April 2022, and the test dataset incorporates data between 1 May 2022 and 30 April 2023. Thus, the training set has 3180 records, and the test set has 624 records, representing an approximate distribution of 83% for training and 17% for testing.

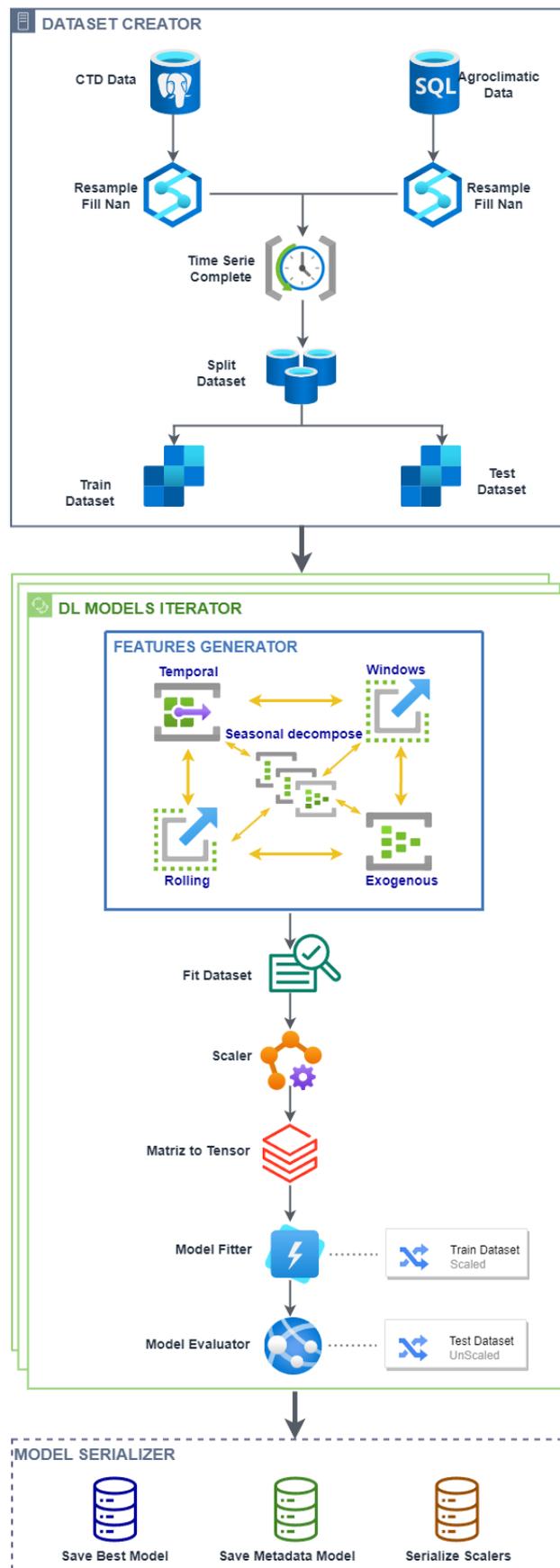
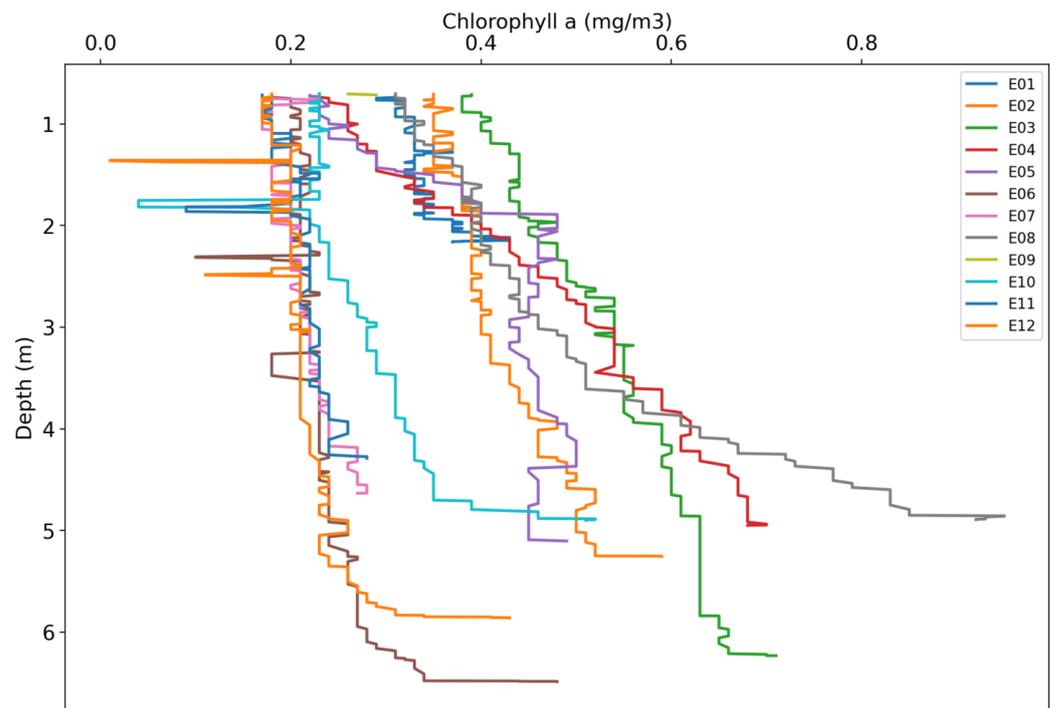
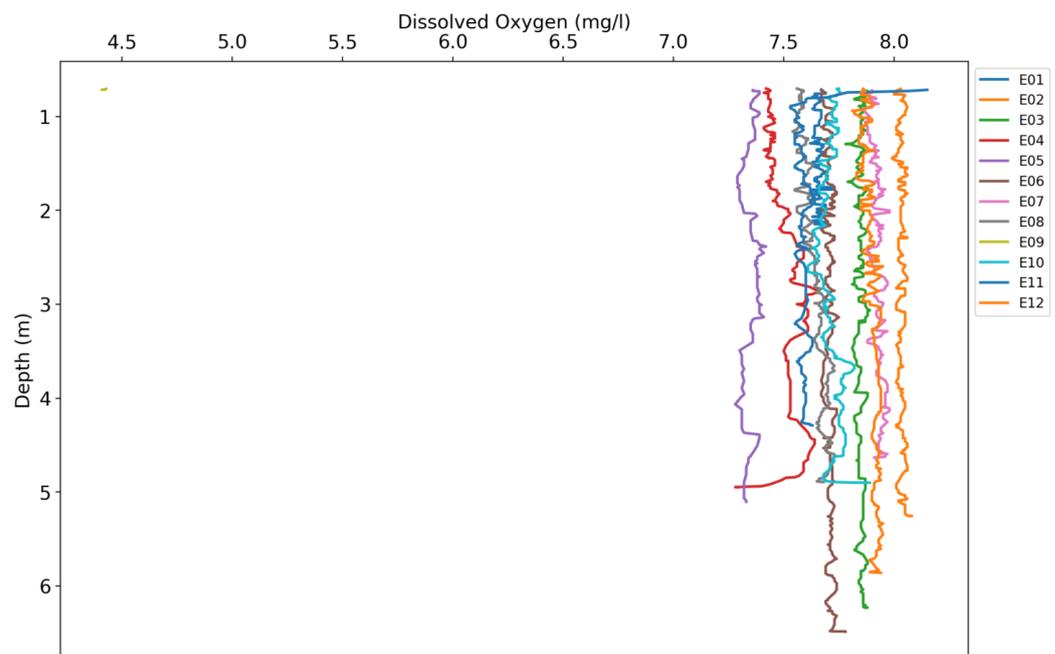


Figure 7. Workflow designed for dissolved oxygen and chlorophyll a forecasting.



(a) CTD chlorophyll a measurement profile.



(b) CTD oxygen measurement profile.

Figure 8. CTD measurement profile.

Finally, by separating the original dataset prior to the feature engineering phase, we ensure that no data leaks occur between the two generated datasets and thus avoid elements that may induce overfitting and optimistic bias on the part of our model.

4.2. Features Generator

Once the training and test datasets have been obtained, this module is responsible for selecting features already existing in these datasets and adding new features derived from the existing ones. The features to be added are grouped by their typology.

The philosophy behind the design of this component is to avoid making a priori assumptions and to find conclusions empirically. Therefore, this module is designed to perform as many combinations of feature sets as possible. It is also designed to be scalable and to be able to add, in a simple way, new groups of features based on their typology.

4.2.1. Temporal Features

These features capture the data's sequentiality and attempt to help the model correctly detect the relationship between data and time. The Features Generator module can generate features based on the calendar, Fourier transform [25], and Radial Basis Functions (RBF) [26]. The latter two allow us to represent time with a continuous cyclic scale.

This feature type has a higher temporal resolution than the dataset. As the base dataset used for this work has a weekly periodicity, the new features generated have a periodicity greater than a week.

The features related to the calendar are the most basic ones, and to obtain them, information is extracted directly from the date, such as year, month, season, and quarter of the year.

Regarding the Fourier transforms, it is generated for the year and month with the sine and cosine. In this way, the seasonality of a time series can be extracted since it is a periodic function.

Finally, as the Fourier transforms, RBF allows time to be represented as a continuous cyclic scale. RBF generates a series of curves that indicate how close we are to a certain time of the year. For example, if we decide to index RBFs at the month level, 12 curves will be represented, with the first one measuring the distance from January. Hence, this curve peaks in the first month and decreases symmetrically as we move away from that month. RBFs are generated at the monthly level, every two months, and every four months.

4.2.2. Window Features

At this point, a predetermined number of previous observations concerning the present are added as features. Thus, if a window of two is indicated for dissolved oxygen, we would have, as features, oxygen_t , oxygen_{t-1} , and oxygen_{t-2} .

4.2.3. Rolling Features

This type of feature is intended to connect the present with an aggregated statistic corresponding to a window in the past. Instead of relying on the observations of the immediately preceding time steps, the last n time steps grouped by a specific statistical function are added to the dataset without using the present observation to avoid data leakage. For example, for a mean grouping function and a window of four previous time steps, a column with the mean of these four-time steps would be added to the dataset. The grouping functions used are mean, median, minimum, maximum, variance, and difference.

4.2.4. Seasonal Decompose Features

Time series primarily combine three components: trend, seasonality, and residuals. Trend refers to the overall movement of the series, seasonality refers to any seasonal pattern found in the series, and residuals are what remains after considering seasonality and trend.

This part is responsible for incorporating, as featured, the three previous components into the dataset, both in additive and multiplicative modes [27].

4.2.5. Exogenous Variables

Exogenous variables are those not influenced by other variables and on which the output variable depends [28]. Making this formal definition more flexible, in this work, a

priori, we consider as exogenous, except for the target variable, all the parameters provided by the CTD sensors, and all the variables provided by the agroclimatic stations.

Introducing this type of feature to the dataset aims to detect temporal cross-correlations [29] between the time series of the endogenous and exogenous variables during the training of our model and thus achieves better performance.

When evaluating the effect of the exogenous variables on the performance of each tested model, eight batches are created to obtain the best possible combination of features. It should be noted that the features derived from the target variable are already incorporated into the dataset:

- CTDOBJ_CTDALL: All CTD variables are incorporated.
- CTDOBJ_CTDCOR: The exogenous variables of the CTD most correlated with the target variable are incorporated.
- CTDOBJ_SIAMALL: All the features provided by the SIAM agroclimatic stations are added.
- CTDOBJ_SIAMCOR: The most correlated variables of the agroclimatic seasons are added.
- CTDALL_SIAMALL: All available exogenous variables are added.
- CTDALL_SIAMCOR: All the exogenous variables of the CTD and the most correlated variables of the agroclimatic stations are incorporated.
- CTDCOR_SIAMALL: All the exogenous variables of the SIAM agroclimatic stations and the most highly correlated variables of the CTD are incorporated.
- CTDCOR_SIAMCOR: The most highly correlated variables from both the CTD and the agroclimatic stations are added.

4.3. DL Models Iterator

Once the Features Generator module generates a dataset for model training, the DL Models Iterator component takes care of the following aspects:

1. **Scaler:** This element is responsible for applying a transformation to the data so that they are in the same unit of measurement and there is no great difference between the ranges of values of the different variables. In this work, we experiment with normalization (transforming the data from their original values to a range between 0 and 1) and standardization of the data (transforming the data so that the mean of the observed values is equal to 0 and their standard deviation is equal to 1) [30].
2. **Matrix to Tensor Converter:** The Features Generator module forms a two-dimensional matrix in which the first dimension represents the different rows, and the second represents each feature or column. In order to supply the dataset to the neural network and make it capable of working with it, the time series data must be reformatted to supervised learning. For this purpose, the sliding window method is employed, whereby the previous time steps are used to forecast the next time step. Two parameters are required for this, the window size and the number of future time steps to be predicted. The window size refers to the number of previous steps needed to forecast the future horizon. In the case of the matrix of independent variables, this component will generate a three-dimensional tensor in which the first dimension will be the number of rows, the second will be the window size or sequence, and the third will be the independent features of the dataset. The matrix with the target variable differs a little from the previous one; in this case, the first two dimensions coincide, but the third dimension represents the forecast for each of the future time steps. The component is designed to parameterize the input window size between the range of values 2, 4, and 6. The output window size is fixed to a future time step with no possibility of being modified. Figure 9 explains the task of this component graphically.
3. **Model Fitter:** A model is generated based on the predefined neural networks, which will be describe below (Section 4.3.1). For all generated models, the Mean Squared Error (MSE) function is defined as a loss function, as optimizer Adam with a learning rate of 0.001, and as metrics with the MSE, the Root Mean Squared Error (RMSE),

the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R2) [31,32]. Regarding the number of epochs, for dissolved oxygen, a maximum number of epochs of 1000 and a patience of 50 is defined, i.e., the model will be adjusted up to 1000 epochs, but if the loss function does not improve during 50 epochs, the training will be terminated. In the case of chlorophyll a, the maximum number of epochs is 2000, and the patient is set to 100. The difference in epochs between dissolved oxygen and chlorophyll a is due to the absence of seasonality, which can be seen in the Figure 5a. Finally, the same seed is used for all models.

4. Model Evaluator: The component is responsible for evaluating the model's performance through the metrics defined in the previous step and the test dataset.

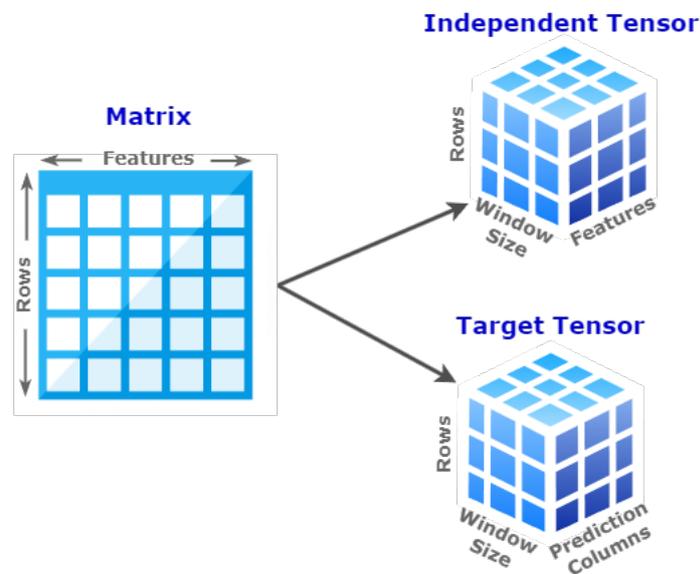


Figure 9. Matrix-to-tensor converter

4.3.1. Neural Networks Architectures

The DL Models Iterator module is executed as many times as the neural networks have been defined. Subsequently, each neural network is adjusted as many times as the datasets generated by the Feature Generator component. When establishing the neural networks to be evaluated, a series of base architectures are defined and then combined as layers. Each of the base architectures is listed below:

- Convolutional Neuronal Network (CNN): This type of network was initially conceived for computer vision applications. Among its advantages is its ability to handle large datasets. Numerous papers in the current literature show that this is a good architecture if properly combined with others specially designed for time series forecasting [33]. Combined with an LSTM architecture, CNN is responsible for extracting the features and LSTM for learning the extracted features over the time series.
- Bidirectional LSTM (Bi-LSTM): LSTM neural networks [34] are specifically designed to process sequential data, such as time series. Based on these, bidirectional LSTMs work by processing the input data in two directions, from the beginning to the end and from the end to the beginning, allowing the model to capture the past and future context of the input data. The output of this neural network is generated by concatenating the outputs of the forward and backward LSTM layers. This fact allows the model to take advantage of information from both directions. Several papers in the literature show that BiLSTM models offer better predictions compared to regular models based on LSTM [35–37]. As disadvantages of BiLSTM concerning LSTM, it is worth noting that

BiLSTM is a slower model due to its greater complexity and requires more training time to reach equilibrium.

- Seq2Seq: Based on recurrent neural networks (RRN), this model processes input sequences and generates output sequences. It is composed of two components: an encoder and a decoder [38]. The encoder processes the input sequence and transforms it into a feature vector, which is used as input for the decoder. The decoder generates the output sequence step by step, taking the feature vector and the previously generated sequence as input. They are capable of handling variable-length sequences and capturing long-term dependencies.
- Mixture Density Network (MDN): Oriented to time series forecasting, they can handle large datasets and detect changes in trend and variability in the data [39,40].
- Temporal Convolutional Network (TCN): It is a variant of the CNN architecture, specially designed for time series forecasting [41–43]. Generally speaking, it has a longer-term memory than recurrent architectures and a better performance than LSTM on large time series.
- Attention: It is proposed as an evolution of the encoder–decoder architecture, which aims to avoid forgetting the first parts of the sequences, especially when faced with large data sequences [44,45]. To do this, the attention mechanism assigns different importance to different input sequence elements and pays more attention to the most relevant inputs. In this way, it can remember all the inputs provided. There are different attention mechanisms; we use the so-called Self-Attention and Multi-Head Attention in this work. The main difference between the two is that in the former, only one attention mechanism is incorporated, and in the latter, there are several attention mechanisms, thus increasing the ability to find the most relevant inputs.
- Transformer: It was created as a model of natural language processing [46]. It is based on the Attention architecture, which allows the network to process variable-length input sequences efficiently, thanks to the use of multiple attention mechanisms, which allow the network to consider multiple elements of the input sequence simultaneously. Although born as a methodology for natural language processing, they can not only be used for natural language processing but have also proven effective for a wide range of other tasks involving data sequences, such as time series forecasting [47]. This architecture is changing the paradigm of artificial intelligence.
- Time2Vec: Its authors acknowledge that the ultimate goal was to develop a general-purpose model-agnostic representation for time [48], which can potentially be used in any architecture [49]. It is not created as a new model for time series analyses but rather aims to provide a representation of time in the form of a vector embedded in order to automate the feature engineering process and model time in a better way.

Combining the architectures listed above or using them individually, the DL Models Iterator component generates up to 12 different neural networks. In addition, this component is based on the scalability principle, which makes it easy to incorporate new neural networks. In this work, the following architectures are evaluated:

1. CNN-BiLSTM
2. Seq2Seq
3. BiLSTM
4. MDN-BiLSTM
5. TCN-BiLSTM
6. CNN-BiLSTM-SelfAttention (CNN-BiLSTM-Att)
7. SelfAttention-Seq2Seq (Seq2Seq-Att)
8. BiLSTM-SelfAttention (BiLSTM-Att)
9. CNN-BiLSTM-MultiHeadAttention (CNN-BiLSTM-MultiHead)
10. BiLSTM-MultiHeadAttention (BiLSTM-MultiHead)
11. Time2Vec-BiLSTM
12. Time2Vec-Transformer

4.4. Model Serializer

This component is responsible for storing on disk the necessary information to reuse the model later and the related metadata. It stores the model in h5 format, the generated scalars, the obtained metrics, and the dataset needed to forecast the future time step.

5. Results

This section presents the metrics obtained by each model, both for dissolved oxygen and chlorophyll a, and the results obtained are discussed in detail. In addition, a comparison between the forecast of the next time horizon out of the dataset used and the real values for each of the 12 measurement points is included.

5.1. Dissolved Oxygen Forecasting

In this section, the results related to dissolved oxygen forecasts are presented. First, the best model obtained is discussed, including temporal features and those related to dissolved oxygen itself, and then the performance of the models evaluated, including exogenous variables, is shown. In addition, the features that contribute more positively to a better performance of the models are discussed, and the forecast generated for the first date out of the dataset is confronted with the real data for that date.

5.1.1. Endogenous Variable Metrics for Dissolved Oxygen Forecast

Table 1 shows the metrics obtained by the 12 models evaluated using only temporal features and features derived from dissolved oxygen itself.

Table 1. Model metrics with temporal and dissolved oxygen-derived features.

| Model | MAE | MSE | RMSE | MAPE | R2 | Epochs | Best Epoch | Time (s) |
|----------------------|-------|-------|-------|-------|-------|--------|------------|----------|
| CNN-BiLSTM | 0.535 | 0.539 | 0.734 | 10.50 | 0.294 | 62 | 12 | 30 |
| Seq2Seq | 0.380 | 0.285 | 0.534 | 7.38 | 0.649 | 196 | 146 | 169 |
| BiLSTM | 0.405 | 0.314 | 0.560 | 8.27 | 0.594 | 109 | 59 | 104 |
| MDN-BiLSTM | 0.216 | 0.076 | 0.259 | 3.30 | 0.922 | 233 | 183 | 512 |
| TCN-BiLSTM | 0.174 | 0.067 | 0.259 | 3.39 | 0.902 | 141 | 91 | 490 |
| CNN-BiLSTM-Att | 0.272 | 0.124 | 0.352 | 4.95 | 0.875 | 445 | 395 | 980 |
| Seq2Seq-Att | 0.342 | 0.232 | 0.482 | 6.69 | 0.734 | 449 | 399 | 2430 |
| BiLSTM-Att | 0.153 | 0.044 | 0.211 | 3.05 | 0.926 | 262 | 212 | 306 |
| CNN-BiLSTM-MultiHead | 0.208 | 0.077 | 0.278 | 3.95 | 0.927 | 942 | 892 | 2043 |
| BiLSTM-MultiHead | 0.207 | 0.079 | 0.282 | 4.33 | 0.892 | 113 | 62 | 345 |
| Time2Vec-BiLSTM | 0.100 | 0.021 | 0.145 | 1.50 | 0.978 | 623 | 573 | 792 |
| Time2Vec-Transformer | 0.146 | 0.047 | 0.218 | 2.17 | 0.951 | 576 | 526 | 360 |

Table 1 shows that the architectures based on Time2Vec have the best metrics. Time2Vec-BiLSTM has better metrics than the Time2Vec-Transformer, but the adjustment time is almost twice as long. Other advanced structures, such as MDN and TCN with BiLSTM or BiLSTM with Attention, have obtained remarkable results. As a general rule, adding an attention mechanism to an architecture makes it experience an improvement with the dataset used in this work.

A plot of the predicted and test values for point E05 of the Time2Vec-BiLSTM model is shown in Figure 10:

Concerning the features used, the temporal features that performed best were those based on the Fourier Transform. They performed very similarly to those based on the Radial Basis Function, although slightly better. The temporal features based on the calendar obtained a poorer performance. The models based on the Time2Vec architecture require special mention since they can detect temporal relationships between data without manually inputting additional temporal features. Table 2 shows the metrics obtained by the two models without manually inputting any temporal features. When compared to the results of Table 1, similar metrics are observed but with lower execution times.

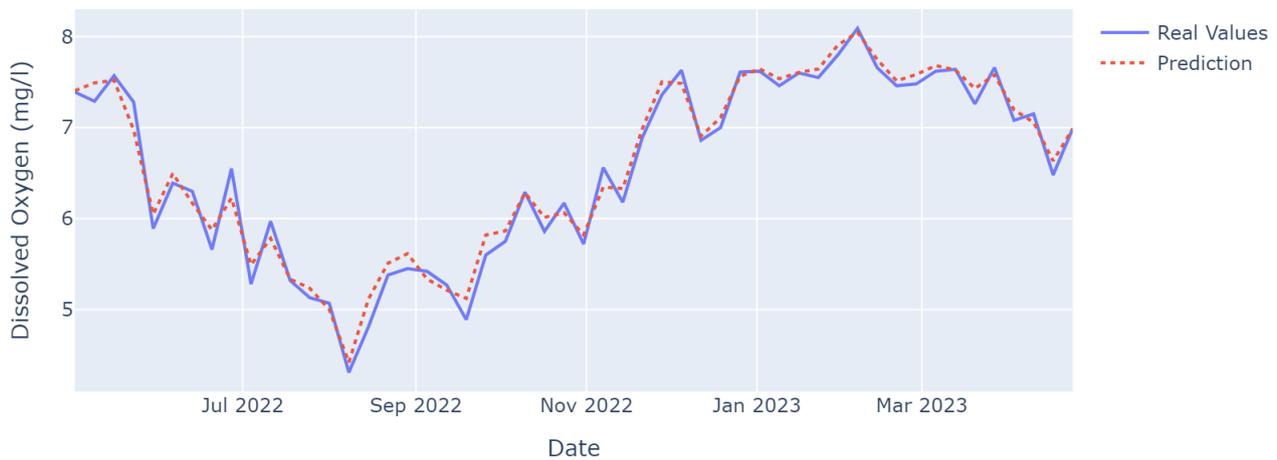


Figure 10. Dissolved oxygen forecast of the Time2Vec-BiLSTM model in the test dataset.

Table 2. Time2Vec model metrics without manually inputting additional temporal features.

| Model | MAE | MSE | RMSE | MAPE | R2 | Epochs | Best Epoch | Time (s) |
|----------------------|-------|-------|-------|------|--------|--------|------------|----------|
| Time2Vec-BiLSTM | 0.125 | 0.029 | 0.170 | 1.88 | 0.970 | 393 | 342 | 484 |
| Time2Vec-Transformer | 0.167 | 0.049 | 0.222 | 2.54 | 0.9471 | 340 | 289 | 219 |

As for the dissolved oxygen-derived features, the Window and Rolling features were used in all models’ best versions. Of the Seasonal Decompose type features, the so-called trend and residuals, in additive mode, were used in 100% of the models evaluated. In contrast, the seasonality component was used only in the models based on the Time2Vec architecture.

Concerning data transformation, as previously mentioned, normalization and standardization are applied in this work so that the data are scaled in the same range. All models, except those based on the Time2Vec, MDN-BiLSTM, and BiLSTM-MultiHead architectures, obtain their best results using normalization as the transformation. In the case of MDN-BiLSTM and BiLSTM-MutiHead, standardization yields metrics to be considered but inferior when normalizing the data. In the case of Time2Vec-based models, applying the two transformation techniques generates similar results. A comparison of the metrics obtained with these two transformations for the best-performing models is shown in Table 3.

Table 3. Standardization vs. Normalization through dissolved oxygen forecasting metrics.

| Scaler | Time2Vec-BiLSTM | | | | | Time2Vec-Transformer | | | | |
|-----------------|-----------------|-------|-------|------|-------|----------------------|-------|-------|------|-------|
| | MAE | MSE | RMSE | MAPE | R2 | MAE | MSE | RMSE | MAPE | R2 |
| Normalization | 0.100 | 0.021 | 0.145 | 1.50 | 0.978 | 0.171 | 0.052 | 0.228 | 2.54 | 0.947 |
| Standardization | 0.151 | 0.037 | 0.193 | 1.20 | 0.971 | 0.146 | 0.047 | 0.218 | 2.17 | 0.951 |

Finally, concerning the input data window, windows of sizes two and four obtain similar results, although those corresponding to a window of four are slightly better. As for the size six window, it generates inferior metrics.

5.1.2. Exogenous Variables Metrics

This work has two data sources that can provide us with exogenous variables. The first is the CTD itself, which provides us with data on temperature, salinity, and turbidity in addition to dissolved oxygen and chlorophyll a parameters. On the other hand, we have the SIAM agroclimatic stations, which, thanks to the sensors they incorporate, provide us with data on temperature, radiation, evapotranspiration, wind speed, and direction, precipitation, and radiation.

At this point, we evaluate the effect of the exogenous variables in each model. To do this, sets of these exogenous variables are added to the dataset obtained in the previous section (Section 5.1.1). For this, we consider the correlation matrix between the target and exogenous variables. Table 4 shows the dissolved oxygen correlation matrix.

Table 4. Correlation matrix of dissolved oxygen with available exogenous variables.

| Parameter | Data Source | Correlation |
|---------------------|-------------|-------------|
| Dissolved Oxygen | CTD | 1.000 |
| Temperature | CTD | −0.734 |
| Salinity | CTD | −0.314 |
| Chlorophyll a | CTD | 0.231 |
| Turbidity | CTD | −0.106 |
| Ambient Temperature | SIAM | −0.703 |
| Evapotranspiration | SIAM | −0.495 |
| Radiation | SIAM | −0.408 |
| Precipitation | SIAM | 0.042 |
| Wind Speed | SIAM | 0.048 |
| Relative Humidity | SIAM | 0.013 |

Taking into account the above, the features included in the sub-lots of exogenous variables (Section 4.2.5) based on the correlation of features are listed below:

- CTDCOR: Temperature, Salinity, and Chlorophyll a
- SIAMCOR: Ambient Temperature, Evapotranspiration, and Radiation

Table 5 shows the metrics obtained by the Time2Vec-BiLSTM model for each of the eight execution batches with exogenous variables.

Table 5. Metrics obtained by Time2Vec-BiLSTM with lots of exogenous variables.

| Exogenous Lot | MAE | MSE | RMSE | MAPE | R2 | Epochs | Best Epoch | Time (s) |
|----------------|-------|-------|-------|------|-------|--------|------------|----------|
| CTDOBJ_CTDALL | 0.124 | 0.031 | 0.176 | 2.00 | 0.968 | 549 | 499 | 689 |
| CTDOBJ_CTDCOR | 0.134 | 0.030 | 0.174 | 1.92 | 0.969 | 419 | 369 | 511 |
| CTDOBJ_SIAMALL | 0.127 | 0.075 | 0.275 | 3.31 | 0.923 | 407 | 357 | 501 |
| CTDOBJ_SIAMCOR | 0.218 | 0.035 | 0.189 | 2.12 | 0.963 | 752 | 702 | 1044 |
| CTDALL_SIAMALL | 0.142 | 0.097 | 0.312 | 3.92 | 0.900 | 488 | 438 | 605 |
| CTDALL_SIAMCOR | 0.259 | 0.073 | 0.271 | 3.13 | 0.925 | 228 | 178 | 293 |
| CTDCOR_SIAMALL | 0.210 | 0.050 | 0.225 | 2.56 | 0.948 | 459 | 409 | 588 |
| CTDCOR_SIAMCOR | 0.175 | 0.050 | 0.223 | 2.67 | 0.949 | 469 | 419 | 594 |

As can be observed, incorporating lots of exogenous variables does not improve the results obtained so far. Considering the eight lots, the variability of the MAPE is between 3.92 and 1.92, and the R2 is between 0.900 and 0.969. If we only look at the lots incorporating correlated exogenous variables, we find an MAPE between 2.67 and 1.92 and an R2 between 0.949 and 0.969. Therefore, the Time2Vec-BiLSTM model, when working with exogenous variables, performs better with those correlated.

Table 6 shows the metrics obtained by the Time2Vec-Transformer model for each of the eight execution lots with exogenous variables.

We can observe that considering the totality of the lots, the variability of the MAPE is between 2.68 and 2.27, and the R2 is between 0.932 and 0.951. At this point, the Time2Vec-Transformer model presents greater stability when introducing exogenous variables that are not correlated with the target variables. In addition, the adjustment time is still considerably shorter.

The rest of the models show a considerable deterioration when adding the available exogenous variables to the dataset, except the MDN-BiLSTM model, which moves in a somewhat worse differential than those presented by the Time2Vec architectures.

Table 6. Metrics obtained by Time2Vec-Transformer with lots of exogenous variables.

| Exogenous Lot | MAE | MSE | RMSE | MAPE | R2 | Epochs | Best Epoch | Time (s) |
|----------------|-------|-------|-------|------|-------|--------|------------|----------|
| CTDOBJ_CTDALL | 0.156 | 0.052 | 0.228 | 2.33 | 0.947 | 668 | 618 | 418 |
| CTDOBJ_CTDCOR | 0.154 | 0.050 | 0.223 | 2.31 | 0.949 | 515 | 465 | 329 |
| CTDOBJ_SIAMALL | 0.181 | 0.059 | 0.243 | 2.68 | 0.939 | 761 | 711 | 471 |
| CTDOBJ_SIAMCOR | 0.151 | 0.047 | 0.219 | 2.27 | 0.951 | 735 | 685 | 448 |
| CTDALL_SIAMALL | 0.166 | 0.058 | 0.242 | 2.50 | 0.940 | 738 | 688 | 466 |
| CTDALL_SIAMCOR | 0.153 | 0.052 | 0.228 | 2.29 | 0.947 | 658 | 608 | 418 |
| CTDCOR_SIAMALL | 0.163 | 0.053 | 0.231 | 2.44 | 0.945 | 650 | 600 | 406 |
| CTDCOR_SIAMCOR | 0.171 | 0.067 | 0.258 | 2.58 | 0.932 | 469 | 419 | 304 |

5.1.3. Dissolved Oxygen Metrics vs. Real Values

Since, at the time of writing this paper, real data were already available for the 12 measurement points for the date following the last date of the dataset used in this paper, we are in a position to predict the first week of May 2023 and compare the values obtained with the real values.

Table 7 shows the relation of values predicted by the best versions of the Time2Vec-BiLSTM and Time2Vec-Transformer models for the 12 measurement points together with the actual values for the prediction date.

Table 7. Dissolved oxygen values predicted by Time2Vec models for the first week of May 2023, together with actual values.

| Point | Real Value | Time2Vec-BiLSTM Forecast | Time2Vec-Transformer Forecast |
|-------|------------|--------------------------|-------------------------------|
| E01 | 5.75 | 5.66 | 6.01 |
| E02 | 6.84 | 6.96 | 6.91 |
| E03 | 6.70 | 6.80 | 6.77 |
| E04 | 6.57 | 6.64 | 6.61 |
| E05 | 6.84 | 7.02 | 6.88 |
| E06 | 6.83 | 6.89 | 6.87 |
| E07 | 7.13 | 7.25 | 7.22 |
| E08 | 7.07 | 7.12 | 7.16 |
| E09 | 7.32 | 7.41 | 7.48 |
| E10 | 7.36 | 7.32 | 7.44 |
| E11 | 6.92 | 7.05 | 7.01 |
| E12 | 6.66 | 6.69 | 6.68 |

Table 8 shows the metrics associated with the forecasts of the previous table. It can be seen that the metrics follow the same line as those obtained in the model testing process.

Table 8. Dissolved oxygen forecast metrics for the first week of May 2023 for Time2Vec-based architectures.

| Model | MAE | MSE | RMSE | MAPE | R2 |
|----------------------|-------|-------|-------|------|-------|
| Time2Vec-BiLSTM | 0.093 | 0.010 | 0.102 | 1.37 | 0.935 |
| Time2Vec-Transformer | 0.102 | 0.012 | 0.113 | 1.54 | 0.920 |

5.1.4. Dissolved Oxygen Forecast Summary

To summarize, in this work regarding dissolved oxygen, after training the 12 different models, we obtain that the models based on the Time2Vec architecture obtain the best performance, with an MAPE and R2 of 2.17 and 0.951 for Time2Vec-Transformer, and 1.50 and 0.978 for Time2Vec-BiLSTM. Other models based on Bi-LSTM and incorporating attention mechanisms have obtained good metrics with the dataset used. When introducing exogenous variables, the architectures based on Time2Vec show the best results. Although they do not improve the model metrics, at least they do not worsen them considerably, as

with the other models. Finally, a prognostic test is performed to forecast the next time step outside the dataset used for training the models, resulting in na MAPE and R2 OF 1.37 and 0.935 for the Time2Vec-BiLSTM model and 1.54 and 0.920 for Time2Vec-Transformer.

5.2. Chlorophyll a Forecasting

Following the methodology of the previous section, this section presents the detailed results obtained by all the models evaluated for the chlorophyll a forecast. Their strengths and weaknesses are highlighted for the best-performing models, and a chlorophyll a forecast is carried out for the first date beyond the dataset used.

5.2.1. Endogenous Variable Metrics for Chlorophyll a Forecast

Table 9 shows the metrics obtained by the 12 models evaluated. The dataset used in this phase only contains temporal features and features derived from chlorophyll a itself.

Table 9. Model metrics with temporal features and chlorophyll a derived.

| Model | MAE | MSE | RMSE | MAPE | R2 | Epochs | Best Epoch | Time (s) |
|----------------------|-------|-------|-------|-------|-------|--------|------------|----------|
| CNN-BiLSTM | 0.252 | 0.191 | 0.437 | 57.66 | 0.238 | 204 | 104 | 247 |
| Seq2Seq | 0.139 | 0.055 | 0.236 | 43.49 | 0.639 | 637 | 537 | 502 |
| BiLSTM | 0.131 | 0.042 | 0.205 | 33.04 | 0.794 | 598 | 498 | 579 |
| MDN-BiLSTM | 0.089 | 0.020 | 0.144 | 13.38 | 0.932 | 793 | 693 | 339 |
| TCN-BiLSTM | 0.122 | 0.036 | 0.191 | 34.23 | 0.820 | 195 | 95 | 673 |
| CNN-BiLSTM-Att | 0.130 | 0.031 | 0.178 | 25.52 | 0.797 | 1234 | 1134 | 569 |
| Seq2Seq-Att | 0.252 | 0.154 | 0.393 | 41.68 | 0.498 | 525 | 425 | 1918 |
| BiLSTM-Att | 0.113 | 0.025 | 0.159 | 21.86 | 0.817 | 1284 | 1184 | 1379 |
| CNN-BiLSTM-MultiHead | 0.162 | 0.114 | 0.338 | 55.45 | 0.259 | 542 | 442 | 1737 |
| BiLSTM-MultiHead | 0.128 | 0.031 | 0.176 | 24.01 | 0.798 | 255 | 155 | 154 |
| Time2Vec-BiLSTM | 0.086 | 0.018 | 0.135 | 10.51 | 0.940 | 797 | 697 | 999 |
| Time2Vec-Transformer | 0.093 | 0.020 | 0.145 | 12.36 | 0.929 | 1370 | 1270 | 817 |

It can be seen that the models based on Time2Vec again show the best metrics. Again, Time2Vec-BiLSTM shows better metrics than Time2Vec-Transformer. The MDN-BiLSTM model also shows good results. In general, the rest of the models obtain worse metrics than in the case of dissolved oxygen, which is to be expected due to the non-seasonality of the data, since no regular fluctuations or changes are observed over time.

For illustrative purposes, Figure 11 shows a plot of the test and predicted values for chlorophyll a from the Time2Vec-Transformer model:

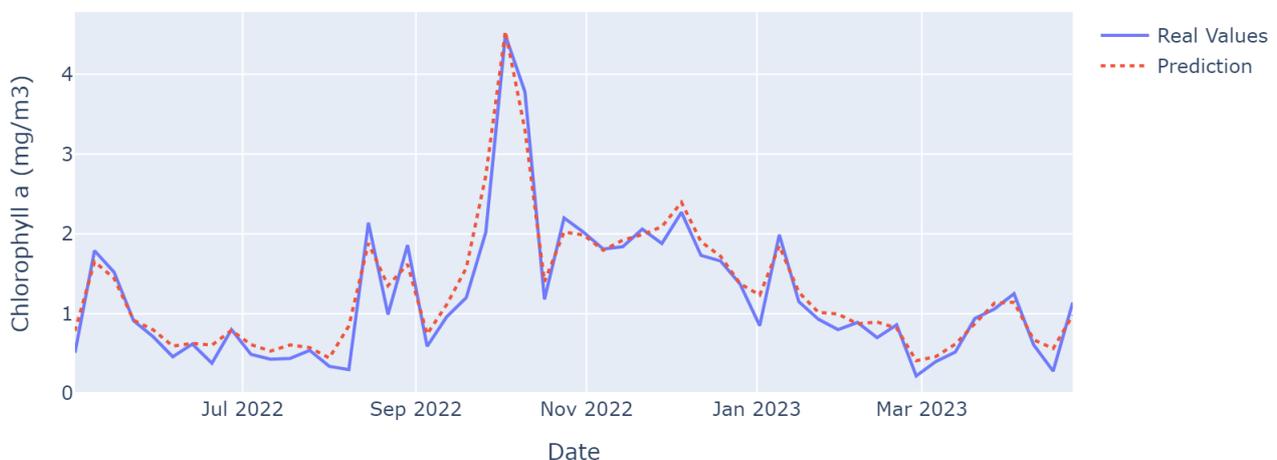


Figure 11. Chlorophyll a forecast of the Time2Vec-Transformer model in the test dataset.

Regarding the features used, focusing on the temporal features, all the models have shown a better performance than those based on the Radial Basis Function, with frequencies

of 2 or 4 months. The temporal features based on the Fourier Transform show inferior performances, and those based on the calendar fail to obtain acceptable metrics. As with dissolved oxygen, Time2Vec-based architectures can also identify the temporal component of the data without the need for manual input of such features. There is less difference in metrics for chlorophyll a than for dissolved oxygen (MAPE of 10.42 and R2 of 0.933 for TimeVec-BiLSTM and MAPE of 12.36 and R2 of 0.934 for TimeVec-Transformer).

As for the chlorophyll a-derived features, the same applies to dissolved oxygen, i.e., as a general rule, all of the Window and Rolling features are used. Of the Seasonal Decompose features, the so-called residuals are discarded.

Concerning data transformation, the models based on the Time2Vec architecture and the MDN-BiLSTM model generate very similar results when applying both transformations. The rest of the models obtain, by far, their best results by applying normalization to the data.

Finally, as with dissolved oxygen, the input windows with a four-date sequence offered slightly higher values than the two-date sequences and higher values than the six-date sequences.

5.2.2. Exogenous Variables Metrics

Table 10 shows the correlation matrix of the parameters used in this work for chlorophyll a.

Table 10. Correlation matrix of chlorophyll a with available exogenous variables.

| Parameter | Data Source | Correlation |
|---------------------|-------------|-------------|
| Chlorophyll a | CTD | 1.000 |
| Dissolved Oxygen | CTD | 0.231 |
| Temperature | CTD | 0.405 |
| Salinity | CTD | −0.174 |
| Turbidity | CTD | 0.617 |
| Ambient Temperature | SIAM | −0.703 |
| Evapotranspiration | SIAM | −0.115 |
| Radiation | SIAM | −0.325 |
| Precipitation | SIAM | 0.397 |
| Wind Speed | SIAM | −0.019 |
| Relative Humidity | SIAM | 0.004 |

According to the above, the exogenous variables included in the sub-lots based on the correlation of features are listed below:

- CTDCOR: Temperature, Turbidity, and Chlorophyll a
- SIAMCOR: Ambient Temperature, Precipitation, and Radiation

Table 11 shows the MAPE and R2 metrics obtained by the three best models.

For all three models, we observed high variability in the metrics when adding exogenous variables. None of them improve with the addition of exogenous variables. The one that offers the best metrics is Time2Vec-BiLSTM. As a remarkable aspect, the dataset with all the variables obtains an R2 of 0.059, lower than the best score obtained by this model, or the combination of all the CTD variables and those most correlated with chlorophyll a obtains an R2 differential of 0.046 concerning the score obtained by this model with the exogenous variables.

5.2.3. Chlorophyll a Metrics vs. Real Values

As was done for dissolved oxygen, Table 12 shows the chlorophyll a forecast made by the three best models for the following week not included in the dataset and for the 12 measurement points.

It can be seen how at points where one model deviates, another model gives a tighter forecast. For example, at point E08, the models based on the Time2Vec architecture present an MAE of up to 0.16, and the MDN-BiLSTM model presents an MAE of a few thousandths.

Table 13 shows the metrics associated with the forecasts of the previous table. It can be seen that the metrics follow the same line as those obtained in the model testing process.

Table 11. Metrics obtained by the Time2Vec type architectures and the MDN-BiLSTM model with lots of exogenous variables for chlorophyll a.

| Exogenous Lot | Time2Vec-BiLSTM | | Time2Vec-Transformer | | MDN-BiLSTM | |
|----------------|-----------------|-------|----------------------|-------|------------|-------|
| | MAPE | R2 | MAPE | R2 | MAPE | R2 |
| CTDOBJ_CTDALL | 13.17 | 0.915 | 13.99 | 0.922 | 23.86 | 0.880 |
| CTDOBJ_CTDCOR | 21.02 | 0.888 | 20.60 | 0.877 | 13.45 | 0.928 |
| CTDOBJ_SIAMALL | 18.49 | 0.882 | 29.78 | 0.814 | 20.96 | 0.887 |
| CTDOBJ_SIAMCOR | 23.73 | 0.854 | 30.40 | 0.785 | 18.52 | 0.901 |
| CTDALL_SIAMALL | 19.98 | 0.881 | 36.10 | 0.757 | 35.24 | 0.812 |
| CTDALL_SIAMCOR | 18.95 | 0.894 | 30.37 | 0.784 | 34.87 | 0.756 |
| CTDCOR_SIAMALL | 19.67 | 0.861 | 38.64 | 0.676 | 47.24 | 0.659 |
| CTDCOR_SIAMCOR | 26.00 | 0.832 | 29.21 | 0.818 | 32.76 | 0.831 |

Table 12. The Time2Vec and MDN-BiLSTM models forecast chlorophyll a values for the first week of May 2023.

| Point | Real Value | Time2Vec-BiLSTM Forecast | Time2Vec-Transformer Forecast | MDN-BiLSTM Forecast |
|-------|------------|--------------------------|-------------------------------|---------------------|
| E01 | 0.35 | 0.25 | 0.25 | 0.28 |
| E02 | 0.23 | 0.20 | 0.21 | 0.25 |
| E03 | 0.32 | 0.28 | 0.32 | 0.31 |
| E04 | 0.34 | 0.32 | 0.33 | 0.36 |
| E05 | 0.24 | 0.20 | 0.31 | 0.26 |
| E06 | 0.58 | 0.45 | 0.48 | 0.59 |
| E07 | 0.86 | 0.78 | 0.74 | 0.93 |
| E08 | 0.95 | 0.79 | 0.82 | 0.95 |
| E09 | 0.39 | 0.36 | 0.38 | 0.47 |
| E10 | 0.19 | 0.18 | 0.22 | 0.29 |
| E11 | 0.38 | 0.32 | 0.37 | 0.46 |
| E12 | 0.25 | 0.26 | 0.34 | 0.35 |

Table 13. Chlorophyll a forecast metrics for the first week of May 2023 from Time2Vec and MDN-BiLSTM-based architectures.

| Model | MAE | MSE | RMSE | MAPE | R2 |
|----------------------|-------|-------|-------|-------|-------|
| Time2Vec-BiLSTM | 0.055 | 0.005 | 0.072 | 12.21 | 0.906 |
| Time2Vec-Transformer | 0.054 | 0.005 | 0.071 | 13.85 | 0.909 |
| MDN-BiLSTM | 0.052 | 0.004 | 0.065 | 17.10 | 0.924 |

5.2.4. Chlorophyll a Forecast Summary

In summary, in this work concerning chlorophyll a, after training the 12 different models, we conclude that the models based on the Time2Vec architecture and the MDN-BiLSTM model obtain the best performance. Time2Vec-BiLSTM obtains an MAPE and R2 of 10.51 and 0.940, respectively, and an MAPE of 12.36 and an R2 of 0.929 for Time2Vec-Transformer. Moreover, MDN-BiLSTM obtains an MAPE of 13.38 and an R2 of 0.932. When introducing exogenous variables, none of the three architectures improves the results obtained with the endogenous variables. Finally, a forecast test is performed to predict the next time step outside the dataset used for training the models, resulting in an R2

of 0.906 for the Time2Vec-BiLSTM model, 0.909 for Time2Vec-Transformer, and 0.924 for MDN-BiLSTM.

6. Discussion

This work aims to make a prediction, with a horizon window of one week, of the concentration of dissolved oxygen and chlorophyll for 12 different points located in the water body of the Mar Menor coastal lagoon, with the highest possible accuracy.

We find works that use ML and DL-based techniques to improve the detection and prediction of the ecological state of the Mar Menor. Diego Gómez et al. [50] propose an ML-based approach for the estimation of chlorophyll a content in the upper part of the water column in the Mar Menor, using Sentinel-2 data and algorithms such as RandomForest, Support Vector Machine, artificial neural networks, and deep neural networks. The case study by Patricia Jimeno-Sáez et al. [51] investigates the potential of ML methods to predict chlorophyll a levels by evaluating Support Vector Regression models and multilayer neural networks using data provided by a CTD. In the work of Javier González-Enrique et al. [52], a case study is carried out for the prediction of chlorophyll a concentration in the Mar Menor using DL and, more specifically, Long Short-term Memory Neural Networks (LSTM) [53] with a prediction horizon of one week. Finally, in a different location, Manuel Valera et al. [54] propose applying ML techniques to predict dissolved oxygen through RandomForest and Support Vector regressors.

Up to 12 DL models specialized in neural network forecasting have been evaluated to achieve this objective. The aim is to obtain one or more models that can forecast dissolved oxygen and chlorophyll a values at 12 different points in the water body, with a forecast horizon of one week. The models are generated by combining atomic structures in the form of layers. Among these structures are LSTM, Attention, Time2Vec, or Transformer. As features of the dataset supplied to the models, the parameters provided by a CTD and agrometeorological stations located near the study area have been used. The implemented process generates several datasets for each model by combining different groups of features and different time sequence sizes until the one with the best results is found. Up to five different metrics have been selected to evaluate the performance of the models. Finally, a forecast is made for the measurement points of the week following the last available week to rule out that the models give random forecasts.

The results confirm the ability of advanced neural networks for time series forecasts to forecast dissolved oxygen and chlorophyll a values in the Mar Menor. To do so, it is necessary to correctly adjust the dataset's features and build the appropriate model. One of the main findings of this work is the potential of the Time2Vec architecture for the forecasting of time series data. Time2Vec offers the best metrics for both dissolved oxygen and chlorophyll a. Time2Vec has proven to be an architecture that correctly identifies the most important features and does not pay attention to those that may introduce uncertainty to the model because they are not directly related to the target variable. It has also demonstrated its ability to identify the temporal relationships of the data, as it is not necessary to provide such features artificially. The MDN-BiLSTM model also has excellent metrics in the case of chlorophyll a prognosis.

On the other hand, it has been observed that the exogenous variables introduced in the data set used have not positively influenced the metrics offered by the different models, which is complicated because the results obtained with the endogenous variables are very high. It is true, however, that in the case of architectures based on Time2Vec, they do not significantly worsen them, especially if these exogenous variables are combined correctly. In the case of extraordinary events that drastically change the behavior of the Mar Menor, as with the DANAs, this fact allows the introduction of exogenous variables to improve the prediction offered in cases where the biophysical parameters of the water are conditioned by external elements, at the cost of obtaining slightly lower metrics.

In contrast, this work is not exhaustive concerning the hyperparameters used for each of the models evaluated, as they could have been customized on a model-by-model basis to

achieve the best possible performance for each model tested. In addition, additional scalers could have been incorporated to transform the input data set, remove noise, and improve the signal in the time series. Finally, the reasons why using exogenous variables a priori related does not improve the metrics needed to be explored in more detail.

This work becomes a process integrated into the OMM monitoring system's ML Model Generator module. Therefore, the idea is that it will be continuously evolving and improving based on future needs. As for new ways of working, the aim is to incorporate the forecast of new parameters such as temperature and salinity or to make forecasts at different depths. The Scaler component plans to incorporate new input data transformers from the power transformer family, including logarithm or square root transforms. It also intends to incorporate other exogenous variables, such as nutrient data or water flow entering the wadis from the watershed. Other studies could be carried out with the hyperparameters [55] of the models in order to improve the models that have not performed so well or to use specialized algorithms for feature selection such as mRMR [56] and thus make the Features Generator component lighter. Also, it is intended to extend the forecast horizon to several weeks. We intend to evaluate new models to forecast more recently created time series, such as N-BEATS [57], NBEATSx [58], N-hits [59], or TFT [60]. Finally, an investigation will be carried out to find out why exogenous variables have not positively impacted the metrics and to try to go into more detail to clarify which mechanisms allow better performance of the models by including these variables.

7. Conclusions

The lagoon of the Mar Menor has a series of features that make it unique and of high ecological and environmental value. Since 2016, a series of episodes of hypoxia and eutrophication have been changing the quality of its waters. In order to understand the dynamics that cause these episodes, it is necessary to monitor the lagoon mass and its watershed. Thanks to the data generated by this monitoring and techniques based on ML and DL, we can easily model highly complex and heterogeneous systems without excessive computational capacity.

Thanks to the fact that all efforts in recent years have been focused on monitoring both the coastal lagoon of the Mar Menor and its watershed, data on biophysical water parameters and agroclimatic variables have been available for this work for the time interval from April 2017 to April 2022. This large amount of heterogeneous data becomes an input dataset for training 12 neural network models specially designed for time series forecasting.

The best-performing models have been those based on the Time2Vec architecture. However, it is true that other models, such as MDN-BiLSTM, have offered significant performance. It has also been observed that the Time2Vec architecture allows detecting the temporal relationships of the data without the need to incorporate time features artificially and that, in addition, they can discriminate quite accurately those less relevant features. Regarding metrics, the best performer is Time2Vec-BiLSTM, with an MAPE of 1.50, an R2 of 0.978 for dissolved oxygen, an MAPE of 10.51, and an R2 of 0.940 for chlorophyll a. Both Time2Vec-Transformer and MDN-BiLSTM provide similar metrics. These metrics are corroborated by forecasting for the next time step outside the dataset used. Furthermore, it is clear from these forecasts that the models can complement each other, compensating for the weaknesses of one with the strengths of the other [61].

In conclusion, we have been able to predict values for two biophysical variables, dissolved oxygen and chlorophyll a, with a horizon of 1 week with promising metrics. Among the models evaluated, those based on the Time2Vec architecture stand out, which, in addition to obtaining the best metrics, can discriminate the less important characteristics and detect the time relationships of the time series provided. As an innovative element, we can highlight that this work will be transformed into a module that will be part of the monitoring platform of the Observatory of the Mar Menor that, with a weekly periodicity, will allow for predicting both the concentration of chlorophyll a and dissolved oxygen.

In addition, it will serve as a decision support tool for scientific groups and responsible entities to improve the water quality of the Mar Menor.

Author Contributions: Conceptualization, F.J.L.-A., M.E. and J.A.L.-M.; methodology, F.J.L.-A., M.E., J.A.L.-M. and J.A.C.-R.; software, F.J.L.-A. and J.A.L.-M.; validation, F.J.L.-A., M.E., J.A.L.-M., Z.H.-G., J.A.C.-R., M.S.-A. and J.F.A.-J.; formal analysis, F.J.L.-A., M.E. and J.A.L.-M.; investigation, F.J.L.-A., M.E., J.A.L.-M. and Z.H.-G.; resources, F.J.L.-A., M.E., J.A.L.-M., Z.H.-G., J.A.C.-R., M.S.-A. and J.F.A.-J.; data curation, F.J.L.-A. and J.A.C.-R.; writing—original draft preparation, F.J.L.-A., M.E., J.A.L.-M., and Z.H.-G.; writing—review and editing, F.J.L.-A., M.E., J.A.L.-M., Z.H.-G., J.A.C.-R., M.S.-A. and J.F.A.-J.; visualization, F.J.L.-A. and J.A.L.-M.; supervision, F.J.L.-A., M.E., J.A.L.-M., Z.H.-G., J.A.C.-R., M.S.-A. and J.F.A.-J.; project administration, M.E.; funding acquisition, M.E. All authors have read and agreed to the published version of the manuscript.

Funding: This study forms part of the ThinkInAzul program and was supported by the Ministerio de Ciencia e Innovación de España (MCIN) with funding from the European Union’s NextGenerationEU (PRTR-C17.I1) and by the Comunidad Autónoma de la Región de Murcia through the project OMAzul, and by the European Regional Development Fund (ERDF), through project FEDER 14-20-25 “Impulso a la economía circular en la agricultura y la gestión del agua mediante el uso avanzado de nuevas tecnologías-iagua”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the Fisheries and Aquaculture Service of General Directorate of Livestock, Fisheries, and Aquaculture, attached to the Regional Ministry of Water, Agriculture, Livestock, and Fisheries of the Autonomous Community of the Region of Murcia, for providing the data collected in the Mar Menor to complete the time series used in this work.

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

References

1. Perez-Ruzafa, A.; Campillo, S.; Fernández-Palacios, J.M.; Garcia-Lacunza, A.; Garcia-Oliva, M.; Ibanez, H.; Navarro-Martinez, P.C.; Pérez-Marcos, M.; Perez-Ruzafa, I.M.; Quispe-Becerra, J.I.; et al. Long-term dynamic in nutrients, chlorophyll a, and water quality parameters in a coastal lagoon during a process of eutrophication for decades, a sudden break and a relatively rapid recovery. *Front. Mar. Sci.* **2019**, *6*, 26. [CrossRef]
2. Soria, J.; Caniego, G.; Hernández-Sáez, N.; Dominguez-Gomez, J.A.; Erena, M. Phytoplankton distribution in Mar Menor coastal lagoon (SE Spain) during 2017. *J. Mar. Sci. Eng.* **2020**, *8*, 600. [CrossRef]
3. Androulakis, D.N.; Banks, A.C.; Dounas, C.; Margaris, D.P. An evaluation of autonomous in situ temperature loggers in a coastal region of the eastern Mediterranean sea for use in the validation of near-shore satellite sea surface temperature measurements. *Remote Sens.* **2020**, *12*, 1140. [CrossRef]
4. Malde, K.; Handegard, N.O.; Eikvil, L.; Salberg, A.B. Machine intelligence and the data-driven future of marine science. *ICES J. Mar. Sci.* **2020**, *77*, 1274–1285. [CrossRef]
5. Zhong, S.; Zhang, K.; Bagheri, M.; Burken, J.G.; Gu, A.; Li, B.; Ma, X.; Marrone, B.L.; Ren, Z.J.; Schrier, J.; et al. Machine learning: New ideas and tools in environmental science and engineering. *Environ. Sci. Technol.* **2021**, *55*, 12741–12754. [CrossRef]
6. López-Andreu, F.J.; López-Morales, J.A.; Atenza Juárez, J.F.; Alcaraz, R.; Hernández, M.D.; Erena, M.; Domínguez-Gómez, J.A.; García Galiano, S. Monitoring System of the Mar Menor Coastal Lagoon (Spain) and Its Watershed Basin Using the Integration of Massive Heterogeneous Data. *Sensors* **2022**, *22*, 6507. [CrossRef]
7. Pérez-Ruzafa, A.; Marcos, C.; Pérez-Ruzafa, I.M.; Pérez-Marcos, M. Coastal lagoons: “Transitional ecosystems” between transitional and coastal waters. *J. Coast. Conserv.* **2011**, *15*, 369–392. [CrossRef]
8. Perez-Ruzafa, A.; Perez-Marcos, M.; Marcos, C. Coastal lagoons in focus: Their environmental and socioeconomic importance. *J. Nat. Conserv.* **2020**, *57*, 125886. [CrossRef]
9. Keras. 2015. Available online: <https://github.com/fchollet/keras> (accessed on 6 June 2023).
10. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th {USENIX} Symposium on Operating Systems Design and Implementation (OSDI 16), Savannah, GA, USA, 2–4 November 2016; pp. 265–283.
11. McKinney, W. Data structures for statistical computing in python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; Volume 445, pp. 51–56.

12. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
13. Tal, A.; Weinstein, Y.; Baïssset, M.; Golan, A.; Yechieli, Y. High resolution monitoring of seawater intrusion in a multi-aquifer system-implementation of a new downhole geophysical tool. *Water* **2019**, *11*, 1877. [[CrossRef](#)]
14. Comunidad Autónoma de la Región de Murcia. Available online: <https://www.carm.es/> (accessed on 6 June 2023).
15. Instituto Murciano de Investigación y Desarrollo Agrario y Medioambiental. Available online: <https://www.imida.es> (accessed on 6 June 2023).
16. Diego-Feliu, M.; Rodellas, V.; Alorda-Kleinglass, A.; Domínguez-Gabarró, J.; Saaltink, M.; Folch, A.; Garcia-Orellana, J. Ephemeral Streams: An overlooked permanent source of groundwater and Nutrients to the Mediterranean Sea. In Proceedings of the EGU General Assembly Conference Abstracts, Online, 19–30 April 2021; pp. EGU21–15871.
17. Almuammar, M.; Fasli, M. Deep learning for non-stationary multivariate time series forecasting. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 2097–2106.
18. Malve, O.; Qian, S.S. Estimating nutrients and chlorophyll a relationships in Finnish lakes. *Environ. Sci. Technol.* **2006**, *40*, 7848–7853. [[CrossRef](#)]
19. Denman, K.L. Covariability of chlorophyll and temperature in the sea. In *Deep Sea Research and Oceanographic Abstracts*; Elsevier: Amsterdam, The Netherlands, 1976; Volume 23, pp. 539–550.
20. Agricultural Information System of the Murcia Region. Available online: <http://siam.imida.es> (accessed on 13 March 2023).
21. Montero-Manso, P.; Hyndman, R.J. Principles and algorithms for forecasting groups of time series: Locality and globality. *Int. J. Forecast.* **2021**, *37*, 1632–1653. [[CrossRef](#)]
22. Salinas, D.; Flunkert, V.; Gasthaus, J.; Januschowski, T. DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *Int. J. Forecast.* **2020**, *36*, 1181–1191. [[CrossRef](#)]
23. Smyl, S. A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *Int. J. Forecast.* **2020**, *36*, 75–85. [[CrossRef](#)]
24. Kuang, L.; Yan, X.; Tan, X.; Li, S.; Yang, X. Predicting taxi demand based on 3D convolutional neural network and multi-task learning. *Remote Sens.* **2019**, *11*, 1265. [[CrossRef](#)]
25. Jeon, Y.E.; Kang, S.B.; Seo, J.I. Hybrid Predictive Modeling for Charging Demand Prediction of Electric Vehicles. *Sustainability* **2022**, *14*, 5426. [[CrossRef](#)]
26. Chen, J.; Li, Q.; Wang, H.; Deng, M. A machine learning ensemble approach based on random forest and radial basis function neural network for risk evaluation of regional flood disaster: A case study of the Yangtze River Delta, China. *Int. J. Environ. Res. Public Health* **2020**, *17*, 49. [[CrossRef](#)]
27. Seabold, S.; Perktold, J. Statsmodels: Econometric and statistical modeling with python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010.
28. Suradhaniwar, S.; Kar, S.; Durbha, S.S.; Jagarlapudi, A. Time series forecasting of univariate agrometeorological data: A comparative performance evaluation via one-step and multi-step ahead forecasting strategies. *Sensors* **2021**, *21*, 2430. [[CrossRef](#)]
29. Wang, H.; Tian, C.; Wang, W.; Luo, X. Temporal cross-correlations between ambient air pollutants and seasonality of tuberculosis: A time-series analysis. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1585. [[CrossRef](#)]
30. Ahsan, M.M.; Mahmud, M.P.; Saha, P.K.; Gupta, K.D.; Siddique, Z. Effect of data scaling methods on machine learning algorithms and model performance. *Technologies* **2021**, *9*, 52. [[CrossRef](#)]
31. Botchkarev, A. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *arXiv* **2018**, arXiv:1809.03006.
32. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [[CrossRef](#)]
33. Rajagukguk, R.A.; Ramadhan, R.A.; Lee, H.J. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies* **2020**, *13*, 6623. [[CrossRef](#)]
34. Graves, A.; Graves, A. Long short-term memory. In *Supervised Sequence Labelling with Recurrent Neural Networks*; Springer: Berlin, Germany; 2012; pp. 37–45.
35. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The performance of LSTM and BiLSTM in forecasting time series. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 3285–3292.
36. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. A comparative analysis of forecasting financial time series using arima, lstm, and bilstm. *arXiv* **2019**, arXiv:1911.09512.
37. Pirani, M.; Thakkar, P.; Jivrani, P.; Bohara, M.H.; Garg, D. A comparative analysis of ARIMA, GRU, LSTM and BiLSTM on financial time series forecasting. In Proceedings of the 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballari, India, 23–24 April 2022; pp. 1–6.
38. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. *arXiv* **2014**, arXiv:1409.3215.
39. Bishop, C.M. *Mixture Density Networks*; Technical Report; Neural Network Research Group, Aston University: Birmingham, UK, 1994; ISBN NCRG/94/004.
40. Zhang, H.; Liu, Y.; Yan, J.; Han, S.; Li, L.; Long, Q. Improved deep mixture density network for regional wind power probabilistic forecasting. *IEEE Trans. Power Syst.* **2020**, *35*, 2549–2560. [[CrossRef](#)]

41. Lea, C.; Vidal, R.; Reiter, A.; Hager, G.D. Temporal convolutional networks: A unified approach to action segmentation. In Proceedings of the Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, 8–10 and 15–16 October 2016; Proceedings, Part III 14; Springer: Berlin, Germany, 2016; pp. 47–54.
42. Wan, R.; Mei, S.; Wang, J.; Liu, M.; Yang, F. Multivariate temporal convolutional network: A deep neural networks approach for multivariate time series forecasting. *Electronics* **2019**, *8*, 876.
43. Remy, P. Temporal Convolutional Networks for Keras. 2020. Available online: <https://github.com/philipperemy/keras-tcn> (accessed on 6 June 2023).
44. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv* **2014**, arXiv:1409.0473.
45. Luong, M.T.; Pham, H.; Manning, C.D. Effective approaches to attention-based neural machine translation. *arXiv* **2015**, arXiv:1508.04025.
46. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *arXiv* **2017**, arXiv:1706.03762
47. Huang, L.; Mao, F.; Zhang, K.; Li, Z. Spatial-temporal convolutional transformer network for multivariate time series forecasting. *Sensors* **2022**, *22*, 841. [[CrossRef](#)]
48. Kazemi, S.M.; Goel, R.; Eghbali, S.; Ramanan, J.; Sahota, J.; Thakur, S.; Wu, S.; Smyth, C.; Poupart, P.; Brubaker, M. Time2vec: Learning a vector representation of time. *arXiv* **2019**, arXiv:1907.05321.
49. Song, J.; Tong, X.; Xu, X.; Zhao, K. A Real-Time Reentry Guidance Method for Hypersonic Vehicles Based on a Time2vec and Transformer Network. *Aerospace* **2022**, *9*, 427. [[CrossRef](#)]
50. Gómez, D.; Salvador, P.; Sanz, J.; Casanova, J.L. A new approach to monitor water quality in the Menor sea (Spain) using satellite data and machine learning methods. *Environ. Pollut.* **2021**, *286*, 117489. [[CrossRef](#)]
51. Jimeno-Sáez, P.; Senent-Aparicio, J.; Cecilia, J.M.; Pérez-Sánchez, J. Using machine-learning algorithms for eutrophication modeling: Case study of Mar Menor Lagoon (Spain). *Int. J. Environ. Res. Public Health* **2020**, *17*, 1189. [[CrossRef](#)]
52. González-Enrique, J.; Ruiz-Aguilar, J.J.; Madrid Navarro, E.; Martínez Álvarez-Castellanos, R.; Felis Enguix, I.; Jerez, J.M.; Turias, I.J. Deep Learning Approach for the Prediction of the Concentration of Chlorophyll a in Seawater. A Case Study in El Mar Menor (Spain). In Proceedings of the 17th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2022), Salamanca, Spain, 5–7 September 2022; Springer: Berlin, Germany, 2022; pp. 72–85.
53. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
54. Valera, M.; Walter, R.K.; Bailey, B.A.; Castillo, J.E. Machine learning based predictions of dissolved oxygen in a small coastal embayment. *J. Mar. Sci. Eng.* **2020**, *8*, 1007. [[CrossRef](#)]
55. Xiao, X.; Yan, M.; Basodi, S.; Ji, C.; Pan, Y. Efficient hyperparameter optimization in deep learning using a variable length genetic algorithm. *arXiv* **2020**, arXiv:2006.12703.
56. Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [[CrossRef](#)]
57. Oreshkin, B.N.; Carpov, D.; Chapados, N.; Bengio, Y. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. *arXiv* **2019**, arXiv:1905.10437.
58. Olivares, K.G.; Challu, C.; Marcjasz, G.; Weron, R.; Dubrawski, A. Neural basis expansion analysis with exogenous variables: Forecasting electricity prices with NBEATSx. *Int. J. Forecast.* **2023**, *39*, 884–900. [[CrossRef](#)]
59. Challu, C.; Olivares, K.G.; Oreshkin, B.N.; Garza, F.; Mergenthaler-Canseco, M.; Dubrawski, A. N-hits: Neural hierarchical interpolation for time series forecasting. *arXiv* **2022**, arXiv:2201.12886.
60. Lim, B.; Arık, S.Ö.; Loeff, N.; Pfister, T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *Int. J. Forecast.* **2021**, *37*, 1748–1764. [[CrossRef](#)]
61. Livieris, I.E.; Pintelas, E.; Stavroyiannis, S.; Pintelas, P. Ensemble deep learning models for forecasting cryptocurrency time-series. *Algorithms* **2020**, *13*, 121. [[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.