



**HAL**  
open science

## Modelling daily Dissolved Oxygen Dynamics in the Sebou River (Morocco): Data-Centric Approaches

Souad Haida, Saad Ablat, Hayat Sibari, Sara El Mrissani, Youssef Brouziyne,  
Hanaa Lamine, Jean-Luc Probst, Anne Probst, Laila Misane

► **To cite this version:**

Souad Haida, Saad Ablat, Hayat Sibari, Sara El Mrissani, Youssef Brouziyne, et al.. Modelling daily Dissolved Oxygen Dynamics in the Sebou River (Morocco): Data-Centric Approaches. 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET 2022), Mar 2022, Meknes, Morocco. pp.1-9, 10.1109/IRASET52964.2022.9737979 . hal-03755515

**HAL Id: hal-03755515**

**<https://hal.science/hal-03755515>**

Submitted on 22 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Modelling daily Dissolved Oxygen Dynamics in the Sebou River (Morocco): Data-Centric Approaches

Souad Haida

*Lab. Natural Resources & Sustainable Development*  
*Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco*  
haida.souad@uit.ac.ma

Saad Ablat

*Lab. Natural Resources & Sustainable Development*  
*Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco*  
saad.ablat@uit.ac.ma

Hayat Sibari

*Lab. Natural Resources & Sustainable Development*  
*Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco*  
sibarih@yahoo.fr

Sara El Mrissani

*Lab. Natural Resources & Sustainable Development*  
*Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco*  
sara.elmrissani@uit.ac.ma

Youssef Brouziyne

*Mohammed VI Polytechnic University (UM6P)*  
*International Water Research Institute, Benguerir, Morocco*  
Youssef.brouziyne@gmail.com

Hanaa Lamine

*Lab. Natural Resources & Sustainable Development*  
*Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco*  
hanaa.lamine@uit.ac.ma

Jean Luc Probst

*Lab. Ecologie Fonctionnelle et Environnement, Université de Toulouse, CNRS, France*  
jean-luc.probst@toulouse-inp.fr

Anne Probst

*Lab. Ecologie Fonctionnelle et Environnement, Université de Toulouse, CNRS, France*  
anne.probst@toulouse-inp.fr

Laila Misane

*Agence du Bassin Hydraulique du Sebou, Fes, Morocco*  
l.misane@abhsebou.ma

**Abstract**— Dissolved oxygen is one of the parameters of river water quality for which the accurate estimation is critical and of significant importance for many environmental applications. The continuous monitoring at high frequency (1 hour) of the water quality of the Sebou at the Mechra Bel Ksiri gauging station began in June 2020 using a multiparameter probe YSI (EXO2) and sensors. Real-time water quality data were collected in the Sebou River for Dissolved Oxygen, pH, temperature, electrical conductivity, turbidity, chlorides, and nitrates. Water level measurements were recorded at the same frequency by the radar. In this study, the high-frequency measurement highlights nycthemeral cycles at the day scale. A sine function model and three machine learning techniques: XGB Regressor, Random Forest Regressor, and Linear Regression, were used to develop data-driven models using temperature, nitrate levels, pH, and electrical conductivity to reproduce the daily fluctuations of dissolved oxygen. Each machine learning model was calibrated for each input scenario using 80% of the data obtained between June and November 2020 and validated using the remaining 20%. The evaluation simulations quality included the calculation of the error indices between simulations and observations (MAE and RMSE) and coefficient of determination ( $R^2$ ). The prediction based on the three machine learning techniques gave good results for XGB Regressor, Linear Regression, and Random Forest Regressor with  $R^2$  values of 0.93, 0.89, and 0.88, respectively, while the sine function was slightly over-predicting the daily cycle of dissolved oxygen ( $R^2=0.80$ ) and indicated a relatively high saturation rate. The error indices were very low, with MAE and RMSE ranging in 0.05–0.06 and 0.06–0.08 respectively, which suggests satisfactory models performance.

**Keywords**— Daily fluctuations, Dissolved oxygen, Sine equation, Machine-learning techniques, Water monitoring, Sebou River.

## I. INTRODUCTION

In water quality monitoring programs, single measurements of dissolved oxygen are usually taken during the daytime and do not represent the real oxygen levels in the river. Only continuous monitoring on 24-hour cycles can lead to an accurate assessment of oxygen status, but this is not always easy to achieve due to logistical and budgetary constraints. Modeling can be a practical alternative to direct measurements. However, water quality models that take into account the nycthemeral oxygen cycle are generally complex to implement. Dissolved oxygen concentration evolves according to the diurnal cycle, which is crucial to consider when assessing the oxygenation status of a river and which can be represented by the high frequency of water quality monitoring [1]. The water quality prediction-related literature review shows that data-driven techniques have taken considerable importance in recent years [2], [3]. In the streams, a simulation of dissolved oxygen dynamics must include temperature, which reflects their effects on dissolved oxygen saturation concentration, re-aeration, respiration, and algal production. To improve the knowledge on the quality of aquatic systems and their prediction, an automated numerical system was able to analyze the high-resolution data-driven. Artificial intelligence may be a practical tool in the prediction of daily, seasonal, and interannual variations [3]. Authors have focused on the application and role of deep learning in understanding the prediction of quality parameters [4]. Among these techniques, machine-learning algorithms have shown their efficiency in several fields such as water quality forecasting [5], [6], [7]. The purpose of this study was to evaluate the feasibility of two types of simulation models and their specificities regarding the available data selected as input variables in the Sebou River. The outcomes of this study are expected to provide a reliable approach to monitoring the water quality aspect in the Sebou River.

---

Funding : Agence de l'eau Artois-Picardie , Doué, in France/ CNRS France, Agence du Bassin Hydraulique du Sebou, Ibn Tofail University, Morocco

## II. MATERIALS AND METHODS

### A. Study area

The Sebou River has its source, at an altitude of 2030 m, stretching over 500 km before reaching the Atlantic Ocean and draining an area of about 40,000 km<sup>2</sup> in the northwest of Morocco “Fig. 1”. The climate is of Mediterranean type with

oceanic influence. The average cumulative annual rainfall varies between 600 mm to 1000 mm and the average annual temperature between 10°C to 20°C, respectively.

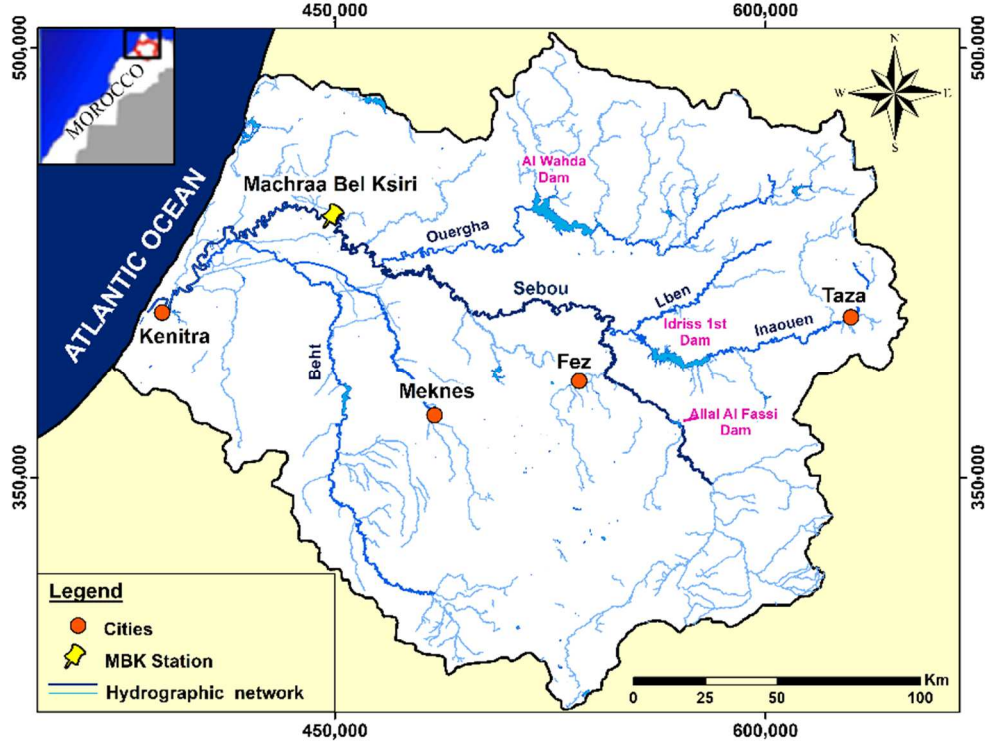


Fig. 1. Location of the study area: the Sebou basin Morocco

The annual surface water inflow of the Sebou River 5600 Mm<sup>3</sup>.year<sup>-1</sup> is very irregular in space and time with the effects of dams [8]. Liquid discharges from domestic wastewater and food processing industries (oil, sugar, dairy products, etc.) generate biological inhibition in the aquatic ecosystem due to the high level of organic matter concentration [9].

The main water physicochemical parameters, such as dissolved oxygen concentration, temperature, pH, electrical conductivity, turbidity, nitrate concentration, and chloride, of the Sebou River were measured with high frequency (every hour) since June 2020 by multi-probe sensors (Exo2) located at Mechra Bel Ksiri, the downstream station “Fig.1”.

### B. Methods

#### -Sine function

Modelling of diurnal oxygen variations was carried out, with a sine function “(1)”, by using the physicochemical variables measured in the river [10].

$$O_2 = C_m + A \sin \left[ \left( 2\pi \cdot \frac{t}{24} \right) - B \right] \quad (1)$$

$O_2$ : dissolved oxygen concentration;  $C_m$ : average dissolved oxygen concentration;  $A$ : amplitude of the cycle;  $B$ : phase of the oscillation;  $t$ : time.

The mean dissolved oxygen concentration ( $C_m$ ) shows, in the Sebou River, a strong correlation, ( $R^2 = 0.78$ ), with the mean temperature ( $T_m$ ) “Fig.2”.

Mean dissolved oxygen concentration ( $C_m$ ) and amplitude were substituted for temperature and nitrate concentration to calculate the diurnal oxygen variation according to “(2)”:

$$O_2 = 12.04 - 0.17 T_m + \left[ (0.31 \times NO_3 - N) - 0.12 \right] \times \sin \left[ \left( 2\pi \cdot \frac{t}{24} \right) - \pi \right] \quad (2)$$

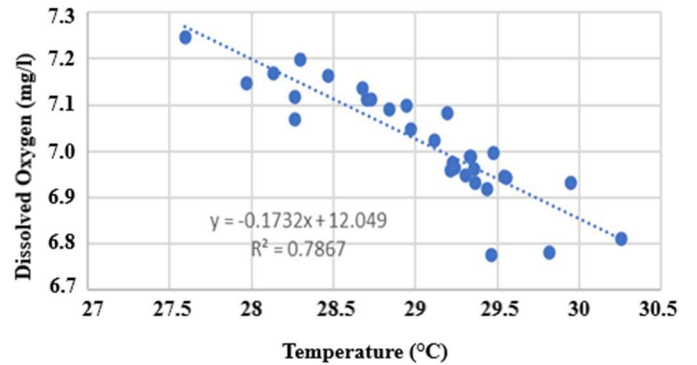


Fig. 2. Relationship between the average daily oxygen concentration (mg/l) and the average daily temperature (°C) in the Sebou River

### -Machine learning algorithms

Input variable selection is one of the various techniques for improving the performance in dissolved oxygen prediction.

The selection input methods include correlation coefficients (such as Pearson's correlation coefficient), mutual information, and MIC [11], [12]. The Mutual information can capture the relationship between two random variables sampled at the same time. In this study, mutual information has been widely applied to select the right combination of inputs and lead to good performance instead of the correlation coefficient.

The selected algorithms are XGB Regressor, Random Forest Regressor, and Linear Regression. Those algorithms are not requiring too much data and are characterized by good performance [13]. The inputs were chosen based on the mutual information defining the relationship between two random variables sampled simultaneously "Fig.3".

Decision Tree Learning in the Random Forest Regressor is one of the predictive modeling approaches in statistics, data mining, and machine learning. It uses a decision tree to go from observations by branches to conclusions in leaves [13].

The Linear Regression algorithm is a statistical method used to model the linear relationship between dependent and independent variables [14] and to predict output values in the absence of input values [15]. In the XGB, the Gradient booster relies on the next model predicted will reduce the predicted errors when mixed with the previous ones. The main idea is to set up a target result for the next model to minimize these errors. It refers to the fact of gradient descent on the loss function to find the model parameters.

The physicochemical parameters used as input are water flow, temperature, pH, Nitrates, and electric conductivity, which form a combination that shows the best results. However, this model does not take into account the temporal scale, as well as the part of the day so the model differs between the day when the light is present and the night when it is not present.

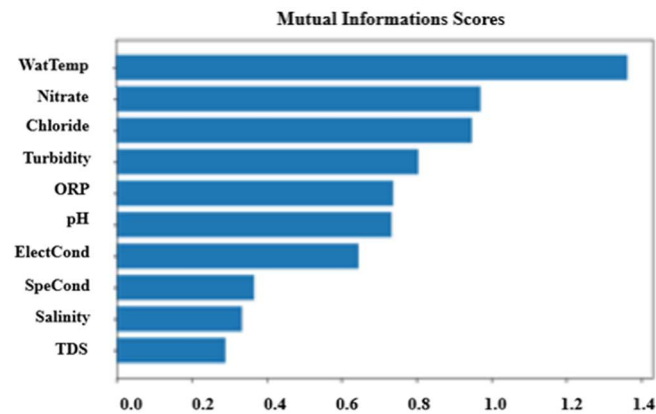


Fig. 3. Mutual information scores between dissolved oxygen and other measured parameters (*WatTemp*: Water Temperature, *Nitrate*, *Chloride*, *Turbidity*, *ORP*: Oxidation-reduction potential, *ElectCond*: Electrical conductivity, *SpeCond*: Specific Conductivity, *Salinity* and *TDS*: Total dissolved Solids)

The variation of dissolved oxygen follows a daily cycle that Machine Learning algorithms cannot reproduce directly, the cyclic data were transformed into sine and cosine functions as follows "(3)":

$$X_{sin} = \sin\left(\frac{(2 \cdot \pi \cdot x)}{\max(x)}\right) \quad \text{et} \quad X_{cos} = \cos\left(\frac{(2 \cdot \pi \cdot x)}{\max(x)}\right) \quad (3)$$

### -Model performance assessment.

In this study, the coefficient of determination ( $R^2$  "(4)"), Mean absolute error (MAE, "(5)"), and root mean square error (RMSE, "(6)") are selected to measure the prediction accuracy and operation efficiency, which are defined as follows:

A perfect model/data match is characterized by an  $R^2$ -value equal to 1, an MAE value equal to 0, and an RMSE value equal to 0. The formulas of the three parameters are given below.

- Coefficient of determination ( $R^2$ )

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (4)$$

- Mean absolute error (MAE)

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - x_i| \quad (5)$$

- Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - x_i)^2} \quad (6)$$

Where  $N$  is the total number of actual points in the data,  $x_i$  is the observed value of period  $i$ ,  $y_i$  is the prediction value of period  $i$ ,  $\bar{x}$  is the average of observed values, and  $\bar{y}$  is the average of prediction values.

## III. RESULTS

### A. Temporal variations of temperature and dissolved oxygen on the downstream of the Sebou River

The rolling mean of the temperature and the dissolved oxygen can highlight the monthly variations "Fig.4". The water temperatures measured during the summer (June, July, and August) increase and reach values between 24.26°C and 30.8°C, with a mean of 28.15°C and a standard deviation of

1.32°C. During the fall season (September, October, and November), the water temperature decreases and reaches values between 18.09°C and 28.34°C, with a mean of 23.24°C and a standard deviation of 2.82°C.

Dissolved oxygen varies the opposite of temperature. The lowest values recorded during the summer are between 6.77 mg/l and 7.53 mg/l, with an average of 7.14 mg/l and a standard deviation of 0.20 mg/l. During the fall season, dissolved oxygen increase levels vary between 7.22 mg/l and 8.78 mg/l, with a mean value of 7.95 mg/l and a standard deviation of 0.44 mg/l.

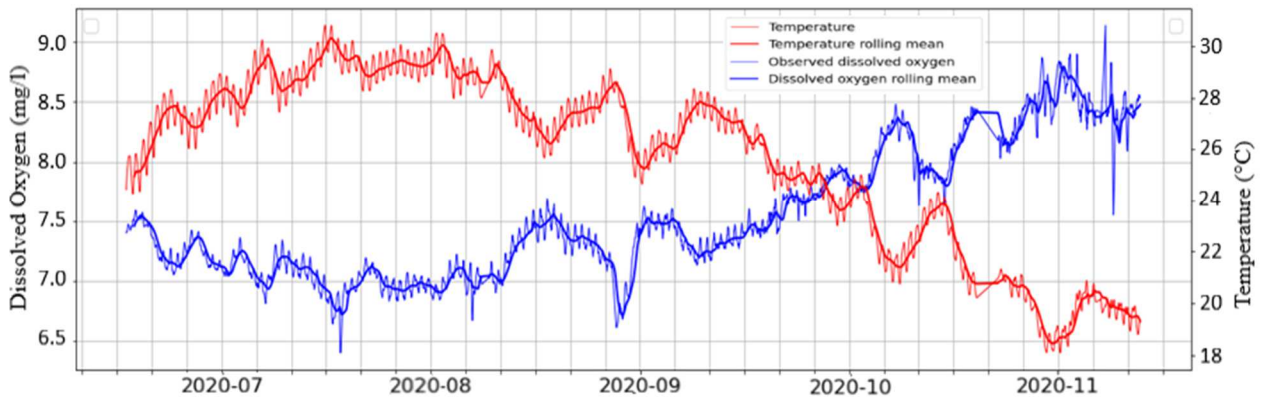


Fig. 4. Hourly variations and the rolling mean of water temperature and dissolved oxygen concentration on the Sebou River

### B. Dissolved oxygen cycles measured in the Sebou River

The daily variations of dissolved oxygen amplitude ( $O_{2max} - O_{2min}$ ) downstream of Sebou River vary between 0.05 mg/l and 1.28 mg/l with an average of 0.22 mg/l and a standard deviation of 0.15 mg/l “Fig.5”. This amplitude was considered low compared to the values of the Loire River in France, characterized by phytoplankton development of 8mg/l in July and a standard deviation of 2mg/l [16].

The value obtained for the Sebou river is close to that of the Huron River in the USA, which shows an amplitude of 1.48 mg/l, mainly related to the regulated nature of the river by the dam and organic pollutant discharges [17]. The daily temperature range ( $T_{max}-T_{min}$ ) varies between 0.29°C and 2.06°C.

Water discharge varies among 5.28 m<sup>3</sup>/s and 32.61 m<sup>3</sup>/s with a maximum of 57.73 m<sup>3</sup>/s. These supplies, related to dam releases, have an apparent effect on the dissolved oxygen daily cycle. The nyctemeral cycle was interrupted; discharge increase induces high concentration variations. This is due to the contribution of waters from a different origin in the catchment (deep water in recession flow and dam’s water release).

The correlation between physicochemical parameters shows a significant relationship between dissolved oxygen concentration and other parameters such as water temperature, pH, nitrate concentration, and total dissolved load (Table 1).

These correlations are positive with pH, nitrates, and Total dissolved solids (TDS) and negative with chloride, turbidity, conductivity, and water flow. Average oxygen concentrations showed a very strong correlation with temperature “ $R^2 = 0.95$ ” and the amplitude of oxygen level variations was strongly correlated with average nitrate concentration “ $R^2 = 0.77$ ”.

Table 1. Correlation matrix between the Physico-chemical parameters measured in the Sebou River (TDS: Total dissolved solids, ORP: Oxidation-reduction Potential)

	Deficit	Dissolved oxygen	Flow rate	Temperature	Conductivity	PH	Turbidity	Nitrate	Chloride	TDS	ORP
Deficit											
Dissolved oxygen	-0.2505										
Flow rate	0.2287	-0.3380									
Temperature	-0.0287	-0.9583	0.2687								
Conductivity	0.0292	-0.3774	-0.1355	0.3988							
PH	-0.3765	0.7647	-0.1291	-0.7116	0.3940						
Turbidity	0.2650	-0.3934	-0.0523	0.3345	0.3597	-0.5483					
Nitrate	0.1019	0.7777	-0.2832	-0.8248	-0.1629	0.5079	-0.1748				
Chloride	0.0676	-0.5840	-0.0959	0.5968	0.2986	-0.8185	0.4827	-0.5554			
TDS	0.0585	0.4655	-0.3430	-0.4811	0.6081	0.2353	0.0571	0.5593	-0.2241		
ORP	-0.0305	0.5397	0.1412	-0.5419	-0.2362	0.6353	-0.4047	0.5605	-0.6915	0.2497	

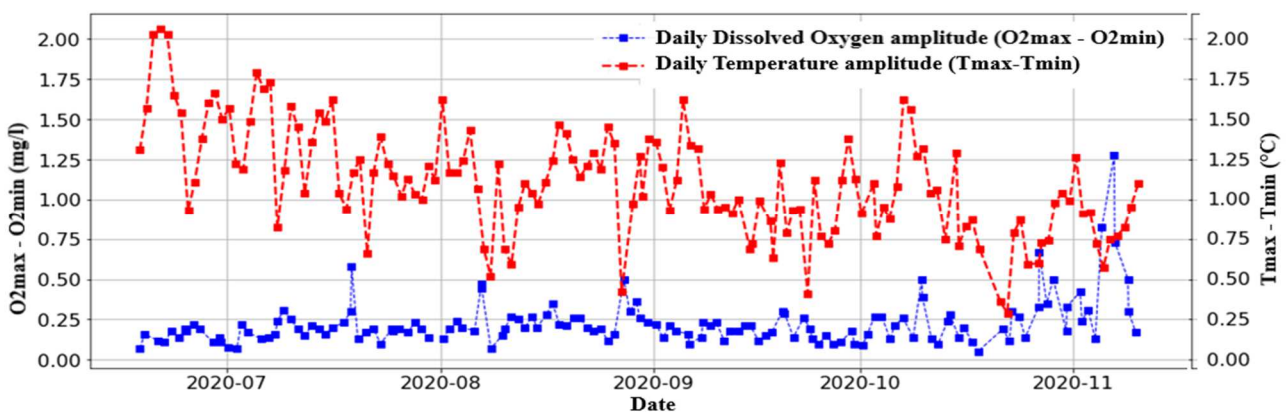


Fig. 5. The daily amplitude of Dissolved Oxygen ( $O_{2max} - O_{2min}$ ) and temperature ( $T_{max}-T_{min}$ ) in the Sebou River

C. Diurnal dissolved oxygen cycle modelling by the sine function

The relations between water temperature and nitrate concentration were used in the sine function and resulted in a

significant correlation between modelled, and measured oxygen concentrations for six daily cycles “Fig.6”.

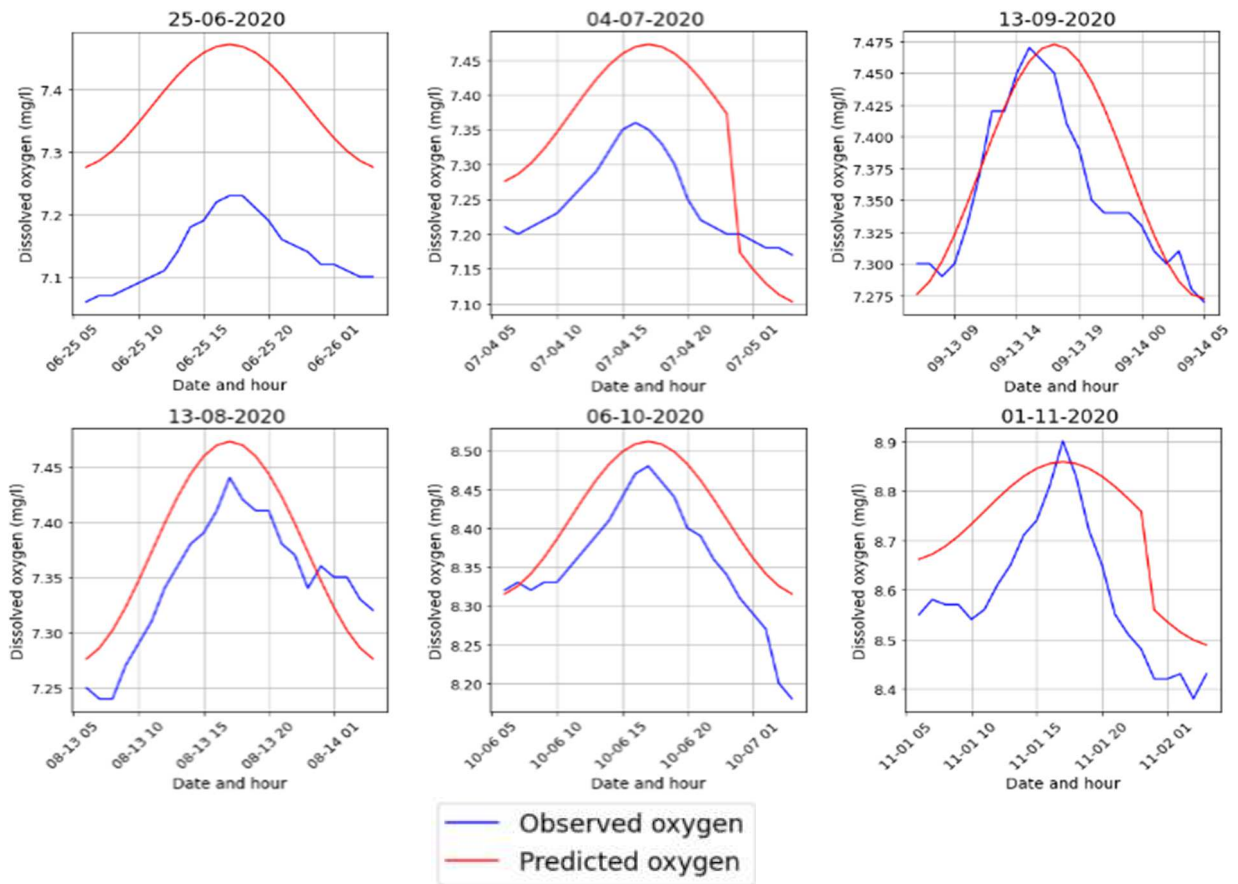


Fig. 6. Comparison of the Nycthemeral cycles of dissolved oxygen measured and estimated with the sine function in the Sebou River

The magnitude of the variations generated by modelling for the several cycles shows a global trend that is relatively close to the dissolved oxygen values measured “Fig. 7”. However, this model encounters difficulties in reproducing

the daily variation and the amplitude of the cycles by generally tending towards an overestimation of the dissolved oxygen concentrations.

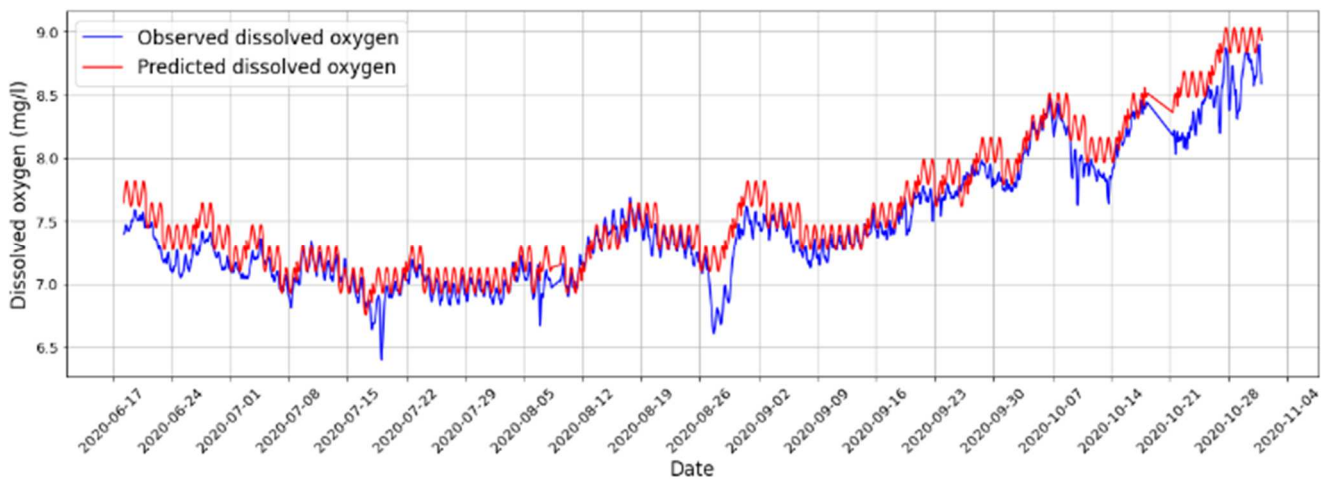


Fig. 7. Comparison between the Dissolved Oxygen daily variations observed and predicted in summer and fall seasons

#### D. Dissolved oxygen Modelling by machine learning

The ecological quality of the Sebou River was predicted by Machine Learning using the XGB Regressor algorithm based on a set of data measured by the probe and compared

with the Linear Regression and Random Forest Regressor. A first modeling attempt used data from the entire period to train the model excluding, the cycle to be modeled “Fig.8”.

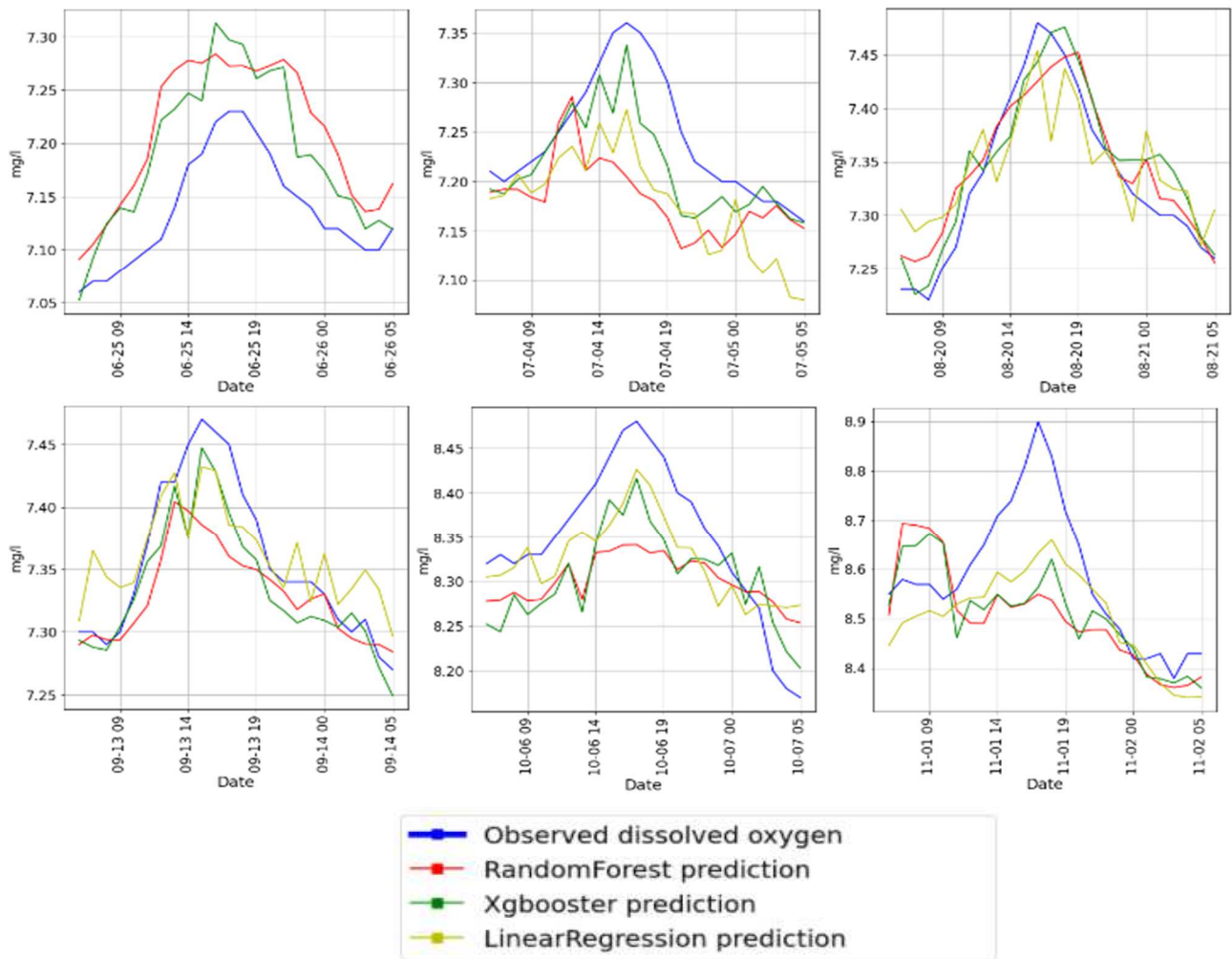


Fig. 8. Comparison of dissolved oxygen Nycthemeral cycles measured and estimated with XGBRegressor, Random Forest Regressor, and Linear Regression in the Sebou River

Algorithmic Random Forest Regressor accepts a restricted database and predicts statistically optimal values when measurements have not been done. The simulation results of dissolved oxygen match with the daily variation measured for the cycle of 2/August/20, underestimated for three cycles of 4/July- 13/October- 1/November, and overestimated for the 25/June cycle. The results obtained with Linear Regression show overall the best estimate of dissolved oxygen content in the first four cycles and an underestimate for the 1/November cycle.

The results obtained with Linear Regression show overall the best estimate of dissolved oxygen content in the first four cycles and an underestimate for the 1/November cycle. The regression coefficient between the data modelled and

measured data for the three algorithms Random Forest, XGB Regressor, and linear Regression varies between 0.92 to 0.2, 0.91 to 0.31, as well as 0.80 to 0.63 respectively.

The simulation results by the three algorithms were compared with the observed variations for the period from 18<sup>th</sup> July 2020 to 31<sup>st</sup> July 2020 “Fig.9”. Dissolved oxygen predicted by XG Bregressor and Random Forest Regressor was similar for most of the daily cycles. However, the comparison with the measured dissolved oxygen shows that those models overestimate the maximum values recorded during the daytime. The predictions calculated by the Linear Regression algorithm show values that are close to those measured, above all when the general trend is increasing.

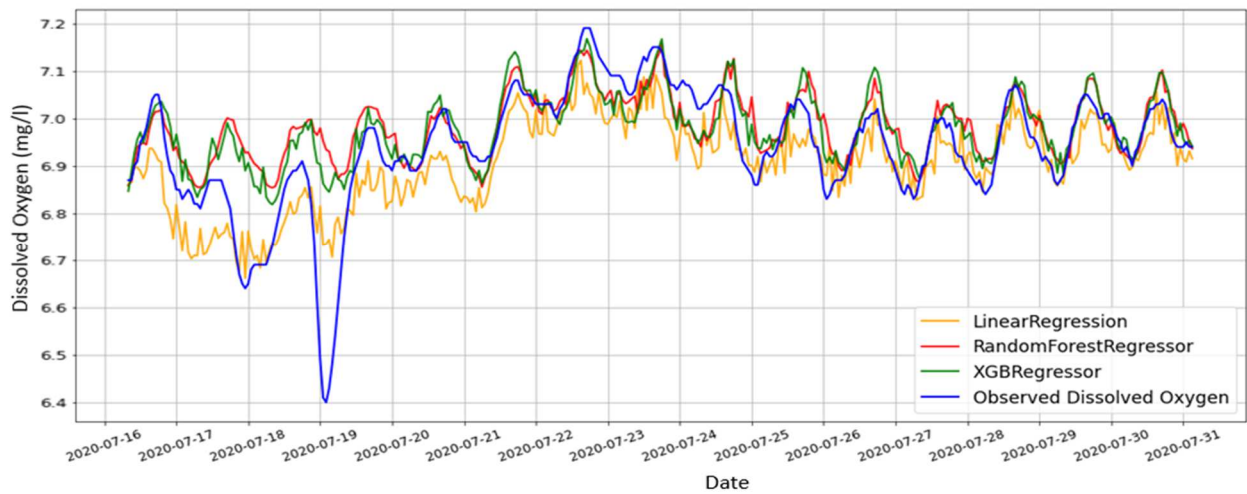


Fig. 9. Comparison between the simulation results of XGB Regressor, Random Forest Regressor, Linear Regression algorithms, and the data measured in July between 18<sup>th</sup> and 31<sup>st</sup> /2020.

### E. Comparisons between the sine function and the Machine Learning models

The results of simulating dissolved oxygen daily cycles by Machine learning algorithms are very close to the real-time field measurements “Fig.10”. The simulation by the sine function overestimates all the daily cycle values and indicates a high saturation rate.

The algorithms and sine function model's performance was evaluated for each data set by calculating the  $R^2$ , the root means square error (RMSE), and the Mean absolute error (MAE). The performance statistics of the three algorithms and sine function were computed and summarized in Table 2.

Table 2. Coefficient of determination “ $R^2$ ”, mean absolute error “MAE”, and root-mean-square error “RMSE” between modelled values and measured data.

	$R^2$	MAE (mg/L)	RMSE (mg/L)
Sine function	0.869	0.151	0.199
XGBRegressor	0.900	0.055	0.069
RandomForest	0.861	0.065	0.081
LinearRegression	0.887	0.058	0.073

The value of  $R^2$  for XGB Regressor, Random Forest, and Linear Regression was between 0.86 and 0.90 (good predictive ability). The error indices were very low, with MAE and RMSE ranging in 0.05–0.06 and 0.06–0.08 respectively, which suggests either satisfactory model

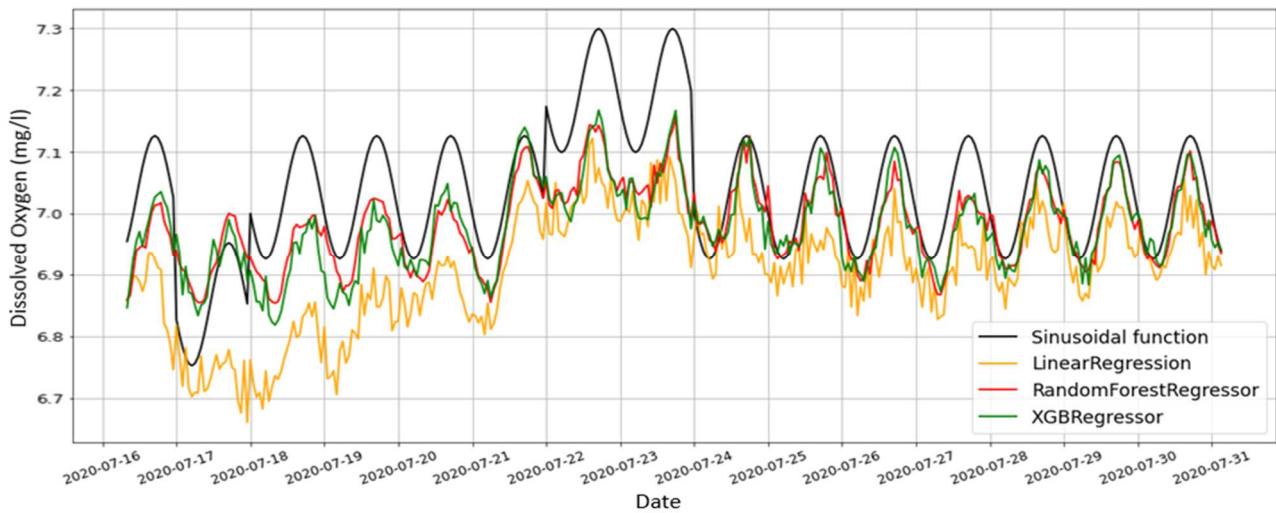


Fig. 10. Comparison between the simulation results of the XGB Regressor, Random Forest Regressor, Linear Regression algorithms, and the data predicted by the sinusoidal function

These performance criteria show nearly identical values for all three models while indicating a slightly higher quality in the results of the XGB Regressor model ( $R^2=0.90$ , MAE=0.05, and RMSE=0.069 mg/L) and the Random Forest and Linear Regression thus, give a similar performance in estimating the diurnal variation of oxygen.

The obtained  $R^2$  value for the sine function model corresponds to 0.86, which is similar compared to the results obtained for Random Forest and Linear Regression. However, the error indices were higher, with MAE and RMSE about, 0.15 and 0.19 respectively.



## IV. DISCUSSION

### A. Effect of temperature and nitrate concentration on dissolved oxygen cycles

The dissolved oxygen dynamic of aquatic ecosystems is a specific index of water quality and trophodynamics [18] and presents a complex aspect that is not yet well studied. The continuous and high-resolution monitoring of the water quality of the Sebou River allowed a better understanding of the diel dynamics. Among Physical, chemical, and biological factors, the temperature is one of the principal factors, which affect the diurnal variability of dissolved oxygen. Temperature influences both the solubility of oxygen in the water and biological activity (photosynthesis and respiration), thus affecting the short-term oxygen concentration balance [19].

Water Temperature and Dissolved Oxygen in the Sebou waters show notable diel variations due to the changing weather conditions and solar radiation driven by respiration and photosynthesis processes [20].

In the Sebou River, dissolved oxygen amplitude ( $O_{2\max} - O_{2\min}$ ) was probably, controlled by photosynthesis, phytoplanktonic respiration, and decomposition of organic matter. The relationship between dissolved oxygen variation and nitrate concentration is mainly due to the nitrification process of ammonia nitrogen ( $NH_3-N$ ) which uses oxygen from the river for the reaction. This process generates an oxygen loss and can affect the average oxygen concentration in the river [21]. The diurnal variability of oxygen concentration in Sebou River probably, controlled by autotrophic and heterotrophic activities, was disturbed by the inflow of water released by dams, which can induce the re-aeration of the river water [22].

### B. Simulation of diel dissolved oxygen fluctuations

The results of the dissolved oxygen simulation showed that both approaches adopted (i.e., mathematical function and Machine Learning algorithms) resulted in a good correlation between the simulated and measured data. Machine learning models using ten input parameters have better performance indices ( $R^2$ , MAE, and RMSE) than the sine function model, which uses only two input parameters. The sine function simulation is not powerful enough to represent the apparent complexity of dissolved oxygen influence on other control parameters besides temperature and nitrate concentration. This simple model overestimates the dissolved oxygen saturation amplitude but reproduces the nycthemeral cycle well over the whole study period. The best predictions are provided by the XGB Regressor algorithm, with a good performance and small error intervals.

This modeling approach is emerging as a priority mitigation measure for restoring the condition of freshwater ecosystems in a watershed [23]. In addition, to test the predictive ability of the long time series of the model proposed in this study, time intervals of one hour was used. A long prediction time interval needs a much longer effective history, which is a challenge for the time series model [24].

## V. CONCLUSION

High-frequency online monitoring of water quality and hydrological parameters provided a wealth of data, which was analyzed and used to simulate the diel fluctuations of dissolved oxygen. The present study demonstrates that it is possible to develop a simple model such as sine function that allowed us to estimate the diurnal variability of oxygen in the river environment and thus eliminate the bias introduced by point measurements. This model can estimate the fluctuations of dissolved oxygen of the Sebou River during a nycthemeral cycle but overestimate its saturation rate. This study was evaluated the predictive ability of a machine learning technique using three algorithms: XGB Regressor, Random Forest Regressor, and Linear Regression. In addition, machine learning models can effectively accommodate complex time-series datasets and provide a robust prediction tool for dissolved oxygen content prediction.

This work has several limitations that require the acquisition of data over a long period to identify the seasonal and annual trends. Indeed, analysis of dissolved oxygen trends was very complex and influenced by many factors. It is not clear how factors such as water temperature and other quality parameters affect the performance of predictive models. Finally, using other feature extraction techniques such as convolutional neural networks can improve the predictive capability.

## ACKNOWLEDGMENT

The authors would like to express gratitude to all the members of the O'Sebou observatory and Special thanks to Agence de l'Eau Artois Picardie (Christine Deric), Agence du Bassin Hydraulique du Sebou, and Qualitas for their pioneering work that contribute to the successful implementation of the high-resolution water quality monitoring in the Sebou River

## REFERENCES

- [1] D. Barthelemy and M. Goubier, "Etude du cycle nyctheméral de la teneur en oxygène dissous en bassins de pisciculture par la méthode des moyennes mobiles. Relation avec la température et l'insolation", *Journal of Water Science*, 4(3), pp. 393–414, 1999. <https://doi.org/10.7202/705107ar>.
- [2] T.M. Tung Tiyasha, and Z.M. Yaseen, "A survey on river water quality modelling using artificial intelligence models: 2000–2020". *J. Hydrol.* 585, 2020. <https://doi.org/10.1016/j.jhydrol.2020.124670>.
- [3] M. H. Ahmed, L-S. Lin, "Dissolved oxygen concentration predictions for running waters with different land use land cover using a quantile regression forest machine learning technique". *Journal of Hydrology*, 597, 2021, p.126213. <https://doi.org/10.1016/j.jhydrol.2021.126213>.
- [4] M. Jancic, J. Kocijan and B. Grasi, "Identification of Atmospheric Variable Using Deep Gaussian processes IFAC". *Papers OnLine* 51-5, pp. 43–48, 2018.
- [5] V. Hull, L. Parrella, and M. Falcucci, "Modelling dissolved oxygen dynamics in coastal lagoons". *Ecological Modelling*, 211(3-4), 2008, pp.468-480.
- [6] J. Huang, S. Liu, S. G. Hassan, L. Xu, L., and C. Huang, "A hybrid model for short-term dissolved oxygen content prediction". *Computers and Electronics in Agriculture*, 186, 2021. p.106216.
- [7] C. Valerio, L. De Stefano, G. Martínez-Muñoz, and A. Garrido, "A machine learning model to assess the ecosystem response to water

policy measures in the Tagus River Basin (Spain)”. *Science of the Total Environment*, 750, 2021, p.141252.

- [8] S. El Mrissani, S. Haida, J.L. Probst, and A. Probst, “Multi-Indices Assessment of Origin and Controlling Factors of Trace Metals in River Sediments from a Semi-Arid Carbonated Basin (the Sebou Basin, Morocco)”. *Water* 13.22 2021, p. 3203.
- [9] H. Hayzoun, C. Garnier, G. Durrieu, V. Lenoble, C. Le Poupon, B. Angeletti, and S. Mounier, “Organic carbon, and major and trace element dynamic and fate in a large river subjected to poorly-regulated urban and industrial pressures (Sebou River, Morocco)”. *Science of the Total Environment*, 502, pp. 296-308, 2015
- [10] V. Villeneuve, S. Lègaré, J. Painchaud, and W. Vincent, “Dynamique et modélisation de l’oxygène dissous en rivière”. *Revue des sciences de l’eau / Journal of Water Science*, 19 (4), 2006, pp. 259–274.
- [11] R. Prasad, R.C., Deo, Y., Li , and T. Maraseni. “Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm”. *Atmos. Res.* 197, pp. 42–63. 2017. <https://doi.org/10.1016/j.atmosres.2017.06.014>.
- [12] C. Xu, X. Chen, and L. Zhang, “Predicting river dissolved oxygen time series based on stand-alone models and hybrid wavelet-based models”. *Journal of Environmental Management*, 295, 2021, p.113085.
- [13] G. Shobha, and R. Shanta. “Machine learning”. *Handbook of statistics*. Vol. 38. Elsevier, 2018. pp. 197-228. <https://doi.org/10.1016/bs.host.2018.07.004>
- [14] M.A. Razi, and K. Athappilly, “A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models”. *Expert Syst Appl*, 29, pp. 65–74, 2005.
- [15] D. Freedman, “Statistical Models: Theory and Practice”. Cambridge University Press, 2009.
- [16] F. Moatar, A. Poirel, and C. Obled, “Analyse de séries temporelles de mesures de l’oxygène dissous et du pH sur la Loire au niveau du site nucléaire de Dampierre (Loiret) : Compréhension des variations temporelles des teneurs en oxygène dissous et du pH en relation avec des données hydrométéorologiques”. *Hydroécologie Appliquée*, 11, 127-151, 1999,
- [17] V. Hull, C. Mocenni, M. Falcucci, and N. Marchettini, “A trophodynamic model for the lagoon of Fogliano (Italy) with ecological dependent modifying parameters”. *Ecol. Model.* 134, 2000, pp. 153–167.
- [18] S. Sabater, J. Armengol, E. Comas, F. Sabater, I. Urrizalqui, and I. Urrutia, “Algal biomass in a disturbed Atlantic river: water quality relationships and environmental implications”. *Sci. Total Environ.*, 263, 2000, pp. 185-195.
- [19] R.R.Chaudhury, J.A.H. Sobrinho, R.M. Wright, and M.Sreenivas, “Dissolved oxygen modeling of the Blackstone River (northeastern United States)”. *Water Res.*, 32,1998, pp.2400-2412.
- [20] B. Tasnim, J. A. Jamily, X. Fang, Y.Zhou and J.S. Hayworth, “Simulating diurnal variations of water temperature and dissolved oxygen in shallow Minnesota lakes”. *Water*, 13(14), (2021) p1980.
- [21] H. Guasch, J. Armengol, E. Mart, and S. Sabater, “Diurnal variation in dissolved oxygen and carbon dioxide in two low-order streams”. *Water Res.*, 32, 1998. Pp.1067-1074.
- [22] V. Hull, L. Parrella, and M. Falcucci, “Modelling dissolved oxygen dynamics in coastal lagoons”. *Ecological Modelling*, 211(3–4), 2008, pp.468-480.
- [23] C. Valerio, L. De Stefano, G. Martínez-Muñoz, and A. Garrido, “A machine learning model to assess the ecosystem response to water policy measures in the Tagus River Basin (Spain)”. *Science of the Total Environment*, 750, 2021, p.141252.
- [24] J. Huang, S. Liu, S. G. Hassan, L. Xu, L., and C. Huang, “A hybrid model for short-term dissolved oxygen content prediction”. *Computers and Electronics in Agriculture*, 186, 2021. p.106216.