Machine Learning Techniques for forecasting the effect of Climate Change and Anthropogenic Pressures on Coastal Wetlands (Ichkeul Lake, Ramsar site)

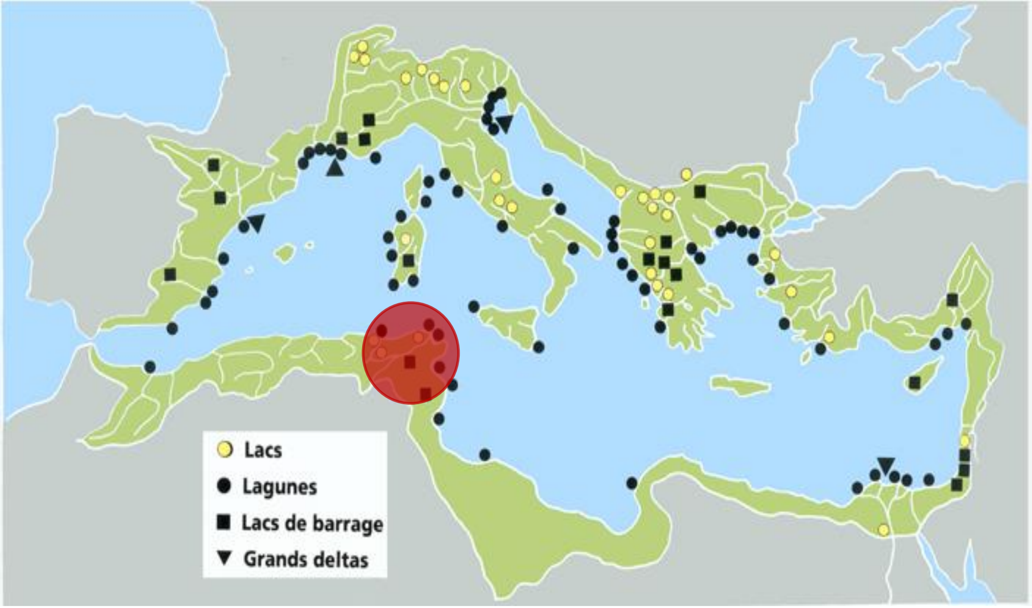
Presented by: Sabrine SAHBANI

Supervised by:

Béchir BEJAOUI , Ennio OTTAVIANI, Sihem Benabdallah, Eva RICCOMAGNO , Enrico PRAMPOLINI , Hechmi MISSAOUI

1. Introduction

- Wetlands have undergone several disruptions, which have been threatening their ecological status , their biodiversity, and their socio-economic services.
- Necessity to assess their sensitivity to anthropogenic pressures and climate change.



Mediterranean Wetlands

2. Study Site

Situation

North of Tunisia

Morphological Characteristics

An area of 133 km² with three units: Lake, marshes, and wooded massive (Djebel Ichkeul).

Continental Hydrology

The catchment (2100 km²) drains a developed network of six rivers.

Connection with the Bizerte Lagoon

The Lake communicate with the Bizerte Lagoon.

Ecological value

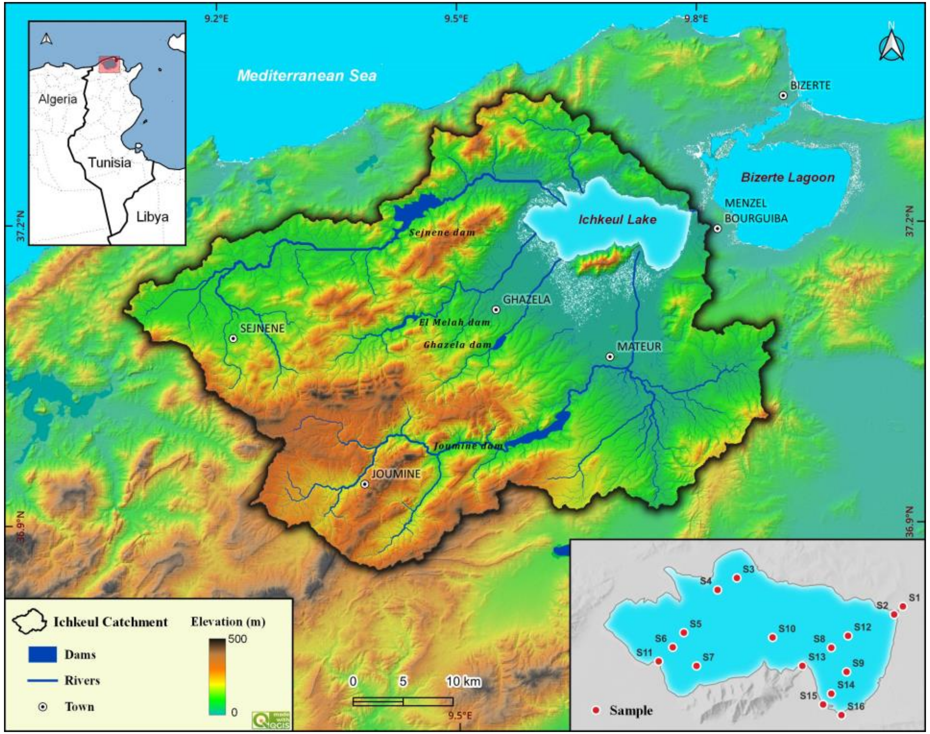
The most productive environment in Tunisia.

Registration in three international conventions.

Ecological fragility

Climate change: increase in T and decrease in P.

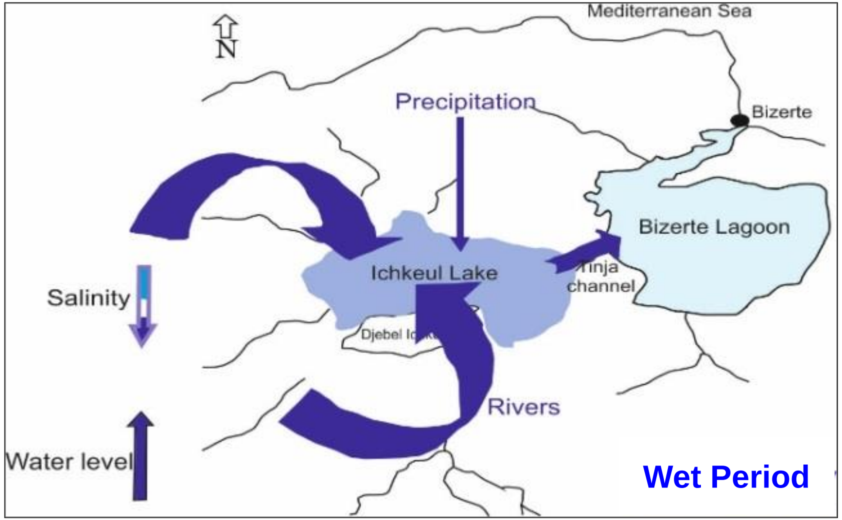
Anthropogenic pressures: construction of dams and locks, over-exploitation of resources, pollution...

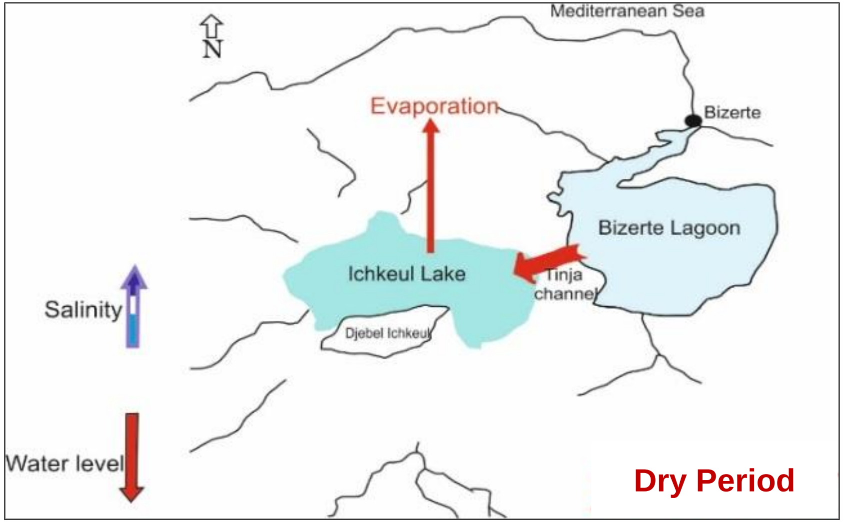


Localisation of Ichkeul Lake and sampling stations

2. Study Site

- The water budget in the Ichkeul Lake is characterized by seasonal variation:





In wet period:

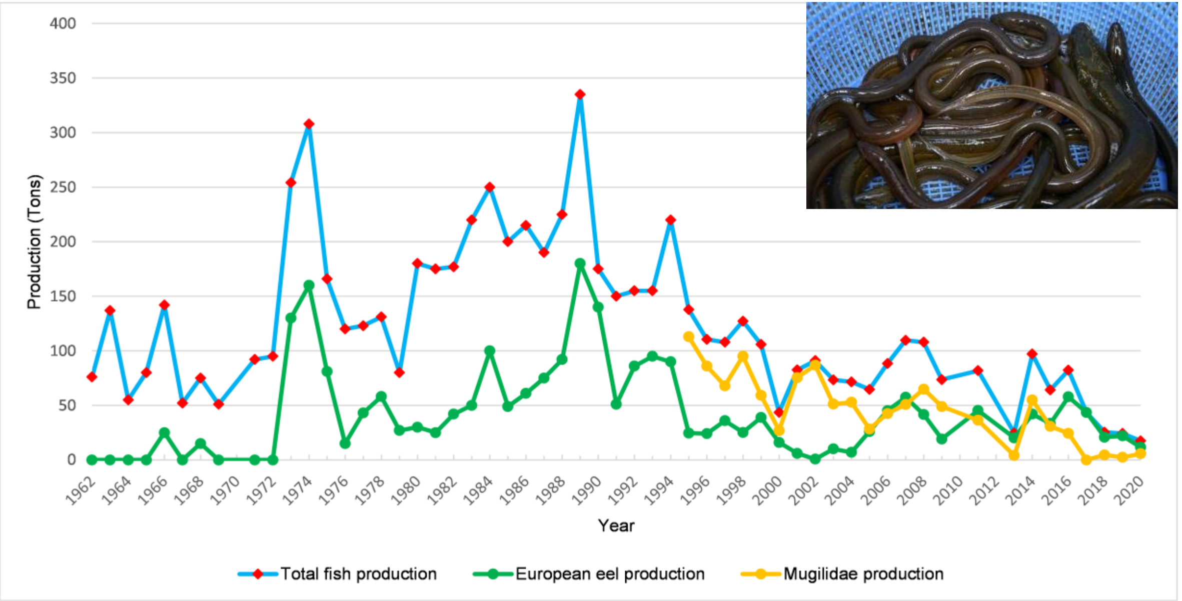
- **Water leaves Ichkeul Lake,**
- **High rainfall,**
- **High water level / low salinity.**

In dry period

- **Water spills off to the Bizerte Lagoon,**
- **High evaporation,**
- **High salinity/low water level.**

2. Study Site

- The fishery resources are threatened by :
- The collapse of some fish species (*Lithognathus mormyrus*...),
- Decrease in stocks of other species (Mugilidae and *A. anguilla*): Eel production has been reduced from 86 tons for the period 1985-1995 to 32 tons for the period 2010-2020.

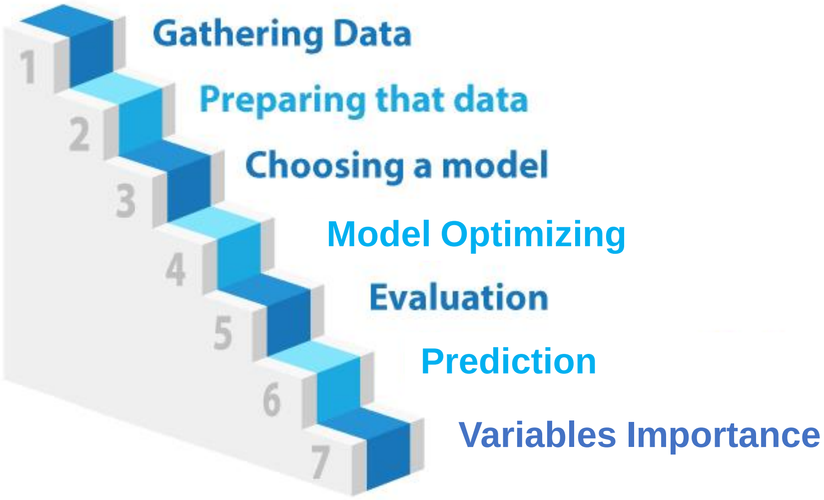


Evolution of European eel and Mugilidae production in relation to total ichthyofauna production in Lake Ichkeul between 1962 and 2020.

3. Objectives

- Evaluate the effect of Climate Change and anthropogenic pressures on European eel yield by developing a predictive approach to determine the most important environmental factors influencing the presence of the species.

7 steps were followed



4. Data Collection

- Dataset = 142 samples of 13 variables (from 2010 to 2020)
- Data sources: From BASSIANA database, and field monitoring.
- Abiotic parameters:
- Meteorological parameters: P, and W.
- Physico-chemical parameters: T, WL, S, DO and Tur.
- Chemical parameters: DIN,DIP, TN.
- Biotic parameter: Chl.a, and Eels landing

5. Data Analysis

- missForest,
- Box-Cox transformation,
- Random Forest model and,
- Cubist model.

Summary of environmental parameters in Ichkeul Lake for the period 2010-2020

Parameter	**Abbreviation and Unit**	Mean	
Period	Pr	Dry	Wet
Precipitation	P (mm)	15.96	82.61
Wind Intensity	W ($m.s^{-1}$)	6.01	5.22
Temperature	T (°C)	21.11	15.39
Water Level	WL (cm)	44.29	67.44
Salinity	S (psu)	44.68	21.42
Dissolved Oxygen	DO ($mg. l^{-1}$)	7.29	7.13
Turbidity	Tur (NTU)	18.9	27.34
Total Nitrogen	DIN (µM)	22.05	29.46
Dissolved Inorganic Nitrogen	TN (µM)	22.66	20.97
Dissolved Inorganic Phosphorus	DIP (µM)	2.08	1.06
Chlorophyll a	Chl.a ($\mu g. l^{-1}$)	6.8	3.36
Eels landing	Eels (tons)	16.84	38.34

5. Data Analysis

5.1. missForest :

- To deal with missing values NA (proportion of missingness = 5 %),
- Prediction of the NA from the non-missing value available from the dataset,
- Performance Evaluation: the Normalized Root Mean Squared Error NRMSE:

$$NRMSE = \sqrt{\frac{mean\left(\left(X^{true} - X^{imp}\right)^2\right)}{var\left(X^{true}\right)}}$$

Where:

X^{true} is the complete data matrix

X^{imp} is the imputed data matrix.

5.2. Box-Cox transformation:

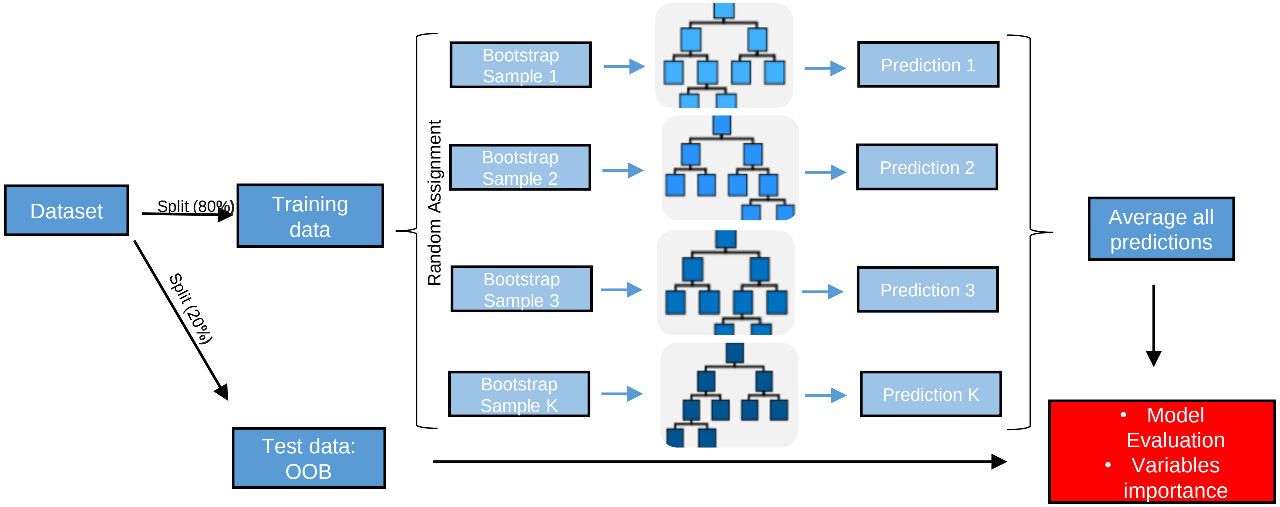
- To conform to the requirement for normality,
- Assessed using Shapiro–Wilk test.

5. Data Analysis

5.3. Machine Learning models

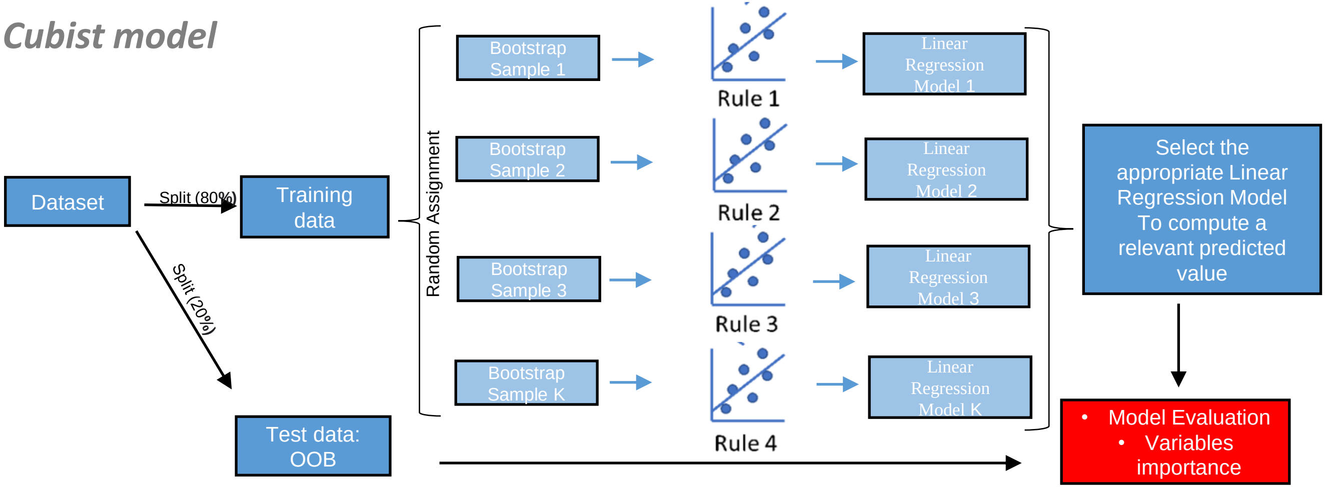
- Evaluate the relationship between the target (Eels landing) and predictors (abiotic parameters).

a. Random Forest model



5. Data Analysis

b. Cubist model



c. Models Optimizing

- Hyperparameters selected for tuning
- 3-folds Cross-Validation Method (for Cubist model) and OOB procedure (for Random forest model): used to determine the final optimum hyperparameters of the models.

Model	Hyper-parameter	Description
RF	*min.node.size*	Number of trees for the forest
	m_{try}	Number of predictors selected at each split
Cubist	neighbors	Number of instances
	committees	Number of rules set

5. Data Analysis

c. Performance metrics

- Coefficient of Determination $$R^2 = \frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y^{-obs}\right)^2 - \sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{pred}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{obs} - Y^{-obs}\right)^2} \in [0,1]$$
- Mean Absolute Error $$MAE = \frac{\sum_{i=1}^{n}\left|Y_i^{obs} - Y_i^{pred}\right|}{n} \in [0,+\infty]$$
- Root Mean Square Error $$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y^{-obs}\right)^2}{n}} \in [0,+\infty]$$

Where

n: the total number of data,
Yi^{pred} : the predicted eels landing of i observation,
Yi^{obs} : the measured eels landing of i observation,
Y^{-obs} : the mean of all observed responses.

- A good model prediction was expected to have low MAE and RMSE (close to 0) as well as an R^2 value close to 1.

IEEE

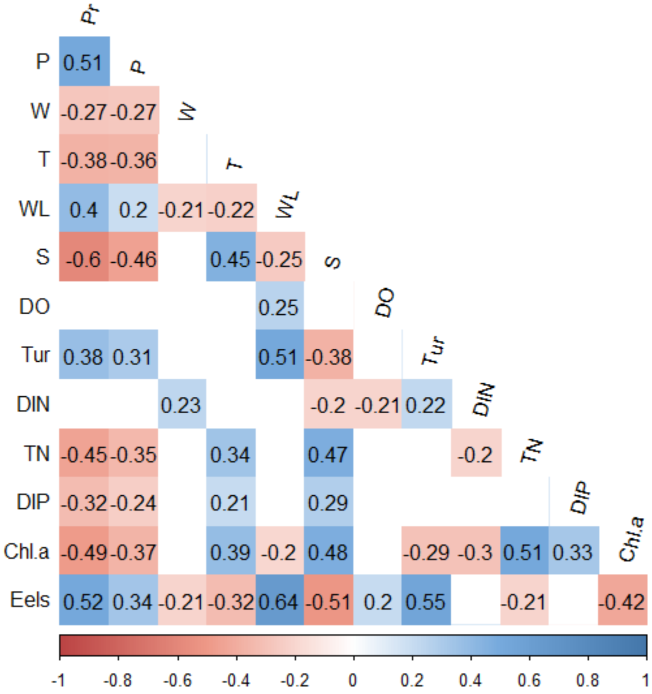
6. Results & Discussions

6.1. MissForest algorithm:

- NRMSE=0.22, indicating a sufficiently good performance

6.2. Pearson Correlation:

- Eel landing was positively correlated with water level (r = 0.64) and turbidity (r = 0.55),
- It was negatively correlated with salinity (r = -0.51) and chlorophyll a (r = -0.42).

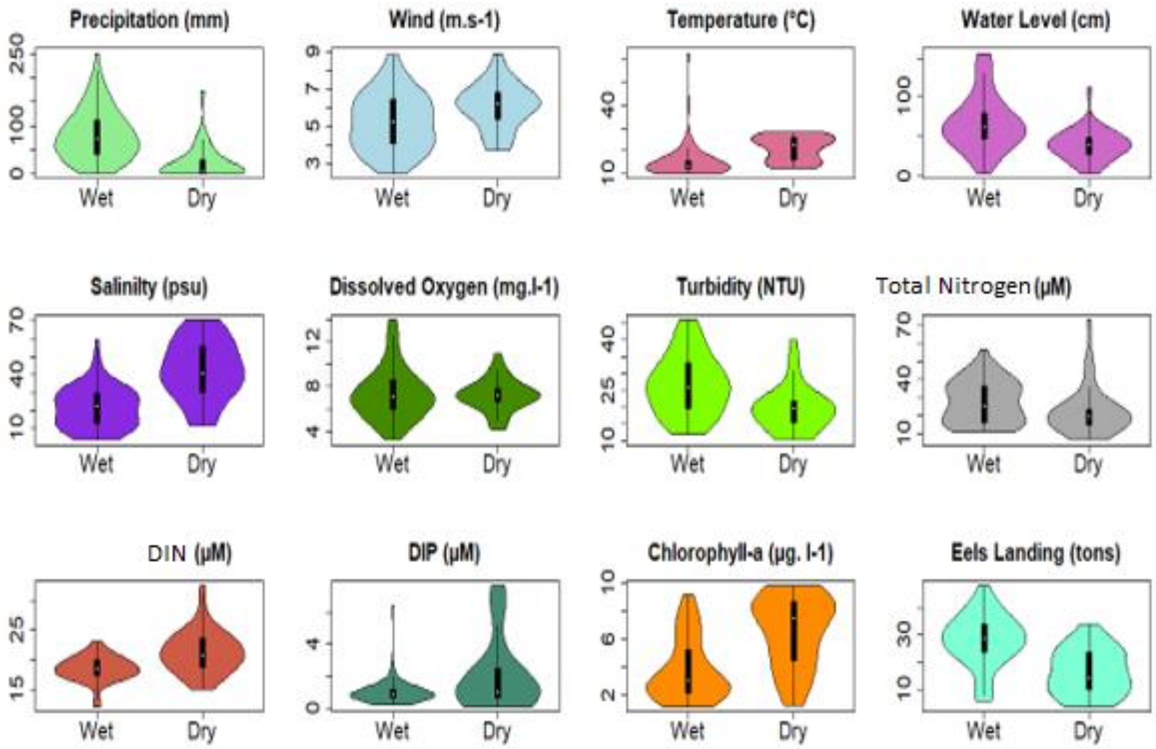


Correlation matrix plot of the database

6. Results & Discussions

6.3. Parameter's properties

- Seasonal variability of environmental parameters.
- The water turbidity was high throughout the year due to the effect of meteorological and morphological characteristics.
- High levels of TP, DIP,TN and DIN during the period of study and,
- A clear variations between the seasons for the Eels landing.

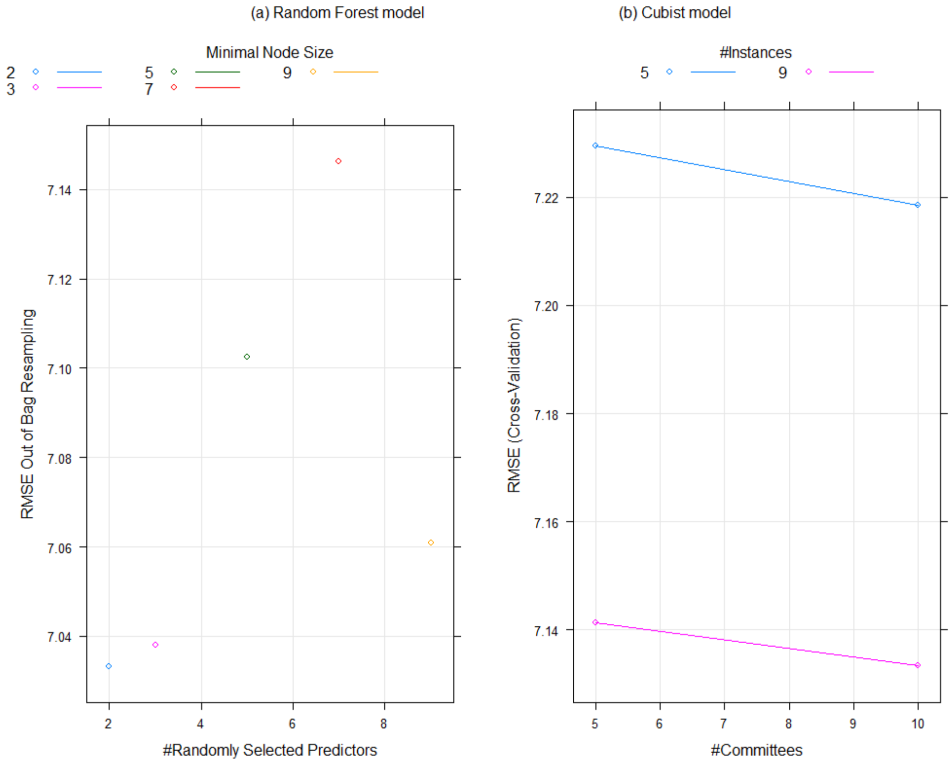


Data visualization with box and violin plots (VIP)

6. Results & Discussions

6.4. Results of hyper-parameters tuning

- For the RF model:
  - The best tuning values of mtry equal to 2 and min.node.size equal to 2,
- The CB model
  - The optimal values were committees equal to 10 and instances equal to 9.



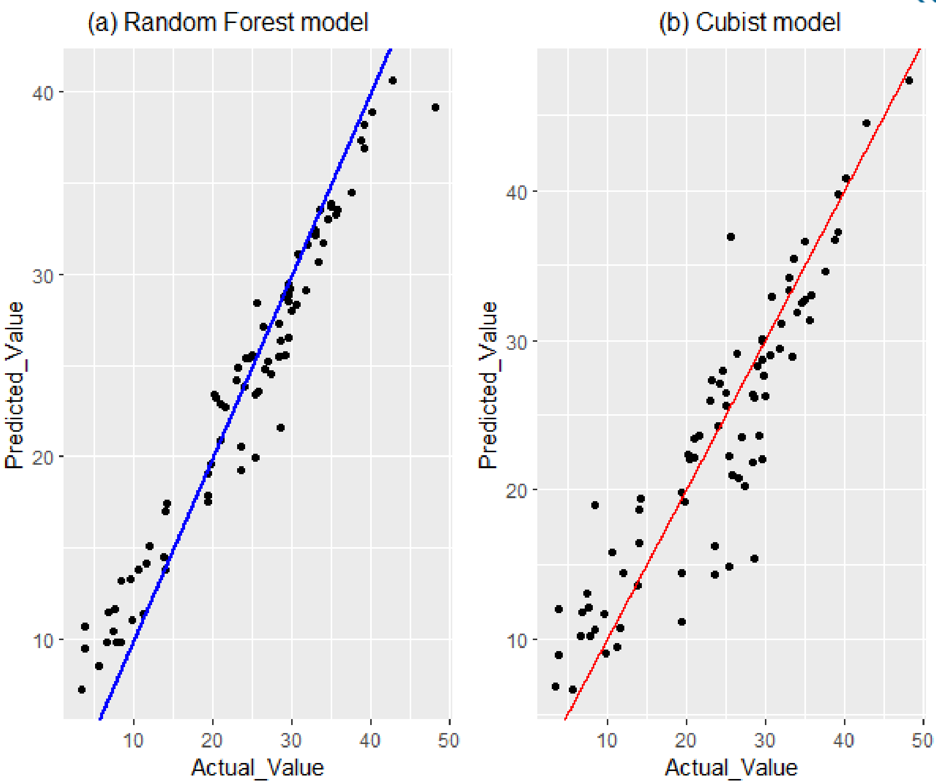
OOB procedure and 3-folds cross-validated RMSE profiles for determining the optimal tuning parameters for (a) RF model and (b) Cubist model

6. Results & Discussions

6.5. Results of performance matrics

Efficiency metric	RF model	Cubist model
R^2	0.55	0.56
MAE	5.69	5.51
RMSE	6.92	6.81

- RF and Cubist models are suitable techniques for predicting different pressures on Eels landing.

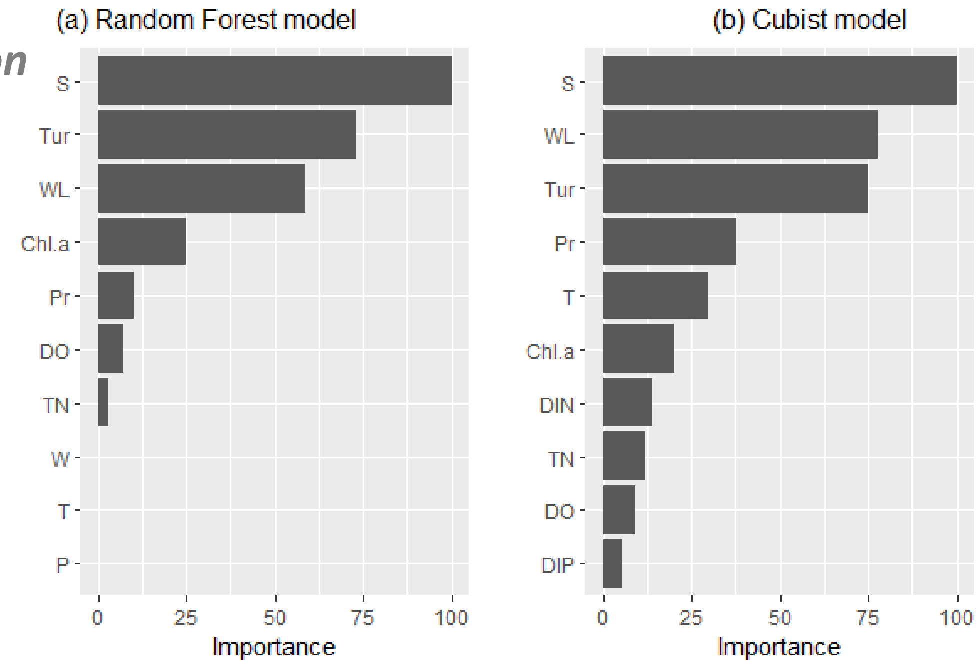


Scatter plot of the observed and the predicted eels landing using (a) the RF model and (b) the Cubist model

6. Results & Discussions

6.6. Variable importance and interpretation

- According to the RF model, the most important predictors are Salinity followed by Turbidity and Water level.
- According to the Cubist model, the most important predictors are Salinity followed by Water level and Turbidy.
- In Ichkeul Lake, the most important predictors of the quantity of European eel appear to be Salinity,Water level and Turbidity.
- Water level and turbidity promote eel migration and facilitate the foraging process, while the salinity plays a key factor during the cycle life of European eels (Lagarde et al., 2021).



Variable importance scores for the 13 predictors in the proposed models for Eels Landing: (a) random forests model, and (b) Cubist model

7. Conclusion

- **Based on the numerical approach:**

  ● **Climate change and anthropogenic pressures have affected the functioning of the lake.**

  ● **Increase of the nutrients loading, the risk of episodic eutrophication, and thus the decrease of Eels population and of biotic resources.**

- **The results of this study suggest that local management agencies can use these smart technologies in the monitoring system of the trophic and ecological status of the lake, as they offer a reliable and efficient means of maintaining its ecological conditions in the future.**

- **Actions required:**

  ● **For the management of the Ichkeul Lake wetland, the key to success lies in maintaining water quality to support the growth and survival of aquatic organisms through improved water management and depollution.**

Thanks for your attention

Sabrine SAHBANI:
National Institute of Marine Science & Technologies.

E-mails:
sahbenisabrine0@gmail.com
sabrine.sahbani@inat.u-carthage.tn

Béchir BEJAOUI:
National Institute of Marine Science & Technologies.

E-mail:
bejaoui.bechir@gmail.com