Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2023.30007

312 (2023) 108–118
November

Application of binary statistical model for the environmental risk assessment of metal contamination through aquatic organisms

Soumia Ferroudj[a,*], Sabrine Boucetta[b,c,*], Zouina Doukha[d], Soumia Kharfouchi[e,*], Mohamed Khelil Aouati[f], Wafa Benchalel[a]

[a]*Ecology of Terrestrial and Aquatic Systems Laboratory (EcoStaq), Badji Mokhtar Annaba University, Annaba, Algeria, email: soumia.ferroudj@yahoo.com (S. Ferroudj)*
[b]*Department of Nature and Life Sciences, University of August 20, 1955 Skikda, Algeria, email: s.boucetta@univ-skikda.dz (S. Boucetta)*
[c]*Laboratory of Ecobiology of Marine and Coastal Environments (EMMAL), Badji Mokhtar Annaba University, Annaba, Algeria*
[d]*RIIMA Laboratory, Computer Science Faculty, University of Science and Technology Houari Boumedien*
[e]*Applied Research Laboratory in Medical and Social Sciences (RASMS), Faculty of Medicine, Constantine 3 University, Algeria, email: s_kharfouchi@yahoo.fr (S. Kharfouchi)*
[f]*Department of Physics and Chemistry, Mentouri Constantine 1 Brothers University, Constantine, Algeria*

Received 16 April 2023; Accepted 27 September 2023

ABSTRACT

This study presents a statistical model that assesses pollution in aquatic ecosystems based on heavy metal concentrations (dry weight). The model uses binary logistic regression to analyze factors like species, stations, seasons, and heavy metals associated with pollution. The method constructs a binary interest factor ("polluted" and "unpolluted") based on World Health Organization standards. The results help to understand risks to human health from heavy metals in coastal areas of East-Algeria and provide a useful tool for monitoring pollution.

Keywords: Heavy metals; Human health; Risk assessment; Biomonitoring molluscs algae; Regression models

1. Introduction

Minerals, essential fatty acids, and proteins are essential nutrients of people's health [1]. Seafood is one of the main sources of these nutrients. However, seafood can also be a significant source of heavy metal contaminants [2]. Heavy metals have the ability to bind to short carbon chains [3] and can be bio-accumulated and easily assimilated by organisms. Some heavy metals, such as nickel, lead, chromium, and cadmium, are particularly toxic even at small levels [4].

In the marine environment, mollusks and algae are frequently used for the biomonitoring of heavy metals pollution [5,6]. The choice of the species used for biomonitoring is constrained by specific requirements: the species must have a significant concentration of the pollutant being studied, must be widely distributed geographically for comparison across different locations, must ingest the pollution without dying from the concentrations it encounters, and must have a pollutant concentration that can be easily connected to the mean concentration in the surrounding environment.

Metals can be accumulated in organisms through soluble fractions, particulates, sediments in seawater, and food. Certain species of mollusks have specialized metabolic defense mechanisms that allow them to manage the amounts of these pollutants in their tissues [7]. In such situations, it is challenging to determine whether the variations reported at the sampling sites reflect actual environmental pollution

* Corresponding authors.

1944-3994/1944-3986 © 2023 Desalination Publications. All rights reserved.

or are influenced by the sample size [8]. Therefore, evaluating the amounts of heavy metals in mollusk tissues is crucial due to the non-biodegradability of these metals.

Algae are well-suited organisms for studying heavy metal contamination in aquatic ecosystems as they are abundant, adaptable to environmental conditions [9,10] and can accumulate high quantities of heavy metals, serving as a sink for these pollutants [11].

In this study, we present a statistical model capable of assessing pollution in terms of metal concentration dry weight (DW). We aim to identify factors associated with the pollution phenomenon, including various potential explanatory variables such as species, stations, seasons, and metals. The study seeks to answer the following questions: (i) Which species are most associated with the pollution phenomenon?; (ii) Does the season play a role in the phenomenon?; (iii) Which station is the most polluted?; (iv) What is the metal that pollutes the most?. This information, combined with a bottom-up mechanical manner linking the concentrations distribution to physiological processes, can lead to a precise evaluation of existing risks. Our database consists of five qualitative variables described as follows:

- Stations: 4 level factors: St.1, St.2, St.3, and St.4, where St.4 is a control station.
- Seasons: 4 level factors: winter, spring, summer, and autumn.
- Species: 4 level factors: *Patella caerulea*, *Ulva lactuca*, *Stramonita haemastoma* and *Phorcus turbinatus*.
- Heavy metals: 6 level factors: zinc (Zn), nickel (Ni), lead (Pb), chromium (Cr), copper (Cu), and cadmium (Cd), where metal concentrations are reported on a dry weight (DW) in $mg \cdot g^{-1}$.
- Pollution: to evaluate the degree of pollution, a binary interest factor ("polluted" and "unpolluted"), based on (DW) measurements through the use of well-established World Health Organization (WHO) standards, is built.

The "pollution" factor is coded as follows: If the DW measurement indicates at least one of the WHO and/or International Atomic Energy Agency (IAEA) risks *T* (toxic for health established by WHO), *N* (dangerous for environment established by WHO), X_n (harmful established by WHO), *C* (corrosive established by WHO), X_i (irritant established by WHO) or *D* (exceed the norm established by IAEA) (indicating the presence of pollution from various sources) the factor is set to 1, if the DW measurement indicates *A* (admissible established by WHO) or *R* (recommended established by IAEA) (indicating the absence of pollution), the factor is set to 2.

Thus, "pollution" becomes the variable of interest (i.e., the factor to be explained) in our analysis. We will explore how the factors "stations", "species", "seasons", and "heavy metals" influence the presence of "pollution" in the aquatic ecosystem. By understanding these relationships, we aim to gain insights into the potential risks posed by "heavy metals" contamination to human health and the environment. In the context of the study on pollution assessment using bio-monitoring tools, we can provide a scientific basis to justify the use of the independent variables season, stations, heavy metals, and different organism (species) in the following way:

- Season scientific basis: Seasonal variations can influence the accumulation and distribution of heavy metals in aquatic ecosystems. Factors like temperature, precipitation, and biological activities can affect metal concentrations. We can include a categorical variable for season with four levels: winter, spring, summer, and autumn.
- Stations scientific basis: Different sampling stations can represent distinct environmental conditions and potential pollution sources, leading to variations in metal concentrations. We can include a categorical variable for stations with four levels (St.1, St.2, St.3, St.4).
- Heavy metals scientific basis: Different metals may have different sources of contamination and behavior in the aquatic environment, leading to variations in metal concentrations. We can include a categorical variable for metals with six levels (Zn, Ni, Pb, Cr, Cu, Cd).
- Species scientific basis: Different organisms may have varying abilities to accumulate and eliminate heavy metals, making them suitable bio-indicators for pollution assessment. We can include a categorical variable for species with four levels (*P. caerulea*, *U. lactuca*, *S. haemastoma*, *P. turbinatus*).

In this regard, a binary logistic regression, on the collected data, is adopted. The binary logistic regression was first introduced in the 1970s to make up for the shortcomings of ordinary least squares (OLS) regression to process binomial outcomes [12]. The best-fitting function is presented depending on the maximum likelihood (ML) approach [13]. It has proven to be an effective tool for interpreting ubiquitous biomarker data, performing new studies to answer specific risk-based questions. It also allowed the integration of the maximum amount of information on the collected data which leads to a better understanding of human health risks [14]. Particularly, in the present study the reasons for binary logistic regression choice are summarized in the following:

- The variable "pollution" to be explained is qualitative.
- There is a sufficient number of events (presence of pollution) against the number of explanatory variables. Indeed, the number of recorded pollution cases is significantly higher than the number of explanatory variables that are significantly associated with the explained variable (the general rule is to have at least 5 to 10 events per explanatory variable).
- The bioaccumulation of heavy metals by aquatic organisms is considered "similar".

2. Materials and methods

2.1. Study area

Biota were sampled at four stations along the Algerian East coast (36°56′44.19″N 6°15′39.08″E) and (36°50′38.18″N 7°49′39.38″E) over a distance of almost 350 km as shown in Fig. 1. The stations St.1 (Bay of Collo: 37°00′07.50″N 6°34′39.93″E) and St.2 (Gulf of Skikda: 36°53′32.33″N 6°53′12.21″E) are near from Collo and Stora Port. They are characterized by intense maritime traffic. Furthermore, these sites are exposed to pollution by PAH due to the presence of a large petrochemical complex [15]. The station St.3

(Gulf of Annaba: 36°55′31.46″N 7°45′44.43″E), is exposed to pesticides and/or heavy metals emitted by the Fertial Factory and Port Operations [16] and it receives organic pollutants, in particular discharges of domestic wastewater [17]. Finally, station 4 (St.4) (Bay of Chetaibi: 37°02′26.27″N 7°24′22.04″E), being far from any anthropogenic, is regarded as the control station [5].

2.2. Sampling and samples pretreatment

The study concerns four stations in the eastern coasts of Algeria (St.1: El Djerda beach, St.2: Military beach, St.3: Rizi Amor beach and St.4: Oued El Ganem beach). The six considered heavy metals are: Cd, Cr, Cu, Pb, Zn and Ni. The species included in the study are: "*P. turbinatus*" Born 1778 (n = 161), "*P. caerulea*" Linné 1758 (n = 244), "*S. haemastoma*" (n = 90) and the algae "*U. lactuca*" (L.) Thivy 1960 (n = 84). Samples were collected in two field trips from Dec. 2011 to Sep. 2019, a month with the greatest average metal contents in adult mollusks [8]. Mollusks were carefully gathered, and only mature specimens within a rather restricted range of size (and weight) were chosen; the same method was followed for station samples. All of these factors contribute to a high amount of heavy metal [18]. Before the examination, algal samples were rinsed with saltwater that had been filtered, transported to polyethylene bags, and then frozen at –20°C. For the algae study, only mature leaves and thalli of comparable length were chosen [7]. Afterwards, samples were combined and subsamples of 0.7 g DW were microwave-mineralized (MDS 2000; CEM, Italy) with ultrapure HNO_3 and H_2O_2 (6 + 2 mL; Merck, Germany) [18,19]. Every month, an average of twenty adult individuals of Monodonta, *Patella* and rapa whelks belonging to size class (30 and 40) mm, are manually collected in the intertidal zone. For 24 h, the collected species were immersed in filtered seawater for purification, to purify the mantle cavity and the digestive tract of the particulate matter residues [19,20].

Subsequently and in order to avoid metal contamination, the extraction of the soft parts from the shell has been achieved using a spatula and plastic hammer, and then washed with deionized Milli-Q water. Once every residue of shell was removed, samples have been kept frozen inside polyethylene bags. Microwave digestion has been applied individually to all the mollusk for their analysis [21]. All microwave-assisted mineralizations of algae were carried out. Separate DW assessments were carried out on the various biota by oven drying them to a constant weight at 105°C (20 replicates for each species). All chemicals utilized in sample treatments were of ultrapure quality. A Millipore Milli-Q system supplied the water used for solution preparation and cleaning. Before use, all glassware was cleaned by soaking in 10% HNO_3 for 24 h and rinsing with Milli-Q water [21] Table 1. For soluble metal analysis, laboratory

Table 1
Limits of detection (LoD)[a] and precision (CV)[b] for analyses conducted by the two techniques used [8]

Metal	Algae		Mollusks	
	(μg·g^{-1} DW)	CV	(μg·g^{-1} DW)	CV
Cd	0.04	2.7	0.05	4.3
Cr	0.71	3.9	0.94	3.6
Cu	0.006[a]	4.5	0.01[b]	4.5
Pb	0.07	2.0	0.09	1.4
Zn	0.13[c]	2.5	0.09[d]	2.5

[a]Calculated on the basis of 10 determinations of blanks as three lines the standard deviation of the blank.
[b]Percentage referred to 10 determinations performed on the same sample.
[c]LoD for Cu and Zn was obtained by flame atomic absorption spectrometry (mg·L^{-1}).
[d]LoD for Cu and Zn was obtained by flame atomic absorption spectrometry (mg·L^{-1}).

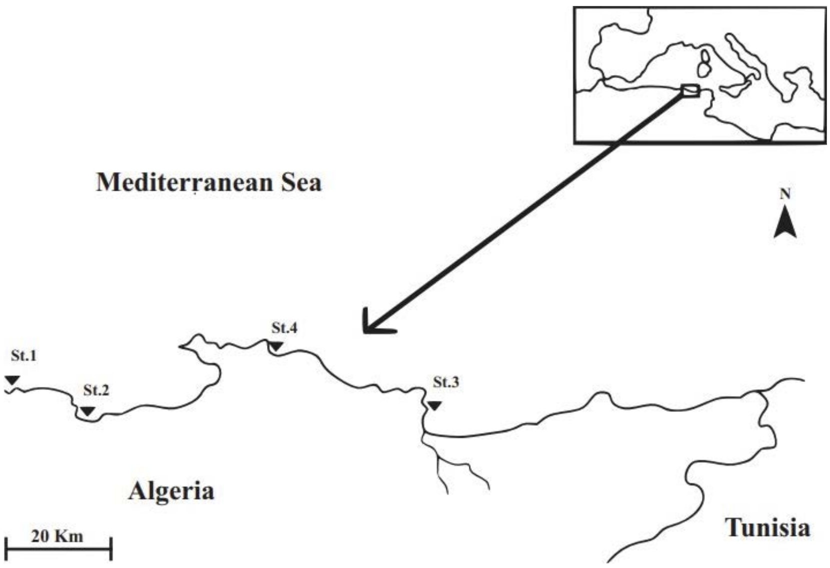


Fig. 1. Study area: (St.1: El Djerda beach, St.2: Military beach, St.3: Rizi Amor beach, and St.4: Oued El Ganem beach) during (Dec. 2011–Dec. 2019).

samples were filtered through an acid-precleared 0.45-lm membrane filter, acidified, and kept at 4°C [21] Table 1.

2.3. Trace metal determination and quality control

Measurements of trace metal concentrations have been taken on biota using a PerkinElmer AAnalyst 300 atomic absorption spectrometer with graphite furnace system (GFAAS) with HGA-800 autosampler and flame atomization (FAAS). FAAS was employed for Cu and Zn, whereas GFAAS with an HGA-800 system was used for Cd, Cr, and Pb. Matrix modifier NH_4-H_2PO_4 at 10% with 0.2 mg of PO_4 was used for Cd and Pb, and $Mg(NO_3)_2$ was employed for Cr; the method of standard additions for calibration was applied [21]. Ammonium pyrrolidine dithiocarbamate (APDC) and methyl isobutyl ketone (MIBK) were used to determine the concentration of dissolved metals [21]. 1% aqueous APDC solution (Eastman Chemical) was prepared each day and purified using an equal volume of MIBK, leading to separating phases. As metal complexes are highly soluble while APDC is only slightly soluble in MIBK, it is easy to purify the reagent in this manner. No further purification has been made throughout for reagent grade MIBK [8]. The control of the accuracy of the whole analytical procedure has been done using certified reference materials: CRM 279 (sea lettuce), ERM-CE 278 (mussel tissue), and CASS3 (Nearshore Seawater; NRCC, Ottawa, Canada). Tables 1 and 2 detail the analytical performance of the employed techniques in terms of limits of detection and precision [18,22].

2.4. Statistical analysis

We aim to study the effect of the factors "species", "stations", "seasons" and "metals" on the presence of pollution. For this purpose, we use an ordinary logistic regression model [23] Subsequently, we first build a binary interest factor ("polluted" and "unpolluted") based on (DW) measurements through the use of well-established WHO joint [24]. The "pollution" to be explained.

The risk ranking for seafood and shellfish according to the reference [24] (food codex) [25] is coded by: *T* (toxic for health), *N* (dangerous for the environment), X_n (harmful), *C* (corrosive), X_i (irritant) and *A* (admissible). The risk classification for the TME (trace metal element) according to the IAEA [26] standard for algae (*U. lactuca*) is coded by: *R* (recommended) and *D* (exceeds the norm).

Thus, we construct the factor "pollution" which will be the variable of interest (i.e., to be explained) that will be equal to 1 if the DW measurement indicates at least one of the risks *T*, *N*, *D*, X_n, *C*, or X_i (testifying to the presence of pollution of various origins) and equal to 2 if it has *A* or *R* (absence of pollution). Therefore, the factors "stations", "species", "seasons", "metals" will be our explanatory variables.

Both logistic regression and chi-square tests are powerful statistical techniques used to analyze categorical data and relationships between variables. In our study, they are commonly applied. Here are the steps for conducting logistic regression and chi-square tests.

2.4.1. Logistic regression

- Data preparation: In this step, data are organized and cleaned.
- Data splitting: Here, the data is split into training set and a testing/validation set. The training set will be used to build the logistic regression model, while the testing set will be used to evaluate its performance.
- Model building: The training set is used to fit a logistic regression model. This involves estimating the coefficients for each independent variable to predict the probability of the dependent variable being 1.
- Model evaluation: The performance of the logistic regression model is evaluated using the testing set. Common evaluation metrics include accuracy, precision, recall, and F1-score.
- Interpretation: Coefficients of the logistic regression model are interpreted to understand the relationship between the independent variables.

2.4.2. Chi-square tests

- Data preparation: Data should be cleaned and organized in a contingency table format.
- Null hypothesis: The null hypothesis should be formulated, stating that there is no significant association between the variables.
- Degrees of freedom: The degrees of freedom are determined for the chi-square test. For a contingency table with *r* rows and *c* columns, the degrees of freedom is $(r-1) \times (c-1)$.
- Expected frequencies: Calculate the expected frequencies for each cell of the contingency table under the assumption of the null hypothesis.
- Chi-square statistic: The chi-square statistic is computed by comparing the observed frequencies in the contingency table to the expected frequencies.
- P-value: The *p*-value associated with the chi-square statistic is calculated. It indicates the probability of observing the data or more extreme results under the assumption of the null hypothesis.
- Conclusion: The *p*-value is compared to a significance level (e.g., 0.05). If the *p*-value is less than the significance level, reject the null hypothesis and conclude that there is a significant association between the variables.

Table 2
Analysis of certified reference materials: certified and found values (mean ± SD) [8][a]

Metal	CRM 279 (sea lettuce) (μg·g⁻¹ DW)		ERM-CE 278 (mussel tissue) (μg·g⁻¹ DW)	
	Certified	Found	Certified	Found
Cd	0.274 ± 0.022	0.272 ± 0.031	0.348 ± 0.007	0.340 ± 0.011
Cr	(10.7)[a]	10.75 ± 0.64	0.78 ± 0.06	0.79 ± 0.03
Cu	13.14 ± 0.37	12.54 ± 0.59	9.45 ± 0.13	9.40 ± 0.21
Pb	13.48 ± 0.36	13.41 ± 0.22	2.00 ± 0.04	2.10 ± 0.05
Zn	51.3 ± 1.2	53.6 ± 1.5	83.1 ± 1.7	82.1 ± 1.6

[a]Not certified values.

2.5. Logistic regression

In a multiple logistic regression analysis, we include multiple independent variables simultaneously to assess their combined impact on the dependent variable (pollution). In our context, the multiple logistic regression equation can be expressed as follows,

Logit(pollution) = $\beta_0 + \beta_1 \times$ (Season.winter) +...

+$\beta_4 \times$ (Season.autumn) + $\beta_5 \times$ (Station.St.1) +...

+β8 × (Station.St.4) + β9 × (Metal.Zn) +...

+β14 × (Metal.Cd) + β15 × (Species.Patella) +...

+β18 × (Species.Phorcus)

The logistic regression analysis will estimate the values of these coefficients based on the data, allowing us to quantify the impact of each independent variable on the likelihood of pollution occurrence (Logit(pollution)). By statistically analyzing the relationships between the independent variables and pollution, we can draw meaningful conclusions about the factors influencing pollution levels in the studied aquatic ecosystem. Indeed, each beta coefficient represents the impact of the corresponding independent variable on the likelihood of pollution occurrence, holding other variables constant. For example, if we estimate that β_1 is positive and statistically significant, it means that the likelihood of pollution occurrence is higher during winter compared to the reference season (e.g., autumn). Similarly, if β_5 is negative and significant, it indicates that station St.1 is associated with a lower likelihood of pollution compared to the reference station (e.g., St.4, which is the control station). The same interpretation can be applied to the coefficients related to metals and species. By conducting the multiple logistic regression analysis on the data, we can obtain the estimates of these coefficients along with their standard errors, *p*-values, and confidence intervals. This information will help us to understand the relative importance of each independent variable in predicting pollution levels in the coastal areas of East-Algeria as well as analyzing the correlation of this analysis technique with pollution-related factors.

Moreover, all the coefficients are determined relative to the reference modality, both for the explained variable and for the explanatory variables. For the explained variable, the reference modality is that of the "no event", in our case "absence of pollution". For the explanatory variables, the choice is made by using the subject knowledge and the way we present the results. Given the objectives initially set, we choose to reference the modality that seems to be the weakest for each variable. We then perform a chi-square test applied to distributions according to the variable "pollution" and each of the explanatory variables. The different tests will allow us to compare the numbers of 1 (presence of pollution) and 2 (absence of pollution) of the different groups of modalities of the explanatory variables in order to verify the hypothesis according to which, within the population, the frequency of 1 is different from the frequency of 2. Therefore, the modalities can be classified according to their importance in each case. The chi-square test results are given in Table 3.

3. Results and discussion

Regarding the variable "seasons", the modality "1" appeared the least frequent with the modality "spring",

Table 3
Chi-square tests between the factor "pollution" and each of the four variables

	Cross-tabulation 'Pollution' ~ 'Heavy metals'								
	Cd	Cr	Cu	Ni	Pb	Zn	χ-squared	df	ρ-value
1	304	118	173	257	228	107	409.43	5	<2.2e^{-16}
2	16	202	147	63	92	213			
	Cross-tabulation 'Pollution' ~ 'Species'								
	Patella caerulea	*Phorcus turbinatus*	*Stramonita haemastoma*	*Ulva lactuca*			χ-squared	df	ρ-value
1	223	299	410	255			176.65	3	<2.2e^{-16}
2	257	181	70	225					
	Cross-tabulation 'Pollution' ~ 'Species'								
	Autumn	Summer	Winter	Spring			χ-squared	df	ρ-value
1	324	295	285	283			9.469	3	0.02366
2	156	185	195	197					
	Cross-tabulation 'Pollution' ~ 'Stations'								
	St.1	St.2	St.3	St.4			χ-squared	df	ρ-value
1	331	298	305	253			27.864	3	3.878e^{-06}
2	149	182	175	227					

with a total count of 283 occurrences compared to the other modalities. Therefore, "spring" was chosen as the reference modality for the "seasons" variable. Lastly, for the variable "stations", the modality "1" appeared the least frequent with the modality "St.4", with a total count of 253 occurrences compared to the other modalities. Thus, "St.4" was chosen as the reference modality for the "stations" variable.

By selecting appropriate reference modalities for each variable, we can make meaningful comparisons between the different groups and interpret the results accurately in subsequent analyses. We are now able to achieve our logistic regression. The results of the logistic regression are given in Table 4.

Before proceeding to the environmental risk assessment of aquatic bio-metallic pollution, let's first illustrate that the binary logistic regression is a suitable approach. Several reasons support this claim. First, the variable "pollution" that needs to be explained is qualitative in nature, as it represents the presence or absence of pollution. Moreover, logistic regression is well-suited for modeling binary outcomes. Second, the study has an adequate number of pollution events (cases) compared to the number of explanatory variables. It is crucial to have a sufficient number of events to ensure the reliability and stability of the logistic regression model. Then, the number of explanatory variables significantly associated with pollution is smaller than the number of pollution events. Having a limited number of explanatory variables relative to the number of events ensures a robust model. Finally, the study assumes that the bio-accumulation of heavy metals by aquatic organisms is similar. This assumption aligns with the binary logistic regression framework, where the relationship between the explanatory variables and pollution is modeled in a unified manner. By employing binary logistic regression, the study can effectively assess the environmental risk of aquatic bio-metallic pollution. The model will provide valuable insights into the factors associated with pollution presence, and the results can be used to inform management strategies and mitigation efforts in the studied coastal areas of East-Algeria. In the context of a logistic model, we do not usually present the coefficients of the model but their exponential value, the latter corresponding to the odds ratios. An odds ratio of 1 means no effect. An odds ratio much higher than 1 corresponds to an increase of the studied phenomenon and an odds ratio much lower than 1 corresponds to a decrease of the studied phenomenon [27].

In the context of our study, the odds ratio helps us understand the extent to which each explanatory variable (species, stations, seasons, and heavy metals) influences the presence of pollution along the Eastern-Algerian coast. A value of the odds ratio greater than 1 indicates that the variable is positively associated with pollution, meaning that an increase in the variable is associated with an increase in the likelihood of pollution occurrence. Conversely, a value less than 1 indicates a negative association, implying that an increase in the variable is associated with a decrease in the likelihood of pollution occurrence.

The p-values associated with the odds ratios are used to determine if the odds ratio is statistically significant. A low p-value (usually less than 0.05) indicates that the odds ratio is significantly different from 1, suggesting a strong association between the variable and pollution presence. On the other hand, a high p-value (greater than 0.05) indicates that the odds ratio is not statistically different from 1, suggesting no significant association. By examining the magnitude and direction of the odds ratios, we can identify the most influential factors contributing to pollution and understand their relative importance in the logistic regression model.

Table 5 depicts the values of the odds ratios obtained with the corresponding p-value (Pr.(>|s|)) which allows knowing if an odds ratio differs significantly from 1. Fig. 2 is a graphical representation that depicts the effects of each model variable. The scale range of the contamination

Table 4
Results of binary logistic regression "pollution vs. species + stations + seasons + heavy metals"

	Estimate	Std. error	z-value	Pr.(>(2))
Intercept	–2.55397	0.23404	–10.913	<2e^{-16}***
Species[*Phorcus turbinatus*]	0.92534	0.15908	5.817	5.99e^{-09}***
Species[*Stramonita haemastoma*]	2.54121	0.18760	13.546	<2e^{-16}***
Species[*Ulva lactuca*]	0.39042	0.15678	2.490	0.012763*
Heavy metals[Cd]	4.17262	0.29942	13.936	<2e^{-16}***
Heavy metals[Cr]	0.19180	0.18691	1.026	0.304824
Heavy metals[Cu]	1.07128	0.18459	5.804	6.49e^{-09}***
Heavy metals[Ni]	2.52953	0.20590	12.285	<2e^{-16}***
Heavy metals[Pb]	1.96239	0.19328	10.153	<2e^{-16}***
Seasons[Autumn]	0.55831	0.16610	3.361	0.000776***
Seasons[Summer]	0.16047	0.16363	0.981	0.326731
Seasons[Winter]	0.02663	0.16319	0.163	0.870366
Stations [St.1]	1.04623	0.16750	6.246	4.21e^{-10}***
Stations [St.2]	0.59317	0.16358	3.626	0.000288***
Stations [St.3]	0.68732	0.16421	4.186	2.84e^{-05}***

Signif. Code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '·' 0.1 '' 1.

levels in Fig. 2 varies for different logistic model variables. The unevenly distributed range of contamination levels in the graphical representation can be attributed to the varying impact of each logistic model variable on pollution levels. The ordinate contamination levels represent the probability of pollution presence or absence, ranging from 0 to 1, with 0 indicating no pollution and 1 indicating complete pollution. In the "species effect plot" and "metals effect plot", the contamination levels range from 0.7 to 0.8. This range indicates that the variables "species" and "metals" have a moderate impact on pollution, resulting in contamination levels that are closer to the middle of the probability scale. On the other hand, in the "season effect plot" and "station effect plot", the contamination levels range from 0.8 to 0.9. This range suggests that the variables "season" and "station" have a stronger impact on pollution, leading to contamination levels that are closer to the higher end of the probability scale. The different scale ranges highlight the varying degrees of contribution of each variable to the overall contamination levels. Variables with a stronger impact result in higher or lower contamination probabilities, while variables with a moderate impact lead to contamination probabilities that are more evenly distributed around the middle of the probability scale.

Table 5
Corresponding odds ratios and ρ-values for all categories

	OR	Pr.(>|z|)
Intercept	0.0777722	<2e^{-16}***
Species[*Phorcus turbinatus*]	2.5227209	5.99e^{-09}***
Species[*Stramonita haemastoma*]	12.6949909	<2e^{-16}***
Species[*Ulva lactuca*]	1.4775977	0.012763*
Heavy metals[Cd]	64.8849225	<2e^{-16}***
Heavy metals[Cr]	1.2114279	0.304824
Heavy metals[Cu]	2.9191163	6.49e^{-09}***
Heavy metals[Ni]	12.5475470	<2e^{-16}***
Heavy metals[Pb]	7.1162908	<2e^{-16}***
Seasons[Autumn]	1.7477204	0.000776***
Seasons[Summer]	1.1740630	0.326731
Seasons[Winter]	1.0269897	0.870366
Stations [St.1]	2.8468871	4.21e^{-10}***
Stations [St.2]	1.8097093	0.000288***
Stations [St.3]	1.9883794	2.84e^{-05}***

Signif. Code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1.

3.1. Identification of variables having an effect

In this subsection, the focus is on assessing the extent to which different variables affect the logistic regression model. The *p*-values associated with the odds ratios help determine if each variable has a significant effect on the outcome (pollution) when compared to the reference category. However, merely looking at individual *p*-values does not give a comprehensive view of the overall effect of the variables on the model. To test the global effect of all the variables in the model, a technique involving the "drop1" function is used. This technique involves removing each variable from the model one at a time and performing an analysis of variance to see if the variance of the model changes significantly. The results obtained from the "drop1" function are presented in Table 6. In this case, all the variables significantly modify the model, indicating that each variable has an effect on the prediction of pollution. This means that the combination of season, stations, metals, and species all contribute significantly to the ability of the logistic regression model to predict pollution levels in the coastal areas of East-Algeria.

3.2. Results interpretation

Tables 4–6 and Fig. 2 provide a series of results that confirm the impact of the 4 explanatory variables on the presence of pollution:

- There is an additional risk of significant pollution associated with all the modalities of the 2 variables "species" and "stations" with respect to each of their reference modalities. For the variable "seasons", we notice that only the "autumn" modality is significantly associated with an additional risk of pollution. For the variable

Table 6
Results of analysis of variance after deletion, in turn, of each variable

	df	Deviance	AIC	LRT	Pr.(>Chi)
<none>		1,799.5	1,829.5		
Species	3	2,051.1	2,075.1	251.61	<2.2e^{-16}***
Heavy metals	5	2,320.1	2,340.1	520.54	<2.2e^{-16}***
Seasons	3	1,814.0	1,838.0	14.52	0.002278**
Stations	3	1,841.4	1,865.4	41.86	4.298e^{-09}***

Signif. Code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1.

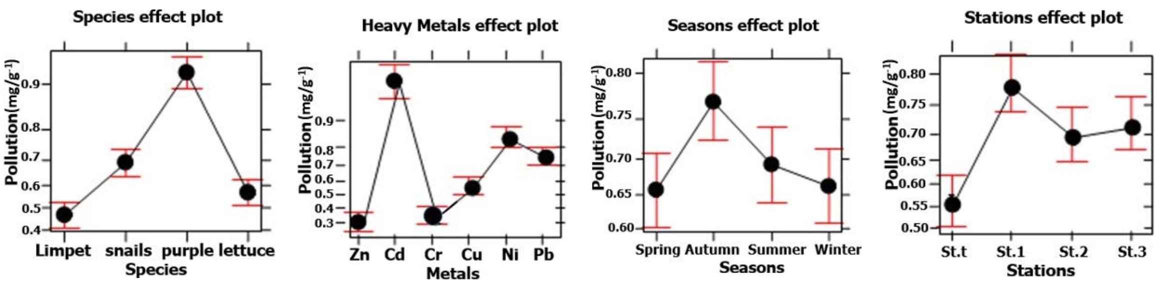


Fig. 2. Graphical representation of the effect of each logistic model variable.

"heavy metals" all the modalities are significantly associated with an additional risk except the Cr modality.

- From the calculation of the odds ratios above we conclude that all things being equal:
- In stations 1, 2, and 3 (St.1, St.2, St.3), the level of pollution risk increases by approximately 3, 2, and 2 times, respectively, compared to the test station (St.4). Moreover, Sampling station one (St.1) is El Djerda island a boathouse (artisanal fishing ports), and anthropogenic wastes are released by local fisherman communities. Station two (St.2) is located about 70 km upstream from the Golfe of Skikda. It is a place where chemical wastes are released by petroleum. Station three (St.3) is located at about 200 km away from the Golfe of Annaba. It is a place where anthropogenic wastes are found. Station four (St.4) is a control station located at about 80 km away from the Chetaibi Bay [18].
- The species most associated with a presence of pollution is "*S. haemastoma*". Its odd ratio is worth 12.695, it indicates then the presence of a risk of pollution almost 13 times higher than the species "*P. caerulea*". Likewise, "*P. turbinatus*" and "*U. lactuca*" were associated to a presence of pollution risk almost 2.5 and 1.5 times higher than the modality "*P. caerulea*". Note that the same results were observed in a previous study on the heavy metals content in veined rapa whelks (*Rapana venosa*) from the Varna Bay. Indeed, authors Stancheva et al. [28] reported Pb, Cd and Hg concentrations of 0.12, 0.008 and 0.08 mg·kg^{-1}, respectively. On the other hand, the contents of Cd and Pb in *R. venosa* from the Black Sea in the research of Jitar et al. [2] were 1.10–1.64 μg·g^{-1} and 0.27–1.29 μg·g^{-1}, respectively.

Moreover, authors Das et al. [29] detected Pb and Cd concentrations of 0.14, 4.63 and 0.050 mg·kg^{-1} in Black Sea veined rapa whelks (*R. venosa*). We can also notice that our results are comparable to by the study of Zhelyazkov et al. [30] that provided evidence for higher Cd (0.02–41.13 μg·g^{-1}) than Pb (0.5 μg·g^{-1}) contents of Black Sea *R. venosa*. The Pb and Cd concentrations in veined rapa whelks (*R. venosa*) caught in the Black Sea ranged from 0.1 to 0.7 mg·kg^{-1}, 0.1 to 1.6 mg·kg^{-1} and 0.4 and 0.7 mg·kg^{-1} in the research of Mülayim and Balkıs [31]. Our results were in agreement with the study of Bat and Öztekin [32] specifying Cd content of 4.4 mg·kg^{-1} and Pb contents 0.05 mg·kg^{-1}.

- The Cd modality of the variable "heavy metals" multiplies by almost 65 the risk of the presence of pollution compared to the modality Zn. Cd accumulates mainly in the human liver and kidneys and is outlined with an exceptionally long half-life. It is nephrotoxic and induces dysfunction of renal tubules characterized with enhanced elimination of low-molecular proteins [24]. Commission Regulation [24] sets maximum content of Cd in bivalve molluscs of 1 mg·kg^{-1} but allowances for rapa whelks are not specified. Then come, by order, the modalities Cu, Pb and Ni showing a risk of pollution almost 3, 7, and 12 times higher than the modality Zn, respectively. This means Cu levels accumulation in *Ulva* St.2 (9.32 ± 0.67 μg·g^{-1} DW) bears a close resemblance with values in *Ulva* from the Gulf of Suez, Red Sea [1,33], and in *Ulva fenestrata* from the Great Bay, Sea of Japan. Also, our results are close to those in *Ulva rigida* from the Venice Lagoon [34] and are higher than values in *Ulva* of Tyrrhenian coastal areas [22] and Turkish coast [35]. Results of the present investigation are in accordance with those reported by the study of Chernova and Kozhenkova [36] in *Ulva fenestrata* from Peter the Great Bay, Sea of Japan, by the study of Al-Shwafi and Rushdi [37] in *E. compressa* from the Gulf of Aden, Yemen. Moreover, Ni values in this study were higher than those from previous studies of algae from the Turkish coast [35] and from the Gulf of Kutch, in the western part of India [33], but lower than most of the algae from different biotopes of the Aegean Sea [38]. Generally, a concentration of 410 mg·g^{-1} (DW) has been considered as a borderline between contaminated and uncontaminated species [39].

Finally, for the variable "seasons", "autumn" is the only modality that influences the risk of the presence of pollution multiplying it by almost twice compared to the other 3 seasons. Optimal growth of green algae *Ulvaceae* in Mediterranean coastal areas has been observed at water temperatures between 12°C and 23°C, while temperatures higher than 24°C are responsible for a halt in growth during summer [40]. In autumn, despite favorable temperatures, the growth of this opportunistic ephemeral seaweed might have been moderate due to low nutrient availability [41]. Hence, that most element contents in this seaweed displayed no distinct seasonal trend from summer to winter-early spring could be at least partly explained by a comparatively low variation in growth rate during this particular period of the year. Environmental factors, metabolic factors, or interactions between both kinds of factors may have contributed to the observed tissue element seasonality. That, as hypothesized, elevated tissue concentrations of some elements (e.g., Ba, Cd, Cr, Mo, Se) in spring or autumn were concurrent with relatively lower salinity values and, also, with markedly elevated levels of these elements in seawater and/or sediment, suggests that, seasonal variation in fluvial and terrestrial inputs may have induced marked changes in element load in the seawater/sediment which potentially influenced to some extent the seasonality of seaweed element contents [42]. That, as also hypothesized, tissue element concentrations generally showed a clear seasonal pattern, mainly characterized by a decrease during spring and/or summer with increasing water temperature and solar irradiance suggests that tissue element seasonality is markedly associated with the seasonal growth pattern of the macroalgae. Higher growth rates during spring and or summer induced from higher temperature and light conditions may have diluted the accumulated elements and thus, reduced their concentrations.

3.3. Goodness of fit

One way to test the quality of a model is to compute a confusion matrix, that is, the cross-tabulation of the observed values and the predicted values by applying the model to the original data. Therefore, we applied our logistic model to the table data and calculated for each

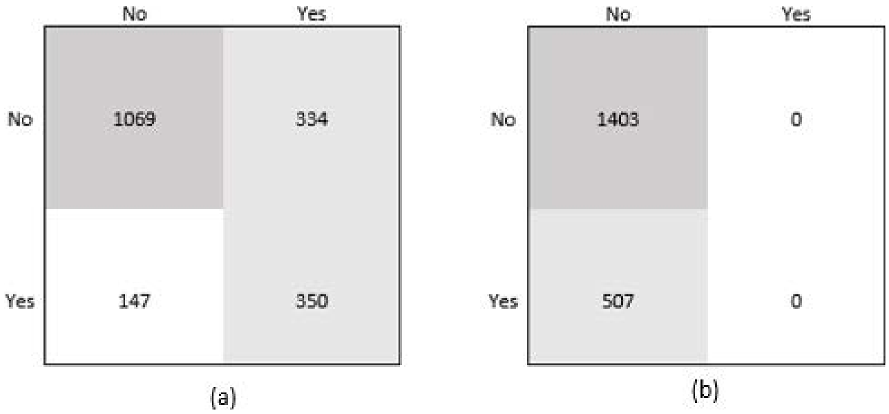


Fig. 3. Confusion matrix of (a) our model and (b) the default model.

individual the probability that he lived the phenomenon under study. Since the studied variable is of binary type, we collect the predicted probabilities in two groups according to whether they are greater or less than half. The confusion matrix is then given by Fig. 3a.

We therefore have 481 (334 + 147) incorrect predictions out of 1910, a misrating of 25.18% denoted E_M. To judge whether the model is good with such an error rate, we will compare it to the error rate of the default model that does not use information from the explanatory variables. The confusion matrix of the default classifier is given by Fig. 3b.

The error rate associated with the default model is 507/1,910 = 26.5% denoted E_D. We can thus derive a measure of performance, denoted R, given by:

R = 1 EM ED (1)

The measure R is interpreted as follows: if R = 1 the model is perfect with a zero-error rate, if R is negative the studied model is worse than the default model and if R is positive the studied model is better than the default model. In our case R = 0.049, logistic regression is better than the default model which means that the fitted statistical model is very suitable for our analysis and highlights the critical role of heavy metals. Findings in green alga "*Ulva rigida*" at Bulgaria (e.g., [43,44]), the present data suggest that the relationships between trace element concentrations in seaweed tissues are a function of environmental variables affecting seaweed growth, and this interferes with the use of macroalgae as biomonitors of trace element contamination.

4. Conclusion

This paper has presented a new study to shed light on the presence of pollution along with four stations in the Eastern-Algerian coast. A statistical model was developed using a logistic regression where the explanatory variables were "species", "stations", "seasons", and "heavy metals". Binary logistic regression revealed that the factors associated with high concentrations depend on each explanatory variable included in the study proving that there is an impact of the four explanatory variables on the presence of pollution. Nevertheless, for the variable "seasons" only the modality "autumn" was significantly associated with an increased pollution risk. On the other hand, the Cr modality of the variable "heavy metals" was the only modality that was not significantly associated with additional risk. Therefore, the logistic model was found to be useful for explaining the pattern of pollution observed at the coasts of Algeria. Overall, the study contributes to our understanding of pollution patterns along the Algerian coast, providing valuable insights for environmental management strategies. The developed logistic regression model offers a promising approach for pollution assessment and holds the potential to be applied in diverse coastal regions for better environmental protection and conservation.

Acknowledgments

I would like to acknowledge my indebtedness and render my warmest thanks to my colleague, Dr. Doukha Zouina (Ph.D.-USTHB, University of Algeria), who made this work possible. Her friendly guidance and expert advice have been invaluable throughout of the writing this paper.

Conflicts of interest

The authors declare no competing interests.

References

[1] Kh.M. El-Moselhy, A.I. Othman, H. Abd El-Azem, M.E.A. El-Metwally, Bioaccumulation of heavy metals in some tissues of fish in the Red Sea, Egypt, Egypt. J. Basic Appl. Sci., 1 (2014) 97–105.

[2] O. Jitar, C. Teodosiu, A. Oros, G. Plavan, M. Nicoara, Bioaccumulation of heavy metals in marine organisms from the Romanian sector of the Black Sea, New Biotechnol., 32 (2015) 369–378.

[3] L. Spada, C. Annicchiarico, N. Cardellicchio, S. Giandomenico, A. Di Leo, Heavy metals monitoring in mussels *Mytilus galloprovincialis* from the Apulian coasts (Southern Italy), Mediterr. Mar. Sci., 14 (2013) 99–108.

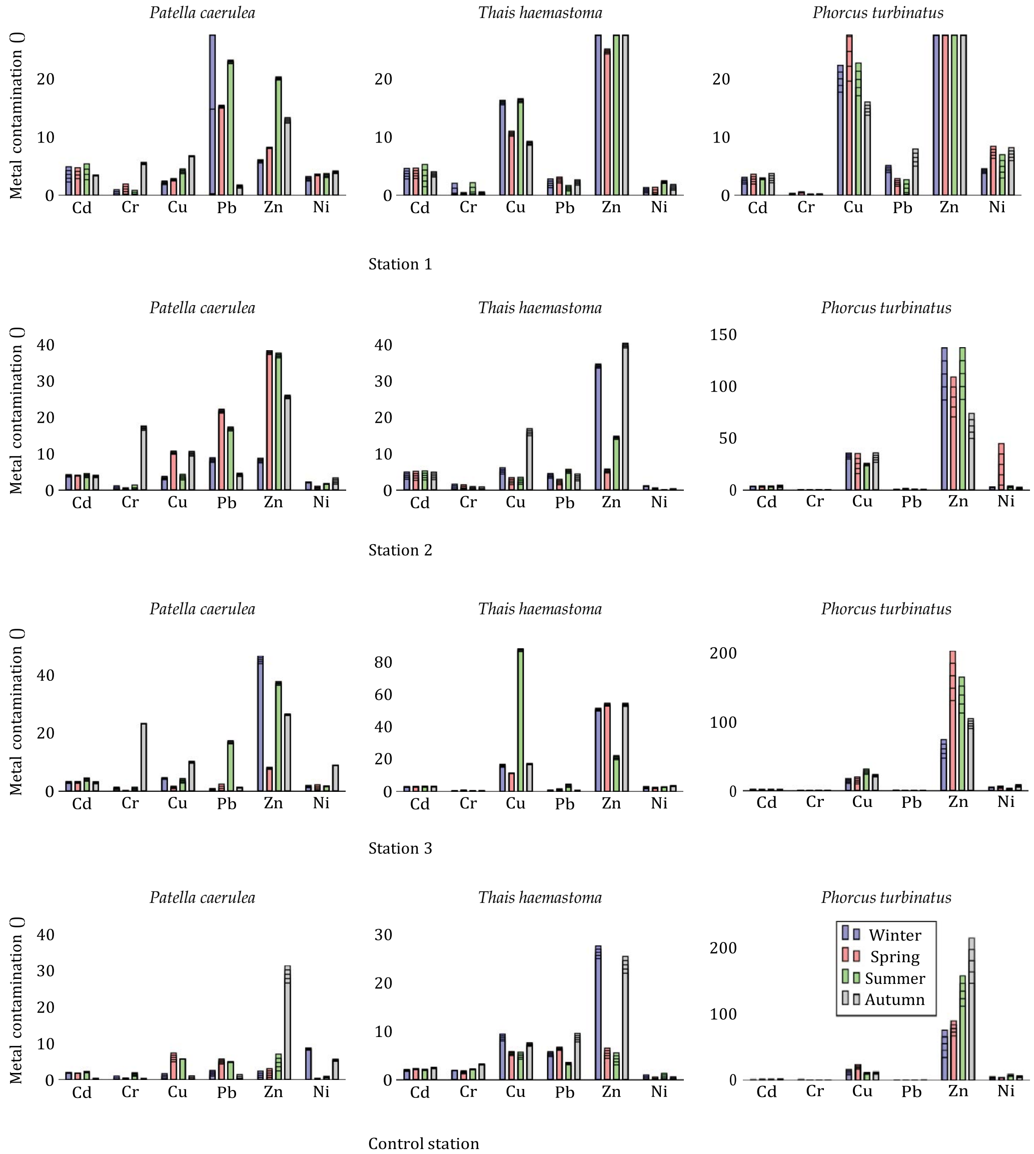
[4] M. Jović, S. Stanković, Human exposure to trace metals and possible public health risks via consumption of mussels *Mytilus galloprovincialis* from the adriatic coastal area, Food Chem. Toxicol., 70 (2014) 241–251.

[5] S. Boucetta, H. Beldi, B. Draredja, Effects of metal pollution on the activities of *Acetylcholinesterase* and glutathione-S-transferase in *Phorcus* (*Osilinus*) *turbinatus* (Gastropoda, Trochidae) of the coast East-Algerian, Adv. Environ. Biol., 10 (2016) 46–61.

[6] O. Rouane-Hacene, Z. Boutiba, M. Benaissa, B. Belhaouari, P. Francour, M.E. Guibbolini-Sabatier, C. Risso-De Faverney,

Seasonal assessment of biological indices, bioaccumulation, and bioavailability of heavy metals in sea urchins *Paracentrotus lividus* from Algerian west coast, applied to environmental monitoring, Environ. Sci. Pollut. Res., 25 (2018) 11238–11251.
[7] L. Volterra, M.E. Conti, Algae as biomarkers, bioaccumulators and toxin producers, Int. J. Environ. Pollut., 13 (2000) 92.
[8] M.E. Conti, M. Iacobucci, G. Cecchetti, The biomonitoring approach as a tool of trace metal assessment in an uncontaminated marine ecosystem: the island of Ustica (Sicily, Italy), WIT Trans. Ecol. Environ., 75 (2004), doi: 10.2495/GEO040311.
[9] M.D. Wallenstein, S.K. Mcmahon, J.P. Schimel, Seasonal variation in enzyme activities and temperature sensitivities in Arctic tundra soils, Global Change Biol., 15 (2009) 1631–1639.
[10] M. Rajfur, A. Kłos, M. Wacławek, Sorption properties of algae *Spirogyra* sp. and their use for determination of heavy metal ions concentrations in surface water, Bioelectrochemistry, 80 (2010) 81–86.
[11] K. Anastasakis, A.B. Ross, Hydrothermal liquefaction of the brown macro-Alga *Laminaria saccharina*: effect of reaction conditions on product distribution and composition, Bioresour. Technol., 102 (2011) 4876–4883.
[12] C.-Y.P. Peng, K.L. Lee, G.M. Ingersoll, An introduction to logistic regression analysis and reporting, J. Educ. Res., 96 (2002) 3–14.
[13] R.P. Burns, R. Burns, Business Research Methods and Statistics Using SPSS, SAGE Open, 2008.
[14] P. Wongsasuluk, S. Chotpantarat, W. Siriwong, M. Robson, Using hair and fingernails in binary logistic regression for bio-monitoring of heavy metals/metalloid in groundwater in intensively agricultural areas, Thailand, Environ. Res., 162 (2018) 106–118.
[15] M. Guemouda, Z. Meghlaoui, T. Daas, O. Daas-Maamcha, P. Scaps, Monitoring pollution in East Algerian coasts using biochemical markers in the polychaete annelid *Perinereis cultrifera*, Ann. Biol. Res., 5 (2014) 31–40.
[16] K. Sifi, S. Chouahda, N. Soltani, Biosurveillance de l'environnement par la mesure de biomarqueurs chez Donax Trunculus dans le golfe d'annaba (Algérie), Mesogee, 63 (2007) 11–18.
[17] S. Hidouci, A.B. Djebar, R. Amara, Et H. Sahraoui, Bacterial quality of coastal waters of Annaba (East Algeria), Eur. J. Sci. Res., 120 (2014) 488–493.
[18] S. Boucetta, W. Benchalel, S. Ferroudj, Z. Bouslama, H. Elmsellem, Trace metal biomonitoring of algae (*Ulva lactuca*), and Mollusks (*Patella caerulea*; *Stramonita haemastoma*; *Phorcus turbinatus*) along the Eastern-Algerian coast, Morrocan J. Chem., 7 (2019) 444–459.
[19] M.E. Conti, B. Bocca, M. Iacobucci, M.G. Finoia, M. Mecozzi, A. Pino, A. Alimonti, Baseline trace metals in seagrass, algae, and mollusks in a southern *Tyrrhenian ecosystem* (Linosa Island, Sicily), Arch. Environ. Contam. Toxicol., 58 (2010) 79–95.
[20] J. Pempkowiak, A. Sikora, E. Biernacka, Speciation of heavy metals in marine sediments vs. their bioaccumulation by mussels, Chemosphere, 39 (1999) 313–321.
[21] B. Bocca, M.E. Conti, A. Pino, D. Mattei, G. Forte, A. Alimonti, Simple, fast, and low-contamination microwave-assisted digestion procedures for the determination of chemical elements in biological and environmental matrices by sector field ICP-MS, Int. J. Environ. Anal. Chem., 87 (2007) 1111–1123.
[22] M.E. Conti, G. Cecchetti, A biomonitoring study: trace metals in algae and molluscs from Tyrrhenian coastal areas, Environ. Res., 93 (2003) 99–112.
[23] F.N.M. El Sanharawi, M. El Sanharawi, F. Naudet, Understanding logistic regression, J. Fr. Ophtalmol., 36 (2013) 710–715.
[24] FAO and WHO Expert Committee on Food Additives and World Health Organization, Toxicological Evaluation of Certain Food Additives and Contaminants, World Health Organization, 1993.
[25] L.F. Lindner, Regulating food safety: the power of alignment and drive towards convergence, Innovation: Eur. J. Soc. Sci., 21 (2008) 133–143.
[26] M. Maluszynski, K. Nichterlein, L. Van Zanten, B.S. Ahloowalia, Officially Released Mutant Varieties - The FAO/IAEA DATABASE, Joint FAO/IAEA Division of Nuclear Techniques in Food and Agriculture and FAO/IAEA Agriculture and Biotechnology Laboratory, Seibersdorf International Atomic Energy Agency, Vienna, 2000, pp. 2–84.
[27] D.W. Hosmer Jr., S. Lemeshow, R.X. Sturdivant, Applied Logistic Regression, 2nd ed., Wiley, 2013.
[28] M. Stancheva, V. Ivanova, K. Peycheva, Determination of heavy metals in black sea *Mytilusgallo provincialis* and *Rapana venosa*, Scr. Sci. Med., 44 (2012) 27, doi: 10.14748/ssm.v44i2.353.
[29] Y.K. Das, A. Aksoy, R. Baskaya, H.A. Duyar, D. Guvenc, V. Boz, Heavy metal levels of some marine organisms collected in Samsun and Sinop coasts of Black Sea, in Turkey, J. Anim. Vet. Adv., 8 (2009) 496–499.
[30] G. Zhelyazkov, T. Yankovska-Stefanova, E. Mineva, D. Stratev, I. Vashin, L. Dospatliev, E. Valkova, T. Popova, Risk assessment of some heavy metals in mussels (*Mytilus galloprovincialis*) and veined rapa whelks (*Rapana venosa*) for human health, Mar. Pollut. Bull., 128 (2018) 197–201.
[31] A. Mülayim, H. Balkıs, Toxic metal (Pb, Cd, Cr, and Hg) levels in *Rapana venosa* (Valenciennes, 1846), *Eriphia verrucosa* (Forskal, 1775), and sediment samples from the Black Sea littoral (Thrace, Turkey), Mar. Pollut. Bull., 95 (2015) 215–222.
[32] L. Bat, H.C. Öztekin, Heavy metals in *Mytilus galloprovincialis*, *Rapana venosa* and *Eriphia verrucosa* from the Black Sea coasts of turkey as bioindicators of pollution, Walailak J. Sci. Technol., 13 (2016) 715–728.
[33] S. Chakraborty, T. Bhattacharya, G. Singh, J.P. Maity, Benthic macroalgae as biological indicators of heavy metal pollution in the marine environments: a biomonitoring approach for pollution assessment, Ecotoxicol. Environ. Saf., 100 (2014) 61–68.
[34] M. Caliceti, E. Argese, A. Sfriso, B. Pavoni, Heavy metal contamination in the seaweeds of the Venice lagoon, Chemosphere, 47 (2002) 443–454.
[35] S. Topcuoğlu, Ç. Kırbaşoğlu, N. Güngör, Heavy metals in organisms and sediments from Turkish Coast of the Black Sea, 1997–1998, Environ. Int., 27 (2002) 521–526.
[36] E.N. Chernova, S.I. Kozhenkova, Determination of threshold concentrations of metals in indicator algae of coastal waters in the northwest Sea of Japan, Oceanology, 56 (2016) 363–371.
[37] N.A. Al-Shwafi, A.I. Rushdi, Heavy metal concentrations in marine green, brown, and red seaweeds from coastal waters of Yemen, the Gulf of Aden, Environ. Geol., 55 (2007) 653–660.
[38] T. Sawidis, M.T. Brown, G. Zachariadis, I. Sratis, Trace metal concentrations in marine macroalgae from different biotopes in the Aegean Sea, Environ. Int., 27 (2001) 43–47.
[39] G. Lozano, A. Hardisson, A.J. Gutiérrez, M.A. Lafuente, Lead and cadmium levels in coastal benthic algae (seaweeds) of Tenerife, Canary Islands, Environ. Int., 28 (2003) 627–631.
[40] V. Rigollet, A. Sfriso, A. Marcomini, M.L. De Casabianca, Seasonal evolution of heavy metal concentrations in the surface sediments of two Mediterranean Zostera marina L. beds at Thau lagoon (France) and Venice lagoon (Italy), Bioresour. Technol., 95 (2004) 159–167.
[41] P. Malea, A. Chatziapostolou, T. Kevrekidis, Trace element seasonality in marine macroalgae of different functional-form groups, Mar. Environ. Res., 103 (2015) 18–26.
[42] S. Haritonidis, P. Malea, Seasonal and local variation of Cr, Ni and Co concentrations in *Ulva rigida* C. Agardh and *Enteromorpha linza* (Linnaeus) from Thermaikos Gulf, Greece, Environ. Pollut., 89 (1995) 319–327.
[43] T. Boubonari, P. Malea, T. Kevrekidis, The green seaweed *Ulva rigida* as a bioindicator of metals (Zn, Cu, Pb and Cd) in a low-salinity coastal environment, Bot. Mar., 51 (2008), doi: 10.1515/BOT.2008.059.
[44] D.J.H. Phillips, Use of Macroalgae and Invertebrates as Monitors of Metal Levels in Estuaries and Coastal Waters, CRC Press, 1990.

Appendix A: Spatio-temporal variations of metallic contents 'Cr, Cd, Cu, Pb, Zn and Ni' in rocky aquatic organisms all along the East-Algerian coasts from December 2011 to December 2019.



Appendix B: Spatio-temporal variations of metallic contents 'Cr, Cd, Cu, Pb, Zn and Ni' in algae "Ulva lactuca" all along the East-Algerian coasts from December 2011 to December 2019.