Data in brief 24 (2019) 103860

Contents lists available at ScienceDirect

Data in brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Data on long-term monitoring programs to assess environmental pressures on coastal area

Matilda Mali [a, *], Nicola Ungaro [b], Angelo Cardellicchio [c], Maria Michela Dell'Anna [a], Giuseppe Romanazzi [a], Piero Mastrorilli [a], Leonardo Damiani [a]

[a] DICATECh, Politecnico di Bari, Via Orabona, 4 I-70125 Bari, Italy
[b] ARPA Puglia, Regional Agency for Environment Prevention and Protection, Corso Trieste, 27, 70126 Bari BA, Italy
[c] DEE, Politecnico di Bari, Via Orabona, 4 I-70125 Bari, Italy

ARTICLE INFO

Article history:
Received 3 February 2019
Received in revised form 11 March 2019
Accepted 11 March 2019
Available online 20 March 2019

ABSTRACT

The concentration of six metals/metalloids, five congeners of high molecular weight Polycyclic Aromatic Hydrocarbons (PAHs), and sum of five congeners of Polychlorinated Biphenyls (PCBs) determined within marine-coastal sediments of the Apulia Coast during a 5-year long-term monitoring program, are reported through tables and radial graphs. The data are referred to the pollutant concentration determined within 70 sites alongside two marine transects (500 m from coastline and 1750 m of coastline) representing different morphologic features of the coast and different pollution stressors loading [1]. Concentration variability during the five monitored years and data generated by the non-parametric correlation analyses among sediment physical-chemical main parameters and metal concentrations are also included.

© 2019 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

DOI of original article: https://doi.org/10.1016/j.ecolind.2019.01.085.
* Corresponding author.
E-mail address: matilda.mali@poliba.it (M. Mali).

https://doi.org/10.1016/j.dib.2019.103860
2352-3409/© 2019 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Specifications table

Subject area	*Environmental Science*
More specific subject area	*Marine-coastal water and sediment hazard assessment.*
Type of data	*Tables, radial graphs, equations, figures, text file*
How data was acquired	*Chemical Determination are performed by standardized methods adopted by ARPA Puglia. Descriptive statistics of pollution concentration, correlation and regression analyses of physicochemical properties are performed using a FOSS software (Free and Open Source Software), available at https://github.com/anhelus/pylab as a Jupyter Notebook.*
Data format	*Tables and Radial Graphs are used for pollutant concentration, while graphs, equations and figures are used for correlation among physical-chemical sediment properties.*
Experimental factors	*The concentration of metals and metalloids (As, Cr, Cd, Ni, Pb) were obtained by inductively coupled plasma-mass spectrometry (ICP-MS), according to EPA 6020A-2007 method. The Hg mercury content was quantified by cold vapor atomic absorption spectrometry (CetacQuickTrace M-6100 Mercury Analyzer). The concentration of five congeners of high weight Polycyclic Aromatic Hydrocarbons (PAHs), (benzo(b)fluoranthene (BbF), benzo(k)fluoranthene (BkF), benzo(a)pyrene (BaP), benzo(g,h,i) perylene (BghiP), indeno(1,2,3,c,d)pyrene (Ind)) are determined the EPA 610 Method for Polynuclear Aromatic Hydrocarbon Mix. The sum of five PCB congeners (28, 47, 99, 100, 153, 154) was determined through extraction in acetone/petroleum solvent followed by analysis by Gas Chromatograph equipped with an Electron Capture Detector. Data on quality assurance and quality control were included in Mali* et al. *2017 [2,3] The Reference standard materials utilized is NIST-2977.*
Experimental features	*The data provide a comprehensive set that describe the pollution trend alongside Adriatic Ionian marine coastal area [1].*
Data source location	*The dataset of pollution concentration in the investigated sites were collected by ARPA PUGLIA and are available in the Scientific Direction in Corso Trieste, 27, 70126 Bari. The coordinates of sampling sites investigated are reported in the Table n. 1. Statistical elaborations are available at DICATECh Department of Politecnico di Bari, via Orabona, 4 I-70125 Bari, Italy.*
Data accessibility	*Data are within this article.*
Related research article	*The data are submitted as a companion paper to the research article of Mali* et al. *2019 [1] entitled "Long-term monitoring programs to assess environmental pressures on coastal area: weighted indexes and statistical elaboration, as handy tools for decision-makers", currently accepted for publication in Ecological Indicator Journal*

Value of the data

The data included in the present paper can be used for:

- Further analyses in risk assessment for marine-coastal areas.
- Further studies focused on the geospatial contamination distribution at regional scale (Geo-mapping Pollution) as well as for studies in contamination transport modelling within marine-coastal areas under different hydrodynamic pressures [4].
- Further studies on interaction of individual pollutant concentration with benthic community in toxic responses

1. Data

Quantitative data on pollutant concentration within marine sediments of Adriatic Coastal Area are presented in re-usable format. Table 1 shows the latitude and longitude of the sampling sites. Tables 2–5 show the concentration of pollutants determined within each monitoring period (respectively within 2010–2011, 2012–2013; 2013–2014, 2014–2015) in 35 marine transects. The pollutants considered are: six trace elements (As, Cd, Cr, Ni, Hg, Pb), five high molecular PAHs congeners, benzo(b) fluoranthene(BbF), benzo(k)fluoranthene (BkF), benzo(a)pyrene (BaP), benzo(g,h,i)perylene (BghiP), indeno(1,2,3,c,d)pyrene) (Ind), sum of five congeners of Polychlorinated biphenyls (PCBs) congeners (PCB28, PCB47, PCB99, PCB100, PCB153, PCB154), TOC and sediment textural features.

Correlation analyses (LARS regression graphs) among TOC and Clay concentration with the six metal concentration are reported in Fig. 1a and b, Fig. 2a and b, Fig. 3, Fig. 4a and b and Fig. 5a and b.

Table 1
Coordinates (latitude and longitude) of the investigated sites.

Site	Site-transect	Code	Latitude	Longitude
Foce Varano-Peschici	Peschici 200	PE01	41°57′10,400″ N	16°1′3200″ E
	Peschici 1750	PE02	41°57′48,909″ N	16°1′8045″ E
Peschici-Vieste	Vieste 200	VI01	41°53′13,900″ N	16°11′11,000″ E
	Vieste 1750	VI02	41°53′46,427″ N	16°6′51,069″ E
Vieste-Mattinata	Mattinatela 200	MI01	41°43′42,187″ N	16°11′51,069″ E
	Mattinatela 1750	MI02	41°43′ 3131″ N	16°7′ 29,603″ E
Mattinata-Manfredonia	Mattinata 200	MT01	41°41′ 40,600″ N	16°4′ 10,300″ E
	Mattinata 1750	MT02	41°41′ 34,652″ N	16°5′ 1793″ E
	Manfredonia SIN 500	MN01	41°38′ 38,000″ N	15°57′ 32,300″ E
	Manfredonia SIN 1750	MN02	41°38′ 2758″ N	15°57′ 57,231″ E
Manfredonia-Torre Cervaro	F. Candelaro 500	FC01	41°35′ 5100″ N	15°53′ 59,500″ E
	F Candelaro 1750	FC02	41°35′ 1733″ N	15°54′ 49,392″ E
Torre Cervaro-Foce Carapelle	F. Carapelle 500	CR01	41°29′ 45,300″ N	15°55′ 53,600″ E
	F. Carapella 1750	CR02	41°30′ 1684″ N	15°56′ 37,674″ E
Foce Carapelle-Foce Aloisa	F. Aloisa 500	AL01	41°26′ 11,571″ N	16°0′ 41,094″ E
	F. Aloisa 1750	AL02	41°26′ 44,253″ N	16°1′ 7913″ E
Foce Aloisa -Margherita di Savoia	F. Camosina 500	CM01	41°24′ 54,300″ N	16°4′ 15,200″ E
	F. Camosina 1750	CM02	41°25′ 33,780″ N	16°4′ 37,080″ E
Margherita di Savoia-Barletta	F. Ofanto 500	FO01	41°21′ 56,400″ N	16°12′ 17,200″ E
	F. Ofanto 1750	FO02	41°22′ 27,442″ N	16°12′ 45,726″ E
Barletta-Bisceglie	Bisceglie 500	BI01	41°14′ 48,300″ N	16°30′ 56,300″ E
	Bisceglie 1750	BI02	41°15′ 23,603″ N	16°31′ 39,090″ E
Bisceglie-Molfetta	Molfetta 500	ML01	41°12′ 10,800″ N	16°36′ 59,900″ E
	Moletta 1750	ML02	41°12′ 45,360″ N	16°37′ 27,874″ E
Molfetta-Bari	Bari Balice 500	BB01	41°8′ 41,600″ N	16°48′ 43,100″ E
	Bari Balice 1750	BB02	41°9′ 22,489″ N	16°49′ 8461″ E
Bari-San Vito Polignano	Bari Trullo 500	BA01	41°6′ 43,500″ N	16°56′ 9700″ E
	Bari Trullo 1750	BA02	41°7′ 20,404″ N	16°56′ 30,450″ E
San Vito Polignano-Monopoli	Mola 500	MA01	41°3′ 21,482″ N	17°7′ 0,198″ E
	Mola 1750	MA02	41°3′ 49,658″ N	17°7′ 25,566″ E
Area Protetta Torre Guaceto	T. Guaceto 500	TG01	40°42′ 29,400″ N	17°48′ 40,900″ E
	T. Guaceto 1750	TG02	40°43′ 24,701″ N	17°49′ 29,575″ E
Limite Sud AMP Torre Guaceto	Punta Penna 100	PP01	40°41′ 10,983″ N	17°56′ 22,482″ E
	Punta Penne 600	PP02	40°41′ 22,300″ N	17°56′ 27,654″ E
Cerano-Le Cesine	Le Cesine 500	CE01	40°32′ 25,500″ N	18°4′ 53,100″ E
	Le Cesine 1750	CE02	40°22′ 14,922″ N	18°21′ 13,244″ E

2. Experimental design, materials, and methods

Five year monitoring program of the marine sediment quality alongside Apulia coastal area (South Italy) has been performed with a yearly frequency starting from March to April. The monitoring program, conducted by ARPA Puglia, includes different matrixes. In the present paper data referring to quality of sediments are shown.

Preliminary data assessment and evaluation is performed by means of standard visual Exploratory Data Analysis **(EDA)** techniques [5,7]. In the Figure S1 of the supporting information are reported Radar Plots of metal concentrations expressed as normalized (min-max algorithm) data. Correlation analysis through sediments data is performed using both Pearson's, Spearman's and Kendall's ; results are visually shown by means of correlation matrices (Fig. 1a and b, Fig. 2a and b, Fig. 3). Step-wise regression analysis is performed, and LARS-LASSO regularization is employed to strengthen results (Fig. 4a and b, Fig. 5a and b) [8,9] The software used for these analysis is based on a set of widely employed FOSS (Free and Open Source Software) libraries for machine learning written in Python, namely Scikit Learn, Pandas, Numpy and TensorFlow. The code is freely available at https://github.com/anhelus/pylab as a Jupyter Notebook.

Table 2
Granulometry and concentration of TOC, P, Metals and organic pollutants for the Period 31st March 2010 − 1st April 2011.

Code	Coarse (%)	Sand (%)	Pelite (%)	TOC (mg/kg ds)	P_{tot} (mg/kg ds)	As (mg/kg ds)	Cd (mg/kg ds)	Cr_{TOT} (mg/kg ds)	Hg (mg/kg ds)	Ni (mg/kg ds)	Pb (mg/kg ds)	ΣPCB (ppm)	B(a)P (ppm)	B(b)F (ppm)	$B_{ghi}P$ (ppm)	B(k)F (ppm)	I(cd)P (ppm)
PE01	0.0	100.0	0.00	0.45	93	3.4	0.075	3.1	0.001	2.5	1.4	5	5.0	5.0	5.0	5.0	5.0
PE02	0.0	31.8	68.2	0.88	185	3.1	0.100	27.3	0.001	12.4	0.6	5	5.0	5.0	5.0	5.0	5.0
VI01	0.0	100.0	0.00	0.13	220	14.9	0.340	12.6	0.001	11.7	4.3	43.6	5.0	5.0	5.0	5.0	5.0
VI02	0.0	49.2	50.8	0.78	554	11.5	0.430	45.4	0.001	35.0	14.7	107	5.0	5.0	5.0	5.0	5.0
MI01	0.0	100.0	0.00	0.10	161	19.9	0.390	5.3	0.001	2.7	3.4	62.2	5.0	5.0	5.0	5.0	5.0
MI02	0.0	62.6	37.5	0.66	668	15.1	0.390	70.5	0.001	38.9	22.2	138	5.0	5.0	5.0	5.0	5.0
MT01	0.0	47.0	53.0	0.44	94	16.6	0.300	66.5	0.001	32.3	15.9	33.8	5.0	5.0	5.0	5.0	5.0
MT02	0.0	84.6	15.4	0.56	679	12.3	0.350	56.7	0.001	31.6	15.9	68.5	5.0	5.0	5.0	5.0	5.0
MN01	0.0	80.0	20.0	0.16	560	6.6	0.290	18.5	0.001	10.1	4.1	193	5.0	5.0	5.0	5.0	5.0
MN02	0.0	51.4	48.6	1.06	559	12.1	0.390	62.4	0.001	32.7	17.7	150	5.0	5.0	5.0	5.0	5.0
FC01	5.8	93.6	0.5	0.63	1117	19.6	0.050	23.4	0.001	32.3	15.3	47.2	5.0	5.0	5.0	5.0	5.0
FC02	0.0	82.0	18.0	0.74	898	12.7	0.050	28.5	0.001	41.3	25.0	52.3	5.0	5.0	5.0	5.0	5.0
CR01	0.0	4.2	95.8	1.13	2700	2.1	0.050	15.1	0.001	42.7	31.4	14.1	5.0	5.0	5.0	5.0	5.0
CR02	0.0	4.8	95.2	1.29	979	6.7	0.050	19.6	0.030	30.0	17.8	105	5.0	5.0	5.0	5.0	5.0
AL01	0.4	82.0	17.6	0.20	835	7.0	0.100	20.0	0.025	13.0	6.0	27.9	0.5	0.5	1.7	1.0	1.1
AL02	25.2	73.3	1.5	0.10	285	19.0	0.100	11.0	0.025	9.0	4.0	31.1	0.5	0.5	0.5	0.5	0.5
CM01	1.1	94.3	4.6	0.10	1296	6.0	0.080	25.0	0.025	12.0	5.0	27.1	0.5	0.5	0.5	0.5	0.5
CM02	4.6	59.6	35.8	0.50	496	11.0	0.080	28.0	0.025	16.0	9.0	27.8	0.5	0.5	2.4	1.3	1.6
FO01	0.1	84.0	15.9	0.20	1824	7.0	0.180	39.0	0.025	29.0	13.0	28.5	0.5	0.5	1.1	0.5	0.5
FO02	0.1	52.3	47.6	0.40	862	6.0	0.090	36.0	0.025	21.0	12.0	29.9	0.5	0.5	2.4	1.1	1.2
BI01	0.0	68.6	31.4	0.20	538	9.0	0.080	10.0	0.025	12.0	6.0	53.6	19.0	8.0	12.0	11.0	10.0
BI02	0.0	13.8	86.2	0.60	687	12.0	0.090	20.0	0.070	21.0	12.0	77.3	6.0	3.0	6.0	3.0	5.0
ML01	5.8	39.4	54.8	0.80	657	11.0	0.130	26.0	0.050	24.0	16.0	24.0	16.0	65.3	11.0	5.0	11.0
ML02	0.0	13.5	86.5	0.80	740	12.0	0.140	31.0	0.260	31.0	19.0	31.0	19.0	58.9	41.0	18.0	28.0
BB01	4.4	33.2	62.4	1.10	669	13.0	0.130	34.0	0.180	29.0	20.0	29.0	20.0	52.5	17.0	6.0	25.0
BB02	0.0	94.7	5.3	0.40	387	5.0	0.110	10.0	0.025	9.0	7.0	9.0	7.0	43.5	49.0	22.0	31.0
BA01	6.5	91.0	2.5	0.60	351	14.0	0.060	7.0	0.025	31.0	7.0	31.0	7.0	34.5	2.0	0.5	2.0
BA02	0.4	28.9	70.7	0.90	640	12.0	0.120	25.0	0.070	24.0	16.0	24.0	16.0	92.4	13.0	6.0	12.0
MA01	0.5	97.9	1.6	0.20	193	10.0	0.060	7.0	0.060	4.0	3.0	4.0	3.0	29.4	62.0	39.0	41.5
MA02	2.6	40.0	57.4	1.10	662	12.0	0.060	42.0	0.090	22.0	13.0	22.0	13.0	95.0	8.9	9.6	11.0
TG01	10.0	90.0	1.0	0.44	416	8.2	0.410	4.7	0.014	2.2	2.3	2.2	2.3	38.8	n.d[a].	n.d.	n.d.
TG02	10.0	90.0	1.0	0.35	292	18.1	0.350	4.9	0.033	3.1	7.0	3.1	7.0	33.1	n.d.	n.d.	n.d.
PP01	5.4	94.6	0.0	0.31	169	4.3	0.050	6.1	0.005	2.3	3.0	2.3	3.0	84.6	n.d.	n.d.	n.d.
PP02	0.6	99.1	0.0	0.19	148	0.19	148	7.7	0.008	2.2	3.2	2.2	3.2	41.4	n.d.	n.d.	n.d.
CE01	6.3	93.6	0.2	0.31	88	0.31	88	5.0	0.100	6.6	4.4	6.6	4.4	48.0	0.3	0.3	0.2
CE02	24.2	71.2	4.7	0.70	103	0.70	103	4.8	0.120	6.5	5.8	6.5	5.8	54.5	0.3	0.2	0.2

[a] n.d. = no detected.

Table 3
Granulometry and concentration of TOC, P, Metals and organic pollutants for the Period 31st March 2012 – 1st April 2013

Code	Coarse (%)	Sand (%)	Pelite (%)	TOC (mg/kg ds)	P_{tot} (mg/kg ds)	As (mg/kg ds)	Cd (mg/kg ds)	Cr_{TOT} (mg/kg ds)	Hg (mg/kg ds)	Ni (mg/kg ds)	Pb (mg/kg ds)	ΣPCB (ppm)	B(a)P (ppm)	B(b)F (ppm)	$B_{ghi}P$ (ppm)	B(k)F (ppm)	I(cd)P (ppm)
PE01	0.0	98.2	1.8	0.13	133	6.3	0.070	5.4	0.005	4.3	3.4	312	5.0	5.0	5.0	5.0	5.0
PE02	0.0	51.0	49.6	0.62	587	2.8	0.110	14.3	0.005	10.0	5.4	n.d.	5.0	5.0	5.0	5.0	5.0
VI01	0.0	98.8	1.2	0.11	94	6.2	0.140	3.9	0.005	4.1	2.4	n.d.	5.0	5.0	5.0	5.0	5.0
VI02	0.0	100.0	0.00	0.70	434	8.0	0.296	41.7	0.005	28.9	16.7	n.d.	5.0	5.0	5.0	5.0	5.0
MI01	0.0	100.0	0.0	0.10	422	7.2	0.080	25.1	0.005	15.1	6.8	21.7	5.0	5.0	5.0	5.0	5.0
MI02	0.0	53.3	46.7	0.66	627	9.9	0.099	16.8	0.005	16.9	5.9	52.9	5.0	5.0	5.0	5.0	5.0
MT01	0.0	80.2	19.8	0.50	154	2.2	0.020	5.6	0.005	3.7	1.9	32.6	5.0	5.0	5.0	5.0	5.0
MT02	0.0	44.5	55.5	0.40	124	1.6	0.020	5.4	0.005	3.7	1.7	50.6	5.0	5.0	5.0	5.0	5.0
MN01	0.0	100.0	0.0	0.16	208	3.6	0.150	6.8	0.005	0.3	3.1	n.d.	5.0	5.0	5.0	5.0	5.0
MN02	0.0	50.2	22.7	1.00	260	6.5	0.182	24.0	0.005	14.8	15.0	n.d.	5.0	5.0	5.0	5.0	5.0
FC01	0.0	23.0	77.0	1.47	287	5.2	0.180	8.4	0.005	12.1	4.6	n.d.	5.0	5.0	5.0	5.0	5.0
FC02	0.0	29.3	70.7	1.40	308	6.3	0.070	5.4	0.005	4.3	3.4	n.d.	5.0	5.0	5.0	5.0	5.0
CR01	0.0	97.7	2.3	0.20	160	2.3	0.020	5.9	0.005	3.9	2.0	n.d.	5.0	5.0	5.0	5.0	5.0
CR02	0.0	78.6	21.4	1.29	145	3.2	0.020	1.7	0.005	2.5	1.2	n.d.	5.0	5.0	5.0	5.0	5.0
AL01	0.1	89.0	11.0	0.10	1718	7.5	0.380	19.0	0.006	15.3	4.1	n.d.	5.0	5.0	5.0	5.0	5.0
AL02	6.5	92.2	1.3	0.73	208	69.9	0.670	13.6	0.009	10.6	8.6	11.8	5.0	5.0	5.0	5.0	5.0
CM01	0.4	89.6	10.0	0.11	594	14.2	0.420	19.2	0.010	18.7	5.3	n.d.	5.0	5.0	5.0	5.0	5.0
CM02	0.9	76.3	22.8	1.00	226	24.4	0.400	23.7	0.020	20.2	6.8	n.d.	11.1	10.7	12.8	5.0	6.8
FO01	0.1	91.3	8.7	0.36	396	6.6	0.380	12.0	0.040	18.0	6.3	n.d.	5.0	5.0	10.5	5.0	10.3
FO02	0.4	72.5	27.1	0.39	576	8.4	0.390	22.6	0.010	21.5	8.3	12.8	5.0	5.0	10.6	5.0	10.4
BI01	0.1	52.1	47.9	0.20	466	17.1	0.410	18.8	0.020	21.7	8.6	41.3	5.0	5.0	5.0	5.0	5.0
BI02	0.1	49.7	50.3	0.51	423	20.1	0.420	30.5	0.030	31.9	13.6	69.8	5.0	5.0	11.0	5.0	10.8
ML01	5.8	38.0	56.2	0.32	50	52.8	0.370	7.7	0.030	5.5	7.3	50.9	10.8	10.4	12.3	5.0	12.1
ML02	0.0	17.0	83.0	0.83	437	22.1	0.470	46.4	0.050	46.5	19.0	320	11.7	11.3	13.5	11.4	13.2
BB01	2.0	95.3	2.7	0.30	110	17.2	0.370	4.9	0.023	3.1	4.0	353	26.3	11.0	11.5	11.0	20.2
BB02	16.0	81.7	2.3	0.40	50	43.9	0.350	4.2	0.036	3.8	6.6	673	5.0	5.0	5.0	5.0	5.0
BA01	17.1	80.8	2.1	0.36	157	51.1	0.340	5.9	0.026	4.9	8.9	14.1	5.0	5.0	5.0	5.0	5.0
BA02	42.1	51.1	6.8	0.58	140	26.3	0.350	6.2	0.038	5.4	6.7	32.6	5.0	5.0	5.0	5.0	5.0
MA01	24.7	73.7	1.6	0.20	50	33.1	0.350	4.2	0.100	3.0	6.7	40.4	5.0	5.0	5.0	5.0	5.0
MA02	32.5	57.3	10.2	1.10	128	13.1	0.350	9.6	0.027	6.9	6.1	44.3	5.0	5.0	5.0	5.0	5.0
TG01	10.0	90.0	1.0	0.44	106	8.2	0.410	4.7	0.014	2.2	2.3	48.2	2.5	2.5	2.5	2.5	2.5
TG02	10.0	90.0	1.0	0.35	50	18.1	0.350	4.9	0.033	3.1	7.0	13.1	2.5	2.5	2.5	2.5	2.5
PP01	21.0	77.0	2.0	0.24	67	8.2	0.188	7.8	0.010	2.4	3.3	25.2	7.0	9.0	12.0	2.5	7.0
PP02	0.1	99.0	0.1	0.18	78	9.4	0.172	8.0	0.007	1.5	2.6	31.2	2.5	2.5	5.7	2.5	2.5
CE01	2.0	98.0	1.0	0.18	125	6.2	0.025	7.8	0.011	4.8	1.6	71.0	2.5	2.5	2.5	2.5	2.5
CE02	3.0	97.0	1.0	0.07	22	3.7	0.025	4.5	0.020	4.3	1.4	45.1	2.5	2.5	2.5	2.5	2.5

Table 4
Granulometry and concentration of TOC, P, Metals and organic pollutants for the Period 31st March 2013 – 1st April 2014.

Code	Coarse (%)	Sand (%)	Pelite (%)	TOC (mg/kg ds)	P_{tot} (mg/kg ds)	As (mg/kg ds)	Cd (mg/kg ds)	Cr_{TOT} (mg/kg ds)	Hg (mg/kg ds)	Ni (mg/kg ds)	Pb (mg/kg ds)	ΣPCB (ppm)	B(a)P (ppm)	B(b)F (ppm)	$B_{ghi}P$ (ppm)	B(k)F (ppm)	I(cd)P (ppm)
PE01	0.0	100.0	0.00	0.10	17	0.8	0.100	0.9	0.050	0.7	2.4	19.1	5.0	5.0	5.0	5.0	5.0
PE02	0.0	100.0	0.00	0.10	23	1.1	0.100	1.2	0.100	0.9	3.1	n.d.	5.0	5.0	5.0	5.0	5.0
VI01	0.0	99.0	1.0	0.20	927	0.1	0.200	10.7	0.050	1.1	5.8	n.d.	5.0	5.0	5.0	5.0	5.0
VI02	0.0	73.0	27.0	0.40	982	0.2	0.300	16.4	0.050	1.6	3.7	n.d.	5.0	5.0	5.0	5.0	5.0
MI01	0.0	98.0	2.0	0.60	1381	1.6	0.200	13.6	0.100	1.6	2.0	n.d.	5.0	5.0	5.0	5.0	5.0
MI02	0.0	78.0	22.0	0.70	825	0.1	0.200	7.4	2.400	2.4	3.6	n.d.	5.0	5.0	5.0	5.0	5.0
MT01	0.0	64.2	35.8	0.20	214	1.5	0.200	21.4	0.100	3.3	2.3	n.d.	5.0	5.0	5.0	5.0	5.0
MT02	0.0	82.8	17.2	0.60	368	1.3	0.100	10.8	0.050	3.6	4.2	n.d.	5.0	5.0	5.0	5.0	5.0
MN01	0.0	86.0	14.0	0.50	95	0.1	0.050	3.9	0.050	0.2	7.2	n.d.	5.0	5.0	5.0	5.0	5.0
MN02	0.0	72.0	28.0	1.10	54	0.1	0.050	2.5	0.050	0.2	13.9	n.d.	5.0	5.0	5.0	5.0	5.0
FC01	0.0	98.6	1.4	1.60	407	10.0	0.200	25.0	0.200	11.4	6.4	n.d.	5.0	5.0	5.0	5.0	5.0
FC02	31.5	59.3	9.2	1.70	433	11.5	0.200	4.5	0.100	9.5	5.5	n.d.	5.0	5.0	5.0	5.0	5.0
CR01	0.0	80.0	20.0	1.00	619	4.0	0.200	10.3	0.100	7.2	8.9	n.d.	5.0	5.0	5.0	5.0	5.0
CR02	2.0	3.5	94.5	1.10	382	3.1	0.100	4.3	0.050	3.5	4.7	n.d.	5.0	5.0	5.0	5.0	5.0
AL01	0.1	91.6	8.4	0.10	880	4.0	0.100	19.0	0.005	13.0	5.0	n.d.	0.5	0.5	0.5	0.5	0.5
AL02	4.5	94.3	1.2	0.10	515	18.0	0.200	11.0	0.005	6.0	4.0	n.d.	0.5	0.5	0.5	0.5	0.5
CM01	0.1	91.7	8.3	0.10	1110	3.0	0.100	19.0	0.005	11.0	4.0	n.d.	0.5	0.5	0.5	0.5	0.5
CM02	9.4	83.6	7.0	0.40	519	11.0	0.200	22.0	0.025	12.0	6.0	n.d.	0.5	0.5	0.5	0.5	0.5
FO01	0.1	92.3	7.7	0.10	1100	2.0	0.100	27.0	0.025	21.0	6.0	n.d.	0.5	0.5	0.5	0.5	0.5
FO02	0.1	60.0	39.9	0.40	620	5.0	0.100	15.0	0.025	9.0	5.0	n.d.	0.5	3.0	3.0	0.5	0.5
BI01	0.1	40.2	59.8	0.30	686	11.0	0.100	34.0	0.005	20.0	11.0	77.3	2.0	3.0	3.0	1.0	2.0
BI02	0.1	29.3	70.7	0.60	810	13.0	0.200	57.0	0.030	29.0	15.0	69.8	2.0	3.0	3.0	1.0	2.0
ML01	25.4	72.6	1.9	0.10	347	19.0	0.100	14.0	0.050	9.0	9.0	n.d.	2.0	2.0	2.0	0.0	2.0
ML02	0.1	20.2	79.8	0.70	771	11.0	0.200	54.0	0.030	30.0	15.0	n.d.	17.0	19.0	11.0	8.0	9.0
BB01	0.1	98.5	1.4	0.20	305	18.0	0.100	12.0	0.005	9.0	7.0	n.d.	0.5	0.5	0.5	0.5	0.5
BB02	13.6	83.9	2.5	0.40	186	2.0	0.300	13.0	0.010	5.0	5.0	n.d.	7.0	19.0	12.0	9.0	12.0
BA01	4.2	92.6	3.2	0.40	350	18.0	0.100	13.0	0.070	9.0	9.0	n.d.	2.0	1.0	1.0	0.5	1.0
BA02	27.0	69.8	3.2	0.40	370	20.0	0.100	14.0	0.040	10.0	10.0	n.d.	1.0	2.0	1.0	0.5	1.0
MA01	14.3	84.0	1.7	0.60	313	23.0	0.200	13.0	0.010	15.0	9.0	n.d.	4.0	6.0	3.0	2.0	3.0
MA02	24.1	71.9	4.0	0.50	462	22.0	0.300	17.0	0.010	11.0	10.0	n.d.	5.0	6.0	4.0	3.0	4.0
TG01	32.0	68.0	0.00	0.50	88	3.4	0.033	2.4	0.006	1.9	6.9	n.d.	2.5	2.5	2.5	2.5	2.5
TG02	20.0	79.0	1.0	0.50	53	4.4	0.037	1.7	0.007	0.8	4.5	n.d.	2.5	2.5	2.5	2.5	2.5
PP01	1.0	99.0	1.0	0.20	126	7.8	0.025	7.5	0.005	2.9	2.9	n.d.	2.5	2.5	2.5	2.5	2.5
PP02	3.0	97.0	1.0	0.20	99	8.3	0.025	6.4	0.005	2.0	2.5	n.d.	2.5	2.5	2.5	2.5	2.5
CE01	10.0	90.0	0.00	0.20	92	1.7	0.062	4.1	0.016	0.7	2.4	33.1	2.5	2.5	2.5	2.5	2.5
CE02	5.0	95.0	0.00	0.20	40	1.6	0.025	2.6	0.030	0.9	3.1	27.1.	2.5	2.5	2.5	2.5	2.5

Table 5
Granulometry and concentration of TOC, P, Metals and organic pollutants for the Period 31st March 2014 – 1st April 2015.

Code	Coarse (%)	Sand (%)	Fine Fraction (%)	TOC (mg/kg ds)	Ptot (mg/kg ds)	As (mg/kg ds)	Cd (mg/kg ds)	Cr Tot. (mg/kg ds)	Hg (mg/kg ds)	Ni (mg/kg ds)	Pb (mg/kg ds)	Sum PCB (ppm)	B(a)P (ppm)	B(b)F (ppm)	B(ghi)P (ppm)	B(k)F (ppm)	I(cd)P (ppm)
PE01	0.0	100.0	0.00	0.1	83.2	3.6	0.1	3.5	0.2	2.9	15.0	13.6	5.0	5.0	5.0	5.0	5.0
PE02	0.0	93.0	7.0	0.5	180.4	2.8	0.2	13.6	0.3	7.6	14.0	42.5	5.0	5.0	5.0	5.0	5.0
VI01	0.0	98.0	2.0	0.2	146.8	8.0	0.1	9.3	0.2	7.4	5.5	28.7	5.0	5.0	5.0	5.0	5.0
VI02	0.0	99.0	1.0	0.2	127.3	5.0	0.1	8.5	0.6	5.6	6.1	5.0	5.0	5.0	5.0	5.0	5.0
MI01	0.0	100.0	0.00	0.1	82.7	5.0	0.2	2.1	0.1	1.1	3.5	11.2	5.0	5.0	5.0	5.0	5.0
MI02	0.0	99.0	1.0	0.2	192.0	1.8	0.1	5.9	0.1	3.3	6.0	5.0	5.0	5.0	5.0	5.0	5.0
MT01	0.0	94.0	6.0	0.5	147.6	3.5	0.1	10.8	0.1	6.3	9.3	5.0	5.0	5.0	5.0	5.0	5.0
MT02	0.0	87.3	12.7	0.5	130.6	3.8	0.1	13.9	0.1	8.6	13.1	5.0	5.0	5.0	5.0	5.0	5.0
MN01	0.0	96.0	4.0	0.5	199.4	4.3	0.1	21.3	0.1	11.1	15.1	12.6	5.0	5.0	5.0	5.0	5.0
MN02	0.0	72.8	27.2	1.1	125.0	4.2	0.1	14.5	0.1	16.6	23.2	5.0	5.0	5.0	5.0	5.0	5.0
FC01	0.0	100.0	0.00	1.3	236.3	0.5	0.1	3.4	0.1	1.7	2.7	23.9	5.0	5.0	5.0	5.0	5.0
FC02	0.0	76.0	24.0	1.3	132.3	0.3	0.1	3.2	0.1	1.7	1.9	5.0	5.0	5.0	5.0	5.0	5.0
CR01	0.0	93.0	7.0	0.1	163.4	8.9	0.1	4.7	0.3	2.5	4.8	18.9	5.0	5.0	5.0	5.0	5.0
CR02	0.0	88.0	12.0	1.3	356.3	5.8	0.2	12.3	0.1	10.9	13.5	25.0	5.0	5.0	5.0	5.0	5.0
AL01	1.2	90.0	8.8	0.2	370.0	15.6	0.1	12.1	0.0	10.1	7.2	5.0	2.5	2.5	2.5	2.5	2.5
AL02	2.9	93.9	3.2	0.1	230.0	10.6	0.1	7.0	0.0	8.9	4.0	5.0	2.5	2.5	2.5	2.5	2.5
CM01	11.0	86.6	2.4	0.1	190.0	30.5	0.1	6.3	0.0	8.5	6.3	12.2	2.5	2.5	2.5	2.5	2.5
CM02	20.6	75.9	3.5	0.2	210.0	25.7	0.1	7.0	0.0	8.5	6.2	5.0	2.5	2.5	2.5	2.5	2.5
FO01	0.1	95.6	4.4	0.1	650.0	3.0	0.1	11.0	0.0	15.0	15.0	5.0	2.5	2.5	2.5	2.5	2.5
FO02	0.1	28.3	71.7	0.6	540.0	7.9	0.1	29.4	0.0	28.7	14.9	49.9	2.5	2.5	2.5	2.5	2.5
BI01	0.1	53.2	46.8	0.2	520.0	11.0	0.1	17.0	0.0	16.0	9.0	66.1	2.5	2.5	2.5	2.5	2.5
BI02	0.1	27.8	72.2	0.7	490.0	13.0	0.1	33.0	0.0	27.0	16.0	68.4	14.8	19.0	12.4	28.8	9.6
ML01	12.5	86.7	0.8	0.2	370.0	32.8	0.1	15.6	0.0	14.6	9.3	12.8	2.5	2.5	2.5	2.5	2.5
ML02	0.1	21.9	78.2	0.7	490.0	12.2	0.1	34.3	0.0	30.7	17.2	103.8	15.3	21.2	13.2	28.0	9.8
BB01	0.2	98.0	1.8	0.2	170.0	25.0	0.1	3.0	0.0	10.0	6.0	5.0	2.5	2.5	2.5	2.5	2.5
BB02	14.8	83.4	1.8	0.5	230.0	40.0	0.1	4.0	0.0	12.0	7.0	10.0	2.5	2.5	2.5	2.5	2.5
BA01	13.5	85.2	1.3	0.3	150.0	25.0	0.1	3.0	0.0	13.0	6.0	17.8	2.5	2.5	2.5	2.5	2.5
BA02	3.8	94.7	1.5	0.1	220.0	26.0	0.1	3.0	0.0	10.0	7.0	24.5	2.5	2.5	2.5	2.5	2.5
MA01	7.1	91.1	1.7	0.2	170.0	26.0	0.1	7.0	0.0	8.0	7.0	11.1	2.5	2.5	2.5	2.5	2.5
MA02	31.8	65.8	2.4	0.3	240.0	40.0	0.1	8.0	0.0	9.0	9.0	10.6	2.5	2.5	2.5	2.5	2.5
TG01	33.0	63.0	4.0	0.4	74.0	6.2	0.0	2.1	0.0	4.0	3.7	10.6	2.5	2.5	2.5	2.5	2.5
TG02	33.0	63.0	4.0	0.3	70.3	11.5	0.0	2.0	0.0	3.4	4.7	28.5	2.5	2.5	2.5	2.5	2.5
PP01	1.0	98.0	1.0	0.2	47.0	5.3	0.1	6.5	0.0	3.1	2.8	5.0	2.5	2.5	2.5	5.4	2.5
PP02	1.0	99.0	0.5	0.1	37.0	6.7	0.0	6.4	0.0	2.3	4.1	13.2	15.5	7.4	8.5	9.4	10.9
CE01	2.0	98.0	0.5	0.2	229.0	7.6	0.1	9.2	0.0	1.8	2.5	29.2	5.3	2.5	10.4	5.5	2.5
CE02	14.0	86.0	0.5	0.4	115.0	9.3	0.1	4.8	0.0	2.1	4.2	24.1	2.5	2.5	2.5	2.5	2.5

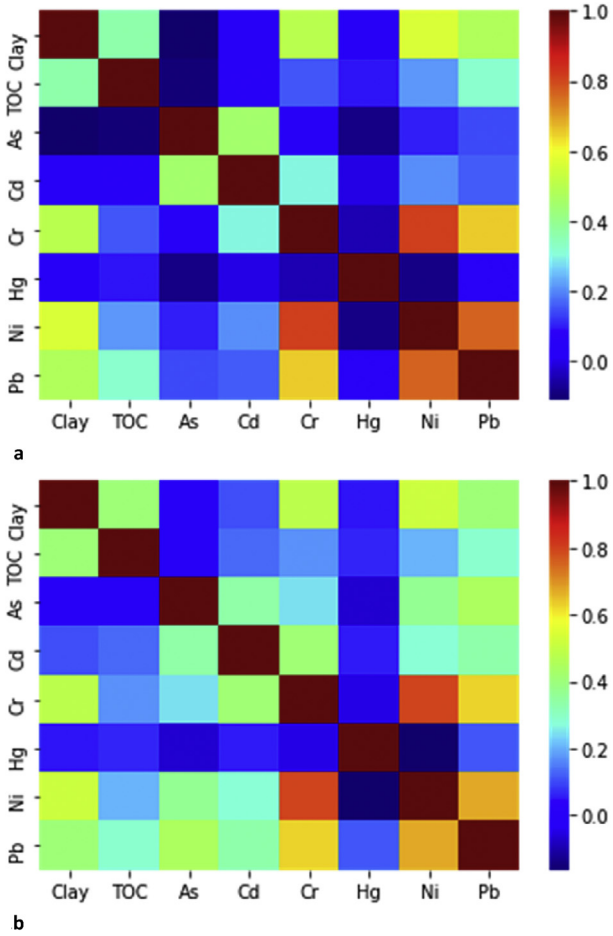


Fig. 1. (a) Correlation analysis using Pearson *r*. (b) Correlation analysis using Spearman ρ showing the strong correlations between nickel, chromium and lead.

3. Correlation analyses (CA)

Correlations over eight different metal concentration (**predictors)**, total organic carbon (TOC) and clay content is investigated through Pearson correlation coefficient*r*, which evaluates linear relationships, and Spearman correlation coefficient ρ, which evaluates monotonic relationships. The results are illustrated graphically for both r and ρ are in (Fig. 1a and b), where warmer colors indicate higher correlation values, while colder colors indicate lower correlation values.

The relative distribution of Cr/Ni (Fig. 2a) and Ni/Pb (Fig. 2b) are shown through two-dimensional scatter plots.

Correlation analysis using the Kendall $\tau$are also reported. The results of the correlation analysis using Kendall τ are shown in Fig. 3.

Least Angle RegreSsion (*LARS*) has been proposed to investigate on relationship of TOC and Clay content with metal concentration. LARS was used to overcome the drawbacks of stepwise regression [6,8,9]. The algorithm works as follow:

- Start with all the coefficients $\beta = 0$;

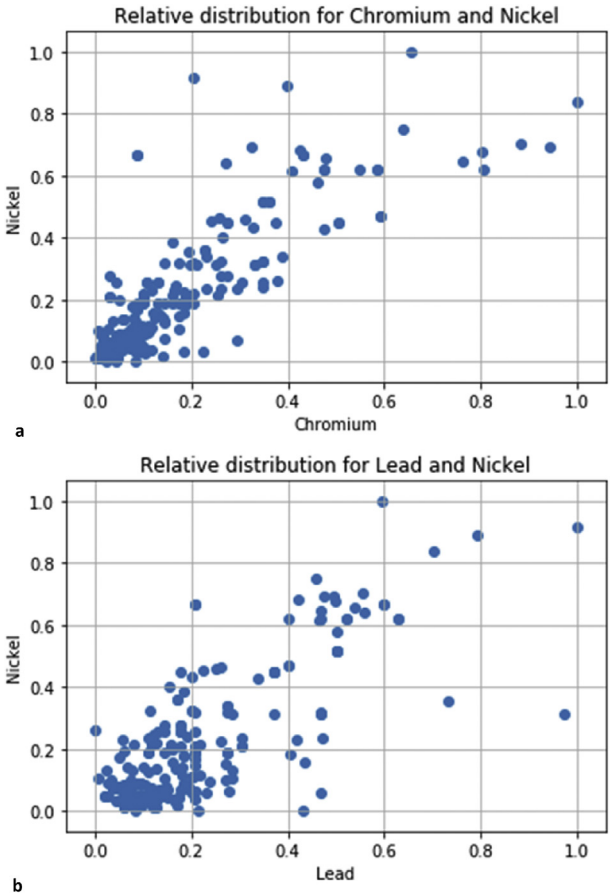


Fig. 2. (a) Relative distribution of Chromium and Nickel. (b) Relative distribution of Lead and Nickel.

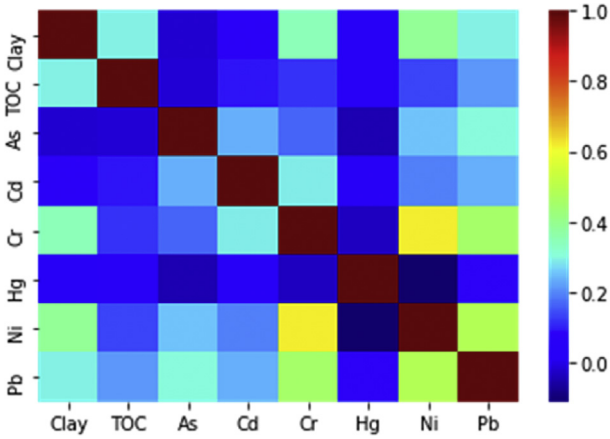


Fig. 3. Correlation analysis using Kendall τ.

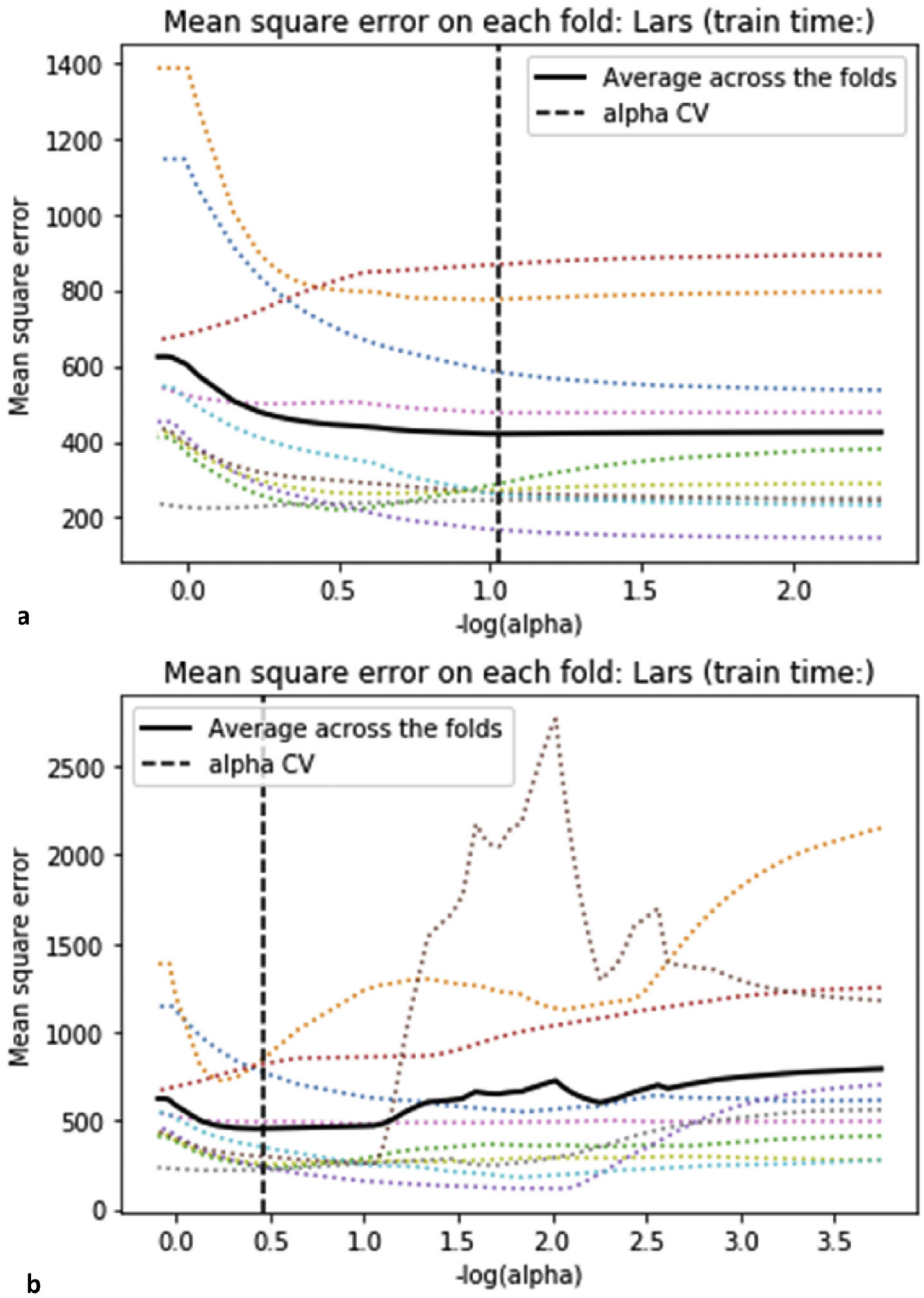


Fig. 4. (a) Cross-validation using a linear regressor for clay. (b) Cross-validation using a polynomial regressor for clay.

- Find the predictor x_i which is most correlated with the residual ρ (that is, the predictor which makes the 'least angle' with the residual); increase β_i accordingly;
- Move in the direction of $x_i$until $r(x_i, \varepsilon) = r(x_j, \varepsilon)$ for a predictor x_j; increase $\beta_j$accordingly;
- Repeat the procedure until all the predictors are in the model.

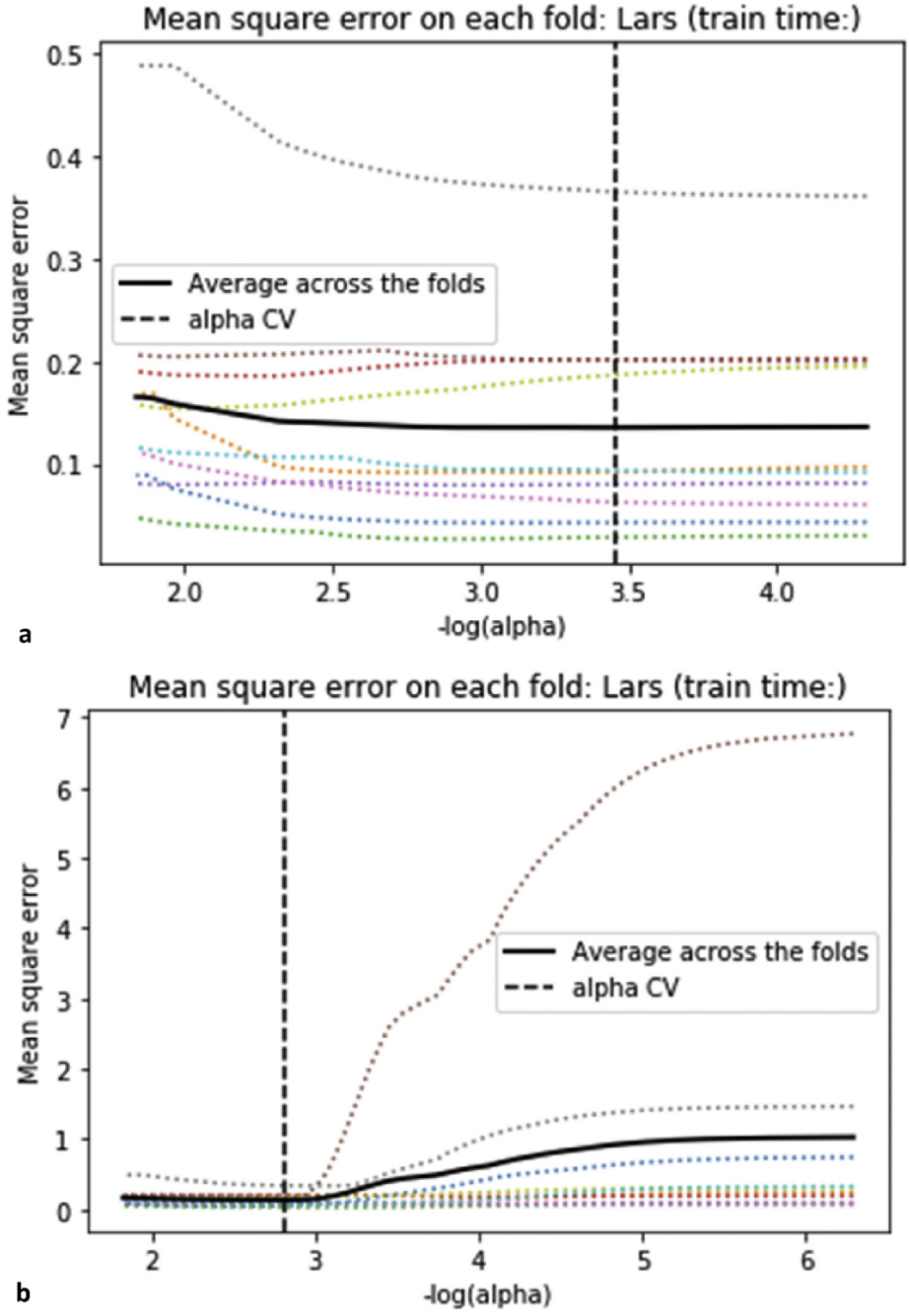


Fig. 5. (a) Cross-validation using a linear regressor for TOC. (b) Cross-validation using a polynomial regressor for TOC.

Results of regression are validated using a standard k-fold cross validation [7]. Specifically, the value of the mean square error between the regression function and the real values will be minimized over the 20 rounds of cross-validation.

Fig. 4a and b show results for the 20 validation rounds using a linear regressor and a polynomial regressor for Clay and TOC. The linear model describing correlation of clay with metals is given by the equations (1) and (2), while for TOC is reported in the equations (3) and (4).

$$clay = -0.27*As + Cr*0.19 + Pb*0.51 + 0.96*Ni \quad (1)$$

$$clay = 0.95*Hg*Ni - 0.17*As + 0.2*Cr + 0.53*Ni + 0.03*Ni*Pb \quad (2)$$

4. Relation between TOC and metals

$$TOC = 0.2*Cd + 0.15*Hg + 0.03*Pb - 0.005*As - 0.002*Cr \quad (3)$$

$$TOC = Hg*Ni*0.06 - As*0.002 + Pb*0.001 \quad (4)$$

Acknowledgments

The authors thank ARPA Puglia (Italy) Laboratories for chemical analyses.

Transparency document

Transparency document associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2019.103860.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2019.103860.

References

[1] M. Mali, N. Ungaro, A. Cardellicchio, M.M. Dell'Anna, G. Romanazzi, P. Mastrorilli, L. Damiani, Long-term Monitoring Programs to Assess Environmental Pressures on Coastal Area: Weighted Indexes and Statistical Elaboration, as Handy Tools for Decision-Makers, Accepted for publication in Ecological Indicators, 2019. https://doi.org/10.1016/j.ecolind.2019.01.085.
[2] M. Mali, M.M. Dell'Anna, P. Mastrorilli, L. Damiani, A.F. Piccinni, Assessment and source identification of pollution risk for touristic ports: heavy metals and polycyclic aromatic hydrocarbons in sediments of 4 marinas of the Apulia region (Italy), Mar. Pollut. Bull. 114 (2) (2017) 768–777.
[3] M. Mali, M.M. Dell'Anna, M. Notarnicola, L. Damiani, P. Mastrorilli, Combining chemometric tools for assessing hazard sources and factors acting simultaneously in contaminated areas, Case study: "Mar Piccolo" Taranto (South Italy) 184 (2017) 784–794.
[4] M. Mali, D. Malcangio, M.M. Dell'Anna, L. Damiani, P. Mastrorilli, Influence of hydrodynamic features in the transport and fate of hazard contaminants within touristic ports, Case study: Torre a Mare (Italy. Heliyon. 4 (1) (2018).
[5] J.W. Tukey, Exploratory Data Analysis, vol. 2, 1977, pp. 152–172.
[6] P.L. Flom, D.L. Cassell, "Stopping stepwise: why stepwise and similar selection methods are bad, and what you should use," NESUG 2007, 2007.
[7] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, Ijcai 14 (2) (1995, August) 1137–1145.
[8] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression, Ann. Stat. 32 (2) (2004) 407–499.
[9] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B (1996) 267–288.