Title: Multiple regression analysis to assess the spatial distribution and speciation of mercury in surface sediments of a contaminated lagoon

Authors: T. Stoichev, E. Tessier, J.P. Coelho, M.G. Lobos Valenzuela, M.E. Pereira, D. Amouroux

PII:	S0304-3894(18)31256-1
DOI:	https://doi.org/10.1016/j.jhazmat.2018.12.109
Reference:	HAZMAT 20143
To appear in:	*Journal of Hazardous Materials*
Received date:	5 October 2018
Revised date:	12 December 2018
Accepted date:	27 December 2018

Please cite this article as: Stoichev T, Tessier E, Coelho JP, Lobos Valenzuela MG, Pereira ME, Amouroux D, Multiple regression analysis to assess the spatial distribution and speciation of mercury in surface sediments of a contaminated lagoon, *Journal of Hazardous Materials* (2018), https://doi.org/10.1016/j.jhazmat.2018.12.109

Multiple regression analysis to assess the spatial distribution and speciation of mercury in surface sediments of a contaminated lagoon

T. Stoichev[1]*, E. Tessier[2], J.P. Coelho[3,4], M.G. Lobos Valenzuela[5], M.E. Pereira[4], D. Amouroux[2]**

[1]Interdisciplinary Center of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, Terminal de Cruzeiros de Leixões, Av. Norton de Matos, 4450-208 Matisinhos, Portugal

* tstoichevbg@yahoo.com; Tel: +351 220402 611

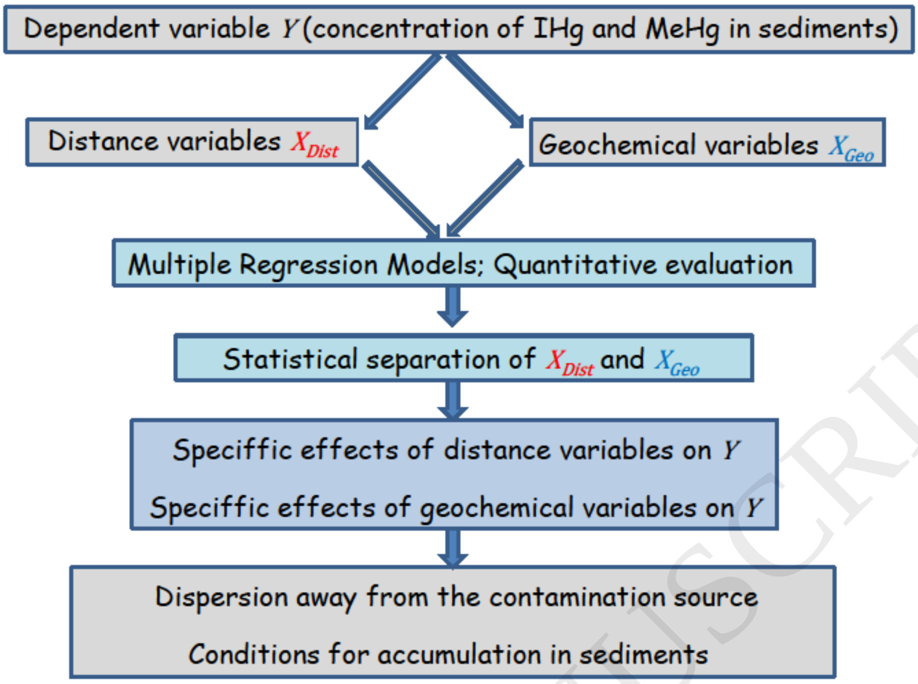
[2]CNRS/ UNIV PAU & PAYS ADOUR/ E2S UPPA, INSTITUT DES SCIENCES ANALYTIQUES ET DE PHYSICOCHIMIE POUR L'ENVIRONNEMENT ET LES MATERIAUX – MIRA, UMR5254, 64000, PAU, France
** david.amouroux@univ-pau.fr; Tel:+33 559407756

[3]Department of Biology & CESAM & ECOMARE, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

[4]Department of Chemistry & CESAM, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

[5]Laboratory of Analytical and Environmental Chemistry, Institute of Chemistry and Biochemistry, University of Valparaíso, Valparaíso, Chile

Graphical abstract



Highlights

- Inorganic mercury and methylmercury studied in sediments of contaminated lagoon
- Multiple regression model demonstrated different behavior for mercury species
- Inorganic mercury (IHg) depends on the distance to the IHg source and Al levels
- Methylmercury (MeHg) depends on IHg levels in contaminated sediments
- MeHg depends also on the levels of Ca (marine proxy), Mn and organic matter

ABSTRACT

The concentrations of inorganic mercury (IHg) and methylmercury (MeHg) in surface sediments from the contaminated Aveiro Lagoon (Portugal) were determined by species-specific isotope dilution analysis. Different behaviour of IHg and MeHg was observed based upon multiple regression analysis, including geochemical characteristics of the surface sediments (fine fraction, concentrations of organic matter and metals) and non-Euclidean distances between sampling points. This data treatment method was valid over the entire concentration range of IHg and MeHg, allowing robust quantitative evaluation with respect to extrapolation. For IHg, there was statistical separation of the dispersion away from the contamination source and of Al concentration in the sediments. The MeHg concentrations followed those of IHg at high concentrations. The geochemical variables, such as concentrations of Ca (marine influence proxy), Mn and organic matter, were necessary to describe the behaviour of MeHg across the whole concentration range. The models for MeHg demonstrated that, close to the mouth of the lagoon, net production of MeHg was higher. In future, multiple regression analysis could be applied to separate and to evaluate quantitatively the effects of geochemistry and dispersion away from the contamination source in sediments contaminated with other substances.

Keywords: Chlor-alkali industry; Mercury species; Coastal lagoon; Sediments; Multiple regression model

1. Introduction

Inorganic mercury (IHg(II)) is the main form of mercury (Hg) in the aquatic environment. However, natural transformations lead to in situ production of toxic methylmercury (MeHg) with a tendency for bioaccumulation and biomagnification in aquatic food webs [1].

Although Hg is a global pollutant, local Hg pollution from the chlor-alkaly industry affects various types of aquatic environments [2-5] but shallow areas could be especially vulnerable [6]. The Aveiro Lagoon has suffered from chlor-alkali Hg pollution coming from a single upstream source [7]. The transport of Hg in shallow coastal environments is strongly site-dependent [8]. Most of the Hg in the Aveiro Lagoon is confined within the sediments of a relatively small zone of the lagoon – Laranjo Bay [9]. In Laranjo Bay, MeHg has only rarely been found in sediments [10-12]. However, the methods previously applied did not allow the determination of MeHg in surface sediments from less contaminated areas in the Aveiro Lagoon. Nevertheless, MeHg, produced in sediments, is enriched in water [13]. As a result, Hg can rapidly accumulate in fish, attaining levels of 8.5 μmol kg^{-1} wet weight in *Dicentrarchus labrax* from Laranjo Bay [9]. Therefore, the behaviour of IHg and, especially, MeHg in sediments deserves further attention.

Multiple regression analysis could be applied to study the dynamics of IHg and MeHg in contaminated environments. A similar statistical approach was used to separate removal from addition processes of IHg and MeHg in water during estuarine mixing [13] as well as to evaluate contaminant biogeochemistry in estuarine water and sediment [14, 15]. This method allows for the verification of the effects of several explanatory variables and their interactions, as well as higher order effects. Upon simplification of the starting models, only a few significant explanatory variables will remain, resulting in minimal adequate models.

The aim of this work was to study spatial and seasonal variations of both IHg and MeHg concentrations in surface sediments from very contaminated shallow tidal environments. An attempt was made to separate and quantitatively evaluate the effects of contaminant dispersion from a point source from those of geochemical processes, using multiple regression analysis. This is possible if additive statistical effects of some explanatory variables (responsible for the contaminant dispersion and the geochemistry of the sediments) on the dependent variables (e.g. Hg species concentrations) are found. Specific behaviour of both IHg and MeHg in surface sediments will be presented.

2. Experimental

2.1. Study area

Aveiro Lagoon is a coastal lagoon in the North of Portugal (Fig. 1) comprised of a network of channels, opening into the Atlantic Ocean by way of a single channel. The lagoon covers an area of 83 km^2 at high tide (HT) and 66 km^2 at low tide (LT) and has an average depth of 1 m. The tidal range is 2 m (minimum 0.6 m during neap tide; maximum 3.2 m during spring tide). The water residence time in the lagoon is approximately 2 days. However, it is more than two weeks in the Hg-contaminated area [16]. The particle residence time is approximately 2-3 days in the lower areas of the lagoon and up to 14 days in the upper ends [17]. The Laranjo Bay received waste discharges of 27 tons of Hg from the 1950s to 1994 [18]. The population near the lagoon is 370 thousand [19] and in the total catchment area (3500 km^2) it is >600 thousand [20]. The main human activities are associated with the port (near the mouth), fisheries, agriculture, aquaculture, salt production, pulp mill and chemical industries.

2.2. Sampling

The sediment samples, named from 1 to 14, were collected in the summer of 2010. The remaining samples were collected during two campaigns (February 2012 and August 2012) in order to estimate eventual seasonal change. Samples were taken at low tide and during daylight from the top 5 cm layer, stored in plastic bags and kept cool until arrival at the laboratory. The sites were located: *i)* close to the lagoon entrance (BAR, GAF, SJA, 11, 13); *ii)* at the edges of the main channels (ARE, CAR, VAG, 10, 12, 14); *iii)* in the middle of the longest channel (TOR, 9); *iv)* near the main freshwater input (RIO); *v)* in the area with the highest levels of Hg pollution (LAR, 1-5); *vi)* in the central area of the lagoon (6–8).

Each sampling point is characterized by geographical coordinates and falls within a simplified polygon border, marked with hyphen line (Fig. 1). For distribution of the contaminants, each point is related to the non-Euclidian distance d_S to Hg source (**S**) and non-Euclidian distance d_M to the mouth of the lagoon (**M**), obtained within the simplified polygon border. Alternatively, each sampling point may be described by the distance ratio $DR_M = d_S/d_M$. Each point was additionally characterized using distance d_R to the reference point **R** (instead of d_M) obtained within the simplified polygon border (Fig. 1) or by reference distance ratio $DR_R = d_S/d_R$.

2.3. Chemical analysis

At the laboratory, macrodetritus were removed and samples were freeze-dried (Unicryo MC-4 L-60 ºC). Fine fraction (FF) content (<63 µm) was quantified by wet sieving. Succinctly, 1g aliquots of each sediment sample were sieved through a calibrated 63 µm mesh under running water and the retained particles dried in a forced air oven (100 ºC) until constant weight. Sediment organic matter (OM) content was analyzed through loss on ignition (6 h combustion at 500 ºC).

Each sample was ground with agate mortar and sieved (300 µm) before dark storage until analysis. Total carbon (TC) and organic carbon (OC, carbonate removed according to ISO 10694) were quantified in a CHNS analyser (Leco Truspec Micro). Sediment digestions for metal determination were performed using microwave field with aqua regia (US EPA 3051). The metals (Ca, Al, Mn) were measured using inductively coupled plasma optical emission spectrometry (ICP-OES; Perkin-Elmer, model ICP Optima 2000 DV). The accuracy was checked using BCSS-1 marine sediment with measured values being between 98% and 102% of the certified ones.

Hg_{TOT} in sediments was determined by Atomic Absorption Spectroscopy (AAS) after thermal decomposition, using an Automated Mercury Analyzer (AMA-254 Altec, Symalab, France) according to the ISO 10694 standard [21]. The results for Hg_{TOT} were used as preliminary information in order to proceed to Hg speciation analysis.

Hg species were extracted from sediments (sample weight 10–290 mg) with 6 M HNO_3 using microwave irradiation (Discover SP-D, CEM Explorer). Between 12 and 600 mg of the extracts were spiked with species-specific stable Hg isotopes (^{199}IHg(II) and ^{201}MeHg) to overcome non quantitative recoveries due to matrix effects or analyte transformation [22, 23]. The Hg species were derivatized with sodium tetraethylborate at pH 3.9 and extracted into 2,2,4-trimethylpentane. The organic layer was analysed by capillary GC (Trace GC Ultra with autosampler TriPlus RSH, Thermo Scientific) coupled to ICP-MS (X Series 2, Thermo Scientific).

The repeatability was determined as the relative standard deviations (RSD) from the triplicate injections (average from all samples). Values 1.2% and 6.5% were obtained for IHg and MeHg, respectively. The results for IHg concentrations in the samples were corrected with the respective blank value (0.518 ± 0.050 pmol). The method limit of detection (LOD) for IHg was calculated as three times the standard deviation (SD) of the concentrations found in the blank extraction samples. Since MeHg was not detected in the blanks, its LOD was estimated from the calculation of the Background noise Equivalent Concentrations (BEC) and expressed as LOD = BEC + 3xSD. For sample weight of 250 mg, the LODs were 0.60 nmol kg^{-1} and 0.11 nmol kg^{-1} for IHg and MeHg, respectively. The concentrations of IHg (4.05 ± 0.08 μmol kg^{-1}) and MeHg (26.4 ± 1.2 nmol kg^{-1}) in reference material estuarine sediment IAEA 405 (n=10) show good agreement with the reference values of 4.01 ± 0.19 μmol kg^{-1} for IHg and 27.4 ± 2.6 nmol kg^{-1} for MeHg. In five of the samples from the Aveiro Lagoon (1, 4, 5, 14, Bar (Summer)), MeHg was not detected and the concentrations, being between 0 and LOD, were replaced by half of LOD to integrate them in the statistical analysis. Further, the sum of the concentrations of IHg and MeHg is referred to here as Hg_{TOT} and was in the same order as Hg_{TOT} measured by AAS ($Hg_{TOT\ (GC\text{-}ICPMS)}=(0.830 \pm 0.042)Hg_{TOT\ (AAS)}$, $R^2=0.981$).

2.4. Statistical treatment

Statistical data treatment was carried out using R software [24]. The dependent variables (Y), such as Hg species concentrations, were represented as functions of q explanatory variables X_i. One type of the explanatory variables is related to geographical coordinates (either both distances d_S, d_M or their ratio DR_M). Other distances (d_S, d_R) or their ratio DR_R were also used. The other part of explanatory variables represents the geochemical characteristics of the samples (FF, OM, OC, TC and concentrations of Al, Ca, Mn). The concentrations of Al and Ca in sediments were proxies for continental and ocean contribution, respectively [17, 25]. Organic matter and oxyhydroxides of Mn(IV) are recognized as scavengers of toxic metals in Aveiro Lagoon [26] and of Hg species in other coastal environments [27]. For the case of MeHg, one of the explanatory variables was IHg concentration (precursor of MeHg). The data is given in the supplementary material (Aveiro_Sediments_DistanceR.xlsx) as well as the R scripts (IHg_Dist_DR_S_M, MeHg_Dist_DR_S_M, IHg_Dist_DR_S_R, MeHg_Dist_DR_S_R).

Both Y and X_i were separately normalized, if necessary, using graphical visualization and Box-Cox transformations [28] to give variables Y_T and $X_{T,i}$. Some of the explanatory variables (Al, Mn, TC, d_S, d_M, d_R) were not altered, with only the following transformations finally being selected:

$$IHg_T = \ln(IHg) \qquad MeHg_T = \ln(MeHg) \qquad Ca_T = \ln(Ca)$$

$$OM_T = \ln(OM) \qquad FF_T = \sqrt{FF} \qquad OC_T = \sqrt{OC} \qquad (1)$$

$$DR_{M,T} = \ln(DR_M) = \ln(d_S/d_M) \qquad DR_{R,T} = (DR_R)^{-0.5} = \sqrt{d_R/d_S}$$

The relationship between the explanatory variables is expressed by L_1, L_2 or L_3, the index representing the highest interaction order, described in the starting model:

$$L_3(X_{T,1}, X_{T,2}, \dots, X_{T,i}, \dots, X_{T,q}) = \sum_{i=1}^{q} a_i\, X_{T,i} + \sum_{i\neq j} a_{i,j}\, X_{T,i}X_{T,j} + \sum_{i\neq j\neq k} a_{i,j,k}\, X_{T,i}X_{T,j}X_{T,k}$$

$$L_2\left(X_{T,1}, X_{T,2}, \dots, X_{T,i}, \dots, X_{T,q}\right) = \sum_{i=1}^{q} a_i\, X_{T,i} + \sum_{i \neq j} a_{i,j}\, X_{T,i} X_{T,j}$$

(2)

$$L_1\left(X_{T,1}, X_{T,2}, \dots, X_{T,i}, \dots, X_{T,q}\right) = \sum_{i=1}^{q} a_i\, X_{T,i}$$

The regression coefficients a_i represent the simple terms for variable X_i. The coefficients a_{ij} and a_{ijk}, account for the double and triple interactions, respectively. The starting models, used to describe IHg concentrations, can be written with the equations:

$$\ln(IHg) = a_0 + L_2(d_S, d) + L_1\left({d_S}^2, d^2, X_{T,i}\right)$$ X_i: Mn, Al, TC, OC (3)

$$\ln(IHg) = a_0 + L_2(d_S, d) + L_1\left({d_S}^2, d^2, FF_T, Ca_T, OM_T, Mn, X_{T,i}\right)$$ (4)

$$\ln(IHg) = a_0 + L_2\left(DR_T, X_{T,i}\right) + L_1\left(DR_T^2\right)$$ X_i: Mn, Al, TC, OC (5)

$$\ln(IHg) = a_0 + L_1\left(DR_T, DR_T^2, FF_T, Ca_T, OM_T, Mn, X_{T,i}\right)$$ (6)

$$\ln(IHg) = a_0 + L_3\left(DR_T, FF_T, X_{T,i}\right) + L_1\left(DR_T^2\right)$$ X_i: OC, TC (7)

Eq. 3 and Eq. 4 represented IHg concentration as a function of d_S and d (which can be either d_M or d_R) while Eqs 5-7 used *DR* (either DR_M or DR_R). Numerous geochemical characteristics were checked simultaneously without accounting for interaction effects (Eqs. 4, 6). Some of the starting equations accounted for interactions between distances (Eqs. 3, 4) or between *DR* and geochemical variables (Eqs 5, 7). Quadratic effects of distances and of *DR* were checked in all starting models. In Eq. 4 and Eq. 6 X_i was Al, OC or TC. The last three variables (Al, OC, TC) were never considered together in order to avoid problems with co-linearity (ESM, R scripts).

The starting models for MeHg concentrations can be expressed with the equations:

$$\ln(MeHg) = a_0 + L_3\left(d_S, d, X_{T,i}\right) + L_1\left({d_S}^2, d^2\right) \quad (8)$$

X_i: IHg, Ca, Mn, FF, OM, OC, TC

$$\ln(MeHg) = a_0 + L_1\left(d_S, d, {d_S}^2, d^2, Ca_T, IHg_T\right) \quad (9)$$

$$\ln(MeHg) = a_0 + L_2\left(d_S, d, IHg_T, X_{T,i}\right) \qquad X_i\text{: OC, TC} \quad (10)$$

$$\ln(MeHg) = a_0 + L_2(d_S, d, FF_T) + L_1(IHg_T, Ca_T) \quad (11)$$

$$\ln(MeHg) = a_0 + L_1\left(DR_{M,T}, DR_{M,T}^2, IHg_T, FF_T, OM_T, Ca_T, Mn, X_{T,i}\right) \quad (12)$$

X_i: OC, TC

$$\ln(MeHg) = a_0 + L_2\left(DR_{M,T}, X_{T,i}\right) + L_1\left(DR_{M,T}^2\right) \qquad X_i\text{: Ca, IHg} \quad (13)$$

$$\ln(MeHg) = a_0 + L_3\left(DR_{M,T}, IHg_T, X_{T,i}\right) + L_1\left(DR_{M,T}^2\right) \quad (14)$$

X_i: Ca, Mn, FF, OM, OC, TC

$$\ln(MeHg) = a_0 + L_1\left(DR_{M,T}, IHg_T\right) + L_2\left(Ca_T, Mn, FF_T, X_{T,i}\right) \quad (15)$$

X_i: OM, OC, TC

In Eqs. 8-11, distances d_S and d (either d_M or d_R) were used, while DR_M was used in Eqs. 12-15. The geochemical characteristics as well as IHg concentration were included in the starting models. Some starting models included double (Eqs. 10, 11, 13, 15) or triple interactions (Eqs. 8, 14). To avoid problems with co-linearity, models using both DR_R and IHg were not developed.

The coefficients a_0 (intercept), a_i, a_{ij}, a_{ijk} were determined using multiple regression analysis [29]. The models were simplified by leaving only the coefficients significantly different from 0 ($p<0.05$). The number of the considered coefficients in the starting equations never exceeded 13 in order to avoid overparametrization.

First, the minimal adequate model was developed using all available data ($n=32$). Second, the applicability of the obtained model was checked after removing values of Y higher than 95th percentile ($P_{0.95}$) of all data and the regression coefficients for the minimal adequate model were determined again with the remaining dataset (short data, *n=30*). The obtained regression equation, using the short data, were extrapolated for data removed (extrapolated data, *n=2*). The minimal adequate model was considered further only if robust, meaning the same significant terms were obtained for all data and for short data.

All models were compared using graphical representation of model values Y_M against experimental ones Y_E and by root mean square deviation (*RMSD*) criteria (for all data, short data and extrapolated data) representing how different the model values are when compared with *n* experimental data of *Y*:

$$RMSD = \sqrt{\frac{\sum\left(Y_{E,i}-Y_{M,i}\right)^2}{n}} \tag{16}$$

3. Results and discussion

3.1. Seasonal variations

Different seasonal trends have been observed for Hg species concentration in surface sediments. The concentrations of MeHg (and %MeHg) were higher during warm seasons in contaminated Tagus estuary [30] and Ravenna Lagoon [31]. However, in surface sediments from the Aveiro Lagoon, general seasonal trend for Hg species concentrations was not observed (Fig. 2) and when, for some sampling sites, winter / summer differences existed, they usually followed concentration changes observed for geochemical variables (Figure SI-1). For both species; the highest concentrations were observed in LAR, near the Hg source. The time differences of IHg concentrations followed those found for Al (ANCOVA, $p<0.001$) and those of MeHg usually followed the trend for IHg (ANCOVA, $p<0.001$). Similar conditions with no clear seasonal difference in contaminated sediments have been observed in Venice Lagoon [32]. For the further treatment of the spatial variation, no seasonal differences were taken into account.

3.2. Spatial variation

The average concentration (range) for IHg in surface sediments from the Laranjo Bay (Fig. 1c) was 48.7 μmol kg^{-1} (5.2-210.0 μmol kg^{-1}, n=7) and from the rest of the lagoon, 0.426 μmol kg^{-1} (0.014-1.670 μmol kg^{-1}, n=25). These results confirmed the gradient observed previously for Hg_{TOT} concentrations [9]. Globally, a similar pattern was observed for MeHg, it being 17.70 (1.65-48.40 nmol kg^{-1}) in Laranjo Bay and 2.78 (0.13-14.90 nmol kg^{-1}) in the rest of the lagoon. The %MeHg was 0.09% (0.01-0.21%) and 0.85% (0.05-2.40%) in Laranjo Bay and in the rest of the lagoon, respectively. In surface sediments from Laranjo Bay, MeHg has previously been found only in one site (66 nmol kg^{-1}), representing <0.1% of Hg_{TOT} [11]. A similar situation was observed in other contaminated coastal environments, such as Tagus estuary [30] and Ravenna Lagoon [31].

The gradients of Hg species concentrations being large, the Aveiro Lagoon has areas that are extremely contaminated as well as almost pristine zones, giving rise to huge concentration ranges (especially for IHg, Table SI-1) compared to other coastal environments in Europe [23, 30, 31, 33-40]. If all the coastal zones, including the Aveiro Lagoon, were considered together, the concentration range of MeHg in surface sediments could be roughly described from the range of IHg (Figure SI-2). The concentrations of MeHg in surface sediments around the globe have been found proportional to IHg but reach a saturation value of 0.277 μmol kg^{-1} in very contaminated sites [41]. However, within each environment, the MeHg concentrations in surface sediments can be determined by IHg [6, 30, 42] though they may be more dependent upon biogeochemical characteristics of the sediments [5, 36, 38, 40, 42, 43].

To explain the spatial variations of Hg species concentrations in surface sediments, minimal adequate models were developed. The dispersion from the contamination source, geochemistry, and the influence of IHg concentrations (the last one being important for MeHg) were evaluated quantitatively.

3.2.1. IHg distribution in contaminated lagoon

Most of the models used to relate IHg concentrations with distances d_S and d_M were different for all data and short data. There were only two minimal adequate models, derived from Eq. 4, that were robust (Table 1). From the studied geochemical variables, only FF and Al concentrations were found to be important. There was significant interaction effect between d_S and d_M, which is difficult to interpret. More models were found robust when using d_S and d_R (Table SI-2). However, the high values of RMSD decreased the prediction of IHg concentration when d_S, d_M or d_R were used in the linear model construction.

When DR_M was used, the obtained models were always robust, with the same regression coefficients for all data and short data. The values of RMSD (Eq. 16) were usually low (Table 1), with the following equations being selected for further evaluation:

$$\ln(IHg) = a_0 - a_3\ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_8[Al] \tag{17}$$

$$\ln(IHg) = a_0 - a_3\ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_4\sqrt{FF} + a_8[Al] \tag{18}$$

$$\ln(IHg) = a_0 - a_3\ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_4\sqrt{FF} + a_7[TC] + a_{4,7}[TC]\sqrt{FF} \tag{19}$$

Equation 17 was derived from Eq. 5, with Al concentration being the most important geochemical predictor of the behavior of IHg in surface sediments (Table 1). Another essential property of the obtained models is that all interaction terms were non-significant. A significant interaction term was found between TC and FF (Eq. 19), meaning the effects of FF and TC on IHg concentrations are not so straightforward. However, the effects of geochemistry were separated from the effects of distance function DR_M. The results from Eqs 17-19 were compared to the experimental values for IHg concentrations in sediments for all, short and extrapolated data (Fig. 3). The extrapolating ability of the obtained models is about one order of magnitude of IHg concentration. The three minimal adequate models showed essentially the same results, and, therefore, Eq. 17 was finally selected due to its simplicity.

When DR_R was used instead of DR_M, several robust models were obtained (Table 2). The transformation of the variable DR_R was different from that of DR_M (Eq. 1). Applying the same (logarithmic) transformation for $\mathrm{DR_R}$ as for DR_M lead to models with much higher RMSD values (Table SI-3) and was therefore not considered further. From the minimal adequate models (Table 2) two were selected for having low RMSD values and for lacking interaction terms:

$$\ln(IHg) = a_0 + a_{3,3}/DR_R + a_8[Al] = a_0 + a_{3,3}d_R/d_S + a_8[Al] \qquad (20)$$

$$\ln(IHg) = a_0 - a_3/\sqrt{DR_R} + a_{3,3}/DR_R + a_4\sqrt{FF} + a_6\sqrt{[OC]} \qquad (21)$$

The results from Eq. 17 are derived using DR_M, which depends on both d_S and d_M. They could eventually describe processes related to mixing of contaminated water from the source with clean ocean water. However, the reference distance ratio DR_R depends on both d_S and d_R and is not related to processes of the mixing of contaminated water with ocean water. Using DR_R, Eq. 20 was obtained, which has similar model robustness and simplicity as Eq. 17.

In a hypothetic one-dimensional case, for an estuarine channel with known linear distance (*LD)* between the source **S** and the mouth **M** (or the reference **R**), each sampling point can be characterized only by its distance from the source as $d_M=LD-d_S$ (or $d_R=LD-d_S$). Then, both Eq. 17 and Eq. 20 can be expressed using only effects of d_S and Al concentration. In both equations, for fixed concentrations of Al (in order to leave only the dependence on d_S), the IHg concentration was inversely proportional to the fourth power of d_S. Therefore, an additional model was developed for two dimensions (Fig. 1) according to the equation:

$$IHg = c_1 exp(c_2[Al])/{d_S}^4 \quad (22)$$

The coefficients c_1 and c_2 were determined using non-linear regression and depended on the intensity of contamination (c_1) and the strength of geochemical associations between IHg and Al (c_2). The RMSDs for Eq. 22 were 5.5, 2.9 and 75.7 for all, short and extrapolated data, respectively, being among the best obtained for IHg. The comparison between model values and experimental ones (Fig. 3d) demonstrated that the model was reliable, as were Eqs 17-19, at least for high and medium concentrations of IHg. The transport of IHg could depend on the considered direction away from the contaminant source [36]. However, in the Aveiro Lagoon, when both geochemistry and distance measures were included in the model, it was possible to describe the dispersion of IHg with simple equations not dependent upon direction. However, in this case, the transport of the contaminant was not dependent on d_M. The dependence on d_S is extremely strong, localizing the Hg pollution very near the contamination source **S** (Fig. 4). A rapid decrease of IHg concentrations in sediments at higher distance from the source was also observed in other contaminated coastal environments, such as Gulf of Trieste [36] and Lavaca Bay [8].

An exponential dependence of IHg on Al concentration was observed in surface sediments from Vigo estuary [34]. Since Al was considered as indicative, not only of the particle size (i.e binding specific surface) but also mineral composition (mainly clay fraction), IHg from Vigo estuary, as in the Aveiro Lagoon, was bound to the finest particles in the sediments. It is known that the surface area of coastal sediments is grain-size dependent and controls the adsorption of metals [44]. Variations of Al levels in the sediments from Aveiro Lagoon have more important effects on IHg concentrations close to source **S** (Fig. 4). The sediments with intermediate Al concentration are contaminated with Hg only when very close to **S** (<6000 m). At higher distance, the contamination persists up to about 12000 m but only in Al-rich fine sediments and decreases rapidly at lower Al levels.

3.2.2. MeHg distribution in surface sediments from contaminated lagoon

The obtained robust minimal adequate models for MeHg are summarized in Table 3. From the developed models, seven were selected for further evaluation:

$$\ln(MeHg) = a_0 + a_{11}\ln(IHg) \quad (23)$$

$$\ln(MeHg) = a_0 + a_{11}\ln(IHg) - a_2 d_M \quad (24)$$

$$\ln(MeHg) = a_0 + a_{11}\ln(IHg) + a_9\ln(Ca) \quad (25)$$

$$\ln(MeHg) = a_0 + a_{11}\ln(IHg) - a_2 d_M + a_9\ln(Ca) \quad (26)$$

$$\ln(MeHg) = a_0 + a_{11}\ln(IHg) + a_4\sqrt{FF} + a_9\ln(Ca) - a_{1,4} d_S\sqrt{FF} \quad (27)$$

$$\ln(MeHg) = a_0 + a_{11}\ln(IHg) + a_9\ln(Ca) - a_{5,10}[Mn]\ln(OM) \quad (28)$$

$$\ln(MeHg) = a_0 + a_{11}\ln(IHg) + I_{GEO}(Ca, Mn, TC, FF) \quad (29)$$

$$I_{GEO} = -a_7[TC] + a_9\ln(Ca) + a_{10}[Mn] + \sqrt{FF}\left(a_4 + a_{4,7}[TC] - a_{4,10}[Mn]\right) \quad (29a)$$

The results from Eqs 23-29 were compared with the experimental values for MeHg concentrations in sediments for all, short and extrapolated data (Fig. 5). The most important explanatory variable was found to be IHg concentration. However, when only IHg is present in the equation (Eq. 23), the model tends to produce higher results at concentrations of MeHg of less than 0.001 μmol kg^{-1} (Fig. 5a). Close to the mouth of the lagoon, other processes might have occurred, increasing the concentrations of MeHg (Eq 24, 26). High %MeHg has been found in sample 11 in summer and in SJA, BAR, RIO both in summer and winter (1.7±0.3%, n=7). All of these samples are situated close to the mouth (Fig. 1b). A possible reason for the apparent effect of Ca on the concentration of MeHg (Eqs. 25-29) could be processes taking place close to the mouth of the lagoon, d_M and Ca being anticorrelated ($r=-0.40$, $p<0.05$). As confirmation, it can be observed that, when, instead of d_M, in the starting equation d_R is included, similar dependences on d_R were not observed (Table SI-4). Similarly, in a well-mixed estuary, higher %MeHg was measured at high salinity [6]. Most probably, the presence of sandy sediments, the oxidation of sulphides by ocean water and/or the input of labile planktonic organic matter from the coastal zones increases the IHg availability / methylation potentials of the sediments or decreases demethylation near the mouth of the lagoon [23, 31, 33, 38, 40-42, 45, 46].

A less probable explanation of the higher %MeHg near lagoon mouth is external supply of particles rich in MeHg via effluents from wastewater treatment, aquatic food industry and other sources [38, 40, 43]. The plume from a wastewater treatment plant, moving from submarine outfall to the mouth of Aveiro Lagoon, has an influence on the lagoon processes [47]. In industrial or urban wastewater, MeHg is associated with the smallest particles and its removal during treatment is more difficult than that of IHg [48]. Although the sedimentation is hindered near the mouth, the contribution from such MeHg-rich particles to increase %MeHg should be stronger for sediments having smaller IHg concentrations and cannot be completely discarded.

From all minimal adequate models, Eq. 29 has the lowest values of RMSD (Table 3), producing reliable model results for all concentrations of MeHg (Fig. 5g). The ratio DR_M, present in the starting model (Eq. 15), was not found to be important in Eq. 29, only IHg, as a precursor of MeHg, and geochemical variables being significant. The marine influence (Ca) on MeHg concentrations is statistically separated from the effects of other geochemical variables (Mn, TC, FF). However, there are interaction effects between the last three variables (Eq. 29a), possibly due to the influence of other conditions not accounted for in the models.

There was no correlation between I_{GEO} (Eq. 29a) and the concentrations of IHg and MeHg. However, the relationship between %MeHg and I_{GEO} was highly significant ($r=0.64$, $p<0.001$, all data). The effects of I_{GEO} on %MeHg were not observed at high concentrations of IHg (Fig. 6). Although the highest MeHg concentrations in surface sediments were well estimated from IHg, the geochemical variables were necessary in order to develop a reliable model over the whole concentration range of MeHg. The results from multiple regression analysis demonstrated that, in a contaminated shallow lagoon, MeHg exhibits different behaviour from that of IHg. Similarly, in Lavaca Bay sediments, affected by chlor-alkali effluents, IHg was bound to sulphides and OC while MeHg had different behaviour, being retained on Fe and Mn oxides [49].

4. Conclusions

Different behaviour of IHg and MeHg was observed in surface sediments, based upon multiple regression analysis including geochemical characteristics of the sediments and non-Euclidean distances between sampling points. This data treatment method is valid across the concentration range of IHg and MeHg allowing robust quantitative evaluation with respect to extrapolation. For IHg, there was statistical separation of the dispersion from the Hg source and of Al concentration. The inclusion of geochemical variables in the model results in equations not dependent on the direction of the contaminants´ dispersion, even in a geometrically complex environment. Usually, the MeHg concentrations follow those of IHg but geochemical characteristics are necessary to describe the behaviour of MeHg across the whole concentration range. The concentrations of both Hg species are more dependent on TC and OC than on OM, determined by loss on ignition. The models for MeHg demonstrate that, close to the mouth of the lagoon, the sediment has higher potential for methylation, lower MeHg degradation capacity or is exposed to external sources of MeHg. In future, multiple regression analysis could be similarly applied in order to quantitatively evaluate the effects of other contaminants in sediments.

Acknowledgement

T. Stoichev is grateful to FCT for fellowship SFRH/BPD/88675/2012, co-financed by Programa Operacional Potencial Humano (POPH) / Fundo Social Europeu (FSE). The financial support of European SUDOE Interreg IVB Programme through the Orque-Sudoe project and of FONDECYT project (1150855) is acknowledged. J.P.Coelho was supported by the Integrated Program of SR&TD ‘Smart Valorization of Endogenous Marine Biological Resources Under a Changing Climate’ (reference Centro-01-0145-FEDER-000018), co-funded by Centro 2020 program, Portugal 2020, European Union, through the European Regional Development Fund.

References

[1] United Nations Environment Program, Global Mercury Assessment 2013: Sources, Emissions, Releases and Environmental Transport, UNEP Chemicals Branch, Geneva, 2013.

[2] L. Carrasco, S. Díez, D.X. Soto, J. Catalan, J.M. Bayona, Assessment of mercury and methylmercury pollution with zebra mussel (Dreissena polymorpha) in the Ebro River (NE Spain) impacted by industrial hazardous dumps, Sci. Total Environ. 407 (2008) 178-184.

[3] A.G. Bravo, C. Cosio, D. Amouroux, J. Zopfi, P.-A. Chevalley, J.E. Spangenberg, V.-G. Ungureanu, J. Dominik, Extremely elevated methyl mercury levels in water, sediment and organisms in a Romanian reservoir affected by release of mercury from a chlor-alkali plant, Water Res. 49 (2013) 391-405.

[4] S. Guédron, L. Huguet, D.A.L. Vignati, B. Liu, F. Gimbert, B.J.D. Ferrari, R. Zonta, J. Dominik, Tidal cycling of mercury and methylmercury between sediments and water column in the Venice Lagoon (Italy), Mar. Chem. 130-131 (2012) 1-11.

[5] J. Yáñez, M. Guajardo, C. Miranda, C. Soto, H.D. Mansilla, A.R. Flegal, New assessment of organic mercury formation in highly polluted sediments in the Lenga estuary Chile, Mar. Pollut. Bull. 73 (2013) 16-23.

[6] D.L. Taylor, J.C. Linehan, D.W. Murray, W.L. Prell, Indicators of sediment and biotic mercury contamination in a southern New England estuary, Mar. Pollut. Bull. 64 (2012) 807-819.

[7] J.P. Coelho, M.E. Pereira, A. Duarte, M.A. Pardal, Macroalgae response to a mercury contamination gradient in a temperate coastal lagoon (Ria de Aveiro, Portugal), Est. Coast. Shelf Sci. 65 (2005) 492-500.

[8] N.S. Bloom, L.M. Moretto, P. Ugo, A comparison of the speciation and fate of mercury in two contaminated coastal marine ecosystems: The Venice Lagoon (Italy) and Lavaca Bay (Texas), Limnol. Oceanogr., 49 (2004), 367-375.

[9] M.E. Pereira, A.I. Lillebø, P. Pato, M. Válega, J.P. Coelho, C.B. Lopes, S. Rodrigues, A. Cachada, M. Otero, M.A. Pardal, A.C. Duarte, Mercury pollution in Ria de Aveiro (Portugal): a review of the system assessment, Environ. Monit. Assess. 155 (2009) 39-49.

[10] E. Ramalhosa, S.R. Segade, E. Pereira, C. Vale, A. Duarte, Mercury cycling between the water column and surface sediments in a contaminated area, Water Res. 40 (2006) 2893-2900.

[11] E. Ramalhosa, S.R. Segade, M.E. Pereira, C. Vale, A.C. Duarte, Monomethylmercury behaviour in sediments collected from a mercury-contaminated lagoon, Int. J. Environ. Anal. Chem. 91 (2011) 49-61.

[12] M. Válega, A.I. Lillebø, M.E. Pereira, W.T. Corns, P.B. Stockwell, A.C. Duarte, M.A. Pardal, Assessment of methylmercury production in a temperate salt marsh (Ria de Aveiro lagoon, Portugal), Mar. Pollut. Bull. 56 (2008) 153-158.

[13] T. Stoichev, E. Tessier, D. Amouroux, C.M. Almeida, M.C.P. Basto, V.M. Vasconcelos, Multiple regression analysis to assess the role of plankton on the distribution and speciation of mercury in water of a contaminated lagoon, J. Haz. Mater. 318 (2016) 711-722.

[14] C.M. Magalhães, A.A. Bordalo, W.J. Wiebe, Temporal and spatial patterns of intertidal sediment-water nutrient and oxygen fluxes in the Douro River estuary Portugal, Mar. Ecol. Prog. Ser. 233 (2002) 55-71.

[15] A. Turner, Trace metal contamination in sediments from UK estuaries: an empirical evaluation of the role of hydrous iron and manganese oxides, Est. Coast. Shelf Sci. 50 (2000) 355-371.

[16] J.M. Dias, J.F. Lopes, I. Dekeyser, A numerical system to study the transport properties in the Ria de Aveiro lagoon, Ocean Dynam. 53 (2003) 220-231.

[17] J.M. Dias, I. Abrantes, F. Rocha, Suspended Particulate Matter Sources and Residence Time in a Mesotidal Lagoon, J. Coastal Res. SI50 (2007) 1034-1039.

[18] A.T. Reis, S.M. Rodrigues, C. Araújo, J.P. Coelho, E. Pereira, A.C. Duarte, Mercury contamination in the vicinity of a chlor-alkali plant and potential risks to local population, Sci. Total Environ. 407 (2009) 2689-2700.

[19] T. Fidélis, T. Carvalho, Estuary planning and management: the case of Vouga Estuary (Ria de Aveiro) Portugal, J. Environ. Plann. Manag. 58 (2015) 1173-1195.

[20] J.F. da Silva, R.W. Duck, T.S. Hopkins, M. Rodrigues, Evaluation of the nutrient inputs to a coastal lagoon: the case of the Ria de Aveiro Portugal, Hydrobiologia 475/476 (2002) 379–385.

[21] C.T. Costley, K.F. Mossop, J.R. Dean, L.M. Garden, J. Marshall, J. Carroll, Determination of mercury in environmental and biological samples using pyrolysis atomic absorption spectrometry with gold amalgamation, Anal. Chim. Acta 405 (2000) 179-183.

[22] M. Monperrus, P.R. Gonzales, D. Amouroux, J.I.G. Alonso, O.F.X. Donard, Evaluating the potential and limitations of double-spiking species-specific isotope dilution analysis for accurate quantification of mercury species in different environmental matrices, Anal. Bioanal. Chem. 390 (2008) 655-666.

[23] S. Bouchet, D. Amouroux, P. Rodriguez-Gonzalez, E. Tessier, M. Monperrus, G. Thouzeau, J. Clavier, E. Amice, J. Deborde, S. Bujan, J. Grall, P. Anschutz, MMHg production and export from intertidal sediments to the water column of a tidal lagoon (Arcachon Bay, France), Biogeochemistry 114 (2013) 341-358.

[24] R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, http://www.R-project.org/

[25] L. Perez, F. Garcia-Rodriguez, T.J.J. Hanebuth, Variability in terrigenous sediment supply offshore of the Río de la Plata (Uruguay) recording the continental climatic history over the past 1200 years, Clim. Past 12 (2016) 623–634.

[26] M.V.A. Martins, M.A. Mane, F. Frontalini, J.F. Santos, F.S. da Silva, D. Terroso, P. Miranda, R. Figueira, L.L.M. Laut, C. Bernardes, J.G.M. Filho, R. Coccioni, J.M.A. Dias, F. Rocha, Early diagenesis and clay mineral adsorption as driving factors of metal pollution in sediments: the case of Aveiro Lagoon (Portugal), Environ. Sci. Pollut. Res. 22 (2015) 10019–10033.

[27] B. Muresan, D. Cossa, D. Jéséquel, F. Prévot, S. Kerbellec, The biogeochemistry of mercury at the sediment – water interface in the Thau lagoon. 1. Partition and speciation, Est. Coast Shelf Sci. 72 (2007) 472-484.

[28] L. Bellanger, R. Tomassone, Exploration de données et méthodes statistiques: data analysis & data mining avec le logiciel R, Ellipses, Paris, 2014.

[29] M.J. Crawley, The R book, Wiley, Chichester, 2007.

[30] J. Canário, V. Branco, C. Vale, Seasonal variation of monomethylmercury concentrations in surface sediments of the Tagus Estuary (Portugal), Environ. Pollut. 148 (2007) 380-383.

[31] C. Trombini, D. Fabbri, M. Lombardo, I. Vassura, E. Zavoli, M. Horvat, Mercury and methylmercury contamination in surficial sediments and clams of a coastal lagoon (Pialassa Baiona, Ravenna, Italy), Cont. Shelf Res. 23 (2003) 1821-1831.

[32] N.S. Bloom, L.M. Moretto, P. Scopece, P. Ugo, Seasonal cycling of mercury and monomethyl mercury in the Venice Lagoon (Italy), Mar. Chem. 91 (2004) 85-99.

[33] W. Baeyens, C. Meuleman, B. Muhaya, M. Leermakers, Behaviour and speciation of mercury in the Scheldt estuary (water, sediments and benthic organisms), Hydrobiologia 366 (1998) 63-79.

[34] J. Canário, R. Prego, C. Vale, V. Branco, Distribution of Mercury and Monomethylmercury in Sediments of Vigo Ria NW Iberian Peninsula, Water Air Soil Pollut. 182 (2007) 21-29.

[35] M. Coquery, Biogeochemistry of mercury in estuarine and coastal environments. Ph.D. thesis (1994) Univ. Paris 06, France.

[36] S. Covelli, J. Faganeli, M. Horvat, A. Brambati, Mercury contamination of coastal sediments as the result of long-term cinnabar mining activity (Gulf of Trieste, northern Adriatic sea), Appl. Geochem. 16 (2001) 541-558.

[37] P.J. Craig, P.A. Moreton, Total mercury, methylmercury and sulphide levels in British estuarine sediments III, Water Res. 20 (1986) 1111-1118.

[38] J. Faganeli, M.E. Hines, S. Covelli, A. Emili, M. Giani, Mercury in lagoons: an overview of the importance of the link between geochemistry and biology, Est. Coast. Shelf Sci. 113 (2012) 126-132.

[39] Z. Kwokal, S. Franĉiŝković-Bilinski, H. Bilinski, M. Branica, A comparison of anthropogenic mercury pollution in Kaŝtela Bay (Croatia) with pristine estuaries in Öre (Sweden) and Krka (Croatia), Mar. Pollut. Bull. 44 (2002) 1152-1169.

[40] T. Stoichev, D. Amouroux, J.C. Wasserman, D. Point, A. De Diego, G. Bareille, O.F.X. Donard, Dynamics of mercury species in surface sediments of a macrotidal estuarineecoastal system (Adour River, Bay of Biscay), Estuar. Coast. Shelf Sci. 59 (2004) 511-521.

[41] D. Cossa, C. Garnier, R. Buscail, F. Elbaz-Poulichet, N. Mikac, N. Patel-Sorrentino, E. Tessier, S. Rigaud, V. Lenoble, C. Gobeil, A Michaelis–Menten type equation for describing methylmercury dependence on inorganic mercury in aquatic sediments, Biogeochemistry 119 (2014) 35-43.

[42] H. Haris, A.Z. Aris, M. bin Mokhtar, Mercury and methylmercury distribution in the intertidal surface sediment of a heavily anthrophogenically impacted saltwater – mangrove - sediment interplay zone, Chemosphere 166 (2017) 323-333.

[43] C.H. Conaway, S. Squire, R.P. Mason, A.R. Flegal, Mercury speciation in the San Francisco Bay estuary, Mar. Chem. 80 (2003) 199-225.

[44] A.R. Mostafa, A.O. Barakat, Y. Qian, T.R. Wade, D. Yuan, An overview of metal pollution in the western harbour of Alexandria Egypt, Soil Sedim. Contam. 13 (2004) 299-311.

[45] C.R. Hammerschmidt, W.F. Fitzgerald, P.H. Balcom, P.T. Visscher, Organic matter and sulfide inhibit methylmercury production in sediments of New York/New Jersey Harbor, Mar. Chem. 109 (2008) 165-182.

[46] C.S. Langer, W.F. Fitzgerald, P.T. Visscher, G.M. Vandal, Biogeochemical cycling of methylmercury at Barn Island Salt Marsh Stonington CT USA, Wetlands Ecol. Manage. 9 (2001) 295-310.

[47] J.F. da Silva, R.W. Duck, T.S. Hopkins, J.M. Anderson, Nearshore circulation revealed by wastewater discharge from a submarine outfall Aveiro Coast Portugal, Hydrol. Earth Syst. Sci. 6 (2002) 983-989.

[48] T. Stoichev, E. Tessier, H. Garraud, D. Amouroux, O.F.X. Donard, D. Tsalev, Mercury speciation and partitioning along municipal sewage treatment plant, J. Balk. Ecol. 12 (2009) 135-145.

[49] N.S. Bloom, G.A. Gill, S. Cappellino, C. Dobbs, L. McShea, C. Driscoll, R. Mason, J. Rudd, Speciation and cycling of mercury in Lavaca Bay Texas sediments, Environ. Sci. Technol. 33 (1999) 7-13.

Figure captions

Fig. 1. Map of **(a)** Localization of the Aveiro Lagoon (Atlantic coast, Portugal) **(b)** Aveiro Lagoon with the sampling points. Each point is characterized by distances (d_S, d_M, d_R) within the simplified border (hyphen line) to Hg source (**S**), to the lagoon mouth (**M**) and to a reference point **R**, respectively **(c)** Laranjo Bay with the sampling points

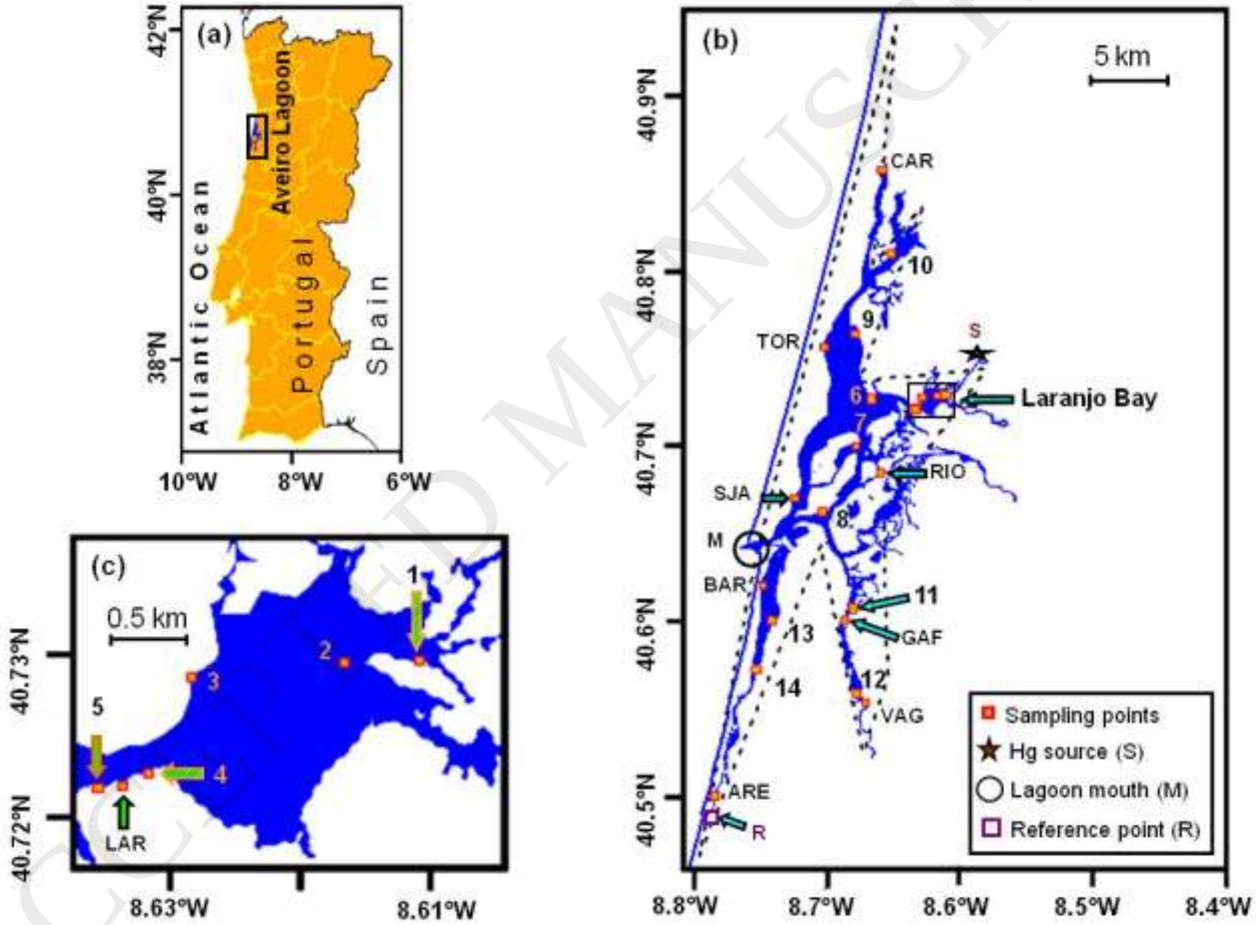


Fig. 1

Fig. 2. Winter-summer variations in surface sediments (Aveiro Lagoon) for concentrations of **(a)** IHg; **(b)** MeHg

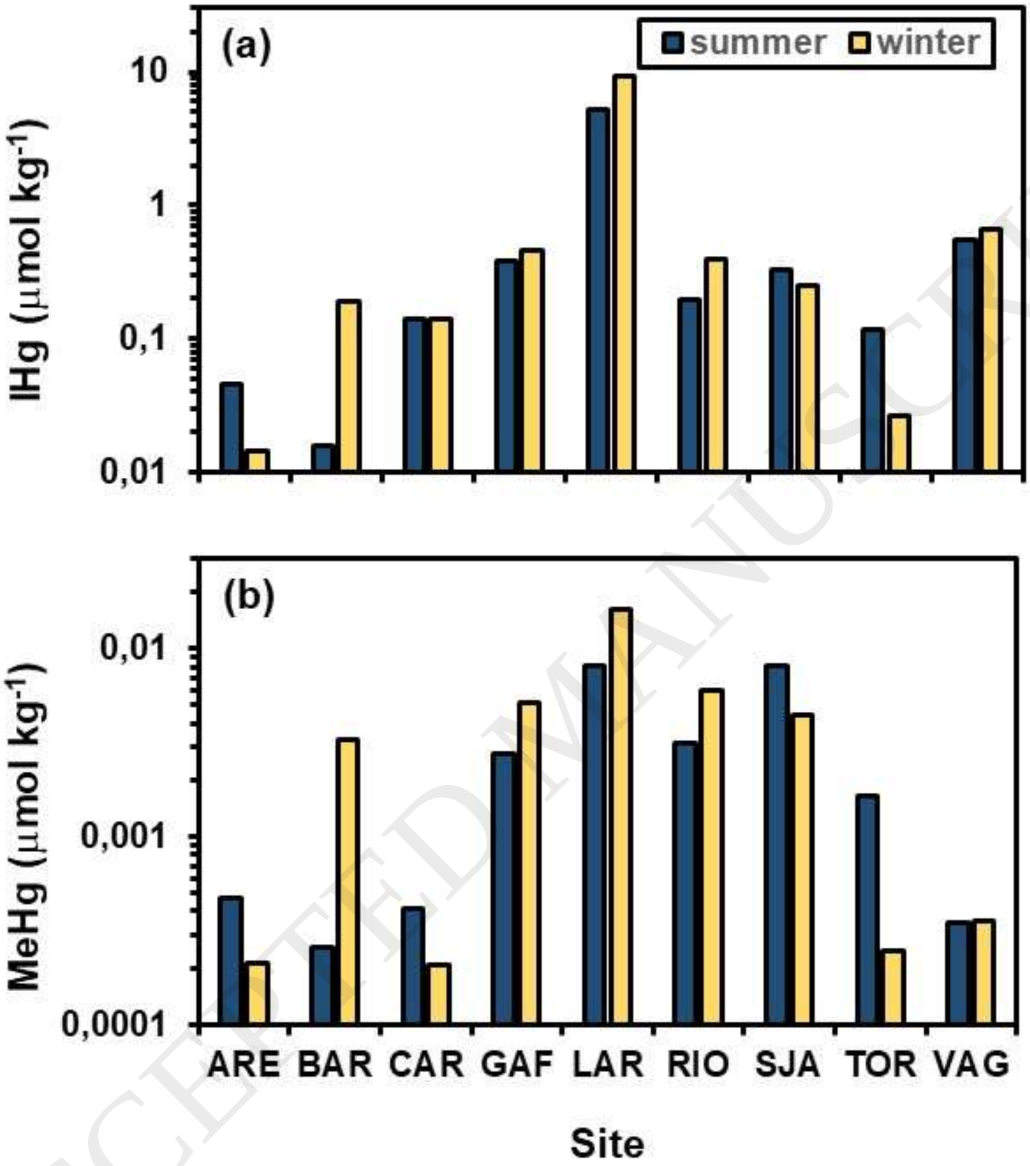


Fig. 2.

Fig. 3. Dependence of model values on experimental ones for IHg concentrations in surface sediments (Aveiro Lagoon) for all data, short (IHg concentrations less than 95th percentile) and extrapolated data. Significant explanatory variables: **(a)** DR_M, Al

(Eq. 17); **(b)** DR_M, Al, FF (Eq. 18); **(c)** DR_M, TC, FF, interaction TC:FF (Eq. 19); **(d)** d_S, Al (Eq. 22). Abbreviations: $DR_M = d_S/d_M$ where d_S is distance to Hg source, d_M is distance to the lagoon mouth, FF: fine fraction, TC: total carbon

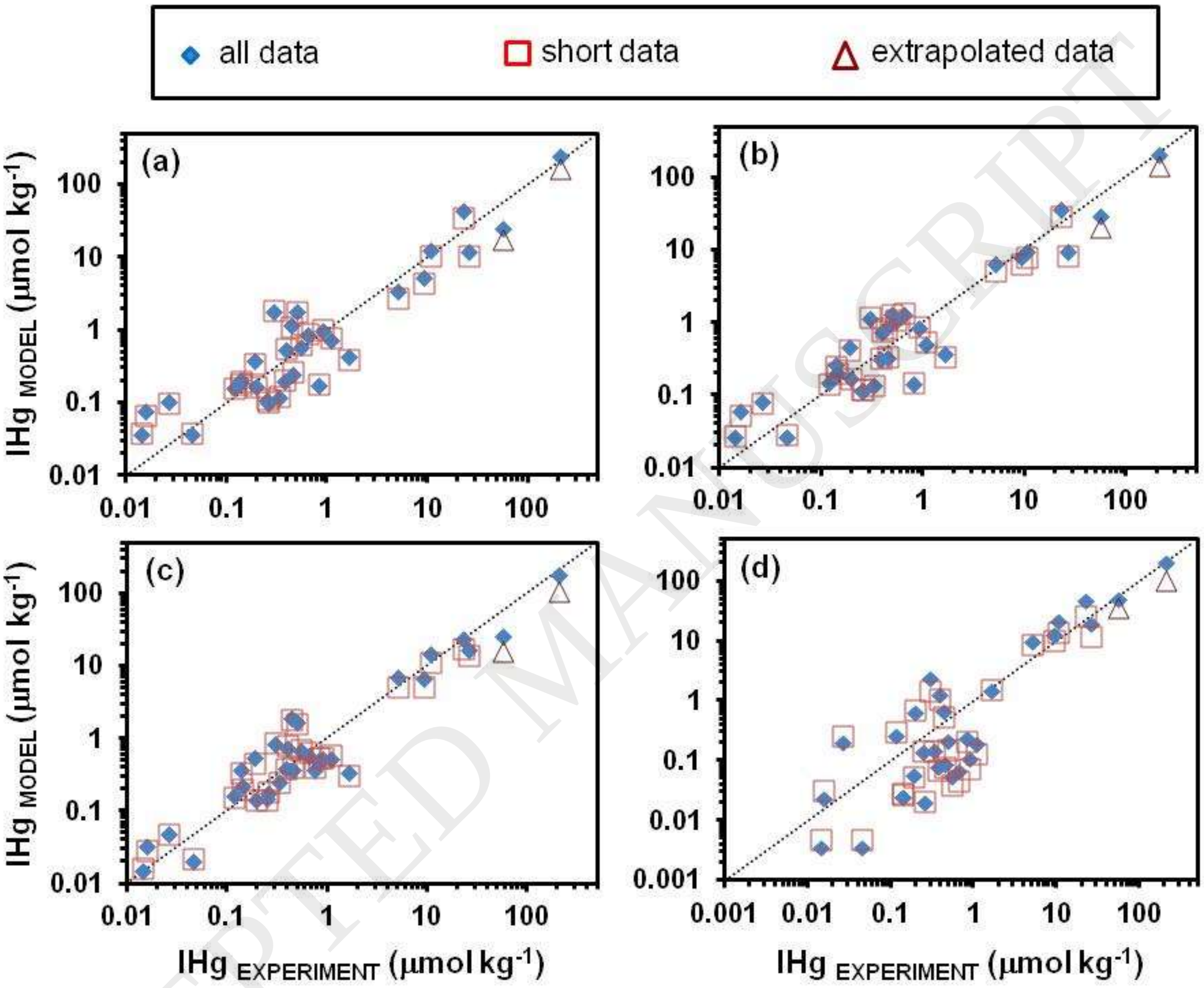


Fig. 3.

Fig. 4. Experimental IHg concentration in surface sediments (Aveiro Lagoon) as a function of distance to Hg source (d_S) and Al concentration

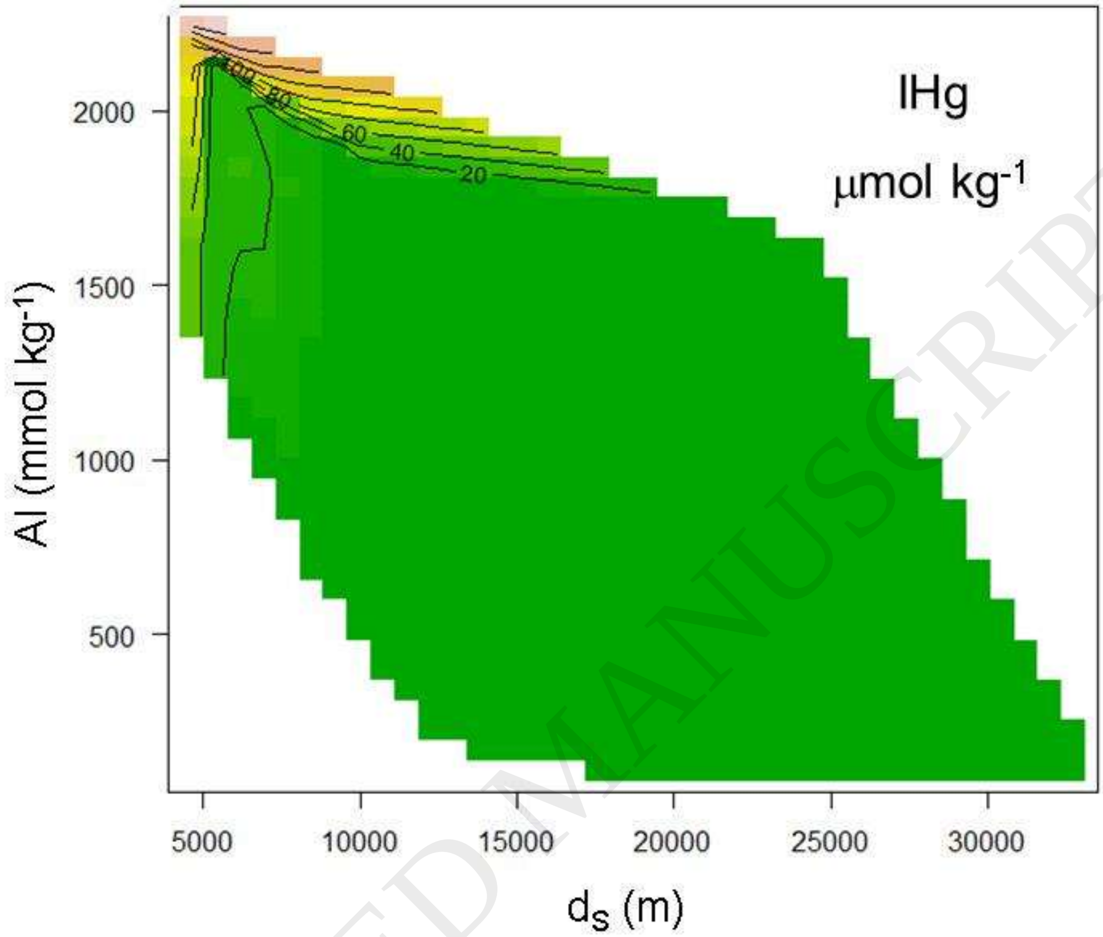


Fig. 4.

Fig. 5. Dependence of model values on experimental ones for MeHg concentrations in surface sediments (Aveiro Lagoon) for all data, short (MeHg concentrations less than 95th percentile) and extrapolated data. Significant explanatory variables: **(a)** IHg (Eq. 23); **(b)** IHg, d_M (Eq. 24); **(c)** IHg, Ca (Eq. 25); **(d)** IHg, Ca, d_M (Eq. 26); **(e)** IHg, Ca, FF, interaction FF:d_S (Eq. 27); **(f)** IHg, Ca, interaction OM:Mn (Eq. 28); **(g)** IHg, Ca, Mn, FF, TC, interactions Mn:FF and TC:FF (Eq. 29). Abbreviations: d_M: distance

to the lagoon mouth, d_S: distance to Hg source; FF: fine fraction, OM: Organic matter (loss on ignition), TC: total carbon

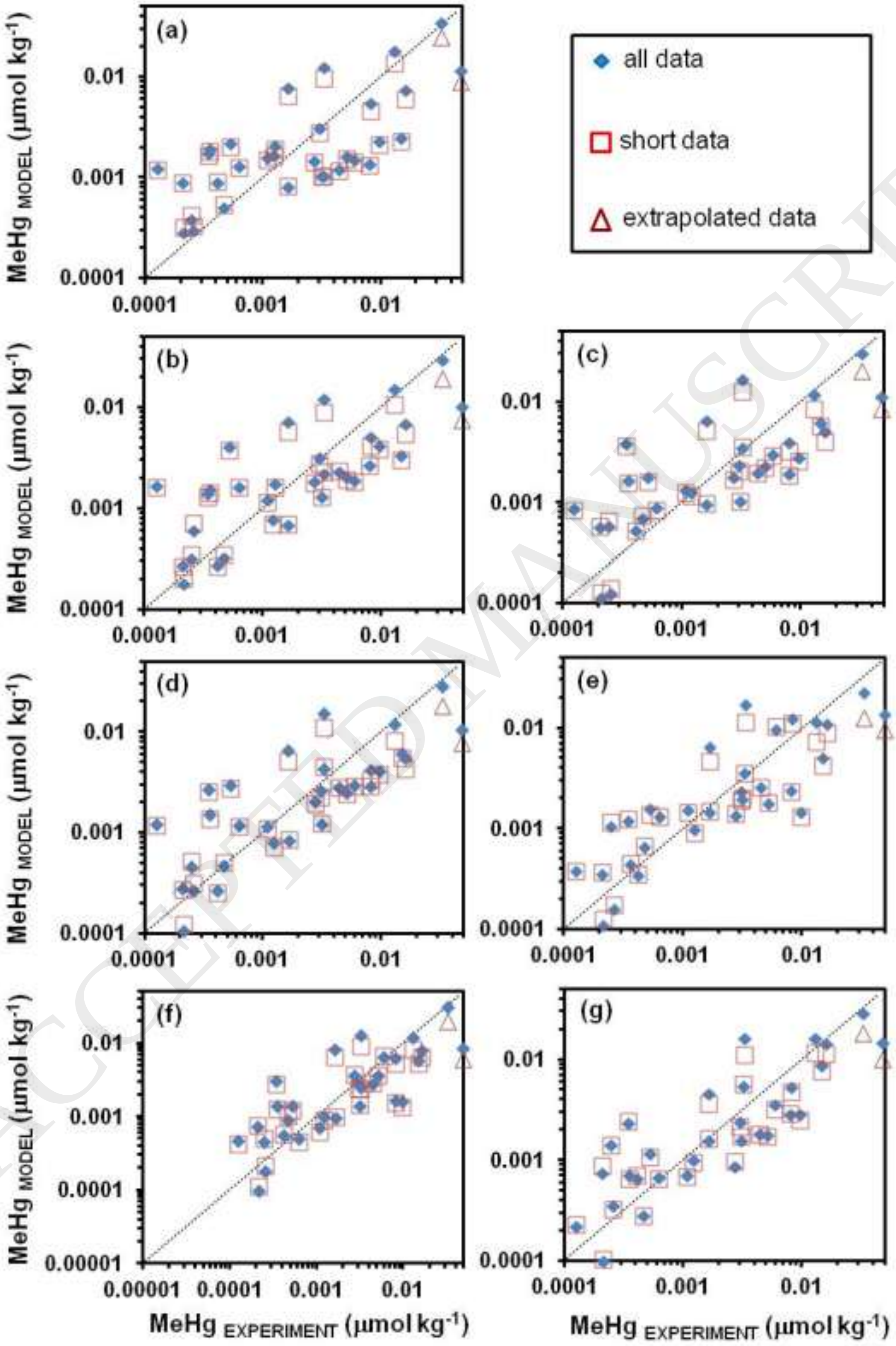


Fig. 5.

Fig. 6. Percentage of MeHg with respect to Hg_{TOT} in surface sediments (Aveiro Lagoon) as a function of IHg concentration and geochemical function I_{GEO} (Eq. 29a).

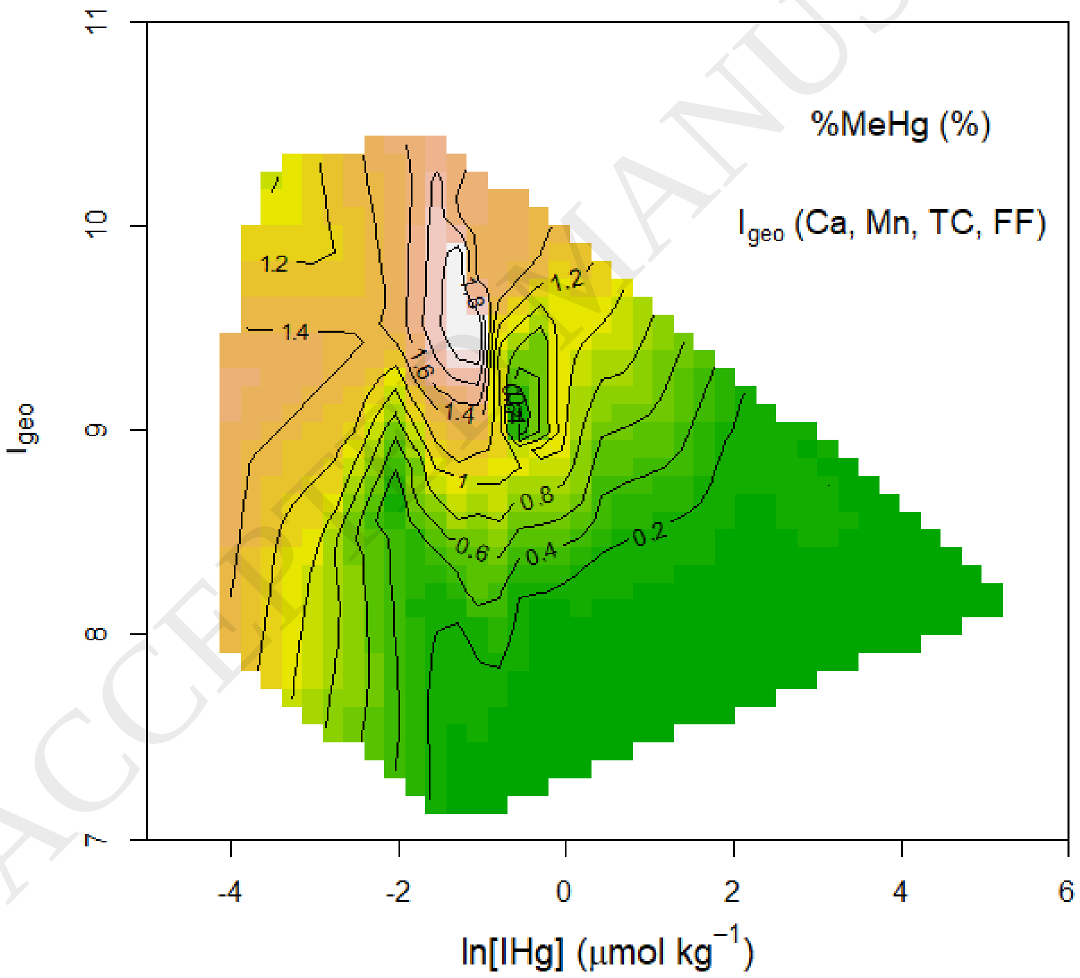


Fig. 6.

Table 1. Minimal adequate models for IHg concentrations in surface sediments from the Aveiro Lagoon as function of distance to source d_S, distance to mouth d_M and geochemistry. Distance ratio $DR_M=d_S/d_M$ was used instead of distances in starting model Eqs 5-7. Root mean square deviation (RMSD) for all data, short (IHg concentrations less than 95th percentile) and extrapolated data also given.

Model Eqs.[(a)]	**Minimal adequate model**	**RMSD** [(b)]		
		all	**short**	**extrapolated**
4	$\ln(IHg) = -a_1 d_S + a_{1,1} {d_S}^2 + a_{2,2} {d_M}^2 - a_{1,2} d_S d_M + a_4\sqrt{FF} + a_8[Al]$	14.7	3.4	95.4
4	$\ln(IHg) = -a_1 d_S + a_{1,1} {d_S}^2 + a_{2,2} {d_M}^2 - a_{1,2} d_S d_M + a_4\sqrt{FF} + a_6\sqrt{OC}$	22.5	2.3	121.1
5	$\ln(IHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_8[Al]$	**8.6**	**3.9**	**43.9**
5	$\ln(IHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_6\sqrt{OC}$	15.3	2.7	107.3
5	$\ln(IHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_7[TC] + a_{3,7}[TC]\ln(DR_M)$	15.1	3.0	96.5
5	$\ln(IHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_{10}[Mn]$	16.5	3.7	103.5
6	$\ln(IHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_4\sqrt{FF} + a_8[Al]$	**6.4**	**3.7**	**54.7**
6, 7	$\ln(IHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_4\sqrt{FF} + a_6\sqrt{OC}$	16.1	3.0	107.0
6	$\ln(IHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_4\sqrt{FF} + a_7[TC]$	8.5	2.4	83.3
7	$\ln(IHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_4\sqrt{FF} + a_7[TC] + a_{4,7}[TC]\sqrt{FF}$	**8.2**	**2.7**	**78.9**

(a) Starting equations described in part 2.4; FF: fine fraction, OC: organic carbon, TC: total carbon

(b) Models with values of RMSD in bold are considered further

Table 2. Minimal adequate models for IHg concentration in surface sediments from the Aveiro Lagoon as function of referent distance ratio $DR_R = d_S/d_R$ and geochemistry. Root mean square deviation (RMSD) for all data, short (IHg concentration less than 95th percentile) and extrapolated data also given. The explanation variable DR_R was transformed according to the equation $DR_{R,T} = (DR_R)^{-0.5}$.

Model Eqs.[a]	**Minimal adequate model**	**RMSD** [b]		
		all	**short**	**extrapolated**
5, 6	$\ln(IHg) = a_0 + a_{3,3}/DR_R + a_8[Al]$	**7.9**	**3.6**	**33.2**
5	$\ln(IHg) = a_0 + a_{3,3}/DR_R + a_{10}[Mn]$	8.8	3.5	58.5
5	$\ln(IHg) = a_0 - a_3/\sqrt{DR_R} + a_{3,3}/DR_R + a_6\sqrt{[OC]} - a_{3,6}\sqrt{[OC]/DR_R}$	6.6	3.3	129.4
5	$\ln(IHg) = a_0 - a_3/\sqrt{DR_R} + a_{3,3}/DR_R + a_7[TC] - a_{3,7}[TC]/\sqrt{DR_R}$	5.6	3.2	30.8
6	$\ln(IHg) = a_0 - a_3/\sqrt{DR_R} + a_{3,3}/DR_R + a_4\sqrt{FF} + a_6\sqrt{[OC]}$	**7.4**	**2.5**	**20.2**
6	$\ln(IHg) = a_0 - a_3/\sqrt{DR_R} + a_{3,3}/DR_R + a_4\sqrt{FF} + a_7[TC]$	14.8	2.2	35.5
7	$\ln(IHg) = a_0 - a_3/\sqrt{DR_R} + a_{3,3}/DR_R + a_4\sqrt{FF} + a_6\sqrt{[OC]} - a_{4,6}\sqrt{FF[OC]}$	11.7	2.8	73.5
7	$\ln(IHg) = a_0 - a_3/\sqrt{DR_R} + a_{3,3}/DR_R + a_4\sqrt{FF} + a_7[TC] - a_{4,7}[TC]\sqrt{FF}$	20.6	2.4	112.3

(a)Starting equations described in part 2.4; d_S: distance to Hg source:, d_R: distance to reference point, OC: organic carbon, TC: total carbon, FF: fine fraction;

(b)Models with values of RMSD in bold are considered further

Table 3. Minimal adequate models for MeHg concentrations in surface sediments from the Aveiro Lagoon as function of distance to source d_S, distance to mouth d_M and geochemistry. Distance ratio $DR_M=d_S/d_M$ was used instead of distances in starting model Eqs 12-15. Root mean square deviation (RMSD) for all data, short (MeHg concentrations less than 95th percentile) and extrapolated data also given.

Model Eqs.[(a)]	**Minimal adequate models**	**RMSD ($\times10^{-3}$)** [(b)]		
		all	**short**	**extrapolated**
8	$\ln(MeHg) = a_0 - a_1 d_S + a_9 \ln(Ca)$	9.7	4.5	37
8, 10	$\ln(MeHg) = a_0 - a_2 d_M + a_{11} \ln(IHg)$	**7.8**	**3.8**	**30**
9, 12, 14	$\ln(MeHg) = a_0 + a_9 \ln(Ca) + a_{11} \ln(IHg)$	**7.9**	**4.2**	**30**
9	$\ln(MeHg) = a_0 - a_2 d_M + a_9 \ln(Ca) + a_{11} \ln(IHg)$	**7.9**	**4.0**	**31**
8	$\ln(MeHg) = a_0 + a_4\sqrt{FF} - a_{1,4} d_S\sqrt{FF} - a_{2,4} d_M\sqrt{FF}$	10.0	4.1	38
11	$\ln(MeHg) = a_0 + a_4\sqrt{FF} + a_9 \ln(Ca) + a_{11} \ln(IHg) - a_{1,4} d_S\sqrt{FF}$	**7.7**	**3.8**	**31**
13	$\ln(MeHg) = a_0 - a_3 \ln(DR_M) + a_{3,3}[\ln(DR_M)]^2 + a_9 \ln(Ca)$	8.2	4.7	31
13, 14	$\ln(MeHg) = a_0 + a_{11} \ln(IHg)$	**7.8**	**4.1**	**28**
15	$\ln(MeHg) = a_0 + a_9 \ln(Ca) + a_{11} \ln(IHg) - a_{5,10}[Mn] \ln(OM)$	**8.0**	**3.7**	**31**
15	$\ln(MeHg) = a_0 - a_7[TC] + a_9 \ln(Ca) + a_{11} \ln(IHg) + a_{10}[Mn] + \sqrt{FF}\left(a_4 + a_{4,7}[TC] - a_{4,10}[Mn]\right)$	**7.0**	**3.1**	**29**

[(a)] Starting equations described in part 2.4; FF: fine fraction, OM: organic matter (loss on ignition), TC: total carbon

[(b)] Models with values of RMSD in bold are considered further