Deriving probabilistic soil distribution coefficients (K_d). Part 3: Reducing variability of americium K_d best estimates using soil properties and chemical and geological material analogues

Oriol Ramírez-Guinart[a], Daniel Kaplan[b], Anna Rigol*[,a], and Miquel Vidal[a]

[a] *Chemical Engineering and Analytical Chemistry Department, Faculty of Chemistry, University of Barcelona, Martí i Franqués 1-11, 08028, Barcelona, Spain.*

[b] *Savannah River National Laboratory, Aiken, South Carolina, USA*

*Corresponding author: Telephone: (+34) 93 403 9274; Fax: (+34) 93 402 1233

Email address: annarigol@ub.edu (A. Rigol)

Acknowledgements

This work was carried out under the auspices of the IAEA MODARIA and MODARIA II programmes. It was supported by the Ministerio de Ciencia e Innovación de España (CTM2014-55191 and CTM2017-87107-R), the Generalitat de Catalunya (AGAUR 2014SGR1277), and the Savannah River National Laboratory's LDRD program (LDRD-2017-00005). Oriol Ramírez-Guinart would like to thank the support of an APIF pre-doctoral fellowship from the University of Barcelona. Authors would also like to thank Dr. Brenda Howard (Center for Ecology and Hydrology, U.K.) for their fruitful discussions and Oriol Toll for his contribution in data treatment.

Deriving probabilistic soil distribution coefficients (K_d). Part 3: Reducing variability of americium K_d best estimates using soil properties and chemical and geological material analogues

Abstract

The solid-liquid distribution coefficient (K_d) is a key input parameter in radioecological risk models. However, its large variability hampers its usefulness in modelling transport processes as well as its accuracy in representing soil-radionuclide interactions. To assist in the selection of K_d values and their cumulative distribution functions for study areas without site specific information, a critically reviewed dataset was developed, containing more than 5000 soil K_d entries for 83 elements and an additional 2000 entries of K_d data for 75 elements gathered from a selection of other geological materials. For the specific case of americium (Am), the dataset contained 109 entries for soils and 33 additional entries for sediment and subsoils. The analysis of the Am K_d soil dataset showed that values varied 4-orders of magnitude, and consequently the resulting Am K_d best estimate (geometric mean (GM): 4.6 x 10^3 L kg^{-1}) lacked sufficient reliability. The objective of this study was to calculate cumulative distribution functions and statistically evaluate this dataset to determine if the Am K_d variability in soils could be reduced by considering various factors, including: 1) measurement methods, 2) key soil properties, 3) the use of chemical analogue data, and 4) the use of analogue data. Accounting for Am K_d experimental method (*i.e.*, sorption vs. desorption; long- vs. short-term experiments) had little effect on reducing variability. However, accounting for key soil factors (*i.e.,* organic matter content (OM), pH, soil texture) succeeded in reducing variability of Am K_d, especially when combining pH and OM. While previous data sets have used 20% OM content as a critical value to distinguish between mineral and organic soils, this study shows that this critical value should be reduced to 10% OM to minimize Am K_d variability. The inclusion in the dataset of Am K_d from other geological materials (e.g., gyttjas, tills, and subsoils) and K_d values from trivalent lanthanides (Ln (III)) and actinides (An (III)) (172 additional entries) did not statistically affect the Am K_d geometric means of the various pH and OM partial datasets. The larger composite dataset (>310 entries), with both chemical analogues and geological material analogues to address data gaps, increased the statistical power for calculating K_d best estimates with lower variability, thereby enhancing their usefulness for radionuclide risk calculations.

Keywords: Solid-liquid distribution coefficient (K_d); soil; americium; lanthanides; actinides; probabilistic modeling

1. Introduction

There is a significant amount of qualitative and quantitative data on the interaction of selected radionuclides (radiocaesium, radiostrontium, and several naturally occurring radionuclides, such as uranium) with soils. However, there are evident knowledge and data gaps describing how some transuranic and actinide elements interact with soils in the environment and in nuclear waste disposal systems. Among these, americium (Am) is of special concern since some Am radioisotopes are long-lived alpha emitters. Other actinides and trivalent lanthanides have been reported to have several similar geochemical behaviours as Am due to similar chemical properties (such as ionic radii, coordination number and oxidation state) (GRS, 2012).

Americium interaction with geological materials is affected by pH, organic matter (OM), and mineralogy, especially the presence of metal amorphous coatings, clay minerals and carbonates (Pavlotskaya et al., 2003; EPA, 2004). These phases not only provide sorption sites of different affinities, but also modify Am speciation in solution, which is pH-dependent and involves hydrolysis and complexation reactions with organic and inorganic ligands, such as fulvic/humic acids and carbonate (Lujaniene et al., 2007; Ho Lee et al., 2011). Sorption data on soils is scarce for actinides and lanthanides. Recent studies confirmed that Am sorption to soils is governed by pH, OM, specific surface area, and Am aqueous speciation, which in turn is affected by the presence of hydroxyl, phosphate and carbonate ligands in solution (Choppin, 2007; Ramírez-Guinart et al., 2016; Ramirez-Guinart et al., 2017).

This is the last in a series of three papers (Ramírez-Guinart et al., 2020a; Ramírez-Guinart et al., 2020b), aimed at deriving sorption data suitable for risk assessment from soil K_d datasets, as well as developing a strategy to reduce and describe K_d variability based on probabilistic models, including the construction of distribution functions to statistically describe K_d values. As explained in parts 1 and 2 in this series of three papers, the TRS 472 dataset (IAEA, 2010) was the starting point of the work. Under the auspices of the IAEA-Modelling and Data for Radiological Impact Assessments (MODARIA I and MODARIA II) projects, the TRS-472 dataset was updated and critically reviewed following agreed acceptance criteria (Ramirez-Guinart et al., 2020a).

Besides obtaining Am K_d distributions and best estimates for soils grouped based on their texture and organic matter content (Gil Garcia et al., 2009), in this work we examine additional strategies to reduce uncertainty, including the impact of: 1) experimental measurement method followed to derive the Am K_d values; 2) redefining the OM content threshold for distinguishing mineral and organic soils; 3) including pH as a grouping factor; and 4) including Am and Am-chemical analogue data to enhance partial datasets and to derive probabilistic K_d values that can be applied for both Am and for other actinides and lanthanides.

2. Data collection and treatment.

2.1. Current status of the Am K_d compilation

The update of the TRS-472 dataset was based on a number of criteria such as rejecting any K_d value originated from parametric regression equations, mass-transport experiments, or compilation K_d reference estimates; pooling values originated by varying non-relevant operational or soil variables; excluding values obtained from experiments not representative for environmental conditions (such as extremely low or high pH); accepting data from stable isotopes obtained at the lowest concentration range (Ramírez-Guinart et al., 2020a). K_d values originating from soils, and not pure phases, such as clay minerals or metal (hydro)oxides, were included in the dataset. Our working definition of soil was the unconsolidated geological material comprising the terrestrial root zone. This definition has the benefit of using the same terminology commonly used by the authors of the original research and excluded those unconsolidated geological materials below the soil layer (aquifer sediments and vadose zone) and beneath waterways (streams, lakes, and oceans). As will be discussed below, we also included some other types of geological materials (subsoils, tills, and gyttjas) to test if their K_d values differed significantly from soil K_d values. Importantly, the dataset includes as much ancillary information of the sample as reported, such pH, OM content, clay and sand contents in the mineral fraction, carbonate content and cationic exchange capacity.

The updated soil dataset contained 109 Am K_d entries, around 45 more than in the previous K_d compilation (IAEA, 2010), varying up to 4 orders of magnitude (the range was 2.7×10^1 – 2.8×10^5 L kg^{-1}). In addition to this, 33 entries for Am K_d in subsoils (from 3 m down to 48 m depth) and surface sediments (mostly from estuaries), were available. Finally, another dataset

(Analogue dataset) was created with data of lanthanides and actinides K_d (specifically, La, Sm, Eu, Gd, Er, Lu and Cm) in soils (116 entries), tills (28 entries), subsoils (4 entries) and gyttja (2 entries).

2.2. Soil factors and developed criteria to group Am K_d data

Dissolved organic matter (DOM), pH and specific surface area play a key role in Am interaction in soils (Ramírez-Guinart et al., 2016). Americium sorption increases in soils with a high specific surface area. The specific surface area is often not available from routine soil characterization data. Nevertheless, because the soil specific surface area is related to the presence of the finest soil particles (< 200 µm), the effect of specific surface area could be approximated by soil texture, a parameter that is commonly measured. Besides, Am sorption can be strongly inhibited by the formation of stable and anionic complexes with DOM that remains in solution because of their low affinity (electrostatic repulsions) for soil surfaces. A grouping criterion should distinguish between those soils in which Am sorption is controlled by the DOM, presenting a lower capacity to sorb Am, from those soils in which the Am sorption is controlled by the mineral fraction, presenting a higher capacity to sorb Am. The amount of DOM would then be a suitable soil factor for Am K_d data grouping. However, the scarcity of available DOM data hampered its use for grouping K_d values. Because DOM is derived from soil OM content, K_d data grouping based on the soil OM could serve as a reasonable surrogate parameter. Therefore, the OM K_d grouping criterion was applied as a first approach to separate Am K_d into two partial datasets, as previously defined (IAEA, 2010). In short, Am K_d values were included in the Organic group if the soil had an OM content ≥ 20%, whereas they were included in the Mineral group if OM content was lower than 20%. Secondly, the Am K_d data contained in the Mineral group were split in three textural groups (Sand, Loam, and Clay), thus defining the OM+Texture criterion. Textural groups were operationally defined as follows: the Sand group had a sand fraction > 65% and a clay fraction < 18%; the Clay group had a clay fraction > 35%; and the Loam group were all other mineral soils.

Previous results (Ramirez-Guinart et al., 2016) suggested that the Am-DOC speciation in solution can be dominant even at low concentration of DOC and for OM contents around 10%. Therefore, the OM threshold to distinguish between Mineral and Organic (20%) groups may not be the most suitable for minimizing Am K_d variability. Accordingly, a second analysis of the OM+Texture criterion was conducted to explore whether a lower OM threshold would be more suitable for Am K_d values. This was accomplished by comparing changes in geometric means (GM) and geometric standard deviations (GSD) of the new mineral and organic datasets created when decreasing the OM threshold from 20% to 15% and 10%.

A pH criterion for reducing Am K_d variability was also explored. Like the non-linear uranium K_d vs. pH dependence (Vandenhove et al., 2009; Ramírez-Guinart et al., 2020a), the Am K_d dataset was split into four partial datasets to reflect established trends of pH-dependent Am speciation in solution and pH-dependent affinity for sorption sites (Kaplan et al., 1996; Choppin, 2007; Ramírez-Guinart et al., 2016):

- pH < 6: presence of positively charged sorption sites and cationic Am species (primarily as Am^{3+}) are expected to result in lower Am K_d values due to electrostatic repulsions.
- 6 ≤ pH < 7.5: presence of deprotonated sorption sites, leading to an increase in the sorption of cationic Am species (primarily as $Am(OH)^{2+}$ and $AmCO_3^{+}$).
- 7.5 ≤ pH < 9: increase of sorption sites due to increase in negative charge resulting from progressive deprotonation of functional groups. High Am K_d values are expected, excepting for soil-water systems with high content of dissolved carbonate with predominance of the anionic $Am(CO_3)_2^{-}$ species.
- pH ≥ 9: unless Am precipitation or co-precipitation occurs, much lower Am K_d are generally expected since anionic and neutral Am species (primarily $Am(CO_3)_3^{-3}$ and $Am(OH)_3$) are predominant.

Multiple linear regressions have been recently proposed to estimate the Am K_d values in soils from properties related to the soil factors mentioned above (Ramirez-Guinart et al., 2016). Therefore, grouping criteria can be developed by combining as many of these soil factors as possible by using soil properties that are frequently available. Since the pH-dependency of the Am sorption in soils remains unclear when DOM is present in the soil solution at concentrations high enough to control Am speciation, a combined grouping criterion involving pH and OM was also tested (the pH+OM criterion). Finally, since the Am sorption in soils is influenced by the soil specific surface area, a final attempt was done to improve Am K_d data grouping by further splitting the previous pH+OM partial datasets into sand, loam, and clay textural classes, thus leading to partial datasets containing Am K_d data only from a given soil texture (OM+pH+Texture grouping criterion).

2.3. Analyses of the influence of experimental approach on Am K_d data variability

As in the other two papers of this series (Ramírez-Guinart et al., 2020a; Ramírez-Guinart et al., 2020b), the influence of the experimental approach was simultaneously evaluated along with relevant soil factors.

The majority of Am K_d entries fell within the "short-term sorption" category (ST-S, that is, Am K_d derived from applying a sorption batch test based on putting in contact for short times (< 1 yr) a non-contaminated soil with a solution spiked with americium), and "short-term desorption" category (ST-D, Am K_d derived from applying an extraction batch test to soils recently (< 1 yr) contaminated with americium) (Ramirez-Guinart et al., 2020a). There were no entries that could be considered as "long-term desorption" data (that is, Am K_d derived from applying an extraction test to long-term contaminated solid materials with americium) (Ramirez-Guinart et al., 2020a). Consequently, the effect of sorption dynamics on Am K_d data could not be checked due to a lack of long-term data.

The data treatment was based on group mean centring (GMC) to minimize the effect of soil factors identified as relevant on Am K_d variability (additional details are provided in Ramírez-Guinart et al., 2020a). Firstly, the overall Am dataset was log-transformed, the log Am K_d data was then grouped according to the OM+Texture or pH criteria, the arithmetic mean (AM) of log Am K_d values of each soil-type group created was calculared and each single log Am K_d value within a given group was corrected by subtracting the AM log Am K_d value of the respective soil-type group. Subsequently, the GMC-corrected log Am K_d datasets were divided according to sorption and desorption data (Ramírez-Guinart et al., 2020a). Then, statistical tests (Fisher's least significant differences (FLSD) test for multiple samples; 95% confidence level; StatGraphics 18) were performed to check whether the GMC-corrected log Am K_d data significantly differed between experimental approaches.

2.4. Construction of cumulative distribution functions to describe Am K_d variability

Cumulative Distribution Functions (CDF) of Am K_d data were constructed to describe their population and variability datasets. Since the K_d parameter is a ratio of concentrations, K_d data are expected to follow a lognormal distribution (Sheppard et al., 2011). For the construction of CDFs, Am K_d data were log-transformed and the presence of possible outlier values in the datasets was examined by performing an exploratory analysis based on box-and-whisker plots. The log Am K_d data within every dataset were sorted by increasing value and an empirical frequency ($f_{exp,i}$) equal to 1/N (where N is the total number of Am K_d entries in the respective dataset) was assigned to each entry. Experimental cumulative frequency distributions were constructed by assigning to each sorted log Am K_d value their corresponding cumulative frequency ($F_{exp,i}$), *i.e.*, the sum of the preceding frequencies $\left(F(K_{d,j}) = \sum_{i=0}^{j} f(K_{d,i})\right)$. The Kolmogorov-Smirnov test was applied to ascertain that underlying frequency distribution in each Am K_d dataset did not differ from a lognormal distribution. As expected, it was confirmed that overall and partial Am K_d datasets followed a lognormal distribution. Consequently, the experimental cumulative frequency distributions constructed with the log Am K_d data were fitted to the theoretical normal CDF equation, and the related geometric mean (GM; 50th percentile) and percentile ranges (5th and 95th) were derived. Additional details are provided by Ramírez-Guinart et al. (2020a).

To properly derive a reliable CDF from a given K_d dataset it is necessary that it contains a minimum number of entries. Although it was generally considered 10 entries as this minimum value (Ciffroy et al., 2009), CDFs were also constructed for those partial datasets containing between 7 and 10 entries. For the rest of cases GM values were calculated directly from the dataset.

3. Analyses of Am K_d distributions

3.1. Influence of the experimental approach on Am K_d data

The overall Am K_d dataset contained K_d data gathered by applying sorption experiments in a short-term scenario (ST-S; n =60), and desorption experiments in a short-term scenario (ST-D; n=36). When the statistical analysis was performed without applying the GMC correction, statistical differences were observed between the two datasets (GM ST-S: 2760 L/kg; GM ST-D: 8713 L/kg), which agrees with reported data in which desorption rates for actinides appear to be slower than sorption rates in geological materials (Kaplan et al., 2004; Wong et al., 2015). However, after applying the GMC correction to the partial datasets created from the application of either the OM or the pH criteria, the FLSD test revealed that there were not significant differences between ST-S and ST-D datasets. This finding suggested that the variability on Am K_d values due to the method applied for its quantification and/or the sorption/desorption scenario was negligible with respect to that caused by the contrasting properties of the soils, that is, pH or organic matter content. Consequently, experimental approach factor was not considered in subsequent statistical analyses of the Am K_d dataset.

The lack of long-term data prevented a proper evaluation of the effect of sorption dynamics on the Am K_d values. However, it can be predicted to be of lower significance than the effect of soil properties, as it has been shown that Am and other trivalent lanthanide are quickly and strongly bound to soils, thus it is expected to have a minor effect on Am sorption dynamics (Ramírez-Guinart et al., 2016).

3.2. Am K_d best estimates and CDFs based on the OM+Texture criterion

3.2.1. Am K_d best estimates and CDFs based on the initial OM+Texture criterion

Table 1 summarises the descriptors of the Am K_d distributions obtained by applying the OM+Texture criterion. Those Am K_d entries originated from the same soil sample, but from sorption and desorption tests, were pooled into a single, mean value for this and subsequent analysis included in this work. A few of the derived partial datasets (Organic and Clay datasets) did not have enough Am K_d entries to construct CDFs. The statistical analyses showed that there were not significant differences among Mineral and Organic datasets. Moreover, the GM among the textural classes did not follow the anticipated sequence of GM Sand < GM Loam < GM Clay. The GM Loam was similar to GM Clay, but the former was statistically higher than GM Sand. Besides, neither GM Loam nor GM Sand differed to the GM of the Organic dataset (this latter dataset with a low number of entries). Thus, it seems as if the initial OM+Texture criterion, with the 20% OM threshold, did not capture the actual mechanisms governing Am sorption. Thus, other OM threshold and other soil factors more specifically related to the Am-soil interaction, such as pH, were evaluated to properly and efficiently group Am K_d data.

3.2.2. Am K_d best estimates and CDFs based on the refined OM+Texture criterion

The differences between the descriptors of the mineral and organic soil distributions were examined for OM thresholds of 20, 15 and 10% (Table 2). The low number of entries in three derived partial datasets weakened the statistical power of Fisher's Least Significant Difference test. However, applying the 10% OM threshold led to a better distinction than the 15% or 20% thresholds for distinguishing between the Mineral and Organic groups, with the GM-Mineral being around 2-fold greater than the GM-Organic, and with decreasing GSD for the Organic groups when decreasing the percent OM thresholds. When applying the 10% threshold, it was possible to confirm statistical differences between the Organic and the Mineral-Loam datasets, whereas no changes were observed for the Clay group still with a very low number of entries. These findings generally agreed with the Am sorption mechanisms, as Am is sorbed in a lesser extent in soils with higher OM content (potentially leading to higher DOM), whereas those soils with higher specific surface area (here indirectly represented by the clay content) present higher Am sorption capacity. Therefore, and despite the low number of entries in the derived datasets, a new 10% OM content threshold was established to distinguish between mineral and organic soils for Am K_d.

Although the variability decreased with respect to that of the overall dataset, the data variability was still very high in the partial datasets despite the redefined OM+Texture criterion, in which K_d data still varied more than 2-3 orders of magnitude. This fact suggests that the redefined OM+Texture criterion does not capture all the factors relevant to the Am sorption, such as pH.

3.3. Am K_d best estimates and CDFs based on soil factors related to Am sorption mechanisms.

3.3.1. The pH criterion

As an alternative approach, the Am K_d overall dataset was split based on the pH, as described in Section 2.2 (Table 3). As no entries with a pH lower than 3 were available, this value defined the lowest pH value of the examined pH ranges. Consistent with basic understanding of trivalent geochemistry (Choppin 2007), the GM of the pH partial datasets gradually increased within the 3 - 9 pH range, and decreased at higher pH values, which was consistent with the Am sorption mechanisms. At pH $\geq$ 9, especially for mineral soils, the lower GM K_d values are consistent with the observation that greater concentrations of dissolved Am-carbonate complexes are formed; these species bond relatively much weaker to mineral or OM surfaces.

The pH criterion was suitable to propose Am K_d best estimates with a lower related variability, sometimes lower than 2-orders of magnitude, as it directly considers one of the parameters with greater relevance for the Am sorption in soils. However, the relatively high GSD values, especially at pH > 9, suggest that other factors than pH may further decrease Am K_d variability.

3.3.2. Hierarchical application of pH, OM and Texture criteria

A combined grouping criterion, considering simultaneously the soil pH and OM (the pH+OM criterion) was applied to further decrease Am K_d variability. Table S1 in the Supplementary Material summarises the Am K_d descriptors of the distributions derived from the pH+OM criterion. The application of the pH+OM criterion led to Mineral - pH partial datasets that did not significantly differ from the respective pH partial datasets derived from the pH criterion, without a further reduction in variability. Thus, it was confirmed the key role of the pH factor in the Am sorption in mineral soils. It was also confirmed that Am GM for pH < 6 and > 9 mineral soils groups were lower than for the intermediate pH ranges. Regarding Organic soils, although conclusions were preliminary due to the low number of entries, the GM of the derived partial datasets indicated a negligible role of the pH, thus confirming that when there is a sufficient amount of organic matter, this factor overcomes pH in the Am interaction.

A final attempt was made to propose Am K_d best estimates with a lower variability by analysing the simultaneous use of all the previously examined factors (OM, pH and Texture) for Am K_d grouping of mineral soils. However, most of the derived partial datasets generally did not contain either enough data (N << 10) to construct CDFs or a single entry was only available. Only the pH - Loam datasets had enough data to reliably construct a CDF.

3.4. Exploring K_d data potentially analogous to soil Am K_d data to decrease variability: consideration of Am analogue elements and soil analogue geological materials

To further address the combined use of pH, OM and texture to decrease Am K_d variability, and to overcome the observed data gaps, the use of K_d data originating from Am in sediments and soils and from analogue elements in soils or in similar geological materials were evaluated. Thus, the dataset was enriched with data from trivalent actinides and lanthanides (especifically, La, Sm, Eu, Gd, Er, Lu and Cm) and with geological materials other than soils (gyttjas, tills, and subsoils), thus defining an Analogue K_d dataset with around 170 entries. To test that the use of these analogues was appropriate, the Analogue K_d dataset was split into partial datasets based on the geological material (soil, till, gyttja, and subsoil). Subsequently, data of each material partial dataset was grouped, when possible, according to the OM and pH criteria.

The statistical tests revealed that soil Am K_d data for the OM and pH partial datasets did not significantly differ from corresponding groups of K_d values created with the analogue data (Table 4). Besides this, the Analogue K_d GMs, as well as the 5^{th}-95^{th} percentile ranges of the CDFs that could be constructed, followed the same trend as Am K_d (i.e., GM tended to increase when increasing pH, until pH values >9), whereas in some cases the higher number of entries permitted to statistically confirm previous conclusions derived from Am K_d (i.e. , GM of the Mineral soils was statistically higher than that for Organic soils; the GM of the materials within the $7.5 \leq pH < 9$ range were statistically higher).

Given these statistical characteristics of the Analogue dataset with respect to the Am K_d soil dataset, it was concluded that the Analogue dataset was suitable for enhancing the soil Am K_d soil partial datasets. Thus, all available Am and Analogue K_d were pooled into a single dataset and the OM, pH and Texture criteria were applied with the intention of creating partial datasets with lower overall variability. Table 5 summarises the K_d descriptors of the enhanced distributions, whereas Figure 1 plots the CDFs of selected scenarios. The inclusion of Analogue K_d data succeeded in enhancing the number of entries for most partial datasets, although there were still data gaps for certain pH - Textural combinations that did not permit the calculation of CDF, especially for pH+Clay datasets (Table 5). Texture criterion was not applied for samples with pH > 9 because, as stated before, sorption mechanisms at this pH range do not depend on texture.

Regarding the OM criterion, the GM of the Mineral and Organic samples statistically differed, with a significant decrease in the K_d variability, especially for the Organic soils, with low GSD and 5^{th}-95^{th} percentiles ranging within only one order of magnitude (Figure 1a). Moreover, the relative sequence of K_d GMs for each textural class being statistically different and they followed the expected sequence: GM Sand < GM Loam < GM Clay (Figure 1b).

Regarding the pH criterion, previous conclusions were verified, as for mineral soils the GM increased when increasing pH <9, with values being statistically different and related variability decreased, especially when excluding organic soils, with values ranging within 2-order of magnitude or less (Figure 1c). Moreover, variability of these pH+Mineral groups was lower than for the overall mineral group, especially when the percentiles of the derived CDF are compared. On the contrary, GM for the Organic groups did not vary among the pH ranges tested with a sufficient number of entries, thus confirming that pH is not a relevant factor to decrease Am K_d variability when dealing with organic soils. The single value of K_d for organic soils with $pH \geq 9$ indicated a lower value than for the previous pH ranges. Going further from the pH+OM criterion by splitting data into textural classes for mineral soils did not lead to further improvements, especially due to the lack of data for clay soils at a given pH range, although in some cases the Loam texture had statistically different GM, with only 1 order of magnitude variability.

Although this work does not aim at examininig multivariate correlations between soil properties and K_d of lanthanides (Ln (III)) and actinides (An (III)) in depth, it is of interest to check whether simple, linear regressions may agree with the main conclusions drawn from distribution descriptors summarized in Table 5. As an example, the correlation between K_d and pH and clay content (% with respect to total soil weight) for mineral soils at pH < 9 was examined. The resulting equation was:

$$\log K_d = 2.00\ (0.25) + 0.34\ (0.14) \times \log \text{clay} + 0.25\ (0.04) \times \text{pH}\ (N=127;\ r = 0.62;\ p = 1.5\times10^{-13}) \quad \text{(Eq. 1)}$$

where the values into brackets indicate the standard error of every coefficient. As this multivariate correlation originates from a highly heterogenous dataset (it includes data from several lanthanides and actinides, from diverse geological materials, and obtained by different experimental methods), it is thus confirmed that these soil properties, often obtained in routine soil analyses, are key factors to describe and thus reducing K_d variability of these elements in mineral soils, and thus they can also be used in predicting the order of magnitude of K_d values of Ln (III) and An (III) in mineral soils.

4. Conclusions and recommendations

The Am K_d values in soils in the current compilation initially varied more than 4 orders of magnitude, primarily because of the contrasting properties of the soils. It was demonstrated that when soils are grouped based on soil properties, statistically different probabilistic Am K_d data are obtained for different soil-types, and generally, with much lower uncertainty. Therefore, it is not recommended the use of a single Am K_d best estimate or CDF to perform radiological assessments related to soils contaminated with Am because it may not be representative of numerous scenarios, and Am K_d best estimates (as calculated from the 50th percentile of CDF distributions) may differ in nearly two orders of magnitude depending on soil characteristics, especially pH and organic matter content.

K_d data available of Am chemical analogues (trivalent actinides and lantanides) gathered from soils and other geological materials were used to fill some of the Am K_d data gaps. These analogue K_d value entries did not statistically differ from Am K_d in the materials tested, which allowed to for enhancement of existing CDFs but also additional comparisons for other pH and texture combinations, and to confirm conclusions based on improved statistical analyses. However, some data gaps still exist and, therefore, it was not possible to unequivocally suggest CDF for all soil groups created based on pH+OM+texture criterion. From the analyses of the K_d of the Am+Analogues it was demonstrated the suitability of interchangeably using Am and Ln (III) and An (III) data in radiological assessments dealing with these families of elements.

The K_d best estimates to be used will eventually depend on the available information of the soil at the study site. Geological materials should be increasingly characterized in terms of specific surface area and DOM for a better examination of Am K_d variability, as OM content and texture may only serve as subrogates for these properties. If only information on pH is available, best estimates can be selected as shown in Table 5 for the overall soils. For soils having an OM content > 10%, a single best estimate value could be recommended for soils having a pH < 9 ($4.5x10^3$ L/kg), with a 5^{th}-95^{th} range of only one order of magnitude, and one order of magnitude lower best estimate ($4.4x10^2$ L/kg) for organic soils with pH ≥ 9. If no pH information is available, the first value could also be used. On the other hand, for those soils with an OM content < 10%, best estimates included in Table 5 can be selected depending on whether pH information is available or not. In some cases, using textural information may lead to best estimates with a much lower uncertainty, especially when distinguishing sand soils from clay and loam soils.

References

Choppin, G.R., 2007. Actinide speciation in the environment. J. Radioanal. Nucl. Ch. 273 (3), 695-703.

Ciffroy, P., Durrieu, G., Garnier, J. M., 2009. Probabilistic distribution coefficients (Kds) in freshwater for radioisotopes of Ag, Am, Ba, Be, Ce, Co, Cs, I, Mn, Pu Ra, Ru, Sb, Sr and Th. Implications for uncertainty analysis of models simulating the transport of radionuclides in rivers, J. Environ. Radioactiv. 100, 785-794.

EPA, 2004. EPA 402-R-04-002C - Understanding variation in partition coefficient, Kd, values. Vol III: Review of geochemistry and available Kd values for americium, arsenic, curium, iodine, neptunium, radium and technetium. Environmental Protection Agency, USA.

Gil-García, C.J., Tagami, K., Uchida, S., Rigol, A., Vidal, M., 2009. New best estimates for radionuclide solid–liquid distribution coefficients in soils. Part 3: miscellany of radionuclides (Cd, Co, Ni, Zn, I, Se, Sb, Pu, Am and others). J. Environ. Radioact. 100, 704–715.

GRS, 2012. Radionuclide Inventory of Vitrified Waste after Spent Nuclear Fuel Reprocessing at La Hague. Gesellschaft für Anlagen- und Reaktorsicherheit, GRS-294, Germany.
Ho Lee, M., Chang, E., Song, K., Hee, Y., Sang, H., 2011. The influence of humic acid on the pH-dependent sorption of americium (III) onto kaolinite. J. Radioanal. Nucl. Chem. 287, 639-645.
IAEA, 2010. TRS-472 - Handbook of Parameter Values for the Prediction of Radionuclide Transfer in Terrestrial and Freshwater Environments. International Atomic Energy Agency, Austria.
Kaplan, D. I., Serne, R. J., Owen, A. T., Conca, J., Wietsma, T. W., Gervais T. L., 1996. PNNL-11485 Radionuclide adsorption distribution coefficients measured in Hanford sediments for the low-level waste performance assessment project. Pacific Northwest national Laboratory, USA.
Kaplan, D.I., Powell, B.A., Demirkanli, D.I., Fjeld, R.A., Molz, F.J., Serkiz, S.M., Coates, J.T. 2004. Influence of oxidation states on plutonium mobility during longterm transport through an unsaturated subsurface environment. Environ. Sci. Technol. 38 (19), 5053–5058.
Lujaniene, G., Motiejunas, S., Sapolaite, J., 2007. Sorption of Cs, Pu and Am on clay minerals. J. Radioanal. Nucl. Chem. 274 (2), 345–353.
Pavlotskaya, F.I., Goryachenkova, T.A., Kazinskaya, I.E., Novikov, A.P., Myasoedov, B.F., Kuznetsov, Y., Legin, V.K., Shishkunova, L.V., 2003. Speciation and migration behaviour of Pu and Am in floodplain soils and bottom sediments of the Yenisei River. Radiochemistry 45 (5), 524–531.
Ramírez-Guinart, O., Vidal, M., Rigol, A., 2016. Univariate and multivariate analysis to elucidate soil properties governing americium sorption in soils. Geoderma 269, 19-26.
Ramírez-Guinart, O., Rigol, A., Vidal, M., 2017. Assessing soil properties governing radiosamarium sorption in soils: can trivalent lanthanides and actinides be considered as analogues? Geoderma 290, 33-39.
Ramírez-Guinart, O., Kaplan, D., Rigol, A., Vidal, M. 2020a. Deriving probabilistic soil distribution coefficients (K_d). Part 1: General approach to decreasing and describing variability and example using uranium K_d values. Submitted to J. Environ. Radioactiv.
Ramírez-Guinart, O., Kaplan, D., Rigol, A., Vidal, M. 2020b. Deriving probabilistic soil distribution coefficients (K_d). Part 2: Reducing caesium K_d uncertainty by accounting for experimental approach and soil properties. Submitted to J. Environ. Radioactiv.
Sheppard, S. C., 2011. Robust Prediction of K_d from Soil Properties for Environmental Assessment. Hum. Ecol. Risk Assess. 17(1), 263-279.

Vandenhove, H., Gil-García, C., Rigol, A., Vidal, M. 2009. New best estimates for radionuclide solid-liquid distribution coefficients in soils. Part 2. Naturally occurring radionuclides. J. Environ. Radioactiv. 100, 697-703.

Wong, J.C., Zavarin, M., Begg, J.D., Kersting, A.B., Powell, B.A., 2015. Effect of equilibration time on Pu desorption from goethite. Radiochim. Acta 103 (10), 695–705.

Table 1

Descriptors of Am K_d (L kg^{-1}) distributions after applying the OM+Texture criterion.

Partial dataset	N	GM	GSD	FLSD[1]	FLSD[1]	5th	95th
Mineral+Organic	55	7.3×10^3	5.8			2.1×10^2	9.1×10^4
Mineral (OM < 20%)	49	8.1×10^3	5.8	a		2.8×10^2	9.1×10^4
Clay*	3	1.1×10^4	15		-	n.a.*	n.a.
Loam	28	1.1×10^4	4.2		a	9.9×10^2	1.0×10^5
Sand	14	5.8×10^3	4.9		b	6.7×10^1	3.7×10^4
Organic (OM ≥ 20%)*	6	4.0×10^3	6.0	a	ab	n.a	n.a.

N = number of observations, GM = geometric mean, GSD = geometric standard deviation
[1] Different letters indicate significant differences between GMs according to the Fisher's Least Significant Differences test.
*CDF not constructed due to lack of Am K_d data (N << 10); n.a.: not applicable

Table 2

Descriptors of Am K_d (L kg^{-1}) distributions for soils grouped according to the redefined OM+Texture criterion.

Partial dataset	N	GM	GSD	FLSD[1]	FLSD[1]	5th	95th
Mineral (OM < 20%)	49	8.1×10^3	5.8	a		2.8×10^2	9.1×10^4
Organic (OM ≥ 20%)*	6	4.0×10^3	6.0	a		n.a.*	n.a.
Mineral (OM < 15%)	46	8.5×10^3	6.1	a		2.8×10^2	9.1×10^4
Organic (OM ≥ 15%)	9	5.0×10^3	4.2	a		2.1×10^2	4.9×10^4
Organic (OM ≥ 10%)	10	5.1×10^3	3.9	a	a	2.1×10^2	4.9×10^4
Mineral (OM < 10%)	45	8.8×10^3	6.2	a		2.8×10^2	9.1×10^4
Clay*	3	1.1×10^4	15			n.a.	n.a.
Loam	25	1.2×10^4	4.3		b	9.9×10^2	1.0×10^5
Sand	13	4.8×10^3	5.2		a	6.7×10^1	3.7×10^4

N = number of observations, GM = geometric mean, GSD = geometric standard deviation
[1] Different letters indicate significant differences between GMs according to the Fisher's Least Significant Differences test.
*CDF not constructed due to lack of Am K_d data (N << 10); n.a.: not applicable

Table 3

Descriptors of Am K_d (L kg^{-1}) distributions for soils grouped according to the pH criterion.

Partial dataset	N	GM	GSD	FLSD[1]	5^{th}	95^{th}
3 < pH < 6	34	3.2×10^3	4.5	ab	2.0×10^2	4.8×10^4
6 ≤ pH <7.5	23	7.3×10^3	5.3	bc	1.8×10^2	3.6×10^4
7.5 ≤ pH < 9	21	1.1×10^4	5.8	c	1.3×10^3	1.0×10^5
pH ≥ 9	9	4.8×10^2	8.7	a	2.7×10^1	2.2×10^4

N = number of observations, GM = geometric mean, GSD = geometric standard deviation
[1] Different letters indicate significant differences between GMs according to the Fisher's Least Significant Differences test.

Table 4

Descriptors of Analogue K_d (L kg^{-1}) distributions from soils and other materials grouped according to the OM and pH criteria.

Material	Partial dataset	N	GM	GSD	FLSD[1]	5^{th}	95^{th}
Soil	Mineral+Organic	116	8.3×10^3	4.2		5.0×10^2	9.6×10^4
	Organic (OM ≥ 10%)	54	6.3×10^3	5.3	a	1.3×10^3	1.5×10^4
	Mineral (OM < 10%)	62	1.4×10^4	2.6	b	3.3×10^2	9.8×10^4
	3 ≤ pH < 6	47	4.1×10^3	3.3	a	5.0×10^2	3.1×10^4
	6 ≤ pH < 7.5	34	8.3×10^3	2.9	a	1.3×10^3	4.0×10^4
	7.5 ≤ pH < 9	31	2.5×10^4	5.9	b	2.9×10^2	1.4×10^5
	pH > 9*	4	1.3×10^4	3.1	-	n.a.*	n.a.
Till	Mineral (OM < 10%)	28	3.3×10^4	4.3		1.2×10^3	7.6×10^4
	3 ≤ pH < 6	11	2.4×10^3	5.4	a	1.2×10^3	8.0×10^4
	6 ≤ pH < 7.5*	6	2.1×10^4	2.2	-	n.a.	n.a.
	7.5 ≤ pH < 9	11	5.1×10^4	1.5	b	2.5×10^4	7.6×10^4
Subsoils	Mineral (OM < 10%)*	4	2.5×10^4	2.2		n.a.	n.a.
Gyttja	Organic (OM ≥ 10%)	24	2.6×10^3	2.1		1.3×10^3	9.0×10^3

N = number of observations, GM = geometric mean, GSD = geometric standard deviation

[1] Different letters indicate significant differences between GMs according to the Fisher's Least Significant Differences test.

*CDF not constructed due to lack of Am K_d data (N << 10). n.a.: not applicable

Table 5

Descriptors of Am+Analogue K_d ($L\ kg^{-1}$) distributions grouped according to the OM, pH and Texture criteria.

Partial dataset		N	GM	GSD	FLSD[1]	5th	95th
Mineral+Organic		240	7.4×10^3	4.8		4.7×10^2	7.8×10^4
Organic		84	4.5×10^3	2.7	a	1.3×10^3	1.5×10^4
Mineral		156	1.1×10^4	5.7	b	3.5×10^2	9.6×10^4
Mineral	Clay	15	5.6×10^4	5.7	a	2.8×10^2	3.9×10^5
Mineral	Loam	78	1.9×10^4	3.4	b	2.7×10^3	9.1×10^4
Mineral	Sand	44	4.9×10^3	5.1	c	2.2×10^2	3.7×10^4
$3 \leq pH < 6$	Overall	127	3.0×10^3	3.6	a	4.1×10^2	3.9×10^4
$6 \leq pH < 7.5$	Overall	76	9.8×10^3	4.3	b	2.8×10^2	6.1×10^4
$7.5 \leq pH < 9$	Overall	71	3.7×10^4	5.1	c	6.9×10^2	1.4×10^5
$pH \geq 9$	Overall	17	8.8×10^3	16.5	ab	2.7×10^1	2.2×10^5
$3 \leq pH < 6$	Mineral	54	2.1×10^3	5.4	a	2.5×10^2	5.2×10^4
$6 \leq pH < 7.5$	Mineral	48	1.0×10^4	5.2	b	1.9×10^2	4.7×10^4
$7.5 \leq pH < 9$	Mineral	62	3.7×10^4	4.6	c	1.3×10^3	1.1×10^5
$pH \geq 9$	Mineral	12	3.8×10^3	10.4	a	2.7×10^1	3.6×10^4
$3 \leq pH < 6$	Organic	69	3.2×10^3	2.5	a	1.3×10^3	1.4×10^4
$6 \leq pH < 7.5$	Organic	20	6.7×10^3	1.9	a	1.7×10^3	1.3×10^4
$7.5 \leq pH < 9$*	Organic	5	5.0×10^3	11.5	a	n.a.*	n.a.
$pH \geq 9$*	Organic	1	4.4×10^2	n.a.	-	n.a.	n.a.
$3 \leq pH < 6$	Loam	24	7.4×10^3	6.1	a	2.5×10^2	6.3×10^4
	Sand	18	2.2×10^3	4.2	a	6.7×10^1	3.1×10^4
$6 \leq pH < 7.5$	Clay*	4	1.7×10^3	20	a	n.a.	n.a.
	Loam	26	1.5×10^4	2.4	b	4.6×10^3	4.9×10^4
	Sand	14	4.6×10^3	5.3	a	1.8×10^2	1.8×10^4
$7.5 \leq pH < 9$	Clay	11	9.2×10^4	2.7	a	4.5×10^3	1.4×10^5
	Loam	35	3.7×10^4	3.1	a	2.7×10^3	1.1×10^5
	Sand	20	4.9×10^3	7.8	b	6.6×10^1	4.7×10^4

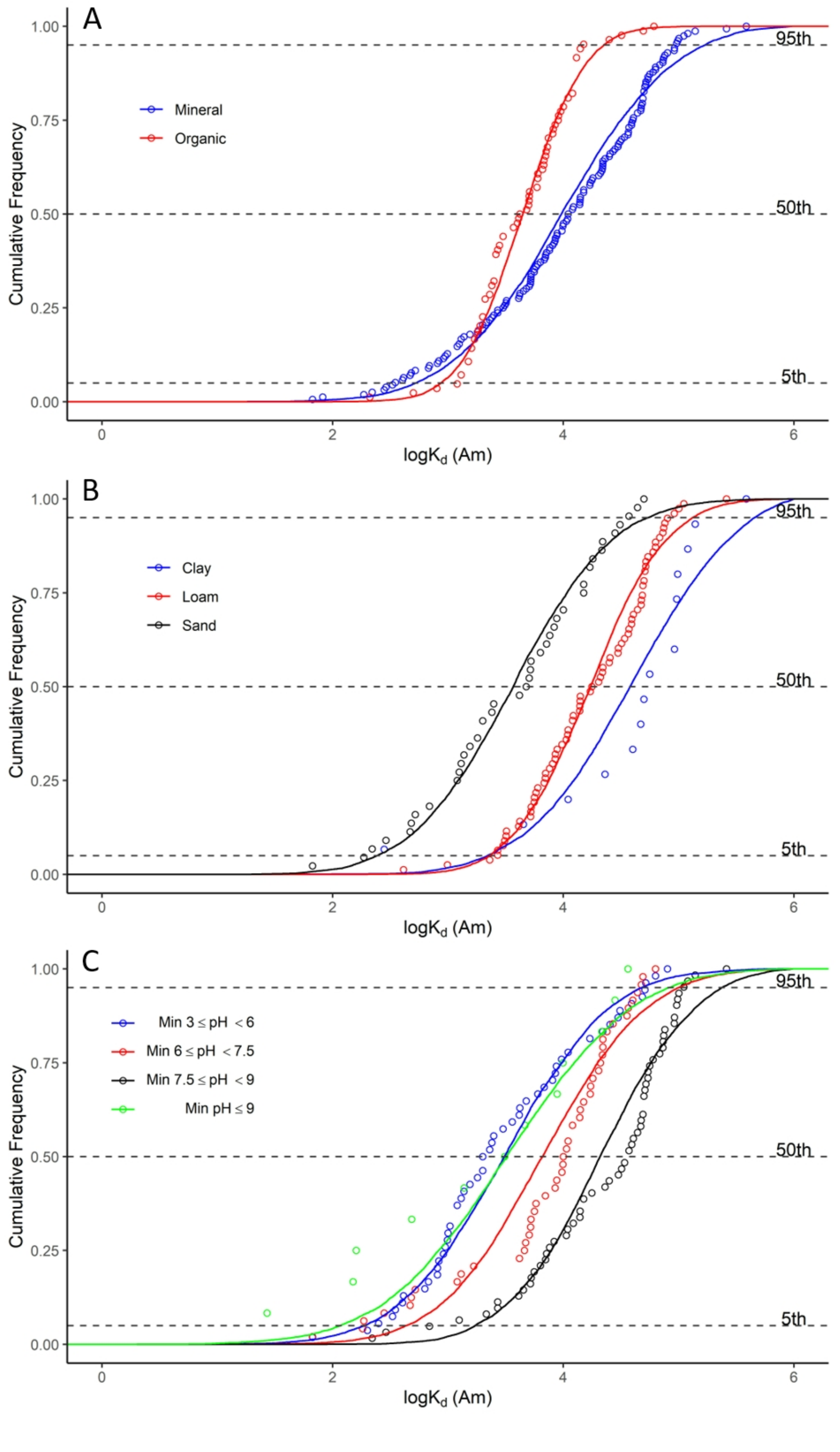
N = number of observations, GM = geometric mean, GSD = geometric standard deviation

[1] Different letters indicate significant differences between GMs according to the Fisher's Least Significant Differences test.

*CDF not constructed due to lack of K_d data (N << 10); n.a.: not applicable

Figure captions

Fig. 1. CDFs of Am+Analogue K_d (L kg^{-1}). 1A. Mineral and Organic datasets. 1B. Clay, Loam and Sand groups of the mineral soils. 1C. Mineral soils within $3 \leq pH < 6$, $6 \leq pH < 7.5$, $7.5 \leq pH < 9$, and $pH \geq 9$ ranges. Points indicate individual dataset values whereas lines indicate the fitted distributions.


A
Mineral
Organic
Cumulative Frequency
95th
50th
5th
logK$_d$ (Am)
B
Clay
Loam
Sand
C
Min 3 ≤ pH < 6
Min 6 ≤ pH < 7.5
Min 7.5 ≤ pH < 9
Min pH ≤ 9

SUPPLEMENTARY MATERIAL

Deriving probabilistic soil distribution coefficients (K_d). Part 3: Reducing variability of americium K_d best estimates using soil properties and chemical and geological material analogues

Table S1

Descriptors of Am K_d (L kg^{-1}) distributions for soils grouped according to pH+OM.

		N	GM	GSD	FLSD[1]	5th	95th
$3 \leq pH < 6$	Mineral	25	2.5×10^3	5.7	ab	2.0×10^2	4.8×10^4
	Organic	9	3.2×10^3	1.6	-	1.6×10^3	7.9×10^3
$6 \leq pH < 7.5$	Mineral	19	8.7×10^3	6.2	b	6.7×10^1	4.7×10^4
	Organic*	4	6.4×10^3	1.2	-	n.a.*	n.a.
$7.5 \leq pH < 9$	Mineral	18	1.4×10^4	4.2	c	1.3×10^3	1.1×10^5
	Organic*	3	2.9×10^3	n.a.	-	n.a.	n.a.
$pH \geq 9$	Mineral	8	8.2×10^2	10	a	2.7×10^1	2.2×10^4
	Organic*	1	4.4×10^2	n.a.	-	n.a.	n.a.

N = number of observations, GM = geometric mean, GSD = geometric standard deviation
[1] Different letters indicate significant differences between GMs according to the Fisher's Least Significant Differences test.
*CDF not constructed due to lack of Am K_d data (N << 10); n.a.: not applicable