

(Un)certainty of overall binding constants of Al with dissolved organic matter determined by the Scatchard approach

Boris Jansen*, Klaas G.J. Nierop, Jasper A. Vrugt, Jacobus M. Verstraten

Centre for Geo-ecological Research (ICG), Institute for Biodiversity and Ecosystem Dynamics (IBED), Physical Geography, Universiteit van Amsterdam, Nieuwe Achtergracht 166, NL-1018 WV Amsterdam, The Netherlands

Received 6 September 2002; received in revised form 15 September 2003; accepted 19 November 2003

Abstract

One of the best approaches to date to obtain overall binding constants (K_o) for Al and dissolved organic matter (DOM) from acidic soil solutions is to collect ‘free’ Al data with diffusive gradients in thin films (DGT) and to infer the K_o values by fitting a continuous distribution model based on Scatchard plots. Although there is clear established literature demonstrating the usefulness of the Scatchard approach, relatively little attention has been given to a realistic assessment of the uncertainties associated with the final fitted K_o values. In this study we present an uncertainty analysis of the fitted K_o values using a synthetic dataset with different levels of random noise and a real data set using DGT data from an acidic soil solution. The parameters in the continuous distribution model and their corresponding upper and lower 95% uncertainty bounds were determined using the Shuffled Complex Evolution Metropolis (SCEM) algorithm. Although reasonable fits of the distribution model to the experimental data were obtained in all cases, an appreciable uncertainty in the resulting K_o values was found due to three main reasons. Firstly, obtaining ‘free’ Al data even with the DGT method is relatively difficult, leading to uncertainty in the data. Secondly, before Scatchard plots can be constructed, the maximum binding capacity (MBC) must be estimated. Any uncertainty in this MBC propagates into uncertainty associated with the final plots. Thirdly, as the final fitted K_o values are largely based on extrapolation, a small uncertainty in the fit of the binding data results in an appreciable uncertainty in the obtained K_o . Therefore, while trends in K_o for Al and DOM could easily be discerned and compared, the uncertainty in the K_o values hinders the application in quantitative speciation calculation. More comprehensive speciation models that avoid the use of K_o seem to fit better for this purpose.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Scatchard; Al; DOM; Stability constants; SCEM; Uncertainty

1. Introduction

Aluminum plays an important role in the biogeochemical processes that take place in ecosystems. It contributes, among others, to the acidification of soils, lakes and streams [1] and is toxic to soil and water organisms when present in too high a concentration [1,2]. Dissolved organic matter (DOM) can interact with Al thereby forming either dissolved or precipitated Al–

DOM complexes [3], which can be very strong and stable due to binding of Al to multiple ligands at the same time [4]. Complexation of Al with DOM greatly influences its mobility, bioavailability and toxicity in soils and water [2,5].

The interactions of Al and other metals with DOM have been the subject of intensive research attention over the last two decades, and several methods have been developed to quantify these interactions ([6] and references cited therein). The two most comprehensive approaches to date are the speciation models: Model VI [7] and NITCA-Donnan [8]. Both models appear to be able to adequately calculate the speciation of a variety of

*Corresponding author. Tel.: +31-20-5257-444; fax: +31-20-5257-431.

E-mail address: bjansen@science.uva.nl (B. Jansen).

metals in solution, including Al, in the presence of DOM. A simpler approach to investigate the interaction between metals and organic ligands in general, is the derivation of overall binding constants (K_o) from experimental data (e.g. [9–11]). The overall binding constant has no rigorous thermodynamic and molecular significance. Instead, it is a conditional binding constant that varies with soil solution variables of which the most important are metal loading and pH [10,12]. Thereby, it reflects the affinity of a given metal for binding to organic ligands under specific environmental conditions.

In spite of the existence of more comprehensive models, many researchers continue to derive and use conditional binding constants to compare relative binding strengths of Al and other metals with different types of ligands and to use them in speciation calculations (e.g. [13–15]). Therefore, an assessment of the robustness of K_o values for the binding of metals to DOM is of great importance. Especially, as there are two main concerns when deriving K_o values for metal binding to organic ligands. The first is obtaining accurate and precise data of the speciation of the metal over dissolved ‘free’ metal and soluble metal–DOM complexes. Here ‘free’ metal is usually defined as the pool of metal in solution that is present in the form of hydrated metal cation, plus its dissolved inorganic complexes. The second concern is finding a robust, objective mathematical method to derive K_o from the experimental data.

Concerning the experimental data, when looking specifically at Al, the problem is that in contrast to metals for which, for instance, ion selective electrodes can be used, experimentally distinguishing between ‘free’ Al and dissolved Al–DOM complexes in acidic soil solutions is difficult. Several analytical methods such as equilibrium dialysis, Clarke’s quickly reacting Al method and the Donnan-membrane technique have been employed for this purpose [16–18]. Unfortunately, most of these methods are prone to measurement error due to the relative complexity of the analyses, which limits the precision of the data obtained. In addition, many methods are time consuming, restricting the number of data points that can be obtained within a given time frame and further restricting the precision of the experimental data. This has limited the available datasets of Al–DOM interactions in solution to date [7]. Furthermore, each method uses its own separation strategy and the results can differ as much as an order of magnitude, limiting the accuracy of the data [11]. One of the most promising methods currently available for the determination of ‘free’ Al in acidic soil solutions is the method of diffusive gradients in thin films (DGT) [19]. This method combines relative robustness with a reasonable time needed per analysis [19].

There are several ways to derive K_o from experimental speciation data of Al and DOM in the concentration

ranges commonly found in soil solution, ground water and surface water. It is generally agreed that the approach of Scatchard plots is most suitable for this purpose [11,20]. Although there is a clear established volume of literature demonstrating the usefulness of the Scatchard approach, relatively little attention has been given to a realistic assessment of the uncertainty associated with the final binding constants. This is of special importance because there are various sources of error contributing to the final uncertainty of the fitted K_o values when using Scatchard plots. One of the variables needed in the Scatchard approach is a measure of the binding capacity of DOM for the metal of interest [21]. The maximum binding capacity (MBC) expresses the maximum amount of the metal of interest that can bind to DOM on a mmol/mol C basis [21]. There are several ways to derive a measure for the binding capacity of DOM for Al, but no matter which approach is used, the uncertainty in this variable will affect the final binding constants [21]. Another potential problem when using Scatchard plots is the lack of consensus on how to fit the curves that are generated to obtain the binding constants of interest [9,10,22]. The choice of a certain method not only influences the values of the binding constants that are obtained, but also influences the uncertainty of those values.

In a previous study we used the Shuffled Complex Evolution Metropolis (SCEM) algorithm to derive K_o for the binding of Cu to DOM using Scatchard plots of DGT data, and assessed the uncertainty of the values that were thus obtained [23]. In the present study we applied a similar procedure to quantify the uncertainty in the overall binding constants for Al and DOM in soil solutions as obtained using the Scatchard approach, and thereby to critically evaluate the value of this commonly used method for the quantification of interactions of Al with DOM. To accomplish this we used experimental data obtained in a previous study of Al–DOM interactions in forest soil solutions [24] and a synthetic perfect dataset to which different levels of random noise were added.

2. Materials and methods

2.1. Al–DOM addition experiments

The speciation data of Al over ‘free’ Al and soluble Al–DOM complexes were obtained during a previous study in which the DGT experiments are described in detail [24].

2.2. The Scatchard approach

The reaction between a metal (M) and organic ligand (L) with successive reaction products (LM,

LM_2, \dots, LM_n) can be described as

$$K_1 = \frac{[LM]}{[L] + [M]}; \quad K_2 = \frac{[LM_2]}{[LM] + [M]};$$

$$K_n = \frac{[LM_n]}{[LM_{n-1}] + [M]}. \quad (1)$$

The extent of binding (v) of the metal to the organic ligand is expressed as follows:

$$v = \frac{M_b}{L_t} = \frac{[LM] + 2[LM_2] + \dots + n[LM_n]}{[L] + [LM] + \dots + [LM_n]}. \quad (2)$$

Here M_b represents the concentration of metal bound to DOM and L_t the total organic ligand concentration that is related to the concentration of functional groups involved in metal binding on the DOM molecules. Rewritten in terms of stability constants, one obtains

$$v = \frac{K_1[M] + 2K_1K_2[M]^2 + \dots + nK_1K_2\dots K_n[M]^n}{1 + K_1[M] + K_1K_2[M]^2 + \dots + K_1K_2\dots K_n[M]^n} \quad (3)$$

which for independent and identical binding sites can be rewritten to yield the Scatchard equation

$$\frac{v}{M_f} = nK_o - vK_o. \quad (4)$$

Here M_f stands for the metal that is not complexed with DOM, i.e. the ‘free’ metal. When v/M_f is plotted against v , the inverse of the slope represents the overall binding constant K_o .

2.3. Calculation of the binding capacity and total organic ligand concentrations

Since L_t in Eq. (2) signifies the total organic ligand concentrations, it must be calculated from the DOM concentrations using a measure for the average MBC. Unfortunately, there is no consensus on how to derive the MBC. Some researchers use the total acidity (i.e. the number of carboxylic and phenolic hydroxyl groups) or just the carboxylic acidity as a measure of the MBC [3,10]. However, different metals have different binding affinities for different functional groups [3,21,25] and consequently organic matter has different MBC values for different metals. This aspect is ignored when MBC values are derived from total or carboxylic acidity data. A better way to obtain MBC values is by the Langmuir adsorption Eq. (1) in, e.g. a one-surface approach [3,21]:

$$x = \frac{(\text{MBC})K_L[M]}{1 + K_L[M]}. \quad (5)$$

In this equation, x represents the concentration of metal bound to 1 mol of DOC, and K_L is a constant that represents the affinity between a metal cation and DOM. $[M]$ stands for the total metal concentration of the metal of interest at equilibrium. The MBC is obtained from the reciprocal of the slope when $[M]/x$ is plotted against $[M]$. Note that K_L differs from K_o and consequently MBC does not equal L_t . Instead, L_t is

calculated from the MBC by multiplying it with the DOC concentrations.

2.4. Calculation of K_o

Al_f was obtained from the DGT data and Al_b by subtraction of Al_f from the total Al concentrations. These variables, together with L_t , were used to calculate v from Eq. (2) and construct Scatchard plots described by Eq. (4).

A problem when using the Scatchard approach is that K_o values are obtained by curve fitting of the curve generated by Eq. (4), which can be done in several ways. Many researchers have fitted one, two or even three lines to the curve to derive as many separate K_o values believed to represent different functional groups or types of complexes [9,21,22]. However, such a procedure can be quite arbitrary and greatly oversimplifies the complexation reactions, severely limiting the usefulness ([10] and references therein). We avoided this problem by attempting the more realistic approach of fitting non-linear regression lines through the data. These lines describe conditional K_o constants at specific pH values as a function of the extent of binding.

In summary, the following stepwise procedure was applied to obtain K_o values for Al:

1. $[Al_b]$ is calculated from $[Al_t]$ and $[Al_f]$ and is divided by $[DOC]$ to obtain x .
2. $[Al_b]/x$ is plotted against x (Eq. (5)) and the reciprocal of the linear regression line yields MBC.
3. MBC is multiplied by $[DOC]$ to obtain L_t , while $[Al_b]$ divided by L_t yields v (Eq. (2)).
4. $v/[Al_f]$ is plotted against v to obtain the final Scatchard plots (Eq. (4)).
5. A regression fitted through the Scatchard plots describes K_o .

2.5. Optimizing the parameters of the distribution model using the SCEM

For optimization and uncertainty assessment of both the MBC values and the binding constants, we used a powerful new general-purpose method: the SCEM algorithm [26]. The SCEM algorithm is a modified version of the original SCE global optimization algorithm developed by Duan et al. [27] and uses a Bayesian inference scheme to arrive at a stationary posterior distribution of the model parameters. It operates by selecting and modifying an initial population of parameter sets merging the strengths of the Metropolis algorithm [28], with the concepts of controlled random search [29], competitive evolution [30] and complex shuffling [27] to evolve the population of initial parameter sets to a stationary posterior target distribution. The SCEM algorithm is different from the original

SCE algorithm in two important ways. First, the trapezoidal probability distribution, used in the original SCE algorithm to ensure that the evolution process is competitive, was replaced by a linear probability distribution. This ensures that the probability that ‘better’ parents (parameter sets) contribute to the generation of offspring is equal to that of ‘worse’ parents. Second, the downhill Simplex method in the Competitive Complex Evolution algorithm outlined by Duan et al. [27], was replaced by a Metropolis—annealing method [28], thereby avoiding a deterministic drift towards a single optimum parameter set. Both these modifications are necessary to prevent the search from becoming mired in a small basin of attraction and thus to arrive at a stationary posterior target distribution.

From the final posterior distribution of the model parameters the 95% confidence intervals for the MBC and K_o values as well as the prediction uncertainty intervals associated with the final Scatchard plots were derived. The uncertainty in the optimized MBC values was taken into account when optimizing the Scatchard plots and assessing their uncertainty.

2.6. Sensitivity assessment using a synthetic dataset

Our experience suggests that the most important source of experimental uncertainty in binding experiments like the ones we described here, is the determination of Al_f . Even though as explained in the introduction, the DGT method represents one of the best methods currently available for determining Al_f in acidic soil solutions, it still involves many experimental steps during which errors may be introduced. For instance, the results strongly depend on the thickness of the diffusion gel layer [31]. Furthermore, while DGT is relatively fast and robust compared to other methods used for measuring Al_f , it is still a fairly complex and time consuming method [19]. Because of this, the manufacturer states in his documentation that DGT units are functioning properly when the results of the recommended performance tests are within 10% of the real value. Based on our performance tests of DGT this seems to be a realistic margin of error [19].

Since the uncertainty in ‘free’ Al measurements was high compared to the total Al and DOC analyses, when testing the sensitivity of the Scatchard approach with a synthetic dataset we focused on the uncertainty in Al_f only. As a consequence, we assumed all experimental data to be known with infinite precision except for Al_f . To investigate the sensitivity of the Scatchard approach for experimental uncertainty, we constructed a perfect dataset using values for total Al and bound Al (Al_b) within the range of values of our actual experimental data. We chose the total concentrations of Al and used an assumed expression for K_o based on the real dataset

to calculate Al_b . We used one realistic constant DOM concentration of 3000 μM . To the Al_b data we added a random noise between 0% and $\pm 2\%$, 0% and $\pm 5\%$ or 0% and $\pm 10\%$ and determined the MBC values and K_o as well as the uncertainty in these values in the same way as with the regular experimental data. The artificial dataset is presented in Table 1. It consists of 20 data points for each situation, which is quite high (but not unrealistic) considering the time generally needed for ‘free’ Al analyses.

3. Results and discussion

3.1. The experimental data

3.1.1. MBC values and their uncertainty

When we tried to derive MBC values for Al from our data, the following problem arose. While there was no significant precipitation of inorganic Al complexes at the pH values of our experiments, according to calculations using the speciation code PHREEQC [32], there was significant precipitation of organic Al complexes when increasing amounts of Al were added. This influenced the Al and DOM concentrations remaining in solution at higher total Al concentrations [24].

One could say that the overall binding constants represent the degree of interaction irrespective of whether the resulting complex remains in solution or precipitates out of solution. In that case in Eq. (5), for M the theoretical metal concentration, which assumes all metal to have stayed in solution, must be used and L_t must be derived using the theoretical DOM concentration. However, this approach will only be accurate if the system is at overall equilibrium, which means that all precipitation/dissolution reactions must be completely reversible and just as fast as interactions in solution, avoiding preferential interactions in solution over solution/precipitation interactions. We call this Scenario I.

An alternative approach is to consider the precipitated fraction as an inactive fraction that is effectively removed from the system. Preliminary experiments showed that the re-dissolution of precipitated organic Al complexes can have very slow kinetics and can even be (partially) irreversible. In that case in Eq. (5) for M the actual total metal concentrations in solution must be used and L_t must be derived using the actual DOM concentrations in solution. This may lead to a change in derived binding constants because of the resulting shift in the metal/DOM ratio and consequently the extent of binding in solution. Furthermore, the average MBC values may change because of preferential precipitation of a certain DOM fraction. We call this scenario where precipitation is assumed to be inert Scenario II. In order to compare the implications of a choice for either

Table 1
Synthetic dataset used in sensitivity analysis of the Scatchard approach

Total Al (μM)	Bound ^a Al (μM) 0% random error	Bound* Al (μM) 2% random error	Bound* Al (μM) 5% random error	Bound* Al (μM) 10% random error
20.00	12.48	12.53	12.01	12.33
26.84	16.33	16.01	16.25	15.80
33.68	19.98	19.60	20.42	19.24
40.53	23.47	23.17	24.39	22.83
47.37	26.78	26.88	26.18	26.21
54.21	29.95	29.42	29.21	30.50
61.05	32.97	32.80	34.18	30.47
67.89	35.86	36.05	34.90	32.55
74.74	38.63	38.97	39.81	38.31
81.58	41.28	41.60	42.97	44.34
88.42	43.82	43.10	42.65	47.63
95.26	46.26	46.18	45.06	44.08
102.11	48.61	48.49	46.42	45.30
108.95	50.86	50.56	48.71	54.65
115.79	53.03	52.29	53.77	50.25
122.63	55.11	55.50	53.41	56.72
129.47	57.12	57.58	59.08	62.45
136.32	59.06	59.60	57.13	61.01
143.16	60.93	60.87	58.92	65.44
150.00	62.73	62.87	65.83	56.58

^aBound Al signifies Al complexed with DOM. The DOC-concentration was 3000 μM C/L in all cases.

Table 2
Optimal MBC values and 95% upper and lower limits for the experimental data in $\text{mmol}_c/\text{mol C}$

	pH = 3.5 I	pH = 3.5 II	pH = 4.0 I	pH = 4.0 II	pH = 4.5 I	pH = 4.5 II
Optimum:	68.49	52.08	196.1	153.8	357.1	416.7
Minimum:	54.95	40.82	166.7	111.1	263.2	126.6
Maximum:	96.15	70.92	238.1	270.3	555.6	n.c.
Least-squares linear regression ^a	71.43	54.95	196.1	153.9	434.8	476.2

I = Scenario I, where metal-organic precipitation interacts with the solution phase; II = Scenario II, where metal-organic precipitation is assumed to be inert. n.c. denotes non-convergence of the fit (see text).

^a Values calculated by simple least-squares linear regression, provided as a frame or reference for the optimized values.

scenario, we used both scenarios in our calculations and compared the results.

The MBC values for the experimental data from our previous study as well as the 95% upper and lower confidence interval limits for both scenarios are presented in Table 2. Some outliers were removed from the experimental data before calculation. The values obtained by simple least-squares regression are provided as frame of reference for the optimized values. As can be seen, the uncertainty in the MBC values for Scenario I is lower than for Scenario II. This is due to the fact that in Scenario I, the total DOM content remains constant whereas the total DOM content decreases in Scenario II. The difference in uncertainty is the highest at pH = 4.5, where the extent of Al–DOM precipitation was also the highest, resulting in the highest decrease in DOM

concentrations in solution [24]. The maximum MBC at pH = 4.5 for Scenario II did not converge to a meaningful value. The optimal Langmuir plot in this situation was nearly horizontal. Combined with the higher uncertainties in the experimental data for Scenario II, this led to the plot of the maximum being completely horizontal. Since the MBC value is the reciprocal of the slope, it would be infinite.

From Table 2 a trend of strongly increasing MBC values with increasing pH is visible for both scenarios. This is in agreement with the results from our previous study, where we found that contrary to Fe(III) there is a significant pH influence on the binding of Al to DOM in the pH range 3.5–4.5 [24]. This suggests that Al is not able to deprotonate all acidic functional groups through coordination binding at pH = 3.5 and 4.0 [24]. Therefore,

a lower number of deprotonated acidic functional at lower pH values leads to a lower MBC. There is no significant difference between the optimum MBC values for the two scenarios at the same pH values.

3.1.2. K_o values and their uncertainty

The most appropriate transition model to fit the experimental data under consideration appeared to be the non-linear equation:

$$\frac{v}{M_f} = av^{-b} \quad (6)$$

from which K_o could be derived by using the inverse of the derivative

$$K_o = -\frac{d(v/M_f)}{dv} = abv^{-(b-1)}. \quad (7)$$

The optimal Scatchard plots as well as the 95% upper and lower uncertainty boundaries are presented in Fig. 1. As with the MBC values, the uncertainty in the plots based on Scenario I is in all cases lower than in the plots based on Scenario II. This is a direct result of the propagation of the lower uncertainty in the MBC values in the former case. The differences in the range of extent

of binding on the x -axes that are apparent from Fig. 1 are mainly caused by the differences in binding at the different pH values. This effect is strengthened in the case of Scenario II by an increased degree of precipitation at higher pH values. Notice that the uncertainty boundaries differ for different data points and are not necessarily evenly distributed above and below the optimal curve. This is a natural consequence of the Bayesian inference scheme implemented in the SCEM algorithm. From Fig. 1 it is very clear that the plots are in all cases curvilinear and that fitting one, two or even three straight lines through these plots would indeed result in arbitrary binding constants.

In Fig. 2 the relation between $\log(K_o)$ and the extent of binding is presented and in Tables 3 and 4, $\log(K_o)$ values for some discrete optimum values for the extent of binding for both scenarios are presented. While a selection was made as an example, all values of v were used in the calculations. The values of parameters a and b (Eqs. (6) and (7)) are also provided in Tables 3 and 4, and since they are more generic they pertain to the entire dataset. A problem when assessing the uncertainty of the K_o values is the common way of expressing them as

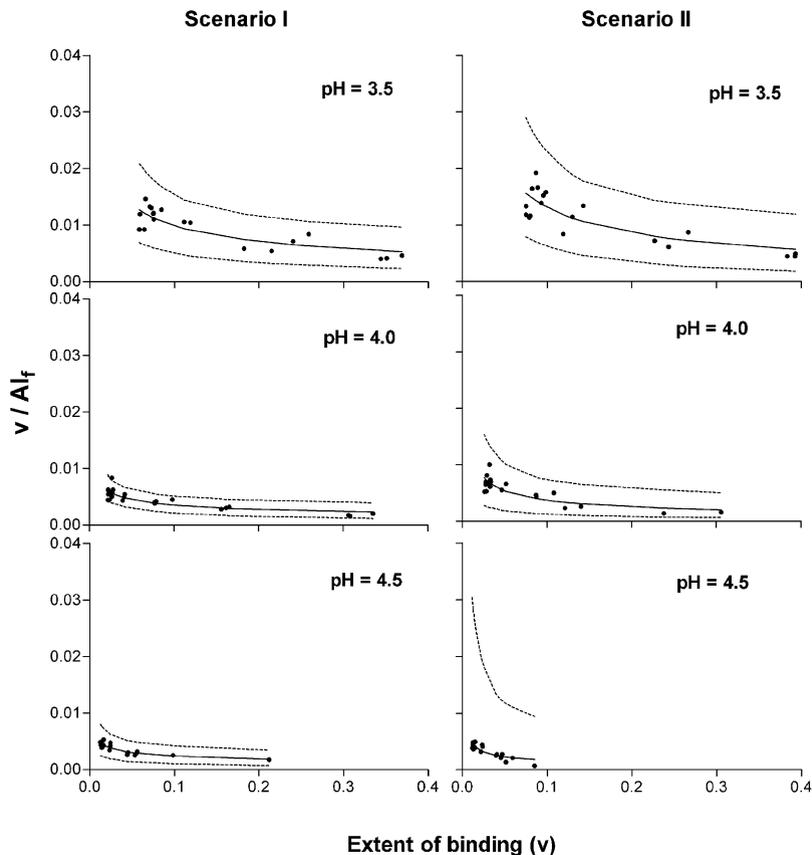


Fig. 1. Optimal Scatchard plots and their 95% confidence interval for Al and DOM based on real DGT data for 'free' Al (Al_f). Scenario I: metal-organic precipitation interacts with the solution phase; Scenario II: precipitation is assumed to be inert.

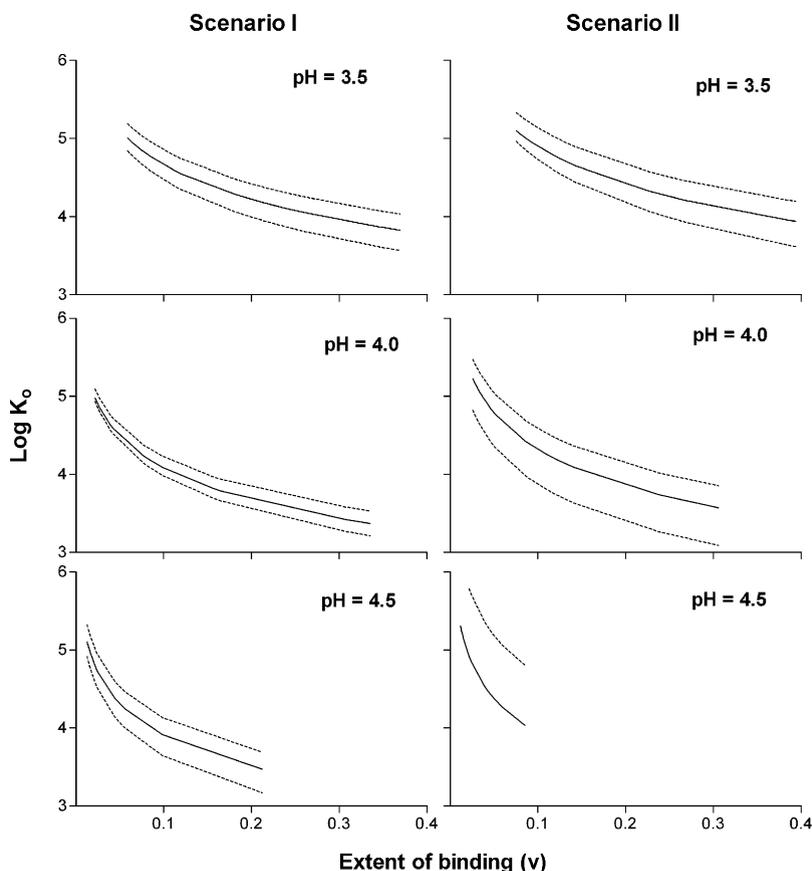


Fig. 2. Plots of the relationship between $\log(K_o)$ and the extent of binding for Al and DOM and their 95% confidence interval, based on real DGT data for 'free' Al (Al_f). Scenario I: metal-organic precipitation interacts with the solution phase; Scenario II: precipitation is assumed to be inert.

logarithmic values. This means that also the logarithmic value of the uncertainty is taken. At first glance this might lead to too optimistic an interpretation of the uncertainties. Therefore, in Tables 3 and 4 in addition to $\log(K_o)$ we included the K_o values themselves and their upper and lower bounds within the 95% confidence interval.

It is clear from Tables 3 and 4 that the uncertainty in the K_o values is appreciable. In the case of Scenario I, the maximum K_o values are $(1.6\text{--}3.3) \times$ the minimum K_o values, whereas for Scenario II, the maximum K_o values are $(2.0\text{--}5.6) \times$ the minimum K_o values for pH = 3.5 and 4.0. This uncertainty makes it difficult to obtain exact information about the affinity of DOM for Al in specific points since changes in K_o by the orders of magnitude as presented here will result in appreciable shifts in the distribution of Al over 'free' and organically complexed Al species.

The number of other studies where K_o was obtained for Al and DOM is small and the number of studies using a similar type of DOM and comparable pH and Al

concentration ranges as used in the present study is even smaller. To our knowledge, only the study by Gerke [10] is similar enough to compare trends in K_o values. Unfortunately, Gerke does not explicitly list the uncertainties in his K_o values, making a full comparison impossible. However, the uncertainty in general of K_o values of metal binding with DOM in some recent studies using the Scatchard approach but other metals than Al, is in the same order of magnitude as in the present study [9,22]. The uncertainty clearly limits the value of K_o values in speciation calculations. On the other hand in spite of the uncertainty, we observed clear trends in K_o with changing v and pH. Consequently, a comparison of trends of binding affinity based on K_o values is possible.

3.1.3. Trends in K_o values

It would be very interesting to compare the absolute K_o values with those obtained in other studies, in order to assess their accuracy. However, for such an assessment extensive datasets of conditional binding constants

Table 3

Some optimal K_o and $\log K_o$ values and 95% confidence interval upper and lower limits for the experimental data of Scenario I (metal-organic precipitation interacts with solution phase)

	Optimum		Minimum		Maximum	
	$a=0.0033$	$b=0.4721$	$a=0.0013$	$b=0.5898$	$a=0.0058$	$b=0.4450$
pH = 3.5						
v	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$
0.05	128172	5.11	89127	4.95	194568	5.29
0.1	46201	4.66	29609	4.47	71463	4.85
0.2	16653	4.22	9837	3.99	26248	4.42
	$a=0.0015$	$b=0.3556$	$a=0.00074$	$b=0.4513$	$a=0.0025$	$b=0.3194$
pH = 4.0						
v	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$
0.05	30955	4.49	25948	4.41	41810	4.62
0.1	12096	4.08	9489	3.98	16753	4.22
0.2	4727	3.67	3470	3.54	6713	3.83
	$a=0.0011$	$b=0.3374$	$a=0.00038$	$b=0.4288$	$a=0.0018$	$b=0.3353$
pH = 4.5						
v	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$
0.05	20396	4.31	11626	4.07	33618	4.53
0.1	8071	3.91	4318	3.64	13323	4.12
0.2	3194	3.50	1604	3.21	5280	3.72

Table 4

Some optimal K_o and $\log K_o$ values and 95% confidence interval upper and lower limits for the experimental data of Scenario II (metal-organic precipitation assumed to be inert)

	Optimum		Minimum		Maximum	
	$a=0.0032$	$b=0.6096$	$a=0.00084$	$b=0.8664$	$a=0.0062$	$b=0.5784$
pH = 3.5						
v	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$
0.05	242289	5.38	195789	5.29	408673	5.61
0.1	79395	4.90	53697	4.73	136845	5.14
0.2	26017	4.42	14727	4.17	45823	4.66
	$a=0.0011$	$b=0.5431$	$a=0.00029$	$b=0.6155$	$a=0.0023$	$b=0.5095$
pH = 4.0						
v	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$
0.05	60798	4.78	22814	4.36	109715	5.04
0.1	20863	4.32	7445	3.87	38535	4.59
0.2	7159	3.85	2430	3.39	13535	4.13
	$a=0.00060$	$b=0.4765$	$a=n.d.$	$b=n.d.$	$a=0.0017$	$b=0.6494$
pH = 4.5						
v	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$	K_o (L/mol)	$\log K_o$
0.05	23833	4.38	n.d.	n.d.	153211	5.19
0.1	8565	3.93	n.d.	n.d.	48839	4.69
0.2	3078	3.49	n.d.	n.d.	15569	4.19

n.d. = not determined, due to non-convergence of the fit of the minimum MBC values (see text).

obtained by another well-established reference method are needed. In addition, the binding constants in such a dataset must have been derived under exactly the same solution chemical conditions (i.e. same pH, DOC concentrations, Al concentrations, ionic solution strength, etc.), since conditional binding constants per definition depend on solution chemistry. Unfortunately, no such databases exist due to the experimental problems until recently to distinguish between ‘free’ Al in solution and dissolved organic Al complexes. Consequently, it is impossible for us to determine what the ‘true’ K_o values are and to what extent the K_o values obtained by us deviate from these ‘true’ values. However, the study by Gerke [10] used conditions similar enough to ours to allow for a comparison of the order of magnitude of the K_o values derived in this study, and their dependence on pH and extent of binding.

Two clear trends can be observed from the K_o values in Tables 3 and 4. Firstly, for both scenarios the $\log(K_o)$ values decreased with increasing extent of binding. This is not surprising because the functional groups on DOM molecules with the highest binding affinity will fill up first, leading to a lower overall binding constant at higher v . Gerke found the same trend in his study, within the same range of pH and v [10]. Secondly, for both scenarios a decrease in $\log K_o$ values with increasing pH was observed. Again this is in agreement with the results from the study by Gerke, in the same pH and v range [10]. However, at pH values higher than the ones used in our previous study, Gerke found the opposite trend.

A study using Cu and Pb instead of Al at higher pH values also found this opposite trend [9,21,22]. A possible explanation is that there is a small number of strong, non-acidic Al binding sites present on the DOM molecules. Possible candidates are nitrogen or sulfur bearing groups that can have stronger binding affinities for specific metals than the common carboxylic acid groups [33,34]. The abundance of such functional groups is much lower than that of carboxylic acid and other acidic oxygen bearing groups [33]. Therefore, the influence of S and N bearing groups at higher pH values, where many carboxylic acid groups are deprotonated would be marginal. However, at lower pH values, where

a large percentage of the carboxylic acid groups is protonated, the influence of these other groups may become more important. A further decrease in pH would then lead to an increase in K_o values due to the increasing importance of these stronger binding groups as more and more acidic groups become protonated.

Due to the uncertainty in the K_o values calculated for the two scenarios, in most cases there is no significant difference between the K_o values obtained for each scenario within the 95% confidence interval. However, there is a trend of the optimum K_o values for Scenario I being lower than those in Scenario II. As the extent of binding increases, the functional groups with the highest affinity for Al will become saturated with Al first. Consequently, DOM molecules that are relatively enriched in these groups will be the first to precipitate out of solution, leaving the solution relatively deprived of these groups, resulting in a lower K_o value. This effect will only be visible in Scenario II because only here precipitation is assumed to be inert.

3.2. The synthetic dataset

The MBC as calculated for the different levels of uncertainty in the synthetic dataset as well as calculated K_o and $\log(K_o)$ for the same range of v as the experimental dataset are presented in Table 5. Figs. 3 and 4 represent the Scatchard plots and the relation between $\log(K_o)$ values and the extent of binding. The MBC value calculated from the perfect dataset was 54.95 $\text{mmol}_c/\text{mol C}$, while the optimum MBC values calculated for the $\pm 2\%$, $\pm 5\%$ and $\pm 10\%$ noise datasets were, respectively, 56.82, 54.05 and 61.72 $\text{mmol}_c/\text{mol C}$. While these values seem close to the optimum value, for the $\pm 2\%$ and $\pm 10\%$ datasets, the percentage of deviation from the optimum value already exceeds the maximum noise percentage. This, of course, is even more so for the minimum and maximum MBC values. Especially, the maximum MBC value for the $\pm 10\%$ noise dataset is well removed from the original value, contrary to the $\pm 5\%$ and $\pm 2\%$ cases. Recall that such an asymmetry in the upper and lower 95% error boundaries was also found in the MBC values of the real dataset at $\text{pH} = 4.5$ in Scenario II. The

Table 5
95% upper and lower limits for the MBC, K_o and $\log K_o$ values of the synthetic datasets

Random noise level:	$\pm 2\%$ minimum	$\pm 2\%$ maximum	$\pm 5\%$ minimum	$\pm 5\%$ maximum	$\pm 10\%$ minimum	$\pm 10\%$ maximum
MBC ($\mu\text{mol}/\text{mmol C}$)	52.74	60.75	43.63	73.05	44.40	166.1
$K_o(v = 0.05)$	105873	143918	72322	226410	24586	204204
$\log K_o (v = 0.05)$	5.02	5.16	4.86	5.35	4.39	5.31
$K_o(v = 0.1)$	37520	51141	25850	78227	8696	72533
$\log K_o(v = 0.1)$	4.57	4.71	4.41	4.89	3.94	4.86
$K_o(v = 0.2)$	13297	18173	9239	27028	3076	25764
$\log K_o(v = 0.2)$	4.12	4.26	3.97	4.43	3.49	4.41

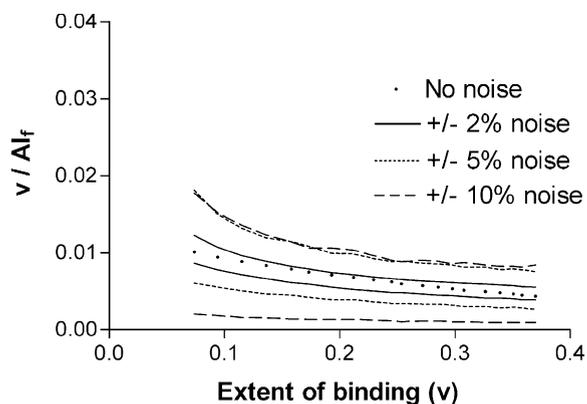


Fig. 3. The 95% confidence interval for Scatchard plots of Al and DOM based on synthetic data for 'free' Al (Al_f) with different levels of random noise added.

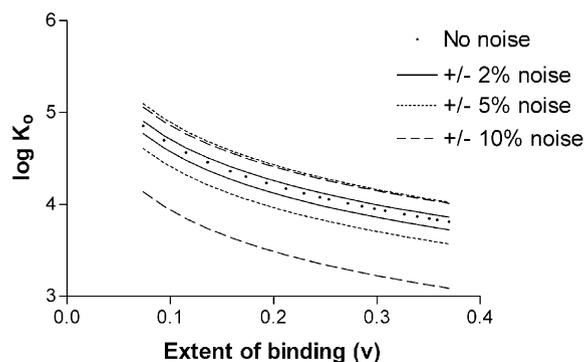


Fig. 4. Plots the 95% confidence interval of the relationship between $\log(K_o)$ and the extent of binding for Al and DOM and based on synthetic data for 'free' Al (Al_f), with different levels of random noise added.

cause for this is the sensitivity of the Langmuir approach used in estimating the MBC values for the maximum values at higher levels of noise, where the slope is close to zero.

From Figs. 3 and 4 and the results in Table 5 it is clear that the uncertainty in the calculated binding constants increases rapidly with increasing levels of noise. Where at $\pm 2\%$ noise the ratio of maximum K_o to minimum K_o is on average 1.4, this ratio increases through an average of 3.0 for $\pm 5\%$ noise to 8.3 for $\pm 10\%$ noise. Fig. 4 also shows that while for $\pm 2\%$ and $\pm 5\%$ the minimum and maximum lines center around the same optimum values for $\log(K_o)$, for the $\pm 10\%$ line, the minimum line deviates much more from the optimal (0% noise) line. This is a direct result of the asymmetric uncertainty boundaries of the MBC values in the $\pm 10\%$ case and again illustrates the sensitivity of the Scatchard approach for errors in the estimation of MBC values. Recall that in the calculations based on the simulated

dataset all but the Al_f data were assumed to be known with infinite certainty. In reality there will be some additional noise in the other data used to construct the plots, causing even higher uncertainty in the calculated K_o .

4. Conclusions

The results from the experimental and simulated datasets clearly show that for moderate, realistic levels of uncertainty in the measured 'free' Al concentrations only, the uncertainty in the calculated K_o is appreciable (see Tables 3–5). Firstly, this is caused by the uncertainty in the estimated MBC, which propagates in the optimal Scatchard plots. Secondly, the uncertainty is caused by the nature of the Scatchard plots. Since K_o values are determined from the slope of the optimal curve, a relatively small uncertainty in the fit of the binding data results in appreciable uncertainty in the obtained stability constants.

When considering natural systems, Scenario II where precipitates are believed to be inert is the most realistic one. Unfortunately, compared to Scenario I, Scenario II is also the scenario with the highest experimental uncertainty. Therefore, the sensitivity of the Scatchard approach for uncertainty in the necessary experimental data, limits the usefulness of the K_o values thus derived for quantitative speciation calculations of Al and DOM. In the light of the availability of more comprehensive speciation models like Model VI and NICCA-Donnan [7,8] that circumvent using K_o values at all, it seems fruitless to attempt to obtain better K_o values for use in quantitative speciation calculations. On the other hand, trends in binding behavior with changes in important soil solution variables such as pH and metal/DOC ratios can be discerned quite well in spite of the uncertainties and the derivation and use of K_o values presents a very simple model for this purpose.

Acknowledgements

We thank Ton van Wijk and Piet Wartenbergh for the chemical analyses they performed.

References

- [1] Sposito G, editor. The environmental chemistry of aluminum. 1 ed. Boca Raton, Florida: CRC Press; 1989.
- [2] Ma FJ, Ryan PR, Delhaize E. Aluminum tolerance in plants and the complexing role of organic acids. *TRENDS Plant Sci* 2001;6:273–8.
- [3] Stevenson FJ. Humus chemistry: genesis, composition, reactions, 2nd ed. New York: Wiley; 1994.

- [4] Tam SH, McColl JG. Aluminum-binding ability of soluble organics in Douglas Fir litter and soil. *Soil Sci Soc Am J* 1991;55:1421–7.
- [5] Jansen B, Nierop KGJ, Verstraten JM. Mobility of Fe(II), Fe(III) and Al in acidic forest soils mediated by dissolved organic matter: influence of solution pH and metal/organic carbon ratios. *Geoderma* 2003;113:323–40.
- [6] Tipping E. Cation binding by humic substances. Cambridge: Cambridge University Press; 2002.
- [7] Tipping E. Humic ion-binding model VI: an improved description of the interactions of protons and metal ions with humic substances. *Aquat Geochem* 1998;4:3–48.
- [8] Kinniburgh DG, Van Riemsdijk WH, Koopal LK, Borkovec M, Benedetti MF, Avena MJ. Ion binding to natural organic matter: competition, heterogeneity, stoichiometry and thermodynamic consistency. *Colloid Surf A* 1999;151:147–66.
- [9] Warwick P, Hall A, King SJ, Zhu J, Van der Lee J. Thermodynamic aspects of nickel humic acid interactions. *Radiochim Acta* 1998;81:215–21.
- [10] Gerke J. Aluminum complexation by humic substances and aluminum species in the soil solution. *Geoderma* 1994;63:165–75.
- [11] Macásek F. Complexation of radionuclides with natural polyelectrolytes—proteins, polysaccharides and humic substances. *J Radioanal Nucl Chem* 1999;240:251–9.
- [12] Browne BA, Driscoll CT. pH-dependent binding of aluminum by a fulvic acid. *Environ Sci Technol* 1993;27:915–22.
- [13] Van Hees PAW, Van Hees AMT, Lundström US. Determination of aluminium complexes of low molecular organic acids in soil solution from forest soils using ultrafiltration. *Soil Biol. Biochem* 2001;33:867–74.
- [14] Pesavento M, Alberti G. Determination of the complexing properties of drinking waters toward copper(II) and aluminum(III) by ligand titration. *Water Res* 2000;34:4482–92.
- [15] Pletnev IV, Zernov V. Classification of metal ions according to their complexing properties: a data-driven approach. *Anal Chim Acta* 2002;455:131–42.
- [16] Glaus MA, Hummel W, Van Loon LR. Trace metal-humate interactions. I. Experimental determination of conditional stability constants. *Appl Geochem* 2000;15:953–73.
- [17] Van Hees P, Tipping E, Lundström US. Aluminum speciation in forest soil solution—modelling the contribution of low molecular weight organic acids. *Sci Total Environ* 2001;278:215–29.
- [18] Temminghoff EJM, Plette ACC, Van Eck R, Van Riemsdijk WH. Determination of the chemical speciation of trace metals in aqueous systems by the Wageningen Donnan Membrane Technique. *Anal Chim Acta* 2000;417:149–57.
- [19] Jansen B, Kotte MC, Van Wijk AJ, Verstraten JM. Comparison of diffusive gradients in thin films and equilibrium dialysis for the determination of Al, Fe(III) and Zn complexed with dissolved organic matter. *Sci Total Environ* 2001;277:45–55.
- [20] Stevenson FJ, Fitch A, Brar MS. Stability constants of Cu(II)–humate complexes: comparison of selected models. *Soil Sci* 1993;155:77–91.
- [21] Logan EM, Pulford ID, Cook GT, Mackenzie AB. Complexation of Cu²⁺ and Pb²⁺ by peat and humic acid. *Eur J Soil Sci* 1997;48:685–96.
- [22] Mizera J, Dolanský J, Beneš P. Radiotracer study of the electrophoretic behavior of Eu and its complexes with humic acid. *Radiochim Acta* 2001;89:529–35.
- [23] Nierop KGJ, Jansen B, Vrugt JA, Verstraten JM. Copper complexation by dissolved organic matter and uncertainty assessment of their stability constants. *Chemosphere* 2002;49:1191–200.
- [24] Jansen B, Nierop KGJ, Verstraten JM. Influence of pH and metal/carbon ratios on soluble organic complexation of Fe(II), Fe(III) and Al in soil solutions determined by diffusive gradients in thin films. *Anal Chim Acta* 2002;454:259–70.
- [25] Evans LJ. Chemistry of metal retention in soils. *Environ Sci Technol* 1989;23:1046–56.
- [26] Vrugt JA, Gupta HV, Bouten W, Sorooshian S. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour Res* 2003; in press.
- [27] Duan Q, Gupta HV, Sorooshian S. A shuffled complex evolution approach for effective and efficient global minimization. *J Optim Theory Appl* 1993;76:501–21.
- [28] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equations of state calculations by fast computing machines. *J Chem Phys* 1953;21:1087–91.
- [29] Price WL. Global optimization algorithms for a CAD workstation. *J Optim Theory Appl* 1987;55:133–46.
- [30] Holland J. Adaptation in natural and artificial systems. Ann Harbor: University of Michigan Press; 1975.
- [31] Zhang H, Davison W. Diffusional characteristics of hydrogels used in DGT and DET techniques. *Anal Chim Acta* 1999;398:329–40.
- [32] Parkhurst DL. User's guide to PHREEQC—A computer program for speciation, reaction-path, advective-transport, and inverse geochemical calculations: US Geological Survey Water-Resources Investigations, 1995. p. 143.
- [33] Martinez CE, McBride MB. Dissolved and labile concentrations of Cd, Cu, Pb and Zn in aged ferrihydrite-organic matter systems. *Environ Sci Technol* 1999;33:745–50.
- [34] Fox TR. The influence of low-molecular-weight organic acids on properties and processes in forest soils. In: McFee WW, Kelly JM, editors. Carbon forms and functions in forest soils. Madison, Wisconsin: Soil Science Society of America; 1995. p. 43–62.