# Derivation of partition relationships to calculate Cd, Cu, Ni, Pb and Zn solubility and activity in soil solutions

# Derivation of partition relationships to calculate Cd, Cu, Ni, Pb and Zn solubility and activity in soil solutions

**P.F.A.M. Römkens**
**J.E. Groenenberg**
**L.T.C. Bonten**
**W. de Vries**
**J. Bril**

**Alterra-report 305**

**Alterra, Wageningen, 2004**

ABSTRACT

P.F.A.M. Römkens, J.E. Groenenberg, L.T.C. Bonten, W. de Vries and J. Bril, 2004. *Derivation of partition relationships to calculate Cd, Cu, Ni, Pb and Zn solubility and activity in soil solution samples.* Wageningen, Alterra, Alterra-report 305. 72 blz. 10 figs.; 12 tables.; 42 refs.

The distribution of heavy metals (Cd, Cu, Ni, Pb and Zn) between soil and soil solution was measured in more than 1400 soil - solution samples. These samples reflect the variability that occurs within the Netherlands and covers all major soil types. The distribution of metals between solid phase (adsorbed) and the soil solution was described using an extended Freundlich equation taking into account the impact of organic matter, clay, pH and Dissolved Organic Carbon. Apart from this, the impact of Fe and Al oxides was tested. The free ionic activities were also calculated from the data and partition relationships were derived for both total dissolved metal contenrations and free metal ion activities. The results show that an extended Freundlich equation is able to represent the wide range of total dissolved concentrations and free ion activities that are encountered in the database. For the metals Pb and Cu, the uncertainty in the predictions is, however, rather high. Further model improvement is required for these metals.

Keywords: partitioning, free metal ion activity, Freundlich, Cd, Cu, Ni, Pb, Zn, solubility, speciation, model

ISSN 1566-7197

This report can be ordered by paying € 23,- to bank account number 36 70 54 612 by name of Alterra Wageningen, IBAN number NL 83 RABO 036 70 54 612, Swift number RABO2u nl. Please refer to Alterra-report 305. This amount is including tax (where applicable) and handling costs.

P.O. Box 47; 6700 AA Wageningen; The Netherlands
Phone: + 31 317 474700; fax: +31 317 419000; e-mail: info.alterra@wur.nl

[Alterra-report 305/May/2004]

# Contents

***Appendices***

# Preface

In polluted soils, heavy metals pose several risks for humanhealth and ecosystems. Uptake of metals by arable crops and natural vegetations, leaching to ground- and surface water as well as impact on soil ecosystems are examples of risks addressed within this context. To assess the magnitude of the risk it is crucial to know exposure pathways, i.e. in what form or through which process are the targets mentioned previously affected by heavy metals. Knowledge of both pathways and processes not only can be used to assess actual risk levels, but can also be used to evaluate the sustainability of land use, or changes therein.

In many cases, risk levels are associated with the chemical and biological availability of metals. The (bio)availability in turn is often related to the solubility and/or the free metal ion activity in the soil solution. In soils, the solubility and speciation or controlled not only by the total heavy metal content but even more so by soil properties like pH and organic matter content. To assess the current and future risk levels it is, therefore, imperative to develop tools that are able to understand the relationship between the soil and the soil solution (and the role of soil properties) in order to estimate for example leaching losses, uptake by crops or exposure to soil micro-organisms.

In this report an approach is presented that enables the user to calculate both the solubility and speciation of heavy metals in soils. One of the main prerequisites was that the resulting approach had to be applicable on a national or international (EU) level. This puts a limitation on the data than can be used to feed the model. Basic soil properties including organic matter, clay content and pH are available throughout Europe whereas, for example, data on the composition of the soil solution are scarce.

The results will be used both on the national and EU level. Examples of applications include:

- Calculation of critical metal loads for Cd and Pb throughout Europe
- Assessment of critical limits soil solution from NOEC data for the soil solid phase and critical limits soil solid phase from limits for ground water and surface water, to be used in ecotoxicological assessments and derivation of Dutch soil quality standards
- Prediction of leaching and off on a national scale to assess the contribution of leaching to the load of heavy metals in surface waters
- Decision support systems: a DSS to evaluate the spreading of dredged sludge on adjacent land (in development) and a DSS to evaluate the possibilities of (contaminated) agricultural land for nature development
- Additional applications in ongoing research projects, among others in the EU-project Phytodec (phytoremediation, www.phytodec.nl)

# Summary

The need for applicable tools to calculate solubility and speciation of heavy metals in soils has increased since it has become clear that risks associated with metals in soils can be linked to their chemical and biological availability. To assess the degree of the chemical availability, the distribution of heavy metals between the solid phase and the soil solution is crucial. Approaches to describe this distribution range from very simple linear Kd models to complex mechanistic models. The goal of this study was to develop a practical approach that can be applied at various scale levels with an emphasis on regional and national applications. Therefore, it was decided to test empirical models that take into account major soil properties.

The use of an extended Freundlich equation, was chosen as the basis for the models tested. It appears that this approach is one of the few that can be applied on a regional or even national scale since the required number of input parameters is limited and can be obtained largely from existing soil information systems. Using data from a large laboratory study (more than 1400 records on soil - solution data), Freundlich-type partition equations were derived by step-wise multiple linear regression. Both models for the prediction of the soil solution concentration as well as the free metal ion activity were derived. To obtain the latter (free ionic activity) it was necessary to calculate the speciation from the composition of the soil solution. This was done using the chemical equilibrium program CHARON. These model calculated free ionic activities were then used to derive partition relationships between solid phase properties and the free metal activity.

The dissolved concentration and the free metal ion activity could be predicted reasonably well by a combination of soil organic matter, pH and clay content and the reactive metal content. Addition of DOC improved the estimates of especially Pb and Cu. The 90% confidence intervals indicated that a considerable ‘noise’ still cannot be explained. The models developed so far, therefore, seem most suitable for the prediction of the dissolved metal concentration for average soil types or high scale levels starting at the field or even regional level. For the prediction of the solubility and activity at a single location, more data need to be included so as to reduce the error margins and obtain a more certain estimate. Especially for Pb error levels are still considerable whereas for Zn, and to a lesser extent also Cd and Ni good estimates are obtained with an acceptable error around the predicted values.

Addition of other soil properties like amorphous Fe and Al oxides, CEC, Ca concentration in the soil solution resulted in an improvement of the model for the prediction of the dissolved concentration. Although $R^2$ values and standard error levels of the estimates indicated that the addition resulted in a reduction of the noise, the improvement was limited. $R^2$ increased with 5% at most (for Cd), whereas se(Y-est) values decreased up to 0.04 (log scale). In most cases the effect of addition of Fe- and Al oxide resulted in a decrease of the significance of the clay fraction. Since national and Europe wide data of Fe- and Al oxides as well as EC, and Ca in solution

are not readily available it was concluded that the model using organic matter, clay and pH is to be used. In those cases where data on dissolved organic carbon are available, these should be used. For the calculation of the free metal ion activity the added value of Fe- and Al-oxides was even less than the one for concentration.

Apart from partition equations to calculate the solubility and speciation of metals in soils, also empirical relationships are included that are able to obtain estimates of the 'reactive' metal content of Cd, Cu, Pb and Zn from the so-called total content (appendix 1). This is essential since part of the heavy metal content in soils is present in non-reactive forms (included in soil minerals and or organic matter. This part will not contribute to the chemical or biological availability in soils. Here, a dilute acid (0.43 mol.$l^{-1}$ $HNO_3$) was chosen as a measure of the reactive content. Multiple linear regression showed that this fraction can be predicted well by a combination of the Aqua Regia extractable metal content and the soil organic matter- and clay content.

# 1 Introduction

***Need for tools to calculate metal solubility***

Heavy metal levels in soils have increased significantly during the last centuries due to human activities. As a result of mining activities, traffic, application of sludge or waste and agricultural activities, total metal levels in the soil have increased. Whether or not these increased levels will lead to higher risk levels (e.g. uptake by plants or animals, leaching to groundwater etc.) mainly depends on the chemical conditions in soils.

The need for applicable tools to calculate the dissolved concentration and or free ion activity in the soil solution has increased. Currently soil quality criteria are still almost exclusively based on the total metal content but ongoing research has shown that in many cases risks related to metals in soils are strongly linked to the metal concentration in the soil solution. Both leaching to the groundwater and uptake by crops (either agricultural or in natural vegetations) are controlled largely by the chemical speciation of metals in the soil solution. Also effects on soil organisms are believed to be controlled (at least partially) by the dissolved concentration although the amount in the solid phase is important as well for higher organisms (e.g. earthworms). To asses risk levels in soils, therefore, it is necessary to be able to quantify the total dissolved metal concentration and/or free metal ion activity that mostly control the risk level. In addition to this, it is imperative to know which parameters control both solubility and speciation.

***Adsorption studies for heavy metals***

Numerous adsorption studies have revealed the importance of soil properties like clay, organic matter, pH, dissolved organic carbon and ionic strength. With an increase in both clay and organic matter, the capacity of the soil to retain metals increases (e.g. McBride, 1994). Similarly, with an increase in pH, metal retention for cationic metals like Cd, Zn, Cu and Pb also increases (Selim and Sparks, 2001). Important solution properties that also affect metal solubility and speciation are ionic strength (or composition of the soil solution) and the presence of dissolved organic carbon. With an increase in the concentration divalent cations like Ca, adsorption of especially Cd and Zn decreases due to cation exchange reactions. On the other hand, dissolved organic carbon increases the concentration of Cu and Pb due to the formation of stable metal-organic complexes.

Various studies have dealt with one or more of these properties or processes but few have investigated the simultaneous impact of all properties. In addition to this, laboratory studies often were designed to study the process of *ad*sorption. This does not necessarily reflect the behaviour of metals already stored in the soils. In many cases, concentration ranges applied exceed those encountered in the field situation by orders of magnitude. This is not surprising because the impact of low dissolved metal concentrations on the adsorbed pool is very small and in order to measure a significant increase in the adsorbed pool, large additions are necessary. As a result of this approach, the information obtained in these kind of studies cannot be used directly to develop or calibrate models that predict the behaviour of metals in field situations. Furthermore studies are often limited to one or a few number of metals. In this study we included five important metals (Cd, Cu, Ni, Pb and Zn).

### *Metal behaviour models*

The solubility and speciation of metals in soils can be calculated in many ways ranging from a simple constant distribution coefficient to complex mechanistic (non)equilibrium models. During the last three decades numerous models have thus been proposed, developed and tested that describe the relation between the amount of metals in the solid phase and that in the soil solution. The majority of these models is based on results from *ad*sorption experiments which were performed under controlled conditions. Results have been used to calibrate both empirical and mechanistic model formulations

Mechanistic models that have been developed to describe various processes like complexation, adsorption on specific sites or planes (high affinity, low affinity, inner-sphere, outer-sphere) are often calibrated for a few specific soils or sites and thus not generally applicable. Furthermore, they usually require a rather large number of (soil-specific) input parameters to be used in a wide range of soils extensive. Although our understanding of the mechanistic behaviour of metal sorption has increased considerably as a result of these kind of studies the applicability of mechanistic models, i.e. model formulations based on process knowledge, to describe the solubility of metals in field conditions is still restricted to well characterised sites or samples used in laboratory studies.

This is due not only to the variability and heterogeneity of soil properties and metals levels in soils but also to the large number of input parameters that are required to run mechanistic models. On a regional, national or even international scale, such data are scarce, or even non-existent. This means that a different type of model approach is required to estimate risks at large scale levels.

### *Need for an alternative approach*

Although detailed mechanistic models are capable of predicting the soil - solution distribution of metals in well defined (laboratory) systems quite accurately they still require a rather large number of parameters that are simply not available under field conditions, such as surface charge properties of soils and soil components. Therefore, models based on ‘commonly’ available soil properties such as pH, soil organic matter and clay have to be used until either data availability increases to feed more mechanistic models or mechanistic models are improved to such an extent that they can be used with a limited number of parameters as well.

Both the emphasis on adsorption processes in laboratory studies and the rather soil-specific mechanistic models stress the need for an alternative, applicable approach that could be used:

1. in a wide range of soils without the need for immediate recalibration, and
2. based on a limited number of input parameters to be applicable on a regional or on a (inter)national scale.

One approach that could fulfil these conditions is the partition approach. In this approach the concentration in the soil solution can be calculated taking into account soil properties. Usually an empirical equation like a Freundlich or Langmuir equation is used to relate these properties to either the dissolved total concentration or the free metal ion activity.

This approach was tested on data from Dutch soils (Otte et al., 2000) using field data only. Although results showed that the approach was able to account for a considerable amount of the measured variation in dissolved and total metal concentrations, it became clear that results from different field studies were difficult to use in a single study.

Differences in sampling techniques and destruction methods resulted in a large amount of 'noise' in the database. Furthermore, the distribution of the data were strongly biased: a large part of the data originated from forest soils with low to very low metal contents (total) and a limited range in pH values.

### *Aim of this study*

In this study we derive and evaluate partition relationships that can be used to describe the behaviour of five important metals (Cd, Cu, Ni, Pb and Zn) by accounting for the influence of important soil properties on the solubility and speciation in soils, like organic matter, clay content and pH. A rather simple model formulation (in this case a scaled Freundlich equation) is used to relate the measured concentrations or model calculated free metal activity in the soil solution to the aforementioned soil properties in combination with the soil metal content itself.

The goal of the study was to derive both a 'minimal' model which is based on basic soil properties only (organic matter, clay, pH and soil metal content) as well as an 'extended' model that uses other soil properties as well like base cation concentration, dissolved organic carbon, and the amorphous Fe and Al content. The reason for doing this is that we wanted to be able to describe metal solubility both on a national/international scale level as well as application on a small scale. In both cases, data availability will be different, and the use of a simple model is probably the only way to achieve an estimate at all. When more data are available, more complex models can and should be used to obtain a more reliable estimate.

To overcome the limitations of samples from field studies (not representative for contaminated soils, different analytical methods used), a laboratory study was performed. In this study a large number of soils samples from both clean and polluted soils were subject to different extractions using dilute salt solutions. It is assumed that this extraction (see appendix A for a more detailed description of the methods used) mimics conditions that prevail under field conditions. In total a database containing more than 1400 records (each record contains data from an extract performed on a soil sample) was created representing Dutch soils (including all major soil types). Since Dutch soils have been developed in unconsolidated sediments only (sand, clay. loess and peat deposits) soils developed from solid parent materials (rocks) are not included.

The results presented here are based on data from well-aerated soils, and models derived from these data therefore are applicable to terrestrial systems only. In systems where reducing conditions prevail, solubility and speciation of metals is not necessarily controlled by adsorption processes anymore. Models that take into account sorption as the main process such as the ones described here cannot be applied to those systems.

### *Outline of the report*

In Chapter 2 the materials and methods are described that were used to get data on metal concentrations in both the solid phase and solution phase and major soil properties. Furthermore, models used in this study to predict both the total dissolved metal concentration and the free metal ion activity are explained in general terms.
In the Chapters 3 and 4 the results from the various regression analyses to calculate the total dissolved metal concentration (Chapter 3) and free metal ion activity (Chapter 4) are presented, followed by a discussion on the applicability and usefulness of the equations described and calibrated in this report.
In Chapter 5 the results from this study are compared with available data from the literature.
Finally, in chapter 6 the major conclusions of this study are presented together with a discussion on the applicability of the partition relationships.
In addition to this report, which focuses on the 'Dutch' database, a comparison between the results obtained here and recent data from other comparable studies from other countries is made. Results from this comparison will be published separately (Groenenberg et al., 2003 and 2004)

In addition to the chapters, 3 appendices are included with information on the:

1. Estimation of the reactive metal content from soil properties and the total (aqua regia extracted) metal content;
2. Estimation of DOC from organic matter and pH;
3. Evaluation of the effect of analytical detection limits on the magnitude of the coefficients of the models tested here.

# 2 Experimental approach and data evaluation

## 2.1 Materials and methods

***Origin of soils***

The results shown in this report are based on two large datasets. One set contains 49 soil samples from the topsoil of various soil types and degree of contamination (De Groot et al., 1998) ranging from non-contaminated sandy soils with low clay and organic matter contents to heavy clay and peat soils. The second set is based on soil samples from 11 locations. In this set samples were taken not only from the topsoil but from all diagnostic soil horizons down to 120 cm. which resulted in 69 different soil samples.

All major soil types from the Netherlands were included such as heavy clay soils (both river- and sea clay soils), peat soils, sandy soils, and loess soils from the southern parts of the Netherlands. Apart from soils that are still used for agricultural purposes (both arable land and pasture), also soils form natural areas (forest, heathland) were included. Both non-polluted and polluted soils were included. Examples of polluted soils include soils from areas with elevated Cd and Zn deposition from smelters (Budel area, province of Brabant), peat soils enriched with waste material (so-called 'toemaakdek' soils from the central western peat area, Province of South-Holland) and soils along rivers with high Pb and Zn contents due to sedimentation of polluted river sediments (river Geul, Province of Limburg). In Table 2.1 an overview of the soil types and location is given.

***Soil sampling and soil analyses***

Soil samples were obtained by collecting field moist soil from the topsoil (dataset 1) and the entire soil profile (dataset 2). The samples from dataset 1 were collected from 50 x 50 cm plots (3 plots from each site). From each 50 x 50 cm plot approximately 10 kg was collected and mixed with the soil from the other 2 plots to obtain a single mixed sample.

The samples from dataset 2 were collected by using a soil auger. From each diagnostic horizon, soil material was sampled from at least 3 to 4 locations at a given site. The material from identical diagnostic horizons was mixed to obtain at least 2 kg of soil material from each site and horizon. To identify the diagnostic horizons an entire profile was dug first before deciding how many horizons were present at a specific site.

*Table 2.1 Overview of soil type, profile, land use and locations of samples in dataset 2*

| Soil type | Land Use | note | sampling depth (cm) | |
|---|---|---|---|---|
| | | | from | to |
| Peat | horticulture | experimental station Boskoop | 0 | 30 |
| Sand (podzol) | grassland | near Cd/Zn smelter “Budel” | 0 | 30 |
| | | prov. Noord-Brabant | 30 | 45 |
| | | | 50 | 70 |
| | | | 70 | 120 |
| Sand (podzol) | grassland | near Cd/Zn smelter “Budel” | 0 | 30 |
| | | prov. Noord-Brabant | 30 | 45 |
| | | | 50 | 70 |
| | | | 70 | 120 |
| Loam | grassland | nature reserve (Geuldal) | 0 | 30 |
| | | next to river | 30 | 60 |
| | | next to Pb, Zn mine | 60 | 90 |
| | | prov. Limburg | 90 | 120 |
| Silt loam | grassland | nature reserve (Geuldal) | 0 | 30 |
| | | approx. 100 m from river | 30 | 60 |
| | | prov. Limburg | 60 | 90 |
| | | | 90 | 120 |
| Sand (podzol) | arable | experimental station IB-DLO Haren | 0 | 30 |
| | wheat | prov. Groningen | 30 | 55 |
| | | | 55 | 90 |
| | | | 90 | 120 |
| Sand (podzol) | arable | experimental station IB-DLO Haren | 0 | 40 |
| | wheat | prov. Groningen | 40 | 65 |
| | | | 65 | 75 |
| | | | 75 | 120 |
| Sand (mollic) | arable | experimental station Horst | 0 | 40 |
| | | prov. Limburg | 40 | 60 |
| | | | 60 | 90 |
| | | | 90 | 120 |
| Sand (mollic) | fruit | experimental station Horst | 0 | 40 |
| | | prov. Limburg | 40 | 60 |
| | | | 60 | 90 |
| | | | 90 | 120 |
| Sand (no profile) | Forest | near smelter but excavated | 0 | 10 |
| | | prov. Gelderland | 10 | 30 |
| Sand (no profile) | heath | near smelter but excavated | 0 | 10 |
| | | prov. Gelderland | 10 | 40 |
| Clayey peat | grassland | prov. Zuid-Holland | 0 | 15 |
| | | | 15 | 30 |
| | | | 30 | 50 |
| | | | 50 | 70 |
| (river) Clay | arable | experimental station AB-DLO | 0 | 40 |
| | potato | prov. Gelderland | 40 | 80 |
| | | | 80 | 120 |
| (river) Clay | arable | experimental station AB-DLO | 0 | 40 |
| | potato | prov. Gelderland | 40 | 80 |
| | | | 80 | 120 |
| Silty loam (“Loess”) | arable | Experimental station Wijnandsrade | 0 | 30 |
| | wheat | prov. Limburg | 30 | 60 |
| | | | 60 | 90 |

| Soil type | Land Use | note | sampling depth (cm) | |
|---|---|---|---|---|
| | | | 90 | 120 |
| Silty loam ("Loess") | arable | Experimental station Wijnandsrade | 0 | 40 |
| | wheat | prov. Limburg | 40 | 60 |
| | | | 60 | 90 |
| | | | 90 | 120 |
| Clayey peat | public garden | next to railway | 0 | 15 |
| | | prov. Zuid-Holland | 15 | 30 |
| | | | 30 | 50 |
| Clayey peat | public garden | next to railway | 0 | 15 |
| | | prov. Zuid-Holland | 15 | 30 |
| | | | 30 | 50 |
| Peat | grassland | experimental station Zegveld | 0 | 10 |
| | | prov. Zuid-Holland | 10 | 40 |
| | | | 40 | 60 |

After transport to the laboratory, the soil material was dried at 40 degrees centigrade for 40 tot 72 hours depending on texture and organic matter content. Finally all soil samples were sieved on a contaminant-free 2 mm sieve. The samples from the clay and peat soils were crushed before sieving on a platinum-coated mechanical crusher to obtain soil material that was able to pass the 2 mm sieve.

In the dried and sieved soil samples the following general soil parameters were determined:

- pH KCl (1 mol.$l^{-1}$ KCl)
- pH $H_2O$
- pH $CaCl_2$
- dry matter content after drying at 105 degrees
- CEC un-buffered ($BaCl_2$ extraction) and base saturation
- CEC buffered (Bascomb) at pH 8.1
- Fe and Al in ammonium-oxalate (also P-ox was measured)
- Total carbon content by loss on ignition
- Clay content (gravimetric methods)

***Heavy metal analysis***

Several extractions were performed to obtain both the 'total' and 'reactive' heavy metal content of the soil samples. As of now only operational definitions exist on what the reactive fraction in soil is and how it should be analysed:

Aqua Regia extraction (total or pseudo - total, $HCl:HNO_3 = 3:1$)

- 2 mol.$l^{-1}$ $HNO_3$ extraction ('pseudo'- total, to compare with Aqua Regia)
- 0.43 mol.$l^{-1}$ $HNO_3$ extraction (reactive heavy metal content, option 1)
- 0.05 mol.$l^{-1}$ EDTA extraction (reactive heavy metal content, option 2)
- 0.05 mol.$l^{-1}$ Ca-EDTA extraction (more readily available fraction): experimental approach.

For these extractions standard methods of the chemical laboratory of Alterra have been used. For further questions concerning the methods please contact P. Römkens at paul.romkens@wur.nl

To account for the fact that part of the metals in the soil are not involved in chemical reactions (the so-called 'inert' fraction, i.e. metals occluded in clay minerals) we used the 0.43 mol.l$^{-1}$ $HNO_3$ extractable metal content as an input for the soil metal content. However, in many cases this analysis is not available. To overcome this limitation, a relationship between the total (Aqua Regia) extractable metal content and the 0.43 mol.l$^{-1}$ $HNO_3$ extractable metal content was derived based on data from the two datasets mentioned above as well as two other databases that contain data on both AR and 0.43 mol.l$^{-1}$ $HNO_3$ (Wiersma et al., 1986; Van Driel et al., 1988). More information is given in Appendix 1.

### ***Extraction and analysis of soil solution extracts***

The samples from both sets were equilibrated (1:2 soil solution ratio) with solutions including de-mineralised water (set 1 only), 0.002 mol.l$^{-1}$$CaCl_2$ (set 1) or $Ca(NO_3)_2$ (set 1 and 2), 0.01 mol.l$^{-1}$ $CaCl_2$ or $Ca(NO_3)_2$ (set 1 and 2), 0.03 mol.l$^{-1}$ $NaNO_3$ (set 1) and 0.1 mol.l$^{-1}$ $CaCl_2$ (set 1). Twenty (20) gram of soil material and 40 ml of the appropriate solution were put in polypropylene centrifuge tubes and closed with a screw-cap.

The set containing 49 soil samples were extracted at the natural soil pH only. The samples from set 2 were extracted (in addition to the natural pH) at two additional pH levels, approximately 0.5 to 1 and 1.5 to 2 pH units lower than the initial soil pH. To obtain this, small amounts of nitric acid (at the same ionic strength as the background electrolyte, i.e. 0.002 or 0.01 N) were added to the soil-solution mixture in the centrifuge tubes. The total volume of solution, however, was always kept constant at 40 ml, so a correction for the acid added on the total volume was taken into account.

The samples were then shaken continuously in an end-over-end shaker to ensure complete contact between soil and solution (approx. 30 rpm). After equilibration (shaking) for 48 hours, the solutions were allowed to settle in the centrifuge tubes. After 1 hour, the pH was measured in the tube directly (before filtration). After measurement of pH, the supernatant-soil mixture was centrifuged at 3000 rpm followed by filtration on a 0.45 μm filter using syringes and screw filters. The solutions were stored at 4 ºC until further analysis.

Total carbon (TC) and inorganic carbon (IC) in solution were determined on a Shimadzu TC-5000. Dissolved organic carbon (DOC) was calculated as the difference between TC and IC. Usually measurement of TC and IC were performed within 3 to 4 days after extraction to avoid microbial decay. EC (electrical conductivity was measured in the filtered supernatant. Total cations and anions (P and S) were measured by ICP-AES. Since the shaking solutions contained either nitrate or chlorine at elevated concentration levels, these elements were not measured in the final extracts. The soil extracts thus obtained were not acidified previous to analysis since this might affect the amount of metals in the extract due to flocculation of organic material and metals attached to it.

### *Correction of soil solution samples below detection limit*

For Cd and Pb, and to a lesser extent also for Ni, a substantial part of the dissolved total metal concentrations was below the detection limit, especially in soils with pH levels higher than 7.0 and clay soils. By skipping these data, the band width of for example pH would decrease which would limit the application of the model in the high pH range (or for soils with a high clay content). This is illustrated for Pb and Cd in Table 2.2 where the distribution of pH, clay, organic matter and metal content is shown for all data (including those below detection limit) and the database corrected for detection limits. For Cu and Zn this problem was less urgent (total number of samples below detection was 10 for Zn and 55 for Cu on a total of 1466).

To correct values below the detection limit, the following procedure was applied:
All data below detection were selected and sorted on (measured) value. This results in an array with values ranging from the minimum value to the detection limit.
The lowest reading was set at a default minimum value (0.01 $\mu g.l^{-1}$)
All values between this default minimum value and the detection limit were re-scaled using a quadratic fit such that the recalibrated values ranged between 0.01 and the detection limit. The original readings were used as X-values to obtained rescaled Y-values.

The idea behind this action was that levels below detection still have a certain value and that simply skipping these values results in a biased dataset where low values have been left out deliberately. However it is clear that the absolute value of this rescaled samples should be interpreted with caution and that predictions in this range, i.e. predictions that are lower than the detection limit should be treated with caution as well.

To evaluate whether or not the use of rescaled data results in significantly different isotherms and a difference in the quality of the predictions, also isotherms based on the dataset without samples below the detection limit have been derived. The results of this exercise are discussed separately in Appendix 2.

*Table 2.2 Distribution of soil properties and concentration in the soil solution: Overview of effect of detection limit on frequency distributions*

| Cd | (n > det: 863) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | pH | | OM (%) | | clay (%) | | DOC (mg.l$^{-1}$) | | Q (mg.kg$^{-1}$) | | C (µg.l$^{-1}$) | |
| Perc. | > det | all | > det | all | > det | all | > det | all | > det | all | > det | all |
| 5 | 2.45 | 2.87 | 0.5 | 0.5 | 0.8 | 1.0 | 5.8 | 4.1 | 0.0 | 0.0 | 1.3 | 0.1 |
| 10 | 3.00 | 3.36 | 0.8 | 0.9 | 1.0 | 1.0 | 8.9 | 6.1 | 0.1 | 0.1 | 1.5 | 0.2 |
| 25 | 3.59 | 3.94 | 1.9 | 1.9 | 2.0 | 2.0 | 17.4 | 12.3 | 0.2 | 0.3 | 2.5 | 0.5 |
| 50 | 4.24 | 4.80 | 4.3 | 3.9 | 3.0 | 5.8 | 33.1 | 26.3 | 0.6 | 1.9 | 6.2 | 1.8 |
| 75 | 5.10 | 6.16 | 6.6 | 7.6 | 11.0 | 19.1 | 66.8 | 63.1 | 4.3 | 6.1 | 28.6 | 9.3 |
| 90 | 6.30 | 7.40 | 13.8 | 26.9 | 21.5 | 36.3 | 128.2 | 167.9 | 10.7 | 10.4 | 97.5 | 55.1 |
| 95 | 7.10 | 7.60 | 35.5 | 37.2 | 29.4 | 46.8 | 251.2 | 293.4 | 13.8 | 13.2 | 190.4 | 112.4 |
| 100 | 7.90 | 7.90 | 45.7 | 74.1 | 51.3 | 55.0 | 1380.4 | 1380.4 | 20.3 | 20.3 | 1073.4 | 1073.4 |
| Zn | (n > det 1458) | | | | | | | | | | | |
| | pH | | OM (%) | | clay (%) | | DOC (mg.l$^{-1}$) | | Q (mg.kg$^{-1}$) | | C (µg.l$^{-1}$) | |
| Perc. | > det | all | > det | all | > det | all | > det | all | > det | all | > det | all |
| 5 | 2.87 | 2.87 | 0.5 | 0.5 | 1.0 | 1.0 | 4.1 | 4.1 | 1.1 | 1.2 | 10.6 | 9.0 |
| 10 | 3.35 | 3.36 | 0.9 | 0.9 | 1.0 | 1.0 | 6.1 | 6.1 | 1.6 | 1.6 | 20.2 | 19.7 |
| 25 | 3.93 | 3.94 | 1.9 | 1.9 | 2.0 | 2.0 | 12.3 | 12.3 | 5.6 | 5.6 | 47.4 | 46.3 |
| 50 | 4.80 | 4.80 | 4.2 | 3.9 | 5.8 | 5.8 | 26.3 | 26.3 | 22.7 | 22.7 | 278.9 | 278.9 |
| 75 | 6.15 | 6.16 | 7.6 | 7.6 | 18.2 | 19.1 | 63.1 | 63.1 | 71.7 | 70.1 | 1297.3 | 1246.0 |
| 90 | 7.40 | 7.40 | 26.9 | 26.9 | 36.3 | 36.3 | 169.8 | 167.9 | 376.3 | 376.3 | 8100.9 | 8044.7 |
| 95 | 7.60 | 7.60 | 37.2 | 37.2 | 46.8 | 46.8 | 296.2 | 293.4 | 1465.3 | 1275.0 | 22752.3 | 22294.8 |
| 100 | 7.90 | 7.90 | 74.1 | 74.1 | 55.0 | 55.0 | 1380.4 | 1380.4 | 9671.9 | 9671.9 | 351165.1 | 351165.1 |
| Cu | (n > det 1421) | | | | | | | | | | | |
| | pH | | OM (%) | | clay (%) | | DOC (mg.l$^{-1}$) | | Q (mg.kg$^{-1}$) | | C (µg.l$^{-1}$) | |
| Perc. | > det | all | > det | all | > det | all | > det | all | > det | all | > det | all |
| 5 | 2.86 | 2.87 | 0.5 | 0.5 | 1.0 | 1.0 | 4.1 | 4.1 | 0.4 | 0.4 | 1.4 | 1.3 |
| 10 | 3.35 | 3.36 | 0.9 | 0.9 | 1.0 | 1.0 | 6.3 | 6.1 | 0.8 | 0.8 | 1.9 | 1.6 |
| 25 | 3.93 | 3.94 | 1.9 | 1.9 | 2.0 | 2.0 | 12.6 | 12.3 | 2.1 | 2.1 | 5.0 | 4.5 |
| 50 | 4.80 | 4.80 | 4.3 | 3.9 | 5.8 | 5.8 | 26.9 | 26.3 | 5.8 | 5.7 | 10.8 | 10.3 |
| 75 | 6.13 | 6.16 | 8.3 | 7.6 | 18.2 | 19.1 | 64.6 | 63.1 | 14.2 | 14.2 | 25.9 | 24.2 |
| 90 | 7.40 | 7.40 | 26.9 | 26.9 | 36.3 | 36.3 | 173.8 | 167.9 | 29.0 | 29.0 | 74.7 | 73.0 |
| 95 | 7.60 | 7.60 | 37.2 | 37.2 | 46.8 | 46.8 | 309.0 | 293.4 | 40.1 | 40.1 | 196.4 | 177.1 |
| 100 | 7.90 | 7.90 | 74.1 | 74.1 | 55.0 | 55.0 | 1380.4 | 1380.4 | 304.2 | 304.2 | 17910.8 | 17910.8 |
| Pb | (n > det 535) | | | | | | | | | | | |
| | pH | | OM (%) | | clay (%) | | DOC (mg.l$^{-1}$) | | Q (mg.kg$^{-1}$) | | C (µg.l$^{-1}$) | |
| Perc. | > det | all | > det | all | > det | all | > det | all | > det | all | > det | all |
| 5 | 2.17 | 2.87 | 0.5 | 0.5 | 0.5 | 1.0 | 5.4 | 4.1 | 6.1 | 0.9 | 15.0 | 0.8 |
| 10 | 2.66 | 3.36 | 0.6 | 0.9 | 0.8 | 1.0 | 8.9 | 6.1 | 9.5 | 2.3 | 16.8 | 1.5 |
| 25 | 3.50 | 3.94 | 2.1 | 1.9 | 1.0 | 2.0 | 18.0 | 12.3 | 19.1 | 6.1 | 29.6 | 3.8 |
| 50 | 4.30 | 4.80 | 4.8 | 3.9 | 2.3 | 5.8 | 39.8 | 26.3 | 41.3 | 23.8 | 77.0 | 9.7 |
| 75 | 5.39 | 6.16 | 7.4 | 7.6 | 8.9 | 19.1 | 87.1 | 63.1 | 103.8 | 55.8 | 303.0 | 38.4 |
| 90 | 6.87 | 7.40 | 15.0 | 26.9 | 15.8 | 36.3 | 218.8 | 167.9 | 413.4 | 122.0 | 1707.7 | 237.9 |
| 95 | 7.40 | 7.60 | 36.3 | 37.2 | 27.5 | 46.8 | 323.6 | 293.4 | 806.1 | 366.4 | 3330.3 | 1087.4 |
| 100 | 7.90 | 7.90 | 74.1 | 74.1 | 55.0 | 55.0 | 1380.4 | 1380.4 | 1571.8 | 1571.8 | 16083.8 | 16083.8 |
| Ni | (n > det 1153) | | | | | | | | | | | |
| | pH | | OM (%) | | clay (%) | | DOC (mg.l$^{-1}$) | | Q (mg.kg$^{-1}$) | | C (µg.l$^{-1}$) | |
| Perc. | > det | all | > det | all | > det | all | > det | all | > det | all | > det | all |
| 5 | 2.70 | 2.87 | 0.5 | 0.5 | 1.0 | 1.0 | 5.8 | 4.1 | 0.0 | 0.0 | 4.3 | 1.1 |
| 10 | 3.17 | 3.36 | 0.8 | 0.9 | 1.0 | 1.0 | 8.9 | 6.1 | 0.1 | 0.1 | 5.6 | 1.6 |
| 25 | 3.72 | 3.94 | 1.8 | 1.9 | 2.0 | 2.0 | 17.8 | 12.3 | 0.2 | 0.3 | 9.3 | 4.3 |
| 50 | 4.45 | 4.80 | 4.4 | 3.9 | 4.0 | 5.8 | 34.7 | 26.3 | 1.5 | 1.9 | 18.1 | 13.6 |
| 75 | 5.30 | 6.16 | 9.5 | 7.6 | 19.1 | 19.1 | 83.2 | 63.1 | 8.1 | 6.1 | 41.5 | 33.0 |
| 90 | 6.18 | 7.40 | 35.5 | 26.9 | 39.8 | 36.3 | 204.2 | 167.9 | 10.9 | 10.4 | 97.4 | 72.2 |
| 95 | 7.00 | 7.60 | 45.7 | 37.2 | 46.8 | 46.8 | 338.8 | 293.4 | 13.8 | 13.4 | 152.3 | 127.0 |
| 100 | 7.90 | 7.90 | 74.1 | 74.1 | 55.0 | 55.0 | 1380.4 | 1380.4 | 20.3 | 20.3 | 2561.9 | 2561.9 |

### *Database characteristics*

In total the final database contains 1466 complete records with both solid phase and solution characteristics. These have been used to derive the partition equations shown in this report.

In Figures 2.1 - 2.3 frequency distributions and minimum, median, average, and maximum are shown of several important soil properties (soil organic matter, clay content and measured pH; Figure 2.1), reactive metal content of the soils (0.43 mol.l$^{-1}$ $HNO_3$ acid extracts; Figure 2.2) and total dissolved measured concentration (filtered through 0.45 µm filters; Figures 2.3).

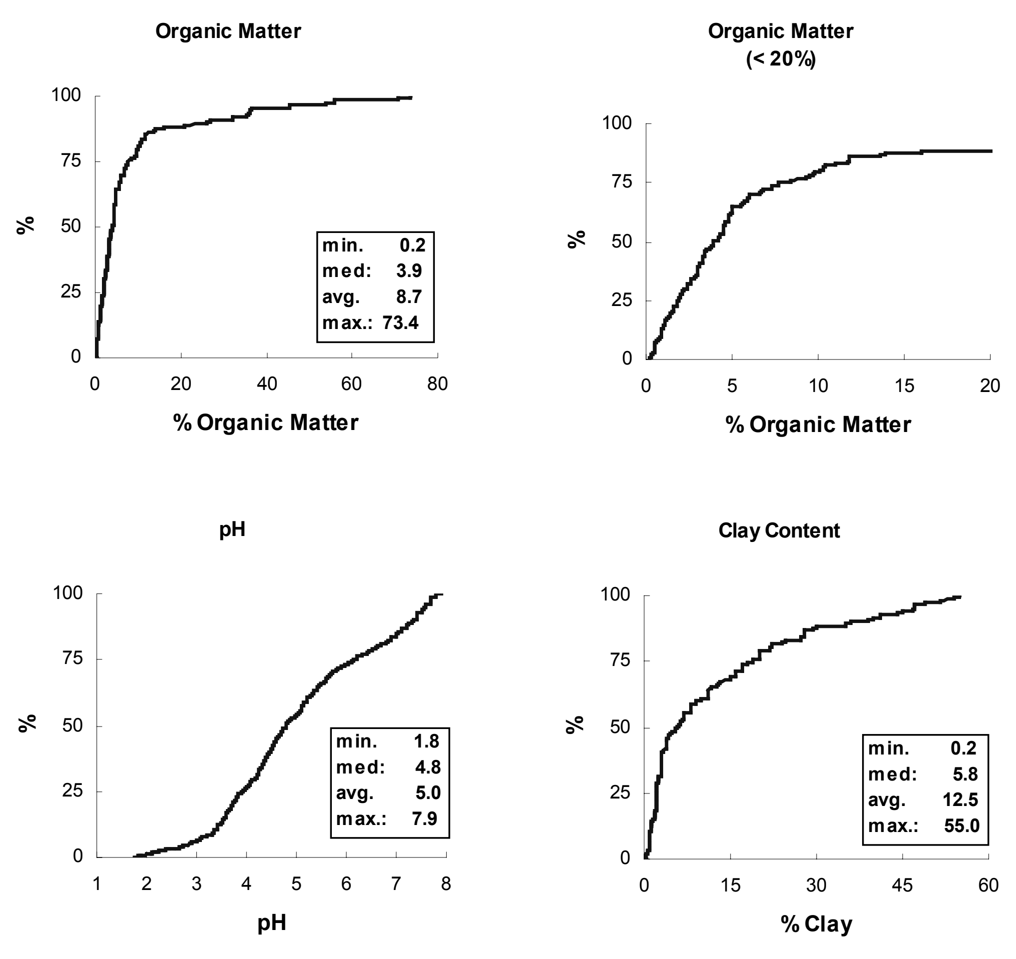

*Figure 2.1 Frequency distribution of the soil organic matter content, pH (extract) and clay content of the soil samples used in the extractions*

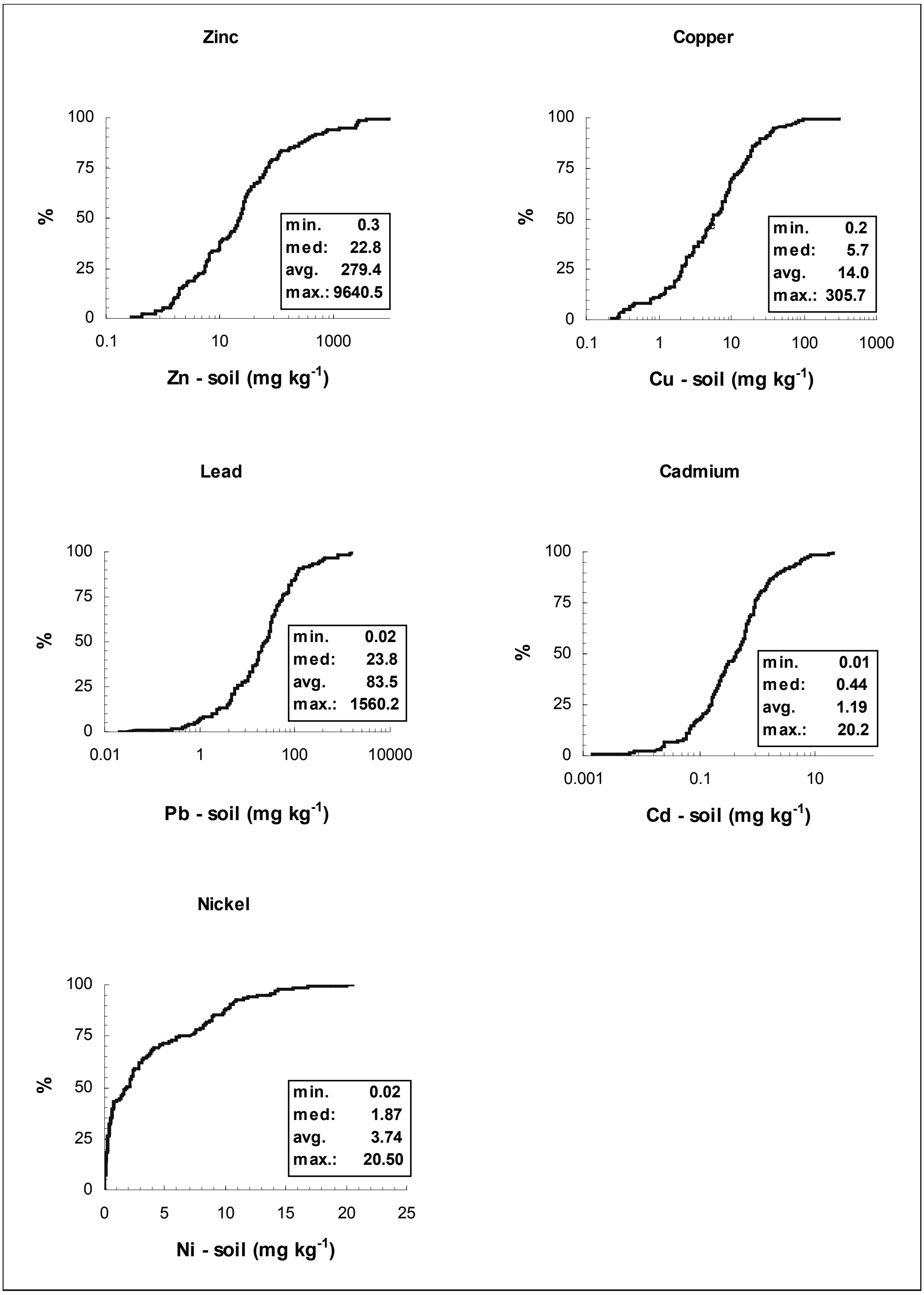

*Figure 2.2 Frequency distribution of the reactive metal content (in mg.kg$^{-1}$) in the soil samples used in the extractions*

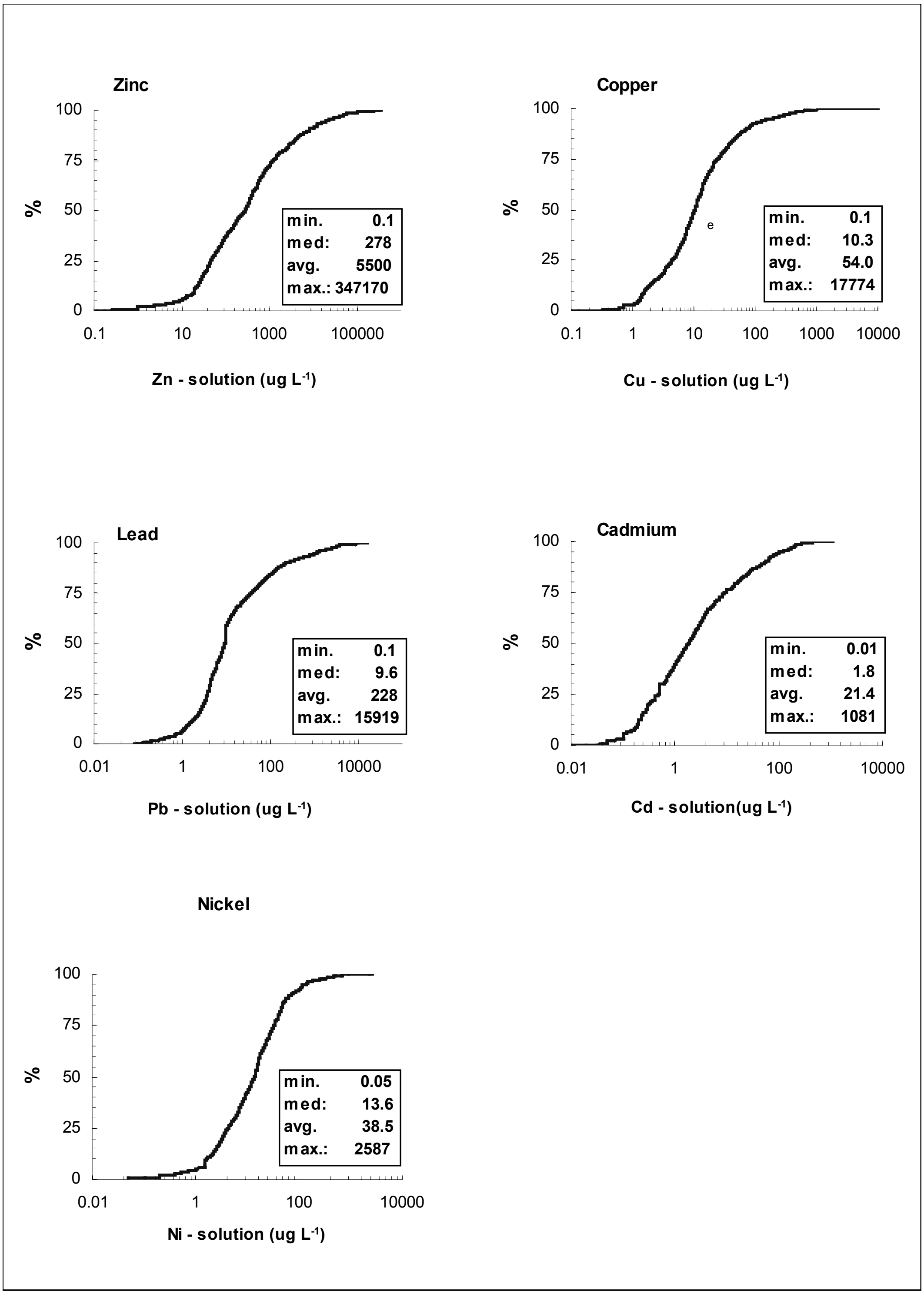

*Figure 2.3 Frequency distribution of the total dissolved metal concentration (in µg.l$^{-1}$) in the soil extracts*

## 2.2 Model derivation

***Freundlich model approach: conceptual advantages and disadvantages***

Although the use of semi-mechanistic or even empirical equation like a Freundlich equation undoubtedly implies that process knowledge is not fully taken into account, it can be shown that the concept of a Freundlich equation comes close to the nature of sorption phenomena in soils. One of the major conceptual shortcomings of the Freundlich equation is that there is no maximum adsorption capacity, so applications should be limited to situations where the degree of saturation of sites with elements of interest is low. Since in most soils the total site density (sum of sites on organic matter and clay) is much higher than the total load of heavy metals adsorbed onto these sites, this is usually not a problem.

A special situation arises, however, when precipitation occurs. Since this cannot be accounted for by a Freundlich model (where concentrations increase with an increase in the adsorbed amount in the soil, the model prediction invariably will fail. In a separate study by Bonten et al. (in prep.), this effect will be illustrated on data where precipitation is likely to occur. It is of importance, however, to realise that under conditions where adsorption is not the major process in control of the solubility (e.g. under reducing conditions), application of a Freundlich based model is not appropriate.

The major advantage of a Freundlich equation is, however, that it is rather easy to incorporate various soil properties that are known to control the adsorption behaviour of metals in soils. In many sorption studies the 'classic' Freundlich equation has been used to describe the experimental results (Elzinga et al., 1997, 1999; Ingwersen et al., 2000; Sauvé et al., 2000):

$$Q_{soil} = K_f \cdot C_{solution}{}^n \text{ or } Q_{soil} = K_f \cdot a_{solution}{}^n \qquad [1]$$

Where $Q_{soil}$ is (or should be) the reactive metal content of the soil, $K_f$ the Freundlich adsorption constant and n the non-linearity term, with $0 < n < 1$. The values for $K_f$ and *n* can be derived easily from a log-transformation of the experimental data. A plot of $\log[Q_{soil}]$ versus $\log[C_{solution}]$ results in a straight line with an intercept that is equal to $\log[K_f]$ and a slope that equals *n*. However, this will always results in soil specific values for $K_f$ and *n* and such constants can only be applied to the soil from which the value has been derived.

Although the Freundlich equation as described by eq. [1], is an empirical model, derived by regression analysis, there is a similarity between more advanced models like the NICA model.

Despite the fact that Freundlich type models can account for adsorption phenomena in soils, care should be taken not to use models derived by statistical analyses beyond the range of soil properties present in the database (but this is true also for mechanistic models, since mechanisms studied are only assumed valid within the boundary conditions of the data).

Equation [1] in the form as given as shown here is valid only for a single soil, since Kf depends on soil properties like pH, organic matter etc. To extend its use to describe sorption onto a variety of soils, however, $K_f$ can be expressed as a combination of various soil properties (e.g. Elzinga et al., 1999; Seuntjens, 2000; Tipping, 2003):

$$K_f = K^* \cdot [SOM]^a \cdot [clay]^b \cdot [H]^c \qquad [2]$$

Once the parameters *a*, *b*, *c* and $K^*$ are known (e.g. by multiple linear regression), adsorption for a given soil can be calculated based on the variation in C, SOM, clay and pH. Often a log-transformation is performed to obtain the values of all constants and the final Q-C or Q-a equation (calculation of Q based on the variation in C or a) is presented in the following form:

$$\mathrm{Log}[Q_{soil}] = \log K^{Q-C} + a \cdot \log[SOM] + b \cdot \log[clay] - c \cdot pH + n \cdot \log[C_{solution}] \qquad [3]$$

or

$$\mathrm{Log}[Q_{soil}] = \log K^{Q-a} + a \cdot \log[SOM] + b \cdot \log[clay] - c \cdot pH + n \cdot \log[a_{solution}] \qquad [4]$$

Activities were calculated from total concentrations in solution, concentrations of major cations, DOC and pH using a chemical speciation program (see end of this paragraph).

***Derivation of model parameters***

In order to use this equation to calculate the change in the dissolved metal concentration in case of pH changes in a certain soils, equation 3 has to be rewritten in order to calculate $C_{solution}$. Although this is done frequently, the estimate of C thus obtained is not right: When deriving equation [3] by multiple linear or stepwise regression, the coefficients in the equation are obtained by minimising the variation in the predicted values of $Q_{soil}$. However, when rearranging the equation to calculate $C_{solution}$ or $a_{solution}$ (In the sequel, C is a term which stands for concentration or activity), the predicted value of C would only be correct when a C-Q and a Q-C isotherm are equal, implying a perfect model, with an $R^2$ of 1. Since this is never the case, there will always be a difference between both isotherms which implies that the predicted value of C based on a Q-C isotherm (i.e. derived from minimising the variation in Q-predicted) is not correct.. This should be achieved only by a regression where C is the Y-variable and soil properties and metal content ($Q_{soil}$) are the predictor variables.

One way to obtain an equation that allows for the calculation of both Q and C (which can be rearranged both ways) is to simultaneously minimise the variance in Q and C. Although complicated statistical tools exist by which this can be done, a different, more practical approach was chosen here. Best estimates of the equation:

$$\mathrm{Log}[Q_{soil} / \{C_{solution}\}^n] = \log K + a' \cdot \log[SOM + b' \cdot \log[clay] - c' \cdot pH \qquad [5]$$

Or, taking into account DOC:

$$\begin{aligned}\mathrm{Log}[Q_{soil} / \{C_{solution}\}^n] &= \log K + a' \cdot \log[SOM + b' \cdot \log[clay] - c' \cdot pH \\ &\quad + d' \cdot \log[DOC]\end{aligned} \qquad [6]$$

were obtained by maximising the F-value of the regression as a function of n. Here units used for SOM and clay are %, pH equals pH solution, $Q_{soil}$ is the amount extractable by 0.43 mol.l$^{-1}$ $HNO_3$ in mol.kg$^{-1}$ and $C_{solution}$ is the measured concentration in the extracts in mmol.l$^{-1}$. To do so, n was varied between 0 and 1 and the optimal value (i.e. the value for n at which F was highest) was chosen as the best model. For all heavy metals considered, the shape of the n vs. F curve was parabolic with only one maximum in the range between 0 and 1. Using the data of n and the values of F, the optimal value of n was derived by calculation of the maximum of the parabolic curve. Apart from this approach we also present results from the "direct" regression of concentration (or activity) as a function of soil properties. Care should be taken not to use these to calculate Q by using these equations in reverse order. The reason for presenting these equations is to compare them with results from other studies and the fact that we can calculate uncertainties directly. When using ($Q/C^n$) as the explanatory variable, the uncertainty calculated from the standard error of the Y-estimate is not equal to either C-estimated or Y-estimated.

Both approaches were used for relations for the total dissolved metal concentration (Chapter 3) as well as for relations for the free metal ion activity (Chapter 4). In case of the total dissolved metal concentration, both the optimal model with and without taking DOC into account was derived.

In many cases however DOC is not available. From the data in this and other databases, a functional relation between soil properties and DOC was obtained (See Appendix 2). For DOC however, the soil - solution ratio appeared to very important, together with the organic matter content and pH:

$$\log(DOC) = 2.66 + 0.70 \cdot \log(\text{organic matter}) - 0.15 \cdot pH + 1.52 \cdot \log(\text{solid} / \text{sol.}) \qquad [7]$$

With DOC in mg C.l$^{-1}$, organic matter in %. Solid/solution is dimensionless and expressed on mass basis.

### *Calculation of free metal ion activity*

In this study, free metal ion activity levels were calculated from the solution composition using the chemical speciation model CHARON (De Rooij and Kroot, 1991). Both inorganic as organic complex formation was taken into account.

Metal complexation with DOC is calculated with a simple 2-pK model (Reinds et al., 1995). This model takes into account the protonation of protons and binding of cations (including metals) to both carboxyl- and phenol-type groups (pKa1 = 4.5, pKa2 = 9.5). Schematically the reactions are described by:

$$\mathrm{Cat}^{n+} + \mathrm{HDOC}^{-} \leftrightarrow \mathrm{CatHDOC}^{(n-1)+} \qquad \mathrm{p}K_1$$

$$\mathrm{Cat}^{n+} + \mathrm{DOC}^{2-} \leftrightarrow \mathrm{CatDOC}^{(n-2)+} \qquad \mathrm{p}K_2$$

Where *Cat* stands for cations (protons, base cations and trace metals) An overview of relevant $\mathrm{p}K_i$ values used in the model is given in Table 2.3. Complexation constants of metals with both groups were taken from literature or (for Ca and Cu) determined experimentally (Römkens, 1998).

*Table 2.3 List of model parameters used to calculate the speciation of the soil solution*

| Element | $\mathrm{p}K_1$ | $\mathrm{p}K_2$ | Element | $\mathrm{p}K_1$ | $\mathrm{p}K_2$ |
|---|---|---|---|---|---|
| $\mathrm{Al}^{1*}$ | 11.4 | 29.4 | $\mathrm{H}^{+}$ | 4.4 | 9.4 |
| $\mathrm{Ca}^{2+}$ | 3.6 | 6.1 | $\mathrm{Pb}^{2+}$ | 5.5 | 10.0 |
| $\mathrm{Cd}^{2+}$ | 4.0 | 7.9 | $\mathrm{Zn}^{2+}$ | 4.0 | 8.2 |
| $\mathrm{Cu}^{2+}$ | 6.5 | 11.4 | $\mathrm{Fe(III)}^{1*}$ | 15.2 | 34.5 |

[1] also mixed Fe-Al-DOC complexes are taken into account

[*] for both Al and Fe(III) the complexes used here are $\mathrm{MeHUM}^{+}$ for $\mathrm{p}K_1$ and $\mathrm{MeHUM(OH)}_{2-}$ for $\mathrm{p}K_2$ respectively.

Based on experimental data the site density on dissolved organic matter is fixed at 10 $\mathrm{meq.gC}^{-1}$ (Dai et al., 1996; Quideau and Bockheim, 1997; Römkens, 1998).

To calculate free metal activities that were used in the regression against solid phase data, the following parameters were included in the speciation model:

1. pH soil solution,
2. total dissolved Ca, Mg, Fe, and Al concentrations
3. total dissolved metal concentrations
4. DOC as an estimate of the total dissolved binding capacity
5. The total anion concentration from the background electrolyte (i.e. nitrate or chlorine) assuming that little or no adsorption/exchange occurred of these elements. In reality this proved to be a good assumption that was verified by the calculation of the charge balance.

# 3 Partition relationships for the calculation of the total dissolved metal concentration

## 3.1 Derivation of C - Q isotherms: direct estimate of concentration

***The partition equations used***

The final database used to derive partition equations contains 1466 samples. Apart from major soil properties (organic matter, clay content, pH, DOC and 0.43 $mol.l^{-1}$ $HNO_3$ extractable metal content) also additional properties like amorphous Fe and Al content ($NH_4$-oxalate extractable), Ionic strength (EC) of the extract, dissolved Ca concentration, and CEC are known for each sample.

Any of these additional parameters has an effect on the distribution of metal between the solid phase and the soil solution. However, often data on Ca in solution, or Fe-Al content are not available. Therefore a minimum relation was derived with the general available parameters pH, organic matter and clay content. To evaluate whether additional soil properties are able to significantly improve the partition equation, a full regression (including all available parameters) was performed as well. This resulted in 3 models for the estimation of the dissolved concentration:

1. Model CI includes all known parameters that were available in the dataset.
2. Model CII only includes the major soil properties (organic matter, clay and pH) and Dissolved Organic Carbon,
3. Model CIII includes only Organic matter, clay, pH and the reactive metal content.

All models are based on the same structure, i.e.:

$$\mathrm{Log}[C_{Me}] = \mathrm{INT} + \alpha_1 \cdot \log\{\text{par-1}\} + \alpha_2 \cdot \log\{\text{par-2}\} + \ldots. + \alpha_n \cdot \log\{\text{par-n}\} \quad [8]$$

With $C_{Me}$ in $mmol.l^{-1}$. All other units as well as the values for INT, $\alpha_1$, $\alpha_2$, $\ldots\alpha_n$ are shown in Table 3.1. Apart from this also the standard error of Y-estimated is shown. This was used to derive the confidence intervals of the estimates according to:

$$n\% \text{ conf. Interval} = \text{Y-est.} \pm t_{n-2} \cdot \text{se(Y-est)}$$

With $t_{n-2}$ equal to 1.64 in case of the 90% interval (student t-distribution at n = 1466). All models were derived using GENSTAT version 6.1. All parameters that did not meet the 95% significance criteria are marked in bold. For all other parameters t-values were usually less than 0.001.

***The regression coefficients obtained***

The results for the statistical analyses are summarised in Table 3.1. For all metals included in the study the contribution of pH, organic matter, clay and metal content was crucial (model C-III). Differences between elements are obvious: elements that are predominantly present in cationic non-organically complexed forms like Cd, Zn and to a lesser extent Ni, pH has a profound impact on the solubility. The coefficient for pH ranges from 0.4 to 0.5 for these three elements whereas the coefficient for Cu and Pb is less pronounced and varies from approx. 0.1 (Cu) to 0.2 (Pb). This is mainly due to the effect of complexation with DOC which masks the effect of pH. This is due to the fact that a decrease in pH both decreases metal binding to the soil solid phase and to DOC in solution. Whereas the first effect increases the free metal ion concentration in solution the second at the same time causes a decrease in the total concentration. For Cd, Zn, and Ni usually an increase in the dissolved concentration is observed when decreasing soil pH, whereas for Cu and Pb this effect is less pronounced and buffered by DOC. The differences in the coefficients for OM and clay are less pronounced and are related mainly to the binding capacity of both soil properties (together they possess the majority of biding sites under normal soil conditions)

For Cu and Pb, therefore, a clear improvement was obtained by including Dissolved Organic Carbon (Model C-II). This effect was less obvious for Cd, Zn and Ni. This is not surprising since most of the Cu and Pb in solution is bound to DOC and the presence of DOC is one of the major factors controlling the solubility of both elements. However, even for Zn and Ni and to a lesser extent for Cd, model CII (including DOC) is significantly better than model C-III. Based on these findings it was concluded that model C-II is to be preferred when data on DOC are available.

Further addition of parameters like CEC, Al-ox, Fe-ox, electrical conductivity and dissolved Ca concentration did not result in a significant improvement of the estimate as expressed by the standard error of the estimates (se-$Y_{est}$). In most cases only a marginal increase of the correlation coefficient was obtained ranging from less than 1% for Zn to approx. 5% for Cd an Cu. Furthermore, not all parameters proved to be significant when all available data were used. This is due to correlation between soil properties (e.g. between CEC on one hand and clay and organic matter on the other). For example when CEC is included the effect of clay is invariably reduced and even becomes insignificant for all elements except for Zn.

The overall goodness of fit as indicated by the correlation coefficient ranged from 50% for lead to 78% for Zn. The model gives the best fit for metals that prevail in their cationic form. Strong complexing metals like Cu an Pb give a less good fit. This may be due to the fact that the total concentration in solution , including metals complexed with DOC, is not directly related to the solid phase concentration which is expected to be in equilibrium with the free metal ion in solution. Including DOC as an explaining variable improves the fit especially for Cu. The model for lead was less good than others and could not be improved further by taking into account additional soil properties (model C-I had an $R^2$ of 54%).

*Table 3.1 Overview of Transfer Functions (according to equation 8) to calculate the total dissolved metal concentration in soil extracts (n=1466)*

| Metal | Parameter | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | INT<br>- | Qme<br>$mol.kg^{-1}$ | OM<br>% | pH<br>pH extract | Clay<br>% | DOC<br>$mg\ C.l^{-1}$ | $Al_{ox}$<br>$mmol.kg^{-1}$ | $Fe_{ox}$<br>$mmol.kg^{-1}$ | CEC<br>$meq.100g^{-1}$ | EC<br>M | Ca<br>$mmol.l^{-1}$ | Se-Y | $R^2$ |
| Cadmium | | | | | | | | | | | | | |
| CI | 5.83 | 1.33 | -0.43 | -0.45 | -0.09 | 0.23 | -0.35 | -0.22 | -0.40 | -0.19 | 0.28 | 0.49 | 72.1 |
| CII | 4.91 | 1.27 | -0.73 | -0.39 | -0.48 | 0.08 | - | - | - | - | - | 0.53 | 66.6 |
| CIII | 5.05 | 1.26 | -0.69 | -0.40 | -0.48 | - | - | - | - | - | - | 0.54 | 66.5 |
| Copper | | | | | | | | | | | | | |
| CI | 0.49 | 0.83 | -0.42 | -0.14 | **0.00** | 0.65 | -0.11 | **-0.05** | -0.36 | **-0.01** | **-0.08** | 0.40 | 60.8 |
| CII | **-0.24** | 0.82 | -0.56 | -0.08 | -0.33 | 0.55 | - | - | - | - | - | 0.43 | 54.7 |
| CII-a | - | 0.86 | -0.57 | -0.09 | -0.33 | 0.53 | - | - | - | - | - | 0.43 | 54.7 |
| CIII | 1.10 | 0.87 | -0.28 | -0.18 | -0.27 | - | - | - | - | - | - | 0.49 | 42.1 |
| Nickel | | | | | | | | | | | | | |
| CI | 2.27 | 0.84 | -0.59 | -0.41 | **0.01** | 0.42 | 0.06 | -0.17 | -0.32 | -0.17 | 0.29 | 0.37 | 70.3 |
| CII | 2.78 | 0.91 | -0.68 | -0.40 | -0.22 | 0.28 | - | - | - | - | - | 0.39 | 67.0 |
| CIII | 3.40 | 0.93 | -0.53 | -0.45 | -0.20 | - | - | - | - | - | - | 0.41 | 64.1 |
| Lead | | | | | | | | | | | | | |
| CI | 0.73 | 0.63 | -0.42 | -0.22 | **0.02** | 0.41 | -0.10 | -0.43 | -0.24 | **0.16** | -0.21 | 0.61 | 53.8 |
| CII | **-0.22** | 0.69 | -0.73 | -0.20 | -0.34 | 0.35 | - | - | - | - | - | 0.63 | 50.4 |
| CII-a | - | 0.72 | -0.75 | -0.21 | -0.33 | 0.33 | - | - | - | - | - | 0.63 | 50.4 |
| CIII | 0.51 | 0.70 | -0.54 | -0.26 | -0.30 | - | - | - | - | - | - | 0.65 | 47.7 |
| Zinc | | | | | | | | | | | | | |
| CI | 4.17 | 1.08 | -0.35 | -0.51 | -0.37 | 0.24 | -0.07 | -0.19 | -0.11 | -0.21 | 0.21 | 0.50 | 78.0 |
| CII | 4.26 | 1.08 | -0.46 | -0.50 | -0.51 | 0.21 | - | - | - | - | - | 0.50 | 77.5 |
| CIII | 4.69 | 1.08 | -0.35 | -0.54 | -0.48 | - | - | - | - | - | - | 0.51 | 76.9 |

Bold numbers indicate significance levels below the 95% level

This means that either other properties, not included in this study, control the solubility of lead or that the way in which the reactive metal content was measured in this study (0.43 mol.l$^{-1}$ $HNO_3$) does not reflect the available metal pool in soils. Especially for lead, the forms in which this metal is present in soil remain not quite understood. Estimates of the reactivity (i.e. the fraction of lead in soils that participates in soil chemical equilibria) range from less than 10% of the total lead content in soils to almost 90%. Also the binding of reactive lead to the functional groups at clay or organic matter has not been resolved completely. In contrast to Cd and Zn which are bound by non-specific electrostatic forces only, lead is able to bind almost irreversibly onto these same groups. This makes that models such as the ones presented here which do not take into account these specific effects cannot grasp the full complexity of the binding behaviour of all elements.

It should be stressed that the equations shown in Table 3.1 should be used *only* for the calculation of the dissolved metal concentration and not for the derivation of Q (see 2.2 derivation of model parameters). This should be done only by equations presented later.

***Comparison of model predictions with laboratory data***

In Figure 3.1 model predictions are shown. Model predictions are based on model C-II for all elements. The goodness of fit and accuracy of the predictions decreases in the order Zn > Ni > Cd ≈ Cu > Pb. It should be stressed here that the partition functions used here cover an extremely wide range in soils, degrees of soil pollution and dissolved metal concentrations. As such the model is able to predict the order of magnitude of the dissolved metal concentration quite reasonably. Only for Pb and to a lesser extent Cu, model improvement seems necessary when the model is used to predict concentration at specific sites. Dissolved metal concentrations for lead were near the detection limit of the ICP (AES) used here. Model improvement by addition of new data with a higher degree of analytical accuracy probably is a first step to improve the model.

## 3.2 Derivation of $K_f$ partition equations based on concentration

***The partition equations used***

Similar to the approach to estimate the total dissolved metal concentration also the value of Kf can be estimated. In literature often estimates of $K_d$ (equal to $Q_{Me}/C_{Me}$) are given as a function of soil properties (Römkens and Salomons, 1998; Holm et al., 1998). When estimating $K_d$, however, it is assumed that the relation between Q and C is independent of the magnitude of C and Q, whereas this is not the case. It assumes that the relation between Q and C is linear (n = 1) whereas it is generally accepted that the relation between Q and C is a non-linear relationship (i.e. n ≠ 1) because of decreasing affinity with increasing concentrations.

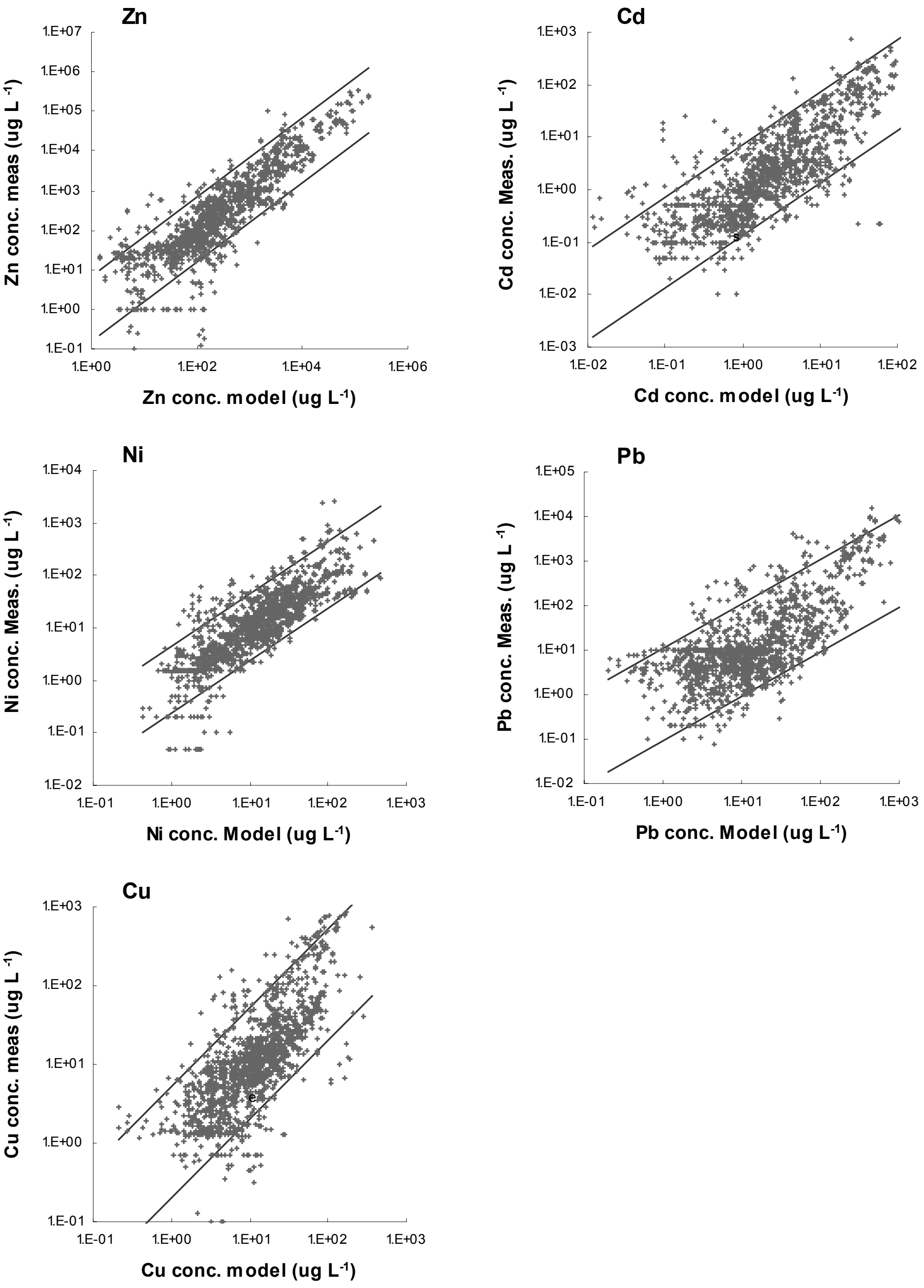

*Figure 3.1 Model predicted dissolved metal concentration using model II (Table 3.1) versus measured data and 90% confidence levels.*

Here, $K_f$ is related to soil properties according to:

$$\log K_f = \log[Q_{Me}/C_{Me}^{n}] = INT + \alpha_1 \cdot \log\{par\text{-}1\} + \alpha_2 \cdot \log\{par\text{-}2\} + \ldots + \alpha_n \cdot \log\{par\text{-}n\} \quad [9]$$

with: C in $mmol.l^{-1}$, all other units are given in Tables 3.2 and 3.3

The advantage of this equation is as stated before that one can use it to calculate both C as a function of Q and Q as a function of C. The latter option is of importance in order to calculate critical soil metal levels at specified critical limits in the soil solution (e.g. toxicity limits or ground/surface water limits). A relation that can be used in both directions can also be used to describe sorption and desorption in dynamic simulation models. The coefficients for the different model parameters were derived as described in chapter 2.2.

### ***The regression coefficients obtained***

In Table 3.2 the equations for the calculation of Kf are shown including the values for se(Y-est) and $R^2$. Together with the value for *n* this can be used to calculate C or Q as well. Here only the models taking into account the major soil properties (pH, organic matter, clay) are used (model CII, as well as the version with DOC (model CI). As was the case for the C - Q isotherms presented in paragraph 3.1, DOC has a marked influence on the model for Cu and Pb and to a lesser extent for Ni and Zn.

*Table 3.2 Overview of transfer functions to calculate $K_f$ (according to equation 9). Bold values indicate that the standard error of the parameter estimate exceeds the value of the parameter estimate itself (low significance)*

| Metal | | *Parameter* | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | INT | OM % | pH extract | Clay % | DOC mg.l$^{-1}$ | n-opt | Se-Y | $R^2$ |
| Cd | CI | -4.75 | 0.61 | 0.26 | 0.29 | -0.05 | 0.54 | 0.33 | 0.80 |
| | CII | -4.85 | 0.58 | 0.27 | 0.28 | | 0.54 | 0.33 | 0.79 |
| Cu | CI | -2.61 | 0.60 | 0.12 | 0.23 | -0.27 | 0.59 | 0.34 | 0.65 |
| | CII | -3.55 | 0.48 | 0.16 | 0.18 | | 0.47 | 0.35 | 0.62 |
| Ni | CI | -4.73 | 0.72 | 0.30 | 0.39 | -0.13 | 0.54 | 0.28 | 0.88 |
| | CII | -5.05 | 0.65 | 0.31 | 0.39 | | 0.51 | 0.28 | 0.88 |
| Pb | CI | -2.38 | 0.95 | 0.22 | 0.07 | -0.23 | 0.73 | 0.55 | 0.59 |
| | CII | -2.96 | 0.83 | 0.25 | **0.02** | | 0.68 | 0.55 | 0.57 |
| Zn | CI | -4.23 | 0.47 | 0.43 | 0.37 | -0.14 | 0.75 | 0.40 | 0.82 |
| | CII | -4.51 | 0.39 | 0.45 | 0.35 | | 0.74 | 0.40 | 0.82 |

The choice of soil properties that are included in the model will have an effect on the magnitude of the parameters. A parameter which is known to affect metal sorption in soils is the amount of amorphous Fe and Al oxides. Usually these are determined by an oxalate extraction. In our database this was measured as well and to evaluate the role of Fe and Al oxides on the outcome of the models tested here, it was included in the regression as well. In Table 3.3 the equations including Fe and Al oxides are shown.

For most metals the effect seems limited, for Ni the contribution proved to be non-significant. Addition of Fe and Al-oxides does not affect the coefficients for pH nor DOC but only suppressed the effect for clay and organic matter. This is not surprising since clay, organic matter and Fe Al-oxides form the major adsorbing compounds in soil. When Fe-Al is excluded from the regression the contribution of this compound is distributed among the two remaining compounds. The fact that for Ni no net effect of Fe-Al oxides on the regression was obtained suggests that Ni mainly sorbs onto organic matter and clay.

*Table 3.3 Transfer functions to calculate $K_f$ (according to equation 9) including Fe and Al oxides*

| Metal | *Parameter* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | INT | OM % | pH | Clay % | DOC $mg.l^{-1}$ | Fe+Al $mmol.kg^{-1}$ | n-opt | Se-Y | $R^2$ |
| Cd | -4.75 | 0.61 | 0.26 | 0.29 | -0.05 | - | 0.54 | 0.33 | 0.80 |
| | -5.33 | 0.46 | 0.25 | 0.16 | -0.05 | 0.36 | 0.52 | 0.31 | 0.81 |
| Cu | -2.61 | 0.60 | 0.12 | 0.23 | -0.27 | - | 0.59 | 0.34 | 0.65 |
| | -3.07 | 0.45 | 0.12 | 0.12 | -0.27 | 0.35 | 0.60 | 0.33 | 0.66 |
| Ni | -4.73 | 0.72 | 0.30 | 0.39 | -0.13 | - | 0.54 | 0.28 | 0.88 |
| | -4.73 | 0.71 | 0.30 | 0.39 | -0.13 | 0.00 | 0.54 | 0.28 | 0.88 |
| Pb | -2.38 | 0.95 | 0.22 | 0.07 | -0.23 | - | 0.73 | 0.55 | 0.59 |
| | -2.97 | 0.71 | 0.23 | -0.09 | -0.25 | 0.59 | 0.79 | 0.56 | 0.61 |
| Zn | -4.23 | 0.47 | 0.43 | 0.37 | -0.14 | - | 0.75 | 0.40 | 0.82 |
| | -4.52 | 0.39 | 0.43 | 0.30 | -0.13 | 0.20 | 0.74 | 0.40 | 0.82 |

### *Comparison of model predictions with laboratory data*

Predictions of the value of Kf are shown in Figure 3.2. Confidence intervals of the estimates are given as well according to:

n% conf. Interval = Y-est. ± $t_{n-2} \cdot$ se(Y-est)

With $t_{n-2}$ equal to 1.64 in case of the 90% interval (student t-distribution at n = 1466). The performance of the models decreases in the order: Ni > Zn > Cd > Cu > Pb. Again, the model to predict the distribution of Pb in soils is worse than those for the other models.

## 3.3 Difference between C-Q isotherms and Kf partition equations

In general the Kf model is not strikingly different from the C - Q model but its major advantage is that it can be used to calculate both C from Q and Q from C. The disadvantage, obviously, is that the Kf model loses some degree of accuracy to predict C from Q compared t he C-Q isotherm. However the differences between both are marginal as is illustrated for Zn in Figure 3.3 Here both models (C - Q from paragraph 3.1 and Kf model) were used on the same data to predict C.

However, there seems to be a difference between the models when looking at the predictions in the low and high range. Compared to the direct C-Q model concentrations calculated with the Kf model are lower in the low concentration range whereas they are higher in the high concentration range. This is visible in Figure 3.3 as the deviation from the 1:1 line (marked bold) where predictions in the high range are consistently larger using the Kf model (and lower in the low range although this difference was less obvious).

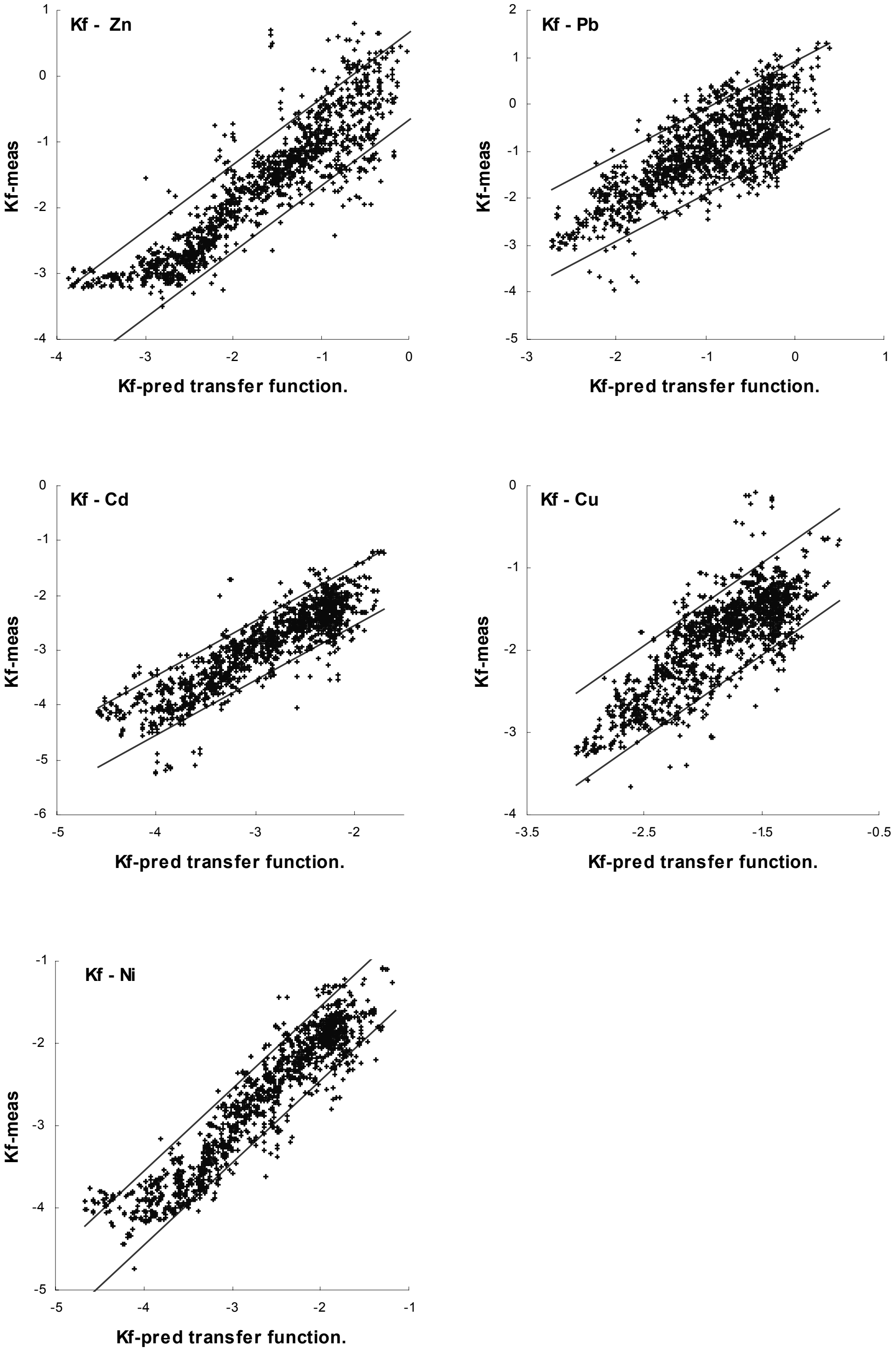

*Figure 3.2 Predicted and measured Kf (Concentration) using model C-I (Table 3.2) with 90% confidence levels*

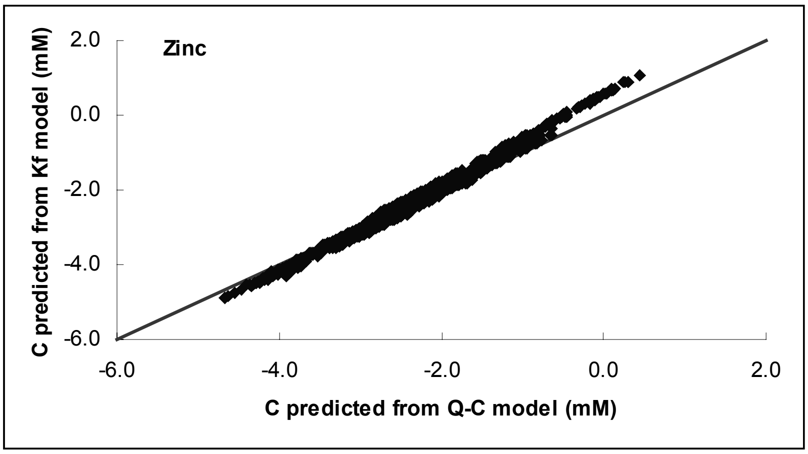

*Figure 3.3 Comparison of predicted dissolved Zn concentrations using the direct C-Q isotherm and the Kf model (the line indicates a 1:1 fit)*

# 4 Partition relationships for the calculation of the free metal ion activity

## 4.1 Derivation of a - Q isotherms: direct estimate of activity

***The partition equations used***

When fitting the calculated speciation data to solid phase data, 2 options were available. Model AI includes all known solid phase parameters that were available in the dataset. Model AII only includes the major soil properties and the reactive metal content. DOC was obviously not included (neither were other soil solution properties except the pH in the soil solution) since this was used to calculate the speciation in the soil solution.

All models are based on the same structure, i.e.:

$$\text{Log}[a_{\text{Me}}] = \text{INT} + \alpha_1 \cdot \log\{\text{par-1}\} + \alpha_2 \cdot \log\{\text{par-2}\} + \ldots + \alpha_n \cdot \log\{\text{par-n}\} \qquad [10]$$

With $a_{\text{Me}}$ = the free ion metal activity in solution in mmol.l$^{-1}$, all other units as well as the values for INT, $\alpha_1$, $\alpha_2$, …$\alpha_n$ are shown in Table 4.1.

Apart from this also the standard error of Y-estimated is shown. This was used to derive the confidence intervals of the estimates according to:

$$\text{n\% conf. Interval} = \text{Y-est.} \pm t_{n-2} \cdot \text{se(Y-est)}$$

With $t_{n-2}$ equal to 1.64 in case of the 90% interval (student t-distribution at n = 1466). In Figure 4.1 the relation between *a*-measured and *a*-estimated is shown graphically. The 90% confidence levels of the estimates are included as well (straight lines).For all metals model A-II was used.

***The regression coefficients obtained***

In Table 4.1 the results from the direct regression of the free metal ion activity as a function of soil properties are shown. In general the regression results in Table 4.1 indicate that the combination of parameters included here (organic matter, clay, pH and metal content) are able to explain the majority of the variation in the results. Not surprisingly, the model for Cu works best. Model calculations of the free metal ion activity from the solution speciation are based on measured data. In general speciation for Cu is well characterised, this in contrast to that of Pb and Ni. Also for Cd there is rather large unexplained variation despite the fact that complexation of Cd with organic ligands is limited which should result in a regression curve similar to that of Zn. A possible explanation is the fact that Cd concentrations are much lower and as a result the analytical error is relatively large.

*Table 4.1 Overview of Transfer Functions to calculate the free metal activity in soil extracts according to equation 10 (n=1466). Bold values indicate that the standard error of the parameter estimate exceeds the value of the parameter estimate itself (low significance)*

| Metal | Parameter | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | INT | Qme mol.kg$^{-1}$ | OM % | pH | Clay % | $Al_{ox}$ mmol.kg$^{-1}$ | $Fe_{ox}$ mmol.kg$^{-1}$ | CEC meq. 100 g$^{-1}$ | Se-Y | $R^2$ |
| **Cadmium** | | | | | | | | | | |
| AI | 6.65 | 1.39 | -0.55 | -0.51 | -0.11 | -0.37 | -0.21 | -0.17 | 0.50 | 73.6 |
| AII | 5.27 | 1.31 | -0.87 | -0.46 | -0.42 | - | - | - | 0.53 | 70.2 |
| **Copper** | | | | | | | | | | |
| AI | 2.91 | 0.80 | -0.57 | -0.80 | 0.11 | -0.14 | 0.09 | -0.47 | 0.42 | 86.6 |
| AII | 1.91 | 0.78 | -0.78 | -0.70 | -0.23 | - | - | - | 0.46 | 83.6 |
| **Nickel** | | | | | | | | | | |
| AI | 3.50 | 0.90 | -0.49 | -0.50 | **-0.05** | **0.04** | -0.17 | -0.14 | 0.37 | 71.7 |
| AII | 3.39 | 0.94 | -0.60 | -0.49 | -0.23 | - | - | - | 0.37 | 70.7 |
| **Lead** | | | | | | | | | | |
| AI | 2.26 | 0.63 | -0.62 | -0.73 | **0.05** | -0.09 | -0.33 | -0.23 | 0.61 | 78.4 |
| AII | 1.67 | 0.68 | -0.90 | -0.70 | -0.23 | - | - | - | 0.62 | 77.5 |
| **Zinc** | | | | | | | | | | |
| AI | 5.34 | 1.13 | -0.42 | -0.65 | -0.35 | **-0.06** | -0.18 | **-0.04** | 0.52 | 79.7 |
| AII | 5.02 | 1.12 | -0.53 | -0.64 | -0.45 | - | - | - | 0.52 | 79.5 |

For the direct regression coefficients the effect of clay in the regression was limited (see Table 4.1). The contribution of clay appeared to be less important (although significant) than organic matter. In general the value of additional soil properties on the regression results appeared to be minimal (maximum of 3% increase in correlation coefficient, and almost n reduction of the se-Yest.). From this it was concluded that for practical purposes, only model AII should be used (based on clay, pH, and organic matter only).

### Comparison of model predictions with laboratory data

In Figure 4.1 the free metal ion activity as calculated with the transfer function is compared to the ones calculated using the soil solution speciation model (2-pK model). Here the modelled values from the soil speciation model are plotted against the ones obtained from the Freundlich regression model. For most elements the agreement between both models is quite good which is also reflected by a rather narrow band between the 90% confidence interval lines.

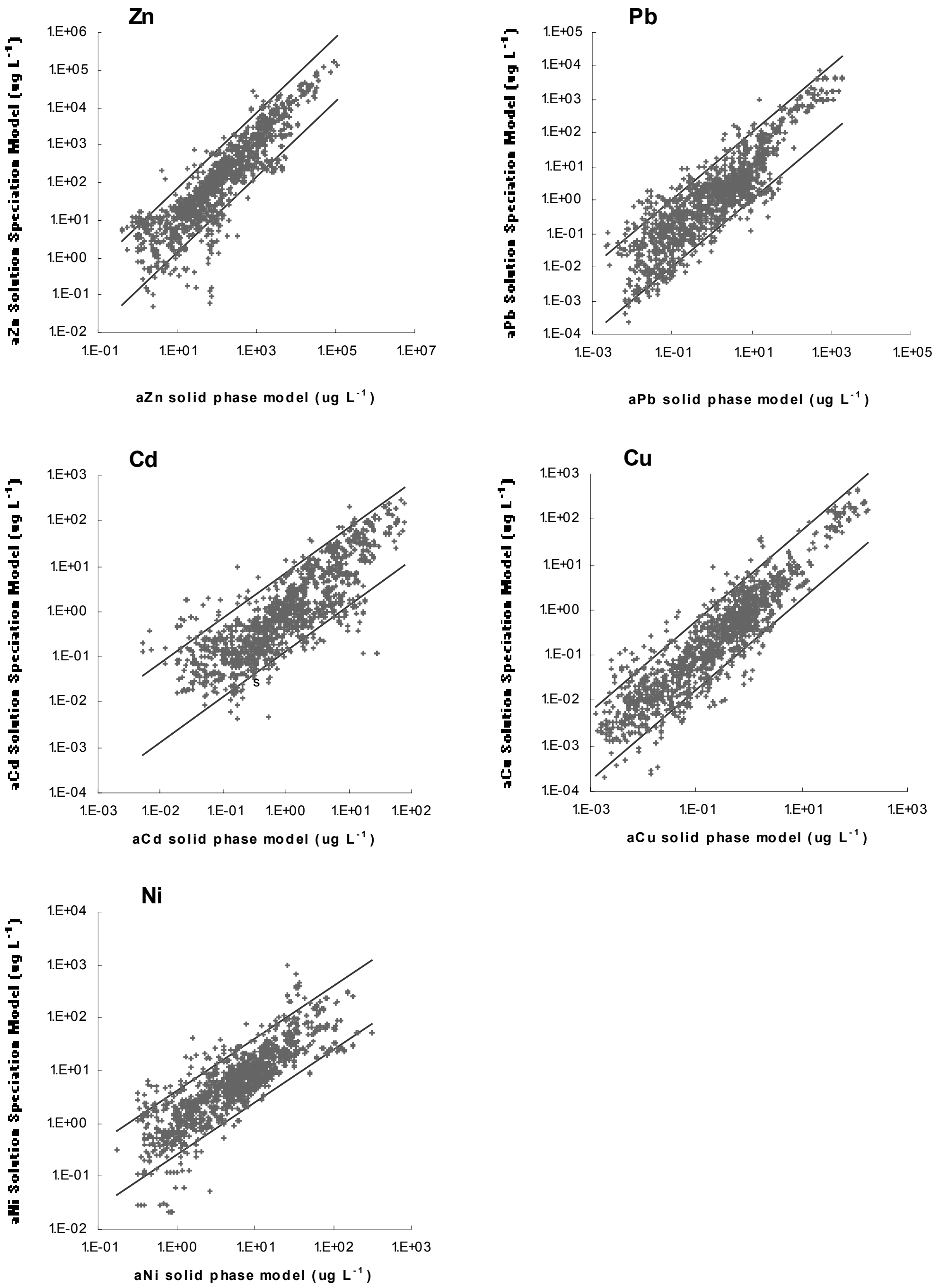

*Figure 4.1 Comparison of predictions of the free metal ion activity based on the transfer function and the solution speciation model. Straight lines mark the 90% confidence limits of the predicted values of the transfer functions*

## 4.2 Derivation of $K_f$ partition equations based on activity

### *The partition equations used and regression coefficients obtained*

Similar to the approach outlined for total dissolved concentrations, the value of Kf can be estimated, expressed in terms of activity. In Table 4.2 the results of the estimate of $K_f$ as a function of soil properties 1 through n is given. Here, $K_f$ is defined as:

$$\log K_f = \log[Q_{Me}/a_{Me}^{n}] = INT + \alpha_1 \cdot \log\{par\text{-}1\} + \alpha_2 \cdot \log\{par\text{-}2\} + \dots + \alpha_n \cdot \log\{par\text{-}n\} \quad [11]$$

with: $a_{Me}$ = the free ion metal activity in solution ($mmol.l^{-1}$), all other units are given in Table 4.2.

Based on the results of the direct regression it was decided not to include the additional soil properties.

*Table 4.2 Overview of transfer functions to calculate $K_f$ based on activity according to equation 11.*

| Metal | | Parameter | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | INT | OM % | pH | Clay % | n-opt | Se-Y | $R^2$ |
| Cd | | -4.76 | 0.66 | 0.29 | 0.25 | 0.55 | 0.32 | 0.83 |
| Cu | | -3.25 | 0.87 | 0.67 | 0.23 | 0.90 | 0.43 | 0.89 |
| Ni | | -4.81 | 0.67 | 0.35 | 0.38 | 0.58 | 0.27 | 0.90 |
| Pb | n-opt. | -3.22 | 1.20 | 0.77 | 0.14 | 1.04 | 0.67 | 0.82 |
| | n=1 | -3.28 | 1.19 | 0.74 | 0.12 | 1.00 | 0.65 | 0.82 |
| Zn | | -4.59 | 0.50 | 0.52 | 0.33 | 0.73 | 0.40 | 0.85 |

For Pb two isotherms are given. The value for n exceeds 1 which is not realistic. It means that Q increases exponentially with an increase in activity. Therefore, also an isotherm with n=1 has been derived.

### *Comparison of model predictions with laboratory data*

In Figure 4.2 the measured vs. predicted Kf value are shown together with the 90% confidence levels. The differences between the elements included here are not as pronounced as for the other approaches shown previously. Both for Zn and Ni an obvious tailing of the curve became apparent in the lower Kf range ($Kf < -3$ for Zn and $Kf < -3.5$ for Ni). This appeared to be a group of samples from soils with an extremely low total reactive metal content (Zn: $< 3$ $mg.kg^{-1}$; Ni $< 0.5$ $mg.kg^{-1}$) in combination with an extremely acid pH (between 2 and 3). As a result of the extremely low pH, dissolved metal concentrations were very high ($> 6000$ $\mu g.l^{-1}$ for Zn), which makes these samples rather exceptional in their composition. However, it shows that at the extreme end of the predictive curve, care should be taken with some of the results from the predictions.

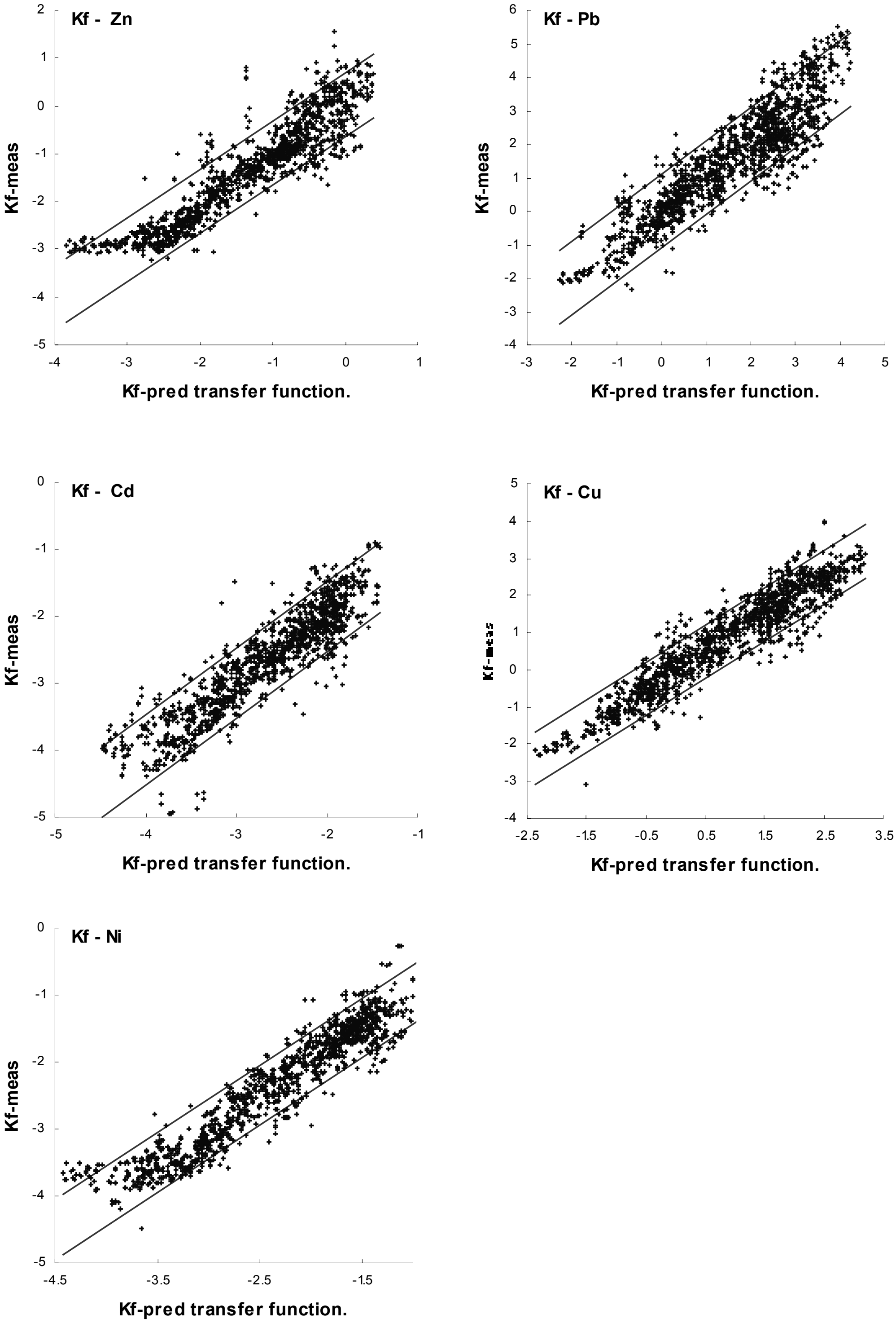

*Figure 4.2 Predicted and measured Kf (Activity) using coefficients from Table 4.2 with 90% confidence levels*

# 5 Comparison of results with approaches and data from the literature

### *Available approaches*

During the last decade various papers have been published describing metal partitioning in soil. The following approaches have been described (here we only mention the empirical models without describing the mechanistic approaches):

- Derivation of $K_d$ (heavy metal in soil/heavy metal in soil solution) values without correction for soil properties (Gooddy et al., 1995);

- Relationships between $K_d$ and soil properties (Römkens and Salomons, 1998; Yin et al., 2002; Tipping et al., 2003);

- Relationships between the adsorbed metal content and soil properties and the concentration in the soil solution (or vice versa; Chardon, 1984; Boekhold, 1992; Elzinga et al., 1997, 1999; Seuntjens, 2000; Ingwersen et al., 2000; Otte et al., 2000)

- Relationships between the free metal ion activity and the adsorbed metal content in combination with soil properties. (Sauvé et al., 1997a, b, 2000a, b; Römkens, 1998; Tipping et al., 2003; Tye et al., 2003).

One of the major problems which complicate the comparison of the results from various studies are differences in the approach:

- Some of studies are adsorption studies (metals are added to the soil in the form of salts), whereas others deal with desorption (similar to the results presented in this report, no metals are added).

- In many cases different methods to obtain the soil solution were used, ranging from soil extracts (1:2 to 1:10 soil - solution ratios) to in situ soil solution obtained by centrifugation of field moist soil or using so-called artificial roots placed in soil containers.

### *Model comparisons made*

To avoid comparison of data which are non-comparable we will compare the results from two of the largest studies that have a comparable approach (desorption only, using extracts). One is a laboratory study performed by Tipping et al. (2003) and the other a compilation of data from the literature that were used in a study by Sauvé et al. (2000b). A major difference between these sources is the range in soil properties present in the databases. In the study by Sauvé a wide range (comparable to the results presented here) in soil properties was present whereas the database from Tipping was filled mainly with soils high in organic matter (upland soils from Wales and England). Another difference between the datasets is the way how solution was extracted.

In Table 5.1 through 5.4 both studies are compared with the results from this study. To compare the results it was necessary to convert the data from our study to the same units used by the authors from the two studies mentioned. This was limited to organic matter for which a conversion factor of 0.58 (%OM into %C) was used and the units for the dissolved concentration and total metal content (μg.l$^{-1}$ and mg.kg$^{-1}$ respectively that were used by Sauvé et al instead of molar units). Also a new regression on the results from our study was performed using the same parameters as those included by the authors (clay and DOC were not included by both Tipping and Sauvé et al.)

### ***Comparisons of total concentration based isotherms***

In general the match between the isotherms based on total concentrations (C-Q) is very good, with the exception of that of Ni. In the database used by Sauvé et al. 2000b only 69 samples were included for Ni which could affect the outcome. However, when looking at the original data in the paper by Sauvé it can be seen that for the fit between Kd and pH (not taking into account OM, Table 2 in Sauvé et al., 2000a) for Ni similar values for the intercept were found as was the case for the other metals. However, after including OM as a second variable, the intercept suddenly differs more than 3 orders of magnitude from those of Cd, Cu, Pb and Zn. It is likely that some error has occurred.

*Table 5.1 Comparison of results from this study with those from Sauvé et al. (2000b) (Coefficients for the direct estimation of dissolved concentration from reactive metal content ($Q_{me}$), pH, and organic matter (OM)) according to equation 8. C in μg.l$^{-1}$*

| Metal | Source | Coefficients | | | | R² | Se(Y) |
|---|---|---|---|---|---|---|---|
| | | Q mg.kg$^{-1}$ | pH | OM % | INT | | |
| Cd | This study | 1.23 | -0.46 | -0.86 | 3.52 | 0.61 | 0.58 |
| | Sauvé et al. | 1.08 | -0.47 | -0.81 | 3.42 | 0.88 | 0.62 |
| Cu | This study | 0.84 | -0.21 | -0.38 | 1.65 | 0.38 | 0.51 |
| | Sauvé et al. | 0.93 | -0.21 | -0.21 | 1.37 | 0.61 | 0.47 |
| Ni | This study | 0.81 | -0.45 | -0.50 | 3.46 | 0.63 | 0.42 |
| | Sauvé et al. | 1.21 | -1.05 | -0.85 | 7.02 | 0.73 | 0.61 |
| Pb | This study | 0.76 | -0.31 | -0.71 | 1.96 | 0.46 | 0.67 |
| | Sauvé et al. | 0.56 | -0.37 | ns | 1.81 | 0.35 | 0.53 |
| Zn | This study | 1.09 | -0.61 | -0.56 | 4.24 | 0.72 | 0.56 |
| | Sauvé et al. | 0.94 | -0.55 | -0.34 | 3.68 | 0.62 | 0.72 |

Also the comparability between Kd isotherms is rather good. For Zn (see Figure 5.1), Pb and Cd (both not shown) the differences between calculated values of Kd based on the isotherm derived by Sauvé and the ones from this study were small. For Cu and Ni the differences were larger, predicted Kd values for Cu using the isotherms of Sauvé were systematically lower than the ones predicted by the isotherms derived from the data (Figure 5.2).

*Table 5.2 Comparison of results from this study with those from Sauvé et al. (2000b) (coefficients for the estimation of $K_d$ from pH and organic matter (OM) according to equation 9 with n=1, C in $ug.l^{-1}$ and Q in $mg.kg^{-1}$)*

| Metal | Source | *Coefficients* | | | $R^2$ | Se(Y) |
|---|---|---|---|---|---|---|
| | | pH | OM | INT | | |
| Cd | This study | 0.43 | 0.70 | -0.18 | 0.64 | 0.59 |
| | Sauvé et al. | 0.48 | 0.82 | -0.65 | 0.61 | 0.62 |
| Cu | This study | 0.24 | 0.48 | 1.32 | 0.45 | 0.51 |
| | Sauvé et al. | 0.21 | 0.51 | 1.75 | 0.42 | 0.55 |
| Ni | This study | 0.49 | 0.68 | -0.73 | 0.81 | 0.42 |
| | Sauvé et al. | 1.02 | 0.80 | -4.16 | 0.76 | 0.61 |
| Pb | This study | 0.33 | 0.89 | 1.17 | 0.54 | 0.68 |
| | Sauvé et al. | 0.37 | 0.44 | 1.19 | 0.56 | 0.59 |
| Zn | This study | 0.58 | 0.51 | -1.22 | 0.74 | 0.56 |
| | Sauvé et al. | 0.6 | 0.21 | -1.34 | 0.57 | 0.72 |

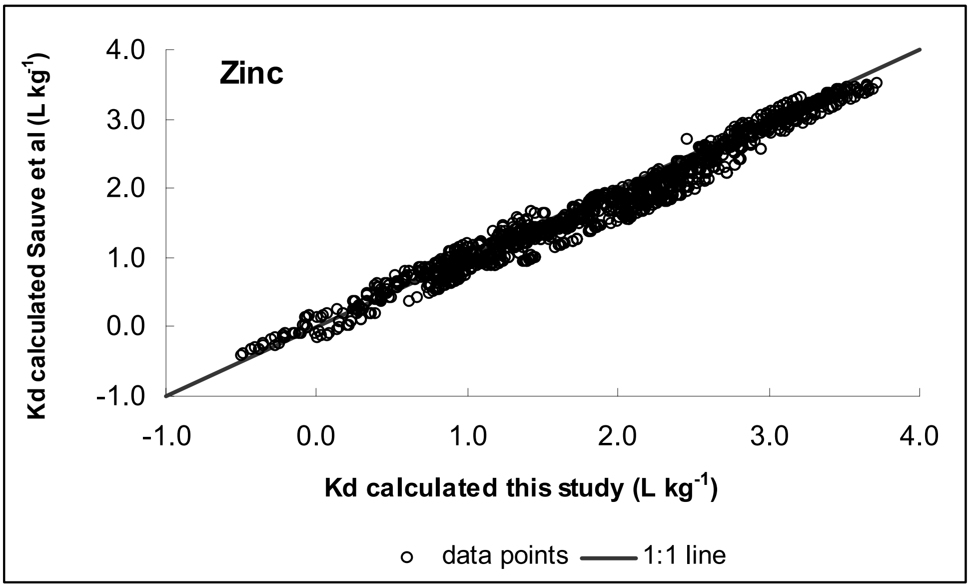

*Figure 5.1 Comparison between calculated Kd for Zn (concentration) using the equation from this study and that of Sauvé et al. (from Table 5.2)*

### *Comparisons of activity based isotherms*

The data in Tables 5.3 and 5.4 show that the transfer functions for activity (both direct regression results and the fit for $K_d$) are less identical than the transfer functions for total concentration (see also Figure 5.2 for a comparison between calculated Kd values for Lead). This might be due to the fact that only organic soils were included in the dataset of Tipping et al. (which is reflected by the higher coefficient for organic matter for Pb and Cu in Table 5.3). Furthermore different speciation models were used to calculate the free ion metal activity in solution which may also have an effect on the derived regression coefficients.

*Table 5.3 Comparison of results from this study with those from Tipping et al. (2003) Coefficients for the direct estimation of the free metal ion activity from reactive metal content ($Q_{me}$), pH, and organic matter (OM) according to equation 10. $a_{Me}$ in mol.l$^{-1}$, Q in mol.g$^{-1}$*

| Metal | Source | Coefficients | | | | $R^2$ | se-Y |
|---|---|---|---|---|---|---|---|
| | | OM | pH | Q | INT | | |
| Cd | this study | -1.03 | -0.51 | 1.28 | 5.95 | 0.66 | 0.56 |
| | Tipping | -0.69 | -0.42 | 0.97 | 5.54 | 0.76 | - |
| Cu | this study | -0.86 | -0.73 | 0.76 | 1.1 | 0.83 | 0.45 |
| | Tipping | -0.40 | -1.15 | 0.54 | 2.96 | 0.94 | - |
| Pb | this study | -1.03 | -0.73 | 0.72 | 1.09 | 0.77 | 0.63 |
| | Tipping | -0.53 | -1.14 | 0.87 | 6.12 | 0.95 | - |
| Zn | this study | -0.73 | -0.70 | 1.13 | 5.51 | 0.76 | 0.56 |
| | Tipping | -0.61 | -0.36 | 0.73 | 3.38 | 0.76 | - |

In general however, the comparison of the results from this study with those from other independent sources shows that the approach described here results in comparable if not almost identical results. This strengthens our hypothesis that this type of models can be used to describe the behavior of metals in a wide range of soils. Especially for applications on a large scale (region, nations) this approach seems very suitable since it uses rather well known soil properties that can be measured quite easily.

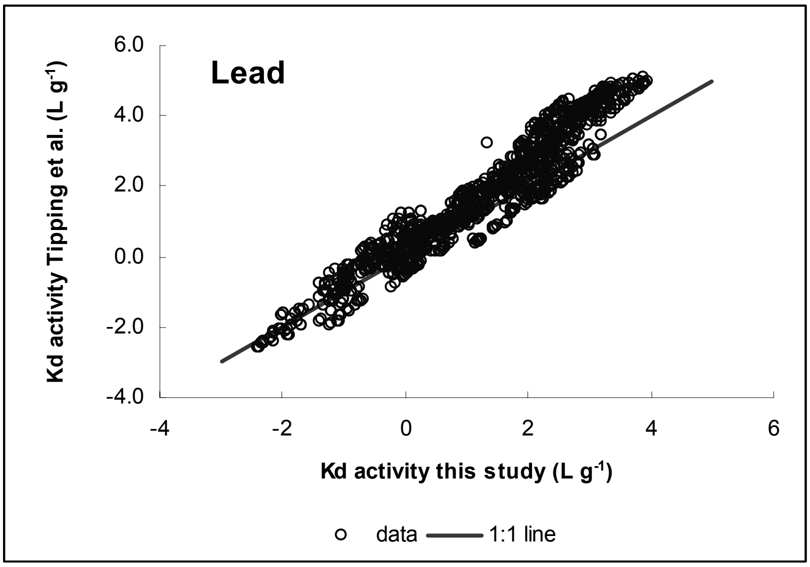

*Figure 5.2 Comparison between calculated Kd for Pb (activity) using the equation from this study and that of Tipping et al. (from Table 5.4)*

More consensus is needed however on the characterization of the reactive metal pool as well as a general methodological approach to obtain and characterize the soil solution. The use of extracts based on dilute salt extractions seems an appropriate way resulting in comparable estimates across a wide range of European and American soils. This, of course should not prevent people from using commonly accepted methods in various countries. The use of a generally accepted method to obtain soil solutions (or extracts mimicking these) apart from nation-specific methods seems highly desirable to enhance the possibility to use results from different sources.

*Table 5.4 Comparison of results from this study with those from Tipping et al. (2003) Coefficients for the estimation of the $K_d$ based on the free metal ion activity from, pH, and organic matter (OM) according to equation 11 with n=1, $a_{Me}$ in mol.$l^{-1}$, Q in mol.$g^{-1}$*

| Metal | Source | Coefficients | | | | Se-Y |
|---|---|---|---|---|---|---|
| | | OM | pH | INT | $R^2$ | |
| Cd | this study | 0.83 | 0.46 | -3.24 | 0.70 | 0.58 |
| | Tipping | 0.71 | 0.43 | -2.93 | 0.73 | - |
| Cu | this study | 1.01 | 0.77 | -3.07 | 0.88 | 0.48 |
| | Tipping | 0.48 | 1.16 | -4.69 | 0.93 | - |
| Pb | this study | 1.23 | 0.76 | -3.30 | 0.82 | 0.66 |
| | Tipping | 0.63 | 1.13 | -4.36 | 0.94 | - |
| Zn | this study | 0.64 | 0.67 | -4.44 | 0.79 | 0.57 |
| | Tipping | 0.90 | 0.39 | -3.49 | 0.69 | - |

# 6 Conclusions and discussion

## 6.1 Conclusions

The model approach presented in this report allows for the calculation of both the total dissolved metal concentration and the free metal ion activity. In its current version model predictions of Cd, Cu, Ni, Pb and Zn solubility and free metal activity can be obtained based on soil pH, organic matter and clay content in combination with the reactive metal content of the soil. Although prediction of the dissolved concentration and activity at extreme conditions ($pH < 3$, low reactive metal content, organic matter > 30%) remains to be improved, the equations are able to represent the measured data quite well, especially when considering the large range in soil properties and degree of pollution.

In general models for zinc, cadmium and to a lesser extent for nickel are working quite well for the prediction of the dissolved metal concentration. Models for copper and lead only obtain reasonable predictions when DOC is included in the equation. This is not surprising considering the large influence of Dissolved Organic Carbon on the speciation of both elements. An additional factor that reduces the fit for lead could be the precipitation of lead-bearing minerals like $Pb_5(PO_4)_3Cl$. In a separate paper (Bonten et al., in prep.) it is shown that lead bearing minerals can be stable under the experimental conditions imposed here. Also the role of manganese oxides can be of importance in describing lead-binding to soils. In general the role of Fe- and Al-oxides as adsorbing compounds has an effect on metal binding. However, the overall goodness of fit, and the standard error of Y-est. do not improve when both Fe and Al oxides are included in the regression analysis.

Addition of soil properties other than clay, organic matter and pH does not lead to significant improvement of the regression results for the metals tested here.

## 6.2 Discussion

***Which type of transfer function to use.***

In this report different types of transfer functions were derived. This includes "direct" relations to calculate the solution concentration or free metal activity from the reactive metal content in the solid phase (C-Q and a-Q relations respectively). These represent more or less the standard Freundlich type relations that have been used extensively by others as well. An additional approach is the derivation of transfer functions that calculate the Freundlich adsorption constant ($K_f$ defined as $Q/[C^n]$) from soil properties. The main criteria for doing this were that Freundlich isotherms are often used in two directions, i.e. used to calculate both the solid metal content (Q) or the dissolved concentration (C) depending on which variable is measured. From a statistical point of view it is then better to include both properties in the variable that is explained from soil properties.

The results of this study however do not give a clear answer to the question which approach (direct versus $K_f$) is to be preferred. Obviously when using a direct relation ship, the fit between measured and modelled data is slightly better than the one based on a $K_f$ approach. On the other hand, a fit of C on Q (to model C) results in a poorer fit when using the relation to model Q (in comparison to the Kf approach, which gives a better overall performance). However these statistical facts are not particularly helpful when making a decision on the application of either approach in modelling studies. The ultimate use of the model determines which one is to be preferred. The same is true for the choice between the calculation of the dissolved concentration of the free metal activity.

Direct (C-Q and a-Q) relations have the disadvantage (from a statistical point of view) that they can only be used in one way, i.e. the direction of their derivation, so either C from Q or Q from C. So in cases where one is solely interested in a prediction of the dissolved metal concentration or the free metal activity, the direct approach should be considered, simply because it works better. However, when transfer functions are to be used in both directions like in the case of dynamic modelling, transfer functions based on the Kf approach are to be preferred. The choice between transfer functions based on C (dissolved concentrations) versus *a* (free metal activity) also depends on available data and, to some extend, theoretical considerations.

The theory of chemical equilibrium thermodynamics indicates that the best relation between the solution and solid phase is based on the free metal ion activity rather then the total solution concentrations. This is mainly due to the fact that the adsorption - desorption reactions involve the free metal ion. This is reflected to some extent by the partition relations for activities, which show the highest explained variance. It should be kept in mind though that the regression of the activity isotherms (solid phase) is based on calculated free metal activity levels in solution. These are based on a chemical equilibrium model, which has been (partly) validated. In both models soil pH is included and this could to a certain extent lead to auto correlation, thereby increasing the goodness of fit between predicted activity levels from solution and from the solid phase. However, using different speciation models did not result in significantly different estimates of the free metal activity. This suggests that the estimate of the free metal activity that was used to obtain the partition equation for the solid phase is rather robust.

Again, the use of the models is an important issue for the choice between concentration or activity isotherms. Metal leaching from soils for example can be estimated directly by the concentration isotherms. On the other hand ecotoxicological studies often relate effects to the free metal activity. One should keep in mind that using activity isotherms, when total concentrations are to be calculated, always calls for the use of speciation models to convert the activity to concentration. The additional model errors in speciation models add up to the uncertainty in the partition equation. An interesting fact, albeit not reported here, is that the goodness of fit between data and model results from partition equations and mechanistic models is similar (Bonten et al, in prep).

The choice between solution and activity isotherms also depends on the metal of interest. For Zn and Cd, inclusion of DOC in the model for concentration hardly improves the goodness of fit. For Cu and Pb on the other hand, the use of models not considering DOC seems pointless, since the levels in solution of both elements is largely controlled by DOC.

In summary, direct or $K_f$ concentration models should be used when concentration is asked for (e.g. leaching). On the other hand direct activity based models are best for applications where activity is asked for. The advantage of a partition equation (solid phase model) is that less data are needed which are also more easily to obtain. In most cases data on organic matter, clay content, pH and metal content are available whereas data on the solution composition (needed for speciation models) are often scarce.

Total concentrations can also be calculated in two steps. Initially the free metal ion activity is calculated with a solid phase transfer function. In a second step the total concentration is calculated from the activity, DOC, pH and major cations using a speciation model. Further (uncertainty) analysis should reveal whether a direct calculation of total concentrations or a two step method gives the most accurate estimates.

### *Applicability of the model in regional scale and site scale approaches*

Despite the fact that the models give a rather good representation of the database as a whole, predictions of individual points are sometimes still erroneous. This is not surprising since it basically means that the model is able to reflect the availability in different soil types (or averages on a field or regional scale). When looking at one particular soil (or soil sample from one field) the capacity to predict either solubility or speciation is limited due to factors currently not included in the model (apart from temporal variability) like spatial and temporal variability on a small scale. In other words the application of models like the ones presented here is meant primarily for studies focussing on larger soil units (field level or higher), and not so much on individual soil samples. Truly generally applicable model (i.e. models that can be applied everywhere, from the micro scale to national scale) still require much more (site and soil specific) input. For example, the nature of organic matter in peat soil (moderately humified organic material) versus that in clay or sandy soil (completely humified material) is different. This difference however cannot be accounted for when the only available parameter for model calculations is 'total organic matter content'. Correction factors that take into account the effect of mineralisation of organic matter on the binding capacity are still lacking or at least not included in the model. The same is true for clay, the chemical properties of clays that are relevant for adsorption - desorption equilibria vary enormously. River clays in the temperate climate zones with a high illite content have a substantially higher CEC that kaolinitic clays in soils derived from parent material in tropical areas.

## *Limitations for European wide applications*

*The use of the concept of the reactive metal pool*

The approach presented here is based on the 'reactive' heavy metal content (here the 0.43 mol.$l^{-1}$ $HNO_3$ extraction was used) as an approximation of the total reactive pool. One can discuss as to whether this is either too strong or too weak. As of now only operationally defined ways to measure the reactive metal content exist. Although more advance techniques like isotopic dilution (Tye et al., 2003) are available, a robust method is needed in order to apply the concept of total vs. reactive on a regional, national or international scale. Compared to, for example, EDTA the differences are minor and even differences with Aqua Regia are, for most soils limited. This, however, is probably strongly related to the origin of metals in the soils. In artificially contaminated soils, the difference between total and reactive is small but in non-contaminated clay soils for example, the total Zn or Ni content can be substantially higher than the reactive Zn or Ni content. This is mostly due to fixation of these elements in the mineral structure of clays.

*Representativity of Dutch data for European applications*

Although the model is based on a large database with a considerable range in soil properties, one cannot neglect the fact that these are samples from Dutch soils only. In general Dutch soils are derived from unconsolidated sediments, either from wind, water or aerial deposits. Soils that originate from rock are virtually not present (with a few exceptions in the province of Limburg where soils developed on limestone are present). Furthermore, in the Netherlands a large part of the metals has been added to the soil, either from atmospheric deposition, manure, fertiliser or industrial and household waste. This means that the difference between reactive and total metal content, which basically accounts for the metals occluded in non-reactive minerals- is limited. This, however, might not be the case in soils derived from parent material as is the case in many European countries where soils derived in situ from either limestone, granite, shale etc are omnipresent. In general it is likely to assume that in soils derived from granite parent material, the reactive metal content, expressed as a part of the total metal content, is lower that in soil derived from sediments.

Also the nature of organic matter will be different in soils from different climatic zones. The Netherlands are located in a mild climatic region with an annual rainfall of approx. 800 mm and an average annual temperature of 11 degrees. This in contrast to soils from Spain, Greece, Italy, or, on the other end, Scandinavia or Scotland. Differences in temperature will influence the nature and amount of organic matter in soils and can therefore affect the binding capacity of organic matter (and clay as well).

*Experimental procedures to derive partition equations*

Another issue that remains to be solved for applications on a European scale is the way soil solutions or extracts are obtained and used for model derivation. Currently different ways to obtain the chemically available fraction are used ranging from water, dilute salts ($CaCl_2$, but also $NaNO_3$) to more concentrated salt solutions like 1 mol.$l^{-1}$ $NH_4NO_3$. The same is true, as stated before, for the reactive metal content. For application on a European scale harmonisation is highly desirable even if only to compare the results from the different methods.

Considering these three facts (distribution total – reactive, impact of climate on binding properties of soils and use of different soil solution extracts) care should be taken to use the models on a European scale without testing the applicability on strongly contrasting soils from other climatic regions.

***Future steps to improve the model***

To improve the general validity of the model, more experimental work is ongoing. Current emphasis is on speciation and solubility of metals in non-polluted soils at low solute concentrations. Actual measurements of soil solution activity using Donnan cells will be used to validate and if necessary re-calibrate the current version of the model. Field measurements are performed at 2 sites where metal solubility is being measured during the year in soil solutions at two depth intervals as well as in groundwater and drainage waters. This will not only yield field data on metal solubility in soils and groundwater but also increase our understanding of the annual fluctuation of metal solubility during the year (seasonal variation).

In addition to the experimental work, the results obtained by various model concepts presented here are compared to those from more mechanistic models An assemblage model, which comprises several models for the different (metal) binding phases in soil, was also used to calculate the metal concentrations in solution for the same data sets used for the transfer functions (Bonten et al., in prep). This model uses the NICA-Donnan model for both soil organic matter and DOC. For clay a non specific Donnan model was used and for adsorption to AL and Fe a diffuse double layer model was used. Model parameters were all taken from literature. In the database sufficient parameters were measured that allow for a comparison of both approaches (i.e. the statistical versus the mechanistic). This comparison will be used to evaluate the applicability of both types of models at different scale levels (Bonten et al., in prep).

Furthermore, the transfer functions derived will be improved in terms of more general European wide applicability by including data from other countries, notably the UK, Russia, Switzerland and Germany (Groenenberg et al., in prep.).

# References

Boekhold, A.E. 1992. Field scale behaviour of cadmium in soil. Ph.D. Thesis Wageningen Agricultural University, Wageningen, the Netherlands.

Bonten, L.T.C., J.E. Groenenberg, and P.F.A.M. Römkens. Comparison of different modelling approaches to estimate solid solution partitioning of heavy metals. Alterra report, In prep.

Chardon, W.J. 1984. Mobility of Cadmium in soils. Ph.D. Thesis Wageningen Agricultural University, Wageningen, the Netherlands.

Dai, K.H., M.B. David, G.F. Vance, J.W. McLaughlin, and I.J. Fernandez. 1996. Acidity characteristics of soluble organic substances in spruce-fir forest floor leachates. Soil Sci., 161: 694-704.

De Groot, A.C., W.J.G.M. Peijnenburg, M.A.G.T. vand en Hoop, R. Ritsema, and R.P.M. van Veen. 1998. Heavy metals in Dutch field soils: an experimental and theoretical study on equilibrium partitioning. Institute for Public Health and the Environment (RIVM). Report no.607220001, Bilthoven, the Netherlands.

De Rooij, N.M., and M.P.J.M. Kroot. 1991. CHARON Manual, Delft Hydraulics, Delft, the Netherlands.

De Vries, W., G. Schütze, S. Lofts, M. Meili, P.F.A.M. Römkens, R. Farret, L. de Temmerman, and M. Jakubowski. 2003. Critical limits for cadmium, lead and mercury related to ecotoxicological effects on soil organisms, aquatic organisms, plants, animals and humans. *In*: Proceedings Expert meeting on Critical Limits for Heavy Metals and Methods for their Application. Berlin, December 2-4, 2002. Umweltbundesambt, Texte no.47-3, pp.29-78

Elzinga, E.J., B. van den Berg, J.J.M. van Grinsven, and F.A. Swartjes. 1997. Freundlich adsorption isotherm equations for Cd, Cu and Zn in soils based on literature data (In Dutch w. English sum.). Institute for Public Health and the Environment (RIVM). Report no.711501001, Bilthoven, the Netherlands.

Elzinga, E.J., J.J.M. Van Grinsven, and F.A. Swartjes. 1999. General purpose Freundlich isotherms for cadmium, copper and zinc in soils. Eur. J. of Soil Sci., 50, 139-149.

Gooddy, D.C., P. Shand, D.G. Kinniburgh and W.H. van Riemsdijk. 1995. Field-based partition coefficients for trace elements in soil solutions. Eur. J. of Soil Sci., 46: 265-285.

Groenenberg, J.E., P.F.A.M. Römkens, E. Tipping, T. Pampura, W. de Vries and G. Schütze. 2003. Transfer functions for the calculation of critical loads for lead and cadmium. In: Proceedings Expert meeting on critical limits for heavy metals and methods for their application. Berlin 2-4 December 2002, held under the UN/ECE Convention on Long-Range Transboundary Air Pollution. UBA-Texte 47/03, Umweltbundesambt, Berlin.

Groenenberg, J.E., P.F.A.M. Römkens, E. Tipping, T. Pampura and W. de Vries, 2004. Transfer functions for the solid solution partitioning of Cd, Cu, Pb and Zn, Review, synthesis of data-sets, derivation and validation. Alterra report, In prep.

Holm, P.E., Th. Christensen, S.E. Lorenz, R.E. Hamon, H.C. Domingues, E.M. Sequeira, and S.P. McGrath. 1998. Mesaured soil water concentrations of cadmium and zinc in plant pots and estimated leaching outflows from contaminated soils. Water, Air Soil Pol., 102: 105-115.

Ingwersen, J., Th. Streck, J. Utermann, and J. Richter. 2000. Ground water preservation by soil protection: determination of tolerable total Cd contents and Cd breakthrough times. J. Plant Nutr. Soil Sci., 163:31-40.

Kros, J., S.C. Bos, P. Domburg, J.H. Faber, J.E. Groenenberg, T.C. Klok, W.C. Ma, W.G.H. Ogg, H.R.G. de Ruiter, W. de Vries, and J.G. Wesseling. 2001. Ontwikkeling van een bodembeoordelingssysteem voor natuurontwikkeling op met nutriënten en zware metalen verontreinigde gebieden; beslissingsondersteunend systeem bodembeoordeling; fase 1: BONANZA. Gouda, SKB, 2001. SKB Rapp. SV-004, 66 p.

McBride, M.B. 1994. Environmental chemistry of soils. Oxford University Press, New York, USA. 406 pp.

Otte, J.G., P.F.A.M. Römkens, A. Tiktak, and W. de Vries. 2000. Partition relationships for heavy metals (Cd, Cu, Pb, Zn) in diffuse polluted Dutch soils (In Dutch w. English sum.). PGBO report no. 30, PGBO - Wageningen, the Netherlands.

Quideau, S.A., and J.G. Bockheim. 1997. Biogeochemical cycling following planting to red pine on a sandy prairie soil. J. Environ. Qual., 26: 1167-1175.

Reinds, G.J., J. Bril, W. de Vries, J.E. Groenenberg, and A. Breeuwsma. 1995. Critical loads and excess loads of cadmium, copper and lead for European forest soils. SC-DLO report no. 96, Wageningen, the Netherlands.

Römkens, P.F.A.M. 1998. Effects of land use change on organic matter and trace metal solubility in soils. Ph.D. thesis University of Groningen, the Netherlands.

Römkens, P.F.A.M. and J. Dolfing. 1998. Effect of Ca on the solubility and molecular size distribution of DOC and Cu binding in soil solution samples. Env. Sci. Technol., 32: 363-369.

Römkens, P.F.A.M., and W. Salomons. 1998. Cd, Cu and Zn solubility in arable and forest soils: consequences of land use changes for metal mobility and risk assessment. Soil Sci., 163(11): 859-871.

Römkens, P.F.A.M., L.T.C. Bonten, R.P.J.J. Rietra, J.E. Groenenberg, A.C.C. Plette, and J. Bril. 2003. Leaching of heavy metals from agricultural soils. Estimation of the contribution of leaching from agricultural soils to the load of surface waters: model setup and results. [In Dutch]. Alterra report 791, Wageningen, The Netherlands. 94 pp.

Sauvé, S., M.B. McBride, W.A. Norvell, and W.A. Hendershot. 1997a. Copper solubility and speciation of in situ contaminated soils: Effects of copper level, pH and organic matter. Water, Air, Soil Pol., 100 (1-2) : 133-149.

Sauvé, S., M.B. McBride, and W.A. Hendershot. 1997b. Speciation of lead in contaminated soils. Env. Pol., 98 (2) : 149-155.

Sauvé, S., W.A. Norvell, M.B. McBride, and W.A. Hendershot. 2000*a*. Speciation and complexation of cadmium in extracted soil solutions. Env. Sci. Technol. 34 (2) : 291-296.

Sauvé, S., W.A.. Hendershot, and H.E. Allen. 2000*b*. Solid-solution partitioning of metals in contaminated soils: dependence on pH, total metal burden, and organic matter. Env. Sci. Technol., 34(7):1125-1131.

Selim, H.M., and D.L. Sparks. 2001. Heavy metals release in soils. Lewis Publishers, Boca Raton, USA.

Seuntjens, P. 2000. Reactive solute transport in heterogeneous porous media: cadmium leaching in acid sandy soils. Ph.D. thesis University of Antwerpen, Belgium, 236 pp..

Tipping, E., J. Rieuwerts, G. Pan, M.R. Ashmore, S. Lofts, M.T.R. Hill, M.E. Farago, and I. Thornton. 2003. The solid-solution partitioning of heavy metals (Cu, Zn, Cd, Pb) in upland soils of England and Wales. Env. Poll. 125:213-225.

Tye, A.M., S.D. Young, N.M.J. Crout, H. Zhang, S. Preston, V.L. Barbosa-Jefferson, D. Davison, S.P. McGarth, G.I. Paton, K. Khilham, and L. Resende. 2003. Predicting the activity of $Cd^{2+}$ and $Zn^{2+}$ in soil pore water from the radio-labile metal fraction. Geochim. Cosmochim. Act. 67:375-385.

Van Driel, W., B. Van Luit, W. Schuurmans, W. de Vries, M.J.J. Stienen, and G. Vos. 1988. Heavy metals in floodplain soils and crops along the river Meuse, Geul and Roer. Institute for Soil Fertility Research, Haren (Gr.), the Netherlands. Maasoever Report part 1,2 and 3.

Wiersma, D., B.J. van Goor, and N.G. van der Veen. 1986. Cadmium, lead, mercury and arsenic concentrations in crops and corresponding soils in the Netherlands. Journal of Agriculture and Food Chemistry, 34, 1067-1074.

Yin, Y., C.A. Impellitteri, S-J. You, and H.E. Allen. 2002. The importance of organic matter distribution and extract soil:solution ratio on the desorption of heavy metals from soils. Sci. Tot. Environ., 287:107-119.

### *Additional sources on partition equations and applications from the Netherlands and Belgium*

Hoop, van den, M.A.G.T. 1995. Metal speciation of Dutch soils; field based partition coefficients for heavy metals at background levels. Institute for Public Health and the Environment (RIVM). Report no. 719101013, Bilthoven, the Netherlands.

Janssen, R.P.T., P.J.Pretorius, W.J.G.M. Peijnenburg, and M.A.G.T. van den Hoop. 1996. Determination of field-based partition coefficients for heavy metals in Dutch soils and the relationships of these coefficients with soil characteristics. Institute for Public Health and the Environment (RIVM). Report no. 719101023, Bilthoven, the Netherlands.

Lee, S.Z., H.E. Allen, C.P. Huang, D.L. Sparks, P.F. Sanders, and W.J.G.M. Peijnenburg. 1996. Predicting soil-water partition coefficients for cadmium. Env. Sci. Technol., 30:3418-3424.

Makaske, G.B., H.A. Vissenberg, J.J.M. van Grinsven, A.. Tiktak, and F.J. Sauter. 1995. METRAS: A one-dimensional model for assessment of leaching of trace metals from soils. Model description and results of a case study for Cd in 'De Kempen'. Institute for Public Health and the Environment (RIVM). Report no. 715501005, Bilthoven, the Netherlands.

Pretorius, P.J., R.P.T. Janssen, W.J.G.M. Peijnenburg, and M.A.G.T van den Hoop. 1996. Chemical equilibrium modelling of metal partitioning in soils. Institute for Public Health and the Environment (RIVM). Report no. 719101024, Bilthoven, the Netherlands.

Temminghoff, E.J.M. 1998. Chemical speciation of heavy metals in sandy soils in relation to availability and mobility. Ph.D. thesis, Wageningen Agricultural University, the Netherlands.

Tiktak, A. 1999. Modelling non-point source pollutants in soils. Ph.D. thesis University of Amsterdam, the Netherlands.

Wezel A.P. van, W. de Vries, M. Beek, P. Otte, J. Lijzen, M. Mesman, P. van Vlaardingen, J. Tuinstra, M van Elswijk, P.F.A.M. Romkens and L.T.C. Bonten. 2003. Soil-use remediation objectives for agriculture, nature and sediments. Technical derivation of values. [In Dutch] RIVM report. 711701031. 161 pp. in Dutch.

# Appendix 1 Estimation of the reactive metal content from the total metal content (Aqua Regia) and soil properties

### *The concept of a reactive fraction in soils*

The rationale to use the reactive metal content to calculate the dissolved metal concentration or the free metal ion activity is based on the fact that part of the metals stored in soils are present in what is called the 'inert' fraction. Metals bound in silicate structures in soil forming minerals or clay particles are not likely to contribute to the metal fraction in the soil solution. This of course depends on the time scale involved. One might argue that due to weathering processes eventually all metals would become available to some extent. The models here however are developed to cover a time span of years to decades and the influence of weathering usually involves time spans of hundreds to thousands of years. This means that soils with a high amount of unstable minerals with high weathering rates (e.g. recently deposited volcanic ashes) are not covered here. In these soils, weathering and its contribution to the reactive metal content is such that it cannot be ignored. In general however, it can be safely assumed that for the majority of European soils, weathering rates are limited and its contribution to the reactive fraction insignificant. The assumption that the reactive metal content as determined by the dilute acid extraction is representative for the model time spans involved here seems, therefore, reasonable.

In this study the 0.43 $mol.l^{-1}$ $HNO_3$ extractable metal content was used as a measure of the reactive metal content in soils. As of now no general consensus has been reached as to what extract is the best or most suitable way to extract those metals that participate in the chemical equilibrium reactions in soil. Alternative methods that could be applicable include the EDTA extraction. Although there is a close correlation between EDTA extractable metals and 0.43 $mol.l^{-1}$ $HNO_3$ extractable metals we chose to use 0.43 $mol.l^{-1}$ $HNO_3$. This is mainly based on the observation that EDTA dissolves most of the organic matter fraction. One could argue that 0.43 $mol.l^{-1}$ $HNO_3$ dissolves most of the amorphous iron and or aluminium but it is believed that this is not unrealistic at extreme pH values in soils. At least this is more likely than assuming that all metal included in the organic matrix will be available. In addition to this the 0.43 $mol.l^{-1}$ $HNO_3$ extraction is rapid, easy and reproducible. The only constraint is the application in the calcareous soils with a high (i.e. > 10%) lime content. In these soils, buffering of the acid by the lime prevents extraction of metals at low pH. In the soils tested so far this did not seem to be a problem (pH was measured in the final extracts and was always lower than 3).

The amount of metals bound in the residual (inert) fraction can be calculated as the difference between the Aqua Regia and 0.43 $mol.l^{-1}$ $HNO_3$ extractable metal content. This means that three different pools can be distinguished based on the extraction used here:

| Abbreviation | Equivalent to | Extractant or calculation |
|---|---|---|
| $M_{total}$ | Total or pseudo total met content | Aqua Regia (AR) |
| $M_{reactive}$ | Reactive metal content | 0.43 $mol.l^{-1}$ $HNO_3$ |
| $M_{inert}$ | Inert or occluded metal pool | AR - 0.43 $mol.l^{-1}$ $HNO_3$ |

***Model used to estimate reactive metal content***

To obtain estimates of the reactive metal content 2 different methods can be used which, in practice result in similar estimates:

1. A direct estimate of $M_{reactive}$ using $M_{total}$ and soil properties.
2. An estimate of $M_{inert}$ using $M_{total}$ and a combination of soil properties followed by the calculation of $M_{reactive}$ from $M_{total}$ and $M_{inert}$.

Here we chose to obtain direct estimates of the reactive fraction of $M_{reactive}$ which was obtained using similar parameters that were used to calculate the dissolved metal content: clay, organic matter and the Aqua Regia metal content:

$$\log(Me_{reactive}) = a + b*\log(\text{organic matter}) + c*\log(\text{clay}) + d*\log(Me_{aqua\ regia}) \qquad [A1]$$

Two reasons can be mentioned why we chose these soil properties:

1. both in the clay and the organic matter fraction a part of the total metal load is occluded in the matrix (either the organic or the mineral matrix); in soils with a similar total metal content but different clay and organic matter content, the amount of the reactive metal content will be different (i.e. decrease with an increase in organic matter and clay).
2. These soil properties are already available from the soil map so the concept can be applied on a national or regional scale.

A log-linear relation was chosen because the use of a linear regression resulted in negative values for the estimate of the reactive metal content in the 'low' range. In fact the following options exist to calculate the reactive fraction in soils:

1. A direct estimate of the reactive fraction as a linear function of soil properties
2. A direct estimate of the reactive fraction as a logarithmic function of soil properties
3. An estimate of the inert fraction as a linear function of soil properties; the reactive fraction then can be calculated as the difference between the total and the inert fraction
4. As option 3 but using log transformed data.

Analysis of the data showed that option 2 resulted in the best estimates of $Me_{reactive.}$ This was concluded based on the sum of the residuals from all four methods which was smallest for option 2 for all metals tested here. The major disadvantage of option 1 and 3 is that estimates of the reactive fraction in the low range are often negative, which is due to either the negative intercept of the linear function of the reactive metal content or due to the fact that the estimate of the inert fraction is larger than that of the total metal content which also renders a negative reactive fraction. Using a standard intercept of zero in the regression function reduces the number of negative estimates but does not eliminate negative estimates. The estimate of the reactive fraction using option 4 resulted in a larger standard error of Y-estimated compared to the direct log transformed estimate of the reactive fraction. It also resulted in a general overestimation of the reactive fraction in the entire range.

***Results obtained***

In Table A1.1 an overview of the regression function obtained using option 2 are shown (direct estimate of $Q_{reactive}$ after log transformation)

Table A1.1 Values for the coefficients in the relationship relating reactive and total soil concentrations of Cu, Zn, Cd and Pb (estimates are in $mg.kg^{-1}$).

| Metal | INT | OM | Clay | Q-AR | $R^2$ | se-$y_{est}$ 1) |
|---|---|---|---|---|---|---|
| Cu | -0.331<br>0.399 2) | 0.023 | -0.171 | 1.152 | 0.93 | 0.13 |
| Zn | -0.703<br>0.428 2) | 0.183 | -0.298 | 1.235 | 0.96 | 0.16 |
| Cd | -0.089<br>0.289 2) | 0.022 | -0.062 | 1.075 | 0.96 | 0.11 |
| Pb | -0.263<br>0.210 2) | 0.031 | -0.112 | 1.089 | 0.92 | 0.16 |

1) on a logarithmic basis

2) using this intercept yields the outcome in $mol.kg^{-1}$ instead of $mg.kg^{-1}$

The log model presented here, however has a few disadvantages as well. An important check on the quality of the prediction is not only the absolute value of the prediction and the deviation from the measured levels but also the ratio between reactive and total content (the percentage of the reactive fraction as part of the total metal content). For all metals this ratio ranges between 0 and 1 (representing 0% reactive and 100% reactive respectively). This ratio changes with an increase in the total metal content: the ratio tends to increases with an increase in the total metal content which indicates that the reactivity of metal in polluted soils is higher than that of unpolluted soils.

However, also at a given total metal content, the ratio varies considerably which reflects the effect of soil type and soil properties (organic matter and clay) on the reactivity of metals in soils. This observation (variable reactivity ratios ate a given total metal content) is illustrated in Table A1.2 for Zn.

Table A1.2 Effect of amount of total soil Zn (Aqua Regia) on the ratio reactive/total.

| | Zn content in soil (total - Aqua Regia in $mg.kg^{-1}$) | | | | |
|---|---|---|---|---|---|
| | 0 to 50 | 50 - 100 | 100-200 | 200 - 500 | > 500 |
| Minimum: | 0.05 | 0.16 | 0.21 | 0.22 | 0.22 |
| Maximum: | 0.92 | 0.83 | 0.80 | 0.96 | 0.96 |
| Median: | 0.30 | 0.34 | 0.42 | 0.49 | 0.61 |
| Number of samples: | 57 | 45 | 69 | 67 | 66 |

The data in Table A1.2 show that for Zn (as well as for the other metals)

1. The reactivity increases with an increase in the aqua regia extractable metal content.
2. The ratio of reactive to total can still increase to almost 1, i.e. 100% reactive in soils with a low total metal content.

This illustrates also that using aqua regia as the sole predictor of the reactive metal content (which also give quite high correlation coefficients) is insufficient to predict the range in reactivity at a given total metal content. To check whether the different model concepts presented earlier to predict the reactive metal content are able to represent the measured range of the ratio reactive to total, all models were used to predict the reactive metal content. These predictions were then used to calculate the ratio reactive to total. The outcome of all models were then grouped into classes of reactivity (ranging from 0 to 1 in steps of 0.05) and plotted in a bar diagram. This is shown for Cu in Figure A1.1. Models tested here include:

1. The logarithmic model described earlier (data shown in Table A1.1) using organic matter, clay and AR as predictors
2. The logarithmic model using AR only (model coefficients not shown here)
3. A combination of the logarithmic and the linear model (model coefficients not shown here).

The linear model alone was not used since it results in negative estimates of the reactive metal content, especially in the 'low' range (for Cu this means Cu -total < 15 mg.kg$^{-1}$). Here a combination of the linear and logarithmic model was used. The reactive fraction was calculated using a linear regression model; in case of negative estimates the logarithmic model was used.

The results presented in Figure A1.1 show that the combination of the linear and logarithmic model is able to reproduce best the measured range in the ratio reactive to total. However, for Cd and Pb it was not possible to reproduce the measured range in ratios of reactive to total (in contrast to Cu as shown before), neither using a logarithmic model nor a combination of the linear and the logarithmic model (data not shown)

This led us to the conclusion that no single concept for all metals exist which gives the best fit for all variables. An additional aspect is that these equations have to be used also to calculate the total metal content once the reactive metal content is known. Calculations of critical metal levels in soil can be based on a critical limit in the soil solution. Using the partition equation presented in the chapters of this report, results in an estimate of the reactive metal content which has to be converted into a total or pseudo total (Aqua Regia) metal content. This is illustrated in Figure A1.2

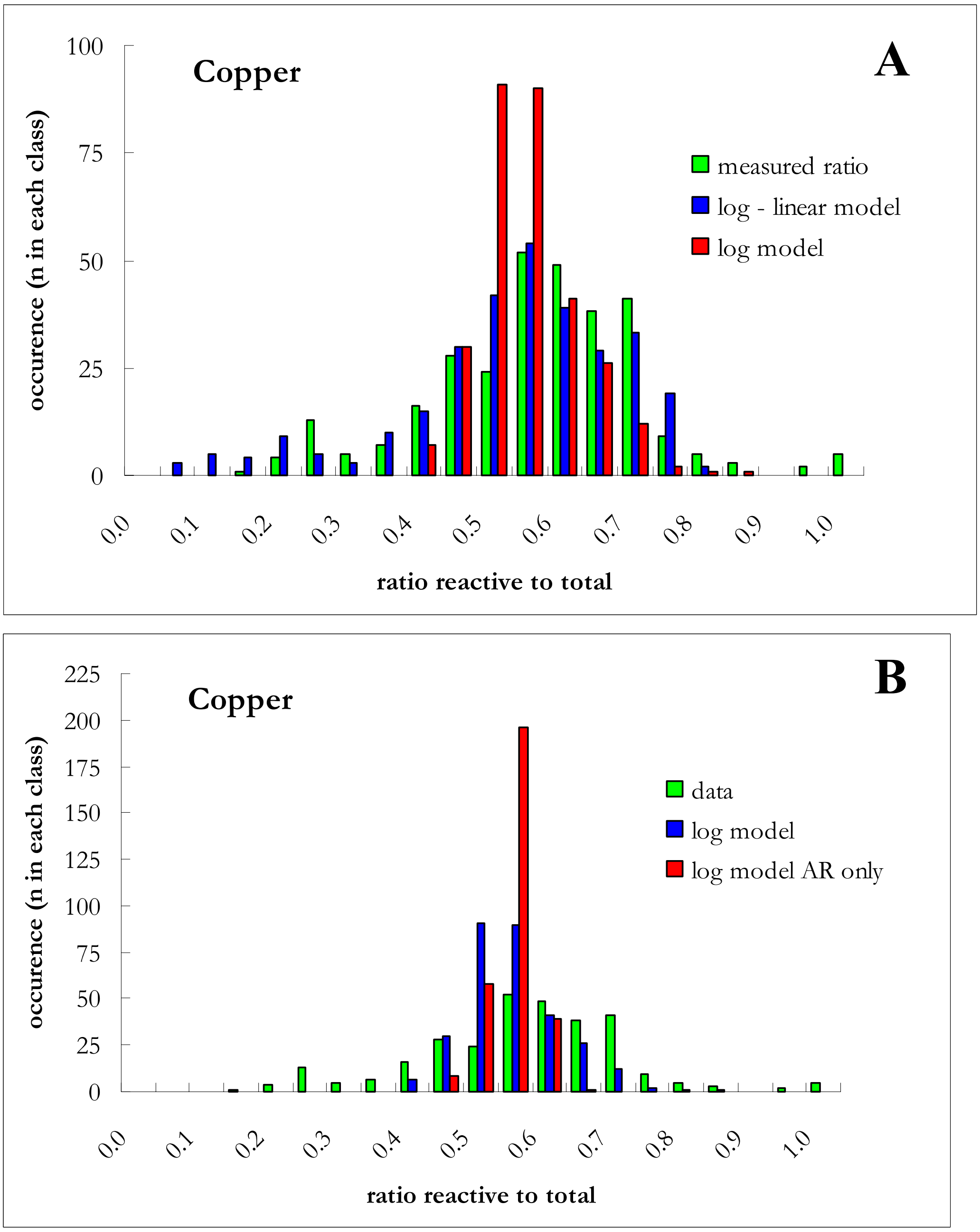

*Figure A1.1Measured versus predicted ratio of reactive tot metal grouped in classes. A: logarithmic model according to equation A1, fig B: logarithmic model using AR only*

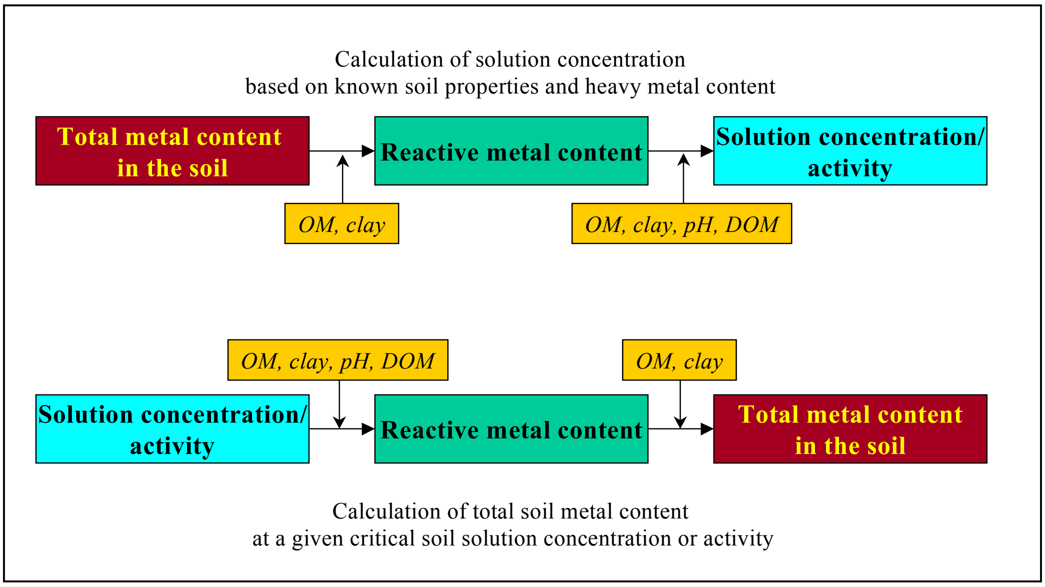

*Figure A1.2 Application of transfer functions to calculate 'back and forth' from solid to solution and vice versa.*

When looking at the estimation of the AR metal content based on equations presented in Table A1.1 it was concluded that the logarithmic regression equation gave the best estimate of the measured AR metal content when using the measured reactive metal content together with the organic matter and clay content as input parameters.

The final conclusion from this analysis was therefore that although the combined linear logarithmic model gave the best estimate of the ratio reactive to total and a good estimate of the reactive metal content, the logarithmic model was chosen as the sole model to be used. This was mainly based on the observation that the sum of squares of the logarithmic model when back calculating the total metal content from the reactive metal content was lower than that of the linear - logarithmic model.

### *Comparison of modelled and measured reactive metal contents*

In Figure A1.3 and A1.4 the predicted reactive metal content and the predicted AR metal content respectively are shown in comparison to measured values using the regression coefficients from Table A.1.1

To obtain an estimate of the reactive metal content by regression the same objections that were raised in the conclusions of the main text about application outside the range of measured date are valid here as well. In sedimentary soils, a considerable part of the metals are adsorbed onto the soil and will be available in the dilute acid. This is not true for metals in minerals that are present in soils derived from parent material in situ (i.e. rocks). Therefore, care should be taken to use these equations for significantly different soils. Validation of the equations therefore for soils developed from slid parent material is necessary.

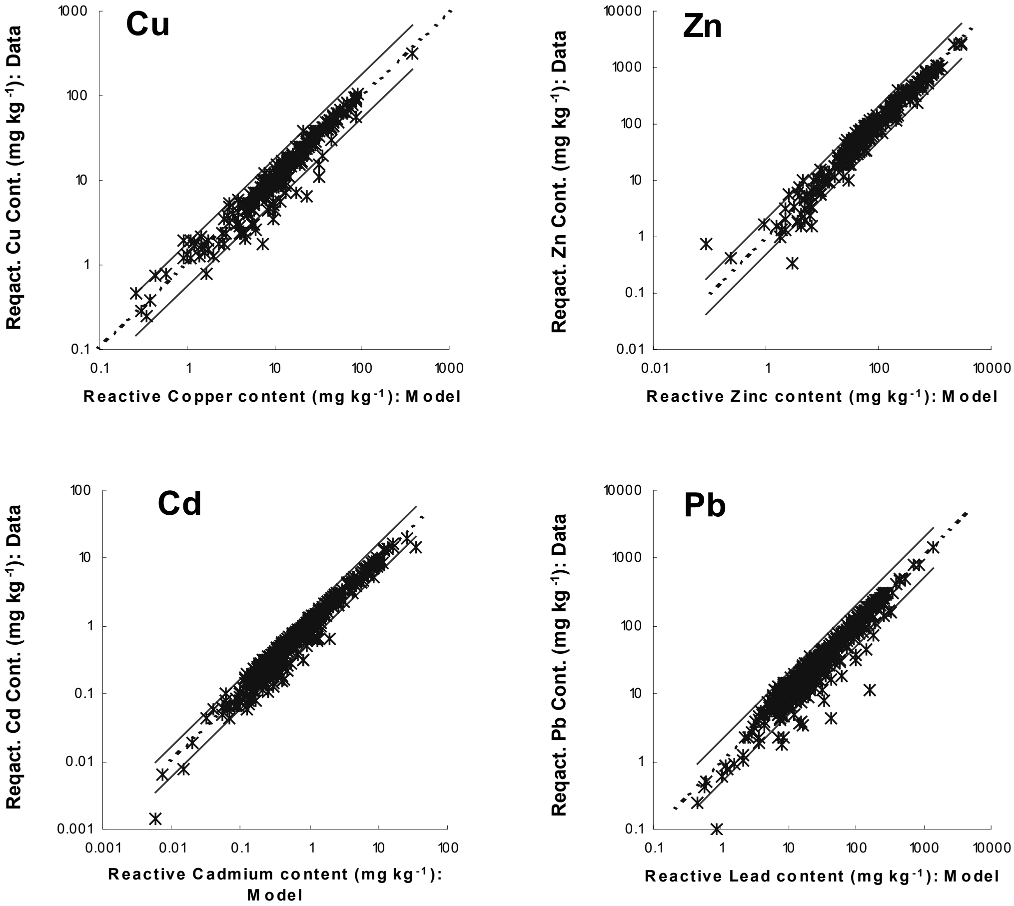

*Figure A1.3. Measured and predicted values of the reactive lead content using equation A1.*

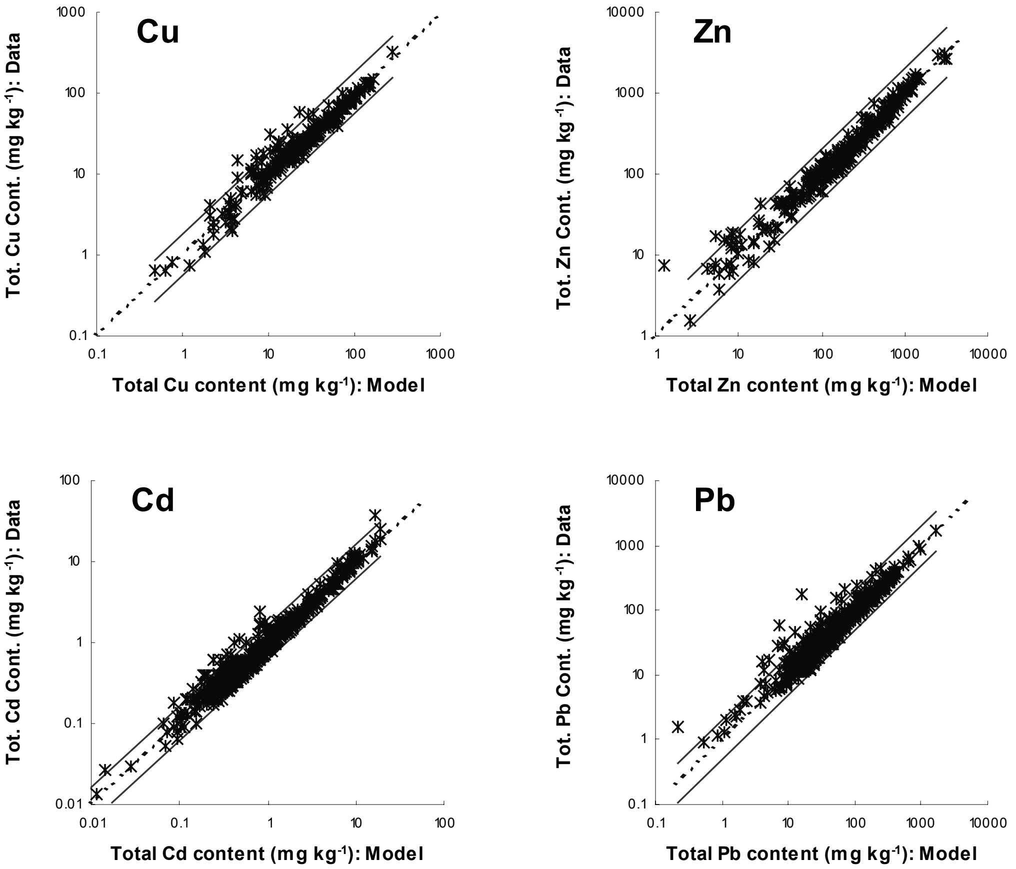

*Figure A1.4 Measured and predicted values of the total heavy metal content using equation A1.*

# Appendix 2 Estimation of DOC from soil parameters

For elements like Cu and Pb it is essential to know the amount of dissolved organic carbon. As was shown in chapter 3, the equation without DOC for Cu and Pb was clearly insufficient. Also for the calculation of the activity, the amount of DOC has to be known. If data on DOC are non-existent, estimates can be made using a combination of soil parameters. Important parameters that are known to influence the concentration of DOC in solution include:

1. organic matter
2. electrical conductivity (EC)
3. pH
4. the amount of Fe and Al (hydr)oxides
5. the ratio soil:solution applied in the extract

In this study data from soil extracts were used. A disadvantage of a database containing extracts with different soil:solution ratios is that this ratio will affect the amount of dissolved organic carbon in solution. Here we derive a few applicable regression equations that can be used to obtain estimates of DOC in solution. Part of the database was based on data where only 1:2 soil:solution extracts were used. A separate part however contained data with varying soil:solution ratios. This set was be used to evaluate the magnitude of this effect.

The reason for presenting different equations is that if data on for example EC are know, they can be used to obtain more precise estimates. In those cases where it is not measured a more simple equation can be used taking into account organic matter only. In Table A2.1 the results from the regression analysis is shown.

Table A2.1 Equations to calculate DOC in solution

| Data Set | n | soil: solution | Int | OM % | pH | EC | Fe/Al mmol.kg$^{-1}$ | Solid* | Se(Y) | R$^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| SEO | 840 | 1:2 | 1.30 | 0.65 | - | - | - | - | 0.37 | 0.55 |
| SEO | 840 | 1:2 | 2.04 | 0.73 | -0.17 | - | - | - | 0.31 | 0.68 |
| SEO | 840 | 1:2 | 2.25 | 0.75 | -0.20 | -0.30 | - | - | 0.30 | 0.71 |
| RIVM/IB | 332 | variable | 1.67 | 0.68 | - | - | - | 1.33 | 0.40 | 0.49 |
| RIVM/IB | 332 | variable | 2.27 | 0.61 | -0.09 | - | - | 1.37 | 0.38 | 0.54 |
| RIVM/IB | 332 | variable | 2.51 | 0.88 | -0.07 | - | -0.30 | 1.39 | 0.37 | 0.56 |
| Combined | 1172 | variable | 2.00 | 0.63 | - | - | - | 1.57 | 0.39 | 0.52 |
| Combined | 1172 | variable | 2.66 | 0.70 | -0.15 | - | - | 1.52 | 0.33 | 0.65 |
| Combined | 1172 | variable | 2.74 | 0.73 | -0.15 | - | -0.06 | 1.53 | 0.33 | 0.65 |
| Combined | 1172 | variable | 3.70 | 0.71 | -0.17 | -0.29 | - | 1.47 | 0.31 | 0.69 |

*Solid: Solution Ratio

In both datasets the coefficient for Organic matter is roughly the same and ranges from 0.61 to 0.75. Using both organic matter and pH improves the fit considerably in both cases. It became clear that the amount of solids used in the extractions is of utmost importance. In the first dataset, a single 1:2 soil:solution ratio was used whereas in the second database various soil:solution ratios were used ranging from very high ratio's in case of in situ soil solution that were obtained by centrifugation to 1:10 (soil:solution) extracts in dilute salt solutions. In the combined data base this effect had to be considered so the ratio soil:solution is included. In this case it was calculated as the amount of solids per litre divided by the weight of the solids and the water:

Solid/solution = gram solids per liter/(gram solids per liter + 1000)

Although the effect of EC is significant, and probably reflects the impact of the bi- and trivalent cations on the solubility of DOC (DOC decreases with an increase in EC which is consistent with observation in the literature, see Römkens and Dolfing (1998), it is often not measured routinely and not available from soil maps. If available, it should be included.

The effect of Fe and Al hydroxides was less than expected. Although it was a significant model parameter when included in the RIVM/IB database, it did not improve the overall estimate of DOC very much as was reflected by the overall $R^2$ and the standard error of the estimate, both of which did increase only marginally. For practical purposes we suggest that the regression with pH and organic matter are to be used as the 'best available' estimate:

Log(DOC) = 2.66 + 0.70*log(organic matter) -0.15*pH + 1.52*log(solid/sol.)

# Appendix 3 Effect of detection limit on partition equations

In the original database all data were scaled such that negative values were corrected for. This was done because elimination of all values below detection would results in a skewed database for some elements. In many cases a considerable part of the high pH samples would have been left out reducing the applicability of the model in that range.

The procedure followed here was that the lowest (measured) value (which had a negative value) was set to zero and all measured values between this most negative value and the detection limit were rescaled using a quadratic regression function with 0 and the detection limit as end-points.

However, for certain elements this results in a considerable noise in the low range. To evaluate what the effect is of leaving out the data below detection limit, a regression was performed using only data with measured values higher than the detection limit. For some elements 2 detection limits were used which was a result of using 2 different ICP instruments.

In the table on the following pages the results of the analysis taking into account the detection limit are shown. The following tables are included:

| | |
|---|---|
| Table A3.1 | Concentration transfer function (direct regression of C-Me versus soil properties) |
| Table A3.2 | Kf transfer function based on concentration |
| Table A3.3 | Kf transfer function based on concentration including Fe-Al oxides |
| Table A3.4 | Activity transfer function (direct regression of *a*Me vs. soil properties) |
| Table A3.5 | Activity transfer function (direct regression of *a*Me vs. soil properties without clay content) |
| Table A3.6 | Kf transfer function based on activity |
| Table A3.7 | Kf transfer function based on activity without clay content |

### *Concentration Transfer function*

Table A3.1 Effect of detection limits on Concentration Transfer function (model II only).

| Metal | Coefficients | | | | | | se-Y | $R^2$ | #[1] |
|---|---|---|---|---|---|---|---|---|---|
| | INT | Q | OM | pH | clay | DOC | | | |
| **Cd** | | | | | | | | | |
| All data | 4.9 | 1.27 | -0.73 | -0.39 | -0.48 | 0.08 | 0.53 | 0.67 | 1466 |
| dt[2] 0.7 | 3.84 | 1.15 | -0.59 | -0.3 | -0.36 | 0.1 | 0.47 | 0.58 | 980 |
| dt 1.1 | 3.65 | 1.13 | -0.59 | -0.28 | -0.3 | 0.08 | 0.45 | 0.57 | 863 |
| **Cu** | | | | | | | | | |
| All data | -0.24 | 0.82 | -0.56 | -0.08 | -0.33 | 0.55 | 0.43 | 0.55 | 1466 |
| dt 1.0 | -0.38 | 0.78 | -0.57 | -0.07 | -0.31 | 0.54 | 0.4 | 0.57 | 1421 |
| **Ni** | | | | | | | | | |
| All data | 2.78 | 0.91 | -0.68 | -0.4 | -0.22 | 0.28 | 0.39 | 0.67 | 1466 |
| dt 1.1 | 2.54 | 0.9 | -0.66 | -0.34 | -0.27 | 0.24 | 0.34 | 0.66 | 1386 |
| dt 3.5 | 2.13 | 0.85 | -0.63 | -0.27 | -0.25 | 0.15 | 0.31 | 0.57 | 1153 |
| **Pb** | | | | | | | | | |
| All data | -0.22 | 0.69 | -0.73 | -0.2 | -0.34 | 0.35 | 0.63 | 0.51 | 1466 |
| dt 7.8 | -0.86 | 0.55 | -0.66 | -0.16 | -0.32 | 0.44 | 0.55 | 0.51 | 813 |
| dt 13 | -1.3 | 0.43 | -0.83 | -0.14 | -0.2 | 0.52 | 0.55 | 0.45 | 535 |
| **Zn** | | | | | | | | | |
| All data | 4.26 | 1.08 | -0.46 | -0.5 | -0.51 | 0.21 | 0.5 | 0.78 | 1466 |
| dt 0.9 | 4.28 | 1.07 | -0.48 | -0.51 | -0.46 | 0.19 | 0.48 | 0.79 | 1458 |
| dt 7.0 | 4.23 | 1.04 | -0.47 | -0.5 | -0.41 | 0.12 | 0.42 | 0.81 | 1405 |

[1]: number of samples remaining in the database

[2] dt: detection limit in µg.l$^{-1}$. In this case only data points with measured concentrations exceeding this level were included in the regression analysis.

Table A3.2 Partition equations for the calculation of Kf based on concentration. Only model CI was tested (including DOC)

| Metal | Coefficients | | | | | | se-Y | $R^2$ |
|---|---|---|---|---|---|---|---|---|
| | INT | DOC | OM | pH | clay | n-opt | | |
| **Cd** | | | | | | | | |
| All data | -4.75 | -0.05 | 0.61 | 0.26 | 0.29 | 0.54 | 0.33 | 0.80 |
| Cd > 1.1 | -4.91 | -0.07 | 0.57 | 0.26 | 0.25 | 0.49 | 0.30 | 0.78 |
| **Cu** | | | | | | | | |
| All data | -2.62 | -0.27 | 0.60 | 0.12 | 0.23 | 0.59 | 0.34 | 0.65 |
| Cu > 1.0 | -2.33 | -0.31 | 0.63 | 0.12 | 0.24 | 0.66 | 0.34 | 0.67 |
| **Ni** | | | | | | | | |
| All data | -4.72 | -0.13 | 0.72 | 0.30 | 0.39 | 0.54 | 0.28 | 0.88 |
| Ni > 3.5 | -4.40 | -0.09 | 0.70 | 0.26 | 0.41 | 0.61 | 0.27 | 0.88 |
| **Pb** | | | | | | | | |
| All data | -2.36 | -0.23 | 0.95 | 0.22 | 0.07 | 0.73 | 0.55 | 0.59 |
| Pb > 13 | -1.11 | -0.61 | 0.96 | 0.22 | 0.33 | 1.00 | 0.63 | 0.58 |
| Pb > 7.8 | -1.02 | -0.45 | 0.91 | 0.23 | 0.31 | 1.08 | 0.65 | 0.56 |
| **Zn** | | | | | | | | |
| All data | -4.23 | -0.14 | 0.47 | 0.43 | 0.37 | 0.75 | 0.40 | 0.82 |
| Zn > 0.9 | -4.23 | -0.13 | 0.47 | 0.44 | 0.35 | 0.77 | 0.39 | 0.83 |

Table A3.3 Regression equations including Fe and Al oxides

| Metal | Coefficients | | | | | | | se(Y) | $R^2$ |
|---|---|---|---|---|---|---|---|---|---|
| | INT | FeAl | clay | OM | pH | DOC | n-opt | | |
| **Cd** | | | | | | | | | |
| All data | -5.33 | 0.36 | 0.16 | 0.46 | 0.25 | -0.05 | 0.52 | 0.31 | 0.81 |
| Cd > 1.1 | -5.33 | 0.32 | 0.12 | 0.43 | 0.25 | -0.04 | 0.50 | 0.29 | 0.79 |
| **Cu** | | | | | | | | | |
| All data | -3.07 | 0.35 | 0.12 | 0.45 | 0.12 | -0.27 | 0.60 | 0.33 | 0.67 |
| Cu > 1.0 | -2.81 | 0.37 | 0.12 | 0.47 | 0.12 | -0.31 | 0.67 | 0.33 | 0.69 |
| **Ni** | | | | | | | | | |
| All data | -4.73 | 0.00 | 0.39 | 0.71 | 0.30 | -0.13 | 0.54 | 0.28 | 0.88 |
| Ni > 3.5 | -4.45 | 0.06 | 0.39 | 0.68 | 0.26 | -0.08 | 0.62 | 0.27 | 0.88 |
| **Pb** | | | | | | | | | |
| All data | -2.97 | 0.59 | -0.09 | 0.71 | 0.23 | -0.25 | 0.79 | 0.56 | 0.61 |
| Pb >13 | -1.74 | 0.65 | 0.09 | 0.61 | 0.22 | -0.43 | 1.11 | 0.64 | 0.59 |
| n=1 >13 | -2.05 | 0.60 | 0.06 | 0.59 | 0.21 | -0.38 | 1.00 | 0.6 | 0.59 |
| **Zn** | | | | | | | | | |
| All data | -4.52 | 0.20 | 0.30 | 0.39 | 0.43 | -0.13 | 0.74 | 0.40 | 0.82 |
| Zn > 0.9 | -4.58 | 0.23 | 0.26 | 0.38 | 0.44 | -0.13 | 0.75 | 0.38 | 0.84 |

### *Activity Based Transfer functions, model II only (based on organic matter, clay and pH)*

First all records were sorted on concentration. All records below detection limit were left skipped and the remaining records (using activity) were used to derive the regression equation. In contrast to the concentration transfer functions little differences were observed between the transfer functions derived from the entire database and the one from the database with values higher than the detection limit. Here only the effect on the model using organic matter, clay and pH are shown.

Table A3.4 Partition equations for the direct estimation of activity based on the metal content, organic matter and clay content and pH.

| Metal | Coefficients | | | | | se-Y | $R^2$ | # |
|---|---|---|---|---|---|---|---|---|
| | Int | Q | clay | OM | pH | | | |
| **Cd** | | | | | | | | |
| All data | 5.26 | 1.31 | -0.42 | -0.87 | -0.46 | 0.53 | 0.70 | |
| Cd > 1.1 | 4.54 | 1.24 | -0.26 | -0.76 | -0.38 | 0.45 | 0.62 | 863 |
| **Cu** | | | | | | | | |
| All data | 1.91 | 0.78 | -0.23 | -0.78 | -0.70 | 0.46 | 0.84 | |
| Cu > 1.0 | 1.80 | 0.75 | -0.21 | -0.79 | -0.70 | 0.44 | 0.85 | 1421 |
| **Ni** | | | | | | | | |
| All data | 3.39 | 0.94 | -0.23 | -0.60 | -0.49 | 0.37 | 0.71 | |
| Ni > 3.5 | 2.56 | 0.89 | -0.25 | -0.62 | -0.34 | 0.29 | 0.64 | 1153 |
| **Pb** | | | | | | | | |
| All data | 1.67 | 0.68 | -0.23 | -0.90 | -0.70 | 0.62 | 0.78 | |
| Pb > 13 | 2.08 | 0.70 | -0.23 | -0.89 | -0.71 | 0.51 | 0.84 | 535 |
| Pb > 7.8 | 1.49 | 0.62 | -0.25 | -0.83 | -0.66 | 0.54 | 0.81 | 813 |
| **Zn** | | | | | | | | |
| All data | 5.02 | 1.12 | -0.45 | -0.53 | -0.64 | 0.52 | 0.80 | |
| Zn > 0.9 | 5.01 | 1.12 | -0.41 | -0.55 | -0.65 | 0.50 | 0.81 | 1458 |

To compare the results with isotherms from other sources, also models without clay have been derived

Table A3.5 Partition relationships based on organic matter, pH and Q only: direct estimate of aMe

| Metal | Coefficients | | | | se-Y | $R^2$ |
|---|---|---|---|---|---|---|
| | INT | Q | OM | pH | | |
| **Cd** | | | | | | |
| All data | 5.11 | 1.28 | -1.03 | -0.51 | 0.56 | 0.66 |
| Cd > 1.1 | 4.25 | 1.19 | -0.81 | -0.40 | 0.46 | 0.59 |
| **Cu** | | | | | | |
| All data | 1.82 | 0.76 | -0.86 | -0.73 | 0.47 | 0.83 |
| Cu > 1.0 | 1.74 | 0.73 | -0.87 | -0.73 | 0.45 | 0.84 |
| **Ni** | | | | | | |
| All data | 2.55 | 0.81 | -0.57 | -0.49 | 0.38 | 0.69 |
| Ni > 3.5 | 1.68 | 0.74 | -0.60 | -0.34 | 0.31 | 0.59 |
| **Pb** | | | | | | |
| All data | 1.93 | 0.72 | -1.03 | -0.73 | 0.63 | 0.77 |
| Pb > 13 | 2.07 | 0.68 | -1.01 | -0.73 | 0.52 | 0.84 |
| Pb > 7.8 | 1.64 | 0.63 | -0.96 | -0.70 | 0.55 | 0.80 |
| **Zn** | | | | | | |
| All data | 5.12 | 1.13 | -0.73 | -0.70 | 0.56 | 0.76 |
| Zn > 0.9 | 5.10 | 1.12 | -0.73 | -0.70 | 0.53 | 0.78 |

Table A3.6 Activity isotherms based on Kf approach

| Metal | Coefficients | | | | | se-Y | $R^2$ |
|---|---|---|---|---|---|---|---|
| | INT | OM | pH | clay | n-opt | | |
| **Cd** | | | | | | | |
| All data | -4.76 | 0.66 | 0.29 | 0.25 | 0.55 | 0.32 | 0.83 |
| Cd > 1.1 | -4.94 | 0.60 | 0.30 | 0.21 | 0.51 | 0.28 | 0.82 |
| **Cu** | | | | | | | |
| All data | -3.25 | 0.87 | 0.67 | 0.23 | 0.90 | 0.43 | 0.89 |
| Cu > 1.0 | -3.20 | 0.92 | 0.70 | 0.22 | 0.94 | 0.43 | 0.90 |
| **Ni** | | | | | | | |
| All data | -4.82 | 0.67 | 0.35 | 0.38 | 0.58 | 0.27 | 0.90 |
| Ni > 3.5 | -4.33 | 0.69 | 0.31 | 0.38 | 0.67 | 0.25 | 0.90 |
| **Pb** | | | | | | | |
| All data | -3.22 | 1.20 | 0.77 | 0.14 | 1.04 | 0.67 | 0.82 |
| Pb < 13 | -3.30 | 1.01 | 0.80 | 0.29 | 1.09 | 0.57 | 0.88 |
| Pb < 7.8 | -2.91 | 1.17 | 0.86 | 0.29 | 1.26 | 0.69 | 0.84 |
| n = 1 | | | | | | | |
| All data | -3.28 | 1.18 | 0.74 | 0.12 | 1.00 | 0.65 | 0.82 |
| Pb < 13 | -3.40 | 0.94 | 0.75 | 0.28 | 1.00 | 0.54 | 0.88 |
| Pb < 7.8 | -3.27 | 1.04 | 0.71 | 0.21 | 1.00 | 0.59 | 0.84 |
| **Zn** | | | | | | | |
| All data | -4.60 | 0.50 | 0.52 | 0.33 | 0.73 | 0.40 | 0.85 |
| Zn < 0.9 | -4.59 | 0.51 | 0.53 | 0.30 | 0.75 | 0.39 | 0.86 |

Table A3.7 Partition relationships based on organic matter, pH and Q only: estimate of Kf (activity)

| Metal | Coefficients | | | | se-Y | $R^2$ |
|---|---|---|---|---|---|---|
| | Int | pH | OM | n-opt | | |
| **Cd** | | | | | | |
| All data | -4.87 | 0.32 | 0.77 | 0.53 | 0.34 | 0.80 |
| Cd > 1.1 | -4.98 | 0.33 | 0.67 | 0.51 | 0.30 | 0.80 |
| **Cu** | | | | | | |
| All data | -3.28 | 0.70 | 0.97 | 0.90 | 0.45 | 0.88 |
| Cu > 1 | -3.21 | 0.73 | 1.02 | 0.95 | 0.45 | 0.89 |
| **Ni** | | | | | | |
| All data | -4.46 | 0.45 | 0.80 | 0.74 | 0.34 | 0.86 |
| Ni > 3.5 | -4.02 | 0.39 | 0.84 | 0.80 | 0.31 | 0.86 |
| **Pb** | | | | | | |
| All data | -3.30 | 0.76 | 1.23 | 1.00 | 0.66 | 0.82 |
| Pb > 13 | -3.40 | 0.86 | 1.20 | 1.12 | 0.60 | 0.87 |
| Pb > 7.8 | -3.07 | 0.89 | 1.29 | 1.23 | 0.69 | 0.84 |
| n=1 | | | | | | |
| Pb > 13 | -3.52 | 0.79 | 1.10 | 1.00 | 0.55 | 0.87 |
| Pb > 7.8 | -3.35 | 0.74 | 1.14 | 1.00 | 0.60 | 0.83 |
| **Zn** | | | | | | |
| All data | -4.66 | 0.55 | 0.64 | 0.71 | 0.42 | 0.83 |
| Zn > 0.9 | -4.66 | 0.56 | 0.65 | 0.72 | 0.41 | 0.84 |