research for
man and environment

RIJKSINSTITUUT VOOR VOLKSGEZONDHEID EN MILIEU
NATIONAL INSTITUTE OF PUBLIC HEALTH AND THE ENVIRONMENT

RIVM report 711401 007

Determination of field-based sorption isotherms for Cd, Cu, Pb and Zn in Dutch soils

J.G.Otte, J.J.M. van Grinsven, W.J.G.M. Peijnenburg and A.Tiktak

November 1999

This investigation has been performed by order and for the account of the Directorate-General for Environmental Protection, within the framework of project 711401, 'Models and Information Systems Soil and Groundwater'.

Laboratory for Soil- and Groundwater Research (LBG)
RIVM, P.O. Box 1, 3720 BA Bilthoven, telephone: 31 - 30 - 274 91 11; telefax: 31 - 30 - 274 29 71

Mailing list

1. Directie RIVM, Directeur Milieu, Prof. Ir. N.D. van Egmond
2. Sectordirecteur Milieuonderzoek, Ir. F. Langeweg
3. Hoofd Laboratorium voor Bodem- en Grondwateronderzoek, Ir. R. v.d. Berg
4. Hoofd Laboratorium voor Ecotoxicologie, Drs. H. Canton
5. Hoofd Laboratorium voor Water- en Drinkwateronderzoek, Ir. A.H.M. Bresser
6. Dr. Ir. J.R.M. Alkemade, RIVM/LBG
7. Dr. Ir. A.F. Bouwman, RIVM/LBG
8. Dr. Ir. J.J.B. Bronswijk, RIVM/LBG
9. Dr. A.L.M. Dekkers, RIVM/CIM
10. Dr.Ir. M.A.G.T. van den Hoop, RIVM/LAC
11. Ir. J.P.A. Lijzen, RIVM/LBG
12. Dhr. P.F. Otte, RIVM/LBG
13. Drs. G.J.B. Overbeek, RIVM/LBG
14. Dr. M. Posch, RIVM/MNV
15. Dr. D.T.H.M Sijm, RIVM/CSR
16. Dr.Ir. F.A. Swartjes, RIVM/LBG
17. Drs. T. Traas, RIVM/ECO
18. Dr. Ir. C.W. Versluijs, RIVM/LBG

19-23. Auteurs

24. Hoofd Voorlichting en Public Relations RIVM, Drs. J.A.M Lijdsman-Schijvenaars
25. Bureau Projecten- en Rapportenregistratie
26. Bibliotheek RIVM

27-41. Bureau Rapportenbeheer

42. Depot Nederlandse Publicaties en Nederlandse Bibliografie
43. Directoraat-Generaal Milieubeheer, Directie Bodem
44. Dr. F. van Assche, International Lead and Zinc Research Organisation
45. Prof. H.E. Allen, University of Delaware
46. Dr. Ir. A.E. Boekhold, DGM/Bodem
47. Drs. D.J. Bakker, Provincie Flevoland, Afdeling Milieuplanvorming
48. Drs. J. Bril, AB-DLO
49. Drs. P. del Castilho, AB-DLO
50. Ir. J.E. Groenenberg, SC-DLO
51. Dr. Ir. T. J. Heimovaara, IWACO West
52. Dr. J. Japenga, SC-DLO
53. Ir. J. Kros, SC-DLO
54. Dr. J.P.G. Loch, Universiteit Utrecht
55. Dr. B. Pedroli, IBN-DLO
56. Ir. G.J. Reinds, SC-DLO
57. Dr. Ir. P.F.A.M. Romkens, AB-DLO
58. Drs. N.M. de Rooij, WL/Delft Hydraulics
59. Dr. C. van der Salm, SC-DLO
60. Prof. Dr. J. Sevink, Uva/FGBL
61. Prof. Dr. N.M. van Straalen, VU/Oecologie en ecotoxicologie
62. Dr. Ir. E.J.M. Temminghof, LUW/Bodemkunde en Plantenvoeding
63. Prof. Dr. J.M. Verstraten, Uva/FGBL
64. Dr. Ir. J.P.M. Vink, RIZA
65. Dr. Ir. W. de Vries, SC-DLO

66. Ir. R. van de Weert, LUW/Bodemkunde en Plantenvoeding
67. Dr. M. van der Weide, DGM/SVS
68. Prof. Dr. Ir. S.E.A.T.M. van der Zee, LUW/Bodemkunde en Plantenvoeding
69. Dr. J.J.G. Zwolsman, RIZA

Abstract

Sorption isotherms for metals in soil obtained in the laboratory generally underpredict the observed metal content in the solid phase in the field. Isotherms based on in-situ data are therefore required. The aim of this study is to obtain field-based sorption isotherms for Cd, Cu, Pb and Zn as input for the heavy-metal accumulation model SOACAS. Two types of sorption isotherms were fitted using field data sets in this research: the Freundlich type, with one solid phase fraction, fitted by Stepwise Linear Regression (LR) and a model with two solid phase mass fractions was fitted: a reactive fraction and a non-reactive fraction fitted with Nonlinear Least Squares Regression (NLSSR). From the results of the LR and the NLLSR fits of the Hoop-Janssen data set, it appears that generally the explained variance for the Zn models is the highest, followed by the Cd models. The explained variance of the Cu and Pb models is lower. The performance of the LR models and the NLLSR models is comparable. The NLLSR fits of the 2-phase isotherms of the Hoop-Janssen data set almost never include a statistically significant inert metal fraction. This implies that a 2-phase model, in most cases, can not be derived from this data set. The LR models derived from field data (this research) predicts the observed metal content in the solid phase in the field best compared with isotherms derived from batch data. The extension of field partition data set and quality improvement is recommended for future research.

Contents

Summary

Sorption isotherms for metals in soil obtained in the laboratory generally underpredict the observations of the solid metal phase in the field. Isotherms based on in-situ data are therefore required. The aim of this study is to obtain field-based sorption isotherms for Cd, Cu, Pb and Zn as input for the heavy-metal accumulation model SOACAS. The partition isotherms obtained may also be of use for ecotoxicological risk analysis, determination of bioavailability, prediction of groundwater spreading behaviour and derivation of standards for soil pollution policy purposes. Partition behaviour should be related to general soil data, in particular pH, clay content and organic matter content.

Two types of sorption isotherms were fitted in this research. The first type is the Freundlich type, with one solid phase fraction, fitted by Stepwise Linear Regression (LR) on log-transformed observations:

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n}$$

where X (mg kg^{-1}) is the total metal content in the solid phase, {CEC} (mmolc kg^{-1}) is the Cation Exchange Capacity, {%OC} is mass percentage of organic carbon, {%clay} is mass percentage of clay, (H^{+}) is the proton activity and {M}(mg l^{-1}) is the total metal concentration in solution or the free metal activity. To confirm the hypothesis that a fraction of the total metal content is in a non-reactive phase, a model with two solid phase mass fractions was fitted: a reactive fraction and a non-reactive fraction:

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n}$$

The model is fitted with Nonlinear Least Squares Regression (NLSSR) on original observations. The operator {M} can be: the free metal activity, the total metal concentration or the amount of metal extracted with $0.01M$ $CaCl_2$. To fit the isotherms, three data sets were used:

- Data from Van den Hoop (1995) and Janssen (1996), with field data of soils in use for nature conservation;
- Data from the National Soil Monitoring Network (LMB), with field data of soils in use for agriculture.

From the results of the LR and the NLLSR fits of the Hoop-Janssen data set, it appears that generally the explained variance for the Zn models is the highest, followed by the Cd models. The explained variance of the Cu and Pb models is lower. The performance of the LR models and the NLLSR models is comparable. The NLLSR fits of the 2-phase isotherms of the Hoop-Janssen data set almost never include a statistically significant inert metal fraction. The explained variance of the total metal concentration models is generally higher than for the free activity models. The isotherms derived from the field data set of Hoop and Janssen do not depend on the clay content and the organic carbon content in soils. The isotherms based on data for soils in use for nature conservation (Hoop-Janssen) and isotherms based on data for soils in use for agriculture (LMB) are not comparable for Cd, Pb and Zn. The isotherms derived from the agricultural data include non-reactive fractions and the pH dependence is much smaller than for the isotherms derived from the Hoop-Janssen data set. This is caused by the fact that the range in pH values is much smaller for the LMB data set than for the

Hoop-Janssen data set. The LR activity models derived from field data (this research) predicts the observed solid metal contents better than isotherms derived from batch data (Bril (1999) and Elzinga et al. (1997)). The LR activity models of Elzinga et al. (1997) predict the observed solid metal contents reasonably accurate for Zn and with slight underestimation for Cd and Cu. The LR activity models of Bril (1999) predict the observed metal content in the solid phase less well.

Based on the abovementioned result that the NLLSR fits for the Hoop-Janssen data set almost never include an inert fraction, it can be concluded that a 2-phase model can, in most cases, not be derived from this dataset. The result that the NLLSR models with total concentration have a higher explained variance than the free activity models is in contradiction with the fact that the free metal activity should describe the sorption behaviour best. The free metal activity is corrected for ion strength and (inorganic and organic) complexation in solution. Therefore free activity models are considered as the most suitable isotherms for SOACAS. In spite of the fact that the explanatory potential of the LR models and the NLLSR models is comparable, the NLLSR fits are preferred. The reason for this is the fact that for the NLLSR procedure untransformed values are used, which give more weight to the high concentrations compared to log-transformed values. For SOACAS the purpose is especially to predict the high metal concentrations well. To judge the validity and applicability of the fitted sorption isotherms, the small number of reliable observations of the data sets for agriculture and nature conservation and non-plausible regression coefficients should be considered. The extension of field partition data set is recommended for future research. In the future sorption isotherms must be derived from large enough, reliable data sets which include a balanced composition of soils in use for agriculture and soils in use for nature conservation. These data sets must also have a large spread of metal concentration and soil characteristics.

Keywords: heavy metals, sorption, partition, Freundlich isotherms

Samenvatting

Sorptie-isothermen voor metalen in de bodem die in het laboratorium verkregen zijn, onderschatten meestal de totale metaalgehalten in het veld. Daarom zijn isothermen gebaseerd op in-situ data nodig. Het doel van dit onderzoek is sorptie-isothermen voor Cd, Cu, Pb en Zn te verkrijgen als invoer voor het zware metalen accumulatiemodel SOACAS. s De verkregen sorptie-isothermen kunnen ook nuttig zijn voor ecotoxicologische risico-analyses, bepaling van de biobeschikbaarheid, voorspelling van verspreidingsgedrag in het grondwater en afleiding van normen voor beleid m.b.t bodemverontreiniging. Partitiegedrag moet gerelateerd worden aan algemene bodemgegevens zoals pH, kleigehalte en organische stofgehalte.

Twee typen sorptie-isothermen zijn afgeleid in dit onderzoek. Het eerste type is het Freundlich type met één vaste fase fractie, afgeleid met Stapsgewijze Lineaire Regressie (LR) op log-getransformeerde waarnemingen:

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n}$$

waarin X (mg kg^{-1}) het totale metaalgehalte in de vaste fase is, {CEC} (mmolc kg^{-1}) is de Kationuitwisselingscapaciteit, {%OC} is het massapercentage organische koolstof, {%clay} is het massapercentage klei, (H^{+}) is de proton activiteit en {M}(mg l^{-1}) is het metaalgehalte in oplosssing. Om de hypothese te bevestigen dat een fractie van het totale metaalgehalte in de vaste fase een niet-reactieve fase is, is een model afgeleid met twee vaste fase metaalfracties, een reactieve fractie en een inerte fractie:

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n}$$

Het model is afgeleid met Niet Lineaire Kleinste Kwadraten Regressie (NLSSR). De parameter {M} kan zijn: de vrije metaalactiviteit, de totale metaalconcentratie of de hoeveelheid metaal geëxtraheerd met $0.01M$ $CaCl_2$. Om de isothermen af te leiden zijn drie datasets gebruikt:

- Data van van den Hoop (1995) en Janssen (1996), met velddata van bodems in gebruik voor natuur;
- Data van het Landelijk Meetnet Bodem (LMB), met velddata van bodems in gebruik voor landbouw.

Uit de resultaten van de LR en de NLLSR fits van de Hoop-Janssen dataset blijkt dat de verklaarde variantie voor de Zn modellen meestal het hoogste is, gevolgd door de Cd modellen. De verklaarde variantie van de Cu en Pb modellen is lager. De performance van de LR en de NLLSR modellen is vergelijkbaar. De NLLSR fits van de 2-fase isothermen van de Hoop-Janssen dataset bevatten bijna nooit een statistisch significante inerte metaalfractie. De verklaarde variantie van de totale metaalconcentratie modellen is over het algemeen hoger dan voor de vrije activiteit modellen. De isothermen die afgeleid zijn van de velddataset van Hoop en Janssen zijn niet afhankelijk van het kleigehalte en het organisch stofgehalte in bodems. De isothermen afgeleid van data van bodems in gebruik voor natuur (Hoop-Janssen) en isothermen afgeleid van data van bodems in gebruik voor landbouw (LMB) zijn niet vergelijkbaar voor Cd, Pb en Zn. De isothermen afgeleid van de landbouw data bevatten niet-reactieve fracties en de pH afhankelijkheid is veel kleiner dan voor de isothermen afgeleid

van de LMB dataset dan voor de Hoop-Janssen dataset. De LR activiteitsmodellen afgeleid van velddata (dit onderzoek) voorspelt de waargenomen totale metaalgehalten het beste vergeleken met de isothermen afgeleid van laboratoriumdata (Bril (1999) en Elzinga et al. (1997)). De LR activiteitsmodellen van Elzinga et al. (1997) voorspellen redelijk goed voor Zn en met een kleine onderschatting voor Cd en Cu. De LR activiteitsmodellen van Bril (1999) voorspellen de waargenomen totale metaalgehalten minder goed.

Gebaseerd op de bovengenoemde resultaten dat de NLLSR fits voor de Hoop-Janssen dataset bijna nooit een significante inerte fractie bevatten, kan geconcludeerd worden dat een 2-fase model, in de meeste gevallen, niet van deze dataset afgeleid kan worden. Het resultaat dat de NLLSR modellen met totale concentratie een hogere verklaarde variantie hebben dan de vrije activiteit modellen is in tegenspraak met het feit dat de vrije metaal activiteit het sorptiegedrag het beste zou moeten beschrijven. De vrije metaalactiviteit is gecorrigeerd voor ionsterkte en (anorganische en organische) complexatie in oplossing. Daarom worden de vrije metaalactiviteit modellen beschouwd als de meest geschikte isothermen voor SOACAS. Ondanks het feit dat de LR modellen en de NLLSR modellen vergelijkbaar zijn wat betreft potentie om te voorspellen, is er een voorkeur voor de NLLSR modellen. De reden hiervoor is het feit dat voor de NLSSR-procedure ongetransformeerde waarden worden gebruikt, in plaats van loggetransformeerde waarden, wat meer gewicht geeft aan de hoge concentraties. Voor SOACAS is namelijk het doel om vooral de hoge concentraties goed te voorspellen. Om de validiteit en de toepasbaarheid van de afgeleide isothermen te beoordelen, moet rekening worden gehouden met het beperkte aantal betrouwbare waarnemingen van de gebruikte datasets en niet-plausibele regressiecoëfficiënten. De uitbreiding van de velddataset wordt aanbevolen voor toekomstig onderzoek. In de toekomst moeten isothermen worden afgeleid van datasets die groot genoeg en betrouwbaar zijn en die een evenwichtige samenstelling bevatten van bodems in gebuik voor landbouw en bodems in gebruik voor natuur. Deze datasets moeten ook een grote spreiding vertonen in metaalconcentratie en bodemeigenschappen.

Trefwoorden: zware metalen, sorptie, partitie, Freundlich-isothermen

1. Introduction

1.1 Background and research aim

Background

Sorption isotherms for heavy metals are needed to estimate the mobility and bioavailability of metals in soils. The sorption models for Cd, Cu, Pb and Zn that are used until now are considered inappropriate because they:

- take no account of the large difference between sorption behaviour under lab and field conditions;
- were not specifically derived and validated for the situation in the Netherlands;
- were not consistently derived for the different metals (Elzinga et al, 1999).

In this research new sorption isotherms are fitted, with a regression procedure, on field partition data obtained from Dutch soils. In this analysis specific attention was given to:

- (statistical) evidence for inert metal fractions in the solid phase leading to sorption isotherms with more than one solid phase metal fraction;
- comparison of field-based sorption isotherms (Cd, Cu, Pb and Zn) derived from soils used for nature conservation (Van den Hoop, 1995; Janssen et al., 1996) to in situ isotherms derived from soils used for agriculture (Lagas and Groot, 1996);
- comparison of sorption isotherms with alternatively (a) free activity of the metal, (b) total metal concentration in solution or (c) the amount of metal extracted with $0.01M$ $CaCl_2$ as independent variable to represent the mobile concentration;
- comparison of sorption isotherms with either (a) organic carbon content and the clay content or (b) the CEC as parameter to correct for soil type;
- the effective concentration (application) ranges on the sorption isotherms.

Research aim

The aim of this research is to obtain field-based sorption isotherms as input for the heavy-metal accumulation model SOACAS. These isotherms should describe equilibrium partitioning between metal contents in the solid phase and the solution phase, taking into account the possible presence of inert metal fractions in the solid phase. SOACAS will be applied for non-point source pollution of Cd, Cu, Pb and Zn in soils in the Netherlands. Additionally the aim is to derive partition isotherms in soils for ecotoxicological risk analysis, determination of bioavailability, prediction of groundwater spreading behavior and derivation of standards for soil pollution policy purposes. Partition behaviour should be related to general soil data in particular pH, clay content and organic matter content.

1.2 Sorption and speciation concepts

The partitioning of heavy metals is crucial for describing transport and accumulation in soil and groundwater. Partitioning is the distribution of heavy metals over the various phases. The total amount of metal in soil gives no information about the bioavailability or the mobility of the metal. Therefore partitioning is used to distinguish between the total metal content in the

solid phase and the metal content in solution. The partition coefficient can be calculated as the ratio of the total metal content in the solid phase and the metal content in solution:

$$K_p = \frac{X}{M} \qquad [1]$$

where K_p (l kg^{-1}) is the partition coefficient, X (mg kg^{-1}) is the total metal content in the solid phase and M (mg l^{-1}) is the metal content in solution.

Speciation is the distribution of metals over the various phases. It considers the various binding forms as the unavailable fraction build in the structure of soil minerals and organic matter, the adsorbed fractions, the precipitated metals and the metals in solution which may be complexed (Figure 1) (Elzinga et al., 1997). Complexation is the binding of metal to anions or negatively charged parts in the soil solution. With respect to the (bio-)availibility of heavy metals, knowledge of their speciation is of great significance, both from the viewpoint of their essentiality for various life forms, which is the case for e.g. Zn, and with respect to their toxicity, as is the case for e.g. Cd and Pb. It is now widely accepted for a number of heavy metals that toxicity is more related to the concentration of the free metal, rather than total metal (Van den Hoop, 1995). The concept of partition is often confused with the concept of sorption. Partition is the observed distribution of a compound over two phases. Sorption is a collective noun for binding processes, which is a wider concept than adsorption and can also comprise processes as precipitation or incorporation in soil minerals and organic matter (Koops et al., 1998).

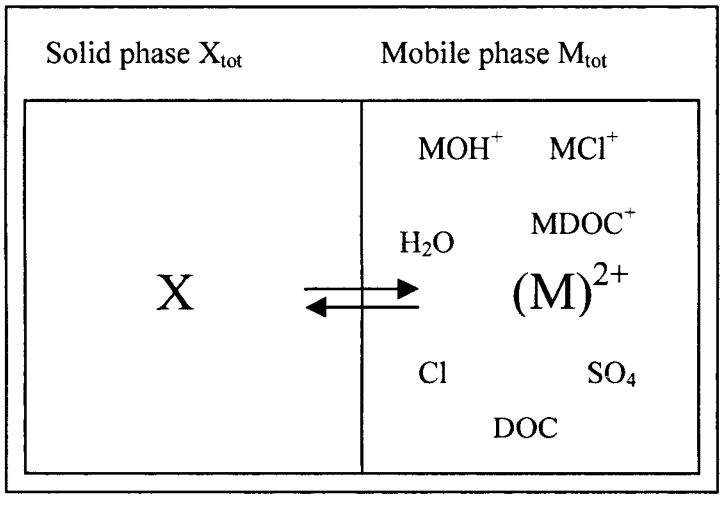


Figure 1: *The binding of metals onto soil components.*

Another difference is that sorption is usually described as a function of metal concentration, which yields a sorption isotherm, whereas partition is a point measurement, that is to say it applies to one concentration. The Freundlich type sorption isotherm is:

$$X = K_f M^n \qquad [2]$$

where K_f (l kg^{-1}) is the Freundlich coefficient, X (mg kg^{-1}) is the total metal content in the solid phase, M (mg l^{-1}) is the metal content in solution and n is the Freundlich exponent.

Sorption isotherms for cations are usually fitted as a function of the metal concentration in the pore water and soil characteristics like pH, organic carbon content, clay content and CEC (Cation Exchange Capacity):

- when the pH decreases, the concentration of metal in solution will increase because the H^+ ion competes with the metal ion for specific sorption sites. Moreover, the dissociation degree of organic matter decreases at a decreasing pH, which will result in a decrease of the amount of sorption sites;
- organic matter, clay minerals and sesquioxides are the most important components for the sorption of metals because of their negative electric charge and large specific area;
- CEC is a measure for the total amount of binding sites. At a high CEC, many cations can be adsorbed in the soil (Elzinga et al., 1997).

The Freundlich type sorption isotherm including soil charasteristics is:

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n} \quad [3]$$

where X (mg kg^{-1}) is the total metal content in the solid phase, {CEC} ($mmol_c$ kg^{-1}) is the Cation Exchange Capacity, {%OC} is mass percentage of organic carbon, {%clay} is mass percentage of clay, (H^+) is the proton activity and {M}(mg l^{-1}) is the metal content in solution.

1.3 Sorption equations

1.3.1 Freundlich isotherms

Two types of sorption isotherms were fitted in this research. The first type is the Freundlich type, with one solid phase fraction:

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n} \quad [3]$$

where X (mg kg^{-1}) is the total metal content in the solid phase, {CEC} ($mmol_c$ kg^{-1}) is the Cation Exchange Capacity, {%OC} is mass percentage of organic carbon, {%clay} is mass percentage of clay, (H^+) is the proton activity and {M}(mg l^{-1}) is the metal content in solution.

1.3.2 Isotherms extended with non-reactive pools

Sorption models based on laboratory batch data systematically underestimate the observed total metal pool in the field. This suggests that only a fraction of the total metal content in the solid phase is reactive (Elzinga et al., 1999). A fraction of the total metal content in the solid phase is probably non-reactive to the pore water and build in the structure of clay minerals and organic matter. To confirm this hypothesis, a second type of model is fitted with two solid phase mass fractions is fitted: a reactive fraction and a non-reactive or inert fraction:

$$X = A + 1phase \quad [4]$$

$$A = a + b\{\%OC\} + c\{\%clay\} \quad [5]$$

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n} \qquad [6]$$

where A (mg kg^{-1}) is the non-reactive solid phase and 1*phase* is the reactive solid phase (mg kg^{-1}).

the operator {M} can be:
(M^{2+}) = free metal activity;
[M] = total metal concentration;
$[M_{CaCl2}]$ = amount of metal extracted with $0.01M$ $CaCl_2$.

1.4 Short description of SOACAS model

The SOACAS model has been developed with three objectives:
- explanation of current metal contents (diagnosis);
- derivation of critical metal loads;
- description of the long-term response of trace metal contents to soil loadings in a regional context (prognosis).

SOACAS describes the fate of a metal in one completely mixed soil compartment. Metal inputs considered are atmosferic deposition, fertilizers and animal manure. After addition, complexation of these metals in the soil solution is calculated and they are partitioned over the mobile and immobile phases. Finally, part of the metal is removed from the system by leaching and harvesting. Leaching and plant-uptake are represented in the model by first-order rate processes, while equilibrium partioning between the mobile and immobile phases is described by a Freundlich isotherm (Tiktak, 1999).

For the application in SOACAS, the isotherms fitted in this research have to come up to the following requirements:
- the isotherms must be derived from field data because sorption models based on laboratory batch data systematically underestimate the observed total metal pool in the field.;
- the isotherms must be derived from data of soils which are in use for nature conservation and soils in use for agriculture, so that SOACAS will be applicable for soils in use for either form of landuse;
- the isotherms must contain regionalizable parameters as the clay content, the organic carbon content and the pH, so that the regional variation in metal content can be predicted using SOACAS;
- the isotherms must be as simple as possible to fit into the model but with a high explained variance;
- the isotherms must contain the free metal activity as independent parameter to represent the mobile fraction. The free metal activity describes sorption behaviour best, compared with the total metal concentration in solution, because a correction for ion strength and (inorganic and organic) complexation in solution has taken place (Section 1.2);
- the possibility of the isotherm to contain an inert solid phase fraction. To confirm the hypothese that a fraction of the total metal content in the solid phase is non-reactive to the pore water, a model with two solid phase mass fractions should be fitted: a reactive

fraction and a non-reactive or inert fraction (Section 1.3.2.). Up till now in SOACAS a regional variable is used to impose a background metal content:

$$X = A + B \qquad [7]$$

where X (mg kg^{-1}) is the total metal content in the solid phase, A (mg kg^{-1}) is the non reactive solid phase fraction and B (mg kg^{-1}) is the reactive solid phase fraction. The reactive solid phase fraction is described by the Freundlich sorption isotherm. A regional variable is used to construct a map of estimated background metal contents.

1.5 Framework

This report consists of four chapters. Chapter 2 describes the materials and methods as the calculation procedures, background information on field observations, background information in regression techniques and data manipulations. Chapter 3 consists of the results and discussion. Section 3.1 to 3.2 describe the results of the LR and NLLSR fits of isotherms. Section 3.4 consists of an evaluation of isotherms based on batch data versus field data. Section 3.5 consists of an evaluation of isotherms with free activity versus total solution concentration. Section 3.6 describes the validation of isotherms and section 3.7 describes the application perspective of isotherms based on $CaCl_2$-extracts. Chapter 4 consists of the conclusions and recommendations

2. Materials and Methods

2.1 Calculation procedures

Two different regression procedures are performed:
a) the 1-phase sorption model was fitted by Stepwise Linear Regression on log-transformed observations (procedure-a);
b) the 2-phase sorption model was fitted with Nonlinear Least Squares Regression on original observations (procedure-b).
In Table 1 a summary of the various types of fitted models is given.

Table 1: *Summary of the various types of fitted models*

Independent parameter	Regression procedure LR	NLLSR
Free metal activity - (M^{2+})	%clay and %OC or CEC as independent parameter	%clay and %OC as independent parameter
	no initial inert phase	initial inert phase
Metal concentration in solution - [M]	%clay and %OC or CEC as independent parameter	%clay and %OC as independent parameter
	no initial inert phase	initial inert phase
Metal content extracted with 0.01*M* $CaCl_2$ - [M_{CaCl2}]		%clay and %OC as independent parameter
		initial inert phase

2.1.1 LR fits of Freundlich isotherms on log-transformed data for soils in use for Nature conservation

LR models are fitted to make a comparison possible with the Freundlich isotherms derived on batch data by Bril (1999) and Elzinga et al. (1997). Linear Regression was performed for different soil and experimental characteristics as independent parameters. The following type of regression equations were fitted by Linear Regression from the log transformed values of the combined Hoop-Janssen data set:

- $X = 10^{p}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}(M^{2+})^{n}$ [8]
- $X = 10^{p}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}[M]^{n}$ [9]
- $X = 10^{p}\{CEC\}^{q}(H^{+})^{m}(M^{2+})^{n}$ [10]

- $X = 10^{p}\{CEC\}^{q}(H^{+})^{m}[M]^{n}$ [11]

where X (mg kg^{-1}) is the total metal content in the solid phase, {CEC} ($mmol_c$ kg^{-1}) is the Cation Exchange Capacity, {%OC} is mass percentage of organic carbon, {%clay} is mass percentage of clay, (H^+) is the proton activity, (M^{2+}) (mg l^{-1}) is the free metal activity in solution and [M] (mg l^{-1}) is the total concentration in solution. The log transformation of data will generate a log-normal distributed dataset. It will decrease the weight of high concentrations in the data set and increase the weight of low concentrations. When performing Linear Regression several data points were left out of the Hoop-Janssen data set because the total metal content in the solid phase or the Dissolved Organic Carbon (DOC) was not measured. For all metals the same observations were left out of the data set for the regression with the total metal concentration as for the regression with the free metal activity.

2.1.2 NLLSR fits of 2-phase isotherms on original data for soils in use for Nature conservation

The Nonlinear Least Squares Regression procedure is used to fit the 2-phase isotherms. These isotherms are nonlinear and can not be fitted using Linear Regression. NLLSR is performed for different soil and experimental characteristics as independent parameters. The following type of regression equations are derived by Nonlinear Least Squares Regression from the untransformed values of the combined Hoop-Janssen data set :

- $X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}(M^{2+})^{n}$ [12]

- $X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}[M]^{n}$ [13]

- $X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}[M_{CaCl2}]^{n}$ [14]

where X (mg kg^{-1}) is the total metal content in the solid phase, {CEC} ($mmol_c$ kg^{-1}) is the Cation Exchange Capacity, {%OC} is mass percentage of organic carbon, {%clay} is mass percentage of clay, (H^+) is the proton activity, (M^{2+}) (mg l^{-1}) is the free metal activity in solution, [M] (mg l^{-1}) is the total concentration in solution and [M_{CaCl2}) (mg l^{-1}) is the metal content extracted with $0.01M$ $CaCl_2$. For the sorption isotherms with the total metal concentration in solution of Cd and Cu the same observations were left out of the data set as for the 2-phase sorption isotherm with the metal content according to a $CaCl_2$ extraction.

For the Nonlinear Least Squares Regression procedure untransformed values are used, which gives more weight to the high concentrations. A consequence of this will be that the NLLSR models will predict the high concentrations well, but overestimate the low concentrations. For the model SOACAS the purpose is especially to predict the high metal concentrations well to find out where there will be a danger of accumulation of metal contents in soils. The CEC coefficient of the NLLSR models is set to zero because from the LR results, it appears that CEC does not explain more variance than the clay content and the organic carbon content. . In this way the danger of overfitting of data is also prevented because CEC and the clay content and the organic carbon content are correlated. Moreover the regional variation of metal content predicted by SOACAS is based on the variation in the clay content and the organic carbon content.

The models with the metal content extracted with $CaCl_2$ as independent parameter were fitted to enable a comparison with models derived from the LMB data set. The models derived from the Hoop-Janssen data set are based on data of soils in use for nature conservation. When these models will be used in SOACAS, they must be applicable for soils in use for agriculture too. To verify this, models are derived from the LMB data set and compared with the models derived from the Hoop-Janssen data set. The LMB data set does however not contain the metal concentration in solution or the free metal activity, therefore these models are fitted with the metal content extracted with $CaCl_2$ as independent parameter.

For a number of samples the amount of metal extracted with $CaCl_2$ is relatively high compared to the amount of metal digested with HNO_3. This is caused by the fact that in case of a sandy soil a larger amount of metal is extracted with $CaCl_2$ for Cd and Zn. This does not apply to Cu and Pb because these metals form stronger complexes which are not destructed with the $CaCl_2$ extraction. For this reason the difference between the HNO_3 digestion and the $CaCl_2$ extraction is used as the total amount of adsorbed metal for Cd and Zn.

When deriving the 2-phase sorption isotherms, several data points were left out of the data sets for the following reasons:

- either the total metal content in the solid phase, the metal content according to the $CaCl_2$ extraction or the Dissolved Organic Carbon (DOC) was not measured for several data points (Hoop-Janssen data set);
- the measured values lie below detection limits for several data points (LMB data set);
- several data points were outliers. This was examined by making several plots:
  1. a plot of the measured versus the predicted total metal content in the solid phase (obtained by Linear Regression);
  2. a plot of the residuals versus the predicted total metal content in the solid phase;
  3. a plot of the residuals versus the quantiles of standard normal;
  4. a plot of the Cook's Distance versus the observation number. The Cook's Distance is a measure of the influence of each data point on the regression equation (Hoop-Janssen and LMB data set).

Each time Nonlinear Regression was performed, the independent parameter at the p-level smaller than 0.05 was left out of the next Nonlinear Regression in a stepwise manner. This was continued until all parameters of the regression equation were significant at the p-level < 0.05. For each metal a maximum of three different fitted models are reported:

1. a model with the highest explained variance;
2. a model with only significant regression coefficients at the p=0.05 level;
3. an intermediate model with a selection of non-significant coefficients but which appear more plausible than (2) from a theoretical point of view.

2.1.3 NLLSR fits of 2-phase isotherms on original data for soils in use for agriculture

NLLSR models are derived from the LMB data set, as mentioned in Section 2.1.2., to compare them with the models derived from the Hoop-Janssen data set to verify if the models derived from the Hoop-Janssen data set are applicable for soils in use for agriculture also. The LMB data set does, as described above, not contain the metal concentration in solution or the free metal activity, therefore these models are fitted with the metal content extracted with

$CaCl_2$ as independent parameter. The following type of regression equation is derived by Nonlinear Least Squares Regression from the untransformed values of the LMB data set (Lagas and Groot, 1996):

- $X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}[M_{CaCl2}]^{n}$ [14]

where X (mg kg^{-1}) is the total metal content in the solid phase, {CEC} ($mmol_c$ kg^{-1}) is the Cation Exchange Capacity, {%OC} is mass percentage of organic carbon, {%clay} is mass percentage of clay, (H^+) is the proton activity and [M_{CaCl2}) (mg l^{-1}) is the metal content extracted with $0.01M$ $CaCl_2$.

2.2 Background information on field observations

For this research three data sets are used:
- the Van den Hoop data set, with field data of soils which are in use for nature conservation;
- the Janssen data set, with field data of soils which are in use for nature conservation;
- the National Soil Monitoring Network (LMB) data set, with field data of soils which are in use for agriculture.

Van den Hoop data set

The objectives of the study of Van den Hoop were:
- to quantify heavy metal concentrations in solid and soil solution in soils under field conditions in order to obtain field partition coefficients (Kp);
- to compare the partition coefficients obtained with values proposed for the harmonization of quality criteria for heavy metals in the Netherlands (Van den Hoop, 1995).

Soil samples were collected from 14 different sites in use for nature conservation in the Netherlands between April 23 and July 20, 1993. The sites were selected on the next set of criteria:
- the heavy metal content should be close to "background" levels;
- the soil samples should vary with respect to organic carbon content, clay content, pH and carbonate content (Van den Hoop, 1995).

Detailed background information on field observations is given in Appendix 1. In Appendix 2 the data of the used data sets are presented. The most important features of the data sets used in this research are presented in Table 2.

Total metal concentrations of the solid phase were obtained by means of GFAAS or ICP-AES after digesting the soil samples with concentrated nitric acid. To define the fraction of (bio-)-available or mobile metals in the soil, the soil samples were extracted by adding 100 ml $0.01M$ $CaCl_2$ solution to 10 g of soil (Van den Hoop, 1995).

Janssen data set

The objectives of the study of Janssen were:
- Determination of field partition coefficients (As, Cd, Cr, Cu, Pb and Zn) for 20 polluted soils sampled in the Netherlands;

- Investigation of the dependence of measured Kp values on soil and pore water characteristics (Janssen et al., 1996).

Soil samples were collected from 20 different sites in use for nature conservation in the Netherlands in October and November 1994. Sampling was carried out under dry weather conditions. The sites were selected on the next set of criteria (Table 2):
- elevated metal concentrations had to be expected due to the location of sites close to sources of pollutants of interest;
- little or no impact from agricultural practices;
- soil characteristics had to vary among sites (Janssen et al., 1996).

The detection limits of the pore water analysis, the HNO_3 destructions and the $CaCl_2$ extractions were not available but according to Groot (1999) the detection limits are comparable with those used at the Laboratory of Anorganic Chemistry (LAC) of the RIVM because similar methods are used. Therefore these detection limits are reported.

To obtain the total amount of metal, the soil samples were destructed with $15M$ HNO_3. To define the fraction of (bio-)available or mobile metals in the soil, the soil samples were extracted by adding 100 ml 0.01 M $CaCl_2$ solution to 10 g of air-dry soil (Janssen et al., 1996).

The National Soil Monitoring Network (LMB) data set

A research project to assess the quality of agricultural soil in the Netherlands has been carried out by the National Institute of Public Health and Environmental Protection in cooperation with the Laboratory of Soil and Crop research (BLGG).

A total of 42 different combinations of soil type (sand, fluvial and marine clay, peat and loam) and soil use (grassland, arable land, maize, flower, bulbs and orchards) were selected from 13 agricultural regions in the Netherlands (Lagas and Groot, 1996). For each combination four representative mixed samples were made to determine average heavy metal and organic compounds contents. The representative samples, composed of 20 parcel samples, were among other parameters analysed for total and bioavailable heavy metals. For grasslands and orchards, samples were taken from the upper 5 cm and for arable land from the upper 25 cm (Lagas and Groot, 1996).

To obtain the total amount of metal, all soil samples were destructed with Aqua Reagia (mixture of nitric acid and hydrochloric acid) according to NEN 6465 (Table 2). For each combination of soil type and soil use, one soil sample has been extracted with $0.01M$ $CaCl_2$ solution in order to define the fraction of (bio-)available or mobile metals in the soil. Extractions were performed by adding 100 ml $0.01M$ $CaCl_2$ solution to 10 g of soil (Lagas and Groot, 1996).

Table 2: *Properties of the data sets used in this research*

Property	Data set Hoop	Janssen	LMB
Number of observations	13	20	42
Landuse	nature conservation	nature conservation	agriculture
Degree of contamination	natural background metal concentrations	large range of elevated metal concentrations	small range of metal concentrations, on average below Target Values
Range of pH pore water	3.2 – 7.85	3.4 - 7.91	5.2 – 7.9
Range of organic carbon content (%)	0.3 – 43.1	1.2 - 12.8	0.76 – 23.35
Range of clay content (%)	1.4 – 64.9	0.79 – 33.83	2.4 – 34.1
Heavy Metals	Cd, Cu, Ni, Pb and Zn	As, Cd, Cr, Cu, Ni, Pb and Zn	As, Cd, Cr, Cu, Ni, Pb and Zn
Determination pore water	GFAAS: Cd, Ni, Cu and Pb ICP-AES: Zn, Ca, Mg, Na, K, Fe and Mn ion chromotography: NO_3, Cl and SO_4 continuous flow analysis: PO_4 SOP LAC/M049: DOC	AAS: Cd, Cr, Ni, Pb, Cu and Zn FIAAS: As ICP-AES: Ca, Mg, Na, K, Zn, Fe, Mn and Al ion chromotography: Cl, NO_3, HCO_3, CO_3 continuous flow analysis: PO_4 TOC-analyzer: DOC	not determined
Detection limits pore water	Cd: 0.1 µg/l Cu: 0.5 µg/l Pb: 0.5 µg/l Zn: 6.5 µg/l	Cd: 0.034 µg/l Cu: 0.635 µg/l Pb: 1.036 µg/l Zn: 6.54 µg/l	not determined
Determination solid phase	digestion with concentrated HNO_3	digestion with concentrated HNO_3	digestion with Aqua Regia
Detection limits solid phase	Cd: 1 µg/kg Cu: 5 µg/kg Pb: 5 µg/kg Zn: 65 µg/kg	Cd: 0.067 mg/kg Cu: 1.270 mg/kg Pb: 2.072 mg/kg Zn: 1.308 mg/kg	Cd: 0.35 mg/kg (0.15 mg/kg graphite furnace) Cu: 3 mg/kg Pb: 8 mg/kg Zn: 6 mg/kg
Detection limits 0.01*M* $CaCl_2$ extraction	Cd: 0.1 µg/l Cu: 0.5 µg/l Pb: 0.5 µg/l Zn: 6.5 µg/l	Cd: 0.337 µg/l Cu: 6.35 µg/l Pb: 10.36 µg/l Zn: 65.4 µg/l	Cd: 0.045 µg/l Cu: 0.64 µg/l Pb: 0.21 µg/l Zn: 10 µg/l
Reference detection limits	Van den Hoop, 1999	De Groot, 1999	Groot, 1998

2.3 Background information on regression techniques

Linear Regression

The option Linear Stepwise Regression of the statistical package SPlus (SPlus 4, 1998) for the metals Cd and Pb was used to fit the log-transformed Freundlich sorption isotherms. Detailed background information about Linear Regression and Nonlinear Least Squares Regression is given in Appendix 3. In Table 3 the regression procedure Linear Regression is compared with Nonlinear Least Squares Regression.

Table 3: *Comparison of regression procedures used in this research*

Regression procedure Linear Regression (LR)	Nonlinear Least Squares Regression (NLLSR)
log-transformed values	untransformed values
unique solution	multiple solutions local solutions
no starting values	starting values
R^2	estimation of R^2 via the Sum of Squares

Nonlinear Least Squares Regression

The option Nonlinear Least Squares Regression of the statistical package SPlus (SPlus 4, 1998) is used to fit the 2-phase sorption isotherms.

Nonlinear models require starting values (SPlus 4, 1998): the data set of Hoop-Janssen and the LMB data set were used to calculate starting values for the non-reactive part of the 2-phase sorption isotherm:

$$X = A + 1phase \quad [7]$$

$$A = a + b\{\%OC\} + c\{\%clay\} \quad [8]$$

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n} \quad [9]$$

where A is the non-reactive solid phase (mg kg^{-1}) phase and $1phase$ is the reactive solid phase (mg kg^{-1}).

With the calculation of each starting value of a coefficient, the other coefficients of the isotherm were assumed to be zero each time. To obtain a starting value for coefficient a, the mean total metal content in the solid phase of the data set was calculated. To obtain a starting value for coefficient b, the ratio of the mean total metal content in the solid phase and the mean mass percentage of organic carbon was calculated (X / %OC), the quotient of the mean

total metal content in the solid phase and the mean mass percentage of clay was calculated (X / %clay) to obtain a starting value for coefficient c.

For the starting values of the reactive part of the 2-phase sorption isotherm the coefficients of the non-reactive part were assumed to be zero. The regression coefficients of the 1-phase Freundlich sorption isotherms as fitted by Elzinga (Elzinga et al., 1997) were used as starting values for the metals Cd, Cu and Zn. For Pb the regression coefficients of the 1-phase Freundlich sorption isotherms derived from the combined Hoop-Janssen data set were used. To examine if each 2-phase sorption isotherm had more than one solution (minimum sum of squares differences between response and prediction), different starting values were used.

The option Nonlinear Least Squares Regression in SPlus does not calculate the explained variance, because with Nonlinear Regression the explained variance has no physical meaning. An estimation of the explained variance is made by calculating the residual sum of squares (RSS_0) of the zero model of each data set with Nonlinear Regression:

$$X = a_0 \qquad [15]$$

where X (mg kg^{-1}) is the total metal content in the solid phase and a_0 is a regression coefficient. In this way the difference in explained variance can be examined between the 2-phase sorption isotherm and a random regression coefficient. Next the residual sum of squares (RSS) of the 2-phase sorption isotherm [2] of the same data set is calculated with Nonlinear Regression. The estimated explained variance is calculated using the following equation:

$$R^2 \equiv \frac{RSS_0 - RSS}{RSS_0} \qquad [16]$$

2.4 Data manipulations

Calculation of soil characteristics

Some of the soil characteristics of the combined Hoop-Janssen data set and the LMB data set, that were necessary to obtain sorption isotherms, were conversed to other units. The total content of metal in the solid phase was conversed from mmol kg^{-1} to mg kg^{-1} and the total concentration of metal in solution was conversed from nanomol l^{-1} or µmol l^{-1} to mg l^{-1}. The total concentrations in solution of all compounds were conversed from mmol l^{-1}, µmol l^{-1} and nanomol l^{-1} to mol l^{-1} for the calculation of the free metal concentration in MINEQL$^+$. The Cation Exchange Capacity was conversed from cmol kg^{-1} to mmol kg^{-1}.

The soil characteristics that were not available, were the organic carbon content in the Janssen data set and the total concentration of organic anions (DOC in mol l^{-1}) in the combined Hoop-Janssen data set. These characteristics were calculated using other soil characteristics. The organic carbon content of the Janssen data set is calculated from the estimated organic matter from LOI by dividing by a value of 1.7 (Janssen et al.,1996). The total concentration of organic anions (µmol_c l^{-1}) in the soil solution was related to the Dissolved Organic Matter (mg l^{-1}) concentration by:

$$[A_t] = [A^{2-}] + [HA^-] + [H_2A^0] = m\ DOC \qquad [17]$$

where m represents the amount of organic anions as a fraction of DOC. m was set to 5.5 $\mu mol_c\ mg^{-1}$ in accordance to Henriksen and Seip (1980) and Tiktak (1999).

Calculation of free metal concentration

The chemical equilibrium program MINEQL$^+$ was used to calculate the free metal concentration using the total concentration of metal in solution. MINEQL$^+$ can calculate the distribution of chemical species in an aqeous system. The following components were selected:

H_2O	Fe^{2+}	SO_4^{2-}
H^+	K^+	Zn^{2+}
Ca^{2+}	Mg^{2+}	DOC^{2-}
Cd^{2+}	Mn^{2+}	
Cl^-	Na^{2+}	
CO_3^{2-}	NO_3^-	
Cu^{2+}	Pb^{2+}	

To determine how the components interact with each other to form chemical species the thermodynamic database was scanned and the total concentration of each component was filled in. As the MINEQL$^+$ database did not contain dissociation constants for organic complexes, these constants were added. The association of organic acids was represented by a diprotic acid analogue:

$$HA^- + H^+ \Leftrightarrow H_2A^0 \qquad \log K = 4.3 \qquad [18]$$

$$A^{2-} + H^+ \Leftrightarrow HA^- \qquad \log K = 9.5 \qquad [19]$$

where A^{2-} represents the organic anion. The association constants were taken from Mantoura et al. (1978) (Table 4).

Table 4: *Association constants for metal complexes used in MINEQL$^+$*

Species	log K			
	Cd	Cu	Pb	Zn
MA^0	6.10	10.00	9.70	6.10
MHA^+	2.60	3.90	3.70	2.70

The temperature of the aqueous solution was set to 15 °C. At this temperature the pore water of the soil samples of the data set of Janssen et al.(1996) was obtained by centrifugation. To fix the partial pressure of CO_2 (P_{CO2}) with the atmosphere, the value of the association constant of P_{CO2} was set to 21.66.

The results of the MINEQL$^+$ calculations for each metal are presented in Appendix 4. The free activity of each species k was linked to the calculated free metal concentration by the activity coefficient, f_k (-)

$$(X) = f_k[X] \quad [20]$$

which was calculated using the Davies extension of the Debye-Hückel equation:

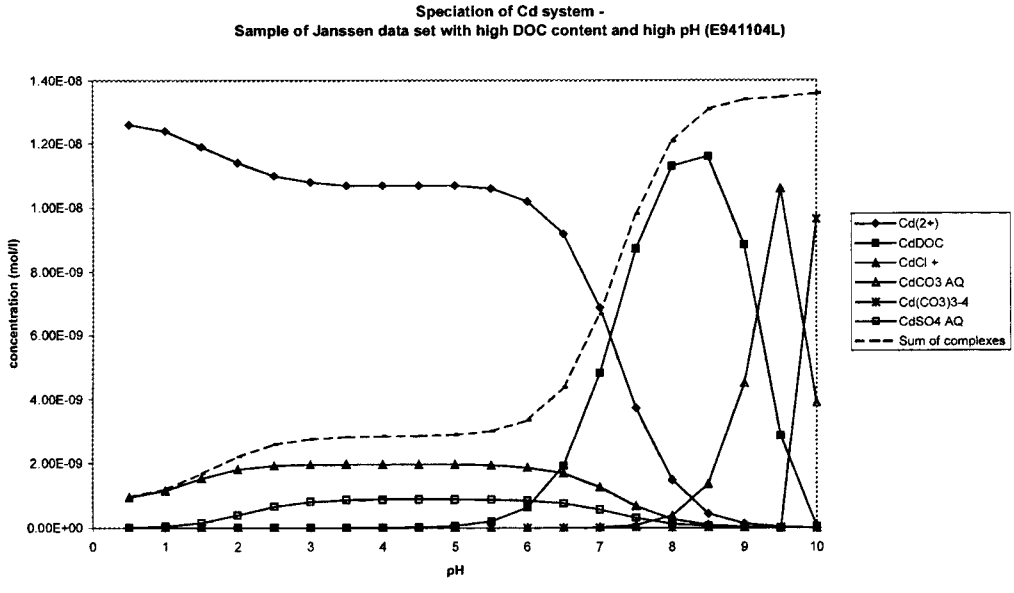
$$\log f_i = -0.51 z_i^2 \left(\frac{\sqrt{I}}{1+\sqrt{I}} - 0.3I \right) \quad [21]$$

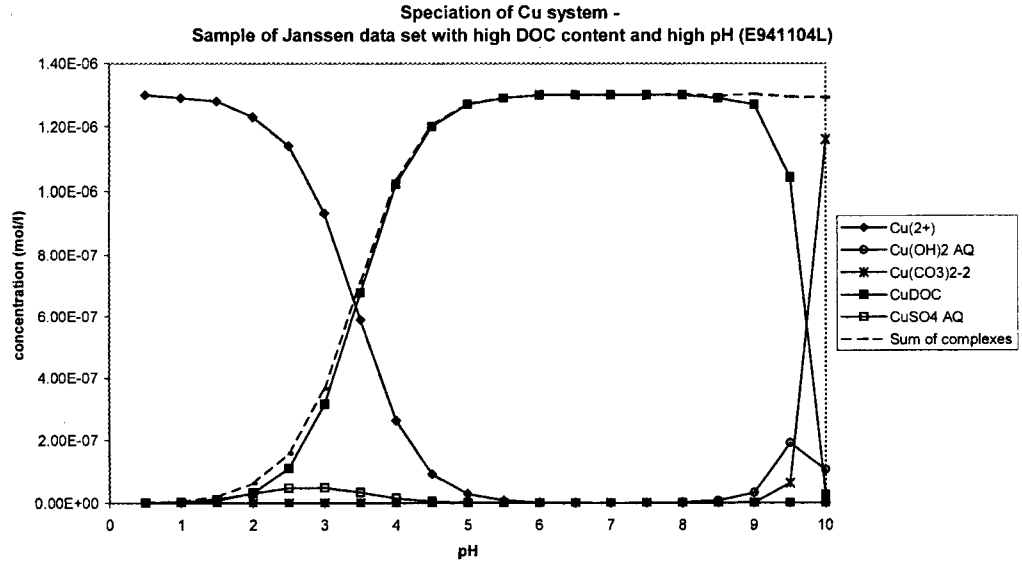
$$I = 0.5 \sum_{i=1}^{m} z_i^2 c_i \quad [22]$$

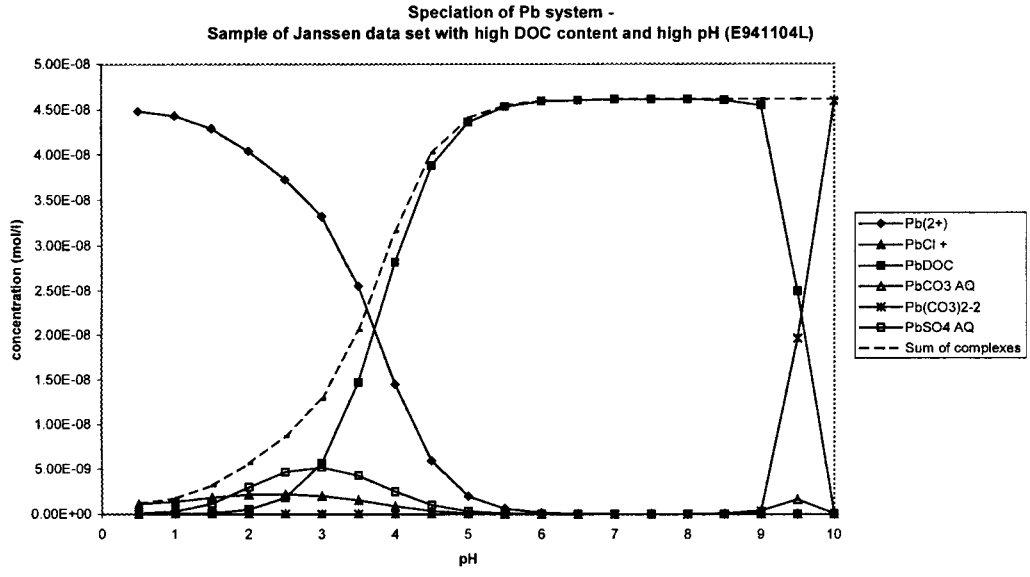
where z_i is valence of ion k in solution, I is ionic strength of the soil solution and m is number of species (Tiktak,1999).

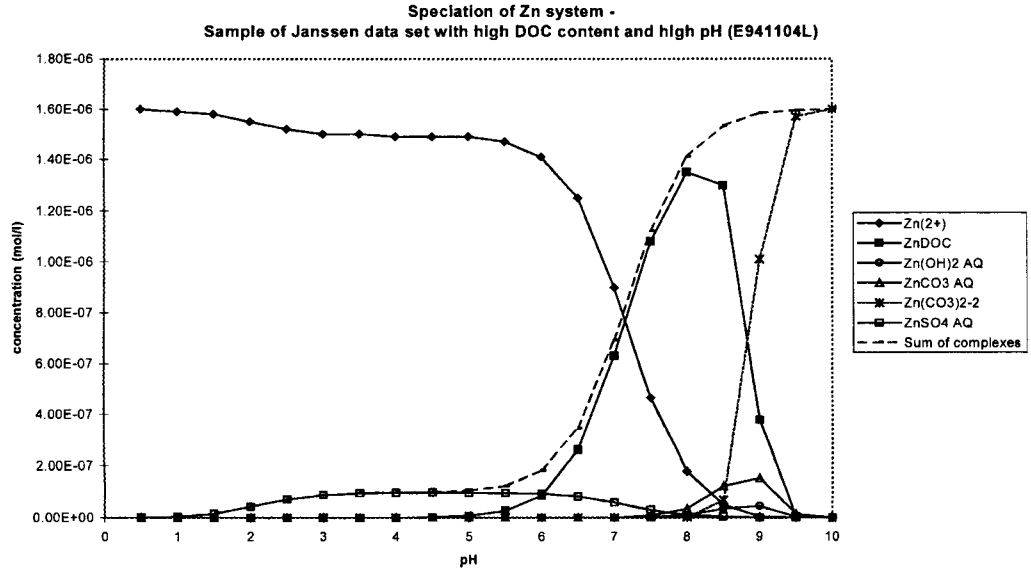
In Figure 2 the speciation of Cd, Cu, Pb and Zn systems calculated with MINEQL$^+$, is presented. A sample of the Janssen data set is used with a high DOC content and a high pH (E941104L).

Figure 2: *Speciation of Cd, Cu, Pb and Zn systems calculated with MINEQL$^+$. Sample of the Janssen data set with a high DOC content and a high pH (E941104L)*




Speciation of Cu system -
Sample of Janssen data set with high DOC content and high pH (E941104L)
concentration (mol/l)
1.40E-06
1.20E-06
1.00E-06
8.00E-07
6.00E-07
4.00E-07
2.00E-07
0.00E+00
0 1 2 3 4 5 6 7 8 9 10
pH
Cu(2+)
Cu(OH)2 AQ
Cu(CO3)2-2
CuDOC
CuSO4 AQ
Sum of complexes


Speciation of Pb system -
Sample of Janssen data set with high DOC content and high pH (E941104L)
concentration (mol/l)
5.00E-08
4.50E-08
4.00E-08
3.50E-08
3.00E-08
2.50E-08
2.00E-08
1.50E-08
1.00E-08
5.00E-09
0.00E+00
0 1 2 3 4 5 6 7 8 9 10
pH
Pb(2+)
PbCl +
PbDOC
PbCO3 AQ
Pb(CO3)2-2
PbSO4 AQ
Sum of complexes


Speciation of Zn system -
Sample of Janssen data set with high DOC content and high pH (E941104L)
concentration (mol/l)
1.80E-06
1.60E-06
1.40E-06
1.20E-06
1.00E-06
8.00E-07
6.00E-07
4.00E-07
2.00E-07
0.00E+00
0 1 2 3 4 5 6 7 8 9 10
pH
Zn(2+)
ZnDOC
Zn(OH)2 AQ
ZnCO3 AQ
Zn(CO3)2-2
ZnSO4 AQ
Sum of complexes

3. Results and Discussion

3.1 LR fits of Freundlich isotherms on log-transformed data for soils in use for Nature conservation

In Table 5 the regression results are presented for the LR Freundlich sorption isotherms of Cd, Cu, Pb and Zn with free metal activity and total metal concentration as one of the independent parameters. The Freundlich isotherms are fitted on log-transformed data for soils in use for nature conservation.

Table 5: *Regression results for the LR Freundlich sorption isotherms of Cd, Cu, Pb and Zn with free metal activity and total metal concentration as one of the independent parameters. Each model is given by:*

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(Ca)^{t}(H^{+})^{m}(M^{2+})^{n}$$

Metal	A/C[a]	Regression coefficients[bc]						R^2	N[d]
		p	q[e]	r	s	m	n		
Cd	A	0.030	= 0			-0.448	0.808	0.674	28
Cu	A	-1.410	= 0			-0.91	0.717	0.570	30
Pb	A	0.442	= 0		-0.223	-0.716	0.556	0.550	29
Zn	A	-0.318	= 0		0.52	-0.445	0.66	0.716	30
Cd	A	0.030		= 0	= 0	-0.448	0.808	0.674	28
Cu	A	-1.410		= 0	= 0	-0.91	0.717	0.570	30
Pb	A	0.874	-0.338	= 0	= 0	-0.712	0.533	0.563	29
Zn	A	-1.123	0.77	= 0	= 0	-0.400	0.72	0.742	30
Cd	C	0.452	= 0			-0.338	0.839	0.665	28
Cu	C	-0.103	= 0		0.47	-0.199	0.26	0.564	30
Pb	C	2.064	= 0			-0.186	0.571	0.591	29
Zn	C	-0.104	= 0		0.64	-0.360	0.77	0.729	30
Cd	C	0.452		= 0	= 0	-0.338	0.839	0.665	28
Cu	C	-0.781	0.58	= 0	= 0	-0.156	0.13	0.536	30
Pb	C	2.064		= 0	= 0	-0.186	0.571	0.591	29
Zn	C	-0.922	0.85	= 0	= 0	-0.300	0.80	0.739	30

[a] Free metal activity (A), total metal concentration (C) as independent parameter in the model
[b] This research uses mg/kg for X, meq/kg for CEC (unbuffered) and mg/l for (M^{2+})
[c] This research uses mg/kg for X, meq/kg for CEC (unbuffered) and mg/l for [M]
[d] N is number of observations
[e] = 0: regression coefficient is set to zero. The regression coefficient of (Ca) is always set to zero.

From Table 5, it appears that:
- the R^2 of the isotherms with CEC and the isotherms with clay and OC are comparable;
- the R^2 of the Zn isotherms are the highest, the R^2 of the Cd are somewhat lower and the R^2 of the Cu and Pb isotherms are the lowest;
- in most of the isotherms only the constant, the free metal activity or the metal concentration and the proton activity (H^+) are significant. The activity isotherms for Pb and Zn and the concentration isotherms for Cu and Zn however include a significant clay content percentage. The organic carbon content percentage is for none of the isotherms significant. Further include the activity isotherms for Pb and Zn and the concentration isotherms for Cu and Zn include a significant CEC.

3.2 NLLSR fits of 2-phase isotherms on original data for soils in use for Nature conservation

In Table 6 the regression results are presented for the NLLSR sorption isotherms of Cd, Cu, Pb and Zn with free metal activity, total metal concentration and the amount of metal extracted with $0.01M$ $CaCl_2$ as one of the independent parameters. The isotherms are fitted on untransformed data for soils in use for nature conservation.

For each metal a maximum of three different fitted models are reported:
1. a model with the highest explained variance;
2. a model with only significant regression coefficients at the p=0.05 level;
3. an intermediate model with a selection of non-significant coefficients but which appear more plausible than (2) from a theoretical point of view. For example the NLLSR concentration isotherm with only significant coefficients for Cu (model 7 in Table 6), does not contain the metal concentration in solution. From a theoretical perspective it is not plausible that the total metal content in the solid phase is completely independent of the total metal concentration in solution. Therefore a model which contains the total metal concentration, with a non-significant coefficient, is also reported (model 6 in Table 6).

The regression results are presented in Appendix 5. A selection is presented in Table 6.

Table 6: *Sorption isotherms of Cd, Cu, Pb and Zn, with free metal activity, total metal concentration or the amount of metal extracted with 0.01M $CaCl_2$ as independent parameter, fitted by Nonlinear Least Squares Regression on the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\{M\}^{n}$$

Model[a]	Metal	A/C/ CA[b]	Regression coefficients[d] a	b	c	10^p	r	s	m	n	R^2	N[c]
1	Cd	A				1.895			-0.341	0.581	0.593	27
2	Cd	A							-0.382	0.585	0.588	27
3	Cd	C							-0.314	0.519	0.529	25
4	Cd	CA							-0.229	0.428	0.801	27
5	Cu	A							-0.471	0.348	0.202	30
6	Cu	C	-25.61	14.072			-1.358		-0.177	-0.344	0.426	26
7	Cu	C	-16.31	11.385					-0.146		0.412	26
8	Cu	CA	10.629		0.298		2.123	1.333	-0.158	1.550	0.889	26
9	Pb	A	53.248						-0.817	0.674	0.445	28
10	Pb	A							-0.829	0.664	0.371	28
11	Pb	C		-1.499		163.75			-0.148	0.481	0.551	29
12	Pb	C				152.63			-0.158	0.511	0.539	29
13	Pb	CA		29.056				0.993	-0.373	1.317	0.938	14
14	Zn	A							-0.470	0.505	0.782	27
15	Zn	C							-0.433	0.588	0.802	30
16	Zn	CA			1.849		-0.832		-0.495	0.547	0.936	28

[a] Model 1, 6, 9 and 11 are models with the most probable independent parameters
Model 2, 3, 4, 5, 7, 8, 10 and 12 - 16 are models with significant regression coefficients at the p=0.05 level

[b] Free metal activity (A), total metal concentration (C) or the amount of metal extracted with 0.01M $CaCl_2$ (CA) as independent parameter in the model

[c] N is number of datapoints

[d] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

From Table 6, it appears that:

- the R^2 of models of type 2 for Cu, Pb and Zn with total metal concentration are higher than for models with free metal activity. For Cd the R^2 of the type 2 with free metal activity is somewhat higher;
- the models of type 2 with free metal activity include only the proton activity (H^+) and the free metal activity. The inert phase is never significant at the p=0.05 level. There is no dependence of clay and organic carbon;
- the models of type 2 for Cd, Zn and Pb with total metal concentration include the proton activity (H^+) and the total metal concentration. The inert part of these NLLSR models is however not significant at the p=0.05 level. Only for Cu the sorption isotherm includes (statistically) significant inert metal fraction. Further it can be mentioned that only for Pb

the sorption isotherm includes a regression constant. Moreover, there is only a organic carbon dependence for Cu and Pb and for none of the metals there is a clay dependence;
- the models of type 2 for Cu with the total metal concentration yields a negative regression coefficient for the inert fraction. This result is not plausible from a theoretical perspective and may lead to errouneous prediction for low values of M;
- the models of type 3 yield slightly higher estimated explained variances than the models of type 2 for each metal except for Pb with free metal activity as independent parameter.

3.3 NLLSR fits of 2-phase isotherms on original data for soils in use for agriculture

In Table 7 the regression results are presented for the NLLSR sorption isotherms of Cd, Cu, Pb and Zn with the amount of metal extracted with $0.01M$ $CaCl_2$ as one of the independent parameters. The isotherms are fitted on untransformed data for soils in use for agriculture.

Table 7: *The NLLSR sorption isotherms of Cd, Cu, Pb and Zn with the amount of metal extracted with 0.01M $CaCl_2$ as independent parameter, derived from the untransformed of the LMB data set are resumed in the undermentioned table. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]	Metal	Regression coefficients[c]									
		a	b	c	10^p	r	s	m	n	R^2	N[b]
1	Cd	-0.728	0.018	0.017			-0.152	-0.027	0.056	0.801	31
2	Cd	0.132	0.015	0.008						0.752	31
3	Cu	9.502		0.289		1.729		-0.271	1.313	0.660	39
4	Pb	7.910	1.109	1.102		8.376		-4.460	11.952	0.812	30
5	Zn	12.889	3.059	4.041				-0.165	1.351	0.855	33
6	Zn		3.615	4.142				-0.145		0.841	33

[a] Model 1 and 5 are models with the most plausible independent parameters
Model 2, 3, 4 and 6 are models with significant regression coefficients at the p=0.05 level
[b] N is number of datapoints
[c] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

From Table 7, it appears that:
- the models derived from the LMB data set include for all metals non-reactive fractions;
- the pH dependence for Cd, Cu and Zn is very small, for Pb is the coefficient for the proton activity (H^+) not plausible from a theoretical perspective. The proton activity should yield a coefficient between –1 and 0;
- the coefficient for the metal content extracted with $0.01M$ $CaCl_2$ is above 1 for Cu, Pb and Zn and for Cd is this coefficient not significant. The metal content extracted with $0.01M$ $CaCl_2$ should yield a coefficient between 0 and 1;
- the R^2–values of all models are relatively high.

3.4 Evaluation of isotherms based on batch data versus field data

The fitted isotherms were compared to alternative isotherm sets for Cd, Cu, Pb and Zn based on batch data, fitted by:

- Bril (Bril, 1999): The sorption data are based on a small selection of literature sources. Only measured metal activities were used, which in fact is the selection criterion for using a published data set. The regression procedure is a three-step procedure. First the mean Freundlich exponent n is determined, which is fixed. Next the data of each individual isotherm are fitted with a fixed Freundlich n coefficient, yielding a regression constant. Finally the regression constants are fitted to the soil and solution parameters;
- Elzinga (Elzinga et al., 1997, 1999): The sorption data are based on a wide selection of literature sources. Data are inferred from the original sources by either digitizing the graphical isotherms or by evaluating isothermplots. The original solution data are concentration data and transformed to free activities by correction for activity and complexation effects. Necessary CEC data are sometimes and DOC data are always inferred from data on clay and organic matter contents. In the regression analysis, all inferred data points are considered as independent observations. Freundlich isotherms are fitted both for concentrations and activities. In general, regression Freundlich coefficients strongly depend on using concentration or activity data, but the explained variance is unaffected.

The CEC values of the data sets of Bril (1999) and Elzinga et al. (1997) are the buffered CEC values determined at a pH of 7 or 8. At such high pH values not only the non-specific but also the specific sorption sites of the solid phase participate in the measurement of CEC (Mott, 1988). The CEC values of the Hoop-Janssen data set were measured unbuffered. An unbuffered CEC measurement is performed at the pH of the soil. The type of sorption sites that participate in this measurement differs for each soil, because the pH value of each soil differs (Elzinga et al., 1997). CEC values at pH=8 were inferred from the Hoop-Janssen data set using a model used by Bril (1999):

$$CEC(eq/kg) = 0.06\{\%OC\}^{0.75} + 0.004\{\%clay\} \qquad [23]$$

A selection has been made of the LR Freundlich sorption isotherms with free metal activity (and concentration as well for this research) as one of the independent parameters fitted by Bril, Elzinga and fitted in this research. This selection is presented in Table 8. The selected isotherms are compared by means of a validation against the field data set of Hoop-Janssen. The results of this comparison are presented in Figure 3 to 6. A selection of the fitted NLLSR isotherms is compared by means of validation and also presented in Figure 3 to 6.

Table 8: *Regression results for the LR Freundlich sorption isotherms of Cd, Cu, Pb and Zn with free metal activity as one of the independent parameters fitted by Bril (Bril, 1999), Elzinga (Elzinga et al., 1997) and fitted in this research. Each model is given by:*

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(Ca)^{t}(H^{+})^{m}(M^{2+})^{n}$$

Research	Metal	Regression coefficients							R^2	N^{ef}
		p	q[g]	r	s	m	n	t[g]		
Bril[a]	Cd	-3.05	1.00		-0.24	-0.50	0.82	-0.41		
	Cu	-3.85	0.52	0.46	-0.14	-0.70	0.55	-0.27		
	Pb	-4.4	0.624	0.462		-0.60	0.55	-0.27		
	Zn	-3.42	1.30			-0.75	0.75	-0.38		
Elzinga[b]	Cd	-1.92	0.659			-0.444	0.831		0.755	1125
	Cu	-2.83	0.861			-0.744	0.733		0.652	386
	Pb									
	Zn	-1.83	0.866	0.216	-0.295	-0.448	0.730		0.859	261
This research[c]	Cd	0.030	=0			-0.448	0.808	=0	0.674	28
	Cu	-1.410	=0			-0.91	0.717	=0	0.570	30
	Pb	0.442	=0		-0.223	-0.716	0.556	=0	0.550	29
	Zn	-0.318	=0		0.52	-0.445	0.66	=0	0.716	30
This research[d]	Cd	0.452	=0			-0.338	0.839	=0	0.665	28
	Cu	-0.103	=0		0.47	-0.199	0.26	=0	0.564	30
	Pb	2.064	=0			-0.186	0.571	=0	0.591	29
	Zn	-0.104	=0		0.64	-0.360	0.77	=0	0.729	30

[a] Bril (1999) uses mol/kg for X, eq/kg for CEC (buffered), mol/m^3 for (M^{2+}) and mol/m^3 for (Ca)

[b] Elzinga et al. (1997) uses mg/kg for X, meq/kg for CEC (buffered) and mg/l for (M^{2+}), the calculated R^2 is adjusted for the number of observations

[c] This research uses mg/kg for X, meq/kg for CEC (unbuffered) and mg/l for (M^{2+})

[d] This research uses mg/kg for X, meq/kg for CEC (unbuffered) and mg/l for [M]

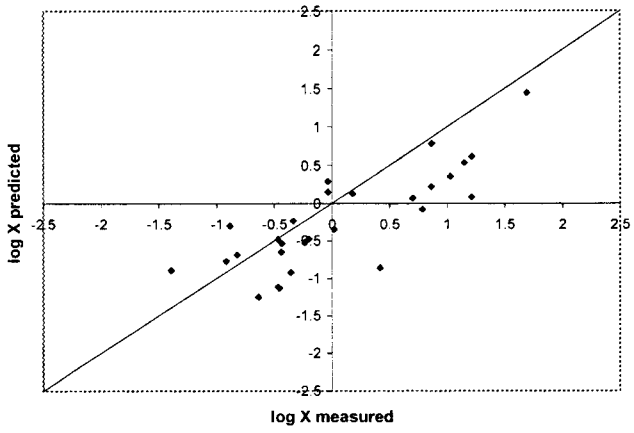
[e] N is number of observations

[f] The models of Bril (1999) consist of several seperately fitted parts. R^2 and N can differ for each part (Reinds at al., 1995)

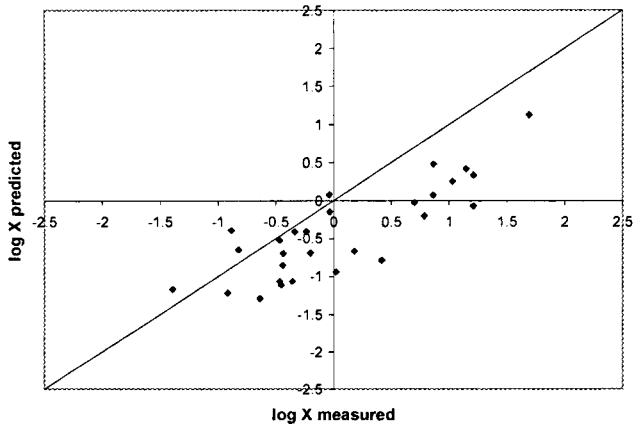
[g] = 0: regression coefficient is set to zero

It is surprising that the coefficient for the proton activity (H^+) is similar for the Elzinga et al. (1997) isotherms and the field-based isotherms because Elzinga uses buffered CEC whereas in this research unbuffered CEC is used. One would expect that the entire effect of pH is incorporated in the coefficient of the proton activity (coefficient m) in the Elzinga isotherms with buffered CEC, which should result in a more negative coefficient m than for the field-based isotherms with unbuffered CEC.

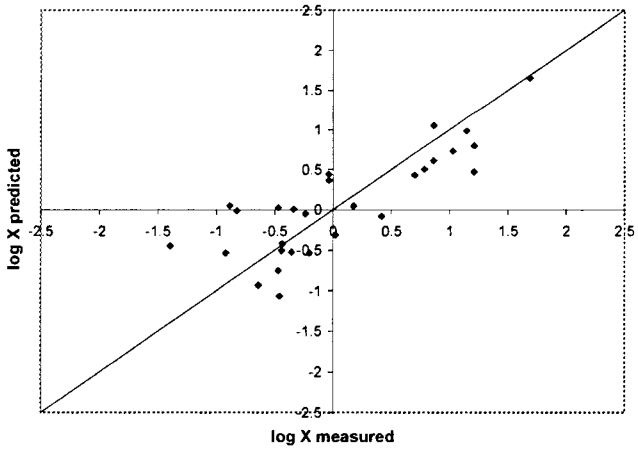
LR-model (Cd^{2+}) Bril: $X = 10^{-3.05}\{CEC\}^{1.00}\{\%clay\}^{-0.24}(Ca)^{-0.41}\left(H^{+}\right)^{-0.5}\left(Cd^{2+}\right)^{0.82}$



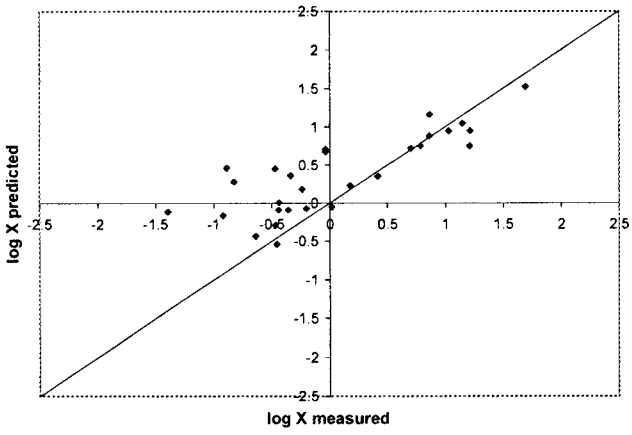
LR-model (Cd^{2+}) Elzinga: $X = 10^{-1.92}\{CEC\}^{0.659}\left(H^{+}\right)^{-0.444}\left(Cd^{2+}\right)^{0.831}$



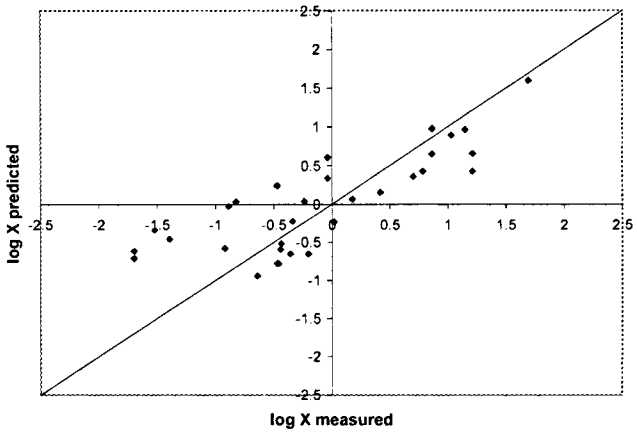
LR-model (Cd^{2+}) field: $X = 10^{0.030}\left(H^{+}\right)^{-0.448}\left(Cd^{2+}\right)^{0.808}$



LR-model [Cd] field: $X = 10^{0.452}\left(H^{+}\right)^{-0.338}[Cd]^{0.839}$



NLLSR-model (Cd^{2+}) field: $X = (H^{+})^{-0.382}\left(Cd^{2+}\right)^{0.585}$



NLLSR-model [Cd] field: $X = (H^{+})^{-0.314}[Cd]^{0.519}$

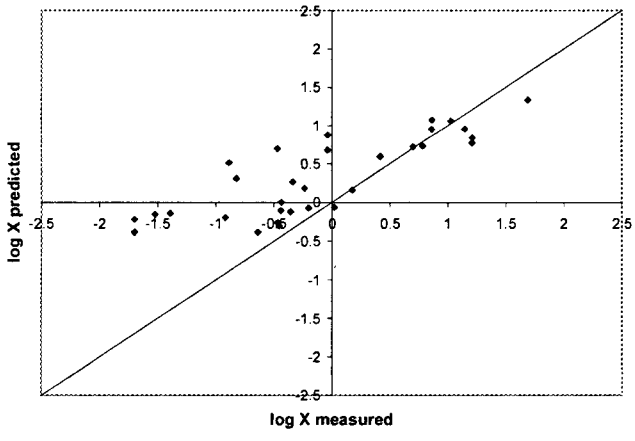
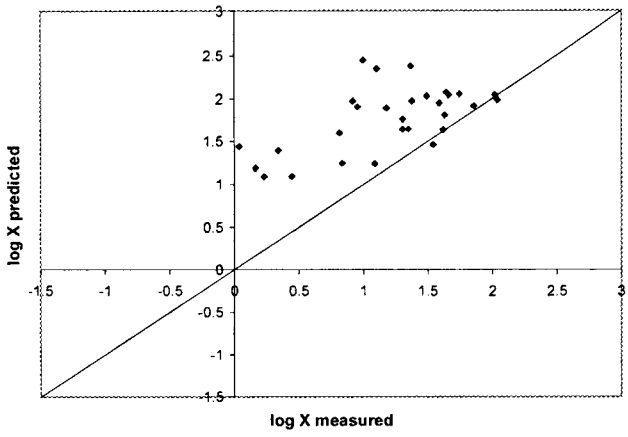
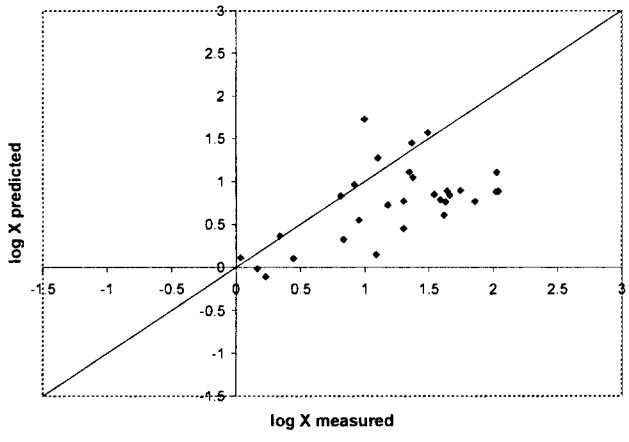


Figure 3: *Comparison of the log-transformed predicted and measured adsorbed Cd content in soils (mg/kg). Cd contents were predicted with batch isotherms (Bril (1999) and Elzinga et al. (1997)) and field isotherms (this research).*

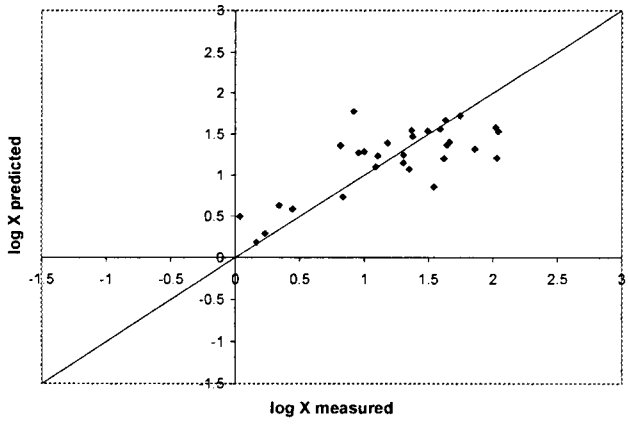
LR-model (Cu^{2+}) Bril: $X = 10^{-3.85}\{CEC\}^{0.52}\{\%OC\}^{0.46}\{\%clay\}^{-0.14}(Ca)^{-0.275}\left(H^{+}\right)^{-0.7}\left(Cu^{2+}\right)^{0.55}$



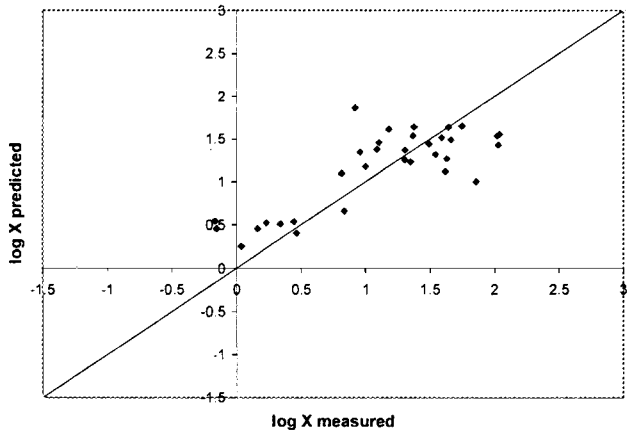
LR-model (Cu^{2+}) Elzinga: $X = 10^{-2.83}\{CEC\}^{0.861}\left(H^{+}\right)^{-0.744}\left(Cu^{2+}\right)^{0.733}$



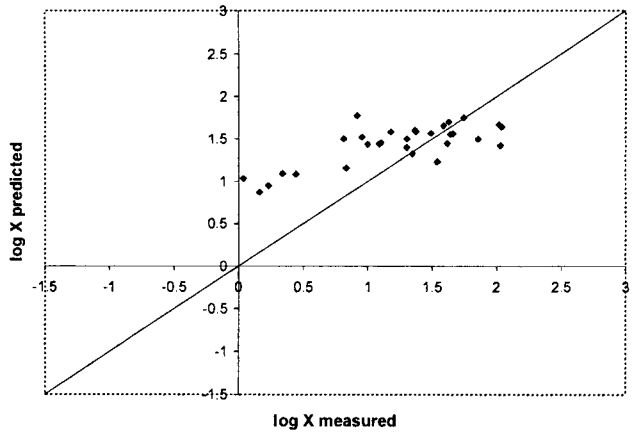
LR-model (Cu^{2+}) field: $X = 10^{-1.410}\left(H^{+}\right)^{-0.91}\left(Cu^{2+}\right)^{0.717}$



LR-model [Cu] field: $X = 10^{-0.103}\{\%clay\}^{0.47}\left(H^{+}\right)^{-0.199}[Cu]^{0.26}$



NLLSR-model (Cu^{2+}) field: $X = (H^{+})^{-0.471}\left(Cu^{2+}\right)^{0.348}$



NLLSR-model [Cu] field: $X = -16.308 + 11.385\{\%OC\} + (H^{+})^{-0.146}$

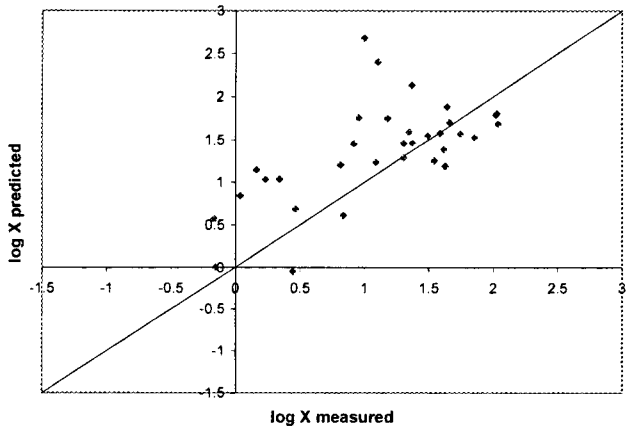
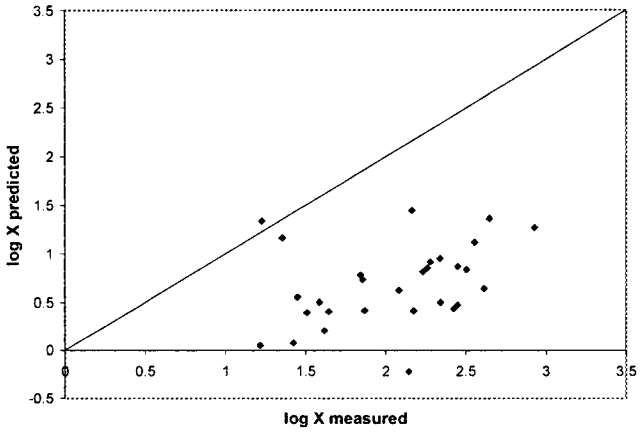
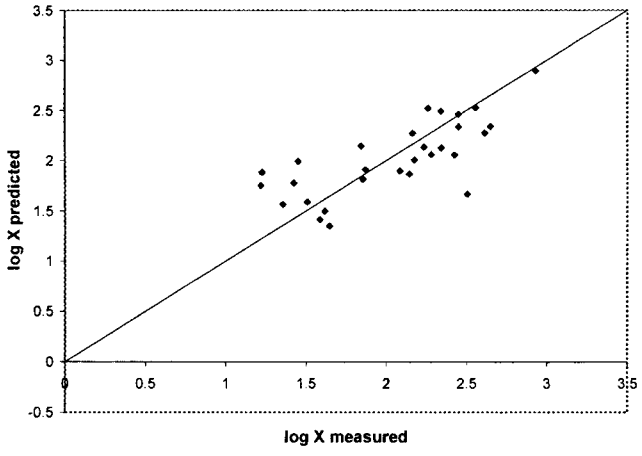


Figure 4: *Comparison of the log-transformed predicted and measured adsorbed Cu content in soils (mg/kg). Cu contents were predicted with batch isotherms (Bril (1999) and Elzinga et al. (1997)) and field isotherms (this research).*

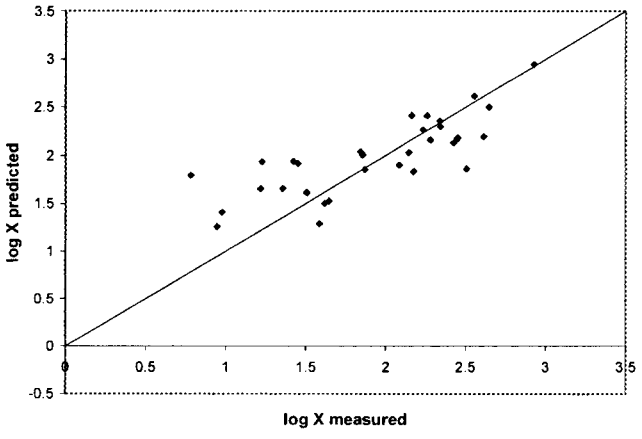
LR-model (Pb^{2+}) Bril: $X = 10^{-4.4}\{CEC\}^{0.624}\{\%OC\}^{0.462}(Ca)^{-0.275}\left(H^{+}\right)^{-0.6}\left(Pb^{2+}\right)^{0.55}$



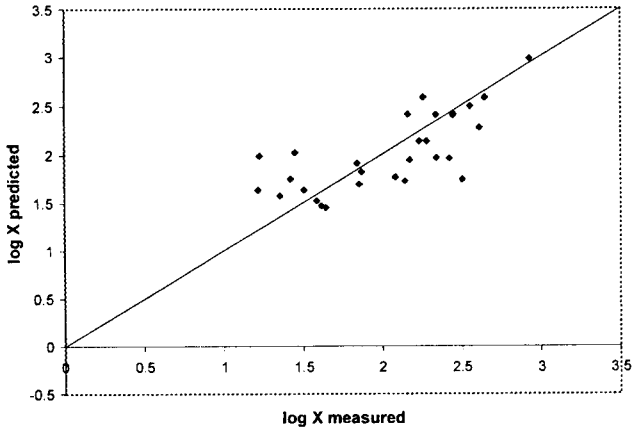
LR-model (Pb^{2+}) field: $X = 10^{0.442}\{\%clay\}^{-0.223}\left(H^{+}\right)^{-0.716}\left(Pb^{2+}\right)^{0.556}$



LR-model [Pb] field: $X = 10^{2.064}\left(H^{+}\right)^{-0.186}[Pb]^{0.571}$



NLLSR-model (Pb^{2+}) field: $X = (H^+)^{-0.829}\left(Pb^{2+}\right)^{0.664}$



NLLSR-model [Pb] field: $X = 152.633(H^+)^{-0.158}[Pb]^{0.511}$

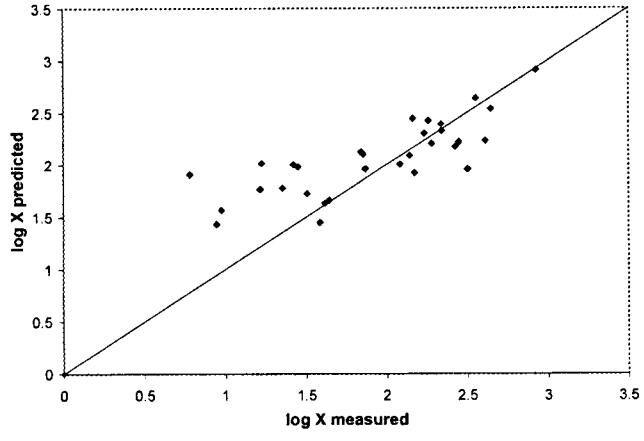
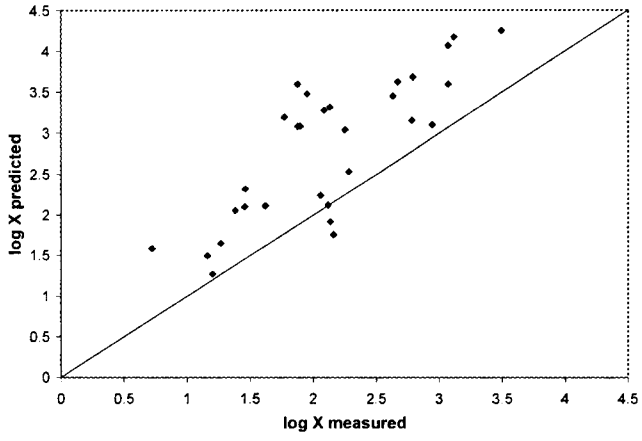
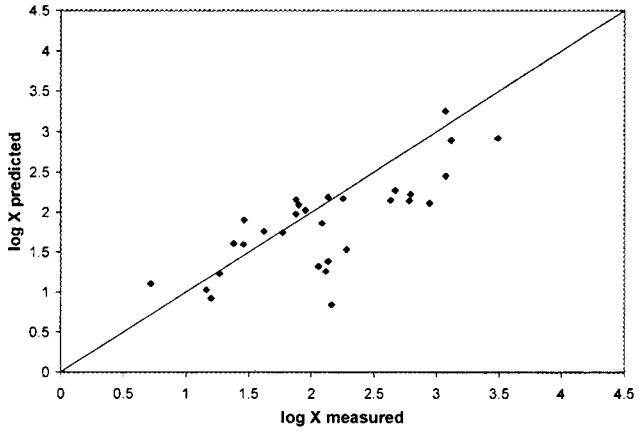


Figure 5: *Comparison of the log-transformed predicted and measured adsorbed Pb content in soils (mg/kg). Pb contents were predicted with a batch isotherms (Bril) and field isotherms (this research).*

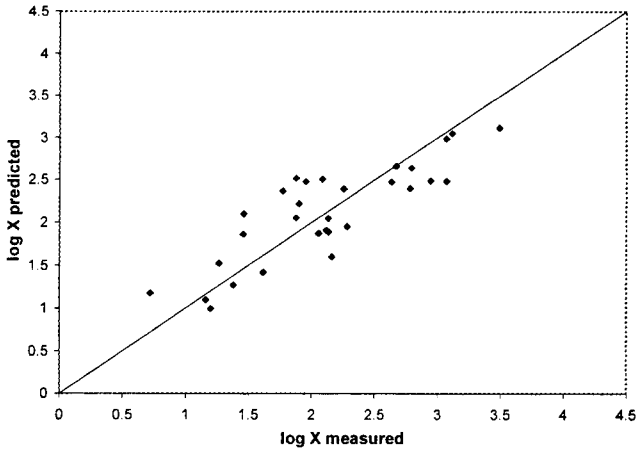
LR-model (Zn^{2+}) Bril: $X = 10^{-3.42}\{CEC\}^{1.3}(Ca)^{-0.375}\left(H^{+}\right)^{-0.75}\left(Zn^{2+}\right)^{0.75}$



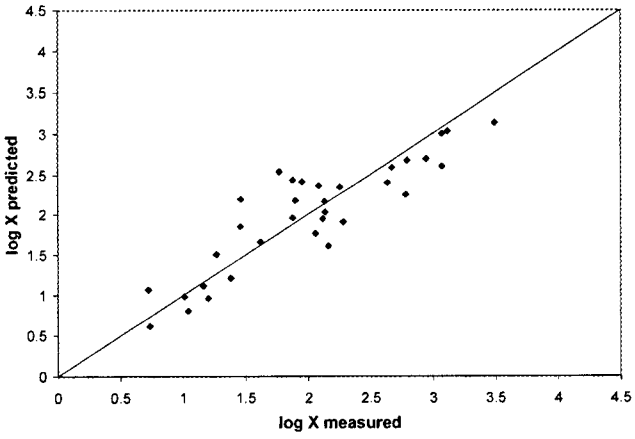
LR-model (Zn^{2+}) Elzinga: $X = 10^{-1.83}\{CEC\}^{0.866}\{\%OC\}^{0.216}\{\%clay\}^{-0.295}\left(H^{+}\right)^{-0.448}\left(Zn^{2+}\right)^{0.73}$



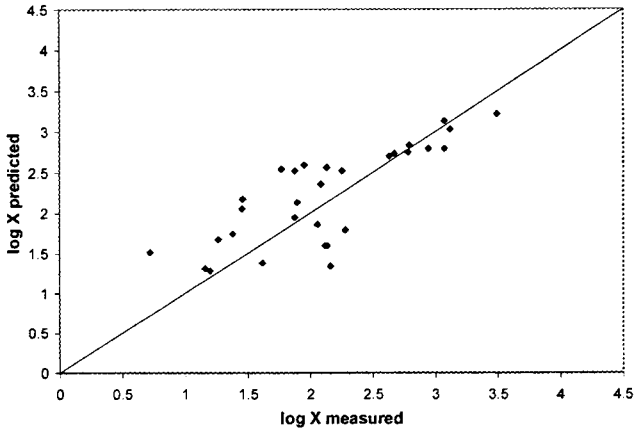
LR-model (Zn^{2+}) field: $X = 10^{-0.318}\{\%clay\}^{0.52}\left(H^{+}\right)^{-0.445}\left(Zn^{2+}\right)^{0.66}$.



LR-model [Zn] field: $X = 10^{-0.104}\{\%clay\}^{0.64}\left(H^{+}\right)^{-0.360}[Zn]^{0.77}$



NLLSR-model (Zn^{2+}) field: $X = (H^{+})^{-0.470}\left(Zn^{2+}\right)^{0.505}$



NLLSR-model [Zn] field: $X = (H^{+})^{-0.433}[Zn]^{0.588}$

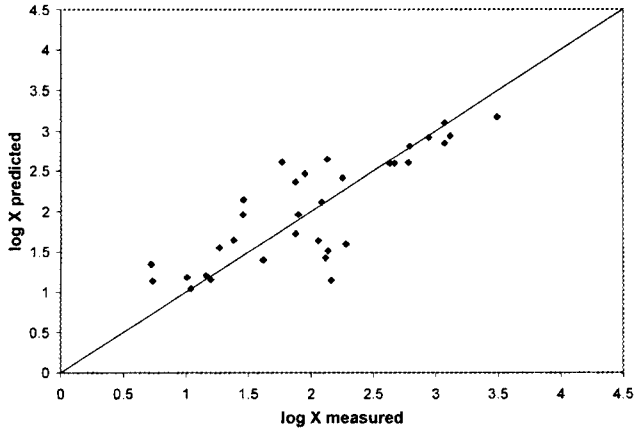


Figure 6: *Comparison of the log-transformed predicted and measured adsorbed Zn content in soils (mg/kg). Zn contents were predicted with batch isotherms (Bril (1999) and Elzinga et al. (1997)) and field isotherms (this research).*

When one compares the LR activity models of Bril (1999) and Elzinga et al. (1997) with the field observations of Hoop and Janssen (Table 9), it appears that:

- the Elzinga models predict the observed metal content in the solid phase in the field reasonably well, there is a slight underestimation at the higher range of the concentrations;
- the Bril models predict the the observed metal content in the solid phase in the field less well. For Cd the observations are only slightly underestimated, but for Pb there is a large underestimation of the observations. For Cu and Zn the Bril model overestimates the observations. This is contrary to expectation because from a theoretical perspective it is probable that models derived from batch experiments underestimate the observations. This is caused by differences in reaction time in batch experiments and in the field. Although rates of metal sorption decrease rapidly within one day, hypothetical sorption or desorption fluxes for additional equilibrium times of tens to hundreds of years may be considerable (Elzinga et al., 1999).

When one compares the LR activity models of Bril (1999) and Elzinga et al. (1997) with the LR activity model derived from the field data of Hoop and Janssen, it appears that:

- the field activity models predicts the observed metal content in the solid phase best. This can partly be explained by the fact that the model is derived from the same data as used in this validation. For Cu however, the prediction is not very well. There is an overestimation at the low concentrations and an underestimation at the high concentrations. This explains the low explained variance of this model.

When one compares the LR models with the NLLSR models derived from the field data of Hoop and Janssen, it appears that:

- the LR models predict the low concentrations well, but underestimate the high concentrations and the NLLSR models predict the high concentrations well, but overestimate the low concentrations. This can be explained by the fact that for the Linear Regression procedure log-transformed values are used, which give more weight to the low concentrations. For the Nonlinear Least Squares Regression procedure however untransformed values are used, which give more weight to the high concentrations.

When one compares the activity models with the concentration models derived from the field data of Hoop and Janssen, it appears that:

- the prediction of the LR concentration models derived from the field data is comparable with the LR activity models. The same applies for the NLLSR models.

Table 9: *Comparison of the observed values and the predicted values with the LR-models of Cd, Cu, Pb and Zn with free metal activity as one of the independent parameters fitted by Bril (Bril, 1999), Elzinga (Elzinga et al., 1997) and the LR- and NLLSR models fitted in this research.*

Model	Metal							
	Cd		Cu		Pb		Zn	
	low conc.	high conc.	low conc.	high conc.	low conc.	high conc.	low conc.	high conc.
LR-model (M^{2+}) Bril	-	-	+	+	-	-	+	+
LR-model (M^{2+}) Elzinga	-	-	0	-	n.a.	n.a.	0	-
LR-model (M^{2+}) field	0	-	+	-	0	0	0	-
LR-model [M] field	0	-	0	-	+	0	0	-
NLLSR-model (M^{2+}) field	+	0	+	0	0	0	+	0
NLLSR-model [M] field	+	0	+	+	+	0	0	0

\+ = overestimation of the observed metal content in the solid phase
\- = underestimation of the observed metal content in the solid phase
0 = no over- or underestimation of the observed metal content in the solid phase
n.a. = not available

3.5 Evaluation of isotherms fitted with LR and with NLLSR

In spite of the fact that models derived from the log-transformed observations with Linear Regression (procedure a) and models derived from the original observations with Nonlinear Regression (procedure b) are comparable, models fitted with procedure b are preferred, regarding the aim of this research to obtain isotherms for the model SOACAS. The reason for this is the fact that for the Linear Regression procedure log-transformed values are used, which give more weight to the low concentrations. For the Nonlinear Least Squares Regression procedure, however, untransformed values are used, which give more weight to the high concentrations. As a consequence the LR models predict the low concentrations well, but underestimate the high concentrations and the NLLSR models predict the high concentrations well, but overestimate the low concentrations. For the model SOACAS the purpose is especially to predict the high metal concentrations well. The performance of alternatively obtained LR and NLLSR sorption models was evaluated by comparing predicted and observed X with Linear Regression. It appears that the explained variances of the NLLSR models are comparable with the explained variances of the LR models (Table 10).

Table 10: *Results for the Linear Regression of the untransformed values of the observed total metal content in the solid phase and the predicted total metal content in the solid phase according to the fitted LR and NLLSR sorption isotherms of Cd. Each model is given by:*

$$X_{measured} = bX_{predicted}$$

Model[a]	Metal	Regr. coeff. b	R^2	N[b]
1	Cd	1.138	0.899	28
2	Cd	1.223	0.901	28
3	Cu	1.223	0.548	30
4	Cu	1.005	0.601	30
5	Pb	1.088	0.853	29
6	Pb	1.054	0.865	29
7	Zn	1.602	0.823	30
8	Zn	1.277	0.853	30

[a] Model 1, 3, 5 and 7 are LR models with free metal activity as independent parameter
Model 2, 4, 6 and 8 are NLLSR models with free metal activity as independent parameter
[b] N is number of datapoints

3.6 Validation of isotherms based on data for soils in use for Nature conservation for application on agricultural soils

The NLLSR sorption isotherms for Cd, Pb and Zn derived from the LMB data set, with the amount of metal extracted with $0.01M$ $CaCl_2$ (Table 7), have a different form and different parameters than the NLLSR isotherms derived from the Hoop-Janssen data set (Table 6). The equations derived from the LMB data set more often include non-reactive fractions and the pH dependence is much smaller than with the equations derived from the Hoop-Janssen data set. For Cu the equations for agricultural soil and nature conservation soil are fairly comparable (Table 11).

When both sets of equations sets are used to calculate the total metal content in the solid phase under standard soil conditions (OM = 10%, clay = 25%) for a pH range of 3 – 9, divided in an immobile and mobile part, it appears that for a pH range of 5 – 6:

- the predicted total metal contents in the solid phase for Cu and Zn are of the same magnitude. The inert part is comparable for Cu and Zn as well;
- the predicted total metal contents in the solid phase for Cd with the isotherm for the "agricultural" data set is one order of magnitude smaller than for the equation derived from the "nature conservation" data set;
- the predicted total metal contents in the solid phase for Pb are not comparable at all;

For the pH range beyond 5 – 6 the results are not comparable, except for Cu.

Table 11: *The NLLSR sorption isotherms of Cu with amount of metal extracted with 0.01M $CaCl_2$ as independent parameter, derived from the untransformed values of the combined Hoop-Janssen data set and the LMB data set are resumed in the undermentioned table. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Metal	Data set	Regression coefficients[a]								R^2	N
		a	b	c	10^p	r	s	m	n		
Cu	Hoop-Janssen	10.629		0.298		2.123	1.333	-0.158	1.550	0.889	26
Cu	LMB	9.502		0.289		1.729		-0.271	1.313	0.660	39

[a] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

3.7 Application perspective of isotherms based on 0.01*M* $CaCl_2$ extracts

The explained variance for the NLLSR with the metal concentration in the 0.01*M* $CaCl_2$ extraction was much higher for Cd, Cu, Pb and Zn than for the regression with the total metal concentration. For Cu, Pb and Zn more model parameters are significant. To compare the use of the total metal concentration in solution to the amount of metal extracted with $CaCl_2$ as an independent parameter in the sorption isotherm, Linear Regression was performed between the total metal concentration in solution and the amount of metal extracted with $CaCl_2$ of the Hoop-Janssen data set for each metal (Table 12). Linear Regression between the pH of the pore water and the pH according to the $CaCl_2$ extraction was also performed (Table 13).

From the Linear Regression between the total metal concentration in solution and the amount of metal extracted with 0.01*M* $CaCl_2$ in the Hoop-Janssen data set, it appears that these parameters correlate well for Cd, not at all for Cu and rather well for Pb and Zn (Table 12). The intercept of the regression equations are for all metals zero.

Table 12: *Regression results for the Linear Regression of the untransformed values of the total metal concentrations and the metal contents according to $CaCl_2$ extraction of the combined Hoop-Janssen data set. Each model is given by:*

$$[M_{CaCl_2}] = a + b[M]$$

Model[a]	Metal	Regr. coeff. a	b	R^2	N[b]
1	Cd	-0.000	1.036	0.992	30
2	Cd	0.000	0.668	0.836	29
3	Cu	0.000	-0.004	0.000	26
4	Pb	0.000	4.204	0.628	15
5	Zn	-0.002	2.082	0.722	31

[a] Model 1, 3, 4 and 5 are models with outliers
Model 2 is the model without outliers
[b] N is number of datapoints

Table 13: *Regression results for the Linear Regression of pH according to the $CaCl_2$ extraction and pH of the pore water of the combined Hoop-Janssen data set. The model is given by:*

$$pH_{CaCl_2} = a + b \times pH_{porewater}$$

Model	Regr. coeff. a	b	R^2	N[a]
1	0.568	0.847	0.936	33

[a] N is number of datapoints

4. Conclusions and recommendations

4.1 Conclusions

Selection of most suitable isotherms for SOACAS

From the results of the NLLSR fits of 2-phase isotherms on original data for soils in use for nature conservation, it can be concluded that:

- the models with free metal activity include only the proton activity (H^+) and the free metal activity. The inert phase is never statistically significant;
- the models for Cd, Zn and Pb with total metal concentration include the proton activity (H^+) and the total metal concentration. The inert part of these NLLSR models is however not significant at the p=0.05 level. Only for Cu the sorption isotherm includes (statistically) significant inert metal fraction. This implies that a 2-phase model, in most cases, can not be derived from this data set.
- the R^2 of models for Cu, Pb and Zn and with total metal concentration are higher than for models with free metal activity. For Cd the R^2 of the model with free metal activity is somewhat higher. These results are in contradiction with the fact that the free metal activity should describe the sorption behaviour best, compared with the total metal concentration in solution, because a correction for ion strength and (inorganic and organic) complexation in solution has taken place. Therefore free activity models are considered as the most suitable isotherms for SOACAS (Table 14).

From the comparison of the LR fits on log transformed data and the NLLSR fits on original dat for soils in use for nature conservation, it can be concluded that: regarding the aim of this research to obtain isotherms for the model SOACAS, the NLLSR fits are preferred, in spite of the fact that the models are comparable. The reason for this is the fact that for the Linear Regression procedure log-transformed values are used, which give more weight to the low concentrations. For the Nonlinear Least Squares Regression procedure however untransformed values are used, which give more weight to the high concentrations. As a consequence the LR models predict the low concentrations well, but underestimate the high concentrations and the NLLSR models predict the high concentrations well, but overestimate the low concentrations. For the model SOACAS the purpose is especially to predict the high metal concentrations well.

Table 14: *Sorption isotherms of Cd, Cu, Pb and Zn, with free metal activity fitted by Nonlinear Least Squares Regression on the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left(M^{2+}\right)^{n}$$

Metal	Regression coefficients[b]									
	a	b	c	10^p	r	s	m	n	R^2	N[a]
Cd							-0.382	0.585	0.588	27
Cu							-0.471	0.348	0.202	30
Pb							-0.829	0.664	0.371	28
Zn							-0.470	0.505	0.782	27

[a] N is number of datapoints

[b] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

From the comparison of isotherms based on data for soils in use for nature conservation and isotherms based on data for soils in use for agriculture, it appears that only for Cu the isotherms are comparable. The isotherms for Cd, Pb and Zn derived from the agricultural soils include non-reactive fractions and the pH dependence is much smaller than for the isotherms derived from the Hoop-Janssen data set. This would imply that the isotherms based on data for soils in use for nature conservation, except for Cu, are not applicable for soils in use for agriculture. This should be kept in mind when using these isotherms.

The isotherms derived from the field data set of Hoop and Janssen do not depend on the clay content and the organic carbon content in soils. A consequence is that when with these isotherms are used as input for SOACAS, little regional variation in metal content is expected. The isotherms derived from the batch data of Bril (1999) and Elzinga et al. (1997) consist of a reactive metal fraction that depends on the clay content and the organic carbon content. Observed regional patterns of heavy metal contents seem to correspond to regional variation of clay and organic matter.

Added value compared to batch isotherms

From a theoretical perspective it is probable that models derived from batch experiments underestimate field observations, because of differences in reaction time in batch experiments and in the field. Although rates of metal sorption decrease rapidly within one day, hypothetical sorption or desorption fluxes for additional equilibrium times of tens to hundreds of years may be considerable (Elzinga et al., 1999).

The performance of isotherms derived from batch data (Bril (1999) and Elzinga et al. (1997)) and from field data (this research) was compared by means of a validation against the field data set of Hoop-Janssen. From this comparison, it appears that the LR activity models derived from field data predicts the observed metal content in the solid phase best. The LR activity models of Elzinga predict reasonably accurate for Zn and with slight underestimation for Cd and Cu. The LR activity models of Bril predict the observations less well. For Cu and Zn the models overestimate the observations, for Pb the model has a large underestimation. On the contrary for Cd the model of Bril slightly underestimates the observations.

Remaining short-coming

To judge the validity and applicability of the fitted sorption isotherms, several points should be considered:

- reliability of the observations:

1. a part of the data sets is less reliable because of values that lie near detection limits. To derive sorption isotherms that are more reliable, data sets with more reliable values are necessary;
2. reliability of pore water concentration data: from the high explained variance of the $CaCl_2$ isotherms in comparison with the total concentration isotherms, it appears that derivation of in situ sorption isotherms is hampered by the "noise" in the pore water data. This noise is probably caused by the tedious and laborious procedure in which the pore water data were determined. The $CaCl_2$ isotherms can not be used directly in SOACAS because LR between the total metal content extracted with $0.01M$ $CaCl_2$ and the total pore water concentration showed that only for Cd these parameters correlate well;

3. reliability of the calculation of free metal concentrations: from a theoretical perspective it is expected that models with the free metal activity have a higher R^2 than the models with the total metal concentration in solution. From the regression results, it appears that for Cu, Pb and Zn the opposite is the case. The use of total solution concentration isotherms for input in SOACAS could be considered but a consequence would be that the mobile concentration is independent of ligand concentrations. All complexation effects are accounted for by pH-correction;
- robustness of the statistical analysis:
- small number of reliable observations of the data sets for agriculture and nature conservation. The Hoop-Janssen data set and the LMB data set consist of a small number of observations;
- large negative correlation between the metal content in solution and the proton activity (H^+) in the sorption isotherms. In a robust sorption isotherm the independent parameters should not have a high correlation. A high correlation of coefficients indicates that many different solutions of these coefficients are possible;
- non-plausible coefficients: the inert phase fraction should yield positive coefficients (coefficients a, b and c). The proton activity should yield a coefficient between –1 and 0 (coefficients m) and the metal content in solution should yield a coefficient between 0 and 1 (coefficient n). The NLLSR isotherm for Cu with the total concentration as independent parameter derived from the Hoop-Janssen data set however yields a negative regression coefficient for the inert fraction. This may lead to errouneous prediction for low values of M. Further yields the NLLSR model derived from the LMB data set a coefficient m below –1 for Pb and a coefficient m which is not significant for Cd. For Cu, Pb and Zn a coefficient n is yielded above 1 and a coefficient n which is not significant Cd. These results would mean that the solid phase content of Cd is independent of the amount of Cd in the soil solution and the pH of the soil solution. These results are not plausible from a theoretical perspective;
- comparison of isotherms fitted for agriculture and nature conservation: the difference between the isotherms could be caused by the fact that (a) the LMB data set is very noisy and unreliable, (b) difference in pH and metal concentration range. The fact that the solid phase contents of the Hoop-Janssen data set were determined by means of a HNO_3 destruction and the solid phase contents of the LMB data set by means of a Aqua Regia digestion can cause a difference in the isotherms as well. Moreover, in soils in use for nature conservation a natural long term equilibrium condition will develop as a result of deposition. In soils in use for agriculture active soil management has taken place with consequences for metal sorption.

4.2 Recommendations

Recommendations for future research are:
- extension of field partition data set and quality improvement: in May 1999 a project will be started, in cooperation with DLO and RIZA on behalf of PGBO (Project Integrated Soil Research), in which new sorption models will be derived from DLO data sets of soils in use for agriculture and nature conservation in the Netherlands. The soil characteristics of these soils are representative for the situation in the Netherlands. The combined Hoop-Janssen data set and the data set of the National Soil Monitoring Network (LMB) consist of a relatively small number of samples. It is plausible that this has affected the results. Further is the LMB data set less reliable than the Hoop-Janssen data set because of the

large amount of values near the detection limits and the small concentration range. Therefore in the future sorption isotherms must be derived from sufficiently large, reliable data sets which include a balanced composition of soils in use for agriculture and soils in use for nature conservation;

- exploration of possibilities to obtain better starting values for model parameters. For example fixing coefficients n, m and t using the results of batch experiments. Another example is fixing coefficients a, b and c using the results of sequential extractions (Bril in prep, 1999);
- further research on sorption models with free metal activity as independent parameter, which distinguish between reactive and non-reactive solid phase, has to be performed. From a theoretical perspective it is expected that a non-reactive fraction of the solid phase is present in soils in the Netherlands, especially for Cu, Pb and Zn. In this research this could not be proven statistically for models with free metal activity as independent parameter. To fit the 2-phase models, the Nonlinear Least Squares Regression technique on untransformed values is recommended. A consequence of the use of untransformed values is that the range of high concentrations will be predicted best instead of the range of low concentrations with the use of log-transformed values;
- derive sorption models for other priority metals for which field partition data are reported, particularly Ni, Cr, As and Hg. This will be difficult because the metals Cr and As occur in anionic form. Other problems will be the complex sorption and speciation behaviour of Hg and the small amount of available information;
- on average metal concentrations in the Hoop-Janssen data set are higher than observed in diffusively contaminated soils in the Netherlands (the validation set for SOACAS). For example, 8 of the 33 observations of the Hoop-Janssen data set have a concentration of Pb (determined with HNO_3 digestion) higher than 250 mg/kg. The highest concentration of Pb of this data set is 848 mg/kg (Sample E941102E). From the validation map for SOACAS, it appears that only in a small part of the Netherland concentrations higher than 80 mg/kg are found. This must be kept in mind with the application of the fitted sorption isotherms.

References

Andreasen, A.H.M. (1928). Kolloidchem. Beihefte, 27, 349.

Bril, J. (1999). Personal communication February 1999 regarding fitted sorption isotherms for Cd, Cu, Pb and Zn based on batch data.

Bril, J. (1999). Development of the soil extraction technique(Chapter 4) (in preparation).

Elzinga, E.J., Van den Berg, B., Van Grinsven, J.J.M. and Swartjes, F.A. (1997). Freundlich-adsorptievergelijkingen voor Cd, Cu en Zn in de bodem op basis van literatuurgegevens. RIVM report no. 711501001, Bilthoven, the Netherlands.

Elzinga, E.J., Van Grinsven, J.J.M. and Swartjes, F.A. (1999). General purpose Freundlich isotherms for Cd, Cu and Zn in soils. European Journal of Soil Science, March 1999, 50 , 139-149.

Gillman, G.P. (1979). Aust. J. Soil Res., 17, 129.

De Groot, A.C. (1999). Personal communication about detection limits of total fraction of metals comparable with those of the Janssen data set (1996).

Groot, M.S.M. (1998). Personal communication about detection limits of total and bioavailable fraction of metals of the research project to assess the quality of agricultural soil in the Netherlands (1996).

Henriksen, A. and Seip, H.M. (1980) Strong and weak acids in surface waters in southern Scotland. Water Res. (14):809.

Van den Hoop, M.A.G.T. (1995). Metal speciation in Dutch soils: Field-based partition coefficients for heavy metals at background levels. RIVM report no. 719101013, Bilthoven, the Netherlands.

Van den Hoop, M.A.G.T. (1999). Personal communication about detection limits of total and bioavailable fraction of metals of the Hoop data set (1995).

Houba, V.J.G., Van der Lee, J.J., Novozamsky, I. and Wallinga, I. (1989). Soil and plant analysis, a serie of syllabi. Part 5, Soil analysis procedures. Department of Soil Science and Plant Nutrition, Wageningen Agricultural University, the Netherlands.

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

Koops, R., Van Grinsven, J.J.M., Crommentuijn, T., Van den Hoop, M.A.G.T., Swartjes, F.A., Kramer, P.R.G. Kramer and Peijnenburg, W.J.G.M. (1998). Evaluatie van door het RIVM gehanteerde partitiecoëfficiënten voor metalen. RIVM report no. 711401005, Bilthoven, the Netherlands.

Lagas, P. and Groot, M.S.M.(1996). Bodemkwaliteitskartering van de Nederlandse landbouwgronden. RIVM report no. 714801003, Bilthoven, the Netherlands.

Mantoura, R.F.C., Dickson, A. & Riley, J. P. (1978). The complexation of metals with humic materials in natural waters. Estuarine and Coastal Marine Science, 10, 1-56

Minitab (1996). MINITAB Reference Manual. Release 11 for Windows.

Mott, C.J.B. (1988). The inorganic components of the soil. p. 264-269. In: A. Wilde (ed.) Russell's Soil Conditions and Plant Growth. 11th edition.

Reinds, G.J., Bril, J., De Vries, W., Breeuwsma, A. and Groenenberg, J.E. (1995) Critical and present loads of Cd, Cu and Pb for European forest soils. SC-DLO report no. 96, Wageningen, the Netherlands.

Schecher, W.D. and McAvoy, D.C. (1991). MINEQL+: A Chemical Equilibrium Program for Personal Computers. User's Manual. Version 2.1

SPlus 4 (1998). Guide to Statistics. Data Analysis Products Division, Mathsoft Inc., Seattle, Washington.

Tiktak, A. (1999). Modeling Non-Point Source Pollutants in Soils. Applications to the Leaching and Accumulation of Pesticides and Cadmium.

Appendix 1 Background information on field observations

Van den Hoop data set

The objectives of the study of Van den Hoop were:
- to quantify heavy metal concentrations in solid and soil solution in soils under field conditions in order to obtain field partition coefficients (Kp);
- to compare the partition coefficients obtained with values proposed for the harmonization of quality criteria for heavy metals in the Netherlands (Van den Hoop, 1995).

Soil samples were collected from 14 different sites in use for nature conservation in the Netherlands between April 23 and July 20, 1993. The sites were selected on the next set of criteria (Table 2):
- the heavy metal content should be close to "background" levels;
- the soil samples should vary with respect to organic carbon content, clay content, pH and carbonate content (Van den Hoop, 1995).

After removing the litter layer, about 8 litres of soil were collected from the toplayer (0-20 cm). Pore water was obtained by centrifugating of the soils at 7500 rpm. The pH of the soil solution was measured immediately after collection. After centrifugation, the soils were dried for one week at room temperature. Then, the dried soil was ground with a mill of either agate or porcelain and homogenized by hand (Van den Hoop, 1995).

The organic carbon content (OC) is calculated as the difference between total carbon content and calcium carbonate content. For soils with pH (H_2O) > 6.5, the carbonate content was determined by adding HCl and measuring the amount of gas developed. Total carbon was determined by dry combustion at approximately 1030 °C with thermal conduction detection of CO_2. The clay content was determined by a pipet method (Andreasen, 1928) after dissolving cementing substances in HCl and oxidizing organic matter by addition of hydrogen peroxide. The cation exchange capacity (CEC) was measured according to Gillman (1979) at soil pH and ionic strength similar to that encountered in the field (Van den Hoop, 1995).

Total metal concentrations of the solid phase were obtained by means of GFAAS or ICP-AES after digesting the soil samples with concentrated nitric acid.

In the resulting solution and the pore water, metal concentrations were obtained by the following methods:
- Cd, Ni, Cu and Pb by GFAAS (Standard Operating Procedures: LAC/M025; LAC/M030; LAC/M035; LAC/M032);
- Zn, Ca, Mg, Na, K, Fe and Mn by radial ICP-AES (Standard Operating Procedures: LAC/M259);
- Nitrate, chloride and sulphate by ion chromotography (Standard Operating Procedures: LAC/M210);
- Fosfate by continious flow analysis (Standard Operating Procedures: LAC/M064);
- Dissolved Organic Carbon (DOC) and Non Purgeable Organic Carbon (NPOC) by
- Standard Operating Procedures: LAC/M049 and ECO/034 respectively (Van den Hoop, 1995).

To define the fraction of (bio-)available or mobile metals in the soil, the soil samples were extracted by adding 100 ml 0.01*M* $CaCl_2$ solution to 10 g of soil. The suspensions were, as for the destruction with HNO_3, shaken for 24 hours at 21±1°C and then centrifuged for 20 minutes at 13000 rpm. The pH of the supernatant was set to 2 by adding an appropriate aliquot of concentrated nitric acid (Van den Hoop, 1995).

Janssen data set

The objectives of the study of Janssen were:
- Determination of field-based partition coefficients (As, Cd, Cr, Cu, Pb and Zn) for 20 polluted soils sampled in the Netherlands;
- Investigation of the dependence of measured Kp values on soil and pore water characteristics (Janssen et al., 1996).

Soil samples were collected from 20 different sites in use for nature conservation in the Netherlands in October and November 1994. Sampling was carried out under dry weather conditions. The sites were selected on the next set of criteria (Table 2):
- elevated metal concentrations had to be expected due to the location of sites close to sources of pollutants of interest;
- little or no impact from agricultural practices;
- soil characteristics had to vary among sites (Janssen et al., 1996).

For each site the upper or gras layer was removed and a total of 20 L of soil from the toplayer (0-20 cm) was collected. Pore water was obtained by centrifugation of the soils at 7500 rpm and 15°C. To separate solid phase and pore water from soils, the pore water was filtered during centrifugation with a 2.5 μm pore size filter. Next the pH of the pore water samples was measured. After centrifugation the soil samples were air-dried and organic remainders like roots were removed. Agglomerates were broken by hand, and gently ground in an agate mortar (Janssen et al., 1996).

The loss-on-ignition (LOI) was determined from the weight loss of approximately 5 g of air-dry soil heated at 550 °C for 2 hours. The carbon content of the soil was calculated from LOI by correction for Fe_2O_3 and <2μm particles (Houba et al., 1989). The organic matter content was calculated from the carbon content by multiplying with a value of 1.7. CEC of the soil was determined in an unbuffered $BaCl_2$ extract. The clay content (parts < 2μm) was obtained by measuring the different sedimentation rate of parts (Houba et al., 1989; Janssen et al., 1996).

The pore water metal concentrations were obtained by the following methods:
- Cd, Cr, Ni and Pb by graphite furnace AAS;
- Cu and Zn by AAS;
- As by FIAAS;
- Ca, Mg, Na, K, Zn, Fe, Mn and Al by ICP-AES;
- Cl^-, NO_3^-, HCO_3^- and CO_3^{2-} by ion chromotography;
- Fosfate by continious flow analysis;
- Dissolved Organic Carbon (DOC) by Dohrmann DC-190 (Janssen et al., 1996)

The detection limits of the pore water analysis, the HNO_3 destructions and the $CaCl_2$ extractions were not available but according to Groot (1999) the detection limits are comparable with those used at the Laboratory of Anorganic Chemistry (LAC) of the RIVM because similar methods are used. Therefor these detection limits are reported.

To obtain the total amount of metal, the soil samples were destructed with 15*M* HNO_3. Therefore 0.500 g of air-dry soil was weighed into a microwave digestion bomb and 2 ml deionized water and 5 ml concentrated nitric acid were added to each bomb. The soil samples were digested in a microwave oven for one hour at 630 W. After the samples were cooled, they were transferred into a volumetric flask and diluted to a final volume of 100 ml with deionized water. The digest was allowed to settle and the next day the clear solution was decanted into a plastic bottle (Janssen et al., 1996).

To define the fraction of (bio-)available or mobile metals in the soil, the soil samples were extracted by adding 100 ml 0.01 M $CaCl_2$ solution to 10 g of air-dry soil. The suspensions were shaken for 24 hours at 150 rpm. The pH in the extracts was measured (pH($CaCl_2$)) and the pH of the supernatant was set to 2 by adding an appropriate aliquot of concentrated nitric acid (Janssen et al., 1996).

The National Soil Monitoring Network (LMB) data set

A research project to assess the quality of agricultural soil in the Netherlands has been carried out by the National Institute of Health and Environmental Protection in cooperation with the Laboratory of Soil and Crop research.

A total of 42 different combinations of soil type (sand, fluvial and marine clay, peat and loam) and soil use (grassland, arable land, maize, flower, bulbs and orchards) were selected from 13 agricultural regions in the Netherlands (Lagas and Groot, 1996). For each combination four representative mixed samples were made to determine average heavy metal and organic compounds contents. The representative samples, composed of 20 parcel samples, were among other parameters analysed for total and bioavailable heavy metals. For grasslands and orchards, samples were taken from the upper 5 cm and for arable land from the upper 25 cm. The pretreatment of the soil samples has taken place according to NEN 5751 (Lagas and Groot, 1996).

The soil samples were, among other soil characteristics, examined for pH, organic matter content and clay content. The pH-H_2O was determined in a suspension of 1 volume part soil sample and 5 volume parts demineralized water (H_2O) after a contact period of 16 hours (NEN 5750). The clay content (parts < 2μm) was obtained by measuring the different sedimentation rate of the parts (sedimentation equation of Stokes) (NEN 5753) (Lagas and Groot, 1996).

To obtain the total amount of metal, all soil samples were destructed with Aqua Reagia (mixture of nitric acid and hydrochloric acid) according to NEN 6465 (Table 2). The metals Cd (NEN 5762), Cu (NEN 5758), Pb (NEN 5761) and Zn (NEN 5759) were determined with the flame-AAS with the principle of atomization of elements in an oxidizing air-acetylene flame. For Cd applies that at amounts smaller than 2.5 ppm, use was made of a graphite furnace (NEN 6458) (Lagas and Groot, 1996).

For each combination of soil type and soil use, one soil sample has been extracted with 0.01*M* $CaCl_2$ solution in order to define the fraction of (bio-)available or mobile metals in the soil. Extractions were performed by adding 100 ml 0.01*M* $CaCl_2$ solution to 10 g of soil. The suspensions were shaken for 2 hours and then centrifuged and filtrated. The fraction of (bio-)available amount of metals were obtained by means of the following methods:

- Cd and Pb by ICP-MS;

- Cu by GF-AAS;
- Zn by AAS.

The (bio-)available amounts of metals were not corrected for errors in weighing (0.5%), the recovery (at least 80%) and blank values. Especially the error in blank values will have a fairly large influence on the final results of most heavy metal determinations (Lagas and Groot, 1996).

Appendix 2 Data sets of Hoop, Janssen and of the National Soil Monitoring Network (LMB)

Table 1: *Combined data set of Hoop and Janssen: Soil characteristics (Van den Hoop, 1995; Janssen et al, 1996)*

Nr.	Sample	pH $CaCl_2$	OC[1]	Clay	CEC[2]	CEC[3]
			(g/100 g)	(g/100 g)	(meq/kg)	(meq/kg)
1	E941018A	3.42	2.1	2.86	35	116
2	E941027B	5.67	12.8	24.53	388	504
3	E941027C	5.35	6.6	33.83	295	382
4	E941102D	7.17	1.9	6.17	109	121
5	E941102E	6.53	3.0	7.41	98	165
6	E941102F	6.51	5.1	11.66	180	249
7	E941102G	6.8	5.8	6.8	205	251
8	E941103H	3.54	1.7	0.79	26	94
9	E941103I	3.02	2.4	3.4	40	129
10	E941103J	4.98	3.7	2.65	83	172
11	E941104K	3.77	2.0	2.68	29	112
12	E941104L	6.83	5.4	3.37	200	225
13	E941122M	3.82	1.5	4.31	43	97
14	E941122N	3.85	1.2	3.07	17	80
15	E941125O	5.26	2.9	7.04	99	161
16	E941128P	6.75	1.8	4.27	98	112
17	E941128Q	6.69	7.2	26.86	309	372
18	E941128R	6.89	4.7	19	251	267
19	E941129S	7.02	3.5	13.6	172	207
20	E941129T	6.95	3.7	10.5	165	200
21	E930423A	4.97	4.1	64.9	432	432
22	E930423B	4.36	43.1	19.4	461	1087
23	E930611C	7.39	5.2	23.6	353	301
24	E930624D	5.19	23	35.8	515	773
25	E930624E	3.23	1.5	2.1	53	90
26	E930629F	4.32	0.3	1.4	46	30
27	E930707G	5.74	2.6	62.6	431	373
28	E930707I	3.54	1.6	2	46	93
29	E930715J	7.21	2.7	21	283	210
30	E930715K	5.81	3.2	41.6	390	310
31	E930715L	3.89	4.5	37.3	238	335
32	E930720M	4.88	2.3	5.9	86	136
33	E930720N	5.34	2.6	19	203	199

[1]Organic carbon content of the Janssen data set was calculated from the estimated organic matter content by dividing by a value of 1.7 (soil samples starting with E94)

[2] Unbuffered CEC

[3] Calculated buffered CEC

Table 2: *Combined data set of Hoop and Janssen: Metal contents of solid phase determined with a concentrated HNO_3 digestion*

Nr.	Sample	Cd	Cu	Pb	Zn
		mmol/kg	mmol/kg	mmol/kg	mmol/kg
1	E941018A	0.000	0.035	0.585	0.243
2	E941027B	0.003	0.367	0.702	1.208
3	E941027C	0.003	1.683	1.542	2.919
4	E941102D	0.023	0.193	0.671	13.438
5	E941102E	0.065	0.317	4.093	47.545
6	E941102F	0.145	0.722	0.826	20.002
7	E941102G	0.095	1.653	1.281	18.055
8	E941103H	0.013	0.017	0.335	0.367
9	E941103I	0.009	0.023	0.346	0.222
10	E941103J	0.438	1.136	1.052	17.986
11	E941104K	0.001	0.027	1.730	0.081
12	E941104L	0.008	0.142	1.059	2.071
13	E941122M	0.000	0.108	0.357	0.283
14	E941122N	0.000	0.044	0.472	0.442
15	E941125O	0.125	0.654	0.720	2.740
16	E941128P	0.054	0.671	1.353	9.246
17	E941128Q	0.008	0.690	0.918	1.150
18	E941128R	0.045	1.727	1.357	7.156
19	E941129S	0.065	0.877	0.877	9.467
20	E941129T	0.145	0.612	1.978	6.567
21	E930423A	0.003	0.488	0.186	1.996
22	E930423B	0.003	0.157	0.109	0.636
23	E930611C	0.001	0.239	0.136	1.367
24	E930624D	0.006	0.200	0.081	1.147
25	E930624E	0.000	0.011	0.046	0.168
26	E930629F	0.000	0.011	0.029	0.083
27	E930707G	0.002	0.548	0.213	2.215
28	E930707I	0.000	0.046	0.043	0.156
29	E930715J	0.003	0.130	0.128	0.899
30	E930715K	0.004	0.374	0.155	1.860
31	E930715L	0.005	0.351	2.140	2.085
32	E930720M	0.001	0.103	0.080	0.438
33	E930720N	0.004	0.317	0.199	1.745

Table 3: *Combined data set of Hoop and Janssen: Metal contents of solid phase determined with a 0.01M $CaCl_2$-extraction (Van den Hoop, 1995; Janssen et al, 1996)*

Nr.	Sample	Cd	Cu	Pb	Zn
		µmol/kg	µmol/kg	µmol/kg	µmol/kg
1	E941018A	0.102	4.693	28.317	14.370
2	E941027B	0.084	-7.898	-1.043	4.202
3	E941027C	0.303	2.910	-0.260	33.240
4	E941102D	0.040	2.480	2.932	14.548
5	E941102E	3.925	1.271	7.884	762.894
6	E941102F	2.813	3.634	-0.295	532.652
7	E941102G	0.866	8.590	-0.351	7.482
8	E941103H	6.719	1.762	38.842	299.197
9	E941103I	3.691	2.121	30.783	107.863
10	E941103J	85.122	n.d.	n.d.	3245.771
11	E941104K	0.300	0.772	112.518	24.600
12	E941104L	0.028	0.280	0.143	3.129
13	E941122M	0.254	3.255	7.637	55.370
14	E941122N	0.183	0.700	26.493	169.360
15	E941125O	9.282	2.262	-0.630	142.812
16	E941128P	0.321	2.882	-0.367	12.956
17	E941128Q	0.042	1.178	-0.211	4.758
18	E941128R	0.298	5.143	-0.696	31.850
19	E941129S	0.491	4.680	-0.591	8.190
20	E941129T	0.727	2.848	-0.640	9.164
21	E930423A	0.167	0.048	-0.002	4.064
22	E930423B	0.050	-4.671	-0.069	17.488
23	E930611C	0.000	0.293	-0.001	-0.092
24	E930624D	0.049	-0.512	0.076	5.466
25	E930624E	0.178	-0.356	2.296	60.111
26	E930629F	0.134	-0.017	1.320	41.617
27	E930707G	0.130	0.242	-0.004	0.960
28	E930707I	0.211	-0.070	1.484	101.561
29	E930715J	0.000	0.283	0.027	-0.092
30	E930715K	0.123	0.483	-0.001	2.438
31	E930715L	1.371	1.484	16.654	81.417
32	E930720M	0.214	0.338	0.050	30.344
33	E930720N	0.079	0.287	0.000	39.254

Table 4a: *Combined data set of Hoop and Janssen: Pore water composition (Van den Hoop, 1995; Janssen et al, 1996)*

Nr.	Sample	pH	DOC	Cl	NO_3	PO_4	SO_4	K	Na
			mg/l	mmol/l	mmol/l	µmol/l	mmol/l	mmol/l	mmol/l
1	E941018A	3.57	51.8	0.2873	0.8792	12.43	0.0815	0.148	0.168
2	E941027B	5.74	75.1	0.9040	1.2618	1.80	1.2383	0.37	0.721
3	E941027C	5.15	88.5	0.8688	0.0010	45.26	0.4361	0.201	0.41
4	E941102D	7.43	113.8	0.7733	0.0295	3.10	14.4545	0.871	3.592
5	E941102E	7.2	67.6	0.0685	0.0717	14.12	0.2876	0.077	0.229
6	E941102F	6.79	88.2	0.1561	0.3889	2.97	0.3891	0.039	0.241
7	E941102G	7.57	75.9	1.1181	0.0402	3.39	4.3365	0.154	1.088
8	E941103H	3.78	87.7	0.1847	0.0226	3.12	0.1222	0.061	0.143
9	E941103I	3.4	146.2	0.2308	0.0676	7.30	0.2639	0.058	0.281
10	E941103J	5.97	58.2	0.5385	0.1551	13.82	1.444	0.157	0.995
11	E941104K	4.21	148.4	<0.0020	<0.0010	21.25	<0.0010	0.246	0.831
12	E941104L	7.45	148.1	3.1834	0.2136	13.01	1.012	1.336	3.628
13	E941122M	4.02	56	0.2078	0.9299	19.66	0.0853	0.218	0.249
14	E941122N	4.01	53.7	0.5673	0.9802	0.55	0.0982	0.958	0.381
15	E941125O	6.09	34.3	1.9727	0.2085	0.89	2.4277	0.595	1.794
16	E941128P	7.09	21.2	4.7110	0.0043	6.83	1.1498	0.099	4.614
17	E941128Q	7.14	54	0.7985	0.7031	15.23	1.1092	0.982	0.504
18	E941128R	7.41	20.4	3.8966	0.0100	0.80	1.8071	0.082	3.903
19	E941129S	7.91	26.4	1.6198	0.2118	4.78	0.9874	0.607	1.334
20	E941129T	7.5	34.7	2.2224	0.7482	7.81	1.1725	1.066	1.781
21	E930423A	4.75	15.16	1.1136	4.5750	0.26	0.3196	0.023	1.266
22	E930423B	4.36	36.01	0.7579	0.7507	0.76	11.1437	0.071	0.824
23	E930611C	7.72	28.8	0.6174	2.1405	0.19	1.6230	0.198	0.472
24	E930624D	5.82	47.17	1.6244	0.2045	0.43	1.1808	0.011	1.445
25	E930624E	3.2	n.d.	0.7607	1.1325	20.80	0.2546	0.465	0.689
26	E930629F	4.19	n.d.	0.2406	0.1327	0.65	0.1584	0.118	0.365
27	E930707G	4.56	65.42	1.5186	0.2334	1.20	0.7458	0.030	1.288
28	E930707I	3.19	n.d.	n.d.	n.d.	n.d.	n.d.	0.187	0.549
29	E930715J	7.85	80.88	3.8378	0.4608	11.68	1.4954	0.031	1.368
30	E930715K	6.5	39.95	0.3726	0.3332	0.15	0.6508	0.010	0.413
31	E930715L	4.14	53.35	0.5435	6.5677	8.47	0.4465	0.542	0.796
32	E930720M	5.39	53.35	2.3925	4.8332	0.24	2.6085	0.063	2.050
33	E930720N	5.35	53.29	0.9520	0.2084	124.17	0.3556	0.128	0.621

n.d.= not determined

Table 4b: *Combined data set of Hoop and Janssen: Pore water composition (Van den Hoop, 1995; Janssen et al, 1996)*

Nr.	Sample	Ca	Mg	Mn	Fe	Cd	Cu	Pb	Zn
		mmol/l	mmol/l	µmol/l	µmol/l	nmol/l	µmol/l	µmol/l	µmol/l
1	E941018A	0.1386	0.058	10.81	7.3	14.591	1	0.1707	3.3
2	E941027B	1.7905	0.292	0.8	58.3	3.025	4	0.2662	2
3	E941027C	0.6898	0.28	1.13	43.2	4.270	2.4	0.0444	1.3
4	E941102D	13.1057	2.564	8	1.8	4.004	0.6	0.0159	4.8
5	E941102E	0.8536	0.158	0.86	43.3	47.331	0.6	0.7707	19.5
6	E941102F	0.9179	0.162	0.58	19	28.737	1.6	0.0660	15.4
7	E941102G	6.9247	0.704	0.22	3.7	26.779	1.6	0.0212	2.9
8	E941103H	0.036	0.012	0.18	10.2	93.060	0.7	0.2529	16.1
9	E941103I	0.0593	0.024	0.5	22.9	58.630	0.6	0.2981	5.4
10	E941103J	0.9799	0.208	0.65	54.8	822.687	1.3	0.1776	116.5
11	E941104K	0.1871	0.203	3.46	22.5	10.587	0.4	1.8896	2.4
12	E941104L	2.9536	1.078	0.39	4.4	13.612	1.3	0.0461	1.6
13	E941122M	0.2932	0.111	12.81	24.7	17.527	0.8	0.1004	7.4
14	E941122N	0.2453	0.159	15.58	91.8	n.d.	0.5	n.d.	76.9
15	E941125O	1.9498	0.598	5.15	8.7	128.292	0.5	0.0195	6.5
16	E941128P	3.287	0.496	12.68	10.6	11.655	0.8	0.0362	2.5
17	E941128Q	1.522	0.281	0.38	28.4	8.630	0.7	0.0331	0.9
18	E941128R	3.5251	0.583	3.18	7.9	7.117	0.4	0.0298	1.4
19	E941129S	2.9051	0.41	0.42	16.8	10.053	0.7	0.0517	1.4
20	E941129T	2.8051	0.561	0.6	10.5	7.918	0.7	0.0292	1.2
21	E930423A	2.1238	0.571	0.66	11.6	3.737	1.7	0.0060	1.3
22	E930423B	9.6861	1.853	47.16	4.7	5.338	2.8	0.0350	2.3
23	E930611C	3.5694	0.310	0.01	3.0	1.868	0.3	0.0080	0.5
24	E930624D	1.5112	0.237	1.56	109.9	1.957	0.9	0.0360	0.7
25	E930624E	0.1919	0.122	4.02	21.8	18.861	4.4	0.0310	4
26	E930629F	0.0921	0.040	2.17	12.5	9.786	0.7	0.0700	1.1
27	E930707G	1.0177	0.273	1.14	4.5	2.847	0.8	0.0180	0.6
28	E930707I	0.2400	0.175	16.63	18.8	52.313	1.4	0.0170	7.1
29	E930715J	4.1557	0.702	0.12	5.4	3.470	2.3	0.0080	0.7
30	E930715K	0.8890	0.159	0.50	3.8	3.381	1.0	0.0060	1
31	E930715L	2.6712	0.808	86.66	9.5	61.833	2.1	1.2651	5.1
32	E930720M	4.2380	1.308	7.59	12.4	19.128	2.0	0.0160	3.6
33	E930720N	0.8281	0.192	0.37	6.3	3.025	1.0	0.0090	1.1

n.d.= not determined

Table 5: *Data set of the National Soil Monitoring Network (LMB): Soil characteristics (Lagas et al., 1996)*

Nr.	Sample	pH H_2O	pH pore water[1]	OM (g/ 100 g)	OC[2] (g/ 100 g)	clay (g/ 100 g)
1	DD 2	6.2	5.9	19.9	11.71	34.1
2	JD 2	6.0	5.7	5.0	2.94	15.6
3	KD 2	6.4	6.1	9.9	5.82	18.1
4	LA 2	6.2	5.9	10.3	6.06	31.6
5	MD 2	6.3	6.0	4.4	2.59	13.0
6	BA 2	5.7	5.3	31.6	18.59	21.9
7	CD 2	5.8	5.4	34.9	20.53	19.1
8	DA 2	5.8	5.4	39.7	23.35	21.7
9	FA 2	5.7	5.3	26.3	15.47	13.5
10	HA 2	5.6	5.2	25.7	15.12	2.4
11	JA 2	6.0	5.7	26.4	15.53	8.5
12	AD 2	6.1	5.8	5.3	3.12	8.1
13	BB 2	5.9	5.5	9.7	5.71	4.5
14	CC 2	6.3	6.0	6.2	3.65	3.4
15	DB 2	5.9	5.5	6.7	3.94	3.3
16	EC 2	6.9	6.7	4.3	2.53	4.1
17	FB 2	5.9	5.5	7.1	4.18	4.4
18	GA 2	6.0	5.7	7.5	4.41	3.6
19	HB 2	5.9	5.5	7.9	4.65	3.5
20	JB 2	6.1	5.8	5.9	3.47	5.4
21	KB 2	6.1	5.8	5.2	3.06	3.9
22	LB 2	6.3	6.0	4.8	2.82	7.0
23	MA 2	6.3	6.0	3.9	2.29	3.0
24	AA 2	6.8	6.5	10.4	6.12	27.6
25	BC 2	6.1	5.8	16.8	9.88	30.8
26	CB 2	7.2	7.0	8.5	5.00	20.6
27	DC 2	6.1	5.8	13.2	7.76	24.9
28	EB 2	6.8	6.5	13.4	7.88	21.5
29	GB 2	5.8	5.4	5.3	3.12	3.0
30	HC 2	5.9	5.5	6.1	3.59	3.2
31	MC 2	6.2	5.9	2.9	1.71	2.8
32	AB 2	7.4	7.2	3.6	2.12	20.2
33	CA 2	7.8	7.7	3.0	1.76	17.3
34	DZ 2	7.9	7.8	4.7	2.76	24.4
35	EA 2	8.0	7.9	2.1	1.24	22.0
36	HD 2	5.9	5.5	19.5	11.47	3.9
37	NA 2	6.7	6.4	6.7	3.94	15.4
38	JC 2	6.1	5.8	3.9	2.29	3.3
39	KC 2	5.8	5.4	3.7	2.18	4.3
40	MB 2	6.3	6.0	3.0	1.76	3.1
41	DF 2	7.2	7.0	1.3	0.76	2.7
42	LC 2	7.1	6.9	3.0	1.76	23.9

[1]pH pore water was calculated from pH H_2O by the regression equation (derived from the combined data set of Hoop and Janssen): $pH(porewater) = -1.040 + 1.116 \times pH(H_2O)$

[2]Organic carbon content was calculated from the organic matter content by dividing by a value of 1.7

Table 6: *Data set of the National Soil Monitoring Network (LMB): Metal contents in solid phase determined with Aqua Regia (Lagas et al.,1996)*

Nr.	Sample	Cd	Cu	Pb	Zn
		(mg/kg)	(mg/kg)	(mg/kg)	(mg/kg)
1	DD 2	0.65	39.8	75	212
2	JD 2	0.23	12.2	24	95
3	KD 2	0.32	22.5	26	127
4	LA 2	0.66	29.3	58	171
5	MD 2	0.29	12.6	30	108
6	BA 2	0.31	21.3	44	93
7	CD 2	0.77	34.1	88	176
8	DA 2	0.68	68.7	212	285
9	FA 2	0.44	12.7	28	78
10	HA 2	0.35	16.1	25	67
11	JA 2	0.49	18.1	38	96
12	AD 2	0.19	8.7	21	33
13	BB 2	0.24	7.2	13	48
14	CC 2	0.21	7.5	29	44
15	DB 2	0.16	21.3	61	70
16	EC 2	0.12	5.9	14	37
17	FB 2	0.24	10.8	20	40
18	GA 2	0.24	11.9	19	38
19	HB 2	0.14	13.3	12	43
20	JB 2	0.16	12.7	17	43
21	KB 2	0.14	20.2	19	39
22	LB 2	0.21	12.6	19	50
23	MA 2	0.25	10.8	19	39
24	AA 2	0.21	18.5	27	76
25	BC 2	0.52	16.2	35	129
26	CB 2	0.19	12.2	22	82
27	DC 2	0.49	27.7	69	164
28	EB 2	0.34	21.1	36	137
29	GB 2	0.29	13.5	22	38
30	HC 2	0.12	9.7	13	27
31	MC 2	0.24	9.5	14	39
32	AB 2	0.09	9.8	21	54
33	CA 2	0.57	13.4	21	84
34	DZ 2	0.30	20.2	34	100
35	EA 2	0.29	24.9	29	81
36	HD 2	0.35	23.4	30	52
37	NA 2	0.67	15.7	34	178
38	JC 2	0.13	9.5	13	30
39	KC 2	0.14	10.5	15	33
40	MB 2	0.21	10.7	12	36
41	DF 2	0.20	7.1	6	36
42	LC 2	0.37	33.5	34	122

Table 7: *Data set of the National Soil Monitoring Network (LMB): Metal contents in solid phase determined with a 0.01M $CaCl_2$ extraction (Lagas et al.,1996)*

Nr.	Sample	Cd	Cu	Pb	Zn
		(µg/l)	(µg/l)	(µg/l)	(µg/l)
1	DD 2	0,99	22.24	0,77	133,4
2	JD 2	3,57	8.9	0,54	56,7
3	KD 2	1,07	15.89	<0,207	124,0
4	LA 2	3,36	20.33	0,62	27,1
5	MD 2	2,45	8.9	<0,207	213,4
6	BA 2	1,83	10.8	1,12	191,2
7	CD 2	1,19	12.07	1,53	117,2
8	DA 2	1,56	19.06	6,09	442,1
9	FA 2	1,67	5.08	1,35	121,8
10	HA 2	0,25	6.35	1,31	100,6
11	JA 2	2,19	6.99	1,18	217,2
12	AD 2	1,34	6.99	1,02	95,6
13	BB 2	1,49	4.45	2,26	27,6
14	CC 2	0,84	5.08	1,24	118,2
15	DB 2	2,00	12.07	4,50	345,4
16	EC 2	0,20	5.08	0,25	<10
17	FB 2	1,23	4.45	1,28	148,5
18	GA 2	1,75	6.99	3,05	191,4
19	HB 2	1,03	5.08	0,56	239,6
20	JB 2	2,46	12.71	5,41	122,8
21	KB 2	0,90	3.18	0,52	117,9
22	LB 2	1,80	9.53	1,91	92,6
23	MA 2	2,72	6.99	1,04	91,8
24	AA 2	0,26	13.98	0,41	<10
25	BC 2	0,89	9.53	0,50	98,9
26	CB 2	<0,045	9.53	<0,207	<10
27	DC 2	1,17	15.25	1,43	75,4
28	EB 2	0,43	17.79	<0,207	<10
29	GB 2	2,03	6.99	2,15	211,1
30	HC 2	1,07	3.18	0,54	141,1
31	MC 2	3,32	6.99	0,50	553,5
32	AB 2	<0,045	6.99	0,44	32,4
33	CA 2	0,08	6.35	<0,207	<10
34	DZ 2	0,06	8.26	<0,207	<10
35	EA 2	0,12	8.26	<0,207	237,8
36	HD 2	1,35	8.26	1,28	259,1
37	NA 2	1,74	13.98	<0,207	97,1
38	JC 2	2,61	8.9	0,64	51,9
39	KC 2	2,09	6.99	0,31	9,8
40	MB 2	2,34	6.99	0,27	147,0
41	DF 2	0,40	6.35	0,44	59,9
42	LC 2	0,37	14.61	<0,207	121,8

Appendix 3 Background information on regression techniques

Stepwise Regression

Stepwise Regression, using the maximum F-statistic, is used to identify a useful subset of the predictors. At each step the procedure calculates an F-statistic for each variable in the model. Suppose the model contains X1,....,Xp. Then the F-statistic for Xi is:

$$\frac{SSE[X1,...,X(i-1),X(i+1),...,Xp]-SSE[X1,...,Xp]}{MSE[X1,...,Xp]}$$

where: SSE = sum of square error;
MSE = mean square error.

and with 1 and n – p – 1 degrees of freedom. If the F-statistic for any variable is less than FREMOVE, the variable with the smallest F is removed from the model. The regression equation is calculated for this smaller model. If no variable can be removed, the procedure tries to add a variable. An F-statistic is calculated for each variable not yet in the model. Suppose the model contains X1,...,Xp. Then the F-statistic for a new variable, X (p + 1) is:

$$\frac{SSE[X1,...,Xp]-SSE[X1,...,Xp,X(p+1)]}{MSE[X1,...,Xp,X(p+1)]}$$

where: SSE = sum of square error;
MSE = mean square error.

The variable with the largest F-statistic is added, provided its F-statistic is larger than FENTER. If no variable can be added, the Stepwise Regression ends (Minitab, 1996).

Nonlinear Least Squares Regression

In Linear Regression the mean surface in sample space is a plane, in Nonlinear Regression it may be an arbitrary curved surface. When using the Nonlinear Least Squares Algorithm in SPlus, the Gauss-Newton algorithm is used with a step factor to ensure that the sum of squares differences between response and prediction decreases at each iteration. A line search method is used and the step direction is determined by a quadratic model. The algorithm proceeds as follows:

1. The residuals are calculated, and the gradient is calculated or approximated at the current parameter values;
2. A linear least-squares fit of the residual on the gradient gives the parameter increment;
3. If applying the full parameter increment increases the sum-of-squares rather than decreasing it, the length of the increment is succesively halved until the sum-of-squares is decreased;
4. The step factor is retained between iterations and started at min {2*previous step factor),1).

If the gradient is not specified analytically, it is calculated using finite differences with forward differencing (SPlus 4, 1998).

Appendix 4 Results of the calculation of free metal activity

Table 1: *Free Cd activity and distribution of the formed species in an aqeous solution*

Nr.	Sample	$[Cd^{2+}]$ (mol/l)	[CdDOC] (mol/l)	$[CdCl^{+}]$ (mol/l)	$[CdCO_3$ AQ] (mol/l)	$[CdNO_3^{+}]$ (mol/l)	$[Cd(SO_4)_2^{2-}]$ (mol/l)	$[CdSO_4$ AQ] (mol/l)	Act. coeff. 2-wrd. ion	(Cd^{2+}) (mol/l)
1	E941018A	1.40E-08	2.48E-12	3.13E-10	9.45E-18	3.53E-11	1.91E-13	2.10E-10	0.78298	1.10E-08
2	E941027B	2.45E-09	7.70E-11	1.42E-10	3.28E-14	7.29E-12	4.13E-12	3.36E-10	0.66524	1.63E-09
3	E941027C	3.76E-09	3.96E-11	2.33E-10	3.50E-15	9.86E-15	1.10E-12	2.39E-10	0.72122	2.71E-09
4	E941102D	1.22E-09	1.99E-09	4.30E-11	3.31E-11	6.05E-14	5.90E-11	6.38E-10	0.40694	4.96E-10
5	E941102E	2.38E-08	2.19E-08	1.19E-10	2.82E-10	4.58E-12	3.12E-12	1.04E-09	0.70781	1.68E-08
6	E941102F	1.87E-08	8.65E-09	2.09E-10	3.33E-11	1.92E-11	4.33E-12	1.07E-09	0.72872	1.36E-08
7	E941102G	8.50E-09	1.47E-08	5.02E-10	4.72E-10	6.65E-13	8.56E-11	2.35E-09	0.48925	4.16E-09
8	E941103H	8.94E-08	4.57E-11	1.35E-09	1.62E-16	6.07E-12	3.16E-12	2.26E-09	0.83959	7.51E-08
9	E941103I	5.49E-08	1.87E-11	9.89E-10	1.69E-17	1.07E-11	8.02E-12	2.70E-09	0.79728	4.38E-08
10	E941103J	6.49E-07	2.79E-08	2.34E-08	2.56E-11	2.48E-10	1.83E-09	1.20E-07	0.69152	4.49E-07
11	E941104K	1.06E-08	2.33E-11	1.63E-12	1.35E-16	3.01E-14	2.09E-17	1.90E-12	0.78273	8.30E-09
12	E941104L	3.29E-09	9.26E-09	6.04E-10	1.10E-10	1.50E-12	2.52E-12	2.74E-10	0.52498	1.73E-09
13	E941122M	1.69E-08	8.95E-12	2.67E-10	8.95E-17	4.41E-11	2.35E-13	2.50E-10	0.79654	1.35E-08
15	E941125O	9.18E-08	2.77E-09	1.09E-08	5.97E-12	4.26E-11	5.26E-10	2.19E-08	0.62206	5.71E-08
16	E941128P	7.49E-09	1.33E-09	2.03E-09	4.76E-11	6.83E-14	7.57E-12	7.16E-10	0.53989	4.04E-09
17	E941128Q	4.86E-09	2.80E-09	2.53E-10	4.14E-11	8.21E-12	7.02E-12	6.28E-10	0.65503	3.18E-09
18	E941128R	4.03E-09	1.41E-09	8.97E-10	1.11E-10	8.48E-14	9.81E-12	5.92E-10	0.54597	2.20E-09
19	E941129S	3.34E-09	5.02E-09	3.31E-10	9.56E-10	1.60E-12	2.86E-12	3.12E-10	0.57783	1.93E-09
20	E941129T	3.82E-09	2.94E-09	5.09E-10	1.64E-10	6.31E-12	4.48E-12	4.09E-10	0.57309	2.19E-09
21	E930423A	3.36E-09	2.11E-12	2.32E-10	4.63E-16	3.51E-11	3.29E-13	1.07E-10	0.66602	2.24E-09
22	E930423B	3.27E-09	1.49E-12	1.16E-10	6.49E-17	4.22E-12	1.64E-10	1.79E-09	0.46296	1.51E-09
23	E930611C	7.88E-10	8.22E-10	2.90E-11	9.28E-11	3.70E-12	1.69E-12	1.13E-10	0.63084	4.97E-10
24	E930624D	1.55E-09	3.71E-11	1.62E-10	3.00E-14	7.56E-13	2.51E-12	2.11E-10	0.68786	1.07E-09
27	E930707G	2.37E-09	4.52E-12	2.43E-10	1.42E-16	1.37E-12	1.77E-12	2.28E-10	0.71944	1.71E-09
29	E930715J	6.53E-10	2.43E-09	1.43E-10	1.37E-10	6.32E-13	1.01E-12	7.64E-11	0.60302	3.94E-10
30	E930715K	2.75E-09	2.92E-10	7.26E-11	1.28E-12	2.39E-12	1.77E-12	2.58E-10	0.75590	2.08E-09
31	E930715L	5.69E-08	2.95E-11	1.83E-09	4.61E-16	8.14E-10	9.17E-12	2.23E-09	0.63308	3.60E-08
32	E930720M	1.43E-08	1.20E-10	1.85E-09	3.50E-14	1.37E-10	6.11E-11	2.63E-09	0.56637	8.10E-09
33	E930720N	2.67E-09	2.70E-11	1.81E-10	6.25E-15	1.46E-12	5.16E-13	1.38E-10	0.74730	2.00E-09

Table 2: *Free Cu activity and distribution of the formed species in an aqeous solution*

Nr.	Sample	$[Cu^{2+}]$	[CuDOC]	$[CuSO_4$ AQ]	Act. coeff.	(Cu^{2+})
		(mol/l)	(mol/l)	(mol/l)	2-wrd. ion	(mol/l)
1	E941018A	4.15E-07	5.80E-07	4.36E-09	0.78298	3.25E-07
2	E941027B	1.63E-08	3.98E-06	1.57E-09	0.66524	1.08E-08
3	E941027C	2.85E-08	2.37E-06	1.28E-09	0.72122	2.06E-08
4	E941102D	4.65E-11	5.99E-07	1.70E-11	0.40694	1.89E-11
5	E941102E	8.21E-11	6.00E-07	2.51E-12	0.70781	5.81E-11
6	E941102F	4.37E-10	1.60E-06	1.75E-11	0.72872	3.18E-10
7	E941102G	1.17E-10	1.60E-06	2.27E-11	0.48925	5.72E-11
8	E941103H	1.38E-07	5.59E-07	2.45E-09	0.83959	1.16E-07
9	E941103I	1.61E-07	4.34E-07	5.55E-09	0.79728	1.28E-07
10	E941103J	3.82E-09	1.30E-06	4.96E-10	0.69152	2.64E-09
11	E941104K	2.16E-08	3.78E-07	2.72E-12	0.78273	1.69E-08
12	E941104L	5.83E-11	1.30E-06	3.41E-12	0.52498	3.06E-11
13	E941122M	1.54E-07	6.44E-07	1.60E-09	0.79654	1.23E-07
14	E941122N	1.04E-07	3.95E-07	1.15E-09	0.75768	7.88E-08
15	E941125O	2.09E-09	4.97E-07	3.49E-10	0.62206	1.30E-09
16	E941128P	5.75E-10	7.98E-07	3.86E-11	0.53989	3.10E-10
17	E941128Q	1.54E-10	6.99E-07	1.40E-11	0.65503	1.01E-10
18	E941128R	1.44E-10	3.98E-07	1.49E-11	0.54597	7.86E-11
19	E941129S	5.87E-11	6.93E-07	3.85E-12	0.57783	3.39E-11
20	E941129T	1.15E-10	6.98E-07	8.65E-12	0.57309	6.59E-11
21	E930423A	2.86E-07	1.38E-06	6.42E-09	0.66602	1.90E-07
22	E930423B	5.73E-07	2.04E-06	2.20E-07	0.46296	2.65E-07
23	E930611C	3.19E-11	2.63E-07	3.23E-12	0.63084	2.01E-11
24	E930624D	4.59E-09	8.70E-07	4.39E-10	0.68786	3.16E-09
27	E930707G	5.17E-08	7.80E-07	3.50E-09	0.71944	3.72E-08
29	E930715J	7.96E-11	2.33E-06	6.55E-12	0.60302	4.80E-11
30	E930715K	1.22E-09	1.02E-06	8.02E-11	0.75590	9.22E-10
31	E930715L	4.04E-07	1.64E-06	1.11E-08	0.63308	2.56E-07
32	E930720M	2.92E-08	1.93E-06	3.78E-09	0.56637	1.65E-08
33	E930720N	1.28E-08	1.02E-06	4.62E-10	0.74730	9.57E-09

Table 3: *Free Pb activity and distribution of the formed species in an aqeous solution*

Nr.	Sample	$[Pb^{2+}]$	$[PbCl^{+}]$	[PbDOC]	$[PbNO_3^{+}]$	$[PbSO_4$ AQ]	Act. coeff.	(Pb^{2+})
		(mol/l)	(mol/l)	(mol/l)	(mol/l)	(mol/l)	2-wrd. ion	(mol/l)
1	E941018A	9.76E-08	7.26E-10	6.86E-08	1.07E-09	3.03E-09	0.78298	7.64E-08
2	E941027B	2.11E-09	4.07E-11	2.63E-07	2.73E-11	6.00E-10	0.66524	1.40E-09
3	E941027C	1.03E-09	2.13E-11	4.32E-08	1.18E-14	1.36E-10	0.72122	7.43E-10
4	E941102D	2.46E-12	2.88E-14	1.59E-08	5.28E-16	2.66E-12	0.40694	1.00E-12
5	E941102E	2.11E-10	3.52E-13	7.71E-07	1.77E-13	1.91E-11	0.70781	1.49E-10
6	E941102F	3.58E-11	1.34E-13	6.60E-08	1.60E-13	4.24E-12	0.72872	2.61E-11
7	E941102G	3.07E-12	6.05E-14	2.12E-08	1.04E-15	1.76E-12	0.48925	1.50E-12
8	E941103H	8.19E-08	4.12E-10	1.66E-07	2.42E-11	4.30E-09	0.83959	6.88E-08
9	E941103I	1.21E-07	7.28E-10	1.64E-07	1.02E-10	1.24E-08	0.79728	9.65E-08
10	E941103J	1.03E-09	1.24E-11	1.77E-07	1.71E-12	3.96E-10	0.69152	7.12E-10
11	E941104K	1.94E-07	1.00E-11	1.70E-06	2.40E-12	7.23E-11	0.78273	1.52E-07
12	E941104L	4.12E-12	2.52E-13	4.61E-08	8.14E-15	7.11E-13	0.52498	2.16E-12
13	E941122M	3.17E-08	1.67E-10	6.68E-08	3.59E-10	9.73E-10	0.79654	2.53E-08
15	E941125O	1.60E-10	6.37E-12	1.92E-08	3.24E-13	7.93E-11	0.62206	9.95E-11
16	E941128P	5.12E-11	4.64E-12	3.61E-08	2.03E-15	1.02E-11	0.53989	2.76E-11
17	E941128Q	1.45E-11	2.51E-13	3.31E-08	1.06E-13	3.88E-12	0.65503	9.50E-12
18	E941128R	2.13E-11	1.58E-12	2.97E-08	1.95E-15	6.50E-12	0.54597	1.16E-11
19	E941129S	8.63E-12	2.86E-13	5.15E-08	1.79E-14	1.67E-12	0.57783	4.99E-12
20	E941129T	9.53E-12	4.24E-13	2.92E-08	6.85E-14	2.12E-12	0.57309	5.46E-12
21	E930423A	1.65E-09	3.80E-11	4.12E-09	7.50E-11	1.10E-10	0.66602	1.10E-09
22	E930423B	8.71E-09	1.03E-10	1.58E-08	4.90E-11	9.88E-09	0.46296	4.03E-09
23	E930611C	1.93E-12	2.36E-14	7.99E-09	3.94E-14	5.76E-13	0.63084	1.22E-12
24	E930624D	3.72E-10	1.30E-11	3.55E-08	7.91E-13	1.05E-10	0.68786	2.56E-10
27	E930707G	2.04E-09	6.99E-11	1.55E-08	5.15E-12	4.08E-10	0.71944	1.47E-09
29	E930715J	5.39E-13	3.93E-14	8.00E-09	2.27E-15	1.31E-13	0.60302	3.25E-13
30	E930715K	1.41E-11	1.25E-13	5.97E-09	5.34E-14	2.75E-12	0.75590	1.07E-11
31	E930715L	3.96E-07	4.26E-09	8.13E-07	2.47E-08	3.22E-08	0.63308	2.51E-07
32	E930720M	4.58E-10	1.97E-11	1.53E-08	1.92E-11	1.75E-10	0.56637	2.59E-10
33	E930720N	2.17E-10	4.91E-12	8.73E-09	5.15E-13	2.33E-11	0.74730	1.62E-10

Table 4: *Free Zn activity and distribution of the formed species in an aqeous solution*

Nr.	Sample	$[Zn^{2+}]$ (mol/l)	[ZnDOC] (mol/l)	$[ZnOH^{+}]$ (mol/l)	$[Zn(OH)_2$ AQ (mol/l)	$[ZnHCO_3^{+}$ (mol/l)	$[ZnCO_3$ (mol/l)	$[Zn(SO_4)_2^{2}$ (mol/l)	$[ZnSO_4$ AQ (mol/l)	Act. coeff. 2-wrd. ion	(Zn^{2+}) (mol/l)
1	E941018A	3.26E-06	5.76E-10	5.56E-12	5.21E-16	5.92E-12	1.75E-15	2.67E-11	3.89E-08	0.78298	2.55E-06
2	E941027B	1.75E-06	5.49E-08	4.00E-10	5.55E-12	4.26E-10	1.86E-11	1.77E-09	1.91E-07	0.66524	1.16E-06
3	E941027C	1.22E-06	1.29E-08	7.60E-11	2.71E-13	8.09E-11	9.08E-13	2.16E-10	6.22E-08	0.72122	8.80E-07
4	E941102D	1.55E-06	2.49E-06	1.46E-08	9.92E-09	1.56E-08	3.33E-08	4.49E-08	6.43E-07	0.40694	6.31E-07
5	E941102E	1.01E-05	8.83E-06	7.09E-08	2.83E-08	7.55E-08	9.50E-08	7.92E-10	3.50E-07	0.70781	7.15E-06
6	E941102F	1.02E-05	4.64E-06	2.78E-08	4.32E-09	2.96E-08	1.45E-08	1.42E-09	4.65E-07	0.72872	7.43E-06
7	E941102G	9.55E-07	1.64E-06	1.34E-08	1.26E-08	1.43E-08	4.22E-08	5.79E-09	2.11E-07	0.48925	4.67E-07
8	E941103H	1.58E-05	8.04E-09	4.46E-11	6.78E-15	4.75E-11	2.27E-14	3.33E-10	3.17E-07	0.83959	1.33E-05
9	E941103I	5.19E-06	1.76E-09	6.00E-12	3.80E-16	6.39E-12	1.27E-15	4.56E-10	2.04E-07	0.79728	4.14E-06
10	E941103J	9.83E-05	4.11E-06	3.90E-08	9.18E-10	4.15E-08	3.08E-09	1.62E-07	1.43E-05	0.69152	6.80E-05
11	E941104K	2.39E-06	5.28E-09	1.78E-11	7.27E-15	1.89E-11	2.44E-14	2.85E-15	3.43E-10	0.78273	1.87E-06
12	E941104L	4.07E-07	1.14E-06	4.53E-09	3.22E-09	4.83E-09	1.08E-08	1.88E-10	2.71E-08	0.52498	2.14E-07
13	E941122	7.31E-06	3.86E-09	3.47E-11	9.17E-15	3.70E-11	3.08E-14	6.09E-11	8.61E-08	0.79654	5.82E-06
14	E941122N	7.59E-05	3.62E-08	3.45E-10	8.92E-14	3.68E-10	2.99E-13	7.53E-10	9.38E-07	0.75768	5.75E-05
15	E941125O	5.30E-06	1.59E-07	2.63E-09	8.18E-11	2.81E-09	2.74E-10	1.83E-08	1.01E-06	0.62206	3.30E-06
16	E941128P	1.96E-06	3.45E-07	9.53E-09	2.96E-09	1.01E-08	9.92E-09	1.19E-09	1.50E-07	0.53989	1.06E-06
17	E941128Q	5.29E-07	3.04E-07	3.07E-09	1.07E-09	3.27E-09	3.59E-09	4.60E-10	5.47E-08	0.65503	3.47E-07
18	E941128R	9.19E-07	3.20E-07	9.30E-09	6.03E-09	9.90E-09	2.02E-08	1.35E-09	1.08E-07	0.54597	5.02E-07
19	E941129S	4.76E-07	7.07E-07	1.58E-08	3.23E-08	1.68E-08	1.08E-07	2.45E-10	3.55E-08	0.57783	2.75E-07
20	E941129T	6.23E-07	4.76E-07	7.95E-09	6.34E-09	8.46E-09	2.13E-08	4.39E-10	5.33E-08	0.57309	3.57E-07
21	E930423A	1.27E-06	7.93E-10	2.92E-11	4.14E-14	3.11E-11	1.39E-13	7.45E-11	3.23E-08	0.66602	8.46E-07
22	E930423B	1.57E-06	7.15E-10	1.28E-11	7.38E-15	1.36E-11	2.47E-14	4.75E-08	6.84E-07	0.46296	7.27E-07
23	E930611C	2.15E-07	2.24E-07	4.54E-09	6.01E-09	4.83E-09	2.02E-08	2.77E-10	2.48E-08	0.63084	1.36E-07
24	E930624D	6.16E-07	1.48E-08	1.71E-10	2.84E-12	1.82E-10	9.53E-12	6.03E-10	6.71E-08	0.68786	4.24E-07
27	E930707G	5.55E-07	1.06E-09	8.63E-12	7.90E-15	9.19E-12	2.65E-14	2.50E-10	4.27E-08	0.71944	3.99E-07
29	E930715J	1.37E-07	5.10E-07	3.81E-09	6.81E-09	4.06E-09	2.28E-08	1.28E-10	1.28E-08	0.60302	8.26E-08
30	E930715K	8.44E-07	8.95E-08	1.17E-09	9.33E-11	1.25E-09	3.13E-10	3.27E-10	6.32E-08	0.75590	6.38E-07
31	E930715L	4.94E-06	2.56E-09	2.73E-11	9.51E-15	2.91E-11	3.19E-14	4.79E-10	1.55E-07	0.63308	3.13E-06
32	E930720	3.10E-06	2.61E-08	2.91E-10	1.80E-12	3.10E-10	6.05E-12	7.98E-09	4.57E-07	0.56637	1.76E-06
33	E930720N	1.04E-06	1.06E-08	1.03E-10	5.80E-13	1.09E-10	1.95E-12	1.22E-10	4.31E-08	0.74730	7.77E-07

Appendix 5 Regression results

Cd

Table 1: *Regression results for LR Freundlich sorption isotherms of Cd derived from the log-transformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}\left(H^{+}\right)^{m}\left(M^{2+}\right)^{n}$$

Model[a]		p	q[e]	r	s	m	n	R^2	N[bc]
1	ß[d]	0.030	= 0			-0.448	0.808		
	St.error	0.470				0.070	0.139	0.674	28
	t ratio	0.063				-6.379	5.815		
2	ß	0.452	= 0			-0.338	0.839		
	St.error	0.524				0.065	0.148	0.665	28
	t ratio	0.863				-5.203	5.680		
3	ß	0.030		= 0	= 0	-0.448	0.808		
	St.error	0.470				0.070	0.139	0.674	28
	t ratio	0.063				-6.379	5.815		
4	ß	0.452		= 0	= 0	-0.338	0.839		
	St.error	0.524				0.065	0.148	0.665	28
	t ratio	0.863				-5.203	5.680		

[a] Model 1 and 3 based on free Cd activity
Model 2 and 4 based on total Cd concentration in solution
[b] N is number of datapoints
[c] Samples 1, 14, 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient
[e] = 0: regression coefficient is set to zero

Table 2: *Regression results for the NLLSR sorption isotherms of Cd, with free metal activity as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left(M^{2+}\right)^{n}$$

Model[a]		a	b	c	10^{p}	r	s	m	n	R^2	N[bc]
1	ß[de]	-2.039	0.007	-0.015	4.369	0.148	0.012	-0.230	0.435		
	St.error	6.385	0.119	0.069	9.020	0.416	0.252	0.212	0.355	0.624	27
	t ratio	-0.319	0.058	-0.212	0.484	0.355	0.048	-1.087	1.226		
2	ß				1.895			-0.341	0.581		
	St.error				2.396			0.107	0.138	0.593	27
	t ratio				0.791			-3.181	4.210		
3	ß							-0.382	0.585		
	St.error							0.060	0.137	0.588	27
	t ratio							-6.495	4.278		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level
[b] N is number of datapoints
[c] Samples 1, 10, 14, 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient
[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 3: *Regression results for the NLLSR sorption isotherms of Cd, with total metal concentration in solution as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}[M]^{n}$$

Model[a]		a	b	c	10^{p}	r	s	m	n	R^2	N[bc]
1	ß[de]	-2.111	0.025	-0.027	5.633	-0.091	0.191	-0.169	0.408		
	St.error	7.805	0.133	0.077	12.455	0.368	0.392	0.199	0.421	0.591	25
	t ratio	-0.271	0.190	-0.350	0.452	-0.247	0.488	-0.851	0.970		
2	ß										
	St.error										
	t ratio										
3	ß							-0.314	0.519		
	St.error							0.067	0.187	0.529	25
	t ratio							-4.667	2.780		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level

[b] N is number of datapoints

[c] Samples 1, 10, 14, 23, 25, 26, 28 and 29 were left out of the data set

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 4: *Regression results for the NLLSR sorption isotherms of Cd, with metal content extracted with 0.01M $CaCl_2$ as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	-1.559	-0.000	-0.010	2.012	0.151	-0.035	-0.162	0.299		
	St.error	1.773	0.053	0.028	1.951	0.237	0.161	0.070	0.110	0.842	27
	t ratio	-0.879	-0.001	-0.367	1.032	0.637	-0.215	-2.318	2.729		
2	ß										
	St.error										
	t ratio										
3	ß							-0.229	0.428		
	St.error							0.016	0.064	0.801	27
	t ratio							-14.07	6.629		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level
[b] N is number of datapoints
[c] Samples 1, 6, 10, 20, 23 and 29 were left out of the data set
[d] ß is regression coefficient
[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 5: *Regression results for the NLLSR sorption isotherms of Cd, with metal content extracted with 0.01M $CaCl_2$ as one of the independent parameters, derived from the untransformed values of the LMB data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	-1.265	0.017	0.009		0.029	-0.030	-0.053	0.057		
	St.error	0.310	0.009	0.007		0.057	0.066	0.023	0.028	0.599	34
	t ratio	-4.079	1.939	1.211		0.511	-0.459	-2.302	2.051		
2	ß	-1.155	0.020	0.006				-0.044	0.053		
	St.error	0.216	0.005	0.003				0.018	0.027	0.594	34
	t ratio	-5.348	3.724	2.560				-2.398	1.963		
3	ß	0.158	0.013	0.008							
	St.error	0.039	0.004	0.002						0.544	34
	t ratio	4.073	3.610	3.727							
4	ß	0.413	0.022	0.032	-0.229	0.031	0.439	0.054	-0.098		
	St.error	0.319	0.009	0.017	0.264	0.108	0.375	0.057	0.067	0.814	31
	t ratio	1.295	2.550	1.841	-0.868	0.289	1.170	0.947	-1.476		
5	ß	-0.728	0.018	0.017			-0.152	-0.027	0.056		
	St.error	0.308	0.004	0.004			0.104	0.024	0.039	0.801	31
	t ratio	-2.360	4.665	4.382			-1.466	-1.159	1.419		
6	ß	0.132	0.015	0.008							
	St.error	0.028	0.003	0.002						0.752	31
	t ratio	4.730	5.632	5.192							

[a] Model 1 is the model with the highest explained variance (with outliers)
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level (with outliers)
Model 3 is the model with only significant regression coefficients at the p=0.05 level (with outliers)
Model 4 is the model with the highest explained variance (without outliers)
Model 5 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level (without outliers)
Model 6 is the model with only significant regression coefficients at the p=0.05 level (without outliers)

[b] N is number of datapoints

[c] Samples 16, 19, 21, 26, 30, 32, 38 and 39 were left out of the data set for model 1, 2 and 3
Samples 16, 19, 21, 24, 26, 30, 32, 33, 37, 38 and 39 were left out of the data set for model 4, 5 and 6

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 6: *Regression results for the Linear Regression of the untransformed and log-transformed values of the measured total metal content in the solid phase and the predicted total metal content in the solid phase according to the fitted LR and NLLSR sorption isotherms of Cd. Each model is given by:*

$$X_{measured} = bX_{predicted}$$

Model[a]	Data		b	R^2	N[bc]	Ref.
1	Untrans-formed	ß[d]	1.138			
		St.error	0.074	0.899	28	Table 1.1
		t ratio	15.478			
2	Untrans-formed	ß	1.223			
		St.error	0.078	0.901	28	Table 2.2
		t ratio	15.701			
3	Log-trans-formed	ß	1.000			
		St.error	0.133	0.676	28	Table 1.1
		t ratio	7.513			
4	Log-trans-formed	ß	0.818			
		St.error	0.175	0.447	28	Table 2.2
		t ratio	4.673			

[a] Model 1 and 3 are LR models with free metal activity as independent parameter
Model 2 and 4 are NLLSR models with free metal activity as independent parameter
[b] N is number of datapoints
[c] Samples 1, 14, 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient

Cu

Table 7: *Regression results for the LR Freundlich sorption isotherms of Cu derived from the log-transformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}\left(H^{+}\right)^{m}\left(M^{2+}\right)^{n}$$

Model[a]		p	q[e]	r	s	m	n	R^2	St.er.	N[bc]
1	ß[d]	-1.410	= 0			-0.91	0.717	0.570	0.382	30
	t ratio					-4.68	3.63			
2	ß	-0.103	= 0		0.47	-0.199	0.26	0.564	0.392	30
	t ratio				2.84	-3.93	0.94			
3	ß	-1.410		= 0	= 0	--0.91	0.717	0.570	0.382	30
	t ratio					-4.68	3.63			
4	ß	-0.781	0.58	= 0	= 0	-0.156	0.13	0.536	0.404	30
	t ratio		2.47			-2.59	0.41			

[a] Model 1 and 3 based on free Cu activity
Model 2 and 4 based on total Cu concentration in solution
[b] N is number of datapoints
[c] Samples 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient
[e] = 0: regression coefficient is set to zero

Table 8: *Regression results for the NLLSR sorption isotherms of Cu, with free metal activity as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left(M^{2+}\right)^{n}$$

Model[a]		a	b	c	10^{p}	r	s	m	n	R^2	N[bc]
1	ß[de]	-3.580	-2.821	0.561		0.609	-0.318	-0.499	0.377		
	St.error	16.883	4.597	0.445		0.528	0.249	0.148	0.155	0.328	30
	t ratio	-0.212	-0.614	1.261		1.155	-1.278	-3.375	2.443		
2	ß										
	St.error										30
	t ratio										
3	ß							-0.471	0.348		
	St.error							0.091	0.131	0.202	30
	t ratio							-5.190	2.653		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level
[b] N is number of datapoints
[c] Samples 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient
[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 9: *Regression results for the NLLSR sorption isotherms of Cu, with total metal concentration in solution as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}[M]^{n}$$

Model[a]		a	b	c	10^{p}	r	s	m	n	R^2	N[bc]
1	ß[de]	-34.29	14.574	0.092	3.682	-1.159	0.080	-0.115	-0.252		
	St.error	89.950	8.019	0.726	48.421	4.488	1.385	0.456	1.547	0.430	26
	t ratio	-0.381	1.817	0.127	0.076	-0.258	0.058	-0.253	-0.163		
2	ß	-25.61	14.072			-1.358		-0.177	-0.344		
	St.error	19.082	4.440			2.301		0.183	0.833	0.426	26
	t ratio	-1.342	3.170			-0.590		-0.968	-0.413		
3	ß	-16.31	11.385					-0.146			
	St.error	12.522	3.416					0.068		0.412	26
	t ratio	-1.302	3.333					-2.148			

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level

[b] N is number of datapoints

[c] Samples 2, 10, 22, 24, 25, 26 and 28 were left out of the data set

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 10: *Regression results for the NLLSR sorption isotherms of Cu, with metal content extracted with 0.01M $CaCl_2$ as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	9.451	-0.815	0.331	27.240	1.211	0.882	-0.066	1.259		
	St.error	7.785	2.931	0.151	57.302	0.524	0.272	0.055	0.344	0.904	26
	t ratio	1.214	-0.278	2.184	0.475	2.311	3.242	-1.191	3.661		
2	ß										
	St.error										
	t ratio										
3	ß	10.629		0.298		2.123	1.333	-0.158	1.550		
	St.error	3.645		0.141		0.354	0.249	0.022	0.312	0.889	26
	t ratio	2.916		2.117		5.995	5.351	-7.101	4.967		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level
[b] N is number of datapoints
[c] Samples 2, 10, 22, 24, 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient
[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 11: *Regression results for the NLLSR sorption isotherms of Cu, with metal content extracted with 0.01M $CaCl_2$ as one of the independent parameters, derived from the untransformed values of the LMB data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	10.723	-0.696	0.439		2.228	-0.579	-0.329	1.307		
	St.error	2.093	0.717	0.111		0.512	0.381	0.182	0.694	0.832	42
	t ratio	5.123	-0.972	3.941		4.352	-1.521	-1.801	1.883		
2	ß										
	St.error										
	t ratio										
3	ß	10.231		0.344		2.415		-0.301	1.878		
	St.error	1.346		0.107		0.412		0.117	0.485	0.811	42
	t ratio	7.599		3.222		5.857		-2.562	3.876		
4	ß	10.005	-0.688	0.426		2.074	-0.522	-0.315	1.195		
	St.error	2.176	1.087	0.124		1.219	0.688	0.325	1.425	0.698	39
	t ratio	4.597	-0.633	3.433		1.701	-0.759	-0.969	0.838		
5	ß										
	St.error										
	t ratio										
6	ß	9.502		0.289		1.729		-0.271	1.313		
	St.error	1.469		0.120		0.623		0.096	0.569	0.660	39
	t ratio	6.471		2.407		2.776		-2.827	2.308		

[a] Model 1 is the model with the highest explained variance (with outliers)
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level (with outliers)
Model 3 is the model with only significant regression coefficients at the p=0.05 level (with outliers)
Model 4 is the model with the highest explained variance (without outliers)
Model 5 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level (without outliers)
Model 6 is the model with only significant regression coefficients at the p=0.05 level (without outliers)

[b] N is number of datapoints

[c] Samples 8, 21 and 35 were left out of the data set for model 4, 5 and 6

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 12: *Regression results for the Linear Regression of the untransformed and log-transformed values of the measured total metal content in the solid phase and the predicted total metal content in the solid phase according to the fitted LR and NLLSR sorption isotherms of Cu. Each model is given by:*

$$X_{measured} = bX_{predicted}$$

Model[a]	Data		b	R^2	N[bc]	Ref.
1	Untrans-	ß[d]	1.223			
	formed	St.error	0.207	0.548	30	Table 7.1
		t ratio	5.924			
2	Untrans-	ß	1.005			
	formed	St.error	0.152	0.601	30	Table 7.1
		t ratio	6.608			
3	Log-trans-	ß	1.024			
	formed	St.error	0.053	0.927	30	Table 8.3
		t ratio	19.189			
4	Log-trans-	ß	0.876			
	formed	St.error	0.054	0.901	30	Table 8.3
		t ratio	16.252			

[a] Model 1 and 3 are LR models with free metal activity as independent parameter
Model 2 and 4 are NLLSR models with free metal activity as independent parameter
[b] N is number of datapoints
[c] Samples 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient

Pb

Table 13: *Regression results for the LR Freundlich sorption isotherms of Pb derived from the log-transformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}\left(H^{+}\right)^{m}\left(M^{2+}\right)^{n}$$

Model[a]		p	q[e]	r	s	m	n	R^2	N[bc]
1	ß[d]	0.442	= 0		-0.223	-0.716	0.556		
	St.error	0.407			0.136	0.140	0.118	0.550	29
	t ratio	1.087			-1.641	-5.121	4.717		
2	ß	2.064	= 0			-0.186	0.571		
	St.error	0.256				0.045	0.097	0.591	29
	t ratio	8.065				-4.178	5.866		
3	ß	0.874	-0.338	= 0	= 0	-0.712	0.533		
	St.error	0.518	0.182			0.138	0.118	0.563	29
	t ratio	1.689	-1.860			-5.162	4.505		
4	ß	2.064		= 0	= 0	-0.186	0.571		
	St.error	0.256				0.045	0.097	0.591	29
	t ratio	8.065				-4.178	5.866		

[a] Model 1 and 3 based on free Pb activity
Model 2 and 4 based on total Pb concentration in solution
[b] N is number of datapoints
[c] Samples 14, 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient
[e] = 0: regression coefficient is set to zero

Table 14: *Regression results for the NLLSR sorption isotherms of Pb, with free metal activity as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left(M^{2+}\right)^{n}$$

Model[a]		a	b	c	10^{p}	r	s	m	n	R^2	N[bc]
1	ß[de]	65.805	-1.799	-1.700	4.271	-0.011	0.077	-0.597	0.500		
	St.error	164.29	3.332	1.934	24.768	0.354	0.221	0.750	0.608	0.497	28
	t ratio	0.401	-0.540	-0.879	0.172	-0.030	0.348	-0.796	0.823		
2	ß	53.248						-0.817	0.674		
	St.error	29.478						0.042	0.061	0.445	28
	t ratio	1.806						-19.37	11.096		
3	ß							-0.829	0.664		
	St.error							0.036	0.051	0.371	28
	t ratio							-22.85	12.884		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level

[b] N is number of datapoints

[c] Samples 5, 14, 25, 26 and 28 were left out of the data set

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 15: *Regression results for the NLLSR isotherms of Pb, with total metal concentration in solution as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}[M]^{n}$$

Model[a]		a	b	c	10^{p}	r	s	m	n	R^2	N[bc]
1	ß[de]	-1.310 e+3	-5.564 e+0	-8.367 e-1	1.490 e+3	2.216 e-2	1.968 e-2	-9.032 e-3	4.122 e-2		
	St.error	6778.7	4.246	2.493	6831.6	0.108	0.086	0.043	0.186	0.592	29
	t ratio	-0.193	-1.310	-0.336	0.218	0.206	0.230	-0.208	0.221		
2	ß		-1.499		163.75			-0.148	0.481		
	St.error		1.823		69.255			0.046	0.108	0.551	29
	t ratio		-0.822		2.364			-3.246	4.441		
3	ß				152.63			-0.158	0.511		
	St.error				66.065			0.047	0.110	0.539	29
	t ratio				2.310			-3.345	4.651		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level

[b] N is number of datapoints

[c] Samples 5, 11, 14 and 20 were left out of the data set

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 16: *Regression results for the NLLSR sorption isotherms of Pb, with metal content extracted with 0.01M $CaCl_2$ as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	-40.54	46.912	-3.218	5.243	-0.890	1.180	-0.269	1.109		
	St.error	25.968	8.783	1.958	6.309	0.834	0.461	0.093	0.304	0.973	14
	t ratio	-1.561	5.341	-1.643	0.831	-1.067	2.559	-2.894	3.651		
2	ß										
	St.error										
	t ratio										
3	ß		29.056				0.993	-0.373	1.317		
	St.error		5.048				0.086	0.020	0.114	0.938	14
	t ratio		5.755				11.593	-18.78	11.520		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level

[b] N is number of datapoints

[c] Samples 2, 3, 5, 6, 7, 10, 15 to 24, 27, 30 and 33 were left out of the data set

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 17: *Regression results for the NLLSR sorption isotherms of Pb, with metal content extracted with 0.01M $CaCl_2$ as one of the independent parameters, derived from the untransformed values of the LMB data set (without values below detection limit). Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	11.883	-7.049	1.144	18.008	1.077	0.035	-0.090	0.331		
	St.error	7.975	11.944	0.538	25.448	0.147	0.101	0.142	0.362	0.920	33
	t ratio	1.490	-0.590	2.125	0.708	7.317	0.345	-0.634	0.912		
2	ß										
	St.error										
	t ratio										
3	ß	15.454				0.740	1.032	-0.287	0.745		
	St.error	3.653				0.146	0.334	0.088	0.103	0.918	33
	t ratio	4.231				5.053	3.089	-3.252	7.248		
4	ß	8.321	1.097	0.928		3.155	2.854	-3.400	8.759		
	St.error	2.760	0.394	0.173		1.496	1.770	1.688	4.742	0.872	30
	t ratio	3.015	2.788	5.354		2.108	1.613	-2.014	1.847		
5	ß										
	St.error										
	t ratio										
6	ß	7.910	1.109	1.102		8.376		-4.460	11.952		
	St.error	3.239	0.481	0.199		8.661		4.273	12.288	0.812	30
	t ratio	2.442	2.303	5.541		0.967		-1.044	0.973		

[a] Model 1 is the model with the highest explained variance (with outliers)
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level (with outliers)
Model 3 is the model with only significant regression coefficients at the p=0.05 level (with outliers)
Model 4 is the model with the highest explained variance (without outliers)
Model 5 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level (without outliers)
Model 6 is the model with only significant regression coefficients at the p=0.05 level (without outliers)

[b] N is number of datapoints

[c] Samples 3, 5, 26, 28, 33, 34, 35, 37 and 42 were left out of the data set for model 1, 2 and 3
Samples 3, 5, 6, 8, 20, 26, 28, 33, 34, 35, 37 and 42 were left out of the data set for model 4, 5 and 6

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 18: *Regression results for the Linear Regression of the untransformed and log-transformed values of the measured total metal content in the solid phase and the predicted total metal content in the solid phase according to the fitted LR and NLLSR sorption isotherms of Pb. Each model is given by:*

$$X_{measured} = bX_{predicted}$$

Model[a]	Data		b	R^2	N[bc]	Ref.
1	Untrans-	ß[d]	1.088			
	formed	St.error	0.085	0.853	29	Table 13.1
		t ratio	12.746			
2	Untrans-	ß	1.054			
	formed	St.error	0.079	0.865	29	Table 14.2
		t ratio	13.404			
3	Log-trans-	ß	0.999			
	formed	St.error	0.029	0.978	29	Table 13.1
		t ratio	35.084			
4	Log-trans-	ß	0.946			
	formed	St.error	0.030	0.972	29	Table 14.2
		t ratio	31.313			

[a] Model 1 and 3 are LR models with free metal activity as independent parameter
Model 2 and 4 are NLLSR models with free metal activity as independent parameter
[b] N is number of datapoints
[c] Samples 14, 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient

Zn

Table 19: *Regression results for the LR Freundlich sorption isotherms of Zn derived from the log-transformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}\left(H^{+}\right)^{m}\left(M^{2+}\right)^{n}$$

Model[a]		p	q[e]	r	s	m	n	R^2	St.er.	N[bc]
1	ß[d]	-0.318	= 0		0.52	-0.445	0.66	0.716	0.382	30
	t ratio				2.88	-7.86	4.61			
2	ß	-0.104	= 0		0.64	-0.360	0.77	0.729	0.373	30
	t ratio				3.39	-7.44	4.86			
3	ß	-1.123	0.77	= 0	= 0	-0.400	0.72	0.742	0.364	30
	t ratio		3.43			-7.42	5.10			
4	ß	-0.922	0.85	= 0	= 0	-0.300	0.80	0.739	0.365	30
	t ratio		3.61			-5.92	5.05			

[a] Model 1 and 3 based on free Zn activity
Model 2 and 4 based on total Zn concentration in solution
[b] N is number of datapoints
[c] Samples 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient
[e] = 0: regression coefficient is set to zero

Table 20: *Regression results for the NLLSR sorption isotherms of Zn, with free metal activity as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left(M^{2+}\right)^{n}$$

Model[a]		a	b	c	10^{p}	r	s	m	n	R^2	N[bc]
1	ß[de]	3.488 e+1	-5.865 e-1	1.354 e+0	5.762 e-6	2.990 e-1	-3.461 e-1	-1.287 e+0	1.000 e+0		
	St.error	7.732 e+1	4.098 e+0	2.154 e+0	3.068 e-5	2.704 e-1	1.981 e-1	3.630 e-1	2.517 e-1	0.859	27
	t ratio	0.451	-0.143	0.629	0.188	1.106	-1.747	-3.544	3.975		
2	ß										
	St.error										
	t ratio										
3	ß							-0.470	0.505		
	St.error							0.006	0.036	0.782	27
	t ratio							-76.19	13.894		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level

[b] N is number of datapoints

[c] Samples 5, 9, 14, 25, 26 and 28 were left out of the data set

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 21: *Regression results for the NLLSR sorption isotherms of Zn, with total metal concentration in solution as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}[M]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	4.486 e+1	-9.303 e-1	1.625 e+0	1.245 e-5	3.665 e-1	-2.116 e-1	-1.140 e+0	1.185 e+0		
	St.error	4.987 e+1	3.370 e+0	1.651 e+0	5.893 e-5	2.141 e-1	2.925 e-1	2.991 e-1	2.868 e-1	0.856	30
	t ratio	0.899	-0.276	0.984	0.211	1.712	-0.723	-3.813	4.131		
2	ß										
	St.error										
	t ratio										
3	ß							-0.433	0.588		
	St.error							0.005	0.038	0.802	30
	t ratio							-89.64	15.566		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level

[b] N is number of datapoints

[c] Samples 5, 6 and 14 were left out of the data set

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 22: *Regression results for the NLLSR sorption isotherms of Zn, with metal content extracted with 0.01M $CaCl_2$ as one of the independent parameters, derived from the untransformed values of the combined Hoop-Janssen data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	-0.029	-0.267	2.707	0.007	-0.766	-0.574	-0.881	0.763		
	St.error	16.592	1.052	0.521	0.009	0.145	0.147	0.088	0.054	0.984	28
	t ratio	-0.002	-0.254	5.195	0.807	-5.269	-3.911	-9.997	14.08		
2	ß										
	St.error										
	t ratio										
3	ß			1.849		-0.832		-0.495	0.547		
	St.error			0.677		0.148		0.013	0.040	0.936	28
	t ratio			2.732		-5.635		-38.17	13.51		

[a] Model 1 is the model with the highest explained variance
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level
Model 3 is the model with only significant regression coefficients at the p=0.05 level

[b] N is number of datapoints

[c] Samples 5, 7, 10, 23 and 29 were left out of the data set

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 23: *Regression results for the NLLSR sorption isotherms of Zn, metal content extracted with 0.01M $CaCl_2$ as one of the independent parameters, derived from the untransformed values of the LMB data set. Each model is given by:*

$$X = a + b\{\%OC\} + c\{\%clay\} + 10^{p}\{CEC\}^{q}\{\%OC\}^{r}\{\%clay\}^{s}(H^{+})^{m}\left[M_{CaCl_2}\right]^{n}$$

Model[a]		a	b	c	10^p	r	s	m	n	R^2	N[bc]
1	ß[de]	19.398	1.927	4.157		2.487	-0.857	-0.409	6.670		
	St.error	8.619	1.093	0.552		29.240	29.848	3.493	54.682	0.827	36
	t ratio	2.251	1.763	7.531		0.085	-0.029	-0.117	0.122		
2	ß	19.596	1.947	4.141		1.904		-0.370	7.053		
	St.error	8.455	1.041	0.537		4.745		1.825	44.667	0.827	36
	t ratio	2.318	1.870	7.706		0.401		-0.203	0.158		
3	ß	41.903	-18.88	3.405		1.318		-0.173			
	St.error	15.845	6.814	0.772		0.113		0.038		0.780	36
	t ratio	2.645	-2.770	4.412		11.645		-4.539			
4	ß	6.854	3.172	4.285		0.076	-2.272	-0.418	0.595		
	St.error	15.670	0.916	0.615		1.407	8.014	0.592	1.057	0.860	33
	t ratio	0.437	3.464	6.964		0.054	-0.284	-0.706	0.563		
5	ß	12.889	3.059	4.041				-0.165	1.351		
	St.error	9.597	0.762	0.381				0.657	10.804	0.855	33
	t ratio	1.343	4.015	10.613				-0.251	0.125		
6	ß		3.615	4.142				-0.145			
	St.error		0.666	0.424				0.053		0.841	33
	t ratio		5.430	9.774				-2.721			

[a] Model 1 is the model with the highest explained variance (with outliers)
Model 2 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level (with outliers)
Model 3 is the model with only significant regression coefficients at the p=0.05 level (with outliers)
Model 4 is the model with the highest explained variance (without outliers)
Model 5 is the model with independent parameters which are plausible from a theoretical perspective and which may include coefficients not significant at the p=0.05 level (without outliers)
Model 6 is the model with only significant regression coefficients at the p=0.05 level (without outliers)

[b] N is number of datapoints

[c] Samples 16, 24, 26, 28, 33 and 34 were left out of the data set for model 1, 2 and 3
Samples 6, 8, 16, 24, 26, 28, 33, 34 and 37 were left out of the data set for model 4, 5 and 6

[d] ß is regression coefficient

[e] The regression coefficient of the Cation Exchange Capacity (CEC) is set to zero

Table 24: *Regression results for the Linear Regression of the untransformed and log-transformed values of the measured total metal content in the solid phase and the predicted total metal content in the solid phase according to the fitted LR and NLLSR sorption isotherms of Zn. Each model is given by:*

$$X_{measured} = bX_{predicted}$$

Model[a]	Data		b	R^2	N[bc]	Ref.
1	Untrans-formed	ß[d]	1.602			
		St.error	0.138	0.823	30	Table 19.1
		t ratio	11.599			
2	Untrans-formed	ß	1.277			
		St.error	0.099	0.853	30	Table 20.3
		t ratio	12.953			
3	Log-trans-formed	ß	0.998			
		St.error	0.030	0.974	30	Table 19.1
		t ratio	33.167			
4	Log-trans-formed	ß	0.951			
		St.error	0.033	0.966	30	Table 20.3
		t ratio	28.652			

[a] Model 1 and 3 are LR models with free metal activity as independent parameter
Model 2 and 4 are NLLSR models with free metal activity as independent parameter
[b] N is number of datapoints
[c] Samples 25, 26 and 28 were left out of the data set
[d] ß is regression coefficient

$CaCl_2$ extraction

Table 25: *Regression results for the Linear Regression of pH according to the $CaCl_2$ extraction and pH of the pore water of the combined Hoop-Janssen data set. The model is given by:*

$$pH_{CaCl_2} = a + b \times pH_{porewater}$$

Model		a	b	R^2	N[a]
1	ß[b]	0.568	0.847		
	St.error	0.234	0.040	0.936	33
	t ratio	2.434	21.261		

[a] N is number of datapoints
[b] ß is regression coefficient

Table 26: *Regression results for the Linear Regression of the untransformed values of the total metal concentrations and the metal contents extracted with 0.01M $CaCl_2$ metal of the combined Hoop-Janssen data set. Each model is given by:*

$$[M_{CaCl_2}] = a + b[M]$$

Model[a]		a	b	R^2	N[bc]	Metal
1	ß[d]	-0.000	1.036			
	St.error	0.000	0.018	0.992	30	Cd
	t ratio	-4.495	57.898			
2	ß	0.000	0.668			
	St.error	0.000	0.057	0.836	29	Cd
	t ratio	-1.979	11.720			
3	ß	0.000	-0.004			
	St.error	0.000	0.067	0.000	26	Cu
	t ratio	2.690	-0.065			
4	ß	0.000	4.204			
	St.error	0.001	0.898	0.628	15	Pb
	t ratio	0.486	4.681			
5	ß	-0.002	2.082			
	St.error	0.006	0.240	0.722	31	Zn
	t ratio	-0.384	8.671			

[a] Model 1, 3, 4 and 5 are models with outliers
Model 2 is the model without outliers
[b] N is number of datapoints
[c] Samples 14, 23 and 29 were left out of the data set for model 1
Samples 10, 14, 23 and 29 were left out of the data set for model 2
Samples 2, 10, 22, 24, 25, 26 and 28 were left out of the data set for model 3
Samples 2, 3, 6, 7, 10, 14 to 23, 27, 30 and 33 were left out of the data set for model 4
Samples 23 and 29 were left out of the data set for model 5
[d] ß is regression coefficient