

# A comparison of characterisation and modelling approaches to predict dissolved metal concentrations in soils

Judith M. Garforth<sup>A,B,D</sup>, Andrew M. Tye<sup>C</sup>, Scott D. Young<sup>B</sup>, Elizabeth H. Bailey<sup>B</sup> and Stephen Lofts<sup>A,\*</sup> 

**Environmental context.** It is useful to know the concentration of 'labile', or chemically active, metal in soils because it can be used to predict metal solubility and environmental impact. Several methods for extracting the labile metal from soils have been proposed, and here we have tested two of these to see how well the resulting data can be used to model metal solubility. Such mixed approaches can be applied to different soil types with the potential to model metal solubility over large areas.

For full list of author affiliations and declarations see end of paper

**\*Correspondence to:**

Stephen Lofts  
UK Centre for Ecology and Hydrology,  
Lancaster Environment Centre, Lancaster,  
LA1 4AP, UK  
Email: [stlo@ceh.ac.uk](mailto:stlo@ceh.ac.uk)

**Handling Editor:**

Kevin Wilkinson

**Received:** 26 July 2023

**Accepted:** 24 October 2023

**Published:** 7 December 2023

**Cite this:**

Garforth JM *et al.* (2023)  
*Environmental Chemistry*  
doi:[10.1071/EN23075](https://doi.org/10.1071/EN23075)

© 2023 The Author(s) (or their employer(s)). Published by CSIRO Publishing.

This is an open access article distributed under the Creative Commons Attribution 4.0 International License ([CC BY](https://creativecommons.org/licenses/by/4.0/))

OPEN ACCESS

## ABSTRACT

**Rationale.** Predicting terrestrial metal dynamics requires modelling of metal solubility in soils. Here, we test the ability of two geochemical speciation models that differ in complexity and data requirements (WHAM/Model VII and POSSMs), to predict metal solubility across a broad range of soil properties, using differing estimates of the labile soil metal concentration. **Methodology.** Using a dataset of UK soils, we characterised basic properties including pH and the concentrations of humic substances, mineral oxides and metals. We estimated labile metal by extraction with  $0.05 \text{ mol L}^{-1} \text{ Na}_2\text{H}_2\text{EDTA}$  and by multi-element isotopic dilution ( $E$ -value). Dissolved concentrations of Ni, Cu, Zn, Cd and Pb were estimated in  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions using the total metal ( $\{M\}_{\text{total}}$ ), the EDTA-extracted pool ( $\{M\}_{\text{EDTA}}$ ) and the  $E$ -value ( $\{M\}_E$ ) as alternative estimates of the chemically reactive metal concentration. **Results.** Concentrations of  $\{M\}_{\text{EDTA}}$  were highly correlated with values of  $\{M\}_E$ , although some systematic overestimation was seen. Both WHAM/Model VII and POSSMs provided reasonable predictions when  $\{M\}_{\text{EDTA}}$  or  $\{M\}_E$  were used as input. WHAM/Model VII predictions were improved by fixing soil humic acid to a constant proportion of the soil organic matter, instead of the measured humic and fulvic acid concentrations. **Discussion.** This work provides further evidence for the usefulness of speciation modelling for predicting soil metal solubility, and for the usefulness of EDTA-extracted metal as a surrogate for the labile metal pool. Predictions may be improved by more robust characterisation of the soil and porewater humic substance content and quality.

**Keywords:** analytical chemistry, chemical extraction, isotopic dilution, modelling, partitioning, soil chemistry, speciation, trace metals.

## Introduction

The dynamics of soil metals are known to be controlled, at least partly, by the solid–solution partitioning of a labile or ‘geochemically active’ pool of soil-associated metal, rather than by the total soil metal concentration (Degryse *et al.* 2004). Partitioning measurements, along with modelled predictions of metal speciation, form the basis of our understanding of metal transport in soils and aquifers. The solid–solution partitioning of soil metals may be predicted using either empirical or mechanistic ‘assemblage’ models. These predict metal adsorption and solution speciation using the concentrations of soil components having proton and metal binding capabilities, including soil organic matter (SOM), Fe, Mn and Al hydroxides, and clay. Over a number of years, studies have compared predicted and measured results to test this approach (Groenenberg *et al.* 2012). The majority of these studies have used either WHAM/Model V–VII (Tipping and Hurley 1992; Tipping 1998; Tipping *et al.* 2011), the NICA–Donnan model (Kinniburgh *et al.* 1996) or the Stockholm Humic Model (SHM) (Gustafsson 2001) as

the sub-model to describe metal binding to organic matter in the assemblage model. Models VI and VII have been incorporated into the assemblage model, WHAM, which also contains models for surface complexation at oxide surfaces and ion exchange on clays (Lofts and Tipping 1998). The NICA-Donnan model has been incorporated into modelling frameworks such as ECOSAT (Keizer and van Riemsdijk 2009) and ORCHESTRA (Meeussen 2003), in order to describe metal binding in whole soils.

Assemblage models require, as an input, a measurement of the concentration of soil metal that is active in adsorption processes. Various measurements have been used to determine estimates of the 'active pool'. These include metal extracted by  $0.05 \text{ mol L}^{-1}$  EDTA (Quevauviller 1998),  $2 \text{ mol L}^{-1}$   $\text{HNO}_3$  (Weng et al. 2001, 2002a),  $0.43 \text{ mol L}^{-1}$   $\text{HNO}_3$  (Tipping et al. 2003; Dijkstra et al. 2004),  $0.22 \text{ mol L}^{-1}$   $\text{HNO}_3$  (Almås et al. 2007) and pseudo-total metal concentration determined by aqua regia digestion (Schröder et al. 2005). It has been noted that discrepancies between predicted and measured results in the aforementioned studies may be partly due to over-estimation or underestimation of the 'active pool' of metal (Weng et al. 2001, 2002a; Almås et al. 2007) by these extraction measurements. However, over the past two decades, methods using isotopic dilution (ID) with radioisotopes (e.g. Tye et al. 2003) and stable metal isotopes (Ahnstrom and Parker 2001; Nolan et al. 2004) have been used to estimate the 'E-value' or isotopically exchangeable metal concentration. Conceptually, such measurements represent the best description of the pool of soil metal that is active in rapid (de)sorption processes. Results have shown that the E-value of a metal is dependent on: (i) the source of contamination, (ii) soil physiochemical properties such as pH and redox potential and (iii) soil-metal contact time (Ayoub et al. 2003). Consideration of contact time is particularly important in measuring E-values to allow the added isotope to equilibrate with the soil while not undergoing any significant removal from the labile pool due to fixation or precipitation (Mao et al. 2017). A further advantage of the ID approach is that, since it involves suspension of soil in a weak electrolyte solution, it allows for simultaneous determination of metal solubility. For example, Marzouk et al. (2013) measured Zn, Cd and Pb solution concentrations in 246 soil suspensions in  $0.01 \text{ mol L}^{-1}$   $\text{Ca}(\text{NO}_3)_2$  when undertaking E-value measurements.

Whilst assemblage modelling is based on sound mechanistic and technical understanding, it requires considerable amounts of input data, and this may be expensive or impractical when analysis of large datasets is required. Simpler, empirical approaches may therefore be useful to obtain the relationships between adsorbed metal and solution variables including the total and free ion concentrations in soil pore waters. The recently published POSSMs model (Lofts 2022) seeks to emulate the outputs of mechanistic models by calculating the free, solution-bound and sorbed

soil metal within a single calculation, whilst restricting input parameters to an estimate of the labile (or geochemically active) metal, SOM content, solution phase pH and dissolved organic matter (DOM) concentrations. The POSSMs model has been largely tested on organic rich and lower pH soils. Previous work on multiphase modelling (Mao et al. 2017) has shown that at higher pH and lower organic matter concentrations, other phases such as the mineral oxides are predicted to become increasingly important in metal binding. The aim of this paper is to compare the outputs of (i) the geochemical assemblage model WHAM/Model VII and (ii) POSSMs, to predict soluble Ni, Cu, Zn, Cd and Pb concentrations using a dataset of soils with a range of pH values similar to that of Mao et al. (2017), but with a wider range of organic matter contents. In particular we compare measured dissolved metal concentrations with model predictions computed using three expressions of the 'geochemically active pool'. These included total soil metal ( $\{M\}_{\text{total}}$ ), metal extracted using the EU recommended extractant of  $0.05 \text{ mol L}^{-1}$  EDTA ( $\{M\}_{\text{EDTA}}$ ) and the labile pool as measured by multi-element ID ( $\{M\}_{\text{E}}$ ) with solution concentrations measured in  $0.01 \text{ mol L}^{-1}$   $\text{Ca}(\text{NO}_3)_2$  soil suspensions.

## Experimental

A glossary of terms is provided in Table 1.

### Soils

The comparison of modelling approaches used a dataset of 109 UK soils, which were selected for their wide range of total soil metal concentrations, metal sources and soil characteristics. A total of 56 soils were collected specifically for this investigation, from field sites chosen because of their metal contamination histories, or underlying geology. The remaining soils were selected from archives held by the British Geological Survey (G-BASE, 44 soils; Rawlins et al. 2012) and the University of Nottingham (nine soils; Thornton et al. 2008; Atkinson 2010).

All soils, whether freshly collected or archived, were air dried and sieved to  $< 2 \text{ mm}$ . Sub samples of both archived and freshly collected soils were agate ball-milled (Fritsch Pulverisette, Germany) for total metal analysis. Loss on ignition (LOI;  $\text{g} (\text{g dry weight soil})^{-1}$ ) was measured after heating soils ( $\sim 1 \text{ g}$ ,  $< 2 \text{ mm}$ ) to  $550^\circ\text{C}$  for 2 h in a muffle furnace, and used as an estimate of SOM content (Karam 1993).

### Soil metal binding phases

To predict soil solution metal concentrations using the chemical speciation model WHAM/Model VII the following soil phases were quantified. Metal oxides ( $\text{MnOx}$  and  $\text{FeOx}$ ) were extracted in a solution containing  $0.07 \text{ mol L}^{-1}$

**Table 1.** Glossary of terms, including units.

Term	Meaning	Units
$\{M\}_{\text{total}}$	Soil metal concentration obtained by extraction with $\text{HNO}_3$ and $\text{HClO}_4$ (2 mL:1 mL) followed by digestion with 2.5 mL HF.	$\text{mol kg}^{-1}$
$\{M\}_{\text{EDTA}}$	Soil metal concentration obtained by extraction with $0.05 \text{ mol L}^{-1} \text{ Na}_2\text{H}_2\text{EDTA}$ .	$\text{mol kg}^{-1}$
$\{M\}_{\text{E}}$	Soil metal concentration obtained by ID ( <i>E</i> -value).	$\text{mol kg}^{-1}$
$\text{pH}_{\text{ca}}$	pH measured in $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$ suspension at a soil/solution ratio of $33 \text{ g L}^{-1}$ .	–
$\text{AM}_{\text{soil}}$	Average atomic mass of metal in the un-spiked control soil.	$\text{g mol}^{-1}$
$\text{AM}_{\text{sp}}$	Average atomic mass of metal in the isotope spike.	$\text{g mol}^{-1}$
$C_{\text{sp}}$	Concentration of metal in the isotope spike.	$\text{mol L}^{-1}$
$V_{\text{sp}}$	Volume of the isotope spike in the spiked soil suspension.	L
$W$	Mass of soil in suspension.	kg
${}^s\text{IA}_{\text{sp}}$	Isotopic abundance of spike isotope in the isotope spike solution.	$\text{g g}^{-1}$
${}^r\text{IA}_{\text{sp}}$	Isotopic abundance of reference isotope in the isotope spike solution.	$\text{g g}^{-1}$
${}^s\text{IA}_{\text{sus}}$	Isotopic abundance of spike isotope in the solution phase of an un-spiked control suspension.	$\text{g g}^{-1}$
${}^r\text{IA}_{\text{sus}}$	Isotopic abundance of reference isotope in the solution phase of an un-spiked control suspension.	$\text{g g}^{-1}$
${}^s\text{IA}_{\text{sp-sus}}$	Isotopic abundance of spike isotope in the solution phase of the spiked suspension.	$\text{g g}^{-1}$
${}^r\text{IA}_{\text{sp-sus}}$	Isotopic abundance of reference isotope in the solution phase of the spiked suspension.	$\text{g g}^{-1}$
${}^{204}\text{Pb cps}/{}^{208}\text{Pb cps}$	Ratio of ICP-MS count rates of the isotopes ${}^{204}\text{Pb}$ and ${}^{208}\text{Pb}$ .	–
NIST ${}^{204}\text{Pb IA}/{}^{208}\text{Pb IA}$	Ratio of isotopic abundances of ${}^{204}\text{Pb}$ and ${}^{208}\text{Pb}$ in NIST certified reference standard SRM-981.	–
FeOx	Iron(III) oxide.	–
MnOx	Manganese oxide.	–
pM	Negative logarithm of solution metal concentration (with units of $\text{mol L}^{-1}$ ).	–
$\Delta\text{pM}$	Measured pM – modelled pM for a single soil.	–
MBE	Mean bias error, mean measured pM – mean modelled pH.	–
RSD	Residual standard deviation.	–

$\text{Na}_2\text{S}_2\text{O}_4$ ,  $0.3 \text{ mol L}^{-1} \text{ C}_6\text{H}_5\text{Na}_3\text{O}_7$  and  $1 \text{ mol L}^{-1} \text{ NaHCO}_3$  ('DCB' solution), shaken for 24 h at  $20^\circ\text{C}$  (Pansu and Gautheyrou 2006). Extraction suspensions were centrifuged at  $2200\text{g}$  and  $10^\circ\text{C}$  for 15 min, and supernatant solutions filtered to  $< 0.2 \mu\text{m}$  (Filtropur S without prefilter). Samples were analysed by inductively coupled plasma–mass spectrometry (ICP-MS). Concentrations of FeOx and MnOx were calculated from the concentrations of Fe and Mn ( $\text{mol kg}^{-1}$ ) in the DCB extractions and converted into relative concentrations of  $\text{Fe}_2\text{O}_3 \cdot \text{H}_2\text{O}$  and  $\text{MnO}_2$ . Humic and fulvic acids (HAs and FAs) were quantified by an extraction procedure adapted from the typical extraction schemes used to remove and fractionate humic substances (HS) from soils (Tipping 2002; Anderson and Schoenau 2007), but simplified by omitting certain steps (e.g. acid pretreatment to remove inorganic C, N, P and S) as the focus was on quantifying HS rather than extracting them for experimental use. Extraction was done by suspending 1–2 g,  $< 2 \text{ mm}$  soil in 20 mL of  $0.1 \text{ mol L}^{-1} \text{ NaOH}$ , and shaking for 24 h at  $20^\circ\text{C}$  to solubilise the HA and FA fractions. Extraction suspensions

were centrifuged at  $2200\text{g}$  and  $10^\circ\text{C}$  for 15 min, and a 5 mL aliquot of the supernatant stored for analysis of the HA and FA fractions. A 10 mL aliquot of the supernatant was acidified with 1 mL of  $1.5 \text{ mol L}^{-1} \text{ HNO}_3$  and left overnight, allowing HA to precipitate out of solution. After centrifugation at  $2200\text{g}$  and  $10^\circ\text{C}$  for 15 min, the supernatant was stored at  $4^\circ\text{C}$  for analysis of the FA fraction. All samples were analysed for dissolved organic carbon (DOC) content (Shimadzu TOC-VCPH, Total Organic Carbon Analyser, Model TNM-1). Particulate FA and HA concentrations were calculated assuming that FA and HA comprise 50% carbon (Weng *et al.* 2002b).

### Soil metal concentrations

Soil total metal concentrations,  $\{M\}_{\text{total}}$ , were determined following a HF digestion procedure in preference to the aqua regia digestion commonly used, since the latter may not be able to access metal held within silicate matrices. Two replicate ball-milled soil samples ( $\sim 0.2 \text{ g}$ ) were extracted

with concentrated HNO<sub>3</sub> and HClO<sub>4</sub> (2 mL:1 mL) at 80°C for 8 h followed by 100°C for 2 h. After addition of concentrated HF (2.5 mL), samples were then digested at 120°C for 1 h, 140°C for 3 h, 160°C for 4 h and 50°C for 1 h. Samples were left overnight, before addition of concentrated HNO<sub>3</sub> (2.5 mL) and ultrapure water (2.5 mL). Samples were warmed to 50°C for 1 h, and then cooled and diluted to 50 mL with high purity water, and stored prior to multi-element analysis by ICP-MS (X-Series II; Thermo Fisher Scientific, Bremen, Germany).

### Isotopic dilution assays

Enriched stable isotopes, with certified isotopic abundances (IAs) of 97.0% <sup>62</sup>Ni, 99.2% <sup>65</sup>Cu, 99.9% <sup>70</sup>Zn, 70.3% <sup>108</sup>Cd and 99.4% <sup>204</sup>Pb, were obtained as metal foils from Isoflex, San Francisco, CA, USA. The foils were dissolved in HNO<sub>3</sub> to create dilute isotope stock solutions of concentrations 603, 840, 154, 192 and 442 μg mL<sup>-1</sup> for <sup>62</sup>Ni, <sup>65</sup>Cu, <sup>70</sup>Zn, <sup>108</sup>Cd and <sup>204</sup>Pb, respectively. Dilute isotope stock solutions were mixed to create multi-element isotope spike solutions specific to each soil. In general, the multi-element isotope spike solution concentrations were intended to increase the values of the natural abundances of the spike isotopes in the soil by between 10 and 30%, based on a spike volume of 0.4 mL. However, for soils with low total metal concentrations and for practical purposes, it was necessary to sometimes use spike concentrations resulting in values outside this range. In the soils used for modelling in this work, the means (and ranges) of the percentages by which the natural abundance of the spike isotope was raised by spiking were 3.60 (0.235–23.3), 7.35 (1.03–20.5), 0.368 (0.0356–1.72), 10.9 (0.152–42.8) and 1.03 (0.0813–6.16) for Ni, Cu, Zn, Cd and Pb, respectively.

The effect of the isotope on the pH of the soil suspensions may influence the derived *E*-value. Sterckeman et al. (2009) found that the spiking approach had a significant effect on the soil suspension pH and recommended adjusting the pH of the spike solution to a sufficiently high value to minimise the possibility of acidification effects. In contrast, Marzouk et al. (2013) investigated the effect on the soil suspension pH of adding variable spike volumes, using a broadly similar approach to spike composition and concentration for <sup>70</sup>Zn, <sup>111</sup>Cd and <sup>204</sup>Pb isotopes as was taken here. They found that the pH values of soil suspensions, spiked to give 10 and 100% changes in the natural abundance of the labile pools of <sup>70</sup>Zn, <sup>108</sup>Cd and <sup>204</sup>Pb, were not significantly different (*P* < 0.05) from suspensions of unspiked soils. Given the similarity of our spiking approach to that of Marzouk et al. (2013), we have assumed that there was no significant effect on suspension pH due to our procedures.

ID assays required six replicate soil samples (~1 g, < 2 mm). Initially, these were each suspended in 30 mL of 0.01 mol L<sup>-1</sup> Ca(NO<sub>3</sub>)<sub>2</sub> solution and shaken for 72 h at room temperature. Three of the replicate suspensions were then

spiked with 0.4 mL of a multi-element isotope spike solution (<sup>62</sup>Ni, <sup>65</sup>Cu, <sup>70</sup>Zn, <sup>108</sup>Cd, <sup>204</sup>Pb), before all replicates were shaken for a further 72 h. Suspensions were centrifuged at 2200g and 10°C for 15 min and supernatant solutions filtered to < 0.2 μm (Filtropur S without prefilter, Sarstedt) to exclude non-labile metal bound to larger colloids. Samples were acidified to 2% HNO<sub>3</sub> prior to isotope ratio analysis by ICP-MS.

For all the metals, it was found that, in some cases, the concentration of metal solubilised by the electrolyte was so low that the measurements of isotopic ratios required to compute *E*-values could be considered unreliable. The numbers of soils affected were 10, 18, 16, 20 and 27 for Ni, Cu, Zn, Cd and Pb, respectively. Therefore, the assays for these soil/metal combinations were repeated using 10<sup>-5</sup> mol L<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>EDTA as the suspending electrolyte. This matrix has previously been shown to increase the concentration of labile metal in solution, due to complexation of metal ions by EDTA, but to mobilise negligible non-labile metal (Degryse et al. 2004; Atkinson et al. 2011; Izquierdo et al. 2013). For comparative purposes, assays using 10<sup>-5</sup> mol L<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>EDTA were also done on a number of soils for which satisfactory results were obtained using Ca(NO<sub>3</sub>)<sub>2</sub> as suspending electrolyte: 21, 18, 15, 16 and 7 for Ni, Cu, Zn, Cd and Pb, respectively.

Values of {*M*}<sub>*E*</sub> (mol kg<sup>-1</sup>) were determined from the IA of the spike isotope (<sup>s</sup>IA) and a reference isotope (<sup>r</sup>IA) measured in three solutions: the isotope spike solution (IA<sub>sp</sub>), the solution phase of the spiked suspension (IA<sub>sp-sus</sub>) and the solution phase of an un-spiked control suspension (IA<sub>sus</sub>). Values of {*M*}<sub>*E*</sub> were calculated from Eqn 1 (Heumann 1988; Gäbler et al. 2007):

$$\{M\}_E = \left( \frac{AM_{\text{soil}}}{W} \right) \left( \frac{10^3 C_{\text{sp}} V_{\text{sp}}}{AM_{\text{sp}}} \right) \left( \frac{{}^s\text{IA}_{\text{sp}} - \left( \frac{{}^s\text{IA}_{\text{sp-sus}}}{{}^r\text{IA}_{\text{sp-sus}}} \right) {}^r\text{IA}_{\text{sp}}}{\left( \frac{{}^s\text{IA}_{\text{sp-sus}}}{{}^r\text{IA}_{\text{sp-sus}}} \right) {}^r\text{IA}_{\text{sus}} - {}^s\text{IA}_{\text{sus}}} \right) \quad (1)$$

where AM<sub>soil</sub> and AM<sub>sp</sub> are the average atomic masses of the metal in the un-spiked control soil and the isotope spike respectively (g mol<sup>-1</sup>), C<sub>sp</sub> and V<sub>sp</sub> are the concentration (g L<sup>-1</sup>) and volume (L) of the isotope spike respectively, and *W* is the mass of the soil (kg).

The IAs of the isotopes in the spiked and un-spiked soil solutions were calculated from the measured isotope ratios, and the natural abundance ratios of the unmeasured isotopes for each metal. For Pb, however, natural abundance ratios were not used, because lead isotope abundance varies depending on the source of the lead in the soil. For this reason, the isotope ratios for all Pb isotope combinations were measured. The isotope ratios <sup>62</sup>Ni/<sup>60</sup>Ni, <sup>65</sup>Cu/<sup>63</sup>Cu, <sup>70</sup>Zn/<sup>66</sup>Zn, <sup>108</sup>Cd/<sup>111</sup>Cd, <sup>204</sup>Pb/<sup>208</sup>Pb, <sup>206</sup>Pb/<sup>208</sup>Pb and <sup>207</sup>Pb/<sup>208</sup>Pb, were measured by quadrupole ICP-MS, operating in CCT-KED (collision cell technology with kinetic energy discrimination) mode, particularly to minimise interference from the chlorine dimer (<sup>35</sup>Cl-<sup>35</sup>Cl) in some



extractions which interferes with measurements of  $^{70}\text{Zn}$ . The isobaric interference of  $^{204}\text{Hg}$  on  $^{204}\text{Pb}$  was corrected by assay of  $^{202}\text{Hg}$ , although this was a very small adjustment. Where necessary, samples were diluted to ensure measurement by the electron multiplier detector in pulse-counting mode. The ions of each isotope of a metal are transported with different efficiencies within the ICP-MS due to their different mass to charge ratios, and this effect is referred to as mass discrimination (MD). Corrections for MD were calculated from measured count rate ratios of ICP-MS calibration standards for Ni, Cu, Zn and Cd and the certified reference standard NIST SRM-981 for Pb. An example for Pb is given in Eqn 2 (Marzouk *et al.* 2013). The standards were run approximately every 20 samples and MD factors implemented as a drift correction.

$$\text{Correction factor} = \frac{\text{measured } ^{204}\text{Pb cps} / ^{208}\text{Pb cps}}{\text{NIST } ^{204}\text{Pb IA} / ^{208}\text{Pb IA}} \quad (2)$$

### Additional measurements for modelling

Soil suspension pH ( $\text{pH}_{\text{Ca}}$ ) was determined in the un-spiked  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions ( $\sim 1 \text{ g}$ ,  $< 2 \text{ mm}$ /30 mL) after 144 h shaking, using a pH meter and combined pH electrode (Radiometer Analytical SAS). The pH electrode was calibrated using pH 1, 4 and 7 buffer solutions (Fixanal, Fluka Analytical) at  $20^\circ\text{C}$ . Solution concentrations of Ni, Cu, Zn, Cd and Pb were also measured in these solutions, for the purpose of testing the model outputs. DOC concentrations were measured on filtered, unacidified aliquots of the un-spiked  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions, after 144 h shaking, using a total organic carbon analyser (Shimadzu TOC-V CPH, Model TNM-1).

### WHAM/Model VII application

Default parameter settings and input files were used for the speciation modelling. The soil concentration was set equal to the soil/solution ratio ( $33 \text{ g L}^{-1}$ ), and the pH to the  $\text{pH}_{\text{Ca}}$  value. The total Ni, Cu, Zn, Cd and Pb concentrations were set to the measured values of either  $\{M\}_{\text{total}}$ ,  $\{M\}_{\text{EDTA}}$  or  $\{M\}_{\text{E}}$ . Soil binding phases used in the computations were particulate HA, FA, MnOx and FeOx. Measured concentrations of particulate HA, FA and DCB-extractable Mn were used to compute the phase concentrations. The concentration of particulate FeOx was computed by inputting the DCB-extractable concentration of iron as the total  $\text{Fe}^{\text{III}}$  and allowing FeOx to precipitate from this iron pool as a particulate solid with a chemically active surface. The solution electrolyte was represented by inputting concentrations of dissolved Ca and  $\text{NO}_3$  ( $0.01$  and  $0.02 \text{ mol L}^{-1}$ , respectively). DOM in the suspensions was represented by colloidal FA concentrations, calculated from the DOC concentrations assuming that 65% of the measured DOC was 'active' FA in respect to binding, and that FA is 50% C (Weng *et al.* 2002b). The temperature was set to  $293 \text{ K}$  ( $20^\circ\text{C}$ ) and for computing carbonate chemistry, the atmospheric partial pressure of  $\text{CO}_2$  was set to  $0.0004$ . The variables used for model application are summarised in Table 2.

A check on the possibility of mineral control of metal solubility was undertaken by separate speciation of the soil supernatants, using the measured solution metal concentrations as input, and calculating saturation indices for a range of hydroxide and carbonate-containing minerals.

An alternative parameterisation of the model was also run, using an assumption that the SOM comprised 50% HA.

**Table 2.** Input soil variables for WHAM/Model VII.

Input parameter	Phase	Unit	Notes
$\{M\}_{\text{E}}$ for Ni, Cu, Zn, Cd, Pb	Solute. Total concentration.	$\text{mol L}^{-1}$	
$\{M\}_{\text{EDTA}}$ for Ni, Cu, Zn, Cd and Pb	Solute. Total concentration.	$\text{mol L}^{-1}$	
Ca, $\text{NO}_3$	Solute. Total concentration.	$\text{mol L}^{-1}$	From background electrolyte.
Fulvic acid	Colloidal. <sup>A</sup>	$\text{g L}^{-1}$	From measured [DOC].
Soil humic acid concentration	Particulate. <sup>B</sup>	$\text{g L}^{-1}$	
Soil fulvic acid concentration	Particulate. <sup>B</sup>	$\text{g L}^{-1}$	
DCB-extractable Fe	Solute. Total concentration.	$\text{g L}^{-1}$	Allowed to precipitate to form particulate iron(III) oxide <sup>B</sup> .
Mn oxide concentration	Particulate. <sup>B</sup>	$\text{g L}^{-1}$	Computed from the concentration of DCB-extractable Mn.
Suspended particulate matter	–	$\text{g L}^{-1}$	Calculated from the suspension soil to solution ratio 1 g/30 mL.
Temperature	–	K	293 K
Atmospheric partial pressure of $\text{CO}_2$	–	–	0.0004
Soil suspension pH	–	–	

<sup>A</sup>Solutes computed to be bound to a colloidal phase are considered to be dissolved.

<sup>B</sup>Solutes computed to be bound to a particulate phase are considered to be particulate. Particulate humic and fulvic acids, and oxides, represent the soil binding phases.

This assumption was previously used by Buekers *et al.* (2008) and Marzouk *et al.* (2013).

The WHAM outputs used were (1) the total aqueous concentrations of Ni, Cu, Zn, Cd and Pb ( $\text{mol L}^{-1}$ ), which were compared to the concentrations measured in solution (section *Additional measurements for modelling*), and (2) the proportion of each metal predicted to be bound to each particulate and colloidal phase.

## POSSMs application

POSSMs was developed (Lofts 2022) to provide a model for the equilibrium soil–solution partitioning of Ni, Cu, Zn, Cd and Pb with minimal data requirements and thus suitable for application at large scale, where the relatively detailed data requirements of WHAM or other geochemical speciation models might not be achievable. The model represents the system of chemical equilibria acting on metals in soils via two ‘lumped’ equilibria: those between the free and adsorbed metal, and between the free and metal complexed in the soil solution. These equilibria are expressed as

$$\{M\}_{\text{ads}} = K_{\text{ads}} [M^{2+}]^A a_{\text{H}}^B \{\text{POM}\}^{\Delta}, \text{ and} \quad (3)$$

$$[M]_{\text{comp,aq}} = K_{\text{comp}} [M^{2+}]^{\alpha} a_{\text{H}}^{\beta} [\text{DOM}]^{\delta}. \quad (4)$$

The terms  $K_{\text{ads}}$ ,  $K_{\text{comp}}$ ,  $A$ ,  $B$ ,  $\Delta$ ,  $\alpha$ ,  $\beta$  and  $\delta$  are empirical parameters obtained by directly fitting the model to data of matched soil solution pH, solid phase and DOM, and geochemically active, soil solution and free metal concentrations. Lofts (2022) parameterised the model by fitting it to datasets where porewaters were extracted from intact soils prior to characterisation, in order to provide the best possible analogue to field porewaters. Where free ion in the porewaters was not measured directly, it was obtained by modelling using WHAM/Model VII. Best fit values of the parameters are provided in Table 3.

## Results

### General soil characteristics

Of the 109 soils studied, 84 were successfully analysed for all the determinants required to perform the desired speciation for all five metals. To allow robust comparison of modelling

outcomes across the metals and modelling approaches, only the results for these soils are presented and analysed here.

Graphical summaries of the ranges of pH, soil LOI and [DOC] (in the  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  suspensions) are shown in Supplementary Fig. S1. Table 4 shows summary statistics for the total and EDTA-extractable metal concentrations, and the  $E$ -values. Soil metal concentrations span several orders of magnitude, reflecting the wide range of soil characteristics, from uncontaminated rural soils to those impacted by urban, industrial and mining activity. Values of  $\{M\}_{\text{total}}$ ,  $\{M\}_{\text{EDTA}}$  and  $\{M\}_E$  for each soil are shown in the Supplementary material (Supplementary Table S1).

### Humic and fulvic acid contents

Supplementary Fig. S2 shows the proportion of SOM (as LOI) present as HS, and the measured ratio of HA to FA. The proportion of HS ranged widely, from below  $0.05 \text{ g g}^{-1}$  to approximately  $0.5 \text{ g g}^{-1}$ . There was no clear relationship between the humic content of the SOM and the LOI, although when the humic content is plotted against  $\text{pH}_{\text{Ca}}$  a trend to lower humic content with higher  $\text{pH}_{\text{Ca}}$  can be seen (correlation coefficient 0.55,  $P < 0.001$ ). The proportion of the base-extractable organic matter (the sum of HA and FA) that is humic in nature increases with increasing SOM, but in a non-linear manner. Correlation between the HA proportion and the logarithm of the LOI has a coefficient of 0.44 ( $P < 0.001$ ). Conversely, there was no clear relationship between the HA proportion and  $\text{pH}_{\text{Ca}}$ , other than a tendency to a higher proportion on average when  $\text{pH}_{\text{Ca}} < 4$ .

### Comparison of $\{M\}_{\text{EDTA}}$ and $\{M\}_E$ as geochemically reactive metal pools

Figs 1 and 2 show the ratios of metal extracted by EDTA in each soil to the corresponding  $E$ -value, as a function of pH in the  $\text{Ca}(\text{NO}_3)_2$  suspension. Standard errors in ratios were computed by bootstrapping using the standard deviations of measured  $\{M\}_E$  and  $\{M\}_{\text{EDTA}}$  concentrations and a sample size of 2000. Median relative standard errors (RSEs) were all 3%, with the exception of Pb for which it was 4%. Maximum RSEs were 54, 25, 15, 21 and 32% for Ni, Cu, Zn, Cd and Pb, respectively.

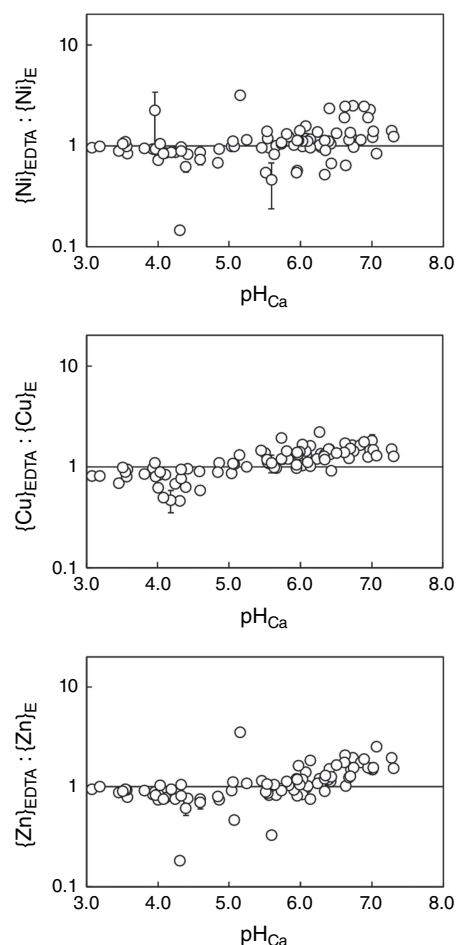
Values of  $\{M\}_{\text{EDTA}}$  and  $\{M\}_E$  represent two estimates of the labile or ‘geochemically active’ metal – by wet chemical

**Table 3.** Metal binding parameters in the POSSMs model.

Metal	$\log K_{\text{comp}}$	$\alpha$	$\beta$	$\delta$	$\log K_{\text{ads}}$	$A$	$B$	$\Delta$
Ni	−1.58	1.00	−0.52	0.98	−1.36	1.00	−0.50	1.28
Cu	−3.80	0.62	−0.60	0.75	−2.90	1.00	−1.02	0.97
Zn	−2.60	0.80	−0.39	0.75	−1.97	1.00	−0.49	0.96
Cd	−1.21	1.00	−0.34	0.70	−1.63	1.00	−0.47	1.08
Pb	−2.90	0.88	−0.83	1.19	−2.72	1.00	−1.05	0.60

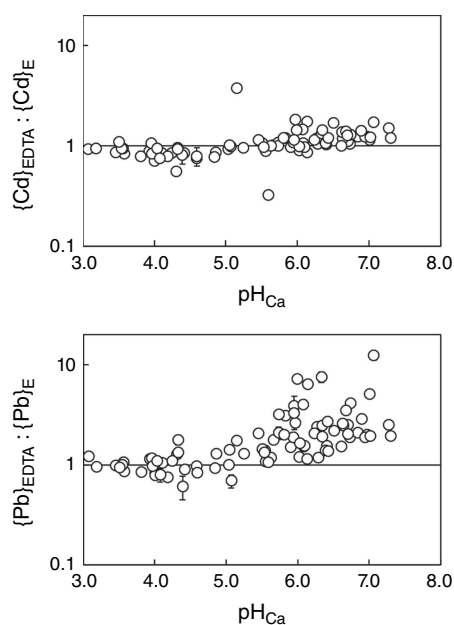
**Table 4.** Variable means, medians and interquartile ranges for  $\{M\}_{\text{total}}$ ,  $\{M\}_{\text{EDTA}}$  and  $\{M\}_E$ . All values are  $\log_{10}$ .

	Ni ( $\text{mol kg}^{-1}$ )			Cu ( $\text{mol kg}^{-1}$ )			Zn ( $\text{mol kg}^{-1}$ )			Cd ( $\text{mol kg}^{-1}$ )			Pb ( $\text{mol kg}^{-1}$ )		
	$\{M\}_{\text{total}}$	$\{M\}_{\text{EDTA}}$	$\{M\}_E$	$\{M\}_{\text{total}}$	$\{M\}_{\text{EDTA}}$	$\{M\}_E$	$\{M\}_{\text{total}}$	$\{M\}_{\text{EDTA}}$	$\{M\}_E$	$\{M\}_{\text{total}}$	$\{M\}_{\text{EDTA}}$	$\{M\}_E$	$\{M\}_{\text{total}}$	$\{M\}_{\text{EDTA}}$	$\{M\}_E$
Geomean	-3.29	-4.40	-4.39	-3.03	-3.69	-2.64	-2.30	-3.11	-3.09	-4.93	-5.25	-5.23	-2.94	-3.48	-3.26
Median	-3.36	-4.61	-4.56	-3.09	-3.74	-3.58	-2.36	-3.15	-3.11	-5.06	-5.33	-5.30	-2.98	-3.45	-3.21
IQR	-3.26	-4.41	-4.55	-2.72	-3.28	-3.17	-1.96	-2.68	-2.62	-4.75	-5.06	-5.03	-2.64	-3.19	-2.92
Min	-4.25	-5.77	-5.76	-4.11	-5.15	-5.28	-3.43	-4.81	-5.02	-5.97	-7.20	-7.29	-3.98	-5.05	-4.74
Max	-1.17	-2.03	-2.02	-0.81	-1.65	-1.69	-0.74	-1.11	-1.04	-2.96	-3.11	-3.12	-0.86	-1.61	-1.40

**Fig. 1.** Values of  $\{M\}_{\text{EDTA}}$  as a proportion of  $\{M\}_E$  as a function of pH in  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$ , for Ni, Cu and Zn. The horizontal line represents equal values of  $\{M\}_{\text{EDTA}}$  and  $\{M\}_E$ . Error bars represent  $\pm 1$  standard error.

extraction and by ID analysis. Values of  $\{M\}_E$  expressed as a proportion of  $\{M\}_{\text{total}}$  ranged from 0.01–0.64 for Ni, 0.02–0.65 for Cu, 0.01–1.08 for Zn, 0.01–1.08 for Cd and 0.03–1.16 for Pb. These ranges exceed the literature values compiled by Izquierdo *et al.* (2013) in contaminated flood plain soils of 0.10–0.33 for Zn, 0.03–1.00 for Cd and 0.31–0.78 for Pb. They also exceed those reported by Mao *et al.* (2017) for urban soils in two UK cities.

On average, Cd was most isotopically labile and Ni least; the mean  $\{M\}_E:\{M\}_{\text{total}}$  ratios were 0.12, 0.25, 0.22, 0.54 and 0.35 for Ni, Cu, Zn, Cd and Pb, respectively. Pairwise *t*-testing showed that only the means for Cu and Zn were not significantly different ( $P < 0.05$ ). Supplementary Fig. S3 shows the distribution of individual ratios for each metal; for Zn, Cd and Pb there is considerable variation in individual ratios, covering almost the entire range of possible values. This is in line with results previously reported by Marzouk *et al.* (2013) and Izquierdo *et al.* (2013), who found that  $\{M\}_E:\{M\}_{\text{total}}$  ratios decreased in the order Cd > Pb > Zn for sets of 246 minespoil-contaminated soils



**Fig. 2.** Values of  $\{M\}_{EDTA}$  as a proportion of  $\{M\}_E$  as a function of pH in  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$ , for Cd and Pb. The horizontal line represents equal values of  $\{M\}_{EDTA}$  and  $\{M\}_E$ . Error bars represent  $\pm 1$  standard error.

and 27 alluvial soils respectively. Supplementary Fig. S4 further compares the mean ratios found here with those found from other studies in UK soils. Although the mean values vary among the studies, there is a consistent pattern to the relative ratios:  $\text{Cd} > \text{Pb} > \text{Cu} \sim \text{Zn} > \text{Ni}$ .

The  $E$ -values were found to exceed  $\{M\}_{\text{Total}}$  concentrations for Pb in soil 65, Zn and Cd in soil 72, and Cd and Pb in soil 86. All these soils were acidic ( $\text{pH}_{\text{Ca}} < 5$ ) with relatively high organic matter contents ( $> 0.24 \text{ g g}^{-1}$  LOI). Previous research on fixation of labile Zn and Cd in soils has shown that the extent of fixation is relatively low in acidic soils (Degryse et al. 2004; Marzouk et al. 2013). It is likely that such  $E$ -values reflect near complete lability of the soil metal coupled with natural heterogeneity of the sampled soil.

Values of  $\{M\}_{EDTA}$  and  $\{M\}_E$  were highly correlated (Table 5). On average, the ratio of  $\{M\}_{EDTA}$  to  $\{M\}_E$  was 1.03, 1.16, 1.22, 1.17 and 1.48 for Ni, Cu, Zn, Cd and Pb, respectively. Overall, the  $\text{Na}_2\text{H}_2\text{EDTA}$  extraction provided a reasonable match to the more conceptually precise isotopically exchangeable pool. For Ni, Cu, Zn and Cd the match between  $\{M\}_{EDTA}$  and  $\{M\}_E$  was reasonable but for Pb the  $\{M\}_{EDTA}$  values were more clearly biased to higher values. Investigation of the relationship between the  $\{M\}_{EDTA}:\{M\}_E$  ratio and  $\text{pH}_{\text{Ca}}$  (Figs 1 and 2) showed that, in general, ratios for Ni, Cu, Zn and Cd were below unity when  $\text{pH}_{\text{Ca}} \leq 5.0$  (0.90, 0.80, 0.82 and 0.86 respectively) and over unity when  $\text{pH}_{\text{Ca}} > 5.0$  (1.23, 1.33, 1.29 and 1.20 respectively). The mean  $\{\text{Pb}\}_{EDTA}:\{\text{Pb}\}_E$  ratio was 1.02 when  $\text{pH}_{\text{Ca}} \leq 5.0$  and 2.51 when  $\text{pH}_{\text{Ca}} > 5.0$ . A single soil (Port Talbot 1) showed a  $\{\text{Pb}\}_{EDTA}:\{\text{Pb}\}_E$  ratio of over 10.

**Table 5.** Pearson correlation coefficients and regression equations for  $0.05 \text{ mol L}^{-1}$  EDTA-extracted concentrations of Ni, Cu, Zn, Cd and Pb ( $\{M\}_{EDTA}$ ;  $\text{mol kg}^{-1}$ ) and  $E$ -values ( $\{M\}_E$ ;  $\text{mol kg}^{-1}$ ). All correlation coefficients are significant ( $P < 0.001$ ).

Metal	Pearson correlation coefficient	Regression equation
Ni	0.996	$\{M\}_{EDTA} = 1.04 \cdot \{M\}_E - 5.63 \times 10^{-6}$
Cu	0.976	$\{M\}_{EDTA} = 0.99 \cdot \{M\}_E + 0.000143$
Zn	0.994	$\{M\}_{EDTA} = 1.19 \cdot \{M\}_E + 0.000121$
Cd	0.978	$\{M\}_{EDTA} = 1.06 \cdot \{M\}_E + 3.44 \times 10^{-6}$
Pb	0.988	$\{M\}_{EDTA} = 1.56 \cdot \{M\}_E - 4.87 \times 10^{-5}$

Overall, it was found that if the non-labile metal pool is smaller at low pH, there is less likelihood of a difference between  $\{M\}_{EDTA}$  and  $\{M\}_E$ . However, some exceptions were found. For example, Ni, Zn and Cd  $\{M\}_{EDTA}$  and  $\{M\}_E$  in the soils Wemyss Mine Horizon 1 and Wemyss Mine Horizon 2, with  $\text{pH}_{\text{Ca}}$  5.6 and 5.1 respectively, showed notable contrasts. The  $E$ -values for Ni, Zn and Cd for the Horizon 1 soil (the upper horizon) were underestimated by the  $\text{Na}_2\text{H}_2\text{EDTA}$  extraction ( $\{M\}_{EDTA}:\{M\}_E$  ratios were 0.45, 0.33 and 0.32 of  $\{M\}_E$  for Ni, Zn and Cd respectively). In contrast, the  $E$ -values for Horizon 2 (the lower horizon) were overestimated by the  $\text{Na}_2\text{H}_2\text{EDTA}$  extraction ( $\{M\}_{EDTA}:\{M\}_E$  ratios were 3.16, 3.53 and 3.76 for Ni, Zn and Cd respectively). Although the underestimation of  $E$ -values by the EDTA extraction was unexpected, a similar finding was made for Pb in acidic organic soils by Marzouk et al. (2013). It may also be that the particular history of such a site, being in the vicinity of an abandoned Pb/Zn mine, has resulted in metals occurring in specific forms that produce the unusual result seen.

No significant relationships were found between the  $\{M\}_{EDTA}:\{M\}_E$  ratio and LOI or DCB-extractable Mn or Fe, for any of the metals (data not shown).

## WHAM/Model VII modelling

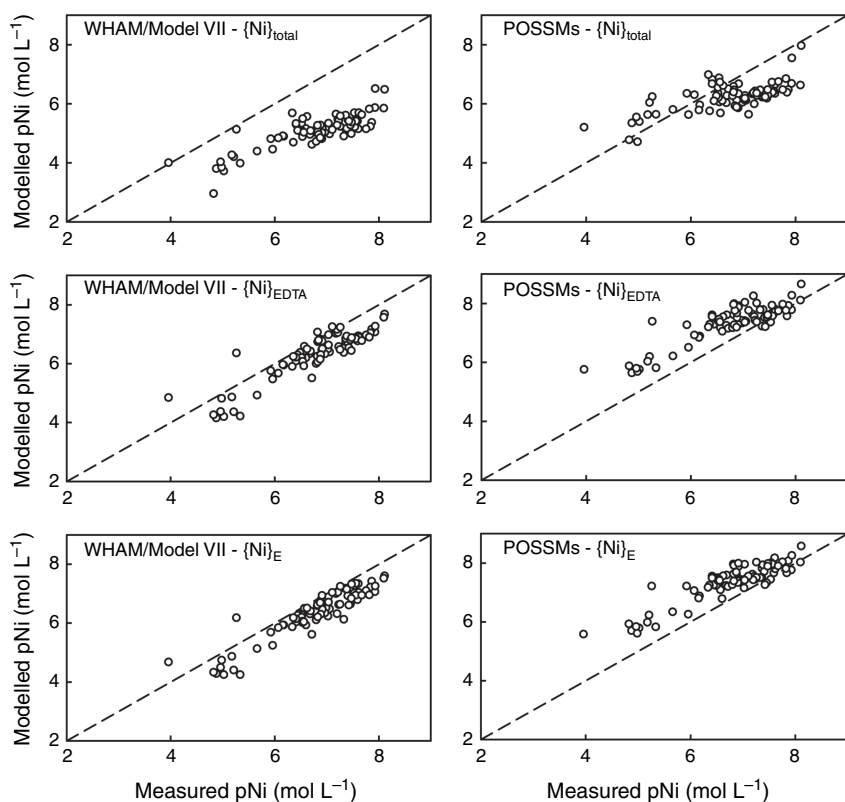
The mineral solubility control check showed that predicted concentrations of metals in the supernatants were all below saturation with respect to any of the mineral phases considered (Supplementary Table S3).

Comparisons were made using WHAM/Model VII, where values of  $\{M\}_{\text{total}}$ ,  $\{M\}_{EDTA}$  and  $\{M\}_E$  were used as the reactive solute concentrations for modelling purposes. Outputs from both models were compared to the measured total soluble Ni, Cu, Zn, Cd and Pb concentrations in the unspiked  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions. Goodness of fit parameters ( $r^2$  values, RMSD and mean bias error (MBE, mean measured pM – mean modelled pM)) are shown in Table 6, and comparisons of observations and predictions are shown in Figs 3–7. When using  $\{M\}_{\text{total}}$  as the estimate

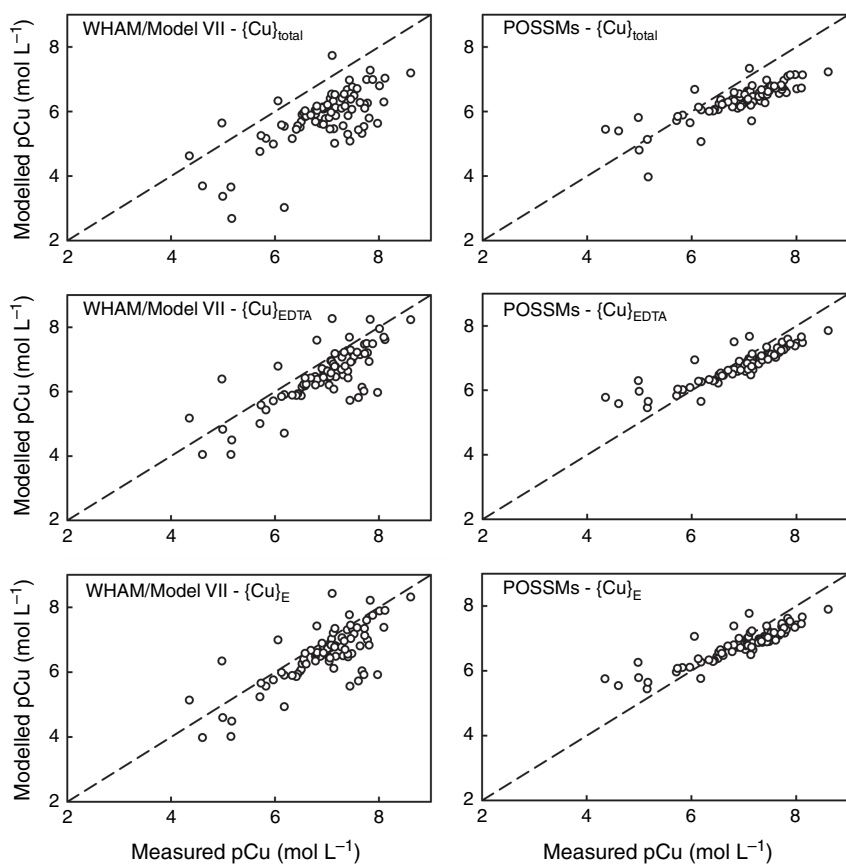


**Table 6.** Goodness of prediction parameters ( $r^2$ , residual standard deviation RSD, mean bias error MBE) for modelled metal solution concentrations pM ( $\text{mol L}^{-1}$ ;  $-\log_{10}$  scale), using WHAM/Model VII and POSSMs, compared to measured metal solution concentrations pM ( $\text{mol L}^{-1}$ ;  $-\log_{10}$  scale). MBE = mean of (measured pM – modelled pM).

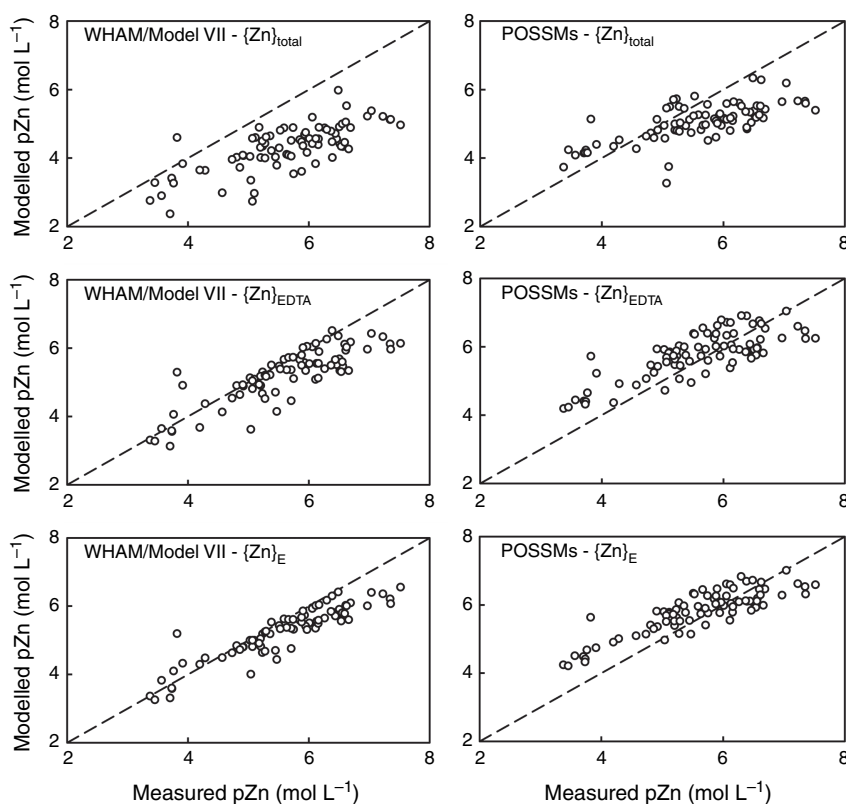
	WHAM/Model VII measured HA, FA			WHAM/Model VII HA = 50% SOM			POSSMs		
	$r^2$	RSD	MBE	$r^2$	RSD	MBE	$r^2$	RSD	MBE
{Ni} <sub>total</sub>	0.68	1.76	1.69	0.70	1.56	1.48	0.48	0.82	0.53
{Ni} <sub>EDTA</sub>	0.81	0.59	0.46	0.86	0.39	0.23	0.75	0.72	-0.57
{Ni} <sub>E</sub>	0.85	0.53	0.42	0.88	0.36	0.20	0.80	0.70	-0.58
{Cu} <sub>total</sub>	0.50	1.27	1.10	0.48	1.07	0.84	0.69	0.75	0.58
{Cu} <sub>EDTA</sub>	0.61	0.70	0.44	0.58	0.60	0.16	0.80	0.45	0.16
{Cu} <sub>E</sub>	0.57	0.70	0.39	0.53	0.64	0.10	0.80	0.45	0.14
{Zn} <sub>total</sub>	0.55	1.48	1.33	0.64	1.20	1.05	0.49	0.91	0.57
{Zn} <sub>EDTA</sub>	0.71	0.66	0.40	0.80	0.46	0.14	0.61	0.63	-0.17
{Zn} <sub>E</sub>	0.84	0.55	0.36	0.87	0.38	0.10	0.75	0.56	-0.19
{Cd} <sub>total</sub>	0.67	1.00	0.88	0.78	0.76	0.66	0.66	0.47	-0.10
{Cd} <sub>EDTA</sub>	0.74	0.61	0.45	0.84	0.39	0.23	0.72	0.58	-0.40
{Cd} <sub>E</sub>	0.84	0.52	0.41	0.90	0.31	0.19	0.82	0.54	-0.42
{Pb} <sub>total</sub>	0.73	1.13	0.87	0.73	1.11	0.84	0.56	1.10	0.66
{Pb} <sub>EDTA</sub>	0.72	0.85	0.40	0.72	0.84	0.39	0.60	0.94	0.37
{Pb} <sub>E</sub>	0.75	0.80	0.13	0.75	0.80	0.12	0.76	0.74	0.17



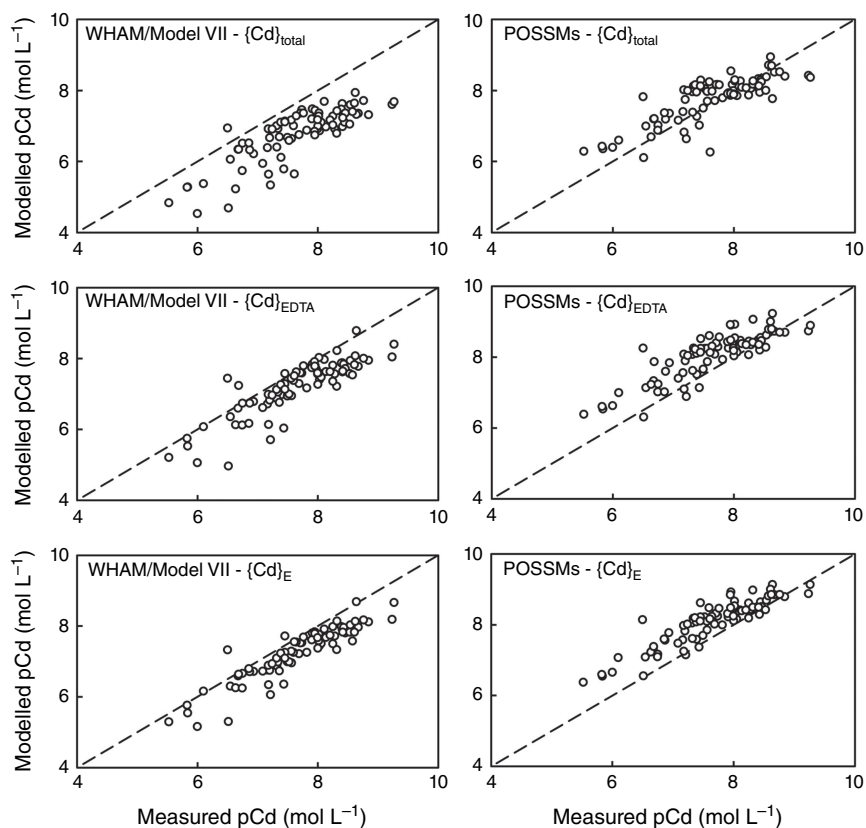
**Fig. 3.** Comparison of measured Ni solution concentrations (pNi,  $-\log_{10}$  of the total solution Ni in  $\text{mol L}^{-1}$ ) in  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions with model outputs using  $\{\text{Ni}\}_{\text{total}}$ ,  $\{\text{Ni}\}_{\text{EDTA}}$  and  $\{\text{Ni}\}_{\text{E}}$  as model inputs. Left: predictions using WHAM/Model VII, right: predictions using POSSMs. The dashed lines indicate the 1:1 line.



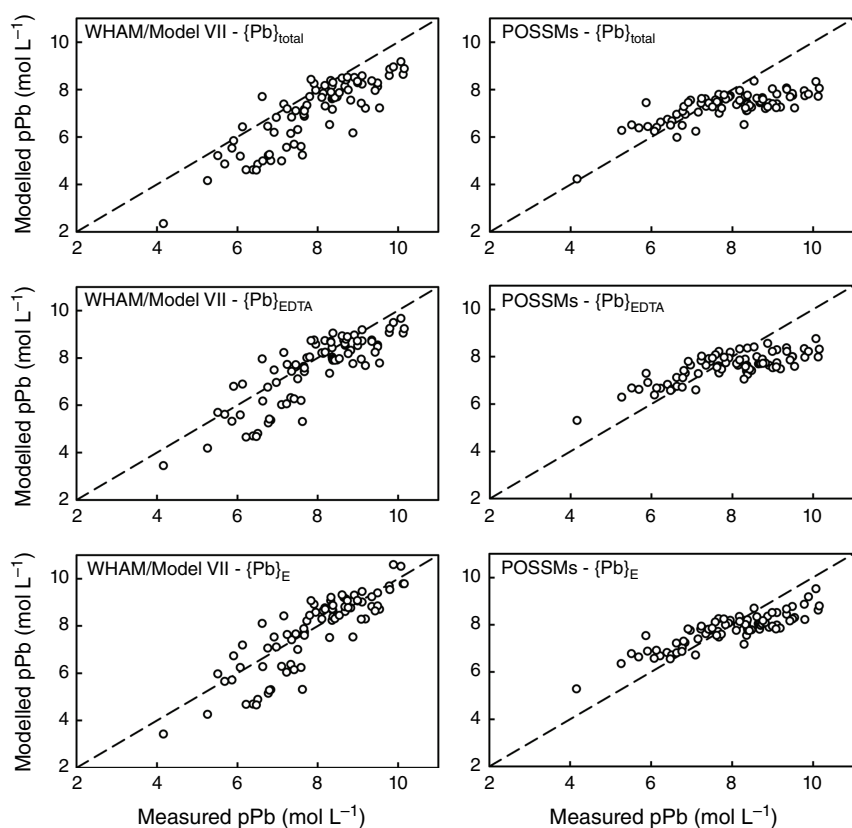
**Fig. 4.** Comparison of measured Cu solution concentrations ( $pCu$ ,  $-\log_{10}$  of the total solution Cu in  $\text{mol L}^{-1}$ ) in  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions with model outputs using  $\{\text{Cu}\}_{\text{total}}$ ,  $\{\text{Cu}\}_E$  and  $\{\text{Cu}\}_{\text{EDTA}}$  as model inputs. Left: predictions using WHAM/Model VII, right: predictions using POSSMs. The dashed lines indicate the 1:1 line.



**Fig. 5.** Comparison of measured Zn solution concentrations ( $pZn$ ,  $-\log_{10}$  of the total solution Zn in  $\text{mol L}^{-1}$ ) in  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions with model outputs using  $\{\text{Zn}\}_{\text{total}}$ ,  $\{\text{Zn}\}_E$  and  $\{\text{Zn}\}_{\text{EDTA}}$  as model inputs. Left: predictions using WHAM/Model VII, right: predictions using POSSMs. The dashed lines indicate the 1:1 line.



**Fig. 6.** Comparison of measured Cd solution concentrations ( $pCd$ ,  $-\log_{10}$  of the total solution Cd in  $\text{mol L}^{-1}$ ) in  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions with model outputs using  $\{Cd\}_{\text{total}}$ ,  $\{Cd\}_{\text{E}}$  and  $\{Cd\}_{\text{EDTA}}$  as model inputs. Left: predictions using WHAM/Model VII, right: predictions using POSSMs. The dashed lines indicate the 1:1 line.



**Fig. 7.** Comparison of measured Pb solution concentrations ( $pPb$ ,  $-\log_{10}$  of the total solution Pb in  $\text{mol L}^{-1}$ ) in  $0.01 \text{ mol L}^{-1} \text{ Ca}(\text{NO}_3)_2$  soil suspensions with model outputs using  $\{Pb\}_{\text{total}}$ ,  $\{Pb\}_{\text{E}}$  and  $\{Pb\}_{\text{EDTA}}$  as model inputs. Left: predictions using WHAM/Model VII, right: predictions using POSSMs. The dashed lines indicate the 1:1 line.

of geochemically reactive soil metal, both models generally overestimated metal solubility ( $MBE > 0$ ), with the exception of POSSMs for Cd where a small underestimation was seen on average. By contrast, when using  $\{M\}_{EDTA}$  or  $\{M\}_E$ , there was a consistent shift towards smaller values of MBE. This is to be expected given that  $\{M\}_{EDTA}$  and  $\{M\}_E$  were, on average, smaller than  $\{M\}_{total}$  for each soil.

The WHAM/Model VII predictions using  $\{M\}_{EDTA}$  or  $\{M\}_E$  all showed positive MBE, i.e. overestimation of metal solubility on average. Generally, the MBE values obtained from WHAM/Model VII were rather similar when using either  $\{M\}_{EDTA}$  or  $\{M\}_E$  as input, though there was a consistent tendency to slightly smaller MBE when using  $\{M\}_E$ . The largest difference in MBE was seen for Pb, which may be related to the differences seen between  $\{M\}_{EDTA}$  and  $\{M\}_E$  for this metal at the higher end of the soil  $pH_{Ca}$  range. Overall goodness-of-prediction, quantified by the RMSD value, showed similar trends to MBE, with the use of  $\{M\}_{EDTA}$  or  $\{M\}_E$  as input generally resulting in lower RMSDs, with the RMSD when using  $\{M\}_E$  generally the same or lower than the RMSD when using  $\{M\}_{EDTA}$ . The exception was Pb, where all the RMSD values were similar. Patterns in  $r^2$  were similar, with the highest  $r^2$  for all the metals except Cu being found when using the  $\{M\}_E$ . Supplementary Fig. S5 shows the patterns in observed and modelled metal solubility as a function of  $pH_{Ca}$ , using the model with  $\{M\}_E$  as input. WHAM/Model VII reproduced the general trends in solubility well using the measured soil HA and FA concentrations, albeit with a general tendency to overestimate the solubility of Ni, Zn and Cd across the  $pH_{Ca}$  range, particularly in the most acidic soils ( $pH_{Ca} < 5.5$ ). Both the relatively weak trend in Cu solubility and the strong trend in Pb solubility were reproduced, however, for both metals the model predicted near complete solubility below  $pH_{Ca} \sim 3.75$ , which overestimated observed solubility by over an order of magnitude in most cases.

The alternative parameterisation of the model, where active soil HA was set to 50% of LOI and active FA not used, increased the modelled HA + FA concentration in the soils by a factor of 2.3 on average, with only five soils having lower HA + FA than under the standard parameterisation. The outcome was consistently lower RSD and MBE values than were found when using the measured HA and FA concentrations (Table 6), with the exception of Pb for which the RSD and MBE under both assumptions were essentially identical. Supplementary Fig. S6 shows the predictions made using the alternative parameterisation and values of  $\{M\}_E$  as input. Supplementary Fig. S5 shows that the improvements in model prediction when assuming soil HA to be 50% of LOI became larger at higher  $pH_{Ca}$  values, and that in the most acidic soils (lowest  $pH_{Ca}$  quintile, 3.08–4.24), the improvement in prediction was small to negligible.

Adjusting the activity of DOM generally had a smaller effect on model goodness-of-prediction. For example,

halving active DOM decreases MBE by 0.01 for Ni and had no effect on RSD, while doubling active DOM increased both RSD and MBE by small amounts (RSD from 0.53 to 0.55 and MBE from 0.41 to 0.43). Similar patterns were found for Zn and Cd. A larger response was seen for Cu: halving DOM activity decreased RSD slightly (0.70–0.68) but reduced MBE by about a half (0.39–0.19), while doubling active DOM increased RSD (0.70–0.80) and MBE (0.39 and 0.60). Changing DOM activity had little effect on the RSD for Pb (0.80–0.84 when halving DOM activity and to 0.81 when doubling it). The MBE changed from 0.13 to  $-0.02$  on halving DOM activity.

Supplementary Fig. S7 shows the predicted distribution of each metal among the solution and the individual solid phases (SOM, iron oxide and manganese oxide) when modelling using values of  $\{M\}_E$  as input. For Cu, Zn and Cd, there was a general trend of increasing adsorption with increasing pH, with organic matter generally being the dominant adsorption phase. The metal oxides become more significant as adsorption phases as pH increases and, for Zn and Cd, iron oxide is the most important adsorbing phase in soils with  $pH_{Ca}$  above  $\sim 6.6$ . Manganese oxide is a relatively unimportant contributor to binding, particularly for Cu, and tended to be most important in soils of  $pH_{Ca} \sim 5.5$ – $6.0$ , becoming less important at higher pH. Somewhat different trends were seen for Ni and Pb. Soil adsorption of Ni exhibited a maximum in the  $pH_{Ca}$  range  $\sim 5.5$ – $6.0$ , in contrast to the other metals. Manganese oxide was also predicted to be a relatively important contributor to Ni adsorption, comparable to that of HA and FA, with iron oxide playing a relatively minor role. Lead adsorption was dominated by the metal oxides, and binding to HA and FA played only a minor role. The relative importance of iron and manganese oxides as Pb sorbents is predicted to vary considerably with pH, with manganese oxide being most important up to a  $pH_{Ca}$  of  $\sim 6.0$ , above which iron oxide is predicted to play an increasingly important role and to dominate adsorption above  $pH_{Ca} \sim 6.6$ .

As the soil pH is a key variable influencing most chemical equilibria, the relationship between  $pH_{Ca}$  and prediction bias ( $\Delta pM$ , measured  $pM$  – modelled  $pM$ ) was investigated further. When  $\{M\}_E$  was used as the input to WHAM/Model VII,  $\Delta pM$  for Ni, Zn and Cd had significant positive correlations with  $pH_{Ca}$ , with Pearson correlation coefficients of 0.26 ( $P = 0.015$ ), 0.47 ( $P < 0.001$ ) and 0.44 ( $P < 0.001$ ) respectively. Copper and lead were found to have significant negative correlations (Pearson correlation coefficients  $-0.51$  and  $-0.52$ , both  $P < 0.001$ ). This indicates a systematic tendency for solubility in soils at lower pH to be overestimated to a greater extent than at higher pH. Calculation of Cu and Pb MBE values for the  $pH_{Ca}$  quintiles shows that there was considerable overestimation of solubility in the lowest quintile (3.08–4.24) (MBE for Cu was 1.08 and for Pb was 1.23). Across the remaining  $pH_{Ca}$  range, MBE for Cu was 0.24 and for Pb was  $-0.15$ . [Buekers et al.](#)

(2008) noted that overestimation of Ni, Cu, Zn and Cd solubility occurred mainly for soils with  $\text{pH} > 6$  and suggested that this may result from an underestimation of sorption on oxyhydroxides, which becomes increasingly important as  $\text{pH}$  increases (Supplementary Fig. S7). Previously, predictions of Pb solubility have been found to be  $\text{pH}$  dependent, with over prediction at low  $\text{pH}$  and under predictions at high  $\text{pH}$  (Marzouk *et al.* 2013).

Similar correlation patterns were found between  $\text{pH}$  and  $\{M\}_{\text{EDTA}}$  concentrations when used as inputs to WHAM/Model VII. For Ni, Zn and Cd, significant positive correlations between  $\Delta\text{pM}$  and  $\text{pH}_{\text{Ca}}$  were found, with coefficients of 0.45, 0.60 and 0.54 respectively ( $P < 0.001$  for all). Copper and lead  $\Delta\text{pM}$  values again had significant negative correlations with  $\text{pH}_{\text{Ca}}$  (correlation coefficients  $-0.35$  and  $-0.30$  respectively,  $P = 0.001$  and  $0.006$  respectively). The Cu and Pb correlations were again driven by overestimation of metal solubility in the lowest  $\text{pH}_{\text{Ca}}$  quintile, with the MBE for Cu being 0.98 and for Pb 1.21, compared with MBEs of 0.30 and 0.20 for the other quintiles.

When the alternative model parameterisation was used, clear changes were seen in the relationship between model bias and  $\text{pH}_{\text{Ca}}$ . Pearson correlation coefficients between  $\Delta\text{pM}$  and  $\text{pH}_{\text{Ca}}$  for Ni, Zn and Cd changed to  $-0.19$  ( $P = 0.088$ ),  $0.11$  ( $P = 0.332$ ) and  $0.00$  ( $P = 0.979$ ); thus, the alternative parameterisation did not result in significant relationships between model bias and  $\text{pH}_{\text{Ca}}$  for these metals. Supplementary Fig. S5 shows that the predicted decrease in metal solubility caused by the change in model assumptions did not occur for the soils in the lowest  $\text{pH}_{\text{Ca}}$  quintile. The change in assumptions thus not only reduced MBE for the whole dataset, but made the model bias more consistent across the whole  $\text{pH}_{\text{Ca}}$  range.

## POSSMs modelling

POSSMs gave broadly similar goodness-of-prediction results compared to WHAM/Model VII, yet there were some notable differences. In particular, when using either  $\{M\}_{\text{EDTA}}$  or  $\{M\}_{\text{E}}$  as input, dissolved Ni and Cd were underestimated on average across the range of  $\text{pH}_{\text{Ca}}$  quintiles, while WHAM/Model VII overestimated solubility on average. When using  $\{M\}_{\text{E}}$  as input, goodness-of-prediction (RSD, Table 6) for Ni was inferior to that for WHAM/Model VII, while for cadmium it was similar. Goodness-of-prediction and MBE for copper were superior to that of WHAM/Model VII, while for zinc RSD was similar, with a lower MBE for POSSMs. Goodness-of-prediction and MBE for lead were similar, with RSD being marginally lower for POSSMs and MBE being marginally higher.

As with the WHAM/Model VII results, we investigated further the relationship between prediction bias ( $\Delta\text{pM}$ ) and  $\text{pH}_{\text{Ca}}$  for POSSMs. When  $\{M\}_{\text{E}}$  was used as the input,  $\Delta\text{pM}$  for Ni, Zn, Cd and Pb had significant positive correlations ( $P < 0.001$ ) with  $\text{pH}_{\text{Ca}}$ , with Pearson correlation coefficients

of 0.56, 0.66, 0.63 and 0.69 respectively. Copper showed a weaker negative correlation coefficient of  $-0.23$  ( $P = 0.034$ ). These trends can also be seen by inspecting Supplementary Fig. S7. With the exception of Cu, POSSMs underestimated the trend in average solubility across the quintiles of  $\text{pH}_{\text{Ca}}$ . This was particularly notable in the case of Pb, where the model underestimated solubility on average in the two lowest  $\text{pH}_{\text{Ca}}$  quintiles, but then overestimated it at higher  $\text{pH}_{\text{Ca}}$ . This  $\text{pH}_{\text{Ca}}$ -dependent bias in predictions can also be seen in the overall solubility prediction (Fig. 7), since there was a strong  $\text{pH}$  dependence of absolute Pb solubility. So, despite the fact that overall goodness-of-prediction measures for Pb were quite similar for POSSMs and WHAM/Model VII, a more detailed look at how goodness-of-prediction varies showed clear patterns that suggest areas in which the model could be improved; in this case, to better describe the dependence of solubility on  $\text{pH}$ . Correlation patterns and significance when using  $\{M\}_{\text{EDTA}}$  as input were very similar, with the exception that the correlation for Cu was not significant ( $P = 0.53$ ).

## Discussion

The approach to assessing and modelling metal solubility in this study is broadly similar to that taken previously by a number of authors (e.g. Izquierdo *et al.* 2013; Marzouk *et al.* 2013; Mao *et al.* 2017), but we have used a range of soils covering a broader spectrum of land use and potential contamination sources. The range of  $\text{pH}_{\text{Ca}}$  values in our study (3.08–7.31) is broader than those of Izquierdo *et al.* (2013) (5.3–8.0) and Mao *et al.* (2017) (3.6–8.1), as is the range of SOM contents (this study  $0.028$ – $0.939 \text{ g g}^{-1}$ ; Izquierdo *et al.* (2013)  $0.04$ – $0.18 \text{ g g}^{-1}$ , Mao *et al.* (2017)  $0.020$ – $0.502 \text{ g g}^{-1}$ ). The ranges of  $\text{pH}_{\text{Ca}}$  and SOM in this study are similar to those of Marzouk *et al.* (2013) (2.58–7.58 and  $0.034$ – $0.961 \text{ g g}^{-1}$ , respectively), however, this study covers a far wider range of locations, land use and metal source types than Marzouk *et al.* (2013), which focused on a single mining-impacted upland catchment.

The proportion of metal in labile form in a given soil is a complex function of factors including contamination history, soil chemistry and mineralogy and metal properties. Factors suggested to influence rates of fixation of labile metal to non-labile forms include the ionic radius (Degryse *et al.* 2009), solid-phase diffusion into metal (hydr)oxides (Bruemmer *et al.* 1988; Trivedi and Axe 2001) or carbonates (Hamon *et al.* 2002; Buekers *et al.* 2007), surface precipitation (Ma *et al.* 2006), substitution into iron (oxy)hydroxides (Manceau *et al.* 2000) or fixation into natural organic matter (Mao *et al.* 2015). Processes may be metal-specific, such as the formation of surface precipitates, or they may be favoured by intrinsic metal properties, such as the role of ionic radius in substitution reactions (Xu *et al.* 2006), which favours metals having smaller ionic radii (Ni, Cu, Zn) over those with larger radii (Cd, Pb). Thus, it is not surprising,



and highly consistent with previous research, to see both high variation in the labile proportions of individual metals and patterns in the degree of mean lability.

Our results suggest that in general the  $0.05 \text{ mol L}^{-1}$  EDTA extraction is a reasonable analogue for isotopically labile metal, although we can see systematic differences between determinations as a function of pH. These differences may arise due to the ability of EDTA to promote ligand-induced dissolution of soil solid phases such as metal oxides and carbonates (Izquierdo et al. 2013). Other researchers have found somewhat different results. Izquierdo et al. (2013) performed isotopic lability measurements and  $0.05 \text{ mol L}^{-1}$  EDTA extractions for Zn, Cd and Pb on floodplain soils of central England, and found a concentration-dependent relationship between  $\{\text{Pb}\}_{\text{EDTA}}$  and  $\{\text{Pb}\}_{\text{E}}$ , with the majority of their soils having EDTA-extractable Pb three to four times higher than the isotopically determined labile Pb. We do not see such a relationship in this work. This is likely due to the broader range of soil types used, which causes any concentration relationship to be swamped by the relationship with pH (Fig. 2). Marzouk et al. (2013) compared isotopic lability and chemical extractability of Zn, Cd and Pb for soils of a mining-impacted catchment in Northern England. They found that while there was reasonable correspondence between isotopically labile metal and metal extracted by  $0.05 \text{ mol L}^{-1}$  EDTA or  $0.43 \text{ mol L}^{-1}$   $\text{HNO}_3$  in acidic organic soils, the extractions overestimated the isotopically labile pool in circumneutral and alkaline soils. Similarly, Garforth et al. (2016) found that extraction with  $0.05 \text{ mol L}^{-1}$  EDTA provided the best agreement with isotopically labile pools of Ni, Cu, Zn, Cd and Pb in four UK soils, compared with  $0.43 \text{ mol L}^{-1}$   $\text{HNO}_3$ ,  $0.43 \text{ mol L}^{-1}$  acetic acid and  $1 \text{ mol L}^{-1}$   $\text{CaCl}_2$ .

The prediction of metal solubilities by WHAM/Model VII clearly demonstrates that the use of total metal as a speciation model input for prediction of metal solubility is inferior to the use of a measure of labile or 'geochemically active' metal. This provides further confirmation of a result already observed by a number of authors (Izquierdo et al. 2013; Marzouk et al. 2013). It also demonstrates that although good predictions of metal solubility are obtained using  $0.05 \text{ mol L}^{-1}$  EDTA-extracted metal as the model input (RSD < 1 for all the metals), the use of isotopically labile metal as input consistently provides predictions that are at least as good, and usually better, than those obtained using EDTA. This is in agreement with the results previously found by Izquierdo et al. (2013) for Zn and Cd in a more restricted set of soils, although these authors did find a lower MBE and similar RSD for Pb when using  $0.05 \text{ mol L}^{-1}$  EDTA. In contrast, we obtained a notable improvement in MBE from 0.40 when using EDTA-extracted Pb, to 0.13 when using isotopically labile Pb. This contrast may well be due to the better agreement between  $\{\text{Pb}\}_{\text{EDTA}}$  and  $\{\text{Pb}\}_{\text{E}}$  in soils where  $\text{pH}_{\text{Ca}} < 5.0$ . In soils in this study with  $\text{pH}_{\text{Ca}} < 5.0$  ( $n = 26$ ) the mean absolute difference in predicted pPb between the

two methods of estimating labile metal was 0.0056, whereas for soils of  $\text{pH}_{\text{Ca}} > 5.0$  it was 0.39.

Comparing the goodness-of-prediction values obtained from the speciation modelling (using  $\{M\}_{\text{E}}$  as input) to those obtained in previous similar studies may be useful in providing clues as to the causes of model error. The most similar previous study to this one is that of Mao et al. (2017), conducted on urban soils. The RSD values obtained in this study were similar to those of Mao et al. (2017) for Ni, superior for Zn and Cd but inferior for Cu and Pb. Conversely, in comparison to Marzouk et al. (2013) this study shows a superior prediction for Pb but inferior predictions for Zn and Cd, while for Izquierdo et al. (2013) this study is superior for Zn and Cd but inferior for Pb. So there is little apparent consistency in the RSD values, which possibly indicates that the RSD is dependent on the nature of the soils within each separate dataset.

Patterns in the mean bias error (MBE) across studies may be indicative of inaccurate formulation of model assumptions, or bias in one or more supporting measurements. In this study we observed a consistent positive MBE for all metals, indicating overestimation of metal solubility on average. The same pattern was observed by Mao et al. (2017), with the exception of Pb, and by Izquierdo et al. (2013) for Zn, Cd and Pb. Marzouk et al. (2013) also observed this pattern for Cd and Pb but found a small overestimation of Zn solubility on average. There is thus a reasonably consistent pattern across studies, most notably between this study and that of Mao et al. (2017). We investigated the possibility of systematic model bias by performing additional WHAM/Model VII simulations with the soil HA concentration set to 50% of LOI, after Marzouk et al. (2013). The results show that the model outcomes are somewhat sensitive to the concentration of HA (and FA) representing SOM, and that setting HA and FA to measured concentrations tends to underestimate solid phase binding for Ni, Cu, Zn and Cd. The reasons for this warrant further examination. A possibility is that the base extraction method used for the determination of soil HA and FA might be underestimating the amount of 'active' HS, possibly because the extraction is not efficient enough to remove the humics from the soil matrix in a single step. Alternatively, there may be components of the SOM active in metal binding that are not extractable using base. Izquierdo et al. (2013) suggested that overestimation may be due to a proportion of the soil metal being adsorbed to soil carbonates, a process not considered in WHAM/Model VII. This was because the soils they worked with were sourced from a limestone-based catchment. However, metal adsorption to carbonate surfaces is unlikely to be important for many of the soils in the current study given their provenance. Interestingly, the apparent underestimation of the 'activity' of particulate organic matter was also seen by Lofts and Tipping (1998), when using base extraction to estimate the HS content of riverine suspended sediment for modelling using WHAM/Model V.

The influence of DOM activity on the predictions was generally more metal-dependent than the influence of SOM activity. The influence was broadly related to the strength of metal–DOM binding, with the relatively weakly DOM-binding metals (Ni, Zn and Cd) being less sensitive to varying the DOM activity than the relatively strongly binding Cu and Pb. The use of a single value of ‘DOM activity’ to estimate concentrations of ‘active’ FA for modelling is a clear simplification. Previous research (Amery *et al.* 2008; Djae *et al.* 2017; Ouédraogo *et al.* 2022) has shown that fitting the DOM activity on an individual soil basis results in a broad range of activities (as low as 10% and as high as 215% in the cited studies, compared to the range of 32.5–130% used for sensitivity testing in this work). Currently, no single approach to estimating DOM activity, based on ancillary quantifications of DOM quality such as fluorescence indices or specific UV absorbance at 254 nm (SUVA<sub>254</sub>), appears to be appropriate for estimating activity. For example, Amery *et al.* (2008) found that SUVA<sub>254</sub> variations could improve the prediction of DOM-bound Cu in soil solutions, but Ouédraogo *et al.* (2022) could not improve the prediction of free copper using either SUVA<sub>254</sub> or fluorescence indices.

An important element of studies such as ours is the choice of the binding surface assemblage to be used in simulation of the solid phase. Other researchers (Bonten *et al.* 2008; Dijkstra *et al.* 2009) have used alternative assemblages including clay and aluminium oxide, which we have not used here. Their use of aluminium oxide was limited by lack of data to the assumption of identical binding parameters as for iron oxide. The results suggested that clay might be important for Zn and Cd binding in soils of particularly high clay content, but it is not possible to tease out the possible importance of aluminium oxide from their results. Nevertheless, the choice of solid phase assemblage components remains key to studies of trace element solubility in soil. A standard approach to soil characterisation for solubility modelling would be a way forward here, allowing a higher degree of data interoperability across studies and thus of model application and comparison. Key here is the fact that other trace elements, particularly those that occur as anions, may have different preferences for binding surfaces to the metals studied here.

Compared to the other metals, predictions of Pb solubility resulted in the highest RSD value (Table 6). These results are in contrast to those of Izquierdo *et al.* (2013) who reported that accuracy of predictions decreased in the order Pb > Zn > Cd. However, weaker predictions of Pb solubility compared to Zn and Cd have previously been reported by Marzouk *et al.* (2013). They suggested that this may be due to Pb binding in the solid phase being more distributed between organic matter and Fe and Mn oxides, compared to Zn and Cd which were predominantly adsorbed to organic matter (Supplementary Fig. S7). The predictions of Pb solubility are therefore sensitive to uncertainties in a greater

number of measurements than the other metals, which may well explain the greater scatter seen in the WHAM/Model VII predictions.

The POSSMs model is intended to provide a means of computing a simplified soil metal speciation with practical minimum input requirements. Therefore, in contrast to WHAM, there is relatively little scope for the application of different assumptions regarding factors such as SOM and DOM activities. The parameter sets used by the model were derived by Lofts (2022) from datasets of British and Dutch soils. The labile metal in the British soils was estimated by extraction with 0.1 mol L<sup>-1</sup> EDTA, and in the Dutch soils by extraction with 2 mol L<sup>-1</sup> HNO<sub>3</sub>. The solution phase was obtained by extraction of porewater following soil incubation; thus the soil/solution ratios were far higher (up to approximately two orders of magnitude) than the ratios employed in this work. The porewater pH range was 3.35–8.28, the SOM range was 0.004–0.98 g g<sup>-1</sup> and the DOM range was 0.0106–0.942 g L<sup>-1</sup>. The pH and DOM ranges of the dataset used in this study extend slightly below those of the training dataset (3.08 and 0.0069 g L<sup>-1</sup> respectively) but the SOM range is within that of the training dataset. The range of land use is narrower in the training dataset, comprising semi-natural UK soils and Dutch soils from an agricultural field experiment.

Overall, the results show that POSSMs provides a useful and plausible alternative approach for modelling metal solubility compared to the more complex and data-intensive WHAM/Model VII and similar models. It may provide a useful alternative approach for situations where the data required for more complex models are not available, for example when there is a need to model solubility using large scale soil geochemical datasets. Although the dataset produced in this study covered a similar range of soil conditions to the POSSMs parameterisation dataset, the parameterisation dataset used reactive metal concentrations obtained by extraction, as opposed to *E*-values. In this context, the quality of the predictions obtained here when using {*M*}<sub>*E*</sub> is encouraging, as are the small differences in predictions obtained when using either {*M*}<sub>EDTA</sub> or the *E*-value as input. The systematic bias in the predictions seen for some metals suggests that broader testing and ultimately reparameterisation of the model to the largest possible dataset is desirable. Parameterisation to data containing metal *E*-values is desirable, however, the results found here suggest that it may have little influence on predictions with the possible exception of Pb.

## Conclusions

- We applied two geochemical speciation models, WHAM/Model VII and POSSMs, to model the solubility of Ni, Cu, Zn, Cd and Pb in a broad range of soils from the UK, using

- $\{M\}_{\text{total}}$ ,  $\{M\}_{\text{EDTA}}$  and  $\{M\}_E$  to represent the labile pool of metal controlling solubility.
- In agreement with previous work, representing the labile metal by the total measured pool resulted in overestimation of metal solubility on average.
  - Estimation of labile soil metal by EDTA extraction provided concentrations largely comparable to the isotopically exchangeable pool, although some systematic deviations were observed, particularly for Pb.
  - Both EDTA-extracted and isotopically exchangeable metal pools provided reasonable predictions of metal solubility using WHAM/Model VII with measured concentrations of soil humic substances. Predictions were generally improved by applying the assumption that soil organic matter comprised 50% humic acid. This suggests that further testing of the method for estimating soil humic substance concentrations may be warranted.
  - The relatively simple POSSMs model provided predictions comparable to WHAM/Model VII, further indicating its potential usefulness for large scale application.
  - Quantification of the variability in the ‘activity’ of the dissolved organic matter in the solution phase may further improve the predictive capabilities of the models.
- ## Supplementary material
- Details of the datasets and a range of additional figures are provided in the Supplementary material [online](#).
- ## References
- Ahnstrom ZAS, Parker DR (2001) Cadmium reactivity in metal-contaminated soils using a coupled stable isotope dilution-sequential extraction procedure. *Environmental Science & Technology* **35**, 121–126. doi:10.1021/es001350o
- Almås ÅR, Lofts S, Mulder J, Tipping E (2007) Solubility of major cations and Cu, Zn and Cd in soil extracts of some contaminated agricultural soils near a zinc smelter in Norway: modelling with a multisurface extension of WHAM. *European Journal of Soil Science* **58**, 1074–1086. doi:10.1111/j.1365-2389.2007.00894.x
- Amery F, Degryse F, Cheyns K, De Troyer I, Mertens J, Merckx R, Smolders E (2008) The UV-absorbance of dissolved organic matter predicts the fivefold variation in its affinity for mobilizing Cu in an agricultural soil horizon. *European Journal of Soil Science* **59**, 1087–1095. doi:10.1111/j.1365-2389.2008.01078.x
- Anderson DW, Schoenau JJ (2007) Soil humus fractions. In ‘Soil sampling and methods of analysis’, 2nd edn. (Eds MR Carter, EG Gregorich) pp. 675–680. (CRC Press: Boca Raton, FL, USA)
- Atkinson NR (2010) Heavy metal geochemistry of contaminated fenland soils in NW England. PhD Thesis, University of Nottingham, Nottingham, England. Available at <https://eprints.nottingham.ac.uk/id/eprint/27795>
- Atkinson NR, Bailey EH, Tye AM, Breward N, Young SD (2011) Fractionation of lead in soil by isotopic dilution and sequential extraction. *Environmental Chemistry* **8**, 493–500. doi:10.1071/EN11020
- Ayoub AS, McGaw BA, Shand CA, Midwood AJ (2003) Phytoavailability of Cd and Zn in soil estimated by stable isotope exchange and chemical extraction. *Plant and Soil* **252**, 291–300. doi:10.1023/A:1024785201942
- Bonten LTC, Groenenberg JE, Weng L, van Riemsdijk WH (2008) Use of speciation and complexation models to estimate heavy metal sorption in soils. *Geoderma* **146**, 303–310. doi:10.1016/j.geoderma.2008.06.005
- Brummer GW, Gerth J, Tiller KG (1988) Reaction kinetics of the adsorption and desorption of nickel, zinc and cadmium by goethite: I. Adsorption and diffusion of metals. *Journal of Soil Science* **39**, 37–52. doi:10.1111/j.1365-2389.1988.tb01192.x
- Buekers J, Van Laer L, Amery F, Van Buggenhout S, Maes A, Smolders E (2007) Role of soil constituents in fixation of soluble Zn, Cu, Ni and Cd added to soils. *European Journal of Soil Science* **58**, 1514–1524. doi:10.1111/j.1365-2389.2007.00958.x
- Buekers J, Degryse F, Maes A, Smolders E (2008) Modelling the effects of ageing on Cd, Zn, Ni and Cu solubility in soils using an assemblage model. *European Journal of Soil Science* **59**, 1160–1170. doi:10.1111/j.1365-2389.2008.01053.x
- Degryse F, Buekers J, Smolders E (2004) Radio-labile cadmium and zinc in soils as affected by pH and source of contamination. *European Journal of Soil Science* **55**, 113–122. doi:10.1046/j.1351-0754.2003.0554.x
- Degryse F, Smolders E, Parker DR (2009) Partitioning of metals (Cd, Co, Cu, Ni, Pb, Zn) in soils: concepts, methodologies, prediction and applications – a review. *European Journal of Soil Science* **60**, 590–612. doi:10.1111/j.1365-2389.2009.01142.x
- Dijkstra JJ, Meeussen JCL, Comans RN (2004) Leaching of heavy metals from contaminated soils: an experimental and modeling study. *Environmental Science & Technology* **38**, 4390–4395. doi:10.1021/es049885v
- Dijkstra JJ, Meeussen JCL, Comans RNJ (2009) Evaluation of a Generic Multisurface Sorption Model for Inorganic Soil Contaminants. *Environmental Science & Technology* **43**, 6196–6201. doi:10.1021/es900555g
- Djaj T, Bravin MN, Garnier C, Doelsch E (2017) Parameterizing the binding properties of dissolved organic matter with default values skews the prediction of copper solution speciation and ecotoxicity in soil. *Environmental Toxicology and Chemistry* **36**, 898–905. doi:10.1002/etc.3622
- Gäbler HE, Bahr A, Heidkamp A, Utermann J (2007) Enriched stable isotopes for determining the isotopically exchangeable element content in soils. *European Journal of Soil Science* **58**, 746–757. doi:10.1111/j.1365-2389.2006.00863.x
- Garforth JM, Bailey EH, Tye AM, Young SD, Lofts S (2016) Using isotopic dilution to assess chemical extraction of labile Ni, Cu, Zn, Cd and Pb in soils. *Chemosphere* **155**, 534–541. doi:10.1016/j.chemosphere.2016.04.096
- Groenenberg JE, Dijkstra JJ, Bonten LTC, de Vries W, Comans ENJ (2012) Evaluation of the performance and limitations of empirical partition–relations and process–based multisurface models to predict trace element solubility in soils. *Environmental Pollution* **166**, 98–107. doi:10.1016/j.envpol.2012.03.011
- Gustafsson JP (2001) Modeling the acid–base properties and metal complexation of humic substances with the Stockholm Humic Model. *Journal of Colloid and Interface Science* **244**, 102112. doi:10.1006/jcis.2001.7871
- Hamon RE, McLaughlin MJ, Cozens G (2002) Mechanisms of attenuation of metal availability in in situ remediation treatments. *Environmental Science & Technology* **36**, 3991–3996. doi:10.1021/es025558g
- Heumann K (1988) Isotope dilution mass spectrometry. In ‘Inorganic mass spectrometry’. (Eds F Adams, R Gijbels, R Van Grieken) pp. 301–376. (Wiley and Sons)
- Izquierdo M, Tye AM, Chenery SR (2013) Lability, solubility and speciation of Cd, Pb and Zn in alluvial soils of the River Trent catchment UK. *Environmental Science: Processes & Impacts* **15**, 1844–1858. doi:10.1039/C3EM00370A
- Karam A (1993) Chemical properties of organic soils. In ‘Soil sampling and chemical analysis’. (Ed. MR Carter) pp. 459–472. (Lewis Publishers: Boca Raton, FL, USA)
- Keizer MG, van Riemsdijk WH (2009) ‘ECOSAT: A computer program for the calculation of speciation and transport in soil–water systems’. p. 81. (Department of Soil Quality, Wageningen University: Wageningen, Netherlands)
- Kinniburgh DG, Milne CJ, Benedetti MF, Pinheiro JP, Filius J, Koopal LK, Vanriemsdijk WH (1996) Metal ion binding by humic acid: application of the NICA-Donnan model. *Environmental Science & Technology* **30**, 1687–1698. doi:10.1021/es950695h
- Lofts S (2022) POSSMs: a parsimonious speciation model for metals in soils. *Environmental Chemistry* **18**, 335–351. doi:10.1071/EN21100



- Lofts S, Tipping E (1998) An assemblage model for cation binding by natural particulate matter. *Geochimica et Cosmochimica Acta* **62**, 2609–2625. doi:10.1016/S0016-7037(98)00183-5
- Ma Y, Lombi E, Nolan AL, McLaughlin MJ (2006) Short-term natural attenuation of copper in soils: effects of time, temperature, and soil characteristics. *Environmental Toxicology and Chemistry* **25**, 652–658. doi:10.1897/04-601r.1
- Manceau A, Schlegel ML, Musso M, Sole VA, Gauthier C, Petit PE, Trolard F (2000) Crystal chemistry of trace elements in natural and synthetic goethite. *Geochimica et Cosmochimica Acta* **64**, 3643–3661. doi:10.1016/S0016-7037(00)00427-0
- Mao LC, Young SD, Bailey EH (2015) Lability of copper bound to humic acid. *Chemosphere* **131**, 201–208. doi:10.1016/j.chemosphere.2015.03.035
- Mao LC, Young SD, Tye AM, Bailey EH (2017) Predicting trace metal solubility and fractionation in Urban soils from isotopic exchangeability. *Environmental Pollution* **231**, 1529–1542. doi:10.1016/j.envpol.2017.09.013
- Marzouk ER, Chenery SR, Young SD (2013) Measuring reactive metal in soil: a comparison of multi-element isotopic dilution and chemical extraction. *European Journal of Soil Science* **64**, 526–536. doi:10.1111/ejss.12043
- Meeussen JCL (2003) ORCHESTRA: An object-oriented framework for implementing chemical equilibrium models. *Environmental Science & Technology* **37**, 1175–1182. doi:10.1021/es025597s
- Nolan AL, Ma YB, Lombi E, McLaughlin MJ (2004) Measurement of labile Cu in soil using stable isotope dilution and isotope ratio analysis by ICP-MS. *Analytical and Bioanalytical Chemistry* **380**, 789–797. doi:10.1007/s00216-004-2816-6
- Ouédraogo F, Cornu J-Y, Janot N, Nguyen C, Sourzac M, Parlanti E, Denaix L (2022) Do DOM optical parameters improve the prediction of copper availability in vineyard soils? *Environmental Science and Pollution Research* **29**, 29268–29284. doi:10.1007/s11356-021-16361-5
- Pansu M, Gautheyrou J (2006) ‘Handbook of soil analysis: mineralogical, organic and inorganic methods.’ (Springer: Berlin, Heidelberg, Germany)
- Quevauviller P (1998) Operationally defined extraction procedures for soil and sediment analysis I. Standardization. *Trends in Analytical Chemistry* **17**, 289–298. doi:10.1016/S0165-9936(97)00119-2
- Rawlins BG, McGrath SP, Scheib AJ, Breward N, Cave M, Lister TR, Ingham M, Gowing C, Carter S (2012) ‘The advanced soil geochemical atlas of England and Wales.’ (British Geological Survey: Keyworth, Nottingham, UK)
- Schröder TJ, Hiemstra T, Vink JPM, van der Zee SEATM (2005) Modeling of the solid-solution partitioning of heavy metals and arsenic in embanked flood plain soils of the rivers Rhine and Meuse. *Environmental Science & Technology* **39**, 7176–7184. doi:10.1021/es048154s
- Sterckeman T, Carignan J, Srayeddin I, Baize D, Cloquet C (2009) Availability of soil cadmium using stable and radioactive isotope dilution. *Geoderma* **153**, 372–378. doi:10.1016/j.geoderma.2009.08.026
- Thornton I, Farago ME, Thums CR, Parrish RR, McGill RAR, Breward N, Fortey NJ, Simpson P, Young SD, Tye AM, Crout NMJ, Hough RL, Watt J (2008) Urban geochemistry: research strategies to assist risk assessment and remediation of brownfield sites in urban areas. *Environmental Geochemistry and Health* **30**, 565–576. doi:10.1007/s10653-008-9182-9
- Tipping E (1998) Humic Ion-Binding Model VI: an improved description of the interactions of protons and metal ions with humic substances. *Aquatic Geochemistry* **4**, 3–47. doi:10.1023/A:1009627214459
- Tipping E (2002) ‘Cation binding by humic substances.’ (Cambridge University Press: Cambridge, UK)
- Tipping E, Hurley MA (1992) A unifying model of cation binding by humic substances. *Geochimica et Cosmochimica Acta* **56**, 3627–3641. doi:10.1016/0016-7037(92)90158-F
- Tipping E, Rieuwerts J, Pan G, Ashmore MR, Lofts S, Hill MTR, Farago ME, Thornton I (2003) The solid-solution partitioning of heavy metals (Cu, Zn, Cd, Pb) in upland soils of England and Wales. *Environmental Pollution* **125**, 213–225. doi:10.1016/S0269-7491(03)00058-7
- Tipping E, Lofts S, Sonke JE (2011) Humic Ion-Binding Model VII: a revised parameterisation of cation binding by humic substances. *Environmental Chemistry* **8**, 225–235. doi:10.1071/EN11016
- Trivedi P, Axe L (2001) Predicting divalent metal sorption to hydrous Al, Fe, and Mn oxides. *Environmental Science & Technology* **35**, 1779–1784. doi:10.1021/es001644+
- Tye AM, Young SD, Crout NMJ, Zhang H, Preston S, Barbosa-Jefferson VL, Davison W, McGrath SP, Paton GI, Kilham K, Resende L (2003) Predicting the activity of Cd<sup>2+</sup> and Zn<sup>2+</sup> in soil pore water from the radio-labile metal fraction. *Geochimica et Cosmochimica Acta* **67**, 375–385. doi:10.1016/S0016-7037(02)01138-9
- Weng LP, Temminghoff EJM, van Riemsdijk WH (2001) Contribution of individual sorbents to the control of heavy metal activity in sandy soil. *Environmental Science & Technology* **35**, 4436–4443. doi:10.1021/es010085j
- Weng LP, Temminghoff EJM, Lofts S, Tipping E, van Riemsdijk WH (2002a) Complexation with dissolved organic matter and solubility control of heavy metals in a sandy soil. *Environmental Science & Technology* **36**, 4804–4810. doi:10.1021/es0200084
- Weng LP, Fest EPMJ, Filus J, Temminghoff EJM, van Riemsdijk WH (2002b) Transport of humic and fulvic acids in relation to metal mobility in a copper-contaminated acid sandy soil. *Environmental Science & Technology* **36**, 1699–1704. doi:10.1021/es010283a
- Xu Y, Boonfueng T, Axe L, Maeng S, Tyson T (2006) Surface complexation of Pb(II) on amorphous iron oxide and manganese oxide: spectroscopic and time studies. *Journal of Colloid and Interface Science* **299**, 28–40. doi:10.1016/j.jcis.2006.01.041

**Data availability.** The data used in this study may be obtained by emailing Stephen Lofts (stlo@ceh.ac.uk).

**Conflicts of interest.** Stephen Lofts is a Guest Editor for the Special Issue of *Environmental Chemistry* in which this paper appears. To mitigate this potential conflict of interest they were blinded from the review process. The authors declare no other conflicts of interest.

**Declaration of funding.** JG was funded under a NERC PhD studentship award.

**Acknowledgements.** The authors dedicate this paper to Professor Edward Tipping with grateful thanks for his major scientific contributions to the field of soil metal chemistry modelling, and for his generous and continuing advice, mentorship, scientific collaboration and friendship over many years. They also thank three anonymous reviewers for their constructive and useful comments on the initial version of the manuscript.

#### Author affiliations

<sup>A</sup>UK Centre for Ecology and Hydrology, Lancaster Environment Centre, Lancaster, LA1 4AP, UK.

<sup>B</sup>School of Biosciences, University of Nottingham, Loughborough, Leicestershire, LE12 5RD, UK.

<sup>C</sup>British Geological Survey, Kingsley Dunham Centre, Keyworth, Nottingham, NG12 5GG, UK.

<sup>D</sup>Present address: The Woodland Trust, Kempton Way, Grantham, Lincolnshire, NG31 6LL, UK.