PARAMETERIZING THE BINDING PROPERTIES OF DISSOLVED ORGANIC MATTER WITH DEFAULT VALUES SKews THE PREDICTION OF COPPER SOLUTION SPECIATION AND ECOTOXICITY IN SOIL

TANALOU DJAE,† § || MATTHIEU N. BRAVIN,⁎|| CÉDRIC GARNIER,§ and EMMANUEL DOELSCH‡
† ADME, Angers, France
‡ CIRAD, UPR Recyclage et Risque, Montpellier, France
§ Université de Toulon, PROTEE, La Garde, France
(Submitted 7 July 2016; Returned for Revision 6 September 2016; Accepted 12 September 2016)

Abstract: Parameterizing speciation models by setting the percentage of dissolved organic matter (DOM) that is reactive (% r-DOM) toward metal cations at a single 65% default value is very common in predictive ecotoxicology. The authors tested this practice by comparing the free copper activity (pCu²⁺ = −log_{10}[Cu²⁺]) measured in 55 soil sample solutions with pCu²⁺ predicted with the Windermere humic aqueous model (WHAM) parameterized by default. Predictions of Cu toxicity to soil organisms based on measured or predicted pCu²⁺ were also compared. Default WHAM parameterization substantially skewed the prediction of measured pCu²⁺ by up to 2.7 pCu²⁺ units (root mean square residual = 0.75–1.3) and subsequently the prediction of Cu toxicity for microbial functions, invertebrates, and plants (root mean square residuals ≤ 9%, 11%, and 17%), respectively. Reparameterizing WHAM by optimizing the 2 DOM binding properties (i.e., % r-DOM and the Cu complexation constant) within a physically realistic value range much improved the prediction of measured pCu²⁺ (root mean square residual = 0.14–0.25). Accordingly, this WHAM parameterization successfully predicted Cu toxicity for microbial functions, invertebrates, and plants (root mean square residual ≤ 3.4%, 4.4%, and 5.8%, respectively). Thus, it is essential to account for the real heterogeneity in DOM binding properties for relatively accurate prediction of Cu speciation in soil solution and Cu toxic effects on soil organisms. Environ Toxicol Chem 2017;36:898–905.

Keywords: Trace elements Dose–response modeling Complexation Fulvic acid Windermere humic aqueous model (WHAM)

INTRODUCTION

Speciation in solution is unequivocally considered to be an effective way to estimate the bioavailability and toxicity of metal cations to soil organisms. Among metal species in solution, free metal cations are considered to be the main bioavailable and toxic species for soil organisms [1,2]. Accordingly, determination of the activity of free metal cations in soil solution is critical for most recently developed ecoxicological models [3–5].

The most direct approach to estimate the activity of free metal cations in soil solution is to determine it analytically. Many analytical techniques, such as the Donnan membrane technique and electrochemical methods based on voltammetry or potentiometry, are still used to measure the activity of metal cations in soil solution [6]. All of these techniques have, however, specific analytical drawbacks, such as being time-consuming, not directly related to the free metal form, and not sensitive enough to deal with the full range of the environmentally relevant activities of free metal cations usually found in real soil solutions.

The activity of free metal cations in soil solution may also be estimated by modeling. Speciation models work with 1) input parameters that can be measured routinely in soil solutions, such as pH and total metal cation and dissolved organic matter (DOM) concentrations, and 2) databases describing metal cation binding with inorganic ligands and DOM. Because of the high affinity of DOM for metal cations and the heterogeneity and complexity of the molecular structure of DOM, the modeling of metal cation binding with DOM is highly challenging [7].

The humic ion-binding model included in the Windermere humic aqueous model (WHAM) [8] and the nonideal competitive adsorption (NICA)-Donnan model [9,10] are currently the 2 models most widely used for simulating metal cation binding with DOM and consequently for estimating free metal cation activity. Because fulvic acids are among the most abundant DOM components in soil and particularly reactive toward metal cations [11–13], these 2 models similarly describe the binding properties of DOM as those of a reference fulvic acid parameterized specifically in each model. Beyond the analytically determined input parameters (e.g., pH, metal cation, and DOM concentrations), the activity of free metal cations is thus estimated in WHAM and the NICA–Donnan model by theoretically adjusting the percentage of DOM reactive toward metal cations (hereafter referred to as % r-DOM), corresponding to the fraction of DOM assumed to have the same reactivity toward metal cations as the defined reference fulvic acid, with the remaining DOM being considered inert. The free metal cation activity could also be successfully predicted by optimizing the binding constant of metal cations of the reference fulvic acid. However, model optimization by adjusting the % r-DOM is a more generic and practical approach [14–17].

Early studies focused on reactive DOM in natural waters revealed that the % r-DOM ranged from 30% to 120% but that the bulk of it was within the 40% to 80% range [14,15].
Vulkan et al. [18] predicted Cu$^{2+}$ activity in the solution of 22 soil samples from 4 areas, with a difference between measured and calculated values below ±0.5 free copper activity (pCu$^{2+}$; $-\log_{10}(\text{Cu}^{2+})$) units in 82% of the samples when the % r-DOM was set at a single value (i.e., 69%). Accordingly, in WHAM and the NICA–Donnan model, the % r-DOM is thereafter very often set at the single default value of 65% to predict the activity of free metal cations in soil solutions [19,20] and subsequently to predict the toxic effect of metal cations on soil organisms [3,4,21–24].

Nolan et al. [25], however, found that WHAM parameterization with a % r-DOM set at 100% or 70% poorly fit the measured activities of Cu$^{2+}$ and Pb$^{2+}$ in the solution of 25 soil samples. This result suggests that the % r-DOM variability in soil solutions could be much greater than previously expected. In agreement, Amery et al. [17] more recently estimated that the % r-DOM ranged from 28% to 152% in soil solutions collected over time in a soil profile. De Schamphelaere et al. [16] also showed that accounting for the variability in the % r-DOM when estimating Cu$^{2+}$ activity in solution improved the prediction of Cu toxicity to Daphnia magna. Accordingly, the very common practice of setting % r-DOM at a single 65% default value is perhaps an approximation. The extent to which this approximation may skew the prediction of metal cation speciation in solution and toxicity to soil organisms is highly questionable.

To test this very common practice, we compared 1) Cu$^{2+}$ activity measured in soil solutions with Cu$^{2+}$ activity predicted by modeling using default settings and 2) the prediction of Cu toxicity to soil organisms calculated from Cu$^{2+}$ activity measured with that calculated from Cu$^{2+}$ activity predicted by modeling using default settings. The present study was applied to 55 soil samples exhibiting a very wide range of physicochemical properties. Finally, uncertainty and sensitivity analyses were performed to highlight the impact on predicted Cu$^{2+}$ activity of 1) analytical errors in measured input parameters (i.e., pH, total Cu concentrations, and DOM concentration), and 2) the variability in parameters other than % r-DOM that are usually set at a single default value (i.e., Cu binding constant of DOM and the solubility products of Al(OH)$_3$ and Fe(OH)$_3$).

MATERIALS AND METHODS

Soil characteristics and incubation

Fifty-five soil samples were collected from agricultural and forest areas in mainland France ($n = 46$) and the 2 French tropical islands of Réunion ($n = 7$) and New Caledonia ($n = 2$). Soil samples were air-dried and sieved to 2 mm. The physicochemical properties of the soil samples were then determined according to French and international standardized procedures. Soil samples were selected to encompass a broad range of physicochemical soil characteristics and different origins of trace elements present (Table 1; Supplemental Data, Tables S1 and S2).

The 55 soil samples were incubated ex situ using the RHIZOtest experimental device. The RHIZOtest is a standardized plant-based test designed to evaluate the bioavailability of trace elements to soil-grown plants [26,27]. Briefly, cabbage (Brassica oleracea), tall fescue (Festuca arundinacea), and tomato (Lycopersicon esculentum) were grown in hydroponic conditions for 14 d. At the same time, the 55 soil samples were incubated at 70% of their water holding capacity with a nutrient solution consisting of 0.05 mM KH$_2$PO$_4$, 2 mM KNO$_3$, 1 mM MgSO$_4$, and 2 mM Ca(NO$_3$)$_2$. Plants were then transferred for 8 d onto a layer of each preincubated soil sample (6 mm thick, equivalent to 9 g dry mass, with a density of 1.2 g/cm$^3$), but a polyamide mesh (30 μm pore size) separated the roots from the soil. During soil–plant contact, a nutrient solution (same composition as previously described) connected to the soil layer via filter paper wicks (grade 544; Whatman) supplied the plants with major nutrients. During soil–plant contact, soil moisture thus increased up to 100% water holding capacity. After 8 d, the soil layer in contact with roots was considered a rhizosphere. Preincubated bulk soil samples were also incubated for 8 d in RHIZOtest devices without plants. The whole experiment was replicated 4 times successively (i.e., experiments 1–4) in similar conditions. In each experiment, 3 soil samples (18, 29, and 47) were replicated 3 times. The present study was focused on the bulk soil results (hereafter referred to as “soil results”); the plant-rhizosphere results will be discussed elsewhere.

The soil solution data analysis (see the section Extraction and analysis of soil solutions) using the Wilcoxon test for matched samples (Statistica, Ver 6; StatSoft) showed that the 4 experiments were significantly ($p < 0.05$) different from each other (Supplemental Data, Table S3). Data from the 4 experiments were consequently analyzed independently and not averaged. The results obtained in the 4 experiments are discussed in the text; for brevity, however, only data from experiment 1 are presented in the tables and figures. Data from experiments 2, 3, and 4 are presented as Supplemental Data.

| Table 1. Physicochemical properties of the 55 soil samples$^a$ |
|-----------------|--------|---------|--------|-------|-------|---------|--------|
| Minimum         | 3      | 3       | 4.4     | 0.0   | 0.1   | 0.9     | 0.1    |
| Quartile 1      | 9      | 17      | 5.8     | 0.1   | 1.3   | 1.1     | 5.2    |
| Median          | 15     | 34      | 6.4     | 0.4   | 2.3   | 1.3     | 10.2   |
| Mean            | 17     | 34      | 6.5     | 2.4   | 2.6   | 1.4     | 13.2   |
| Quartile 3      | 22     | 46      | 7.3     | 1.7   | 3.4   | 1.5     | 18.9   |
| Maximum         | 53     | 75      | 8.2     | 22.1  | 12.6  | 3.1     | 73.8   |

$^a$n = 55, except for CaCO$_3$, for which $n = 14$.  
$^b$According to the Association Francaise de Normalisation [39].  
$^c$According to the International Organization for Standardization [40].  
$^d$According to the International Organization for Standardization [41].  
$^e$According to the International Organization for Standardization [42].  
$^f$According to the International Organization for Standardization [43].  
$^g$According to the Association Francaise de Normalisation [44].  
$^h$Extracted with (NH$_4$)$_2$C$_2$O$_4$ at pH 3 according to Tamm [45].  
$^i$According to the International Organization for Standardization [46].  
CEC = cation exchange capacity.
After 14 ± 8 d of soil incubation with the nutrient solution (composition given in the section Soil characteristics and incubation), the soil solution was considered to be in equilibrium with the nutrient solution composition [28]. At harvest, each of the 55 moist soil samples was stirred for 2 h with the nutrient solution (dry soil, 1:10 nutrient solution ratio). The mixture of soil sample and nutrient solution was then centrifuged at 5000 g for 10 min and the supernatant was filtered at 0.22 µm with cellulose acetate filters prerinsed with ultrapure water. The first milliliter of filtrate was discarded. The soil solutions obtained were subdivided. Subsamples were stabilized with either 1 mM NaN₃ or HNO₃ 2.5 % (v/v) and finally stored at 4 °C until subsequent analysis.

Free Cu activity was measured using a Cu-selective electrode (Thermo Scientific Orion; 9629BNWP). The electrode was calibrated on a daily basis within the 13.0 to 4.9 range of pCu²⁺ in soil solutions buffered with iminodiacetic acid, as detailed elsewhere [29, 30]. The calibration curve showed a 95% slope efficiency compared with the theoretical Nernstian slope (β² = 0.997, n = 100). The pH was measured using a combined electrode (Mettler Toledo; InLab 418). In addition to calibration, the quality of pH and pCu²⁺ measurements was validated by repeating in each daily series a sample from the previous analytical series. Dissolved organic carbon (DOC) was determined using a carbon analyzer in nonpurgeable organic carbon mode (TOC-L; Shimadzu). Samples were bubbled for 90 s with 1.5% HCl to eliminate carbonates before injection into the oven (680 °C). The quality of DOC measurements was validated using the certified reference material Mississippi-03 that was recovered within 105 ± 12%. Total Cd, Cr, Cu, Ni, Pb, and Zn concentrations were determined by inductively coupled plasma–mass spectrometry (ICP-MS; Q-ICP-MS X Series II + CXTTM; Thermo Fischer). The quality of ICP-MS measurements was validated using the certified reference material SLRS-5. Concentrations of Ca, Mg, K, and Na were determined by ICP–atomic emission spectroscopy (Vista pro model; Varian). Concentrations of NH₄⁺ and Cl⁻ were determined by colorimetry (model Evolution II; Alliance). Concentrations of NO₃⁻ and SO₄²⁻ were determined by ion chromatography (DX-600; Dionex). All analytical determination findings are given as Supplemental Data (Table S4).

Modeling free Cu activity

Free Cu activity in soil solutions was predicted using WHAM (version VII), which was designed to simulate the competitive binding of proton and metal cations on fulvic and humic acids [31]. We used WHAM to assess the impact of the default settings on the prediction of pCu²⁺ in solution. The input parameters were the concentrations of major cations (Ca, Mg, Na, K, and NH₄⁺) and anions (Cl⁻, NO₃⁻, and SO₄²⁻); total Cu, Cd, Zn, Pb, Ni, and Cr concentrations; Al³⁺ and Fe³⁺ activities; concentration of DOM reactive toward metal cations; pH; CO₂ partial pressure (arbitrarily assigned at 10⁻³.5 atm); and temperature (25 °C). Free Al and Fe activities were calculated as commonly done by assuming their equilibrium with Fe(OH)₃ and Al(OH)₃ with solubility products (log Kₛ) of 2.7 and 8.5, respectively [20]. The binding properties of DOM toward proton and metal cations were simulated by using the reference fulvic acid defined by default in WHAM [31]. In WHAM, the Cu binding constant for type A sites of fulvic acid (log K_Cu) is set at 2.16 and used to calculate the Cu binding constant for type B sites (see Tipping et al. [31] for rationale). In the absence of information on the DOC/DOM ratio, the default option consisted of calculating the concentration of DOM from DOC based on the assumption that DOM contained 50% of C [14]. The concentration of DOM reactive toward metal cations was determined by multiplying the DOM concentration by % r-DOM. Free Cu activity was calculated with WHAM by using the default settings defined above, as is commonly done in predictive ecotoxicology, or by adjusting % r-DOM, log K_{Cu}, log K_{Fe}, or log K_{Al} for each soil to nearly fit the measured pCu²⁺. The deviation between measured and predicted pCu²⁺ was determined by calculating the root mean square residual. The uncertainty in predicted pCu²⁺ resulting from analytical uncertainties in pH (±0.3), total Cu concentration (±1.3%), and DOM concentration (±12.3%) was calculated using the Monte Carlo procedure (1997 samples) included in WHAM. The analytical uncertainties were determined by calculating the standard deviation on calibration standards and certified material. Sensitivity analyses were performed to determine the impact of the variability of log K_{Cu}, log K_{Fe}, and log K_{Al} on pCu²⁺ predicted with WHAM.

Modeling Cu toxicity on soil organisms

The free ion effective dose model was developed and parameterized to estimate the Cu toxicity at 3 soil organism levels: microorganisms, invertebrates, and higher plants [3]. Seven toxicological endpoints were thus targeted, which included the inhibition of potential nitrification, maize residue mineralization, glucose-induced respiration, the reproduction of Folsomia candida and Eisenia fetida, and the growth of Hordeum vulgare roots and L. esculentum shoots. The free ion effective dose model was described in detail by Lofts et al. [3]. Briefly, the toxic response (R) of soil organisms (R = 100, no toxic effect; R = 50 and 0, toxic effect equivalent to a 50% and 100% reduction in the biological endpoint, respectively) was calculated from the effective dose of toxicant in soil, which in turn was calculated from the Cu²⁺ concentration and pH in soil solution. A specific equation was used for each toxicological endpoint according to model 2b proposed by Lofts et al. [3].

We used the free ion effective dose model to theoretically assess the impact on R of the default settings defined in WHAM and used to predict Cu speciation in soil solution. The R values for the 7 toxicological endpoints were calculated for the 55 soil solutions from the pH and Cu²⁺ concentration. The Cu²⁺ concentration was calculated from pCu²⁺ measured or predicted with WHAM by using the default settings, as is commonly done in predictive ecotoxicology. The R deviation estimated from the predicted Cu²⁺ was compared with the R estimated from measured Cu²⁺ by calculating the root mean square residual.

RESULTS AND DISCUSSION

Wide variation in soil solution chemical properties

Although significant differences were noted between some of the 4 replicated experiments (Supplemental Data, Table S3) regarding the 4 main parameters (i.e., pH, total copper concentration [pCu_total], pCu²⁺, and DOM concentration) measured in the 55 soil solutions, the similarity of the trend exhibited between each replication of these 4 parameters was checked by assessing the correlations between each experiment and the mean values of the 4 experiments (Supplemental Data, Figure S1 and Tables S5 and S7). The pH, pCu_total, and pCu²⁺ values for the 55 soil solutions of each experiment were very well correlated (r²_adj ≥ 0.91) with the mean values of the 4
experiments, with a relative standard deviation of <4.2%, <5.1%, and <12.5%, respectively. Considering that the DOM concentration only varied over 1 order of magnitude, the DOM concentration values (expressed as log10) for the 55 soil solutions of each experiment were reasonably well correlated (0.51 < r² < 0.79) with the mean values of the 4 experiments and exhibited a relative standard deviation of <33.8%.

The chemical properties of the solutions extracted from the 55 soil samples varied widely in the 4 experiments (Figure 1; Supplemental Data, Figure S3). In experiment 1, for instance, the DOM concentration ranged from 2.7 mg/L to 94 mg/L, while the pH ranged from 5.5 to 7.9 and the total dissolved Cu, Pb, Ni, Zn, and Cd concentration ranged from 10⁻¹⁰ to 10⁻⁴.5 M, respectively. Accordingly, pCu²⁺ varied over 7 orders of magnitude, from 12.7 to 5.9. The Cu²⁺/total Cu concentration ratio ranged from 0.01% to 44%, which means that pCu²⁺ was not solely dependent on the pCuT variation. Our data set encompassed the variability in chemical parameters commonly found in the soil solution [18,32] and was therefore relevant for assessing the impact of reactive DOM on Cu speciation in solution and the ecotoxicity in soil.

**Default WHAM parameterization skews the prediction of Cu speciation in soil solution**

Default WHAM parameterization substantially skewed the prediction of measured pCu²⁺, with the root mean square residual ranging from 0.8 to 1.3 in the 4 experiments (Figure 2; Supplemental Data, Figures S4–S6, panel a). Considering that the measured pCu²⁺ varied widely (i.e., over approximately 7 orders of magnitude), the predicted pCu²⁺ with the default WHAM parameterization was poorly correlated with the measured pCu²⁺ (r²_adj = 0.61–0.73). The maximal deviation between predicted and measured pCu²⁺ was 2.7 pCu²⁺ units, 2.5 pCu²⁺ units, 1.8 pCu²⁺ units, and 2.1 pCu²⁺ units in experiments 1, 2, 3, and 4, respectively. The measured pCu²⁺ was mainly overestimated in experiments 1 and 2, and it was mainly underestimated in experiments 3 and 4.

The uncertainty in predicted pCu²⁺ as a result of analytical uncertainties in the measured pH, pCuT, and the DOM concentration used as input data in WHAM ranged from 0.32 pCu²⁺ units to 0.90 pCu²⁺ units in the 4 experiments (Figure 2; Supplemental Data, Figures S4–S6, panel a). These analytical uncertainties do not fully explain the discrepancy observed between the predicted and measured pCu²⁺ when WHAM was parameterized by default. The present results thus showed that parameterizing WHAM with default values, including % r-DOM set at 65% (as is common practice in predictive ecotoxicology), substantially skewed the prediction of the actual Cu speciation in soil solution, particularly when the investigated soil solutions exhibited a very broad range of chemical properties.

**Default WHAM parameterization skews the prediction of Cu ecotoxicity in soil**

In the 4 experiments, the toxic response (R) to Cu estimated from the default WHAM parameterization deviated by up to 59%, 45%, and 36% for plants (H. vulgare and L. esculentum),

---

Figure 1. Metal and dissolved organic matter (DOM) concentrations, pH, and Cu²⁺ activities (pCu²⁺) measured in the 55 soil solutions of experiment 1. The values represented from the bottom to the top indicate the minimum, first quartile, median, third quartile, and maximum values. pCdᵣ = total cadmium concentration; pCuᵣ = total copper concentration; pNiᵣ = total nickel concentration; pPbᵣ = total lead concentration; pZnᵣ = total zinc concentration.

Figure 2. Free Cu activity (pCu²⁺) predicted with the Windermere humic aqueous model VII (WHAM VII) as a function of the measured pCu²⁺ in the 55 soil solutions of experiment 1. Free Cu activity was predicted with WHAM VII using either (a) a single default value for the percentage of dissolved organic matter reactive toward metal cations (% r-DOM = 65%) or (b) an optimized % r-DOM and log Kₑ₆₅ value within a range restricted to physically meaningful values for each soil solution. Vertical error bars represent the 95% confidence interval of pCu²⁺ predicted with WHAM VII when considering the analytical uncertainty on pH and total Cu and DOM concentrations (see Materials and Methods for rationale). Empty data points represent soil samples 18, 29, and 47 that were replicated thrice with their standard deviation as a horizontal error bar. The thin solid line represents the regression line. The thick solid line represents: (a) the 1:1 line ± 0.5 pCu²⁺ unit; r²_adj = adjusted r²; RMSR = root mean square residual.
invertebrates (F. candida and E. fetida), and microbial functions (potential nitrification, maize residue mineralization, and glucose-induced respiration), respectively, from $R$ estimated from the measured $pCu^{2+}$ (Figure 3; Supplemental Data, Figures S7–S9, panels a, c, and e). These results showed that parameterizing WHAM by default, as is common practice in predictive ecotoxicology, substantially skewed the estimation of Cu ecotoxicity in soil, which was a result of the skewed prediction of the actual Cu speciation in soil solution.

The $R$ deviation between the default WHAM parameterization and the measured values showed the same trend as the sensitivity of soil organisms to Cu toxicity: plants > invertebrates > microbial functions. According to free ion effective dose parametrization [3], the concentration causing a 50% effect (EC50) increases in the following order: plants < invertebrates < microbial functions. This suggests that Cu toxicity has a greater impact on plants than on invertebrates and microbial functions. In the 4 experiments (Figure 3; Supplemental Data, Figures S7–S9, panels a, c, and e), $R$ deviations between the measured and predicted Cu activity ($pCu^{2+}$) were calculated with the free ion effective dose model from free Cu activity ($pCu^{2+}$) predicted with the Windermere humic aqueous model VII (WHAM VII) (y axis) or measured (x axis) in the 55 soil solutions of experiment 1. The toxic endpoints are the growth of Hordeum vulgare roots (Hv; blue circles) and Lycopersicon esculentum shoots (Le; red circles) for plants; the reproduction of Folsomia candida (Fc; blue triangles) and Eisenia fetida (Ef; red triangles) for invertebrates; and the inhibition of potential nitrification (PN; blue squares), maize residue mineralization (MRM; red squares), and glucose-induced respiration (GIR; green squares) for microbial functions. R values of 100 and 0 represent the minimal and maximal Cu toxicity, respectively. Free Cu activity was predicted with WHAM VII by using either (a,c,e) a single default value for the percentage of dissolved organic matter reactive toward metal cations ($\% r-DOM = 65\%$), the Cu binding constant of reactive DOM ($log K_{Cu} = 2.16$), and the solubility constants of Fe(OH)$_3$ ($log K_{sFe} = 2.7$) and Al(OH)$_3$ ($log K_{sAl} = 8.5$) for all solutions, or (b,d,f) an optimized $\% r-DOM$ and $log K_{Cu}$ value within a range restricted to physically meaningful values for each soil solution. Error bars represent the standard deviation ($n = 3$) for soil samples 18, 29, and 47. Solid line represents the 1:1 line. Dashed lines represent the 1:1 line $\pm 10\%$. DOM = dissolved organic matter; $% r-DOM$ = percentage of dissolved organic matter reactive toward metal cations; RMSR = root mean square residual.

Figure 3. Toxic response ($R$) of plants (a,b), invertebrates (c,d), and microbial functions (e,f) calculated with the free ion effective dose model from free Cu activity ($pCu^{2+}$) predicted with the Windermere humic aqueous model VII (WHAM VII) (y axis) or measured (x axis) in the 55 soil solutions of experiment 1. The toxic endpoints are the growth of Hordeum vulgare roots (Hv; blue circles) and Lycopersicon esculentum shoots (Le; red circles) for plants; the reproduction of Folsomia candida (Fc; blue triangles) and Eisenia fetida (Ef; red triangles) for invertebrates; and the inhibition of potential nitrification (PN; blue squares), maize residue mineralization (MRM; red squares), and glucose-induced respiration (GIR; green squares) for microbial functions. R values of 100 and 0 represent the minimal and maximal Cu toxicity, respectively. Free Cu activity was predicted with WHAM VII by using either (a,c,e) a single default value for the percentage of dissolved organic matter reactive toward metal cations ($\% r-DOM = 65\%$), the Cu binding constant of reactive DOM ($log K_{Cu} = 2.16$), and the solubility constants of Fe(OH)$_3$ ($log K_{sFe} = 2.7$) and Al(OH)$_3$ ($log K_{sAl} = 8.5$) for all solutions, or (b,d,f) an optimized $% r-DOM$ and $log K_{Cu}$ value within a range restricted to physically meaningful values for each soil solution. Error bars represent the standard deviation ($n = 3$) for soil samples 18, 29, and 47. Solid line represents the 1:1 line. Dashed lines represent the 1:1 line $\pm 10\%$. DOM = dissolved organic matter; $% r-DOM$ = percentage of dissolved organic matter reactive toward metal cations; RMSR = root mean square residual.
panels a, c, and e), the root mean square residual of $R$ ranged from 10% to 18% for plant growth ($H. vulgare$ and $L. esculentum$ combined), from 7% to 11% for invertebrate reproduction ($F. candida$ and $E. fetida$ combined), and from 1% to 9% for microbial functions (potential nitrification, maize residue mineralization, and glucose-induced respiration combined). An $R$ deviation of $>10\%$ between the default WHAM parameterization and the measured values was observed for 39% of the data points for plant growth ($H. vulgare$ and $L. esculentum$ combined), 16% of the data points for invertebrate reproduction ($F. candida$ and $E. fetida$ combined), and 8% of the data points for microbial functions (potential nitrification, maize residue mineralization, and glucose-induced respiration combined). This indicated that the $R$ deviation between the default and measured values increased at a given level of Cu exposure with the sensitivity of soil organisms to Cu toxicity.

The $R$ deviation between the default WHAM parameterization and the measured values tended to be maximal at approximately a 50% $R$ value but tended to be minimal for $R$ higher than approximately 80% and lower than approximately 20% for the 7 toxicological endpoints (Figure 3; Supplemental Data, Figures S7–S9, panels a, c, and e). This was directly related to the log-logistic dose–response relationship used in the free ion effective dose model to predict toxicological responses. Because of the sigmoidal shape of this relationship, the slope coefficient is highest at a 50% $R$ value and tends to be 0 when $R$ is 0% and 100% [33]. Accordingly, the skewed prediction of Cu speciation in soil solution in a pCu$^{2+}$ range close to the effective concentration causing an effect close to 0% or 100%, such as the widely used 10% effect concentration, would have less impact on the toxic response prediction. However, skewed prediction of Cu speciation in soil solution in a pCu$^{2+}$ range close to the EC50 would result in a high deviation in the toxic response prediction. Considering the widespread use of EC50 in predictive ecotoxicology, we feel that it is critical to optimize the default WHAM parameterization by including the optimization of % r-DOM when calculating Cu speciation in soil solution to avoid substantial deviation in the prediction of toxic effects on soil organisms.

### Table 2. Deviation calculated as the root mean square residual between measured and predicted free Cu$^{2+}$ activities (pCu$^{2+}$) in the 55 soil solutions of experiment 1

| Parameters | pCu$^{2+}$ predicted by default$^b$ | pCu$^{2+}$ predicted with unrestricted optimization$^a$ | pCu$^{2+}$ predicted with restricted optimization$^a$ |
|------------|-----------------------------------|--------------------------------------------------|--------------------------------------------------|
| % r-DOM    | 65                                | 0.8 to 7550                                      | 35 to 215$^c$                                     |
| Log $K_{Cu}$ | 2.16                              | $\leq -1.33$ to 2.99                            | 1.84 to 2.46$^d$                                  |
| Log $K_{Al}$ | 2.7                               | $\leq -6.3$ to 14.2                            | 2.5 to 5$^e$                                      |
| Log $K_{sAl}$ | 8.5                                | $\leq 1.3$ to 16.4                             | 8.0 to 9.0$^h$                                    |

$^a$Free Cu$^{2+}$ activities were predicted with Windermere humic aqueous model VII using either a single, default value for the percentage of dissolved organic matter reactive toward metal cations (% r-DOM), the Cu binding constant of reactive dissolved organic matter (log $K_{Cu}$), and the solubility constants of Fe(OH)$_3$ (log $K_{Fe}$) and Al(OH)$_3$ (log $K_{sAl}$) for all soil solutions, or an optimized value of % r-DOM, log $K_{Cu}$, log $K_{Fe}$, or log $K_{sAl}$ with or without restriction for each soil solution.

$^b$Free Cu$^{2+}$ was calculated by setting the 4 parameters at the default values given in model VII.

$^c$Free Cu$^{2+}$ was calculated by optimizing a single parameter within an unrestricted range of values, while the 3 other parameters were set at the default values given in model VII.

$^d$Free Cu$^{2+}$ was calculated by optimizing a single parameter within a range restricted to physically meaningful values, while the 3 other parameters were set at the default values given in model VII.

$^e$According to Tipping [8].

$^f$According to Baes and Mesmer [35].

$^g$According to Tipping et al. [36].

RMSR = root mean square residual.
complexation to fulvic acids reviewed by Tipping when parameterizing WHAM [8], physically meaningful log $K_{Cu}$ values should range from 1.84 to 2.34. According to Baes and Mesner [35], Tipping et al. [36], and Nordstrom et al. [37], physically meaningful log $K_{A}$ and log $K_{AI}$ values should range from 2.5 to 5 and from 8 to 9, respectively. This highlights the need to look for a more realistic optimization of % r-DOM, log $K_{Cu}$, log $K_{A}$, and log $K_{AI}$ while improving the ability of WHAM to adequately predict pHCu²⁺.

The WHAM parameterization by optimizing % r-DOM, log $K_{Cu}$, log $K_{A}$, or log $K_{AI}$ within a range restricted to physically realistic values substantially improved the prediction of measured pHCu²⁺ in comparison with the prediction obtained with default WHAM parameterization (Table 2; Supplemental Data, Tables S9–S11). For the 4 experiments, the root mean square residual calculated between predicted and measured pHCu²⁺ ranged from 0.41 to 0.73, 0.28 to 0.59, 0.35 to 1.21, and 0.73 to 1.24 when WHAM was parameterized by optimizing % r-DOM, log $K_{Cu}$, log $K_{A}$, or log $K_{AI}$, respectively. These WHAM prediction improvements remained, however, inadequate when we optimized 1 parameter at a time. This prompted us to further parameterize WHAM by optimizing 2 parameters simultaneously. The 2 best candidates for this simultaneous optimization were % r-DOM and log $K_{Cu}$, which also correspond to the 2 main parameters describing dissolved organic matter binding properties in WHAM. Moreover, WHAM predictions parameterized by optimizing the % r-DOM over an unrestricted range, while then modifying the log $K_{Cu}$, log $K_{A}$, or log $K_{AI}$ within physically realistic values range, showed high sensitivity with log $K_{Cu}$, lower sensitivity with log $K_{A}$, and insensitivity with log $K_{AI}$ (Supplemental Data, Figures S10–S13).

The WHAM parameterization by simultaneously optimizing both the % r-DOM and log $K_{Cu}$ within a range restricted to physically realistic values fitted very well the measured pHCu²⁺ with a root mean square residual less than 0.25 for the 4 experiments (Figure 2; Supplemental Data, Figures S4–S6 panel b). Consequently, the toxic response (R) estimated from the WHAM parameterization by simultaneously optimizing the % r-DOM and log $K_{Cu}$ values for each soil also very well fitted the R estimated from the measured pHCu²⁺ for the 7 endpoints in the 4 experiments, with root mean square residuals always <6% (Figure 3; Supplemental Data, Figures S7–S9, panels b, d, and f). These results showed that parameterizing WHAM by simultaneously optimizing both the % r-DOM and log $K_{Cu}$ within a realistic range of values substantially improved the prediction of Cu solution speciation and hence Cu ecotoxicity in soil.

The major shortcoming of the present work is that the DOM binding properties (% r-DOM and log $K_{Cu}$) were estimated by model optimization instead of by performing analytical determinations, thus questioning the empirical relevance and the practical application of our conclusions. The next step would be to analytically recover the % r-DOM and log $K_{Cu}$ values in some of the most contrasted soil solutions to check if they would confirm the WHAM predictions. Two analytical approaches seem promising to analytically estimate DOM binding properties in soil solution. The first approach is based on the chemical fractionation of DOM into humic acid, fulvic acid, hydrophilic acid, and hydrophobic neutral organic matter, with the great benefit of directly serving the parameterization of models such as WHAM or NICA-Donnan [13,38]. The second approach is based on ultraviolet absorbance, with the great benefit to be routinely applied to large sets of solution samples [17]. Overall, the present findings underline the need for further analytical investigations of the DOM binding properties to improve the parameterization and optimization procedure for geochemical models used in predictive ecotoxicology.

CONCLUSION

Parameterization speciation models such as WHAM with default values, particularly by setting the % r-DOM at 65%, has become very common practice for predicting metal cation speciation in solution and toxicity on soil organisms. However, the present results led us to conclude that it is critical to account for the real heterogeneity in DOM binding properties, including the % r-DOM and log $K_{Cu}$ when calculating Cu speciation in soil solution to avoid substantial deviation in the prediction of toxic effects on soil organisms.

In the 55 soil samples we investigated, WHAM parameterization by optimizing the % r-DOM, log $K_{Cu}$, log $K_{A}$, or log $K_{AI}$ over an unrestricted range led to much better prediction of measured pHCu²⁺ when optimizing either of the DOM binding properties (% r-DOM or log $K_{Cu}$) in comparison with the optimization of either log $K_{A}$ or log $K_{AI}$. However, unrestricted optimization led to much larger % r-DOM and log $K_{Cu}$ variation ranges than previously reported in the literature. More interestingly, WHAM parameterization by optimizing the 2 DOM binding properties simultaneously within a range restricted to physically realistic values predicted the measured pHCu²⁺ very well and, accordingly, much improved the prediction of Cu toxicity on soil organisms.

From a regulatory perspective, the present results more generally suggest that the derivation of soil quality limits based on quantitative relationships integrating soil properties should better account for the quality of organic matter as related to its binding properties toward metal cations.

Supplemental Data—The Supplemental Data are available on the Wiley Online Library at DOI: 10.1002/etc.3622.

Acknowledgment—The PhD scholarship of T. Djae was funded by ADEME and CIRAD. The French National Research Agency (ANR) via the NormaRHIZO project (09-CES-01) funded the present study. The authors kindly acknowledge J.P. Ambrosi (CNRS), D. Baize (INRA), L. Denaix (INRA), G. Faure (Université Jean Monnet), C. Jolivet (INRA), M. Mench (INRA), and F. van Oort (INRA) for providing some of the soil samples. The authors thank P. Cazeveille, C. Chevassus-Rosset, L. Lemal, J.-L. Marger, M. Montes, and C. Pradier (CIRAD) for the RHIZOtest experiments and the soil solution extraction and analyses; G. Durrieu (Université de Toulon) for dissolved organic matter analyses; and R. Freydier (CNRS) for ICP-MS analyses.

Data availability—Data, associated metadata, and calculation tools are available from the corresponding author (matthieu.bravin@cirad.fr).

REFERENCES
1. Parker D, Pedler J. 1997. Reevaluating the free-ion activity model of trace metal availability to higher plants. Plant Soil 196:223–228.
2. Ardestani MM, van Straalen NM, van Gestel CAM. 2014. The relationship between metal toxicity and biotic ligand binding affinities in aquatic and soil organisms: A review. Environ Pollut 195:133–147.
3. Lofst S, Cripl P, Janssen CR, Lock K, McGrath SP, Oorts K, Rooney CP, Smolders E, Spurgeon DJ, Svendsen C, Van Eekhout H, Zhao F-Z. 2013. Modelling the effects of copper on soil organisms and processes using the free ion approach: Towards a multi-species toxicity model. Environ Pollut 178:244–253.
4. Wang P, Kinraide TB, Smolders E, Zhou D-M, Menzies NW, Thakali S, Xia W-W, Hao X-Z, Peijnenburg WJGM, Kopittke PM. 2013. An electrostatic model predicting Cu and Ni toxicity to microbial processes in soils. Soil Biol Biochem 57:720–730.
