Environmental Toxicology and Chemistry—Volume 37, Number 1—pp. 80–90, 2018
Received: 8 February 2017 | Revised: 3 April 2017 | Accepted: 14 July 2017

Multiple Linear Regression Models for Predicting Chronic Aluminum Toxicity to Freshwater Aquatic Organisms and Developing Water Quality Guidelines

David K. DeForest,[a,*] Kevin V. Brix,[b,c] Lucinda M. Tear,[a] and William J. Adams[d]

[a]Windward Environmental, Seattle, Washington, USA
[b]EcoTox, Miami, Florida, USA
[c]Department of Marine Biology and Ecology, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida, USA
[d]Red Cap Consulting, Lake Point, Utah, USA

Abstract: The bioavailability of aluminum (Al) to freshwater aquatic organisms varies as a function of several water chemistry parameters, including pH, dissolved organic carbon (DOC), and water hardness. We evaluated the ability of multiple linear regression (MLR) models to predict chronic Al toxicity to a green alga (*Pseudokirchneriella subcapitata*), a cladoceran (*Ceriodaphnia dubia*), and a fish (*Pimephales promelas*) as a function of varying DOC, pH, and hardness conditions. The MLR models predicted toxicity values that were within a factor of 2 of observed values in 100% of the cases for *P. subcapitata* (10 and 20% effective concentrations [EC10s and EC20s]), 91% of the cases for *C. dubia* (EC10s and EC20s), and 95% (EC10s) and 91% (EC20s) of the cases for *P. promelas*. The MLR models were then applied to all species with Al toxicity data to derive species and genus sensitivity distributions that could be adjusted as a function of varying DOC, pH, and hardness conditions (the *P. subcapitata* model was applied to algae and macrophytes, the *C. dubia* model was applied to invertebrates, and the *P. promelas* model was applied to fish). Hazardous concentrations to 5% of the species or genera were then derived in 2 ways: 1) fitting a log-normal distribution to species-mean EC10s for all species (following the European Union methodology), and 2) fitting a triangular distribution to genus-mean EC20s for animals only (following the US Environmental Protection Agency methodology). Overall, MLR-based models provide a viable approach for deriving Al water quality guidelines that vary as a function of DOC, pH, and hardness conditions and are a significant improvement over bioavailability corrections based on single parameters. *Environ Toxicol Chem* 2018;37:80–90. © 2017 SETAC

Keywords: Aluminum; Bioavailability; Multiple linear regression model; Water quality guideline

INTRODUCTION

Aluminum (Al) can be present in a variety of forms in aquatic environments, including Al^{3+} and inorganic hydroxy species ($AlOH^{2+}$, $Al[OH]_2^{+}$, $Al[OH]_3^{0}$, and $Al[OH]_4^{-}$), inorganic complexes with fluoride (F^-) and sulfate (SO_4^{2-}), organic species based on weak and strong complexes with organic material that tend to keep Al in solution, and as an exchangeable fraction with soils, sediments, and precipitated organic material [1]. Aluminum is poorly soluble at neutral pH, but solubility, and toxicity, greatly increases at pH levels below 6 and above 8 [2]. The impacts of Al on freshwater ecosystems became apparent in approximately 1980, when it was realized that acid rain was increasing the solubility of Al and it was identified as an important factor affecting aquatic organisms in acidic aquatic habitats [1].

The bioavailability of Al to aquatic organisms depends on several water chemistry parameters. For example, pH determines both Al solubility and speciation, and protons (H^+) and other cations, such as calcium (Ca^{2+}), compete with Al^{3+} for uptake by aquatic biota [1]. Organic Al complexes, such as those formed with humic and fulvic acids, tend to increase total Al concentrations in solution but reduce bioavailability of Al to aquatic organisms [2].

Existing regulatory water quality criteria or guidelines for Al are dated. In 1988, the US Environmental Protection Agency (USEPA) released acute and chronic ambient water quality criteria (AWQC) of 750 and 87 μg/L, respectively [3]. These criteria were intended to be implemented using measurement of

This article includes online-only Supplemental Data.
* Address correspondence to DavidD@windwardenv.com
Published online 22 August 2017 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.3922

acid-soluble Al, which was operationally defined as the Al that passes through a 0.45-μm filter after the sample is acidified to a pH of 1.5 to 2.0 using nitric acid. In practice, however, most US states adopted Al criteria measured as total recoverable Al, which includes Al from suspended solids in the water samples. In 1987, the Canadian Council of Ministers of the Environment released Al water quality guidelines (expressed as total Al) that varied categorically depending on pH: 5 μg/L at pH <6.5 and 100 μg/L at pH ≥6.5 [4]. The Australian and New Zealand Environment and Conservation Council recommended Al guidelines based on the same pH categories, which are 0.8 μg/L as a "low reliability trigger value" at pH <6.5 and 55 μg/L as a "moderate reliability trigger value" at pH ≥6.5 [5].

All 3 of the above regulatory jurisdictions recommend Al criteria or guidelines that are ≤100 μg/L. However, the USEPA noted that field data indicate that many high-quality waters in the United States contain >87 μg Al/L, when either total recoverable or dissolved is measured [6]. Further, in one study of acid mine drainage–related stressors, benthic invertebrate communities in 46 stream sites were always identified as unimpaired at filtered (≤0.45 μm) Al concentrations up to 160 μg/L based on a stream condition index [7]; and a review of studies in acidified waters suggested that adverse effects on benthic invertebrates often occur in the 100 to 300 μg Al/L range and sometimes higher [8]. In addition, all 3 of these regulatory jurisdictions noted additional factors that can modify Al bioavailability and toxicity, such as organic matter and hardness (Ca^{2+} in particular), but the toxicity data were considered insufficient for development of criteria or guidelines that varied as a function of these parameters. Because of these issues regarding measuring toxicologically inappropriate forms of Al and not considering the effects of water quality parameters on Al bioavailability, existing Al criteria and guidelines in all jurisdictions are of questionable reliability.

Recently, Al toxicity to a green alga (*Pseudokirchneriella subcapitata*), a cladoceran (*Ceriodaphnia dubia*), and a fish (*Pimephales promelas*) was evaluated over varying pH, dissolved organic carbon (DOC), and hardness conditions, to more rigorously demonstrate the relative importance of these key parameters on Al bioavailability [9]. In the present study, these new data were used to evaluate whether chronic Al toxicity to aquatic biota could be predicted as a function of key water chemistry parameters (namely DOC, pH, and hardness) using multiple linear regression (MLR) models that could then be applied in the development of chronic Al guidelines following European Union and USEPA methodologies. The use of MLR models for copper (Cu) was recently proposed as a bioavailability-based approach for deriving water quality criteria [10].

METHODS

Data sources

The chronic toxicity of Al has been most robustly tested over a wide range of water chemistries for a green alga (*P. subcapitata*), a cladoceran (*C. dubia*), and the fathead minnow (*P. promelas*; Supplemental Data, Table S1). Many of these tests are described in Cardwell et al. [11] and Gensemer et al. [9]; toxicity data for these and other species were augmented from a literature search. Acceptable chronic toxicity data generally met USEPA guidelines for AWQC development [12]. One exception was inclusion of the 7-d survival and growth test for *P. promelas*. This is not an early-life stage test as defined in USEPA guidelines [12], but testing demonstrated that this test had a comparable sensitivity to Al as an early-life stage test (Supplemental Data, Table S1). All Al toxicity data used in the present evaluation are expressed as total Al because this includes both dissolved and precipitated forms of Al that contribute to potential toxicity in laboratory waters [9,11,13].

MLR analysis

The MLR analysis was conducted on the chronic toxicity data for *P. subcapitata*, *C. dubia*, and *P. promelas* (all data in Supplemental Data, Table S1 for these species were included in the analysis). We focused on DOC, pH, and hardness as the parameters that most influence Al toxicity and, hence, the magnitudes of 10 and 20% effective concentrations (EC10s and EC20s) [1,9,11]. To inform the MLR analysis, we first evaluated whether any of the relationships between Al EC20s and DOC, pH, and hardness were nonlinear (for simplicity, we focus on EC20s when discussing relationships between Al toxicity and water chemistry; the same patterns were observed for EC10s). Aluminum EC20s were plotted against each independent variable (DOC, pH, and hardness) using data from those laboratory studies in which one parameter was varied whereas the other 2 were held constant. A single, robust study with a $3 \times 3 \times 3$ factorial design was available for *P. subcapitata* using DOC concentrations of 0.3, 2, and 4 mg/L; pH levels of 6.1 to 6.2, 7.0 to 7.1, and 7.8 to 8.0; and hardness concentrations of 22 to 24, 60 to 61, and 120 to 121 mg/L (as $CaCO_3$). Aluminum EC20s in the present study generally increased with increasing DOC concentration across all pH and hardness conditions (Figure 1A–C), increased from pH approximately 6.2 to pH approximately 7.0 and decreased from pH approximately 7.0 to pH approximately 8.0 regardless of DOC and hardness concentration (Figure 1D–F), and increased with hardness in pH approximately 6.2 waters while decreasing with hardness in pH approximately 8 waters (Figure 1G–I). The changes in slope at some combinations of DOC, pH, and hardness indicate that these variables interact in their effects on *P. subcapitata*. For *C. dubia* and *P. promelas*, fewer tests have been conducted in which one of the water chemistry parameters was varied and the other 2 were held constant; however, based on the studies available, it appears that both *C. dubia* and *P. promelas* EC20s generally increase with each independent variable (DOC, pH, and hardness) regardless of the levels of the other 2 variables (Figures 2 and 3).

Following Brix et al. [10], we conducted initial stepwise MLR analyses in which ln(DOC), pH, and ln(hardness) were all introduced to the model as independent variables with either ln[EC10] or ln[EC20] as the dependent variable; EC10s were used to evaluate hazardous concentrations to 5% of the species or genera (HC5s) using the European Union approach, and EC20s were used to evaluate HC5s using the USEPA approach.

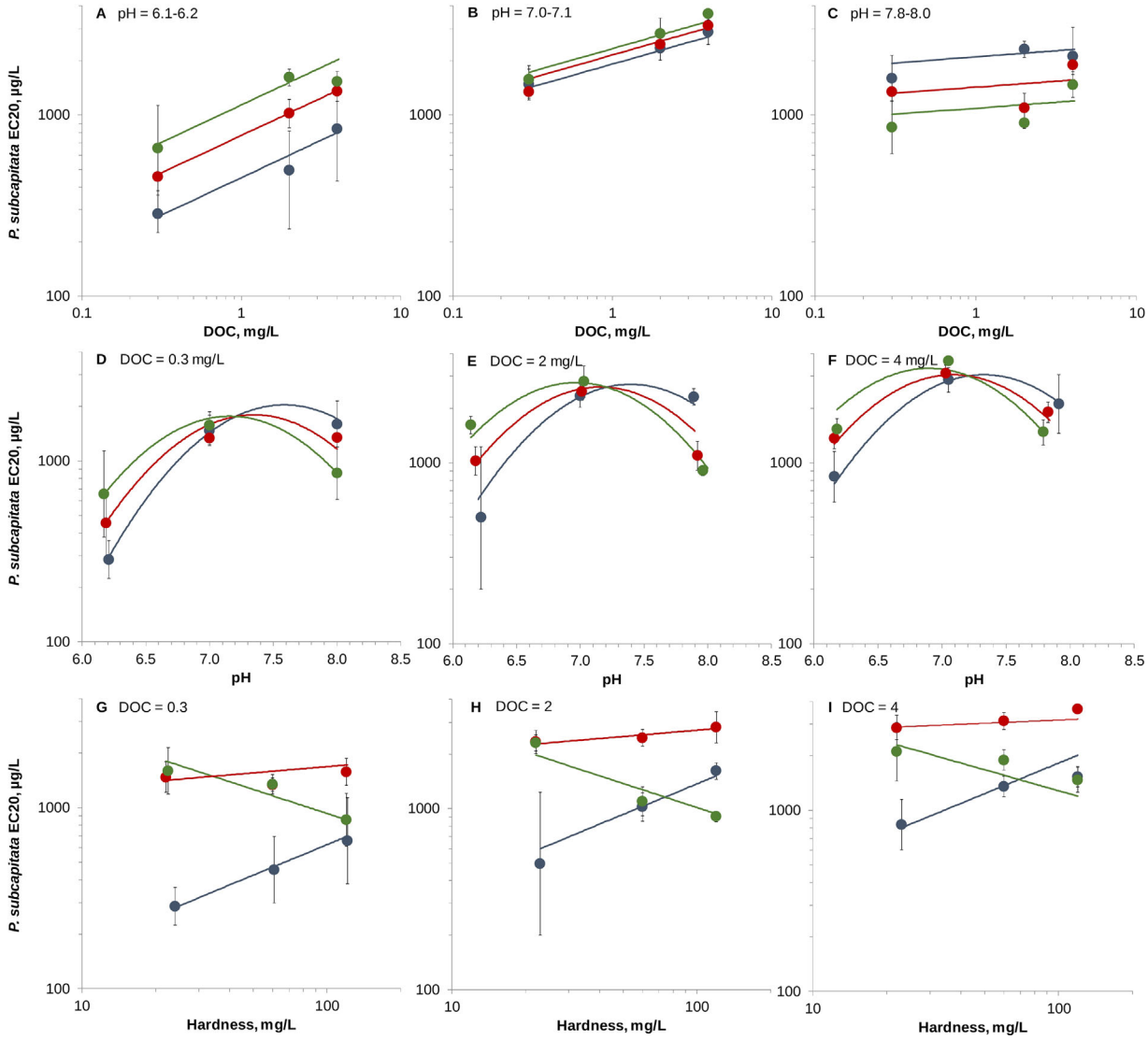

FIGURE 1: Observed and multiple linear regression–predicted (regression lines) total Al 20% effect concentrations (±95% confidence limits) for *Pseudokirchneriella subcapitata* growth rate where one water chemistry parameter was varied. (**A–F**) Hardness of 22 to 24, 60 to 61, and 120 to 121 mg/L (blue, red, and green symbols, respectively). (**G–I**) pH of 6.1 to 6.2, 7.0 to 7.1, and 7.8 to 8.0 (blue, red, and green symbols, respectively). EC20 = 20% effect concentration.

The MLR model without interaction terms is as follows:

$$\ln(\mathrm{ECX}) = b_0 + b_1 \times \ln[\mathrm{DOC}] + b_2 \times \mathrm{pH} + b_3 \times \ln[\mathrm{H}] \quad (1)$$

where ECX = EC10 or EC20 (micrograms of Al per liter); b_0 is the intercept; b_1, b_2, and b_3 are the slopes for ln(DOC), pH, and ln(H), respectively; and DOC and hardness (H) are in units of milligrams per liter.

We also conducted a second phase of stepwise MLR analyses that included the main independent variables as well as all 2-way interactions (ln[DOC]×pH, ln[DOC]×ln[H], ln[H]×pH) among the main independent variables to capture some of the interactions apparent in the graphical plots. Specific to Al, we also considered a pH^2 term given its unusual chemistry and the peak in EC20s at intermediate pH. For example, a negative pH^2 term would help account for the fact that Al bioavailability decreases from pH 6 to pH 7 and then increases from pH 7 to pH 8. Likewise, a negative ln(DOC)×pH term would characterize the mitigating effect of DOC on Al bioavailability, which tends to decrease as pH increases; a negative ln(DOC)×ln(H) term would reflect that the mitigating effect of DOC on Al bioavailability tends to decrease as hardness increases; and a negative ln(H)×pH term would reflect that the mitigating effect of hardness on Al bioavailability tends to decrease as pH increases. The MLR model with interaction terms is as follows:

$$\begin{aligned}\ln(\mathrm{ECX}) = {} & b_0 + b_1 \times \ln[\mathrm{DOC}] + b_2 \times \mathrm{pH} + b_3 \times \ln[\mathrm{H}] \\ & + b_4 \times \mathrm{pH}^2 + b_5 \times (\ln[\mathrm{DOC}] \times \mathrm{pH}) + b_6 \times (\ln[\mathrm{DOC}] \\ & \times \ln[\mathrm{H}]) + b_7 \times (\ln[\mathrm{H}] \times \mathrm{pH}) \end{aligned} \quad (2)$$

Because results of MLR models can be difficult to interpret when independent variables are correlated, variance inflation factors were calculated to assess collinearity among the 3 main independent variables in the first phase of modeling. The variance inflation

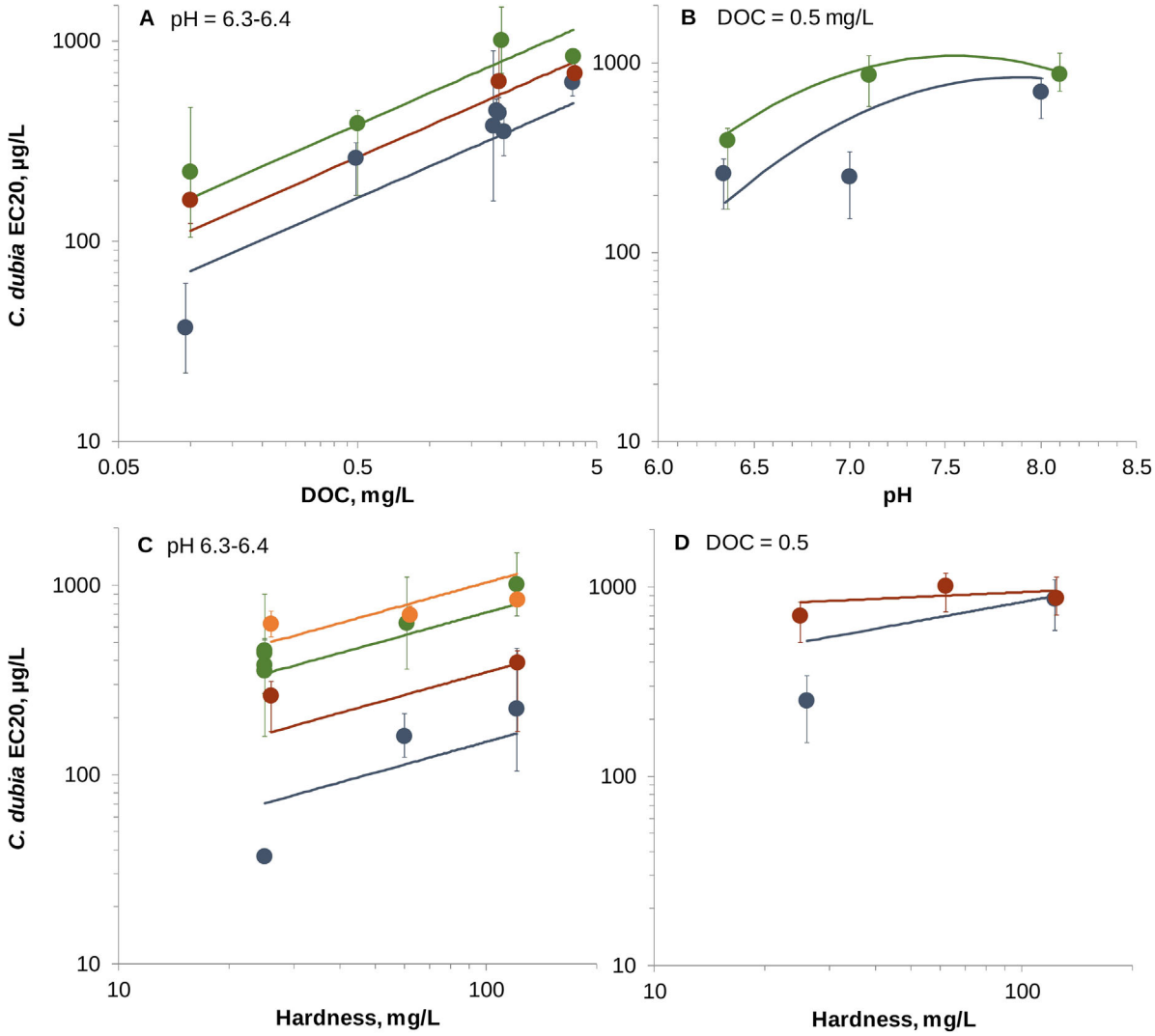

FIGURE 2: Observed and multiple linear regression–predicted (regression lines) total Al 20% effect concentrations (±95% confidence limits) for *Ceriodaphnia dubia* reproduction where one water chemistry parameter was varied. (**A**) Hardness of 25 to 26, 60 to 61, and 121 to 122 mg/L (blue, red, and green symbols, respectively). (**B**) Hardness of 25 to 26 and 122 to 123 mg/L (blue and green symbols, respectively). (**C**) Dissolved organic carbon of 0.1, 0.5, 2, and 4 mg/L (blue, red, green, and orange symbols, respectively). (**D**) pH of 7.0 to 7.1 and 8.0 to 8.1 (blue and red symbols, respectively). EC20 = 20% effect concentration; DOC = dissolved organic carbon.

factor is a measure of the degree to which the variance of a covariate in a multivariate linear model is inflated by its correlation with other variables in the model. The higher the variance in the covariate, the less likely it is to be considered significantly different from zero and the less able the stepwise approach is to quantify the true effect of each covariate on the dependent variable. A variance inflation factor of 10 indicates severe multicollinearity and almost complete correlation between a variable and a linear combination of other independent variables in the model [14]. Variance inflation factors in the range of 1 indicate that correlation between a covariate and other independent variables in the model is less than 0.1. In factorial experiments in which the levels of each independent variable are carefully controlled, correlation among independent variables is not anticipated, whereas in field studies covariance of independent variables is often an issue.

We employed the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) to determine which combination of terms resulted in the most parsimonious models for predicting Al EC10s or EC20s for each species [15]. Both the AIC and the BIC achieve parsimony and balance specificity and generality by penalizing a model's goodness-of-fit term by a factor related to the number of parameters (p) in the model. The BIC also considers sample size (n). The penalty used by the AIC is $2p$ (for all n) and that for the BIC is $\ln(n)p$ so that the full goodness of fit function to be minimized, where L = log likelihood, is

$$\text{AIC} = -2L + 2p \tag{3}$$

$$\text{BIC} = -2L + \ln(n)p \tag{4}$$

When $\ln(n) < 2$ (i.e., when $n < 8$), the AIC penalty is greater than the BIC penalty and the AIC will tend to retain fewer

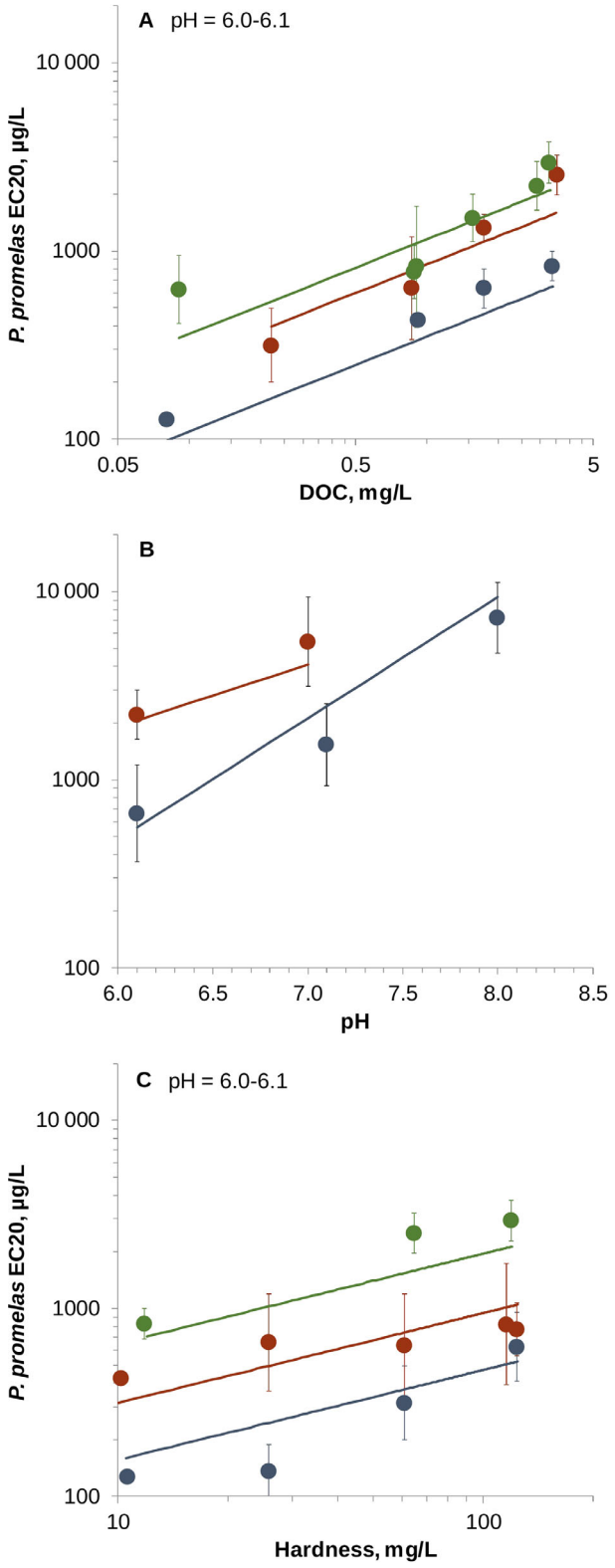

FIGURE 3: Observed and multiple linear regression–predicted (regression lines) total Al 20% effect concentration (±95% confidence limits) for *Pimephales promelas* biomass where one water chemistry parameter was varied. (**A**) Hardness of 10 to 12, 60 to 65, and 116 to 124 mg/L (blue, red, and green symbols, respectively). (**B**) Dissolved organic carbon (DOC) of 0.7 to 0.8 mg/L and hardness of 26 to 29 mg/L (blue symbols); DOC of 2.5 to 2.9 mg/L and hardness of 122 to 123 mg/L (red symbols). (**C**) Dissolved organic carbon of 0.1 to 0.3, 0.7 to 0.9, and 3.3 to 3.5 mg/L (blue, red, and green symbols, respectively). EC20 = 20% effect concentration.

independent variables in the model. When $n > 8$, the BIC penalty is larger and the BIC model will tend to retain fewer independent variables. Including too many terms in a model runs the risk of overextrapolating from the specific data set so that the model is less general; including too few terms may reduce predictive precision [16]. When the AIC and BIC recommended different models, we also considered our understanding of Al bioavailability to assess which was most appropriate.

Stepwise regressions were run in R [17] using the stepAIC() function from the MASS library (e.g., for BIC model, stepAIC {model, direction = c["both"], k = 1n[n]}). Variance inflation factors were calculated in R using the vif() function in the usdm library.

The predicted values from each stepwise MLR model were plotted against the observed EC10s or EC20s to visually assess the goodness-of-fit of each final model. Absolute residuals for each model were plotted versus single independent variables (DOC, pH, hardness) to look for patterns in the predictions under varying water chemistry conditions. Horizontal grid lines were included to show ±2 standard deviations in the residuals. In addition, the percentages of predicted EC10s or EC20s that were within a factor of 2 of the observed EC10s or EC20s were evaluated as a measure of stepwise MLR model accuracy.

Species and genus sensitivity distributions and HC5s

The MLR models were used to derive chronic Al HC5s following both European Union and USEPA approaches [12,18]. Among the key differences between these 2 approaches are 1) use of EC10s in the European Union versus EC20s in the United States; 2) use of species-mean EC10s in the European Union versus genus-mean EC20s in the United States; 3) inclusion of algae and plant data in the EU species sensitivity distribution (SSD) versus exclusion of algae and plant data in the US genus sensitivity distribution (GSD); and 4) use of different statistical models for calculating the HC5. Although the United States does not include algae and plant data in determining the HC5, the sensitivity of algae and plant species relative to the HC5 is evaluated to ensure it is protective.

To adjust the chronic Al SSDs (following the European Union approach) and GSDs (following the USEPA approach) for water chemistry characteristics of interest, the MLR models developed for *P. subcapitata*, *C. dubia*, and *P. promelas* were applied to the chronic toxicity data summarized in Supplemental Data, Table S1. The final *P. subcapitata* MLR model was applied to algae and macrophytes, the final *C. dubia* MLR model was applied to invertebrates, and the final *P. promelas* model was applied to fish. We recognize there is uncertainty in applying an MLR model for one species and endpoint to another species and potentially a different endpoint, but this is an uncertainty common to hardness- and biotic ligand model (BLM)–based approaches for bioavailability-based adjustments to SSDs and GSDs. The empirical Al EC10s or EC20s can be adjusted to the desired DOC, pH, and hardness concentrations, at a site of

interest for example, using the MLR modeling results applied to the following basic equation:

$$\begin{aligned}\ln(ECX_{norm}) = {} & \ln(ECX_{meas}) - b_1[\ln(DOC_{meas}) - \ln(DOC_{site})] \\ & - b_2[pH_{meas} - pH_{site}] - b_3[\ln(H_{meas}) - \ln(H_{site})] \\ & - b_4[pH^2_{meas} - pH^2_{site}] - b_5[\ln(DOC_{meas}) \times pH_{meas} \\ & - \ln(DOC_{site}) \times pH_{site}] - b_6[\ln(DOC_{meas}) \times \ln(H_{meas}) \\ & - \ln(DOC_{site}) \times \ln(H_{site})] - b_7[\ln(H_{meas}) \times pH_{meas} \\ & - \ln(H_{site}) \times pH_{site}] \end{aligned} \quad (5)$$

In Equation 5, ECX_{norm} is the Al EC10 or EC20 (micrograms of Al per liter) normalized to the site DOC (milligrams per liter), pH, and hardness (milligrams per liter) concentrations; ECX_{meas} is the empirically measured Al EC10 or EC20; b_1 is the slope of the relationship between ln(ECX) values and ln(DOC) concentrations from the MLR model; DOC_{meas} is the empirically measured DOC concentration in the toxicity test; DOC_{site} is the DOC concentration in the site water of interest, with analogous terms for the remaining 6 independent variables.

We applied the MLR models based on EC10s to the EC10 data used to develop the SSDs following the EU approach and the MLR models based on EC20s to the EC20 data used to develop the GSDs following the USEPA approach. However, we note that it may be reasonable to assume common slopes to have a single model that is applicable to both EC10 and EC20 data.

Once EC10s or EC20s were adjusted to the DOC, pH, and hardness concentrations of interest, chronic Al SSDs were derived based on species-mean EC10s for all species (algae, plants, and animals) and chronic GSDs were derived based on genus-mean EC20s for animals. Chronic EC20s for algae and plants were also calculated for each water chemistry of interest. A log-normal distribution was used to calculate the HC5 (5th percentile EC10) following the European Union approach, and the HC5 (5th percentile EC20) based on the USEPA approach was calculated following Stephan et al. [12].

The relative ranking of species in the SSD and of genera in the GSD varies with water chemistry because the MLR models for *P. subcapitata*, *C. dubia*, and *P. promelas* do not respond the same to changes in water chemistry. This is different from the USEPA's hardness-based criteria for metals, in which a pooled hardness slope is applied to all species, and BLM–based Cu criteria, in which only the sensitivity term varies between species—in these cases the relative ranking of species and genera does not change with water chemistry. It is, however, consistent with the European Union approach for developing BLM–based SSDs in which separate algal, invertebrate, and fish BLMs are applied to the toxicity data for different water chemistry conditions.

RESULTS AND DISCUSSION

MLR models

The variance inflation factors for ln(DOC), pH, and ln(H) in the *P. subcapitata*, *C. dubia*, and *P. promelas* EC10 and EC20 models ranged from 1.0 to 1.2, indicating very low correlation between independent variables, as would be expected in controlled experiments. The Al toxicity data used to derive the MLR models were from tests in which an individual parameter was varied (e.g., DOC was varied as pH and hardness were held constant). In contrast, variance inflation factors in the Cu MLR models described in Brix et al. [10] ranged from 1.0 to 5.6, with higher variance inflation factors associated with Cu toxicity studies conducted in natural waters where water chemistry parameters may covary.

Pseudokirchneriella subcapitata. When ln(DOC), pH, and ln(H) were included as the independent parameters in the initial stepwise MLR analyses for *P. subcapitata* EC10s and EC20s, the AIC and BIC both retained DOC and pH in the models but not hardness (Table 1). The adjusted R^2 values were 0.312 and 0.335 for the EC10s and EC20s, respectively. The predicted EC10s and EC20s were within a factor of 2.0 of observed in 85 and 89% of the tests, respectively (Supplemental Data, Figure S1A for EC10s; Figure 4A for EC20s).

When the 4 interaction terms were added to the models, the final model using both the AIC and BIC fits included all but the ln(DOC)×ln(H) term. The adjusted R^2 values for the EC10 and EC20 models increased to 0.940 and 0.956, respectively (Table 1), and all predicted EC10s and EC20s were within a factor of 2.0 of observed (Supplemental Data, Figure S1A for EC10s; Figure 4A for EC20s). No clear patterns in the residuals were observed over a wide range of Al bioavailability conditions as expressed by the magnitudes of the EC10s and EC20s (Supplemental Data, Figure S2A,B) or relative to single independent variables (Supplemental Data, Figures S3A,B; S4A,B; and S5A,B). Lastly, inclusion of the interaction terms allowed for modeling of the conspicuous Al toxicity trends observed in the *P. subcapitata* data set, namely reduced toxicity from pH approximately 6.2 to 7 and increased toxicity from pH 7 to 8 (Figure 1D–F). This was important for obtaining accurate predictions of toxicity. (Also, see Figure 1A–C and G–I for examples of how the MLR model for *P. subcapitata* EC20s performs relative to empirical toxicity data under conditions where DOC or hardness were varied, respectively).

Ceriodaphnia dubia. When ln(DOC), pH, and ln(H) were included in the initial stepwise MLR analysis for *C. dubia*, the AIC and BIC retained all 3 parameters in both the EC10 and EC20 models; adjusted R^2 values were 0.658 and 0.674, respectively (Table 1). The predicted EC10s and EC20s were within a factor of 2.0 of observed in 91% of the tests (Supplemental Data, Figure S1B for EC10s; Figure 4B for EC20s).

We initially evaluated the same 4 interaction terms included in the *P. subcapitata* model. Both the AIC and BIC retained the pH^2 and ln(DOC)×pH terms, and the AIC also retained the ln(DOC)×ln(H) term. The adjusted R^2 values were 0.756 and 0.738 based on the AIC and BIC, respectively. A conspicuous observation from these models was prediction of increasing Al toxicity (i.e., decreasing EC20s) with increasing DOC at a pH of 8 (Supplemental Data, Figure S2). The empirical toxicity data for *C. dubia*, at a DOC concentration of 0.5 mg/L, indicate that Al EC20s between a pH of 7 and 8 either increase (in soft water) or plateau (in hard water) (Figure 2B). Unfortunately, there are no empirical *C. dubia* toxicity data at a pH of 8 and DOC

TABLE 1: Aluminum multiple linear regression model statistics for *Pseudokirchneriella subcapitata*, *Ceriodaphnia dubia*, and *Pimephales promelas* 10 and 20% effective concentrations[a]

						Slopes							
Species	Endpoint statistic	*n*	Model	AIC/BIC	Adj. r^2	ln(DOC) [b_1]	pH [b_2]	ln(H) [b_3]	pH^2[b_4]	ln(DOC)× pH [b_5]	ln(DOC)× ln(H) [b_6]	ln(H)×pH [b_7]	ln(Intercept [b_0]
P. subcapitata	EC10	27	No interactions (AIC and BIC)	56.56	0.312	0.349 ($p = 0.004$)	0.346 ($p = 0.05$)	—	ne	ne	ne	ne	4.372
			With interactions (AIC and BIC)	–6.36	0.940	2.342 ($p < 0.001$)	20.923 ($p < 0.001$)	4.560 ($p < 0.001$)	–1.274 ($p < 0.001$)	–0.288 ($p < 0.001$)	—	–0.628 ($p < 0.001$)	–77.283
	EC20	27	No interactions (AIC and BIC)	44.63	0.335	0.276 ($p = 0.005$)	0.327 ($p = 0.02$)	—	ne	ne	ne	ne	4.837
			With interactions (AIC and BIC)	–25.48	0.956	1.678 ($p < 0.001$)	17.019 ($p < 0.001$)	4.007 ($p < 0.001$)	–1.020 ($p < 0.001$)	–0.204 ($p < 0.001$)	—	–0.556 ($p < 0.001$)	–61.952
C. dubia	EC10	23	No interactions (AIC and BIC)	43.94	0.658	0.536 ($p < 0.001$)	0.871 ($p < 0.001$)	0.375 ($p = 0.02$)	ne	ne	ne	ne	–1.311
			With interactions (AIC)	41.49	0.719	1.208 ($p = 0.02$)	13.338 ($p = 0.05$)	3.639 ($p = 0.04$)	–0.737 ($p = 0.11$)	ex	–0.155 ($p = 0.20$)	–0.490 ($p = 0.06$)	–51.420
			With interactions (BIC)	42.79	0.685	0.571 ($p < 0.001$)	2.282 ($p = 0.02$)	2.900 ($p = 0.08$)	—	ex	—	–0.365 ($p = 0.12$)	–11.055
	EC20	23	No interactions (AIC and BIC)	36.66	0.674	0.455 ($p < 0.001$)	0.780 ($p < 0.001$)	0.366 ($p = 0.01$)	ne	ne	ne	ne	–0.420
			With interactions (AIC)	34.17	0.726	0.525 ($p < 0.001$)	11.282 ($p = 0.05$)	2.201 ($p = 0.10$)	–0.663 ($p = 0.09$)	ex	—	–0.264 ($p = 0.17$)	–41.026
			With interactions (BIC)	34.80	0.710	0.502 ($p < 0.001$)	10.709 ($p = 0.07$)	0.378 ($p = 0.006$)	–0.695 ($p = 0.09$)	ex	—	—	–35.571
P. promelas	EC10	22	No interactions (AIC and BIC)	35.60	0.822	0.495 ($p < 0.001$)	0.966 ($p < 0.001$)	0.354 ($p = 0.008$)	ne	ne	ne	ne	–0.696
			With interactions (AIC and BIC)	32.70	0.850	0.545 ($p < 0.001$)	2.794 ($p = 0.006$)	3.201 ($p = 0.03$)	—	ex	—	–0.460 (0.06)	–12.009
	EC20	22	No interactions (AIC and BIC)	35.49	0.845	0.449 ($p < 0.001$)	1.166 ($p < 0.001$)	0.383 ($p = 0.004$)	ne	ne	ne	ne	–1.869
			With interactions (AIC and BIC)	31.70	0.874	0.503 ($p < 0.001$)	3.131 ($p = 0.002$)	3.443 ($p = 0.02$)	—	ex	—	–0.494 ($p = 0.04$)	–14.029

[a] Independent variables: ln(DOC), pH, and ln(H) in multiple linear regression analyses with no interaction terms; pH^2, ln(DOC)×pH, ln(DOC)×ln(H), and ln(H)×pH added in multiple linear regression analyses with interaction terms included. The Akaike information criterion and Bayesian information criterion were used to identify the best models. — indicates slope not provided for Akaike information criterion– or Bayesian information criterion–recommended models because it did not improve the model fit.

AIC = Akaike information criterion; BIC = Bayesian information criterion; EC10/EC20 = 10 and 20% effective concentrations; H = hardness; ex = excluded as an interaction term (see text); ne = not evaluated.

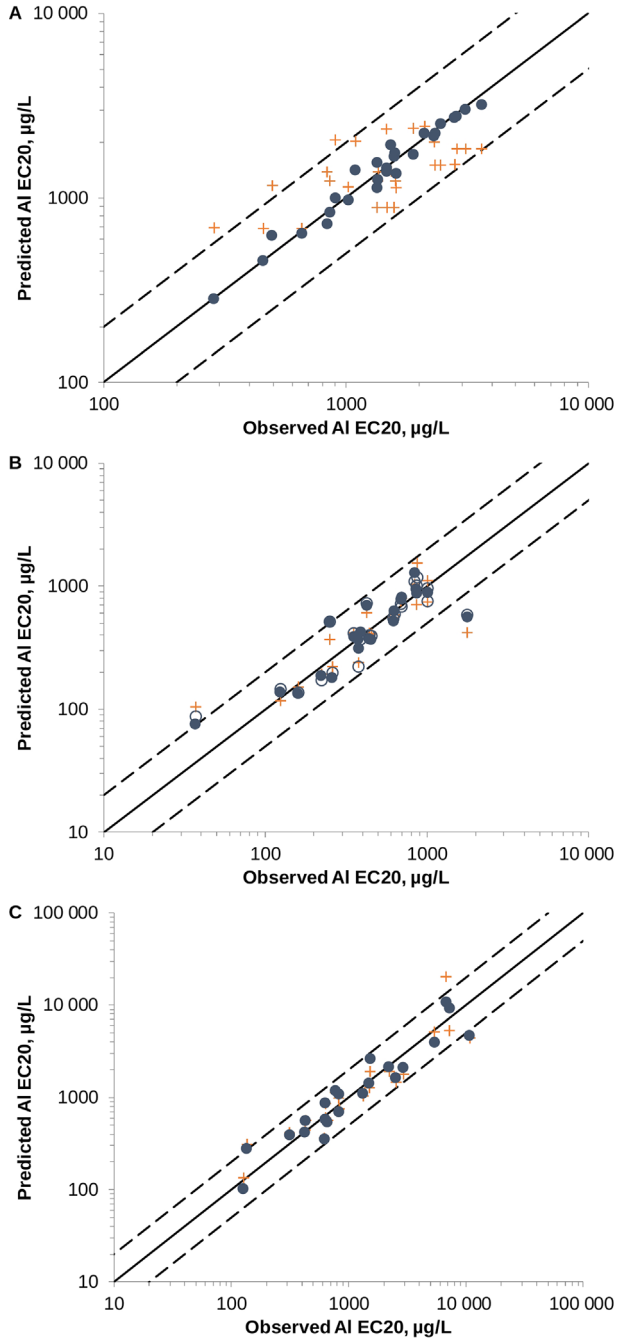

FIGURE 4: Comparison of stepwise multiple linear regression models with and without interaction terms as determined by the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Predicted versus observed total Al 20% effect concentrations for (**A**) *Pseudokirchneriella subcapitata*, (**B**) *Ceriodaphnia dubia*, and (**C**) *Pimephales promelas*. Models without interaction terms shown by "plus" symbols; models with interaction terms shown by circles. In (**A**) and (**C**), AIC and BIC determined the same models. In (**B**), AIC model predictions are filled circles and BIC model predictions are open circles. Solid line represents 1:1 agreement; dashed lines represent a factor of ±2 agreement. EC20 = 20% effect concentration.

concentrations >0.5 mg/L. The increasing toxicity prediction with increasing DOC concentration at high pH was attributed to the negative ln(DOC)×pH term, in which the combined influence of high DOC and pH resulted in rapidly decreasing predicted Al EC20s. Given the absence of toxicity data at combinations of high pH and DOC >0.5 mg/L, we concluded that the ln(DOC)×pH term required extrapolation beyond the empirical data for *C. dubia*; and it was removed as an input to the model.

When the remaining 3 interaction terms (pH^2, ln[DOC]×ln[H], ln[H]×pH) were provided to the *C. dubia* stepwise analysis, the AIC and BIC again retained different interaction terms (Table 1). The adjusted R^2 values for the final EC10 and EC20 models were 0.719 and 0.726, respectively, based on the AIC and 0.685 and 0.710 based on the BIC (Table 1). The predicted EC10s and EC20s were within a factor of 2.0 of observed in 91% of the tests (Supplemental Data, Figure S1B for EC10s; Figure 4B for EC20s). Also, as for *P. subcapitata*, no clear patterns were observed in the residuals over a wide range of Al bioavailability conditions as expressed by the magnitudes of the EC10s and EC20s (Supplemental Data, Figure S3C,D) or relative to single independent variables (Supplemental Data, Figure S4C,D; S5C,D; and S6C,D). Examples of how the MLR model for *C. dubia* EC20s performs relative to empirical toxicity data under conditions where either DOC, pH, or hardness was varied are provided in Figure 2.

Pimephales promelas. In the initial stepwise MLR analysis for *P. promelas*, ln(DOC), pH, and ln(H) were all retained in both AIC and BIC models for both the EC10 and EC20. The adjusted R^2 values were 0.822 and 0.845, respectively (Table 1), and 91% of predicted EC10s and 86% of predicted EC20s were within a factor of 2.0 of observed (Supplemental Data, Figure S1C for EC10s; Figure 4C for EC20s).

Like *C. dubia*, Al toxicity data are limited for combinations of high DOC and pH, so the ln(DOC)×pH interaction term was excluded from the model analysis. When the same 3 interaction terms added to the *C. dubia* model were added to the *P. promelas* model, the AIC and the BIC both included ln(H)×pH for the EC10 and EC20 models. The adjusted R^2 values for the EC10 and EC20 models were 0.850 and 0.874, respectively (Table 1); 95% of predicted EC10s and 91% of predicted EC20s were within a factor of 2.0 of observed (Supplemental Data, Figure S1C for EC10s; Figure 4C for EC20s). As for *P. subcapitata* and *C. dubia*, no clear patterns were observed in the residuals over a wide range of Al bioavailability conditions as expressed by the magnitudes of the EC10s and EC20s (Supplemental Data, Figure S3E,F) or relative to single independent variables (Supplemental Data, Figure S4E,F; S5E,F; and S6E,F). Examples of how the MLR model for *P. promelas* EC20s performs relative to empirical toxicity data under conditions where either DOC, pH, or hardness was varied are provided in Figure 3. In the Cu MLR evaluation described in Brix et al. [10], inclusion of interaction terms in the models only moderately improved the adjusted R^2 values (e.g., 0.76 without interaction terms and 0.80 with interaction terms for *P. promelas*). Consequently, Cu MLR models without interaction terms were recommended. For the Al MLR models for *P. promelas*, inclusion of the ln(H)×pH interaction term likewise moderately improved the adjusted R^2 values for the EC10 and EC20 models (adjusted R^2 values increased from 0.822 to 0.850 for the EC10 model and from 0.845 to 0.874 for the EC20 model). We would normally not

recommend use of an interaction term for a small increase in the R^2 value, but in this case the interaction term was retained given the unique chemistry of Al that supports a reduced hardness effect at higher pH levels [9]. To empirically verify this, one Al toxicity test was conducted with *P. promelas* at a pH of 8 and in a high hardness water (127 mg/L); the EC10 and EC20 from this test were overpredicted by a factor of 2.7 and 3.0, respectively, when the interaction term was not retained (i.e., the effect of hardness at pH 8 was greatly overestimated when the ln(H)×pH interaction term was not included).

Final Al MLR models

Following are the final selected Al MLR models for *P. subcapitata*, *C. dubia*, and *P. promelas*. The models are expressed as in Equation 2, with slopes and intercepts from Table 1. The EC10s and EC20s are in units of micrograms of Al per liter, and DOC and hardness (H) are in units of milligrams per liter.

The final selected Al MLR models for *P. subcapitata* EC10s and EC20s were those determined based on the AIC and BIC, with interaction terms included:

$$\begin{aligned} P.\ subcapitata\ \text{Al ln(EC10)} = {} & -77.283 + 2.342 \times \ln[\text{DOC}] \\ & + 20.923 \times \text{pH} + 4.560 \times \ln[\text{H}] - 1.274 \times \text{pH}^2 \\ & - 0.288(\ln[\text{DOC}] \times \text{pH}) - 0.628 \times (\ln[\text{H}] \times \text{pH}) \end{aligned} \quad (6)$$

$$\begin{aligned} P.\ subcapitata\ \text{Al ln(EC20)} = {} & -61.952 + 1.678 \times \ln[\text{DOC}] \\ & + 17.019 \times \text{pH} + 4.007 \times \ln[\text{H}] - 1.020 \times \text{pH}^2 \\ & - 0.204 \times (\ln[\text{DOC}] \times \text{pH}) - 0.556 \times (\ln[\text{H}] \times \text{pH}) \end{aligned} \quad (7)$$

For *C. dubia*, the AIC and BIC selected different EC10 and EC20 MLR models when interaction terms were included, with the AIC model including more terms. Both models had similar adjusted R^2 values and provided a comparable level of accuracy in EC10 and EC20 predictions. We selected the AIC model because it provided slightly more accurate predictions in high Al bioavailability waters:

$$\begin{aligned} C.\ dubia\ \text{Al ln(EC10)} = {} & -51.420 + 1.208 \times \ln[\text{DOC}] \\ & + 13.338 \times \text{pH} + 3.639 \times \ln[\text{H}] - 0.737 \times \text{pH}^2 - 0.155 \\ & \times (\ln[\text{DOC}] \times \ln[\text{H}]) - 0.490 \times (\ln[\text{H}] \times \text{pH}) \end{aligned} \quad (8)$$

$$\begin{aligned} C.\ dubia\ \text{Al ln(EC20)} = {} & -41.026 + 0.525 \times \ln[\text{DOC}] \\ & + 11.282 \times \text{pH} + 2.201 \times \ln[\text{H}] - 0.663 \times \text{pH}^2 - 0.264 \\ & \times (\ln[\text{H}] \times \text{pH}) \end{aligned} \quad (9)$$

The final selected Al MLR models for *P. promelas* EC10s and EC20s were those determined based on the AIC and BIC, with interaction terms included:

$$\begin{aligned} P.\ promelas\ \text{Al ln(EC10)} = {} & -12.009 + 0.545 \times \ln[\text{DOC}] \\ & + 2.794 \times \text{pH} + 3.201 \times \ln[\text{H}] - 0.460 \times (\ln[\text{H}] \times \text{pH}) \end{aligned} \quad (10)$$

$$\begin{aligned} P.\ promelas\ \text{Al ln(EC20)} = {} & -14.029 + 0.503 \times \ln[\text{DOC}] \\ & + 3.131 \times \text{pH} + 3.443 \times \ln[\text{H}] - 0.494 \times (\ln[\text{H}] \times \text{pH}) \end{aligned} \quad (11)$$

Species and genus sensitivity distributions and HC5s

Aluminum EC10s were available for 14 species (one alga, one macrophyte, 9 invertebrates, and 3 fish), and EC20s were available for 8 genera of invertebrates and 3 genera of fish (Supplemental Data, Table S1). Using Equation 5, the *P. subcapitata*, *C. dubia*, and *P. promelas* MLR model slopes presented above (Table 1, Equations 6–11) were used to develop chronic Al SSDs and GSDs that were adjusted to a range of DOC, pH, and hardness combinations. The ranges of DOC, pH, and hardness conditions were constrained to those in the toxicity tests used to derive the MLR models: DOC of 0.1 to 5.0 mg/L, pH of 6.0 to 8.1, and hardness of 10 to 127 mg/L. The lower bounds for DOC and hardness encompass most natural waters [19], whereas the upper bounds for DOC and hardness would be conservative for waters with higher DOC and hardness concentrations. The pH range of 6.0 to 8.1 likewise captures the range of most natural waters. There is greater uncertainty in whether assuming a pH of 8.1 is conservative for waters with a pH >8.1 because Al solubility increases above a pH of 8 and there is some evidence that Al is more toxic at a pH of 8 than circumneutral pH. This is particularly apparent for *P. subcapitata* (Figure 1), which has been more rigorously tested at pH 8. The pattern is not clear for *C. dubia* and *P. promelas*, but data for these species at high pH are limited (Figures 2 and 3).

To evaluate how the Al HC5s vary with DOC, pH, and hardness concentration, HC5s were calculated as a function of one parameter being varied and the other 2 held constant. In these examples, HC5s were calculated using the EC20 models and following the USEPA approach. The most noticeable observations are that the HC5s consistently increase with increasing DOC (Figure 5A–C) and with increasing pH (Figure 5D–F). The influence of hardness on HC5s is variable depending on pH. Overall, HC5s increase with increasing hardness at pH 6, remain essentially constant at pH 7, and show a variable pattern at pH 8 (Figure 5G–I). These trends generally follow the empirical data where available, which is not unexpected given that the MLR models were derived solely from those data. However, fewer empirical toxicity data are available to evaluate the HC5 trends at pH 8. For example, the observation that HC5s at pH 8 decrease with increasing hardness appears to be consistent with data for *P. subcapitata* (Figure 1G–I) but less clearly so for *C. dubia* based on more limited data (Figure 2), and insufficient data are available for *P. promelas*. A spreadsheet for calculating MLR-normalized HC5s for DOC, pH, and hardness conditions of interest is provided as Supplemental Data, Table S2.

The other issue influencing the relationships between HC5s and water chemistry parameters is that the relative sensitivities of species vary with water chemistry. In general, brook trout (*Salvelinus fontinalis*) is the most sensitive species over the greatest range of DOC, pH, and hardness conditions (Supplemental Data, Figure S7A–C). Under some conditions, the amphipod *Hyalella azteca* is more sensitive than brook trout, particularly at pH 8 and hardness concentrations up to about 90 mg/L (Supplemental Data, Figure S7D–F). Further, toward the

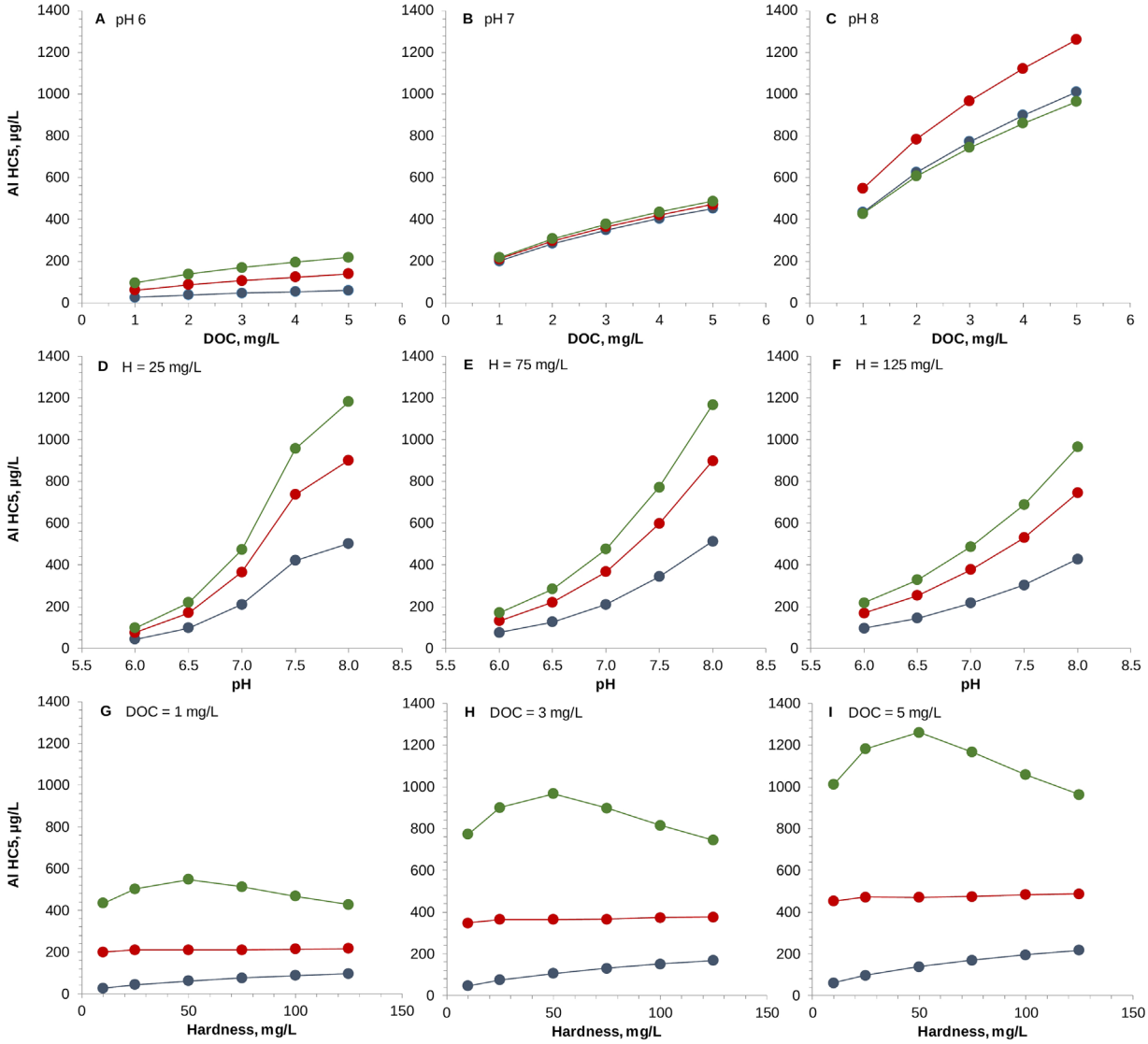

FIGURE 5: Total Al 5% hazardous concentrations as a function of dissolved organic carbon (DOC) concentration (**A–C**), pH (**D–F**), and hardness (**G–I**). (**A–C**) Hardness of 10, 50, and 125 mg/L (blue, red, and green symbols, respectively). (**D–F**) Dissolved organic carbon of 1, 3, and 5 mg/L (blue, red, and green symbols, respectively). (**G–I**) pH of 6, 7, and 8 (blue, red, and green symbols, respectively). H = hardness; HC5 = 5% hazardous concentration.

upper bounds of the DOC, pH, and hardness conditions for the model, the alga *P. subcapitata* is slightly underprotected by the HC5 (Supplemental Data, Figure S7F). Overall, as an evaluation on the protectiveness of the models, the MLR-based HC5s were protective of empirical EC20s for all species (Supplemental Data, Figure S8). Additional chronic Al toxicity testing at higher pH levels would be useful for further validating and extending the MLR models.

As noted earlier, the USEPA's current chronic criterion for Al is 87 µg/L [3]. That chronic criterion is based on a 60-d brook trout study and a 7-d striped bass study in which 88 and 87.2 µg Al/L, respectively, were identified as no-effect concentrations. The 60-d brook trout study [20], which was included in the present evaluation, had a test water with a DOC concentration of 1.9 mg/L, a pH of 6.5, and a hardness of 14.7 mg/L. The EC20 derived for that study was 156 µg/L (the no-effect concentration of 88 µg/L was associated with a 4% reduction in weight; the low-effect concentration was 169 µg/L, which was associated with a 24% reduction in weight). Under these DOC, pH, and hardness conditions, we calculated an MLR-based HC5 of 117 µg/L following the USEPA methodology. Reducing the DOC concentration to 1 mg/L results in an MLR-based HC5 of 85 µg/L. As such, it is interesting that the MLR-based approach predicts HC5s for low DOC, pH, and hardness conditions that are consistent with the USEPA's current chronic criterion.

Lastly, it is critical to note that although the Al toxicity data used to develop the MLR models and HC5s described in the present evaluation are based on total Al concentrations in laboratory waters, it is inappropriate to analyze total Al concentrations in natural waters for comparison. This is because many natural waters contain Al in mineral forms, such as clays and other suspended particles, which are nonbioavailable [1]. Methods for analyzing dissolved Al and colloidal precipitates in natural waters are needed. Research is currently under way for

developing such a method (P.H. Rodriguez, Centro de Investigación Minera y Metalúrgica, Santiago, Chile, personal communication).

In summary, MLR models could predict, within a factor of 2.0, chronic Al EC10s and EC20s, as a function of DOC, pH, and hardness concentration, in ≥91% of the tests with an alga (*P. subcapitata*), a cladoceran (*C. dubia*), and the fathead minnow (*P. promelas*). These MLR models were then used to adjust species and genus sensitivity distributions, following European Union and USEPA methodologies, respectively, to derive chronic HC5s that varied depending on the DOC, pH, and hardness concentrations of interest. Aluminum MLR models provide a viable method for predicting Al toxicity to freshwater aquatic organisms over a range of water chemistry conditions.

Supplemental Data—The Supplemental Data are available on the Wiley Online Library at DOI: 10.1002/etc.3922.

Acknowledgment—We gratefully acknowledge the thoughtful comments from anonymous reviewers that improved the manuscript. We also thank B. Santore and A. Ryan for their comments on earlier versions of the manuscript and Rio Tinto for supporting this effort.

Data availability—Please contact the corresponding author (DavidD@windwardenv.com) for any metadata or calculation tools not already provided as Supplemental Data.

REFERENCES

[1] Gensemer RW, Playle RC. 1999. The bioavailabity and toxicity of aluminum in aquatic environments. *Crit Rev Env Sci Tec* 29:315–450.

[2] Wilson RW. 2012. Aluminum. In Wood CM, Farrell AP, Brauner CJ, ed, *Homeostatis and Toxicology of Non-Essential Metals, Vol 31A—Fish Physiology*. Elsevier, Amsterdam, The Netherlands, pp 67–123.

[3] US Environmental Protection Agency. 1988. Ambient water quality criteria for aluminum. EPA 440/5-86-008. Washington, DC.

[4] Canadian Council of Ministers of the Environment. 2007. Canadian water quality guidelines for the protection of aquatic life: Summary table. Winnipeg, MB, Canada.

[5] Australian and New Zealand Environment and Conservation Council. 2000. Australia and New Zealand guidelines for fresh and marine water quality. Canberra, Australia.

[6] US Environmental Protection Agency. 1999. National recommended water quality criteria—Correction. EPA 822-Z-99-001. Washington, DC.

[7] Freund JG, Petty JT. 2007. Response of fish and macroinvertebrate bioassessment indices to water chemistry in a mined Appalachian watershed. *Environ Manage* 39:707–720.

[8] Herrmann J. 2001. Aluminum is harmful to benthic invertebrates in acidified waters, but at what thresholds? *Water Air Soil Poll* 130:837–842.

[9] Gensemer RW, Gondek JC, Rodriguez PH, Arbildua JJ, Stubblefield WA, Cardwell AS, Santore RC, Ryan AC, Adams WJ, Nordheim E. 2018. Evaluating the effects of pH, hardness, and dissolved organic carbon on the toxicity of aluminum to aquatic organisms under circumneutral conditions. *Environ Toxicol Chem* 37:49–60 (*this issue*).

[10] Brix KV, DeForest DK, Tear L, Grosell M, Adams WJ. 2017. Use of multiple linear regression models for setting water quality criteria for copper: A complementary approach to the biotic ligand model. *Environ Sci Technol* 51:5182–5192.

[11] Cardwell AS, Adams WJ, Gensemer RW, Nordheim E, Santore RC, Ryan AC, Stubblefield WA. 2018. Chronic toxicity of aluminum, at a pH of 6, to freshwater organisms: Empirical data for the development of international regulatory standards/criteria. *Environ Toxicol Chem* 37:36–48 (*this issue*).

[12] Stephan CE, Mount DI, Hansen DJ, Gentile JR, Chapman GA, Brungs WA. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. PB85-227049. US Environmental Protection Agency, National Technical Information Service, Springfield, VA.

[13] Santore RC, Ryan AC, Kroglund F, Rodriquez P, Stubblefield WA, Cardwell AS, Adams WJ, Nordheim E. 2018. Development and application of a biotic ligand model for predicting the toxicity of dissolved and precipitated aluminum to aquatic organisms. *Environ Toxicol Chem* 37:70–79 (*this issue*).

[14] Zuur AF, Ieno EN, Elphick CS. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1:3–14.

[15] Burnham KP, Anderson DR. 2004. Multimodel interference: Understanding AIC and BIC in model selection. *Sociol Method Res* 33:261–304.

[16] Dziak JJ, Coffman DL, Lanza ST, Li R. 2012. Sensitivity and specificity of information criteria. Technical Report Series 12-119. Pennsylvania State University, State College, PA, USA.

[17] R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[18] European Chemicals Bureau. 2003. Technical guidance document on risk assessment. Ispra, Italy.

[19] US Environmental Protection Agency. 2016. Draft technical support document: Recommended estimates for missing water quality parameters for application in EPA's biotic ligand model. EPA 820-R-15-106, Washington, DC.

[20] Cleveland L, Little EE, Wiedmeyer RH, Buckler DR. 1989. Chronic no-observed-effect concentrations of aluminum for brook trout exposed in low-calcium, dilute acidic water. In Lewis TE, ed, *Environmental Chemistry and Toxicology of Aluminum*. Lewis, Chelsea, MI, pp 229–245.

