Environmental Science & Technology

Article

pubs.acs.org/est

Use of Multiple Linear Regression Models for Setting Water Quality Criteria for Copper: A Complementary Approach to the Biotic Ligand Model

Kevin V. Brix,*[,†,‡] David K. DeForest,[§] Lucinda Tear,[§] Martin Grosell,[‡] and William J. Adams[∥]

[†]EcoTox, 2263 SW 37th Avenue, #816, Miami, Florida 33145, United States

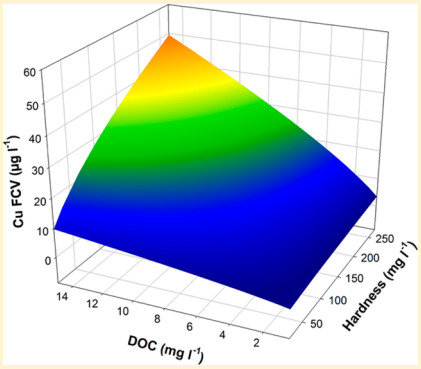
[‡]University of Miami, RSMAS 4600 Rickenbacker Causeway, Miami, Florida 33149, United States

[§]Windward Environmental, 200 West Mercer Street, Suite 401, Seattle, Washington 98119, United States

[∥]Red Cap Consulting, 7760 North Boulder Drive Lake Point, Utah 84074, United States

Supporting Information

ABSTRACT: Biotic Ligand Models (BLMs) for metals are widely applied in ecological risk assessments and in the development of regulatory water quality guidelines in Europe, and in 2007 the United States Environmental Protection Agency (USEPA) recommended BLM-based water quality criteria (WQC) for Cu in freshwater. However, to-date, few states have adopted BLM-based Cu criteria into their water quality standards on a state-wide basis, which appears to be due to the perception that the BLM is too complicated or requires too many input variables. Using the mechanistic BLM framework to first identify key water chemistry parameters that influence Cu bioavailability, namely dissolved organic carbon (DOC), pH, and hardness, we developed Cu criteria using the same basic methodology used by the USEPA to derive hardness-based criteria but with the addition of DOC and pH. As an initial proof of concept, we developed stepwise multiple linear regression (MLR) models for species that have been tested over wide ranges of DOC, pH, and hardness conditions. These models predicted acute Cu toxicity values that were within a factor of ±2 in 77% to 97% of tests (5 species had adequate data) and chronic Cu toxicity values that were within a factor of ±2 in 92% of tests (1 species had adequate data). This level of accuracy is comparable to the BLM. Following USEPA guidelines for WQC development, the species data were then combined to develop a linear model with pooled slopes for each independent parameter (i.e., DOC, pH, and hardness) and species-specific intercepts using Analysis of Covariance. The pooled MLR and BLM models predicted species-specific toxicity with similar precision; adjusted R^2 and R^2 values ranged from 0.56 to 0.86 and 0.66−0.85, respectively. Graphical exploration of relationships between predicted and observed toxicity, residuals and observed toxicity, and residuals and concentrations of key input parameters revealed many similarities and a few key distinctions between the performances of the two models. The pooled MLR model was then applied to the species sensitivity distribution to derive acute and chronic criteria equations similar in form to the USEPA's current hardness-based criteria equations but with DOC, pH, and hardness as the independent variables. Overall, the MLR is less responsive to DOC than the BLM across a range of hardness and pH conditions but more responsive to hardness than the BLM. Additionally, at low and intermediate hardness, the MLR model is less responsive than the BLM to pH, but the two models respond comparably at high hardness. The net effect of these different response profiles is that under many typical water quality conditions, MLR- and BLM-based criteria are quite comparable. Indeed, conditions where the two models differ most (high pH/low hardness and low pH/high hardness) are relatively rare in natural aquatic systems. We suggest that this MLR-based approach, which includes the mechanistic foundation of the BLM but is also consistent with widely accepted hardness-dependent WQC in terms of development and form, may facilitate adoption of updated state-wide Cu criteria that more accurately account for the parameters influencing Cu bioavailability than current hardness-based criteria.



INTRODUCTION

Over the past 25 years, the level of sophistication with which the toxicity and bioavailability of metals to aquatic organisms has been assessed, and corresponding environmental regulations derived have progressively increased. In the mid-1980s, the US Environmental Protection Agency (USEPA) released a series of water quality criteria (WQC) for metals. For many of the metals (e.g., Cd, Cu, Ni, Pb, Zn), WQC were expressed as a function of water hardness using a simple linear equation of log-transformed independent (hardness) and dependent (toxicity) variables. This equation was derived by first

Received: November 3, 2016
Revised: March 30, 2017
Accepted: April 14, 2017
Published: April 14, 2017

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

developing species-specific regressions of toxicity versus hardness. Then, for species for which a significant relationship was observed, data were normalized as a function of the mean toxicity value, and then all normalized data were pooled to derive a slope applicable to all species.[1] These hardness-dependent WQC were implemented by State regulatory agencies across the US and in many other countries.

More recently, the Biotic Ligand Model (BLM) has been developed to consider water quality variables other than hardness (e.g., dissolved organic carbon [DOC], Ca, Mg, pH).[2−5] This quasi-mechanistic model considers water quality parameters that affect metal speciation (e.g., pH), can complex free metal ion (e.g., DOC), or can compete with metals at a site of uptake by the organism (e.g., Ca^{2+}, H^+). Using data for these parameters, the model predicts metal accumulation at an undefined site of toxic action (assumed to be the gill or homologous respiratory organ) in the organism and corresponding toxic effects.[3]

Numerous versions of the BLM have been developed for use in regional risk assessments in Europe for Cu, Pb, Ni and Zn,[6,7] and development of BLMs for other metals is ongoing (e.g., Al, Co). In the United States, USEPA has derived BLM-based WQC for Cu.[8] However, despite these WQC having been recommended by USEPA more than 8 years ago, few states have fully adopted the BLM-based WQC into state water quality standards, although a number of states now allow use of the BLM to set site-specific water quality standards. Consequently, Cu and many other metals continue to be regulated using the previously developed hardness-dependent WQC. Similarly, in Canada, despite recent revisions to water quality guidelines for several metals at both the national and provincial level, no BLM-based water quality guideline has been proposed. Numerous reasons likely exist for regulatory agencies to not adopt BLM-based WQC, but a significant factor appears to be the perception that the BLM is too complicated and not sufficiently transparent for regulators to understand and implement. Additionally, the BLM currently requires measurement of ten water quality parameters compared with the hardness-based criteria that require only one. We believe there is also a perception that the BLM will result in less stringent WQC although recent analyses suggest that the BLM-based WQC for Cu are more protective than the hardness-based WQC, which may underprotect important species and endpoints under certain water chemistry conditions.[9,10]

The current study develops a statistical model for Cu that considers the key water quality parameters identified by the BLM but is sufficiently simplified both in terms of the mechanics of the model and number of water quality parameters considered, that regulatory agencies might consider its implementation. We selected a simple, multiple linear regression (MLR) approach that considers 3 water quality parameters (hardness, pH, DOC) and uses the same conceptual framework currently used for deriving hardness-based WQC. The final model is similar in form to existing hardness-based WQC but with 3 parameters instead of one. We then compared the performance of the MLR-based WQC to the more mechanistically robust BLM-based WQC for Cu.

METHODS AND MATERIALS

Toxicity Data Sets. Acute and chronic toxicity data sets for Cu were assembled using the most recent updates to USEPA WQC.[8] These data sets were amended with additional data from recent publications that met USEPA guidelines for test acceptability,[1] with several additional studies identified for Cu[11−17] (Tables S2 and S4). If a given toxicity test lacked information on water quality parameters identified to affect bioavailability, the data were excluded from further analysis.

Data Analysis. Previous efforts to use multiple linear regression to describe the effects of water quality variables on metal toxicity have generally relied on a stepwise approach in which all water quality variables for which data were available were entered into the model.[18,19] We chose not to use this approach for water quality parameter selection, but rather used information from existing Cu BLMs[8,20] to *a priori* select the water quality variables that would be assessed. This approach retains the mechanistic underpinnings provided by the BLM, ensuring that parameters known to be important were included, even if some data sets did not provide a sufficient range to allow for statistical discrimination of their importance, and excludes parameters that might be retained in the MLR due to artifacts such as covariance with a parameter known to effect toxicity. We concluded that hardness, DOC, and pH were the key water quality parameters that met these criteria. For the development of species-specific models, we required toxicity tests span a minimum range of water chemistries: 100 mg L^{-1} for hardness, 5 mg L^{-1} for DOC, and 1.5 pH units.

MLR models were developed both for individual species using species-specific data and for all species using a pooled, all-species data set. Relationships between toxicity and hardness or DOC tend to be linear, while for pH (which is a log-transformed variable) a linear relationship is observed when toxicity is log-transformed. Log transformation also has the effect of homogenizing variance in the residuals around the model along the range of the data. Consequently, toxicity and water quality parameters were log transformed, except pH.

For both the species-specific and pooled models, stepwise linear regression methods were used to determine the minimum number of significant independent variables in models with just the main independent variables, ln(DOC), ln(hardness), and pH, and in models including the main independent variables and all two-way interactions. For a species-specific model, the basic model form was

$$\ln(\text{toxicity}) = b_0 + b_1{*}\ln(\text{DOC}) + b_2{*}\ln(\text{hardness}) + b_3{*}\text{pH} + \text{error} \quad (1)$$

and the model with interactions was of the form

$$\ln(\text{toxicity}) = b_0 + b_1{*}\ln(\text{DOC}) + b_2{*}(\text{hardness}) + b_3{*}\text{pH} + b_4{*}[\ln(\text{DOC}) \times \text{pH}] + b_5{*}[\ln(\text{DOC}) \times \ln(\text{hardness})] + b_6{*}[\ln(\text{hardness}) \times \text{pH}] + \text{error} \quad (2)$$

where DOC and hardness are expressed in mg L^{-1}, and toxicity is expressed in μg L^{-1} Cu.

An Analysis of Covariance (ANCOVA) was conducted to test for differences between species-specific coefficients and the respective mean coefficient of all species in the basic model. The analysis was considered exploratory only, since it is not clear how differences among species-specific coefficients for any individual model term would influence toxicity predictions from the full, multi-term model.

The form of this ANCOVA involves the addition of a "species" term to the model and is run using the data for all species. The request for species-specific coefficients is specified in R by adding an interaction between species and each independent variable. We specified a deviance contrast matrix

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

for species that would test the value of each species-specific coefficient against the mean of the other species' coefficients for each term (see Supplemental Table S10 for model code). The model produces the same species-specific coefficients as the individual basic models (eq 1), along with additional information about how the species' coefficients differ and can be stated as

$$\ln(\text{toxicity}) = (b_0 + b_{0i}) + (b_1 + b_{1i})*\ln(\text{DOC}) + (b_2 + b_{2i}) *\ln(\text{hardness}) + (b_3 + b_{3i})*\text{pH} + \text{error} \quad (3)$$

where coefficients b_0–b_3 define the mean coefficient of all species, and coefficients b_{0i}–b_{3i} define the difference between the mean coefficients across species and the coefficient of each species (i).

Finally, two ANCOVAs, a basic model and a model with interactions, were conducted to select a model to be used in WQC calculations. These models contained an all-species slope for each independent variable and species-specific intercepts (see Supplemental Table S11 for model code).

$$\ln(\text{toxicity}) = (b_0 + b_{0i}) + b_1*\ln(\text{DOC}) + b_2 *\ln(\text{hardness}) + b_3*\text{pH} \quad (4)$$

$$\ln(\text{toxicity}) = (b_0 + b_{0i}) + b_1*\ln(\text{DOC}) + b_2*\ln(\text{hardness}) + b_3*\text{pH} + b_4*\ln(\text{DOC})*\text{pH} + b_5*\ln(\text{DOC})*\ln(\text{hardness}) + b_6 *\ln(\text{hardness})*\text{pH} \quad (5)$$

The results of these models are very similar to model results computed using the data-normalizing approach described in ref 1 but do not require the complication of back-transforming results to compute a species-specific intercept and provide additional statistical information about the species-specific differences in coefficients (see the Supporting Information for a more detailed comparison).

Best-fitting models were identified using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).[21] Similar to the familiar adjusted R^2, each information criterion balances model specificity and generality by penalizing a model's goodness of fit term ($-2l$) by a factor related to the number of parameters (p) in the model in order to achieve parsimony − a model that uses the fewest terms to fit the data set while remaining adequately general so as to also fit other data sets from the same population. The penalty used by the AIC, $2p$, includes only the number of parameters (eq 6), while the penalty used by the BIC ($\ln(n)p$) also includes sample size (eq 7)

$$\text{AIC} = -2l + 2p \quad (6)$$

$$\text{BIC} = -2l + \ln(n)p \quad (7)$$

where l = log likelihood (model fit term).

The model with the smallest AIC or BIC is considered the best-fitting model. Depending on sample size, the AIC and BIC can result in the same or different final models and differences of opinion exist in the literature as to which is the more valid criterion in different situations. Here we simply report when the criteria concur and when they do not. Stepwise regressions were run in R using the stepAIC() function from the MASS library (e.g., for BIC model: stepAIC (model, direction = c("both"), k = 1n(n)).

Variance Inflation Factors (VIFs) were calculated for the main independent variables (see the Supporting Information) to assess collinearity. VIFs quantify the increase in variance in the estimate of a covariate that is created by its correlation with other variables in the model.[22] Correlation among independent variables makes it difficult to determine the driving factor among the collinear variables, and focused model analysis is required to interpret model results when correlated independent variables are included in a model. VIFs were calculated in R using the vif() function in the usdm library.

Comparison of MLR and BLM Predictions and Goodness of Fit. Fits of the species-specific BLM and MLR models were assessed visually by plotting predicted vs observed toxicity for both models along a one-to-one line for each species. Differences between models in the magnitude and shape of the dispersion around the one-to-one line and proportion of data falling outside the range were visually compared. For each species, variance explained by the two models was compared using adjusted R^2 for the pooled MLR model and R^2 for the BLM model. Adjusted R^2 for the MLR model was calculated using r = correlation between log-transformed predicted and observed values, n = species sample size, and p = 3. The R^2 for the BLM was not adjusted because there are no fitted parameters in the model. Model fits of the pooled MLR and BLM were also assessed by plotting model residuals against observed toxicity and against each of the independent variables by species. Patterns in the residuals around the dependent and independent variables provide information about situations in which the models might be over- or underpredicting toxicity. Tests of model assumptions for the MLR, including homogeneity of variance and normality of residuals, were assessed using the R functions ncvTest() and Shapiro.test(), respectively. QQ plots of species-specific residuals (around the pooled model predictions) were also reviewed to provide insight about normality of residuals.

To allow direct comparison of BLM and MLR model predictions for the pooled models, differences between the predicted values for the two models ($\text{MLR}_{\text{predicted}} - \text{BLM}_{\text{predicted}}$) were plotted against observed toxicity. The species-specific and pooled MLR model fits were also compared by plotting the differences between the predicted values for the two models ($\text{MLR}_{\text{species}} - \text{MLR}_{\text{pooled}}$) as a function of observed toxicity.

Derivation of MLR-Based Copper Criteria. The slopes of the pooled MLR model for acute toxicity were then used to derive acute MLR-based criteria following the same methodology for deriving hardness-based criteria, as described in ref 1. The USEPA's recommended WQC are derived using a species sensitivity distribution (SSD) approach. To calculate an acute criterion, 48- to 96-h LC50 values for aquatic animals are first compiled (Table S2). Species mean acute values (SMAVs) are calculated as the geometric mean LC50 for each species, and then genus mean acute values (GMAVs) are calculated as the geometric mean of SMAVs (Table S3). The acute criterion is the fifth percentile of the GMAVs, divided by two to provide an "EC-low" effect level.

Each acute toxicity value was first standardized to a common water chemistry condition using the DOC, hardness, and pH slopes from the pooled model (target DOC = 0.5 mg L^{-1}, target hardness = 85 mg L^{-1}, and target pH = 7.5)

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

Table 1. Species-Specific MLR Model Coefficients[a]

species	interactions	statistic	ln(Int.)	ln(DOC)	ln(Hard)	pH	ln(DOC) × pH	ln(DOC) × ln(Hard)	ln(Hard) × pH	*n*	adj. r^2
Acute											
C. dubia	no	AIC	−3.901	0.625		0.877				87	0.63
	no	BIC	−3.901	0.625		0.877				87	0.63
	yes	AIC	−9.535	6.703	0.144	1.511	−0.776			87	0.78
	yes	BIC	−9.082	6.431		1.536	−0.742			87	0.77
D. magna	no	AIC	−7.094	0.941	0.503	1.042				307	0.86
	no	BIC	−7.094	0.941	0.503	1.042				307	0.86
	yes	AIC	−4.005	0.947	−0.254	0.628			0.101	307	0.87
	yes	BIC	−7.094	0.941	0.503	1.042				307	0.86
D. obtusa	no	AIC	−2.853	0.843	0.233	0.670				53	0.82
	no	BIC	−2.853	0.843	0.233	0.670				53	0.82
	yes	AIC	−6.245	4.224	0.139	1.131	−0.353	−0.171		53	0.89
	yes	BIC	−6.245	4.224	0.139	1.131	−0.353	−0.171		53	0.89
D. pulex	no	AIC	−5.280	0.875	0.406	0.838				35	0.81
	no	BIC	−5.280	0.875	0.406	0.838				35	0.81
	yes	AIC	−9.932	6.931	0.172	1.502	−0.782			35	0.92
	yes	BIC	−12.423	7.386		1.894	−0.836			35	0.92
P. promelas	no	AIC	−6.000	0.670	1.030	0.855				206	0.76
	no	BIC	−6.000	0.670	1.030	0.855				206	0.76
	yes	AIC	−7.467	2.434	1.040	1.030	−0.130	−0.197		206	0.80
	yes	BIC	−6.744	1.620	1.065	0.925		−0.241		206	0.80
Chronic											
D. magna	no	AIC	0.200	0.848	0.235	0.172				77	0.87
	no	BIC	0.200	0.848	0.235	0.172				77	0.87
	yes	AIC	−5.512	0.858	1.351	0.908			−0.145	77	0.87
	yes	BIC	0.200	0.848	0.235	0.172				77	0.87
Pooled Acute											
	no	AIC	−6.52	0.786	0.582	0.966				688	0.81
	no	BIC	−6.52	0.786	0.582	0.966				688	0.81
	yes	AIC	−7.66	2.22	0.682	1.087	−0.118	−0.132		688	0.82
	yes	BIC	−7.66	2.22	0.682	1.087	−0.118	−0.132		688	0.82

[a]AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion, DOC = Dissolved Organic Carbon, Int. = Intercept, Hard = Hardness. Empty cells are variables or interaction terms that were evaluated but not significant for the model.

$$\text{standardized LC50} = \exp[\ln(\text{EC50}_{\text{meas}}) - \text{DOC}_{\text{slope}}(\ln[\text{DOC}_{\text{meas}}] - \ln[\text{DOC}_{\text{target}}]) - \text{Hard}_{\text{slope}}(\ln[\text{Hard}_{\text{meas}}] - \ln[\text{Hard}_{\text{target}}]) - \text{pH}_{\text{slope}}(\text{pH}_{\text{meas}} - \text{pH}_{\text{target}})] \quad (8)$$

where $\text{LC50}_{\text{meas}}$ = observed LC50; $\text{DOC}_{\text{slope}}$ = DOC slope from pooled MLR model; DOC_{meas} = test water DOC; $\text{DOC}_{\text{target}}$ = standardized DOC (0.5 mg L^{-1}); pH_{slope} = pH slope from pooled MLR model; pH_{meas} = test water pH; $\text{pH}_{\text{target}}$ = standardized pH (7.5); $\text{Hard}_{\text{slope}}$ = hardness slope from pooled MLR model; $\text{Hard}_{\text{meas}}$ = test water hardness; $\text{Hard}_{\text{target}}$ = standardized hardness (85 mg L^{-1}).

The standardized LC50s were then used to calculate SMAVs, GMAVs, and the fifth percentile GMAV. The standardized fifth percentile GMAV (divided by 2; see above) and pooled slopes for DOC, hardness, and pH were then used to calculate the intercept, which results in an acute criteria equation with the following form:

$$\text{acute criterion} = \exp(\text{DOC}_{\text{slope}}(\ln[\text{DOC}_{\text{meas}}]) + \text{Hard}_{\text{slope}}(\ln[\text{Hard}_{\text{meas}}]) + \text{pH}_{\text{slope}}(\text{pH}_{\text{meas}}) + \text{Intercept}) \quad (9)$$

The same general approach may be used for deriving chronic criteria if sufficient chronic toxicity data are available to meet the USEPA 8 family guideline rule.[1] Alternatively, an acute-chronic ratio (ACR) may be used to estimate the chronic fifth percentile GMAV. Chronic criteria based on both empirical chronic toxicity data and the ACR approach were explored (see Results) (Tables S4 and S5).

Comparison of MLR and BLM Models. Once the MLR-based WQC equation was developed, we compared its performance to the BLM-based WQC across a range of water quality conditions. A matrix of varying water chemistry was developed for hardness (17−271 mg L^{-1}), DOC (0.5−15 mg L^{-1}), and pH (6.0−8.5) (Table S6). The matrix was developed from an initial moderately hard synthetic water (Ca = 14 mg L^{-1}, Mg = 12.1 mg L^{-1}, Na = 26.3 mg L^{-1}, K = 2.1 mg L^{-1}, SO4 = 81.4 mg L^{-1}, Cl = 1.9 mg L^{-1}, alkalinity =65 mg L^{-1}). All ion concentrations, and alkalinity, were then proportionally increased or decreased to derive a range of water chemistry conditions. Consequently, major ions used by the BLM covaried within this matrix to test whether the BLM, which considers all of these ions in calculating a WQC, responded differently from the MLR, which only considers hardness, DOC, and pH. The influence of alkalinity that varied from hardness by a factor of ±1.5 was also evaluated for a subset of the data and found not to have a strong influence (generally 2−3%) on BLM-based CMCs (data not shown).

RESULTS

Species-Specific Models. Sufficient data were available to develop species-specific acute toxicity MLR models for

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

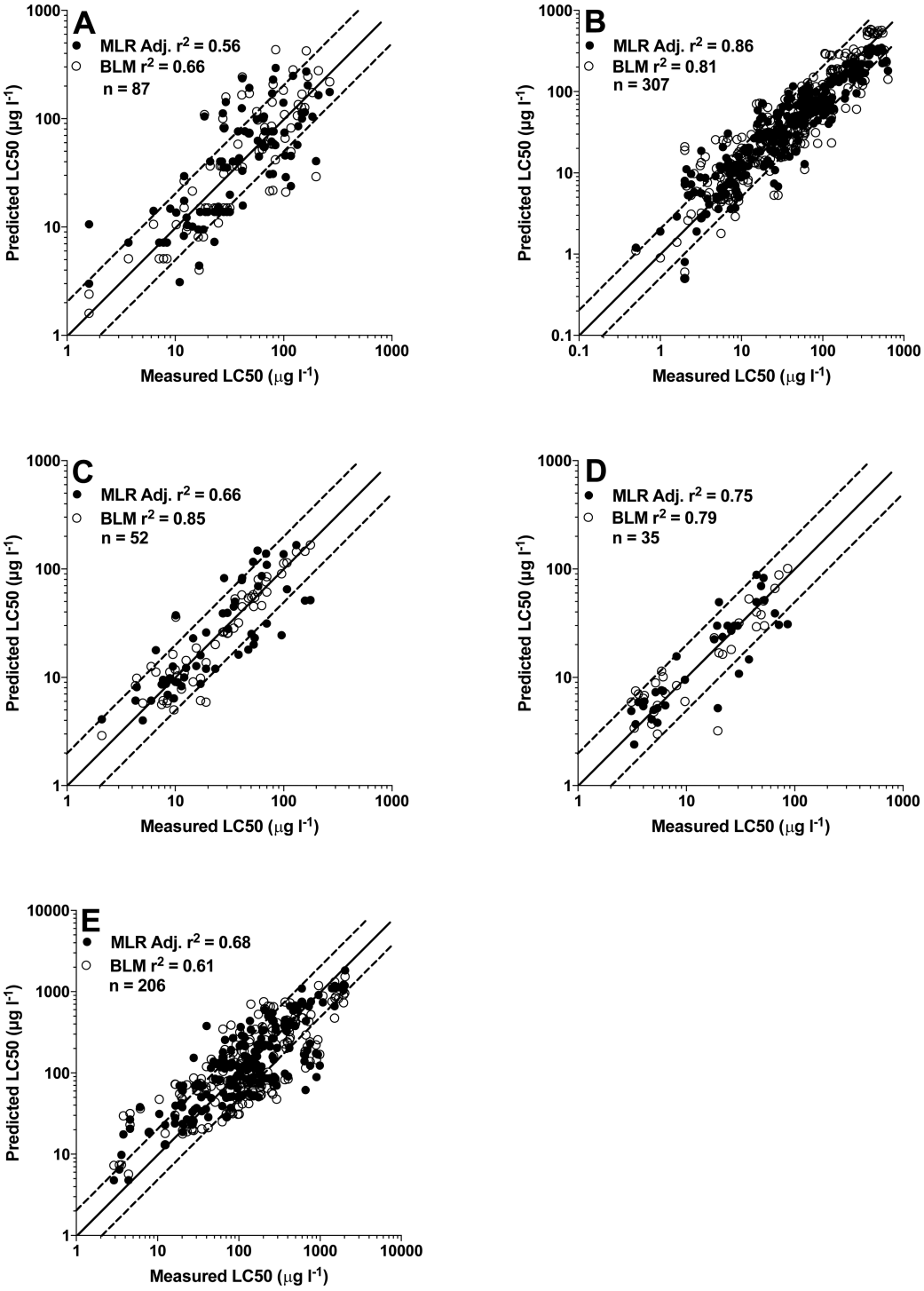


Figure 1. Comparison of measured versus MLR- and BLM-predicted LC50s for A) *C. dubia*, B) *D. magna*, C) *D. obtusa*, D) *D. pulex*, and E) *P. promelas*. All MLR predictions are based on the pooled model. Solid line represents line of perfect agreement between measured and predicted LC50s. Dashed lines represent ± a factor of 2 from the line of perfect agreement.

Ceriodaphnia dubia, Daphnia magna, Daphnia pulex, Daphnia obtusa, and *Pimephales promelas* (Table 1). Only *Daphnia magna* had sufficient chronic toxicity data to develop a species-specific chronic MLR model (Table 1). Both DOC and pH were retained in all models evaluated, while hardness was excluded from the acute *C. dubia* model without interactions and the BIC model for acute *D. pulex* with interactions (Tables 1, S7). VIFs were generally <2 for most variables in the species-specific models, indicating low correlation among the independent variables. Hardness and pH in the *D. obtusa* and *D. pulex* models had VIFs of 2.6−2.8 and 4.5−5.6, respectively (Table S8). These higher VIFs reflect the correlation between hardness and alkalinity in many natural and synthetic waters and the relationship between alkalinity and pH.

In general, species-specific models performed well in predicting metal toxicity, with adjusted R^2 values (based on log-transformed observed and predicted values) ranging between 0.63 and 0.92 (Table 1). As might be expected,

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

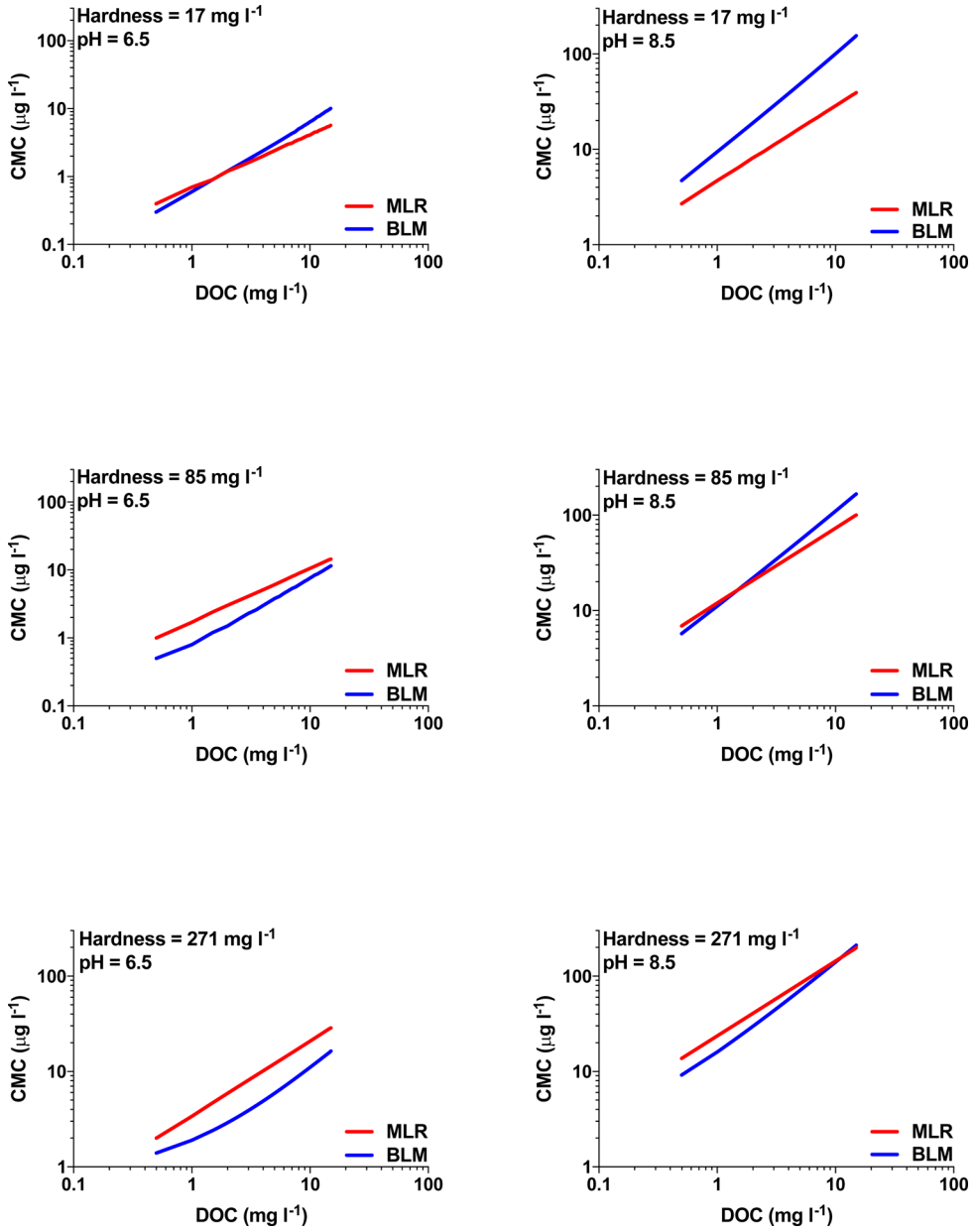


Figure 2. Criterion maximum concentrations (CMCs) for the MLR model and USEPA (2007) BLM updated with more recent toxicity data as a function of DOC at three hardness levels (17, 85, and 217 mg L^{-1}) at pH 6.5 and 8.5.

models with interaction terms performed better than models without interactions, but the improved performance was generally modest with an increase in adjusted R^2 ranging from 0.00 to 0.15. For all of the species-specific models without interactions, the AIC and BIC retained the same variables since all variables were either highly significant or nonsignificant. For models with interaction terms, significance levels were more widely distributed, and the AIC tended to retain more independent variables than the BIC (Table S7). Given the modest improvement in the adjusted R^2 when interaction terms were included, we focused on developing a pooled model without interaction terms in the remainder of our assessment.

Derivation of Pooled Model: Comparison to Species-Specific Models and the BLM. A total of five species-specific MLRs (4 invertebrates and 1 fish) were available for development of the pooled acute MLR model and derivation of an acute criterion equation. The exploratory ANCOVA performed to test for differences between species-specific slopes and intercepts and mean slopes and intercepts revealed that a few species had slopes different from the mean slope (Table S10). Specifically, the DOC slopes for *C. dubia and D. magna*, as well as the hardness slope for *C. dubia*, differed significantly from the mean MLR slope. While it is important to keep in mind these identified differences, the pooled model was still able to explain variability in Cu toxicity for individual species at a comparable level to that explained by the individual species models, as discussed further below.

The final pooled model (adj. $R^2 = 0.81$) with a slope for each independent variable and separate intercept for each species was

$$\ln(\mathrm{LC50}_{\mathrm{predicted}}) = \mathrm{intercept}_i + 0.786 \times \ln(\mathrm{DOC}) + 0.582 \times \ln(\mathrm{Hard}) + 0.966 \times \mathrm{pH} \quad (10)$$

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

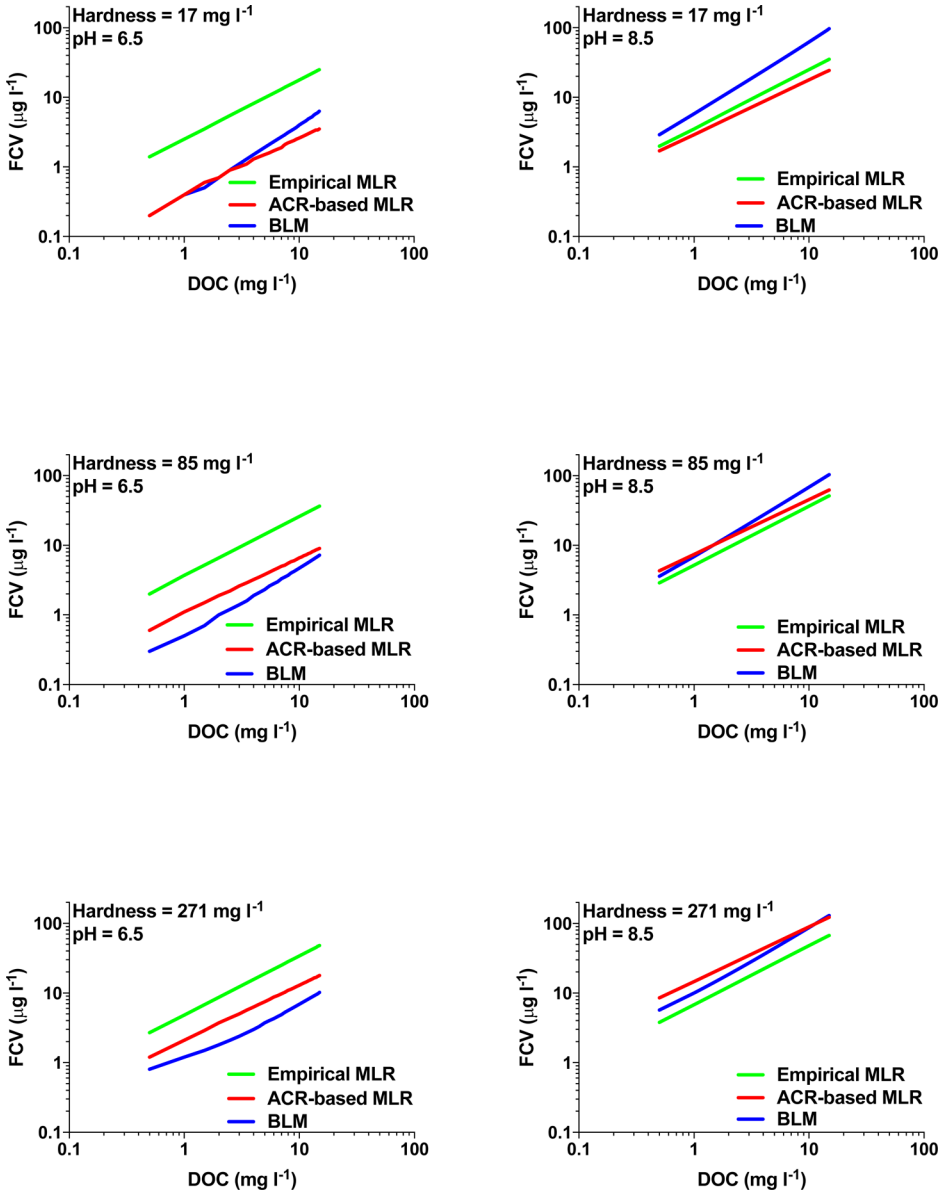


Figure 3. Final chronic values (FCVs) for 1) empirical chronic toxicity data and chronic MLR model, 2) acute toxicity data, pooled acute MLR model, and acute-chronic ratio (ACR) of 3.22, and 3) USEPA (2007) BLM updated with more recent toxicity data. FCVs plotted as a function of DOC at three hardness levels (17, 85, and 217 mg L^{-1}) at pH 6.5 and pH 8.5.

where the intercepts were −7.37 for *C. dubia*, −6.70 for *D. magna*, −6.46 for *D. obtusa*, −6.98 for *D. pulex*, and −5.11 for *P. promelas*.

The residuals of the pooled model had homogeneous variance (p = 0.92) but were not normally distributed based on the Shapiro Wilks test (p < 0.001). QQ plots of the pooled model residuals show deviation from the normal expectation predominantly in the tails of the distribution (Table S11).

The results of a stepwise regression using the pooled model with 2-way interaction terms among the independent variables did not substantially improve the fit of the model; adj. R^2 increased from 0.81 with just the main independent variables to 0.82 with the addition of ln(DOC)*pH and ln(DOC)*ln-(hardness) (Table S11). As would be expected, adjusted R^2 values for predictions of Cu toxicity to each of the individual species using the pooled model were lower than adjusted R^2 values from the species-specific models, but reductions were modest, ranging from 0.02 to 0.16 (Table 1, Figure S1). Both the ratio of pooled MLR to BLM predictions and species-specific to pooled MLR model predictions as a function of observed toxicity indicate there is no systematic bias in the differences between model predictions (Figures S2 and S3). The Shapiro Wilks tests indicated that species-specific residuals around both the pooled MLR and BLM predictions were normally distributed for *C. dubia* and *D. obtusa* (but not for *D. magna*, *D. pulex*, and *P. promelas* (Table S9)). QQ plots of the species-specific residuals around the pooled models show that the central portions of most of the species-specific residual distributions lie along the one-to-one line of expected normal quantiles with the tail portions of the distributions deviating slightly from expectations for some species (Figure S4).

Pooled MLR model and BLM fits differ among species; patterns in residuals with respect to observed toxicity are similar within a species for the two models but are not consistent across all species, indicating that neither model consistently over- or underestimates toxicity for all species (Figure S5). For

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

example, for both models, residuals for *C. dubia*, *D. obtusa*, and *D. pulex* appear relatively homogeneously distributed along the range of observed toxicity values (though variance is slightly higher for *C. dubia*). In contrast, residuals for *P. promelas* and, to a lesser extent, *D. magna* tend to be negative at lower EC50s (the model is overpredicting toxicity) and positive at higher EC50s (the model is underpredicting toxicity).

Residual patterns as a function of water chemistry also tend to be similar within a species for the two models and to vary between species, though there are exceptions. Most relationships between residuals and water chemistry are not homogeneously distributed, and some exhibit strong patterns (Figures S6–S8). For example, for both the MLR and BLM, the relationship of residuals to ln(hardness) is opposite for *C. dubia* and *P. promelas*. The relationship differs between models for *D. magna*, *D. obtusa*, and *D. pulex*. Inconsistent patterns of over- and underestimation across species indicate species may be responding differently to water chemistry and that neither model is fully accounting for all these differences.

Still, the precision of the pooled MLR model was quite comparable to the BLM, with differences in R^2 (BLM) versus adjusted R^2 (MLR) ranging from 0.05 to 0.20, with the pooled MLR model performing slightly better for 2 out of 5 species (compared to 3 out of 5 species for the species-specific MLR models) (Figure 1). We concluded that the pooled slopes model performed well and that the model was suitable for WQC development.

CMC Derivation and Comparisons between the MLR and BLM. The DOC, hardness, and pH slopes from the pooled model were used to standardize all LC50s in the acute SSD to the target water chemistry of interest (Table S2), as shown in eq 8. Following the methodology summarized above, and described in detail in ref 1, the resulting acute criteria equation is as follows (see also Table S3):

$$\text{acute Cu criteria} = \exp(0.786 \times \ln(\text{DOC}) + 0.582 \times \ln(\text{Hard}) + 0.966 \times (\text{pH}) - 8.32) \quad (11)$$

We then evaluated the MLR- and BLM-based acute criteria over a range of water quality conditions (0.5–15 mg L^{-1} DOC, 17–271 mg L^{-1} hardness, pH 6.5 and 8.5). The MLR-based acute criterion was comparable to the BLM-based criterion at low hardness and pH over a range of DOC concentrations. However, with increasing hardness the MLR-based criterion became increasingly greater than the BLM-based criterion at this same low pH (6.5). In contrast, at high pH (8.5), the BLM-based criterion was greater than the MLR-based criterion at low hardness, but the two models were comparable in intermediate and high hardness waters over a range of DOC concentrations (Figure 2).

Derivation of Chronic Criteria Equation. As noted above, only *D. magna* had sufficient data for development of a chronic MLR model. Following ref 1, however, data on the effects of water chemistry on toxicity for a single species is sufficient for deriving a WQC that varies as a function of water chemistry. The chronic MLR model for *D. magna* differed from the acute MLR model for this same species, mainly in that it had shallower slopes for hardness and pH (Table 1). While technically this meets the minimum data requirements in the USEPA guidelines we decided to derive chronic WQC based on application of an acute-chronic ratio (ACR) to the MLR-based acute WQC as well as directly from the chronic toxicity data. The available chronic Cu toxicity data do not meet the 8 family rule as data are lacking for a benthic crustacean, but there is precedent for deriving chronic WQC based on empirical chronic toxicity data even when the rule is not met.[23] Hence, we make such a calculation below to compare with the ACR-based approach.

For the ACR-based approach, we used the final Cu ACR of 3.22 from USEPA (2007), which resulted in the following chronic criteria eq (Table S3):

$$\text{chronic Cu criteria (ACR approach)} = \exp(0.786 \times \ln(\text{DOC}) + 0.582 \times \ln(\text{Hard}) + 0.966 \times (\text{pH}) - 8.79) \quad (12)$$

For the empirical chronic toxicity approach, the DOC, hardness, and pH slopes from the chronic *D. magna* model were used to standardize all chronic values in the chronic SSD to the target water chemistry of interest (Table S4). The resulting chronic criteria equation is as follows (Table S5):

$$\text{chronic Cu criteria (empirical chronic data approach)} = \exp(0.848 \times \ln(\text{DOC}) + 0.235 \times \ln(\text{Hard}) + 0.172 \times \text{pH} - 0.870) \quad (13)$$

Over a range of DOC (0.5–15 mg L^{-1}) and hardness (17–271 mg L^{-1}) conditions, at pH 6.5 and 8.5, the empirical FCV was greater than the ACR-based FCV and BLM at pH 6.5 but comparable to the other two models at pH 8.5 (Figure 3).

DISCUSSION

Use of MLRs for describing the effects of water chemistry on metal toxicity is not a novel concept, and several examples can be found in the literature using this or similar statistical approaches.[19,24−28] In this paper, we developed this concept further by, for the first time to the best of our knowledge, using species-specific MLRs to derive a MLR-based WQC. The approach we used was analogous to the existing USEPA procedure for pooling species-specific water hardness models to develop a hardness-dependent WQC for metals.[1]

Comparison of the MLR Model and BLM. A comparison of the pooled MLR model and BLM indicates that MLR model performance is comparable to the BLM across a wide range of water chemistries. The pooled MLR model performs better than the BLM for species with large data sets (*D. magna* and *P. promelas*) which dominate the pooled MLR data set but not as well as the BLM for species with smaller data sets (Figure 1). Evaluation of model residuals as a function of observed toxicity and water quality parameters generally show very similar species-specific patterns for the MLR model and BLM in many cases (Figures S5–S8), suggesting deviations between observed and predicted values may be caused by factors that are not accounted for in either model.

Comparison of MLR- and BLM-Based Acute Criteria. We also compared the MLR-based Cu criteria derived here to BLM-based Cu criteria. The BLM-based Cu criteria were released in 2007, and there are more recent Cu toxicity data included in the current evaluation. To directly compare the MLR- and BLM-based criteria based on a common toxicity data set, the new toxicity data were added to the species-sensitivity distribution (SSD) from the 2007 AWQC document, and updated BLM-based criteria were derived. The updated BLM-based criterion maximum concentrations (CMCs) based on our analysis are ~20% lower than those derived by USEPA[8] (Figure 2), largely as a result of including recent data on the freshwater mussel *Villosa iris* (Tables S4 and S5).

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

There are several important differences between the MLR model and BLM responses to water quality parameters. The MLR model is less responsive than the BLM to DOC across a range of hardness and pH conditions (Figures 2 and 3). In contrast, the MLR model is more responsive to hardness than the BLM. Additionally, at low and intermediate hardness, the MLR model is less responsive than the BLM to pH, but the two models respond comparably at high hardness. The net effect of these different response profiles is that under many typical water quality conditions, the MLR- and BLM-based criteria (both acute and ACR-based chronic) are quite comparable. Indeed, conditions where the two models differ most (high pH/low hardness and low pH/high hardness) are relatively rare in natural aquatic systems.

Comparison of MLR- and BLM-Based Chronic Criteria. Specific to chronic criteria, we also evaluated empirical MLR-based final chronic values (FCVs), which are derived from the chronic MLR model for *D. magna* and an empirical chronic SSD. Empirical-based FCVs are greater than the ACR-based FCVs at low pH and comparable to or slightly lower at high pH. The empirical MLR has a shallower hardness slope (0.235) than the pooled acute MLR model (0.589). This difference is in part driven by the steeper hardness slope for *P. promelas* compared to the slopes for daphnids (Table 1) but also by differences in response to hardness between acute and chronic Cu toxicity (Figure S9). More important though, the pH slope for the chronic empirical MLR model (0.172) is substantially lower than the pH slope in the pooled MLR model (0.966) which results in the CMC being lower than the FCV at pH < 7.9 if the empirical MLR-based model is used for setting chronic criteria (Figure S10). Acute and chronic toxicity data for *D. magna* in which only pH was varied do not support the empirical MLR model, as mean (±SEM) pH slopes for acute (0.86 ± 0.13, $n = 9$) and chronic (0.87 ± 0.10, $n = 5$) data sets were very similar (Figure S11). It is not clear why the chronic empirical MLR model does not accurately characterize the apparent pH slope of data sets where just pH was varied, but it may in part be related to the composition of the overall data set with other factors (e.g., DOC) dominating response variation. Regardless, given the empirical MLR-based FCV is based on data for a single species, there are issues with the model for this species with respect to pH, and use of this model would result in CMCs less that FCVs for waters with pH < 7.9, we recommend use of the ACR-based MLR model for setting water quality criteria.

This analysis highlights the importance of critically evaluating MLR models using appropriate subsets of the data to ensure that the final model of all data reasonably reflects responses to water quality variables that may be best characterized in subsets of the data where only a single water quality variable is changed.

Differences in Response to Water Quality Variables in MLR Models vs BLM. As discussed above, the pooled MLR model and BLM respond differently to the three key water quality parameters − DOC, hardness, and pH. While the net effect of these differences is relatively modest in terms of derived water quality criteria across a range of water chemistries, it is worth exploring the differential responses of the two models to individual water quality parameters. Inspection of residual plots as a function of a single water quality parameter can be informative (Figure S6−S8), but these can be confounded by variation in the other water quality parameters.

To explore this further, we evaluated a subset of toxicity data for a single species (*D. magna*) in which each parameter was varied in toxicity tests while other water chemistry parameters were held constant. While several relatively small data sets exist to support this analysis, the data sets from ref 29 (DOC and pH) and ref 30 (hardness) were considered optimal for this analysis. In the Ryan[29] study, acute Cu toxicity tests with *D. magna* were conducted at DOC concentrations ranging from <1 mg L^{-1} to ~15 mg L^{-1} across pH levels of 6, 7, and 8, with other water chemistry parameters held constant at each of these pH levels. In the Villavicencio et al.[30] study acute toxicity tests with *D. magna* were conducted at hardnesses ranging from 43 to 167 mg L^{-1} across DOC concentrations of ~0.1, 2.1, and 4.6 mg L^{-1}.

We first calibrated the pooled MLR model to the sensitivity of *D. magna* in ref 29 or ref 30 which entailed back-calculation of the intercept based on the DOC, pH, and hardness slopes in the pooled model and the reported EC50 for each test. The species mean intercept was then calculated as the arithmetic mean of back-calculated intercepts for all tests at a given pH. The arithmetic mean (rather than the geometric mean) was used in this case because the intercepts are already log-transformed values. This species mean intercept was then used in the pooled MLR model. Likewise, the BLM was calibrated to the sensitivity of *D. magna* in each study by first running the BLM in speciation mode for all tests, which provided the LA50 for each test. The geometric mean LA50 was then used to define the sensitivity of *D. magna* in ref 29 and in ref 30. The resulting MLR model and BLM, calibrated to the sensitivity of *D. magna* in each study, were used to predict acute Cu EC50s that could be compared to the measured EC50s in the two studies.

In assessing the response of models to changing DOC concentrations it is important to keep in mind that DOC from different sources have variable compositions which can significantly affect complexation with Cu and other metals.[31,32] Neither the MLR model or BLM currently takes this variability into account (the BLM can account for percentage of humic acid content, but it is rarely measured). In the current assessment, we have constrained the analysis to a single DOC source (DOC isolated from Ogeechee River, Georgia, USA), while in the larger data set there are multiple DOC sources that could be differentially influencing predictions for the MLR model and BLM.

In the analysis of the Ryan[29] data, variable patterns in Cu LC50s as a function of DOC concentration were observed at the different pH levels. At pH 6, the MLR tended to more accurately predict toxicity over the range of DOC concentrations than the BLM, with the BLM generally consistently overpredicting toxicity (Figure 4). At pH 7, the two models are largely comparable in their ability to predict toxicity, with both models underpredicting toxicity at low DOC (Figure 4). At pH 8, the two models are comparable at low DOC but with increasing DOC concentrations separate and bound the observed toxicity data with the BLM underpredicting toxicity and the MLR overpredicting toxicity. These data suggest that part of the difference between MLR and BLM predictions may be the result of a DOC × pH interaction, though it is unclear which model is better at predicting toxicity. Inspection of model residuals as a function of DOC concentration does not provide clarification as both models appear to perform similarly when all data are evaluated (Figure S6).

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

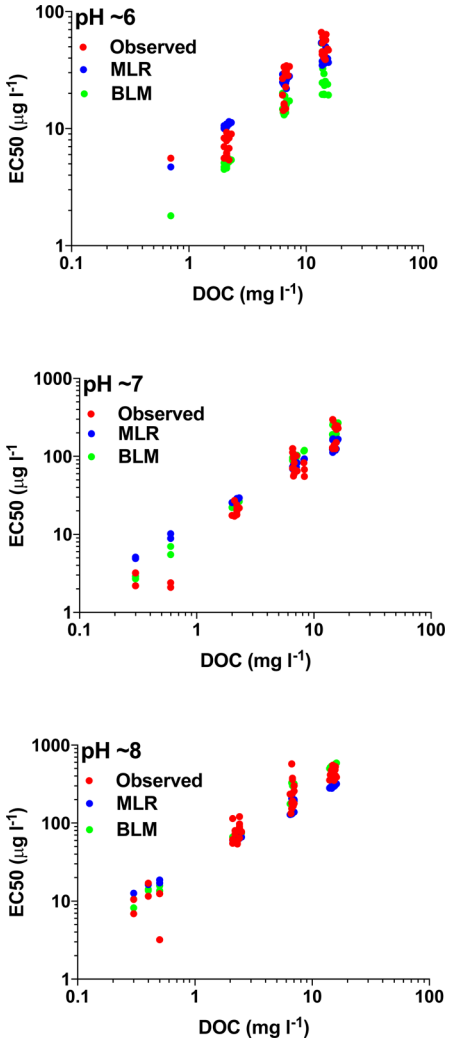


Figure 4. Effect of DOC on acute Cu toxicity to *D. magna* with comparison to MLR and BLM predictions at pH 6, 7, and 8. Data from ref 29.

When Cu LC50s were plotted versus pH with DOC and hardness held constant, MLR and BLM predictions were generally consistent or differed only slightly (Figure S12). Additionally, predictions for both models were generally consistent with observed toxicity. Consequently, this data set did not provide any insight into which model more accurately accounts for the effect of pH on Cu toxicity. Inspection of model residuals again show similar patterns between the two models (Figure S8).

When Cu LC50s were plotted versus hardness with DOC and pH held constant, both models significantly underpredicted Cu LC50s in low DOC waters. In higher DOC waters, the MLR accurately predicted toxicity, while the BLM consistently underpredicted toxicity, though to a lesser extent than in low DOC waters (Figure S13). Model residuals support these observations for some species but not others (Figure S7). For the large data sets of *D. magna* and *P. promelas*, the BLM underestimates the effect of hardness, while the MLR generally does not. In constrast, the MLR overestimates the effect of hardness on Cu toxicity for *D. obtusa* while the BLM does not, while both models overestimate the effect of hardness on Cu toxicity to *C. dubia*.

The BLM provides a robust scientific tool for improving our ability to regulate metals in the environment, and the mechanistic foundation of the model provides confidence that it can predict toxicity over a wide range of water qualities. Unfortunately, there has been reluctance by some regulatory authorities to adopt the BLM for routine use in setting water quality standards. In this paper, we developed a statistical model, analogous to the hardness-based models commonly used in water quality standards prior to development of the BLM, for setting WQC for Cu. We selected water quality parameters identified by the BLM as variables in our statistical model to retain a mechanistic foundation. Overall, the MLR model performance is comparable to the BLM in terms of predicting toxicity across a wide range of water chemistries and species. While the Cu MLR model responds less to DOC and more to hardness than the BLM, available data suggest the BLM may underpredict the hardness response for some species, and it is unclear from available data which model better characterizes the true DOC response.

Despite the differences in the ways the two models predict effects from individual water quality variables, MLR- and BLM-based WQC are generally within a factor of 2 except in low hardness (approximately <40 mg L^{-1}), high DOC (approximately >10 mg^{-1}), and high pH (>8) waters where the MLR is more conservative by up to a factor of 3. Given this, we conclude the MLR is a viable alternative approach for setting water quality standards for Cu in scenarios where there is reluctance to use the BLM.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.6b05533.

Figures S1−S13, a comparison of statistical approaches for deriving the pooled MLR model, and references (PDF)

Tables S1−S11 and references (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 305 773 8347. E-mail: kevinbrix@icloud.com.

ORCID

Kevin V. Brix: 0000-0003-3127-7435

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This analysis was supported by funding from Rio Tinto. Funding for M. Grosell was provided through a Maytag Professor of Ichthyology endowment.

■ REFERENCES

(1) Stephan, C. E.; Mount, D. I.; Hansen, D. J.; Gentile, J. H.; Chapman, G. A.; Brungs, W. A. *Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses*; NTIS No. PB85-227049; U.S. Environmental Protection Agency, Environmental Research Laboratory: Duluth, MN, 1985; pp 1−98.

(2) De Schamphelaere, K. A. C.; Lofts, S.; Janssen, C. R. Bioavailability models for predicting acute and chronic toxicity of zinc to algae, daphnids, and fish in natural surface waters. *Environ. Toxicol. Chem.* **2005**, *24* (5), 1190−1197.

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192

(3) Di Toro, D. M.; Allen, H. E.; Bergman, H. L.; Meyer, J. S.; Paquin, P. R.; Santore, R. C. A biotic ligand model of the acute toxicity of metals. I. Technical basis. *Environ. Toxicol. Chem.* **2001**, *20* (10), 2383−2396.
(4) Niyogi, S.; Wood, C. M. Biotic ligand model, a flexible tool for developing site-specific water quality guidelines for metals. *Environ. Sci. Technol.* **2004**, *38* (23), 6177−6192.
(5) Paquin, P. R.; Gorsuch, J. W.; Apte, S. C.; Batley, G. E.; Bowles, K. C.; Campbell, P. G. C.; Delos, C. G.; Di Toro, D. M.; Dwyer, R. L.; Galvez, F.; Gensemer, R. W.; Goss, G. G.; Hogstrand, C.; Janssen, C. R.; McGeer, J. C.; Naddy, R. B.; Playle, R. C.; Santore, R. C.; Schneider, U.; Stubblefield, W. A.; Wood, C. M.; Wu, K. B. The biotic ligand model: a historical overview. *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2002**, *133* (1−2), 3−35.
(6) Schlekat, C. E.; Van Genderen, E. J.; De Schamphelaere, K. A. C.; Antunes, P. M. C.; Rogevich, E. C.; Stubblefield, W. A. Cross-species extrapolation of chronic nickel Biotic Ligand Models. *Sci. Total Environ.* **2010**, *408*, 6148−6157.
(7) Van Sprang, P. A.; Verdonck, F. A. M.; Van Assche, F.; Regoli, L.; De Schamphelaere, K. A. C. Environmental risk assessment of zinc in European freshwaters: a critical appraisal. *Sci. Total Environ.* **2009**, *407*, 5373−5391.
(8) USEPA *Aquatic Life Ambient Freshwater Quality Criteria - Copper*; EPA/822/R-07/001; U.S. Environmental Protection Agency, Office of Water: Washington, DC, February 2007, 2007; pp 1−204.
(9) Meyer, J. S.; Adams, W. J. Relationship between biotic ligand model-based water quality criteria and avoidance and olfactory responses to copper by fish. *Environ. Toxicol. Chem.* **2010**, *29* (9), 2096−2103.
(10) Brix, K. V.; DeForest, D. K.; Burger, M.; Adams, W. J. Assessing the relative sensitivity of aquatic organisms to divalent metals and their representation in toxicity data sets compared to natural aquatic communities. *Hum. Ecol. Risk Assess.* **2005**, *11* (6), 1139−1156.
(11) De Schamphelaere, K. A. C.; Vasconcelos, F. M.; Tack, F. M. G.; Allen, H. E.; Janssen, C. R. Effect of dissolved organic matter source on acute copper toxicity to *Daphnia magna*. *Environ. Toxicol. Chem.* **2004**, *23* (5), 1248−1255.
(12) Long, K. E.; Van Genderen, E. J.; Klaine, S. J. The effects of low hardness and pH on copper toxicity to *Daphnia magna*. *Environ. Toxicol. Chem.* **2004**, *23* (1), 72−75.
(13) Markich, S. J.; Batley, G. E.; Stauber, J. L.; Rogers, N. J.; Apte, S. C.; Hyne, R. V.; Bowles, K. C.; Wilde, K. L.; Creighton, N. M. Hardness corrections for copper are inappropriate for protecting sensitive freshwater biota. *Chemosphere* **2005**, *60*, 1−8.
(14) Ryan, A. C.; Tomasso, J. R.; Klaine, S. J. Influence of pH, hardness, dissolved organic carbon concentration, and dissolved organic matter source on the acute toxicity of copper to *Daphnia magna* in soft waters: implications for the biotic ligand model. *Environ. Toxicol. Chem.* **2009**, *28* (8), 1663−1670.
(15) Ryan, A. C.; Van Genderen, E. J.; Tomasso, J. R.; Klaine, S. J. Influence of natural organic matter source on copper toxicity to larval fathead minnows (*Pimephales promelas*): implications for the biotic ligand model. *Environ. Toxicol. Chem.* **2004**, *23* (6), 1567−1574.
(16) Sciera, K. L.; Isely, J. J.; Tomasso, J. R.; Klaine, S. J. Influence of multiple water-quality characteristics on copper toxicity to fathead minnows (*Pimephales promelas*). *Environ. Toxicol. Chem.* **2004**, *23* (12), 2900−2905.
(17) Wang, N.; Mebane, C. A.; Kunz, J. L.; Ingersoll, C. G.; May, T. W.; Arnold, W. R.; Santore, R. C.; Augspurger, T.; Dwyer, F. J.; Barnhart, M. C. Evaluation of acute copper toxicity to juvenile freshwater mussels (Fatmucket, *Lampsilis siliquoidea*) in natural and reconstituted waters. *Environ. Toxicol. Chem.* **2009**, *28* (11), 2367−2377.
(18) De Schamphelaere, K.; Janssen, C. R. A biotic ligand model predicting acute copper toxicity for *Daphnia magna*: the effects of calcium, magnesium, sodium, potassium, and pH. *Environ. Sci. Technol.* **2002**, *36*, 48−54.
(19) Rogevich, E. C.; Hoang, T. C.; Rand, G. M. The effects of water quality and age on the acute toxicity of copper to the Florida apple snail. *Arch. Environ. Contam. Toxicol.* **2008**, *54*, 690−696.
(20) De Schamphelaere, K. A. C.; Heijerick, D. G.; Janssen, C. R. Refinement and field validation of a biotic ligand model predicting acute copper toxicity to *Daphnia magna*. *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2002**, *133* (1−2), 243−258.
(21) Burnham, K. P.; Anderson, D. R. Multimodel inference: understanding AIC and BIC in model selection. *Soc. Meth. Res.* **2004**, *33* (2), 261−304.
(22) Zuur, A. F.; Ieno, E. N.; Elphick, C. S. A protocol for data exploration to avoid common statistical problems. *Meth. Ecol. Evol.* **2010**, *1* (1), 3−14.
(23) USEPA *Ambient water quality criteria for cadmium - 1984*; NTIS No. PB85-227031; U.S. Environmental Protection Agency: Washington, DC, 1985; 126 pp.
(24) Erickson, R. J.; Benoit, D. A.; Mattson, V. R. *A prototype toxicity factors model for site-specific water quality criteria*; U.S Environmental Protection Agency: Duluth, MN, 1987; pp 1−40.
(25) Esbaugh, A. J.; Brix, K. V.; Mager, E. M.; De Schamphelaere, K. A. C.; Grosell, M. Multi-linear regression analysis, preliminary biotic ligand modeling, and cross species comparison of the effects of water chemistry on chronic lead toxicity in invertebrates. *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2012**, *155*, 423−431.
(26) Esbaugh, A. J.; Brix, K. V.; Mager, E. M.; Grosell, M. Multi-linear regression models predict the effects of water chemistry on acute lead toxicity to *Ceridaphnia dubia* and *Pimephales promelas*. *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.* **2011**, *154*, 137−145.
(27) Fulton, B. A.; Meyer, J. S. Development of a regression model to predict copper toxicity to *Daphnia magna* and site-specific copper criteria across multiple surface-water drainages in an arid landscape. *Environ. Toxicol. Chem.* **2014**, *33* (8), 1865−1873.
(28) Welsh, P. G.; Parrott, J. L.; Dixon, D. G.; Hodson, P. V.; Spry, D. J.; Mierle, G. Estimating acute copper toxicity to larval fathead minnow (*Pimephales promelas*) in soft water from measurements of dissolved organic carbon, calcium, and pH. *Can. J. Fish. Aquat. Sci.* **1996**, *53*, 1263−1271.
(29) Ryan, A. C. *Influence of dissolved organic matter source and chemical characteristics on acute copper toxicity*; Clemson University: 2005.
(30) Villavicencio, G.; Urrestarazu, P.; Carvajal, C.; De Schamphelaere, K. A. C.; Janssen, C. R.; Torres, J. C.; Rodriguez, P. H. Biotic ligand model prediction of copper toxicity to daphnids in a range of natural waters in Chile. *Environ. Toxicol. Chem.* **2005**, *24* (5), 1287−1299.
(31) Luider, C. D.; Crusius, J.; Playle, R. C.; Curtis, P. J. Influence of natural organic matter source on copper speciation as demonstrated by Cu binding to fish gills, by ion selective electrode, and by DGT gel sampler. *Environ. Sci. Technol.* **2004**, *38*, 2865−2872.
(32) Richards, J. G.; Curtis, P. J.; Burnison, B. K.; Playle, R. C. Effects of natural organic matter source on reducing metal toxicity to rainbow trout (*Oncorhynchus mykiss*) and on metal binding to their gills. *Environ. Toxicol. Chem.* **2001**, *20* (6), 1159−1166.

DOI: 10.1021/acs.est.6b05533
Environ. Sci. Technol. 2017, 51, 5182−5192