Freshwater Science, 2012, 31(3):709–723
© 2012 by The Society for Freshwater Science
DOI: 10.1899/11-133.1
Published online: 5 June 2012

Estimating risks to aquatic life using quantile regression

Travis S. Schmidt[1,2,5], William H. Clements[3,6], AND Brian S. Cade[4,7]

[1] *Water Resources Division, US Geological Survey, Fort Collins Science Center, 2150 Centre Avenue Building C, Fort Collins, Colorado, 80226 USA*

[2] *Mineral Resources Team, US Geological Survey, Denver Federal Center, Denver, Colorado, 80225 USA*

[3] *Fish, Wildlife, and Conservation Biology Department, Colorado State University, Fort Collins, Colorado, 80523 USA*

[4] *Biological Resources Division, US Geological Survey, Fort Collins Science Center, Fort Collins, Colorado, 80226 USA*

Abstract. One of the primary goals of biological assessment is to assess whether contaminants or other stressors limit the ecological potential of running waters. It is important to interpret responses to contaminants relative to other environmental factors, but necessity or convenience limit quantification of all factors that influence ecological potential. In these situations, the concept of limiting factors is useful for data interpretation. We used quantile regression to measure risks to aquatic life exposed to metals by including all regression quantiles (τ = 0.05–0.95, by increments of 0.05), not just the upper limit of density (e.g., 90th quantile). We measured population densities (individuals/0.1 m^2) of 2 mayflies (*Rhithrogena* spp., *Drunella* spp.) and a caddisfly (*Arctopsyche grandis*), aqueous metal mixtures (Cd, Cu, Zn), and other limiting factors (basin area, site elevation, discharge, temperature) at 125 streams in Colorado. We used a model selection procedure to test which factor was most limiting to density. *Arctopsyche grandis* was limited by other factors, whereas metals limited most quantiles of density for the 2 mayflies. Metals reduced mayfly densities most at sites where other factors were not limiting. Where other factors were limiting, low mayfly densities were observed despite metal concentrations. Metals affected mayfly densities most at quantiles above the mean and not just at the upper limit of density. Risk models developed from quantile regression showed that mayfly densities observed at background metal concentrations are improbable when metal mixtures are at US Environmental Protection Agency criterion continuous concentrations. We conclude that metals limit potential density, not realized average density. The most obvious effects on mayfly populations were at upper quantiles and not mean density. Therefore, we suggest that policy developed from mean-based measures of effects may not be as useful as policy based on the concept of limiting factors.

Key words: quantile regression, population, metals, risk, biological assessment.

Elevated concentrations of metals in streams draining mineralized and mined basins are common in Colorado (USA) and globally (Runnells et al. 1992, Clements et al. 2000). Metals transported into aquatic habitats cause declines in some metal-sensitive populations. Resultant communities are less diverse and are dominated by metal-tolerant populations (Clements 1994, Clements et al. 2000). These predictable responses are useful for evaluating the ecological effects of metal pollution in aquatic ecosystems.

[5] E-mail addresses: tschmidt@usgs.gov

[6] willc@warnercnr.colostate.edu

[7] cadeb@usgs.gov

Mayflies, particularly taxa in the families Heptageniidae and Ephemerellidae, are among the aquatic insects most sensitive to metal pollution (Clements et al. 2000, Cain et al. 2004). Field surveys, microcosm experiments, and field manipulations have consistently demonstrated that *Rhithrogena* spp. and *Drunella* spp. densities decline when exposed to metals (Clements et al. 2000, Clements 2004, Clark and Clements 2006). In contrast, some caddisflies, including *Arctopsyche grandis,* are thought to be tolerant of metal pollution (Clements 1994, Cain et al. 2004). Because of its tolerance, *A. grandis* can be collected along a gradient of metal concentrations and used to evaluate exposure to metal by comparing tissue concentrations across sites (Kiffney and Clements

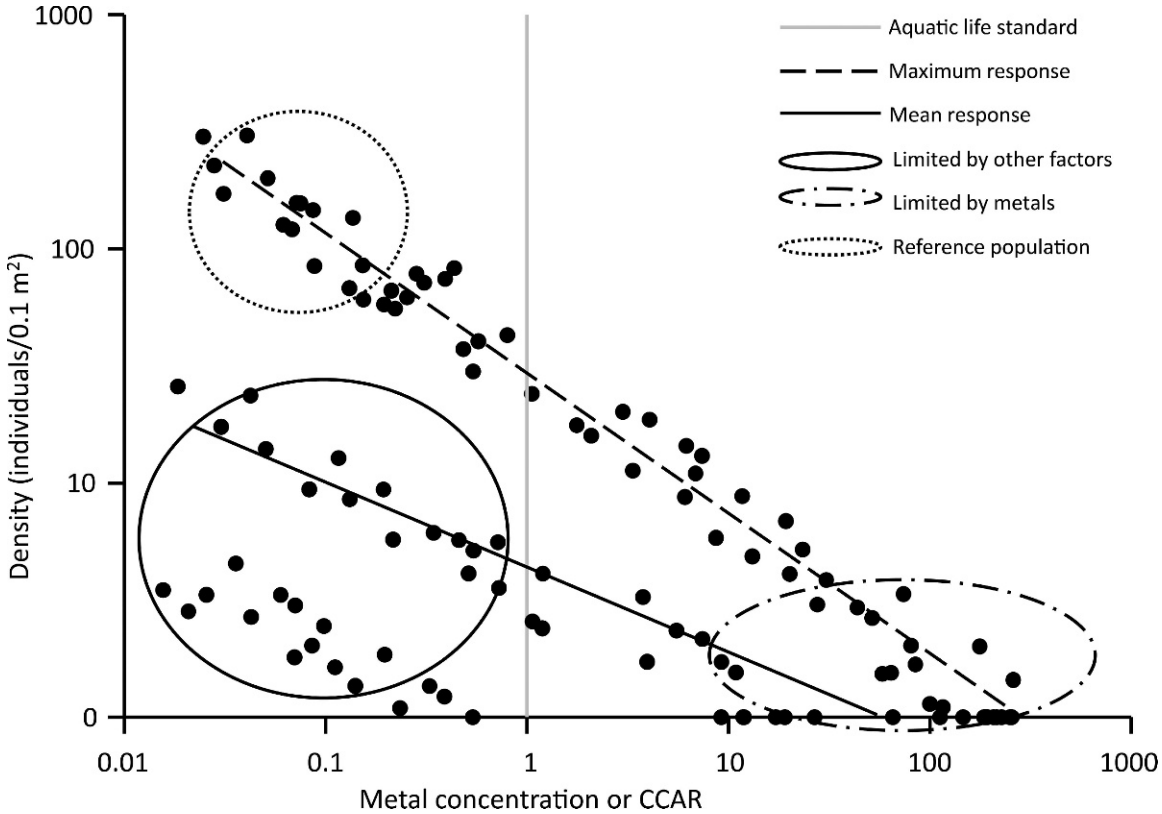

FIG. 1. Conceptual diagram showing a wedge-shaped distribution commonly observed in ecological studies (modified from Cade and Noon 2003). Using metals as an example, we hypothesized that at low levels of metals aquatic-insect density (in this example, *Drunella* spp.) is not limited by metals but is limited by natural processes (i.e., habitat and biotic interactions). However, as metals increase and become the primary limiting factor, high-density populations are not observed, even when other potentially limiting conditions might be optimal. Maximum quantiles of density (τ = 0.90, 0.95, etc.) are less affected by confounding variables and, therefore, are less biased than models based on means. Evaluation of effects of metals on high quantiles allows estimates of effects on resources in highly productive habitats. CCAR = chronic criterion accumulation ratio.

1993, Cain et al. 2004). The divergent properties of these mayfly and caddisfly taxa and their ubiquity in the intermountain west make them attractive biological indicators of the effects of abandoned mine lands on aquatic ecosystems (Ward et al. 2002).

Detecting the effects of metals on biological indicators can be complicated by the many other limiting factors that control population size in streams (Clements 1994, Carlisle and Clements 1999). Natural factors, such as physiographic variables (elevation, stream order) (Kiffney and Clements 1994, Clements and Kiffney 1995); biotic interactions, such as the presence of fish (Forrester 1994); or even life-history characteristics (Kiffney and Clements 1996) can limit local population densities. Stream ecologists generally recognize the importance of these and other factors, but quantifying all limiting factors is not possible. When the effects of contaminants, such as metals, on the biological indicator are estimated as a bivariate relationship, the influences of these other limiting factors cause variability in responses that is independent of the effects of metals. This variability and variability associated with sampling error or stochastic processes can obscure or bias statistical relationships between the contaminant and the biological response. Thus, detecting the adverse effect of metals, particularly at low concentrations, on macroinvertebrate populations can be difficult.

When aquatic insect density is plotted against a gradient of metal concentration, insect density often is distributed in a wedge-like pattern across the gradient (Fig. 1). Maximum density occurs at low metal concentrations and decreases with increasing metal concentrations, as do the mean and variance of density. At the same time, lower densities and absences are observed across all metal concentrations. Ecologists studying ecological responses to a well characterized limiting factor in the presence of other unknown or poorly characterized limiting factors have made similar observations (Terrell et al. 1996,

Cade and Noon 2003). The maximum observable density can be constrained by a strong and well characterized limiting factor, while at some locations, the presence of other limiting factors may not permit this maximum value. This additional constraint causes observed densities at some locations to fall below those that could occur if other conditions were not limiting. The effect of these colimiting factors is to reduce the rate of change in density of aquatic insects for unique portions (e.g., $\tau = 0.90$ vs $\tau = 0.50$) of the response function (cf. slopes of lines at mean vs maximum density values; Fig. 1).

For contaminants like metals, low concentrations are not expected to limit density, and other natural phenomena (e.g., physiographic factors, biotic interactions, phenology, and life history) introduce variability in the density of an organism observed at different sites with similar metal concentrations. However, as metal concentrations increase, maximum density is limited to levels below that expected at low metal concentrations, and natural phenomena may further constrain the maximum potential density at some sample locations. This variability caused by unmeasured factors reduces statistical power and sensitivity of biological monitoring endpoints when comparing means across different levels of contamination (Carlisle and Clements 1999, Clements et al. 2000). These inherent weaknesses of observational data combined with the lack of appropriate statistical methods have limited the utility of field monitoring data when setting aquatic-life criteria (Pacheco et al. 2005, Linton et al. 2007).

Quantile regression is a tool that can be used to estimate the rate of change for any quantile of a response variable to a limiting factor (Fig. 1) (Koenker and Bassett 1978, Pacheco et al. 2005, Linton et al. 2007). Regression quantiles are ascending sequences of planes that are above an increasing proportion of sample observations as values of the quantiles (τ) increase (Koenker and Bassett 1978, Cade and Noon 2003). This property of regression quantiles facilitates estimation of a rate of change (slope) for any quantile of the data, not just along the central tendency as in ordinary least squares (OLS) regression. Each fraction of the response variable or quantile could be limited by any of a suite of measured and unmeasured factors. For example, maximum density (e.g., $\tau = 0.90$ or 0.95) could be limited by metals, while lower quantiles (e.g., $\tau = 0.10$) could be limited by other factors. Unequal variance in the distribution of data plotted across a gradient is consistent with interactions of unmeasured limiting factors with the measured factor (Cade and Noon 2003, Cade et al. 2005). OLS is not effective in this situation because it describes mean rates of change (mean slope) by averaging all quantile slopes. As a result, OLS slopes will be similar to some quantiles (e.g., $\tau = 0.50$), but probably will fail to describe real non-0 rates of change in other quantiles (Terrell et al. 1996, Cade and Noon 2003).

Most investigators have used quantile regression to determine the functional relationship between a stressor and the response variable at a limited number of high quantiles (e.g., $\tau = 0.90$ or 0.95) (Pacheco et al. 2005, Linton et al. 2007). Ecologists focus on change in these quantiles because it evokes the ecological concept of limiting factors as constraints on high resource values most reliably affected by the limiting factor. However, restricting inference to a single quantile severely limits the information available to scientists and decision makers and could result in less useful interpretation of results. By evaluating multiple quantiles, one can assess which factor might limit low values vs which factor might limit high values of the biological response across similar levels of the stressor (Fig. 1). This capability is important because, at some level of the stressor where both high densities and low densities are observable, other factors unrelated to the stressor drive lower densities and obscure our ability to detect change at low levels of the stressor. As the basis for making inference, considering several quantiles can enable an investigator to identify which regression quantile has the greatest rate of change and to identify portions of the probability distribution most limited by the stressor. Furthermore, investigators can evaluate the sampling variation (90% confidence intervals [CIs]) in statements about the rate of change associated with each quantile and how those rates of change might influence decision-making. Last, investigators can use these regression quantile relationships to make risk statements about the likelihood of observing a given value of the biological indicator at selected levels of the stressor by estimating the empirical conditional cumulative distribution function as the inverse of the quantile estimates. This capability provides a new way to communicate results to decision makers.

We evaluated 3 aquatic insect populations as biological indicators of metal contamination in the Central Colorado Rocky Mountains. This region is affected by natural and anthropogenic releases of metals in the presence of strong natural gradients (e.g., elevation, discharge, basin area, water temperature) and provided an opportunity to assess the effect of metals on aquatic insect densities in the presence of other potentially limiting factors. We explored the response of each taxon using quantile

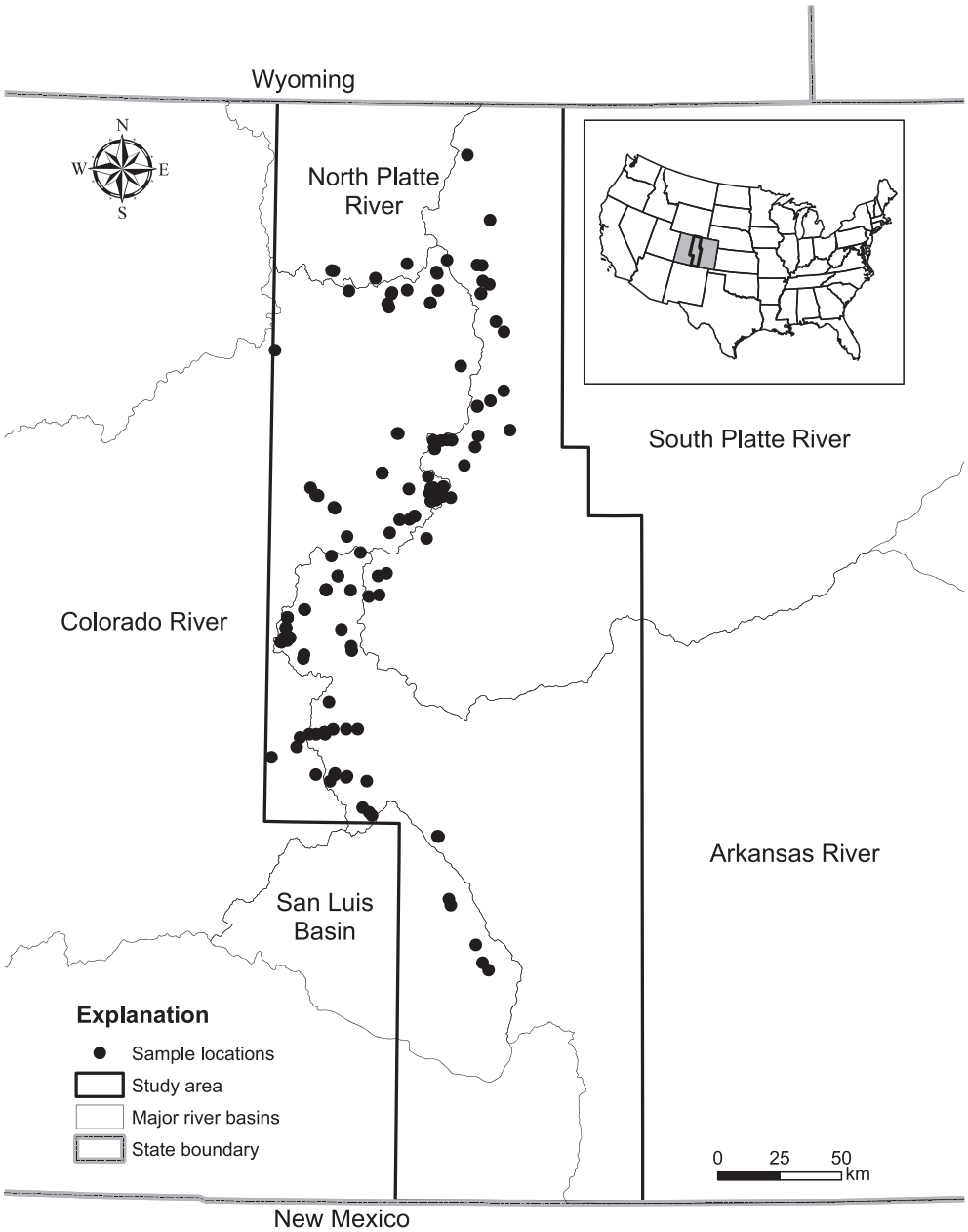

FIG. 2. Map of the study area in Colorado, USA. Streams were sampled throughout the mountainous region of Colorado from the Wyoming border to the north to the New Mexico border to the south.

regression to quantify how they responded to colimiting factors. Specifically, we assessed which quantiles of density were limited by metals or other factors, determined the interval of quantiles across which metals were the primary limiting factor, observed which quantile had the maximum rate of change (steepest slope), developed an understanding of how sampling variation in the slope estimates changed across all quantiles, and assessed how risks to aquatic insect populations changed based on the quantile evaluated and level of exposure to metals. Last, we discuss the ecological and management implications of our findings. For simplicity, we limited our analysis to bivariate relationships between the biological indicator and a single limiting factor. We could have developed models that were more complicated, but we think the scenario we modeled is a common scenario for state and federal biologists, and the simpler model facilitates easy visualization of the ideas we present.

TABLE 1. Summary statistics of potential limiting factors observed at stream sites where density was measured. CCAR = chronic criterion accumulation ratio where values < 1.0 are presumed protective of aquatic life.

Variable	Median ± SD	Measured range
CCAR	0.18 ± 60.03	0.02–268.35
Temperature (°C)	9.5 ± 2.7	1–17.8
Basin area (km^2)	17 ± 45	2–292
Discharge (m^3/s)	0.20 ± 0.31	0.01–2.07
Site elevation (m)	2993 ± 287	2329–3547

Methods

Study design

The study area was central Colorado from Wyoming to New Mexico, an area of ~55,000 km^2 that includes most of the Rocky Mountains in Colorado and represents ~20% of the land area in the state (Fig. 2). This area includes a geographic feature called the Colorado Mineral Belt that has been exploited for the past 150 y for its mineral resources. Sampling sites in this study are at high elevation, ranging from ~2330 to 3550 m asl (Table 1). The climate of the study area is temperate continental, with generally >50 cm of precipitation per year, especially at higher altitudes. Much of this precipitation occurs as winter snow or as summer rain. Vegetation ranges from deciduous cover at lower altitudes and in riparian zones to coniferous forests and open tundra at the highest altitudes (Mutel and Emerick 1992). Soils in the study area are thin (rarely >10 cm) to nonexistent in areas dominated by bedrock outcrops. Thicker (up to 1 m or more) immature soils and unconsolidated overburden occur intermixed at lower elevations and along streams (Soil Survey Staff 1999).

We sampled small basins (1st–3rd-order) predominantly underlain by a single rock type and categorized them based on mineral-deposit criteria. The purpose of this sampling strategy was to target a variety of water-quality conditions caused by interaction with the underlying rocks of the basin and to develop geochemical and biological baselines based on rock type (Schmidt et al. 2012). We collected geochemical and benthic macroinvertebrate community samples during summer baseflow conditions (July–September) from 2003 through 2007. All samples used for the purposes of our study (n = 149) were affected by limited or no other anthropogenic stressors except the influence of historical mining practices. Of these 149 samples, 125 were discrete locations (Fig. 2). We sampled 12 of these locations annually (24 additional samples) to capture interannual variability.

Physicochemical variables

We used field methods that meet the requirements of the biotic ligand model (Wilde and Radtke 1998, Wilde et al. 1998, HydroQual 2007) to collect water samples in 2003–2007. The biotic ligand model predicts the free-ion metal concentration available to accumulate on the respiratory surface of aquatic organisms and accounts for the influence of water quality on metal toxicity. The amount of metal bound to the respiratory surface, modeled as a biotic ligand using thermodynamic equilibrium, is correlated with mortality in fish and aquatic invertebrates (Di Toro et al. 2001).

At each sampling site, we made instantaneous measurements of pH, water temperature, specific conductance, and stream discharge (Rantz et al. 1982, Wilde and Radtke 1998, Wilde et al. 1998). We filtered 1 aliquot of water through a 0.45-μm filter and acidified it with ultrapure HNO_3 to a pH of ~1 for cation analysis. We also collected a filtered, unacidified aliquot for anion analyses. We collected separate aliquots for analysis of alkalinity (unfiltered, unacidified) and dissolved organic C (DOC; 1 μm glass-fiber filtered, HCl acidified to pH ~1, stored in amber glass bottle). We refrigerated unacidified samples and stored acidified samples at room temperature.

Sample analysis was conducted at the US Geological Survey's (USGS) Geologic Discipline Laboratory in Denver, Colorado. Concentrations of major cations (Na^+, K^+, Mg^{2+}, and Ca^{2+}) were analyzed by inductively coupled plasma-atomic emission spectrometry (Perkin–Elmer 5300 Optima Dual-View ICP-AES; Perkin–Elmer, Waltham, Massachusetts), metals (Cd, Cu, Zn) were analyzed by inductively coupled plasma mass spectrometry (Perkin–Elmer Sciex Elan 6000 ICP-MS), and major anions (Cl^-, F^-, NO_3^-, SO_4^{2-}) were measured by ion chromatography (Taggart 2002). Analytical methods used in 2003 differed from above in that major cations and metals were analyzed by furnace atomic absorption spectrophotometry (Perkin–Elmer model 372) at the Department of Fish, Wildlife, and Conservation Biology, Colorado State University. The minimum reported detection limits between sampling periods were: Cd (0.01 μg/L in 2003 vs 0.02 μg/L in 2004–2007), Cu (0.01 μg/L in 2003 vs 0.5 μg/L in 2004–2007), and Zn (2 μg/L in 2003 vs 0.5 μg/L in 2004–2007). We used a Shimadzu 5000A (Shimadzu, Kyoto, Japan) to quantify DOC concentrations with a detection limit of 0.3 mg/L. When running the biotic ligand model, ½ the detection limit was substituted for constituents below the analytical limits of detection. Schmidt et al. (2010)

presented a more complete description of sample methods and quality-assurance/quality-control protocols and results.

Chronic criterion accumulation ratio

Most metal-polluted streams in Colorado are affected by mixtures of metals (Cd, Cu, Zn) at chronic concentrations, so a measure of cumulative toxicity was necessary (Clements et al. 2000, Schmidt et al. 2010). The chronic criterion accumulation ratio (CCAR) is a metric that relies on the biotic ligand model to predict the cumulative toxicity of metal mixtures while accounting for site-specific influences of water quality on metal toxicity to aquatic organisms (HydroQual 2007, Schmidt et al. 2010). CCAR is the ratio of the free metal ion concentration available to accumulate on a biological surface (a biotic ligand) to that accumulated at the US EPA chronic criterion value, summed for all metals (Cd + Cu + Zn). A CCAR value ≥ 1 indicates that a free metal ion concentration (or in this case, a combination of metals) under site-specific water chemistry is predicted to accumulate on the biotic ligand at concentrations above that which accumulates on the biotic ligand at US EPA continuous chronic criterion value(s) under a standardized water-chemistry mix, and thus, toxicity might occur. A value <1 indicates that all combinations of free metal ions accumulated on the biotic ligand under site-specific water chemistry are below that expected to accumulate at US EPA continuous chronic criterion values given standardized water chemistry, and thus, toxicity is not expected. See Schmidt et al. (2010) for a complete method description and comparison of CCAR with chronic criterion units.

Density estimates

At each sampling location, we collected 5 replicate benthic samples with a 0.1-m^2 Hess sampler (350-µm mesh net) from shallow riffle areas ($<$0.5 m). We processed these samples in the field and laboratory separately (Schmidt et al. 2010). We scrubbed overlying substrate, disturbed it to a depth of ~10 cm, and washed the remaining material through a 350-µm-mesh sieve. We preserved all retained organisms in 80% ethanol in the field and counted them in the laboratory. In the laboratory, we removed 300 organisms (±10%) from the sample with methods described by Moulton et al. (2000). We identified invertebrates to the lowest practical taxonomic level (genus or species) (Merritt and Cummins 1996, Ward et al. 2002). We adjusted subsamples proportionately to estimate the total number of organisms found in the whole sample. We estimated a single mean density (individuals [ind]/0.1 m^2) from the 5 replicate density values at each site (n = 125 discrete locations; Fig. 2).

Statistical analysis

Model fit and model selection.—We estimated quantile regression coefficients and 90% CIs (*quantreg* package, version 4.50; R Development Core Team, Vienna, Austria) for quantiles (τ = 0.05–0.95 by increments of 0.05) of the bivariate relationship between density and a suite of 5 predictors (CCAR, basin area [km^2], site elevation [m], discharge [m^3/s], and stream temperature [°C]) thought to influence density (Koenker and Bassett 1978, Koenker 2005). We developed CIs for the quantile regression coefficients from the inverted rank score test assuming errors were not independent and identically distributed (Cade et al. 2005, Koenker 2005). We developed slopes and 90% CIs for OLS regression models for the bivariate relationships between insect density and the suite of predictors described above to assess how the average of all quantile slopes compared to the 19 individual quantile estimates.

Akaike information criteria (AIC) were used to select which of the 5 competing models best fit each of the 19 quantile planes. We calculated AIC as follows:

$$\mathrm{AIC} = -2(\mathrm{loglikelihood}) + 2(k+1)$$

where k is the number of regressors in the model. We normalized all models by the AIC value of the CCAR model to derive delta-AIC (Δ_i) because we expected metals to be the strongest and, thus, most likely limiting factor. The lowest Δ_i value indicates the most likely candidate model or the most limiting factor for each quantile. $\Delta_i < |2|$ indicates candidate models that are indistinguishable in their likelihood of being the best candidate model, whereas $\Delta_i > |2|$ indicates models that are less likely candidates (Burnham and Anderson 2002). All statistics were developed using R software (version 2.10.0).

Drawing inference from quantile regression.—We developed an approach to measure limitation of density caused by metals based on all of the data rather than only an extreme upper quantile (e.g., τ = 0.90 or 0.95). We estimated the density observed for each quantile (0–1) at selected levels of metal exposure to generate empirical distribution functions describing how the probability of observing a given density changed with different exposure scenarios. Specifically, we used the quantile regression models

TABLE 2. Spearman rank correlations (r) among measured colimiting factors that potentially limit density of aquatic insects. CCAR = chronic criterion accumulation ratio.

Variable	Temperature (°C)	Discharge (m^3/s)	Site elevation (m)	Basin area (km^2)
CCAR	0.13	−0.01	0.26	0.05
Temperature (°C)		−0.11	0.02	−0.04
Discharge (m^3/s)			−0.02	0.64
Site elevation (m)				−0.41

(τ = 0.05–0.95 by increments of 0.5) to calculate the number of individuals expected at CCAR = 0.1 (background) and CCAR = 1.0, the value thought to be protective of aquatic ecosystems (Schmidt et al. 2010). Each quantile estimate of density at a given value of CCAR can be used to provide an empirical cumulative distribution function of densities at the value of CCAR. This distribution function can be used to determine the proportion of the population having densities $\leq$ a specified value. For example, if the upper regression quantile (Fig. 1) were τ = 0.95, then 95% of the populations observed would have $\leq$100 individuals (ind)/0.1 m^2 at CCAR = 0.1. These probabilities are derived from an empirical distribution function of the data, not from an assumed distributional form (e.g., normal or lognormal) (Koenker and Bassett 1978, Cade and Noon 2003). We developed 90% CIs for each point estimate using the standard *xy*-pair bootstrap method (Koenker 2005). As a result, a biologist would be able to express changes in the indicator value as a risk statement that might be informative to resource managers interested in risks to aquatic ecosystems posed by different concentrations of a contaminant.

Results

Metal toxicity and physiographic variables

Detailed information about stream chemistry and the chemical constituents used to calculate CCAR was published by Schmidt et al. (2010). CCAR ranged from 2 orders of magnitude below the presumptively protective threshold (CCAR = 1.0) to 3 orders of magnitude above this threshold (Table 1). The median value of CCAR in basins with little or no mineralization or mining activity was 0.1 (background value; Schmidt et al. 2012). Instantaneous stream temperatures ranged from near freezing to almost 18°C, typical of alpine and subalpine streams during midsummer. Basin area (median = 17 km^2) and stream discharge (median = 0.20 m^3/s) ranged over 3 orders of magnitude (2–292 km^2 and 0.01–2.07 m^3/s respectively), and differences in site elevation were >1000 m (median = 2993 m, range 2329–3547 m).

Regression analyses

With the exception of basin area and stream discharge, Spearman rank correlations among predictors showed relatively weak ($r < 0.25$) relationships (Table 2). Not all regression slopes were significantly different from 0, but this depended on taxon and the model considered (Fig. 3A–E). For *A. grandis*, regression quantile slopes were not estimated for $\tau < 0.45$ because the slopes were not different from 0. However, for $\tau > 0.45$, a number of quantile slopes for CCAR, basin area, and elevation were significantly different from 0 (Fig. 3A–C). The greatest differences in slopes across quantiles were observed for basin area and elevation where rates of change in the extreme high quantiles (e.g., τ = 0.90 or 0.95) were nearly 2× that observed at the mean (OLS estimate) or median (τ = 0.50). This result suggests that estimating effects based on mean *A. grandis* response would underestimate effects relative to effects estimated by many regression quantiles above the mean. For *Drunella* spp., most regression quantile slopes for CCAR, basin area, and elevation were different from 0 (Fig. 3A–C). The greatest differences in slopes across quantiles were observed for elevation. In general, individual regression quantile slopes for all factors except CCAR were within the 90% CI of the mean (OLS) slope estimate except for regression quantiles at extreme low and high quantiles. For CCAR, intermediate and high quantile (τ = 0.60–0.75 and 0.95) slopes for *Drunella* spp. density were somewhat different from the mean slope. This result suggests that mean effects in *Drunella* spp. density might be a reasonable approximation for effects across most regression quantiles but would underestimate effects at intermediate and high quantiles. For *Rhithrogena* spp., models with CCAR or temperature generally produced slopes different from 0, whereas the other models did not (Fig. 3A–E). The greatest differences in slopes for *Rhithrogena* spp. were observed for CCAR where slopes ranged from −0.12 to −0.44, exceeding the OLS estimate (−0.20) in the case of most of the higher quantiles ($\tau > 0.50$) by a factor of 2. This suggests that greater rates of change in *Rhithrogena* spp. density were observed at high quantiles than in the mean and

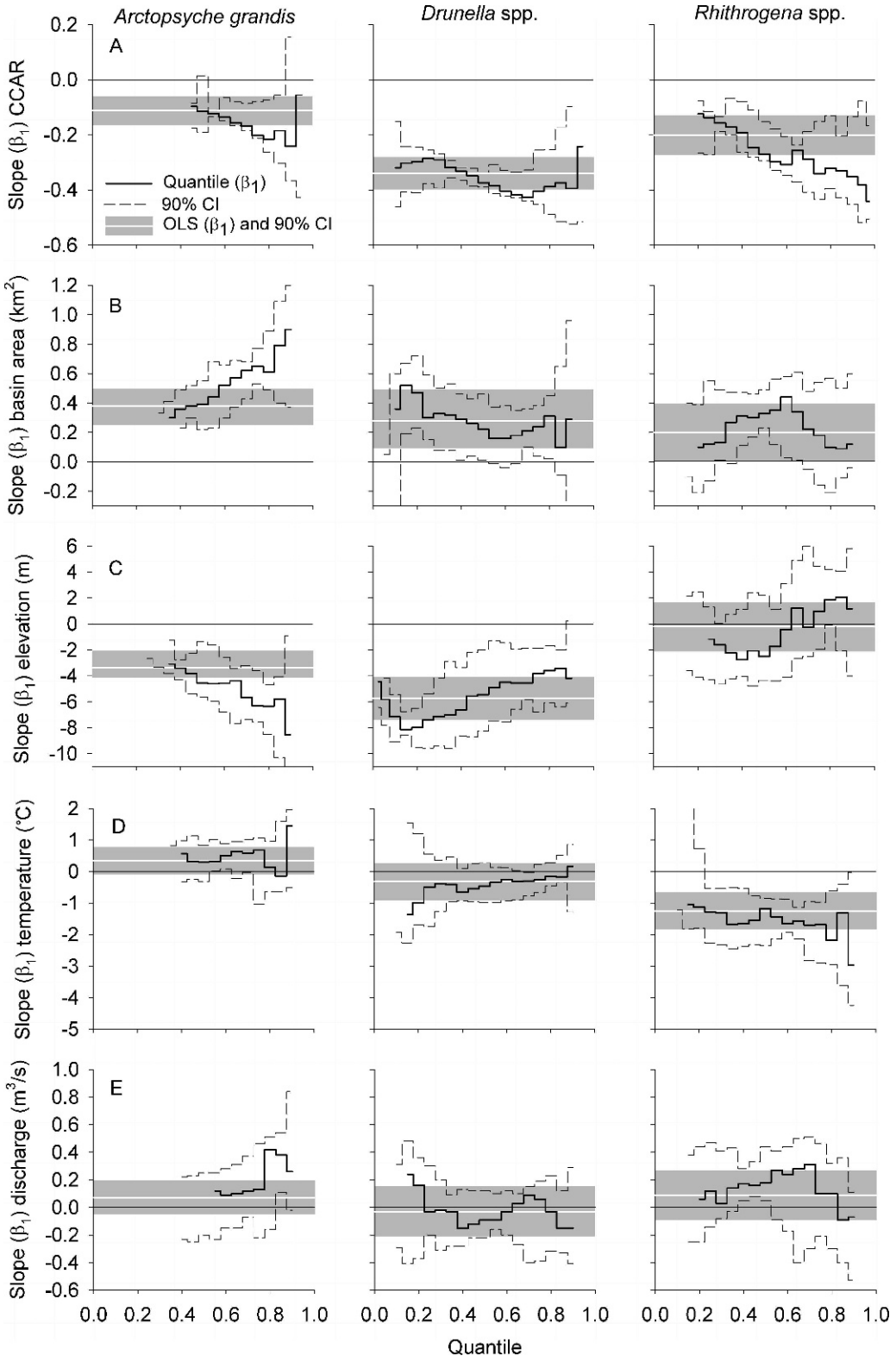

FIG. 3. Model slopes and confidence intervals (CI) of quantile regressions and ordinary least squares (OLS) regressions for density vs chronic criterion accumulation ratio (CCAR) (A), basin area (B), elevation (C), $\log_{10}$(temperature) (D), and discharge (E). To observe differences in slopes, compare the OLS slope estimates (white line in gray box [90% CI]) to a quantile slope estimate (black horizontal line) and 90% CI (dashed black lines).

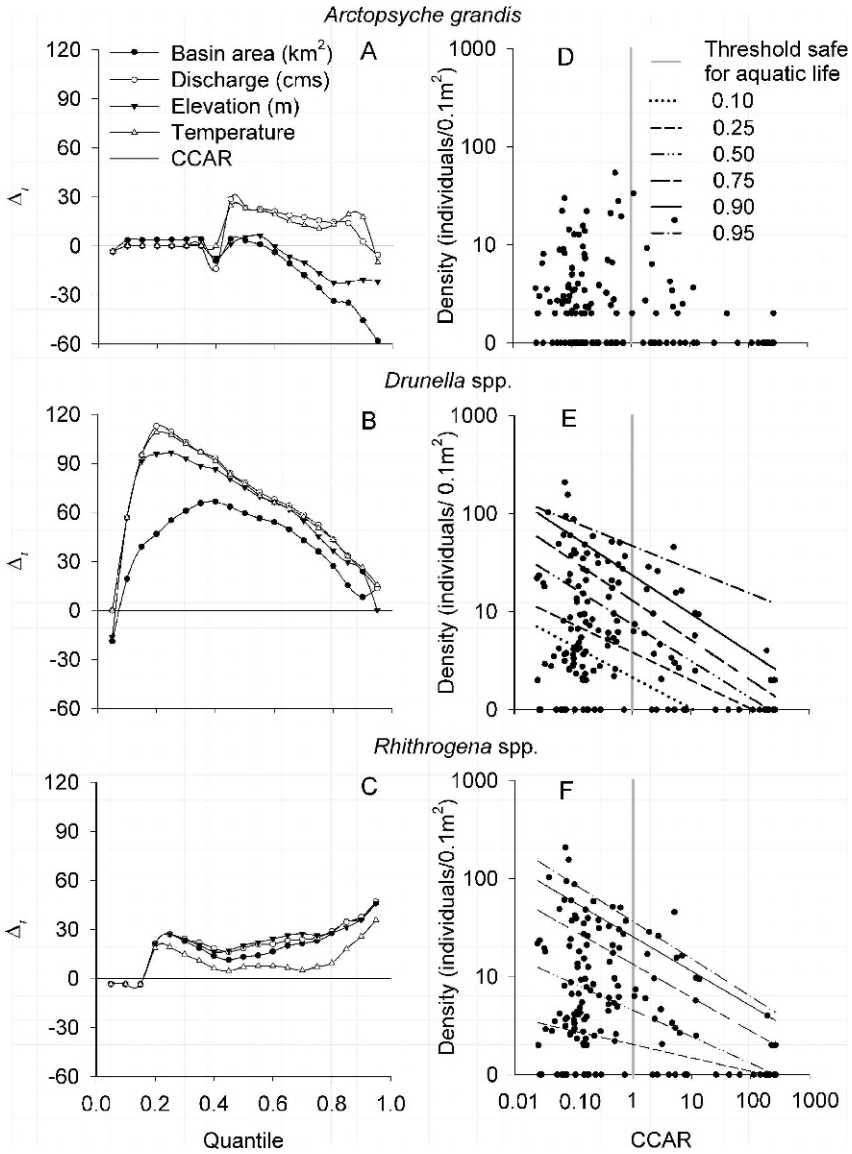

FIG. 4. Summary of Akaike information criteria model-selection results (A, B, C) and plots of quantile regressions of the chronic criterion accumulation ratio (CCAR) vs density (D, E, F) of *Arctopsyche grandis* (A, D), *Drunella* spp. (B, E), and *Rhithrogena* spp. (C, F). Δ_i values are calculated as $\Delta_i = \mathrm{AIC}_{\mathrm{model}i} - \mathrm{AIC}_{\mathrm{CCAR}}$. The lowest Δ_i value for each quantile is the best model. Quantile regression lines ($\log_{10}$-scale) plotted are those quantiles where CCAR was the top model (lowest Δ_i value) and slope $\neq 0$ for quantiles ($\tau = 0.10, 0.25, 0.50, 0.75, 0.90$). Threshold is represented by the vertical gray line at CCAR = 1, where metal mixtures are at criterion continuous concentrations and assumed to be safe for aquatic life (Schmidt et al. 2010). No regression quantiles are plotted for *A. grandis* because no models satisfied the conditions stated above.

lower quantiles. In comparison, the range of quantile slopes for the other factors were generally within the 90% CI for the OLS slope estimate. Intercepts were expected to increase with quantile, but we did not present them because they have little interpretive value for our purposes.

Results of the model-selection procedure (Δ_i) are presented in Fig. 4A–C. The top candidate model for each quantile is the model with the lowest Δ_i value. AIC for CCAR is always = 0 because of the way Δ_i was calculated. For example, CCAR was the most probable candidate model for most regression quantiles of *Drunella* spp. (exceptions $\tau = 0.05$ and 0.95; Fig. 4B). The same was also found for *Rhithrogena* spp. (exceptions $\tau \leq 0.15$; Fig. 4C), but not for *A. grandis* (Fig. 4A). Physiographic factors were the top model or among the top models for all regression quantiles of *A. grandis* density except $\tau = 0.45$ to 0.55 where CCAR was indistinguishable ($\Delta_i < |2|$) from site elevation and basin area as the top model (Burnham and

Anderson 2002). Basin area was the top model for quantiles $\tau = 0.60$ to 0.95 of *A. grandis*, whereas no model was the best model for quantiles below $\tau = 0.40$ where model slopes were not different from 0. Physiographic variables were the top-candidate models for *Drunella* spp. quantiles 0.05 and 0.95 and quantiles $\tau \leq 0.15$ for *Rhithrogena* spp., but none of these models had slopes different from 0. Fig. 4D–F depicts regression quantiles ($\tau = 0.10, 0.25, 0.50, 0.75, 0.90$) where CCAR was the top model and had a slope significantly different from 0. The highest and lowest quantile depicted in these graphs shows the interval over which metals were the most limiting factor. For *Drunella* spp., that interval was from $\tau = 0.10$ to 0.90 (Fig. 4E), whereas for *Rhithrogena* spp., the interval ranged from $\tau = 0.25$ to 0.90 (Fig. 4F). Other limiting factors also caused densities to fall from the maximum expected in the absence of other limiting factors to below the median quantile ($\tau = 0.50$), a result supporting the idea that other limiting factors not included in statistical models predicting mean indicator response can attenuate the observed rate of change or slope. Δ_i was not calculated for OLS models because we were not interested in making comparisons among these models.

Deriving inference from multiple quantiles

Evaluating the effects of metals on density in terms of risk can offer an alternative means for communicating results to managers and show how risks are quantile dependent. We can describe changes in density in terms of the proportion of sites having densities within some interval of values by estimating the conditional cumulative distribution function from the regression quantile estimates. For example, the 90th quantile of density is a density $\geq$ densities observed in 90% of the streams sampled. We can evaluate how the differences in densities change depending on a particular level of acceptable risk by evaluating the difference in expected densities among quantiles or associated risks to streams. Presenting results in terms of risk shows that stream insect densities that are typical at background metal concentrations can be rare or improbable at the threshold for aquatic life (Fig. 5A–F). In 95% of the streams sampled, density (ind/0.1 m^2) of *Drunella* spp. was $\leq$81 (90% CI = 55–111) at CCAR = 0.1 (Fig. 5B) and $\leq$46 (CI = 25–60) at CCAR = 1 (Fig. 5E). Similarly, in 95% of the streams sampled, the density of *Rhithrogena* spp. was $\leq$87 (90% CI = 57–138) at CCAR = 0.1 (Fig. 5C) and $\leq$35 (CI = 27–47) at CCAR = 1 (Fig. 5F).

Limiting inference to the upper-most regression quantile can produce misleading results. For example, % reductions in the predicted density of *Rhithrogena* spp. between CCAR = 0.1 and CCAR 1.0 were 57% and 59% for regression quantiles $\tau = 0.90$ and 0.95, respectively (Fig. 5C, F). The difference in the % reduction in *Drunella* spp. density was much greater (60% and 43% for $\tau = 0.90$ and 0.95, respectively; Fig. 5B, E). There is less certainty (e.g., wider 90% CI) in the risk estimates in these upper extreme regression quantiles. Moreover, had we not investigated other regression quantiles, we would not have learned that effects in $\tau > 0.70$ are greater than those at or below the median and mean.

Discussion

Our aim was to use quantile regression as a means to learn what drives changes in biological indicator values beyond estimating change in the upper-most extreme quantile. We found that metals limited metal-sensitive mayflies, whereas basin area and other factors limited the metal-tolerant caddisfly. Slopes or rates of change differed among quantiles, and in many cases, the mean rate of change (OLS slope estimate) underestimated effects on high quantiles of density. These results indicated that measured and unmeasured limiting factors (e.g., life history, seasonality, presence of fish, antecedent conditions) contributed to variance in density responses. These other factors mostly affected quantiles below the mean, thereby obscuring the detection of change in the response by causing low values of mayfly density at low metal concentrations, whereas the greatest effects were observed at quantiles above the mean. Thus, the effect of metals on aquatic populations varied greatly among streams because these other factors limited response outcomes.

We identified the data space where biological indicators were not limited by metals, and this approach improved our certainty regarding seemingly contradictory indicator responses. For example, the interval between the minimum and maximum regression quantiles where CCAR was the top model and where regression quantile slopes were significantly different from 0 is the data space most reliably limited by metals. Both *A. grandis* and *Rhithrogena* spp. were absent or occurred at low densities at levels of metals where we also observed high densities of these taxa (Fig. 4D, F). By recognizing that many of the low densities and absences at low levels of metals (CCAR < 1) were caused by other factors, we improved our certainty about what level of metals is safe for these populations. We also found that low densities of *Drunella* spp. were less common in background or reference streams (Figs 1, 4E). Therefore, density or presence/absence of this taxon was a more reliable

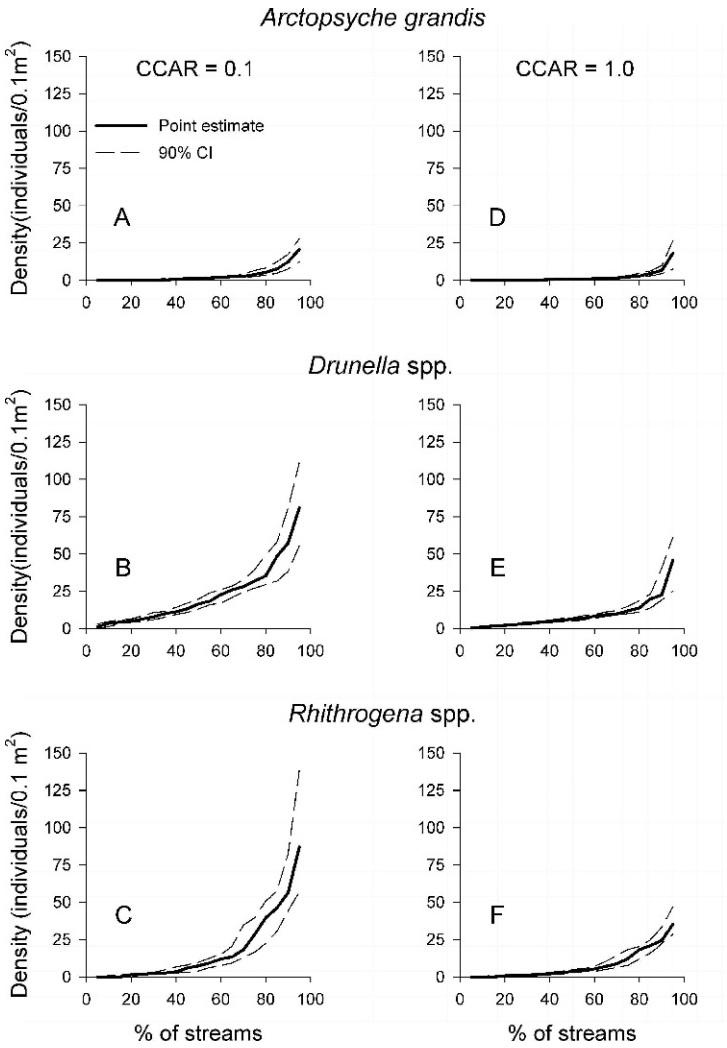

FIG. 5. Point estimates (90% confidence intervals [CIs]) of density of *Arctopsyche grandis* (A, D), *Drunella* spp. (B, E), and *Rhithrogena* spp. (C, F) densities estimated from quantile regression models at background (chronic criterion accumulation ratio [CCAR] = 0.1) (A, B, C) and at the toxic threshold (CCAR = 1.0) (D, E, F). Point estimates were derived from quantile regression models (τ = 0.05–0.95 by increments of 0.05; Fig. 3A). *x*-axis (% of streams) is the inverse of the quantile (τ).

indicator of metals pollution than density or presence/absence of the other taxa evaluated.

Natural variability in aquatic insect communities resulting from stream physiography can limit detection of subtle changes in aquatic communities at low metal concentrations (Kiffney and Clements 1994, Clements and Kiffney 1995). An important factor not included in our study is the effect of population phenology on heterogeneity in density responses (Kiffney and Clements 1994, 1996, Clark and Clements 2006). Aquatic insect populations can be composed of multiple cohorts of the same taxa, and smaller, immature instars are more sensitive to metals than are older more mature individuals (Kiffney and Clements 1994, 1996, Clark and Clements 2006). In a mesocosm experiment, Kiffney and Clements (1994) observed a 70% decline in the density of small instars and no change in large instars of *Drunella grandis* exposed to metal mixtures of Cd, Cu, and Zn near US EPA chronic aquatic life criteria. This high variability in a controlled laboratory experiment shows that some of the heterogeneity in aquatic insect densities in our study probably was caused by defining these populations based on taxonomy without regard for life history. We probably sampled different instars of *Drunella* spp. and *Rhithrogena* spp., and less mature and more sensitive larvae probably were abundant at some sites with high densities of mayflies. We did not measure instars and did not account for differences in sensitivity among instars, so differences in population structure among sites could account for variability in our mayfly densities.

We question the traditional approach of investigating only mean responses of populations and communities to contaminants. We expected to see maximum densities at sites where contaminants were not limiting. However, densities of mayflies at these sites were often limited by other factors, and in some cases, deviated significantly from expected values. Large decreases in mean density at higher metal concentrations were required to detect a statistical difference because of high variance in density at low metal concentrations. However, we did not observe large decreases in mean density as metal concentrations increased because other limiting factors attenuated regression slopes, particularly for quantiles $\tau \leq 0.70$ (Figs 3A, 4D–F, 5A–F). We should not expect to observe great changes in the mean response until the effect of the contaminant is so great that all other potential limiting factors are statistically unimportant. In other words, we should expect differences in mean responses only when contaminant concentrations are high enough to dominate other ecological processes. Furthermore, we expect multiple chemicals or physical disturbances to affect these responses as much as natural limiting factors do. Multiple stressors probably would attenuate the relationship between contaminants and response by acting on both low and high quantiles of the response. As a result, the effect of multiple chemical stressors on streams will become increasingly difficult to detect.

We should not assume that the effect of metals (or perhaps any contaminant) will be the same across all streams. This assumption can greatly affect scientific inference and decision making. In our study, metals caused greater declines in high than in low quantiles of mayfly densities (Figs 3A, 5A–F). These differences among quantiles probably arose because the effects of other limiting factors, possibly differences in habitat, were not equal across sites. Thus, the effect of metals on density will not be the same across these habitats. Mean-based regression approaches that assume homogeneity of variance and that all quantiles of the response have parallel slopes would not have detected this important characteristic of the data. In fact, had we analyzed these data only with a mean-based statistical model, the heterogeneity observed in the data probably would have been perceived as a problem, either a violation of parametric model assumptions (i.e., heterogeneous variance) or as residual error (i.e., poor model fit).

The practice of limiting inference to a few regression quantiles can be problematic. For example, inferences made solely on $\tau = 0.90$ for *Drunella* spp. would have resulted in a much larger loss estimate than inferences made on $\tau = 0.95$ (Fig. 4E). By analyzing multiple regression quantiles, we demonstrated that most of the losses in mayfly densities occurred in the intermediate-to-higher regression quantiles ($\tau \geq 0.70$) and not just the upper limits of density. Furthermore, 90% CIs generally were wider at extreme quantiles (e.g., $\tau = 0.05$ or 0.95) than at intermediate quantiles (e.g. $\tau = 0.50$; Fig. 3A–E), suggesting greater uncertainty in these extreme quantiles (Cade and Noon 2003). Moreover, by considering all the quantiles ($\tau = 0$–0.95), we were able to express model predictions in probabilistic terms allowing us to compare risks among exposure scenarios (Fig. 5A–F). Our results showed that high densities of mayflies commonly observed at background metal concentrations are rare at concentrations of metals previously thought safe for aquatic life (CCAR = 1). Had we arbitrarily selected the most extreme quantile to derive statistical inference and not considered other quantiles, we would have missed these important findings. We strongly suggest considering multiple quantiles of the response to understand better the effects of contaminants on populations, and we caution against arbitrary selection of the most extreme quantile that can be estimated as the way to draw inference from the data.

Quantile regression is not a universal tool that can solve all ecological data problems, but it does have several advantages over other statistical methods (Cade and Noon 2003). Some of these advantages are estimating effects on characteristics of the response other than the mean, relaxed parametric model assumptions (Cade and Noon 2003), and a model form that links observed change in resources to a key concept from ecology and ecotoxicology, limiting factors. Other techniques, such as propensity scores and hierarchical linear modeling, are appropriate statistical techniques for dealing with data heterogeneity. However, these model forms measure change only in means, they do not necessarily measure change where it is the greatest, and they do not link well with the theory of limiting factors (Cade and Noon 2003).

By necessity or convenience, not all the factors that substantially influence biomonitoring endpoints can be measured in biomonitoring studies. Thus, the concept of limiting factors can be very useful when interpreting variability in endpoints. This concept helped to clarify that although metals limited high quantiles of mayflies, other factors (e.g., habitat) also limited response outcomes. This effect was especially strong for *A. grandis.* High quantiles of *A. grandis* density were limited by basin area (more precisely, by stream factors correlated with basin area), but many other quantiles ($\tau < 0.40$) were unrelated to the metals

gradient and the other limiting factors we measured (e.g., temperature). Unmeasured factors caused the density of *A. grandis* to deviate from the value expected were basin area the only limiting factor.

Responses to metals are not limited to measures of ecosystem structure. In addition to reducing the proportion of metal-sensitive taxa and altering other measures of community structure, metals can alter ecological functioning (Clements et al. 2000, Schmidt et al. 2010). For example, contaminant-induced losses in the density of aquatic insect larvae can disrupt detritus processing, invertebrate secondary production, and the flow of energy into aquatic food webs (Carlisle and Clements 2005). In-stream disturbances that reduce larval densities also reduce the density of emerging aquatic-insect adults that subsidize riparian consumers (Paetzold et al. 2011). Our results indicate that the effects of contaminants on ecosystem structure and function should be most obvious on high quantiles of the response.

Our results indicate that scientists and managers have been trying to detect contaminant-induced changes in biological indicators with models that do not necessarily measure change where it is most obvious. Metals limit ecological potential, but they do not necessarily change the realized average potential of ecosystems. Advances in stream and ecosystem science will be limited if the model used to measure change cannot detect it. Effectively implemented policy cannot be derived from ineffective models, and confidence in policy will be eroded if outcomes are vastly different from expectation.

Our study and others show that quantile regression can make biomonitoring data more useful for establishing regulatory limits for pollutants in aquatic ecosystems (Pacheco et al. 2005, Linton et al. 2007). Water-quality criteria are derived from laboratory toxicity tests that determine species-specific responses to a contaminant. Genus-mean responses are selected from a number of families so that information from a theoretical community of organisms can be used to develop a value protective of 95% of the species observed in nature (Stephan et al. 1985). However, effects in the field are being observed increasingly often at concentrations of contaminants below aquatic life standards (Schmidt et al. 2010, 2011). This mismatch probably reflects the limitations of extrapolating results of laboratory toxicity tests with aquatic invertebrates to field situations (Buchwalter et al. 2007), although Wang et al. (2009) was able to test sensitive life stages of sensitive taxa. The US EPA considers field-monitoring data when developing water-quality criteria (Stephan et al. 1985), but variability and reliability of these data in the presence of other factors probably have limited their utility. Quantile regression can help overcome this limitation, and its use for development or validation of standards based on field data probably will increase as recognition of limiting factors and available software increase.

Acknowledgements

We thank James Terrell, Christopher Konrad, and anonymous reviewers for insightful comments that greatly improved this manuscript. Support for this research came from the USGS Central Colorado Assessment Project, the Mendenhall Post-Doctoral Fellowship program, Central Region Mineral and Environmental Resources Science Center, and the Colorado Water Science Center. Disclaimer: any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

Literature Cited

BUCHWALTER, D. B., D. J. CAIN, W. H. CLEMENTS, AND S. N. LUOMA. 2007. Using biodynamic models to reconcile differences between laboratory toxicity tests and field biomonitoring with aquatic insects. Environmental Science and Technology 41:4821–4828.

BURNHAM, K. P., AND D. R. ANDERSON. 2002. Model selection and multimodel inference. 2nd edition. Springer, New York.

CADE, B. S., AND B. R. NOON. 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1:412–420.

CADE, B. S., B. R. NOON, AND C. H. FLATHER. 2005. Quantile regression reveals hidden bias and uncertainty in habitat models. Ecology 86:786–800.

CAIN, D. J., S. N. LUOMA, AND W. G. WALLACE. 2004. Linking metal bioaccumulation of aquatic insects to their distribution patterns in a mining-impacted river. Environmental Toxicology and Chemistry 23:1463–1473.

CARLISLE, D. M., AND W. H. CLEMENTS. 1999. Sensitivity and variability of metrics used in biological assessment of running waters. Environmental Toxicology and Chemistry 18:285–291.

CARLISLE, D. M., AND W. H. CLEMENTS. 2005. Leaf litter breakdown, microbial respiration and shredder productions in metal-polluted streams. Freshwater Biology 50: 380–390.

CLARK, J. L., AND W. H. CLEMENTS. 2006. The use of in situ and stream microcosm experiments to assess population- and community-level responses to metals. Environmental Toxicology and Chemistry 25:2306–2312.

CLEMENTS, W. H. 1994. Benthic invertebrate community responses to heavy metals in the Upper Arkansas River Basin, Colorado. Journal of the North American Benthological Society 13:30–44.

CLEMENTS, W. H. 2004. Small-scale experiments support causal relationships between metal contamination and macroinvertebrate community responses. Ecological Applications 14:954–967.

CLEMENTS, W. H., D. M. CARLISLE, J. M. LAZORCHAK, AND P. C. JOHNSON. 2000. Heavy metals structure benthic communities in Colorado mountain streams. Ecological Applications 10:626–638.

CLEMENTS, W. H., AND P. M. KIFFNEY. 1995. The influence of elevation on benthic community responses to heavy metals in Rocky Mountain streams. Canadian Journal of Fisheries and Aquatic Sciences 52:1966–1977.

DI TORO, D. M., H. E. ALLEN, H. L. BERGMAN, J. S. MEYER, P. R. PAQUIN, AND R. C. SANTORE. 2001. Biotic ligand model of the acute toxicity of metals. 1. Technical basis. Environmental Toxicology and Chemistry 20:2383–2396.

FORRESTER, G. E. 1994. Influences of predatory fish on the drift, dispersal, and local-density of stream insects. Ecology 75:1208–1218.

HYDROQUAL. 2007. Biotic ligand model: Windows interface user's guide and reference manual, Version 2.2.3. HydroQual, Mahwah, New Jersey.

KIFFNEY, P. M., AND W. H. CLEMENTS. 1993. Bioaccumulation of heavy metals by benthic invertebrates at the Arkansas River, Colorado. Environmental Toxicology and Chemistry 12:1507–1517.

KIFFNEY, P. M., AND W. H. CLEMENTS. 1994. Structural responses of benthic invertebrate communities from different stream orders to zinc. Environmental Toxicology and Chemistry 13:389–395.

KIFFNEY, P. M., AND W. H. CLEMENTS. 1996. Size-dependent response of macroinvertebrates to metals in experimental streams. Environmental Toxicology and Chemistry 13:389–395.

KOENKER, R. 2005. Quantreg: quantile regression. Version 4.17. R Project for Statistical Computing, Vienna, Austria. (Available from: http://cran.r-project.org/web/packages/quantreg/index.html)

KOENKER, R., AND G. BASSETT. 1978. Regression quantiles. Econometrica 46:33–50.

LINTON, T. K., M. A. W. PACHECO, D. O. MCINTYRE, W. H. CLEMENT, AND J. GOODRICH-MAHONEY. 2007. Development of bioassessment-based benchmarks for iron. Environmental Toxicology and Chemistry 26:1291–1298.

MERRITT, R. W., AND K. W. CUMMINS (EDITORS). 1996. An introduction to the aquatic insects of North America. 3rd edition. Kendall/Hunt, Dubuque, Iowa.

MOULTON, S. R., J. L. CARTER, S. A. GROTHEER, T. F. CUFFNEY, AND T. M. SHORT. 2000. Methods of analysis by the U.S. Geological Survey national water quality laboratory: processing, taxonomy, and quality control of benthic macroinvertebrate samples. Open-File Report 00-212. US Geological Survey, Reston, Virginia.

MUTEL, C. F., AND J. C. EMERICK. 1992. From grassland to glacier: the natural history of Colorado and the surrounding region. Johnson Printing, Boulder, Colorado.

PACHECO, M. A. W., D. O. MCINTYRE, AND T. K. LINTON. 2005. Integrating chemical and biological criteria. Environmental Toxicology and Chemistry 24:2983–2991.

PAETZOLD, A., M. SMITH, P. H. WARREN, AND L. MALTBY. 2011. Environmental impact propagated by cross-system subsidy: chronic stream pollution controls riparian spider populations. Ecology 92:1711–1716.

RANTZ, S. E., H. H. BARNES, T. J. BUCHANAN, R. W. CARTER, F. A. KILPATRICK, G. F. SMOOT, E. D. COBB, H. F. MATTHAI, M. A. BENSON, A. F. PENDLETON, T. DALRYMPLE, H. HULSING, C. E. KINDSVATER, G. L. BODHAINE, H. J. TRACY, J. DAVIDIAN, AND J. F. WILSON. 1982. Measurement and computation of streamflow. Volume 1: measurement of stage and discharge. U.S. Geological Survey Water-Supply Paper 2175. US Geological Survey, Reston, Virginia.

RUNNELLS, D. D., T. A. SHEPHERD, AND E. E. ANGINO. 1992. Metals in water: determining natural background concentrations in mineralized areas. Environmental Science and Technology 26:2316–2323.

SCHMIDT, T. S., W. H. CLEMENTS, K. MITCHELL, S. E. CHURCH, R. B. WANTY, D. L. FEY, P. L. VERPLANCK, AND C. A. SAN JUAN. 2010. Development of a new toxic-unit model for the bioassessment of metals in streams. Environmental Toxicology and Chemistry 29:2432–2442.

SCHMIDT, T. S., S. E. CHURCH, R. B. WANTY, P. L. VERPLANCK, C. A. SAN JUAN, D. FEY, B. W. ROCKWELL, E. H. DEWITT, AND T. J. KLEIN. 2012. Geologic processes influence the effects of mining on aquatic ecosystems. Ecological Applications 22:870–879.

SCHMIDT, T. S., W. H. CLEMENTS, R. E. ZUELLIG, K. A. MITCHELL, S. E. CHURCH, R. B. WANTY, C. A. SAN JUAN, M. ADAMS, AND P. J. LAMOTHE. 2011. Critical tissue residue approach linking accumulated metals in aquatic insects to population and community-level effects. Environmental Science and Technology 45:7004–7010.

SOIL SURVEY STAFF. 1999. Soil taxonomy: a basic system of soil classification for making interpretive soil surveys. Agricultural handbook 436. US Department of Agriculture, Washington, DC.

STEPHAN, C. E., D. I. MOUNT, D. J. HANSEN, J. H. GENTILE, G. A. CHAPMAN, AND W. A. BRUNGS. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. PB85-227049. National Technical Information Service, Springfield, Virginia.

TAGGART, J. E. (EDITOR). 2002. Analytical methods for chemical analysis of geologic and other materials. U.S. Geological Survey Open-File Report 02-223. US Geological Survey, Reston, Virginia.

TERRELL, J. W., B. S. CADE, J. CARPENTER, AND J. M. THOMPSON. 1996. Modeling stream fish habitat limitations from wedge-shaped patterns of variation in standing stock. Transactions of the American Fisheries Society 125: 104–117.

WANG, N., C. A. MEBANE, J. L. KUNZ, C. G. INGERSOLL, T. W. MAY, W. R. ARNOLD, R. C. SANTORE, T. AUGSPURGER, J. DWYER, AND M. C. BARNHART. 2009. Evaluation of acute copper toxicity to juvenile freshwater mussels (fatmucket, *Lampsilis siliquoidea*) in natural and reconstituted waters. Environmental Toxicology and Chemistry 28: 2367–2377.

WARD, J. V., B. C. KONDRATIEFF, AND R. E. ZUELLIG. 2002. An illustrated guide to the mountain stream insects of Colorado. University of Colorado Press, Niwot, Colorado.

WILDE, F. D., AND D. B. RADTKE. 1998. Field measurements. Book 9, Chapter A6 *in* F. D. Wilde, D. B. Radtke, J. Gibs, and R. T. Iwatsubo (editors). National field manual for the collection of water-quality data. U.S. Geological Survey Techniques of Water Resources Investigations. US Geological Survey, Reston, Virginia. (Available from: http://pubs.water.usgs.gov/twri9A)

WILDE, F. D., D. B. RADTKE, J. GIBS, AND R. T. IWATSUBO. 1998. Field measurements. Book 9, Chapter A4 *in* F. D. Wilde, D. B. Radtke, J. Gibs, and R. T. Iwatsubo (editors). National field manual for the collection of water-quality data. U.S. Geological Survey Techniques of Water Resources Investigations. US Geological Survey, Reston, Virginia. (Available from: http://pubs.water.usgs.gov/twri9A)

Received: 4 October 2011
Accepted: 5 March 2012