

Multilevel regression models describing regional patterns of invertebrate and algal responses to urbanization across the USA

Authors: Cuffney, Thomas F., Kashuba, Roxolana, Qian, Song S., Alameddine, Ibrahim, Cha, Yoon Kyung, et al.

Source: Journal of the North American Benthological Society, 30(3) : 797-819

Published By: Society for Freshwater Science

URL: <https://doi.org/10.1899/10-140.1>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Multilevel regression models describing regional patterns of invertebrate and algal responses to urbanization across the USA

Thomas F. Cuffney¹

US Geological Survey, 3916 Sunset Ridge Road, Raleigh, North Carolina 27607 USA

**Roxolana Kashuba², Song S. Qian³, Ibrahim Alameddine⁴,
YoonKyung Cha⁵, AND Boknam Lee⁶**

*Nicholas School of the Environment and Earth Sciences, Duke University, Durham,
North Carolina 27708 USA*

James F. Coles⁷

US Geological Survey, 331 Commerce Way, Suite 2, Pembroke, New Hampshire 03275 USA

Gerard McMahon⁸

US Geological Survey, 3916 Sunset Ridge Road, Raleigh, North Carolina 27607 USA

Abstract. Multilevel hierarchical regression was used to examine regional patterns in the responses of benthic macroinvertebrates and algae to urbanization across 9 metropolitan areas of the conterminous USA. Linear regressions established that responses (intercepts and slopes) to urbanization of invertebrates and algae varied among metropolitan areas. Multilevel hierarchical regression models were able to explain these differences on the basis of region-scale predictors. Regional differences in the type of land cover (agriculture or forest) being converted to urban and climatic factors (precipitation and air temperature) accounted for the differences in the response of macroinvertebrates to urbanization based on ordination scores, total richness, Ephemeroptera, Plecoptera, Trichoptera richness, and average tolerance. Regional differences in climate and antecedent agriculture also accounted for differences in the responses of salt-tolerant diatoms, but differences in the responses of other diatom metrics (% eutraphenic, % sensitive, and % silt tolerant) were best explained by regional differences in soils (mean % clay soils). The effects of urbanization were most readily detected in regions where forest lands were being converted to urban land because agricultural development significantly degraded assemblages before urbanization and made detection of urban effects difficult. The effects of climatic factors (temperature, precipitation) on background conditions (biogeographic differences) and rates of response to urbanization were most apparent after accounting for the effects of agricultural development. The effects of climate and land cover on responses to urbanization provide strong evidence that monitoring, mitigation, and restoration efforts must be tailored for specific regions and that attainment goals (background conditions) may not be possible in regions with high levels of prior disturbance (e.g., agricultural development).

Key words: Multilevel hierarchical regression, urbanization, benthic macroinvertebrates, algae, diatoms, agriculture, multiple spatial scales.

¹ E-mail addresses: tcuffney@usgs.gov

² rkashuba@usgs.gov

³ song@duke.edu

⁴ ima4@duke.edu

⁵ yoonyung.cha@duke.edu

⁶ boknam.lee@duke.edu

⁷ jcoles@usgs.gov

⁸ gmcMahon@usgs.gov

An increasing proportion of the global population is concentrated in expanding urban areas that affect the environment at local, regional, and global scales in a manner disproportionate to their representation in the landscape (Grimm et al. 2008, Pickett et al. 2008). The deleterious effects of urbanization on stream ecosystems have been well documented, but only a few investigators have studied large-scale patterns of responses to urbanization by comparing stream

conditions among metropolitan areas qualitatively (Paul and Meyer 2001, Walsh et al. 2005, Wenger et al. 2009) or quantitatively (Cuffney et al. 2005, 2010, Potopova et al. 2005, Brown et al. 2009, Coles et al. 2009). Comparisons among metropolitan areas have shown that responses can vary substantially among urban areas and biological indicators (fish, benthic macroinvertebrates, and benthic algae), as can the environmental stressors (e.g., water chemistry, hydrology, temperature, and habitat) associated with biological responses (Jacobson et al. 2001, Sponseller et al. 2001, Morse et al. 2003, Vølstad et al. 2003, Mahler et al. 2005).

Detecting, understanding, and predicting large-scale patterns of responses to urbanization involve integrating effects that are driven by factors acting at multiple spatial scales (e.g., reach, basin, and regional). Simple regression (linear or generalized regression) analyses cannot effectively incorporate variables operating at different spatial scales because of problems incorporating cross-scale interactions. Multivariate analyses (e.g., ordinations) can incorporate multiscale environmental factors and identify environmental factors that are associated with changes in biological assemblages (Kratzer et al. 2006). However, multivariate analyses are not a particularly effective method for estimating model parameters that can be used to predict responses based on regional predictors. Mixed-effect models, such as multilevel hierarchical regression, provide a more effective method (Gelman and Hill 2007) for incorporating multiscale predictor variables in a model that can be used to predict responses (slopes and intercepts) at multiple, hierarchically related scales (e.g., basin and region). Multilevel hierarchical regression provides an effective method for developing models for understanding and predicting responses that arise from the interplay of fine- and broad-scale processes, which is fundamental to understanding ecosystem dynamics (Peters et al. 2008).

We used multilevel hierarchical regression to incorporate both basin- and region-scale predictor variables into models of invertebrate and algal responses to urbanization in 9 metropolitan areas across the continental USA (Boston, Massachusetts [BOS]; Raleigh, North Carolina [RAL]; Atlanta, Georgia [ATL]; Birmingham, Alabama [BIR]; Milwaukee-Green Bay, Wisconsin [MGB]; Denver, Colorado [DEN]; Dallas-Fort Worth, Texas [DFW]; Salt Lake City, Utah [SLC]; and Portland, Oregon [POR]; Fig. 1). Each of these metropolitan areas is an aggregation of cities, towns, and villages that represent the process of urbanization in regions of the country that potentially differ in natural vegetation, temperature, precipitation, basin relief, elevation, and basin slope (Table 1). The objectives of our study were

to determine: 1) if selected regional variables could explain the differences in the responses of invertebrates and algae to urbanization among these 9 metropolitan areas (Brown et al. 2009, Coles et al. 2009, Cuffney et al. 2010), and 2) the relevance of regional factors for understanding large-scale patterns in the effects of urbanization. This study was conducted as a part of the US Geological Survey (USGS) National Water-Quality Assessment (NAWQA) Program.

Methods

The 9 urban studies had a common study design (Tate et al. 2005) that used nationally available geographic information system (GIS) data (Falcone et al. 2007) to define a population of candidate basins (typically basins drained by 2nd- or 3rd-order streams) from which study basins were selected to represent a gradient of urbanization. Natural environmental variability was minimized within each metropolitan area by dividing candidate basins into groups with relatively homogeneous natural environmental features (e.g., ecoregion, climate, elevation, stream size). Urban intensity was quantified based on a multimetric index (metropolitan area national urban-intensity index [MA-NUII]) that combined housing density, % basin area in developed land cover, and road density into an index scaled to range from 0 (little or no urban) to 100 (maximum urban) within each metropolitan area (Cuffney and Falcone 2008). Once groups of basins with relatively homogeneous environmental features were defined (candidate basins), 28 to 30 basins were selected to represent the gradient of urbanization in each metropolitan area (Fig. 2). This spatially distributed sampling network was intended to represent changes in urbanization through time (i.e., substitute space for time).

Conditions in each basin were verified by field reconnaissance. If conditions in a basin deviated substantially from what was expected, a sampling reach (150-m stream section at the base of the basin) was disturbed by local-scale effects (e.g., channelization, road or building construction), or land owners denied access, then an alternate basin from the same group or a group with similar natural environmental characteristics was selected to represent the same level of urban intensity. Studies were conducted during 1999–2000 (BOS, BIR, and SLC), 2000–2003 (ATL, DEN, and RAL), and 2003–2004 (DFW, MGB, and POR). Details of the study designs can be found in Coles et al. (2004), Cuffney et al. (2005, 2010), Tate et al. (2005), and Giddings et al. (2009).

The SLC design differed from that of the other metropolitan areas in that many of the basins were nested one within another. This modification was

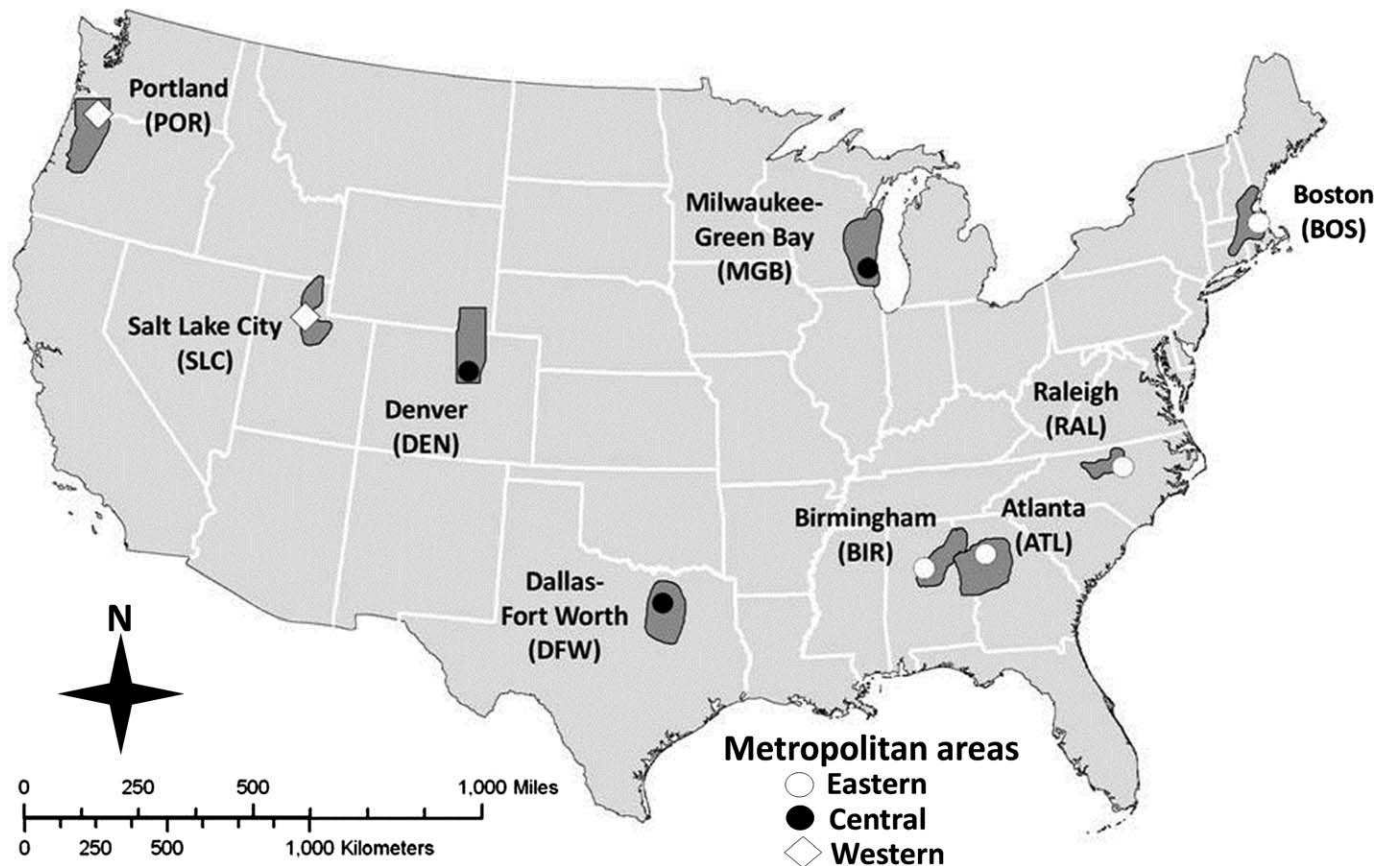


FIG. 1. Nine metropolitan areas in which urban studies were conducted. Shaded areas show the spatial extent of each metropolitan area.

necessary because of the very small number of streams that are present in SLC and was possible only because urban development in the SLC basins has progressed upstream over time ensuring that urban intensity increases downstream. The SLC landscape characterizations were restricted to the portions of the basins that were in the Central Basin and Range ecoregion (Omernik 1987). The portions in the Wasatch and Uinta Mountains ecoregion were excluded because no urban development has occurred in this area, and the biology and geomorphology of the streams in this ecoregion are very different from the Central Basin and Range ecoregion. Large reservoirs in DEN constituted major discontinuities that effectively isolated the upper and lower portions of many of the candidate basins. Consequently, landscape characterizations in DEN were restricted to the portions of the basins that were below the major reservoirs.

Macroinvertebrate and algae data

The NAWQA Program sampling protocols were used to collect quantitative benthic macroinvertebrate

and algae samples over a 1- to 4-wk period during summer base flows as described in Giddings et al. (2009). The USGS Invertebrate (IDAS) and Algal Data Analysis Systems (ADAS) (Cuffney and Brightbill 2011) were used to resolve taxonomic ambiguities (Cuffney et al. 2007) and to calculate assemblage metrics and diversity measures. Invertebrate tolerances from 4 different regions of the country (Barbour et al. 1999, NCDENR 2006) were used to calculate invertebrate tolerance metrics for the 9 metropolitan areas as follows: mid-Atlantic (BOS), southeast (ATL, BIR, RAL, DFW), midwest (MGB), and northwest (DEN, SLC, POR). Algal-assemblage metrics were based on national traits information compiled by Porter (2008). Data were converted to densities (no./m²) prior to resolving ambiguous taxa and calculating assemblage metrics. Nonmetric multidimensional scaling (Clarke and Gorley 2006) was used to obtain axis-1 site scores based on 4th-root-transformed density data and Bray-Curtis similarity. Axis-1 site scores were rescaled, where necessary, to standardize the direction of response so that site scores decreased as urban intensity increased (Brown et al. 2009, Cuffney et al.

TABLE 1. Major environmental characteristics of the 9 metropolitan areas (Falcone et al. 2007). Environmental characteristics used as regional variables in the hierarchical models are shown as mean ($\pm 95\%$ CI). TEMP = mean annual temperature; PRECIP = mean annual precipitation; AAG = antecedent agriculture; CLAY = % clay soil in basin; Relief = basin relief (range in elevation); BOS = Boston, Massachusetts; RAL = Raleigh, North Carolina; ATL = Atlanta, Georgia; BIR = Birmingham, Alabama; MGB = Milwaukee-Green Bay, Wisconsin; DEN = Denver, Colorado; DFW = Dallas-Fort Worth, Texas; SLC = Salt Lake City, Utah; POR = Portland, Oregon; NECZ = New England Coastal Zone; PIED = Piedmont; SWA = Southwest Appalachians; SEWT = Southeastern Wisconsin Till Plains; HP = High Plains; TBP = Texas Blackland Prairies; CBR = Central Basin and Range; WV = Willamette Valley.

| Metropolitan area | Predominant ecoregion | Natural vegetation | TEMP (°C) | PRECIP (cm) | AAG | CLAY | Relief (m) | Basin elevation (m) | Basin slope (%) |
|-------------------|-----------------------|--------------------|----------------|--------------|------------|----------------|------------|---------------------|-----------------|
| BOS | NECZ | Forest | 8.7 \pm 0.3 | 123 \pm 2 | 10 \pm 1 | 5.1 \pm 0.4 | 188 | 113 | 5.9 |
| RAL | PIED | Forest | 14.9 \pm 0.1 | 119 \pm 1 | 24 \pm 1 | 37.1 \pm 0.8 | 83 | 180 | 5.4 |
| ATL | PIED | Forest | 16.3 \pm 0.2 | 134 \pm 2 | 17 \pm 2 | 34.0 \pm 0.3 | 116 | 278 | 6.1 |
| BIR | SWA | Forest | 16.0 \pm 0.2 | 147 \pm 1 | 15 \pm 1 | 34.4 \pm 1.4 | 197 | 234 | 10.6 |
| MGB | SEWT | Forest | 7.6 \pm 0.2 | 86 \pm 1 | 79 \pm 2 | 33.7 \pm 2.1 | 64 | 236 | 2.2 |
| DEN | HP | Grass/shrub | 9.2 \pm 0.3 | 43 \pm 1 | 88 \pm 3 | 26.8 \pm 2.2 | 323 | 1704 | 6.2 |
| DFW | TBP | Grass/shrub | 18.3 \pm 0.1 | 104 \pm 2 | 82 \pm 1 | 48.4 \pm 1.8 | 79 | 172 | 2.3 |
| SLC | CBR | Grass/shrub | 9.7 \pm 0.5 | 68 \pm 4 | 12 \pm 8 | 17.5 \pm 2.8 | 253 | 1487 | 17.8 |
| POR | WV | Forest | 10.8 \pm 0.3 | 152 \pm 11 | 20 \pm 1 | 32.9 \pm 3.3 | 418 | 220 | 13.4 |

2010). Ordination analysis also was used to identify and remove outliers from the data. On this basis, 1 Denver site (413659104370001 Bear Creek above Little Bear Creek, near Phillips, Wyoming) was dropped from the analyses. The 4 metrics used to represent invertebrate and algal responses (Table 2) were chosen on the basis of previous analyses that had established these metrics as responsive to urbanization (Brown et al. 2009, Cuffney et al. 2010).

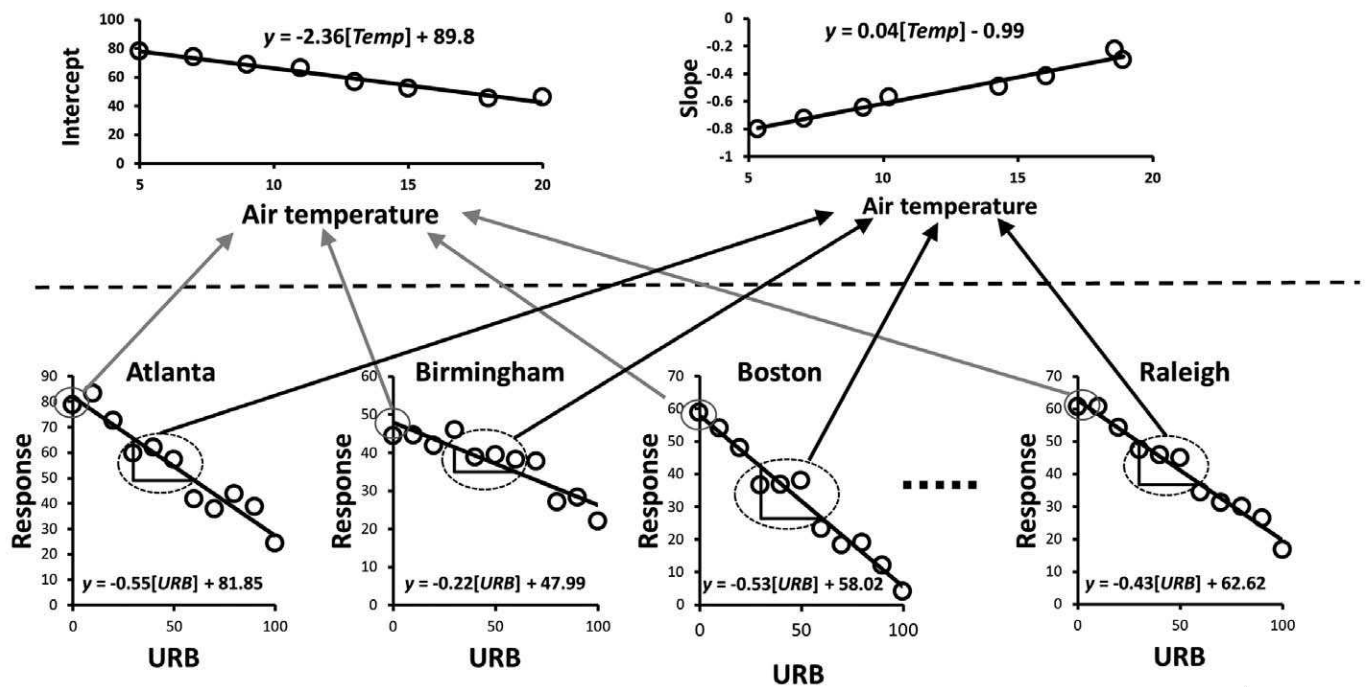
Environmental data

Land-cover data were based on the National Land Cover Data 2001 (NLCD01) data set and processing protocols (USGS 2005, Falcone et al. 2007). Antecedent agriculture (AAG) estimates in each region were obtained by averaging % basin area in row crop, pastures, and grasslands for the population of candidate basins that had low levels of urbanization (MA-NUII \leq 10). This procedure assumes that the mix of land cover that exists today is representative of the mix of land cover that existed when the basins were first developed. AAG could be defined only for basins with low levels of urbanization, and it showed much more variance among regions than within (Table 1), so AAG was categorized as a regional variable. Our previous work (Cuffney et al. 2010) has indirectly linked AAG to differences in invertebrate responses to urbanization and has shown that AAG divides metropolitan areas into 2 categories: high ($\geq 79\%$ basin area) and low ($< 25\%$) AAG (Cuffney et al. 2010). For this reason, AAG was modeled as both a continuous variable and as a categorical variable.

Soils data (% clay soils in the basin) were summarized from STATSCO data (Falcone et al. 2007). The region-level clay soils variable (CLAY) was obtained by averaging the basin-level values within each metropolitan area (region). CLAY was selected over other soil characteristics because it provides a measure of the potential for streams to be affected by fine sediment if perturbed by urbanization.

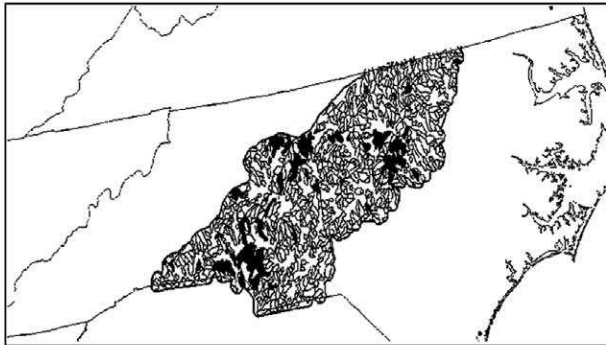
Mean annual precipitation (cm) and air temperature (°C) were derived for each of the study basins on the basis of 1-km-resolution DAYMET model data (Daymet 2005). These data represented 18-y (1980–1997) temperature and precipitation means obtained from terrain-adjusted daily climatological observations (Falcone et al. 2007). Region-level mean annual temperatures (TEMP) and precipitation (PRECIP) were obtained by averaging mean annual temperatures for the study basins in each metropolitan area. TEMP and PRECIP represent the influence of regional climatic factors on the responses of invertebrates and algae.

**Region-level response (basin-level slope or intercept) =
 $f(\text{region-level mean air temperature})$**

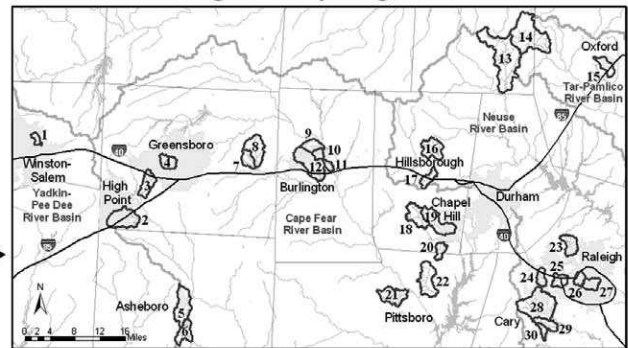


Basin-level response (abundance or richness) = $f(\text{basin-level URB})$

North Carolina Piedmont candidate basins



Raleigh sampling basins



Sampling design: Raleigh metropolitan area

FIG. 2. Schematic diagram showing the relations between the study design and the basin- and regional-scale analyses of the multilevel hierarchical regression. The study design (Raleigh [RAL]) involves identification of candidate basins and selection of sampling sites that represent a gradient of urbanization. Slopes and intercepts describing the relation between % developed land cover (URB) and invertebrate and algal responses are used to describe the basin-scale responses for each metropolitan area. The basin-scale responses (intercepts and slopes) are then used to describe the relations between regional-scale factors (e.g., air temperature) and the intercepts and slopes of the basin-scale responses.

TABLE 2. Variables used to model the responses of invertebrate and algal assemblages.

| Assemblage | Response variable |
|-----------------|---|
| Invertebrates | |
| NMDS1 | Nonmetric multidimensional scaling axis-1 ordination site scores based on density (no./m ²) |
| RICH | Total taxon richness |
| EPT | Ephemeroptera, Plecoptera, Trichoptera taxon richness |
| RICHTOL | Average tolerance of taxa ($\sum_{i=1}^n \frac{TV_i}{n}$ where TV_i = tolerance value of taxon i , n = number of taxa in the sample) |
| Algae (diatoms) | |
| HALO | Halobiontic (salt tolerant, 0.9–9 g/L salinity) diatoms (Van Dam et al. 1994), proportion of total diatom density (cells/m ²) |
| EUTR | Eutrathenic (mesotrophic + eutrophic + polytrophic) diatoms (Van Dam et al. 1994), proportion of total diatom density (cells/m ²) |
| SILT | Silt tolerant (motile) diatoms, proportion of total diatom density (cells/m ²) |
| SENS | Sensitive diatoms (class 3a, Lange-Bertalot 1979), proportion of total diatom density (cells/m ²) |

Simple linear regression models

Invertebrate and algal responses to urbanization were modeled for each metropolitan area using simple linear regression in R (lm and glm; R Foundation for Statistical Computing, Vienna, Austria). Count variables (richness [RICH] and Ephemeroptera, Plecoptera, Trichoptera taxon richness [EPT]) were analyzed using Poisson regression (glm, family = poisson, link = log) to adjust for the skewness of the count data distributions and to avoid negative predicted values. Continuous variables (nonmetric multidimensional scaling axis 1 [NMDS1], richness-weighted invertebrate tolerance [RICHTOL], and algal metrics) were analyzed using lm. Algal metrics (% salt-tolerant diatoms [HALO], % eutrathenic [EUTR], % silt tolerant [SILT], and % sensitive [SENS]), which were expressed as proportions of total abundance, were logit-transformed prior to analysis to ensure equal uncertainty across all proportions. Urbanization in the regression analyses was represented by % basin area in developed land (URB) rather than the MA-NUII used in our previous work (Cuffney and Falcone 2008) to facilitate comparison of urban intensity in this study with other studies. URB is 1 of the 3 variables used to calculate the MA-NUII and is strongly correlated with that index ($r = 0.95$ – 0.99).

Multilevel hierarchical regression models

Multilevel hierarchical regression was used to incorporate hierarchically structured (multiple basins within multiple regions) predictor variables into the models of invertebrate and algal responses to urbanization. This regression method provides a hierarchical framework for identifying the effects of group-level

(regional) predictors on individual level outcomes (basin-level urbanization). From a statistical perspective, mixed-effect models (e.g., multilevel hierarchical regression) are more efficient at modeling multitiered effects than are fixed-effects models (e.g., multiple linear or generalized linear regression) that must use dummy variables and interaction terms to represent multitiered effects (Gelman and Hill 2007).

Multilevel hierarchical models estimate the effects of regional predictors (e.g., TEMP) by relating the slopes and intercepts of the basin-level responses to URB ($y_{ij} = \alpha_j + \beta_j[\text{URB}_{ij}]$) to regional predictors ($P_{\beta j}$ for slopes, $P_{\alpha j}$ for intercepts) using separate linear models for slopes ($\beta_j = \gamma_{\beta 0} + \gamma_{\beta 1} P_{\beta j}$) and intercepts ($\alpha_j = \gamma_{\alpha 0} + \gamma_{\alpha 1} P_{\alpha j}$). The linear predictors ($\gamma_{\alpha 0}$ and $\gamma_{\alpha 1}$ for intercepts, $\gamma_{\beta 0}$ and $\gamma_{\beta 1}$ for slopes) are hyperparameters that describe the relation between the region-level predictors and the basin-scale intercepts (α_j) and slopes (β_j) that describe the biology-to-urbanization relationships (Fig. 2). The hyperparameters $\gamma_{\alpha 0}$ and $\gamma_{\alpha 1}$ describe how background conditions (URB = 0) are affected by the regional predictor(s). For example, if TEMP is the regional predictor, the hyperparameter $\gamma_{\alpha 0}$ estimates background conditions at a TEMP of 0°C, and $\gamma_{\alpha 1}$ estimates the rate at which background conditions change as TEMP changes. The hyperparameters $\gamma_{\beta 0}$ and $\gamma_{\beta 1}$ describe how the rate of response to URB (slope) is affected by the regional predictor(s). For the TEMP example, $\gamma_{\beta 0}$ describes the estimated rate of response to URB at 0°C, and $\gamma_{\beta 1}$ estimates the rate at which the rate of response to URB changes as TEMP changes. Collectively, the hyperparameters can be used to predict basin-level biotic responses to URB based on region-level predictors.

Multilevel hierarchical models incorporate partial pooling (Appendix I, Eq. 1) in the estimation of regional intercepts (α_j) and slopes (β_j). Partial pooling

TABLE 3. Basin- and regional-scale variables used to define the 9 multilevel hierarchical regression models. The templates (1–3) that describe the model structure are presented in Appendix 1. URB = % developed land cover in the drainage basin, PRECIP = mean annual precipitation, TEMP = mean annual air temperature, AAG = antecedent agriculture, CLAY = mean % clay soils in basin.

| Model | Model structure | Basin-level predictor | Region-level predictor(s) | |
|---------|-----------------|-----------------------|------------------------------|--------------------------|
| | | | Basin intercept predictor(s) | Basin slope predictor(s) |
| Model 1 | Template 1 | URB | None | None |
| Model 2 | Template 2 | URB | PRECIP | PRECIP |
| Model 3 | Template 2 | URB | TEMP | TEMP |
| Model 4 | Template 2 | URB | TEMP | PRECIP |
| Model 5 | Template 2 | URB | Continuous AAG | Continuous AAG |
| Model 6 | Template 3 | URB | PRECIP, categorical AAG | PRECIP, categorical AAG |
| Model 7 | Template 3 | URB | TEMP, categorical AAG | TEMP, categorical AAG |
| Model 8 | Template 3 | URB | TEMP, categorical AAG | PRECIP, categorical AAG |
| Model 9 | Template 2 | URB | CLAY | CLAY |

represents a compromise between complete pooling (all metropolitan areas combined) and no pooling (separate regressions for each metropolitan area) of the data. The advantage of partial pooling is that the unexplained variance at each level (basin and region) can be reduced by using both individual-(basin) and group-level (region) predictors to estimate model parameters. This approach is particularly useful when sample sizes in some groups are very low.

Three model templates were used to develop the 9 multilevel hierarchical regression models (Table 3). Each template uses URB as the basin-scale predictor of urban intensity, but each template differs in how regional predictors are incorporated. Template 1 does not incorporate region-level predictors. Instead, the regional intercepts (α_j) and slopes (β_j) are based on the means of the intercepts (μ_α) and slopes (μ_β) for the 9 regions (Appendix I, Eqs 2, 3). Template 2 uses a single continuous regional predictor to estimate regional responses (intercepts and slopes) to URB (Appendix I, Eqs 4, 5). Template 3 uses 2 regional predictors, a continuous predictor and a categorical (high, low) representation of AAG to estimate regional response (slopes and intercepts) to URB (Appendix I, Eqs 6, 7). Regional AAG was included both as a continuous (model 5) and categorical (models 6, 7, and 8) predictor because AAG in metropolitan areas tended to be either high ($\geq 79\%$ of basin area in MGB, DEN, and DFW) or low ($< 25\%$ in BOS, RAL, BIR, ATL, SLC, and POR) (Cuffney et al. 2010). Detailed descriptions of the models are given in Appendix I, and the statistical concepts behind the use of these multilevel hierarchical models for analyzing responses along urban gradients are given in Kashuba et al. (2010) and Qian et al. (2010).

Most of the multilevel hierarchical regression models (2, 3, 5, 6, 7, and 9) use the same predictors to model

both the intercepts and slopes of the responses to URB (Table 3). However, different regional predictors can be used for slopes (P_{β_j}) and intercepts (P_{α_j}) as illustrated in models 4 and 8. The estimated regional parameters (α_j and β_j) were derived using the lmer (linear mixed-effects regression) and glmer (general linear mixed-effects regression) functions in the lme4 package in R (Bates and Maechler 2010). As with the simple linear regressions, responses derived from count data were modeled using Poisson multilevel hierarchical models (glmer), continuous response variables were modeled using lmer, and algal response measures (HALO, EUTR, SILT, and SENS) were logit-transformed prior to analysis with lmer.

Deviance information criteria (DIC) were used to compare the fit of the multilevel models derived for each assemblage metric. Comparisons of DIC values were simplified by expressing DIC relative to model 1 (no regional predictors) by subtracting DIC values for models 2 to 9 from the DIC value for model 1 (the larger the value, the better the fit relative to model 1 [no regional predictors]). DIC provides a mechanism for comparing the fit of models with differing numbers of predictor variables, but it is a relative and not an absolute measure of fit (Spiegelhalter et al. 2002), so comparisons of DIC are restricted to comparisons of models for each metric and cannot be used to compare models for different metrics.

Results

Simple regression (unpooled) models

The intercepts of the simple linear regression models provide an estimate of conditions prior to urbanization (URB = 0), i.e., an estimate of background conditions. The models of invertebrate responses showed that background conditions varied

TABLE 4. Summary of regressions between % developed land cover (URB) and invertebrate and algal response variables for each metropolitan area (unpooled regressions). Response variables are as in Table 2. NMDS1 and RICHTOL are based on linear regressions; RICH and EPT are based on Poisson regressions; HALO, EUTR, SILT, and SENS are based on logistic regressions. * = $p < 0.05$, ** = $p < 0.001$.

| Region | Invertebrates | | | Algae (diatoms) | | |
|--------|---------------|-----------|----------|-----------------|-----------|---------|
| | Response | Intercept | Slope | Response | Intercept | Slope |
| ATL | NMDS1 | 2.814 | -3.072** | HALO | -3.233 | 0.509 |
| BIR | | 2.418 | -2.311** | | -3.333 | 0.847* |
| BOS | | 2.557 | -3.572** | | -3.657 | 1.283** |
| DEN | | 1.666 | -0.486 | | -1.968 | 0.875 |
| DFW | | 2.121 | -0.972* | | -0.736 | -0.334 |
| MGB | | 1.886 | -0.705* | | -1.921 | 1.761* |
| POR | | 2.685 | -1.839** | | -2.813 | 0.174 |
| RAL | | 2.908 | -2.103** | | -3.168 | 1.376* |
| SLC | | 2.495 | -1.758** | | -3.209 | 1.811* |
| ATL | | RICH | 3.792 | | -0.620** | EUTR |
| BIR | 3.578 | | -0.282* | -0.150 | 0.514 | |
| BOS | 3.938 | | -1.509** | -2.675 | 4.299* | |
| DEN | 3.504 | | -0.161 | 0.534 | 0.570 | |
| DFW | 3.500 | | -0.165 | 1.304 | -0.025 | |
| MGB | 3.452 | | -0.287* | 2.611 | -0.918* | |
| POR | 3.473 | | -0.241* | 0.457 | 2.129* | |
| RAL | 3.644 | | -0.504** | -0.981 | 1.754* | |
| SLC | 3.712 | | -0.588** | -0.424 | 1.877* | |
| ATL | EPT | | 2.753 | -1.924** | SILT | |
| BIR | | 2.654 | -1.502** | -1.533 | | 1.346 |
| BOS | | 3.044 | -2.780** | -3.245 | | 2.574** |
| DEN | | 1.803 | -0.040 | -1.850 | | 1.532* |
| DFW | | 1.924 | -0.344 | 0.913 | | -0.159 |
| MGB | | 1.992 | -0.646* | -0.429 | | 0.595 |
| POR | | 2.548 | -1.653** | -1.906 | | 0.784 |
| RAL | | 2.489 | -1.574** | -2.123 | | 1.944* |
| SLC | | 2.834 | -1.800** | -2.379 | | 2.011* |
| ATL | | RICHTOL | 5.221 | 1.751** | | SENS |
| BIR | 4.555 | | 2.112** | 1.786 | -1.116 | |
| BOS | 4.048 | | 3.054** | 3.521 | -5.155** | |
| DEN | 6.033 | | 0.155 | 1.180 | -1.195* | |
| DFW | 6.937 | | -0.064 | -1.210 | -0.102 | |
| MGB | 5.360 | | 0.351 | 0.911 | -1.556* | |
| POR | 4.367 | | 1.982** | 1.366 | -1.148 | |
| RAL | 5.144 | | 1.566** | 0.171 | -1.277* | |
| SLC | 3.931 | | 2.831** | 2.557 | -1.903* | |

among metropolitan areas (Table 4). Areas in eastern (BOS, RAL, BIR, ATL) and western (POR, SLC) regions (East and West, respectively) tended to have higher RICH, EPT, and NMDS and lower RICHTOL scores than in the midwestern (Midwest) region (MGB, DEN, DFW). The slopes of the regressions (Table 4) estimate the rates at which invertebrates responded to urbanization. Slopes were significant ($p < 0.05$) for all response variables in the East (BOS, RAL, BIR, ATL) and West (SLC, POR), but either were not significantly different from 0 or were much lower in the central US (MGB, DEN, DFW). Slopes of the regressions also indicated that invertebrates in metropolitan areas of the East and West tended to change

more rapidly per unit of URB than did metropolitan areas in the Midwest.

The intercepts for the algal regressions (Table 4) indicated that HALO at background sites in the Midwest (MGB, DEN, DFW) were higher (12–32%) than for background sites in the East and West (3–6%) as were EUTR (6–61% in East and West, 63–93% in Midwest) and SILT (4–18% in East and West, 14–70% in Midwest). SENS showed large overlaps between intercepts in the Midwest (25–76%) and the East and West (44–97%). Intercepts for HALO, SILT, and EUTR showed the same regional differences for algae as were observed for invertebrates. Metropolitan areas in the Midwest tended to have background conditions,

TABLE 5. Relative deviance information criteria (DIC) for the multilevel hierarchical regression models used to describe invertebrate and algal responses to urbanization. DIC is expressed relative to the DIC value for model 1 (e.g., $DIC_{\text{model 1}} - DIC_{\text{model 2}}$). Model 5 (continuous antecedent agriculture [AAG]) is shown in italics to emphasize that AAG is more appropriately modeled as a categorical variable (models 6–8). The model with the best fit (excluding model 5) is shown in bold for each response variable. Response variables are as in Table 2. TEMP = mean annual temperature, PRECIP = mean annual precipitation, CLAY = % clay soils in basin.

| Model | Region-level predictors | | Invertebrate responses | | | | Algal responses | | | |
|-------|------------------------------|-------------|------------------------|------------|-------------|-------------|-----------------|-------------|-------------|-------------|
| | Continuous (intercept/slope) | Categorical | NMDS1 | RICH | EPT | RICHTOL | HALO | EUTR | SILT | SENS |
| 1 | None | None | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | PRECIP/PRECIP | None | 26.7 | 0.5 | 5.1 | 18.3 | 24.2 | 15.1 | 16.9 | 13.7 |
| 3 | TEMP/TEMP | None | 9.6 | 4.9 | 0.1 | 18.1 | 11.4 | 9 | 9.5 | 12 |
| 4 | TEMP/PRECIP | None | 15.5 | 7.1 | 3.2 | 30.7 | 18.4 | 16 | 17.8 | 11.7 |
| 5 | <i>AAG</i> | <i>None</i> | <i>37.3</i> | <i>4.9</i> | <i>22</i> | <i>41.3</i> | <i>36</i> | <i>23.6</i> | <i>23.1</i> | <i>17.1</i> |
| 6 | PRECIP/PRECIP | AAG | 28.1 | 0.5 | 11.2 | 25.6 | 29.4 | 14.7 | 18.2 | 11.7 |
| 7 | TEMP/TEMP | AAG | 20 | 4.9 | 9.8 | 33 | 22.2 | 10.9 | 14.8 | 12.6 |
| 8 | TEMP/PRECIP | AAG | 25.1 | 8.3 | 11.3 | 39.3 | 26.3 | 15.6 | 20.2 | 11 |
| 9 | CLAY/CLAY | None | 17.7 | 8.4 | 4 | 3.5 | 17.5 | 16.4 | 22.8 | 31 |

relative to urbanization, that suggested poorer-quality algal assemblages (lower SENS; higher HALO, SILT and EUTR) than in the West or East.

Slopes of the algal responses to URB varied among regions, and none of the response variables showed statistically significant slopes for all metropolitan areas (Table 4). The regional patterns in slopes that were observed for invertebrates were not observed for algae, i.e., the range of slopes in the Midwest overlapped with slopes in the East and West and nonsignificant slopes were associated with metropolitan areas in the East, West, and Midwest. Slopes varied considerably among metropolitan areas, but HALO, EUTR, and SILT diatom metrics all had positive slopes for all metropolitan areas except DFW (HALO, SILT) and MGB and DFW (EUTR) indicating that the percentage of diatoms composing these metrics tended to increase with increasing urbanization. SENS diatoms showed the opposite trend, and the percentage decreased as urbanization increased except in ATL.

Regional differences (Midwest vs East and West) in responses to urbanization (slopes and intercepts) were evident for invertebrates. However, regional patterns in algal responses to urbanization, particularly slopes, were not as evident. The simple linear regressions indicated regional differences in responses to urbanization, but they were not able to provide a direct explanation for these regional patterns.

Multilevel regional models

Hyperparameters (regional slopes and intercepts) derived from the multilevel hierarchical regressions (Appendix 2 for invertebrates, Appendix 3 for algae)

define how the relations between invertebrate responses and URB within each region (slopes and intercepts) change in response to regional predictors. Each of the multilevel models describes regional regressions for slopes and intercepts based on a regional predictor (Models 2–9) or regional averages (Model 1). Models 6, 7, and 8 incorporate AAG as an additional regional predictor in metropolitan areas with high AAG (MGB, DEN, DFW). Consequently, models 6, 7, and 8 have separate regression lines that define the responses of intercepts and slopes for regions with high and low AAG with the differences between the lines determined by the hyperparameters $\delta_{\alpha\text{AG}}$ and $\delta_{\beta\text{AG}}$. In addition to describing the regional responses, the hyperparameters presented in Appendices 2 and 3 can be used to predict a slope or intercept in another metropolitan area based on the regional variables in the model.

The relative fits of the models are described in Table 5, which compares DIC scores relative to model 1 (no regional predictors). The DIC scores indicated that the models with region-level predictors (models 2–9) provided better overall fit than did models without regional predictors (model 1) for both invertebrates and algae. Model 5 (continuous AAG) provided the largest improvement in fit for most invertebrate (NMDS1, EPT, RICHTOL) and algal (HALO, EUTR, SILT) responses. However, the plot of model 5 (Fig. 3) confirmed earlier concerns that the discontinuous distribution of regional AAG violates model assumptions and exaggerates the fit of the model caused by the statistical implications of predicting a line with essentially 2 points. Given these problems, it was decided that AAG should be represented only as a categorical predictor (models

Model 5: EPT richness

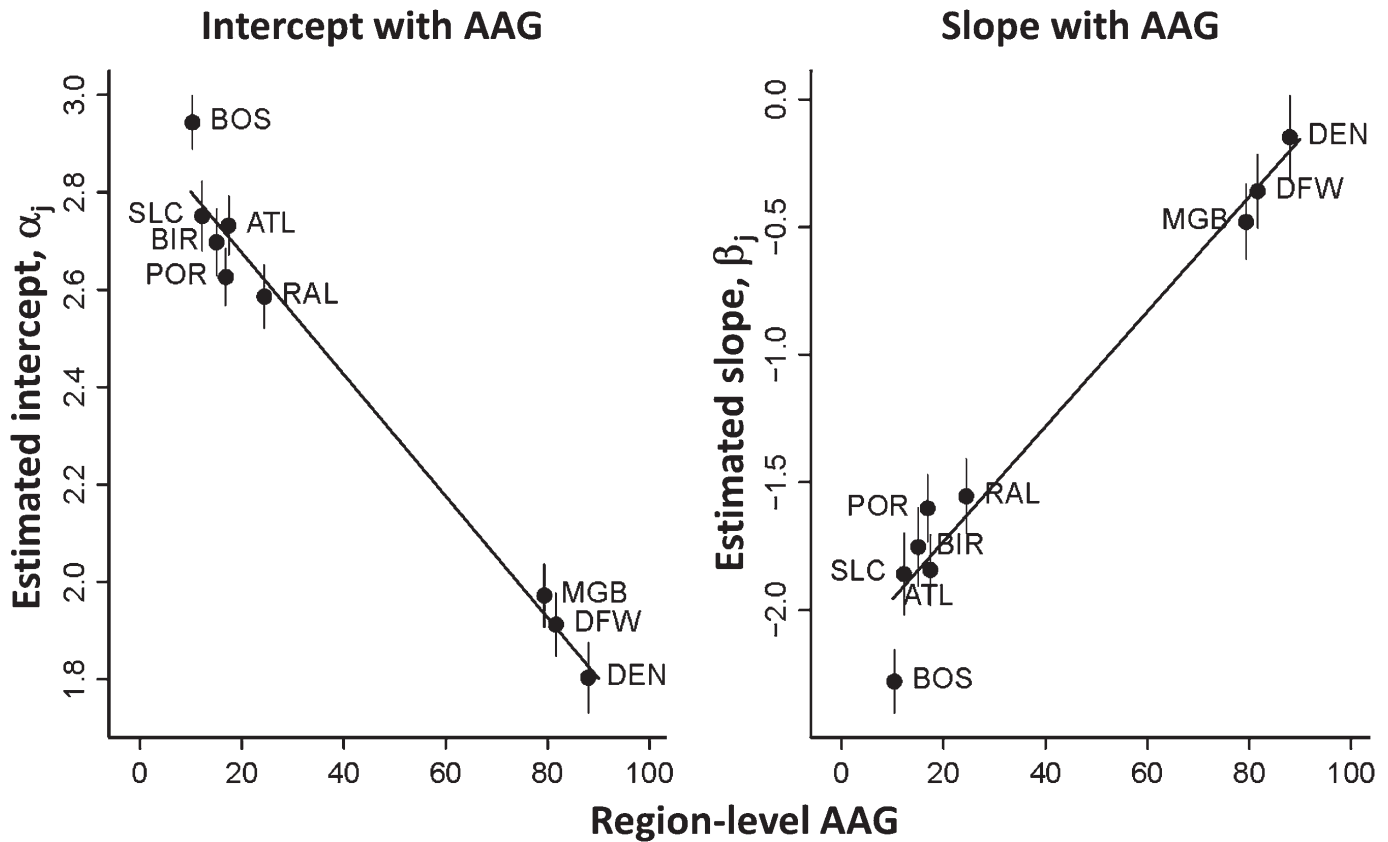


FIG. 3. Slopes (± 1 SD) and intercepts (± 1 SD) of the response of Ephemeroptera, Plecoptera, Trichoptera (EPT) taxon richness to % developed land cover (URB) predicted using continuous region-level antecedent agriculture (AAG) (model 5). Regional mean values for AAG fall into 2 categories (high: $\geq 79\%$, low: $< 25\%$) violating the assumption of the regression model. Abbreviations of metropolitan areas are given in Table 1.

6, 7, and 8) and the results for model 5 should be ignored.

Models that incorporated categorical AAG and TEMP or PRECIP generally were better for predicting invertebrate responses than were models that used only a single predictor. The responses of RICHTOL to TEMP (model 3) and to the combination of TEMP and categorical AAG (model 7) illustrate how incorporating the effects of AAG improved the fit of the model and the ability to discern the effects of TEMP. The models based solely on regional TEMP (Fig. 4A) contained much greater scatter around the regression lines than did the models that incorporated categorical AAG (Fig. 4B). The influence of categorical AAG in model 7 (Fig. 4B) is evidenced by the distance between the regression lines, which is determined by the hyperparameters $\delta_{\alpha_{AG}}$ (intercepts) and $\delta_{\beta_{AG}}$ (slopes). The influence of TEMP after accounting for categorical AAG is observable by examining the slopes of the regression lines ($\gamma_{\alpha 1}$ and $\gamma_{\beta 1}$). In this example, the

effect of AAG on the intercepts (1.65 units higher for high AAG) and slopes (2.05 units higher for low AAG) of the response to URB is large when compared to the response across the range of TEMP (8–20°C) for intercepts (1.56 unit increase) and slopes (0.91 unit decrease). The model hyperparameters provide quantitative evidence of the influence of AAG on the responses of invertebrates that could only be inferred in previous studies (Cuffney et al. 2010).

The DIC scores for the invertebrate models indicate that model 8 (categorical AAG, TEMP predicting intercepts, PRECIP predicting slopes) provides a good fit for all invertebrate responses (i.e., improvement in DIC relative to model 1 is close to the maximum observed for each response variable). This result suggests commonality in the responses of the 4 invertebrate metrics, but some noteworthy differences were found among invertebrate response metrics. The model based on RICH had relatively little improvement in DIC scores compared to the other metrics, a

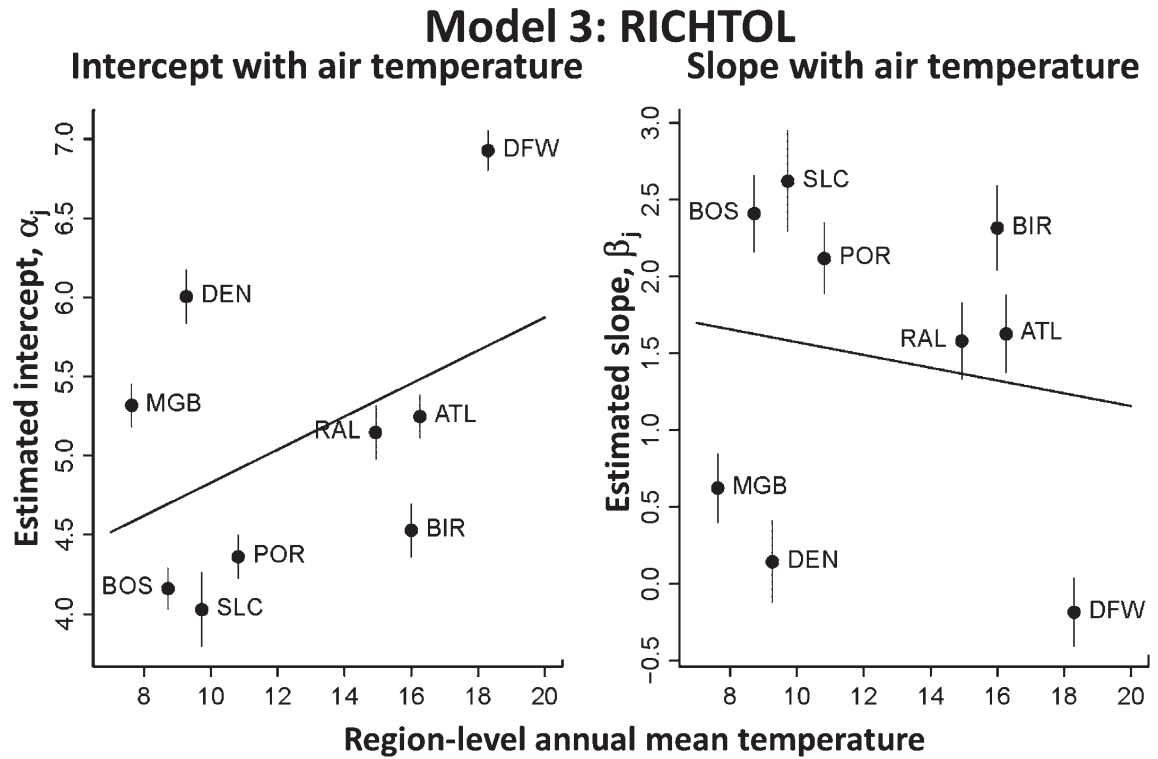
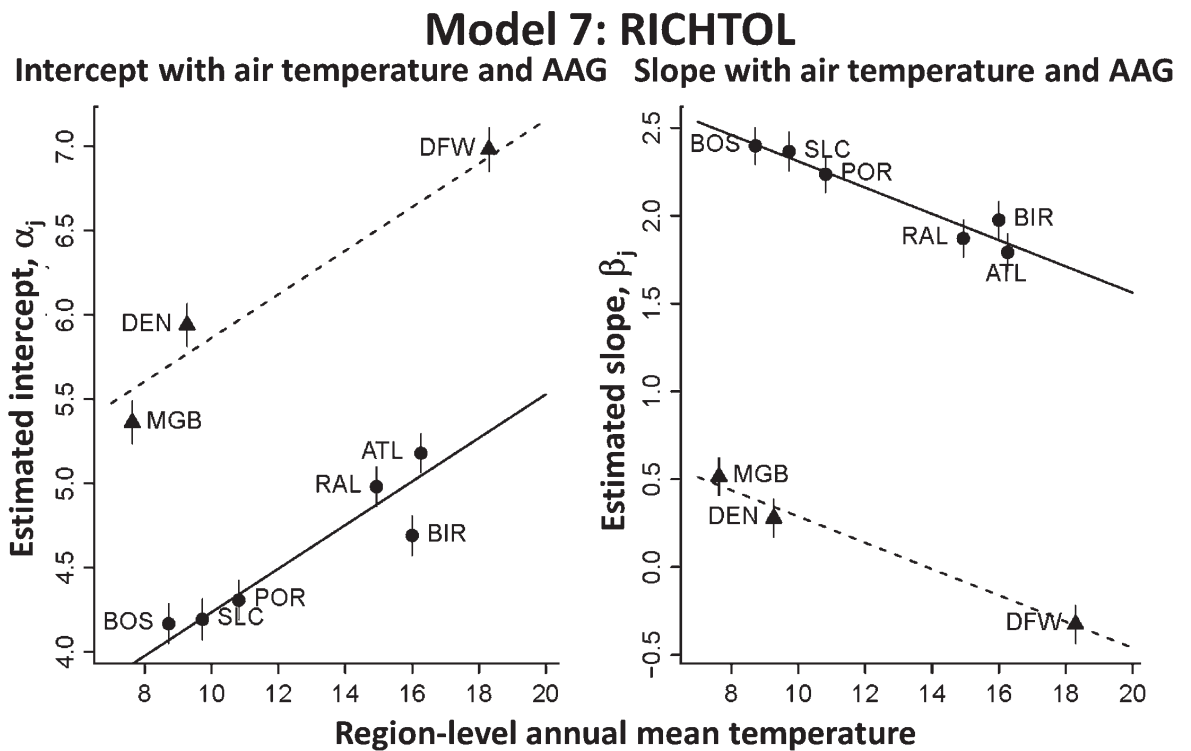
A**B**

FIG. 4. Slopes (± 1 SD) and intercepts (± 1 SD) of the response of richness-weighted invertebrate tolerance (RICHTOL) to % developed land cover (URB) predicted using mean annual air temperature (TEMP) (A) and TEMP combined with categorical antecedent agriculture (AAG) (B). The influence of TEMP is discernible after accounting for AAG.

TABLE 6. Comparison of the influences of categorical antecedent agriculture (AAG), mean annual air temperature (TEMP), and mean annual precipitation (PRECIP) on invertebrate responses based on model 8 hyperparameters. The effects of AAG on estimates of intercepts ($\delta_{\alpha AG}$) and slopes ($\delta_{\beta AG}$) are compared to the change in intercepts ($\gamma_{\alpha 1}$) and slopes ($\gamma_{\beta 1}$) over the gradient of temperature (8–20°C) and precipitation (40–160 cm). The signs of the AAG values indicate the difference between the high and low AAG regression lines as defined by the difference in intercepts (high AAG – low AAG) of the regional models (Fig. 4B). Negative and positive values for TEMP and PRECIP indicate that the invertebrate responses are decreasing (–) or increasing (+) across the TEMP or PRECIP gradient. Response variables are as in Table 2.

| Regional model | Regional predictor (model hyperparameter) | Invertebrate response | | | |
|----------------|---|-----------------------|---------|---------|---------|
| | | NMDS1 | RICH | EPT | RICHTOL |
| Intercepts | AAG ($\delta_{\alpha AG}$) | –0.7125 | –0.2008 | –0.8085 | 1.6334 |
| | TEMP ($\gamma_{\alpha 1}$) | 0.3204 | 0.1608 | 0.0312 | 1.2708 |
| Slopes | AAG ($\delta_{\beta AG}$) | 1.3842 | 0.2924 | 1.3347 | –2.2361 |
| | PRECIP ($\gamma_{\beta 1}$) | –0.660 | –0.372 | –0.396 | –0.792 |

result suggesting that it was not a good indicator of response to URB. NMDS1, which incorporated information on the entire invertebrate assemblage, showed nearly as good a fit with PRECIP alone as it did in combination with categorical AAG.

Model 8 indicated that conditions prior to urbanization (background conditions) were better (higher NMDS1, RICH, and EPT, lower RICHTOL) in regions with low AAG compared to regions with high AAG at the same TEMP as indicated by $\delta_{\alpha AG}$ (in Table 6, negative for NMDS1, RICH, and EPT and positive for RICHTOL). This hyperparameter ($\delta_{\alpha AG}$) represents the difference in intercepts (Appendix 2, Table A1) between high and low AAG (high AAG–low AAG). All background conditions increased as TEMP increased (positive $\gamma_{\alpha 1}$ values for intercepts). The changes in background conditions arising from AAG ($\delta_{\alpha AG}$) were greater than the changes in the value of the intercepts ($\gamma_{\alpha 1}$) over the TEMP gradient (8–20°C).

The rates (slopes) of response to URB in model 8 were strongly affected by AAG. The difference in slope between metropolitan areas with high and low AAG ($\delta_{\beta AG}$) exceeded the change in slope ($\gamma_{\beta 1}$) over the PRECIP gradient (40–160 cm) for NMDS1, EPT, and RICHTOL, but not RICH (Table 6). The influence of regional factors (PRECIP) on the rates at which slopes change must be interpreted carefully by considering how the regional ($\gamma_{\beta 1}$) slopes affect the basin-level (β_i) slopes and how these changes affect the values of the metrics and the assessment of condition (e.g., better conditions associated with lower RICHTOL and higher NMDS1, RICH, and EPT). The regional slopes ($\gamma_{\beta 1}$) are negative for the 4 metrics, whereas the basin-scale slopes are negative for NMDS1, RICH, and EPT and positive for RICHTOL. Consequently, negative regional slopes indicate that the rate of decline increases (i.e., becomes more negative) for NMDS1, RICH, and EPT as PRECIP increases and decreases (i.e., become less

positive) for RICHTOL. Collectively, these changes indicate that assemblage conditions represented in all 4 metrics degrade more rapidly per unit of URB as PRECIP increases.

The combination of categorical AAG and regional climate (TEMP and PRECIP) provided a means to explain the poor correspondence between invertebrate responses and URB that were observed in the unpooled regressions for DEN, DFW, and MGB (Table 4). High AAG degrades the condition of the background assemblages and reduces the rates of response across the URB gradient relative to other regions. Consequently, high AAG can mask the effects of urbanization and make detecting assemblage changes more difficult.

In contrast to the invertebrate models, the DIC values for the algal multilevel models did not suggest a common model for the responses of all 4 metrics. Three (EUTR, SENS, SILT) of the 4 algal metrics were best fit by model 9 (CLAY), and only HALO showed a best fit with a model that included AAG (model 6). HALO responded in a manner similar to the NMDS1 invertebrate metric in that models that incorporated PRECIP alone (Model 2) or in combination with AAG (models 6 and 8) showed good improvement in fit compared to model 1 (no regional predictors). Similar to RICHTOL, the response of HALO contained considerable scatter when predicted using PRECIP alone (Fig. 5A) and considerable improvement in fit when combined with categorical AAG (Fig. 5B). Responses of other algal metrics were best fit by CLAY in the region (Fig. 6). Some similarities existed among invertebrate and algal responses (e.g., RICHTOL and HALO), but the algal assemblages were less responsive to AAG than were the invertebrate assemblages and were more responsive to CLAY.

Background conditions (intercepts) of HALO (model 6) differed greatly between regions with high AAG

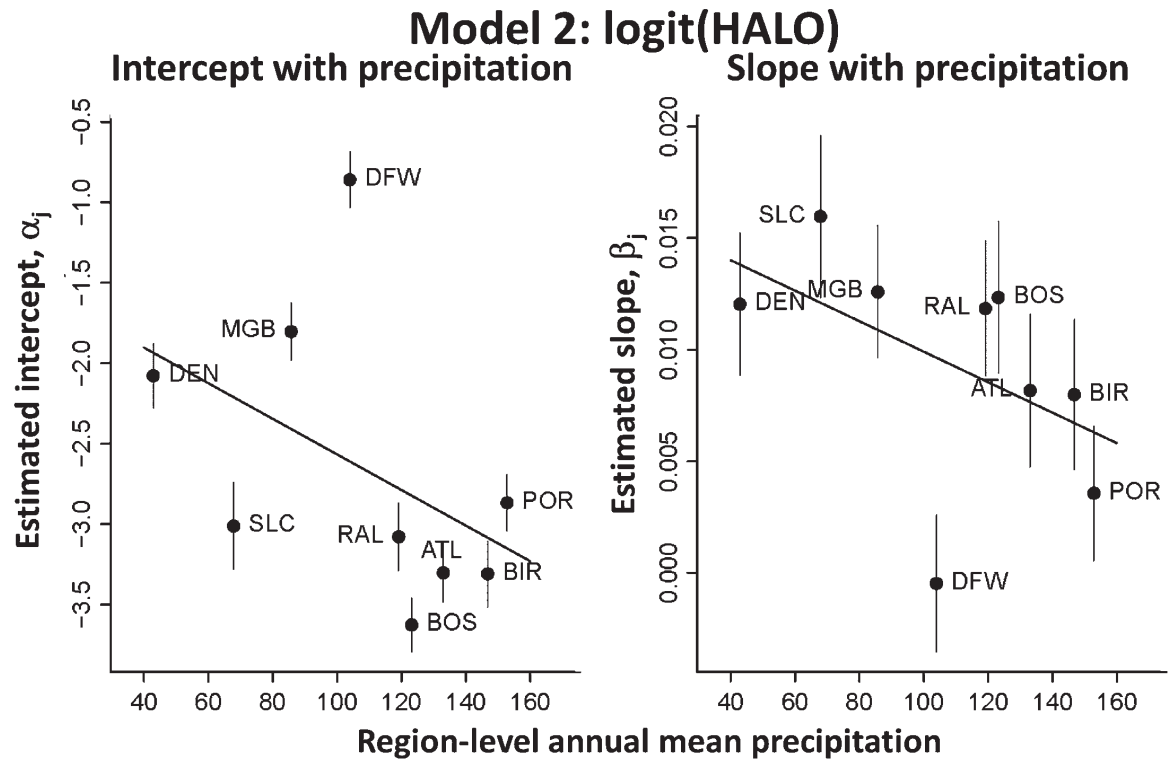
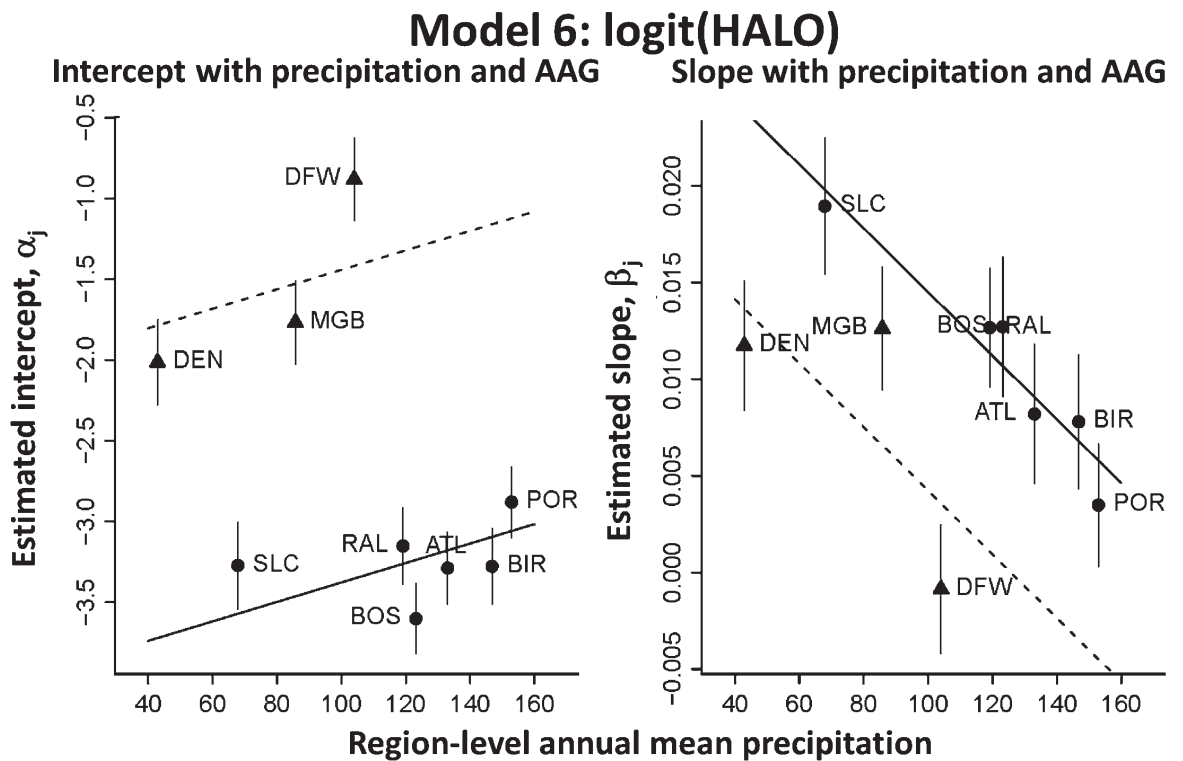
A**B**

FIG. 5. Slopes (± 1 SD) and intercepts (± 1 SD) of the response of % salt-tolerant diatoms (HALO) to % developed land cover (URB) predicted using mean annual precipitation (PRECIP) (A) and mean annual air precipitation (PRECIP) combined with categorical antecedent agriculture (AAG) (B). The influence of PRECIP is discernible after accounting for AAG.

Model 9: logit(SENS)

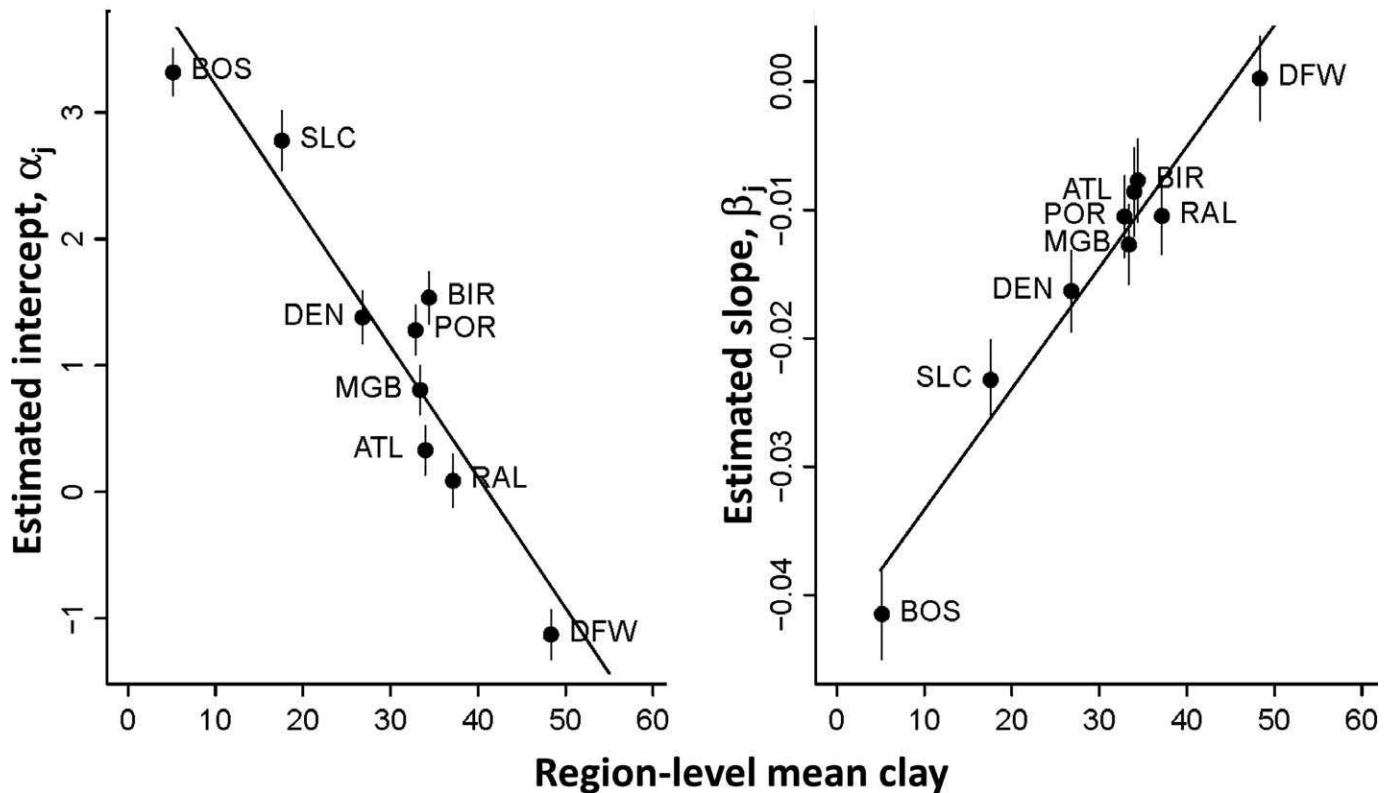


FIG. 6. Slopes (± 1 SD) and intercepts (± 1 SD) of SENS (% sensitive diatoms) to % developed land cover (URB) predicted using region-level % clay soil (CLAY) (model 9).

and low AAG (Fig. 5B). Regions with high AAG had a higher HALO in the absence of urbanization than did regions with low AAG, and HALO in background sites tended to increase with increasing PRECIP. The slope of the response of HALO to URB tended to decrease as PRECIP increased and to be higher in low AAG sites than in high AAG sites for a given level of PRECIP. These results indicate that, in the absence of urbanization, high levels of AAG produce conditions that favor halophilic algae that probably are a result of salts washing into streams from irrigation, drainage, and use of fertilizer.

The responses of EUTR, SILT, and SENS algae to URB (intercepts and slopes) were best represented by CLAY in each region (model 9). The proportion of SENS diatoms at background sites (intercepts) decreased as the regional CLAY increased, and the rate of response increased as CLAY increased in the region (Fig. 6). The opposite response was observed for EUTR and SILT diatoms. The proportion of each at background sites tended to increase as the regional

CLAY increased (e.g., EUTR at urban background sites = $-2.3119 + 0.0741[\text{CLAY}]$; Appendix 3, model 9). The rates at which EUTR and SILT responded to URB (slopes) both decreased as CLAY increased in the region (e.g., EUTR rate of response to urbanization = $3.8218 - 0.0862[\text{CLAY}]$; Appendix 3, model 9). These trends were consistent with a general decline in conditions associated with an increase in fine substrates in regions with high CLAY.

Discussion

Multilevel hierarchical regression models directly incorporated environmental variables that act at basin (URB) and regional (AAG, TEMP, PRECIP, and CLAY) scales into relatively simple models that explained the variation in invertebrate and algal responses across the USA and that could be used to investigate causal factors and to make predictions. These models revealed that invertebrate responses to urbanization tended to be more sensitive than algal responses to

regional climatic (TEMP and PRECIP) and land cover (AAG) disturbances. Algae tended to be influenced more by regional soil characteristics (CLAY). Models that incorporated only basin-scale predictors were not able to detect these large-scale patterns.

Regional temperature and precipitation acted together to influence important ecological characteristics, such as natural vegetation (forest, shrub-, and grasslands), riparian conditions, hydrology, material cycling, and energy flow. Collectively, these factors determined the lotic communities (background conditions) that are affected by agriculture and urbanization. Regions of the country where temperature and precipitation favored the development of forest lands (ATL, BIR, BOS, MGB, RAL, and POR) tended to have assemblages with higher richness and more intolerant forms than areas where climate favored the development of grass- (DEN, DFW) or shrub-lands (SLC). Consequently, regions where forests were being converted to urban land (BOS, RAL, ATL, BIR, POR) tended to have background invertebrate assemblages with higher richness and more intolerant forms than did areas of the country where agricultural or grazing lands were being converted to urban lands (MGB, DEN, DFW). Biological conditions also tended to degrade at faster rates in forested areas than in areas with high AAG, in part, because higher initial background values led to greater changes over the gradient. The hyperparameters derived from multilevel hierarchical regression were able to quantify the effects of natural (TEMP and PRECIP) and anthropogenic regional factors (AAG) that directly affected the background conditions (assemblage structure) upon which urbanization acts. These models were particularly useful for documenting the effects of high levels of AAG on the invertebrate assemblages prior to the onset of urbanization. Degradation by AAG reduced the ability to detect effects arising from urbanization.

Models that incorporated only basin-scale variables underestimated or failed to detect the effects of urbanization when these effects were masked by factors that varied regionally (e.g., AAG). This masking led to erroneous conclusions regarding the condition of streams, the effectiveness of mitigation efforts, or the causes of degradation. Our use of multilevel hierarchical regression models to incorporate predictors measured at multiple, hierarchically related spatial scales and encompassing multiple metropolitan areas with contrasting environmental settings led to a much greater understanding of the regional factors driving basin-scale responses to urbanization and helped us avoid conclusions that were obviously erroneous when viewed at larger spatial scales. For example, the extent of agriculture in

the Midwest made it more difficult to establish background conditions against which change could be assessed. Multilevel hierarchical regression models provided a mechanism for addressing this problem by incorporating regional predictors that could be used to establish expected conditions and predict changes arising from climatic conditions, agriculture, and urbanization. Hierarchical regression provides a mechanism for incorporating multiscale predictors in predictive models that can complement more complex modeling procedures, such as structural equation models (Grace 2006) and multivariate models (e.g., River InVertebrate Prediction and Classification System [RIVPACS], Clarke et al. 2003).

The distribution of AAG observed in our study was interesting because it was essentially categorical and either represented relatively modest (<25% of basin area) or very high ($\geq 79\%$) levels of agricultural development. Whether this pattern is consistent for all metropolitan areas across the continental USA has yet to be established. However, AAG in other metropolitan areas studied in the NAWQA Program (Anchorage, Alaska; Dayton, Ohio; and Chicago, Illinois) followed the same pattern. The effects of AAG on the response to URB were large, often exceeding the effect of regional climate variables (TEMP and PRECIP). AAG, which encompassed both row-crop and grazing lands, affected biota through changes in natural vegetation, sedimentation associated with land disturbance, nutrient enrichment, pesticide contamination, and flow alteration (drains, irrigation). These effects also are commonly associated with urbanization (Paul and Meyer 2001, Walsh et al. 2005) making differentiation of AAG and URB effects more difficult.

Climate factors (TEMP and PRECIP) also were important in determining differences in the responses of invertebrate assemblages. However, the effects of these variables were most apparent when they were modeled in combination with AAG (models 6–8). This result implies that large-scale studies of the effects of climate change may fail to detect effects if they do not also address changes in regional land use. Metropolitan areas with lower TEMP or higher PRECIP tended to have better background conditions (intercepts were higher for RICH, EPT, NMDS1; lower for RICHTOL) than did metropolitan areas with higher TEMP or lower PRECIP, particularly when compared within categories of AAG (high, low). Better background conditions were associated with lower TEMP and higher PRECIP because these conditions favor the development of forests that present less harsh environmental conditions (e.g., generally more moderate temperature regimes, more abundant and

predictable precipitation, and greater variety of C sources) than grass- or shrub-lands. The ability of multilevel hierarchical regression to separate the effects of regional climate (TEMP and PRECIP) factors from anthropogenic factors (AAG) provided an important mechanism for understanding and predicting background conditions across the country. Regional changes in background conditions reflected both natural biogeographic differences (e.g., climatic differences) and anthropogenic differences.

The rates at which assemblages changed in response to increasing URB also were lower in areas with higher TEMP or lower PRECIP, in part, because the differences in background conditions (intercepts) dictated the maximum amount of change that could occur over the gradient. The results from these models suggested that in regions where climate change will result in increased temperatures or decreased precipitation, background conditions are likely to become degraded relative to the original conditions as temperature (Sweeney and Vannote 1978) and flow (Poff et al. 2010) regimes become suboptimal for the species that compose the assemblage. If these types of changes in the background assemblages are not taken into account, then the effects of urbanization on invertebrate assemblages may be underestimated or missed entirely. This may be particularly problematic in areas where reference sites no longer exist and assessments depend on comparisons with historic conditions.

HALO diatoms displayed responses similar to those of invertebrates, i.e., the responses (intercepts and slopes) were affected by AAG and PRECIP. Most of the diatom response indicators (EUTR, SILT, SENS) were better explained by CLAY than by AAG. The lower the CLAY in the region, the better the condition of the background diatom assemblage (i.e., higher SENS taxa, lower EUTR and SILT forms) and the higher the rate of response to URB. The presence of clay soils in the region probably was indicative of the sedimentation potential in the streams. Areas with high-clay soils were likely to have finer and less stable substrates that may have favored disturbance-tolerant diatoms. The slopes of diatom responses were lower in areas with higher CLAY, in part, because background conditions were degraded relative to those at sites with less CLAY. Thus, the amount of change that could occur along the gradient was less than in areas with low CLAY (e.g., BOS) that received most of their sediments from urbanization.

The multilevel hierarchical regressions confirmed our earlier work that suggested that AAG plays a role in determining regional background conditions and rates of response for invertebrates (Brown et al. 2009,

Cuffney et al. 2010). However, our previous work could link invertebrate responses only indirectly to AAG by comparing the intercepts and slopes of the unpooled regressions for regions with high and low AAG. This indirect method did not provide a means for quantifying the relation between classes of AAG and URB. That is, no means existed to estimate the expected response (slopes and intercepts) for a region outside the 9 that were studied. Multilevel hierarchical regression also helped reveal that climate (TEMP, PRECIP) was important in determining invertebrate response and that CLAY was important in determining algal responses, results that were not detected in our previous work.

The primary utility of the multilevel models is incorporation of larger-scale (regional) predictors to explain the differences in slopes and intercepts of the models that are based on smaller-scale (basin) predictors. Incorporating regional predictors does not improve the fit of the basin-scale models as might be expected when adding additional basin-scale predictors in a multiple regression model. For example, the slopes of the basin-scale models for EPT in MGB, DEN, and DFW were not significantly different from 0 and the incorporation of regional variables did not change this result. The incorporation of regional predictors (e.g., AAG, TEMP, PRECIP) in the multilevel models did provide the ability to explain what circumstances led to low or high slopes or intercepts. This information was extremely valuable for interpreting the basin-scale regressions and for understanding the ecological processes governing the responses at both the basin and regional scales. A nonsignificant regression can arise from a number of factors including problems with implementing the study design and obtaining and processing samples. The ability to explain the low values for slopes in MGB, DEN, and DFW on the basis of a common set of regional predictors that also explained high slopes helped to ascertain that the observed differences were ecologically meaningful and not the result of experimental error.

Models serve 2 important functions. They can provide insight into the environmental factors that are important in driving responses (understanding), and they can provide the ability to extrapolate to regions outside of the study (prediction). The primary contribution of the models described in our paper is to the understanding of the regional factors that modify responses to urbanization across the continent. The use of these models for prediction purposes has yet to be established, but given the relatively small sample size (9 metropolitan areas, 6 representing high AAG and 3 representing low AAG), data

from more regions probably will be needed before these models can be used to predict responses across the country reliably.

The use of multilevel hierarchical regression models helped to demonstrate clearly the influence of regional-scale environmental factors, both natural (climate and soils) and anthropogenic (agricultural), in determining basin-scale responses to urbanization. This influence provided strong evidence against a one-size-fits-all approach to monitoring and mitigating the effects of urbanization across the USA and provided rationale for developing regionally tailored solutions to these problems. Continuing the development and refinement of regional models is critical to understanding how urbanization affects invertebrate and algal assemblages at local, regional, and continental scales and for the development of effective programs for monitoring, regulating, and mitigating the effects of urbanization (Grimm et al. 2008).

Acknowledgements

We gratefully acknowledge the many colleagues who have worked to collect samples and analyze the data that support this paper. Mary Freeman, Daren Carlisle, and 2 anonymous referees provided many helpful suggestions that greatly improved this paper. We also thank the many property owners and municipalities for granting us permission to access and sample these streams. Without their generous support, none of this work would have been possible. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

Literature Cited

- BARBOUR, M. T., J. GERRITSEN, B. D. SNYDER, AND J. B. STRIBLING. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. EPA 841-B-99-002. Office of Water, US Environmental Protection Agency, Washington, DC.
- BATES, D., AND M. MAECHLER. 2010. lme4: linear mixed-effects models using Eigen and Eigen. R package version 0.999375-33. R Foundation for Statistical Computing, Vienna, Austria. (Available from: <http://CRAN.R-project.org/package=lme4>)
- BROWN, L. R., T. F. CUFFNEY, J. F. COLES, A. H. BELL, J. T. MAY, F. FITZPATRICK, G. MCMAHON, AND J. STEUER. 2009. Urban streams across the United States: lessons learned from studies in nine metropolitan areas. *Journal of the North American Benthological Society* 28:1051–1069.
- CLARKE, K. R., AND R. N. GORLEY. 2006. PRIMER v6: user manual/tutorial. PRIMER-E Ltd., Plymouth, UK.
- CLARKE, K. R., J. F. WRIGHT, AND M. T. FURSE. 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. *Ecological Modelling* 160:219–233.
- COLES, J. F., A. BELL, B. SCUDDER, AND K. CARPENTER. 2009. The effects of urbanization and other environmental gradients on algal assemblages in nine metropolitan areas across the United States. U.S. Geological Survey Scientific Investigations Report 2009-5022. US Geological Survey, Reston, Virginia.
- COLES, J. F., T. F. CUFFNEY, AND G. MCMAHON. 2004. The effects of urbanization on the biological, physical, and chemical characteristics of coastal New England streams. U.S. Geological Survey Professional Paper 1695. US Geological Survey, Reston, Virginia.
- CUFFNEY, T. F., M. D. BILGER, AND A. M. HAIGLER. 2007. Ambiguous taxa: effects on the characterization and interpretation of invertebrate assemblages. *Journal of the North American Benthological Society* 26:286–307.
- CUFFNEY, T. F., AND R. A. BRIGHTBILL. 2011. User's manual for the National Water-Quality Assessment Program Invertebrate Data Analysis System (IDAS) software, version 5. U.S. Geological Survey Techniques and Methods 7-C4. US Geological Survey, Reston, Virginia. (Available from: <http://pubs.usgs.gov/tm/7c4/>)
- CUFFNEY, T. F., R. A. BRIGHTBILL, J. T. MAY, AND I. R. WAITE. 2010. Responses of benthic macroinvertebrates to environmental changes associated with urbanization in nine metropolitan areas of the conterminous United States. *Ecological Applications* 20:1384–1401.
- CUFFNEY, T. F., AND J. F. FALCONE. 2008. Derivation of nationally consistent indices representing urban intensity within and across nine metropolitan areas of the conterminous United States. U.S. Geological Survey Scientific Investigations Report 2008-5095. US Geological Survey, Reston, Virginia.
- CUFFNEY, T. F., H. ZAPPALÀ, E. M. P. GIDDINGS, AND J. F. COLES. 2005. Effects of urbanization on benthic macroinvertebrate assemblages in contrasting environmental settings: Boston, Massachusetts; Birmingham, Alabama; and Salt Lake City, Utah. Pages 361–407 in L. R. Brown, R. M. Hughes, R. Gray, and M. R. Meador (editors). *Effects of urbanization on stream ecosystems*. Symposium 47. American Fisheries Society, Bethesda, Maryland.
- DAYMET. 2005. Numerical Terradynamic Simulation Group. University of Montana, Missoula, Montana. (Available from <http://www.daymet.org>)
- FALCONE, J. F., J. STEWART, S. SOBIESZCZYK, J. DUPREE, G. MCMAHON, AND G. BUELL. 2007. A comparison of natural and urban characteristics and the development of urban intensity indices across six geographic settings. U.S. Geological Survey Scientific Investigations Report 2007-5123. US Geological Survey, Reston, Virginia.
- GELMAN, A., AND J. HILL. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York.
- GIDDINGS, E. M. P., A. H. BELL, K. M. BEAULIEU, T. F. CUFFNEY, J. F. COLES, L. R. BROWN, F. A. FITZPATRICK, J. FALCONE, L. A. SPRAGUE, W. L. BRYANT, M. C. PEPLER, C. STEPHENS, AND G. MCMAHON. 2009. Selected physical, chemical,

- and biological data used to study urbanizing streams in nine metropolitan areas of the United States, 1999–2004. U.S. Geological Survey Data Series 423. US Geological Survey, Reston, Virginia.
- GRACE, J. B. 2006. Structural equation modeling and natural systems. Cambridge University Press, New York.
- GRIMM, N. B., D. FOSTER, P. GROFFMAN, J. M. GROVE, C. S. HOPKINSON, K. J. NADELHOFFER, D. E. PATAKI, AND D. P. C. PETERS. 2008. The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. *Frontiers in Ecology and the Environment* 6:264–272.
- JACOBSON, R. B., S. R. FEMMER, AND R. A. MCKENNEY. 2001. Land-use changes and the physical habitat of streams—a review with emphasis on studies within the U.S. Geological Survey Federal-State cooperative program. U.S. Geological Survey Circular 1175. US Geological Survey, Reston, Virginia.
- KASHUBA, R., C. YOONKYUNG, I. ALAMEDDINE, B. LEE, AND T. F. CUFFNEY. 2010. Multilevel hierarchical modeling of benthic macroinvertebrate responses to urbanization in nine metropolitan regions across the conterminous United States. U.S. Geological Survey Scientific Investigations Report 2009-5243. US Geological Survey, Reston, Virginia.
- KRATZER, E. B., J. K. JACKSON, D. B. ARSCOTT, A. K. AUFDENKAMPE, C. L. DOW, L. A. KAPLAN, J. D. NEWBOLD, AND B. W. SWEENEY. 2006. Macroinvertebrate distribution in relation to land use and water chemistry in New York City drinking-water-supply watersheds. *Journal of the North American Benthological Society* 25:954–976.
- LANGE-BERTALOT, H. 1979. Pollution tolerance of diatoms as a criterion for water quality estimation. *Nova Hedwigia* 64:285–304.
- MAHLER, B. J., P. C. VAN METRE, T. J. BASHARA, G. T. WILSON, AND D. A. JOHNS. 2005. Parking lot sealcoat—an unrecognized source of urban polycyclic aromatic hydrocarbons. *Environmental Science and Technology* 39:5560–5566.
- MORSE, C. C., A. D. HURY, AND C. CRONAN. 2003. Impervious surface area as a predictor of the effects of urbanization on stream insect communities in Maine, U.S.A. *Environmental Monitoring and Assessment* 89:95–127.
- NCDENR (NORTH CAROLINA DEPARTMENT OF ENVIRONMENT AND NATURAL RESOURCES). 2006. Standard operating procedures for benthic macroinvertebrates. Biological Assessment Unit, Division of Water Quality, North Carolina Department of Environment and Natural Resources, Raleigh, North Carolina. (Available from: <http://www.esb.enr.state.nc.us/BAUwww/benthossop.pdf>)
- OMERNIK, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77:118–125.
- PAUL, M. J., AND J. L. MEYER. 2001. Streams in the urban landscape. *Annual Review of Ecology and Systematics* 32:333–365.
- PETERS, D. P. L., P. M. GROFFMAN, K. J. NADELHOFFER, N. B. GRIMM, S. L. COLLINS, W. K. MICHENER, AND M. A. HUSTON. 2008. Living in an increasingly connected world: a framework for continental-scale environmental science. *Frontiers in Ecology and the Environment* 6:229–237.
- PICKETT, S. T. A., M. L. CADENASSO, J. M. GROVE, P. M. GROFFMAN, L. E. BAND, C. G. BOONE, W. R. BURCH, C. S. B. GRIMMOND, J. HORN, J. C. JENKINS, N. L. LAW, C. H. NILON, R. B. POUYAT, K. SZLAVECZ, P. S. WARREN, AND M. A. WILSON. 2008. Beyond urban legends: an emerging framework of urban ecology, as illustrated by the Baltimore Ecosystem Study. *BioScience* 58:139–150.
- POFF, N. L., B. D. RICHTER, A. H. ARTHINGTON, S. E. BUNN, R. J. NAIMAN, E. KENDY, M. ACERMAN, C. APSE, B. P. BLEDSOE, M. C. FREEMAN, J. HENRIKSEN, R. B. JACOBSON, J. G. KENNEN, D. M. MERRITT, J. H. O'KEEFE, J. D. OLDEN, K. ROGERS, R. E. THARME, AND A. WARNER. 2010. The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. *Freshwater Biology* 55:147–170.
- PORTER, S. D. 2008. Algal attributes—an autecological classification of algal taxa collected by the National Water-Quality Assessment Program. U.S. Geological Survey Data Series 329. US Geological Survey, Reston, Virginia. (Available from: <http://pubs.usgs.gov/ds/ds329/>)
- POTAPOVA, M., J. F. COLES, E. M. P. GIDDINGS, AND H. ZAPPIA. 2005. A comparison of the influences of urbanization in contrasting environmental settings on stream benthic algal assemblages. Pages 333–359 in L. R. Brown, R. M. Hughes, R. Gray, and M. R. Meador (editors). *Effects of urbanization on stream ecosystems*. Symposium 47. American Fisheries Society, Bethesda, Maryland.
- QIAN, S. S., T. F. CUFFNEY, I. ALAMEDDINE, G. MCMAHON, AND K. H. RECKHOW. 2010. On the application of multilevel modeling in environmental and ecological studies. *Ecology* 91:355–361.
- SPIEGELHALTER, D. J., N. G. BEST, B. P. CARLIN, AND A. VAN DER LINDE. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 64:583–616.
- SPONSELLER, R. A., E. F. BENFIELD, AND H. M. VALETT. 2001. Relationships between land use, spatial scale and stream macroinvertebrate communities. *Freshwater Biology* 46:1409–1424.
- SWEENEY, B. W., AND R. L. VANNOTE. 1978. Size variation and the distribution of hemimetabolous aquatic insects: two thermal equilibrium hypotheses. *Science* 200:444–446.
- TATE, C. M., T. F. CUFFNEY, G. MCMAHON, E. M. P. GIDDINGS, J. F. COLES, AND H. ZAPPIA. 2005. Use of an urban intensity index to assess urban effects on streams in three contrasting environmental settings. Pages 291–315 in L. R. Brown, R. M. Hughes, R. Gray, and M. R. Meador (editors). *Effects of urbanization on stream ecosystems*. Symposium 47. American Fisheries Society, Bethesda, Maryland.
- USGS (US GEOLOGICAL SURVEY). 2005. National Land Cover Database 2001 (NLCD 2001). US Geological Survey, Reston, Virginia. (Available from: http://www.mrlc.gov/mrlc2k_nlcd.asp)
- VAN DAM, H., A. MERTENS, AND J. SINKELDAM. 1994. A coded checklist and ecological indicator values of freshwater

diatoms from the Netherlands. *Netherlands Journal of Aquatic Ecology* 28:117–133.

VØLSTAD, J. H., N. E. ROTH, G. MERCURIO, M. T. SOUTHERLAND, AND D. E. STREBEL. 2003. Using environmental stressor information to predict the ecological status of Maryland non-tidal streams as measured by biological indicators. *Environmental Monitoring and Assessment* 84:219–242.

WALSH, C. J., A. H. ROY, J. W. FEMINELLA, P. D. COTTINGHAM, AND P. M. GROFFMAN. 2005. The urban stream syndrome: current knowledge and search for a cure. *Journal of the North American Benthological Society* 24:706–723.

WENGER, S. J., A. H. ROY, C. R. JACKSON, E. S. BERNHARDT, T. L. CARTER, S. FILOSO, C. A. GIBSON, W. C. HESSION, S. S. KAUSHAL, E. MARTÍ, J. L. MEYER, M. A. PALMER, M. J. PAUL, A. H. PURCELL, A. RAMÍREZ, A. D. ROSEMOND, K. A. SCHOFIELD, E. B. SUDDUTH, AND C. J. WALSH. 2009. Twenty-six key research questions in urban stream ecology: an assessment of the state of the Science. *Journal of the North American Benthological Society* 28:1080–1098.

Received: 22 October 2010

Accepted: 6 May 2011

APPENDIX 1. Model descriptions

Partial pooling

Partial pooling represents a compromise between 2 extremes: complete pooling (no categorical predictor, all regions combined) and no pooling (separate models for each region). The pooled and unpooled estimates are weighted by group (region) sample size and variation within and between groups to obtain the estimate of the partially pooled, multilevel estimate of the mean (Eq. 1) for the region (Gelman and Hill 2007):

$$\hat{\alpha}_j^{multilevel} \approx \frac{\frac{n_j}{\sigma_y^2} \bar{y}_j + \frac{1}{\sigma_\alpha^2} \bar{y}_{all}}{\frac{n_j}{\sigma_y^2} + \frac{1}{\sigma_\alpha^2}} \quad [1]$$

Where $\hat{\alpha}_j^{multilevel}$ is the partially pooled, multilevel parameter mean estimate for region j , \bar{y}_j is the unpooled estimate of the mean from region j , \bar{y}_{all} is the completely pooled estimate of the mean for all regions, n_j is sample size of region j , σ_y^2 is the within-region variance, and σ_α^2 is the between-region variance.

The multilevel estimate $\hat{\alpha}_j^{multilevel}$ is close to a complete pooling regression model when the region-level variation (σ_y^2) and sample (n_j) size are small, whereas it is close to a no-pooling regression model when between-region variance (σ_α^2) and sample size are large. Unlike classical regression models, multilevel hierarchical regression models can reduce the

unexplained variance at each level (basin and region) by using both individual- (basin) and group-level (region) predictors to estimate model parameters. This approach is particularly useful when sample sizes in some groups are very low. The 9 metropolitan areas had similar sample sizes (28–30) and large between-region variance, so the regression models based on partial pooling were similar to the regression models based on no pooling (Fig. A1).

Description of the templates used in the multilevel hierarchical regressions (see Table A1 for description of model parameters)

Template 1.—No region-level predictors

$$y_{ij} \sim N(\alpha_j + \beta_j x_{ij}, \sigma_y^2) \quad [2]$$

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho \sigma_\alpha \sigma_\beta \\ \rho \sigma_\alpha \sigma_\beta & \sigma_\beta^2 \end{pmatrix} \right) \quad [3]$$

Template 2.—One continuous region-level predictor:

$$y_{ij} \sim N(\alpha_j + \beta_j x_{ij}, \sigma_y^2) \quad [4]$$

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \sim N \left(\begin{pmatrix} \gamma_{\alpha 0} + \gamma_{\alpha 1} P_{\alpha j} \\ \gamma_{\beta 0} + \gamma_{\beta 1} P_{\beta j} \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho \sigma_\alpha \sigma_\beta \\ \rho \sigma_\alpha \sigma_\beta & \sigma_\beta^2 \end{pmatrix} \right) \quad [5]$$

Template 3.—Two region-level predictors, one continuous and one categorical

$$y_{ij} \sim N(\alpha_j + \beta_j x_{ij}, \sigma_y^2) \quad [6]$$

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \sim \begin{pmatrix} \gamma_{\alpha 0} + \gamma_{\alpha 1} P_{\alpha j} \\ \gamma_{\beta 0} + \gamma_{\beta 1} P_{\beta j} \end{pmatrix} + \begin{pmatrix} \delta_{\alpha j} \\ \delta_{\beta j} \end{pmatrix} + \begin{pmatrix} \delta_{\alpha k} \\ \delta_{\beta k} \end{pmatrix} \quad [7]$$

$$\begin{pmatrix} \delta_{\alpha j} \\ \delta_{\beta j} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha j}^2 & \rho_j \sigma_{\alpha j} \sigma_{\beta j} \\ \rho_j \sigma_{\alpha j} \sigma_{\beta j} & \sigma_{\beta j}^2 \end{pmatrix} \right) \quad [8]$$

$$\begin{pmatrix} \delta_{\alpha k} \\ \delta_{\beta k} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\alpha k}^2 & \rho_k \sigma_{\alpha k} \sigma_{\beta k} \\ \rho_k \sigma_{\alpha k} \sigma_{\beta k} & \sigma_{\beta k}^2 \end{pmatrix} \right) \quad [9]$$

*Template 1.—*The intercepts (α_j) and slopes (β_j) are allowed to vary by region (Eq. 2) to account for regional differences in urban effects. This template

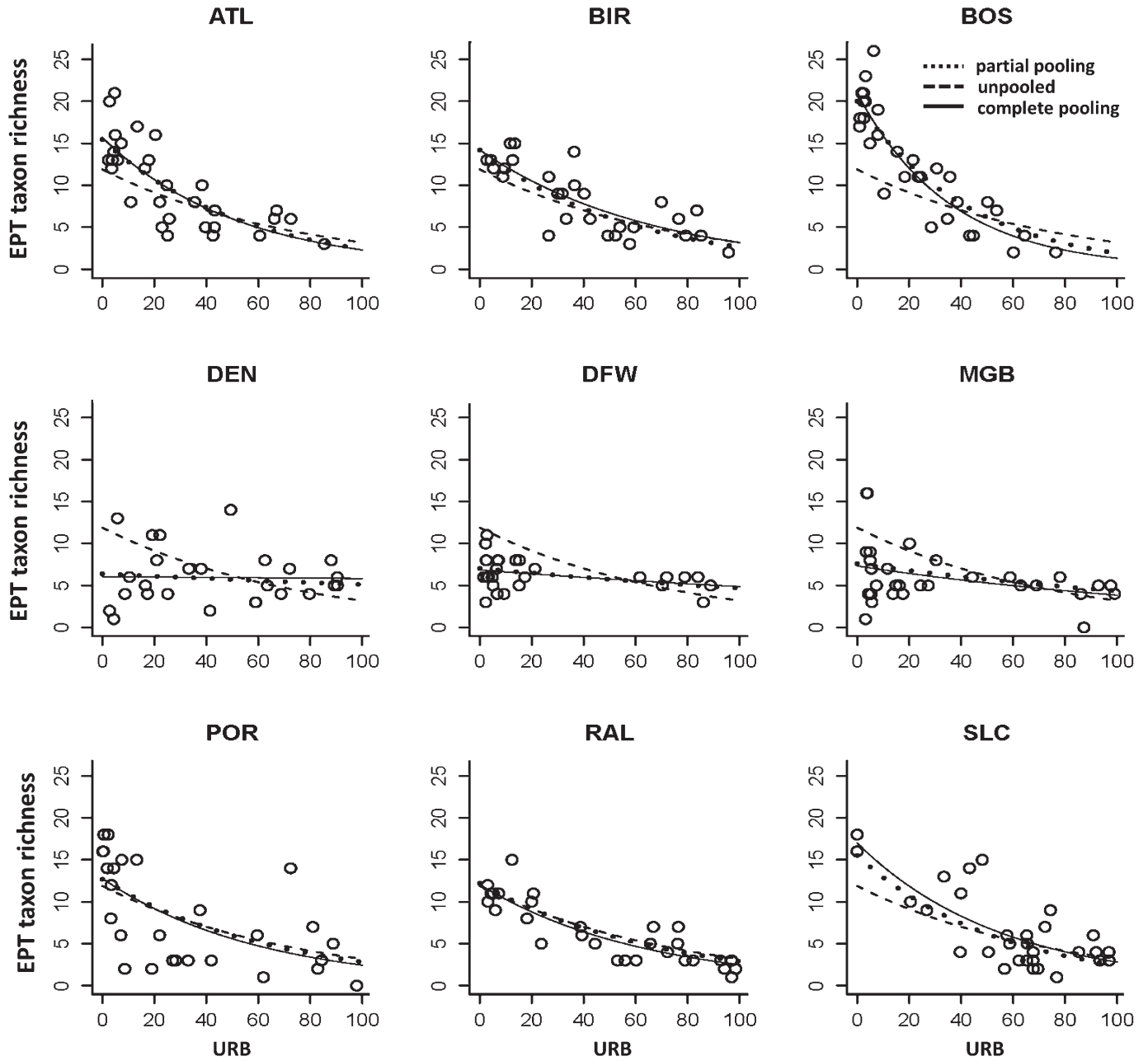


FIG. A1. Scatterplots of Ephemeroptera, Plecoptera, Trichoptera (EPT) richness compared to % developed land cover (URB) for each metropolitan area showing results for model 1 (no regional predictors) based on complete pooling, no pooling, and partial pooling of the data. Abbreviations for metropolitan areas are given in Fig. 1.

does not incorporate region-level predictors, so the estimates of the regional intercepts (α_j) and slopes (β_j) are modeled as a joint bivariate normal distribution centered on a vector of constant means (μ_α and μ_β ; Eq. 3). The error term σ_y^2 represents variations within regions and between basins beyond what is explained by % basin area in developed land (URB) (x_{ij}). The errors terms σ_α^2 and σ_β^2 represent unexplained variation between regions. The term ρ accounts for

the correlations between α_j and β_j because an increasing intercept typically is associated with a decreasing slope and vice versa.

Template 2.—The intercepts (α_j) and slopes (β_j) are allowed to vary by region (Eq. 4). However, intercepts and slopes are each modeled as a linear function of a single continuous regional predictor (P_{α_j} for intercepts, P_{β_j} for slopes) that estimates the mean intercept and slope for each region as a linear

TABLE A1. Description of the terms that define the parameters of the multilevel models (Eqs 1–9). AAG = antecedent agriculture, NMDS1 = nonmetric multidimensional scaling axis 1, RICHTOL = richness-weighted invertebrate tolerance, URB = % developed land cover, PRECIP = mean annual precipitation, TEMP = mean annual temperature.

| Model parameter | Description |
|----------------------|--|
| i | Number of basins (1...28 or 30) |
| j | Number of regions (1...9) |
| k | Categorical level of AAG, 1 for high level and 0 for low level |
| y_{ij} | Basin-level ecological response (e.g., NMDS1 or RICHTOL) in basin i of region j |
| x_{ij} | Basin-level URB in basin i of region j |
| α_j | Region-level intercept (estimated invertebrate or algal response at URB = 0) |
| β_j | Region-level slope (estimated change in invertebrate or algal response per unit change in URB) |
| σ_{y2} | Within-region variance in invertebrate or algal response |
| $\sigma_{\alpha2}$ | Between-region variance in intercepts |
| $\sigma_{\beta2}$ | Between-region variance in slopes |
| $\sigma_{\alpha j2}$ | Between-region variance of the regional effect on the intercept |
| $\sigma_{\beta j2}$ | Between-region variance of the regional effect on the slope |
| $\sigma_{\alpha k2}$ | Between-AAG group variance of the AAG effect on the intercept |
| $\sigma_{\beta k2}$ | Between-AAG group variance of the AAG effect on the slope |
| μ_{α} | Mean of region-level intercepts |
| μ_{β} | Mean of region-level slopes |
| ρ | Correlation of model coefficients α_j and β_j |
| ρ_j | Correlation of model coefficients $\delta_{\alpha j}$ and $\delta_{\beta j}$ |
| ρ_k | Correlation of model coefficients $\delta_{\alpha k}$ and $\delta_{\beta k}$ |
| $\gamma_{\alpha 0}$ | Hyperparameter intercept that along with $\gamma_{\beta 0}$ describes the relation between a region-level predictor and α_j |
| $\gamma_{\beta 0}$ | Hyperparameter slope that along with $\gamma_{\alpha 0}$ describes the relation between a region-level predictor and α_j |
| $\gamma_{\alpha 1}$ | Hyperparameter intercept that along with $\gamma_{\beta 1}$ describes the relation between a region-level predictor and β_j |
| $\gamma_{\beta 1}$ | Hyperparameter slope that along with $\gamma_{\alpha 1}$ describes the relation between a region-level predictor and β_j |
| $P_{\alpha j}$ | Region-level conditions (e.g., PRECIP, TEMP, or AAG) that predict the intercept α_j |
| $P_{\beta j}$ | Region-level conditions (e.g., PRECIP, TEMP, or AAG) that predict the slope β_j |
| $\delta_{\alpha j}$ | Effect of the region on the intercept |
| $\delta_{\beta j}$ | Effect of the region on the slope |
| $\delta_{\alpha k}$ | Effect of AAG group on the intercept ($k = 0$ for low AAG, 1 for high AAG) |
| $\delta_{\beta k}$ | Effect of AAG group on the slope ($k = 0$ for low AAG, 1 for high AAG) |

function of the regional predictor. These intercepts and slopes are modeled as a bivariate normal distribution centered on a different mean vector for each region (Eq. 5). The error term σ_y^2 represents variations within regions, and σ_{α}^2 and σ_{β}^2 represent unexplained variation between regions and between basins beyond what is explained by the region-level predictors. The term ρ accounts for the correlations between α_j and β_j . The linear predictors of α_j ($\gamma_{\alpha 0}$, $\gamma_{\alpha 1}$) and β_j ($\gamma_{\beta 0}$, $\gamma_{\beta 1}$) are hyperparameters that describe the relation between the region-level predictor(s) and α_j and β_j .

Template 3.—As in templates 1 and 2, the intercepts (α_j) and slopes (β_j) are allowed to vary by region (Eq. 6). However, this template uses 2 regional predictors—1 continuous and 1 categorical (AAG)—to estimate the regional intercepts and slopes as a linear function of the 2 regional predictors (Eq. 7). The intercepts and slopes are modeled as a joint bivariate distribution centered on a different mean vector for each region–AAG group combination. The region-varying components of the intercept and slope ($\delta_{\alpha j}$

and $\delta_{\beta j}$) are modeled as a bivariate normal distribution with the prior centered on 0 and with $\sigma_{\alpha j}^2$ and $\sigma_{\beta j}^2$ accounting for between-region variance of the regional effect on intercept and slope, respectively (Eq. 8). The AAG-varying components of the intercept and slope ($\delta_{\alpha k}$ and $\delta_{\beta k}$) also are modeled as a bivariate normal distribution with the prior centered on 0 and with $\sigma_{\alpha k}^2$ and $\sigma_{\beta k}^2$ accounting for between-AAG group variance of the AAG effect on intercept and slope, respectively (Eq. 9). The terms ρ_j and ρ_k account for the correlation between regional effects on slope and intercept and AAG effect on slope and intercept, respectively. The error term σ_y^2 represents variation within regions. This template addresses the discontinuous distribution of the AAG predictor, which is either high ($\geq 79\%$ of basin area in Milwaukee-Green Bay, Denver, and Dallas-Fort Worth) or low values ($< 25\%$ in Boston, Raleigh, Birmingham, Atlanta, Salt Lake City, and Portland). The linear predictors of α_j ($\gamma_{\alpha 0}$, $\gamma_{\alpha 1}$) and β_j ($\gamma_{\beta 0}$, $\gamma_{\beta 1}$) are hyperparameters that describe the relation between the region-level predictor(s) and α_j and β_j .

APPENDIX 2. Multilevel regression coefficients describing the responses of benthic macroinvertebrates to % developed land cover (URB) based on models with (models 2–9) and without (model 1) regional predictors. Response variables are as in Table 2. PRECIP = mean annual precipitation (cm), TEMP = mean annual temperature (°C), AAG = antecedent agriculture (% basin area). Model parameters are explained in Appendix 1, Table A1.

| Model | Parameter | Response variable | | | |
|--|---|-------------------|---------|---------|---------|
| | | NMDS1 | RICH | EPT | RICHTOL |
| Model 1: No region predictor | μ_α | 2.3773 | 3.6181 | 2.4420 | 5.0786 |
| | μ_β | -1.8051 | -0.4672 | -1.3130 | 1.4819 |
| Model 2: $P_{\alpha j} = \text{PRECIP}_j$ | $\gamma_{\alpha 0}$ | 1.4071 | 3.5097 | 1.7286 | 6.0320 |
| | $\gamma_{\alpha 1}$ | 0.0088 | 0.0010 | 0.0066 | -0.0088 |
| $P_{\beta j} = \text{PRECIP}_j$ | $\gamma_{\beta 0}$ | 0.2396 | -0.2255 | 0.1655 | 0.0783 |
| | $\gamma_{\beta 1}$ | -0.0187 | -0.0022 | -0.0138 | 0.0130 |
| Model 3: $P_{\alpha j} = \text{TEMP}_j$ | $\gamma_{\alpha 0}$ | 1.9011 | 3.6015 | 2.4444 | 3.7859 |
| | $\gamma_{\alpha 1}$ | 0.0383 | 0.0011 | -0.0002 | 0.1043 |
| $P_{\beta j} = \text{TEMP}_j$ | $\gamma_{\beta 0}$ | -1.1988 | -0.7234 | -1.3728 | 1.9889 |
| | $\gamma_{\beta 1}$ | -0.0492 | 0.0215 | 0.0049 | -0.0416 |
| Model 4: $P_{\alpha j} = \text{TEMP}_j$ | $\gamma_{\alpha 0}$ | 1.9478 | 3.4422 | 2.2694 | 3.9375 |
| | $\gamma_{\alpha 1}$ | 0.0347 | 0.0140 | 0.0144 | 0.0924 |
| $P_{\beta j} = \text{PRECIP}_j$ | $\gamma_{\beta 0}$ | -0.8360 | -0.2725 | -0.8864 | 1.9624 |
| | $\gamma_{\beta 1}$ | -0.0090 | -0.0017 | -0.0042 | -0.0049 |
| Model 5: $P_{\alpha j} = \text{AAG}_j$ | $\gamma_{\alpha 0}$ | 2.8182 | 3.7416 | 2.9257 | 4.1712 |
| | $\gamma_{\alpha 1}$ | -0.0106 | -0.7177 | -0.0125 | 0.0239 |
| $P_{\beta j} = \text{AAG}_j$ | $\gamma_{\beta 0}$ | -2.7835 | -0.0032 | -2.1826 | 2.6561 |
| | $\gamma_{\beta 1}$ | 0.0247 | 0.0065 | 0.0225 | -0.0310 |
| Model 6: $P_{\alpha j} = \text{PRECIP}_j$, high AAG | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][1]$ | 1.6839 | 3.6024 | 1.9973 | 5.5260 |
| $P_{\alpha j} = \text{PRECIP}_j$, low AAG | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][0]$ | 2.2717 | 3.8787 | 2.8556 | 3.6469 |
| | $\gamma_{\alpha 1}$ | 0.0029 | -0.0015 | -0.0010 | 0.0074 |
| $P_{\beta j} = \text{PRECIP}_j$, high AAG | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][1]$ | -0.2622 | -0.3921 | -0.2842 | 0.7075 |
| $P_{\beta j} = \text{PRECIP}_j$, low AAG | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][0]$ | -1.5920 | -0.9023 | -1.7099 | 3.0463 |
| | $\gamma_{\beta 1}$ | -0.0063 | 0.0024 | -0.0011 | -0.0074 |
| Model 7: $P_{\alpha j} = \text{TEMP}_j$, high AAG | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][1]$ | 1.6081 | 3.5040 | 2.0850 | 4.5714 |
| $P_{\alpha j} = \text{TEMP}_j$, low AAG | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][0]$ | 2.3207 | 3.7099 | 2.8988 | 2.9244 |
| | $\gamma_{\alpha 1}$ | 0.0262 | -0.0020 | -0.0143 | 0.1301 |
| $P_{\beta j} = \text{TEMP}_j$, high AAG | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][1]$ | -0.4887 | -0.5155 | -0.6903 | 1.0356 |
| $P_{\beta j} = \text{TEMP}_j$, low AAG | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][0]$ | -2.1013 | -0.9494 | -2.1721 | 3.0818 |
| | $\gamma_{\beta 1}$ | -0.0232 | 0.0278 | 0.0291 | -0.0756 |
| Model 8: $P_{\alpha j} = \text{TEMP}_j$, high AAG | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][1]$ | 1.5994 | 3.3225 | 1.8811 | 4.8601 |
| $P_{\alpha j} = \text{TEMP}_j$, low AAG | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][0]$ | 2.3119 | 3.5233 | 2.6896 | 3.2267 |
| | $\gamma_{\alpha 1}$ | 0.0267 | 0.0134 | 0.0026 | 0.1059 |
| $P_{\beta j} = \text{PRECIP}_j$, high AAG | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][1]$ | -0.3259 | 0.0486 | -0.1068 | 0.6651 |
| $P_{\beta j} = \text{PRECIP}_j$, low AAG | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][0]$ | -1.7101 | -0.2438 | -1.4415 | 2.9012 |
| | $\gamma_{\beta 1}$ | -0.0055 | -0.0031 | -0.0033 | -0.0066 |
| Model 9: $P_{\alpha j} = \text{CLAY}_j$ | $\gamma_{\alpha 0}$ | 2.4471 | 3.3949 | 3.8838 | 3.0545 |
| | $\gamma_{\alpha 1}$ | -0.0024 | 0.0559 | -0.0087 | -0.0203 |
| $P_{\beta j} = \text{CLAY}_j$ | $\gamma_{\beta 0}$ | -2.8883 | 3.3809 | 3.8837 | -2.4182 |
| | $\gamma_{\beta 1}$ | 0.0362 | -0.0629 | 0.0328 | 0.0360 |

APPENDIX 3. Multilevel regression coefficients describing the responses of benthic algae to % developed land cover (URB) based on models with (models 2–9) and without (model 1) regional predictors. Response variables are as in Table 2. PRECIP = mean annual precipitation (cm), TEMP = mean annual temperature (°C), AAG = antecedent agriculture (% basin area). Model parameters explained in Appendix 1, Table A1.

| Model | Parameter | Response variable | | | |
|---|---|-------------------|---------|---------|---------|
| | | HALO | EUTR | SILT | SENS |
| Model 1: No region predictor | μ_{α} | -2.6552 | -0.0695 | -1.5694 | 1.1225 |
| | μ_{β} | 0.9317 | 1.2006 | 1.3068 | -1.3726 |
| Model 2: $P_{\alpha j} = \text{PRECIP}_j$ $P_{\beta j} = \text{PRECIP}_j$ | $\gamma_{\alpha 0}$ | -1.5041 | 1.5509 | -1.2475 | 1.2900 |
| | $\gamma_{\alpha 1}$ | -0.0107 | -0.0149 | -0.0030 | -0.0016 |
| | $\gamma_{\beta 0}$ | 1.7147 | -0.1461 | 1.4812 | -1.5794 |
| Model 3: $P_{\alpha j} = \text{TEMP}_j$ $P_{\beta j} = \text{TEMP}_j$ | $\gamma_{\beta 1}$ | -0.0072 | 0.0123 | -0.0017 | 0.0019 |
| | $\gamma_{\alpha 0}$ | -3.1649 | 0.4107 | -3.2213 | 4.1460 |
| | $\gamma_{\alpha 1}$ | 0.0412 | -0.0396 | 0.1332 | -0.2443 |
| Model 4: $P_{\alpha j} = \text{TEMP}_j$ $P_{\beta j} = \text{PRECIP}_j$ | $\gamma_{\beta 0}$ | 2.4013 | 2.2227 | 2.2287 | -4.1730 |
| | $\gamma_{\beta 1}$ | -0.1183 | -0.0815 | -0.0740 | 0.2269 |
| | $\gamma_{\alpha 0}$ | -2.4833 | 1.8113 | -2.7275 | 2.6823 |
| Model 5: $P_{\alpha j} = \text{AAG}_j$ $P_{\beta j} = \text{AAG}_j$ | $\gamma_{\alpha 1}$ | -0.0150 | -0.1520 | 0.0927 | -0.1267 |
| | $\gamma_{\beta 0}$ | 2.3382 | 0.4574 | 2.1038 | -2.0707 |
| | $\gamma_{\beta 1}$ | -0.0129 | 0.0071 | -0.0074 | 0.0065 |
| Model 6: $P_{\alpha j} = \text{PRECIP}_j$, high AAG $P_{\alpha j} = \text{PRECIP}_j$, low AAG $P_{\beta j} = \text{PRECIP}_j$, high AAG $P_{\beta j} = \text{PRECIP}_j$, low AAG | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][1]$ | -2.1338 | 0.6847 | -2.0498 | 1.8816 |
| | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][0]$ | -4.2127 | -2.1632 | -4.7285 | 4.0733 |
| | $\gamma_{\alpha 1}$ | 0.0077 | 0.0103 | 0.0206 | -0.0204 |
| | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][1]$ | 2.1517 | 0.6404 | 2.1124 | -1.9803 |
| | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][0]$ | 3.2937 | 3.1270 | 4.0235 | -3.2883 |
| Model 7: $P_{\alpha j} = \text{TEMP}_j$, high AAG $P_{\alpha j} = \text{TEMP}_j$, low AAG $P_{\beta j} = \text{TEMP}_j$, high AAG $P_{\beta j} = \text{TEMP}_j$, low AAG | $\gamma_{\beta 1}$ | -0.0179 | -0.0100 | -0.0190 | 0.0134 |
| | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][1]$ | -2.3959 | 1.5171 | -2.3713 | 3.4461 |
| | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][0]$ | -4.1823 | -0.8421 | -4.2370 | 4.9518 |
| | $\gamma_{\alpha 1}$ | 0.0737 | -0.0031 | 0.1643 | -0.2687 |
| | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][1]$ | 2.2953 | 1.1335 | 1.7366 | -3.8034 |
| Model 8: $P_{\alpha j} = \text{TEMP}_j$, high AAG $P_{\alpha j} = \text{TEMP}_j$, low AAG $P_{\beta j} = \text{PRECIP}_j$, high AAG $P_{\beta j} = \text{PRECIP}_j$, low AAG | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][0]$ | 2.7440 | 3.3416 | 2.8872 | -4.7022 |
| | $\gamma_{\beta 1}$ | -0.1322 | -0.1121 | -0.0945 | 0.2446 |
| | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][1]$ | -2.1570 | 2.6635 | -1.9711 | 2.3795 |
| | $\gamma_{\alpha 0} + [\delta_{\alpha \text{AAG}}][0]$ | -3.9220 | 0.4039 | -3.8202 | 3.7975 |
| | $\gamma_{\alpha 1}$ | 0.0534 | -0.1000 | 0.1307 | -0.1786 |
| Model 9: $P_{\alpha j} = \text{CLAY}_j$ $P_{\beta j} = \text{CLAY}_j$ | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][1]$ | 1.8359 | -0.3625 | 1.4241 | -1.1891 |
| | $\gamma_{\beta 0} + [\delta_{\beta \text{AAG}}][0]$ | 2.7281 | 1.5901 | 2.8533 | -1.9739 |
| | $\gamma_{\beta 1}$ | -0.0138 | 0.0025 | -0.0100 | 0.0031 |
| | $\gamma_{\alpha 0}$ | -4.0937 | -2.3119 | -3.9358 | 4.2965 |
| | $\gamma_{\alpha 1}$ | 0.0478 | 0.0741 | 0.0789 | -0.1053 |
| | $\gamma_{\beta 0}$ | 2.0459 | 3.8218 | 2.9506 | -4.3464 |
| | $\gamma_{\beta 1}$ | -0.0365 | -0.0862 | -0.0545 | 0.0971 |