Science Advances
AAAS

Supplementary Materials for

Density declines, richness increases, and composition shifts in stream macroinvertebrates

Samantha L. Rumschlag *et al.*

Corresponding author: Samantha L. Rumschlag, rumschlag.samantha@epa.gov

Sci. Adv. **9**, eadf4896 (2023)
DOI: 10.1126/sciadv.adf4896

This PDF file includes:

Tables S1 to S7
Figs. S1 to S13

Table S1. Federal and regional project codes, associated with agencies, included in the present analyses. Data from six EPA federal projects and 64 federal and regional USGS projects were included.

Agency	Project Codes
EPA	EMAP, EMAP-W, WSA, NRSA0809, NRSA1314, NRSA1819
USGS	RSQA-NESQA, RSQA-SESQA, RSQA-MSQA, meskwakeco, ATG, WindCave, WhiteRiver, Hydro, BigCrkHF, FountainCk, RSQA-CSQA, RSQA-PNSQA, IdahoQWBI, BigWoodRiv, SilverTNC, Moab10, NECB, CONN, LINJ, HDSN, DELR, PODL, LSUS, ALBE, SANT, GAFL, ACFB, SOFL, MOBL, ALMN, KANA, WHMI, TENN, WMIC, MMSD Eco, LERI, REDN, UMIS, EIWA, CNBR, UIRB, LIRB, YELL, SPLT, OZRK, MISE, MS NAWQA, TRIN, ACAD, SCTX, RIOG, UCOL, NVBR, CAZB, GRSL, SOCA, SANJ, SACR, PUGT, NROK, CCYK, USNK, BoiseR WQ, WILL

Table S2. Results of linear mixed effects models primarily evaluating the effects of year on site-level density and α diversity, and regional $\bar{\alpha}$, γ_{est}, and β diversities for macroinvertebrate communities. Type II sums of squares were used in all models.

Endpoints and Source of Variation	*df*	*residual df*	*F*	*p*
Density, macroinvertebrates identified to genera				
Proportion of specimens identified to genus	1	8022.0	1310.3	<0.0001
Proportion of sample identified	1	7914.7	18764.9	<0.0001
Area sampled	1	7173.3	2169.7	<0.0001
Year	1	7585.9	8.4	0.004
Ecoregion	8	4361.7	11.2	<0.0001
Agency	1	4368.1	0.2	0.631
Ecoregion*Year	8	7530.5	1.0	0.424
Ecoregion*Agency	8	4068.7	14.6	<0.0001
Density, all macroinvertebrates in a sample				
Proportion of sample identified	1	7903.3	19426.5	<0.0001
Area sampled	1	7007.400	2253.2	<0.0001
Year	1	7463.3	0.9	0.355
Ecoregion	8	4322.0	10.0	<0.0001
Agency	1	4258.1	5.8	0.016
Ecoregion*Year	8	7481.6	1.3	0.238
Ecoregion*Agency	8	3963.9	15.3	<0.0001
α diversity (no. genera across at sites, rarefied)			χ^2	
Proportion of specimens identified to genus	1	-	127.3	<0.0001
Proportion of sample identified	1	-	7.6	0.006
Area sampled	1	-	41.4	<0.0001
Year	1	-	68.0	<0.0001
Ecoregion	8	-	930.7	<0.0001
Agency	1	-	715.1	<0.0001
Ecoregion*Year	8	-	28.9	0.0003
Ecoregion*Agency	8	-	56.1	<0.0001
α diversity (no. genera across at sites, not rarefied)				
Proportion of specimens identified to genus	1	-	84.3	<0.0001
Proportion of sample identified	1	-	0.1	0.798
Area sampled	1	-	23.0	<0.0001
Year	1	-	49.6	<0.0001
Ecoregion	8	-	916.3	<0.0001
Agency	1	-	1318.1	<0.0001
Ecoregion*Year	8	-	66.0	<0.0001
Ecoregion*Agency	8	-	81.4	<0.0001
$\bar{\alpha}$ diversity (mean no. genera across sites within ecoregions)			*F*	
Proportion of specimens identified to genus	1	262.3	3.1	0.080
Year	1	23.0	4.1	0.056
γest diversity (total no. of estimated genera within ecoregions)				
Proportion of specimens identified to genus	1	180.8	14.6	0.0002
Year	1	13.6	0.7	0.424
β diversity (spatial turnover within ecoregions)				
Proportion of specimens identified to genus	1	183.9	14.2	0.0002
Year	1	12.1	1.0	0.344

Table S3. Results of linear mixed effects models primarily evaluating the effects of year on site-level density and α diversity, and regional γ_{est} and β diversities for insect communities. Type II sums of squares were used in all models.

Endpoints and Source of Variation	*df*	*residual df*	*F*	*p*
Density				
Proportion of specimens identified to genus	1	8093.9	1381.6	<0.0001
Proportion of sample identified	1	7999.4	14242.7	<0.0001
Area sampled	1	7518.1	1589.8	<0.0001
Year	1	7705.9	49.6	<0.0001
Ecoregion	8	4491.4	22.6	<0.0001
Agency	1	4657.5	17.2	<0.0001
Ecoregion*Year	8	7672.8	0.8	0.644
Ecoregion*Agency	8	4383.4	16.3	<0.0001
α diversity (no. genera across at sites, rarefied)			χ^2	
Proportion of specimens identified to genus	1	-	209.7	<0.0001
Proportion of sample identified	1	-	13.4	0.0002
Area sampled	1	-	41.5	<0.0001
Year	1	-	17.8	<0.0001
Ecoregion	8	-	1034.4	<0.0001
Agency	1	-	448.1	<0.0001
Ecoregion*Year	8	-	25.0	0.002
Ecoregion*Agency	8	-	85.6	<0.0001
γest diversity (total no. of estimated genera within ecoregions)			*F*	
Proportion of specimens identified to genus	1	171.9	13.1	0.0004
Year	1	17.9	2.7	0.115
β diversity (spatial turnover within ecoregions)				
Proportion of specimens identified to genus	1	175.2	9.4	0.002
Year	1	12.6	1.4	0.256

Table S4. Results of linear mixed effects models primarily evaluating the interactive effects of year and land use on year on site-level density and α diversity, and regional γ_{est} and β diversities for macroinvertebrate communities. Type III sums of squares were used in all models.

Endpoints and Source of Variation	*df*	*residual df*	*F*	*p*
Density				
Proportion of specimens identified to genus	1	8025.4	1299.0	<0.0001
Proportion of sample identified	1	7925.7	18654.0	<0.0001
Area sampled	1	7160.8	2156.1	<0.0001
Year	1	7524.7	0.7	0.418
Ecoregion	8	7258.7	2.7	0.006
Agency	1	4643.4	1.8	0.177
Land use	3	6529.3	0.4	0.748
Ecoregion*Year	8	7491.0	1.0	0.464
Ecoregion*Agency	8	4078.3	14.6	<0.0001
Land use*Year	3	7160.1	0.4	0.717
α diversity (no. genera across at sites, rarefied)			χ^2	
Proportion of specimens identified to genus	1	-	114.7	<0.0001
Proportion of sample identified	1	-	3.8	0.051
Area sampled	1	-	39.2	<0.0001
Year	1	-	6.5	0.011
Ecoregion	8	-	162.6	<0.0001
Agency	1	-	19.5	<0.0001
Land use	3	-	65.7	<0.0001
Ecoregion*Year	8	-	33.0	<0.0001
Ecoregion*Agency	8	-	46.2	<0.0001
Land use*Year	3	-	9.4	0.024
γest diversity (total no. of estimated genera within ecoregions)			*F*	
Proportion of specimens identified to genus	1	357.3	4.4	0.038
Year	1	463.9	2.7	0.103
Agency	1	8.3	96.0	<0.0001
Land use	3	52.9	1.3	0.299
Land use*Year	3	451.7	4.7	0.003
β diversity (spatial turnover within ecoregions)				
Proportion of specimens identified to genus	1	388.4	0.002	0.967
Year	1	464.4	0.02	0.878
Agency	1	8.6	54.7	<0.0001
Land use	3	66.1	0.5	0.655
Land use*Year	3	454.9	2.0	0.111

Table S5. Results of linear mixed effects models primarily evaluating the interactive effects of year and land use on year on site-level density and α diversity, and regional γ_{est} and β diversities for insect communities. Type III sums of squares were used in all models.

Endpoints and Source of Variation	*df*	*residual df*	*F*	*p*
Density				
Proportion of specimens identified to genus	1	8090.8	1356.7	<0.0001
Proportion of sample identified	1	8002.2	14236.6	<0.0001
Area sampled	1	7488.2	1597.7	<0.0001
Year	1	7741.5	0.1	0.732
Ecoregion	8	7376.8	4.8	<0.0001
Agency	1	5032	12.5	0.0004
Land use	3	6699.5	2.2	0.086
Ecoregion*Year	8	7625.3	0.8	0.629
Ecoregion*Agency	8	4372	16.1	<0.0001
Land use*Year	3	7286.8	0.8	0.476
α diversity (no. genera across at sites, rarefied)			χ^2	
Proportion of specimens identified to genus	1	-	206.2	<0.0001
Proportion of sample identified	1	-	6.3	0.012
Area sampled	1	-	36.9	<0.0001
Year	1	-	0.01	0.928
Ecoregion	8	-	141.6	<0.0001
Agency	1	-	2.4	0.119
Land use	3	-	106.2	<0.0001
Ecoregion*Year	8	-	24.2	0.002
Ecoregion*Agency	8	-	67.5	<0.0001
Land use*Year	3	-	7.6	0.054
γest diversity (total no. of estimated genera within ecoregions)			*F*	
Proportion of specimens identified to genus	1	373.0	2.2	0.143
Year	1	530.4	0.3	0.608
Agency	1	8.3	56.8	<0.0001
Land use	3	47.2	3.7	0.017
Land use*Year	3	447.2	2.8	0.038
β diversity (spatial turnover within ecoregions)				
Proportion of specimens identified to genus	1	389.5	1.2	0.274
Year	1	529.2	0.2	0.674
Agency	1	8.7	52.3	<0.0001
Land use	3	61.6	1.8	0.166
Land use*Year	3	454.3	1.5	0.224

Table S6. Results of a partial distance-based redundancy analysis (dbRDA) evaluating how composition of macroinvertebrate communities at the ecoregion scale changed through time and with land use. The dbRDA is conditional upon agency, ecoregion, and a categorical variable of pre- and post-2004, a time when composition shifted due to improvements in taxonomists abilities to make identification at the genus level. Composition was defined as the presence or absence of families for all non-chironomid macroinvertebrates within ecoregions or subfamilies for all chironomids within ecoregions.

Endpoint and Source of Variation	*df*	*F*	*p*
Composition			
Land use	3	23.717	0.0001
Year	1	6.394	0.0001
Land use*Year	3	1.729	0.02
Residual	518		

Table S7. Results of a partial distance-based redundancy analysis (dbRDA) evaluating how composition of insect communities at the ecoregion scale changed through time and with land use. The dbRDA is conditional upon agency, ecoregion, and a categorical variable of pre- and post-2004, a time when composition shifted due to improvements in taxonomists abilities to make identification at the genus level. Composition was defined as the presence or absence of families for all non-chironomid macroinvertebrates within ecoregions or subfamilies for all chironomids within ecoregions.

Endpoint and Source of Variation	*df*	*F*	*p*
Composition			
Land use	3	27.321	0.0001
Year	1	1.752	0.098
Land use*Year	3	2.257	0.004
Residual	516		

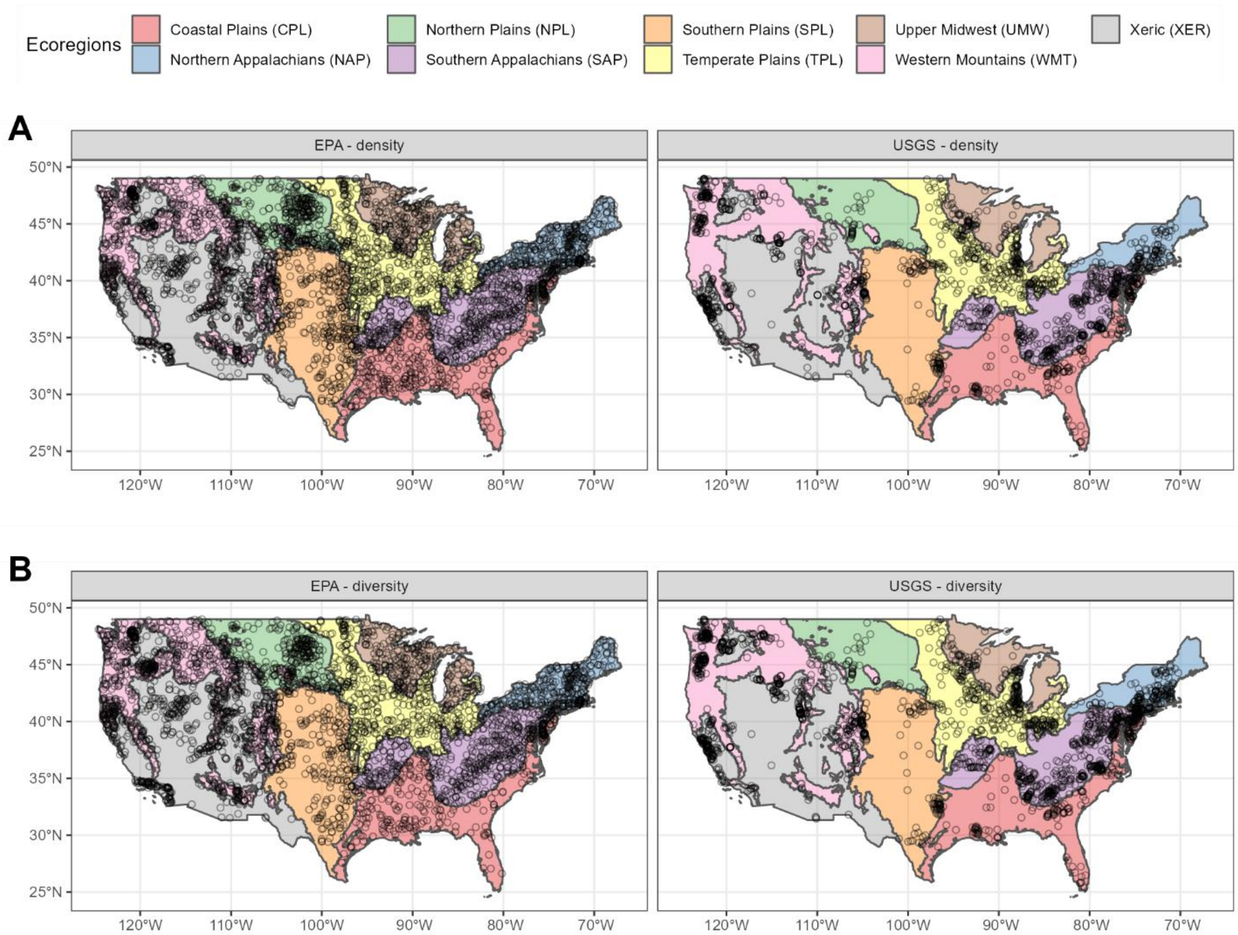

Figure S1. Extension of Fig. 1 showing the spatial distribution of sites within ecoregions that were used to calculate **A)** site-level biodiversity metrics, total densities and α diversity, and **B)** regional biodiversity metrics, including the average number of genera at sites within ecoregions ($\bar{\alpha}$ diversity), the total number of genera within ecoregions (γ_{est} diversity), and the spatial turnover (β diversity). There are fewer sites for regional biodiversity metrics than density because the calculation of regional biodiversity metrics required having at least 300 organisms for rarefaction to account for differences in sampling effort at the site level (see Methods).

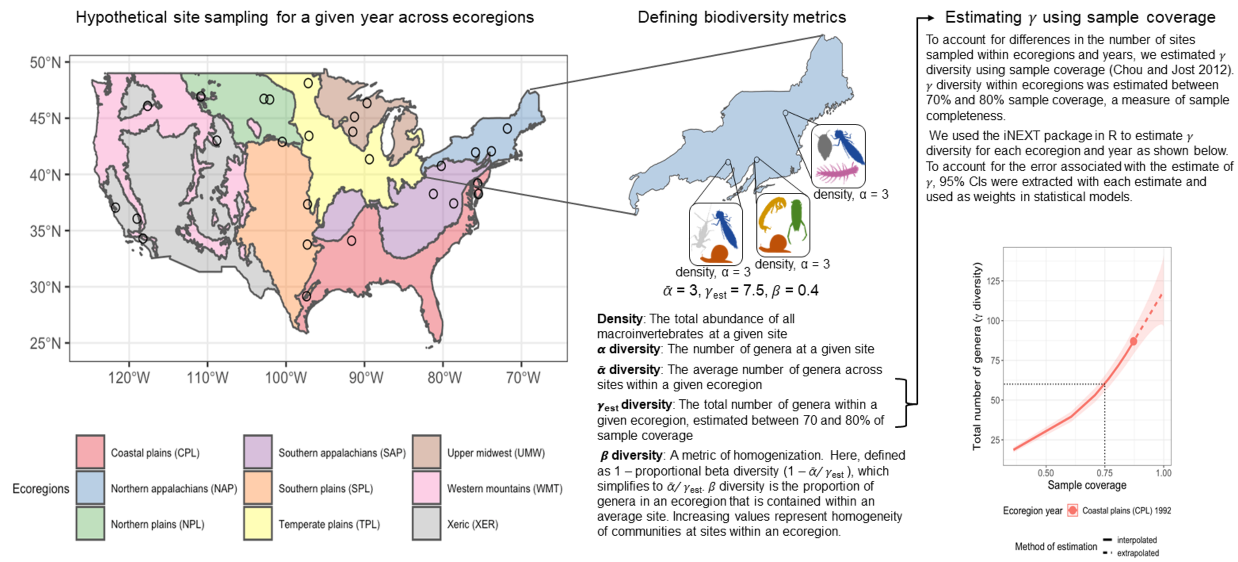

Figure S2. Conceptual model depicting and defining biodiversity metrics used in the current study. Total density of macroinvertebrates and α diversity are site level metrics, while $\bar{\alpha}$, γ_{est}, and β diversity are at the region level. γ_{est} diversity is the total number of genera within a given ecoregion estimated using sample coverage.

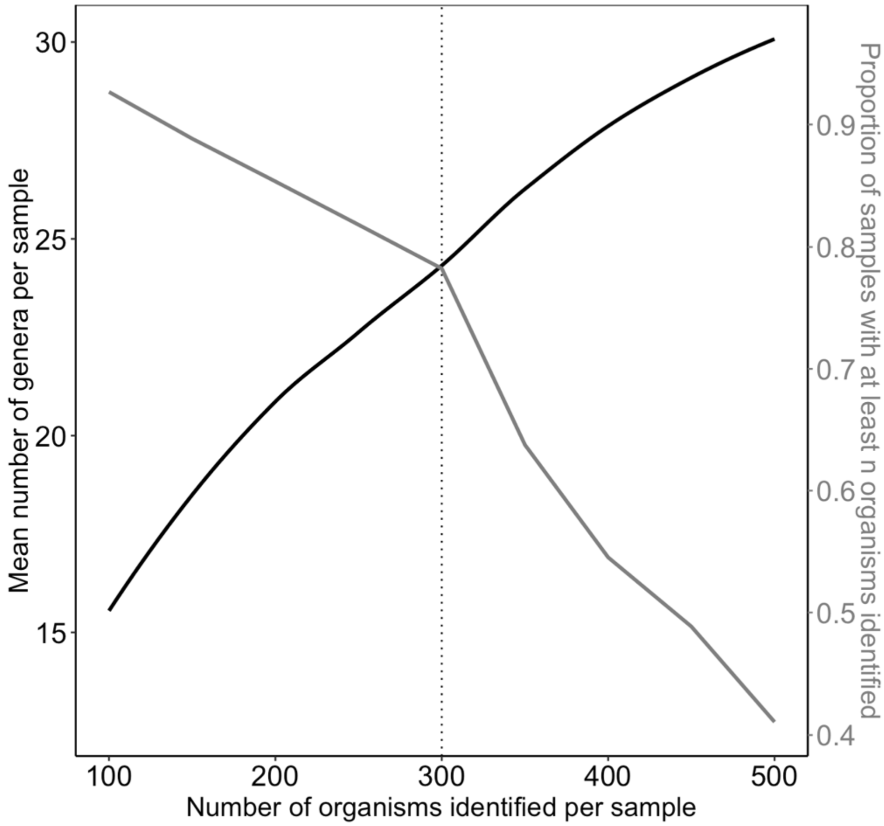

Figure S3. Visualization of the relationship between the mean number of genera gained, the proportion of the samples with at least n number of organisms identified, and the number of organisms identified per sample to justify a 300 threshold for rarefaction in the calculation of total densities at sites.

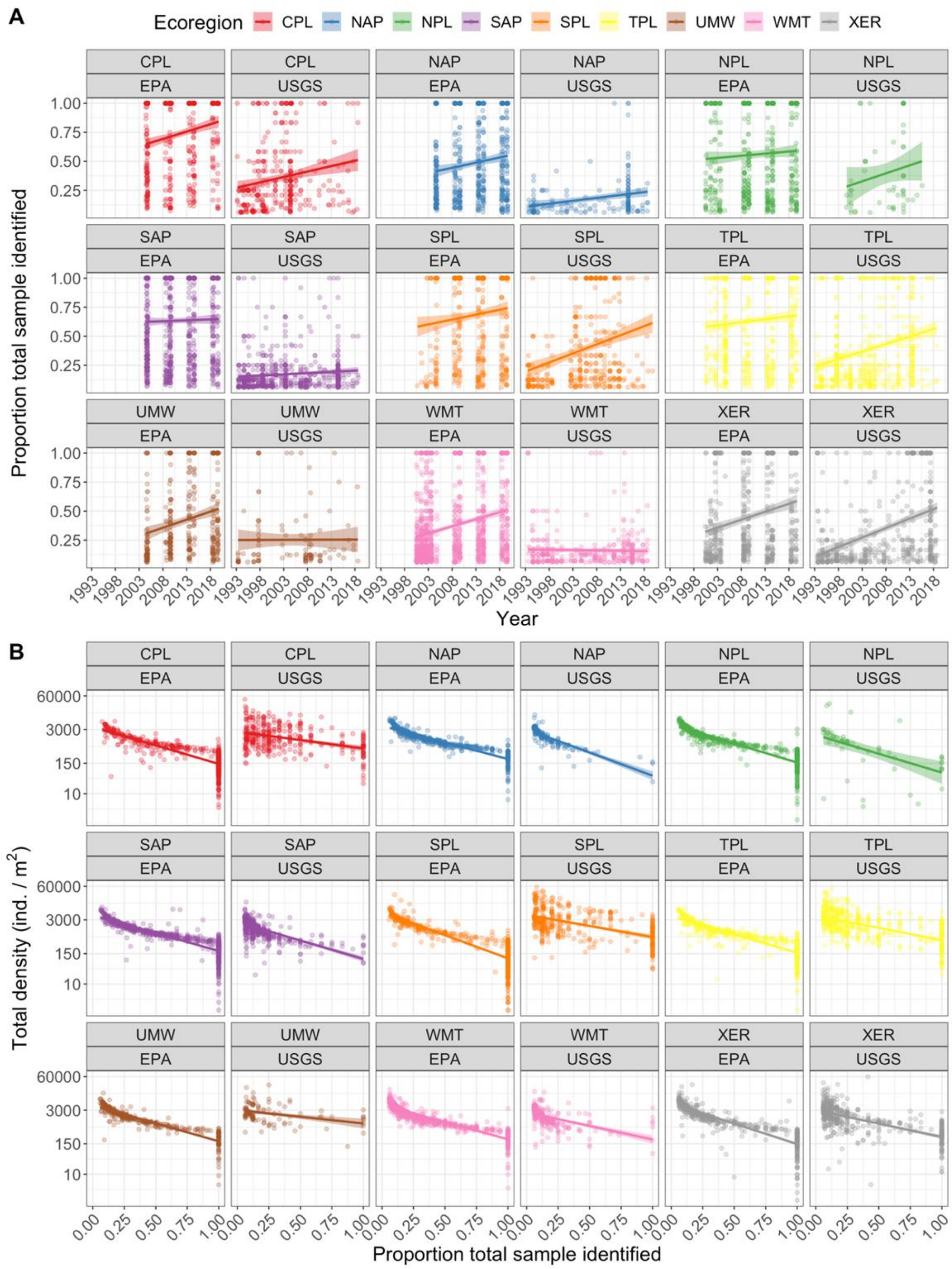

Figure S4. Plots showing relationships among proportion of the total sample that was identified, total density, and time. **A)** Within most ecoregions and agencies, as time proceeds the proportion of the samples identified at sites increases. **B)** At the same time, as the proportion of the sample identified increases, the total density of macroinvertebrates decreases. Together, these figures suggest the potential that early years (with low proportions of the sample identified) could be biased to high density measurements and later years (with high proportions of the sample identified) could be biased to low density measurements, which might suggest declines in total densities over time. To attempt to account for this bias, we include proportion of the sample identified as a covariate in models.

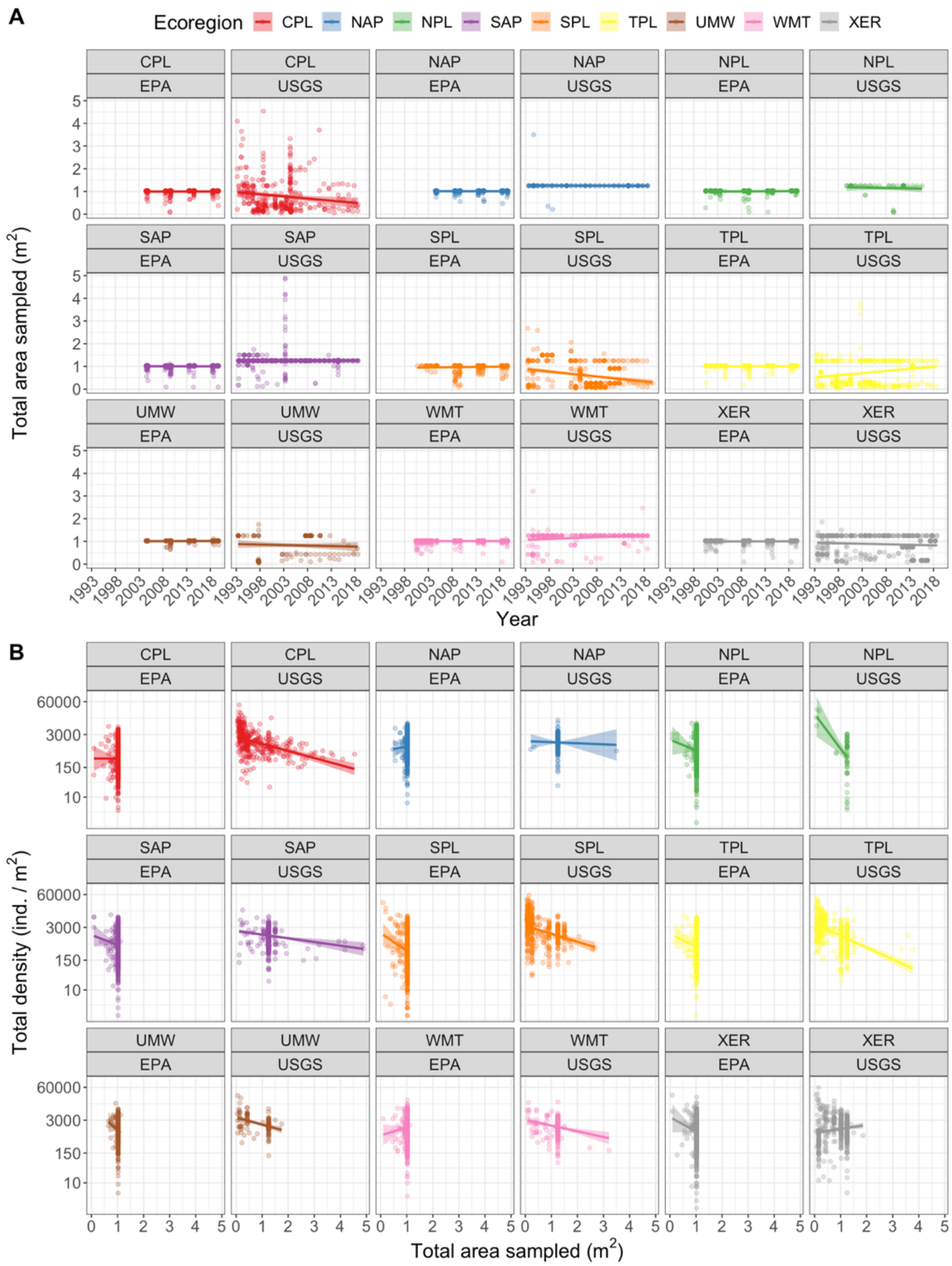

Figure S5. Plots showing relationships among total area sampled, total density, and time. **A)** Within some ecoregions for USGS sites, as time proceeds the total area sampled decreases. **B)** At the same time, as the total area sampled increases, the total density of macroinvertebrates decreases. Together, these figures suggest the potential that early years (with high areas sampled) could be biased to low density measurements and later years (with low areas sampled) could be biased to high density measurements, which might suggest increases in total densities over time for USGS sites. To account for this bias, we include total area sampled at sites as a covariate in models.

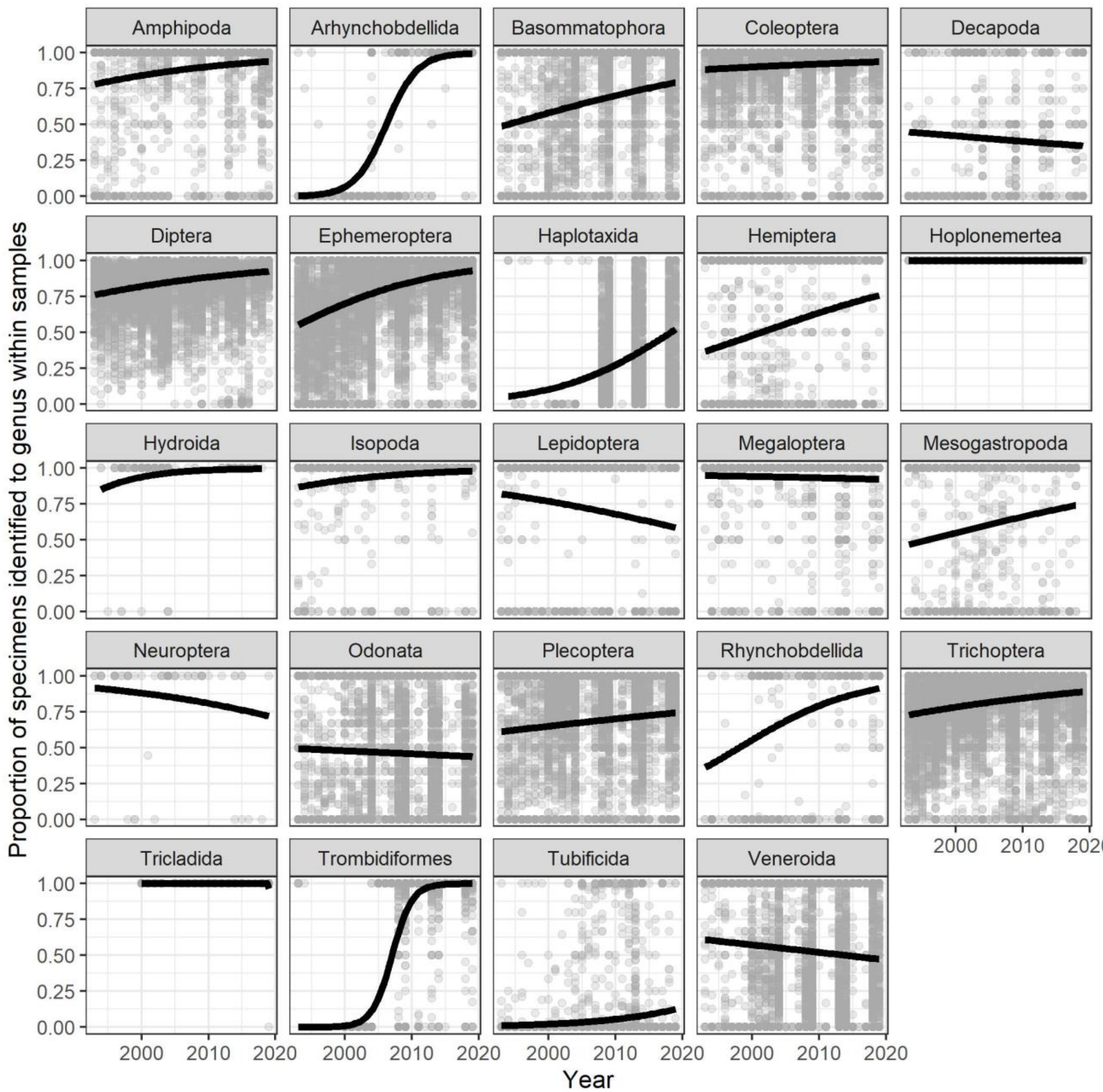

Figure S6. Plots showing relationships between time and the proportion of specimens identified to genus within samples. Positive trends likely indicate that within a given order taxonomists improved in their abilities to make identifications to the genus level. Given these trends, we included proportion of specimens identified to genus as a covariate in all models.

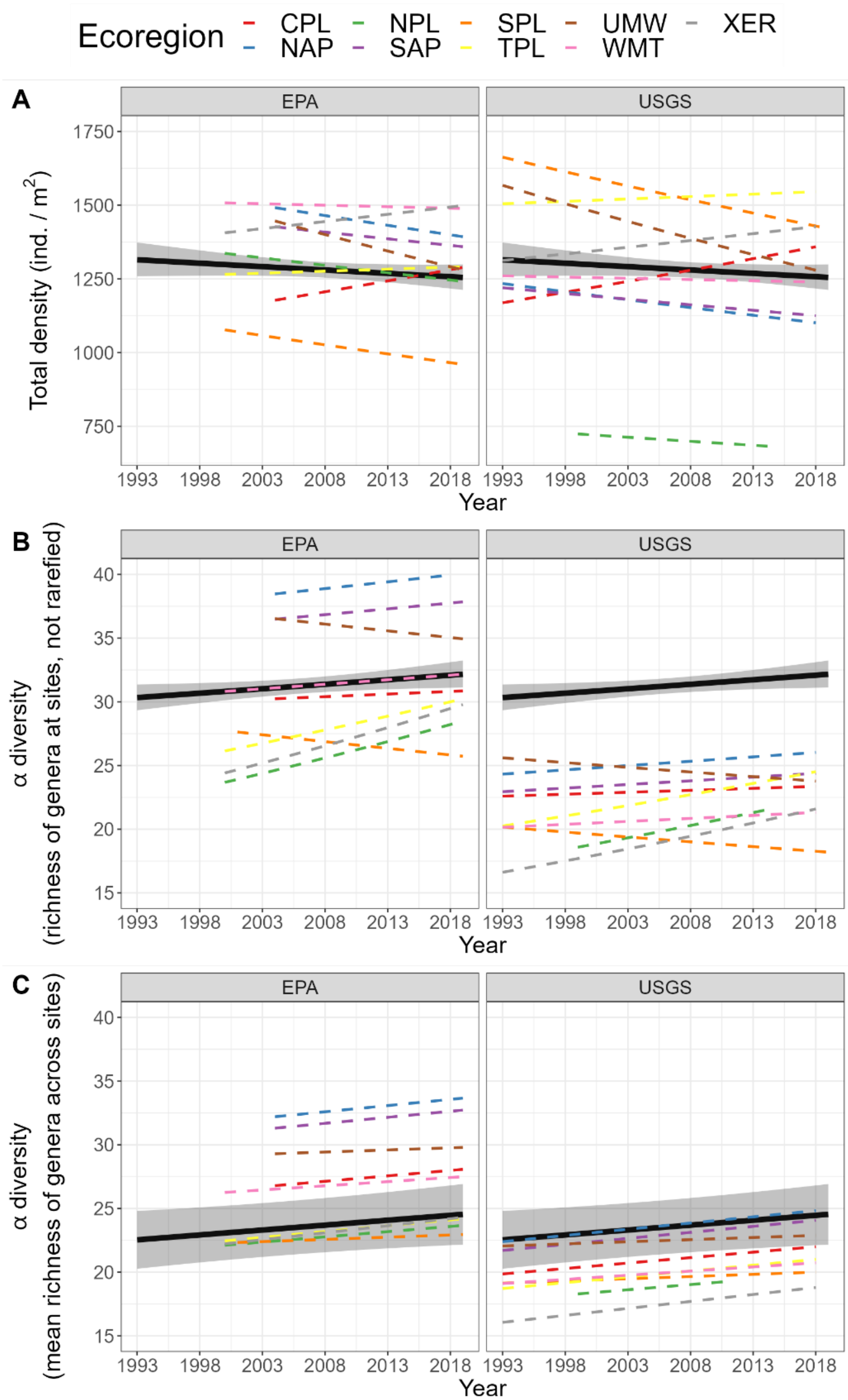

Figure S7. Temporal biodiversity trends of macroinvertebrate communities showing **A)** no change in total density derived from all organisms in samples (conditional $R^2 = 0.81$, marginal $R^2 = 0.77$, F(Year) = 0.9, p(Year) = 0.355), **B)** increases in α diversity generated without rarefication (conditional $R^2 = 0.74$, marginal $R^2 = 0.39$, χ^2 (Year) = 49.6, p(Year) <0.0001), and **C)** no change in $\bar{\alpha}$ diversity (conditional $R^2 = 0.72$, marginal $R^2 = 0.3$, F(Year) = 4.1, p(Year) = 0.056).

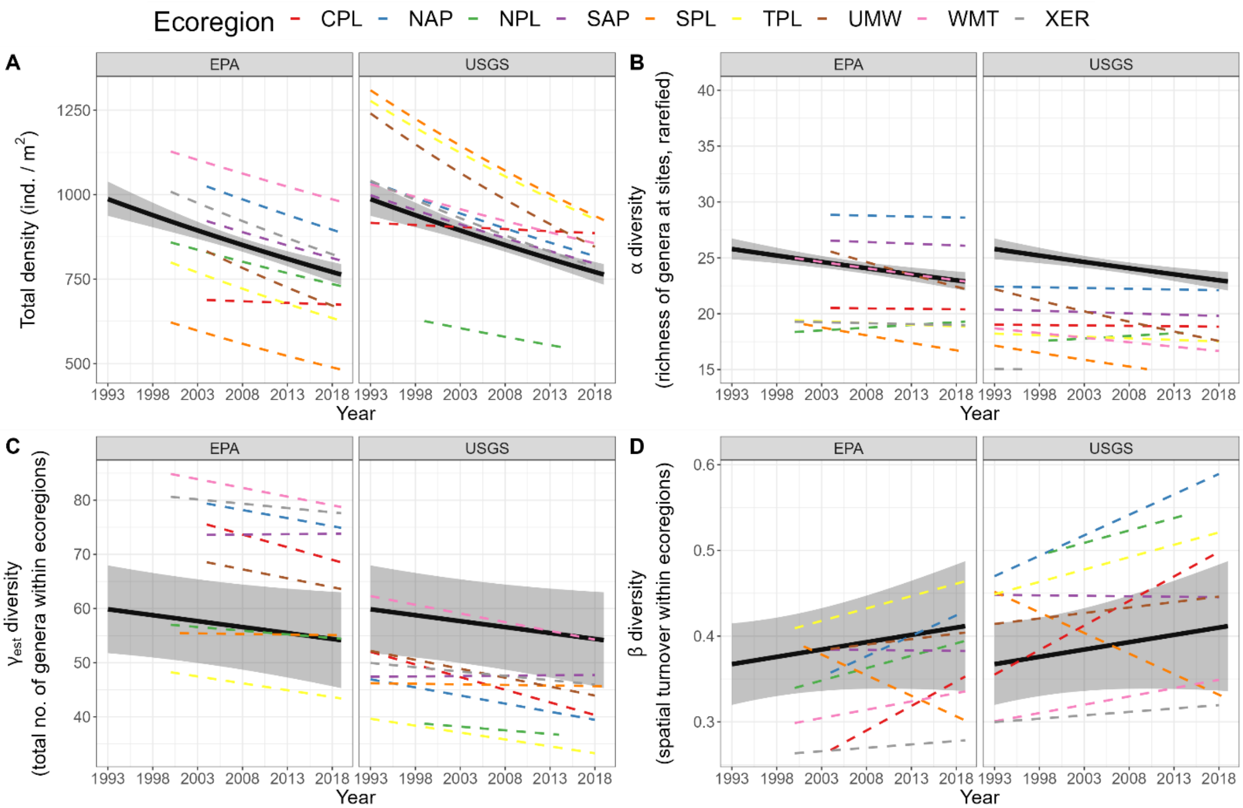

Figure S8. Temporal biodiversity trends of benthic insect communities showing **A)** decreases total density (conditional $R^2 = 0.79$, marginal $R^2 = 0.74$, F(Year) = 49.6, p(Year) < 0.0001), **B)** decreases in α diversity (conditional $R^2 = 0.67$, marginal $R^2 = 0.29$, χ^2(Year) = 17.8, p(Year) < 0.0001), **C)** no change in γ_{est} diversity (conditional $R^2 = 0.80$, marginal $R^2 = 0.2$, F(Year) = 2.7, p(Year) = 0.115), and **D)** no change in β diversity, spatial turnover within ecoregions, suggesting that communities are neither homogenizing or differentiating (conditional R^2 =, marginal R^2 =, F(Year) = 1.4, p(Year) = 0.256). Additional statistical output provided in Table S3.

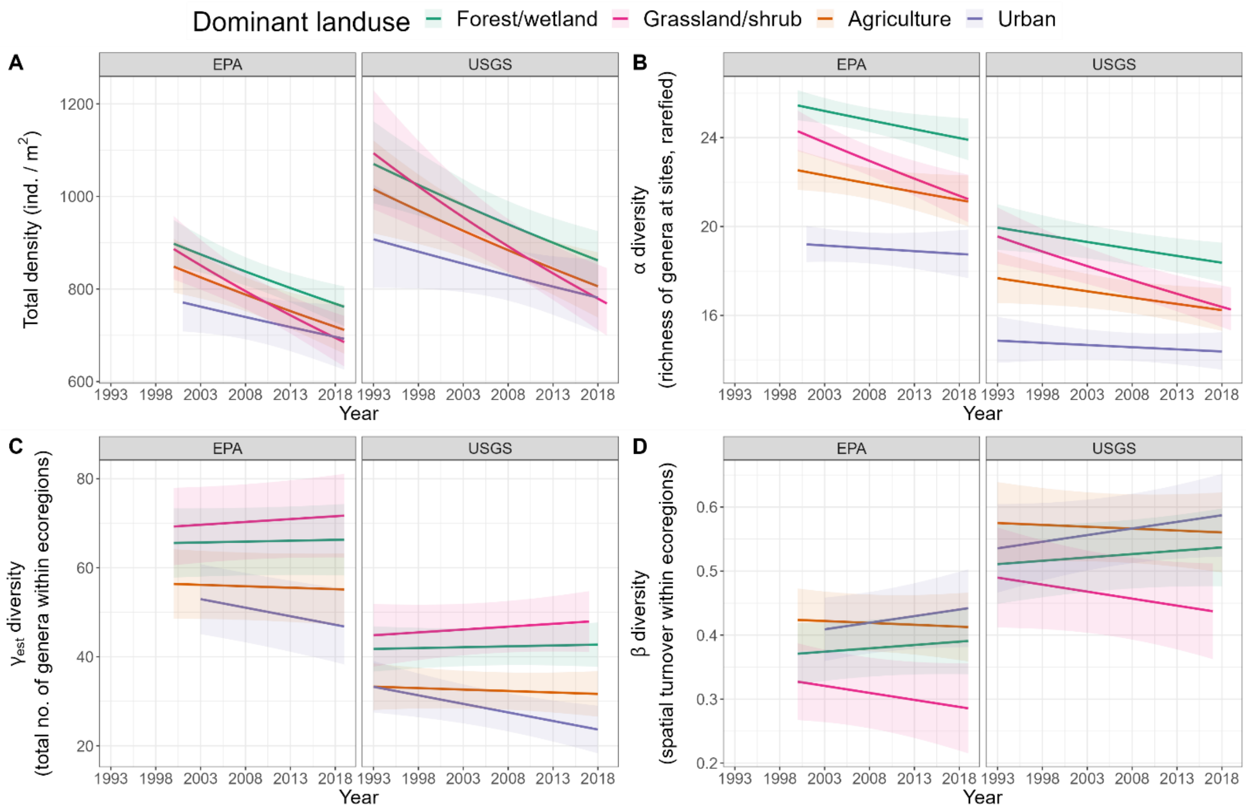

Figure S9. Temporal biodiversity trends of benthic insect communities across dominant land use types showing **A)** no interaction of time and land use on total density (conditional $R^2 = 0.79$, marginal $R^2 = 0.74$, *F*(Year) = 0.1, *p*(Year) = 0.732, *F*(Land use) = 2.2, *p*(Land use) = 0.086, *F*(Year*Land use) = 0.8, *p*(Year*Land use) = 0.476), **B)** a main effect of land use on α diversity, but no interaction of land use and time (conditional $R^2 = 0.63$, marginal $R^2 = 0.34$, χ^2 (Year) = 0.01, *p*(Year) = 0.928, χ^2(Land use) = 106.2, *p*(Land use) < 0.0001, χ^2(Year*Land use) = 7.6, *p*(Year*Land use) = 0.054). α diversity is lower in urban and agricultural streams compared to other land use types and all land use types are different from each other (pairwise comparisons with time held at mean, t-ratio ≥ -3.0, $p \leq 0.003$). **C)** γ_{est} diversity decreased in urban streams compared to all other land use types (conditional $R^2 = 0.70$, marginal $R^2 = 0.61$, *F*(Year) = 0.3, *p*(Year) = 0.608, *F*(Land use) = 3.7, *p*(Land use) = 0.017, *F*(Year*Land use) = 2.8, *p*(Year*Land use) = 0.038, pairwise comparisons of slopes, t-ratio ≥ 2.5, $p \leq 0.015$). **D)** The temporal trend in β does not vary across land use types (conditional $R^2 = 0.46$, marginal $R^2 = 0.29$, *F*(Year) = 0.2, *p*(Year) = 0.674, *F*(Land use) = 1.8, *p*(Land use) = 0.166, *F*(Year*Land use) = 1.5, *p*(Year*Land use) = 0.224). Additional statistical output provided in Table S5.

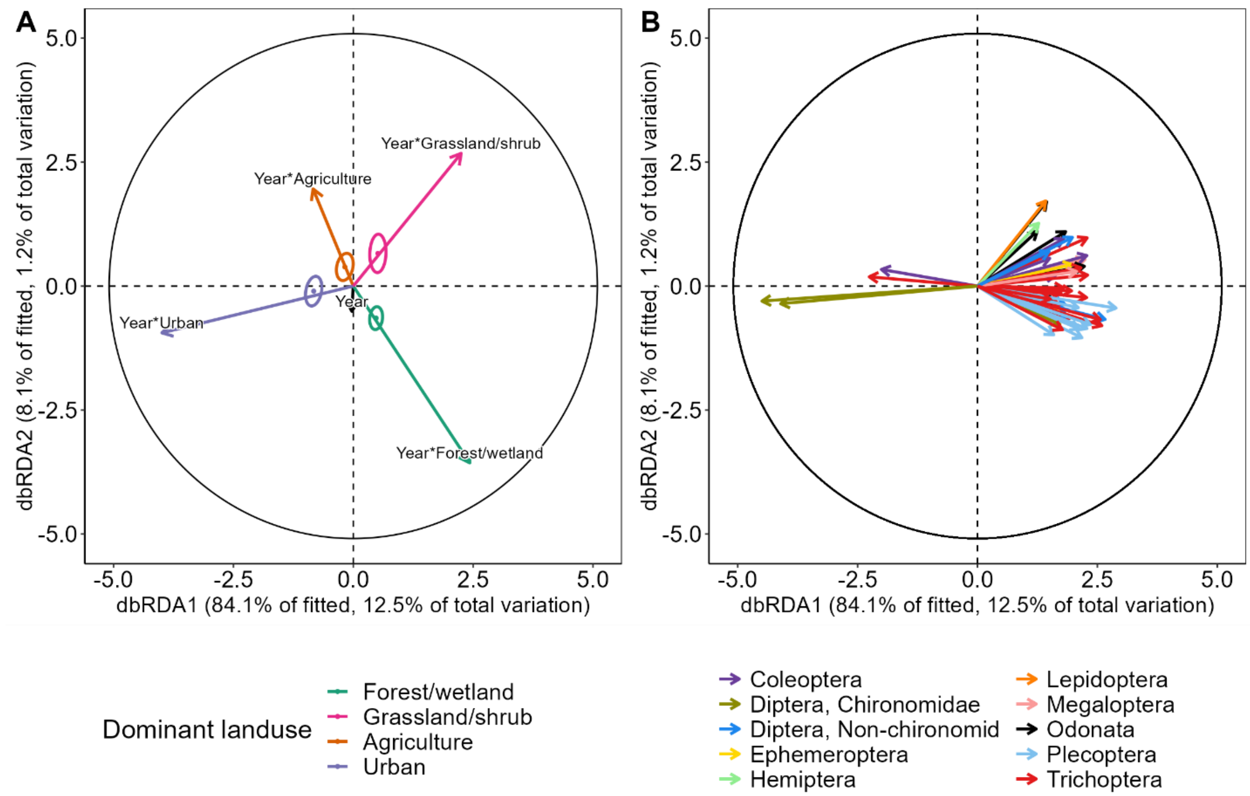

Figure S10. Partial distance-based redundancy analysis (partial dbRDA) demonstrating composition of insect communities changed differentially through time according to dominant land use. All non-insect macroinvertebrates have been excluded. Urban community composition diverges through time compared to both forest/wetland and grassland/shrub composition. Urban streams lose sensitive taxa (e.g. Ephemeroptera, Plecoptera, Trichoptera, Megaloptera) and gain tolerant taxa (e.g. Chironomidae) through time. **A)** Plot of model predictors showing the additive and interactive effects of land use and year. Individual points and circles are the centroids and 95% confidence intervals of ecoregion-year combinations according to dominant land use. **B)** The corresponding vector overlay of model responses. Individual vectors are either families for all non-chironomid macroinvertebrates or subfamilies for chironomids. As shown in the legend, non-chironomid insect vectors are colored by order and chironomid vectors are colored by family. Vectors with lengths less than 0.3 have been excluded. In both plots, the black circles correspond to vector lengths that would have a correlation coefficient of one with each axis. The entire dbRDA model explains 41.89% of the variance. Additional statistical output provided in Table S7.

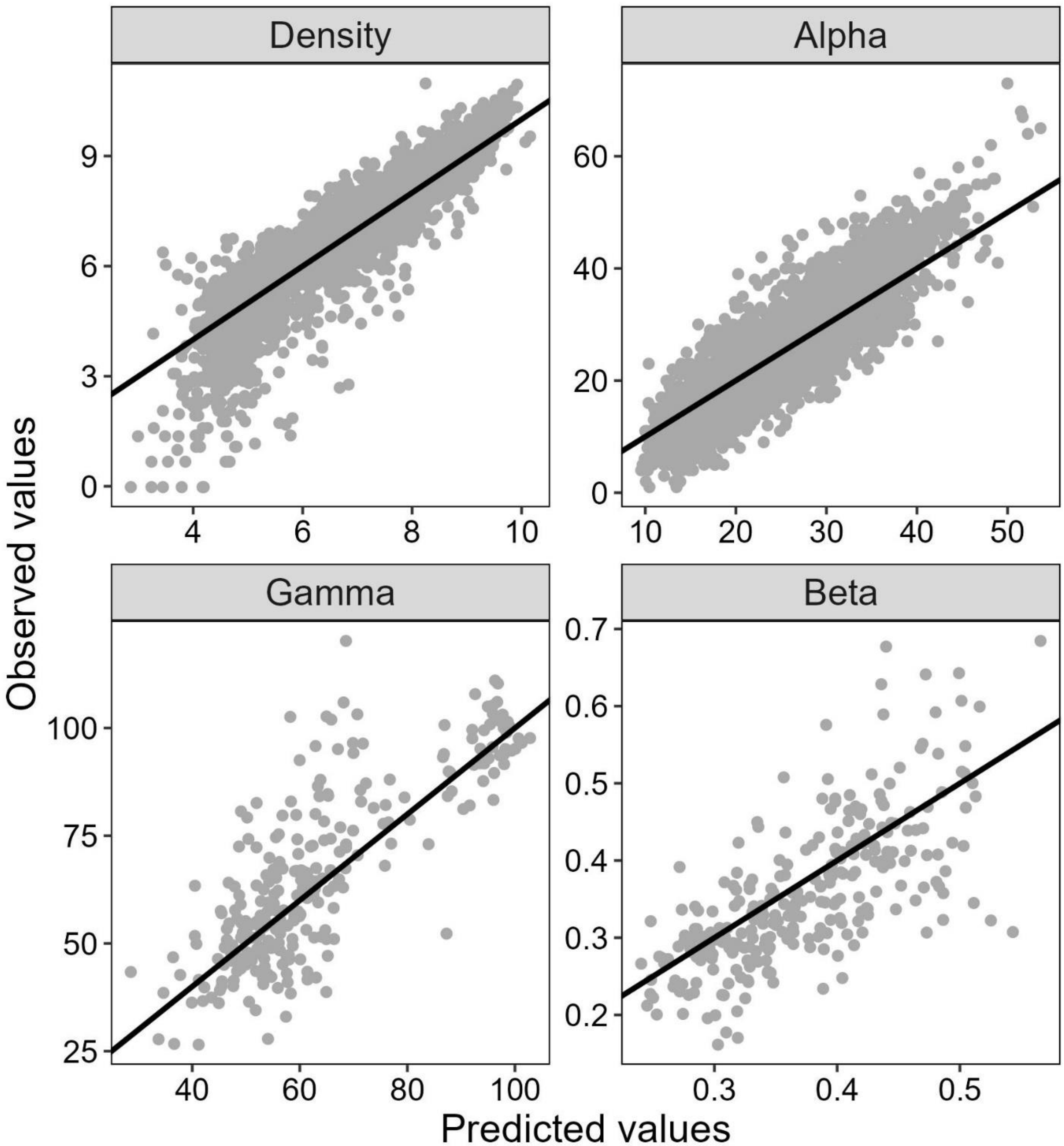

Figure S11. Correlations between predicted and observed values of temporal biodiversity trends in Fig. 2.

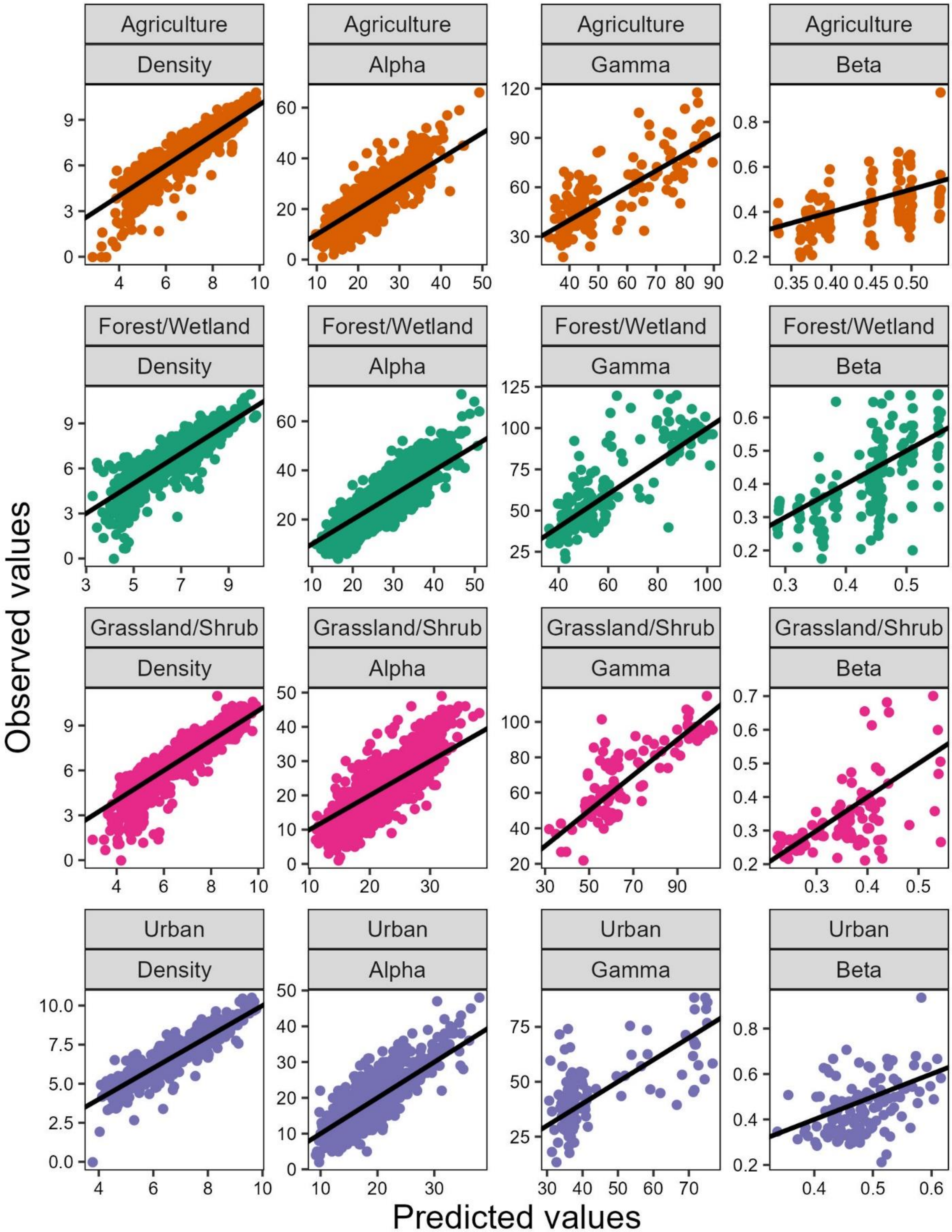

Figure S12. Correlations between the predicted and observed values of the temporal biodiversity trends across land use in Fig. 3.

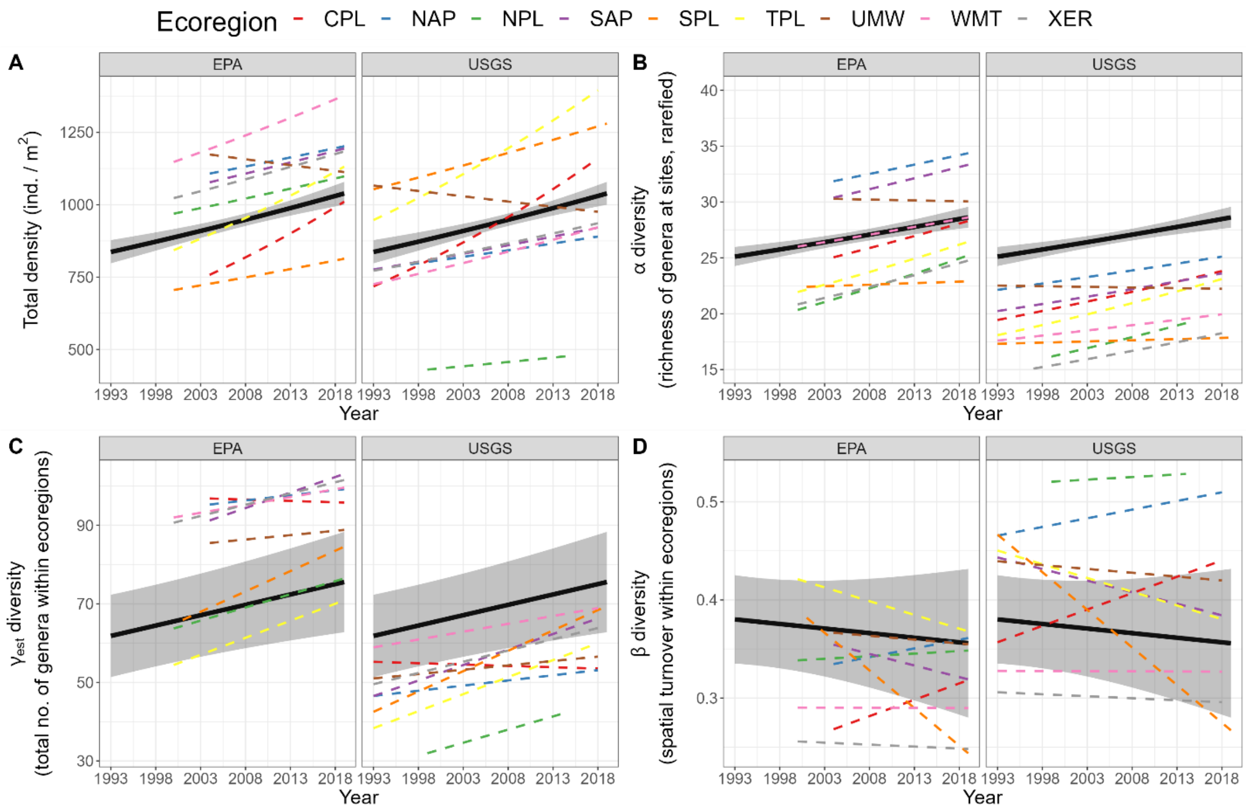

Figure S13. Temporal biodiversity trends of macroinvertebrate communities without accounting for improvements in the ability of taxonomists to make genera-level identifications showing **A)** increases in total density (conditional $R^2 = 0.79$, marginal $R^2 = 0.73$, $F(\text{Year}) = 49.857$, $p(\text{Year}) < 0.0001$), **B)** increases in α diversity (conditional $R^2 = 0.69$, marginal $R^2 = 0.33$, $\chi^2(\text{Year}) = 146.633$, $p(\text{Year}) < 0.0001$), **C)** increases in γ_{est} diversity (conditional R^2 =, marginal R^2 =, $F(\text{Year}) = 9.545$, $p(\text{Year}) = 0.016$), and **D)** no change in β diversity, spatial turnover within ecoregions, suggesting that communities are neither homogenizing or differentiating (conditional $R^2 = 0.72$, marginal $R^2 = 0.003$, $F(\text{Year}) = 0.507$, $p(\text{Year}) = 0.496$).