Length-mass equations for freshwater unionid mussel assemblages: Implications for estimating ecosystem function

Carla L. Atkinson[1,4], Thomas B. Parr[2,5], Brian C. van Ee[1,6], Daniel D. Knapp[1,7], Monica Winebarger[1,8], Kyle J. Madoni[1,9], and Wendell R. Haag[3,10]

[1]Department of Biological Sciences, University of Alabama, 300 Hackberry Lane, Tuscaloosa, Alabama 35487 USA
[2]Oklahoma Biological Survey, University of Oklahoma, 111 Chesapeake Street, Norman, Oklahoma 73019 USA
[3]United States Forest Service, Southern Research Station, 3761 Georgetown Road, Frankfort, Kentucky 40601 USA

Abstract: Biomass is often used to scale the contributions of individuals and their functional traits to a community or ecosystem. However, accurate biomass measurements can require destructive sampling, which is detrimental to long-lived organisms such as unionid mussels. We amassed a database of 6684 measurements of length and soft tissue dry mass (STDM) or shell dry mass (SHDM) from 43 species of unionid mussels to reduce the need for destructive sampling. We used these data to produce regression equations that relate maximum shell length to mass (either STDM or SHDM) at 3 taxonomic levels: family, phylogenetic tribe, and species. For 2 widely-distributed unionid species, *Amblema plicata* and *Elliptio complanata*, we present length–STDM regression equations from 6 waterbodies and basins. We also used bootstrapping resampling at the family level to develop a universal regression equation for unionid mussels. We compared models within all 3 taxonomic levels to determine if multiplicative (log–log transformation) or additive (non-linear parameter estimation) error structures provided better fits for biomass prediction. Models with multiplicative errors best fit length–STDM data at the family level (STDM = $6.63 \times 10^{-6} \times L_{max}^{2.89}$; $r^2 = 0.94$), for 83% of tribes ($n = 5$, average $r^2 = 0.95 \pm 0.03$), and for 90% of species ($n = 33$, average $r^2 = 0.93 \pm 0.07$). For length–SHDM, models with multiplicative errors best fit the family level (SHDM = $2.98 \times 10^{-4} \times L_{max}^{2.98}$; $r^2 = 0.86$), all tribes ($n = 5$, average $r^2 = 0.88 \pm 0.13$), and all species ($n = 27$, average $r^2 = 0.94 \pm 0.09$). Models with multiplicative errors also provided the best fit ($r^2 > 0.76$) for our 2 wide-ranging species, *A. plicata* and *E. complanata*. Finally, we present a case study based on data collected from 19 river sites in Alabama and Oklahoma, USA, to determine the performance of our power relationships (bootstrap resampling versus length–STDM regressions). In both river systems, tribe- and species-specific equations improved the prediction of unionid STDM over the family-level regression by 14%. Finer taxonomic resolution equations produce more accurate mass predictions, but where accurate taxonomic identifications are lacking, our family-level regression for STDM will produce acceptable estimates, which are key parameters when estimating mussel contributions to ecosystem services. Our study provides a toolkit that will allow scientists and managers to non-destructively quantify biomass (with uncertainty) of freshwater unionid mussels for secondary production, ecosystem function, and services estimates.

Key words: freshwater mussels, allometric scaling, length–mass regression, biomass, ecosystem function, river, stream, lake

Conserving and restoring ecosystem services is a critical component of sustainable ecosystem management. Recent research highlights the role of animal functional traits in providing or shaping those services (Duffy et al. 2007, Dirzo et al. 2014, Atkinson et al. 2017). Functional traits are often reported as a rate of activity/unit organism mass (Atkinson et al. 2017). Consequently, the mass of an individual organism (biomass) is frequently used to scale the contributions of individuals and their traits to a community or ecosystem (e.g., McIntyre et al. 2008, Atkinson et al. 2018). For many

E-mail addresses: [4]clatkinson@ua.edu; [5]thomas.parr@ou.edu; [6]bcvanee@crimson.ua.edu; [7]dknapp@crimson.ua.edu; [8]mmwinebarger@crimson.ua.edu; [9]kjmadoni@crimson.ua.edu; [10]wendell.haag@usda.gov

DOI: 10.1086/708950. Received 8 August 2019; Accepted 25 November 2019; Published online 15 June 2020.
Freshwater Science. 2020. 39(3):000–000.

animals, accurate and non-lethal biomass sampling is difficult and may cause long-term damage to a population. Long-lived animal populations, such as unionid mussels, may take years or decades to recover from destructive sampling. Allometric scaling equations that relate morphology (e.g., length) to tissue mass provide a non-destructive method for estimating individual- and assemblage-level biomass (Jenkins 2015, Straus and Aviles 2018). However, natural variation or errors in the biomass measurement can affect the uncertainty of regression estimates and are directly proportional to the uncertainty in estimates of derived ecosystem function, services, and the conservation value placed on those communities and ecosystems. Our study provides a toolkit that will allow scientists and managers to non-destructively quantify biomass (with uncertainty) of freshwater unionid mussels for secondary production, ecosystem function, and services estimates.

Unionid freshwater mussels (family Unionidae) once dominated (up to 200 ind/m^2) the benthic biomass of many rivers (Parmalee et al. 1998, Haag 2012), but they are now among the most endangered fauna, with 72% of the ~898 known species imperiled and a predicted extinction rate of ~6%/decade (Ricciardi et al. 1998). Freshwater mussels are particularly vulnerable to population declines because of their relative immobility, long lifespan (6–100+ y), and complex reproductive strategy in which larval unionids parasitize fish hosts as a dispersal mechanism (e.g., Haag and Warren 1997, 1998, Strayer 2008). Unionid mussels provide critical ecosystem services in freshwater ecosystems including water filtration (Malmqvist et al. 2001, Vaughn and Hakenkamp 2001), nutrient cycling that shapes nutrient stoichiometry (Atkinson et al. 2013, Atkinson and Vaughn 2015), and habitat modification that creates habitat for other species (Allen and Vaughn 2009, Atkinson et al. 2014a, Sansom et al. 2018).

Conservation efforts to restore lost species or prevent future species loss often struggle to quantify animals' traits and functions. Whereas past and ongoing management and conservation efforts typically emphasize quantifying all individuals (Carter and Resh 2001), more recent initiatives have focused on the functions provided by species (Strayer and Dudgeon 2010, Geist 2011). Freshwater mussels are often sampled both quantitatively and qualitatively to estimate individual density (Strayer and Smith 2003), but it can be challenging to non-destructively estimate the biomass and production of soft tissue and shell. Individual organism mass can increase by several orders of magnitude over their life span, so abundance alone may lead to over- or underestimation of individual contributions to ecosystem function and derived services. Capturing all inter/intraspecific variation is fundamental to anticipating and managing the ecosystem effects of species loss and additions (see Vaughn et al. 2015). However, it is not always possible to euthanize individuals (i.e., Bogan and Roe 2008) because of mussels' functional role in ecosystems, their slow growth, and their highly-endangered status. As such, it is important to determine the extent to which generalized regressions can predict biomass across species and locality without unnecessary mortality of rare individuals. Therefore, we propose the application of power-law equations to predict mussel biomass from non-lethal sampling methods (i.e., counting and measuring the total length of individuals found during traditional surveys).

Power-law relationships are ubiquitous and are used to explain natural and anthropogenic phenomena, both physical and biological (Leopold et al. 1964, Solé et al. 2002, Brown et al. 2004). In biology, they predict relationships between body size, physiological rates, and life-history traits (Brown et al. 2004). These relationships are frequently used to scale biological processes from cells to individuals to ecosystems and are of the general form $Y = aX^b$, where X and Y are measured properties and a and b are, respectively, the intercept and slope that describe the allometric relationship between X and Y. Their error structure can be multiplicative, which means the absolute magnitude of the error increases with the magnitude of the measurement (Kerkhoff and Enquist 2009):

$$Y = aX^b \times e^{\varepsilon} \text{or} \quad \log(Y) = \log(a) + \log(b)X + \varepsilon \quad \text{(Eq. 1)}$$

Alternatively, the error structure can be additive, which means the absolute magnitude of the error is constant with the magnitude of measurement (Ritz and Streibig 2008):

$$Y = aX^b + \varepsilon \quad \text{(Eq. 2)}$$

Linear regressions are typically assumed to have multiplicative errors (Eq. 1), whereas non-linear regressions are typically assumed to have additive errors (Eq. 2). Non-linear regressions are computationally intensive, so power-law relationships have primarily been investigated with linear regression and the assumption of multiplicative error. However, recent work suggests that non-linear regression may provide superior parameter estimates (Packard et al. 2011). The choice between linear regression on log-transformed data or non-linear regression on untransformed data should depend on the distribution of statistical error present in the data (Xiao et al. 2011).

We compiled published and unpublished data (Tables 1, 2) for 43 species of unionid mussels to further expand the available data and tools for invertebrate biomass scaling (as in Smock 1980, Golightly and Kosinski 1981, Benke et al. 1999). We used these data to answer 2 questions: 1) what scaling coefficients should be used to relate soft tissue dry mass (STDM) and shell dry mass (SHDM) to length (see Table S3 for wet mass) across all sampled unionid taxa and across phylogenetic tribes? and 2) do these relationships vary spatially within broadly-distributed species? In addition, we conducted a case study to understand how length–mass relationships at different levels of taxonomic resolution affect total biomass

Table 1. The length–mass equations (STDM = $a\ L_{max}{}^{b}$, where STDM is soft tissue dry mass [g], L_{max} is the maximum axial shell length [mm], and a and b are constants) for Unionidae (family level), unionid tribes (e.g., Anodontini), and species. All a and b constants are based on a log–log linear regression with multiplicative error structure. We used bootstrap resampling (1000×) to generate data for each species that was evenly distributed across the size ranges observed. n is the number of observations from which those bootstrapped data were drawn. n_{reg} is the number of individual resampled measurements used for the length–mass regression. The numbers of the 1000 resampling runs that were provided by a multiplicative linear regression (LR) error structure, an additive non-linear regression (NLR) error structure, or both error structures equally (model averaging, MA) are indicated. Reported coefficients (a, b) and confidence intervals (CI) are the average of the coefficient estimates for each of the 1000 bootstrap resamplings. The range is the range of length values available for each taxonomic level.

Classification	Common Name	LR	MA	NLR	n	n_{reg}	a	a 95% CI	b	b 95% CI	r^2	Range (mm)
Unionidae		1000	0	0	5802	990	9.48×10^{-6}	$7.50\times10^{-6} \pm 1.16\times10^{-5}$	2.82	2.77 ± 2.88	0.91	3.5–167
Amblemini												
Amblema plicata	Threeridge	1000	0	0	626	40	1.19×10^{-5}	$9.60\times10^{-6} \pm 1.48\times10^{-5}$	2.79	2.74 ± 2.84	0.99	6.2–167
Anodontini		1000	0	0	533	150	6.71×10^{-6}	$4.56\times10^{-6} \pm 9.44\times10^{-6}$	2.83	2.75 ± 2.92	0.97	12.5–136.6
Alasmidonta varicosa	Brook floater	999	0	1	68	40	6.36×10^{-6}	$3.80\times10^{-6} \pm 1.03\times10^{-5}$	2.91	2.78 ± 3.04	0.98	13.5–67.1
Anodonta implicata	Alewife floater	1000	0	0	386	40	1.13×10^{-6}	$7.94\times10^{-7} \pm 1.93\times10^{-6}$	3.2	3.05 ± 3.35	0.98	12.5–121.4
Pyganodon grandis	Giant floater	999	1	0	36	35	4.20×10^{-7}	$9.40\times10^{-8} \pm 2.13\times10^{-6}$	3.41	3.06 ± 3.77	0.92	83.5–136.6
Strophitus undulatus	Creeper	1000	0	0	26	35	1.66×10^{-5}	$7.61\times10^{-6} \pm 2.75\times10^{-5}$	2.65	2.49 ± 2.8	0.97	24.7–79.4
Lampsilini		1000	0	0	743	410	6.66×10^{-6}	$4.68\times10^{-6} \pm 9.52\times10^{-6}$	2.92	2.83 ± 3.01	0.91	3.5–129
Actinonaias ligamentina	Mucket	0	0	1000	76	40	4.31×10^{-7}	$1.36\times10^{-7} \pm 7.27\times10^{-7}$	3.53	3.38 ± 3.67	0.82	66.9–129
Lampsilis ornata	Southern pocketbook	1000	0	0	21	40	2.73×10^{-7}	$9.52\times10^{-8} \pm 8.56\times10^{-7}$	3.67	3.41 ± 3.94	0.95	34.4–107.5
Lampsilis straminea	Southern fatmucket	999	0	1	16	35	7.58×10^{-7}	$2.54\times10^{-7} \pm 2.21\times10^{-6}$	3.49	3.24 ± 3.75	0.96	27.4–89
Lampsilis teres	Yellow sandshell	1000	0	0	11	25	8.44×10^{-7}	$4.09\times10^{-7} \pm 2.02\times10^{-6}$	3.34	3.14 ± 3.54	0.98	19.8–180.6
Leptodea ochracea	Tidewater mucket	1000	0	0	57	35	2.40×10^{-6}	$1.21\times10^{-6} \pm 5.13\times10^{-6}$	3.15	2.95 ± 3.34	0.97	9–72.9
Medionidus acutissimus	Alabama moccasinshell	17	70	913	15	30	6.15×10^{-6}	$2.42\times10^{-6} \pm 9.79\times10^{-6}$	2.86	2.69 ± 3.03	0.99	7.6–34.3
Obliquaria reflexa	Threehorn wartyback	1000	0	0	82	30	1.76×10^{-6}	$7.69\times10^{-7} \pm 4.06\times10^{-6}$	3.33	3.09 ± 3.56	0.96	12–65.7
Obovaria unicolor	Alabama hickorynut	1000	0	0	80	35	1.58×10^{-6}	$9.61\times10^{-7} \pm 4.14\times10^{-6}$	3.45	3.23 ± 3.67	0.97	11.8–49.3
Ptychobranchus fasciolaris	Kidneyshell	1000	0	0	18	40	1.72×10^{-7}	$1.85\times10^{-8} \pm 1.23\times10^{-6}$	3.69	3.29 ± 4.08	0.9	49.4–97.8
Ptychobranchus occidentalis	Ouachita kidneyshell	796	152	52	34	40	1.58×10^{-5}	$2.23\times10^{-6} \pm 1.43\times10^{-4}$	2.65	2.2 ± 3.09	0.78	67.8–104.5
Truncilla truncata	Deertoe	0	1	999	16	30	4.95×10^{-3}	$4.93\times10^{-3} \pm 4.97\times10^{-3}$	1.5	1.5 ± 1.5	1	33.1–58.8
Villosa nebulosa	Alabama rainbow mussel	1000	0	0	103	30	6.98×10^{-6}	$3.00\times10^{-6} \pm 1.41\times10^{-5}$	2.87	2.6 ± 3.13	0.94	8.5–45
Pleurobemini		1000	0	0	3520	230	6.84×10^{-6}	$4.08\times10^{-6} \pm 1.10\times10^{-5}$	2.87	2.74 ± 3	0.9	4.2–103.6
Elliptio arca	Alabama spike	997	1	2	54	30	2.70×10^{-6}	$1.28\times10^{-6} \pm 5.05\times10^{-6}$	2.98	2.81 ± 3.15	0.98	8.3–89.3
Elliptio complanata	Eastern elliptio	1000	0	0	3222	40	4.11×10^{-6}	$3.06\times10^{-6} \pm 5.35\times10^{-6}$	2.93	2.78 ± 3.07	0.98	4.2–100
Elliptio dilatata	Spike	999	1	0	30	40	4.53×10^{-7}	$9.88\times10^{-8} \pm 2.86\times10^{-6}$	3.45	3.03 ± 3.87	0.88	46.5–103.6
Fusconaia cerina	Gulf pigtoe	990	7	3	28	40	8.49×10^{-6}	$4.00\times10^{-6} \pm 1.62\times10^{-5}$	2.92	2.75 ± 3.1	0.97	23.7–70
Fusconaia flava	Wabash pigtoe	323	199	478	95	40	3.44×10^{-5}	$-2.22\times10^{-5} \pm 5.61\times10^{-5}$	2.63	2.24 ± 3.01	0.81	23.7–72.3
Pleurobema decisum	Southern clubshell	1000	0	0	37	40	2.94×10^{-5}	$1.65\times10^{-5} \pm 5.46\times10^{-5}$	2.44	2.28 ± 2.59	0.96	15.6–67
Quadrulini		1000	0	0	290	160	3.95×10^{-5}	$2.54\times10^{-5} \pm 6.39\times10^{-5}$	2.54	2.42 ± 2.66	0.92	5.2–119.6
Cyclonaias asperata	Alabama orb	1000	0	0	101	40	1.40×10^{-5}	$7.43\times10^{-6} \pm 2.94\times10^{-5}$	2.86	2.6 ± 3.13	0.92	24.3–59.4
Cyclonaias pustulosa	Pimpleback	1000	0	0	68	40	5.19×10^{-6}	$4.15\times10^{-6} \pm 6.38\times10^{-6}$	3.18	3.08 ± 3.28	0.99	5.2–68.8
Quadrula rumphiana	Ridged mapleleaf	963	16	21	65	40	2.60×10^{-5}	$7.69\times10^{-6} \pm 5.75\times10^{-5}$	2.67	2.35 ± 2.98	0.89	33.9–77.8
Tritogonia verrucosa	Pistolgrip	1000	0	0	56	40	3.69×10^{-6}	$7.18\times10^{-7} \pm 2.51\times10^{-5}$	2.98	2.59 ± 3.38	0.85	52.5–119.6

Table 2. The length–mass equations (SHDM = $a\ L_{max}{}^{b}$, where SHDM is shell dry mass [g], L_{max} is the maximum axial shell length [mm], and a and b are constants) for Unionidae (family level), unionid tribes (e.g., Amblemini), and species. All a and b constants are based on a log–log linear regression with multiplicative error structure. Bootstrap resampling (1000×) was used to generate data for each species that was evenly distributed across the size ranges observed. n is the number of observations from which those bootstrapped data were drawn. n_{reg} is the number of individual resampled measurements used for regression. The numbers of the 1000 resampling runs that were provided by a multiplicative linear regression (LR) error structure, additive non-linear regression (NLR) error structure, or both error structures equally (model averaging, MA) are indicated. Reported coefficients and confidence intervals are the average of the coefficient estimates for each of the bootstrap resamplings. The range is the range of length values available for each taxonomic level.

Classification	Common Name	LR	MA	NLR	n	n_{reg}	a	a 95% CI	b	b 95% CI	r^2	Range (mm)
Unionidae		1000	0	0	2063	770	2.98×10^{-4}	$1.97\times10^{-4} \pm 4.38\times10^{-4}$	2.79	2.69 ± 2.89	0.80	3.5–129.0
Amblemini												
Amblema plicata	Threeridge	1000	0	0	598	40	3.52×10^{-4}	$2.66\times10^{-4} \pm 5.93\times10^{-4}$	2.87	2.76 ± 2.97	0.99	6.2–125.8
Lampsilini		1000	0	0	622	340	5.81×10^{-5}	$3.61\times10^{-5} \pm 9.52\times10^{-5}$	3.13	3.01 ± 3.25	0.89	3.5–129.0
Actinonaias ligamentina	Mucket	986	10	1	29	40	1.28×10^{-5}	$1.46\times10^{-6} \pm 7.17\times10^{-5}$	3.52	3.11 ± 3.92	0.89	66.9–129.0
Lampsilis ornata	Southern pocketbook	1000	0	0	20	35	2.17×10^{-6}	$6.27\times10^{-7} \pm 5.29\times10^{-6}$	3.81	3.59 ± 4.02	0.97	34.4–107.5
Lampsilis straminea	Southern fatmucket	1000	0	0	16	35	7.01×10^{-6}	$3.91\times10^{-6} \pm 1.17\times10^{-5}$	3.67	3.52 ± 3.81	0.99	27.4–89.0
Lampsilis teres	Yellow sandshell	977	14	9	11	25	6.86×10^{-6}	$3.24\times10^{-6} \pm 1.52\times10^{-5}$	3.42	3.24 ± 3.61	0.98	19.8–108.6
Medionidus acutissimus	Alabama moccasinshell	1000	0	0	15	30	3.31×10^{-5}	$2.38\times10^{-5} \pm 4.75\times10^{-5}$	3.07	2.95 ± 3.20	0.99	7.6–34.3
Obliquaria reflexa	Threehorn wartyback	1000	0	0	82	30	1.00×10^{-4}	$6.87\times10^{-5} \pm 1.43\times10^{-4}$	3.27	3.16 ± 3.38	0.99	12.0–65.7
Obovaria unicolor	Alabama hickorynut	949	24	27	80	35	1.35×10^{-5}	$5.25\times10^{-6} \pm 2.26\times10^{-5}$	3.92	3.69 ± 4.14	0.98	11.8–49.3
Ptychobranchus fasciolaris	Kidneyshell	1000	0	0	18	40	5.39×10^{-5}	$1.62\times10^{-5} \pm 1.64\times10^{-4}$	3.15	2.87 ± 3.42	0.93	49.4–97.8
Ptychobranchus occidentalis	Ouachita kidneyshell	981	16	3	34	40	2.49×10^{-4}	$4.92\times10^{-5} \pm 1.55\times10^{-3}$	2.74	2.24 ± 3.23	0.76	67.8–104.5
Villosa nebulosa	Alabama rainbow mussel	961	25	14	103	30	1.16×10^{-4}	$8.39\times10^{-5} \pm 1.81\times10^{-4}$	2.81	2.70 ± 2.92	0.99	8.5–45.0
Pleurobemini		1000	0	0	392	230	1.55×10^{-3}	$6.00\times10^{-4} \pm 4.36\times10^{-3}$	2.32	2.09 ± 2.56	0.63	8.3–103.6
Elliptio arca	Alabama spike	1000	0	0	54	30	1.86×10^{-5}	$1.13\times10^{-5} \pm 3.07\times10^{-5}$	3.36	3.22 ± 3.50	0.99	8.3–89.3
Elliptio complanata	Eastern elliptio	988	12	0	154	40	5.89×10^{-6}	$1.79\times10^{-6} \pm 1.55\times10^{-5}$	3.40	3.12 ± 3.68	0.94	43.6–83.4
Elliptio dilatata	Spike	995	1	4	30	40	1.09×10^{-4}	$2.89\times10^{-5} \pm 4.50\times10^{-4}$	2.92	2.68 ± 3.16	0.94	46.5–103.6
Fusconaia cerina	Gulf pigtoe	733	145	122	22	40	3.32×10^{-3}	$1.86\times10^{-3} \pm 6.52\times10^{-3}$	2.39	2.19 ± 2.59	0.94	23.7–70.0
Fusconaia flava	Wabash pigtoe	968	28	4	95	40	2.81×10^{-5}	$9.79\times10^{-6} \pm 9.97\times10^{-5}$	3.35	3.06 ± 3.63	0.93	23.7–72.3
Pleurobema decisum	Southern clubshell	1000	0	0	37	40	1.24×10^{-3}	$8.50\times10^{-4} \pm 1.78\times10^{-3}$	2.54	2.47 ± 2.62	0.99	15.6–67.0
Quadrulini		1000	0	0	283	160	3.33×10^{-3}	$2.21\times10^{-3} \pm 5.09\times10^{-3}$	2.35	2.24 ± 2.46	0.92	5.2–119.6
Cyclonaias asperata	Alabama orb	1000	0	0	96	40	1.28×10^{-3}	$7.73\times10^{-4} \pm 2.34\times10^{-3}$	2.67	2.53 ± 2.81	0.97	24.3–59.4
Cyclonaias pustulosa	Pimpleback	1000	0	0	68	40	4.31×10^{-4}	$3.81\times10^{-4} \pm 6.71\times10^{-4}$	2.96	2.87 ± 3.05	0.99	5.2–68.8
Quadrula rumphiana	Ridged mapleleaf	984	12	4	65	40	1.17×10^{-3}	$4.59\times10^{-4} \pm 3.68\times10^{-3}$	2.65	2.45 ± 2.85	0.95	33.9–77.8
Tritogonia verrucosa	Pistolgrip	1000	0	0	54	40	1.73×10^{-3}	$4.07\times10^{-4} \pm 5.24\times10^{-3}$	2.39	2.12 ± 2.66	0.89	52.5–119.6

estimates when scaling from individuals to ecosystems. This case study compares sites dominated by lower-biomass, higher-diversity beds (Sipsey River in Alabama, USA) with higher-biomass, lower-diversity beds in the Kiamichi and Little Rivers of Oklahoma, USA. Our analysis and dataset provide the most comprehensive set of tools available to researchers and managers for estimating unionid mussel biomass.

METHODS

Data sources and dataset

In this analysis, we used published and unpublished morphometric data from 6694 observations of field-collected unionid mussels spanning 30 y (1988–2018) and 10 first authors (Tables S1, S2). We also recorded habitat type, waterbody, and general geographic region for each observation. Our analysis primarily focused on the relationships between maximum axial shell length (hereafter length or L_{max}; Fig. 1A, B) and either STDM or SHDM. We report the relationships between length and total wet mass in Table S3.

Modeling and statistical analyses

All statistical analyses were done in R (version 3.5.1; R Project for Statistical Computing, Vienna, Austria). We used regression approaches to estimate the scaling relationships between length–STDM and length–SHDM at 3 taxonomic levels: family (i.e., all species combined), phylogenetic tribe, and species. We estimated scaling coefficients between untransformed length and mass and log-transformed length and log-transformed mass (i.e., STDM and SHDM) with both linear regression (R function lm; Eq. 1) and non-linear regression (R function nls; Eq. 2). We evaluated the appropriateness of the error structure of each model for our data (i.e., whether the error structure is additive vs multiplicative) with code provided in Xiao et al. (2011).

We used stratified Monte Carlo simulations (bootstrap resampling) to randomly generate length and STDM or SHDM datasets from our collected data from which we estimated slope (b) and intercept (a) regression parameters. We used a stratified approach to ensure that parameters were estimated from an equal number of observations among species (tribe- and family-level regressions) and that observations were evenly distributed across the range of lengths observed within a species. We stratified the data by subdividing the length observations for each species into 8 equally-spaced bins (R function cut), resampling 5 observations from each bin to produce a length and mass dataset with $n = \sim 40$ (R function sample), and fitting regression models (R function lm). This stratified resampling approach minimizes the bias that could be introduced to parameter estimates by species with disproportionately-high observation numbers or uneven distributions of observations across the size spectrum. Some of the size bins may have been empty for species with clumped or sparse distributions along their empirical size

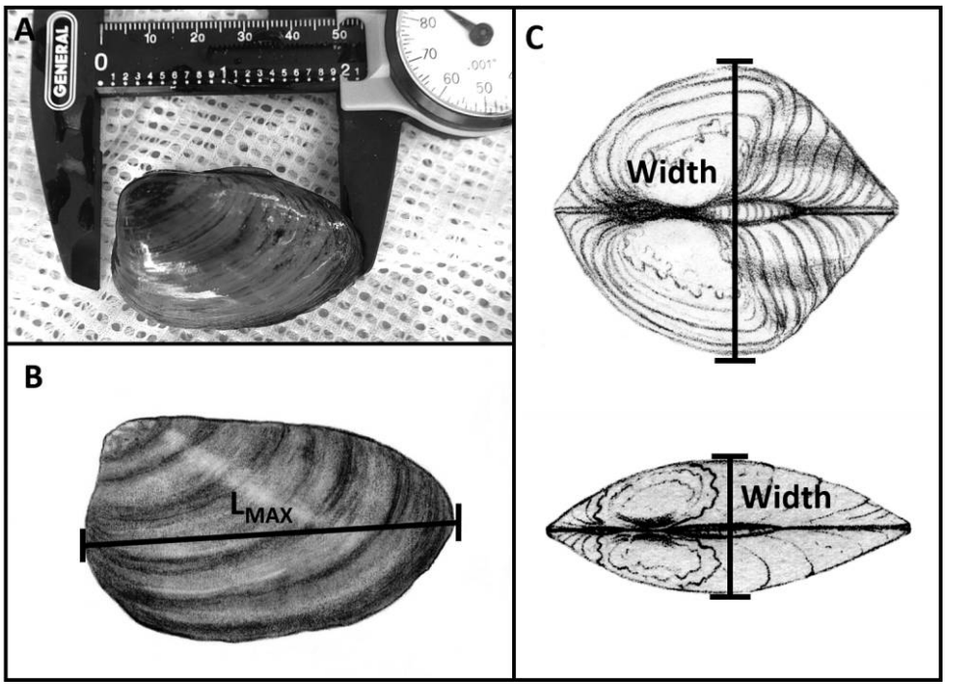

Figure 1. A.—Maximum axial shell length (L_{max}) measured in the field with a caliper (Photo credit: CLA). B.—Maximum axial shell length is measured as the greatest distance between 2 points of the shell from a lateral exterior view (Illustration: K. L. Lambert). C.—Maximum axial shell width is measured as the greatest distance between 2 points of the shell from a dorsal exterior view (Illustration: I. Lea 1931, 1942). The top individual represents a greater inflation of shell morphology compared to the more compressed morphology of the bottom individual, with inflation defined as a greater ratio of total width:total length (Ortmann 1920).

Table 3. Length–mass equations (STDM and SHDM) for *Amblema plicata* (STDM = $a\ L_{max}^{b}$, where STDM is soft tissue dry mass [g], SHDM is shell dry mass [g], L_{max} is the maximum axial shell length [mm], and a and b are constants) and *Elliptio complanata* (STDM) across multiple systems. All a and b constants are based on a log–log linear regression with multiplicative error structure. Bootstrap resampling (1000×) was used to generate data for each species that was evenly distributed across the size ranges observed. n is the population from which n_{reg} is resampled. The n_{reg} represents the number of individual measurements that went into generating the regression, and we report the size range of individuals measured and the 95% confidence intervals (CI) around the mean a and b coefficients. The range is the range of length values available for each taxonomic level.

Waterbody	n	n_{reg}	a	a 95% CI	b	b 95% CI	r^2	Range (mm)
Amblema plicata								
STDM								
Kiamichi–Little Basin	32	40	2.04×10^{-5}	$1.01\times10^{-6} \pm 4.12\times10^{-5}$	2.64	2.48 ± 2.82	0.87	50–105
Tombigbee Basin	94	40	9.80×10^{-6}	$5.70\times10^{-6} \pm 1.69\times10^{-5}$	2.81	2.69 ± 2.93	0.93	54.2–104.6
Upper Mississippi River Basin	500	40	6.04×10^{-6}	$3.06\times10^{-6} \pm 1.20\times10^{-5}$	2.98	2.82 ± 3.13	0.90	50.26–104.82
SHDM								
Kiamichi–Little Basin	32	40	9.83×10^{-4}	$2.26\times10^{-4} \pm 4.28\times10^{-3}$	2.62	2.29 ± 2.96	0.87	50–105
Tombigbee Basin	91	40	6.06×10^{-4}	$1.72\times10^{-4} \pm 2.14\times10^{-3}$	2.75	2.46 ± 3.04	0.90	54.2–104.6
Upper Mississippi River Basin	475	40	1.93×10^{-3}	$8.02\times10^{-4} \pm 4.62\times10^{-3}$	2.47	2.27 ± 2.68	0.93	50.26–104.82
Elliptio complanata								
STDM								
Hudson River, New York	2641	40	5.29×10^{-6}	$6.61\times10^{-7} \pm 4.23\times10^{-5}$	2.86	2.36 ± 3.36	0.76	43.6–83.4
Lake Opeongo, New York	154	40	4.66×10^{-6}	$1.30\times10^{-6} \pm 1.67\times10^{-5}$	2.81	2.50 ± 3.12	0.88	43.6–83.4
Neversink River, New York	427	40	1.45×10^{-5}	$3.55\times10^{-6} \pm 5.93\times10^{-5}$	2.65	2.31 ± 2.99	0.85	43.6–83.0

spectra, which resulted in $n < 40$ (See n_{reg}, in Tables 1–3). This resampling process produced $n = 40$ for 50% of the species and $n = 30$ or 35 observations for the other 50%. We report the average coefficients and fit (i.e., r^2) from 1000 resamplings. Finally, to provide a method for converting between ash-free dry mass (AFDM) and STDM (2 common biomass measurement units), we used our stratified Monte Carlo approach to develop a family-level STDM–AFDM equation (e.g., Benke et al. 1999).

We also used this stratified Monte Carlo technique to test for spatial variation in length–mass relationships for 2 species (*Amblema plicata* and *Elliptio complanata*) that are widely-distributed across 3 hydrologically-disconnected basins. For these analyses, we selected observations with the same range of shell lengths from each of the 3 basins (50–105 mm for *A. plicata* and 43.6–83.4 mm for *E. complanata*). We then reported the estimated length–STDM and length–SHDM coefficients across the 3 basins, compared them visually, and examined overlap using confidence intervals.

We assessed model fit for the different taxonomic levels by inspecting residual plots for positive or negative bias. We used a non-linear smooth function to highlight the patterns in the residual plots and used generalized additive models (GAMs) for their ability to fit nonlinear data. We selected 3 as the minimum number of knots required to adequately capture the patterns in the residuals without overfitting any of the species, therefore, all GAMs were fit using 3 knots. Better fitting models had lower relative error (coefficient of variation [CV], %) and minimized bias. We also inspected mean absolute errors and CVs for high and low biomass individuals.

Case study

An important application of these equations is scaling biomass-linked organismal functions (e.g., biofiltration or nutrient excretion rates) from individuals to ecosystems. Management planning can suffer when errors in biomass estimates occur, just as it can suffer from errors in population variability estimates (Strayer and Smith 2003). To demonstrate how the use of equations developed at different levels of taxonomic resolution may affect biomass estimates at the assemblage level, we used field survey data (i.e., quadrat sampling) from 19 sites in 2 distinct systems: the Sipsey River in Alabama (12 sites) and the Kiamichi and Little River system of the middle Red River Basin in Oklahoma (7 sites). The Sipsey River has a diverse unionid mussel assemblage (43 species) of smaller-bodied species in which local assemblages (individual beds) tend to have high evenness and richness (Ward et al. 2005, Haag and Warren 2010). The Kiamichi River (31 species) and Little River (35 species) are also diverse, but individual mussel beds tend to be dominated by a few larger-bodied species (Vaughn et al. 1996, Matthews et al. 2005, Galbraith et al. 2008, Spooner and Vaughn 2009). We used these data to estimate the areal STDM and SHDM at the family, tribe, and lowest available taxonomic level (species-specific) levels. In most cases, species was the lowest level of information available, but we lacked species-specific regressions for some rare taxa. In those cases, we

used a tribe-level regression in place of the species-specific regression.

RESULTS

Error structure and coefficients

A multiplicative error structure (Eq. 1; log–log linear regression) provided the best fit for our STDM and SHDM datasets in most cases, so our results are based on those models (Tables 1, 2). Multiplicative error structures provided the best fit for length–STDM scaling relationships at the family level (Unionidae; Fig. 2A), tribe level (Fig. 2B), and for 85% of species (Table 1). We found that STDM = $9.48 \times 10^{-6} \times L_{max}^{2.82}$ at the family level. At the tribe level, a averaged $1.41 \times 10^{-5} \pm 1.43 \times 10^{-5}$ and b averaged 2.80 ± 0.15. At the species level, a averaged $7.70 \times 10^{-6} \pm 9.56 \times 10^{-6}$ and b averaged 3.08 ± 0.36 (excluding *Truncilla truncata*; Table 1). Multiplicative error structures also provided the best fit for length–SHDM scaling relationships at the family, tribe, and species levels (Table 2). We found that SHDM = $2.98 \times 10^{-4} \times L_{max}^{2.79}$ at the family level (Fig. 3A). At the tribe level for SHDM (Fig. 3B), mean a and b were $1.32 \times 10^{-3} \pm 1.49 \times 10^{-3}$ and 2.67 ± 0.40, respectively. At the species level, mean a and b were $4.89 \times 10^{-4} \pm 8.35 \times 10^{-4}$ and 3.09 ± 0.45, respectively (Table 2).

Length–STDM model evaluation

At the family level, length–STDM regressions used 5463 individuals. For the 6 phylogenetic tribes with available data, n ranged from 269 for Quadrulini to 3502 for Pleurobemini. For the 28 species, n ranged from 11 for *Lampsilis teres* to 3222 for *E. complanata*. We found that soft tissue AFDM was an average of $79.6 \pm 0.33\%$ of STDM in mussels and was well predicted by the equation AFDM = $0.81 \times \text{STDM}^{1.01}$ ($n = 585$, $r^2 = 0.99$, $p < 0.0001$).

Length explained most of the variation in STDM across taxonomic levels (average $r^2 = 0.93$; Table 1). The model fit most of the species-specific regressions well, and in 20 of the 26 species, $r^2 \geq 0.90$. Length–STDM equations at finer taxonomic resolution improved and reduced residual variation relative to coarser-resolution models (Fig. 4A–C). Across the 3 taxonomic levels considered, several species exhibited a tendency to over- or underestimate the STDM (Fig. 4A–C). Among small-bodied individuals (length <20 mm, ~5 mg STDM), the average CV and the absolute value of residual mass error was 32% (2.2 mg) for the species-specific model, 74% (4.3 mg) for the tribe-level model, and 111% (6.0 mg) for the family model. All model levels fit large-bodied (length >90 mm, ~3 g) individuals well, and the average CV absolute value of the residual mass error was 20.1% (0.89 g) for the species-specific model, 24.1% (0.97 g) for the tribe-level model, and 27.3% (1.16 g) for the family-level model. Additionally, species-specific regressions reduced the presence or magnitude of bias in predictions (Fig. 4A–C). The average absolute value of the residual mass error for species-specific length–STDM models was 41% lower than the family-level model and 34% lower than the tribe-level model.

Taxonomic variation in length–STDM relationships

At the tribe level (Fig. 2B), all STDM equations exhibited r^2 values ≥0.89 (Table 1). Model fit for most species was

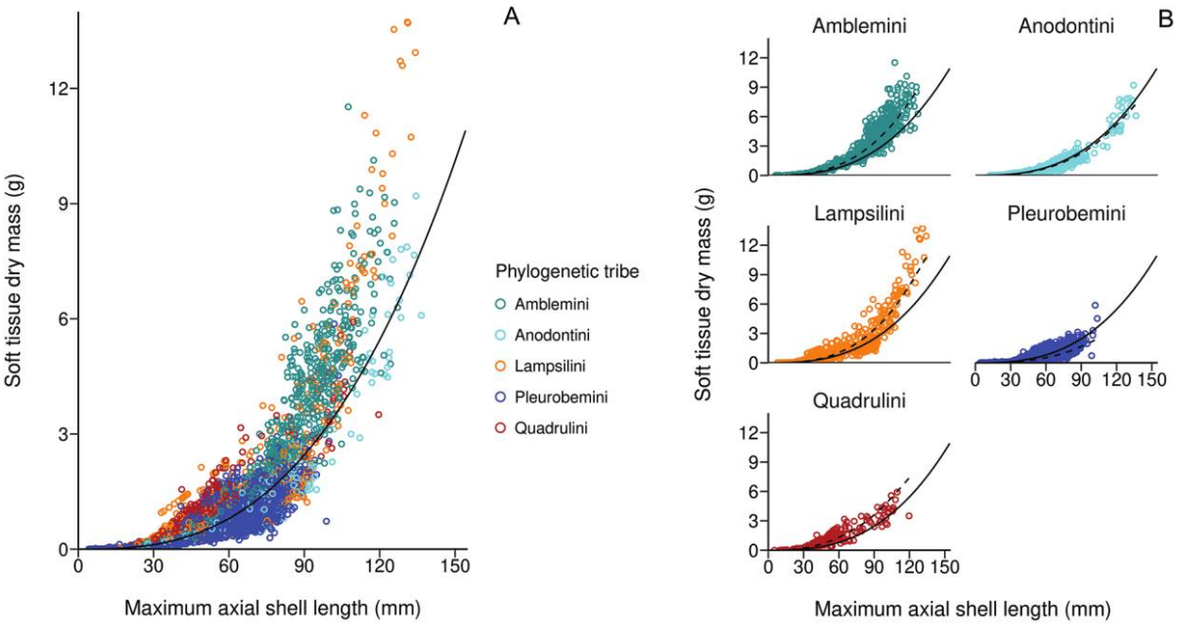

Figure 2. Soft tissue dry mass as a function of maximum axial shell length. A.— Family-level log–;log regression for the combined 5 unionid phylogenetic tribes ($n = 5802$). B.— Log–log regressions (dotted line) for individual phylogenetic tribes (Amblemini [$n = 626$], Anodontini [533], Lampsilini [743], Pleurobemini [3520], and Quadrulini [290]) compared with the family-level regression (solid line).

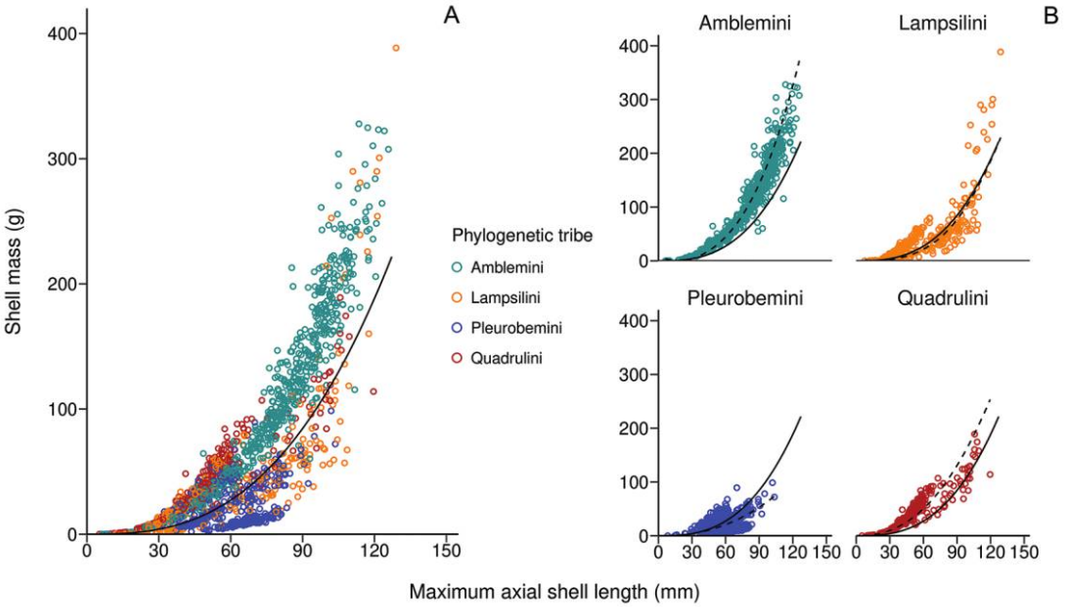

Figure 3. Shell mass as a function of maximum axial shell length. A.—Family-level log–log regression for the combined 4 unionid phylogenetic tribes (n = 4126). B.—Log–log regressions (dotted line) for individual phylogenetic tribes (Amblemini [n = 598], Lampsilini [622], Pleurobemini [447], and Quadrulini [283]) compared with the family-level regression (solid line).

also good, and 21 of the 27 species had r^2 values ≥0.90. The intercept (a) and slope (b) coefficients that described the relationship between length and STDM differed between some of our phylogenetic tribes because the 95% confidence intervals did not overlap for some comparisons (Table 1). In particular, slope (b) for Quadrulini did not overlap with any of the other tribes. Also, Quadrulini's intercept (a) was an order of magnitude greater than all other tribes except Amblemini. Species within the same phylogenetic tribe also had high CVs, and confidence intervals did not overlap among all species within tribes.

Length–SHDM model evaluation

Family-level length–SHDM regressions used 1973 individuals. For the 5 phylogenetic tribes with available data, n ranged from 23 for Anodontini to 622 for Lampsilini. For the 28 species, n ranged from 11 for *L. teres* to 598 for *A. plicata*.

Similar to the length–STDM regression coefficients, finer taxonomic resolution length–SHDM equations improved model fit relative to coarser-resolution models (Fig. 5A–C). Across the 3 taxonomic levels considered, several species exhibited a tendency to over- or underestimate the SHDM of large-bodied individuals, especially *A. plicata* (Fig. 5C). Among small-bodied individuals (length <20 mm, ~472 mg SHDM), the average CV and the absolute value of residual mass error was 18% (0.88 mg) for the species-specific model, 78% (212 mg) for the tribe-level model, and 127% (369.1 mg) for the family-level model. On average, species-specific and tribe-level models fit the large-bodied (length >90 mm, ~172 g) individuals better than family-level models. The absolute value of the residual mass error was an average of 16.3% (25.7 g) for the species-specific model, 21.9% (30.1 g) for the tribe-level model, and 35.2% (30.1 g) for the family-level model. Broadly, the use of species-specific regressions reduced the presence or magnitude of this bias in the prediction of SHDM from length (Fig. 5A–C). Within species the average absolute value of the residual mass error for species-specific length-SHDM models was 60% lower than the family-level model and 51% lower than the tribe-level model. The tribe-level STDM scaled similarly with length, but the SHDM-scaling relationships were generally more tribe- and species-specific (Table 2, Fig. 3A, B).

Taxonomic variation length–SHDM relationships

At the tribe level (Fig. 3B), Pleurobemini, which exhibit diverse shell morphologies, had the lowest r^2 of 0.66 (n = 190). The remaining tribes all had r^2 values ≥0.89 (Table 2). Most species also had models that fit well, and 18 of the 21 species had r^2 values of 0.93 or higher. The a and b coefficients that described the relationship between length and SHDM differed between some of the phylogenetic tribes because their 95% confidence intervals did not overlap (Table 2). In particular, b was greater for Lampsilini than for any other tribe. Also, Lampsilini's a coefficient was an order of magnitude smaller than any other tribe. Species in

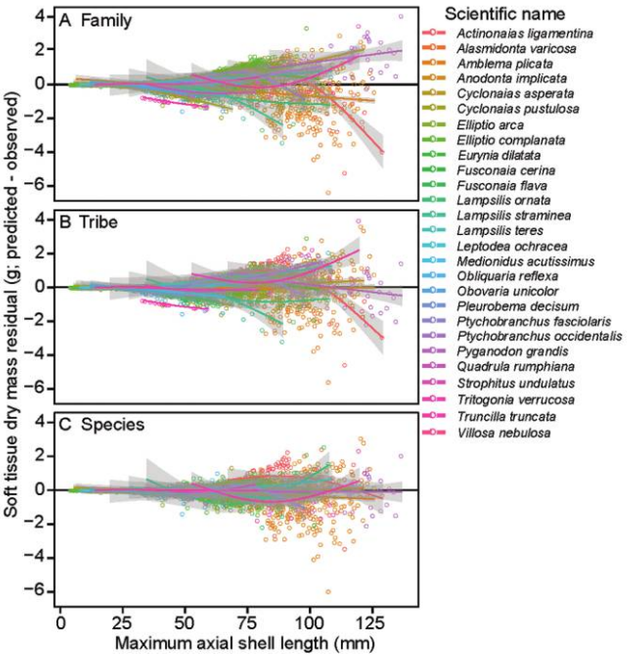

Figure 4. Residual for soft tissue dry mass estimation for family- (A), tribe- (B), and species- (C) level regression equations. Lines are generalized additive model smoothers (knots = 3) that trend or pattern the mean of the residuals across the range of lengths for species in our dataset. Gray bands represent 95% confidence intervals around the generalized additive model smoothed lines.

the Quadrulini tribe had the highest intercept coefficient of any tribe. As with STDM, species within tribes had high variation in their *a* and *b* coefficients, and SHDM confidence intervals did not always overlap. (e.g., *Fusconaia cerina* and *F. flava*).

Geographic variation in scaling relationships

Amblema plicata and *E. complanata* length–mass scaling coefficients varied among 3 hydrologically-disconnected basins. We compared *A. plicata* individuals ranging from 50 to 105 mm across all sites reporting data. *Amblema plicata* length–STDM regression was $2.04 \times 10^{-5} \times L_{max}^{2.64}$ ($n = 32$, $r^2 = 0.87$) for the Kiamichi–Little River basin, $9.80 \times 10^{-6} \times L_{max}^{2.81}$ ($n = 94$, $r^2 = 0.93$) for the Tombigbee basin, and $6.04 \times 10^{-6} \times L_{max}^{2.98}$ ($n = 500$, $r^2 = 0.90$; Fig. 6A) for the Upper Mississippi River basin. Length–SHDM was $9.83 \times 10^{-4} \times L_{max}^{2.62}$ ($n = 32$, $r^2 = 0.87$) for the Kiamichi–Little River basin, $6.06 \times 10^{-4} \times L_{max}^{2.81}$ ($n = 91$, $r^2 = 0.93$) for the Tombigbee basin, and $1.93 \times 10^{-3} \times L_{max}^{2.47}$ ($n = 40$, $r^2 = 0.90$; Fig. 6A) for the Upper Mississippi River basin (see Table 3). We restricted the size range for *E. complanata* to between 43.6 and 83.4 mm for all sites. The length–STDM regression was $5.29 \times 10^{-6} \times L_{max}^{2.86}$ ($n = 2641$, $r^2 = 0.76$) for the Hudson River basin, $4.66 \times 10^{-6} \times L_{max}^{2.81}$ ($n = 154$, $r^2 = 0.88$) for Lake Opeongo, and $1.45 \times 10^{-5} \times L_{max}^{2.65}$ ($n = 427$, $r^2 = 0.85$; Fig. 6B) for the Neversink River basin. SHDM data were not available for *E. complanata*.

Case study

Our model evaluations indicated that species-level models provided the most accurate estimates of STDM and SHDM biomass (Figs 4–5), so we assumed that species-level models represent the most accurate estimates of assemblage biomass. In both the Sipsey River and Kiamichi–Little River systems, models with higher taxonomic resolution produced higher estimates of mussel assemblage areal STDM (g STDM/m^2). In the Kiamichi–Little River system, which is dominated by larger-bodied mussels, average areal soft tissue biomass was 104 ± 71 g/m^2 based on species-level models, which was 26 ± 6% higher than family-level estimates (82 ± 56 g/m^2) and 14 ± 5% higher than tribe-level estimates (90 ± 62 g/m^2). In the Sipsey River, which is dominated by smaller-bodied species, the effect of classification level of the regressions was less apparent. Average areal soft tissue biomass was 11 ± 6 g/m^2 with species-level models, which was 12 ± 12% higher than family-level estimates (10 ± 5 g/m^2) and 8 ± 9% higher than tribe-level estimates (10.5 ± 5.5 g/m^2).

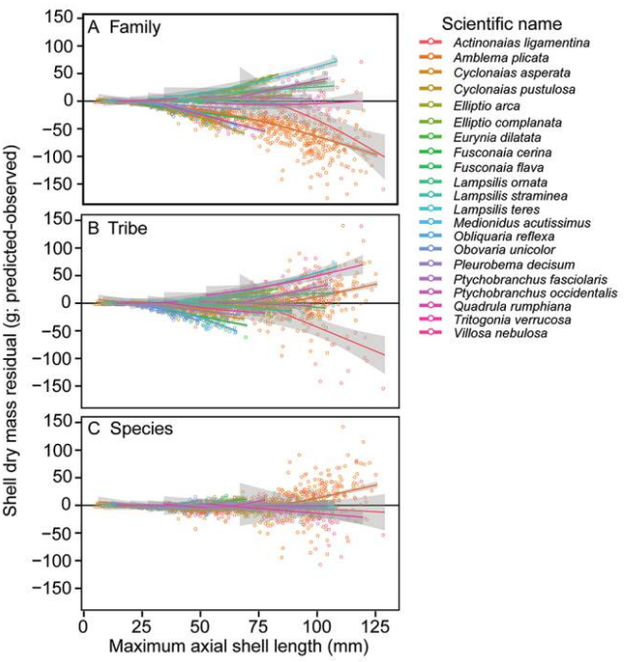

Figure 5. Residual for shell dry mass estimation for family- (A), tribe- (B), and species- (C) level regression equations. Lines are generalized additive model smoothers (knots = 3) that trend or pattern the mean of the residuals across the range of lengths for species in our dataset. Gray bands represent 95% confidence intervals around the generalized additive model smoothed lines.

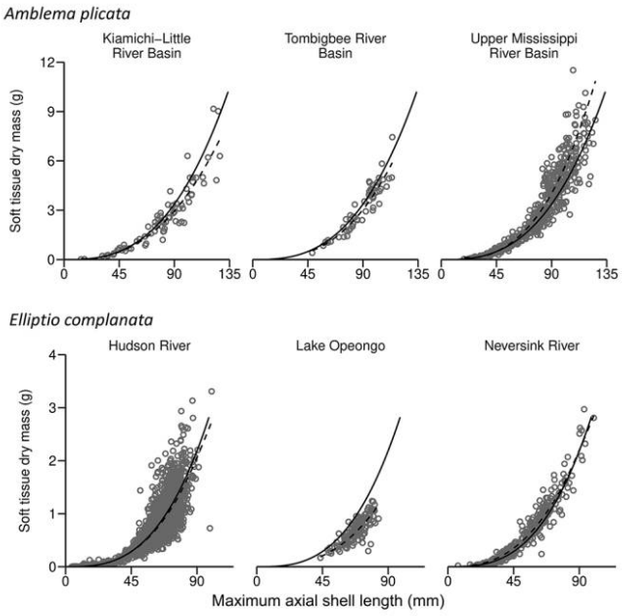

Figure 6. Log–log regressions of soft tissue dry mass as a function of maximum axial shell length for 2 widely-distributed mussel species. A.—*Amblema plicata* ($n = 626$) within 3 basins: Kiamichi–Little River Basin, Oklahoma, USA (32); Tombigbee Basin, Alabama, USA (91); and the Upper Mississippi River Basin, Minnesota/Wisconsin, USA (475). B.—*Elliptio complanata* (3222) within 3 water bodies: Hudson River, New York, USA (2641); Lake Opeongo, Ontario, Canada (154); and the Neversink River, New York (427).

Species-specific equations also predicted SHDM with higher accuracy than did tribe-level and family-level equations. In the Kiamichi–Little River basin, species-level models increased the estimated areal SHDM (3207 g/m^2) 45 ± 15% over the family-level model (2246 g/m^2) and 17 ± 8% over tribe-level models (2685 g/m^2). In the Sipsey River, species-level models increased the estimated areal SHDM (441 g/m^2) by 57 ± 15% over the family-level model (282 g/m^2) and 43 ± 13% for tribe-level models (308 g/m^2).

DISCUSSION

Soft tissue biomass and shell biomass are key parameters in estimating mussel contributions to the ecosystem services of water filtration, nutrient cycling, and nutrient storage. Our analysis showed that natural variation in the length–mass relationship led to a multiplicative error structure in most of the species we assessed. Therefore, the magnitude of variation (and uncertainty) in the absolute contributions of an individual and a community to ecosystem functions will increase as assemblage composition or population size structure shifts toward larger-bodied individuals. Family-level length–STDM equations consistently underpredicted the areal soft tissue biomass of an assemblage to within 12 to 26% of the biomass predicted by species-level equations. Tribe-level length–STDM equations estimated biomass to within 9 to 14% of the true biomass. Species-specific equations provided the most accurate estimates of total areal soft and shell tissue biomass. Our analysis of unionid length–mass relationships will enable researchers and managers to accurately estimate individual STDM, AFDM, and SHDM at the family, tribe, or species level and scale these estimates to areal densities. These tools are essential for estimating critical ecosystem services, such as water filtration, as well as for developing functional biodiversity concepts in community ecology and integrating animal contributions to biogeochemical cycles.

Our mass–length intercept patterns varied among species and between length–STDM and length–SHDM. Intercept parameters, not surprisingly, differed by 2 orders of magnitude between shell (2.98×10^{-4}) and soft tissue (9.48×10^{-6}) as shell is much heavier than soft tissue. However, the relative variances were similar ($CV_{SHDM} = 154\%$, $CV_{STDM} = 124\%$). The intercept of the length–shell biomass relationship with Pleurobemini and Quadrulini was greater than the intercept of the other tribes. The overall relationship, not surprisingly, overestimates shell biomass for thin-shelled species (e.g., Lampsilini) and underestimates shell biomass in heavier-shelled species (e.g., Amblemini). Therefore, it is important to use species- or tribe-specific relationships to provide accurate estimates of soft tissue and, especially, shell biomass.

Species had variable mass–length relationships, but the average b coefficients at the family level were similar for both shell ($b = 3.03 \pm 0.47$) and soft tissue (2.84 ± 0.36) biomass. In invertebrates, length–mass scaling relationships of powers (b) of 3 indicate that body shape and specific gravity remain constant as individuals mature (Benke et al. 1999). In insects, $b < 3$ suggests that insects either become proportionately narrower or their specific gravity declines with increasing length (Meyer 1989, Burgherr and Meyer 1997, Benke et al. 1999). In contrast, $b > 3$ indicates that organisms become proportionately wider or their specific gravity increases as length increases (Towers et al. 1994, Benke et al. 1999). At the family level, the $b < 3$ for the relationship between both STDM (2.82; Table 1) and SHDM (2.79; Table 2), and their confidence intervals did not include 3 (Table 1). This result suggests that as mussel taxa become older or larger, their adaptive response to flow is to become narrower or heavier to maintain position in the sediment.

Challenges in predicting biomass

Accurately estimating biomass is important for scaling to system-wide ecological functions in rivers over time (decadal time scale) and space (Atkinson et al. 2014b, 2018, Atkinson and Vaughn 2015, Vaughn et al. 2015, Benke and Huryn 2017). Part of the difficulty in estimating biomass, particularly shell biomass, results from the natural interspecific and intraspecific variability in shell morphologies (e.g., shape, sculpturing, and thickness). Interspecific variability

may arise from evolutionary adaptation to environmental conditions. Certain shell shapes and sculpturing are thought to aid in burrowing and reduce dislodgement during extreme hydrologic events (Stanley 1981, Watters 1994). For example, compressed and elongated shells allow for more efficient burrowing by reducing drag in the sediment compared to inflated or spherical shell morphologies (Stanley 1988). In addition, shell sculpturing (e.g., tubercles, ridges) is thought to assist anchoring in the sediment by increasing friction as well as reducing scour by decreasing turbulent flow around the shell (Watters 1994). Further, lighter-shelled species tend to be found in lentic or soft-sediment environments, which presumably enhance buoyancy and prevent sinking into anoxic sediments (Watters 1994).

Intraspecific differences in shell morphologies may also affect the prediction accuracy for a given species within and across systems. In lentic systems, shell length and thickness can vary with sediment characteristics and water chemistry within the same waterbody (Ortmann 1920, Hinch et al. 1989, Haag 2012). In lotic systems, Ortmann's (1920) law of stream position describes the characteristic change in a mussel species' shell morphology along a river continuum, with flat and compressed shells in the headwaters that become progressively more inflated as river order increases (Fig. 1C). These intraspecific differences that arise because of location within the river continuum, water chemistry, and sediment characteristics can cause prediction accuracy to vary with geographic location. Shell forms and thickness can also vary within the same stream system in response to water velocity, and animals located in high-current velocities often have greater shell erosion (resulting in lower shell mass) than those in lower-velocity habitats (Ortmann 1920, Hornbach et al. 2010). For example, our models consistently overpredicted *A. plicata* SHDM in larger-bodied individuals (Fig. 6A). In the field, we frequently observe heavier erosion of *A. plicata* than most other species in the same habitat (CLA and TBP, personal observations). This trend may be because *A. plicata* are generally long lived (Sansom et al. 2016), so larger individuals are older and their shells have eroded over time. Additionally, *A. plicata* has been documented to move less than other species (Allen and Vaughn 2009), so they may be more likely to undergo shell erosion because they may not burrow in response to high flows. For species inhabiting both lotic and lentic environments (e.g., *Actinonaias ligamentina*, *Pyganadon grandis*) there may be greater uncertainty in estimates of their biomass because of the greater intraspecific variation in shell morphology.

Finally, some tribes of freshwater mussels exhibit sexual dimorphism. For example, Lampsilini females are typically more inflated (i.e., thicker dorsally) than males of the same length (Haag 2012). Our analyses did not differentiate females and males. Thus, when sexual dimorphism is present, our estimates will represent an average that underestimates one sex and overestimates the other. This imbalance can introduce error into assemblage-level biomass estimates because sex ratios can vary widely across populations (Garner et al. 1999, Haag and Staton 2003, Galbraith and Vaughn 2011). However, sexual dimorphism can be difficult to recognize and therefore difficult to account for in models. Our analysis consequently reduces the error that can arise from the misidentification of a mussel's sex, and many of our sexually-dimorphic taxa (Lampsilini tribe) had the highest predictive power (i.e., high r^2).

From individual length to assemblage biomass

Models based on finer taxonomic resolution improved the accuracy of soft tissue and shell biomass predictions at the assemblage and individual levels in both the Sipsey River and Kiamichi–Little Rivers. The effect appears to be most significant when assemblages are composed of large-bodied individuals. The Sipsey system was composed of smaller-bodied animals with lower areal biomass (15 g/m^2) and dominated by species in the Pleurobemini and Quadrulini tribes, whereas the Kiamichi–Little system was dominated by larger individuals with higher areal biomass (100 g/m^2), mostly from the Amblemini and Lampsilini tribes. STDM field-survey data showed that substituting family- or tribe-level equations for species-level equations likely underestimates STDM relative to the true biomass present in a system. Previous work in insects suggested that family-level length–mass equations were sufficient for predicting individual insect biomass in the absence of species-level equations (Benke et al. 1999). Our results suggest that family-level models of mussel biomass may introduce systematic biases that could cause biomass to be underestimated at the assemblage level, even though these models fit the data well. Thus, we advise the use of species-level equations whenever possible for both soft tissue and shell biomass. However, when taxonomic certainty is lacking, the family- or tribe-level equations can provide a reasonably-accurate estimate, particularly when estimating STDM for smaller-bodied individuals.

From biomass to ecosystem function

Accurate biomass estimates are essential for estimating nutrient storage and rates of ecological function such as water filtration and nutrient recycling (e.g., Atkinson and Vaughn 2015). Errors in shell or soft tissue biomass are directly proportional to errors in estimates of mussel nutrient storage in those 2 compartments, such that a 30% error in SHDM is a 30% error in an estimate of nitrogen or phosphorous storage (Atkinson and Vaughn 2015, Vaughn et al. 2015). In contrast, this error scales nonlinearly for estimating metabolic processes like excretion. Scaling to metabolically-linked functions nests mass scaling errors (mass $= a \times L_{\max}{}^{b} \times e^{\varepsilon\text{mass}}$) within the function describing metabolic scaling (metabolism $= c \times \text{mass}^{d} \times e^{\varepsilon\text{metabolism}}$) such that error in length–mass estimates scale as $e^{(d \times \varepsilon\text{mass} + \varepsilon\text{excretion})}$ (a and c are

nomalization constants for metabolic scaling). Therefore, in practice, overestimating mass in a small-bodied individual would underestimate the mass-specific rates of function because mass-specific rates of metabolism (and associated functions) decline with increasing body size (Wen and Peters 1994, Hall et al. 2007). Conversely, underestimating mass in a large-bodied individual would overestimate mass-specific rates of function. For the median dense aggregations (20 mussels/m^2) in the Sipsey and Kiamichi–Little Rivers, a single large-bodied individual can account for 40 to 50% of the total biomass and 28 to 33% of the total nitrogen (as ammonium) excretion.

Toolkit for estimating biomass

To facilitate the use of these models by other investigators, we provide a simple R script (Appendix S4) that uses the equation coefficients we report here to calculate STDM and SHDM at the desired level of taxonomic resolution. This script will output estimates for STDM, soft tissue AFDM, or SHDM as well as their 95% confidence intervals when run in R. A tutorial (Appendix S5) with example input data files (Appendix S6) are provided to guide the user through the application of the script. We hope that this resource is useful for both researchers and managers in making accurate assessments of biomass of this highly-imperiled taxonomic group.

Implications for management

Most legal frameworks that support management and monitoring activities call for the protection of ecosystem structure and function, and most monitoring programs rely on structural indicators of ecosystem integrity (Carter and Resh 2013). These biotic structural indicators can be converted to functional indicators of the ecosystem by integrating biomass estimates and functional traits of foundational taxa (e.g., Atkinson et al. 2018, Hopper et al. 2018). Indeed, biomass measurement and prediction are central to our emerging understanding of the role of animal functional traits in ecosystem functions like elemental cycling (Vanni and McIntyre 2016, Atkinson et al. 2017). The length–mass relationships presented in this paper are intended to aid researchers and agencies in the determination of potential ecosystem services provided by unionid mussels based on non-destructive survey methods, an important aspect of species conservation and ecosystem management (Southwick and Loftus 2017, Strayer 2017). For example, the National Science Foundation-funded National Ecological Observatory Network, a next-generation ecological research program intended to inform fundamental research and policy, is collecting length data for monitored animal species. These datasets, once converted to biomass, will allow cross-system analysis of animal contributions to ecosystem function as well as provide a scientific basis for concomitant monitoring of animal assemblage structure and function as derived from biomass.

ACKNOWLEDGEMENTS

Author contributions: CLA and WRH conceived the manuscript. TBP wrote the R script for the management toolkit. BVE assisted with the initial data collection, and all authors shared equally in further collection and analysis of data and the writing of the manuscript.

We appreciate the sharing of data for this manuscript from all contributing researchers. We appreciate helpful reviews of previous versions of this manuscript by Art Benke, Garrett Hopper, and Caryn Vaughn. BVE was supported by grants from United States Fish and Wildlife Service and the Weyerhaeuser Corporation. TBP was supported by National Science Foundation DEB-1457542. CLA was supported by National Science Foundation DEB-1831512.

LITERATURE CITED

Allen, D. C., and C. C. Vaughn. 2009. Burrowing behavior of freshwater mussels in experimentally manipulated communities. Journal of the North American Benthological Society 28:93–100.

Atkinson, C. L., K. A. Capps, A. T. Rugenski, and M. J. Vanni. 2017. Consumer-driven nutrient dynamics in freshwater ecosystems: From individuals to ecosystems. Biological Reviews 92:2003–2023.

Atkinson, C. L., J. P. Julian, and C. C. Vaughn. 2014a. Species and function lost: Role of drought in structuring stream communities. Biological Conservation 176:30–38.

Atkinson, C. L., J. F. Kelly, and C. C. Vaughn. 2014b. Tracing consumer-derived nitrogen in riverine food webs. Ecosystems 17:485–496.

Atkinson, C. L., B. J. Sansom, C. C. Vaughn, and K. J. Forshay. 2018. Consumer aggregations drive nutrient dynamics and ecosystem metabolism in nutrient-limited systems. Ecosystems 21:521–535.

Atkinson, C. L., and C. C. Vaughn. 2015. Biogeochemical hotspots: Temporal and spatial scaling of the impact of freshwater mussels on ecosystem function. Freshwater Biology 60:563–574.

Atkinson, C. L., C. C. Vaughn, K. J. Forshay, and J. T. Cooper. 2013. Aggregated filter-feeding consumers alter nutrient limitation: Consequences for ecosystem and community dynamics. Ecology 94:1359–1369.

Benke, A., and A. D. Huryn. 2017. Secondary production and quantitative food webs. Pages 235–254 *in* F. Hauer and G. A. Lamberti (editors). Methods in stream ecology. 3rd edition. Academic Press, Cambridge, Massachusetts.

Benke, A. C., A. D. Huryn, L. A. Smock, and J. B. Wallace. 1999. Length-mass relationships for freshwater macroinvertebrates in North America with particular reference to the southeastern United States. Journal of the North American Benthological Society 18:308–343.

Bogan, A. E., and K. J. Roe. 2008. Freshwater bivalve (Unioniformes) diversity, systematics, and evolution: Status and future directions. Journal of the North American Benthological Society 27:349–369.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771–1789.

Burgherr, P., and E. I. Meyer. 1997. Regression analysis of linear body dimensions vs. dry mass in stream macroinvertebrates. Archiv fur Hydrobiologie 139:101–112.

Carter, J. L., and V. H. Resh. 2013. Analytical approaches used in stream benthic macroinvertebrate biomonitoring programs of State agencies in the United States. Report 2013–1129. United States Department of the Interior, United States Geological Survey, Reston, Virginia.

Carter, J. L., and V. H. Resh. 2001. After site selection and before data analysis: Sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. Journal of the North American Benthological Society 20:658–682.

Dirzo, R., H. S. Young, M. Galetti, G. Ceballos, N. J. B. Isaac, and B. Collen. 2014. Defaunation in the anthropocene. Science 345: 401–406.

Duffy, J. E., B. J. Cardinale, K. E. France, P. B. McIntyre, E. Thébault, and M. Loreau. 2007. The functional role of biodiversity in ecosystems: Incorporating trophic complexity. Ecology Letters 10:522–538.

Galbraith, H. S., D. E. Spooner, and C. C. Vaughn. 2008. Status of rare and endangered freshwater mussels in southeastern Oklahoma. Southwestern Naturalist 53:45–50.

Galbraith, H. S., and C. C. Vaughn. 2011. Effects of reservoir management on abundance, condition, parasitism and reproductive traits of downstream mussels. River Research and Applications 27:193–201.

Garner, J., T. Haggerty, and R. Modlin. 1999. Reproductive cycle of *Quadrula metanevra* (Bivalvia: Unionidae) in the Pickwick Dam tailwater of the Tennessee River. The American Midland Naturalist 141:277–283.

Geist, J. 2011. Integrative freshwater ecology and biodiversity conservation. Ecological Indicators 11:1507–1516.

Golightly, C. G., and R. J. Kosinski. 1981. Estimating the biomass of freshwater mussels (Bivalvia: Unionidae) from shell dimensions. Hydrobiologia 80:263–267.

Haag, W. R. 2012. North American freshwater mussels: Ecology, natural history, and conservation. Cambridge University Press, New York, New York.

Haag, W. R., and J. L. Staton. 2003. Variation in fecundity and other reproductive traits in freshwater mussels. Freshwater Biology 48:2118–2130.

Haag, W. R., and M. L. Warren. 1997. Host fishes and reproductive biology of 6 freshwater mussel species from the Mobile Basin, USA. Journal of the North American Benthological Society 16:576–585.

Haag, W. R., and M. L. Warren. 1998. Role of ecological factors and reproductive strategies in structuring freshwater mussel communities. Canadian Journal of Fisheries and Aquatic Sciences 55:297–306.

Haag, W. R., and M. L. Warren. 2010. Diversity, abundance, and size structure of bivalve assemblages in the Sipsey River, Alabama. Aquatic Conservation: Marine and Freshwater Ecosystems 20:655–667.

Hall Jr, R. O., B. J. Koch, M. C. Marshall, B. W. Taylor, and L. M. Tronstad. 2007. How body size mediates the role of animals in nutrient cycling in aquatic ecosystems. Pages 286–305 *in* A. G. Hildrew, R. Edmonds-Brown, and D. Raffaelli (editors). Body size: The structure and function of aquatic ecosystems. Cambridge University Press, New York, New York.

Hinch, S. G., L. J. Kelly, and R. H. Green. 1989. Morphological variation of *Elliptio complanata* (Bivalvia: Unionidae) in differing sediments of soft-water lakes exposed to acidic deposition. Canadian Journal of Zoology 67:1895–1899.

Hopper, G. W., K. B. Gido, C. C. Vaughn, T. B. Parr, T. G. Popejoy, C. L. Atkinson, and K. K. Gates. 2018. Biomass distribution of fishes and mussels mediates spatial and temporal heterogeneity in nutrient cycling in streams. Oecologia 188:1133–1144.

Hornbach, D. J., V. J. Kurth, and M. C. Hove. 2010. Variation in freshwater mussel shell sculpture and shape along a river gradient. American Midland Naturalist 164:22–36.

Jenkins, D. G. 2015. Estimating ecological production from biomass. Ecosphere 6:1–31.

Kerkhoff, A. J., and B. J. Enquist. 2009. Multiplicative by nature: Why logarithmic transformation is necessary in allometry. Journal of Theoretical Biology 257:519–521.

Leopold, L. B., M. G. Wolman, and J. P. Miller. 1964. Fluvial processes in geomorphology. Dover Publications, New York, New York.

Malmqvist, B., R. S. Wotton, and Y. X. Zhang. 2001. Suspension feeders transform massive amounts of seston in large northern rivers. Oikos 92:35–43.

Matthews, W. J., C. C. Vaughn, K. B. Gido, and E. Marsh-Matthews. 2005. Southern plains rivers. Pages 282–325 *in* A. C. Benke and C. E. Cushing (editors). Rivers of North America. Elsevier, Burlington, Maryland.

McIntyre, P. B., A. S. Flecker, M. J. Vanni, J. M. Hood, B. W. Taylor, and S. A. Thomas. 2008. Fish distributions and nutrient cycling in streams: Can fish create biogeochemical hotspots? Ecology 89:2335–2346.

Meyer, E. 1989. The relationship between body length parameters and dry mass in running water invertebrates. Archiv fur Hydrobiologie 117:191–203.

Ortmann, A. E. 1920. Correlation of shape and station in freshwater mussels (Naiades). Proceedings of the American Philosophical Society 59:269–312.

Packard, G. C., G. F. Birchard, and T. J. Boardman. 2011. Fitting statistical models in bivariate allometry. Biological Reviews 86: 549–563.

Parmalee, P. W., A. E. Bogan, and A. P. Farms. 1998. The freshwater mussels of Tennessee. University of Tennessee Press, Knoxville, Tennesse.

Ricciardi, A., R. J. Neves, and J. B. Rasmussen. 1998. Impending extinctions of North American freshwater mussels (Unionoida) following the zebra mussel (*Dreissena polymorpha*) invasion. Journal of Animal Ecology 67:613–619.

Ritz, C., and J. C. Streibig. 2008. Nonlinear regression with R. Springer, New York, New York.

Sansom, B. J., C. L. Atkinson, and C. C. Vaughn. 2016. Growth and longevity estimates for mussel populations in three Ouachita mountain rivers. Freshwater Mollusk Biology and Conservation 19(2):19–26.

Sansom, B. J., S. J. Bennett, J. F. Atkinson, and C. C. Vaughn. 2018. Long-term persistence of freshwater mussel beds in labile river channels. Freshwater Biology 63:1469–1481.

Smock, L. A. 1980. Relationships between body size and biomass of aquatic insects. Freshwater Biology 10:375–383.

Solé, R. V., D. Alonso, and A. McKane. 2002. Self-organized instability in complex ecosystems. Philosophical Transactions of the Royal Society of London B: Biological Sciences 357:667–681.

Southwick, R., and A. Loftus. 2017. Investigation and monetary values of fish and freshwater mollusk kills. American Fisheries Society Special Publication 35, Bethesda, Maryland.

Spooner, D. E., and C. C. Vaughn. 2009. Species richness and temperature influence mussel biomass: A partitioning approach applied to natural communities. Ecology 90:781–790.

Stanley, S. M. 1981. Infaunal survival: Alternative functions of shell ornamentation in the Bivalvia (Mollusca). Paleobiology 7:384–393.

Stanley, S. M. 1988. Adaptive morphology of the shell in bivalves and gastropods. Pages 105–141 *in* E. R. Trueman and M. Clarke (editors). The mollusca. Form and function. Academic Press, San Diego, California.

Straus, S., and L. Aviles. 2018. Estimating consumable biomass from body length and order in insects and spiders. Ecological Entomology 43:69–75.

Strayer, D. L. 2008. Freshwater mussel ecology: A multifactor approach to distribution and abundance. University of California Press, Berkeley, California.

Strayer, D. L. 2017. What are freshwater mussels worth? Freshwater Mollusk Biology and Conservation 20:103–113.

Strayer, D. L., and D. Dudgeon. 2010. Freshwater biodiversity conservation: Recent progress and future challenges. Journal of the North American Benthological Society 29:344–358.

Strayer, D. L., and D. R. Smith 2003. A guide to sampling freshwater mussel populations. American Fisheries Society, Bethesda, Maryland.

Towers, D. J., I. M. Henderson, and C. J. Veltman. 1994. Predicting dry weight of New Zealand aquatic macroinvertebrates from linear dimensions. New Zealand Journal of Marine and Freshwater Research 28:159–166.

Vanni, M. J., and P. B. McIntyre. 2016. Predicting nutrient excretion of aquatic animals with metabolic ecology and ecological stoichiometry: A global synthesis. Ecology 97:3460–3471.

Vaughn, C. C., C. L. Atkinson, and J. P. Julian. 2015. Drought-induced changes in flow regimes lead to long-term losses in mussel-provided ecosystem services. Ecology and Evolution 5:1291–1305.

Vaughn, C. C., and C. C. Hakenkamp. 2001. The functional role of burrowing bivalves in freshwater ecosystems. Freshwater Biology 46:1431–1446.

Vaughn, C. C., C. M. Mather, M. Pyron, P. Mehlhop, and E. K. Miller. 1996. The current and historical mussel fauna of the Kiamichi River, Oklahoma. Southwestern Naturalist 41:325–328.

Ward, G., P. Harris, and A. Ward. 2005. Gulf Coast rivers of the southeastern United States. Pages 125–162 *in* A. Benke and C. E. Cushing (editors). Rivers of North America. Elsevier, Burlington, Maryland.

Watters, G. T. 1994. Form and function of unionoidean shell sculpture and shape (Bivalvia). American Malacological Bulletin 11:1–20.

Wen, Y. H., and R. H. Peters. 1994. Empirical models of phosphorus and nitrogen excretion rates by zooplankton. Limnology and Oceanography 39:1669–1679.

Xiao, X., E. P. White, M. B. Hooten, and S. L. Durham. 2011. On the use of log-transformation vs. nonlinear regression for analyzing biological power laws. Ecology 92:1887–1894.