*North American Journal of Fisheries Management*

ISSN: 0275-5947 print / 1548-8675 online
DOI: 10.1002/nafm.10070

MANAGEMENT BRIEF

# Quantile Regression Estimates of Body Weight at Length in Walleye

**Steven H. Ranney***
*1546 Tempest Court, #105, Bozeman, Montana 59718, USA*

***Abstract***
**Quantile regression is a method of estimating fish weight at length for alternate portions of a probability distribution, but it is an approach that has not received much attention in fisheries literature. Quantile regression can provide estimates of any quantile of weight at length without bias (including the 75th quantile, which was often the focus of standard weight [$W_s$] equations), and this is more advantageous than previously defined $W_s$ equations derived from linear or quadratic regression methods. The goal of this study was to demonstrate the utility of quantile regression as a tool to assess fish weight at length at various portions of the probability distribution without bias using Walleye *Sander vitreus* as a case study. Quantile regression models at the 75th quantile were developed for three randomly selected Walleye populations from Georgia and South Dakota and compared with a large ($N$ = 33,589) reference population. Bootstrap resampling procedures indicated that only one population from the state of Georgia had an intercept and slope similar to the reference population. For the one population that had similar intercept and slope to the reference population, predictions of weight at various lengths still fell below the 95% CIs for predicted weights of the reference population, suggesting that slight differences in intercepts and slopes in allometric relationships can result in predicted weights that still differ at some lengths. Predicted weights of Walleye derived from the 10th, 25th, 50th, 75th, and 90th quantiles were used to demonstrate how individuals and populations may be compared at different management targets. Overall, this study demonstrates the relative ease with which quantile regression may be used to compare fish body condition between populations without bias.**

Standard weight ($W_s$) equations and the concept of relative weight ($W_r$) were developed so that fisheries managers would have a “quick, inexpensive, and useful way of obtaining and interpreting fishery data for management purposes” (Wege and Anderson 1978). However, many $W_s$ equations exhibit length-related biases, which not only hamper the evaluations of fish body condition but also influence a manager’s ability to accurately assess fish body condition at all lengths (Gerow et al. 2005; Ranney et al. 2010, 2011). Additionally, there are a number of statistical concerns relating to estimating $W_s$ equations, estimating the length-related biases associated with $W_s$ equations (e.g., Gerow 2011; Ranney et al. 2011), and interpreting $W_r$ (Brendan et al. 2003; Pope and Kruse 2007). Despite these concerns, the use of $W_r$ as an evaluation tool has expanded beyond the scope envisioned by the originators (Cade et al. 2008; Ranney et al. 2011). Standard weight equations and $W_r$ have been used to test whether body condition differs among or within fish populations, and parametric tests are frequently used to compare $W_r$ data (Murphy et al. 1990; Hyatt and Hubert 2001; Brendan et al. 2003). However, Cade et al. (2008) suggested that other means of comparing fisheries populations can provide a higher level of statistical rigor and be more relevant to the questions being asked. One tool to better compare changes in weight–length relationships is quantile regression (Cade and Noon 2003; Cade et al. 2008, 2011; Crane et al. 2015; Crane and Farrell 2017).

Quantile regression is a method to estimate different quantiles of a response variable distribution (Koenker and Bassett 1978; Cade and Noon 2003). In most regression applications, the mean response variable is estimated as a function of the predictor variable (Cade and Noon 2003). For example, the linear weight-as-a-function-of-length model, common in fisheries applications,

$$\log_{10}(W) = a + b \cdot \log_{10}(\text{TL}),$$

is an estimate of the mean response of $\log_{10}$-transformed weight ($W$) as a function of $\log_{10}$-transformed TL. Quantile regression estimates use individual quantiles of the response variable (i.e., $W$) as a function of the predictor variable (i.e., TL). Thus, for a given quantile τ, $Q_W(\tau|\text{TL})$

*E-mail: steven.ranney@gmail.com
Received January 17, 2018; accepted April 4, 2018

is the τth quantile $Q$ as a function of TL (Cade et al. 2008). Quantile regression allows for the estimation of all quantiles of $W$ as a function of TL from $\tau = 0.01$ to 0.99 (Cade et al. 2008).

Length-biased $W_s$ equations may lead to low-biased estimates of $W_r$ at the upper length ranges for many species. Indeed, Ranney et al. (2010, 2011) found that $W_s$ equations developed for Walleye *Sander vitreus* using either the regression-line percentile (RLP) technique (Murphy et al. 1990) or the empirical percentile (EmP) method (Gerow et al. 2005) were biased low. Regressing weight as a function of length and estimating the 75th quantile of weight for this functional relationship can generate unbiased predictions of Walleye weight at the 75th quantile of weight (Cade et al. 2008). Using direct estimates from any quantile of fish weight—from 0.01 to 0.99—would provide a more statistically valid means of comparing fish populations or individual fish to a reference population (Cade et al. 2008). However, little attention has been given to this method in the fisheries literature to date.

The goal of this study was to demonstrate the utility of quantile regression as a tool to assess fish weight at length without bias using Walleye as a case study. Direct comparison of population-specific quantile regression lines—based on well-established statistical theory—show how managers can directly compare fish populations with each other or to a reference population. Predicted weights of Walleye derived from $\tau = 0.10$, 0.25, 0.50, 0.75, and 0.90 across 10-mm length-classes demonstrate how populations may be compared at different management targets. Walleye was considered an ideal candidate species for this study based on previous analyses (see Ranney et al. 2010, 2011), availability of data from a large geographic area, and their recreational and economic importance (Aiken 2011).

## METHODS

To generate regression models, I used Walleye data from Ranney et al. (2010, 2011), which were filtered before analyses with the filtering methods described in Ranney et al. (2010) because data quality can influence parameter estimation, model fit, and predictive ability (Belsley et al. 1980). I refer to these data and any quantile regressions from this data set as the "reference" data set and regressions, respectively. I solicited additional Walleye weight and length data from state fisheries management agencies in Georgia and South Dakota and randomly selected three Walleye populations from each state. Sample size was not a consideration when populations were randomly selected. I refer to these data and the quantile regressions from these data sets as the "state" data sets and regressions.

I used the linear quantile regression function rq() in the quantile regression package "quantreg" (Koenker 2017) in R version 3.3.3 (R Development Core Team 2017) to regress $\log_{10}W$ as a function of $\log_{10}$TL to estimate 75th quantile intercepts ($\beta_0$) and slopes ($\beta_1$) from the reference and state population data. I used $\tau = 0.75$ because this is similar to the established value used in $W_s$ equations (though $W_s$ is not a true estimate of weight at length at $\tau = 0.75$) and could be considered an estimator of fish with "above average" body weight. Previous quantile regression analyses (i.e., Cade et al. 2008, 2011) have shown that a single linear model can provide estimates for separate population-level $\beta_0$ and $\beta_1$ values when a categorical factor (i.e., reference and each state population identifier) and its interaction with the continuous predictor variable (i.e., $\log_{10}$TL) are included in the same model. I used the model

$$Q_{\log_{10} W}(\tau|\log_{10}\mathrm{TL}, I_j) = \beta_{0j}(\tau)I_j + \beta_{1j}(\tau)I_j\log_{10}\mathrm{TL},$$

where $\beta_{0j}$ and $\beta_{1j}$ are the intercepts and slopes for each reference and state population, and $I_j$ is the population identifier. I estimated the SE of $\beta_{0j}$ and $\beta_{1j}$ by resampling, with replacement, 5,000 $\log_{10}W$–$\log_{10}$TL pairs 1,000 times. I calculated 95% CIs around each $\beta_{0j}$ and $\beta_{1j}$ with the $t$ statistic = 1.960.

I used a reparametrized version of the linear model described above to estimate which state populations had $\beta_0$ and $\beta_1$ values different from those of the reference population. To do this, I set the reference population as the base level then set the contrasts to be among the rest of the categorical factors (population identifier and its interaction with the continuous predictor variable). That model then becomes

$$\begin{aligned}Q_{\log_{10} W}(\tau|\log_{10}\mathrm{TL}, I_j) = {} & \log_{10}\beta_0(\tau) + \beta_1(\tau)\log_{10}\mathrm{TL} \\ & + \beta_{0j}(\tau)I_j + \beta_{1j}(\tau)I_j\log_{10}\mathrm{TL},\end{aligned}$$

where $\beta_0$ and $\beta_1$ are intercepts and slope for the reference population, and $\beta_{0j}$ and $\beta_{1j}$ then become the proportionate differences in intercept and slope, respectively, for each state population $I_j$ from the reference population. In a manner similar to that for estimating the SE of the $\beta_0$ and $\beta_1$ for each population-level model, I estimated the SE of the differences in $\beta_{0j}$ and $\beta_{1j}$ from the reference population $\beta_0$ and $\beta_1$ with resampling. I resampled 5,000 $\log_{10}W$–$\log_{10}$TL pairs 1,000 times and calculated the 95% CIs around the differences in $\beta_0$ and $\beta_1$ with the $t$ statistic = 1.960. Alpha for all statistical tests was set to 0.05.

I compared quantiles of weight at specific lengths from each state population to the reference population by regressing $\log_{10}W$ as a function of $\log_{10}$TL with $\tau = 0.05$, 0.10, … up to 0.95 in increments of 0.05 for each state population. I then estimated the weight and 95% CIs of weight for the midpoints of the length categories for Walleye (midpoints: substock = 125 mm; stock–quality = 315 mm; quality–preferred = 445; preferred–memorable = 570 mm; memorable–trophy = 695 mm: Gabelhouse 1984).

To demonstrate how individual fish may be compared at different management targets, I used the reference data set to create linear quantile regressions of $\log_{10}W$ as a function of $\log_{10}$TL at $\tau = 0.10$, 0.25, 0.50, 0.75, and 0.90 and predicted total weight values of each $\tau$ at 10-mm length-classes from 155 to 745 mm. I pooled data from 102 populations ($N$ = 33,597 Walleyes) and assigned equal weights to all populations regardless of contributed sample size (Cade et al. 2008). These values of $\tau$ are appropriate benchmarks of comparison for population weight data. Estimating the 90th quantile of weight as a function of length corresponds to a fish of "excellent" body weight; the 75th quantile allows for a comparison of fish to the reference population "above average" body weight; the 50th quantile allows for comparison of fish to a reference population "average" body weight; the 25th quantile allows for comparison of fish to a reference population "below average" body weight; and the 10th quantile allows for comparison of fish to a reference population "poor" body weight. I exponentiated the predicted values of $\log_{10}W$ to convert $W$ back to body weight in grams [e.g., $10^{(\log_{10}W)}$].

A repository that includes supplementary data and R code used in this manuscript is available at https://github.com/stevenranney/waeQuantiles.

## RESULTS

The filtered reference data set included 102 Walleye populations and 33,597 individual observations of weight and length. The three randomly sampled populations from each state contained $N$ = 313, 795, and 199 individuals for the GA1, GA2, and GA3 populations, respectively, from Georgia and $N$ = 392, 280, and 140 for the SD1, SD2, and SD3 populations, respectively, from South Dakota. Individuals from populations in Georgia were collected with fall gillnetting and late-winter electrofishing. Individuals from South Dakota populations were collected with summer gill nets. The method used in the collection of individuals in the reference data set was unknown.

The 75th quantile regression model $\beta_0$ and $\beta_1$ estimates for the reference data set were −5.702 and 3.277, respectively (Table 1). The 95% CI for the reference data set was $\beta_0$ = −5.717 to −5.687 and $\beta_1$ = 3.271 to 3.282. The lowest $\beta_0$ was in the GA1 population (−5.713) and the largest was in the SD2 population (−5.174; Table 1). The lowest $\beta_1$ value in the state population data was in the SD2 population (3.070) and the largest in the GA1 population (3.275; Table 1).

Bootstrapped 95% CIs of 75th quantile $\beta_0$ and $\beta_1$ values for the reference data set overlapped the 75th quantile estimates of slope and intercept for only one of three Georgia populations and none of the South Dakota populations (Table 2). Estimates and 95% CIs of weight at TL equal to the midpoints of length categories across all quantiles ranging from $\tau = 0.05$ up to 0.95 by increments of 0.05 overlapped considerably for all three populations from Georgia and the reference population, especially at TL = 125 mm (Figure 1A). Overlap in the CIs for all Georgia populations and the reference population was evident across the remaining four weight-at-total-length estimates, though it became less pronounced at the upper length ranges for GA2. There was less overlap in the CI bands for the South Dakota populations at lower lengths (TL = 125, 315, and 445 mm) than at higher lengths (TL = 570 and 695 mm; Figure 1B). Estimates of Walleye weight at $\tau = 0.10$, 0.25, 0.50, 0.75, and 0.90 from the reference population for each 10-mm length class ranged from 23.6, 25.3, 27.3, 29.8, and 32.1 g at the 155-mm length class to 4,107.1, 4,385.1, 4,730.0, 5,114.9, and 5,532.0 g at the 745-mm length class (Table 3).

TABLE 1. Estimates of intercept and slope values from bootstrapped replicates of quantile regression with $\tau = 0.75$ for seven different data sets of Walleye. Bootstrap replicates were estimated by resampling with replacement 5,000 $\log_{10}W$ and $\log_{10}$TL pairs 1,000 times.

| | Intercept ($\beta_0$) | | | Slope ($\beta_1$) | | |
|---|---|---|---|---|---|---|
| Population | 2.5% | Estimate | 97.5% | 2.5% | Estimate | 97.5% |
| Reference | −5.717 | −5.702 | −5.687 | 3.271 | 3.277 | 3.282 |
| GA1 | −5.954 | −5.713 | −5.471 | 3.186 | 3.275 | 3.364 |
| GA2 | −5.551 | −5.451 | −5.352 | 3.135 | 3.173 | 3.212 |
| GA3 | −5.678 | −5.508 | −5.339 | 3.135 | 3.201 | 3.266 |
| SD1 | −5.497 | −5.425 | −5.353 | 3.140 | 3.169 | 3.197 |
| SD2 | −5.356 | −5.174 | −4.992 | 3.000 | 3.070 | 3.139 |
| SD3 | −5.692 | −5.603 | −5.514 | 3.186 | 3.223 | 3.259 |

TABLE 2. Differences and 95% CIs around the differences in intercepts and slope of each state population from the reference population. The *P*-value is from a test of the null hypothesis that the intercept and slope from the quantile regression, where $\tau = 0.75$ for each state population, is equal to the reference population intercept and slope. An asterisk (*) indicates populations with significantly different intercept and slope from the reference population.

| | Intercept ($\beta_0$) | | | | Slope ($\beta_1$) | | | |
|---|---|---|---|---|---|---|---|---|
| Population | 2.5% | Estimate | 97.5% | *P*-value | 2.5% | Estimate | 97.5% | *P*-value |
| GA1 | −0.2484 | −0.0107 | 0.2271 | 0.9299 | −0.0898 | −0.0019 | 0.0861 | 0.9668 |
| GA2* | 0.1486 | 0.2507 | 0.3528 | <0.0001 | −0.1423 | −0.1032 | −0.064 | <0.0001 |
| GA3* | 0.0062 | 0.194 | 0.3819 | 0.0429 | −0.1483 | −0.0757 | −0.0032 | 0.0407 |
| SD1* | 0.2024 | 0.2772 | 0.352 | <0.0001 | −0.1375 | −0.108 | −0.0784 | <0.0001 |
| SD2* | 0.3549 | 0.5281 | 0.7014 | <0.0001 | −0.273 | −0.207 | −0.141 | <0.0001 |
| SD3* | 0.0129 | 0.0991 | 0.1853 | 0.0242 | −0.089 | −0.054 | −0.019 | 0.0025 |

## DISCUSSION

Quantile regression of fisheries weight–length data is a means to compare individual fish weights and population regression data without the biases and statistical limitations inherent in $W_s$ equations and $W_r$ values (Cade et al. 2008). Quantile regression of a reference data set(s) provides fisheries scientists and managers a means by which conspecific fish populations and individuals can be compared with statistical rigor (Cade et al. 2008; Ranney et al. 2011). Fisheries scientists and managers may consider that $W_r$ and $W_s$ still have a use in fisheries science; however, the biases inherent in $W_s$ equations (see Ranney et al. 2010, 2011 for a detailed discussion) do not provide the statistical validity with which to compare individuals and populations. Even when comparing the results of local management actions (e.g., prey augmentation: Cade et al. 2008), quantile regression was demonstrated as superior to $W_r$ and there was no need to have a $W_s$ equation from a large reference population. Those working with weight–length data should carefully consider whether $W_r$ is relevant to the questions they are asking. Linear quantile regression provides the capability to compare unbiased estimates of quantiles and is rooted in standard linear model procedures rather than the ad hoc methods in use today to compare $W_r$ values.

Estimates of weight at length from quantile regression can be used in place of $W_r$ in nearly all instances in which $W_r$ is subject to statistical analysis. Linear quantile regression is a more statistically rigorous tool that can be used to compare changes in weight–length relationships before and after specific treatments or management actions, or to compare the weight–length relationships of multiple populations (e.g., Cade et al. 2008, 2011; Crane et al. 2015; Crane and Farrell 2017). Populations that are monitored on a yearly basis (e.g., trophy fisheries or threatened and endangered populations) are also well suited to quantile regression analysis.

Slight differences in $\beta_0$ and $\beta_1$ in allometric relationships can result in large differences in predictions of weight at various lengths. For example, though GA1 and the reference population had similar $\beta_0$ and $\beta_1$ values (Table 1), the predictions of weight at the 75th quantile for the GA1 population at higher lengths (Figure 1A) fell outside the 95% CIs for the reference data set. Similarly, differences in $\beta_0$ and $\beta_1$ in allometric relationships do not imply that predicted weights will be different across length ranges. At TL = 315 mm (Figure 1B), the SD1 and SD2 populations had almost identical predictions of weight at the 75th quantile compared with that from the reference population. This suggests that the joint effect of $\beta_0$ and $\beta_1$ is what matters most and that it is still difficult to make conclusive statements about how predictions (and populations) differ until actual predictions of weight have been compared with each other. I could have achieved better precision in the CIs around the predictions of weight at length at various quantiles for both the state and reference populations had I resampled more than 5,000 weight–length pairs and bootstrapped the regressions more than 1,000 times. However, resampling pairs to the size of the entire data set ($N$ = 35,716) and running 10,000 bootstrap regressions would have required significant additional computing time.

With unbiased estimates of any quantile of weight for 10-mm length categories, a fisheries manager (or angler) can compare the length and weight of an individual fish to the reference values in Table 3 to determine how that individual compares with the reference population. For example, if a 580-mm Walleye weighs 2,100 g, the fisheries manager can compare those values to the 585-mm length class in Table 3. Thus, a fish that weighs 2,100 g and is 580 mm long is somewhere between the 25th and 50th quantile estimate of weight for the reference population. If interested, a fisheries manager could determine at which $\tau$ a 580-mm fish weighing 2,100 g lies. Further, any additional quantile weights between $\tau = 0.01$ and 0.99 and their CIs may be readily calculated and included in tabular form as a quick reference for fish body weight to identify potential areas for growth bottlenecks. For example, a

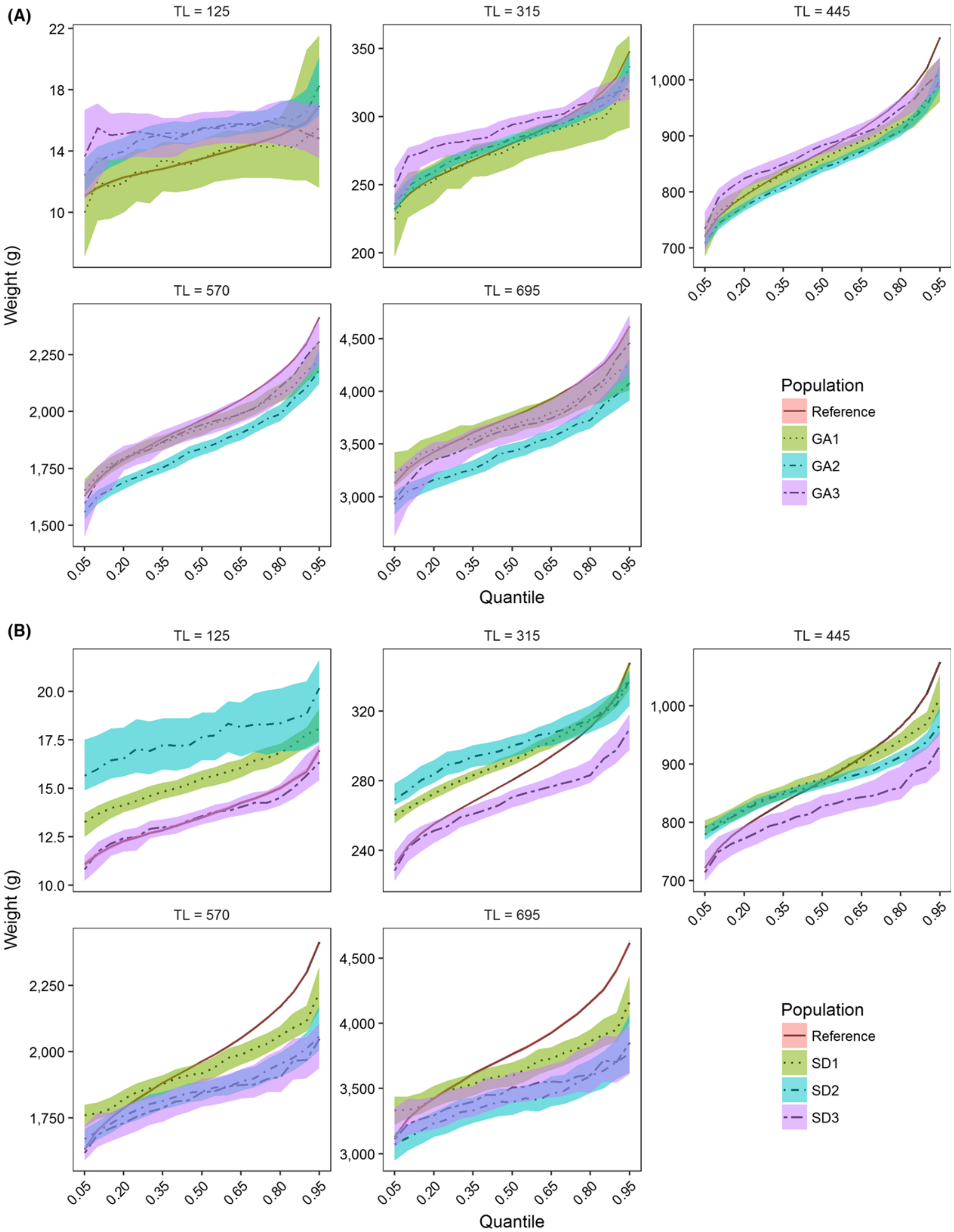


FIGURE 1. Quantile regression ($\tau = 0.05$ up to 0.95 by increments of 0.05) predictions and 95% CIs of weight for the midpoint of each length category (substock, stock–quality, quality–preferred, preferred–memorable, memorable–trophy: Gabelhouse 1984) for the reference data set and **(A)** three populations of Walleye from Georgia and **(B)** three from South Dakota. Confidence bands around the reference population are narrow and appear as a solid line.

TABLE 3. Weight (g) estimates of Walleye at 10-mm length classes from 155 to 745 mm for τ = 0.10, 0.25, 0.50, 0.75, and 0.90 of the reference data set. Estimates of weight were derived from linear quantile regressions of $\log_{10}W$ as a function of $\log_{10}$TL.

| | Quantile (τ) | | | | |
|---|---|---|---|---|---|
| TL (mm) | 0.10 | 0.25 | 0.50 | 0.75 | 0.90 |
| 155 | 23.6 | 25.3 | 27.3 | 29.8 | 32.1 |
| 165 | 28.9 | 31.0 | 33.6 | 36.6 | 39.5 |
| 175 | 35.1 | 37.7 | 40.7 | 44.4 | 47.8 |
| 185 | 42.1 | 45.2 | 48.9 | 53.3 | 57.4 |
| 195 | 50.1 | 53.7 | 58.1 | 63.3 | 68.2 |
| 205 | 59.0 | 63.3 | 68.5 | 74.6 | 80.4 |
| 215 | 69.1 | 74.0 | 80.1 | 87.2 | 94.0 |
| 225 | 80.2 | 86.0 | 92.9 | 101.2 | 109.1 |
| 235 | 92.5 | 99.1 | 107.2 | 116.7 | 125.8 |
| 245 | 106.1 | 113.7 | 122.9 | 133.7 | 144.2 |
| 255 | 121.0 | 129.7 | 140.2 | 152.5 | 164.4 |
| 265 | 137.3 | 147.1 | 159.0 | 172.9 | 186.6 |
| 275 | 155.1 | 166.1 | 179.6 | 195.3 | 210.6 |
| 285 | 174.4 | 186.8 | 201.9 | 219.5 | 236.8 |
| 295 | 195.4 | 209.2 | 226.1 | 245.8 | 265.2 |
| 305 | 218.0 | 233.4 | 252.2 | 274.1 | 295.8 |
| 315 | 242.4 | 259.5 | 280.4 | 304.7 | 328.8 |
| 325 | 268.6 | 287.6 | 310.7 | 337.6 | 364.3 |
| 335 | 296.8 | 317.7 | 343.2 | 372.8 | 402.4 |
| 345 | 326.9 | 349.9 | 378.0 | 410.5 | 443.1 |
| 355 | 359.1 | 384.3 | 415.2 | 450.8 | 486.6 |
| 365 | 393.4 | 421.0 | 454.8 | 493.8 | 533.1 |
| 375 | 430.0 | 460.1 | 497.0 | 539.5 | 582.5 |
| 385 | 468.8 | 501.7 | 541.8 | 588.1 | 635.0 |
| 395 | 510.1 | 545.7 | 589.4 | 639.6 | 690.7 |
| 405 | 553.8 | 592.4 | 639.8 | 694.2 | 749.7 |
| 415 | 600.0 | 641.8 | 693.1 | 752.0 | 812.1 |
| 425 | 648.9 | 694.0 | 749.5 | 813.0 | 878.0 |
| 435 | 700.4 | 749.1 | 808.9 | 877.4 | 947.6 |
| 445 | 754.7 | 807.2 | 871.6 | 945.2 | 1,021.0 |
| 455 | 812.0 | 868.3 | 937.5 | 1,016.6 | 1,098.1 |
| 465 | 872.1 | 932.6 | 1,006.9 | 1,091.7 | 1,179.3 |
| 475 | 935.3 | 1,000.1 | 1,079.7 | 1,170.5 | 1,264.5 |
| 485 | 1,001.6 | 1,070.9 | 1,156.1 | 1,253.2 | 1,353.9 |
| 495 | 1,071.1 | 1,145.1 | 1,236.2 | 1,339.9 | 1,447.6 |
| 505 | 1,143.9 | 1,222.9 | 1,320.1 | 1,430.6 | 1,545.7 |
| 515 | 1,220.1 | 1,304.2 | 1,407.8 | 1,525.6 | 1,648.4 |
| 525 | 1,299.7 | 1,389.3 | 1,499.6 | 1,624.8 | 1,755.7 |
| 535 | 1,382.9 | 1,478.1 | 1,595.4 | 1,728.4 | 1,867.8 |
| 545 | 1,469.7 | 1,570.8 | 1,695.4 | 1,836.5 | 1,984.7 |
| 555 | 1,560.2 | 1,667.4 | 1,799.6 | 1,949.3 | 2,106.6 |
| 565 | 1,654.5 | 1,768.2 | 1,908.2 | 2,066.7 | 2,233.7 |
| 575 | 1,752.8 | 1,873.0 | 2,021.4 | 2,189.0 | 2,366.0 |
| 585 | 1,855.0 | 1,982.1 | 2,139.1 | 2,316.3 | 2,503.6 |
| 595 | 1,961.3 | 2,095.6 | 2,261.4 | 2,448.5 | 2,646.7 |
| 605 | 2,071.8 | 2,213.5 | 2,388.6 | 2,586.0 | 2,795.4 |
| 615 | 2,186.5 | 2,336.0 | 2,520.6 | 2,728.7 | 2,949.8 |
| 625 | 2,305.6 | 2,463.1 | 2,657.7 | 2,876.8 | 3,110.0 |
| 635 | 2,429.1 | 2,594.9 | 2,799.8 | 3,030.3 | 3,276.1 |
| 645 | 2,557.1 | 2,731.5 | 2,947.1 | 3,189.5 | 3,448.4 |
| 655 | 2,689.8 | 2,873.1 | 3,099.8 | 3,354.4 | 3,626.8 |
| 665 | 2,827.1 | 3,019.7 | 3,257.8 | 3,525.2 | 3,811.6 |
| 675 | 2,969.3 | 3,171.4 | 3,421.4 | 3,701.9 | 4,002.8 |
| 685 | 3,116.4 | 3,328.3 | 3,590.6 | 3,884.6 | 4,200.5 |
| 695 | 3,268.5 | 3,490.5 | 3,765.5 | 4,073.5 | 4,405.0 |
| 705 | 3,425.7 | 3,658.2 | 3,946.3 | 4,268.8 | 4,616.3 |
| 715 | 3,588.0 | 3,831.4 | 4,133.0 | 4,470.4 | 4,834.5 |
| 725 | 3,755.6 | 4,010.2 | 4,325.8 | 4,678.5 | 5,059.8 |
| 735 | 3,928.6 | 4,194.8 | 4,524.7 | 4,893.3 | 5,292.2 |
| 745 | 4,107.1 | 4,385.1 | 4,730.0 | 5,114.9 | 5,532.0 |

fisheries manager responsible for the SD1 and SD2 populations can determine that the 75th quantile of fish at the stock–quality midpoint (TL = 315 mm in Figure 1B) are tracking well with the reference data set. However, at the memorable–trophy midpoint (TL = 695 mm in Figure 1B), Walleyes from those same populations fall well below the species-wide standards at all estimated quantiles except τ = 0.15 and below. This could be an indication that prey availability at length categories beyond the quality–preferred category may be limiting Walleye growth. Similarly, the manager responsible for all Georgia populations can see that the populations of Walleye at the quality–preferred category midpoint (TL = 445 mm in Figure 1A) are all tracking well with the species-wide reference populations. Values in Table 3 could easily be converted to imperial measurements if so desired.

Though I have refrained from referring to quantile regression predictions of weight as estimates of "condition," I did select the quantiles shown in Table 3 for a reason. Historically, $W_s$ equations have been estimated at the 75th percentile (Wege and Anderson 1978) and if the ratio of individual fish weight to $W_s \times 100$ was greater than 100, then that individual is assumed to be in good or "above average" condition (Wege and Anderson 1978; Neumann et al. 2012). Here, because quantile regression is estimating the response of $\log_{10}$-transformed fish weight as a function of $\log_{10}$-transformed TL at the quantile specified, we can use phrases like "excellent," "above average," "average," "below average," and "poor" weight for τ = 0.90, 0.75, 0.50, 0.25, and 0.10, respectively, because that is one interpretation of what those quantiles represent. Other quantiles and more refined definitions could be established to identify benchmarks for comparison in other species or for different management priorities. Lower values of $\tau$ for nongame fishes or threatened or endangered species may be a more appropriate benchmark for comparison. State and regional fisheries scientists and managers can determine which quantile or set of quantiles best fits the question they are asking or the species they are investigating. The ease with which quantile models can be estimated and compared makes the need to set a "standard" quantile (or set of quantiles) superfluous.

I have used a reference data set that spans a large portion of the geographic range of Walleye. Given that this paper is meant as an introduction to using quantile regression to estimate fish body weight at length, this seemed like a reasonable first step. This is not meant to suggest, however, that fisheries scientists and managers should always use such a robust data set to which they can compare their own fish populations. Indeed, though fisheries managers in Georgia may want to compare their populations to those of the upper U.S. Midwest periodically, it seems more likely that managers in one region would be most interested in making within-region comparisons. As a result, regional reference data sets could be established and used as reference populations, depending upon the comparison of interest. Fisheries scientists and managers could easily define a useful organization of geography and habitat (e.g., reservoirs, lakes, rivers) so that consistency within a region is maintained. Deciding which regions and which habitats should be grouped together is challenging; however, the well-developed statistical method by which those comparisons can be made should allow fisheries scientists and managers to spend more time considering the question of organization rather than considering how to conduct that analysis.

The minimum sample size needed for a reference data set has not been investigated. However, given that CI estimates are made with the *t* statistic (Zar 1999), if sample sizes are low, then CIs of any range will be less precise. The minimum number of populations that should be included in a reference data set has also not been investigated. Provided that the reference data set contains enough populations to be representative of the species across the spatial extent of the comparison(s) being made, the number of populations in the reference data set is likely of little concern because all populations are given equal weights, regardless of contributed sample size (Cade et al. 2008). If there is a relevant weighting scheme based on survey design, weights can be incorporated into the linear quantile regression model by implementing appropriate weighting arguments. Averaging quantile estimates across multiple populations—based on constructing appropriate contrasts in the linear model—is an alternative approach to having a single "reference" population (Cade et al. 2011).

Many state and provincial management agencies already have standard methods in place to collect fisheries weight–length data, and therefore alternative methods are unnecessary for "standardized" collections. For management agencies that do not have a standard method by which to collect fisheries length–weight data for a given species, I recommend establishing protocols as soon as is reasonable (see Bonar et al. [2009] for guidelines). Recognizing that time of year may affect the weight of an individual fish (e.g., prespawn versus postspawn), state reference data sets that are compiled from a standardized collection method would be consistent. For larger, regional reference data sets, the variability inherent across standardized collection methods from contributing management agencies (e.g., prespawn versus late summer versus under ice) would likely negate any biases that could be inherent in the data (e.g., heavier fish collected during the prespawn season, lighter fish collected in midwinter). Additionally, resampling a data set considered to be representative of the species across the spatial extent under consideration when estimating the $\beta_0$ and $\beta_1$ coefficients of the quantile regression models would serve to reduce the influence of the collection method on estimates of $\beta_0$, $\beta_1$, or the differences between the $\beta_0$ and $\beta_1$ of the reference population and any other populations being compared.

Frequency of reference data set compilation should be determined by fisheries scientists and managers. Climate change may impact different geographic regions at different rates. As a result, defining the frequency with which reference data sets should be established is beyond the scope of this paper. Climate change could affect growth of fishes across geographic regions and latitudes, and reference data sets should reflect those changes. Regardless of how climate change affects fish growth, the ease with which population comparisons can be made with quantile regression will allow fisheries scientists and managers more time to think about their research questions and management goals rather than the best way to conduct analyses of their data.

Quantile regressions of fisheries $\log_{10}$-transformed weight–length data provide a statistically valid means of comparing any quantile estimates of weight. While I do not advocate eliminating $W_s$ and $W_r$ as a means of evaluating fish condition in situations appropriate to management, quantile regression provides a simple, tractable, and statistically valid means of directly evaluating fisheries weight at length for alternate portions of a probability distribution (Cade and Noon 2003). Quantile regression procedures are available in common data analysis software: XLStat for Excel, PROC QUANTREG in SAS, and the "statsmodels" library in Python. It is my hope that quantile regression will become a common means of evaluating various distributions of fish weight at length and will help fisheries scientists and managers compare weight at length of disparate populations without the length biases of $W_s$ and the ad hoc methods used to analyze $W_r$ data.

## ACKNOWLEDGMENTS

I thank the contributors of the state population data sets used in these analyses: Bill Couch, Anthony Rabern, and Dennis Schmitt of the Georgia Department of Natural Resources, Wildlife Resources Division, and Brian Blackwell of the South Dakota Department of Game, Fish, and Parks. I also thank Melissa Wuellner from the Department of Natural Resource Management, South Dakota State University, and Mandy Kauffman from Western Ecosystems Technologies, Inc., in Laramie, Wyoming, both of whom provided necessary and useful reviews of this manuscript early in its development. I also thank Brian Cade of the U.S. Geological Survey for providing particularly helpful reviews that included substantial technical support regarding quantile regression and two additional anonymous reviewers that provided thoughtful guidance on this manuscript. There is no conflict of interest declared in this article.

## REFERENCES

Aiken, R. 2011. Net economic values for wildlife-related recreation in 2011: addendum to the 2011 national survey of fishing, hunting, and wildlife-associated recreation. U.S. Fish and Wildlife Service Report 2011-8.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. Regression diagnostics: identifying influential data and sources of collinearity. Wiley, New York.

Bonar, S. A., W. A. Huber, and D. W. Willis, editors. 2009. Standard methods for sampling North American freshwater fishes. American Fisheries Society, Bethesda, Maryland.

Brendan, T. O., B. R. Murphy, and J. B. Birch. 2003. Statistical properties of the relative weight ($W_r$) index and an alternative procedure for testing $W_r$ differences between groups. North American Journal of Fisheries Management 23:1136–1151.

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

Cade, B. S., J. W. Terrell, and B. C. Neely. 2011. Estimating geographic variation in allometric growth and body condition of Blue Suckers with quantile regression. Transactions of the American Fisheries Society 140:1657–1669.

Cade, B. S., J. W. Terrell, and M. T. Porath. 2008. Estimating fish body condition with quantile regression. North American Journal of Fisheries Management 28:349–359.

Crane, D. P., and J. M. Farrell. 2017. Trends in body condition of Smallmouth Bass and Northern Pike (1982–2013) following multiple ecological perturbations in the St. Lawrence River. Canadian Journal of Fisheries and Aquatic Sciences 74:1158–1172.

Crane, D. P., J. M. Farrell, D. W. Einhouse, J. R. Lantry, and J. L. Markham. 2015. Trends in body condition of native piscivores following invasion of Lakes Erie and Ontario by the Round Goby. Freshwater Biology 60:111–124.

Gabelhouse, D. W. Jr. 1984. A length-categorization system to assess fish stocks. North American Journal of Fisheries Management 4:273–285.

Gerow, K. G. 2011. Comment: assessing length-related biases in standard weight equations. North American Journal of Fisheries Management 31:656–660.

Gerow, K. G., R. C. Anderson-Sprecher, and W. A. Hubert. 2005. A new method to compute standard-weight equations that reduces length-related bias. North American Journal of Fisheries Management 25:1288–1300.

Hyatt, M. W., and W. A. Hubert. 2001. Statistical properties of relative weight distributions of four salmonid species and their sampling implications. North American Journal of Fisheries Management 21:666–670.

Koenker, R. 2017. quantreg: Quantile regression. R package version 5.34. Available: http://cran.r-project.org/package=quantreg. (April 2018).

Koenker, R., and G. Bassett. 1978. Regression quantiles. Econometrica 46:33–50.

Murphy, B. R., M. L. Brown, and T. A. Springer. 1990. Evaluation of the relative weight ($W_r$) index, with new applications to Walleye. North American Journal of Fisheries Management 10:85–97.

Neumann, R. M., C. S. Guy, and D. W. Willis. 2012. Length, weight, and associated indices. Pages 637–676 *in* A. V. Zale, D. L. Parrish, and T. M. Sutton, editors. Fisheries techniques, 3rd edition. American Fisheries Society, Bethesda, Maryland.

Pope, K. L., and C. G. Kruse. 2007. Condition. Pages 423–471 *in* C. S. Guy and M. L. Brown, editors. Analysis and interpretation of freshwater fisheries data. American Fisheries Society, Bethesda, Maryland.

R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available: http://www.r-project.org/. (April 2018).

Ranney, S. H., M. J. Fincel, M. R. Wuellner, J. A. VanDeHey, and M. L. Brown. 2010. Assessing length-related bias and the need for data standardization in the development of standard weight equations. North American Journal of Fisheries Management 30:655–664.

Ranney, S. H., M. J. Fincel, M. R. Wuellner, J. A. VanDeHey, and M. L. Brown. 2011. Assessing length-related biases in standard weight equations: response to comment. North American Journal of Fisheries Management 31:661–665.

Wege, G. J., and R. O. Anderson. 1978. Relative weight ($W_r$): a new index of condition for Largemouth Bass. Pages 79–91 *in* G. D. Novinger and J. G. Dillard, editors. New approaches to the management of small impoundments. American Fisheries Society, Special Publication 5, Bethesda, Maryland.

Zar, J. H. 1999. Biostatistical analysis, 4th edition. Prentice Hall, Upper Saddle River, New Jersey.