North American Journal of Fisheries Management
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ujfm20

Assessing Length-Related Biases in Standard Weight Equations: Response to Comment

Steven H. Ranney [a], Mark J. Fincel [b], Melissa R. Wuellner [b], Justin A. VanDeHey [b] & Michael L. Brown [b]

[a] Gateway Environmental, Post Office Box 103, Gallatin Gateway, Montana, 59730-0103, USA

[b] Department of Wildlife and Fisheries Sciences, South Dakota State University, Box 2140B, Northern Plains Biostress 138, Brookings, South Dakota, 57007-1696, USA

Available online: 25 Aug 2011

To cite this article: Steven H. Ranney, Mark J. Fincel, Melissa R. Wuellner, Justin A. VanDeHey & Michael L. Brown (2011): Assessing Length-Related Biases in Standard Weight Equations: Response to Comment, North American Journal of Fisheries Management, 31:4, 661-665

To link to this article: http://dx.doi.org/10.1080/02755947.2011.608611

North American Journal of Fisheries Management 31:661–665, 2011

ISSN: 0275-5947 print / 1548-8675 online
DOI: 10.1080/02755947.2011.608611

COMMENT

Assessing Length-Related Biases in Standard Weight Equations: Response to Comment

Length and weight data are among the most frequently collected data and provide some of the most inexpensive and important information available to fisheries scientists (Anderson and Neumann 1996). The concept of condition in fisheries stems directly from length–weight relationships and has been around for over 100 years (Fulton 1904, cited by Nash et al. 2006). More recently, the concepts of standard weight (W_s) and relative weight (W_r) (Wege and Anderson 1978) were established as a "quick, inexpensive, and useful way of obtaining and interpreting fishery data for management purposes" (Anderson 1978). Further, the creation of a metric that is readily and easily comparable across populations permits the establishment of quantifiable objectives, provides a means of evaluating management actions, and permits standardization (Legler 1977; Anderson 1978; Wege and Anderson 1978; Brouder et al. 2009). Since the inception of W_s equations, there have been advances in their development (Willis 1989; Murphy et al. 1990; Gerow et al. 2004).

In 1990, Murphy et al. proposed a new technique for the development of W_s equations: the regression line percentile (RLP) method. This method was widely used for over a decade and remains popular to this day (e.g., Rennie and Verdon 2008; Ranney et al. 2010). Recently, however, W_s equations developed using the RLP method have been criticized for exhibiting length-related bias, especially for longer fish (Gerow et al. 2004, 2005). Gerow et al. (2005) offered a new method (the empirical percentile [EmP] method) to produce W_s equations that they claim reduces the over- or underestimation of W_r values with increasing length. To investigate the claims made by Gerow et al. (2005), Ranney et al. (2010) compared W_s equations using historical validation methods (i.e., Willis 1989; Murphy et al. 1990). They found that W_s equations developed by both the RLP and EmP methods perform similarly and suggested that fisheries managers and scientists continue to use the RLP method to develop such equations.

Gerow (2011) was directly critical of Ranney et al. (2010) on two specific points. First, that our use of the third quartile (Q_3) residuals by 10-mm length-class to detect length-related bias was incorrect, and second, that we used the incorrect parameter (i.e., the mean of W_r values) when we regressed W_r on length to evaluate W_s equations for length-related biases. Below, we address these criticisms and discuss other issues raised by Gerow (2011).

CHOICE OF DISTRIBUTIONS AND PARAMETERS

To visually determine the goodness or lack of fit of W_s equations developed by both the RLP and EmP methods, Ranney et al. (2010) plotted Q_3 residuals by 10-mm length-class (Figure 3 in Ranney et al. 2010). In these plots, the lack of fit was evident for both black crappies *Pomoxis nigromaculatus* and walleyes *Sander vitreus*, from which Ranney et al. (2010) concluded that there was length-related bias in the W_s equations developed for both species by both the RLP and EmP methods. Gerow (2011) criticizes our approach, stating that

> the distribution of individual fish weights from any set of samples from different populations will be different from (usually wider than) the distribution of mean weights from those samples. It is the latter distribution for which Q_3 estimates are computed. . . .

Because of this criticism, we reevaluated our plots for goodness of fit from Ranney et al. (2010). For our "hold-out" data sets (see Ranney et al. 2010 for the geographic location and sample sizes), we calculated residuals for each fish by subtracting W_s from the observed weight. We then calculated the mean residuals by 10-mm length-class within each population. Within each 10-mm length-class, we calculated the Q_3 of the mean residuals. We then plotted the Q_3 residuals as a function of 10-mm length-class (Figure 1).

Despite the alternative analysis suggested by Gerow, our original conclusion remains the same. Interestingly, the length bias associated with both methods appears more pronounced than the length bias from the original analysis from Ranney et al. (2010). For both walleyes and black crappies, it appears that the W_s equations developed by the RLP method have less length-related bias in the upper length-classes even though the W_s equations developed by both methods still contain length-related biases (Figure 1).

INCORRECT PARAMETER

In an attempt to identify whether the EmP method was better than the RLP method at reducing length bias in W_s equations, Ranney et al. (2010) independently compared the two W_s development methods using analytical methods that are supported in the fisheries literature (e.g., Willis 1989; Murphy et al. 1990). Bonar et al. (2009) defined bias as "a *systematic* tendency of

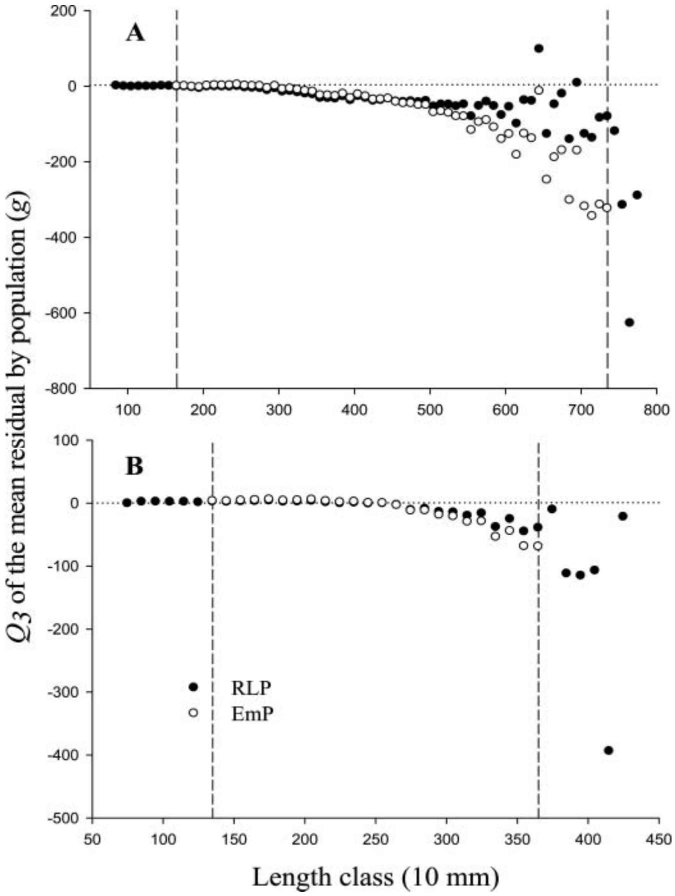

FIGURE 1. Third-quartile (Q_3) residuals from 10-mm length-class population means as a function of length-class from W_s equations developed according to the regression line percentile (RLP) and empirical percentile (EmP) methods. The residuals were calculated according to Gerow (2011) from independent data sets for (**A**) walleyes ($N = 64{,}987$) and (**B**) black crappies ($N = 21{,}317$) from nine different states across the United States. The dashed vertical lines represent the upper and lower limits of the applicable length range for the EmP method; the dotted horizontal lines indicate residuals of zero.

a statistic or estimate derived by sampling to differ from the population value" (emphasis ours). Contrary to the conclusions of Gerow et al. (2005), Ranney et al. (2010) found that the RLP and EmP methods had similar amounts of length-related bias. Further, the W_r estimates—the end products used by fisheries managers—were very similar between methods (see Figure 5 in Ranney et al. 2010). Based on this similarity, Ranney et al. (2010) concluded that fisheries researchers should continue to use the RLP method and suggested that newly developed equations be evaluated against an equally large, fully independent data set collected from across the geographic range of the species (see Ranney et al. 2010 for details).

Gerow (2011) suggests that the length biases that Ranney et al. (2010) found in their study were artifacts of using an improper method (i.e., an "incorrect parameter") to determine length-related bias. Following Gerow's suggestion, we again conducted our analyses on the same randomly selected populations of walleyes and black crappies stratified by state (the "hold-out" data sets, to use Gerow's term), except that we regressed Q_3 on length rather than mean W_r (see Ranney et al. 2010 for details). We used the statistical programming language R, version 2.12.1 (R Development Core Team 2010) to conduct our analyses and the quantile regression package Quantreg, version 4.54 (Koenker 2011) to conduct the Q_3 regressions. For all statistical analyses, α was set to 0.05.

As in Ranney et al. (2010), there were no significant differences ($P = 0.66$) between the number of slopes >0 or <0 or the number of W_r intercepts >100 or <100 ($P = 0.16$; Table 1) for walleye W_r based on the filtered RLP W_s equation (see Ranney et al. 2010 for details on the data filtering methods and hypotheses tested). However, there were significantly more slopes <0 ($P = 0.044$) but no significant difference between the number of intercepts >100 or <100 ($P = 0.15$; Table 1) for walleye W_r based on the filtered EmP W_s equation. For the black crappie W_r based on the filtered RLP W_s equation, there were significantly more slopes <0 ($P = 0.012$) and significantly more W_r intercepts >100 ($P = 0.035$; Table 1). Lastly, for black crappie W_r based on the filtered EmP W_s equation, there were significantly more slopes <0 ($P = 0.002$) and significantly more intercepts <100 ($P = 0.007$; Table 1). From these data, we conclude that the filtered RLP W_s equation exhibited length-related bias only for black crappies, while the filtered EmP W_s equation exhibited length-related bias for both species (Table 1). Our analyses based on Q_3 regression of W_r on length support the conclusions in Ranney et al. (2010) regarding length-related biases. In other words, both the RLP and EmP methods exhibit some length biases and the differences in predicted W_r (the end use of any W_s equation) from the two methods are not sufficient to support redeveloping how W_s equations are generated or used.

EmP VERSUS RLP

Our attempts to address the two criticisms of Gerow (2011) with the previous analyses of Ranney et al. (2010) still demonstrate length-related bias for both the EmP- and RLP-derived W_s equations for both walleyes and black crappies. We believe that a couple of issues with the EmP method still have to be addressed. The first relates to data cleaning. Ranney et al. (2010) suggested that data be filtered to remove aberrant data points. We saw no discussion of data filtering by Gerow in any of his associated papers (Gerow et al. 2004, 2005; Gerow 2010, 2011). Ranney et al. (2010) found that the removal of aberrant data points can have significant influences on the development of W_s equations. Further, Gerow (2011) does not disclose where the data on black crappies and channel catfish *Ictalurus punctatus* were collected. Were these data simulated based on a smaller data set? Are they fully independent of the development model? These are questions that should be explicitly answered in any modeling paper.

TABLE 1. Number of slopes (positive [Pos] and negative [Neg]) that were significantly different from zero and number of intercepts (>100 and <100) that were significantly different from 100 when third-quartile relative weights (W_r) derived from empirical percentile (EmP) and regression line percentile (RLP) W_r equations (using filtered data) were plotted against total length for stratified, randomly selected populations by state. The *P*-values are from chi-square tests of the hypotheses that the number of slopes that were positive would equal the number that were negative and the number of intercepts >100 would equal the number <100. This table is modeled after Ranney et al. (2010).

		RLP				EmP			
		Slopes		Intercepts		Slopes		Intercepts	
Species	*N*	Pos	Neg	>100	<100	Pos	Neg	>100	<100
Walleye	35	13	10	6	15	4	17	10	9
		($P > 0.05$)		($P > 0.05$)		($P < 0.05$)		($P > 0.05$)	
Black crappie	32	3	20	18	4	1	22	3	22
		($P < 0.05$)		($P < 0.05$)		($P < 0.05$)		($P < 0.05$)	

The second issue relates to sample size. During the development of the EmP method, Gerow et al. (2005) failed to satisfy their own criteria. For example, Gerow et al. (2005) suggested that at least 50 fish per length-class be sampled to develop a W_s equation and that at least 20 fish from at least 50 different populations be sampled to eliminate "distortion" bias. Of the 15 species used by Gerow et al. (2004, 2005), only 7 had at least 50 populations. Additionally, many of the species used did not meet the requirement for the minimum number of fish per length-class. For example, the splake (lake trout *Salvelinus namaycush* × brook trout *S. fontinalis*) data set had only 10 populations with a total of 249 fish, and the bull trout *S. confluentus* data set had only 13 populations and 1,033 fish (734 of which were from one population; Hyatt and Hubert 2000). The lack of samples from across the species' range, coupled with low sample sizes, makes the development and interpretation of W_s equations from these data suspect (Murphy et al. 1990; Brown and Murphy 1996; Gerow et al. 2005; Brenden and Murphy 2006). Further, because Gerow et al. (2005) resampled from small data sets, any biases inherent in these data would be entrained by the resampling process. Creating a new data set by resampling a small data set that may be biased would represent pseudovalidation rather than true model validation (Haefner 2005). In his response to Ranney et al. (2010), Gerow (2011) suggests that simulating a data set creates a population for which the truth is known. However, if sample sizes are small in the first and third quartiles, how can a researcher know the "truth"?

Resampling a data set has sound theoretical foundations (Simon 1997; Chernick 1999); however, the length–weight data from one population are not representative of the species as a whole. To prove this point, we conducted an informal resampling exercise using our data. Our independent walleye data set had 64,987 individual measurements of length and weight from across nine states. We selected one population from Georgia that had a sufficiently large population ($N = 749$) and resampled from it (with replacement) until we had 64,987 values for weight. In Figure 2, we present a histogram of both the original data set and our resampled data set. Though the distributions appear to be similar, the mean for the original data set (Figure 2A) is 711.5 (SD, 634.2) and the mean for the resampled data

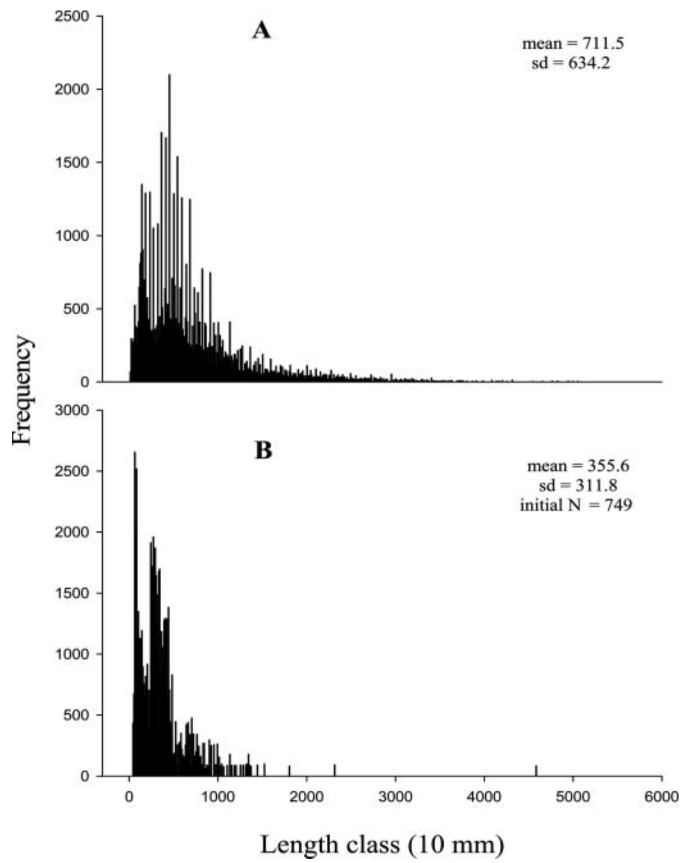

FIGURE 2. Histogram of weights from (**A**) the original independent data set for walleyes ($N = 64{,}987$) and (**B**) a data set that was generated by resampling (with replacement) the initial data set ($N = 749$) to obtain one with 64,987 observations. The difference in means between the two populations is 355.9 g, which suggests that resampling the data from a population that may not be representative of the entire growth form of the species does not create an unbiased independent data set.

set (Figure 2B) is 355.6 (SD, 311.8). The difference in means between these two data sets shows that resampling one population with replacement does not provide an unbiased data set on which to base decisions.

Model validation shows that specified performance standards have been met for a given model (Rykiel 1996). Model validation helps a user decide whether a model is acceptable for its intended purpose. Historically, the operational validation methods for W_s equations have included statistically testing whether or not the number of $\beta_0 > 100$ is significantly different from the number of $\beta_0 < 100$ (Willis 1989; Murphy et al. 1990). Given that W_r values should theoretically model whether or not a population is in good condition (i.e., $\beta_0 = 100$), and given that there should be a greater number of populations that are not in "good" condition, we believe that the statistical test evaluating whether or not the number of $\beta_0 > 100$ and $\beta_0 < 100$ is equal should no longer be used. Perhaps a more reasonable method (i.e., one designed to evaluate management-related significance) for comparing slope values with zero would be to determine equivalence to zero rather than difference from it. We have not investigated how testing for the equivalence of β_1 to zero rather than for its difference from zero would affect the development of W_s equations, but given that a statistical hypothesis can always be rejected with a large enough N, perhaps testing for equivalence is better from a management point of view than testing for differences.

CONCLUSION

We applied the suggestions of Gerow (2011) to new analyses of the data used in Ranney et al. (2010) and still found that EmP-derived W_s equations do not resolve length-related bias issues better than RLP-derived equations. Indeed, there were even cases in which the length-related bias was worse for the EmP-derived equations. While the reduction of length-related bias is a noble goal, recent suggestions of ways to address the issue appear to have taken us away from the fundamental use of W_r. Wege and Anderson (1978) first developed W_r "to provide a quick, convenient means of evaluating management actions." Since that time, many studies have attempted to use W_r for other purposes, including statistical comparisons. This has led to criticisms of W_r indices for unnecessary reasons. We remind the reader that W_r is just one of many tools we can use to determine the impacts of management decisions. Other, more statistically rigorous tools can be used to compare changes in weight–length relationships pre- and posttreatment [e.g., quantile regression (Koenker 2005); see Cade and Noon 2003 and Cade et al. 2008 for examples] or to compare the weight–length relationships of two separate populations (see Pope and Kruse 2007 for discussion of the use of analysis of covariance to test for differences in length–weight regression lines). Thus, RLP-derived W_s equations and the resulting calculations of W_r remain a valuable management tool, but those working with length–weight data should consider whether this tool is relevant to the questions they are asking.

Steven H. Ranney*

Gateway Environmental,
Post Office Box 103,
Gallatin Gateway, Montana 59730-0103, USA

Mark J. Fincel, Melissa R. Wuellner,
Justin A. VanDeHey, and Michael L. Brown

Department of Wildlife and Fisheries Sciences,
South Dakota State University,
Box 2140B, Northern Plains Biostress 138,
Brookings, South Dakota 57007-1696, USA

*Corresponding author: steven.ranney@gateway-environmental.com

REFERENCES

Anderson, R. O. 1978. New approaches to recreational fishery management. Pages 73–78 *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.

Anderson, R. O., and R. M. Neumann. 1996. Length, weight, and associated structural indices. Pages 447–482 *in* B. R. Murphy and D. W. Willis, editors. Fisheries techniques, 2nd edition. American Fisheries Society, Bethesda, Maryland.

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

Brenden, T. O., and B. R. Murphy. 2006. Variance–covariance estimation of standard weight equation coefficients. Journal of Freshwater Ecology 21:1–7.

Brouder, M. J., A. C. Iles, and S. A. Bonar. 2009. Length frequency, condition, growth, and catch per effort indices for common North American fishes. Pages 231–282 *in* S. A. Bonar, W. A. Hubert, and D. W. Willis, editors. Standard methods for sampling North American freshwater fishes. American Fisheries Society, Bethesda, Maryland.

Brown, M. L., and B. R. Murphy. 1996. Selection of a minimum sample size for application of the regression line percentile technique. North American Journal of Fisheries Management 16:427–432.

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 M. T. Porath. 2008. Estimating fish body condition with quantile regression. North American Journal of Fisheries Management 28:349–359.

Chernick, M. R. 1999. Bootstrap methods: a practitioner's guide. Wiley, Oxford, UK.

Fulton, T. W. 1904. The rate of growth of fishes. Pages 141–241 *in* 22nd Annual Report of the Fishery Board of Scotland 1904.

Gerow, K. G. 2010. Biases with the regression line percentile method and the fallacy of a single standard weight. North American Journal of Fisheries Management 30:679–690.

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

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.

Gerow, K. G., W. A. Hubert, and R. C. Anderson-Sprecher. 2004. An alternative approach to detection of length-related biases in standard weight equations. North American Journal of Fisheries Management 24:903–910.

Haefner, J. W. 2005. Modeling biological systems: principles and applications, 2nd edition. Springer, New York.

Hyatt, M. W., and W. A. Hubert. 2000. Proposed standard weight (W_s) equations for kokanee, golden trout, and bull trout. Journal of Freshwater Ecology 15:559–564.

Koenker, R. 2005. Quantile regression. Cambridge University Press, New York.

Koenker, R. 2011. Quantreg: quantile regression. R package version 4.54. Available: www.r-project.org. (October 2010).

Legler, R. E. 1977. New indices for well-being for bluegill. Master's thesis. University of Missouri, Columbia.

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.

Nash, R. D. M., A. H. Valencia, and A. J. Geffen. 2006. The origin of Fulton's condition factor: setting the record straight. Fisheries 31:236–238.

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 Development Core Team. 2010. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

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 30:655–666.

Rennie, M. D., and R. Verdon. 2008. Development and evaluation of condition indices for the lake whitefish. North American Journal of Fisheries Management 28:1270–1293.

Rykiel, E. J., Jr. 1996. Testing ecological models: the meaning of validation. Ecological Modelling 90:229–244.

Simon, J. L. 1997. Resampling: the new statistics. Resampling Stats, Arlington, Virginia.

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.

Willis, D. W. 1989. Proposed standard length–weight equation for northern pike. North American Journal of Fisheries Management 9:203–208.