Black Grouper Length and Weight Relationships for SEDAR 48

Joe O'Hop
Florida Fish and Wildlife Conservation Commission
Fish and Wildlife Research Institute
St. Petersburg, Florida

Introduction

Fisheries research projects, surveys, and monitoring programs in the South Atlantic and Gulf of Mexico regions routinely collect length and weight data as part of providing basic information on species encountered. The types of measurements employed vary among the programs (e.g., Table 1), and potentially can provide information on conversions between measurement types when multiple length or weight types are collected from individual fish. Conversion equations are sometimes necessary to standardize the observed information across surveys and prepare data for stock assessments (e.g., Lombardi-Carlson, 2014).

Methods

Length and weight data for SEDAR 19 (O'Hop and Beaver 2009) were obtained from various sources (Crabtree and Bullock 1998; NMFS Trip Interview Program (TIP), NMFS Southeast Region Headboat Survey (SRHS), NMFS Marine Recreational Information Program [formerly known as the Marine Recreational Fishery Statistics Survey (MRFSS)] NMFS Observer Programs (Shark Bottom Longline Observer Program (SBLOP) and Gulf Reef Fish Observer Program (GOP)), opportunistic samples from NMFS offshore fishery independent surveys, state fisheries programs (North Carolina Division of Marine Fisheries (NCDMF), Florida Fish and Wildlife Conservation Commission (FWC) Fisheries Dependent (FDM) and Fisheries Independent Monitoring (FIM) programs and archived data). Each of these programs may also collect biological samples (typically otoliths) which are submitted to regional laboratories (NMFS Panama City, NMF Beaufort, FWC Fish and Wildlife Research Institute) for processing. The FDM Biostatistical program may collect additional samples (e.g., otoliths for ageing, fin clips for genetic samples, muscle tissue for mercury determinations) from measured specimens which may have originated through sampling for other programs (e.g., TIP, SRHS).

For SEDAR 48, length and weight data used to develop the meristic relationships for SEDAR 19 (Table 2) were re-analyzed and additional sample data were compiled from available sources (TIP, SRHS, MRFSS/MRIP, NCDMF, FDM Biostat, FIM, etc.). Because of some overlap between the data used in SEDAR 19 and the other sources, data were cross-checked using the sampling identifiers of the various programs to prevent duplication. In addition, because of the overlap of data sources, length and weight fields were re-matched to the measurements used in SEDAR 19 to incorporate any revisions that may have occurred since those data were compiled in 2008.

Simple linear regressions (SAS Institute, 2016) were used to analyze the length-length relationships and, after $\log_e$-transformation, the length-weight relationships. Measurements (Table 3) were examined using two rounds of outlier detection, and standardized residuals were computed during each round following methods outlined in O'Hop and Beaver (2009). In the first round of outlier detections, measurements with standardized residuals of 8 or larger (possibly data recording or entry errors) were removed from datasets. The remaining data were re-analyzed, and measurements with standardized residuals of 4 or larger were considered outliers and removed from the analysis, prior to estimating regression coefficients (Table 4) and plotting model fits and residuals (Figs 1 and 2). Outliers identified in these analyses were investigated and some revisions were made to correct data entry errors in the TIP and FDM Biostatistical databases.

Simple linear regression techniques are used in this report to develop the conversion equations. Among the assumptions for these analyses are that there are functional relationships between the variables measured and that measurement error is random for each of the variables. Regression assumes that independent variables are measured without error, and all the error is assigned to the measured "dependent" variable. This simplifying assumption is made to develop the conversion equations. Because of time pressures and other environmental factors in the field, there can be errors in measurement technique, the actual measurement, and recording and entry of data (e.g., Bunch et al. 2013). With the existing survey data available, it is not possible to examine the error associated with making an individual measurement, but it is possible to estimate the degree of relationship between the measurement types and produce suitable conversions between them.

Results and Discussion

While analyses of length-length and length-weight relationships for species are routine, it is good practice to examine data for outliers and potential biases in measurement. The outliers identified from these analyses usually accounted for 1.6% or fewer of the records available for analyses (Table 3) with the exception that a larger percentage (7.5%) of records were excluded for the fork length (FL) versus whole weight regression. Relatively few (0-5) of the outliers had standardized residuals >= 8. Plots of the regression fits and raw residuals (Figs. 1, 2) were typical for length-length and log-transformed length-weight regressions. Upon examination of the FL versus whole weight regression (Fig. 3a), there were unusual runs of weights for specimens under 400 mm in FL (Fig. 3a, "red box"). These "runs" were from specimens (95% of them prior to 1991) measured from the MRFSS/MRIP survey, and may have resulted from the choice of scales used to measure the weight of the specimens (i.e., weight range for the scale was more suitable for heavier specimens). Although similar outlier analyses were performed during SEDAR 19, these patterns went unnoticed when the regressions for SEDAR 19 were estimated (Table 2). The whole weight versus FL regression from SEDAR 19 was like the new regression before the MRFSS/MRIP specimens below 400 mm were removed. Because of the leverage exerted by some of the recorded weights of these smaller specimens, the regression over-stated the weight of specimens under 700 mm and understated the weight of larger specimens compared with the fit after these measurements were removed (Fig. 3c). Many of the excluded specimens were within the 99% prediction interval of the FL-whole wt. regression (Fig. 3c, inset).

The SEDAR 48 relationships (Table 4) differ little from those recommended for SEDAR 19 except for the whole weight versus FL regression. Because the regressions for SEDAR 48 contain the same measurements as were available in SEDAR 19 (and to some extent the research of Crabtree and Bullock 1998) and were augmented with measurements made subsequently, the results are unsurprising.

Acknowledgements

Numerous field biologists from states in southeastern Florida and the National Marine Fisheries Service (NMFS, including their contractors) contributed the essential raw information summarized in this report. This effort and dedication to collecting basic biological data on specimens often goes unrecognized. Thanks also goes to the supervisors of these field programs and database support in making these data available. I thank the NMFS Southeast Fisheries Science Center, particularly Larry Beerkircher (Miami

Laboratory) for the Trip Interview Program data and Kenneth Brennan and Kelly Fitzpatrick (Beaufort Laboratory) for the Southeast Region Headboat Survey data, Beverly Barnett and Dr. Linda Lombardi of the NMFS Panama City Lab, and Dr. Jennifer Potts of the Beaufort Laboratory, and the NMFS staff in Silver Spring, MD, of the Marine Recreational Fisheries Statistics Survey (pre-2004) and Marine Recreational Information Program (2004 and later) for maintaining and making available the national marine recreational survey data. Kelley Kowal, Julia Reeves, Bridget Cermak, Steve Brown, Chris Bradshaw, and Dr. Richard Cody of the FWRI Fisheries Dependent Monitoring Program provided valuable assistance in chasing down information recorded in databases and on datasheets and verifying or revising database entries when necessary. Anne Markwith (North Carolina Division of Marine Fisheries) provided field information and lengths for specimens collected by their biologists. Special thanks to Dr. Roy Crabtree and Lewis Bullock for making their research data available for these analyses.

Literature cited

Bunch, A.J., C.J. Walters, and L.G. Coggins, Jr. 2013. Measurement error in fish lengths: evaluation and management implications. Fisheries 38: 320-326.

Crabtree, R.E. and L.H. Bullock. 1998. Age, growth, and reproduction of black grouper, *Mycteroperca bonaci*, in Florida waters. Fishery Bulletin 96: 735-753. http://fishbull.noaa.gov/964/crabtree1.pdf

Garcia-Cagide, A. and T. Garcia. 1996. Reproduccion de *Mycteroperca bonaci* y *Mycteroperca venenosa* (Pisces: Serranidae) en la plataforma cubana. Rev. Bio. Trop. 44: 771-780.

Hubbs, C.L. and K.F. Lagler. 1964. Fishes of the Great Lakes Region. Univ. Michigan Press, Ann Arbor, MI. 213 p.

Lombardi-Carlson, L. 2014. An age and growth description of Red Grouper (*Epinephelus morio*) from the northeastern Gulf of Mexico: 1978-2013 for SEDAR 42. SEDAR 42-DW-10. South Atlantic Fishery Management Council. North Charleston, SC. http://sedarweb.org/sedar-42-data-workshop

O'Hop, J. and R. Beaver. 2009. Age, growth, and maturity of Black Grouper (*Mycteroperca bonaci*) – Crabtree and Bullock (1998) revisited. SEDAR 19-DW-09. South Atlantic Fishery Management Council. North Charleston, SC. http://sedarweb.org/sedar-19-data-workshop

Saari, C. and L. Beerkircher. 2014. User's Guide for the TIP Trip Interview Program. Version 6. NOAA Fisheries. Southeast Fisheries Science Center, Fisheries Statistics Division. 75 Virginia Beach Drive. Miami, FL 33149. https://www.sefsc.noaa.gov/docs/TIP_UsersGuide_Ver6.pdf

SAS Institute, Inc. 2016. SAS/STAT 14.2. User's Guide. Cary, NC.

SEDAR. 2010. Gulf of Mexico and South Atlantic Black Grouper. SEDAR 19 Stock Assessment Report. South Atlantic Fishery Management Council. Charleston, SC. http://sedarweb.org/sedar-19-final-stock-assessment-report-south-atlantic-and-gulf-mexico-black-grouper

SEDAR. 2016. SEDAR Data Best Practices: Living Document – September 2016. SEDAR. North Charleston, SC. http://sedarweb.org/sedar-data-best-practices

Zar, J. H. 1996. Biostatistical Analysis. Third edition. Prentice-Hall, Englewood Cliffs, New Jersey.

Table 1. A comparison of measurement types and typical units used by selected South Atlantic and Gulf of Mexico fishery monitoring programs and previous Black Grouper research projects.

Program/Project	Standard Length (SL)	Fork Length[1] (FL)	"Natural" Total Length[2] ($TL_{natural}$)	"Maximum" Total length (TL_{max})[3]	Whole weight (TW)	Gutted weight (GW)
Crabtree and Bullock 1998	mm	mm	-------[4]	mm[4]	lbs or kg	lbs or kg
TIP	inches, cm, or mm	inches, cm, or mm	-------[5]	inches, cm, or, mm[5]	lbs, kg, or g[6]	lbs, kg, or g[6]
SRHS	-------	mm	mm	-------	grams	-------
MRFSS/MRIP	-------	mm	-------	-------	kg	-------
FDM Biostat	mm	mm	mm	mm	kg	kg
FIM	mm	mm	-------	mm	kg or g	-------

[1] In fish without forked tails, this measurement is the straight-line distance from the tip of the snout to the rear center edge of the tail. It is also referred to as the "mid-line length".

[2] This measurement is made with the tail in its "natural" or flat orientation.

[3] This measurement follows the method described in Hubbs and Lagler (1964) where the measurement is made by compressing the tips of the tail to its "maximum" length

[4] This project measured the tail following the method described in Hubbs and Lagler (1964).

[5] The recommended measurement method described in the TIP manual (v.6) follows the method described in Hubbs and Lagler (1964). Recently the TIP database has added codes to distinguish between a "natural" TL and "maximum" TL measurements. Recommended precision of measurements was to 1 mm for lengths, and 1 gram or 0.01 pounds for weights.

[6] The TIP database currently stores a whole weight or a gutted weight, but not both.

Table 2. Length-length, whole weight to gutted weight, and length-weight regressions developed for SEDAR 19 [see SEDAR (2010) Table 2.14.7].

Length-Length and Whole Weight-Gutted Weight

Source	**Y**	***a***	***b***	**X**	***n***	**Min X**	**Max X**
SEDAR 19	SL (mm)	-23.85	0.8827	FL[1] (mm)	1320	238	1495
	TL$_{natural}$[2] (mm)	21.29	1.1595	SL (mm)	74	435	1135
	TL$_{natural}$ (mm)	-2.84	1.0193	FL (mm)	137	387	1336
	TL$_{max}$[3] (mm)	26.96	1.1630	SL (mm)	1338	51.2	1325
	TL$_{max}$ (mm)	-1.44	1.0276	FL (mm)	1339	238	1495
	TL$_{max}$ (mm)	4.30	1.0097	TL$_{natural}$ (mm)	78	528	1252
	TW[4](kg)	-----[5]	1.061	GW[4] (kg)	636	0.44 kg	58.64 kg
	TW (kg)	1.006	1.055	GW (kg)	636	0.44 kg	58.64 kg
Crabtree and Bullock 1998	TW (kg)	0.0815	1.056	GW (kg)	638	0.47 kg	61.59 kg

Length-Weight

Source	**Ln(Y)**	**Ln(a)**	**b**	**Ln(X)**	***n***	**Min**	**Max**
SEDAR 19	TW[4] (kg)	-18.5545	3.0843	FL[1] (mm)	2552	206 mm	1495 mm
	TW (kg)	-19.2684	3.1863	TL$_{max}$[3] (mm)	904	77.5 mm	1525 mm
	GW[4] (kg)	-18.8956	3.1306	FL (mm)	2420	314 mm	1495 mm
	GW (kg)	-19.0487	3.1438	TL$_{max}$ (mm)	1075	334 mm	1525 mm

[1] FL – Fork length (mm; in this species, the straight-line distance from the tip of the snout to the rear-center edge of the tail, also called a "mid-line" length).
[2] TL$_{natural}$ - Tail flat (mm), in its natural state
[3] TL$_{max}$ - Tail compressed to its maximum length (mm)
[4] TW=whole weight (kg), GW=gutted weight (kg)
[5] no intercept model

Table 3. Available records and numbers of outliers excluded for various length-length and length-weight regressions for SEDAR 48 and SEDAR 19.

dependent variable	units	independent variable	units	available records*	outliers excluded	records excluded (%)	SEDAR 19 regressions
SL	mm	FL	mm	1573	17	1.1%	1320
TL$_{natural}$	mm	FL	mm	516	3	0.6%	137
TL$_{natural}$	mm	TL$_{max}$	mm	167	0	0.0%	78
TL$_{max}$	mm	SL	mm	1548	16	1.0%	1338
TL$_{max}$	mm	FL	mm	1683	12	0.7%	1339
TW	kg	FL	mm	3220	242**	8.1%	2552
TW	kg	TL$_{natural}$	mm	1055	15	1.4%	------
TW	kg	TL$_{max}$	mm	982	6	0.6%	904
GW	kg	FL	mm	3404	29	0.9%	2420
GW	kg	TL$_{natural}$	mm	130	2	1.6%	------
GW	kg	TL$_{max}$	mm	1596	15	0.9%	1075

*before exclusion of outliers
**MRFSS/MRIP FL measurements less than 400 mm excluded (n=226) because of measurement/scale issues.

Table 4. Length-length (mm) and Length-weight relationships developed for Black Grouper (*Mycteroperca bonaci)*. Linear regressions are in the form $Y = a + bX$. SL: standard length (mm); FL: fork length (mm); TL: total length (mm); TW: total weight (kg), GW: gutted weight (kg).

LENGTH-LENGTH

Source	Y (mm)	*a*	*b*	X (mm)	*n*	Min X (mm)	Max X (mm)	Avg. X* (mm)	MSE*	r^2	Σx^2*	Σxy*	Σy^2*
SEDAR 48	SL	-24.681	0.883	FL[1]	1556	238	1495	776.44	46.880	0.99	88234626.27	77920795.14	68885409.52
	TL$_{natural}$[2]	8.429	1.011	FL	513	387	1347	725.86	81.574	0.99	7329205.89	7409609.11	7532578.56
	TL$_{natural}$	-2.886	0.993	TL$_{max}$	167	534	1270	725.11	37.916	0.99	1946234.06	1931882.26	1923892.36
	TL$_{max}$[3]	26.860	1.164	SL	1532	51.2	1260	663.74	81.923	0.99	71735840.75	83503806.48	97327599.60
	TL$_{max}$	-2.080	1.029	FL	1671	238	1495	777.00	22.834	0.99	91866919.09	94487308.81	97220551.61
Crabtree and Bullock 1998	SL	-23.712	0.883	FL	1134					0.99			
	TL$_{max}$	26.186	1.164	SL	1141					0.99			
	TL$_{max}$	-1.317	1.028	FL	1150					0.99			
Garcia-Cagide y Garcia 1996	TL	17.8	1.10	SL	209								

LENGTH-WEIGHT

Source	Ln (Y [kg])	Ln(a)	b	Ln (X[mm])	*n*	Min [mm]	Max [mm]	Avg. Ln (X[mm])	MSE	r^2	Σx^2	Σxy	Σy^2
SEDAR 48	TW[4]	-19.2391	3.1896	FL	2978	238	1495	6.493	0.01925	0.97	171.0718	545.6546	1797.7224
	TW	-18.5636	3.0722	TL$_{natural}$	1040	260	1600	6.536	0.01566	0.97	66.8133	205.2621	646.85420
	TW	-19.1945	3.1742	TL$_{max}$	976	63	1518	6.521	0.01023	0.99	114.1667	362.3831	1160.2217
	GW[4]	-18.8323	3.1217	FL	3375	327	1495	6.755	0.00782	0.99	240.0079	749.2236	2365.2037
	GW	-19.0112	3.1528	TL$_{natural}$	128	602	1184	6.575	0.00697	0.96	1.96605	6.198593	20.421136
	GW	-18.8867	3.1195	TL$_{max}$	1581	332	1518	6.707	0.00864	0.99	136.4381	425.6203	1341.3646
Crabtree and Bullock 1998	TW	-19.473[5]	3.218	TL$_{max}$	772	177	1518			0.99			

*Avg. X, MSE, Σx^2, Σxy, Σy^2 - Mean of independent variable (X), mean square error and corrected sums of squares (CSS) for the independent variable (X), corrected sum of cross-products for XY, and CSS for the dependent variable (Y); used for generating prediction intervals and for analysis of covariance (Zar 1996), and MSE also used for bias corrections for the means of log-transformed data [e.g., Haddon (2001)]. Usually, lengths were measured to the nearest millimeter, and weight to the nearest 0.02 kg. However, some data may have been taken using length measurements to the nearest 0.5 cm or in fractions of inches and weight measurements to the nearest 0.1 or 0.01 pound. Estimates derived from the above equations should be rounded to the nearest 0.5 centimeter and nearest 0.02 kg. The number of decimals shown in the table were meant solely to reduce rounding errors for calculation of prediction intervals and for generating sums of squares and cross-products needed for analysis of covariance.

[1] FL – Fork length (mm; in this species, the straight-line distance from the tip of the snout to the rear-center edge of the tail, also called a “mid-line” length).

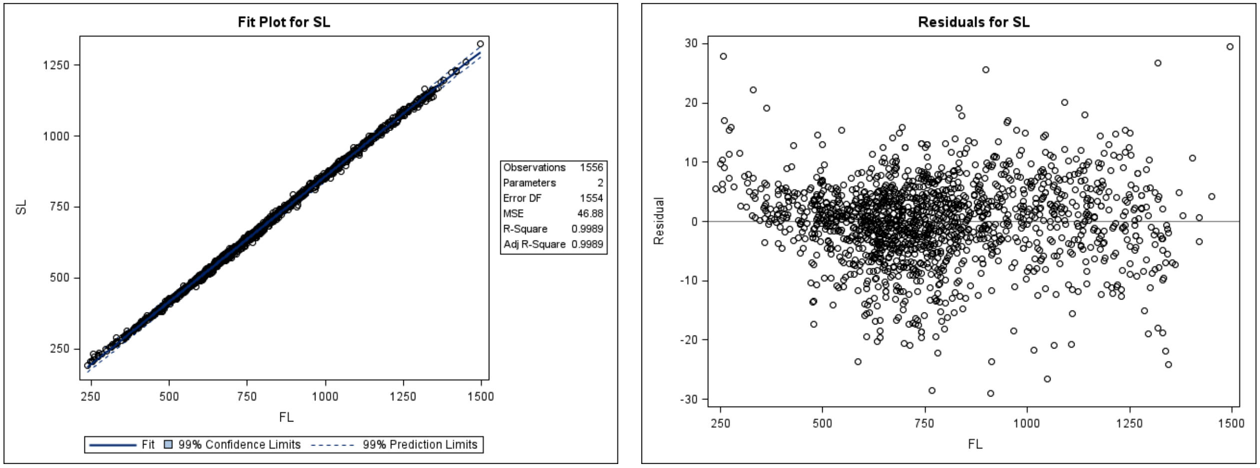
[2] TL$_{natural}$ - Tail flat (mm), in its natural state

[3] TL$_{max}$ - Tail compressed to its maximum length (mm)

[4] TW=whole weight (kg), GW=gutted weight (kg)

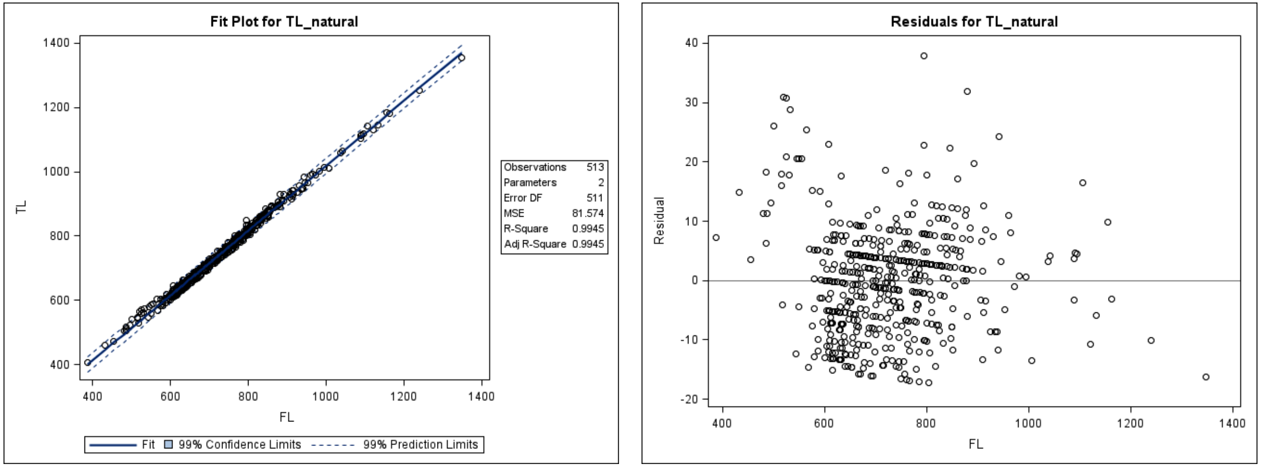
[5] converted from common logarithms and weight in grams to natural logarithms and weight in kg

Figure 1. Predicted fits and raw residuals (in mm) for length-length regressions.



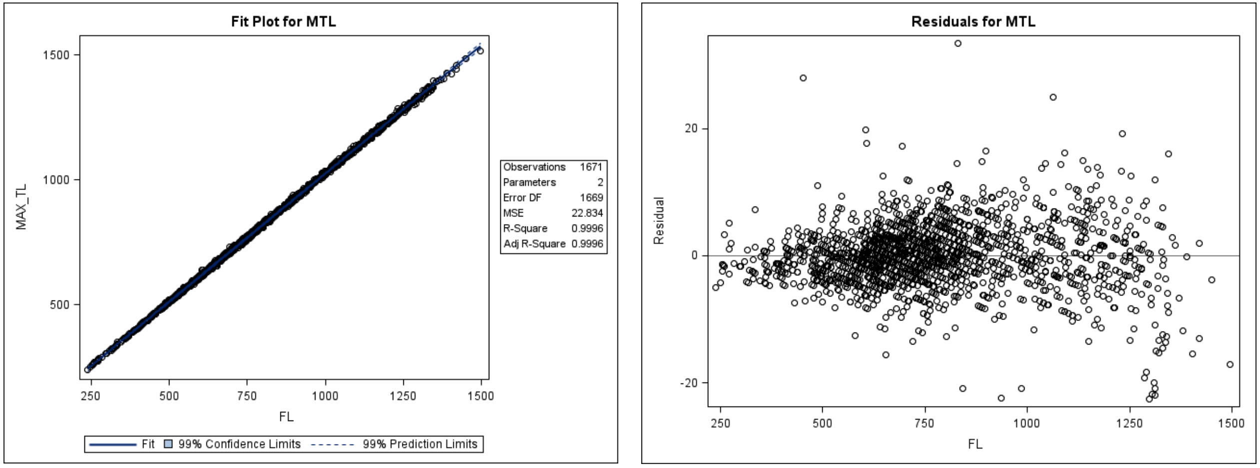
a. SL vs. FL predicted fit

b. SL vs. FL residual plot



c. TL $_{natural}$ vs. FL predicted fit

d. TL $_{natural}$ vs. FL residual plot



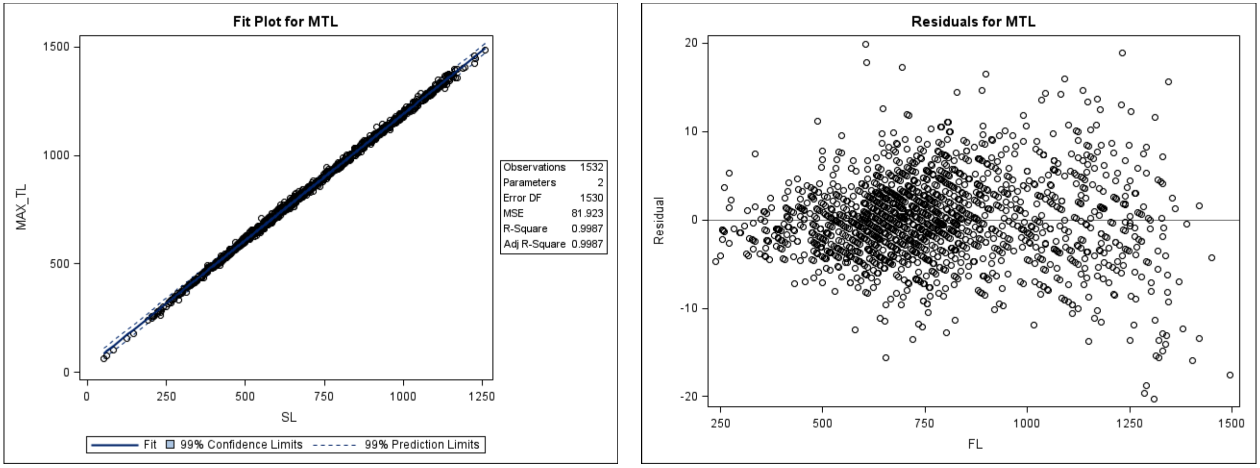
e. TL $_{max}$ vs. FL predicted fit

f. TL $_{max}$ vs. FL residual plot

Figure 1. (continued) Predicted fits and raw residuals (in mm) for length-length regressions.

g. TL $_{max}$ vs. SL predicted fit

h. TL $_{max}$ vs. SL residual plot



i. TL $_{natural}$ vs. TL $_{max}$ predicted fit

j. TL $_{natural}$ vs. TL $_{max}$ residual plot

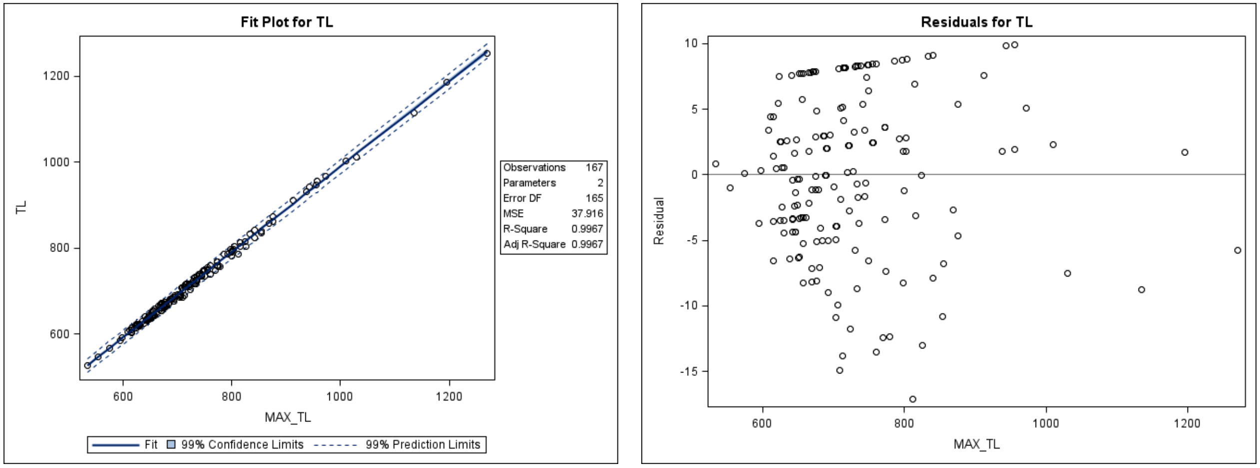
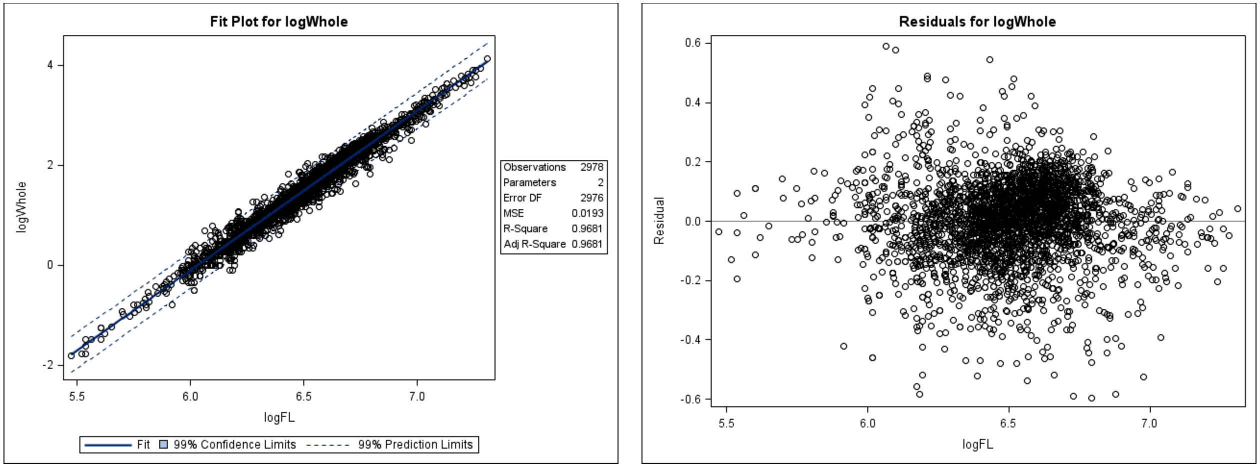
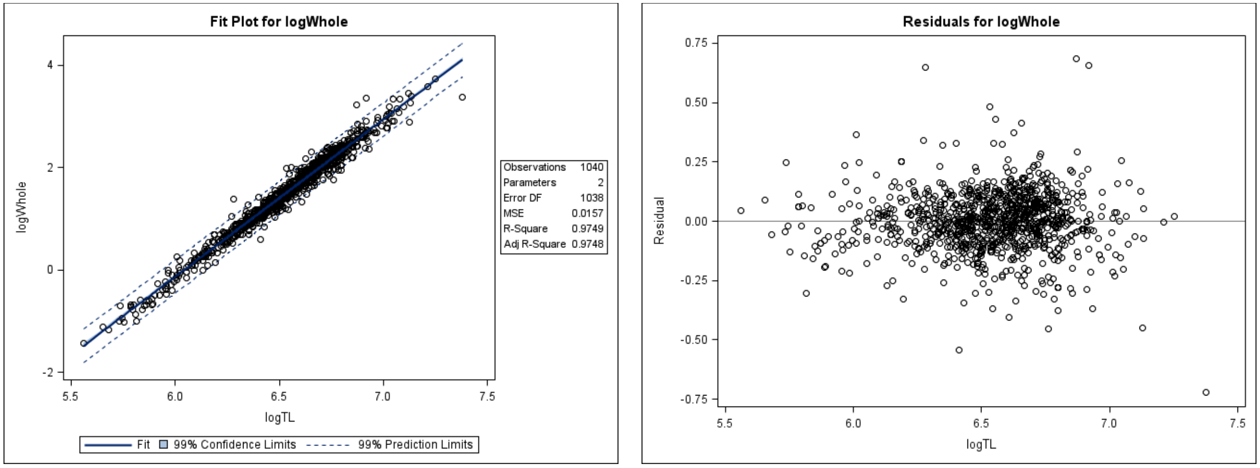


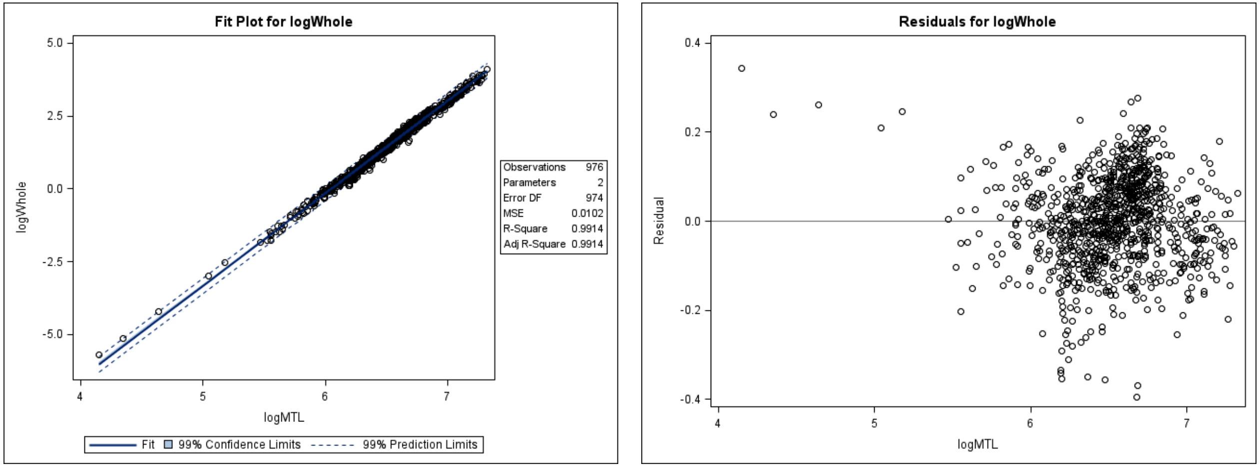
Figure 2. Predicted fits and raw residuals (in log$_e$ kg) for log$_e$-transformed length-weight regressions.


a. Log$_e$(Whole wt.) vs. Log$_e$(FL) predicted fit

b. Log$_e$(Whole wt.) vs. Log$_e$(FL) residual plot

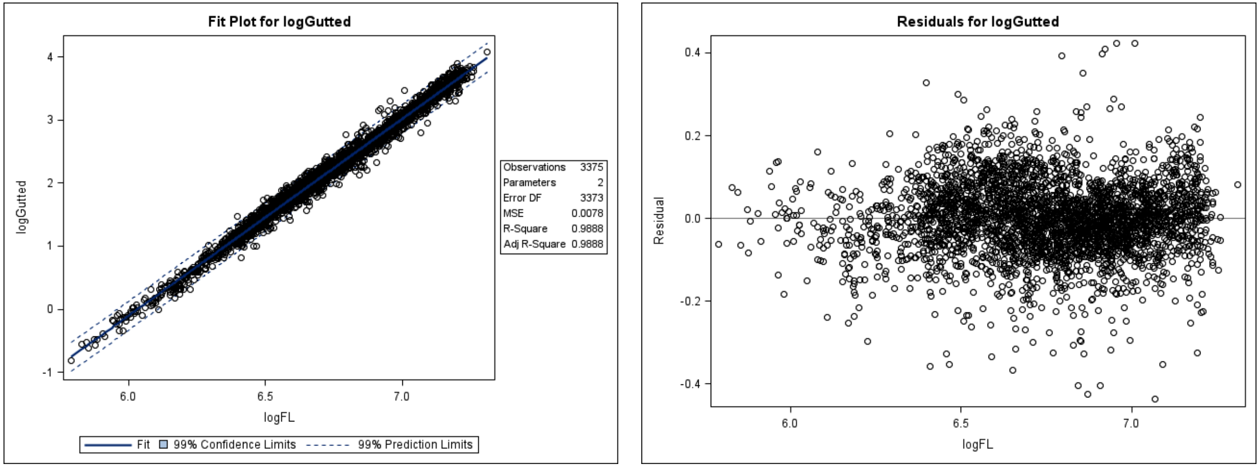

c. Log$_e$(Whole wt.) vs. Log$_e$(TL $_{natural}$) predicted fit

d. Log$_e$(Whole wt.) vs. Log$_e$(TL $_{natural}$) residual plot

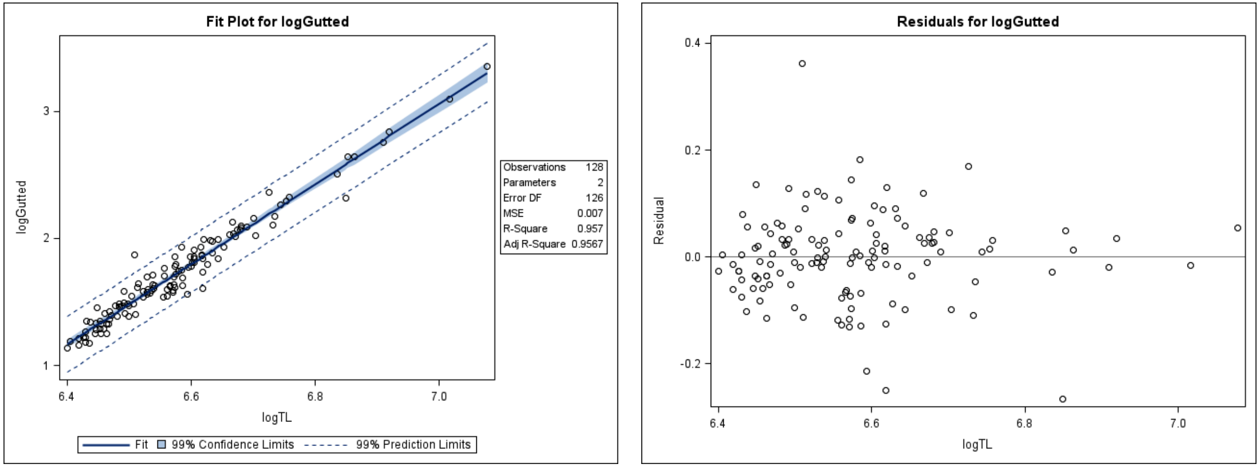

e. Log$_e$(Whole wt.) vs. Log$_e$(TL $_{max}$) predicted fit

f. Log$_e$(Whole wt.) vs. Log$_e$(TL $_{max}$) residual plot

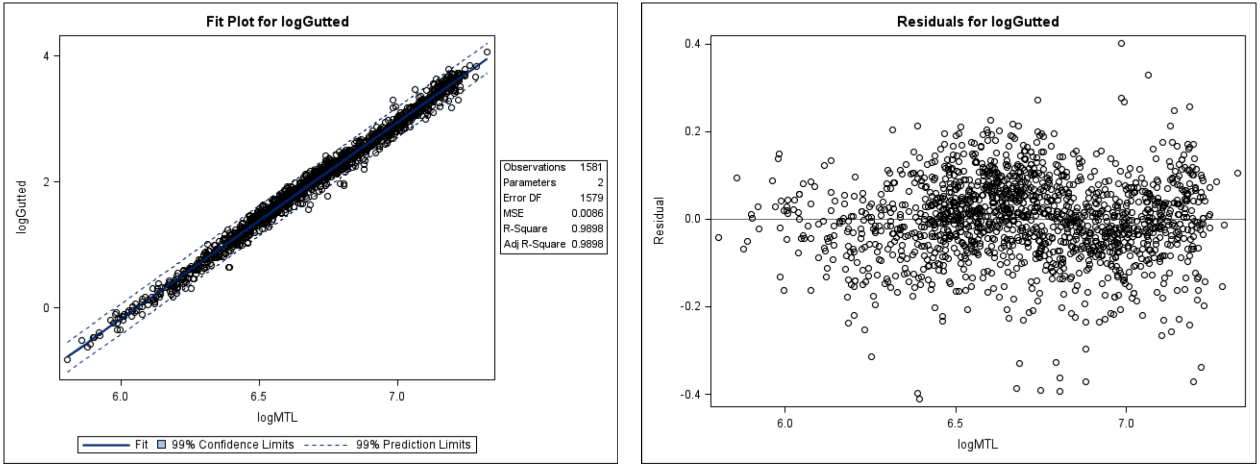
Figure 2. (continued) Predicted fits and raw residuals (in $\log_e$ kg) for $\log_e$-transformed length-weight regressions.


a. Log_e(Gutted wt.) vs. Log_e(FL) predicted fit

b. Log_e(Gutted wt.) vs. Log_e(FL) residual plot

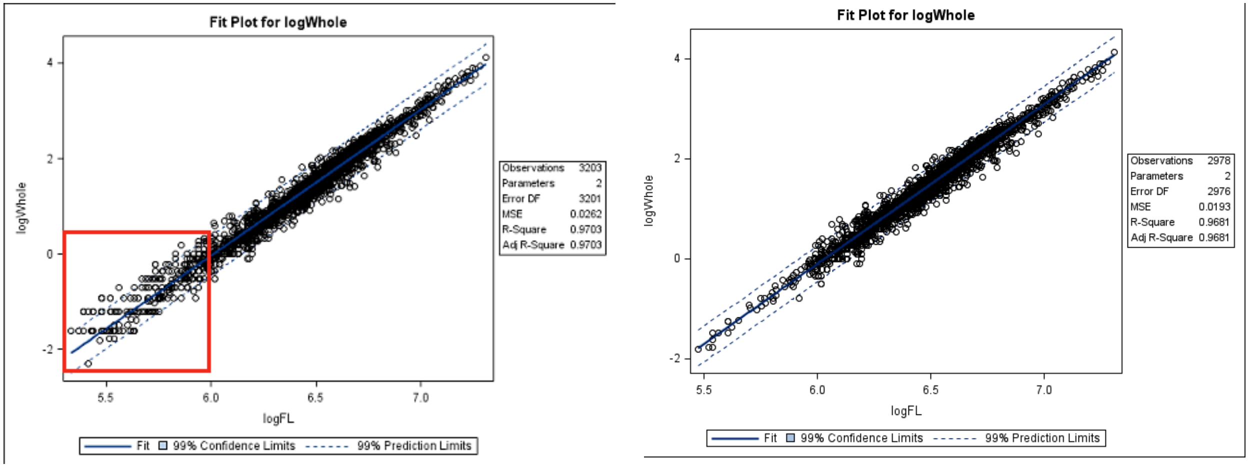

c. Log_e(Gutted wt.) vs. Log_e($\text{TL}_{\text{natural}}$) predicted fit

d. Log_e(Gutted wt.) vs. Log_e($\text{TL}_{\text{natural}}$) residual plot


e. Log_e(Gutted wt.) vs. Log_e(TL_{max}) predicted fit

f. Log_e(Gutted wt.) vs. Log_e(TL_{max}) residual plot

Figure 3. Predicted fits for $\log_e$-transformed FL (mm)- whole wt. (kg) regressions before and after deleting MRFSS/MRIP specimens < 400 mm.



a. Log_e(Gutted wt.) vs. Log_e(FL) predicted fit with all specimens. Highlighted box frames specimens < 400 mm FL

b. Log_e(Gutted wt.) vs. Log_e(FL) predicted fit after deleting MRFSS/MRIP specimens with FL< 400 mm.

c. Comparison of back-transformed regression lines before and after excluding MRFSS/MRIP specimens < 400 mm. Inset shows locations of all excluded specimens in relation to the SEDAR 48 predicted regression line, and most were within the 99% prediction interval.