1-1-1988

Specification and Estimation of Weight-Length Relationships and Fishery Regulations

William D. DuPaul
Virginia Institute of Marine Science

James E. Kirkley
Virginia Institute of Marine Science

SPECIFICATION AND ESTIMATION OF WEIGHT-LENGTH RELATIONSHIPS AND FISHERY REGULATIONS

BY

WILLIAM D. DUPAUL AND JAMES E. KIRKLEY[1]

Marine Resource Report No. 88-2

---

[1]William D. DuPaul is Director, Marine Advisory Services, Virginia Institute of Marine Science, College of William and Mary. James E. Kirkley is Assistant Professor, School of Marine Science, Virginia Institute of Marine Science, College of William and Mary, Gloucester Point, VA 23062

Introduction

The customary practice of determining the statistical relationship between weight and length of individual fish is to estimate the logarithmic version of the allometric weight-length equation by ordinary-least-squares (OLSQ). The estimation method and application for prediction have both been criticized.

Pienaar and Thomson (1969) demonstrated that the assumption of a multiplicative error term, which is required for OLSQ estimation, may be incorrect. An additive error was suggested. Ricker (1973) offered that the prediction of weight conditional on length should be based on the geometric-mean (GM) functional regression. More recently, Cohen and Fishman (1980) argued that the traditional approach does not acknowledge stochastic variation and that the conditional mean or forecasting equation should include the residual variance in multiplicative form. Last, it has been suggested that the estimated relationship may vary over time, space, and sex (Manzer 1972, Ricker 1975; Gulland 1976; Cohen and Fishman 1980).

Thus, there is substantial evidence to recommend against using the traditional allometric model and estimation method. Yet, the practice is widespread in current research and in the determination of age-at-capture regulations. There is a need for an examination of the criticisms of the traditional approach.

However, there is also a need for an examination of alternative functional specifications which may more closely conform to expectations about the relationship between weight and length. The traditional allometric model imploses the implasible conditions that weight globally increases as length increase, and the precentage change in weight associated with a one-percent increase in length is a constant and independent of the length. Additional research on the specification of functional form does not appear to have been a major concern (Richards 1959; Sillman 1967).

In this paper, the traditional allometric model is further examined by applying methods presented in the literature to estimate the weight-length relationship for mid-Atlantic sea scallops (Placopecten magellanicus) during August of 1987. The weight-length relationship is also estimated for alternative functional specifications. Predictions based on different methods of estimation and functional forms are subsequently made and examined. The standard allometric approach is then related to the design of fishery regulations and the need for confidence and tolerance intervals is illustrated.

The Allometric Weight-Length or Traditional Model

The traditional approach or allometric model equates weight (W_i) of an individual fish to the product of a constant (α) and length (L_i) raised to an unknown value (β):

$$(1) \qquad W_i = \alpha L_i^{\beta} \exp(U_i)$$

where U_i is a random error assumed to be $N(0, \sigma_u^2)$. Model (1) is intrinsically linear. That is, the model is nonlinear in the variables but linear in the parameters. Thus, it is possible to convert model (1) into an equation suitable for estimation by linear regression.

The conversion or most common transformation is the log of base e transformation. Model (1) is then estimated by applying linear regression methods to

$$(2) \qquad \ln W_i = \ln \alpha + \beta \ln L_i + U_i.$$

If the error term is distributed as $N(0, \sigma_u^2)$, Eq. (2) may be estimated by ordinary-least squares. This estimator is the best-linear-unbiased estimator for $\ln \alpha$ and β.

Problems with the Traditional Approach

The literature, however, suggests that the OLSQ version of model (1), Eq. (2), may have several problems or be inappropriately used to estimate the conditional mean weight. First, there is the potential problem that the transformation of variables required for the linear regression model may overcompensate for the anticipated increase in the variance of weight for larger fish (Pinaar and Thomson 1969). In this case, ordinary-least-squares will not yield minimum variance estimates. Second, the estimator of α, which is exp(ln α), is not an unbiased estimate; it is, though, a consistent estimate (Kelejian and Oates 1981). Third, model (1) is often incorrectly used in that the conditional median is used to predict weights (Cohen and Fishman 1980). Fourth, Ricker (1973) has argued that GM functional regression should be used to predict weight. Jolicoeur (1975), however, argues that the standard regression is preferred.

Error Specification

Pienaar and Thomson suggest that the allometric model with an additive error term may mitigate the problem of heteroscedasticity or non-constant variance of the logarithm of weight. The alternative model of Pienaar and Thomson is

$$(3) \qquad W_i = \alpha L_i^{\beta} + U_i$$

where U_i is $N(0, \sigma^2)$. The model given by Eq. (3) is intrinsically nonlinear and must be estimated by nonlinear methods. Pienaar and Thomson, however, demonstrate how the linearization technique of Draper and Smith (1967) may be used to estimate Eq. (3).

Alternatively, Goldfelt and Quandt (1971) offer procedures for estimating the multiplicative function when both multiplicative and addditive errors are possible. They also note that a multiplicative error is likely in the case of omitted variables or if the constant (α) varies over individual observations. An additive error term is justified if the only source of a stochastic term in model (1) is the fact that weight, but not length, is observed with error.

Heteroscedasticity

The method outlined by Pienaar and Thomson is not the only procedure for dealing with nonconstant variance. Amemiya (1973) demonstrated that if the dependent variable ($\ln W_i$) has a lognormal distribution, its variance is proportional to the square of its mean. That is, given Eq. 2, the variance of the logarithm of weight is proportional to the square of its mean:

$$(4) \qquad \mathrm{Var}\,(\ln W_i) = \theta^2 (\ln \alpha + \beta \ln L_i)^2$$

Estimation procedures are summarized in Amemiya. Additional forms of heteroscedasticity may also characterize the residuals; methods for testing and estimating are further discussed in Maddala (1977).

Use of the OLSQ Allometric Model

Estimates of the allometric model, Eq. (2), are used to obtain estimates of weight conditional on length:

$$(5) \qquad W_i = \exp(\ln \alpha)\, L_i^{\beta}$$

In the literature, these estimates are often referred to as average or mean weights (Haynes 1966; Pienaar and

Thomson 1969; Serchuk and Rak 1983; MacDonald and Bourne 1987). However, the OLSQ estimates of Eq. (5) do not directly yield the conditional mean weights. They are conditional median weights. The conditional mean weight for the OLSQ estimates of model (1) is given by

$$(6) \qquad W_i = \exp(\ln \alpha)\ L_i^{\beta}\ \exp(\sigma^2/2)$$

where σ^2 is the estimated residual variance (Goldfeld and Quandt 1972; Cohen and Fishman 1980). Interestingly, $\exp(\sigma^2/2)$ appears to be quite close to one for many weight-length relationships, and thus, implying that the conditional median may not be significantly different than the conditional mean. This may not, though, apply to the case of ~~larger fish in which there is~~ extreme variability in weight and length.

GM Regression

Ricker (1973, 1975) states "The GM functional regression should be used rather than the predictive regression which has commonly been employed in the past". The rationale for functional regression is that the values of the independent variables are subject to natural variability and are a symmetrical sample from a real or imaginary distribution. The GM regression does not appear to have been widely used to examine the weight-length relationship. Jolicoeur, ~~however,~~ moreover, argues strongly against this approach. In view that the GM approach does not appear to have been widely used and Jolicoeur disputes its applicability, it is not further considered in this paper.

Specification of the Weight-Length Relationship

In practice, the traditional allometric model and OLSQ estimate appear to provide a reasonable empirical estimate of the weight-length relationship over a wide variety of sample data (Pienaar and Thomson 1969; Manzar 1972, Ricker 1975; Cohen and Fishman 1980; Serchuk and Rak 1983; MacDonald and Bourne 1987). However, the traditional approach imposes unnecessary restrictions on the relationship between weight and length.

First, model (1) imposes the global condition of a constant rate of increase in weight for increases in shell size. Second, if the estimated parameter, ß, exceeds one in value, a one-percent increase in length will always yield more than a one-percent increase in weight, regardless of the length. Third, model (1) in the absence of a length constraining equation imposes the condition of no maximum.

Several alternative specifications which do not a priori impose the above restrictions are available. Three possible specifications are the polynomial, translog, and transcendental. These are as follows:

(7) Polynomial--$W_i = \alpha_1 L_i + \alpha_2 L_i^2 + \ldots + \alpha_k L_i^k$

(8) Translog----$W_i = \alpha L_i^{\beta_1} L_i^{\beta_2 \ln L_i}$

$L_i^{\beta_2}$

(9) Transcendental--$W_i = \alpha L_i^{\beta_1} \exp(\beta_2 L_i)$

All three functions allows for different growth rates over different lengths. The specifications, like the standard allometric model, may be estimated by ordinary least squares. The translog and transcendental may also be estimated by nonlinear methods for an additive error. The translog and transcendental, however, have the same statistical limitations as the standard allometric model (e.g., heterocedasticity, biased constant, and the need to multiply by the exponential value of the variance).

An alternative approach for determining functional form is the method of Box-Cox (1962). This approach requires estimation of data transformations which minimize the maximum likelihood function; alternative methods of estimation are discussed in Spitzer (1982). In the case of the weight-length equation, the following equation might be specified and estimated:

(10) $$\left(W_i^t - 1\right)/t = \alpha + \alpha_1\left(\left(L_i^t - 1\right)/t\right)$$

where t is the transformation. If t equals zero, the logarithmic transformation is implied; if t equals one, the standard linear model applies. In addition, each variable in a model can have different values of t (e.g., $t_1, t_2, \ldots, t_n$ and n equals the number of variables in the equation). The procedure has been widely applied in economic analysis. However, it is quite difficult to derive the conditional mean for all values of t other than zero or one (Smallwood and Blaylock 1986). This approach is not further investigated in this paper. It is introduced only because it is one alternative for determining the functional form.

Estimating the Weight-Length Relationship

In this section, the weight-length relationship for Mid-Atlantic sea scallops is estimated using data obtained from an on-going sea scallop study (DuPaul and Kirkley 1987). The data are for August of 1987. Spatial and temporal differences or the need to consider sex are not considered. The need to consider differences over time, area, and sex were discussed in Posgay (1953) and Manzer (1972); they are important but ~~are~~ beyond the concerns of this paper.

The emphasis of this section is on demonstrating alternative estimation procedures and specifications which may have important ramifications for predicting weight conditional on length. However, there are many other important aspects of sea scallops and statistical analysis which also have important ramifications for predicting weight. These other aspects are also excluded from this paper. All estimates presented in this section were done on a 640 K RAM personal computer using either SST (Dubin and Rivers 1986) or LIMDEP (Greene 1986).

The traditional allometric model, Eq. (1), is estimated by ordinary-least-squares subject to the standard assumptions about the error term. Similarly, the translog and transcendental are estimated. The polynomial, Eq. (7), requires no data transformations other than raising the ~~value~~ variable length to a power; estimation is accomplished by OLSQ. Estimates and associated statistics appear in table 1.

The results in table 1 indicate that the mathematical elasticitiy of weight conditional on length, as given by

$$(11) \qquad \delta \ln W_i / \delta \ln L_i$$

is not likely to be a constant. The hypothesis that $\beta_2 = 0$ is rejected at any reasonable level of significance.

The latter three forms also allow for a maximum weight; however, none of the estimated maximums appear reasonable. The transcendental form yields the only estimate which may be possible:

$$(12) \qquad \text{Transcendental--}\delta W_i / \delta L_i = W_i [(\beta_1 / L_i) + \beta_2] = 0$$

$$\text{Maximum } L_i = 332.6 \text{ mm and } W_i |_{L_i = 332.6} = 152.4$$

Norton (1931) noted a maximum shell size of 230 mm.

Alternatively, the literature suggests that the three logarithmic specifications may have additive errors rather than multiplicative errors. In this case, estimation requires a nonlinear approach. As an example, the standard allometric model and the transcendental model are estimated using the linearization technique of Draper and Smith (1967). This was necessary since it was not possible to achieve convergence using the standard maximum likelihood routine available in SST; this may be due to not having a math coprocessor. The nonlinear estimates are as follows:

$$(13) \qquad W_i = .00007676746\ L_i^{2.705}$$

$$(14) \qquad W_i = .00000016905\ L_i^{4.37226}\ \exp(-.0151983\ L_i)$$

The nonlinear estimates are not offered as the preferred estimates. They did not, in fact, provide as good an estimate of weight as did the OLSQ estimates. As a consequence, additional attention is not given to the nonlinear estimates. However, it should be remembered that there may be situations in which the nonlinear estimates provide better estimates of weight.

A different problem for estimating the coefficients of the standard allometric and transcendental models is that of heteroscedasticity. As shown by Amemiya (1973), if the dependent variable has a lognormal distribution, estimation should be by generalized-least-squares (GLS). Parameter estimates obtained from Amemiya's algorithm and available in LIMDEP for the allometric and transcendental models are

$$(15) \qquad W_i = .00002103758\ L_i^{2.98212}$$

$$(166.79) \qquad (205.45)$$

(16) $$W_i = .0000004941632\ L_i^{4.05959}\ \exp(-.0120377\ L_i)$$

(34.83) (33.23) (8.13)

where numbers in parentheses are the t-statistics for GLS estimates of the logarithmic transformations.

Estimating Weight Conditional on Length

The major purpose of estimating a weight-length relationship is to estimate the expected weight of a fish conditional on a given length (i.e., $E(W_i | L_i = L_i^o)$). In practice, it appears that when the standard allometric model is used and estimated by OLSQ, the conditional median rather than the conditional mean is estimated. For example, Pienaar and Thomson (1969) title a table in their paper as "Estimated mean weights of Pacific cod"; the first column is the mean weight as estimated by the OLSQ estimates of the standard allometric model. Simarly, estimates from Serchuk and Wood (1981) used in the management plan for sea scallops implies that the OLSQ estimates without adjustment for the residual variance yield average or mean meat weights for a given length (New England Fishery Management Council 1982, p. 28).

The conditional median for the standard allometric model and the transcendental are given by

(17) $$W_i = .00003598746\ L_i^{2.86672}$$

(18) $$W_i = .0000004688935\ L_i^{4.07722}\ \exp(-.012259\ L_i)$$

In comparison, the conditional mean or prediction equations are given by

$$(19) \qquad W_i = .000033598746 \, L_i^{2.86672} \, (1.009642)$$

$$(20) \qquad W_i = .0000004688935 \, L_i^{4.07722} \exp(-.012259 \, L_i) \cdot (1.009495)$$

where the constants in parentheses equal $\exp(\sigma^2/2)$. The conditional expectation of the polynominal simply equals the estimated equation; no adjustment is necessary if the standard assumptions about the error term are correct.

Estimates of weight conditional on shell sizes of 80, 89, 96.03 (mean length), 100, 130, and 150 mm. are in table 2. Similar estimates of weight using the GLS coefficients appear in table 3. Heteroscedasticity in the polynominal model was not considered.

As illustrated in tables (2) and (3), the conditional median and mean for the OLSQ and GLS estimates of the standard allometric model and the transcendental model are quite close. However, all models and estimates tend to overestimate weight for large lengths.

In terms of model performance, there are no universally accepted criteria for model selection (Johnston 1984). If making point predictions over the range of observed data is the only objective of estimating the relationship, a third or fourth order polynomial will likely provide good point predictions (MacDonald and Bourne 1987). Moreover, if multicollinearity presents a problem for estimating the individual parameters, it does not present a problems for prediction when the value of all right hand side or independent variables are known (Kelejian and Oates 1981).

Hypothesis testing, however, usually requires that the error be $N(0,\sigma^2)$ and the estimators be minimum variance. Thus, both multicollinearity and heteroscedasticity present problems for hypothesis testing.

Regulations and Confidence and Tolerance Intervals

The weight-length relationship is often used to design age-at-capture regulations. Consider the U.S. sea scallop fishery in which it was determined that a minimum shell size of 3.50 inches was necessary to meet the objectives of the management plan. Alternatively, a 30 meat-count (number of meats per pound) was believed to be equivalent to a 3.50 inch shell and used to regulate fishermen who shucked at sea.

In 1986, the fishery, though, appeared to be dominated by the 82 and 83 year classes which were between 3.50 and 3.75 inches. The industry experienced substantial compliance problems since enforcement did not monitor or consider the median or average meat count for a given length. In addition, there was an inequity between shell stockers and shuckers since the shuckers could not legally harvest the 3.50 inch scallops which did not yield a total of 30 meats per pound.

Management allowed a 10-percent tolerance because of the difficultly in precisely measuring meat-count at sea. However, there were still several violations. In actuality, a violation is a violation. There is no statistical analysis which can be used to demonstrate that a violation was not, in fact, a violation. However, managers should be aware that a given shell size may yield many different weights, or in the case of scallops, several meat counts. Alternatively, management should be aware that a given meat-count may occur for several shell sizes. These occurences should be considered in designing the tolerances for the regulations.

A possible way to better consider the tolerances is to examine the confidence and tolerance intervals of the point estimates. In the case of the sea scallop fishery, a 3.50 inch shell height yields a point estimate of 32.63 meats per pound for the standard allometric model. However, the 95-percent confidence interval as calculated by the formula in Kmenta (1971) is 22.99 to 46.28. In addition, the interval is not symmetric as shown in Haynes (1966); this is because predictions based on the antilog will be asymmetric. In any event, an allowable percentage adjustment based on the confidence interval is in excess of 50-percent.

In comparison, consider a regulation designed to yield a 30 meat-count and shell size equivalency. Based on the estimates for the standard allometric model, the equivalent point estimate of shell size is 3.60 inches. However, the 95-percent tolerance interval, estimated by the method of Fieller (1944), yields an interval of 3.04 to 3.74 inches. In this case, the 30 meat-count regulation might require a minimum shell size of 3.60 inches with a 15-percent tolerance.

The example, while seemingly simple and limited by inadequate attention to spatial and temporal variability, is consistent with the manner in which the sea scallop regulations were determined. That is, a standard allometric model~~s~~ was estimated by ordinary-least-squares using survey data which typically covered several years but few months. The spatial and temporal variability were ignored. Possible problems shuch as heteroscedasticity also were not considered. Estimated weights were based on the conditional median and not the conditional mean. Adjustments or the calculation of tolerance limits were not based on statistical criteria such as the confidence and tolerance intervals. The procedures, thus, used to determine the regulations were quite limited in scope.

Table 1. Parameter estimates and associated statistics for four weight-length specifications

Specification[a]	Parameter estimates						
	α	β_1	β_2	β_3	β_4	R^2	N
Standard allometric	-10.23	2.867				.851	2371
	(90.97)[b]	(116.21)					
Translog	-22.42	8.19	-.58			.853	2371
	(11.20)	(9.38)	(6.10)				
Transcendental	-14.57	4.08	-.01			.853	2371
	(20.35)	(20.51)	(6.14)				
Polynomial		.31	-.01	.0001	-.0000004	.866	2371
		(3.06)	(3.51)	(4.81)	(4.83)		

[a]The four forms are

(1) $\ln W_i = \alpha + \beta_1 \ln L_i$
(2) $\ln W_i = \alpha + \beta_1 \ln L_i + \beta_2 (\ln L_i)^2$
(3) $\ln W_i = \alpha + \beta_1 \ln L_i + \beta_2 L_i$
(4) $W_i = \beta_1 L_i + \beta_2 L_i^2 + \beta_3 L_i^3 + \beta_4 L_i^4$

[b]The numbers in parentheses are t-statistics for the OLSQ estimates.

Table 2. Estimates (OLSQ) of weight of sea scallops conditional on selected lengths

Length (mm)	Conditional median		Conditional mean		
	Standard allometric Eq. (17)	Transcendental Eq. (18)	Allometric Eq. (19)	Transcendental Eq. (20)	Polynomial Eq. (21)
80	10.275	10.108	10.374	10.204	10.253
89	13.948	13.981	14.082	14.114	14.037
93.06 (mean)	17.340	17.490	17.510	17.650	17.603
100	19.481	19.649	19.669	19.836	19.838
130	41.329	39.646	41.727	40.022	40.216
150	62.290	55.605	62.890	56.133	53.748

*Observed mean weights are 9.64 (80 mm), 13.79 (89 mm), 18.19 (sample mean), 19.77 (100 mm), 44.5 (130 mm), 48.5 (150 mm).

Table 3. Estimates (GLS) of weight of sea scallops conditional on selected lengths

Length (mm)	Conditional median		Conditional mean	
	Standard allometric	Transcendental	Allometric	Transcendental
80	10.000	10.032	10.017	10.045
89	13.749	13.877	13.767	13.895
93.06 (mean)	17.249	17.361	17.272	17.384
100	19.465	19.510	19.491	19.535
130	42.576	39.444	42.632	39.495
150	65.247	55.427	65.333	55.498

*Observed mean weights are 9.64 (80 mm), 13.79 (89 mm), 18.19 (sample mean), 19.77 (100 mm), 44.5 (130 mm), 48.5 (150 mm).