Aquaculture Nutrition **2000 6;** 91–102

Experimental design, statistical analysis and modelling of dietary nutrient requirement studies for fish: a critical review

K.D. SHEARER
Northwest Fisheries Science Centre, NMFS, NOAA, Seattle, WA, USA

Abstract

A limited survey of published reports on dietary nutrient requirement estimates for fish (three journals, 46 papers) indicates that broken-line analysis and analysis of variance (ANOVA) are often used to estimate nutrient requirements from dose–response data. The application of regression models using published treatment mean values to re-evaluate estimates was possible using 33 of these reports. Re-evaluation suggests that the broken-line method and ANOVA frequently underestimate the requirement. Regression produced estimates that averaged approximately twice, but were up to five times the published requirement. Additional problems that prevented re-evaluation or produced errors in the original estimates were: failure to include nutrient levels high enough to produce a maximum response, failure to space nutrient input levels closely enough to adequately model the dose–response relationship, an apparent failure to screen data before analysis, and insufficient model diagnosis. Examples from the literature are presented to illustrate how design, method of analysis and the choice of model affect the requirement estimate. The effects of measurement frequency and the experiment duration on the resulting requirement estimate are discussed. A set of protocols is presented to help improve nutrient requirement estimates.

KEY WORDS: experimental design, modelling, requirements, statistical analysis

Received 18 May 1999, accepted 17 September 1999

Correspondence: Karl D. Shearer, Northwest Fisheries Science Center, NMFS, NOAA, 2725 Montlake Boulevard, East, Seattle, WA 98112, USA. E-mail: karl.d.shearer@noaa.gov

Introduction

Dietary nutrient requirements in fish are usually estimated empirically by feeding graded levels of a specific nutrient (dose), in a basal diet containing a deficient level of that nutrient, and then measuring growth, feed intake, body nutrient stores or other variables (response). The experiment is usually conducted for a sufficient period of time to produce differences in the response variable. The dose–response relationship is then examined using one or more methods, and the nutrient requirement is estimated from the level that produces the maximum response. The general design of these experiments has remained relatively unchanged since the early days of fish nutrition research (Wilson 1994). Methods available for examining the dose–response relationship have, however, evolved owing to improvements in statistical methodology, availability of personal computers and inexpensive statistical software, and the development of new mathematical models to evaluate dose–response relationships. A number of authors have addressed statistical analysis of nutrient requirement studies in fish and other animals (Zeitoun *et al.* 1976; Mercer *et al.* 1978; Robbins *et al.* 1979; Mercer 1982, 1992; Robbins 1986; Baker 1986; Mercer *et al.* 1989; Cowey 1992). Baker (1986) stressed that the choice of statistical method may have a large effect on the estimated value of the requirement. Despite the fact that these authors agreed that regression produces a more accurate estimate than alternative methods of examining the dose–response relationship, neither this technique, nor the newer dose–response models (Mercer *et al.* 1989, 1993), have been consistently used in fish nutrition studies.

The purpose of this paper is to present the results of a limited survey of the methods of dose–response analysis used during the last few years in nutrient requirement studies with fish, and to show the effect of re-evaluation of the data from these studies, using regression, on the estimated nutrient requirements. Examples of some common design and analysis problems are shown and protocols for preventing these problems are presented. The approach taken is conceptual rather than computational, and references discussing pertinent topics in detail are provided to aid in rectifying specific problems.

Materials and methods

Selection of papers for re-evaluation

The papers chosen for review and re-evaluation were selected from three journals which publish fish nutrition papers: *Aquaculture* (Elsevier Publishers BV, Amsterdam, The Netherlands), *Aquaculture Nutrition* (Blackwell Science, Oxford, UK) and *Journal of the World Aquaculture Society* (World Aquaculture Society, Baton Rouge, LA, USA). The papers chosen presented the results of studies where a single macro- or micro-nutrient was evaluated. For papers presenting more than one experiment, only the experiment resulting in an optimum dietary level was re-evaluated. For each paper, the method of evaluation of the dose–response data was noted, along with the optimum dietary inclusion level of the nutrient, as determined by the chosen method. The data used for the re-evaluation of the relationship between dietary inclusion level and the response were obtained from values (dose–response pairs) tabled in the paper. In all cases, only treatment mean values were available. Some papers contained estimates of experimental error (standard deviations or standard errors of the means). It was also noted whether or not the results were presented graphically.

Re-evaluation

From the reviews set forth above, three models were selected as best suited for analysing nutrient requirements. An attempt was made to fit these three models to each data set: the five-parameter saturation kinetics model (5-SKM, Mercer *et al.* 1989):

$$r = \frac{b(K_{0.5})^n + R_{max}I^n + bI^{2n}/(K_s)^n}{(K_{0.5})^n + I^n + I^{2n}/(K_s)^n}, \qquad (1)$$

the four-parameter saturation kinetics model (4-SKM, Mercer *et al.* 1984)

$$r = \frac{b(K_{0.5})^n + R_{max}I^n}{(K_{0.5})^n + I^n}, \qquad (2)$$

and a quadratic model (second-order polynomial)

$$r = B_0 + B_1I + B_2I \qquad (3)$$

where r = physiological response, I = dietary concentration, b = intercept on r axis, R_{max} = maximum theoretical

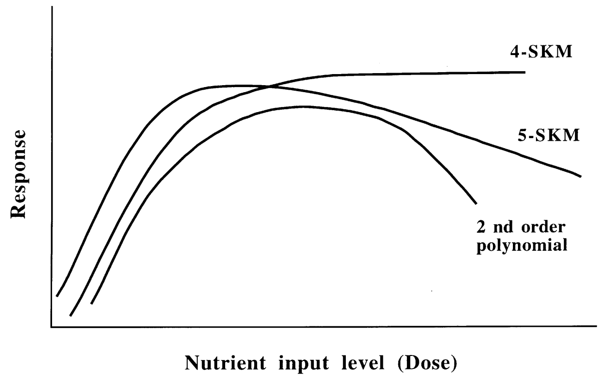

Figure 1 Dose–response curves used to re-evaluate data from nutrient requirement studies; a second-order polynomial, four-parameter (4-SKM; Mercer 1982) and five-parameter (5-SKM; Mercer *et al.* 1989) saturation kinetics models.

response, n = apparent kinetic order, $K_{0.5}$ = concentration for 1/2 of (R_{max} + b), K_s = inhibition constant, B_1 and B_2 = regression coefficients and B_0 = intercept on the response (Y) axis

Equations 1 and 2 can also be written for use in a curve fitting program as:

$$(1) r = (b * (a\hat{}n) + d * (I\hat{}n) + ((b * (I\hat{}(2 * n))))/(c\hat{}n)/((a\hat{}n) + (I\hat{}n) + ((I\hat{}(2 * n))/(c\hat{}n))$$

and

$$(2) r = (b * (a\hat{}n) + d * (I\hat{}n))/((a\hat{}n) + (I\hat{}n))$$

where: a = $K_{0.5}$, c = K_S, d = R_{max}, and n, b, r and I are as above.

The three models were fitted using SigmaPlot (1992 version 5, Abacus Concepts, Berkeley, CA, USA) which uses the Marquardt–Levenberg algorithm that finds the parameters that minimize the sum of squared differences between the values of the observed and predicted values of the dependent variable. The 5-SKM produces an asymmetric curve; the 4-SKM, a sigmoid-shaped curve which reaches a plateau; and the second-order polynomial, a symmetric parabola (Fig. 1). The magnitude of the differences between reported and re-evaluated requirements (estimated at the maximum response) was determined by dividing the re-evaluated requirement by the published requirement (or optimum) and converting this to a percentage. For example, re-evaluation = 42 mg kg^{-1}, published value = 24 mg kg^{-1}; the difference is 42/24 = 175%, so that 100% indicates no difference. Where the requirement was listed as a range, the difference was based on the midpoint of the range.

Table 1 Journals used for selection of articles to determine the statistical methods used to evaluate dose–response data from dietary nutrient requirement studies

Journal	Issues	Dates	Number of papers
Aquaculture	109–154	1993–97	32
Aquaculture Nutrition	1–3	1995–97	5
Journal of the World Aquaculture Society	22–28	1991–97	9
Total			46

Table 2 The results of a limited survey of statistical methods used to estimate the dietary nutrient requirements of fish

Method used	Number of cases[1]
Broken line	22
Regression:	13
Quadratic	6
Sigmoid	3
SKM-5 parameter[2]	2
Linear	2
ANOVA	13
Nonparametric	1

[1] More than one method used in several cases.
[2] Five-parameter saturation kinetics model (Mercer *et al.* 1989).

Results of the survey and re-evaluation

A total of 46 papers fit the selection criteria (Table 1). The most common statistical method used (22 cases) was the broken-line plot, followed by analysis of variance (ANOVA) with various multiple-mean comparison tests, and regression (13 cases each) (Table 2). When regression was used, the requirement was estimated using the second-order polynomial (eight cases), a curve that plateaued (three cases) or the 5-SKM (two cases). Two studies fitted a straight line and one used a nonparametric test. In several studies more than one method was used. Two papers were excluded from the re-evaluation because the treatments produced no differences in the response variables examined. The results were presented graphically in only 20 of the 46 papers.

Eight of the papers contained data sets that could not be modelled with any of the three models used for the re-evaluation. The data in these papers either showed no discernible pattern, or the responses were best modelled with a straight line. In the papers where re-evaluation was possible (30 cases), the second-order polynomial (Equation 3) most often provided the best fit based on residual analysis (18 cases), followed by the 4-SKM (Equation 2) (eight cases) and the 5-SKM (Equation 1) (four cases).

Table 3 Results of re-evaluation of nutrient requirement estimates using regression and the best fitting of three regression models[1]

Original method	n	Re-evaluation estimate[2]/ published estimate	Range
Broken line	18	200%	120–500%
ANOVA	10	195%	120–300%
Regression	5	126%	100–230%

[1] See the text for additional details on the models used and the methods of re-evaluation.
[2] Mean of all re-evaluations using a specific original method (original estimate/re-evaluation estimate) $\times$ 100. A value of 100% means equal estimates.

Re-evaluation of the studies that used broken-line analysis (Table 3) produced estimates averaging approximately 200% of the published values (range 120–500%). Re-evaluation of the studies that used ANOVA to determine the requirement produced estimates that were, on average, about 195% of the published estimate (range 120–300%). On average, estimates for those using curvilinear regression were 126% of published values (range 100–230%). This was entirely accounted for by one study, where the re-evaluated requirement was about 230% of the reported value; the remaining four studies showed no difference between the reported estimate and the re-evaluation estimate.

Discussion

It should be stressed that the re-evaluations were conducted on treatment mean values and not on the original data. It was therefore impossible to perform thorough model diagnosis. The present work therefore violated many of the protocols presented in the following discussion. The purpose of the re-evaluations was simply to demonstrate that alternative, often more appropriate, methods of analysis are available and that these methods may lead, in many cases, to higher requirement estimates.

Why are there often such large differences among requirements estimated by the various methods? An understanding of this problem can be achieved, in part, by examining the basic statistical elements of a nutrient requirement study. A researcher is concerned with three distinct elements of analysis: (1) proper design of the experiment, (2) collection of accurate and precise data, and (3) selection of the appropriate methods and models for statistical analysis (Zolman 1993). The three are interconnected. Use of the appropriate statistical analysis can not salvage a poorly designed study and inappropriate statistical analysis will produce erroneous conclusions, even when the experiment is well designed and the data are of good quality. Some problems associated with each of these three areas are best illustrated using examples from the papers reviewed and from additional fish nutrition literature.

Design problems

As indicated earlier, the basic plan of a nutrient requirement study involves feeding graded levels of a nutrient in a basal diet that is nutritionally balanced but contains as low a level as possible of the nutrient under study. One or more responses, usually growth, is chosen and the relationship between the independent (dose) and dependent (response) variables is examined. If sufficiently low and high levels of the nutrient are fed, a typical dose–response relationship is produced (Fig. 1). At a very low level of input there may be no growth, or even weight loss. As the nutrient level is increased, weight gains improve. At the optimum level of inclusion, maximum weight gain is observed. At high levels of inclusion the nutrient may become toxic or suppress feed intake, or the metabolic cost of excretion of the excess nutrient reduces growth.

One common design problem is that the highest nutrient dose used is below the level capable of producing the maximum response. A study designed to determine the protein level to optimize weight gain and feed efficiency in golden shiners *Notemigonus crysoleucas* (Lochmann & Phillips 1994) illustrates this problem. The authors fed five levels of protein, between 21.2% and 34.5% of the diet, to triplicate groups of fish for 8 weeks, and then measured weight gain and feed efficiency. They concluded, based on ANOVA, that this species required only 29% protein for optimal performance. The data were not presented graphically. If the data are graphed (Fig. 2), it can readily be seen that the highest protein level fed was too low to determine the level for maximum growth; that is, the levels of nutrient fed were still in the ascending portion of the dose–response curve. The same problem was found in other papers that were re-evaluated (for example: Brecka *et al.* 1995; Carmona-Osalde *et al.* 1996; Webster *et al.* 1997).

Another design problem results from spacing input levels too far apart, preventing accurate determination of the requirement (Kim 1993). This study was designed to determine the phenylalanine requirement of the rainbow trout *Oncorhynchus mykiss*. Seven levels of phenylalanine (0.26, 0.35, 0.45, 0.55, 0.65, 0.75 and 1.75%) were fed for 6 weeks to four replicate groups, and weight gain was determined. The requirement was estimated using broken-line analysis and a sigmoid regression relationship; the requirement was

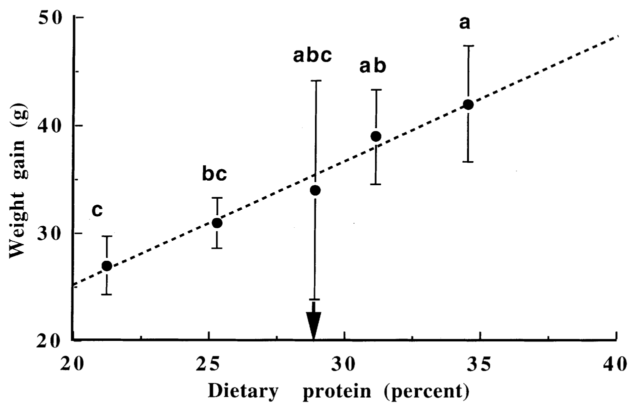

Figure 2 An example illustrating insufficient treatment range; the maximum level of nutrient was not high enough to achieve maximum response. The protein requirement associated with the maximum response was estimated at about 29% based on analysis of variance and multiple mean comparison (data points with different letters were reported as significantly different using ANOVA and a multiple-mean-comparison test with $P < 0.05$). The design is inadequate, there appears to be heteroscedasticity, an inappropriate statistical test was used and test power appears low because of few replicates and the large variance. The dotted line is shown to suggest the possible relationship between the variables. (Data from Lochmann & Phillips 1994).

estimated to be 0.7% of the diet. However, because of the large gap between the inclusion levels of 0.75% and 1.75%, the optimum phenylalanine level may not have been accurately estimated. For example, a second-order polynomial, which fit the data equally well, indicated that the requirement might be as high as 1.1% (Fig. 3).

It should be clear from the preceding examples that the dose–response observations, and therefore the estimated requirement, are very dependent on study design. This point will be examined subsequently in more detail. It appears that very few studies include levels of nutrient input high enough to suppress growth. Yet it is desirable to do so in order to ensure that the maximum response has been achieved. The failure to include sufficiently high nutrient levels is a common design flaw which results in an experiment which cannot be salvaged even by the best statistical test.

Data quality

In studies that are properly designed and conducted, the data generated show the typical dose response curve (Arzel *et al.* 1995; Andersen *et al.* 1996; Rodehutscord 1996). However, in some studies it is very difficult to discern a pattern in the data. A study designed to determine the optimum protein requirement of young Arctic charr *Salvelinus alpinus* (Gurure

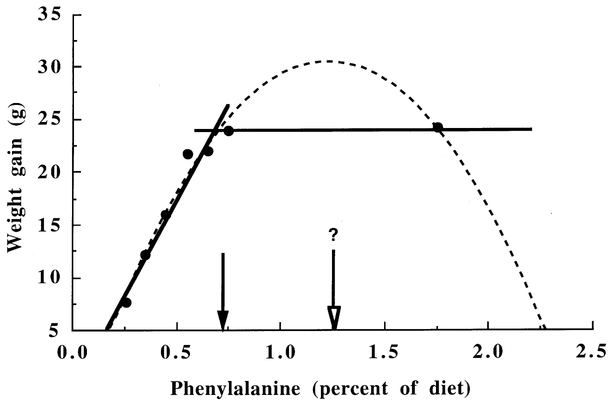

Figure 3 An example illustrating the potential effect of large gaps between nutrient levels. The study was designed to determine the phenylalanine requirement of rainbow trout. Using a broken-line plot the requirement was estimated to be 0.7% of the diet. A second-order polynomial (dotted line) fits the data equally well and suggests that the requirement may be considerably higher, but the lack of input levels between 0.75 and 1.75% prevents proper evaluation of this model. (Data from Kim 1993).

et al. 1995) illustrates this problem: no curve adequately fits these data points (Fig. 4). It appears that an additional source of variation was influencing the response. A similar problem was observed with another data set (Olivera-Novoa *et al.* 1996). Residual analysis and outlier detection should be performed. Moreover, the fact that a significant regression coefficient (r^2) was obtained in these studies indicates only that there is a discernible pattern in the data, not that the correct model was applied.

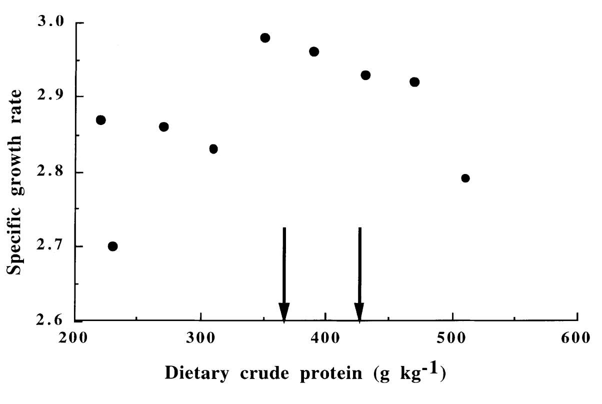

Figure 4 The relationship between dietary protein and specific growth rate of Arctic charr *Salvelinus alpinus*. The requirement estimates based on the average of broken line, second-order polynomial and the 5-SKM are shown as a range (between the arrows). Re-evaluation failed to produce an adequate fit to this data. It appears that an additional variable was confounding growth. (Data from Lochmann & Phillips 1994).

Analytical method

ANOVA: ANOVA is frequently used to estimate nutrient requirements from dose–response data. The use of ANOVA to analyse this type of data is inappropriate for a number of reasons (Kirk 1982; Dawkins 1983; JSFA 1988). It is inappropriate when multiple levels of an independent variable are used. When ANOVA is used, nutrient levels are treated as discrete rather than continuous, so that the optimum nutrient level is stated as a range between two input levels. Test power is often low, even when variances are small, as only two or three replicates are normally used in most fish nutrition studies (Searcy-Bernal 1994). A good example can be seen in a study that was designed to determine the protein requirement of brown trout *Salmo trutta* fry (Arzel 1995). In this study, dietary protein levels of 38, 43, 48, 53, 57, 60, and 65% were fed and weight gain (in grams) was used as the response. Despite the relatively small standard errors of the means (0.3–0.13 g) and three replicates per treatment, a multiple means test (Neuman-Keuls, $P < 0.05$) detected no differences between 53% and 65% protein. Re-examination of the dose–response relationship, using a second-order polynomial, indicated that the maximum response may occur at about 64% (Fig. 5).

A more recent example of misinterpretation of results, when an inappropriate test is used, can be seen in Lim *et al.* (1998). The study was designed to determine the effects of substituting canola meal for soybean meal in channel catfish diets. From 0 to 61.6% of the soybean meal was

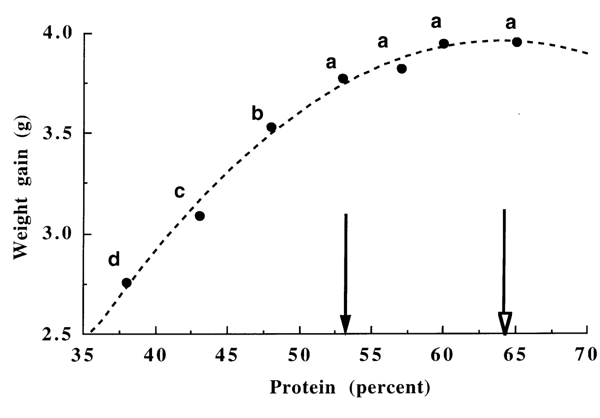

Figure 5 An example showing that even when three replicates are used and the standard errors of the means are small, test power is often not adequate to separate treatments using ANOVA and a multiple mean comparison (points with the same letters were reported as not significantly different; the estimated requirement is shown with the solid arrow). A curve fitted using a second-order polynomial is shown (dotted line) along with the re-evaluation estimate (open arrow). (Data from Arzel 1995).

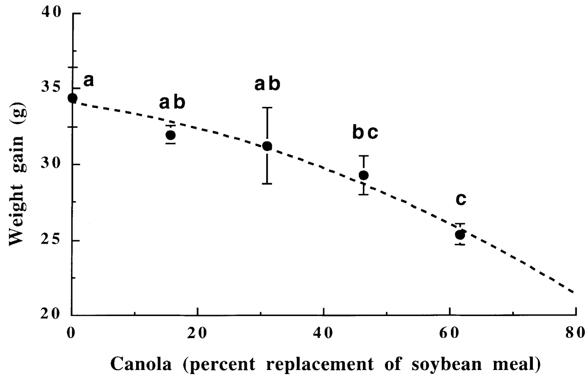

Figure 6 The effect on growth of replacing soybean meal with canola meal in channel catfish diets. The authors concluded, based on ANOVA, that up to 31% of the soybean meal could be replaced by canola with no reduction in growth (data points with different letters were reported as significantly different using ANOVA and a multiple-mean-comparison test with $P < 0.05$). It appears that a Type II error occurred. When the data are analysed with polynomial regression it appears that growth is reduced at any level of canola inclusion. (Data from Lim *et al.* 1998).

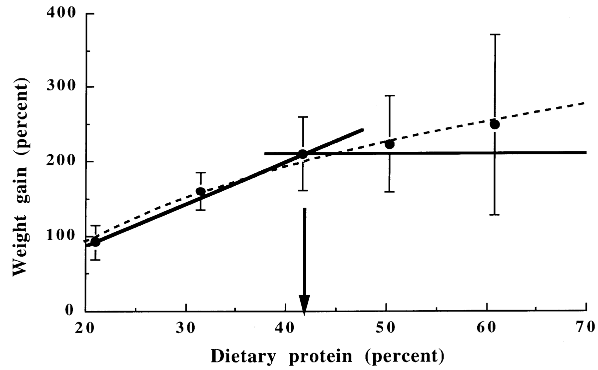

Figure 7 Estimated optimum percentage of protein in kuruma prawn diets using a broken line plot (published requirement estimate) or a second-order polynomial (re-evaluation, dotted line). (Data from Koshio *et al.* 1993).

replaced with canola and weight gain and other responses were compared. The authors concluded, based on ANOVA and multiple mean comparison, that up to 31% of the soybean meal could be replaced without reducing growth. However, if the data are examined using regression, an entirely different interpretation is achieved (Fig. 6). The large variance at the 31% canola inclusion level produces a Type II error. It appears that growth is reduced with any level of soybean meal replacement. The decision on the amount of canola meal to substitute could be based on economic grounds (i.e. is the reduction in weight gain compensated for by the use of a less-expensive ingredient), but it is an error to conclude that no reduced growth occurred at less than 31% canola inclusion. A Type II error like this often occurs when ANOVA is used in nutrition studies owing to low test power (Searcy-Bernal 1994). Additional examples where ANOVA understates the requirement can easily be found (Koshio *et al.* 1993; Dougall *et al.* 1996).

Broken line: The broken-line plot appears to be the most widely used method of evaluating dose–response data in nutrient requirement studies with fish. This technique involves using two straight lines to model the dose–response relationship. The ascending line represents increases in response with increasing nutrient intake, while the horizontal line (the one-slope method) represents nutrient sufficiency. Some authors fit a descending line after the break point (the two-slope method). The break point, using either method, corresponds to the nutrient requirement or minimum nutrient level that will produce the maximum response. The majority of authors using the broken-line method cited Zeitoun *et al.* (1976), Robbins *et al.* (1979) or Robbins (1986). It is interesting that these authors stress that, although some data sets can be modelled using the broken-line method, this method frequently underestimates the requirement so that curvilinear models are preferred. Baker (1986) and Fuller & Garthwaite (1993) also noted that the broken-line method understates the requirement. This was found to be the case with the majority of data sets using the broken-line method when they were re-evaluated. For example, in a study designed to determine the optimum protein level to feed kuruma prawns *Penaeus japonicus* (Koshio *et al.* 1993), prawns in two replicate tanks were fed diets containing from 21.0 to 60.7% protein and growth was chosen as the measure of response. The broken-line analysis indicated that 42% protein was adequate for maximum growth. When polynomial regression is used, the estimate appears to be greater than 61% (Fig. 7). Similar results were obtained when other data sets using broken-line analysis were re-evaluated (He & Lawrence 1993; Fox *et al.* 1995).

It appears to be more than coincidental that the average differences between the re-evaluated curve fits for ANOVA (195%) and the broken-line plots (200%) are similar. Zeitoun *et al.* (1976) showed how to use an iterative process to determine the breakpoint. This process appears to have been used in only a few cases. In the majority of data sets that used the broken-line method, the data were first subjected to ANOVA, and then the break point was placed at the lowest nutrient level above which no significant difference among

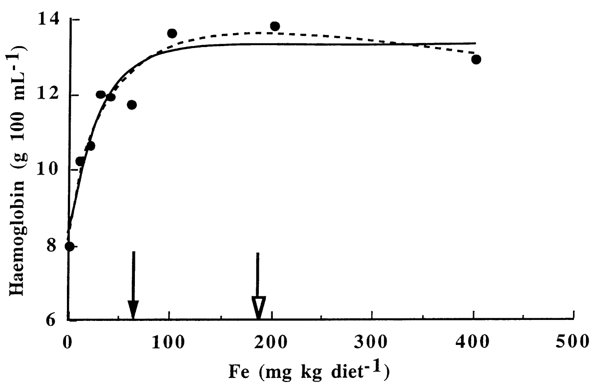

Figure 8 Comparison of estimated iron requirements in Atlantic salmon using the model: $y = u + v(1 - e^{wx})$ and Equation 1, the five-parameter saturation kinetics model (5-SKM) model of Mercer *et al.* (1989). The published estimate (solid line and arrow) and re-evaluated estimate (dotted line and open arrow) are shown. (Data from Andersen *et al.* 1996).

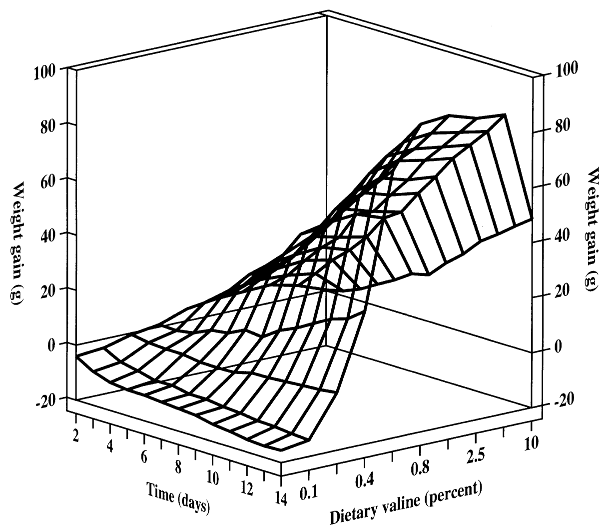

Figure 9 A dose–response surface showing weight vs. time of rats fed graded levels of valine. The data are from Mercer *et al.* (1993).

points could be found. Using this method, broken-line analysis and ANOVA produce equivalent estimates of the requirement.

Regression: In most cases where regression was used, the model chosen was appropriate for the data set. The models chosen appeared to have been selected because the equation produced the desired curve shape (an empirical vs. a mechanistic model). In only a few cases was there mention that an attempt had been made to obtain the model of best fit. In several cases, it was evident from a plot of the data that a better fit could be obtained using a different model. An example of the use of an inappropriate model can be seen in a study designed to determine the iron requirement of Atlantic salmon parr (Andersen *et al.* 1996). In this study, duplicate groups of fish were fed from 0 to 400 mg iron kg^{-1} diet, and blood haemoglobin and hepatic iron were the response variables measured. It is readily apparent, when the published plot is examined, that at the highest level of iron inclusion (400 mg kg^{-1}), blood haemoglobin levels were below those observed at 200 mg kg^{-1} (Fig. 8). The fact that the haemoglobin levels were above the fitted line at 100 and 200 mg iron kg^{-1} diet also suggests that the relationship is not correctly modelled. When this data set is examined using the 5-SKM (Mercer 1992) the estimate of the dietary iron level producing the maximum response is increased from the 60–100 mg kg^{-1} reported to approximately 200 mg kg^{-1}. Re-evaluation of the dietary iron vs. hepatic iron relationship also increased the estimated iron requirement from 100 to 200 mg kg^{-1}. A similar failure to test alternative models can be seen in another recently published study (Rodehutscord 1996)

Time: A component of nutrition studies that has received little attention is the duration of the experiment. Mercer *et al.* (1993) showed that requirement estimates should be considered a response surface, using dietary nutrient level, weight gain and time as the three dimensions rather than a two-dimensional relationship (Fig. 9). This requires that the response variable be measured at various times during the experiment, and that measurement and evaluation continue until a constant estimate of the requirement is obtained (Fig. 10). This type of analysis may prove difficult if the true nutrient requirement is dynamic; that is, it changes as the fish grows. If an experiment is of too short a duration, the requirement will be underestimated. This is probably most true when growth is used as the response variable, since a subclinical deficiency of some nutrients, such as phosphorus, will have no effect on growth for a relatively long period (Baeverfjord *et al.* 1998). In a study designed to estimate the magnesium (Mg) requirement of the rainbow trout, the estimated requirement increased each week when dose-response curves were constructed using weight gain as the response (Shearer 1989). When whole-body Mg was used as the response, it was possible to estimate the requirement after 2 weeks, and further estimates at weeks 4 and 6 produced similar values. Differences in the estimated requirement for a given nutrient in the same species may result from conducting experiments for differing periods of time (Shearer 1995; Åsgård & Shearer 1997).

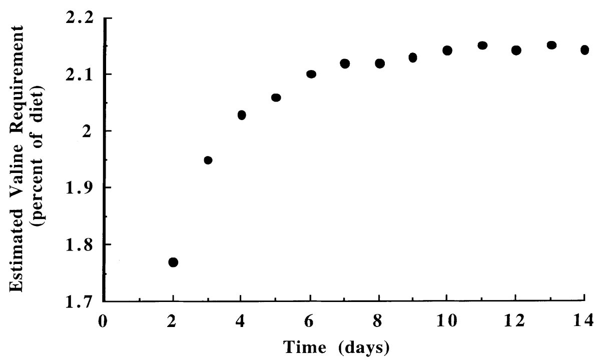

Figure 10 A time vs. estimated valine requirement plot for rats fed graded levels of valine (see Fig. 9). The data is from Mercer *et al.* (1993).

Protocols for avoiding some of the problems

The design, analysis and modelling steps that should take place in a nutrient requirement study are outlined in Fig. 11. This outline and the discussion that follows are highly simplified (for a more detailed outline see Daniel & Wood 1980 or Gilchrist 1984) as each topic is the subject of considerable statistical literature and each topic is actively undergoing research. A more detailed introduction to design, modelling and regression analysis of the types of experiments performed by fish nutritionists, along with a list of references on specific topics, can be found in Ryan (1997) and Freund & Wilson (1998). Additional references, on the major topics shown in Fig. 11, are included in the appropriate sections of the following discussion.

Design: The selection of dietary input levels is critical to estimation of the requirement. Unfortunately, without prior knowledge of the outcome of the experiment, optimum placement of the dose levels is impossible and application of a sequential strategy is recommended. Ryan (1997) discusses some of the procedures involved in trying to optimize the design of regression experiments. Mead (1988) presents a more in-depth discussion of this problem and indicates that

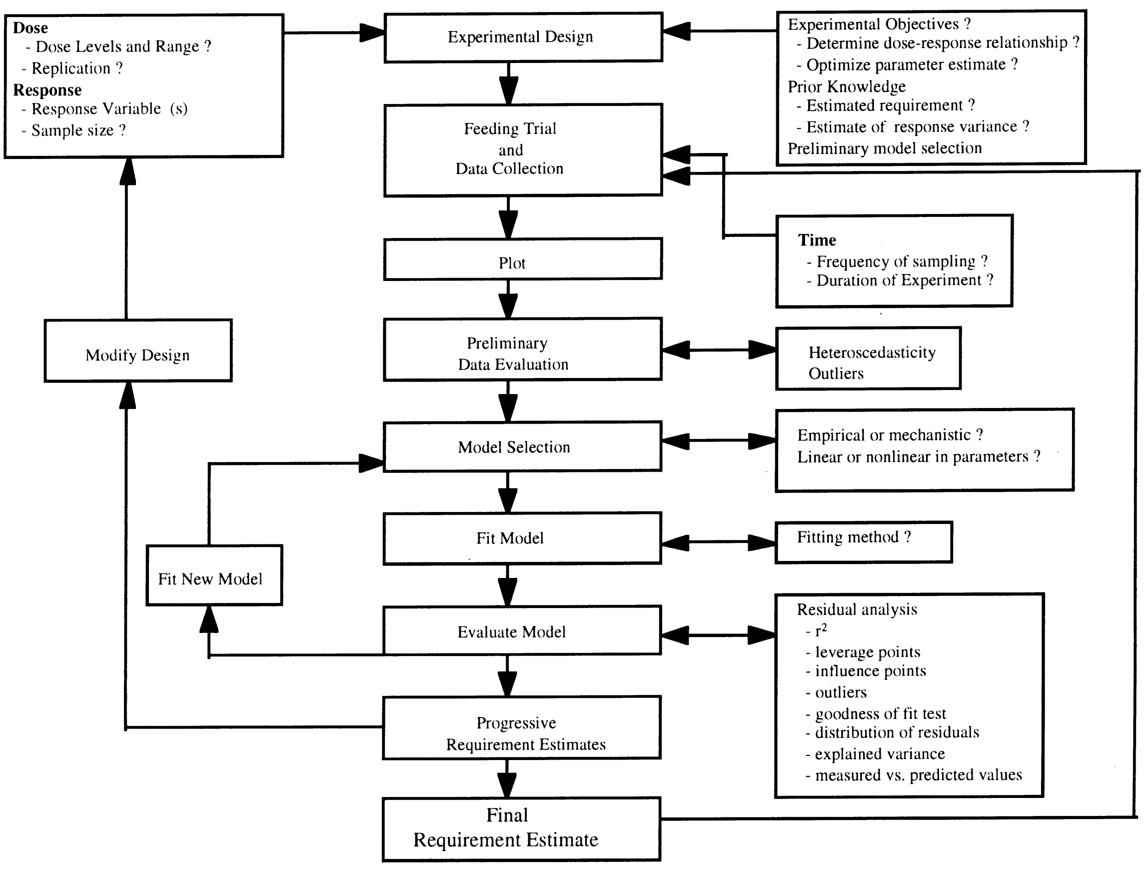

Figure 11 Flow chart of design, analysis and modelling steps in a nutrient requirement study. This shows a sequential strategy for determining an estimate.

the design of the study is dependent on the specific objectives of the experiment. If the primary objective is to establish the dose–response relationship, then the best strategy is to allocate nutrient input levels as follows: 1/3 near the lower end of the dose–response curve, 1/3 near the high point of the curve, and 1/3 at a high enough level so that the inhibition portion can be defined. If, however, the purpose of the experiment is to estimate the optimum nutrient level (the level corresponding to the maximum response), then the points should be allocated 1/4 in the ascending portion of the curve, 1/2 near the estimated requirement and 1/4 where the curve begins to decline. Positioning of these points is dependent on prior knowledge or a best guess about the shape of the dose–response curve. The number of nutrient levels should be equal to, or preferably (for robustness) one more than, the number of parameters in the model (Mead 1988). So, for the 5-SKM, there should be at least six levels, and for the 4-SKM five levels. Under ideal conditions, where resources are not limited, one would use many replicates at many levels. Increasing the number of replicates will improve the accuracy of the model and thus reduce the confidence intervals on the parameter estimates. It also makes it possible to check for heteroscedasticity, and allows a lack of fit test to be performed. However, based purely on statistical grounds, when regression is used, replication is not required; in fact there is little improvement in model confirmation or parameter estimation when the total number of observations (number of levels × replicates) is greater than 20 (Mead 1988). Methods of selecting dose levels and appropriate numbers of replicates are discussed in Seber (1977) and Mead (1988). Further discussion of experimental design, for the types of experiments performed by fish nutritionists, can be found in Ford *et al*. (1989), and Seber & Wild (1989).

Plot and initial data evaluation: Carroll & Ruppert (1988) indicate that routine data can be expected to contain between 1 and 10% gross errors, primarily due to errors in measurement or data recording. Judging from the data sets examined in this limited survey, outliers appear to be a common occurrence. Methods of dealing with outliers are explained by Rousseeuw & Leroy (1987), Chatterjee & Price (1991), Ryan (1997) and Freund & Wilson (1998). Ryan (1997) asserts that heteroscedasticity is common with nonlinear models and suggests that it must be moderately severe before a transformation is required. He also states that if the data cannot be adequately transformed, then robust regression or a nonparametric approach can be used.

Model selection: Freund & Wilson (1998) indicate that regression is the most common method of statistical analysis chosen when the true relationship between the variables is not known. They also observe that most true models are intrinsically nonlinear. Ryan (1997) notes that although the polynomial model does not correspond to a physical process, it can often be used to model physical data. In this survey, it appeared to model many of the studies that were re-evaluated reasonably well. Whether the models presented by Mercer *et al*. (1989) have broad applicability in modelling nutrient requirement studies will require further evaluation with designs that include nutrient levels where inhibitory responses are observed.

As can be seen in Fig. 1, the three models used for re-evaluation are similar in their ascending portions, but differ markedly as the nutrient level exceeds the maximum response. The polynomial and 5-SKM are similar up to, and just beyond, the maximum response. The 5-SKM appears to model growth responses to nutrient input level well, if a broad range of levels is fed, but it may not be applicable for all nutrients, or where a response other than growth is used to determine the requirement. The 4-SKM fits some data sets satisfactorily but, since it reaches a plateau with the predicted asymptote (maximum response) often well beyond the range of the nutrient levels fed, the researcher must arbitrarily select a nutrient level corresponding to 95 or 99% of the maximum response as the estimated requirement (see Fig. 8). In contrast, when a level of nutrient high enough to produce a growth reduction is fed, and the 5-SKM or quadratic model is used, a unique optimum response is obtained. Model choice and the advantages of mechanistic vs. empirical models are discussed in Box *et al*. (1978). Fry (1993) works through a number of examples of model fitting in a step-by-step manner. Ratkowsky (1990) provides examples of a number of models that can be used for nonlinear relationships, and Bates & Watts (1988) also provide guidelines for model selection. A group of appropriate models for use in dietary nutrient studies for fish should become available as different models are evaluated and published.

Model fitting: Most models are fitted with ordinary least squares (Ryan 1997). As indicated earlier, when model assumptions are violated, then robust methods or nonparametric methods can be used. Both topics are discussed by Freund & Wilson (1998). Ryan (1997) also discusses the use of weighted least-squares when the residuals are not normally distributed. Both Ryan (1997) and Freund & Wilson (1998) compare the strong and weak points of some of the computer software that can be used for statistical modelling and testing model adequacy.

Model evaluation: The widespread availability of personal computers and relatively inexpensive statistical software have vastly reduced the effort involved in data analysis. Unfortunately, their use can also lead to erroneous conclusions if the underlying assumptions of the statistical test are not met. In several of the papers examined it appears that only a single model was fitted, as there was no mention that the adequacy of the model was evaluated. Ryan (1997) stresses that care must be taken with the interpretation of r^2 when nonlinear regression is used. He also states that what constitutes a 'good' r^2 value is relative. Most tests of model adequacy involve examination of residual plots. Chatterjee & Price (1991) discuss the various residual plots and their use in determining if the model is adequate. More detailed discussions of residual plots can be found in Ryan (1997) and Freund & Wilson (1998). A lack of fit test can also be conducted if there is sufficient replication, but nutrient requirement studies generally contain small data sets with limited replication. The model should also be verified with an additional data set if possible. Finally, how does the researcher confirm that the correct model has been fitted to the data? Unfortunately, absolute confirmation is impossible, since all statistical (stochastic) models are approximations to the true or deterministic model. Ryan (1997) quotes George Box as stating 'All models are wrong, but some are useful.' All requirements that are determined empirically should therefore be considered as estimates.

Time: The nutrient requirement estimate obtained should be compared with estimates previously obtained during the experiment by plotting the estimates over time. The experiment should be continued until a constant value for the estimate is obtained.

Responsibility

Zolman (1993) estimated that over half of the papers published in scientific journals contain incorrect statistical methodology, and that in at least 25% of published papers this leads to the conclusions being compromised. The responsibility for ensuring that the published nutrient requirements are meaningful rests primarily with the scientist conducting the experiments. However, it is also the responsibility of the reviewers and editors to ensure that appropriate design and analysis procedures are followed. Useful guidelines for reviewing statistical methods in research papers can be found in Bausell (1994), Girden (1996) and Huck & Cormier (1996). Editorial guidelines for appropriate statistical methodology would help eliminate many problems. Some journals have already taken this step (JSFA 1988). Problems could also be reduced if one reviewer with statistical expertise is included when papers of this nature are evaluated for publication. It should be a requirement that the dose–response data and the model fitted be presented graphically; if they are not, the data should be plotted by the reviewer. A poor-fitting model is often readily apparent. Finally, authors of text and reference books should be encouraged to include a section on experimental design and statistical methodology. Despite the proliferation of books on fish nutrition in recent years, none has yet discussed how to conduct a nutrient requirement study or evaluate the inclusion level of an ingredient.

Summary

In summary, a number of published studies designed to estimate the dietary nutrient requirements in fish were re-evaluated using regression. The results of the re-evaluation suggest that many published papers seriously underestimate the nutrient requirement. The primary cause of the underestimation appears to be the use of inappropriate statistical methods to examine the dose–response relationship. Flaws in experimental design, failure to screen data and poor model choice due to a failure to test for adequate model fit are also responsible. An additional source of error may arise from conducting experiments for too short a period of time, but it was impossible to evaluate this potential in the present study. Before published requirements are accepted, the design of the study and the statistical methods used should be critically evaluated. It is the joint responsibility of the authors, reviewers and editors to ensure that valid statistical procedures are followed in papers submitted for publication.

Acknowledgements

This paper was inspired by an excellent paper on statistical power by Searcy-Bernal (1994) and spirited discussion with many colleagues over the years. The author thanks Dr Trond Storebakken and two anonymous referees for their comments and Ms Smith for her editorial assistance.

References

Andersen, F., Maage, A. & Julshamn, K. (1996) An estimation of the dietary iron requirement of Atlantic salmon, *Salmo salar* L., parr. *Aquacult. Nutr.*, **2,** 41–47.

Arzel, J., Metailler, R., Kereguer, C., Le Delliou, H. & Guillaume, J. (1995) The protein requirement of brown trout (*Salmo trutta*) fry. *Aquaculture*, **130,** 67–78.

Åsgård, T. & Shearer, K.D. (1997) The phosphorus requirement of juvenile Atlantic salmon *Salmo salar* L. *Aquacult. Nutr.*, **3,** 17–23.

Baeverfjord, G., Åsgård, T. & Shearer, K.D. (1998) Development and detection of phosphorus deficiency in Atlantic salmon, *Salmo salar* L., parr and post-smolts. *Aquacult. Nutr.*, **4,** 1–11.

Baker, D.H. (1986) Problems and pitfalls in animal experiments designed to establish dietary requirements for essential nutrients. *J. Nutr.*, **116,** 2339–2349.

Bates, D.M. & Watts, D.G. (1988) *Nonlinear Regression and its Application.* Wiley, New York, NY.

Bausell, R.B. (1994). *Conducting Meaningful Experiments.* Sage Publications, London, UK.

Box, G.E.P., Hunter, W.G. & Hunter, J.S. (1978) *Statistics for Experimenters — an Introduction to Design, Data Analysis and Model Building.* John Wiley & Sons, New York, NY.

Brecka, B.J., Kohler, C.C. & Wahl, D.A. (1995) Effects of dietary protein concentration on growth, survival, and body composition of muskellunge *Esox masquinongy* and tiger muskellunge *Esox masquinongy* × *E. luscius* fingerlings. *J. World Aquacult. Soc.*, **26,** 416–425.

Carmona-Osalde, C., Olivera-Novoa, M.A., Rodríguez-Serna, M. & Flores-Nava, A. (1996) Estimation of the protein requirement for bullfrog (*Rana catesbeiana*) tadpoles, and its effect on metamorphosis ratio. *Aquaculture*, **114,** 223–231.

Carroll, R.J. & Ruppert, D. (1988) *Transformation and Weighting in Regression.* Chapman & Hall, New York, NY.

Chatterjee, S. & Price, B. (1991) *Regression by Example.* John Wiley and Sons, New York, NY.

Cowey, C.B. (1992) Nutrition: estimating requirements of rainbow trout. *Aquaculture*, **100,** 177–189.

Daniel, C. & Wood, F.S. (1980). *Fitting Equations to Data.* John Wiley & Sons, New York, NY.

Dougall, D.S., Woods III, L.C., Douglass, L.W. & Soares, J.H. (1996) Dietary phosphorus requirement of juvenile striped bass *Morone saxatilis. J. World Aquacult. Soc.*, **27,** 82–91.

Dawkins, H.C. (1983) Multiple comparisons misused: Why so frequently in response-curve studies? *Biometrics*, **39,** 789–790.

Ford, I., Kitsos, C.P. & Titterington, D.M. (1989) Recent advances in nonlinear experimental design. *Technometrics*, **31,** 49–60.

Fox, J.M., Lawrence, A.L. & Li-Chan, E. (1995) Dietary requirement for lysine by juvenile *Penaeus vannamei* using intact and free amino acid sources. *Aquaculture*, **131,** 279–290.

Freund, R.J. & Wilson, W.J. (1998) *Regression Analysis: Statistical Modeling of a Response Variable.* Academic Press, San Diego, CA.

Fry, J.C. (1993) Bivariate regression. In: *Biological Data Analysis: a Practical Approach.* (Fry, J.C. ed.), pp. 81–125. Oxford University Press, Oxford, UK.

Fuller, M.F. & Garthwaite, P. (1993) The form of response of body protein accretion to dietary amino acid supply. *J. Nutr.*, **123,** 957–963.

Gilchrist, W. (1984) *Statistical Modeling.* John Wiley and Sons, Chichester, UK.

Girden, E.R. (1996) *Evaluating Research Articles from Start to Finish.* Sage Publications, London, UK.

Gurure, R.M., Moccia, R.D. & Atkinson, J.L. (1995) Optimal protein requirements of young Arctic charr (*Silvelinus alpinus*) fed practical diets. *Aquacult. Nutr.*, **1,** 227–234.

He, H. & Lawrence, A.L. (1993) Vitamin C requirements of the shrimp *Penaeus vannamei. Aquaculture*, **114,** 305–316.

Huck, S.W. & Cormier, W.H. (1996) *Reading Statistics and Research.* Harper Collins, New York, NY.

JSFA (1988) Recommendations of the Editorial Board on the use of statistics in papers submitted to JSFA — Guidelines to Authors (as formulated by A.W.A. Murray). *J. Sci. Food Agric.*, **42,** unnumbered.

Kim, K. (1993) Requirement for phenylalanine and replacement value of tyrosine for phenylalanine in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture*, **113,** 243–250.

Kirk, R.E. (1982) *Experimental Design: Procedures for the Behavioral Sciences.* Brooks/Cole Publishing Co, Belmont, CA.

Koshio, S., Teshima, S., Kanazawa, A. & Watase, T. (1993) The effect of dietary protein content on growth, digestion efficiency and nitrogen excretion of juvenile kuruma prawns, *Penaeus japonicus. Aquaculture*, **113,** 101–114.

Lim, C., Klesius, P.H. & Higgs, D.A. (1998) Substitution of canola meal for soybean meal in diets for channel catfish *Ictalurus punctatus. J. World Aquacult. Soc.*, **29,** 161–168.

Lochmann, R.T. & Phillips, H. (1994) Dietary protein requirement of juvenile golden shiners (*Notemigonus crysoleucas*) and goldfish (*Carassius auratus*) in aquaria. *Aquaculture,* **128,** 277–285.

Mead, R. (1988) *The Design of Experiments: Statistical Principles for Practical Application.* Cambridge University Press, Cambridge, UK.

Mercer, L.P. (1982) The quantitative nutrient–response relationship. *J. Nutr.*, **112,** 560–566.

Mercer, L.P. (1992) The determination of nutritional requirements: mathematical modeling of nutrient–response curves. *J. Nutr.*, **122,** 706–708.

Mercer, L.P., Flodin, N.W. & Morgan, P.H. (1978) New methods for comparing the biological efficiency of alternative nutrient sources. *J. Nutr.*, **108,** 1244–1249.

Mercer, L.P., Gustafson, J.M., Hogbee, P.T., Geno, C.E., Schweisthal, M.R. & Cole, T.B. (1984) Control of physiological response in the rat by dietary nutrient concentration. *J. Nutr.*, **114,** 144–152.

Mercer, L.P., May, H.E. & Dodds, S.J. (1989) The determination of nutritional requirements in rats: mathematical modeling of sigmoidal, inhibited nutrient response curves. *J. Nutr.*, **119,** 1465–1471.

Mercer, L.P., Yi, T. & Doods, S.J. (1993) Determination of nutritional requirements in rats: variation with time of weight gain responses to indispensable amino acids. *J. Nutr.*, **123,** 964–971.

Olivera-Novoa, M.A., Gasca-Leyva, E. & Martinez-Palacios, C.A. (1996) The dietary protein requirements of *Cichlasoma synspilum* Hubbs, 1935 (Pices: Cichlidae) fry. *Aquacult. Res.*, **27,** 167–173.

Ratkowsky, D.A. (1990) *Handbook of Nonlinear Regression Models.* Marcel Dekker, New York, NY.

Robbins, K.R. (1986) *A Method, SAS Program, and Examples for Fitting the Broken Line to Growth Data. University of Tennessee Res. Rep. 86–09.* University of Tennessee Agric. Exp. Sta., Knoxville, KY.

Robbins, K.R., Norton, H.W. & Baker, D.H. (1979) Estimation of nutrient requirements from growth data. *J. Nutr.*, **109,** 1710–1714.

Rodehutscord, M. (1996) Response of rainbow trout (*Oncorhynchus mykiss*) growing from 50 to 200 g to supplements of dibasic sodium phosphate in a semipurified diet. *J. Nutr.*, **126,** 324–331.

Rousseeuw, P.J. & Leroy, A.M. (1987) *Robust Regression and Outlier Detection.* John Wiley and Sons, New York, NY.

Ryan, T.P. (1997) *Modern Regression Methods.* John Wiley and Sons, New York, NY.

Searcy-Bernal, R. (1994) Statistical power and aquaculture research. *Aquaculture*, **127,** 371–388.

Seber, G.A.F. (1977) *Linear Regression Analysis.* Wiley, New York, NY.

Seber, G.A.F. & Wild, C.J. (1989) *Nonlinear Regression*. Wiley, New York, NY.

Shearer, K.D. (1989) Whole body magnesium concentration as an indicator of magnesium status in rainbow trout (*Salmo gairdneri*). *Aquaculture*, **77,** 201–210.

Shearer, K.D. (1995) The use of factorial modeling to determine the dietary requirements for essential elements in fishes. *Aquaculture*, **133,** 57–72.

Webster, C.D., Tiu, L.G. & Tidwell, J.H. (1997) Growth and body composition of juvenile hybrid bluegill *Lepomis cyanellus* × *L. macrochirus* fed practical diets containing various percentages of protein. *J. World Aquacult. Soc.*, **28,** 230–240.

Wilson, R.P. (1994) Amino acid requirements of finfish. In: *Amino Acids in Farm Animal Nutrition*. (D'Mello, J.P.F. ed.), pp. 377–399. CAB International, Wallingford, UK.

Zeitoun, I.H., Ullrey, D.E., Magee, W.T., Gill, J.L. & Bergen, W.G. (1976) Quantifying nutrient requirements in fish. *J. Fish. Res. Bd. Can.*, **33,** 167–172.

Zolman, J.F. (1993) *Biostatistics: Experimental Design and Statistical Inference*. Oxford University Press, Oxford, UK.