Two New Approaches for Examining Multivariate Aquaculture Growth Data: The "Extended Bayley Plot" and Path Analysis*

MARK PREIN
and
DANIEL PAULY
International Center for Living Aquatic Resources Management MCPO Box 2631, 0718 Makati Metro Manila, Philippines

Prein, M. and D. Pauly. 1993. Two new approaches for examining multivariate aquaculture growth data: the "extended Bayley plot" and path analysis, p. 32-49. *In* M. Prein, G. Hulata and D. Pauly (eds.) Multivariate methods in aquaculture research: case studies of tilapias in experimental and commercial systems. ICLARM Stud. Rev. 20, 221 p.

Abstract

Two new approaches for the multivariate analysis of fish growth in aquaculture are presented. The first of these two approaches, the "extended Bayley plot" is a multiple regression expansion of an existing bivariate method, which permits the inclusion of environmental and treatment variables when estimating the parameters L_∞ and K of the von Bertalanffy growth function, given precise measurements of fish length *and* weight at different ages. The derived regression model which must be based on a Type II, or "functional" regression, can be used to predict fish growth under anticipated conditions and thus identify appropriate farm management options. The differences of this method with the related "extended Gulland-and-Holt plot" are discussed. The second approach pertains to the application of "path" (or "causal") analysis in the context of aquaculture. Causal path diagrams, based on either extended "Bayley" or extended "Gulland-and-Holt" plots, can be used to put into a rigorous framework hypothesized networks of interacting variables controlling, for example, tilapia growth in ponds. Both methods were applied to a dataset based on pond growth experiments with Nile tilapia *Oreochromis niloticus*, conducted in Muñoz, Philippines.

Introduction

Two new approaches for the multivariate analysis of pond growth experiments are presented here:

(i) the multivariate extension of the Bayley plot, a method for estimating the parameters L_∞ and K of the von Bertalanffy growth function (VBGF) for length and weight growth data, and
(ii) the application of "path analysis" (also known as "causal analysis") to data from aquaculture experiments.

A rationale for the application of multivariate methods in aquaculture is given in Prein et al. (this vol.) and this topic need not be discussed here, where we shall limit ourselves to presenting new variants of existing techniques. It is our hope that these new variants will serve in highlighting those aspects of aquaculture datasets which traditional methods, and/or the methods discussed in the other contributions included in this volume, may fail to highlight.

We shall first discuss the theory behind the proposed new approach, then apply them to a dataset derived from growth experiments on Nile tilapia *Oreochromis niloticus*, conducted in Muñoz, Philippines from August 1979 to June 1981, and also used and documented by Prein (this vol.).

The Simple and Extended Bayley Plots

The Bivariate Model

The method to be discussed was proposed by Bayley (1977) as an approach for the estimation of the parameters L_∞ and K of the VBGF via a new linearizing transformation of this nonlinear function. The VBGF has for length the form

$$L_t = L_\infty(1-\exp-(K(t-t_0)) \qquad \ldots 1)$$

where

- L_t is the length at age t,
- L_∞ the mean length the fish would reach if they were to grow indefinitely;
- K the instantaneous rate at which L_∞ is approached; and
- t_0 fixes the origin on the time axis, and will be ignored in this contribution.

Given a length-weight relationship of the form

$$W = v \cdot L^m \qquad \ldots 2)$$

the VBGF for weight becomes

$$W_t = W_\infty(1-\exp-K(t-t_0))^m \qquad \ldots 3)$$

*ICLARM Contribution No. 909.

where

W_t is the predicted weight at age t,

W_∞ the weight corresponding to L_∞, and the other parameters are as defined above.

Bayley (1977) when presenting his new method pointed out that instantaneous growth rate in weight, G, is defined by the differential equation:

$$G = \frac{d(\ln W)}{dt} \quad \text{...4)}$$

which is approximated, for short time intervals, by the difference equation:

$$G = \frac{\ln W_2 - \ln W_1}{t_2 - t_1} \quad \text{...5)}$$

Given growth processes correctly described by the VBGF, Fig. 1 depicts the exponential decrease of the instantaneous weight growth rate with age, and the reciprocal of length *vs.* time. By incorporating a length-weight relationship into equation (4), the process of growth can be formulated in terms of weight on one side and length on the other. For short intervals this growth relationship takes the form:

$$\frac{\ln W_2 - \ln W_1}{t_2 - t_1} = \frac{\ln v + m\ln L_2 - \ln v - m\ln L_1}{t_2 - t_1} \quad \text{...6)}$$

or

$$\frac{\ln W_2 - \ln W_1}{t_2 - t_1} = \frac{m(\ln L_2 - \ln L_1)}{t_2 - t_1} \quad \text{...7)}$$

This, in terms of a difference equation, takes the form:

$$\frac{\Delta(\ln W)}{\Delta t} = \frac{m}{L} \cdot \frac{\Delta L}{\Delta t} \quad \text{...8)}$$

and in terms of a differential equation, the form:

$$\frac{d(\ln W)}{dt} = m \cdot \frac{d(\ln L)}{dt} \quad \text{...9)}$$

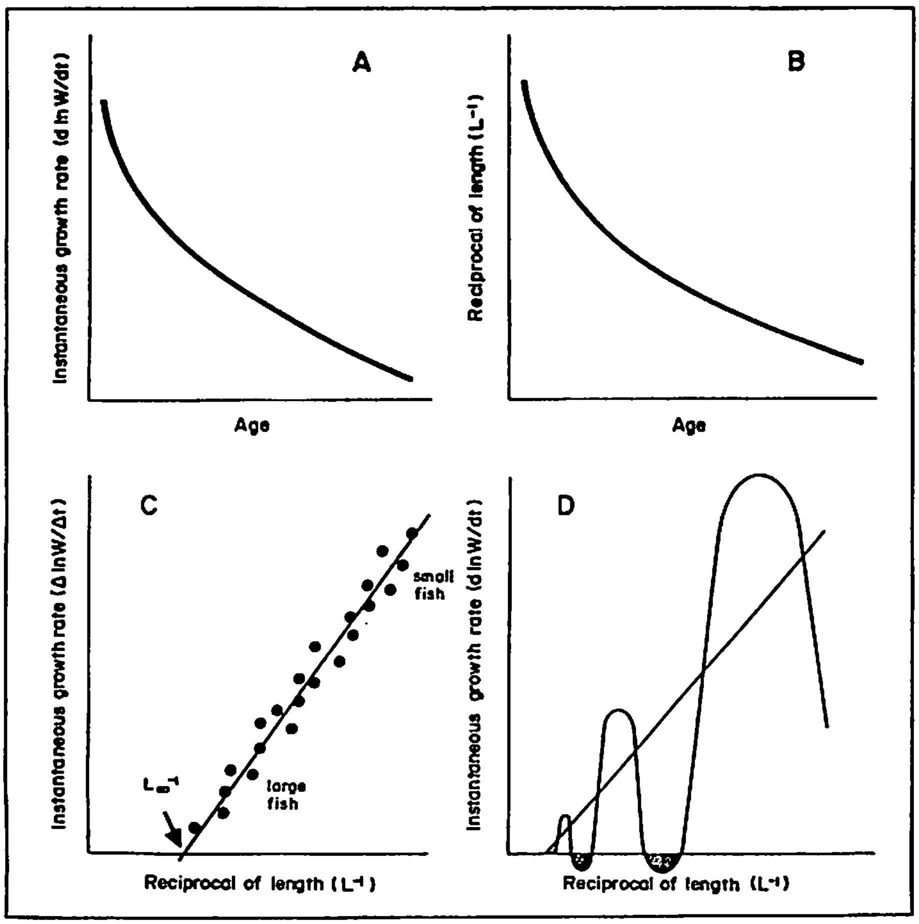

Fig. 1. Schematic representation of basic processes considered in new method (extended Bayley plot). A) The non-linear decrease of instantaneous growth rate in weight with age; B) The nonlinear decrease of the reciprocal of length with age; C) The Bayley plot. The relationship between instantaneous rate of growth in weight and the reciprocal of length can be described by a linear relationship, starting from the upper right end of the line for small fish, down towards the intercept with the abscissa for large fish; D) The Bayley plot for seasonally oscillating growth, which (in contrast to the Gulland-and-Holt plot, and assuming no shrinkage in length) permits negative values (i.e. loss of weight; hatched area) for the dependent variable. The residual variance around the regression line can be explained by including seasonally oscillating variables into a multiple regression.

Rearranged, this gives

$$\frac{d(\ln W)}{dt} = \frac{m}{L} \cdot \frac{dL}{dt} \qquad ...10)$$

Equation (8) is equal to the instantaneous growth rate G. By combining equations (4) and (8), instantaneous growth rate can be reexpressed in terms of the parameters L_∞ and K of the VBGF (equation 1), i.e.,

$$\frac{d(\ln W)}{dt} = mK \cdot \frac{L_\infty}{\overline{L}} - 1 \qquad ...11)$$

or:

$$\frac{\ln W_2 - \ln W_1}{t_2 - t_1} = -mK + mKL_\infty(1/L) \qquad ...12)$$

which has the form of a linear regression, where the expression on the left hand side of the equation is the dependent variable (y), L^{-1} is the independent variable, -mK is the intercept (a) and mKL_∞ is the slope (b).

Therefore

$$\Delta(\ln W)/\Delta t = a+b(L^{-1}) \qquad ...13)$$

Thus, the parameters of the VBGF can be estimated from successive measurements of length and weight and the parameter m of equation (2), from

$$K = -a/m \qquad ...14a)$$

and $$L_\infty = b/-a \qquad ...14b)$$

The relationship between instantaneous growth rate and the reciprocal of length is illustrated in Fig. 1C. Bayley (1977), who developed this method, gives an approach for estimating the variance of K. Estimating the variance of L_∞ can be done according to Snedecor and Cochran (1980).

The Multivariate Extension

The method discussed above relies on the relationship between growth rate in weight and the reciprocal of the average length during a given growth increment. Of these two, the variable showing the greatest amount of variance as a result of environmental effects will be growth rate in weight, which can, at times, have negative values (Fig. 1D). When plotting data from different experiments, with many different treatments, the variance around the regression line can be attributed, at least in large part, to environmental and treatment factors. To include these factors explicitly into one's analysis, equation (13) can be extended into a multiple linear regression equation of the form:

$$\frac{d\ln W}{dt} = a + b_1 (1/L) + b_2X_2 + ... + b_nX_n \qquad ...15)$$

where $a = b_1 \approx -m/k$ and whose parameters can be obtained through multiple regression analysis. The VBGF parameter L_∞ and K are obtained from

$$L_\infty = b_1/(-a + b_2X_2 + ... + b_nX_n) \qquad ...16)$$

and, for K, from -a/m.

The method embodied in equation (15) relates growth increments over short time periods to environmental or treatment effects measured during and averaged for these time intervals (Table 1). This is similar to the 'extended Gulland-and-Holt plot' (Pauly et al., this vol.), but requires that both the length and the weight of individual fish be recorded at the sampling events.

As for the "extended Gulland-and-Holt plot" the data requirements are therefore:

1. A cultured population of aquatic organisms must be sampled in length and weight at regular, short intervals. For shorter intervals, growth in length will be difficult to detect and sampling stress may result. For longer intervals information will be lost. The sample sizes should cover a representative portion of the population. The data for each individual organism should be recorded;
2. All environmental and treatment variables of interest should be measured at regular intervals with the appropriate frequency to obtain representative values for these intervals;
3. In the design of factorial experiments for analysis by these methods, a wide range of values of each variable should be covered:
   a) from small to large organisms, so that a representative growth model based on the VBGF can be fitted correctly;

Table 1. Extended Bayley method: data table organised according to experiment duration and individual measurements during intervals.

			BIOMETRICS				ENVIRONMENT			
	DATE		WEIGHTS		LENGTHS		VARIABLE-1		VARIABLE-2	
STOCKING	t1		W1		L1					
		Δt		$\bar{L}$		ΔL	xi : xi	$\bar{X}1$	xi : xi	$\bar{X}2$
1st SAMPLING	t2		W2		L2					
		Δt		$\bar{L}$		ΔL	xi : xi	$\bar{X}1$	xi : xi	$\bar{X}2$
2nd SAMPLING	t3		W3		L3					
⋮	⋮		⋮		⋮		⋮		⋮	
HARVEST	tn		Wn		Ln					

b) from low to high values of environmental and treatment variables, including an adequate number of zero-treatment (control) experiments, so that the regression can detect environmental and other effects on growth.

From the sampling intervals, mean values are calculated for all variables measured during the interval, together with the time interval in days, the instantaneous growth rate in weight and the reciprocal of the average length (Table 1). The data of all treatments and ponds are then organized in a data matrix ready to be used for multiple regression analysis. For the first pond and treatment the interval numbers are also the case numbers (Table 2). With this data matrix a multiple regression analysis can be performed.

Use of Type II Regression

Since the extension of the Bayley plot to a multivariate method was originally proposed (Prein 1990), the tendency of this method to overestimate L_∞ and underestimate K (see Prein 1990, section 3.82 and Fig. 4.6) has led us to reexamine the contention of Bayley (1977) that a Type I regression is appropriate for use in conjunction with his method, and by extension to equation (15).

Recall that fitting a Type I, *predictive* regression involves minimizing the squares of the vertical distance between the regression line and the observations. Thus, when plotting Y on X, it must be assumed that the X values are estimated (more or less) without error, (almost) all measurement errors being associated to Type I regression. Therefore, Type I (or arithmetic mean = AM) regressions, i.e., those taught in most statistics courses and built into most statistical computer packages, tend to have slopes whose values decline when the variance of the data points increase - a result of the way fitting is done.

Aquaculture growth data obtained as described above will tend to be "messy", with a large amount of unexplained variance remaining, whatever the method of fitting. Hence, the slope will tend to be strongly biased downward.

In a Bayley plot (see Fig. 1C) this will have the effect of underestimating the value of the intercept of the regression line with the X axis (i.e., L^{-1}) - and hence to overestimate L_∞ (see equation 14b).

One straightforward approach to reducing this bias is to use a Type II or functional regression. Such regression, also called geometric mean (GM) regression, may be seen as the geometric mean (hence the name) of two regressions, one with Y plotted against X, the other with X plotted against Y (each still minimizing the square of the vertical distance between line and residuals). The parameters (a', b') of a Type II can be obtained from those of a Type I regression (a,b) via

Table 2. Extended Bayley method: data matrix organised in final form appropriate for multiple regression analysis.

CASE	Y	X1	X2	...	Xn
1	$\Delta \ln W/\Delta t$	$1/\bar{L}$	$\overline{VAR1}$	...	$\overline{VARn}$
2	•	•	•	...	•
3	•	•	•	...	•
.	•	•	•	...	•
.	•	•	•	...	•
.	•	•	•	...	•
n	•	•	•	...	•

mean values of environmental variables in fish growth intervals.

$$b' = b/|r| \qquad \text{...17a)}$$

and $$a' = Y\text{-}b'X \qquad \text{...17b)}$$

where $|r|$ is the absolute value of the correlation coefficient linking Y and X.

Thus, a Bayley plot fitted with a Type II regression will always produce estimates of L_∞ lower than a Bayley plot fitted with a Type I regression, and the difference between the two estimates of L_∞ will increase as $|r|$ decreases. Applying these considerations to equation (15), i.e., to the multivariate extension of the Bayley plot is not straightforward, however, because an equation analogous to (17a) is not available.

The job can be done, however, by estimating the parameters of a number of multiple regression models, then computing their geometric mean.

This is best explained using an example involving three variables: Y the *real* dependent variable, and two independent variables, X and Z. In this case:

i) estimate the slopes and intercepts of three equations (i, j, k) making each of the variables act in turn as the "dependent" variable:

$$Y = a_i+b_{1i}X+b_{2i}Z \qquad \text{...(i)}$$
$$X = a_j+b_{1j}Y+b_{2j}Z \qquad \text{...(j)}$$
$$Z = a_k+b_{1k}X+b_{2k}Y \qquad \text{...(k)}$$

ii) solve equations (j) and k) for Y, i.e., for the *real* dependent variable:

$$Y = -(a_j/b_{ij})+(1/b_{1j})X-(b_{2j}/b_{1j})Z$$
$$Y = -(a_k/b_{2k})-(b_{1k}/b_{2k})X+(1/b_{2k})Z$$

iii) estimate geometric mean values of b_1 and b_2 (i.e., b'_1 and b'_2) via

$$b'_1 = \sqrt[3]{b_{1j} \cdot \left(\frac{1}{b_{1j}}\right) \cdot \left(\frac{b_{1K}}{b_{2K}}\right)}$$

and $$b'_2 = \sqrt[3]{b_{2i} \cdot \left(\frac{b_{2j}}{b_{1j}}\right) \cdot (1b_{2K})}$$

iv) estimate the corresponding intercept (a'), in analogy to equation (17b) from

$$a' = Y\text{-}(b_1X+b'_2Z) \qquad \text{...18)}$$

[The extension of this approach to more variables, although tedious, is quite straightforward, but is not shown here; see Pauly (1986) for an example involving five variables].

Using the a' and b'_1 values in equation (15) instead of a and b_1 values will lead to less biased estimates of L_∞ and K, as will be shown below.

Methods to estimate the variance of Type II multiple regression parameter estimates are not known to us; indeed no such methods appear to exist even for the bivariate case (Ricker 1975).

Path Analysis

History and Theory

The method of path analysis, also called causal analysis, was developed by the geneticist Sewall Wright (1921, 1923, 1934) for the analysis and interpretation of effects of heredity (Land 1969; Li 1975). Later applications were made in genetics

(Cloninger 1980), econometrics (Backhaus et al. 1989), political sciences (Sanders 1980), social sciences (Weede 1970; Boyle 1970; Kang and Seneta 1980; Blalock 1985a, 1985b), psychology (Brandstädter 1976; Brandstädter and Bernitzke 1976), agriculture (Dörfel and Neumann 1973; Rasch 1983), marine biology (Schwinghamer 1983) and fisheries biology (Davidson et al. 1943; Coelho and Rosenberg 1984; Robinson and Doyle 1988). Eknath and Doyle (1985) used the LISREL VI approach to causal analysis (Jöreskog and Sörbom 1984) to estimate unobserved variables from scale data of Indian carp. Here only a short overview of the method can be presented, partly adapted from Prein (1985). For an extensive description of technical procedures see Turner and Stevens (1959), Dörfel (1972a, 1972b), Kim and Kohout (1975), Li (1975), Heise (1969, 1975), Draper and Smith (1981), Backhaus et al. (1989), Jöreskog and Sörbom (1984).

General Concept

The general approach to the application of path analysis is:

1. The researcher has to formulate an *a priori* causal hypothesis, which requires that the examined system is adequately understood. Also the researcher must have a hypothesis of the interactions of the variables in the system based on knowledge and reasoning. Mostly, several different hypotheses are formulated and tested in an interactive process over several runs.
2. With path analysis one can examine, but not test, a causal hypothesis.
3. Analysis is done by:
   a) calculating a multiple regression equation, and then
   b) graphically and visually analyzing a path diagram.

Requirements

As in multiple regression (on which path analysis is based), the relationships among the variables must be linear. Thus, nonlinear processes must be linearized. In the present case, the Bayley plot served for linearization of the growth process. Similarly, the Gulland-and-Holt plot may serve as a basis for growth curve linearization and use with path analysis, as demonstrated in Prein (this vol.). In path analysis the variables must be used in a standardized form.

Standardization of Variables

Standardization of all variables in the analysis is done by subtracting the mean of each variable from each individual value and dividing by its standard deviation (Li 1975; Heise 1975; Backhaus et al. 1989):

$$SV_x = \frac{X - \overline{X}}{S.D._x} \qquad ...19)$$

Through this procedure the mean of each standardized variable becomes zero and its standard deviation becomes equal to unity. Therefore the effects of different factors can be compared directly between all independent variables in terms of their relative strength. With these variables multiple regression equations are calculated.

The regression coefficients of standardized variables are called beta coefficients (Blalock 1972). The beta coefficients (also termed 'beta weights') can also be determined directly from the regression coefficients (Norusis 1985) using:

$$beta_x = b_x \frac{S.D._x}{S.D._y} \qquad ...20)$$

where $beta_x$ is the beta coefficient of the independent variable x, b_x is the regression coefficient of the independent variable x, $S.D._x$ is the standard deviation of the independent variable x, and $S.D._y$ is the standard deviation of the dependent variable.

The independent variables are termed "cause" variables, the dependent variable is termed "effect" variable:

"cause" variables —> "effect" variable
$(x_1, x_2, x_3...x_n)$ (y)

Path Diagrams

The basis of path analysis is the design of a path diagram and the insertion of the beta coefficients (now termed path coefficients) at the respective paths (arrows) in the diagram. From a two-variable example:

$$y = a + b_1X_1 + b_2X_2 \qquad ...21)$$

the resulting path diagram is shown in Fig. 2.

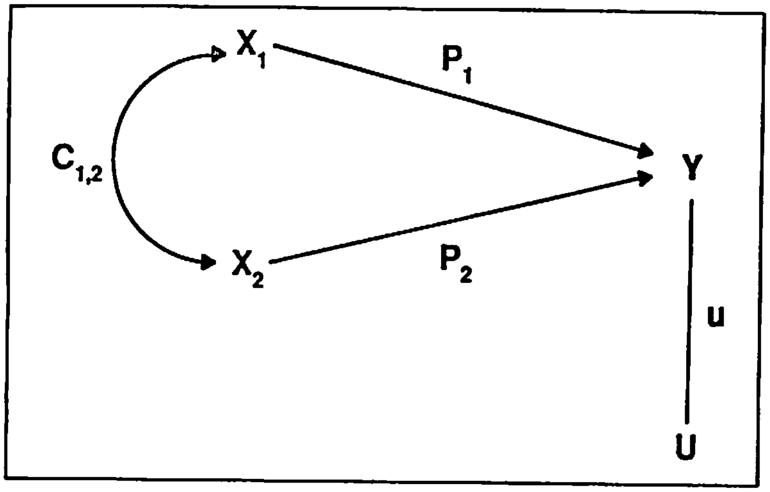

Fig. 2. Theoretical path diagram for a two-variable example. P_1 and P_2 = path coefficients (Beta coefficients of X_1 and X_2); $C_{1,2}$ = the correlation between X_1 and X_2; U = residual effect; u = amount of unexplained variance which is $1\text{-}R^2$.

Examination of the path diagram reveals the direction and strength of influences among variables. By following the paths in several steps over different variables, combined effects can be described. These combined effects are called 'compound paths'. The coefficients have no meaning in an absolute sense. Their relative comparison though allows for identification of the strength of direct and indirect influences. These causal relations and interrelations may be localized and described through path analysis. There are several rules for the interpretation of path diagrams which were summarized by Coelho and Rosenberg (1984) as follows:

"1) Cause-and-effect relationships are unidirectional and are shown by arrows with heads pointing at the dependent variable;
2) All hypothesized factors (predictors) which contribute to the variation of the dependent variable(s) are included in the diagram;
3) Direct paths are the direct connection between two variables;
4) Compound paths with component paths are the result of several individual paths;
5) The overall coefficient for a compound path is the product of the coefficients of its component paths;
6) The correlation between two variables is the sum of all paths by which they are connected. Correlations which imply no causality are shown with double headed arrows;
7) The residual coefficient ($1 - R^2$), which is a composite of unknown sources of variation, is indicated by a simple line;
8) The amount of variance explained by the model for any dependent variable is the sum of all complete circuits among the independent variables which affect the dependent variable. Alternatively, this value can be defined as one minus the square of the residual coefficient."

Since path analysis is based on multiple regression, the application of this method to the "extended Gulland-and-Holt" method and "extended Bayley" method can be expected to generate useful insights. Here, path analysis is demonstrated in combination with the extended Bayley plot. The methods and analyses presented herein were partly included in Prein (1985, 1990). The results will be compared with those of analysis based on the extended Gulland-and-Holt plot (Prein, this vol.).

Applications of the New Approaches

The Data Used

Data from the ICLARM/CLSU experiments (Hopkins and Cruz 1982) contained some recordings of individual lengths and weights of Nile tilapia during the sampling events (Prein, this vol.). Together with the length/weight relationship derived there, these were applied to test the new method proposed above using the data in the file PHILSAMP.WK1 (see Appendix II). The results of the analysis with the extended Bayley method are compared with those produced with the extended Gulland-and-Holt method (Prein, this vol.).

The usefulness of predictive multiple regression models for production planning and farm management has been pointed out (Prein, this vol.), which goes beyond the purpose of analytical identification and quantification of governing effects.

Testing of the Model

The analysis performed here was based on the same randomly sampled part of the dataset which was used in the extended Gulland-and-Holt method. To conform with the procedures of statistical model building, the derived equation was then tested on the unused, remaining part of the dataset (Prein, this vol.). The obtained set of regression coefficients should not differ significantly from that initially developed.

Extended Bayley Analysis

An ordinary Bayley plot of the entire ICLARM-CLSU dataset is given in Fig. 3. With the sample dataset of 200 cases, the following equation is obtained:

$$\Delta \ln W \Delta t = -0.03947 + 0.8678\ (ML^{-1}) \qquad ...22)$$

with n = 184, $r^2 = 0.628$, SEE = 0.0178, $P < 0.001$, $K = 0.01212^{-day}$ and $L_\infty = 22.0$ cm

An extended Bayley plot employing the same set of variables as in the extended Gulland-and-Holt plot produces the following regression equation:

		mean	range
$\Delta\ln W \Delta t$ =			
0.79907	$(\text{mean length})^{-1}$	0.07	0.19-1.0
$+1.151 \cdot 10^{-4}$	manure input $kg \cdot ha^{-1} \cdot day^{-1}$	73	0-221
-0.03338	SQRT stocking density $kg \cdot m^{-3}$	0.36	0.08-0.8
$+1.797 \cdot 10^{-5}$	pond area m^2	811	400-1000
$+5.873 \cdot 10^{-5}$	solar radiation $ly \cdot day^{-1}$	364	133-633
-0.06801			
			...23)

with n = 184, $R^2 = 0.675$, SEE = 0.0168, $P < 0.001$, $K = 0.0209^{-day}$ and $L_\infty = 35.0$ (22.46 to 88.01) cm, where SQRT is the square root.

The percentage of total explained variation represented by each of the independent variables, together with their 95% confidence limits is:

		lower	upper
$(\text{mean length})^{-1}$	= 35.1 %	0.63851	0.95962
manure input	= 3.4 %	$2.503 \cdot 10^{-5}$	$2.051 \cdot 10^{-4}$
stocking density	= 2.8 %	-0.06233	$-4.428 \cdot 10^{-3}$
pond area	= 6.5 %	$7.907 \cdot 10^{-6}$	$2.803 \cdot 10^{-5}$
solar radiation	= 7.1 %	$2.735 \cdot 10^{-5}$	$9.011 \cdot 10^{-5}$
CONSTANT		-0.09468	-0.04134

TEST OF THE MODEL

As in the test of the extended Gulland-and-Holt equation described in Prein (this vol.), the remaining part of the divided dataset was used to compute the regression equation for the extended Bayley plot. The coefficients of the derived equation were not significantly different (at the 95% level) from the coefficients determined with the sample dataset. Further, the signs of the independent variables were the same.

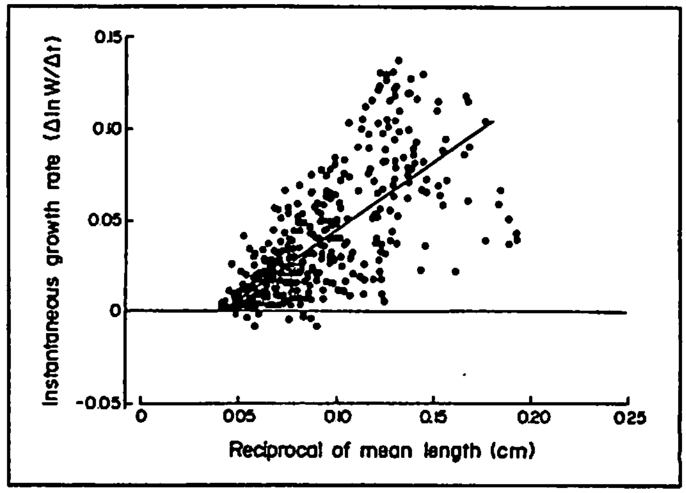

Fig. 3. Bayley plot of Nile tilapia grown in ICLARM-CLSU experiments at Muñoz, Philippines, in 1978-1981, n = 616. See text for regression equation. Note heteroskedasticity as discussed in text.

COMPARISON OF BOTH METHODS

The extended Gulland-and-Holt method and the extended Bayley method performed similarly in identifying variables which are influential on fish growth. The extended Bayley method resulted in a higher R^2 value (i.e., a higher amount of explained variance) than the extended Gulland-and-Holt method.

With the dataset used here, both methods showed the same sensitivity and identified the same set of variables as significant predictors of Nile tilapia growth rate. The sign of the variables was the same, except for the reciprocal of mean length, which was positive compared to untransformed mean length. Regarding the contribution of the auxiliary variables to the total amount of explained variance, their relative strength is very similar in both methods.

In the tests on the part of the dataset not used for model derivation, both methods showed the same stability and precision. The estimated models from both parts of the dataset were not significantly different.

The coefficient of determination was considerably higher for the equation determined with the extended Bayley method ($R^2 = 0.66$) than for the extended Gulland-and-Holt method ($R^2 = 0.40$, Prein, this vol.), based on the same dataset and the same variables. This is a consequence of the close relationship between reciprocal length and weight growth rate. The relative contribution of the auxiliary variables remained similar.

In the analyses performed here, some variables could not be included in the regression models. In some cases, the independent variables were insignificant, i.e. there was no correlation between them and the dependent variables. This can result if the variable in question does not vary in the dataset (i.e., due to experimental design), or the variance is not related to growth rate.

In other cases, variables could not be entered into the regression due to multicollinearity with other variables. Through disturbing effects, a highly significant variable already in the equation may become insignificant when a further, collinear variable is entered. This leads to the main problem, and disadvantage, of the methods applied here. Mean length must be entered as an independent variable for the methods to work, since these are based on the ordinary Gulland-and-Holt and Bayley plots. Therefore, any variables highly correlated with mean length in the datasets cannot be included. Variables which are normally considered to be important predictors for fish growth cannot be included if, due to experimental design, these were not varied according to factorial design principles. Particularly in the dataset from Dor station (Prein, this vol.), treatment variables such as stocking density, manure and pellet input, but also solar radiation and water temperature were highly collinear with mean length. In such cases, the datasets cannot be extensively analyzed with these two methods, limiting the amount of information that can be extracted from them. These restrictions are due to the rules of multiple regression. Correlation tables and values of the 'tolerance' statistic must be checked for compliance with acceptable limits. Multicollinear variables do not improve R^2, but rather inflate the standard errors of the regression coefficients (Norusis 1985). Besides this, the 'parsimony-principle' of regression model building should generally be followed, i.e., fewer variables in a regression model are better, making it more robust (Weisberg 1980).

Derived Growth Parameters

In the determination of the VBGF growth parameters K and L_∞, different values were obtained with both methods. The ordinary Gulland-and-Holt plot results in an estimate of K = 0.00994day^{-1} with lower and upper 95% confidence limits of 0.00773day^{-1} and 0.01215day^{-1}, respectively. The value for K obtained by the ordinary Bayley method is 0.01215day^{-1}, which is within the limits. It is therefore not significantly different.

The value for L_∞ derived with the ordinary Gulland-and-Holt plot was 25.4 cm with lower and upper confidence limits of 22.3 and 28.6 cm, respectively. According to Sparre et al. (1989), the confidence limits for L_∞ are only approximations. The value of L_∞ obtained with the ordinary Bayley plot is 22 cm, which is beyond the lower limit.

With the extended Gulland-and-Holt method, a value of K = 0.00652day^{-1} was estimated, with lower and upper confidence limits of 0.00162day^{-1} and 0.01141day^{-1}, respectively. The value of K computed with the extended Bayley method is 0.0209day^{-1}, which is beyond the upper limit, and is therefore significantly different.

The values for L_∞ derived with the extended Gulland-and-Holt method are 30.8, 23.2 and 38.3 cm, based on the average, minimum and maximum values of the independent variables, respectively. The lower and upper 95% confidence limits are 11.3 and 33.5 cm, obtained by inserting the average values into the lower and upper confidence limits of the regression coefficients. The average value for L_∞ calculated with the extended Bayley plot is 35.0 cm, which is beyond the upper limit.

It should be noted that the dataset for the extended Bayley method contains an entirely different variable, which is also the dependent variable (weight) and was measured separately on the fish. Differences in estimation of the equations and the VBGF growth parameters may be due to this fact. A more adequate test for the precision of the two methods is performed when the values of fish weight for the Bayley method (ordinary and extended) are computed with a length-weight relationship. In this case, all differences in the obtained equations and VBGF growth parameters are attributable to the methods. On the other hand, if the parameters were not significantly different, this would prove that the differences between VBGF parameters found above are due to the fish weights actually measured.

In the sample dataset which was used for the derivation of the equations described above, Nile tilapia weights were computed from the measured lengths with the length-weight relationship. An ordinary Bayley plot resulted in the following equation:

$$\Delta \ln W \Delta t = -0.03653 + 0.88455\ (ML^{-1}) \quad ...24)$$

with n = 193, $r^2 = 0.601$, SEE = 0.0170, $P < 0.001$, $K = 0.0112 day^{-1}$ and $L_\infty = 24.2$ cm

The obtained VBGF parameters are not significantly different from those estimated with the ordinary Gulland-and-Holt plot. A regression analysis run with the extended Bayley method with the same set of variables as used above produces:

		mean,	range
$\Delta \ln W \Delta t$ =			
0.80246	(mean length)$^{-1}$	0.07	0.19-1.0
$+1.254 \cdot 10^{-4}$	manure input $kg \cdot ha^{-1} \cdot d^{-1}$	73	0-221
-0.03326	SQRT stocking density $kg \cdot m^{-3}$	0.36	0.08-0.8
$+1.718 \cdot 10^{-5}$	pond area m^2	811	400-1000
$+5.384 \cdot 10^{-5}$	solar radiation $ly \cdot day^{-1}$	364	133-633
-0.06199			

...25)

with n = 193, $R^2 = 0.663$, SEE = 0.0158, $P < 0.001$, $K = 0.0190 day^{-1}$ and $L_\infty = 25.7$ (15.9 to 83.3).

The estimate of K is beyond the upper limit and is therefore significantly different. In contrast, the value for L_∞ is not significantly different. Thus, in the present case, the extended Bayley method produced slightly higher estimates of K and similar estimates of L_∞ compared to the extended Gulland-and-Holt plot.

Extended Gulland-and-Holt Method

In a derived equation based on the extended Gulland-and-Holt method (Pauly et al., this vol.), all environmental effects are incorporated in the VBGF parameters K and L_∞, although to a different extent (Prein, this vol.). While a single value of K is computed for the entire dataset, a separate value of L_∞ results for each individual case. Therefore, the value of K is influenced by the average environmental and treatment conditions, while L_∞ is highly flexible and responds to changes in the environmental variables (if these are included as a variable in the equation).

The method is based entirely on length measurements. If only weights are available, these have to be transformed with a length-weight relationship. This procedure though, may lead to negative values for growth rate, since weight loss may occur. A loss in length can be nearly excluded for fish under aquaculture conditions. Therefore, negative values in a dataset with length measurements can be usually identified as measurement errors. On the other hand, since fish size is the "instrument" for detecting environmental and treatment effects, the time interval between samplings must be long enough for fish size to respond in form of growth in length. Minor effects may not be measurable or may be hidden within the error range of measurement and will therefore not be detected. Weight of fish is much more responsive to environmental and treatment influences and can be regarded as much more sensitive than length, particularly on a short time scale.

EXTENDED BAYLEY METHOD

For the estimation of K and L_∞ with the extended Bayley method, the same as said above applies here too, with the exception that both also contain information on the influence of environmental and treatment effects on the relationship between length and weight. In this method, weight increments are used as the 'instrument' to detect environmental influences on growth. Length data are also necessary for the method to work, since the reciprocal of mean length per growth interval is the first (and obligatory) predictor variable. Both methods used here are applicable to size increment data collected at unequal time intervals.

The wide ranges for the derived VBGF parameters, based on the extended Bayley method, are due to the bulk weighings in the ICLARM-CLSU dataset. Individual fish weighings should give more precise values, which represent better the true relationship between length and weight. Based on such data, the extended Bayley method produces more reasonable VBGF parameters, as shown by a test vased on a subset of the ICLARM-CLSU data.

Svärdson (1984) showed that the ordinary Gulland-and-Holt plot was sensitive to growth variation in the smaller fish sizes (i.e., in the ascending limb of length growth curves). Measurement errors in small fish lead to more biased estimates of K and L_∞. Both methods presented here are based on a linearization of a nonlinear function. The necessary transformations have consequences for parameter estimation (Svärdson 1984). The Bayley method is based on a 'strong' transformation, leading to a higher value for r^2 than the Gulland-and-Holt plot, when applied to the same dataset. The average estimates of K and L_∞ are similar in both methods. However, based on the multiple regression version, the Bayley plot produces a much wider range of L_∞ values. Particularly, low growth rates lead to a flattening of the slope, which produces unrealistically high values of L_∞, a result of the reciprocal transformation. Thus the extended Bayley method may be more sensitive, but consequently less robust, than the extended Gulland-and-Holt method.

Unexplained Variance

The amount of total variance in growth rate in the ICLARM-CLSU dataset which was explained by the developed regression models was 67% in the extended Bayley analysis. Although highly significant, this indicates that part of the variance in tilapia growth rate was due to effects which were not included in the equations. Several reasons may be responsible for this fact.

In some cases, variables with strong influence could not be incorporated into the model due to high correlations with other variables. Their possible contribution to variance explanation is lost (as discussed above).

A second possible source of unexplained variance may be caused by cases in which important key variables are not measured. For example, in the presently analyzed datasets, the amount of natural food available to the fish (e.g., water samples measured as plankton content, chlorophyll *a*, or protein content) was not measured directly in the ponds in any case. Also, a longer duration of the low-oxygen periods in the early morning could possibly outbalance the positive effect of higher food availability and reduce growth.

A third possibility may lie in the imprecision and inaccuracy of average values per interval. Some of the measured parameters are highly variable, such as D.O. or pH. In the present approach, their average effect on fish growth during a growth interval is expressed in form of a single, mean value, based on the available individual measurements in the interval. If these measurements are not taken frequently and representatively, the averages will not adequately reflect the true effects. For example, AMDO varies daily, depending on meteorological conditions; two single measurements during a two-week interval are thus inadequate. The pH of pond water is often measured only in the morning, while higher afternoon values can lead to toxic conditions for the fish through ammonia conversion from the ionized to the molecular form (Steinmann and Surbeck 1922; Schaeperclaus 1952; Wuhrmann and Woker 1953; Ball 1967; Sousa et al. 1974; Redner and Stickney 1979; Alabaster and Lloyd 1980; Chetty et al. 1980; Spehar et al. 1982). Today, such problems can be overcome with the application of continuous measurement of parameters with long-standing electrodes and dataloggers (Piedrahita et al. 1987).

A further, considerable source of variance may be due to errors in determining fish sizes at the sampling events. In the methods applied here, growth rate is used as the 'instrument' to detect environmental and treatment effects and is used as dependent variable in the regression procedure. Any errors in the determination of average fish size during sampling procedures will introduce variance into the growth rate variable. This amount of variance cannot be accounted for by any independent variable, which reduces the value of all the effort invested into their measurement. Therefore, it should be of highest priority to strive for the highest possible precision at the sampling events when designing and performing sampling procedures in fishponds.

The common method applied to fishpond studies is to seine a sample of fish from the pond with a small-meshed net at regular intervals (two to four weeks). The sample is either bulk-weighed and counted, or each fish is measured and weighed individually. Often, sampling of a nonrepresentative portion of the pond population may lead to erroneous estimates. A sample size of 10% of the pond population is common, but Lovshin (1984) showed that even larger sample sizes of 20% can have a 20% error. Even large sample sizes may be biased since tilapia are known to cause sampling errors during netting operations through their evasive behavior, which they learn quickly (Kelly 1957). A further source of sampling error is nonrepresentative capture performance by the net, caused by jumping and hiding of the fish and incorrect net handling by personnel (Barthelmes 1960; Yashouv 1969). In the case of tilapia reproduction in ponds, young fish may cause the size distribution to 'smear' if experiment durations are long, since young fish can grow quickly and catch up with the smallest sizes of the adult stock.

Random variance in fish growth is a factor leading to the natural size distributions in fish populations. Reactions of fish populations in different ponds may not be the same, when confronted with the same treatment. Variation in the growth rates may be due to such effects. In most experiments, 'zero' treatment controls are often not performed. Therefore it is not possible to assess to which extent observed variations are due to treatments or due to natural variability. It is not well understood as to what extent the size distribution of a fish population grown under culture conditions is influenced by intrinsic (e.g. genetic or behavioral), or extrinsic factors (e.g. stocking density, age, sex, food availability, environmental conditions), or by interactions between them (Buschkiel 1937; Wohlfarth and Moav 1969; Wohlfarth 1977; Brett 1979; Nakanishi and Onozato 1987; Hepher et al. 1989). Only few studies exist in which these effects have been investigated for fish under aquaculture conditions (Kawamoto et al. 1957; Nakamura and Kasahara 1955, 1956, 1957, 1961; Yamagishi 1962, 1969; Yamagishi et al. 1988; Yamagishi and Ishioka 1989).

The outlined effects all have consequences for the derivation of regression models and VBGF parameters.

Heteroskedasticity

The transformation of the length and weight variables in the methods used here, and in other contributions in this volume (Prein; Prein and Milstein), has consequences for the performance of the methods. In both the "extended Gulland-and-Holt" (Pauly et al. this vol.) and "extended Bayley" plots, the points belonging to the fish of medium and larger sizes are clustered near the abscissa, close to L_∞, and have a small amount of variance. The points belonging to smaller fish cover nearly 50% of the entire data range and show a considerably larger amount of variance. Also, much fewer points are located in the data range covered by smaller fish. The residuals of a regression through these points show a trumpet-shaped distribution, indicating heteroskedasticity. One of the main requirements in regression is that of homogeneous variance of the data over the entire data range. Both methods violate this rule.

As a consequence, for estimation of VBGF parameters, the variance in growth rates of smaller fish sizes have a high influence on the estimation of K, as discussed above. Some points may thus have a considerable 'leverage'. The effect is worse in the Bayley plot, since the transformation involved there is radical. Additionally, in the Bayley plot, L_∞ is subject to greater variance, since a minimal change in slope leads to a large change in L_∞, due to the reciprocal transformation of mean length. Heteroskedasticity leads to an inflation of the confidence intervals of the regression coefficients (Norusis 1985). These limitations of the methods must be considered in applications of the regression models.

Comparative Sensitivity Analyses of the Bayley Plot and the Gulland-and-Holt Plot

The behavior of the Bayley plot and the Gulland-and-Holt plot can be investigated and compared through sensitivity analysis using the data in the file PHILSAMP.WK1 (see Appendix II). The procedure adopted here (for the case of the simple versions only) was to use the slopes of both regressions (obtained on a random sample dataset, n = 198, with arithmetic mean regressions) as reference and vary their values in steps of ±10% (Majkowski 1982). Resultant values of K, L_∞ and ϕ' were thus computed and their responses studied (Fig. 4).

In the case of the Gulland-and-Holt plot (Fig. 4A), the change in slope (i.e., the response to variance in the dataset) has only a limited effect on the growth parameters, K and L_∞, with ϕ' compensating the diverging effect. Here the Type I regression is appropriate.

In the case of the Bayley plot (Fig. 4B), the change in slope has a strong effect on the growth parameters, particularly on L_∞ (the latter is due to the nonlinearity of the x-scale). This effect cannot be compensated by ϕ', which suffers a strong bias at slope changes below 30% of the true value.

Hence, given the tendency for a Type I (=AM, or predictive) regression to have a low slope when variance is high, there is a tendency for the Bayley plot to overestimate L_∞ and ϕ', and to underestimate K. This effect can be partly counteracted by using a Type II (=GM, or functional) regression with the Bayley plot. This leads to lower estimates of L_∞ and higher estimates of K and ϕ'. In this case, the GM slope is 26% higher than the AM slope leading to a ϕ' that is 4.7% greater.

The Bayley plot is capable of extracting more information from a dataset, as it uses an additional variable (i.e., weight). In spite of this, the above leads to the conclusion that the Gulland-and-Holt plot is more robust and is easier to use, i.e., (a) does not require individual fish weight which is often not measured, and (b) is directly computed by most statistical software packages.

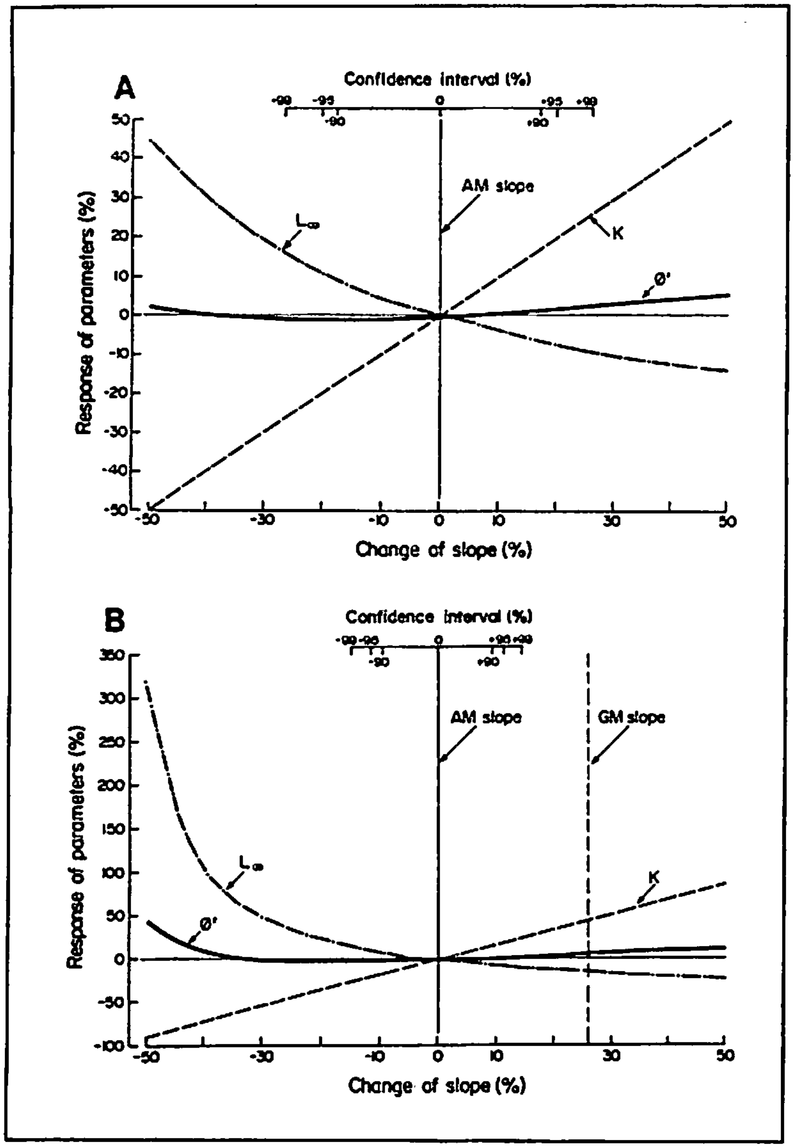

Fig. 4. Sensitivity analysis of A) ordinary Gulland-and-Holt plot, and B) ordinary Bayley plot, based on an AM regression on a random sample dataset (n = 198). In the Gulland-and-Holt plot, ϕ' compensates the effects of slope changes on growth parameters. In the Bayley plot, lower slopes have extreme effects on growth parameters, which ϕ' cannot compensate. Here the GM regression is appropriate. Note large difference of ordinate scales.

On the other hand, the Bayley method should be used only with (a) data with a low amount of variance, and (b) the Type II regression.

Path Analysis

A path diagram for the ordinary Bayley plot is shown in Fig. 5A. The amount of variance explained by the Bayley plot is larger (63%) than in the Gulland-and-Holt plot (28%), denoted by the smaller path coefficient for the unexplained effects (residual term). The slope of the path coefficient for reciprocal mean length is positive, indicating a proportional increase in growth rate with the reciprocal of average fish length.

Based on the same set of variables as used in the extended Gulland-and-Holt plot, the causal path diagram for the extended Bayley plot is shown in Fig. 5B. Through inclusion of auxiliary variables, a greater amount of variance in Nile tilapia growth rate could be accounted for (68%), compared to the ordinary Bayley plot. The structure of the path diagram is the same as for the extended Gulland-and-Holt plot, since the same set of variables was found to be significant. The amount of variance explained by the extended Bayley plot is higher than the amount explained by the extended Gulland-and-Holt plot (40%), although both have the same set of auxiliary variables. In the extended Bayley plot, a much higher portion of the total variance is explained by mean length. Correspondingly, the auxiliary variables participate to a lesser extent in the explanation of variance in growth rate, which is denoted by the smaller values of their path coefficients. The independent treatment variables are also correlated in the extended Bayley plot, yet to a lesser degree.

As auxiliary variables in explaining further variance, three treatment variables are significant in controlling Nile tilapia growth in the manure-fed ponds. These were stocking density (here in a transformed state as the square root of $kg \cdot m^{-3}$), manure loading rate (in form of $kg \cdot ha^{-1} \cdot day^{-1}$) and the pond surface area in m^2. A further variable, solar radiation, reflects uncontrollable environmental effects on fish growth. Three variables have a certain degree of positive correlation among each other (manure input, stocking density, and mean length). This is due to the fact that in all experiments, all three variables increased with experiment duration, due to the experimental design. Solar radiation and pond area are not correlated with any of the other variables.

Taking advantage of the large number of variables available in the ICLARM-CLSU dataset, more detailed causal path models could be designed and tested. Although some of the variables are not significant in directly explaining variance in fish growth, they can be used to reflect secondary causal relationships with other, significant variables. This requires a stepwise process of hypothesis formulation, path diagram design, and multiple regression computation, followed by drawing of the path diagram and inspection of the path coefficients. In case of statistical inconsistencies or implausibilities in terms of biological theory, the process must be repeated again until a correct and explicable model is derived.

After numerous trials, the causal path diagram shown in Fig. 5C was obtained, based on the extended Bayley plot and a reduced set of variables. The diagram represents the same pattern as that built with the extended Gulland-and-Holt method. In the present path diagram, the path coefficients, correlations and residual effects are different only for the part concerning growth rate in weight (W-GRO). This path diagram comprises 10 variables, including length growth rate, where WIND is the cumulative run of wind, CLOUD is the cloud covering, WATEM is the water temperature, and OXY is the early morning dissolved oxygen concentration (as saturation in per cent). Growth rate in weight is influenced by four variables directly, of which two are treatment variables. The variables are the reciprocal mean length, stocking density, early morning oxygen saturation, and pond area. Together, these five variables explain 68% of the total variance in weight growth rate. Two of them are treatment variables. Individually, the contributions of the variables towards explaining total variance are: 39% (1/ML), 5% (POND), 11% (OXY) and 3% (DENS).

The strong influence of early morning dissolved oxygen concentration can be further analyzed with path analysis. Five variables were found significant in predicting OXY. One is a treatment variable (manure input), three are uncontrollable meteorological variables (solar radiation, wind run, and cloud covering) and one is an uncontrollable variable of the pond environment (water temperature). These variables explain 58%

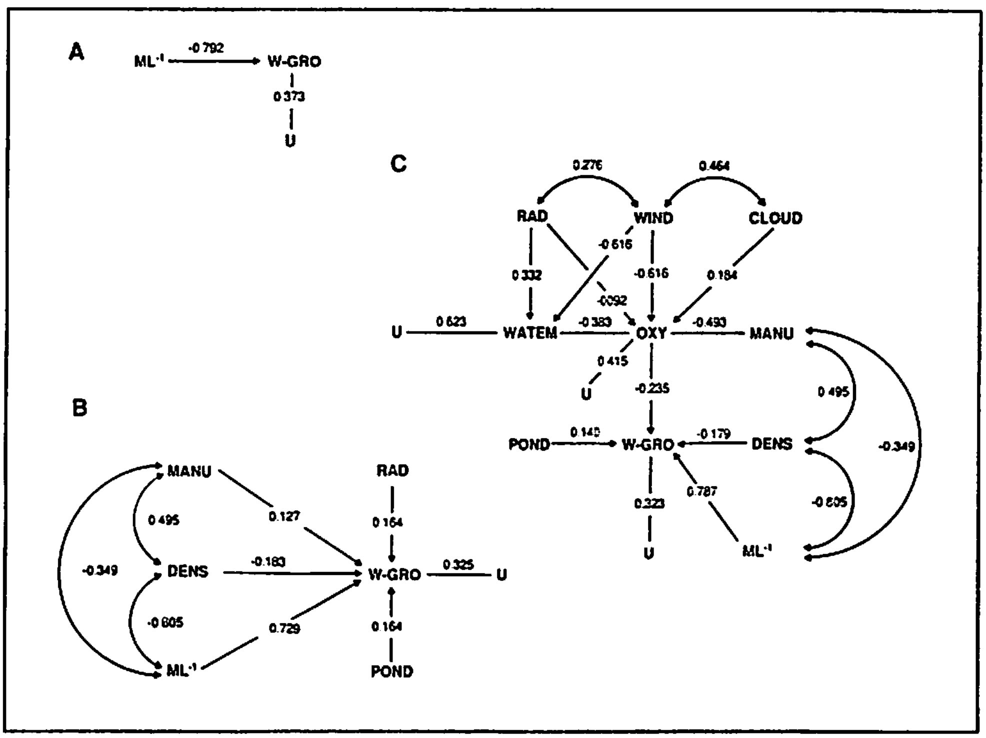

Fig. 5. Path diagrams for A) the ordinary Bayley plot, B) the extended Bayley plot with five predictor variables, and C) the extended Bayley plot with four direct predictor variables and five further explanatory variables. W-GRO = growth rate in weight; ML^{-1} = reciprocal of mean length; MANU = livestock manure input (kg dry weight $ha^{-1} \cdot day^{-1}$); DENS = square root of stocking density ($kg \cdot m^{-3}$); POND = pond area (m^2); RAD = solar radiation ($ly \cdot day^{-1}$); OXY = early morning dissolved oxygen content (% saturation); WATEM = early morning water temperature (°C); WIND = cumulative run of the wind ($km \cdot day^{-1}$); CLOUD = cloud cover (decals); U = residual effect or unexplained variance.

of the total variation in OXY. Their individual amounts of explanation are: MAN 35%, TW 16%, RAD 1.5%, WIND 13% and CLOUD 5%. The meteorological variables show positive correlation. The manure input is correlated with stocking density and mean length, for reasons given above.

Water temperature has a relatively strong influence on OXY. The variance in water temperature can be explained to 38% by solar radiation and wind. Individually, they are responsible for 14% (RAD) and 36% (WIND) of the total variation. Solar radiation increases water temperature, while wind reduces water temperature through evaporative cooling. Cloud covering had an implausible sign. In the present path diagram, solar radiation and wind act twice as predictors (for WATEM and for OXY). Water temperature was not significant as a direct predictor for fish growth.

In the previous path diagrams, solar radiation and manure input were used to directly explain variance in Nile tilapia growth rate. In those models, the amount of explained variance was lower. For management purposes under field conditions in developing countries, manure input and solar radiation are easier to handle in terms of growth prediction than oxygen concentration. In the present detailed path diagram, OXY was incorporated as an intermediate variable. In terms of path analysis, RAD, WIND, CLOUD form compound paths, contributing individually and in a combined manner to the variation in growth rate.

Discussion of Path Analysis

With path analysis, the effects discovered and quantified with the regression methods can be visualized in form of path diagrams. This is possible through their connection with the extended Gulland-and-Holt and extended Bayley methods, which are linear models. Additional analyses can be made through the development of detailed causal path diagrams of the culture systems, based on the available variables and significant relationships between them.

Numerous different path diagrams can be hypothesized with the same dataset, yet there are rules of path analysis and regression which will limit the outcome in terms of plausibility. Newer developments in path analysis allow for the consideration of unmeasured variables but require considerable computational efforts (Blalock 1985a). Further developments have widened the theoretical foundation of path analysis, with the inclusion of effects such as feedback loops (Heise 1975; Jöreskog and Sörbom 1984). With more extensive datasets from aquaculture systems, more detailed analyses with path analysis may be performed in the future, based on the LISREL-approach (Jöreskog and Sörbom 1984) which is a combination of factor analysis and multiple regression.

Conclusions and Recommendations

In the present study, flexible regression models were derived with the "extended Bayley" method and path analysis. With these, and with the "extended Gulland-and-Holt plot" (Pauly et al., this vol.; Prein, this vol., Prein and Milstein, this vol.), growth can be predicted over a wide range of culture conditions, if these are included as parameters in the model. Within the rules of regression, the main influential variables controlling fish growth can be identified and their effects quantified in form of regression coefficients. These are combined in form of VBGF growth parameters.

Depending on the source and quality of the data, considerable efforts may be necessary in the preparation of datasets for analysis, particularly if some variables were not measured. Data from different sources may be merged into one dataset for combined analysis if the species and variables match each other (Prein 1990). The methods are useful analytic tools when the datasets have well-spread variances and wide data ranges for all environmental and treatment variables of interest, as is the case in well-designed factorial experiments. As a whole, the strategy of reanalyzing 'old' data with different new methods has proved rewarding and beneficial, particularly in view of the low costs of such research (involving essentially only personnel cost).

Recommendations for Further Applications

The further successful application of the methods to other 'old' data will depend on the quality of the datasets. These should be inspected for consistency with the rules of multiple regression, but also with the particular requirements of the methods. For example, the extended Bayley method requires precise measurements of both weight and length. Both methods used here cannot accept collinearity among predictor variables. High variance in the datasets due to imprecision or measurement errors cannot be explained by the methods.

It would be rewarding to find and analyze datasets which contain detailed information on pond biology. These variables could not be studied with the datasets analyzed here since they were not measured. Further, simulation studies could provide a better understanding of the sensitivity of the methods towards different amounts of variance in the data.

For the design of new experiments that are to be analyzed with the multivariate methods presented here and in Pauly et al. (this vol.), the following conclusions may be drawn. The main aim should be to have as much variance in the variables as possible in order to avoid collinearity among environmental, treatment and target variables (here fish size). The experimental layout should be in a factorial form, where the fish sizes range from small to large. A wide range of stocking densities is required for all fish sizes and treatments used. This means that small fish would have to be stocked at high densities and, conversely, large fish at low densities. Only with such a spread in the data can the regression describe the effects precisely. Similar requirements of wide data ranges can be made for other treatments, such as manure and feed inputs and, as far as controllable through scheduling, environmental variables such as solar radiation and water temperature.

The experiments do not have to be of long duration. A few, 14-day intervals over a total period of six to eight weeks would suffice for each treatment. It is more important to have a wide range of conditions than many intervals repeating the same few conditions. Due to the distortions of the fish size data caused by the transformations, it must be concluded that more experiments should be made with smaller fish and that these should be sampled at shorter intervals. Greatest care must be taken to obtain precise estimates of average fish size, since the influence of measurement error is greatest in small fish. Larger fish can be sampled at greater intervals. All environmental and treatment variables should be measured at such frequencies that a representative average value can be obtained from them, which reliably reflects the true conditions during the interval.

The present study has shown that the multivariate analysis methods presented here can be used to derive empirical models of fish growth in aquaculture systems. The degree of detail of the models and the accuracy of growth predictions depend on the quality of the datasets used to build the models. More detailed and accurate datasets are more rewarding and permit deeper insights into the qualitative and quantitative relationships governing the growth of tilapia in ponds.

Acknowledgements

We wish to thank Kevin Hopkins and Roger Pullin for providing the raw data for these analyses and Carmela Janagap for performing the computations in the sensitivity analysis. This study was supported in part by the Federal Republic of Germany-Israel Fund for Agricultural Research in Third World Countries (GIARA), grant No. 1-2/86, which is financed by the German Ministry Cooperation (BMZ).

References

Alabaster, J.S. and R. Lloyd, Editors. 1980. Water quality criteria for freshwater fish. Butterworths, London. 297 p.

Backhaus, K., B. Erichson, W. Plinke, C. Schuchard-Ficher and R. Weiber. 1989. Multivariate Analysemethoden, eine anwendungsorientierte Einführung. Fifth edition. Springer Verlag, Berlin. 418 p.

Ball, I.R. 1967. The relative susceptibilities of some species of fresh-water fish to poisons. I. Ammonia. Water Research 1:767-775.

Barthelmes, D. 1960. Probefänge in Karpfenteichen und ihre Bedeutung. Deutsche Fischerei-Zeitung 7:19-23.

Bayley, P.B. 1977. A method for finding the limits of application of the von Bertalanffy growth model and statistical estimates of the parameters. J. Fish. Res. Board Can. 34:1079-1084.

Blalock, H.M. 1972. Social statistics. McGraw-Hill, New York. 583 p.

Blalock, H.M. 1985a. Causal models in the social sciences. Second edition. Aldine, New York. 448 p.

Blalock, H.M. 1985b. Causal models in panel and experimental designs. [Causal models in the social sciences. Companion volume.] Aldine, New York. 287 p.

Boyle, R.P. 1970. Path analysis and ordinal data. Am. J. Sociol. 75:461-480.

Brandstädter, J. 1976. Soziale Schicht, Umwelt und Intelligenz: Eine Pfadanalyse der Korrelationsbefunde von Marjoribanks. Psychologische Beiträge 18:35-53.

Brandstädter, J. and F. Bernitzke. 1976. Zur Tecknik der Pfadanalyse. Ein Beitrag zum Problem der nichtexperimentellen Konstruktion von Kausalmodellen. Psychologische Beiträge 18:12-34.

Brett, J.R. 1979. Environmental factors and growth. p. 559-675. *In* W.S. Hoar, D.J. Randall and J.R. Brett (eds.) Fish physiology. Vol. 8. Bioenergetics and growth. Academic Press, New York. 786 p.

Buschkiel, A.L. 1937. Lehren aus tropischer Teichwirtschaft, Z. Fisch. Hilfswiss. 35:181-197.

Chetty, C.S., R.C. Naidu, Y.S. Reddy, P. Aruna and K.S. Swami. 1980. Tolerance limits and detoxification mechanisms in the fish *Tilapia mossambica* subjected to ammonia toxicity. Indian J. Fish. 27:177-182.

Cloninger, C.R. 1980. Interpretations of intrinsic and extrinsic structural relations by path analysis: theory and applications to assortive mating. Genetic Research 36:133-145.

Coelho, M.L. and A.A. Rosenberg. 1984. Causal analysis of some biological data for *Illex illecebrosus* from the Scotian Shelf. NAFO Sci. Coun. Stud. 7:61-66.

Davidson, F.A., E. Vaughan and S.J. Hutchinson. 1943. Factors influencing the upstream migration of the pink salmon (*Oncorhynchus gorbuscha*). Ecology 24:149-168.

Dörfel, H. 1972a. Pfadkoeffizienten und Strukturmodelle. Biometrische Z. 14:12-26.

Dörfel, H. 1972b. Schätzen von Pfadkoeffizienten. Biometrische Z. 14:209-226.

Dörfel, H. and U. Neumann. 1973. Untersuchung eines Systems der Trieb-und Ertragsbildung bei Apfelbäumen mit Hilfe von Pfadkoeffizienten. Biometrische Z. 15:527-543.

Draper, N.R. and H. Smith. 1981. Applied regression analysis. Second edition. John Wiley, New York. 762 p.

Eknath, A.E. and R.W. Doyle. 1985. Maximum livelihood estimation of 'unobservable' growth and development rates using scale data: application to carp aquaculture in India. Aquaculture 49:55-71.

Heise, D.R. 1969. Problems in path analysis and causal inference. p. 38-73. *In* E.F. Borgatta and G.W. Bohrnstedt (eds.) Sociological methodology 1969. Jossey-Bass, San Francisco. 302 p.

Heise, D.R. 1975. Causal analysis. John Wiley, New York. 301 p.

Hepher, B., A. Milstein, H. Levener and B. Teltsch. 1989. The effect of fish density and species combination on growth and utilization of natural food in ponds. Aquacult. and Fish. Manage. 20:59-71.

Hopkins, K.D. and E.M. Cruz. 1982. The ICLARM-CLSU integrated animal-fish farming project: final report. ICLARM Tech. Rep. 5, 96 p.

Jöreskog, K.G. and D. Sörbom. 1984. LISREL VI: analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods. Third edition. University of Uppsala, Uppsala.

Kang, K.M. and E. Seneta. 1980. Path analysis: an exposition. Dev. Stat. 3:217-246.

Kawamoto, N.Y., Y. Inouye and S. Nakanishi. 1957. Studies on effects by the pond-areas and the densities of fish in the water upon the growth rate of carp (*Cyprinus carpio* L.). Rep. Fac. Fish., Univ. Mie Prefecture 2:437-447.

Kelly, H.D. 1957. Preliminary studies on *Tilapia mossambica* Peters relative to experimental pond culture. Southeastern Association of Game and Fish Commissioners, Proceedings of the 1956 Annual Conference 10:139-149.

Kim, J.-O. and F.J. Kohout. 1975. Multiple regression analysis: subprogram REGRESSION. Chapter 20, p. 320-367. *In* N.H. Nie, C.H. Hull, J.G. Jenkins, K. Steinbrenner and D.H. Bent (eds.) SPSS statistical package for the social sciences. Second edition. McGraw-Hill, New York. 675 p.

Land, K.C. 1969. Principles of path analysis. p. 3-37. *In* E.F. Borgatta and G.W. Bornstedt (eds.) Sociological methodology 1969. Jossey-Bass, San Francisco. 302 p.

Li, C.C. 1975. Path analysis: a primer. The Boxwood Press, Pacific Grove, California. 347 p.

Lovshin, L.L. 1984. Letter to the editor concerning the article by D. Pauly and K.D. Hopkins: "A method for the analysis of pond growth experiments". ICLARM Newsl. 7(2):30.

Majkowski, J. 1982. Usefulness and applicability of sensitivity analysis in a multispecies approach to fisheries management, p. 149-165. *In* D. Pauly and G.I. Murphy (eds.) Theory and management of tropical fisheries. ICLARM Conf. Proc. 9, 360 p.

Nakamura, N. and S. Kasahara. 1955. A study on the phenomenon of the tobi koi or shoot carp. I. On the earliest stage at which the shoot carp appears. Bull. Jap. Soc. Sci. Fish. 21:73-76. (In Japanese with English summary).

Nakamura, N. and S. Kasahara. 1956. A study on the phenomenon of the tobi koi or shoot carp II. On the effect of particle size and quantity of food. Bull. Jap. Soc. Sci. Fish. 21:1022-1024. (In Japanese with English summary).

Nakamura, N. and S. Kasahara. 1957. A study on the phenomenon of the tobi koi or shoot carp III. On the results of culturing the modal group and the growth of carp fry reared individually. Bull. Jap. Society Sci. Fish. 22:674-678. (In Japanese with English summary).

Nakamura, N. and S. Kasahara. 1961. A study on the phenomenon of the tobi koi or shoot carp. IV. Effects of adding a small number of larger individuals to the experimental batches of carp fry and culture density upon the occurence of shoot carp. Bull. Jap. Soc. Sci. Fish. 27:958-962. (In Japanese with English Summary).

Nakanishi, T. and H. Onozato. 1987. Variability in the growth of isogeneic crucian carp *Carassius gibelo langsdorfii*. Nippon Suisan Gakkaishi 53:2099-2104.

Norusis, M.J. 1985. SPSS-X advanced statistics guide. McGraw-Hill, New York and SPSS Inc., Chicago. 505 p.

Pauly, D. 1986. A simple method for estimating the food consumption of fish populations from growth data and food conversion experiments. U.S. Fish. Bull. 84(4):829-842.

Piedrahita, R.H., J.H. Ebeling and T.M. Losordo. 1987. Use of data acquisition systems in aquaculture, p. 259-262. *In* J.G. Balchen (ed.) Automation and data processing in aquaculture. Proceedings of the International Federation of Automatic Control (IFAC) Symposium, Trondheim, Norway, 18-21 August 1986. Pergamon Press, Oxford. 280 p.

Prein, M. 1985. The influence of environmental factors on fish production in tropical ponds investigated with multiple regression and path analysis. Mathematisch-Naturwissenschaftliche Fakultät, Christian-Albrechts-Universität Kiel, 91 p. M.S. thesis (Diplomarbeit).

Prein, M. 1990. Multivariate analysis of tilapia growth experiments in ponds: case studies from the Philippines, Israel, Zambia and Peru. Mathematisch-Naturwissensschaftliche Fakultät, Christian-Albrechts-Universität Kiel, 125 p. Dissertation.

Rasch, D., Editor. 1983. Biometrie, Einführung in die Biostatistik. VEB Deutscher Landwirtschaftsverlag, Berlin, 276 p.

Redner, B.D. and R.R. Stickney. 1979. Acclimation to ammonia by *Tilapia aurea*. Trans. Am. Fish. Soc. 108:383-388.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191. Ottawa, 282 p.

Robinson, B.W. and R.W. Doyle. 1988. Behavior phenotypes of juvenile tilapia and their correlations with individual growth rates. p. 503-507. *In* R.S.V. Pullin, T. Bhukuswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Sanders, D. 1980. Path analysis/causal modelling. Quality and Quantity 14:181-204.

Schaeperclaus, W. 1952. Fischerkrankungen und Fischsterben durch Massenentwicklung von Phytoplankton bei Anwesenheit von Ammonium. Z. Fisch. Hilfswiss. 1:29-44.

Schwinghamer, P. 1983. Generating ecological hypotheses from biomass spectra using causal analysis: a benthic example. Mar. Ecol. (Prog. Ser.) 13:151-166.

Snedecor, G.W. and W.G. Cochran. 1980. Statistical methods. Seventh edition. The Iowa State University Press, Ames, Iowa. 507 p.

Sousa, R.J., T.L. Meade and R.E. Wolke. 1974. Reduction of ammonia toxicity by salinity and pH manipulation. Proc. World Maricult. Soc. 5:343-354.

Sparre, P., E. Ursin and S.C. Venema. 1989. Introduction to tropical fish stock assessment. Part 1. Manual. FAO Fish. Tech. Pap. 306.1, 337 p.

Spehar, R.L., G.M. Christensen, C. Curtis, A.E. Lemke, T.J. Norberg and Q.H. Pickering. 1982. Effects of pollution on freshwater fish. J. Water Pollut. Control Fed. 54:877-922.

Steinmann, P. and G. Surbeck. 1922. Fischvergiftungsversuche mit frischem und fauligem Harn, mit Ammoniak und Ammoniumverbindungen. Schweizerische Fischerei-Zeitung 30:123-129, 145-152.

Svärdson, P. 1984. A Monte Carlo study of three methods for estimating the parameters in the von Bertalanffy growth equation. J. Cons. CIEM 41:248-258.

Turner, M.E. and C.D. Stevens. 1959. The regression analysis of causal paths. Biometrics 15:236-258.

Weede, E. 1970. Zur Methodik der kausalen Abhängigkeitsanalyse (Pfadanalyse) in der nicht-experimentellen Forschung. Kölner Z. Soziol. Sozialpsychol. 22:532-550.

Weisberg, S. 1980. Applied linear regression. John Wiley, New York. 283 p.

Wohlfarth, G.W. 1977. Shoot carp. Bamidgeh 29(2):35-55.

Wohlfarth, G.W. and R. Moav. 1969. The genetic correlation of growth rate with and without competition in carp. Verh. Int. Ver. Limnol. 17:702-704.

Wright, S. 1921. Correlation and causation. J. Agric. Res. 20:557-585.

Wright, S. 1923. The theory of path coefficients - a reply to Niles's criticism. Genetics 8:239-255.

Wright, S. 1934. The method of path coefficients. Annals Math. Stat. 5:161-215.

Wuhrmann, K. and H. Woker. 1953. Über die Giftwirkung von Ammoniak- und Zyanidlösungen mit verschiedener Sauerstoffspannung und Temperatur auf Fische. Schweizer Z. Hydrol. 15:235-260.

Yamagishi, H. 1962. Growth relation in some small experimental populations of rainbow trout fry, *Salmo gairdneri* Richardson with special reference to social relations among individuals. Jap. J. Ecol. 12:43-53.

Yamagishi, H. 1969. Postembryonal growth and its variability of the three marine fishes with special reference to the mechanism of growth variation in fishes. Res. Popul. Ecol. 11:14-33.

Yamagishi, H. and K. Ishioka. 1989. Simulation study on the growth variation of fish with special reference to the sexual dimorphism in growth. Bull. Tokai Reg. Fish. Res. Lab. 127:1-16.

Yamagishi, H., K. Ishioka and T. Matsumoto. 1988. Simulation study on the growth variation of fish with special reference to the exponential and asymptotic growth. Bull. Tokai Reg. Fish. Res. Lab. 125:15-31.

Yashouv, A. 1969. The fish pond as an experimental model for study of interactions within and among populations. Verh. Int. Ver. Limnol. 17:582-593.