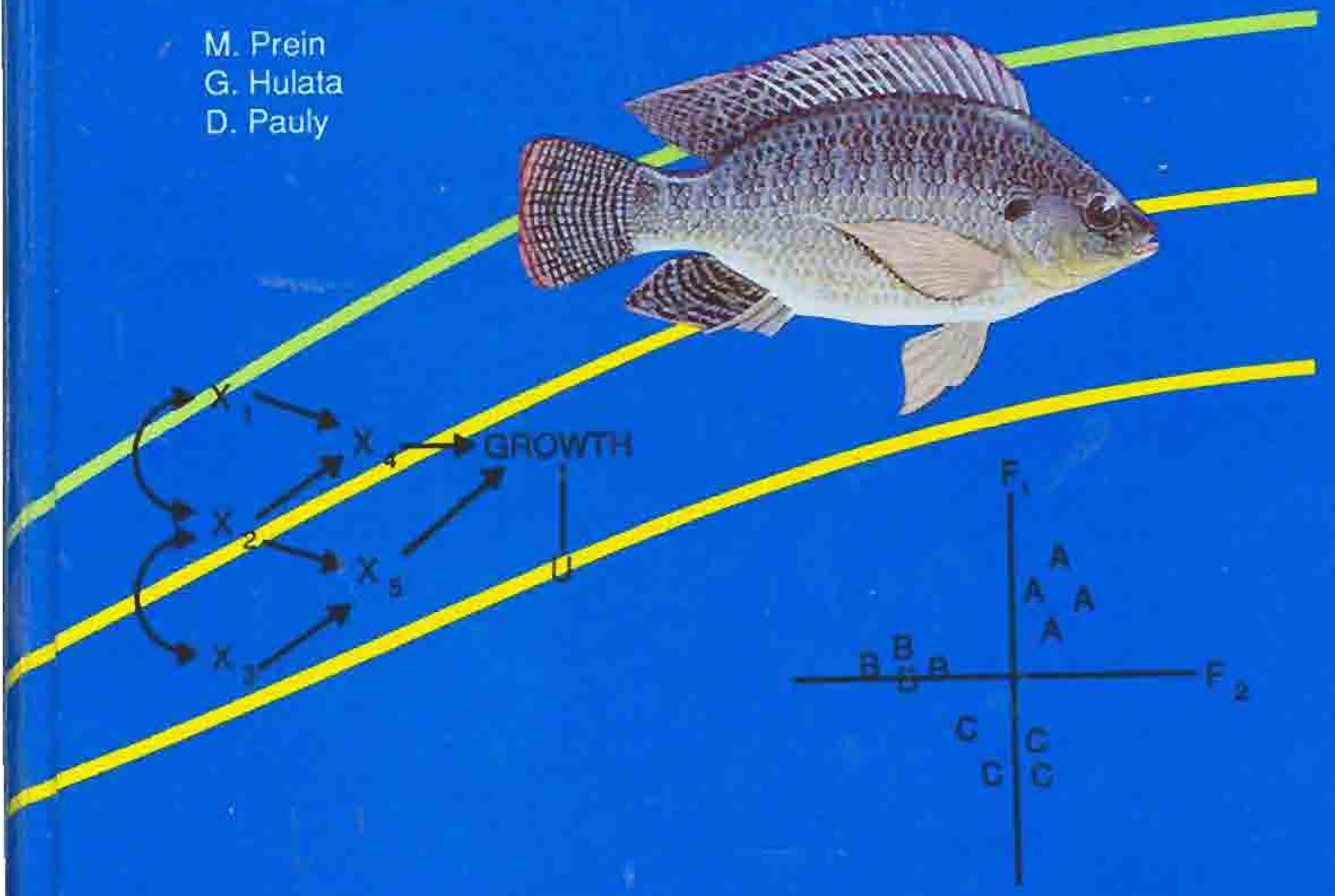
Multivariate Methods in Aquaculture Research: Case Studies of Tilapias in Experimental and Commercial Systems

Edited by

M. Prein
G. Hulata
D. Pauly



ICLARM INTERNATIONAL CENTER FOR LIVING AQUATIC RESOURCES MANAGEMENT, PHILIPPINES

AGRICULTURAL RESEARCH ORGANIZATION, ISRAEL

BMZ BUNDESMINISTERIUM FÜR WIRTSCHAFTLICHE ZUSAMMENARBEIT UND ENTWICKLUNG (BMZ), FEDERAL REPUBLIC OF GERMANY

Multivariate Methods in Aquaculture Research: Case Studies of Tilapias in Experimental and Commercial Systems

Edited by

M. Prein
G. Hulata
D. Pauly

1993

INTERNATIONAL CENTER FOR LIVING AQUATIC RESOURCES MANAGEMENT
PHILIPPINES

AGRICULTURAL RESEARCH ORGANIZATION
ISRAEL

BUNDESMINISTERIUM FÜR WIRTSCHAFTLICHE ZUSAMMENARBEIT UND ENTWICKLUNG (BMZ)
FEDERAL REPUBLIC OF GERMANY

Multivariate Methods in Aquaculture Research: Case Studies of Tilapias in Experimental and Commercial Systems

EDITED BY

M. PREIN
G. HULATA
D. PAULY

1993

Published by the International Center for Living Aquatic Resources Management (ICLARM), MCPO Box 2631, 0718 Makati, Metro Manila, Philippines; Agricultural Research Organization (ARO), The Volcani Center, P.O. Box 6, Bet Dagan 50250, Israel; and Bundesministerium für Wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), Karl-Marx-Strasse 4-6, D-5300 Bonn, Federal Republic of Germany.

Printed in Manila, Philippines

Prein, M., G. Hulata and D. Pauly, Editors. 1993. Multivariate methods in aquaculture research: case studies of tilapias in experimental and commercial systems. ICLARM Stud. Rev. 20, 221 p.

ISSN 0115-4389
ISBN 971-1022-85-0

Cover design by Mark Prein and Angela Egold. Picture of Nile tilapia (*Oreochromis niloticus*) is a computerized image of this species as provided in FishBase, ICLARM's comprehensive database on fishes of the world (digitization artist: Robbie Cada).

ICLARM Contribution No. 669

Dedicated to the memory
of
Dr. Balfour Hepher
(1925-1988)

Contents

ICLARM Foreword

One of the goals of ICLARM is the development of appropriate quantitative methods for the general area of tropical fisheries and aquaculture research.

Outside of ICLARM, the need for the development of such methods is widely acknowledged with regard to fisheries, but this is not so for aquaculture, perhaps because it is often perceived as requiring methods and approaches that are site- or species-specific.

The present volume, the result of fruitful cooperation between the Agricultural Research Organization (ARO) and ICLARM, and of generous support by the German Ministry for Economic Cooperation and Development (BMZ), may help redress the methodological imbalance between fisheries and aquaculture research, and initiate more work on methods applicable to both fields. This volume emphasizes the following points:

- Although aquaculture can be very site-specific, there are many questions that can be tackled on a comparative basis, across sites - given that suitable concepts and quantitative methods are applied.
- Aquaculture experiments and commercial production systems generate a large amount of data that are not only *suitable* for computer-based quantitative analysis, but actually *require* such analysis, just as with the massive catch and catch composition datasets generated by commercial fisheries.
- The standard methods used by aquaculturists to analyze their experimental results - Latin square and related designs for the experimental layout, ANOVA for analysis of yields, absolute or relative growth rates for analysis of fish growth in ponds, etc. - may generally fail to extract important information, mainly because they cannot account for the multifactor, synergistic effects which only the analysis of large datasets with multivariate methods can make visible.

ICLARM has given much emphasis, in its various aquaculture projects, to the tilapias, both as crucial elements in aquaculture-agriculture integration schemes, and as test animals for genetic improvement of organisms used in tropical aquaculture. I am therefore particularly pleased to see that this book documents our emphasis on both methodology development and on tilapias in one volume.

I hope that the methods illustrated in this book will find wide application and be further developed. To encourage this, a set of four diskettes has been assembled, which contains the data files used by the authors of the various contributions assembled here. These diskettes are available for a nominal fee from ICLARM.

Finally, I take this opportunity to thank ARO, particularly Gideon Hulata and Ana Milstein, for their cooperation in this project and BMZ for its unflagging support of the project which led to this book.

DR. ROGER S.V. PULLIN
Director
Aquaculture Program
ICLARM

ARO Foreword

It is generally considered that aquaculture has the potential to generate additional, diverse sources of protein for the growing human population, particularly in developing countries. With the aim to strengthen the research base on tropical pond aquaculture systems, ARO of Israel's Ministry of Agriculture joined in with ICLARM. On the Israeli side, this work was performed by the staff of the Fish and Aquaculture Research Station at Dor. This work was part of a larger Aquaculture Project, funded by the Federal Republic of Germany-Israel Fund for Agricultural Research in Third World Countries (GIARA), which is financed by BMZ.

The aim of the project was to retrieve available data from experiments and commercial farms from tropical countries and Israel and to analyze these with new multivariate statistical methods. Israel has a long tradition of pond aquaculture with fish species used in many tropical countries, notably tilapia and carp. The knowledge and expertise gathered in Israel may be valuable for adaptation to the conditions prevailing in developing countries. The results of the present analyses may in turn be of interest to Israeli fish farmers. I appreciate, together with the authors, their participation and willingness to make their farm records available to the present project.

I thank Dr. Ana Milstein and Dr. Gideon Hulata for their highly successful efforts in coordinating and performing this project, and for cooperating so productively with the ICLARM counterparts, Dr. Daniel Pauly and Dr. Mark Prein, to produce this fine volume.

Finally, I deeply appreciate the decision of the editors to dedicate this book to the memory of the late Dr. Balfour Hepher, who has contributed significantly to the development of aquaculture and has forwarded the idea of applying multivariate analysis to aquaculture research.

PROFESSOR YESHAY FOLMAN
Chief Scientist
Ministry of Agriculture

BMZ Foreword

Research for the benefit of developing countries is being supported by the Bundesministerium für Wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) in a wide range of fields and with a broad range of partners. The incentive behind one line of such activities was the cooperation with Israeli scientists in the field of agriculture, in this particular case of aquaculture, to develop aquaculture systems appropriate for tropical and subtropical countries. This was performed in terms of the Aquaculture Project of the Federal Republic of Germany-Israel Fund for Agricultural Research in Third World Countries (GIARA). A further cooperating institution was the International Center for Living Aquatic Resources Management (ICLARM) in Manila. The results presented in this volume are to a greater part the outcome of Subproject 1: Optimal Management of Aquaculture Pond Systems in Developing Countries, Part II: Multivariate Analysis of Existing Farm Datasets.

The project was suggested and negotiated by Dr. Martin Bilio of the Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH and supported by Mr. Thomas Schurig of the BMZ. The smooth administrative processing was done by Mrs. Miriam Bar-Lev of the Centre for International Development Coordination (CINADCO) on the Israeli side and Dr. Claudia Wiedey of the German Council for Tropical and Subtropical Agricultural Research (ATSAF), on the German side. All their efforts towards a successful functioning of the project are gratefully acknowledged.

We are pleased that the activities of the cooperating scientists in Israel and from ICLARM harmonized so well and led to this productive outcome, underlining the value of the results for scientists in developing countries. The book contains a section in which the different methods used herein are described in detail, so as to facilitate the application of these methods by other researchers, notably in the tropics and subtropics. The examples in which the methods are applied cover a wide range of aquaculture types and systems.

From a donor's point of view, the cost- and resource- saving approach is welcomed as to sharing and using available datasets as a basis for analysis work with the aim of extracting more information from them than has already been done through the application of new methods or to use different analysis methods on the same data.

We therefore are happy that the project has been so successful and that the results are presented in such a fine comprehensive form. We hope that these will find due appreciation in the scientific community, and are proud that we have contributed to it.

Dr. Hans-J. de Haas
BMZ

On the Use of Multivariate Statistical Methods in Aquaculture Research*

Mark Prein, *International Center for Living Aquatic Resources Management, MCPO Box 2631, 0718 Makati, Metro Manila, Philippines*

Gideon Hulata, *Agricultural Research Organization, Fish and Aquaculture Research Station, Dor, Mobile Post Hof Hacarmel 30820, Israel*

Daniel Pauly, *International Center for Living Aquatic Resources Management, MCPO Box 2631, 0718 Makati, Metro Manila, Philippines*

Prein, M., G. Hulata and D. Pauly. 1993. On the use of multivariate statistical methods in aquaculture research, p. 1-12. *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

This contribution presents the rationale for the other chapters in this book, to which it serves as introduction. We discuss sequentially: our emphasis on tilapia (*Oreochromis* spp., Fam. Cichlidae), modelling of pond aquaculture, exploratory and confirmatory methods for multivariate analysis, the relationships between multiple regression and analysis of variance and the rationale for sharing and analyzing "old", public domain data, all with emphasis on tropical/subtropical aquaculture systems.

Introduction

This contribution presents the rationale for the application of multivariate statistics, especially of multiple regression (MR), for analysis of data from aquaculture systems, and for our choice of commercial and experimental culture of tilapia (*Oreochromis* spp., Fam. Cichlidae) for the case studies included in this book.

The version of MR given most emphasis in this volume is the extended Gulland and Holt plot, i.e., a multidimension extension of a bivariate plot, the "Gulland and Holt plot", commonly used by fisheries scientists to estimate the parameters of the von Bertalanffy growth function (VBGF)(Gulland and Holt 1959; Pauly 1984).

This extension was initially used for the analysis of coral fish growth (Pauly and Ingles 1981), and the first application of this method to aquaculture data was presented by Pauly and Hopkins (1983, see Appendix I). The method needed further validation, however, and this was performed by M. Prein in the context of an MS thesis (Prein 1985) which used data from the Philippines.

The conclusions were encouraging, and this led to a decision to continue this work in the form of a doctoral thesis, completed five years later (Prein 1990). The character of the second phase of this investigation of the applicability of multivariate methods to aquaculture was strongly shaped, however, by a chance meeting in Lima, Peru, in early March 1985, of Dr. B. Hepher with D. Pauly, where the former informed the latter of his encouragement of similar work being conducted at Dor Station in Israel by Ms. Ana Milstein and associates.**

We decided to cooperate and found a far-sighted donor, in Germany, to support that cooperation.

Why Tilapia for This Book?

Aquaculture of tilapia (Pisces, Cichlidae) has progressed much in the past three decades (Balarin and Hatton 1979; Pullin and Lowe-McConnell 1982; Fishelson and Yaron 1983; Pullin et al. 1988). The main regions of production increase are several Asian countries and Israel. The culture of tilapia has so far expanded little in Africa (Huisman 1986; Balarin 1988), although these fish are endemic to the continent (Fryer and Iles 1972; Trewavas 1983; Pullin 1988), and in Latin America (Verreth et al. 1987), where it was introduced. However, interest in tilapia culture is widespread throughout the tropics (Pullin 1985, 1991).

Tilapia can tolerate a wide range of environmental conditions, such as temperatures (Kutty and Sukumaran 1975; Balarin 1988) and dissolved

*ICLARM Contribution No. 443.

**Our dedication of this book to Dr. B. Hepher is thus not only to express our respect: he *is* the one who made this book happen.

oxygen levels (Dusart 1963; Peer and Kutty 1981; Fernandes and Rantin 1987; Ojda 1987; Tsadik and Kutty 1987). They can be grown commercially at high stocking densities (Hepher and Pruginin 1981), tolerate frequent handling and are resistant to fish diseases. Tilapia thrive on diets such as detritus (Vaas and Hofstede 1952; Bowen 1981, 1982; Hofer and Schiemer 1983; Salvadores and Guzman 1983; Moriarty and Pullin 1987; Pauly et al. 1988), phytoplankton and zooplankton (Le Roux 1956; Doha and Haque 1966; Moriarty 1973; Saha and Dewan 1979; Hepher 1988). The locations of the tilapia culture experiments and commercial production systems analyzed in this volume are depicted in Fig. 1.

A marked disadvantage of tilapia for aquaculture is their ability to reproduce early in life. This leads to overpopulation in the ponds and "dwarfing" of the fish (Lowe-McConnell 1958; Iles 1973). These can be overcome either through monosex culture, based on manual sexing or hormonal sex inversion (Sanico 1975; Wohlfarth and Hulata 1983), through polyculture with predators (Hopkins et al. 1982) or hybridization with other species (Wohlfarth and Hulata 1983).

Tilapias are excellent model species for genetic research because of their worldwide importance in warmwater aquaculture and their short generation. These qualities make them very attractive for investigating the application of genetics in aquaculture, from conservation of genetic resources to breeding programs (Pullin et al. 1991).

Tilapia can be reared in a wide range of culture systems, ranging from normal earthen ponds or sewage stabilization ponds (Hey 1955; Edwards 1992; Edwards and Pullin 1990; Bartone et al. 1985; Gaigher and Toerien 1985; Baozhen 1987; Cointreau 1987), net cages (Armbrester 1971), flowthrough tanks (Uchida and King 1962) and raceways (Balarin and Haller 1983) to recirculation systems in tanks and silos (Mironova and Skvortsova 1967; Meske 1980; Spotts 1983; Zentrum für Angepasste Technologie und Sozialökonomie Langenbruck 1983; Rennert and Steffens 1985; Provenzano and Winfield 1987). They also tolerate both freshwater and seawater (Hopkins et al. 1986; Cheong et al. 1987). The bulk of the commercially produced tilapia though, is grown in freshwater, earthen ponds at locations in the intertropical belt, where these fish serve as

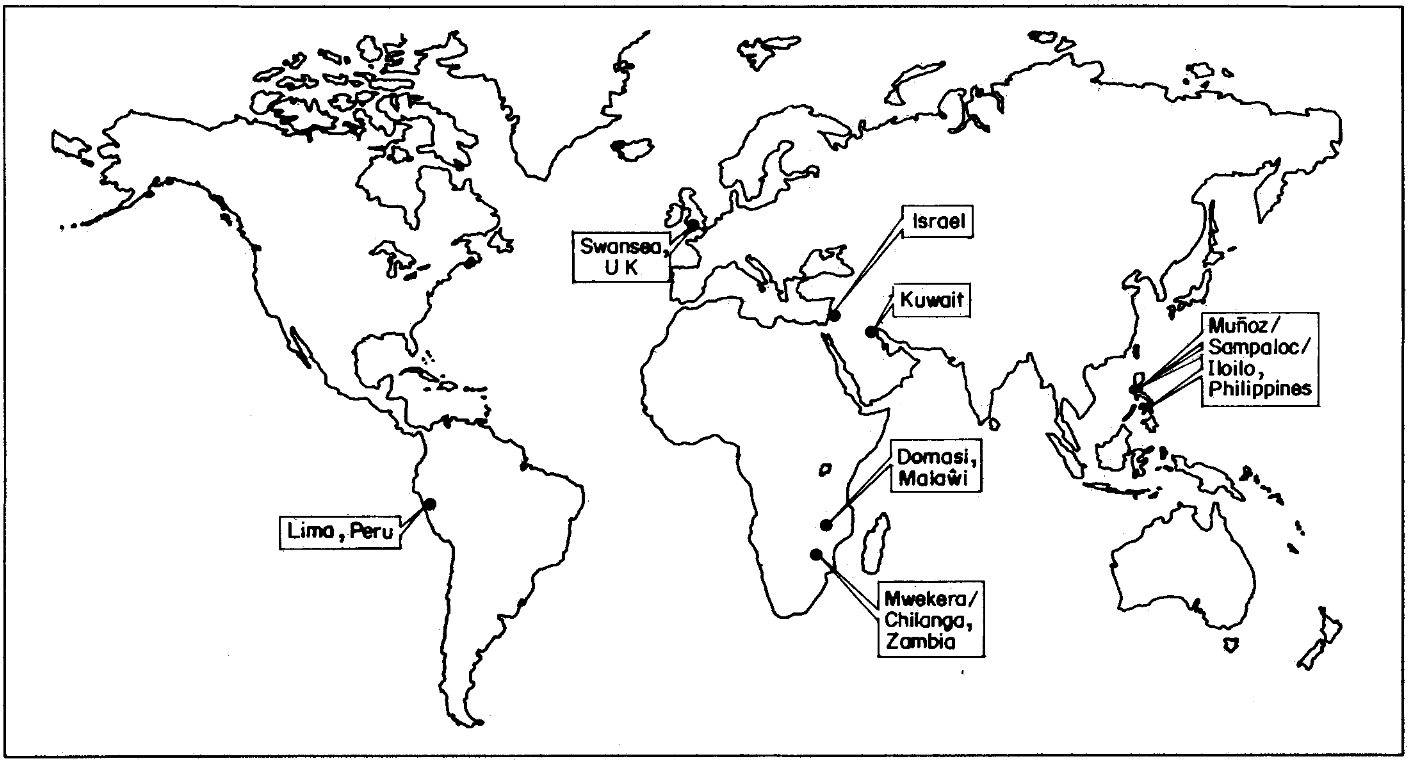


Fig. 1. Locations of data sources for multivariate analyses of tilapia growth presented in this volume (dots): Israel, Kuwait, Malaŵi, Peru, Philippines, United Kingdom and Zambia.

an important protein source for the local human population. Three types of systems of tilapia culture in earthen ponds can be distinguished (Pullin 1989):

1. *Extensive pond systems* receive no or only low levels of energy and nutrient input and low management requirements. These are conducted in monoculture or in polyculture with various species at low densities, receiving no or only little fertilization in form of organic manure. These systems rely mainly on the natural productivity of the pond environment. Some forms of integrated crop-livestock-fish farming fall into this category (Schroeder 1986; Little and Muir 1987).
2. *Semi-intensive pond systems* receive intermediate levels of energy inputs (mainly in form of labor) and management efforts. Ponds are well-kept by their operators, with intermediate stocking densities, and receive sufficient applications of manure, fertilizer and simple feeds in form of rice bran, sorghum or grain meals. Most forms of integrated crop-livestock-fish farming systems are of the semi-intensive type (Pullin and Shehadeh 1980; Little and Muir 1987; Edwards et al. 1988).
3. *Intensive pond systems* require high levels of energy inputs (in form of electricity or fuel), manure and fertilizer applications, with direct feeding of the fish with artificial feeds, i.e., pellets. Partly, aeration is applied on a nightly schedule. Tilapia are grown in polyculture with different species, such as carp, mullet, catfish, gourami and prawn (Hepher and Pruginin 1981).

Models in Pond Aquaculture

A multitude of factors interact dynamically in fishponds and influence fish production. Some of the factors are uncontrollable (e.g., meteorological variables), while others may be regulated through management (e.g., stocking density, manure, fertilizer and feed applications, aeration, pond design and water exchange). Their interactions and their resulting effects on fish growth are only partly understood. Pond managers and scientists strive to make the behavior of these systems more predictable, particularly as the systems themselves undergo internal changes over a culture cycle. The ultimate goal is to have a mathematical model which describes the fishpond system to a required degree of precision. With such a model, predictions of fish growth and production can be made, while simultaneously considering variations, for example, of water temperature and stocking density. Management decisions can be made and culture strategies can be planned according to economic constraints and market demand (e.g., fish size or harvest date).

Several modelling applications to aquaculture systems exist (Allen et al. 1984; Bernard 1986; Balchen 1987; Cuenco 1989). Usually, when these applications deal with fish growth, the models pertain to species of high market value, such as catfish (Miller 1985; Cuenco et al. 1985a, 1985b, 1985c; Machiels and Henken 1986, 1987; Machiels and van Dam 1987; Machiels 1987), trout (Sparre 1976; Jϕrgensen 1976; Sperber et al. 1977; Elliott 1979; From and Rasmussen 1984), eel (Houvenaghel and Huet 1987), gilthead seabream (Unger 1983), lobster (Botsford et al. 1974, 1975; Botsford 1977; Rauch et al. 1975; Botsford and Gossard 1978), or shrimp (Azizan 1983; Huang et al. 1976; Polovina and Brown 1978; Griffin et al. 1981, 1984; Leung et al. 1984).

Bioenergetic models of Nile tilapia (*Oreochromis niloticus*) growth were developed by Melard (1987) and Ross et al. (1988). For studying the production of Nile tilapia in ricefields, van Dam (1989) suggested two modelling approaches, i.e., an empirical multiple regression model and a dynamic simulation model. Bhattarai et al. (1986) published a model of Nile tilapia growth in septage-fed ponds, based on nitrogen loading rates. McDonald (1984) developed a set of equations describing phytoplankton ingestion by blue tilapia (*O. aureus*) and its effect on the algal concentrations. Griffin et al. (1980) presented a bioeconomic model of tilapia production in a chicken manure-fed wastewater treatment system.

Common to most of these models is the need for detailed information on the quality and quantity of nutrients available to the fish in form of feeds or phytoplankton and on oxygen content. Models exist to represent such variables in the pond environment (Piedrahita 1988). These models enable estimation, for example, of plankton availability to the fish (Patten 1968; Piedrahita 1984; Piedrahita et al. 1984; Svirezhev et al. 1984; Laws et al. 1985; Smith and Piedrahita 1988) and diurnal oxygen dynamics (Busch et al. 1977; Boyd et al. 1978; Romaire 1979; Romaire and Boyd 1979; Meyer and Brune 1982; Meyer et al. 1983; Piedrahita et al. 1984; Piedrahita 1988). The variables required to parametrize these models are numerous though, and are not routinely measured in most pond experiments. Further, published models are usually calibrated using the very dataset from which they were developed and are therefore valid only for very similar environments. When aquaculture systems are not well-studied (i.e., when detailed measurements of the pond environment at short intervals are not made), models cannot be developed without making highly simplifying assumptions.

Multivariate Analysis of Pond Aquaculture

An alternative to these mechanistic, or "internally descriptive" models (Piedrahita 1988), is the development of empirical models, based on statistical relationships. In pond experiments, only a limited number of variables can be measured, usually inputs and outputs, i.e., the internal processes of the pond ecosystem are treated as "black boxes" (Pauly and Hopkins 1983; Piedrahita 1988). Given some knowledge on the pond ecosystem, however, causal hypotheses linking inputs and outputs can be formulated, which can then be tested for statistical significance, based on measured data.

In aquaculture research, the statistical methods used for the derivation of empirical relationships are mostly univariate or bivariate in nature (e.g., t-test, correlations, linear regression, etc.). The common approach of the analyses is to deduct key relationships between variables in the systems or to reject/confirm hypotheses formulated *a priori*, through pairwise comparisons of variables.

To move beyond this, two approaches are possible. Either available data from experiments initially designed for analysis with univariate or bivariate methods can be assembled and standardized, then reanalyzed using multivariate methods. This approach is the one that was used throughout most of this book. The alternative is to design and conduct new experiments for the explicit purpose of analysis with a particular multivariate method. The advantage of the former approach is that large amounts of data from all kinds of experiments, species and systems are already available in the desk drawers of many laboratories all over the world. The disadvantage is that the "experimental design" will have shortcomings; these will generally reduce the extent to which the insights gained are as innovative and far-reaching as initially hoped for. The latter approach permits the researcher to exploit the full analytical potential of a given method. However, the experiments that are required will be costly and time-consuming.

Overview of Methods for Multivariate Analysis

Multivariate analysis techniques can be classified in two different ways (Backhaus et al. 1989): the first of these distinguishes the individual methods according to the requirements towards the data to be analyzed. The second classifies the methods according to application-oriented criteria.

With the first type of classification, the data required for analysis either exist or are still to be collected. The quality of the data depends, among others, upon the mode of measurement. Some variables are relatively easy to measure (e.g., fish length) and others are extremely difficult (e.g., natural food availability in a fishpond).

The numbers assigned to a certain variable are arranged on a certain scale, which is either nonmetric (with nominal and ordinal scales) or metric (with interval or ratio scales). In some cases, nominal variables can be expressed metrically in form of dummy variables (see below).

The second way of classification of analysis methods distinguishes between mainly exploratory (or structure-identifying) methods and mainly confirmatory (or structure-testing) methods (Mosteller and Tukey 1977; Tukey 1977; Backhaus et al. 1989). These two criteria can be described as follows:

1. *Exploratory methods* are multivariate analysis methods aimed primarily at identifying interrelationships and structures among the variables.

   In this case, at the beginning of the analysis, the researcher does not have a precise understanding as to which relational interdependencies exist in a dataset. Methods which can identify possible relationships in a dataset are factor analysis/principal component analysis, cluster analysis and multidimensional scaling (Backhaus et al. 1989).
2. *Confirmatory methods* have the primary aim of testing the hypothesized relationships between variables. The researcher has an understanding of the interrelationships based on theoretical reasoning or logical conclusions from known facts, and these lead to the formulation of hypotheses which are then tested and quantified. The

relevant multivariate methods are correlation analysis, MR analysis, canonical correlation, analysis of variance (AV), analysis of covariance (ACV), discriminant analysis, conjoint analysis and the LISREL (Linear Structural Equation Modelling) approach to causal (or path) analysis (Morrison 1976; Marinell 1977; Haf and Cheaib 1985). Application of these methods to aquaculture may be found in Morita (1977), Jana and De (1983), Prikyrl and Kepr (1984), Eknath and Doyle (1985), Milstein et al. (1988, 1989) and Prein (1985, 1987, 1990).

Exploratory Methods

Exploratory methods have in common that they are mostly applied to datasets with numerous variables collected to answer a specific question. Here, the methods are used to reduce the large amount of individual information embedded in distinct variables and objects into a few key components representing the major elements or relationships in the data. Factor analysis is used to reduce the number of variables in a dataset to few, central, "zero-loading" variables. Cluster analysis combines traits of objects, and not variables, to describe groups (or clusters) in which the objects are very similar to each other in relation to their descriptive traits (variables). Within a cluster the objects should be homogeneous, whereas the clusters should be heterogeneous among each other. Multidimensional scaling (MDS) only searches for global similarities between objects (which are not described by variables) and attempts to depict the measured similarities in a space of least possible dimensions (Backhaus et al. 1989).

Confirmatory Methods

In all hypothesis-testing methods, the variables are divided into dependent and independent variables, and the researcher is assumed to have an *a priori* understanding of the basic relationships among the variables before conducting the analysis. In MR, hypothesized relationships between one dependent variable and several independent variables can be tested.

If the independent variables are measured on a nominal scale, and the dependent variable on a metric scale, then analysis of variance is applied. On the other hand, if the dependent variable is on a nominal scale and the independent variable is on a metric scale, then discriminant analysis is used. The latter method can also be used to test the results of cluster analysis, where the object grouping (clusters) are regarded as dependent (nominally scaled) variable (Backhaus et al. 1989).

The methods listed above assume that all variables are real and measurable. In cases where experimental or natural conditions preclude the measurement of important key variables, the LISREL approach to path (or causal) analysis is advisable (Jöreskog and Sörbom 1984). Here unmeasurable variables are accounted for through hypothetical constructions, termed "latent variables".

A notable example here is the study by Eknath and Doyle (1985). A LISREL model was developed by specifying *a priori* the interrelationship between measurable (or observed) physical features of fish scales such as number of growth checks, circuli counts and spacing between circuli and the "latent" life history variables to be estimated, namely, pre-maturation growth rate, actual age and age at maturation. The estimated life history variables were subsequently used for genetic analysis in fish farms where such fundamental data were not available.

Another method for use when certain variables cannot be measured is tetrachoric correlation analysis (Kendall and Stuart 1974) for estimation of the "repeatability" of relative size of individuals reared communally when size records are obtained in the form of dichotomies, i.e., whether individual sizes are greater or less than population means at each census and obtained in situations where individual marking is not possible (Eknath and Doyle 1990).

The above-mentioned methods are not exhaustive and their classifications not always clear-cut. Here, the focus is on applicability of the methods to aquaculture research, which does not preclude that in the future, other methods will successfully be applied to this field (see below). Details of the individual methods mentioned here are given in the respective contributions.

The Relationships between Multiple Regression and Analysis of Variance

This contribution, and by extension, the volume of which this is the "lead article", reiterate that multivariate methods exist that are suitable for analysis of data obtained from aquaculture systems. Reiterating this appears necessary in view of the continued tendency, by "field-orientated" aquaculturists, to underutilize the data they obtain from the costly experiments they conduct, by applying methods that are not powerful enough, or which cannot account for more than one or two effects at a time.

Particularly, we would like to convince the readers that multiple regression (MR) and its derivatives (e.g., path analysis) and variants (e.g., canonical analysis) should be methods of choice for aquaculturists, rather than analysis of variance (AV), presently taught as the basic tool for analysis of aquaculture experiments and often used alone.

Some statisticians see MR and AV as equivalent "general linear models", and hence, we may seem as quibbling about the details of a long resolved issue. This would be wrong, because MR and AV, while sharing a large part of their assumption and mathematics, are, in fact, taught and therefore used differently, with a strong tendency for the users of MR to quickly perform the transformations, contrast coding and other robust procedures that lead to *predictions* while AV users seem forever to argue about whether the rigid assumptions of AV are met, even when it is used only for *description*. Thus, for example Underwood (1981), in a much-cited review paper, estimated that only 12% of 151 published papers he examined had used AV "correctly". He thus wrote that "an inescapable conclusion must be that the techniques are insufficiently understood by many workers, despite the many excellent and eminently readable texts which discuss analysis of variance."

However, given the otherwise high standard of the authors of the 151 papers he examined, Underwood (1981) could also have concluded - equally "inescapably" - that something must be wrong with a technique which its practitioners cannot seem to ever use correctly.

This conclusion is supported by Cohen (1968), who recalled the trade-off behind AV: hard-to-meet assumptions which markedly facilitate computations.

With the advent of personal computers, this trade-off is no longer necessary; Cohen (1968) wrote on this:

"The MR procedure, in general, requires the computation and inversion of the matrix of correlations (or sums of squares and products) among the independent variables, a considerable amount of computation for even a modest number of independent variables. It is true that *classical* AV, whose main effects, interactions, polynomial trend components, etc., are mutually orthogonal, capitalizes on this orthogonality to substantially reduce the computation required. Whatever computational reduction there is in AV or MR depends directly on the orthogonality of the independent variables, which we have seen is restricted to manipulative experiments, and is by no means an invariate feature of such experiments.

However, given the widespread availability of electronic computer facilities, the issue of the *amount* of computation required in the analysis of data [...] dwindles to the vanishing point [...]. The typical [users] of a typical computer facility require that a computer program which will analyze their data be available in the program library. Such programs will have been either prepared or adapted for the particular computer configuration of that facility. Unfortunately, it is frequently the case that the available AV program or programs will not analyze the particular fixed AV design which investigators bring. Some AV programs are wanting in capacity in number of factors or levels per factor, some will handle only orthogonal designs, some will handle only equal cases per cell, some will do AV but not ACV, some of those that do handle ACV can handle only one or two covariates. Many will not handle special forms of AV, for example, Latin squares.

On the other hand, even the most poorly programmed scientific computer facility will have at least one good MR program, if for no other reason than its wide use in various technologies, particularly engineering. All the standard statistical program packages contain at least one MR program. [...] In contrast to the constraints of AV programs, the very general MR program can be particularized for any given design by representing (coding) those aspects of the independent variables of interest to the investigator according to the principles which have been described."

We leave it to the reader to consult Cohen (1968) for further aspects of the interrelationships between MR and AV.

A Note on Data Manipulations

Cohen (1968), in his discussion of the versatility of MR, emphasized dummy variables, contrast coding and linearizing transformations, i.e., techniques of which some forms have been used throughout this book.

We have given little emphasis, however, to some recent (and some not so recent) elaborations of these techniques, of which some would probably have improved our analyses. We mention these techniques here, in the hope that readers may be inspired to apply them to our data (see below on their availability). These techniques are:

(i) least median squares (LMS) methods of Rousseeuw and Leroy (1987), along with their PROGRESS software; and
(ii) the ACE algorithm of Breiman and Friedman (1985).

The approach in (i) enables MR analysis to be unaffected by outliers, as the technique only uses the "best" 50% of available observations. In the case of "messy" data such as those frequently obtained from aquaculture experiments, this procedure has considerable advantages, especially when the datasets are large, and lots of time must be devoted to identification of outliers. This technique was used in only one of the contributions in this book (Prein, this vol.); we recommend its general use.

The approach in (ii), a very recent development in statistics, enables the user to identify the linearizing transformation most appropriate for a given dataset.

Here, the result is not a set of parameters defining a given MR, but a set of scattergrams (one per variable, see Fig. 2) describing the "optimal transformations", i.e., those ensuring linearity of the variables included in the method.

This approach was only recently introduced from pure statistics into fisheries research (Mendelssohn and Mendo 1987; Cury and Roy 1989) and previous applications to aquaculture have not come to our attention. We recommend its use for datasets such as documented in Appendix II.

Colleagues planning to reanalyze the data in Appendix II and/or analyze similar data might wish to consider two other techniques, known since two decades, but which we didn't use in the MR analyses presented here:

(iii) using a dummy variable to code for "filled in" values of independent variables;
(iv) using geometric mean or "functional" MR instead of the predictive model that is incorporated in most software packages.

Filling in values of missing data points is necessary to prevent substantial data losses in cases where one value of the dependent variable (which is never "filled in" when absent) is matched with a large number of values for the dependent variables, of which only one might be missing (Prein, this volume). The various procedures used to fill in data are outlined in the contributions where data were filled in. However, in none of these contributions were these observations coded as suggested by Cohen (1968), which would have enabled evaluation of the technique used for filling in the missing data (i.e., of the bias induced by the filling-in procedure itself). We recommend that this procedure be used in future analyses.

Item (iv) is related to the fact that the MR routine incorporated in software packages uses, in virtually all cases, plots of Y_i on X_i, on the assumption that only the Y_i are measured with error

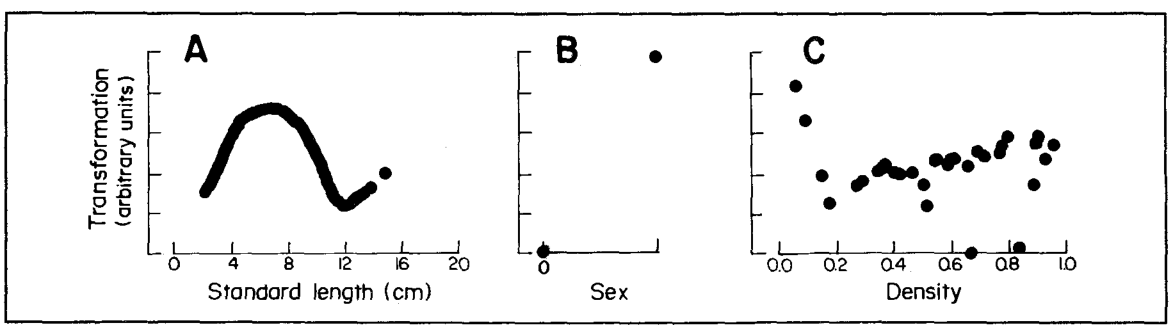


Fig. 2. Transformations that would be required for a multiple *linear* regression of growth rate (GR) *vs* length (A), sex (B) and density (C) of the Nile tilapia data as obtained by applying the ACE algorithm of Breiman and Friedman (1985) to a subset (N=500) of the Nile tilapia data in Mair and Pauly (this vol.); the transformation in A suggests that the extended Gulland and Holt plot did not fully linearize the relationship between GR and length; the two-point transformation in B is as expected from a dummy variable; the transformation in C suggests that density, when low, is negatively related to GR, but that its relationship to GR changes at higher densities. See text for implications.

(Ricker 1975). This is not a realistic assumption, and when the available data scatter widely around the trend lines, a marked bias occurs, i.e., the partial regression coefficients are underestimated.

This may perhaps be alleviated by using an MR extension of the approach advocated by Schnute (1984) (which we have not investigated, however), or by assuming similar errors in the Y_i *and* X_i, in which case geometric mean MR can be used (Pauly 1986; Prein and Pauly this vol.). We have not followed up on this issue, however, but feel that this would be worth pursuing, given the wide scatter of some of the datasets in Appendix II.

Rationale for Sharing Data

On Using 'Old' Data

As stated above, one of the incentives behind this book is to demonstrate the cost effectiveness, in terms of research output of using data that are already existing, collected in the context of earlier experiments (Pauly 1988), instead of conducting new, costly and time-consuming experiments. In most aquaculture research institutes, large amounts of experimental data are routinely produced, and these accumulate over the years. The experiments are conducted for various purposes with different approaches, but mostly consist of factorial designs for testing the effects of different variables on fish growth. Mostly, these data are analyzed superficially, only to answer the questions which initially invoked the experiments. However, these data could serve for further analyses for which more information could be extracted and further questions answered, particularly if several datasets are combined, e.g., from different locations, but with the same species or strains (Prein and Milstein 1988), or in the context of meta-analyses (Rosenthal 1984; Wolf 1986; Mann 1990).

'Public Domain' Data

Given the above, it will generally be useful to reanalyze "old", existing data, particularly with the new methods mentioned above. However, this requires a common understanding among aquaculture scientists to share with colleagues the raw data they have produced, as has been the case, for decades, in oceanographic and meteorological research and as practiced by the USAID-funded Aquaculture Collaborative Research Support Program (CRSP), which has deposited hard copies of their standardized dataset in the ICLARM Library, and which also will make summaries of these same data available through FishBase, the computerized encyclopedia on fishes (Pauly and Froese 1991). In spite of problems associated with differences in experimental design, data measurement, documentation standards, recording formats etc., this approach has a high benefit/cost ratio compared to conducting new experiments. This approach may also help fellow scientists formulate more sophisticated hypotheses, since they can base their designs on information extracted from available datasets. Also, scientists without access to experimental pond facilities can cooperate with others that have accumulated data and thereby still contribute to aquaculture research. This approach may also help to reduce costly duplication of research. In line with this argumentation, all data (and original software) used in this "data-rich book" (Pauly 1993) are available at nominal cost to anyone who asks for them either in summary form, through FishBase, or in full (see Appendix II for details). We encourage interested colleagues to "squeeze" these data even more than we have been able to.

Acknowledgements

The authors thank Mr. T. Schurig (BMZ), and Drs. H.-J. de Haas (BMZ), M. Bilio (GTZ), C. Wiedey (ATSAF) and Mrs. M. Bar-Lev (CINADCO) for their support of this new approach through the funding of the GIARA project, as well as the publication of this book; further, we thank Dr. A. Eknath for his review of and suggestion for the improvement of a draft of this contribution; we also thank Dr. R.S.V. Pullin for his contribution to the design and for his support of the project which led to this book, for reviewing several contributions and for his kind foreword.

References

Allen, P.G., L.W. Botsford, A.M. Schuur and W.E. Johnston. 1984. Bioeconomics of aquaculture. Dev. Aquacult. Fish. Sci. 13, 351 p. Elsevier, Amsterdam. 351 p.

Armbrester, W. 1971. The growth of caged *Tilapia aurea* (Steindachner) in fertile farm ponds. Proc. Annu. Conf. Southeast. Assoc. Fish Wildl. Agencies 25:446-451.

Azizan, Z. 1983. Using simulation methods to determine the optimal harvesting period for the cultured Malaysian freshwater prawns, *Macrobrachium rosenbergii*. University of Illinois at Urbana-Champaign, Urbana, Illinois. 196 p. M.S. thesis.

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.

Balarin, J.D. 1988. Development planning for tilapia farming in Africa, p. 531-538. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. [redacted] 15, 623 p.

Balarin, J.D. and R.D. Haller. 1983. Commercial tank culture of tilapia, p. 473-483. *In* L. Fishelson and Z. Yaron (comps.) Proceedings of the International Symposium on Tilapia in Aquaculture, 8-13 May, Nazareth, Israel. Tel Aviv University, Tel Aviv. 624 p.

Balarin, J.D. and J.P. Hatton. 1979. Tilapia: a guide to their biology and culture in Africa. Institute of Aquaculture, University of Stirling, Stirling, Scotland. 174 p.

Balchen, J.G., Editor. 1987. Automation and data processing in aquaculture. Proceedings of the International Federation on Automatic Control Symposium, Trondheim, Norway, 18-21 August 1986, Pergamon Press, Oxford. 280 p.

Baozhen, W. 1987. The development of ecological wastewater treatment and utilization systems (EWTUS) in China. Water Sci. Technol. 19:51-63.

Bartone, C.B., M.L. Esparza, C. Mayo, O. Rojas and T. Vitko. 1985. Monitoring and maintenance of treated water quality in the San Juan lagoons supporting aquaculture (UNDP/WORLD BANK/GTZ Integrated Resource Recovery Project GLO/80/004): final report of phases I-II. Pan American Center for Sanitary Engineering (CEPIS/PAHO), Lima, Peru. 74 p.

Bernard, D.R. 1986. A survey of the mathematical models pertinent to fish production and tropical pond aquaculture, p. 231-243. *In* J.E. Lannan, R.O. Smitherman and G. Tchobanoglous (eds.) Principles and practices of pond aquaculture. Oregon State University Press, Corvallis, Oregon. 252 p.

Bhattarai, K.K., C. Polprasert and B.N. Lohani. 1986. Models for aquacultural treatment of septage. Water Sci. Technol. 18:103-112.

Botsford, L.W. 1977. Current economic status of lobster culture research. Proc. World Maricult. Soc. 8:723-739.

Botsford, L.W. and T.W. Gossard. 1978. Implications of growth and metabolic rates on costs of aquaculture. Proc. World Maricult. Soc. 9:413-423.

Botsford, L.W., H.E. Rauch and R. Shleser. 1974. Application of optimization theory to the economics of aquaculture. Proc. World Maricult. Soc. 5:387-401.

Botsford, L.W., A.M. Schuur, R. Shleser and H.E. Rauch. 1975. An economically optimum aquaculture facility. Proc. World Maricult. Soc. 6:407-420.

Bowen, S.H. 1981. Digestion and assimilation of periphytic detrital aggregate by *Tilapia mossambica*. Trans. Am. Fish. Soc. 110:239-245.

Bowen, S.H. 1982. Feeding, digestion and growth - qualitative considerations, p. 141-156. *In* [redacted]e- [redacted]he biol[redacted] 432 p.

Bo[redacted]nd E. Johnston. 1978. Predicting early morning dissolved oxygen concentrations in channel catfish ponds. Trans. Am. Fish. Soc. 107:484-492.

Breiman, L. and J.H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. J. Am. Stat. Assoc. 80:580-619.

Busch, C.D., C.A. Flood, J.L. Koon and R. Allison. 1977. Modelling catfish pond nighttime dissolved oxygen levels. Trans. Am. Soc. Agric. Eng. 20:394-396.

Cheong, L., F.K. Chan, F.J. Wong and R. Chou. 1987. Observations on the culture of red tilapia (*Oreochromis niloticus* hybrid) in seawater under intensive tank conditions using a biodrum. Singapore J. Primary Ind. 15:42-56.

Cohen, J. 1968. Multiple regression as a general data-analytic system. Psychol. Bull. 70(6):426-443.

Cointreau, S.J. 1987. Aquaculture with treated wastewater: a status report on studies conducted in Lima, Peru. Applied Resource and Technology (WUDAT) Technical Note No. 3, 57 p. Integrated Resource Recovery Project, Water Supply and Urban Development Department, The World Bank, Washington, DC.

Cuenco, M.L. 1989. Aquaculture systems modelling: an introduction with emphasis on warmwater aquaculture. ICLARM Stud. Rev. 19, 46 p.

Cuenco, M.L., R.R. Stickney and W.E. Grant. 1985a. Fish bioenergetics and growth in aquaculture ponds: I. Individual fish model development. Ecol. Modelling 27:169-190.

Cuenco, M.L., R.R. Stickney and W.E. Grant. 1985b. Fish bioenergetics and growth in aquaculture ponds: II. Effects of interactions among size, temperature, dissolved oxygen, unionized ammonia and food on growth of individual fish. Ecol. Modelling 27:191-206.

Cuenco, M.L., R.R. Stickney and W.E. Grant. 1985c. Fish bioenergetics and growth in aquaculture ponds: III. Effects of intraspecific competition, stocking rate, stocking size and feeding rate on fish productivity. Ecol. Modelling 28:73-95.

Cury, P. and C. Roy. 1989. Optimal environmental window and pelagic fish recruitment success in upwelling areas. Can. J. Fish. Aquat. Sci 46(4):670-680.

Doha, S. and A. Haque. 1966. Studies on the biology of tilapia (*Tilapia mossambica* Peters). I. The food and feeding habits. Pak. J. Sci. 18:205-215.

Dusart, J. 1963. Contribution à l'étude de l'adaptation des tilapia (Pisces, Cichlidae) à la vie en milieu mal oxygèné. Hydrobiologia 21:328-341.

Edwards, P. 1992. Reuse of human wastes in aquaculture - a technical review. Water and Sanitation Report 2, 350 p. UNDP-World Bank Water and Sanitation Program, The World Bank, Washington, DC.

Edwards, P. and R.S.V. Pullin, Editors. 1990. Wastewater-fed aquaculture. Proceedings of the International Seminar on Wastewater Reclamation and Reuse for Aquaculture, Calcutta, India, 6-9 December 1988. Environmental Sanitation Information Center, Asian Institute of Technology, Bangkok. 296 p.

Edwards, P., R.S.V. Pullin and J.A. Gartner. 1988. Research and education for the development of integrated crop-livestock-fish farming systems in the tropics. ICLARM Stud. Rev. 16, 53 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.

Eknath, A.E. and R.W. Doyle. 1990. Repeatability of relative size of individuals under communal stocking: Implications for size grading in aquaculture, p. 45-48. *In* R. Hirano and I. Hanyu (eds.) The Second Asian Fisheries Forum. Asian Fisheries Society, Manila.

Elliott, J.M. 1979. Energetics of freshwater teleosts, p. 29-61. *In* P.J. Miller (ed.) Fish phenology: anabolic adaptiveness in teleosts. Proceedings of a Symposium held at the Zoological Society of London, 6-7 April 1978. Symposia of the Zoological Society London, Academic Press, London. 449 p.

Fernandes, M.N. and F.T. Rantin. 1987. Respiratory responses of *Oreochromis niloticus* (Pisces, Cichlidae) to environmental reductions of dissolved oxygen. Bol. Fisiol. Anim. (Sao Paulo) 11:51-60.

Fishelson, L. and Z. Yaron, Compilers. 1983. Proceedings of the International Symposium on Tilapia in Aquaculture, 8-13 May 1983, Nazareth, Israel. Tel Aviv University, Tel Aviv. 624 p.

From, J. and G. Rasmussen. 1984. A growth model, gastric evacuation, and body composition in rainbow trout, *Salmo gairdneri* Richardson, 1836. Dana 3:61-139.

Fryer, G. and T.D. Iles. 1972. The cichlid fishes of the great lakes of Africa: their biology and evolution. T.F.H. Publications, Neptune City, New Jersey. 641 p.

Gaigher, I.G. and D.F. Toerien. 1985. Cage culture of Mozambique tilapia, *Oreochromis niloticus*, without artificial feeding in maturation ponds of the Phuthabitjhaba sewage system. Water SA 11:19-24.

Griffin, W.L., J.S. Hanson, R.W. Brick and M.A. Johns. 1981. Bioeconomic modelling with stochastic elements in shrimp culture. J. World Maricult. Soc. 12:94-103.

Griffin, W.L., R.G. Anderson, R.R. Stickney and R.E. Whitson. 1980. Bioeconomic assessment of a poultry sewage and tilapia aquaculture system. Texas J. Sci. 32:311-318.

Griffin, W.L., W.E. Grant, R.W. Brick and J.S. Hanson. 1984. A bioeconomic model of shrimp maricultural systems in the USA. Ecol. Modelling 25:47-68.

Gulland, J.A. and S.J. Holt. 1959. Estimation of growth parameters for data at unequal time intervals. J. CIEM 25:47-49.

Haf, C.M. and T. Cheaib. 1985. Multivariate Statistik in den Natur- und Verhaltenswissenschaften. Vieweg, Braunschweig. 155 p.

Hepher, B. 1988. Nutrition of pond fishes. Cambridge University Press, Cambridge. 388 p.

Hepher, B. and Y. Pruginin. 1981. Commercial fish farming, with special reference to fish culture in Israel. John Wiley, New York. 261 p.

Hey, D. 1955. A preliminary report on the culture of fish in the final effluent from the New Disposal Works, Athlone, S. Africa. Verh. Int. Ver. Theor. Angew. Limnol. 12:737-742.

Hofer, R. and F. Schiemer. 1983. Feeding ecology, assimilation efficiencies and energetics of two herbivorous fish: *Sarotherodon* (Tilapia) *mossambicus* (Peters) and *Puntius filamentosus* (Cuv. et Val.), p. 155-164. *In* F. Schiemer (ed.) Limnology of Parakrama Samudra, Sri Lanka: a case study of an ancient man-made lake in the tropics. W. Junk Publishers, The Hague. 236 p.

Hopkins, K., M. Hopkins, D. Leclercq and A.A. Al-Ameeri. 1986. Tilapia culture in Kuwait: a preliminary economic analysis of production systems. Kuwait Bull. Mar. Sci. 7:45-64.

Hopkins, K.D., D. Pauly, E.M. Cruz and J.H. van Weerd. 1982. An alternative to predator-prey ratios in predicting recruitment. Meeresforschung/Rep. Mar. Res. 29:125-135.

Houvenaghel, G.T. and T. Huet. 1987. Computer-aided production modelling in fish farming. An application to the eel, p. 105-108. *In* J.G. Balchen (ed.) Automation and data processing in aquaculture. Proceedings of the IFAC Symposium, 18-21 August 1986, Trondheim, Norway. Pergamon Press, Oxford. 280 p.

Huang, W.C., J.K. Wang and T. Fujimura. 1976. A model for estimating prawn populations in ponds. Aquaculture 8:57-70.

Huisman, E.A., Editor. 1986. Aquaculture research in the Africa region. Proceedings of the African Seminar on Aquaculture, 7-11 October 1985, Kisumu, Kenya. Pudoc, Wageningen. 274 p.

Iles, T.D. 1973. Dwarfing and stunting in the genus *Tilapia* (Cichlidae) as a possibly unique recruitment mechanism. Rapp. P.-V. Réun. CIEM 164:247-254.

Jana, B.B. and U.K. De. 1983. Primary productivity of phytoplankton, environmental covariates and fish growth in West Bengal ponds under a polyculture system. Int. Rev. Gesamt. Hydrobiol. 68:45-58.

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.

Jorgensen, S.E. 1976. A model of fish growth. Ecol. Modelling 2:303-313.

Kendall, M.G. and A. Stuart. 1974. Advanced theory of statistics. Vol. 2. Statistical influence and statistical relationship. Rev. 3rd. ed. Hafner, New York.

Kutty, M.N. and N. Sukumaran. 1975. Influence of upper and lower temperature extremes on the swimming performance of *Tilapia mossambica*. Trans. Am. Fish. Soc. 104:755-761.

Laws, E.A., D.R. Jones, K.L. Terry and J.A. Hirata. 1985. Modifications in recent models of phytoplankton growth: theoretical developments and experimental examination of predictions. J. Theor. Biol. 114:323-341.

Le Roux, P.J. 1956. Feeding habits of the young of four species of tilapia. S. Afr. J. Sci. 53:33-37.

Leung, P.S., L.A. Fallon and Y.C. Shang. 1984. Bioeconomic modelling of freshwater prawn production: the Hawaiian experience. ISEM J. (International Society for Ecological Modelling, Copenhagen) 6:25-37.

Little, D. and J. Muir. 1987. A guide to integrated warmwater aquaculture. Institute of Aquaculture, University of Stirling, Stirling, Scotland. 238 p.

Lowe-McConnell, R.H. 1958. Observations on the biology of *Tilapia nilotica* Linne in East African waters (Pisces: Cichlidae). Rev. Zool. Bot. Afr. 57:129-170.

Machiels, M.A.M. 1987. A dynamic simulation model for growth of the African catfish, *Clarias gariepinus* (Burchell 1822). IV. The effect of feed formulation on growth and feed utilization. Aquaculture 64:305-323.

Machiels, M.A.M. and A.A. van Dam. 1987. A dynamic simulation model for growth of the African catfish, *Clarias gariepinus* (Burchell 1822). III. Effect of body composition on growth and feed intake. Aquaculture 60:55-71.

Machiels, M.A.M. and A.M. Henken. 1986. A dynamic simulation model for growth of the African catfish, *Clarias gariepinus* (Burchell 1822). I. Effect of feeding level on growth and energy metabolism. Aquaculture 56:29-52.

Machiels, M.A.M. and A.M. Henken. 1987. A dynamic simulation model for growth of the African catfish, *Clarias gariepinus* (Burchell 1822). II. Effect of feed composition on growth and energy metabolism. Aquaculture 60:33-53.

Mann, C. 1990. Meta-analysis in the breech. Science 249:476-480.

Marinell, G. 1977. Multivariate Verfahren. R. Oldenbourg, Munich. 155 p.

McDonald, M.E. 1984. Trophic interactions and carbon budgets for the sub-reproductive adult phytoplanktivorous fish, *Tilapia aurea*, fed three unialgal cultures; with models for an algae-phytoplanktivorous fish wastewater nutrient removal system. North Carolina State University, Raleigh. 92 p. Ph.D. thesis.

Melard, C. 1987. Recherches sur la biologie *d'Oreochromis* (Tilapia) *niloticus* L. (Pisces Cichlidae) en elevage experimental: reproduction, croissance, bioenergetique. Cah. Ethol. Appl. 6(3):1-224.

Mendelssohn, R. and J. Mendo. 1987. Exploratory analysis of anchoveta recruitment off Peru and related environmental series, p. 294-306. *In* D. Pauly and I. Tsukayama (eds.) The Peruvian anchoveta and its upwelling ecosystem: three decades of change. ICLARM Stud. Rev. 15, 351 p.

Meske, C. 1980. Erfahrungsbericht über die Aufzucht von tropischen Warmwasserfischen in Ahrensburg. Arb. Dtsch. Fisch.-Verb. 30:109-122.

Meyer, D.I. and D.E. Brune. 1982. Computer modelling of the diurnal oxygen levels in a stillwater aquaculture pond. Aquacult. Eng. 1:245-262.

Meyer, D.I., D.E. Brune and B. Haight. 1983. Modelling the diel oxygen response of a small fish pond. J. World Maricult. Soc. 14:336-343.

Miller, G.E. 1985. An evaluation of a confinement catfish production facility incorporating recirculating culture systems with fish barn: a simulation for aquaculture analysis. Purdue University. 226 p. Ph.D. thesis.

Milstein, A., G. Hulata and G.W. Wohlfarth. 1988. Canonical correlation analysis of relationships between management inputs and fish growth and yields in polyculture. Aquacult. Fish. Manage. 19:13-24.

Milstein, A., M. Feldlite, N. Mozes and Y. Avnimelech. 1989. Limnology of reservoirs used for fish farming and crop irrigation, with integrated free and cage fish culture. Isr. J. Aquacult./Bamidgeh 41:12-22.

Mironova, N.V. and T.A. Skvortsova. 1967. [Growth of *Tilapia mossambica* Peters and its relationship to feeding.] Problemy Kosmicheskoy Biologii (Problems of Space Biology) 7:498-505. (In Russian).

Moriarty, D.J.W. 1973. The physiology of digestion of blue-green algae in the cichlid fish *Tilapia nilotica*. J. Zool. (Lond.) 171:25-39.

Moriarty, D.J.W. and R.S.V. Pullin, Editors. 1987. Detritus and microbial ecology in aquaculture. ICLARM Conf. Proc. 14, 420 p.

Morita, S.K. 1977. An econometric model of prawn pond production. Proc. World Maricult. Soc. 8:741-746.

Morrison, D.F. 1976. Multivariate statistical methods. McGraw-Hill, New York. 415 p.

Mosteller, F. and J.W. Tukey. 1977. Data analysis and regression. Addison-Wesley, Reading, Massachusetts. 588 p.

Ojda, L. 1987. Verhalten von *Oreochromis niloticus* (Teleostei: Cichlidae) bei Hypoxie. Hamburg University. 81 p. M.S. thesis (Diplomarbeit).

Patten, B.C. 1968. Mathematical models of plankton production. Int. Rev. Gesamt. Hydrobiol. 53:357-408.

Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Stud. Rev. 8, 325 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.

Pauly, D. 1988. On using other people's data. Naga, ICLARM Q. 11(1):6-7.

Pauly, D. 1993. Data-rich books. BioScience 43(3):167-168.

Pauly, D. and R. Froese. 1991. FISHBASE: assembling information on fish. Naga, ICLARM Q. 14(4):10-11.

Pauly, D. and J. Ingles. 1981. Aspects of the growth and mortality of exploited coral reef fishes, p. 89-98. *In* E. Gomez, C.E. Birkeland, R.W. Buddemeyer, R.E. Johannes, J.A. Marsh, Jr. and R.T. Tsuda (eds.) Proceedings of the Fourth International Coral Reef Symposium, Manila. Vol. 1. Marine Science Center, University of the Philippines, Quezon City. 725 p.

Pauly, D. and K.D. Hopkins. 1983. A method for the analysis of pond growth experiments. ICLARM Newsl. 6(1):10-12.

Pauly, D., J. Moreau and M.L.D. Palomares. 1988. Detritus and energy consumption and conversion efficiency of *Sarotherodon melanotheron* (Cichlidae) in a West African lagoon. J. Appl. Ichthyol. 4:190-193.

Peer, M.M. and M.N. Kutty. 1981. Respiratory quotient and ammonia quotient in *Tilapia mossambica* (Peters) with special reference to hypoxia and recovery. Hydrobiologia 76:3-9.

Piedrahita, R.H. 1984. Development of a computer model for the aquaculture pond ecosystem. University of California, Davis, California. 162 p. Ph.D. thesis.

Piedrahita, R.H. 1988. Introduction to computer modelling of aquaculture pond ecosystems. Aquacult. Fish. Manage. 19:1-12.

Piedrahita, R.H., D.E. Brune, G. Tchobanoglous and G.T. Orlob. 1984. A general model of the aquaculture pond ecosystem. J. World Maricult. Soc. 15:355-366.

Polovina, J. and H. Brown. 1978. A population dynamics model for prawn aquaculture. Proc. World Maricult. Soc. 9:393-404.

Prein, M. 1985. The influence of environmental factors on fish production in tropical ponds investigated with multiple regression and path analysis. Mathematisch-Naturwissenschaftliche Fak., Inst. für Meereskunde an der Christian-Albrechts-Univ. Kiel, Part I. - 92 p.; Part II (Data volume). 248 p. Master's thesis.

Prein, M. 1987. Multivariate models of tilapia and carp growth as an application of microcomputers in aquaculture research. Naga, ICLARM Q. 10(2):7-8.

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

Prein, M. and A. Milstein. 1988. Techniques for handling large pond and farm datasets. Aquabyte 1(2):4-5.

Prikryl, I. and T. Kepr. 1984. Statistical analysis of data on tench culture. Bull. VURH Vodnany 20:17-21.

Provenzano, A.J. and J.G. Winfield. 1987. Performance of a recirculating fish production system stocked with tilapia hybrids. Aquacult. Eng. 6:15-26.

Pullin, R.S.V. 1985. Tilapias: 'everyman's fish'. Biologist 32:84-88.

Pullin, R.S.V., Editor. 1988. Tilapia genetic resources for aquaculture. ICLARM Conf. Proc. 16, 108 p.

Pullin, R.S.V. 1989. Third-world aquaculture and the environment. Naga, ICLARM Q. 12(1):10-13.

Pullin, R.S.V. 1991. Cichlids in aquaculture, p. 280-309. *In* M.H.A. Keenleyside (ed.) Cichlid fishes - behaviour, ecology and evolution. Fish and Fisheries Series 2, Chapman and Hall, London. 378 p.

Pullin, R.S.V. and R.H. Lowe-McConnell, Editors. 1982. The biology and culture of tilapias. ICLARM Conf. Proc. 7, 432 p.

Pullin, R.S.V. and Z.H. Shehadeh. 1980. Integrated agriculture-aquaculture farming systems. ICLARM Conf. Proc. 4, 258 p.

Pullin, R.S.V., T. Bhukaswan, K. Tonguthai and J.L. Maclean, Editors. 1988. The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Pullin, R.S.V., A.E. Eknath, T. Gjedrem, M.M. Tayamen, J. Macaranas and T.A. Abella. 1991. The Genetic Improvement of Farmed Tilapias (GIFT) project: the story so far. Naga, ICLARM Q. 14(2):3-6.

Rauch, H.E., L.W. Botsford and R.A. Shleser. 1975. Economic optimization of an aquaculture facility. IEEE Trans. Automatic Control (Institute of Electrical and Electronics Engineers) 20:310-319.

Rennert, B. and W. Steffens. 1985. Erstmalige Aufzucht des tropischen Buntbarsches *Oreochromis mossambicus* in der DDR in einem geschlossenen Kreislauf. Fortschr. Fischereiwiss. 4:45-52.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191, 382 p.

Romaire, R.P. 1979. Modelling the dynamics of dissolved oxygen in channel catfish production ponds. Auburn University, Auburn. 93 p. Ph.D. thesis.

Romaire, R.P. and C.E. Boyd. 1979. Effects of solar radiation on the dynamics of dissolved oxygen in channel catfish ponds. Trans. Am. Fish. Soc. 108:473-478.

Rosenthal, R. 1984. Meta-analytic procedures for social research. Applied social research methods series, Vol. 6. Sage Publications Newberry Park, California, 148 p.

Ross, L.G., R.W. McKinney and B. Ross. 1988. Energy budgets for cultured tilapias, p. 83-89. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Rousseeuw, P.J. and A.M. Leroy. 1987. Robust regression and outlier detection. John Wiley, New York. 329 p.

Saha, S.N. and S. Dewan. 1979. Food and feeding habits of *Tilapia nilotica* (Linnaeus) (Perciformes: Cichlidae). I. Types and amount of food taken by the fish and its size and patterns of feeding. Bangladesh J. Zool. 7(1):53-60.

Salvadores, M.L. and M. Guzman. 1983. Contenido estomacal de la tilapia *Sarotherodon aureus* Steindachner, en la presa "Vicente Guerrero", Guerrero, Edo. de Mexico, Mexico. Biotica (Instituto Nacional de Investigaciones sobre recursos bioticos, Xalapa, Mexico) 8(1):59-70.

Sanico, A.F. 1975. Effects of 17-ethynyltestosterone and estrone on sex ratio and growth of *Tilapia aurea* (Steindachner). Philipp. J. Sci. 13:1-15.

Schnute, J. 1984. Linear mixtures: a new approach to bivariate trend lines. J. Amer. Stat. Assn. 79(385):1-8.

Schroeder, G.L. 1986. Integrated fish farming: an international effort. Am. J. Alternative Agric. 1:127-130.

Smith, D.W. and R.H. Piedrahita. 1988. The relation between phytoplankton and dissolved oxygen in fish ponds. Aquaculture 68:249-265.

Sparre, P. 1976. A Markovian decision process applied to optimization of production planning in fish farming. Medd. Danm. Fisk. Havunders. N.S. 7:111-197.

Sperber, O., J. From and P. Sparre. 1977. A method to estimate the growth rate of fishes, as a function of temperature and feeding level, applied to rainbow trout. Medd. Danm. Fisk. Havunders. N.S. 7:275-317.

Spotts, D. 1983. High density tilapia culture, Taiwan. Aquacult. Mag. 9(6):21-23.

Svirezhev, Y.M., V.P. Krysanova and A.A. Voinov. 1984. Mathematical modelling of a fish pond ecosystem. Ecol. Modelling 21:315-337.

Trewavas, E. 1983. Tilapiine fishes of the genera *Sarotherodon, Oreochromis* and *Danakilia*. British Museum (Natural History), London. 583 p.

Tsadik, G.G. and M.N. Kutty. 1987. Influence of ambient oxygen on feeding and growth of the tilapia, *Oreochromis niloticus* (Linnaeus). Working Paper No 10, 13 p. African Regional Aquaculture Centre, Port Harcourt, Nigeria.

Tukey, J.W. 1977. Exploratory data analysis. Addison-Wesley, Reading, Massachusetts. 688 p.

Uchida, R.N. and J.E. King. 1962. Tank culture of tilapia. Fish. Bull. US Fish Wildl. Serv. 199:1-52.

Underwood, A. 1981. Techniques of analysis of variance in experimental marine biology: a review. Oceanogr. Mar. Biol. Ann. Rev. 19:513-605.

Unger, G.E. 1983. A simulation model of an integrated pond mariculture system. Hebrew University of Jerusalem, Jerusalem, Israel. 87 p. M.S. thesis.

Vaas, K.F. and A.E. Hofstede. 1952. Studies on *Tilapia mossambica* Peters (ikan mudjair) in Indonesia. Pemberitaan Balai Penjelidikan Perikanan Darat, Djakarta-Bogor, No. 1. 68 p.

Van Dam, A.A. 1989. An example of the potential application of modelling in aquaculture: production of the Nile tilapia *Oreochromis niloticus* in rice fields, Appendix. p. 36-46. *In* M.L. Cuenco. Aquaculture systems modelling: an introduction with emphasis on warmwater aquaculture. ICLARM Stud. Rev. 19, 46 p.

Verreth, J.A.J., M. Carillo, S. Zanuy and E.A. Huisman, Editors. 1987. Aquaculture research in Latin America. Proceedings of the Latin American Seminar on Aquaculture held in Lima, Peru. Pudoc, Wageningen. 451 p.

Wohlfarth, G.W. and G. Hulata. 1983. Applied genetics of tilapias. Second, revised edition. ICLARM Stud. Rev. 6, 26 p.

Wolf, F.M. 1986. Meta-analysis, quantitative methods for research synthesis. Quantitative applications in the social sciences, 59. Sage Publications Beverly Hills, California. (En).

Zentrum für Angepasste Technologie und Sozialökologie Langenbrück 1983. Fischzucht in bepflanzten Solartonnen. AT Verlag, Stuttgart, Federal Republic of Germany. 70 p.

Multiple Regression Analysis of Aquaculture Experiments Based on the "Extended Gulland-and-Holt Plot": Model Derivation, Data Requirements and Recommended Procedures*

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

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

KEVIN D. HOPKINS, *University of Hawaii at Hilo, College of Agriculture, 523 W. Lanikaula St., Hilo, Hawaii 96720-4091, USA*

Pauly, D., M. Prein and K.D. Hopkins. 1993. Multiple regression analysis of aquaculture experiments based on the "extended Gulland-and-Holt Plot": model derivation. data requirements and recommended procedures p. 13-23. *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

A method for the multivariate analysis of fish growth in aquaculture is presented. It is derived from a linearized version of the von Bertalanffy growth function (VBGF), which, in its original form, is a bivariate regression termed the Gulland-and-Holt plot. Here, a version in form of a multiple regression equation is presented. The "extended Gulland-and-Holt plot" permits to identify and quantify the key variables controlling fish growth and permits the inclusion of these environmental and treatment variables to explain variance in growth of fish. Von Bertalanffy growth parameters K and L_∞ are obtained, which contain the combined environmental effects on fish growth and reflect the range of culture conditions. By computing the index of growth performance (ϕ'), the obtained regression models can be used for growth prediction and decisionmaking in fish farm management and production under a wide range of environmental and management conditions. Recommendations for the design of experiments, preparation of data for analysis and actual analysis procedures are given.

Introduction

For aquaculturists, the processes of main interest are fish growth and production (Bardach et al. 1972; Steffens 1981; Hepher and Pruginin 1981). When conducting experiments, the normal procedure is to compare the length of time needed to reach "market" sizes (i.e., length or weight), growth rates (e.g., $g \cdot day^{-1}$) or "production" (actually: standing stock in, e.g., $kg \cdot ha^{-1}$) at the end of a growth period. This requires that all variables are kept constant except for the treatments under investigation. Influences on the growth process itself during experiments are neglected. Several problems arise here:

1. Only one treatment at a time can be directly compared. Possible associated effects must be neglected. In ordinary growth experiments the final weights are compared using ANOVA or t-tests. This requires complete control of all variables and equal fish sizes at the beginning. The information contained in the growth curve itself is lost since only the single last value is used for interpretation.

   For example, in single-factor experiments with a simple ANOVA table for analysis, only one variable can be compared (e.g., final fish sizes or average growth rates) and all that can be inferred is whether a positive or negative effect was observed. The disadvantage is that this method does not permit to quantify how much better one type of treatment was compared to the other. Further, associated effects among the variables are neglected, so that they remain undiscovered.
2. Uncontrollable and varying effects such as those caused by meteorological factors (seasonality of temperature and rain, environment, etc.) and different experiment locations are difficult to account for in the analysis.
3. The growth of fish is not linear. Thus, it is important to have (a) identical fish sizes at the beginning of the experiment, and (b) equal duration of experiments.

For the aquaculture scientist it is important to know the growth performance of certain species for the planning of aquaculture projects, the

*ICLARM Contribution No. 720.

development of industries at new locations, and the correct management of farms during different seasons or conditions (Reay 1979).

When conducting large experiments with many treatments (i.e., multifactor experiments), the problems mentioned above are worsened (Gomez and Gomez 1984). The amount of data increases exponentially with every further variable measured (Hopkins et al. 1988). Data handling and management become very complicated. With the ordinary methods of data analysis, the interpretation of results and the testing of hypotheses is very difficult. Today, the availability of computers with data management and statistical analysis software can help to solve many of these problems (Vakily 1989) even with sophisticated methods of data analysis. An approach along this line is a method developed by Pauly and Hopkins (1983, see Appendix I, this vol.) which will be explained in detail here. This may be regarded as representing only the beginning of this type of analysis. In the future, more methods of statistical analysis and modeling will be developed since aquaculture is undermathematized compared to agriculture or fisheries.

The Process of Growth

Curves describing growth in length or in weight both approach an asymptotic value towards the end of an animals' life span. Length growth can usually be modelled using an asymptotic curve which tapers off with increasing age. Weight growth is roughly sigmoid, i.e., the weight increment increases gradually up to an inflection point from where it then gradually decreases again. The growth rates change constantly, which imposes problems when using these parameters for interpretation of experiments (Fig. 1). For different fish sizes the absolute growth increments will be different (Fig. 1).

A method for the mathematical description of fish growth is the von Bertalanffy growth function

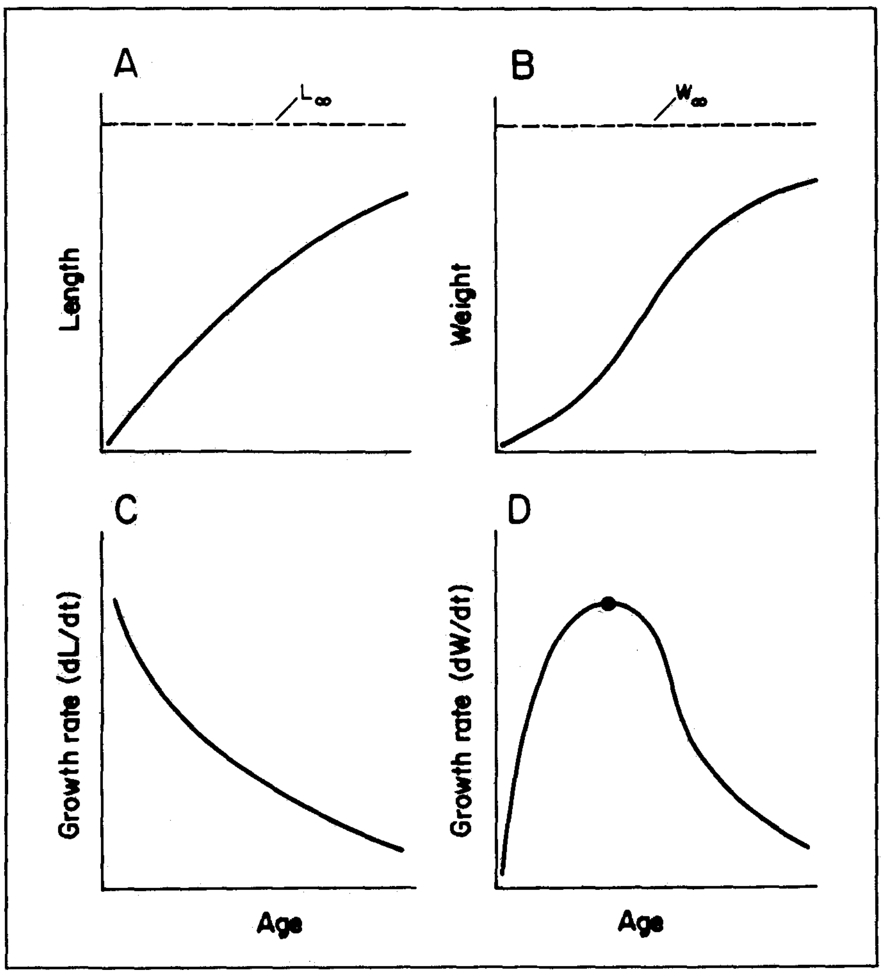


Fig. 1. Schematic example of growth curves. A) for length, and B) for weight, which has an inflection point. Both curves approach asymptotic size. C) Schematic representation of the change of absolute growth rate with time in length and D) in weight. Note the position of the inflection point in weight growth rate. The growth rate in weight of fish at different ages (sizes) cannot be compared.

(VBGF) (Pütter 1920; von Bertalanffy 1934, 1938), which is widely used in fisheries research. This function has several advantages as will be shown later. The VBGF, as modified from von Bertalanffy by Beverton and Holt (1957) is:

$$L_t = L_\infty \cdot \left(1 - e^{-K(t-t_o)}\right) \qquad \text{...1)}$$

for length, and

$$W_t = W_\infty \cdot \left(1 - e^{-K(t-t_o)}\right)^m \qquad \text{...2)}$$

for weight, where

- K = growth constant
- L_∞ = asymptotic length, i.e., the (mean) length the fish would reach if they were to grow indefinitely
- W_∞ = asymptotic weight, i.e., the weight corresponding to L_∞
- t = age of the fish in days/months/years
- t_o = theoretical (generally negative) "age" of the fish at zero size
- m = exponent of length-weight relationship

For the case of isometric growth, i.e., when m = 3, the VBGF takes the form:

$$W_t = W_\infty \cdot \left(1 - e^{-K(t-t_o)}\right)^3 \qquad \text{...3)}$$

With these parameters the growth of a single fish or the mean of a whole fish population can be described. For reasons to be shown later we will concentrate on the length-based version of the VBGF.

Estimation of VBGF Parameters

GULLAND-AND-HOLT PLOT

From aquaculture experiments the VBGF growth parameters L_∞ and K can be estimated, among several other methods, by a method termed the 'Gulland-and-Holt plot' (Gulland and Holt 1959), a method already described in von Bertalanffy's original derivation of his growth equation (von Bertalanffy 1934). This provides an approximative method to estimate the VBGF parameters by a simple linear regression technique. Here growth rates in length are regressed upon their corresponding average lengths during the intervals. The intervals need not be of equal duration. Also, different fish sizes and growth rates can be used.

This method is used in fisheries research for the analysis of tagging data and of length-frequency data (Gulland 1967, 1983; Pauly and Ingles 1981; Pauly 1984; Sparre et al. 1989). In aquaculture experiments we are in a situation similar to a tag-recapture experiment. Initially the ponds are stocked. During the experiment the fish are sampled periodically and finally the population is harvested. The Gulland-and-Holt method is based entirely on length measurements. Unfortunately this is in contrast to the procedure in most aquaculture experiments where only weights are measured.

The length growth curve is characterized by a gradual increase in length over time and a gradual decrease in growth rate (indicated by the slope of the tangent) over time. As length (L) approaches its maximum (i.e., L_∞) the growth rate becomes zero. This relationship between length and growth rate is linear and thus can be used to estimate the two parameters L_∞ and K. Growth rate ($\Delta L/\Delta t$) of an experiment interval is plotted over the mean length in that interval. It should be noted that this is a linear description of a nonlinear growth process (Fig. 2). The differential form is:

$$dl/dt = K(L_\infty - \bar{L}) \qquad \text{...4)}$$

or, in terms of growth increments per interval (length L_1 and L_2):

$$(L_2 - L_1)/(t_2 - t_1) = a + b\,(L_1 + L_2)/2 \qquad \text{...5)}$$

It should be noted that (4) is a differential equation and (5) is a difference equation, yet both

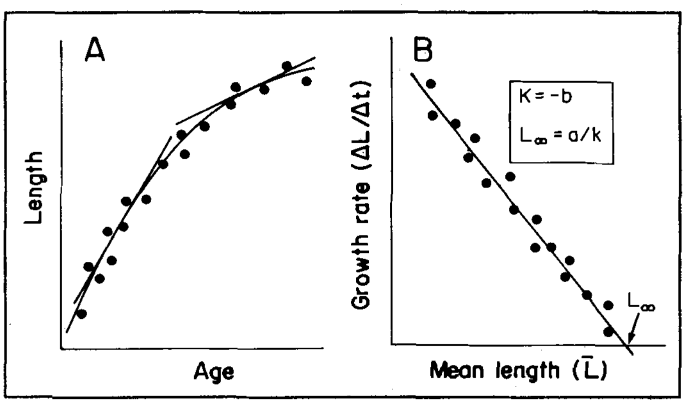


Fig. 2. A) Example of a length growth curve through data points. Tangents indicate the decrease in slope with age (i.e., size). B) Representation of a Gulland-and-Holt plot, describing the decrease in growth rate with increasing size. The VBGF parameters K and L_∞ are derived from the regression coefficients.

give similar results for small intervals. The growth parameters are obtained according to:

$$K = -b \quad \text{...6)}$$

and

$$L_\infty = a/K \quad \text{...7)}$$

Gulland and Holt (1959) pointed out that the discrepancy between difference and differential equations comes into effect at larger time intervals (i.e., several months to years). Therefore they suggested a correction factor with which the y-values are to be multiplied before computing the regression. For short intervals of 14 days this factor is very close to unity and can consequently be neglected (e.g., for $K = 0.01482\ day^{-1}$ and $\Delta t = 14$ days, the correction factor is 1.0033).

CONFIDENCE LIMITS FOR K

Since K = -b, the confidence limits (CL) for K, derived with the Gulland-and-Holt plot, are the same as those for 'b', only with the sign changed (Sparre et al. 1989):

$$CL_{(K)} : [K - s_{(K)} \cdot t(n\text{-}2),\ K + s_{(K)} \cdot t(n\text{-}2)] \quad \text{...8)}$$

where

$CL_{(K)}$	=	confidence limits of K
$s_{(K)}$	=	standard deviation of K (here with n-2 df)
$t(n\text{-}2)$	=	fractiles of Student's t-distribution, here based on n-2 degrees of freedom

CONFIDENCE LIMITS FOR L_∞

According to Sparre et al. (1989) the confidence limits for L_∞ derived with the Gulland-and-Holt plot cannot be straightforwardly derived. Rather they are "conditioned" on a previously determined value for K together with the confidence limits for 'a' in form of a ratio. The confidence limits for 'a' are obtained using:

$$CL_{(a)} : [a - s_{(a)} \cdot t(n\text{-}2),\ a + s_{(a)} \cdot t(n\text{-}2)] \quad \text{...9)}$$

where $s_{(a)}$ = standard deviation of 'a' (here with n-2 df).

Since $L_\infty = -a/b$, the procedure suggested by Sparre et al. (1989) is to divide the lower and upper confidence intervals of 'a' by the value of K:

$$CL_{(a)} : [(a - s_{(a)} \cdot t(n\text{-}2))/K,\ (a + s_{(a)} \cdot t(n\text{-}2))/K] \quad \text{...10}$$

where $CL_{(L\infty)}$ = confidence limits for L_∞.

Sparre et al. (1989) point out that, strictly, the confidence limits of a ratio are not defined. An alternative method is to regard L_∞ as a predicted value of y (here zero), estimated from a given x-value (here the x-intercept). According to Snedecor and Cochran (1982), it is possible to estimate the confidence interval for L_∞ (in form of the x-axis intercept), as also suggested for the "extended Bayley plot" in Prein and Pauly (this vol.).

CONSIDERATIONS FOR APPLICATION

For fish size data based on weight, the weight values must be converted to lengths, using a length-weight relationship. Therefore, a length-weight relationship should always be determined for the fish stock used in the experiments, or a set of length-weight parameters should be used, estimated from a stock of fish kept under similar conditions.

When dealing with several experiments, a separate growth curve could be derived for each experiment, producing a range of combinations of VBGF parameters (Fig. 3). The VBGF parameters could then be related to the different treatments.

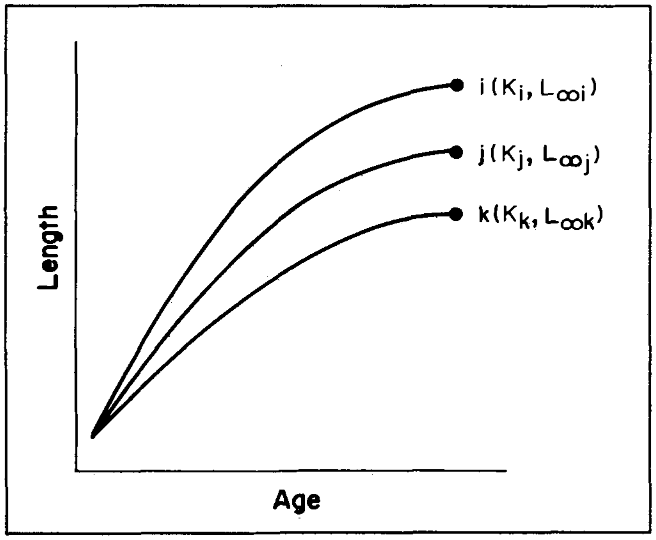


Fig. 3. Schematic example of fish growth (in length) under three different treatments (I, II and III). Usually, only sizes at the end of experiments, or overall yields, are compared (given equal sizes at stocking). In this way, any variations in growth occurring during the culture period are neglected. One possibility is to estimate a set of VBGF parameters for each pond or treatment.

The parameter combinations thus obtained could then be used for the prediction of growth under the assumption that all environmental and treatment factors are constant during the prediction range.

From the obtained pairs of VBGF growth parameters K and L_∞ the index ϕ' (Pauly 1979; Pauly and Munro 1984) can be computed to compare the growth performance of different fish stocks. It has been shown that the index of growth performance ϕ' can be used to compare the growth of tilapias under a wide range of culture conditions, within and between species and between sexes (Pauly et al. 1988). The growth performance index is computed according to:

$$\phi' = \log_{10}K + 2\log_{10}L_\infty$$

with K put on annual basis, and L_∞ referring to total length, in cm.

Multivariate Extension of the Gulland-and-Holt Plot

MULTIVARIATE APPROACH

Plotting the data of a single experiment in form of a Gulland-and-Holt plot, we will have a single regression line with some variance around the line. If, for example, from factorial experiments we have many different treatments and plot these all in the same Gulland-and-Holt plot, we will have a large scatter of points. Essentially, there are as many growth curves as experiments hidden in the cloud of points and these could be described by numerous VBGF parameter combinations. If we calculated a single regression from this scattered plot this would result in a pair of growth parameters expressing the central tendency in the whole of the combined experiments. Growth performance is governed by the concert of environmental variables encountered by the fish. The variance above and below the regression line can be attributed to the different factors or experimental treatments. The aim is to identify these variables, their interplay and to quantify their effects on growth, since they govern the shape of the growth curve (and the value of the corresponding VBGF parameters).

To explain more of the variance around the regression line and when numerous factors are involved, the bivariate regression can be extended to a multiple linear regression of the form:

$$y = a + b_1X_1 + b_2X_2 + ... + b_nX_n$$

Thus, the Gulland-and-Holt plot can be extended into a multiple regression form (Pauly and Ingles 1981; Pauly and Hopkins 1983; Hopkins et al. 1988), permitting environmental and treatment variables to be considered simultaneously in the same analysis:

$$(L_2 - L_1)/(t_2 - t_1) = a + b_1\bar{L} + b_2X_2 + ... + b_nX_n \quad ...11)$$

where

$$\bar{L} = (L_1 + L_2)/2 \quad ...12)$$

and where X_2 ... X_n are environmental and treatment variables simultaneously recorded during the growth increments.

The idea is to relate all environmental effects (controlled treatments and uncontrolled variables) to the mean growth of fish during the investigated interval. Fish growth is used as an instrument to assess the effects of external variables, their direction (positive or negative) and their strength.

This requires that additional variables are measured simultaneously with fish size data. The VBGF parameters are estimated from:

$$K = -b_1 \quad ...13)$$

where b_1 is the multiple regression coefficient of mean length. The value of L_∞ is the intercept of the multidimensional surface with the X_1-axis:

$$L_\infty = (a + b_2X_2 + ... + b_nX_n)/-b_1 \quad ...14)$$

The information on the influence of environmental and treatment factors is now 'imprinted' in L_∞ and K.

CONFIDENCE LIMITS FOR K

The confidence limits for K derived with the multiple regression version of the Gulland-and-Holt plot are derived in the same way as for the simple Gulland-and-Holt plot, but with degrees of freedom = n-u, where n is the number of cases in the regression, and u is the sum of dependent and independent variables included in the regression.

CONFIDENCE LIMITS FOR L_∞

Since L_∞ is obtained from a combination of the intercept, slopes and actual data values, the confidence limits are approximated using the method suggested by Sparre et al. (1989), described earlier in this paper.

Length-Weight Relationships

In cases where only the weight of fish was taken, these values may be transformed to lengths with a length-weight relationship.

The relationship between length and weight of a fish species is generally described by the relationship (Ricker 1975; Pauly 1984):

$$W = v \cdot L^m \qquad \text{...15)}$$

or in a linearized form:

$$\log W = \log v + m \cdot \log L \qquad \text{...16)}$$

where

- W = weight of individual fish (individual values should cover a large range of weights)
- L = length of individual fish (individual values should cover a large range of lengths)
- v = intercept with ordinate
- m = exponent of length-weight relationship

The logarithms used are usually based 10, but base e can also be used. The exponent m usually varies around three (Carlander 1969; Pauly 1984). When the exponent is exactly three, growth is called "isometric", otherwise it is called "allometric". The coefficients describing the relationship of weight to length vary according to biological and environmental factors (Ricker 1975). For each stock or population under investigation the length-weight relationship should be determined separately over a size range as wide as possible.

We have not followed the suggestion of Ricker (1975) to use a "functional" (= Type II, or GM) regression for the length-weight relationship in this and other contributions in this volume and used instead a predictive (= Type I, or AM) regression.

Our reasons for using Type I relationships (Sokal and Rohlf 1969) are:

- we require length-weight relationships largely for predictive purpose;
- we consider W to vary more than L and measurement errors to affect W more strongly than L (the GM regression assumes X and Y to have similar variability and measurement errors);
- most software packages for statistical analysis incorporate Type I but not Type II regression routines; and
- when r is close to unity (as is generally the case when W and L values covering a wide range of sizes are related), the difference between the two regression types vanishes.

For conversions from weight to length, equation 15 can be rearranged to:

$$L = \sqrt[m]{\frac{W}{v}} \qquad \text{...17)}$$

For the calculation of the allometric relationship (equation 15), the common practice is to linearize the variables through logarithmic transformation (equation 16) and to use linear regression to estimate the parameters v and m. This transformation introduces, however, a small systematic bias into the calculation when using the relationship for conversions. This can be accounted for with a correction factor (Finney 1941; Baskerville 1972; Beauchamp and Olson 1973; Whittaker and Marks 1975; Sprugel 1983), obtained from

$$F = e^{\left(\frac{SEE^2}{2}\right)} \qquad \text{...18)}$$

where SEE is the standard error of the estimate. It should be noted that this is based on natural logarithms, which should also be used in the regression. If the regression is determined on base-10 logarithms, this SEE should be transformed to natural base by multiplying the SEE with ln(10) = 2.303 before use in equation (18).

The correction procedure is to calculate predicted values with the allometric (here: length-weight) relationship and then to multiply each of these values with the correction factor (F) to eliminate log-transformation bias. Alternatively, the intercept v of the length-weight relationship can be multiplied with the correction factor (Vakily et al. 1986).

Data Requirements

The method presented above relates growth increments over short time periods to environmental or treatment effects measured during and averaged for these time intervals (Table 1).

The data requirements therefore are:

1. A cultured population of aquatic organisms must be sampled in length and weight at regular, short intervals. For shorter intervals (e.g., less than two weeks at tropical

Table 1. Extended Gulland-and-Holt plot: data table organized according to experiment duration and individual measurements during intervals.

	Date			Length		Environment			
						Variable-1		Variable-2	
Stocking	t1			L1					
		Δt	L		ΔL	xi : xi	$\bar{X}1$	xi : xi	$\bar{X}2$
1st sampling	t2			L2					
		Δt	L		ΔL	xi : xi	$\bar{X}1$	xi : xi	$\bar{X}2$
2nd sampling	t3			L3					
.	.			.		.		.	
.	.			.		.		.	
.	.			.		.		.	
Harvest	tn			Ln					

temperatures), growth in length will be difficult to detect and sampling stress may result. For longer intervals (e.g., six weeks or more at tropical temperatures), 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 and the model can be valid over a large size range of the species;
   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. Otherwise variables may become significant in the regression only due to natural or random variance; and
   c) the range should be large in relation to the variance induced by the treatment variables.
4. Data must conform to basic assumptions of regression [see Hopkins et al. (1988) for a discussion of how these assumptions apply to the extended Gulland-and-Holt method].

Computing the Extended Gulland-and-Holt Plot

It is necessary to assemble all data in form of a data table for final multiple regression analysis. The first step is to tabulate all data according to sampling intervals (Table 1).

From the sampling intervals mean values are calculated for all variables during the interval, together with the time interval in days, the average length and the length increment. 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 equal to the case numbers (Table 2). With this data matrix a multiple regression analysis is performed.

Hardware and Software Requirements

All steps of data compilation, processing and analysis can be performed on microcomputers. It is

Table 2. Extended Gulland-and-Holt plot: data matrix organized in final form appropriate for multiple regression analysis.

Case	Y	X1	X2	. . .	Xn
1	$\Delta L/\Delta t$	L	$\overline{\text{VAR1}}$	. . .	$\overline{\text{VARn}}$
2	"	"	"	. . .	"
3	"	"	"	. . .	"
.	"	"	"	. . .	"
.	"	"	"	. . .	"
.	"	"	"	. . .	"
n	"	"	"	. . .	"
			mean values of environmental variables in fish growth intervals		

recommendable to employ spreadsheet software that is widely used (e.g., Lotus 123, Borland Quattro, Microsoft Excel) to construct several small spreadsheets of identical format for ease of data handling. These can later be combined for analysis. Several statistical packages for personal computers exist that perform multiple regression analysis and supply necessary statistics for regression diagnostics (e.g., SPSS/PC+, SAS, PC-Statistician, etc.). These all enable scientists in developing countries, without access to mainframe computers, to perform these analyses, and to exchange the data with colleagues for comparative analyses.

Testing of the Model

The usual procedures and tests of statistical significance applied in least-squares regression should be applied. For a 'clean' hypothesis-testing approach (Prein et al. this vol.), each dataset was split into two parts by random sampling. One part contains 1/4 to 1/2 of the dataset and is used for model derivation; the other larger portion of the dataset is used for model testing (Norusis 1985).

Both portions of the datasets should produce regression models with the same set of significant variables. The signs and values of the regression coefficients should be the same, which can be tested by comparing the confidence intervals of the regression coefficients.

Use of the Model

In determining the model with multiple regression analysis two goals can be achieved:

1. From a number of measured variables those with the strongest influence on fish growth should be identified. The rest will fall out as insignificant during analysis. Usually there are more insignificant variables than significant ones, leaving only few variables in the final equation.
2. The effects of the variables governing fish growth are quantified, each for itself and all together "in concert". The relative strengths of the variables can be compared.

With the Gulland-and-Holt plot, "t_0" is not estimated since absolute ages are not known. Therefore a "recursive" form of the VBGF is used for predictions of growth (as originally published by von Bertalanffy in 1934):

$$L_2 = L_1 \cdot e^{-K\Delta t} + L_\infty(1 - e^{-Kt}) \qquad \text{...19)}$$

where L_1 and L_2 are the initial and final lengths, respectively, and Δt is the time interval of prediction.

With such a model the relative strengths of each of the variables can be used to deduce design and management implications in further culture operations. The whole model in itself can be used to make short- and long-term predictions of fish growth under certain, anticipated conditions (stocking density, feeding, temperature, etc.). Under given constraints, such as minimum market-size requirements or economic demands, the correct design and management scheme of a commercial culture operation may be determined using a biologically founded growth model incorporating environmental effects (Fig. 4).

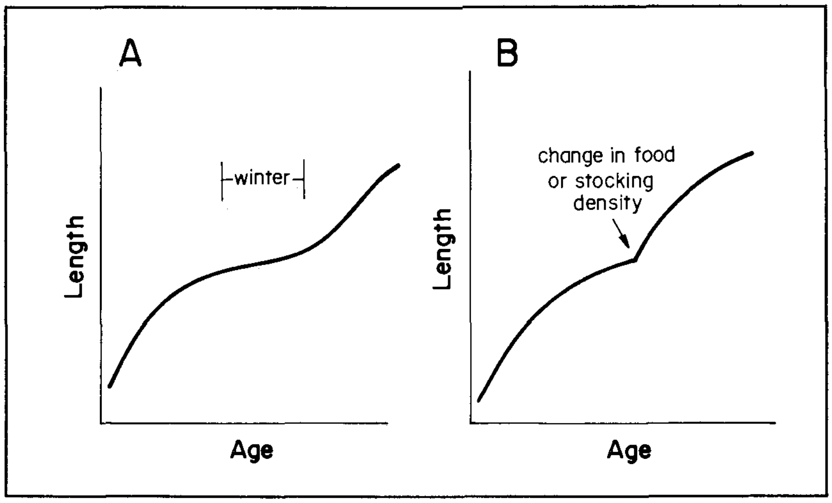


Fig. 4. Schematic examples for applications of the VBGF with inclusion of environmental and treatment variables to permit growth predictions under anticipated or planned conditions. A) High growth in summer and low growth in winter. B) Changes in feeding rate or stocking density can be modelled by the extended Gulland-and-Holt method.

Table 3. Some common problems encountered in aquaculture growth experiments and "solutions".

Problems	"Solutions"	Reference[a]
1. Changing environmental conditions during the culture period	a. Terminate the experiment when the change occurs	Maguire and Hume (1982); Rappaport and Sarig (1979)
	b. Average growth over the whole period	Sin (1982)
2. Loss of a replicate due to circumstances unrelated to the experiment	a. Drop the replicate	Hopkins and Cruz (1982)
	b. Use missing plot procedures	Gomez and Gomez (1976)
3. Limited number of ponds/culture units.	a. Limit the number of treatments and/or replicates.	b
	b. Use complex experimental designs such as fractional factorials, etc.	Montgomery (1976)
4. Limited numbers of fish of one size for stocking	a. Multiple nursing	Wohlfarth and Moav (1972)
	b. Limit the number of treatments and/or replicates	b
5. Changing growth responses as the fish grow larger	a. Conduct experiments over short size ranges only	Knights (1983); Bryant and Matty (1981)
	b. Average response over the whole size range	Eldani and Primavera (1981)
6. Comparison of experiments which used different sizes of fish	a. Average growth rate using g/day or daily rates of increase, etc.	Coche (1982); Halevy (1979)
7. Estimation of size at time t	a. Use an average growth rate such as $g \cdot day^{-1}$	Huguenin and Rothwell (1979)
	b. Use "sophisticated" growth equations such as von Bertalanffy, etc.	Gates and Mueller (1975); Elliot (1975)

[a]Papers in which the "solution" is used or described.
[b]A standard technique which is usually not specifically mentioned by authors.

If mortality is negligible, fish yield at harvest is equal to the product of the number of fish stocked times the sum of measured growth increments. Alternatively, instantaneous mortality can be estimated (see paper by Hopkins and Pauly, this vol.) and included in the model.

Discussion

Growout experiments play an essential role in aquaculture research in estimating the growth potential of various species and strains, in assessing the value of different feeds or treatments, etc. The problems with pond growout experiments are, however, that it is generally very difficult to *control* effectively the "control variables" and generally impossible to control extraneous variables (e.g., climatic factors) likely to affect the results of such experiments, resulting in "experiments" that in fact *cannot* be duplicated. Finally, since experiments are extremely costly in both time and resources, long-term aquaculture experiments that are intended to represent the time scale of commercial pond operations generally tend to be too limited to obtain secured and generalizable results (Table 3).

Explicit statements of this problem are few. Similarly, few papers are available in which optimal experimental designs for aquaculture research are discussed. This situation contrasts markedly with that prevailing in agriculture research, where experimental designs and analytical formats have traditionally benefitted each other, to the extent that whole chapters of statistical textbooks are devoted to them (Steel and Torrie 1960; Prowse 1968; Gropp 1979; Gomez and Gomez 1976).

In this paper, a method for the reduction and analysis of data from growout experiments was presented in form of the extended Gulland-and-Holt plot, which allows one to overcome the problems listed in Table 3. It is based on two assumptions whose validity can be assessed for any given set of experiments: a) that the growth of the fish can be described by the von Bertalanffy growth function, and b) that the effects of treatments and environment express themselves through changes in fish growth.

We conclude that the extended Gulland-and-Holt plot can be applied when the growth rate (dl/dt) of the fish in the experiments can be described by the von Bertalanffy growth function (VBGF) and when mainly environmental or treatment effects govern fish growth.

The applicability and benefit of this method for the analysis of aquaculture experiments has been demonstrated (Prein 1985, 1990, this vol.; Aquino-Nielsen et al., this vol.; Prein and Pauly, this vol.; Hopkins et al. 1988).

References

Bardach, J.E., J.H. Ryther and W.O. McLarney. 1972. Aquaculture, the farming and husbandry of freshwater and marine organisms. John Wiley, New York. 868 p.

Baskerville, G.L. 1972. Use of logarithmic regression in the estimation of plant biomass. Can. J. Forest Res. 2:49-53.

Beauchamp, J.J. and J.S. Olson. 1973. Corrections for bias in regression estimates after logarithmic transformation. Ecology 54:1403-1407.

Beverton, R.J.H. and S.J. Holt. 1957. On the dynamics of exploited fish populations. Fish. Investigations (Series II) 19:1-533. UK Ministry for Agriculture, Fisheries and Food, London.

Bryant, P.L. and A.J. Matty. 1981. Adaptation of carp (*Cyprinus carpio*) larvae to artificial diets. 1. Optimum feeding rate and adaptation age for a commercial diet. Aquaculture 23:275-286.

Carlander, K.D. 1969. Handbook of freshwater fishery biology. Vol. I, Iowa State University Press, Ames, Iowa. 752 p.

Coche, A.G. 1982. Cage culture of tilapias, p. 205-246. *In* R.S.V. Pullin and R.H. Lowe-McConnell (eds.) The biology and culture of tilapias. ICLARM Conf. Proc. 7, 432 p.

Eldani, A. and J.H. Primavera. 1981. Effect of different stocking combinations on growth, production and survival of milkfish (*Chanos chanos* Forskal) and prawn (*Penaeus monodon Fabricius*) in polyculture in brackish water ponds. Aquaculture 23:59-72.

Elliot, J.M. 1975. The growth rate of brown trout (*Salmo trutta* L.) fed on reduced ratios. J. Anim. Ecol. 44:823-842.

Finney, D.J. 1941. On the distribution of a variate whose logarithm is normally distributed. J.R. Stat. Soc. (Series B) 7:155-161.

Gates, J.M. and J.J. Mueller. 1975. Optimizing the growth and marketing of fish in a controlled environment. Mar. Technol. Soc. 9(5):13-16.

Gomez, K.A. and A.A. Gomez. 1976. Statistical procedures for agricultural research with emphasis on rice. International Rice Research Institute, Los Baños, Philippines. 294 p.

Gomez, K.A. and A.A. Gomez. 1984. Statistical procedures for agricultural research. Second edition. John Wiley, New York. 680 p.

Gropp, J. 1979. Standard fish husbandry designs of experiments and their evaluation, p. 445-452. *In* J.E. Halver and K. Tiews (eds.) Finfish nutrition and fish feed technology. Vol. II. Heenemann Verlagsgesellschaft, Berlin.

Gulland, J.A. 1967. Manual of methods for fish stock assessment. Part I. Fish population analysis. FAO Fish. Ser. 3, 154 p.

Gulland, J.A. 1983. Fish stock assessment. A manual of basic methods. John Wiley & Sons, Chichester. 223 p.

Gulland, J.A. and S.J. Holt. 1959. Estimation of growth parameters for data at unequal time intervals. J. Cons. CIEM 25:47-49.

Halevy, A. 1979. Observations on polyculture of fish under standard farm pond conditions at the Fish and Aquaculture Research Station, Dor, during the years 1972-1977. Bamidgeh 31(4):96-104.

Hepher, B. and Y. Pruginin. 1981. Commercial fish farming, with special reference to fish culture in Israel. John Wiley, New York. 261 p.

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

Hopkins, K.D., M.L. Hopkins and D. Pauly. 1988. A multivariate model of tilapia growth, applied to seawater tilapia culture in Kuwait, p. 29-39. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Huguenin, J.E. and G.N. Rothwell. 1979. The problems, economic potentials and system design of large future tropical marine fish cage systems. Proc. World Maricult. Soc. 10:162-181.

Knights, B. 1983. Food particle-size preferences and feeding behaviour in warmwater aquaculture of European eel *Anguilla anguilla* (L.). Aquaculture 30:173-190.

Maguire, G.B. and I.D. Hume. 1982. A study of the nutritional requirements of school prawns *Metapenaeus macleayi* (Haswell) in some Australian brackish water farming ponds. Aquaculture 29:261-278.

Montgomery, D.C. 1976. Design and analysis of experiments. John Wiley, New York. 418 p.

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

Pauly, D. 1979. Gill size and temperature as governing factors in fish growth: a generalization of von Bertalanffy's growth formula. Ber. Inst. Meereskd. Univ. Kiel No. 63. 156 p.

Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Stud. Rev. 8, 325 p.

Pauly, D. and J. Ingles. 1981. Aspects of the growth and natural mortality of exploited coral reef fishes, p. 89-98. *In* E.D. Gomez, C.E. Birkeland, R.W. Buddemeier, R.E. Johannes, J.A. Marsh and R.T. Tsuda (eds.) The reef and man. Proceedings of the Fourth International Coral Reef Symposium, 18-22 May 1981, Manila, Philippines. Vol. I. Marine Science Center, University of the Philippines, Quezon City. 725 p.

Pauly, D. and K.D. Hopkins. 1983. A method for the analysis of pond growth experiments. ICLARM Newsl. 6(1):10-12.

Pauly, D. and J.L. Munro. 1984. Once more on growth comparisons in fish and invertebrates. Fishbyte 2(1):21.

Pauly, D., J. Moreau and M. Prein. 1988. A comparison of overall growth performance of tilapia in open waters and in aquaculture, p. 469-479. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 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. Mathemathisch-Naturwissenschaftliche Fakultät, Christian-Albrechts-Universität Kiel. 125 p. Doctoral dissertation.

Prowse, G.A. 1968. Standardization of statistical methods in fish culture research, p. 386-396. *In* T.V.R. Pillay (ed.) Proceedings of the FAO Symposium on Warmwater Pond Culture. FAO Fish. Rep. 44. Vol. 4.

Pütter, A. 1920. Studien über physiologische Ähnlichkeit. VI. Wachstumsähnlichkeiten. Pflügers Archiv für die gesamte Physiologie 180:298-340.

Rappaport, U. and S. Sarig. 1979. The effect of population density of carp in monoculture under conditions of intensive growth. Bamidgeh 31(2):26-34.

Reay, P.J. 1979. Aquaculture. Studies in Biology No. 106. Edward Arnold Publishers, London. 60 p.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191. Department of Fisheries and Oceans, Ottawa, Canada. 382 p.

Sin, A.W. 1982. Stock improvement of the common carp in Hong Kong through hybridization with the introduced Israeli race 'Dor-70'. Aquaculture 29:299-304.

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

Sokal, R.R. and F.J. Rohlf. 1969. Biometry. The principles and practice of statistics in biological research. W.H. Freeman, New York. 776 p.

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

Sprugel, D.G. 1983. Correcting for bias in log-transformed allometric equations. Ecology 64:209-210.

Steel, R.G.D. and J.H. Torrie. 1960. Principles and procedures of statistics with special reference to the biological sciences. McGraw-Hill, New York. 481 p.

Steffens, W., Editor. 1981. Industriemässige Fischproduktion. VEB Deutscher Landwirtschaftsverlag, Berlin. 376 p.

Vakily, J.M. 1989. Microcomputer use in experimental aquaculture: what software is needed? Aquabyte 2(3):2-4.

Vakily, J.M., M.L. Palomares and D. Pauly. 1986. Computer programs for fish stock assessment: applications for the HP 41CV calculator. FAO Fish. Tech. Pap. 101. Supplement 1. 255 p.

von Bertalanffy, L. 1934. Untersuchungen über die Gesetzlichkeit des Wachstums. I. Teil: Allgemeine grundlagen der Theorie; mathematische und physiologische Gesetzlichkeiten des Wachstums bei Wassertieren. Roux' Archiv für Entwicklungsmechanik 131:613-652.

von Bertalanffy, L. 1938. A quantitative theory of organic growth (enquiries on growth laws, 2). Human Biol. 10:181-213.

Whittaker, R.H. and P.L. Marks. 1975. Methods of assessing terrestrial productivity, p. 55-118. *In* H. Lieth and R.H. Whittaker (eds.) Primary productivity of the biosphere. Ecological Studies 14, Springer, Berlin. 339 p.

Wohlfarth, G. and R. Moav. 1972. The regression of weight gain on initial weight in carp. I. Methods and results. Aquaculture 1:7-28.

Factor and Canonical Correlation Analyses: Basic Concepts, Data Requirements and Recommended Procedures

ANA MILSTEIN, *Agricultural Research Organization, Fish and Aquaculture Research Station, Dor, Mobile Post Hof Hacarmel 30820, Israel*

Milstein, A. 1993. Factor and canonical correlation analyses: basic concepts, data requirements and recommended procedures, p. 24-31. *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

Factor analysis and canonical correlations are multivariate statistical methods appropriate to explore data from complex systems in which there is a high degree of interaction allowing to consider simultaneously the different variables involved. These methods are presented through examples of application to aquaculture data, in order to introduce the aquaculture-oriented reader into the multivariate statistical world.

Introduction

Factor analysis and canonical correlations are multivariate statistical methods which allow to study patterns of interrelationships within one set of variables and between two sets of variables, respectively. Both methods are based on the analysis of linear relationships, which express the simplest relationships between variables. These techniques are appropriate to explore data from complex systems, such as aquaculture ponds, in which there is a high degree of interaction between the different elements of the system (fish biology, pond ecology and management procedures). The methods are described in standard texts on the subject (e.g., Horst 1965; Morrison 1967; Malinvand 1970; Cooley and Lohnes 1971; Lawley and Maxwell 1971; Mulaik 1972; Tatsuoka 1972; Comrey 1973; Harman 1976) and in texts for users without strong mathematical background (e.g., Seal 1964; Lefebvre 1976; Levine 1977; Jeffers 1978; Kim and Mueller 1978a, 1978b). The present paper briefly presents the methods through examples of application to analyses of aquaculture data in order to introduce the aquaculture-oriented reader into the world of multivariate statistics.

Basic Concepts

Factor Analysis

Factor analysis refers to a variety of statistical techniques, from which the simplest and the one on which this paper concentrates is the Principal Components method. The common objective of factor analysis techniques is to reduce the number of variables into a smaller number of new, hypothetical variables. An everyday example of such a hypothetical variable is "weather condition", whose values (from "bad" to "good") are a combination of temperature, humidity, rain, cloudiness and a number of other measurable variables.

Factor analysis assumes that the observed variables are linear combinations of some underlying (unobservable) factors, independent of one another, which generally reflect an ecological or operational process. Data handling steps involved in a factor analysis are shown in Fig. 1. The starting point for the calculation of the factors is generally the correlation matrix between variables but the variance-covariance matrix can also be used. From this matrix the method computes the linear combination of the original variables, which accounts for most of the variance in the dataset called first factor (FACTOR1). This is followed by the calculation of a second linear function (FACTOR2), which is independent of the first and accounts for most of the remaining variance, and so on. The number of factors that can possibly be calculated equals the number of variables included in the analysis. However, since most of the variability will be accounted for in the first few combinations, the last ones can be neglected. The factors have no units and are normally distributed standardized variables, with mean = 0 and variance = 1. The value of each factor for each observation of the original variables can be calculated and used as a new variable in plots, histograms and statistical analyses such as ANOVA or regressions. In matrix terminology, the extraction of factors from a data matrix implies the calculation of the eigenvalues and eigenvectors of that matrix. The proportion of variance accounted for by each factor is calculated from the corresponding eigenvalue. Each eigenvector contains the coefficients of the linear

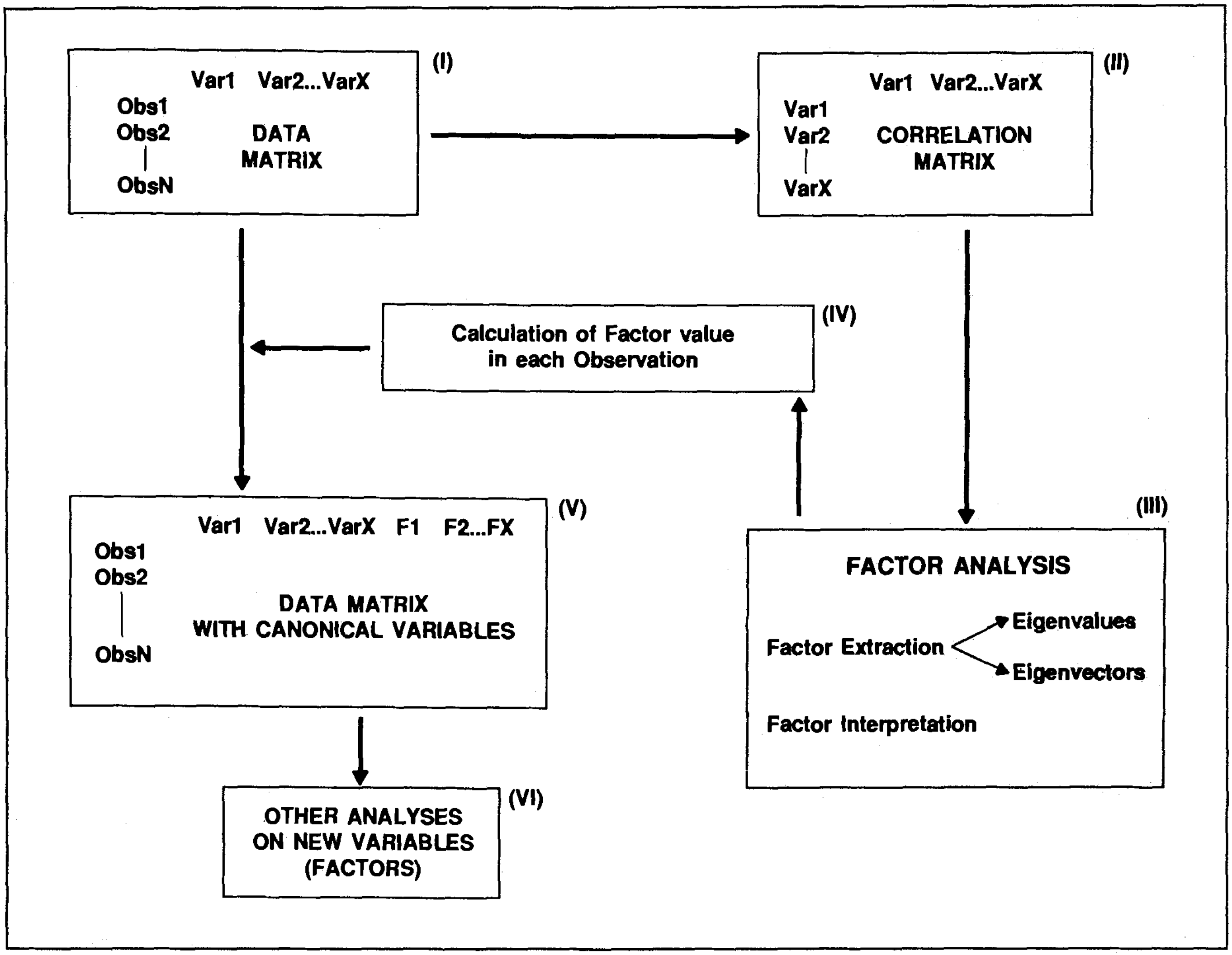


Fig. 1. Factor analysis: steps in data handling. Roman numbers indicate order of steps. Var_i = variable. Obs_i = observation. F_i = factor variable.

combination corresponding to each original variable. The interpretation of the factor (hypothetical variable) is then performed based on the relative size and sign of these coefficients.

As an example of the application of this technique to aquaculture, the data analyzed by Milstein et al. (1989) are presented here. Table 1 shows the computer output of a factor analysis of water quality data from dual-purpose reservoirs used for fish culture and crop irrigation. The analysis was run on 143 observations, each one including measurements of dissolved oxygen (DO), ammonium (NH4), nitrite + nitrate (NO2_3), total nitrogen (N_TOT), dissolved phosphorus (P_DIS), total phosphorus (P_TOT), organic carbon (C_ORG) and pH (PH). The section of the output related to the eigenvalues presents the elements used to determine how many factors should be retained. The two criteria most commonly used are: factors with eigenvalues larger than 1 and differences between eigenvalues of relatively large consecutive factors. According to these criteria, the first three factors should be retained, since the difference between the third and fourth eigenvalues is already relatively small. Sometimes when the decision is not so easy as in this case then the proportion of variance explained is used.

The factor pattern section of the output presents the coefficients (eigenvector) of the first three factors associated with each original

Table 1. Computer output of a factor analysis.

SAS

FACTOR PROCEDURE

Factor Analysis

INITIAL FACTOR METHOD: PRINCIPAL COMPONENTS

PRIOR COMMUNALITY ESTIMATES: ONE

EIGENVALUES OF THE CORRELATION MATRIX: TOTAL = 8 AVERAGE = 1

	1	2	3	4
EIGENVALUE	2.964939	1.654771	0.999581	0.877779
DIFFERENCE	1.310168	0.655189	0.121802	0.231288
PROPORTION	0.3706	0.2068	0.1249	0.1097
CUMULATIVE	0.3706	0.5775	0.7024	0.8121

	5	6	7	8
EIGENVALUE	0.646492	0.417729	0.299791	0.138918
DIFFERENCE	0.228763	0.117939	0.160872	
PROPORTION	0.0808	0.0522	0.0375	0.0174
CUMULATIVE	0.8929	0.9452	0.9826	1.0000

3 FACTORS WILL BE RETAINED BY THE NFACTOR CRITERION

FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
DO	0.55818	0.68666	-0.19709
NH4	-0.65032	-0.03791	0.50630
NO2_3	-0.48793	0.53387	0.46167
N_TOT	0.77939	0.03107	0.46398
P_DIS	-0.66744	0.14019	-0.31859
P_TOT	0.38987	-0.23277	-0.30873
C_ORG	0.86092	-0.20592	0.26691
PH	0.21512	0.88295	-0.08901

variable. To interpret the factors, only the coefficients with the highest values are considered, usually those larger than 0.5. For example, FACTOR1 includes high positive contributions of C_ORG, N_TOT and DO, as well as high negative contributions of dissolved nutrients (P_DIS, NH4 and NO2_3). This combination points to "algal activity" as the hypothetical underlying variable we are looking for, since as algae develop, photosynthesis, algal biomass and nutrient absorption increase. Accordingly, there is an increase in DO and in algal biomass measured as particulate organic carbon and total nitrogen (high positive coefficients of these variables) and a decrease in nutrients in the water body due to their uptake by the algae (high negative coefficients). The same analysis applied to FACTOR2 shows high contributions of pH, DO and also oxygenated inorganic nitrogen compounds. This combination points to "oxygenation" of the water resulting from processes other than algal activity (already accounted for by the FACTOR1) as the hypothetical underlying variable we are looking for. This factor reflects the combined effects of heterotrophic activity and wind action. Processes like respiration and decomposition lower pH and DO concentrations, while wind action increases DO concentration. In a general way, under aerobic conditions nitrification is promoted and nitrate accumulates in the water, while a low DO level retards this process and promotes denitrification.

Canonical Correlation Analysis

The multivariate model of canonical correlations is an extension of the Principal Component method which enables to summarize and explore complex relationships between two sets of variables rather than within one set. This analysis gives a direct measurement of how much of the variability of the two sets can be accounted for by the relationships between them. One set consists of explanatory (independent) variables (e.g., fish stocking parameters, nutritional inputs, etc.) and the other of response (dependent) variables (e.g., fish growth rate or yields). Each set may contain several variables. Multiple and simple regressions are special cases of canonical correlation in which one or both sets contain a single variable, respectively. Data handling steps involved in a canonical correlation analysis are shown in Fig. 2, where the sets of variables are separated by broken lines. The starting point of calculations is again the correlation matrix between variables, but the variance-covariance matrix can also be used. From this matrix the canonical correlation method computes a linear combination for each variable set, called a canonical variable, such that the correlation between the two canonical variables is maximized. This is the first canonical correlation, which accounts for most of the variance in the dataset. The coefficients of the linear combinations are canonical coefficients, which are normalized so

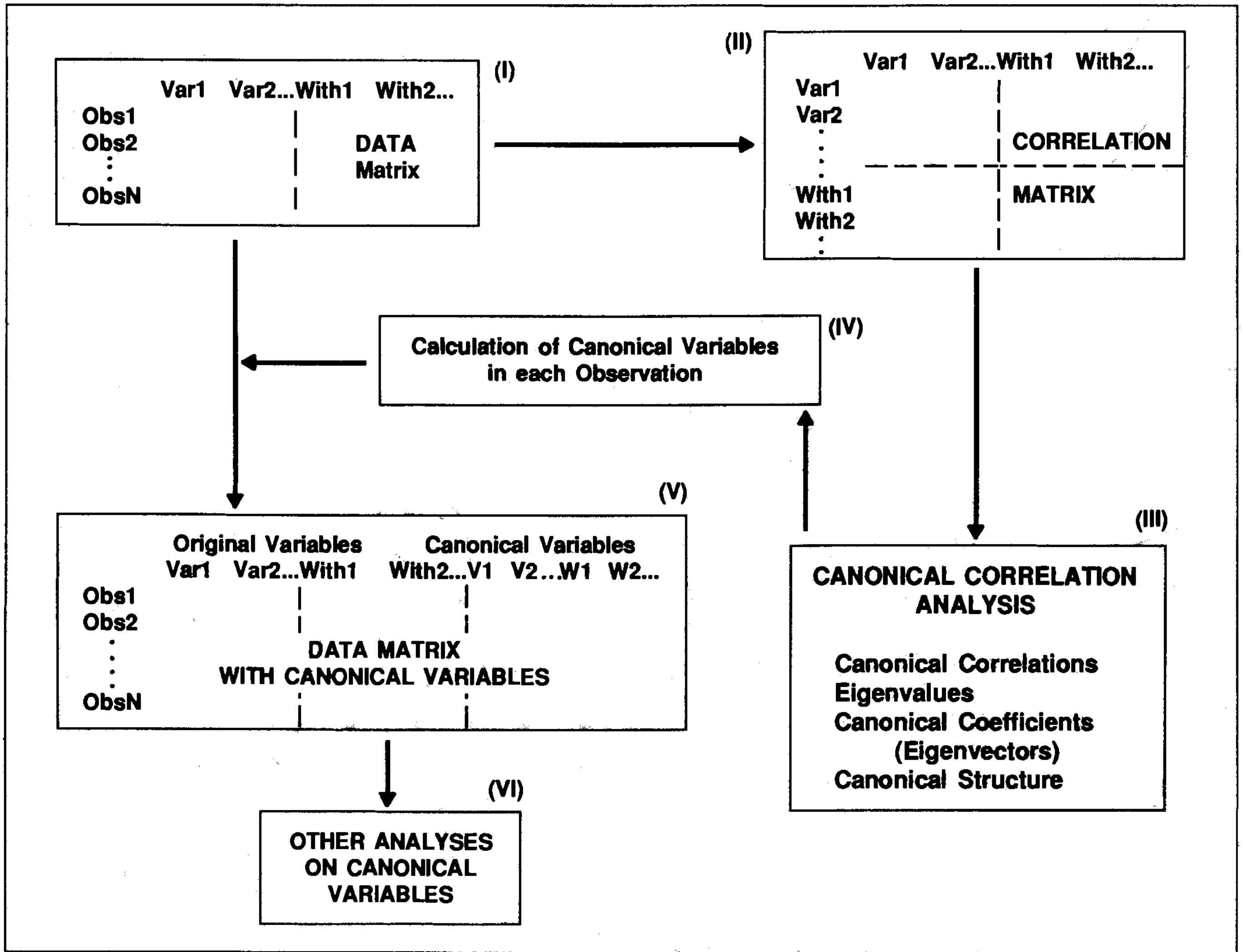


Fig. 2. Canonical correlation analysis: steps in data handling. Roman numbers indicate order of steps. Broken line separates areas related to different sets of variables. Var_i = explanatory variable. $With_i$ = response variable. V_i and W_i = canonical variables. Obs_i = observation.

that each canonical variable has a variance of one. The second set of canonical variables, uncorrelated with the first pair, produces the second highest correlation coefficient. This second canonical correlation accounts for most of the residual variance remaining after calculating the first one. The process of constructing canonical variables continues until the number of pairs of canonical variables equals the number of variables in the smaller set.

As an example for application of this technique to aquaculture, the data analyzed by Milstein et al. (1988) are presented here. Table 2 shows the computer output of a canonical correlation analysis on fish growth and management inputs in a polyculture system. The output consists of three pages, the first one relates to the canonical correlation coefficients, the second page to the canonical coefficients for the two sets of variables, and the third page to the canonical structure. The analysis was run on 105 observations, each one including as explanatory variables (VAR variables in the output) feed pellets (PEL_FISH), organic manure (MANURE), numbers and weights of common carp, tilapia and silver carp at stocking

Table 2. Computer output of a canonical correlation analysis.

(page 1)

SAS

CANCORR PROCEDURE

Canonical Correlation Analysis

	Canonical Correlation	Adjusted Canonical Correlation	Approx Standard Error	Squared Canonical Correlation
1	0.838525	0.800545	0.029111	0.703125
2	0.817808	.	0.032476	0.668811
3	0.752131	.	0.042586	0.565702

Eigenvalues of INV(E)*H
= CanRsq/(1-CanRsq)

	Eigenvalue	Difference	Proportion	Cumulative
1	2.3684	0.3490	0.4162	0.4162
2	2.0194	0.7169	0.3549	0.7711
3	1.3026	.	0.2289	1.0000

Test of H0: The canonical correlations in the current row
and all that follow are zero

	Likelihood Ratio	Approx F	Num DF	Den DF	Pr > F
1	0.04270104	22.3853	24	273.2298	0.0001
2	0.14383502	22.2129	14	190	0.0001
3	0.43429833	20.8410	6	96	0.0001

S=3 M=2 N=46.5

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilk's Lambda	0.04270104	22.3853	24	273.2298	0.0001
Pillai's Trace	1.93763710	21.8867	24	288	0.0001
Hotelling-Lawley Trace	5.69040419	21.9713	24	278	0.0001
Roy's Greatest Root	2.36841963	28.4210	8	96	0.0001

Note: F Statistics for Roy's Greatest Root is an upper bound.

continued

Table 2 continued

(page 2) SAS

CANCORR PROCEDURE

Canonical Correlation Analysis

Standardized Canonical Coefficients for the 'VAR' Variables

	V1	V2	V3
PEL_FISH	0.5655	-0.0543	-0.6454
MANURE	-0.0906	-0.4400	-0.9025
N_CARP	0.2218	-0.2464	0.1623
WI_CARP	0.1595	-0.1318	-0.1295
N_TILAP	-0.0400	0.0081	0.7452
WI_TILAP	-0.2269	1.1169	0.0892
N_SILV	-0.8910	0.0665	-0.1260
WI_SILV	0.0228	-0.3779	-0.0726

Standardized Canonical Coefficients for the 'WITH' Variables

	W1	W2	W3
GR_CARP	0.4962	-0.1901	-1.0936
GR_TILAP	-0.0298	0.9666	-0.3297
GR_SILV	0.6368	0.1112	1.0481

(page 3) SAS

CANCORR PROCEDURE

Canonical sturcture

Correlations between the 'VAR' variables and their canonical variables

	V1	V2	V3
PEL_FISH	0.5892	-0.0215	-0.3692
MANURE	-0.4486	-0.0280	-0.3023
N_CARP	-0.3246	-0.4293	0.0517
WI_CARP	0.2676	0.1857	0.1098
N_TILAP	0.0300	-0.0715	0.6483
WI_TILAP	-0.0031	0.8559	0.1016
N_SILV	-0.7243	-0.2971	0.1739
WI_SILV	0.4196	0.0804	0.0161

Correlations between the 'WITH' Variables and their canonical variables

	W1	W2	W3
GR_CARP	0.8523	-0.1451	-0.5025
GR_TILAP	0.0630	0.9877	-0.1431
GR_SILV	0.9091	0.1593	0.3849

continued

Table 2 continued

Correlations between the 'VAR' variables and the canonical variables of the 'WITH' variables

	W1	W2	W3
PEL_FISH	0.4941	-0.0176	-0.2777
MANURE	-0.3762	-0.0229	-0.2273
N_CARP	-0.2722	-0.3511	0.0389
WI_CARP	0.2244	0.1519	0.0826
N_TILAP	0.0025	-0.0585	0.4876
WI_TILAP	-0.0026	0.6999	-0.0764
N_SILV	-0.6073	-0.2429	-0.1308
WI_SILV	0.3518	0.0657	0.0121

Correlations between the 'WITH' variables and the canonical variables of the 'VAR' variables

	V1	V2	V3
GR_CARP	0.7147	-0.1187	-0.3779
GR_TILAP	0.0528	0.8078	-0.1076
GR_SILV	0.7623	0.1303	0.2895

(N_CARP, WI_CARP; N_TILAPIA, WI_TILAPIA; and N_SILV, WI_SILV; respectively), and as response variables (WITH variables in the output) daily growth rates of each fish species (GR_CARP, GR_TILAP GR_SILV). Since the smaller set has three variables, three canonical correlations can be calculated. The first page of the output shows, among other things, the three canonical correlation coefficients (0.838, 0.818 and 0.752), the three eigenvalues and corresponding proportion of variance explained, and four tests of whether there is any significant link between the sets (at the bottom of the page. All F tests are highly significant in the example, meaning that there are links between the sets); the test of significance of each correlation (all three highly significant in the example) implies that there was a probability of only one in 10,000 to obtain the corresponding canonical correlation coefficient just by chance.

The second page of the output presents the standardized canonical coefficients for both sets of variables. As in the factor analysis, the highest coefficients (generally those greater than 0.5) are used for interpretation. For example, the first pair of canonical variables (V1 and W1) shows that the growth rates of both common carp and silver carp (positive and rather similar coefficients in W1) increase mainly with decreased silver carp stocking density (N_SILV highly negative coefficient in V1), and secondly with increasing amounts of feed pellets (PEL_FISH positive but lower value in V1). This example also shows that the computer output alone is not enough to understand the processes in the system from which the data were obtained. From the V1-W1 coefficient just analyzed, it could be concluded that at low silver carp density the growth rates of both fish species increase with the amount of feed pellets. But silver carp does not eat pellets. This variable therefore mainly affects the growth of common carp and in the data analyzed it happened that more pellets were supplied when silver carp density was low. The resulting interactions between these two species are such that the canonical correlation analysis failed to isolate growth of each species in separate canonical correlations.

The analysis of the other sets of canonical coefficients is done in a similar way. Thus, the second pair of canonical variables (V2 and W2) shows tilapia growth rate correlated mainly to its stocking weight and secondly (and lagging well behind) with manuring rate and silver carp stocking weight.

The third page of the computer output presents the canonical structure matrix, that is the correlations between the original variables and the canonical variables. Its analysis allows to check for multicollinearity (intercorrelations within each set of variables). If two variables are closely correlated with each other, once one of the two has made its contribution to the canonical variable, the other has no additional autonomous contribution to make. The first variable's coefficient will be high and the second will be near zero because it is hidden or suppressed by the first one. In our example, the correlations between the VAR variables and the V1 and W1 canonical variables show that manure and silver carp stocking weight were suppressed by feed pellets and silver carp density. Small silver carp are generally stocked at higher densities than large ones, where this negative correlation accounts for the inclusion of N_SILV but not of WI_SILV. A similar situation holds for the nutritional inputs since the data included experiments on replacement of feed pellets by manure, which explains their negative relationship.

Data Requirements and Recommended Procedures

The first step for running multivariate analyses is to construct a data matrix in which each column contains a variable and each line an observation. This step may be very simple when all the data come from a single experiment, or very complex, e.g., when a database is assembled from different sources. In both cases, but especially in the second, the documentation of all details is absolutely essential (Juanico 1989). Documentation includes all relevant details on data sources, details on measurement, units of all recorded variables, equations and/or transformations for all calculated variables, definitions and ranges of variables, criteria for selection of variables, all changes (e.g., corrections) of values, procedures used, etc. Documentation is often neglected by inexperienced users, who then realize - only too late - that "what is not written does not exist".

In the construction of the data matrix for factor analysis or canonical correlations, several specific points should be taken into account:

1. All the variables must be continuous, or, if discrete, they should increase by sufficiently small intervals of measurement to be regarded as approximately continuous. Categorical variables (such as treatment, farm, sex, etc.) can be "added" in an indirect way at a second step: after the value of the factors or canonical variables in each observation is calculated, these new variables can be used as any "regular" variable in ANOVA or other tests which handle discrete variables. Examples for this are presented in Milstein and Hulata (this vol.).
2. No ratios or linear functions of the original variables should be added to the variables included in the analysis.
3. No missing values are allowed. A missing value in one variable causes the automatic elimination of the whole line of the entire observation from the analysis. Thus, an effort should be made to replace missing values by computed ones, using suitable techniques (see for example Prein, this vol.).
4. The data should be as homogeneous as possible. If a variable (or group of variables) include a (set of) value(s) well outside the range defined in the other values, the variance introduced by this exceptional (set of) value(s), or outlier(s) will be accounted for in the first factors or canonical correlations, which will be "true" but of little interest. For instance, in an analysis of a polyculture system in which tilapia was present in only 3% of the ponds, the first factor comprised all and only the variables related to this fish. Deletion of the few observations related to tilapia (outliers in this case) led to a homogenous dataset which could then be analyzed straightforwardly.

Once the data matrix is ready, the analyses are carried out using statistical packages. Multivariate analyses are standard procedures in most of them and are available for mainframe and personal computers. The outputs presented as examples above were produced by the SAS (1985) procedures FACTOR and CANCOR run on PC, and analyses of the larger dataset presented by Milstein and Hulata (this vol.) were run with the same package on a mainframe computer. Among other packages which include procedures for factor (or principal component) analysis and canonical correlations are BMD, DATATEXT, OSIRIS, SPSS and STATGRAPH.

Discussion

The methods presented herein allow to study relationships in complex systems and simultaneously consider the different variables involved. Factor analysis is very useful as an exploratory technique when all the variables are of the same nature (water quality, or meteorological data, or plankton counts, etc.). It allows to identify the major processes impacting on a system and to discover outliers among the data points. Canonical correlations allow to explain the *simultaneous* variance in a group of dependent variables in terms of a combination of independent variables. This technique is related to the widely used regression methods, differing in (a) the amount of variables in each set and (b) the nature of the analysis. (a) Canonical correlation (CA) explores relationships between a set of several dependent variables with a set of several independent variables. When the set of dependent variables contains only one variable, CA reverts to multiple regression analysis; when the independent set also contains only one variable, CA reverts to simple regression analysis. (b) Regressions are usually predictive (but see Prein and Pauly, this vol.), while canonical correlation may be exploratory or predictive. For a predictive analysis the variance-covariance matrix should be the starting step of calculations. The use of that matrix makes the analysis very sensitive to the units used and is not recommended when the ranges of the variables are very different, as often happens (e.g., fish weights in the order of few to tens of grams, fish densities in the order of thousands). Regression and canonical correlations have different objectives, and can be used on the same dataset to obtain different insights into the complex relationships embedded in that dataset.

References

Comrey, A.L. 1973. A first course in factor analysis. Academic Press, New York.

Cooley, W.W. and P.R. Lohnes 1971. Multivariate data analysis. John Wiley, New York.

Harman, H.H. 1976. Modern factor analysis. University of Chicago Press, Chicago.

Horst, P. 1965. Factor analysis of data matrices. Holt, Rinehart and Winston, New York.

Jeffers, J.N.R. 1978. An introduction to system analysis: with ecological applications. Edward Arnold, London.

Juanico, M. 1989. A database for a multi-institutional environment monitoring program. Environ. Monit. Assess. 12:181-190.

Kim, J.O. and C.W. Mueller. 1978a. Introduction to factor analysis: what it is and how to do it. Sage University papers on quantitative applications in the social sciences, 07-013. Sage Publications, Beverly Hills and London.

Kim, J.O. and C.W. Mueller 1978b. Factor analysis: statistical methods and practical issues. Sage University papers on quantitative applications in the social sciences, 07-014. Sage Publications, Beverly Hills and London.

Lawley, D.N. and A.E. Maxwell. 1971. Factor analysis as a statistical method. Butterworth and Co., London.

Lefebvre, J. 1976. Analyses statistiques multidimensionnelles. Masson, Paris.

Levine, M.S. 1977. Canonical analysis and factor comparison. Sage University papers on quantitative applications in the social sciences, 07-006. Sage Publications, Beverly Hills and London.

Malinvand, E. 1970. Statistical methods of econometrics. Elsevier, New York.

Milstein, A., M. Feldlite, N. Mozes and Y. Avnimelech. 1989. Limnology of reservoirs used for fish farming and crop irrigation, with integrated free and cage fish culture. Isr. J. Aquacult./Bamidgeh 41(1):12-22.

Milstein, A., G. Hulata and G.W. Wohlfarth. 1988. Canonical correlation analysis of relationships between management inputs and fish growth and yields in polyculture. Aquacult. Fish. Manage. 19:13-24.

Morrison, D.F. 1967. Multivariate statistical methods. McGraw-Hill, New York.

Mulaik, S.A. 1972. The foundations of factor analysis. McGraw-Hill, New York.

SAS. 1985. SAS user's guide: statistics. Version 5 edition. Statistical Analysis Systems Institute Inc., Cary, North Carolina.

Seal, H.L. 1964. Multivariate statistical analysis for biologists. Methuen, London.

Tatsuoka, M.M. 1972. Multivariate analysis. Wiley, New York.

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

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

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

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

the VBGF for weight becomes

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

*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:

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

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

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

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:

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

or

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

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

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

and in terms of a differential equation, the form:

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

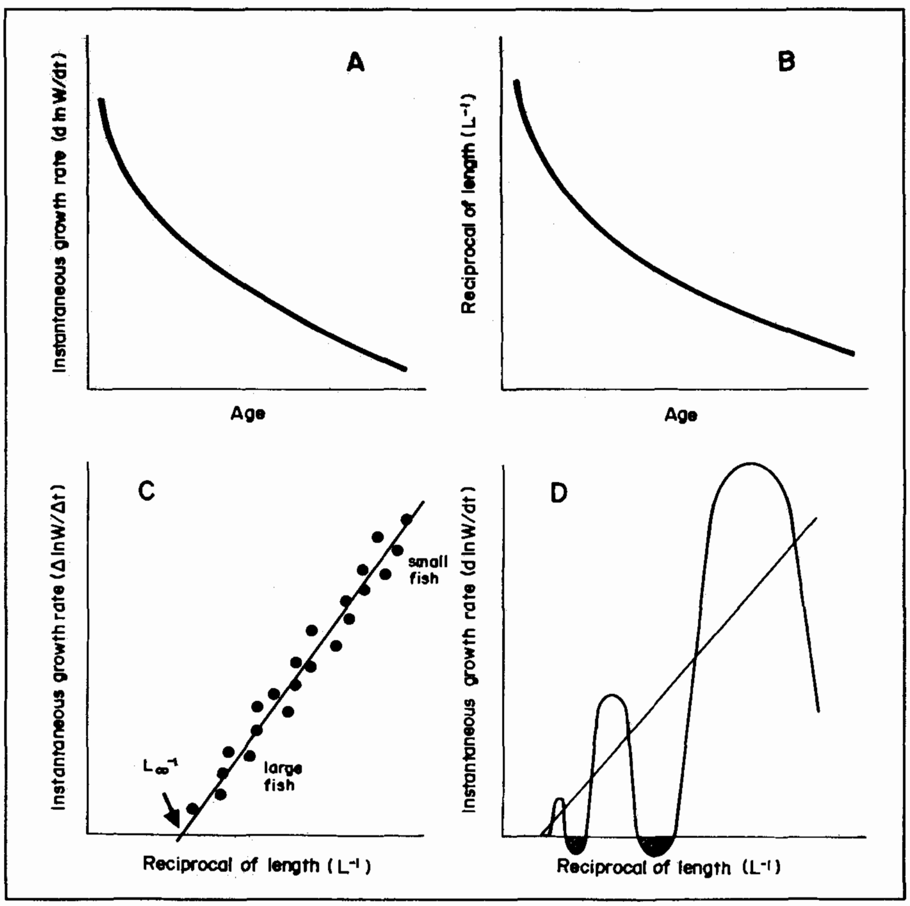


Fig. 1. Schematic representation of basic processes considered in new method (extended Bayley plot). A) The non-linear decrease of instantaneous growth rate in weight with age; B) The nonlinear decrease of the reciprocal of length with age; C) The Bayley plot. The relationship between instantaneous rate 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	$\Delta \ln W/\Delta t$	$1/\overline{L}$	$\overline{\text{VAR1}}$	...	$\overline{\text{VARn}}$
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):

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

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:

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

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:

$$y = a + b_1X_1 + b_2X_2 \quad \text{...21)}$$

the resulting path diagram is shown in Fig. 2.

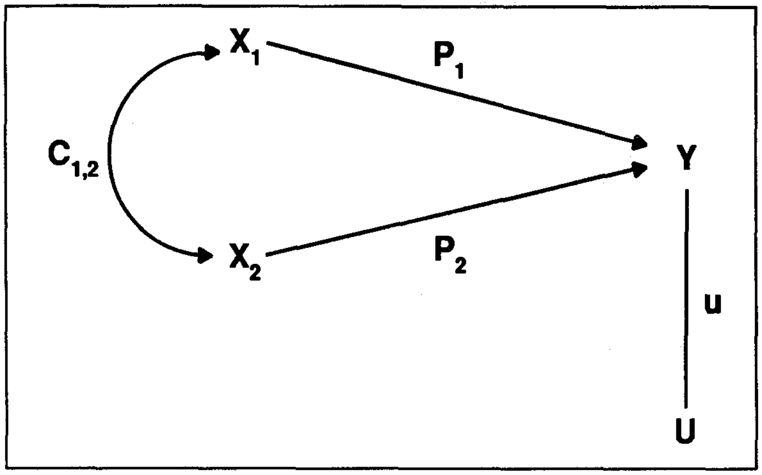


Fig. 2. Theoretical path diagram for a two-variable example. P_1 and P_2 = path coefficients (Beta coefficients of X_1 and X_2); $C_{1,2}$ = the correlation between X_1 and X_2; U = residual effect; u = amount of unexplained variance which is $1\text{-}R^2$.

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:

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

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.

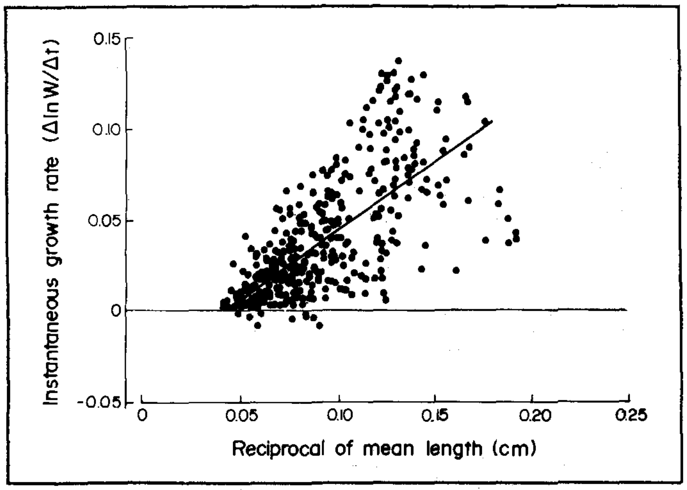


Fig. 3. Bayley plot of Nile tilapia grown in ICLARM-CLSU experiments at Muñoz, Philippines, in 1978-1981, n = 616. See text for regression equation. Note heteroskedasticity as discussed in text.

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:

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

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·ha^{-1}·d^{-1}	73	0-221
-0.03326	SQRT stocking density kg·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·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.

Särdson (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 with n = 193, r^2 = 0.601, SEE = 0.0170, P < 0.001, K = 0.0112day^{-1} and L_∞ = 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·ha^{-1}·d^{-1}	73	0-221
-0.03326	SQRT stocking density kg·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·day^{-1}	364	133-633
-0.06199			

...25)

with n = 193, R^2 = 0.663, SEE = 0.0158, P < 0.001, K = 0.0190day^{-1} and L_∞ = 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.

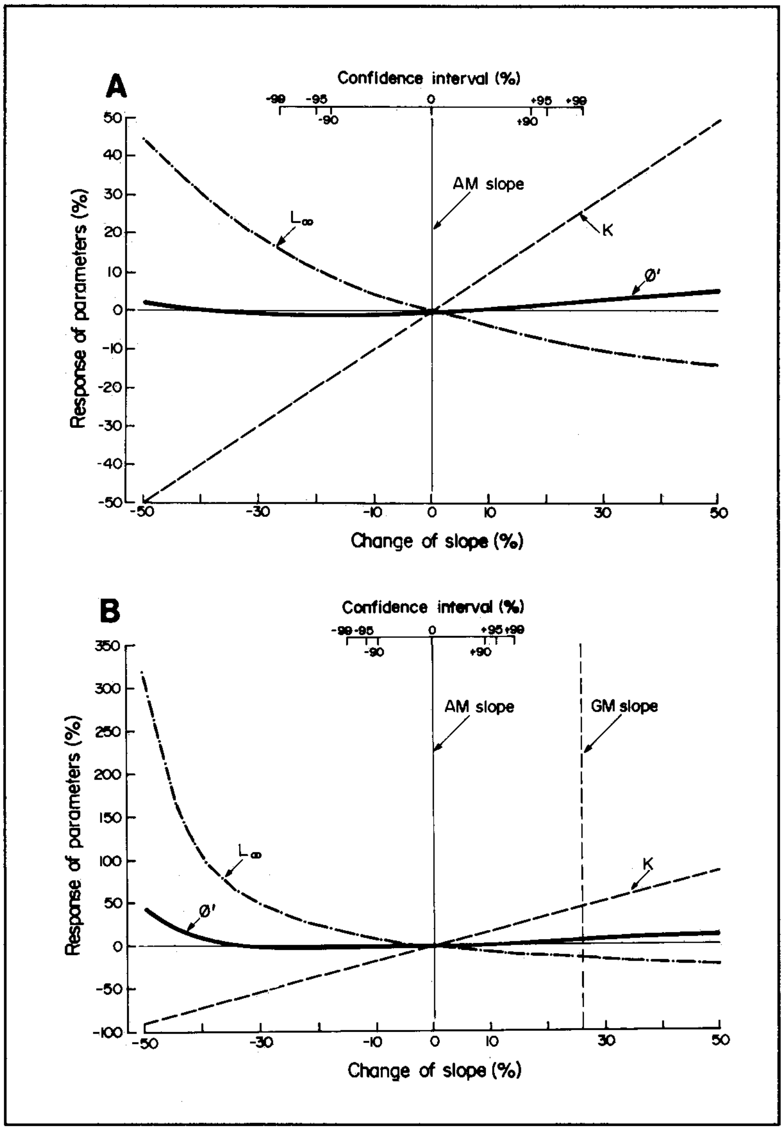


Fig. 4. Sensitivity analysis of A) ordinary Gulland-and-Holt plot, and B) ordinary Bayley plot, based on an AM regression on a random sample dataset (n = 198). In the Gulland-and-Holt plot, ϕ' compensates the effects of slope changes on growth parameters. In the Bayley plot, lower slopes have extreme 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%

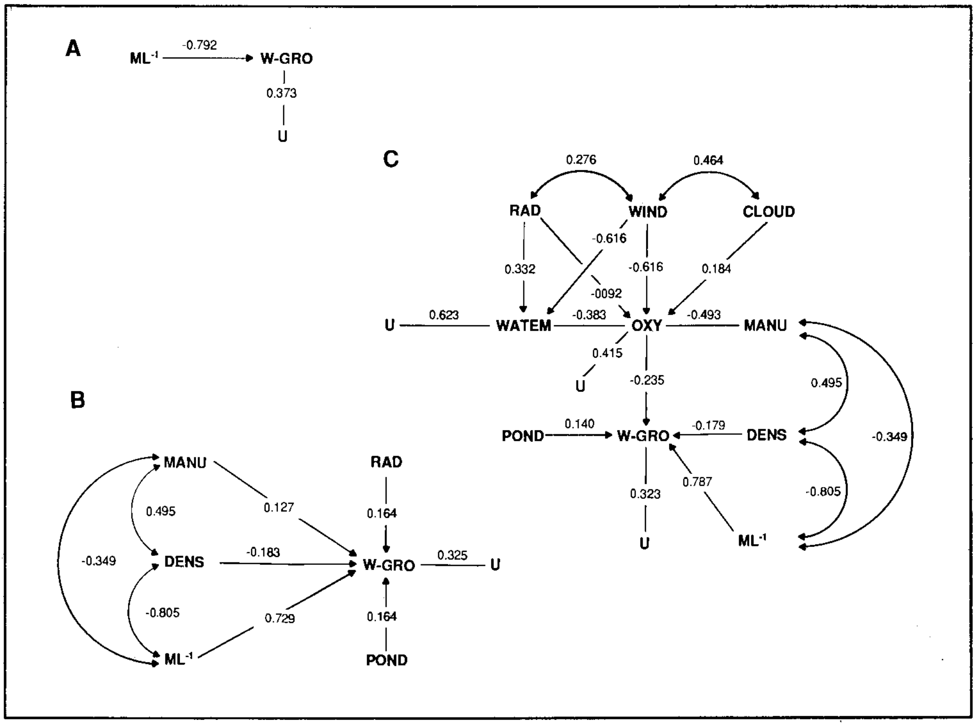


Fig. 5. Path diagrams for A) the ordinary Bayley plot, B) the extended Bayley plot with five predictor variables, and C) the extended Bayley plot with four direct predictor variables and five further explanatory variables. W-GRO = growth rate in weight; ML^{-1} = reciprocal of mean length; MANU = livestock 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.

Multiple Regression and Path Analysis of Nile Tilapia Growth in Integrated Livestock-Fish Culture in the Philippines*

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

Prein, M. 1993. Multiple regression and path analysis of Nile tilapia growth in integrated livestock-fish culture in the Philippines, p. 50-66. *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

A large dataset (64 variables, 713 cases) originating from four-year integrated livestock-fish farming experiments with mixed-sex populations of Nile tilapia (*Oreochromis niloticus* Fam: Cichlidae) conducted in Muñoz, Philippines, was analyzed with the multivariate statistical methods of multiple regression analysis based on the extended Gulland-and-Holt method and path analysis (or causal analysis).

Besides fish size, the methods identified four environmental or treatment variables governing Nile tilapia growth in the experimental ponds, i.e., the amount of pig, duck or chicken manure input, stocking density, pond area (m^2) and the amount of total solar radiation, which jointly explained 40% of the variance in Nile tilapia growth rate. The parameters of the von Bertalanffy growth function were estimated for a range of environmental conditions.

Procedures for the assembly of the dataset are outlined, together with a description of methods to compute meteorological variables. The results obtained are analyzed using path diagrams and sensitivity analysis.

Introduction

In line with the rationale to subject available historic data from earlier experiments to reanalysis with recent multivariate methods (Prein et al., this vol.), the present contribution demonstrates the application of two such methods to a large comprehensive set of aquaculture data from Muñoz, Philippines.

Materials and Methods

Description of the ICLARM/CLSU Experiments

Results from a four-year study at the Freshwater Aquaculture Center of the Central Luzon State University (CLSU) near Muñoz, Philippines, were used as a source for data on low-input aquaculture. The aim of the project was to develop economically viable small-scale integrated livestock-fish culture methods suitable for the Philippines, based entirely on livestock manure inputs, i.e., without fish feed. It was intended to produce fish in 3-6 months at a marketable size of 60 to 150 g. Details of the experimental design and preliminary results are given in Hopkins and Cruz (1980, 1982), Cruz and Shehadeh (1980), Hopkins et al. (1981, 1982, 1983) and PCARRD (1982). Only a general overview is given here.

Experiment Design

One hundred sixteen growth experiments in 18 experiment groups were conducted from August 1979 to June 1981 in 24 backyard-size ponds (12 of 0.04 ha and 12 of 0.1 ha, with average depths of 0.7 to 0.9 m). Treatments were always duplicated or triplicated.

Ponds were stocked in a polyculture of 85% Nile tilapia (*Oreochromis niloticus*) as the main culture crop, 14% common carp (*Cyprinus carpio*) as a bottom stirrer and 1% predator, either snakehead (*Ophicephalus striatus* = *Channa striata*) or Thai catfish (*Clarias batrachus*) (Table 1). Nile tilapia were of the "Ghana" strain, introduced via Dor Fish and Aquaculture Research Station in Israel (Pullin 1988). The design anticipated tilapia reproduction since mixed sexes were stocked. Since fry were competitors for space and food, a predator was stocked to limit fry biomass (Hopkins et al. 1982). Overall stocking densities were either 10,000 or 20,000 fish/ha. The average size of tilapia at stocking was usually 2.5 cm.

Nutrient inputs to the ponds were either inorganic fertilizer (IF) or fresh manure from pigs, ducks or chickens kept in stalls on the pond dikes. The manure from pigs and ducks were washed into the ponds during the daily cleaning of the pens. Chicken manure was collected from the

*ICLARM Contribution No. 729

Table 1. Summary of treatments, livestock and fish stocking rates in the ICLARM-CLSU experiments (adapted from Hopkins and Cruz 1982).

Density[a] (animals per hectare)		
Pigs	40, 60, 80, 100, 120, 140	
Ducks	750, 1,000, 1,250, 1,500	
Chickens	250, 500, 750, 1,000, 3,000, 5,000, 7,500, 10,000	
Inorganic fertilizers	3.6 (16-20-0; N-P-K)	
Nile tilapia	17,000	(85% of 20,000 fish·ha^{-1})
Nile tilapia	8,500	(85% of 10,000 fish·ha^{-1})
Common carp	2,800	(14% of 20,000 fish·ha^{-1})
Common carp	1,400	(14% of 10,000 fish·ha^{-1})
Predator[b]	200	(1% of 20,000 fish·ha^{-1})
Predator[b]	100	(1% of 10,000 fish·ha^{-1})

[a]All in pieces, except inorganic fertilizer, which is expressed in kg·day^{-1}.
[b]Thai catfish or snakehead.

stalls and applied to the ponds three times weekly. No additional feed were given, nor was any aeration applied.

The location guaranteed a year-round water temperature of 20 to 30°C. Natural hazards occurring were typhoons causing flooding of ponds and/or loss of livestock.

The experiment duration was set to match the period of 90 days which tilapias take to attain the local market size of approximately 60 g under the given conditions. Since pigs and Peking ducks reach their market size in 180 days, two fish culture cycles were serially combined with one pig or duck growing cycle. Chickens were either grown in short cycles of 15 days or for the entire growing periods of 45 days. Various strategies were chosen to maintain more or less constant chicken manure inflow into the ponds in the desired amounts [see Hopkins and Cruz (1982) for details]. Overall, the amount of manure applied to individual ponds was regulated through the number and size of livestock animals in the stall above each pond.

Data Collection

Fish size data were obtained by sampling the ponds at stocking, during the growing period at approximately biweekly intervals and at harvest. The total length of 20 to 50 individual tilapia was measured for each sampling and pond. During the first experiments, the individual weights of these fish were also recorded to establish a length-weight relationship. In later experiments only bulk weights were taken for large tilapia, tilapia recruits, carp and predators, respectively.

Average daily manure input was determined on a dry weight basis and were converted into chemical constituents using the conversion factors given in Hopkins and Cruz (1982).

Water quality parameters measured in the early morning (AM) were water temperature and dissolved oxygen. The following other measurements were made, although these were not performed continuously during the entire project period. At mid-morning (0900 to 1100 hours) pH and total ammonia (NH_3-NH_4) were determined. Alkalinity, conductivity, nitrite, nitrate and phosphate were sporadically measured.

Measurements referring to pond biology were Secchi disk visibility, primary production and plankton counts (106 µm and 38 µm mesh) (Hopkins and Cruz 1982). These parameters however were not recorded over the entire project period.

Meteorological data were obtained from a weather station (CLSU/PAGASA Agromat Station) situated 1 km away from the project site. Variables recorded were air temperature (minimum/maximum), solar radiation, bright sunshine duration, rainfall, wind (cumulative run), wind (direction), humidity and evaporation.

All nonfish data were averaged and expressed as mean values per day over the respective interval between fish sampling dates. This resulted in up to 11 intervals per experiment. All procedures that refer to data collection and primary processing (i.e., averaging of interval data) were conducted by project personnel in the Philippines.

A subset of the raw data of the project pertaining to 25 variables was published by Hopkins and Cruz (1982) in their final project report. These and data on 14 further variables were kindly supplied to the author on computer tape (EBCDIC format) by Dr. Kevin Hopkins.

In the course of preliminary analyses (Prein 1985), it was found that many of the data points and several variables had either missing values or had not been entered into the file supplied on tape. Under the auspices of the BMZ/Israel/ICLARM/GIARA 86-1/2 project a visit was made to ICLARM headquarters in Manila from 20 January to 16 May 1987 to retrieve as much information as possible from the original raw data recording sheets. This was complemented by visits to the experimental site at CLSU in Muñoz. The ponds, livestock and poultry stalls and laboratories were visited twice, in March 1983 and April 1987, and personnel involved in the experiments were interviewed and additional raw data were retrieved.

Unfortunately, many variables remained incomplete due to lack of either individual raw data sheets or even entire experiments, leaving blanks to be accounted for in data preparation for the final analysis (Table 2).

FISH GROWTH AND PRODUCTION DATA:
O. NILOTICUS (GHANA STRAIN)

The experiments conducted at CLSU were initially designed for analysis with a multivariate method of the "extended Gulland-and-Holt" type. Therefore, individual fish lengths were measured at regular, short intervals. Also, during the interval, important environmental variables were recorded frequently and then averaged to represent the mean during the interval.

Table 2. Overview of available data from ICLARM/CLSU experiments conducted in the Philippines (+ = data available; - = data missing; i = incomplete raw data; s = data summaries only; * = not included in dataset on tape but some raw data retrieved).

Experiment group	1	2	3	4	5	6	7*	8	9	10	11	12	13*	14	15*	16*	17*	18*
Fish data	+	+	+	+	+	+	+	+	+	s	-	-	i	+	s	+	-	-
Water quality	+	+	+	+	+	+	-	+	+	+	-	-	-	-	-	-	-	-

Further raw data retrieved for some ponds and experiments were: oxygen profiles; recordings of dissolved oxygen dynamics of up to 48 hours; the original four-year daily weather dataset from the meteorological station; proximate analysis of livestock feeds, livestock manure, and fish flesh; plankton counts; manure outputs of livestock; length-weight relationships and data for *Oreochromis niloticus* and *Ophicephalus striatus*; and measurements of individual fish lengths at the sampling events.

Fish production during a culture period usually depends strongly on stocking size and density. A previous analysis of the data (Prein 1985) failed to detect any density effects, probably due to an inappropriate transformation of the relevant variable and/or the masking effect of unexplained variance (see below).

Treatment of Data

The computer tape containing the ICLARM/CLSU experiment data was read at the Computer Center of Kiel University and stored to disk. Here, preliminary editing and analyses were performed which involved thorough error checking and missing value replacement procedures.

LENGTH-WEIGHT RELATIONSHIP, CONDITION FACTOR

The length-weight relationships were recalculated from the supplied raw data from CLSU with both length and weight as predicted variables to enable comparison with other published length-weight relationships from wild and aquaculture populations.

Data Handling and Processing

Central to all data handling and editing procedures was the establishment of several small spreadsheets on microcomputers. From these raw data files the required average values per growth interval were computed and inserted into the main datafile. Regression analyses were performed with the SPSS software package (Norusis 1985). The 95% significance level was used for all tests. The final dataset of the ICLARM-CLSU experiments contains 64 variables in 713 cases (file format: LOTUS 1-2-3, Ver. 2.1, MS-DOS; filename: PHILALL.WK1, data disk No. 1, described in Appendix II).

DETECTION OF MISSING VALUES AND DATA ENTRY ERRORS

To detect missing values in the datasets, all variables were printed out as datalists and each missing point was recorded. Errors from data entry were discovered through plotting the variables over the duration of the culture period (Tukey 1977). Errors in form of outliers, i.e., values deviating considerably from the general data range of the variables, were found most easily.

The intended methods for data analysis require a complete data matrix (Weisberg 1980; Norusis 1985). With just one missing value in a case containing several other variables, the entire information contained in that case would be omitted from the analysis. Instead of losing the information in an entire case, it is therefore better to fill in the missing values with approximations of known precision.

Fish size is used to compute growth rate which is the dependent variable in these analyses. It is not valid to fill in missing values for the dependent variable. These values would have no meaning in the regression procedure, since they are not part of the sample population. Therefore, no attempt was made to fill in missing values in fish size. As a consequence, some cases had to be omitted to establish the final and complete data matrix.

Environmental Variables

One aim of this study is to explain growth variations due to the effect of environmental variables. Therefore, it was necessary to obtain representative values of candidate variables during the culture periods. If there are missing values or if the variables were not measured at all, certain methods can be devised to fill in blanks in a data matrix or even generate entire new, unmeasured variables through computational models (Weisberg 1980). In the former case such a procedure should be indispensable because the analysis methods used require complete data matrices without missing values. A missing value in a case containing several other variables would lead to the loss of the entire information contained in those variables. Therefore it is worth the extra effort to find methods to fill in the blanks with values approximately representing those missing. Methods to do so are mostly case-specific. In the present case several different approaches had to be taken to replace missing values. They are described in detail below for the individual variables.

Wholly unmeasured variables are more difficult to recreate if precise functional relationships are not already known (such as for un-ionized ammonia from pH, temperature and total ammonia, or for per cent oxygen saturation from temperature, salinity and oxygen content). If the variables are regarded as very important, these may be derived from empirical relationships of known precision. The specific methods applied here are given below in detail.

Meteorological Variables

A number of points were missing in the meteorological variables. Methods which are described below were developed to fill in missing values.

The nearest station to the project site (15.68°N, 120.91°E), for which a climatic diagram was available, was Cabanatuan, located 25 km south of Muñoz in Central Luzon. The climate of the area is tropical and humid, with two distinct seasons: a four-month dry winter season from January to April and one rainy summer season from June to December. At an altitude of 30 m above sea level the average annual temperature is 27.6°C and the average annual amount of precipitation is 1,995 mm. Average monthly air temperatures range between 25 and 30°C, with the hottest month in May and the coolest in December. Average monthly rainfall ranges from zero in February to 300 mm in August.

Solar Radiation

Solar radiation is an important variable governing the productivity of terrestrial and aquatic ecosystems. Due to the tilt angle of the earth's axis relative to the sun, the possible daily amount of solar radiation reaching the top of the atmosphere above a given location is unevenly distributed around the globe (Fig. 1), also causing seasonal differences between northern and southern hemispheres. Depending on the location (latitude) the annual distribution pattern may differ considerably and can have direct consequences on fishpond productivity (Brock 1981).

The data from the meteorological station (CLSU-PAGASA Agromat) contained several missing values that had to be filled in. Besides solar radiation, the amount of bright sunshine in minutes per day also contained blanks. For the calculation of the possible daily amount of bright sunshine and solar radiation on the top of the atmosphere, a computer program was written, ('LIGHT') based on astronomical and meteorological equations (Prein and Gayanilo 1992).*

---

*Available on a 3.5" diskette for IBM compatible personal computers from the ICLARM Software Project, MC P.O. Box 2631, 0718 Makati, Metro Manila, Philippines.

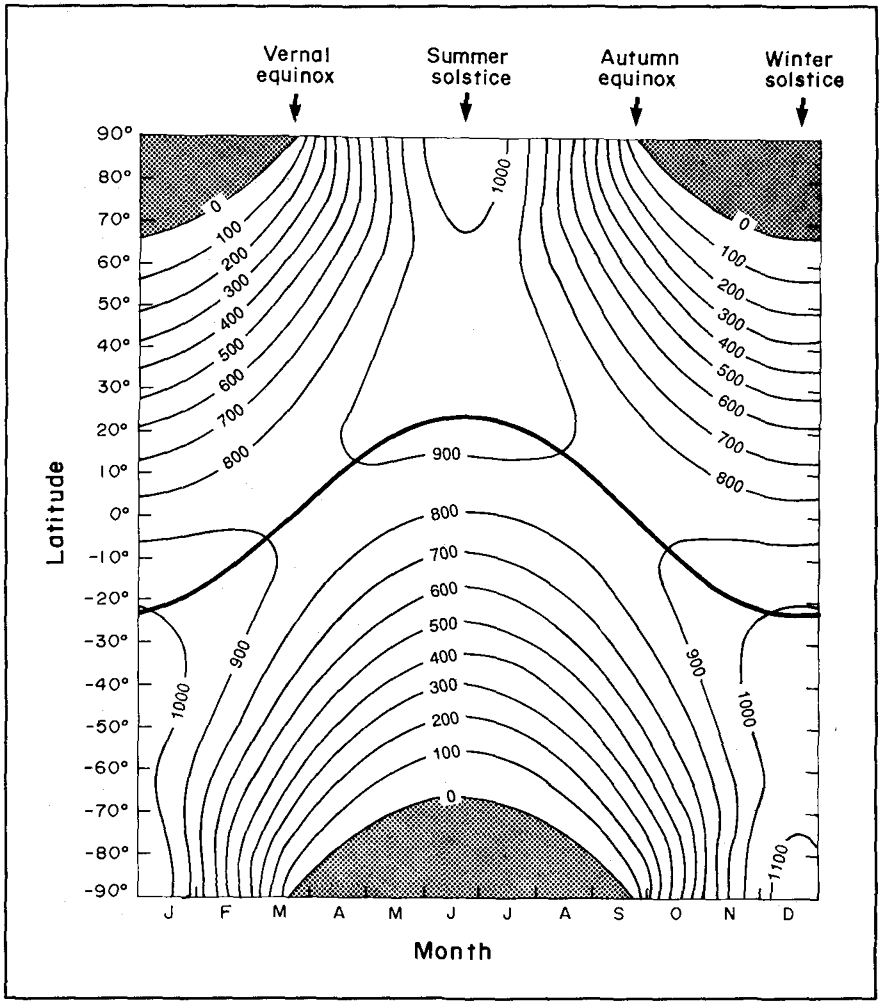


Fig. 1. Leighly chart: solid curves represent total daily solar radiation on a horizontal surface at the top of the the atmosphere, measured in cal·cm^{-2}. Shaded areas represent regions of continuous darkness while the thick solid line represents the declination of the sun (modified from List 1963).

As radiation passes through the atmosphere, it is attenuated due to the atmosphere's physical properties and condition. For completion of radiation data two procedures were tried. The first was to apply an existing relationship known in the literature as the Ångström-Prescott equation (List 1963; Coulson 1975). This relates the amount of radiation on the earth's surface (QGR) to the calculated solar flux on the top of the atmosphere (QT) as a function of the transmissivity of the atmosphere.

Transmissivity can be expressed by the ratio of available (n) to possible (N) bright sunshine per day:

$$QGR = QT(a - b \cdot n/N) \qquad ...1)$$

Rearranging (1) gives a simple linear regression of the form:

$$QGR/QT = a - b \cdot n/N \qquad ...2)$$

The constant (a) is known as the "atmospheric transfer coefficient". Both constants (a) and (b) are known as Ångström coefficients. These have been determined for many locations on the globe and characterize the prevailing type of climate (Lof et al. 1966; Duffie and Beckman 1980).

The second method used here was to derive an empirical regression equation. Relationship (2) was extended to a multiple regression form for which several combinations of variables were tried. This resulted in a better fit than that achieved by the ordinary Ångström-Prescott equation. It should be noted that these missing value treatments for the Philippine dataset apply to averages of intervals and not to individual daily values.

Computer Program for Calculation of Daily Solar Radiation

Relationships governing daylength and amount of radiative insolation are known in astronomy and meteorology. Bright sunshine on the surface of the earth is a value smaller than the theoretically possible amount of sunshine, due to sky haziness at dawn and dusk, and to cloud covering. The theoretical daylength varies according to the time of the year and can be calculated with the common equation for solar elevation (Milankovitch 1930):

$$\cos Z = \sin LAT \cdot \sin D + \cos LAT \cdot \cos D \cdot \cos h \quad \text{...3)}$$

where

Z	=	solar zenith angle
LAT	=	latitude of the location, in degrees
D	=	solar declination in radians
h	=	hour angle in degrees

At sunrise or sunset, cos Z = 0 and h = H, which is the half-day length in degrees. It follows (Sellers 1965) that:

$$\cos H = \tan LAT \cdot \tan D \quad \text{...4)}$$

and

$$H = \arccos(- \tan LAT \cdot \tan D) \quad \text{...5)}$$

Since one hour equals 15 degrees, the astronomical half-day length for any given location on the globe can be calculated. Doubling the value produces the possible time of bright sunshine for any particular day of the year.

The solar declination (D) depends only on the day of the year, for which an exact approximation is given by Spencer (1971):

$$\begin{aligned} D = & -0.399912 \cdot \cos A \\ & +0.070257 \cdot \sin A \\ & -0.006758 \cdot \cos 2A \\ & +0.000907 \cdot \sin 2A \\ & -0.002697 \cdot \cos 3A \\ & +0.001480 \cdot \sin 3A \quad \text{...6)} \end{aligned}$$

The angle A is obtained by converting the day of the year J (1 January = 0, 31 December = 364) to radians:

$$A = J \cdot (2 \cdot \pi / 365) \quad \text{...7)}$$

Through multiple regression analysis of the available weather data, an empirical relationship was developed to predict the number of minutes of bright sunshine per day to fill in the missing values.

Radiation is measured as the amount of energy per unit area, here in terms of Langleys (ly), which is equal to 1 gCalorie/cm^{-2}. The time unit used here is the whole day. The incident radiation on the ground is dependent upon the possible amount reaching the atmosphere from space (QT). This amount fluctuates over the year depending upon distance to the sun and the earth's relative position. The amount of radiation available on top of the atmosphere can be calculated with some of the constants also used for sunshine prediction. The following equation integrates the amount over the whole day (List 1963; Sellers 1965):

$$QT = 1440/\pi \cdot S \cdot CD \cdot (H - \tan H) \cdot \sin \cdot LAT \cdot \sin D \quad \text{...8)}$$

where

QT	=	daily total solar radiation incident on a horizontal surface on top of the atmosphere
S	=	solar constant (=1.94 ly·min^{-1})
CD	=	distance correction between the earth and the sun
H	=	half-day length in degrees
D	=	solar declination in radians

The distance correction between the earth and the sun lies between 0.97 and 1.03 depending on the time of the year and can be calculated (Paltridge and Platt 1976) with:

$$\begin{aligned} CD = 1.00011 & + 0.034221 \cdot \cos A \\ & + 0.001280 \cdot \sin A \\ & + 0.000719 \cdot \cos 2A \\ & + 0.000077 \cdot \sin 2A \quad \text{...9)} \end{aligned}$$

These computations are also included in the LIGHT program (Prein and Gayanilo 1992). With this program, the theoretical total solar radiation values reaching the top of the atmosphere over the respective locations were calculated. These were compared with the actually measured values for Muñoz, Philippines (Fig. 2). This graph shows the generally observed reduction of solar radiation by 30 to 40% (for clear skies) as it passes through the atmosphere.

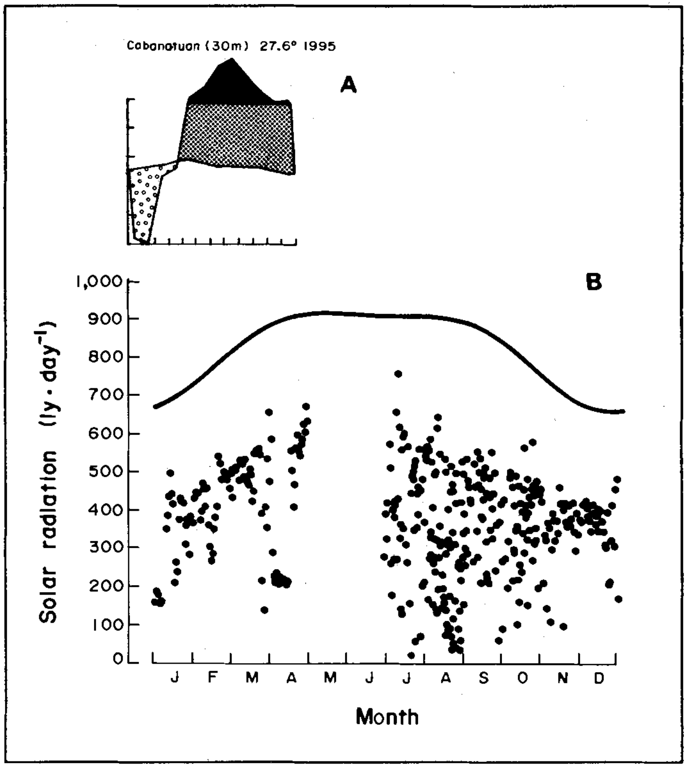


Fig. 2. Meteorological and climatological characteristics of experimental site in Central Luzon, Nueva Ecija, Philippines. A) Climatic diagram of meteorological station at Cabanatuan, from Walter and Lieth (1969) as described in Prein (this vol.). B) Theoretical values (line) of total solar radiation reaching the top of the atmosphere over Muñoz, compared to actually measured values (full dots) in 1978 to 1981.

Air Temperature

The dataset contained water temperature data, so air temperature from the meteorological station did not require further processing here for the analysis.

Wind and Cloud Cover

In the ICLARM-CLSU dataset, other available meteorological variables were evaporation, cloud cover, wind velocity and direction, and humidity. For the latter three, daily values were partly available as 0800 and 1400 hours readings. Cloud cover was recorded in decals, i.e., as a fraction of 10.

Water Quality Variables

WATER TEMPERATURE

For cases where water temperature data were unavailable, Hopkins and Cruz (1982) used an empirically derived equation to compute missing values of early morning water temperature:

$$TW = -1.6675 + 1.055\ TA \qquad ...10)$$

with $n = 47$ and $r^2 = 0.634$

where
TW = early morning water temperature (°C)
TA = average of maximum and minimum air temperature read at 0800 hours the preceding day

Therefore the ICLARM-CLSU dataset had no missing values of water temperature (Fig. 3). In the course of the exploratory data analysis, strong correlations with other variables were found. A series of multiple regression equations were tried, of which the following was the best predictive and explanatory model:

$$TW = 17.09398 + 0.01396\ QT - 0.02100\ WIND \qquad ...11)$$

with $n = 54$, $R^2 = 0.769$, $P < 0.001$

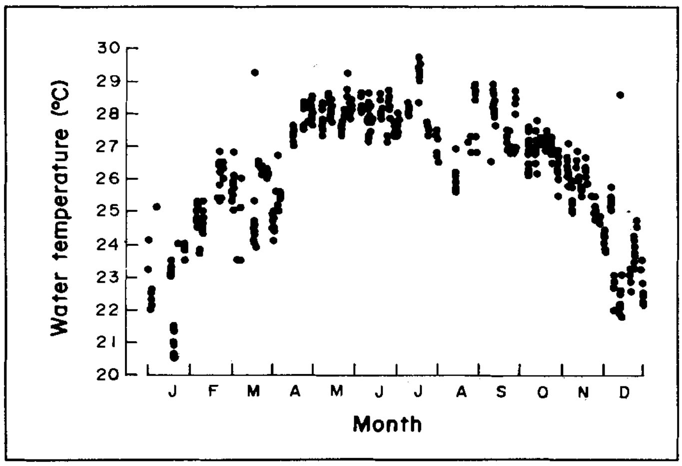


Fig. 3. Seasonal variation of daily early morning (AM) water temperature in experimental ponds at Muñoz, Philippines. Data from 1978 to 1981.

This suggests that water temperature increases with solar radiation but is decreased by wind-driven evaporative cooling.

DISSOLVED OXYGEN

Of the 713 cases in the final dataset, 67 had missing values for early morning dissolved oxygen. About 15 different runs using different multiple regressions were tried to predict missing cases (Prein 1985) for early morning dissolved oxygen content (AMDO). The equation selected contains 10 variables:

$$
\begin{aligned}
\text{AMDO} = 10.48084 &+ 0.03859 \cdot (\text{MANURE BOD}_5)^2 \\
&- 1.87377 \cdot \text{POND AGE} \\
&- 0.00071 \cdot \text{'30-DAY-MANURE'} \\
&- 0.00343 \cdot \text{RADIATION} \\
&- 0.18191 \cdot \text{WATEMP} \\
&- 0.85029 \cdot \text{DUMMY CHICKEN (MANURE TYPE)} \\
&+ 0.05789 \cdot \text{RAIN} \\
&- 0.00008 \cdot \text{CUMULATIVE MANURE} \\
&+ 0.00967 \cdot \text{WIND} \\
&- 0.0000001 \cdot \text{WIND}^3 \qquad \text{...12)}
\end{aligned}
$$

with $n = 646$, $R^2 = 0.665$, $P << 0.001$.

Here, the aim was not to arrive at a biologically optimal equation but to determine a predictive model with a good fit to the data in order to fill in the 67 missing values.

Once all missing oxygen values were filled in, % saturation values were calculated. Their distribution over the year (Fig. 4) shows a marked seasonal pattern. Critical saturation values may exist all year-round, but the dry season has prevailing levels of usually under 40% dissolved oxygen saturation. In the course of the rainy season the situation improves, yet still bearing the danger of adverse oxygen regimes. The overall trend of oxygen content during experiments (Fig. 5) is a gradual decrease from a saturation of 37% at the beginning to 10% at the end of the longest experiments although considerable fluctuations were observed. The shorter experiments have wide variations from near zero to over 90% of saturation, whereas in the second half of the long experiments early morning oxygen conditions below 30% prevailed.

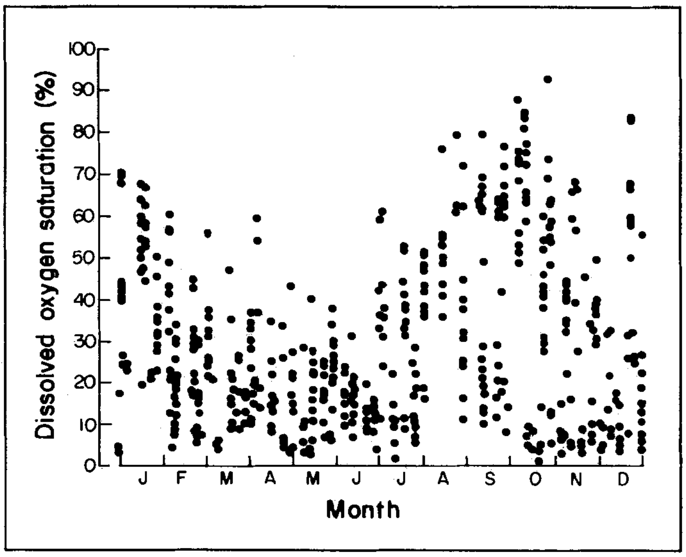


Fig. 4. Seasonal variation of early morning dissolved oxygen saturation in experimental ponds at Muñoz, Philippines. Data from 1978 to 1981. Largest variation is observed during the rainy season (July to October). Generally low values are found at the end of the dry season (May to June).

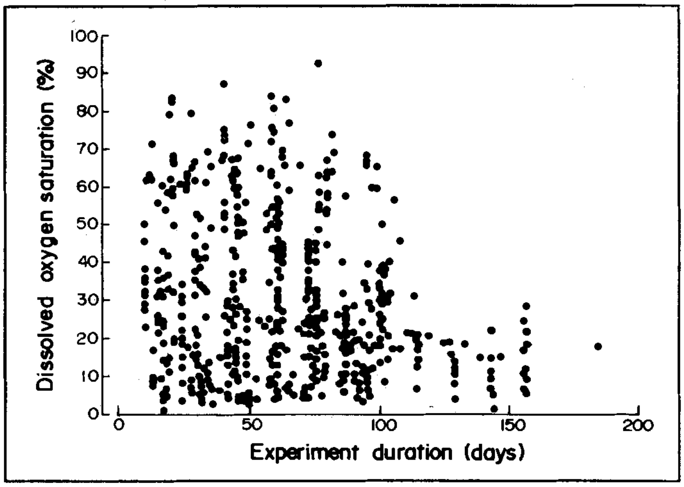


Fig. 5. Distribution of early morning dissolved oxygen (AMDO) saturation with experiment duration in experimental ponds at Muñoz, Philippines. With longer duration the AMDO is reduced. Data from 1978 to 1981.

The percentage of days in an interval in which the early morning dissolved oxygen values were under 2 mg·l^{-1} (OXY_2) was predicted using:

$$
\begin{aligned}
\text{OXY}_2 = 78.37863 &- 18.49817 \cdot \text{AM DISSOLVED OXYGEN} \\
&- 0.02691 \cdot \text{BRIGHT SUNSHINE} \\
&+ 0.03464 \cdot \text{SOLAR RADIATION} \\
&+ 19.55743 \cdot \text{DUMMY INORGANIC FERTILIZER} \\
&+ 0.39645 \cdot \text{MANURE BOD}_5 \text{ CONTENT} \\
&- 0.34398 \cdot (\text{MANURE P}_2\text{O}_5 \text{ CONTENT})^2 \\
&+ 8.18144 \cdot \text{DUMMY CHICKEN (MANURE TYPE)} \\
&+ 0.48848 \cdot \text{MANURE FIBER CONTENT} \qquad \text{...13)}
\end{aligned}
$$

with $n = 646$, $R^2 = 0.836$, $P << 0.001$.

Here the dependent variable was transformed: due to the extreme U-distribution of the data, it was necessary to apply a variance-stabilizing transformation because all values did not lie within the bounds of 30 and 70% (Snedecor and Cochran 1980; Weber 1980; Sachs 1984). The per cent values (days during which DO was below n $mgO_2 \cdot l^{-1}$) were converted to fraction data and their square roots were arcsine transformed. The procedures for OXY_1 and $OXY_{0.5}$ were the same, though resulting in different equations (see below). Percentages of days in which early morning dissolved oxygen was below 1.0 $mg \cdot l^{-1}$ (OXY_1) were predicted using:

$$
\begin{aligned}
OXY_1 = 64.30304 & - 13.55003 \cdot \text{AM DISSOLVED OXYGEN} \\
& + 2.42163 \cdot \text{MANURE BOD}_5 \text{ CONTENT} \\
& - 0.02058 \cdot \text{BRIGHT SUNSHINE} \\
& - 0.36646 \cdot (\text{MANURE P}_2\text{O}_5 \text{ CONTENT})^2 \\
& + 18.58377 \cdot \text{DUMMY INORG.FERTIL.} \\
& + 26.93918 \cdot \text{DUMMY CHICKEN (MANURE TYPE)} \\
& + 0.01373 \cdot \text{POND SIZE} \\
& - 17.05879 \cdot \text{POND AGE} \\
& - 0.29010 \cdot \text{DAYS SINCE EXPERIMENT BEGIN} \qquad \textbf{...14)}
\end{aligned}
$$

with n = 646, $R^2 = 0.7178$, $P << 0.001$.

The predictive equation for the percentage of days in which early morning dissolved oxygen was below 0.5 $mg \cdot l^{-1}$ ($OXY_{0.5}$) includes 13 significant variables and takes the form:

$$
\begin{aligned}
OXY_{0.5} = 17.93177 & + 3.07459 \cdot \text{MANURE BOD}_5 \text{ CONTENT} \\
& - 7.11624 \cdot \text{AM DISSOLVED OXYGEN} \\
& - 0.12299 \cdot \text{MANURE DRY MATTER CONTENT} \\
& - 2.21578 \cdot \text{FISH} \\
& - 0.02399 \cdot \text{BRIGHT SUNSHINE} \\
& + 0.03562 \cdot \text{SOLAR RADIATION} \\
& - 0.00231 \cdot \text{CUMULATIVE MANURE} \\
& + 0.04221 \cdot \text{POND SIZE} \\
& - 32.74644 \cdot \text{DUMMY PIG (MANURE TYPE)} \\
& - 0.00367 \cdot \text{'45-DAY-MANURE'} \\
& - 5.83230 \cdot \text{MANURE P}_2\text{O}_5 \text{ CONTENT} \\
& - 16.32583 \cdot \text{DUMMY INORG.FERTIL.} \\
& + 0.43539 \cdot \text{RAINFALL} \qquad \textbf{...15)}
\end{aligned}
$$

with n = 646, $R^2 = 0.5952$, $P << 0.001$.

where the variable FISH is the sum of the log-transformed standing stocks (in $kg \cdot ha^{-1}$) of Nile tilapia, tilapia recruits, carp and predators.

Generally, for the above three regressions, as the level of oxygen to be predicted decreased, it became more difficult to find a predictive relationship. The number of variables had to be increased, yet the overall coefficient of determination for each equation obtained decreased with the predicted variables $OXY_{0.5}$ to OXY_2.

pH AND AMMONIA

In ponds receiving high amounts of nutrient inputs (inorganic fertilizer, manure, feed), high levels of total ammonia in the pond water may develop, since nitrification rates are slow. Intensive plankton blooms associated with high afternoon pH values cause the transformation of ammonia from the ionized form (NH_4^+) to the un-ionized form (NH_3) which is toxic to fish. Sublethal levels have growth-depressing effects while higher levels may lead to massive fish kills in a few hours. In the ICLARM-CLSU dataset, 33% of the intervals lacked an entry for ammonia. It was attempted here to fill in these missing values with such derived from empirical relationships in order to gain another important variable for consideration in the analysis; however, no significant and/or plausible relationship could be identified, and this plan had to be abandoned.

Length-Weight Relationship of Nile Tilapia

The length-weight data from the ICLARM-CLSU experiments were recomputed in form of a functional regression:

$$TL = \sqrt[3.258]{\frac{W}{0.01065}} \qquad \textbf{...16)}$$

with n = 612, $r^2 = 0.973$, SEE = 0.215, $P < 0.001$, size range: 0.7 to 210.8 g, 4.3 to 22.0 cm

where W is in g and total length (TL) is in cm (Fig. 6).

Although proposed e.g. by Sprugel (1983), correction for log-transformation bias was not applied here, to allow direct comparisons with other length-weight relationships for Nile tilapias in which correction factors were not incorporated (Pauly et al., this vol.)

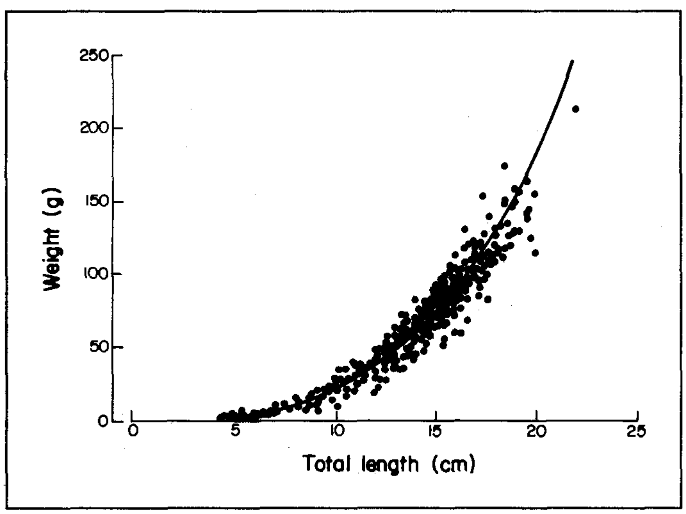


Fig. 6. Length-weight relationship of Nile tilapia grown in ICLARM-CLSU experiments in Muñoz, Philippines, in 1978 to 1981. n = 612. See text for regression equation.

Exploratory Data Analysis

At several stages of the data handling and model building process, techniques of exploratory data analysis (Mosteller and Tukey 1977; Tukey 1977; Daniel and Wood 1980; Weisberg 1980) were applied to ascertain the structure and behavior of the variables and detect interrelationships among them. Frequency distributions, data plots, multiple regressions and residuals plots were investigated, starting with the raw data up to the final completed dataset. The experience gained hereby and the knowledge of localized interdependencies was brought into the final process of model development and proved to be extremely valuable.

Multivariate Analysis Methods

The basic concepts and data requirements of the methods used here are described in Pauly et al. (this vol.) for the "extended Gulland-and-Holt method" and in Prein and Pauly (this vol.) for path analysis (or causal analysis). Discussions of the methods and an outline of sensitivity analysis are presented in Prein and Pauly (this vol.).

Least Median Squares Regression

Classical least squares regression (LS) is sensitive to outliers in the data, resulting in a distortion of the regression estimate. A single outlier can have such a strong influence (termed "leverage") that a regression equation may result which does not "correctly" represent the general trend in the rest of the data. In a different case, the data may be clustered in such a manner that ordinary least squares regression results in a totally "incorrect" regression line, e.g., a negative slope instead of a positive one. In both problem cases, the sum of the squared residuals, which is minimized in LS, causes aberrant values to have a strong influence on the regression estimate.

To overcome this problem, Rousseeuw (1987) suggested the method of least median squares regression (LMS), which is more robust towards contaminated data. This method uses the median of the squared residuals instead of their sum. LMSR may be regarded as an analytical tool or filter for the purpose of identifying influential outliers and finding the unbiased general trend in the data. Thereafter, ordinary least squares regression can be used for final model estimation. The method is available in form of a program for MS-DOS microcomputers called PROGRESS (Program for RObust reGRESSion; Rousseeuw and Leroy 1987).

Fish growth datasets usually contain several outliers. This is particularly the case if the data is originally in form of weights which are transformed to lengths. Mostly, these outliers are attributable to sampling errors. For their identification and primary analysis of the data, the PROGRESS software was used here. After deletion of the discovered outliers, the data were finally analyzed with the ordinary least squares regression procedure in the SPSS package (Norusis 1985).

Results

Extended Gulland-and-Holt Analysis

An ordinary Gulland-and-Holt plot of the entire ICLARM-CLSU data is given in Fig. 7. A random sample of 200 cases of this dataset results in the following equation for the Gulland-and-Holt plot:

$$\Delta L/\Delta t = 0.2528 - 9.9396 \text{ ML} \quad \ldots 17)$$

where ML is mean length, with $n = 200$, $r^2 = 0.284$, SEE = 0.0613, $P < 0.001$ and $K = 0.00994 \cdot \text{day}^{-1}$, $L_\infty = 25.4$ cm

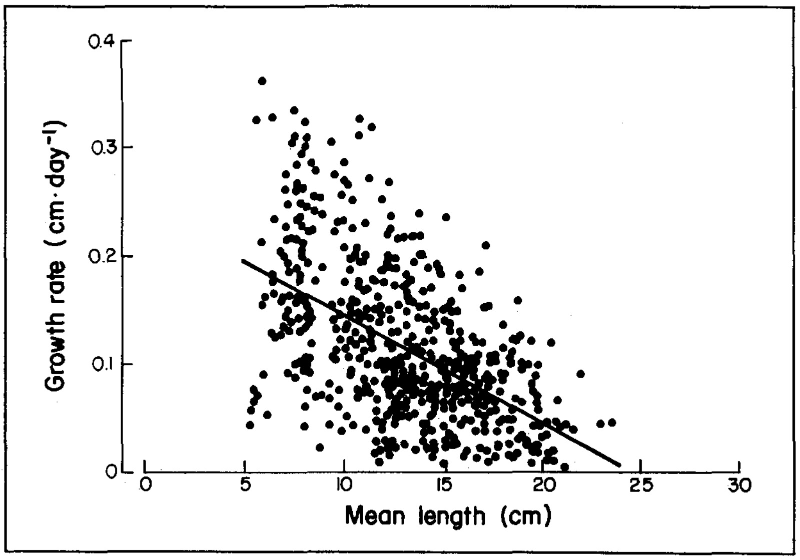


Fig. 7. Gulland-and-Holt plot of Nile tilapia grown in ICLARM-CLSU experiments in Muñoz, Philippines, in 1978 to 1981, n = 616. See text for regression equation.

With the extended Gulland-and-Holt plot, the following regression equation was obtained:

		mean	range
$\Delta L/\Delta t$ =			
$-6.516 \cdot 10^{-3}$	mean length cm	13.7	5.3-22.1
$+4.374 \cdot 10^{-4}$	manure input $kg \cdot ha^{-1} \cdot day^{-1}$	73	0-221
-0.16656	SQRT stocking density $kg \cdot m^{-3}$	0.36	0.08-0.8
$+5.966 \cdot 10^{-5}$	pond area m^2	811	400-1000
$+1.720 \cdot 10^{-4}$	solar radiation $ly \cdot day^{-1}$	364	133-633
+0.11745			

...18)

with n = 200, R^2 = 0.398, SEE = 0.0569, P < 0.001, K = 0.00652$\cdot day^{-1}$ and L_∞ = 30.8 (23.2 to 38.3) cm, where the latter values are computed for the mean, minimum and maximum values of the variables in the dataset.

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

		lower	upper
mean length	= 3.4 %	-0.0114	$-1.618 \cdot 10^{-3}$
manure input	= 5.8 %	$1.876 \cdot 10^{-4}$	$6.871 \cdot 10^{-4}$
stocking density	= 3.6 %	-0.28896	-0.04415
pond area	= 7.0 %	$2.894 \cdot 10^{-5}$	$9.038 \cdot 10^{-5}$
solar radiation	= 5.1 %	$6.688 \cdot 10^{-5}$	$2.771 \cdot 10^{-4}$
constant		0.06077	0.17413

TEST OF THE MODEL

To test the validity of the obtained model, a regression was determined, from the remaining part of the dataset (which was not used for model building), using the same variables. The obtained equation was not significantly different (at 95% level) from that determined with the sample dataset (based on comparisons of the confidence intervals of the individual regression coefficients). The signs of the independent variables were also the same.

As a further test of the model, the equation originally obtained was used to predict the growth rates at the end of the intervals in the rest of the data based on the parameters actually measured. In case of good approximation, a plot of the predicted vs. the actually measured growth rates should produce a relationship with a slope near unity and an intercept near zero (Fig. 8).

MODEL WITH POLYNOMIAL TERM

With the sample dataset of 200 cases, a further model was developed based on the hypothesis that an overload of manure has negative effects on fish growth by inducing an adverse pond environment:

		mean	range
$\Delta L/\Delta t$ =			
$-6.153 \cdot 10^{-3}$	mean length cm	13.7	5.3-22.1
-0.17327	SQRT stocking density $kg \cdot m^{-3}$	0.36	0.08-0.8
$+6.835 \cdot 10^{-3}$	water temperature °C	26.3	21.3-29.4
$+3.797 \cdot 10^{-4}$	wind run $km \cdot day^{-1}$	115.5	32-275
$+8.702 \cdot 10^{-4}$	manure input $kg \cdot ha^{-1} \cdot day^{-1}$	73	0-221
$-3.471 \cdot 10^{-6}$	manure input^2	5329	0-48841
$-6.541 \cdot 10^{-3}$			

...19)

with n = 195, R^2 = 0.390, SEE = 0.058, P < 0.001, K = 0.00615day^{-1} and L_∞ = 32.4 cm (mean values for variables).

Path Analysis

A path diagram for the ordinary Gulland-and-Holt plot is shown in Fig. 9. Here ML is the mean length, L-GRO is the length growth rate and U is the combined indicator for the amount of unexplained variance or residual effect. This shows that slightly more than 28% of the total variance in the data is explained by average fish size. The rest (72%) is due to other effects or variables not accounted for in the ordinary Gulland-and-Holt plot analysis. The sign of the path coefficient for

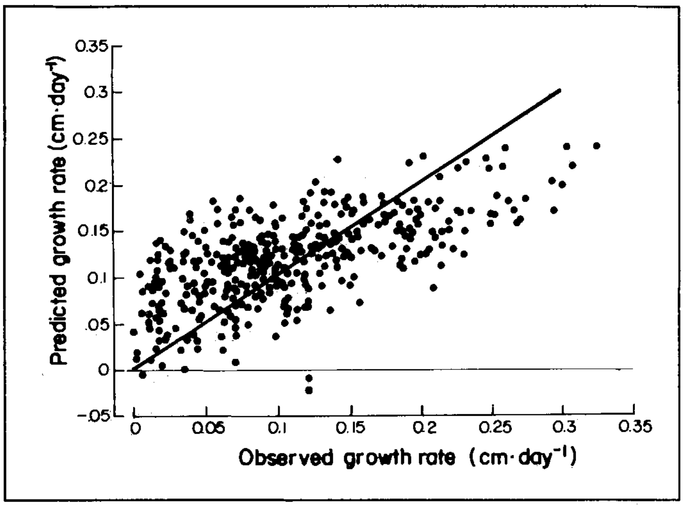


Fig. 8. Plot of predicted Nile tilapia growth rate vs. observed growth rate. Based on data and regression model for ICLARM-CLSU experiments in Muñoz, Philippines. n = 430. Line denotes a 1:1 relationship. A regression through the points gives: y = -0.00186+0.928x, r^2 = 0.410, SEE = 0.0511, P<0.001.

mean length is negative, describing a decrease in growth rate with increasing fish size.

A causal path diagram for the extended Gulland-and-Holt plot, as derived above, depicts the effect of five independent variables on Nile tilapia growth rate (Fig. 9). Here MANU is the manure input, DENS is the stocking density, RAD is the solar radiation and POND is the pond area.

The amount of total variance explained by the model is higher than in the ordinary Gulland-and-Holt plot (40%), because three treatment variables are significantly affecting 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.

After numerous trials, the causal path diagram in Fig. 9 was obtained, based on the extended Gulland-and-Holt plot and a reduced set of variables. 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 length is influenced by four variables directly, of which two are treatment variables. The variables

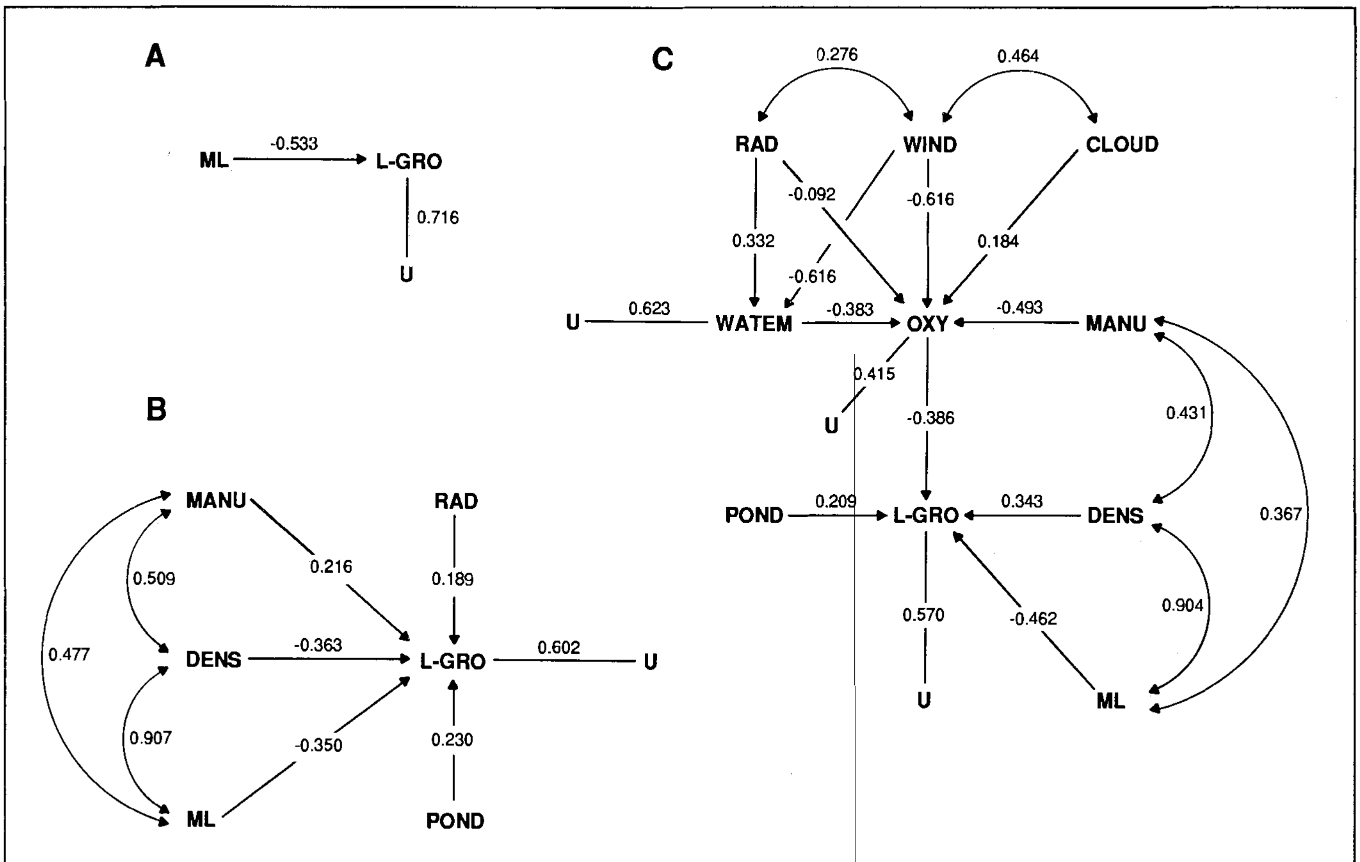


Fig. 9. Path diagrams for A) the ordinary Gulland-and-Holt plot, B) the extended Gulland-and-Holt plot with five predictor variables and C) the extended Gulland-and-Holt plot with four direct predictor variables and five further explanatory variables. L-GRO = growth rate in length ($cm{\cdot}day^{-1}$); ML = mean length (cm); DENS = square root of stocking density ($kg{\cdot}m^{-3}$); POND = pond area (m^2); MANU = manure input ($kg{\cdot}ha^{-1}{\cdot}day^{-1}$); OXY = early morning dissolved oxygen content (% saturation); WATEM = early morning water temperature (°C); RAD = solar radiation ($ly{\cdot}day^{-1}$); WIND = cumulative run of the wind ($km{\cdot}day^{-1}$); CLOUD = cloud cover (decals); U = residual effect or unexplained variance.

are mean length, stocking density, early morning oxygen saturation and pond area. Together these explain 43% of the variance in length growth rate. The percentage of total variation explained by the individual variables are: ML 6%, OXY 16%, DENS 4% and POND 7%.

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 cover) and one is an uncontrollable variable of the pond environment (water temperature). These variables explain 58% of the total variation in OXY. Their individual amount of explanation is: MAN 35%, TW 16%, RAD 1.5%, WIND 13% and CLOUD 5%. 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 cover 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 explain directly variance in Nile tilapia growth rate. In those models, the amount of explained variance was lower. For management purposes under field conditions, manure input and solar radiation are easier to handle than oxygen concentration. In the detailed path diagram, OXY was therefore incorporated as an intermediate variable between tilapia growth rate on one side and manure input and meteorological parameters on the other. In terms of path analysis, RAD, WIND, CLOUD form compound paths, contributing individually and in a combined manner, to the variation in length growth rate.

OTHER VARIABLES

Several other variables were tested for inclusion in the causal path models. These were either insignificant (at the 95% level) or had implausible signs. For example, Secchi disk visibility was neither predictable with any plausible combination of variables, nor was it plausible itself when tested for causal relationship with Nile tilapia growth rate. It had a positive sign in all cases.

Spurious correlations, i.e., significant correlations arrising by chance only were also found; the corresponding variables were excluded.

The main problem in design of path diagrams and regression model building was the fact that some variables were intercorrelated (i.e., multicollinearity). This was the case in mean length, stocking density (in various forms, e.g, $n \cdot ha^{-1}$, $kg \cdot ha^{-1}$, $kg \cdot m^{-3}$, for individual and all species) and manure input (as dry matter or as via its chemical components). In such cases, the variable which could be obtained most straightforwardly was chosen. This was done mainly in view of future applications of the models in the field.

Fig. 10 shows the results of a sensitivity analysis based on the regression equation obtained with the extended Gulland-and-Holt method and the ICLARM-CLSU data from the Philippines. Since the equation is a multiple linear regression, all effects are also linear. Solar radiation, pond area and manure loading rate show positive effects, while stocking density and mean length have negative effects. Mean length causes the strongest change in growth rate (max. 40%) when varied from -50% to +50% from its average value. Manure loading rate has, in the model derived here, the weakest effect on Nile tilapia growth rate.

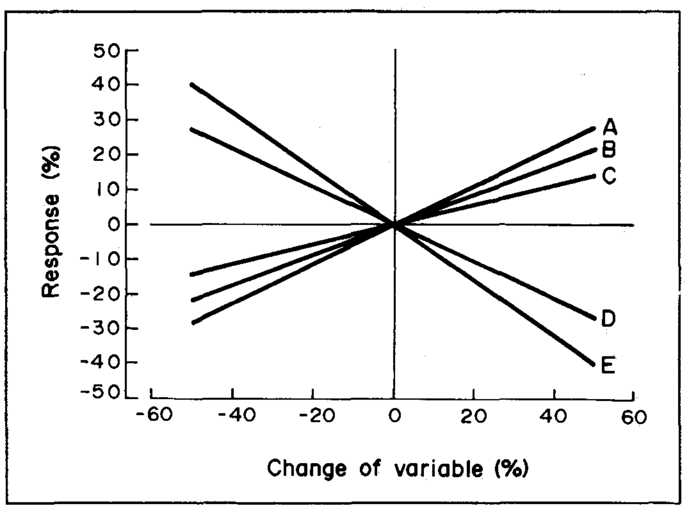


Fig. 10. Sensitivity analysis of the regression model for *Oreochromis niloticus* growth rate, developed with data from ICLARM-CLSU experiments in Muñoz, Philippines, based on the extended Gulland-and-Holt method. The % response from the average growth rate (ordinate) is shown as a result of % change from the average value of each independent variable (abscissa). Predictor variables: A = total solar radiation, B = pond area, C = manure loading rate, D = stocking density, E = mean length.

Discussion

The Regression Models and Path Diagrams

With all datasets, significant regression models could be derived based on the multivariate methods presented here and in all cases, length had a significant relationship, of the appropriate sign, with growth rate. Hence, the underlying von Bertalanffy growth function was validated by this study.

With the extended Gulland-and-Holt method, four independent variables, in addition to size itself, could be identified as significantly affecting Nile tilapia growth. These were manure loading rate, stocking density, pond area and solar radiation. In the experiments, livestock manure was the only source of external nutrient supply to the ponds. Manure loading rate was increased as the fish grew, with the consequence that both variables were correlated. In spite of this, the variation in experiment design was large enough to prevent high collinearity. In the latter case, this would have caused manure loading rate to be excluded from the equation.

Stocking density was varied at two densities in the experimental design. With fingerling-size tilapia, the density effects in terms of biomass per unit volume of water do not appear under normal pond culture conditions. As tilapia weight increases, density effects (e.g., territorial behavior, food limitation) become apparent and negatively affect tilapia growth (Chen and Prowse 1964; Oberst et al. 1983, Höner et al. 1987). Therefore, it can be concluded that density effect can only be detected by regression after larger sizes have been reached. Stocking density was expressed as total biomass per m^3 (i.e., per volume of water). Additionally, the square root of these values was taken as a variance-stabilizing transformation (Weisberg 1980). In this form, stocking density explained a higher percentage of total variance than in form of $kg{\cdot}ha^{-1}$ (also with square root transformation). In the latter case, manure input became insignificant. These methods thus allowed the density effect of tilapia growth in ponds to be detected and quantified.

Pond area was either 400 or 1,000 m^2. The positive slope of the regression coefficient indicates that Nile tilapia growth was higher in larger ponds than in smaller ones. This is in line with the findings of Chen and Prowse (1964), who observed that tilapia grew better in larger ponds when kept at the same density. A certain minimum population size is necessary to enable the development of a more "peaceful" social hierarchy in tilapia (Baerends and Baerends-van Roon 1950; Chen and Prowse 1964; Antoniou 1989).

Solar radiation explained a higher percentage of variance in tilapia growth than pond water temperature. Solar radiation affects tilapia growth indirectly by increasing water temperature and consequently by enhancing fish metabolism and synthesis of body tissues. Besides, solar radiation controls primary production in the ponds, i.e., the amount of natural food available to Nile tilapia (Bowen 1982).

Together, all variables explain 40% of the total variation in Nile tilapia growth rate. The remaining 60% must be attributed to unmeasured effects, to errors in the dependent and independent variables, or to random variance. Beyond the determined set of variables, all other factors in the dataset turned out to be either insignificant or to be highly correlated with other variables already included in the model (e.g., water temperature, carp or predator biomass).

The empirically determined relationships could be studied further using path analysis. In the form of path coefficients, the effects of the (standardized) independent variables on tilapia growth could be compared in terms of their magnitude and interrelationships. Correlations between variables are shown and indicate possible interactions and combined effects. A path diagram was designed, including as many variables as possible, in an effort to represent as many components of the fishpond ecosystem as was possible with the available data.

Taking advantage of the large number of variables available in the ICLARM-CLSU dataset, more detailed causal path models may be designed and tested. Although some of the variables are not significant in directly explaining variance in fish growth, they can be used to quantify 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 biological implausibilities, the process must be repeated again until a correct and explicable model is derived.

Early morning dissolved oxygen (AMDO) was one of the key variables influencing Nile tilapia growth - albeit in a manner that is at first puzzling. The negative sign of the path coefficient indicates that Nile tilapia growth is *higher* when AMDO is *low*. AMDO is directly controlled by the amount of phytoplankton in the pond, besides other factors (Boyd 1982). Phytoplankton is in turn the main nutritional component for tilapia (Harbott 1982). With a higher plankton abundance, the amplitude of the diurnal dissolved oxygen variations increases, characterized by highest levels reaching 300% saturation in the early afternoon and lowest levels near zero at dawn (Hopkins and Cruz, unpubl. data). Tilapia can not only withstand these low levels for certain periods of time (Fernandes and Rantin 1987), but also grow satisfactorily (Tsadik and Kutty 1987). This is because the short-term negative effect is compensated by better growth due to a higher food amount during the rest of the day.

Plankton content or primary production was not measured. However, it can be shown, using path analysis, that other variables indirectly affect AMDO. These are solar radiation, wind run and cloud covering as meteorological variables, water temperature as pond environment variable and manure input as treatment variable. Although the data are based on averages over intervals, the percentage of variation in AMDO that can be explained by indirect effects is 50%.

Based on the available data, and including only "permissible" variables, the qualitative and quantitative relationships in livestock manure-fed tilapia ponds can be studied and described with the two methods used here.

Growth Parameters and ϕ'

The obtained values of the VBGF growth parameters K and L_∞ were inversely related as expected (Pauly 1979). Thus, both must be used when comparisons of growth are to be performed. The growth performance index ϕ' (= $\log_{10}K + 2\log_{10}L_\infty$) is a convenient and robust tool for the comparison of growth parameters from different datasets (Moreau et al. 1986; Pauly et al. 1988). Here it was not the aim to derive "representative" VBGF parameters to characterize stocks in terms of fisheries biology and population dynamics. The parameters obtained here are only valid for the tilapia strain and range of conditions in the experiments (Prein, this vol.b).

The VBGF parameters obtained for the ordinary Gulland-and-Holt plot produced a value for ϕ' of 3.37. With the extended Gulland-and-Holt method, an average value for ϕ' of 3.35 was obtained ranging from 3.11 for "worst" to 3.54 for "best" pond culture conditions as represented by the extreme values of all variables in the dataset.

Heteroskedasticity

The residual plots of derived regression models showed some evidence of heterogeneity of error variances or heteroskedasticity. The effects and implications of this phenomenon on models such as presented here are discussed in Prein and Pauly (this vol.).

Conclusions

When interpreting the results from the analyses of "historic" data, it is important to have as much information as possible which describe the details of the experiments. In the present case, the report describing the experiments were closely studied and several of the persons directly involved were interviewed. Also, the project site was visited twice by the author. The aim therein was to obtain the best possible understanding of the experiments and environmental conditions to make up for the fact that the data were not collected personally by the author.

The data analysis was based on multiple linear regression, which is available in most statistical software packages. As some of the datasets contained considerable amounts of variance, a robust approach in form of least median squares regression (LMS) was chosen to detect potentially harmful outliers in the datasets (Rousseeuw 1987; Rousseeuw and Leroy 1987). It turned out that in most cases, the same outliers were detected by LMS and LS (least squares regression). The LMS procedure is very helpful if the general trend of the regression line in the data is not known *a priori* and therefore must be determined empirically from the data. In the present case, the theory underlying the Gulland-and-Holt plot dictated the sign of the regression coefficients (slopes) associated with length or its transform. Extreme outliers could therefore be detected by inspection of the ordinary Gulland-and-Holt plot. Thereafter, normal techniques of outlier detection with regression diagnostics were applied (Daniel and Wood 1980; Velleman and Welsch 1981; Norusis 1985). A discussion of the performance of the extended Gulland-and-Holt method and path analysis is given in Prein (this vol.) and Prein and Pauly (this vol.).

Acknowledgements

I gratefully thank Kevin D. Hopkins and Roger S.V. Pullin for making available the raw data of the ICLARM/CLSU experiments. Daniel Pauly introduced me to the field of multivariate statistical analysis and further guided this work, which was part of a dissertation presented in 1990 at Kiel University, Germany. 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 for Economic Cooperation (BMZ).

References

Antoniou, E. 1989. Der Einfluß der Gruppengröße auf den Stoffwechsel und das Verhalten von Fischen. Mathematisch-Naturwissenschaftliche Fakultät, Christian-Albrechts-Universität Kiel. 69 p. M.S. thesis (Diplomarbeit).

Baerends, G.P. and J.M. Baerends-van Roon. 1950. An introduction to the study of the ethology of cichlid fishes. Behaviour (Supplement 1):1-242.

Bowen, S.H. 1982. Feeding, digestion and growth - qualitative considerations, p. 141-156. *In* R.S.V. Pullin and R.H. Lowe-McConnell (eds.) The biology and culture of tilapias. ICLARM Conf. Proc. 7, 432 p.

Boyd, C.E. 1982. Water quality management for pond fish culture. Dev. Aquacult. Fish. Sci. 9. Elsevier, Amsterdam. 318 p.

Brock, T.D. 1981. Calculating solar radiation for ecological studies. Ecol. Modelling 14:1-19.

Chen, F.Y. and G.A. Prowse. 1964. The effect of living space on the growth rate of fish. Ichthyologica 3:11-20.

Coulson, K.L. 1975. Solar and terrestial radiation. Methods and measurements. Academic Press, New York. 322 p.

Cruz, E.M. and Z.H. Shedadeh. 1980. Preliminary results of integrated pig-fish and duck-fish production tests, p. 225-238. *In* R.S.V. Pullin and Z.H. Shedadeh (eds.) Integrated agriculture-aquaculture farming systems. ICLARM Conf. Proc. 4, 258 p.

Daniel, C. and F.S. Wood. 1980. Fitting equations to data. Computer analysis of multifactor data. Second edition. John Wiley, New York. 458 p.

Duffie, J.A. and W.A. Beckman. 1980. Solar engineering of thermal processes. Wiley Interscience, New York. 762 p.

Fernandes, M.N. and F.T. Rantin. 1987. Respiratory responses of *Oreochromis niloticus* (Pisces, Cichlidae) to environmental reductions of dissolved oxygen. Bol. Fisiol. Anim. Univ. Sao Paulo 11:51-60.

Harbott, B.J. 1982. Studies of the feeding activity of *Sarotherodon niloticus* (L.) in Lake Turkana, p. 1357-1368. *In* A.J. Hoplon (ed.) Lake Turkana. A report on the findings of the Lake Turkana Project 1972-1975. Government of Kenya and The Ministry of Overseas Development, London.

Honer, G., H. Rosenthal and G. Krüner. 1987. Growth of juvenile *Sarotherodon galilaeus* in laboratory aquaria. J. Aquacult. Trop. 2:59-71.

Hopkins, K.D. and E.M. Cruz. 1980. High yields but still questions: three years of animal-fish farming. ICLARM Newsl. 3(4):12-13.

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

Hopkins, K.D., E.M. Cruz, M.L. Hopkins and K.C. Chong. 1981. Optimum manure loading rates in tropical freshwater fish ponds receiving untreated piggery wastes, p. 15-29. *In* ICLARM. The ICLARM-CLSU integrated animal-fish farming project: poultry-fish and pig-fish trials. ICLARM Tech. Rep. 5, 29 p.

Hopkins, K.D., P. Inocencio and E.M. Cruz. 1983. Water quality in integrated livestock-fish ponds. J. World Maricult. Soc. 14:495-504.

Hopkins, K.D., D. Pauly, E.M. Cruz and J.H. van Weerd. 1982. An alternative to predator-prey ratios in predicting recruitment. Meeresforschung/Rep. Mar. Res. 29:125-135.

List, R.J. 1963. Smithsonian meteorological tables. Sixth edition. Smithsonian Institution, Washington, DC. 540 p.

Lof, G.O.G., J.A. Duffie and C.O. Smith. 1966. World distribution of solar radiation. Solar Energy 10:27-37.

Milankovitch, M. 1930. Handbuch der Klimatologie. Vol. I. Part A. Mathematische Klimalehre, Berlin. Cited in Heyer, E. 1972. Witterung und Klima, eine allgemeine Klimatologie. Second edition. Teubner, Leipzig. 458 p.

Moreau, J., C. Bambino and D. Pauly. 1986. Indices of overall fish growth performance of 100 tilapia (Cichlidae) populations, p. 201-206. *In* J.L. Maclean, L.B. Dizon and L.V. Hosillos (eds.) The First Asian Fisheries Forum. Asian Fisheries Society, Manila, Philippines. 727 p.

Mosteller, F. and J.W. Tukey. 1977. Data analysis and regression. Addison-Wesley, Reading, Massachusetts. 588 p.

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

Oberst, S., W. Villwock and H. Rosenthal. 1983. Growth and food conversion in *Tilapia* under two different rearing conditions, p. 374-383. *In* L. Fishelson and Z. Yaron (comps.) Proceedings of the International Symposium on Tilapia in Aquaculture, 8-13 May 1983, Nazareth, Israel. Tel Aviv University, Tel Aviv. 624 p.

Paltridge, G.W. and C.M.R. Platt. 1976. Radiative processes in meteorology and climatology. Developments of Atmospheric Science. Vol. 5, Elsevier, New York. 318 p.

Pauly, D. 1979. Gill size and temperature as governing factors in fish growth: a generalization of von Bertalanffy's growth formula. Ber. Inst. Meereskd. Univ. Kiel. No. 63. 156 p.

Pauly, D., J. Moreau and M. Prein. 1988. A comparison of overall growth performance of tilapia in open waters and in aquaculture, p. 469-479. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia aquaculture. ICLARM Conf. Proc. 15, 623 p.

PCARRD. 1982. Integrated fish-pig farming. Technology 4(9):1-16. Philippine Council for Agriculture and Resources Research Development, Los Baños, Laguna, Philippines.

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

Prein, M. and F.C. Gayanilo Jr. 1992. SUNLIGHT: a computer program for calculation of daily solar radiation. NAGA, the ICLARM Quarterly. 15(3):30-31.

Pullin, R.S.V., Editor. 1988. Tilapia genetic resources for aquaculture. ICLARM Conf. Proc. 16, 108 p.

Rousseeuw, P.J. 1987. Least median of squares regression. J. Am. Stat. Assoc. 79:871-880.

Rousseeuw, P.J. and A.M. Leroy. 1987. Robust regression and outlier detection. John Wiley, New York. 329 p.
Sachs, L. 1984. Angewandte Statistik, Anwendung statistischer Methoden. Sixth edition. Springer Verlag, Berlin. 552 p.
Sellers, W.D. 1965. Physical climatology. University of Chicago Press, Chicago. 272 p.
Snedecor, G.W. and W.G. Cochran. 1980. Statistical methods. Seventh edition. The Iowa State University Press, Ames, Iowa. 507 p.
Spencer, J.W. 1971. Fourier series representation of the position of the sun. Search 2:172.
Sprugel, D.G. 1983. Correcting for bias in log-transformed allometric equations. Ecology 64(1):209-210.
Tsadik, G.G. and M.N. Kutty. 1987. Influence of ambient oxygen on feeding and growth of the tilapia, *Oreochromis niloticus* (Linnaeus). Working Paper No. 10. African Regional Aquaculture Centre, Port Harcourt, Nigeria. 13 p.
Tukey, J.W. 1977. Exploratory data analysis. Addison-Wesley, Reading, Massachussets. 688 p.
Velleman, P.F. and R.E. Welsch. 1981. Efficient computing of regression diagnostics. Am. Stat. 35:234-242.
Walter, H. and H. Lieth. 1969. Klimadiagramm Weltatlas. Gustav Fisher, Jena. pag. var.
Weber, E. 1980. Grundriss der biologischen Statistik. Anwendung der mathematischen Statistik in Forschung, Lehre und Praxis. Gustav Fischer, Jena. 625 p.
Weisberg, S. 1980. Applied linear regression. John Wiley, New York. 283 p.

Factor and Canonical Correlation Analysis of Nile Tilapia Production in Integrated Livestock-Fish Culture in the Philippines*

ANA MILSTEIN, *Agricultural Research Organization, Fish and Aquaculture Research Station, Dor, Mobile Post Hof HaCarmel, 30820, Israel*

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

Milstein, A. and M. Prein. 1993. Factor and canonical correlation analysis of Nile tilapia production in integrated livestock-fish culture in the Philippines, p. 67-74. *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

A dataset from integrated livestock-fish culture experiments conducted in the Philippines was analyzed with two multivariate statistical techniques: factor and canonical correlation analyses. Nile tilapia (*Oreochromis niloticus*), common carp (*Cyprinus carpio*) and snakehead (*Channa striata*) as predator were cultured in ponds at densities of 10,000 and 20,000$\cdot$ha^{-1} and pig, duck or chicken manure. Eight combinations of variables were identified as the main factors affecting the system, which together account for 86% of the overall variability of yields. These were: overall fish performance, tilapia stocking characteristics, snakehead production, carp performance, density effect on carp and predator growth, carp-predator interaction, predator survival, and tilapia survival. The canonical correlation analysis identified the factors affecting yield, growth rate or survival of each species, when all species were considered together. Topics concerning livestock type, manuring rate, early morning dissolved oxygen concentration in the pond water, uncontrolled tilapia spawning and predator-prey relationships are discussed.

Introduction

From 1978 to 1981 a series of integrated livestock-fish culture experiments were carried out at the Freshwater Aquaculture Center of the Central Luzon State University (CLSU), Muñoz, in cooperation with the International Center for Living Aquatic Resources Management (ICLARM) in Manila, both in the Philippines (Hopkins and Cruz 1982). Part of the database produced in the course of this project is analyzed herein, using two multivariate statistical techniques - factor and canonical correlation analyses - as appropriate for analysis of complex systems with multiple interactions (Milstein, this vol.).

Materials and Methods

The dataset analyzed here is an edited version of Appendix B of Hopkins and Cruz (1982), in which errors were corrected and to which further data extracted from the original raw data sheets were added (Prein, this vol.). This dataset is available on diskette No. 1 (Filename: PHILPROD.WK1), the contents of which are described in Appendix II.

The experiments were conducted in ponds of 400 or 1,000 m^2 during periods of 90 to 180 days. The fish species stocked were mixed-sex populations of herbivorous Nile tilapia (*Oreochromis niloticus*), bottom-feeding common carp (*Cyprinus carpio*), and snakehead (*Channa striata*) as a predator. These were stocked at densities of 10,000 and 20,000 fish$\cdot$ha^{-1} in ratios of 85:14:1 of tilapia, carp and snakehead, respectively. Nutritional input to the ponds was either manure from pigs, ducks or chicken raised on the ponds' embankments, or inorganic fertilizers. Further details are given in Hopkins and Cruz (1982) and Prein (this vol.).

In the present analyses each observation summarizes the inputs and outputs of an entire culture period (i.e., experiment) and not individual growth increments over short intervals, as analyzed by Prein (this vol.). The variables are expressed on a per-hectare and per-day basis where applicable. Fish growth rates for tilapia, carp and snakehead were calculated as average daily weight gains during the experiment. The variable MARKDAY is the net yield of fish of marketable size expressed on a daily basis. The values of manuring rates were computed by Hopkins and Cruz (1982) from individual livestock weights based on regressions describing the manure output

*ICLARM Contribution No. 730

based on regressions describing the manure output from animals of different sizes. Some missing values of livestock biomass (when chickens were stocked) were replaced by estimates based on an assumed weight of 1.5 kg per chicken.

The data were analyzed through factor and canonical correlation analyses, described in Milstein (this vol.). Furthermore, the effect of livestock type on the factors and on the extracted canonical response variables was analyzed through the general linear model used as unbalanced ANOVA. The procedures FACTOR, CANCORR and GLM of the SAS (1985) package were used.

Results and Discussion

Factor Analysis

Table 1 presents the factor analysis of 34 variables in 97 cases. Eight factors were selected which altogether account for 86% of the total variation.

Table 1. Factor analysis. Only large coefficients included (n = 97).

Variable[a]	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6	FACTOR7	FACTOR8
DAYS	.	-0.60	.	-0.48	-0.42	.	.	.
TDENi	.	0.73	.	.	-0.39	.	.	.
TWTi	.	0.66	.	.	.	.	0.43	.
TBIOi	.	0.75	.	.	.	.	0.45	.
TDENo	0.45	0.65	.	.	.	.	.	.
TSURV	0.42	.	.	.	.	.	.	-0.48
TBIOo	0.84	.	.	-0.42	.	.	.	.
TWTo	0.64	-0.50	.	.	.	.	.	.
TYIELDAY	0.86	.	.	.	.	.	.	.
TGROWTH	0.69	.	.	.	.	.	.	.
TRECR	.	-0.47	.	-0.44	-0.51	.	.	.
CDENi	.	.	0.40	.	-0.64	.	.	.
CWTi	.	.	.	0.74	.	.	.	.
CBIOi	.	.	.	0.70	.	.	.	.
CDENo	.	.	0.50	.	-0.68	.	.	.
CSURV	.	.	.	0.41	.	0.56	.	.
CBIOo	0.70	.	.	0.52	.	.	.	.
CWTo	0.62	.	.	.	.	.	.	.
CYIELDAY	0.56	.	.	0.49	.	0.54	.	.
CGROWTH	0.48	.	.	.	0.52	.	.	.
PDENi	.	0.61	.	.	.	.	0.48	.
PWTi	0.41	.	.	.	.	-0.54	.	0.49
PBIOi	0.54	.	.	.	.	-0.41	.	0.50
PDENo	.	0.46	0.45	.	.	.	.	.
PSURV	.	.	0.56	.	.	.	-0.48	.
PBIOo	0.56	.	0.65	.	.	.	.	.
PWTo	.	-0.42	0.51	.	.	.	.	.
PYIELDAY	0.44	.	0.74	.	.	.	.	.
PGROWTH	.	.	0.57	.	0.43	.	.	.
MARKDAY	0.89	.	.	.	.	.	.	.
TOTYIDAY	0.84	.	.	.	.	.	.	.
LSBIO	0.47	0.51	-0.42	.	.	.	-0.36	.
LSDEN	.	-0.63	.	.	.	.	-0.39	.
MANUDAY	0.46	.	-0.54	.	.	.	-0.42	.
Variance explained	23%	15%	11%	10%	9%	7%	5%	4%

[a]DAYS = number of days in experiment; TDENi = Nile tilapia initial stocking density (n·ha^{-1}); TWTi = tilapia mean individual weight at stocking (g); TBIOi= tilapia initial biomass (kg·ha^{-1}); TDENo = tilapia output density (at harvest, n·ha^{-1}); TSURV = tilapia survival (%); TBIOo = tilapia biomass at harvest (kg·ha^{-1}); TWTo = tilapia mean individual weight at harvest (g); TYIELDAY = tilapia net yield per day (kg·ha^{-1}·day^{-1}); TGROWTH = tilapia growth rate (g·day^{-1}); TRECR = tilapia recruitment (kg·ha^{-1}·experiment^{-1}). Same variables for common carp start with the letter C, and those for predator with P. MARKDAY = marketable fish yield per day (kg·ha^{-1}·day^{-1}); TOTYIDAY = total fish yield per day (kg·ha^{-1}·day^{-1}); LSBIO = mean livestock biomass (kg·ha^{-1}·experiment^{-1}); LSDEN = livestock density (n·ha^{-1} of fish pond); MANUDAY = manure application rate (kg·dry matter·ha^{-1}·day^{-1}).

FACTOR1 accounts for 23% of the overall variability of the dataset. It represents fish production; herein total and marketable yield depend mainly on tilapia, and secondarily on carp. Tilapia harvested biomass and yield are correlated with growth rate and harvesting size and, to a lesser degree, with density. Common carp harvested biomass and, to a lesser degree, yields are correlated with harvesting size. The correlation of carp growth rate with these variables is not strong. Total and marketable yield and tilapia and carp performances are also affected by the biomass (at stocking and at harvest) of the predator and by the amount of manure supplied, i.e., the biomass of livestock per pond. The ANOVA analysis showed that 26% of the variance of FACTOR1 is accounted for by livestock type and that higher yields occurred in ponds supplied with pig or chicken manure. Ponds supplied with duck manure have FACTOR1 values not significantly different from ponds supplied with chicken manure. The lowest values of this factor occurred in ponds with inorganic fertilizers and no livestock.

FACTOR2 describes tilapia stocking characteristics. It accounts for 15% of the variability of the data independently of the variability due to yield (FACTOR1). The larger the tilapia inputs (biomass, density and individual weight), the larger were predator density, livestock density and biomass. This generally occurred in short growing periods, which also resulted in small tilapia at harvest and small amounts of uncontrolled spawning in the ponds. Forty per cent of the variability of tilapia stocking characteristics (FACTOR2) is accounted for by livestock type. Higher values of FACTOR2 occurred when chicken was the livestock type.

FACTOR3 represents snakehead production and accounts for another 11% of the variability. Higher predator yield, harvesting biomass, survival and individual size, and growth rate were obtained when common carp density was high and manuring rate was low. This factor is not correlated with the type of livestock.

FACTOR4 is related to carp and accounts for another 10% of overall variability. It shows that larger carp weight and biomass at stocking, grown for relatively short periods, resulted in high carp biomass and yield at harvest, while small carp weight and biomass at stocking resulted in low biomass and yield at harvest, even when the latter were grown for longer periods. The longer the experiment, the higher was the carp mortality and hence the lower the yield. Additionally, longer experiments were associated with higher tilapia biomass at harvest and massive spawning. Twenty-five per cent of the variability of carp performance (FACTOR4) is accounted for by livestock type, wherein higher values occurred in ponds with ducks than in ponds with the other livestocks, while ponds with fertilizers were not significantly different from the ones with ducks or pigs.

FACTOR5 represents the density effect on carp and predator growth, and accounts for a further 9% of the overall data variability. The higher the carp and tilapia densities (the latter mainly due to uncontrolled spawning), the lower was the growth of carp and snakehead. This was correlated with the length of the culture period. It is only in the longer experiments, when the tilapias had enough time to produce significant amounts of recruits that the growth of the other fish species was affected. There was no correlation with any livestock type.

FACTOR6, due to carp-snakehead interactions, accounts for a further 7% of the overall variability in the data. The higher the individual stocking weight and total biomass of snakehead, the lower were carp survival and yield.

FACTOR7, which represents conditions related to the survival of snakehead, explains further 5% of overall variability. Predator survival was low in association with high tilapia stocking weight and biomass, high stocking density of snakehead, a low manuring rate (livestock biomass and number, and daily manure input). Larger tilapia were not eaten by the predator nor, it seems, was their spawning affected by snakehead.

FACTOR8, tilapia survival, explains 4% of variability. The larger the individual size of the predator at stocking, the lower was the survival of tilapia. Large predators prey on the stocked tilapia, at least in the beginning of the experiments. This is contrary to the intention that the predator should only prey on tilapia recruit resulting from uncontrolled spawning (Hopkins et al 1982).

Canonical Correlation Analysis

Three canonical correlation analyses were run on 102 observations, relating several explanatory variables (management variables) to the mean daily yield of each species, mean daily growth rate and survival. The results are presented in Table 2. The difference between the number of observations

Table 2. Results of canonical correlation analyses on fish yield, growth and survival of which only significant correlations are presented. Only high coefficients included. The sign (+) points to a variable correlated to the canonical variables (from canonical structural analysis). n = 102. CC1 = first canonical correlation. CC2 = second canonical correlation. E1 and E2 = explanatory variables. R1 and R2= response variables. See Table 1 for explanation of variables.

Analysis on:	Yields	Growth		Survival	
	CC1	CC1	CC2	CC1	CC2
Canonical correlation coefficient	0.72	0.65	0.57	0.48	0.46
Variance accounted for (%)	86	57	38	46	39
Standardized canonical coefficients for the explanatory variables					
	E1	E1	E2	E1	E2
DAYS	.	.	0.40	.	-0.97
TDENi	1.03	.	-0.40	.	-0.58
TWTi	-0.47	-0.43	.	0.82	.
CDENi	-0.53	-0.79	0.65	.	0.88
CWTi	.	.	.	-0.72	-0.57
PDENi	.	0.41	.	.	.
PWTi	.	.	-0.50	0.66	.
MANUDAY	0.77	0.52	0.84	0.51	.
Standardized canonical coefficients for the response variables					
	R1	R1	R2	R1	R2
TYIELDAY	1.11				
CYIELDAY	.				
PYIELDAY	.				
TGROWTH		(+)	1.09		
CGROWTH		0.80	-0.90		
PGROWTH		.	-0.51		
TSURV				.	.
CSURV				-0.92	-0.42
PSURV				.	0.97

in the factor and canonical analyses is due to five observations with missing values in variables included in the factor analysis, but excluded from the canonical correlations.

ANALYSIS OF DAILY YIELDS

Only one of the canonical correlations on daily yields was significant, accounting for 86% of the overall yield variability. It shows that tilapia daily yield was positively correlated with its stocking density and with manuring rate and negatively correlated with its stocking weight and carp stocking density.

ANALYSIS OF DAILY GROWTH RATES

In this analysis two correlations are significant, accounting for 57% and 38% of variability. The first canonical correlation is related to common carp growth, which is, on one hand, negatively correlated with its own stocking density and tilapia stocking weight and, on the other hand, positively correlated with manuring rate and snakehead stocking density. The canonical structure shows that tilapia growth rate is also positively correlated with this canonical variable. This means that the same combination of explanatory variables affects both carp and tilapia growth. The ANOVA model of this response variable shows that 17% of common carp growth variability is accounted for by livestock type, wherein pig, duck and chicken led to higher carp growth rates than inorganic or no fertilizer.

The second growth canonical correlation is related to all three species. Good tilapia growth together with low growth of carp and of snakehead were positively correlated with manuring rate and carp stocking density and negatively correlated with the size of the predator at stocking.

ANALYSIS OF SURVIVAL RATES

In this analysis two correlations are significant, accounting for 46% and 39% of the variability, respectively. The first canonical correlation shows carp survival positively correlated with its own stocking weight and negatively with the other fish stocking weights and with manuring rate.

The second canonical correlation shows high predator survival, associated with low carp survival, in short culture periods when carp was stocked at small size and at high density and when tilapia density was low. This suggests an interaction effect between carp and snakehead, in which small stocked carp were preyed upon by snakehead. This is contrary to the intention that snakehead should prey on tilapia recruits.

General Discussion

The multivariate analyses presented here describe some of the interrelationships in integrated livestock-fish polyculture systems. The present approach characterizes the system as a whole, in contrast to the approaches of Hopkins and Cruz (1982) and Prein (this vol.) which concentrated, respectively, on bivariate or multivariate analysis of the growth of only one species (Nile tilapia). Generally, each approach confirmed the results of the others, while providing further insights into the functioning of the system.

Prein (this vol.) found that the growth rate of tilapia was governed by mean early morning dissolved oxygen content of pond water (AMDO), food availability (i.e., manuring rate), tilapia stocking density and pond size. These findings are confirmed by the present analyses, in which FACTOR1 shows that not only tilapia growth, but also other fish performance variables are related to manuring rate (the effects associated with dissolved oxygen will be discussed further below). The positive influence of manuring rate on tilapia growth and yield is also shown in the canonical correlation analysis, along with the interactions with the other fish species.

In the factor analyses, two variables of moderate strength in FACTOR1 are manure loading rate and livestock biomass, which reflect the same information. This only moderate correlation is due to a nonlinear relationship as seen in Fig. 1. Fish production (FACTOR1) increases with manuring rate (or livestock biomass) up to a rate of approximately 100 kg dry matter·ha^{-1}·day^{-1}. Beyond this, the plot indicates that manuring rate exerts a depressing effect on fish production. This supports the findings of Hopkins and Cruz (1982; Figs. 4.2, 4.5 and 4.6) that amounts of pig and chicken manure above 100 to 110 kg dry matter·ha^{-1}·day^{-1} depressed tilapia yield and that chicken manure over 50 to 60 kg dry matter·ha^{-1}·day^{-1} depressed carp yield.

Fish production in the investigated ponds was related to the early morning dissolved oxygen content. The plot of FACTOR1 (fish production) over AMDO shows an inverse relationship (Fig. 2A). High AMDO values were related to low total fish production (tilapia, carp and snakehead) and low manuring rates. The early morning oxygen content of the ponds reflects the minimum point of the diurnal oxygen cycle. The lower the early morning oxygen level, the higher the amplitude, which means higher primary production during the day and in turn higher nighttime respiration. This is driven by the manure input, which provides nutrients for the autotrophic and heterotrophic food chains. The same trend is visible in the canonical correlation analyses of yields and growth, wherein

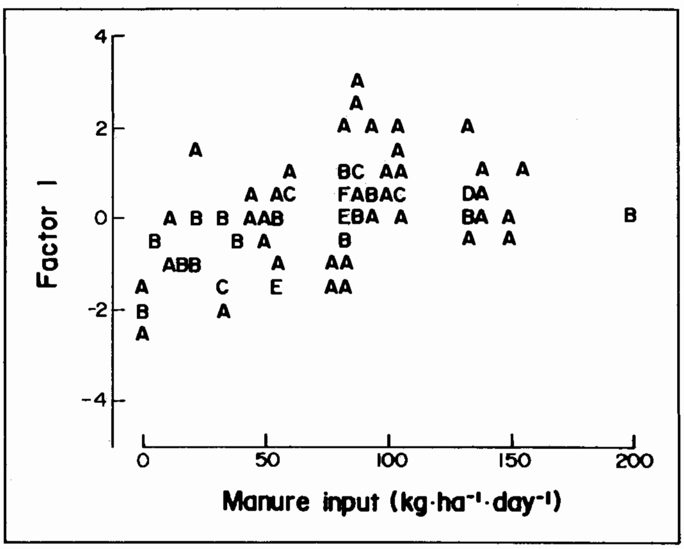


Fig. 1. Relationship between average daily manuring rate (MANUDAY) and overall fish production (FACTOR1) in integrated livestock-fish polyculture experiments in Muñoz, Philippines. Fish production increases with manuring rate up to a maximum production at around 100 to 110 kg dry matter·ha^{-1}·day^{-1}. Beyond this manuring rate a depressing effect on fish production is observed. A = 1 observation, B = 2 observations, etc.

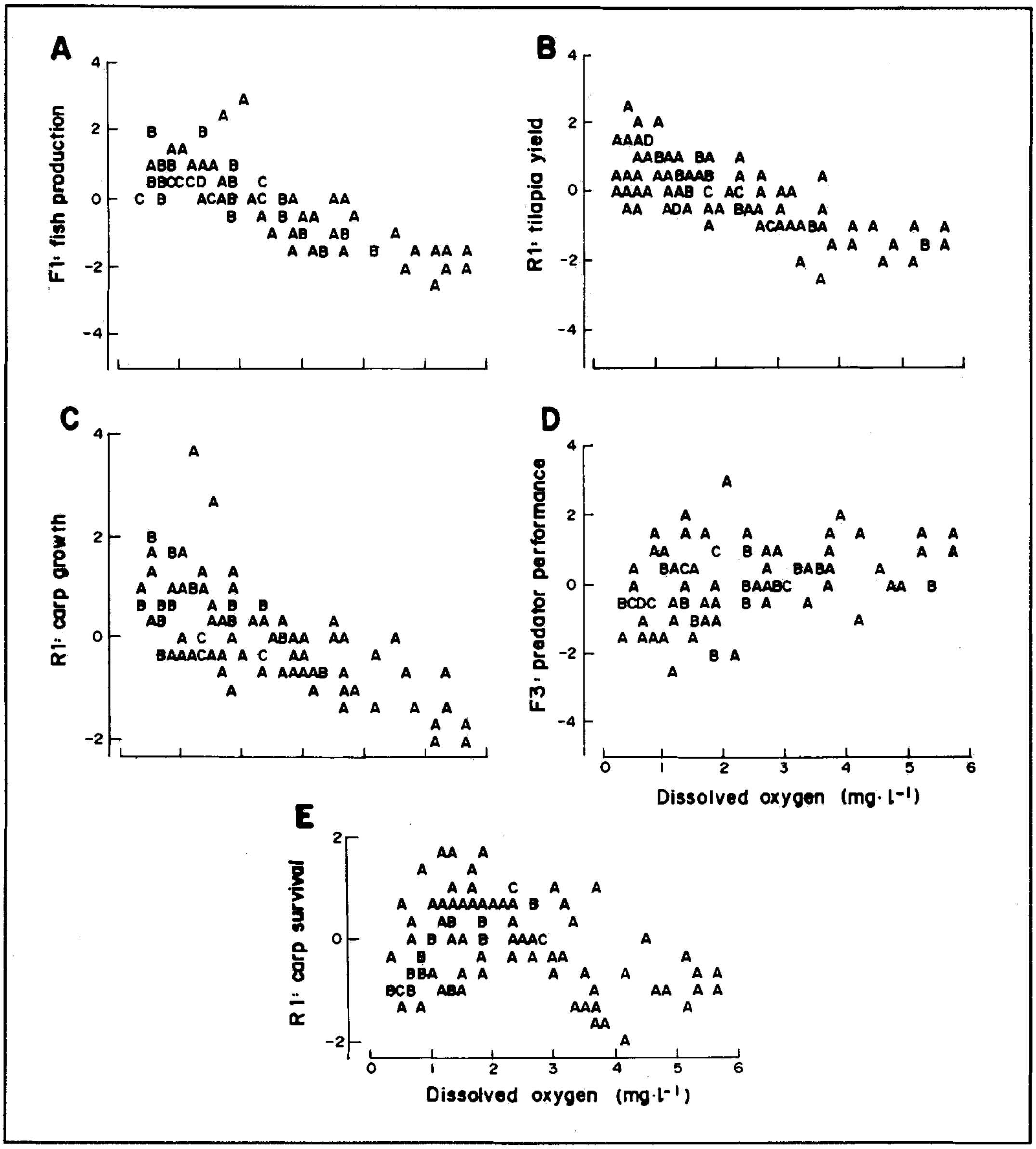


Fig. 2. Results of factor and canonical correlation analyses of integrated livestock-fish polyculture experiments in Muñoz, Philippines, in relation to early morning dissolved oxygen content (AMDO). A) *Factor analysis*: decrease of fish production (F1) with increased AMDO, reflecting the nutritional status of the pond for the fish. B) *Canonical correlation*: decrease of tilapia yield (R1: first response variable of canonical correlation analysis) with increasing AMDO. Higher tilapia yields were achieved in ponds in which the average AMDO per culture period was below 2 $mg \cdot l^{-1}$. C) *Canonical correlation*: decrease of carp growth with increasing AMDO. Lowest carp growth performance results in ponds with average AMDO values above 2 $mg \cdot l^{-1}$. D) *Factor analysis*: positive trend of predator (*Channa striata*) performance (F3) with increasing AMDO. Higher growth and survival at higher AMDO reflects the better prey visibility for the snakehead. E) *Canonical correlation*: effect of AMDO on carp survival (R1) where high R1-values represent low carp survival and *vice versa*. High AMDO directly affects (i.e., increases) carp survival, while low AMDO values increase natural food availability which, in turn, increases carp survival. A = 1 observation, B = 2 observations, etc.

the first response variables (tilapia yield and carp growth) were negatively correlated with AMDO (Fig. 2B and 2C).

Besides the above-mentioned effect, AMDO can have a further indirect influence on fish. A plot of FACTOR3 (reflecting predator performance) over early morning dissolved oxygen content of the pond water shows a linear trend (Fig. 2D). Low AMDO values correspond with low predator performance. The snakehead is capable of air breathing. At high AMDO values, the pond water contains less plankton and is therefore less turbid. This enables the snakehead to find prey more easily, leading to better growth and survival, i.e., performance.

The effect of AMDO on carp survival is shown in Fig. 2E. The figure suggests that carp survival under conditions of high and low AMDO is affected by the two different processes described above. High AMDO directly and positively affects carp survival. On the other hand, carp survival is also increased by availability of natural food, related to low AMDO values.

One of the key questions in conducting the ICLARM-CLSU experiments was to ascertain the effect of manure from different livestock types on fish production. Therefore, an analysis of variance was conducted which revealed that 24% of the variability in marketable fish biomass (expressed on a daily basis) is accounted for by livestock type. The marketable fish yield was 15 $kg \cdot ha^{-1} \cdot day^{-1}$ when chicken or pig were the livestock type, as compared to 10 and 4 $kg \cdot ha^{-1} \cdot day^{-1}$ for ducks and inorganic fertilizers, respectively. Similarly, 27% of the variability in total fish yield is accounted for by livestock type, where 18 $kg \cdot ha^{-1} \cdot day^{-1}$ were obtained with chicken or pig as livestock type and again 10 and 4 $kg \cdot ha^{-1} \cdot day^{-1}$ for ducks and inorganic fertilizers, respectively. Hopkins and Cruz (1982) also found that chicken and pig manure produced higher yields than ducks and inorganic fertilizers. However, when all variables are considered simultaneously, overall fish production (FACTOR1, which includes marketable and total daily yield) was still higher when pig or chicken were the livestock type. (Note, however, the ducks did not differ significantly from the chicken).

Uncontrolled spawning of tilapias is considered a nuisance in pond culture (e.g., Hepher and Pruginin 1981; Guerrero 1982). It may affect target fish performance, as shown by FACTOR4 for carp yield and FACTOR5 for carp and snakehead growth rate. The number of tilapia recruits occurring in the ponds is a variable of moderate strength in FACTOR2 (tilapia stocking characteristics), showing a linear relationship with this factor over only part of its range (Fig. 3). High numbers of wild spawns occurred only with low values of FACTOR2, that is in long culture periods when tilapia biomass, density and individual weight at stocking were low, along with predator density and livestock biomass and density. Zero or low numbers of tilapia recruits are associated with a

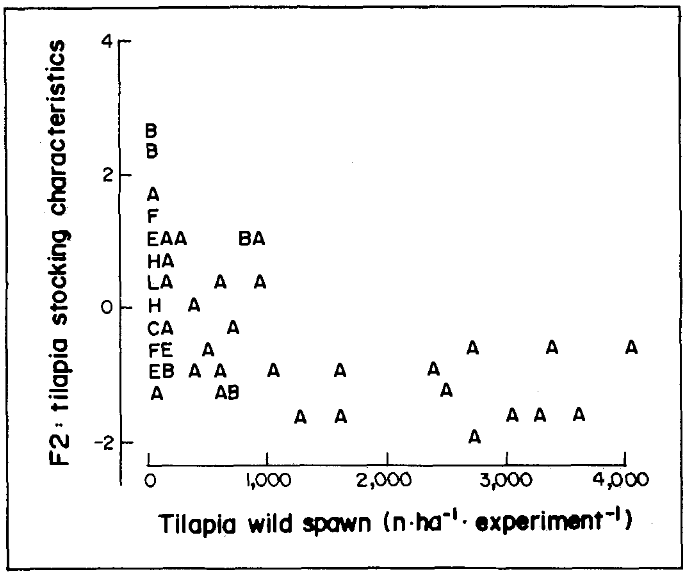


Fig. 3. Nonlinear relationship of Nile tilapia stocking characteristics (F2) with tilapia wild spawn production. Large amounts of recruits were produced in the ponds when culture periods were long, when Nile tilapia biomass, density and individual weight at stocking were low, and when these were associated with low predator density and low livestock density and biomass. A = 1 observation, B = 2 observations, etc.

wide range of values for FACTOR2. To control the recruits, predators are sometimes stocked in polyculture with tilapia (Guerrero 1982; Wohlfarth and Hulata 1983). Hopkins and Cruz (1982) concluded that the predator-prey ratios they used were ineffective in controlling tilapia recruitment. In the present analyses, survival of the stocked tilapia (FACTOR8), but not of the recruits, was affected by the predator. Besides this failure of snakehead to control tilapia recruitment, several predator-prey relationships were identified, notably between the common carp and the snakehead (FACTOR3, FACTOR5 and FACTOR6 and the second canonical correlation on survival).

We conclude by noting that factor and canonical correlation analyses, as applied here, not only confirmed previous results but also provided new insights into the understanding of the complex systems analyzed herein and which uni- or bivariate methods would have failed to even describe adequately.

Acknowledgement

This study was supported 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 for Economic Cooperation (BMZ).

References

Guerrero, R.D. III. 1982. Control of tilapia reproduction, p. 309-316. *In* R.S.V. Pullin and R.H. Lowe-McConnell (eds.) The biology and culture of tilapias. ICLARM Conf. Proc. 7, 432 p.

Hepher, B. and Y. Pruginin 1981. Commercial fish farming, with special reference to fish culture in Israel. Wiley Interscience, New York. 261 p.

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

Hopkins, K.D., D. Pauly, E.M. Cruz and J.H. van Weerd. 1982. Recruitment control: an alternative to predator-prey ratios. Meeresforschung/Rep. Mar. Res. 29(3):125-135.

SAS. 1985. SAS user's guide: statistics. Version 5 edition. 957p. SAS Institute Inc., Cary, Connecticut.

Wohlfarth, G.W. and G. Hulata. 1983. Applied genetics of tilapias. ICLARM Stud. Rev. 6, 26 p.

Multivariate Analysis of Tilapia Growth Experiments in Israel, Zambia and Peru*

Mark Prein, *International Center for Living Aquatic Resources Management, MCPO Box 2631, 0718 Makati, Metro Manila, Philippines*

Prein, M. 1993. Multivariate analysis of tilapia growth experiments in Israel, Zambia and Peru, p. 75-87. *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

Datasets from tilapia pond experiments in Israel, Zambia and Peru were analyzed with path analysis and multiple regression. Growth of mixed-sex Nile tilapia (*Oreochromis niloticus* Fam. Cichlidae) in experiments conducted in polyculture ponds in Israel, in 1982 and 1983, was positively affected by water temperature and percentage of males in the population. Other treatment variables such as chicken manure and food pellet applications and meteorological variables did not result as appropriate predictor variables for Nile tilapia growth in these experiments. Growth of *O. andersonii* (both sexes) in experiments conducted in Zambia from 1955 to 1957 was mainly controlled by water temperature. Food applications, stocking densities and other tilapia species in polyculture were insignificant variables. Growth of all male Nile tilapia in culture experiments performed in Lima, Peru, was positively affected by manure input and water temperature and negatively affected by stocking density.

The regression models developed were analyzed using causal path diagrams and sensitivity analysis. Also, the parameters of the von Bertalanffy growth equation were estimated for each of the tilapia species and environmental settings.

Introduction

The wide distribution of tilapias, many different geographic locations and climate types of the world has led to numerous experiments for testing the growth performance of their different species, hybrids and strains in various systems, subjected to a range of different treatments. Some of these experiments date back to the 1950s and have been conducted and analyzed according to the then prevailing knowledge, available methods and research aims. With increasing knowledge in the component disciplines of aquaculture, the "historic" data collected in previous experiments represent a valuable source of new information, which can be obtained from reanalyses based on modern methods. Additionally, this procedure proves to be very economic in that much information can be obtained from experiments that have already been conducted.

This paper presents three applications of new multivariate methods to historic data. The data sources are Nile tilapia culture experiments at (i) Dor, Israel, (ii) Lima, Peru, and (iii) experiments with *Oreochromis andersonii* in Zambia. In the first case, raw data were supplied by the original experimenters; in the other two cases, raw data were extracted from published reports.

Materials and Methods

The multivariate methods used here are a) multiple regression analysis of growth rate in form of the "extended Gulland-and-Holt plot" and b) path analysis, which are described in Pauly et al. (this vol.) and in Prein and Pauly (this vol.), respectively.

Since the fish sizes were recorded in form of weight, while the analysis methods require lengths, the former were transformed from weights using length-weight relationships. The weights of the Nile tilapia in the experiments in Dor, Israel, and Lima, Peru, were converted to lengths with the following relationship, determined for Nile tilapia grown in ponds in the Philippines (Prein, this vol.).

$$TL = (W/0.01065)^{1/3.258} \qquad \ldots 1)$$

where W is in g and total length (TL) is in cm and for which n = 612, r^2 = 0.973, SEE = 0.215, P < 0.001, size range: 0.7 to 211 g, 4.3 to 22.0 cm.

In the case of the Nile tilapia at Dor these fish were of the same strain as those cultured in the Philippines.

For the experiments with *Oreochromis andersonii* conducted by Mortimer (1960) in Zambia, a new relationship was determined, i.e.,

*ICLARM Contribution No. 731

$$TL = (W/0.01069)^{1/3.173} \quad ...2)$$

where W is in g and total length (TL) is in cm, with n = 39, $r^2 = 0.997$, SEE = 0.058, $P < 0.001$; size range: 6 to 272 g, 7.7 to 25.0 cm (Fig. 1).

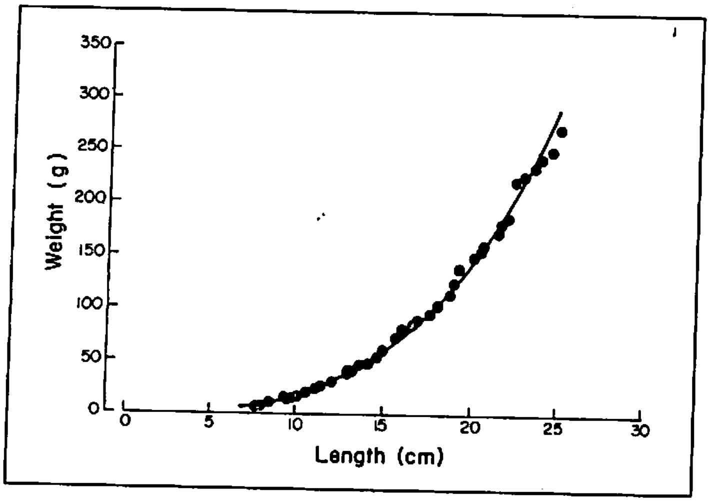


Fig. 1. Length-weight relationship of *Oreochromis andersonii* based on data from experiments conducted in Zambia, 1955-1957(n = 39). See text for regression equation.

Sensitivity Analysis

The individual variables in a multiple regression equation differ in the impact they have on the dependent variable. These effects can be compared and studied with sensitivity analysis. The approach chosen here, termed ordinary sensitivity analysis (Majkowski 1982), was to compute the percentage of change in growth rate caused by the independent variables when these were individually varied in steps, up to 50% from their average values in the dataset. While one variable was changed, all others were kept at their average values. The change in growth rate caused by the independent variables is expressed in terms of % change relative to the average response when computed with average values for all variables.

Comparison of VBGF Parameters

The parameters K and L_∞ of the von Bertalanffy growth function (VBGF) are inversely related and they must be used jointly when assessing the growth performance of fishes. The growth performance index ϕ' (= $\log_{10}K + 2\log_{10}L_\infty$) is a convenient and robust tool for the comparison of growth parameters from different datasets (Pauly 1979; Moreau et al. 1986; Pauly et al. 1988).

Values of these parameters were computed, and used for comparisons between culture systems and strains/species.

Description of Experiments at Dor, Israel

Growth experiments with pure Nile tilapia (*O. niloticus*; Ghana strain) were conducted at the Dor research station in 1982 and 1983. Summary results were published in Hulata et al. (1986). The tilapias were of the same stock as those used in the ICLARM/CLSU experiments in the Philippines. Twenty different ponds of 400 m^2 each were used.

In the experiments both sexes of tilapias were grown from fingerling size to over 150 to 200 g in polyculture with common carp, silver carp, grass carp and giant freshwater prawns at total stocking densities of $14{,}000 \cdot ha^{-1}$. The main source of nutrients was dry chicken manure applied at 50 to 175 $kg \cdot ha^{-1} \cdot day^{-1}$ six days per week. Additionally, small amounts of pellets (25% protein) were fed daily (Hulata et al. 1986).

The experiments were not originally designed to be analyzed by multivariate methods based on length with the inclusion of environmental variables. Therefore, the dataset required specific treatments for a final dataset of adequate format and content to emerge, enabling a multivariate statistical analysis of the type envisaged here.

Water temperature is the only water quality parameter, and air temperature is the only meteorological parameter available for Dor station. These were recorded by Mr. Amir Halevy as daily minimum-maximum values since 1978 and were made available to the author. Further meteorological variables were generated from meteorological data recorded at the meteorological station at 'En Hahoresh (34°56'E, 32°23'N), which is located 12 km south of Dor (34°56'E, 32°37'N). Solar radiation data were obtained from the Meteorological Services at Bet Dagan (34°49'E, 32°00'N) near Tel Aviv.

Description of Mortimer's Experiments

The results of numerous growth experiments with *O. andersonii* (both sexes) conducted from 1955 to 1957 at the research stations at Chilanga and Mwekera were published by Mortimer (1960). Both stations are located near Lusaka, Zambia.

O. andersonii were grown in polyculture with *O. macrochir* and *Tilapia rendalli*. Mortimer

(1960) originally identified the latter species as "*T. melanopleura*", but this was corrected to *T. rendalli* by Thys van den Audenaerde (1968).

The aim of the experiments was to obtain data on the growth of *O. andersonii* under different densities, different species in polyculture and different feeding rates. For the experiments at the Mwekera research station, 20 ponds of 200 to 240 m^2 area and 1 to 1.25 m depth were used. At Chilanga research station, experiments were conducted in six ponds of 1,000 to 1,200 m^2 and depths sloping from 0.15 to 1.5 m. Stocking sizes were smaller than 10 cm. At least 10% of the pond population were sampled with a net of 2.5 or 3.8 cm stretched mesh. The duration of experiments ranged from 10 to 15 months.

The feed consisted of maize bran or ground maize (6-10% protein) given at rates ranging from 17 to 68 $kg \cdot ha^{-1} \cdot day^{-1}$ according to size of fish. During the last 6 to 7 months of the 1956/1957 experiments, shredded grass was added to the ponds at a rate of 340 $kg \cdot ha^{-1} \cdot week^{-1}$, mainly as food for the herbivorous *T. rendalli*.

Fish were stocked as fingerlings. Sizes of *O. andersonii* at harvest were in the range of 96 to 252 g. Extrapolated net yields of all three species combined ranged from 7 to 619 $kg \cdot ha^{-1} \cdot year^{-1}$ with an average of 267 $kg \cdot ha^{-1} \cdot year^{-1}$.

Mortimer (1960) published average monthly water temperatures for 24 months in Chilanga ponds and 17 months in the Mwekera ponds. These ranged from 14.2 to 23.9°C for Chilanga ponds and 17.3 to 23.7°C in the Mwekera ponds. These data were used to calculate average water temperatures during the fish growth intervals.

Data on growth, length-weight relationship, stocking density, feeding rates and water temperature were presented in tables and graphs in Mortimer (1960) from where they were read off and entered into microcomputer spreadsheets.

Description of Delgado's Experiments in Lima, Peru

A series of nine experiments was conducted over two years from October 1980 to August 1982 at the aquaculture research station of the Instituto del Mar del Peru (IMARPE) at Huachipa, located on the eastern outskirts of Lima. The data were published in the final report by Delgado (1985). The station at Huachipa was visited by the author in September 1988.

Integrated livestock-fish culture experiments were conducted in concrete ponds of 113 m^2 and 0.5 m depth and in earthen ponds ranging from 253 to 1,500 m^2 at 1.0 m depth. All-male (hand-sexed) Nile tilapia were the only species stocked at densities ranging from 6,000 to 24,000 $\cdot ha^{-1}$. Pig and duck manure were applied individually, or both mixed, at loading rates ranging from 67 to 205 $kg \cdot ha^{-1} \cdot day^{-1}$. Fish were not given any supplementary feed. The length of the fish culture periods ranged from 177 to 347 days. Average fish sizes at stocking and at harvest ranged from 13 to 82 g and 147 to 248 g, respectively. Net yields extrapolated to a whole year, ranged from 761 to 4,272 $kg \cdot ha^{-1} \cdot year^{-1}$ with an average of 1,996 $kg \cdot ha^{-1} \cdot year^{-1}$.

A sample of 5 to 10% of the pond population was taken monthly and every individual fish sampled was measured and weighed. Only weights are given in the report. Aside from water temperature, no further data for water quality or meteorological variables were available in the report.

Delgado (1985) published average monthly water temperatures, in Huachipa, covering the experimental period which were incorporated in the data analysis. The average and range of monthly water temperature was 21.3 (18.0 to 24.0) and 22.4 (19.3 to 26.3) °C for earthen and concrete ponds, respectively.

Data Handling and Processing

Central to all data handling and editing procedures was the establishment of several small spreadsheets on microcomputers. From these raw data files the required average values per growth interval were computed and inserted into the main datafile as described in Pauly et al. (this vol.). Regression analyses were performed with the SPSS software package (Norusis 1985). The 95% significance level was used for all tests. Procedures for data handling and processing were done as described in Pauly et al. (this vol.). The regression type used here was a Type I, or "predictive" regression (but see Prein and Pauly, this vol.). The datasets analyzed are described in Appendix II.

Climatic Diagrams

The influences of climate are among the main criteria for comparing the productivity of different regions of the earth. Instead of studying large tables of different climatic variables for different regions, the examination of climatic diagrams provide a means to assess quickly the climatic conditions of a certain region, since these are standardized and available for a large number of locations (Walter and Lieth 1969). Climatic diagrams contain information on two variables: air temperature and precipitation. These are of particular importance for plant growth, but helpful insights and characteristics relevant to aquaculture may also be gained from the data. A short description is given here, adapted from Walter and Lieth (1969).

"Monthly averages of air temperature (k = thin line) and precipitation (l = thick line) curves are set in relation by standardized scales on the diagrams (Fig. 2). Ten degrees centigrade correspond to a precipitation of 20 mm. In this arrangement one can characterize an arid period when precipitation drops below the temperature curve (m = open-dotted area) and a humid period when monthly precipitation exceeds temperature (n = finely dotted area). Precipitations exceeding 100 mm are shown in a 1:10 scale (o = black area). In certain cases it has proven useful to include an additional precipitation curve (p = dotted line) in the diagram in a 1:3 scale. Through this method of presentation, unfavorable seasons causing water shortage can be characterized. Occurrence of frost is depicted by special blocks on the abscissa for each month. Blocks are black (q) in case the average minimum of a month falls below zero degrees centigrade, and hatched (r) in case only the absolute minimum lies below zero. Each descriptive number has its defined position in the diagram. Its absence indicates that appropriate data were not available. To facilitate direct comparisons, the abscissa begins in January for locations on the northern hemisphere, and in July for locations on the southern hemisphere.

For the interpretation of climatic diagrams for ecological and biological studies, certain restrictions must be considered. All measurements by the meteorological stations are conducted under protection from direct solar radiation. Thus, straightforward evaluation of water and energy budgets and of the photosynthetic activity of plants is difficult. The recording instruments are located 2 m above the ground, which is not representative of the near-bottom air layer. The

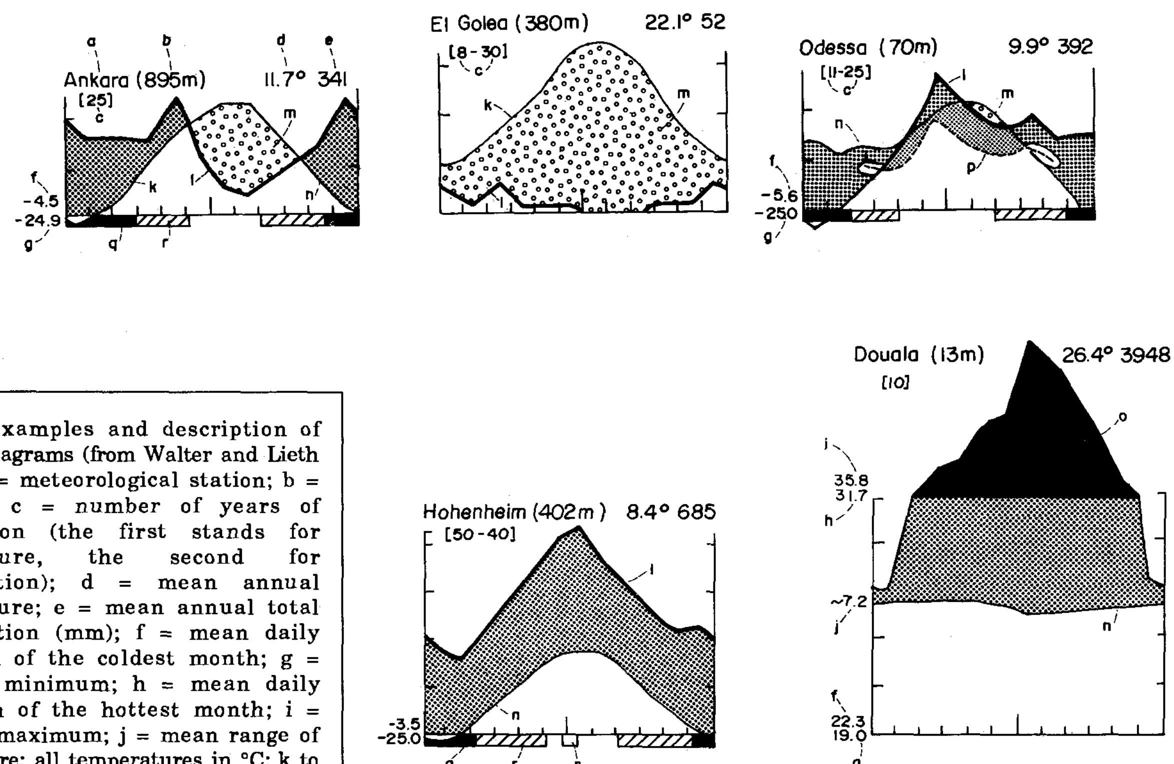


Fig. 2. Examples and description of climatic diagrams (from Walter and Lieth 1969): a = meteorological station; b = altitude; c = number of years of observation (the first stands for temperature, the second for precipitation); d = mean annual temperature; e = mean annual total precipitation (mm); f = mean daily minimum of the coldest month; g = absolute minimum; h = mean daily maximum of the hottest month; i = absolute maximum; j = mean range of temperature; all temperatures in °C; k to r and shades are described in text.

diagrams are based on long-term averages, yet this "normal climate" practically never occurs. Unperiodic deviations and large variances are of high importance for the survival of plants and animals."

The data from Israel, Zambia and Peru are given in the files DORGROW.WK1, ZAMB.WK1 and LIMA.WK1 and are described in Appendix II.

Results

Comparison of Climatic Regions

ISRAEL

The climatic type of Israel is Mediterranean, with dry summers and rain in winter (Fig. 3). The station in Natanya was chosen as representative for the climate in Dor. In Israel, air (and water) temperature reaches levels below 15°C which are lethal to tilapias. Further, there is the disadvantage of limited amounts of water in Israel, restricting aquaculture management and expansion (Hepher 1985).

PERU

The experiments were conducted in Huachipa (76°56'W, 12°00'S) on the eastern outskirts of Lima, for which the climatic diagram is given in Fig. 3. The climate is arid year-round with very low precipitations; indeed there is no clear rainy season. The average annual temperature is 19.3°C, ranging from 16 to 23°C. During the cooler winter months (May to September) the area is cloudy and foggy, which is the only source of precipitation.

ZAMBIA

The climate of Zambia is characterized as tropical with a dry season from October to April and summer rains from May to September. The location of the Chilanga experiment station is near Lusaka (28°17'E, 15°34'S), that of Mwekera is near Kitwe (28°13'E, 12°49'S). A climatic diagram is available for Lusaka (Fig. 3). The average annual temperature is 20.6°C, ranging from 16 to 24°C.

Israel: Dor Station Experiments

EXTENDED GULLAND-AND-HOLT PLOT ANALYSIS

An ordinary Gulland-and-Holt plot of the entire Nile tilapia data from Dor station is given in Fig. 4A. The corresponding regression equation is:

$$\Delta L/\Delta t = 0.402 - 0.0178 \ ML \qquad ...3)$$

where $\Delta L/\Delta t$ are the standardized growth increments, and ML the mean lengths of the fish; also $n = 156$, $r^2 = 0.50$, SEE = 0.0556, $P < 0.001$, $K = 0.0178 \ day^{-1}$ and $L_\infty = 22.5$ cm.

For the data from Dor, the multiple regression, obtained with the extended Gulland-and-Holt plot, takes the form:

$\Delta L/\Delta t =$		mean,	range
-0.01321	mean length (cm)	15.8	9-20
+0.01046	water temperature (°C)	28.1	22.6-30.1
$+9.710 \cdot 10^{-4}$	per cent males	46.5	25-74
$-9.735 \cdot 10^{-3}$			

...4)

with $n = 156$, $R^2 = 0.569$, SEE = 0.0519, $P < 0.001$, $K = 0.01321 \ day^{-1}$ and $L_\infty = 24.9$ (range 19.0 to 28.5 cm).

This multiple regression model explained 57% of the total variance in Nile tilapia growth rate; the individual contributions of the variables were: mean length 28%, water temperature 11% and per cent males 4%.

It was not possible to establish a biologically plausible and statistically sound extended Gulland-and-Holt plot including more variables with the data from Dor station. Most treatment variables were highly collinear, i.e., increased or decreased gradually as the fish grew. This resulted in erroneous signs for the regression coefficients. Similarly, none of the meteorological variables could be identified as significant predictors.

PATH ANALYSIS

A causal path diagram of the Nile tilapia growth rate in the Dor station experiments (based on the ordinary Gulland-and-Holt plot) is shown in Fig. 5A. Here, 50% of the variation in growth rate is left unexplained. A path diagram based on the extended Gulland-and-Holt plot incorporates three further variables (Fig. 5B). Here PERC is the percentage of males in the tilapia population, determined after final harvest.

Water temperature and the percentage of males have significant positive effects on Nile tilapia growth in ponds. Together with mean length, these variables explain 57% of the total variation. None of the variables are treatment variables. Solar radiation has a significantly positive effect on water temperature, responsible for 88% of the

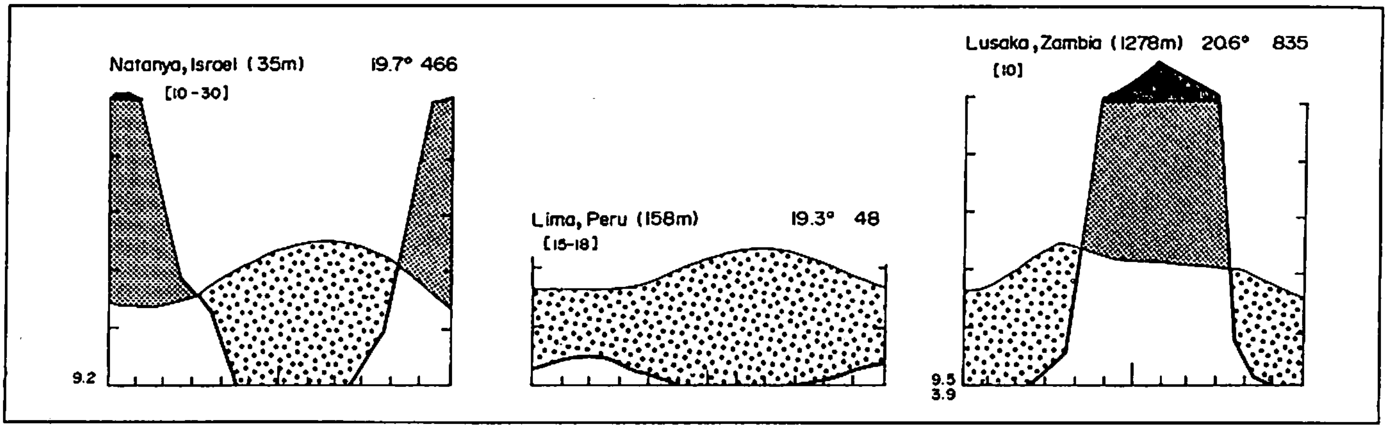


Fig. 3. Climatic diagrams for experimental locations from which data was analyzed here. From Walter and Lieth (1969).

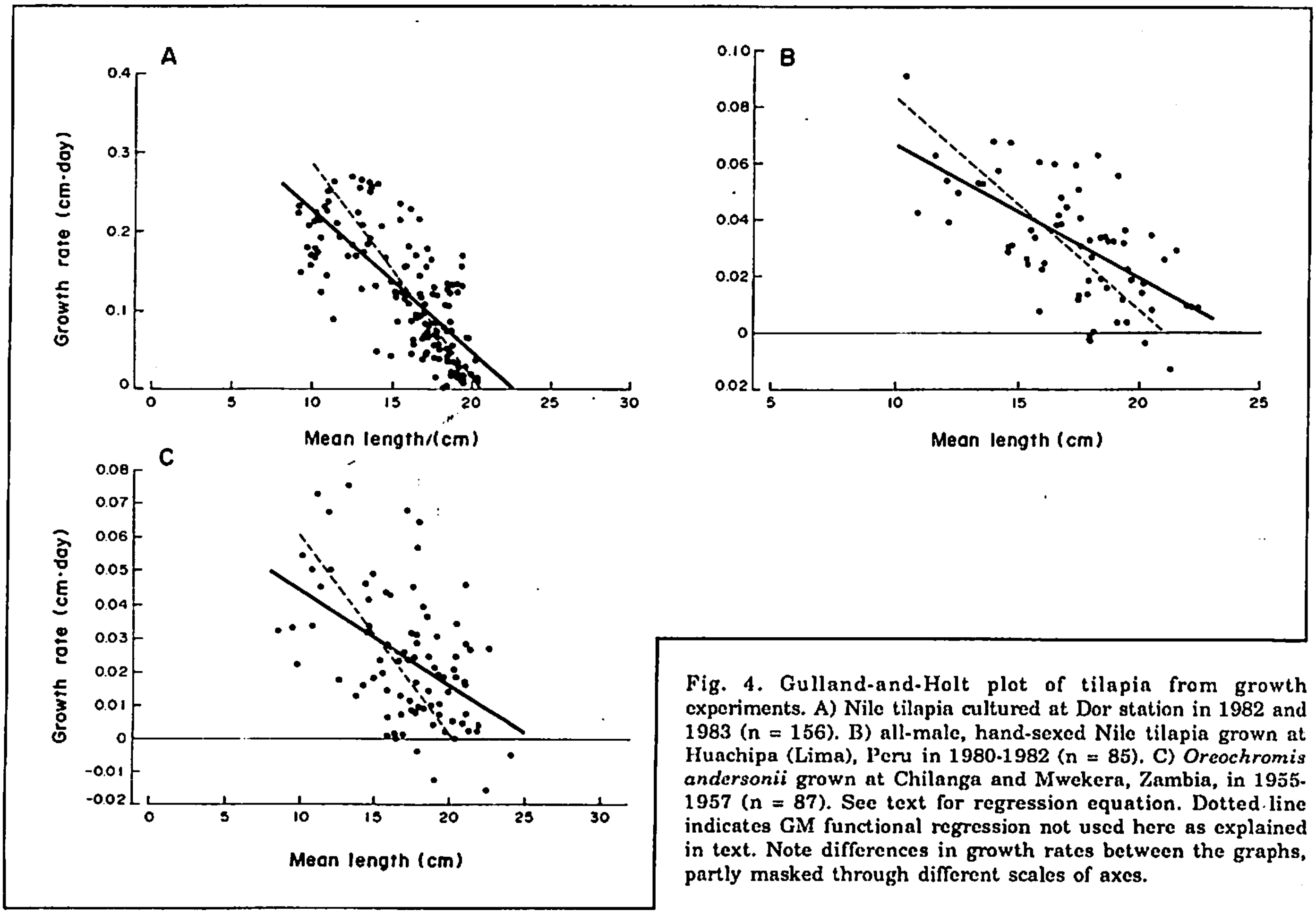


Fig. 4. Gulland-and-Holt plot of tilapia from growth experiments. A) Nile tilapia cultured at Dor station in 1982 and 1983 (n = 156). B) all-male, hand-sexed Nile tilapia grown at Huachipa (Lima), Peru in 1980-1982 (n = 85). C) *Oreochromis andersonii* grown at Chilanga and Mwekera, Zambia, in 1955-1957 (n = 87). See text for regression equation. Dotted line indicates GM functional regression not used here as explained in text. Note differences in growth rates between the graphs, partly masked through different scales of axes.

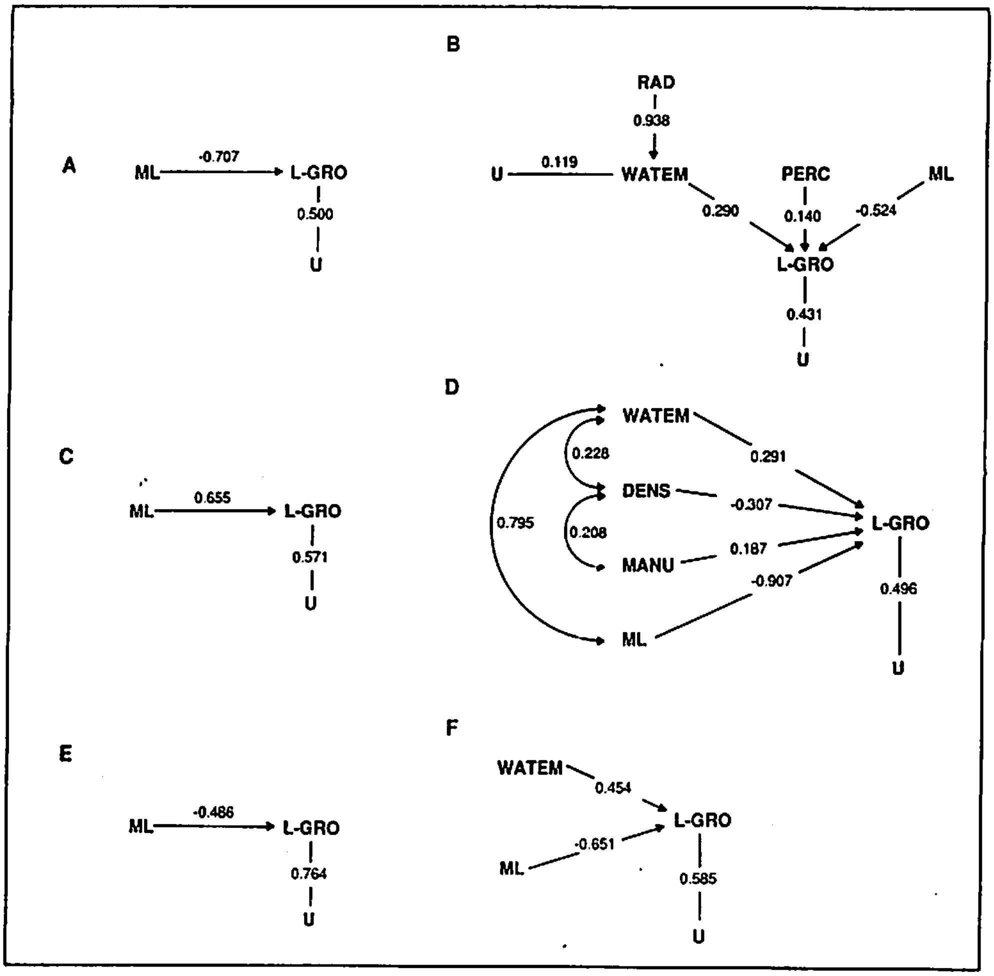


Fig. 5. Causal path diagrams of tilapia growth in the experiments analyzed here, based on A) the ordinary and B) extended Gulland-and-Holt plot of Nile tilapia growth at Dor Station, Israel; C) ordinary, and D) extended Gulland-and-Holt plot of Nile tilapia growth in Huachipa/ Lima, Peru; E) ordinary, and F) extended Gulland-and-Holt plot of *Oreochromis andersonii* growth in experiment stations in Zambia. L-GRO = growth rate in length ($cm \cdot day^{-1}$); ML = mean length (cm); PERC = percentage of male tilapia in the pond population; WATEM = early morning water temperature (°C); RAD = solar radiation ($ly \cdot day^{-1}$); DENS = stocking density ($n \cdot ha^{-1}$); MANU = wet manure input ($kg \cdot ha^{-1} \cdot day^{-1}$).

variation. Usually, wind velocity is understood to have a cooling effect on pond water temperature and a mixing effect on the water body. In the present case though, wind velocity had the wrong sign and cloud cover was not significant when tested for inclusion in the model.

Peru

EXTENDED GULLAND-AND-HOLT ANALYSIS

An ordinary Gulland-and-Holt plot of the Nile tilapia data from the experiments conducted in Huachipa, Peru (Fig. 4B), results in:

$$\Delta L/\Delta t = 0.1128 - 0.00473 \text{ ML} \qquad \text{...7)}$$

with $n = 85$, $r^2 = 0.429$, SEE = 0.0161, $P < 0.001$, $K = 0.00473 \text{ day}^{-1}$ and $L_\infty = 23.9$ cm.

The amount of explained variance was increased to 50% through the inclusion of auxiliary variables. The amount of total variation in Nile tilapia growth rate explained by the independent variables is: mean length 31%, water temperature 4.2%, stocking density 12.5% and manure input 5%. No variable representing pond dynamics or nutrient availability for tilapias in the ponds was available.

PATH ANALYSIS

A causal path diagram of the Nile tilapia growth at Huachipa, based on the ordinary Gulland-and-Holt plot, is shown in Fig. 5C. The extended Gulland-and-Holt plot results in a more detailed causal path diagram (Fig. 5D).

For the data from Peru, the multiple regression, based on the extended Gulland-and-Holt plot, results in the following equation:

		mean	range
$\Delta L/\Delta t$ =			
$-6.534 \cdot 10^{-3}$	mean length cm	17	9-22
$+2.778 \cdot 10^{-3}$	water temperature (°C)	22.7	18.8-26.3
$-1.303 \cdot 10^{-6}$	stocking density n·ha	12	$6\text{-}24 \cdot 10^3$
$+1.012 \cdot 10^{-4}$	manure input (wet) $kg \cdot ha^{-1} \cdot day^{-1}$	77	33-205
+0.08829			

...8)

with n = 85, R^2 = 0.504, SEE = 0.0153, P < 0.001, K = 0.00654 day^{-1} and L_∞ = 25.5 (range 23.2 to 32.6 cm).

Three auxiliary variables, two of them treatment variables, increase the amount of explained variation by 7%. Water temperature and wet manure input positively affect growth rate, while increasing density reduces tilapia growth rate. No correlations exist between water temperature, stocking density and manure input. Water temperature and mean length display a certain amount of correlation, attributable to the increase of water temperature as the fish grew.

Zambia

EXTENDED GULLAND-AND-HOLT ANALYSIS

Fig. 4C depicts an ordinary Gulland-and-Holt plot of *O. andersonii* growth in Zambian fishponds. The amount of variation is large, since the data are based on weight and the sampling intervals were large (53 to 151 days; mean = 82 days). The linear regression results in:

$$\Delta L/\Delta t = 0.0721 - 0.00281\ ML \qquad \text{...9)}$$

with n = 84, r^2 = 0.236, SEE = 0.0171, P < 0.001, K = 0.00281 day^{-1} and L_∞ = 25.7 cm.

The total variability in growth rate explained by the model was increased from 24% to 42% through the inclusion of water temperature. The individual contribution of the independent variables to the total explained variation is: mean length 41% and water temperature 23%.

For the data from Zambia, a multiple regression analysis, based on the extended Gulland-and-Holt plot, results in the following equation:

		mean	range
$\Delta L/\Delta t$ =			
$-3.766 \cdot 10^{-3}$	mean length cm	17.1	8-24
$+5.406 \cdot 10^{-3}$	water temperature (°C)	20.4	17.5-23.4
+0.02190			

...10)

with n = 84, R^2 = 0.415, SEE = 0.015, P < 0.001, K = 0.00377 day^{-1} and L_∞ = 23.5 (range 19.3 to 27.8 cm).

PATH ANALYSIS

A causal path diagram of an ordinary Gulland-and-Holt plot describes the decrease in growth rate with increasing fish size (Fig. 5E). The path diagram based on the extended Gulland-and-Holt plot is shown in Fig. 5F. Nearly 60% of the total variation is left unexplained. Water temperature has a positive influence on the growth rate of *O. andersonii*, leading to a proportional increase in growth rate. Neither feed input nor stocking density were significant variables when tested for inclusion. No variables describing pond dynamics were recorded.

Sensitivity Analysis

In the regression model based on the extended Gulland-and-Holt method and the data from Dor station in Israel, water temperature and mean length had the strongest effects (Fig. 6A). Water temperature caused up to 120% change in Nile tilapia growth rate, followed by mean length with 80% change at maximum. The third variable in the equation, the percentage of male tilapia, caused maximally 20% of change in growth rate. The strong effect of water temperature in this regression occurred because at Dor station temperature fluctuates in a wide range over the season. Significant effects on fish growth are therefore detected by the regression procedure and are appropriately quantified by the regression coefficients.

Fig. 6B shows the sensitivity analysis of the equation describing all-male Nile tilapia growth rate in the dataset from Peru. Change of mean length effected the strongest change in growth rate followed by water temperature. The weakest effects resulted from changes in stocking density and manure loading rate. The range of values for the variables was lower in the experiments at Huachipa. In the sensitivity analysis, the changes in growth rate were therefore 100% at most.

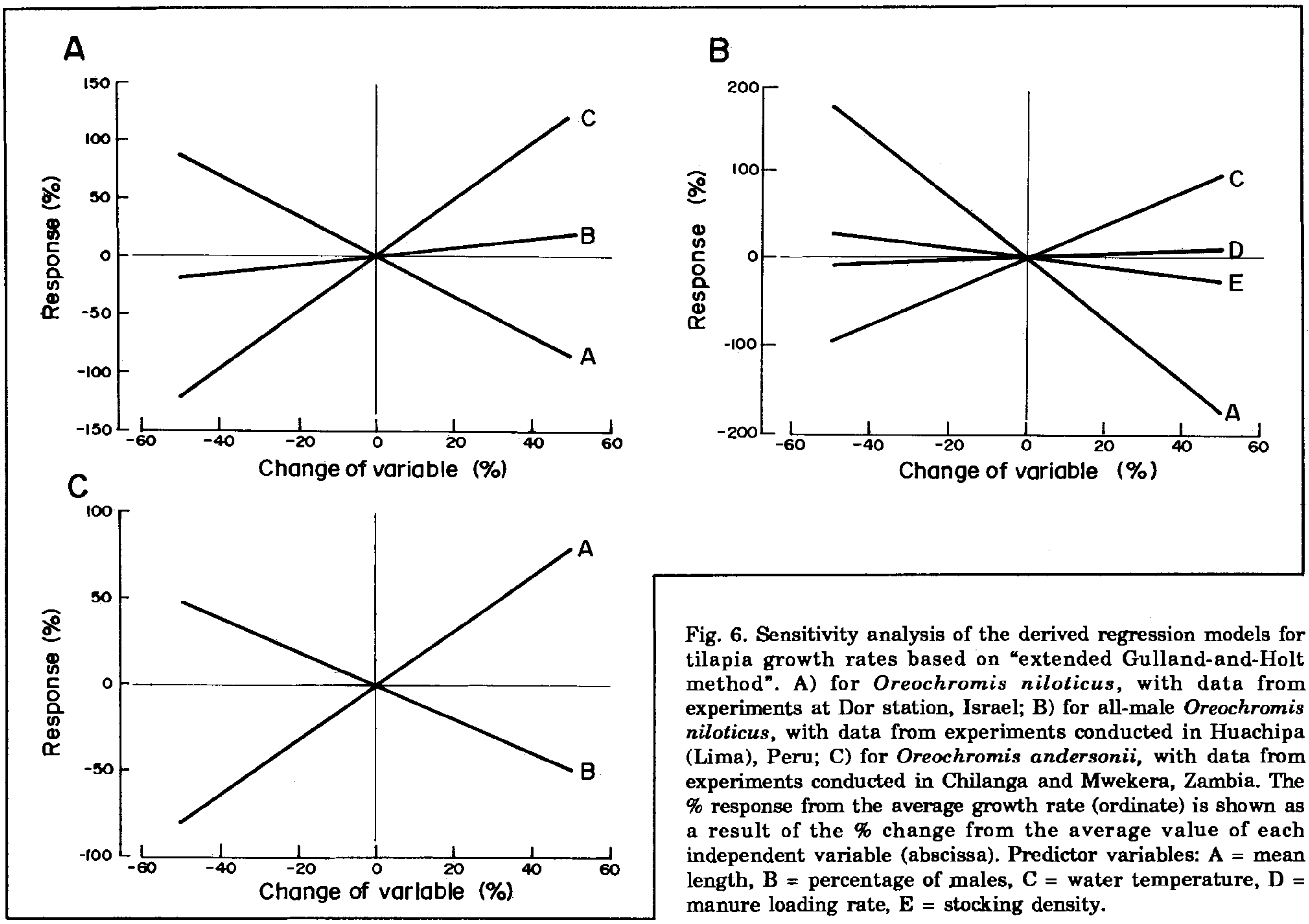


Fig. 6. Sensitivity analysis of the derived regression models for tilapia growth rates based on "extended Gulland-and-Holt method". A) for *Oreochromis niloticus*, with data from experiments at Dor station, Israel; B) for all-male *Oreochromis niloticus*, with data from experiments conducted in Huachipa (Lima), Peru; C) for *Oreochromis andersonii*, with data from experiments conducted in Chilanga and Mwekera, Zambia. The % response from the average growth rate (ordinate) is shown as a result of the % change from the average value of each independent variable (abscissa). Predictor variables: A = mean length, B = percentage of males, C = water temperature, D = manure loading rate, E = stocking density.

The regression model for Mortimer's experiments in Zambia contained only two variables, as the only auxiliary variable found to be significant was water temperature. This caused an 80% change in growth rate when varied by 50%, in contrast to mean length which showed only a 50% change (Fig. 6C).

Comparison of Growth Parameter Estimates

COMPARISON OF GROWTH CURVES

Fig. 7 shows growth curves in length and weight for different tilapia datasets. These growth curves were calculated with the average values of the environmental parameters in the respective datasets. The curves therefore reflect the combined effects of the different culture conditions on tilapia growth.

The two curves describing the lowest growth refer to the experiments with *O. niloticus* in Peru and with *O. andersonii* in Zambia. The tilapia grown at Dor station and the commercial farms showed superior growth (under average culture conditions) compared to those in the Philippines. At Dor station and in Israeli farms, the fish were fed pellets. Also, the tilapia cultured at the farms were all-male hybrids, while those at Dor station and in the Philippines were Nile tilapia of mixed sexes.

COMPARISON OF ϕ' VALUES

Values of the growth performance index ϕ' were computed for all estimates of K and L_∞, based on the average, minimum and maximum values of the environmental and treatment variables (Fig. 8). The minimum and maximum values are theoretical combinations only, based on extremes in the data, and do not necessarily occur under real culture conditions. Rather, the wide range of VBGF parameters reflects the flexibility of the derived equations for modelling tilapia growth.

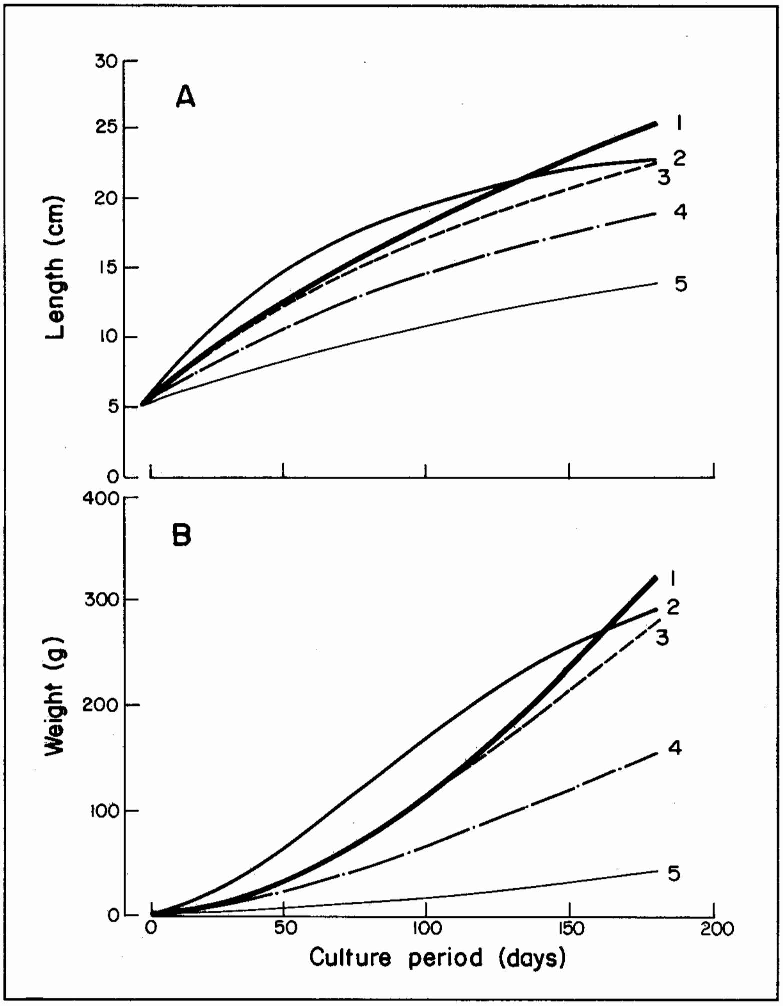


Fig. 7. Theoretical VBGF growth curves in A) length and B) weight, computed over 180 days for the different datasets analyzed with the extended Gulland-and-Holt method. The VBGF parameters are based on the average values for the variables in the datasets. Curves computed with a starting length of 5 cm. Numbers refer to datasets. 1 = farms, Israel (from Prein and Milstein, this vol.); 2 = Dor station, 3 = Philippines (from Prein, this vol.); 4 = Peru; 5 = Zambia. Tilapia species are: 1 = all-male tilapia hybrids, 2 and 3 = mixed-sex Nile tilapia, 4 = all-male Nile tilapia, 5 = mixed-sex *Oreochromis andersonii*.

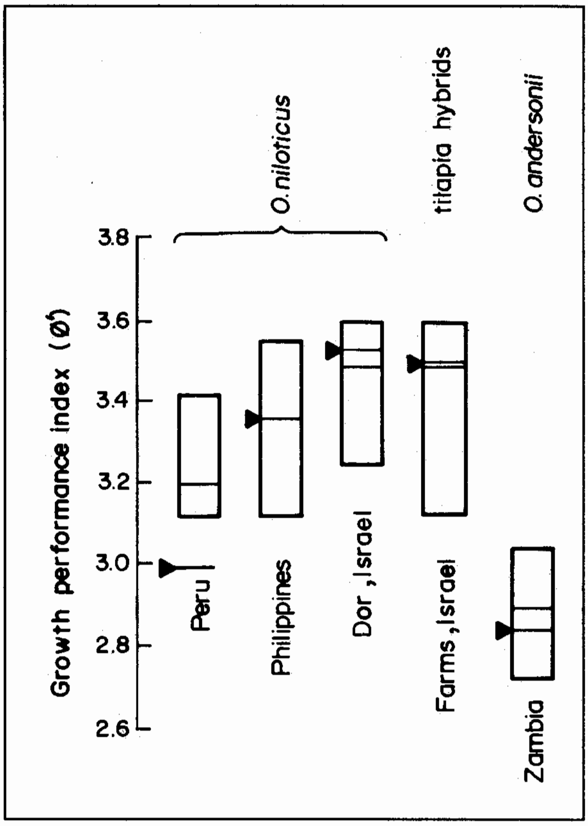


Fig. 8. Averages and theoretically possible ranges of growth performance index ϕ' for different data sources and species. Bars denote minimum, maximum and averages of ϕ' values estimated with extended version of Gulland-and-Holt plot. Marker denotes value of ϕ' estimated with ordinary version of Gulland-and-Holt plot. Data sources are: Huachipa (Peru), Muñoz (Philippines, from Prein, this vol.), Dor station (Israel; from Prein and Milstein, this vol.), commercial farms (Israel), and Chilanga and Mwekera (Zambia).

For Nile tilapia, the overall mean ϕ' is 3.34, with a range from 3.19 to 3.48. The lowest mean value of 3.19 was obtained for the dataset from Peru although this was an all-male monoculture. The value for the Dor station experiments is highest with 3.48; these fish were fed pellets, which may explain the high growth performance.

The ϕ' values determined for the tilapia hybrids grown on the Israeli commercial farms are similar to those obtained for the Dor station experiments. One can consider the conditions at the farms to be far better for fish growth, with feeding of pellets and sorghum, intensive manuring and applied aeration when necessary, than under low energy-input conditions, i.e., in the Philippines (see Prein, this vol.). The lowest value obtained for ϕ' pertain to *O. andersonii* in Zambia.

When ϕ' values are computed using the ordinary Gulland-and-Holt plot, these do not necessarily lie in the same range as those computed with the extended Gulland-and-Holt plot. In the case of the data from Peru, the inclusion of explanatory variables strongly changed the estimated values. The influence of environmental effects is not considered in the ordinary version.

Discussion

With all datasets, significant regression models could be derived based on the multivariate methods presented here and in all cases, length or its transform had a significant relationship, of the appropriate sign, with growth rate. Hence, the underlying von Bertalanffy growth function was validated by this study.

Israel: Dor Station Experiments

The dataset from the Dor station experiments proved to be of limited adequacy for multivariate analysis with the extended Gulland-and-Holt method. This was due to the fact that the experiments were not initially designed for the evaluation of environmental and treatment effects on tilapia growth (Hulata et al. 1986). The experiments were started in mid-summer (July) and lasted until the end of the growing season (November). With increasing fish weight, more dry chicken manure and pellets were applied. No parallel treatments were run with reduced rates of manuring, feeding or stocking density at larger fish sizes. This led to collinearity, i.e. to a close correlation between growth rate, fish sizes, stocking density (as biomass), and manuring and feeding rate. This resulted, for example in the case of pellet feeding rate, in a negative sign for the regression coefficient. This would have meant that feed applications decrease growth rate, obviously an artifact caused by the data (i.e., experimental design). Similarly, solar radiation and water temperature decreased from highest summer values towards low winter values. This also led to a close correlation with growth rate, making difficult the selection of variables for the regression. In the best model obtained, the percentage of males in the Nile tilapia population had a positive influence on overall growth rate, besides water temperature. Male tilapia usually have a higher growth performance than females (Pauly et al. 1988).

These combined effects caused an additional reduction in growth rate with increasing size. This effect can be seen in the ordinary Gulland-and-Holt plot (Fig. 4), which shows a strong concentration of points below the intercept with the regression line, i.e., the estimate of L_∞.

Peru: Experiments in Huachipa

Delgado's experiments in Peru were based on a monoculture of hand-sexed all-male Nile tilapia grown in ponds and concrete tanks which received pig and duck manure as the only nutrient input (Delgado 1985). The net yields achieved ranged from 2 to 12 $kg \cdot ha^{-1} \cdot day^{-1}$ with an average of 5.5 $kg \cdot ha^{-1} \cdot day^{-1}$; these are low to moderate values for manure-based systems (Schroeder 1978). The auxiliary variables which were identified as significantly controlling Nile tilapia growth rate in the experiments were: manure loading rate, stocking density and water temperature.

Huachipa is often covered by clouds and fog, which limit phytoplankton production (Delgado 1985). Although this was not explicitly stated by Delgado (1985), it must be assumed here that the amount of manure is given as wet weight. Pig manure contains approximately 70% moisture (Hopkins and Cruz 1982). On a dry weight basis therefore, the average manure input into the ponds was 23 $kg \cdot ha^{-1} \cdot day^{-1}$, which is considerably lower than the values applied in the experiments in Dor. This explains, in part at least, the low growth rates of tilapias grown at Huachipa.

Zambia: Mortimer's Experiments

The experiments at Chilanga and Mwekera in Zambia with *O. andersonii* were conducted as a polyculture with two other species of tilapia (*O. macrochir* and *T. rendalli*) of mixed sexes. Reproduction occurred early in the other two species, but late (after several months) in *O. andersonii* (Mortimer 1960). Only the growth of *O. andersonii* was analyzed here, since this was the species of prime interest in Mortimer's experiments. No manure was applied in these experiments, i.e., ground maize, maize bran and freshly cut grass were the only nutrient inputs to the ponds. Ground maize was applied at average rates of 36 $kg \cdot ha^{-1} \cdot day^{-1}$ and maize bran at 10 $kg \cdot ha^{-1} \cdot day^{-1}$. Stocking densities ranged from 350 to 26,000 $fish \cdot ha^{-1}$ (all three tilapia species combined) with an average of $4{,}600 \cdot ha^{-1}$, although in most experiments these were reduced by predation. To support good tilapia growth at these densities, the nutrient inputs were too low, at least when compared to the data from Israel. The water temperatures were lower (mean of 20.4°C), compared to locations in Israel (mean of 27.8°C). Together, these facts suggest much poorer growth conditions for tilapia in Mortimer's experiments. The experimental design aimed at determining the possible fish production based on the natural fertility and productivity of the ponds with little or no additional feeding. Net daily yields were consequently low, averaging 0.7 kg·ha^{-1}·day^{-1} with ranges from 0.02 to 1.7 kg·ha^{-1}·day^{-1}.

Only water temperature emerged as a significant predictor of *O. andersonii* growth rate, besides mean length. Neither stocking density nor maize or grass inputs were significant variables. Intervals between sampling events were large (mean 82 days, range 53 to 151 days), leading to reduced resolution of the environmental and treatment effects on growth increment. The weight determinations showed considerable variation, indicating inadequate sampling of the fish. Overall, 60% of the total variance in *O. andersonii* growth rate remained unexplained.

Growth Parameters and ϕ'

The obtained values of the VBGF growth parameters and L_∞ are summarized in Table 1.

as those for all-male Nile tilapia determined by Pauly et al. (1988) who computed ϕ' values for 65 Nile tilapia 'stocks' reared under aquaculture conditions and compared these values with those for natural stocks assembled by Moreau et al. (1986). The values for L_∞ are considerably higher for the commercially produced fish, but this is balanced by smaller values of K.

The values of the growth performance index determined for *O. andersonii* in the Zambian experiments were considerably lower than those determined for all other species (mean 2.88). Pauly et al. (1988) had used the same dataset for computations of ϕ' for individual experiments. Their results match those determined here. A further ϕ' source is given in Pauly et al. (1988) for which a value of 3.4 is estimated. This indicates that this species has a large growth potential under aquaculture conditions, which was not exploited in Mortimer's experiments.

Table 1. Summary of derived values for VBGF growth parameters and index of growth performance ϕ' for the analyzed datasets.

Species**	K(day^{-1})	L_∞(cm; TL)			ϕ'			Source of data	Method*
		min	mean	max	min	mean	max		
O. niloticus M	0.00472		23.89			2.99		Peru	ord. G. & H.
O. niloticus M	0.00654	23.21	25.46	32.64	3.41	3.19	3.41	Peru	ext. G. & H.
O. niloticus	0.01780		22.53			3.52		Dor, Israel	ord. G. & H.
O. niloticus	0.01321	18.99	24.93	28.53	3.24	3.48	3.59	Dor, Israel	ext. G. & H.
O. niloticus	0.00994		25.42			3.37		Philippines	ord. G. & H.
O. niloticus	0.00652	30.76	30.76	38.28	3.11	3.35	3.54	Philippines	ext. G. & H.
tilapia hybrids M	0.00524		39.57			3.48		farms, Israel	ord. G. & H.
tilapia hybrids M	0.00411	29.40	44.50	50.30	3.11	3.47	3.58	farms, Israel	ext. G. & H.
O. andersonii	0.00281		25.66			2.83		Zambia	ord. G. & H.
O. andersonii	0.00377	19.30	23.47	27.80		2.88	3.03	Zambia	ext. G. & H.

*Method used for derivation of parameters (ordinary or extended version of Gulland-and-Holt method).
**M = all-male, otherwise mixed sexes.

The grand mean of the ϕ' values obtained for *O. niloticus* is 3.34, where the dataset from Peru produced the lowest value of ϕ', and the dataset from Dor station the highest value. These reflect the gradient of combined environmental and treatment effects which governed Nile tilapia growth in the datasets. This range of conditions is incorporated in the regression equation and can be used as a flexible management tool for predictions.

The ϕ' values obtained here are in the medium to upper range of published values of ϕ' in *O. niloticus*, suggesting that all three systems investigated here took at least some advantage of the high growth potential of Nile tilapia for aquaculture (Pauly et al. 1988).

The values of ϕ' determined for mixed-sex Nile tilapia in the present study are in the same range

Acknowledgements

Ana Milstein, Gideon Hulata and Giosa Wohlfarth of the Dor Fish and Aquaculture Research Station (ARO) kindly made available their raw experimental data and translated from Hebrew whatever required translating. Amir Halevy kindly provided the water temperature recordings. Further, I wish to acknowledge the Israel Meteorological Service, Bet Dagan, for supplying the climatological data used in this study, through my cooperation with G. Hulata, ARO. 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 for Economic Cooperation (BMZ).

References

Delgado, S. 1985. Crecimiento y produccion de *Tilapia nilotica* (Pisces, Cichlidae) en el laboratorio de Huachipa. Bol. Inst. Mar Peru 9:29-64.

Hepher, B. 1985. Aquaculture intensification under land and water limitations. GeoJournal 10:253-259.

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

Hulata, G., G.W. Wohlfarth and A. Halevy. 1986. Mass selection for growth rate in the Nile tilapia (*Oreochromis niloticus*). Aquaculture 57:177-184.

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.

Moreau, J., C. Bambino and D. Pauly. 1986. Indices of overall growth performance of 100 tilapia (Cichlidae) populations, p. 201-206. *In* J.L. Maclean, L.B. Dizon and L.V. Hosillos (eds.) The First Asian Fisheries Forum. Asian Fisheries Society, Manila, Philippines. 727 p.

Mortimer, M.A.E. 1960. Observations on the biology of *Tilapia andersonii* (Castelnau) (Pisces, Cichlidae), in Northern Rhodesia, p. 42-67. *In* P.B.N Jackson (ed.) Joint Fisheries Research Organisation, Annual Report No. 9, 1959. Government Printing Office, Lusaka. 71 p.

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

Pauly, D. 1979. Gill size and temperature as governing factors in fish growth: a generalization of von Bertalanffy's growth formula. Ber. Inst. Meereskd. Univ. Kiel. No. 63. 156 p.

Pauly, D., J. Moreau and M. Prein. 1988. A comparison of overall growth performance of tilapia in open waters and in aquaculture, p. 469-479. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia and Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Schroeder, G.L. 1978. Autotrophic and heterotrophic production of microorganisms in intensely-manured fish ponds, and related fish yields. Aquaculture 14:303-325.

Thys van den Audenaerde, D.F.E. 1968. An annotated bibliography of tilapia (Pisces, Cichlidae). Zoologische Documentatie 14, Koninklijk Museum voor Midden-Afrika, Tervuren, Belgium. 406 p.

Walter, H. and H. Lieth. 1969. Klimadiagramm Weltatlas. Gustav Fischer, Jena. pag. var.

Multiple Regression Analysis of Growth of *Tilapia rendalli* in Polyculture with *Oreochromis shiranus* as Affected by Water Quality and Pond Inputs*

BARRY A. COSTA-PIERCE and ANNE A. VAN DAM,
ICLARM/GTZ Africa Aquaculture Project, P.O. Box 229, Zomba, Malawi

MICHAEL V. KAPELETA, Malawi Department of Fisheries, Domasi Experimental Fish Farm, P.O. Box 44, Domasi, Malawi

Costa-Pierce, B.A., A.A. van Dam and M.V. Kapeleta. 1993. Multiple regression analysis of growth of *Tilapia rendalli* in polyculture with *Oreochromis shiranus* as affected by water quality and pond inputs, p. 88-96. *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

Multiple regression analyses were done to quantify the combined impacts on growth of two native tilapias (*Tilapia rendalli* and *Oreochromis shiranus*) used for aquaculture in Malawi. Water quality changes and a diverse combination of low cost pond inputs widely available on smallholder farms were analyzed to help structure future on-station pond and water quality research programs.

A model was identified which related specific growth rate (SGR; %/day in fortnightly sampling periods) of *T. rendalli* to initial weights, pond stirring and applications of wood ash, urea, maize bran and napier grass (adjusted $R^2 = 0.69$; $P < 0.001$; $n = 64$). The corresponding model for *Oreochromis shiranus*, identified maize bran, napier grass and fish initial weights as significantly related to SGR (adjusted $R^2 = 0.15$; $P < 0.05$; $n = 72$).

Other models, relating stirring and inputs as independent variables to dissolved oxygen, pH, conductivity, total hardness and total alkalinity showed that dissolved oxygen and pH were significantly affected by maize bran and grass ($R^2 = 0.59$ and 0.43; both $P < 0.001$). Ash and stirring significantly affected conductivity ($R^2 = 0.77$); total hardness ($R^2 = 0.39$); and total alkalinity ($R^2 = 0.86$) (all models $P < 0.001$; all predictors, $P < 0.05$). A model fitting growth data for *O. shiranus* showed dissolved oxygen and fish initial weight as the most significant predictor variables (adjusted $R^2 = 0.10$; $P < 0.05$); and total hardness, dissolved oxygen, temperature and fish initial weight for *T. rendalli* (adjusted $R^2 = 0.47$; $P < 0.001$).

*ICLARM Contribution No. 732. An earlier version of this paper was presented by M.V. Kapeleta at the Conference on "Research for the Development of Tropical Aquaculture Technology Appropriate for Implementation in Rural Africa" 2-6 April 1990, ICLARM-GTZ/Malawi Department of Fisheries/University of Malawi, Zomba, Malawi

Introduction

Average fish yields in aquaculture ponds in Malaŵi are low, between 400 and 500 $kg \cdot ha^{-1} \cdot year^{-1}$ (Msiska 1985; Satia 1989). A number of reasons could explain this low production, such as the low alkalinity of natural waters, high costs and inavailability of suitable pond inputs, household economic factors and sociocultural constraints (ICLARM and GTZ 1991), or genetic deterioration caused by poor farming practices. Detailed analyses of water quality variables to assess the suitability of Malaŵi's natural waters for aquaculture and scientific studies of potential pond inputs available on farms have heretofore not been attempted.

The potential of multivariate models in aquaculture research has been demonstrated (Pauly and Hopkins 1983; Hopkins et al. 1988; van Dam 1989, 1990a; Capili et al. 1990). While there is an urgent need to identify a compatible spectrum of low cost, on-farm resources for use in smallholder aquaculture ponds in Malaŵi, the potential combinations that need to be evaluated scientifically by traditional "treatment testing" are endless.

The objective of conventional aquaculture experimental design is to remove as much as possible the inherent variability between treatments so that treatment effects can be observed. More often than not, fish yield characteristics obtained by testing various treatments are so variable that conventional parametric and nonparametric statistics cannot be used. Some reasons for inconclusive results are the inherent variabilities of the aquatic ecosystem, time variabilities of biological processes, fish growth depensation, water quality differences, and a host of other factors. Conventional experimental designs are very expensive in research time and space, and often have little power (van Dam 1990b).

Multivariate analyses can effectively relate water quality, time, input and other predictor variables to fish growth (Pauly and Hopkins 1983). Multivariate models attempt to explain and make use of observed variances by combining and explaining the variability involved in all production processes rather than eliminating variability. Construction of separate treatments with large numbers of replicates are secondary considerations.

In this paper, multiple regression analyses of water quality parameters, a novel combination of isonitrogeneous inputs derived from, and widely available on, smallholder farms and pond stirring were accomplished to determine factors affecting the growth of *Tilapia rendalli* and *Oreochromis shiranus* in polyculture in experimental ponds in Malaŵi.

Materials and Methods

Combinations of on- and off-farm resources and pond stirring were used as inputs to 200-m^2 aquaculture ponds at the Domasi Experimental Fish Farm (DEFF). One hundred mixed-sex fingerlings of *Tilapia rendalli* and *Oreochromis shiranus* were stocked at a combined mean size (±1 SD) of 50.9 ± 1.3 g in 27 ponds at a 1:1 ratio in nine treatments with three replicates per treatment. Individual mean fish weights at stocking ranged from 31.2 g to 52.9 g ± 0.35 g for *T. rendalli* and 46.0 g to 55.6 g ± 3.6 g for *O. shiranus*. Treatments were: (i) stirring; (ii) urea; (iii) maize bran; (iv) napier grass; (v) wood ash; (vi) an isonitrogeneous combination of all the above (including stirring); (vii) a combination without urea; (viii) a combination without stirring; and (ix) a zero input control treatment. Inputs were proportionally combined or used alone to attain an isonitrogeneous loading rate of approximately 13 $gN{\cdot}day^{-1}$ in each 200 m^2 pond (Table 1).

Fish growth in all ponds was evaluated fortnightly, i.e., fifteen fish of each species from seine net samples were individually weighed. Six water quality and physical parameters were measured at 0600-0800 hours in one pond of each treatment at fortnightly intervals using instruments and methods detailed in Table 2; however, the monitoring of pH was not continued after day 44. Water samples for total hardness and total alkalinity were taken by hand using clean, inert plastic bottles at 3 to 5 cm water depth. Analyses were completed within one hour of sampling under cover at the pond site.

Table 1. Application rates to achieve approximately isonitrogeneous input in small-scale aquaculture ponds.

Treatments	Application rates ($kg{\cdot}200\ m^{-2}{\cdot}day^{-1}$)				$g{\cdot}200\ m^{-2}{\cdot}day^{-1}$	$kgN{\cdot}ha^{-1}{\cdot}day^{-1}$
	Maize bran	Grass	Urea	Wood ash		
Stirring/urea/maize bran/ napier grass/wood ash	0.5	2.3	0.019	1.0	13.6	0.68
Urea/maize bran/wood ash/ napier grass	0.5	2.3	0.019	1.0	13.6	0.68
Maize bran	1.4	-	-	-	13.0	0.65
Napier grass	-	7.0	-	-	13.6	0.68
Urea	-	-	0.056	-	13.2	0.66
Maize bran/napier grass/ wood ash/stirring	0.7	3.5	-	1.0	13.3	0.67
Wood ash	-	-	-	1.0	-	-

Nitrogen levels calculated from protein in maize bran after Kadongola (1990); protein in napier grass after Gohl (1975); and from manufacturers' specifications in urea.

Table 2. Water quality methods and instruments used, at fortnightly intervals, from 0600-0800 hours, during the experimental period in Zomba, Malaŵi.

Parameters	Units	Method
Total alkalinity	$mg{\cdot}l^{-1}$ as $CaCO_3$	Hach digital titrator
Total hardness	$mg{\cdot}l^{-1}$ as $CaCO_3$	Hach digital titrator
Dissolved oxygen	$mg{\cdot}l^{-1}$	YSI meter
Temperature	°C	YSI meter
Conductivity	µS	WTW LF 91 meter
pH	unit	Horiba meter

Meters with probes were used *in situ* at 3-5 cm water depth.

Pond stirring was done twice a week at 1400-1500 hours with the use of a 1-m wide weighted pond rake. For each pond stirred, pond bottoms were raked with five "pulls" (Fig. 1; Costa-Pierce 1991). The rationale for simple pond stirring was to increase nutrient, detrital and direct foods to the pelagic zone of the ponds; potential benefits of stirring have been discussed in detail by Costa-Pierce and Pullin (1989).

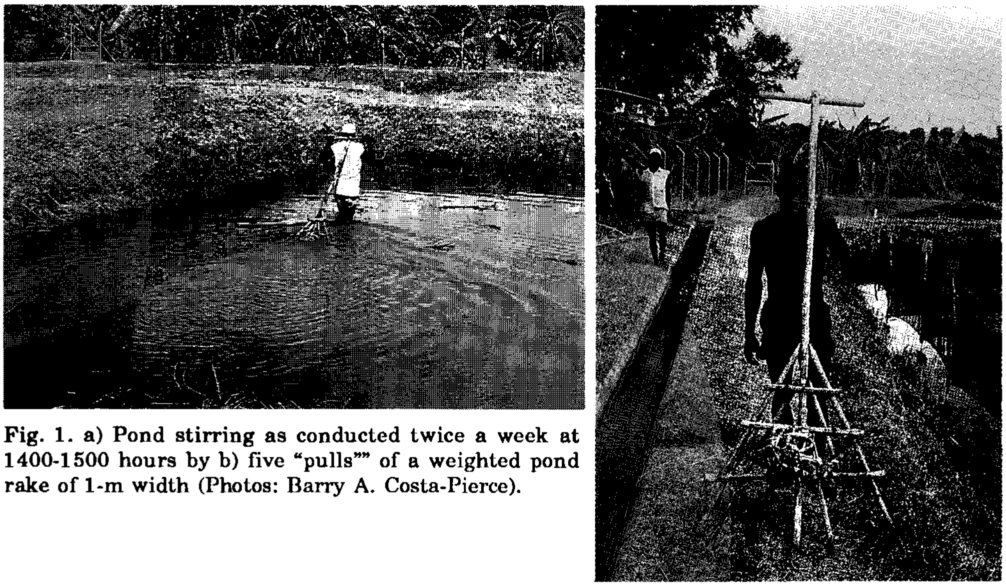


Fig. 1. a) Pond stirring as conducted twice a week at 1400-1500 hours by b) five "pulls"" of a weighted pond rake of 1-m width (Photos: Barry A. Costa-Pierce).

All water quality parameters and body weights for both fish species were measured eight times in 16 weeks, i.e., fortnightly. For predicting pH after day 44, a multiple regression equation with temperature and dissolved oxygen (DO) as predictor variables was computed using pH data before day 44. Fish specific growth rates (SGR) in per cent were calculated according to Brett and Groves (1979):

$$SGR = \frac{(\ln W_2 - \ln W_1) \cdot 100}{t_2 - t_1} \quad \text{...1)}$$

Water quality, pond inputs and fish growth parameters were entered into Microstat (Ecosoft, Inc.) software on an IBM compatible personal computer. Stirring was added as a dummy variable (Yamane 1973). All ponds receiving mixtures of inputs were entered together with ponds receiving single inputs into two datasets comprising all water quality, input and SGRs for 27 ponds. Dummy variables were used to complete the dataset where individual inputs were missing. Therefore, a complete dataset for each fish species contained 72 cases (8 samplings x 9 ponds) including specific growth rate, water quality and pond input values for every two-week sampling period.

After inputting, the frequency distributions of the dependent variables (SGRs for both fish) were plotted, and a Chi-square goodness of fit test was run to test for normality. The SGR for *O. shiranus* was found to be normally distributed ($P<0.05$), but the SGR for *T. rendalli* was not ($P > 0.05$). Further examination of the dataset for *T. rendalli* showed eight cases were negative specific growth rates (-0.40 to -1.14%·day^{-1}), followed by high values of SGR (2.03 to 4.71%·day^{-1}). It was decided to eliminate these values from the dataset because of presumed sampling bias, since it was known previously that specific growth of *T. rendalli* rates using these low quality inputs never reached these levels at Domasi (Chikafumbwa 1990) and that fish shrinkage to the extent suggested by these negative growth rates was highly unlikely. For these reasons, eight cases were deleted from the dataset for *T. rendalli*. After deletion, a Chi-square test was run for 64 cases which showed that the data did not significantly differ from normality ($P < 0.05$).

A complete correlation matrix was accomplished for both datasets to determine if predictor variables were correlated, which would pose a danger of multicollinearity (Hopkins et al. 1988).

The multiple regression approach used here was an "exploratory data analysis" (EDA) as opposed to a "confirmatory" one (Prein et al., this vol.). It was our intention to fit a regression model to a relatively small dataset where a large range of inputs was used to infer which pond inputs were significant in accounting for the observed variances in SGR and water quality parameters. The EDA approach was used to give insight into new areas for on-station research and to better plan water quality monitoring programs and experimental designs.

Three groups of models were constructed as follows: (i) models relating the initial weight of fish at each period and pond inputs to the SGR of both species. Fish initial weights were included because differences in initial fish weights would help explain differences in growth rates; (ii) models relating pond inputs to individual water quality parameters; and (iii) models relating water quality parameters to SGR of both fish species.

Standardized partial regression coefficients or β-weights, were computed for all models from:

$$\beta = b \cdot (sd_x/sd_y) \qquad \text{...2)}$$

where sd_x and sd_y are the standard deviations of the independent and dependent variables, respectively, and b is the partial regression coefficient. β-weights are presented in form of bar graphs to show the relative importance of independent variables that are expressed in different units (Yamane 1973).

For every model, the Durbin-Watson statistic was calculated to test for positive or negative serial correlation. Where comparisons between models for *T. rendalli* and *O. shiranus* were done, the adjusted R^2 was used to eliminate any bias due to the different numbers of cases included in the models. Residuals of every model were plotted against independent and predicted variables to check the assumptions of zero mean error and constant error variance (Yamane 1973; Dillon and Goldstein 1984).

The data used in these analyses are available in files TRENDALL.WK1 and OSHIRAN.WK1 and are described in Appendix II.

Results

Descriptive statistics of all variables are given in Table 3. Overall, water quality was not stressful for either fish species during the period; however, early morning dissolved oxygen values (DO) measured at 0500-0700 hours did drop as low as 1.6 $mg \cdot l^{-1}$ (Table 3).

Correlation matrices between all 20 variables (72 cases for *O. shiranus* [Table 4] and 64 for *T. rendalli* [Table 5]) showed that significant correlations were present between conductivity, total hardness and total alkalinity; and between DO and pH. Significant correlations also existed between stirring and ash inputs and conductivity; total hardness and total alkalinity; and maize bran, dissolved oxygen and pH.

A model of SGR against stirring, inputs and initial weight of *T. rendalli* showed that growth was significantly and negatively affected by

Table 3. Descriptive statistics of all variables used in the present analysis.

Variable	Units	N	Mean	Std.dev.	Min.	Max
Stirring	(dummy)	72	0.2	0.42	0.0	1.0
Ash	($kg \cdot day^{-1}$)	72	5.6	5.00	0.0	10.0
Urea	($kg \cdot day^{-1}$)	72	0.1	0.18	0.0	0.6
Grass	($kg \cdot day^{-1}$)	72	20.7	22.68	0.0	70.0
Maize bran	($kg \cdot day^{-1}$)	72	4.2	4.55	0.0	14.0
Conductivity	(μS)	72	52.8	20.93	24.3	104.1
Total hardness	($CaCO_3$, $mg \cdot l^{-1}$)	72	19.6	8.53	5.0	41.0
Total alkalinity	($CaCO_3$, $mg \cdot l^{-1}$)	72	22.7	9.15	8.8	47.1
Dissolved oxygen	($mg \cdot l^{-1}$)	72	4.8	2.04	1.6	9.1
Temperature	(°C)	72	26.4	0.75	24.4	27.5
pH	-	72	7.4	0.44	6.7	8.9
Initial TR weight[a]	(g)	64	64.5	15.35	31.0	100.6
Initial OS weight[a]	(g)	72	66.0	13.86	32.8	100.7
T. rendalli SGR	(%/day)	64	0.41	0.85	-0.9	3.7
O. shiranus SGR	(%/day)	72	0.29	0.88	-2.6	3.8

[a]TR = *T. rendalli*; OS = *O. shiranus*.

Table 4. Correlation matrix of all 21 variables used in the present analysis for derivation of model for specific growth rate of *Oreochromis shiranus* (n = 72).[a)]

Variable[a)]	stir	ash	urea	grass	m.bran	conduct	ammonia	th	ta	do	temp	pH	iwoss	fwoss	sgr
stir	1.000														
ash	.478	1.000													
urea	-.028	-.178	1.000												
grass	.198	.126	-2.950	1.000											
m. bran	.210	.143	-.285	-.071	1.000										
conduct	.665	.808	-.204	.269	.225	1.000									
ammonia	.146	-.014	.103	-.041	-.003	.068	1.000								
th	.489	.572	-.145	.128	.111	.570	.422	1.000							
ta	.716	.852	-.208	.234	.246	.917	.013	.615	1.000						
do	-.305	-.189	.303	-.429	-.601	-.356	-.709	-.433	-.302	1.000					
temp	.051	-.085	-.121	-.011	.091	-.036	.307	.386	-.102	-.360	1.000				
pH	-.146	-.045	.262	-.373	-.508	-.128	-.040	-.253	-.087	.816	-.178	1.000			
iwoss	-.175	-.080	-.237	.338	.344	-.037	.071	.127	-.125	-.614	.498	-.568	1.000		
fwoss	-.216	-.131	-.268	.333	.345	-.037	.071	.127	-.135	-.617	.466	-.586	.865	1.000	
sgr	-.072	-.090	-.037	.063	.073	-.008	-.001	-.097	-.029	.024	-.124	-.010	-.296	.189	1.000

CRITICAL VALUE (1-TAIL, .05) = +/- .196
CRITICAL VALUE (2-TAIL, .05) = +/- .232

[a)]conduct = conductivity; th = total hardness; ta = total alkalinity; do = dissolved oxygen; temp = water temperature; m.bran = maize bran; iwoss = initial weight *O. shiranus*; fwoss = final weight *O. shiranus*.

Table 5. Correlation matrix of all 21 variables used as initial dataset for derivation of model for specific growth rate or *Tilapia rendalli* (n = 64).[a)]

Variable[a)]	stir	ash	urea	grass	m.bran	conduct	ammonia	th	ta	do	temp	pH	fwtr	iwtr	sgr
stir	1.000														
ash	.520	1.000													
urea	-.030	-.191	1.000												
grass	.238	.227	-.294	1.000											
m. bran	.186	.130	-.324	-.054	1.000										
conduct	.676	.814	-.221	.351	.211	1.000									
ammonia	.131	-.021	.051	.000	-.050	.030	1.000								
th	.482	.621	-.204	.239	.070	.556	.350	1.000							
ta	.745	.857	-.231	.315	.234	.914	-.033	.621	1.000						
do	-.256	-.218	.354	-.511	-.604	-.342	.016	-.366	-.293	1.000					
temp	-.029	-.036	-.199	.044	.055	-.056	.219	.334	-.111	-.289	1.000				
pH	-.114	-.051	.279	-.370	-.510	-.106	.004	-.212	-.066	.825	-.167	1.000			
fwtr	-.275	.038	-.202	.524	.308	-.001	-.007	.200	-.070	-.622	.342	-.512	1.000		
iwtr	-.269	.015	-.189	.474	.158	-.048	.018	.265	-.103	-.557	.441	-.432	.912	1.000	
sgr	.036	.025	.041	-.053	.264	-.117	-.196	-.323	-.115	.052	-.467	-.010	-.178	-.542	1.000

CRITICAL VALUE (1-TAIL, .05) = +/- .208
CRITICAL VALUE (2-TAIL, .05) = +/- .246

[a]conduct = conductivity; th = total hardness; ta = total alkalinity; do = dissolved oxygen; temp = water temperature; m.bran = maize bran; iwtr = initial weight *T. rendalli*; fwtr = final weight *T. rendalli*.

stirring and fish initial weight, but positively affected by wood ash, urea, maize bran and napier grass (adjusted R^2 = 0.69; model, $P < 0.001$; predictors, $P < 0.05$; N = 64). A similar model for *O. shiranus* showed significant negative growth responses only for initial weight, but positive effects of maize bran and napier grass (adjusted R^2 = 0.15; model and predictors, $P < 0.05$; N = 72) (Table 6). In both models, initial weight was the most important. Grass and maize bran were stronger predictor variables than ash and urea (see β-weights, Fig. 2).

Both pH and DO were negatively affected by maize bran and napier grass inputs (R^2 = 0.43 and 0.59; both predictors and models, $P < 0.001$) (Table 7 and Fig. 3). Conductivity (R^2 = 0.77), total hardness (R^2 = 0.39) and total alkalinity (R^2 = 0.86) were positively affected by stirring and wood ash (all models, $P < 0.001$; all predictors $P < 0.05$). Urea did not have a significant effect on any water quality parameter.

A multiple regression of growth versus water quality parameters showed that dissolved oxygen had a consistently negative relationship with growth rates of both *T. rendalli* and *O. shiranus* (Table 8 and Fig. 4). Growth of *T. rendalli* was also significantly affected by total hardness and temperature ($P < 0.05$).

Residual plots showed no systematic patterns, therefore it was assumed that the zero mean error and constant error variance assumptions were not violated. The Durbin-Watson test showed that there was no serial correlation of the residuals in five of our models, while the test was inconclusive in the other four. Three models (for DO, pH and alkalinity) showed positive serial correlation. For example, in the model of DO with stirring and pond inputs, the Durbin-Watson statistic was 1.1675, which is smaller than the significance level d_L = 1.49 from the Durbin-Watson table (e.g., in Yamane 1973, N = 75 and five independent variables; α = 0.05), but also smaller than $4\text{-}d_U$ (4-1.77 = 2.23). This means that there is a positive but no negative serial correlation (α = 0.05, Yamane 1973). Serial correlation may have an effect on the variance of the partial regression coefficients (b's) but it does not affect the estimation of the b itself. As most b's were significantly different from zero (at α = 0.001), we did not attempt to remove the effect of the serial correlations.

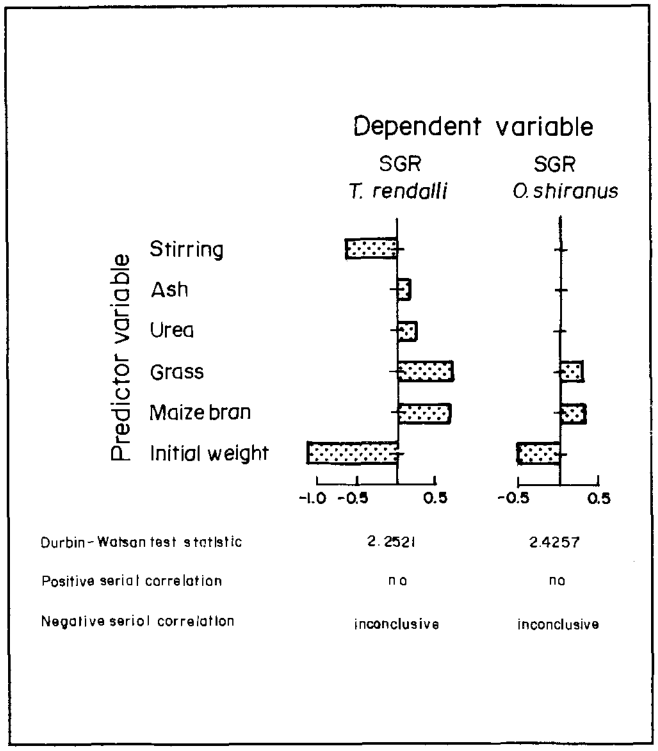


Fig. 2. Standardized partial regression coefficients (beta-weights) for variables in the models of specific growth rate as a result of pond inputs (see Table 6). Only β-weights of significant variables (α = 0.05) are shown.

Table 6. Multiple regression models for growth of *Tilapia rendalli* (n = 64) and *Oreochromis shiranus* (n = 72), with initial fish weight, stirring and inputs as independent variables. The dependent variable is SGR. Significance levels of the independent variables are indicated with stars (* = 5%, ** = 1% *** = 0.1%); b = regression coefficient; s.e. = standard error of the regression coefficient.

	Tilapia rendalli		*Oreochromis shiranus*	
	b	s.e.	b	s.e.
Independent variables				
Initial weight fish	-0.0605	0.0053***	-0.0333	0.0082***
Maize bran	0.1186	0.0150***	0.0604	0.0250*
Grass	0.0266	0.0038***	0.0114	0.0050*
Urea	1.1929	0.3668**	-0.0492	0.6034
Ash	0.0301	0.0144*	-0.0176	0.0223
Stirring	-1.2489	0.1994***	-0.5045	0.2865
Constant (a)	3.2782		2.2184	
Adjusted R^2	0.6906		0.1520	
F-value	24.4710		3.1220	
Durbin-Watson statistic	2.2521		2.4257	
Probability	<0.001		<0.05	

Table 7. Multiple regression models for water quality parameters with stirring and inputs as independent variables (n = 72). Significance levels of the independent variables are indicated with stars (* = 5%, ** = 1% *** = 0.1%).

	DO		pH		Conductivity		Total hardness		Total alkalinity	
	b	s.e.	b	s.e.	b	s.e.	b	s.e.	b	s.e.
Independent variables										
Stirring	-0.3960	0.4571	0.0188	0.1154	16.4747	3.4827***	7.3218	2.8713*	8.2736	1.1885***
Ash	-0.0024	0.0372	0.0069	0.0094	2.5954	0.2833***	0.8938	0.2336	1.1796	0.0097***
Urea	-0.1439	1.0076	-0.0199	0.2544	-3.4163	7.6771	-4.2511	6.3293	-2.1889	2.6199
Grass	-0.0414	0.0078***	-0.0083	0.0020***	0.1126	0.0590	-.0034	0.0487	0.0284	0.0202
Maize bran	-0.2778	0.0386***	-0.0535	0.0098***	0.3079	0.2942	-0.0760	0.2426	0.1348	0.1004
Constant (a)	6.9574		7.7441		31.4832		11.3555		13.376	
Adjusted R^2	0.5915		0.4340		0.7747		0.3924		0.8627	
F-value	19.115		10.122		45.385		8.523		82.963	
Durbin-Watson	1.1675		1.3402		2.2706		1.8100		1.1720	
Probability	<0.001		<0.001		<0.001		<0.001		<0.001	

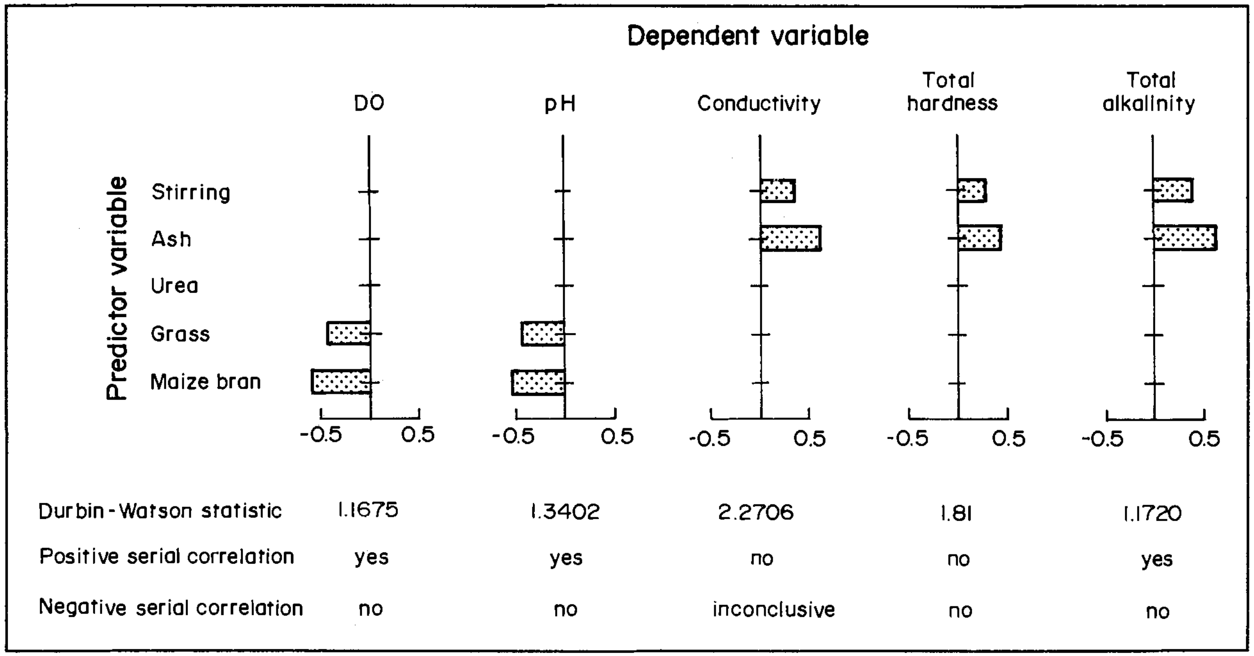


Fig. 3. Standardized partial regression coefficients (beta-weights) for variables in the models of water quality against pond inputs (see Table 7). Only β-weights of significant variables (α = 0.05) are shown.

Table 8. Multiple regression models for growth of *Tilapia rendalli* (n = 64) and *Oreochromis shiranus* (n = 72), with water quality parameters as independent variables. Significance levels of the independent variables are indicated with stars (* = 5%, ** = 1% *** = 0.1%).

	Tilapia rendalli b	*Tilapia rendalli* s.e.	*Oreochromis shiranus* b	*Oreochromis shiranus* s.e.
Independent variables				
Total hardness	-0.0210	0.0088*	-0.0158	0.1120
Dissolved oxygen	-0.1882	0.0476***	-0.1487	0.0679*
Temperature	-0.2950	0.1283*	0.0953	0.1630
Initial weight	-0.0346	0.0065***	-0.0327	0.0101**
Constant (a)	11.7114		0.9193	
Adjusted R^2	0.4723		0.1020	
F-value	15.0950		3.0160	
Durbin-Watson	1.9552		2.3624	
Probability	<0.001		<0.05	

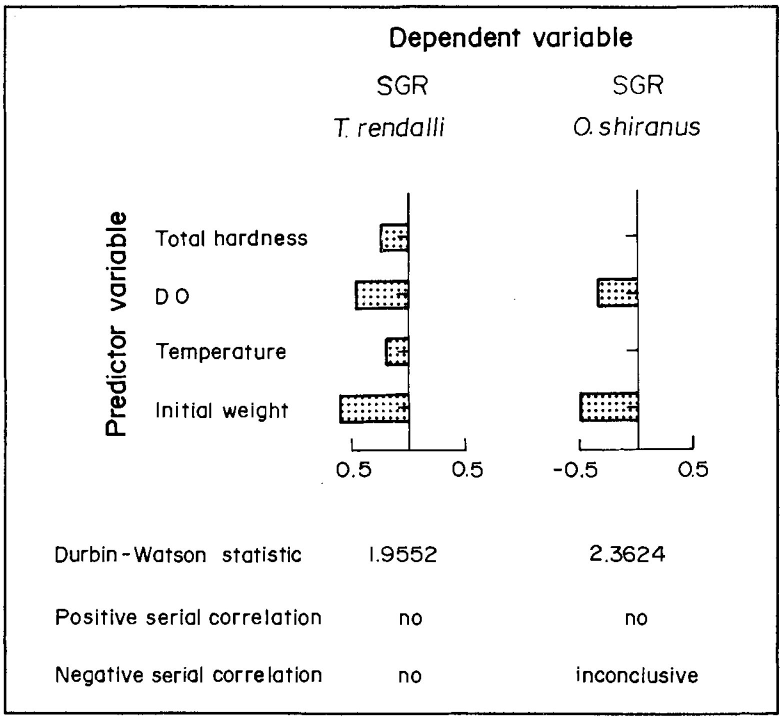


Fig. 4. Standardized partial regression coefficients (beta-weights) for variables in the models of fish growth against water quality (see Table 8). Only β-weights of significant variables (α = 0.05) are shown.

Discussion

Balarin (1988) stated that *Tilapia rendalli* is likely the second most frequently cultured species in Africa after *Oreochromis niloticus*. *T. rendalli* has been classically described as a macrophytic herbivore and has been imported into many nations for controlling aquatic weeds (Junor 1969; Bell-Cross 1976; Rashidi, in press). Our multiple regression model of growth in *T. rendalli* explained 69% of the variance in SGR. The regression model with inputs as predictor variables identified napier grass as the most important predictor variable of SGR in *T. rendalli*. Maize bran is the most commonly used feed for fish in Malaŵi (Banda 1991). In a previous experiment, fish production (*T. rendalli* and *O. shiranus* in a 50:50% polyculture) showed no difference in production in 200-m^2 ponds fed napier grass or maize bran as the sole pond input (Chikafumbwa 1990).

Maize bran and napier grass significantly reduced DOs and pH, while wood ash and stirring increased the ionic content of the water as noted by the increased conductivity, total hardness and total alkalinities. Wood ash has been shown to affect pH and total alkalinity significantly in the laboratory at rates encompassing those used in this study (Jamu 1990). Due to the strong intercorrelations and responses in the models noted here as a result of the suite of inputs used, future water quality sampling programs will need to measure only one of these water quality parameters (conductivity, total hardness and total alkalinity).

DO was significantly decreased by additions of maize bran and napier grass in the models. Maize bran is normally added to fishponds in Malaŵi at 3% fish body weight per day (BWD) (Kadongola 1990). In this study, maize bran was fed at a very high rate (>20% BWD). This was done in order to attempt isonitrogeneous loadings in all treatments, except for wood ash, which only supplies traces of nitrogen. Maize bran has a low nitrogen content, approximately 1% N by dry weight (Kadongola 1990), so the very high rate applied likely accounted for the impact on DO. Napier grass, on the other hand, was fed at a lower rate than recommended by Chikafumbwa (1990), so likely affected DO levels to a lesser degree.

Stirring a highly significant negative impact on the SGR of *T. rendalli*. A similar negative effect was noted on *O. shiranus*, but this was not significant (P > 0.05). Costa-Pierce and Pullin (1989) hypothesized that pond stirring would be beneficial where ponds were previously loaded with organic matter and nutrients and benthic foods were concentrated in a stratified pond hypolimnion. In this experiment, the sediments from a previous pond experiment were completely removed from ponds before the experiment started. Here, stirring could have leached ions from soils suspended into the pond water column, thus increasing total alkalinity, hardness and conductivity. The negative effects of stirring on fish growth could also have resulted from suspended soil particles which decreased light penetration and primary production of phytoplankton, thereby decreasing fish growth, or by suspending nutrient poor particles in the water, affecting fish health and food assimilation. Thus, stirring may be effective as a means to increase SGR only in organically rich ponds.

Acknowledgements

We are grateful for the professional assistance of Dr. Reg Noble in the water quality monitoring program. The experiment was made possible through financial assistance from the Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH, Federal Republic of Germany for the ICLARM/Malaŵi Department of Fisheries/ University of Malaŵi project, "Research for the Development of Aquaculture Technologies Appropriate for Rural Africa". A.A. van Dam is seconded to ICLARM by the Directorate General for International Cooperation, Government of the Netherlands.

References

Balarin, J.D. 1988. Development planning for tilapia farming in Africa, p. 531-538. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Banda, G.A. 1991. Farm household decisionmaking with reference to fish farming in Zomba District, Malaŵi, p. 9-10. *In* B.A. Costa-Pierce, C. Lightfoot, K. Ruddle and R.S.V. Pullin (eds.) Aquaculture research and development in rural Africa. ICLARM Conf. Proc. 27, 52 p.

Bell-Cross, G. 1976. The fisheries of Rhodesia. National Museum and Monuments of Rhodesia, Salisbury, Rhodesia. 262 p.

Brett, J.R. and T.D.D. Groves. 1979. Physiological energetics, p. 280-352. *In* W.S. Hoar, D.J. Randall and J.R. Brett (eds.) Fish physiology. Vol. VIII. Bioenergetics and growth. Academic Press, New York.

Capili, J.B., S.M. Luna and M.L.D. Palomares. 1990. The growth of three strains of tank-reared tilapia *Oreochromis niloticus*, p. 425-428. *In* R. Hirano and I. Hanyu (eds.) The Second Asian Fisheries Forum. Asian Fisheries Society, Manila. 991 p.

Chikafumbwa, F.J.K.T. 1990. Studies on napier grass (*Pennisetum purpureum*) as a pond input for the culture of *Tilapia rendalli* and *Oreochromis shiranus*. Department of Biology, University of Malaŵi. 177 p. M.Sc. thesis.

Costa-Pierce, B.A. 1991. Some impacts of manual pond stirring on aquaculture ponds in Malaŵi, p. 29. *In* B.A. Costa-Pierce, C. Lightfoot, K. Ruddle and R.S.V. Pullin (eds.) Aquaculture research and development in rural Africa. ICLARM Conf. Proc. 27, 52 p.

Costa-Pierce, B.A. and R.S.V. Pullin. 1989. Stirring ponds as a possible means of increasing aquaculture production. Aquabyte 2(3): 5-7.

Dillon, W.R. and M. Goldstein. 1984. Multivariate analysis. Methods and applications. John Wiley, New York. 587 p.

Gohl, B. 1975. Tropical feeds. FAO, Rome. 661 p.

Hopkins, K.D., M.L. Hopkins and D. Pauly. 1988. A multivariate model of tilapia growth, applied to seawater tilapia culture in Kuwait, p. 29-39. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. MacLean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

ICLARM and GTZ. 1991. The context of small-scale integrated agriculture-aquaculture systems in Africa: a case study of Malaŵi. ICLARM Stud. Rev. 18, 302 p.

Junor, F.J.R. 1969. *Tilapia melanopleura* Dam. *In* Artificial lakes and dams in Rhodesia with special reference to its undesirable effects. Rhodesian J. Agric. Res. 7:61-68.

Jamu, D. 1990. Studies on ash as a liming agent in fish ponds. Department of Biology, University of Malaŵi. 193 p. M.Sc. thesis.

Kadongola, W.K. 1990. Maize (*Zea mays*, Linnaeus) bran as a supplemental feed in the culture of *Tilapia rendalli* (Boulenger) and *Oreochromis shiranus* (Boulenger). University of Malaŵi. 177 p. M.Sc. thesis.

Msiska, O.V. 1985. Aquaculture and rural development in Malaŵi. FAO Aquaculture for Local Community Development Programme: Technical Consultation in Rural Development, Lusaka, Zambia, 27-30 October 1987. 13 p. (mimeo).

Pauly, D. and K.D. Hopkins 1983. A method for the analysis of pond growth experiments. ICLARM Newsl. 6:10-12.

Rashidi, B.B.A. A synoptic review of the biology and culture of *Tilapia rendalli* Boulenger (1896). FAO Fish. Synop. Ser. (In press).

Satia, B. 1989. A regional survey of the aquaculture sector in Africa south of the Sahara. Aquaculture Development and Coordination Programme ADCP\REP\89\36. 36 p.

van Dam, A.A. 1989. Appendix: An example of the potential application of modeling in aquaculture: production of the Nile tilapia, *Oreochromis niloticus*, in ricefields, p. 36-46. *In* M.L. Cuenco. Aquaculture systems modeling: an introduction with emphasis on warmwater aquaculture. ICLARM Stud. Rev. 19, 46 p.

van Dam, A.A. 1990a. Multiple regression analysis of accumulated data from aquaculture experiments: a rice-fish culture example. Aquacult. Fish. Manage. 21:1-15.

van Dam, A.A. 1990b. Is ANOVA powerful enough for analyzing replicated pond experiments? Aquabyte 3(3):3-5.

Yamane, T. 1973. Statistics, an introductory analysis. 3rd edition. Harper and Row, New York.

Multivariate Analysis of Growth in Juvenile Tilapia *Oreochromis aureus* and *O. niloticus*, Cichlidae, Reared in Recirculating Systems*

Graham C. Mair, *School of Biological Sciences, University College of Swansea, Swansea SA2 8PP, Wales, UK (Present address: FAC/CLSU-University of Wales Swansea Research Project on Genetic Improvement of Tilapia, Freshwater Aquaculture Center, Central Luzon State University, Muñoz, Nueva Ecija, Philippines 3120)*

Daniel Pauly, *International Center for Living Aquatic Resources Management, MC P.O. Box 2631, 0718 Makati, Metro Manila, Philippines*

Mair, G.C. and D. Pauly. 1993. Multivariate analysis of growth in juvenile tilapia *Oreochromis aureus* and *O. niloticus*, Cichlidae), reared in recirculating systems, p. 97-104. *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

A multiple regression analysis of growth in juvenile tilapia *Oreochromis niloticus* and *O. aureus* (Cichlidae) is presented based on experiments conducted from 1984 to 1987 in recirculating systems at the Tilapia Genetics Laboratory of the University College of Swansea, Wales, UK. One regression model was derived for *O. niloticus*, one for *O. aureus* and one for both species jointly. Overall fit, although not high, was significant (R^2 = 0.500, 0.493 and 0.423, respectively for n = 2,446, 3,631 and 6,077, respectively). A surprisingly high number of variables were found to correlate with growth rate; of 19 hypothesized variables (17 of which were dummies, coded 0 or 1), only 4 were not significant in the *O. niloticus* model, 2 in the *O. aureus* model and 4 in the joint model (α = 1%).

The discussion emphasizes the usefulness of dummy variables for decomposing into primary factors the impacts on fish growth of complexly integrated breeding experiments and rearing systems.

Introduction

This study presents an attempt to apply a multiple regression method for analysis of fish growth to data initially assembled for a completely different purpose, i.e., elucidating the genetics of sex determination of the commercially important tilapia (*Oreochromis niloticus* and *O. aureus*) (Mair 1988; Mair et al. 1991a, 1991b).

Thus, this study may be considered as a contribution toward assessing the utility of a method for extracting information on growth from data collected for other purposes.

Another aim is to demonstrate how dummy variables (i.e., variables coded either 0 or 1) can be defined such that complex systems - illustrated here by a recirculating system providing the fish with a large number of qualitatively different microenvironments - can be decomposed, for the purpose of multiple regression analysis, into a number of factors (i.e., dummy variables) permitting quantitative analysis. It should be noted that, in contrast to other applications of multiple regression analysis presented in this volume, the present approach is strictly for the purpose of hypothesis-testing. This is regarded here as a complementary approach to the methods usually applied to these questions such as, for example, analysis of variance.

Materials and Methods

Identity of Fish Stocks

The stocks of *O. niloticus* (Linnaeus) and *O. aureus* (Steindachner) used in this study were obtained as laboratory strains from the Institute of Aquaculture, University of Stirling, UK, in 1982-83. Both originate from Lake Manzala, Egypt, where they occur sympatrically. Electrophoretic and morphometric analyses (McAndrew and Majumdar 1983; Mair 1988) provided no evidence to suggest that these strains are introgressed with each other or with any other species of tilapia. It is assumed therefore that these strains are "pure". Note, however, that they have been through a number of genetic bottlenecks and are somewhat inbred (Mair 1988).

Breeding and Growth Experiments

The experiments analyzed here were conducted from October 1984 to September 1987 in closed recirculating systems at the Tilapia Genetics Laboratory of the University College of Swansea, Wales, UK. They were primarily directed toward elucidating the genetics of sex determination in the two species mentioned. The study utilized a number of techniques of hormonal and genetic manipulation (Mair 1988) and included a study of temperature effects on progeny sex ratio (Mair et

*ICLARM Contribution No. 733

al. 1990). The data analyzed here were gathered at the termination of experiments when progeny from test crosses were sacrificed in order to determine accurately their sex using the gonad squash technique of Guerrero and Shelton (1974). Prior to sacrifice all fish were held, for at least part of the growout period, in plastic bins in one of four closed recirculating systems.

Definition of Continuous Variables

For each sacrificed fish, the following data were recorded:

(i) AGE : age at sacrifice, in days (AGE = 0 refers to first feeding fry);

(ii) SL : standard length at sacrifice, in cm; and

(iii) W : wet weight at sacrifice, in g.

These data were used to compute derived variables, i.e.,

(iv) GR : instantaneous growth rate, as given by

$$GR = (\ln W - \ln 0.06)/AGE \qquad \text{...1)}$$

where W and AGE are as given above, ln indicates natural logarithms and 0.06 g is the mean initial weight of the first feeding fry as used for the experiments. Note that GR serves as *dependent* variable in all models presented below. Also computed was

(v) CF : condition factor, defined by

$$CF = 100 \cdot W/(SL)^3 \qquad \text{...2)}$$

CF is hypothesized to display a positive (partial) correlation with GR. The number of fish stocked (N_1) and the number recovered (N_2) from each experiment were also used to define:

(vi) MORTALITY : fraction of fish dying per day in a given bin, as defined by

$$MORTALITY = (1/AGE)\cdot(- \ln (N_2/N_1)) \qquad \text{...3)}$$

We hypothesized that MORTALITY is negatively correlated with GR.

The continuous variables SL, W and AGE, were used to derive other variables but were not included in the model. In addition, the variable DENSITY (N_2/N_1) was dropped from the model due to its collinearity with MORTALITY (see Table 1).

Definition of Dummy Variables

A number of variables are defined below which take only the values 0 or 1; these "dummy variables" can obviously be used to code inherently dichotomous characteristics such as sex, but also to decompose a complex series of experiments and rearing systems into a number of quantifiable factors. In the case of the experiments and systems mentioned above, this approach led to the following dummy variables pertaining to each of the experimental fish (with "DEFAULT" indicating cases when a dummy variable was set at zero):

(vii) SPECIES : *Oreochromis aureus* = 1 (n = 3,631); *O. niloticus* = 0 (n = 2,446). This variable is used only for the two-species model defined below; for this case, a negative sign is hypothesized for the corresponding partial regression coefficient or "slope" because *O. niloticus* is known to have a higher growth performance than *O. aureus* (Pauly et al. 1988);

(viii) SEX : Sex of individual fish, with ♀ = 1 and ♂ = 0; slope hypothesized to have a negative sign because tilapia males generally have a higher growth performance than females (Wohlfarth and Hulata 1983);

(ix) ALLMALES : Variable indicating sex of brood; 1 = fish were part of an all-male brood, 0 = fish were not part of such brood. Slope hypothesized to have a positive sign due to the absence of complex behavioral interactions between sexes in monosex broods;

Table 1. Pearson correlation coefficients for all variables used in the analysis of JOINT.DAT

	SL	W	AGE	GR	CF	SEX	DENS	MORT	SPEC	SYS1	SYS2	SYS3	WINT
SL	1.00	0.93	0.67	-0.44	-0.04	-0.03	-0.13	-0.20	-0.03	0.04	-0.09	-0.14	-0.01
W	0.93	1.00	0.61	-0.38	-0.01	-0.06	-0.13	-0.16	-0.02	0.07	-0.11	-0.14	0.01
AGE	0.67	0.61	1.00	-0.86	-0.01	0.03	-0.21	-0.27	0.06	-0.21	-0.04	-0.02	-0.02
GR	-0.44	-0.38	-0.86	1.00	0.00	-0.02	0.13	0.38	0.02	0.27	0.04	0.02	0.09
CF	-0.04	-0.01	-0.01	0.00	1.00	-0.02	-0.02	0.01	0.02	-0.01	-0.02	-0.01	0.01
SEX	-0.03	-0.06	0.03	-0.02	-0.02	1.00	-0.10	0.04	-0.07	-0.05	-0.02	0.01	-0.01
DENS	-0.13	-0.13	-0.21	0.13	-0.02	-0.10	1.00	-0.71	-0.11	0.14	0.13	-0.16	-0.07
MORT	-0.20	-0.16	-0.27	0.38	0.01	0.04	-0.71	1.00	0.08	0.00	-0.06	0.14	0.08
SPEC	-0.03	-0.02	0.06	0.02	0.02	-0.07	0.11	0.08	1.00	-0.05	0.27	0.13	-0.04
SYS1	0.04	0.07	-0.21	0.27	-0.01	-0.05	0.14	0.00	-0.05	1.00	-0.37	-0.17	0.00
SYS2	-0.09	-0.11	-0.04	0.04	-0.02	-0.02	0.13	-0.06	0.27	-0.37	1.00	-0.24	-0.07
SYS3	-0.14	-0.14	-0.03	0.02	-0.01	0.01	-0.16	0.14	0.13	-0.17	-0.24	1.00	0.11
WINT	-0.01	0.01	-0.02	0.09	0.01	-0.01	-0.07	0.08	-0.04	0.00	-0.07	0.11	1.00
SPRG	-0.01	-0.02	0.06	-0.06	-0.02	0.06	-0.05	-0.04	0.03	-0.18	-0.01	-0.11	-0.31
SUMM	-0.06	-0.04	-0.03	-0.01	0.02	0.01	0.06	-0.00	-0.04	0.09	-0.03	0.01	-0.30
REAR	0.12	0.14	-0.20	0.29	-0.03	-0.01	-0.02	0.10	-0.24	0.13	-0.33	-0.19	-0.09
PAR1	-0.01	-0.04	0.07	-0.07	0.03	-0.35	0.07	-0.06	0.47	0.11	0.01	0.07	0.03
PAR2	-0.03	-0.03	0.01	-0.06	-0.01	0.14	-0.03	-0.01	-0.14	-0.06	0.09	-0.04	-0.05
PAR3	0.15	0.16	-0.06	0.11	-0.01	-0.19	0.12	0.00	-0.28	0.19	-0.01	0.01	-0.08
SPAWN	-0.13	-0.15	0.02	-0.09	0.03	-0.07	0.04	0.02	0.49	-0.18	0.39	0.01	-0.08
ALLM	0.05	0.07	-0.02	0.02	0.04	-0.38	0.02	0.04	0.12	0.13	0.01	-0.07	-0.03
ALLF	-0.05	-0.05	-0.06	0.07	-0.01	0.24	-0.08	0.13	-0.07	0.06	0.01	-0.05	-0.05
GYNO	-0.03	-0.06	-0.06	0.05	-0.01	0.16	-0.10	0.10	-0.12	0.09	-0.09	-0.02	-0.02
LOTEMP	0.01	-0.03	0.06	-0.12	0.01	0.03	-0.16	0.05	-0.17	-0.13	-0.04	0.04	0.37
HITEMP	0.00	-0.04	0.07	-0.10	-0.01	0.01	-0.04	-0.04	0.14	-0.09	0.24	-0.06	-0.08

	SPRG	SUMM	REAR	PAR1	PAR2	PAR3	SPAWN	ALLM	ALLF	GYNO	LOTEMP	HITEMP
SL	-0.01	-0.06	0.12	-0.01	-0.03	0.15	-0.13	0.05	-0.05	-0.03	0.01	0.00
W	-0.02	-0.04	0.14	-0.04	-0.03	0.16	-0.15	0.07	-0.05	-0.06	-0.03	-0.04
AGE	0.06	-0.03	-0.20	0.07	0.01	-0.06	0.02	-0.02	-0.06	-0.06	0.06	0.07
GR	-0.06	-0.01	0.29	-0.07	-0.06	0.11	-0.09	0.02	0.07	0.05	-0.12	-0.10
CF	-0.02	0.02	-0.03	0.03	-0.01	-0.01	0.03	0.04	-0.01	-0.01	0.01	-0.01
SEX	0.06	0.01	-0.01	-0.35	0.14	-0.20	-0.07	-0.38	0.24	0.16	0.03	0.01
DENS	-0.05	0.06	-0.02	0.07	-0.03	0.12	0.04	0.02	-0.08	-0.10	-0.16	-0.04
MORT	-0.04	0.00	0.10	-0.06	-0.01	0.00	0.02	0.04	0.13	0.10	0.05	-0.04
SPEC	0.03	-0.14	-0.24	0.47	-0.14	-0.30	0.49	0.19	-0.07	-0.12	-0.17	0.14
SYS1	-0.18	0.09	0.13	0.11	-0.06	0.19	-0.13	0.13	0.06	0.09	-0.13	-0.09
SYS2	-0.01	-0.03	-0.33	0.01	0.09	-0.01	0.39	0.01	0.01	-0.09	-0.04	0.24
SYS3	-0.11	0.01	-0.19	0.07	-0.04	0.01	0.01	-0.07	-0.05	-0.02	0.04	-0.06
WINT	-0.31	-0.30	-0.09	0.03	-0.05	-0.08	-0.08	-0.03	-0.05	-0.02	0.37	-0.08
SPRG	1.00	-0.45	-0.02	-0.11	-0.08	-0.03	0.09	-0.17	-0.12	0.03	-0.17	0.09
SUMM	-0.45	1.00	-0.06	-0.18	0.06	0.15	-0.04	0.07	0.14	0.01	-0.02	0.06
REAR	-0.02	-0.06	1.00	-0.07	-0.06	-0.01	-0.47	0.02	-0.03	-0.02	0.01	-0.27
PAR1	-0.11	-0.18	-0.07	1.00	-0.07	-0.16	0.14	0.50	-0.12	-0.08	-0.16	0.02
PAR2	-0.08	0.06	-0.06	-0.07	1.00	-0.03	-0.07	-0.05	0.40	-0.02	-0.03	-0.02
PAR3	-0.03	0.15	0.01	-0.16	-0.03	1.00	-0.15	0.29	0.05	-0.04	-0.06	-0.04
SPAWN	0.09	-0.04	-0.47	0.14	-0.07	-0.15	1.00	-0.03	-0.12	-0.08	0.00	0.29
ALLM	-0.17	0.07	0.02	0.50	-0.05	0.29	-0.03	1.00	-0.09	-0.07	-0.11	-0.00
ALLF	-0.12	0.14	-0.03	-0.12	0.40	-0.05	-0.12	-0.09	1.00	0.38	-0.05	-0.03
GYNO	0.03	0.01	-0.02	-0.08	-0.02	-0.04	-0.08	-0.07	0.38	1.00	-0.04	-0.03
LOTEMP	-0.17	-0.02	0.01	-0.16	-0.03	-0.06	0.00	-0.11	-0.05	-0.04	1.00	-0.04
HITEMP	0.09	0.06	-0.27	0.02	-0.02	-0.04	0.29	0.00	-0.03	-0.03	-0.04	1.00

ALLFEMALES: Other variable indicating sex of brood; 1 = fish were part of an all-female brood; 0 = fish were not part of such brood. Slope hypothesized to have a positive sign;

DEFAULT : mixed sex, i.e., fish were part of a mixed brood.

x) PARENT1 : Variable referring to parents of fish in question with 1 = Δ♀ x N , i.e., delta female by normal male; 0 = not as above, and where a Δ♀ is a genetic male hormonally sex reversed to female;

PARENT2 : Same as above but 1 = N x Δ♂, 0 = not as above, and where a Δ♂ is a genetic female hormonally sex reversed to male;

PARENT3 : 1 = Sup x N, i.e., one parent was a super-female or a super-male, and 0 = not as above; super females or males are novel homogametic genotypes, WW female in *O. aureus* or YY male in *O. niloticus* (Mair et al. 1991a, 1991b).

DEFAULT : parents were normal female and normal male (i.e., PARENTS 1, 2 and 3 = 0)

(xi) GYNOGEN : 1 = Fish was gynogenetic (i.e., having only maternal inheritance); or 0 = fish received genes from both parents. Gynogens were produced by suppression of meiosis or mitosis in eggs fertilized with ultra-violet irradiated sperm according to Mair et al. (1987). The slope is hypothesized to have a negative sign due to manifestation of inbreeding depression.

(xii) SPAWNING : 1 = "natural" spawning (in aquaria); 0 = stripping, artificial fertilization and incubation. We hypothesized that artificial incubation may reduce subsequent growth due to the potentially poorer condition of fry compared to those incubated by the female parent.

The following dummy variables were devised to describe the major characteristics of the four recirculating systems and the conditions under which the experimental fish were kept:

(xiii) REARING : Variable set at 1 when the fish have been all the time in a recirculating system, and at 0 when they have spent some time in aquaria (usually the initial period). The corresponding slope is hypothesized to have a positive sign due to reduced growth of fish in aquaria.

(xiv) SYSTEM 1 : variable set at 1 when fish were kept in first (of four) recirculating systems, otherwise at 0;

SYSTEM 2 : second system = 1, otherwise 0;

SYSTEM 3 : third system = 1, otherwise 0;

DEFAULT : fourth recirculating system (i.e., SYSTEM 1, 2 and 3 = 0).
Recirculating systems 1,2 and 3 consisted of 22-28 plastic bins (27-l capacity) which were initially stocked at 1.4-1.9 $fish \cdot l^{-1}$. System 4 had larger bins (47 l) stocked at 1.6 - 2.3 $fish \cdot l^{-1}$.

(xv) WINTER : One of four seasons during which midpoint of growth period occurred. Set at 1 when season is winter (Jan., Feb., Mar.), otherwise 0;

SPRING : set at 1 when season is spring (Apr., May, Jun.), otherwise 0;

SUMMER : set at 1 when season is summer (Jul., Aug., Sep.), otherwise 0;

DEFAULT : autumn (Oct., Nov., Dec.).

(xvi) LOTEMP : 1 = fish grew (at some point, usually at the beginning) below 26°C; 0 = fish always grew above 26°C. The slopes of these variables are hypothesized to have a negative sign as these temperatures are outside the optimum for growth of these tilapia;

HIGHTEMP : 1 = fish grew (at some point, usually at the beginning) above 30°C; 0 = fish always grew below 30°C. Slope hypothesized to have a negative sign;

DEFAULT : fish grew at all times between 26°C and 30°C.

Three files were created, using an IBM compatible personal computer, with appropriate values for the above variables; one for *O. niloticus* ("NILE.DAT"), one for *O. aureus* ("AUR.DAT") and one for both species ("JOINT.DAT").[a]

Hypothesized Multiple Regression Models

The approach in the present analysis is to formulate a hypothetical regression model for each dataset, based on biological reasoning, and to test

[a]See Appendix II, p. 201.

this in a single regression run on the entire amount of data available. It should be noted that this is strictly a hypothesis-testing approach, in contrast to other uses of multiple regression for exploratory data analysis or predictive model building, as in other presentations in this volume.

For NILE.DAT and AUR.DAT, the hypothesized multiple regression models had the form

$$\begin{aligned}GR = {} & INTERCEPT + b_1CF + b_2MORTALITY \\ & + b_3SEX + b_4ALLMALES + b_5ALLFEMALES \\ & + b_{6\text{-}8}PARENT1\text{-}3 + b_9GYNOGEN \\ & + b_{10}REARING + b_{11}SPAWNING \\ & + b_{12\text{-}14}SYSTEM1\text{-}3 + b_{15\text{-}17}SEASON1\text{-}3 \\ & + b_{18}LOTEMP + b_{19}HITEMP \\ & + ERROR\ TERM \qquad ...4)\end{aligned}$$

For JOINT.DAT, the hypothesized model differed from (4) only in that "SPECIES" (see above) was added as 20th variable before the error term.

Parameter estimates (intercepts, slopes and their related statistics) were obtained using the SAS/STAT package (SAS 1988).

The partial regression coefficients (b_i, or slopes) were used to compute standardized slopes, or β coefficients according to

$$\beta = b_i \cdot \frac{sd\ (variable)_i}{sd(GR)} \qquad ...5)$$

which allow comparison of effects on GR measured in different units (including comparisons between dummy and continuous variables; Blalock 1972).

Results and Discussion

Table 2 shows the means and standard deviations for all variables (including SL, W, AGE and DENSITY). These show that growth rates were very low in comparison with those normally observed in pond culture situations, the fish growing to a mean weight of only 19 g by a mean age of 185 days. Mortality rates were high at 0.002-0.004 fish per day resulting in a mean loss of 35% of stocked fish over the growout period.

Points of interest regarding the frequency of occurrence of 0 and 1 values within the dummy variables are that 71% of fish were grown entirely in recirculating systems and all *O. niloticus* crosses were made using artificial fertilization. Variables PARENT 2 and 3, ALLFEMALE, GYNOGEN, LOTEMP, HITEMP and SYSTEM 3 were all set at 1 in less than 10% of observations.

Table 2. Mean and standard deviation for all variables.

Variable	*O. aureus* Mean	*O. aureus* SD	*O. niloticus* Mean	*O. niloticus* SD	Both species Mean	Both species SD
SL	6.67	2.88	6.94	2.86	6.78	2.87
W	18.20	23.23	20.09	24.13	18.97	23.59
AGE	188.39	97.72	178.01	70.93	184.21	87.38
GR	0.06	0.02	0.06	0.01	0.06	0.02
CF	4.07	8.35	3.86	0.83	3.99	6.48
SEX	0.38	0.49	0.45	0.50	0.41	0.49
DENSITY	0.63	0.23	0.68	0.24	0.65	0.24
MORTALITY	0.004	0.004	0.003	0.004	0.003	0.004
SPECIES	1.00	0.00	0.00	0.00	0.60	0.49
SYSTEM 1	0.20	0.40	0.24	0.42	0.21	0.41
SYSTEM 2	0.44	0.50	0.18	0.38	0.34	0.47
SYSTEM 3	0.13	0.34	0.05	0.22	0.10	0.30
WINTER	0.16	0.37	0.19	0.39	0.17	0.38
SPRING	0.33	0.47	0.30	0.46	0.32	0.47
SUMMER	0.24	0.43	0.38	0.49	0.30	0.46
REARING	0.62	0.49	0.84	0.36	0.71	0.45
PARENT 1	0.46	0.50	0.03	0.16	0.29	0.45
PARENT 2	0.00	0.00	0.03	0.18	0.01	0.11
PARENT 3	0.00	0.00	0.15	0.36	0.06	0.24
SPAWNING	0.44	0.50	0.00	0.00	0.26	0.44
ALLMALE	0.23	0.42	0.09	0.29	0.17	0.38
ALLFEMALE	0.03	0.16	0.05	0.23	0.04	0.19
GYNOGEN	0.01	0.10	0.05	0.21	0.03	0.16
LOTEMP	0.03	0.16	0.11	0.31	0.06	0.24
HITEMP	0.05	0.22	0.00	0.00	0.03	0.17

Table 3 summarizes the parameter estimates of the models derived from the NILE.DAT, AUR.DAT and JOINT.DAT files. The model derived from the NILE.DAT file, pertaining to *O. niloticus* explains half of the variance in the dataset, (R^2 = 0.500) and all but four of the variables included in the model were found to be significant (α = 1%). The model based on the AUR.DAT file, pertaining to *O. aureus* also explained approximately half of the variance (R^2 = 0.493), and only two of the variables were not significant (α = 1%).

Table 3. Variable estimates and significance levels for NILE.DAT, AUR.DAT and JOINT.DAT.

		O. niloticus			*O. aureus*			Both species	
Variable	Estimate	α(%)	β	Estimate	α(%)	β	Estimate	α(%)	β
INTERCEPT	0.0417	<0.01		0.0154	<0.01		0.0221	<0.01	
CF	-0.0005	13.81	-0.0225	0.0001	1.15	0.0301	0.0001	0.57	0.0271
MORTALITY	1.2874	<0.01	0.2775	1.6281	<0.01	0.2936	1.6568	<0.01	0.3166
SEX	-0.0007	22.73	-0.0194	-0.0020	0.36	-0.0413	-0.0019	0.02	-0.0427
ALLMALE	0.0054	<0.01	0.0902	-0.0026	0.51	-0.0454	-0.0002	80.79	-0.0032
ALLFEMALE	0.0054	0.27	0.0701	0.0065	0.05	0.0440	0.0057	<0.01	0.0509
PARENT1	-0.0274	<0.01	-0.2543	-0.0021	1.10	-0.0435	-0.0042	<0.01	-0.0899
PARENT2	-0.0099	<0.01	-0.1015	-	-	-	-0.0066	0.17	-0.0354
PARENT3	0.0006	58.99	0.0131	-	-	-	0.0015	17.89	0.0166
GYNOGEN	0.0072	<0.01	0.0897	-0.0126	<0.01	-0.0513	0.0000	99.29	0.0001
REARING	0.0164	<0.01	0.3429	0.0224	<0.01	0.4603	0.0199	<0.01	0.4242
SPAWNING	-	-	-	0.0031	<0.01	0.0655	0.0018	0.43	0.0380
SYSTEM1	0.0190	<0.01	0.4677	0.0226	<0.01	0.3818	0.0215	<0.01	0.4136
SYSTEM2	0.0141	<0.01	0.3132	0.0213	<0.01	0.4486	0.0192	<0.01	0.4260
SYSTEM3	0.0114	<0.01	0.1433	0.0171	<0.01	0.2446	0.0160	<0.01	0.2236
WINTER	-0.0151	<0.01	-0.3421	0.0214	<0.01	0.3352	-0.0139	<0.01	0.2472
SPRING	-0.0002	79.57	-0.0066	0.0071	<0.01	0.1406	0.0074	<0.01	0.1612
SUMMER	-0.0114	<0.01	-0.3204	0.0132	<0.01	0.2393	0.0053	<0.01	0.1144
HITEMP	-	-	-	-0.0061	<0.01	-0.0557	-0.0056	<0.01	-0.0445
LOTEMP	0.0046	0.02	0.0825	-0.0204	<0.01	-0.1387	-0.0135	<0.01	-0.1507
SPECIES	-	-	-	-	-	-	-0.0002	78.46	-0.0039

Dashes (-) refer to variables for which there was an insufficient number of 0 and 1 cases.

The results for both species based on the JOINT.DAT file explained a slightly smaller amount of variance (R^2 = 0.423), but only four variables were nonsignificant (α = 1%): ALLMALE, PARENTS, GYNOGEN and SPECIES.

From the estimates of slopes in JOINT.DAT, it can be seen that the MORTALITY variable has a large effect on GR. However the nature of this analysis prohibits the comparison of the magnitude of this effect with that of any of the dummy variables. Direct comparison of the effects of the dummy variables shows that REARING, LOTEMP and the SEASON and SYSTEM variables have the greatest effect on GR with GYNOGEN, ALLMALE and SPECIES having the smallest effect. Comparison of the β coefficients of the most important variables shows that the continuous variable MORTALITY also has a considerable effect on GR.

Ranking of the contributions of each variable to R^2 in Table 4 shows that most of the variables having the greatest effect on GR (MORTALITY, REARING, SYSTEM1-3, WINTER and LOTEMP) also account for 94% of R^2.

From the above it can be concluded that MORTALITY, REARING and the SYSTEM variables have the strongest, and statistically most reliably estimated, effect on growth rate. Mortality exerts its effect through a reduction in stocking density in bins. This strong positive correlation of mortality with growth disproves our hypothesis which was based on the assumption that mortality was indicative of suboptimal rearing environment which would also adversely affect growth rate. The high relative importance of the rearing environment confirms our hypothesis that growing fish initially in aquaria suppresses growth which is then not compensated when fish are moved to bins in recirculating systems*. The importance of the

*Similarly the effect of initial rearing of fry at low temperature prior to transfer to recirculating systems had a predictable effect of suppression of growth.

Table 4. Ranked contribution to R^2, sign of correlations and agreement with hypothesis of variables in JOINT.DAT file.

Variables	Contribution to R^2 (%)	Significance of estimate	Sign of correlation	Agreement with hypotheses
MORTALITY	34.85	***	+	No
REARING	18.19	***	+	Yes
SYSTEM2	13.45	***	+	N/A
SYSTEM1	12.15	***	+	N/A
SYSTEM3	6.71	***	+	N/A
WINTER	4.97	***	+	N/A
LOTEMP	4.13	***	-	Yes
PARENT3	1.99	N.S.	+	N/A
PARENT1	1.10	***	-	N/A
PARENT2	0.97	**	-	N/A
SPRING	0.63	***	+	N/A
HITEMP	0.35	***	-	Yes
ALLFEMALE	0.21	***	+	Yes
SEX	0.13	***	-	Yes
SPECIES	0.08	N.S.	-	Yes
SUMMER	0.03	***	+	N/A
SPAWNING	0.02	**	+	Yes
CF	0.00	**	+	Yes
GYNOGEN	0.00	N.S.	+	No
ALLMALE	0.00	N.S.	-	No

** a < 1%; *** a < 0.1%; N.S. = not significant; N/A = not applicable.

SYSTEM variables indicates there are large differences between growth of fish in the different recirculating systems. Since the design of these systems are similar, it is likely that these effects are brought about by water quality differences due to variable effectivity of biofilters. The importance of the season variable, especially WINTER, is harder to interpret as the water temperature and light variables were kept relatively constant throughout the year. However, these variables explain a relatively small amount of the total variation and should not be used to formulate strong hypotheses regarding their effect on growth.

The lack of significant correlation of GYNOGEN with growth rate is surprising as established genetic theory suggests that growth is likely to be affected adversely by the level of inbreeding induced by gynogenesis. There was a significant effect of GYNOGEN on GR in *O. niloticus*, but this had a positive sign, contrary to our hypothesis. The absence of the predicted negative correlation of GYNOGEN with growth may be due to the fact that these laboratory strains have become highly inbred during domestication, limiting the deleterious effects of further inbreeding induced by gynogenesis.

A more surprising observation was the low significance of the small effect of SEX on GR, this variable also explaining a very small proportion of R^2. One possible explanation for the limited effect of sex on growth rate is that the majority of fish were sacrificed at a small size, perhaps before the effect of sexual maturity exerted their differential influence on the growth of the respective sexes. However, the mean age of the fish was high and it would be expected that many of these fish would be sexually mature by the time of sacrifice. A more plausible explanation is that the environmental and physiological factors that play an important role in the differential growth of the sexes (territorial and courtship behavior, gamete production, and egg incubation by females) are largely absent in recirculating systems. Fish were confined at high densities preventing dominant males from establishing territories and preventing spawning in bins containing sexually mature fish of both sexes. The inference here is that environmental factors play a very important role in the well established phenomenon of differential growth of the sexes in tilapia. A similar explanation could be applied to the result from the ALLMALE variable which, contrary to our hypothesis, produced a nonsignificant estimate and explains virtually none of the variance in GR.

The hypotheses made regarding the remaining variables were confirmed by this analysis. No hypotheses were made regarding the relative effects of the grouped variables of PARENT, SEASON and SYSTEM. The effect of PARENT variables are difficult to interpret as their effects should be felt chiefly through the sex of the progeny. Correlations between PARENT1, 2, and 3 with sex were fairly high but less than 0.5 (Table 1). Correlations between PARENT1 with ALLMALE, PARENT2 with ALLFEMALE, and PARENT3 with ALLMALE and ALLFEMALE are high. These variables should be analyzed further for each species separately as their effects differ between species (e.g., PARENT1 usually gives 1:0 (M:F) sex ratios in *O. aureus* and 3:1 ratios in *O. niloticus*). Nevertheless the effect on GR of PARENT1 in *O. niloticus* is surprising given the limited influence of sex on GR.

One problem with the use of series of dummy variables in this analysis is that it is difficult to make inferences regarding DEFAULT. For example it is not obvious that the sign of the slope estimates for SYSTEMS 1, 2 and 3 indicate that an estimate for SYSTEM4 would have had a negative slope, and that the ranking of these four variables in their effect on GR would have remained the same had another default been used.

This problem, along with other questions raised by this analysis will, however, have to be addressed in future contributions. Mosteller and Tukey (1977) suggest that instead of using datasets with a large number of cases, only a random sample of approximately 200 cases should be used for predictive model building. Although the approach used in this analysis is strictly one of hypothesis-testing, a high number of cases caused an inflation of the significance levels for R^2 and of the individual predictor variables. This may result in variables, whose significance in the present case is due only to the large number of cases, becoming insignificant in a regression based on a smaller number of cases. Furthermore, the extensive use of dummy variables in replacement for unmeasured continuous variables may introduce artificial random effects such as wrong significances, signs, slopes and intercorrelations.

Acknowledgements

The authors acknowledge Carmela Janagap of ICLARM for her invaluable contribution to the computational analyses of data. In addition, we thank Eric Roderick, University College of Swansea, for his technical assistance. This work was supported by grants from the Overseas Development Administration and the EEC Tropical Aquaculture Programme to J.A. Beardmore and D.O.F. Skibinski.

References

Blalock, H.M. 1972. Social statistics. McGraw-Hill, San Francisco. 383 p.

Guerrero R.D. and W.L. Shelton. 1974. An aceto-carmine squash method for sexing juvenile fishes. Prog. Fish.-Cult. 36: 56.

Mair, G.C. 1988. Studies on sex determining mechanisms in *Oreochromis* species. University College of Swansea, Wales. 326 p. Ph.D. thesis.

Mair, G.C., J.A. Beardmore and D.O.F. Skibinski. 1990. Experimental evidence for environmental sex determination in *Oreochromis* species, p. 555-558. *In* R. Hirano and I. Hanyu (eds.) The Second Asian Fisheries Forum. Asian Fisheries Society, Manila. 991 p.

Mair, G.C., A.G. Scott, J.A. Beardmore and D.O.F. Skibinski. 1987. A technique for induction of diploid gynogenesis in *Oreochromis niloticus* by suppression of the first mitotic division, p. 289-299. *In* K. Tiews (ed.) Proceedings of the World Symposium on Selection, Hybridization and Genetic Engineering in Aquaculture, Bordeaux, 27-30 May 1986. Vol. II. Heenemann, Berlin.

Mair, G.C., A.G. Scott, D.J. Penman, J.A. Beardmore and D.O.F. Skibinski. 1991a. Sex determination in the genus *Oreochromis*. I: Sex reversal, gynogenesis and triploidy in *O. niloticus* L. Theor. Appl. Genet. 82:144-152.

Mair, G.C., A.G. Scott, D.J. Penman, D.O.F. Skibinski and J.A. Beardmore. 1991. Sex determination in the genus *Oreochromis*. II: Sex reversal, hybridisation, gynogenesis and triploidy in *O. aureus* Steindachner. Theor. Appl. Genet. 82:153-160.

McAndrew, B.J. and K.D. Majumdar. 1983. Tilapia stock identification using electrophoretic markers. Aquaculture 30:249-261.

Mosteller, F. and J.W. Tukey. 1977. Data analysis and regression. Addison-Wesley, Reading, Massachusetts. 588 p.

Pauly, D, J. Moreau and M. Prein. 1988. A comparison of overall growth performance of tilapia in open waters and aquaculture, p. 469-479. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

SAS. 1988. SAS/STAT User's guide, Release 6.03 edition. Statistical Analysis System Institute, Inc., Cary, North Carolina.

Wohlfarth, G.W. and G. Hulata. 1983. Applied genetics of tilapias. Second edition. ICLARM Stud. Rev. 6, 26 p.

...stantaneous Mortalities ...d Multivariate Models: ...plications to Tilapia Culture ... Saline Water*

...IN D. HOPKINS, *College of Agriculture, University of ...vaii at Hilo, 523 W. Lanikaula St., Hilo, Hawaii ...20-4091, USA*

...IEL PAULY, *International Center for Living Resources ...nagement, MC P.O. Box 2631, 0718 Makati, Metro ...nila, Philippines*

...kins, K.D. and D. Pauly. 1993. Instantaneous mortalities and multivariate models: applications to tilapia culture in saline water, p. 105-111. *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

The "per cent mortalities" commonly used by aquaculturists ...ot allow separation of the different components of fish mor...y between stocking and harvesting in aquaculture experi...ts. It is shown that "instantaneous" or exponential ...alities, as used in fish population dynamics, have the prop...s required for such separation, especially when used in con...tion with a multiple regression model. Examples drawn from ...ia experiments conducted in seawater tanks in Kuwait and ...kishwater ponds in the Philippines are presented.

Introduction

Survival and mortality are of great importance ...aquaculturists because, in combination with ...wth, they determine net yield. However, this ...ortance is not reflected in the amount of effort ...anded by aquaculturists to analyze survival ...l mortality. Aquaculturists typically analyze ...vival by computing the per cent survival at the ... of each experiment and test for significant dif...ences between treatments using ANOVA. Fur...r analyses of mortality are usually not con...ted. This situation is in distinct contrast to ...eries biology where substantial portions of ...tbooks deal with methods for analyzing survival ...l mortality (Pauly 1984; Sparre et al. 1989).

Aquaculturists have a strange dual perspective ... fish mortality. First, at a visceral level, ...aculturists feel that mortality occurs only when ...nething goes wrong (e.g., anoxia or a disease ...break). One hundred per cent survival is a goal ...strive for. At the same time, production models use standard survival rates for particular life stages (e.g., Clifford 1985 uses a 75% survival rate from postlarvae to juvenile shrimp in semi-intensive culture). How can this dual perspective exist? A possible reason is that mortality is usually hidden from view unless a major catastrophe occurs. Therefore, the aquaculturist feels mortalities are caused by the catastrophes and the analysis of mortality is complete if the catastrophe has been identified. However, the number of fish harvested is almost always less than the number stocked even when no catastrophes have occurred. There is usually no indication of when or how the fish disappeared. The missing fish are a mystery which remains hidden by the standard expressions of survival rate.

This paper presents a methodology which can be used in our attempts to solve the mystery of the missing fish. It reviews the computation and attributes of two numerical representations of mortality, percentage mortalities and instantaneous rates of mortality. A multiple regression method to quantify the effects of external factors on instantaneous mortality is then presented. Examples using data from tilapia culture in seawater tanks and brackishwater ponds are given.

Mortality Rates

Aquaculturists typically use percentage mortalities (%M) which are computed as follows:

$$\%M = (N_o - N_t)/N_o \times 100 \qquad ...1)$$

where N_o is the number at the start of the culture period and N_t is the number present at time t. Percentage mortalities have the following characteristics:

1. They are restricted to time periods for which they were computed. For example, 20% over 100 days is not the same as 10% over 50 days. Thus, %M from experiments with different durations cannot be directly compared;
2. Percentage mortalities are not additive. A 10% rate attributed to one cause plus 10% from a second cause does not give a total relative mortality rate of 20%; and
3. Percentage mortalities are binomially distributed. Thus, depending on the range of values, square root or arcsin

...LARM Contribution No. 734

transformations may be needed before statistical analyses which require normal distributions can be used (Gomez and Gomez 1976).

Fisheries biologists, on the other hand, typically use instantaneous rates of mortality. These are also known as exponential rates of mortality. They are based on the following relationship.

$$N_t = N_o e^{-Zt} \qquad \text{...2)}$$

where N_t = the number at time t, N_o = the initial number, and Z is the instantaneous rate of total mortality (Pauly 1984). Instantaneous rates of mortality have the following characteristics:

1. Instantaneous rates of mortality assume that the number of fish in a particular cohort decline at a rate proportional to the number of fish alive at any point in time (Everhart and Youngs 1981). This exponential decline (Fig. 1) has been shown in many situations and species of fish;
2. As long as the unit of time used in the computation of Z is the same, instantaneous rates from experiments with different durations can be directly compared. Also, if the time units are different, the units for one Z can be converted into the units of another by simply multiplying with the appropriate time conversion factor (e.g., multiply $Zday^{-1}$ by 30 to equal $Zmonth^{-1}$);
3. Instantaneous rates are additive; and
4. Instantaneous rates of mortality usually are normally distributed.

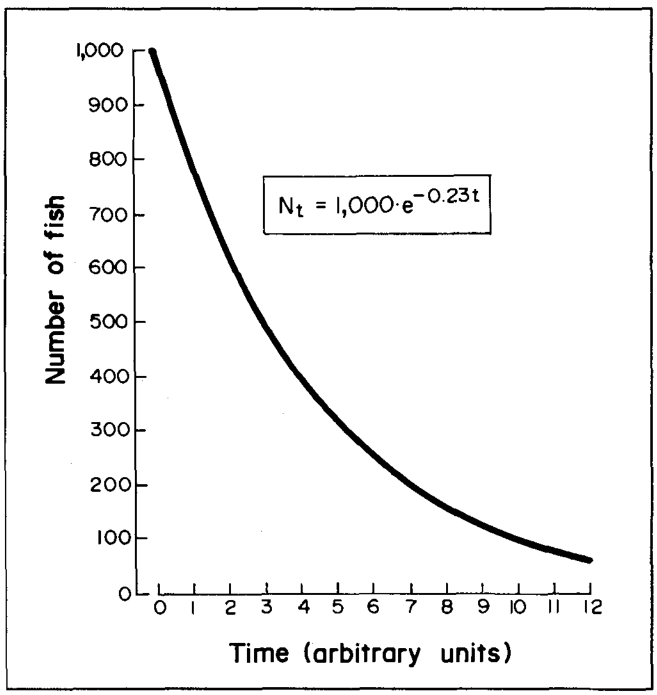


Fig. 1. Negative exponential decline of an initial population of fish (N_o = 1000) exposed to a total mortality (Z = 0.23) per time unit.

Based on the characteristics listed above, instantaneous rates are considerably superior to percentage mortalities when conducting comparisons.

Similarly to the growth analyses discussed in much of the rest of this volume, analyzing mortality data from several experiments simultaneously has the potential for furthering our understanding of the factors affecting mortality. To quantify the effects of external factors on mortality, we propose a multiple linear regression of the form:

$$Z = a + b_1X_1 + b_2X_2 + ... + b_nX_n \qquad \text{...3)}$$

where X_1 ... X_n are environmental and treatment variables recorded during the culture period.

Example 1 - **Oreochromis niloticus** *in Brackishwater Ponds*

This example is representative of a typical pond experiment in which the only available mortality data are the difference between numbers stocked and harvested. The experiments were conducted at the Brackishwater Aquaculture Center of the University of the Philippines in the Visayas as part of the USAID-supported Pond Dynamics/Aquaculture Collaborative Research Program (PD/A CRSP). Descriptions of the pond facilities can be found in Egna et al. (1987).

Data from experiments were pooled for the analyses for this paper. The experiments were designed to determine the effects of various nutrient input regimes (e.g., inorganic fertilizer, manure, feed and various combinations of these inputs) and season on fish yields and water quality (Table 1). The first two experiments, labeled cycle 1, used "impure" *Oreochromis niloticus* stocks which were felt to be contaminated with *O. mossambicus*. The other three experiments used a "purer" strain of *O. niloticus*. Details of the experimental protocols can be found in the program's work plans (PD/A CRSP, undated; PD/A CRSP 1984; PD/A CRSP 1985).

The data were extracted from the CRSP database, a large computerized database containing information from CRSP experiments throughout the world (Hopkins et al. 1988a). RBase for DOS was used for data extraction and computation of average salinities. Regression analyses, including residual analyses, were conducted using SPSS/PC+.

A cursory view of the data indicated a positive relationship between salinity and mortality (Fig. 2) which appears to be approximately linear over the range 23 to 36 ppt salinity. However, two problems are apparent with such a simple linear relationship. First, it would have a negative y-intercept. As mortality cannot be negative, this is impossible. This effect can be overcome by taking the logarithm of Z. Second, stock "purity" seems to affect the response. We took accounts of both of these problems via the more general model:

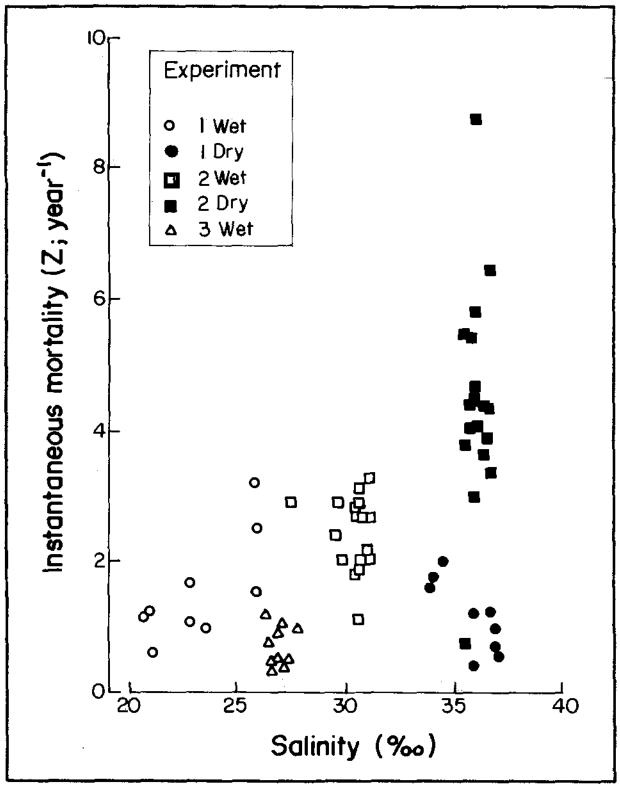


Fig. 2. Scatterplot of instantaneous mortalities as a function of salinity in brackishwater ponds.

$$\ln Z = a + b_1S + b_2P + b_3(S \times P) \quad \text{...4)}$$

where S is the average salinity, P is a dummy variable indicating relative stock "purity" (1 = "impure" and 0 = "pure"), and S x P is an interaction term. Taking the natural logarithm of Z ensures that predicted values of Z cannot be less than zero.

The results of the regression are shown in Table 2. All three terms are significant. The resulting equation is

$$\ln Z = -4.976 + 0.181\ S + 5.692\ P - 0.200\ (S \times P) \quad \text{...5)}$$

A graphical representation of the results is presented in Fig. 3. The salinities encountered during these experiments were extremely deleterious to the "purer" *O. niloticus*. The fish contaminated with *O. mossambicus* actually showed a decrease in mortality as the salinity increased. More detailed analysis to examine the effects of other water quality parameters and growth rates on mortality could be performed by simply expanding the model (the data to perform such analyses are presented in Appendix II).

Example 2 - Red Tilapia in Seawater Tanks

Data for this example were drawn from a set of experiments conducted at the Mariculture and Fisheries Department of the Kuwait Institute for Scientific Research. In these experiments, several tilapia species and hybrids were screened for their culture potential in seawater. As the experiments

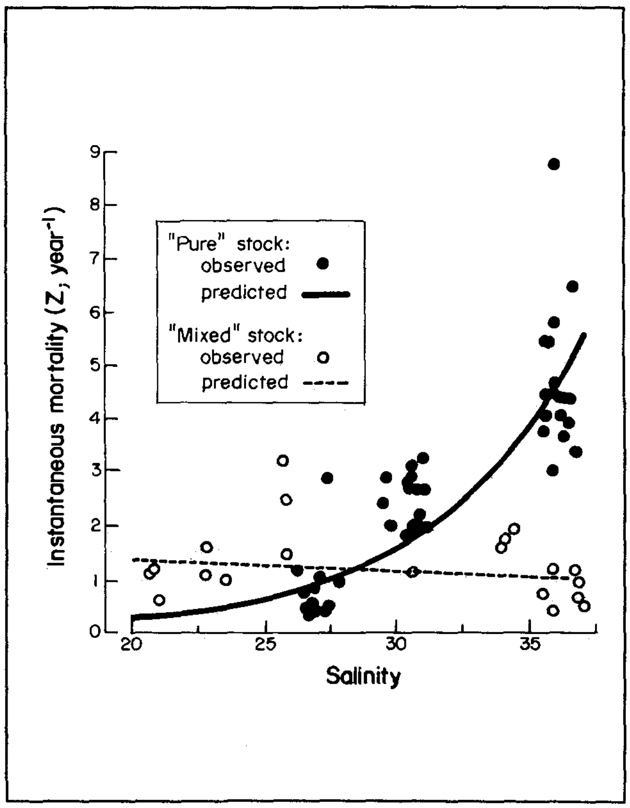


Fig. 3. Observed and predicted rates of instantaneous mortality in brackishwater tilapia ponds.

Table 1. Summary of brackishwater tilapia culture experiments (CRSP, Philippines).

Experiment	Stocking date	Harvest date	Number of ponds	Treatments	Average water temperature (°C)	Average salinity (ppt)	Average mortality Z·year^{-1}
Cycle 1 Wet season	22 Jul 83	12 Dec 83	9	Inorganic fertilizer	25.1	23.3	1.53
Cycle 1 Dry season	1 Feb 84	3 Jul 84	9	Inorganic fertilizer	23.8	35.8	1.15
Cycle 2 Dry season	24 Nov 84	25 Apr 85	21	Manure only Inorganic fertilizer only Manure & inorganic fertilizer Manure & feed Manure & inorganic fertilizer & feed No inputs	25.9	36.0	4.55
Cycle 2 Wet season	3 Oct 85	28 Feb 86	21	Manure only Inorganic fertilizer only Manure & inorganic fertilizer Manure & feed Manure & inorganic fertilizer & feed No inputs	25.0	30.4	2.48
Cycle 3 Dry season	3 Sep 86	13 Nov 86	12	Chicken manure at 125, 250, 500 or 1,000 kg dry matter·ha^{-1}·week^{-1}	28.7	26.9	0.66

Table 2. Statistics of multiple regression model (equation 5) for predicting Z as a function of species "purity"[a] and salinity[b].

Independent Variables	Estimate	Stand. Error	Beta Coefficients[c)]	t
Intercept	-4.976	0.6047	-	-8.23
Salinity (S)	0.181	0.0189	1.044	9.6
Species purity[c] (P)	5.692	0.8221	3.066	6.92
Interaction (SxP)	-0.200	0.0264	-3.283	-7.58

[a]Species purity is a dummy variable with 0 = pure *O. niloticus* and 1 = *O. niloticus* contaminated with *O. mossambicus* genes.

[b]Overall model statistics are: multiple R: 0.79; R^2: 0.63; adjusted R^2: 0.61; standard error: 0.51; degrees of freedom due to regression: 3; due to residuals: 68.

[c]See Prein and Pauly (this vol.) for a definition.

were conducted in tanks, observations of daily mortalities were possible. The experimental methodology and some preliminary analyses are presented in Hopkins et al. (1989). The growth data were also used in the detailed presentation of the derivation, computation and attributes of the expanded Gulland and Holt plot (Hopkins et al. 1988b).

Data from a single experiment using a red tilapia from Taiwan are used in this example. These data were selected because they appeared to meet most of the assumptions for regression (see Hopkins et al. 1988b). Testosterone-treated red tilapia were stocked as 1-g fingerlings into 2-m^3 tanks in August 1983. They were harvested in March 1984 at an average weight of 132 g. Salinity was 38 to 41 ppt and temperatures ranged from 20 to 28°C. Mortalities were noted daily.

Observations during the experiments indicated a substantial increase in mortality immediately following sampling for fish size and when the temperatures decreased, in winter. To quantify the influences of handling and temperature on mortality, each sampling period was divided into two parts: a) the week immediately following sampling, and b) the remainder of the period. The number of mortalities during each part of the sampling period was determined from the daily mortality records. Instantaneous rates were computed for each subperiod and compared.

The method described in the previous paragraph is quite straightforward when all of the mortalities are accounted for. However, a few fish were unaccounted for at the end of the experiment. Thus, two instantaneous rates of mortality were required to account for all of deaths. Z_u, the rate of unrecorded mortality is defined as the mortality rate which accounts for deaths not recorded during the culture period. Z_u is assumed to be constant over the whole experiment. Z_r is the rate of recorded mortality. The total instantaneous rate of mortality during a particular time period, Z_t is the sum of the unrecorded and recorded mortality rates:

$$Z_t = Z_u + Z_r \qquad \textbf{...6)}$$

where Z_r is the instantaneous rate of recorded mortality during the period t.

The procedure used for computing Z_t and Z_u was:

1. Compute total mortality for the entire culture period on a daily basis:

$$Z = -\ln(N_h/N_s)\Sigma d \qquad \textbf{...7)}$$

   where N_h is the number of fish at harvest, N_s is the number of fish stocked, and Σd is the total duration of the experiment in days.

2. Compute Z_u by using the additive property of instantaneous rates. To do this, multiply Z_t by the per cent of total mortalities which were unrecorded:

$$Z_u = Z[1-\Sigma N_r/(N_s-N_h)] \qquad \textbf{...8)}$$

   where ΣN_r is the total number of mortalities recorded during the entire experiment.

After computing Z and Z_u, Z_r was computed for each part of the sampling periods. As the time periods were relatively short, the computations were done in a sequential fashion. For longer periods, the computational method for concurrent mortalities should be used (Ricker 1975).

3. Compute the number of fish included in unrecorded mortality (N_u) during the first week after sampling:

$$N_u = N_i - N_i \cdot e^{-Zut} \qquad \textbf{...9)}$$

   where N_i is the number of fish at the start of the period, Z_u is the rate of unrecorded mortality form equation 8 and t = 7 days.

4. Compute the number of fish remaining at the end of the first week after sampling:

$$N_t = N_i - N_u - N_r \qquad \textbf{...10)}$$

   where N_t is the number of fish at the end of period and N_r is the number of deaths during the period.

5. Compute the total rate of daily mortality (Z_t) during the first week after stocking:

$$Z_t = -\ln(N_t/N_i)/t \qquad \textbf{...11)}$$

6. Obtain Z_r by subtraction:

$$Z_r = Z_t - Z_u \qquad \textbf{...12)}$$

7. The process is then repeated for the second part of the sample period with N_t from the

preceding part of the period becoming the N_i for the subsequent part. The value of t changes depending upon the length of the period being examined. These sequential computations can be easily accomplished with a spreadsheet program.

All of the Z_r were converted to an annual basis and the effects of handling, temperature and fish weight on these rates were examined by computing the following multiple regression:

$$\ln(Z_r+1) = a + b_1H + b_2T \qquad \text{...13)}$$

where H is a dummy variable indicating the presence or absence of handling and T is the average daily water temperature (°C) during each period. As in the first example, the natural logarithm of Z was used to cause Z to level out near zero. The addition of 1 to Z_r was required because some of the values of Z were zero and the logarithm of zero is undefined. The resulting regression equation was:

$$\ln(Z_r+1) = 4.157 + 0.265H - 0.159T \qquad \text{...14)}$$

The regression output is presented in Table 3. The greater variability with corresponding decrease in the R^2 exhibited during this experiment compared with the pond experiment is probably a result of the shorter time periods. Also, preliminary analyses of the screening trials with other species indicated that mortality rates from short time periods may not always be normally distributed.

The analysis of the red tilapia data allowed the quantification of the effects of temperature and handling on mortality (Fig. 4). This information can be extremely useful when determining if the value of information from additional samples outweighs the resulting increase in mortality. Also, the positive effects of heating the water on mortality could be compared to the cost of heating.

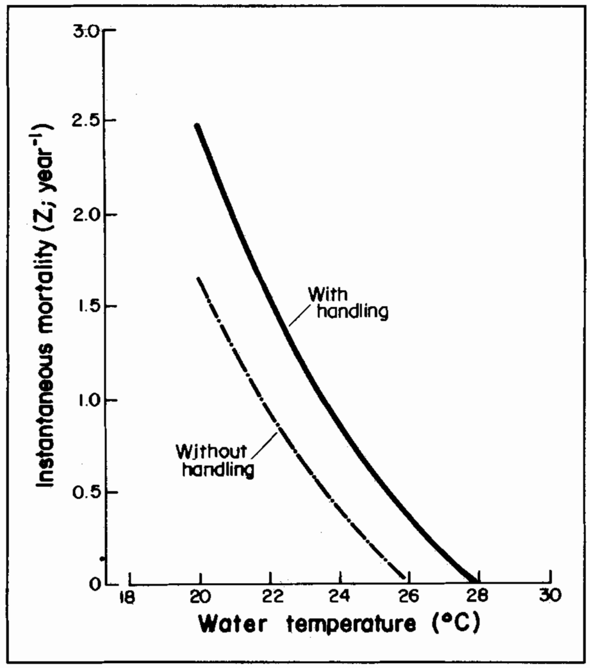


Fig. 4. Estimated effects of temperature and handling on the mortality of red tilapia in seawater.

Table 3. Statistics of multiple regression model [equation (14)] for predicting Z as a function of handling[a] and temperature[b].

Independent variables	Estimate	Standard error b	Standardized beta	T	Significance level of T
Constant	4.157	0.653	-	6.36	<0.0001
Temperature	-0.159	0.028	-0.657	-5.80	<0.0001
Handling	0.265	0.145	0.207	1.83	<0.0750

[a]Handling is a dummy variable with 0 = no handling at the start of the period and 1 = some handling.

[b]Overall model statistics are: multiple R: 0.69; R^2: 0.47; adjusted R^2: 0.45; standard error: 0.48; degrees of freedom due to regression: 2; due to residuals: 41.

Conclusion

Instantaneous rates of mortality can be used with multiple regression procedures to quantify the effects of external factors on mortality. These quantified effects can then be used in decisionmaking and for bioeconomic modelling.

Data collected over long periods (i.e., several months) is much easier to use than data collected over very short periods (i.e., 1 to 3 weeks). When working with short period data, care must be taken to ensure that the data meet the assumptions of linear regression.

Acknowledgements

The assistance of Carole Miura and Margarita Hopkins in reviewing the data and preliminary analyses is gratefully acknowledged. This paper is accession number 1054 of the Collaborative Research Support Program in Pond Dynamics/Aquaculture. Its preparation was supported in part by Grant No. DAN 4023-G-00-0031 from the US Agency for International Development.

References

Clifford, H.C. 1985. Semi-intensive shrimp farming, p. IV-15 to IV-42. *In* G. Chamberlain, M. Haby and R. Miget (eds.). Texas shrimp farming manual. An update on current technology. Texas Agricultural Extension Service, Corpus Christi, Texas.

Egna, H., N. Brown and M. Leslie. 1987. General reference: site descriptions, materials and methods for the global experiment. Pond Dynamics/Aquaculture Collaborative Research Data Reports. Vol. 1. PD/A CRSP, Oregon State University, Corvallis, Oregon.

Everhart, W.H. and W.D. Youngs. 1981. Principles of fishery science. Second edition. Cornell University Press, Ithaca, New York. 349 p.

Gomez, K.A. and A.A. Gomez. 1976. Statistical procedures for agricultural research, with emphasis on rice. International Rice Research Institute, Los Baños, Philippines. 294 p.

Hopkins, K.D., J. Lannan and J. Bowman. 1988a. Managing a database for pond research data - the CRSP experience. Aquabyte 1(1):3-4.

Hopkins, K.D., M.L. Hopkins and D. Pauly. 1988b. A multivariate model of tilapia growth, applied to seawater tilapia culture in Kuwait, p. 29-39. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Hopkins, K.D., M.Ridha, D. Leclercq, A.A. Al-Ameeri and T. Al-Ahmad. 1989. Screening tilapia for culture in seawater in Kuwait. Aquacult. Fish. Manage. 20:389-397.

Pauly, D. 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Stud. Rev. 8, 325 p.

PD/A CRSP. undated. CRSP work plan: first experimental cycle. Vol. 1. experimental protocol and methods. Oregon State University, Marine Science Center, Newport, Oregon.

PD/A CRSP. 1984. CRSP work plan: second experimental cycle. Oregon State University, Marine Science Center, Newport, Oregon.

PD/A CRSP. 1985. CRSP work plan: third experimental cycle. Oregon State University, Marine Science Center, Newport, Oregon.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191. 382 p.

Sparre, P., E. Ursin and S. Venema. 1989. Introduction to tropical fish stock assessment. Part 1 - Manual. FAO Fish. Tech. Pap. 306/1. FAO, Rome. 337 p.

Multiple Regression Analysis of Relationships Between Management Inputs and Fish Yield or Profit in Fish Polyculture Experimental Ponds*

GIDEON HULATA AND ANA MILSTEIN, *Agricultural Research Organization, Fish and Aquaculture Research Station, Dor, Mobile Post Hof Hacarmel 30820, Israel*

AMOS GOLDMAN, *Agricultural Research Organization, Department of Natural Resources, The Volcani Center, P.O. Box 6, Bet Dagan 50250, Israel*

Hulata, G., A. Milstein and A. Goldman. 1993. Multiple regression analysis of relationships between management inputs and fish yield or profit in fish polyculture experimental ponds, p. 112-118. *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

The multiple regression analysis presented in this paper was applied to data from experiments carried out during 1974-1985 at the Fish and Aquaculture Research Station, Dor, Israel. The objective of these experiments was to test the possibility of complete or partial substitution of supplemented feed by manure. Due to the large number of variables affecting fish yield, or profit, multivariate statistical methods were used.

Nonlinear relationships existed between fish yield (or profit) and each of the management variables (amounts of feed and manure, density and weight at stocking of common carp and tilapia). Multiple regression analysis identified stocking density of common carp and an interactive combination of amounts of feed and manure as the major factors affecting yield and profit. Manure had a stronger effect on profit than did the feeds. The highest profit was obtained in treatments in which a low ration of supplemental feed was added to a high amount of manure and not in the treatment in which only supplemented feed was applied.

The analysis presented here indicates the potential of multiple regression for the purpose of extracting information from fish culture data and identifying predictors of increased production and profit.

*ARO Contribution No. 40-H. Translation of an article published originally in Hebrew, in Israel Agresearch 4(1): 73-84, 1990.

Introduction

Polyculture of fish occurs in a complex ecological system, in which species with different feeding habits are stocked together in the same pond, with the aim of increasing fish production through optimal utilization of the various components of the natural food produced in the pond (Tang 1970). Herein, the amount of available natural food and the growth rate of the different fish species depend on environmental and management factors. Only the latter, such as stocking time, stocking density and size at stocking of the different fish species, quality and quantity of supplemented feed, and activities such as fertilization and manuring which increase natural food production (Wohlfarth and Hulata 1987), are controlled by the fish farmer. The large number of variables affecting fish production suggest the use of multivariate statistical methods. Such techniques were suggested or applied by, for example, Hopkins and Cruz (1982), Pauly and Hopkins (1983) for aquaculture data, by Winnans (1984) and by Milstein et al. (1985a, 1985b) for other areas of aquatic ecology, and by Ziv and Goldman (1987) for wheat culture data, among others.

In the analysis presented here, multiple regression was applied to data collected over 10 years of pond experiments at the Fish and Aquaculture Research Station, Dor. These data stem from experiments not specifically designed for the present study; rather they were generated in the frame of a study dealing with (partial) substitution of supplemental feeds by manure (Wohlfarth and Hulata 1987). Milstein et al. (1988) applied canonical correlation analysis to the dataset and reached the following conclusions: (a) the yield of each fish species was affected mainly by its own stocking density followed by interactions with other species; (b) the best yield and growth of tilapia were obtained with individual stocking weights of over 13 g; and (c) common carp (*Cyprinus carpio*) was affected negatively by silver carp (*Hypophthalmichthys molitrix*) density, and positively by nutrient inputs; its best performance was obtained at silver carp density below 1,000$\cdot$ha^{-1}.

The multiple regression treatment of the same dataset, presented in this paper, aimed at extracting the quantitative relations between the target variables (yield or profit) and the management variables in a way which enables prediction of the response to further changes in management variables.

Materials and Methods

Data from 105 polyculture ponds, used in different experiments carried out during 1974-1985 (including a total of 170 ponds), formed the analyzed dataset. These experiments were carried out in 400-m^2 ponds to which either 25% protein feed pellets and/or manure (cow or poultry) were applied as nutrient inputs. The dataset also included data from four 1,000-m^2 ponds in which mineral fertilizer (ammonium sulfate and superphosphate) was the only nutrient input (1983, treatments M and N, see Table 1). Ponds were not included in the analyzed dataset when mortality of fish was over 20%, or partial harvesting occurred during the culture season, or two size groups of any fish species were stocked. Treatments receiving sorghum as the nutrient input were also excluded from the final dataset due to their small number.

All experiments were initiated in July and terminated after approximately 4 months. The fish stocked included common carp (*Cyprinus carpio* L.), tilapia hybrids [mainly *Oreochromis niloticus* (L.) females crossed with either *O. aureus* (Steindachner) or *O. urolepis hornorum* (Trewavas) males], and silver carp [*Hypophthalmichthys molitrix* (Valenciennes)]. During 1979-1980, silver carp was substituted in some ponds by either bighead carp [*Aristichthys nobilis* (Richardson)] or silver carp X bighead carp hybrid. All these were pooled as "silver carp" in the dataset. Grass carp [*Ctenopharyngodon idella* (Valenciennes)] was also stocked in most of these experiments but is included only in the total yield due to its low stocking density. The fish were stocked at different weights and densities and different feeding and manuring rates were used for different experiments (Table 1). Further details of the experiments can be found in Milstein et al. (1988) and Wohlfarth and Hulata (1987).

For the analyses reported here, the amount of pelleted feed was calculated as the average daily amount per fish (carp and tilapia only, since silver carp does not feed directly on pellets, see Spataru 1977). The amount of manure is expressed in kg·ha^{-1}·day^{-1}. Growth rates and net yields (the difference between the biomass of harvested and stocked fish) were expressed as mean daily values, i.e., as g·day^{-1} per fish and kg·ha^{-1}·day^{-1}, respectively. These are the most important and meaningful parameters for the fish farmers.

Analyses were carried out using the GLM procedure included in the SAS (1985) package. The first response variable in the dataset was mean total yield (kg·day^{-1}) and the explanatory variables included the following management variables: mean daily amount of feed (g·fish^{-1}·day^{-1}), mean daily amount of manure (kg·ha^{-1}·day^{-1}), stocking density (no.·ha^{-1}) and mean weight at stocking of common carp, tilapia and silver carp. Stocking dates and the lengths of growing seasons were not included as variables since they were very similar in all experiments. The dataset did not contain all possible combinations among the management variables (since the experiments were not originally designed for a comprehensive multivariate analysis) and, furthermore, some of the management variables were significantly correlated (Table 2).

The GLM procedure used to estimate the regression function of yield (or profit) handles class variables, which have discrete levels, as well as continuous variables, which measure quantities. It also reports the explanation level (R^2) of the multiple regression model, the significance of the effect of each variable included in the model, its relative weight and its partial slope in the multiple regression equation. Models containing various combinations of management variables were tested and the criteria for their evaluation were: (i) the effects of all variables in the model are significant ($\alpha < 0.05$); (ii) the model is balanced, i.e., the Y-intercept is close to the mean, the estimated effects of the included explanatory variables are not extreme and the signs are plausible; and (iii) the multiple coefficient of determination (R^2) is high and the coefficient of variation (CV) of the partial slopes are low.

The procedure tests linear relationships of the target variable with the continuous variables. Since, in biological systems, nonlinear relationships are very common, the nature of the relationships was tested in several ways: through transformation of the values of the management variables (x_i) to their natural logarithms ($\ln x_i$) and/or inclusion of squared terms; and transformation of each continuous variable into a class variable with the range of values divided into three to five class levels, where the differences among these levels indicated by the analysis suggest the nature of the relationship. Attempts were made to develop models explaining

Table 1. Mean values (per treatment) of data used in the analysis.

Treatment	Year	Duration of test (days)	No. of ponds	Feed (kg·ha^{-1}·day^{-1})	Feed (g·fish^{-1}·day^{-1})	Manure (kg·ha^{-1}·day^{-1})	Carp Density (no·ha^{-1})	Carp Initial weight (g)	Carp Growth rate (g·day^{-1})	Carp Yield (kg·ha^{-1}·day^{-1})	Tilapia Density (no·ha^{-1})	Tilapia Initial weight (g)	Tilapia Growth rate (g·day^{-1})	Tilapia Yield (kg·ha^{-1}·day^{-1})	Silver Carp[a] Density (no·ha^{-1})	Silver Carp[a] Initial weight (g)	Silver Carp[a] Growth rate (g·day^{-1})	Silver Carp[a] Yield (kg·ha^{-1}·day^{-1})	Total Yield[b] (kg·ha^{-1}·day^{-1})
A	1974	126	4	125	7.6	0	11,450	22.4	3.68	33.2	5,000	4.0	1.15	3.9	2,500	133	6.50	9.3	49.7
B	1974	126	4	98	20.4	0	3300	22.4	8.35	21.8	1,500	4.0	1.50	1.7	1,250	132	9.05	6.6	31.6
C	1974	126	2	0	0	28	3050	22.4	3.60	10.7	1,500	4.0	1.20	1.0	1,250	109	8.95	7.1	19.8
D	1977	124	8	0	0	112	4,000	34.2	5.59	19.7	3,000	14.8	1.77	5.4	1,000	98	6.60	6.0	32.2
E	1978	123	7	0	0	97	4,000	25.6	4.67	18.2	3,000	14.5	1.79	4.7	1,000	104	6.54	5.8	30.1
F	1978	123	8	0	0	108	5,000	23.0	3.01	3.3	9,000	12.8	1.50	11.9	1,500	101	5.97	7.7	34.7
G	1979	109	6	23	1.5	160	5,300	35.3	5.07	23.8	9,200	10.2	1.23	9.2	2,000	111	5.33	8.5	44.7
H	1979	109	6	23	3.3	137	4,000	34.5	7.33	26.8	3,000	11.2	1.43	3.5	1,000	112	6.73	6.0	38.7
I	1980	115	9	24	1.7	111	5,200	24.0	3.79	16.7	9,000	9.3	1.4	8.9	1,200	131	5.88	4.8	33.0
J	1981	110	5	62	10.3	21	3,020	23.7	6.50	18.4	3,000	26.4	2.37	6.2	1,000	275	9.06	9.0	36.0
K	1982	117	9	26	2.4	113	3,050	40.3	4.02	11.0	7,500	20.7	1.91	13.7	1,000	76	6.93	6.4	36.1
L	1983	110	8	20	1.7	104	3,000	44.1	4.19	11.5	9,000	17.4	2.20	17.7	1,000	82	7.90	6.2	35.9
M	1983	154	2	0	0	0	1,500	23.1	2.85	3.8	1,500	0.5	1.75	2.5	1,300	56	5.40	6.5	12.8
N	1983	154	2	0	0	0	1,500	23.6	2.25	2.4	1,500	0.5	1.45	1.8	2,600	52	2.35	5.8	10.1
O	1984	103	8	41	3.4	40	3,000	36.4	7.00	20.3	9,000	1.0	1.16	8.1	500	117	10.70	5.0	34.7
P	1984	103	10	41	3.4	40	3,000	37.6	5.86	16.5	9,000	17.2	1.82	13.1	500	150	10.63	4.7	35.6
Q	1985	89	7	13	1.2	97	2,500	51.6	4.12	11.5	9,000	18.2	1.4	10.7	500	230	7.76	3.7	28.3

[a]In treatments G, H and I these were either silver carp, bighead carp or their hybrid.
[b]Including grass carp.

Table 2. Correlation matrix for management variables.

Management variables	Silver carp	Tilapia		Common carp			
	Density	Density	Weight	Density	Weight	Manure	Feed
Manure							-0.56***
Common carp density						0.03	0.08
Common carp weight					-0.41**	0.30***	-0.24***
Tilapia density				0.45***	0.02	0.31***	-0.31***
Tilapia weight			0.12***	0.36***	-0.25***	0.30	-0.09
Silver carp density		-0.33***	-0.22***	-0.52***	0.65***	0.01	0.05
Silver carp weight	-0.29***	-0.25***	-0.05	0.09	0.05	0.34***	0.34***

** = P<0.01; *** = P<0.001.

most of the variance in the target variables by inclusion of further variables, testing the interactions among them, and testing the residuals from the models (the values of the observations minus the values estimated by the model).

The economic value of yields in polyculture systems depends on the prices of the different species and their relative contribution to total yields. Profits also depend on the marginal effect of the production factors and on the price relationship between outputs and inputs. For our analyses, a new target variable, profit, was computed for each observation by summing the income from each of the three species (yield X price) and deducting from it the production costs, i.e., the sum of variable costs (quantity X price, for each species) + fixed costs (per ha and season) + marketing costs (per tonne). Prices used were supplied by Mr. G. Zohar (chief extension officer for fish culture in Israel) and are correct as of the end of December 1988. Since some of the above values are given as annual figures, the profits of all observations were standardized to a culture season of 250 days.

The data are described in Appendix II (Filename: DORPROD.WK1).

Results

Relationships Between Target Variables (Yield and Profit) and Management Variables

Pearson correlation coefficients (r) were calculated between each target variable and each management variable (x_i) and its transformations ($\ln x_i$ and x_i^2). Table 3 presents the 23 highest correlation coefficients found. It is seen that most correlation coefficients were increased by the log transformation. Hence, the relationships between management variables and the yield or profit are nonlinear indicating diminishing returns. The only two exceptions are the relationships between yield and manure and between profit and pellet input where significant negative correlations occurred only after squared terms for the amounts of manure or pellets were added.

Multiple Regression Analyses of the Target Variables (Yield and Profit)

Table 4 presents the two multiple regressions (one for yield and one for profit) which were selected from among many models with various combinations of the explanatory variables as those best representing the investigated systems. Density of common carp and the interaction of pelleted feed and manure were identified as the main predictors of the target variables (here, the nonlinearity of the relationships was handled by transforming the continuous explanatory variables into class variables).

The best yield model accounts for approximately 78% of the variance in yield whereas the best profit model accounts for only 65% of the variance in profit. The interaction terms for 'feed x manure' explains approximately 66% and 75% of the variance in yield and profit, respectively.

Table 5 presents the regression equations of the best models for yield and profit. The former shows a 1:1 substitution ratio between feed and manure at all levels. The latter shows that the interaction of feed and manure tends to increase the effect of manure on profit while the effect of feed is lower and becomes negative at high feeding levels. The density of common carp had a similar effect on both yield and profit. [Note, however, that the presumed positive role of a density of $11,450 \cdot ha^{-1}$ is based on only four observations from one experiment (treatment A in Table 1) and may thus not be conclusive. Also, no difference

Table 3. Values of the 23 highest estimates of Pearson's correlation coefficients (r) between target variables and the explanatory variables and/or their transforms[a]. (NIS = New Israeli Shekel; US$ 1 = NIS 2.)

Explanatory variable	Correlation coefficients with profit (NIS·0.1 ha^{-1}·$season^{-1}$)			Correlation coefficients with yield (kg·0.1 ha^{-1}·day^{-1})		
Transformation:	x_i^2	x_i	log x_i	x_i^2	x_i	log x_i
Density (no.·0.1 ha^{-1}) : Carp		.325	.490	.504	.568	.632
Tilapia			.371		.313*	.405
Feed : kg·0.1 ha^{-1}·day^{-1}				.380	.449	.474
g·$fish^{-1}$·day^{-1}	-.358	-.314				.354
Manure : kg·0.1 ha^{-1}·day^{-1}	.515	.557	.579	.308*		
g·$fish^{-1}$·day^{-1}	.450	.526	.561			
Initial weight : Tilapia			.400			.317*

[a]Values with an * indicate a probability of less than 0.001 that a result was a random effect. In all unmarked cases, this probability was less than 0.0001.

Table 4. Summary of GLM analyses of the best models for the target variables yield and profit.

Target variable: Description	Mean value	Explanatory variables and their weight in model: Variable	d.f.	SSD III[a]	Goodness of model: R^2	CV(%)[b]
Yield (kg·0.1ha^{-1}·day^{-1})	3.4	Manure level x Feed level	6	19.1	.784	11
		Common Carp density level	3	9.7		
Profit (NIS·0.1ha^{-1}·$season^{-1}$)	177	Manure level x Feed level	6	3,076	.651	9.8
		Common carp density level	3	925		

[a]Type III marginal sum of squares contributed to the model by each variable (after deduction of the effect of other variables).
[b]Coefficient of variation = [(mean standard deviation from model) x 100] / (mean value of target variable).

was found among the density levels of 3,000 to 5,200·ha^{-1}].

Each of the other management variables (i.e., its log transform in most cases) not included in the best models presented here had a positive effect in some of the other models tested: stocking density of common carp and tilapia generally affected both yield and profit positively while silver carp density and initial weight generally had negative effects. A negative effect was shown also for length of season, i.e., culture seasons terminating in September had higher mean daily yields than those terminating in October or later. The negative effect on profit was even stronger.

Discussion

The analyses presented are based on data from experimental ponds collected over 10 years. The major drawback of the dataset is its relatively small size and the unbalanced and incomplete representation of all management variables. Not all orthogonal combinations existed and, furthermore, strong collinearity existed among some of the explanatory variables (see Table 2) due to the discrete design of the experiments. A larger and more balanced dataset may lead to more conclusive results. Such a dataset, based on commercial fish farms, similar in nature to that analyzed by

Table 5. Summary of multiple regression coefficients of models (from Table 4) on target variables (yield and profit).

Explanatory variables	Combinations of levels: Variables — Feed	Manure	No. of observations	Yield model: Coeff. of regression equat.[b]	PR>|T|[a]	Profit model: Coeff. regression equat.[c]	PR>|T|[a]
Intercept				5.0	0.0001	249	0.0005
Feed x Manure ($kg \cdot 0.1\ ha^{-1} \cdot day^{-1}$)	0	0	4	-1.3	0.0001	58	0.7
	0	3	2	-1.8	0.0003	113	0.5
	0	11	23	0.1	0.7	656	0.0001
	2	11	39	0.4	0.08	537	0.0001
	2	14	6	1.3	0.0001	777	0.0001
	5	3	23	0.3	0.2	494	0.0001
	11	0	8	0[d]	0[d]	—	—
	Density levels						
Density of common carp ($no. \cdot 0.1\ ha^{-1}$)	200		11	-2.5	0.0001	-800	0.0001
	300		40	-1.7	0.0001	-583	0.0001
	520		50	-1.8	0.0001	-548	0.0001
	1,145		4	0[d]	0[d]	—	—

[a]The probability that the deviation from the reference level (zero, the last one) is due to random effects.
[b]In yield units, i.e., kg = 0.1 $ha^{-1} \cdot day^{-1}$.
[c]In profit units, i.e., NIS·0.1 $ha^{-1} \cdot season^{-1}$ (of 250 days).
[d]Reference level for class variables.

Ziv and Goldman (1987), is analyzed by Milstein et al. (this vol.).

The discrete correlation analysis (Table 3) identified density of common carp and the amounts of pelleted feed and of manure applied to the ponds as the strongest explanatory variables. The multivariate analysis gave a slightly different, more complex picture. The best models explained the system quite well, i.e., led to high R^2. The relations are mostly nonlinear, indicating decreasing marginal effects of explanatory variables. The multivariate analyses also indicated interactions among explanatory variables and the strongest one - that between amounts of feed and manure - was included in the models selected. Feed and manure seem to complement one another and combinations of these gave better performances than either of them separately. Note, however, that only two treatments (representing a total of eight ponds) received feed only. As an average of several models tested, feed and manure together contributed between 60% and 70% of the marginal SSD (indication of their weight in the model) while the contribution of each separately ranged between 0 (rare) and 60%. In yield models the effect of feed was generally higher, whereas in profit models the effect of manure was higher.

The variables "feed per ha" and "manure per fish" did not improve the models compared with "feed per fish" and "manure per ha" nor were they better correlated with the target variable. Manure may serve as direct feed for fish or act indirectly by boosting the natural food web (Wohlfarth and Schroeder 1979). In either case, it turned out to affect yields as much as feed did and had a stronger effect than feed on reduction of feeding costs (as indicated by the stronger effect on profit). The relatively low contribution of feed to profits, compared with manure, is due to the fact that the marginal effect of feed decreases when its amount is increased. Given the price ratio of feed and manure and the input to output ratio in present systems analyzed here, direct costs increase faster than the saving resulting from the fixed costs being divided over a higher yield when the amount of feed is increased. With manure, marginal production decreases more slowly when the amount of manure increases while cost is much lower than that for feed. Thus, feed and manure levels of 20 and 140 kg·ha^{-1}·day^{-1}, respectively, are more profitable than 20 and 110 kg·ha^{-1}·day^{-1}, respectively (Table 5).

The effect, though relatively small, of the length of culture season is obviously due to temperature which drops in autumn and reduces the growth rate.

The Israeli fish culture industry is characterized by strong differences in yield among regions, among farms of the same region (Sarig 1990) and even among ponds of the same farm. An understanding of the factors governing these fluctuations may lead to changes of management procedures to utilize all production elements optimally and to increase profitability. Regular and consistent data recording of culture conditions, preferably coupled with computer storage and analysis, and multivariate analysis as exemplified in this and other papers in this volume, should benefit farmers, extension workers and researchers.

References

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

Milstein, A., B. Hepher and B. Teltsch. 1985a. Principal component analysis of interactions between fish species and ecological conditions in fish ponds. I. Phytoplankton. Aquacult. Fish. Manage. 16:305-317.

Milstein, A., B. Hepher and B. Teltsch. 1985b. Principal component analysis of interactions between fish species and ecological conditions in fish ponds. II. Zooplankton. Aquacult. Fish. Manage. 16:319-330.

Milstein, A., G. Hulata and G. Wohlfarth. 1988. Canonical correlation analysis of relationships between management inputs and fish growth and yields in polyculture. Aquacult. Fish. Manage. 19:13-24.

Pauly, D. and K.D. Hopkins. 1983. A method for the analysis of pond growth experiments. ICLARM Newsl. 6:10-12.

Sarig, S. 1990. The fish culture industry in Israel in 1989. Isr. J. Aquacult./Bamidgeh 42:39-45.

SAS. 1985. SAS user's guide: statistics. Version 5 edition. Statistical Analysis Systems Institute Inc., Cary, North Carolina. 957 p.

Spataru, P. 1977. Gut contents of silver carp, *Hypophthalmichthys molitrix* (Val.), and some trophic relations to other fish species in a polyculture system. Aquaculture 11:137-146.

Tang, Y.A. 1970. Evaluation of balance between fishes and available fish foods in multispecies fish culture ponds in Taiwan. Trans. Am. Fish. Soc. 99:708-718.

Winnans, G.A. 1984. Multivariate morphometric variability in Pacific salmon: technical demonstration. Can. J. Fish. Aquat. Sci. 41:1150-1159.

Wohlfarth, G.W. and G. Hulata. 1987. Use of manure in aquaculture, p. 353-367. *In* D.J.W. Moriarty and R.S.V. Pullin (eds.) Detritus and microbial ecology in aquaculture. ICLARM Conf. Proc. 14, 420 p.

Wohlfarth, G.W. and G.L. Schroeder. 1979. Use of manure in fish farming - a review. Agric. Wastes 1:279-299.

Ziv, O. and A. Goldman. 1987. Comparison of methods for statistical analysis of wheat growth factors in Israel. Israel Agresearch 1(2):53-66. (In Hebrew, with English summary).

Factor Analysis and Canonical Correlation Analysis of Fish Production in Commercial Farms in Israel

ANA MILSTEIN and GIDEON HULATA, *Agricultural Research Organization, Fish and Aquaculture Research Station, Dor, Mobile Post Hof Hacarmel 30820, Israel*

Milstein, A. and G. Hulata. 1993. Factor analysis and canonical correlation analysis of fish production in commercial farms in Israel, p. 119-160. *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

Israeli commercial farms data were analyzed through multivariate analyses to study the relationships among fish species grown in a wide range of stocking combinations, culture seasons and management practices in different regions and farms. Data analyzed included 1,124 observations from well managed growout ponds from 31 farms located in four regions.

The results of the factor and canonical analyses applied to the dataset show that: a) yield of each species was mainly affected by its own stocking density; b) feed pellets mainly affected common carp with little or no effect on other species; c) tilapia accounted for most of the variability in the polyculture system. Pond size and depth had a relatively low influence on fish production, compared to the abovementioned variables. Some variation among regions, mainly attributed to different climatic conditions, were indicated.

Introduction

Polyculture of several fish species that feed on different natural resources is an important management technique to increase fish production. A successful polyculture system depends on a proper combination of ecologically different species, at adequate densities, which will efficiently utilize the available resources, maximize the synergistic fish-fish and fish-environment relationships, and minimize the antagonistic ones (Milstein 1990).

Originating in China during the Tang dynasty (618-907 AD) and widespread in the Far East (Tang 1970; Bardach et al. 1972; Chang 1987, 1989), polyculture was adopted by Israeli fish farmers in the 1970s (Hepher and Pruginin 1981; Sarig 1988). The species stocked into Israeli polyculture ponds are: common carp (*Cyprinus carpio*), hybrid tilapia (mainly *Oreochromis niloticus* x *O. aureus*), silver carp (*Hypophthalmichthys molitrix*), mullet (*Mugil cephalus*, and to a lesser extent, *M. capito*), and occasionally grass carp (*Ctenopharyngodon idella*). On national average they contribute about 60%, 20 to 30%, 5 to 8%, 6% and less than 2%, respectively, to the total fish production (Sarig 1990).

Attempting to increase production from polyculture systems requires a better understanding of the interactions among operational variables and their effects on fish performances. The large number of variables involved suggests the use of multivariate statistical methods. Milstein et al. (1988) applied canonical correlation to data from experimental ponds, accumulated over a period of 10 years at the Fish and Aquaculture Research Station, Dor. This study was limited by the number of observations and the uniformity of culture conditions (including season and fish combination). Nevertheless, it showed the usefulness and appropriateness of such methods to obtain new insights into the complexity of the system through analysis of existing data. The study presented here is an extension of such an application to commercial fish farm data, including a wide range of stocking combinations, culture seasons, and management practices in different regions and farms.

Materials and Methods

Data of Israeli commercial farms were analyzed through multivariate analysis to study the relationships among fish species cultured in balanced systems. Thus, the data included only well managed ponds in which mortality was relatively low (up to 20%).

The compilation of raw data from hand written data sheets (Fig. 8.24 in Hepher and Pruginin 1981) of commercial fish farms took two years and was finished with 1,725 culture season summaries from four geographical areas (Fig. 1), 31 farms and 11 years. Of these, about 400 are of carp or tilapia monoculture nursing ponds which were excluded from the present study. Growout ponds which were also used for nursing tilapia for a short period at the end of the season were not eliminated from the database. After extensive editing the final file contained 1,124 observations, not including nursing ponds, carp winter storage

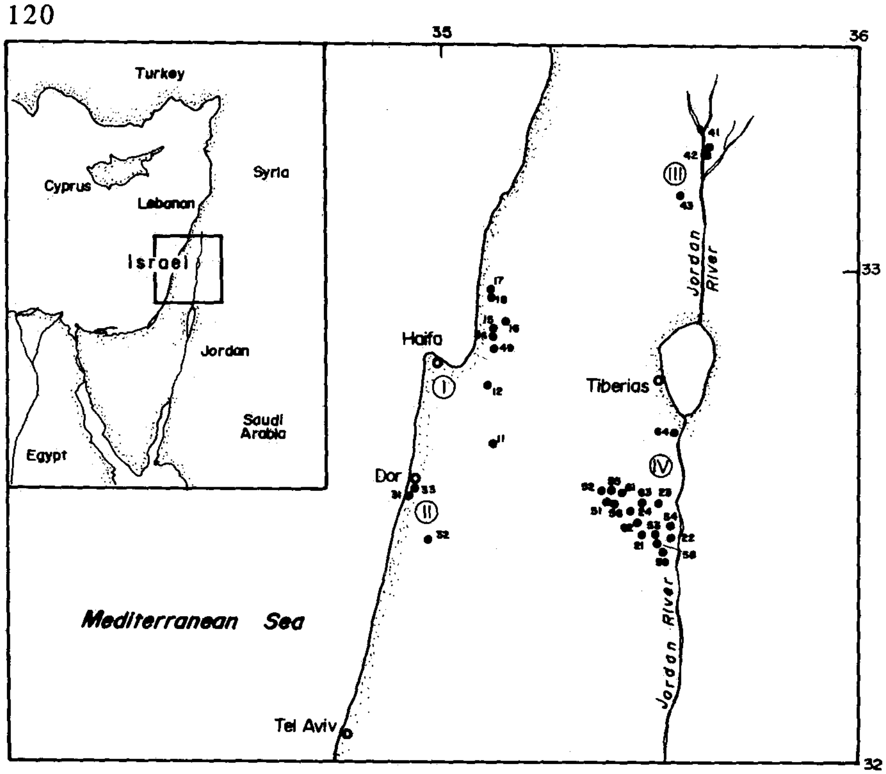


Fig. 1. Location of the fish farms in Israel. Regions: I = Western Galilee. II = Bet She'an Valley. III = Coastal Plain. IV = Upper Galilee. Farms: 11 = Hazorea, 12 = Kfar Macabi, 14 = Kfar Masarik, 15 = Yas'ur, 16 = Ein Hamifratz, 17 = Lohamei Hagetaot, 18 = Shomrat, 19 = Afek, 21 = Reshafim, 22 = Kfar Rupin, 23 = Hamadia, 24 = Nir David, 51 = Hefzibah, 52 = Ein Harod, 53 = Ein Hanatziv, 54 = Ma'oz Hayim, 55 = Tel Yosef, 56 = Bet Alfa, 57 = Agudat Mayim, 58 = Sde Eliyahu, 59 = Tirat Zvi, 61 = Bet Hashita, 62 = Mesilot, 63 = Sde Na'hum, 64 = Afikim, 31 = Ma'agan Michael, 32 = Gan Shemuel, 33 = Ma'yai Zvi. 41 = Daphna, 42 = Dan, 43 = Misgav Am.

ponds, outliers and incomplete records. This data set is available on diskette No. 4, the contents of which are described in Appendix II.

File Structure

Each observation consists of original and calculated values (Table 1), which were converted to units per 0.1 ha (=1 dunam, area unit used in Israel) and day, where applicable. Total yield includes grass carp and does not include wild spawning. Yields and growth rates were calculated as mean daily values for the season, considering only the days of effective growth, that is excluding the cold period between 15 November and 15 March of each year. The amounts of pelletized feed and sorghum were calculated as average daily amounts per fish (not including silver carp), weighted by each species' density and days of effective growth. The amount of manure is expressed as kg·0.1 ha·day^{-1}, since we believe its major contribution is in stimulating the pond's natural food web rather than as a direct feed (Wohlfarth and Schroeder 1979; Wohlfarth and Hulata 1987). Four groups of ponds were established, according to depth, which include ponds of 1 m and 2.5 m depth and reservoirs of 3.5 m and 5 m depth. These two categories differ in their water management, since in most cases reservoir water is used for crop irrigation in summer and therefore depth decreases gradually (Hepher 1985).

Observations were also classified according to the seasons (Table 2) and to fish species combinations (Table 3). These two classifications refer to the period the pond was operated and to the maximum number of fish species present during that period. Not all species in the polyculture are stocked at the same time and often differences of several months occur among stocking dates.

Different amounts of data were available from each geographical region (Table 4), most of them from the Western Galilee where shallow ponds prevail and the least from the Bet She'an Valley where most of the reservoirs for fish culture of the country are located. This is also reflected in the frequency distribution of depth (Table 4). Data from all ponds per farm and year were compiled, except for the Bet She'an Valley data from where the regional extension officer supplied selected data (i.e., excluding poor ponds of some farms).

Statistical Methods

Data were analyzed through factor analysis, canonical correlation and general linear model, using the SAS (1985) procedures FACTOR, CANCOR and GLM. These techniques allow to explore complex datasets and analyze interrelationships among variables. The rationale for using factor analysis and canonical correlations is discussed in Milstein (this vol.), and the use of GLM as a multiple regression model is shown in Milstein et al. (this vol.).

Table 1. List of variables and units.

Variable name	Definition	Units
Variables taken directly from farm records		
Region - farm - pond	**Identifiers**	
AREA		0.1 ha
DEPTH	• mean pond depth when full	m
DATEI	• stocking date of first fish (input)	
DATEO	• harvesting date (output)	
DAYS	• dateo minus datei	
CGRDAYS TGRDAYS MGRDAYS SGRDAYS	• number of effective growing days of carp, tilapia, mullet and silver carp, respectively (excluding days between 15 November and 15 March).	
CWTI TWTI MWTI SWTI	• initial weight of each fish	g
CWTO TWTO MWTO SWTO	• final weight of each fish	g
CDEN TDEN MDEN SDEN	• density of each fish (all fish harvested during the growing season)	fish·0.1 ha^{-1}
CYIELD TYIELD MYIELD SYIELD	• yield of each fish during the growing season	kg·0.1 ha^{-1}
SORG PEL MANU	• amounts of sorghum, feed pellets and manure applied during the growing season	kg·0.1 ha^{-1}
WILDKG	• wild spawning harvested	kg·0.1 ha^{-1}
GRASSKG	• grass carp yield	kg·0.1 ha^{-1}
Variables calculated		
TOTDEN	• total density	fish·0.1 ha^{-1}
TOTYIELD	• total yield (excl. wild spawning)	kg·0.1 ha^{-1}
CYIELDAY TYIELDAY MYIELDAY SYIELDAY	• yield of each species per day of effective growth	kg·0.1 ha^{-1}·day^{-1}
TOTYIDAY	• total yield per day of effective growth	kg·0.1 ha^{-1}·day^{-1}
CGROWTH TGROWTH MGROWTH SGROWTH	• daily growth rate of each fish species	g·day^{-1}
MANUDAY	• mean daily amount of manure applied during the days of effective growth	kg·0.1 ha^{-1}·day^{-1}
SORGFISH PELFISH	• mean daily amount of sorghum or feed pellets per fish (not including silver carp) during the days of effective growth	g·$fish^{-1}$·day^{-1}
SEASGRUP	• identifier of growing period	
FISHES	• identifier of fish species combination	

Table 2. Culture period identifier: values of variable SEASGRUP. wi = winter (ineffective growth period),sp = spring, su = summer, au = autumn. Asterisks denote season included in *all* cases.

```
SEASGRUP    wi sp su au wi sp su au
a           ===**
c           ===*****
                 *****
d           ===********
            ===********===
                 ********
                 ********===
e                    **
f                    *****
                     *****===
g                         **
                          **===
h                   ********====**
                        *****====**
                        *****====*****
                             **====**
                             **====*****
                             **====********
```

Table 3. Fish combination identifier: values of variable FISHES.

	Common carp	Tilapia	Mullet	Silver carp
FISHES				
1	*			
3	*	*		
4	*		*	
5	*			*
7	*		*	*
8	*	*		*
9	*	*	*	*

The general linear model (GLM) was used here as an unbalanced ANOVA to analyze the joint effect of categorical variables (FISHES, SEASGRUP, stocking month, geographical region and pond type) on the original variables, the extracted factors and the canonical variables. The ANOVA models test if there is a significant effect of the categorical variables, when considered together, on the dependent variable. They also give the proportion of total variance accounted for and indicate the relative contribution of each categorical variable to the variance. The ANOVA models are not the objective of the present analyses but rather the differences among levels of those categorical variables which are called "main effects". The differences between means of those main effects were tested with the Duncan multicomparison test. Since multicomparison tests of means take into consideration the dispersion of the data around the mean, no standard deviations are included in the corresponding results tables. In each column of these tables (Tables 5 to 11 and part b of the remaining ones) different letters identify means which are significantly different at the 95% level, 'a' indicating a group with larger mean than 'b', 'b' larger than 'c', etc. The letters are arranged in a way which allows an easy visualization of trends, almost simulating a graph. Note that the multicomparisons are univariate tests which relate the dependent variable to one categorical variable of the ANOVA model without considering the effect of the other variables of the model (thence their name 'main effects').

Results

The following sections contain tables presenting the results of the various analyses performed and their interpretation.

Descriptive Statistics of Israeli Fish Culture System

STOCKING CHARACTERISTICS

Fish are mainly cultured in short seasons of half a year (SEASGRUP c and f), or in a culture period of one year (SEASGRUP d). Besides, some very short culture cycles may take place during spring (a), summer (e) or autumn (g), or fish stocked late in the year may be kept in the ponds during the winter and continue growing in the next season (h) (Table 4).

The most frequent fish combinations used are carp-tilapia and carp-tilapia-silver carp followed by carp monoculture and carp-silver carp polyculture (Table 4). The inclusion of mullets in the polyculture is limited by the availability of this species for stocking.

The ANOVA model with five categorical variables (SEASGRUP, FISHES, stocking month of first species, geographical region and pond type) accounts for 45% of total density variability, in which FISHES is the variable that contributes the most. Total density (Table 5) is highest when common carp and tilapia but not mullets are stocked and lowest when common carp alone or with silver carp are stocked. It is significantly lower in short

Table 4. Frequency and percentage of obervations in each depth,fish combination (FISHES), region and culture period (SEASGRUP). c = common carp, t = tilapia, m = mullet, s = silver carp. In SEASGRUP, dashes indicates winter (either with or without) and x indicates growing period of variable length before and after winter.

Depth	Frequency	Percent	Region	Frequency	Percent
1	866	77.0	Western Galilee	482	42.9
2.5	135	12.0	Bet She'an Valley	117	10.4
3.5	51	4.5	Coastal Plain	269	23.9
5	72	6.5	Upper Galilee	256	22.8

FISHES	Frequency	Percent	SEASGRUP	Frequency	Percent
1 (c)	196	17.4	a (-sp)	100	8.9
3 (ct)	270	24.0	c (-spsu)	397	35.3
4 (cm)	65	5.8	d (-spsua)	230	20.5
5 (cs)	126	11.2	e (su)	61	5.4
7 (cms)	98	8.7	f (suau-)	215	19.1
8 (cts)	259	23.0	g (au-)	51	4.5
9 (ctms)	110	9.8	h (xx-xx)	70	6.2

Table 5. Mean density (fish·0.1 ha^{-1}) of each fish species in each culture period (SEASGRUP) and fish combination (FISHES). Number of observations (n) in each group in parenthesis. Total density and common carp have the same n in each SEASGRUP and all fish species in each FISHES group. For each species, analysis include only observations where that fish was present. Total n is: total density = 1,124; common carp = 1,124; tilapia = 639; mullet = 273; silver carp = 593. In each column, means with the same letter are not significantly different at the 0.05 level. a>b> etc.

	Total density			Carp		Tilapia			Mullet			Silver carp		
SEASGRUP														
a (- sp)	441	d	(100)	315	d	613	b	(14)	312		(2)	143	a	(20)
c (- sp su)	1,036	b c	(397)	453	a	860	a b	(240)	193		(79)	43	b	(210)
d (- sp su au -)	1,156	a b	(230)	478	a	860	a b	(122)	349		(127)	41	b	(154)
e (su)	997	c	(61)	379	b c	877	a b	(42)	85		(1)	32	b	(23)
f (su au -)	1,216	a	(215)	429	a b	961	a b	(168)	135		(18)	43	b	(124)
g (au -)	560	d	(51)	370	c	1,027	a	(8)	367		(2)	50	b	(14)
h (x x -x x)	1,049	b c	(70)	430	a b	724	a b	(45)	184		(44)	52	b	(48)
FISHES														
1 (c)	418	d	(196)	418	a b									
3 (c t)	1,409	a	(270)	382	b	1,027	a							
4 (c m)	1,034	b	(65)	434	a				600	a				
5 (c s)	541	d	(126)	468	a							72	a	
7 (c m s)	686	c	(98)	464	a				167	b		55	b	
8 (c t s)	1,361	a	(259)	454	a	871	b					34	c	
9 (c t m s)	1,149	b	(110)	456	a	509	c		147	b		37	c	
STOCKING MONTH OF FIRST SPECIES														
I	919	d e	(61)	411	a b	879	a b c d	(22)	363	a b	(28)	51	a b	(30)
II	939	c d e	(107)	466	a	625	d	(44)	432	a	(47)	47	b	(58)
III	980	c d	(237)	472	a	795	b c d	(124)	223	a b	(67)	45	b	(148)
IV	1,062	c d	(110)	407	a b	959	a b c	(65)	283	a b	(29)	38	b	(39)
V	1,249	a b	(106)	431	a b	1,116	a	(72)	318	a b	(12)	48	b	(53)
VI	1,346	a	(97)	411	a b	1,053	a b	(83)	246	a b	(5)	37	b	(57)
VII	1,109	b c	(113)	420	a b	863	a b c d	(86)	101	b	(10)	46	b	(55)
VIII	1,032	c d	(72)	432	a b	931	a b c	(44)	115	b	(6)	40	b	(39)
IX	724	f	(51)	379	b	755	b c d	(19)	185	a b	(11)	48	b	(25)
X	893	d e	(54)	410	a b	706	c d	(30)	153	b	(23)	42	b	(31)
XI	804	e f	(62)	419	a b	595	d	(29)	176	a b	22)	77	a	(32)
XII	784	e f	(54)	415	a b	787	b c d	(21)	141	b	(13)	51	b	(26)
REGION														
West Galilee	1,116	b	(482)	427	b	817	b	(353)	452	a	(79)	42	b	(200)
Bet She'an Valley	1,391	a	(117)	437	a b	995	a	(91)	255	b	(71)	37	b	(74)
Coastal Plain	734	d	(269)	467	a	646	c	(74)	136	b	(108)	47	a b	(185)
Upper Galilee	964	c	(256)	402	b	1,101	a	(121)	204	b	(15)	57	a	(134)
PONDTYPE														
Pond	1,019	a	(1,002)	425	b	921	a	(557)	295	a	(196)	47	a	(491)
Reservoir	1,016	a	(122)	482	a	572	b	(82)	180	b	(77)	41	a	(102)

seasons (SEASGRUP a and g), and higher when the culture period includes summer and autumn (SEASGRUPs d and f). Ponds whose stocking begins in May-June are stocked at a higher density than those started in the cold months. The highest total densities are practised in the Bet She'an Valley. These are about twice as high as the ones practised in the Coastal Plain, which are the lowest of the four regions. Total density in ponds and reservoirs are not significantly different per unit area.

Common carp is cultured in all SEASGRUP periods and in all fish combinations included in this study. It is stocked at a wide range of densities and initial weight combinations (Fig. 2A), which are not strongly dependent on the five categorical variables tested here (only 12% of its density and 39% of its stocking weight variability are accounted for by the five categorical variables in the ANOVA models, to which SEASGRUP is the variable that contributes the most). Mean density of common carp is lower in the short culture periods (SEASGRUPs a, e and g), when grown with tilapia (FISHES=3), in the ponds than in the reservoirs, and in the Galilees than in the Coastal Plain (Table 5). Common carp initial weight (Table 6) differ according to stocking months and culture periods. In the long culture seasons (SEASGRUP d) the

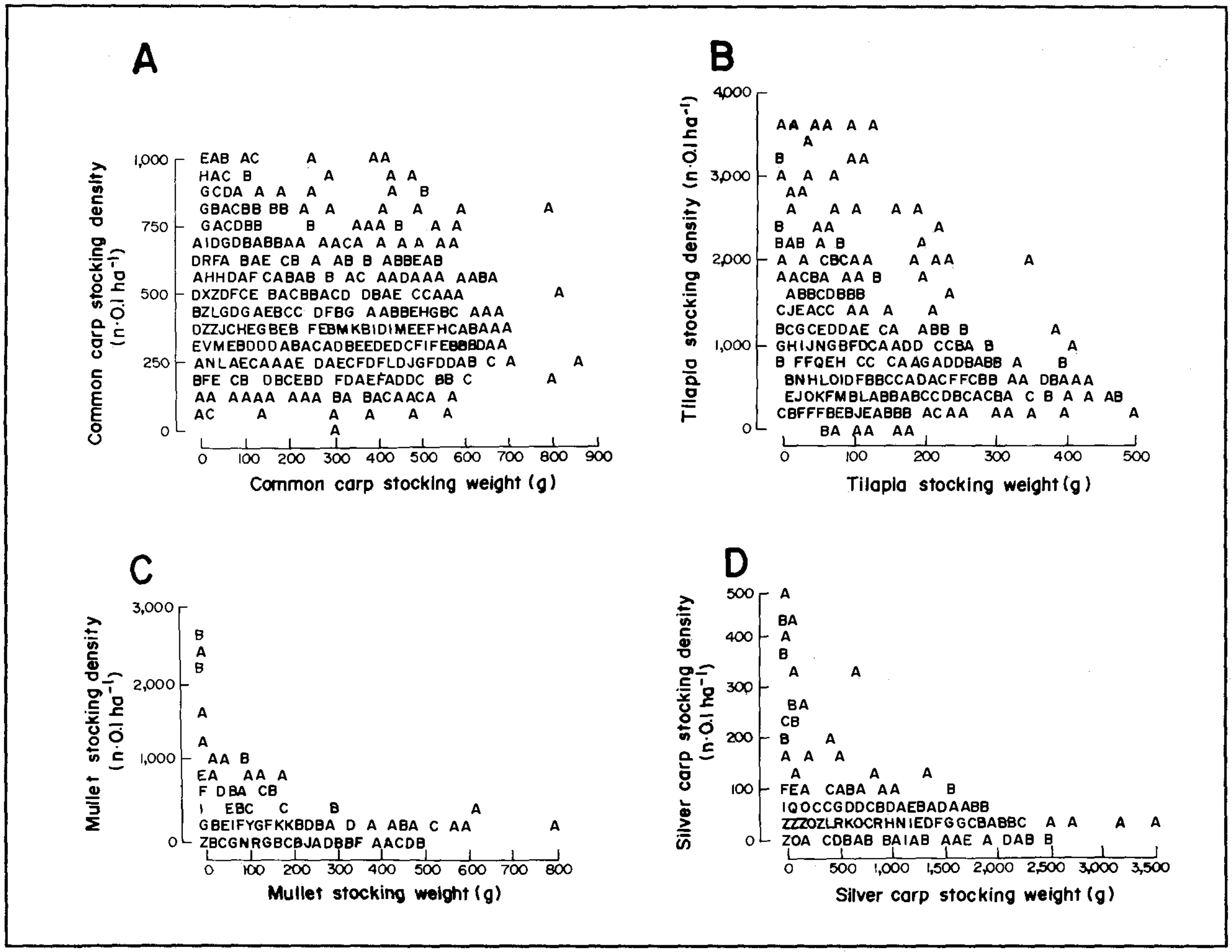


Fig. 2. Relationships between stocking weights and stocking densities grown in polyculture in commercial farms in Israel. A. Common carp. B. Tilapia. C. Mullet. D. Silver carp.

smallest carp (about 70 g) are stocked. Small carp are also stocked during the hot period and allowed to continue growing after the winter (SEASGRUP h). Larger carp (170 g) stocked in winter or spring are cultured in half-year seasons (SEASGRUP c). Carp of about 300 g are stocked in summer and grown for very short (e) or half-year (f) periods, or kept in the ponds during winter and are allowed to grow for a very short period in spring (a). The largest carp stocked are those which did not reach commercial size by the end of the hot period and are allowed to grow during autumn before harvesting (g). Initial weight also is largest (about 300 g) in monoculture and smallest (about 100 g) when mullets are also stocked (FISHES=4, 7, 9). Reservoirs are predominantly stocked with smaller carp in contrast to ponds. In the Bet She'an Valley stocking sizes of carp are smaller compared to the other regions.

Tilapia was present in 57% of the analyzed ponds. Only 22% of its density and 42% of its stocking weight variability are accounted for by the five categorical variables in the ANOVA models, to which region and stocking month, respectively, are the variables that contribute the most. Tilapia densities (Table 5) in the Bet She'an Valley and Upper Galilee are higher than in the western regions. Tilapia density is higher when stocked in the hot months but is not significantly different among culture periods except for a few observations in autumn which include some tilapia nurseries together with other fish. When stocked with common carp alone, tilapia density is higher than when two or three species are stocked in the same pond, partly due to the inclusion of some tilapia nurseries in carp ponds and partly to its replacement by the other fish as a management policy. Tilapia density in ponds is almost twice as high as in reservoirs. Tilapia of a wide range of initial sizes are stocked at densities up to 15,000 fish·ha^{-1}. Above that density, only fish of less than 300 g are cultured (Fig. 2B). These two variables are not inversely correlated as might be expected, and most of the observations concentrate in the area of small fish (up to 100 g) at low densities (up to 10,000 fish·ha^{-1}). Tilapia initial weight is larger when these are stocked with common carp only, in contrast to cases when other species are present (Table 6). The smallest tilapia are stocked in spring (overwintered from the previous year) and cultured over different season lengths; tilapia stocking in spring is done also in ponds which were already stocked in winter with other fish species (SEASGRUP h). Larger tilapia are stocked in summer and cultured for very short periods or during the second half of the year. As in the case with common carp, the largest tilapia stocked are those which did not reach commercial size by the end of the hot period and are allowed to grow during autumn before harvesting (SEASGRUP g). Larger tilapia are stocked into ponds than into reservoirs. Among the regions, the largest tilapia are stocked in the Western Galilee and the smallest in the Coastal Plain and Upper Galilee.

Table 6. Mean initial weight (g) of each fish species in each culture period (SEASGRUP) and fish combination (FISHES). Number of observations (n) in each group in parentheses. For each species, analyses include only observations where that fish was present. Total n is: total density = 1,124; common carp = 1,124; tilapia = 639; mullet = 273; silver carp = 593. In each column, means with the same letter are not significantly different at the 0.05 level. a>b> etc.

	Common carp			Tilapia			Mullet			Silver carp		
SEASGRUP												
a (- sp)	321	b	(100)	105	c	(14)	92		(2)	352	d	(20)
c (- sp su)	176	c	(397)	71	c d	(240)	157		(79)	383	d	(210)
d (- sp su au -)	76	e	(230)	54	d	(122)	113		(127)	299	d	(154)
e (su)	337	b	(61)	164	b	(42)	100		(1)	677	b c	(23)
f (su au -)	304	b	(215)	165	b	(168)	235		(18)	857	b	(124)
g (au -)	405	a	(51)	252	a	(8)	215		(2)	1,198	a	(14)
h (x x - x x)	128	e	(70)	79	c d	(45)	140		(44)	459	c d	(48)
FISHES												
1 (c)	314	a	(196)									
3 (c t)	233	b c	(270)	115	a	(270)						
4 (c m s)	94	d	(65)				97	b	(65)			
5 (c s)	250	b	(126)							530	a	(126)
7 (c m s)	92	d	(98)				126	b	(98)	409	a	(98)
8 (c t s)	203	c	(259)	94	b	(259)				515	a	(259)
9 (c t m s)	107	d	(110)	91	b	(110)	175	a	(110)	490	a	(110)
STOCKING MONTH OF FIRST SPECIES												
I	180	d e	(61)	65	d e	(22)	189	a b c	(28)	495	c d	(30)
II	148	e f	(107)	64	d e	(44)	100	c d	(47)	302	d	(58)
III	116	f	(237)	57	e	(124)	136	a b c d	(67)	280	d	(148)
IV	188	d e	(110)	64	d e	(65)	79	d	(29)	524	c d	(39)
V	223	c d	(106)	71	d e	(72)	90	c d	(12)	297	d	(53)
VI	267	b c	(97)	116	b c	(83)	181	a b c d	(5)	557	c	(57)
VII	329	a	(113)	193	a	(86)	241	a	(10)	913	a b	(55)
VIII	339	a	(72)	197	a	(44)	214	a b	(6)	1,119	a	(39)
IX	306	a b	(51)	133	b	(19)	162	a b c d	(11)	844	b	(25)
X	222	c d	(54)	94	c d e	(30)	167	a b c d	(23)	447	c d	(31)
XI	186	d e	(62)	98	c d	(29)	124	b c d	(22)	367	c d	(32)
XII	181	d e	(54)	73	d e	(21)	172	a b c d	(13)	458	c d	(26)
REGION												
West Galilee	214	a	(482)	119	a	(353)	139	a	(79)	514	a	(200)
Bet She'an Valley	144	b	(117)	98	b	(91)	165	a	(71)	527	a	(74)
Coastal Plain	217	a	(269)	78	c	(74)	129	a	(108)	549	a	(185)
Upper Galilee	221	a	(256)	71	c	(121)	65	b	(15)	381	b	(134)
PONDTYPE												
Pond	226	a	(1,002)	107	a	(557)	124	b	(196)	515	a	(491)
Reservoir	68	b	(122)	66	b	(82)	172	a	(77)	408	b	(102)

Mullets occurred in 24% of the studied ponds. About 41% of its density and 20% of its stocking weight variability are accounted for by the five categorical variables in the ANOVA models, to which FISHES and stocking month, respectively, are the variables that contribute the most. Mullets are stocked at higher densities in the carp-mullet combination than when other fish are also present (Table 5). This situation prevails to a greater extent in the Western Galilee than in other regions and also more in ponds than in reservoirs. Mullets of larger size (Table 6) are stocked when all four fish species participate in the polyculture and in the reservoirs. Among the regions, in the Upper Galilee smaller mullets are stocked than in the other regions. Mullets of all sizes are stocked at densities of up to 4,000 fish·ha^{-1}. Above that density, only fish smaller than 200 g are stocked (Fig. 2C). Multicomparisons by SEASGRUP are not sensible due to the highly unbalanced distribution of observations in each SEASGRUP (Tables 5 and 6).

Silver carp was present in 53% of the analyzed ponds. Only 22% of its density and 27% of its stocking weight variability are accounted for by the five categorical variables in the ANOVA models, to which SEASGRUP is the variable that contributes the most. It is stocked at the highest density (Table 5) when cultured in the very short spring period, when stocked in November, and when tilapia are not present. Silver carp density was not significantly different among pond types. Higher densities are practised in the Upper Galilee than in the Western Galilee and Bet She'an Valley. Fish of all sizes are stocked at densities of up to 1,300 fish·ha^{-1} but fish smaller than 500 g are cultured also at higher densities (Fig. 2D). The smallest silver carp (less than 500 g) are stocked in winter or spring, larger ones (700 to 800 g) in summer and the largest (more than 1 kg) in autumn (Table 6). There are no differences in mean silver carp stocking weight among the different fish combinations. Silver carp stocked in ponds are larger than those stocked in reservoirs. In the Upper Galilee, smaller silver carp are stocked than in the other regions.

NUTRITIONAL INPUTS

Feed pellets were applied in 86% of the analyzed ponds. Although in this analysis feed pellets are considered as a single input, there are several different types of pellets (Hepher 1989). Farmers generally apply low protein level (12%) pellets during spring, since there are large amounts of natural food in the ponds in this season. From May until about August 25% protein pellets are given to the fish and until the end of the culture season fish are fed with 25% or 30% protein pellets.

Since the information on pellet type given to each pond is not available, the variable PELFISH represents the total amount of all types applied. Only 23% of PELFISH variability is accounted for by the five categorical variables in the ANOVA models to which FISHES is the variable that contributes the most followed by SEASGRUP. Feed pellets are given at the highest rate when carp, mullets and silver carp (FISHES=7) are present. High rates are also given in the fish combinations which include tilapia (Table 7). It should be noted that the numbers of the fish and not their weight are considered in the calculation of PELFISH. Higher amounts of pellets per fish are used in the second half of the year (SEASGRUP e, f and d), since in summer and autumn the amount of natural food in the ponds is small due to the large grazing pressure. This is also seen in the distribution by stocking month. Similar feeding rates are applied in ponds and reservoirs and lower rates are given in the Upper Galilee than in the other regions.

Sorghum was applied in 70% of the analyzed ponds. About 40% of SORGFISH variability is accounted for by the five categorical variables in the ANOVA models to which region is the variable that contributes the most followed by SEASGRUP. Strong differences occur among regions (Table 7), where rates of sorghum application in the Upper Galilee are almost twice as high as those of the Coastal Plain, which are in turn almost twice as high as those of the Bet She'an Valley, which are more than thrice as high as those of the Western Galilee. This nutritional input is given at high rates in spring (SEASGRUP a), lower rates in autumn (SEASGRUP g), and at the lowest in the other SEASGRUPs. It is mainly used in carp monoculture (FISHES=1) followed by carp with silver carp (FISHES=5) and the combinations of these fishes with mullets (FISHES=7 and 9). The lowest amounts are applied when tilapia are cultured in the ponds. Higher rates of sorghum are applied in ponds than in reservoirs.

Manure (Table 7) was applied in 48% of the analyzed ponds. About 34% of MANUDAY variability is accounted for by the five categorical variables in the ANOVA models to which region is the variable that contributes the most followed by

Table 7. Mean inputs (feed pellets and sorghum in g·fish^{-1}·day^{-1}, manure in kg·0.1 ha^{-1}·day^{-1} in each tested categories. Total number of observations is 1,124. n = number of observations in each group. In each column, means with the same letter are not significantly different at the 0.05 level. a>b>etc.

	n	PELFISH	SORGFISH	MANUDAY
SEASGRUP				
a (- sp)	100	1.4 d	10.1 a	0.9 d
c (- spsu)	397	5.9 c	2.4 c	2.3 b c
d (- spsuau-)	230	7.1 b c	2.1 c	2.9 b c
e (su)	61	9.0 a	2.4 c	3.3 a
f (suau-)	215	7.8 b	1.8 c	2.9 a b
g (au-)	51	6.5 c	5.1 b	1.9 c
h (xx - xx)	70	6.1 c	2.3 c	1.8 c
FISHES				
1 (c)	196	3.7 d	7.0 a	1.2 c d
3 (c t)	270	6.7 b c	1.0 d	4.1 a
4 (c m)	65	5.9 c	1.7 d	3.0 b
5 (c s)	126	5.6 c	4.9 b	0.5
7 (c ms)	98	9.5 a	3.1 c	1.4 c
8 (c t s)	259	6.6 b c	1.6 d	2.0 c
9 (c t ms)	110	7.7 b	2.9 c	3.1 b
STOCKING MONTH OF FIRST SPECIES				
I	61	4.4 f g	5.2 b	2.1 b c
II	107	5.5 d e f	3.1 c d	2.5 b c
III	237	5.8 c d e f	2.0 d e	1.8 b c
IV	110	6.5 b c d	2.0 d e	1.9 b c
V	106	7.1 b c	3.1 c d	2.8 a b
VI	97	7.8 a b	2.0 d e	2.7 a b c
VII	113	8.6 a	1.5 e	3.5 a
VIII	72	7.3 a b c	2.8 c d e	2.5 b c
IX	51	6.6 b c d	3.7 c	1.8 b c
X	54	6.1 c d e	3.4 c d	2.4 b c
XI	62	4.9 e f g	6.7 a	2.1 b c
XII	54	3.5 f g	6.5 a	1.7 c
REGION				
West Galilee	482	7.1 a	0.6 d	4.2 a
Bet She'an Valley	117	6.5 a	2.1 c	2.5 b
Coastal Plain	269	6.3 a	3.8 b	0.2 c
Upper Galilee	256	4.8 b	7.1 a	0.7 c
PONDTYPE				
Pond	1,002	6.3 a	3.1 a	2.3 a
Reservoir	122	6.4 a	2.5 b	2.4 a

FISHES. Strong differences occur among regions since manure is hardly used in the Coastal Plain and Upper Galilee while being applied more heavily in the Western Galilee than in the Bet She'an Valley. Manure is mainly used in the carp-tilapia combination (FISHES=3) followed by all four fishes and carp-mullet (FISHES=9 and 4). It is given at the highest rates in summer (SEASGRUP e) and the lowest in spring (SEASGRUP a). It is applied at similar rates in ponds and reservoirs on a per-area basis.

FISH OUTPUTS

The overall mean daily total yield was 26 kg·ha^{-1}·day^{-1}. Differences in total yield among several categorical divisions are presented in Table 8. The highest total daily yields were obtained when fish were stocked in summer, cultured during summer (SEASGRUP e), in polyculture including tilapia (FISHES=3, 8 and 9), and in the Bet She'an Valley area. The lowest total daily yields were obtained during the short cold seasons (SEASGRUP a and g), when stocking was in the cold months, in carp monoculture or in polyculture without tilapia, and in the Coastal Plain. No differences in total daily yield occurred between reservoirs and regular ponds.

Common carp (Table 8) mean daily yield was 15 $kg \cdot ha^{-1} \cdot day^{-1}$. It was very similar in all the analyzed categories, although there are some significant differences. Carp mean daily growth rate was 4 $g \cdot day^{-1}$. It was higher when grown during short SEASGRUPs (a, e and g), when stocked between August and December, in monoculture, in the Upper Galilee, and in shallow regular ponds. Mean harvest weight of carp was 680 g. The largest common carp were harvested when cultured during long culture seasons (SEASGRUPs d and h), in polyculture with mullet alone followed by polyculture of at least three species, in reservoirs and in the Bet She'an Valley (where most of the reservoirs are located).

Tilapia (Table 9) mean yield was 15 $kg \cdot ha^{-1} \cdot day^{-1}$. It was higher in summer, when stocking was between April and July, in the Bet

Table 8. Duncan's multicomparison test of means. Mean total yield ($kg \cdot 0.1\ ha^{-1} \cdot day^{-1}$) and common carp yield ($kg \cdot 0.1\ ha^{-1} \cdot day^{-1}$) growth ($g \cdot day^{-1}$) and harvest weight (g) in each culture period (SEASGRUP), fish combination (FISHES), stocking month, pond type and region. n = number of observations in each group. Total n = 1,124. In each column, means with the same letter are not significantly different at the 0.05 level. a>b> etc.

	n	TOTYIDAY	CYIELDAY	CGROWTH	CWTO
SEASGRUP					
a (- sp)	100	1.7 d	1.4 b	5.4 a	644 c d
c (- spsu)	397	2.6 b c	1.6 a b	3.8 b	600 d
d (- spsuau-)	230	2.6 b c	1.6 a b	3.6 b	829a
e (su)	61	3.3 a	1.8 a	5.4 a	672 c
f (suau-)	215	2.3 b	1.4 b	3.7 b	687 c
g (au-)	51	1.9 d	1.6 a b	5.3 a	650 c d
h (xx- xx)	70	2.4 c	1.4 b	3.6 b	767 b
FISHES					
1 (c)	196	1.7 d	1.7 a	5.1 a	634 d
3 (c t)	270	3.1 a	1.3 c	4.0 b c	640 d
4 (c m)	65	2.0 c d	1.4 b c	3.6 c d	788a
5 (c s)	126	1.9 c d	1.5 a b c	4.0 b c	672 c d
7 (c m s)	98	2.1 c	1.4 b c	3.4 d	720 b c
8 (c t s)	259	3.2 a	1.6 a b	4.1 b	708 b c
9 (c t m s)	110	2.8 b	1.4 c	3.3 d	746 b c
STOCKING MONTH OF FIRST SPECIES					
I	61	2.1 d	1.3 c	3.9 d e f	736a b c
II	107	2.3 c d	1.5 a b c	3.6 e f	686 b c d e f
III	237	2.4 b c d	1.5 a b c	3.5 f	628 f
IV	110	2.7 b	1.4 b c	4.0 d e f	649 e f
V	106	3.0 a	1.6 a b c	4.2 c d	705a b c d e
VI	97	3.2 a	1.5 a b c	4.1 c d e	712a b c d e
VII	113	3.0 a	1.4 b c	3.9 d e f	780 c d e f
VIII	72	2.6 b	1.5 a b c	4.3 b c d	661 d e f
IX	51	2.2 c d	1.5 a b c	4.7 a b c	683 c d e f
X	54	2.3 c d	1.5 a b c	4.3 b c d	730a b c d
XI	62	2.5 b c	1.7 a	4.8 a b	760a
XII	54	2.2 c d	1.6 a b	5.1 a	757a b
REGION					
West Galilee	482	2.8 b	1.5 b c	3.9 b	634 d
Bet She'an Valley	117	3.3 a	1.6 a b	4.1 b	780a
Coastal Plain	269	1.9 d	1.3 c	3.4 c	678 c
Upper Galilee	256	2.5 c	1.7 a	5.1 a	740 b
PONDTYPE					
Pond	1,002	2.6 a	1.5 a	4.1 a	670 b
Reservoir	122	2.6 a	1.6 a	3.6 a	797a

Table 9. Duncan's multicomparison test of means. Mean tilapia yield (kg·0.1 $ha^{-1} \cdot day^{-1}$) growth ($g \cdot day^{-1}$) and harvest weight (g) in each culture period (SEASGRUP), fish combination (FISHES), stocking month, pond type and region. n = number of observations in each group. Total n = 639. In each column, means with the same letter are not significantly different at the 0.05 level. a>b> etc.

	n	TYIELDAY	TGROWTH	TWTO
SEASGRUP				
a (- sp)	14	1.2 b c	2.2 b	227 b
c (- spsu)	240	1.6 a b	2.2 b	264 b
d (- spsuau-)	122	1.4 b c	1.9 b	373 a
e (su)	42	2.0 a	2.8 a	353 a
f (suau-)	168	1.5 a b	2.0 b	357 a
g (au-)	8	1.6 a b	2.1 b	344 a
h (xx- xx)	45	1.0 c	2.0 b	382 a
FISHES				
3 (c t)	270	1.9 a	2.2 a	320 b
8 (c t s)	259	1.4 b	2.0 b	293 c
9 (c t m s)	110	0.9 c	2.1 a b	403 a
STOCKING MONTH OF FIRST SPECIES				
I	22	1.5 a b c d	2.0 a	360 a b
II	44	1.2 d e	2.1 a	289 c
III	124	1.4 b c d	2.0 a	265 c
IV	65	1.7 a b c	2.1 a	291 c
V	72	1.9 a	2.1 a	293 c
VI	83	1.7 a b	2.0 a	306 b c
VII	86	1.7 a b c	2.3 a	402 a
VIII	44	1.3 b c d e	2.1 a	360 a b
IX	19	1.2 c d e	2.1 a	370 a
X	30	0.9 e	2.0 a	374 a
XI	29	1.2 d e	2.2 a	396 a
XII	21	1.3 b c d e	2.2 a	385 a
REGION				
West Galilee	353	1.6 b	2.2 a	322 b
Bet She'an Valley	91	1.8 a	2.3 a	423 a
Coastal Plain	74	0.8 c	1.7 b	294 b c
Upper Galilee	121	1.5 b	1.7 b	271 c
PONDTYPE				
Pond	557	1.6 a	2.1 a	312 b
Reservoir	82	1.1 b	2.1 a	404 a

She'an Area, in ponds. In relation to fish combination, each fish species added to the polyculture decreased tilapia daily yield. Tilapia mean growth rate was 2.1 $g \cdot day^{-1}$. It was higher when cultured during summer and in the Western Galilee and Bet She'an Valley. Mean tilapia harvest weight was 320 g. It was lower when grown during spring or spring-summer (SEASGRUPs a and c), stocked in the first half of the year, in the Upper Galilee and Coastal Plain. It was higher when all four fish species were present, in the Bet She'an Valley, and in reservoirs.

Mullets (Table 10) were present in very few observations in several SEASGRUPs and stocking months, which will not be considered in the following analyses. Mean yield of mullet was 3 $kg \cdot ha^{-1} \cdot day^{-1}$. Higher values occurred when stocked with common carp alone (FISHES=4), in the Western Galilee and Bet She'an Valley and in reservoirs. Mullet mean growth rate was 2.1 $g \cdot day^{-1}$. It was higher in the presence of silver carp (FISHES=7 and 9), in Bet She'an Valley and the Coastal Plain and in reservoirs. Mullet mean harvest weight was 520 g. It was higher in the presence of silver carp, when stocking occurred in winter months, in the Bet She'an Valley and in reservoirs.

Silver carp (Table 11) mean daily yield was 3 $kg \cdot ha^{-1} \cdot day^{-1}$. It was higher when tilapia was not present and in reservoirs. Silver carp mean growth

Table 10. Duncan's multicomparison test of means. Mean mullet yield ($kg \cdot 0.1\ ha^{-1} \cdot day^{-1}$) growth ($g \cdot day^{-1}$) and harvest weight (g) in each culture period (SEASGRUP), fish combination (FISHES), stocking month, pond type and region. n = number of observations in each group. Total n = 273. In each column, means with the same letter are not significantly different at the 0.05 level. a>b> etc.

	n	MYIELDAY	MGROWTH	MWTO
SEASGRUP				
a (- sp)	2	0.4 a b	1.8 b	240 a
c (- sp su)	79	0.3 a b	2.4 a b	500 a
d (- sp su au-)	127	0.4 a b	1.9 b	522 a
e (su)	1	0.1 b	1.4 b	200 a
f (su au-)	18	0.2 a b	1.7 b	410 a
g (au-)	2	0.5 a	3.8 a	423 a
h (xx- xx)	44	0.4 a b	2.4 a b	618 a
FISHES				
4 (c m)	65	0.6 a	1.6 b	382 b
7 (c ms)	98	0.2 b	2.3 a	533 a
9 (c t ms)	110	0.3 b	2.3 a	590 a
STOCKING MONTH OF FIRST SPECIES				
I	28	0.4 a b	1.9 b c	566 a b c
II	47	0.4 a b	1.9 b c	458 b c d
III	67	0.3 b c d	2.1 b c	542 a b c
IV	29	0.3 b c	2.2 b	418 b c d
V	12	0.3 b c	1.7 c	392 c d
VI	5	0.1 d	1.2 c	302 d
VII	10	0.2 c d	2.0 b c	437 b c d
VIII	6	0.2 c d	1.9 b c	470 b c d
IX	11	0.5 a	3.2 a	728 a
X	23	0.3 b c d	2.6 a b	638 a b
XI	22	0.3 b c d	2.3 b	591 a b c
XII	13	0.3 b c	2.2 b	547 a b c
REGION				
West Galilee	79	0.4 a	1.7 b c	439 b c
Bet She'an Valley	71	0.5 a	2.6 a	657 a
Coastal Plain	108	0.2 b	2.2 a b	509 b
Upper Galilee	15	0.1 b	1.4 c	377 c
PONDTYPE				
Pond	196	0.3 b	2.0 b	467 b
Reservoir	77	0.4 a	2.5 a	654 a

Table 11. Duncan's multicomparison test of means. Mean silver carp yield ($kg \cdot 0.1\ ha^{-1} \cdot day^{-1}$) growth ($g \cdot day^{-1}$) and harvest weight (g) in each culture period (SEASGRUP), fish combination (FISHES), stocking month, pond type and region. n = number of observations in each group. Total n = 593. In each column, means with the same letter are not significantly different at the 0.05 level. a>b> etc.

	n	SYIELDAY	SGROWTH	SWTO
SEASGRUP				
a (- sp)	20	0.4 a	4.6 c	679 d
c (- sp su)	210	0.3 a b	8.2 b	1,304 c
d (- sp su au-)	154	0.3 a	9.1 a b	2,122 a
e (su)	23	0.3 a	11.6 a	1,413 b c
f (su au-)	124	0.4 a	8.6 b	1,711 a b c
g (au-)	14	0.2 b	9.6 a b	1,763 a b
h (xx- xx)	48	0.4 a	9.6 a b	2,017 a
FISHES				
5 (c s)	126	0.4 a b	7.1 b	1,299 b
7 (c ms)	98	0.4 a	9.3 a	2,009 a
8 (c t s)	259	0.3 b	8.8 a	1,492 b
9 (c t ms)	110	0.3 b	9.5 a	2,121 a
STOCKING MONTH OF FIRST SPECIES				
I	30	0.4 b	9.4 a	2,243 a
II	58	0.3 b c	8.3 a	1,700 b c d
III	148	0.3 b c	8.2 a	1,409 d e
IV	39	0.3 b c	8.5 a	1,652 b c d
V	53	0.3 b c	7.6 a	1,196 e
VI	57	0.3 b c	9.8 a	1,536 c d e
VII	55	0.3 b c	8.5 a	1,722 b c d
VIII	39	0.5 a	10.3 a	1,901 a b c
IX	25	0.2 c	9.3 a	2,021 a b
X	31	0.4 b c	8.2 a	1,815 a b c d
XI	32	0.4 b	9.1 a	1,977 a b c
XII	26	0.3 b c	8.0 a	1,990 a b
REGION				
West Galilee	200	0.4 a	9.1 b	1,540 b
Bet She'an Valley	74	0.4 a	10.7 a	2,355 a
Coastal Plain	185	0.2 a	7.8 b	1,652 b
Upper Galilee	134	0.3 a	8.1 b	1,438 b
PONDTYPE				
Pond	491	0.3 b	8.2 b	1,499 b
Reservoir	102	0.4 a	10.9 a	2,398 a

rate was 8.6 g·day^{-1}. It was higher in the Bet She'an Valley, in reservoirs and when other fish besides common carp were present. Silver carp mean harvest weight was 1,650 g. It was higher when cultured for long periods or during the second half of the year, fish were stocked between August and January, mullet was present, in the Bet She'an Valley and in reservoirs.

Analyses on the Complete Array of Culture Systems

RELATIONSHIPS AMONG VARIABLES

Relationships among 27 variables were studied through factor analysis (Table 12a). The joint effect of SEASGRUP, FISHES, stocking month and region on the factors extracted were tested through the ANOVA model. The proportion of variance explained by the joint model and the main variable contributing to it are indicated in the text. The differences of the factors' value in each level of these main effects are presented in Table 12b.

Five factors account for 65% of the overall variance. The main source of variation (FACTOR1, 24%) is due to tilapia, which grew better, had higher yields and reached larger size when total stocking density was high, and when tilapia (being the main fish) was stocked at rather large size. This performance was achieved when little or no sorghum was used but was not dependent on other inputs to the pond (pellets and manure). Most of the variance of this factor (89%) is accounted for by the ANOVA model tested to which FISHES contributes the most. Tilapia (FACTOR1) had higher performance when grown with carp than when other species were present. Performance of tilapia was better when it was cultured in hot short seasons starting in summer than in long or short cold ones, and when stocking occurred in the hotter months. Its performance ranking among regions was: Bet She'an Valley > Western Galilee > Upper Galilee > Coastal Plain.

The second factor (FACTOR2, 18%) is associated with silver carp. This fish grew better, had higher yields and reached a larger size when cultured at high density (1,000 to 2,000·ha^{-1}, see Dis-

Table 12a. Factor analysis of the complete array of culture systems. Only large coefficients (>0.45) are included.

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
DEPTH	.	0.52	.	.	.
AREA	.	0.52	.	.	.
SORGFISH	-0.47	.	.	.	.
PELFISH	.	.	.	0.63	.
MANUDAY	.	.	.	.	.
TOTDEN	0.73	.	.	.	0.48
CWTI	.	.	0.52	.	.
CDEN	.	.	-0.56	.	.
PERCCDEN	0.87	.	.	.	.
CGRDAYS	.	0.64	.	.	.
CWTO	.	.	.	0.57	.
CYIELDAY	.	.	.	0.70	.
CGROWTH	.	.	.	0.64	.
TWTI	0.61	.	.	.	.
TDEN	0.79	.	.	.	.
PERCTDEN	0.92	.	.	.	.
TGRDAYS	0.79	.	.	.	.
TWTO	0.83	.	.	.	.
TYIELDAY	0.85	.	.	.	.
TGROWTH	0.82	.	.	.	.
SWTI	.	.	.	.	.
SDEN	.	0.50	.	.	0.54
PERCSDEN	.	.	0.54	.	0.43
SGRDAYS	.	0.88	.	.	.
SWTO	.	0.85	.	.	.
SYIELDAY	.	0.76	.	.	.
SGROWTH	.	0.74	.	.	.
Variance explained (%)	24	18	8	8	7

Table 12b. Analysis of the complete array of culture systems. Duncan's multicomparison test of means of FACTORs in each culture period (SEASGRUP), fish combination (FISHES), stocking month and region. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

	n	Duncan grouping				
		Factor 1 (Tilapia)	Factor 2 (Silver carp)	Factor 3 (Carp)	Factor 4 (Carp)	Factor 5 (Silver Carp)
SEASGRUP						
a (- sp)	100	f	d	a	c	a b
c (- spsu)	397	d	c	b	c	a
d (- spsuau-)	230	c	b	d	c	a
e (su)	61	b	d	a	a	a b
f (suau-)	215	a	c	a	b	a b
g (au-)	51	e	d	a	b	b c
h (xx- xx)	70	c	a	c	c	c
FISHES						
1 (c)	146	f	e	c	a	d
3 (c t)	270	a	e	c	b c	c
4 (c m)	65	c	d	d	c	b c
5 (c s)	126	e	b	a	a b	a
7 (c ms)	98	d	a	c	a b	b
8 (c t s)	259	b	c	b	a	b c
9 (c t ms)	110	b	a	c	a b	e
STOCKING MONTH OF FIRST SPECIES						
I	61	f g	a b	e f g	e	d e
II	107	f g	b c	g	c d e	b c d
III	237	e	b c	f g	d e	a b
IV	110	d	f	e f g	c d e	b c d
V	106	c	f	b c d e	b c	a
VI	97	a	e f	b c	a b	a b c
VII	113	b	f	a b	a	d e f
VIII	72	d	e f	a	a	d e
IX	51	g	d e	c d	a b	e f
X	54	e	a	e f g	b c d e	f
XI	62	f	b c	d e	a b	c d
XII	54	g	c d	e f	b c d	d e
REGION						
West Galilee	482	b	d	b	a	b
Bet She'an Valley	117	a	a	c	a	c
Coastal Plain	269	d	b	b	b	b
Upper Galilee	256	c	c	a	a	a

cussion) for long periods together with carp and in large deep ponds. This combination occurred mainly when mullets were also stocked (but not tilapia), in long SEASGRUPs and when fish were stocked in the cold months. In the Bet She'an Valley, where most of the deep reservoirs are located, silver carp performance was highest followed by Coastal Plain > Upper Galilee > Western Galilee. Most of the variance of this factor (82%) is accounted for by the ANOVA model tested to which FISHES contributes the most.

The first two factors are related to fishes which were stocked in only part of the analyzed ponds, hence their presence or absence in the observations has a strong weight in the respective factors. Common carp, being present in all the observations, accounts for a smaller part of the overall variability and more than one factor explains different aspects of its performance. Thus, the next two factors (16% together) are related to carp. The third factor (FACTOR3) shows stocking characteristics of carp. Large carp were stocked at low density when silver carp represented a large proportion of total density and vice versa. The ANOVA model tested accounts for 45% of the variance of this combination to which FISHES and

SEASGRUP contribute the most. The factor has higher values for very short SEASGRUPs than for long ones and when stocking took place in the hotter months. When only carp and silver carp were stocked (FISHES=5), this combination had the highest value followed by the combination which also includes tilapia (FISHES=8). The inclusion of mullets (FISHES=7 and 9) decreases the percent contribution of silver carp so that the factor has lower values.

Carp were stocked at larger sizes and at lower densities in the Upper Galilee and at smaller sizes and at higher densities in the Bet She'an Valley.

The fourth factor (FACTOR4) shows carp harvesting characteristics. The carp grew better, had higher yields and reached bigger harvesting size when high amounts of pellets were applied. The joint effect of the categorical variables tested accounts for only 14% of the variance of this combination to which all four variables contribute about the same. This combination had higher values for SEASGRUPs which develop in the second half of the year and also when fish were stocked between June and November. Higher carp performance and amounts of pellets given to the fish occurred when carp was grown in monoculture or in polyculture including silver carp. Better carp performance characterizes the culture operations in the Galilee and Bet She'an Valley than it does in the Coastal Plain.

The fifth factor (FACTOR5) describes silver carp absolute and relative contribution to total density. Only 18% of the variance of this combination is accounted for by the ANOVA model to which FISHES contribute the most.

In the preceding analysis tilapia and silver carp dominated in the first two factors since their absence in about 50% of the observations increased their impact on variability. From the Israeli farmer's point of view, silver carp, being a cheap fish, is more a sanitary fish than a production fish. Tilapia and common carp performances are the main concern of fish farmers. To reduce silver carp importance in the analysis and gain a better insight into common carp-tilapia relationships, factor analysis was run without the silver carp variables (Tables 13a and 13b). When silver carp variables were excluded, the first factor was still related to tilapia and a correlation of common carp stocking characteristics with pond size appeared in the second factor. Small carp were stocked at higher density for long culture seasons in large and deep ponds, and vice versa. A large part of the variance of this combination (66%) is

Table 13a. Factor analysis of the complete array of culture systems, without SILVER CARP variables. Only large coefficients are included.

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
DEPTH	.	0.67	.	.	.
AREA	.	0.64	.	.	.
SORGFISH	-0.47	.	.	.	.
PELFISH	.	.	0.65	.	.
MANUDAY	.	.	.	.	.
TOTDEN	0.73	.	.	-0.52	.
CWTI	.	-0.67	.	.	.
CDEN	.	0.45	.	-0.58	.
PERCCDEN	-0.88	.	.	.	.
CGRDAYS	.	0.85	.	.	.
CWTO	.	.	0.59	.	0.55
CYIELDAY	.	.	0.81	.	.
CGROWTH	.	.	0.59	.	0.52
TWTI	0.62	.	.	.	-0.40
TDEN	0.80	.	.	.	.
PERCTDEN	0.93	.	.	.	.
TGRDAYS	0.78	.	.	.	.
TWTO	0.82	.	.	.	.
TYIELDAY	0.86	.	.	.	.
TGROWTH	0.82	.	.	.	.
Variance explained	32%	14%	11%	9%	8%

Table 13b. Analysis of the complete array of culture systems. Duncan's multicomparison test of means of FACTORs (calculated without SILVER CARP variables) in each culture period (SEASGRUP), fish combination (FISHES), stocking month and region. Groups with the same letter are not significantly different. n = number of observations. a>b> etc.

		Duncan grouping				
	n	Factor 1 (Tilapia)	Factor 2 (Carp)	Factor 3 (Carp)	Factor 4 (Carp)	Factor 5 (Carp-Tilapia)
SEASGRUP						
a (- sp)	100	f	d	d	b	a
c (- spsu)	397	d	b	c d	d	b c
d (- spsuau-)	230	c d	a	b c	c	a
e (su)	61	b	d	a	b c	b c
f (suau-)	215	a	c	b c	b c	c
g (au-)	51	e	d	b	b c	b
h (xx- xx)	70	c	a	c d	a	a
FISHES						
1 (c)	146	f	e	a	b	b c
3 (c t)	270	a	d	a	b	c
4 (c m)	65	c	b	a	c	a
5 (c s)	126	e	c	a	b c	b c
7 (c ms)	98	d	b	a	b	b c
8 (c t s)	259	b	c	a	b	a b
9 (c t ms)	110	b	a	a	a	c
STOCKING MONTH OF FIRST SPECIES						
I	61	f g	a	d	b	a b
II	107	f g	a	b c d	c d	c
III	237	e	a b	c d	d	c
IV	110	d	c d	b c d	c d	c
V	106	c	e f	a b c	d	b c
VI	97	a	f g	a b c	c d	c
VII	113	b	g	a	b	d
VIII	72	d	g	a b c	b c	d
IX	51	g	d e	a b c	a b	c
X	54	e	a	a b c d	a	b c
XI	62	f	b	a b	a b	a b
XII	54	g	b c	a b c d	a b	a
REGION						
West Galilee	482	b	c	a	c	d
Bet She'an Valley	117	a	a	a	a	b
Coastal Plain	269	d	b	b	b	c
Upper Galilee	256	c	d	a	b c	a

accounted for by the ANOVA model to which SEASGRUP contributes the most. The factor has higher values for long SEASGRUPs than for short ones and when stocking took place in colder months. When all four species were stocked (FISHES=9) this combination of conditions had the highest value followed by the other FISHES which also includes mullet then followed by the remaining ones which include silver carp and finally by carp with tilapia, and carp alone. Thus, monoculture of carp is generally done with large fish at low densities during short periods, in rather shallow and small ponds. These carp input characteristics had higher values in the Bet She'an Valley where deep reservoirs are concentrated followed by the Coastal Plain, by the Western Galilee and finally by the Upper Galilee.

DAILY YIELDS

Canonical correlation analysis (n = 1,124) of 14 explanatory variables on yields of common carp, tilapia and silver carp are presented in Table 14a, together with multicomparison tests of the

Table 14a. Analysis of the complete array of culture systems. Results of canonical correlation analysis of yield data (YCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structure analysis). n = 1,124.

	YCC1	YCC2	YCC3
Canonical correlation coefficient	0.87	0.66	0.59
Variance accounted for (%)	70	18	12
Standardized canonical coefficients for the explanatory variables			
	YE1	YE2	YE3
DEPTH	.	.	.
AREA	.	.	.
SORGFISH	.	.	.
PELFISH	.	.	0.56
MANUDAY	.	.	.
TOTDEN	0.37	.	0.87
CWTI	.	.	.
PERCCDEN	(-)	-0.44	1.16
CGRDAYS	.	-0.47	.
TWTI	(+)	.	.
PERCTDEN	0.94	-0.44	.
TGRDAYS	(+)	.	.
SWTI	.	(+)	.
SGRDAYS	.	0.91	.
Standardized canonical coefficients for the response variables			
	YR1	YR2	YR3
CYIELDAY	.	.	1.00
TYIELDAY	0.99	.	.
SYIELDAY	.	1.00	.

Table 14b. Analysis of the complete array of culture systems. Duncan's multicomparison test of means of yield response variables according to culture period (SEASGRUP), fish combination (FISHES), stocking month and region. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping		
	n	YR1 (Tilapia)	YR2 (Silver carp)	YR3 (Common carp)
SEASGRUP				
a (- sp)	100	e	c	b c
c (- spsu)	397	c	b	a b c
d (- spsuau-)	230	d	a	a b c
e (su)	61	a	b	a
f (suau-)	215	b	a	b c
g (au-)	51	e	c	a b
h (xx- xx)	70	d	a	c
FISHES				
1 (c)	146	d	c	a
3 (c t)	270	a	b	b c
4 (c m)	65	d	c	b c
5 (c s)	126	d	a	b c
7 (c ms)	98	d	a	b c
8 (c t s)	259	b	a	b
9 (c t ms)	110	c	a	c
STOCKING MONTH OF FIRST SPECIES				
I	61	d e	c	c
II	107	d e	b c d	a b c
III	237	c d	b	a b c
IV	110	b	d e	a b c
V	106	a	b c	a b c
VI	97	a	a b	a b c
VII	113	a	b c	b c
VIII	72	c	a	b c
IX	51	e	e	a b c
X	54	d e	b	a b c
XI	62	c d e	b c	a
XII	54	d e	c d e	a b
REGION				
West Galilee	482	b	c	b
Bet She'an Valley	117	a	a	a b
Coastal Plain	269	c	b	c
Upper Galilee	256	d	b c	a

response variables in Table 14b. The first yield canonical correlation (YCC1) explains 70% of the variance in the dataset. It shows tilapia daily yield (YR1) correlated with its proportion and with total density (YE1). The canonical structure shows that tilapia initial weight and culture days are also correlated with both the explanatory and the response canonical variables. Their effect on tilapia yield is not seen in the canonical coefficients since they are confounded with tilapia density. A large part of the variance of this combination (61%) is accounted for by the ANOVA model to which FISHES contribute the most. Tilapia density was lower when more fish species were stocked (Table 5). Accordingly, its daily yields were the highest when stocked only with common carp, were lower when silver carp was added and were still lower when also mullets were present. Of course, the YR1 value was the lowest when no tilapia were stocked. Higher YR1 occurred when the culture season was short and included the summer period followed by long seasons and finally by very short and colder seasons. Also, tilapia yields were higher when fish were stocked in the hot months (May to July). The highest values of this factor occurred in the Bet She'an Valley followed by Western Galilee > Upper Galilee > Coastal Plain.

The second yield canonical correlation (YCC2) accounts for another 18% of the variance. It shows silver carp daily yield (YR2) correlated with its culture days and negatively correlated with carp and tilapia per cent contribution to total density (YE2), which means positive correlation with its own density. The canonical structure also shows that silver carp initial weight correlated with the explanatory and response variables. About 52% of the variance of the response variable is accounted for by the ANOVA model to which FISHES contribute the most. Silver carp daily yield (YR2) was not affected by the other species present in the pond but only by presence or absence of silver carp. Higher values occurred in the long seasons and the opposite occurred in very short cold seasons. Higher silver carp yields occur in the Bet She'an Valley and lower in the Galilee.

The third yield canonical correlation (YCC3) accounts for the remaining 12% of the variance. It shows that common carp daily yields (YR3) positively correlated with its proportion in total density, with total density and in the third place with amount of feed pellets (YE3). The ANOVA model only accounts for 8% of the YR3 variability to which FISHES and SEASGRUP contribute the most. Daily yields (YR3) were higher in monoculture than in polyculture. This points to negative effects of other species on the common carp, since mean carp density in each fish combination was similar but total density was not (Table 5).

The yield analysis shows that yield of each species depends mainly on its own density and in the second place on its stocking weight and length of culture period. The common carp is the only species whose yield was strongly affected by the amount of feed pellets and this in the second place after its own parameters. Fig. 3A represents these relationships graphically.

DAILY GROWTH

Canonical correlation analysis (n = 1,124) of the same 14 explanatory variables on daily growth rate of common carp, tilapia and silver carp is presented in Table 15a and the multicomparison tests of the response variables in Table 15b. The first growth canonical correlation (GCC1) accounts for 71% of the variance. It shows that tilapia daily growth (GR1) correlated with the length of its culture period, stocking weight and its per cent contribution to the system (GE1). The canonical structure shows that total density and carp per cent density are also correlated with the explanatory and response variables. The ANOVA model accounts for 75% of the GR1 variability to which FISHES contributes the most. Tilapia growth rate (GR1) was lower when silver carp was present but not when silver carp and mullets were stocked. Tilapia growth rate was the highest when the culture season was during summer followed by the short seasons, the long seasons and finally the very short and cold seasons. Accordingly, it grew better when stocked in hot months. The best growth occurred in the Bet She'an Valley followed by Western Galilee > Upper Galilee > Coastal Plain.

The second growth canonical correlation (GCC2) accounts for 23% of the variance. It shows that silver carp daily growth (GR2) also correlated with the length of its culture period and stocking weight. Carp growth days and pond depth are also correlated with this component. The ANOVA model accounts for 55% of the GR2 variability to which FISHES contributes the most. Silver carp growth was better in the presence of mullets but was not affected by the presence or absence of tilapia. It had a higher growth rate in long seasons than in short and very short ones. It grew

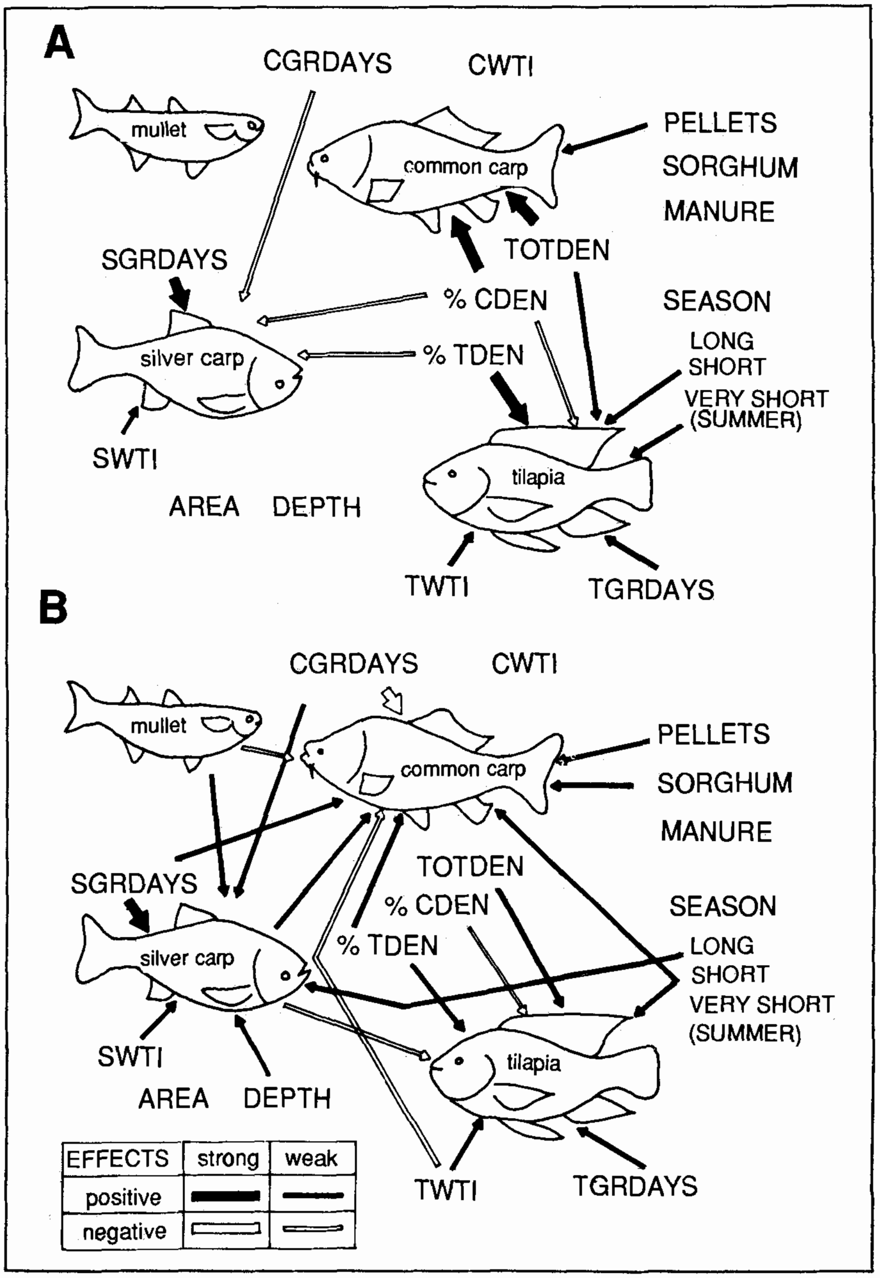


Fig. 3. Relationships affecting (A) fish YIELDs, and (B) fish GROWTH, as obtained from results of the complete array of culture systems analyzed through canonical correlations. Arrows indicate major effects of specific variables on each fish species. TOTDEN is total density. %CDEN and %TDEN are percentage of common carp and tilapia, respectively, in the total density. CWTI, TWTI and SWTI are initial weight of carp, tilapia and silver carp, respectively. CGRDAYS, TGRDAYS and SGRDAYS are days of effective growth of each species.

better in the Bet She'an Valley and Coastal Plain than in the Galilee.

The third growth canonical correlation (GCC3) only explains the remaining 6% of the variance. It shows common carp daily growth (GR3) negatively correlated with the length of culture period and to a lesser extent also with tilapia stocking weight,and positively correlated with the amount of feed pellets and sorghum, silver carp culture period and per cent of tilapia. The ANOVA model accounts for 23% of the GR3 variability to which SEASGRUP and region contribute the most. Common carp daily growth (GR3) was higher in the very short seasons than in the others and when stocked in the last months of the year than in the earlier ones. Common carp growth was the highest in monoculture or in polyculture with silver carp (with or without tilapia). The inclusion of mullets (with or without tilapia) reduced the GR3 but the exclusion of silver carp reduced it even more. In the polyculture systems, the presence of tilapia did not affect carp growth.

The growth analysis shows that growth of each species depends mainly on its own parameters. The common carp is the only fish species whose growth was strongly affected by the amount of nutritional inputs. Growth rates are much more affected than yields by the presence of other species in the ponds. Carp grows better in monoculture than in polyculture. Common carp and tilapia do not affect each other but are affected in opposite ways by silver carp and mullets: mullets have a negative effect on common carp and a positive one on tilapia, while silver carp have a positive effect on common carp and negative on tilapia. These interactions are summarized graphically in Fig. 3B.

Analyses of the Carp-Tilapia Culture System

In the analyses on the whole dataset, tilapia accounted for most of the explained variance and common carp for the least due in part to tilapia being present only in about half of the observations while common carp was present in all of them. Thus, the following analyses were run only

Table 15a. Analysis of the complete array of culture systems. Results of canonical correlation analysis of growth data (GCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated to the canonical variates (canonical structure analysis). n = 1,124.

	GCC1	GCC2	GCC3
Canonical correlation coefficient	0.85	0.68	0.42
Variance accounted for (%)	71	23	6
Standardized canonical coefficients for the explanatory variables			
	GE1	GE2	GE3
DEPTH	.	(+)	.
AREA	.	.	.
SORGFISH	.	.	0.48
PELFISH	.	.	0.66
MANUDAY	.	.	.
TOTDEN	(+)	.	.
CWTI	.	.	.
PERCCDEN	(-)		.
CGRDAYS	.	(+)	-0.93
TWTI	0.37	.	-0.37
PERCTDEN	0.36	.	0.39
TGRDAYS	0.43	.	.
SWTI	.	0.43	.
SGRDAYS	.	0.76	0.42
Standardized canonical coefficients for the response variables			
	GR1	GR2	GR3
CGROWTH	.	.	0.98
TGROWTH	1.00	.	.
SGROWTH	.	0.96	0.31

Table 15b. Analysis of the complete array of culture systems. Duncan's multicomparison test of means of growth response variables according to culture period (SEASGRUP), fish combination (FISHES), stocking month and region. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping		
	n	GR1 (Tilapia)	GR2 (Silver carp)	GR3 (Common carp)
SEASGRUP				
a (- sp)	100	e	d	a
c (- sp su)	397	c	b	b
d (- sp su au-)	230	d	a	b
e (su)	61	a	c	a
f (su au-)	215	b	b	b
g (au-)	51	e	c	a
h (xx- xx)	70	c	a	b
FISHES				
1 (c)	146	c	d	a
3 (c t)	270	a	d	b c
4 (c m)	65	c	c	c
5 (c s)	126	c	b	a
7 (c ms)	98	c	a	b
8 (c t s)	259	b	b	a
9 (c t ms)	110	a b	a	b
STOCKING MONTH OF FIRST SPECIES				
I	61	f	a b	d e
II	107	e f	a b	e
III	237	d e	a	e
IV	110	b c	c	d e
V	106	b	b c	c d e
VI	97	a	a	c d
VII	113	a	a b c	d e
VIII	72	b c	a	b c d
IX	51	f	a b c	a b c
X	54	c d	a b c	b c d
XI	62	d e	a b c	a b
XII	54	e f	c	a
REGION				
West Galilee	482	b	b	d
Bet She'an Valley	117	a	a	b
Coastal Plain	269	d	a	c
Upper Galilee	256	c	b	a

for those observations (n = 639) where both main fish species were present (FISHES = 3, 8 and 9).

RELATIONSHIPS AMONG VARIABLES

Table 16a shows the relationships among variables in this system studied through factor analysis and Table 16b, the multicomparison tests of the factors. Four factors account for 64% of the variance of which the first two (42%) are mainly related to tilapia and the others to common carp. FACTOR1 (25%) defines the common carp - tilapia density combination which affects tilapia yield. The higher the absolute and relative contribution of tilapia to the total density, the higher the tilapia yields (negative coefficients of the factor). This is also associated with low carp percentage contribution, short carp and tilapia culture periods, small tilapia at harvest and growth in shallow small ponds (positive coefficients of the factor). The ANOVA model tested accounts for only 34% of this factor variability to which FISHES contributes the most. The carp-tilapia combinations which lead to high tilapia yields (low FACTOR1 values) occurred mainly when only the two species occurred (the more species present, the higher the factor value), in short and very short SEASGRUPs, when fish were stocked between March and August and in the West and Upper Galilee regions.

The second factor (FACTOR2) (17% of variance) defines conditions for tilapia growth. Tilapia grew better when large amounts of pellets were available, carp and tilapia were stocked at rather large sizes, and these two fish species were cultured for short periods. The ANOVA model tested accounts for 51% of this factor variability to which SEASGRUP contributes the most. The combination of conditions which led to good tilapia growth occurred mainly in very short SEASGRUPs followed by short SEASGRUPs and the smallest fish with lower tilapia growth rates were stocked to be cultured in the long SEASGRUPs. This factor also shows significant differences according to stocking month so that higher tilapia growth rate is related to stocking in summer followed by spring and autumn and finally winter (first species in the polyculture stocked in winter but tilapia added in spring). The presence of more fish species occurred when carp and tilapia were stocked at rather small sizes so that FACTOR2 has lower values the more fish species are present in the ponds. Finally, this factor has higher values in the Western Galilee than in the other regions.

The third factor (FACTOR3) (12% of variance) defines carp output characteristics (weight at harvest and growth rate) correlated to tilapia harvest size. Only 37% of the variance of this combination

Table 16a. Factor analysis of the carp-tilapia culture system. Only large coefficients are included. n = 639.

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
DEPTH	0.47	.	.	.	.
AREA	0.45	.	.	.	.
SORGFISH	.	.	.	.	.
PELFISH	.	0.57	.	.	.
MANUDAY	.	.	.	.	.
TOTDEN	-0.67	.	.	.	.
CWTI	.	0.54	.	.	-0.44
CDEN	.	.	.	0.68	.
PERCCDEN	0.85	.	.	.	.
CGRDAYS	0.52	-0.65	.	.	.
CWTO	.	.	0.77	.	-0.50
CYIELDAY	.	.	.	0.80	.
CGROWTH	.	.	0.60	.	-0.45
TWTI	.	0.75	.	.	.
TDEN	-0.83	.	.	.	.
PERCTDEN	-0.91	.	.	.	.
TGRDAYS	0.45	-0.55	.	.	.
TWTO	0.46	0.45	0.47	.	.
TYIELDAY	-0.64	.	.	.	0.49
TGROWTH	.	0.61	.	.	0.41
Variance explained	25%	17%	12%	10%	8%

Table 16b. Carp-tilapia culture system. Duncan's multicomparison test of means of FACTORs according to culture period (SEASGRUP), fish combination (FISHES), stocking month and region. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping				
	n	Factor 1 (T. yield)	Factor 2 (T. growth)	Factor 3 (C. growth)	Factor 4 (C. yield)	Factor 5 (C. vs T.)
SEASGRUP						
a (- sp)	100	b	b c	d	a	a
c (- sp su)	397	b c	c	d	b	a
d (- sp su au-)	230	a	d	a	b	a
e (su)	61	b c	a	a b c	b	a
f (su au-)	215	b c	b	c	b	a
g (au-)	51	c	a	b c	b	b
h (xx- xx)	70	a	d	a b	c	a
FISHES						
3 (c t)	270	c	a	b	b	a
8 (c t s)	259	b	b	b	a	b
9 (c t ms)	110	a	c	a	c	a
STOCKING MONTH OF FIRST SPECIES						
I	61	a	f	b c	b c	a
II	107	a	c d	e	a	a b
III	237	b	c d e	e	a b	a b
IV	110	b	b c	d e	a b	a b c
V	106	b	b c	c d	a	a b c
VI	97	b	b	c d	a b	b c
VII	113	b	a	c d	b c	a b c
VIII	72	b	a	d e	a b	c
IX	51	a	c d e	b c	c	c
X	54	a	d e	b c	c	b c
XI	62	a	c d e	b	a b	a b c
XII	54	a	e f	a	a	a
REGION						
West Galilee	482	b	a	c	b	a
Bet She'an Valley	117	a	b	a	d	a
Coastal Plain	269	a	b	d	c	b
Upper Galilee	256	b	b	b	a	b

is accounted for by the ANOVA model tested to which SEASGRUP and region contribute the most. Carp performance was better in long SEASGRUPs followed by the ones covering only the second half of the year and finally the spring-summer period. It was also better when fish were stocked during the autumn or winter months followed by summer and then by spring months. Polyculture of the four species enhanced carp growth. Carp growth was better in the internal valleys of the country (Bet She'an Valley better than Upper Galilee) than in the coastal regions (Western Galilee better than in the Coastal Plain).

The fourth factor (FACTOR4) (10% of the variance) is associated with carp yield which depends on carp density. The ANOVA model tested accounts for only 27% of this factor variability to which region contributes the most. The highest carp stocking densities and yields occurred in the Upper Galilee and the lowest in the Bet She'an Valley. Carp density and yield were the highest when silver carp was added to the main two fishes (FISHES=8) and the lowest when all four fishes were stocked (FISHES=9).

Each of the four factors identifies the conditions related to carp or tilapia growth or yields in polyculture with and without silver carp and mullets. Input characteristics of the ponds affect tilapia performance more than carp performance. The categorical variables analyzed affect the growth and yields of the two main species studied in a different way.

Table 17. Carp-tilapia culture system. Analyses with different tilapia densities (n·ha^{-1}). Results of canonical correlation analysis of yield data (YCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structural analysis).

	FISHES 3,8,9 all data		FISHES 3,8,9 2000<TDEN<6000		FISHES 3,8 2000<TDEN<15000	
Number of obs.	639		232		436	
	YCC1	YCC2	YCC1	YCC2	YCC1	YCC2
Canonical correlation coefficient	0.76	0.58	0.72	0.60	0.73	0.71
Variance accounted for (%)	72	28	68	32	52	48
Standardized canonical coefficients for the explanatory variables						
	YE1	YE2	YE1	YE2	YE1	YE2
DEPTH	.	.	.	.	.	.
AREA	.	.	.	.	.	.
SORGFISH	.	.	.	.	.	.
PELFISH	0.30	0.67	.	0.58	0.37	0.56
MANUDAY	.	.	.	.	0.31	.
TOTDEN	0.51	0.74	0.40	0.61	0.48	0.70
CWTI	-0.32	.	-0.40	-0.37	-0.30	.
CGRDAYS	.	.	.	-0.52	.	.
TWTI	.	.	.	.	.	.
PERCTDEN	0.63	-0.76	1.06	.	0.65	-0.67
TGRDAYS	.	-0.43	-0.35	.	.	-0.33
Standardized canonical coefficients for the response variables						
	YR1	YR2	YR1	YR2	YR1	YR2
CYIELDAY	.	1.00	.	0.97	.	1.00
TYIELDAY	1.00	.	0.97	.	1.00	.

DAILY YIELDS

Results of canonical correlations on yield data are presented in Table 17. The ANOVA models show no significant effect of the categorical variables tested, thus the corresponding table is not presented. In this system, tilapia accounts for most (72%) of the daily yield variability. Tilapia daily yield depends mainly on its own proportion in the polyculture combination, then on total density, and in the third place on the amount of feed pellets and (negatively) on carp stocking weight. The presence of species other than carp and tilapia negatively affects tilapia yields since they reduce the tilapia proportion in total density.

Most observations of tilapia density, when stocked with carp alone or with carp and silver carp, fall between 2,000 and 15,000 tilapia·ha^{-1} while when mullets are also present tilapia density was lower (2,000 to 6,000 fish·ha^{-1}). If the canonical correlation analyses are run without the very high densities of this fish (in which common carp was present at low densities), the first correlations still show the same pattern, either for the 2,000 to 6,000 range or for the 2,000 to 15,000 one (Table 17).

Carp daily yield accounts for the remaining variance in yields. It mainly depends on the tilapia proportion in the polyculture (complementary to its own), total density and amount of feed pellets, the latter with a stronger influence than in the tilapia canonical variable.

This analysis shows that the yields of the two major species are mainly affected by total density and the proportion of each species in it and by the amount of pellets given to the fish (mainly for carp). These relationships are shown graphically in Fig. 4A.

DAILY GROWTH

Results of canonical correlations on growth data are presented in Table 18. The ANOVA models show no significant effect of the categorical

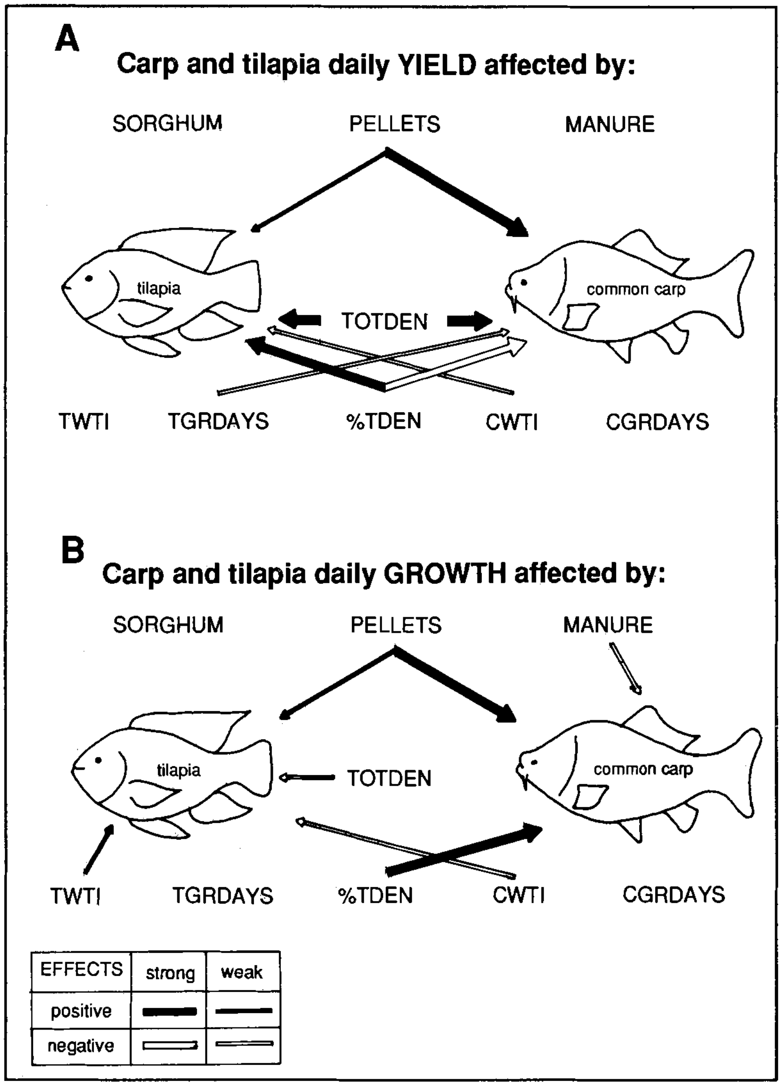


Fig. 4. Relationships affecting (A) fish YIELDs, and (B) fish growth, as obtained from results of the carp-tilapia culture systems analyzed through canonical correlations. Arrows indicate major effects of specific variables on each fish species. TOTDEN is total density. %CDEN and %TDEN are percentage of common carp and tilapia, respectively, in the total density. CWTI and TWTI are initial weight of carp and tilapia, respectively. CGRDAYS and TGRDAYS are days of effective growth of each species.

variables tested, thus the corresponding table is not presented. Tilapia accounts for most (72%) of the variability of fish growth rate. Tilapia daily growth depends on the amounts of feed pellets applied, and (negatively) on carp stocking weight and total density. Its own stocking weight is also correlated with this combination of variables according to the canonical structure analysis.

Carp daily growth is strongly correlated with the amount of feed pellets applied and the proportion of tilapia in the system. The higher the tilapia proportion the lower the carp proportion (density) and hence the higher the carp growth. A negative effect of daily amount of manure on carp growth is also indicated by the canonical correlation.

The analysis of growth in observations with 2,000 to 15,000 tilapia·ha^{-1} is similar to the one discussed above and show no silver carp effect either on carp or tilapia growth. On the other hand, the analysis in observations with 2,000 to 6,000 tilapia·ha^{-1}, which includes ponds with silver carp and mullets, did not isolate the effect of the explanatory variables on each fish species growth rate. The first canonical correlation shows good growth rate of both, tilapia and carp, when small carp were stocked, under high feeding rate, large proportion of tilapia which grew in the ponds for relatively short periods. The second component shows good growth of one species correlated with bad growth of the other one. High tilapia and low carp growth are related to large tilapia and small carp at stocking together with high manuring rate and low feeding rate with pellets, and vice versa. This relationship is not affected by polyculture combination.

The above discussed relationships are presented graphically in Fig. 4.

Analyses by Geographical Regions

Factor analysis by geographical regions revealed strong similarity between the data of the Western Galilee and the Coastal Plain and those of the complete array (except for deviations due to specific effects of the other regions in the latter).

The other two regions - Upper Galilee and Bet She'an Valley - differ remarkably in culture conditions. The Upper Galilee is characterized by lower mean water temperatures during most months of the year while in the Bet She'an Valley deep ponds and reservoirs are more common than in the other regions. Indeed, the results of the analyses on data from these two regions differ from those of the other two and of the complete array. Thus, a detailed presentation of the analyses within each of these two regions follows.

BET SHE'AN VALLEY

In this region 61% of the ponds analyzed included mullets, thence the variables related to this species were explicitly included. The ANOVA

Table 18. Carp-tilapia culture system. Analyses with different tilapia densities (n·ha^{-1}). Results of canonical correlation analysis of growth data (GCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structural analysis).

	FISHES 3,8,9 all data		FISHES 3,8,9 2000<TDEN<6000		FISHES 3,8 2000<TDEN<15000	
Number of obs.	639		232		436	
	GCC1	GCC2	GCC1	GCC2	GCC1	GCC2
Canonical correlation coefficient	0.61	0.43	0.55	0.37	0.61	0.44
Variance accounted for (%)	72	28	74	26	73	27
Standardized canonical coefficients for the explanatory variables						
	GE1	GE2	GE1	GE2	GE1	GE2
DEPTH	.	.	0.37	-0.34	.	.
AREA	.	.	.	.	.	.
SORGFISH	.	.	.	.	.	.
PELFISH	0.53	0.72	0.55	-0.43	0.67	0.67
MANUDAY	.	-0.32	.	0.58	.	-0.45
TOTDEN	-0.42	.	.	0.39	-0.32	.
CWTI	-0.43	.	0.75	-0.52	.	.
CGRDAYS	.	.	(-)	0.41	.	.
TWTI	(+)	.	0.32	0.65	.	.
PERCTDEN	.	0.88	0.40	.	.	0.83
TGRDAYS	.	.	0.59	.	-0.36	.
Standardized canonical coefficients for the response variables						
	GR1	GR2	GR1	GR2	GR1	GR2
CGROWTH	.	1.00	0.49	-0.92	.	0.97
TGROWTH	0.98	.	0.74	0.73	0.91	-0.44

models of this region include as categorical variables SEASGRUP, FISHES, stocking month and pond type. The amount of data available from this region is rather low (117 observations), which led to a small number of observations in each level of the main effects tested; thus, the multicomparison results are considered here only as a general indication.

Relationships Among Variables

Table 19 presents factor analysis and multicomparisons of Bet She'an Valley data. FACTOR1 shows that 28% of the data variability is related to pond size and silver carp and mullet performances. It reflects the main differences between fish culture in reservoirs and in ponds. Fish culture in the Bet She'an Valley is carried out in deep reservoirs and in shallow ponds. About half of the observations in our dataset belong to each category. Reservoirs are filled with water and stocked with small common carp, silver carp and mullet during the winter, and tilapia is added in spring (Hepher 1985; Sarig 1988). During the summer the water of many reservoirs is used for irrigation so that the water level decreases and fish are harvested toward the end of the summer or in autumn. The FACTOR1 shows that in the reservoirs, as opposed to ponds, higher manuring rates are practised; silver carp is stocked at rather higher densities and attain larger size, growth rate and yield; mullets are stocked at large sizes and attain a higher weight, growth rate and also yield; tilapia is stocked at higher density and its yield is somewhat lower. In the shallow ponds, polyculture systems with fewer species and shorter seasons are practised, total density is higher, carp is stocked at a rather large size, silver carp and mullet performances are poorer, and tilapia daily yield is somewhat higher. Most of the variability (92%) of Bet She'an FACTOR1 is accounted for by the four categorical variables of the ANOVA model in which FISHES is the variable that contributes the most.

Table 19a. Factor analysis of Bet She'an data. Only large coefficients (>0.45) are included. n = 117.

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
DEPTH	0.69	.	.	.	.
AREA	0.53	.	.	.	.
SORGFISH	.	-0.48	.	.	.
PELFISH	.	.	.	.	.
MANUDAY	0.49	.	.	.	.
TOTDEN	-0.49	0.50	.	0.49	.
CWTI	-0.62	.	.	.	.
CDEN	.	.	.	0.51	0.57
PERCCDEN	.	-0.74	-0.53	.	.
CGRDAYS	0.79	.	.	.	.
CWTO	.	.	.	.	.
CYIELDAY	.	.	.	.	0.67
CGROWTH	.	.	.	.	.
TWTI	.	0.59	.	.	.
TDEN	0.50	0.68	.	.	.
PERCTDEN	.	0.90	.	.	.
TGRDAYS	0.49	0.62	.	.	.
TWTO	.	0.74	.	.	.
TYIELDAY	-0.47	0.78	.	.	.
TGROWTH	.	0.76	.	.	.
SWTI	.	.	.	.	.
SDEN	0.55	.	.	.	.
PERCSDEN	0.53	.	-0.48	.	.
SGRDAYS	0.85	.	.	.	.
SWTO	0.81	.	.	.	.
SYIELDAY	0.62	.	.	.	.
SGROWTH	0.69	.	.	.	.
MWTI	0.58	.	.	.	.
MDEN	.	-0.46	0.60	.	.
MGRDAYS	0.84	.	.	.	.
MWTO	0.85	.	.	.	.
MYIELDAY	0.49	.	0.58	.	.
MGROWTH	0.80	.	.	.	.
Variance explained	28%	18%	8%	7%	6%

The second factor (FACTOR2) variability source in this geographical area is due to tilapia performance. These fish attain a larger size and better growth rate and yield when stocked at rather large sizes and at high density, when they occupy a large proportion in the polyculture population, when total density was high, and when little or no sorghum feed was supplied. The presence of silver carp did not affect tilapia performance, but mullets did, negatively. Most of the variability (88%) of FACTOR2 is accounted for by the four categorical variables of the ANOVA model, in which FISHES is the variable that contributes the most. No differences between pond types occurred in this factor since the part of tilapia performance related to it was already accounted for by the first factor.

The third factor (FACTOR3) shows that mullet daily yield was correlated with its stocking weight and proportion in the polyculture (shown through the negative correlation with carp and silver carp proportion in polyculture). Only 57% of FACTOR3 variability is accounted for by the four categorical variables of the ANOVA model in which FISHES is the variable that contributes the most. No differences between pond type levels occurred since mullet performance related to it is accounted for by the first factor.

The next two factors are related to common carp. FACTOR4 shows correlations with total density and carp density, not related to FISHES, SEASGRUP or stocking month. FACTOR5 is the correlation between carp density and its daily yield.

Table 19b. Bet She'an data analysis. Duncan's multicomparison test of means of FACTORs according to culture period (SEASGRUP), fish combination (FISHES), stocking month and pond type. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping				
	n	Factor 1 (Silver carp + mullet)	Factor 2 (Tilapia)	Factor 3 (Mullet)	Factor 4 (Carp)	Factor 5 (Carp)
SEASGRUP						
a (- sp)	1	c	c	c	a	a
c (- spsu)	38	b	a b	a b	a	a
d (- spsuau-)	31	a	a	a	a	a
e (su)	2	c	c	c	a	a
f (suau-)	10	c	a b	a b	a	a
g (au-)	2	c	a b	a b	a	a
h (xx- xx)	33	a	a b	a b	a	a
FISHES						
1 (c)	9	d	e	c	b	a
3 (c t)	24	d	a	b	b	b
4 (c m)	10	c	d e	a	b	b
5 (c s)	1	b	c	d	a	b
7 (c ms)	6	a	c d	b	b	b
8 (c t s)	12	b c	a b	c	b	a b
9 (c t ms)	55	a	b	b	b	a b
STOCKING MONTH OF FIRST SPECIES						
I	11	b	b	a b	a	a b c
II	3	b	e	a b	a	c
III	14	c	c d	a	a	b c
IV	13	d	b c	c	a	a b
V	7	d	d e	a b	a	a b c
VI	6	d e	c d	a b c	a	a b c
VII	5	e	a	b c	a	b c
VIII	5	b c	c d	a b c	a	a b
IX	11	a	c d	a b c	a	b c
X	19	b	c d	b c	a	a b c
XI	15	b	d e	c	a	a b c
XII	8	b	c d	a b	a	a
PONDTYPE						
Pond	64	b	a	a	b	a
Reservoir	53	a	a	a	a	a

Daily Yields

Table 20 shows results of canonical correlations on Bet She'an Valley yield data. Most of the variability (62%) is due to tilapia and mullets whose yields are negatively correlated. Tilapia had higher daily yields when total density and tilapia proportion in it were high, little or no sorghum feed was supplied, and mullet daily yield was low. This occurred when tilapia was cultured without mullets, in the half-year SEASGRUPs (c and f), when stocking was in the hot months, and in ponds more than in reservoirs. The opposite situation favors mullet daily yields. The ANOVA model with the four categorical variables accounts for 79% of the YR1 variability, in which FISHES is the variable that contributes the most.

The second source of yield variability (19%) is also related to mullet and tilapia yields, which were positively correlated here. Good yields of both species (mainly mullet) occurred when the proportions of common carp and tilapia in the polyculture were low (which allows for a larger mullet share) and mullets were cultured for long periods. A large part of the variability (60%) of

Table 20a. Bet She'an analysis. Results of canonical correlation analysis of yield data (YCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structure analysis). n = 117.

	YCC1	YCC2	YCC3	YCC4
Canonical correlation coefficient	0.94	0.83	0.75	0.69
Variance accounted for (%)	62	19	11	8
Standardized canonical coefficients for the explanatory variables				
	YE1	YE2	YE3	YE4
DEPTH	.	.	.	0.44
AREA	.	.	.	-0.48
SORGFISH	(-)	.	.	.
PELFISH	.	.	.	.
MANUDAY	.	.	.	.
TOTDEN	0.35	.	0.98	0.60
CWTI	.	.	.	.
PERCCDEN	0.37	1.12	.	1.07
CGRDAYS	.	.	.	.
TWTI	.	.	.	.
PERCTDEN	0.84	0.68	-1.21	0.51
TGRDAYS	.	.	.	0.35
SWTI	.	.	.	.
PERCSDEN	.	.	-0.37	0.65
SGRDAYS	.	.	.	(+)
MWTI	.	.	.	.
MGRDAYS	(-)	-0.38	.	0.65
Standardized canonical coefficients for the response variables				
	YR1	YR2	YR3	YR4
CYIELDAY	.	.	0.92	0.31
TYIELDAY	0.81	-0.80	.	.
SYIELDAY	.	.	.	1.00
MYIELDAY	(-)	-1.03	0.38	.

this YR2 variable is accounted for by the ANOVA model tested in which FISHES is the variable that contributes the most and no differences in the levels of pond type or SEASGRUP occurred.

Another 11% of the yield variability is due to common carp yield, which depends on total density and the proportion of tilapia in the polyculture. This combination shows higher common carp yields when this fish is grown in monoculture (or with mullets). The remaining 8% of the variability is accounted for by silver carp yield, which was higher in reservoirs than in shallow ponds. The ANOVA models tested account for 38% and 54% of the variability of these canonical variables, respectively in which FISHES is the variable that contributes the most.

Daily Growth

Table 21 shows results of canonical correlations on Bet She'an Valley growth data. Most of the variability (45%) is accounted for by mullet and silver carp daily growth. Good growth of both species occurred in deep ponds (reservoirs), under high manuring rates, long culture seasons, and when mullets were stocked at rather large sizes and common carp at small sizes. The ANOVA models tested account for 79% of the variability of this canonical variable in which FISHES is the variable that contributes the most.

The second source of growth variability also explains a large part of the overall variability (42%). It is due to tilapia, which grew better when its own proportion in the polyculture was large and that of common carp was small, when it was

Table 20b. Bet She'an data analysis. Duncan's multicomparison test of means of yield response variables according to culture period (SEASGRUP), fish combination (FISHES), stocking month and pond type. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping			
	n	YR1 (Tilapia - mullet)	YR2 (Mullet + tilapia)	YR3 (Carp)	YR4 (Silver carp)
SEASGRUP					
a (- sp)	1	b c	a	a	a b
c (- sp su)	38	a b	b	a	a b
d (- sp su au-)	31	b c	b	a	a b
e (su)	2	a	b	a	a
f (su au-)	10	a	b	a	a b
g (au-)	2	c	b	a	b
h (xx- xx)	33	b c	b	a	a
FISHES					
1 (c)	9	b c	a	a	b c
3 (c t)	24	a	c d	b c	b c
4 (c m)	10	d	d	a b	c
5 (c s)	1	d	a b	c	a
7 (c ms)	6	d	c d	b c	b c
8 (c t s)	12	a b	b c	b c	b
9 (c t ms)	55	c d	c d	b c	b
STOCKING MONTH OF FIRST SPECIES					
I	11	c d e	b c d e	b	a b
II	3	g	b c d e	b	b
III	14	e f g	b c d e	a b	b
IV	13	b c	a b	a b	b
V	7	b c d	c d e	a b	b
VI	6	b	a b c d	a b	b
VII	5	a	e	b	b
VIII	5	f g	a	b	a
IX	11	f g	d e	b	b
X	19	f g	a b c d e	b	b
XI	15	f g	a b c	a b	b
XII	8	d e f	d e	a	a b
PONDTYPE					
Pond	64	a	a	a	b
Reservoir	53	b	a	a	a

cultured for long periods and was stocked at rather large sizes. The ANOVA models tested account for 70% of the variability of this canonical variable, in which FISHES is the variable that contributes the most. This combination did not present significant differences among SEASGRUPs or type of pond.

Another 8% of the growth variability is due to silver carp and mullet growth. Good silver carp growth and bad growth of mullet occurred when fish were cultured in deep ponds, at high total density, when silver carp was stocked at rather large sizes, and common carp at small sizes, common and silver carp proportions in the polyculture were large, and the culture season was long. The ANOVA models tested account for only 42% of the variability of this canonical variable, in which SEASGRUP and stocking month are the variables that contribute the most. The remaining 5% of the variability is accounted for by common and silver carp growth.

UPPER GALILEE

In this region only a few ponds analyzed included mullet, thence the variables related to this species were not included. The ANOVA models of this region include as categorical variables SEASGRUP, FISHES and stocking month.

Table 21a. Bet She'an analysis. Results of canonical correlation analysis of growth data (GCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structure analysis). n = 117.

	GCC1	GCC2	GCC3	GCC4
Canonical correlation coefficient	0.87	0.85	0.59	0.49
Variance accounted for (%)	45	42	8	5
Standardized canonical coefficients for the explanatory variables				
	GE1	GE2	GE3	GE4
DEPTH	(+)	.	(+)	0.69
AREA	.	.	-0.44	-0.39
SORGFISH	.	.	.	.
PELFISH	.	.	.	-0.36
MANUDAY	(+)	.	.	.
TOTDEN	.	.	0.47	-0.34
CWTI	(-)	.	(-)	.
PERCCDEN	.	(-)	0.78	0.64
CGRDAYS	-0.56	.	0.86	-1.51
TWTI	.	0.40	.	.
PERCTDEN	.	0.31	.	0.78
TGRDAYS	(+)	0.42	0.52	.
SWTI	.	.	0.70	0.31
PERCSDEN	.	.	0.33	.
SGRDAYS	(+)	.	0.45	.
MWTI	(+)	.	.	.
MGRDAYS	0.93	-0.44	.	1.02
Standardized canonical caefficients for the response variables				
	GR1	GR2	GR3	GR4
CGROWTH	.	.	-0.38	0.98
TGROWTH	.	0.99	.	.
SGROWTH	(+)	.	1.05	0.59
MGROWTH	0.81	-0.36	-0.80	.

Table 21b. Bet She'an analysis. Duncan's multicomparison test of means of growth response variables according to culture period (SEASGRUP), fish combination (FISHES), stocking month and pond type. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping			
	n	GR1 (Mullet + silver carp)	GR2 (Tilapia - mullet)	GR3 (Silver carp - mullet)	GR4 (Carp + silver carp)
SEASGRUP					
a (- sp)	1	c	b	a	b
c (- sp su)	38	a b	a	a	a b
d (- sp su au-)	31	a	a b	a	b
e (su)	2	b c	a b	a	a b
f (su au-)	10	b c	a	a	b
g (au-)	2	b c	b	a	a
h (xx- xx)	33	a	a	a	a b
FISHES					
1 (c)	9	c	c	b c	a
3 (c t)	24	b c	a b	b c	a
4 (c m)	10	b	c	c	a
5 (c s)	1	b c	c	a	a
7 (c ms)	6	a	c	b c	a
8 (c t s)	12	b c	a	a b	a
9 (c t ms)	55	a	b	b c	a
STOCKING MONTH OF FIRST SPECIES					
I	11	a b	a b c	b	a b c
II	3	a b	d	b	b c
III	14	b c	c	b	a b
IV	13	d	a b	b	a b
V	7	c d	b c	b	a b
VI	6	d	b c	b	a b c
VII	5	d	a	b	c
VIII	5	a b	c	a	a
IX	11	a	c	b	a b c
X	19	a b	c	b	a b
XI	15	a b	c	b	a b c
XII	8	a b	c	b	a b
PONDTYPE					
Pond	64	b	a	b	a
Reservoir	53	a	a	a	a

Relationships Among Variables

Table 22 presents factor analysis of Upper Galilee data. The first factor (FACTOR1) accounts for 25% of the overall variability. It shows total yield (yield of all species) correlated with fish density (not of silver carp), feed inputs (pellets positive and sorghum negative), length of culture period, large proportion of tilapia and small proportion of common carp, large tilapia and small carp at stocking, conditions which also affect positively the tilapia and silver carp growth rates. Most of the variability of this combination (95%) is accounted for by the ANOVA model tested to which FISHES contribute the most. The highest total yield (FACTOR1) occurred when tilapia and the two carp species were present (FISHES 8 and 9), when fish were cultured in a year-long season (SEASGRUP d) and when fish were stocked in the spring months. On the other hand, the lowest values of FACTOR1 occurred in spring and autumn, in which large carp which had not yet reached commercial size are kept in the ponds at low densities and are fed sorghum until they can be marketed.

The second and third factors (26% of variance altogether) are related to silver carp. FACTOR2 shows silver carp performance negatively correlated with tilapia density, percentage and yield. Most of the variability of this combination (79%) is accounted for by the ANOVA model tested to which FISHES contributes the most due to the absence of tilapia or silver carp in some of the polyculture systems. High silver carp and low tilapia yields occurred in the SEASGRUPs in which growth starts in spring or when the first fish species in the pond was stocked in winter months. High tilapia and low silver carp yields occurred when fish were cultured in SEASGRUPs which start in the second part of the year. FACTOR3 shows that larger relative and absolute amounts of silver carp were cultured in small ponds with low common carp proportion in the polyculture. About 45% of the variability of this factor is accounted for by the ANOVA model tested to which FISHES

Table 22a. Factor analysis of Upper Galilee data. Only large coefficients (>0.4) are included. n = 256.

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
DEPTH	.	.	.	-0.42	0.46
AREA	.	.	-0.40	-0.43	0.43
SORGFISH	-0.73	.	.	.	.
PELFISH	0.60	.	.	0.45	.
MANUDAY	.	.	.	.	.
TOTDEN	0.74	.	.	.	.
CWTI	-0.65	.	.	.	.
CDEN	0.60	.	.	.	-0.41
PERCCDEN	-0.80	.	-0.47	.	.
CGRDAYS	0.78	.	.	.	.
CWTO	.	.	.	.	0.55
CYIELDAY	0.52	.	.	0.41	.
CGROWTH	.	.	.	0.50	0.56
TWTI	0.53	.	.	0.42	.
TDEN	0.63	-0.48	0.40	.	.
PERCTDEN	0.81	-0.43	.	.	.
TGRDAYS	0.88	.	.	.	.
TWTO	0.82	.	.	.	.
TYIELDAY	0.74	-0.42	.	.	.
TGROWTH	0.80	.	.	.	.
SWTI	.	.	.	.	-0.41
SDEN	.	0.61	0.63	.	.
PERCSDEN	.	0.65	0.59	.	.
SGRDAYS	0.75	0.45	.	.	.
SWTO	0.73	0.46	.	.	.
SYIELDAY	0.52	0.61	0.41	.	.
SGROWTH	0.69	0.42	.	.	.
Variance explained	25%	15%	11%	7%	6%

Table 22b. Upper Galilee data analysis. Duncan's multicomparison test of means of FACTORs according to culture period (SEASGRUP), fish combination (FISHES) and stocking month. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping				
	n	Factor 1 (Total yield)	Factor 2 (Silver carp)	Factor 3 (Silver carp)	Factor 4 (Carp)	Factor 5 (Carp)
SEASGRUP						
a (- sp)	67	f	a b	a b	a b	a
c (- spsu)	50	c	a b	a b	a b	b
d (- spsuau-)	58	a	a	c	b	a b
e (su)	14	d	c	a	a	a b
f (suau-)	47	b	b c	a	a b	b
g (au-)	15	e	b c	a	a	a b
h (xx- xx)	5	c	b c	b c	b	a b
FISHES						
1 (c)	92	d	c	b c	a b	a
3 (c t)	28	b	d	a	b	a
4 (c m)	2	c	b c	c	b	a
5 (c s)	37	c	a	a	a b	a
7 (c ms)	4	b	a	c	a b	a
8 (c t s)	84	a	b c	a b	a b	a
9 (c t ms)	9	a	b	c	a	a
STOCKING MONTH OF FIRST SPECIES						
I	24	e f	a b	b c d	c d	a
II	15	c d	b c	b c d	a b c	a b c
III	20	b	a	d	d	c
IV	6	c	d e	a b c	d	a b c
V	45	b	c d	a b c	a b c d	a b c
VI	36	a	b c	a b	a b c d	b c
VII	11	c d	e	a b c	b c d	a b c
VIII	14	d	c d	a	a	a b c
IX	15	e	b c	a b c	a b	a b c
X	5	f	c d	c d	b c d	a b c
XI	34	c d	a	b c d	a b c	a b
XII	31	e f	a	b c d	b c d	a b

and SEASGRUP contribute the most. High values of this factor occurred when mullets were not included in the ponds, in very short and short SEASGRUPs, and when stocking was between April and September.

The next two factors (13% of variance) are related to common carp and the ANOVA model tested accounts for less than 25% of each factor variance. FACTOR4 shows common carp growth and yield positively correlated with feed pellet inputs and tilapia stocking weight and negatively correlated with pond size (depth and area). This combination had higher values in the very short, hot SEASGRUPs than in the long ones. FACTOR5 shows higher common carp growth and weight at harvest when cultured in large ponds at low common carp density and with small silver carp.

Daily Yields

Table 23 shows canonical correlation results of yield data from the Upper Galilee region. Most of the variability (65%) in the yield data of this region is related to tilapia yield. Tilapia yield is positively correlated with total density and tilapia proportion in it, its stocking weight, feed pellet inputs, and length of culture period, and negatively with common carp stocking weight and feed sorghum inputs. About 67% of the variability of the YR1 is accounted for by the ANOVA model tested to which FISHES contributes the most since tilapia was not present in all the polyculture combinations. The presence of other fish species in the polyculture did not affect tilapia yield (no significant differences between FISHES 3, 8 and 9). Higher tilapia daily yields occurred when cultured

Table 23a. Upper Galilee analysis. Results of canonical correlation analysis of yield data (YCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structure analysis). n = 256.

	YCC1	YCC2	YCC3
Canonical correlation coefficient	0.88	0.73	0.63
Variance accounted for (%)	65	22	13
Standardized canonical coefficients for the explanatory variables			
	YE1	YE2	YE3
DEPTH	.	.	.
AREA	.	.	.
SORGFISH	(-)	.	.
PELFISH	(+)	.	0.63
MANUDAY	.	.	0.42
TOTDEN	0.34	.	0.61
CWTI	(-)	.	-0.56
PERCCDEN	(-)	-1.68	0.59
CGRDAYS	(+)	-0.54	-0.36
TWTI	(+)	.	.
PERCTDEN	(+)	-1.88	-0.71
TGRDAYS	(+)	.	0.31
SWTI	.	0.38	.
SGRDAYS	(+)	0.65	.
Standardized canonical coefficients for the response variables			
	YR1	YR2	YR3
CYIELDAY	.	.	1.0
TYIELDAY	0.84	-0.47	-0.45
SYIELDAY	.	1.01	.

Table 23b. Upper Galilee analysis. Duncan's multicomparison test of means of yield response variables according to culture period (SEASGRUP), fish combination (FISHES) and stocking month. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping		
	n	YR1 (Tilapia)	YR2 (Silver carp)	YR3 (Common carp)
SEASGRUP				
a (- sp)	67	d	a b c	a
c (- spsu)	50	a b	a	a
d (- spsuau-)	58	a	a b	a
e (su)	14	b c	a b c	a
f (suau-)	47	a	a b	a
g (au-)	15	c	c	a
h (xx- xx)	5	b c	b c	a
FISHES				
1 (c)	92	b	c	a b
3 (c t)	28	a	d	b
4 (c m)	2	b	c	a b
5 (c s)	37	b	a	a b
7 (c ms)	4	b	a b	a
8 (c t s)	84	a	b c	a b
9 (c t ms)	9	a	b c	a
STOCKING MONTH OF FIRST SPECIES				
I	24	d e	a b c d	a b
II	15	b c d	a b c d	a b
III	20	b c	a	a b
IV	6	c d e	a b c d	b
V	45	a b	a b c d	a b
VI	36	a	a b c	a b
VII	11	b c	c d	a b
VIII	14	c d e	a b c d	a b
IX	15	c d e	d	a b
X	5	e	b c d	a b
XI	34	c d e	a b	a
XII	31	d e	a b c d	a b

in half or complete year cycles (SEASGRUPs c, d, f) and when stocked in spring or summer.

Another 22% of the yield variation in this geographic area is due mainly to silver carp yield which was negatively correlated with tilapia yield. High silver carp and low tilapia yields are related to small common carp and tilapia proportions in the polyculture (which indicates a larger proportion of silver carp in it), short common carp and long silver carp culture periods, and rather large silver carp at stocking.

The last 13% of variability are accounted for mainly by common carp yield, which was negatively correlated with tilapia yield. Higher carp yield occurred when large amounts of feed pellets and manure were applied, total density was high, carp was stocked at small sizes and its proportion in the polyculture was large while that of tilapia was small.

Daily Growth

Most of the variability (75%) of the Upper Galilee growth data is accounted for by tilapia and silver carp (Table 24). Conditions in which the growth of both species was good include large amounts of feed pellets and little or no sorghum, high total density and tilapia proportion in it, large tilapia and small common carp at stocking, and long culture periods of all species. About 78% of the variability of the GR1 is accounted for by the ANOVA model tested to which FISHES contributes the most since tilapia and/or silver carp were not present in all the polyculture combinations. Conditions for good growth of both species occurred mainly in the year-long season (SEASGRUP d), and when stocking was in the spring months.

Another 21% of the variability is also due to silver carp and tilapia growth, which were here negatively correlated. Growth of each species is positively correlated with its stocking weight and negatively with that of the other species and also length of silver carp culture season is involved. This correlation mainly shows differences in fish species combinations in which each species grew better when the other was absent.

The last 4% of variability is due to common carp, which grew better in short culture periods, at low total density of which tilapia represents a large proportion, when stocked at small sizes and under high nutritional input rates (mainly sorghum but also feed pellets and manure).

Analysis of Reservoir Data

RELATIONSHIPS AMONG VARIABLES

Factor analysis was applied to 118 observations from deep reservoirs. Results of this analysis and the differences of each factor according to SEASGRUP, FISHES, stocking month and region are presented in Table 25.

The first factor (FACTOR1) accounts for 20% of the overall variability of the data. It shows tilapia performance (growth, yield and harvesting size) correlated to the total fish density and the absolute and relative density of this fish species but not to its stocking size. When tilapia is the main fish in the polyculture, tilapia performance is better and common carp reaches larger sizes at harvest. The ANOVA model tested accounts for 85% of the variance of this factor to which FISHES contributes the most. Mullet reduced tilapia performance (FISHES 9<8) while silver carp did not affect it (FISHES 3=8). Better tilapia performance occurred when cultured in long seasons (SEASGRUPs d and h). The best tilapia performance is in the Upper Galilee.

The second factor (FACTOR2), which accounts for another 17% of the variance, is due to mullet. Mullet performance was better when its culture period was long and it was stocked at a large size and in higher densities, hence common carp participation was lower. This is associated with larger silver carp at harvest. The ANOVA model tested accounts for 88% of the variance of this factor to which FISHES contributes the most. The combination of characteristics indicated by this factor which led to better mullet performance occurred when the polyculture included the four species. Mullet had also better performance in reservoirs which were stocked in summer or autumn and harvested in the spring or summer of the next year (SEASGRUP h). This timing and polyculture combination is mainly used in the Bet She'an region while in the Upper Galilee mullet culture is not very extended.

The third factor (FACTOR3, 12% of variance) shows silver carp performance related to its absolute and relative density and to be inversely correlated with common carp growth rate. The ANOVA model tested accounts for 64% of this factor variance to which FISHES contributes the most. Better silver carp performance and lower common carp growth rate occurred in the long culture seasons, and when tilapia was not included in the reservoirs, and was not affected by the presence of mullets.

Table 24a. Upper Galilee analysis. Results of canonical correlation analysis of growth data (GCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structure analysis). n = 256.

	GCC1	GCC2	GCC3
Canonical correlation coefficient	0.89	0.72	0.43
Variance accounted for (%)	75	21	4
Standardized canonical coefficients for the explanatory variables			
	GE1	GE2	GE3
DEPTH	.	.	.
AREA	.	.	.
SORGFISH	(-)	.	0.64
PELFISH	(+)	.	0.36
MANUDAY	.	.	0.37
TOTDEN	(+)	.	-0.54
CWTI	(-)	.	-0.54
PERCCDEN	(-)	.	.
CGRDAYS	(+)	.	-1.57
TWTI	(+)	-0.49	.
PERCTDEN	(+)	.	0.74
TGRDAYS	0.42	.	.
SWTI	.	0.54	.
SGRDAYS	(+)	0.85	0.76
Standardized canonical caefficients for the response variables			
	GR1	GR2	GR3
CGROWTH	.	.	1.01
TGROWTH	0.89	-0.68	.
SGROWTH	(+)	1.07	0.33

Table 24b. Upper Galilee analysis. Duncan's multicomparison test of means of growth response variables according to culture period (SEASGRUP), fish combination (FISHES) and stocking month. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping		
	n	GR1 (Tilapia)	GR2 (Silver carp - tilapia)	GR3 (Common carp)
SEASGRUP				
a (- sp)	67	d	b c	a b
c (- sp su)	50	b	a b c	b c
d (- sp su au-)	58	a	a	a b c
e (su)	14	b c	a b c	a
f (su au-)	47	b	a b	a b c
g (au-)	15	d	c	a b
h (xx- xx)	5	c	a b c	c
FISHES				
1 (c)	92	c	d	a
3 (c t)	28	b	e	a
4 (c m)	2	c	c d	a
5 (c s)	37	c	a b	a
7 (c ms)	4	c	a	a
8 (c t s)	84	a b	b c	a
9 (c t ms)	9	a	c d	a
STOCKING MONTH OF FIRST SPECIES				
I	24	c d	a b c	a b
II	15	a b	a b c	a
III	20	a	a b	b
IV	6	b c	a b	a b
V	45	a	a b c	a b
VI	36	a	a	a b
VII	11	c d	b c	a b
VIII	14	b c d	c	a b
IX	15	c d	c	a
X	5	d	c	a b
XI	34	b c	a b c	a
XII	31	c d	a b c	a b

Table 25a. Factor analysis of reservoir data. Only large coefficients (>0.40) are included.

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
AREA	.	.	.	.	.
SORGFISH	.	.	.	.	.
PELFISH	.	.	.	.	.
MANUDAY	.	.	.	.	.
TOTDEN	0.65	.	.	0.65	.
CWTI	.	.	.	.	.
CDEN	.	.	.	0.63	.
PERCCDEN	-0.65	-0.53	.	.	.
CGRDAYS	.	.	0.49	.	.
CWTO	0.57	.	.	.	.
CYIELDAY	.	.	.	0.58	.
CGROWTH	.	.	-0.61	.	.
TWTI	.	.	.	-0.48	.
TDEN	0.82	.	.	.	.
PERCTDEN	0.92	.	.	.	.
TGRDAYS	0.78	.	.	.	.
TWTO	0.72	.	.	.	.
TYIELDAY	0.87	.	.	.	.
TGROWTH	0.73	.	.	.	.
MWTI	.	0.57	.	.	.
MDEN	.	0.63	.	.	.
PERCMDEN	-0.52	0.55	.	.	.
MGRDAYS	.	0.85	.	.	.
MWTO	.	0.85	.	.	.
MYIELDAY	.	0.72	.	.	.
MGROWTH	.	0.79	.	.	.
SWTI	.	.	.	.	0.49
SDEN	.	.	0.56	.	-0.58
PERCSDEN	.	.	0.60	.	-0.51
SGRDAYS	.	.	0.62	.	.
SWTO	.	0.51	0.52	.	0.44
SYIELDAY	.	.	0.69	.	.
SGROWTH	.	0.46	.	.	.
Variance explained	20%	17%	12%	9%	6%

The fourth factor (FACTOR4, 9% of variance) shows common carp yield related to its own and to total density and inversely correlated with tilapia stocking weight. The ANOVA model tested only accounts for 42% of the variability of this factor. Higher values of FACTOR4 occurred when fish were cultured in a long season not interrupted by winter (SEASGRUP d). Carp yield and density were higher in monoculture or in polyculture with mullet but without tilapia and lower when tilapia or silver carp were present. No differences among regions occurred.

The fifth factor (FACTOR5, 6% of variance) is due to silver carp size which is inversely correlated with its density. Larger silver carp at lower densities were grown in the Coastal Plain and during half year culture seasons.

DAILY YIELDS

Table 26 shows canonical correlation results of daily yield data in reservoirs. Daily yield of each species in this analysis is mainly correlated with total density and proportion of the species in the polyculture. In the case of common carp for which data of monoculture and polyculture systems are available, the canonical correlation points to higher yields in monoculture.

Most of the daily yield variability in reservoirs (73%) is due to tilapia which gave higher yields when besides constituting a large part of the polyculture, were stocked at rather large sizes and cultured for long periods. Under these combinations, there is also a positive correlation between tilapia and common carp daily yields. The ANOVA model tested accounts for 56% of the YR1 variabil-

Table 25b. Analysis of reservoir data. Duncan's multicomparison test of means of FACTORs according to culture period (SEASGRUP), fish combination (FISHES), stocking month and region. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping				
	n	Factor 1 (Tilapia)	Factor 2 (Mullets)	Factor 3 (Silver carp)	Factor 4 (Carp)	Factor 5 (Silver carp)
SEASGRUP						
c (- spsu)	28	b	b	b	b	a
d (- spsuau-)	58	a	b	a	a	a b
h (xx- xx)	32	a	a	a	b	b
FISHES						
1 (c)	5	c	e	d	a	a
3 (c t)	5	a b	d	e	c	a
4 (c m)	9	d	b	e	a	a
5 (c s)	6	c d	d	a	b c	a
7 (c ms)	17	d	b	a b	a b	a
8 (c t s)	26	a	c	b c	b c	a
9 (c t ms)	50	b	a	c d	b c	a
STOCKING MONTH OF FIRST SPECIES						
I	16	a	c d	b c	a	a
II	18	b	d e	b	a b	a
III	18	a b	e f	b c	a b	a
IV	4	a	f	b c	a b	a
VIII	2	a	a	a	b	a
IX	11	a	b	b c	a b	a
X	19	a b	b c	b c	b	a
XI	15	a	d e	c	a b	a
XII	15	a	e f	b c	a	a
REGION						
West Galilee	33	c	c	b	a	b
Bet She'an Valley	53	b	a	b	a	b
Coastal Plain	11	b	b	a	a	a
Upper Galilee	21	a	d	a b	a	b

ity of which FISHES contributes the most. Higher tilapia (and common carp) yields occurred in the Upper Galilee and lower in the Western Galilee. None of the differences due to culture period were significant.

DAILY GROWTH

Table 27 presents results of daily growth analysis in reservoirs. Most of the data variance (79%) is due to tilapia growth rate. High tilapia growth rate occurred when it constituted a large proportion of the polyculture, when it was stocked at rather large sizes and was cultured for long periods. Under these conditions growth rate of common carp was also high. Most of the GR1 variability (77%) is accounted for by the ANOVA model tested to which FISHES contributes the most. This combination had higher values when stocking was carried out in the second half of the year, when effective growth took place before and after winter or only up to half a year after winter (SEASGRUPs h, c), and in the Bet She'an and Upper Galilee regions.

High common carp together with low tilapia and silver carp growth rates in reservoirs are related to the combination of high tilapia and low carp proportions in the polyculture, short culture periods, small tilapia and silver carp sizes at stocking, and large levels of feed pellets. This combination mainly occurred in the fish combinations which did not include silver carp, when the effective culture period started in spring and in both Galilee regions.

Table 26a. Analysis of reservoir data. Results of canonical correlation analysis of yield data (YCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structure analysis). n = 118.

	YCC1	YCC2	YCC3
Canonical correlation coefficient	0.91	0.77	0.46
Variance accounted for (%)	73	23	4
Standardized canonical coefficients for the explanatory variables			
	YE1	YE2	YE3
AREA	.	.	.
SORGFISH	.	.	.
PELFISH	.	.	.
MANUDAY	.	.	-0.43
TOTDEN	0.53	0.95	0.53
CWTI	.	.	.
PERCCDEN	(-)	0.68	.
CGRDAYS	.	.	(+)
TWTI	(+)	.	.
PERCTDEN	0.86	-0.61	-0.83
TGRDAYS	(+)	0.36	.
SWTI	.	.	.
SGRDAYS	.	.	0.70
Standardized canonical coefficients for the response variables			
	YR1	YR2	YR3
CYIELDAY	(+)	0.98	.
TYIELDAY	0.89	-0.52	.
SYIELDAY	.	.	1.00

Table 26b. Analysis of reservoir data. Duncan's multicomparison test of means of yield response variables according to culture period (SEASGRUP), fish combination (FISHES), stocking month and region. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping		
	n	YR1 (Tilapia)	YR2 (Carp)	YR3 (Silver carp)
SEASGRUP				
c (- spsu)	28	a	a	b
d (- spsuau-)	58	a	a	b
h (xx- xx)	32	a	b	a
FISHES				
1 (c)	5	b	a	b
3 (c t)	5	a	b	b
4 (c m)	9	b	b	b
5 (c s)	6	b	b	a
7 (c ms)	17	b	b	a
8 (c t s)	26	a	b	a
9 (c t ms)	50	a	b	a
STOCKING MONTH OF FIRST SPECIES				
I	16	a	a b	a
II	18	b	a b	a
III	18	a b	a b	a
IV	4	a b	a b	a
VIII	2	a b	a b	a
IX	11	a b	b	a
X	19	a b	a b	a
XI	15	a b	a	a
XII	15	a	a	a
REGION				
West Galilee	33	c	a	a
Bet She'an Valley	53	b	a	a
Coastal Plain	11	b	a	a
Upper Galilee	21	a	a	a

Table 27a. Analysis of reservoir data. Results of canonical correlation analysis of growth data (GCC-). Only coefficients higher than 0.3 are included. Signs (+) and (-) point to variables correlated with the canonical variates (canonical structure analysis). n = 118.

	GCC1	GCC2	GCC3
Canonical correlation coefficient	0.90	0.69	0.38
Variance accounted for (%)	79	18	3
Standardized canonical coefficients for the explanatory variables			
	GE1	GE2	GE3
AREA	.	.	.
SORGFISH	.	.	.
PELFISH	.	-0.33	.
MANUDAY	.	.	.
TOTDEN	.	.	-0.46
CWTI	.	.	.
PERCCDEN	(-)	0.31	.
CGRDAYS	.	0.55	-0.39
TWTI	0.50	0.34	-0.52
PERCTDEN	0.40	-0.42	0.34
TGRDAYS	0.34	0.41	.
SWTI	.	0.36	0.43
SGRDAYS	.	(+)	0.59
Standardized canonical coefficients for the response variables			
	GR1	GR2	GR3
CGROWTH	(+)	-0.99	0.45
TGROWTH	0.97	0.44	.
SGROWTH	.	(+)	1.00

Table 27b. Analysis of reservoirs data. Duncan's multicomparison test of means of growth response variables according to culture period (SEASGRUP), fish combination (FISHES), stocking month and region. Groups with the same letter are not significantly different. a>b> etc. n = number of observations.

		Duncan grouping		
	n	GR1 (Tilapia)	GR2 (Carp)	GR3 (Silver carp)
SEASGRUP				
c (- spsu)	28	a	b	a
d (- spsuau-)	58	b	a b	a
h (xx- xx)	32	a	a	a
FISHES				
1 (c)	5	c	b c d	b
3 (c t)	5	a	d	b
4 (c m)	9	c	c d	b
5 (c s)	6	c	a	a
7 (c ms)	17	c	a b	a
8 (c t s)	26	b	a b c	a
9 (c t ms)	50	b	a	a
STOCKING MONTH OF FIRST SPECIES				
I	16	b c	b c	b
II	18	c	b	b
III	18	b c	b c	b
IV	4	a b c	b c	b
VIII	2	a	a	a
IX	11	a b	b	b
X	19	a b	b c	b
XI	15	a b	c	b
XII	15	a b c	b c	b
REGION				
West Galilee	33	c	b	a b
Bet She'an Valley	53	a	a b	a
Coastal Plain	11	b	a	b
Upper Galilee	21	a	b	a b

Discussion and Conclusions

The results of the analyses on commercial farm data presented in this paper are basicaly similar to those of Milstein et al. (1988) on experimental station data. This is in spite of the fact that the latter was based on a much more homogeneous system than the farms dataset, regarding fish combinations, culture season, duration of culture period, pond size and covered range of variables. Specifically, the presence or absence of a fish species in the polyculture systems in the farms dataset increased the variance and those species not present in all observations had a stronger weight in the factors and canonical correlations obtained than the species present in all observations. Nevertheless, in both analyses, the yield of each species was mainly affected by its own stocking density; feed pellets mainly affected common carp with little or no effect on other species; and most of the variance in the polyculture system is accounted for by tilapia.

Some conclusions of Milstein et al. (1988) were not confirmed in the present analyses, also due to the heterogeneous nature of the farms dataset and the different ranges of some of the variables. The stocking weight range of tilapia was much lower (1 to 26 g) in the experimental ponds than in the farms dataset (0 to 500 g). Thus, the conclusion from the experimental dataset that better performance of tilapia occurs when stocked over 13 g is irrelevant here. Similarly, the conclusion from the experimental dataset that common carp performance is negatively affected by silver carp density when it is over 1,000$\cdot$ha^{-1}, is irrelevant in the farms dataset since here silver carp density is rarely over 1,000$\cdot$ha^{-1}. However, the opposite is herein supported, i.e., at the low silver carp densities of the farm dataset a positive effect of this fish on common carp was evidenced in some analysis (see Figs. 3 and 4).

Factor analysis and canonical correlations are techniques to explore LINEAR relationships among variables. Relationships which are not linear but their deviations from linearity are small will be detected but will appear with lower coefficients. Relationships which are linear in a large part of their range of values and nonlinear in the rest will also have lower coefficients as shown in Milstein and Prein (this vol.). Relationships strongly deviating from linearity will not be detected. This accounts for some apparent discrepancies between the results presented here and the analysis of the same dataset through the GLM procedure presented in Milstein et al. (this vol.). The GLM procedure allows to include continuous and categorical variables so that if linearity is doubtful the problem can be avoided by transforming the adequate continuous variables into categorical ones. This is the case for manuring rate, which in the daily yield GLM was used as a discrete variable and is an important element of the model, while in the factor and canonical models presented here generally has low coefficients.

Application of different multivariate statistical methods to the same dataset reflects different approaches to the understanding of complex systems. The factor and canonical correlation analyses presented here are oriented towards the search of multiple interactions among one or two groups of variables while the (GLM) multiple regression method presented in Milstein et al. (this vol.) aims to explain one variable in terms of several others. The different approaches are seen in the selection of the response variables included in the models when multiple regression and canonical correlation methods are compared. In the former case, with only one response variable allowed per model, the variables selected (total yield, daily yield or profit) are a measure of the general behavior of the system. In the case of canonical correlations, the variables selected allow the study of the role of each species and their interactions in the yield or growth within the system.

The data analyzed are from well managed systems where the different polyculture combinations are well balanced. This implies no significant negative interactions among the different fish species so that the yield and growth limitations for each species depend mainly on each species parameters (especially on density) as shown by the factor and the canonical correlation analyses. Of both parameters, growth rates of each species are more affected than their yields by the presence of other species in the ponds. The common carp is the only species whose yield and growth were strongly affected by the amount of feed pellets but only in the second place after the effects of its own parameters.

The set of analyses presented here show that common carp performance (yield, growth rate and size at harvest) is mainly related to amount of feed pellets per fish applied, contrary to tilapia and silver carp, the performance of which is mainly related to their relative densities. In the dataset analyzed, common carp stocking characteristics vary from "many small carp cultured for long periods" to "few large carp cultured for short periods" (FACTOR3, complete array of culture systems, Table 12a), and for each level of these stocking characteristics the higher the amounts of feed pellets the better the carp performance (FACTOR4, Table 12a). When considering yields and growth separately (canonical correlation analyses) density effects on each species are evidenced but feed pellets still show a strong effect only on common carp. Being the first and the main species cultured in Israel, common carp culture technology reached a level in which the management practices stabilized and vary little over the country. This explains the low contribution of carp to the models presented in which factors and canonical correlations related to this species account for only 10% to 20% of the overall data variability.

Tilapia higher yield, growth rate and size at harvest are mainly related to high absolute and relative tilapia density. Large tilapia at stocking and to some extent also large common carp, favor tilapia growth rate. Tilapia performance is better when the growing period includes summer. It is not so strongly affected by nutritional inputs as the common carp, except for a negative influence of sorghum application (mainly in the Bet She'an Valley and Upper Galilee areas). A positive effect of feed pellets on tilapia performance can be seen in the analysis of the more homogeneous carp-tilapia system when other stronger effects are neutralized.

Silver carp yield, growth rate and size at harvest are mainly related to its density. Silver carp densities applied in the Israeli polyculture are rather low (300 to 700 silver carp·ha^{-1}), since this fish is used as a sanitary fish rather than for flesh production. As a sanitary fish, it positively affects common carp (Tang 1970; Milstein 1990), when its density is up to 1,000·ha^{-1} (Milstein et al. 1988), as is the case in Israel. In the Chinese polyculture system where it is one of the important species for flesh production, its stocking density varies from few hundreds to many thousands per hectare (Bardach et al. 1972).

Tilapia and silver carp, at the density ranges analyzed here, show a strong linear correlation of their performance with their density. For these fish species, the maximum density at which growth starts to decrease was not achieved, contrary to the case of the common carp.

Pond size had a relatively low influence on fish production. In all cases where fishpond area and depth were included in the models the coefficients were low and were mainly related to silver carp and mullet performances. These species represent a small proportion of the total Israeli fish culture, the former due to its low marketable value and the latter due to its scarcity. When only the deep reservoirs are analyzed, the main source of variability is related to tilapia as in the general analyses.

The relationships among variables found for the whole dataset are also found for the Western Galilee and the Coastal areas. In the Bet She'an Valley the differences between ponds and reservoirs are stronger. In the reservoirs, silver carp and mullet show better performance while tilapia does in ponds. In this region, higher manuring rates are applied to reservoirs than to ponds, common carp was positively affected by manure and tilapia were negatively affected by sorghum. In the Upper Galilee strong interactions among fish species occured, unlike in the other regions. The lower total yields in the Coastal Plain and the Western Galilee (see also Sarig 1990) are due to higher proportions of common carp and lower of tilapia compared to the other regions, leading to lower total stocking densities. These differences are mainly due to climatic conditions.

Acknowledgements

We are indebted to the fish farmers who kindly allowed us to use their data; to the Regional Fish Culture Extension Officers F. Svirski, I. Peleg and A. Ben Ari for their encouragement and help in gathering the information; and to P. Haviv for inputting the data into computer spreadsheets. This study was supported 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 for Economic Cooperation (BMZ).

References

Bardach, J.E., J.H. Ryther and W.O. McLarney, Editors. 1972. Aquaculture: the farming and husbandry of freshwater and marine organisms. Wiley Interscience, New York. 868 p.

Chang, W.Y.B. 1987. Fish culture in China. Fisheries 12(3):11-15.

Chang, W.Y.B. 1989. Integrated lake farming for fish and environmental management in large shallow Chinese lakes: a review. Aquacult. Fish. Manage. 20:441-452.

Hepher, B. 1985. Aquaculture intensification under land and water limitations. GeoJournal 10(3): 253-259.

Hepher, B. 1989. Principles of fish nutrition, p. 121-141. *In* M. Shilo and S. Sarig (eds.) Fish culture in warm water systems: problems and trends. CRC Press Inc., Boca Raton, Florida. 259 p.

Hepher, B. and Y. Pruginin. 1981. Commercial fish farming, with special reference to fish culture in Israel Wiley Interscience, New York. 261 p.

Milstein, A. 1990. Fish species interactions. EIFAC/90/Symp.R.5. 20 p.

Milstein, A., G. Hulata and G. Wohlfarth 1988. Canonical correlation analysis of relationships between management inputs and fish growth and yields in polyculture. Aquacult. Fish. Manage. 19:13-24.

Sarig, S. 1988. The development of polyculture in Israel: a model of intensification, p. 302-332. *In* C.J. Shepherd and N.R. Bromage (eds.) Intensive fish farming. BSP Professional Books, Oxford. 404 p.

Sarig, S. 1990. The fish culture industry in Israel in 1989. Isr. J. Aquacult./Bamidgeh 42(2):39-45.

SAS. 1985. SAS user's guide: statistics. Version 5 edition. Statistical Analysis Systems Institute Inc., Cary, Connecticut. 957 p.

Tang, Y.A. 1970. Evaluation of balance between fishes and available fishfoods in multispecies fish culture ponds in Taiwan. Trans. Am. Fish. Soc. 99:708-718.

Wohlfarth, G.W. and G. Hulata. 1987. Use of manures in aquaculture, p. 353-367. *In* D.J.W. Moriarty and R.S.V.Pullin (eds.) Detritus and microbial ecology in aquaculture. ICLARM Conf. Proc. 14, 420 p.

Wohlfarth, G.W. and G.L. Schroeder. 1979. Use of manure in fish farming - a review. Agricultural Wastes 1:279-299.

Multiple Regression Analysis of Fish Yield and Profit in Commercial Fish Farms in Israel

ANA MILSTEIN, *Agricultural Research Organization, Fish and Aquaculture Research Station, Dor, Mobile Post Hof Hacarmel 30820, Israel*

AMOS GOLDMAN, *Agricultural Research Organization, Department of Natural Resources, The Volcani Center, P.O. Box 6, Bet Dagan 50250, Israel*

GIDEON HULATA, *Agricultural Research Organization, Fish and Aquaculture Research Station, Dor, Mobile Post Hof Hacarmel 30820, Israel*

Milstein, A., A. Goldman and G. Hulata. 1993. Multiple regression analysis of fish yield and profit in commercial fish farms in Israel, p. 161-177. *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

Multiple regression analysis presented in this paper was applied to a dataset containing 1,124 observations from Israeli commercial farms. Analyses were run for the target variables total yield, daily yield and profit using the general linear model (GLM), which allows the inclusion of continuous and discrete independent (explanatory) variables as predictors. The total yield, daily yield and profit models account for 87%, 77% and 65% of the variability, respectively. The main management variables included in the yield models are supplementary and natural food inputs, followed by few fish-related variables, such as common carp and tilapia densities and stocking weights. In the profit model the single strongest factor was the farm, indicating strong differences in production efficiency among farms. These differences are partly due to climatic conditions in the geographical regions. Following the farm variable, further variables in the profit model relate to management practices (such as stocking densities and applied feeding regimes) and timing of culture period.

The GLM regression models integrate the effects of many independent variables. The coefficients of these effects are often different (and seem closer to the real ones) when compared to coefficients derived from univariate or bivariate analyses. Several procedures to increase the explanatory level of the models in order to improve their predictive value for users are presented.

Introduction

Fish culture, as any branch of agricultural production, is the outcome of a complex system of interacting variables that are manipulated by the producer to attain specific goals. Management know-how is based on acquired information from the common fund of knowledge and from experience. Agricultural research is directed to provide management information to be used to improve production of the farms. However, the heterogeneity of production conditions and the complexity of the production system produces many feedbacks and interactions that make difficult the direct extrapolation of standard management directives or tested research results to specific operating conditions. Consequently, much information is used ineffectively, reducing the beneficial impact of agricultural research on production goals. While the nature of the problem does not allow for simple and comprehensive solutions, much can be done to identify those factors that significantly influence production efficiency or research relevance in the context of farming reality. Often, data on the agricultural production process are available but are seldom analyzed comprehensively because of the large number of factors involved and because of the uneven and complex data distribution. On the other hand, data obtained in research experiments, which usually address the question "*What* is the effect of variables A-B-C on a specified target variable", lead to results which are relevant to the experimental conditions only. Thus, different experiments on the same subject often give conflicting results (Ziv and Goldman 1987a), and systematic differences in the effect of treatments between research experiments and commercial applications are common (e.g., Simmonds 1980; Ziv and Goldman 1987b). Data from research experiments are generally analyzed in the planned context only. Very seldom are data of a large number of experiments, which were conducted on the same subject and under variable conditions, integrated in such a way as to extract from them more of the valuable information which they contain. An integrative approach, including data from different sources to cover a wide range of conditions and the appropriate statistical methods, enables to address the questions "*Which* variables affect a specified target variable and *How much*". In the last years this integrative approach expanded into a new branch of statistics termed meta-analysis, which is developing especially in the medical sciences (Sacks et al. 1987).

The use of integrative techniques enables to deal more efficiently with the numerous factors acting simultaneously in the real system leading to more realistic and useful conclusions for management purposes. Integrative analysis of large datasets has become feasible with the introduction of user-friendly, multifactorial analysis packages. Such analyses have significant implications for management, extension and research policy. However, the nature of the data makes such analyses daunting and treacherous for the inexperienced, because of the high dimensionality, non-orthogonality, collinearity and inaccuracy of much of the data. Such data need to be checked carefully and evaluated critically before drawing conclusions, especially as statistical models used to describe the data can be notoriously unstable.

This paper presents examples of how to deal with (a) the problems involved in treating the available information on a complex culture system and the specifications of an appropriate computer program designed to analyze such systems, (b) the extent to which the general linear model (GLM) used simultaneously as multiple regression and analysis of variance meets these specifications, (c) certain measures which improve GLM use for such integrative analyses, and (d) the role of an integrative analysis in improving the 'technology transfer' system (research-extension-farm) and its capability to improve the efficiency of agricultural production. Examples of application of the GLM procedure to fish culture data from experimental facilities and commercial fish farms are given by Hulata et al. (this vol.) and Milstein et al. (this vol.), respectively.

Materials and Methods

Data from 1,124 commercial fishponds, from 31 farms, operated during the years 1976-1987 were assembled in a database described in Milstein and Hulata (this vol.). This database contains stocking parameters, nutritional inputs and production figures of the various fish species stocked into Israeli polyculture ponds: common carp (*Cyprinus carpio*), hybrid tilapia (mainly *Oreochromis niloticus* x *O. aureus*), silver carp (*Hypophthalmichthys molitrix*) or "namsif" (hybrid of silver carp with bighead carp, *Aristichthys nobilis*), mullet (*Mugil cephalus* and, to a lesser extent, *M. capito*), and occasionally grass carp (*Ctenopharyngodon idella*). Values were converted to units per 0.1 ha (=1 dunam, area unit used in Israel) where applicable. Total yield includes grass carp, but not wild spawning. Daily yields were calculated as mean daily values for the season, considering only the days of effective growth, that is, excluding the cold period between 15 November and 15 March of each year. Daily amounts of pelletized feed and sorghum were calculated as average daily amounts per fish (not including silver carp), weighted by each species' density and number of days of effective growth. The daily amount of manure applied is expressed as $kg \cdot 0.1\ ha^{-1} \cdot day^{-1}$. Four groups of ponds were established according to depth, which include ponds and reservoirs. These two categories differ in their water management scheme, since in most cases reservoir water is used for crop irrigation in summer and therefore depth decreases gradually (Hepher 1985). Whether silver carp or the hybrid "namsif" is stocked in the ponds is coded in form of the dummy variable NAMSIF, which has a value of zero or one, respectively.

Economical variables (prices, costs, etc.) were treated normatively, applying the prices of December 1988 to all ponds and years of the database. All economical variables are expressed in New Israeli Shekels (NIS) (NIS 2 = US$ 1). Variable costs were calculated according to the nutritional inputs given. Marketing costs are proportional to yield obtained and fixed costs (NIS $1{,}500 \cdot 0.1\ ha^{-1} \cdot year^{-1}$) depend on the number of days of the culture period. A special cost for water in the period March-November was estimated by differences in water balance (evaporation minus rain), amounting to NIS $30 \cdot 0.1\ ha^{-1}$ in August. The income for each pond was calculated from the yields of the different species and their prices. All these were used to calculate the variable PROFIT as:

PROFIT = income - (fixed costs
+ variable costs
+ marketing costs
+ special water cost).

The data are not uniformly distributed in the different levels of each variable, e.g., there are some years from which only a few observations are available while in some other years they amount to over 150. Also, two farms contributed only one observation each and one farm four observations, while four other farms contributed 131, 136, 138 and 160 observations. To reduce this problem, 13 farms with less than ten observations each were pooled into one "farm".

This database is used herein to analyze the effects of several independent variables on three dependent (target) variables:

1) total fish yield during the entire culture period, measured in kg·0.1 ha^{-1};
2) mean daily yield of fish, measured in kg·0.1 $ha^{-1}\cdot day^{-1}$;
3) profit per unit pond area for the entire culture period, measured in NIS.

The data were analyzed by multiple regression using the SAS (1985) procedure GLM (general linear model) according to the specifications given in the annex/supplement to this paper. The GLM procedure uses the method of least squares to fit general linear models such as regression, analyses of variance and covariance, and others. It can handle continuous as well as discrete variables. It was used herein to determine the regression equation of each dependent variable on several continuous and discrete independent (explanatory) variables.

For each dependent variable many models with various combinations of the independent variables were tested, starting with a model including all independent variables and eliminating nonsignificant variables in subsequent runs until one model was selected as the representative model. The results presented in this paper refer to these representative models, although the other models are also mentioned. The criteria used in the process of selecting the representative models are based on several parameters of the GLM outputs:

1) TYPE III SS (sum of squares for hypotheses which do not depend on the order in which the independent variables enter the model). The sum of squares associated with each variable of the model served as an approximation of the relative weight that this variable has in the total explanation of the model.
2) Pr > F (probability under F test). This indicates the significance level of the effect of each independent variable on the dependent variable. Variables with high values for this test were gradually omitted in successive runs, until only those with F values lower than 0.05 were kept ("significant run").
3) R^2 (coefficient of determination of multiple regression). This is an overall estimate of the success of the multiple regression model in explaining the variability in the dependent variable.
4) ESTIMATE. The estimates of the coefficients of the intercept and the independent variables in the multiple regression equation. They measure the effect of each independent variable on the dependent one in units of the latter. For discrete variables the GLM procedure sets the highest level to zero (reference level) and the values of the other levels are given as deviations from it, in units of the dependent variable. A model with intercept value close to the mean of the dependent variable was preferred.

Continuous independent variables were also treated in some models as discrete variables, to overcome their nonlinear nature. This was done by dividing the entire range of their values into three to seven levels, so that each level contained at least a few dozen observations. The variable was left discrete if this increased the R^2 value and if the ESTIMATE values of adjacent levels did not change sign or had extreme values as compared to the mean of the dependent variable.

Results

Table 1 presents the general parameters of the representative GLM models of total yield, mean daily yield and profit, which will be described in the following paragraphs.

Representative Model of Total Yield per Culture Period

Table 2 presents the estimates of the relative weights of the variables affecting the total fish yield per culture period and the significance levels of the effects. Total amounts of feed pellets applied was the main factor determining the total yield obtained, accounting for about 50% of the variability in yields (TYPE III SS, Table 2). Further identified factors, (and the amount of explained variability in yields) are: farm (12%), length of culture period (10%), amount of manure applied (7%), common carp density (6%) and tilapia density (4%). Other variables, i.e., tilapia, common carp and silver carp stocking weights, year, winter months, amount of sorghum applied and pond depth, contributed less than 2,5% each to the explanation of the model. The barely significant variables NAMSIF (which indicate whether silver carp or "namsif" was the filter feeder fish stocked) and silver carp density contributed less than 0.3% each. Altogether, added nutritional inputs (pellets,

Table 1. General parameters of the three GLM representative models analyzed.

Dependent variable	Units	Mean	Range		R^2	CV (%)	Root MSE
			min	max			
Total yield per culture period	kg·0.1 ha^{-1}	331	1	1530	0.872	24.5	81
Daily yield	kg·0.1 $ha^{-1}\cdot day^{-1}$	2.22	0.02	8.4	0.775	27.7	0.62
Profit per culture period	NIS·0.1 ha^{-1}	179	-1678	+2779	0.647	189	339

R^2= coefficient of determination
CV= coefficient of variation
Root MSE= square root of the mean square error

Table 2. Relative weights of independent variables in explaining total yield variability in the culture period.

Independent variable	Degrees of freedom	TYPE III SS (x 10,000)	P>F
Feed pellets	1	467	0.0000
Farm	18	111	0.0001
Length of culture period	5	96	0.0001
Manure	1	62	0.0001
Common carp density	1	55	0.0001
Tilapia density	1	37	0.0001
Tilapia stocking weight	4	22	0.0001
Year	11	21	0.0001
Winter months	1	19	0.0001
Sorghum	1	19	0.0001
Common carp stocking weight	1	10	0.0001
Silver carp stocking weight	1	10	0.0019
Pond depth	3	9	0.0036
Namsif (filter feeding fish)	1	3	0.0407
Silver carp density	1	2	0.0643
Total		943	

	Proportion of model explained (%)	Accumulated (%)
Added nutritional inputs	57	57
Natural food chain	13	70
Fish related variables	14	84

sorghum and manure) contributed about 60% to the model. The variables 'length of culture period', 'winter months' and 'pond depth' are interpreted as representing the contribution of the natural food chain in the pond, contributing together 13% to the model. Thus, feed factors were responsible for about 73% of the variability of the fish yields in the commercial ponds. The seven variables relating to the fishes themselves contributed only about 14% to the model (10% of them through fish densities).

The estimates of the effect of each independent variable on the representative model were as follows:

Intercept: 337 kg·0.1 ha^{-1}, which is very close to 331 kg·0.1 ha^{-1}, the mean total yield per culture period in the dataset. This represents the yield obtained if all the effects of the independent variables in the model are zero.

Variables: The regression coefficients and the units of the three added nutritional variables are feed pellets (0.280 kg·0.1 ha^{-1}), sorghum (0.126 kg·0.1 ha^{-1}) and manure (0.080 kg·0.1 ha^{-1}).

This means that for each 1.0 kg·0.1 ha^{-1} of feed pellets, sorghum or manure added during the culture period, fish yield was increased by 0.280, 0.126 or 0.080 kg·0.1 ha^{-1}, respectively, over the mean (intercept) total yield in the dataset.

The variable 'length of culture period' varied greatly, ranging from 25 to 506 days. In the representative model it was used as a discrete variable in order to reduce the extreme effect of a few ponds with very short or very long periods, which were grouped in the levels '50' and '375' days, respectively. The following effects were found:

Length of culture period (days)	Number of observations	Effect (kg·0.1 ha^{-1})
50	139	-205
100	359	-188
150	249	-141
200	125	-123
275 *	165	-60
375 **	86	0 (reference level)

*This level includes 96 observations of the '250' level and 69 of the '300'.

**This level includes 66 observations of the '350' level, 16 of the '400', one of the '450' and one of the '500' level.

In the range of 100-200 culture days (about 65% of the data) the contribution to total yield of each culture day after the 100th day is 0.65 kg·0.1 ha^{-1}·day^{-1} (the difference between their coefficients, 188-123 kg·0.1 ha^{-1}·day^{-1}, divided by 100 days).

'Pond depth' was also used as a discrete variable in the model, separated into four levels, resulting in:

Pond depth (m)	Number of observations	Effect (kg·0.1 ha^{-1})
1	866	-61
2.5	135	-52
3.5	51	-65
5	72	0 (reference level)

Consequently, the 5-m deep reservoirs yielded about 52-65 kg·0.1 ha^{-1} more than the shallower ponds. In further models tested, a positive effect of depth was also found for the 1-3.5 m levels.

The effect of the 'month' was tested by defining each month as a dummy variable, setting its value to 1 if the culture period included that month and 0 if not. No single month was found to affect fish yield significantly. Yet, when December, January and February were grouped into one variable (winter months), it was found significant in all runs, with effect coefficients varying from -12 to -25 kg·0.1 ha^{-1}. In the representative model, fish yield was reduced by 17 kg·0.1 ha^{-1} for each of these three winter months included in the culture period. July, not included in the representative model, had a significant effect in a considerable number of models, with positive coefficients between 10 and 23 kg·0.1 ha^{-1}. In some models, June and August also had positive significant effects on total fish yield (from 5 to 15 kg·0.1 ha^{-1}). It is reasonable to assume that the effects of months are due to differences in temperature and light, which affect the primary productivity and hence the natural food web, besides affecting fish metabolism.

Fish-related variables 'density' and 'stocking weight' of common carp, tilapia and silver carp entered the representative model and most of the other models as well. The density ranges and the estimated effect on total yield of stocking one more fish of each species are:

Species	Density effect (kg·0.1 ha^{-1})	Density parameters		Standard deviation
		Range	Mean	
Common carp	0.141	30-1,000	432	194
Silver carp	0.119	0- 500	24	47
Tilapia	0.045	0-3,760	498	647

Only a few observations were in the higher densities of each species and in those cases usually the densities of the other species were low. In models where density was used as a discrete variable, it was found that under 100 common carp·0.1 ha^{-1} the decrease in yield was sharper than the average decrease. The same occurred for tilapia density under 200·0.1 ha^{-1} and silver carp density under 20·0.1 ha^{-1}. Also, the marginal yield increase in the higher densities of the three species was somewhat more moderate than the average increase. It seems that the general nature of the response of total yield to increasing fish density reflects diminishing returns. The inclusion of the density variables as discrete ones did not increase R^2 markedly, probably due to the small number of observations in the extremes, which caused a bigger error in the mean of a density level with a wide range. Therefore, the density variables were left continuous in the representative model. A larger and more balanced set of observations may justify their inclusion as discrete variables.

Farms: the range of fish yield differences between farms was 196 kg·0.1 ha^{-1} according to the representative model (160-210 kg·0.1 ha^{-1} in other models). Part of the farm effect stems from climatic conditions associated with geographical region. Farms located in the hotter region (Bet She'an Valley) had about 50 kg·0.1 ha^{-1} higher yields than those in the cooler regions. Yet, in the same region a difference of 130 kg fish·0.1 ha^{-1} between two farms could be found. A second source accounting for regional differences is some selectivity in including ponds from the Bet She'an region when the database was built (see Milstein and Hulata, this vol.). Other possible sources for differences between farms may be factors not included in this study (e.g., aeration, water quality, parasites, predators, staff experience, etc.) or the balance in the variables included in the representative model.

Year: this was used as a discrete variable in the representative model. The effect of each year on total yield, taking 1987 as the reference, is as follows:

	Year					
	1976	'77	'78	'79	'80	'81
Effect (kg·0.1 ha^{-1}):	-145	-94	-57	-84	-84	-74
Number of observations:	3	5	15	74	159	201

	Year					
	1982	'83	'84	'85	'86	87
Effect (kg·0.1 ha^{-1}):	-72	-60	-52	-43	-41	0
Number of observations:	121	159	131	160	92	4

A clear general tendency of yield increase occurred from 1976 to 1987, except for 1978 in which total yield was about 30 kg·0.1 ha^{-1} higher than the adjacent years. Excluding the relatively steep increases between the first two and the last two years (having very few observations), the mean total yield increase from 1977 to 1986 is about 5 kg·0.1 ha^{-1}·$year^{-1}$. In other models, in which year was used as a continuous variable, its coefficient varied from 6 to 8.5 kg·0.1 ha^{-1}·$year^{-1}$. The tendency of increasing yields with time probably reflects technological improvements and other factors not explicitly included in the dataset, such as improvement of feed pellet composition and aeration practices, genetic improvement of fish, etc.

Some variables which were not included in the representative model had a significant effect in other models. Pond area entered some models with a negative effect: total yield decreased by 0.3-0.4 kg·0.1 ha^{-1} for each 0.1 ha of increased fishpond area. Yield of grass carp and wild spawning increased total yield in a few models.

Representative Model of Daily Yield

Table 3 gives the estimates of the relative weights of the variables affecting daily fish yield and the significance levels of these effects. As in the representative model of total yield, feed pellets (in this model on a daily basis) is the main factor determining the daily yield obtained, accounting for about 45% of the variability (TYPE III SS, Table 3). The other variables are: tilapia density

Table 3. Relative weights of independent variables in explaining daily yield variability.

Independent variable	Degrees of freedom	TYPE III SS (x 10,000)	P>F
Feed pellets per day	1	213	0.0000
Tilapia density	3	82	0.0001
Winter months	1	53	0.0001
Farm	18	52	0.0001
Length of culture period	5	26	0.0001
Common carp density	1	19	0.0001
July	1	11	0.0001
Common carp stocking weight	1	10	0.0001
Manure per day	2	9	0.0001
Year	1	3	0.0077
Silver carp density	1	1.5	0.0501
Total		476	

(about 17%), winter months (about 11%) and farm (about 11%). Total added nutritional inputs (pellets and manure) accounted for 47% of the variability of the model (as compared to 60% in the total yield model). No significant effect of sorghum was found. The variables reflecting the contribution of the natural food chain (length of culture period, culture periods including winter months or July) together accounted for 19% of the variance. Thus, supplemented and natural food explained about 60% of the daily yield variability, as compared to 73% of the total yield during the culture period. Of the four variables related to the fish species which entered the daily yield representative model, the densities (of tilapia, common carp and silver carp) together accounted for 23% of the variability and silver carp stocking weight accounted for another 2%.

The estimates of the effect of each independent variable on the daily yield representative model were as follows:

Intercept: 2.05 kg·0.1 $ha^{-1}·day^{-1}$, which is close to 2.22 kg·0.1 $ha^{-1}·day^{-1}$, the mean daily yield in the dataset.

Variables:

Feed pellets: for each 1.0 kg·0.1 $ha^{-1}·day^{-1}$ of pellets applied, an increase over the mean daily fish yield (intercept) of 0.232 kg·0.1 $ha^{-1}·day^{-1}$ is obtained (within the feed pellet range of 0-21.25 kg·0.1 $ha^{-1}·day^{-1}$).

Manure: this was used as discrete variable in the model.

Manure (kg·0.1 $ha^{-1}·day^{-1}$)	Number of observations	Effect (kg·0.1 $ha^{-1}·day^{-1}$)
0	580	-0.30
3	274	-0.25
10	270	0 (reference level)

Length of culture period: this had a negative effect on daily yield, which was larger the shorter the culture period was. There is a high correlation between length of culture period and winter months included. The direct effect of the winter months is already accounted for in an independent variable (reducing daily yield). Ponds with long culture periods included also more months with a relatively small contribution to daily yield, which while not having significant effects as individual months, when accumulated they did have an effect.

Length of culture period (days)	Number of observations	Effect (kg·0.1 ha^{-1})
50	139	0.74
100	359	0.41
150	249	0.27
200	125	0.17
275	165	0.29
375	86	0 (reference level)

Winter months: -0.25 (kg·0.1 $ha^{-1}·day^{-1}$)

July: 0.29 (kg·0.1 $ha^{-1}·day^{-1}$)

That is, each of the winter months (December to February) included in the culture period reduced daily yield by 0.25 kg·0.1 ha^{-1} and when the culture period included July the daily yield was increased by 0.29 kg·0.1 ha^{-1}. These values amount to about 12% of the mean daily yield, while the effects of months in the total yield model are only 5% of its mean value.

Tilapia density: this is the most important of the fish-related variables and was used as a discrete variable:

Tilapia density (fish·0.1 ha^{-1}) Group	Real range	Number of observations	Effect (kg·0.1 $ha^{-1}·day^{-1}$)
0	0-200	517	-0.93
400	200-600	219	-0.53
800	600-1,000	183	-0.31
2,000	1,000-3,500	205	0 (reference level)

The mean daily yield in ponds with less than 200 tilapia·0.1 ha^{-1} was 0.40 kg·0.1 $ha^{-1}·day^{-1}$ lower than in those with 200-600 tilapia·0.1 ha^{-1}, but in those around 800 tilapia·0.1 ha^{-1} it was only 0.31 kg·0.1 $ha^{-1}·day^{-1}$ lower than in those with 1,000-3,500 tilapia·0.1 ha^{-1}, showing a diminishing rate of return to the density.

Other fish-related variables: their effect on daily yield did not differ markedly from linearity, so they were included in the model as continuous variables. Their effects were:

Variable	Unit	Range	Effect (kg·0.1 $ha^{-1}·day^{-1}$)
Common carp density	fish·0.1 ha^{-1}	30-1,000	0.00078
Silver carp density	fish·0.1 ha^{-1}	0- 500	0.00081
Common carp stocking weight	g	0- 850	0.0070

Farms: the maximum difference in mean daily yield between farms was 0.9 kg·0.1 ha^{-1}·day^{-1}. Again, part of the farm effect stems from the geographical region, although not so clearly as in the total yield model.

Year: in this model this variable was included in its continuous form. During the period 1976-1987, in each year after the first one, mean daily yield increased by 0.026 kg·0.1 ha^{-1}·day^{-1}. There was an increase of 15% in the mean daily yield over those 12 years.

Representative Model of Profit per Culture Period

Table 4 presents the estimates of the relative weights of the variables affecting the profit of a unit pond area per culture period and the significance levels of the effects. All independent variables which affect yield were tested. The profit is expressed in New Israeli Shekels (NIS). The representative model accounts for only 65% of the profit variability among the 1,124 ponds, which ranges from NIS-1,687 to +2,779·0.1 ha^{-1}·culture $period^{-1}$ (Table 1). This is a low level of explanation as compared to the yield models. The relative weights of the independent variables included in the profit model are quite different from those in the yield models. Farm is the most important single variable here (22% of the TYPE III SS). The variables related to the fish (tilapia and common carp densities, common carp and silver carp stocking weights) accounted for about 31% of the variance, while nutritional inputs (pellets and manure) together account for another 17%.

Table 4. Relative weights of independent variables in explaining profit variability.

Independent variable	Degrees of freedom	TYPE III SS (x 1,000,000)	P>F
Farm	18	21.9	0.0001
Tilapia density	3	16.7	0.0001
Feed pellets	4	13.3	0.0000
Winter months	1	13.0	0.0001
Common carp density	4	11.7	0.0001
Length of culture period	5	10.4	0.0001
Year	11	6.0	0.0001
Manure	4	4.2	0.0001
Common carp stocking weight	1	1.7	0.0001
Grass carp yield	1	1.0	0.0035
Silver carp stocking weight	1	0.7	0.0162
Total		100.6	

The estimates of the effects of each independent variable on the profit representative model were as follows:

Intercept: NIS 181·0.1 ha^{-1}, which is very close to NIS 179·0.1 ha^{-1}, the mean profit in the dataset.

Farms: the maximum difference between farms was NIS 954·0.1 ha^{-1}.

Year: profit showed a general tendency to increase over the years. The lowest profit (in 1977) was NIS 847·0.1 ha^{-1} lower than in 1987, the reference year.

Other variables: Table 5 shows the effects of the continuous variables in the profit model and Table 6 those of the discrete variables which have a clear physical or biological meaning. In the latter, the level zero of pellets, manure and tilapia density included many cases where the absolute values were zero, when no feed inputs were applied or no tilapia were stocked and also values up to 200 units of each variable.

In an attempt to increase the explained part of profit variability between ponds, the residuals from the model (difference between the values observed and those predicted by the model) were calculated and analyzed. Significant correlations were found between the residuals and the fixed costs ($r = -0.11$, $P < 0.0001$) or income ($r = +0.33$, $P < 0.0001$). Significant differences between the residuals for certain ponds were found in each farm where most or all the residuals over years were always positive for one pond and always negative for another pond. "Correcting" profit values by the average value of the residuals for such ponds (reversing the sign) before running the GLM procedure, increased R^2 from 0.647 in the representative model to 0.686 in the "corrected" one. Time relations for residuals were also found in a considerable number of farms: some farms showed a high stability (small residuals) while in others residuals increased or decreased with the years. Some interactions were tested as well. Highly significant interactions were found between pairs of factors, which increased by 10% the contribution of each such pair to TYPE III SS of the model as compared to the contribution of the two factors themselves. However, their contribution to an increase of the explanatory level of the model (R^2) did not exceed 1 to 2% and therefore they were not included in the representative model for profit.

Table 5. Estimates of the effects of continuous variables on profit (n=1,124 ponds). (Profit range: NIS -1,678 to +2,779·0.1 ha^{-1}·culture period; weight of fish at stocking).

Independent variable	Coefficient (NIS·0.1 ha)	Unit	Range of variable
Winter months	-132	month	0 - 3
Common carp weight	-0.294	g	0.7 - 850
Grass carp yield	5.52	kg·0.1 ha^{-1}	0 - 86
Silver carp weight	-0.058	g	0 - 3,500

Table 6. Estimates of the effects of categorical variables on profit (n= 1,124 ponds). (m = number of observations in each level; * = reference level; Profit range: NIS -1,678 to +2,779·0.1 ha^{-1}·culture period^{-1}).

Independent variable				Coefficient (NIS·0.1 ha^{-1})
Name	Unit	Level	m	
Feed pellets	kg·0.1 ha^{-1}	0*	330	0
		400	515	10
		800	209	184
		1,200	55	475
		2,000	15	710
Manure	kg·0.1 ha^{-1}	0*	675	0
		400	225	94
		800	128	93
		1,200	62	204
		1,600	37	406
Carp density	fish·0.1 ha^{-1}	0*	4	0
		150	133	180
		300	340	332
		450	365	447
		600	282	561
Tilapia density	fish·0.1 ha^{-1}	0*	517	0
		400	219	249
		800	183	317
		2,000	205	401
Length of culture period	no. of days	50	139	245
		100	359	241
		150	249	336
		200	125	281
		275	165	-44
		375*	87	0

Discussion and Conclusions

The polyculture fishpond is a complex biological and economic system. Each of its target variables is simultaneously affected by many factors, part of which were accounted for in the analyses presented herein. Table 1 shows that the highest level of explanation was obtained for the total yield model and the lowest for the profit model. The low explanatory level of the profit model reflects the higher complexity of this variable, which is affected by economic, as well as human factors, besides biological and physical ones. The predictive value of each model is related to its explanatory level of past data: the closer its R^2 value is to 1.0, the higher the probability that the coefficients of the effects of the different independent variables on the target variable are actually near the true coefficients, to be repeated in the future in the same population. In the representative model of total yield, the coefficients of the main variables may be considered acceptable. In the profit model the coefficients should only be considered as indications for directions to follow.

This leads to the conclusion that more efforts are needed to improve the models. Several procedures may be tried in order to increase explanatory power:

- a thorough system analysis may reveal more factors affecting the target variables, such as aeration, water quality, sanitation, genetics, etc. These data should be added to the existing dataset and the models run again;
- a more detailed characterization of important independent variables, such as different feed pellet types or feeding methods (automatic or demand feeders) may be entered in the models as separate variables;
- new variables may be created from the ones existing in the dataset, which will enable the detection of significant effects that are insignificant in their present form. An example to this is the variable "winter months" already discussed;
- inclusion of interactions between variables. This is possible only in a large and diversified database;
- variables like "year" or "farm", included in the model to reduce error (like blocks in an experiment), cannot contribute to prediction. However, it is possible to attribute biological and/or economic factors to these variables generally by studying the extreme differences between their levels: and

- the study of the residuals of the model in relation to additional variables, like tracing certain ponds which are systematically "good" or "poor", may enable finding the biological or economical characteristics underlying the statistical relationship.

Variability in productivity among farms and regions is a remarkable feature of the Israeli fish culture industry (Sarig 1990). These differences can be attributed to various climatic and management variables. The integrative models presented herein reflect the complexity of the interrelationships among these variables in the polyculture ponds and their effects on yield and profit, better than univariate analyses, as will be exemplified in this and the following paragraphs. Table 7 presents the coefficients of the effects of "year" and "farm" in the representative models for profit and total yield, together with the corresponding averages of these variables and the main independent variables. The "reference level" for the differences between levels in the models are framed in the table. Instead of using deviations from zero (the value in the GLM model) it shows absolute values obtained by adding the average in that year or farm to each level. For example, in the profit model the reference year 1987 got the value 990, which is the simple profit average of that year; 990 was added to the model value in 1986 (-400), so that in comparison to 1987 its value is -400 + 990 = 590. The same procedure was applied to all levels. The table shows an increase from 1976 to 1987 in profit (model and average), total yield (model and average) and the independent variables pellets, manure and tilapia density. However, almost all variables are interrelated, often with very high and significant correlation coefficients (Table 8). In certain cases it seems apparently possible to explain the value of a target variable by means of the values of the independent variables. For example, farm 24 has the highest values for profit (Table 7), and also the highest values of pellets and manure while those of carp and tilapia densities are among the highest. However, the daily yield in this farm is lower than the average in the dataset and the value of winter months (negatively affecting the target variables) is the highest among all farms. Farm 33, with the lowest average profit, has the lowest values of pellets and daily yields, and low total yield, manure and tilapia density but the values of other parameters are higher than the average. Thus, it is difficult to explain the differences in target variable values by separate comparisons with each of the independent variables. A farm applying in all its ponds a higher level of a certain variable, compared to another which is at a low level, will not obtain the predicted yield effect by the variable in excess and will have a negative farm effect relative to all farms and especially to a balanced farm.

Table 7 also shows differences between the values of profit and total yield in the same year when calculated by a simple average and when calculated through the integrative GLM model. For example, from 1978 to 1986 (excluding years with few observations) the profit increased by NIS 649·0.1 ha^{-1} according to the simple average, but by only NIS 187·0.1 ha^{-1} according to the integrative model, while the respective values for total yield are 262 and 16 kg·0.1 ha^{-1}. This occurs since the simple average assigns to one discrete variable (year) all the effects of all other variables correlated to it, while the integrative model separates the effects and assigns to each variable only its share.

A similar situation occurs with farms. For example, farm 11 has a model coefficient of NIS 490·0.1 ha^{-1} for profit and an average profit of NIS -45·0.1 ha^{-1} while the respective numbers for farm 53 are NIS 463 and 666·0.1 ha^{-1}. The relative high values of the average profit may be explained by the utilization in that farm of relatively high levels of the variables positively correlated with profit and low levels of those negatively correlated with it and the opposite for low average profit values. On the other hand, relatively high values of the model coefficient point to a better balance in the selection of the levels of the different variables.

Table 9 presents the distribution of several variables according to the interaction 'pellets X year' and the corresponding distribution by pellet levels. All extreme differences between subgroups in the interaction section (a) are larger than those in the one variable section (b). Namely, the division into subgroups increased the explanation level of the variance of the target variables. The interaction section of the table also allows tracing changes with time: in 1981 the differences between the groups '0' and '1,200' kg pellets·0.1 ha^{-1} amount to about NIS 400·0.1 ha^{-1} profit and about 350 kg·0.1 ha^{-1} fish yield (the latter also in 1980), while in 1984-1986 they increased to about NIS 1,100·0.1 ha^{-1} and 650 kg·0.1 ha^{-1}, respectively. However, in each year entirely different conclusions may be extracted in relation to the effects of 400, 800 and 1,200 kg pellets·0.1 ha^{-1} on yield and

Table 7. Coefficients of the effects of "year" and "farm" in the representative models for total profit and yield, the corresponding averages of these variables, and the averages of the main independent variables. "Reference levels" for the differences between levels in the models are framed.

	No. of obs.	Profit (NIS·0.1 ha^{-1})		Total yield (kg·0.1 ha^{-1})		Daily yield (kg·0.1 ha^{-1}·day^{-1})	Pellets (kg·0.1 ha^{-1})	Manure (kg·0.1 ha^{-1})	Sorghum (kg·0.1 ha^{-1})	Days	Winter months	Density (fish·0.1 ha^{-1})			Stocking weight (g)			Depth (m)
		model	average	model	average							Common carp	Tilapia	Silver carp	Common carp	Tilapia	Silver carp	
YEAR																		
1976	3	252	-417	415	111	0.70	188	0	56	160	1.66	436	236	53	358	166	969	2.50
1977	5	149	-704	466	186	0.81	210	0	191	219	2.20	398	129	48	111	31	338	1.60
1978	15	403	-217	503	147	0.94	123	53	108	151	1.26	341	179	22	280	23	255	1.76
1979	74	370	86	476	239	2.21	355	118	94	124	0.77	425	470	15	220	56	256	1.23
1980	159	364	-6	476	273	1.82	315	225	142	163	1.08	388	377	37	216	50	239	1.67
1981	201	373	90	486	271	2.16	338	351	112	138	0.82	418	470	19	217	65	297	1.25
1982	121	400	165	488	293	2.12	363	205	163	146	0.76	431	514	12	232	40	225	1.24
1983	159	482	210	500	336	2.22	429	187	128	161	0.92	450	476	32	212	56	215	1.48
1984	131	471	210	508	413	2.45	538	380	123	191	1.22	435	647	22	189	65	308	1.89
1985	160	538	382	517	436	2.53	495	514	136	185	0.88	453	566	25	179	68	286	1.92
1986	92	590	432	519	409	2.55	565	221	54	169	0.71	488	548	20	196	53	199	1.49
1987	4	990	990	560	560	4.40	861	493	40	136	0.25	313	671	7	276	123	675	2.50
FARM																		
0	40	558	48	251	513	1.68	365	598	166	308	2.45	402	281	33	57	63	456	4.57
11	18	490	-45	250	366	1.66	455	327	70	226	1.44	347	37	74	41	13	434	2.90
12	27	476	432	293	531	2.48	665	848	41	230	1.11	539	135	18	54	3	113	1.92
14	94	76	107	156	266	2.40	412	604	0	126	0.68	460	584	21	275	91	200	1.11
15	25	342	46	235	255	1.87	359	260	17	140	0.80	331	496	1	219	44	4	1.50
16	138	435	393	251	316	2.86	357	164	71	112	0.16	432	683	26	232	90	466	1.03
17	131	347	292	235	342	2.70	413	710	20	144	0.67	386	729	5	215	123	64	1.37
18	34	221	101	202	251	2.02	337	462	48	134	0.55	479	367	0	217	51	0	1.07
19	15	229	279	196	316	2.71	449	540	34	122	0.40	444	952	0	127	87	0	1.16
21	11	565	242	270	552	2.04	437	653	143	275	1.54	259	548	34	37	93	263	2.45
24	18	930	746	353	764	2.14	763	1,109	193	339	2.50	507	781	32	60	91	507	3.13
31	90	428	186	231	381	2.29	623	4	128	191	0.70	493	28	33	102	3	138	1.00
32	19	310	127	214	484	2.06	627	439	61	256	1.89	472	650	18	185	82	332	2.50
33	160	223	-233	191	191	1.22	216	17	202	159	1.14	451	205	33	285	24	516	1.88
41	96	366	139	214	256	2.06	343	47	226	124	0.90	394	617	23	212	24	153	1.00
42	136	490	153	244	299	1.90	373	131	223	152	1.20	388	387	34	246	37	251	1.00
43	24	722	704	294	622	2.92	675	87	203	240	1.41	516	884	22	113	47	85	4.35
53	17	463	666	285	553	3.48	745	376	134	177	0.58	407	777	22	195	74	428	2.64
64	31	547	547	499	440	3.70	705	3	153	133	0.83	520	1,483	2	313	77	45	1.00
Average of the 1,124 obs.			179		331	2.22	415	289	124	161	0.92	432	498	24	209	58	262	1.55

Table 8. Pearson correlation coefficients among selected variables (n = 1,124 observations). Numbers in the column headings correspond to variables in the row headings. The figure in each cell is the correlation coefficient and the stars below are the significance level (** = 0.01, *** = 0.001). Zero or nonsignificant coefficients were omitted.

	2	3	4	5	6	7	8	9	10	11	12	13
1 Profit	.70 ***	.77 ***	.50 ***	.22 ***			-.42 ***	.32 ***	.40 ***	-.23 ***	.14 ***	
2 Total yield		.54 ***	.79 ***	.39 ***	.18 ***	.60 ***		.38 ***	.33 ***	-.51 ***		.35 ***
3 Daily yield			.53 ***	.15 ***	-.10 ***	-.24 ***	-.50 ***	.24 ***	.53 ***	-.09 **	.32 ***	-.08 **
4 Pellet				.23 ***		.39 ***		.41 ***	.27 ***	-.33 ***	.07 *	.14 ***
5 Manure					-.20 ***	.29 ***			.13 ***	-.18 ***	.15 ***	.17 ***
6 Sorghum						.26 ***	.12 ***	.15 ***		-.17 ***	-.22 ***	.08 **
7 Days							.61 ***	.17 ***	-.11 ***	-.51 ***	-.19 ***	.50 ***
8 Winter months								-.19 ***	-.08 **	-.16 ***	.30 ***	
9 Carp density								-.11 ***	-.23 ***	-.10 **	.09 **	
10 Tilapia density										.28 ***		
11 Carp weight										.25 ***	-.23 ***	
12 Tilapia weight												
13 Depth												

profit, which points to the difficulty of explaining the values of target variables by only one or two independent variables, and to the low predictive value of such analyses.

Fish farmers and extension officers should have a special interest in the integrative analysis over all the ponds and in the specific results of each farm and pond. On one hand, these focus their attention to specific problems (deviations from the averages) and can aid in improving their production. On the other hand, by studying the extreme deviations, they can trace and identify variables not included in the integrative analysis and improve the model by adding them to the analysis.

A specific problem of interest to the farmers is shown by analyzing the number of observations in each 'pellet X year' subgroup in Table 9a. There is a slow tendency of an increase in the numbers of ponds through the years in the 800 kg·0.1 ha^{-1} and higher levels of pellets and of a decrease in the numbers of ponds in the zero level. However, even in the latter years, most of the ponds belong to the low "400" level and no more than 10% to the higher level, despite the high correlation between pellets and total yield and profit. This shows that the highest potential of yield and profit related to pellets is actually obtained only in a few ponds. The same holds for other independent variables highly correlated with the target variables (GLM models and Tables 7, 8, 9). Few are the ponds which received high levels of manure and high densities of carp and tilapia and still fewer those with high values of all together. Indeed, the percentage of ponds with very high yields and profit is low.

Table 9a. Means of target and major explanatory variables according to pellet level and year (n= no. of observations in each group).

Pellet level (kg· 0.1 ha^{-1})	Year	n	Profit (NIS· 0.1 ha^{-1})	Total yield (kg· 0.1 ha^{-1})	Tilapia density (fish· 0.1 ha^{-1})	Carp density (fish· 0.1 ha^{-1})	Manure (kg· 0.1 ha^{-1})	Carp weight (g)
0		41	-292	119	158	319	63	258
400	1976	43	3	238	453	454	69	216
800	to	9	458	403	1,129	568	230	233
1,200	1979	2	618	471	0	698	185	54
1,600		2	1,121	755	1,365	352	757	25
0		66	-214	140	216	327	121	285
400	1980	66	94	296	476	372	312	205
800		23	362	546	603	558	316	75
1,200		4	-359	499	112	623	0	53
0		68	-124	134	237	359	104	261
400	1981	100	197	293	571	429	447	216
800		28	194	484	737	491	609	142
1,200		5	305	505	111	596	357	79
0		53	-106	136	264	338	81	308
400	1982	39	281	339	629	455	350	208
800		23	573	493	988	538	273	134
1,200		6	258	627	148	695	85	87
0		46	-82	140	260	338	61	290
400	1983	74	179	318	520	462	210	219
800		24	390	524	576	506	181	121
1,200		11	758	602	728	589	64	107
1,600		4	1,586	1,074	846	808	1603	46
0		17	-167	194	220	396	146	324
400	1984	68	15	325	615	359	388	209
800		32	419	526	585	555	493	125
1,200		9	1,033	832	856	689	396	51
1,600		5	1,316	876	2,541	384	314	120
0		26	-33	170	223	408	289	277
400	1985	85	249	366	576	389	551	198
800		34	697	649	598	569	624	111
1,200		13	1,149	832	835	639	507	60
1,600		2	1,137	729	2,283	601	0	42
0		13	79	155	451	409	0	187
400	1986	40	374	322	628	476	169	251
800	and	36	515	518	532	447	375	167
1,200	1987	5	1,446	934	598	758	415	109
1,600		2	974	842	0	968	0	60

Table 9b. Means of target and major explanatory variables according to pellet level (n = number of observations in each group).

Pellet level (kg· 0.1 ha^{-1})	n	Profit (NIS· 0.1 ha^{-1})	Total yield (kg· 0.1 ha^{-1})	Tilapia density (fish· 0.1 ha^{-1})	Carp density (fish 0.1 ha^{-1})	Manure (kg 0.1 ha^{-1})	Carp weight (g)
0	330	-134	141	237	349	105	279
400	515	170	314	557	418	345	213
800	209	460	532	667	523	419	132
1,200	55	776	706	572	655	283	76
1,600	15	1,293	888	1,559	599	633	69

The study of extreme deviations shows, for example, that the maximum value of yield is four times higher than its average and the maximum profit is 15 times higher than its average. There are limitations to reach the maximum possible yield and profit, some of which are difficult to overcome, like climatic conditions. Winter months yield less than other months, reducing the potential yield by up to 50-100 kg·0.1 ha^{-1} and the profit by NIS 400·0.1 ha^{-1} or more (Table 5). Being aware of this limitation and having a good estimate of its effect, it may be possible to minimize it by applying existing know-how (like genetic differences in cold sensivity of fish species) or by direct research to treat the limitation.

Fish losses are a limitation not accounted for in the present database since it does not include ponds with losses higher than 20%. Characterization of this factor in the data, and a more representative sample, will enable to include it in further analysis, identify the magnitude and causes of the problem, and approach the potentials of yield and profit.

Marketing may also be an important limitation. While in growout ponds the farmer's target is to maximize yield and profit, in ponds under marketing limitations the target may be to minimize losses. The characterization of such ponds (when and how long they are under such limitations) should enable to quantify their influence on yield and profit and may indicate the need of a different management practice. It is reasonable to assume that not all decisions made by farmers were optimal for such cases. Water quality and its manipulation are other examples of factors not accounted for in the present study that may also limit yield and profit.

GLM analyses on a variety of independent variables affecting yield and profit of several systems were carried out in Israel, including experimental fishponds (Milstein et al., this vol.), commercial and experimental wheat fields (Ziv and Goldman 1987a, 1987b, 1987c) and on forest fires (Zohar et al. 1988). A quantitative and integrative analysis of results of many experiments and from different sources is developing lately as a special statistical discipline called "meta-analysis" (Sacks et al. 1987). Although the abovementioned GLM analyses are not exactly meta-analysis, they likewise deal with the integration of a large volume of diversified information from different sets of conditions, and their quantitative treatment considering many variables related to the target variable.

The above discussion treated part of the problems related to fish culture and the possibilities of handling them and does not claim to cover the entire subject. What we hoped to show was that the vast information available on such a complex system deserves a complex integrative treatment, which is more appropriate than a simple one-way or unifactorial analysis. The potential of computers enables nowadays to extract conclusions from the data, which are more applicable to the farmers, extension officers and researchers. The more we engage in such integrative treatment, the more improved will be the tools for management, the more qualified the people dealing with it, and the higher the returns to everybody involved.

Acknowledgement

This study was supported 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 for Economic Cooperation (BMZ).

References

Hepher, B. 1985. Aquaculture intensification under land and water limitations. GeoJournal 10(3): 253-259.

Sarig, S. 1990. The fish culture industry in Israel in 1989. Isr. J. Aquacult./Bamidgeh 42:39-45.

Sacks, H.S., J. Berrier, D. Reitman, V.A. Ancona-Berk and T.C. Chalmers. 1987. Meta-analyses of randomized controlled trials. Special article. N. Engl. J. Med. 316(8):451-455.

SAS. 1985. SAS user's guide: statistics. Version 5 edition. Statistical Analysis Systems Institute, Inc., Cary, North Carolina. 957 p.

Simmonds, N.W. 1980. Comparisons of the yields of four potato varieties in trials and in agriculture. Exp. Agric. 16:393-398.

Ziv, O. and A. Goldman. 1987a. General lessons from wheat survey: a. Nitrogen fertilization. Hasadeh 67:1548-1549. (In Hebrew).

Ziv, O. and A. Goldman. 1987b. General lessons from wheat survey: b. Weed control. Hasadeh 67:2460-2463. (In Hebrew).

Ziv, O. and A. Goldman. 1987c. A comparison of methods for statistical analysis of factors affecting wheat crops in Israel. Isr. Agresearch 1(2):53-66. (In Hebrew).

Zohar, Y., A. Weinstein, A. Goldman and A. Genizi. 1988. Fire behaviour in conifer plantations in Israel. Forêt Mediterranée 2:423-426.

Annex

Specifications for an Integrative Analysis Method

The aim of an integrative regression analysis is to define a model that efficiently and validly explains the target variable Y (e.g., fish yield or profit), in terms of a set of explanatory variables, X_1, X_2 ... X_n (e.g., fish density, nutritional inputs, water quality, etc.). The following specifications must be met in order to conduct an efficient analysis of such systems:

1. *Data*

1.1. *Size of the database*: The more complex the system, the bigger and the more diversified is the required database. The program must be capable of handling with ease large databases containing many (hundreds, thousands) of observations and many types of variables.

1.2. *Data distribution*: Data are often unequally distributed between variables, combinations of variables and different levels of discrete variables. The program must be able to handle such heterogeneity.

1.3. *Missing data*: In complex datasets, an observation is composed of many variables. Some statistical analysis programs discard every observation for which not all variables are present, losing information on the remaining variables that have been recorded. An efficient program should not lose such information.

1.4. *Errors and outliers*: In complex datasets, errors are inevitable. The program should be able to detect and identify data points that are highly deviant.

2. *Explanatory variables*

2.1. *Discrete and continuous variables*: Some variables are discrete (e.g., farm, pond, fish species, feed type); others are continuous (e.g., amount of inputs, temperature). The program should be able to handle both types of variables simultaneously.

2.2. *Linear and nonlinear effects*: Many relationships between variables and outputs are nonlinear and should be treated as such. Otherwise distorted models will be the result. The program should be able to handle both linear and nonlinear effects of continuous variables.

2.3. *Correlations between explanatory variables*: Confounded effects and collinearity is a common problem in multifactorial analysis and cannot be ignored. An appropriate program should be able to detect collinearity between correlated variables and be able to separate the specific individual effects among them.

2.4. *Interactions*: These can occur both between and within continuous and discrete variables and the program should be capable of handling both.

3. *Output*

3.1. *General indices*: Certain indices are standard for any analysis of variance: coefficient of determination (R^2), residual standard deviation (RSD), coefficient of variation (CV), degrees of freedom for model and error. A good explanatory model will have both high R^2 and low RSD (or CV).

3.2. *Significance*: The level of significance of a variable in a multifactorial analysis is important to screen the many independent variables which may affect the target variable. Only those with a significant effect on the target variable are allowed to enter the model. The level of significance calculated to test the marginal effect of each variable in relation to all the other variables is the most appropriate.

3.3. *Relative weight*: An index of relative weight of each independent variable in the model is an indication of the attention it deserves.

3.4. *Distribution of observations*: The distribution of observations within each class level or around a continuous variable should be explicitly indicated.

4. *Model selection*: A model is a combination of variables that together explain the variability of a target variable. When there are many variables, there are many possible combinations of variables and of interactions between them. Choosing an efficient model then becomes a matter of skill and judgement. The process of choosing the most efficient model is not an intrinsic part of existing analysis packages.

The General Linear Model (GLM)

The general linear model is the basis of numerous statistical analysis packages. One that we have found very appropriate for our purposes is the GLM procedure of the SAS (1985) package, which is available for mainframe and personal computers. This procedure meets many of the specifications listed above: it can handle large databases (1.1), non-uniform data distribution between explanatory variables (1.2), discrete and continuous variables (2.1), correlations between explanatory variables (2.3), interactions (2.4) and it has all the standard statistical indices (3.1; 3.2). It also has deficiencies: it does not explicitly handle errors and outliers (1.4), nonlinear effects (2.2), or distribution of variables (3.4). It also does not choose the appropriate model for the specific dataset (4). Some of these deficiencies can be overcome by a suitable protocol. Others require the development of supplementary software. Some of the solutions that have been found useful are presented below.

Missing data (1.3). When missing values on one or a small number of variables result in the elimination of an observation for which there is much information on other variables, it is often worth the effort to try to rescue that observation with an estimate of the missing value. The simplest estimate for a continuous variable is either an average or median value. Missing values for a particular discrete variable can be defined as a separate class. When there are many missing values for a continuous variable, it is often best to eliminate the variable and check its effect on the residuals at a later stage in the analysis.

Replacing missing values with estimates does not add information, it only helps to use other information that would otherwise be lost. In order to account for the fact that no extra information or degrees of freedom have been added, some adjustment on the degrees of freedom for error should be made. Where missing values are defined as a class, no adjustment of degrees of freedom for error is necessary.

Nonlinearity (2.2). Where the nature of the functional response of the target variable Y to an independent variable X is known to be nonlinear, appropriate transformations can be imposed on X (square, log, power, etc.). As a rule, the nature of the functional relationship is not known, but where nonlinearity is suspected, it can be taken into account by dividing the continuous independent variable into several (generally between 3 to 7) discrete levels. The nature of the nonlinearity will then be apparent from the regression coefficients for the series of levels; if there is a monotonous increase or decrease of the coefficients through the successive levels the relationship is linear, and the independent variable can be kept as continuous in the model; otherwise it is nonlinear and the variable should be included as discrete in the model.

Correlations between independent variables (2.3). The capability of GLM to handle correlations was tested by simulation. For continuous variables, the true coefficients were obtained both in a model with two highly correlated independent variables and in a model in which uncorrelated independent variables were added to the system. For discrete variables, GLM sets missing values in the coefficient estimates of confounded levels (when all observations in a certain level of a variable are in one level of another variable and vice versa).

Index of the relative weight of the effects of the independent variables (3.3). A practical but not strictly statistical procedure to measure the relative weight of the effect of each independent variable on the target variable may be derived from the sum of squares calculated to test the marginal effect of each variable (TYPE III SS in SAS terminology). The TYPE III SS of each variable is determined while considering simultaneously the effects of all other independent variables in the model. Thus, a variable with a high value of this parameter contributes more to the explanatory level of the entire model (R^2) than one with a lower value. The index SX_i is obtained considering that the sum of the TYPE III SS of all independent variables in the model is the 100% reference level and calculating the per cent contribution of each individual TYPE III SS (SS_i) to it:

$$SX_i = \frac{\text{TYPE III } SS_i}{\text{Total TYPE III SS}} \times 100$$

Model selection (4). For the target variable considered, several models with various combinations of the independent variables are tested. The procedure starts with a model including all independent variables. In the subsequent runs linearity of suspected variables is checked and nonsignificant variables are eliminated, until a model with high R^2, low RSD and all independent variables with significant effect on the target variable is obtained. The criteria used in this selection process are based on common sense and understanding of the system and on the statistical indices and considerations discussed below.

Interpretation of the model. The estimates of the coefficients of the intercept and the independent variables in the multiple regression equation obtained measure the effect of each independent variable on the target variable in units of the latter. For discrete variables the GLM procedure sets the highest value to zero (reference level) and the values of the other levels are given as deviations from it. In a good model the intercept value should be close to the mean of the target variable since the intercept represents the value of the target variable if all the effects of the independent variables in the model are zero.

APPLICATION OF THE MODEL

The GLM model is provided, e.g., within the SAS software package (SAS 1985) to identify and quantify the simultaneous effects of several independent variables on one target variable. Applied to data originating from farms, this method can be used as a powerful management tool. After collecting an adequate amount of data and obtaining a first model, estimates of the effects of several elements in the system are obtained and predictions based on them can be tried. This first model will not account for all the variability of the target variable (R^2 smaller than 1), and to improve it, new variables and new data must be added. Running the model again including the added information will result in new conclusions, with this iterating process leading to improved production and increased knowledge of the system.

Multiple Regression and Path Analysis of Tilapia Growth in Commercial Fish Farms in Israel*

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

ANA MILSTEIN, *Agricultural Research Organization, Fish and Aquaculture Research Station, Dor, Mobile Post Hof Hacarmel 30820, Israel*

Prein, M. and A. Milstein. 1993. Multiple regression and path analysis of tilapia growth in commercial fish farms in Israel, p. 178-188. *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 methods for the multivariate analysis of aquaculture experiments (the extended Gulland-and-Holt plot and path analysis) were used to analyze the growth of tilapia hybrids (*Oreochromis niloticus* x *O. aureus*) in commercial fish farms in Israel. The aim was to identify key variables governing hybrid tilapia growth in intensive polyculture ponds at high stocking densities, receiving manure, sorghum and pellet feed and aeration. Data analyzed were from fish farms from three fish culture regions in Israel: the Western Galilee, the Upper Galilee and the Bet She'an Valley. The extended Gulland-and-Holt method identified stocking density and water temperature (or solar radiation) as the main external variables controlling hybrid tilapia growth. Manure and feed applications were nonsignificant variables as these did not vary in a range adequate to experimental practice. The multiple regressions that were obtained are further described with path analysis; also, the parameters L_∞ and K of the von Bertalanffy growth function were estimated for a range of environmental conditions. Benefits and limitations of the application of multivariate methods to non-experimental data are discussed.

Introduction

The usual approach to the analysis of fish production in aquaculture operations is to conduct experiments at appropriate research facilities. Here, the necessary measures for controlling and maintaining the experimental design are given. The results obtained should then straightforwardly answer the questions initially formulated and provide insights which ultimately will improve fish production. Unfortunately, the difference in scale, level of control and management between experimental pond facilities and commercially producing fish farms are considerable. The individual contribution of the variables governing fish growth in experimental ponds of say 200 m^2 may differ from those affecting fish growth in and commercial ponds of several hectares.

It is therefore of interest to study details of fish production in systems operating on a commercial scale. Unfortunately, the dimensions and managerial procedures, among others, do not facilitate data collection largely at a level of quality sufficient to meet standards of scientific experimentation.

In the case of good recording standard of pond input and output data and managerial actions in a larger number of fish farms, a different approach may be taken. Based on the reasoning that a high number of cases, in spite of a larger amount of variance prevalent in commercial farm data, may provide additional insights into these systems by analyzing these data with appropriate methods. One such possibility is presented in the present paper, in which data from commercial fish farms in Israel are analyzed with multiple regression (here based on the extended Gulland-and-Holt plot) and path analysis (Prein, this vol.). These methods can extract information from datasets containing a larger amount of variance, simultaneously considering the interrelationships of several influential variables. The commercial fish farms in Israel have a good standard of data recording. It was possible to obtain these data from a number of different farms from three of the four fish culture regions in Israel. The present contributions present the application of the multivariate analysis methods mentioned above to these data.

Materials and Methods

In Israeli production farms, fish are generally grown in polyculture and the species composition varies between farms and even between the ponds of one farm (Hepher and Pruginin 1982; Hepher

*ICLARM Contribution No. 735

1985). Generally, three species are always present, although in varying proportions: common carp (*Cyprinus carpio*), all-male tilapia hybrid [*O. aureus* male x *O. niloticus* female; sometimes sex inversed (Hulata 1988)] and silver carp (*Hypophthalmichthys molitrix*). Further species cultured are bighead carp (*Aristichthys nobilis*), a hybrid between bighead and silver carp locally called "namsif", grey mullet (*Mugil* sp.), and the Malaysian freshwater prawn (*Macrobrachium rosenbergii*). Common carp constitutes 60% of the Israeli fish production followed by tilapia hybrids (20-30%) (Sarig 1990). Tilapia are stocked at sizes ranging from 20 g to 300 g or larger. The preferred market size for tilapia in Israel is over 400 g (G. Hulata, pers. comm.).

Due to a marked seasonality, with winter water temperatures sometimes dropping to 10°C or lower, the growing season in Israel ranges from mid-March to mid-November (8 months). Tilapia are stocked into production ponds when water temperatures are above 17°C. Thus, their culture period is shorter than that of carp and mullet. Tilapia are overwintered in separate facilities such as greenhouse-covered ponds or deepwater ponds (Hepher and Pruginin 1981; Hepher 1985).

All farms apply feed with automatic feeders at moderate to high levels and provide different types of pellets. The only other nutrient input is in form of dry chicken manure. This is applied in amounts of 100 to 150 $kg \cdot ha^{-1} \cdot day^{-1}$ up to April or May, thereafter at intervals and rates depending on pond conditions.

Culture cycles may range over the entire season, with several partial harvests to market the large fish, or two or three short cycles with total draining of the ponds, sorting of fish and restocking. Individual management practices vary between farms and regions.

Also, all farms aerate their ponds with paddlewheel aerators, especially in summer. The aerators are generally placed near the feeders close to the monk at the deepest part of the pond and are usually switched on only during the night hours when oxygen levels become critically low. The details of fish culture practices outlined above were kindly provided by S. Rothbard, F. Svirski, A. Ben Ari and I. Peleg (pers. comm.). Further information can be found in Milstein and Hulata (this vol.).

Data Collection

Data were obtained for three of the four fish culture regions in Israel. These are the Upper Galilee, Western Galilee and the Bet She'an Valley (Fig. 1). For the Coastal Plain no adequate data could be obtained since the few but very large farms had data recording methods causing too many problems to enable a sufficiently accurate reconstruction of pond management procedures, inputs, samplings and harvests. In the three abovementioned regions, the cooperating farmers provided their data from 1980 to 1986.

The farmers maintain record sheets for each individual pond and production cycle (Hepher and Pruginin 1981) which were translated from Hebrew and entered into computer files. All fish stocking, periodic weighing and harvesting events and amounts are recorded together with average and total weights, amounts of manure and feed applied, and other observations of concern. However, few water quality parameters are recorded. Ponds are sampled at regular intervals (usually every two weeks) to determine average fish weight in order to adjust feeding rates (Hepher 1985). The mesh size ranges from 10 to 25 mm for

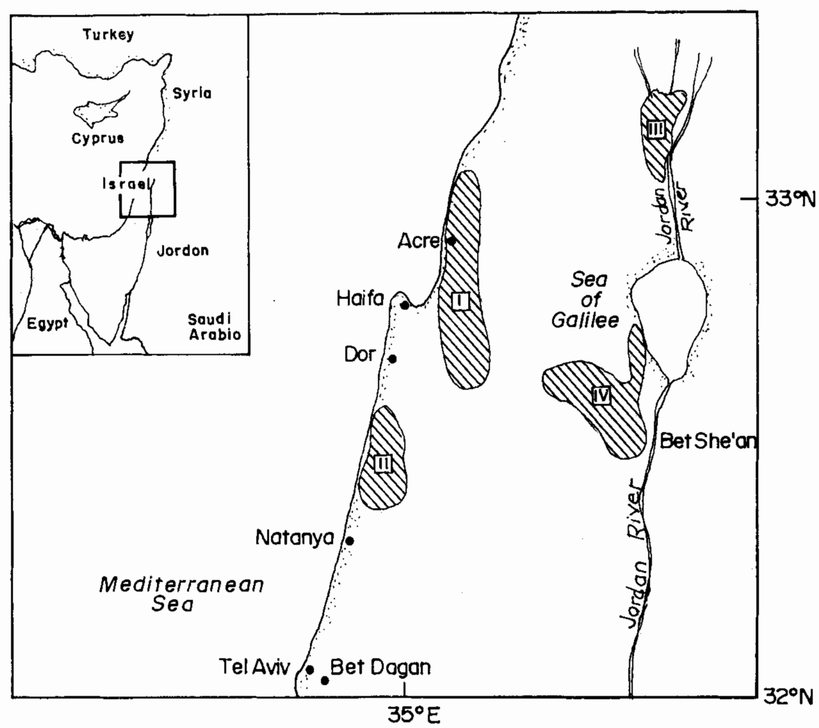


Fig. 1. Locations of the four fish culture regions in Israel. I. = Western Galilee, II. = Coastal Plain, III. = Upper Galilee, IV. = Bet She'an Valley.

sampling procedures and 20 to 35 mm for harvesting. These samplings and the intervals between them conform to the data requirements of the extended Gulland-and-Holt method of analysis.

Several visits were made to a total of 15 farms in all regions to interview fish farm managers and regional extension officers on details of their procedures. After examining samples of the data recording sheets maintained by the farm managers, the data formats required for analysis were designed. When inputting data, considerable error checking was necessary.

Data Handling and Processing

Commercial farm production data and meteorological data were supplied on MS-DOS floppy disks. Details of the editing procedures are given below. The final dataset for analysis comprised 21 variables in 960 cases with a file size of 361 KB (filename FARMSALL.WK1, see Appendix II).

Central to all data handling and editing procedures was the establishment of several small spreadsheets on microcomputers. From these raw data files the required average values per growth interval were computed and inserted into the main datafile. Regression analyses were performed with the SPSS software package (Norusis 1985). The 95% significance level was used for all tests. Procedures of data handling and processing and preparation of environmental variables are described in Prein (this vol.).

TILAPIA LENGTH-WEIGHT RELATIONSHIP

All fish size data obtained from Israeli farms were in form of weight, i.e., length data were not available. Since the main method devised for final data analysis requires fish size to be expressed as length (Pauly et al., this vol.), a large number of Israeli tilapia hybrids (n = 1,094) was measured and weighed at the Dor research station and at Kibbutz Gan Shmuel.

For the present analysis, length-weight relationships were estimated and were regarded as representative for all tilapia hybrids in Israel. While other strains may be used by different breeders, the hatchery at Nir David is the largest supplier of stocking material in Israel. No other length-weight relationship was ever established for these or other Israeli tilapia hybrids. The mean weights in the Israeli tilapia-hybrid dataset were converted to mean lengths.

WATER TEMPERATURE

Water temperature is one of the important variables in fish culture. This variable is not recorded by the fish farmers. In order to be able to include this factor in the analysis, water temperature had to be estimated for the three fish culture regions. For this purpose, the available water temperatures from Dor station (Prein, this vol.) were used as dependent variable and the meteorological data from the nearby meteorological station at 'En Hahoresh were used as independent variables to develop the following multiple regression relationship:

$$TW = 1.8169 + 0.70266\ TA + 0.01655\ QGR \qquad ...1)$$

with $n = 2096$, $R^2 = 0.786$, $SEE = 3.041$, $P < 0.001$

where

TW = early morning water temperature at Dor (°C);

TA = average of minimum and maximum air temperature on the preceding day, by Meteorological Services (°C); and

QGR = Total global solar radiation in Bet Dagan ($ly \cdot day^{-1}$)

This function was used to predict the missing water temperatures in the three fish culture regions from the available meteorological data. These daily values were averaged according to the respective fish sampling intervals to represent the mean water temperature in the sampling intervals. Besides water temperature no water quality variables were available.

METEOROLOGICAL VARIABLES

Climatic diagrams were studied (Fig. 2) for a comparison of the four different fish culture regions.

To obtain the necessary meteorological variables the Meteorological Services for Israel in Bet Dagan (near Tel Aviv) were approached, since they maintain a network of stations distributed over the country.

For the years 1980 to 1987, daily values of air temperature, cloud cover, wind speed and total solar radiation were supplied on floppy disks, comprising approximately 1.2 megabytes of raw data files. Table 1 gives an overview of all sources and stations from where meteorological data were obtained.

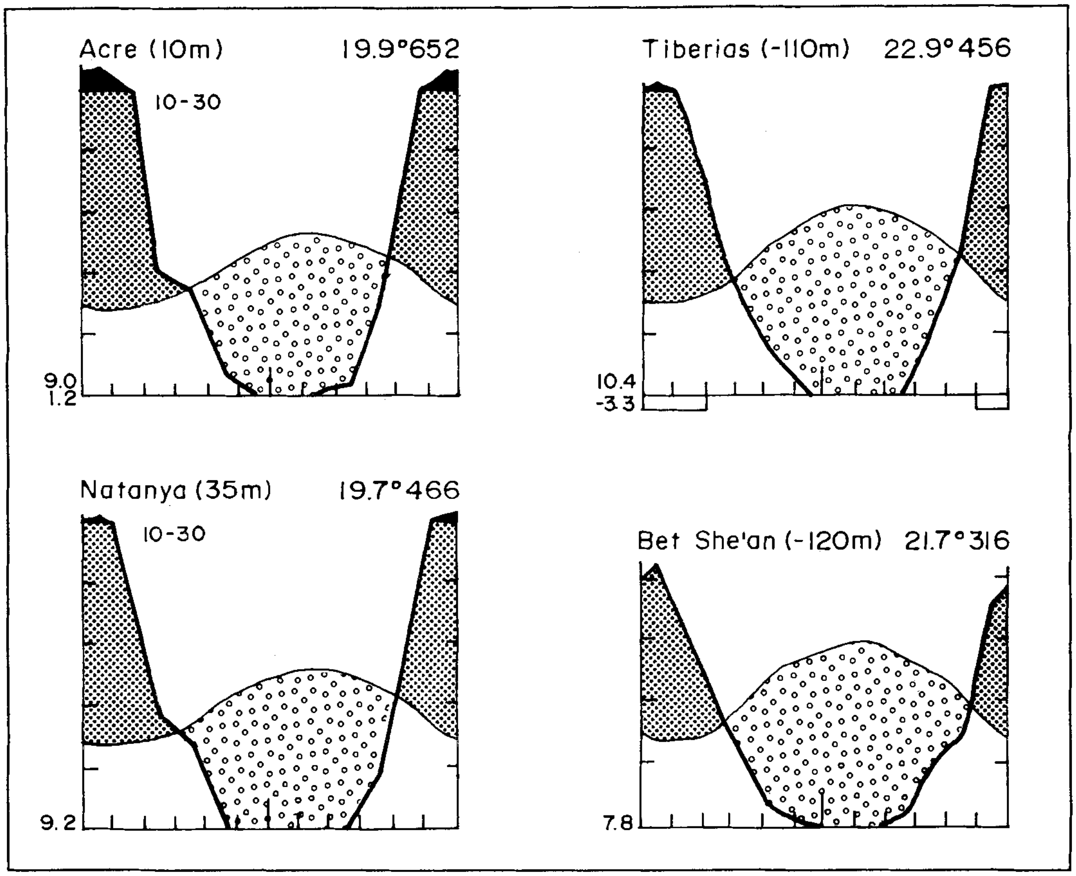


Fig. 2. Climatic diagrams for meteorological stations representative of the four fish culture regions in Israel. From Walter and Lieth (1969). See Prein (this vol.) for interpretation of diagrams.

Table 1. Sources of meteorological data for Israel used in this study. X = data available for all eight years of data. (X) = incomplete data.

Region	Station	Coordinates	Climatic diagrams	Air temperature	Water temperature	Meteorological data	Solar radiation
I	Acre	35°04'E 32°55'N	X				(X)
	Naharya	35°06'N 33°01'N		(X)		(X)	
II	Dor	34°56'E 32°37'N		X	X		
	Ein Hahoresh	34°56'E 32°23'N				X	(X)
	Natanya	34°51'E 32°19'N	X				
III	Hulata	35°36'E 33°03'N					(X)
	Kfar Blum	35°36'E 33°10'N				X	
	Tiberias	35°32'E 32°47'N	X				
IV	Bet She'an	35°30'E 32°29'N	X				
	Tirat Zvi	35°31'E 32°25"N				X	(X)
	Bet Dagan	34°49'E 32°00'N		X		X	X

The daily values of the meteorological variables were averaged according to the respective fish sampling intervals of each pond to represent their mean effect during the fish growth interval. In the case of air temperature, the average of the daily minimum and maximum values were used.

From comparisons of long-term monthly averages of solar radiation of stations in the four fish culture regions and the centers' main station in Bet Dagan, it was found that solar radiation does not differ significantly between regions (Fig. 3). Therefore, it was decided to use the data recorded by the Meteorological Services for Israel at Bet Dagan for all fish culture regions.

The relationship between the theoretically available amount of total solar radiation reaching the top of the atmosphere and the actually measured amount reaching the ground is depicted in Fig. 4. The theoretical values were computed with the SUNLIGHT computer program of Prein and Gayanilo (1992). The graph shows the generally observed reduction of solar radiation by 30 to 40% (for clear skies) when it passes through the atmosphere. Particularly cloudless conditions prevail in the summer months of April to October in Israel.

The supplied dataset comprised eight years (1980 to 1987) of daily total global radiation values in Langleys, measured on a plane surface. Some values were missing and were filled in according to the methods described above or through interpolating between two neighboring values if these were of similar magnitude, suggesting a stable weather period.

In the dataset obtained, other meteorological variables were available and processed for inclusion in the final datasets. These were cloud coverage, wind velocity and direction. For the latter three variables, daily values were partly available as 0800, 1400 and 2000 hours readings. Cloud coverage was recorded in octals, i.e., the total cover of the sky as a fraction of 8.

Data Analysis Methods

The assembled dataset was analyzed with the extended Gulland-and-Holt plot (Pauly et al., this vol.) and path analysis (Prein and Pauly, this vol.). The derived regression models were subjected to sensitivity analysis (Majkowski 1982; Prein, this vol.).

Results

Length-Weight Relationship of Tilapia Hybrids

The functional length-weight relationship estimated for Israeli tilapia hybrids has the form:

$$W = 0.01442 \cdot TL^{3.0948} \qquad ...2)$$

with n = 1094, r^2 = 0.996, SEE = 0.0438, $P < 0.001$; size range: 0.1 to 838 g, 1.9 to 33.6 cm, where W is in g and total length (TL) is in cm (Fig. 5). In the present case, weights were transformed to lengths using:

$$TL = \sqrt[3.0948]{\frac{W}{0.01442}} \qquad ...3)$$

Total and standard lengths (SL) can be converted with (functional regression):

$$TL = 0.1078 + 0.1082\ SL \qquad ...4)$$

with n = 1,094, r^2 = 0.999, SEE = 0.00692, $P < 0.001$; size range: 1.5 to 26.9 cm SL, 1.9 to 33.6 cm TL.

Extended Gulland-and-Holt Analysis

An ordinary Gulland-and-Holt plot of the entire dataset of Israeli tilapia hybrids grown on

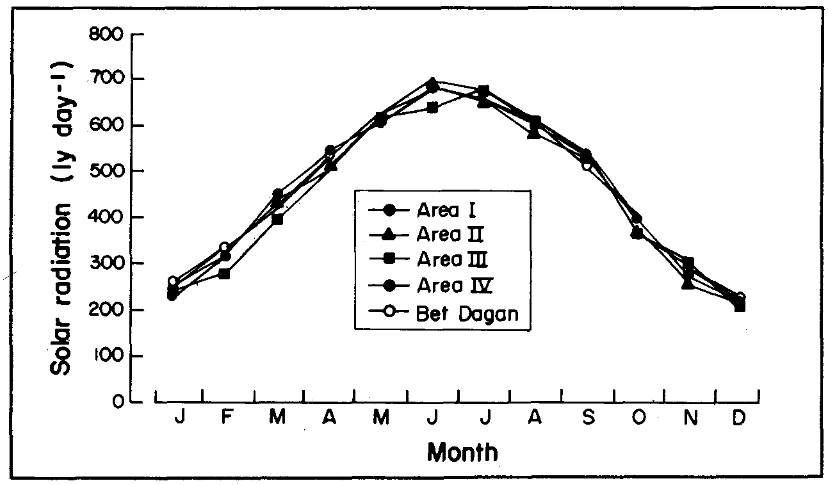


Fig. 3. Average monthly total solar radiation for five locations in Israel. Representative meteorological stations for the four fish culture regions are: I = Acre (Western Galilee), II = Ein Hahoresh (Coastal Plain), III = Hulata (Upper Galilee), IV = Tirat Zvi (Bet She'an Valley). Bet Dagan near Tel Aviv is shown for comparison. Data are averages over 10 years (unpublished data, Israel Meteorological Services, 1987).

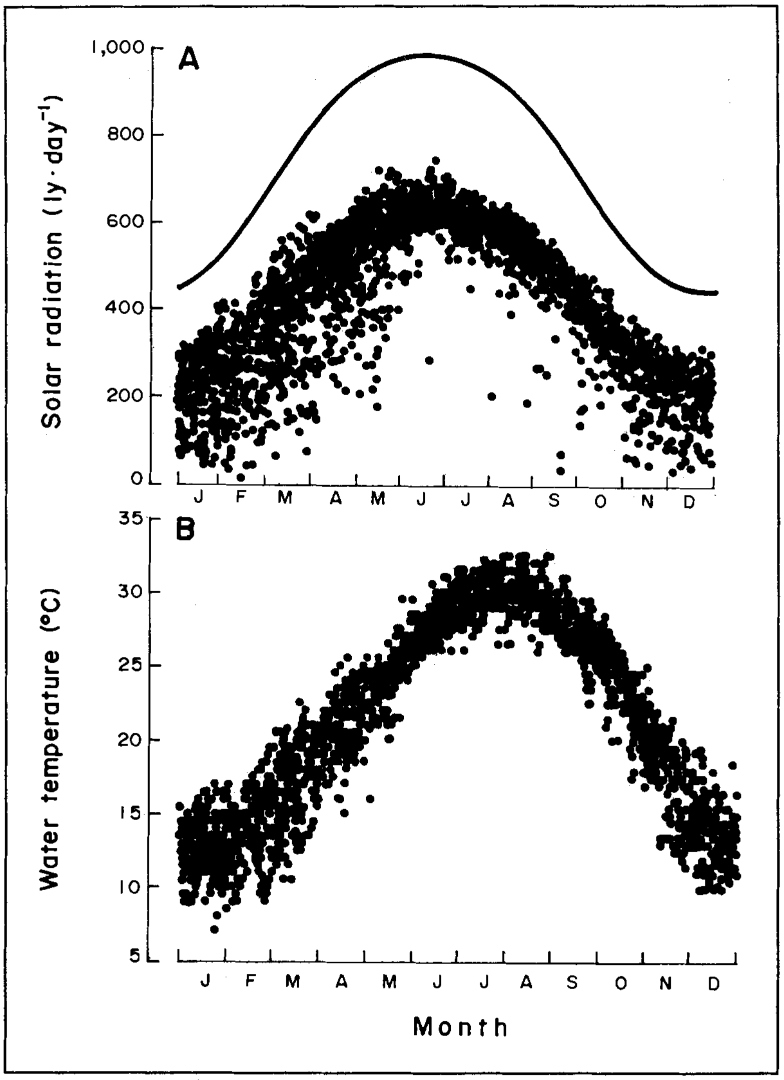


Fig. 4. A) Annual pattern of solar radiation and water temperature in the Coastal Plain, Israel.
A) Theoretical values (line) of total solar radiation reaching the top of the atmosphere over Dor Station, Israel, and daily values actually measured by the Meteorological Services for Israel at Bet Dagan near Tel Aviv (points) for 1980 to 1987.
B) Seasonal variation of daily average water temperature in experimental pond at Dor Research Station, Israel. Data from 1978 to 1987. Note lag in temperature during summer months, compared to solar radiation curve.

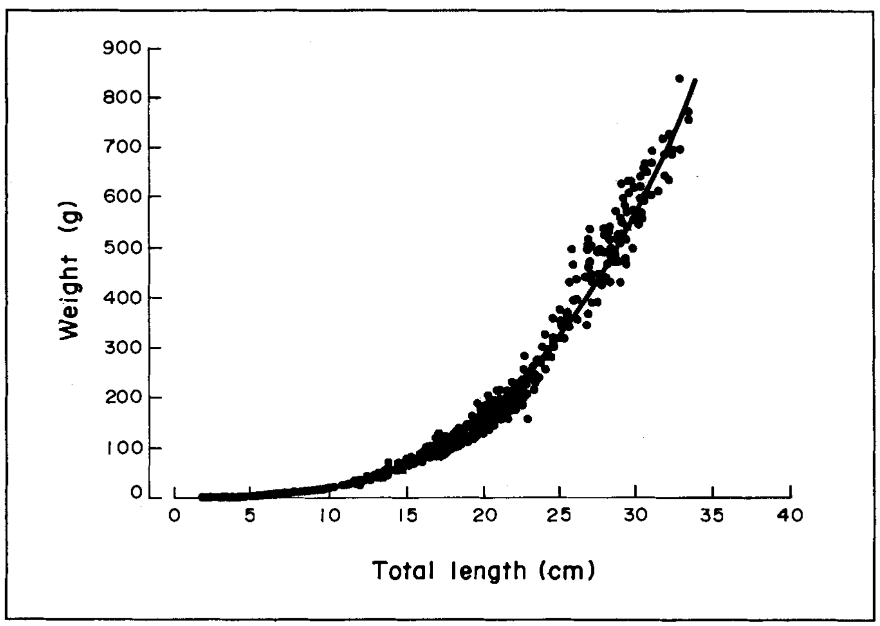


Fig. 5. Length-weight relationship of all-male (hormonally sex-inversed) tilapia hybrids cultured in Israel. n = 1,094. See text for regression equation.

commercial farms is shown in Fig. 6. A sample dataset for derivation of a linear regression model yielded an ordinary Gulland-and-Holt plot of the form:

$$\Delta L/\Delta t = 0.207 - 0.00524\ ML \qquad \text{...5)}$$

with n = 300, r^2 = 0.1195, SEE = 0.0811, P < 0.001, K = 0.00524 day^{-1} and L_∞ = 39.6 cm, where K is the growth constant of the von Bertalanffy growth function, derived as the slope in (5) and L_∞ is the asymptotic length, i.e., the (mean) length the fish would reach if it were to grow indefinitely, derived as the intercept of the regression line with the abscissa.

The equation obtained with the extended Gulland-and-Holt plot takes the form:

		mean	range
$\Delta L/\Delta t$ =			
$-4.108 \cdot 10^{-3}$	mean length (cm)	21.8	4.8-35.3
$-4.981 \cdot 10^{-4}$	SQRT stocking density ($\text{kg} \cdot \text{ha}^{-1}$)	3,026	123-10,877
$+3.296 \cdot 10^{-4}$	solar radiation ($\text{ly} \cdot \text{day}^{-1}$)	526	276-677
-0.03517			

...6)

with n = 300, R^2 = 0.244, SEE = 0.0754, P < 0.001, K = 0.004108 day^{-1} and L_∞ = 44.5 (29.4 to 50.3) cm, where K and L_∞ are obtained according to Pauly et al. (this vol.). 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
mean length	=	8.3 %	$-5.671 \cdot 10^{-3}$	$-2.546 \cdot 10^{-3}$
stocking density	=	1.2 %	$-1.014 \cdot 10^{-3}$	$+1.796 \cdot 10^{-5}$
solar radiation	=	11.3 %	$+2.237 \cdot 10^{-4}$	$+4.355 \cdot 10^{-4}$
constant			-0.04038	+0.11071

For management purposes, a regression model based on variables which can be obtained more conveniently takes the form:

		mean	range
$\Delta L/\Delta t$ =			
$-6.797 \cdot 10^{-3}$	mean length (cm)	21.8	4.8-35.3
$-3.766 \cdot 10^{-7}$	stocking density ($\text{n} \cdot \text{ha}^{-1}$)	22,367	875-352,000
$+7.364 \cdot 10^{-3}$	water temperature (°C)	27.8	15.1-34.0
-0.04503			

...7)

with n = 300, R^2 = 0.209, SEE = 0.0771, P < 0.001, K = 0.00680 day^{-1} and L_∞ = 35.4 cm.

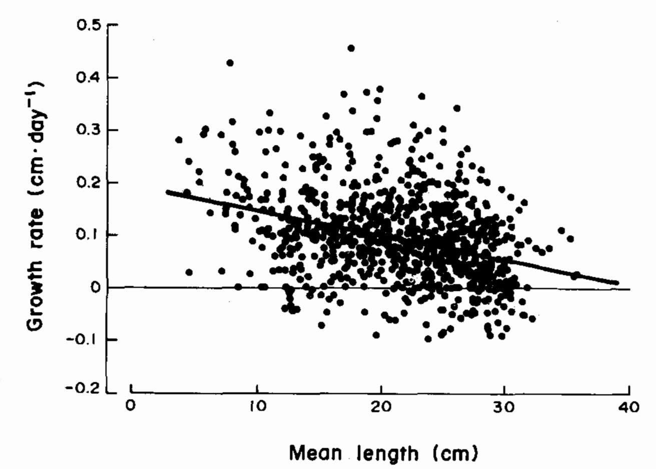


Fig. 6. Gulland-and-Holt plot of all-male, sex-inversed tilapia hybrids grown in commercial farms in Israel in 1980 to 1987. n = 960. See text for regression equation.

A large amount of unexplained variance remained. Since the data are based on weights, negative values were theoretically possible and therefore left in the dataset. Also, very high values for daily growth rate, which may not be realistic, are included in the dataset. This leads to a larger amount of variance.

The variables are responsible for the following portions of the explained variance: mean length (13.7%), stocking density (1.8%) and water temperature (8.2%).

It was not possible to incorporate any further variables into the latter two equations for the explanation of tilapia hybrid growth rate. The main treatment variables, pellet, sorghum and manure inputs, were not significant or were negatively correlated with fish growth. Pellet application was increased as the fish grew. In the dataset, the applications of the input variables often occurred irregularly. This resulted in cases (i.e., intervals) of zero inputs, followed by cases of very high inputs. These irregularities prohibit the detection of any effect by the regression. Wind velocity and cloud cover also were non-significant predictors of tilapia hybrid growth. Their effects should become apparent when related to variables of the pond environment, which were not measured here. Finally, dummy variables for the three different regions were included in the analysis. These were thought to represent any differences among regions in pond management, environment, geology, etc., which were not included in the available data. These were also not significant.

Test of the Derived Model

With the remaining 660 cases of the dataset, the same model as above was computed. The equation obtained takes the form:

		mean	range
$\Delta L/\Delta t =$			
$-3.427 \cdot 10^{-3}$	mean length (cm)	21.6	3.9-35.8
$-6.527 \cdot 10^{-4}$	stocking density ($kg \cdot ha^{-1}$)	3,067	145-18,602
$+2.819 \cdot 10^{-4}$	water temperature (°C)	27.8	14.5-33.6
-0.04503			

...8)

with n = 660, $R^2 = 0.224$, SEE = 0.0719, $P < 0.001$.

The regression coefficients are not significantly different from those determined from the sample dataset.

PATH ANALYSIS

Based on the sample dataset of 300 cases, a causal path diagram based on the ordinary Gulland-and-Holt plot is shown in Fig. 7A. Through inclusion of further variables in a multiple regression, a more detailed path diagram results (Fig. 7B). Only two variables besides mean length were significant predictors of hybrid tilapia growth rate. These were solar radiation and tilapia stocking density in $kg \cdot ha^{-1}$. Solar radiation has a positive effect on growth rate, while stocking density negatively influences growth rate. Water temperature was also identified as a significant predictor of growth rate, yet it was highly correlated with solar radiation and was less influential than solar radiation. Therefore, it was decided to incorporate only solar radiation into the path diagram, since this variable also acts through its influence on primary production (i.e., natural food availability to tilapias). No information on plankton content in the ponds was available, though.

Water temperature can be predicted with three meteorological variables. These are wind velocity at noon, cloud cover at noon, and total daily solar radiation. Solar radiation and wind are negatively correlated with cloud cover, while wind and solar radiation are positively correlated.

A causal path diagram for a model intended for management purposes is shown in Fig. 7C. Instead of solar radiation, water temperature is incorporated here as a positive predictor of growth rate. Stocking density is in the form of $n \cdot ha^{-1}$ instead of $kg \cdot ha^{-1}$. Together though, this set of variables leaves 79% of the total variance in the dataset unexplained.

SENSITIVITY ANALYSIS

The large amount of variance in the dataset from the commercial farms in Israel also found its expression in the sensitivity analysis of the derived regression model (Fig. 8). The variable causing the most change with each 10% change was water temperature. A reduction of 40%, compared to the mean, led to a nearly 500% lower growth rate. A 50% change in stocking density caused 250% change in growth rate, while a 50% change in mean length led to a 430% change in growth rate. In the dataset, the ranges of these independent variables and the dependent variable were large. The resulting strong effects were adequately described by the regression.

Discussion

The tilapia grown in the commercial fish farms are all-male, sometimes sex-inversed hybrids (Hepher and Pruginin 1981). High amounts of chicken manure, pellet feed and sorghum are applied (Eren et al. 1977), accompanied by aeration where necessary. High nutrient inputs lead to increased dynamics of pond water chemistry (Hepher 1959; Abeliovitch 1967), resulting in high fish production (Hepher 1962; Hepher and Pruginin 1981; Rimon and Shilo 1982). Natural food webs play a key role in the nutrition of tilapia in ponds (Schroeder 1978, 1983). Based on these high-intensity operational criteria, higher tilapia growth rates can be expected compared to mixed-sex, nonfed systems.

The von Bertalanffy growth parameters obtained with the extended Gulland-and-Holt method for the tilapia hybrids commercially cultured in Israel reflect the range of culture conditions represented in the data. Depending on the stocking density and amount of solar radiation, L_∞ ranged from 29.4 to 50.3 cm, with a value of 44.5 cm for average conditions. The value of K for the range of different conditions is 0.00411 day^{-1}. These parameters can be used for growth prediction of Israeli tilapia hybrids if the anticipated conditions fall within the range covered by the data used here, i.e., from which the parameters were derived. Then the conditions may vary even on a daily basis, permitting to model fish growth in daily steps.

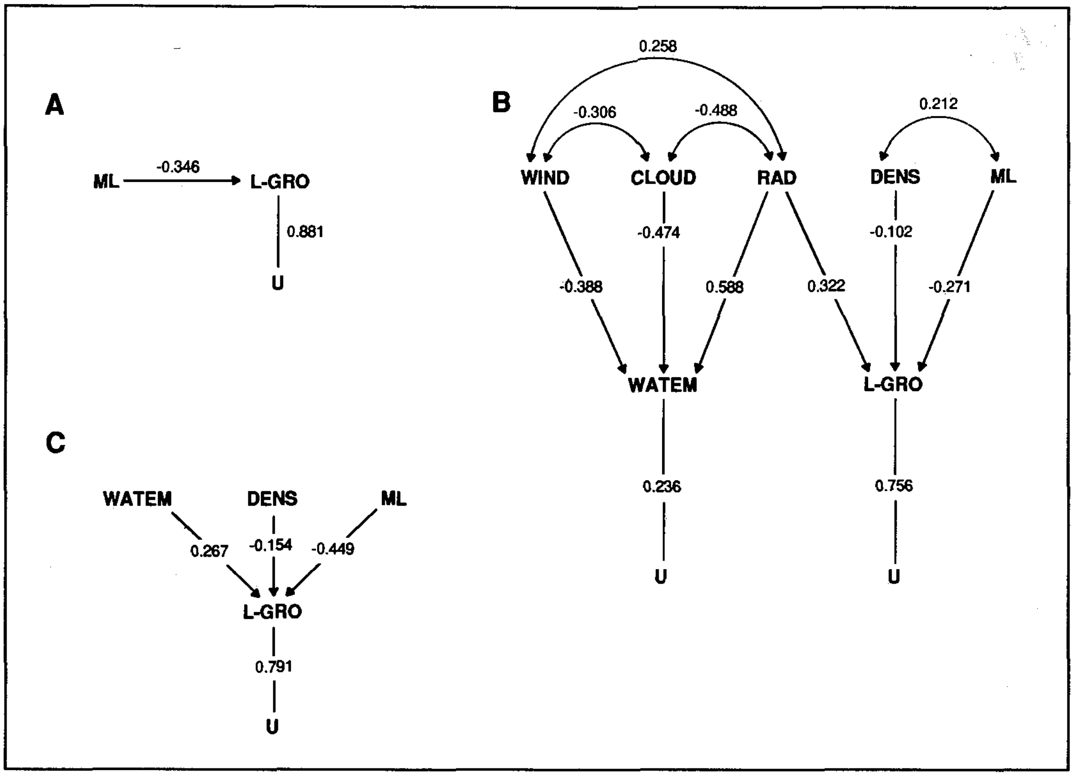


Fig. 7. Path diagrams for a) the ordinary Gulland-and-Holt plot, b) the extended Gulland-and-Holt plot with three direct predictors of tilapia growth and two of water temperature, c) the extended Gulland-and-Holt plot for a model intended for management purposes. ML = mean length (cm); L-GRO = growth rate in length (cm·day^{-1}); WATEM = early morning water temperature (°C); DENS = square root of stocking density (kg·ha^{-1}); WIND = cumulative run of the wind (km·day^{-1}); CLOUD = cloud cover (octals); RAD = solar radiation (ly·day^{-1}); U = residual effect or unexplained variance.

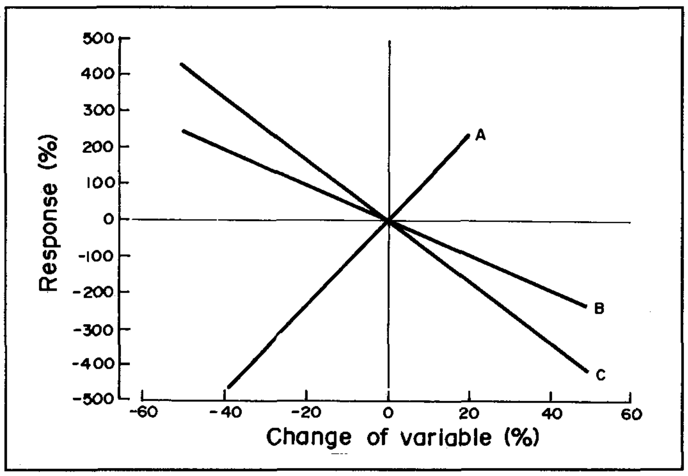


Fig. 8. Sensitivity analysis of the regression model for growth rate of all-male tilapia hybrids grown in commercial farms in Israel, based on the extended Gulland-and-Holt method. The % response from the average growth rate (ordinate) is shown as a result of % change from the average value of each independent variable (abscissa). Predictor variables: A = water temperature (range shortened to that covered by actual data), B = stocking density, C = mean length.

To compare the growth parameters from different populations, the growth performance index ϕ' (= $\log_{10}K + 2\log_{10}L_{\infty}$) is a convenient and robust tool (Moreau et al. 1986; Pauly et al. 1988). The values of ϕ' for the tilapia hybrids grown at the commercial farms in Israel (mean ϕ' = 3.47) are in the same range as those determined for mixed-sex Nile tilapia determined in the Philippines (Prein, this vol.), and for all-male Nile tilapia determined by Pauly et al. (1988) who computed ϕ' values for 65 Nile tilapia stocks reared under aquaculture conditions and compared them with those for natural stocks (Moreau et al. 1986; Pauly et al. 1988). The values for L_{∞} are considerably higher for the commercially produced fish, which is balanced by smaller values of K. This means that tilapia hybrids in the commercial farms grow larger, but faster than Nile tilapia in manure-fed ponds. This is a consequence of better culture conditions, mainly food (pellet feeding and manuring) and oxygen supply (aeration) and reduced territorial behavior (through monosex culture).

In the compiled dataset, collinearity between meteorological or management variables and tilapia growth rate did occur. A possible solution is suggested by Stergiou (1989) although its application precludes the derivation of VBGF growth parameters and the option of forward growth prediction. The data covered the entire year, including overwintering cycles, where temperatures were low and fish were not fed. This led to wide ranges for the environmental and management variables. In spite of this, solar radiation or water temperature were the only significant predictors of tilapia growth. Other variables such as pellet, sorghum and manure application rates were not significant, since they did not vary in a range adequate to experimental practice. Pellets and sorghum were applied according to fish weight, where the rates used were similar in all farms. Manuring rates also did not vary much. Thus, farmers keep nutritional input rates at levels optimal for fish production.

Furthermore, the applications of these inputs were not distributed regularly over the intervals. Generally, the amounts given in the farmers' record sheets were bulk inputs applied (i.e., loaded into large demand feeders) at a certain date, and which had a long-lasting residual effect in the ponds, thereby covering several intervals in which no applications were given. These 'zero-application' intervals were treated as such in the regression, although there might have been residual effects from applications in previous intervals. In some cases, the farmers could not clearly attribute the individual applications of manure, etc. to specific intervals. Only total sums over several intervals were available. Both types of error lead to a blurring of the relationships between nutrient inputs and growth rate. This is an unfortunate tradeoff since the precision of data recordings that is adequate for commercial production cannot meet the demands of the scientific method presented herein.

A further considerable source of variance may be due to errors in determined fish sizes at the sampling events. In the extended Gulland-and-Holt method, 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.

Since fish growth (i.e., productivity of pond environments) at different locations is analyzed here, a rewarding approach is to study and compare the climatic diagrams (Prein, this vol.) available for these locations.

For Israel the climatic type is Mediterranean, with dry summers and rain in winter. The meteorological stations in Acre, Natanya, Tiberias and Bet She'an (Fig. 2) were chosen as representative for fish culture regions I, II, III and IV, respectively. Characteristics are summarized in Table 2.

Conspicuous are the differences in altitude between the two regions on the coastal plain (Regions I and II) and the two regions of the inner (Jordan and Bet She'an) valleys. The inner valleys, which lie more than 100 m below sea level, have a slightly higher average annual temperature which is due to higher summer temperatures. Also, the dry season lasts approximately one month longer.

The location of Tiberias does not correctly represent the region of the Upper Galilee, since the latter is approximately 25' further north, in a fairly narrow valley surrounded by mountains (except from the south). Temperatures are known to

Table 2. Climatic characteristics of the four fish culture regions (I-IV) in Israel.

Region	Station	Coordinates	Altitude m	Mean temp /year °C	Mean rain /year mm/mo	Temperature °C range	Temperature month min	Temperature month max	Rain mm/mo range	Rain month min	Rain month max	Season dry	Season rainy
I	Acre	35°04'E 32°55'N	10	19.9	652	3-27	JAN	AUG	0-185	JUN	JAN	APR-OCT	NOV-MAR
II	Natanya	34°51'E 32°47'N	35	19.7	466	13-27	JAN	AUG	0-130	JUL	JAN	APR-OCT	NOV-MAR
III	Tiberias	35°32'E 32°47'N	-110	22.9	456	13-31	JAN	AUG	0-125	JUL	JAN	APR-NOV	DEC-MAR
IV	Bet She'an	35°30'E 32°29'N	-120	21.7	316	13-30	JAN	AUG	0-83	JUL	JAN	APR-NOV	DEC-MAR

be lower and precipitation is above 500 mm·year^{-1}. Unfortunately no other climatic diagram was found to better represent the region.

In spite of the differences pointed out above, no differences in tilapia growth were found between the fish culture regions, based on the dummy variables assigned to them. Wind velocity and cloud cover were also not found to be significant predictors for fish growth. The length of the intervals averaged 16 days but a few lasted up to 205 days. Longer intervals cannot precisely reflect the changes of variables and effects on fish growth.

Acknowledgements

We gratefully acknowledge the good cooperation of numerous fish farm managers in Israel, too numerous to mention personally. Thanks are due to Aza Ben Ari, Feifke Svirski and Israel Peleg (Regional Fishery Extension Officers) for providing numerous details and informative discussions. Gideon Hulata (Dor Research Station, ARO) is thanked for help in negotiations with the fish farmers and numerous fruitful discussions. Amir Halevy kindly provided the water temperature recordings. Further, we wish to acknowledge the Israel Meteorological Services, Bet Dagan, for supplying the meteorological data used in this study through our cooperation with G. Hulata, ARO. This study was supported by the Federal Republic of Germany-Israel Fund for Agricultural Research in Third World Countries (GIARA), Grant 1-2/86, financed by the German Ministry for Economic Cooperation (BMZ).

References

Abeliovitch, A. 1967. Oxygen regime in Beit-Shean fish ponds related to summer mass fish mortalities. Preliminary observations. Bamidgeh 19:3-15.

Eren, Y., T. Tsur and Y. Avnimelech. 1977. Phosphorus fertilization of fishponds in the Upper Galilee. Bamidgeh 29:87-93.

Hepher, B. 1959. Chemical fluctuations of the water of fertilized and unfertilized fishponds in a subtropical climate. Bamidgeh 11(1):3-22.

Hepher, B. 1962. Ten years of research in fish ponds fertilization in Israel. I. The effect of fertilization on fish yields. Bamidgeh 14(2):29-38.

Hepher, B. 1985. Aquaculture intensification under land and water limitations. GeoJournal 10:253-259.

Hepher, B. and Y. Pruginin. 1981. Commercial fish farming, with special reference to fish culture in Israel Wiley Interscience, New York. 261 p.

Hepher, B. and Y. Pruginin. 1982. Tilapia culture in ponds under controlled conditions, p. 185-203. *In* R.S.V. Pullin and R.H. Lowe-McConnell (eds.) The biology and culture of tilapia. ICLARM Conf. Proc. 7, 432 p.

Hulata, G. 1988. The status of wild and cultured tilapia genetic resources in various countries: Israel, p. 48-51. *In* R.S.V. Pullin (ed.) Tilapia genetic resources for aquaculture. ICLARM Conf. Proc. 16, 108 p.

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.

Moreau, J., C. Bambino and D. Pauly. 1986. Indices of overall growth performance of 100 tilapia (Cichlidae) populations, p. 201-206. *In* J.L. Maclean, L.B. Dizon and L.V. Hosillos (eds.) The First Asian Fisheries Forum. Asian Fisheries Society, Manila. 727 p.

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

Pauly, D. J. Moreau and M. Prein. 1988. A comparison of overall growth performance of tilapia in open waters and in aquaculture, p. 469-479 *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Prein, M, and F.C. Gayanilo Jr. 1992. SUNLIGHT = a computer program for calculation of daily solar radiatioin. 15(3):30-31. Naga, ICLARM Q.

Rimon, A. and M. Shilo. 1982. Factors which affect the intensification of fish breeding in Israel. Bamidgeh 34:87-100.

Sarig, S. 1990. The fish culture industry in Israel in 1989. Isr. J. Aquacult./Bamidgeh 42(2):39-45.

Schroeder, G.L. 1978. Autotrophic and heterotrophic production of microorganisms in intensely-manured fish ponds, and related fish yields. Aquaculture 14:303-325.

Schroeder, G.L. 1983. Stable isotope ratios as naturally occurring tracers in the aquaculture food web. Aquaculture 30:203-210.

Stergiou, K.I. 1989. A method to cope with collinearity of ecological data sets in community studies. Coenoses 4:91-94.

Walter, H. and H. Lieth. 1969. Klimadiagramm Weltatlas. Gustav Fischer, Jena. pag. var.

Multiple Regression Analysis of Growth and Production of *Oreochromis niloticus* in Net Cages in Lake Sampaloc, Philippines*

LOLITA V. AQUINO-NIELSEN**, *Vibevangen 13, DK-3520 Farum, Denmark*

AMALIA S. MANRIQUE-PEMPENGCO, *Institute of Environmental Science and Management (IESAM), University of the Philippines in Los Baños, Laguna, Philippines*

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

Aquino-Nielsen, L.V., A.S. Manrique-Pempengco and M. Prein. 1993. Multiple regression analysis of growth and production of *Oreochromis niloticus* in net cages in Lake Sampaloc, Philippines, p. 189-198. *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

The results of two previous studies on growth and production of Nile tilapia (*Oreochromis niloticus*) in net cages in Lake Sampaloc, Luzon, Philippines, are presented. A dataset from Nile tilapia growth experiments, conducted in 1980, was reanalyzed with the "extended Gulland-and-Holt" method. Multiple regression analysis was used to test hypotheses related to the impact of environmental, management and socioeconomic variables on annual production of Nile tilapia in net cages in 1986. Secchi disk visibility was found to be a better predictor than gross primary production of those environmental effects which determine growth rate of Nile tilapia. Occurrence of hydrogen sulfide in the epilimnion was shown to have a negative effect on fish growth. The morphoedaphic index of Lake Sampaloc was computed and used to compare fish production from this Lake with that of African Lakes.

Introduction

Lake Sampaloc is situated in San Pablo City in Laguna Province, Luzon, Philippines (Fig. 1). It is of volcanic origin, of nearly circular form and has a maximum depth of 27 m. With its surface area of 102.5 ha (Sutherland 1974), it is the largest of seven lakes in the region (Talavera 1932). The lake supports a capture fishery, documented by Manrique (1988).

Floating net cages for the culture of Nile tilapia (*Oreochromis niloticus*) and Mozambique tilapia (*O. mossambicus*) were introduced into Lake Sampaloc in 1976. Tilapia fingerlings were reared in several cycles over the year, each cycle having a duration of 3 to 6 months, resulting in an extrapolated total annual production of around 60 $t \cdot ha^{-1}$ in the lake in 1976 (Fermin 1978). In these early years, fish in cages fed to a large extent on the natural food available in the water column and only a few cage operators gave additional food in form of rice bran. In later years, fish growth was reduced in all lakes of the region where Nile tilapia grew to only 80 g in a one-year culture period (BFAR extension workers, pers. comm.).

The present contribution analyzes Nile tilapia production in Lake Sampaloc. It is assembled from (i) unpublished parts of a study on primary production and Nile tilapia growth in net cages (Aquino 1982), (ii) a reanalysis of data presented in Aquino (1982) and Aquino and Nielsen (1983), and (iii) unpublished parts of a study of total fish yields in net cages and Lake Sampaloc itself (Manrique 1988).

The main objectives of the present contribution is to apply different methods of multivariate statistical analysis to the data obtained in these studies and use their results to identify and quantify the factors which affect fish production - both natural and through cage culture - in Lake Sampaloc.

From 1979 to 1980, the lake had a total net cage area of 2.7 ha in 134 cages. It was observed by the net cage farmers that a decline in the fish production occurred in the lake after a short period of increased fish production. Aquino (1982) attributed this to either: a) depletion of nutrients in the lake as a result of continuous fish harvest, or b) depletion of phytoplankton and other natural food sources due to overgrazing. Studies were undertaken to test these hypotheses (Aquino 1982; Aquino and Nielsen 1983).

Additionally, a holistic approach for yield estimation was used (Manrique 1988). It relies on the morphoedaphic index (MEI), an empirically derived method to assess approximately the potential

*ICLARM Contribution No. 736
**Sequence of authors appears in alphabetical order.

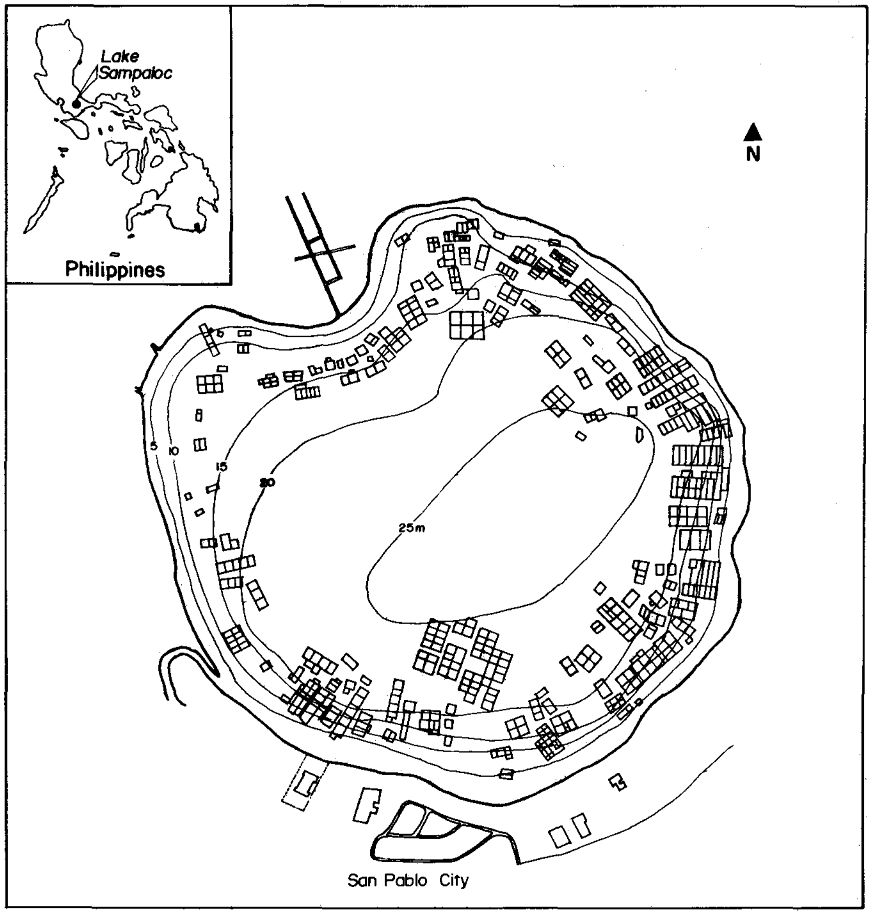


Fig. 1. Location, bathymetry and distribution of fish cages in Lake Sampaloc on Luzon Island, Philippines. Modified from Manrique (1988).

productivity of a lake in terms of fish yield, initially derived for unexploited Canadian lakes (Ryder 1965). For the tropical latitudes the MEI has been applied to lakes in Africa (Henderson and Welcomme 1974). As to our present knowledge, in the Philippines, the MEI has been applied only to lake Mainit (Pauly et al. 1990).

Materials and Methods

Nile Tilapia Growth Experiments in Net Cages

In a first study, six series of Nile tilapia (*Oreochromis niloticus*) culture trials were conducted in triplicate in Lake Sampaloc (Aquino 1982; Aquino and Nielsen 1983). The studies were started in March, April, May, July, August and September 1980 and lasted three to four months. Cages measuring 1 x 1 x 1.5 m were installed in the southwestern portion of the lake. The nets were tied between bamboo rafts such that the internal volume of water was 1 m^3. Nile tilapia fingerlings were stocked at a density of 50 fish·m^{-3}. Monthly samples of 10 fish per cage were measured, weighed and returned to the cages.

Gross primary production (GPP) and Secchi disk visibility were measured two to three times monthly from November 1979 to October 1980. Secchi depth was measured and water was collected at noon, exposed for approximately six hours (until sunset) at different depths (Vollenweider 1974). The estimate of GPP was obtained according to Gaarder and Gran (1927).

Multiple Regression Analyses

The growth data obtained from the fish culture experiment given in Aquino (1982) were analyzed using the "extended Gulland-and-Holt" method (Pauly and Ingles 1981; Pauly and Hopkins 1983; Hopkins et al. 1988; Pauly et al. this vol.). The weight values of the three replicate net cages were used to compute average weights per treatment. These were transformed to total length (TL) based on the length-weight relationship for Nile tilapia given in Prein (this vol.) i.e.,

$$TL=(W/0.01065)^{1/3.258} \quad ...1)$$

Water temperature data from a depth of two meters were expressed as averages per growth interval (i.e., between sampling events). The same was done for Secchi depth and GPP.

In the absence of actual measurements of H_2S in the lake water around the net cages, an approximation was used by coding highly probable H_2S occurrences in form of a dummy variable. This was done by inspection of the available temperature isopleths (Fig. 2) for the depth range between zero and two meters from the surface. The value of one was assigned to an upwelling event reaching the surface during the culture interval and a value of zero to all other intervals.

Data were arranged in form of a matrix suitable for analysis with the following multiple regression equation (Pauly et al., this vol.):

$$\Delta L/\Delta t = a + b_1 ML + b_2\overline{X}_2 + ... + b_n\overline{X}_n \quad ...2)$$

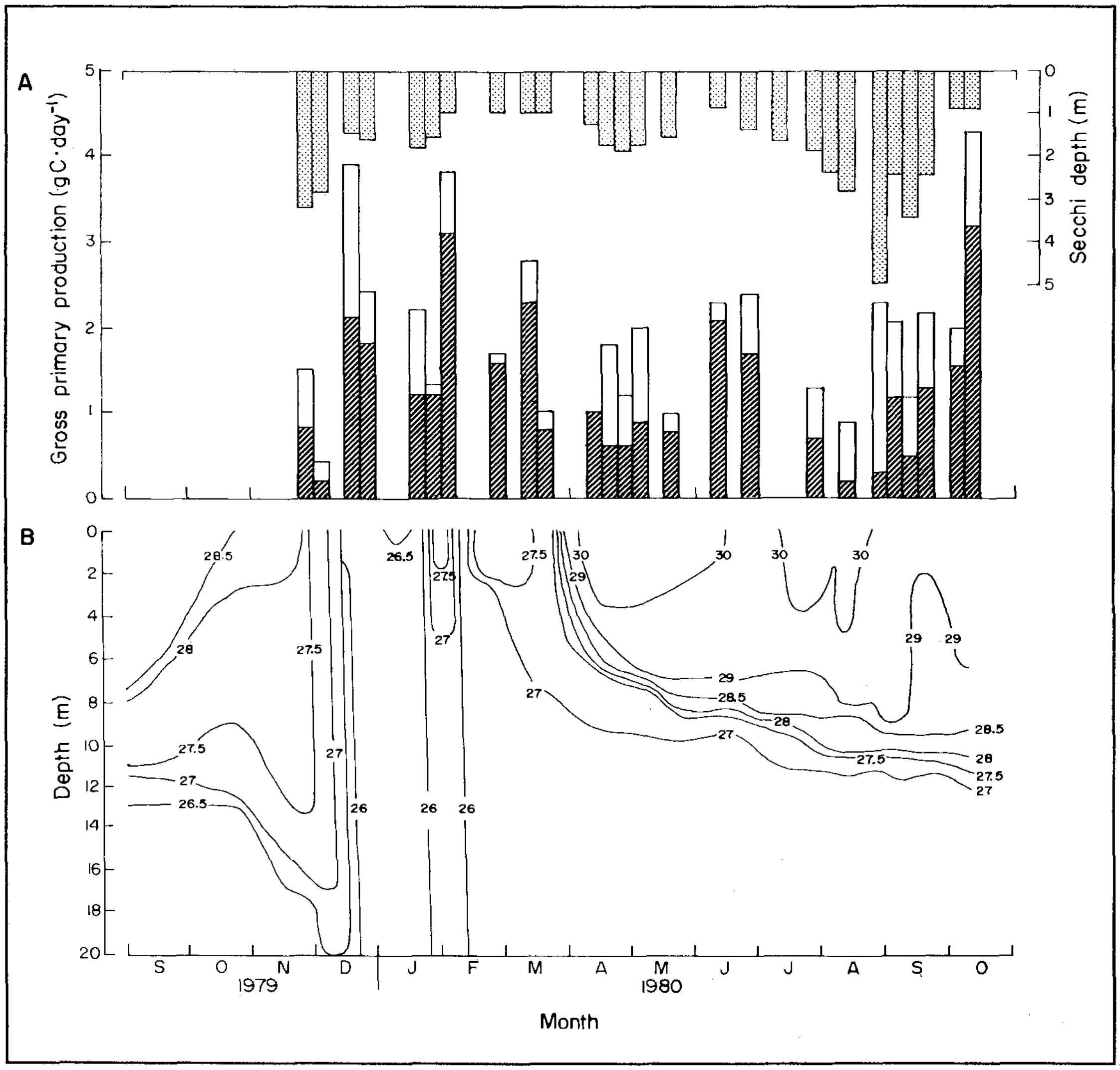


Fig. 2. Characteristics of Lake Sampaloc, Philippines, from September 1979 to October 1980. A) Depth of Secchi disk visibility and amount of gross primary production·m^{-2}·day^{-1} for water column (open bar) and gross primary production·m^{-3}·day^{-1} for upper cubic meter (hatched bar) and for total water column (total height of bar). Adapted from Aquino (1982). B) Diagram of isotherms. Measurements taken in the center of the lake. Dots denote depths of measured samples, according to Aquino (1982). Isotherms drawn by hand.

where

$\Delta L/\Delta t$ = growth rate in length in the interval between samplings, in $cm \cdot day^{-1}$;
$\overline{ML}$ = mean length in interval, cm;
$\overline{X_2}..\overline{X_n}$ = mean values per interval for environmental variables;
$b_1..b_n$ = regression coefficients; and
a = intercept.

In the course of the regression analysis, a set of significant and biologically plausible independent variables was derived according to common regression procedures (Mosteller and Tukey 1977; Draper and Smith 1981). The significance level of $\alpha = 0.05$ was used for all tests.

Parameter estimates of K and L_∞ of the von Bertalanffy growth function

$$L_2 = L_1 \cdot e^{-K\Delta t} + L_\infty(1 - e^{-K\Delta t}) \qquad ...3)$$

can be then obtained according to

$$K = -b_1 \qquad ...4)$$

and

$$L_\infty = \frac{a + b_2X_2 + ... + b_nX_n}{-b_1} \qquad ...5)$$

where

L_1 = total length at the beginning of a growth interval (in cm);
L_2 = total length at the end of a growth interval (in cm);
K = growth constant of the VBGF (here day^{-1}); and
L_∞ = asymptotic length, i.e. mean size the fish would reach if they were to grow indefinitely.

Based on the multiple regression equation obtained, K is estimated from equation (4), while L_∞ must be calculated according to equation (5), separately for each interval over which growth is to be predicted based on average values of the predictor variable during the interval in consideration. Therefore, a single value of K is obtained but as many values of L_∞ can be computed as there are cases in the dataset. The different values of L_∞ reflect variability of growth in response to different treatments or environmental changes and also demonstrate the flexibility of the method to identify and quantify these influences and their changes. The data are available in the file SAMPALOC.WK1 (see Appendix II).

Nile Tilapia Production in Net Cages

In another study (Manrique 1988), multiple regression was used as a diagnostic tool to identify the key variables affecting gross fish yields in the existing net cage industry in 1986. The dependent variable was total fish harvest from the net cages in $kg \cdot m^{-2} \cdot year^{-1}$. The independent, explanatory variables were grouped into three major categories: physical, technological and socioeconomic variables (Table 1). In the absence of precise data or, to enable the inclusion of heterogenous data, some of these variables are coded as dummy (or as dichotomous) variables. In addition to the linear form, a multiple log-linear regression was computed, of the form

$$\log Y = \log a + b_1 \log X_1 + b_n \log X_n \qquad ...6)$$

which results to

$$Y = aX_1 ...X_n \qquad ...7)$$

This production function was determined in a straightforward, full regression approach to test hypotheses represented by a combination of predictive variables. The data for these analyses are available in the files CAGE.WK1 and LNCAGE.WK1 (see Appendix II).

For the computation of the morphoedaphic index (MEI), data on total dissolved solids and conductivity were collected during 10 months in 1986. For Lake Sampaloc the MEI was calculated with the following equation:

$$MEI = \frac{\text{conductivity } (\mu S)}{\text{mean depth (m)}} \qquad ...8)$$

Results and Discussion

Limnology of Lake Sampaloc

Lake Sampaloc is a warm monomictic lake, with the average water temperature in the lake varying from 26°C in January to 30°C during April to August. Temperature profiles with measurements at 1- and 2-m intervals down to 20-m depth indicate thermal stratification in the lake from March to October where the anoxic hypolimnion had temperatures of 26 to 27°C and

Table 1. List of variables used for multiple regression analysis of annual gross Nile tilapia production of net cages in Lake Sampaloc, Philippines, in 1986. Details are given in Manrique (1988).

Variable	Unit	Description
Dependent		
HARVEST	$kg \cdot m^{-2} \cdot year^{-1}$	Gross annual fish yield
Physical		
LOCATION	dichotomous, 1 or 2	Geographic location: 1 = NE, SE, SW and 2 = NW portion of lake.
DISTANCE	m	Distance of net cages from shore
DEPTH	m	Water depth at location of net cages
Technological		
SEASON	dummy, 0 or 1	Season in which fish were stocked in cage. 0 = March to May, 1 = June to Feb.
STOCKDEN	$n \cdot m^{-2} \cdot year^{-1}$	Stocking density of fingerlings
FEEDING	dichotomous, 1 or 2	Relative content of crude protein in feeds used. 1 = <10.5%, 2 =>16%
PROCYCLE	months, n	Number of months from stocking to harvest
HARSYS	dummy, 0 or 1	Harvest system. 0 = single-event total harvest, 1 = multiple event partial harvest
FISHSIZE	$n \cdot kg^{-1} \cdot year^{-1}$	Relative individual weight of fish at harvest; pieces per kg
Socioeconomic		
DOMTEN	dummy, 0 or 1	Dominant tenure status. 0 = owner operator, 1 = non-owner operator
EDUC	n	Education of operator: number of years in school
EXPERIEN	n	Experience of operator: number of years in fish culture.

the epilimnion 28.5 to 31.5°C. The thermocline varied from 5 to 14 m over the stratified period with a declining trend from March to September (Fig. 2). Major upwelling and complete mixture of the water column occurred from late November to February characterized by lowest temperatures recorded throughout the whole year (26°C). During the period of study, two further upwelling events were observed, one in March/April and the other in June/July 1980.

GPP was inversely related to the annual temperature trend and the depth of Secchi disk visibility (Fig. 2). During stratified conditions, higher water temperatures and Secchi depths of 1 to 5 m, GPP usually ranged from 0.6 to 1.2 $gC \cdot m^{-3} \cdot day^{-1}$ but reached 1.7 to 2.1 $gC \cdot m^{-3} \cdot day^{-1}$ during small upwelling events. Associated with complete mixing in October to February, GPP increased to above 3 $gC \cdot m^{-3} \cdot day^{-1}$, coincided by low water temperatures of 26°C and low Secchi depths of 0.8 to 1.8 m. The following equation to describe the inverse nonlinear relationship between Secchi depth (X) and GPP was determined (Aquino 1982; Aquino and Nielsen 1983):

$$GPP = 1.94 - 2.13 \log_{10} X \qquad \text{...9)}$$

$n = 25$, $r^2=0.50$, $P < 0.001$, SEE=0.68 $gC \cdot m^{-3} \cdot day^{-1}$.

The upwelling events coincided by with observations of fish performing emergency surface respiration. Throughout the study, olfactory testing of water samples from below the thermocline identified H_2S. The hypolimnial water of eutrophic lakes is often anaerobic after prolonged periods of stratification, frequently containing levels of CH_4, NH_3 and H_2S that are toxic to fish, besides containing higher amounts of nutrients originating from anaerobic decomposition (Ruttner 1975; Wetzel 1975). These are brought to the surface during turnover of the water column and lead either to sublethal effects, e.g., reduced growth or to fish kills (Sutherland 1974). During the study period, several farmers experienced fish kills in the period from August to December. On the other hand, the supply of nutrients also enhances phytoplankton production.

Analysis of Nile Tilapia Growth

The measured weight increments were different for the six conducted series. Roughly, growth increments were low in July, August and September, but reached higher values in the months during and after turnover, i.e., from November to June.

Multiple regression analysis was used to relate gross primary production and average initial weight (Wi) of Nile tilapia fingerlings to specific growth rate (SGR):

$$SGR = \frac{\ln(W_2) - \ln(W_1)}{\Delta t} \quad \text{...10)}$$

where

- W_1 = weight of fish at beginning of growth interval (g);
- W_2 = weight of fish at end of growth interval (g);
- ln = base e logarithm; and
- Δt = duration of growth interval, days.

The derived multiple regression equation takes the form:

$$SGR = 12.54 - 6.50 \log_{10}Wi + 9.16 \log_{10}GPP \quad \text{...11)}$$

with n = 38, R^2 = 0.54, P = < 0.001, SEE=1.6%·day^{-1}

The initial size of the fish (W_i) had to be included in the equation since SGR decreases with increasing fish size. GPP had a positive effect on SGR and this relationship can be used to compute the SGR of a fish population of known size at a certain level of primary production. The obtained value for SGR is reliable only for short-term predictions, however.

Correlations between possible predictor variables and length growth rate showed that, as hypothesized, growth rate decreases with increasing mean length, although this relationship was not found to be significant here (Tables 2 and 3). Both water temperature and GPP were not correlated with length growth rate. In contrast, Secchi disk visibility was significantly (negatively) correlated with length growth rate; the negative sign indicates that high plankton biomasses coincide with high growth rates (Table 3). Further, a significant negative correlation was found between GPP and Secchi depth. H_2S was not correlated with Nile tilapia growth rate.

The ordinary Gulland-and-Holt plot (Gulland and Holt 1959; Pauly et al., this vol.) takes the form (Fig. 3):

$$\Delta L/\Delta t = 0.2123 - 0.0089ML \quad \text{...12)}$$

n = 17, r^2 = 0.209, P = 0.0653, SEE = 0.0543 cm·day^{-1}

with K = 0.0089·day^{-1} and L_∞ = 23.8 cm, where ML is the mean length of the fish.

In terms of the "extended Gulland-and-Holt" method, the inclusion of H_2S as a dummy variable and GPP increased the explanatory power of the equation:

$$\Delta L/\Delta t = 0.203 - \underset{P<0.05}{0.108\,ML} + \underset{P<0.05}{0.378\,GPP} - \underset{P<0.05}{0.0517\,H_2S} \quad \text{...13)}$$

n = 18, R^2 = 0.616, P < 0.001, SEE = 0.0404 cm·day^{-1} with K = 0.0108·day^{-1} and L_∞ = 21.5, 15.8 and 30.0 for average mean, worst and best environmental conditions. H_2S had a stronger correlation with growth rate; the signs were negative in both cases (Table 2). With the "extended Gulland-and-Holt" method, the variable reflecting environmental effects and having the highest correlation with growth rate was Secchi disk visibility (in meters):

$$\Delta L/\Delta t = 0.373 - \underset{P<0.001}{0.0132\,ML} - \underset{P<0.001}{0.0655\,SECCHI} \quad \text{...14)}$$

where the mean and range of the variables was ML = 11.4; 7.0 to 16.3 cm and SECCHI = 1.7; 0.8 to 3.4 m, and n = 17 (a few outliers were excluded), R^2 = 0.839, P < 0.001, SEE = 0.0254 cm·day^{-1} with K = 0.0132·day^{-1} and L_∞ ranging from 11.4 to 24.3 (mean = 19.8 cm). The Durbin-

Table 2. Correlation matrix for data used for derivation of equation (13). n = 18. $\Delta L/\Delta t$ = length growth rate, ML = mean length, H_2S = dummy variable for hydrogen sulfide occurrence.

Variables	$\Delta L/\Delta t$	ML	GPP
ML	-0.4847		
GPP	0.2006	0.4319	
H_2S	-0.5411	0.3657	0.2064

Critical value for 1-tailed test (α = 0.05) = +/- 0.4010
Critical value for 2-tailed test (α = 0.05) = +/- 0.4670

Table 3. Correlation matrix for data used for derivation of equation (14). n = 17. $\Delta L/\Delta t$ = length growth rate, ML = mean length, WATEMP = water temperature in two meters depth, SECCHI = depth of Secchi disk visibility, GPP = gross primary production; all variables (except $\Delta L/\Delta t$) expressed as average values per interval between sampling events.

Variables	$\Delta L/\Delta t$	ML	WaTemp	Secchi
ML	-0.4568			
WATEMP	-0.2810	0.0654		
SECCHI	-0.6437	-0.2661	0.3485	
GPP	0.3954	0.3849	-0.3466	-0.7205

Critical value for 1-tailed test (α = 0.05) = +/- 0.4134
Critical value for 2-tailed test (α = 0.05) = +/- 0.4807

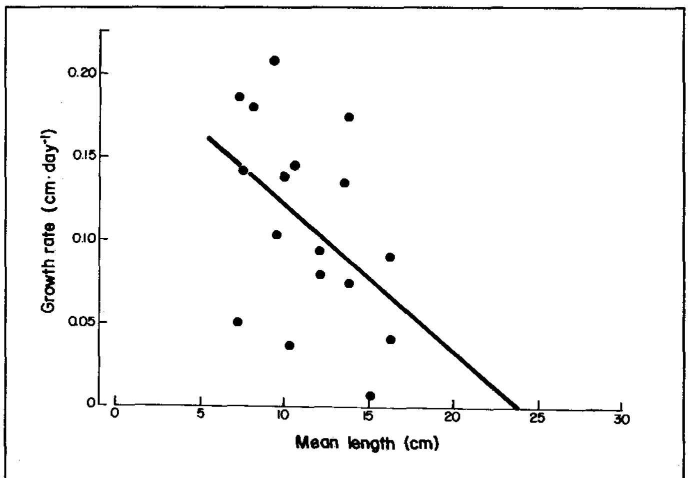


Fig. 3. Gulland-and-Holt plot for growth of mixed-sex *Oreochromis niloticus* in net cages without supplemental feeding in Lake Sampaloc, Philippines; n = 17. See text for regression equation and VBGF parameters.

Watson test showed that the residuals were not autocorrelated (Fig. 4).

Inspite of the close correlation between Secchi depth and GPP, the latter could not be identified as an equally strong predictor for length growth rate in Nile tilapia with this method.

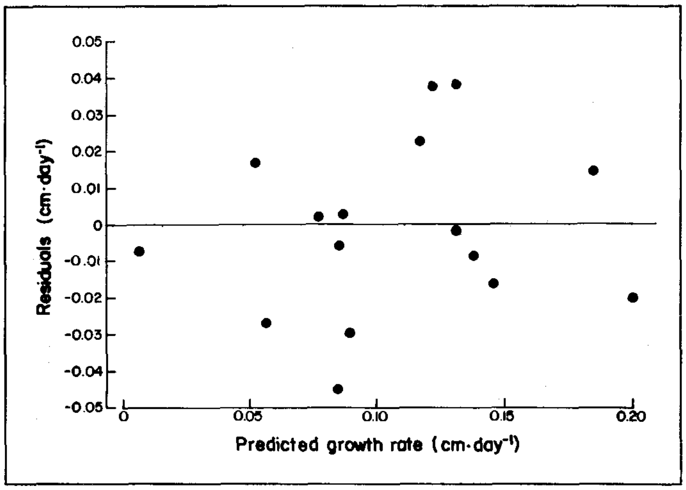


Fig. 4. Plot of residuals (Y) *vs.* predicted values daily growth rate (X) from multiple regression equation derived for *Oreochromis niloticus* grown in net cages in Lake Sampaloc, Philippines; n = 17. Based on equation (10).

The multiple regression equation based on the "extended Gulland-and-Holt" method, which included Secchi depth as a predictor, explained 84% of the total variance in length growth rate of *O. niloticus*. Indeed, Secchi disk visibility turned out to be the best environmental predictor for growth rate. The negative sign of the regression coefficient is plausible since a larger Secchi depth is related to less natural food.

As evident from the table of zero order correlation coefficients, GPP was not significantly correlated with growth rate, in contrast to the findings of Aquino (1982). On the one hand, when included in a multiple regression equation together with mean length and H_2S occurrence, GPP is a significant predictor for Nile tilapia growth rate. In equation (10), H_2S occurrence reflects the main source of negative environmental influence on growth rate. On the other hand, H_2S could not be included in equation (11) as a further predictor variable.

Secchi disk visibility depends on phytoplankton and zooplankton density, on the amount of particulate organic matter and on the density aggregates of detritus, bacteria, and ciliates. All the above contribute to the food spectrum of filter-feeding Nile tilapia, particularly when the fish have no other food source. Therefore, indications of primary production alone do not entirely reflect the food supply to tilapias. The percentage of variance in Nile tilapia growth rate explained by the equation derived by Aquino (1982) was 54% compared to the 84% in equation (14).

In the present analysis, average size computed from three replicates were used to calculate growth rate (the dependent variable), instead of using the original values from each replicate. This was done because all treatments were exposed to the same environmental conditions in the net cages. An inclusion of the original data would have only added to the variance in the dependent variable but would not be complemented by more variance in the predictor variables. Thus, the within-group variance (i.e., between three net cages) was stabilized through averaging, which reduced the amount of bias through sampling errors and natural variability in growth. (A different situation is given in the case of experiments in individual ponds where variability among the predictor variables may also be high).

Water temperature was not a significant predictor variable. Due to low variability over the year (26.0 to 31.5°C), fish size must be monitored more closely to be able to detect an effect, in comparison to food availability, which, in the present case, gave a much stronger signal. This is underlined by the "wrong" sign in Table 1 for the (nonsignificant) correlation between water temperature and growth rate.

To assess the applicability of the derived relationship, theoretical growth curves were computed for each of the six series starting from the sizes actually stocked and using the available and measured values for Secchi disk samplings. Fig. 6 depicts the comparison of observed and predicted growth curves. Negative growth occurred in some cases. This is understood here to reflect weight loss due to unavailability of food. Depending on

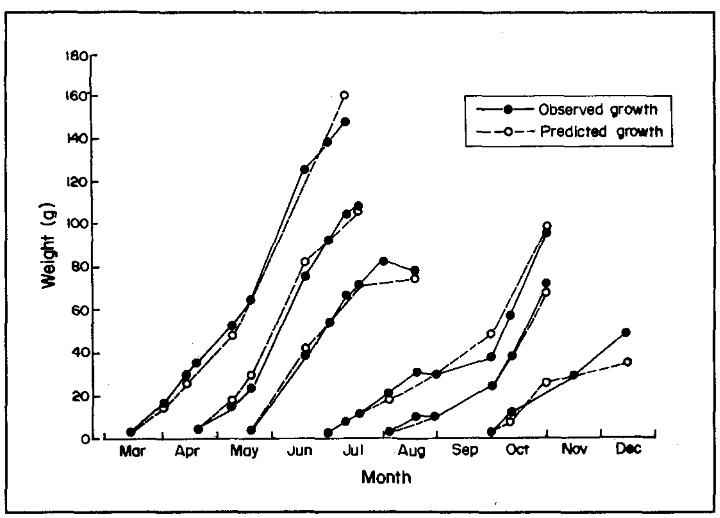


Fig. 5. Comparison of measured and computed growth curves for *Oreochromis niloticus* grown without supplemental feeding in net cages in Lake Sampaloc, Quezon Province, Philippines. Based on the "extended Gulland-and-Holt" method.

environmental conditions, Nile tilapia in net cages reached the minimum market size of 80 g in three to four months.

In accordance with Aquino (1982), the method applied here identified natural food availability as governing the growth of unfed Nile tilapia in net cages. The predictions show the high capability of the method to model growth under varying conditions, given the fact that these effects have been identified and quantified in form of a regression equation. The equation even permits to model weight loss due to adverse conditions. However, the comparisons presented here between actually measured and predicted fish sizes is not a rigorous test of the model. For this purpose, the predictions should be tested on data that were not originally used to develop the regression equation.

Overall, the upwelling of nutrients from the lake hypolimnion controls the annual variability of natural food production, which in turn governs growth of Nile tilapia in Lake Sampaloc. This effect was identified and quantified with a regression model that, in combination with the von Bertalanffy growth function, can be used to forecast fish production in net cages in the lake based on actual environmental conditions, here natural food supply. Negative, growth-retarding effects can originate from upwelling events. If these are slight (i.e., lead to sublethal H_2S concentrations in the cages) but occur frequently over a period of several months, severe reductions in fish growth may result. Aquino (1982) concluded that overgrazing of natural food in the lake and depletion of nutrients from the water column may occur locally and to a limited extent of time. The present study identified upwelling events (i.e., H_2S occurrence) to have a reducing effect on Nile tilapia growth rate. Thus, the best period for Nile tilapia culture in net cages in Lake Sampaloc is from March to November since during the rest of the year fish kills due to upwelling events pose a direct economic threat to the fish farmers. During the culture period, equation (14) can be used to predict Nile tilapia growth and production over the possible period and to design the number of culture cycles.

Analysis of Gross Yield of Nile Tilapia in Net Cages

The multiple linear regression equation of gross annual yields of Nile tilapia in net cages in Lake Sampaloc in 1986 revealed that six of the hypothesized eleven predictor variables were significant ($\alpha = 0.05$):

$$
\begin{aligned}
\text{HARVEST} = -0.438 & - 0.084 \quad \text{DOMTEN} \\
& - 0.397 \quad \text{SEASON} \\
& + 0.842 \quad \text{HARSYS*} \\
& + 0.080 \quad \text{EDUC*} \\
& - 0.094 \quad \text{EXPERIEN*} \\
& + 0.035 \quad \text{DEPTH} \\
& + 0.094 \quad \text{STOCKDEN*} \\
& + 0.191 \quad \text{PROCYCLE*} \\
& - 0.223 \quad \text{FISHSIZE*} \\
& + 0.183 \quad \text{FEEDING} \\
& + 0.252 \quad \text{LOCATION} \quad \text{...15)}
\end{aligned}
$$

with n = 80, R^2 = 0.653, P<0.01,* = significant at α= 0.05.

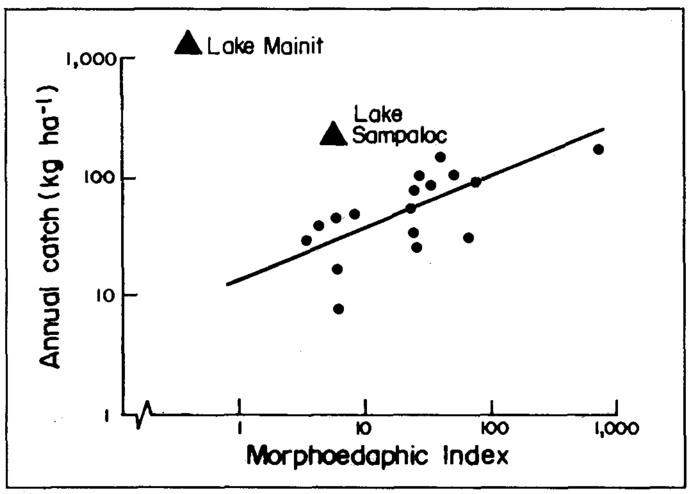


Fig. 6. Relationship between recorded annual catch from capture fisheries and culture operations without supplimental feeding (Y) and morphoedaphic index (X) in 17 African (•) and two Philippine lakes. African data (from Henderson and Welcomme 1974) pertain to lakes with no more than one fisher·km^{-2} and lead to the model $Y = 14.3X^{0.468}$ which predicts catches well below the catch level of Lakes Sampaloc and Mainit. Adapted from Pauly et al. (1990).

The variables "DISTANCE" and "DEPTH" were highly correlated (see Table 4). Therefore only "DEPTH" was used in the regressions.

A second form of the model was fitted using logarithmic transformations which is based on the hypothesis that the explanatory variables behave in a nonlinear fashion. A better fit was noted, R^2 = 0.666, with five variables significant at α = 0.05.

Four of the six technological variables appeared to be significant predictors of gross Nile tilapia production: stocking density; length of the culture period; harvest system and size of fish at harvest. The signs of the regression coefficients indicate that yield is higher at higher stocking densities, longer culture periods, partial harvesting and smaller individual fish size at harvest. The crude protein content of the food given and the season of stocking were not significant (by 1986 most operators were supplying food to their fish).

Of the three socioeconomic variables, education and experience of the net cage operators were significant predictors. The signs reveal that operators with longer school education had higher annual production but those with more experience in fish culture produced less fish on a per area basis. The latter could be caused by the fact that more experienced operators produce fish of larger individual size (which are of higher market value) than the inexperienced operators. On the other hand, education was negatively correlated with experience (although not significantly, r = 0.100), suggesting that less education implies more on-farm experience. Finally, the test of the present hypothesis concludes that whether the operator is the owner of the net cages or an employee has no effect on production.

Only one dummy variable and none of the dichotomous variables became significant predictors.

Table 4. Correlation matrix for data used for derivation of equation (15). Definitions for the variables are given in Table 1; n = 80.

	Variables	Educ.	Exp.	Dist.	Depth	Stock density	Prod. cycle	Fish size at harvest	Location	Stock season	Feeding	Harvest system	Harvest
1	Education	1.000											
2	Experience	-0.100	1.000										
3	Distance	-0.048	0.115	1.000									
4	Depth	-0.058	0.080	0.719	1.000								
5	Stocking density	-0.006	-0.022	0.203	0.094	1.000							
6	Production cycle	-0.077	0.122	-0.147	0.073	-0.144	1.000						
7	Fish size at harvest	-0.298	0.210	0.107	0.302	-0.137	0.166	1.000					
8	Location	-0.342	-0.330	-0.092	-0.048	-0.125	-0.172	0.102	1.000				
9	Stocking season	-0.107	0.077	-0.067	-0.096	-0.306	0.130	0.117	0.082	1.000			
10	Feeding	-0.350	0.125	0.126	-0.040	-0.295	0.004	0.359	0.114	0.397	1.000		
11	Harvest system	-0.164	0.124	0.076	-0.158	-0.144	-0.282	-0.309	-0.242	0.161	-0.133	1.000	
12	Harvest	-0.247	-0.228	0.112	0.075	-0.638	0.004	-0.411	-0.148	-0.295	-0.388	0.116	1.000
13	Dominant tenure	-0.349	0.368	-0.097	-0.196	-0.106	-0.018	-0.426	-0.504	0.026	-0.150	-	-
14	Residual	-0.000	-0.187	-0.078	0.000	0.000	0.000	0.000	-0.057	0.000	0.000	-	-

None of the physical variables hypothesized here turned out to be significant predictors in this combined set of eleven predictor variables.

MEI and Gross Annual Fish Production in Lake Sampaloc

Manrique (1988) estimated the annual production of the capture fishery in Lake Sampaloc in 1986 at 428 $kg \cdot ha^{-1} \cdot year^{-1}$. This comprised fish traps or pens, gill netting and hook and line activities landing mainly tilapia and a silvery theraponid (local name: *ayungin*). The Nile tilapia culture operations in the net cages produced 540 $kg \cdot ha^{-1}$. This leads to a total fish harvest of 968 $kg \cdot ha^{-1} \cdot year^{-1}$ of which 44% was produced by the capture fisheries and 56% by net cage culture operations. Thus total fish output from the lake was 99.2 t or 968 $kg \cdot ha^{-1}$ in 1986. An MEI of 9.0 was estimated for Lake Sampaloc (Manrique 1988). This is based on a mean depth of 17.47 m and values for conductivity ranging from 0.2 to 0.9 µS.

A comparison of the MEI of Lake Sampaloc with that of African lakes reveals that fish production is much higher in this Philippine lake. A similar result was obtained by Pauly et al. (1990) for Lake Mainit, another Philippine Lake (Fig. 6). This theme is not followed upon here because more than two points from Philippine Lakes are needed for detailed comparison with African lakes.

Acknowledgements

L.V. Aquino-Nielsen acknowledges Bent Henning Nielsen who sparked the idea of relating environmental parameters to growth of tilapia in the lake and Flemming Petersen who helped in the statistical analysis of the data.

A.S. Manrique-Pempengco is thankful to IDRC for funding this thesis research through the UPLB-Renewable Resource Economics and Management Research Program. Appreciation is also extended to Dr. Pauly of ICLARM, and Dr. Chona Cruz and Dr. Jimmy Valera of UPLB for having participated in this thesis work as Guidance Committee members.

The contribution of Mark Prein was supported in part by GIARA - the Germany-Israel Agricultural Research Agreement for Developing Countries, which is funded by the Bundesministerium für Wirtschaftliche Zusammenarbeit (BMZ).

References

Aquino, L.V. 1982. Some ecological considerations relative to cage culture of *Tilapia nilotica* in Sampaloc Lake. University of the Philippines in the Visayas, Iloilo City, Philippines, 77 p. M.S. thesis.

Aquino, L.V. and B.H. Nielsen. 1983. Primary production and growth rate of *Tilapia nilotica* in cages in Sampaloc Lake, Philippines, p. 455-463. *In* L. Fishelson and Z. Yaron (comps.) Proceedings of the International Symposium on Tilapia in Aquaculture, Nazareth, 8-13 May 1983. Tel Aviv University, Tel Aviv, Israel.

Draper, N.R. and H. Smith. 1981. Applied regression analysis. Second edition. New York, John Wiley. 709 p.

Fermin, P.D. 1978. Culture of *Tilapia mossambica* or *T. nilotica* in fish cages (fish pen style) in Laguna de Bay. Bureau of Fisheries and Aquatic Resources, Quezon City, Philippines.

Gaarder, T. and H.H. Gran. 1927. Investigations of the production of plankton in the Oslo Fjord. Rapp. P.-V. CIEM 42:1-48

Gulland, J.A. and S.J. Holt. 1959. Estimation of growth parameters for data at unequal time intervals. J. Cons. CIEM 25:47-49.

Henderson, H.F. and R.L. Welcomme. 1974. The relationship of yield to morphoedaphic index and number of fishermen in African fisheries. Committee on Inland Fisheries of Africa. Occas. Pap. 1. FAO, Rome. 19 p.

Hopkins, K.D., M.L. Hopkins and D. Pauly. 1988. A multivariate model of tilapia growth, applied to seawater tilapia culture in Kuwait, p. 29-39. *In* R.S.V. Pullin, T. Bhukaswan, K. Tonguthai and J.L. Maclean (eds.) The Second International Symposium on Tilapia in Aquaculture. ICLARM Conf. Proc. 15, 623 p.

Manrique, A.S. 1988. Simple productivity indices for lake ecosystems: the case of Sampaloc lake. University of the Philippines, Los Baños, 121 p. M.S. thesis.

Mosteller, F. and J.W. Tukey. 1977. Data analysis and regression, a second course in statistics. Reading, Massachusetts, Addison-Wesley. 588 p.

Pauly, D. and K.D. Hopkins. 1983. A method for the analysis of pond growth experiments. ICLARM Newsl. 6(1):10-12.

Pauly, D. and J. Ingles. 1981. Aspects of the growth and mortality of exploited coral reef fishes, p. 81-98 *In* E.D. Gomez, C.E. Birkeland, R.W. Buddemeier, R.E. Johannes, J.A. Marsh and R.T. Tsuda (eds.) The reef and man. Proceedings of the Fourth International Coral Reef Symposium, 18-22 May 1981, Manila, Philippines. Vol. I, 725 p. Marine Sciences Center, University of the Philippines, Quezon City.

Pauly, D., M. Small, R. Vore and M.L. Palomares. 1990. Fisheries yields and morphoedaphic index of Lake Mainit, Philippines. p. 835-838. *In* R. Hirano and I. Hanyu (eds.) The Second Asian Fisheries Forum. Asian Fisheries Society, Manila. 991 p.

Ruttner, F. 1975. Fundamentals of limnology. Third edition. University of Toronto Press, Toronto. 307 p.

Ryder, R.A. 1965. A method for estimating the potential fish production of North temperate lakes. Trans. Am. Fish. Soc. 94:214-218.

Sutherland, B.G. 1974. Preliminary report on the 1973-1974 limnological investigation of Sampaloc Lake. Bureau of Fisheries and Aquatic Resources, Quezon City, Philippines.

Talavera, F. 1932. The fisheries of Lake Sampaloc, San Pablo, Laguna Province. Philipp. J. Sci. 50(1):1-49.

Vollenweider, R.A., Editor. 1974. A manual on methods for measuring primary production in aquatic environments. IBP Handbook No. 12. Blackwell Scientific Publications, London. 213 p.

Wetzel, R.G. 1975. Limnology. Saunders, Philadelphia. 743 p.

Appendix I

A Method for the Analysis of Pond Growth Experiments*

DANIEL PAULY and KEVIN D. HOPKINS
ICLARM

Pond growth experiments play an essential role in aquaculture research in assessing the growth potential of the various species and strains in all those experiments needed to transform aquaculture from the art it is now into the *science* it should become.

A major problem with pond growth experiments is the extreme difficulty of effectively controlling not only the "control" variables (e.g., food supply to the fish in integrated pig-fish culture experiments) but also extraneous variables (e.g., climatic factors) capable of affecting the results. Because they are rather costly in time and resources, pond growth experiments are often not replicated sufficiently, leading to results of dubious validity.

Usually, growth experiments are run for a set period, at the end of which the total yields are compared with that of a control to infer the effect of the treatments. Such treatments include:

- different stocking rates
- different stocking sizes
- different feeds
- different strains of a given species

In polyculture systems, different treatments include:

- different species ratios (at stocking time)
- different predator species for a given prey species

while in integrated systems (e.g., pig-fish), there are:

- different sizes (or numbers) of pigs, and
- different forms of transfer of pig wastes to the ponds

Additionally, nature itself and the vagaries of life may provide such "treatments" as:

- floods that wash the fish out of a few ponds
- ponds with different bottom type and productivity
- pumps that break down with all fish dying one month before harvesting
- all fish stolen, one week before harvesting

Since they can't deal with all these problems at the same time, aquaculturists have tended to concentrate on one or two of the variables believed likely to affect yields.

Such experiments need a lot of ponds. For example, a set of four treatments (+ control) with five replicates requires 25 ponds. Indeed, experimental designs based on the analysis of final yields—the black-box approach, see Fig. 1—are essentially wasteful of time and other resources because they make no use of the information that can be extracted from the *growth*

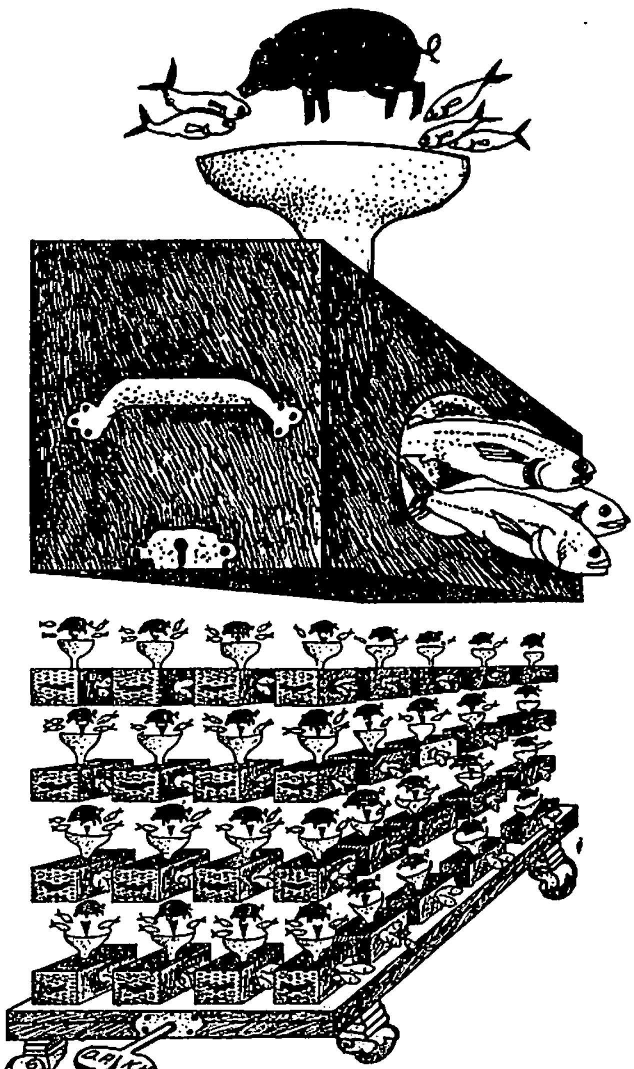


Fig. 1. The traditional method to analyze a pond growth experiment is based on a big *black box,* with a set of inputs, and one output (final yield); our method opens the box and uses the large number of *black boxes* (and their inputs and output) that can be obtained by breaking up the overall growth of the stocked fish into a number of growth increments.

*Reprinted from the ICLARM Newsletter 6(1):10-12(1983).

process which leads to the final yields.

The final yield of a growth experiment can be conceived as the sum of a number of *growth increments* (Fig. 2), as could be assessed by weighing fish in a given pond at regular intervals. Moreover, the final yield is the sum of growth increments of individual fish, each of which can also be conceived to have grown incremently, as also shown in Fig. 2. These two features of the yield of an aquaculture experiment have led us to propose a new method of conducting and analyzing pond growth experiments.

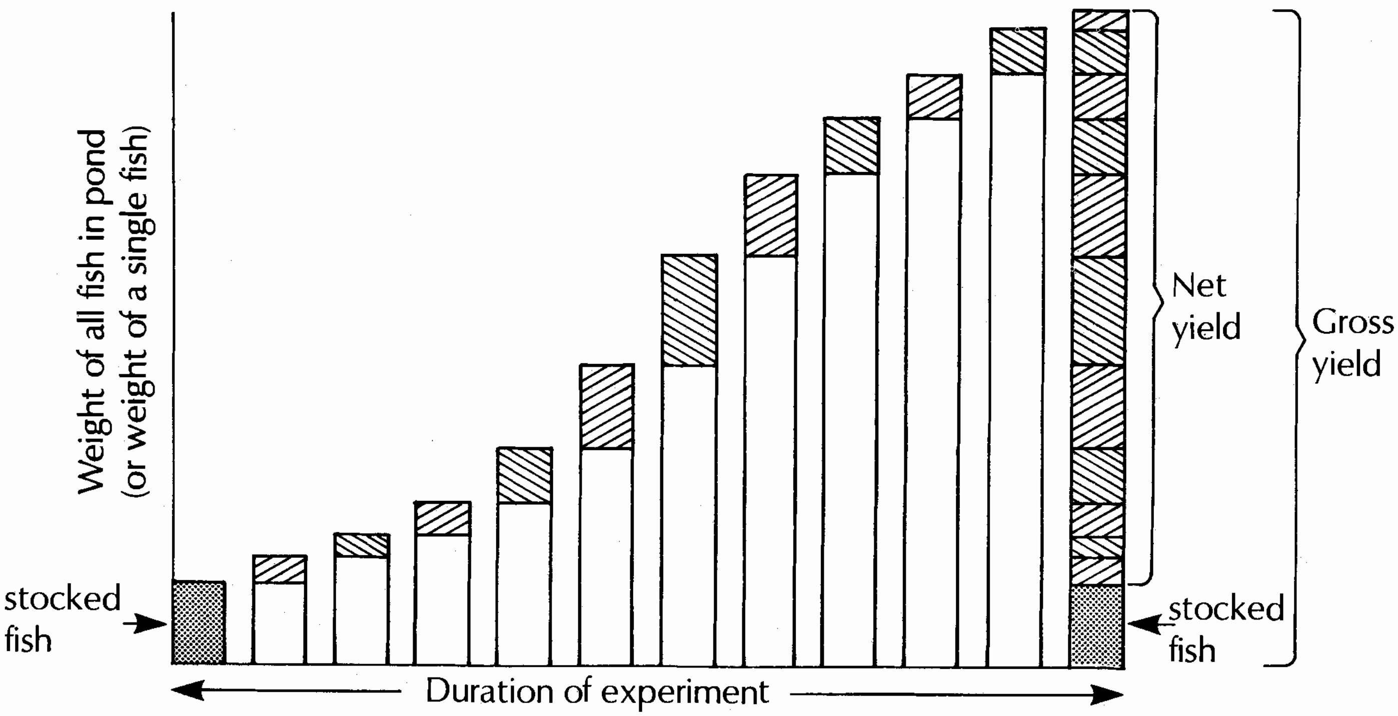


Fig. 2. The final yield of a growth experiment can be viewed as the sum of a number of growth increments, either of a single fish or the whole population of the pond.

The Method

All that is required is to measure the fish in the pond(s) preferably (but not necessarily) on a regular basis. Sometime between fish measurements, measure the values of those variables you think are likely to affect fish growth. Use dummy variables for items that cannot be quantified (e.g., 0 for ponds in site A, 1 for ponds in site B).

Next, calculate the mean growth increments of the fish per day ($\Delta L_i/\Delta t_i$) for each time interval and tabulate these against the mean length ($\bar{L}_i$) and the values of the corresponding variables. Use all tabulated values to estimate the parameters of a multiple regression as shown in box, p. 12. The number of variables which can be included in the analysis is limited in principle only by the available number of $\Delta L_i/\Delta t_i$ and $\bar{L}_i$ values: the more frequently the fish in the experiment have been measured, the more data sets will be available for the multiple regression. In practice, the number of variables will be limited to those that can be monitored concurrently with the growth of the fish.

An Example

From 1978 to 1981, pond growth experiments were conducted in a cooperative project on animal-fish culture between Central Luzon State University and ICLARM. The project experiments involved fish grown with pigs or chickens, the fish consisting of various combinations and stocking rates of tilapia *(Oreochromis niloticus)*, carps *(Cyprinus carpio)* and a predator *(Channa striata)*. Different numbers of pigs and chickens were also used, and, as the animals grew during the course of the various experiments, their inputs (fecal matter and urine) differed within and between the various growth experiments. Moreover, as is always the case with outdoor experiments, climatic factors (such as rain, wind, light, affecting temperature and oxygen) changed within and between experiments, not to mention those factors (e.g., floods) which caused experiments to be interrupted prematurely (for a detailed narrative, see "The ICLARM-CLSU Integrated Animal-Fish Farming Project: Final Report" by K.D. Hopkins and E.M. Cruz. 1982. ICLARM Technical Reports 5.)

Altogether, 117 experiments were completed. A very large number of variables were, in the course of the, experiments, hypothesized to affect yields. No conventional experimental design existed with which, using the yield values available, one could have tested the effect of such a large number of variables.

Throughout the experiments, however, samples of fish had been seined and measured at biweekly (generally) intervals, from which mean values of $\Delta L_i/\Delta t_i$ and L_i could be obtained; these measurements also provided a framework for the computation and tabulation of mean values of the variables pertaining to each of the separate intervals Δt. In this fashion, 713 data sets were obtained.

Four of the hypothesized variables turned out to have a significant impact on the growth increments, as shown in the Table. Of these, mean length contributed most of the explained variance (31%) while pig manure, tilapia biomass and tilapia recruitment contributed to a lesser extent (9.4%, 6.0% and 2.4%, respectively).

Advantages and Potential

The advantages of the method proposed here are, we believe, five-fold:

- it uses more of the data generated during growth experiments;
- it replaces the rather inflexible analysis of variance generally used for pond growth experiments by a much more versatile and powerful method. Multiple regression allows (i) analysis of residuals to test for departures from linearity of the equation (4); (ii) use of dummy variables for non-quantifiable effects; and (iii) use of *beta* coefficients to compare the effects on growth of variables expressed in different units (see Table);
- it allows for a linkup of the results obtained in growth experiments with growth models used in the general field of fishery biology and population modelling;
- it can be used for any fish production experiments in which there are many variables which influence the results: not just integrated farming/pond experiments;
- it permits a new approach to designing experimental aquaculture facilities, since it offers an alternative to replication of treatments.

Equation (1) is a differential form of the von Bertalanffy growth equation, while equation (3) is also a method for estimating the parameters K and L_∞ of that equation.

The method proposed might thus help, in addition to facilitating analysis of aquaculture data, to bridge the gap separating fishery biologists working with wild populations and aquaculturists working with confined populations.

A longer paper, covering in detail the various aspects of the new methods will be submitted for publication in a scientific journal in due course. ●

Variables significantly affecting growth rate of tilapias in pig-fish growth experiments, 1978 to 1981.

Variable	Slope	Standardized or *beta* slope[a]	Variance explained (%)
Length of tilapias	—0.0111	0.3753	30.52
$\log_e$ pig manure input	+0.0223	+0.2930	9.36
$\log_e$ total weight of tilapias	—0.0550	—0.2940	6.02
$\log_e$ tilapia recurits	—0.0118	—0.3052	2.35

[a] *Beta* coefficients are standardized slopes which allow comparison of variables expressed in different units. Thus, in the present case, it can be assessed that tilapia recruitment and tilapia biomass have as much *negative* effects on tilapia growth as manure input has a positive effect.

Derivation

The new method makes the assumptions (1) that mortality of fish stocked is nil or negligible, such that the final yield is (at least approximately) equal to the number of fish stocked times the mean weight of fish at harvesting; and (2) that the growth rate of fish (in length) decreases linearly as the fish get larger, expressed by

$$dl/dt = a+bL \qquad (1)$$

where a and b are empirically determined constants.

The validity of the first assumption is easily assessed in a given set of experiments and requires no further comment. The second is known to apply to most fish past their fingerling stage. The relationship between length (L) and weight (W) of fish is generally proportional to length raised to a power of between 2.5 and 3.5, generally close to 3; weight growth data can always be rendered approximately proportional to length-growth data by taking the cubic root of weight ($\sqrt[3]{W}$).

These concepts imply that the final yield of a growth experiment can be viewed as a function (f) of the length-growth increments of the fish in the pond, i.e.,

$$Y = f\,[dl/dt] \qquad (2)$$

The *differential* equation (1) can however, for short time increments, be replaced by the *difference* equation

$$\frac{\Delta L_i}{\Delta t_i} \approx a + b\bar{L}_i \qquad (3)$$

where ΔL is a length increment, i.e., the difference between the length at the beginning (L_1) and end (L_2) of a given time period Δt_i, while $\bar{L}_i$ is the mean of the L_1 and L_2 values.

From equation (2), any factor increasing the $\Delta L_i/\Delta t_i$ values of a given growth experiment will increase the yield at harvesting. Equation (3) suggests that it is only length itself which affects the $\Delta L_i/\Delta t_i$ values; however nothing prevents us from expanding the regression equation (3) into a multiple regression of the form

$$\Delta L_i/\Delta t_i = a + b\,\bar{L} + b_1 X_1 + b_2 X_2 \ldots b_n X_n \qquad (4)$$

where a number of variables ($X_1, X_2 \ldots X_n$) are conceived as affecting growth rates and hence yield at harvesting. Therefore, given measurements of the variables $X_1, X_2 \ldots X_n$ likely to affect growth rates and pertaining to the time periods for which the $\Delta L_i/\Delta t_i$ and $\bar{L}_i$ values apply, those variables affecting growth (and hence yields) can be identified as those that have slopes ($b_1, b_2 \ldots b_n$) significantly different from zero, while the values of the slope quantify the effects. (The parameters of equation (4) may be estimated using any standard program for multiple regression.)

Letter to the Editor*

The article "A method for the analysis of pond growth experiments" by D. Pauly and K.D. Hopkins promted letters by Drs. M. Pedini, Fishery Resources Officer at FAO and L. Lovshin, Auburn University, U.S.A. Dr. Lovshin's letter, slightly abbreviated, is reprinted here; Dr. Pedini's comments and questions were virtually the same. Dr. Pauly's answer follows.

Dear Dr. Pauly:

As so often happens, aquaculture researchers are very limited in their knowledge of statistics and mathematicians are limited in the nuances of pond culture. As a researcher and pond culturist perhaps I can give you some insight into the problems of pond culture research and sampling that should be given consideration.

1. I don't believe you can assume that mortality is nil because we always have some mortality. Of course, mortality can be determined at final harvest when the pond is drained. I would assume that your samples can then be corrected for mortality. The problem is to determine when the mortality occurs. Unfortunately, we don't always know when mortality occurs because we can't always see the dead fish especially when the fish are small and/or predators are abundant.
2. We don't really know the effects of sampling on the growth of fish. How frequently can we sample before growth is adversely affected? Some fish can be sampled twice a week without any apparent effects. Other species can't be handled once a month without the risk of mortality or growth reduction.
3. I have found it very hard to get a good estimate of growth and yield by sampling ponds with a net for many species of fish. First, many species are very good at avoiding a seine even in a small pond. Often, you can't get a good sample without repeated seining which can cause water quality problems if the pond bottoms are rich in organic matter. Second, even a decent sample, 20% of your population, can be very misleading because of size bias. Often, you get a high percentage of large or small individuals because of their ability to escape capture. For example, a day before pond draining, I have sampled a number of tilapia ponds with a seine, capturing a normal 20% of the population. After draining, a comparison of the sample average fish weight and harvest average weight was made. I often had a difference of 20%. The best I ever got was a 4% difference.
4. Your statement that a "new approach to designing experimental aquaculture facilities is offered since an alternative to replication of treatments is available", is misleading. No matter what the type of statistical analyses used we are always better off to replicate. Your method may explain what is the cause of variance in a fish pond but it does not eliminate the variance. If you use only one pond per treatment and have high mortality due to an O_2 deficiency how can a good comparison of yield between 2 treatments be made? Even if analysis of variance is not used but some form of analysis of curves, I believe replication is wise to give more secure answers.

We need the help of interested mathematicians to design research that can determine some of the sampling biases I have mentioned. Most tilapia species are ideal experimental animals but should not be used as the standard for all aquatic animals. Some good culture species are not born to be excessively handled and sampled. Good luck, your work is cut out for you.

Sincerely,
Dr. Leonard L. Lovshin
Associate Professor
International Center for Aquaculture
Auburn University
36849, Alabama, U.S.A.

*Reprinted from the ICLARM Newsletter 7(2):30 (1984).

Reply to Dr. Lovshin's letter*

First of all, I would like to thank Dr. Lovshin for his comments on our Newsletter article. Also, I have to mention here that my answer to his questions should not be perceived as a 'rebuttal' but rather as a late statement of things which should have been included in our article but were not, due to lack of space and sheer negligence. I now answer Dr. Lovshin's points in turn:

1. *Actually, when we wrote that mortality must be nil as one of the assumptions that must be met for our method to be applicable, we took a shortcut which we hope would help the reader concentrate, when assessing our method, on cases where mortality is indeed negligible. In fact, the method can be used when mortality is non-negligible, granted it is not linked to any factor affecting growth, i.e., not caused by any of the treatments. Put differently, variables which the method identifies as enhancing growth will also enhance yields if they are not simultaneously the cause of mortalities.*
2. *I agree that sampling itself will have an effect on growth, and this effect should generally be negative. On the other hand, there is nothing preventing us from using sampling frequency itself as a variable and hence to eliminate its effect on the estimated statistics. For example, in coral fish caught with traps, and whose growth is negatively affected by trapping, the application of our method allowed both the identification and the removal of the effects of trapping on the fish whose growth was studied.***
3. *There is obviously no simple solution to the problem of obtaining representative sizes from pond samples, and it will be the task of the investigator concerned to choose a sampling gear which is as non-selective as possible.*
4. *With regard to replication, let me state that if one pond per treatment is used and there is a high mortality in one of the ponds, the method we propose allows use of the growth increments up to the time when the mortality occured-which isn't possible when comparing final yields only.*
   I concede that it was mistaken to suggest that "no replication is needed", because something equivalent to replication is needed when multiple linear regression models are used, namely that (a) each variable considered must be represented by a wide range of values and (b) that the variables considered must not vary together (these two requirements imply a 'replication' of some sort, but different from what is required e.g., in ANOVA).

We were aware, when we wrote our article that a more detailed presentation will be needed for the advantages of his new method to be fully appreciated, but went ahead with a preliminary publication because we assumed that some readers of the ICLARM Newsletter would be interested in the preliminary version. As it turned out, several colleagues after reading the article have expressed interest in the method, and in designing their pond growth experiments such that they will be able to apply and test it. We have decided to await the results of these tests, and of an in-depth analysis of the experiments conducted in the original ICLARM-supported project [...] to prepare a comprehensive paper on the method.

Daniel Pauly

*Reprinted from the ICLARM Newsletter 7(2):30 (1984).
**D. Pauly and J. Ingles. 1981. Proc. 4th Int. Coral Reef Symp., Manila.

Postscript (Dec. 1992): This book is the above-mentioned "comprehensive paper."

Appendix II

Documentation of Available 5-1/4" MSDOS Data Diskettes on the Analyzed Data

All data used for the analyses presented in this book are contained in 24 files on four diskettes. The files are in Lotus 1-2-3 format and have the extension ".WK1". The diskettes are in MS-DOS format (5-1/4", high density, 1.2 MB capacity).

The diskette numbers, file names, file sizes, and contents are given in Table 1.

Table 1. Contents of the four data diskettes containing 24 data files analyzed in the contributions in this book. Authors of contributions are given.

Filename	Filesize (KB)	Author(s)
Disk # 1		
PHILALL.WK1	577427	Prein
PHILPROD.WK1	81365	Milstein & Prein
LIMA.WK1	26378	Prein
ZAMB.WK1	35501	Prein
SAMPALOC.WK1	9534	Aquino-Nielsen, Manrique-Pempengco & Prein
CAGE.WK1	20762	Aquino-Nielsen, Manrique-Pempengco & Prein
LNCAGE.WK1	20762	Aquino-Nielsen, Manrique-Pempengco & Prein
RPMORTAL.WK1	10417	Hopkins & Pauly
REDTILAP.WK1	10784	Hopkins & Pauly
TRENDALL.WK1	22713	Costa-Pierce, Van Dam & Kapeleta
OSHIRAN.WK1	25294	Costa-Pierce, Van Dam & Kapeleta
NIL.WK1	323833	Mair & Pauly
Disk # 2		
NIL3.WK1	467164	Mair & Pauly
AUR1.WK1	323352	Mair & Pauly
AUR2.WK1	323287	Mair & Pauly
DORGROW.WK1	51580	Prein
DORPROD.WK1	34042	Hulata, Milstein & Goldman
Disk # 3		
AUR3.WK1	645864	Mair & Pauly
AUR4.WK1	429986	Mair & Pauly
POND.WK1	25781	Milstein & Hulata
Disk # 4		
FARMSALL.WK1	360733	Prein & Milstein
ISRAPRO1.WK1	227067	Milstein & Hulata
ISRAPRO2.WK1	223177	Milstein & Hulata
ISRAPRO3.WK1	173069	Milstein & Hulata
PHILSAMP.WK1	211421	Prein & Pauly

Experiments at Central Luzon State University, Philippines

Filename: *PHILPROD.WK1*

The raw data from Philippine experimental fish culture summaries used in the factor and canonical analyses (paper by Milstein and Prein) is contained in the file PHILPROD.WK1. This contains the fish stocking and harvest summaries from Appendix B of Hopkins and Cruz (1982) and from unpublished raw data (Experiments 13 and 16). A description is given in Table 2.

Table 2. Fish stocking and harvest summaries of Nile tilapia culture experiments conducted at the Central Luzon State University (CLSU) in Muñoz, Philippines, by Hopkins and Cruz.

Variable	Description
Expt No.	Experiments 1-6, 8-12, 14, 18, from Appendix B of Hopkins and Cruz (1982). See Chapter 2 for experiment descriptions. Experiments 13 and 16 are additional unpublished data retrieved from data sheets from incomplete experiments.
Pond No.	Ponds 1-12 are 1,000 m^2, ponds 13-24 are 400 m^2.
Days	Length of culture period in days (not always consistent with Appendix C).
Livestock	Dummy variables of livestock/fertilizer type (0 = no, 1 = yes): IF = inorganic fertilizer, P = pig, D = duck, C = chicken. Number of animals per hectare.
On-ST n/ha	*Oreochromis niloticus* at stocking (number of fish initially stocked per hectare).
On-ST Wi_g	*Oreochromis niloticus* (average weight of individual fish at stocking in grams).
On-ST kg/ha	*Oreochromis niloticus* (total biomass initially stocked, $kg{\cdot}ha^{-1}$)
On-HRV n/ha	*Oreochromis niloticus* at harvest (number of fish harvested per hectare).
On-SVL %	*Oreochromis niloticus* survival over entire experiment period in per cent.
On-HRV kg/ha	*Oreochromis niloticus* at harvest (total biomass at harvest, excluding wild spawn/recruits, $kg{\cdot}ha^{-1}$).
On-HRV Wi_g	*Oreochromis niloticus* (average weight of individual fish at harvest in grams).
On-YLD kg/ha	*Oreochromis niloticus* net yield over experiment period, kg/ha.
On-dYLD	*Oreochromis niloticus* net daily yield, $kg{\cdot}ha^{-1}{\cdot}day^{-1}$.
On-RS kg/ha	*Oreochromis niloticus* reproduction (biomass of wild spawn/recruits captured and removed during sampling, $kg{\cdot}ha^{-1}$).
On-RH kg/ha	*Oreochromis niloticus* (biomass of wild spawn/ recruits captured at harvest, $kg{\cdot}ha^{-1}$).
OnTOTR kg/ha	*Oreochromis niloticus* (total biomass of recruits captured and removed during sampling and harvest, $kg{\cdot}ha^{-1}$). In Experiment 2, no data regarding recruits were contained in the records so zero was entered into this table.
Cc	*Cyprinus carpio*, for further description see above.
Cs	*Channa striata* (*Ophicephalus striatus*), for further description see above.
Market-Y	Net yield over experiment period of all market size fish (all species), $kg{\cdot}ha^{-1}$.

(Continued)

Table 2. (Continuation)

Variable	Description
Market-dY	Net daily yield of all market size fish (all species), $kg \cdot ha^{-1} \cdot day^{-1}$.
Total-Y	Total net yield over experiment period of market sized and undersized fish (all species) plus wild spawn/recruits of *O. niloticus*, $kg \cdot ha^{-1}$.
Total-dY	Total daily net yield of market sized and undersized fish (all species) plus wild spawn recruits of *O. niloticus*, $kg \cdot ha^{-1} \cdot day^{-1}$.
Livestock	Average livestock biomass over experiment period, $kg \cdot ha^{-1}$.
dryMan-SUM	Cumulative total of manure (dry weight) added to ponds during experiment ($kg \cdot$dry matter$\cdot ha^{-1}$). This can be converted to nitrogen, phosphate, potash, fiber, or biochemical oxygen demand by multiplying with the appropriate constant (Table 3.2 in Hopkins and Cruz 1982). *** Inorganic fertilizer: Experiment 1, ponds 1, 5, 8, 11: the nutrient input levels for inorganic fertilizer were 0.53 kg N/day and 0.67 kg PO_4/day (i.e. 50kg of 16-20-0, N-P K, every 14 days).
dryMan	Daily application rate of manure (dry weight) added to ponds ($kg \cdot ha^{-1} \cdot day^{-1}$). *** Inorganic fertilizer: see above.
D.O. mg/l	Average early morning (6:30 A.M.) dissolved oxygen content in pond water during experiment period ($mg \cdot l^{-1}$).
Temp °C	Average early morning (6:30 A.M.) water temperature during experiment period (degrees Centigrade).
Tgrowth	Mean daily tilapia growth rate ($g \cdot day^{-1}$).
Cgrowth	Mean daily common carp growth rate ($g \cdot day^{-1}$).
Chgrowth	Mean daily *Channa striata* growth rate ($g \cdot day^{-1}$).
Lstype	Livestock type. 0= nothing, 1= inorganic fertilizer, 2= chicken, 3= duck, 4= pig.

REMARKS

1) Experiment 13: the animal biomass and manure output was estimated according to data in Hopkins and Cruz (1982), p. 10.
2) Experiment 16: the only data available for ZERO treatment (neither manure nor inorganic fertilizer applied, i.e. baseline for natural productivity of ponds). Only three ponds are given here since others were disrupted by typhoon. Further ZERO treatment experiment (Expt. 7) was also terminated by typhoon shortly after begin (no data available).

Filename: *PHILALL.WK1*

Beyond the stocking and harvest data, this file contains recordings of all intermediate sampling events of fish sizes, and environmental and water quality variables. In their final report, Hopkins and Cruz (1982) published the data of 25 variables. Here additional 14 unpublished variables from raw record sheets are included. Table 3 gives the descriptive statistics for all variables, including those additionally generated and described in the paper by Prein.

Table 3. List of 65 variables and descriptive statistics for the dataset in the file PHILALL.WK1 from ICLARM-CLSU experiments conducted at Muñoz, Philippines. START and END refer to culture intervals.

VARIABLE	MEAN	STD DEV	MINIMUM	MAXIMUM	n	LABEL
VAR01			1.0	14.0	713	EXPERIMENT NUMBER
VAR02			1.0	24.0	713	POND NUMBER
VAR03					713	START DATE OF INTERVAL
VAR04					713	END DATE OF INTERVAL
VAR05	15.8	4.8	6.0	46.0	713	DAYS OF INTERVAL
VAR06	12.581	4.313	2.74	24.01	690	TILAPIA LENGTH START OF INTERVAL cm
VAR07	14.213	3.686	4.77	24.03	698	TILAPIA LENGTH END OF INTERVAL cm
VAR09	10757.604	4379.192	150.0	21100.0	713	TILAPIA HARVESTED n/ha
VAR10	708.105	693.635	0.31	3804.0	713	TILAPIA BIOMASS kg/ha
VAR11	160.241	126.079	0.0	865.0	713	CARP BIOMASS kg/ha
VAR57	0.034	0.180	0.0	1.0	713	DUMMY LIVESTOCK TYPE INORG. FERTILIZER
VAR58	0.617	0.486	0.0	1.0	713	DUMMY LIVESTOCK TYPE PIG
VAR59	0.168	0.374	0.0	1.0	713	DUMMY LIVESTOCK TYPE DUCK
VAR60	0.177	0.382	0.0	1.0	713	DUMMY LIVESTOCK TYPE CHICKEN
VAR13	681.015	1422.599	0.0	10000.0	713	LIVESTOCK NUMBER ANIMALS/ha
VAR14	3448.610	2502.315	0.0	11403.0	587	LIVESTOCK ANIMAL BIOMASS kg/ha
VAR15	68.734	37.960	0.0	221.0	713	MANURE DRY MATTER kg/ha/day
VAR16	4165.951	3447.211	0.0	14702.0	713	MANURE CUMULATIVE DRY MATTER kg/ha/n days
VAR17	1934.174	1148.615	0.0	6273.0	713	MANURE RESIDUAL LAST 30 DAYS BEFORE
VAR18	2681.049	1757.498	0.0	14556.0	713	MANURE RESIDUAL LAST 45 DAYS BEFORE
VAR19	1.546	1.041	0.0	8.05	713	MANURE N kg/ha/day
VAR20	1.342	1.453	0.0	10.16	713	MANURE TOTAL P2O5 kg/ha/day
VAR21	0.644	0.644	0.0	4.64	713	MANURE K2O kg/ha/day
VAR22	14.179	7.885	0.0	42.55	713	MANURE FIBER kg/ha/day
VAR23	8.827	6.826	0.0	47.29	713	MANURE BOD5 kg/ha/day
VAR24	26.248	1.827	20.5	29.7	713	AM WATER TEMPERATURE Celsius
VAR25	436.184	141.633	44.0	675.5	713	BRIGHT SUNSHINE min/day
VAR26	361.831	73.082	133.0	632.6	713	SOLAR RADIATION Langleys/day
VAR27	4.738	5.869	0.0	21.5	713	RAINFALL mm/day
VAR28	111.635	43.512	32.0	275.1	713	WIND CUMULATIVE km/day
VAR29	780.365	289.238	400.0	1000.0	713	POND SIZE square meters
VAR30	0.798	0.402	0.0	1.0	713	POND AGE NEW = 0 OLD = 1

Continued

Table 3. (Continuation)

VARIABLE	MEAN	STD DEV	MINIMUM	MAXIMUM	n	LABEL
VAR31	2.389	1.732	0.05	7.4	713	AM DISSOLVED 0XYGEN mg/l
VAR32	52.767	42.455	0.0	100.0	713	PERCENT DAYS AM DO <2.0 mg/l
VAR33	30.954	37.551	0.0	100.0	713	PERCENT DAYS AM DO <1.0 mg/l
VAR34	14.742	27.379	0.0	100.0	713	PERCENT DAYS AM DO <0.5 mg/l
VAR35	0.193	0.371	0.0	4.6	480	TOTAL AMMONIA NH3-NH4 mg/l
VAR38	149.581	386.737	0.0	2468.0	713	TILAPIA RECRUIT BIOMASS kg/ha
VAR39	567.318	228.493	66.0	1134.0	129	PRIMARY PRODUCTION mg C/m^2/hour
VAR40	21.137	7.047	5.9	55.6	448	SECCHI DISK VISIBILITY cm
VAR41	332.442	219.656	105.0	905.0	43	PLANKTON UNITS MESH No 150 106 um
VAR42	318.816	231.876	117.0	973.0	38	PLANKTON UNITS MESH No 400 38 um
VAR49	61.351	34.667	11.0	185.0	713	CUMULATIVE NUMBER OF DAYS OVER EXPERIMENT
VAR50	15.704	26.464	0.0	264.3	713	PREDATOR BIOMASS kg/ha
VAR51	357.0	205.970	1.0	713.0	713	SEQUENTIAL NUMBER OF INTERVAL
VAR52	187.612	109.172	2.0	362.0	713	JULIAN DATE START OF INTERVAL
VAR53	184.173	109.741	1.0	361.0	713	JULIAN DATE END OF INTERVAL
VAR54	187.677	109.743	0.0	365.0	713	JULIAN DATE MEAN OF INTERVAL
VAR55	8.066	0.274	7.579	8.987	713	100% OXYGEN SATURATION mg/l
VAR56	29.568	21.390	0.630	92.259	713	OXYGEN SATURATION %
VAR61	0.122	0.116	0.002	0.643	713	TOTAL STOCKING DENSITY kg/m^3
dLdt	0.110	0.082	-0.206	0.414	681	TILAPIA GROWTH RATE cm/day
mL	13.413	3.966	5.0	24.05	681	TILAPIA MEAN LENGTH IN INTERVAL cm
VAR64	3.983	2.412	1.0	11.0	713	SAMPLING FREQUENCY INCLUDING STOCKING
VAR65	1033.631	1037.523	21.03	5784.0	713	STANDING STOCK (all species) kg/ha
VAR66	89.071	107.200	1.057	578.4	713	STANDING STOCK (all species) kg/pond
VAR67	7.84	0.451	6.68	9.73	574	pH MID MORNING
VAR68	55.99	50.313	1.57	297.82	629	TILAPIA WEIGHT START OF INTERVAL g
VAR69	72.23	51.897	2.17	302.47	632	TILAPIA WEIGHT END OF INTERVAL g
VAR70	13.97	4.083	6	23	713	WIND DIRECTION degrees/10, 14:00
VAR71	8.68	2.366	4	15	713	WIND VELOCITY knots, 14:00
VAR72	5.44	2.086	2	9	713	CLOUD COVER n/10, 14:00
VAR73	67.10	8.118	54	93	713	HUMIDITY RELATIVE %, 14:00
VAR74	5.03	1.822	2	9	713	EVAPORATION TOTAL mm/day, 14:00

Filename: *PHILSAMP.WK1*

This file contains a subset of 198 cases of the PHILALL.WK1 file which was used for the analyses with the "extended Bayley method" and the comparative sensitivity analysis with the extended "Gulland-and-Holt" method in the paper by Prein and Pauly. See Table 3 for explanation of variables.

Experiments in Lima, Peru

Filename: *LIMA.WK1*

This file contains the data from experiments with *O. niloticus* conducted by Delgado (1985) at the Aquaculture Research Station of the Instituto del Mar del Peru at Huachipa, Lima. The data are given as average values for the treatment and environmental variables within growth intervals and are described in Table 4 and in the paper by Prein.

Table 4. List of 21 variables and descriptive statistics for the dataset contained in the file LIMA.WK1 from experiments conducted at Huachipa/Lima, Peru. START and END refer to culture intervals.

VARIABLE	MEAN	STD DEV	MINIMUM	MAXIMUM	n	LABEL
CASE			1	90	85	CASE NUMBER
DT	31	6.76	28	92	85	DAYS IN INTERVAL
TL1	16.9	3.16	8.9	22.3	85	NILE TILAPIA LENGTH START OF INTERVAL cm
TL2	17.8	2.75	11.5	22.5	85	NILE TILAPIA LENGTH END OF INTERVAL cm
W1	118.99	63.70	13.3	261.8	85	NILE TILAPIA WEIGHT START OF INTERVAL g
W2	137.59	62.33	30.3	271.0	85	NILE TILAPIA WEIGHT END OF INTERVAL g
DLDT	0.03	0.02	-0.014	0.090	85	NILE TILAPIA GROWTH RATE cm/d
ML	17.35	2.94	10.32	22.39	85	NILE TILAPIA MEAN LENGTH cm
DENS	11560	4988.7	6017	23982	85	STOCKING DENSITY n/ha
DUCK	0.05	0.21	0	1	85	DUMMY MANURE TYPE DICK
PIG	0.27	0.45	0	1	85	DUMMY MANURE TYPE PIG
PD	0.68	0.47	0	1	85	DUMMY MANURE TYPE MIXED PIG + DUCK
MANURE	77.24	39.20	33	205	85	MANURE LOADING RATE kg/ha/d
SOIL	0.41	0.50	0	1	85	DUMMY EARTHEN POND
CEM	0.59	0.50	0	1	85	DUMMY CEMENT POND
DEPTH	0.70	0.25	0.5	1	85	POND DEPTH m
AREA	328.4	377.86	113	1500	85	POND AREA m^2
POZO	0.59	0.58	0	1	85	DUMMY POND WATER SOURCE: WELL
RIO	0.18	0.38	0	1	85	DUMMY POND WATER SOURCE: RIVER
MIXED	0.24	0.43	0	1	85	DUMMY POND WATER SOURCE: BOTH MIXED
WTEMP	22.8	2.22	18.8	26.3	85	WATER TEMPERATURE °C

Experiments at Chilanga and Mwekera, Zambia

Filename: *ZAMB.WK1*

The data of the growth experiments on *O. andersonii* conducted by Mortimer (1960) at the research stations of Chilanga and Mwekera in Zambia are presented here. Data are given in form of growth intervals with corresponding averages per growth interval for the treatment and environmental variables (Table 5) and are further described in the paper by Prein.

Table 5. List of 29 variables and descriptive statistics for the dataset contained in the file ZAMB.WK1 from experiments conducted at Chilanga and Mwekera, Zambia. START and END refer to culture intervals.

VARIABLE	MEAN	STD DEV	MINIMUM	MAXIMUM	n	LABEL
CASE			2	104	84	CASE NUMBER
DT	82.2	23	53	151	84	DAYS IN INTERVAL
TLAN1	16.16	3.83	6.9	24.2	84	*O. andersonii* LENGTH START OF INTERVAL cm
TLAN2	18.13	3.06	9.4	24.2	84	*O. andersonii* LENGTH END OF INTERVAL cm
WAN1	86.68	52.93	5.0	261.0	84	*O. andersonii* WEIGHT START OF INTERVAL g
WAN2	115.93	52.94	19.3	261.0	83	*O. andersonii* WEIGHT END OF INTERVAL g
WME1	33.09	35.44	0	131.0	84	*T. rendalli* WEIGHT START OF INTERVAL g
WME2	38.39	37.67	0	131.0	83	*T. rendalli* WEIGHT END OF INTERVAL g
WMA1	38.38	41.61	0	149.0	84	*O. macrochir* WEIGHT START OF INTERVAL g
WMA2	48.88	47.66	0	149.2	84	*O. macrochir* WEIGHT END OF INTERVAL g
DLDT	0.02	0.02	-0.016	0.075	84	*O. andersonii* GROWTH RATE cm/d
ML	17.15	3.37	8.45	24.02	84	*O. andersonii* MEAN LENGTH cm
AREA	483	425.7	200	1200	84	POND AREA m2
DENTOT	4609	4241.4	350	25883	84	TOTAL STOCKING DENSITY n/ha
DENSAN	1991	1453.6	100	7892	84	*O. andersonii* STOCKING DENSITY n/ha
DENSME	1318	1737.7	0	8800	84	*T. rendalli* STOCKING DENSITY n/ha
DENSMA	1274	1726.2	0	9192	84	*O. macrochir* STOCKING DENSITY n/ha
BAN1	140.7	104.92	18.5	577.9	84	*O. andersonii* BIOMASS START INT. kg/ha
BAN2	189.0	118.66	19.5	577.9	84	*O. andersonii* BIOMASS END INT. kg/ha
BME1	57.3	75.32	0	329.0	84	*T. rendalli* BIOMASS START INT. kg/ha
BME2	68.3	83.31	0	329.0	84	*T. rendalli* BIOMASS END INT. kg/ha
BMA1	60.4	70.37	0	277.7	84	*O. macrochir* BIOMASS START INT. kg/ha
BMA2	77.1	86.27	0	311.7	84	*O. macrochir* BIOMASS END INT. kg/ha
BTOT	334.4	234.55	54.7	957.8	84	TOTAL TILAPIA BIOMASS kg/ha
TREPR	40989	43217	0	118800	17	TOTAL REPRODUCTION n/ha/experiment
MAIZEGR	35.6	34.17	0	68	84	GROUND MAIZE INPUT kg/ha/d
MAIZEBR	10.4	17.00	0	57	84	MAIZE BRAN INPUT kg/ha/d
GRASS	5.8	15.83	0	48.6	84	NAPIER GRASS INPUT, FRESH kg/ha/d
WATEMP	20.4	1.64	17.5	23.4	84	WATER TEMPERATURE °C

Farm and experiment data from Lake Sampaloc, Philippines

Filename: *SAMPALOC.WK1*

This data file contains 13 variables in 20 cases from the studies conducted by Aquino (1982) on *O. niloticus* growth in net cages in Lake Sampaloc, Philippines.

Filename: *CAGE.WK1* and *LNCAGE.WK1*

Manrique (1988) conducted studies on *O.niloticus* production in net cages in Lake Sampaloc, of which the 14 variables in 80 cases are given in the file CAGE.WK1. The file LNCAGE.WK1 contains the base-e log-transformed data of the file CAGE.WK1. The analyses of all three files with data on Lake Sampaloc are presented in the paper by Aquino-Nielsen et al.

Mortality data from the Philippines and Kuwait

Filename: *RPMORTAL.WK1*

This file contains the mortality data of tilapia in experiments conducted at the Brackishwater Aquaculture Center of the University of the Philippines in the Visayas, analysed in the paper by Hopkins and Pauly. The file contains eight variables in 72 cases.

Filename: *REDTILAP.WK1*

The file contains the mortality data of red tilapia cultured in experiments at the Mariculture and Fisheries Department of the Kuwait Institute of Scientific Research. The 13 variables in 44 cases were analysed by Hopkins and Pauly.

Experiments at Domasi, Malaŵi

Filename: *TRENDALL.WK1* and *OSHIRAN.WK1*

Costa-Pierce et al. analysed the growth of *Tilapia rendalli* and *O. shiranus* in experiments conducted at the Domasi Experimental Fish Farm, Malaŵi. TRENDALL.WK1 contains 21 variables in 64 cases, and OSHIRAN.WK1 contains 21 variables in 72 cases.

Recirculating system experiments at Swansea, UK

Filename: *NIL.WK1, NIL3.WK1, AUR1.WK1, AUR2.WK1, AUR3.WK1, AUR4.WK1*

In their paper, Mair and Pauly analysed data of growth experiments with juveniles of *O. niloticus* and *O. aureus* conducted in a recirculating system at the Tilapia Genetics Laboratory of the University College of Swansea, UK. Since the files were larger than could possibly fit on a standard 360 KB 5-1/4" MS-DOS diskette, the files had to be split into parts.

NIL.WK1 and NIL3.WK1 should be combined for analysis and together have 23 variables in 2446 cases. NIL.WK1 contains the first 1000 cases and NIL3.WK1 contains the remaining cases 1001 to 2446.

The files AUR1.WK1, AUR2.WK1, AUR3.WK1 and AUR4.WK1 should be combined to a file of 26 variables with 3666 cases. The file AUR1.WK1 contains the first 1,000 cases, AUR2.WK1 the cases 1001 to 2000, AUR3.WK1 the cases 2001 to 3000, and AUR4.WK1 the remaining cases 3001 to 3666.

Experiments at Dor Research Station, Israel

Filename: *DORGROW.WK1*

The file DORGROW.WK1 is analysed in the paper by Prein. It contains the data from experiments conducted at the Fish and Aquaculture Research Station, Dor, as described in the paper by Hulata, Milstein and Goldman. The data are arranged in form of growth intervals with corresponding average values for treatments and environmental variables (Table 6).

Table 6. List of 22 variables and descriptive statistics for the dataset contained in the file DORGROW.WK1 from experiments conducted at Dor Research Station, Israel. START and END relate to culture intervals.

VARIABLE	MEAN	STD DEV	MINIMUM	MAXIMUM	n	LABEL
INTVL			1	165	156	INTERVAL NUMBER
SEQNO			1	7	156	SEQUENTIAL NUMBER OF INTERVAL
PONDNO			15	34	156	POND NUMBER
DLDT	0.12	0.08	0.001	0.328	156	NILE TILAPIA GROWTH RATE cm/d
ML	15.79	3.1	9.1	20.4	156	NILE TILAPIA MEAN LENGTH cm
TOTBIO	1255.1	632.8	367.4	2750.3	156	TOTAL BIOMASS OF POLYCULTURE kg/ha
CBIO	677.5	353.5	173.4	1588.2	156	COMMON CARP BIOMASS kg/ha
TBIO	108.4	33.4	26.3	172.3	156	NILE TILAPIA BIOMASS kg/ha
TL1	14.79	3.7	7.0	20.2	156	NILE TILAPIA LENGTH START OF INTERVAL cm
TL2	16.80	2.6	10.5	20.9	156	NILE TILAPIA LENGTH END OF INTERVAL cm
TW1	83.42	50.6	6.1	191.0	156	NILE TILAPIA WEIGHT START OF INTERVAL g
TW2	113.53	49.5	22.8	213.5	156	NILE TILAPIA WEIGHT END OF INTERVAL g
PERC	46.5	11.3	24.7	73.7	156	PERCENTAGE MALES IN NILE TILAPIA STOCK
SBIO	431.8	232.8	58.8	972.2	156	SILVER CARP BIOMASS kg/ha
THBIO	87.1	46.2	5.6	184.4	67	TILAPIA HYBRID BIOMASS kg/ha
MANU	9.96	4.2	3.5	18.0	156	DRY CHICKEN MANURE INPUT kg/ha/d
PEL2	1.8	1.1	0.4	3.9	156	PELLET FEED (TYPE 2) kg/ha/d
WATEMP	28.1	2.2	22.6	30.1	156	WATER TEMPERATURE °C
RAD	509.2	70.5	338	603	156	TOTAL SOLAR RADIATION Langleys/d
CLOUD	2.9	1.1	1	5	156	CLOUD COVERING n/8
WINDIR	28.4	1.4	26	30	67	WIND DIRECTION °/10
WINVEL	5.4	0.67	4	6	67	WIND VELOCITY knots

Filename: *DORPROD.WK1*

The raw data from fish culture experiments at the Fish and Aquaculture Research Station, Dor, used in the GLM multiple regressions (paper by Hulata, Milstein and Goldman) are contained in the file DORPROD.WK1 and are described in Table 7.

Table 7. Data from fish culture experiments at the Fish and Aquaculture Research Station, Dor, used in the GLM multiple regressions by Hulata, Milstein and Goldman.

Variable	Description
TREAT	Treatment code
YEAR	Year of experiment
POND	Pond number
AREA	Pond area, in 0.1 ha (= dunam)
DAYS	Length of culture period, in days
PELLET	Mean daily amount of feed pellets, in kg·0.1 ha^{-1}·day^{-1}
MANURE	Mean daily amount of dry chicken manure, in kg dry matter·0.1 ha^{-1}·d^{-1}
TOTYIELD	Mean total daily yield, in kg·0.1 ha^{-1}·day^{-1}
CDEN	Common carp density, in fish·0.1 ha^{-1}
CWTi	Common carp stocking (initial) weight, in g
CGROWTH	Mean common carp growth rate, in g·$fish^{-1}$·day^{-1}
CYIELD	Mean common carp daily yield, in kg·0.1 ha^{-1}·day^{-1}
TDEN	Same as for common carp, but for Nile tilapia
TWTi	Same as for common carp, but for Nile tilapia
TGROWTH	Same as for common carp, but for Nile tilapia
TYIELD	Same as for common carp, but for Nile tilapia
SDEN	Same as for common carp, but for silver carp
SWTi	Same as for common carp, but for silver carp
SGROWTH	Same as for common carp, but for silver carp
SYIELD	Same as for common carp, but for silver carp

Israeli commercial fish farm data

Filename: *ISRAPRO1.WK1, ISRAPRO2.WK1, ISRAPRO3.WK1*

The raw data from Israeli commercial fish farms used in the factor and canonical analyses (paper by Milstein and Hulata) and the GLM multiple regressions (paper by Milstein, Goldman and Hulata) are contained in three files. The file ISRAPRO1.WK1 contains the data from the Western Galilee, ISRAPRO2.WK1 from the Upper Galilee and Bet She'an Valley, and ISRAPRO3.WK1 from the Coastal Plain. The file POND.WK1 contains data to identify farms and ponds. Tables 8 and 9 present the definitions and units of the variables included in those files. Definitions and units of some new variables created with the SAS program are given in Table 10.

Table 8. Variables in file POND.WK1

Name	Description
FARM	name of farm
CODE	code used to identify the farm
POND	code used to identify the pond. The first number also indicates the region: 1= West Galilee; 2, 5 and 6= Bet She'an Valley; 3= Coastal Plain; 4= Upper Galilee.
AREA	surface of pond, in 0.1ha (dunam)
DEPTH	mean depth of pond, in m
OBS	observations

Table 9. Variables in files ISRAPROn.WK1

Name	Description
FARM	code of farm
POND	code of pond
AREA	area of pond, in 0.1 ha
DATEi	stocking date (input) of the first fish species
DATEo	harvesting date (output) of all species
DAYS	number of days from stocking of first species to harvest
CDAYS	number of days from common carp stocking date to harvesting date
CWTi	stocking weight of common carp, in grams
CWTo	harvesting weight of common carp, in grams
CDEN	common carp density, including partial harvesting, in fish/0.1 ha
CYIELD	common carp net yield: biomass harvested (including partial harvesting) minus biomass stocked, in kg·0.1 ha^{-1}culture $period^{-1}$
TDAYS	same as for common carp, but for Tilapia
TWTi	same as for common carp, but for Tilapia
TWTo	same as for common carp, but for Tilapia
TDEN	same as for common carp, but for Tilapia
TYIELD	same as for common carp, but for Tilapia
MDAYS	same as for common carp, but for Mullet
MWTi	same as for common carp, but for Mullet
MWTo	same as for common carp, but for Mullet
MDEN	same as for common carp, but for Mullet
MYIELD	same as for common carp, but for Mullet
SDAYS	same as for common carp, but for Silver carp
SWTi	same as for common carp, but for Silver carp
SWTo	same as for common carp, but for Silver carp
SDEN	same as for common carp, but for Silver carp
SYIELD	same as for common carp, but for Silver carp
MANU	dry chicken manure, in kg·0.1 ha^{-1}culture $period^{-1}$
SORG	sorghum, in kg·0.1 ha^{-1}culture $period^{-1}$
PEL	feed pellets, in kg·0.1 ha^{-1}culture $period^{-1}$
GRASS	net yield of grass carp, in kg·0.1 ha^{-1}
NAMSIF	dummy variable: 0= silver carp variables refer to silver carp, 1= silver carp variables refer to "namsif" (hybrid between silver carp and bighead carp)
WILD	wild spawning or yield of a secondary species stocked, in kg·0.1 ha^{-1}. The fish species is indicated in OBS (observations). If no species is indicated it refers to tilapia spawns. This includes the species of wild spawns or secondary species stocked, as well as other data. Wild spawns are of common carp (carp) or tilapia (til). If nothing is written in OBS and there is a number in WILD, this refers to tilapia wild spawn. Secondary species may be catfish, prawns, ornamental fish (including koi), or a second size group of common carp or silver carp

Table 10. Some of the variables created with the SAS program from the raw data (not included in the ISRAPROn.WK1 files).

Name	Description
REGION	1= Western Galilee, 2= Bet She'an Valley, 3= Coastal Plain, 4= Upper Galilee
SEASGRUP	group of seasons covered by the culture period. See Table 2 of Milstein and Hulata for definitions
FISHES	code for fish species combination in the pond. See Table 3 of Milstein and Hulata for definitions
PONDTYPE	P= pond, R= reservoir
MONTHi	stocking month of first species
TOTDEN	= CDEN + TDEN + MDEN + SDEN. Total density, in fish·0.1 ha^{-1}
TOTYIELD	= CYIELD + TYIELD + MYIELD + SYIELD + GRASS. Total yield (without wild spawning) during the culture period, in kg·0.1 ha^{-1}
TOTYIDAY	= TOTYIELD/_GRDAYS of the first species stocked. Mean daily total yield during the effective growth period, in kg·0.1 ha·day^{-1}
PERCCDEN	= CDEN*100/TOTDEN, percent of carp in total density
CGRDAYS	number of days of effective growth of common carp (excluding the winter period between 15 November and 15 March)
CYIELDAY	= CYIELD/CGRDAYS. Common carp mean daily yield during the effective growth period, in kg·0.1 ha·day^{-1}
CGROWTH	= (CWTo-CWTi)/CGRDAYS. Common carp mean daily growth rate during the effective growth period, in g·day^{-1}
PERCTDEN	same as for common carp, but for Tilapia
TGRDAYS	same as for common carp, but for Tilapia
TYIELDAY	same as for common carp, but for Tilapia
TGROWTH	same as for common carp, but for Tilapia
PERCMDEN	same as for common carp, but for Mullet
MGRDAYS	same as for common carp, but for Mullet
MYIELDAY	same as for common carp, but for Mullet
MGROWTH	same as for common carp, but for Mullet
PERCSDEN	same as for common carp, but for Silver carp
SGRDAYS	same as for common carp, but for Silver carp
SYIELDAY	same as for common carp, but for Silver carp
SGROWTH	same as for common carp, but for Silver carp
MANUDAY	= MANU/_GRDAYS of the first fish species stocked (excluding silver carp). Mean daily amount of dry chicken manure during the effective growth period, in kg· 0.1 ha^{-1}·day^{-1}
SORGDAY	= SORG/_GRDAYS of the first fish species stocked (excluding silver carp). Mean daily amount of sorghum during the effective growth period, in kg· 0.1 ha^{-1}·day^{-1}
SORGFISH	= (SORGDAY * 1000) / (((CDEN * CGRDAYS) + (TDEN*TGRDAYS) + (MDEN * MGRDAYS)) / (CGRDAYS + TGRDAYS + MGRDAYS)). Mean daily amount of sorghum per fish, weighted by densities and days of effective growth of common carp, tilapia and mullet, in g·$fish^{-1}$·day^{-1}
PELDAY	= PEL/_GRDAYS of the first fish species stocked (excluding silver carp). Mean daily amount of feed pellets during the effective growth period, in kg· 0.1 ha^{-1}·day^{-1}
PELFISH	= (PELDAY*1000) / (((CDEN * CGRDAYS) + (TDEN * TGRDAYS) + (MDEN * MGRDAYS)) / (CGRDAYS + TGRDAYS + MGRDAYS)). Mean daily amount of feed pellets per fish, weighted by densities and days of effective growth of common carp, tilapia and mullet, in g·$fish^{-1}$·day^{-1}
INCOME	yield of each species * price of each species, in Israeli money NIS·0.1 ha^{-1}·culture $period^{-1}$. [NIS = New Israeli Shekel]
PROFIT	= INCOME - (fixed costs + variable costs + marketing costs + special water costs), in Israeli money NIS·0.1 ha^{-1}·culture $period^{-1}$. See Milstein, Goldman and Hulata for more details

Filename: *FARMSALL.WK1*

The data from the Israeli commercial fish farms was organised in form of growth intervals with the addition of environmental variables (Table 11). The analysis is presented in the paper by Prein and Milstein.

Table 11. List of 23 variables and descriptive statistics for the dataset contained in the file FARMSALL.WK1 from commercial fish farms in Israel. START and END refer to culture intervals.

VARIABLE	MEAN	STD DEV	MINIMUM	MAXIMUM	n	LABEL
FARM			17	64	960	FARM CODE NUMBER
POND			1	250	960	POND CODE NUMBER
AREA	3.28		0.1	22	960	POND AREA ha
DATE1					960	START DATE OF INTERVAL
DATE2					960	END DATE OF INTERVAL
DT	16	19.39	1	205	960	DAYS IN INTERVAL
TL1	21.1	6.18	1.87	35.69	960	HYBRID TILAPIA LENGTH START OF INTERVAL cm
TL2	22.2	5.74	4.92	35.94	960	HYBRID TILAPIA LENGTH END OF INTERVAL cm
DLDT	0.09	0.08	-0.1	0.454	960	HYBRID TILAPIA GROWTH RATE $cm \cdot day^{-1}$
ML	21.65	5.93	3.88	35.81	960	HYBRID TILAPIA MEAN LENGTH cm
W1	227.3	162.3	0	920	960	HYBRID TILAPIA WEIGHT START OF INTERVAL g
W2	256.8	166.0	2	940	960	HYBRID TILAPIA WEIGHT END OF INTERVAL g
TDENS	23441	37661	875	412500	960	HYBRID TILAPIA STOCKING DENSITY $n \cdot ha^{-1}$
TBIO	3054.2	2207.7	123	18602	960	HYBRID TILAPIA BIOMASS $kg \cdot ha^{-1}$
MAN	47.4	114.9	0	2250	960	DRY CHICKEN MANURE INPUT $kg \cdot ha^{-1} \cdot day^{-1}$
SORG	9.7	17.1	0	169	960	SORGHUM INPUT $kg \cdot ha^{-1} \cdot day^{-1}$
PEL	66.3	76.98	0	925	960	PELLET FEED INPUT $kg \cdot ha^{-1} \cdot day^{-1}$
RAD	531.2	83.3	215	677	960	TOTAL SOLAR RADIATION $Langleys \cdot day^{-1}$
TWAT	27.8	3.11	14.5	34.0	960	WATER TEMPERATURE °C
WIND	4.8	3.98	0	13.0	596	WIND VELOCITY knots
CLOUD	2.7	2.18	0	8	819	CLOUD COVERING n/8
R3	0.32	0.47	0	1	960	DUMMY VARIABLE REGION 3
R4	0.30	0.46	0	1	960	DUMMY VARIABLE REGION 4

REFERENCES

Aquino, L.V. 1982. Some ecological considerations relative to cage culture of *Tilapia nilotica* in Sampaloc Lake. University of the Philippines in the Visayas, Iloilo City, Philippines. 77 p. M.S. Thesis.

Delgado, S. 1985 Crecimiento e produccion de *Tilapia nilotica* (Pisces, Cichlidae) en el laboratorio de Huachipa. Boletin del Instituto del Mar del Peru 9:29-64.

Hopkins, K.D. and E.M. Cruz. 1982. The ICLARM-CLSU integrated animal-fish farming project: final report. ICLARM Technical Reports 5. 96 p. International Center for Living Aquatic Resources Management, Manila, and the Freshwater Aquaculture Center, Central Luzon State University, Nueva Ecija, Philipines.

Manrique, A.S. 1988. Simple productivity indices for lake ecosystems: the case of Sampaloc lake. University of the Philippines, Los Baños, 121 p. M.S. Thesis.

Mortimer, M.A.E. 1960. Observations on the biology of *Tilapia andersonii* (Castelnau)(Pisces, Cichlidae), in Northern Rhodesia. p. 42-67 *In* P.B.N. Jackson (ed.) Joint Fisheries Research Organisation, Annual Report No. 9, 1959. Government Printing Office, Lusaka, 71 p.

Appendix III - Indexes*

Author Index

Prepared by F.S.B. Torres, Jr.

Geographic Index

Species Index

This index excludes references to the word "tilapia."

International Center for Living Aquatic Resources Management

The International Center for Living Aquatic Resources Management (ICLARM) is an autonomous, nongovernmental, nonprofit, international scientific and technical center which has been organized to conduct, stimulate and accelerate research on all aspects of fisheries and other living aquatic resources.

The Center was incorporated in Manila in March 1977. It became a member of the Consultative Group for International Agricultural Research (CGIAR) in May 1992.

ICLARM is an operational organization, not a granting entity. Its program of work is aimed to resolve critical, technical and socioeconomic constraints to increased production, improved resource management and equitable distribution of benefits in economically developing countries. The Center's work focuses in tropical developing countries on three resource systems - inland aquatic (mainly ponds and rice floodwaters), coastal and coral reef - in which research is carried out on their dynamics, on investigating alternative management schemes, and on improving the productivity of key species. The work includes cooperative research with institutions in developing countries, and supporting activities in information and training. The programs of ICLARM are supported by a number of private foundations and governments.

Policies are set by a Board of Trustees with members drawn from the international community. Direction of ICLARM, under the policies set by the Board, is the responsibility of the Director General.

Agricultural Research Organization (ARO), Israel

The Agricultural Research Organization (ARO) is the research branch of the Ministry of Agriculture of Israel. The organization consists of seven institutes with several farms and research stations throughout the country, employing over 1,000 scientists and support personnel. The ARO's mission is to further Israel's agricultural potential through scientific research. Therein, three principal lines of work are involved: responding to the needs of farmers or the problems of current production, spearheading long-term research in national priority areas, as defined by the Ministry of Agriculture, and pursuing basic research to expand the frontiers of agricultural science.