PROCEEDINGS OF THE SYMPOSIUM ON COASTAL AQUACULTURE

Held at Cochin

From January 12 to 18, 1980

PART 4: CULTURE OF OTHER ORGANISMS, ENVIRONMENTAL STUDIES, TRAINING, EXTENSION AND LEGAL ASPECTS

(Issued in December 1986)

MARINE BIOLOGICAL ASSOCIATION OF INDIA

POST BOX NO. 1023, COCHIN-682 031, INDIA

Price : Rs. 400.00

SYMPOSIUM SERIES 6

Abbreviation

Proc. Symp. Coastal Aquaculture, Pt. 4.

PRINTED IN INDIA BY A. D. THOMAS STEPHEN AT THE DIOCESAN PRESS, MADRAS 7 AND PUBLISHED BY E. G. SILAS ON BEHALF OF THE MARINE BIOLOGICAL ASSOCIATION OF INDIA, COCHIN-682 031

Proc. Symp. Coastal Aquaculture, 1986, 4 : 1152-1159

STATISTICAL DESIGNING OF AQUACULTURE EXPERIMENTS

T. JACOB

Central Marine Fisheries Research Institute, Cochin-682 031

ABSTRACT

Experiment is the main tool of a researcher. The effect of factors or treatments can be assessed only by testing them experimentally. Several factors like composition of stock, density of stock, feeding level, depth of water and salinity come into play in aquaculture experiments. Sometimes the problem may be one of developing a low-cost technology which can be gainfully employed by small and marginal farmers. Here again a number of factors under study act and interact. The main difficulty in such situations is that effects of extraneous factors very often mask the real treatment effect. One way to overcome this difficulty is to use homogeneous experimental material. However there are very many practical limitations to achieve such an end. A statiscally planned experiment attempts to minimise effects of heterogeneity in experimental units from treatment comparison, reduce experimental error, provide unbiased estimates and ensure validity of test of significance. The paper discusses at length the various statistical requirements in experimentation with special reference to coastal aquaculture systems.

INTRODUCTION

FOR FOOD or for fun, fish farming has been in practice since ancient times. In the course of centuries the area of farming expanded and technological innovations found their place. From earlier pond system culturing have grown to include flowing water and enclosure systems not only in the inland waters but also in estuaries, backwaters and open sea. Today, in several countries of the world, scientifically managed fish farming has become an accepted method for augmenting fish production (Brown, 1977).

Compared to the inland waters, marine environment is often aggressive and the task of evolving mariculture methods suiting different ecosystems is quite formidable and a good deal of research in this direction is in progress (Bardch *et al.*, 1971). In India too, scientists are deeply involved in research to improve the traditional methods of coastal aquaculture and to develop production technologies which are feasible and economically viable.

The main tool of a research worker is experiment. The effect of factors or treatments can be assessed only by testing them experimentally. Several aspects like depth of water, salinity, composition of stock, density of stock and feeding levels come into play in aquaculture experiments. As many factors act and interact a difficulty faced is that effects of factors extraneous to those under study mask the treatment comparisons under test. A statistically planned experiment attempts to reduce the effects of extraneous factors from treatment comparisons and has many desirable properties. The paper explains the need for statistical designing and discusses various situations in coastal aquaculture experimentation where statistical designs can be gainfully employed.

NEED FOR STATISTICAL DESIGNING

It may look strange that while statisticians themselves do not generally conduct experiments they express their claim to have a say at the stage of planning as well as at later stages. It is the scientific nature of an experiment which compels the need for meeting statistical requirements (Kempthorne, 1972). Statistical designing of an experiment involves the formulation of a scheme or a lay out plan, where the placements of treatments in experimental units are specified to suit the requirements of the particular problem keeping in view the requisite statistical principles. The term 'treatments' is used here in a general way and may mean species, levels of feeding, doses of stimulant, stocking densities, etc.

Variability in experimental material is an inevitable freature in any field of research. Consider for example two prawn culture ponds kept under conditions as similar as possible with the same extent of area, species, stocking density, etc. At the time of harvesting one would find that the yield of one pond is different from the other. This may be attributed to the uncontrolled variation inherent in the production process. Consider another two ponds again kept under almost identical conditions except that in one pond supplementary feeding is given. Here again, at the time of harvesting the yields would be found to be different. Can we attribute the difference straightaway to the effect of levels of feeding? We cannot. May be the supplementary feeding did not contribute anything to the difference in yield and the difference could be purely due to the inherent uncontrolled variation. Differences are expected even when similarity is maintained in the two ponds. One might then say that if the difference is quite high it can be attributed to the difference in the level of feeding. But how high the difference should be to attribute it to the level of feeding? The answer becomes quite subjective. Thus variation introduces a degree of uncertainty into the conclusions that are drawn from the results. A mathematical measure of uncertainty is probability, the theory of which enables us to make numerical statements about uncertain outcomes. But unless the planning is made taking into consideration statistical principles it would not be apt to make such statements regarding the outcome.

The formulation and testing of hypothesis are the main features of a scientific method (Kempthorne, 1972). A researcher postulates a hypothesis which he would like to verify. This verification necessitates the collection of observations through an experiment and the designing of experiment is concerned with the pattern of observations to be collected which should be relevant to his hypothesis. It may be stressed here that when we draw inductive inferences from experimental data, every statement of results is subject to some error, an error about which probability statements may be made with the aid of mathematical statistics. Only a statistically designed experiment can permit a valid test of significance involving probability statements whether a particular difference is due to chance causes or can be attributed to the real differences between two treatments.

As indicated earlier the results of an experiment are affected not only by the action of treatments, but also extraneous variation which tend to mask the effects of treatments. This extraneous variation is conventionally termed as 'experimental error' (or sometimes called 'error') where the word 'error' is not synonymous with mistakes, but indicates all types of extraneous variation (Cochran and Cox, 1973). There are two sources of experimental error, one refers to the inherent variability in the experimental material or units to which the treatments are applied and the other type refers to the failure to standardise the experimental technique. It is desirable that the experimental error is kept as minimum as possible as otherwise only a large difference in the treatment means will be detected as significant. Reduction of experimental error automatically increases the precision. One way to reduce the error is by ensuring uniformity in the conduct of the experiment. Two other methods to increase the precision of the estimates are one by providing more replications and the other by skillful grouping of units in such a way that the units to which one treatment is applied are closely comparable with those to which another treatment is applied. Some of the general principles governing these methods and other related aspects are elaborated.

Two primary requisites in designing experiments are replication and randomisation. Replication or repetition of treatments provides stability to the mean, but more than that makes it possible to estimate the experimental error. It also increases the precision of the estimates of both the treatment mean and the experimental error.

Randomisation which means random allocation of treatments to various experimental units, ensures that a treatment will not be unduly favoured or handicapped in successive replications. It ensures unbiasedness of the estimates of experimental error and provide for valid treatment comparisons against the experimental error (Fisher, 1949). When treatments are replicated and allocated randomly to the various units we are in a position to test the significance of observed treatment differences by the use of test of significance procedures. Thus it is essential to provide for adequate number of replications and ensure proper randomisation at the planning stage (Panse and Sukhatme, 1964).

Grouping of units often help in reducing the experimental error. Consider for instance an experiment with a number of replications, all the treatments being tried in each replicate. The error from any replicate can arise from sources of variation that affects the units within the replicate. Variation between replicates do not contribute to the error. Thus if the experimental units form a very heterogeneous set, try to group them so that units in the same replicate is as homogeneous as possible while variation between replicates could be large. By this process, from the total variation in the observations the variation between replicates can be removed resulting in the reduction in the error variance (experimental error). The device of reducing errors through suitable groupings is called local control. Looking from another angle, if treatments are allotted to a replication with homogeneous units the observed differences would reflect the real differences between the treatment effects. The principle of local control is the basis for experimental designs such as 'randomised blocks' and 'latin squares'. When the number of treatments to be accommodated in a replicate becomes large, the homogeneity within a replicate tends to be lost and can be restored by dividing the replication into smaller blocks which is the basis of 'confounding' in factorial experiments and also various 'incomplete block designs'.

To sum up, statistical designing of experiments attempts to minimise the effects of heterogenity in experimental units from treatment comparisons, reduce experimental error, provide unbiased estimates and ensure validity of test procedures. The test of significance emanating from the design exerts a sobering influence on the type of experimenter who jumps to exciting conclusions that can as well be ascribed to the natural variation inherent in the experiment (Cochran and Cox, 1973).

Some Useful Designs

(i) Randomised block

One of the most commonly used plans is the randomised block design where experimental material is divided into blocks each of which constitute a single replicate in such a way that the units within a block is as homogeneous as possible. The treatments are now randomly allotted to the experimental units within a block. This increases the comparability of treatment effects as they act under conditions which are similar except for the treatments. For instance in an experiment to select an economic supplementary feed mixture from among 4 prepared mixtures for prawn culture, 4 ponds all located by the side of the main water body like the backwater or estuary could be grouped as one block or replication and allot treatments at random. The next 4 could be ponds running parallel to the first set, but more inside the land so that within a block salinity and associated features are likely to be similar. This arrangement takes care to a good extent salinity gradient likely to be reducing when moved away from the main water body. In the experiment if there are 5 replications there will be total 20 ponds. If all the 20 ponds are more or less similar no blocking or stratification is required and the treatments could be randomly allotted over the entire range of the 20 ponds. Such a design is called completely randomised design. However if heterogeneity in the features of the ponds is suspected it is desirable to provide blocks which may help in reducing the experimental error.

(ii) Latin square

In randomised blocks one-way restriction is imposed. If heterogeneity is suspected in two directions the experimental area can be divided into, say, rows and columns and treatments are applied in such a way that each treatment appears only once in a row and once in a column. Such an arrangement is called a latin square design. Through elimination of row and column effects the residual error variance may be very much reduced. Consider the question of finding out the best spat attachment material from among say, 4 materials like tile, brick, concrete and asbestos pieces. A raft with 4 poles in rows and 4 in columns can be fabricated and at the 16 junctions hang ropes on which the material is tied. The placement of the 4 materials can be fixed following a 4×4 latin square arrangement. Effects of two-way variation like current direction or nearness to the exterior portion of the floating raft could be reduced from comparison between the quantity of spat collected on the different materials. If three-way and higher-way grouping of treatments is required designs like graeco-latin squares and hyper graeco-latin squares can be attempted (Federer, 1967). The procedure of using latin and graeco-latin cubes also could be explored. With two-way stratification the latin square controls more variation than randomised block design resulting often in a smaller error mean square. However the number of treatments is limited to the number of rows or columns and for large number of treatments latin square design is not preferred.

(iii) Factorial experiments in complete and incomplete blocks

Consider an experiment to study the effect of different levels of protein and energy on weight of fish in culture ponds. If there are say, 2 levels for each factor there will be in all 4 (2^2) treatment combinations. A group of treatments which contains two or more levels of two or more factors in all combination is known as the factorial arrangement. The different combinations could be allotted as in a randomised block design. The experimenter could try a one-factor-at-a-time approach. But the advantage in a factorial experiment is that not only the main effects, but also the interactions between factors can be studied and tested for statistical significance.

If the number of factors and levels are large say 3 factors, salinity, temperature and oxygen content at 3 levels each, the number of treatment combinations will be 27. It may be difficult to get 27 experimental ponds, which are more or less homogeneous with regard to factors other than being tested so that the principle of stratification to reduce experimental error cannot be implemented. An ingenious device to overcome this situation is called confounding where a homogeneous block will not accommodate the full replication. One replication is divided into say, 3 compact blocks such that the units in the smaller blocks are homogeneous. The 27 treatment combinations can be divided into 3 groups of 9 each and allotted to the 3 compact blocks. However some of the treatment comparisons will not be distinguishable from block differences or in other words, get confounded with block differences. Thus some sacrifices have to be made. But at the planning stage this aspect can be considered and the scheme can be so formed that all major and important comparisons are kept free from block differences. Factorial set-up can be easily superimposed in polyculture experiments in pens or in ponds.

(iv) Split-plot

In the usual factorial trial, the effects are estimated with the same degree of precision. It is quite possible that some factors may produce much larger differences than others and in the factorial set-up the precision attained could be sufficient to bring out the significance of the differences between levels of factors capable of showing large difference, but may not permit detection of smaller difference, between the levels of the other factor, (Panse and Sukhatme, 1964). A device known as 'split plot' where the levels of the 2nd factor which we wish to compare with greater precision are assigned to contiguous plots under a common level of the first factor. This effects greater local control. Thus each block is divided into main plots where the level of the first factor is allotted at random and next subdivide each main-plot to subplots to which are allotted at random the levels of the second factor. For instance if 3 species of fish are to be compared in an experiment in pens fabricated in the inshore water for growth with 2 types of supplementary feeding, instead of putting all the 6 combinations at random in a block of 6 pens, arrange the three species of fish to three pens and next subdivide each pen into two sub-pens to be randomly allotted to the two feeds. Enough replications will have to be provided.

(v) General incomplete block

There are situations where a large number of treatments which are not of factorial type are to be tried. Then the factorial type arrangement with or without confounding may not be suitable and one has to look for other procedures for reducing the block size. The general incomplete block designs come handy in such cases. Here the number of units in a block will be less than the number of treatments. When homogeneous number of units in a block is quite small these design are useful. For instance in induced breeding of say, *Sillago sihama* by injecting pituitory harmones one may like to study the effect of the level of the dose injected. Similar type of fish suited for the study may be available at a time only in small numbers say 3 and if 5 doses are to be tried we can have a plan where number of units in a block of homogeneous units is 3. The pattern of allotment of the 5 doses in the blocks of 3 homogeneous units each can be taken from the catalogue of plans available (Cochran and Cox, 1973). The method of analysis of the data collected will be complex compared to some of the designs mentioned earlier.

(vi) Switch-over

There are occasions in which treatments are applied in sequence over several periods on a group of individuals. Consider an experiment to study the effect of mineral supplementation of two types in lobsters kept in artificial tanks. If there are say six groups of lobsters separated and kept in tanks with sub-partitioning then the two types of supplementation, are given such that half the groups received say, type A and the other half type B in period 1. The lobsters receiving type A in period 1 will be switched over to get type B in period 2 and vice versa. Such a design is called switch-over or change-over design (Federer, 1967). On the other hand if a time trend is expected in the character under study a switch-back or a double reversal design will have to be used. In these procedures a rest period is to be provided between two treatment periods so that there is no carry-over effect or residual effect influencing the treatment during the second period. However if a reasonably long rest period is not feasible or the residual effect is itself a topic of interest the procedure is to be modified so that direct and residual effects of treatments can also be measured.

(vii) Rotatable

In factorial experiments with quantitative variables like temperature, salinity, amount of oxygen and level of nutrients the yield or response say y, can be considered as a function of these variables. If we assume a linear multiple regression of y on the variables, $x_1, x_2, \ldots x_k$ the levels of which are controlled by the experimenter then the relationship comes under the model

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots.\beta_k x_k+\epsilon$$

Where β's are the regression coefficients and ϵ random component. From observed data it would be possible to estimate the coefficients by least square procedure and then the function gives a complete summary of the results of the experiment. A design which is specially oriented to study, such response functions is called rotatable design which has some additional desirable properties compared to the usual designs. The lay-out plans for such designs are available in many publications (Cochran and Cox, 1973). A second order rotatable design which is not very complex can be used for the determination of the levels of factors like amount of oxygen and salinity needed for optimum production.

Statistical Analysis

Once a design is fixed the procedure of analysis follows. We can postulate a mathematical model to represent a design. As an example the linear additive model underlying the randomised block design can be written as

$$y_{ij} = \mu + \tau_i + \rho_{ij} + \epsilon_{ij}$$

where μ represents the general mean, τ the effect of treatment, ρ the effect of block and ϵ the random component. Under a set of assumptions it is possible to estimate the treatment and block effects. The components of variation due to treatments, blocks and error can be computed and the analysis of variance Table can be prepared which facilitates test of significance of treatment differences and drawing conclusions so that appropriate decisions can be taken. Similar procedures are available for other designs as well.

Some experimenters do not bother to follow a design, but try to analysis the data statistically. Some others follow a design, but do not care to follow the appropriate procedure of analysis. It is essential in scientific experimentation to follow a suitable design and analyse the data through appropriate statistical procedures.

Much has been talked about the problem of getting homogeneous experimental units for allotment of treatments by using local control methods to reduce the experimental error. In addition there is a purely statistical procedure to reduce the error variance called 'analysis of covariance' where information on a suitably chosen auxiliary variable is used to adjust the error sum of squares by subtracting from it the sum of squares due to the regression of say yield, on the auxiliary variable (Snedecor and Cochran, 1967). The technique provides a powerful method of reducing error variance and should be availed of by the experimenter wherever possible.

Discussion

With increase in the volume of research in mariculture involving prawns, lobsters, crabs, mussels, oysters, clams, finfishes and seaweeds there is an excellent scope for the employment of statistical designs. It is an accepted fact that mariculture problems have to be dealt with a multididiciplinary approach and statisticians working in fisheries research have an important responsibility. It is imperative that any scientific experiment need to be statistically planned so that the emanating data become amenable to statistical analysis for arriving at valid conclusions needed for decision-making.

One aspect need to be stressed here, namely, the provision of enough number of replications in an experiment. Consider the example of 4 feed mixtures which are tried for economic evaluation in prawn culture. If the mixtures are allotted only one each in four ponds without replications we will get only a single figure on, say, cost of production of a unit weight, for one mixture. Thus with four treatments the character under study will have only four figures, a single figure for each, and no statistical analysis is possible. One way is to partition the ponds into 4 sub-ponds which may provide 16 figures 4 each for one treatment for analysis. It my be stressed that apart from reducing experimental error replication of treatments alone can provide an estimate of the experimental error essentially needed for treatment comparisons.

The question of minimum number of replications required is of great importance in coastal aquaculture experiments because of the cost involved and the inherent special problems, compared to experiments on land. An important consideration in determining the minimum number of replications is that the test of significance should be sufficiently sensitive to detect real difference between treatments as distinguished from variation due to chance causes. The sensitiveness of the test will depend primarily on the magnitude of variation in the experimental units with regard to the character under study. If the magnitude is known the number of replications required for detecting a particular difference with a certain level of confidence can be worked out (Federer, 1967; Panse and Sukhatme, 1964). In the absence of any knowledge regarding the magnitude of variability the number of replications provided should be at least sufficient to ensure 12 degrees of freedom for error. This is inferred from the fact that the tabulated value of 'F' at the conventional level of significance of 5 per cent ceases to fall off rapidly for degrees of freedom beyond 12. On this basis the minimum number of replications can be worked out for a particular design (Jacob and Marutiram, 1975).

In the preceeding paragraphs some guidelines in planning experiments have been dealt with. Some of the standard designs available in literature which can be used with advantage in coastal aquaculture experiments have been enumerated. It may however be observed that considering the experimental resources available and the special nature of problems some amount of 'tailoring' may have to be resorted to for suiting particular situations.

References

BARDCH, J. E., J. H. RYTHER AND W. O. MCLARNEY 1971. *Aquaculture*. Wiley, New York, USA.

BROWN, E. E. 1977. *World Fish Farming: Cultivation and Economics*. The AVI Publishing Co. Connecticut, U.S.A.

COCHRAN, W. G. AND G. M. COX 1973. *Experimental Designs*. Asia Publishing House, New Delhi (for Wiley).

FEDERER, W. T. 1967. *Experimental Designs*. Oxford Publishing Co., New Delhi.

FISHER, R. A. 1954. *Design of Experiments*. Oliver and Boyd, London.

JACOB, T. AND B. MARUTIRAM 1975. Planning of experiments in animal sciences research — Statistical requirements. *Agricultural Research Communication Centre, Karnal.*

KEMPTHORNE, O. 1972. *The Design and Analysis of Experiments*. Wiley, India.

PANSE, V. G. AND P. V. SUKHATME 1964. *Statistical Methods for Agricultural Workers*. I.C.A.R., New Delhi.

SNEDECOR, G. M. AND W. G. COCHRAN 1967. *Statistical Methods*. Oxford Publishing Co., New Delhi.