The Open Access Israeli Journal of Aquaculture – Bamidgeh

As from **January 2010** The Israeli Journal of Aquaculture - Bamidgeh (IJA) will be published exclusively as **an on-line Open Access (OA)** quarterly accessible by all AquacultureHub (http://www.aquaculturehub.org) members and registered individuals and institutions. Please visit our website (http://siamb.org.il) for free registration form, further information and instructions.

This transformation from a subscription printed version to an on-line OA journal, aims at supporting the concept that scientific peer-reviewed publications should be made available to all, including those with limited resources. The OA IJA does not enforce author or subscription fees and will endeavor to obtain alternative sources of income to support this policy for as long as possible.

Published under auspices of
The Society of Israeli Aquaculture and Marine Biotechnology (SIAMB),
University of Hawaii at Manoa Library
and
University of Hawaii Aquaculture Program in association with
AquacultureHub
http://www.aquaculturehub.org

ISSN 0792 - 156X

PUBLISHER:
Israeli Journal of Aquaculture - BAMIGDEH -
Kibbutz Ein Hamifratz, Mobile Post 25210,
ISRAEL
Phone: + 972 52 3965809
http://siamb.org.il

The Israeli Journal of Aquaculture – Bamidgeh 60(2), 2008, 134-153.

Determining Variability, Confidence, and Statistical Power in Aquaculture Experiments

Pedro Araujo*
National Institute of Nutrition and Seafood Research (NIFES), P. O. Box 2029, Nordnes, N-5817 Bergen, Norway

(Received 26.7.07, Accepted 13.10.07)

Key words: sources of variation, error decomposition, experimental design, confidence of the design, statistical power

Abstract

This article deals with the lack of appreciation among marine biologists of basic statistics in aquaculture experiments. Topics include calculation of sources of variation, the importance of true replication rather than pseudo-replication to test for treatment effects, estimation and visualization of the confidence of an experimental design used in feeding trials, and calculation of its statistical power. An arbitrary example is used to illustrate how the described theory can be applied in practice. The approach demonstrates that nutritional experiments with a reduced number of tanks can be a valid strategy as long as certain experimental considerations are taken into account. In addition, it shows that information on the relative sizes of 'between' and 'within' sources of variation can be used to design more efficient experiments by minimizing the effects of the stronger sources of variance. The approaches used in this article are applicable to large and small-scale experiments.

Introduction

Global aquaculture production is traditionally dependent upon fishmeal and fish oil as the major source of dietary protein and lipid. The fact that fish oil is a major source of highly unsaturated fatty acids makes identification and development of less expensive and more readily available alternatives a major challenge.

Experimental design is a highly developed field of statistics that can assist aquaculture researchers in coping with that challenge. There is ample literature on experimental designs (Cochran and Cox, 1953; Finney, 1955; Box et al., 1978), the properties of which make them important tools that can provide answers to problems in science and industry as well as specialties such as psychology or chemistry.

Nutritional aquaculture experiments are influenced by several factors. A good design helps determine the significance of each factor on the results, the size of each source of error, and how to minimize the effects of systematic and stochastic sources of error. Application of reduced de-

* Tel.: +47-95285039, fax: +47-55905299, e-mail: pedro.araujo@nifes.no

signs, such as the Plackett-Burman design (Plackett and Burman, 1946; Kaufmann, 1995), allow the researcher to study, for example, the effect of five variables by performing only eight experiments. This is particularly important in large scale studies or investigations with a high degree of restrictions.

Response surface methodology (Deming and Morgan, 1987) is valuable for systematic and simultaneous optimization of multivariate systems such as feed experiments (Ruohonen et al., 2003; Vielma et al., 2003). Randomized designs are important for comparing more than two treatments. However, in spite of the multiple advantages of experimental design techniques in industries such as aquaculture, they have not enjoyed widespread popularity, probably because of the few publications that report on efficient designs using a reduced number of ponds in aquaculture experiments (Riley and Edwards, 1998; Smart et al., 1998). The lack of literature regarding sources of variation and their estimation to obtain representative results also may account for the limited application of experimental designs in aquaculture research (Riley and Edwards, 1998).

The National Institute of Nutrition and Seafood Research (NIFES) in Norway participates in many large scale and pilot aquaculture studies with other countries including the UK, France, Spain, Italy, Belgium, Sweden, Greece, Hungary, Estonia, USA, India, China, and Australia. As a result of this collaboration, project management resources must be allocated. Experimental designs are of great help at this stage. A proposed experimental design can be altered just before implementation as a result of changes in experimental circumstances. For example, if the number of replicates per variable parameter must be reduced due to costs or a lack of tanks, experimental design principles can be used to analyze the confidence of the modified experiment and results can be interpreted in the same way as for a completely balanced design.

In this paper, a full and a reduced experimental design are used in a theoretical large scale aquaculture feeding trial. The designs are studied and analyzed in terms of the sources of variation involved in the experiments. The confidence of the designs and their statistical powers are calculated. The article concentrates on the decomposition, calculation, and description of the sources of variation and estimation of their significance. Calculation and visualization of the confidence for a multifactor experimental domain and the retrospective statistical power of the experimental design are also described. Simulated data are used to illustrate how the theory can be applied in practice.

Materials and Methods

1. Experimental design – An optimal experimental design was prepared for a theoretical nutritional study in aquaculture (Table 1). The aim of the study was to compare a standard control diet with three experimental diets under the influence of four varying factors: location, diet, tank, and subsampling. The study aimed to determine the significance of each factor and the influence of each source of error on the quality of the results. In the equations below, the locations (A, B, C) are designated by the letter L, the diets (control, diet 1, diet 2, diet 3) by the letter D, the tanks (I, II, III) by the letter T, and the subsampling area (α, β, γ) by the symbol $\mathcal{S}$. Thus, a total of $N_{ldt\mathcal{S}}$ fish were sampled from three areas of each tank and the i measurement at location l, diet d, tank t, sampled in the β part, is recorded as $Y_{ldt\beta i}$. Subsamplings are designated by the symbol $\mathcal{S}$ instead of the letter S to avoid confusion with the variance symbol S^2 to be used later.

2. Fixed and random effect models – It is generally accepted in aquaculture that diets are replicable variables that will produce the same average effect in each replicate. Such an assumption is called a fixed effects model. Examples of fixed effects models include studies on the effects

Table 1. An ideal experimental design for comparing a control diet to three experimental diets with four variable study conditions - location (L), diet (D), tank (T), and sampling area (*S*).

L	*D*	*T*	*S*		
			Area α	*Area* β	*Area* γ
A	Control	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 1	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 2	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 3	I	n	n	n
		II	n	n	n
		III	n	n	n
B	Control	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 1	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 2	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 3	I	n	n	n
		II	n	n	n
		III	n	n	n
C	Control	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 1	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 2	I	n	n	n
		II	n	n	n
		III	n	n	n
	Diet 3	I	n	n	n
		II	n	n	n
		III	n	n	n

of an omega-3 rich fish oil versus an omega-6 rich plant oil on the chemical composition of fish or studies that compare feeder machines, pellet sizes, growth at different temperatures, etc. The common feature of all these examples is that the variables are fixed by the researcher. In contrast, when factors cannot be assumed from a known set of variables and are random samples of a larger number of potential treatments, the model is called a random effects model. Random effects models are not fixed by the experimenter; they are sampled from a population of possible samples instead. For example, when comparing a control with three alternative diets in three locations as in Table 1, the locations are random samples of an infinite number of locations in the world, hence they are random factors.

A model can be fixed, random, or mixed. In the design showed in Table 1, location and tank are random effects and diet is a fixed effect. Such combinations of effects generate a mixed model described by the equation:

(1)

$$y = \mu + r_L + r_D + r_T$$

where y is the measurement, μ is the overall mean of the measurement, and the r values are location, diet, and tank residual errors. However, the hierarchical structure of the design in Table 1 also describes the following nested model:

(2)

$$y = \mu + v_L + v_{D(L)} + v_{T(LD)} + r_L + r_D + r_T$$

where V_L represents the effect of the location, $V_{D(L)}$ is the effect of the diet nested with each location, and $V_{T(LD)}$ is the effect of the tank nested with each location and diet.

3. Sources of variability – Some sources of variability can be calculated as mean square errors. In Table 1, the four sources of variability (subsampling, tank, diet, and location) are estimated and the correspondence between and within variances are computed as explained below.

a. *Subsampling*. In Table 1, each tank is divided into smaller units and a sample of n units is selected. This technique is called subsampling since the tank is not measured as a whole but is sampled from different locations to determine whether different parts of the tank are representative of the whole. For example, weight and size variability among fish in the same tank indicate that social interaction plays an important role in fish development (Goldan et al., 1998). Consequently, variation in the fatty acid profiles of the fish could be expected. However, genetic and environmental parameters are also implicated in weight and length variability (Winkelman and Peterson, 1994). According to the design in Table 1, three areas (α, β, γ) of each tank can be sampled, n fish can be collected from each area, and variance terms associated with the representativeness of the subsampling can be calculated using the following equations. Note, estimation of the subsampling variability gives no indication of the significance of the treatment effects.

(3)

$$S^2_{ws} = \frac{\sum_{s=1}^{S_{ldt}} \sum_{i=1}^{N_{ldts}} \left(y_{ldtsi} - \bar{y}_{ldts}\right)^2}{\sum_{s=1}^{S_{ldt}} \left(N_{ldts} - 1\right)}$$

and
(4)

$$S_{bs}^2 = \frac{\sum_{s=1}^{S_{ldt}} \left[N_{ldts} \times \left(\bar{y}_{ldts} - \bar{y}_{ldt} \right)^2 \right]}{S_{ldt} - 1}$$

where S^2_{ws} and S^2_{bs} are the 'within' and 'between' subsampling root mean square errors, respectively, S_{ldt} represents the number in areas (α, β, γ) into which each tank is divided, N_{ldts} represents the total number of sampled fish, y_{ldtsi} is the *i* measurement at location *l*, diet *d*, tank *t*, subsample *S*, and $\bar{y}_{ldts}$ and $\bar{y}_{ldt}$ represent the averages in each sampling (α, β, γ) and in each tank (*I*, *II*, *III*), respectively. These averages can be calculated by the expressions:

(5)

$$\bar{y}_{ldts} = \frac{\sum_{i=1}^{N_{ildts}} y_{ldtsi}}{N_{ldts}}$$

and
(6)

$$\bar{y}_{ldt} = \frac{\sum_{s=1}^{S_{ldt}} \bar{y}_{ldts}}{S_{ldt}}$$

b. *Tank.* The relative sizes of variances within and between tanks are indicated as S^2_{wt} and S^2_{bt} and determined by the following equations. Equation 8 shows how the use of data on diet effects from experiments where tanks are not replicated are arithmetically forbidden. The denominator in Equation 8 would become zero if T_{ld} equals 1 (no tank replication) and, thus, Equation 8 would have no meaning in ordinary statistics.

(7)

$$S_{wt}^2 = \frac{\sum_{d=1}^{D_l} \sum_{t=1}^{T_{ld}} \sum_{s=1}^{S_{ldt}} \sum_{i=1}^{N_{ldts}} \left(y_{ldtsi} - \bar{y}_{ldt} \right)^2}{\left(\sum_{t=1}^{T_{ld}} \sum_{s=1}^{S_{ldts}} N_{ldts} \right) - T_{ld}}$$

and

(8)

$$S_{bt}^2 = \frac{\sum_{d=1}^{D_l}\sum_{t=1}^{T_{ld}}\sum_{s=1}^{S_{ldt}}\left[N_{ldts} \times \left(\bar{y}_{ldt} - \bar{y}_{ld}\right)^2\right]}{T_{ld} - 1}$$

where T_{ld} is the number of tanks per diet at each location and $\bar{y}_{ld}$ is the average at each diet level, calculated as:

(9)

$$\bar{y}_{ld} = \frac{\sum_{t=1}^{T_{ld}} \bar{y}_{ldt}}{T_{ld}}$$

c. *Diet.* To compare alternative diets with the control and determine whether feed performance indicators refer to the entire population, regardless of diet, mean square errors within diets (S^2_{wd}) and between diets (S^2_{bd}) are calculated by the equations:

(10)

$$S_{wd}^2 = \frac{\sum_{d=1}^{D_l}\sum_{t=1}^{T_{ld}}\sum_{s=1}^{S_{ldt}}\sum_{i=1}^{N_{ldts}}\left(y_{ldtsi} - \bar{y}_{ld}\right)^2}{\left(\sum_{d=1}^{D_l}\sum_{t=1}^{T_{ld}}\sum_{s=1}^{S_{ldt}} N_{ldts}\right) - D_l}$$

and

(11)

$$S_{bd}^2 = \frac{\sum_{d=1}^{D_l}\sum_{t=1}^{T_{ld}}\sum_{s=1}^{S_{ldt}} N_{ldts} \times \left(\bar{y}_{ld} - \bar{y}_{l}\right)^2}{D_l - 1}$$

where D_l represents the number of diets tested at each location and $\bar{y}_l$ represents the average at each location, calculated using the following formula:

(12)

$$\bar{y}_{l} = \frac{\sum_{d=1}^{D_l} \bar{y}_{ld}}{D_l}$$

d. *Location.* When comparing the influence of location, factors such as temperature, salinity, weather conditions, etc., must be taken into account. These factors are not considered in the design portrayed in Table 1 which does not provide a basis for induction from data but only allows the researcher to check simple answers such as “yes” or “no” in a pre-existing test protocol. The variances within (S^2_{wl}) and between (S^2_{bl}) locations are determined by the following equations:

(13)

$$S^2_{wl} = \frac{\sum_{l=1}^{L}\sum_{d=1}^{D_l}\sum_{t=1}^{T_{ld}}\sum_{s=1}^{S_{ldt}}\sum_{i=1}^{N_{ldts}}\left(y_{ldtsi} - \bar{y}_l\right)^2}{\left(\sum_{l=1}^{L}\sum_{d=1}^{D_l}\sum_{t=1}^{T_{ld}}\sum_{s=1}^{S_{ldt}} N_{ldts}\right) - L}$$

and

(14)

$$S^2_{bl} = \frac{\sum_{l=1}^{L}\sum_{d=1}^{D_l}\sum_{t=1}^{T_{ld}}\sum_{s=1}^{S_{ldt}} N_{ldts} \times \left(\bar{y}_l - \bar{y}\right)^2}{L-1}$$

where L represents the number of locations (L = A, B, C, i.e., 3) and $\bar{y}$ is the average between locations, calculated as:

(15)

$$\bar{y} = \frac{\sum_{l=1}^{L} \bar{y}_l}{L}$$

The numerators of S^2_{ws}, S^2_{bs}, S^2_{bt}, S^2_{bd}, and S^2_{bl} are the sums of squares within the sampling error E_{ws}, between the sampling error E_{bs}, between the tank error E_{bt}, between the diet error E_l, and between the location error E_{bl}, the sum of which represents the overall design error E_o:

(16)

$$E_o = E_{ws} + E_{bs} + E_{bt} + E_{bd} + E_{bl}$$

4. Design confidence – Even when the known sources of variability or error have been calculated, there will be uncertainty associated with the experimental design. A quantitative indication of the confidence that can be attributed to the design without performing any experiment can be determined using Working-Hotelling confidence limits (Working and Hotelling, 1929):

(17)

$$y_{\pm} = \left[s_0^2 \times m \times F_{m.m-n} \times \mathbf{x_n} (\mathbf{X^T X})^{-1} \mathbf{x_n^T} \right]^{0.5}$$

The root mean square overall error (S_0) does not depend on the design and its magnitude is estimated after completion of the trials. The terms y, *m*, *n,* and *F* represent the measurement, number of parameters in the model, total number of experiments, and Fisher variance ratio, respectively. The expression $x_n (X^TX)^{-1}x^T_n$ measures the confidence of the design before the trials take place. It is an exclusive function of the design and model and therefore does not depend on experimental error. This term is usually called leverage and is a scalar quantity designated by *h*. Thus, Equation 17 can be rewritten as follows:

(18)

$$y_{\pm} = \left[s_0^2 \times m \times F_{m.m-n} \times h \right]^{0.5}$$

Equation 18 reveals an inverse relationship between confidence and leverage (*h*). When $h = 1$ at any particular experimental point, there is a total lack of confidence in the design and the model and the confidence will be far from the experimental value y. Conversely, when $h = 0$, the design and model have complete influence on the predicted value and confidence bands will be very tight at this point. In other words, it is expected that the model will predict the experimental response with a high degree of confidence.

The generation and use of $x_n(X^TX)^{-1}X_n^T$ in calculating confidence can be explained by using the design in Table 2 where a 4-parameter model of the form described in Equation 1 ($x = \mu + r_L + r_D + r_T$) is proposed. A series of 27 (3×3×3) experiments can study three locations, three diets per location, and three tanks per diet. The 4-parameter model generates the four column design matrix **X** shown in Table 3. The columns of this design show the studied parameters and the rows show every experimental combination $\mathbf{x_n}$. By using **X** and $\mathbf{x_n}$ and performing the matrix operations (multiplication, transposing, and inversion) involved in the expression $h = x_n(X^TX)^{-1}X_n^T$, it is possible to obtain 27 confidence values, one for each experimental point of a 3×3×3 design.

At certain stages of an experiment, it may be better to compute the confidence or leverage over the entire experimental domain instead of at an individual point. Two reasons are: (1) to rapidly monitor how the design influences the model and (2) to obtain a graphic display of the variation of the leverage when a parameter in the design is altered, even when the response is not to be measured. The expression $(X^TX)^{-1}$ is a more general equation that allows calculation of the confidence or leverage over an entire 3×3×3 experimental domain. Table 3 shows that the diagonal elements of $(X^TX)^{-1}$ are the intercept and the squared coefficients with magnitudes of 0.037 and 0.056, respectively. The off-diagonal terms *L*, *D*, *T* and *L*×*D*, *L*×*T*, *D*×*T* are the linear and crossed coefficients, respectively, with a null magnitude for the design in Table 2. Thus, the general leverage equation is:

(19)

$$h = 0.037 + 0.056 \times \left(L^2 + D^2 + T^2 \right)$$

By using Equation 19, it is possible to visualize the leverage as a function of the diet (*D*) and tank (*T*) by keeping the location (*L*) constant (Fig. 1a). The leverage is circular and symmetrical with

Table 2. An experimental 3 × 3 × 3 design for studying the confidence of the model $Y = \mu + \gamma_L + \gamma_D + \gamma_T$.

Location	*Diet*	*Tank*
A	1	I
		II
		III
	2	I
		II
		III
	3	I
		II
		III
B	1	I
		II
		III
	2	I
		II
		III
	3	I
		II
		III
C	1	I
		II
		III
	2	I
		II
		III
	3	I
		II
		III

a minimum at the center of the design (high confidence) and increases as the variables move away from the center. By altering the levels in Table 2 the researcher can estimate how leverage changes with the design. For example, if diets cannot be replicated in locations A and C of Table 2, the leverage for this 15-experiment trial (3 × 3 × 5/3) will be:

(20)

$$h = 0.067 + 0.167 \times \left(L^2 + T^2\right) + 0.100 \times D^2$$

The confidence graph for this equation shows that elimination of diet replication at locations A and

Table 3. Design matrix and derivation of a general expression for leverage (*h*) using the matrix $(X^T X)^{-1}$.

	Average	*Location*	*Diet*	*Tank*
X	1	-1	-1	-1
	1	-1	-1	0
	1	-1	-1	1
	1	-1	0	-1
	1	-1	0	0
	1	-1	0	1
	1	-1	1	-1
	1	-1	1	0
	1	-1	1	1
	1	0	-1	-1
	1	0	-1	0
	1	0	-1	1
	1	0	0	-1
	1	0	0	0
	1	0	0	1
	1	0	1	-1
	1	0	1	0
	1	0	1	1
	1	1	-1	-1
	1	1	-1	0
	1	1	-1	1
	1	1	0	-1
	1	1	0	0
	1	1	0	1
	1	1	1	-1
	1	1	1	0
	1	1	1	1
$(X^T X)^{-1}$	0.037	0.000	0.000	0.000
	0.000	0.056	0.000	0.000
	0.000	0.000	0.056	0.000
	0.000	0.000	0.000	0.056
	1	*L*	*D*	*T*
1	0.037	0.000	0.000	0.000
L	0.000	0.056	0.000	0.000
D	0.000	0.000	0.056	0.000
T	0.000	0.000	0.000	0.056
	0.037	0 L	0 D	0 T
	0 L	0.056 L^2	0 L×D	0 L×T
	0 D	0 L×D	0.056 D^2	0 D×T
	0 T	0 L×T	0 D×T	0.056 T^2
h	0.037 +	0.056 L^2 +	0.056 D^2 +	0.056 T^2

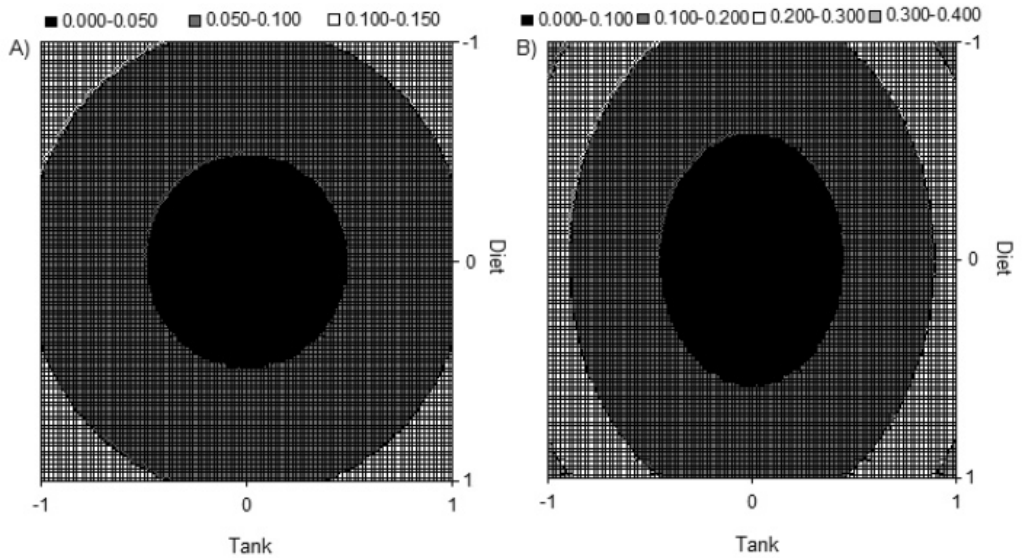

Fig. 1. A confidence graph as a function of the variables in the design presented in Table 2: (a) design 3×3×3, Equation 19, (b) Design 3×3×5/3, Equation 20. The location is kept constant at the center of the design ($L = 0$) while the variables diet (D) and tank (T) vary between +1 and -1 and h is computed.

C influences the shape and magnitude of the confidence (Fig. 1b). The leverage is no longer circular and values are higher than in Fig. 1a. Comparison of Equations 19 and 20 shows that the uncertainty at the center of the design increases 81% ([0.067-0.037]/0.037), while the corners increase a dramatic 125%. Such a comparison shows that confidence and associated h values can be used to compare the merits of different designs prior to performance of an experiment.

5. Power analysis – The study described in Table 2 tests three null hypotheses: (1) there is no difference between tanks at each diet level, (2) there is no difference between diets, and (3) there is no difference between locations. The study also seeks to identify sources of variability by comparing means. The probability that a null hypothesis will be rejected (false) defines the statistical power of the design. The importance and calculation of statistical power in aquaculture by using simple one-way ANOVA have already been reported (Searcy-Bernal, 1994). Power analysis examines whether the effects of various parameters are large enough to detect particular differences or effects. For example, the main goal of the theoretic feeding trials in Tables 1 and 2 is to compare alternative lipid sources to fish oil in fish feed. The random effects (tanks and locations) are not of intrinsic interest but are sources of variation. A small effect for tanks and locations and a large effect for diet would be expected, in other words, power levels would be low for some effects and high for others. The power variability approach in this kind of design was studied by Murphy and Myors (2004).

Although hierarchical designs are frequently used in aquaculture experiments, articles regarding statistical power are few (Ruohonen, 1998; Ling and Cotter, 2003), perhaps due to the complexity of calculating their power and the lack of appropriate computer programs (Ruohonen, 1998).

The statistical power of a design is a function of the degree of freedom, the level of significance, and the variance ratio *F*. By performing the feeding trials and calculating these parameters, the retrospective power of the design in Tables 1 or 2 can be determined by using Laubscher's square root normal equation of a non-central *F* distribution (Laubscher, 1960):

$$z_\beta = \frac{\sqrt{(2w-1)} \times \sqrt{\frac{b \times F_t}{w}} - \sqrt{2b + 2b \times F) - \frac{(b + 2b \times F)}{(b + b \times F)}}}{\sqrt{\frac{bF_t}{w} + \frac{(b + 2b \times F)}{(b + b \times F)}}} \quad (21)$$

where *b* and *w* represent the 'between' and 'within' degrees of freedom of the studied source of variation, also called the numerator and denominator degrees of freedom of an *F*-distribution, F_t is the tabulated Fisher ratio at a 5% significance level, and *F* is the experimental variance ratio of the studied sources of variability. The power (Z1–β) of the design can be computed using the normal percentile value for *Z* β from reported normal distribution tables or by means of the standard normal density function in Excel using the syntax "100*(1-NORMSDIST(Z_1–β))". The term *b* x *F* is generally known as the non-central parameter lambda (λ = *b* x *F*) which measures departure from the null hypothesis (Patnaik, 1949).

6. Application – Aquaculture nutrition experiments largely depend on inferential statistical analysis, implying that the results arise from many independent replicates. If this requirement is not fulfilled, there are strong grounds for questioning whether the study should be undertaken in the first place. Replicate experiments involve extra equipment, laboratory space, animals, feeding formulations, screening of materials entering the experimental units, logistics, labor, and costs. Four approaches have been suggested to cope with this complication: (1) using a microcosm experiment as a model of a large-scale system, (2) focusing on testable predictions within a limited spatial scale, (3) replicating the control but not the treatments, and (4) performing non-replicated experiments (Oksanen, 2001). The first and second approaches are popular in aquaculture experiments, but when problems such as a limited number of tanks, cost of the treatments, logistics, etc., arise, the remaining approaches seem to be practical alternatives and their results are interpreted in the same way as for conventional designs. The only difference is in the quality, or precision, of the mathematical model (Goupy, 1996).

a. *Design selection.* An example of such a study is given in Table 4. Here, feeds containing alternative lipid sources were fed to groups A, B, and C at three levels for each parameter (location, diet, tank). To demonstrate that experimental design principles can be used to analyze the confidence of experiments even with an unequal number of replicates for each treatment and that results can be interpreted in the same way as for a balanced design, group C was tested in duplicate rather than triplicate tanks. Feed performance was indicated by the increase of eicosapentaenoic acid (EPA, C20:5n-3) in the fish liver, estimated by gas chromatography. For the sake of simplicity, the subsampling error was assumed to be negligible. Despite the fact that the researchers in this example were interested only in comparing alternative diets with a fish oil control, such studies are usually performed by statistical geneticists and fish breeders who are interested in effects of tank and location, genetic strains, and their interaction with the sources of variation in the study.

Table 4. Hypothetical study of the increment of eicosapentaenoic acid (EPA, C20:5n-3) in the liver of fish fed a diet containing fish oil (control, diet 1) or substitutes to fish oil (diets 2,3), sampled from three locations in three tanks (three variables), using only two replicates for the control in location C.

Location	*Diet*	*Tank*	*EPA Increment (%)*		
A	1	I	9.35	9.70	10.70
		II	10.38	9.59	10.59
		III	8.34	9.34	10.22
	2	I	8.79	8.65	9.54
		II	8.22	9.20	9.38
		III	9.01	8.34	8.53
	3	I	2.73	2.47	2.84
		II	2.63	2.54	3.66
		III	2.90	2.65	2.92
B	1	I	9.20	9.45	8.75
		II	8.92	9.92	8.46
		III	7.94	8.94	9.00
	2	I	9.00	9.53	8.78
		II	8.78	9.45	8.46
		III	8.91	9.03	9.01
	3	I	2.75	2.84	2.13
		II	2.87	2.63	2.45
		III	2.53	2.37	2.53
C	1	I	9.10	9.67	10.05
		II	8.60	9.60	9.61
	2	I	8.80	9.05	8.55
		II	8.95	8.54	9.02
		III	8.60	9.60	8.66
	3	I	3.03	2.35	2.35
		II	2.56	2.56	2.53
		III	3.02	3.00	2.51

b. *Estimation of the confidence of the design*. The confidence in the overall experiment shown in Table 4 was estimated using the principles described in Section 4, above:

(22)

$$h = 0.039 + 0.059 \times \left(L^2 + D^2\right) + 0.056 \times T^2 + 0.005 \times \left(L - D\right) - 0.007 \times LD$$

As in Fig. 1a, the graphic display of leverage as a function of D and T with $L = 0$ is symmetric and circular with the minimum at the center (Fig. 2). Comparison of the magnitudes of h in Figs. 1a

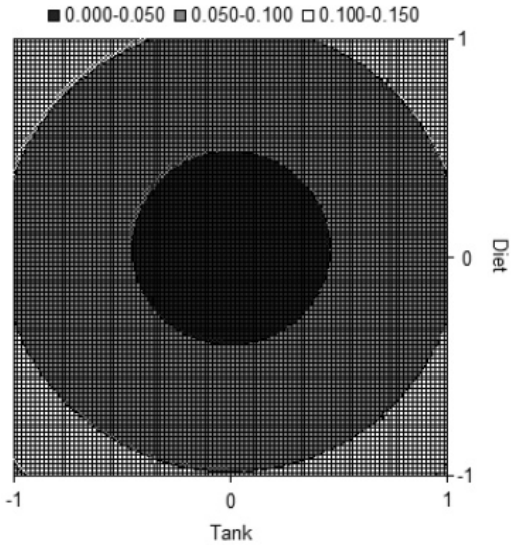

Fig. 2. Confidence graph as a function of diet and tank for the hypothetic design in Table 4. The variable, location, in Equation 22 is kept constant at the center of the design ($L = 0$).

and 2 shows that elimination of one tank inflates h by 5.4% at the center and 7% and 0.7% at the corners. Based on these results, it is the control diet that should be tested in the reduced number of tanks and in the corner of the design where the major increase in h is expected; the alternative diets should be tested in areas of the design where narrow confidence bands are expected (lower h values). Graphic displays of leverage as a function of L and D with $T = 0$ and as a function of L and T with $D = 0$ (graphs not shown) were symmetric and circular with a minimum at the center, similar to Figs. 1a and 2.

c. *Quantification of the sources of variation.* The sources of variation (location, diet, tank) were quantified using Equations 5-15 (Table 5). Assuming that subsampling variation was negligible, Equations 3 and 4 were not used. In Equations 5, 7, 10, and 13, y_{ldtsi} corresponds to every experimental result (% EPA increase) recorded in Table 4. By using the % EPA increase and the equations described above, the magnitudes of variance associated with the tanks, diets, and locations were calculated (Table 6). Variance due to subsampling was not calculated on the assumption that there are no statistical discrepancies among subsamples taken from different parts of the tanks at the different locations.

The F ratios for tank variability indicate that the results from different tanks per diet belong to the same population regardless of the tank. EPA in fish liver increased with all diets and tank average is more representative of the population as a whole at each diet level. The variance between diets largely exceeds the variance within diets at different locations at the 95% level. Major increases of EPA were observed with diets 1 and 2. By omitting diet 3, differences between diets 1 and 2 can be attributed to variations in random sampling at the 95% confidence level (not presented within this article). There were no significant differences between EPA values for each

Table 5. Parameters and averages used to calculate variance for data given in Table 4. Equations used in connection with each average are given in parentheses.

L	D_i	T_{ij}	S_{ijk}	$\bar{y}$ (Eq. 15)	$\bar{y}_{i}$ (Eq. 12)	$\bar{y}_{ij}$ (Eq. 9)	$\bar{y}_{ijt}$ (Eq. 6)
A	$D_A = 3$	$T_{A1} = T_{A2} = T_{A3} = 3$	$S_{A1I} = S_{A1II} = S_{A1III} = 3$	6.90	7.16	9.80	9.92
			$S_{A2I} = S_{A2II} = S_{A2III} = 3$			8.85	10.19
			$S_{A3I} = S_{A3II} = S_{A3III} = 3$			2.81	9.30 9.00 8.93 8.63 2.68 2.94 2.82
B	$D_B = 3$	$T_{B1} = T_{B2} = T_{B3} = 3$	$S_{B1I} = S_{B1II} = S_{B1III} = 3$		6.84	8.95	9.14
			$S_{B2I} = S_{B2II} = S_{B2III} = 3$			9.00	9.10
			$S_{B3I} = S_{B3II} = S_{B3III} = 3$			2.57	8.63 9.11 8.90 8.98 2.57 2.65 2.48
C	$D_C = 3$	$T_{C1} = 2$	$S_{C1I} = S_{C1II} = S_{C1III} = 3$		6.68	9.44	9.61
		$T_{C2} = T_{C3} = 3$	$S_{C2I} = S_{C2II} = S_{C2III} = 3$			8.86	9.27
			$S_{C3I} = S_{C3II} = S_{C3III} = 3$			2.66	8.80 8.84 8.95 2.58 2.55 2.85

location at the different diet levels. Such a degree of consistency between locations allows us to confidently postulate that diet 2 is a good alternative to the fish oil control.

d. *Retrospective power calculation.* In general, researchers use analysis of variance for data analysis and to determine significant *F* values. However, these indicators do not provide information on the statistical power of the study. The statistical power of the design described in Table 4 was estimated using Equation 21 and the results in Table 6 (Table 7). Although the

Table 6. Main errors and significance estimated for data in Table 4. Equations used in connection with each source of error and $F_{calculated}$ are given in parentheses.

		Variance							$F_{calculated}$		
		Tank			Diet		Location		Tank	Diet	Location
L	D	Within (Eq. 7)	Between (Eq. 8)	D	Within (Eq. 10)	Between (Eq. 11)	Within (Eq. 13)	Between (Eq. 14)	(Eq.8;7)	(Eq.11;10)	(Eq.14;13)
A	1	0.550	0.620	1	0.302	129.290	11.163	1.513	1.127	427.977	0.136
	2	0.246	0.117	2	0.162	123.173			0.473	760.249	
	3	0.149	0.051	3	0.137	117.104			0.345	857.818**	
B	1										
	2	0.345	0.243						0.704		
	3	0.135	0.033						0.242		
		0.068	0.023						0.346		
C	1										
	2	0.283	0.169						0.599*		
	3	0.150	0.019						0.126		
		0.078	0.080						1.023		

$F_{2/6}$= 5.143 and $F_{1/4}$= 7.709* (tank) , $F_{2/24}$= 3.403 and $F_{2/21}$= 3.467** (diet), $F_{2/75}$= 3.120 (location) at 95% confidence.

Table 7. Retrospective power analysis of the results presented in Table 6.

Tank						*Diet*						*Location*					
b	*w*	*F*	*Ft*	*Zβ*	*Power (%)*	*b*	*w*	*F*	*Ft*	*Zβ*	*Power (%)*	*b*	*w*	*F*	*Ft*	*Zβ*	*Power (%)*
2	6	1.127	5.143	0.94	17	2	24	427.977	3.403	-24.99	100	2	75	0.136	3.120	1.52	6
2	6	0.473	5.143	1.27	10	2	24	760.249	3.403	-34.13	100						
2	6	0.345	5.143	1.34	9	2	21	857.818	3.467	-35.94	100						
2	6	0.704	5.143	1.14	13												
2	6	0.242	5.143	1.40	8												
2	6	0.346	5.143	1.34	9												
1	4	0.599	7.709	1.27	10												
2	6	0.126	5.143	1.49	7												
2	6	1.023	5.143	0.99	16												

Zβ values were calculated by means of Eq. 21 and transformed into Power (%) values by using the syntax “100*(1-NORMSDIST(*Zβ*))” in an Excel spreadsheet.

hypothetical researchers were not interested in the effects of tank and location, they are included in the statistical power calculation for the sake of thoroughness. By substituting the corresponding degrees of freedom, *F*-calculated and *F*-tabulated at the 5% significant level, the retrospective statistical power of the different parameters can be calculated. Table 7 shows that the alternative lipid sources were compared to fish oil with a statistical power of 100% and that the experimental design given in Table 4 cannot be used to detect differences between tanks or locations. Murphy and Myors (2004) demonstrated that experimental designs may have more power to detect some effects and less power to detect others.

Retrospective power analyses are no substitute for proper planning of research (Cohen, 1990). It is reasonable to change the sampling design or completely reformulate study goals only in the planning stage.

Discussion and Conclusions

The presented application shows that a variety of experiments can be conducted according to the principles of experimental design discussed above, and that results can be interpreted in the same way as for conventional designs. The main advantage of the described approach is that it can be used to estimate independent additive values for each source of variation. Information on the relative sizes of 'between' and 'within' variances for location, diet, tank, and subsampling can be used to design more efficient trials by minimizing the effects of the stronger sources of variance.

Experience suggests that more replicates (at least three) are better. However, in spite of the few degrees of freedom involved in location C, the present study shows that the use of two tanks instead of three can be an acceptable and valid strategy as long as certain considerations are taken into account, such as allocating the control diet to the treatment with the reduced number of tanks.

The experimental design presented in Table 1 and the equations derived from this design can be implemented in large scale nutritional aquaculture experiments. They can be modified to allow participation of a number of countries, diets, tanks, and additional levels of study. An important feature of this design is the visualization of the sources of variation and their interrelation. Visualization of leverage through graphs is an interesting way to study how a given experiment can model the factors. Thus, it can be used to explore the usefulness of designs and contrast approaches.

The design and variability decomposition approaches can be used to study parameters different from those considered in this work. For instance, it could be interesting to estimate variance components within and between genetic strains or lines of aquaculture species and their interaction with the sources of variation included in this study. The error decomposition equations are particularly useful in cases where only the control but not the treatments are replicated or when non-replicated experiments are performed since, with or without replication, the main goal of inferential statistics is to distinguish patterns from scatter. Further, inferential statistics provide an estimate of the probability of obtaining differences between diets or locations as a consequence of sampling or measurement error and random within-site variation.

The appropriate application of power analysis and confidence of an experimental design can be used to obtain maximum information from limited resources and for making critical assessments on what the results of an experiment might tell and not tell about the questions being asked in a study.

Acknowledgements

The author wishes to thank Nico Laubscher for revisiting his square root normal equation after 47 years, Gro-Ingunn Hemre for helpful comments, and Rune Waagbø for his constructive suggestions.

References

Box G.E.P., Hunter W.G. and J.S. Hunter, 1978. *Statistics for Experimenters, An Introduction to Design, Data Analysis and Model Building*. Wiley, NY.

Cochran W.G. and G.M. Cox, 1953. *Experimental Designs*. John Wiley and Sons, London.

Cohen J., 1990. Things I have learned (so far). *Am. Psychol.*, 45(12):1304-1312.

Deming S.N. and S.L. Morgan, 1987. *Experimental Design: A Chemometric Approach*. Elsevier, Amsterdam. 182 pp.

Finney D.J., 1955. *Experimental Design and its Statistical Basis*. Cambridge Univ. Press, London.

Goldan O., Popper D., Kolkovski S. and I. Karplus, 1998. Management of size variation in juvenile gilthead sea bream (*Sparus aurata*). II. Dry food type and live/dry food ratio. *Aquaculture*, 165(3-4):313-320.

Goupy J., 1996. Unconventional experimental designs: theory and application. *Chemometr. Intell. Lab. Syst.*, 33(1):3-16.

Kaufmann P., 1995. Multivariate optimization strategy for liquid chromatography. II. Exploring and finding optimal conditions in the search area of the multidimensional solvent space. *Chemometr. Intell. Lab. Syst.*, 27(1):105-114.

Laubscher N.F., 1960. Normalizing the noncentral *t* and *F* distribution. *Ann. Math. Stat.*, 31(4): 1105-1112.

Ling E.N. and D. Cotter, 2003. Statistical power in comparative aquaculture studies. *Aquaculture*, 224(1-4):159-168.

Murphy K.R. and B. Myors, 2004. *Statistical Power Analysis*. Lawrence Erlbaum Assoc. Inc., NJ. 73 pp.

Oksanen L., 2001. Logic of experiments in ecology: is pseudoreplication a pseudoissue? *Oikos*, 94(1):27-38.

Patnaik P.B., 1949. The non-central χ^2 and *F*-distribution and their applications. *Biometrika*, 36(1-2):202-232.

Plackett R.L. and J.P. Burman, 1946. The design of optimum multi-factorial experiments. *Biometrika*, 33(4):305-325.

Riley J. and P. Edwards, 1998. Statistical aspects of aquaculture research: pond variability and pseudoreplication. *Aquac. Res.*, 29(4):281-288.

Ruohonen K., 1998. Individual measurements and nested designs in aquaculture: a simulation study. *Aquaculture*, 165(1-2):149-157.

Ruohonen K., Koskela J., Vielma J. and J. Kettunen, 2003. Optimal diet composition for European whitefish (*Coregonus lavaretus*): analysis of growth and nutrient utilization in mixture model trials. *Aquaculture*, 225(1-4):27-39.

Searcy-Bernal R., 1994. Statistical power and aquaculture research. *Aquaculture*, 127(4):371-388.

Smart T.S., Riley J. and P. Edwards, 1998. Statistical aspects of aquaculture research: sample sizes for pond experiments. *Aquac. Res.*, 29(5):373-379.

Vielma J., Koskela J., Ruohonen K., Jokinen I. and J. Kettunen, 2003. Optimal diet composition for European whitefish (*Coregonus lavaretus*): carbohydrate stress and immune parameter responses. *Aquaculture*, 225(1-4):3-16.

Winkelman A.M. and R.G. Peterson, 1994. Heritabilities, dominance variation, common environmental effects and genotype by environment interactions for weight and length in chinook salmon. *Aquaculture*, 125(1-2):17-30.

Working H. and H. Hotelling, 1929. Applications of the theory of error to the interpretation of trends. *J. Am. Stat. Assoc.*, 24 (March supplement):73-85.