UNITED NATIONS UNIVERSITY
Fisheries Training Programme

unuftp.is Final Project 2009

STATISTICAL POWER ANALYSIS IN FISH GROWTH STUDIES

Godfrey Kawooya Kubiriza
Department of Zoology,
Makerere University,
P. O. Box, 7062, Kampala, Uganda
kubirizag@yahoo.com; gkubiriza@zoology.mak.ac.ug

Supervisors:

Helgi Thorarensen
helgi@holar.is
Albert K. Imsland
imsland@vortex.is
Holar University College
Akvaplan Niva,Akralind 4, 201 Kópavogi

ABSTRACT

The objective of this project was to estimate the variance in growth studies of different species of fish and use this information to model the expected effects of number of replications and individuals within replication on the statistical power of fish growth studies. Coefficient of variation amongst fish in the same tank (CV-error) and among tanks receiving the same treatment (CV-within) were calculated from 24 independent growth studies on Arctic charr, Atlantic cod, turbot, halibut and tilapia. This information was then used to estimate the effect of number of replications and fish within each tank (n) on the statistical power when effect size was 15% of the grand mean. Furthermore, the effect of replication, and *n* the minimum effect size that would give 80% statistical power was estimated. Both CV-error and CV-within increased during the experiments and in most studies they stabilised. The mean final CV-error was 30.6±1% and ranged from 15% to 56%. The CV-within treatment ranged from 0% to 12% with a mean value of 4.5±0.04%. Both CV-error and CV-within decrease with fish body mass ($P < 0.05$); and their values at termination are positively correlated ($P < 0.05$) with those at stocking. As expected, statistical power increased when *n* and level of replication increased. The statistical power nearly doubled when the level of replication increased from duplicate to triplicate. However, further increase in replication had smaller effect on statistical power. When CV-error and CV-within were at mean levels, the statistical power increased with sample size up to an *n* of 50-100. Beyond that, comparatively less effect is realised. Effect size less than 20% of the grand mean with statistical power over 80% is only attainable when the CV-within is less than 5%. Effect size less than 10% is attainable when CV-within is low and treatments are tested at least in triplicates. The results suggest that when CV-error and CV-within are medium or low the optimum design of growth experiments is to test treatments in triplicates or quadruplicates with *n* of 50-100 individuals. Special effort should be made to keep variance between tanks receiving the same treatment at minimum in order to minimize effect size.

TABLE OF CONTENTS

LIST OF FIGURES **3**

LIST OF TABLES **4**

2 INTRODUCTION **5**
2.1 NATURE OF FISH GROWTH TRIALS IN UGANDA 6
2.2 OVERALL OBJECTIVE OF THE PROJECT 6
2.2.1 Specific objectives *6*

3 LITERATURE REVIEW **7**
3.1 STATISTICAL POWER ANALYSIS AND BIOLOGICAL RESEARCH 7
3.2 MISUSE OF POWER ANALYSIS 7
3.2.1 Retrospective (post hoc/observed) power analysis *7*
3.2.2 Prospective power analysis *8*

4 MATERIALS AND METHODS **8**
4.1 ARTIC CHARR (*SALVELINUS ALPINES*) GROWTH TRIALS 8
4.2 ATLANTIC HALIBUT (*HIPPOGLOSSUS HIPPOGLOSSUS*) GROWTH TRIAL 9
4.3 ATLANTIC COD (*GADUS MORHUA*) GROWTH TRIAL 9
4.4 TURBOT (*SCOPHTHALMUS MAXIMUS* L.) 9
4.5 TILAPIA (*OREOCHROMIS SHIRANUS:* TREWAVAS, 1983) GROWTH TRIAL 9
4.6 DATA ANALYSIS 10

5 RESULTS **11**
5.1 DEVELOPMENT OF CV-ERROR AND CV-WITHIN IN FISH GROWTH STUDIES 11
5.1.1 Development of CV-error over time in fish growth studies *11*
5.1.2 Development of CV-within treatment over time in fish growth studies *14*
5.1.3 Relationship between CV-error and CV—within treatment and fish body mass at experimental termination in Atlantic cod and Arctic charr *17*
5.1.4 Relationship between variance at stocking and Variance at termination of Fish growth studies *18*
5.1.5 Assessment of statistical power in fish growth studies *19*
5.2 THE DIFFERENT SCENARIOS OF CV-ERROR AND CV-WITHIN USED TO MODEL POWER TO DETECT EFFECT SIZE OF 15% OF THE GRAND MEAN USING DIFFERENT SAMPLE SIZE AND REPLICATION LEVELS. 22
5.3 THE MINIMUM EFFECT SIZE TO REACH 80% POWER 24

6 DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS **25**
6.1 DISCUSSION 25
6.2 CONCLUSION 26

ACKNOWLEDGEMENTS **28**

REFERENCES **29**

LIST OF FIGURES

Figure 1: Development of CV-error coefficient of variation as a function of factorial increase in size of Atlantic halibut, turbot and tilapia.....12
Figure 2: Development of CV-error with factorial increase in size in growth studies of Arctic charr. (The different lines represent separate studies) 12
Figure 3: Development of CV-error with factorial increase in size in growth studies of Atlantic cod. (The different lines represent separate studies).....13
Figure 4: Development of CV-within with factorial increase in size of Atlantic halibut, turbot and tilapia. (The different lines represent separate studies).....14
Figure 5: Development of CV-within with factorial increase in size in growth studies of Arctic charr. (The different lines represent separate studies) 15
Figure 6: Development of CV-within with factorial increase in size in growth studies of Atlantic cod. (The different lines represent separate studies).....16
Figure 7: CV-error and CV-within treatment vs. weight of fish at experimental termination in Atlantic cod and Arctic charr growth trials.....17
Figure 8: Relationship between CV-error at stocking and at termination in different fish growth studies 18
Figure 9: Relationship between CV-within at stocking and at termination in fish growth studies 18
Figure 10: The predicted power of a growth study when five treatments are tested under different scenario of CV-error and CV-within (a, b, c, d & e) and effect size of 15% of the grand mean.....23
Figure 11: The effect size (% of grand mean) in growth studies at a statistical power of 80% for five treatments under different scenarios of CV-error and CV-within (a, b, c, d & e) 24

LIST OF TABLES

Table 1: Summary of the experimental protocols (Columns: 2-6) and key results at termination of the different studies (Columns: 7 – 12).....20
Table 2: Different scenarios of CV-error and CV-within used to model statistical power and effect size in fish growth studies.....22

1 INTRODUCTION

Statistical analysis and interpretation is an integral part of biological research (Festing 2006, Ling 2007), and should be well thought-out when experiments are designed. Once considered at the planning stage, the possibility of adequately arriving at the correct conclusions is enhanced (Festing 2006). However, many researchers only consider the statistics at the end of the experiment when mistakes committed at the planning stage are irreversible.

In fish growth studies, statistical analyses are generally based on mixed model ANOVA with hierarchal (nested) design (Scheffé 1959, Imsland *et al.* 2008). In these studies, comparison is made on the effects of different treatments, such as feed type or environmental factors, on the growth of fish. Given the great variability in the responses of fish to treatments in growth studies, more than one fish tank per treatment are used to increase the reliability of these studies (Ling 2007). In fact peer reviewed publications usually do not accept papers that are based on studies without replication. One reason why replication in growth studies is important is that sometimes, significant differences in growth are encountered among replicated tanks that are supposedly exposed to the same treatment. These random incidents may occur without any apparent difference in conditions or treatment. Hence, it is possible that findings derived from one experimental unit could be misleading. Besides, the likelihood of human errors in procedure, observation, recording, computation and/or reporting makes researchers distrustful of un-replicated studies (Rosenthal 1990).

The null hypothesis tested in these growth studies is that all treatment groups come from the same population and that there is no significant difference among the means of different treatment groups. There are two types of errors associated with statistical analysis (Scheffé 1959, Kutner *et al*. 2005, Zuur *et al.* 2010). These are, the Type I error; where a correct null hypothesis is rejected, i.e. a difference observed when in fact there is none (Whalberg 1984, Hoenig and Heisey 2001). A significance level α (critical p value) is usually used to describe the probability of committing Type I error and the fiducial limits are set at less than 5% ($P < 0.05$). Type II error is the second type of error where an incorrect null hypothesis is not rejected, i.e. failure to recognise a difference when it actually exists.

The probability of Type II error is β and statistical power (the probability of detecting a difference when it truly exists) is 1- β. The design of growth experiments should aim for a power of 80% (Hoenig and Heisey 2001, Araujo and Frøyland 2007). Statistical analysis of biological studies has traditionally mainly focused on significance levels and Type I error (Kutner *et al*. 2005, Festing 2006, Araujo and Frøyland 2007) while less or no consideration is given to statistical power and Type II error (Whalberg 1984, Ling and Cotter 2003, Festing, 2006, Ling 2007). This could be because Type II error requires an investigation of the power of the test and thus more difficult to quantify (Araujo and Frøyland 2007).

The statistical power of mixed hierarchical models depends primarily on five factors: (1) The difference in means between treatment groups (effect size), (2) the variance of the data, both of fish within a tank and among tanks within the same treatment, (3) the number of replicate tanks, (4) the number of fish within each tank and the number of treatments tested (Ling and Cotter 2003, Deng 2005, Kutner *et al.* 2005, António *et al.* 2009). The statistical power increases with increased effect size; number of replications, number of individuals, and number of treatments tested and reduced variance (Kutner *et al.* 2005). As a result, it is of importance to design the experiments with as many replications and individuals as possible to secure acceptable statistical power. However, the available facilities and funding to run growth studies are always a constraint. The number of available tanks is always limited and the cost of building

and running research facilities for fish growth studies is high. Moreover, cost of fish for experiments and the staff costs for husbandry and measurements are also high. Another consequence is that utilizing large sample size would make the tests too sensitive to small deviations; and this contradicts the central limit theory (Zuur *et al.* 2010). Therefore, experimental design must strike a balance between acceptable power and available recourses.

1.1 Nature of fish growth trials in Uganda

As in other countries, fish growth trials in Uganda involve confinement of experimental organism (fish) to ponds, tanks, hapas, and/or cages; which are expensive and in limited supply. Furthermore, although labour costs are comparatively low in Uganda, they do add to the total cost of conducting a growth trial. Therefore, it is as important in Uganda as in other countries to optimize the designs of experiments by using the right number of replicates and experimental subjects (fish), if we are to generate reliable results with minimal resources at our disposal.

Another pitfall for the case of Uganda is that most trials are done in ponds and partial sampling is often done because of the difficulty in taking a total harvest. With partial sampling, researchers tend to target at least 30 individuals (Matsiko *et al.* in prep); but these may not necessarily be enough for detecting a significant effect size at the conventional power of 80%. Besides being expensive and difficult to sample, ponds require very big space which is usually not available. Hence, having many replications in pond-based trials, which are common in Uganda, may be impractical. However, increasing the number of fish sampled per pond could improve the power to detect type II error (Ling and Cotter 2003, Deng 2005, Ling 2007) and thus, the reliability of results.

1.2 Overall objective of the Project

The overall objective of this project was to estimate the variance (expressed here as coefficients of variance) in growth studies of different fish species and use this information to model the effects of replication levels and individuals within a replicate/tank on the statistical power and effect size for 80% power in fish growth studies.

1.2.1 Specific objectives

- To determine how the variance (CV-error) and within treatment (CV-within) change during the experiments with increasing body mass of different fish species.
- To compare the CV-error and CV-within for different species and examine their relationship with fish body mass at the termination of experiments.
- To model statistical power for different replication and sample size per tank under differing scenarios of CV-error and CV-within in fish growth studies.
- To assess the minimum effect size required for statistical power of 80%.

It is anticipated that the results of this project shall be useful when estimating power in fish growth studies for many species and will be applicable to studies both in Iceland or Uganda, and elsewhere. For Uganda, which is in its infant stages of aquaculture development, scientists have an uphill task of conducting further research. Thus, information on statistical power generated in this study will provide benchmark information for optimising experimental design for fish studies in nutrition, genetics and environmental effect growth trials in the country.

2 LITERATURE REVIEW

2.1 Statistical power analysis and biological research

The null hypotheses tested in fish growth studies are that there is no difference among treatment groups. Most studies consider primarily the Type I error, which is the possibility of rejecting a true null hypothesis. On the other hand, Type II error, which is the possibility of not rejecting a wrong null hypothesis, is rarely considered (Hoenig and Heisey, 2001, Deng 2005, Ling 2007). The conclusions are primarily based on the significance level (α). However, failure to reject a null hypothesis may be the result of low statistical power when an important effect actually exists and the null hypothesis (of no effect) is in fact false (Hoenig and Heisey 2001). This is a common mistake in biological sciences, because lack of detectable effect may not sufficiently be established by failure to demonstrate statistical significance based on p-values.

The widely accepted practice is to protect the investigator from falsely concluding that a treatment has an effect when indeed it has none (Hoenig and Heisey 2001). However, the extent to which such a conclusion could be reliably taken has always been neglected (Thomas 1997). Consequently, nearly similar growth studies done on the same fish species, under relatively comparable conditions have yielded different inferences when the number of replicates and sample sizes used differ. For example studies by Rana (1981), Macintosh and De Silva (1984), Dambo and Rana (1992), Gall and Bakar (1999) and Abdel-Fattah (2002) drew contentious conclusions about the optimal stocking densities and feeding regimes of Nile tilapia (*Oreochromis niloticus*). These discrepancies are with no doubt a result of setting growth trials with differing statistical power (Ling and Cotter 2003, Ling 2007); but none of these studies mentions power. Thus, there is an increasing demand for scientific backing (particularly based on statistical power) whenever such inconsistent inferences are stated.

2.2 Misuse of power analysis

2.2.1 Retrospective (post hoc/observed) power analysis

Like the name suggests, retrospective power analysis involves looking backwards and examining the potential of a concluded study in detecting a difference that truly existed (Thomas, 1997; Ling and Cotter, 2003; Kutner *et al.,* 2005). This is computed based on the observed value of the test statistic. However, Hoenig and Heisey (2001) strongly warned that retrospective power should be used with prudence. These authors suggested that observed power does not entirely provide evidence for the null hypothesis being true for non-significant statistical tests. This is supported by Steidl *et al.* (1997) and Thomas, (1997) who affirmed that it is misleading when non significant results associated with high observed power, are reported as having high support for the null hypothesis. Steidl *et al.* (1997) and Thomas (1997) reported a one-to-one relationship between p-value and observed power. Thus, a non-significant *p*-value will always correspond to low observed power.

Since observed power is determined by the set *p*-value, it adds nothing to the interpretation of results, once a p-value has been used. Hence, computing the observed power after observing the *p*-value should not be done because it does not change the interpretation derived from the *p*-value (Hoenig and Heisey 2001, Steidl *et al.* 1997). In this regard, retrospective analyses (observed power) are no substitute for involvement of statistical power knowledge in the planning and design of biological growth studies. Consideration of statistical power at experimental planning and design is called prospective power analysis.

2.2.2 Prospective power analysis

The expected power in studies should be estimated in cases where experiments are planned by employing the appropriate sample size and number of replicates to yield biologically significant results whenever they exist (Lenth 2001, Steidl *et al*. 1997). It is also called a priori test where values of minimum effect of size considered biologically significant are set (Thomas 1997), and the likelihood of detecting them by the set study established. This approach enables the researchers to plan experiments with acceptable power using the minimum available resources to gain reliable findings. Moreover, it helps the researcher to reach sound conclusions and to detect difference where it truly exists even where the effect is comparatively small and conversely support hypotheses that are in fact true. Thus, routine incorporation of statistical power analyses into research planning efforts is the only sure way of increasing experimental design efficiency (Steidl *et al*. 1997, Thomas 1997, Lenth 2001).

3 MATERIALS AND METHODS

Data derived from 24 independent fish growth trials on Arctic charr (*Salvelinus alpinus*), Atlantic halibut (*Hippoglossus hippoglossus*), Atlantic cod (*Gadus morhua*), turbot (*Scophthalmus maximus*) and tilapia (*Oreochromis shiranus*) were used in this project. Of the 24 growth studies, 10 were for Arctic charr, 11 for Atlantic cod, while Atlantic halibut, turbot and tilapia each had only one. A brief description of the nature of these growth studies is given below.

3.1 Artic charr (*Salvelinus alpines*) growth trials

All data for Arctic charr were obtained from feed trials conducted at Holar University College, research station. For a series of trials, feeds were tested on fish of six different size classes (0-4g, 4-10g, 30-90g, 90-240g, 240-690g, and 450+ to ≥ 900g) in separate experiments. All size groups of the charr were fed one of six isoenergetic feed formulations containing 34.7-49.2% crude protein (CP). Growth rate, feed conversion ratio (where possible), chemical composition and sensory quality parameters were compared among groups of fish fed different formulations (Sigurgeirsson *et al.* 2008).

In another study, big fish of initial weight 660g were fed on six different types of feed. The ratio of herring fishmeal, herring bone meal and different plant meals were varied in the feed formulations. Individual growth measures of all fish every forth week and the growth response and feed conversion efficiency was estimated (Ólafur Sigurgeirsson pers. com.).

In the third category of the study, feeds with different protein source (fish meal vs. plant proteins) were used. In this study, the groups were not homogeneous initially (difference in initial weight), thus percentage increase in weight of the biomass was considered (Árnason *et al.*, 2008)

In a similar study, juveniles were maintained on 16 different feeds with diverse protein sources in triplicates. Random sampling was done on 250 fish per tank to estimate final weight. In another study, 17 types of feeds with different oil sources, but similar fat contents were investigated (Árnason *et al.* 2008).

3.2 Atlantic halibut (*Hippoglossus hippoglossus*) growth trial

Atlantic halibut data were from a study done at Holar University College that explored the effect of oxygen saturation on the performance (growth, survival, FCR, proximate analysis etc). In this study, fish of relatively uniform initial body mass were maintained at five different levels of oxygen saturation (57%, 84%, 100%, 120%, and 150%) throughout the study time (Thorarensen *et al.* in prep.). Fish were maintained under identical conditions of feeding, temperature and other water parameters.

3.3 Atlantic cod (*Gadus morhua*) growth trial

All the feed growth trials for Atlantic cod used in this study were done at Holar University College. One study examined the effect of five feeds with different protein contents and triplicates were used. A sample size of 12 fish per replicate was used to assess the performance of Atlantic cod (Helgi Thorarensen pers. com.).

In another trial, four treatments were tested involving high and low density of fish and clumped or dispersed food distribution. (Steingrímsson 2010a).

In yet another study, three treatments, were used; in the first one fish could "smell" and see a fish of a similar size eating normal fish feed. In the second, fish could smell and see large fish eating normal fish feed, while in the third, fish could smell and see a large fish eating flesh of other cod. (Steingrímsson 2010b).

Two other studies dealt with two treatments on Juvenile cod; (1) temporally clumped food (fed 4 times over 1 hour), and (2) temporally dispersed food **(**fed 4 times between 9:00-16:30). Fifty fish were maintained at 13°C in each tank (Edelsparre 2004, Pálsson 2006)

3.4 Turbot (*Scophthalmus maximus* L.)

Fish were reared under three different light regimes: continuous light (LD24:0), extended photoperiod (LD16:8) and switched photoperiod. The group undergoing this last treatment was divided into three sub-groups where the fish were first reared under extended photoperiod and then switched to continuous light for 3 months (LD16:8 to LD24:0, then back to LD16:8) at the ages of 9, 13 and 16 months (Le Deuff *et al.* 2010).

3.5 Tilapia (*Oreochromis shiranus:* Trewavas, 1983) growth trial

This study was conducted at Bunda College, University of Malawi to compare performance (survival, growth and FCR among others) of *O. shiranus* at various constant temperatures (20, 25, & 30^{o}C) (See Kassam, *et al.* In press.).

In all the above studies, the effect of several treatments on the growth rate of fish were investigated. Different, but uniform numbers of replicates, treatment levels and sampling intervals were maintained in each growth trial. Sample size used differed amongst experiments.

3.6 Data analysis

Data were analysed using mixed model ANOVA in SPSS. Mean sums of square (MS) for tanks nested within treatments and the error MS were obtained for every data set. Mean sum of squares for error (MS_{error}) from the ANOVA table constituted the error variance (σ_{ε}^2), while the variance among the replicate tanks receiving the same treatment (σ_{β}^2) was computed using the formula stated by Ling and Cotter (2003) as follows.

$$E(\mathrm{MS}_{B(A)}) = \sigma_{\varepsilon}^2 + n\sigma_{\beta}^2$$

Where $E(MS_{B(A)})$ is the corrected denominator for ANOVA mean square among subgroups (tanks under a similar treatment) used to calculate F in a nested model.

The coefficients of variation were then calculated as follows (Ling and Cotter, 2003):

$$\mathrm{CV-error} = \frac{\sqrt{\sigma_{\varepsilon}^2}}{\overline{X}}$$

$$\mathrm{CV-within} = \frac{\sqrt{\sigma_{\beta}^2}}{\overline{X}}$$

where $\overline{X}$ is the treatment grand mean.

The effect size (d), i.e. the maximum difference between treatment means and grand treatment mean were calculated. The treatment levels (a), number of replicates (b) and sample size (n) were noted for each growth trial. These values were used to estimate the variance of the group means as shown below.

$$s_{\bar{Y}}^2 = \frac{\mathrm{MS}_{B(A)}}{nb} = \frac{\hat{\sigma}_{\varepsilon}^2}{nb} + \frac{\hat{\sigma}_{\beta}^2}{b}$$

The group mean variance above was then used to derive the non-centrality parameter (λ) using the formula below (Ling and Cotter, 2003).

$$\phi = \sqrt{\frac{\lambda}{a}}.$$

The value of $\emptyset$ in this formula was also computed as follows

$$\phi = \sqrt{\frac{d^2}{2a\sigma_{\varepsilon}^2/n}}.$$

Where d is the separation of the means (the maximum separation in the case of more than two treatments), and σ_{ε}^2 is the error variance, while a and n are the treatment levels and sample size respectively (Ling and Cotter 2003).

The estimated power was calculated using the non-centrality parameter with the corresponding degrees of freedom (df) obtained from the treatment level and number of replicates (Ling and Cotter, 2003) in G*Power 3.0 software (www.softpedia.com).

Using the results of the statistical analysis the expected power in growth studies was calculated using different levels of replication (2 – 6) and number of fish in each tank (10 – 1000). The estimated power was calculated by first calculating the non-centrality parameter based on different values for CV-error and CV-within and then using G*Power 3.0.

Furthermore, the minimum effect size was to give a power of 80% was calculated by first using G*Power to estimate the λ required to get 80% power for a given degree of freedom and then estimating the effect size required to give this λ using a model written in Excel.

4 RESULTS

4.1 Development of CV-error and CV-within in fish growth studies

Data were available from several studies to estimate how CV-error and CV-within developed over time in growth experiments. The variance in growth studies was compared by calculating the coefficient of variation of fish within tanks (CV-error) and the coefficient of variation for tanks within treatments (CV-within). The CV-error and CV-within were calculated at different times during growth trials to examine how or if these values differ during growth studies.

4.1.1 Development of CV-error over time in fish growth studies

CV- error in halibut, turbot and tilapia studies

In Atlantic halibut the CV-error increased from 23% at the beginning of the experiment to 33% and remained fairly constant at 32% as the experiment progressed till termination (Figure 1). The CV-error for tilapia increases progressively from 11% at stocking to 37% at the end of the experiment (Figure 1). However, it is likely that CV-error in tilapia growth trial did not reach a stable level before the experiment ended. Contrary to the development of CV-error in halibut and tilapia, the CV-error in turbot remained relatively stable (26 – 28%) throughout the study period (Figure 1).

CV- error in Arctic charr studies

Several growth studies on Arctic charr were analysed (Figure 2). In these studies, the CV-error at stocking ranged from 12% to 17% (Figure 2). Across studies, CV-error initially increased slightly and then remained constant. However, in one growth study the CV-error exceeded 30% and appeared not to have stabilized at termination when the factorial increase in fish body mass was about 1.5 times the stocking size (Figure 2). The rest of the growth studies for Arctic charr had a fairly constant CV-error in the range of 12 to 21%; with a mean of 23.2 ± 6.9% at termination.

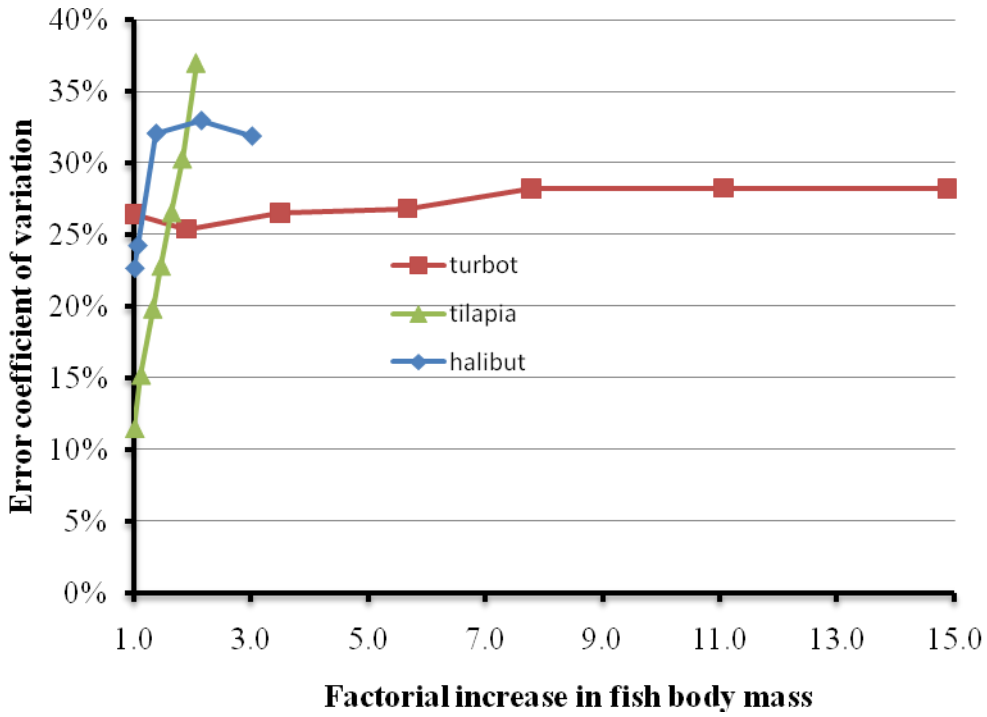

Figure 1: Development of CV-error coefficient of variation as a function of factorial increase in size of Atlantic halibut, turbot and tilapia.

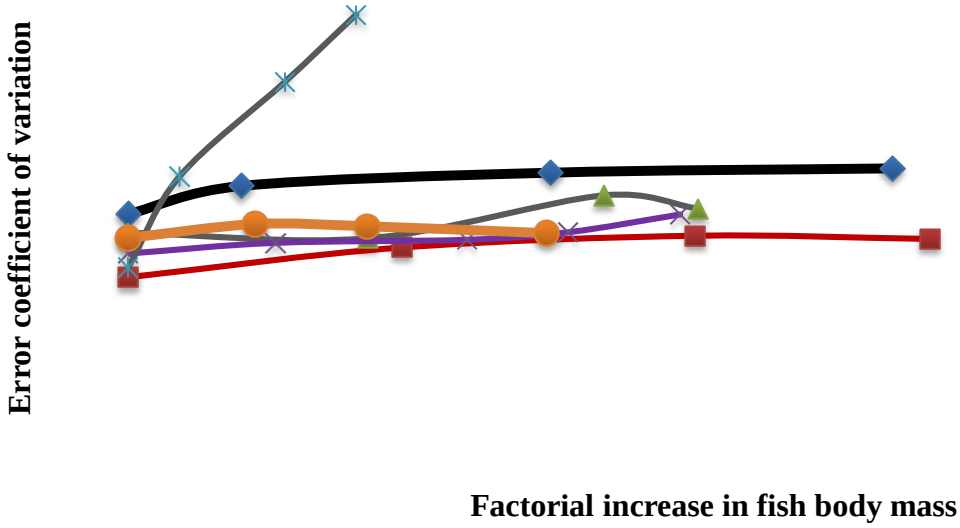

Figure 2: Development of CV-error with factorial increase in size in growth studies of Arctic charr (The different lines represent separate studies).

CV- error in Atlantic cod studies

In Atlantic cod growth studies, initial CV-error ranged from 26% to 29% and increased steadily with factorial increase in size to a maximum in different studies (Figure 3). One growth study had a relatively high CV-error right from the beginning and consistently exhibited a CV-error higher than the rest of the studies over time.

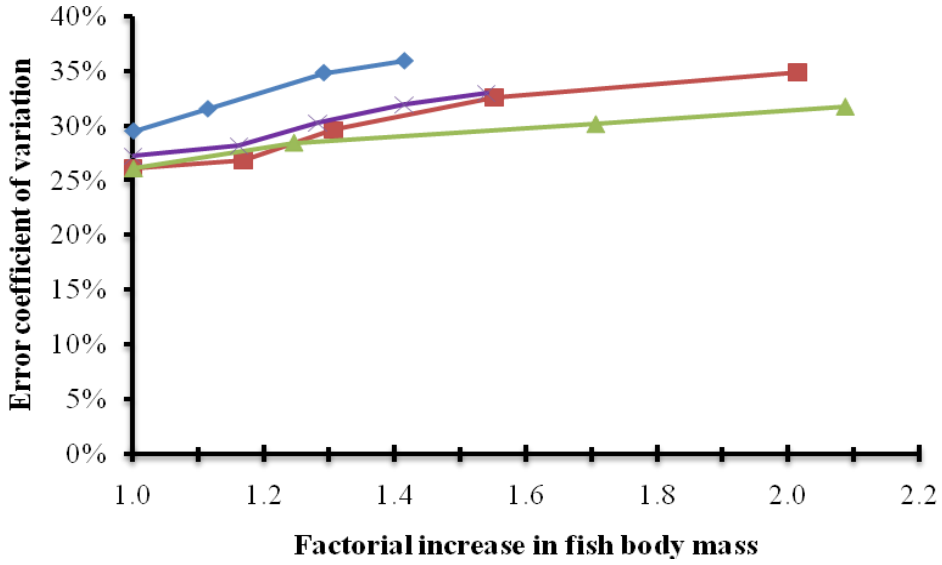

Figure 3: Development of CV-error with factorial increase in size in growth studies of Atlantic cod (The different lines represent separate studies).

At termination, the mean CV-error was 38.3 ± 7.8% and ranged from 32 to 56%. In all studies, the CV-error increased with factorial increase in fish body mass (Figure 3). Stable levels of CV-error were between 30 and 40% in Atlantic cod and were reached when the fish size was about 1.5 times the body mass at stocking.

4.1.2 Development of CV-within treatment over time in fish growth studies

CV-within treatment in halibut, turbot and tilapia studies

There was a general increase in CV-within treatment in the growth studies of halibut, turbot and tilapia to a maximum; after which it dropped and remained fairly stable (Figure 4). Overall, CV-within treatment was higher in growth studies for both turbot (3 – 11%) and tilapia (2 – 11%) than in those for halibut (0 – 2%) as fish increased in size over time (Figure 4).

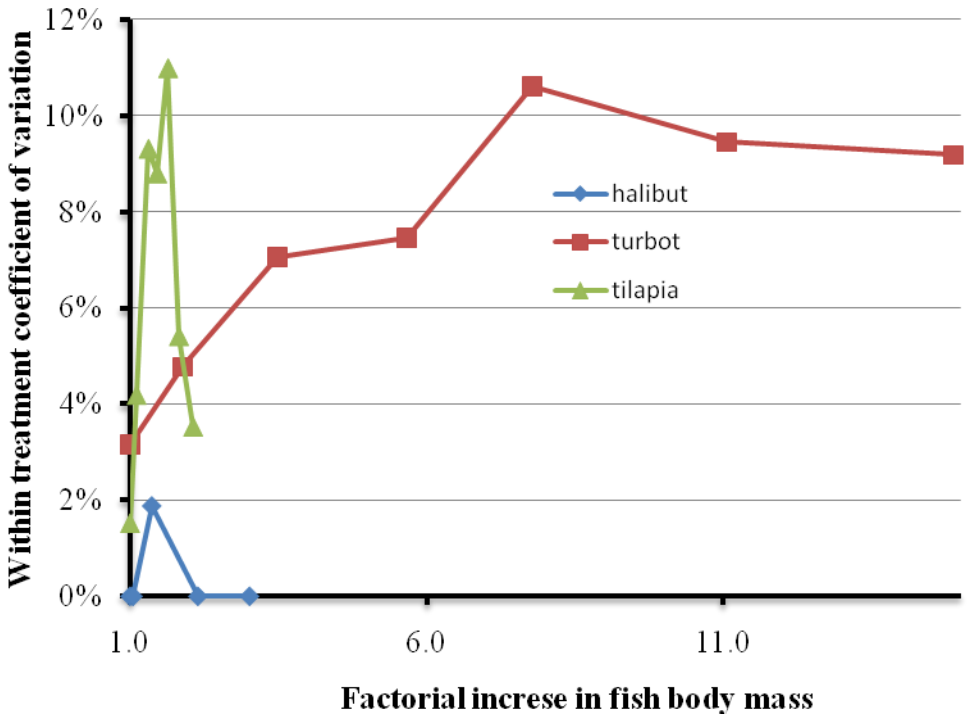

Figure 4: Development of CV-within with factorial increase in size of Atlantic halibut, turbot and tilapia (The different lines represent separate studies).

The growth studies for the three species exhibited fluctuating CV-within treatment over time. In halibut, CV-within (%) was zero at stocking, two when fish were 1.4 times the initial weight and zero from when fish doubled in weight onwards. In tilapia it was two at stocking, maximum (11%) when fish was about 1.6 times the stocking body mass and four at termination when the body mass doubled. Turbot growth trial exhibited a CV-within (%) of three at stocking, maximum of 11% when the fish was eight times the initial body mass and stabilized at nine from when the body mass was 11 times the initial body mass to termination. There was a significant tank effect ($P < 0.05$) both in the turbot and tilapia growth trials but not in the halibut trial ($P > 0.05$). In tilapia, significant tank effect set in when it hit 4 -11% and remained significant ($P < 0.05$) for a factorial increase in fish body mass in the range of 1.1 – 1.6, but did not completely mask the treatment effect. Within treatment variance and therefore tank effect dropped to 5 – 4% and became non-significant ($P > 0.05$) for factorial increase in fish body mass of 1.8 and 2.1, respectively. Only treatment effect remained significant ($P < 0.05$) at this stage. In turbot, significant tank effect ($P < 0.05$) vs. non-significant treatment effect ($P > 0.05$) set in when the fish body mass was about two times that at stocking.

CV-within treatment in Arctic charr studies

In all Arctic charr growth studies, CV-within (%) ranged between zero and three at stocking and zero and nine at termination. Five of the six growth studies were terminated with CV-within in the range of zero to five; and one study with nine percent (Figure 5).

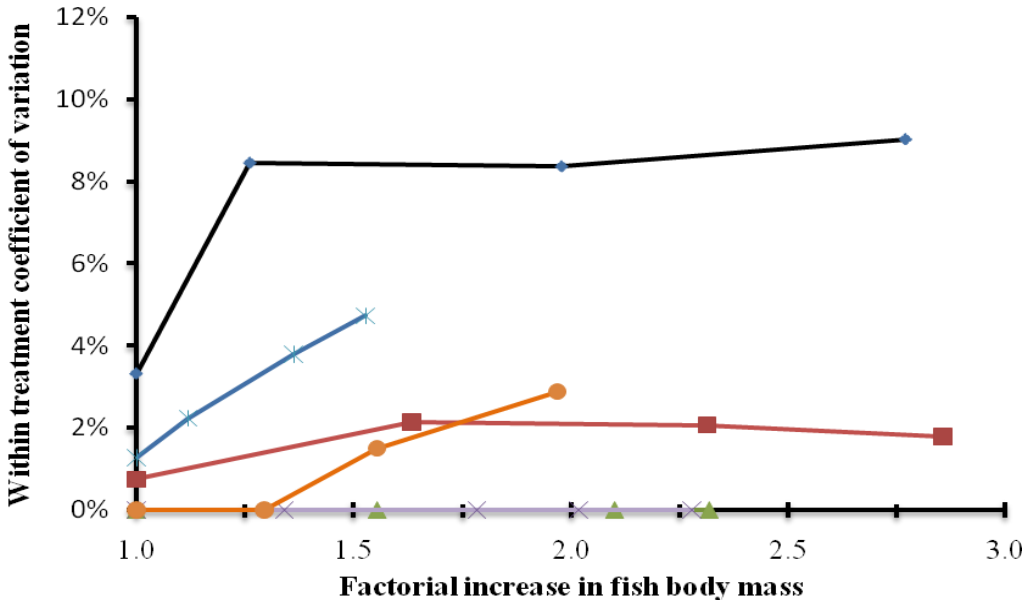

Figure 5: Development of CV-within with factorial increase in size in growth studies of Arctic charr (The different lines represent separate studies).

In one study, a within treatment tank effect of 3% was significant ($P < 0.05$) from the onset, and its effect increased to nearly 9% throughout the study. Studies set with within treatment variance less than two exhibited significant tank effect ($P < 0.05$) when fish body mass 1.5 – 2 times that at stocking, and later ceased.

CV-within treatment in Atlantic cod studies

In all Atlantic cod growth trials used in this project, CV-within treatment values were less than 10% (Figure 6). The CV-within was between zero and five percent (Figure 6). Average CV-within was 5.5 ±

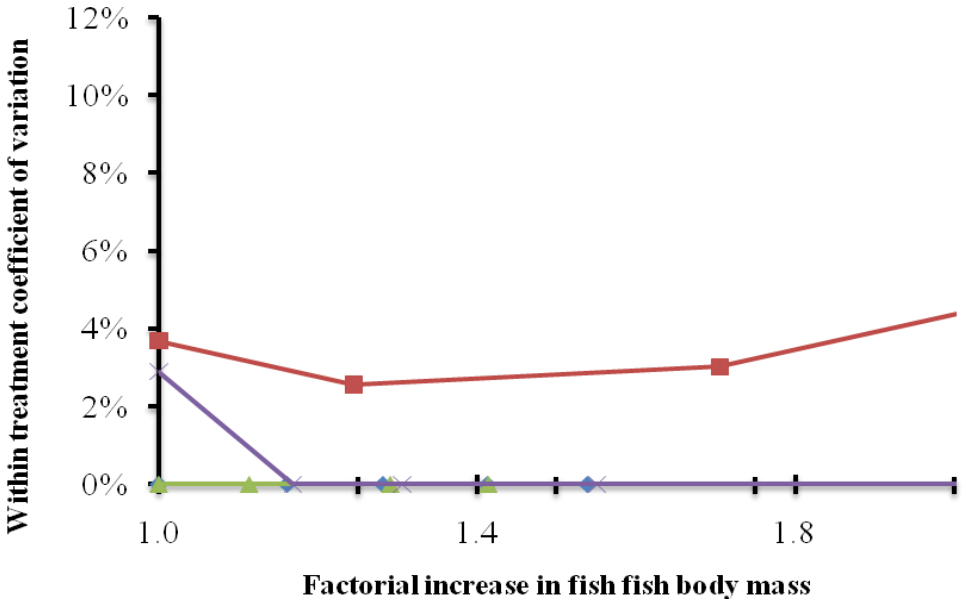

Figure 6: Development of CV-within with factorial increase in size in growth studies of Atlantic cod (The different lines represent separate studies).

4.1.3 *Relationship between CV-error and CV—within treatment and fish body mass at experimental termination in Atlantic cod and Arctic charr*

Both the CV-error and CV-within treatment tended to decrease with increased body mass at the termination of the experiments in data derived for Atlantic cod and Arctic charr (Figure 7). For both species, there was a significant decrease ($P < 0.05$) in CV-error and CV-within treatment with increasing body mass.

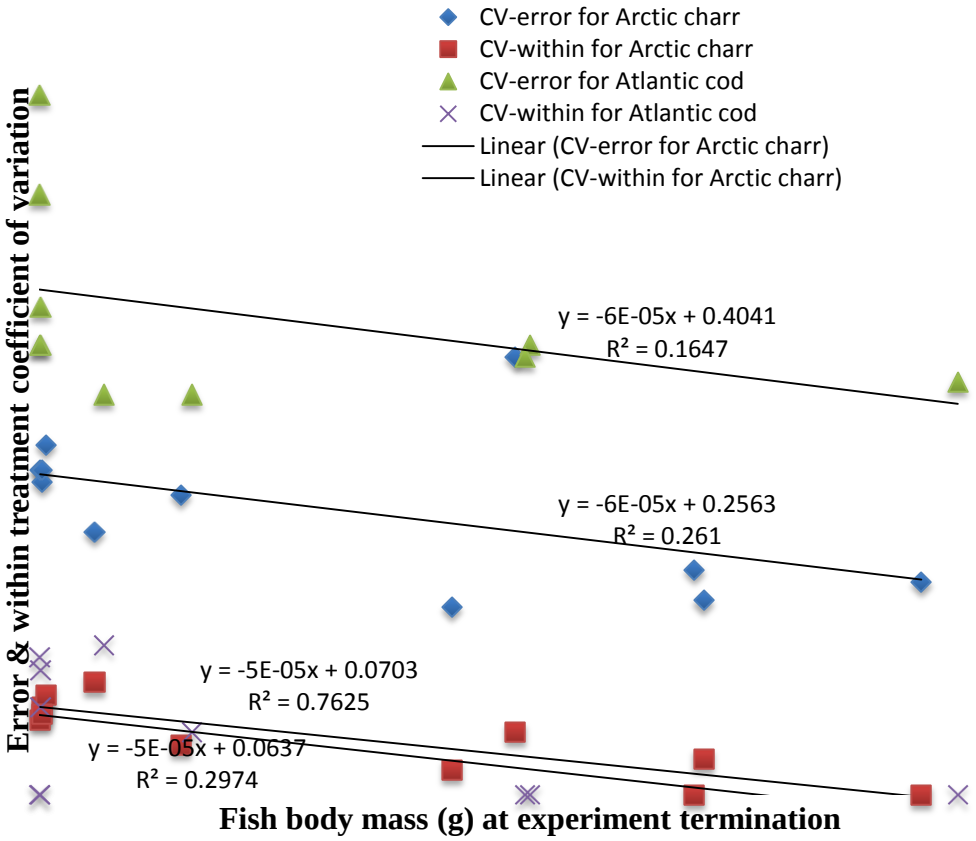

Figure 7: CV-error and CV-within treatment vs. weight of fish at experimental termination in Atlantic cod and Arctic charr growth trials.

ANCOVA showed significantly higher CV-error in Atlantic cod (mean ±SD; n = 10) = 38.3 ±7.8% ($P < 0.05$) than in Arctic charr (mean ±SD; n =11) = 22.8 ± 6.1% at termination. There were no significant differences in CV-within between species ($P > 0.05$). The slopes of the lines (Figure 7) for CV-within were not significantly different ($P > 0.05$); indicating that CV-within treatment (mean ± SD) for Atlantic cod (4.5 ± 5.1%) and Arctic charr (4.6 ± 3%) growth studies were similar.

4.1.4 Relationship between variance at stocking and Variance at termination of Fish growth studies

Regardless of the experimental protocol (level of replication, treatment levels, and fish per tank) and/or body mass of fish at stocking, growth trials with high CV-error at stocking correspondingly exhibited high CV-error at termination (Figure 8).

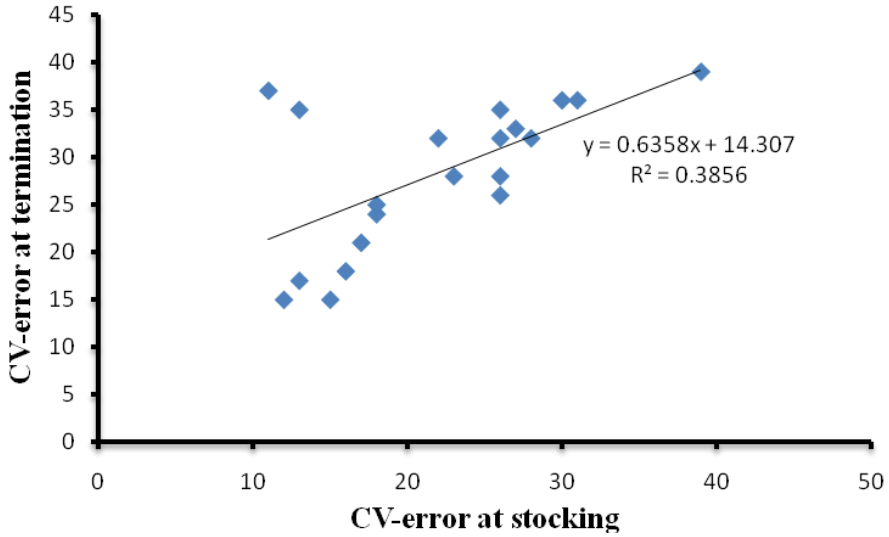

Figure 8: Relationship between CV-error at stocking and at termination in different fish growth studies.

The results suggest that the CV-error at stocking affects the CV error at termination (Figure 8). Similarly, once variability amongst tank under the same treatment is high from the onset, growth trials will yield high variation at termination (Figure 9). Studies with low CV-within at stocking will equally have low variance at termination. Keeping other factor constant, CV-within at termination approximately has a 1:1 positive linear relationship with CV-within at stocking (Figure 9).

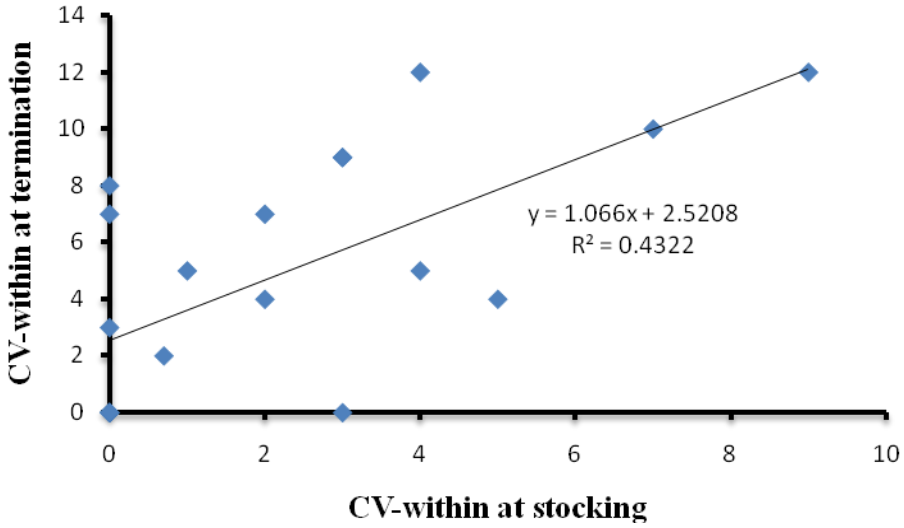

Figure 9: Relationship between CV-within at stocking and at termination in fish growth studies.

The relationship in Figures 8 & 9 affirms that initial variability amongst fish in a given tank and tanks under similar treatment at stocking greatly affect CV-error and CV-within treatment at termination, respectively. Increasing either sample size/level of replication or both may perhaps not compensate for the high variance at the start of a growth trial. Hence, the need to grade fish and ensure tank uniformity at experimental set up is emphasized.

Fish body mass at stocking does not seem to affect variance at termination. The rank of CV-within tanks for different growth studies on Arctic charr growth trials was 32 - 90 g > 558 – 776 g > 235 – 672 g > 550 – 1082 g > 613 – 1438 g = 460 –1068 g. These values revealed no clear correlation between fish body mass at stocking and development of either CV-error or CV-within treatment during the study. Similar findings existed in Atlantic cod growth trials. Thus, fish body mass at stocking appears not to influence the development of CV-error and CV-within treatment during the study.

4.1.5 Assessment of statistical power in fish growth studies

Results from analysis of growth data derived from 24 growth trials is summarised in Table 1 below. The data in the table show a wide range of values for the error and within treatment coefficients of variation. However, CV-error and CV-within values show no apparent relationship (Table 1).

The variance in Atlantic cod growth trials was generally high (CV-error >30). For Arctic charr growth trials, fish were less variable (CV-error <30). However, the values do not show a clear relationship between CV-error and effect size.

The CV-error ranged from 15% to 56% at the end of the studies (Table 1), with mean of 30.6 ± 1.0% for all growth trials. Slight differences existed amongst fish species with cod and tilapia growth studies having the highest CV-errors of 38.3 ± 7.3% and 37% at termination, respectively. Growth studies for Arctic charr and turbot showed the least variance of 21.6 ± 7.3% and 28%, respectively. The overall CV-within treatment ranged between 0 and 12% (Table 1) with a mean value of 4.5 ± 0.04%. There was no apparent difference in CV-within among species.

Observed power (Table 1) ranged from 12 to 100%, with a mean of 53.9 ± 0.3%, while minimum detectable effect size ranged between 4 and 56% of the grand mean with an average of 18.1 ±12.8%.

Table 1: Summary of the experimental protocols (Columns: 2-6) and key results at termination of the different studies (Columns: 7 – 12).

Trial	Nature of experiment	Species	Level	#tanks	n	Average final wt (g)	d (%) of mean wt	CV-error	CV-within	Observed power	Min. effect size (%) at 80% power
1	Effect of Oxygen saturation	halibut	5	3	47	122.0	24	0.32	0.00	0.99	11
2	Effect of photoperiod	turbot	3	3	36	330.3	30	0.28	0.09	0.44	36
3	Thermal effects on growth	tilapia	3	6	16	11.3	56	0.37	0.04	1.00	33
4	Feeds with different Crude protein content	Arctic charr	7	4	50	4.7	30	0.25	0.07	1.00	22
5	Feeds with different Crude protein content	Arctic charr	7	4	39	10.9	17	0.28	0.08	0.49	28
6	Feeds with different Crude protein content	Arctic charr	6	4	50	90.0	12	0.21	0.09	0.23	32
7	Feeds with different Crude protein content	Arctic charr	6	3	35	230.8	11	0.24	0.04	0.34	21
8	Feeds with different Crude protein content	Arctic charr	6	3	132	672.8	4	0.15	0.02	0.40	8
9	Feeds with different Crude protein content	Arctic charr	6	3	64	1067.9	4	0.18	0.00	0.20	9
10	Feeds with different Crude protein content	Arctic charr	6	3		1437.5	10	0.17	0.00	0.98	15
11	Feeds with different Crude protein content	Arctic charr	6	3	96	886.7	17	0.39	0.06	0.55	27
12	Feeds with different Crude protein content	Arctic charr	16	3	30	2.3	37	0.26	0.06	1.00	33
13	Feeds with different fat content	Arctic charr	6	3	90	1082.9	6	0.16	0.03	0.23	12
14	Feeds with different oil sources	Arctic charr	16	4	151	4.7	19	0.26	0.06	0.97	23

15	Feeding experiment	Atlantic cod	5	3	12	1497.3	13	0.33	0.00	0.60	6
16	Feeding experiment	Atlantic cod	5	3	46	248.7	7	0.32	0.05	0.12	24
17	Different lipid levels (10-26%)	Atlantic cod	5	3	13	800	18	0.36	0.00	0.41	31
18	Feeding trial	Atlantic cod	6	3	15	791.8	20	0.35	0.00	0.46	32
19	Feeding trial	Atlantic cod	6	3	32	105.2	37	0.32	0.12	0.37	55
20	Feeding experiment	Atlantic cod	3	6	56	1.9	16	0.36	0.07	0.38	17
21	Feeding experiment	Atlantic cod	2	9	105	1.8	17	0.39	0.10	0.92	14
22	Clumped/dispersed food	Atlantic cod	2	5	31	0.23	13	0.48	0.11	0.29	28
23	Food clumped/dispersed in time	Atlantic cod	2	5	35	0.52	8	0.36	0.00	0.44	12
24	Food clumped/dispersed in time	Atlantic cod	2	5	14	0.08	13	0.56	0.00	0.13	31

Columns 2 &3 show the origin of the data & fish species investigated. Columns 4, 5 & 6 show the experimental arrangement—numbers of levels, the total number of tanks and fish sampled per tank (n). Column 7 is the grand mean at termination over all tanks of fish weights in grams. Column 8 contains the maximum distance of separation (d) between treatments. Columns 9 and 10 show the error and within group coefficients of variation. Column 11 shows minimum effect size (%) of the experiment at 80% power, while column 12 is the retrospective power of the experiment at termination.

4.2 The different scenarios of CV-error and CV-within used to model power to detect effect size of 15% of the grand mean using different sample size and replication levels.

The results of the previous sections were used to estimate the expected statistical power in growth studies. Five different scenarios were chosen (Table 2).

Table 2: Different scenarios of CV-error and CV-within used to model statistical power and effect size in fish growth studies.

Scenario	CV-error (%)	CV-within (%)
a (mean values)	30.6	4.5
b	35	10
c	15	10
d	35	1
e	15	1

The statistical power was estimated as a function of replications and number of fish in each tank. Estimates were based on a 15% difference between the largest and smallest treatment means (effect size) under the four scenarios (Figures 10a, b, c, d & e). Power and effect size estimates were done for levels of replications (b) from 2 to 6 and number of fish sampled per tank (n) as 10, 25, 50, 100, 250, 500 and 1000.

Generally, the results show (as expected), that power increases with both the number of replications and fish sampled per replicate. It is clear that CV-error and in particular CV-within affect the power. When CV-error and CV-within are average (Figure 10a) 60-80% statistical power is reached when replicate tanks are three or four and the n is about 100. When CV-within increases the statistical power is reduced. CV-error appears to have less effect on power and CV-within as low as 1% (Figures 10d & e).

In growth studies where both CV-error and CV-within are high (Figure 10b), and in those with low CV-error but high CV-within (Figure 10c), statistical power of 80% cannot be achieved for effect size of 15% regardless of sample size for a range replicates between 2 and 6. In fact where CV-error is low, but CV-within is high, the minimum effect size is 20-40% for replication levels from three to six and near 60% for duplicates (Figure 11c).

Increasing the number of fish sampled has minimal effect on the achievable power (Figures 10b & c). An increase in sample size yields slight increase in power when replication is five onwards. The results demonstrate that when CV-within is high increased n has little effect on statistical power. However, when CV-within is low both increased sample size and number of replicates improves statistical power (Figures 10d & e). Power more than doubles when the replication level is increased from a duplicate to triplicate; for sample size ranging from 25 to 100 in studies with high CV-error, but low CV-within (Figure 10d).

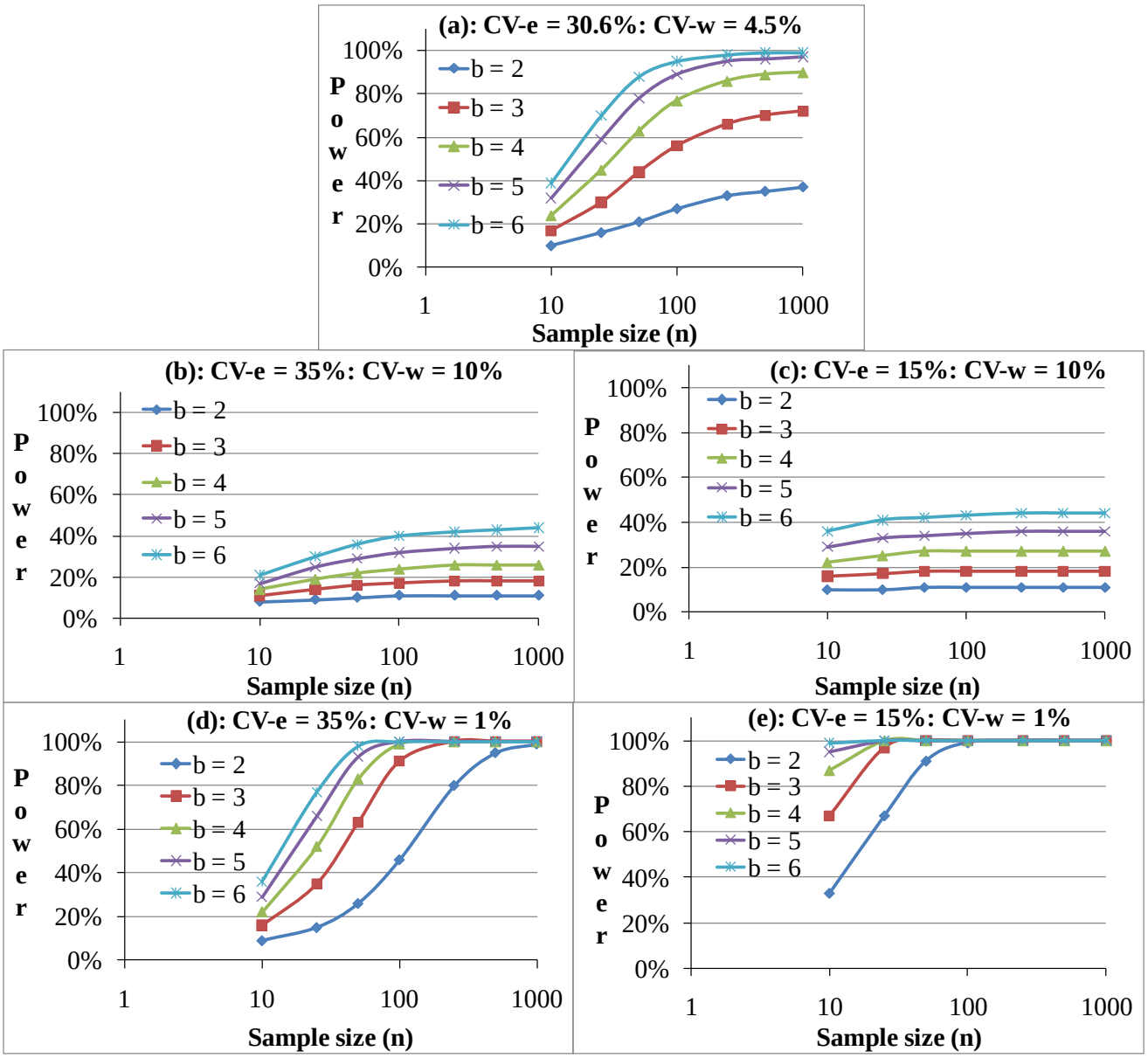

Figure 10: The predicted power of a growth study when five treatments are tested under different scenario of CV-error and CV-within (a, b, c, d & e) and effect size of 15% of the grand mean.

The results show that in studies with high CV-error and minimal CV-within a power of 80% can be achieved with duplicates when 250 fish per tank are sampled (Figure 10d). In instances of both minimal CV-error and CV-within (Figure 10e), power of 80% with effect size of 15% can be achieved by sampling about 80fish per tank when duplicates are used. Higher replication beyond duplicates in such studies yields power of more than 80% when a sample size slightly higher than 25 individuals is taken from every tank. In instances of either average (Figure 10a) or minimal (Figures 10c & e) CV-within, marked increase in power is achieved when the replication level is increased from a duplicate to triplicate; regardless of the magnitude of CV-error. Further replication reduces the sample size required to achieve power of 80% with effect size of 15%. The increase in power is however much less when the level of replication is beyond triplicates. Similarly, increasing the number of fish above 100 individuals per tank may not improve power much regardless of the level of replication.

4.3 The minimum effect size to reach 80% power

The minimum effect size required to reach 80% power was calculated using the same scenarios as before. Generally, the minimum effect size (with a power of 80%) is reduced as the number of replications increases (Figures11a, b, c, d & e) and it is reduced by a third when the number of replications is increased from duplicate to triplicate. The effect size is not significantly reduced when the N is increased above 100. Lower minimum effect size can be detected in growth studies with average CV-error and CV-within (Figure 11a), than in studies with high CV-within. In studies with minimal CV-within in combination with either high or low CV-error (Figures 11d & e), much less effect size can be detected than in studies with average values CV-within.

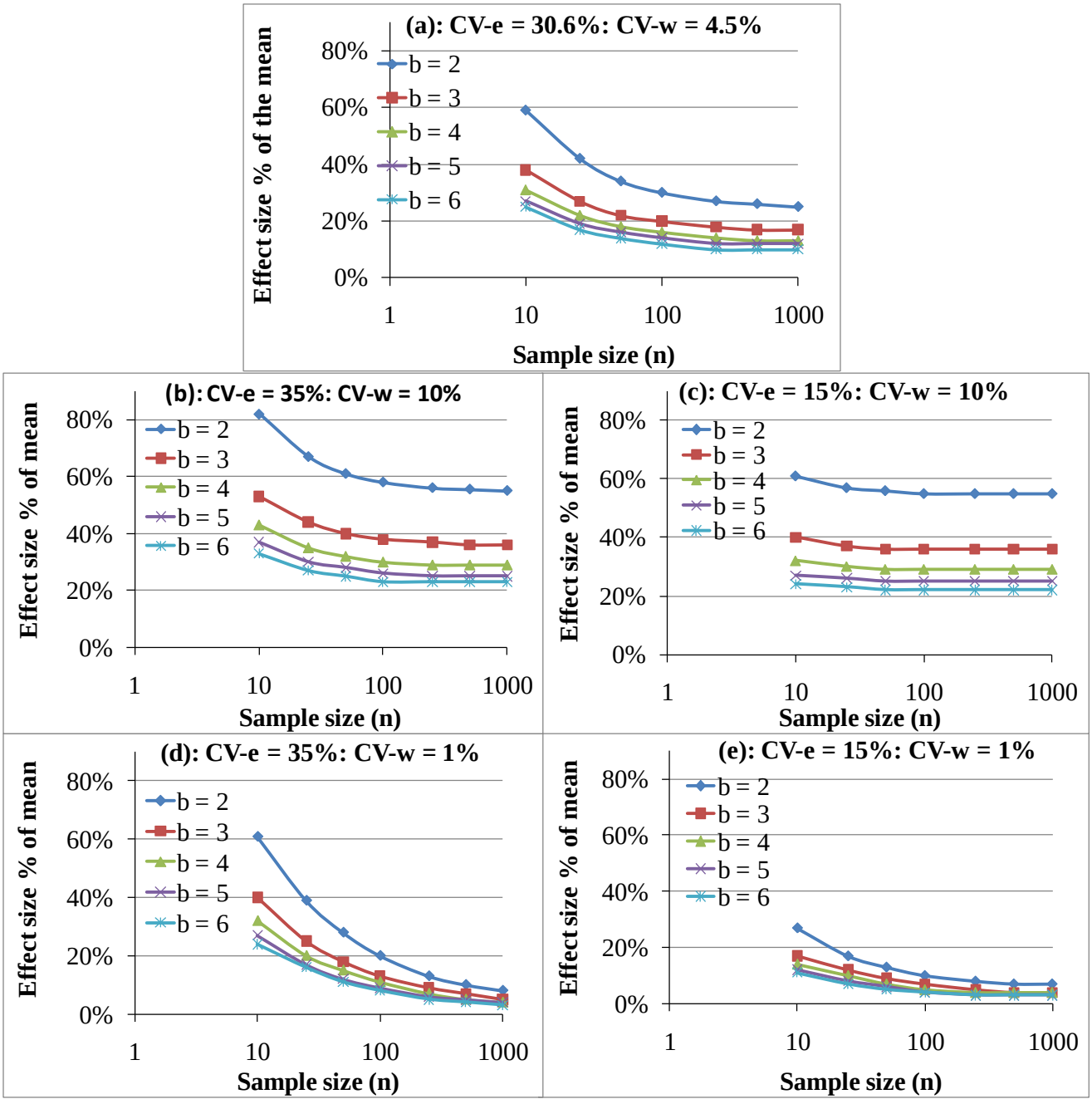

Figure 11: The effect size (% of grand mean) in growth studies at a statistical power of 80% for five treatments under different scenarios of CV-error and CV-within (a, b, c, d & e).

Regardless of CV-error magnitude, effect size less than 20% of grand mean is not possible in studies with high CV-within (Figures 11b & C) even with replication level as high as six. Under these scenarios, increasing the number of fish sampled per tank does not lower the effect size.

A combination of high CV-error and minimal CV-within (Figure 11d) yields effect size of less than 10% when about 100 fish are sampled per tank when replication is quadruplicate or more. Under this scenario, less than 20% effect size is equally possible for a sample size of 100 fish in both duplicates and triplicates.

In studies where both CV-error and CV-within are minimal (Figure 11e), effect size far less than 10% can be realised when less than 100 fish are sampled per tank for triplicate replication level or more. With two replicates, sampling close to 100 fish would equally yield effect size less than 10% in such scenarios.

Overall, the minimum effect size decreases greatly when replication is increased from two to three while increases beyond triplicate or quadruplicates has less effect on minimum effect size. The optimum sample size per replicate to achieve a minimum effect size is 100 fish per tank, regardless of the replication level. Increasing n to more than 100 fish does not notably lower effect size.

5 DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

5.1 Discussion

This study provides information about variance of data in fish growth studies, which is important background information for efficient design of growth studies. The CV-error ranged from 15% to 56% with a mean value of 30.6 ± 1%; while CV-within treatment was 0 – 12% with a mean value of 4.5 ± 0.04%. These values are similar values as reported by Ling and Cotter (2003) for growth studies on Atlantic salmon where the CV-error was 14 – 41% with a mean of 28 ± 8.6%, and CV-within 0 – 7% with a mean of 3.2 ± 1.9%. Thus, the present findings and those by Ling and Cotter (2003) are comparable and applicable to other fish species. However, there may be some differences among species. For example the variance in the data on Atlantic cod and tilapia were larger than in data on turbot and Arctic charr. It is not clear if this difference reflects random occurrences between studies on the same species or if this reflects actual differences among species possibly caused by differences in behaviour and individual competition or human error.

The variance in data on Arctic charr and Atlantic cod was reduced with increasing body mass and this may suggest that variance within a species could be size dependent. This is could be linked to proportionately higher accuracy when larger fish are weighed compared with small fish. Besides, bigger fish tend to grow consistently at lower growth rate compared to the smaller ones. Thus, the bigger the fish from which measurements are taken, the lower are the individual variability and the possibility of the random human errors. Therefore, regardless of the fish species being investigated, researchers may need a bigger sample size to compensate for the high variance in smaller fish than in larger ones.

Results from several growth trials demonstrated that variability (CV-error) increases from experimental set up but tends to stabilize as the experiment progresses. CV-within treatment displays the same pattern, but tends to stabilize at a much lower level. In most studies, significant tank effect is realised when CV-within treatment exceeds four percent. The final CV-error and CV-within at termination clearly depend on their initial values. Therefore, a special emphasis should be placed on having CV-error and particularly CV-within as low as possible at the beginning of the experiment. This can partly be achieved

by stocking fish that are graded, ensuring equal size and size distribution of fish in all tanks and also by ensuring uniform rearing conditions from experimental onset.

It may also be possible to increase the fidelity of growth studies by using a different statistical approach such as a three-way hierarchical design where data from several measurement dates is analysed together (Ling 2007). Repeated measures design of individually tagged fish or mixed linear models (Imsland *et al*. 2010) can also enhance the reliability of fish growth studies results especially where significant tank effect is suspected. Alternatively, running growth studies for relatively longer time would as well increase the effect size and, thus statistical power.

The optimum design of growth experiments is where adequate statistical power is reached at a minimum cost. The present findings suggest that the minimum resources required to reach a statistical power of 70 – 80% are to test treatments in triplicates or quadruplicates with 50-100 individuals in each tank. Utilizing this design seems to be optimal because statistical power increases with sample size up to an n of 50-100 individuals per tank, and nearly doubles when replication is increased from duplicate to triplicate. Designs with replication levels and n beyond triplicates and 100 individuals are more expensive and yet they do not improve statistical power appreciably.

5.2 Conclusion

Conclusions from this study were drawn and presented based on the set objectives, as follows.

How does variance (CV-error) and within treatment (CV-within) change during the experiments with increasing body mass in different fish species?
Both CV-error and CV-within increase from stocking and tend to stabilize as the experiment progresses. CV-error ranged from 15% to 56% in different fish growth studies with a mean value of 30.6 ± 1%, while CV-within treatment was 0 – 12% with a mean value of 4.5 ± 0.04%.

How does CV-error and CV-within for different species compare and how does it relate with fish body mass at the termination of experiments?
Arctic charr and turbot had considerably lower CV-error relative to other species used in the present study; however, CV-within was fairly constant across treatments. Thus, CV-error appears to differ amongst species, but CV-within treatment is species independent.

For Arctic charr and Atlantic cod, both CV-error and CV-within decrease with increasing fish size. Size of fish at stocking does not influence the subsequent variance as fish increases in size over time. Individual fish variation (CV-error) is both species and fish size dependant.

How does variation in CV-error and CV-within affect the expected statistical power when different replication and sample size per tank are used in fish growth studies?
Increased variance among fish tanks under similar treatment (CV-within) greatly reduces the statistical power and detectable effect size in fish growth studies. The variance amongst fish in a given tank (CV-error) appears to have less effect on statistical power.

With minimum CV-within, statistical power increases with sample size up to an n of 50-100 individuals per tank, and nearly doubles when replication is increased from duplicate to triplicate. Further increases in n and replication levels beyond 100 individuals do not improve statistical power appreciably.

In most studies acceptable power is reached when treatments are tested with either triplicates or quadruplicates with 50-100 individuals in each tank. This design seems to strike the balance between available resources and acceptable statistical power.

What is the minimum effect size required for statistical power of 80%?
In growth studies where both CV-error and CV-within are high and in those with low CV-error but high CV-within, statistical power of 80% cannot be achieved for effect size of 15% regardless of sample size for a range replicates between two and six.

In growth studies with low CV-error coupled with high CV-within, the minimum effect size is 20-40% of the grand mean to reach statistical power of 80% for replication levels from three to six. For duplicates the minimum effect size it is even larger.

Regardless of the magnitude of CV-error, effect size of ≤10 – 15% is possible at 80% power when 50 – 100 individuals are sampled per tank when CV-within is minimal.

ACKNOWLEDGEMENTS

I am indebted to the United Nation University – Fisheries Training Programme (UNU-FTP) for funding this project. Special thanks are due to Dr. Tumi Tomasson-the Programme Director, Mr. Thor Asgeirsson-the deputy Programme Director and Ms. Sigridur Kr. Ingvarsdottir-the Programme Manager, for their immense support during the course.

I acknowledge the interest and patience of my main advisor, Professor: Helgi Thorarensen for constantly shaping my thoughts. Professor Albert Imsland is also thanked for the valuable comments on this write up.

The research team at Holar University College (particularly Ólafur Sigurgeirsson, Stefán Óli Steingrímsson and Professor Helgi Thorarensen) and Messrs Maurice Ssebisubi and Lazaro W.W. Jere of Bunda College, University of Malawi; are thanked for providing their experimental data.

Besides, nearly everyone at Holar University College constituted a friendly and supportive family away from my family! I am indebted to you all.

I am grateful to all the lecturers (too many to mention by names), for sharing their wealth of knowledge with us during the six months of training in Iceland. The private and Icelandic government institutions visited during the course are also greatly acknowledged.

Makerere University, particularly the head of Zoology Department Dr. Anne M. Akol is greatly thanked for supporting and allowing me take six months leave to undertake this valuable course.

I thank my wife and children for keeping fond memories of a husband and a father respectively, whilst I was a way!

To the 2009/2010 UNU-FTP fellows, you have been added to the list of my friends!

LIST OF REFERENCES

Abdel-Fattah, M. E. (2002). Effect of stocking density and feeding levels on growth and feed efficiency of Nile tilapia (*Oreochromis niloticus* L.*)* fry. *Aquaculture research* 33: 621 – 626.

António T. Álvaro R. and Teresa C. (2009). Statistical Power Analysis with Microsoft Excel: Normal Tests for One or Two Means as a Prelude to Using Non-Central Distributions to Calculate Power. *Journal of Statistics Education* 17(1):1 – 21.

Araujo, P. and Frøyland, L. (2007). Statistical power and analytical quantification. *Journal of Chromatography B* 847:305–308.

Árnason, J., Sigurgeirsson Ó. I., Jónasson, B., Thorarensen, H. and Björnsdóttir, R. 2008. Plöntuhráefni í bleikjufóðri í stað fiskimjöls og lýsis. [February, 2010] <http://avs.is/media/avs/Skyrsla_10-08.pdf>.

Dambo, W. B. and Rana, K. J. (1992). Effect of stocking density on growth and survival of Nile tilapia (*Oreochromis niloticus* L.) fry in the hatchery. *Aquaculture and Fisheries management* 23: 71-80.

Deng, H. (2005). Does It Matter If Non-Powerful Significance Tests Are Used in Dissertation Research? *Practical Assessment Research and Evaluation* 10 (16): 1- 13.

Edelsparre, H. A. (2004). Áhrif fóðrunartíðni á vöxt, stærðarbreytileika og sjálfrán meðal þorsksseiða (*Gadus morhua*) "The effect of feeding frequency on growth, size variability, and cannibalism in juvenile Atlantic cod (*Gadus morhua*)." A final report for the Icelandic Student Innovation Fund.

Festing, M. F. W. (2006). Design and Statistical Methods in Studies Using Animal Models of Development. *International League of Associations for Rheumatology*, 47 (1):5-14.

Gall, G. A. E. and Bakar, Y. (1999). Stocking density and tank size in the design of breed improvement programs for body size of tilapia. *Aquaculture* 173:197 – 205.

Hoenig, J. and Heisey, D. M. (2001). The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis. *The American Statistician* 55:1-6.

Imsland, A.K., Gústavsson, A., Gunnarsson, S., Foss, A., Árnason, J., Jónsson, A., Sráradóttir, H., Arnarson, I. and Thorarensen, H. (2008). Effects of reduced salinities on growth feed conversion efficiency and blood physiology of juvenile Atlantic halibut (*Hippoglossus hippoglossus* L.). *Aquaculture* 274: 254-259.

Imsland, A.K., Gunnarsson, S., Ásgeirsson, Á., Kristjánsson, B., Árnason, J., Jónsson, A.F., Sráradóttir, H. and Thorarensen, H. (2010). Long term rearing of Atlantic halibut at intermediate salinities: effect on growth and blood physiology. *Journal of World Aquaculture Society* 41:1-9.

Kassam, D., Ssebisubi, M. and Likongwe, J.S. (In press.). Effect of temperature and feed rate on the body shape of *Oreochromis shiranus* (Trewavas, 1983); a widely-cultured tilapia species in Malawi. (In press: *Journal of Aquaculture Research*).

Kutner, H. M., Nachtsheim, J. C., Neter, J. and Li, W. (2005). *Applied Linear Statistical Models.* McGraw-Hill/Irwin. New York, USA.

Le Dauff, S., Thorarensen, H., Imsland, A. K. and Gunnarsson, S. (2010). Growth and feed utilization in juvenile turbot (*Scophthalmus maximus* L.) under extended, switched and continuous photoperiods. Fraedathing landbunadarins. Hotel Soga, Reykjavik. February 18-19, 2010. PP233-234.

Lenth, R. V. (2001). Some Practical Guidelines for Effective Sample-Size Determination. Department of Statistics, University of Iowa. [December, 2009] <http://www.stat.uiowa.edu/techrep/tr303.pdf>.

Ling, N. E. (2007). Efficient analysis of growth trial data. *Aquaculture Research* 38:728 – 732.

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

Macintosh, D. J. and De Silva, S. S. (1984). The influence of stocking density and food ration on fry survival and growth in *Oreochromis mossambicus* and *O. niloticus x O. aureus* male hybrids reared in a closed circulated system. *Aquaculture* 41:345 – 358.

Matsiko, D. S., Kubiriza, K. G., Karisa, H. C., Balirwa, J.and Rurangwa, E. (In prep.). Growth performance, body composition and feed utilization efficiency in all male Nile tilapia (*Oreochromis niloticus* L.) fed on two commercial feeds in Uganda.

Pálsson, F.M. (2006). Áhrif fóðrunartíðni á vöxt þorskseiða (Gadus morhua) [The influence of feeding frequency on the growth of juvenile cod (*Gadus morhua*)]. A final report the Icelandic Student Innovation Fund.

Rana, K. J. (1981). Effect of rearing conditions, age and size on the performance of normal hormone treated *Sarotherodon mossambicus* fry, with observations on fry abnormalities. MSc. Thesis, University of Sterling, Sterling.

Rosenthal, R. (1990). Replication in Behavioral Research. In Neuliep, J.W. (Ed.). *Handbook of replication research in the behavioral and social sciences.* [Special Issue.] *Journal of Social Behavior and Persona.* 5(4):1-30.

Sigurgeirsson, Ó. I., Árnason, H. J., and Ólafsdóttir, A. (2008). Protein requirements of Arctic charr. Lokaskýrsla til AVS-rannsóknarsjóðs. Verkefni *nr*: R040-07-08 : [February, 2010] <http://avs.is/media/avs/Proteintorf_bleikju.pdf>.

Scheffé, H. (1959). *The analysis of variance*. John Wiley and Sons, New York, USA.

Steidl, R. J., Hayes, J. P. and Schauber, E. (1997). Statistical Power Analysis in Wildlife Research. *The Journal of Wildlife Management* 61(2):270-279.

Steingrímsson, Ó. S. (2010a). Áhrif fóðrunartíðni og þéttleika á ung þorskseiði. [Effect of food distribution and density on young cod]. A final report to AVS RandD Fund of Ministry of Fisheries in Iceland.

Steingrímsson, Ó. S. (2010b). Áhrif sjálfráns á atferli og vöxt þorskseiða. [Effect of cannibalism on the behavior and growth of juvenile Atlantic cod]. A final report to VS The fisheries project fund.

Thomas, L. (1997). Retrospective Power Analysis. *Conservation Biology* 11(1): 276-280.

Whalberg, H. J. (1984). Improving the productivity of America's schools. *Educational Leadership* 41(8):19-27.

Zuur, A.F., Elena N., Ieno, E. N. and Elphick, C. N. (2010). A protocol for data exploration to avoid common statistical problems. British Ecological Society: *Methods in Ecology and Evolution* 1:3 – 14.