

NOTE

Sampling effort required to obtain repeatable average size estimates of juvenile fish

Damian Moran*

Department of Biology, Lund University, Sölvegatan 35, 22362 Lund, Sweden

ABSTRACT: Larval and juvenile fish cohorts often have skewed size distributions due to inter-individual variation in growth rates. A resampling analysis was used to investigate how much sampling effort is required to achieve repeatable average size estimates in populations of juvenile yellowtail kingfish *Seriola lalandi* Valenciennes. One hundred juveniles (30 d post-hatch) were measured for total length and wet weight from 9 cohorts reared in captivity. From an analysis of cohort size distribution the mean was determined to be a suitable measure of the average for length, while the median was best for weight. One thousand resamples of mean length and median weight were performed for each cohort data set, with sample sizes ranging from 2 to 100 individuals. The point at which increasing sample size resulted in only a negligible increase in precision (variance between resamples <0.1%) corresponded to $n = 19$ for mean length and $n = 30$ for median weight. Since yellowtail kingfish display a magnitude of size heterogeneity similar to other marine species with small pelagic larvae reared in captivity, the sample sizes determined in the present study can act as a general guideline for larval fish studies.

KEY WORDS: Sampling effort · Sample size · Size heterogeneity

— Resale or republication not permitted without written consent of the publisher —

INTRODUCTION

Experimental studies of fish growth should ideally be carried out using individuals that can be identified and re-measured, as this longitudinal sampling method best captures inter-individual variation. It is often impractical or impossible to follow individual larval or juvenile fish over time, due to their small size; hence, there is a strong reliance on population statistics such as mean individual length or weight to study growth and survival experimentally. Experimental studies of larval or juvenile fish growth have used a wide range of sample sizes for average length or weight measurements (5 to 300 individuals), but typical sample sizes range between 10 and 30 individuals (e.g. Folkvord et al. 1996, Baras et al. 2003, Clemmesen et al. 2003, Kestemont et al. 2003, Smith & Fuiman 2003, Moran 2007).

Despite the reliance on average size measurements in larval and juvenile fish studies, little has been pub-

lished on the relationship between sample size and statistical precision. Ling (2007) and Ling & Cotter (2003) provide a useful approach in calculating how much sampling effort should be invested in fish growth studies to provide adequate statistical power to detect differences; however, the power analysis techniques they describe rely on the assumption of normally distributed size variation. The size variance of newly hatched marine fish larvae is relatively low (coefficient of variation of weight, CV_W : 8 to 20%; Kestemont et al. 2003); and a normal distribution is a reasonable approximation of the size distribution. However, the highly variable growth rates of larvae or juveniles mean that cohorts often develop into skewed size distributions, with CV_W values typically in the range of 20 to 50% or higher (Folkvord & Otterå 1993, Baras et al. 2000, 2003, Kestemont et al. 2003, Moran 2007, Imsland et al. 2009). The variable and skewed distributions of juvenile fish prevent the use of standard error

*Email: d.moran@microcosmofscience.com

minimization and power analysis to assess the effect of sample size on estimation accuracy. Resampling can be used to test the effect of sample size on estimation accuracy in non-normal distributions if one has a large enough subsample of a population (Maindonald & Braun 2003). This bootstrapping approach was used in the present study to test the effect of sample size on average length and weight estimates in independently reared populations of juvenile yellowtail kingfish *Seriola lalandi*. Yellowtail kingfish produce large numbers of small transparent pelagic larvae (Moran et al. 2007) that exhibit considerable size heterogeneity from metamorphosis onward (CV_W : 35 to 55%; Moran et al. 2007, 2011, this issue). This developmental profile is characteristic of many pelagic marine fish species (Kunz 2004); therefore, the results of the present study can act as a general indication of the sampling effort that should be expended to achieve repeatable average size estimates.

MATERIALS AND METHODS

The data used to analyse sampling effort came from an experiment reported by Moran et al. (2011), which tested the effect of conspecific density on growth and mortality of juvenile yellowtail kingfish. The juveniles were reared from eggs obtained from captive broodstock maintained at the National Institute of Water and Atmospheric Research Bream Bay Aquaculture Park, located in the north of New Zealand. Although the parentage was unknown, previous observations had shown that eggs were almost certainly from a single female and 1 or 2 males (Moran et al. 2007). Larvae were reared in 9 replicate tanks using standard live feed protocols for pelagic marine fish until 30 d post-hatch (DPH) (Moran et al. 2011). Metamorphosis occurred at 20 to 23 DPH, at which point fish were termed juveniles. The juveniles were not fed on the final day (30 DPH) when the size measurements occurred. At the end of the experiment, all fish in each replicate were collected and a random subsample of 100 individuals was taken. The fish in the subsample were dispatched via an overdose of anaesthetic, and the total length (L_T) was measured under a dissecting microscope at 2 \times to 6 \times magnification to the nearest 0.1 mm. Individual blotted wet weight (W_W) then recorded to the nearest 1 mg using a microbalance (Model WA205, Oertling). The measurement precision for L_T was approximately 0.5% and for W_W approximately 0.7%. The range of CV_W values recorded for the 9 cohorts was from 35 to 55% (Moran et al. 2011, see their Table 1).

Analysis of the cohort length and weight distributions showed that mean L_T was an appropriate measure of

average length; however, the weights had a strong positive skew and an approximate log-normal distribution; therefore, the median W_W was deemed the most appropriate average for weight (Moran et al. 2011). The geometric mean could also be used as an appropriate average for individual weight, as the population geometric mean is the population median for a log-normal distribution (Shih & Binkowitz 1987). However, according to an analysis by Shih & Binkowitz (1987), the sample median is a better measure of the average than the geometric mean at moderate sample sizes ($n = 30$ to 50) for non-ideal log-normal distributions (i.e. containing outliers or observations from mixed distributions). Given that the individual weights analysed by Moran et al. (2011) were not perfect log-normal distributions, the median was deemed to be more appropriate than the geometric mean. To determine the sampling effort required to obtain repeatable average length and weight estimates, a resampling analysis was carried out for mean L_T and median W_W from the 100 observations recorded per cohort. The resampling analysis (with replacement) was carried out in Microsoft Excel using the methods described by Christie (2004). Sample size varied from 2 to 100, and for each value of n , 1000 resamples were carried out for mean L_T and median W_W for each cohort. The 1000 values resulting from the resampling were assumed to be normally distributed given the high number of resamples carried out. This was checked by performing a Shapiro-Wilk W -test of normality on the resampled data set when $n = 2$, with the expectation that if the 1000 resamples were normally distributed for $n = 2$, then the same would be true when $n = 3$ to 100. None of the Shapiro-Wilk W -tests were significant ($p > 0.05$) for the 9 cohorts with $n = 2$, so the mean and standard deviation of the 1000 resamples was then calculated. The CV was used as a relative measure of variance to evaluate the effect of increasing n on the variability between resamples. The rate of change of CV with sample size was evaluated by calculating the ΔCV for each value of n , which was subsequently plotted against n . An exponential decrease function ($y = ae^{-bx}$) was used to infer the value of n where the rate of change in CV was 0.1%, and this was interpreted as the point at which increasing sample size resulted in negligible reduction in the variation between average length and weight measurements.

RESULTS

The degree of variance in average L_T and W_W in the resampling analysis diminished rapidly with increasing sample size (Fig. 1). The variability between resamples was considerably higher for W_W compared to L_T , as can be seen in the higher CV values for W_W in

Fig. 1. The point at which increasing sample size resulted in only a negligible increase in precision ($\Delta CV = 0.1\%$) corresponded to $n = 19$ for mean L_T and $n = 30$ for median W_W .

DISCUSSION

The measurement precisions for L_T and W_W were comparable, and the degree of precision sufficient to allow individual size variation to significantly outweigh measurement error. Another potential source of variation that may have influenced the spread of data was the variation in gastrointestinal fullness. Although gastrointestinal evacuation rates have not been published for yellowtail kingfish larvae or juveniles, evacuation rates for other marine species at comparable developmental stages are in the range of from 4 to 10 h (Bochdansky & Deibel 2001). The juveniles in the current study were not fed on the day of sampling; therefore, it is reasonable to assume that the gastrointestinal tract was mostly purged and not a significant source of variance among individuals.

The resampling analysis showed that less sampling effort was required to achieve a similar level of repeatability for mean L_T ($n = 19$) compared to median W_W ($n = 30$). This was to be expected given that Moran et al. (2011) reported that W_W distributions were more skewed than L_T distributions for juveniles of this species. These sample size estimates do not say how much sampling effort should be expended to achieve a particular level of statistical precision, as statistical precision is concerned with predicting sample size in order to achieve a particular level of variance, and this term pre-supposes that the population variable of interest is normally distributed. The sample sizes predicted in the current study give estimates of the sampling effort required to achieve a repeatable average, and, in this way, the present study differs from that of Ling & Cotter (2003) and Ling (2007), which used normally distributed fish size populations to calculate the sampling effort needed to detect significant differences between

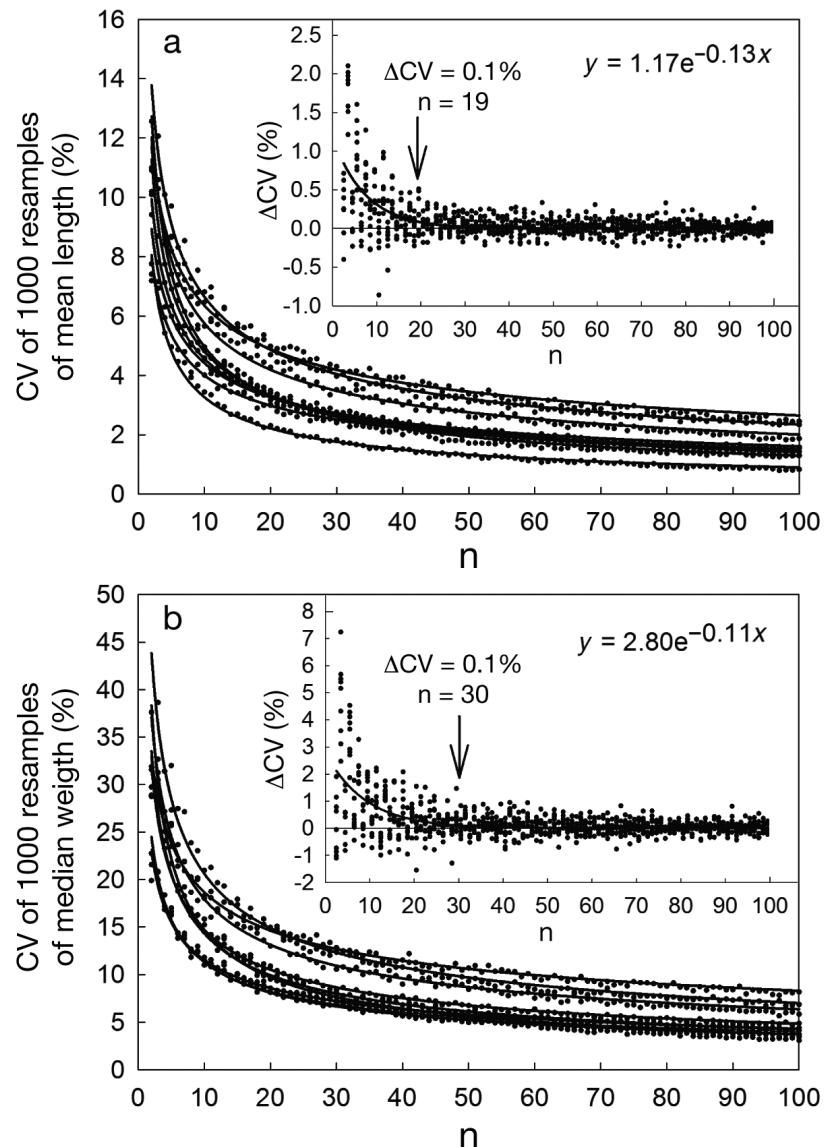


Fig. 1. *Seriola lalandi*. Resampling analysis of (a) mean total length (L_T) and (b) median wet weight (W_W). Each data point in the large graphs represents the coefficient of variation (CV) of 1000 resamples of mean L_T or median W_W from different sample sizes (n). The 9 curves represent different populations for which 100 L_T and W_W measurements were made and constitute the pool of data from which the resamples with differing values of n were taken. The insets represent the rate (ΔCV) at which the variance between resamples decreases with n . A single curve was fitted to the ΔCV data to estimate the value of n for which $\Delta CV = 0.1\%$ (arrow), the point at which increasing n was deemed to give a negligible reduction in the variation between average L_T and W_W measurements

treatments. An important assumption to bear in mind when applying the findings from the present study is that the individuals should be selected at random from a population. This can be a logically difficult task if the juveniles are distributed unevenly throughout the water column (e.g. if smaller individuals seek refuge around objects).

The populations of juvenile *Seriola lalandi* surveyed in the present study can be considered highly skewed, given the magnitude of CV_W and the large difference between mean and median W_W (Moran et al. 2011). The CV_W values for the 9 cohorts studied (30 to 55 %) were comparable to the CV_W values reported for other species at a similar life stage (Folkvord & Otterå 1993, Sakakura & Tsukamoto 1998, Baras et al. 2000, 2003, Kestemont et al. 2003, Imsland et al. 2009); therefore, the sample size estimates from the current study can act as a guideline for researchers who would like some assurance that they have adequately sampled juvenile populations in order to gain a repeatable average size measurement.

In conclusion, resampling analysis of 9 replicate populations of juvenile *Seriola lalandi* indicated that, to obtain repeatable average size estimates of mean L_T and median W_W , it would be necessary to randomly sample at least 19 and 30 individuals, respectively, from a cohort. Providing a random sample of individuals can be taken, these sample sizes can act as a guideline for researchers studying other fish species with similarly distributed juvenile populations.

Acknowledgements. I thank C. K. Smith, P. S. Lee and S. J. Pether for the experimental work, and M. Wellenreuther for improving earlier versions of the manuscript. I also thank the anonymous reviewers for their suggestions. This work was funded by the New Zealand Foundation for Research Science and Technology, and I was supported by a Marie Curie Intra-European Fellowship.

LITERATURE CITED

- Baras E, Maxi MYJ, Nda M, Melard C (2000) Sibling cannibalism in dorada under experimental conditions. II. Effect of initial size heterogeneity, diet and light regime on early cannibalism. *J Fish Biol* 57:1021–1036
- Baras E, Kestemont P, Melard C (2003) Effect of stocking density on the dynamics of cannibalism in sibling larvae of *Perca fluviatilis* under controlled conditions. *Aquaculture* 219:241–255
- Bochdansky AB, Deibel D (2001) Consequences of model specification for the determination of gut evacuation rates: redefining the linear model. *Can J Fish Aquat Sci* 58: 1032–1042
- Christie D (2004) Resampling with Excel. *Teach Stat* 26:9–14
- Clemmesen C, Buhler V, Carvalho G, Case R and others (2003) Variability in condition and growth of Atlantic cod larvae and juveniles reared in mesocosms: environmental and maternal effects. *J Fish Biol* 62:706–723
- Folkvord A, Otterå H (1993) Effects of initial size distribution, day length, and feeding frequency on growth, survival, and cannibalism in juvenile Atlantic cod (*Gadus morhua* L.). *Aquaculture* 114:243–260
- Folkvord A, Ystanes L, Johannessen A, Moksness E (1996) RNA:DNA ratios and growth of herring (*Clupea harengus*) larvae reared in mesocosms. *Mar Biol* 126:591–602
- Imsland AK, Jenssen MD, Jonassen TM, Stefansson SO (2009) Best among unequal? Effect of size grading and different social environments on the growth performance of juvenile Atlantic halibut. *Aquacult Int* 17:217–227
- Kestemont P, Jourdan S, Houbart M, Melard C and others (2003) Size heterogeneity, cannibalism and competition in cultured predatory fish larvae: biotic and abiotic influences. *Aquaculture* 227:333–356
- Kunz YW (2004) Developmental biology of teleost fishes. Springer, Dordrecht
- Ling EN (2007) Efficient analysis of growth trial data. *Aquacult Res* 38:728–732
- Ling EN, Cotter D (2003) Statistical power in comparative aquaculture studies. *Aquaculture* 224:159–168
- Maindonald J, Braun J (2003) Data analysis and graphics using R: an example-based approach. Cambridge University Press, New York, NY
- Moran D (2007) Size heterogeneity, growth potential and aggression in juvenile yellowtail kingfish (*Seriola lalandi Valenciennes*). *Aquacult Res* 38:1254–1264
- Moran D, Smith CK, Gara BG, Poortenaar CW (2007) Reproductive behaviour and egg development in yellowtail kingfish (*Seriola lalandi Valenciennes*, 1833). *Aquaculture* 262:95–104
- Moran D, Smith CK, Lee PS, Pether SJ (2011) Mortality structures population size characteristics of juvenile yellowtail kingfish *Seriola lalandi* reared at different densities. *Aquat Biol* 11:229–238
- Sakakura Y, Tsukamoto K (1998) Effects of density, starvation and size difference on aggressive behaviour in juvenile yellowtails (*Seriola quinqueradiata*). *J Appl Ichthyology* 14:9–13
- Shih WJ, Binkowitz B (1987) Median versus geometric mean for lognormal samples. *J Statist Comput Simulation* 28: 81–83
- Smith ME, Fuiman LA (2003) Causes of growth depensation in red drum, *Sciaenops ocellatus*, larvae. *Environ Biol Fishes* 66:49–60

Editorial responsibility: Roderick Finn,
Bergen, Norway

*Submitted: October 13, 2010; Accepted: December 22, 2010
Proofs received from author(s): January 19, 2011*