Catch-per-unit-effort: which estimator is best?

Petrere Jr., M.[a]*, *Giacomini, HC.*[a] *and De Marco Jr., P.*[b]

[a]Departamento de Ecologia, Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP,
Campus de Rio Claro, Av. 24-A, 1515, CP 199, CEP 13506-900, Rio Claro, SP, Brazil

[b]Laboratório de Ecologia Teórica e Síntese, Departamento de Biologia Geral, Universidade Federal de Goiás – UFG,
Rodovia Goiânia-Nerópolis, Km 5, Campus II, Setor Itatiaia, CP 131, CEP 74001-970, Goiânia, GO, Brazil

*e-mail: mpetrere@rc.unesp.br

(With 6 figures)

Abstract

In this paper we examine the accuracy and precision of three indices of catch-per-unit-effort (CPUE). We carried out simulations, generating catch data according to six probability distributions (normal, Poisson, lognormal, gamma, delta and negative binomial), three variance structures (constant, proportional to effort and proportional to the squared effort) and their magnitudes (tail weight). The Jackknife approach of the index is recommended, whenever catch is proportional to effort or even under small deviations from proportionality assumption, when a ratio estimator is to be applied and little is known about the underlying behaviour of variables, as is the case for most fishery studies.

Keywords: CPUE, probability distributions, accuracy, Jackknife, Monte Carlo.

Captura por unidade de esforço: qual estimador é melhor?

Resumo

Neste trabalho, examinamos a acurácia e precisão de três índices de captura por unidade de esforço (CPUE). Foram feitas simulações, nas quais foram gerados dados de captura de acordo com seis distribuições de probabilidade (normal, Poisson, lognormal, gama, delta e binomial negativa), três estruturas de variância (constante, proporcional ao esforço e proporcional ao quadrado do esforço), e magnitudes (*tail weight*). É recomendado o uso do método Jackknife para os índices, sempre que a captura for proporcional ao esforço ou até em casos de pequenos desvios do pressuposto de proporcionalidade, quando se deseja utilizar um estimador de razão e pouco é conhecido sobre o real comportamento das variáveis, como é o caso da maioria dos estudos de pesca.

Palavras-chave: CPUE, distribuições de probabilidade, acurácia, Jackknife, Monte Carlo.

1. Introduction

In many instances catch-per-unit-effort (cpue) is taken as an estimate of stock size. The cpue is especially useful if the relationship between catch and effort is linear through the origin (strict proportionality) (Gulland. 1954; 1964; Garrod, 1964; Ricker, 1975; Lima et al., 2000). Richards and Schnute (1986) examined the question, both experimentally and theoretically concluding that strict proportionality is best for single-species data collected over uniform weather conditions and that cpue has less value as an index of abundance when data are combined irrespective of species and fishing conditions, as is the case of commercial fisheries. In these last cases, it is worth mentioning that the cpue may also be affected by differences in catchability among samples from different fishing vessels, gear and methods, which might confound indexes of abundance. Even so, when one needs to report catch data it is usual to express cpue totalled by fishing site (quadrats in the sea) and/or sea-

son. This procedure, besides reducing the huge amount of data, as is usual in fisheries, decreases the noise resultant from environmental stochasticity (Petrere, 1986). In practice, there are instances in which these assumptions do not hold, and one of them is that it is possible that non-independence of sampling units generates distinct error structures other than the normal one.

Traditionally there are two ways to define cpue (Equations 1 and 2):

$$cpue_1 = \frac{\sum \frac{C_i}{f_i}}{n} = \overline{\left(\frac{C}{f}\right)} \qquad (1)$$

$$cpue_2 = \frac{\sum C_i}{\sum f_i} = \frac{\overline{C}}{\overline{f}} \qquad (2)$$

where C_i is i[th] catch (usually expressed in weight), f_i is its respective fishing effort. They are both ratio estima-

Braz. J. Biol., 2010, vol. 70, no. 3, p. 483-491

483

tors (Thompson, 1992). Griffiths (1960) named $cpue_1$ as weighted index of density and $cpue_2$ as unweighted index of density reasoning that the ratio $cpue_2/cpue_1$, which he called index of concentration would exceed unity in areas of higher-than-average density. If fishing is at random the index of concentration shall be 1, and if fishing effort is applied in those areas with less than density average, the ratio should be less than 1 (Griffiths, 1960; Calkins, 1961). Another ratio estimator, that does not have much intuitive appeal to fish biologists, is presented by Snedecor and Cochran (1967) (Equation 3):

$$cpue_3 = \frac{\sum C_i.f_i}{\sum f_i^2} = \frac{\overline{C.f}}{\overline{f^2}} \qquad (3)$$

Whenever C be proportional to f, the regression line between them statistically goes through the origin, and can be fitted by the simple model $C_i = \beta f_i + \varepsilon_i$. All the three equations are unbiased estimates of the population ratio β in normally distributed populations. The choice among the three is a matter of precision: the most precise among the three is Equation 3, 2 and 1 if the variance of ε is constant, proportional to f or to f^2, respectively. If the variance of ε increases moderately with f, $cpue_2$ is still expected to perform well. Smith (1980) examined these three estimators, using real data sets assuming that $C_i = \beta f_i + \varepsilon_i$ is the right model. When it is not the case, as frequently happens with real data set, he claims that the best strategy is to jackknife $cpue_2$. In case of inadequate choice of cpue, mainly if there is no strict proportionality between C and f, the three indices above are biased estimates of abundance due to lack-of-fit components introduced when ignoring the intercept. This certainly would lead to biased stock size and poor management.

Even considering the importance of Smith's conclusions, it would be helpful to further investigate how these estimators behave in a wider range of situations, mainly when the error term ε_i does not conform to a normal probability distribution, and also how robust they are under a wide variety of intercept values. Computably intensive Monte Carlo simulations, with a sort of heuristically chosen error structure, seem to be good candidates for this purpose. So, the main objective of this paper is to compare the accuracy, precision and robustness of five distinct cpue indices (three MLE and two Jackknife estimators) under six distinct error distributions, three variance structures, three variance magnitudes, and five intercept values with simulated and real data.

2. Materials and Methods

All simulations reported here are based on simple assumptions. We generated effort data from a uniform distribution from 1 to 10 units. Catch data were simulated in a two-step procedure: (i) we assumed that catch is linearly related to the effort with slope 2 and a given intercept value (0, 0.25, 0.50, 0.75 or 1). It would be interpreted that the expected catch-per-unit-effort must be 2 (unit of biomass or abundance)/(unit of fishing effort).

(ii) We added a random error component to the previous catch and effort data. We examined six random error structures, namely: normal, Poisson, lognormal, delta (Syrjala, 2000), gamma (Ávila-da-Silva, 2002) and negative binomial (Andrade and Teixeira, 2002). So, these common errors structures were chosen just on a heuristic basis and from examples from the literature. Some functions are parameterised according to Mood et al. (1974). In order to produce comparable error structures among the simulations, we chose convenient parameters for these distributions to generate random numbers with mean zero and variances that could or not change with effort, according to Equation 4:

$$V = \left(\frac{a}{3^b}\right) \cdot f^b \qquad (4)$$

where V is the variance of the error term for a given value of effort (f), "a" is a constant and "b" is an exponent dictating how the variance of the error term is related to effort. When b is zero, the variance is constant, irrespective of f values. When b is 1, the variance is proportional to f, and when it is 2, the variance is proportional to the square of f. The term 3^b was just used to scale the variances so that at f = 3 they attain the same value (= a), whatever the value of b. Consequently, for f < 3 the variances generated by "b" = 1 or 2 attain values lower than those with b = 0, increasing to higher values when f > 3. The constant "a" scales the general magnitude of the variances. Here we used three values for "a": 1, 2 and 4. For a given "b", doubling "a" means doubling the variance of the error term over the entire effort range. So, we have a total of 54 different situations, which are combinations of 6 error structures, 3 values of "b" and 3 values of "a". For each combination, we simulated 1000 data sets, each representing 30 landings (30 catch-effort coordinates, from which the cpue estimates were calculated).

To generate random numbers with normal distribution, we simply used the parameters $\mu = 0$ and $\sigma^2 = V$ (which depends on the specific effort value). For lognormal, we set $\mu = 1$ and used the equation for the variance (V) of this distribution (Mood et al., 1974) to solve for the parameter σ (the parameters μ and σ are those related to the normal distribution from which the correspondent lognormal is originated). Each combination of parameters leads to different expected means which, in the case of lognormal distribution, are necessarily higher than zero. We used the equation from the Appendix, Table 2, page 540 in Mood et al., (1974) to calculate these means and subtracted them from the original random numbers, ensuring that the final means of errors be zero. The same procedure used for lognormal was applied also to gamma, Poisson and negative binomial. For the first two, the mean and variance were equal to V, and from the generated numbers we subtracted this same value. The negative binomial was simulated using r = 81, and p varying according to the equation for the variance of this distribution (using again V as the variance for a given effort and conditions set by Equation 4). The delta distribution

484

Braz. J. Biol., 2010, vol. 70, no. 3, p. 483-491

is currently used in fisheries to analyse catch data, as in the case of trawl surveys, where zero catches are fairly common (Syrjala, 2000). It is characterised by having a positive probability of zero values, and a lognormal probability distribution for positive values (Smith, 1988). So, it has three parameters: μ and σ, from the constituent lognormal distribution, and the parameter delta (δ), which is the probability of zero occurrences. To generate random numbers with delta distribution, we first employed uniformly distributed pseudo random numbers, together with lognormal distributed numbers. Every lognormal number associated with a uniform number less than the parameter δ was then turned to be zero. In this way, on average, a proportion equal to δ became zero and a proportion equal to 1- δ remained positive lognormally distributed. In the simulations, the value of δ was fixed as 0.1. The parameters of lognormal distribution were chosen in such a way to give the desired means and variances for the overall delta distribution (the variances obeying Equation 4, depending on the case). We just rearranged the equations for the mean and variance of delta distribution, presented in Smith (1988), to find the correspondent μ and σ. As μ and σ must be positive values, there are some forbidden combinations for the mean (K) and variance (V) values. For the cases in which the variance of errors was constant (b = 0), we used K = 1.5 and V = a = 1, 2 or 4. Later on, we subtracted 1.5 from the generated delta numbers, ensuring that the errors have mean zero. For the cases in which the variances were proportional to f or f^2, we also had to change the K values according to f. This procedure was necessary to make the parameters conform within the constraints. So, when the variances were proportional to f, we made K increasing linearly from 1 (for f = 1, the minimum value) to 3 (for f = 10, the maximum value), and V varying according to Equation 4. Differently, when the variances were proportional to f^2, K increased linearly from 2 to 3. In all cases, the calculated means were subtracted from the generated numbers, always resulting in errors with final mean equals to zero.

For each data set consisting of 30 catch-effort coordinates, we calculated a 95% confidence interval (CI) for cpue based on t-distribution. The variances for each estimator were calculated as in Snedecor and Cochran (1967). In addition, we applied the Jackknife procedure for $cpue_2$ and $cpue_3$, considering each of 30 simulated landings as a replicate according to the current literature in the subject (Efron, 1982; Manly, 1991). Because $cpue_1$ is a mean of ratios, the Jackknife was not used for this estimator (a Jackknifed mean is exactly the mean, so nothing has to be gained). Here the performance of the estimators was analysed by means of three indicators: i) Confidence Interval (CI) coverage (%) and ii) Relative deviation (%) from the real value, both being measures of accuracy and iii) Coefficient of Variation (%) of simulated estimates, this last being a measure of precision. The CI coverage is calculated as the percent of cases (comprising 1000 simulations) in which the actual cpue

(2) is located within the estimated confidence interval. In all simulations, the CI were based on the t-distribution, irrespective of the actual probability distribution. Although not strictly correct, this was done for practical purposes. We think that if someone is interested in calculating a confidence interval for a real data set will do this using the t-distribution, as in most situations little is known about the actual probability distribution which characterises the observed catch variability. The relative deviation is calculated as the module of the difference between the estimate and the real value, divided by this last and multiplied by 100, giving the difference as a percentage, which is easier to interpret and to compare. To test for robustness against violations of strict proportionality assumption, we simulated data with different intercept values: 0, 0.25, 0.50, 0.75 and 1.00. For each intercept, we repeated all procedures described above. We used only positive values for intercepts, as it is expect that negative intercepts would only promote deviations with symmetric effects when compared to positive ones.

2.1. Real data analysis

Real data are from the pink-shrimp "camarão rosa" *Farfantepenaeus brasiliensis* (Latreille, 1817) and *Farfantepenaeus paulensis* (Pérez-Farfante, 1967) landed at Santos/Guarujá (SP) in May 1997. The sample size is 34 landings and the effort is the number of hauls. We adapted the collector curve (Pielou, 1977), in order to determine the effect of increasing sample size upon the five estimators. We generated sets of samples of increasing sizes, 3, 4..., until the maximum of 34 of a sole sample. For each sample size, we took just 50 random combinations (each one sampled without replacement from the original data points), except for the last. For each random combination, we calculated the values for each of the five cpue estimators. We based the analysis only on the relative deviations of them (the absolute differences from the value estimated for the total sample size, divided by this same value). This last was calculated assuming that, for each cpue estimator, the real value was that given by the maximum sample size (or, in other words, the estimate from the original data set).

3. Results

The ordering among estimator performance was most affected by the variance structure (whether constant, proportional to effort or to squared effort). Considering the CI coverage (Figure 1), there is no clear pattern emerging from the graphs. The ordering shows frequent reversals, being affected also by the probability distribution. $Cpue_2$ had higher coverage in all cases for constant variances, although not necessarily closer to 95%, and reasonable coverage for other cases depending on the error distribution. The coverage for $cpue_1$ were generally better when the variances were proportional to the squared effort, and also for some cases when variances were proportional to effort. MLE $cpue_3$ had the worst coverage in almost all
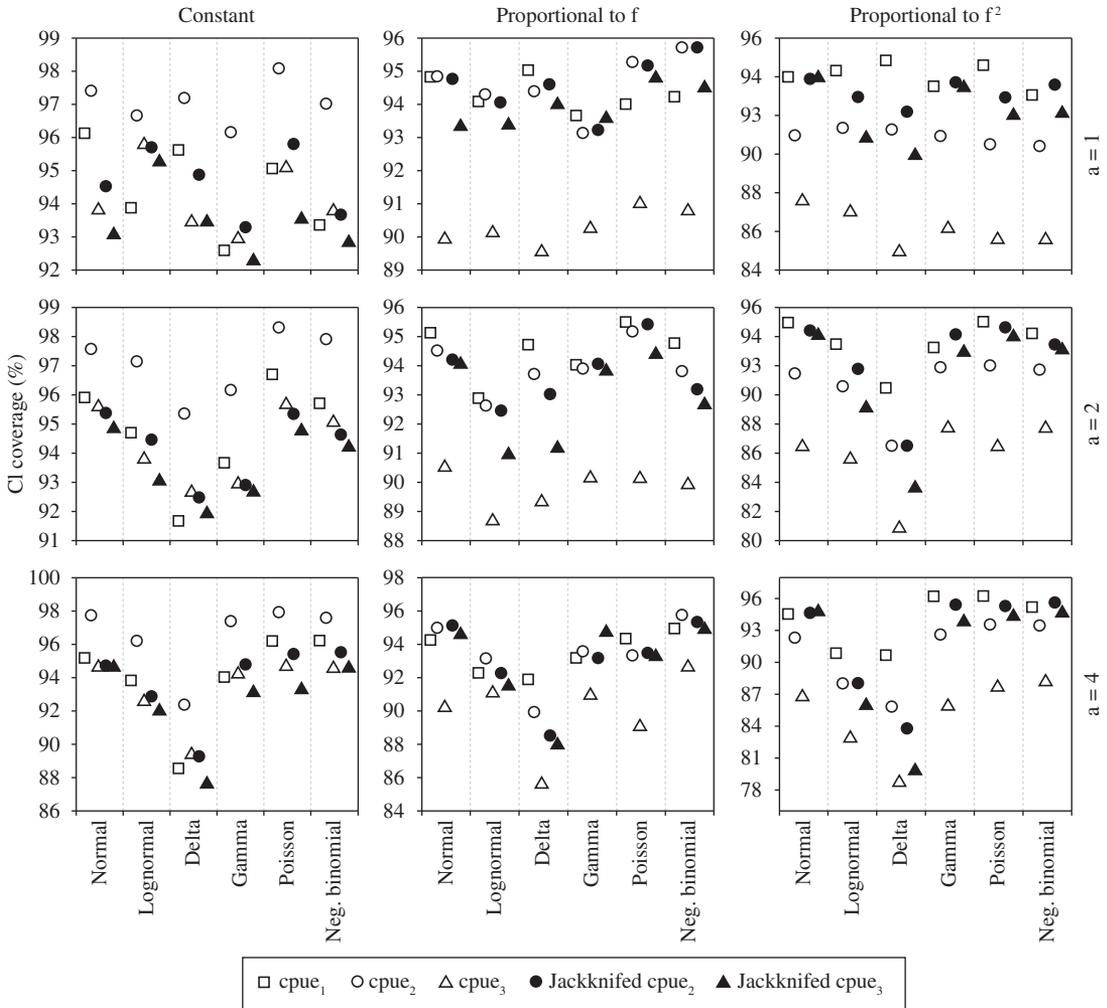
Figure 1. Confidence Interval coverage (ordinate) for the five estimators and six error structures (abscissa). The columns represent different ways in which the error variances changed with effort (f). The lines represent different variance magnitudes, given by the coefficient "a" (Equation 4). Each point gives the proportion of cases (comprising 1000 simulations) in which the true cpue (2) was contained inside the estimated confidence interval.

cases. Jackknifing improved $cpue_2$ and $cpue_3$ coverage when $V \sim f^2$ for the first, and when $V \sim f$ or $V \sim f^2$ for the second.

The most discrepant probability distribution was delta, followed by lognormal. These two distributions depart from the others by producing narrower confidence intervals, but having lower coverage (Figure 1), what is due to their higher asymmetry. These differences were larger the larger the variances. Moreover, this was the only substantial effect of changing the variance magnitudes (the values of the coefficient "a", in Equation 4).

The relative deviations and the coefficients of variation give more clear and similar patterns (Figure 2 and 3). The ordering depended substantially on the variance structure: $cpue_3$ have the best performances when variance is constant; $cpue_2$ and Jackknifed $cpue_3$, when the variance is proportional to effort; $cpue_1$ and Jackknifed $cpue_2$, when variance is proportional to squared effort.

Among them, $cpue_2$ and Jackknifed $cpue_3$ were those which present more regularity, being the least affected by variance structure, and presenting an overall better performance. The probability distribution exerted little or no influence.

To avoid confusion, due to the large amount of data, the results from simulations with different intercepts were grouped and presented on a single graph for each performance indicator (Figure 4). Those graphs contain the means values, maxima and minima, along all kinds of variance structure, magnitude and probability distribution tested. For example, all points present in the graph of Figure 1 (54 values of CI coverage for each cpue estimator) are condensed on a single sequence of mean, maxima and minima for the five estimators in the top graph of Figure 4 (in the region of intercept = 0.00). The same were done for the other intercept values, and repeated for relative deviations and coefficients of vari-
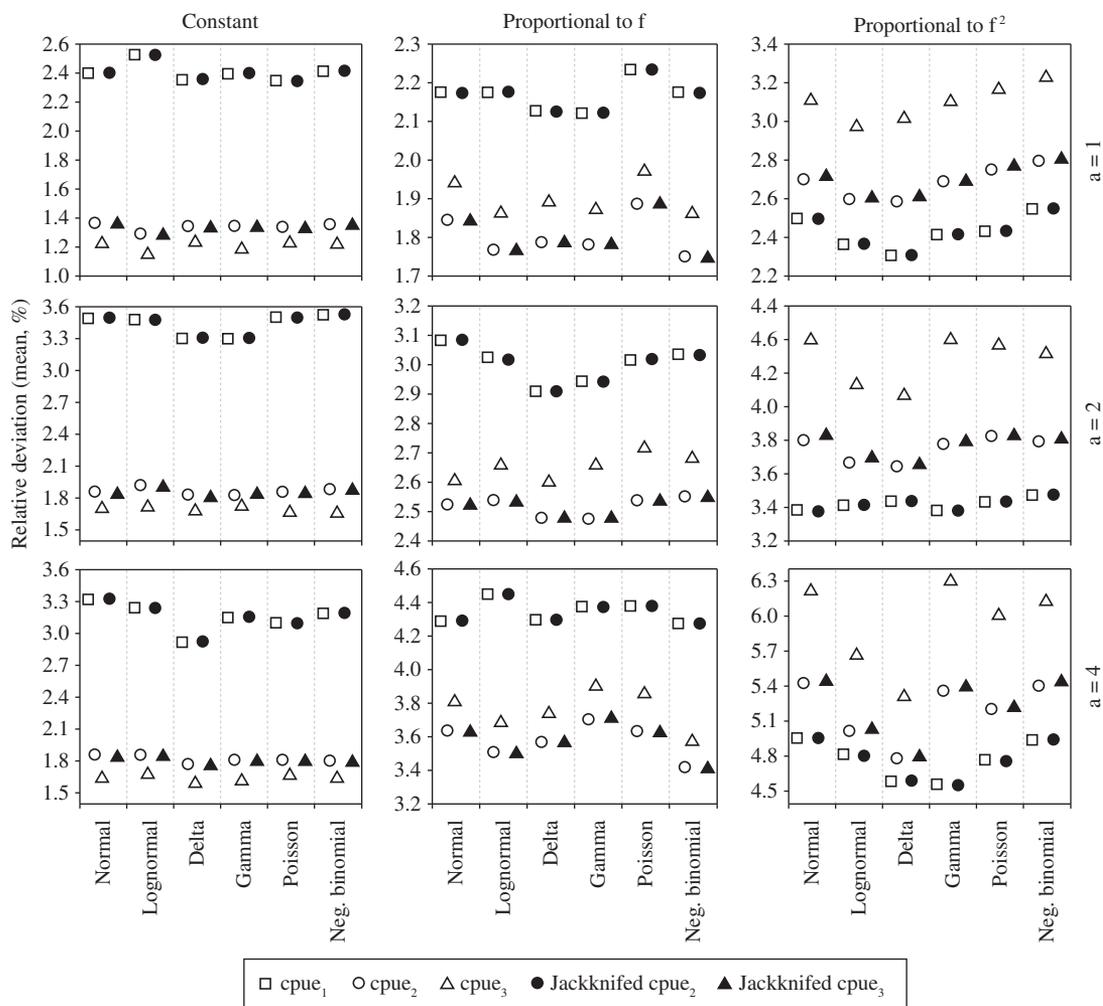
Figure 2. Means of relative deviations (absolute differences of the estimates from the real value, divided by this value) (ordinate) for the five estimators and six error structures (abscissa). The columns represent different ways in which the error variances changed with effort (f). The lines represent different variance magnitudes, given by the coefficient "a" (Equation 4). Each point gives a mean calculated from 1000 simulations.

ation (middle and bottom graph, respectively). This procedure has the disadvantage of ignoring variations due to factors like variance structure or probability distribution, but has the great advantage of showing a clear picture of broader patterns, when they exist.

For intercept = 0, we see that $cpue_1$ and $cpue_2$ have slightly better CI coverage (Figure 4), regarding the entire variety of situations shown in Figure 1. $Cpue_2$ and Jackknifed $cpue_3$ have better performances when considering relative deviations and coefficients of variation. Nevertheless, the differences among estimators for this single situation are rather subtle. When the analysis is extended to a wider range of intercept values, the very subtleness and the context dependence of the estimator differences vanish. Considering the CI coverage, the Jackknifed estimators present an overwhelmingly better performance, mainly for intermediate intercept values (Figure 4). Jackknifing is able to correct almost

completely a bias of up to 0.25 in the value of intercept. Obviously, this capacity decreases as the amount of bias increases, and it is expected that CI coverage would eventually reaches zero. Along intercept values, Jackknifed $cpue_3$ is the one that performs better, even when compared with Jackknifed $cpue_2$.

The relative deviations reinforce Jackknifed $cpue_3$ as the best estimator, followed by MLE $cpue_3$ for this performance index (Figure 4). Contrary to CI coverage, the differences among estimators increase monotonically as intercept values increase. Another still different picture arises from the coefficients of variation (Figure 4). In this case, the intercept does not seem to affect the estimates, which show little variability. MLE $cpue_2$, followed by Jackknifed $cpue_2$, appears as the most precise, although its differences to MLE and Jackknifed $cpue_3$ are negligible.
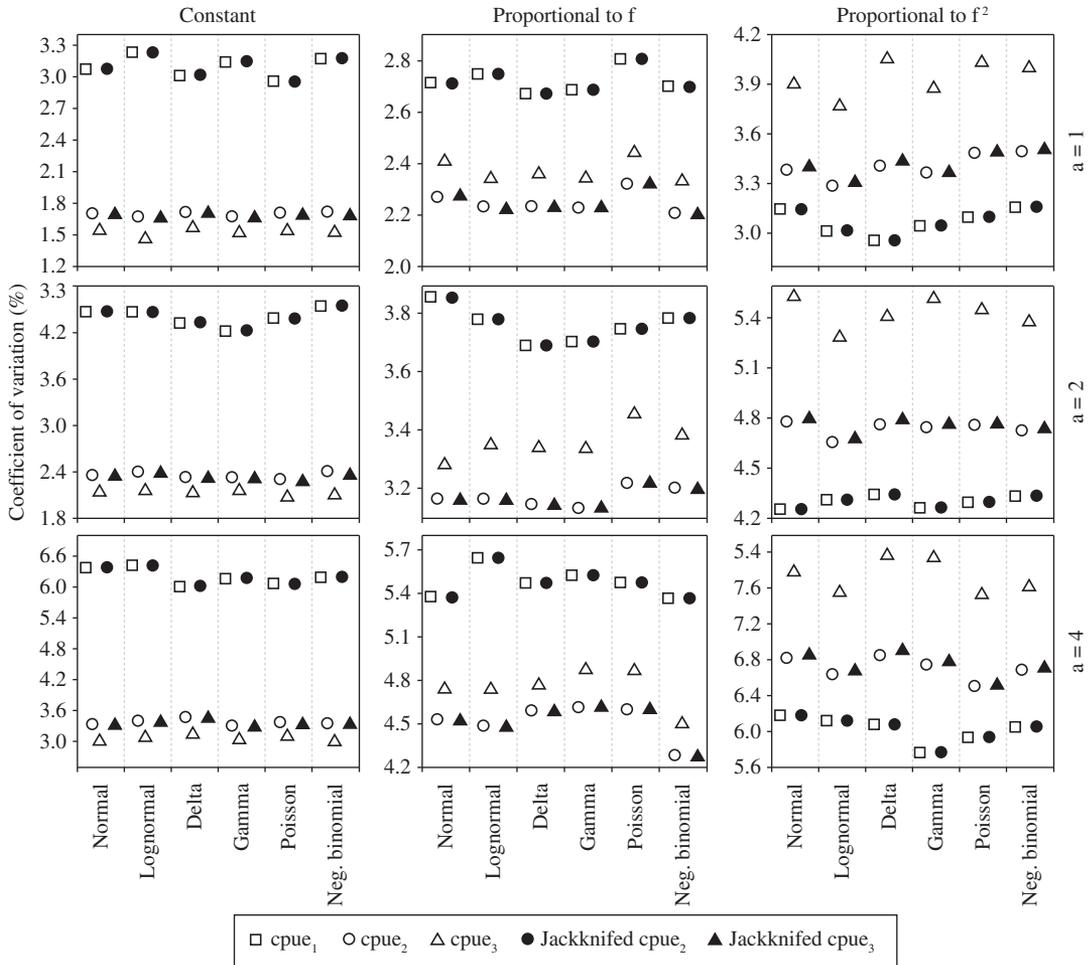
Figure 3. Coefficient of variation (ordinate) of estimates for the five estimators and six error structures (abscissa). The columns represent different ways in which the error variances changed with effort (f). The lines represent different variance magnitudes, given by the coefficient "a" (Equation 4). Each point gives the CV calculated from 1000 simulations.

In the real data of Figure 5 there is a statistically significant relationship between catch and effort (C = 15.898f + 62.764; r² = 0.573; $F_{1, 32}$ = 42.985; p < 0.001). As the intercept is not significantly different from zero (t_{32} = 0.872; p = 0.390), we could assume that the regression pass by the origin, and doing this, we recalculated the relationship (C = 17.8.f; r² = 0.563; $F_{1, 33}$ = 337, p < 0.001).

The cpue estimates (kg/haul) calculated from the shrimp data are: $cpue_1$ = 19.12; $cpue_2$ = 18.21; $cpue_3$ = 17.84; Jackknifed $cpue_2$ = 18.19; Jackknifed $cpue_3$ = 17.83. A quick inspection of Figure 6 reveals that the best estimators for the shrimp data seems to be the Jackknifed and MLE of $cpue_3$, as their relative deviations are lower for almost the entire sample size spectrum (nevertheless, $cpue_2$ is better for the smallest sample size). As expected, the pattern for all estimators is of decreasing relative deviations with increasing sample sizes (as the estimates approach the true values).

4. Discussion

The cpue is a result of several assumptions which may be taken in different combinations: (i) the fish vulnerability to the fishing gear; (ii) stock spatial distribution which determines fishing effort distribution (Paloheimo and Dickie, 1964), *viz* if the stock is randomly spaced this will lead to fishing effort randomly distributed, otherwise we will have a non-random search by the fisher, who behaves like an optimal predator; (iii) independence of each fishing operation; if not, the fishing process would be a Markovian one, as the fisher tends to repeat his operation in those spots with higher abundance; (iv) the proportionality between cpue and abundance N at time t is conventionally expressed as $cpue_t = qN_t$, where q is the catchability coefficient. Lately, a generalisation of this relationship has been examined, the simplest being the power curve $cpue_t = qNt^\beta$ (Harley et al., 2001). Historically catch-per-unit-effort has been suspicious of not always
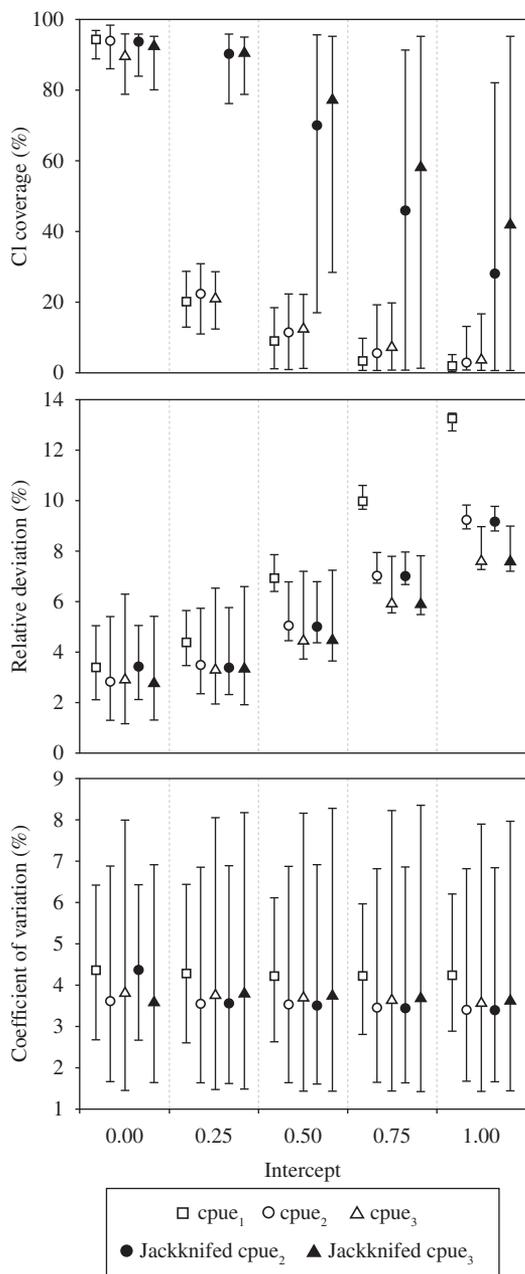
Figure 4. Confidence Interval coverage, Relative Deviation means and Coefficient of Variation along five values of intercepts. The midpoints are the means, and the whiskers represent maxima and minima. For each intercept, data were grouped irrespective of the variance pattern, variance magnitude or error structure, and the mean, maximum and minimum were drawn from 54 values (3 variance patterns X 3 variance magnitudes X 6 probability distributions, as they appear in Figure 1, 2 or 3), for each estimator.

being proportional to abundance and recent studies have been set in order to examine it in depth (Punt et al., 2000; Bigelow et al., 2002; Walters, 2003). Harley et al. (2001) analized the proportionality between cpue and abundance and found real situations where cpue was most likely to
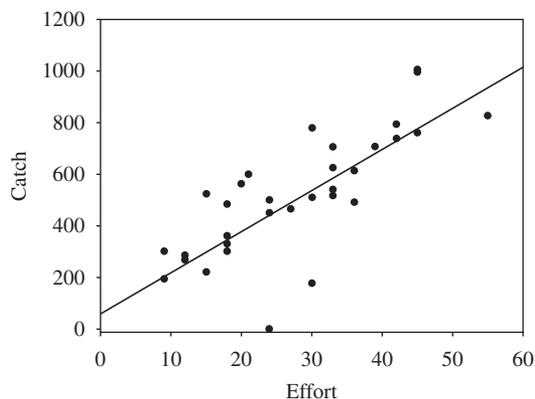
Figure 5. Relationship between the prawn "camarão rosa" *Farfantepenaeus brasiliensis* and *F. paulensis* catches (kg) and fishing effort (number of hauls), C = 17.835f (r² = 0.954; p < 0.001), landed at Santos/Guarujá in May, 1997.
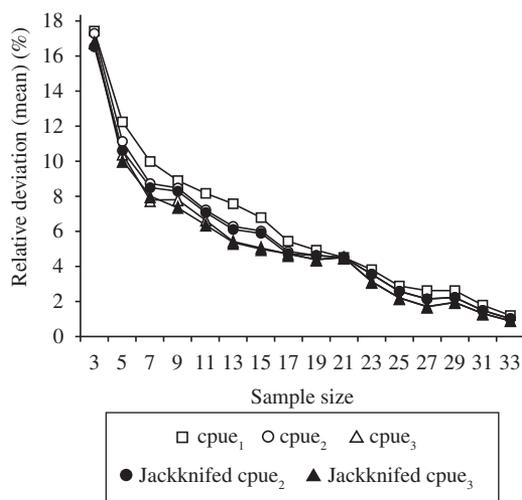
Figure 6. Means of relative deviations for the data of "camarão rosa" *Farfantepenaeus brasiliensis* and *F. paulensis* in May, 1997. In the five methods the mean is a result from 50 randomisations with sample size increasing from 3 to 34 landings (here are presented the data only for odd sample sizes).

remain high while abundance declines ($\beta < 1$ in the power curve $cpue_t = qNt^\beta$, a situation named hyperstabilty). Their simulations also showed that estimates of β have an approximately 10% positive bias. It seems that hyper stability is a common phenomenon in fisheries, a consequence of non-random search (Labelle et al., 1997). Anyway even if $\beta = 1$, it is not a guarantee that c/f is a good abundance index. We must be concerned also with the catchability coefficient q. In practice, for the cpue to be used properly as an index of abundance, the catchability must be reasonably constant among samples, not coming from very different fishing fleets and/or gears. This is a general assumption, on which the conclusions from the present models and simulations also may depend.

When we do not have strict proportionality, a possible way out is try to use an analysis of covariance

(ANCOVA) strategy, with catch as the response variable, the fishing effort as the covariate together with eventual factors as vessel category, season, area, etc. (Allen and Punsly, 1984; Petrere, 1986). This analysis is not biased by the intercept, as it is included explicitly for model estimation. Nevertheless, traditionally people like to express units as ratios, like cpue. It has the appeal of being simpler, more intuitive and does not need the amount of data usually required to perform an ANCOVA. Beyond this, much of the already published work on fisheries and similar research present data as ratios between catch and effort. Any comparison among them and current work would have to adopt the same method.

In simulation experiments, as the ones performed here, we deal with a twin problem: (i) determine the expected behaviour of an estimator (a "population" problem) and (ii) evaluate its behaviour in each performance of the experiment (a "sample" problem). It is important to realise that from the point of view of the field scientist, the behaviour of each sample is more relevant than the expectation of the whole simulation data, which justifies the approach here adopted.

Assuming strict proportionality, our results give evidence that the better choice among the five cpue estimators depends much on the way the catch variance is related to effort and also on the measure of performance, whether the CI coverage, relative deviation, or coefficient of variation. Considering the overall picture for strict proportionality, $cpue_2$ seems to perform better, having relatively high accuracy (high CI coverage and low relative deviations), and precision (low coefficient of variation). Although Calkins (1961) indicated that $cpue_2$ fluctuates more than $cpue_1$ with higher coefficient of variation for the same data set, when reanalysing his data in his original Table 1, catch do not bear any relationship with effort ($r^2 = 0.094$, n = 36, p = 0.068) and in his Table 7, the C × f is not of strict proportionality (C = 10541.58 + 3.328f, s_a = 3278.61, t_a = 4.309**, n = 36, r^2 = 0.353, p < 0.001) making his statements suspicious. Moreover, his conclusion is not supported by our simulations even for nonzero intercepts (Figure 4), but results analogous to his could arise if the variance of data is proportional to the squared effort (Figure 3). The prawn stock analysis showed a better performance of $cpue_3$ (MLE and Jackknifed). Probably this was due to the catch variance, which seems to be constant, or even decreasing with effort (Figure 5). This is consistent with simulation results, as can be seen in Figure 2.

Restricting the analysis only to cases when catch is proportional to effort is, however, of little reaching and application. A zero intercept is a particular case which, although desirable, does not correspond to every real situation in fisheries sampling (Punt et al., 2000; Bigelow et al., 2002). So the robustness of an estimator is of crucial importance. By considering the CI coverage, we see that Jackknifed $cpue_3$ is the least affected by violations of proportionality assumption (Figure 4). This estimator is then the most recommended when one is interested in hypothesis testing, as

it generally makes use of confidence intervals. Estimators that give CI coverage far from the expected value (95%) are prompt to reject true hypothesis or to accept false ones, thus being undesirable. On the other hand, when the estimate itself is of greater importance, the relative deviation is a more adequate index (inverse) of estimator performance. Even in this case, the Jackknifed $cpue_3$ remains as the most accurate estimator. It also has a reasonable precision, not differing much from the other estimators with respect to its coefficient of variation.

The simulations performed here included a wide variety of situations, changing the variance structure, magnitude, and error probability distributions along with the intercept values. By grouping indiscriminately the results from different values of the first three factors above, during the analysis of estimator performance, we assumed that any combination of those factors' values is of equal importance. It implies, for example, that using normal distribution, constant variance with low magnitude has the same weight for the present analysis as using delta distribution, variance proportional to squared effort, high magnitudes, and so on. It could be argued that some combinations of those factors might be more probable to be found in nature. Nevertheless, there is still no comprehensive work that quantifies the frequency distributions of those factors for the overall fisheries data, which is the information needed to ponder the calculation of mean values used to compare the estimators, as presented in Figure 4. Besides this, the only factor that seems to affect the results is the variance structure, what narrows the possibility of mistakes. By covering uniformly a given range of situations, our analysis presumes a practical situation characterised by pure ignorance about underlying variables affecting cpues. Of course, the more quantity and quality of data a researcher has, the greater is the chance of knowing the real behaviour of studied variables (catch, effort or any other). On the other hand, the greater would be the chance of the same researcher not using ratio estimators like cpue, as a lot of alternatives and more sophisticated analysis will become available to his/her data. In the case of less detailed information about the true nature of variables, as occurs with the majority of studies, the cpue still remains popular, and sometimes is the only alternative left. If this is really the case, our conclusions apply adequately, pointing to Jackknifed $cpue_3$ as the most robust and the best ratio estimator of catch-per-unit-effort.

References

ALLEN, R. and PUNSLY, R., 1984. Catch rates as indices of abundance in yellowfin tuna, *Thunnus albacares*, in the Eastern Pacific Ocean. *Inter-American Tropical Tuna Committee Bulletin*, vol. 18, no. 4, p. 301-379.

490

Braz. J. Biol., 2010, vol. 70, no. 3, p. 483-491

ANDRADE, HA. and TEIXEIRA, JA., 2002. A probability model for the catch per unit of effort (cpue) of the skipjack tuna (*Katsuwonus pelamis*) pole and line fishery in the Southwest Atlantic. *Notas Técnicas da FACIMAR*, vol. 6, no. 1, p. 1-5.

ÁVILA da SILVA, AO., 2002. *A evolução da pesca de linha-de-fundo e a dinâmica de população do peixe-batata* Lopholatilus villarii *(Teleostei: Malacanthidae) na margem continental da costa brasileira entre os paralelos 22⁰ e 28⁰ S.* São Paulo: Universidade de São Paulo. 218 p. [Tese de Doutorado].

BIGELOW, KA., HAMPTON J. and MIYABE, N., 2002. Application of habitat-based model to estimate effective longline fishing effort and relative abundance of Pacific bigeye tuna (*Thunnus obesus*). *Fisheries Oceanography*, vol. 11, no. 3, p. 143-155.

CALKINS, TP., 1961. Measures of population density and concentration of fishing effort for yellowfin and skipjack tuna in the eastern tropical Pacific Ocean, 1951-1959. *Inter. American Tropical Tuna Comission Bulletin*, vol. 6, no. 3, p. 71-125.

EFRON, B., 1982. *The jackknife, the bootstrap and other resampling plans*. Philadelphia: Society for Industrial and Applied Mathematics. 128 p.

GARROD, DJ., 1964. Effective fishing effort and the catchability coefficient q. *Rapport et process verbaux des réunions du Conseil International pour l´Exploration de la Mer*, vol. 155, no. 14, p. 66-70.

GRIFFITHS, RC., 1960. A study of measures of population density ando f concentration of fishing effort in the fishery for yellowfin tuna, *Neothunnus macropterus*, in the Eastern tropical Pacific Ocean, from 1951 to 1956. *Inter-American Tropical Tuna Comission Bulletin*, vol. 4, no. 3, p. 41- 96.

GULLAND, JA., 1954. A note on the statistical distribution of trawl catches. *Rapport et process-verbaux des réunions du Conseil International pour l´Exploration de la Mer*, vol. 140, no. 26, p. 28-29

GULLAND, JA., 1964. Catch per unit of effort as a measure of abundance. *Rapport et process- verbaux des réunions du Conseil International pour l´Exploration de la Mer*, vol. 155, no. 1, p. 8- 14.

HARLEY, SJ., MYERS RA. and DUNN, A., 2001. Is catch-per-unit-effort proportional to abundance? *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 58, no. 9, p. 1760-1772.

LABELLE, M., HOCH, T., LIOURZOU, B. and BIGOT, JL., 1997. Indices of bluefin tuna (*Thunnus thynnus thynnus*) abundance derived from sale records of French purse seine catches in the Mediterranean sea. *Aquatic Living Resources*, vol. 10, no. 6, p. 329-342.

LIMA, AC., FREITAS, CEC., ABUABARA, MA., PETRERE M. and BATISTA, VS., 2000. On the standardization of the fishing effort. *Acta Amazonica*, vol. 30, no. 1, p. 167-169.

MANLY, BFJ., 1991. *Randomization and Monte Carlo Methods in Biology*. 1 ed. London: Chapman and Hall.

MOOD, AM., GRAYBILL, FA. and BOES, DC., 1974. *Introduction to the Theory of Statistics*. Singapore: McGraw-Hill.

PALOHEIMO, JE. and DICKIE, LM., 1964. Abundance and fishing success. *Journal du Conseil. International pour l´. Exploration de la Mer,* vol. 155, no. 1, p. 152-163.

PETRERE, M., 1986. Amazon fisheries: I - Variations in the relative abundance of Tambaqui (*Colossoma macropomum* Cuvier, 1818) based on catch and effort data of the gill - net fisheries. *Amazoniana*, vol. 9, no. 4, p. 527-547.

PIELOU, EC., 1977. *Mathematical ecology*. New York: Wiley.

PUNT, AE., WALKER, TI., TAYLOR, BL. and PRIBAC, F., 2000. Standardization of catch and effort data in a spatially-structured shark fishery. *Fisheries Research*, vol. 45, no. 2, p. 129-145.

RICHARDS, LJ. and SCHNUTE, JT., 1986. An experimental and statistical approach to the question: is cpue an index of abundance? *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 43, no. 10, p. 1214-1227.

RICKER, WE., 1975. Computation and interpretation of biological statistics of fish populations. *Buletin of. the Fisheries. Research Board Canada*, vol. 191(único), p. 1-382

SMITH, SJ., 1980. Comparison of two methods of estimating the variance of the estimate of catch per unit effort. *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 37, no. 12, p. 2346-2351.

SMITH, SJ., 1988. Evaluating the efficiency of the Δ-Distribution mean estimator. *Biometrics*, vol. 44, p. 485-493.

SNEDECOR, GW. and COCHRAN, WG., 1967. *Statistical methods*. 6 ed. Ames: The Iwoa State University Press.

SYRJALA, SE., 2000. Critique on the use of the delta distribution for the analysis of trawl survey data. *ICES Journal of Marine Science*, vol. 57, no. 4, p. 831-842.

THOMPSON, SK., 1992. *Sampling*. New York: Wiley.

WALTERS, C., 2003. Folly and fantasy in the analysis of spatial catch rate data. *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 60, no. 9, p. 1433.1436.

Braz. J. Biol., 2010, vol. 70, no. 3, p. 483-491

491