SOME COMMENTS ON PROCEDURES FOR ANALYZING CATCH-AND-EFFORT DATA

Alejandro Anganuzzi

IOTC Secretariat

ABSTRACT MISSING

These comments are intended as a starting point for a discussion of some issues around CPUE analyses. In some cases, they come from the experience of other groups dealing with similar problems. In other cases, they include proposals that could be tested through simulation, and shown to be simply the product of staying up too late at night.

The term 'indices of abundance' is used in a generic way, to include any statistic derived from catch-and-effort data for a particular fishery and that is assumed to hold some relationship (usually linear) to the abundance of a segment of the population. The nominal CPUE obtained by dividing the nominal catches by nominal effort constitutes such an index of abundance, although unsatisfactory for well-known reasons. Usually, the nominal catch-and-effort data are subject to some procedure to remove the dependency of the index on effects not related to abundance. For historical reasons, this procedure is called standardization of effort but, in most cases, the main objective of the procedure is to obtain an index of abundance. The comments assumed that the main objective of the analyses is to obtain such an index, and that the standardization procedure is based on a GLM, ANOVA or GAM models.

SEPARATE OR INTEGRATED STANDARDIZATION?

Indices of abundance are rarely used in isolation in a stock assessment. They are usually the main inputs to (age-aggregated) production model estimating procedures or provide the basic statistical structure in traditional catch-at-age or catch-at-size models. Therefore the first major decision for the analyst is to use the CE data to obtain an index of abundance or to integrate the analysis of CE data directly into the modelling exercise.

Several authors have proposed the direct incorporation of the CE data straight in the modelling and estimate the parameters estimating the effect size jointly with the parameters of the assessment model. Advantages of this procedure include the possibility of including models other than linear into the standardization procedure; an inspection of possible confounding between productivity parameters and standardization effects, and the possibility of carrying forward the covariance structure of the parameters in a more complete way.

The disadvantages are that there are valuable insights to be obtained from an independent data exploratory analysis of the CE data. Especially the identification of factors that could be relevant can be more easily done if the analysis is done outside the modelling exercise.

A possible strategy to test could be to combine both approaches, by carrying out a separate standardization (e.g. a GLM in the usual way) and then use the estimated parameters as initial values for a standardization integrated within the assessment model. In this way, the effect of the additional information brought about in the structure of the production model (as the estimating procedure will modify some of the parameters to improve the consistency with the data) can be better compared with the initial estimates of the trend. This should lead to a better assessment of the contributions of the structural restrictions in the model and the data in the standardized abundance trend. In most assessments, this could be a routine sensitivity run.

SIGNIFICANCE VERSUS INFLUENCE IN CANDIDATE EFFECTS

Usually, in selecting the final structure of any standardizing model with a sufficiently large number of observations, almost any factor entered will be significant. For the same reason, almost any distribution of residuals would depart significantly from the desired normality. Therefore, these results cease to be of practical interest at large sample sizes.

Conversely, if sample sizes are small, the ability to detect important effects will be somewhat limited. The significance levels have to be interpreted in terms of the power of the test. The more observations available, the higher the power of the test and the more likely that significant effects will be detected, even if they are small in magnitude. In that respect, usual model selection procedures would tend to build very large models that don't differ much from the results of more parsimonious models. This is particularly applicable to AIC or BIC criteria for model selection which tend to yield large models.

This suggests that when we consider variables for inclusion into a model we should also look at the influence that the inclusion of a variable has on our perception of the trend in addition to its significance. If sample sizes are large enough, an effect could be highly significant but the overall effect in the trend is negligible. A very simple way at exploring the influence would be to plot the trend resulting from the full model together with those trends obtained after dropping each of the covariates in turn.

DILUTION OF THE ABUNDANCE SIGNAL

Obviously the environment affects the distribution of the resource and its catchability. The problem is that environment could also affect year-class strength, through dependency of survival, growth or both. Therefore, it is again important to consider with care the inclusion of environmental variables in a GLM with time lags. Time lags associated with the period in which environment affects survival or growth can lead to, inadvertently, dilute the information related to abundance. In this case, an environmental variable is not correlated with catchability, but with abundance itself.

The problem of inadvertently removing the abundance signal from the year effects is not restricted to environmental variables. If a covariate or factor in the model has a time trend correlated with the abundance signal, its inclusion might remove the trend in abundance. In an extreme case, we would see a totally flat trend combined with some unusual results, like a decrease in apparent catchability.

TIME-AREA INTERACTIONS

When incorporating time-area interactions, the relative size of the areas affects the probability of getting a significant interaction term. If sample sizes are sufficiently large, the probability of detecting a significant time-area interaction grows with the number of areas considered, depending on the characteristics of the data sets and the homogeneity of the areas. A decision on the number of areas to use in the GLM model could be explored by looking at the effect on the trend as the number of areas changes and looking at the size of the coefficient of the interaction term. Again, there is a trade-off between the gains of incorporating interactions and the complications brought about by their inclusion.

When there are significant time-area interactions we have to collapse the space information through some integration mechanisms. The usual procedure is to use an area-weighted (and sample-size weighted) average index for the year. It is not clear that there is much advantage is partition the space to later collapse it again. Obviously, if the analysis model (VPA or production modelling) has spatial structure then we can input the different indices by area. Weighting by sample size is only a proxy for the proper weighting with should be the inverse of the standard error. As the standard error depends on the square of the number of observations, an appropriate approximation should be weighting by the square root of the number of observations.

In any case, further exploration of the individual trends by area can provide useful insight. For example, they might reflect the abundance of different age groups if fish of different ages segregate spatially. This can be explored using a simple procedure by which we look at the indices per age groups. An inspection of the length-frequency distribution evolution over time can indicate whether there are large differences between year-class strengths. Such differences combined with spatial segregation by size would yield significant time-area interactions.

In general, it seems that there would be a benefit in looking more into developing alternative explanations for the interactions and explicitly mentioned these as working hypotheses. By doing this, we could expose more easily inconsistencies with assumptions arising from other data sources or implicit in the structure of further analyses down the assessment chain.