



Choosing the observational likelihood in state-space stock assessment models

Albertsen, Christoffer Moesgaard; Nielsen, Anders; Thygesen, Uffe Høgsbro

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Albertsen, C. M., Nielsen, A., & Thygesen, U. H. (2015). *Choosing the observational likelihood in state-space stock assessment models*. Abstract from ICES Annual Science Conference 2015, Copenhagen, Denmark.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Choosing the observational likelihood in state-space stock assessment models

Christoffer Moesgaard Albertsen, Anders Nielsen, Uffe Høgsbro Thygesen

National Institute of Aquatic Resources, Technical University of Denmark, Charlottenlund, Denmark.

Presenter contact details: cmoe@aqu.dtu.dk.

Summary

By implementing different observational likelihoods in a state-space age-based stock assessment model, we are able to compare the goodness-of-fit and effects on estimated fishing mortality for different model choices. Model fit is improved by estimating suitable correlations between age-groups. We show by simulations that modelling catch as numbers-at-age is more suitable than proportions if the uncertainty of age-classifications is small.

Introduction

Stock assessment models often rely on aggregated data, such as surveys, or uncertain data, such as landings-at-age which relies on age classification of (effectively) few individuals (Aanes and Pennington 2003). Further, the data can often both be modelled as numbers-at-age or proportions-at-age along with total weight or numbers. This makes it difficult to find the optimal (or true) distribution of the observations a priori. Most currently used age based stock assessment models either considers numbers or proportions at age and only one or few observational likelihoods (ICES 2010). By implementing several observational likelihoods in an age based state-space stock assessment model, we are able to compare the model fit and suitability for each choice of likelihood along with the implications for key outputs such as spawning stock biomass and average fishing mortality. Through a simulation study, we further investigate the difference between modelling numbers-at-age and proportions-at-age.

Materials and Methods

We implement state-space age-based stock assessment models (Nielsen and Berg 2014) with 13 different observational likelihoods. The observational likelihoods included seven distributions for numbers-at-age with no correlation between age groups: Log-normal distribution, gamma distribution, generalized gamma distribution, normal distribution, truncated normal distribution, Student's t-distribution on log-scale and a multivariate log-normal distribution. Further, three distributions for proportions-at-age was included: Additive logistic-normal distribution, multiplicative logistic-normal distribution, and the Dirichlet distribution. All of the three proportions-at-age models were combined with both a log-normal distribution on the total catch in numbers and the total catch in weight. All models were fitted with different parameters for all ages, and the goodness-of-fit was compared by AIC. To compare the AICs, while avoiding excessive hypothesis testing, an upper and lower bound for the minimal AIC was attained. The upper bound was attained by the AIC for the fitted model. The lower bound was calculated as the AIC that would be obtained by the smallest possible model if the negative log-likelihood would not increase compared to the fitted model. Further, estimates of spawning stock biomass, average fishing mortality, and recruitment were compared between the models. Catch and survey data for North Sea cod was used as a case

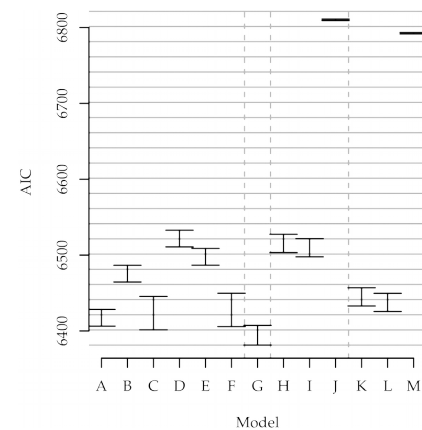


Figure 1: Upper and lower bounds for the minimal attainable AIC in the case study for the models: Log-normal (A), gamma (B), generalized gamma (C), normal (D), truncated normal (E), Student's t (F), multivariate log-normal (G), additive logistic-normal (catch in numbers) (H), multiplicative logistic-normal (numbers) (I), Dirichlet (numbers) (J), Additive logistic-normal (weight) (K), multiplicative logistic-normal (weight) (L), and Dirichlet (weight) (M).

study. The data included catch-at-age from 1963 to 2011 for ages one to seven, and first quarter bottom trawl survey from 1983 to 2012 for ages one to five.

To investigate the difference between considering catch-at-age as numbers and as proportions, we conduct a simulation study. In the simulation study we only consider the multivariate log-normal and the additive logistic normal (with log-normal total catch), since proportions of multivariate log-normal numbers are distributed according to the additive logistic normal distribution. We simulate multivariate log-normal catch, subsample age-proportions from a multinomial distribution and fit the two observational models to the total sample divided by age according to the subsampled proportions.

Results and Discussion

For North Sea cod, the multivariate log-normal achieved the lowest AIC (Fig. 1). The AIC interval for this distribution only barely overlaps with the intervals for the generalized gamma, log-normal, and t-distributions, indicating that the multivariate log-normal fits the data better than the other distributions considered. The AIC intervals for the generalized gamma, log-normal, and t-distributions overlaps with each other, and with the intervals for the logistic normals with log-normal total catch in weight. Hence, the goodness-of-fit of these distributions can not be distinguished without further investigation. However, these distributions perform better than the normal distributions and the Dirichlet distributions. The Dirichlet distribution performs worse than any of the other distributions, confirming earlier results (Francis 2014). This is expected since the Dirichlet is the least flexible model, as it has no free parameters.

The choice of model also influences key outputs such as average fishing mortality (Fig. 2). The estimated average fishing mortality differs as much as 0.1 between the log-normal models and the models for proportions-at-age, although, the difference is not big relative to the uncertainty of the estimates.

Some of the difference in goodness-of-fit between models on proportions- and number-at-age can be attributed to the relative allocation of uncertainty between proportions-at-age estimates and estimates of total catch. The simulation study shows, that if the CV of the observed total is high compared to the variance of the observed proportions, the additive logistic-normal outperforms the multivariate log-normal (Fig. 3).

References

Aanes, S., and Pennington, M. 2003. On estimating the age composition of the commercial catch of Northeast Arctic cod from a sample of clusters. *ICES Journal of Marine Science*, 60: 297–303.

Francis, R. I. C. C. 2014 Replacing the multinomial in stock assessment models: A first step. *Fisheries Research* 151: 70-84

ICES. 2010. Report of the Workshop on Reviews of Recent Advances in Stock Assessment Models Worldwide: "Around the World in AD Models" (WKADSAM), 27 September – 1 October 2010, Nates, France. ICES CM 2010/SSGSUE:10. 122 pp.

Nielsen, A., and Berg, C. W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research* 158: 96–101.

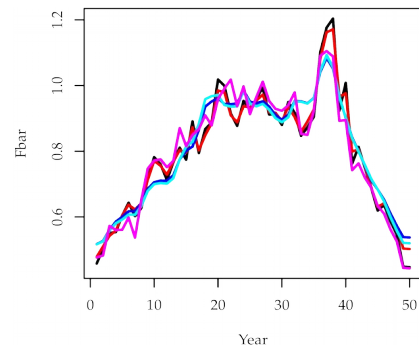


Figure 2: Estimated average fishing mortality (ages 2-4) with the models: log-normal (black), multivariate log-normal (red), Additive logistic normal (weight) (blue), Multiplicative logistic normal (weight) (cyan), and Dirichlet (weight) (magenta).

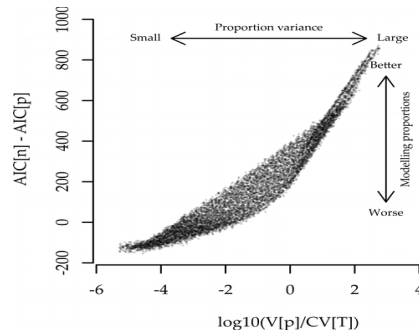


Figure 3: Difference in AIC between the model on numbers and the model on proportions as a function of the log10 difference between the variance of proportions and CV of total numbers. Each dot is the result of a simulation.