



SC14 Meeting Report

Annex 15: Agreed Diagnostics for the Evaluation of Jumbo Flying Squid Stock Assessment Models

A scenario was considered acceptable when no retrospective pattern was detected, more than half of the indices passed each residual diagnostic test, and the median value of the shape parameter occurred lower than Schaefer.

Non-exhaustive list of diagnostics

Retrospective pattern:

- a. Mohn' rho (-0.15,0.2)

Index residual:

- a. Temporal trend test
- b. Normality test
- c. Autocorrelation test

Shape parameter:

- a. The median value should be lower than Schaefer

Testing the predictive ability of the models: Models should be evaluated on their ability to predict the observations that management advice depends on. It is therefore recommended that each candidate stock assessment model be subjected to hindcast cross-validation. This process should involve:

- a. Sequential removal of the terminal years of data (three years)
- b. Refitting the model to each truncated time series.
- c. Projecting the fitted model forward over the withheld years setting any process error to zero. The projected biomass should be multiplied by the estimated catchability to generate a predicted value of each withheld index observation, and this prediction compared directly with the value actually observed.
- d. Calculating the mean absolute scaled error (MASE)