SCRS/2013/117 Collect. Vol. Sci. Pap. ICCAT, 70(5): 2111-2119 (2014)

AN EVALUATION OF DIFFERENT APPROACHES FOR MODELLING UNCERTAINTY IN ASPIC AND BIOMASS DYNAMIC MODELS

Laurence T. Kell[1], Josetxu Ortiz de Urbina[2], Paul De Bruyn[1], Iago Mosqueira[3], Arni Magnusson[4]

SUMMARY

In this paper we compare the estimates of uncertainty obtained for a biomass dynamic stock assessment model. We do this for approaches based on simulation (i.e. the bootstrap, jack knife and Monte Carlo Markov Chain), likelihood profiling and the delta method.

RÉSUMÉ

Le présent document compare les estimations de l'incertitude obtenues pour un modèle d'évaluation des stocks dynamique de la biomasse. Nous réalisons ceci pour des approches basées sur la simulation (c.-à-d. bootstrap, "jack knife" et MCMC-Monte Carlo Markov Chain), profilage des vraisemblances et méthode delta).

RESUMEN

En este documento comparamos las estimaciones de incertidumbre obtenidas para un modelo de evaluación de stock de dinámica de biomasa. Esto se realiza para enfoques basados en simulación (es decir, bootstrap, jack knife y cadena Monte Carlo Markov), perfiles de verosimilitud y el método delta.

KEYWORDS

Albacore, Assessment, Biomass Dynamic, Bootstrap, Bayes, Likelihood Profile, Delta Method, Uncertainty

1. Introduction

A main management objective of ICCAT is to maintain the populations of tuna and tuna-like fishes at levels which will permit the maximum sustainable catch. Scientific advice within ICCAT is therefore based on MSY-based reference points. In common with other tuna Regional Fisheries Management Organisations (tRFMOs) advice is presented showing the probabilities of a stock being greater than B_{MSY} and fishing mortality being less than F_{MSY} in the historical assessment and for different management options projected into the future. To estimate these probabilities a variety of methods have been used by assessment working groups, e.g. bootstrapping and Bayesian simulation. Magnusson *et al*. [2012] compared three methods for estimating uncertainty in aged based stock assessment (i.e. the bootstrap, delta method and Monte Carlo Markov Chain simulation); they showed through simulation that all three methods generated too narrow confidence intervals, underestimating the true uncertainty. In this paper we compare estimates of uncertainty from biomass dynamic models, e.g. ASPIC Prager (1992), and discuss the consequences for advice provided by the SCRS.

2. Material and Methods

There are three main ways to estimate uncertainty within stock assessment models, frequentist, Bayes and likelihood. The frequentist approach treats a parameter as an unknown that poses a true value and estimates confidence intervals rather than probabilities.

[1] ICCAT Secretariat, C/Corazón de María, 8. 28002 Madrid, Spain; Laurie.Kell@iccat.int; Phone: +34 914 165 600
[2] Instituto Español de Oceanografía IEO- CO Málaga, Pto. Pesquero s/n, 29640 Fuengirola (Málaga), Spain; urbina@ma.ieo.es; Phone: +34 952 19 71 24 Fax: +34 952 46 38 08.
[3] European Commission, Joint Research Center, IPSC/Maritime Affairs Unit, Via E. Fermi 2749, 21027 Ispra VA, Italy
[4] Marine Research Institute, Skulagata 4, PO Box 1390, 121 Reykjavik, Iceland

In the likelihood and Bayes approaches parameters are considered to be random variables (Wade, 2000). Bayes methods provide a formal mechanism for updating belief as new data becomes available. Bayesian methods also require a formal description of prior beliefs that reflect current knowledge. Likelihood methods do not need such priors, though as for all methods they are dependent on belief about the appropriate representation of processes and form of error distributions Patterson *et al*. [2001]. Six methods for estimating parameter uncertainty were compared, i.e. the bootstrap, jack knife, Monte Carlo simulation of the input data, Bayesian estimation using Markov Chain Monte Carlo (MCMC) simulation, the delta method and likelihood profiling.

2.1 Data

The data used were time series of the total catch biomass and a single index of abundance as used in the last assessment for North Atlantic swordfish (SCRS/2009/016). The index with the 10th and 90th is plotted in in **Figure 1**.

2.2 Methods

Dynamics were modelled by a biomass dynamic model, where the biomass of a stock next year (B_{t+1}) is equal to the biomass this year B_t less the catch (C_t) plus the surplus production (P_t) where P is given by the Pella-Tomlinson surplus production function.

2.3 Estimation

Estimation of parameters was conducted using the biodyn package based on ADMB and implemented in R. It was assumed that the index was a proxy for stock biomass

2.3.1 Bootstrap

Bootstrapping is a method for providing estimates of a sampling distribution for a statistic. It is routinely used by the SCRS to provide estimates of uncertainty from assessment packages such as ASPIC or VPA2box using the residuals from the original fit.

2.3.2 Jackknife

The jackknife or leave one out procedure is a cross-validation technique to estimate the bias and variance of an estimator. It is similar to bootstrapping but the statistic is recomputed by leaving out one or more observations at a time from the sample set. The new set of replicates of the statistic then allow estimates of the variance and the bias of a statistic to be calculated.

2.3.3 Delta Method

The delta method can provide estimates of the standard deviations for derived quantities in many (but not all) situations. It is based on finding approximations based on Taylor series expansions to the variance of functions of random variables.

As part of the fitting process the Hessian (i.e. the matrix of second-order partial derivatives) is calculated by ADMB. The inverse of the Hessian matrix approximates the variance/covariance matrix of the parameter estimates. The standard errors of derived parameter, i.e. statistics that are not actual parameters in the model but derived from them, can also be estimated.

2.3.4 Likelihood Profile

The delta method can be used to calculate standard deviations for derived quantities in many (but not all) situations. It is based on finding approximations based on Taylor series expansions to the variance of functions of random variables.

2.3.5 MCMC

Markov chain Monte Carlo (MCMC) is a method for simulating a probability distribution for a statistic. It is used to approximate the posterior distribution of estimated parameters. One of the main difficulties with MCMC methods is ensuring that simulations have converged to a stationary distribution. The equilibrium distribution of the chain is the required posterior distribution but how do we know that the chain has reached equilibrium? A burn-in period where initial values are discarded helps. However, in complicated cases, e.g. where there is more than one local maximum or the posterior distribution is in the form of a ridge a chain, it can take a long time to move around the parameter space and a very long burn-in period may be required. While very large sample may have to be taken to ensure that the chain has not just become temporarily stuck in one part of the parameter space.

For these reasons a variety of diagnostics are used to check convergence; i.e. that a stationary distribution has been reached e.g.

- Autocorrelation Plots measure the correlation between fit and fit+1 variable in a chain
- Correlation Plots can show if parameters are confounded
- Gelman-Rubin Diagnostic tests that the burn-in is adequate and requires that multiple starting points be used.
- Geweke Diagnostic , if burn-in is adequate, then the mean of the posterior distribution of from the first half of the chain should equal the mean from the second half of the chain

3. Results

First the results and residual diagnostics from the assessment are presented and then we compare the different estimates of uncertainty for B : B_{MSY} and F : F_{MSY}.

3.1 Assessment

Time series of estimated harvest rate and stock biomass compare and the assumed catch are presented in **Figure 2**, for both the aspic and biodyn R packages. The results were identical.

3.2 Residual Diagnostics

Inspection of residual diagnostics are important to check that the assumptions are met both when fitting a model and conducting simulations. A main assumption of the assessment model used is that the CPUE series is a proportional to stock biomass, therefore we plot the observations against the fitted values in **Figure 3.** If the assumption is correct then the points should be distributed about the y = x line. However, it can be seen that the 95% CI of a linear regression (blue shaded area) fitted to the data do not contain the y = x line and that there appears to be autocorrelation between points.

Figure 4 plots the residuals against year to evaluate whether there is any systematic pattern that may suggest the index is not a proxy for stock biomass. It was also assumed that the residuals are normally distributed and there is no autocorrelation. Therefore the Q-Q plots in **Figures 5** compare a sample of data on the vertical axis to a statistical population on the horizontal axis, in this case a normal distribution. If the points follow a strongly nonlinear pattern this will suggest that the data are not distributed as a standard normal i.e. X N(0; 1). Any systematic departure from a straight line may indicate skewness or over or under dispersion. In **Figure 6** the residuals are plotted against each other with a lag of 1 to identify autocorrelation. There are significant autocorrelations. There appears to be significant positive autocorrelation. It is also assumed that variance does not vary with the mean, i.e. there is no heteroscedasticity, this assumption is evaluated in **Figure 7** where the residuals are plotted against the fitted values.

3.3 Densities

The densities of stock relative to B_{MSY} and of harvest rate relative to F_{MSY} are plotted for each method of estimating parameter uncertainty in **Figures 8 and 9**. The Kobe phase plots, with points showing stock relative to B_{MSY} and harvest rate relative to F_{MSY}, are then shown in **Figure 10** for the various method of estimating parameter uncertainty.

4. Discussion

Estimates of stock relative to B_{MSY} and harvest rate relative to F_{MSY} were compared for 6 different procedures. Estimates of uncertainty obtained from the same data and stock assessment model vary depending on the method used to estimate the uncertainty, i.e. their confidence intervals or probability distributions. The densities obtained for B : B_{MSY} varied, with Monte Carlo simulation of the CPUE Index giving the narrowest estimate of uncertainty and MCMC and profiling giving the broadest. Similar results were seen for F : F_{MSY}.

.Violation of the assumptions with respect to the indices used for fitting may result in biased estimates of estimated parameters, and hencereference points and stock trends. For variance estimates obtained from bootstrapping often assume that residuals are Independently and Identically Distributed (i.i.d.). Although uncertainty estimates was compared for a variety of methods, unlike Magnusson *et al*. (2012) we did not compare the estimates to known values through simulation, this should be done using Management Strategy Evaluation (MSE) to evaluate how to use estimates of uncertainty from stock assessment as part of a harvest control rule (HCR).

Bibliography

L. Kell, I. Mosqueira, P. Grosjean, J. Fromentin, D. Garcia, R. Hillary, E. Jardim, S. Mardle, M. Pastoors, J. Poos, *et al*. Flr: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science: Journal du Conseil, 64(4):640, 2007.

A. Magnusson, A. E. Punt, and R. Hilborn. Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and mcmc. Fish and Fisheries, 2012.

K. Patterson, R. Cook, C. Darby, S. Gavaris, L. Kell, P. Lewy, B. Mesnil, A. Punt, V. Restrepo, D. W. Skagen, *et al.* Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries, 2(2):125{157, 2001.

M. Prager. Aspic-a surplus-production model incorporating covariates. Coll. Vol. Sci. Pap., Int. Comm.Conserv. Atl. Tunas (ICCAT), 28:218{229, 1992.

P. R. Wade. Bayesian methods in conservation biology. Conservation Biology, 14(5):1308{1316, 2000.

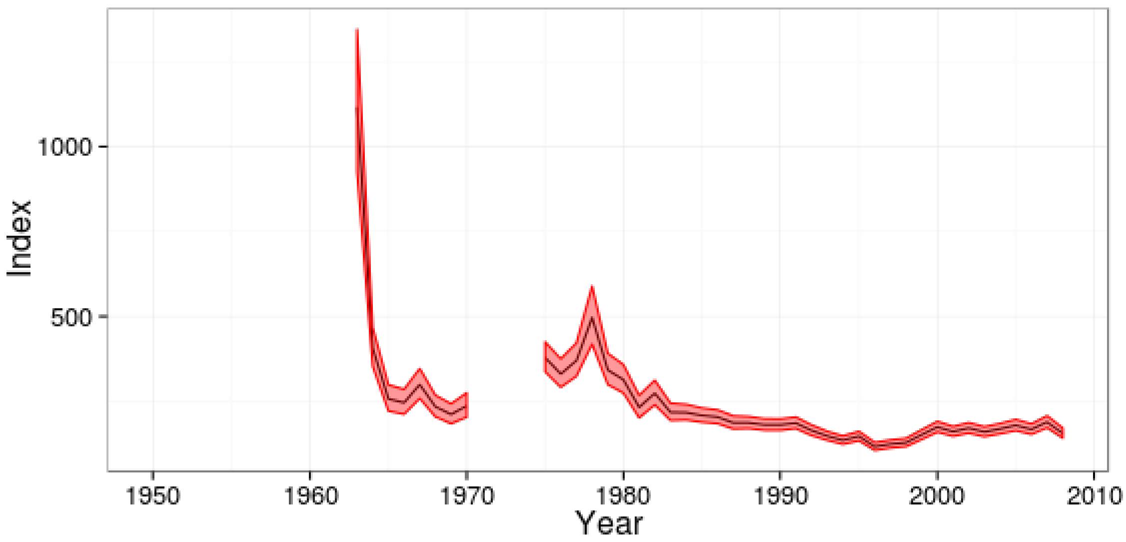

Figure 1. Catch per unit effort index with 10th and 90th confidence interval.

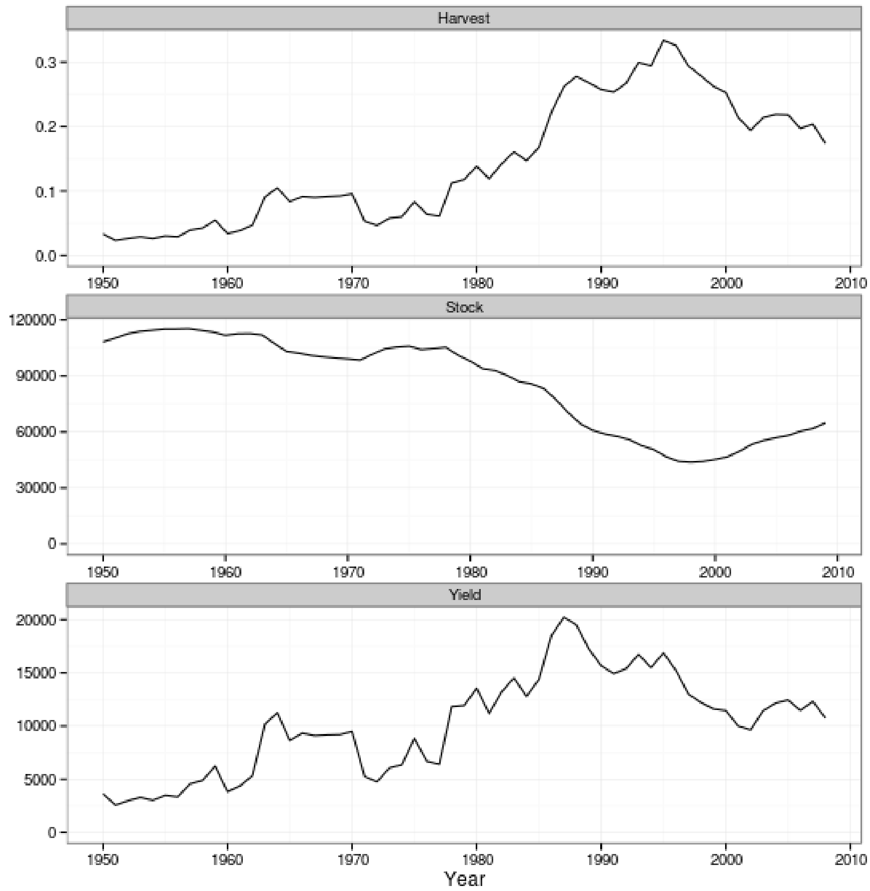

Figure 2. Time series of estimates of stock biomass, harvest rate and yield from ASPIC and biodyn assessments

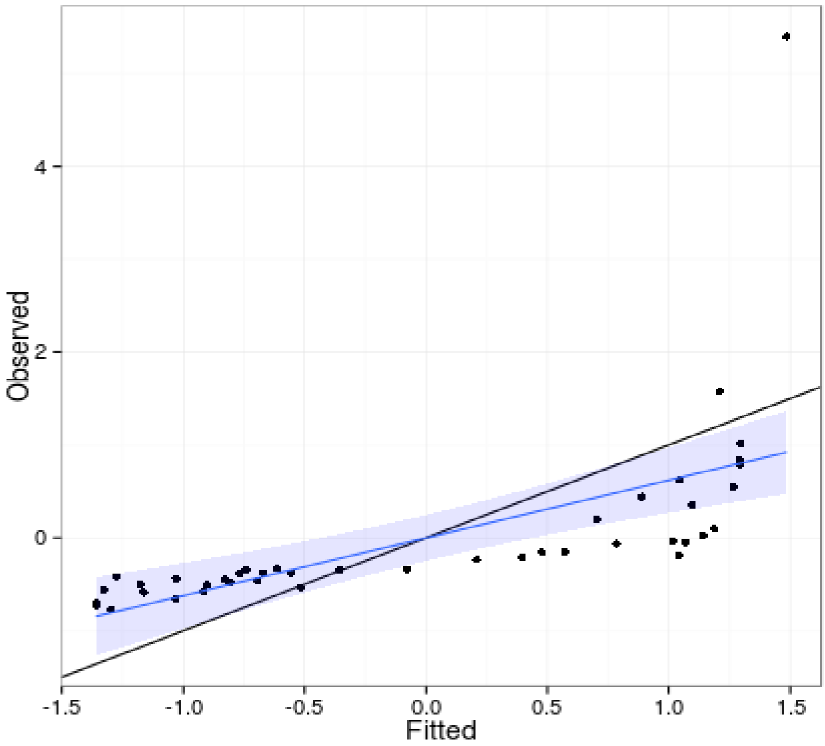

Figure 3. Observed CPUE verses fitted, blue line is a linear regression fitted to points, black the y=x line.

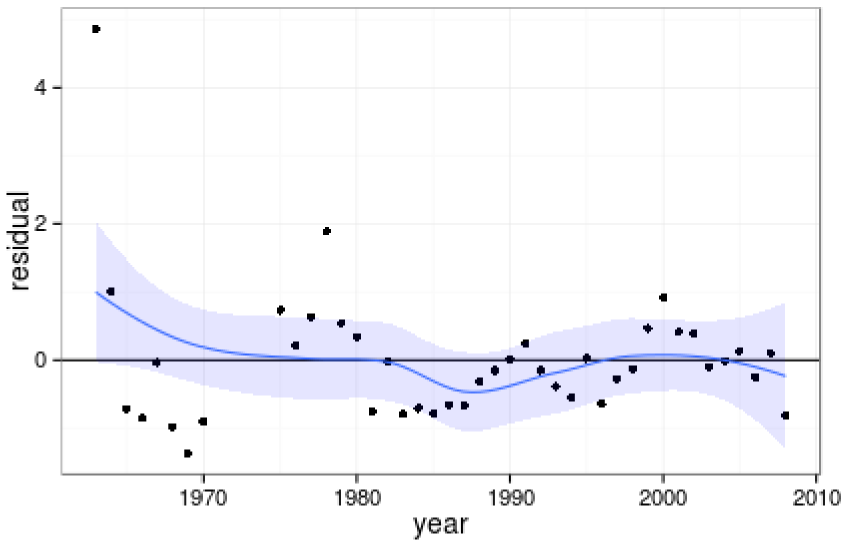

Figure 4. Residuals by year, with lowess smoother and SEs.

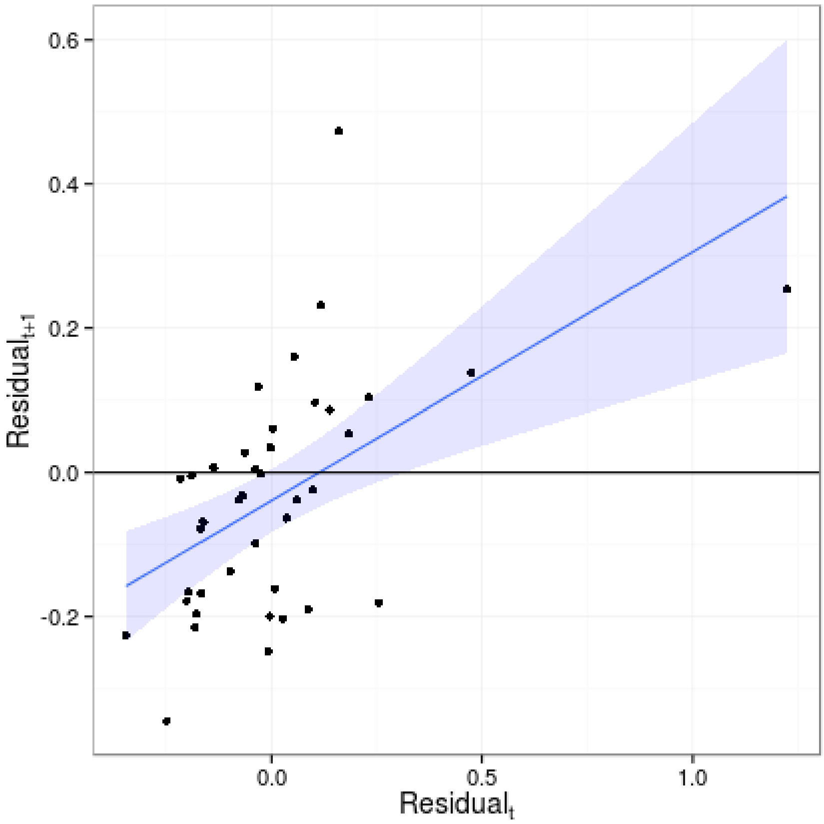

Figure 5. Quantile-quantile plot to compare residual distribution with the normal distribution.

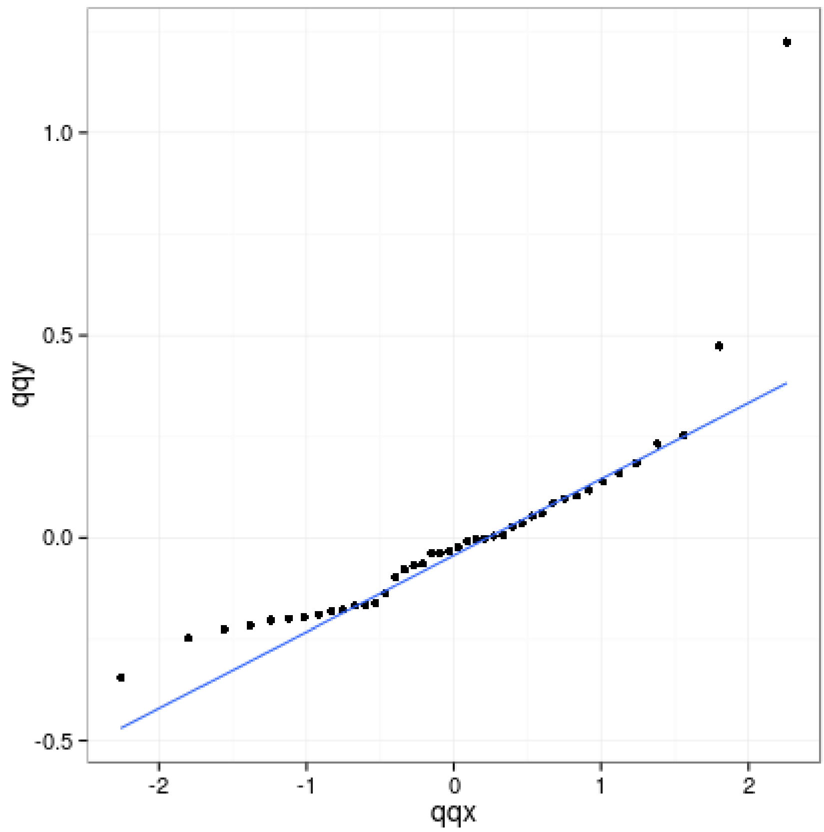

Figure 6. Plot of autocorrelation, i.e. residual$_{t+1}$ verses residual$_t$.

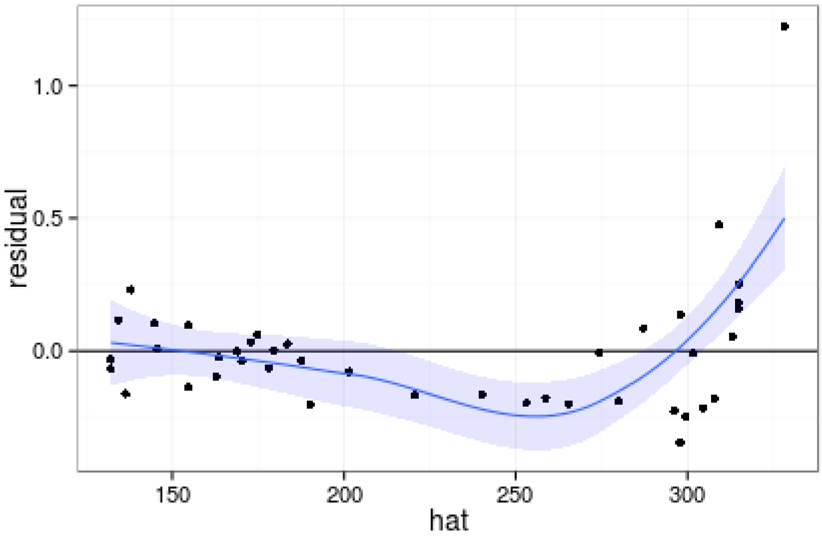

Figure 7. Plot of residuals against fitted value, to check variance relationship.

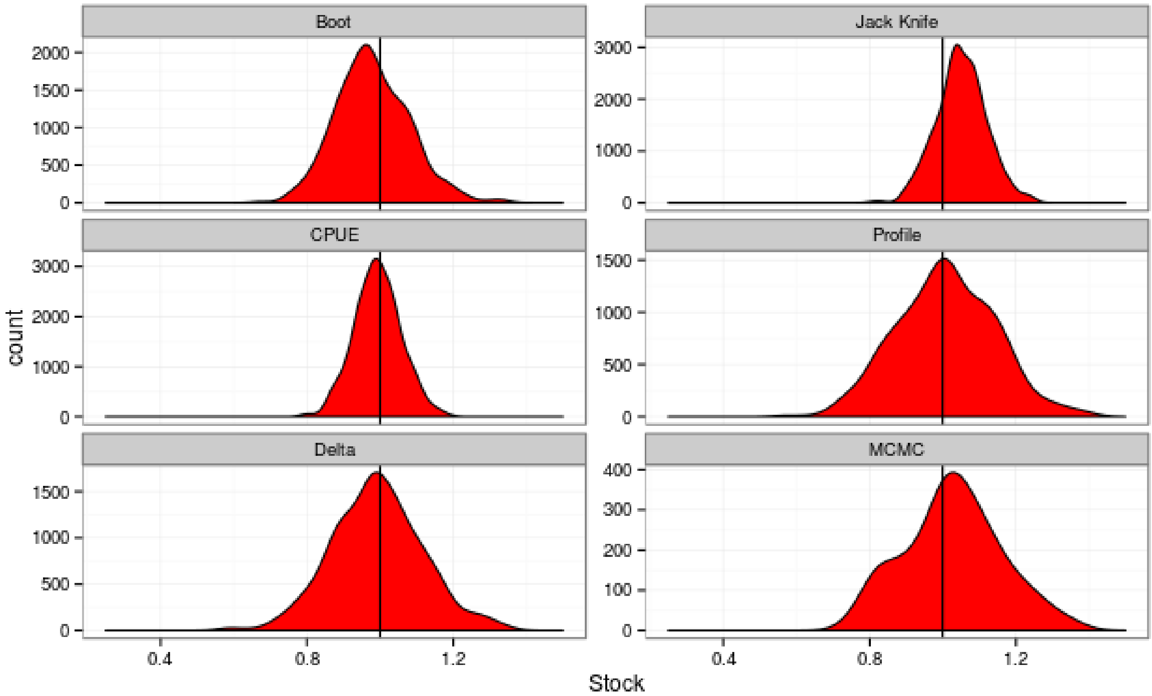

Figure 8. Densities of stock relative to B_{MSY} for each method of estimating parameter uncertainty.

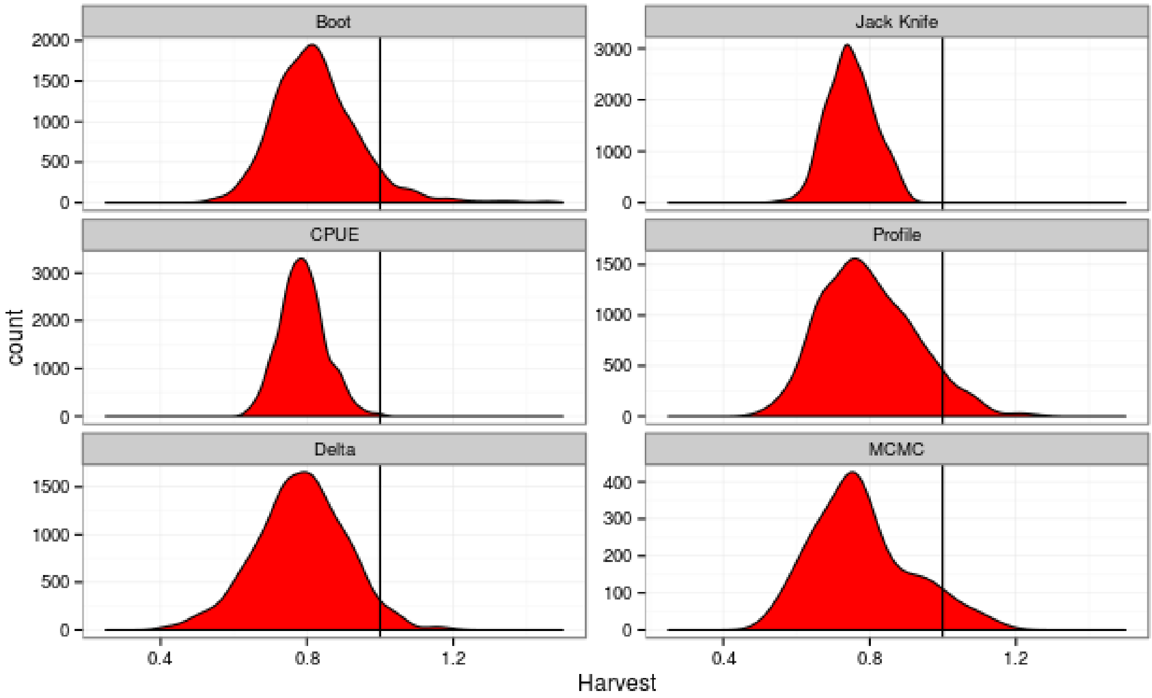

Figure 9. Densities of harvest rate relative to F_{MSY} for each method of estimating parameter uncertainty.

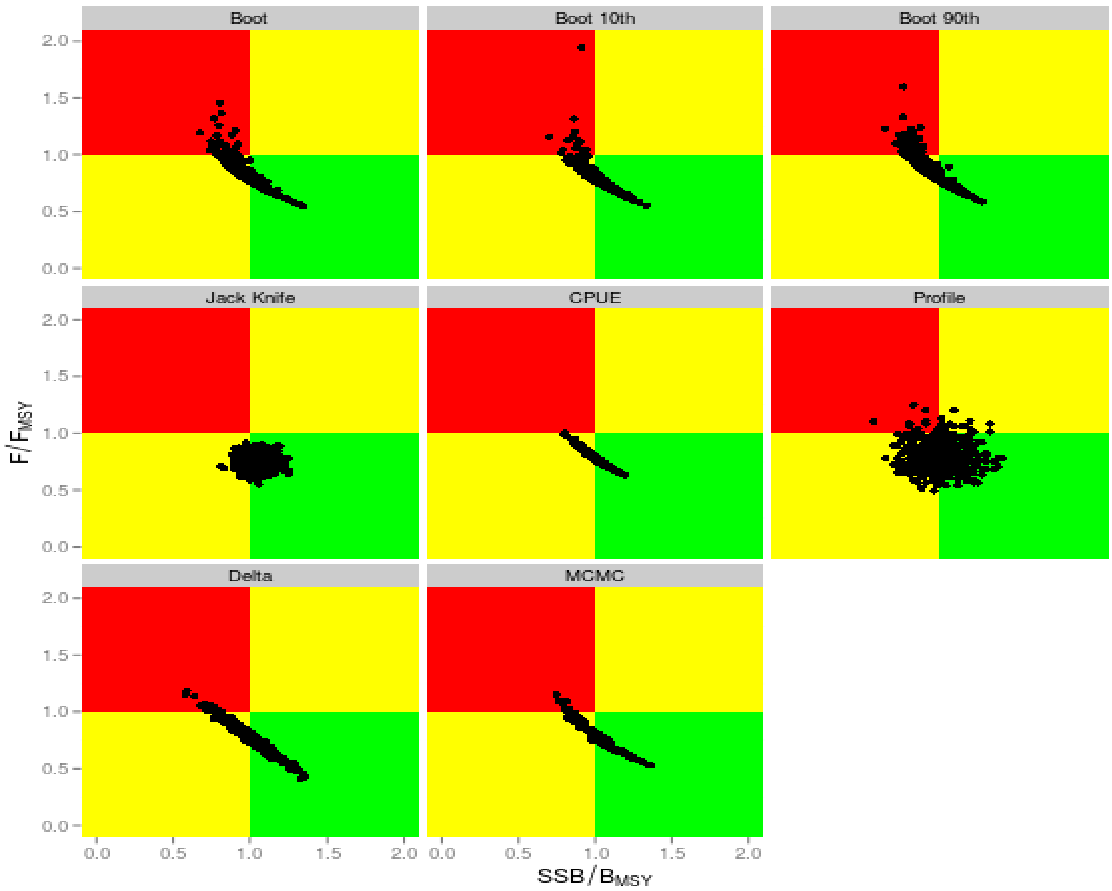

Figure 10. Kobe phase plot with points showing stock relative to B_{MSY} and harvest rate relative to F_{MSY} for each method of estimating parameter uncertainty.