SCRS/2015/073 Collect. Vol. Sci. Pap. ICCAT, 72(1): 245-265 (2016)

STOCK ASSESSMENT DIAGNOSTICS FOR ATLANTIC BIGEYE TUNA

L. Kell[1,] and G. Merino

SUMMARY

*As the complexity of models increase diagnostics become more important to understand the robustness of estimates and how they propagate through to advice. Diagnostics also make the stock assessment process more transparent and help identify where more knowledge and better data are required. A generic strategy for conducting stock assessment was proposed at the Atlantic bigeye (*Thunnus obesus*) data preparatory meeting. This was to i) agree in advance hypotheses to test; ii) check for convergence; iii) identify violation of assumptions by plotting residuals; iv) use methods such as the jack knife or bootstrap to identify problems with the data and model specifications; and v) conduct hindcasts to evaluate predictive ability and hence robustness of advice. Although the diagnostics presented here are for a biomass dynamic model they are generic and applicable to models that use different datasets and a variety of structures.*

RÉSUMÉ

*Plus la complexité des modèles augmente, plus les diagnostics acquièrent d'importance pour comprendre la solidité des estimations et la façon dont elles se reflètent dans l'avis. Les diagnostics rendent également le processus d'évaluation des stocks plus transparent et contribuent à identifier à quels endroits une plus grande quantité et de meilleures données sont nécessaires. Une stratégie générique pour mener l'évaluation des stocks avait été proposée à la réunion de préparation des données sur le thon obèse (*Thunnus obesus*) de l'Atlantique. Celle-ci consistait à i) convenir à l'avance des hypothèses à tester ; ii) vérifier la convergence ; iii) identifier la non-application des hypothèses en représentant les valeurs résiduelles ; iv) utiliser des méthodes, telles que l'eustachage (« jack knife ») ou le bootstrap pour identifier les problèmes avec les données et les spécifications du modèle et v) réaliser des simulations pour évaluer la capacité prédictive et, de fait, la solidité de l'avis. Même si les diagnostics présentés dans ce document concernent un modèle dynamique de biomasse, ils sont génériques et applicables aux modèles qui utilisent différents jeux de données et une variété de structures.*

RESUMEN

*A medida que aumenta la complejidad de los modelos, los diagnósticos se vuelven más importantes para entender la robustez de las estimaciones y el modo en que se reflejan en el asesoramiento. Los diagnósticos también hacen que el proceso de evaluación de stock sea más transparente y contribuyen a identificar dónde se requieren más conocimientos y mejores datos. En la Reunión de preparación de datos sobre patudo del Atlántico (*Thunnus obesus*) se propuso una estrategia genérica para realizar la evaluación de stock. Dicha estrategia consiste en: i) acordar previamente la hipótesis que se tiene que probar, ii) comprobar la convergencia; iii) identificar los supuestos que no se cumplen realizando un diagrama de residuos; iv) utilizar métodos como "jack knife" o bootstrap para identificar problemas con los datos y las especificaciones del modelo y v) realizar una simulación retrospectiva para evaluar la capacidad predictiva y, por tanto, la robustez del asesoramiento. Aunque los diagnósticos presentados correspondían a un modelo de dinámica de biomasa, dichos diagnósticos son genéricos y pueden aplicarse a modelos que utilizan diferentes conjuntos de datos y una variedad de estructuras.*

KEYWORDS

Bigeye, Diagnostics, Stock Assessment

[1] ICCAT Secretariat, C/Corazón de María, 8. 28002 Madrid, Spain; Laurie.Kell@iccat.int; Phone: +34 914 165 600; Fax: +34 914 152 612.

Introduction

As the complexity of models increase diagnostics become more important to understand the robustness of estimates and how they propagate through to advice. Diagnostics also make the stock assessment process more transparent and help identify where more knowledge and better data are required (Kell et al., 2013a, 2013b). A generic strategy for conducting stock assessment was proposed at the Atlantic bigeye (*Thunnus obesus*) data preparatory meeting, i.e. i) agree in advance hypotheses to test; ii) check for convergence; ii) identify violation of assumptions by plotting residuals; iii) use methods such as the jack knife or bootstrap to identify problems with the data and model specifications; and iv) conduct hindcasts to evaluate predictive ability and hence the robustness of advice. Although the diagnostics presented are for a biomass dynamic model they are generic and applicable to models that use different datasets and a variety of structures.

In 2009, assessment advice for bigeye was based on ASPIC (A Stock Production Model Incorporating Covariates, Prager, 1992) a non-equilibrium biomass dynamic model. ASPIC uses time series of indices of abundance and catch biomass to estimate stock status and bootstrapping to construct sampling distribution for statistics of interest, e.g. stock status, maximum sustainable yield (MSY) and the biomass and fishing rate (B_{MSY} and F_{MSY} respectively) that would produce it at equilibrium. A number of assessment scenario were run in 2009, each assessment used a single catch per unit effort (CPUE) time series based on a generalised linear model (GLM) that combined the individual CPUE series that had been standardised by CPCs. Several stock assessment specifications were considered by the working group e.g. i) initial guesses for estimated parameters, ii) form of the production function and iii) the assumed B0/K ratio. In total 96 specifications were run of which three runs were chosen to take forward to provide management advice.

Biomass dynamic models have been criticised (e.g. Maunder, 2003) as being too simplistic to capture the actual population dynamics. However, if a simple model can provide advice on stock status relative to reference points and predict the response of a stock to management why use anything more complicated (Ludwig and Walters, 1985)?

Material and Methods

Data

The data used were CPUE series in biomass as standardised by CPCs and time series of total catch biomass (Task I). At the data preparatory meeting rather than using a combined index it was decided to base assessment scenarios on the individual CPUE series standardised by CPCs. This allows hypotheses about stock and fishery dynamics to be proposed and the use of goodness of fit diagnostics to be used to accept or reject assessment scenarios.

Methods

The assessment is conducted using a biomass dynamic stock assessment model, which assumes that the biomass next year (B_{t+1}) is the sum of the biomass this year (B_t), less the catch (C_t) plus the surplus production (P_t) i.e.

$$B_{t+1}=B_t-C_t+P_t$$

The production function simplifies recruitment, growth and natural mortality into a single function. Life history arguments have been used to show that the Schaefer production model is probably not appropriate for tunas (Maunder, 2003), and that a Pella-Tomlinson form where $B_{MSY}<0.5B_0$ is probably more realistic due to high steepness. We therefore use the Pella-Tomlinson (1969) surplus production function i.e.

$$r/pB(1-(B/K)^p)$$

Productivity and hence reference points are determined by the intrinsic population growth rate at low population sizes (r) and the shape of the production function (p). If p=1 then MSY is found halfway between 0 and K, as p increases MSY shifts to the right.

Since there is seldom sufficient information in the catch data to estimate even the limited parameters of the production function, additional data such as time series of relative abundance based on CPUE or surveys are required for calibration.

The ASPIC and biodyn R packages (http://www.flr-project.org/) were used for fitting, plotting, examining goodness of fit diagnostics, simulation, estimating uncertainly in stock status relative to reference points, deriving quantities used in management (Kell *et al*., 2007). The packages can also be used to run projections, simulating Harvest Control Rules (HCRs), conducting Management Strategy Evaluation (MSE), (2014).

Analysis

CPUE series

To help agree in advance what hypotheses to test a hierarchical cluster analysis (Murtagh and Legendre, 2014) was used to group the CPUE series; where the distance between two clusters is the ANOVA sum of squares between them summed over all the variables. Strong year-classes may show up as lags in the correlations between CPUE indices that target different age-classes, therefore cross-correlations were examined to look for evidence of tracking of cohorts.

Likelihood profiles by CPUE series (i.e. data component) were plotted to evaluate the information in each series in relation to the estimated parameters.

Likelihood profiles

Likelihood profiles are used to check that a solution has actually been found and to evaluate the information content of the data. It is not uncommon for indices to contain insufficient information to estimate the parameters of a stock assessment model. Indices may also be conflicting and fitting therefore involves weighting averages of contradictory trends. This generally produces parameter estimates intermediate to those obtained from the data sets individually. Schnute and Hilborn (1993) pointed out that the most likely parameter values are not intermediary to conflicting values; instead, they occur at one of the apparent extremes.

Residual Analysis

Patterns in the residuals from the fits to the CPUE, may result in biased estimates of parameters, reference points and stock trends. Therefore when fitting a model the residuals should be checked to identified violations of the assumptions.

To check the distribution of the residuals the observed quantiles can be plotted against predicted quantiles from the assumed distribution. Q-Q plots do this by comparing 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 assumed; any systematic departure from a straight line may indicate skewness or over or under dispersion.

In the assessment model it is assumed that an index is proportional to the stock, so when plotting the observed against the fitted values the points should fall around the y=x line. If they do not, then the index may not be a good proxy for the stock trend. Patterns in the residuals, e.g. by year due to changes in fishing practice, can be identified by plotting the residuals against year. It is also assumed that variance does not vary with the mean, this assumption can be checked by plotting the residuals against the fitted values.

Estimates of variance obtained from bootstrapping and other techniques assume that residuals are independently and identically distributed (i.i.d.). This assumptions can be validated by inspection of the residuals above. It is assumed that the residuals are not auto-correlated; plots of the residuals against each other with a lag of 1 to identify autocorrelation. Significant autocorrelations could be due to an increase in catchability with time; which may result in a more optimistic estimate of current stock status as any decline in the stock is masked by an increase in catchability.

Fits

After fitting methods such as the jack-knife or bootstrap can check for problems with the data and model specifications; e.g. due to highly correlated or ill-defined parameters. Where parameters are fixed due to lack of information in the data, e.g. the shape parameter (p) and initial depletion (B_0), a sensitivity analysis can be conducted where an assessment is run for a range of fixed values.

An objective of stock assessment is to provide advice on the response of a stock to management. This requires that the prediction ability of models is validated. If a data series is regarded as being representative of the dynamics of the stock then it can be used as a model-free validation measure using a hindcast. A model is fitted to the first part of a time series and then projected over the omitted period; the best performing model can then be identified by comparing the predictions with observations. Methods like Akaike's Information Criterion, which are commonly used to compare and choose between multiple model fits, require a common dataset. However, in stock assessment the evaluation of alternative model structures and assumptions often requires different datasets to be used. In contrast hindcasting has few parametric or theoretic assumptions, is conceptually simple and can be used with different models and datasets. Also when evaluating the robustness of management advice it is predictions that are of primary interest not fits to historic data. It provides an objective technique for evaluating the impact of model choice and assumptions on predictions, i.e. a red face test. The approach also helps to identify what data are informative and the impact of alternative scenarios.

Statistical Uncertainty

There are a variety of methods for estimating variance in parameter estimates and propagating this into the estimates of stock status and projections (i.e. the Kobe phase plot and strategy matrix), see Kell *et al.*, (2014c)

The bootstrap 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. It can also estimate bias and so can be used to identify model mis-specification and problems with the data (Canty, 2006).

An alternative is the jackknife, or leave one out procedure, 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.

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. If as part of the fitting process the Hessian (i.e. the matrix of second-order partial derivatives) is estimates, 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.

Another approach is Markov chain Monte Carlo (MCMC) 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 (Gelman and Hil, 2007; Wade 2000.) 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

Profiling can be used to estimate confidence intervals.

Results

CPUE Series

The available biomass CPUE are plotted in **Figure 1**, the points are the standardised indices, the red line is the prediction from a GAM fitted to all the indices and the blue line the smooth values for an individual index. The correlation matrix is plotted in **Figure 2,** blue indicate a positive correlation and red negative, the order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities.

The cross correlations between the indices are plotted in **Figure 3**.

Likelihood profiles

The 2009 ASPIC runs used to provide advice are compared to a fit to the 2015 catch data where the population parameters (K & MSY) were fixed using the 2009 run 2 estimates (**Figure 4**). A profile of the residual sum of squares was then conducted for K (**Figure 5)** and MSY **(Figure 7)** using all the indices**.** This shows that the estimates of both K and MSY based on all the indices are lower than obtained in 2009 (red point). Next profiles are presented by data component (**Figure 6 and 8);** this shows that all the indices support a lower value of K and MSY than estimated in 2009, although not all display a minimum.

In order to explore residual diagnostics, a single index, Chinese Taipei, was chosen as it is one of the longer time series and showed a minimum. **Figure 9** compares the fit to those obtained above and **Figure 10** and **Figure 11** present the profiles of K and MSY to check the fit.

Residual Analysis

Patterns in residuals of the fits to the CPUE and stock abundance may indicate a violation of models assumptions. Which in turn may result in biased estimates of parameters, reference points and stock trends.

Figure 12 plots the observed CPUE against the fitted values, the blue line is a linear regression fitted to points, black the y=x line. If the index is a good proxy for stock abundance the two lines should coincide. The residuals are then plotted against year along with a lowess smoother (**Figure 13**).

In addition variance estimates obtained from bootstrapping assume that residuals are Independently and Identically Distributed (i.i.d.). **Figure 14** shows a Quantile-quantile plot to compare residual distribution with the normal distribution. **Figure 15,** the residuals are plotted against the fitted value, to check variance relationship. It is assumed that the residuals are not autocorrelated. Plots of the residuals against each other with a lag of 1 to identify autocorrelation. Significant autocorrelations could be due to an increase in catchability with time; which may result in a more optimistic estimate of current stock status as any decline in the stock is masked by an increase in catchability, **Figure 16** therefore checks for autocorrelation.

Fits

The estimated stock trend is compared to the CPUE index in **Figure 17,** the hindcast is presented in **Figure 19,** and the stock/catch compared to the **s**urplus production curve in **Figure 18**.

A main objective of stock assessment is to estimate uncertainly in stock status (Magnusson, 2012). This requires estimates of distributions as well as point estimates. There are various ways to estimate undercertainty in parameter estimates and quantities derived from them, i.e. use the covariance matrix provided by a maximum likelihood fit, bootstrapping, the jack knife or Bayesian methods such as Monte Carlo Markov Chain,

For fitting we use the biodyn package as this incorporates many more tools for estimation of uncertainty. First the ASPIC and biodyn fits are compared in **Figure 20**. Then the sensitivity of time series estimates to assumed shape of the Pell-Tomlinson production function in **Figures 21 and 22** and to assumed initial depletion level (**Figure 23**).

The senstivity of the estimates of B/B_{MSY} and F/F_{MSY} in 2014 to each index point is then evaluated using the Jack Knife (**Figure 24).**

Figure 25 presents the autocorrelation in MCMC samples.

The densities of the estimates of current biomass to B_{MSY}, derived from the covariance matrix, MCMC, bootstrap and jack-knife are presented in (**Figure 26**) and for F_{MSY} in

Figure 27. Figure 28 show the corresponding kobe phase plots with marginal densities.

Discussion

The diagnostics are generic and can be applied to models that use other datasets and estimate more parameters and so can be used to compare models with different structures. As the complexity of models increase diagnostics become more important to understand the robustness of estimates and how they propagate through to advice. Diagnostics also make the stock assessment process more transparent and help identify where more knowledge and better data are required.

Biomass dynamic models have been criticised (e.g. Maunder, 2003) as being too simplistic to capture the actual population dynamics. However, if a simple model can provide advice on stock status relative to reference points and predict the response of a stock to management why use anything more complicated (Ludwig and Walters, 1985)? For example the Pella-Tomlinson model is used to set catch limits for baleen whales by the IWC. Neither the form of the model nor its parameters are meant to provide an accurate representation of the dynamics of the population. Rather, it has been demonstrated by MSE that when used as an integral part of a management strategy with a HCRs it allows the robust calculation and setting of catches limits (Butterworth and Punt, 1999), see http://iccat-mse.github.io/ for the North Atlantic MSE that is evaluating the current advice framework base on a biomass dynamic model.

Conclusion

The diagnostics in this paper are generic and could be applied to all assessments methods.

A proposed strategy for running the assessment

1. Use the Pella-Tomlinson production function, since the logistic production model is probably not appropriate for tunas (Maunder, 2003). There is seldom sufficient information, however, in stock assessment data sets to estimate the shape parameter. Therefore use the life history relationships to estimate the ratio between B_{MSY} and K (Kell, *et al*., 2013defg).
2. Choose CPUEs to run as separate scenarios based on hypotheses that could eventually be tested, since it is not uncommon for indices to contain insufficient information to estimate the parameters of a stock assessment model. Indices may also be conflicting and fitting therefore involves weighting averages of contradictory trends. This generally produces parameter estimates intermediate to those obtained from the data sets individually. Schnute and Hilborn (1993) pointed out that the most likely parameter values are not intermediary to conflicting values; instead, they occur at one of the apparent extremes.
3. Check robust of the fits using jackknife/bootstrap and residual diagnostics.
4. Conduct projections as done for North Atlantic Albacore (Kell *et al*., 2013a) and Swordfish (Kell *et al*., 2014b).

References

Canty, A. J., Davison, A. C., Hinkley, D. V., and Ventura, V. (2006). Bootstrap diagnostics and remedies. Canadian Journal of Statistics, 34(1):5–27.

Gelman A. and Hill. J. (2007). Data analysis using regression and multilevel/hierarchical models, volume Analytical methods for social research. Cambridge University Press, New York.

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.

Kell, L., de Bruyn P., Merino G. and Ortiz de Urbina J. (2013a). Implementation of a harvest control rule for northern Atlantic albacore. ICCAT Collect. Vol. of Sci. Pap. 70: 1355-1364.

Kell, L., de Urbina J.M., O., and de Bruyn P. (2013b). Stock assessment dagnostics for north Atlantic swordfish. ICCAT Collect. Vol. of Sci. Pap., (70):1954–1963.

Kell, L., de Urbina J.M., O., and de Bruyn P. (2013c). Stock assessment diagnostics for south Atlantic swordfish. ICCAT Collect. Vol. of Sci. Pap., (70):1964–1969.

Kell, L.T., De Bruyn, P. (2013d). An Evaluation Of The Importance Of The Assumed Biological Parameters For Management Advice: Atlantic Bigeye. Collect. Vol. Sci. Pap. ICCAT, 69(5): 2052-2058.

Kell, L.T., Scott, F., De Bruyn, P., Mosqueira, I., (2013e). A Generic Population Simulator Based On Life History Theory. Collect. Vol. Sci. Pap. ICCAT, 69(5): 2172-2180.

Kell, L.T., 1 De Bruyn, P., 2 Scott, F., and Mosqueira, I., (2013f). Sensitivity of The Kobe II Strategy Matrix To Life History Assumptions. Collect. Vol. Sci. Pap. ICCAT, 69(5): 2181-2188

Kell, L.T., De Bruyn, P., and Mosqueira, I., (2103g). The Use of Life History Theory For Stock Assessment: An Albacore Example. Collect. Vol. Sci. Pap. ICCAT, 69(5): 2189-2200.

Kell, L.T., 1 , Ortiz de Urbina, J., 2 , De Bruyn, P. 1 , Mosqueira, I., 3 , Magnusson, A. (2014a). An Evaluation of Different Approaches For Modelling Uncertainty In Aspic And Biomass Dynamic Models. Collect. Vol. Sci. Pap. ICCAT, 70(5): 2111-2119.

Kell, L.T., Ortiz de Urbina, J., De Bruyn, P., P. (2014b). Kobe II Strategy Matrices For North Atlantic Swordfish Based On Catch, Fishing Mortality And Harvest Control Rules. Collect. Vol. Sci. Pap. ICCAT, 70(4): 2009-2016.

Kell, L. T., Merino, G., De Bruyn, P., Ortiz de Urbina, J., Arrizabalaga, H., Santiago. J. and Murua. H., (2014c). An Example Management Strategy Evaluation of A Harvest Control Rule. Collect. Vol. Sci. Pap. ICCAT, 70(5): 2096-2110.

Ludwig, D. and Walters, C. J. (1985). Are age-structured models appropriate for catch-effort data? Canadian Journal of Fisheries and Aquatic Sciences, 42(6):1066–1072.

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

Maunder, M. N. (2003). Is it time to discard the schaefer model from the stock assessment scientist's toolbox? Fish. Res., 61(1):145–149.

Murtagh, F. and Legendre, P. (2014). Wards hierarchical agglomerative clustering method: Which algorithms implement wards criterion? Journal of Classification, 318 31(3):274–295.

Pella, J. and Tomlinson, P. (1969). A generalized stock production model. Inter-American Tropical Tuna Commission.

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

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

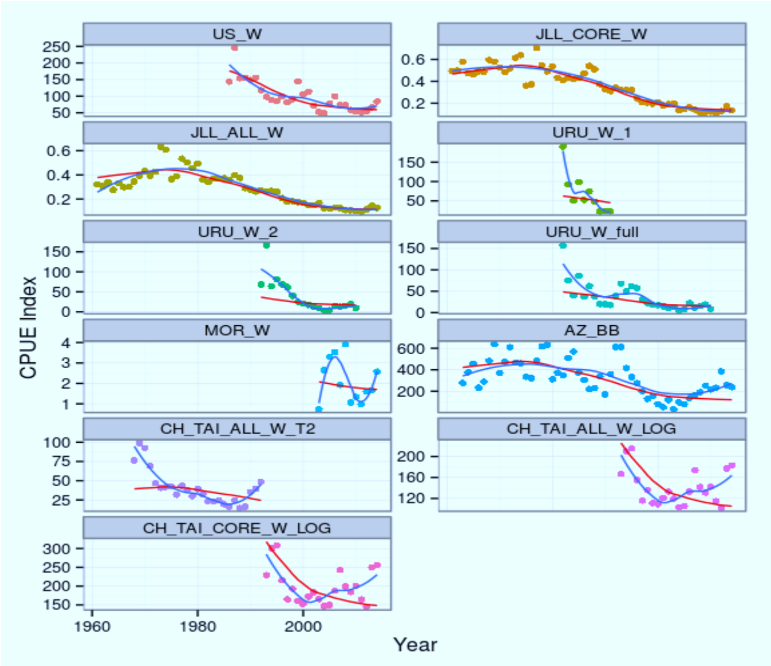

Figure 1. CPUE Series, points are the standardised indices, red line the prediction from a GAM fitted to all the indices and the blue line the smooth values by index.

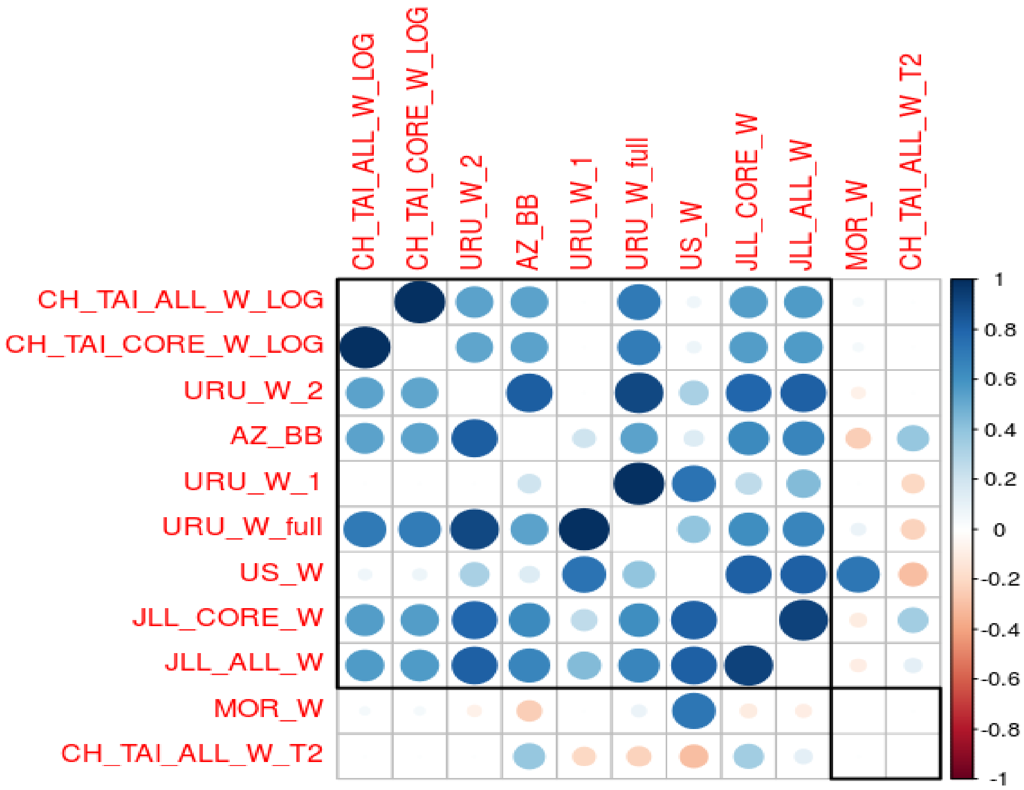

Figure 2. Hierachical cluster analysis of CPUE indices based on correlations between the indices.

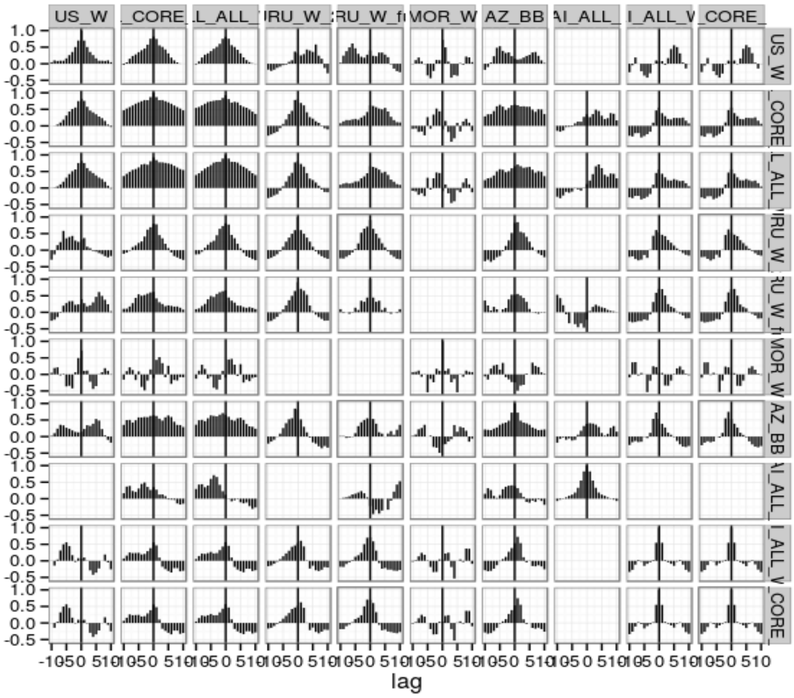

Figure 3. Cross correlations between indices.

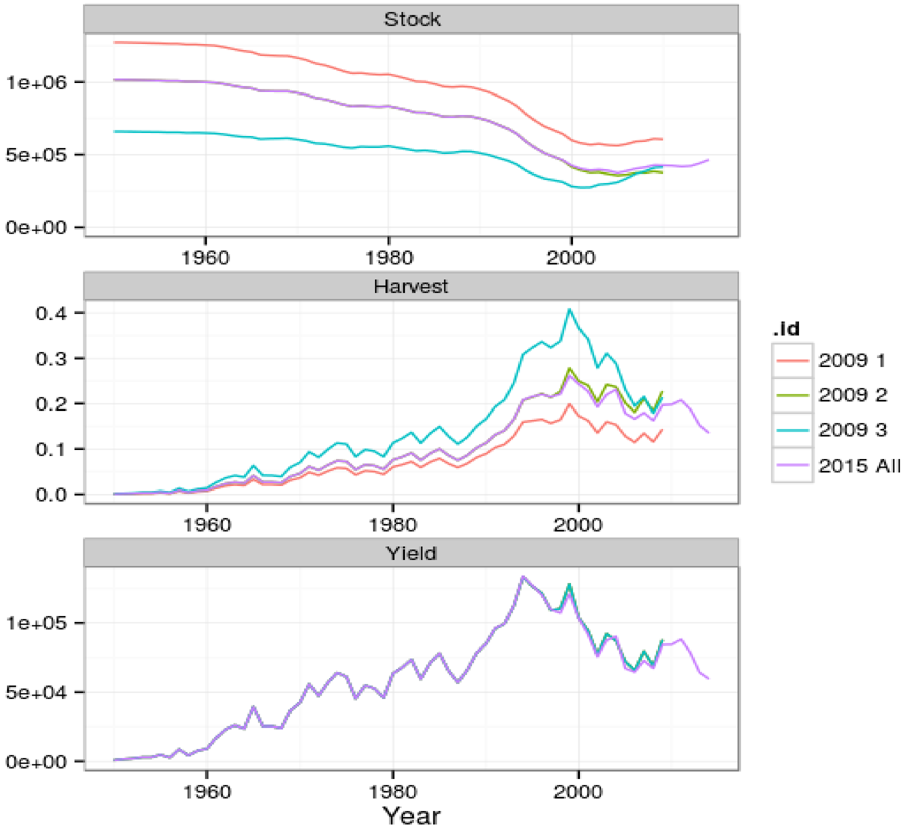

Figure 4. Comparison between 2009 ASPIC runs used to provide advice with a fit to 2015 catch data where population parameters (K & MSY) were fixed to the 2009 run 2 estimates.

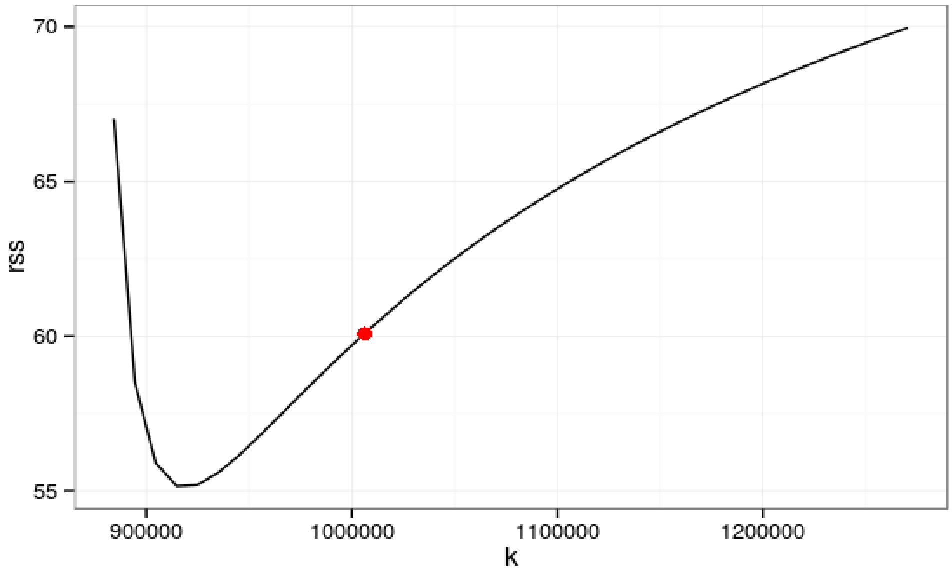

Figure 5. Profile of residual sum of squares for K, red point corresponds to 2009 ASPIC run 2.

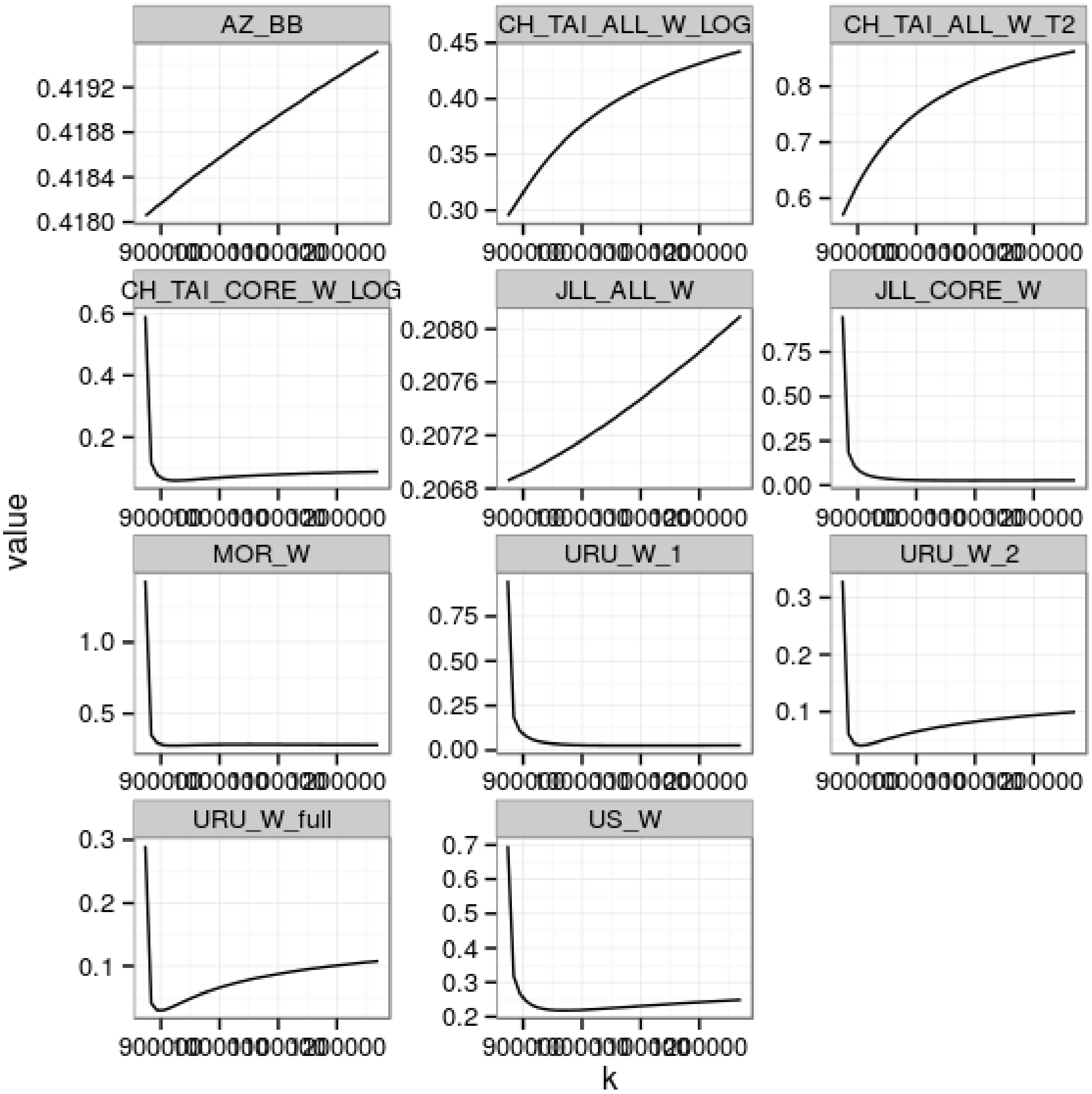

Figure 6. Profile of residual sum of squares for K by data components.

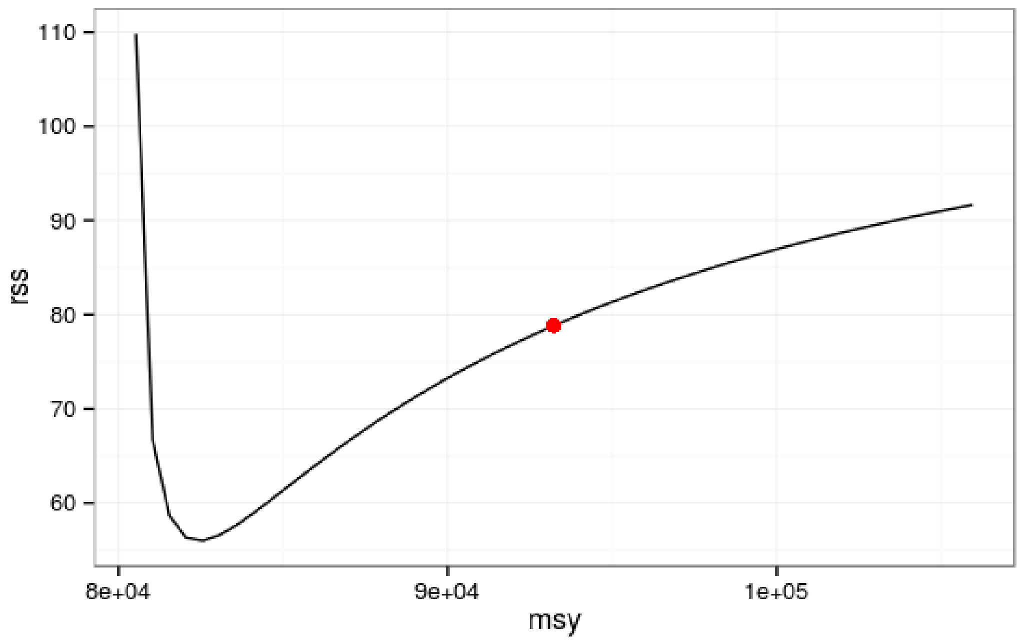

Figure 7. Profile of residual sum of squares for MSY, red point corresponds to 2009 ASPIC run 2.

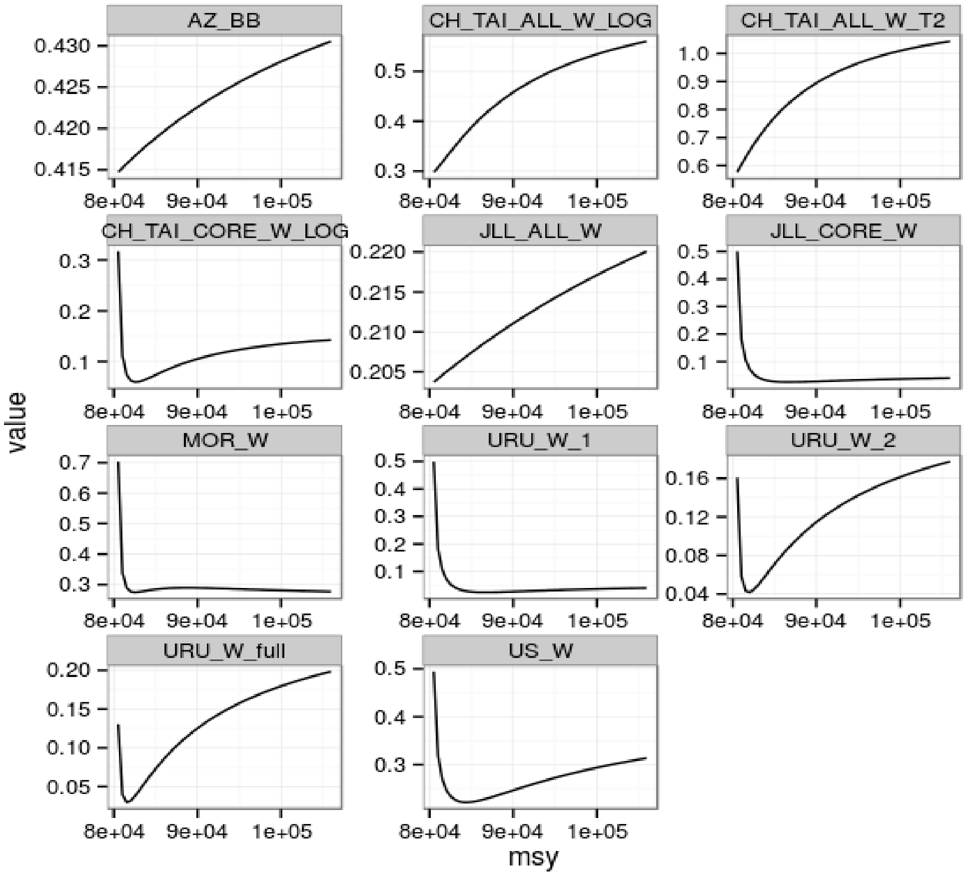

Figure 8. Profile of residual sum of squares for MSY by data component.

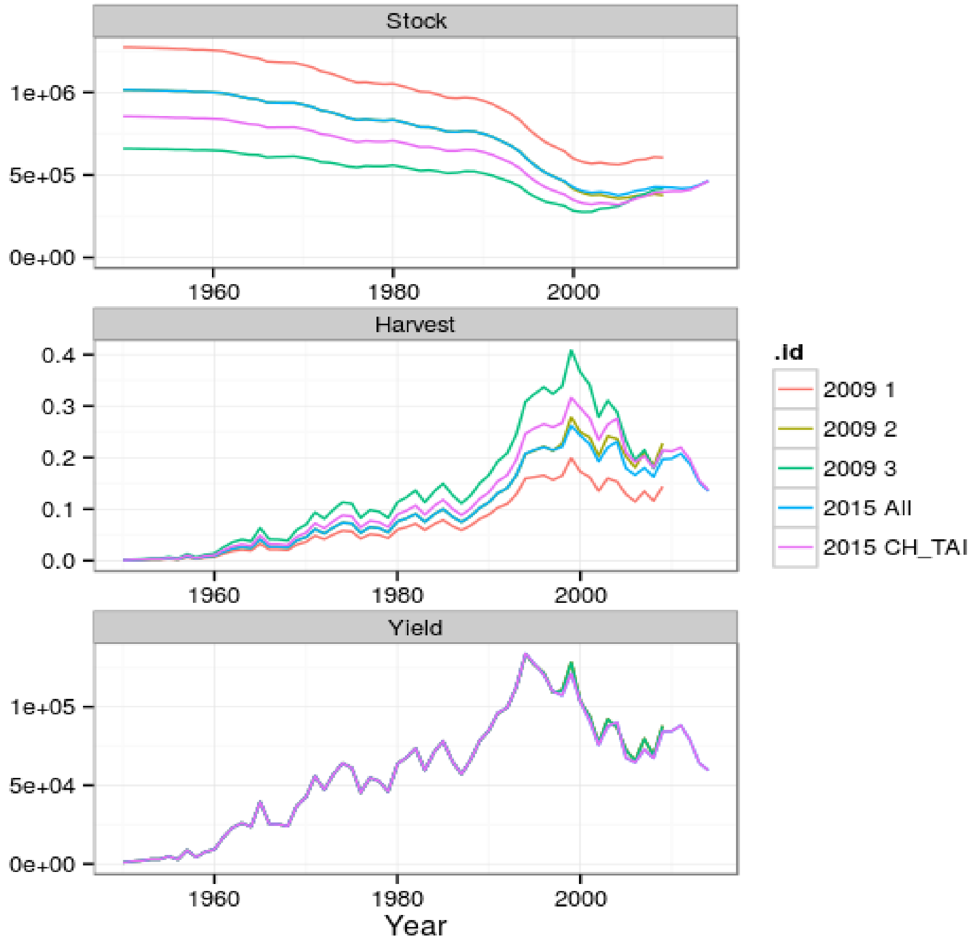

Figure 9. 2009 ASPIC runs and 2015 fit.

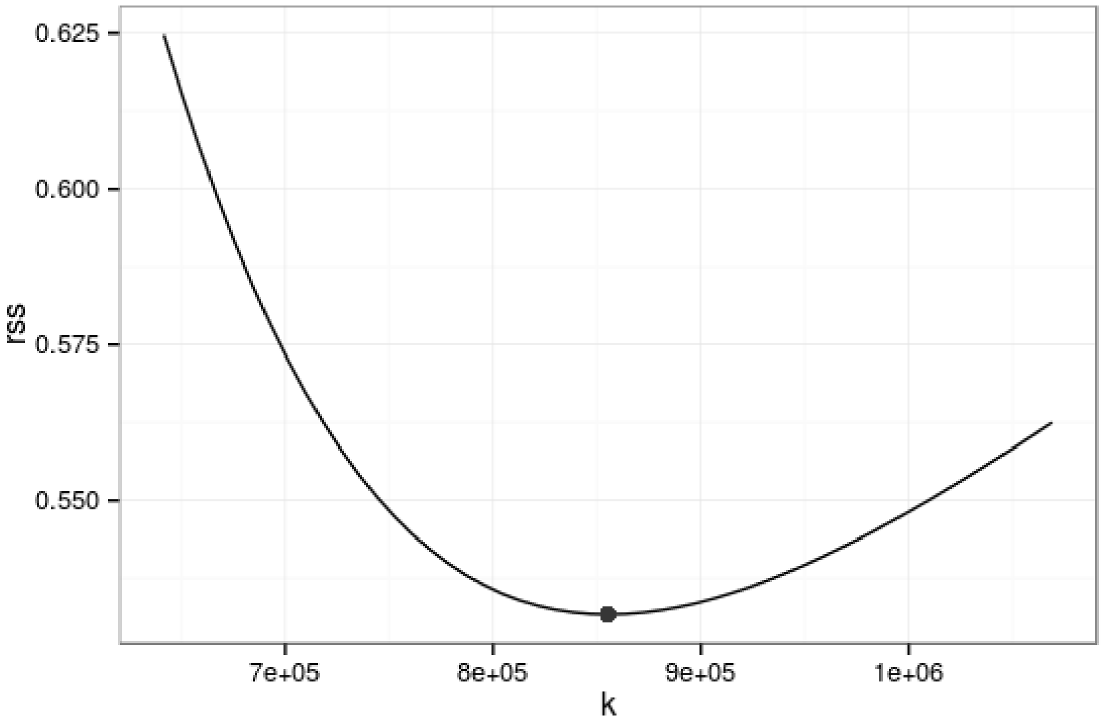

Figure 10. Profile of K for 2015 fit.

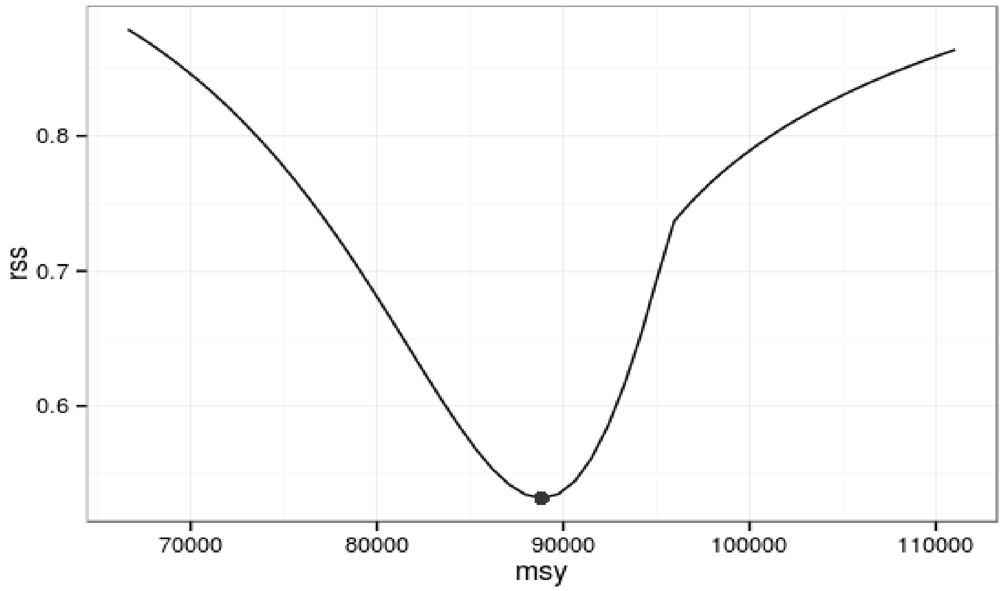

Figure 11. Profile of MSY for 2015 fit.

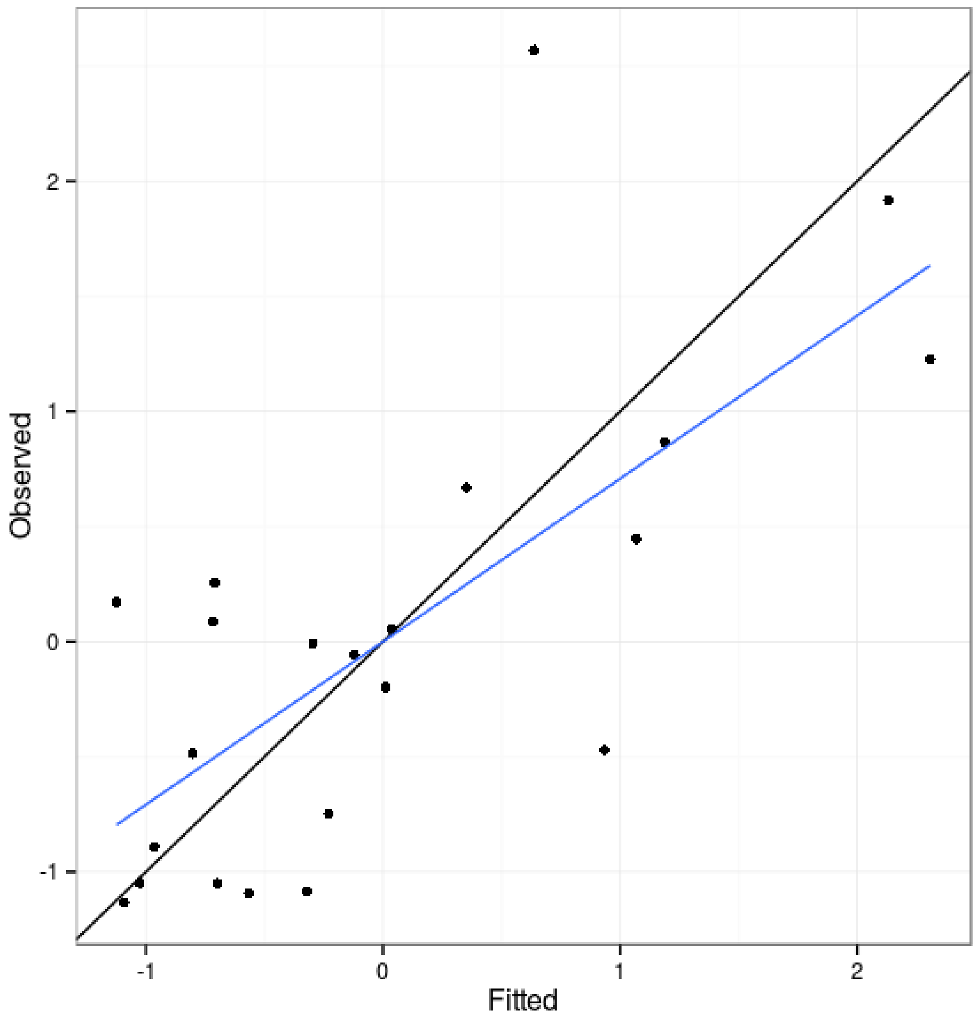

Figure 12. Observed CPUE verses fitted, blue line is a linear resgression fitted to points, black the y=x line.

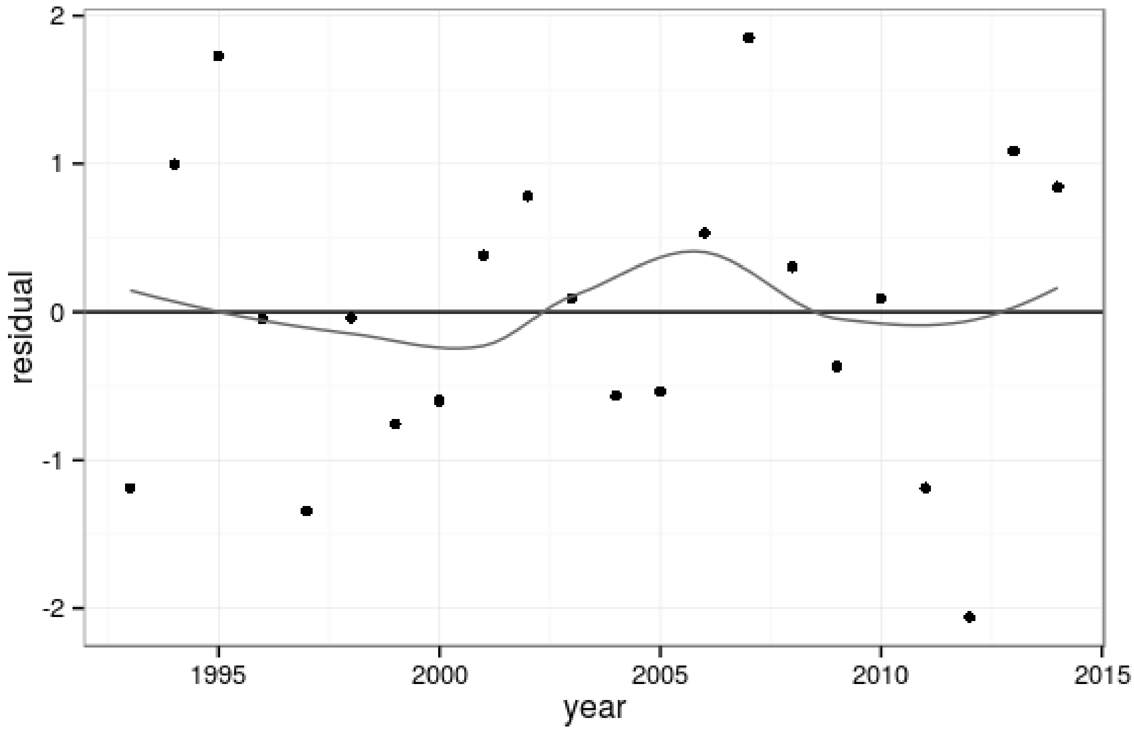

Figure 13. Residuals by year, with lowess smoother and SEs.

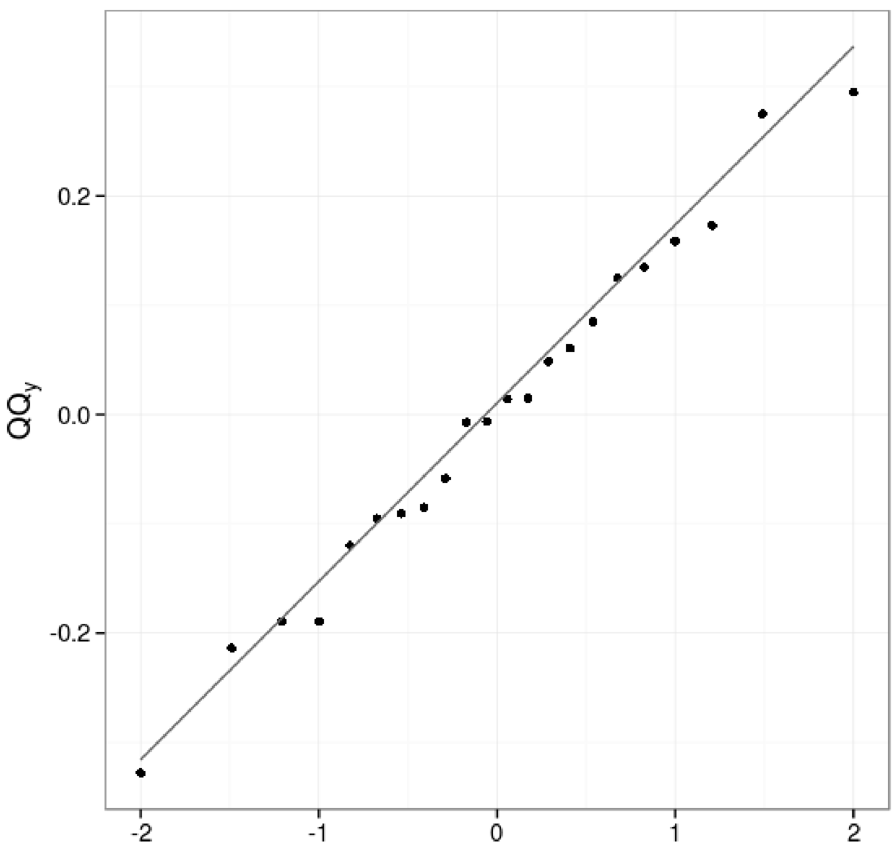

Figure 14. Quantile-quantile plot to compare residual distribution with the normal distribution.

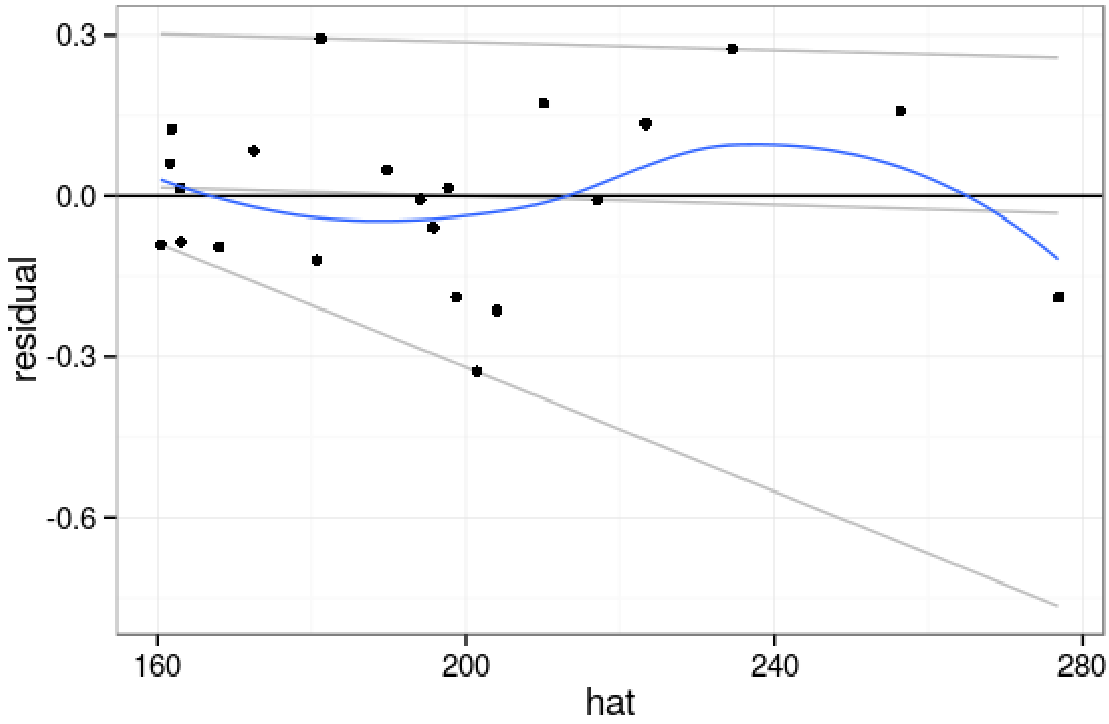

Figure 15. Plot of residuals against fitted value, to check variance relationship, blue line is a lowess smoother and lines the 5th, 50th and 95th percentiles estimated from a percentile regression.

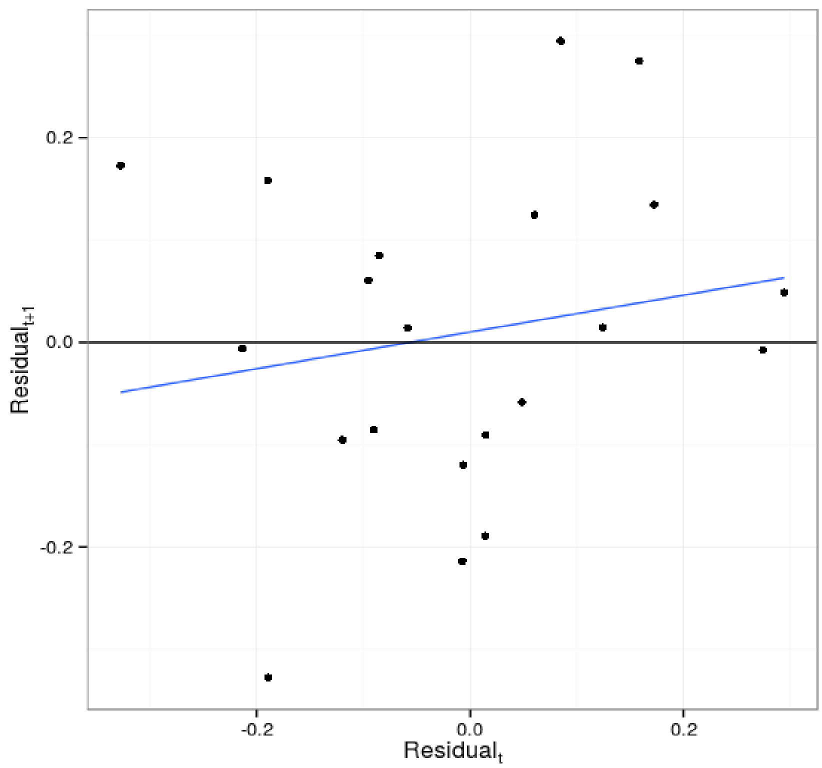

Figure 16. Plot of autocorrelation, i.e. $residual_{t+1}$ verses $residual_t$.

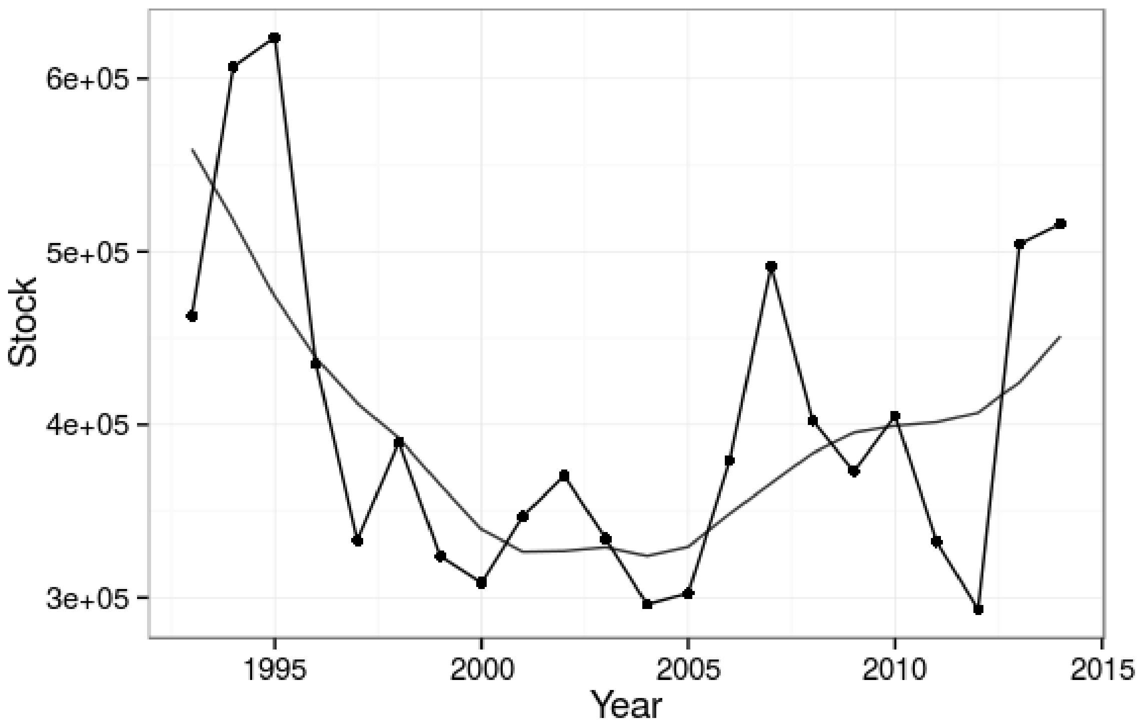

Figure 17.

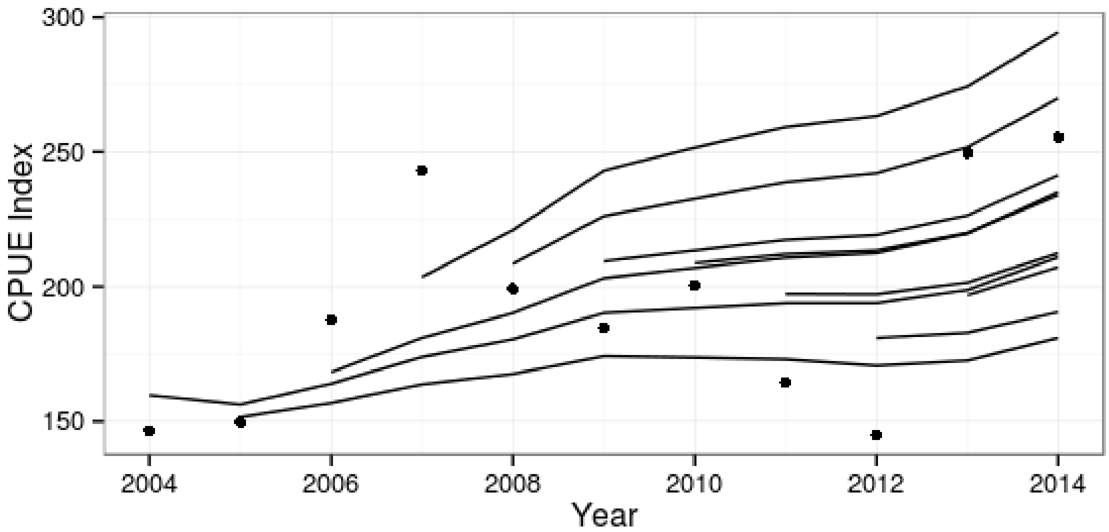

Figure 18. Hindcast

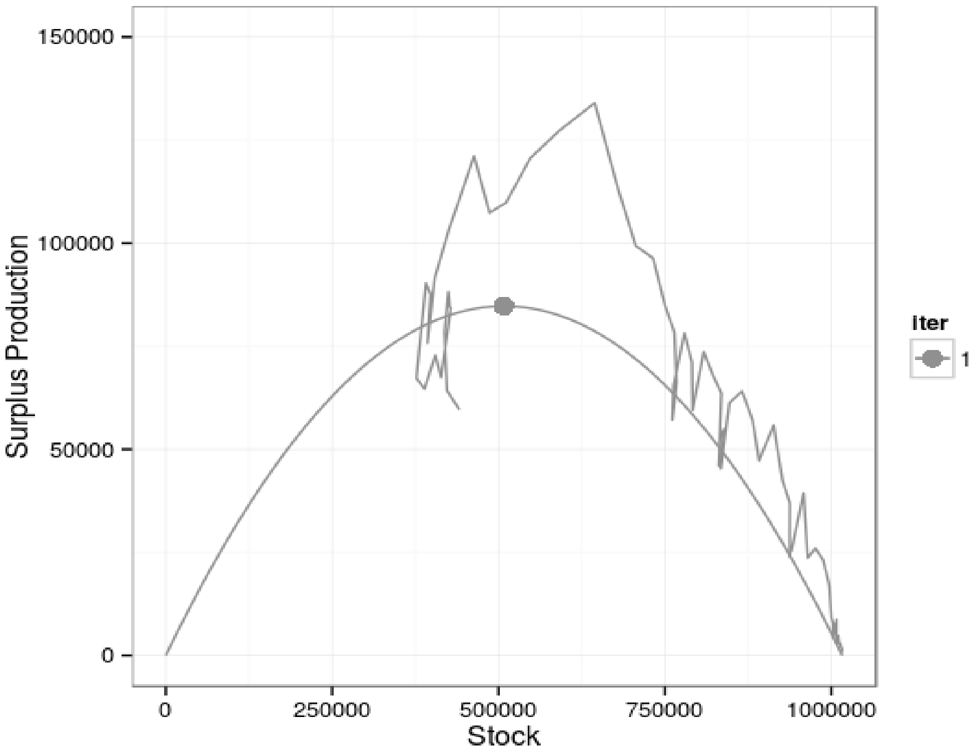

Figure 19 Surplus production curve.

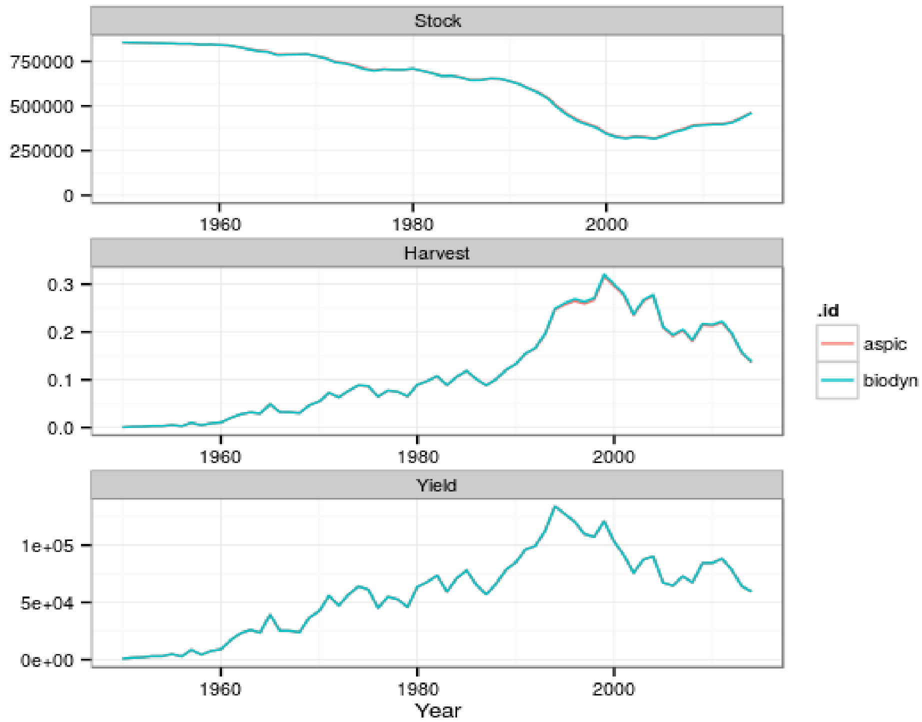

Figure 20. Comparison of ASPIC and biodyn fits.

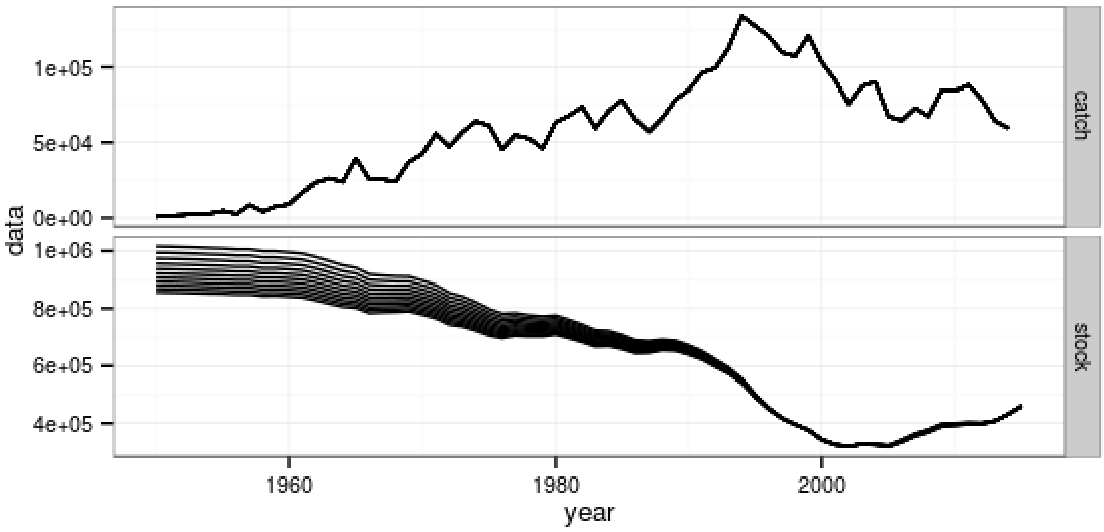

Figure 21. Sensitivity of time series estimates to assumed shape of the Pell-Tomlinson production function.

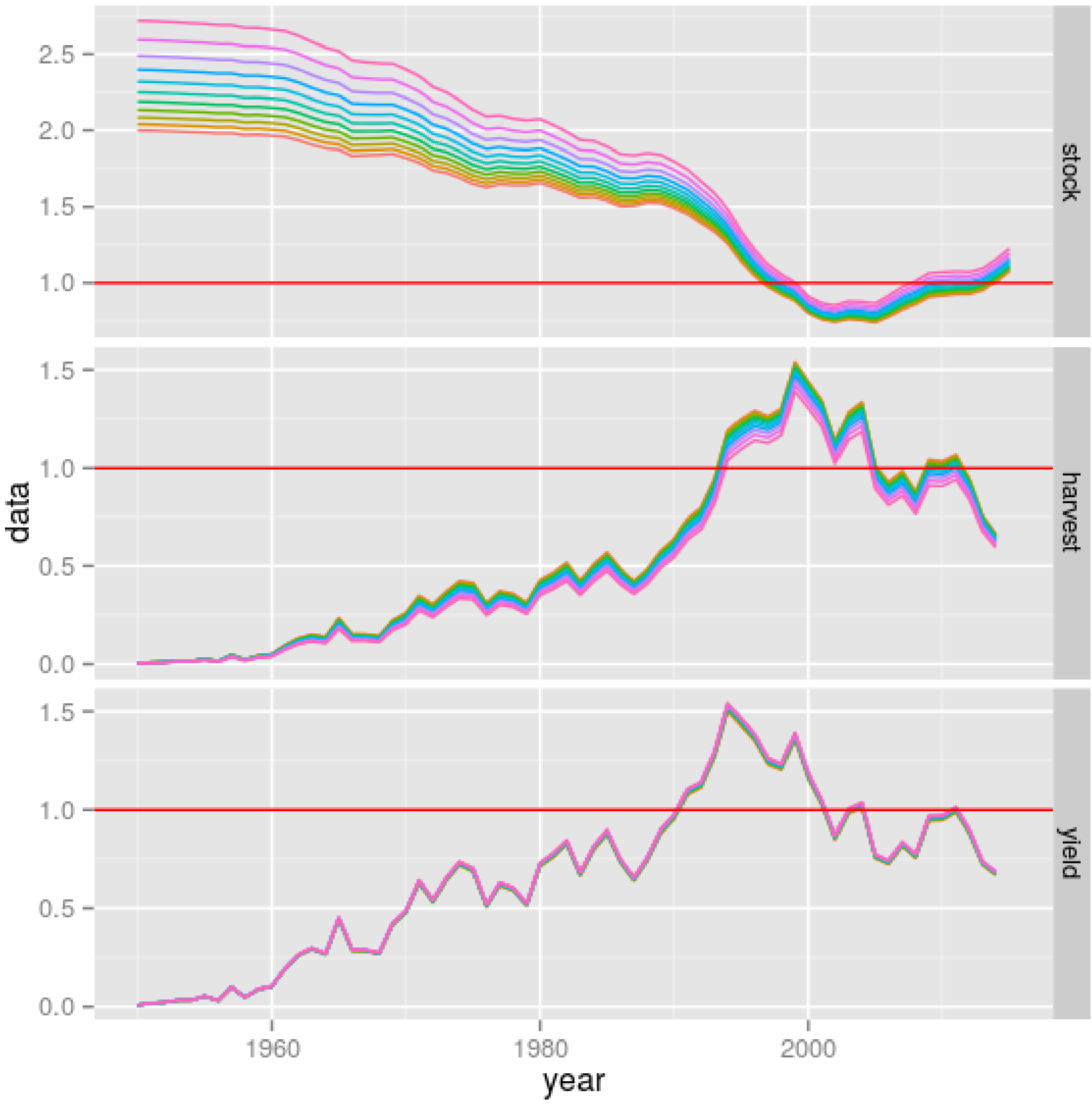

Figure 22. Sensitivity of time series estimates to assumed shape of the Pell-Tomlinson production function.

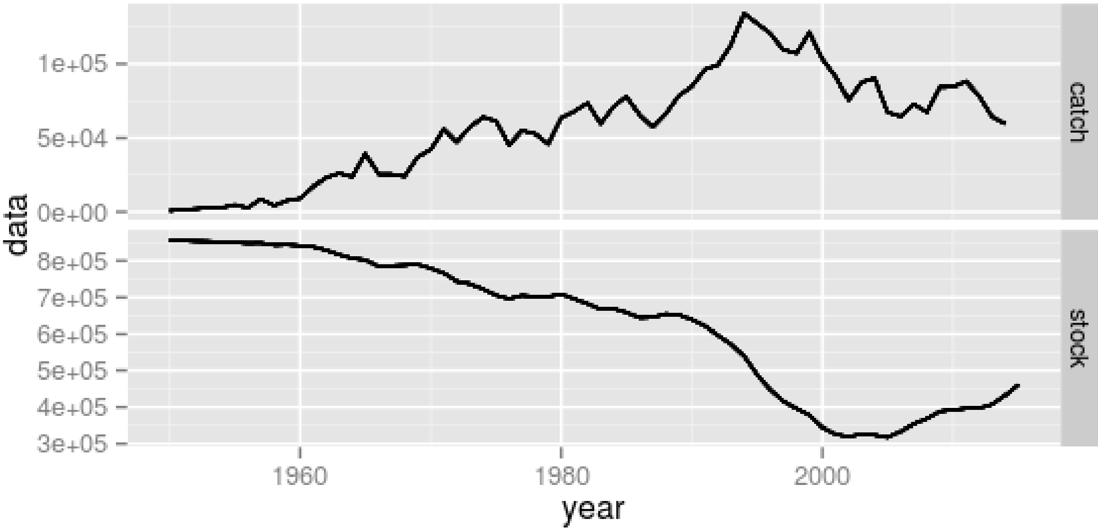

Figure 23. Sensitivity of time series estimates to assumed initial depletion level.

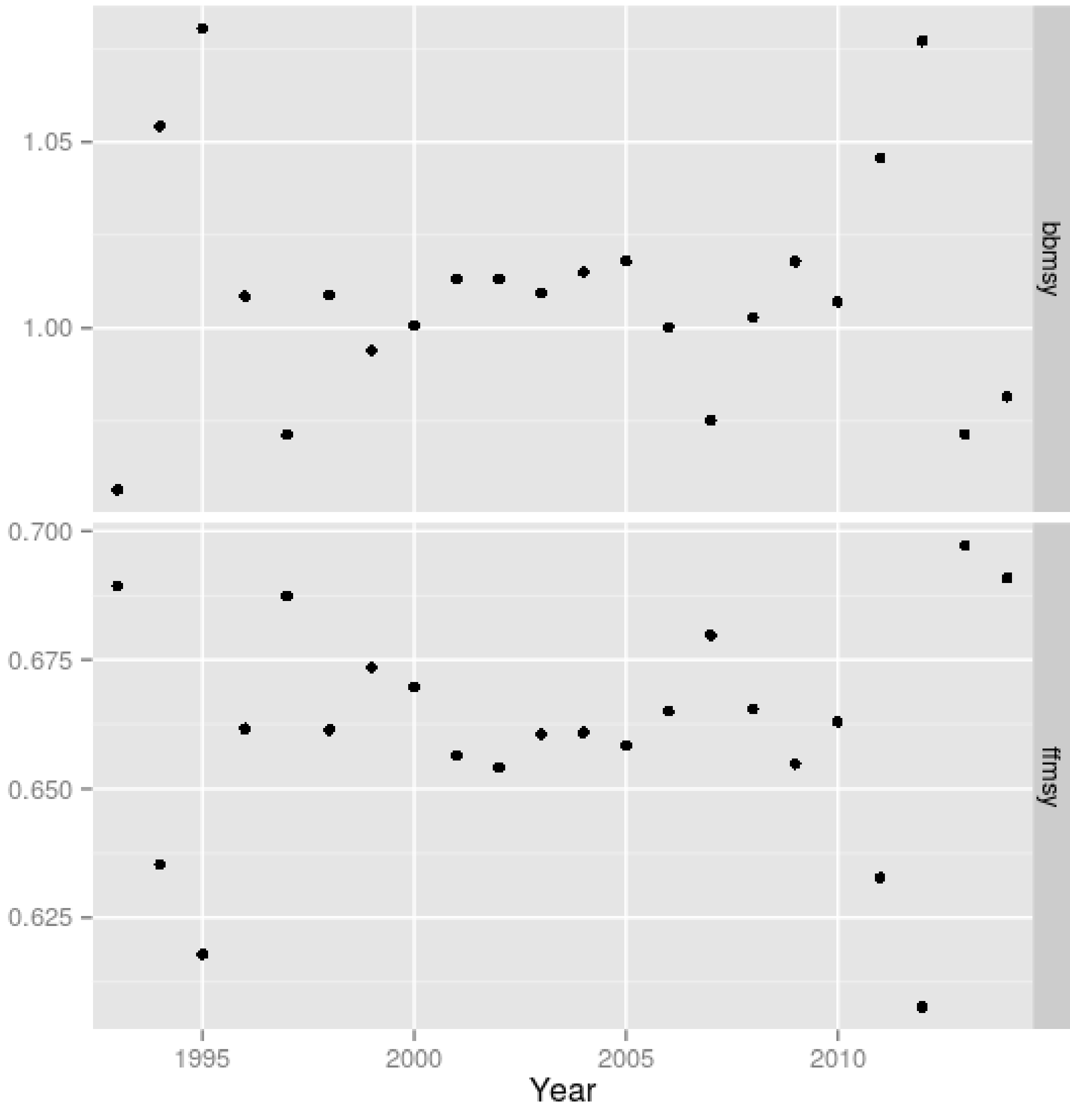

Figure 24. Estimates of B/B_{MSY} and F/F_{MSY} by Jack Knife replicate.

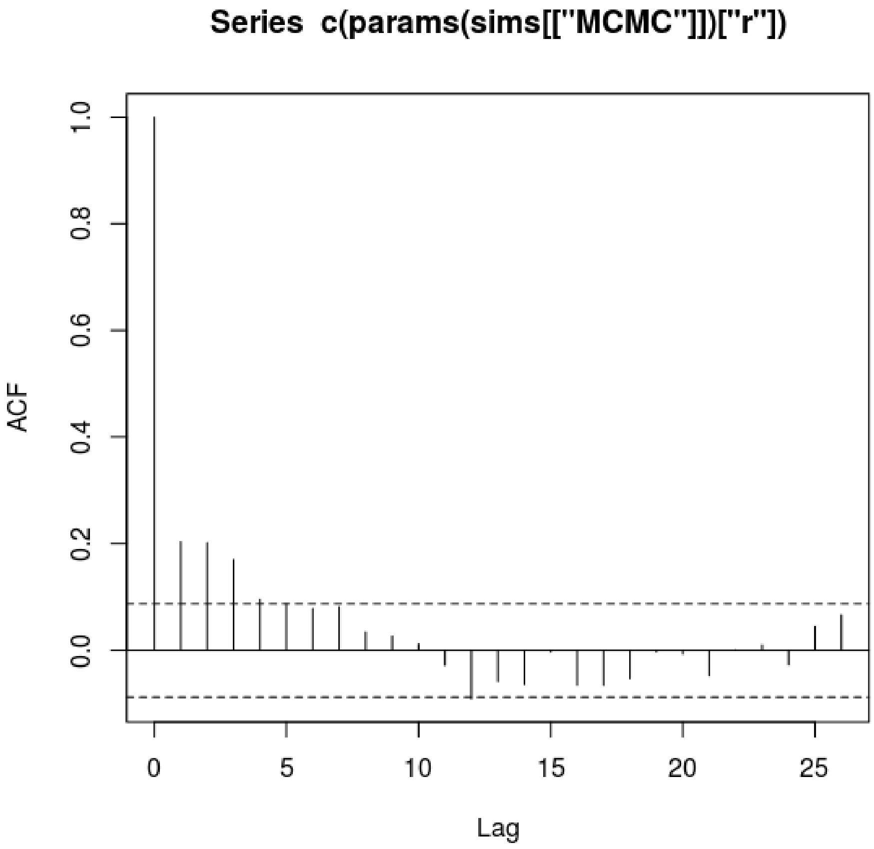

Figure 25. Autocorrelation between MCMC samples.

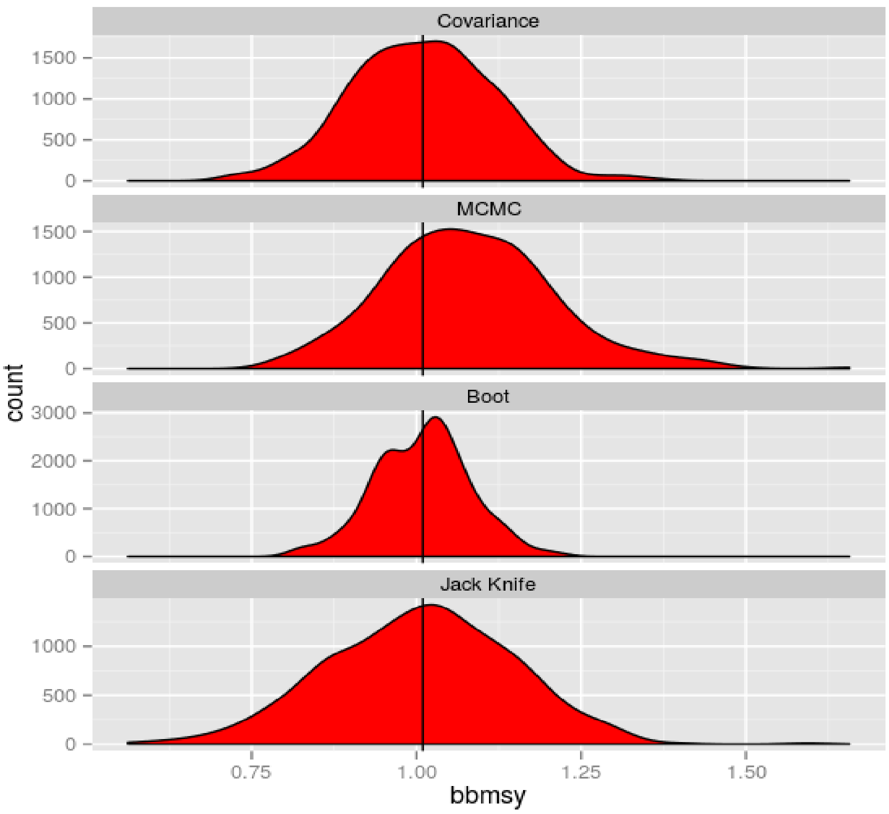

Figure 26. Densities of estimates of current biomass to B_{MSY}, derived from the covariance matrix, MCMC, bootstrap and jack-knife,

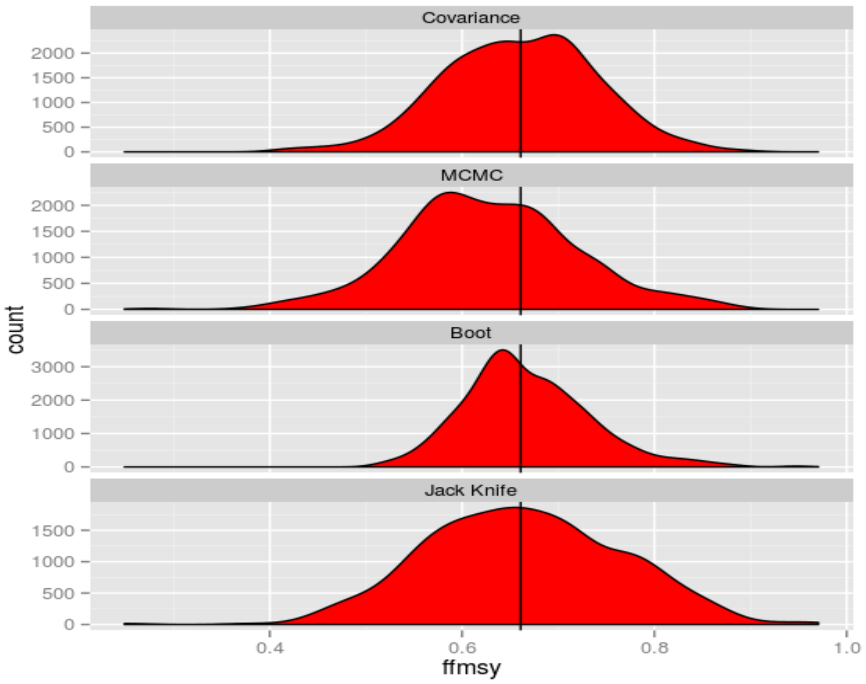

Figure 27. Densities of estimates of current harvest rate to F_{MSY}, derived from the covariance matrix, MCMC, bootstrap and jack-knife.

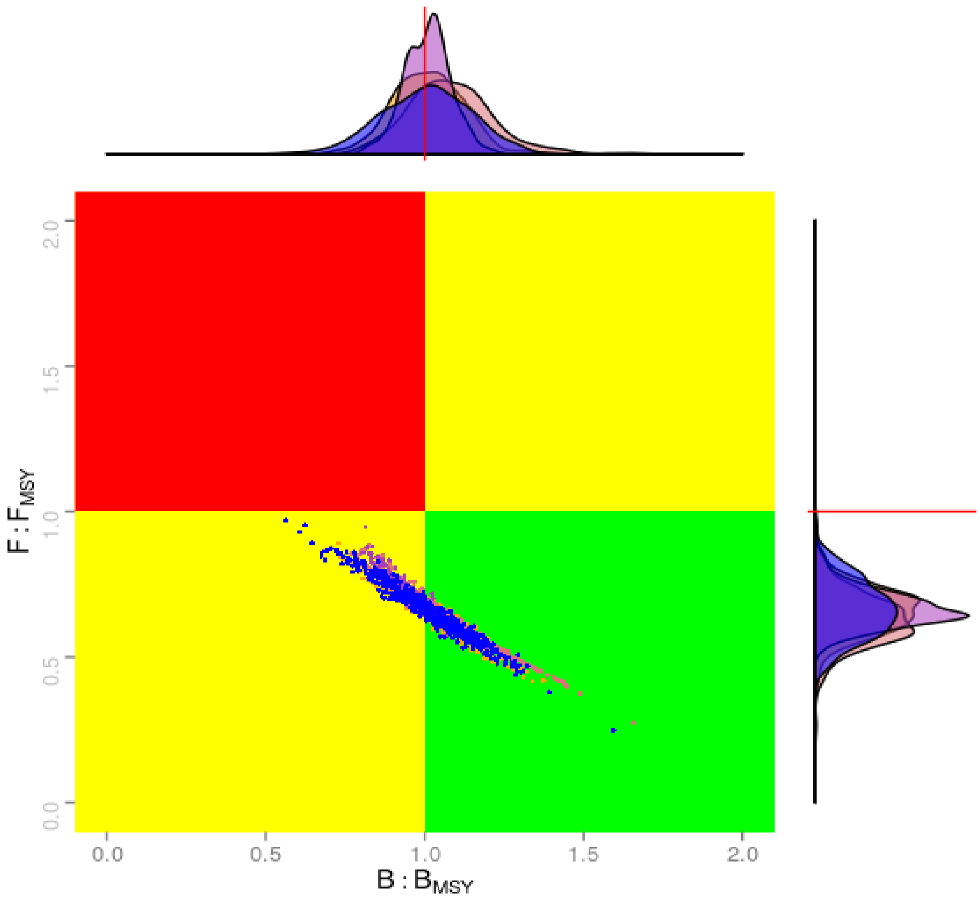

Figure 28. Kobe phase plots with marginal densities, derived from the covariance matrix, MCMC, bootstrap and jack-knife,