SCRS/2016/028 Collect. Vol. Sci. Pap. ICCAT, 73(4): 1377-1396 (2017)

A FIVE STEP PROCEDURE FOR CONDUCTING A STOCK ASSESSMENT; AN EXAMPLE BASED ON NORTH ATLANTIC ALBACORE USING A BIOMASS DYNAMIC MODEL

Laurence Kell[1], Haritz Arrizabalaga[2], and Paul De Bruyn[1], Gorka Merino[2], Iago Mosqueira[3], Rishi Sharma[4] and Jose-Maria Ortiz de Urbina[5]

SUMMARY

Diagnostics are important for evaluating the robustness of models used to estimate stock status and for understanding how uncertainties propagate through into advice. Diagnostics also make the stock assessment process more transparent and help to identify where more knowledge and better data are required. Here we adopt a generic strategy to conduct a preliminary stock assessment for North Atlantic albacore, based on five steps. The steps are to i) agree in advance hypotheses to test; ii) check for convergence; iii) identify violation of assumptions; iv) use simulation methods such as the jack knife or bootstrap to identify problems with the data and model specifications; and v) conduct hindcasting to evaluate prediction ability.

RÉSUMÉ

Des diagnostics sont importants pour évaluer la solidité des modèles utilisés pour estimer l'état des stocks et pour comprendre la façon dont les incertitudes se propagent dans l'avis. Les diagnostics rendent également le processus d'évaluation des stocks plus transparent et contribuent à identifier les domaines qui nécessitent plus de connaissances et de meilleures données. Une stratégie générique a été adoptée dans le présent cas pour réaliser une évaluation préliminaire du stock de germon de l'Atlantique Nord en suivant cinq étapes. Ces étapes étaient les suivantes : i) convenir à l'avance des hypothèses à tester ; ii) vérifier la convergence ; iii) identifier la non-application des hypothèses ; 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 rétrospectives pour évaluer la capacité prédictive.

RESUMEN

Los diagnósticos son importantes para evaluar la robustez de los modelos utilizados para estimar el estado del stock y para comprender el modo en que las incertidumbres se propagan a través del 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. Para este documento se ha adoptado una estrategia genérica para realizar una evaluación preliminar del stock de atún blanco del Atlántico norte, basada en cinco pasos. Estos pasos son: i) acordar previamente la hipótesis que se tiene que probar, ii) comprobar la convergencia; iii) identificar los supuestos que no se cumplen; iv) utilizar métodos de simulación como jack knife o bootstrap para identificar problemas con los datos y especificaciones del modelo y v) realizar una simulación retrospectiva para evaluar la capacidad predictiva.

KEYWORDS

Albacore, Biomass Dynamic, 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.
[2] AZTI -Tecnalia, Herrera Kaia Portualdea, 20110, Pasaia, Spain
[3] European Commission Joint Research Centre (JRC), Institute for the Protection and Security of the Citizen (IPSC), Maritime Affairs Unit, Via E. Fermi 2749, 21027, Ispra, VA, Italy
[4] U.S. National Marine Fisheries Service, Southeast Fisheries Center, Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, FL, 33149-1099, USA
[5] Instituto Español de Oceanografía (IEO- CO Málaga).

Introduction

Diagnostics are important for evaluating stock assessment fits, the robustness of estimates of stock status and how uncertainties propagate through into advice. Diagnostics also make the stock assessment process more transparent and help to identify where more knowledge and better data are required. Here we adopt a generic strategy to conduct a preliminary stock assessment for North Atlantic albacore, using five steps. Where the steps are to i) agree in advance hypotheses to test; ii) check for convergence; iii) identify violation of assumptions; iv) use simulation methods such as the jack knife or bootstrap to identify problems with the data and model specifications; and v) conduct hindcasting to evaluate prediction ability.

Material and Methods

Although the stock assessment method employed, based on a biomass dynamic model, is relatively simple the same procedure can be used for a wide range of models.

Material

The CPUE Series used as indices of relative stock abundance in the 2013 biomass dynamic stock assessment are shown in **Figure 1**; points are the standardised values, lines the prediction from a GAM fitted either to all the indices with year as a smoothing term and index as a factor (red) and by index individually (blue).

Methods

In a biomass dynamic model the stock in thenext year (B_{t+1}) is the sum of the current biomass (B_t) less the catch (C_t) plus the surplus production (P_t) i.e.

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

There are various forms of production functions (Pt), e.g. the symmetric logistic (Hassell, 1975) or the generalised Pella and Tomlinson (1969) forms. The logistic production function is probably not appropriate for tuna species, due to high steepness (Maunder, 2003) and a Pella-Tomlinson form with $B_{MSY} < 0.5B_0$ is perhaps more realistic, e.g.

$$\frac{r}{p} \cdot \; B(1 - (\frac{B}{K})^p)$$

where (r) is the intrinsic rate of increase, (K) the carry capacity (p) the shape of the surplus production function. If p < 1 then the curve is skewed to the left.

Diagnostic Steps

Scenarios

The first step is to agree the hypotheses to test. In a stock assessment, hypotheses are represented by scenarios where a scenario is a possible or plausible, internally consistent, but not necessarily probable historic and future developments of a stock.

The SCRS normally defines a Base Case, i.e. a set of model assumptions and specifications assumed to be the most likely. Then to represent the main uncertainties a range of scenarios are considered. In the case of a biomass dynamic assessment, scenarios are normally limited to the choice of CPUE and the values of fixed parameters. In the former case this is because series are often conflicting (e.g. negative correlations are seen between series), and in the later because there is seldom sufficient information in the data to estimate parameters such as the shape of the production function.

The CPUE series are shown in **Figure 1** and the correlations between these series are shown in **Figure 2**; blue indicates positive and red negative correlations. The order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities. Two clusters were seen, i.e. the Japanese longline and the split Chinese-Taipei indices and Combined Troll. Three scenarios were chosen, 1) Chinese-Taipei; 2) Japanese longline; and 3) Troll. The cross correlations between indices, are shown in **Figure 3** to identify potential lags due to year-class and age effects.

Convergence

Before examining diagnostics a check should be made that the solution has been found (SCRS/2016/027). Therefore likelihood profiles were calculated for each scenario by fixing carrying capacity (K) at a range of values and then estimating the other parameters.

Residual Patterns

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

Simulation

Simulation methods such as the jack knife or bootstrap identify problems with the data and model specifications; and after fitting these methods 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 (B0), 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.

Prediction Skill

A major uncertainty in stock assessment is the difference between models and reality. The validation of model predictions is difficult, however, as fish stocks can rarely be observed and counted. We therefore conducted a hindcast to prediction skill. In a hindcast, a model is fitted to the first part of a time series and then projected over the period omitted in the original fit. Prediction skill can then be evaluated by comparing the projection predictions with observations (Kell *et al.* accepted).

Results

The fits obtained using the biodyn algorithm is shown in **Figure 4**, and the residuals from the CPUE in **Figure 5**. To check convergence and the information content of the data **Figure 6** plots the sums of squares by scenario and data series (i.e. component). The observed values are plotted against the fitted values to check whether the CPUE series are a good index of relative abundance (**Figure 7,** blue line is a linear regression fitted to points, black the y=x line. Next the residuals are plotted to check for violation of model assumptions. The residuals are plotted by year in **Figure 8** to check patterns in time that may suggest the index is not tracking the stock.

Figure 9 shows quantile-quantile plots to compare residual distribution with the normal distribution. **Figure 10** plots the residuals against fitted value with to check the variance relationship. Next **Figure 11** plots the residuals with a lag of one against each other to check for autocorrelation, and finally the predicted stock trend by index (points), with biomass estimates (blue) and a local regression (black) are presented in **Figure 12**. The production functions and time series of yield v stock biomass are presented in **Figure 13.** Next a jack knife is conducted to evaluate the influence of individual points and potential bias; **Figure 14** shows the jack knife of r and **Figure 15** the jack knife of K. The hindcasts are presented in **Figures 16, 17, 18, 19 and 20**; it can be seen that the predictive ability of the Scenario 1 is good while that of Scenario 2 is poor.

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.

A proposed strategy for running the 2016 assessment is

1. Use the Pella-Tomlinson production function, since the logistic production model is probably not appropriate for tunas. Since there is seldom sufficient information, however, in stock assessment data sets to estimate the shape parameter. Therefore life history relationships could be used to estimate the ratio between BM SY and K.
2. Choose CPUEs to run as separate scenarios based on hypotheses that could eventually be tested.
3. Check robustness of the fits using jackknife/bootstrap.
4. Conduct a hindcast to test prediction skill.

References

M. Hassell. Density-dependence in single-species populations. The Journal of animal ecology, pages 283-95, 1975.

L. Kell, P. De Bruyn, M. Maunder, K. Piner, and I. Taylor. 2014. Likelihood component profiling as a data exploratory tool. ICCAT Collect. Vol. Sci. Pap., 70(3): 1288-1293.

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

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

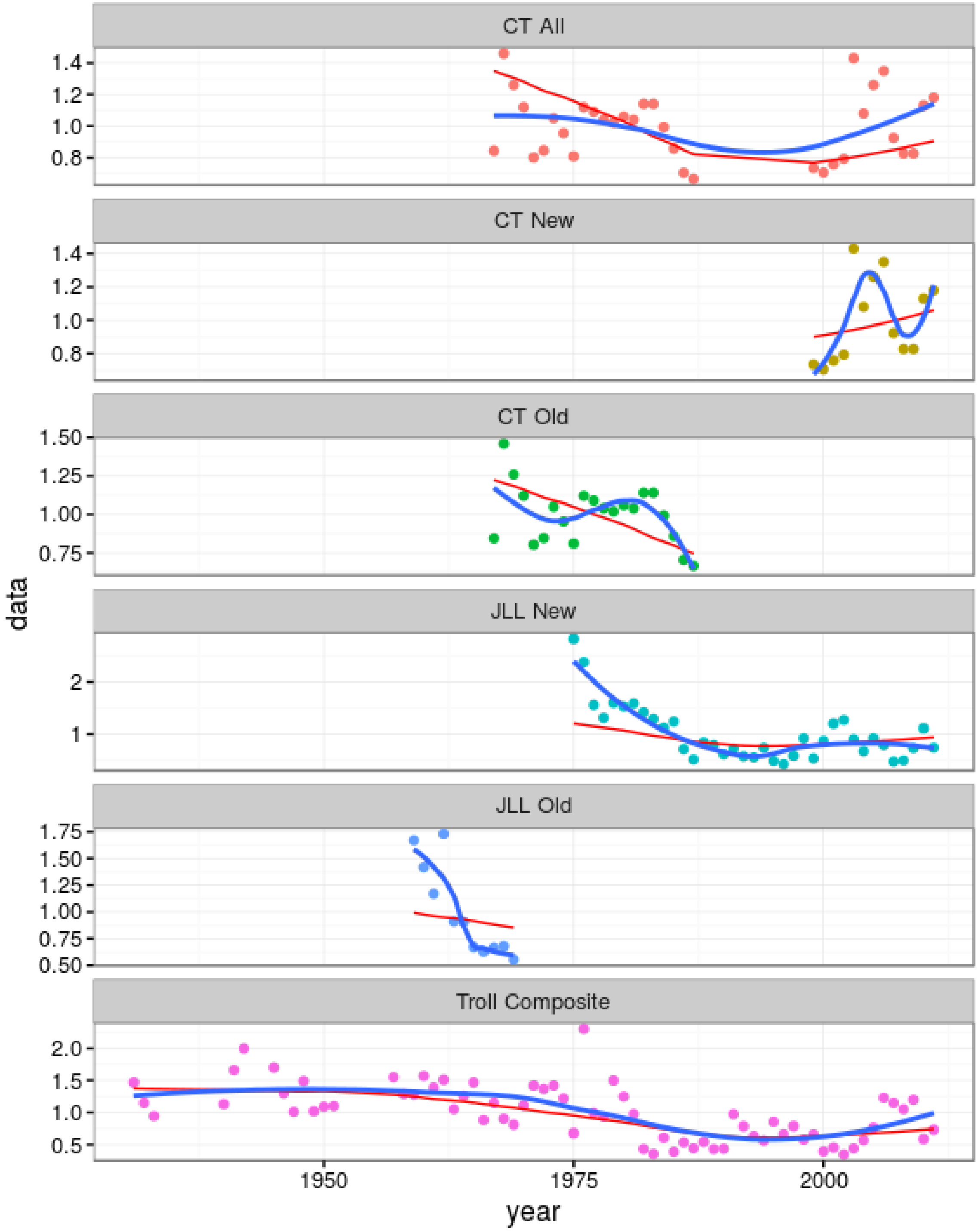
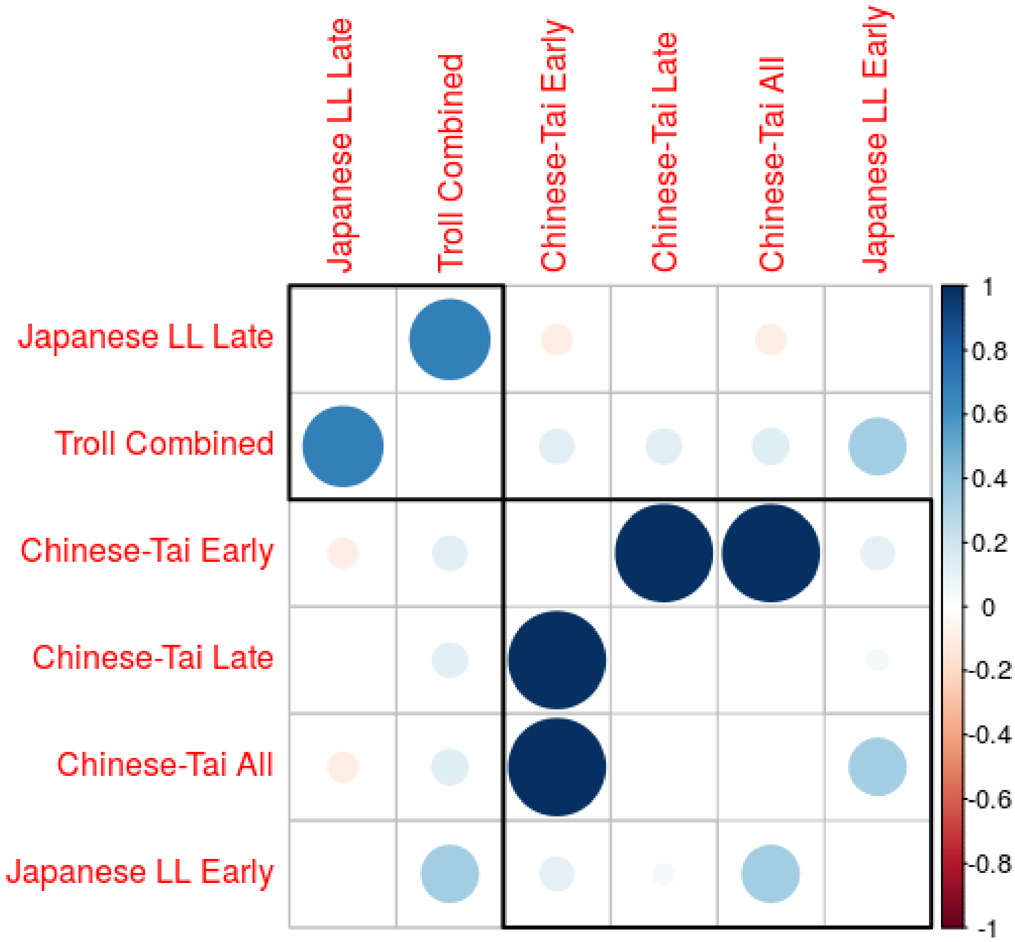


Figure 1. CPUE Series used in the 2013 biomass dynamic stock assessment as potential proxies for stock abundance; points are the standardised values, lines the prediction from a GAM fitted either to all the indices with year as a smooth term and index as a factor (red) and by index individually (blue).



NULL

Figure 2. Correlation matrix for the 2013 indices; blue indicate positive and red negative correlations, the order of the indices and the rectangular boxes are chosen based on a hierarchical cluster analysis using a set of dissimilarities.

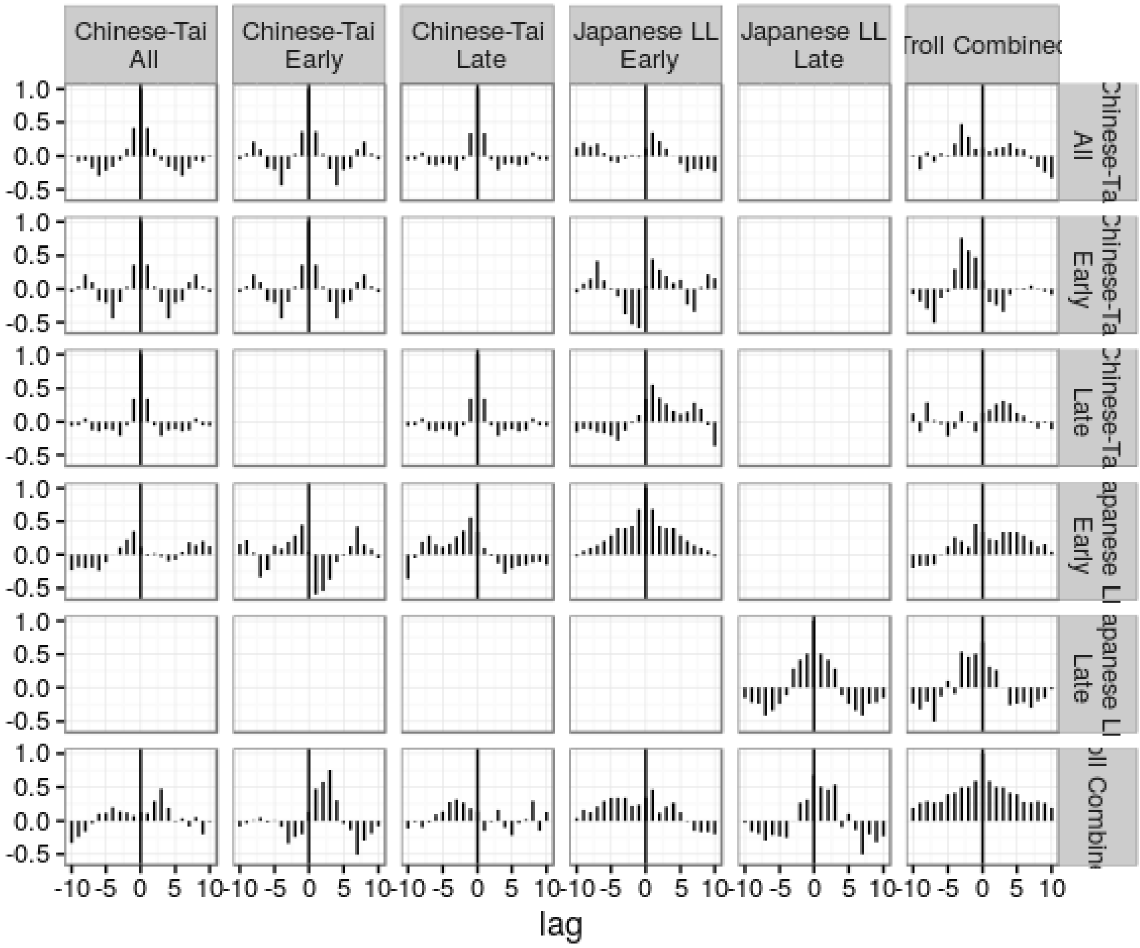


Figure 3. Cross correlations between indices, to identify potential lags due to year-class effects.

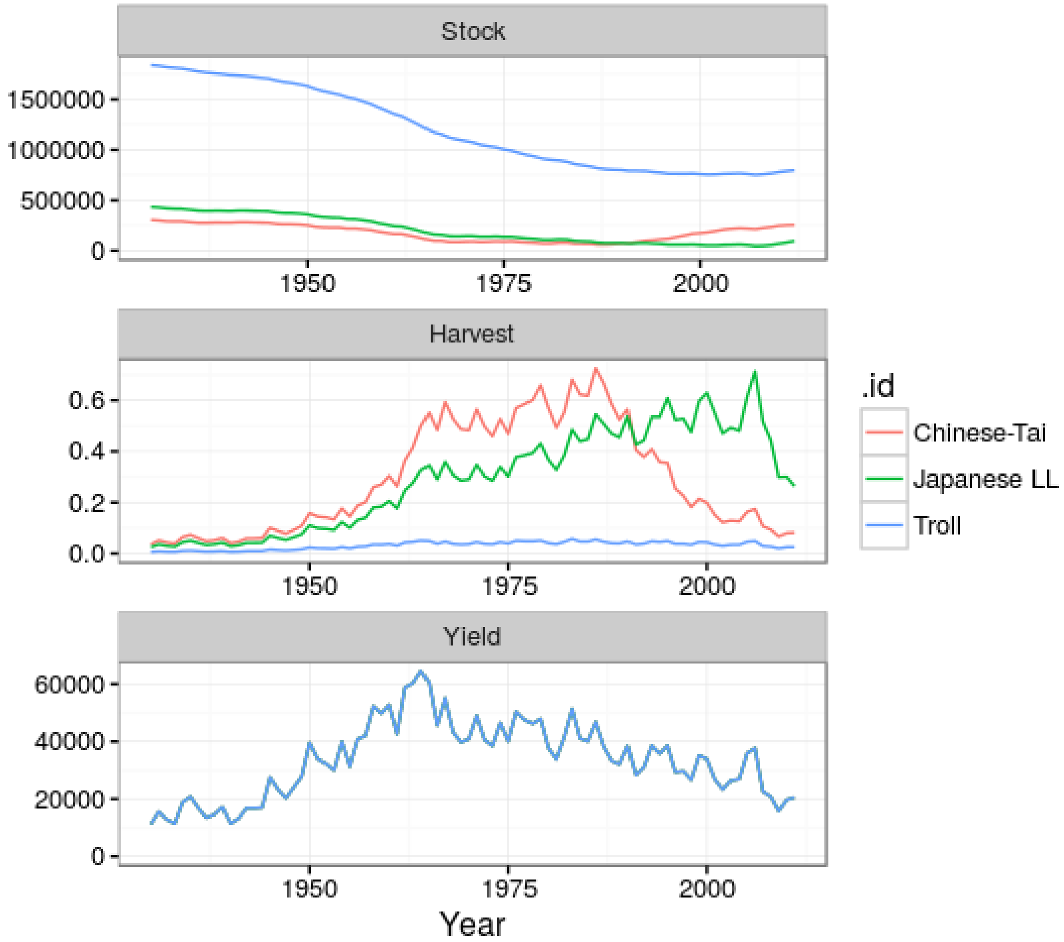


Figure 4. Fits using the biodyn algorithm.

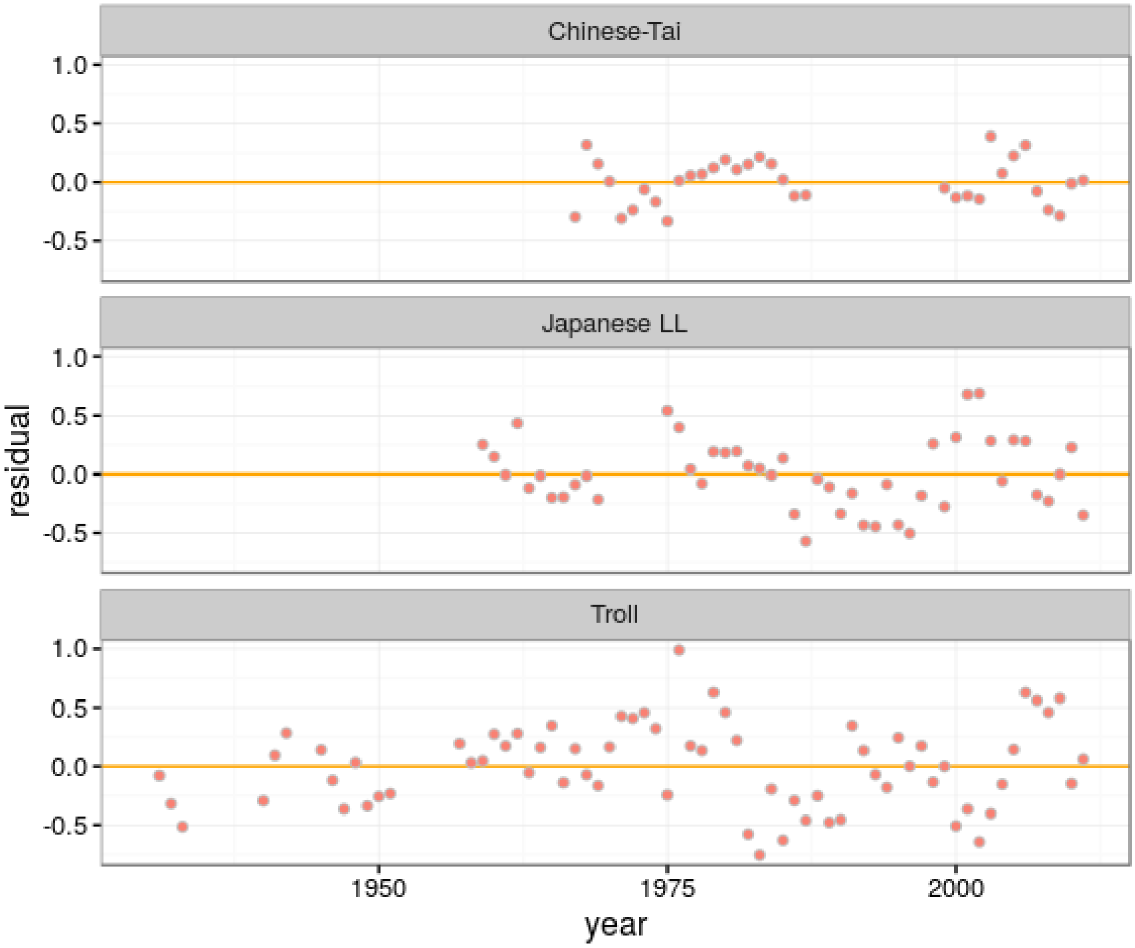


Figure 5. CPUE residuals from biodyn fits.

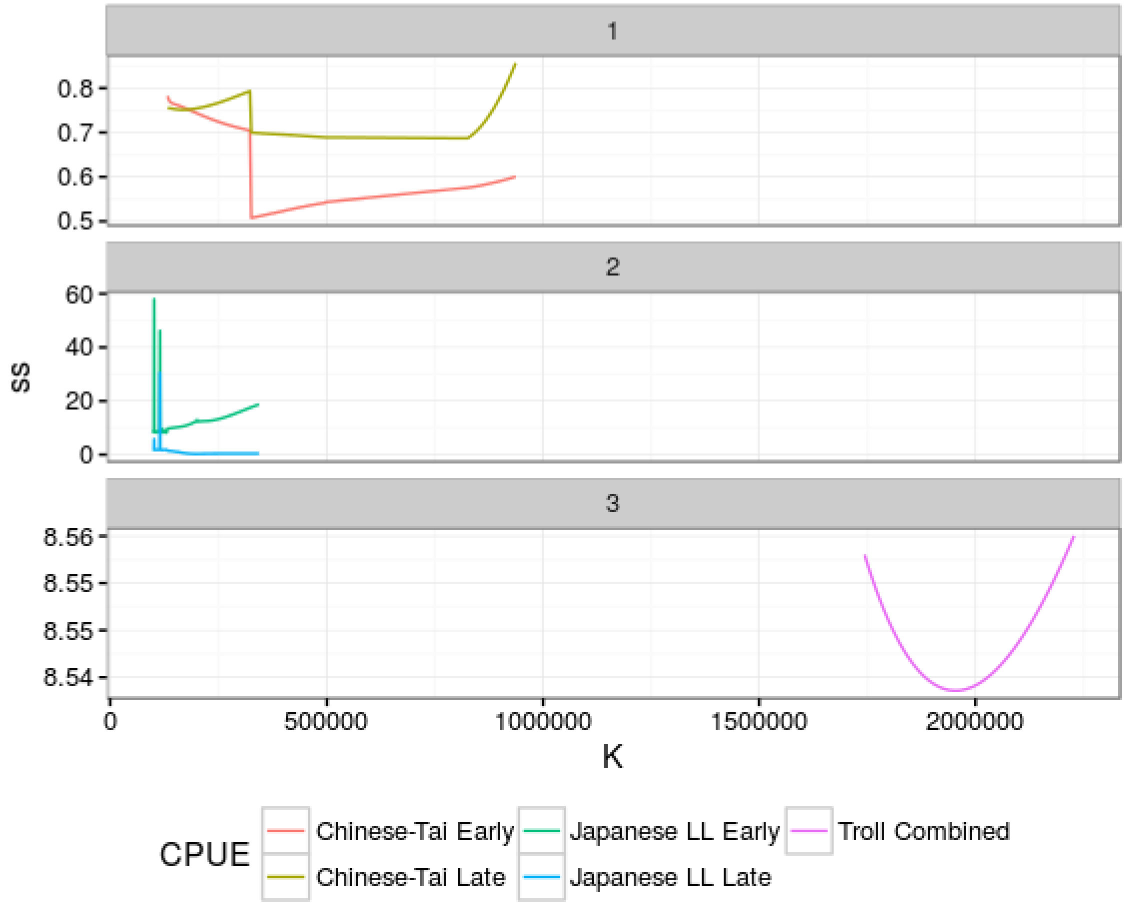


Figure 6. Profiles of sums of squares by stock assessment scenario and data component.

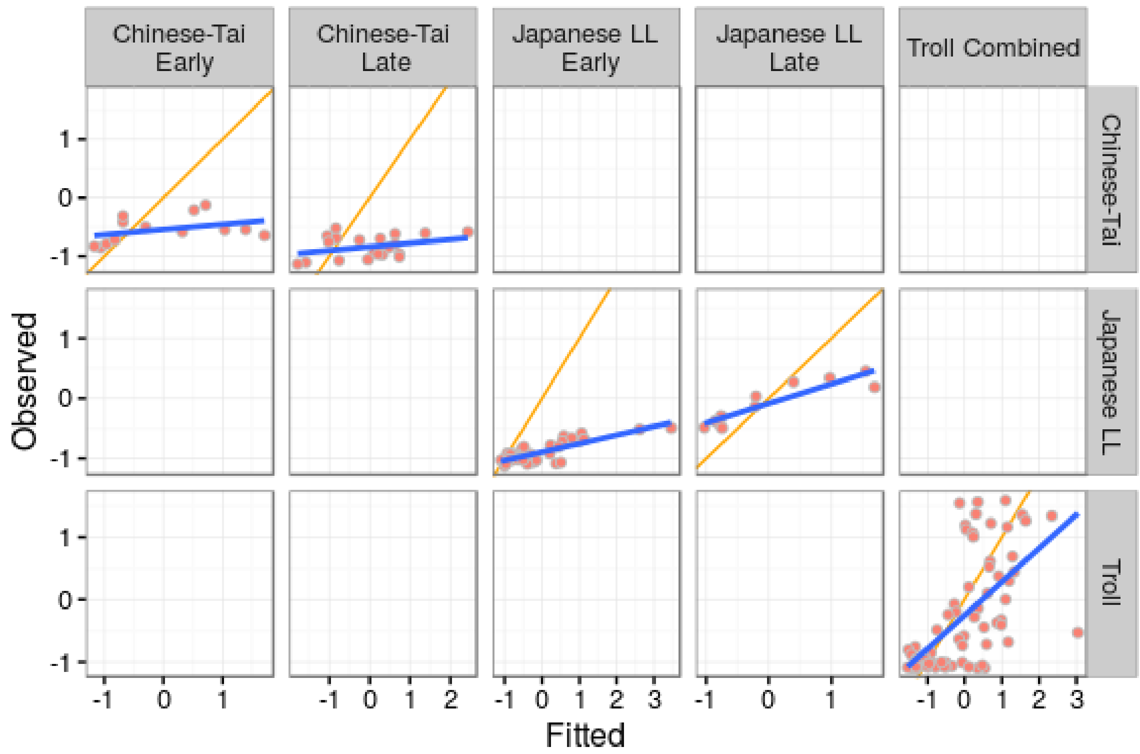


Figure 7. Observed CPUE verses fitted, blue line is a linear regression fitted to points, black the y=x line.

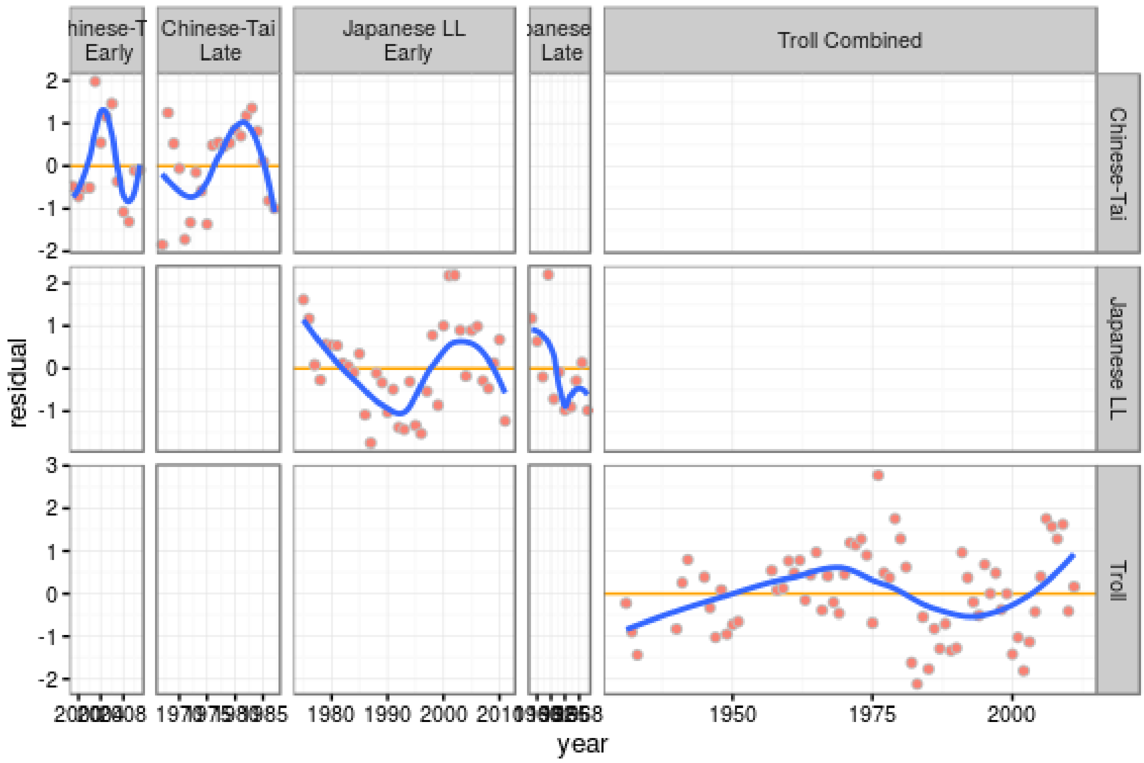


Figure 8. Residuals by year, with lowess smoother (blue).

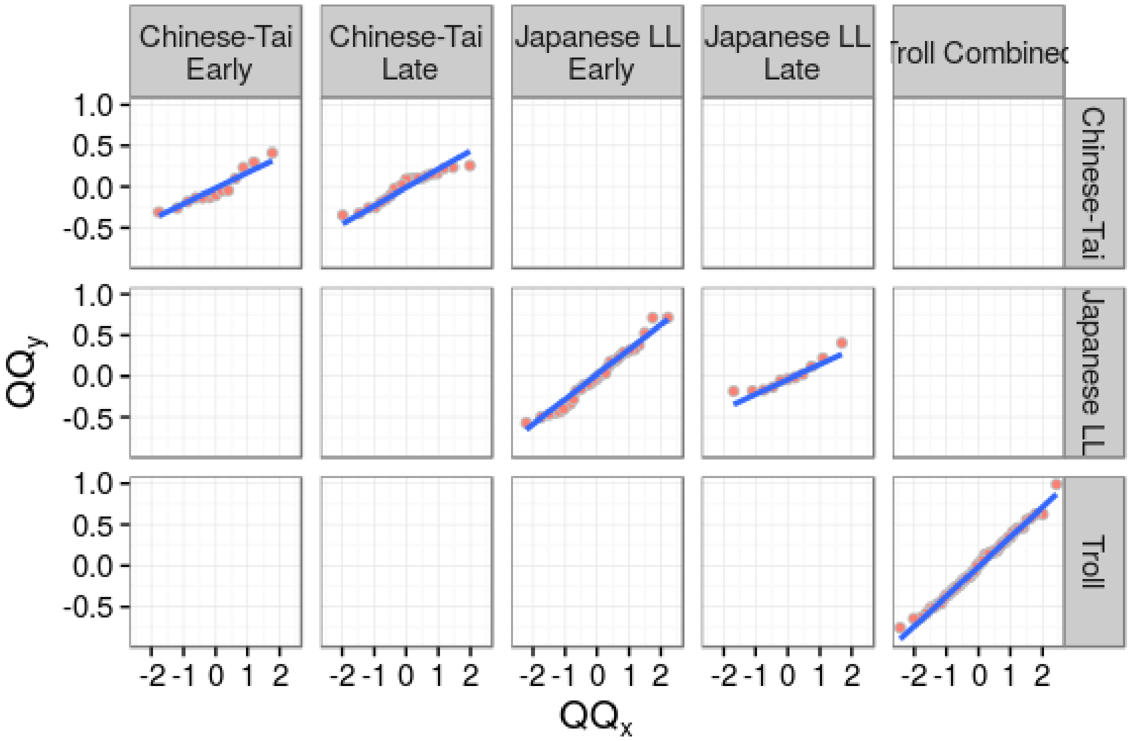


Figure 9. Quantile-quantile plot to compare residual distribution with the normal distribution.

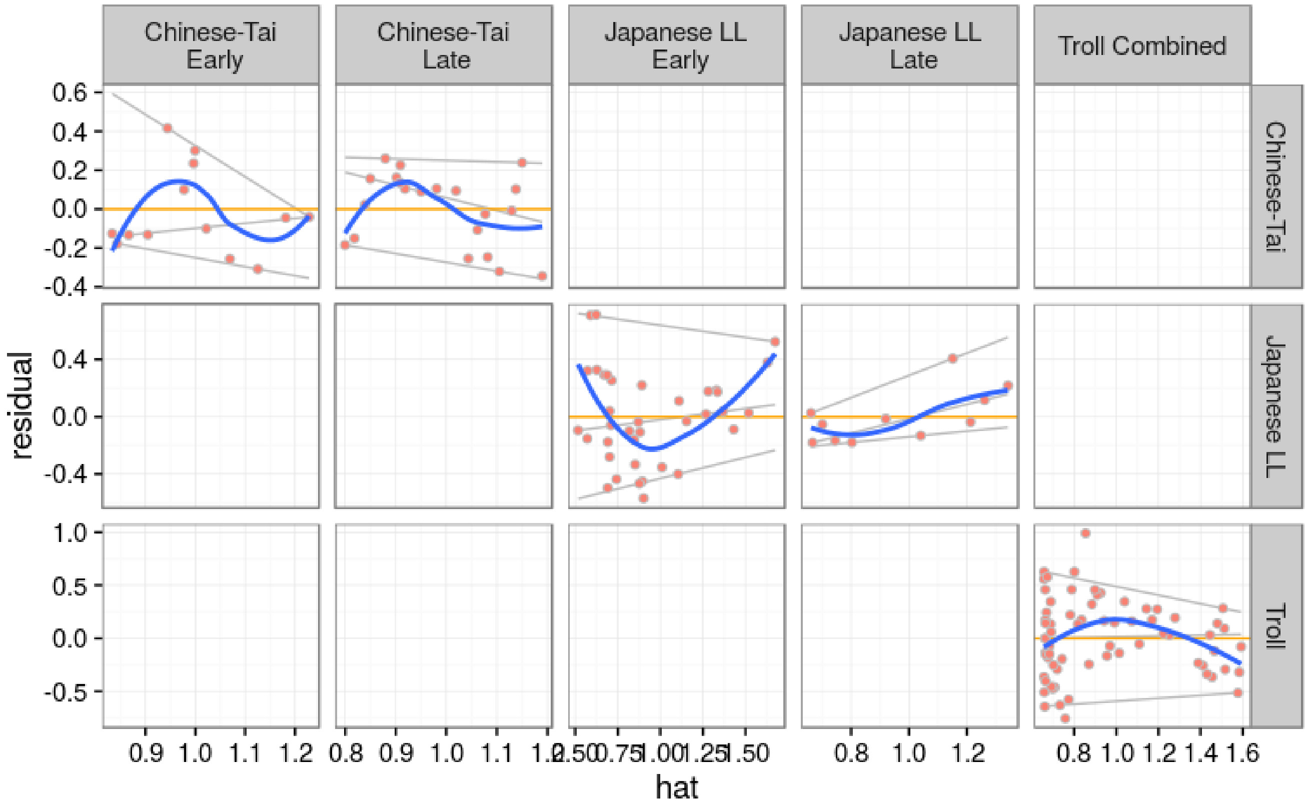


Figure 10. Plot of residuals against fitted value with 5^{th} & 95^{th} percentiles to check variance relationship.

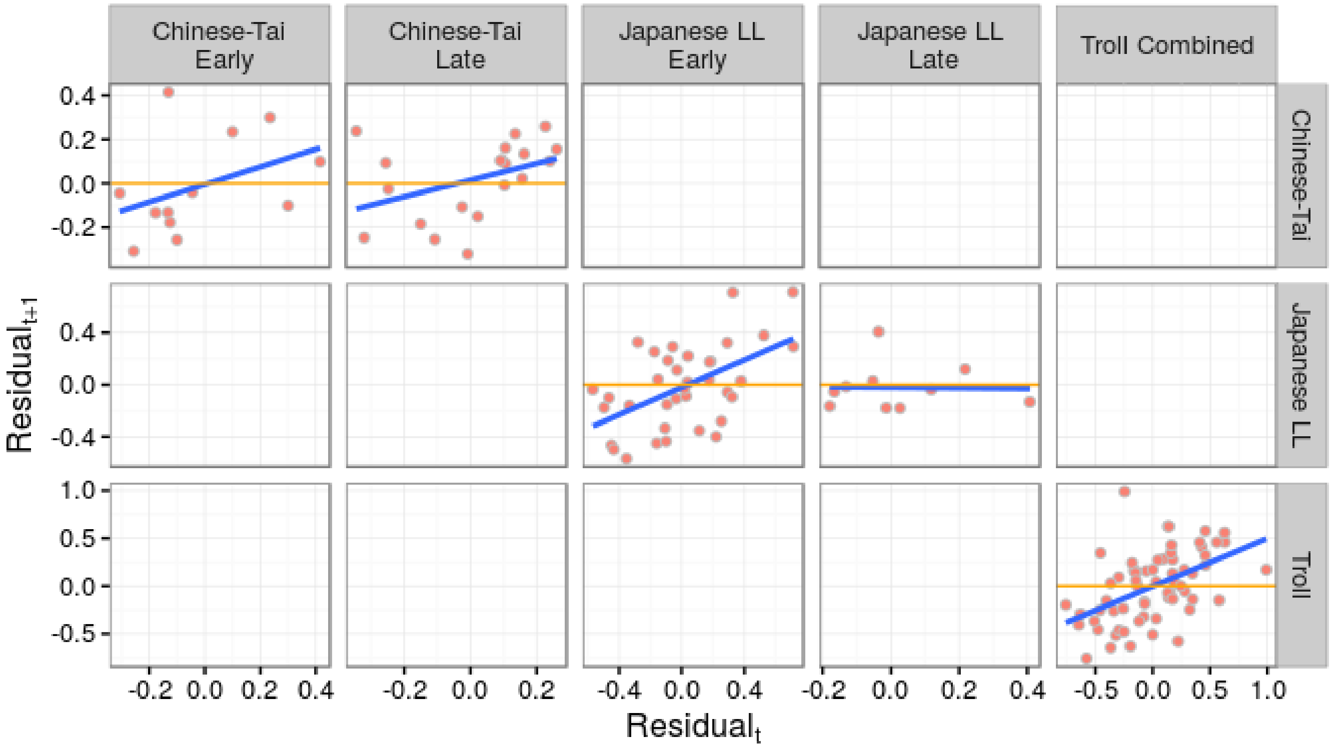


Figure 11. Plot of autocorrelation, i.e. $residual_{t+1}$ verses $residual_t$.

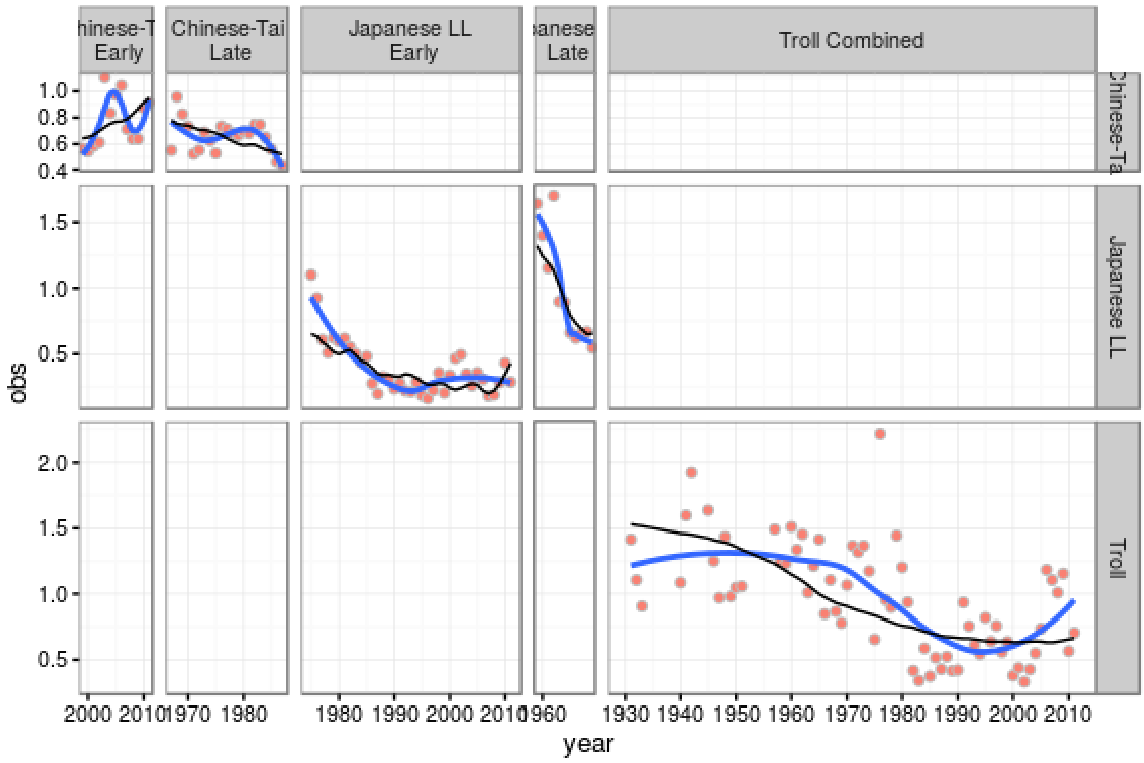


Figure 12. Predicted stock trend by index (points), with biomass estimates (blue) and a local regression (black).

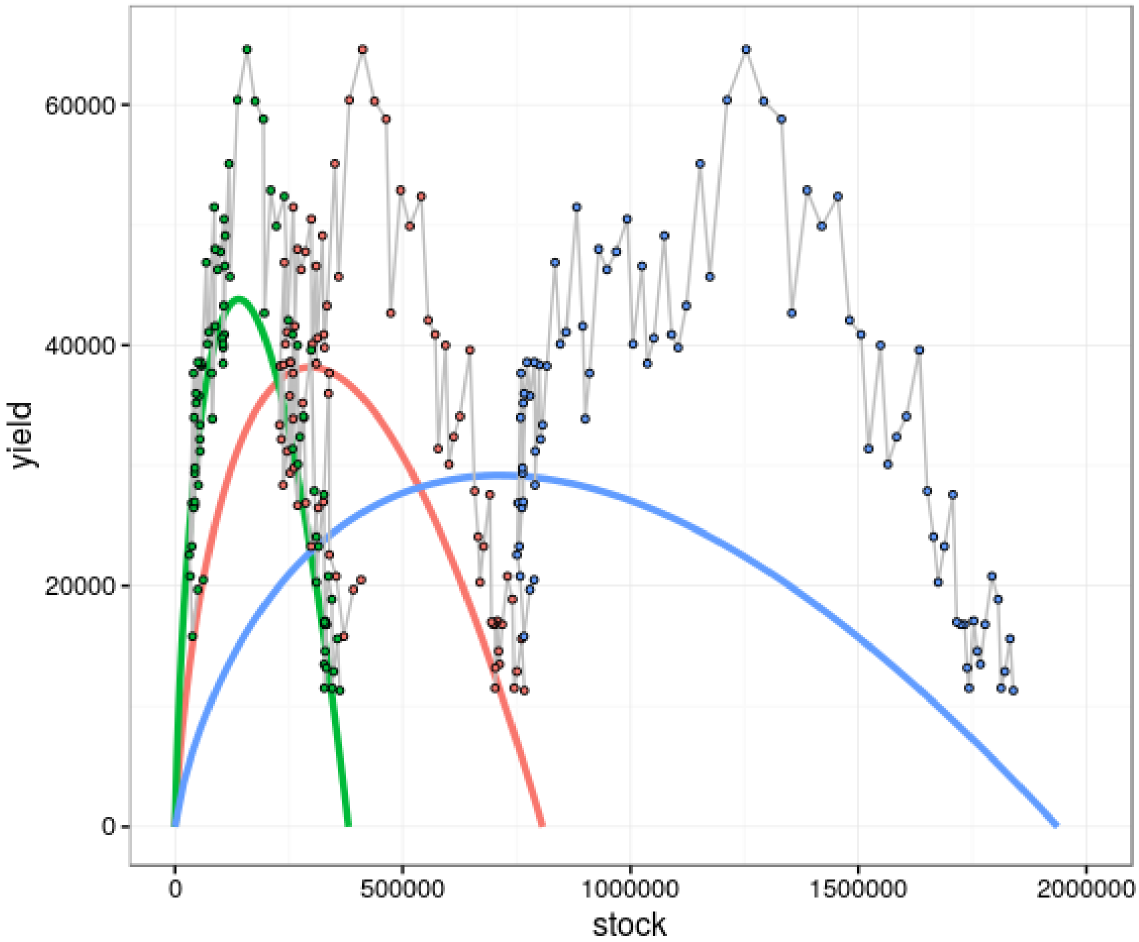


Figure 13. Production functions and time series of yield v stock biomass.

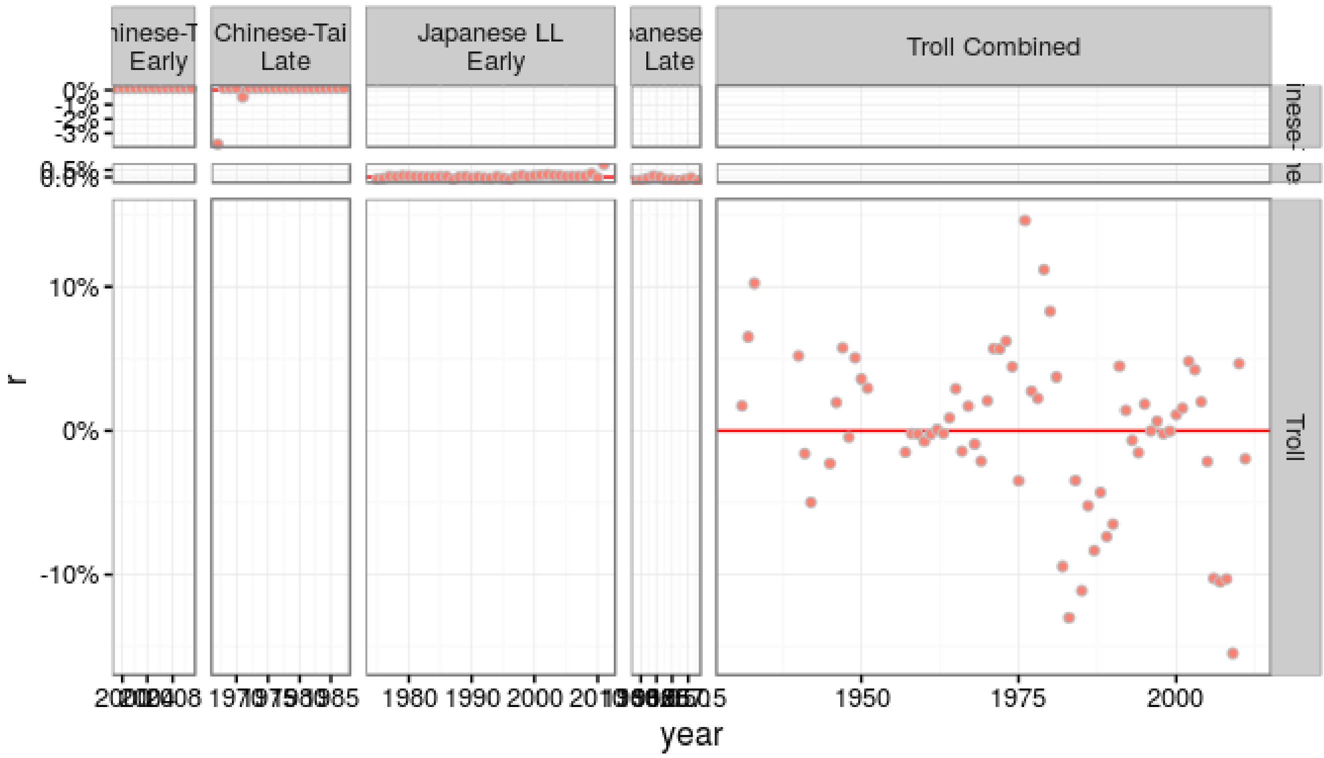


Figure 14. Jackknife of r to check for influence of individual CPUE points.

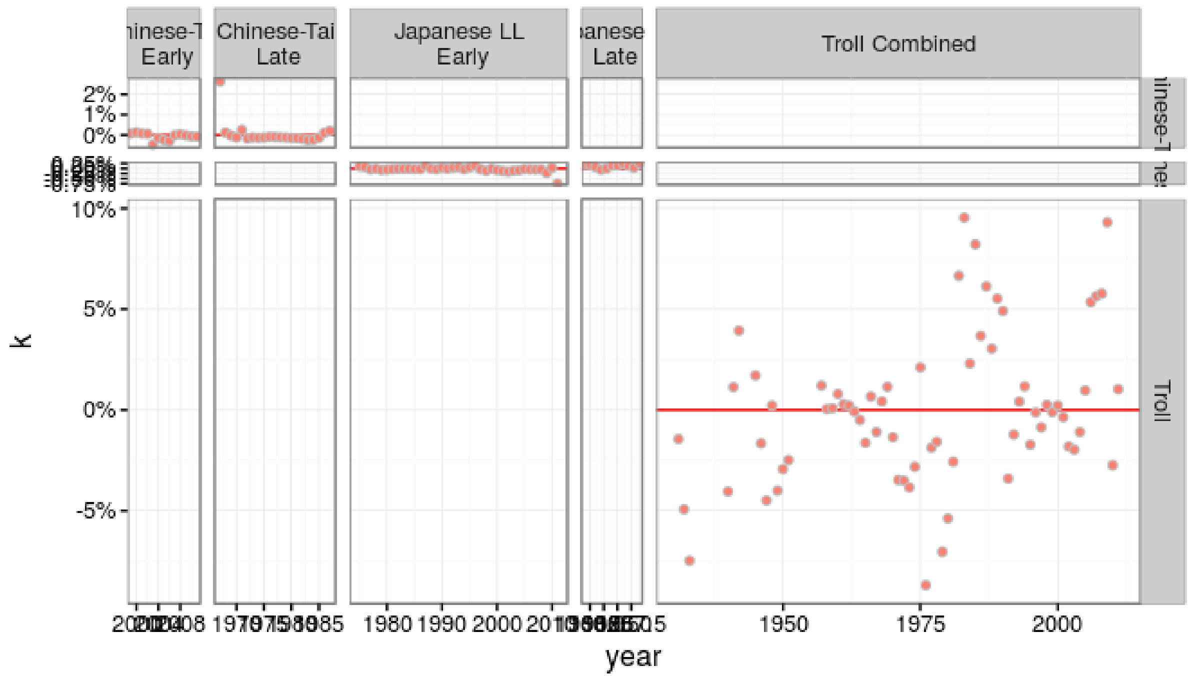


Figure 15. Jackknife of K to check for influence of individual CPUE points.

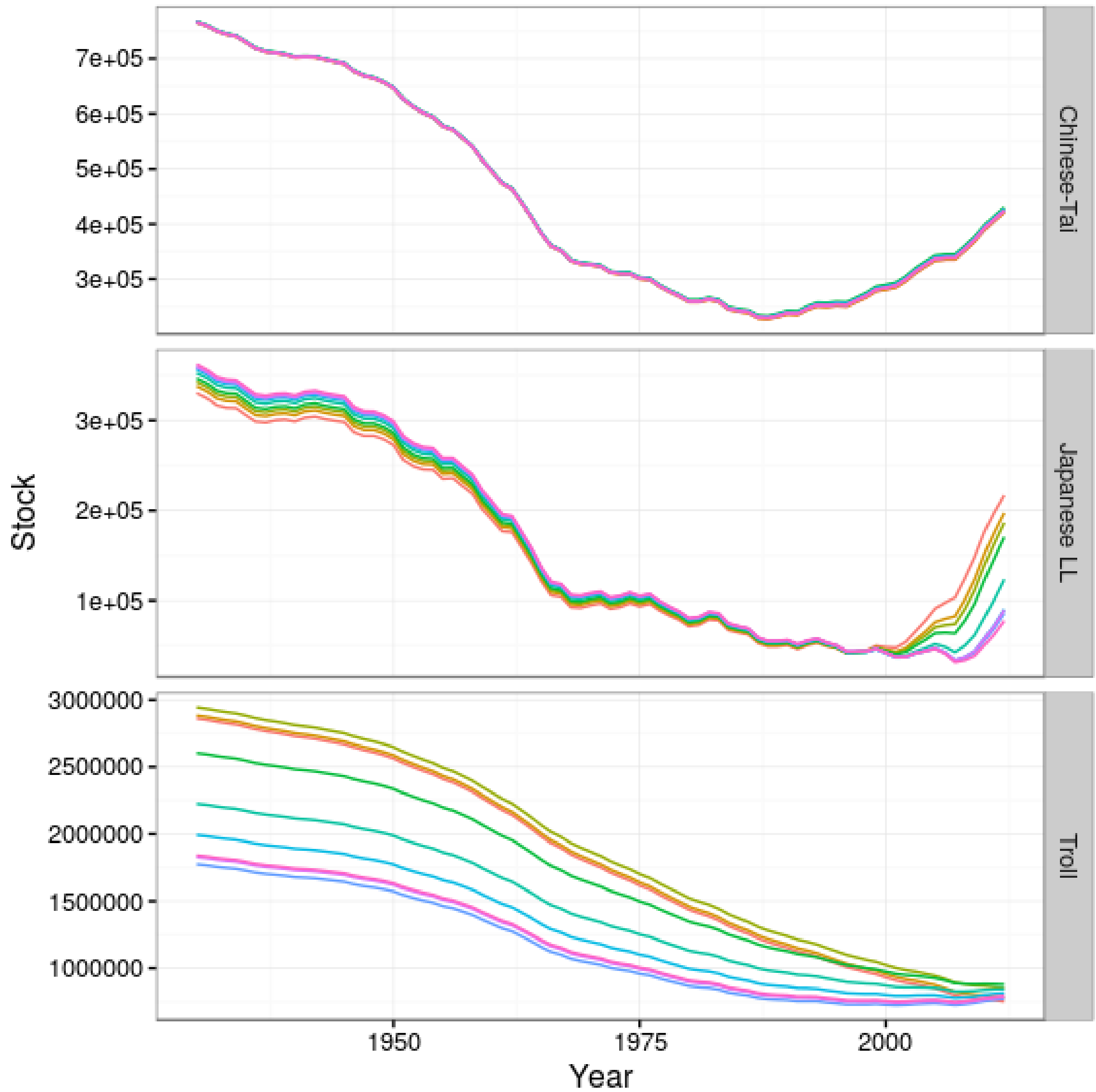


Figure 16. Hindcast of stock.

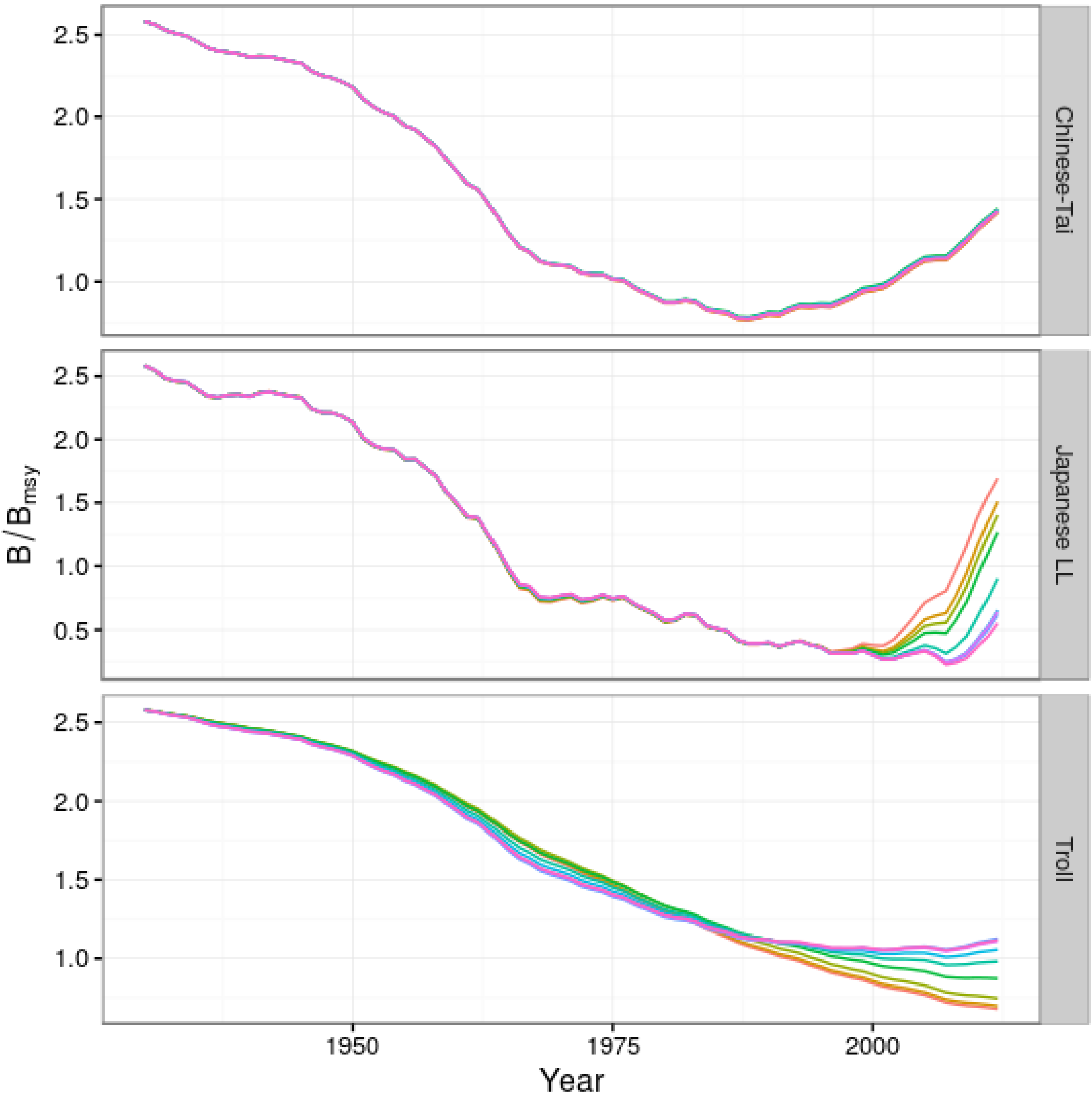


Figure 17. Hindcast of stock relative to B_{MSY}

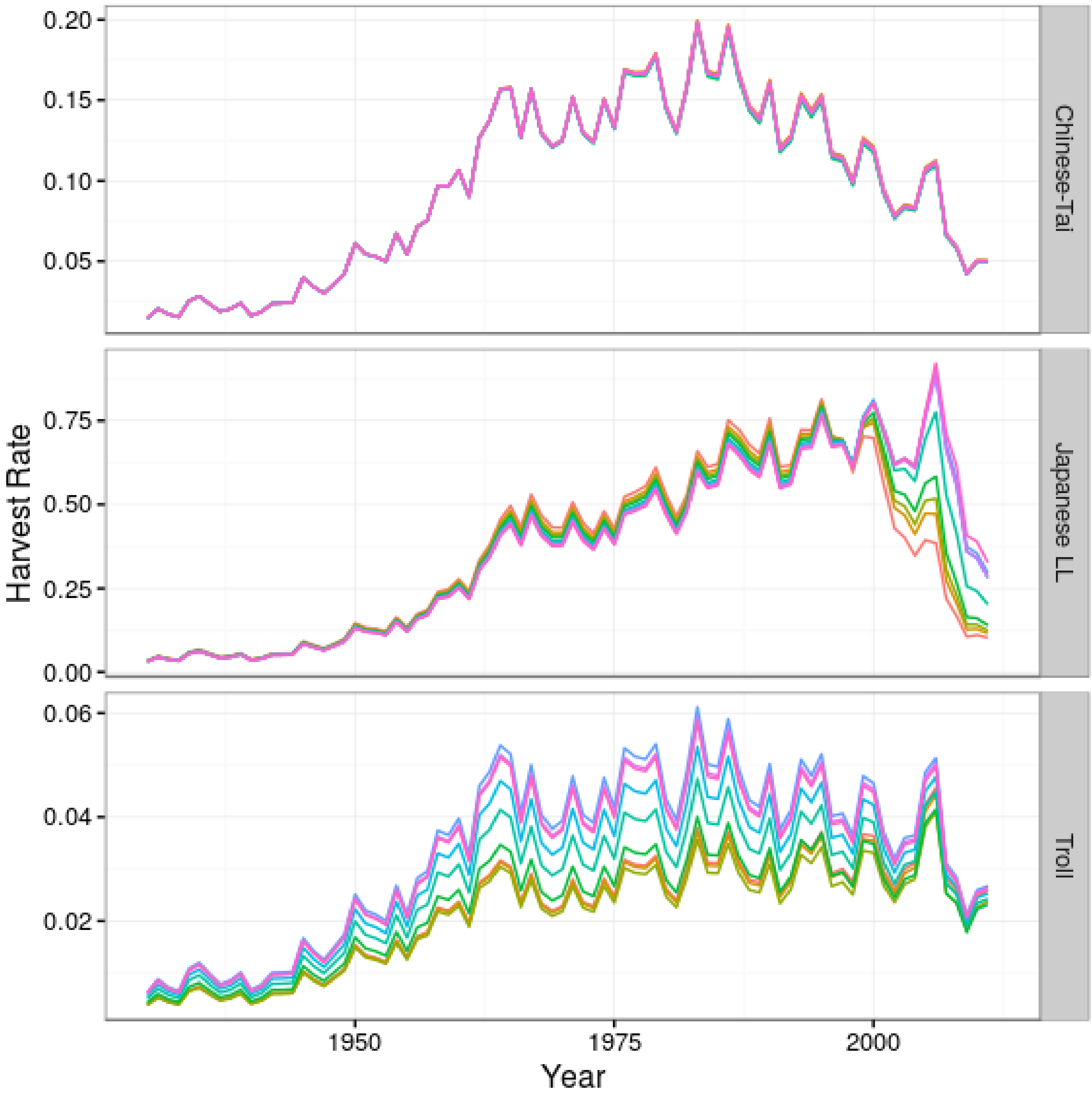


Figure 18. Hindcast of harvest rate.

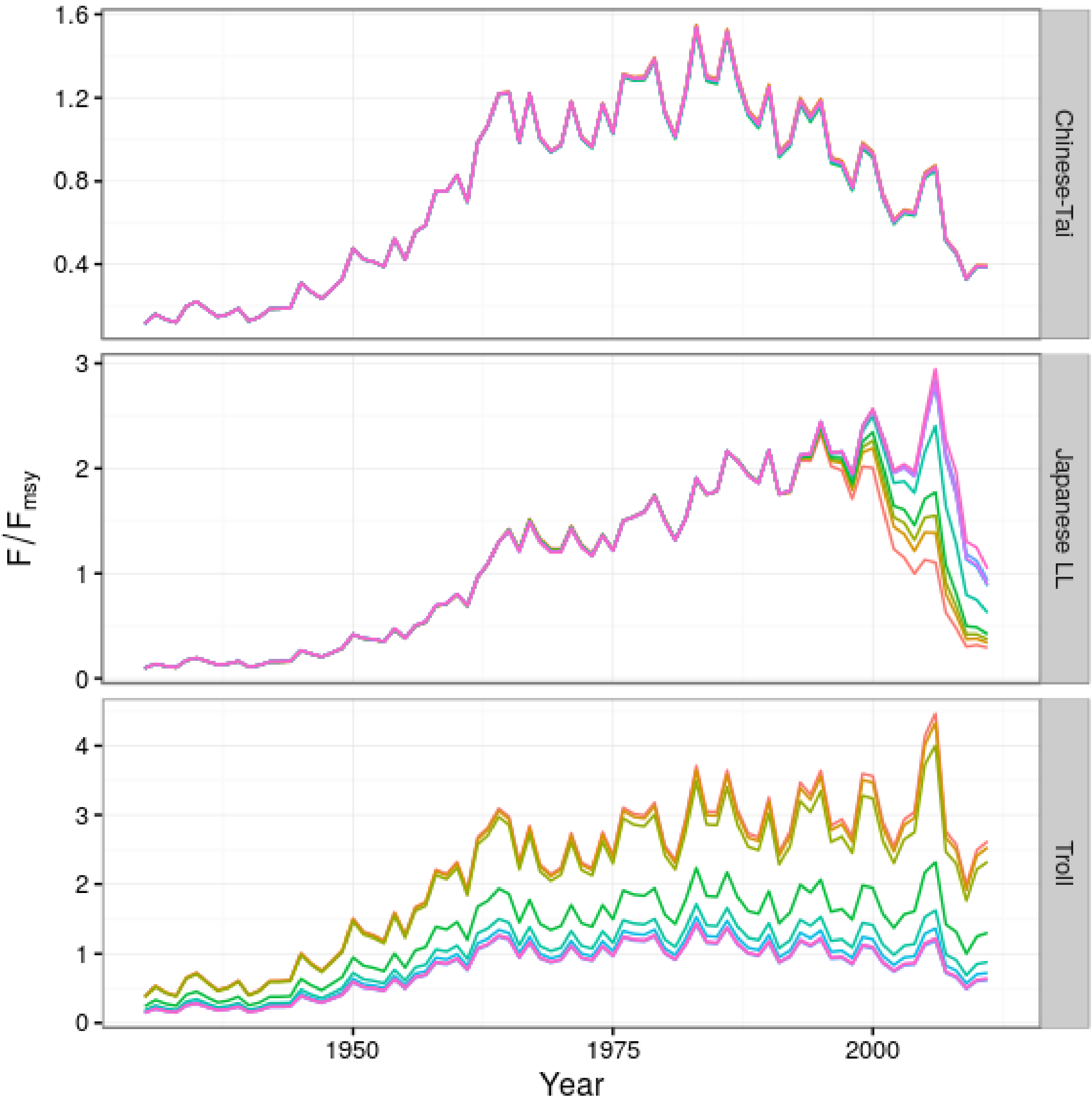


Figure 19. Hindcast of harvest rate relative to F_{MSY}

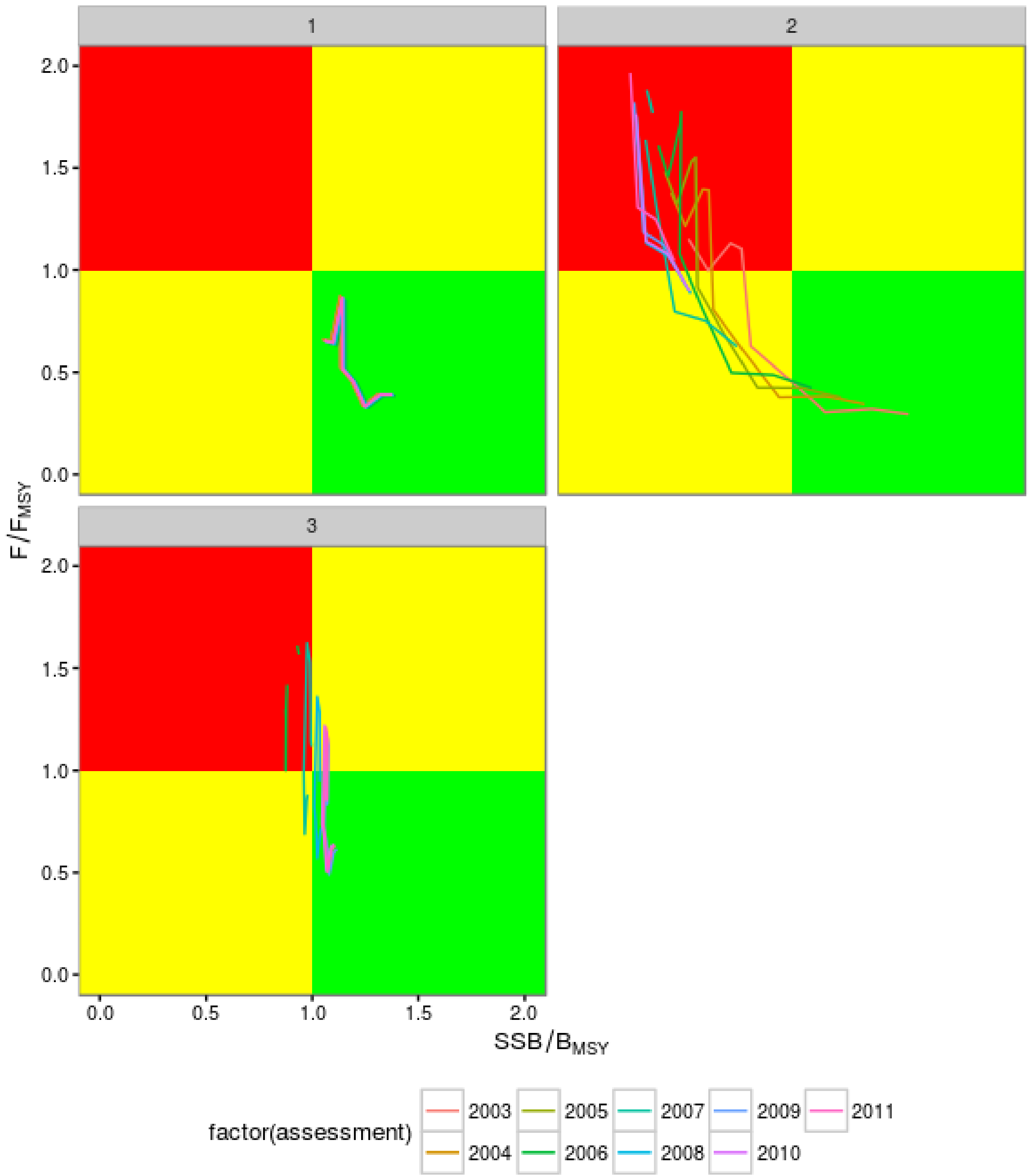


Figure 20. Hindcast in kobe phase plot form.