Fisheries New Zealand

Tini a Tangaroa

Evaluating the effects of changes in the frequency of research trawl surveys and age sampling on the hoki, hake, and ling stock assessments

New Zealand Fisheries Assessment Report 2018/43

C. Marsh,
A. McKenzie,
R.I.C.C. Francis,
I Doonan

ISSN 1179-5352 (online)
ISBN 978-1-98-857118-8 (online)
November 2018

New Zealand Government

Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140
Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries websites at:
http://www.mpi.govt.nz/news-and-resources/publications
http://fs.fish.govt.nz go to Document library/Research reports

CONTENTS

.

EXECUTIVE SUMMARY

Marsh, C.; McKenzie, A.; Francis, R.I.C.C.; Doonan, I. (2018). Evaluating the effects of changes in the frequency of research abundance trawl surveys and age frequency sampling on the hoki, hake, and ling stock assessments.

***New Zealand Fisheries Assessment Report 2018/43.* 60 p.**

Under the Ministry for Primary Industries 10-Year Research Programme, the frequency of the Chatham Rise and Sub-Antarctic trawl surveys has been reduced from annual to biennial to ease pressure on fisheries research budgets, and to focus on priorities for other New Zealand fisheries. This change has the potential to affect the robustness of, and uncertainty associated, with biomass estimates for the hoki, hake, and ling stock assessments, as well as the future assessments of species that are not currently assessed. The effect of the reduction in trawl surveys is investigated in this report using three types of computer simulation: (a) generic simulations, (b) retrospective analyses, and (c) forward projection simulations. The retrospective analyses and forward projections simulations were conducted for the hoki, hake, and ling stock assessments in the Chatham Rise and Sub-Antarctic.

The generic simulations demonstrated that, under very simple scenarios, when moving from annual to biennial surveys we expect to get less precision, but no bias, in estimated biomass changes. For the retrospective simulations, there were small differences in current biomass estimates (%B_0) between annual and biennial survey scenarios (both accuracy and precision), although for the hoki stock there was a larger difference between the two different biennial scenarios, which is attributed to an induced three-year gap between surveys due to a pre-existing two-year gap.

The forward projection simulation analysis looked at the effect on biomass estimates of a survey frequency change from annual to biennial for each stock, under three alternative population trajectories (increasing, decreasing and stable). The size of the differences between annual and biennial scenario varied between stocks and population trajectories, but were small, with the annual scenario having greater accuracy and slightly higher precision than the biennial scenarios.

1. INTRODUCTION

Historically, fishery-independent surveys that indexed the abundance of hoki (and other “Tier one” deepwater species, which are regularly quantitatively assessed) were considered a high priority, especially when hoki stock size was relatively low. From the 2001 to 2010 fishing years (New Zealand fishing years run from 1 October to 30 September, but are normally referenced by the year-ending), Chatham Rise and Sub-Antarctic trawl surveys were completed annually, with the Cook Strait acoustic survey completed in all but one year from 2001 to 2009 (Ministry for Primary Industries, 2017). The Chatham Rise and Sub-Antarctic trawl survey data from these and other years have also been used in the ling and hake stock assessments, for which they inform estimates of biomass and year class strengths (Figures 1 and 2).

Under the Ministry for Primary Industries 10-Year Research Programme, the frequency of the Chatham Rise and Sub-Antarctic research trawl surveys has been reduced to ease pressure on fisheries research budgets, and to focus on priorities for other New Zealand fisheries. This change has the potential to affect the robustness of and uncertainty associated with the outputs of the hoki, hake, and ling stock assessments. Understanding the potential impacts of this change on the quality of the scientific advice will allow managers to determine the appropriate timing and frequency of future research surveys.

Computer simulation experiments were used to investigate different scenarios related to the frequency of abundance surveys and sampling for age frequency. Three different types of simulation experiments were conducted, which we briefly describe here:

1. *Generic simulations* where a biomass change over time is assumed (increasing, stable, declining). From the known biomass trajectory, survey biomass estimates are simulated, on either an annual or biennial basis. The simulated survey biomass estimates are then used to estimate the biomass change. This is not intended to replace stock assessment simulations, but as an illustration of the underlying considerations between the frequency of surveys, and the effect on accuracy of the estimated biomass changes. A simplified model is used to lessen the assessment complexities that may hinder the understanding of the underlying principle.

2. *Retrospective biomass estimations*. In these simulations, current biomass estimates and five year projections are compared using all data, and with data retrospectively removed biennially from the trawl survey observations over a ten year period. The aim of these simulations is to take the current stock assessment, and investigate the consequences if the trawl survey data used in them had been biennial instead of annual.

3. *Forward projection simulations* using current stock assessments implemented in CASAL (Bull et al. 2012). Different scenarios for assumed year class strengths and plausible future catch histories are used to obtain known projected biomasses in the future: decreasing, stable, increasing. From these known projected biomasses, survey biomass estimates are simulated on an annual or biennial survey basis, and a series of assessment biomass estimates done in the projected years. The aim of these simulations is to compare possible future scenarios of rapidly decreasing or increasing biomass, where moving from annual to biennial surveys may matter more.

The work reported here addresses objectives 1 and 2 of the Ministry for Primary Industries project DEE2015-02, with the overall objective: *To evaluate the implications of changing the frequency of abundance surveys and sampling for age frequency ageing, and stock assessments for selected Tier 1 deepwater stocks*.

The specific objectives are:

1. *To evaluate the implications of conducting abundance surveys for hoki at intervals more than annually in terms of the reliability of and uncertainty associated with stock status and population projections.*

2. *To evaluate the impact of less frequent sampling for age frequency for hake and ling stocks in terms of the reliability of and uncertainty associated with stock status and population projections.*

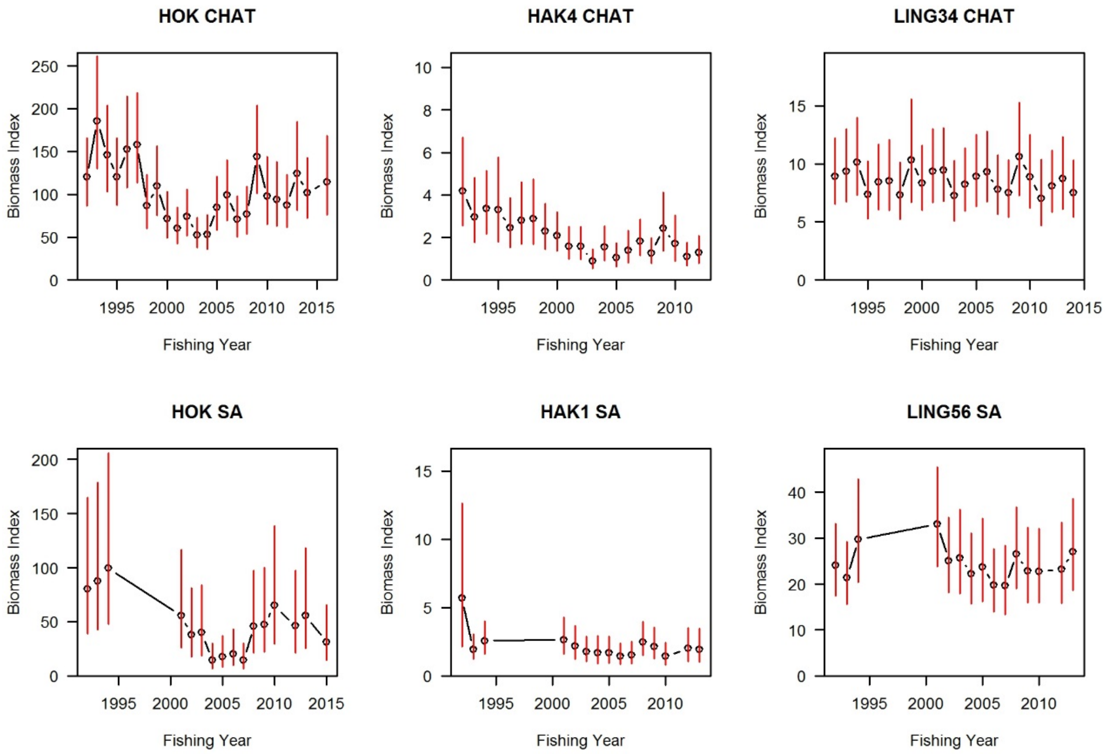

Figure 1: Biomass indices (thousands of tonnes) used in the most recent hoki, hake, and ling stock assessments, with 95% confidence intervals for the observation error. Top row is Chatham Rise trawl survey indices, and the bottom row is Sub-Antarctic trawl survey indices. The hoki stock assessment (HOK) incorporates survey indices from both the Chatham Rise trawl survey (CHAT) and Sub-Antarctic trawl survey (SA). The hake (HAK) and ling (LIN) stock assessments cover specific quota management areas (e.g. HAK 4 for area 4) or combined areas (e.g. LIN 3&4 for areas 3 and 4 combined). See Ministry for Primary Industries (2017) for further details.

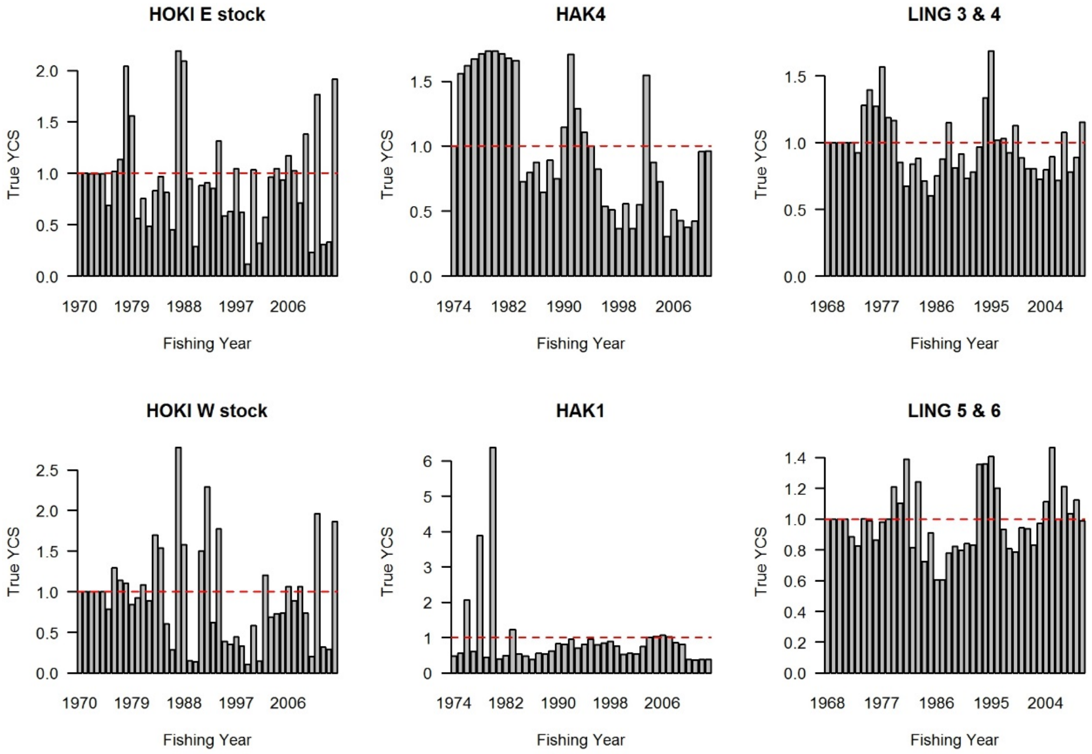

Figure 2: True year class strength estimates from the most recent hoki, hake, and ling stock assessments. Top row is for stocks in the Chatham Rise, and the bottom row is for stocks in the Sub-Antarctic. See Figure 1 for further explanation of the stock and areas titles for each graph.

2. METHODS

There were three types of simulations conducted (a) generic simulations, (b) retrospective analyses, and (c) forward projection simulations. We describe each of these in turn.

2.1 Generic simulations

These simple simulations are used to illustrate the impact of changing the frequency of biomass surveys and are similar to those of Francis & Horn (2005). Three different biomass trajectories over time are assumed (increasing, stable, declining), and from these biomass trajectories survey biomass estimates are simulated on either an annual or biennial basis. The simulated survey biomass estimates are used to re-estimate the biomass trajectory, and the difference in accuracy of estimates between annual and biennial surveys is evaluated.

The assumed biomass trajectories all start in 2011 with an initial biomass of 150 000 t and finish five years later in 2016. Three different biomass trajectories are assumed: (a) increase of 100%, (b) stable, and (c) decline of 50%. The biomass trajectories are assumed to be exponential following the equation

$$B_t = B_0 \exp(r\Delta \mathrm{t})$$

where B_0 is the biomass in the initial year, B_t is the biomass in year t, r the growth rate parameter, and Δt the number of years since 2011. The three assumed biomass trajectories correspond to rate parameter r values of 0.14, 0, and -0.14 respectively (Figure 3).

For each of the three assumed biomass trajectories we assume that we can collect a representative sample of biomass through time (i.e. survey biomass observations over time). The sampling frequencies we assumed were annual and two biennial scenarios (either starting on the initial year, Biennial 1 or second year, Biennial 2).

The simulation survey biomass observations were assumed to follow a lognormal sampling error distribution with an expectation that is centred on the biomass trajectory and a CV of 0.25. For each of the three trajectories and sampling frequencies 10 000 sets of simulated observations were generated (Figure 4).

For each set of simulated observations the trajectory is re-estimated using the same model that simulated the data, hence the model parameters B_0 and r are estimated for each set of simulated observations. The negative log-likelihood was evaluated and minimised using the optim() function in R (R Core Team 2016**)**. For a simulated observation (o_i) and expectation (E_i) with known CV (c_i) the negative loglihood ($-\log(L)$) is as follows

$$-\log(L) = \sum_{i=1}^{n}\left(\log(\sigma_i) + 0.5\left(\frac{\log(\frac{o_i}{E_i})}{\sigma_i} + 0.5\sigma_i\right)^2\right)$$

Where $\sigma_i = \sqrt{\log(1 + c_i^2)}$

For each set of simulated observations, and associated re-estimated trajectories, the percentage change in biomass from the initial to the last year was calculated as

$$\hat{B}_{change} = \frac{\hat{B}_{2016}}{\hat{B}_o} 100$$

The estimates of percentage changes in biomass are compared to the actual changes under the three different biomass trajectory scenario, and the different sampling frequencies for biomass (annual, and two biennial).

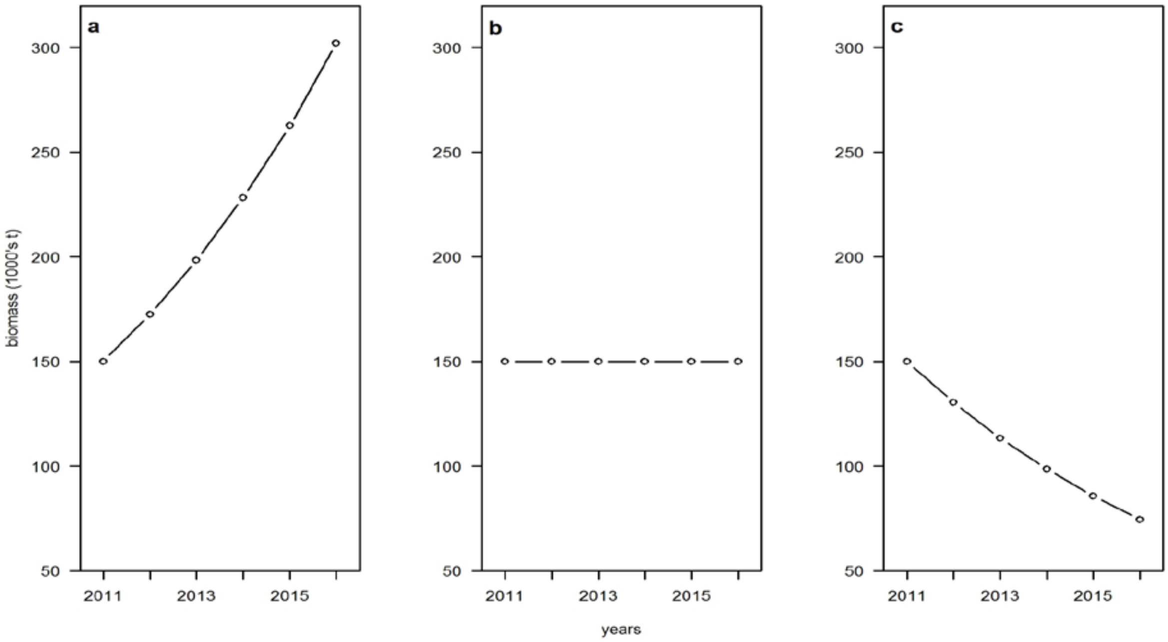

Figure 3: The three biomass trajectories all with B_{2011}= 150 000 t; a) exponentially increasing population with a growth rate parameter of r = 0.14 b) constant trajectory, c) decreasing population with a growth rate parameter of r = -0.14.

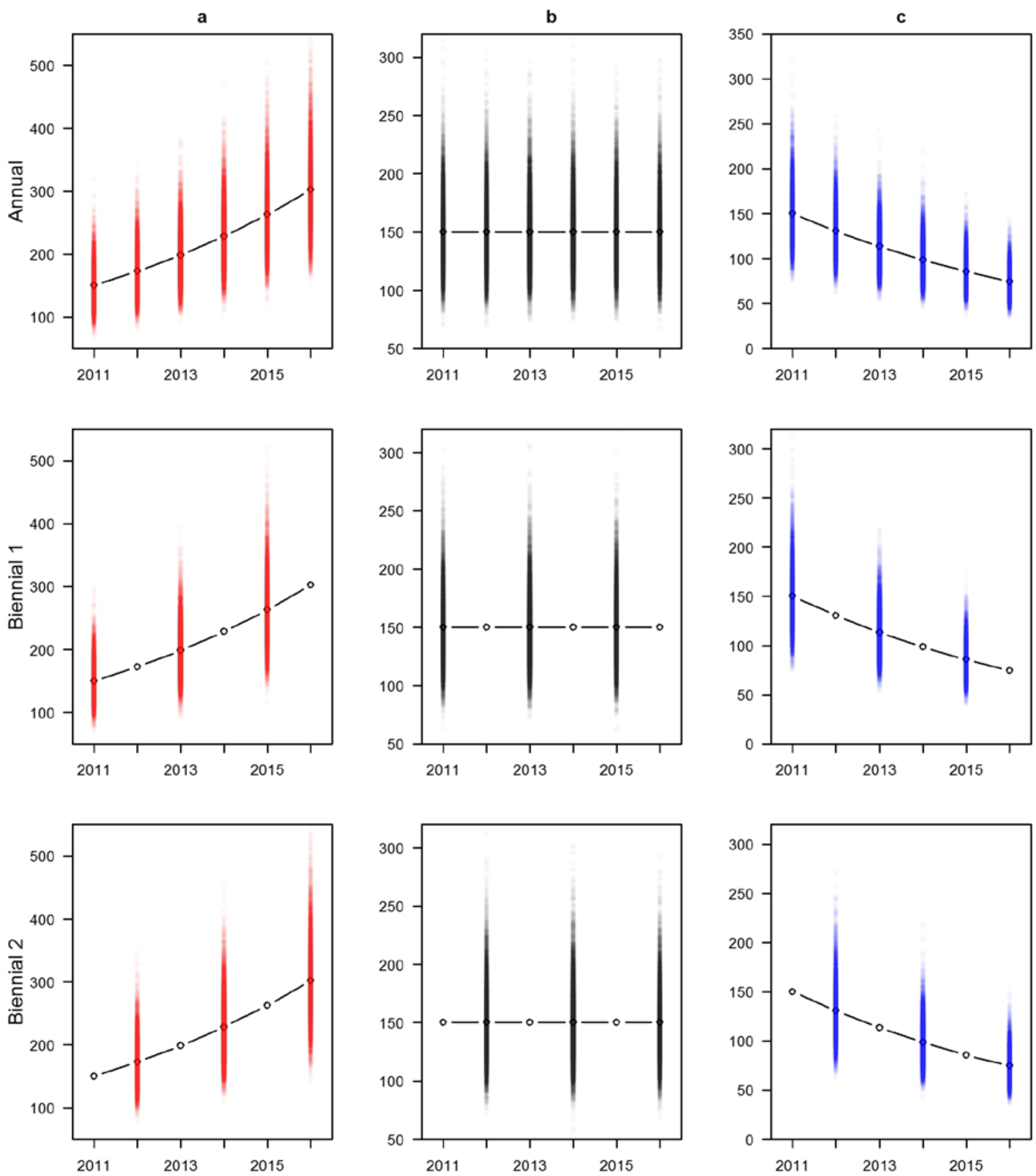

Figure 4: Simulated data for each of the combinations of the three trajectories (increasing, stable, decreasing) and temporal sampling frequencies (Annual, Biennial 1, Biennial 2).

2.2 Retrospective biomass estimations

Retrospective MCMC biomass estimates were calculated using the most recent assessment model for each stock (Table 1), comparing model runs with all data, to those where the Sub-Antarctic (SA) and Chatham Rise (CR) trawl surveys and their associated age composition data were excluded on a biennial basis. Five-year projections were also completed for each retrospective model run. Any catch sampling or other observations informing stock status were kept the same in the simulations. The stocks investigated and assessment models used are summarised in Table 1.

Table 1: Retrospective analyses stocks and assessment models used with references.

Stock	Associated area	Assessment year	Reference
HAK 4	CR	2013	Horn 2013
HAK 1	SA	2015	Horn 2015
LIN 3 & 4	CR	2015	McGregor 2015
LIN 5 & 6	SA	2015	Roberts 2016
HOK 1	CR & SA	2015	McKenzie 2016

The time frame for the biennial scenarios starts in 2005 and finishes in 2013, with two alternating biennial scenarios (see Table 2**).** The rationale behind this time interval being chosen and the two biennial scenarios is explained below.

The reason for the simulation studies is to evaluate the impact of moving from a sequence of annual trawl surveys to biennial trawl surveys. To mimic this in the biennial retrospective simulations one needs a sequence of annual trawl surveys, followed by biennial surveys. While the CR has a sequence of annual trawl surveys from 1992 onwards, in the SA these stopped and recommenced on an annual basis in 2001 (see Table 2). For this reason, 2005 was chosen as the start year in the biennial scenarios, in part to ensure these scenarios included a reasonable sequence of annual surveys. However, one should not choose a start year that is too late, as then there will be little data dropped in the biennial scenarios.

The finish year 2013 for the biennial surveys was chosen for a number of reasons, the main being that after then the CR and SA surveys began alternating biennially, so one of the biennial scenarios provides no further contrast as it matches what has already happened (see Scenario 2 in Table 2). This is particularly pertinent for the HAK 1 and LIN 5 & 6 stocks that have data just from the SA trawl survey.

Data after 2013 is dropped for both the annual and biennial scenarios, as after 2013 the CR and SA trawl surveys are biennial.

Implementing a retrospective analysis that covers nine years from 2005 to 2013 ensures that multiple cohorts are incorporated. For example, cohorts in the hoki fishery persist for around five to six years (Figure 5). The retrospective analyses run over nine years so we will see multiple cohorts which is more representative of a dynamic system (as opposed to a period dominated by the same cohorts).

For each stock and sampling frequency an MCMC estimation was carried out using CASAL v2.30 (Bull et al 2012), along with five-year projections. Each MCMC chain was diagnosed for convergence and reference points calculated. The reference points summarised were $B_{current}, B_{future}$ where both are spawning stock biomass as a percentage of virgin biomass, current is in 2013, and future refers to the projected biomass in 2018. These are compared for each stock across the three proposed sampling frequencies (annual, and two biennial scenarios).

A retrospective analysis looks back at the available data and makes changes to see the effects of changing assumptions, in this case the assumption of sampling frequencies. The benefit of a retrospective analysis is that it uses real data compared to simulated data, and MCMC simulations are computationally feasible. However, the retrospective analysis may not be informative of future scenarios where assessments will have more annual data than is expressed in a retrospective study. This is the reason we have done a third analysis, a forward projections simulation study, which includes all the current available data.

Table 2: Trawl survey data for the Sub-Antarctic (SA) and the Chatham Rise. Shown is the fishing year in which the trawl surveys occurred and data used in two biennial scenarios for retrospective biomass estimates. Data after 2013 is dropped for the retrospective analyses (annual, and both biennial scenarios). The lines encompass the range of years for which trawl survey data is dropped in the biennial scenarios.

Fishing year	Observed data		Biennial (Scenario 1)		Biennial (Scenario 2)	
Survey	SA	CR	SA	CR	SA	CR
1992	●	●	✓	✓	✓	✓
1993	●	●	✓	✓	✓	✓
1994	●	●	✓	✓	✓	✓
1995		●		✓		✓
1996		●		✓		✓
1997		●		✓		✓
1998		●		✓		✓
1999		●		✓		✓
2000		●		✓		✓
2001	●	●	✓	✓	✓	✓
2002	●	●	✓	✓	✓	✓
2003	●	●	✓	✓	✓	✓
2004	●	●	✓	✓	✓	✓
2005	●	●		✓	✓	
2006	●	●	✓			✓
2007	●	●		✓	✓	
2008	●	●	✓			✓
2009	●	●		✓	✓	
2010	●	●	✓			✓
2011		●		✓		
2012	●	●	✓			✓
2013	●	●		✓	✓	
2014		●				
2015	●					
2016		●				

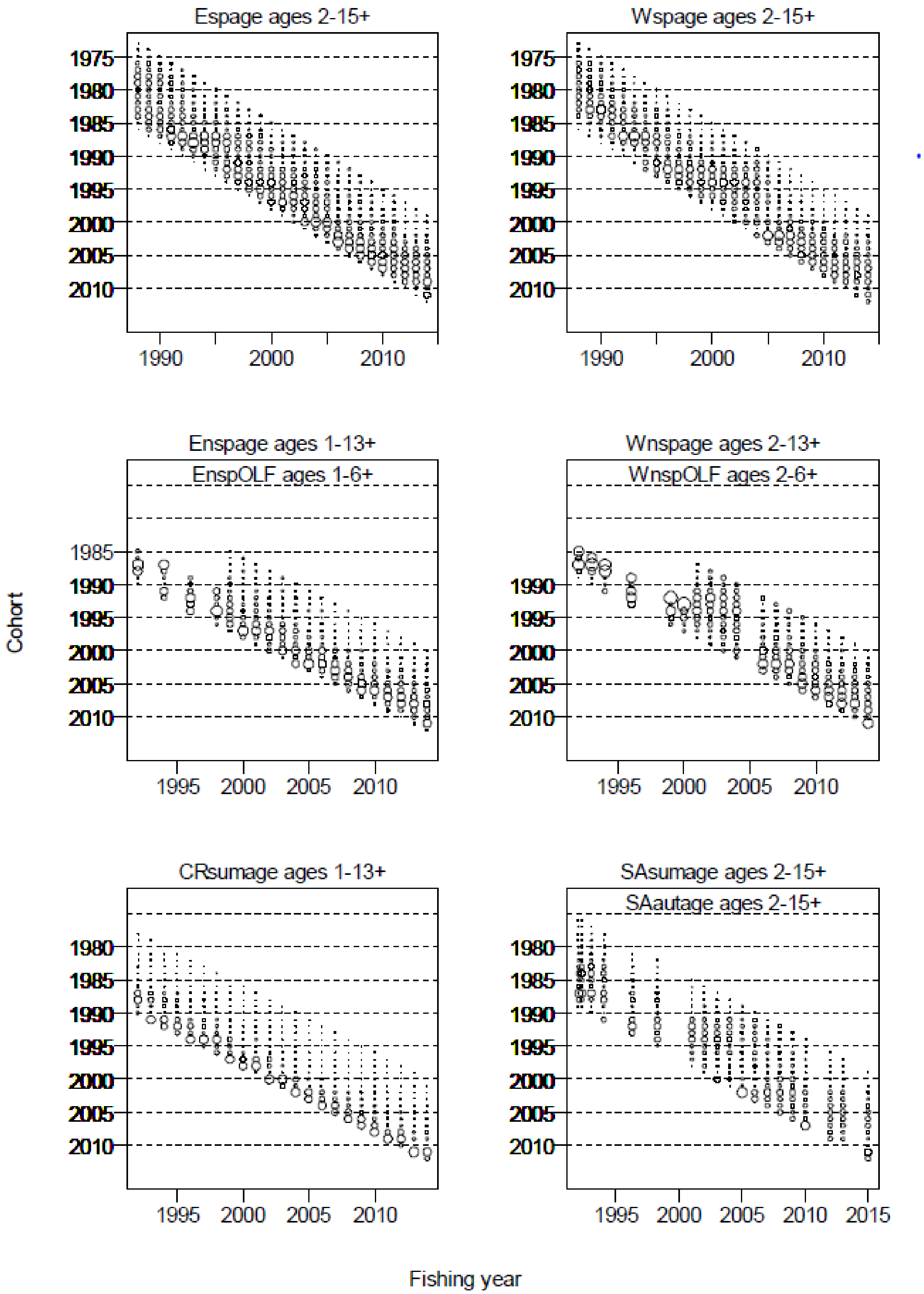

Figure 5: Proportions-at-age data, plotted by cohort and fishing year, with both sexes combined for the HOK 1 stock. The area of each circle is proportional to the associated proportion at age. Circle positions for the SAautage data in 1992 have been offset horizontally to allow them to be plotted on the same panel as the SAsumage data.

2.3 Forward Projection Simulations

2.3.1 Overview

The nature of the forward projection simulations are similar to the generic illustrative simulations, but are more closely based on the current data and state for a fish stock (e.g. hoki, hake, ling) with control over assumed year class strengths (YCSs) in the future.

Due to the complexity of the operating model, which includes time varying parameters, it was implemented in CASAL2 (Doonan et al. 2016). CASAL2 is NIWA's new population model dynamics platform which is an extension of CASAL.

Different scenarios for assumed future year class strengths are used to obtain known projected biomasses. For abundance estimates, survey biomass estimates are simulated on an annual or biennial basis from these known projected biomasses. For age data, simulated random sampling is conducted on the projected age frequencies. Assessment maximum posterior density (MPD) biomass estimates are done in the projected years, and compared for the annual and biennial scenarios.

For stocks where the current quota gives a stable biomass in projections, using YCSs of one in forward projections is unlikely to lead to much change in biomass. In this situation, it is unlikely that changing to biennial instead of annual surveys would matter much. Where it is more likely to matter is when the YCSs are very much below or above average, leading to steep biomass increases or declines. In one set of simulations the YCSs are set up to induce a declining biomass (and hence a potential lag in catch quota decreases), and in another set of simulations an increasing biomass (and a potential lag in catch quota increases to utilise the extra productivity).

We attempted to generate future data that is as realistic as possible. A common criticism of simulation studies is that they use data from a simplified population model that does not reflect a real and complex system, so inferences from them are poor.

In the sections that follow we summarise the operating model used to simulate future data, the scenarios investigated in forward projections, and the MPD re-estimation procedure used for each set of simulations.

2.3.2 Finding an Operating Model

The starting point for the operating model for all stocks was the base case stock assessment model that is currently accepted by the working groups (see Table 1). To add more realistic properties into the simulated data we investigated adding process error in the form of allowing traditionally fixed parameters in our model to vary randomly by year (time-vary). The two sets of parameters that we initially investigated to time-vary were selectivity and natural mortality (M) parameters. Both parameters were independently and identically drawn from the normal distribution to vary between each year. With time-varying M we found that this was unsatisfactory because (a) it required unrealistically large variation in M and (b) the simulations sometimes failed because with higher M there was insufficient biomass to allow the catch to be taken.

Selectivity parameters that could time-vary were the location parameter of the selectivity (a_{50} for logistic and the mean for the double normal). These time-varying selectivities have ideal properties for introducing process error, because they directly contribute to expectations of observations, which simulated data are based on. They are also used in fishing mortality processes, so add more realism into the operating model. The time-varying location parameters were assumed to follow independent draws from the normal distribution of the form

$$Sel_{\mu,t,s} \sim N(Sel_{\mu,mpd,s}, \sigma_s)$$

where, $Sel_{\mu,t,s}$ is the location parameter (μ) in year t for the s^{th} selectivity, $Sel_{\mu,mpd,s}$ is the estimated location parameter from the base case assessment, and σ_s the standard deviation of the normal distribution, referred to from here as the process error term.

The other component of the operating model (other than process error) that we investigated was the observational error for compositional datasets. Observational error was parameterised for each data set by a weighting factor (ω_i), where i is the ith dataset. The likelihood chosen to simulate compositional data was the logistic-normal likelihood. This was preferred to the commonly used multinomial because it can account for correlations between ages or lengths, which is often observed in compositional data (Figures 6–7). Another reason was that it did not generate zero compositional values, unlike the multinomial which was found to generate 5–20% zeros in initial simulations. Maximum likelihood estimation was used to identify the appropriate correlation parameter values that define the logistic-normal likelihood, given the observed and expected values from the base case assessment for a stock (Francis 2014).

The logistic-normal likelihood uses the logistic transformation of the Multivariate-Normal distribution. A logistic-normal distribution is formed by applying a logistic transformation to a multivariate normal vector. Specifically, let **O** be a composition vector for a year where the values sum to one, then **O** is logistic-normal with parameters $\{\mathbf{E},\mathbf{\Sigma}\}$ if $O_b = \exp(X_b) \,/\, \sum \exp(X_b)$, where the vector $\mathbf{X}$ is multivariate normal with mean log(**E**) and covariance matrix $\mathbf{\Sigma} = \ \omega_i \rho_1 \rho_2$.

A multivariate normal distribution can have a range of covariance structures, each of these corresponding to different assumed covariance structures for the logistic-normal distribution. Some common correlation structures coded into CASAL2 are LN1, LN2 and LN3 where the integer part of the label is the number of parameters used to describe the covariance structure as described in Francis (2014). For more information on the logistic-normal distribution see Francis (2014) and Aitchison (2003).

For simulated biomass data (i.e. trawl surveys) the lognormal likelihood was used with CV from the last year of the respective observation. We did initially investigate the observation error on biomass observations. However, we found that all assessments investigated in this study incorporated additional process error into the biomass observation error, which was deemed to be enough variation to produce realistic simulated biomass observations. Auto-correlation in simulated biomass data was investigated, but was not incorporated as there appeared to be little in the data (Appendix A).

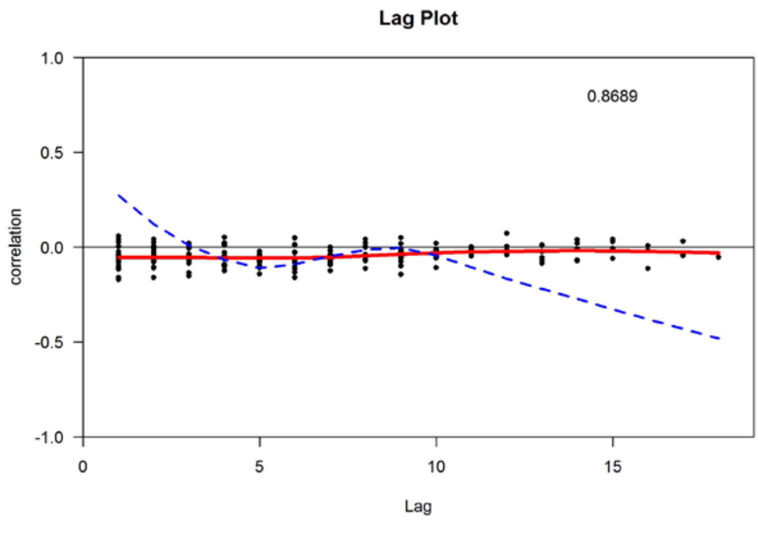

Figure 6: Residual correlation structure between age groups for the HAK 1 stock base case assessment for observer age dataset. Blue dashed line is a loess fit to the observed residual correlation observed in the base case assessment, black dots and red line (loess fit) are the residual correlation when using a multinomial likelihood for compositional data.

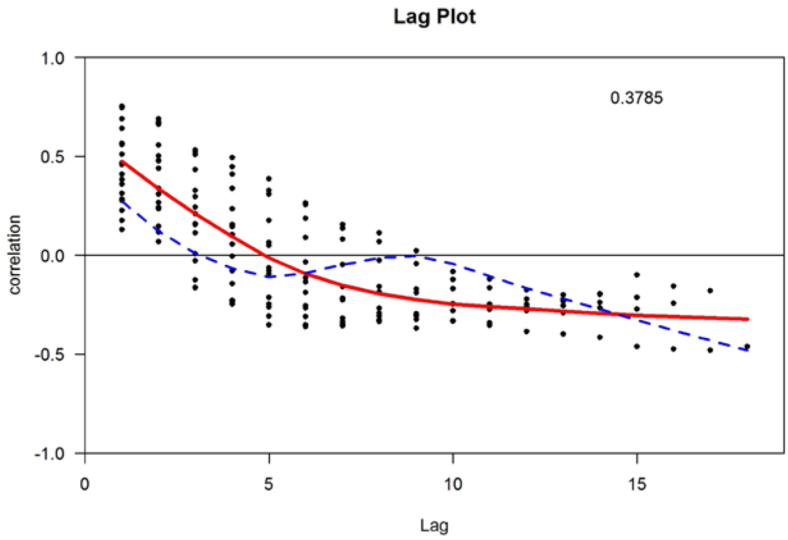

Figure 7: Residual correlation structure for the HAK 1 stock base case assessment for observer age data. The blue line is a loess fit to the observed residual correlation, black dots and red line (loess fit) are the residual correlation when using a logistic-normal likelihood for compositional data.

The residuals distribution patterns for each stock were used to choose the process and observational error parameters σ_s and ω_i. This involved looking at the residual patterns $(\boldsymbol{O_{past}} - \boldsymbol{E_{past}})$ that exist between our past observed data $(\boldsymbol{O_{past}})$ and the base case model fits to that data $(\boldsymbol{E_{past}})$. The goal in selecting operating model parameters σ_s and ω_i is to create simulated data $\boldsymbol{S_{future}}$ that have the same residual variance and correlation patterns as $(\boldsymbol{O_{past}} - \boldsymbol{E_{past}})$. This was tested by simulating data in the past $(\boldsymbol{S_{past}})$ and comparing these to the base case fit $(\boldsymbol{E_{past}})$.

How we determine σ_s and ω_i for each observation set is outlined in the following algorithm:

1. Find the combinations of σ_s and ω_i that satisfy $var(\boldsymbol{O_{past}} - \boldsymbol{E_{past}}) \approx var(\boldsymbol{S_{past}} - \boldsymbol{E_{past}})$.

2. From the combinations in step one, find the single combination σ_s and ω_i that best recreates the residual pattern in the compositional data (including the correlation structure).

In step one the residual variance refers to residual mean age or length for a compositional data time series and the residual abundance or biomass for their relevant time series. An example of such a variance plot is shown in Figure 8. From this figure, we can see that any intersection of the horizontal dashed black line ($var(\boldsymbol{O_{past}} - \boldsymbol{E_{past}})$) are considered candidate pairs of σ_s and ω_i for the operating model. Linear interpolation is used to find the optimal weighting factor for each of the sigma curves that intercept the horizontal dashed line.

In step two for each of the candidate σ_s and ω_i we investigated the inter-age or inter-length correlation structure using lag diagnostics, to find the final values of σ_s and ω_i that are used in the operating model. Lag diagnostics is a plot for visualising the residual intra-bin correlations (bin refers to either an age or length bin). Let there be a matrix of observed proportions $P_{y,a}$ and expected proportions $\widehat{P_{y,a}}$ where y is the year and a is the age. To investigate inter-bin correlations we look at the raw residuals $(e_{y,a})$ which are defined as

$$e_{y,a} = P_{y,a} - \widehat{P_{y,a}}$$

For ease of notation let $\boldsymbol{e_a}$ represent a vector of residuals for age $\boldsymbol{a}$ for all years i.e. $\boldsymbol{e_a} = \{e_{1,a}, e_{2,a} \dots e_{Y,a}\}^T$, where Y is the last year in the series. Inter-bin correlations can be viewed by looking at the correlations of residuals of bins close together, for example between ages 3 and 4 cor($\boldsymbol{e_3}, \boldsymbol{e_4}$), where cor() is Pearson's correlation coefficient. These ages have a difference of 1, which is called the lag. If we plot the correlations between all possible combinations of bins, we can see that amongst most of our compositional data there are correlations (Appendix B). An example of a lag plot is shown for the observer catch at age (subaOBSage) and survey proportions at age (subaTANage) from the HAK 1 stock assessment (Figure 9).

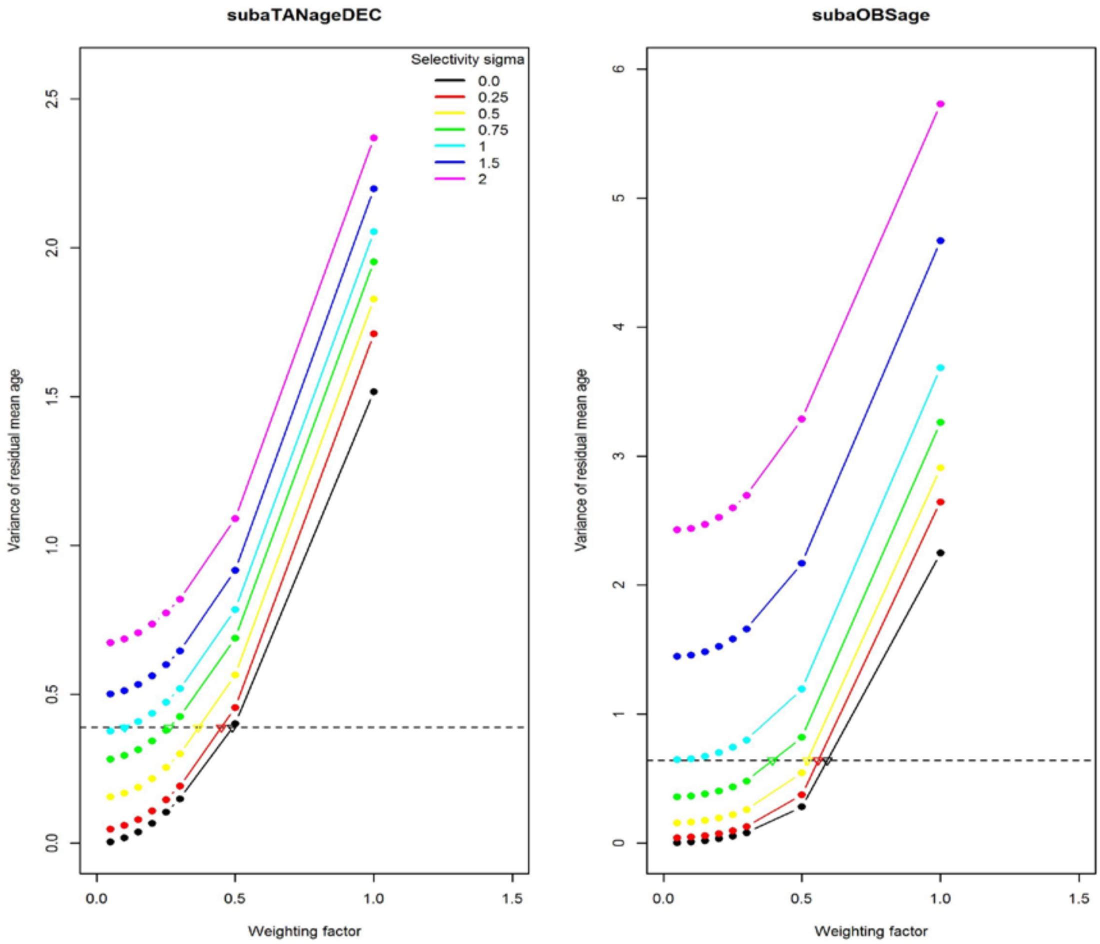

Figure 8: Residual variance plots for two compositional observations from the HAK 1 assessment, over a range of sigma values (separate coloured lines) and weighting factors (x-axis). The horizontal dashed black line is the residual variance found in the base cases assessment.

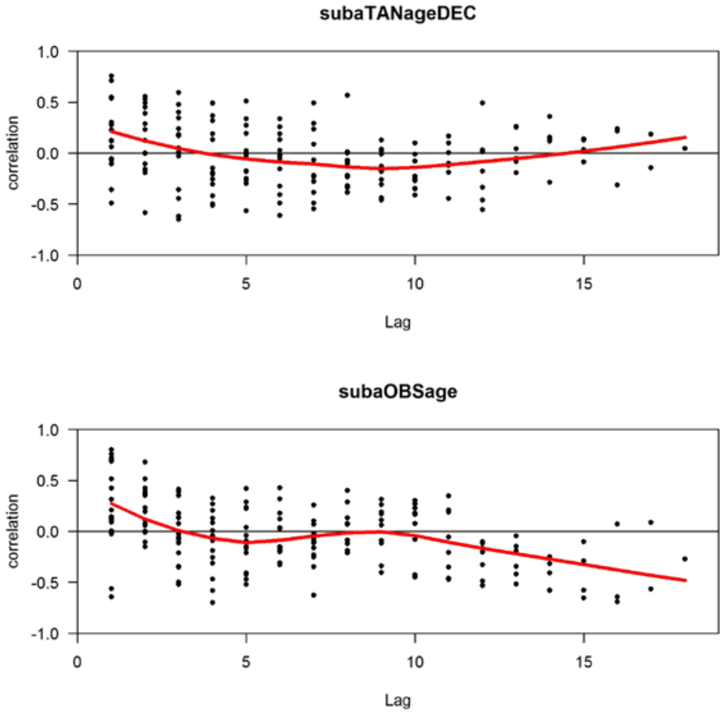

Figure 9: Lag plot of Pearson's correlation for two at age compositional datasets from the HAK 1 stock: subaTANageDEC (Sub-Antarctic survey proportions at age) and SubOBSage (observer catch at age dataset). The solid fit red line is a loess line fit.

We fitted a loess to the lag diagnostics for $\boldsymbol{O_{past}} - \boldsymbol{E_{past}}$ and also $\boldsymbol{S_{past}} - \boldsymbol{E_{past}}$ then used minimisation of the sum of squares for the difference between the loess curves to find the final combination of σ_s and ω_i. In particular σ_s and ω_i were determined from the combinations from step one by minimising the following term with respect to them

$$SSE = \sum_{i=1}^{Y-1} (L_i - \hat{L}_i)^2$$

where, L_i is the loess fit for lag i and $\hat{L}_i$ is the loess fit for a proposed simulating model comprised of σ_s and ω_i, and Y is the number of years of composition data. The final operating model was the combination of σ_s and ω_i that had the smallest SSE term.

The final decision regarding the operating model was choosing the observational sampling error term for future years. For simplicity, this was set for all future years to the last error value from the base case assessment.

Figure 10 illustrates how both process and sampling error can contribute to deviations from a model's expectation and shows how the operating models incorporates both for more realism in the simulations.

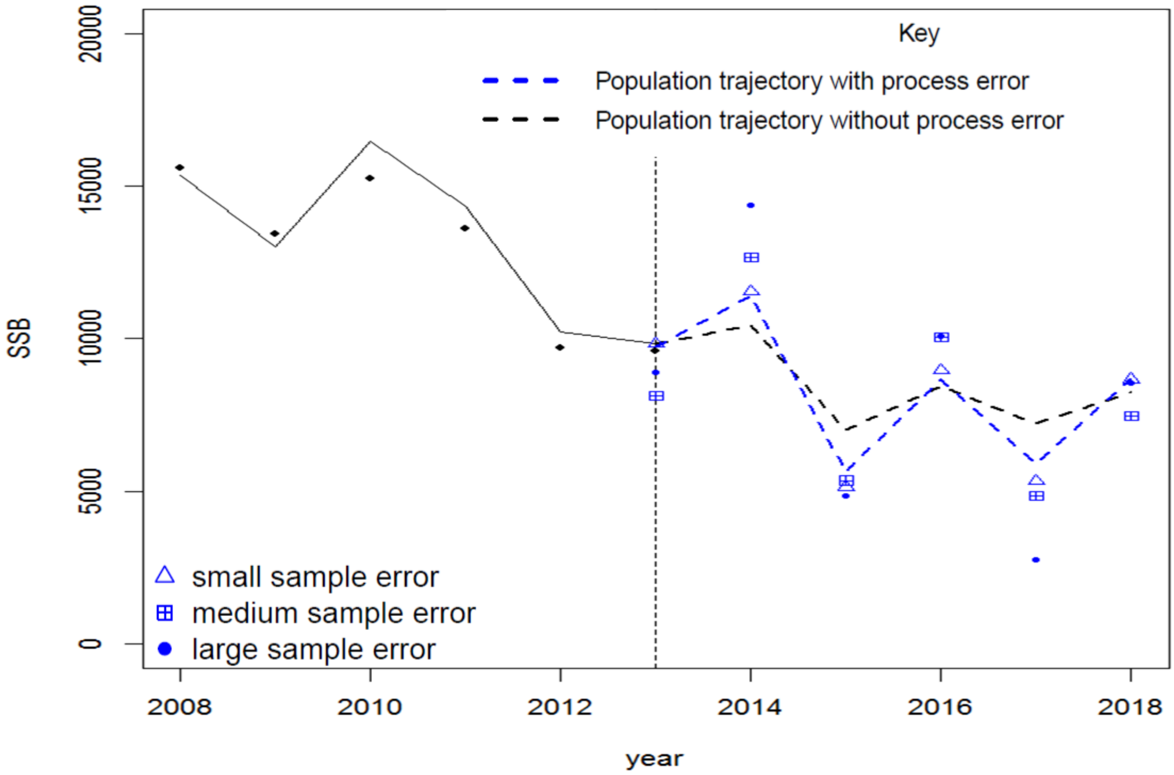

Figure 10: A conceptual plot of how both process and sampling variability can affect simulated observations. To the left hand side of the vertical dashed line, the dots are observed data O_{past} and the line represents the expectation E_{past}. On the right hand side of the vertical dashed line the dashed black line is the (E_{future}) model and blue points are simulated data (S_{future}).

2.3.3 Scenarios

The forward projection simulation study is for five stocks (HAK 1, HAK 4, LIN 3&4, LIN 5&6 and HOK 1), with three alternate biomass trajectories (increasing, decreasing and stable), and two alternative trawl survey frequencies (annual, biennial). This results in 30 [=5 × 3 × 2] scenarios, where for each scenario there are 500 random replicates. We explain here how the operating model is used to project the stock forward into the future.

For the operating model, changes were made to the recruitment dynamics to project the stocks forward ten years under three different trajectories: declining, stable or increasing population. We choose three contrasting trajectories because we hypothesise that the effect of reducing survey frequency will be greater in situations where the biomass is not stable. The biologically plausible mechanism of recruitment changes was used to induce the future trajectories, as has been observed before for some stocks e.g. for the hoki western stock there was a period of low recruitment in the late 1990s which contributed to a subsequent decline in biomass (McKenzie 2016). Other possible scenarios include mass mortality events. However, given the time constraints we went with the most likely process to affect future stock status.

For long-term simulations, a harvest control rule would be needed to make the simulations more realistic. However, no harvest control rule exists for any of the fisheries, and in any case changes to catch often involve factors other than stock status (so any harvest control rule driven only by stock status would be unrealistic). Instead the forward simulations are done for a relatively short time of ten years, with constant catch, using the last year of catch in the base case assessment. In lieu of a changing future catch driven by a harvest control rule, current catches seem to be a plausible near-term management target for the hoki, hake, and ling fisheries.

Different trajectories are induced by scaling future year class parameters by a ratio of current R_0 compared with an assumed future R_f:

$$Y'_y = Y_y \frac{R_f}{R_0}$$

Where, Y'_y is the future year class effect, R_f is the future average recruitment, and Y_y is a year class that is generated from using the projection method for generating year class values in the base case stock assessment. This mimics a regime shift in productivity or a change in carrying capacity, and maintains realistic trends in future year class effects which is better than having a constant year class effect going into the future.

Because this study covered five different stocks, that all have different current stock status relative to B_0, a criterion was developed to choose the value for R_f for each stock. For an increasing and decreasing future trend, an R_f is chosen that creates a 25% change in %B_0 from the base case assessment year, to ten years into the future. For example, if current status for %B_0 is 50% then we want this to be 75% B_0 after a ten projection into the future under an increasing scenario. For the stable scenario there would be no change after a projection ten years into the future.

When simulating data, we don't standardise over future YCSs, instead we keep the standardisation the same as the base case assessment model, with assumed future YCSs not standardised. The reason for this is demonstrated in Figure 11, where if we standardise over future YCSs, the expectation using the current base case assessment parameters change, which we did not want to happen.

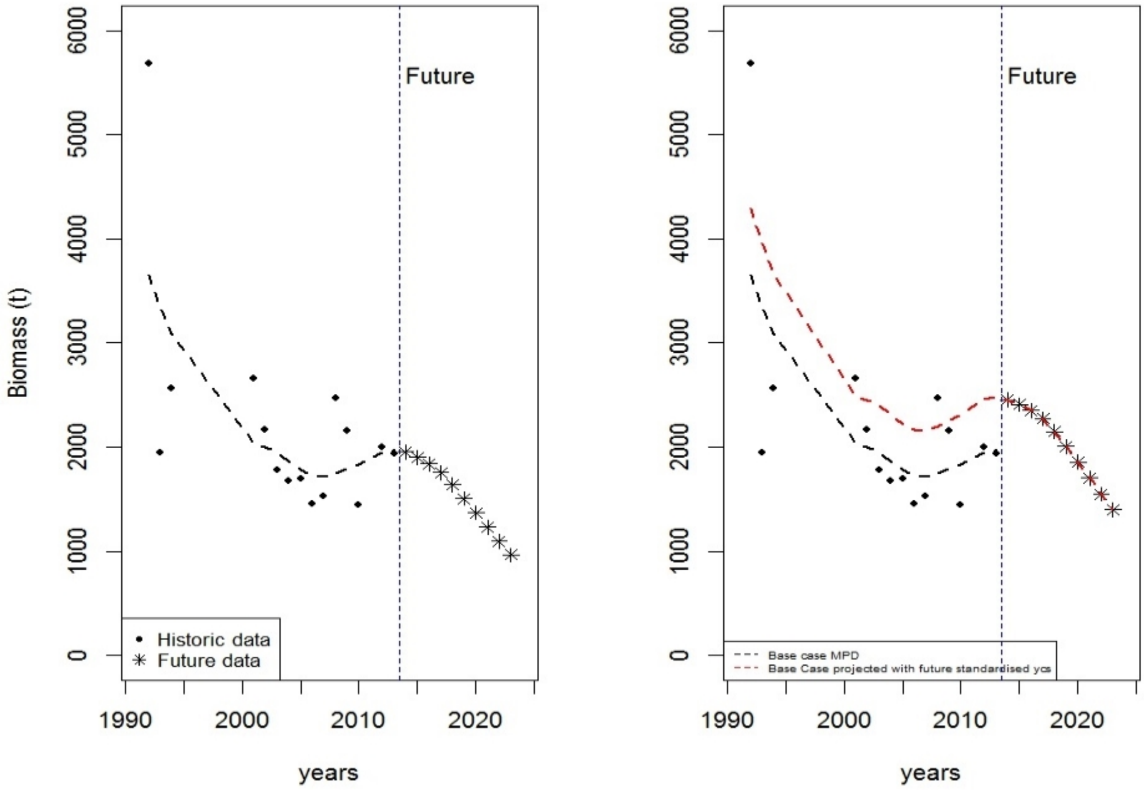

Figure 11: A comparison of future expected trajectories and also mean simulated data under two different standardisation procedures for YCSs. On the left panel we have kept the standardisation as specified in the base case assessment model (future YCS are not standardised). On the right panel we have standardised YCSs over future years as well. The standardisation procedure on the left was chosen for the operating model.

For datasets for which there are no strategic future plans, we set a blanket rule that if a dataset was used in each of the final three years, it was assumed to continue into the future on an annual basis. There are many reasons why the frequency of other observations used in the assessment may not be as consistent as trawl surveys such as un-representative samples, or high uncertainty around estimates which cannot be predicted in the future, hence the use of this simplified rule. For a summary of the temporal frequency for future observations for HAK 1, HOK 1 and LIN 3 & 4 see Tables 3–6.

Along with having three different trajectories (increasing, decreasing, stable) we ran two alternative survey scenarios: (i) annual surveys, and a (ii) biennial survey scenario (which follows the current pattern of biennial surveys).

If the trawl surveys are biennial, a possible future scenario is that they alternate with fishery dependent commercial catch-at-age observations, which is investigated in a sensitivity run for the hoki stock assessment. The observations for this sensitivity run are shown in Table 7.

Table 3: Frequency of HOK 1 datasets in the forward simulation, CSacous is assumed biennial based on O'Driscoll et al. (2016). A black dot means the observation was simulated in that year. Espage = Eastern stock spawning commercial age frequency, Enspage = Eastern stock non-spawning commercial age frequency, Wspage and Wnspage are the same as the two previously mentioned except they apply to the western stock, WCacous is the west coast acoustic survey and CSacous is the cook straight acoustic survey.

Year	Espage	Enspage	Wspage	Wnspage	WCacous	CSacous
2016	●	●	●	●		●
2017	●	●	●	●		
2018	●	●	●	●		●
2019	●	●	●	●		
2020	●	●	●	●		●
2021	●	●	●	●		
2022	●	●	●	●		●
2023	●	●	●	●		
2024	●	●	●	●		●
2025	●	●	●	●		
2026	●	●	●	●		

Table 4: HAK 1 Observer catch at age frequency for projected years.

Year	Commercial age frequencies (subOBSage)
2014	●
2015	●
2016	●
2017	●
2018	●
2019	●
2020	●
2021	●
2022	●
2023	●

Table 5: LIN 3 & 4 observer catch at age frequency for projected years.

Year	Trawl Catch at age
2014	●
2015	●
2016	●
2017	●
2018	●
2019	●
2020	●
2021	●
2022	●
2023	●

Table 6: Annual vs biennial frequencies for the Chatham Rise (CR) and Sub Antartic (SA) surveys for future years.

Year	Annual		Biennial	
	CR	SA	CR	SA
2012	✓	✓	✓	
2013	✓	✓		✓
2014	✓	✓	✓	
2015	✓	✓		✓
2016	✓	✓	✓	
2017	✓	✓		✓
2018	✓	✓	✓	
2019	✓	✓		✓
2020	✓	✓	✓	
2021	✓	✓		✓
2022	✓	✓	✓	
2023	✓	✓		✓
2024	✓	✓	✓	
2025	✓	✓		✓
2026	✓	✓	✓	

Table 7: Observation frequency for the sensitivity run for the HOK 1 stock assessment.

Year	Espage	Enspage	Wspage	Wnspage	WCacous	CSacous	SA survey	CR survey
2016			●	●		●		●
2017	●	●					●	
2018			●	●		●		●
2019	●	●					●	
2020			●	●		●		●
2021	●	●					●	
2022			●	●		●		●
2023	●	●					●	
2024			●	●		●		●
2025	●	●					●	
2026			●	●		●		●

2.3.4 Re-Estimation model

The re-estimation model for a stock was done in CASAL and was the same as the base case assessment model (refer to the respective assessment report in Table 1 for more details). The changes that are needed to the base case model are simply extending ten years into the future for the observational data, and refitting the model. Priors were kept the same as in the base model and future catches the same as the last year of the base model. Following the procedure in the base model the YCSs were standardised.

The base model is stepped forward in time incrementally, during which parameters were re-estimated under various scenarios at each increment. This was done for 10 years which was the end of the simulation period.

In detail the algorithm for re-estimation follows:

- For each assessment from the end of the base case assessment year ($base_{year}$) step forward two years and do an MPD fit, keeping all the data for the annual situation and dropping out the necessary data for the biennial situation (see Tables 3–6).

- For each MPD apply the Francis TA1.8 re-weighting method (Francis 2011a) after the first MPD and re-run the MPD.

- Summarise the estimation results.

- Step forward until another two years and repeat the above process, until ten years in the future is reached.

We summarise here the entire decision-making process that entails finding the operating model, setting future trajectories, and re-estimation.

1. Pick a sub set of σ_s and ω_i that satisfy $var(\boldsymbol{O_{past}} - \boldsymbol{E_{past}}) \approx var(\boldsymbol{S_{past}} - \boldsymbol{E_{past}})$ for all simulated composition observations.

2. From the subset of σ_s and ω_i found in step 1, search for the combination that creates the smallest SSE term between the loess fits in the residual correlations.

3. Find a value for the future average recruitment (R_f) to get either a 25% increase, decrease, or 0% change in biomass within ten years.

4. Simulate N sets of observations for each stock and trajectory.

5. Re-estimation, for the annual situation keeping data for all years or for the biennial situation taking every second year for the relevant datasets. Re-estimation model has constant future catches and standardises over future YCS parameters.

6. Summarise the outputs.

Goodness of fit is always a difficult task with so many model outputs. We randomly selected five assessment fits, and if the Pearson's residuals were within the 95% confidence interval they were deemed to be satisfactory fits. These five fits were assumed to be representative of the 495 remaining re-estimation outputs.

All re-estimation model outputs were summarised by spawning stock biomass as a percentage of equilibrium spawning stock biomass ($\%B_{current}$)

$$\%B_{current} = \frac{B_{current}}{B_0}$$

These were presented for all future years that had survey data available, runs comparing the annual survey frequency to the biennial frequency are presented in the following results section.

3. RESULTS

In the sections that follow each set of results is considered in turn: (a) generic simulations, (b) retrospective, and (c) forward simulations. Overall, there was little difference between the annual and biennial survey results.

3.1 Generic Simulations

For the generic simulations there are three biomass trajectory scenarios: (a) increasing, (b) constant, and (c) decreasing. These simulations demonstrate that going from annual to biennial for simulated observed data, there is no change in the estimate of the percentage change in biomass over the trajectory, but an increase in the variance for this estimate (Figure 12). The increase in the variance was asymmetric, with greater increase for the right-hand tail of the distribution. There was very little difference between the two biennial scenarios.

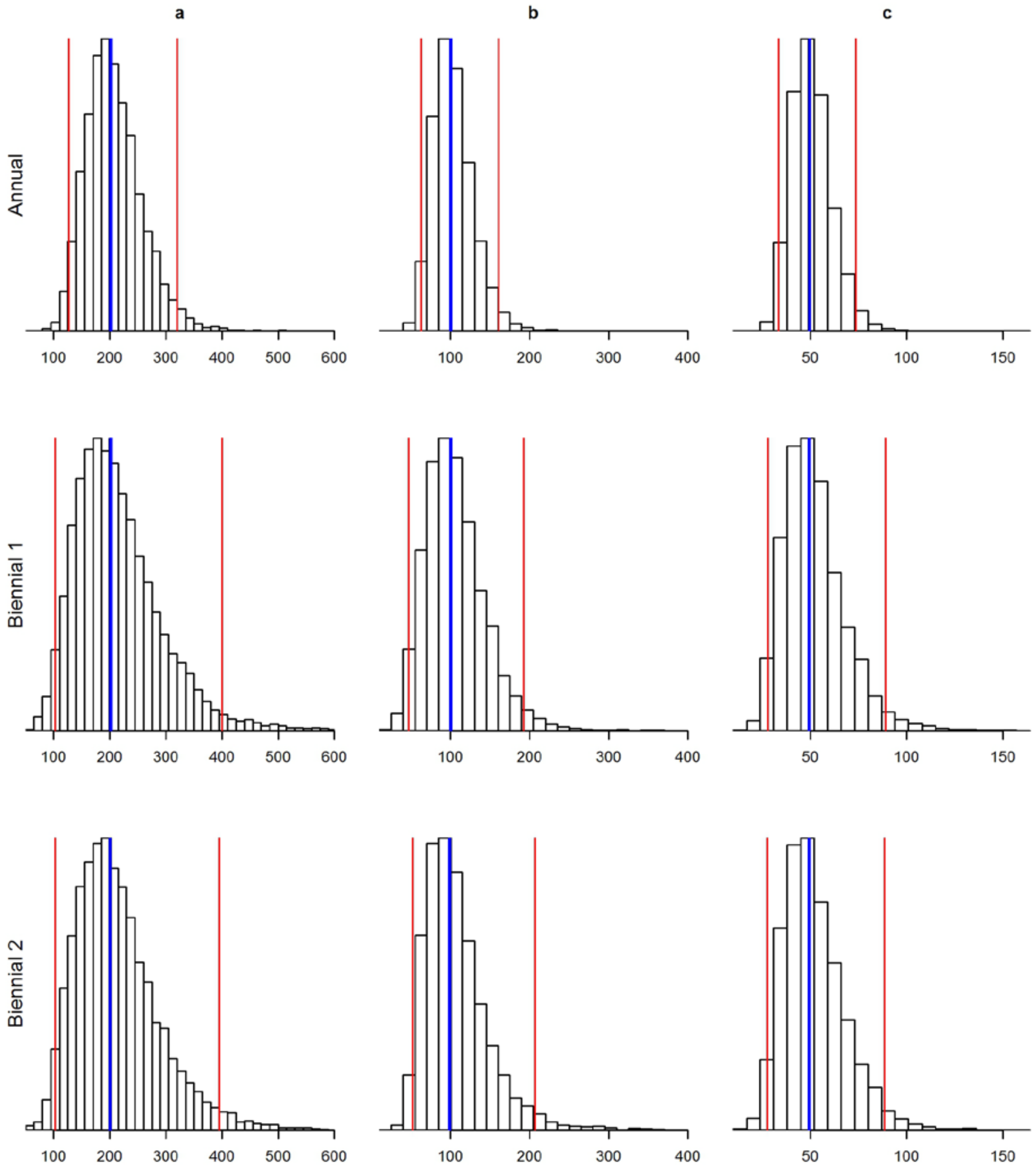

Figure 12: Histogram of $\widehat{B}_{change}$ (percentage changes in biomass from 2011 to 2016) for different scenarios and biomass trajectories. The three scenarios (shown in the rows) are annual and two biennial scenarios. The alternative trajectories (shown in the columns) are: a) exponentially increasing population by a rate of exp(0.14), b) constant trajectory, c) decreasing population by a rate of exp(-0.14). The red vertical lines denote the 95% quantiles for the histograms.

3.2 Retrospective biomass estimates

In these simulations, under annual and two biennial scenarios, estimates are made of biomass (%B_0) in 2013 ($B_{current}$) and with a five-year projection (B_{future}).

Except for the western hoki stock and the second biennial scenario, there is little difference in the estimate of current biomass, for either the median value or variance (Figures 13–18). The difference noted for the western hoki stock is hypothesised to be due to the three-year gap for the Sub-Antarctic trawl survey under the second biennial scenario (see Table 2). Differences in current biomass estimates are magnified under five-year projections, with the biggest difference being for the western hoki stock.

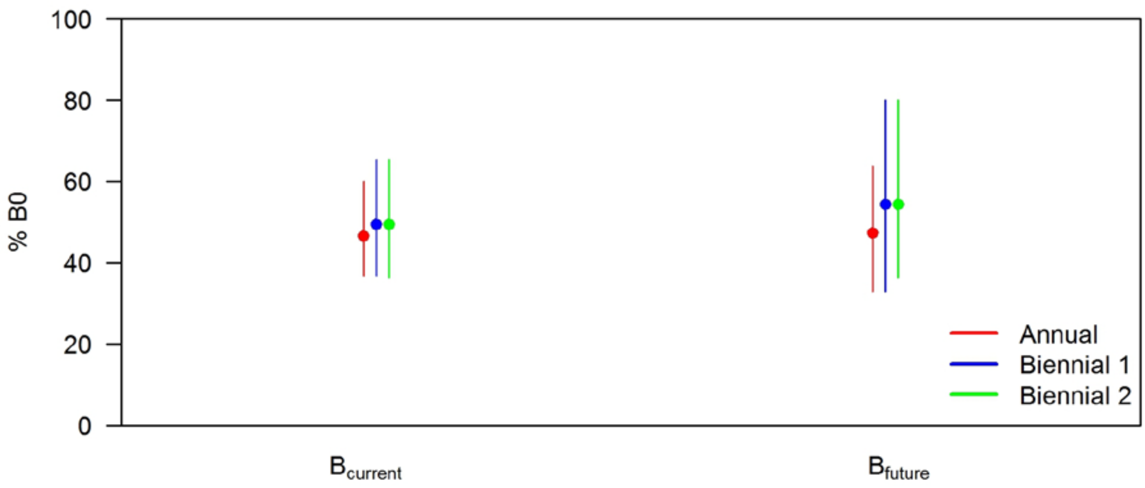

Figure 13: Comparisons of reference points $B_{current}$ and B_{future} from the HAK 4 stock assessment between the three alternative scenarios.

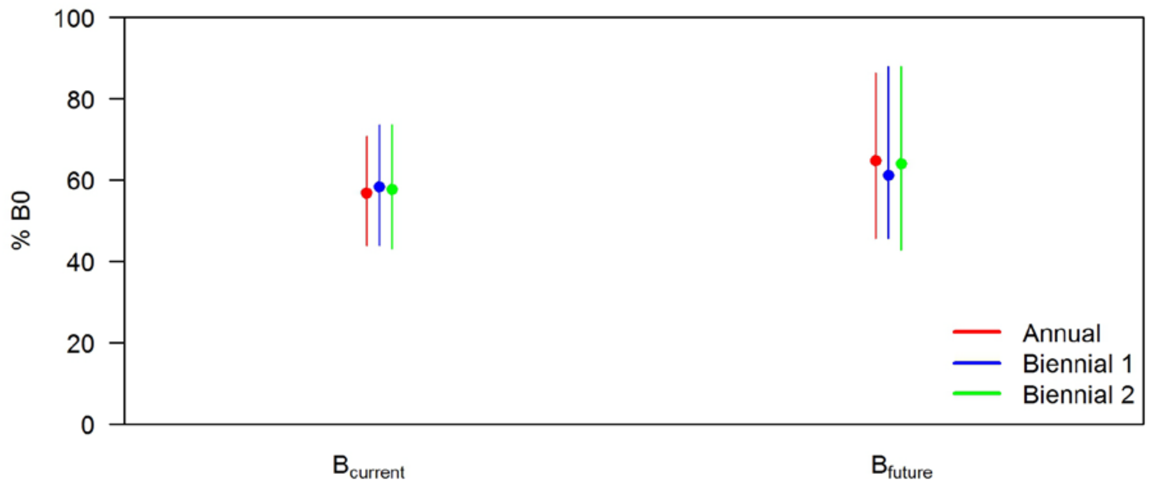

Figure 14: Comparisons of reference points $B_{current}$ and B_{future} from the HAK 1 stock assessment between the three alternative scenarios.

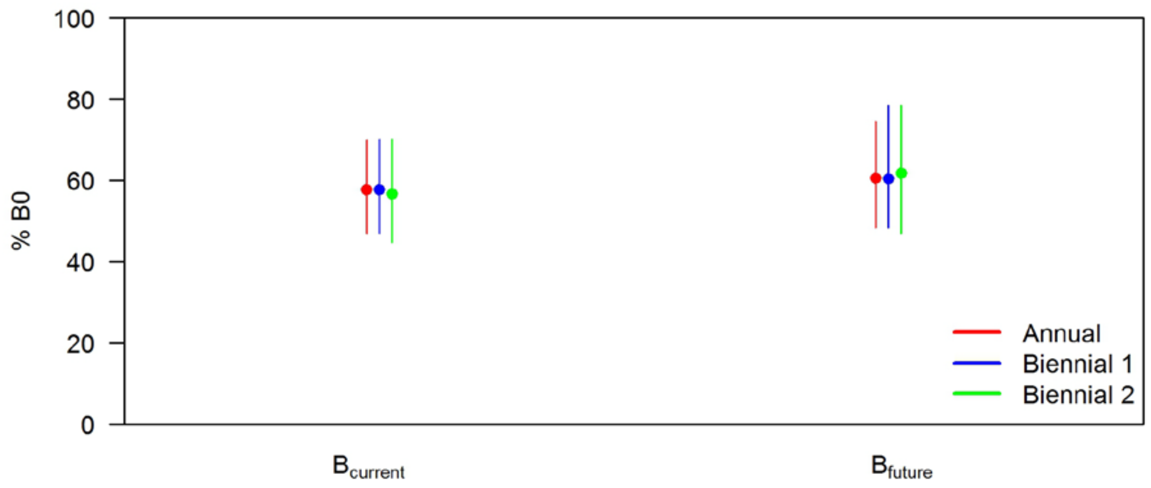

Figure 15: Comparisons of reference points $B_{current}$ and B_{future} from the LIN 3 & 4 stock assessment between the three alternative scenarios.

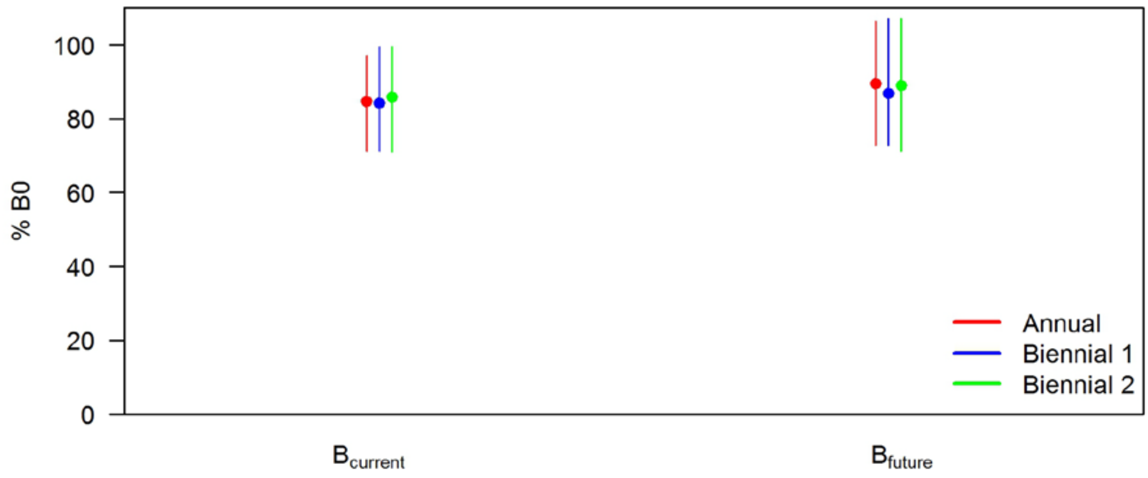

Figure 16: Comparisons of reference points $B_{current}$ and B_{future} from the LIN 5 & 6 stock assessment between the three alternative scenarios.

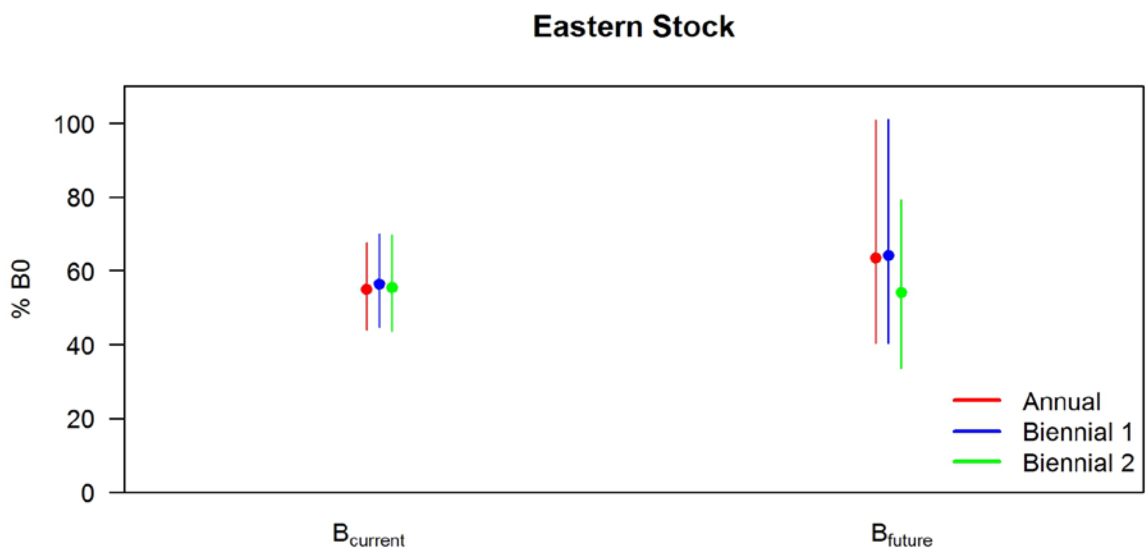

Figure 17: Comparisons of reference points $B_{current}$ and B_{future} for the eastern stock from the HOK 1 stock assessment between the three alternative scenarios.

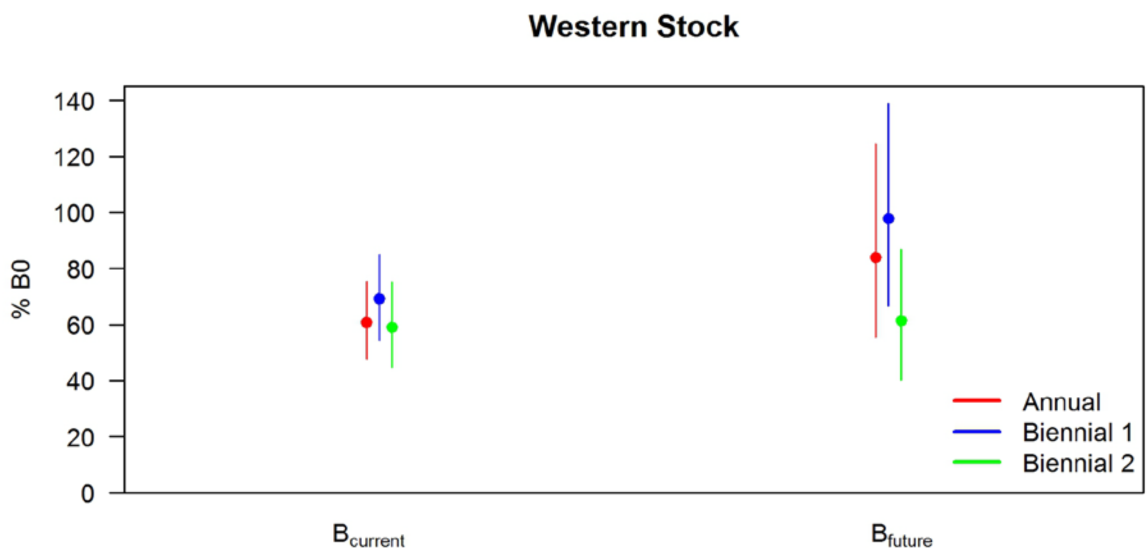

Figure 18: Comparisons of reference points $B_{current}$ and B_{future} for the western stock from the HOK 1 stock assessment between the three alternative scenarios.

3.3 Forward Projection Simulations

For forward projection simulations there was somewhat larger variance in stock status estimates for biennial than annual (Figures 19–36, Tables 9–14). But overall there was little difference between the three scenarios re-estimated, though for certain scenarios and stocks the difference was more pronounced (e.g. HAK 1 and the decreasing trajectory – see Figure 20).

An interesting finding from this analysis was that the expectation could not be re-estimated by both the annual and biennial scenario. This was investigated and found to be due to a difference existing between the operating model and the re-estimation model. The operating model simulated future data based on un-standardised year class parameters. When we re-estimated the simulated data, the year class parameters were standardised over all years (a common practice). This meant that both the annual and biennial scenarios couldn't re-estimate the underlying known biomass when the population was exhibiting change. This highlights the effect of the assumption of forcing relative year class strengths to average one, when in fact this might not be true. This suggests that under periods of low recruitment our models may be optimistic, and under periods of high recruitment our models are pessimistic.

3.3.1 HAK 1

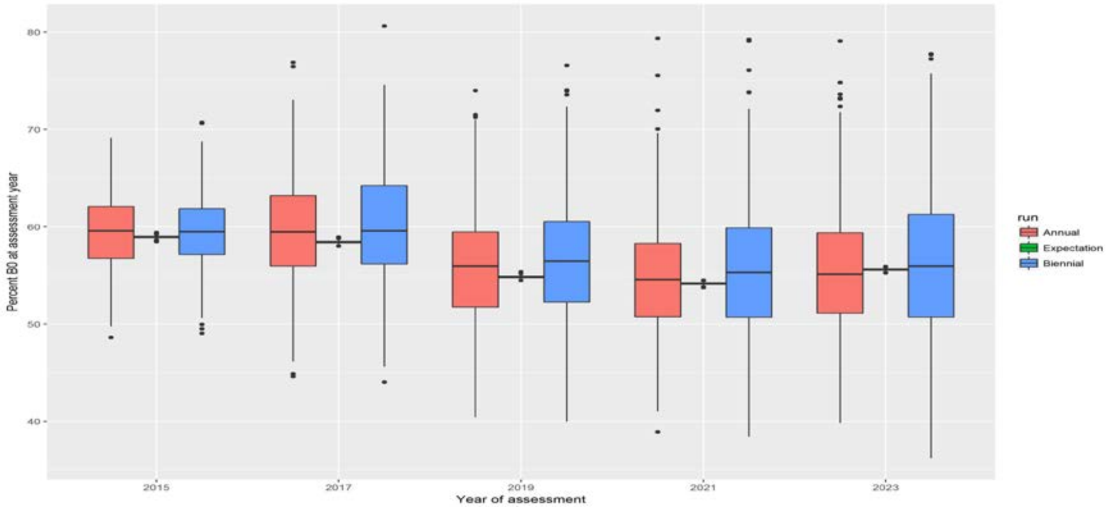

Figure 19: HAK 1 and stable trajectory for the operating model. Current biomass estimates ($\%B_0$) for forward simulation assessments done every second year. Shown are the estimates from 500 MPD fit scenarios: annual, expectation of the operating model, and biennial data.

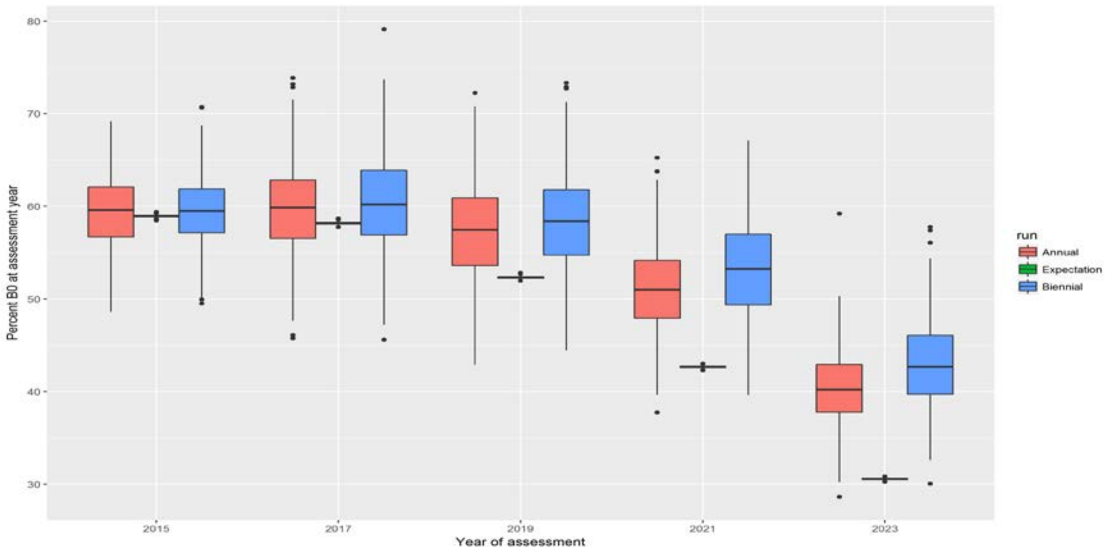

Figure 20: As in Figure 19, but for the decreasing trajectory.

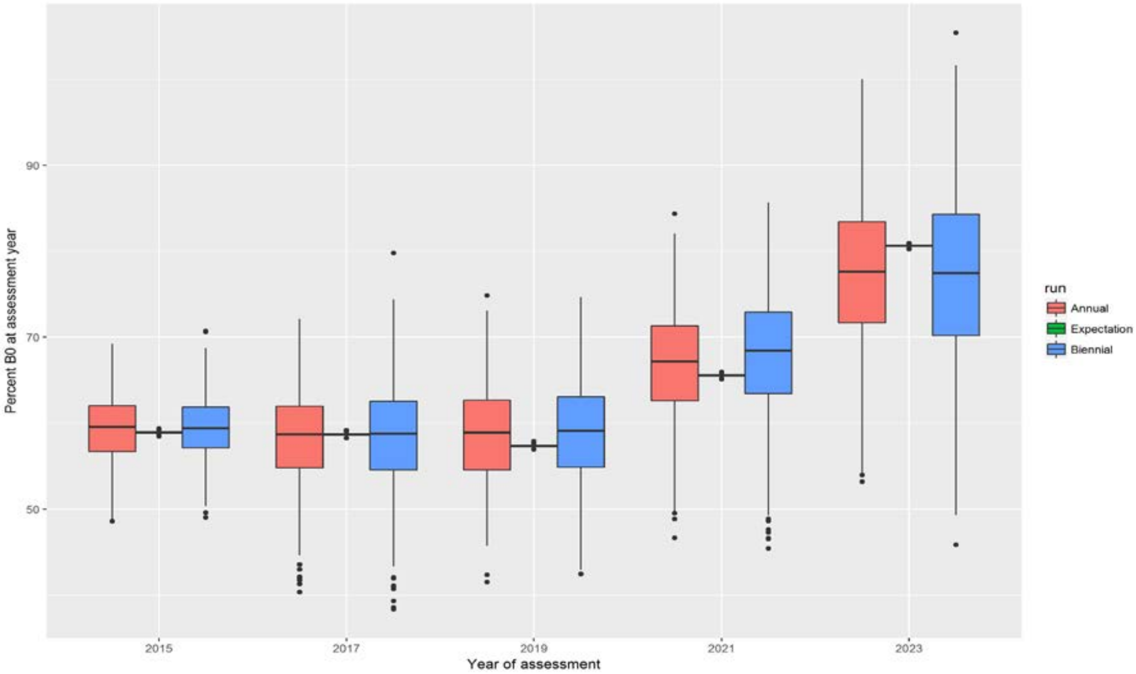

Figure 21: As in Figure 19, but for the increasing trajectory.

Table 8: HAK 1 and different trajectories for the operating model. Standard deviation of $\%B_0$ estimates for the annual and biennial frequency for each year of assessments.

Trajectory	Survey Frequency	2015	2017	2019	2021	2023
Stable	Annual	3.8	5.4	5.7	5.8	6.6
	Biennial	3.7	5.7	6.4	6.9	8.0
Decreasing	Annual	3.8	4.9	5.1	4.7	4.1
	Biennial	3.7	5.1	5.3	5.2	4.6
Increasing	Annual	3.8	5.8	5.7	6.8	8.6
	Biennial	3.7	6.6	6.1	7.4	10.3

3.3.2 HAK 4

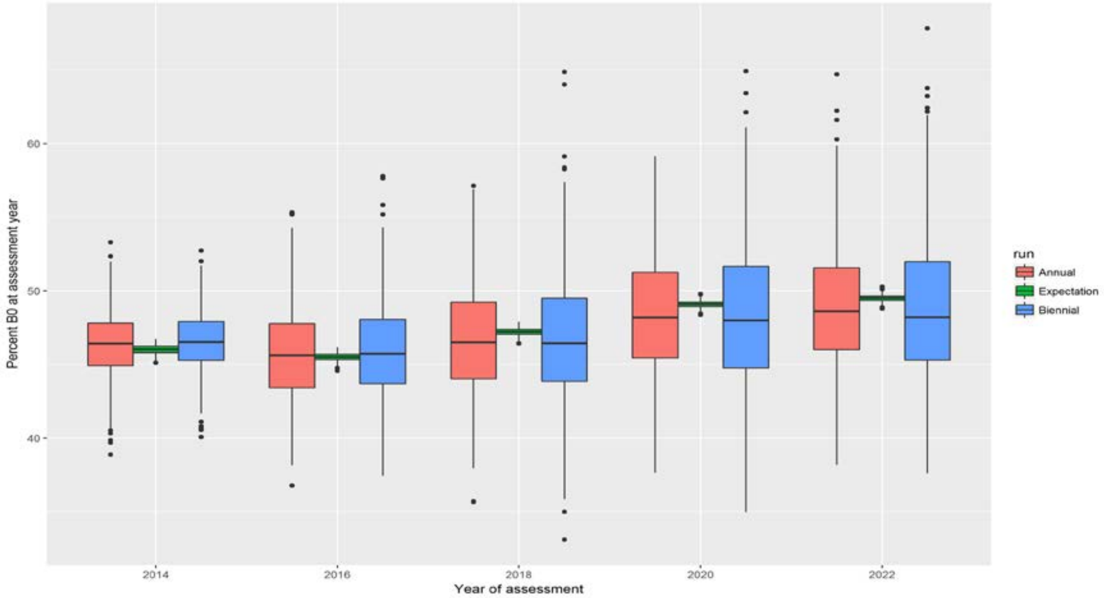

Figure 22: HAK 4 and stable trajectory for the operating model. Current biomass estimates ($\%B_0$) for forward simulation assessments. Shown are the estimates from 500 MPD fit scenarios: annual, expectation of the operating model, and biennial data.

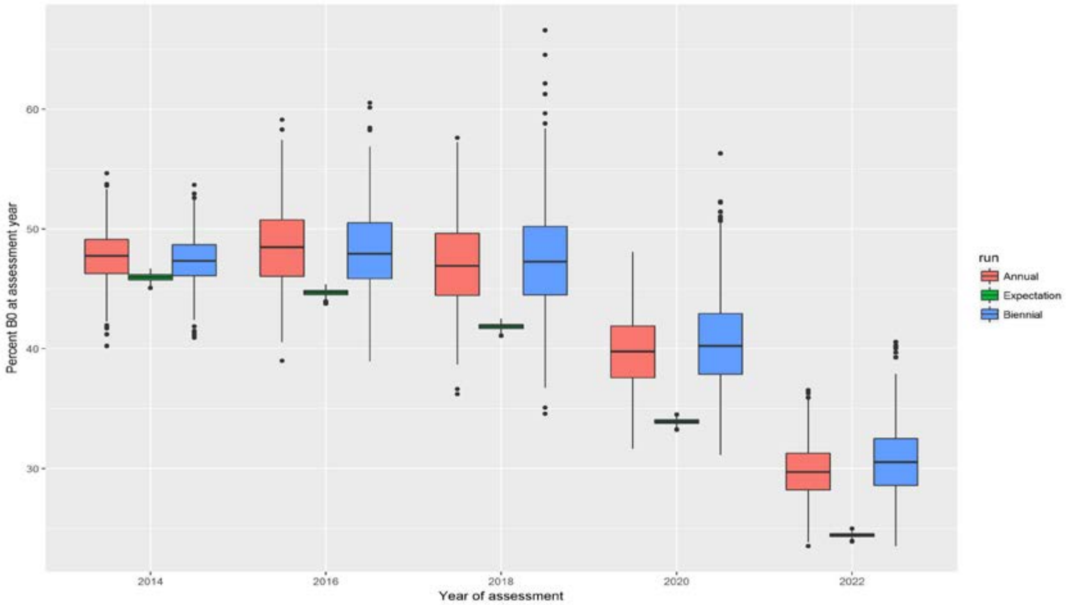

Figure 23: As in Figure 22, but for the decreasing trajectory.

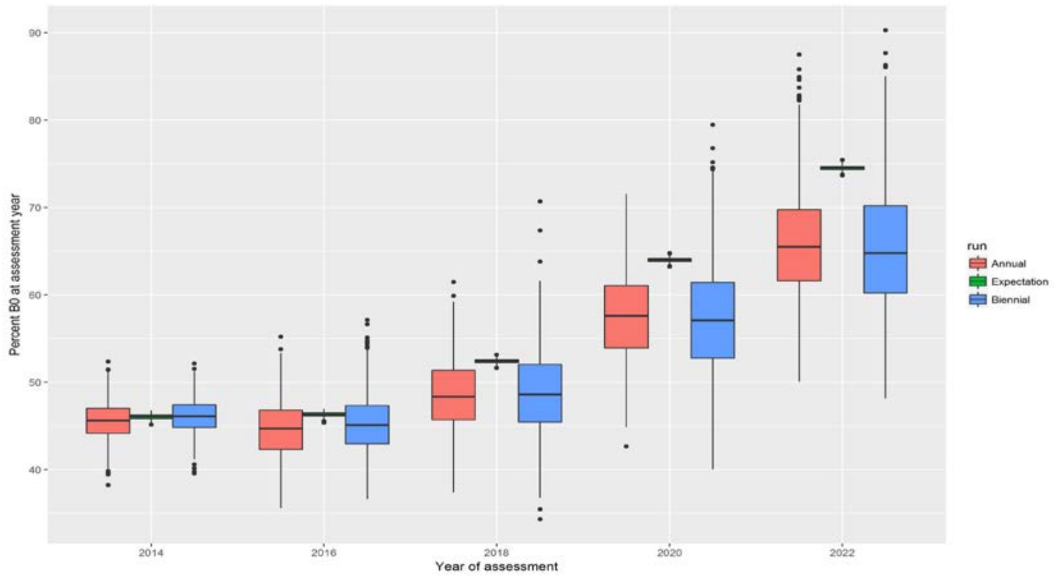

Figure 24: As in Figure 22, but for the increasing trajectory.

Table 9: HAK 4 and different trajectories for the operating model. Standard deviation of $\%B_0$ estimates for the annual and biennial frequency for each year of assessments.

Trajectory	Survey Frequency	2014	2016	2018	2020	2022
Stable	Annual	2.3	3.2	3.8	4.2	4.1
	Biennial	2.0	3.4	4.4	4.9	4.9
Decreasing	Annual	2.3	3.4	3.7	3.1	2.3
	Biennial	2.0	3.6	4.6	3.8	2.9
Increasing	Annual	2.3	3.3	4.2	5.3	6.2
	Biennial	2.0	3.4	4.9	6.1	7.3

3.3.3 LIN 3 & 4

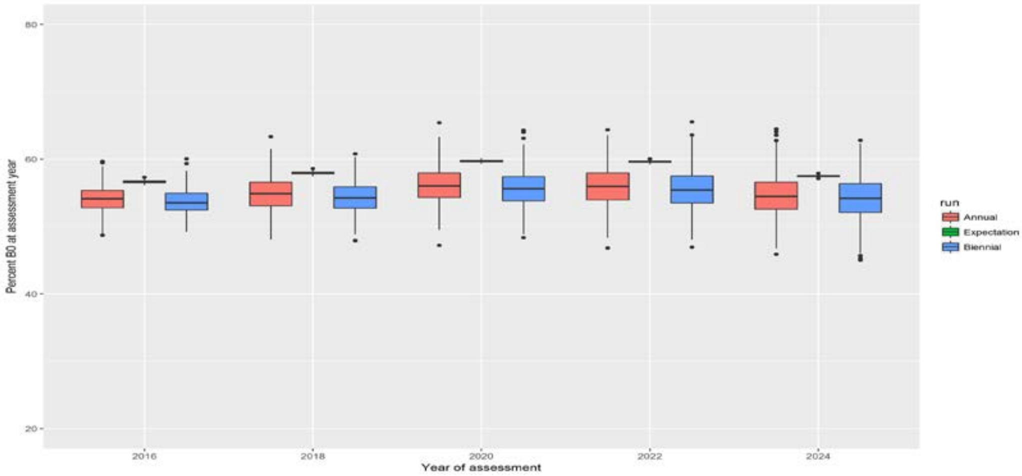

Figure 25: LIN 3 & 4 and stable trajectory for the operating model. Current biomass estimates (%B_0) for forward simulation assessments. Shown are the estimates from 500 MPD fit scenarios: annual, expectation of the operating model, and biennial data.

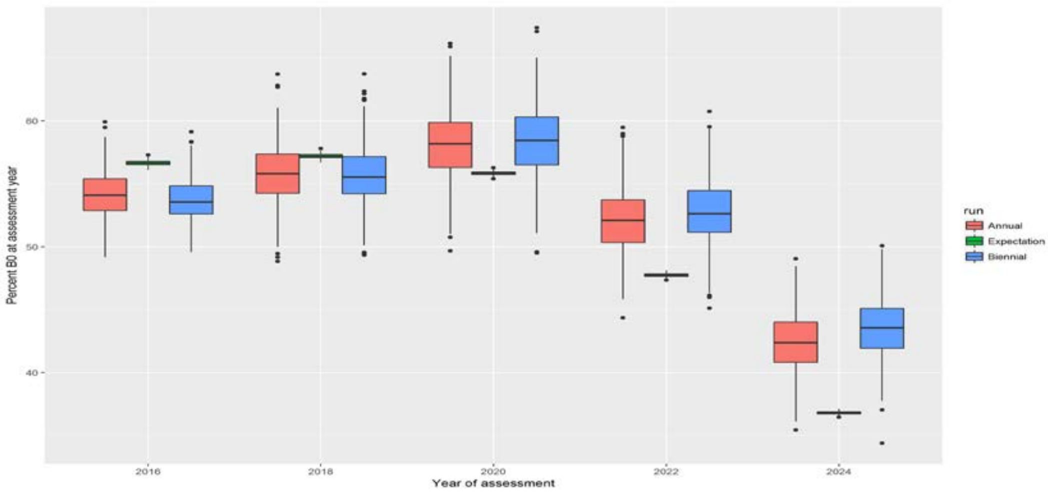

Figure 26: As in Figure 25, but for the decreasing trajectory.

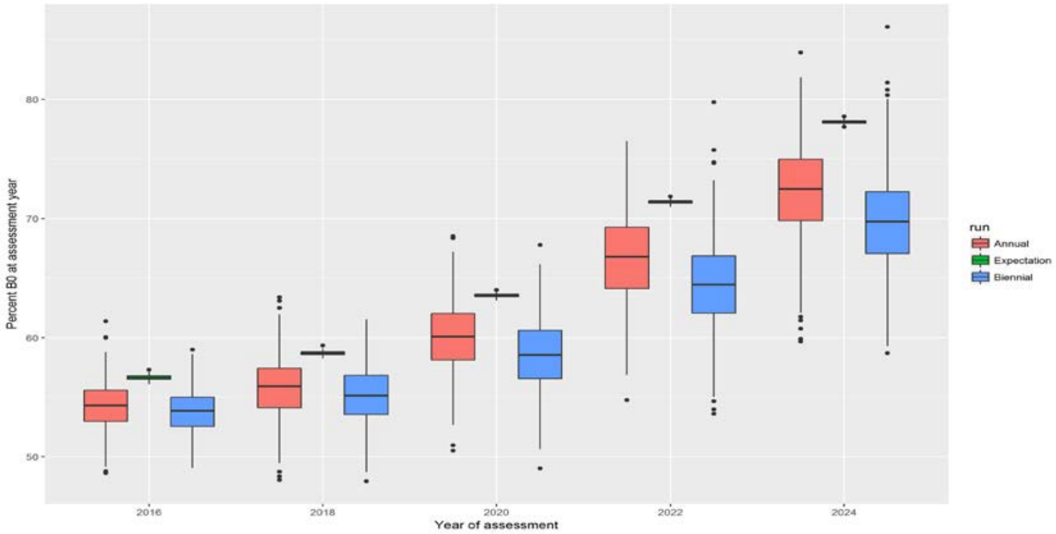

Figure 27: As in Figure 25, but for the increasing trajectory.

Table 10: LIN 3 & 4 and different trajectories for the operating model. Standard deviation of $\%B_0$ estimates between the annual and biennial frequency for each year of assessments.

Trajectory	Survey Frequency	2014	2016	2018	2020	2022
Stable	Annual	1.9	2.5	2.8	2.9	3.1
	Biennial	1.8	2.3	2.7	2.9	3.0
Decreasing	Annual	1.8	2.4	2.8	2.6	2.3
	Biennial	1.6	2.3	2.8	2.5	2.3
Increasing	Annual	1.9	2.6	3.1	3.8	4.0
	Biennial	1.8	2.3	2.9	3.6	4.0

3.3.4 LIN 5 & 6

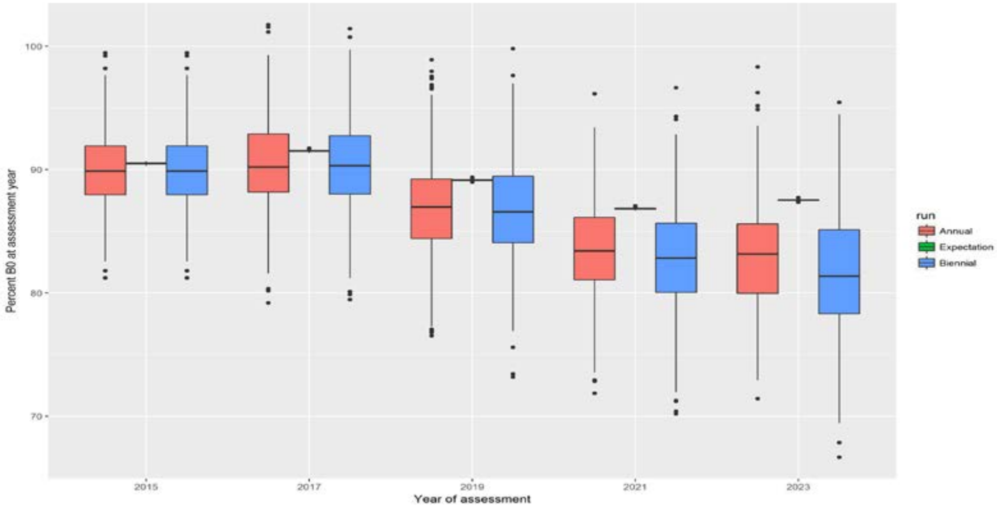

Figure 28: LIN 5 & 6 and stable trajectory for the operating model. Current biomass estimates (%B_0) for forward simulation assessments. Shown are the estimates from 500 MPD fit scenarios: annual, expectation of the operating model, and biennial data.

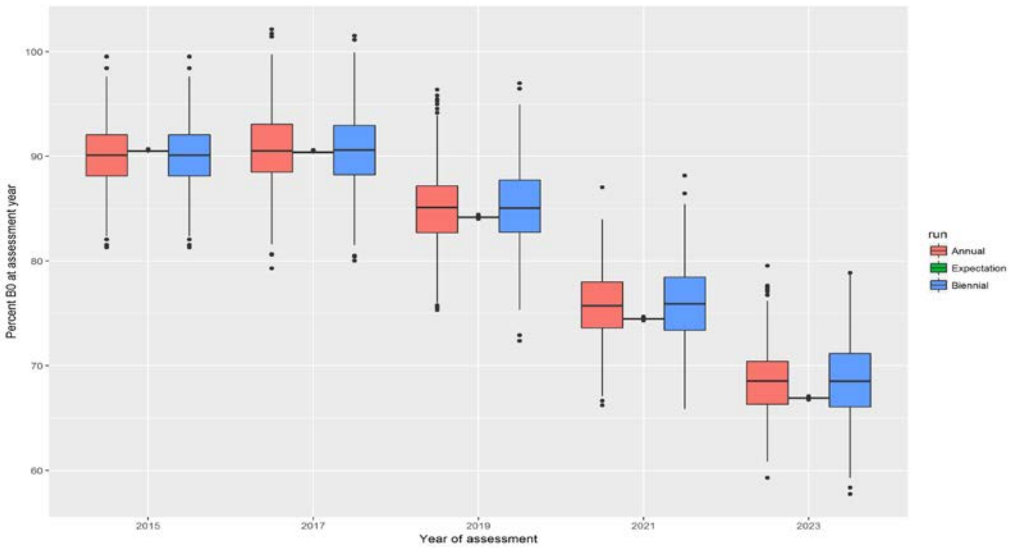

Figure 29: As in Figure 28, but for the decreasing trajectory.

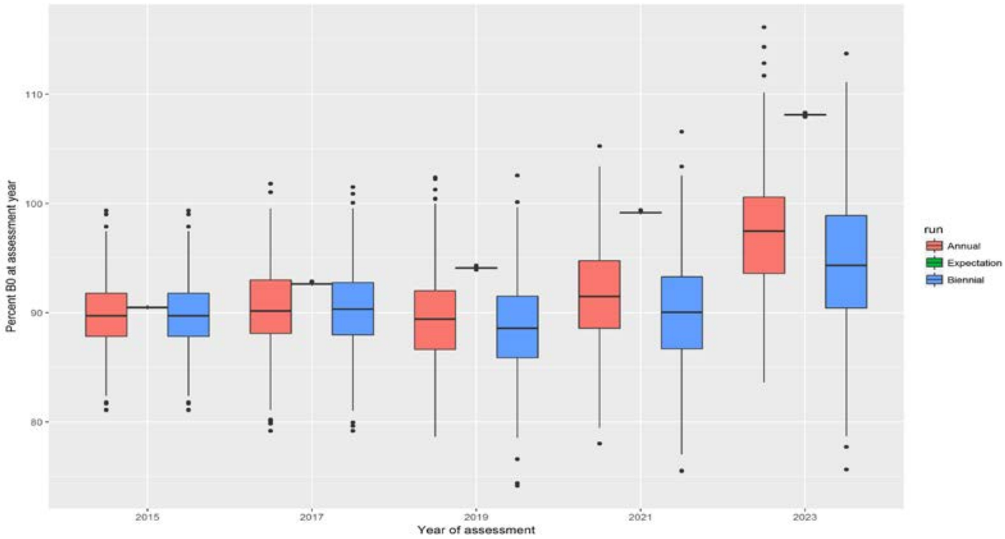

Figure 30: As in Figure 28, but for the increasing trajectory.

Table 11: LIN 5 & 6 and different trajectories for the operating model. Standard deviation of $\%B_0$ estimates between the annual and biennial frequency for each year of assessments.

Trajectory	Survey Frequency	2014	2016	2018	2020	2022
Stable	Annual	3.0	3.6	3.9	4.0	4.3
	Biennial	3.0	3.6	4.1	4.4	5.0
Decreasing	Annual	3.0	3.6	3.6	3.4	3.2
	Biennial	3.0	3.6	3.8	3.7	3.8
Increasing	Annual	3.0	3.6	4.1	4.7	5.4
	Biennial	3.0	3.6	4.3	5.1	6.2

3.3.5 HOK 1

For the HOK 1 stock, along with annual and biennial trawl survey scenarios, we ran another scenario whereby we dropped alternative fishery dependent age compositional data in alternating years to the survey. This meant if there was no survey there were fishery dependent age compositional data, and if there was a survey there were no fishery dependent age compositional data. This sensitivity run is shown as the purple box and whisker graph in the following figures. This is labelled 'Biennial Commercial Sensitivity' in the following figures (see Figures 31–36) and gave very similar biomass to the Biennial scenario, but with slightly more uncertainty.

3.3.5.1 HOK 1 Eastern Stock

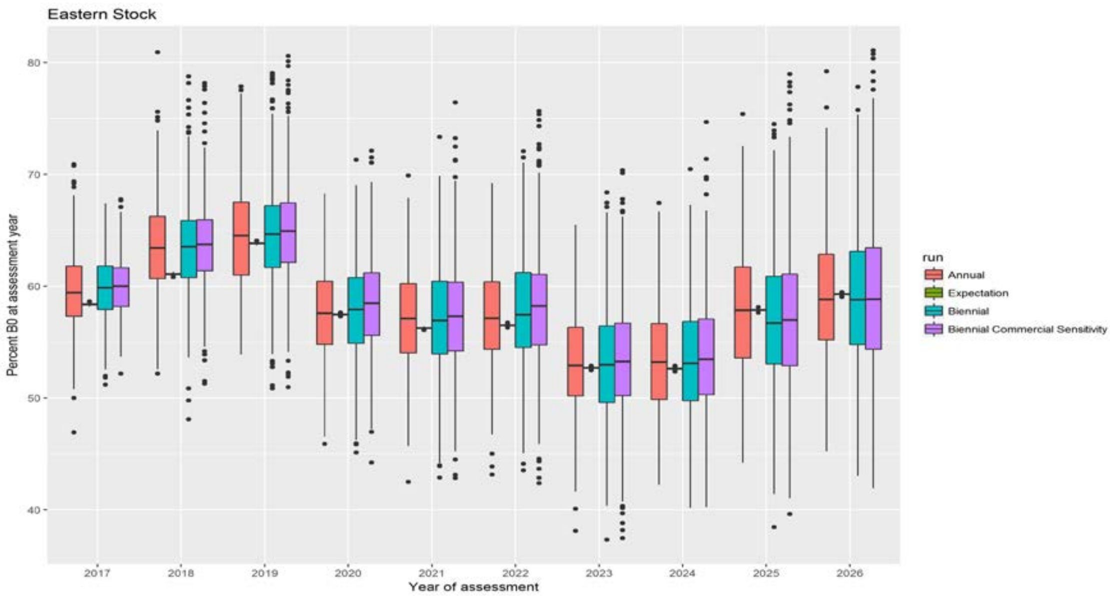

Figure 31: Eastern HOK 1 stock and stable trajectory for the operating model. Current biomass estimates ($\%B_0$) for forward simulation assessments. Shown are the estimates from 500 MPD fit scenarios: annual, expectation of the operating model, biennial data, and biennial with a commercial at-age data sensitivity.

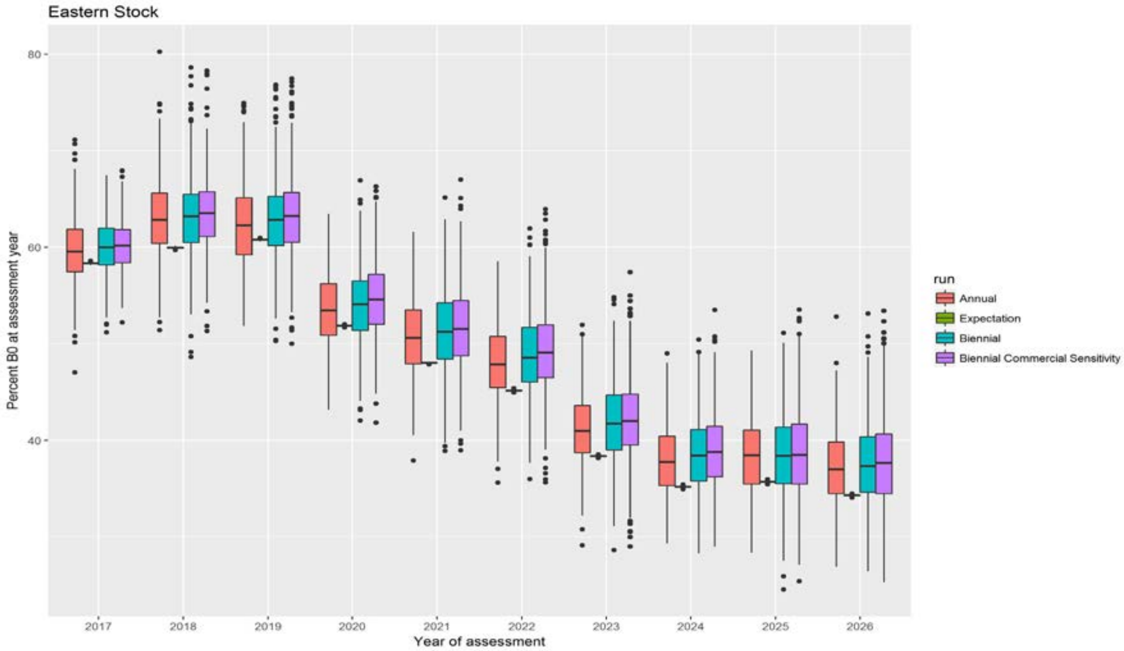

Figure 32: As in Figure 31, but for the decreasing trajectory.

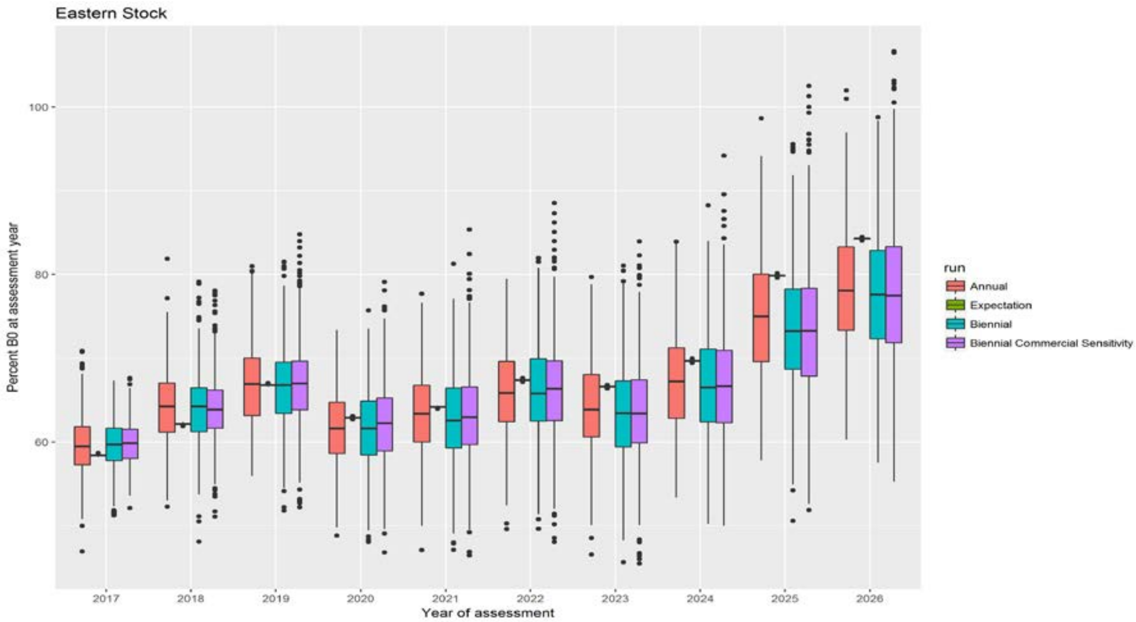

Figure 33: As in Figure 31, but for the increasing trajectory.

Table 12: Standard deviation of $\%B_0$ estimates between the annual and biennial frequency for each year of assessments for the stable trajectory.

Trajectory	Survey Frequency	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026
Stable	Annual	3.6	4.3	4.5	4.2	4.4	4.6	4.7	4.8	5.8	5.7
	Biennial	2.8	4.1	4.6	4.3	4.9	4.9	5.1	5.0	6.0	6.0
Decreasing	Annual	3.5	4.2	4.2	3.8	3.9	4.0	3.8	3.6	4.1	4.0
	Biennial	2.8	4.0	4.3	3.9	4.3	4.3	4.2	3.9	4.4	4.4
Increasing	Annual	3.6	4.5	4.9	4.6	5.1	5.3	5.6	6.0	7.5	7.3
	Biennial	2.8	4.3	4.9	4.7	5.4	5.5	5.9	6.1	7.4	7.5

3.3.5.2 HOK 1 Western Stock

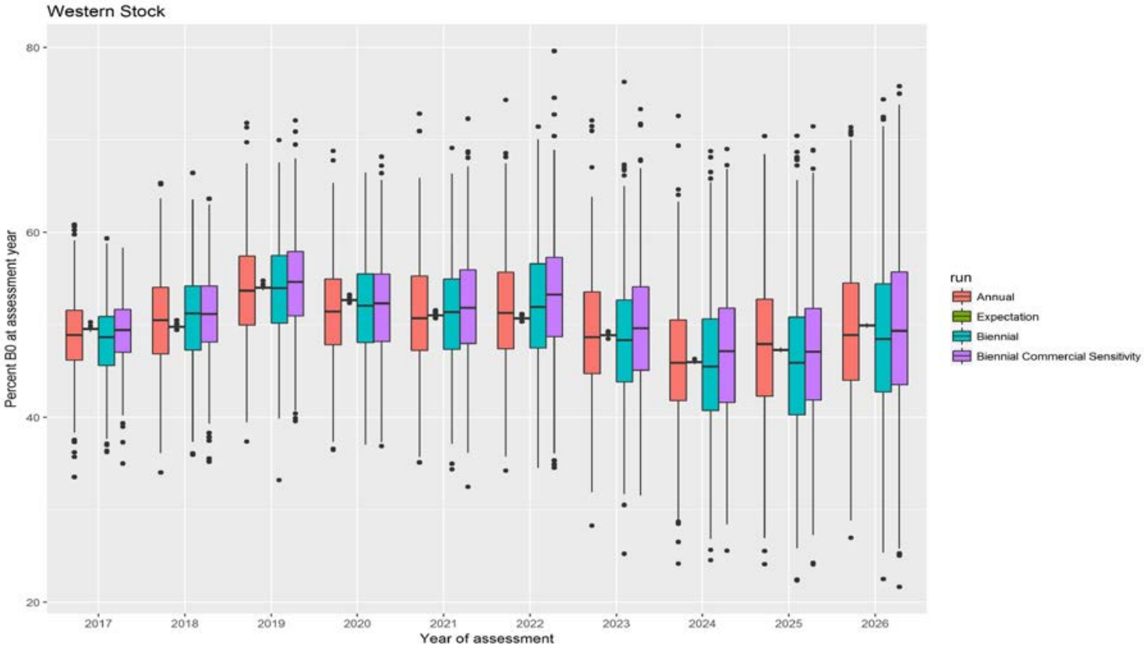

Figure 34: Western HOK 1 stock and stable trajectory for the operating model. Current biomass estimates (%B_0) for forward simulation assessments. Shown are the estimates from 500 MPD fit scenarios: annual, expectation of the operating model, biennial data, and biennial with a commercial at-age data sensitivity.

Table 13: Standard deviation of %B_0 estimates between the annual and biennial frequency for each year of assessments for the stable trajectory.

Trajectory	Survey Frequency	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026
Stable	Annual	4.5	5.2	5.5	5.4	5.9	6.2	6.5	6.9	7.5	7.6
	Biennial	4.3	5.2	5.3	5.4	5.9	6.6	6.9	7.4	8.1	8.5
Decreasing	Annual	4.4	5.0	5.2	5.0	5.3	5.5	5.5	5.4	5.7	5.6
	Biennial	4.2	5.0	5.1	5.0	5.4	5.8	5.8	5.9	6.2	6.4
Increasing	Annual	4.7	5.4	5.9	5.8	6.5	7.0	7.6	8.3	9.1	9.2
	Biennial	4.3	5.3	5.6	5.8	6.5	7.5	7.9	8.8	9.6	10.2

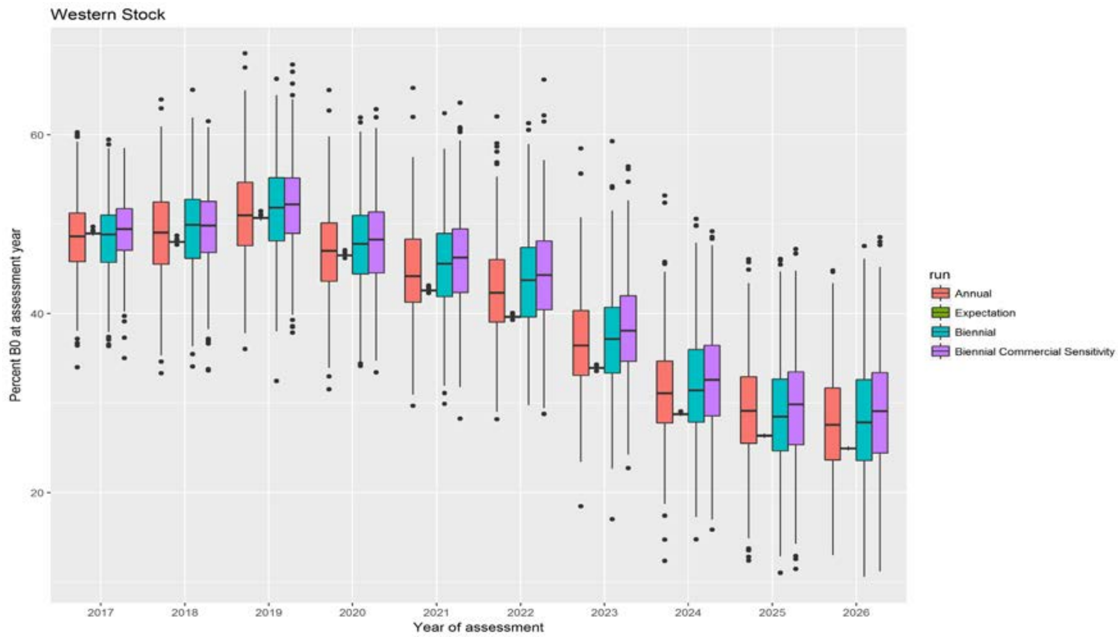

Figure 35: As in Figure 34, but for the decreasing trajectory.

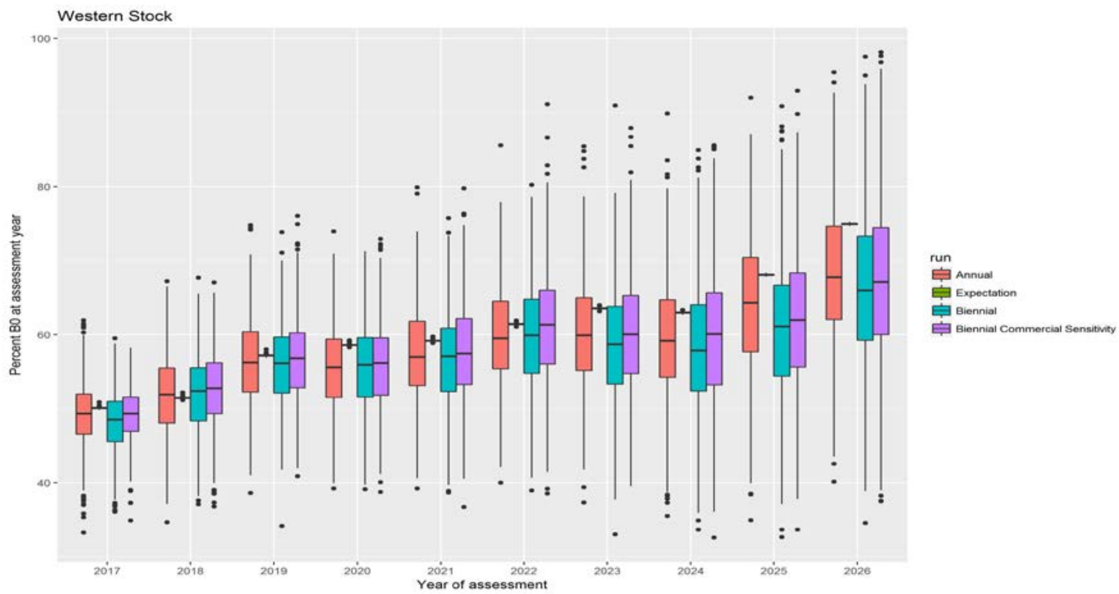

Figure 36: As in Figure 34, but for the increasing trajectory.

4. DISCUSSION

In summary, the generic simulations demonstrated that under very simple scenarios that you expect to get greater variance but no bias in biomass estimation when comparing annual observations versus biennial observations. For the retrospective simulations, there were small differences in biomass estimates between annual and biennial scenarios (both median and uncertainty), although for the hoki stock there were some differences in biomass estimates between the two biennial scenarios, which is thought to be attributed to an induced three-year gap between surveys. For the forward projection analysis, differences between annual and biennial scenario were again small, with the annual scenario biomass estimates closer to the expected values.

In this study, the generic simulation was essentially used for illustration. The retrospective analysis used actual observations, which made no assumptions on the structure of the data, compared to simulated data. Another advantage of retrospective analyses is that MCMC estimations could be done, which arguably explores parameter uncertainty and thus derived quantity uncertainty better, and is more reflective of uncertainty that would propagate and inform management decisions. A disadvantage is that it ignores data that could inform key productivity parameters in future assessments.

For the retrospective analyses, the largest difference from the annual scenario was seen for the hoki assessment under biennial 2 scenario (see Figures 17–18). There were small differences for the current reference quantity $B_{current}$, and when we projected the population forward that difference was magnified for B_{future}. We hypothesise that this was caused by the three year gap that occurred due to there being no Sub-Antarctic trawl survey in 2011. Other than that one case, there were relatively small differences between the annual and biennial survey frequencies in the retrospective analysis.

The forward projection analyses had the opposite advantages and disadvantages compared to the retrospective analysis. They retained all data for current assessments, and made assumptions for future data. However, due to the number of simulations needed to generate an informative range of observed data, it was not practical to do MCMC estimations. Assumptions that were imposed by the operating model were consistent with stock assessment model theory. One assumption of the operating model was that productivity parameters were constant through time (time-invariant) with selectivity parameters being the only time-varying parameters. Although these were strong assumptions, a great deal of effort was put into generating simulated data, consistent with residuals that have been generated from real data and stock assessment fits. Our defence for imposing these assumptions is that the objective of the operating model was to simulate “realistic” data. There are many mechanisms for generating “realistic” data and we used the combination of process error and observation error to add plausible variance and correlation in the operating model and thereby in the simulated data.

For the forward projection analysis, for some stocks and population trajectories the annual scenario re-estimated stock status was on average closer to the expectation of the operating model than the biennial scenario (defined as having a median %B_0 that was 5% closer to the operating model). These stocks and population trajectories were HOK 1 western stock increasing, LIN 5&6 increasing, LIN 3&4 increasing and decreasing and HAK 1 decreasing. However, overall there was little difference between the annual and biennial scenarios, and a more comprehensive investigation would have to be conducted to understand why there were differences for specific stocks and trajectories. The biggest surprise was the minor increase in uncertainly in going from annual to biennial surveys.

Finally, another consideration that merits a mention for the forward simulation component is the mechanism that was used to alter the state of future populations. This used a change in recruitment, and consequently we observed gradual change in biomass estimates because it takes two to three years for new cohorts to emerge into observations. An alternative mechanism for altering the state of future populations could be more abrupt or immediate changes, such as a disease or catchability event, that would have quite different characterisations in the data and may yield different results.

5. ACKNOWLEDGEMENTS

We would like to thank the Deepwater Working Group for useful discussions and direction during the project. This work was funded under Ministry for Primary Industries project DEE201502. Thank you to Charlie Edwards and Matt Dunn for reviewing the manuscript.

6. REFERENCES

Aitchison, J. (2003). The Statistical Analysis of Compositional Data. The Blackburn Press, Caldwell, New Jersey (this is a revision and updating of the 1986 book of the same authorship and title published by Chapman and Hall).

Bull, B.; Francis, R.I.C.C.; Dunn, A.; McKenzie, A.; Gilbert, D.J.; Smith, M.H.; Bian, R.; Fu, D. (2012). CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.30-2012/03/21. *NIWA Technical Report 135*. 280 p.

Doonan, I.; Large, K.; Dunn, A.; Rasmussen, S.; Marsh, C.; Mormede, S. (2016). Casal2: New Zealand's integrated population modelling tool. *Fisheries Research 183:* 498–505.

Francis, R.I.C.C.; Horn, P.L. (2005). Information relevant to the possible continuation of the winter series of east coast South Island trawl surveys. Note to the Inshore Working Group 8-8-05. INS-WG-2005/30.

Francis, R.I.C.C. (2011a). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences 68*: 1124–1138.

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

Horn, P.L. (2013). Stock assessment of hake (*Merluccius australis*) on the Chatham Rise (HAK 4) and off the west coast of South Island (HAK 7) for the 2012–13 fishing year. *New Zealand Fisheries Assessment Report 2013/31*. 58 p.

Horn, P.L. (2015). Stock assessment of hake (*Merluccius australis*) in the Sub-Antarctic (part of HAK 1) for the 2014–15 fishing year. *New Zealand Fisheries Assessment Report 2015/29*. 55 p.

McGregor, V. (2015). Stock assessment of ling (*Genypterus blacodes*) on the Chatham Rise (LIN 3&4) for the 2014–15 fishing year. *New Zealand Fisheries Assessment Report 2015/82*. 54 p.

McKenzie, A. (2016). Assessment of hoki (*Macruronus novaezelandiae*) in 2015. *New Zealand Fisheries Assessment Report 2016/01*. 88 p.

Ministry for Primary Industries (2017). Fisheries Assessment Plenary, May 2017: stock assessments and stock status. Compiled by the Fisheries Science Group, Ministry for Primary Industries, Wellington, New Zealand: https://fs.fish.govt.nz/Page.aspx?pk=113&dk=24474.

O'Driscoll, R.L.; Ladroit, Y.; Dunford, A.J.; MacGibbon, D.J. (2016). Acoustic survey of spawning hoki in Cook Strait during winter 2015 and update of acoustic q priors for hoki stock assessment modelling. *New Zealand Fisheries Assessment Report 2016/44*. 55 p.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roberts, J. (2016). Stock assessment of ling (*Genypterus blacodes*) in the Sub-Antarctic (LIN 5&6) for the 2014–15 fishing year. *New Zealand Fisheries Assessment Report 2016/05*. 39 p

7. APPENDIX A (Analysis of autocorrelation for trawl survey data)

Autocorrelations for the trawl survey data were investigated by looking at lagged correlations for the normalised residuals for the base model fits. Both autocorrelation and partial autocorrelation were calculated (using the *acf* and *pacf* functions from the R stats package). Although some auto-correlations and partial autocorrelations were apparently significant, there was no systematic pattern across the surveys, and it was decided not to include autocorrelation in the simulations (Figures 37–38).

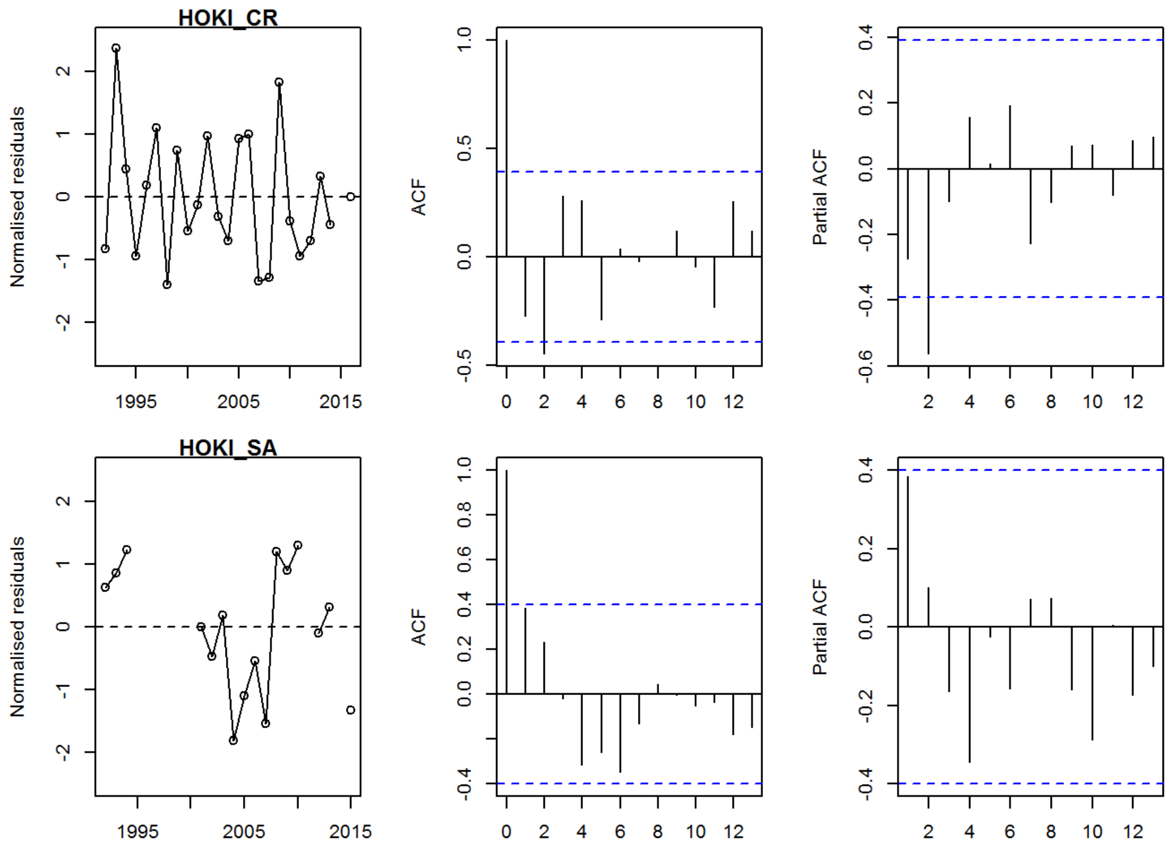

Figure 37: Hoki assessment auto-correlation analysis for the Chatham Rise and Sub-Antarctic trawl surveys. Normalised residuals (left panel), auto-correlation function versus lag (middle panel), and partial auto-correlation function versus lag (right panel). The blue dashed lines represent approximate 95% confidence intervals for the calculated values when it is assumed there is no correlation.

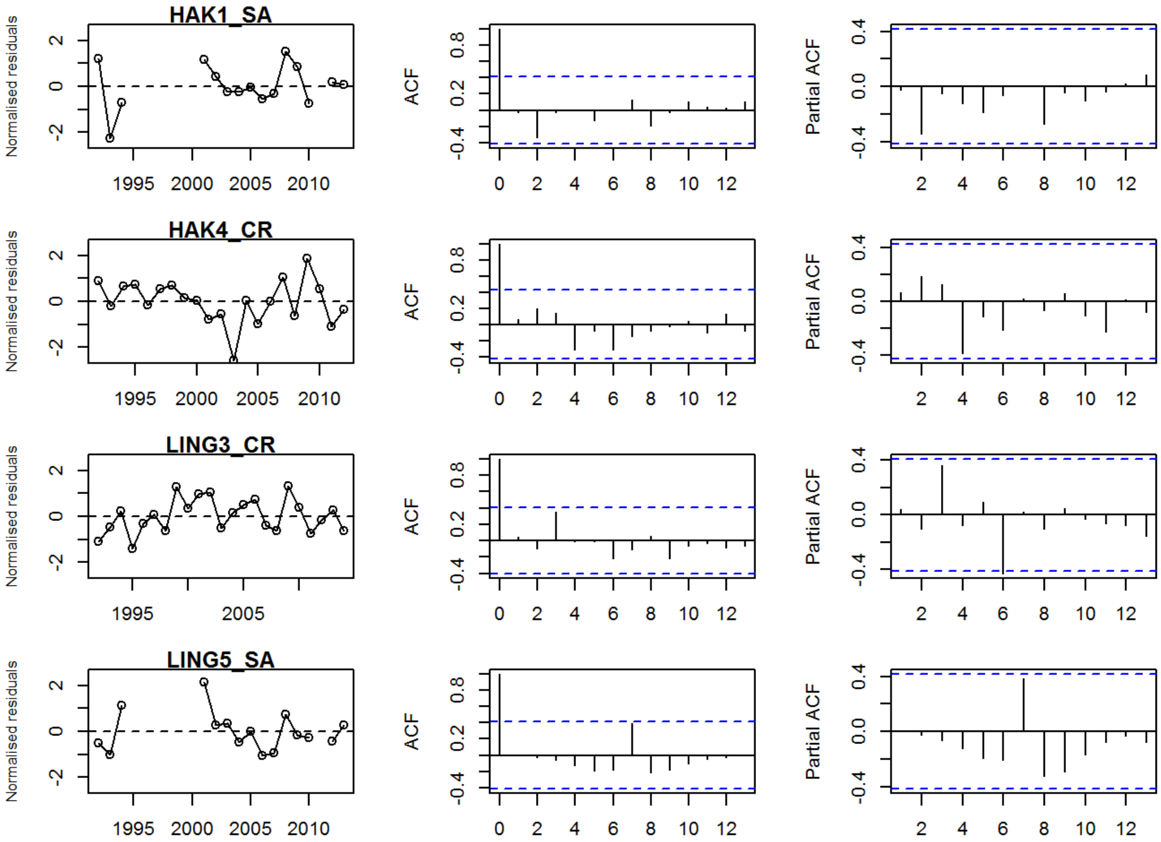

Figure 38: Hake and ling assessments auto-correlation analysis for the Chatham Rise and Sub-Antarctic trawl surveys. Normalised residuals (left panel), auto-correlation function versus lag (middle panel), and partial auto-correlation function versus lag (right panel). The blue dashed lines represent approximate 95% confidence intervals for the calculated values when it is assumed that there is no correlation.

8. APPENDIX B (Forward simulation)

Diagnostics from the operating models used to simulate data for re-estimation.

8.1 HOK 1

A unique element of the hoki assessment model is that the western spawning fishery selectivity already time varies, based on the exogenous variable the median fishing day (McKenzie 2016). This makes it different from other future selectivities which we allow to vary randomly. For this specific selectivity we assumed that the future was randomly drawn as described in the method, but to identify the correct amount of process error we removed the exogenous variable. In the final operating model the exogenous variable was retained, but with future value assumed to be normally distributed.

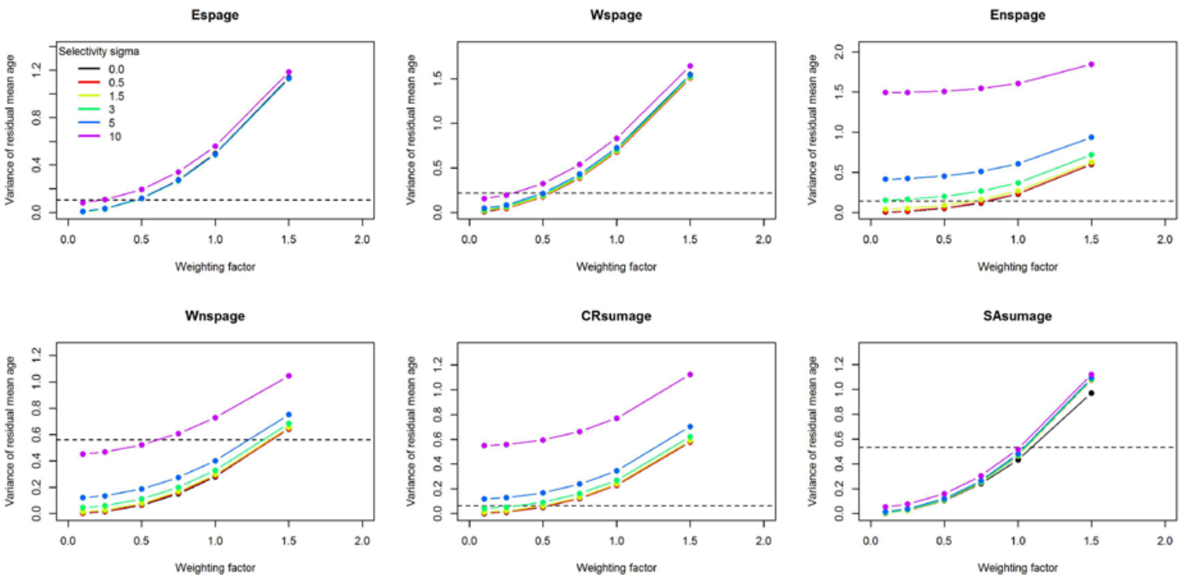

Figure 39: Residual mean age variance plotted for combinations of σ_s and ω_i for each compositional dataset used in the final operating model.

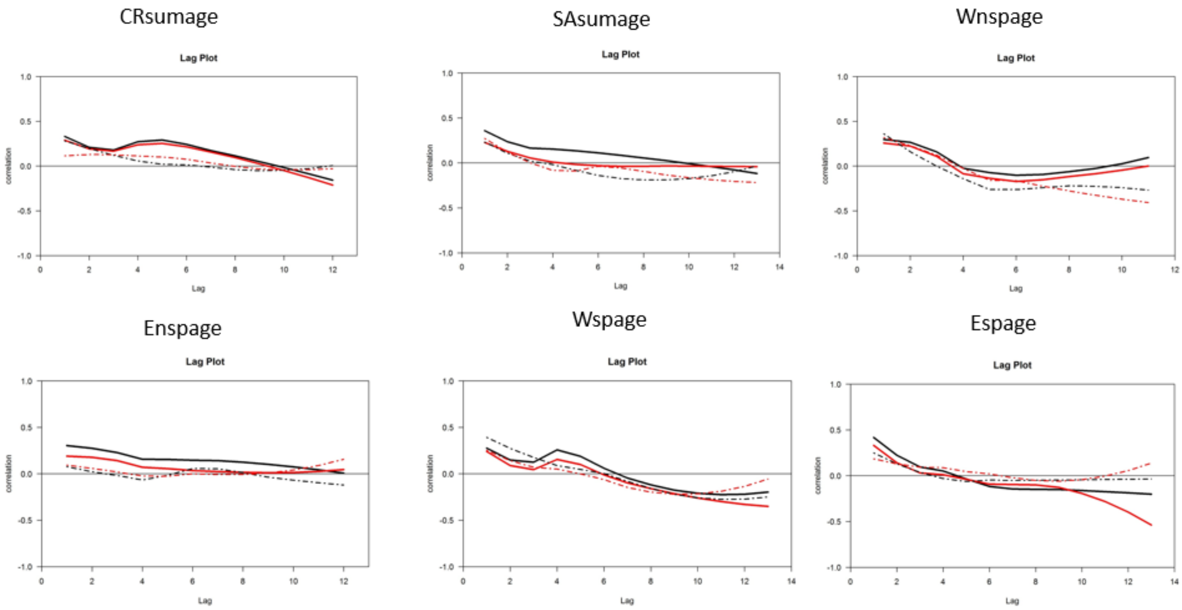

Figure 40: Residual correlation by age (Lag, x-axis) and sex (black = male and female = red) of the base case stock assessment and true observations (dashed lines) and the average residual correlation from 50 simulated datasets when compared with the base stock assessment expectation (solid lines).

Table 14: The combination of process error (σ_s), weighting factors (ω_i) and correlation parameters used in the final operating model for each observation and selectivity.

Selectivity	σ_s	Observation	ω_i	ρ_1	ρ_2
Wnsp fishery	10	Wnspage	0.616	0.18	0.37
Ensp fishery	0.5	Enspage	0.793	0.17	0.32
Wsp fishery	5	Wspage	0.502	0.45	0.14
Esp fishery	5	Espage	0.459	0.74	-0.50
CR survey	0.5	CRsumage	0.536	0.80	-0.61
SA survey	10	SAsumage	1.05	0.19	0.18

8.2 HAK 1

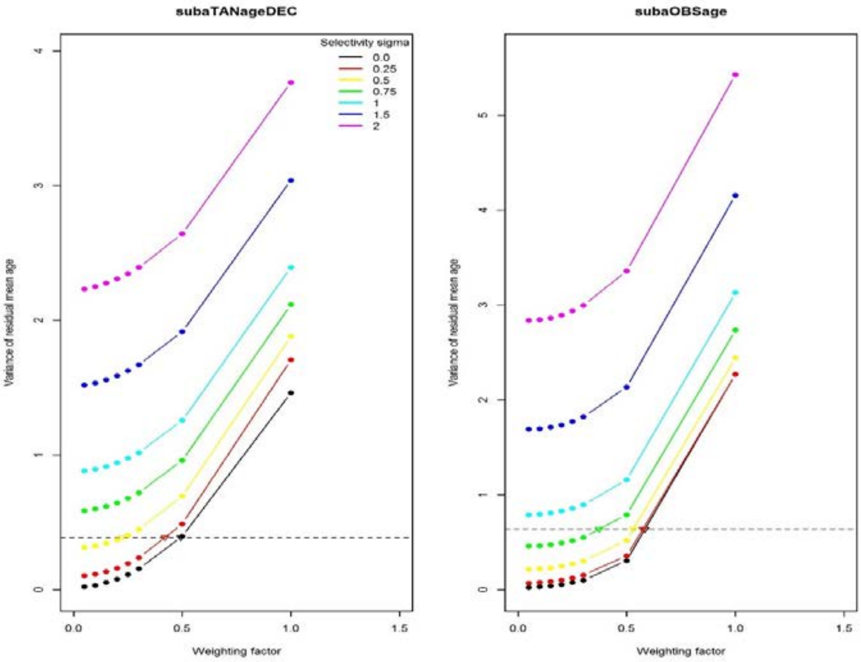

Figure 41: Residual mean age variance plotted for combinations of σ_s (separate lines) and ω_i (x-axis) for each compositional dataset used in the final operating model.

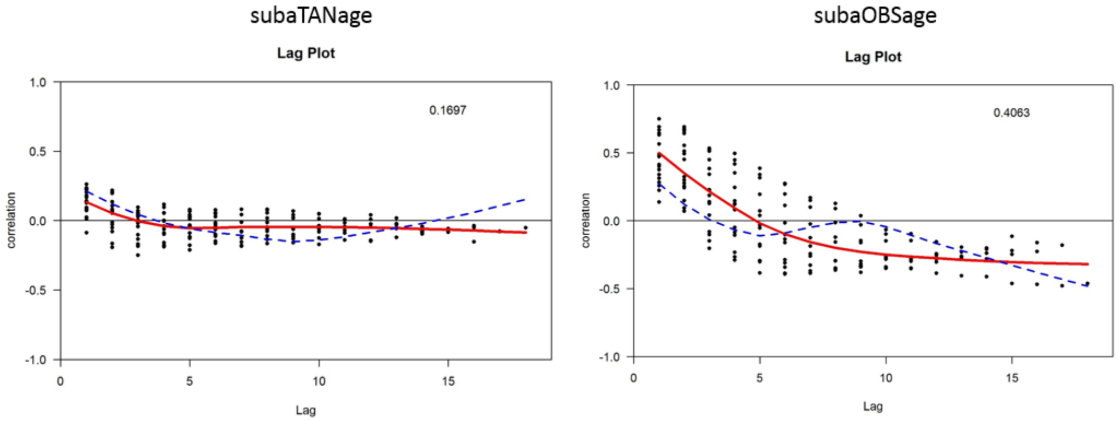

Figure 42: Residual correlation by age (Lag, x-axis) of the base case stock assessment and true observations (dashed blue lines) and the average residual correlation from 50 simulated datasets when compared with the base stock assessment expectation (solid red lines). The points are the averaged residuals from the simulated data when compared to the base case expectation.

Table 15: The combination of process error (σ_s), weighting factors (ω_i) and correlation parameters used in the final operating model for each observation and selectivity.

Selectivity	σ_s	Observation	ω_i	ρ_1	ρ_2
Survey	0.25	subaTANage	0.454	0.22	0.10
Fishery	0.25	subaOBSage	0.588	0.49	0.46

8.3 HAK 4

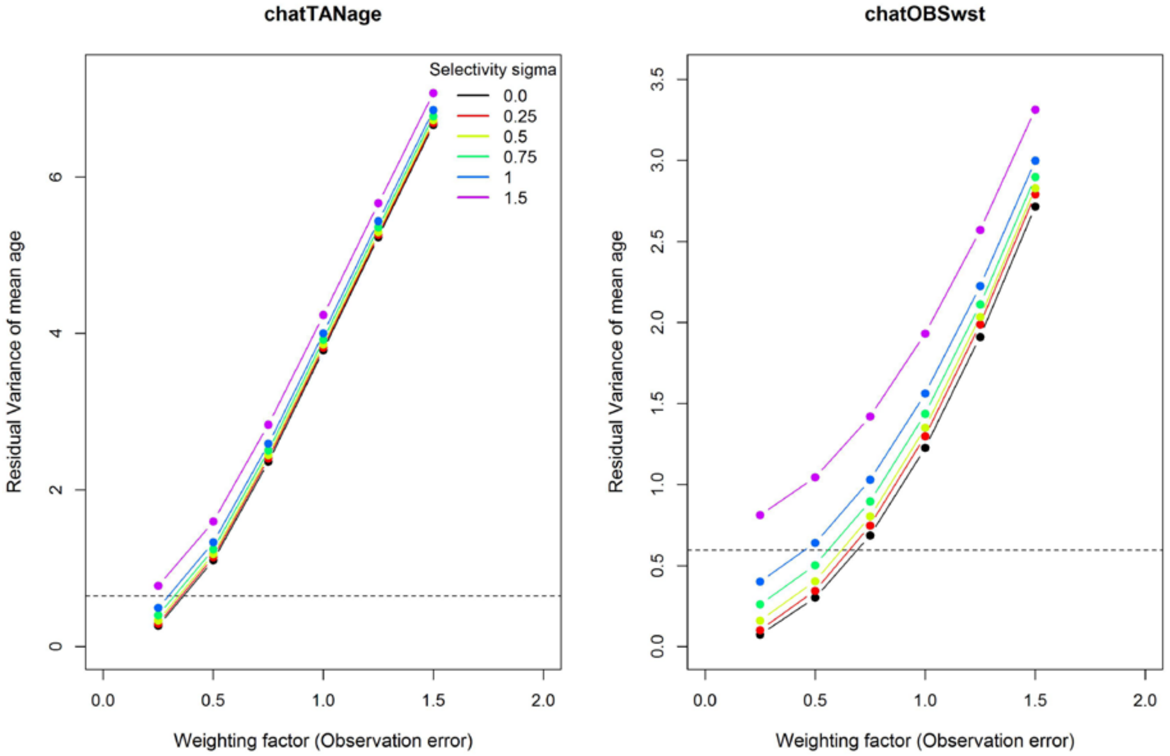

Figure 43: Residual mean age variance plotted for combinations of σ_s (separate lines) and ω_i (x-axis) for each compositional dataset used in the final operating model.

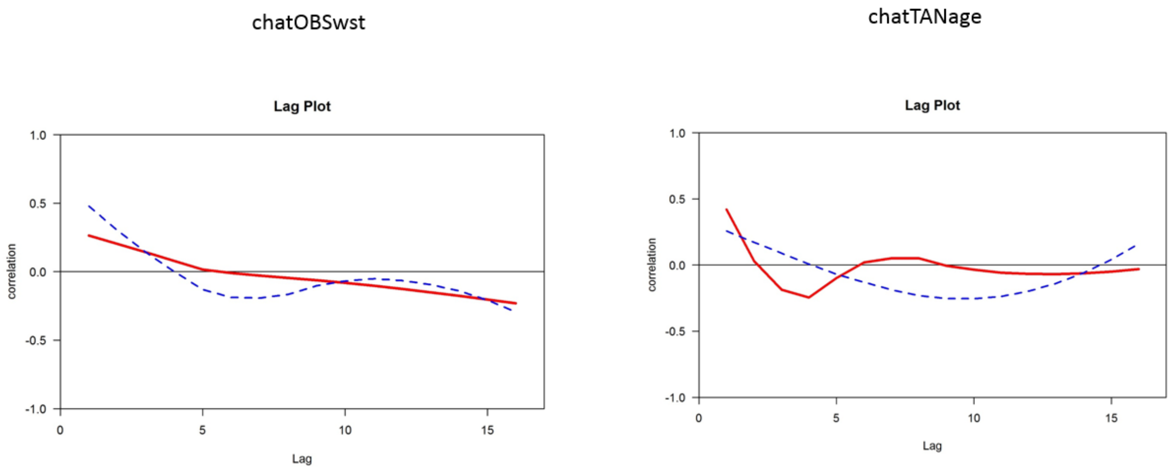

Figure 44: Residual correlation by age (Lag, x-axis) of the base case stock assessment and true observations (dashed blue lines) and the average residual correlation from 50 simulated datasets when compared with the base stock assessment expectation (solid red lines).

Table 16: The combination of process error (σ_s), weighting factors (ω_i) and correlation parameters used in the final operating model for each observation and selectivity.

Selectivity	$\boldsymbol{\sigma_s}$	**Observation**	$\boldsymbol{\omega_i}$	$\boldsymbol{\rho_1}$	$\boldsymbol{\rho_2}$
Fishery	0.75	chatOBSwst	0.572	0.22	0.38
Survey	0.5	chatTANage	0.623	0.97	-0.63

8.4 LIN 5 & 6

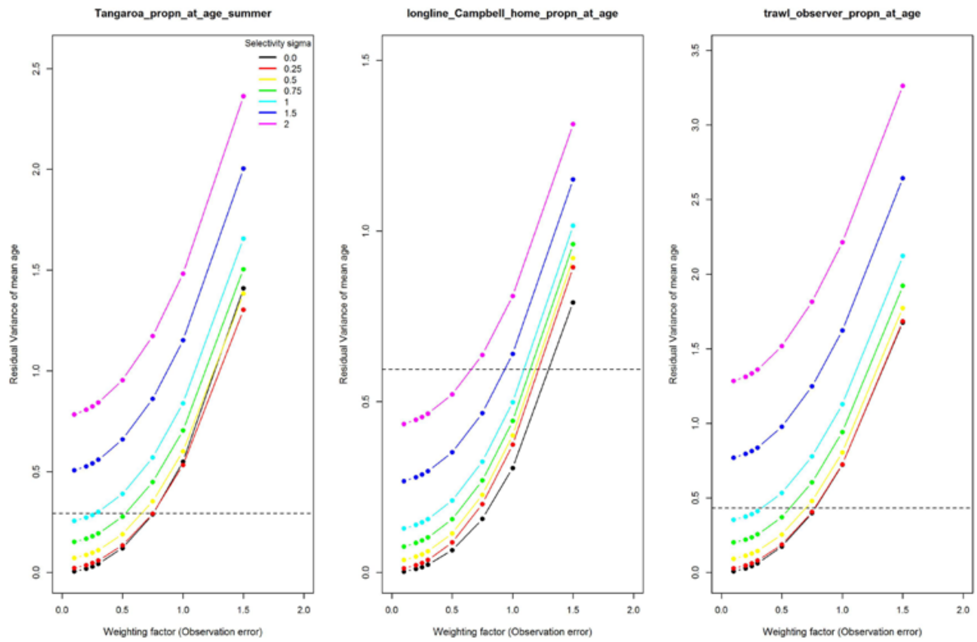

Figure 45: Residual mean age variance plotted for combinations of σ_s (separate lines) and ω_i (x-axis) for each compositional dataset used in the final operating model.

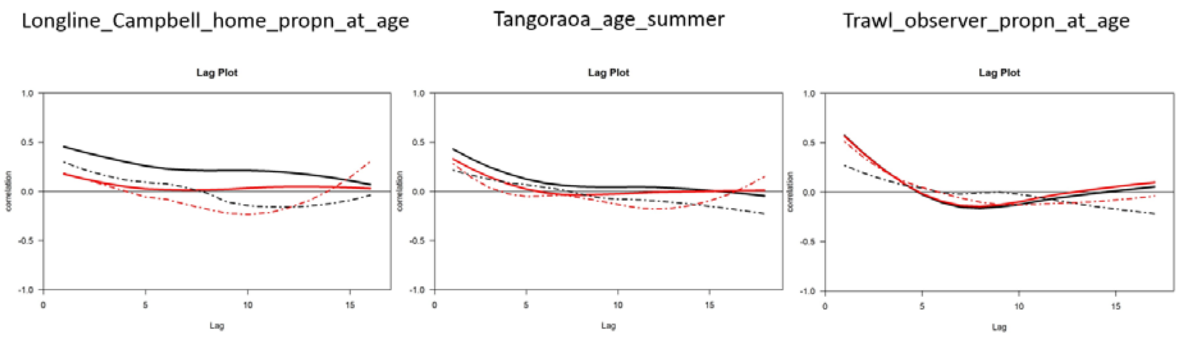

Figure 46 Residual correlation by age (Lag, x-axis) and sex (black = male and female = red) of the base case stock assessment and true observations (dashed lines) and the average residual correlation from 50 simulated datasets when compared with the base stock assessment expectation (solid lines).

Table 17: The combination of process error (σ_s), weighting factors (ω_i) and correlation parameters used in the final operating model for each observation and selectivity.

Selectivity	σ_s	**Observation**	ω_i	ρ_1	ρ_2
Survey	0.25	Tangoraoa_age_summer	0.751	0.79	-0.55
Longline Fishery	0.25	Longline_Campbell_home_propn_at_age	1.213	0.064	0.13
Trawl Fishery	0.5	Trawl_observer_propn_at_age	0.697	0.80	-0.32

8.5 LIN 3 & 4

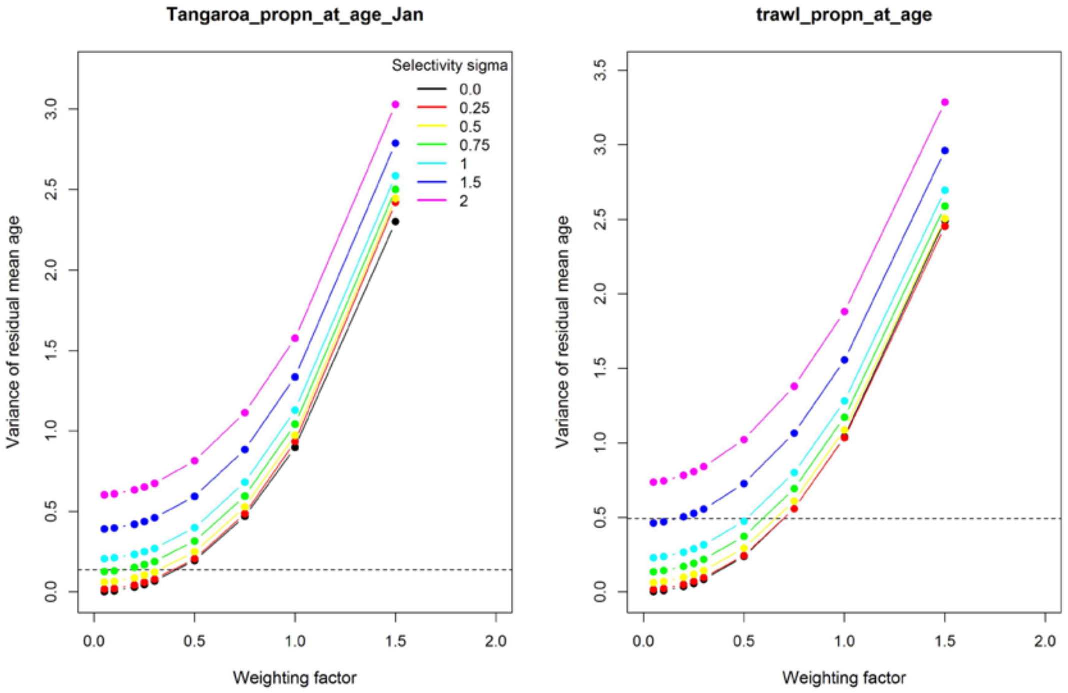

Figure 47: Residual mean age variance plotted for combinations of σ_s (separate lines) and ω_i (x-axis) for each compositional dataset used in the final operating model.

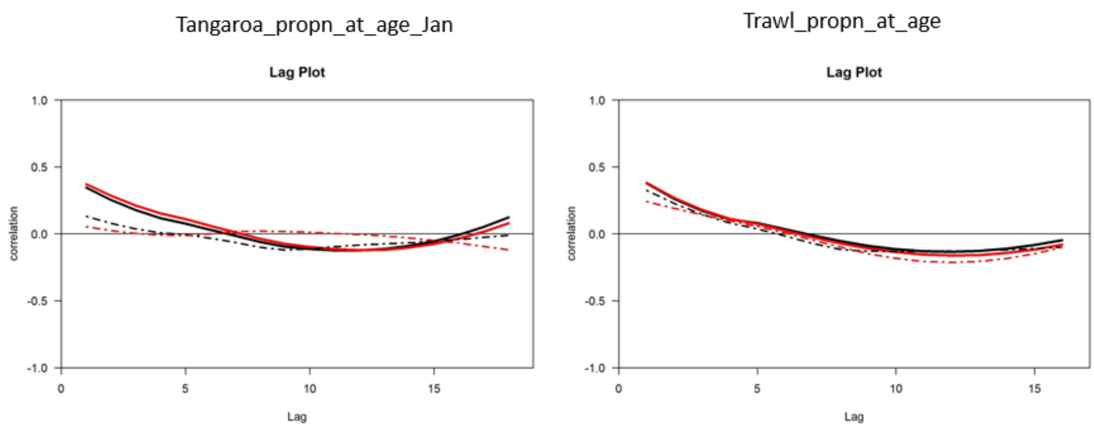

Figure 48 Residual correlation by age (Lag, x-axis) and sex (black = male and female = red) of the base case stock assessment and true observations (dashed lines) and the average residual correlation from 50 simulated datasets when compared with the base stock assessment expectation (solid lines).

Table 18: The combination of process error (σ_s), weighting factors (ω_i) and correlation parameters used in the final operating model for each observation and selectivity.

Selectivity	σ_s	**Observation**	ω_i	ρ_1	ρ_2
Survey	0.25	Tangaroa_propn_at_age_Jan	0.391	0.69	-0.58
Fishery	0.75	Trawl_propn_at_age	0.593	0.81	-0.53

9. APPENDIX C (MCMC diagnostics for the retrospective MCMCs)

9.1 HOK 1

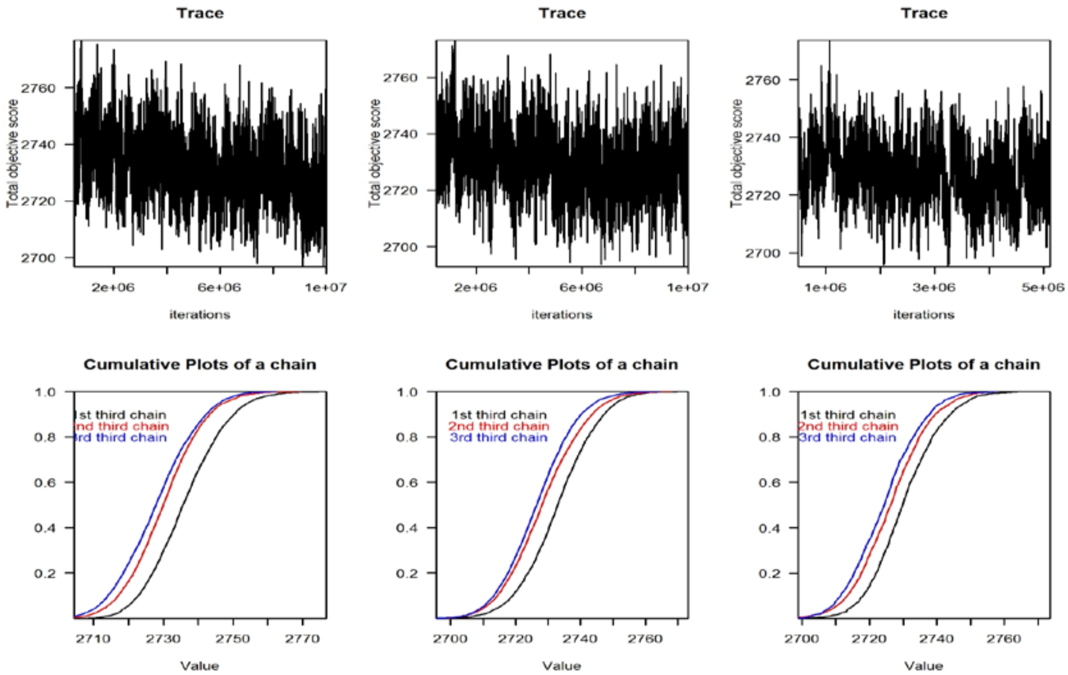

Figure 49: Trace and cumulative plots for the Annual scenario of the total objective score. Each column is an independent MCMC chains, the cumulative plots (bottom right) have chopped each chain into thirds and plotted up over each other.

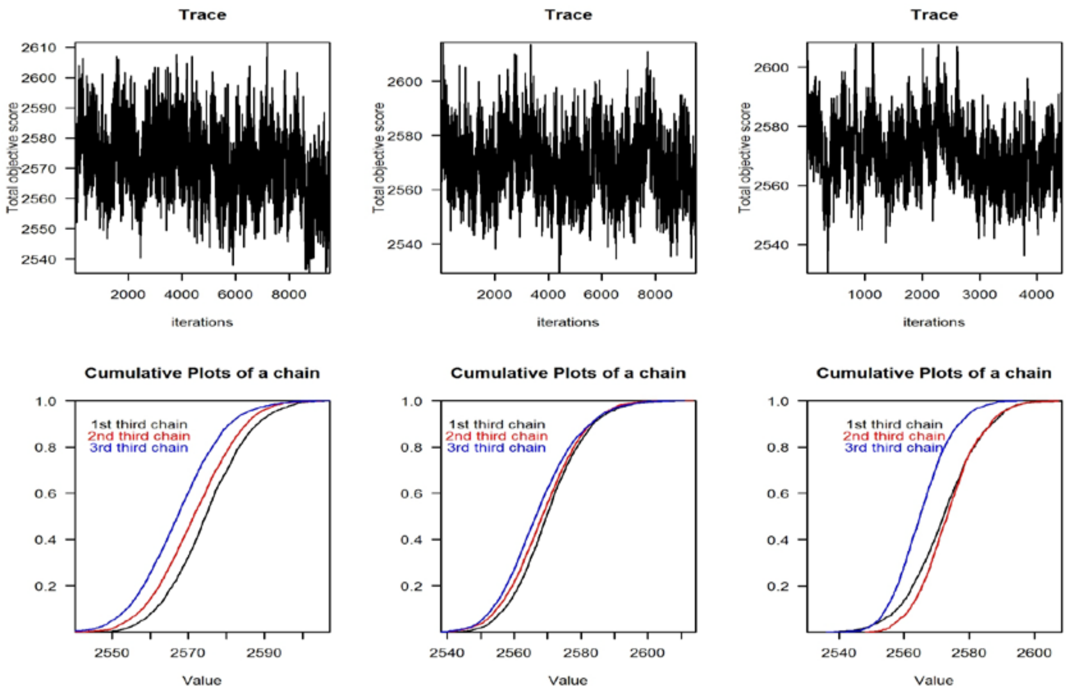

Figure 50: Trace and cumulative plots for the Biennial 1 scenario of the total objective score. Each column is an independent MCMC chains, the cumulative plots (bottom right) have chopped each chain into thirds and plotted up over each other.

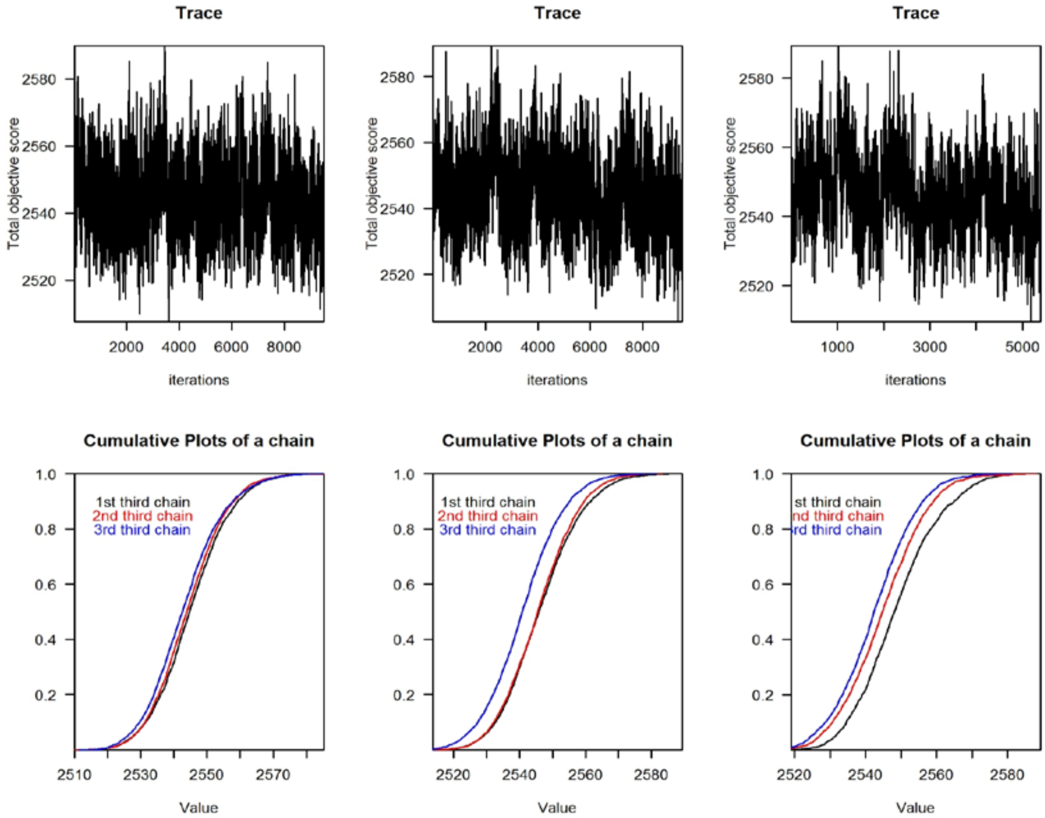

Figure 51: Trace and cumulative plots for the Biennial 2 scenario of the total objective score. Each column is an independent MCMC chains, the cumulative plots (bottom right) have chopped each chain into thirds and plotted up over each other.

9.2 HAK 1

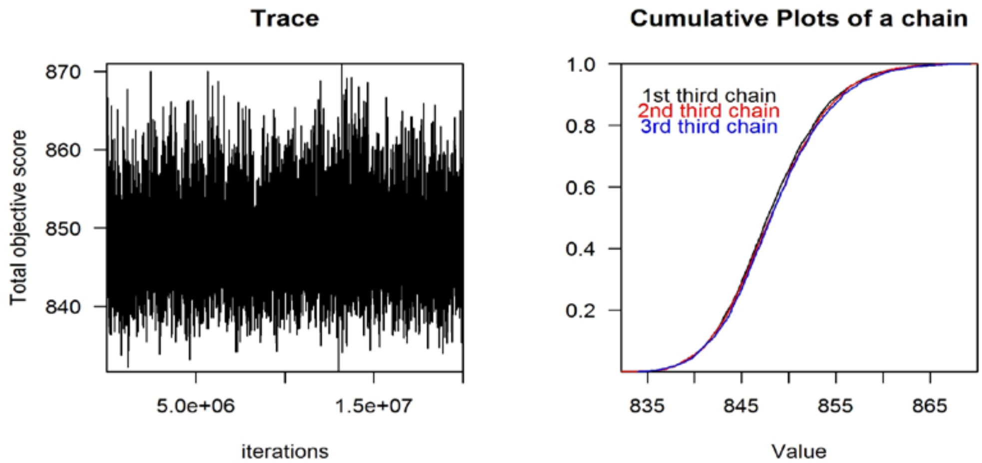

Figure 52: Trace and cumulative plots for the Annual scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

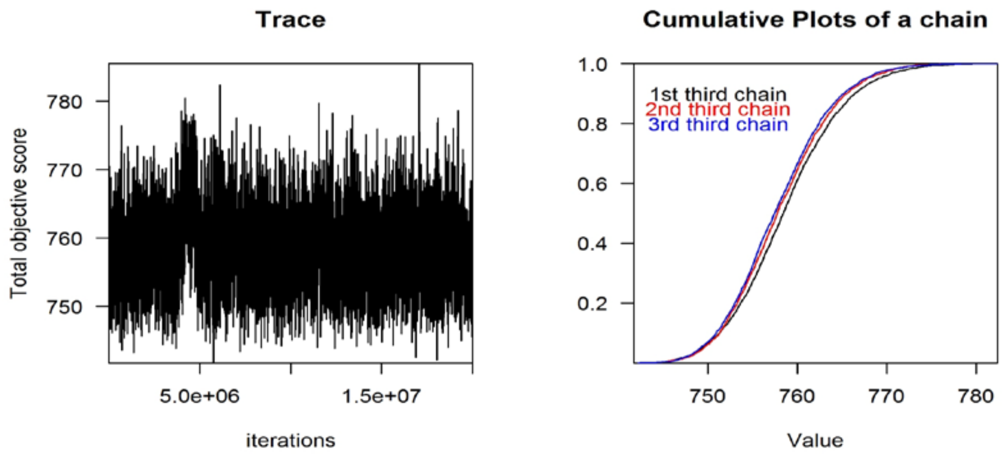

Figure 53 Trace and cumulative plots for the Biennial 1 scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

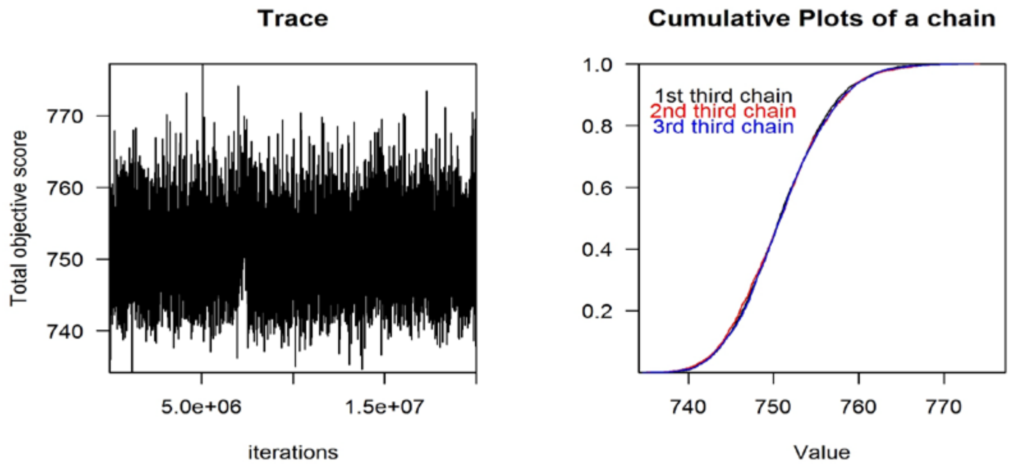

Figure 54: Trace and cumulative plots for the Biennial 2 scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

9.3 HAK 4

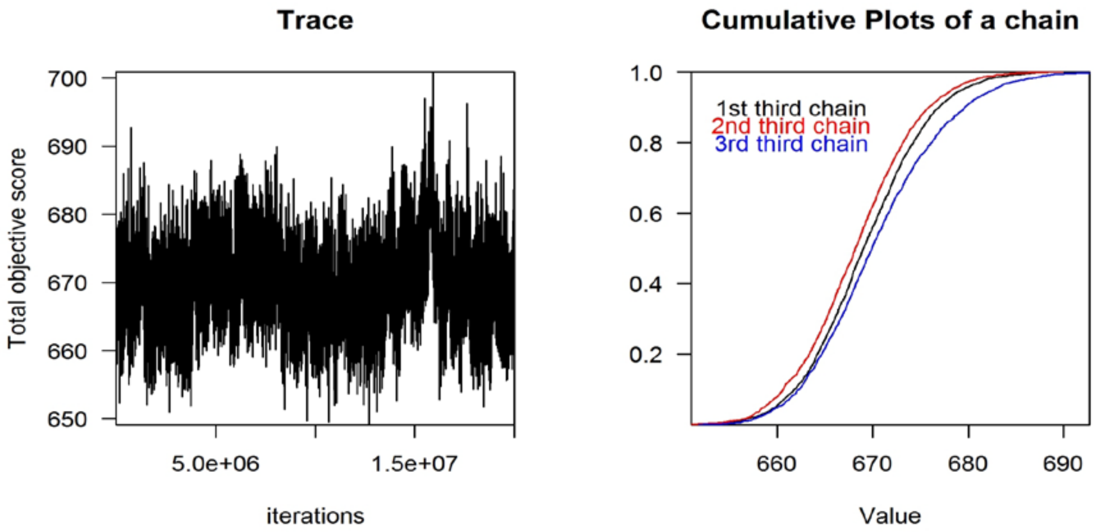

Figure 55: Trace and cumulative plots for the Annual scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

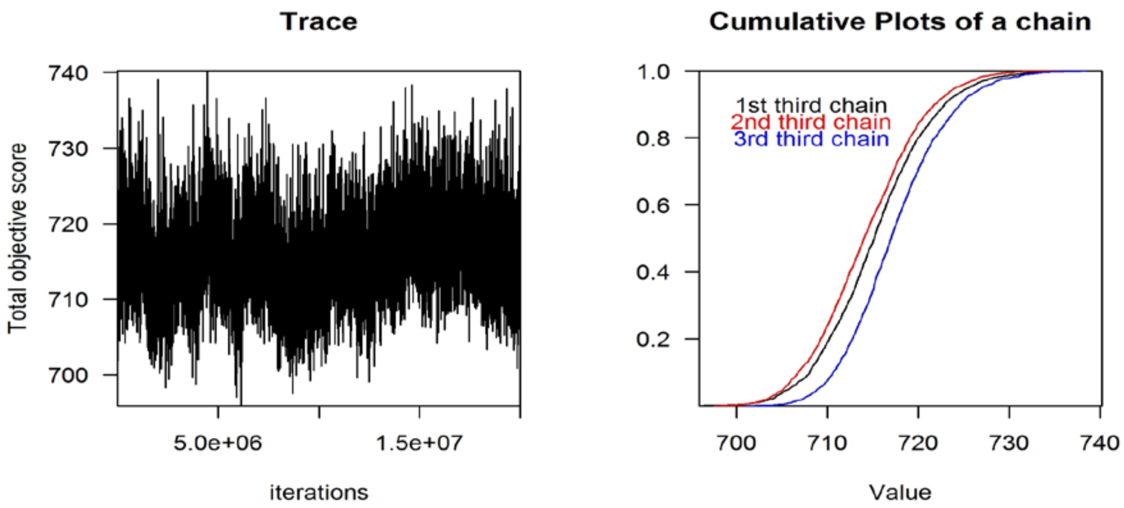

Figure 56: Trace and cumulative plots for the Biennial 1 scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

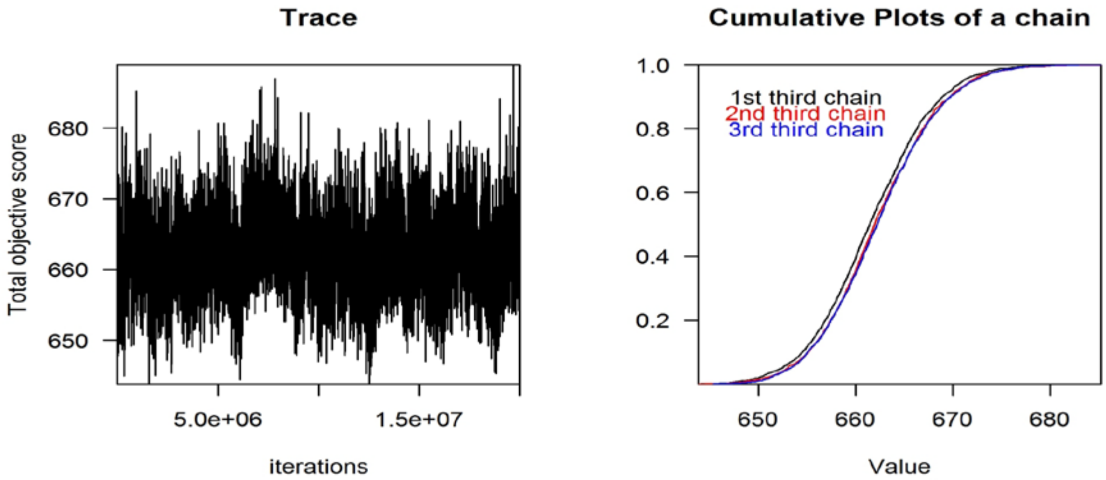

Figure 57: Trace and cumulative plots for the Biennial 2 scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

9.4 LIN 3 & 4

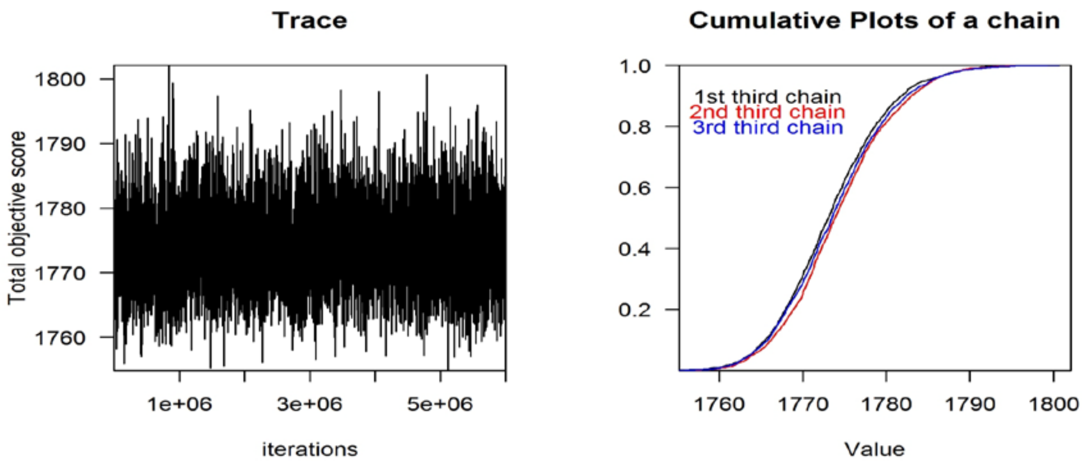

Figure 58: Trace and cumulative plots for the Annual scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

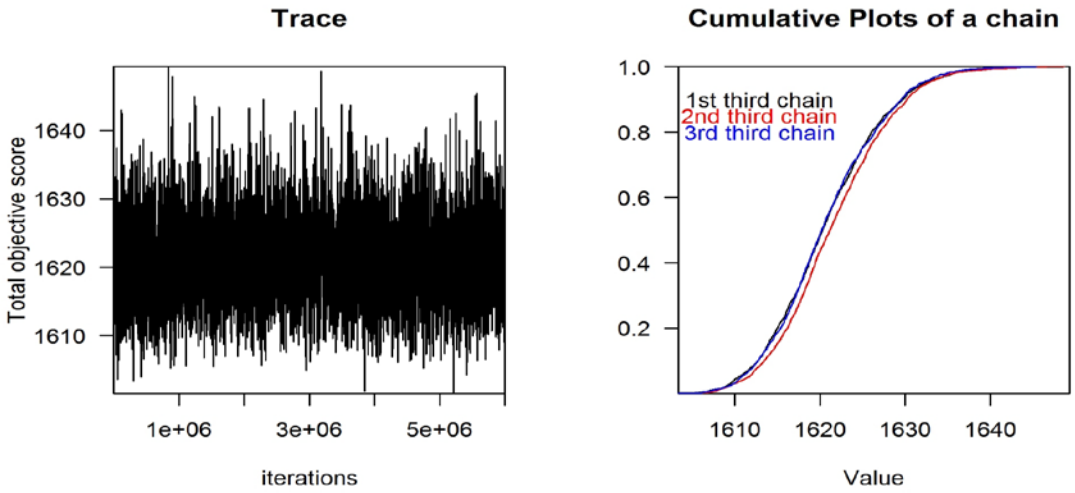

Figure 59: Trace and cumulative plots for the Biennial 1 scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

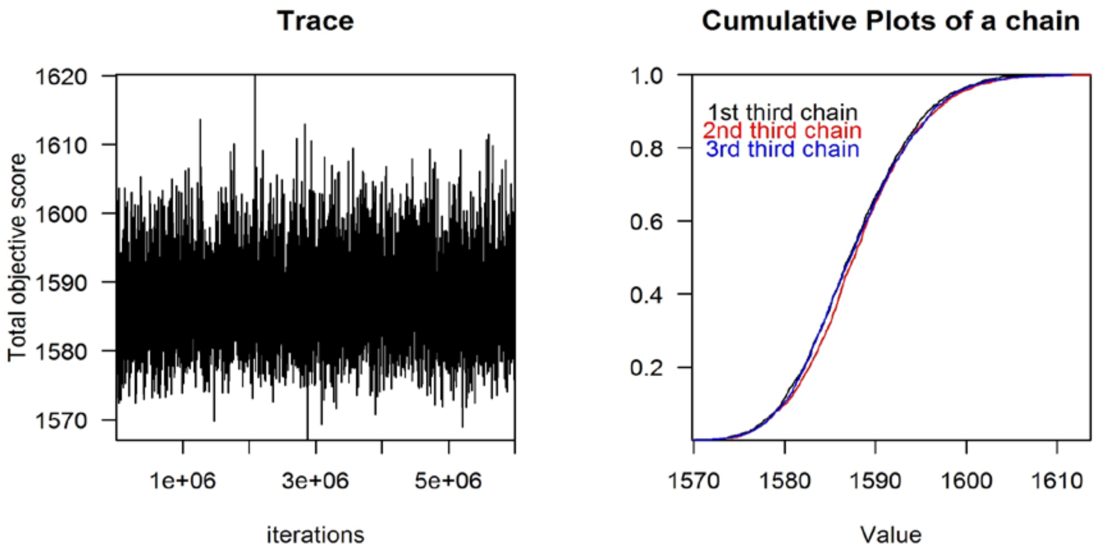

Figure 60: Trace and cumulative plots for the Biennial 2 scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

9.5 LIN 5 & 6

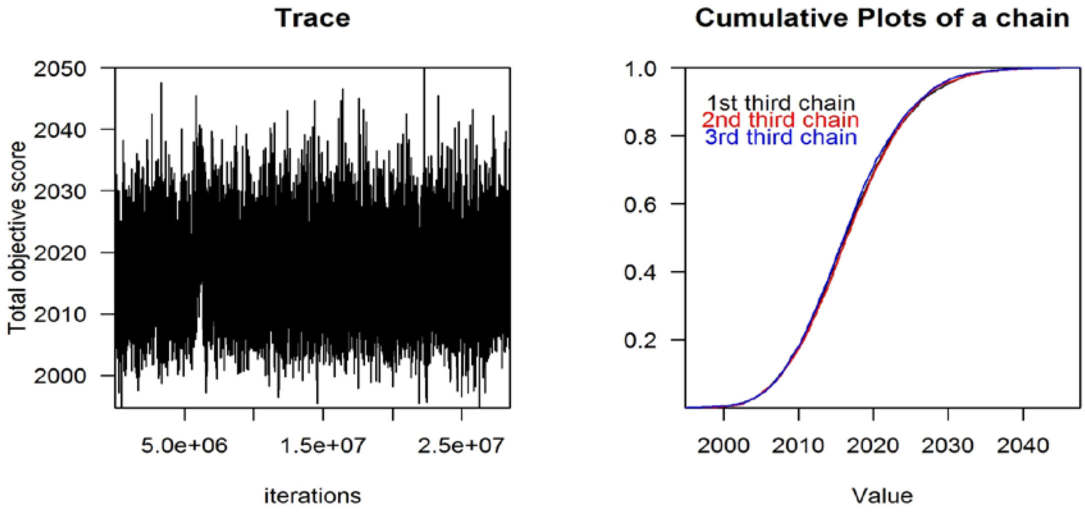

Figure 61: Trace and cumulative plots for the Annual scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

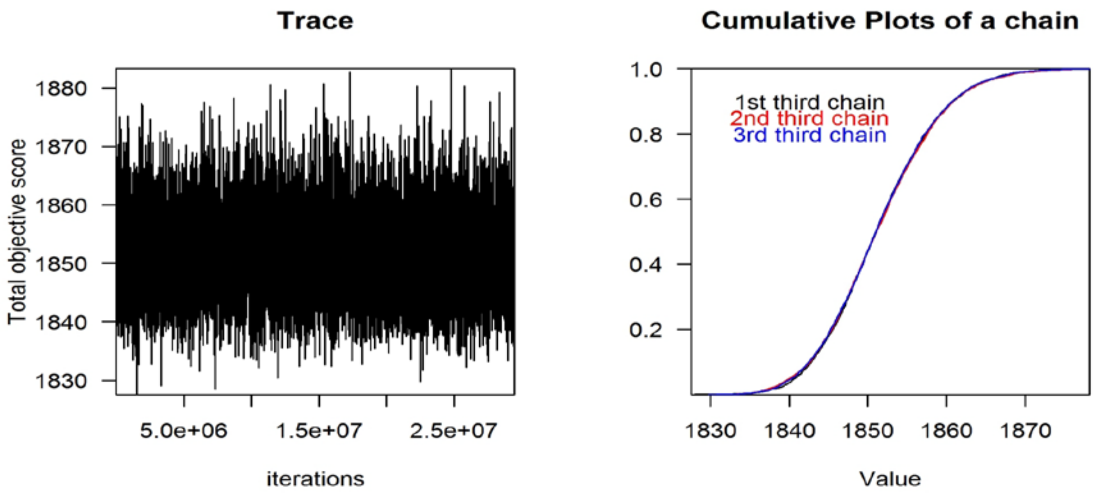

Figure 62: Trace and cumulative plots for the Biennial 1 scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.

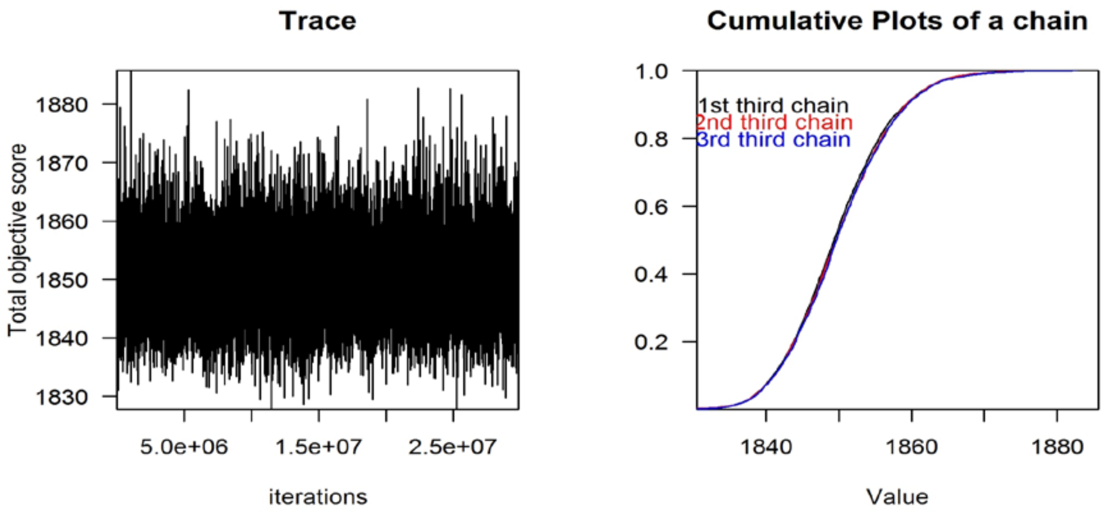

Figure 63: Trace and cumulative plots for the Biennial 2 scenario of the total objective score. The cumulative plot (right) have chopped the chain into equal thirds and plotted up over each other, to check for stationarity.