Operating models for plaice (*Pleuronectes platessa*) in Division 7.e (western English Channel – ple.27.7e) – Working document for WKBPLAICE 2024[I]

Simon H. Fischer[1,2]*

[1] Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, Suffolk, NR33 0HT, UK

[2] Centre for Environmental Policy, Imperial College London, Weeks Building, 16–18 Prince's Gardens, South Kensington, London, SW7 1NE

[*] Contact: e-mail: simon.fischer@cefas.gov.uk

Working document from August 22, 2024

Contents

[I] Version 3. Added clarifications to the captions of Figure 1 and Table 1, Sections 1.8.3, 2.2.1, and 2.5, but no changes were made to the operating models.

1 Operating model

1.1 Operating model conditioning

The conditioning of the age-structured stochastic operating models for plaice followed the approach set out by ICES (2019) and Fischer et al. (2023). This means that the operating models were based on the model fits of the state-space assessment model SAM (Nielsen & Berg, 2014) to data. SAM estimates processes (stock numbers at age, recruitment, fishing mortality), observations (catch numbers, survey indices), as well as uncertainties (parameter uncertainty, process error, observation error) and uncertainty structures of estimated parameters, and this can be used in the conditioning of the operating models.

The uncertainty was implemented into the operating models by generating 1,000 different but self-consistent simulation replicates. Each replicate represented one parameter set derived by the sampling from the variance-covariance matrix of the SAM model fit and represents one possible outcome of the model (ICES, 2019).

The baseline operating model was based on the SAM model fit considered most plausible by WKBPLAICE. The following sections describe the operating model and are based on this baseline operating model. The creation of the alternative operating models is described in Section 2.

The framework for conducting the management strategy evaluation (MSE) was based on the Fisheries Library in R (FLR; Kell et al., 2007) and its `mse` package (https://github.com/flr/mse). Essentially, the work for this benchmark is an update of Fischer et al. (2023) and this document is based on the Supplementary Material of Fischer et al. (2023).

The full code and input data for creating the operating models is available on GitHub at https://github.com/shfischer/WKBPLAICE2024_ple.27.7e_MSE.

1.2 Population dynamics

Population dynamics were simulated with age-structured operating models. Population dynamics mimicked the internal dynamics of the state-space SAM model (Nielsen & Berg, 2014). Stock numbers were calculated as

$$N_{a,y,i} = \begin{cases} R_{y,i} & a = 1 \\ N_{a-1,y-1,i}\ e^{-F_{a-1,y-1,i}-M_{a-1,y-1,i}}\ e^{\varepsilon_{a,y,i}} & 1 < a < A \\ \left(N_{a-1,y-1,i}\ e^{-F_{a-1,y-1,i}-M_{a-1,y-1,i}} + N_{a,y-1,i}\ e^{-F_{a,y-1,i}-M_{a,y-1,i}}\right)\ e^{\varepsilon_{a,y,i}} & a = A \end{cases} \tag{1}$$

where $N_{a,y,i}$ are stock numbers for age a, year y and simulation replicate i, with $a = 1$ being the first age class, A the last age class (plusgroup), R are recruits, F is the fishing mortality, M the natural mortality, and ε the multivariate normal survival process error, $\varepsilon_{a,y,i} \sim N(0, \sigma_i^2)$, where σ is estimated by SAM.

For this plaice stock, the survival process error was set independently for the recruits (age 2) and a common value was used for all other ages (ages 3–10+). For the historical part of the operating model (1980–2023), the survival process error was already included in the simulation replicates derived by sampling from the variance-covariance matrix of the SAM model fit. For the projection period, it was included following Equation 1. The median survival process error for the recruits in the baseline operating model was $\sigma \approx 0.4577$ and $\sigma \approx 0.1359$ for all other ages. In the projected period of the MSE, the survival process error for the recruitment age was superseded by a recruitment process error (see Section 1.3).

Catch numbers were calculated following the Baranov catch equation (Sharov, 2021):

$$C_{a,y,i} = \frac{F_{a,y,i}}{F_{a,y,i} + M_{a,y,i}} \; N_{a,y,i} \; \left(1 - e^{-F_{a,y,i} - M_{a,y,i}}\right) \tag{2}$$

The simulation was based on total catches, combining landings and discards. However, catch numbers at age were also split into landings and discards based on the landings ratio, so that alternative discarding scenarios could be explored (see Section 2).

Figure 1 illustrates the historical population dynamics (catch, recruitment, fishing mortality, and SSB) of the baseline operating model in comparison to the SAM model fit on which it is based. Figure 2 shows the first five of the 1000 simulation replicates.

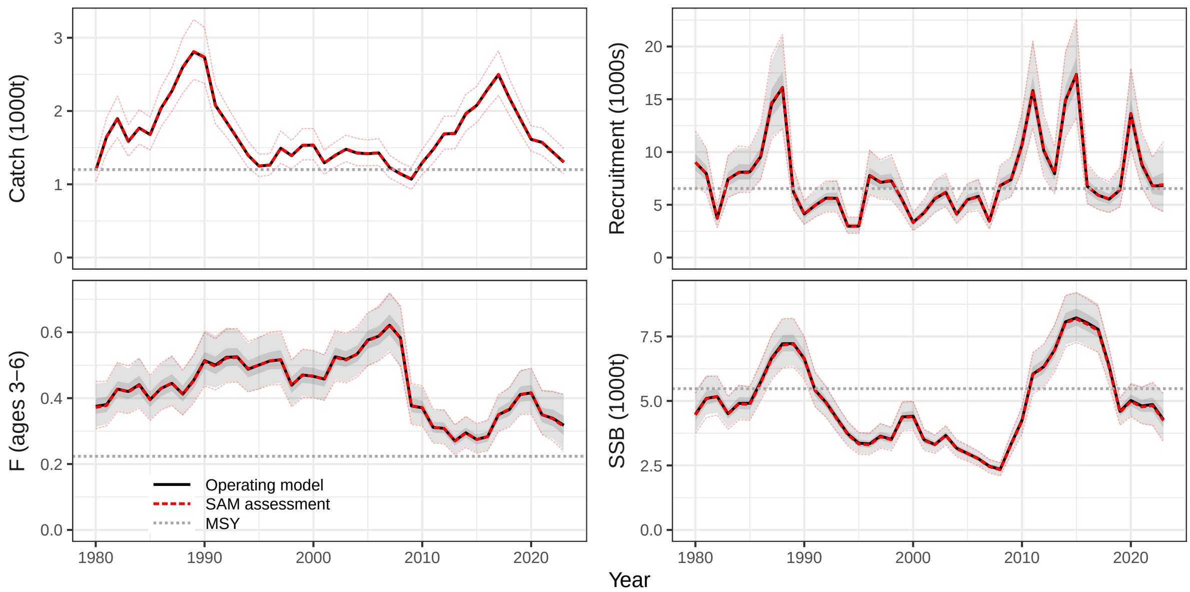

Figure 1: Comparison of the baseline operating model metrics to the output from the SAM model fit. For the operating model, the median (black curve) and 50% confidence intervals (shaded in dark grey) and 95% confidence intervals (shaded in light grey) are shown. For the SAM model fit, the point estimates (red curve) and 95% confidence intervals (dashed red curves) are shown. Please note that the uncertainty of the operating model catch appears much smaller than the uncertainty of the SAM model catch; however, different metrics are shown. The operating model catch is the modelled catch consistent with the stock numbers at age and the fishing mortality, derived by sampling from the variance-covariance matrix of the SAM model. The SAM model fit catch shows the catch and its uncertainty as returned by the SAM model. For details on how catch observations are generated in the MSE projection, see Section 1.8.1 for details).

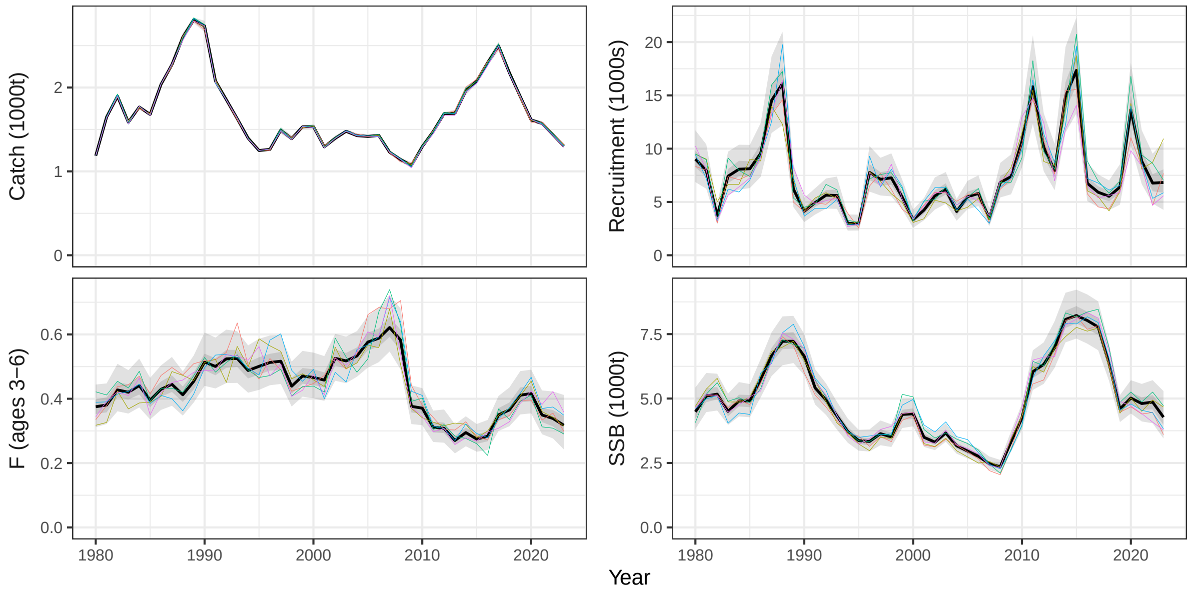

Figure 2: The baseline operating model and the first five simulation replicates (coloured curves).

1.3 Recruitment modelling

Recruitment was modelled by fitting a stock-recruitment model to historical SSB-recruitment pairs. For this plaice stock, a Beverton-Holt model was used:

$$R_{y,i} = \frac{\alpha_i \ \mathrm{SSB}_{y,i}}{\beta_i + \mathrm{SSB}_{y,i}} \ e^{\gamma_{y,i}}, \tag{3}$$

where γ is the recruitment process error (recruitment residuals, see following paragraphs for details).

The Beverton-Holt model was chosen because no clear stock-recruitment relationship is evident for this plaice stock. A hockey-stick model was considered unrealistic because it implies a discrete breakpoint below which recruitment is impaired, but such a value is unknown, and the model was considered biologically implausible.

The recruitment model was fitted individually to each of the 1,000 simulation replicates (Figure 3).

Variability in future recruitment values (process error, implemented with recruitment residuals) was introduced by taking model log-residuals (the difference between the recruitment observations and the modelled recruitment for a given SSB), fitting a kernel density smoother to the residuals (R function `density()`), and sampling from this distribution (Figure 4). This process allowed a wider range of residuals to be generated compared to bootstrapping residuals. Auto-correlation of future residuals was included if lag-1 auto-correlation was statistically significant in historical residuals (as tested with R's `acf()`) as a first-order auto-regressive (AR1) process (ICES, 2019; Kaitala and Ranta, 2001):

$$\gamma'_{y,i} = \rho \ \gamma'_{y-1,i} + \sqrt{1-\rho^2} \ \gamma_{y,i}, \tag{4}$$

where γ are the default recruitment residuals, ρ the auto-correlation, and γ' the recruitment

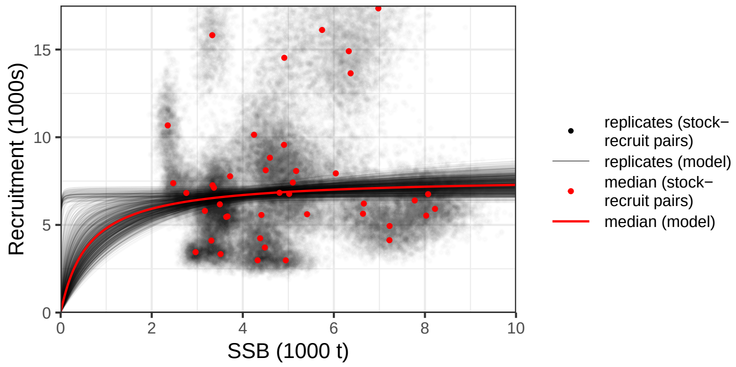

Figure 3: Recruitment model (Beverton-Holt) of the baseline operating model. Red points indicate the medians of the SSB-recruitment pairs and are surrounded by small black points representing the SSB-recruitment pairs for the 1,000 simulation replicates. The thin black curves represent the recruitment models for the 1,000 simulation replicates, and the red curve represents a recruitment model based on the medians of the model parameters.

residuals after including the auto-correlation. For the baseline operating model, residuals exhibited a statistically significant auto-correlation ($\rho \approx 0.5439$) and so the residuals were modified to also include this auto-correlation.

The model fitting and generation of recruitment residuals were done independently for each simulation replicate.

Recruitment is one of the main factors defining the productivity of the stock in the MSE. Consequently, it is important to check that modelled recruitment values are similar to values observed in the past (Figure 5). For the comparison in Figure 5, recruitment was modelled with the approach deployed in the MSE projection for historical years for which recruitment estimates existed (Equations 3 and 4), and showed that the modelled recruitment and its distribution were similar to historical data and therefore appropriate for use in the MSE.

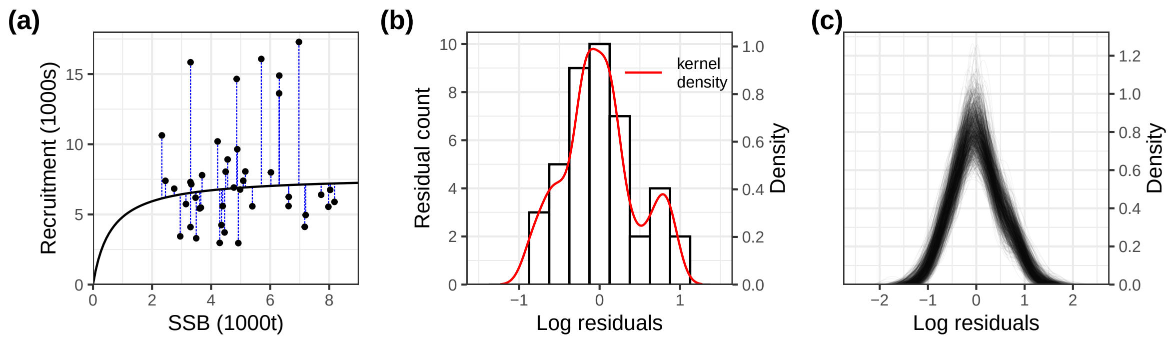

Figure 4: Visualisation of recruitment modelling for the baseline operating model. Panels (a) and (b) illustrate the process for the median stock-recruit pairs. Shown are (a) the Beverton-Holt recruitment model fit (solid black curve; points are stock-recruit pairs) and (b) the distribution of log residuals (histogram). Residuals for the projection are sampled from the kernel density distribution (red curve in b). Panel (c) illustrates the kernel densities for all 1,000 simulation replicates.

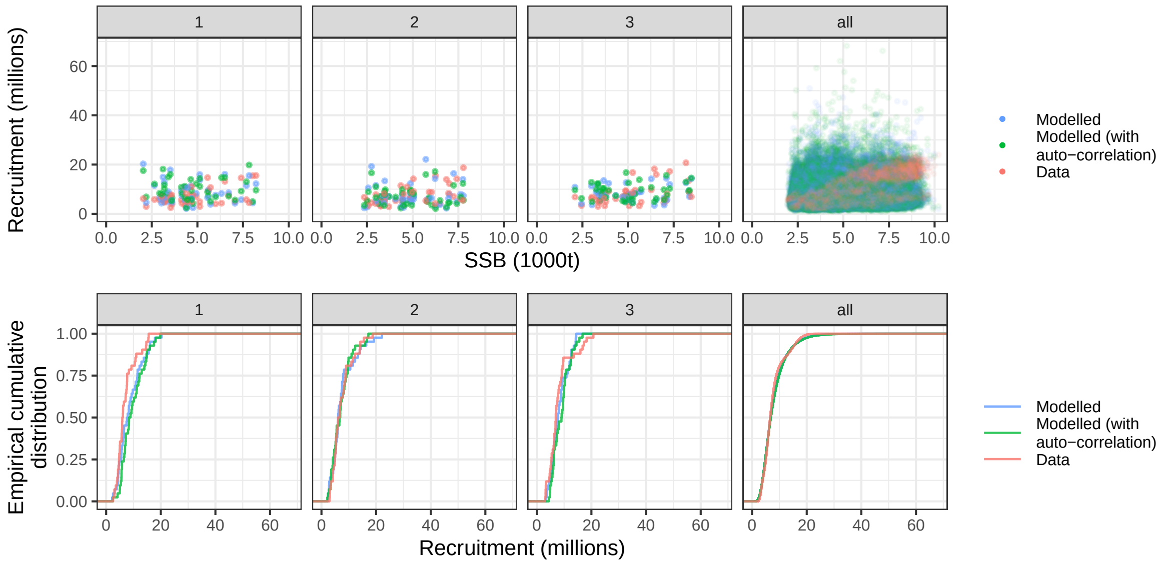

Figure 5: Comparison of recruitment data as estimated by SAM and recruitment values generated by the recruitment modelling approach as used in the MSE simulation for the baseline operating model. The top row shows the stock-recruitment pairs, and the bottom row shows the empirical cumulative distribution of the recruitment values. Columns 1–3 show individual simulation replicates, and in the last column, all 1000 were replicates combined.

1.4 Biological data and fisheries selectivity

Biological data (weights at age for the stock and catch, natural mortality, and maturity) of the operating model for the historical period (for which data existed and are used in the ICES stock assessments) were identical to those used in the stock assessment and the same for each simulation replicate. Time-varying fishery selectivity (assuming a single fleet) is estimated by SAM. Consequently, the selectivity differed by simulation replicate in the historical period.

For the projected period (20 years) in the closed-loop simulation, variability was introduced for biological parameters and fishery selectivity by resampling from the historical period. Values were resampled from the last five historical years (Figure 6). The discard weights for age 10+ were very low and lower than those for younger ages. This was caused by very low discards for age 10+ and consequently very low sampling but it did not affect the MSE because discards were very low.

The resampling process was implemented by randomly selecting a year from the pool of years (the last five historical years, i.e., 2019–2023, indicated by the area highlighted in red in Figure 6) and taking all biological parameters for all ages in this year so that possible correlations between ages or different biological parameters were maintained. This process was repeated for fishery selectivity, independently for each simulation replicate and separate from the biological resampling (Figure 7).

1.5 Intermediate year

The input data for the SAM model on which the operating models were conditioned used data until 2023. The implementation of the management procedures in the MSE loop started in 2025,

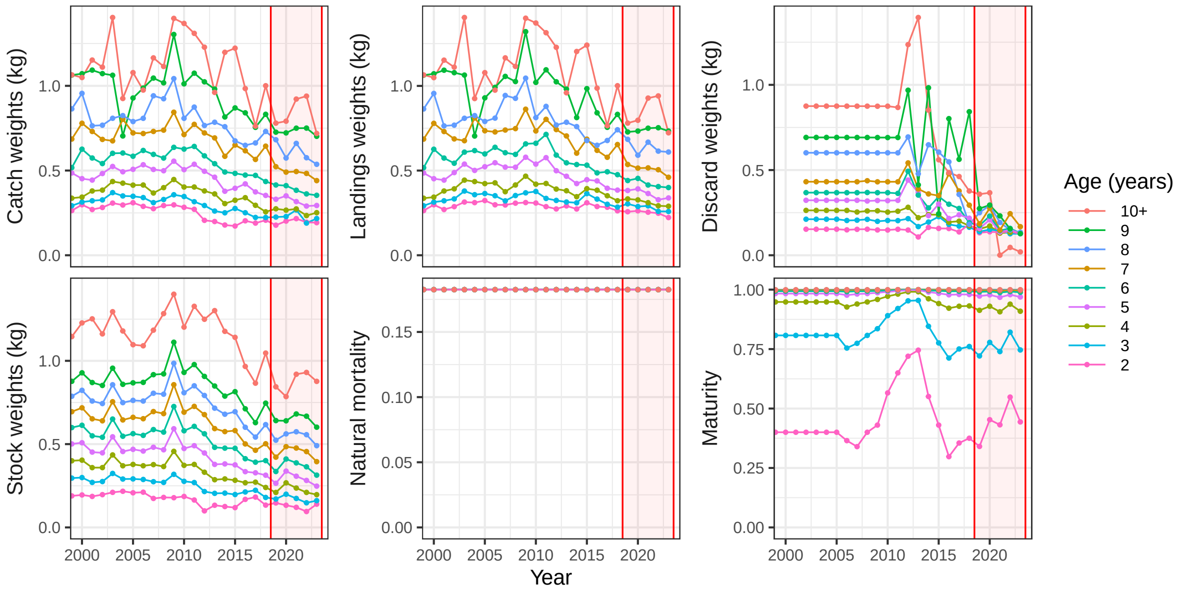

Figure 6: Historical biological data (catch and stock weights at age, natural mortality, maturity). Early years in the time series before 1999 are cut off in the plot. The area highlighted in red (2019–2023) indicates the time period from which values were randomly resampled for the MSE projection.

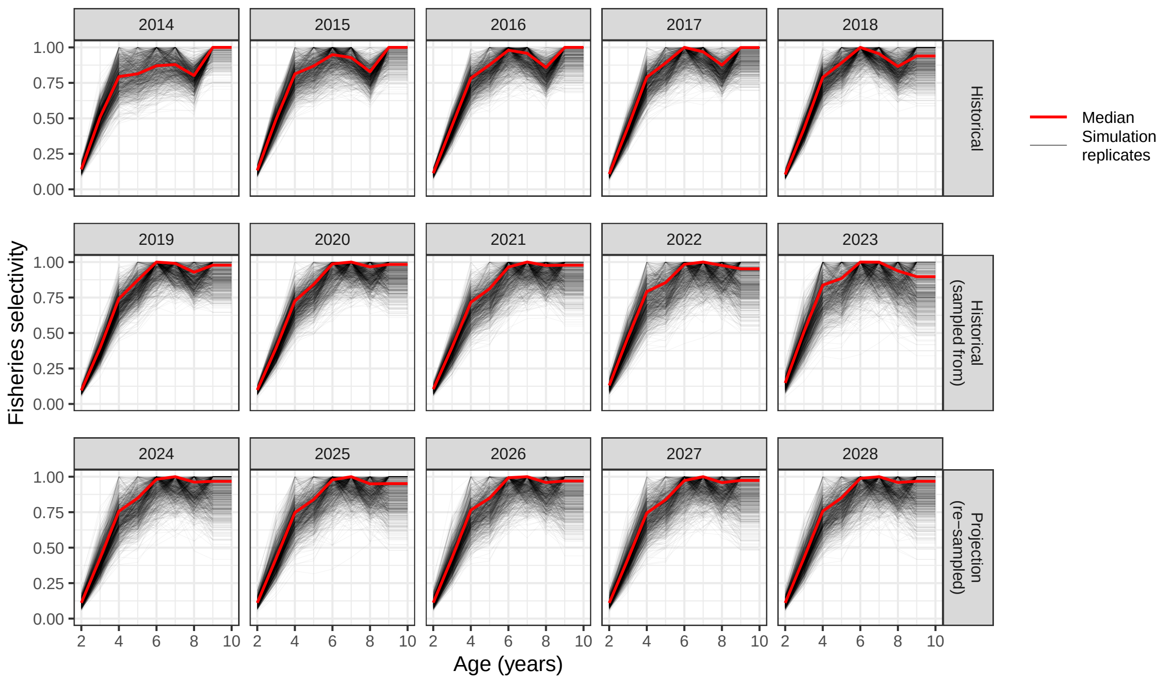

Figure 7: Fisheries selectivity for some historical years (top two rows) and the first five projected years (bottom row). The selectivity in the projected years (2024 and later) was randomly resampled from the last five historical years (2019-2023, middle row).

the year after the benchmark. This means that 2024 was an intermediate year for which no data were available. This year was modelled by extending the operating model to 2024 by specifying a target catch value and following the stock dynamics as described in the previous sections, including the survival process error (Equations 1–3), and the fisheries selectivity described in the previous section (Figure 7). This approach of extending the operating model for the intermediate year was the same as deployed by ICES (2019). The catch target was set to the ICES catch advice for 2024 for plaice in Division 7.e of 1219 tonnes. This catch target was then reduced for the discard survival, if relevant for the respective operating models. The catch was split into landings and discards based on the average discarding pattern of the last 5 years (2019–2023, 21.6%), i.e. the total catch of 1219 tonnes was assumed to consist of 956.0 tonnes of landings and 263.0 tonnes of discards. For example, the baseline operating model assumed a discard survival of 50%, which means that the total 2024 catch was 1087.5 tonnes (956.0 tonnes of landings and 131.5 tonnes of discards).

1.6 Uncertainty characterisation and MCMC

The operating models were conditioned on the model fit of a SAM model and simulation replicates were generated by sampling from the variance-covariance matrix of the model fit. An alternative approach is to use a Markov chain Monte Carlo (MCMC) approach to sample from the model fit (i.e. the probability distribution of estimated parameters) to generate the simulation replicates.

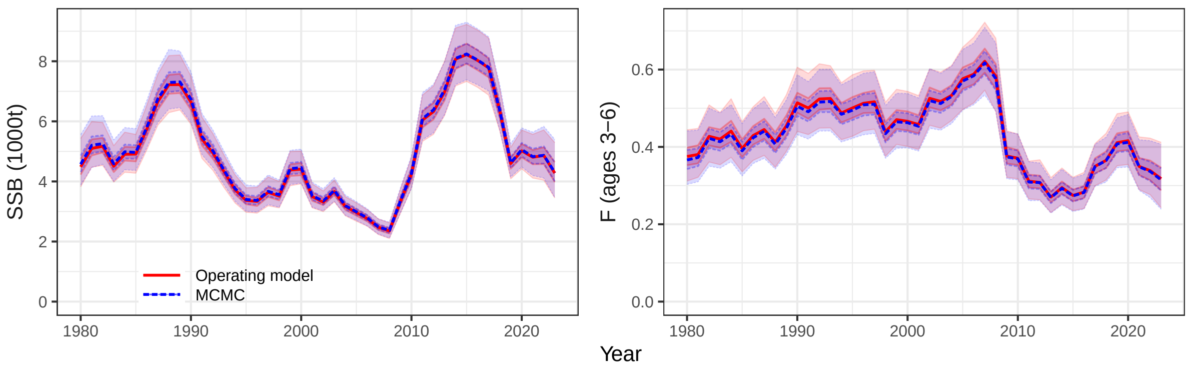

Figure 8: Comparison of SSB and F of the operating model (generated by sampling from the variance-covariance matrix of a SAM model fit) and an alternative approach using Markov chain Monte Carlo (MCMC). Shown are the median, surrounded by 50% and 95% confidence intervals (shaded areas).

An MCMC approach was used to compare it to the default approach. This was implemented with the R package `tmbstan` and run with 10 chains, with 10,000 warmup iterations and then generating 1,000 sampling iterations. Figure 8 shows a comparison of SSB and F from the MCMC approach and the variance-covariance approach. The results are very similar, indicating that the approach adopted for generating the operating model for plaice is appropriate for use in the MSE.

1.7 Reference points

1.7.1 MSY reference points

Operating model maximum sustainable yield (MSY) reference points were estimated using the simulation framework (including the survival process error and recruitment variability) and projecting forward for 100 years with constant Fs.

Table 1: Operating model reference points for the baseline operating model. Shown are fishing mortality (arithmetic mean over ages 3-6, F), catch, spawning stock biomass (SSB), total stock biomass (TSB), and Recruitment (R). F is the value targeted in a 100-year projection to determine MSY; Catch, SSB, TSB and R are the medians over the last 10 years and 1,000 simulation replicates.

	F	**Catch (t)**	**SSB (t)**	**TSB (t)**	R **(1,000)**
MSY	0.233	1,201	5,476	6,349	6,533
Unfished	0	0	21,020	22,031	7,192

MSY was derived by maximising the long-term catch when fishing at constant F. The long-term catch was defined as the last 10 years of a 100-year projection and calculated as the median over this time period and the 1,000 simulation replicates. This means a single value was used for all simulation replicates. The F that led to the highest catch ($F_{\text{MSY}} = 0.233$, MSY = 1201t) was found with a golden-section search (R function `optim()`) and is illustrated in Figure 9. Table 1 summarises the MSY reference points.

Figure 10 illustrates the 100-year projection for $F = F_{\text{MSY}}$ and Figure 11 for $F = 0$. These figures are also a good check to see if the MSE projections make sense and only include process error (survival process error, recruitment variability) but no observation error. For the $F = F_{\text{MSY}}$ projection (Figure 10), B_{MSY} quickly reached its terminal value after only a few years and then stayed relatively stable. For the $F = 0$ projection (Figure 11), the stock dynamics stabilised after around 20 years, which is also the suggested duration for the implementation of the management procedures.

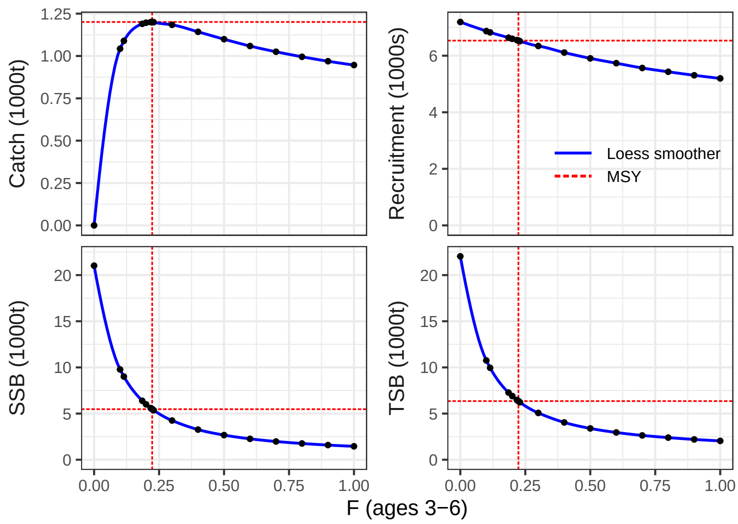

Figure 9: Visualisation of the MSY estimation for the baseline operating model. The points are the long-term averages (median of the last 10 years and 1,000 simulation replicates of a 100-year stochastic projection), and the blue curves are a loess smoother fitted to these values. MSY estimates are highlighted with red dashed lines.

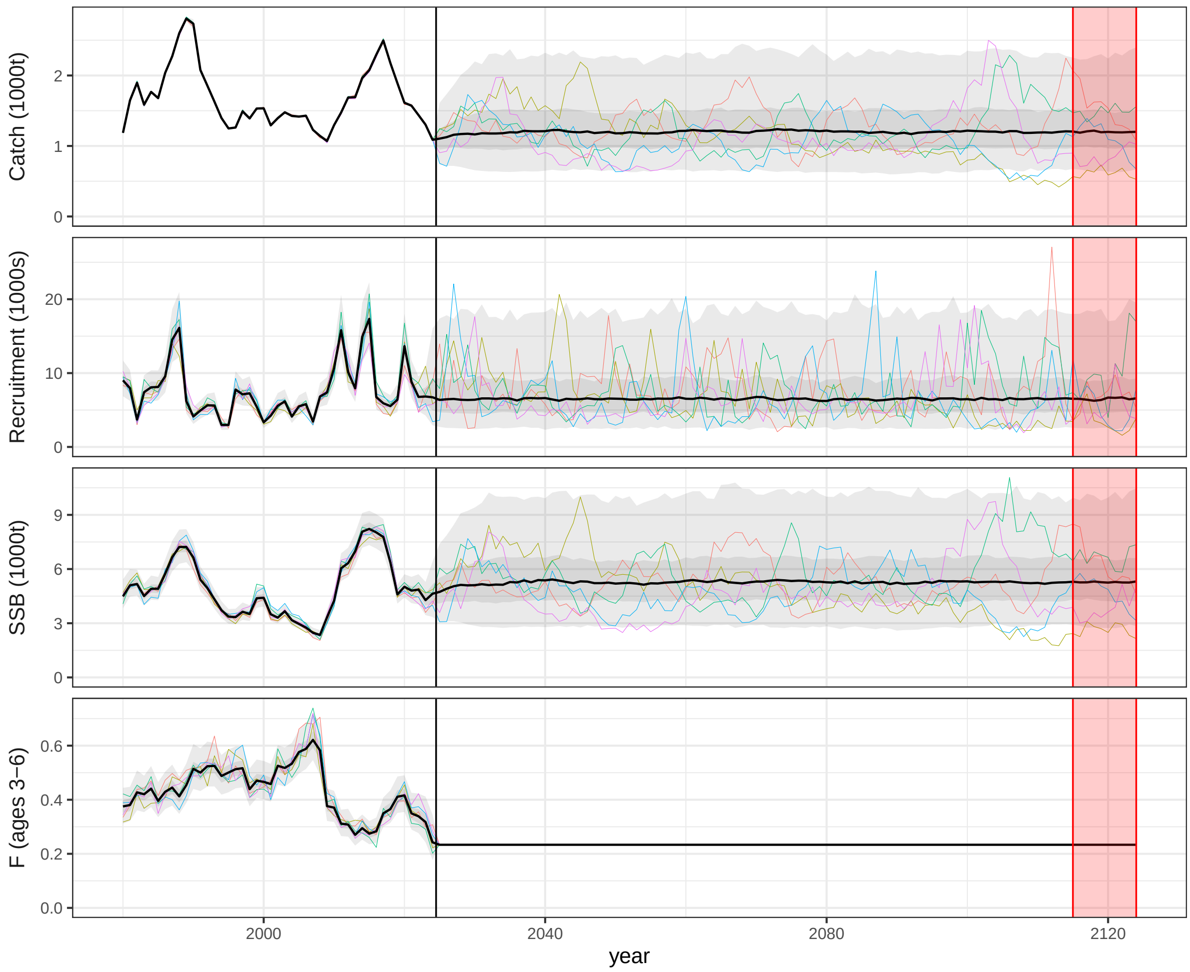

Figure 10: 100-year projections with constant $F = F_{\mathrm{MSY}}$ of the baseline operating model. Shown are the median (black curves), surrounded by 50 and 95% confidence intervals. The black vertical lines indicate the beginning of the projection. The red areas highlight the last 10 years, which were used for MSY estimations. Coloured curves represent five simulation replicates.

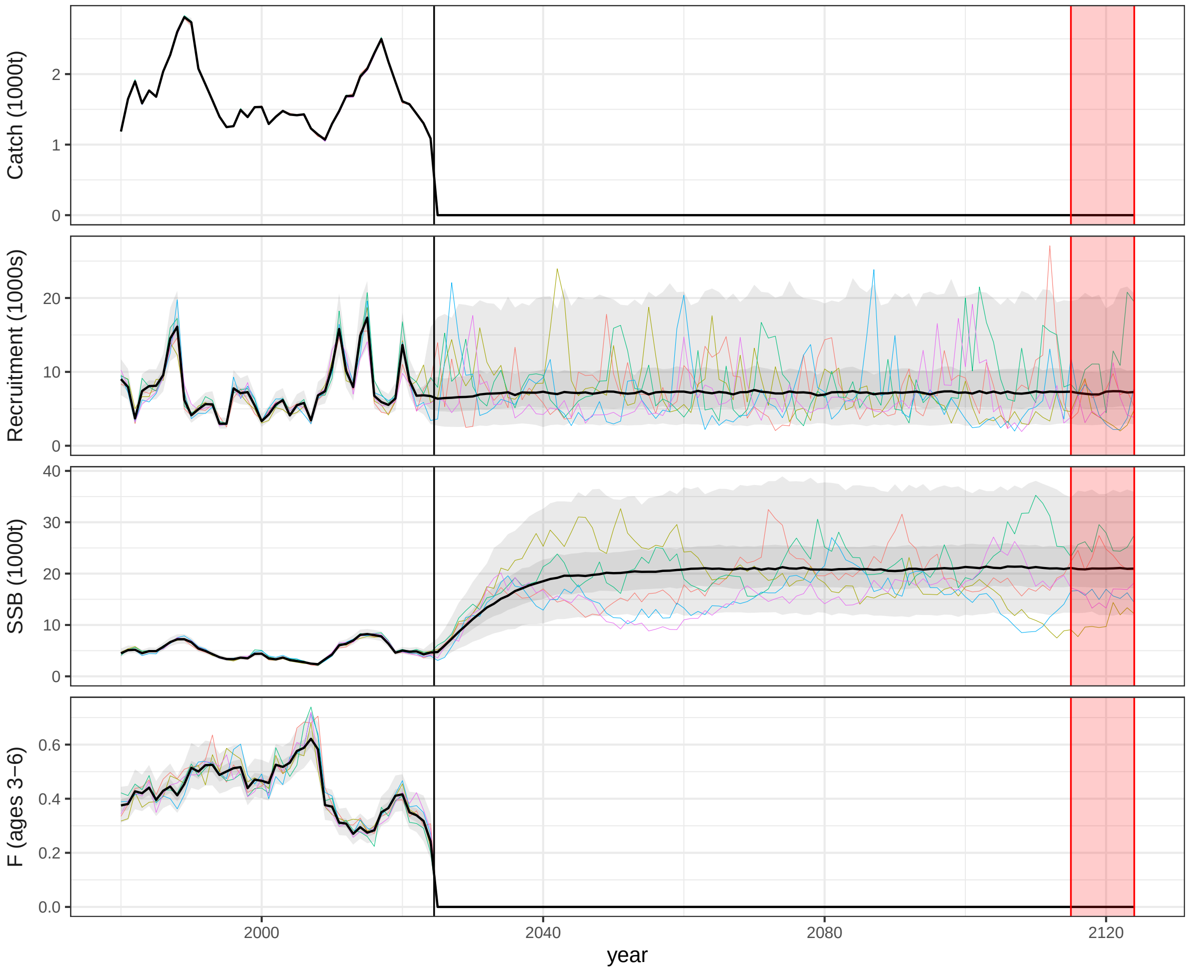

Figure 11: 100-year projections with constant $F = 0$ of the baseline operating model. Shown are the median (black curves), surrounded by 50 and 95% confidence intervals. The black vertical lines indicate the beginning of the projection. The red areas highlight the last 10 years, which were used for MSY estimations. Coloured curves represent five simulation replicates.

1.7.2 $\mathbf{B_{lim}}$

The biomass limit reference point (B_{lim}) requires careful consideration because it is used in the calculation of risk in the ICES precautionary approach. B_{lim} is meant to represent the SSB below which recruitment is impaired (ICES, 2021a). In this MSE, recruitment was modelled with a Beverton-Holt stock-recruitment model, which does not have any obvious breakpoint on which B_{lim} could be based.

Following the current ICES guidelines on calculating reference points for data-rich stocks (ICES, 2021a), this plaice stock can be classified as stock type 5, i.e. a stock showing no evidence of impaired recruitment or with no clear relation between stock and recruitment (see Figure 3). In this case, ICES (2021a) suggest using the lowest observed SSB value B_{loss} as the basis for B_{lim} and this approach was adopted here. The lowest SSB occurred in 2008 and led to $B_{\text{lim}} = 2352$t (Figure 12). This SSB value corresponds to $0.112B_0$ (proportion of unfished SSB) and $0.430B_{\text{MSY}}$. The equilibrium recruitment value at B_{lim} (from the stock-recruitment model) corresponded to $0.794R_0$ (proportion of unfished recruitment), i.e. indicated a recruitment impairment of 21%. Furthermore, the observed recruitment value corresponding to B_{lim} (Figure 3) was fairly high and above the median (86th percentile) of the recruitment distribution.

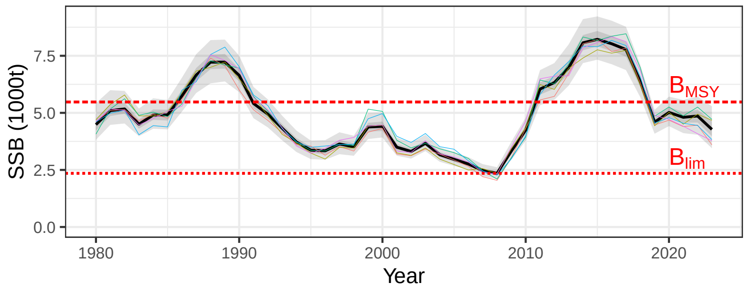

Figure 12: Basis for the biomass limit reference point B_{lim}, defined as the lowest SSB observed in 2008.

There are currently considerations to revise the definition of reference points within ICES as part of the WKREF (ICES, 2022c,d) and WKNEWREF (meeting in 2024, report not finalised in time for WKBPLAICE) workshops. WKNEWREF did not come up with a definite definition of reference points and noted that deviations from the guidelines are possible if justified. However, the workshop proposed that for a type 5 stock, B_{lim} may be based on a proportion of B_0 (yet to be defined, but should not be below 0.1 B_0) or could be estimated empirically as the lowest SSB resulting in good (above median) recruitment. Although this is only a preliminary proposal of WKNEWREF and may not be implemented by ICES (another workshop is planned), B_{lim} for this plaice stock already meets the new conditions suggested by WKNEWREF and also follows the previous guidelines.

1.8 Observations

Observations were generated from the operating model and passed to the management procedure. The data generated by the operating model for the historical period were identical to the data observed in reality. For the projection years, the error structure from SAM was used to model the observations, and these observations were added to the historical observations. All operating model uncertainty estimates (residuals) were generated before the simulations so that they were identical in all simulations of the same operating model to facilitate comparisons between different management procedures.

1.8.1 Catch observations

Catch observations were based on the operating model catch numbers of Equation 2 and included an observation error term:

$$C^{\text{obs}}_{a,y,i} = C_{a,y,i}\ e^{\varepsilon_{a,y,i}}, \tag{5}$$

with $\varepsilon_{a,y,i} \sim N(0, \sigma^2_{a,i})$ and σ being the age-specific observation standard deviation estimated by SAM. The SAM model configuration for this plaice stock specified that all ages were coupled and used the same standard deviation (median $\sigma \approx 0.1737$ for the baseline operating model).

Following ICES (2019), the catch residuals for historical years ($e^{\varepsilon_{a,y,i}}$ in Equation 5 for $y \leq 2023$) were adapted so that these catch observations exactly matched the values observed in reality, as estimated by the ICES assessment working group. This ensured that in the first year of the MSE projection, the management procedures received the same data as would be available in reality.

The total catch C^T was derived by multiplying catch numbers with the individual weight (W^C) and aggregating over all operating model age classes:

$$C^T_{y,i} = \sum_{a=1}^{A} C^{\text{obs}}_{a,y,i}\ W^C_{a,y,i}. \tag{6}$$

1.8.2 Indices

Two surveys (Q1SWBeam and UK-FSP) were included in the SAM assessment and modelled in the MSE.

Survey index observations J were generated from the operating model with

$$J_{a,y,s,i} = q_{a,s,i}\ N_{a,y,i}\ e^{-t_s(F_{a,y,i}+M_{a,y,i})}\ e^{\varepsilon_{a,y,s,i}}, \tag{7}$$

where subscript s indicates the survey (Q1SWBeam and UK-FSP), q the catchability (time-independent, see Figure 13), t the timing of the survey during the year ($t = 0.125$ for the quarter 1 survey Q1SWBeam and $t = 0.775$ for the quarter 3 UK-FSP), and $\varepsilon_{a,y,s,i} \sim N(0, \sigma^2_{a,s,i})$ the observation error with standard deviation σ estimated by SAM. As with catch observations, age coupling is common, and the SAM model for this plaice stock specified that all ages were coupled and used the same standard deviation (median $\sigma \approx 0.5305$ for Q1SWBeam and $\sigma \approx 0.5315$ for UK-FSP for the baseline operating model). The SAM model did not specify a covariance structure of the survey observation error between ages within the same years, so a covariance structure was consequently also not included in the MSE.

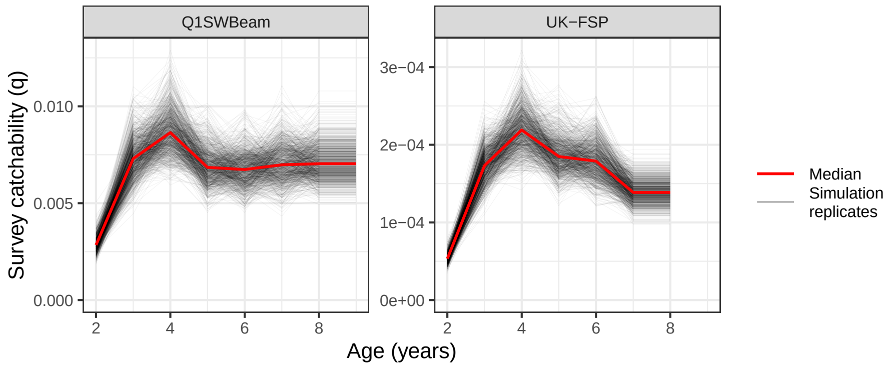

Figure 13: Survey catchabilities (q) for the two surveys (constant over time, as estimated by SAM).

Biomass indices I were generated from the surveys with

$$I_{y,s,i} = \sum_{a=a_{\min,s}}^{a_{\max,s}} J_{a,y,s,i} \; W^{\mathrm{I}}_{a,y,s,i}, \tag{8}$$

where $a_{\min}$ and $a_{\max}$ are the minimum and maximum ages for survey s ($a_{\min} = 2$ for both surveys, $a_{\max} = 9$ for Q1SWBeam and $a_{\max} = 8$ for UK-FSP), J the observed index values from Equation (7), and W^I the weights at age in the index. Weights at age for the projection period were derived from the last five years of data (2019–2023, Figure 14).

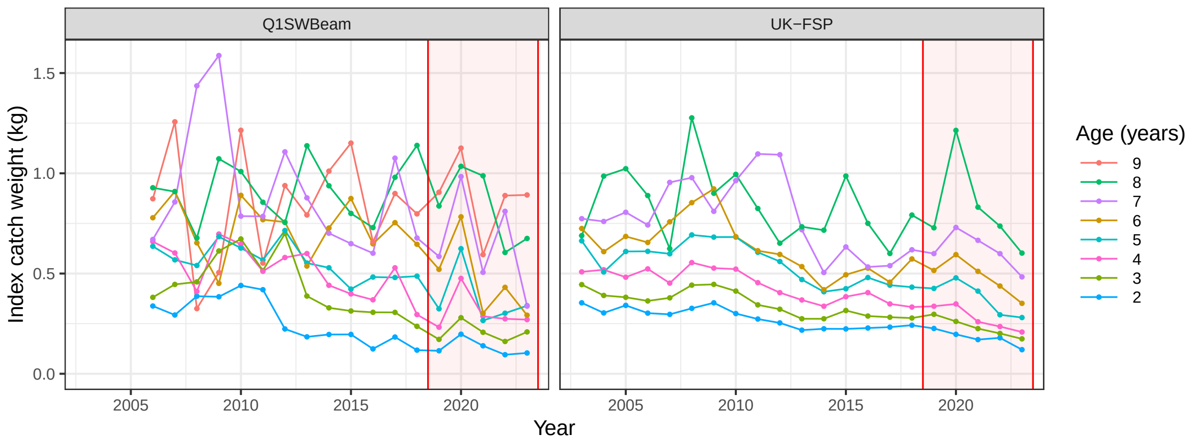

Figure 14: Survey catch weights at age for the two survey indices. The area highlighted in red (2019–2023) indicates the time period from which values were randomly resampled for the MSE projection.

Figure 15 illustrates the survey index generation for the historical years by the observation error model of the MSE in comparison to the values observed in reality (as estimated by the ICES assessment working group). This comparison shows that the survey index modelling appears appropriate for use in the MSE and the uncertainty covers the range observed in reality.

Following ICES (2019), the survey index residuals for historical years ($e^{\varepsilon_{a,y,s,i}}$ in Equation 7 for $y \leq 2023$) were adapted so that these survey index observations exactly matched the values observed in reality, as estimated by the ICES assessment working group. This ensured that in the first year of the MSE projection, the management procedures received the same data as would be available in reality. For projected years, survey observations were generated following Equations 7 and 8.

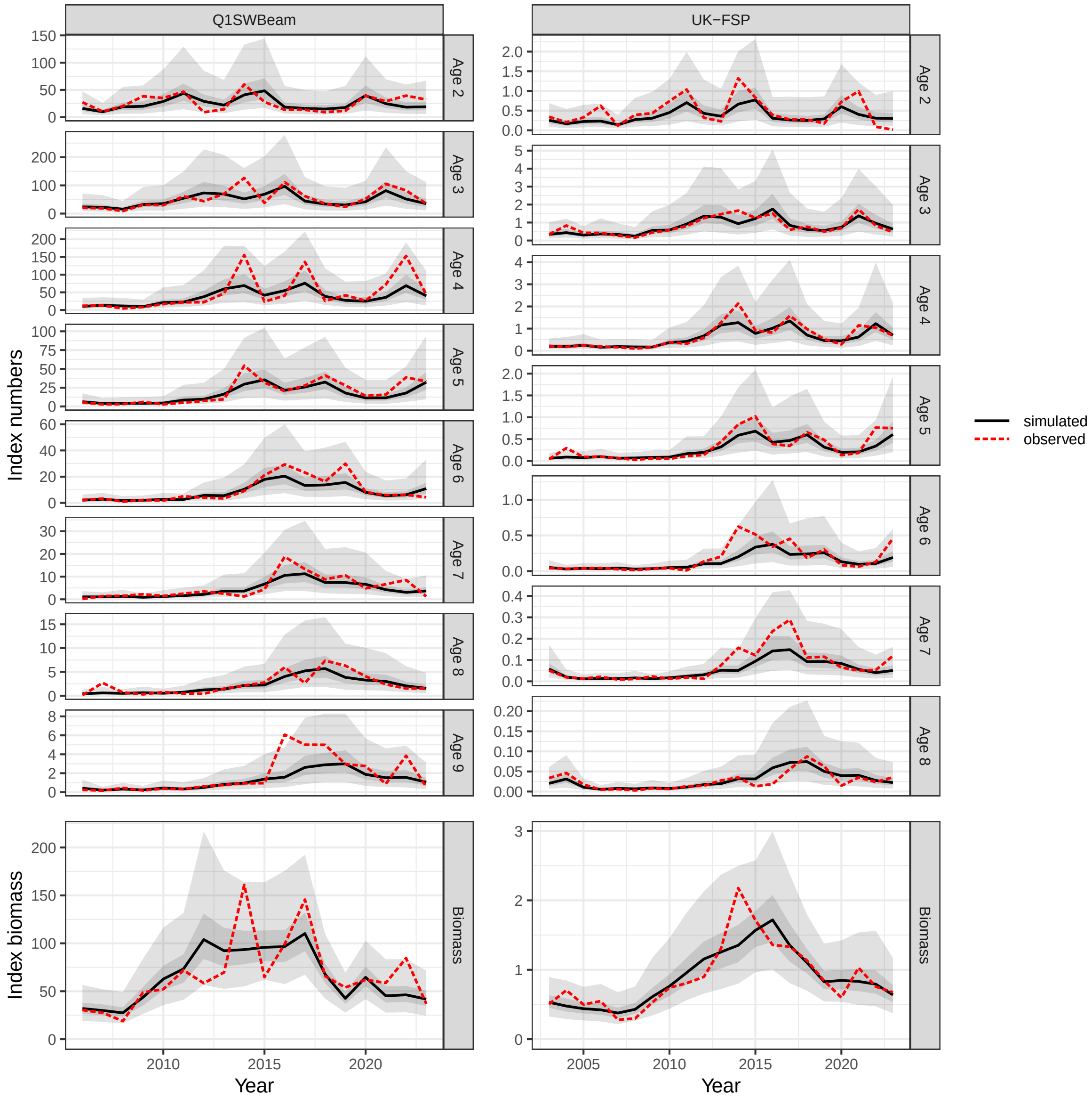

Figure 15: Illustration of the survey index generation. The red curves show the index values observed in reality and passed on to the management procedure. The black curves show the index values (including observation error) generated from the operating model (surrounded by 50% and 95% confidence intervals shaded in grey) as generated by the observation error model in the MSE.

1.8.3 Length data

The operating models were age-structured and did not use lengths internally. However, some data-limited management procedures require length data from the catch (in this MSE, only the rfb rule requires length data). Therefore, the age frequencies of the catch were converted into length frequencies with stock-specific age-length keys (ALKs). The ALKs describe the distribution of lengths for each age class of the operating model. These ALKs were applied to the catch at age data and the catch numbers aggregated by length class to generate the catch at length distribution:

$$C^{\text{obs}}_{L,y,i} = \sum_a C^{\text{obs}}_{a,y,i}\ \kappa_{a,L,y,i}, \tag{9}$$

where C^{obs}_L are the catch numbers at length L, C^{obs}_a the observed numbers at age from Equation (5), and $\kappa_{a,L}$ is the proportion of length L in age a.

For historical years for which yearly ALKs were available, these were used. For the remaining historical years, the available ALKs were combined into a pooled ALK. For the projected years in the closed-loop simulation, length distributions were derived by randomly choosing from the ALKs of the last five historical years (2019–2023), the same years from which other biological parameters were sampled (Section 1.4). This was done separately for each simulation year and replicate.

The final observed length distributions were generated by sampling from the operating model length distribution (described in the previous paragraphs). In each year, 2,000 length samples were drawn, which is a typical sampling level for plaice in previous years (ICES, 2021b).

Subsequently, the mean length of the catch $\bar{L}$ was calculated as the average of the length classes (L) above the length of first capture (L_c), weighted by the number of fish in these length classes (C_L):

$$\overline{L_y} = \frac{\sum_{\ L>L_c} L\ C^{obs}_{L,y,i}}{\sum_{\ L>L_c} C^{obs}_{L,y,i}}. \tag{10}$$

1.8.4 Biological data

Biological data (weights at age, M, maturity, etc.) were passed from the operating model to the management procedure. For the historical period (prior to implementing the management procedures), the biological data passed to the management procedure were identical to those observed in reality. For the projection period, the biological data were the average of the values of those years from which the operating model biological values were sampled (last 5 historical years).

2 Alternative operating models

2.1 Alternative operating model approach

The baseline operating model described in the previous section was conditioned on a specific SAM model fit. For this SAM model, specific decisions about the input data (e.g. about discard mortality) and assumptions (e.g. about natural mortality M) were made. However, there is a degree of uncertainty about these decisions and further specifications of the MSE simulation. To address this, a range of alternative operating models were developed (Table 2) to ensure that the proposed management procedure is robust to such uncertainties. This approach is a best practice in MSE (Punt et al., 2016) and recommended by ICES (ICES, 2020). For the alternative operating models, one characteristic (e.g. catch) was changed at a time and they were not combined (factorial design) to reduce computational complexity.

Table 2: Alternative operating models.

#	Category	ID	Difference	SAM refit?	Operating model type
1	Baseline	*Baseline*	–	–	Baseline
2	Catch	*Catch: no discards*	No discards (100% survival)	Yes	Reference
3		*Catch: 100% discards*	All discards die (0% survival)	Yes	Reference
4		*Catch: +10%*	Catch always 10% above TAC	No	Robustness
5		*Catch: -10%*	Catch always 10% below TAC	No	Robustness
6		*Catch: no migration*	Catch from 7.d excluded	Yes	Reference
7	Natural mortality (M)	*M: -50%*	M 50% below *baseline*	Yes	Reference
8		*M: +50%*	M 50% above *baseline*	Yes	Reference
9		*M: Gislason*	Age-dependent M following Gislason et al. (2010)	Yes	Reference
10	Recruitment	*R: no AC*	No auto-correlation (AC) in recruitment resiudals	No	Reference
11		*R: failure*	Recruitment failure in years 2025–2029	No	Robustness
12		*R: +20%*	Recruitment always 20% higher	No	Robustness
13		*R: -20%*	Recruitment always 20% lower	No	Robustness
14	Uncertainty	*Uncertainty: index*	Higher uncertainty (observation error) for index	No	Robustness

The previously described baseline operating model was considered the most plausible operating model by WKBPLAICE 2024. This operating model is part of a "reference" set of more plausible operating models. Furthermore, there is a separate "robustness" set that includes less plausible operating models that are meant to test the robustness of management procedures with more extreme (less plausible) scenarios.

The main changes for the different operating models happened inside the operating model, but the data passed to the management procedure by the observation (error) model did not pass on these changes. This approach is meant to ensure testing of the robustness of the management procedures if reality is not what it is thought to be and follows the typical approach of MSEs, e.g. by ICES (2019). This meant, for example, when an alternative operating model had an alternative natural mortality, the management procedure assumed natural mortality did not change. The same principle was applied to all observations of biological and fishery data.

For some operating models, the SAM model on which the models were conditioned had to be refitted. This was the case if the input data (e.g. catch) changed. Other operating models only concerned elements during the MSE projection (e.g. higher observation error), in which case the baseline operating model could be re-used.

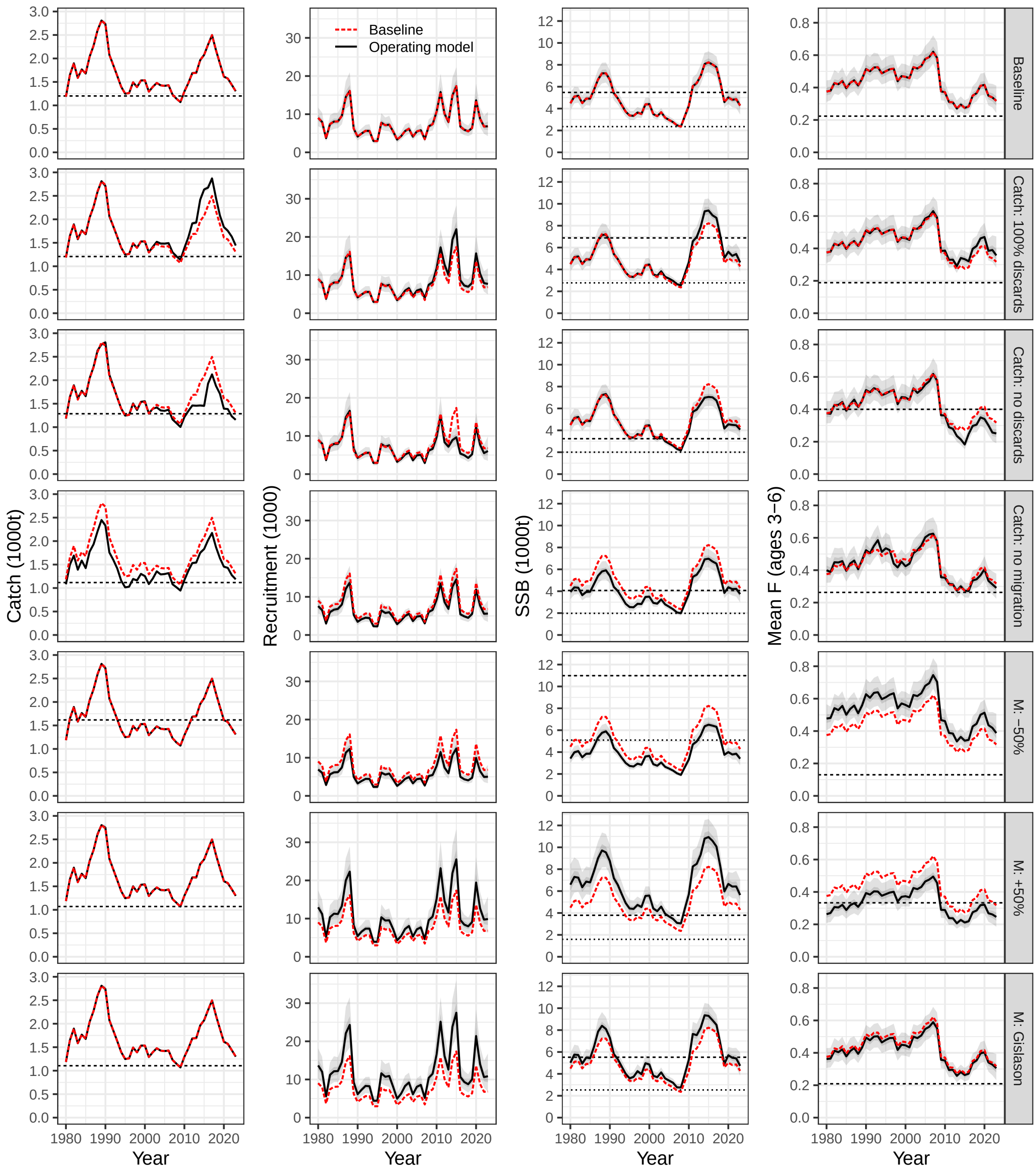

Figure 16: Comparison of the alternative operating models to the baseline operating model. Shaded areas are 50% and 95% confidence intervals. Horizontal dashed lines indicate MSY reference values and horizontal dotted lines B_{lim}. Operating models that were based on the same historical data as the baseline operating model are excluded from the figure (but identical to the top row).

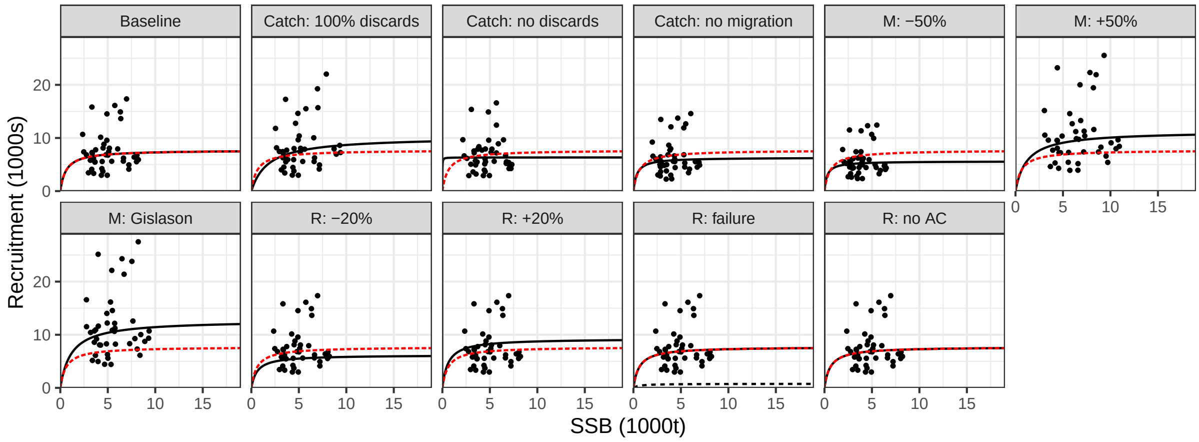

Figure 17: Recruitment models of the alternative operating models. Shown are the median SSB-recruitment pairs (points) and the median recruitment models (curves). For the recruitment failure scenario, the dashed curve indicates the recruitment model during the years of recruitment failure and the solid curves the recruitment model for the remaining years. In all panels, the baseline recruitment model is shown in red for comparison.

A comparison of the alternative operating models is illustrated in Figure 16 and Figure 17 shows the corresponding recruitment models.

2.2 Catch

2.2.1 Discard survival

The baseline operating model assumed that 50% of the discards die and the remaining 50% survive. This assumption was the best estimate of discard survival of WKBPLAICE based on published scientific research (see the working document on the data for this stock). Two alternative operating models with the two extremes for discard survival were developed:

- *Catch: no discards* (reference set) – Assume all discards survived, i.e. only landings contribute to the catch
- *Catch: 100% discards* (reference set) – Assume all discards die (100% discard mortality)

For these two operating models, the input data to the SAM model differed, so SAM was refit. The changes in discards also affected the total catch weights, which were calculated as the average of the landings and discards weights at age, weighted by the landings and discards numbers.

The best estimate of discard survival of WKBPLAICE was 50%, and so the data passed to the management procedure retained the assumption of 50% discard survival (Table 3). This was done because, in reality, an ICES assessment and advice would be based on the best available science (i.e. assume 50% discard survival). The two operating models with alternative assumptions, therefore, explored what happens if the discard survival assumption is wrong. This means there was a discrepancy between the catch of the operating model and the observed catch. The observed catch for *Catch: no discards* was higher than the operating model catch

Table 3: Comparison of operating model catches and catch observations passed to the management procedure for the operating models with different discard survival assumptions.

Operating model	**Discard survival assumption**	
	Operating model	**Management procedure**
Baseline	50% die, 50% survive	50% die, 50% survive
Catch: no discards	0% die, 100% survive	50% die, 50% survive
Catch: 100% discards	100% die, 0% survive	50% die, 50% survive

and lower for *Catch: 100% discards* than the operating model catch (Table 3). These changes were implemented into the MSE as an observation error for the catch.

2.2.2 Implementation error

Two implementation error operating models were developed:

- *Catch:* +10% (robustness set) – Catch always 10% above TAC (overcatch)
- *Catch:* −10% (robustness set) – Catch always 10% below TAC (undercatch)

These changes only affected the operating model, but the catch observations were adjusted with an observation error, so this change was not visible to the management procedure.

2.2.3 Migration

The baseline operating model catches include catches from the adjacent Division 7.d, which are assumed to belong to the 7.e stock. One operating model was developed that did not include these catches:

- *Catch: no migration* (reference set) – Use only catches from Division 7.e; 7.d catches excluded

This change was implemented by treating the 7.e catches (discards and landings) as landings and 7.d catches (discards and landings) as discards. This allowed setting a 100% discard survival, i.e. removing the 7.d catches, and adjusting the catch weights at age accordingly. The observed catches passed to the management procedure included an observation error so that the observed catches still included the catches from 7.d. This operating model required a refit of SAM.

2.3 Natural mortality

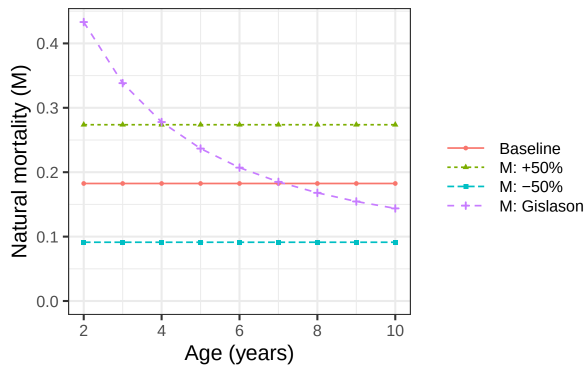

Figure 18: Natural mortality of the baseline operating model and alternative operating models.

Natural mortality M of the baseline operating model was around 0.1825 for all ages. Three alternative operating models were developed with alternative assumptions about M and all three required a refit of SAM (Figure 18). The first two were changes to the age-independent M of the baseline operating model:

- M*:* +50% (reference set) – M 50% higher, i.e. $M = 0.2737$
- M*:* −50% (reference set) – M 50% lower, i.e. $M = 0.0912$

The third operating model used an age-dependent M following Gislason et al. (2010):

- *M: Gislason* (reference set) – age-dependent M

2.4 Recruitment

Four alternative operating models with alternative recruitment modelling choices were created. The first one considered the generation of the recruitment residuals:

- *R: no AC* (reference set) – exclude auto-correlation for recruitment residuals

and meant that ρ in Equation 4 was set to 0. A robustness test was included by simulating a recruitment failure:

- *R: failure* (robustness set) – recruitment failure in years 2025–2029

and meant that in the first five years of the projection (2025–2029) recruitment was reduced by 90%. Two further robustness tests with consistent changes to recruitment were created:

- *R: +20%* (robustness set) – recruitment 20% higher

- *R: −20%* (robustness set) – recruitment 20% lower

This affected all projected years and meant that the stock's productivity was higher (*R: +20%*) or lower (*R: −20%*) compared to the baseline operating model, also affecting MSY reference points.

2.5 Uncertainty

Finally, an operating model with changes to the level of observation error was created. The only observation routinely used in the application of the chr rule (after it has been set up) is a biomass index. Therefore, an operating model with changes to the survey index uncertainty was created:

- *Uncertainty: index* (robustness set) – higher observation error for the index (default: +20%)

Survey observations were created following Equation 7. To increase the uncertainty, the standard deviation σ defining ε in Equation 7 was increased by 20%. Additionally, if time permits, a sensitivity analysis with a range of values for σ (0–1) will be conducted for the final management procedure parameterisation.

2.6 Reference points

MSY Reference points for alternative operating models were calculated following the approach set out in Section 1.7. For the operating models that were based on the baseline operating model and did not change operating model characteristics (recruitment failure or alternative observation error levels), the baseline reference points were retained.

Figure 19 illustrates the calculation of MSY reference points for the alternative operating models. F_{MSY} estimates differed between the operating models, and the levels of SSB and Catch (B_{MSY} and MSY) also showed substantial differences. Table 4 summarises the reference points.

Table 5 shows several options for calculating B_{lim} for the alternative operating models. B_{lim} of the baseline operating model was defined as the lowest SSB, which was observed in 2008 (Section 1.7.2). Selecting B_{lim} based on a year or the year with the lowest SSB is potentially arbitrary because it is not based on biological processes or relationships and may refer to different levels of depletion or productivity, depending on the operating model (Table 5).

One exploration to estimate B_{lim} was to link it to recruitment impairment relative to unfished recruitment (R_0). In the baseline operating model, the recruitment at B_{lim} corresponded to $R_{(\mathrm{SSB}=B_{\mathrm{lim}})} = 0.79R_0$, i.e. recruitment was impaired by 21%. This impairment and the operating model-specific stock-recruitment model can be used to define B_{lim} (Table 5). However, this led to some implausible values. For example, for the operating model without discards (*Catch: no discards*), the recruitment steepness is high (Figure 17) and so B_{lim} would have been extremely low (28 tonnes). For other alternative operating models, B_{lim} would have been above B_{MSY} (Table 5).

The final B_{lim} values were calculated with the depletion from the baseline operating model ($B_{\mathrm{lim}} = 0.11B_0$) and applied to the operating model specific B_0. This means that B_{lim} is based on depletion and B_{lim} values ranged from $0.42B_{\mathrm{MSY}}$ to $0.62B_{\mathrm{MSY}}$, which seemed a reasonable range (Figure 17, Table 5).

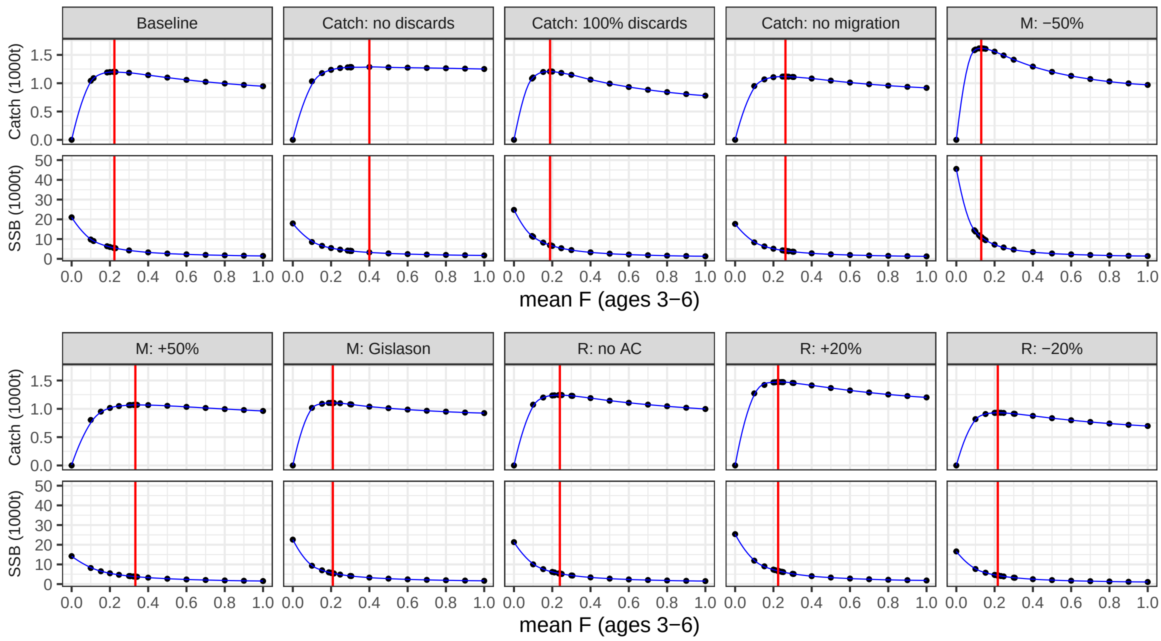

Figure 19: Visualisation of the MSY estimation for all operating models. The points are the long-term averages (median of the last 10 years of a 100-year stochastic projection), the blue curves are a loess smoother fitted to these values. MSY estimates are highlighted with the red vertical lines.

Table 4: Reference points for all operating models. B_0 and R_0 are unfished spawning stock biomass (SSB) and recruitment; F_{MSY}, MSY B_{MSY}, and R_{MSY} are fishing mortality, catch, SSB, and recruitment at MSY; and B_{lim} is the biomass (SSB) limit reference point.

Operating model	B_0 **(t)**	R_0 **(1000)**	F_{MSY}	**MSY (t)**	B_{MSY} **(t)**	R_{MSY} **(1000)**	B_{lim} **(t)**
Baseline	21020	7192	0.223	1201	5476	6533	2352
Catch: no discards	17916	6258	0.400	1286	3234	5979	2005
Catch: 100% discards	24789	8475	0.188	1208	6882	7170	2774
Catch: no migration	17677	6000	0.263	1117	4069	5454	1978
M: −50%	45529	5494	0.130	1617	10983	5313	5095
M: +50%	14217	9573	0.333	1070	3789	7907	1591
M: Gislason	22594	11278	0.209	1105	5519	9572	2528
R: no AC	21325	7132	0.240	1243	5338	6416	2386
R: +20%	25387	8703	0.224	1474	6696	8004	2841
R: −20%	16613	5685	0.217	932	4359	5087	1859

Table 5: Several options for calculating B_{lim} values for all operating models, including B_{lim} relative to B_{MSY} and B_0. Options shown are the B_{lim} based on the lowest observed SSB (B_{loss}), recruitment impairment relative to unfished recruitment ($0.79R_0$), and depletion relative unfished SSB ($0.11B_0$). The final B_{lim} values used in the MSE (Table 4) were those based on depletion. SSB in tonnes.

Operating model	B_0	B_{MSY}		B_{lim} **based on** B_{loss}				B_{lim} **based on** $0.79R_0$			B_{lim} **based on** $0.11B_0$		
	SSB	SSB	SSB/B_0	Year	SSB	SSB/B_0	SSB/B_{MSY}	SSB	SSB/B_0	SSB/B_{MSY}	SSB	SSB/B_0	SSB/B_{MSY}
Baseline	21020	5476	0.261	2008	2352	0.112	0.430	2352	0.112	0.430	2352	0.112	0.430
Catch: no discards	17916	3234	0.181	2008	2156	0.120	0.667	28	0.002	0.009	2005	0.112	0.620
Catch: 100% discards	24789	6882	0.278	2008	2547	0.103	0.370	6914	0.279	1.005	2774	0.112	0.403
Catch: no migration	17677	4069	0.230	2008	1988	0.112	0.489	1434	0.081	0.352	1978	0.112	0.486
M: −50%	45529	10983	0.241	2008	1916	0.042	0.174	1068	0.023	0.097	5095	0.112	0.464
M: +50%	14217	3789	0.266	2008	3048	0.214	0.804	4639	0.326	1.224	1591	0.112	0.420
M: Gislason	22594	5519	0.244	2008	2754	0.122	0.499	4391	0.194	0.796	2528	0.112	0.458
R: no AC	21325	5338	0.250	2008	2352	0.110	0.441	2352	0.110	0.441	2386	0.112	0.447
R: +20%	25387	6696	0.264	2008	2349	0.093	0.351	2352	0.093	0.351	2841	0.112	0.424
R: −20%	16613	4359	0.262	2008	2349	0.141	0.539	2352	0.142	0.540	1859	0.112	0.426

3 References

Fischer, S. H., De Oliveira, J. A. A., Mumford, J. D., & Kell, L. T. (2023). Risk equivalence in data-limited and data-rich fisheries management: An example based on the ICES advice framework. *Fish and Fisheries*, *24*, 231–247. doi:10.1111/faf.12722

Gislason, H., Daan, N., Rice, J. C., & Pope, J. G. (2010). Size, growth, temperature and the natural mortality of marine fish. *Fish and Fisheries*, *11*, 149–158. doi:10.1111/j.1467-2979.2009.00350.x

ICES. (2019). Workshop on North Sea Stocks Management Strategy Evaluation (WKNSMSE). *ICES Scientific Reports*, *1*(12), 378 pp. doi:10.17895/ices.pub.5090

ICES. (2020). The third Workshop on Guidelines for Management Strategy Evaluations (WKGMSE3). *ICES Scientific Reports*, *2*(116), 112 pp. doi:10.17895/ices.pub.7627

ICES. (2021a). ICES fisheries management reference points for category 1 and 2 stocks. In *ICES Technical Guidelines. ICES Advice 2021* 19 pp.. doi:10.17895/ices.advice.7891

ICES. (2021b). Working Group for the Celtic Seas Ecoregion (WGCSE). ICES Scientific Reports, *3*(56), 1082 pp. doi:10.17895/ices.pub.8139

ICES. (2022c). Workshop on ICES reference points (WKREF1). *ICES Scientific Reports*, *4*(2). 70 pp. 10.17895/ices.pub.9822

ICES. (2022d). Workshop on ICES reference points (WKREF2). *ICES Scientific Reports*, *4*(68). 96 pp. 10.17895/ices.pub.20557008

Kaitala, V., & Ranta, E. (2001). Is the Impact of Environmental Noise Visible in the Dynamics of Age-Structured Populations? *Proceedings: Biological Sciences*, *268*, 1769–1774. http://www.jstor.org/stable/3067544

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M. A., Poos, J. J., Scott, F., & Scott, R. D. (2007). FLR: an open-source framework for the evaluation and development of management strategies. *ICES Journal of Marine Science*, *64*, 640–646. doi:10.1093/icesjms/fsm012

Nielsen, A., & Berg, C. W. (2014). Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research*, *158*, 96–101. doi:10.1016/j.fishres.2014.01.014

Punt, A. E., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A. A., & Haddon, M. (2016). Management strategy evaluation: best practices. *Fish and Fisheries*, *17*, 303–334. doi:10.1111/faf.12104

Sharov, A. (2021). The unknown Baranov. Forty years of polemics over the formal theory of the life of fishes. *ICES Journal of Marine Science*, *78*, 743–754. 10.1093/icesjms/fsaa075