



**COMMISSION
SIXTEENTH REGULAR SESSION**
Port Moresby, Papua New Guinea
5 – 11 December 2019

**USING PERFORMANCE INDICATORS TO SELECT A MANAGEMENT PROCEDURE FOR
SKIPJACK**

**WCPFC16-2019-10
19 November 2019**

SPC-OFP

Oceanic Fisheries Programme, The Pacific Community

Executive Summary

Developing a harvest strategy is a stakeholder led process and stakeholder engagement is a key component of the harvest strategy approach. An important area of stakeholder engagement is the communication of results. When developing a harvest strategy, candidate management procedures (MPs) are evaluated using management strategy evaluation (MSE) (Punt et al., 2014; Scott et al., 2016, 2018a). Performance indicators are used by stakeholders to evaluate the expected relative performance and trade-offs between candidate MPs, allowing preferred MPs to be identified (Scott et al., 2018b). Not all indicators will be of interest to all stakeholders and different stakeholders will likely focus on different sets of indicators depending on their priorities for the fishery. It is important that the results from the evaluations, including the indicators, are communicated to stakeholders in a way that allows them to select their preferred MPs.

For the evaluations for the WCPO tuna stocks, each of the performance indicators will be calculated over three time periods (short-, medium- and long-term) (Scott et al., 2018b). Additionally, each indicator reports a distribution of values to reflect the uncertainty in the results and not a single value (except for the probability based indicators, such as the probability of $SB/SB_{F=0}$ being above the limit reference point). As well as the performance indicators, other sources of information can be used to compare the relative performance of the MPs, for example Majuro plots. This means that a lot of information can be presented through the performance indicators and other results from the MSE, making selecting a preferred MP difficult. For example, for the current preliminary results for the skipjack evaluations seven performance indicators are calculated (more will be added later) and each indicator is calculated over the three time periods. This gives 21 indicators to consider for each candidate MP, nearly all of which are distributions rather than single values. As more MPs are evaluated, the amount of information that needs to be communicated to stakeholders can become very large, potentially making it difficult to select a preferred MP.

To assist in the communication of the evaluation results an interactive software tool is being developed (Performance Indicators and Management Procedures Explorer - PIMPLE). The aim of the tool is to facilitate the interactive exploration of the evaluation results, thereby making it easier to compare and evaluate the relative performance of candidate MPs. By selecting and deselecting from the list of available performance indicators and candidate MPs it is possible to 'drill down' into the results. This makes it easier for each stakeholder to focus on the key indicators of interest and consequently identify the preferred MP. As different stakeholders have different preferences for how the results are presented, the tool includes a range of different plot types including bar charts, box plots, time series plots and radar plots as well as summary tables. The development version of the tool can be seen at this address: <https://ofp-sam.shinyapps.io/pimple/>.

This paper is a modified version of WCPFC-SC15-2019/MI-WP-09 (Scott et al., 2019). Changes to this paper include focussing only on the PIMPLE tool and an updated PIMPLE user guide.

Acknowledgments

We gratefully acknowledge funding for this work from the New Zealand Ministry of Foreign Affairs and Trade (MFAT) funded project "Pacific Tuna Management Strategy Evaluation"; In addition we thank the Center for High Throughput Computing (CHTC UW-Madison) for generously providing access to their computing resources.

References

- Punt, A. E., Butterworth, D., de Moor, C., De Oliveira, J., and Haddon, M. (2014). Management strategy evaluation: best practices. *Fish and Fisheries*, (DOI:10.1111/faf12104).
- Scott, F., Scott, R. D., Davies, N., Pilling, G., and Hampton, S. (2018a). Technical developments in the MSE modelling framework. WCPFC-SC14-2018/MI-IP-02, Busan, South Korea, 5–13 August 2018.
- Scott, F., Scott, R. D., Pilling, G., and Hampton, S. (2018b). Performance indicators for comparing management procedures using the MSE modelling framework. WCPFC-SC14-2018/MI-WP-04, Busan, South Korea, 5–13 August 2018.
- Scott, F., Scott, R. D., Pilling, G., Yao, N., and Hampton, S. (2019). Harvest strategy engagement tools. WCPFC-SC15-2019/MI-WP-09, Pohnpei, Federated States of Micronesia, 12–20 August 2019.
- Scott, R., Davies, N., Pilling, G. M., and Hampton, J. (2016). Retrospective forecasting of the 2014 WCPO bigeye tuna stock assessment. WCPFC-SC12-2016/SA-WP-02, Bali, Indonesia, 3–11 August 2016.

A PIMPLE Users Guide

Performance Indicators and Management Procedures Explorer - PIMPLE

User Guide

Finlay Scott - OFP, SPC

2019-18-11

Contents

Introduction

What is it?	
How do I use PIMPLE?	

Compare performance

Bar charts	
Box plots	
Time series plots	
Radar plots	
Table	

Explore indicators

PI 1 & 8: Biomass	
PI 3: Catches	
PI 4: Relative CPUE	
PI 6: Catch stability	
PI 7: Relative effort variability	

Appendix 1: Performance indicators

Summary table of objectives and corresponding indicators	
Indicator $SB/SB_{F=0}$	
Indicator 1. Probability of $SB/SB_{F=0} > LRP$	
Indicator 3. Catch relative to the average catch in 2013-2015	
Indicator 4. CPUE relative to the CPUE in 2010	
Indicator 6. Catch stability	

Indicator 7. Effort stability

Indicator 8. The average distance of $SB/SB_{F=0}$ from the TRP

Introduction

What is it?

Performance Indicators and Management Procedures Explorer (PIMPLE) is a tool for exploring and comparing the relative performance of alternative candidate management procedures (MPs) and harvest control rules (HCRs). The relative performance can be explored using a range of different plots and tables. This allows trade-offs between the different HCRs to be evaluated. This guide gives a brief overview of how to use PIMPLE to explore the results and some guidance on how to interpret the various plots.

The performance of each HCR is measured using a range of *performance indicators* (PIs). More details of the PIs are given in the Appendix. Not all of the desired PIs are currently available and will be added soon.

Note that when evaluating the performance of different HCRs we are more interested in comparing their relative performance (e.g. is HCR 1 better than HCR 2?) rather than comparing the performance of the HCRs to some period in the past (e.g. is HCR 1 better than some situation 10 years ago?).

The demonstration results included with PIMPLE are based on skipjack tuna.

In the current data base there are 13 HCRs, including some ‘status quo’ options with different levels of constant fishing effort. The indicators are calculated over three different time periods: short-, medium- and long-term. The length of each period is three management cycles, where one management cycle is three years. The year range of each period is: short-term (2016:2024) medium-term (2025:2033) long-term (2034:2042).

Note that the values and figures in the current data base are for demonstration purposes only.

To use PIMPLE you need to install the *AMPLE* package from CRAN.

How do I use PIMPLE?

Under the tabs at the top of the screen are the two main tabs of PIMPLE: **Compare performance** and **Explore indicators**. The **Compare performance** tab allows HCRs to be compared across the full range of PIs. The **Explore indicators** tab allows detailed inspection of each type of PI and metric; These tabs are explained in more detail in the rest of this user guide.

Depending on which tab is open, you will be able to select different candidate HCRs and different PIs using the selection boxes on the left of the screen.

For a description of each indicator see the Appendix at the end of this document. A description of the different plot types is found in the **Comparing MPs** section below.

Compare performance

The plots and tables in this tab enable the comparison of the candidate HCRs across multiple PIs and time periods. This allows trade-offs between PIs and time periods to be evaluated.

At the time of writing there were five subtabs under the **Comparing MPs** tab. Each subtab offers a different method of presenting the information. These tabs are:

- Bar charts
- Box plots
- Time series plots
- Radar plots
- Table

These are described in more detail below.

The HCRs under inspection can be chosen using the **HCR selection** boxes on the left. The PIs of interest can also be chosen using the **PI selection** boxes on the left. As you choose different PIs and HCRs the plots and tables will update to reflect the selection. It may take a little time to update the plots. By selecting different HCRs and PIs it is possible to 'drill down' into the results.

The catch based indicators can be presented as either the total catches in the whole region or only for the purse seine catches in assessment regions 2, 3 and 5.

Bar charts

The bar charts in PIMPLE gives the *median* value of the indicator. This means that they do not consider the range of values for each indicator and therefore do not consider the uncertainty. The median value is calculated for each of the time periods (short-, medium- and long-term).

The results from the different HCRs have different colours, as shown by the legend.

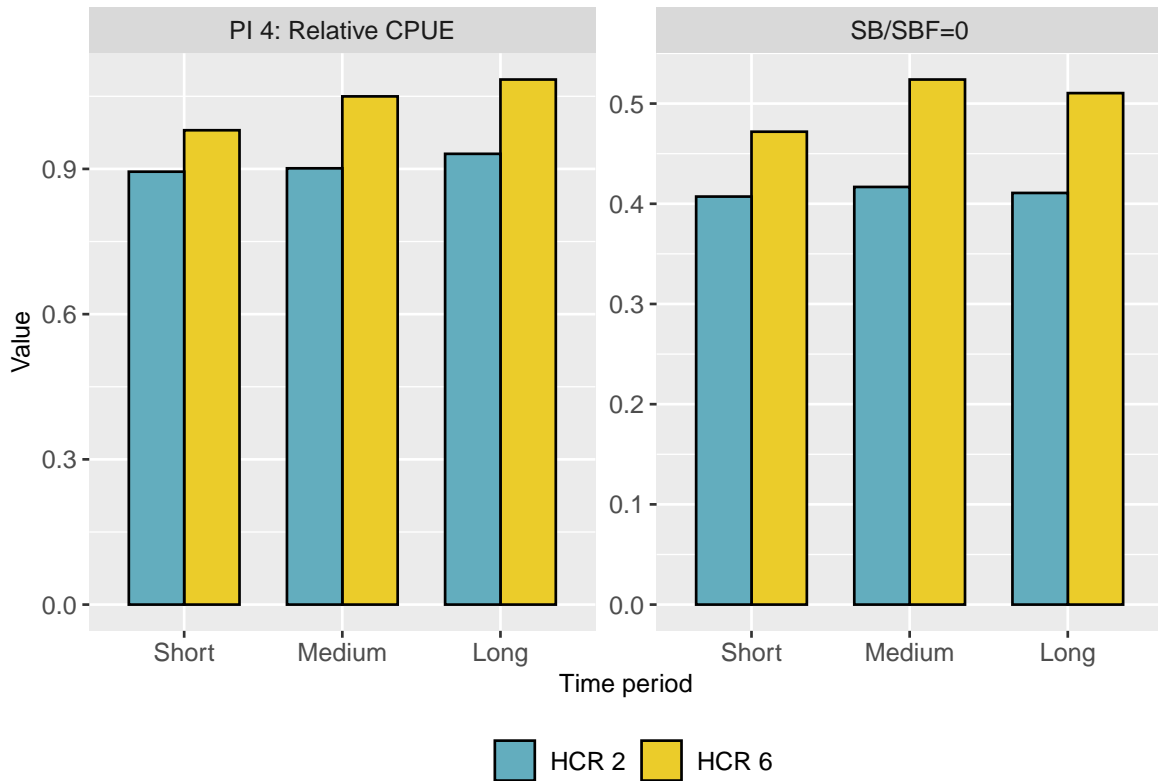


Figure 1: Example bar chart showing median values of two indicators for two HCRs across the three time periods.

Box plots

The box plots allow the uncertainty of the different indicators and metrics to be explored. The coloured box covers the 20-80th percentile. The vertical whisker covers the 5-95th percentile. The horizontal line shows the median. The median line will be in the same position as the height of the bar charts. The larger the box and the longer the whisker, the greater uncertainty there is about the value of the indicator.

As with the bar charts the results from the selected HCRs have different colours, as shown by the legend. The value in each of the time periods (short-, medium- and long-term) is also shown.

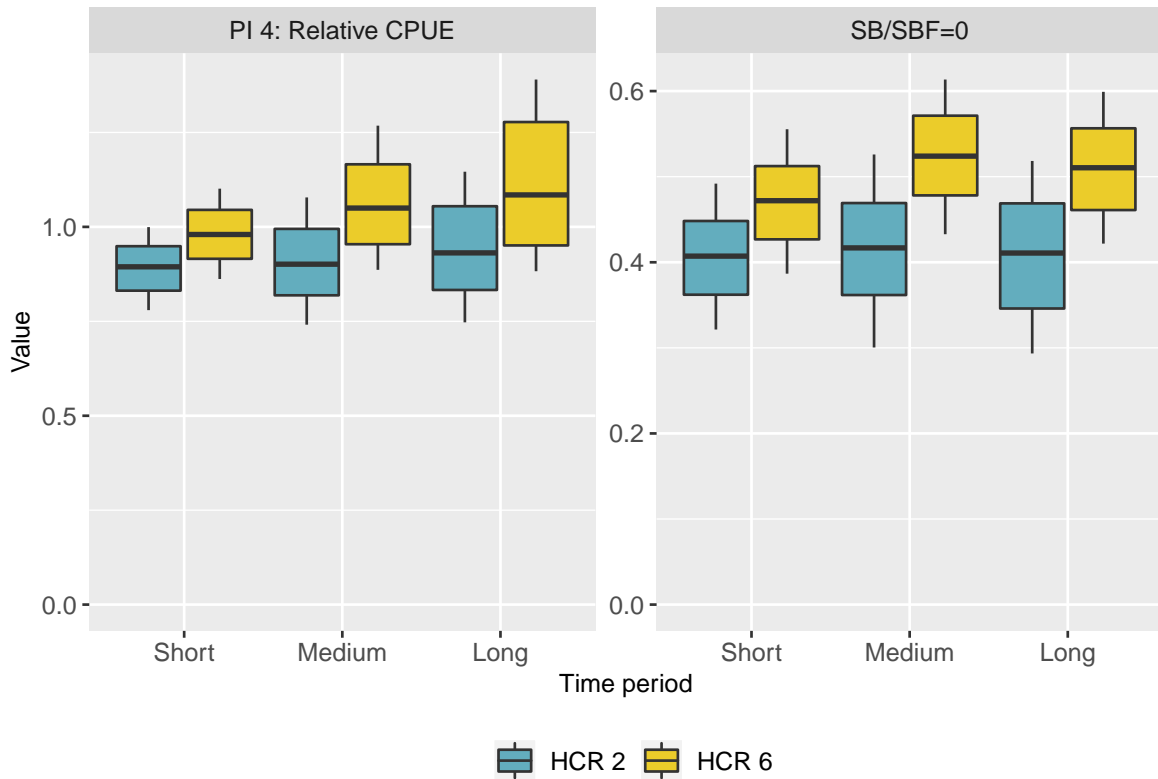


Figure 2: Example box plot showing median values of two indicators for two HCRs across the three time periods.

Time series plots

The time series plots show a coloured envelope for each HCR. The envelope contains the 10-90th percentile of the values. The black dashed line shows the median value. The vertical lines mark the short-, medium- and long-term periods. Several individual trajectories (known as worms or spaghetti) can be shown as illustrations. The results from the selected HCRs have different colours as shown by the legend.

The **Time series plots** tab shows time series plots for *SB/SBF=0*, *Catch* and *Relative CPUE*.

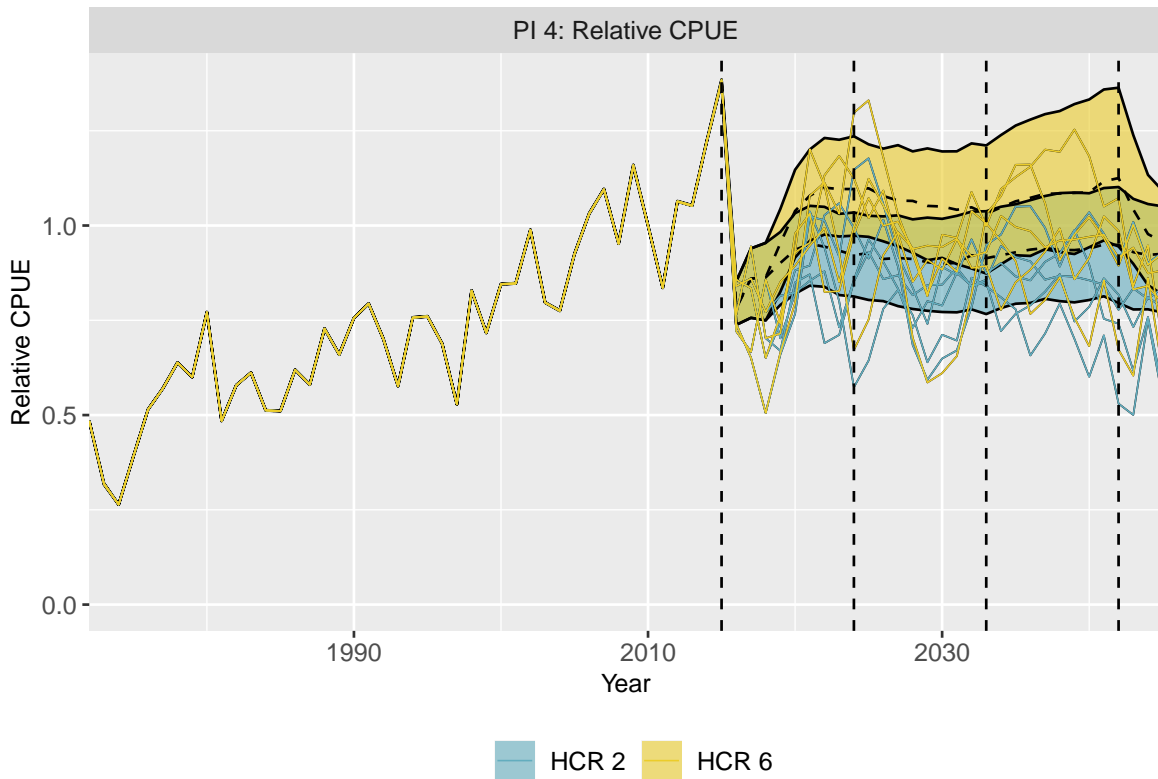


Figure 3: Example time series plot of the relative CPUE for three HCRs. The envelope contains the 10-90th percentile of the values. The black dashed line shows the median value. The coloured lines are example individual trajectories.

Radar plots

Radar plots look quite interesting but they must be handled with care. The indicators are spread around a circle and the median value for each indicator for each HCR is shown. The bigger the area covered by each HCR the better it is performing. This assumes that each PI is of equal importance.

A radar plot is shown for each time period. As with the bar charts and box plots the results from the selected HCRs have different colours, as shown by the legend. Only the selected PIs are shown which means that you need at least three PIs selected for the plot to make any sense.

Note that only the indicators for which 'bigger is better' are included on the radar plot.

As the indicators have different ranges of values it is necessary to scale the values to allow comparison on the same plot. The values within each indicator are scaled by the maximum value within each indicator. The relative values of the indicators are preserved. This means that if the values of an indicator are similar for all the HCRs, the difference between them will be difficult to see.

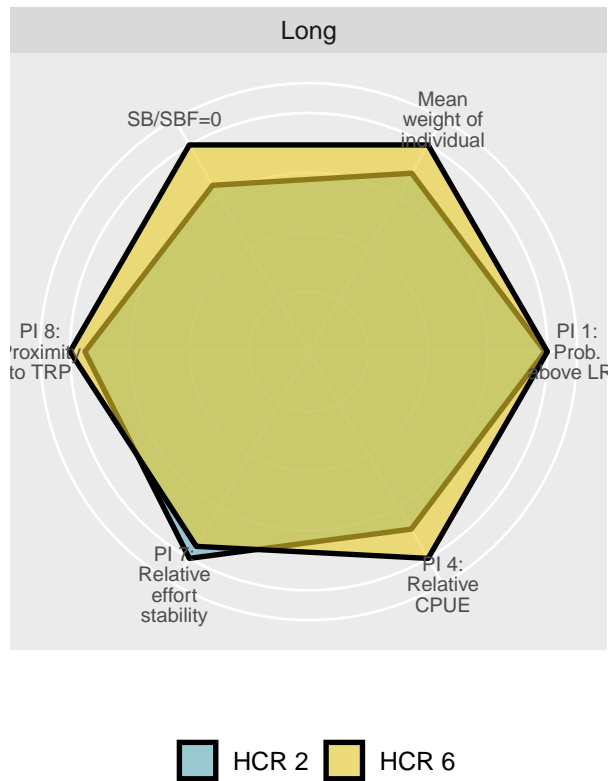


Figure 4: Example radar plot for three HCRs and 8 indicators with results being scaled by the max. Note that there is not much contrast between the HCRs. Only the long-term period is shown.

Table

The **Table** subtab shows the values of the selected indicators and HCRs in each of the three time periods. The indicators are presented as the median value, and then two more values inside parentheses e.g. *0.46 (0.42, 0.5)*. The values inside the parentheses are the 10-90th percentile respectively.

PI 1 (the probability of being above the LRP) only has a single value because it is a probability.

Explore indicators

This tab offers the ability to explore the performance of the HCRs in detail. There are several subtabs, each one offering different plots of the same indicator or theme. The HCRs under inspection can be chosen using the **HCR selection** boxes on the left.

The plots are a mixture of time series plots, bar charts and box plots. For more information on these plot types see the *Comparing performance* section later on.

PI 1 & 8: Biomass

This tab shows metrics and indicators related to biomass (PIs 1 and 8).

There are time series plots, bar charts and box plots of $SB/SBF=0$ and the proximity of $SB/SBF=0$ to the TRP (PI 8) in the different time periods. There is also a bar chart showing PI 1 (the probability of $SB/SBF=0$ being below the Limit Reference Point (LRP)).

Note that for PI 8, a value of 1 indicates being at the TRP and a value of 0 indicates being far from the TRP in any direction. A higher value is therefore considered better than a low value.

It is possible to include sample individual trajectories by selecting the *Show trajectories* option on the left.

PI 3: Catches

This tab shows metrics and indicators related to the expected catches (PI 3), relative to the average catch in 2013-2015.

Using the options on the left of the screen, it is possible to plot a time series, as a bar chart or a box plot of the average catches in the different time periods. Additionally, either the catches relating to the whole region, only purse seine catches in areas 2, 3, and 5 of the stock assessment model, or total catches in individual model areas can be shown.

PI 4: Relative CPUE

This tab is concerned with the CPUE relative to the CPUE in 2010 (PI 4). Note that the CPUE is only calculated for purse seine fisheries in assessment regions 2, 3 and 5 excluding the associated purse seines in region 5. This tab has a time series plot of relative CPUE as well as a bar chart and box plot of the average relative CPUE in the different time periods.

PI 6: Catch stability

The catch stability through time is presented as either the *variability* and the *stability* (PI 6). The higher the variability, the lower the stability. A stability of 1 means that the catches do not change over time. The lower the value of stability, the more catches change over time. Both the variability and stability are presented as bar charts and box plots in the three time periods.

Like the **PI 3: Catches** tab, the catches used to calculate the stability and variability are from either the total catches in the whole region, or only for the purse seine catches in assessment regions 2, 3 and 5 or in individual model areas.

Note that the size of the boxes in the box plots do not reflect the uncertainty in the catches, but uncertainty in what the stability or variability of the catches will be.

PI 7: Relative effort variability

Relative effort is the effort relative to the effort in 2010 (PI 7). Note that the relative effort is only calculated for purse seine fisheries in regions 2, 3 and 5 of the stock assessment excluding the associated purse seines in region 5.

The relative effort variability is presented as either the *variability* or the *stability*. A stability of 1 means that the relative effort does not change over time. The lower value of stability, the more the relative effort changes over time. Both the variability and stability are presented as bar charts and box plots.

Appendix 1: Performance indicators

This section presents details on the performance indicators that are currently calculated. The calculations for these indicators were initially presented at the WCPFC Scientific Committee Fourteenth Regular Session (SC14) in Busan, Korea (WCPFC-SC14-2018/MI-WP-04). These calculations are summarised and, where appropriate, updated here. Note that only those indicators which are currently calculated are included. It is anticipated that other indicators will be included later.

Summary table of objectives and corresponding indicators

PI	Objective type	MOW4 Strawman	Performance Indicator (WP14)
	Biological		$SB/SB_{F=0}$
1	Biological	Maintain SKJ (and YFT and BET) biomass at or above levels that provide fishery sustainability throughout their range	Probability of $SB/SB_{F=0} > LRP$
3	Economic	Maximise economic yield from the fishery	Catch relative to the average catch in 2013-2015
4	Economic	Maintain acceptable CPUE	CPUE relative to the CPUE in 2010
6	Economic	Catch stability	Average annual variation or stability in catch
7	Economic	Stability and continuity of market supply	Average annual variation or stability in relative effort
8	Economic	Stability and continuity of market supply	The average distance of $SB/SB_{F=0}$ from the TRP

Table 1: Summary of performance indicators currently calculated for the purse seine skipjack tuna fishery and the associated objectives.

Indicator $SB/SB_{F=0}$

The indicator $SB/SB_{F=0}$ is the ratio of the adult biomass to the adult biomass in the absence of fishing (sometimes known as depletion).

The model for evaluating the MPs reports the biomass (SB) and the unfished biomass ($SB_{F=0}$) by season and model area. The indicator is based on the mean annual biomass which is calculated by taking the mean biomass over the seasons. Although it is possible to calculate the indicator for each of the five areas in the model, here only the total biomass across all areas is considered. The unfished biomass is taken as the mean of the unfished biomass over a ten year moving window. $SB/SB_{F=0}$ is a measure of stock status and can be compared to the Limit Reference Point (0.2) and the Target Reference Point (0.5, interim) (see performance indicators 1 and 8).

Indicator 1. Probability of $SB/SB_{F=0} > LRP$

This indicator is linked to the objective: Maintain SKJ biomass at or above levels that provide fishery sustainability throughout their range. It measures the probability of $SB/SB_{F=0} > LRP$ where 0.2 is the LRP for skipjack.

The probability of $SB/SB_{F=0} > 0.2$ in a particular year is calculated as the proportion of the replicates in a year in which $SB/SB_{F=0} > 0.2$. The indicator in each time period is then calculated as the average probability over the different time periods (short-, medium-, long-term).

Indicator 3. Catch relative to the average catch in 2013-2015

This indicator is linked to the objective: Maximise economic yield from the fishery (average expected catch). It is based on the average expected catch, relative to the average catch in the years 2013-2015.

The model used for evaluating the MPs reports catches at the area and fishery level. It is therefore possible to calculate this indicator for each fishery and each region in the model. %However, this would result in a large number of values for this indicator, making it challenging to understand the overall relative performances of the candidate MPs. Here the indicator is calculated for several different area and fishery groupings:

- Total catch over all areas and fisheries;
- Purse seine catches in areas 2, 3 and 5;
- Catches from all fisheries in individual areas.

The indicator is calculated by taking the average catch, relative to the average catch in 2013-2015, of each replicate over the desired year range (short-, medium- and long-term) for each area and fishery grouping. This gives a distribution of expected relative catches in each time period for each area and fishery grouping.

Indicator 4. CPUE relative to the CPUE in 2010

This indicator is linked to the objective: Maintain acceptable CPUE. It is based on the predicted skipjack CPUE relative to the CPUE in reference period levels. Here, the reference period level is taken to be 2010.

The MP evaluation model reports effort and catch at the fishery level which means it could be possible to calculate the indicator at the fishery, area and total level. However, it is not always possible to combine the fishing effort from the different fisheries in the model as the historical effort of each fishery may or may not have been standardised (this does not affect the simulations as internally the effort is scaled). This indicator is therefore calculated in each year for the purse seine fisheries operating in model regions 2, 3 and 5 excluding the associated purse seine fishery in region 5 which has a standardised effort index:

Indicator 6. Catch stability

This indicator is concerned with the average annual variation and stability in catch. As with indicator 3, it is possible to calculate this indicator by fishery and area. Here it is calculated using the same fishery and area groups as indicator 3.

This indicator is calculated in two ways: the variability and the stability. The variability is calculated by taking the absolute annual difference of the catch for each simulation and in each year. This measures the variation in the catch, i.e. the higher the value of the indicator, the less stable the catch is. It is assumed that an MP that produces a lower catch variation, i.e. a low value of the indicator, would be preferred.

As well as the variability in the catch, the stability (the inverse of the variability) can be calculated. This involves rescaling the variability so that it is between 0 and 1. A stability of 1 implies that the catch does not change at all over time, i.e. it is completely stable. A stability of 0 means that the catch is very variable in time.

To calculate stability it is necessary to decide what level of variability gives a stability of 0. Here it is assumed that the variability that implies a stability of 0 is the 95th percentile of the observed catch variability across all time periods and MPs, for each fishery and area grouping. This assumption means that the stability can only be interpreted in relative terms (i.e. when comparing between MPs and time periods).

The advantage of using stability instead of variability is that a large value is preferred to a small value. This is in line with the other indicators where a large value is generally preferred. For this reason, the stability is used in PIMPLE as the indicator.

Indicator 7. Effort stability

This indicator is linked to the objective: Stability and continuity of market supply.

This indicator is concerned with the variation of relative effort (effort relative to the effort in a reference period). Here the reference period is taken to be 2010. As with performance indicator 4, this indicator is calculated for the purse seine fisheries operating in regions 2, 3 and 5 excluding the associated purse seine fishery in region 5 which has a standardised effort index.

The indicator is calculated in a similar way to performance indicator 6. The absolute annual difference of the effort relative to the base effort in 2010 is calculated for each replicate in each year. Like indicator 6, this indicator measures the variation in the relative effort, i.e. the higher the value of the indicator, the less stable the relative effort is. It is assumed that MPs that produce a low value of this indicator are considered to be perform better.

An alternative indicator that measures the stability (the inverse of the variability) is also calculated in the same way as the stability of the catches (indicator 6). Stability is used in PIMPLE as it is in line with the other indicators, where bigger is better.

Indicator 8. The average distance of $SB/SB_{F=0}$ from the TRP

This indicator is also linked to the objective: Stability and continuity of market supply (along with indicator 7). It is concerned with maintaining the stock size around the TRP levels (where the interim

TRP for skipjack is $SB/SB_{F=0} = 0.5$). It is assumed that the further away $SB/SB_{F=0}$ is from 0.5, the worse the HCR can be thought to be performing, i.e. it is better to have $SB/SB_{F=0}$ close to 0.5 on average.

The deviation is calculated as the absolute difference between $SB/SB_{F=0}$ and the TRP, scaled by the maximum distance from the TRP (for skipjack this is 0.5). This value is subtracted from 1 to rescale it so that an indicator value of 1 implies that $SB/SB_{F=0}$ is exactly at the TRP and a value of 0 is as far from the TRP as possible.

Note that this indicator is slightly different to comparing the average value of $SB/SB_{F=0}$ to the TRP. If the average value of $SB/SB_{F=0}$ over a range of years is close to the TRP it does not necessarily mean that $SB/SB_{F=0}$ was close to the TRP during those years. For example, if $SB/SB_{F=0}$ oscillated between 0.3 and 0.7 each year, the average value would be 0.5 but it would never have been very close to 0.5 in any year.