



CSIRO
MARINE RESEARCH

**AN OVERVIEW OF POTENTIAL GRAPHICS FOR
EVALUATING THE PERFORMANCE OF CANDIDATE
MANAGEMENT PROCEDURES FOR SOUTHERN BLUEFIN
TUNA**

**Paige Eveson
Daniel Ricard**

**Prepared for the CCSBT 2nd Management Procedure Workshop
7-9 and 13-14 April 2003, Queenstown, New Zealand, CCSBT-MP/0304/05**

Table of Contents

Abstract.....	1
Introduction.....	1
Worm plots.....	1
Comparison plots.....	2
Summary plots.....	2
Discussion.....	3

Abstract

Potential graphical methods for evaluating the performance of candidate management procedures (MP's) are presented. Three types of graphs are proposed: 1) worm plots, which show the projected biomass and catch trajectories; 2) comparison plots, which compare the performance statistics obtained from multiple MP's for a single operating model scenario or from all eight operating model scenarios for a single MP; and 3) summary plots, which compare multiple MP's across all eight operating model scenarios. Each graph type is described in detail, and its features and drawbacks discussed.

Introduction

In order to facilitate the evaluation and comparison of candidate MP's, we have developed some graphical methods (using statistical software package R 1.6.2) for summarising and displaying the results.

First it is useful to specify some terminology. We refer to the general principle for setting the TAC as a "decision rule" (e.g. keep the TAC at a constant C). The decision rule will most likely have tuning parameters (e.g. C), and a specific set of tuning parameters is referred to as a "version" of the rule. A decision rule plus a version define an MP.

Each MP is evaluated under eight different operating model scenarios (h6mcmc is not currently being run) and three simulation hierarchies (the full posterior case is not currently being run). For hierarchies 1 to 3 (denoted H1, H2 and H3), there are also three options for the maximum posterior density (MPD) fit; however, for the eight model scenarios being considered, only option 1 is being implemented at this stage.

All graphs are based on data in the .sum files produced by the CCSBT management procedures software. These files contain the median, 10th and 90th percentiles over all projection replicates of the 8 performance statistics listed in Attachment 4, CCSBT-SC/0209/04. For H2 and H3, we elected to always run 100 projection replicates (only 1 replicate needs to be run for H1 since this hierarchy assumes no error). In addition, the .sum files contain the median, 10th and 90th percentiles of the biomass projections for years 2002 to 2021 and the catch projections for years 2002 to 2020. Up to ten individual realizations of the biomass and catch trajectories (referred to as "bworms" and "cworms") are also given.

Worm plots

At the most specific level, we created graphs of the biomass and catch trajectories for a particular MP under a particular operating model scenario and hierarchy (recall that for the MPD fit, only option 1 is being used at this point). The median values are plotted along with error bars extending to the 10th and 90th percentiles. The plotting function has the option to include anywhere from 0 to 10 bworms and cworms. Figure 1 gives an example of such a plot with 10 bworms and 5 cworms.

Comparison plots

Next we created graphs to compare the performance of either multiple MP's under a single operating model scenario and hierarchy, or a single MP across all eight operating model scenarios (for a specific hierarchy). A plot is created for each performance statistic showing the median value along with error bars extending to the 10th and 90th percentiles. Figure 2 shows an example for option a), with a selection of MP's along the x-axis. Figure 3 shows an example for option b), with the eight operating model scenarios along the x-axis. The plots of the performance statistics are found in the upper eight panels of the figures, with catch-related statistics on the left and biomass-related statistics on the right. The catch statistics labelled C 5yr.avg and C 20yr.avg, which are the average catch over the first 5 years of the projection period and over the entire 20 years of the projection period, have been rescaled such that they are relative to the 2001 catch level (i.e. the median, 10th and 90th percentile values have been divided by 15965.9 MT since this is the 2001 total catch value¹). As such, a value of 1.0 can be used as a reference point for how an MP performs with respect to current catch levels. Note that the average ratio of the surface catch to the total catch (denoted avg(Surf/TotC)) remains the same for all model scenarios and MP's. This is because we are not setting fishery-dependent TAC's in any of our decision rules to date. Although this is not currently an interesting performance statistic to look at, we include it in the figures in case fishery-dependent decision rules are investigated in the future.

The correlation between the two average catch statistics and among the four biomass statistics is often high. If this is the case, we can reduce the number of performance statistics that needs consideration by looking at composite statistics, such as averages. In the bottom left panel of Figures 2 and 3 we have plotted a composite catch indicator (IC), which is the average of the two average catch statistics. The correlation between the two catch statistics (calculated using whichever MP's and model scenarios are included on the figure) is written in the upper right corner of the plot. The closer the correlation is to 1.0, the more meaningful the composite value. Similarly, in the bottom right panel we have plotted a composite biomass indicator (IB), which is the average of the four biomass statistics. There are six possible correlations between the four indicators, so the minimum of these correlations is written in the upper right corner of the plot. In practice, we found that the correlation between the two catch statistics is most often very high, however the correlation between the non-spawning stock biomass statistic and any of the other spawning stock biomass statistics can often be quite low. In such cases, the composite biomass indicator must be interpreted with caution.

A final note regarding these figures—the y-axis scale for each panel has been kept fixed for comparability between figures.

Summary plots

The graphs produced so far contain a lot of information that can be difficult to process without careful inspection. For example, we often have a selection of MP's whose performance we want to compare across operating model scenarios; this would currently require comparing the output from multiple graphs. Furthermore, from the

¹ The total catch in weight for a year is generated by the operating model and differs depending on the operating model scenario being assumed, thus we used the average value from the eight scenarios.

graphs presented so far, it is difficult to evaluate the trade-off between catch performance and conservation (i.e. biomass) performance. Consequently, we have created summary graphs that concentrate on just one measure of catch (IC) and one measure of biomass (usually IB, but other performance statistics are also considered) to assist with such assessments.

Figure 4a shows a summary graph comparing multiple MP's across all eight operating scenarios. At the top of the figure are boxplots of IC and IB, where the box for a specific MP gives the median and interquartile range of the eight IC or IB values obtained under the eight different operating model scenarios. Below the boxplots are two plots showing the trade-off between IC and IB; they contain the same information, just expressed in different ways. On the left-hand side is what we term a "star plot", with each star corresponding to a particular MP. The centroid of the star represents the mean IC and mean IB value across the eight model scenarios. It is marked with a circle whose radius is proportional to the inter-annual variability (IAV) in the catch, so that a smaller circle is desirable. Each star has eight rays, where each ray extends to the (IC, IB) coordinate obtained under one of the scenarios. The star plot does not allow one to distinguish which ray corresponds to which scenario, which is necessary if one wishes to compare performance of MP's under a particular scenario (say h3m10). This gives rise to the plot on the right-hand side of the page, where the results from a given model scenario are connected by a line and labelled accordingly.

As already noted, IB may not be the most meaningful indicator if the four biomass performance statistics are not highly correlated. It may be preferable to consider individual performance statistics instead. Of particular interest are the ratio of the spawning stock biomass in 2020 to 1980 (SB2020:1980), and the ratio of the spawning stock biomass at the end of the projection period to the start of the period (SB2022:2002). In Figure 4b, plots equivalent to those on the bottom of Figure 4a are shown but with SB2020:1980 and SB2022:2002 replacing IB on the y-axis.

Note that on all of the summary plots, the results from running an MP that sets the TAC to zero in all years is included as a reference (labelled "no catch"). The no catch case shows the maximum conservation performance that can be attained under each of the operating model scenarios, and serves as a realistic basis for evaluating the conservation level achieved by other MP's. For example, under model scenario h3m10, the maximum value of SB2020:1980 that can be attained (with no catch) is approximately 0.5; therefore, other MP's cannot be expected to do any better under this scenario. This is an important observation because it means our management objective goal of achieving a value of 1.0 for this statistic cannot be met under certain scenarios.

Discussion

The summary plots can be particularly useful for identifying key features of an MP and key differences between MP's. However, in considering only the summary plots, one may overlook some important observations. For example, the summary plots use only the median values of the performance statistics and give no indication of the amount of variability in these statistics. To get this information, one needs to look at the more detailed comparison plots (as in Figures 2 and 3). Neither the summary plots nor the comparison plots can be used to assess the trend in the biomass and catch

trajectories. For example, two MP's may lead to the same average catch over the 20-year projection period but one may achieve this with a constant catch level and another with a steadily increasing trend in catch. Or the biomass performance statistic SB2020:1980 may have a value close to one, but the trend in the biomass at the end of the projection period may be going down. The worm plots are important for investigating these issues.

Figure 1.

Projections for decision rule Fox v7
using model h6m10 hierarchy H3 and MPD1

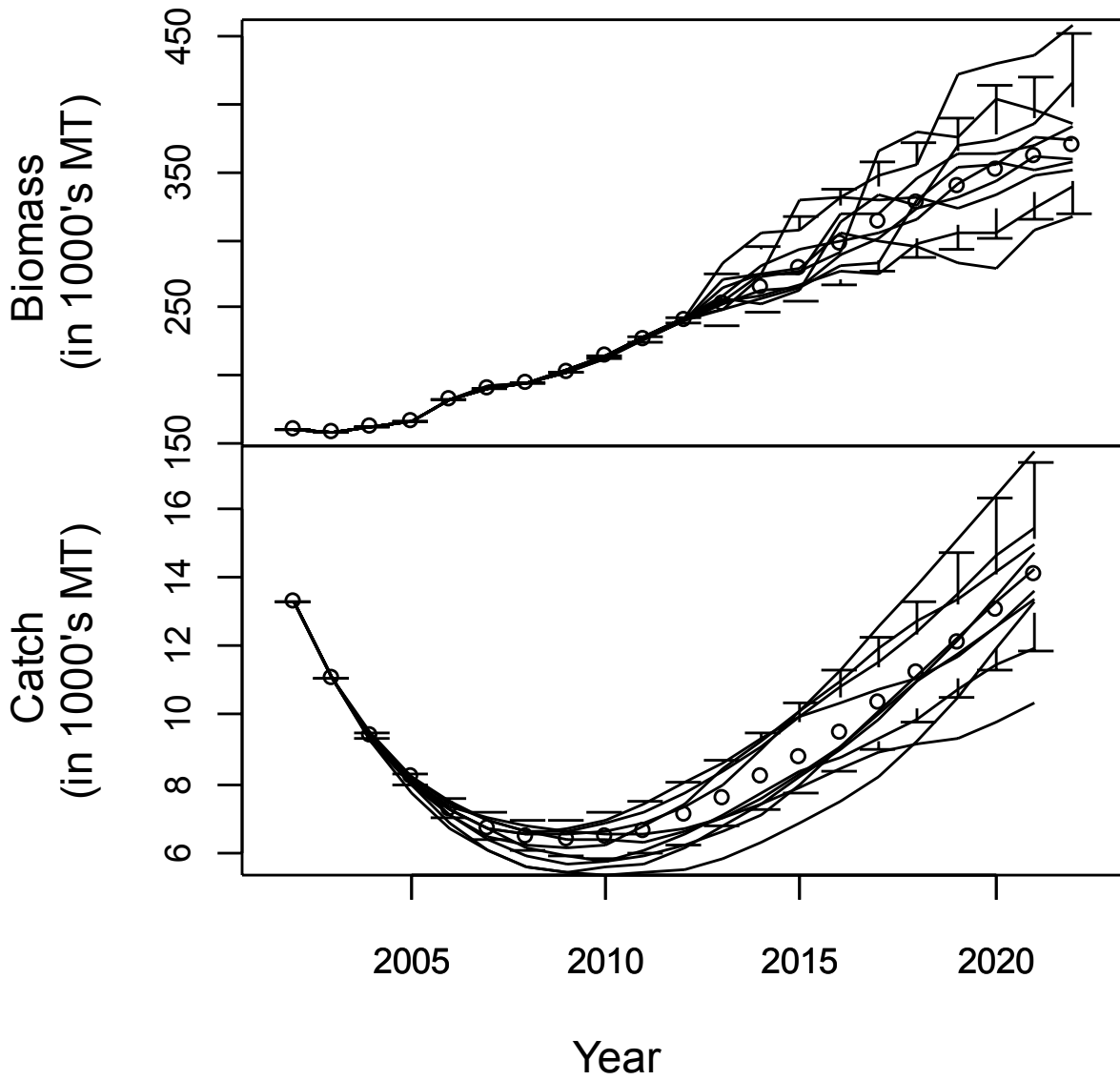


Figure 2.

Model h6m10
(hierarchy H3 and MPD1)

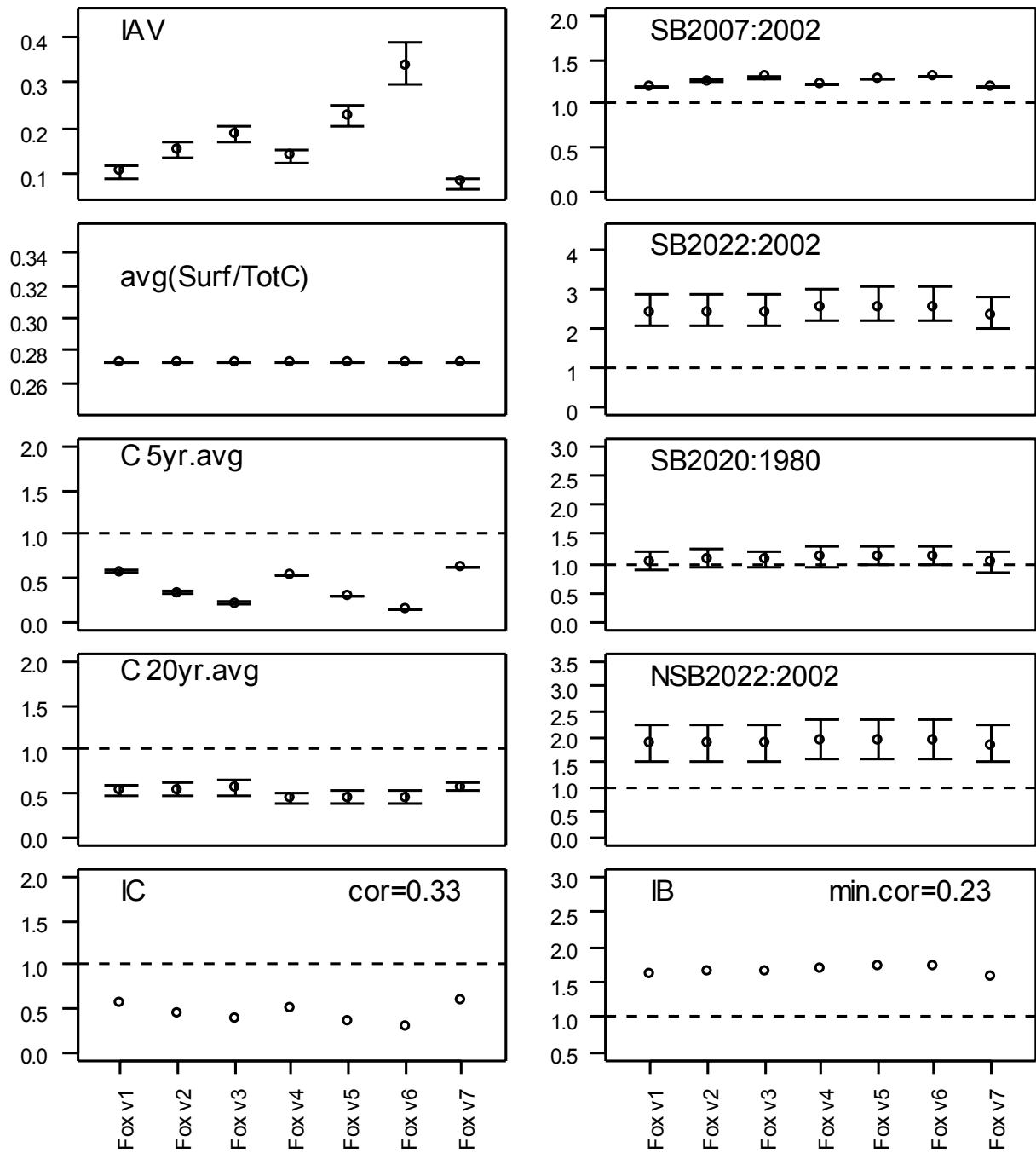


Figure 3.

Decision rule Fox version v7
(hierarchy H3 and MPD1)

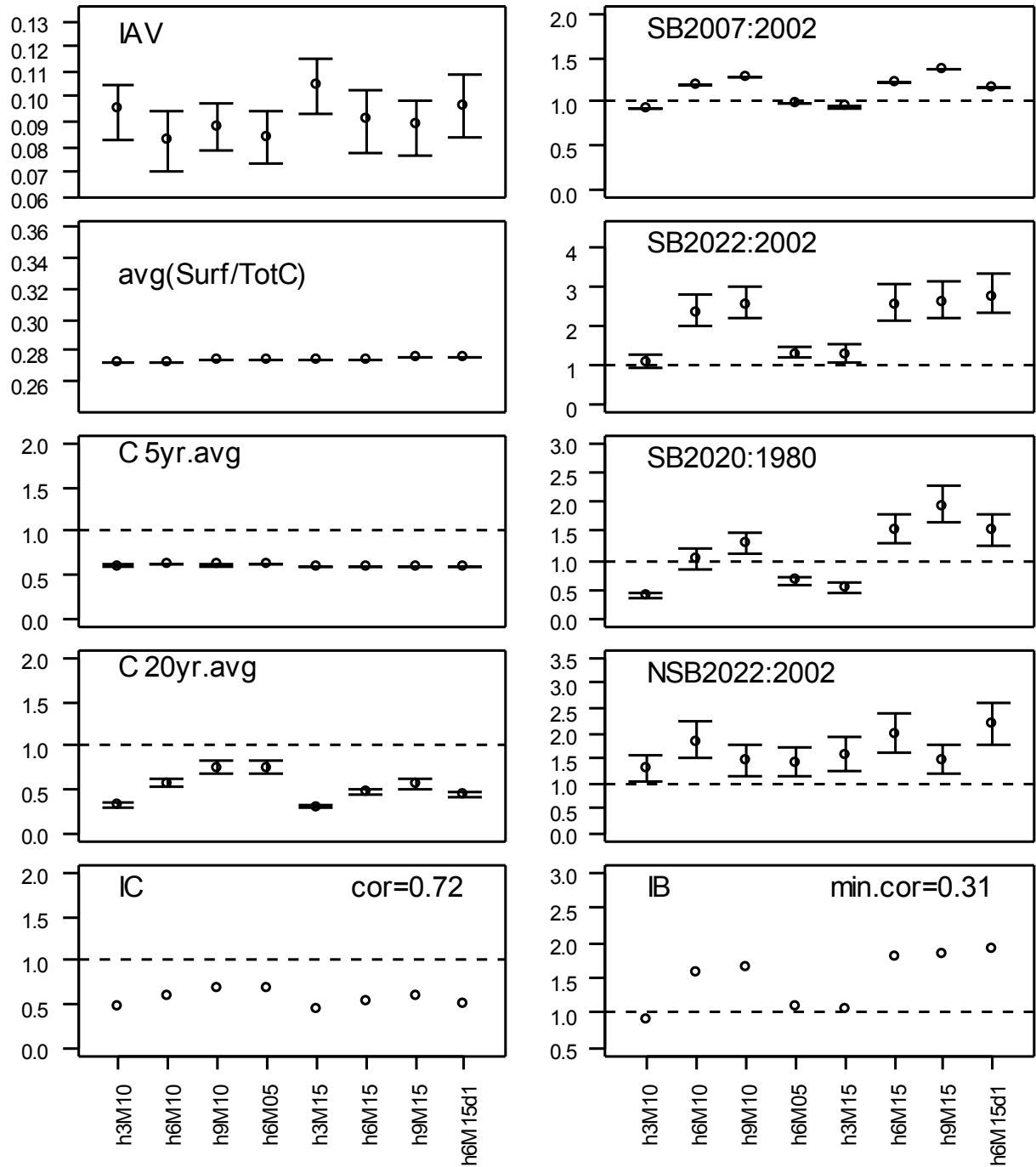


Figure 4a.

Summary over all operating model scenarios

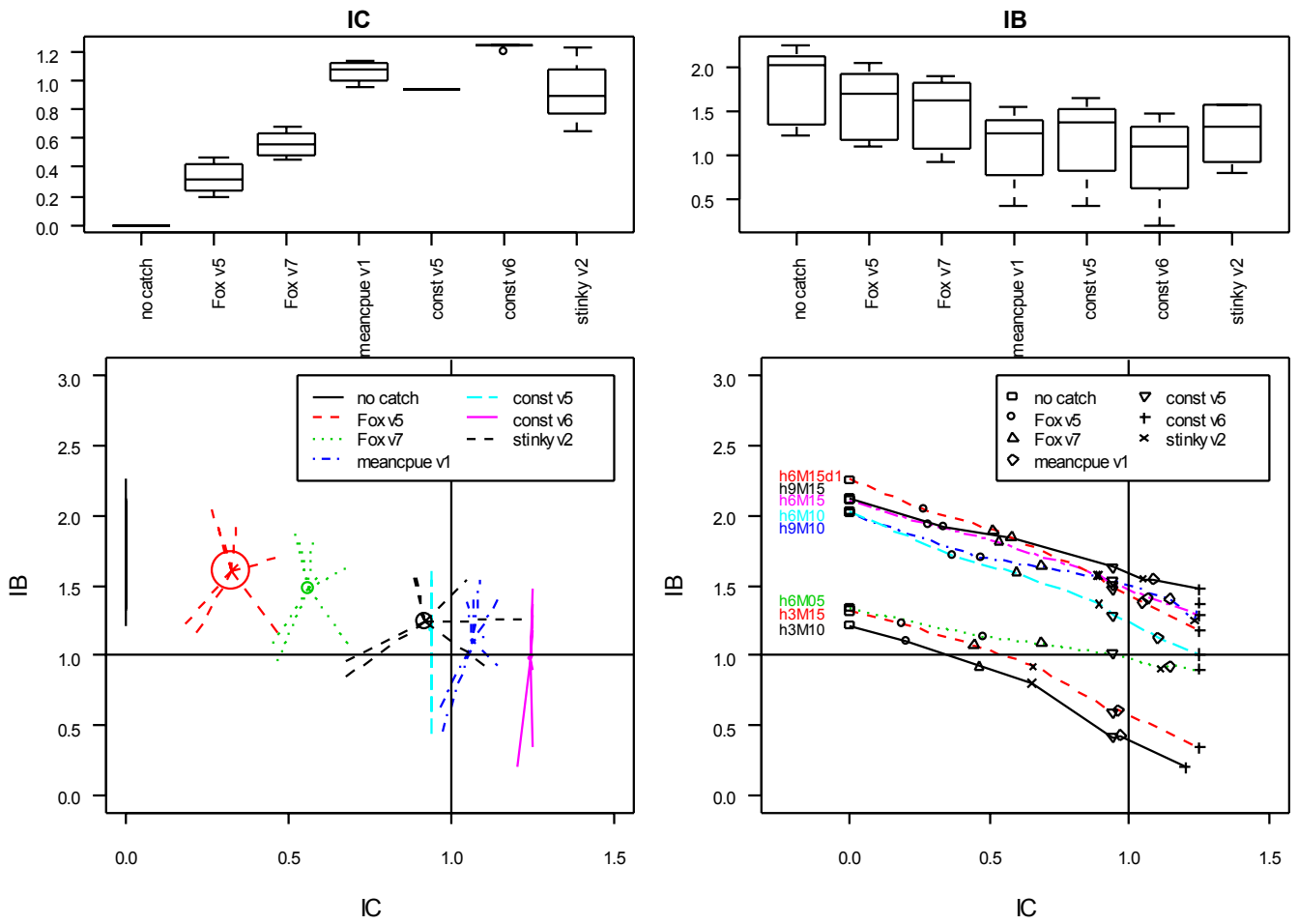


Figure 4b.

