

## A stroll through the automated plots for ASAP3

(asap3\_plots.r April 2014)

In order to facilitate model building and model selection in a consistent and informed manner, a standardized approach to model diagnostics has been taken. The output from any particular model configuration is used to generate the same series of plots to evaluate and compare the fit with earlier model configurations. These plots are aggregated into a series of **pdf** files (Data, Diagnostics, Results, Reference Points, and MCMC chain diagnostics and distributions) and provide an electronic record of the various models attempted. What follows is a brief narration describing the plots that are generated for each pdf. Naming conventions and plot interpretation will be emphasized. Plots are referred to by descriptive type, rather than page number, as the number of indices or fleets modeled could change the total page length. Therefore, the numbering below is relative to the distinct plot types and not the number of plots. In addition to **pdf** files, individual **png** files and several **csv** spreadsheets are generated.

### Diagnostic Plots

1. Overall model fit: The final objective function value and number of parameters are given for the file name identified at the bottom of the page. Contributions to the objective function by individual components are represented by horizontal bar plot.
2. Number of high correlations: The user defines a value for what constitutes a “high” correlation value. The number of estimated parameters that have a correlation with another parameter whose absolute value is above this “high” value are indicated. **A .csv file is generated that lists the parameter pair with the high correlation as well as the value.**
3. Summary of standardized residuals: The component and the number of residuals calculated is compared to the root mean squared error (RMSE). An “ideal” fit would have all  $RMSE=1$ .
4. RMSE for Indices and Catch: A separate plot is generated for indices and catch showing the RMSE versus the number of residuals. This demonstrates the asymptotic “ideal” value of 1, and how that criterion weakens the lower the number of residuals.
5. Fit to Catch/Landings/Discards (4-panel plot): Depending on fleet definition, these plots show the fit to the total catch/landings/discards. A lognormal error distribution is assumed. Observation and lognormal scale fits are shown on the top row. The top right has error bars on the observed values (open circles) based on the user input CV. The bottom row shows the time trend in residuals and the standardized residuals. Ideally, one would not see runs in residuals by year, and the distribution of log-scale standardized residuals should follow a normal distribution.
6. Annual age comp fits (Catch): For each fleet and each year, a small plot of observed (open circle) and predicted (solid line) proportion at age is plotted.
7. Bubble plots for Catch age composition residuals: Residuals are calculated as (Observed-Predicted) for each age and year. The size of the bubble reflects the size of the residual, and the overall scale is indicated by the legend in the top right of the plot. Text between the two legends indicates the maximum value of all residuals in the plot.

8. Overall Effective Sample Size (Catch): Input effective sample size (open circles) are plotted against the calculated effective sample size given the ASAP model input (the user specifies the input effective sample size). One interpretation of this plot is that the user-input ESS should more or less intersect the calculated effective sample size (solid line).
9. Predicted Mean Age (Catch): These plots are referred to colloquially as “Francis plots” as they were motivated by a recent paper by Chris Francis (Francis, R.I.C.C. 2011. Data weighting in statistical fisheries assessment models. *Can. J. Fish. Aquat. Sci.* 68:1124-1138). The top plot compares observed mean age, with 95% CI, with predicted mean age (solid blue line)—ideally the CI will overlap the predicted mean age line. The effective sample size (ESS) is either given in the title (if it is constant for all years) or it is indicated by green bars for each year (secondary y-axis gives scale). The bottom plot compares residuals (standard deviation of normal residuals, SDNR, and RMSE)—ideally, both numbers should be close to one. These plots are a recent addition to the ASAP diagnostics, and we are still evaluating their information content.
10. QQ-plots for observed vs predicted mean age (catch): This plot shows how closely the standardized residuals align with a 1:1 line, assuming that age composition data is multinomially distributed. In a perfect fit, the red and blue lines would lie on top of each other.
11. Fit to Indices (4-panel plot): Same interpretation as for (3) above. A lognormal error distribution is assumed for the aggregate survey. Observation and lognormal scale fits are shown on the top row. Error bars on observed index values (open circle) on the top left plot are based on user-input CVs. The bottom row shows the time trend in residuals and the standardized residuals. Ideally, one would not see runs in residuals by year, and the distribution of log-scale standardized residuals should follow a normal distribution. There are as many plots as there are indices in the model.
12. Bubble plots for Index age composition residuals: Residuals are calculated as (Observed-Predicted) for each age and year. The size of the bubble reflects the size of the residual, and the overall scale is indicated by the legend in the top right of the plot. The value of the maximum residual is printed at the top of the plot, between the two legends.
13. Overall Effective Sample Size (Indices): Input effective sample size (open circles) are plotted against the calculated effective sample size given the ASAP model input (the user specifies the input effective sample size). One interpretation of this plot is that the user-input ESS should more or less intersect the calculated effective sample size (solid line).
14. Predicted Mean Age (Indices): These plots are referred to colloquially as “Francis plots” as they were motivated by a recent paper by Chris Francis (Francis, R.I.C.C. 2011. Data weighting in statistical fisheries assessment models. *Can. J. Fish. Aquat. Sci.* 68:1124-1138). The top plot compares observed mean age, with 95% CI, with predicted mean age (solid blue line)—ideally the CI will overlap the predicted mean age line. The effective sample size (ESS) is either given in the title (if it is constant for all years) or it is indicated by green bars for each year (secondary y-axis gives scale). The bottom plot compares residuals (standard deviation of normal residuals, SDNR, and RMSE)—ideally,

both numbers should be close to one. These plots are a recent addition to the ASAP diagnostics, and we are still evaluating their information content.

15. QQ-plots for observed vs predicted mean age (indices): This plots shows how closely the standardized residuals align with a 1:1 line, assuming that age composition data is multinomially distributed. In a perfect fit, the red and blue lines would lie on top of each other.

### **Results pdf**

16. Fleet selectivities: You will see one plot for each fleet. For a given fleet, selectivity blocks are plotted as separate colored lines and the first year that a selectivity operates is indicated in the legend. For example, a legend entry of "1982" indicates that the first selectivity block begins in 1982, and "1991" indicates that the second selectivity block for that fleet begins in 1991.
17. Fmult by fleet: This plot is the time trend in fully selected F for each fleet for all years.
18. Index selectivities: For fishery independent indices, selectivity is assumed to be constant throughout the whole time series. Therefore, there is one selectivity curve for each index in the model.
19. Catch curve estimates of Z (Fleets and Indices): Catch curves are plotted by year class beginning with either the peak age (age with largest  $\ln(\text{Catch\_at\_age})$ ) or for an age specified by the user in the R script. The top panel shows the descending limb of the catch curve for each cohort, while the bottom panel shows the corresponding estimate of Z (slope of linear regression) with 80% confidence interval. A separate plot is made for each fleet and for each index. Note: if there are only 2 ages available for the descending limb of the catch curve, then no error bars can be calculated (you will only see solid circle in bottom plot). This will produce a warning at the end of the R script but all plots and pdfs will still be produced.
20. Correspondence-correlation plots (Fleets and Indices): Log-scale catch at age is plotted for consecutive ages in a cohort in the top half of these plots for both observed and predicted values. For example, the  $\log(\text{observed catch at age 2})$  would be plotted against the  $\log(\text{observed catch at age 3})$  one year later in order to determine if cohorts track consistently over the time series. The red line is the fitted line from a simple linear regression and the two blue lines are confidence intervals. The age is identified along the diagonal of the plot layout. Below the diagonal, the correlation is printed.
21. Index q estimates: For most models, q is assumed to be constant, and a bar plot for each q estimate (with error bar) is produced. ASAP has the capability to estimate a random walk in q such that annual estimates of q are possible. In this case,  $q_{i,\text{year}}$  is plotted as a line with year on the x-axis. If there are a lot of indices with a random walk for q, the plots are broken up so that there are only 5 lines per plot.
22. SSB and F time trends (2 panel plot): The top plot shows the temporal trend in SSB and the bottom plot shows the temporal trend in both "F.report" and "F.Full." F.report is the total F from all fleets, averaged over a user-specified age range. F.Full is the fully selected total fishing mortality, across all fleets. If the user-specified age range for F.report includes ages that are not fully selected, then F.report will be lower than F.Full.

23. Biomass Comparisons: 3 Biomass series are compared based on the value at the start of the year (January 1). Total Biomass is the sum of all numbers at age times weight at age. SSB is the sum of all mature numbers at age times spawning biomass at age. Exploitable Biomass is the sum of numbers at age times the combined selectivity (across all fleets) times the catch weight at age.
24. Spawning Biomass at age (absolute scale): Standard bar plot of spawning biomass of mature ages over time.
25. Spawning Biomass at age (proportion): Same plot as above, but expressed as a proportion each year.
26. Estimated Numbers at age, NAA (absolute scale): Standard bar plot of numbers at age over time.
27. Estimate NAA (proportion): Same plot as above, but expressed as a proportion each year.
28. Recruitment deviations (2 panel plot): The top plot shows the deterministic predicted recruitment (solid line) as well as the predicted recruitment with deviations (solid lines with filled circles) over time. The bottom plot shows the time trend of recruitment deviations in terms of the log-scale residuals.
29. Recruits vs SSB (symbol is 2 digit year): This is a scatter plot of recruitment versus SSB. The symbol is a two-digit indicator of the model year. The most recent model estimate is highlighted by a solid orange circle.
30. Recruits vs SSB (predicted and replacement line): This is the same scatter plot, but the scale is expanded to allow plotting of the predicted line (solid red).
31. SSB versus R (barplot with superimposed line plot): A standard plot included in our assessment summary reports. Recruitment is given by vertical bars, while the SSB that produced the recruitments is plotted in a solid line. Year is on the x-axis and refers to the spawning year. If recruits are age 1, then the year would also refer to the cohort year.
32. CV in model estimates by year: Hessian based CVs for model estimates of SSB, F\_report, and R by year.
33. Retrospective plots: If you performed a retrospective run, then you will see plots of the absolute and relative retrospective pattern. Plots of F (F.report), SSB, and Recruitment are grouped together; January 1 Biomass, Exploitable Biomass, and Total Stock N are grouped on a second plot; and finally, age-specific numbers at age are plotted. The relative retrospective patterns are plotted in the right column, and the average retrospective value (Mohn's rho) is printed in the top left inside the plot. **A .csv file is generated that gives this same information.**

#### **Reference Points pdf**

34. YPR-SPR: For F values in [0, 0.01, 0.02,...,2.0], the value of YPR and SPR is plotted. The F value is a fully selected F. The user-specifies the number of years over which selectivity, maturity, mortality, and weights are averaged.
35. A Table of YPR-SPR values is printed to the pdf. **A .csv file is generated that gives this same information.**

36. SPR Target Reference Points: For SPR values in [0.20, 0.25, 0.30, ...,0.80], the value of F that produces that SPR and the corresponding YPR are plotted.
37. A Table of SPR target values is printed to the pdf. **A .csv file is generated that gives this same information.**
38. Selectivity and Maturity at age: For a user-defined number of years to average over, the resulting values of selectivity at age (over all fleets) and maturity at age are plotted.
39. Expected Spawning and %SPR: For a user-defined number of years to average over, both the expected number of spawning opportunities and the %SPR are plotted for F values in [0.01, 0.02, ..., 2.0]. Expected spawning is per recruit and is calculated as  $\sum_{ages} [mat_{age} \prod_{i=1}^{age-1} \exp(-Z_{age} \Delta t_{spawn})]$ .
40. A Table of expected spawning and %SPR values is printed to the pdf. **A .csv file is generated that gives this same information.**
41. Annual F(%SPR): For %SPR values of 20%, 30%, 40%, and 50%, the value of fully selected F is plotted each year. SPR is calculated using the annual weight, maturity, and selectivity values from the model.
42. Annual YPR(%SPR): For each F on the previous plot, the corresponding value of YPR is plotted.
43. Histograms of Annual F(%SPR): All of the annual values are displayed as a histogram.
44. Histograms of Annual YPR(%SPR): All of the annual values are displayed as a histogram. ***IF you estimated a stock-recruit curve you will see 8 additional plots***
45. Histogram of Annual MSY Reference Points: Full F, MSY,  $SSB_{MSY}$ , and  $R_{MSY}$ . These annual MSY values come from calculating the replacement line from annual weight, maturity, and mortality values as well as the annual selectivity from the model.
46. Second Histogram of Annual MSY Reference Points: steepness,  $SPR_{MSY}$ , unexploited SSB ( $SSB_0$ ) and unexploited recruitment ( $R_0$ ). These annual MSY values come from calculating the replacement line from annual weight, maturity, and mortality values as well as the annual selectivity from the model.
47. Steepness vs  $SPR_{MSY}$ : The annual steepness and  $SPR_{MSY}$  values are plotted as filled circles. The value for the first and the last year are indicated in the legend.
48. Annual MSY Reference Points: The annual values of  $F_{MSY}$ , MSY,  $\%SPR_{MSY}$ ,  $SSB_{MSY}$ , and  $R_{MSY}$  are plotted individually for each year. **A .csv file is generated that gives this same information.**

**MCMC pdf (only produced if you performed mcmc)**

49. Trace for SSB: A plot of the chain for SSB in the first and last year is plotted. The “trace” (the plotted line) should not show any pattern. If it shows a pattern at the beginning, you can use a longer “burn-in” period by dropping the iterations that show that pattern.
50. Trace for F.report: Same as for SSB.
51. Autocorrelation plot for SSB in first and last year: This plot looks at the correlation between successive values of the chain at lags of 0 to 10. The correlation should be inside the dashed horizontal red lines. If not, then you should thin at a higher rate to reduce correlation.
52. Autocorrelation plot for F.report in first and last year: Same as for SSB.

53. Distribution of SSB in first and last year (same plot for Freport, Full.F and January 1 Biomass) : The distribution is plotted and for comparison the point estimate from the model is indicated by a dashed red vertical line.
54. Probability interval plot: Values of the 5<sup>th</sup>, 50<sup>th</sup>, and 95<sup>th</sup> percentile are plotted versus year for SSB, F.report, Full.F, and January 1 Biomass. For comparison, the annual point estimate from the model is plotted as a green line with filled green triangles.

### **Data pdf**

55. Observed Catch (absolute scale): This plots the total magnitude of catch. If there is more than one fleet, the catch of each fleet in a given year is distinguished by a different color.
56. Observed Catch (proportion): Proportion of total catch in a given year by fleet. If there is only one fleet, this plot is not very interesting.
57. Age composition by fleet: Bubble plots of the catch age comp are plotted for each year that those data are available.
58. Indices: All indices are rescaled to their mean and plotted together on both arithmetic and logarithmic scales.
59. Age composition by index: Bubble plots of the index age comp are plotted for each year that those data are available.
60. WAA matrices: Each WAA matrix is plotted versus year, with a separate line for the weight of each age. A dashed line indicates the mean for each weight at age. The legend identifies the data series that are associated with a given WAA matrix.
61. M versus year: If M is time-invariant, a single line is plotted for each age. If M is time-varying, the mean value is plotted. Colored open circles are jittered to show the spread of annual values about the mean. The first and last year are indicated in the legend.
62. Maturity versus year: If maturity is time-invariant, a single line is plotted for each age. If maturity is time-varying, the mean value is plotted. Colored open circles are jittered to show the spread of annual values about the mean. The first and last year are indicated in the legend.