



NOAA
FISHERIES

Diagnostics in Stock Synthesis

Ian Taylor

CAPAM Diagnostics Workshop
2 February 2022



Outline:

- Overview of diagnostics for SS3
- A few diagnostics of interest that I think we should use more often
- Future directions

Overview: four places to get diagnostics for SS3

- Diagnostics in Stock Synthesis output
- Diagnostics from core r4ss functions
- Additional r4ss functions for diagnostics that aren't called by `SS_plots()`
- Additional R packages

Overview: diagnostics in SS3 output

- Messages during run time and warning files provide initial check for problems
- Improved messages about parameters on bounds in ADMB on the way thanks to Cole Monnahan github.com/admb-project/admb/issues/221

```
Estimating row 44 out of 44 for hessian
5 468 -log(L): 402.017 Spbio: 2524.55 1987.88
5 469 -log(L): 402.017 Spbio: 2524.55 1987.88
5 470 -log(L): 402.017 Spbio: 2524.55 1987.88
5 471 -log(L): 402.017 Spbio: 2524.55 1987.88
do benchmark and forecast if requested in sdphase
  got Fspr 0.107403 0.5
  got_Btgt 0.0722269 0.4
  got Fmsy 0.0763359 702.688
finished benchmark, forecast, and sdreporting
5 472 -log(L): 402.017 Spbio: 2524.55 1987.88

In final section
Finish time: Tue Feb 01 12:52:04 2022
Elapsed time: 0 hours, 3 minutes, 37 seconds.
Iterations: 472 -log(L): 402.017
Final gradient: 7.17341e-05

finished COVAR.SSO
finished forecast
finished StdDev quantities
write mcmc headers
finished posteriors
finished SS_summary.sso
finished SS_summary.sso
writing big output now
finished report.sso
data.ss_new with N replicates: 1 finished
Write new starter file
Write new forecast file
Write new control file
dynamic Bzero: finished
SPR_profile: finished
Global_MSX: finished
```

!! Run has completed !!

See warning.sso for N warnings: 5

Overview: diagnostics from core r4ss functions

- The `SS_output()` function reads model output from text files into a list in R
 - Some diagnostics messages are reported while it runs, but these are buried in a sea of too much information and routinely ignored
 - Returns tables with information such as related to parameters on bounds, data weighting, and tuning variances
- The `SS_plots()` function creates a default set of plots
 - Results displayed in HTML format
 - Additional tables of diagnostics slowly getting moved into the HTML view where they are easier to access
 - Figures probably too numerous but it's easier to add than take away
 - Can be posted to web via github pages (IATTC did similar in 2016, now easier)

Overview: diagnostics from additional functions

- Some diagnostic plots require calling independent `r4ss` functions
- These either haven't been integrated into set called by `SS_plots()` or involve comparing across multiple models
- Examples:
 - jitters
 - likelihood profiles
 - retrospectives
 - Mohn's rho

Side note: multiple versions of Mohn's rho

`r4ss::SSmohnsrho()` function thanks to
Chantel Wetzel, Carey McGilliard, and Brian Langseth

Output

SSB	-2.02
Rec	-0.55
Bratio	-1.73
F	3.68
WoodHole_SSB.all	-1.98
WoodHole_Rec.all	-0.60
WoodHole_Bratio.all	-1.68
WoodHole_F.all	3.50
AFSC_Hurtado_SSB	-0.39
AFSC_Hurtado_Rec	-0.13
AFSC_Hurtado_F	0.68
AFSC_Hurtado_Bratio	-0.33

Terminal year
Sums of vertical lines

All years (but can set
a different year
using `startyr`)
Sum of averages

Terminal year
Average of vertical lines

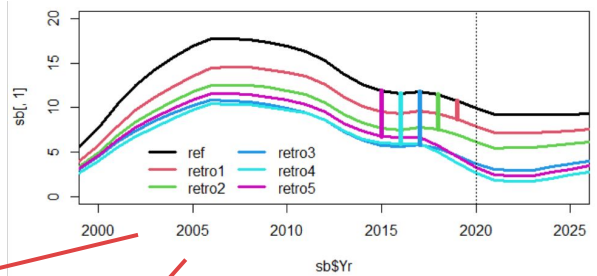


Table of differences from reference

Yr	Ref	Peels (summed)				
		1	2	3	4	5
2015	0	averaged	averaged	averaged	averaged	averaged
2016	0			averaged	averaged	averaged
2017	0					
2018	0					
2019	0					
2020	0					

Values are for 2021 Oregon quillback rockfish,
illustration by Brian Langseth

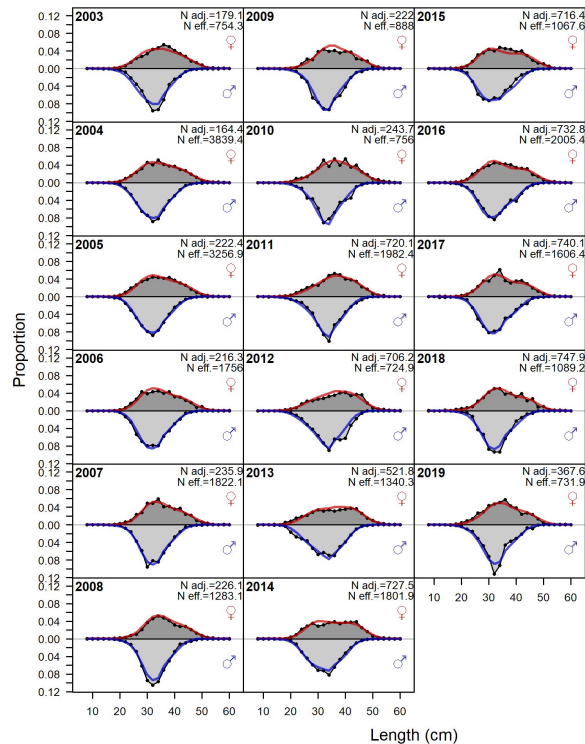
Overview: diagnostics from additional R packages

- ss3diags, see next talk by Megumi Oshima
 - see github.com/jabbamodel/ss3diags
 - generalized to work with any SS3 model
- Pacific Hake stock assessment
 - see github.com/pacific-hake/hake-assessment
 - includes Bayesian analogs to many standard r4ss plots of MLE output
 - meets specific needs of annual assessment
 - longstanding interest in generalizing Bayesian diagnostics (see github.com/r4ss/r4ss/issues/11) but there are still few fully Bayesian SS3 models
- Empirical selectivity
 - hopefully will get integrated into r4ss eventually
- Numerous agencies have their own R scripts or packages



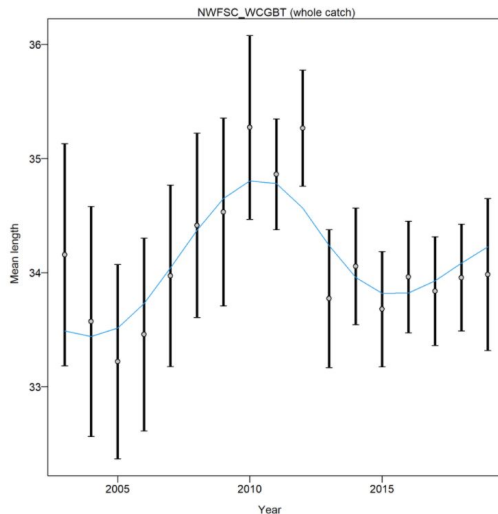
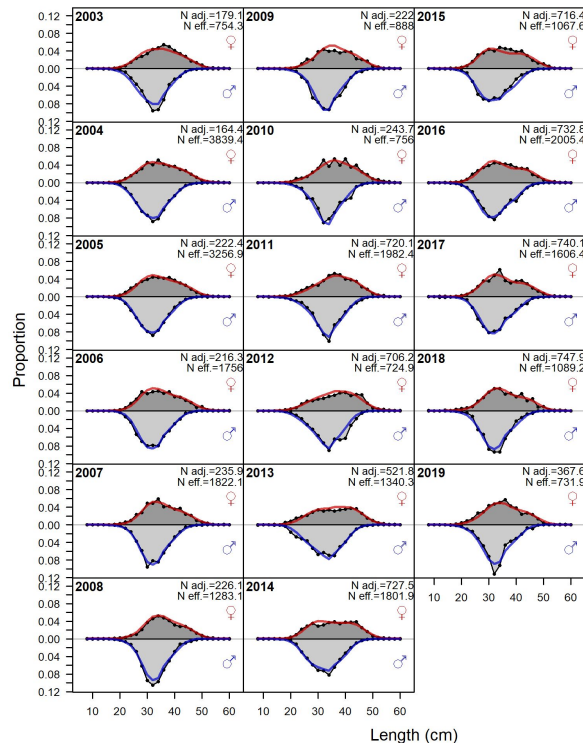
Diagnostics we should use more often

Time series of mean length or age in comp data



- Fit to these length comps is great
- Not clear how these data influence the model other than informing selectivity (and estimation of R_0 as shown by likelihood profiles)

Time series of mean length or age in comp data



Mean length for NWFSC_WCGBT with 95% confidence intervals based on current samples sizes.

Francis data weighting method TA1.8: thinner intervals (with capped ends) show result of further adjusting sample sizes based on suggested multiplier (with 95% interval) for len data from NWFSC_WCGBT:
1.0022 (0.5392-3.6767)

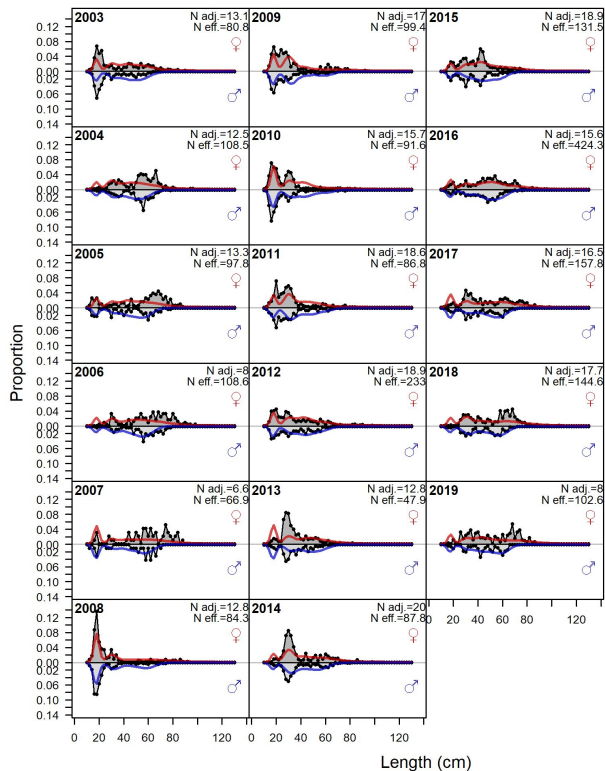
For more info, see

Francis, R.I.C.C. (2011). Data weighting in statistical fisheries stock assessment models. *Can. J. Fish. Aquat. Sci.* 68: 1124-1138.
<https://doi.org/10.1139/f2011-025>

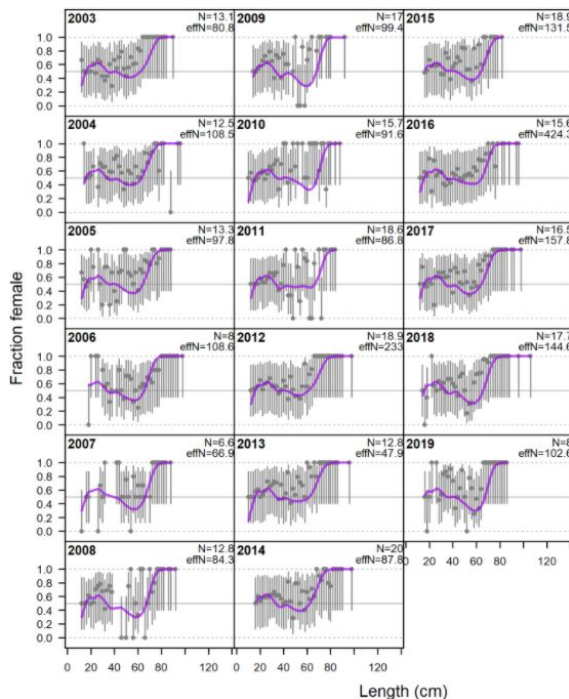
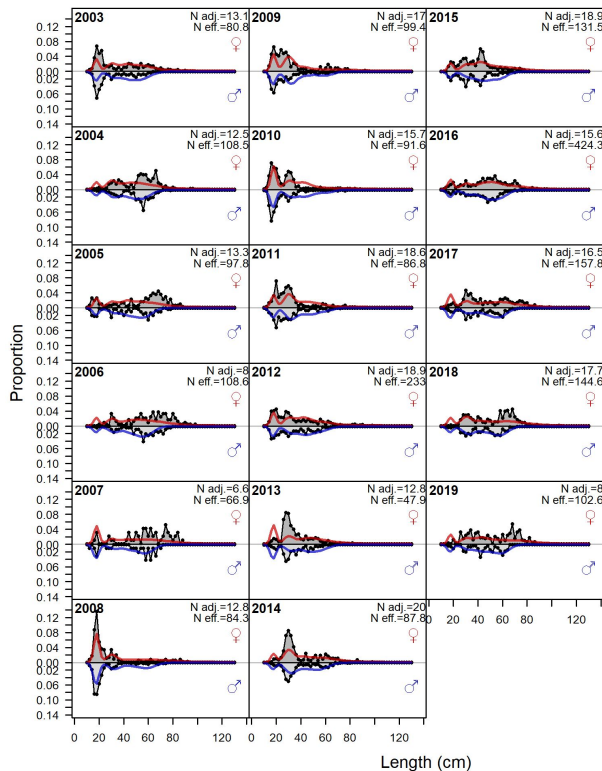
- Figure from Francis weighting is valuable to see what information comp data have about recruitment or population scale, regardless of chosen weighting method
- Would be nice to have a option to compare fit of multiple models to these time series

Diagnostic in r4ss thanks to
Chris Francis and André Punt

Sex ratios in composition data



Sex ratios in composition data

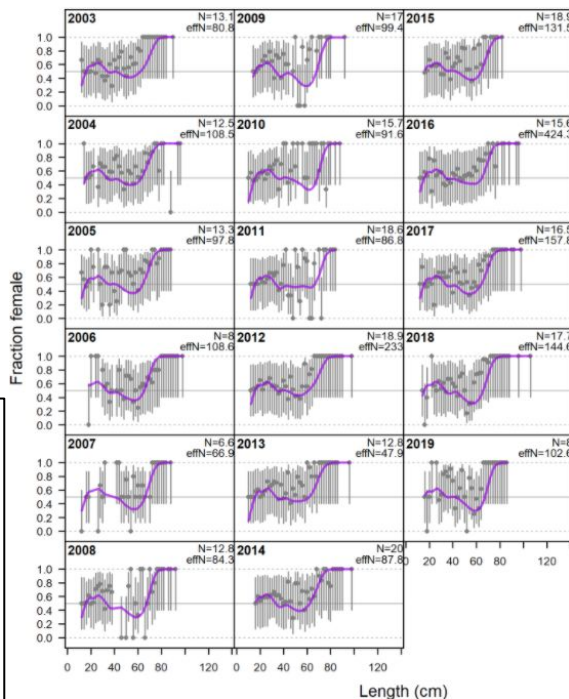
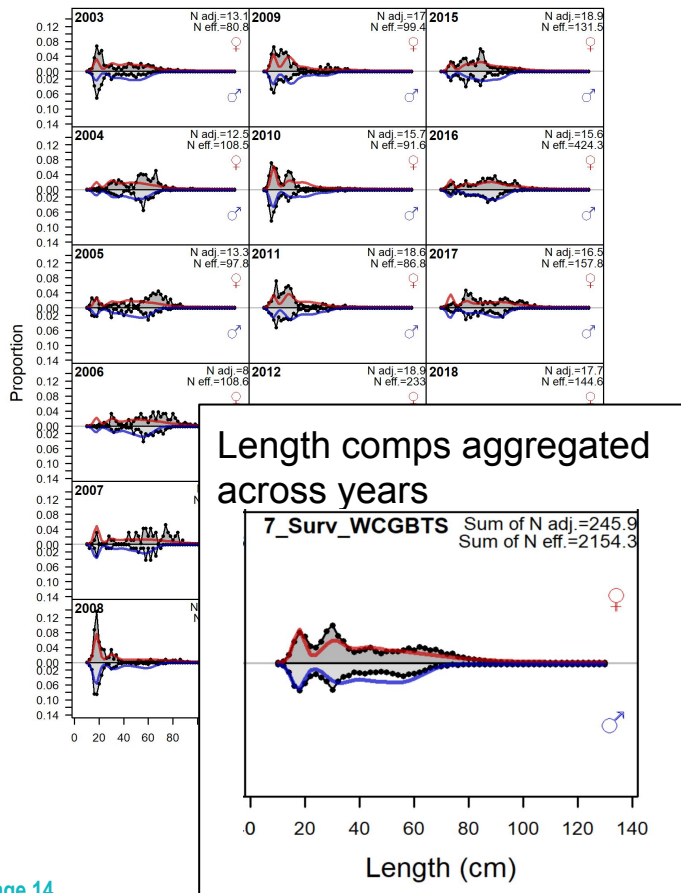


- Sex ratio plots provide a different view of the comp data
- Helps judge reasonableness of estimates of dimorphic growth
- Helps identify when sex-specific selectivity is needed

Sex ratios for length comps, whole catch, 7_Surv_WCGBTS.
Observed sex ratios (points) with 75% intervals (vertical lines) calculated as a [Jeffreys interval](#) based on the adjusted input sample size. The model expectation is shown in the purple line.
file: [sexratio_len_flt7mkt0.png](#).

Diagnostic in r4ss thanks to
Cole Monnahan and Ian Stewart

Sex ratios in composition data

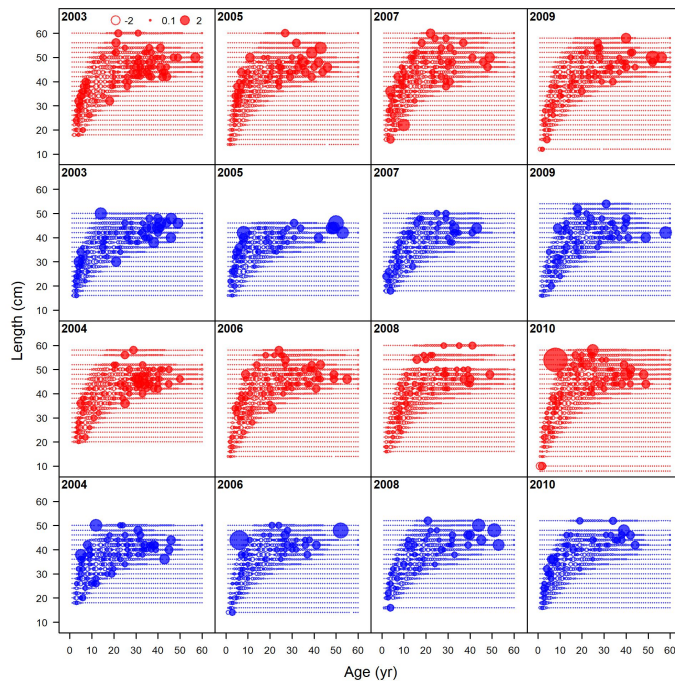
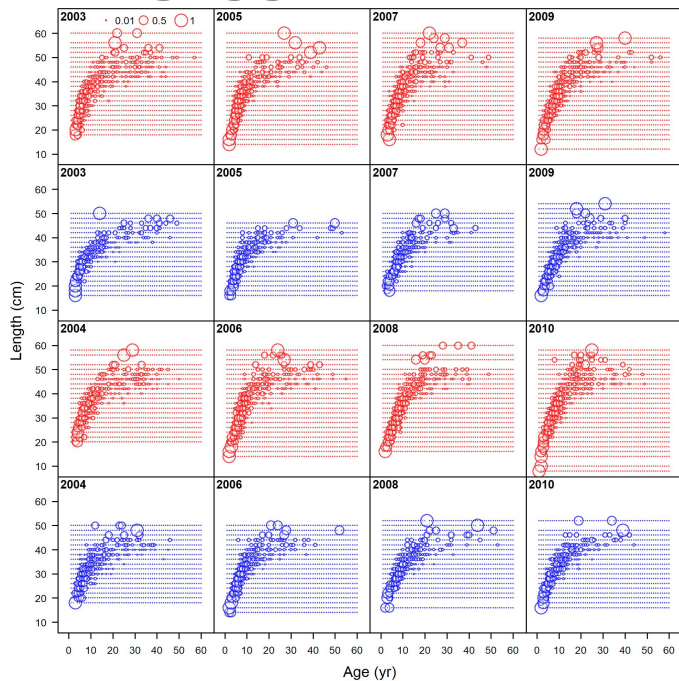


Sex ratios for length comps, whole catch, 7_Surv
Observed sex ratios (points) with 75% intervals
calculated as a [Jeffreys interval](#) based on the ad
size. The model expectation is shown in the pur
file: [sexratio_len_flt7mkt0.png](#).

- Sex ratio plots provide a different view of the comp data
- Helps judge reasonableness of estimates of dimorphic growth
- Helps identify when sex-specific selectivity is needed
- Sex ratio of aggregated comps would be useful

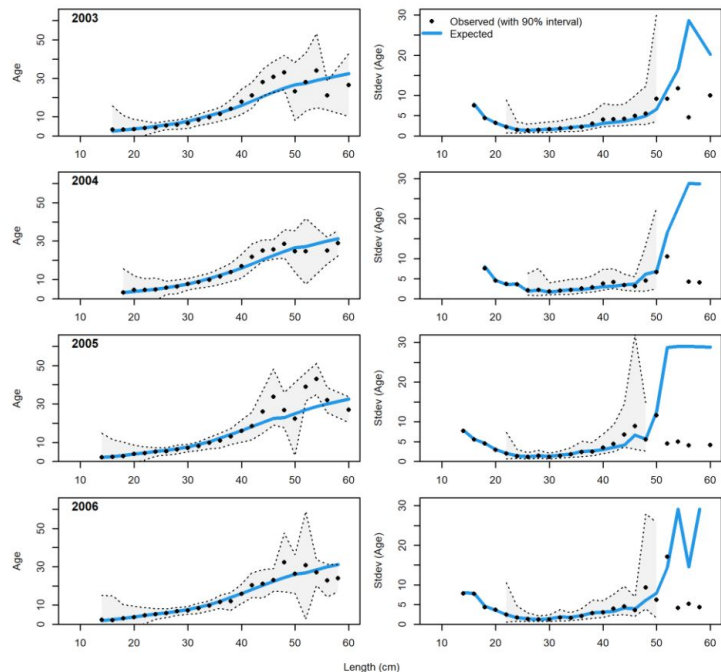
Conditional age-at-length fits

- Estimating growth within a model using CAAL is the best way to get accurate estimates of growth and account for uncertainty in those estimates
- Judging goodness of fit is hard



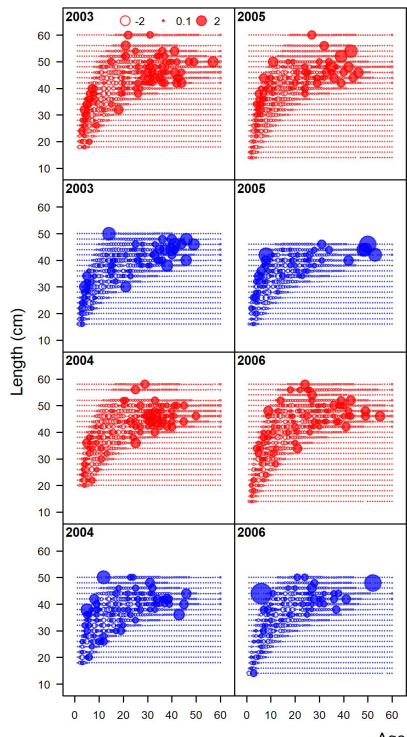
Example is U.S. West Coast Groundfish Bottom Trawl Survey from 2021 Dover sole assessment (Chantel Wetzel and Aaron Berger)

Conditional age-at-length fits



Conditional AAL plot, whole catch, NWFSC_WCGBT (plot 1 of 5)
 These plots show mean age and std. dev. in conditional A@L.
 Left plots are mean A@L by size-class (obs. and pred.) with 90% CIs based on adding 1.64 SE of mean to the data.
 Right plots in each pair are SE of mean A@L (obs. and pred.) with 90% CIs based on the chi-square distribution.
 file: [comp_condAALfit Andre_plotsflt6mkt0_page1.png](#)

Example is U.S. West Coast Groundfish Bottom Trawl Survey from 2021 Dover sole assessment (Chantel Wetzel and Aaron Berger)



- Summarizing in terms of expected and observed mean age and std. dev. provides a useful objective diagnostic
- Diagnostic remains confusing

Diagnostic in r4ss thanks to
 André Punt

Conditional age-at-length fits

We're not used to age dependent on length. Should the plot be flipped?

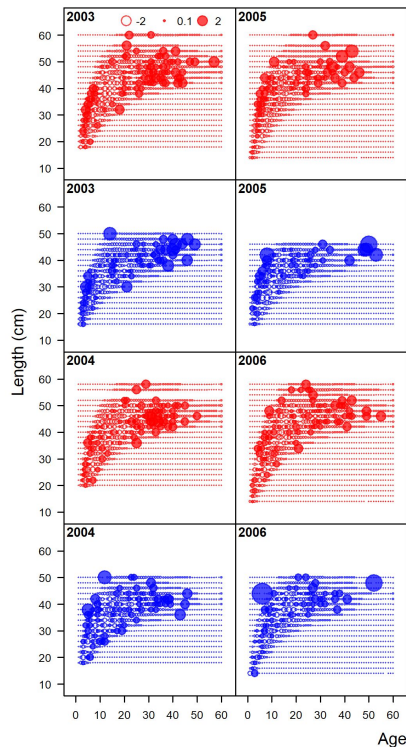
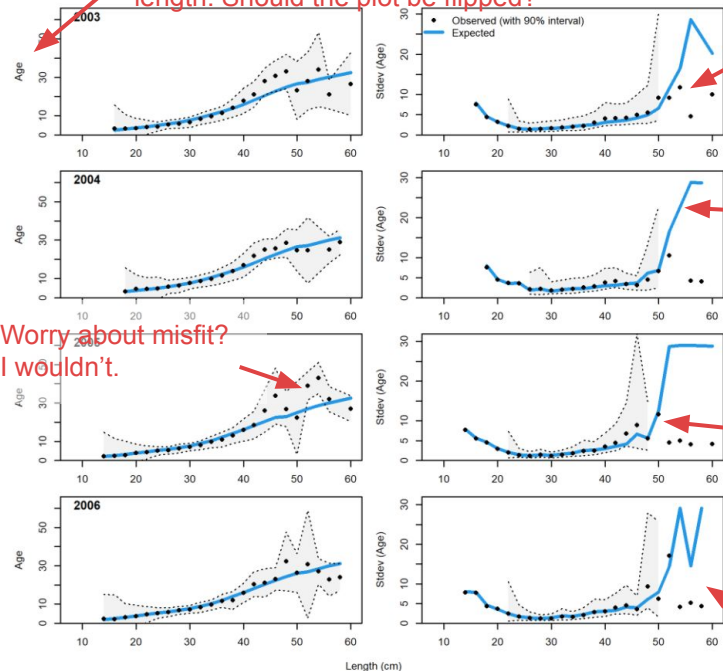
Uncertainty intervals around observed std. dev. unavailable for length bins with adjusted input sample size < 1.0.

Overestimation of std. dev. for larger length bins is common. Can be caused by unmodeled variation over space or time as well as ageing error.

Plot aggregates across sexes so expected and observed values jump around as ratio of sample sizes vary among sexes. Should these plots be sex-specific?

Don't know what's happening with zig-zag here (obs. > 50cm are all females in 2006, so expected value should be smooth).

Worry about misfit? I wouldn't.



Conditional AAL plot, whole catch, NWFSC_WCGBT (plot 1 of 5)
 These plots show mean age and std. dev. in conditional A@L.
 Left plots are mean A@L by size-class (obs. and pred.) with 90%
 CIs based on adding 1.64 SE of mean to the data.
 Right plots in each pair are SE of mean A@L (obs. and pred.)
 with 90% CIs based on the chi-square distribution.
 file: [comp_condAALfit Andre_plotsflt6mkt0_page1.png](#)

Diagnostic in r4ss thanks to
 Andre Punt

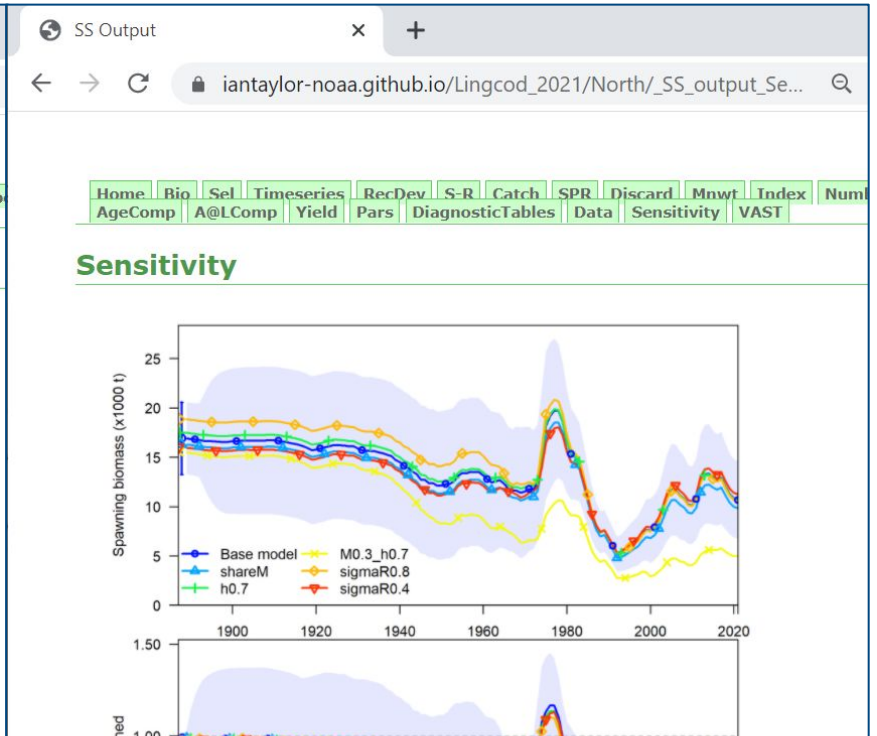
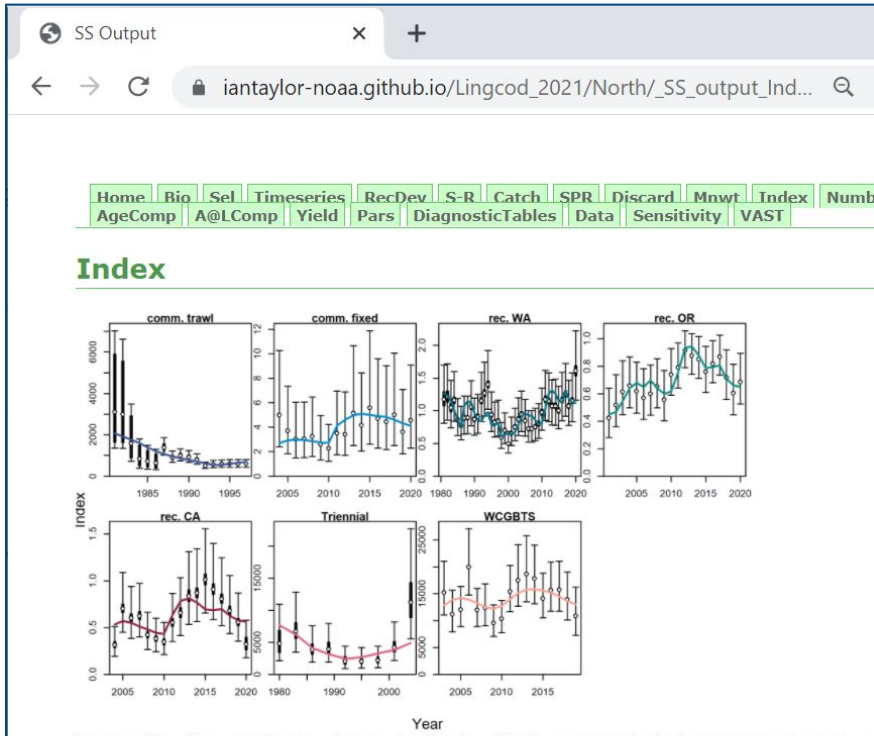
Example is U.S. West Coast Groundfish Bottom Trawl Survey from 2021 Dover sole assessment (Chantel Wetzel and Aaron Berger)

Future directions



Future: adding custom figs and posting to web


- Info available later this year
- See example at iantaylor-noaa.github.io/Lingcod_2021/



Future: wrappers for existing functions

Old approach: 488 line script with repeated calls to `SS_profile()`, `SSgetoutput()`, `SSsummarize()`, `SSplotProfile()`, and `PinerPlot()`

BigSkate_Doc / R / BigSkate_profiles.R

 iantaylor-NOAA update some R scripts


 History

 1 contributor

488 lines (429 sloc) | 20.8 KB

```
1  ### notes on running profiles and making associated plots
2  ### for 2019 Big Skate assessment
3
4  #stop("\n This file should not be sourced!") # note to stop Ian from accidental sourcing
5
6  # define directory on a specific computer
7  if(Sys.info()["user"] == "Ian.Taylor"){
8    dir.outer <- c('c:/SS/skates/models')
9  }
10
11  require(r4ss)
12  require(SSutils) # package with functions for copying SS input files
13  #devtools::install_github('r4ss/SSutils')
14
15  # load model output into R
16  # read base model from each area
17  #mod <- 'bigskate74_spawnbio_3.30.13.02'
18  #mod <- 'bigskate82_base_May13'
```

Lingcod_2021 / unfit / diags.R

 kellijohnson-NOAA add both model names to diagnostic script ✓

 History

 1 contributor

16 lines (13 sloc) | 424 Bytes

```
1  # Source this file using the command line via
2  # Rscript --vanilla .\unfit\diags.R 1
3  # where the 1 at the end is the args you want,
4  # so here args could be 1:3
5
6  args <- commandArgs(trailingOnly=TRUE)
7
8  setwd("c:/stockassessment/lingcod_2021")
9  library(devtools)
10 load_all()
11
12  run_investigatemodel(
13    "2021.n.023.001_fixWAreccatchhistory",
14    # "2021.s.018.001_fixTri3",
15    run = c("profile", "retro", "jitter")[as.numeric(args)]
16  )
```

Future: wrappers for existing functions

New approach:

tiny script

stock-specific function with settings
for a particular assessment

region/agency-specific function or R package to
meet local Terms of Reference
(e.g. <https://github.com/nwfsc-assess/nwfscDiag>)

functions in r4ss and other packages for
profiles, jitters, retrospectives, etc.

Lingcod_2021 / unfit / diags.R



kellijohnson-NOAA add both model names to diagnostic script ✓

History

1 contributor

16 lines (13 sloc) 424 Bytes

```
1 # Source this file using the command line via
2 # Rscript --vanilla .\unfit\diags.R 1
3 # where the 1 at the end is the args you want,
4 # so here args could be 1:3
5
6 args <- commandArgs(trailingOnly=TRUE)
7
8 setwd("c:/stockassessment/lingcod_2021")
9 library(devtools)
10 load_all()
11
12 run_investigatemodel(
13   "2021.n.023.001_fixWareccatchhistory",
14   # "2021.s.018.001_fixTri3",
15   run = c("profile", "retro", "jitter")[as.numeric(args)]
16 )
```

Future: wrappers for existing functions

- Drawback of wrapper functions is less control
 - individual functions allow customization
 - profile fig at right showing choice of alternative states of nature

```
SSplotProfile(...,
  add_cutoff = TRUE,
  cutoff_prob = 0.75)
```
 - Running one step gives opportunity to do things like dealing with non-converged likelihood profiles
- However, we don't have time for those details
- We can automate some of those fixes

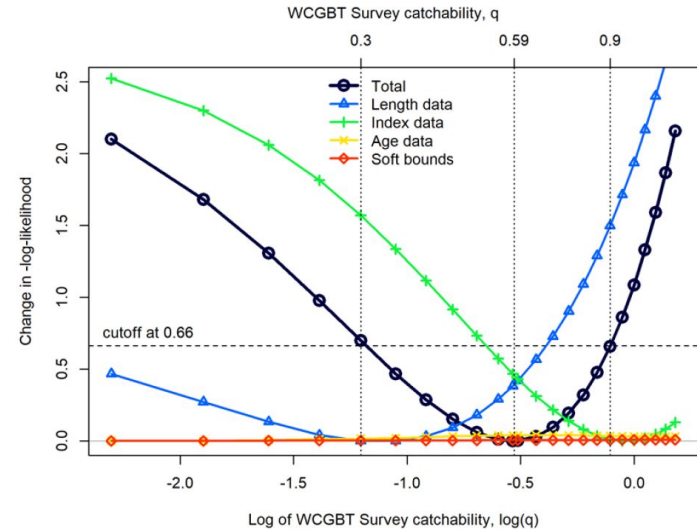


Figure 129. Likelihood profile over $\log(q)$ showing contributions of likelihood components. All values are represented as the change relative to the lowest negative log-likelihood for that component within the range of $\log(q)$ values shown in the figure.

Future: Adding new diagnostics to r4ss

- Ideal process:
 - create a new issue to discuss design and implementation
 - contribute code in a fork of the r4ss package
 - create a pull request
- Also acceptable:
 - emailed code, scribbles on a napkin, etc.
- See github.com/r4ss/r4ss#contributing-to-r4ss

and Fish. Res. paper
associated with 2019
CAPAM workshop



Beyond visualizing catch-at-age models: Lessons learned from the *r4ss* package about software to support stock assessments

Ian G. Taylor^{a,*}, Kathryn L. Doering^b, Kelli F. Johnson^a, Chantel R. Wetzel^a, Ian J. Stewart^c

^a Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd E, Seattle, WA, 98112, USA

^b Caelum Research Corporation Under Contract to Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd E, Seattle, WA, 98112, USA

^c International Pacific Halibut Commission, 2320 West Commodore Way, Suite 300, Seattle, WA, 98199, USA



Future: creating new independent packages

- See ss3diags github.com/jabbamodel/ss3diags
- Benefits:
 - more flexibility for authors,
 - less dependence on r4ss developers,
 - keeps packages smaller.
- Challenges:
 - maintaining compatibility may require more work,
 - less visibility than adding diagnostics to established tools,
 - ?

Future: FIMS



- NOAA folks, with support from lots of others, are developing a next-generation modeling system
- Expected to eventually replace Stock Synthesis and various legacy tools
- Model will be coded in TMB & C++
- Diagnostics will be coded in R
- Developing and refining diagnostics is harder than porting to new platform, so still reasonable to work on additional SS3 diagnostics

Fisheries Integrated Modeling System

fisheries.noaa.gov/national/population-assessments/fisheries-integrated-modeling-system

NOAA FISHERIES

Search NOAA Fisheries

Find A Species Fishing & Seafood Protecting Marine Life Environment Regions Resources & Services

POPULATION ASSESSMENTS

Fisheries Integrated Modeling System

An evolution of stock assessment models to provide better information for fishery managers to support sustainable fishery management.

National

The Fisheries Integrated Modeling System is a next-generation framework of stock assessment models, assisting fishery managers with the goal of achieving sustainable fisheries. This system, when completed in a few years, offers the NOAA Fisheries and global fisheries science communities an advanced set of stock assessment models. These tools can be used separately or in combination to incorporate ecosystem and socioeconomic data and models, as well as climate effects and other drivers within the marine environment, into stock assessment models.

For many years, NOAA Fisheries has relied on regionally developed [stock assessment models](#) to conduct stock assessments for fisheries management. Many of these models are hosted in the [Fisheries Integrated Toolbox](#). This new effort represents a system of tools that capitalizes on the expertise of NOAA Fisheries scientists and software development best practices in consultation with the broader fisheries management community. The system is designed to be modular, maintainable, and scalable. The system also enables stock assessment tools to leverage technological developments, such as:

[Send Us Your Feedback](#)

More Information

- > [Center for Advancement of Population Assessment Methodology](#)
- > [Ecosystem-Based Fisheries Management Road Map \(PDF, 50 pages\)](#)
- > [Fisheries New Zealand](#)
- > [NOAA Fisheries Integrated Toolbox](#)
- > [Population Assessments](#)
- > [Stock Assessment Improvement Plan \(PDF, 137 pages\)](#)

<https://www.fisheries.noaa.gov/national/>

[population-assessments/fisheries-integrated-modeling-system](https://www.fisheries.noaa.gov/national/population-assessments/fisheries-integrated-modeling-system)

Thank you!

Acknowledgements:

- CAPAM and IATTC
- Rick Methot & SS3 development team
- NOAA NWFSC assessment team
- The numerous contributors of SS3 diagnostics

Contact: Ian.Taylor@noaa.gov

