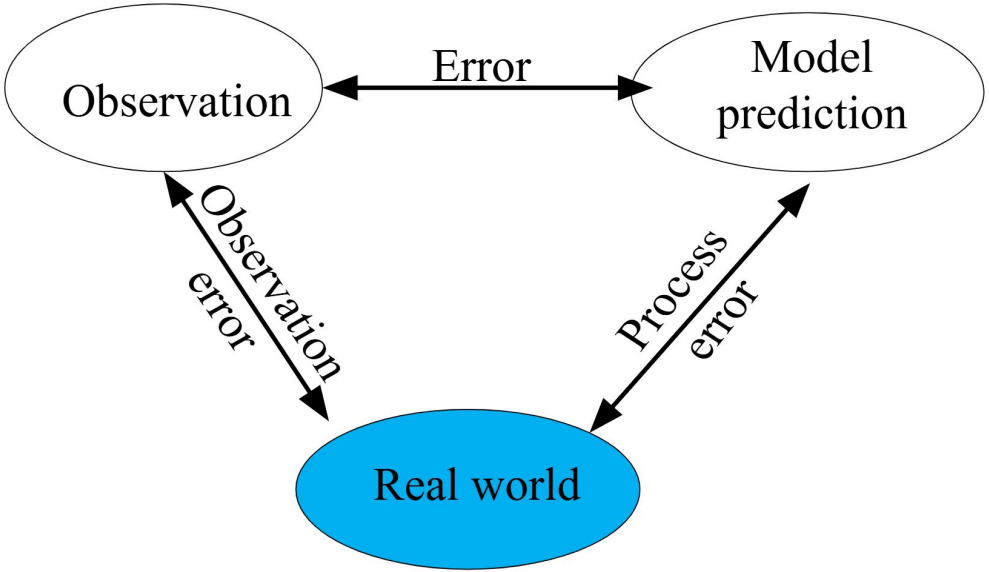
# Data weighting (a.k.a. model tuning)

FISH 576, Week 9

# What is “correct” data weighting?

The weight given to each observation should describe the likely size of its error (and thus its information content)




slide from Chris Francis

# What is “correct” data weighting?

- Integrated models combine likelihoods from disparate data sources
- Data sources are often provide somewhat conflicting information about the estimated parameters
- Without appropriate weighting of the likelihood components, it’s possible for a noisy data source to have more influence than a reliable one.
- If two data sources are in direct conflict it’s recommended to remove one or change model rather than just downweight due to lack of fit

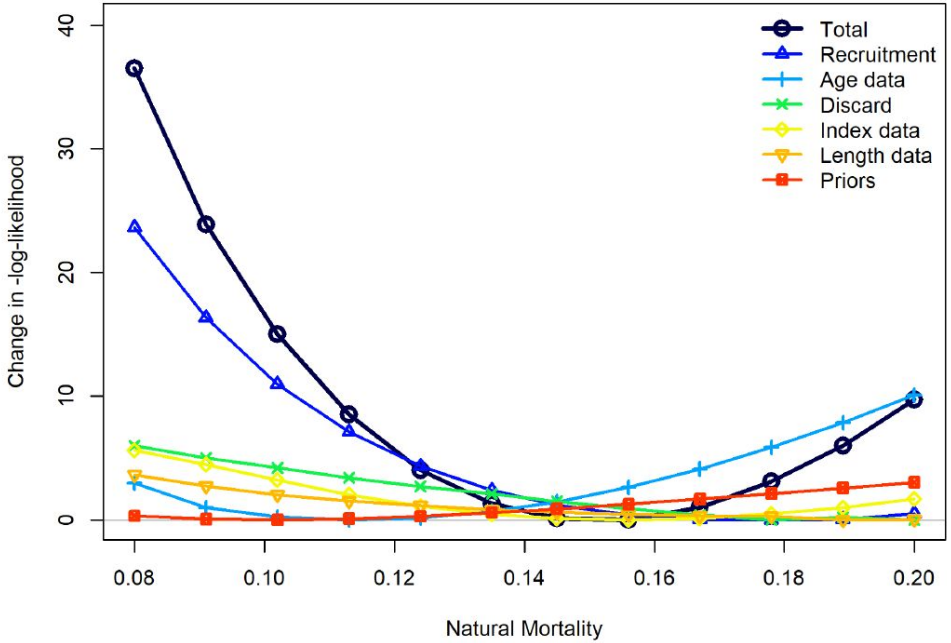


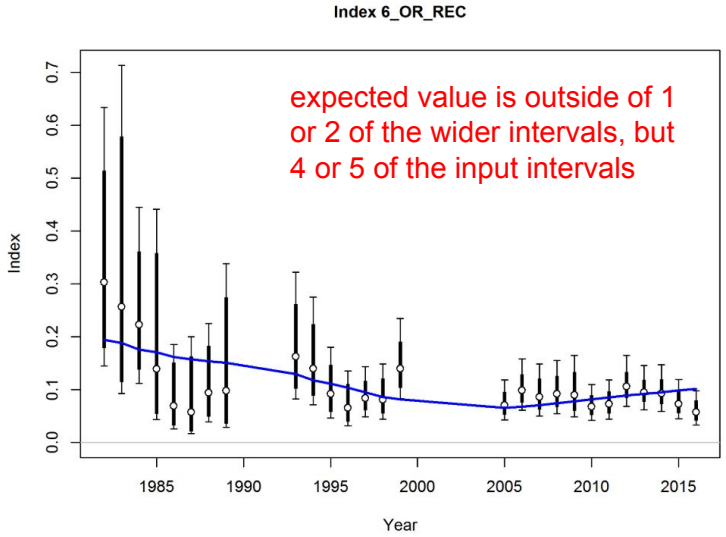

**Figure 70: Likelihood components in the likelihood profile for natural mortality (*M*). Note: male and female natural mortality are set to the same value.**

# What’s in this talk

- Focus is on data sources that are in Yelloweye and Widow Rockfish assessment models and weighting approaches used for PFMC assessments
- Update assessments should be weighted using the same methods is original assessment

# Tuning indices of abundance

- The influence of an index on a model is related to its uncertainty.
- Input uncertainty values come from index standardization outside of SS3
- SS3 has an option to estimate an “extraSD”* parameter which is added to the standard deviation of every index value
- The parameter ensures that the expectation intersects with about 95% of the 95% intervals.




Fit to index data for 6_OR_REC. Lines indicate 95% uncertainty interval around index values. Thicker lines (if present) indicate input uncertainty before addition of estimated additional uncertainty parameter.
*file: index2_cpuefit_6_OR_REC.png*

* should probably be SE: fisheries scientists mix up “standard deviation” and “standard error” all the time

# Tuning indices of abundance

“Although standard model tuning practices, including the estimation and use of added variance parameters within stock synthesis, represent best current practices, STATs should do their best to be cautious about adding variability to an index as a means of resolving model structure issues such as conflicts among data sources. Rather, STATs should identify an error structure appropriate for the data. When including additional variance to indices, one should look for possible overinflation of variance due to conflicts with other data (e.g., biological compositions). In those instances, it may be more appropriate to determine which data sources contain the most representative population signal and justify the need to add variance to index values. Sensitivity analyses should be conducted to evaluate assumptions about which data sets and types are most representative.”

from Accepted Practices Guidelines for Groundfish Stock Assessments
https://www.pcouncil.org/documents/2025/03/accepted-practices-guidelines-for-groundfish-stock-assessments-accepted-practices-guidelines-for-groundfish-stock-assessments-in-2025-and-2026-march-2025.pdf/

```
#_Q_setup for fleets with cpue or survey data
#_1:  fleet number
#_2:  link type: (1=simple q, 1 parm; 2=mirror simple q, 1 mirrored parm; 3=q and power,
#_3:  extra input for link, i.e. mirror fleet# or dev index number
#_4:  0/1 to select extra sd parameter
#_5:  0/1 for biasadj or not
#_6:  0/1 to float
#_   fleet      link link_info  extra_se   biasadj     float  #  fleetname
        1         1         0         1         0         1  #  BottomTrawl
        3         1         0         1         1         0  #  Hake
        6         1         0         1         0         1  #  JuvSurvey
        7         1         0         1         1         0  #  Triennial
        8         1         0         1         0         1  #  NWFSC
        9         1         0         1         0         1  #  ForeignAtSea
-9999 0 0 0 0 0
```

```
 PHASE    env-var    use_dev   dev_mnyr   dev_mxyr     dev_PH      Block    Blk_Fxn  #  parm_name
    -1          0          0          0          0          0          0          0  #  LnQ_base_BottomTrawl(1)
     2          0          0          0          0          0          0          0  #  Q_extraSD_BottomTrawl(1)
     1          0          0          0          0          0         10          1  #  LnQ_base_Hake(3)
     2          0          0          0          0          0          0          0  #  Q_extraSD_Hake(3)
    -1          0          0          0          0          0          0          0  #  LnQ_base_JuvSurvey(6)
     2          0          0          0          0          0          0          0  #  Q_extraSD_JuvSurvey(6)
     2          0          0          0          0          0          9          1  #  LnQ_base_Triennial(7)
    -2          0          0          0          0          0          0          0  #  Q_extraSD_Triennial(7)
    -1          0          0          0          0          0          0          0  #  LnQ_base_NWFSC(8)
    -2          0          0          0          0          0          0          0  #  Q_extraSD_NWFSC(8)
    -1          0          0          0          0          0          0          0  #  LnQ_base_ForeignAtSea(9)
     2          0          0          0          0          0          0          0  #  Q_extraSD_ForeignAtSea(9)
```

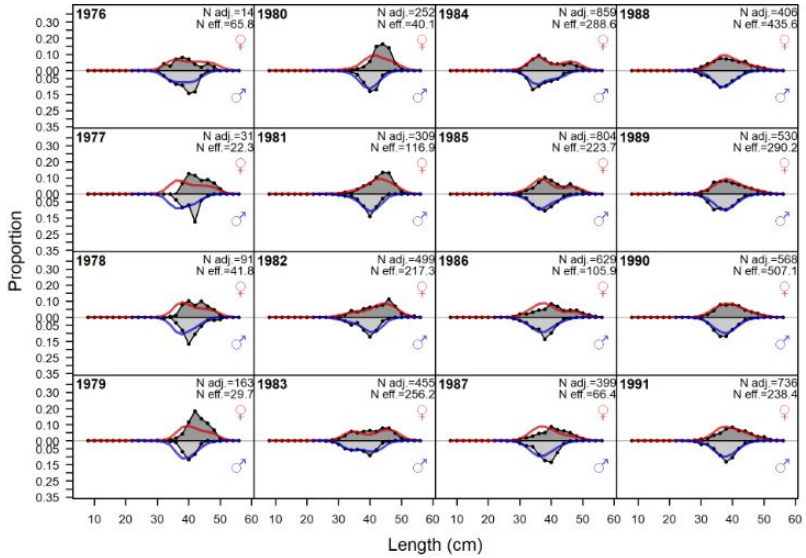
# Tuning composition data

- This is the most complex task, has been the focus of the most research, and has many methods available
- There are 3 options supported by SS3 and r4ss:
  - McAllister-Ianelli (“M-I”, used in Widow base and Yelloweye sensitivity)
  - Francis method (“Francis”, used in Yelloweye base and Widow sensitivity)
  - Dirichlet-Multinomial (“D-M”, introduced in 2017, but not used by either model)
- The first two require iterative tuning of input sample sizes, Dirichlet-Multinomial adds an estimated parameter
- The Francis model has been the most common choice in recent assessments (McAllister-Ianelli didn’t perform as well in simulations and D-M often seems to provide too much weight to comps leading to less plausible model results)
- Tuning applies to all observations for a given fleet, although D-M approach can apply separately to retained and discard comps.

# Tuning composition data: McAllister-Ianelli

- Based on ideas in: McAllister, M. K., and J. N. Ianelli. 1997. “Bayesian Stock Assessment Using Catch-Age Data and the Sampling - Importance Resampling Algorithm.” Can. J. Fish. Aquat. Sci. 54: 284–300.
- SS3 calculates an effective sample size based on the difference between observed and expected for each composition vector which can be compared to the input sample size
- Algorithm requires adjusting the input sample size so that the harmonic mean of the effective sample size matches the arithmetic mean of the adjusted input sample size for each fleet.

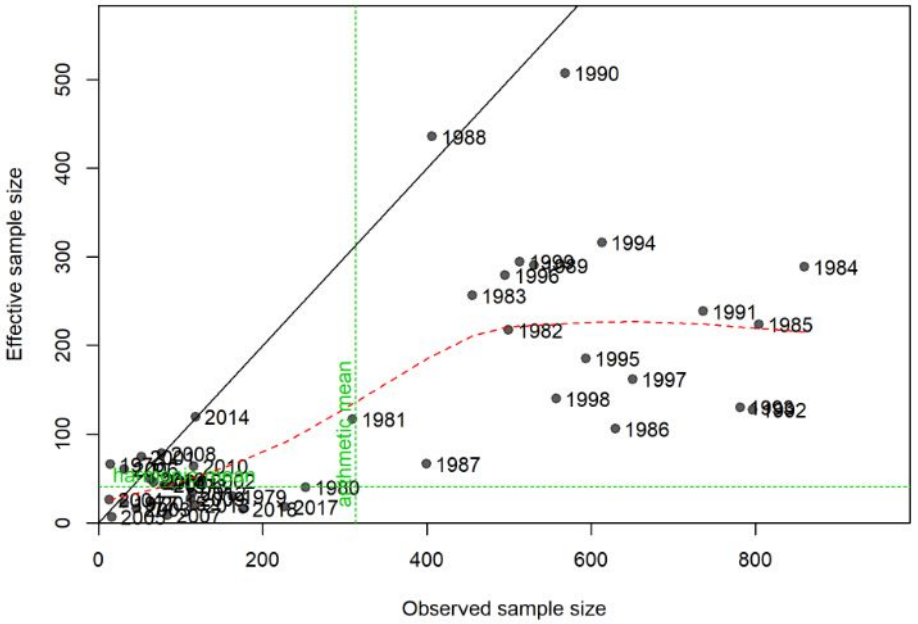
# Tuning composition data: McAllister-Ianelli




Length comps, retained, BottomTrawl (plot 1 of 3).

'N adj.' is the input sample size after data-weighting adjustment. N eff. is the calculated effective sample size used in the McAllister-Iannelli tuning method.

*file: comp_lenfit_flt1mkt2_page1.png*



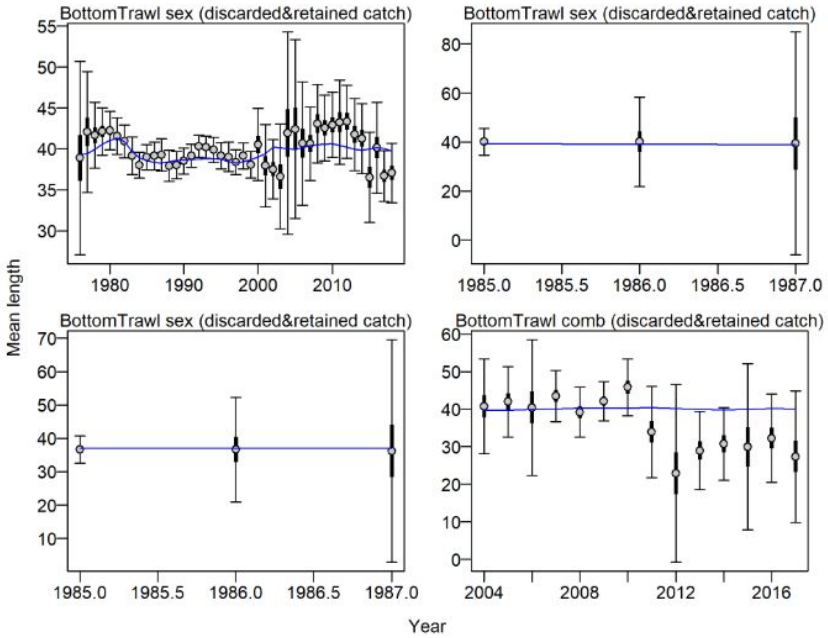

N-EffN comparison, Length comps, retained, BottomTrawl

*file: comp_lenfit_sampsize_flt1mkt2.png*

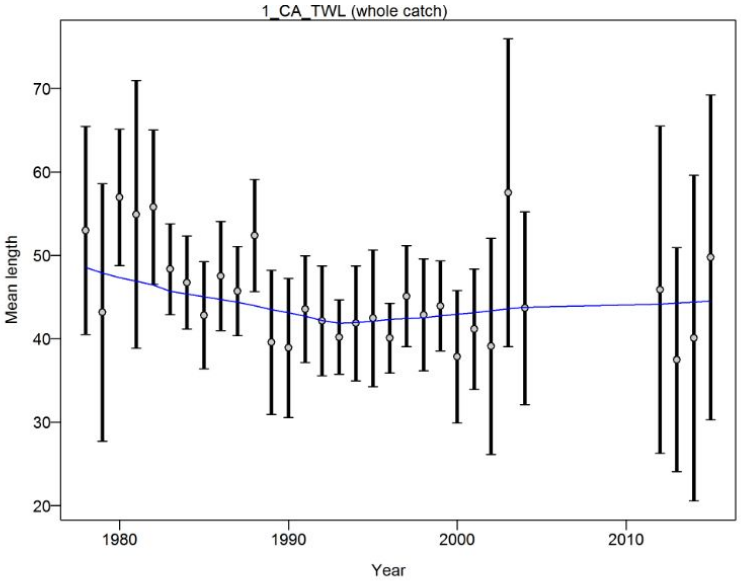
# Tuning composition data: Francis method

- Based on Francis, R. I. C. Chris, and Ray Hilborn. 2011. “Data Weighting in Statistical Fisheries Stock Assessment Models.” Can. J. Fish. Aquat. Sci. 68: 1124–38. https://doi.org/10.1139/f2011-025.
- Attempts to allow the model to fit the indices of abundance by “right-weighting” the comp data so that the variability among observations matches what you would expect from the adjusted input sample sizes.
- Often results in smaller adjusted input sample sizes than other methods
- Can be sensitive to outliers in the comp data

# Tuning composition data: Francis method




Mean length for BottomTrawl 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 BottomTrawl:
0.0559 (0.0425-0.0805)




Mean length for 1_CA_TWL 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 1_CA_TWL:
1.0053 (0.7033-1.8697)

# Implementing M-I and Francis methods

- SS3 output includes tables “Length_Comp_Fit_Summary” and “Age_Comp_Fit_Summary” which include the information required for the M-I method.
- The `r4ss::SS_plots()` function calls on `SSMethod.TA1.8()` and `SSMethod.Cond.TA1.8()` to create the Francis diagnostic plots and recommended tunings
- Easiest approach by far is to use `r4ss::tune_comps()` where the `option` argument can be one of "Francis",  "MI", "none", and "DM". Also see `niters_tuning` argument among others.
- However, tuning used in Widow model is complex

# Tuning comps for Yelloweye Rockfish

- Sample size multipliers are near the bottom of the control file, and can be automatically updated by `r4ss::tune_comps(option = "Francis")`

```
295  # Input variance adjustments factors:
296   #_1=add_to_survey_CV
297   #_2=add_to_discard_stddev
298   #_3=add_to_bodywt_CV
299   #_4=mult_by_lencomp_N
300   #_5=mult_by_agecomp_N
301   #_6=mult_by_size-at-age_N
302   #_7=mult_by_generalized_sizecomp
303  #_Factor  Fleet  Value
304       4      1   0.84204
305       4      2   0.35975
306       4      3  0.600663
307       4      4  0.182385
308       4      5  0.486995
309       4      6  0.395623
310       4      7  0.892214
```

# Tuning comps for Widow Rockfish

- Older versions of SS3 had a lower limit of 1.0 on the adjusted input sample size. This limit could be influential for conditional-age-at-length data where some length bins have small numbers of ages prior to tuning
- The 2015 widow assessment avoided this limit by applying the sample size multipliers to the likelihood weights (lambda) instead of the variance adjustment factors.
- The 2015 widow assessment further down-weighted the comp data by multiplying the lambdas by 0.5 “to account for the potential double use of data since length and age are observed from the same fish.”
- In the current SS3 version, this lower limit can be controlled by the user, so it’s likely the same results could be replicated using sample size multipliers instead of lambdas and the `tune_comps()` function could be used, but the 0.5 multiplier would need to be applied after running that function.

# A few more thoughts on double counting

- The approach of down-weighting by 0.5 to account for the double-counting of lengths and ages from the same pool of fish was inconsistently applied across assessments.
- The 0.5 weight added to the McAllister-Ianelli tuning will generally brings the adjusted input sample sizes closer to those estimated via the Francis method which might be a good thing.
- The adjusted input sample size is typically much lower than the number of fish measured, so the sum of the adjusted inputs for length and age will always be lower than the number of sampled fish.
- An alternative approach to avoid double-counting is to remove the fish with ages from the set of fish used to calculate length compositions, ensuring independence.

# Tuning catch and discards

- Catch data is never tuned (input uncertainty is typically 0.01) because it can be fit perfectly through calculated annual F rates
- Discard data are rarely tuned and no tuning was using for Widow

# Links

- SS3 User Manual section on data weighting https://nmfs-ost.github.io/ss3-doc/SS330_User_Manual_release.html#sec:DataWeight
- Unofficial PFMC assessment handbook section on data weighting https://pfmc-assessments.github.io/pfmc_assessment_handbook/03-stock-synthesis.html#model-tuning
- CAPAM workshop on “Data conflict and weighting, likelihood functions, and process error” in 2015 (includes recorded presentations): https://www.capamresearch.org/data-weighting/workshop
- Fisheries Research special issue associated with CAPAM meeting: https://www.sciencedirect.com/journal/fisheries-research/vol/192