"*and he saith unto them, Follow me, and I will make you fishers of [data]*"
(i.e., how to sort and weigh data)

Photo: Chris Miller, csmphotos.com

James Thorson (with collaborators noted)

Earliest fisheries survey

Mark 6:40 – 6:44

[A crowd] sat down in ranks, by hundreds, and by fifties.

"How many loaves do you have?"

"Seven," they replied, "and a few small fish."

…

he had taken the five loaves and the two fishes … and the two fishes divided he among them all. And [five thousand men] did all eat, and were filled.

Large outlier

(was recorded in meeting minutes)

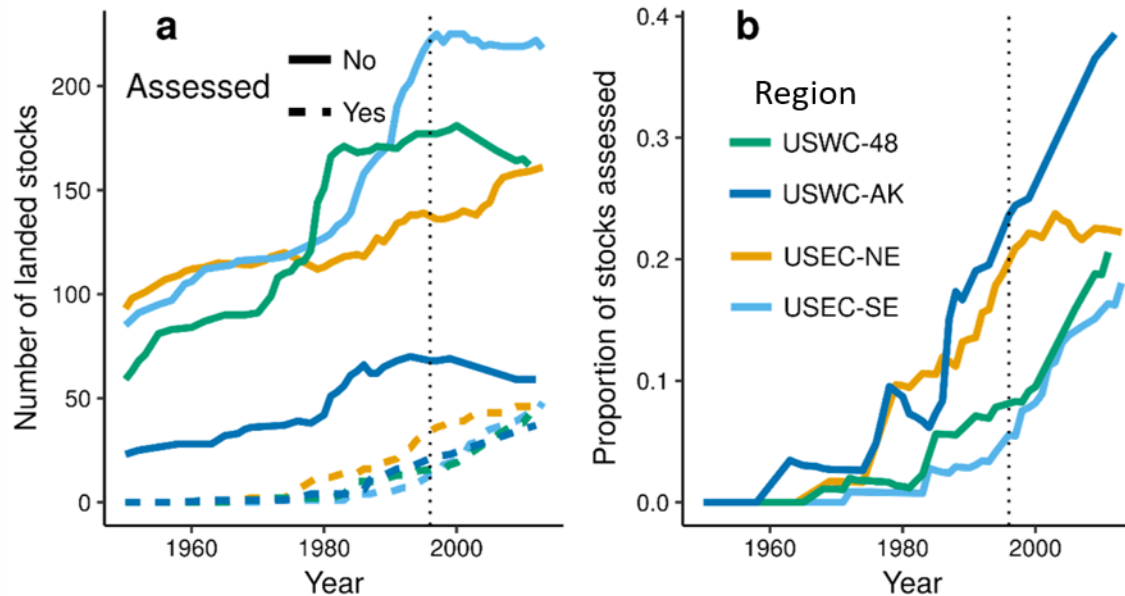(some small disagreement between rapporteurs Mark and John)

Estimate proportion of stocks (species landed in region) with quantitative assessment (pop-dy model)

1. **NOAA landings database**
   - Combine states into four regions (Alaska, West Coast, NE US, SE US)
   - Includes Federal and State landings

2. **RAM management database**
   - Stock assessment ≡ Population model with management targets

Neubauer, Thorson, Melnychuk, Methot, Blackhart, (2018). Drivers and rates of stock assessments in the United States. *PLOS ONE*, *13*(5), e0196483. https://doi.org/10.1371/journal.pone.0196483

Message	Solution
Never use Pearson residuals	Simulation residuals (Dunn-Smyth using PIT) are always better
Leave no variance unexplained	1. Estimate variance of data 2. Estimate variance of residuals 3. Develop scenarios (process errors, covariates and model changes) such that #2 approaches #1
Reduce cost for doing stuff	Predict trade-off in changing age data

Discussion bonus: Percent deviance explained, likelihood residuals, and saturated/null models

Defining residuals

Pearson:

1. Calculate $\varepsilon = \dfrac{Obs - Pred}{Predicted.SD}$

2. Compare ε with normal distribution

Dunn-Smyth:

1. Simulate new values given fitted model, Obs_r^*

2. Compare true observation with simulated distribution and calculate quantile

3. Apply `qnorm` to quantiles, and compare with normal distribution

**Probability integral transform (PIT):**

1. Same as Dunn-Smyth, except randomize quantile between lower and upper bound when simulating from a probability mass function

In preparation. Kelli F. Johnson, Cole C. Monnahan, Ian J. Stewart, Ian G. Taylor, James T. Thorson. Simulation residuals provide more meaningful, accurate diagnostics for composition data than the status quo Pearson residuals.

Defining residuals

Probability integral transform (PIT):

1. Same as Dunn-Smyth, except randomize quantile between lower and upper bound when simulating from a probability mass function
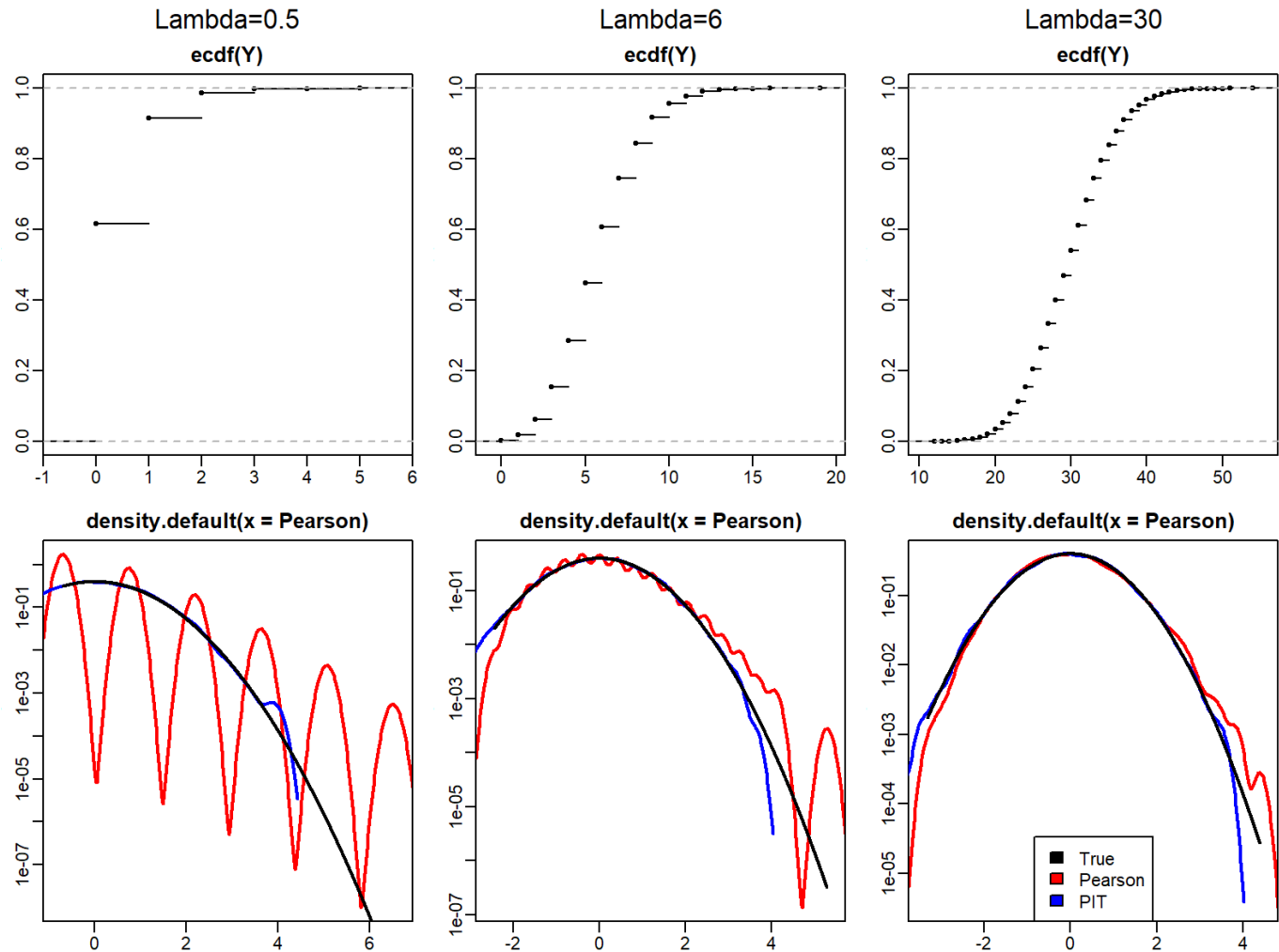
Never use Pearson residuals

Demonstration:

- Compare PIT and Pearson residuals for a catch-curve
- Use a 2x2 experiment involving:
  - Sample size: Small (N=100) vs. very large (N=10,000)
 - Specification: Mis-specified (sigmaR=0.7) vs. correctly specified (sigmaR=0)

Details:

- Logistic selectivity in operating and estimation model
- Z = 0.2
- Age at 50% selection = 3
- A_max = 30
- Selectivity slope at A50% = 1

Never use Pearson residuals

Pearson has weird artefacts even in well-specified model given low sample size (top-left panel)

PIT residuals follow normal distribution when model is correct (top row), and non-normal when it is not (bottom row)

**Both** identify model mis-specification at large sample size (bottom right)
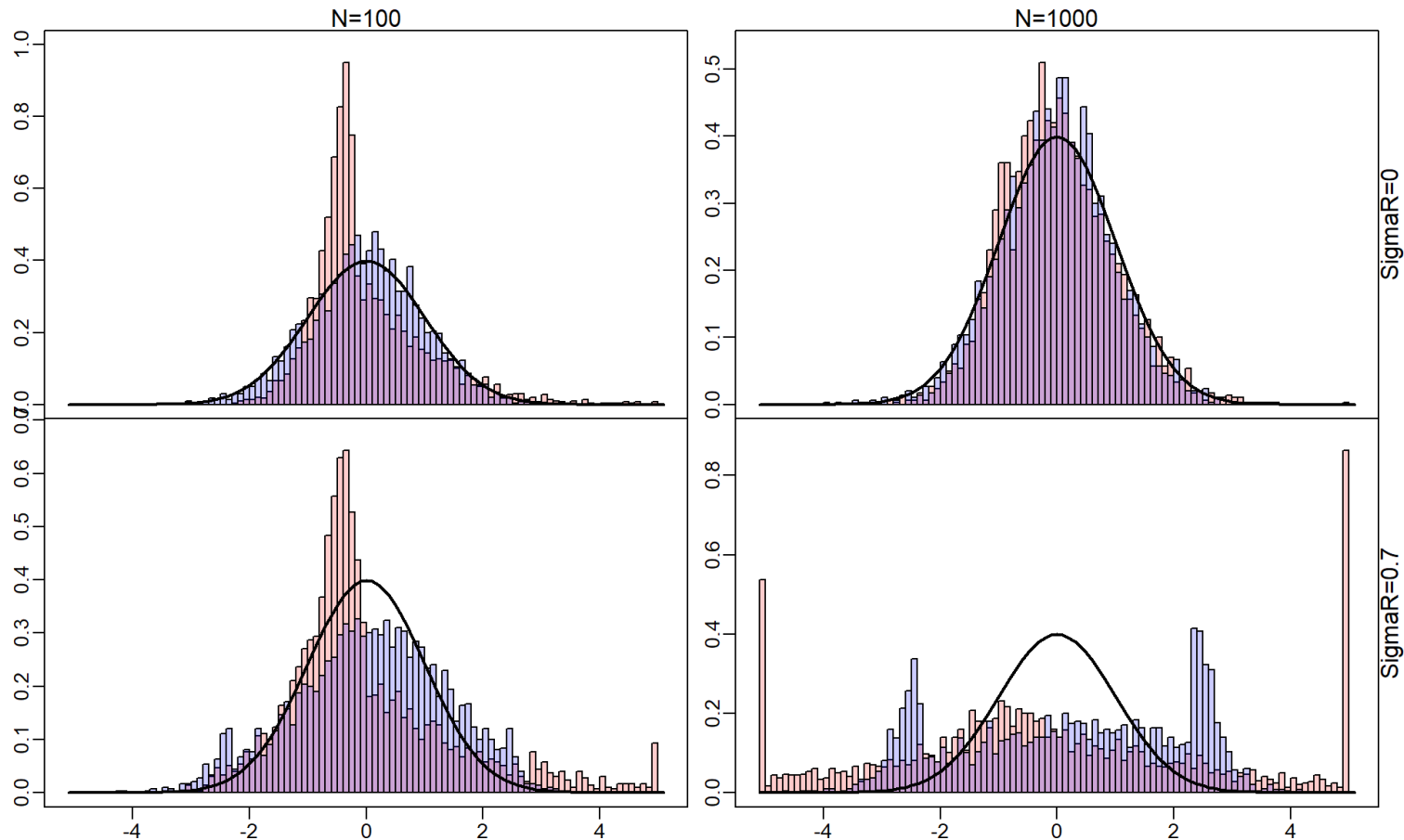
Suggestion #2:
Leave no variance unexplained

Photo: Chris Miller, csmphotos.com

Thorson, & Haltuch. (2018). Spatiotemporal analysis of compositional data: increased precision and improved workflow using model-based inputs to stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, *76*(3), 401–414. https://doi.org/10.1139/cjfas-2018-0015

Proposed workflow overview
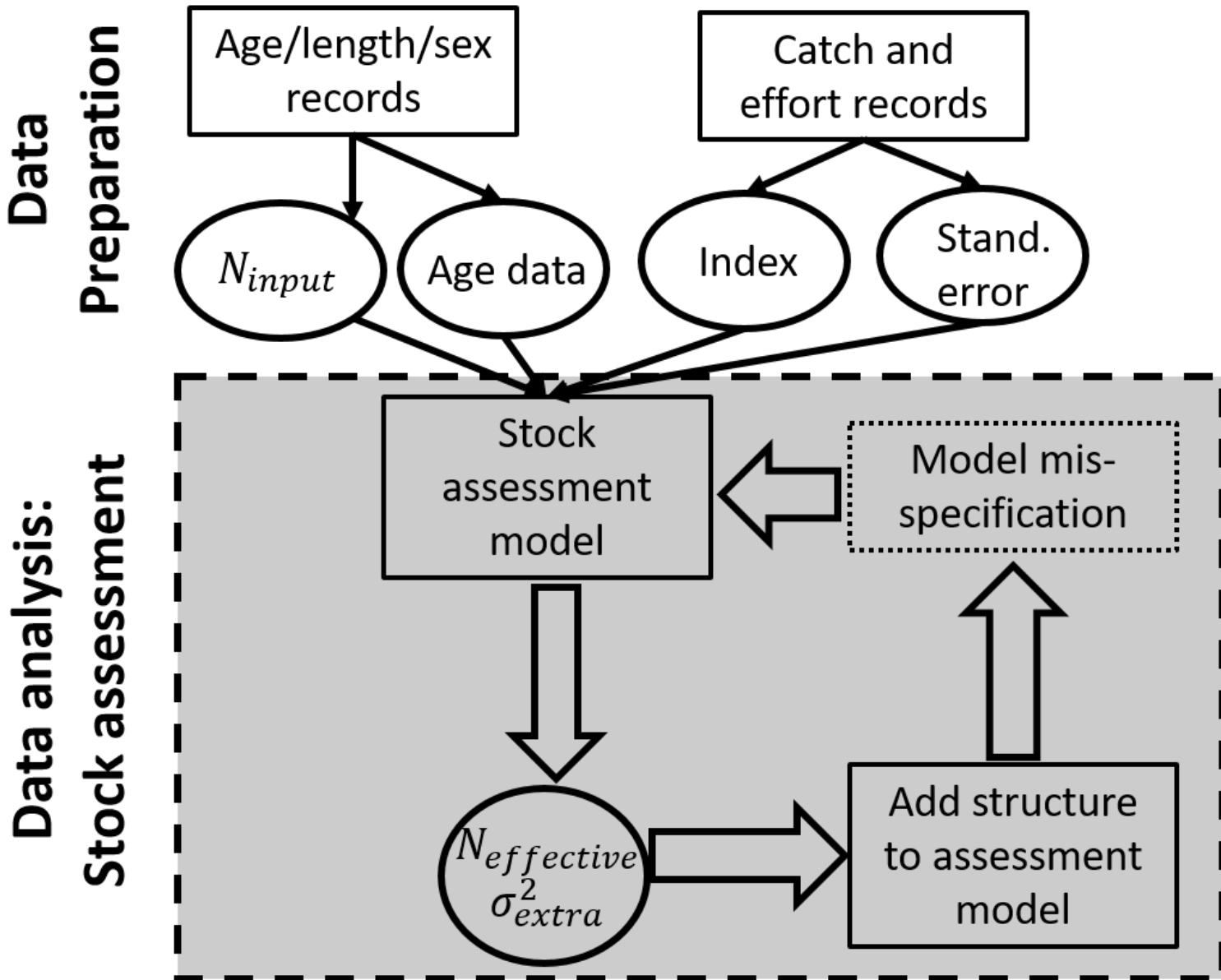
Calculate N_input

1. Fit model predicting abundance/biomass $I_c(t)$ for each bin and year

2. Calculate standard error for proportions

$$\widehat{SE}[\hat{P}_c(t)]^2 \approx \frac{\hat{I}_c(t)^2}{\hat{I}(t)^2}\left[\frac{\widehat{SE}[\hat{I}_c(t)]^2}{\hat{I}_c(t)^2} - 2\frac{\widehat{SE}[\hat{I}_c(t)]^2}{\hat{I}_c(t)\hat{I}(t)} + \frac{\sum_{c=1}^{n_c}\widehat{SE}[\hat{I}_c(t)]^2}{\hat{I}(t)^2}\right]$$
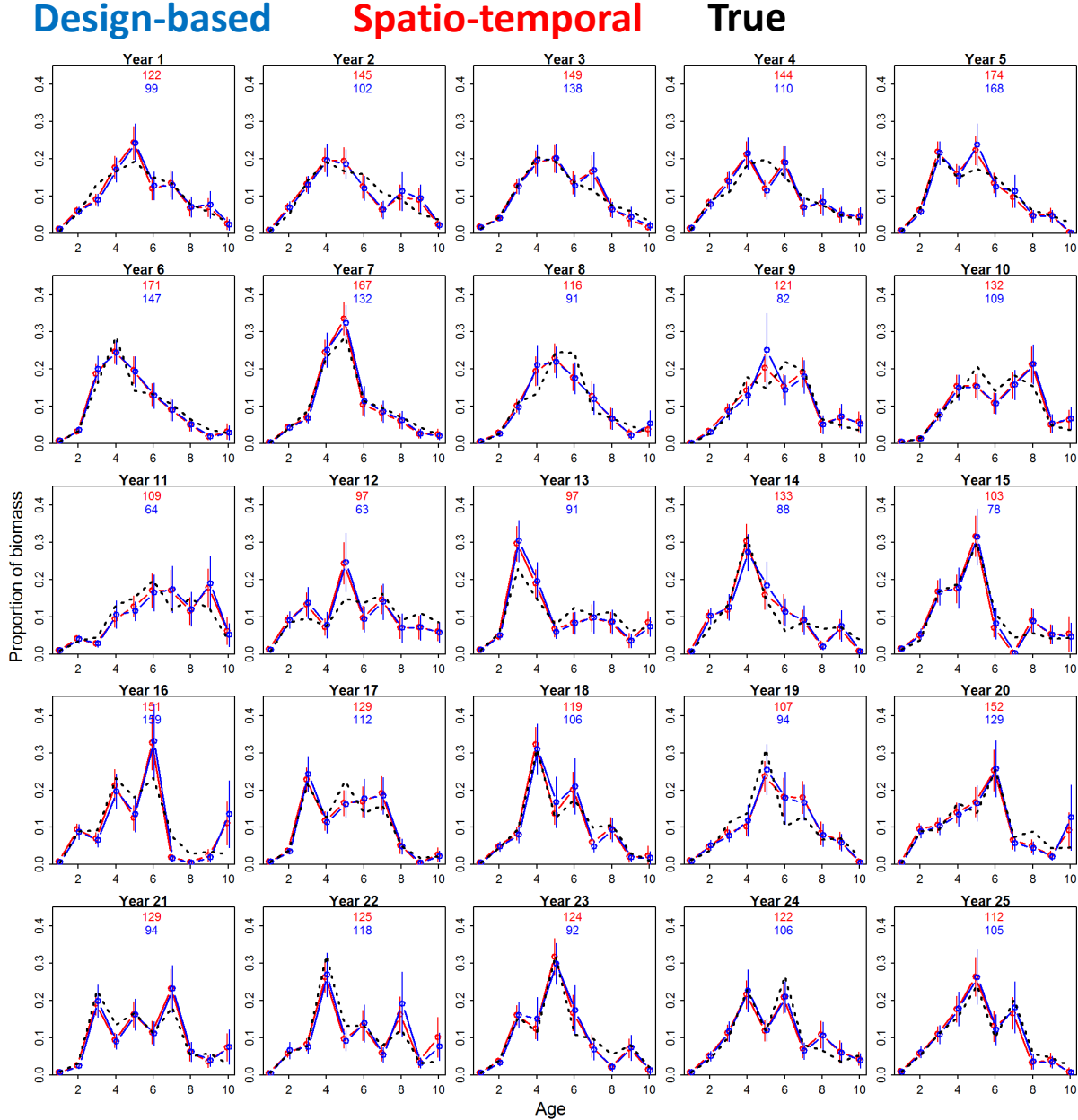
3. Calculate "input sample size" $N_{input}(t)$

$$N_{input}(t) = \text{Median}_c\left[\frac{\hat{P}_c(t)\left(1 - \hat{P}_c(t)\right)}{\widehat{SE}[\hat{P}_c(t)]^2}\right]$$

Spatio-temporal comp-expansion

Simulation results

- Design and spatial provide similar results

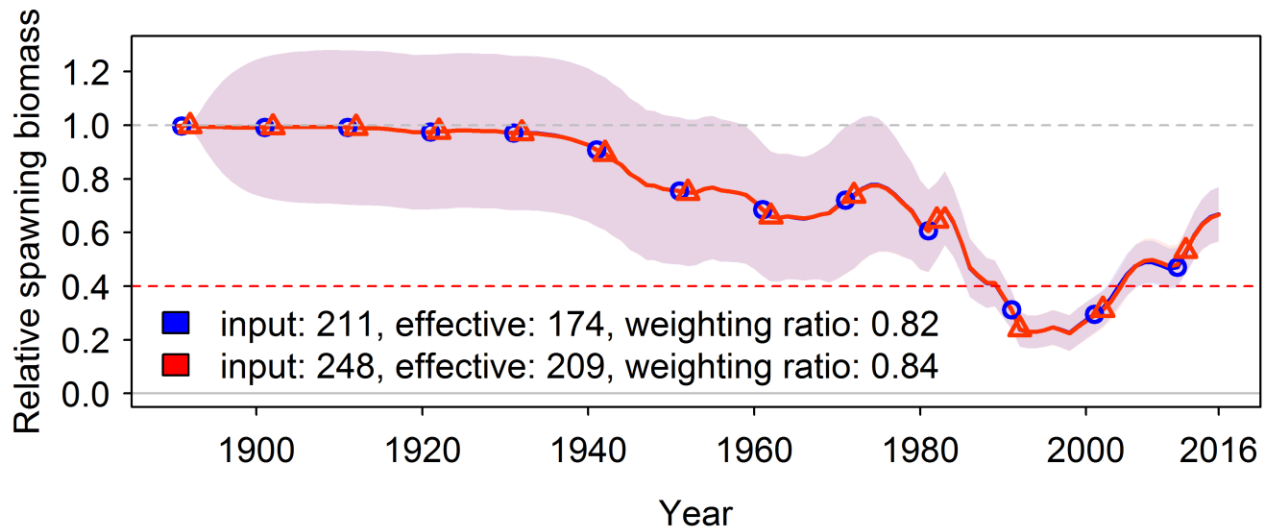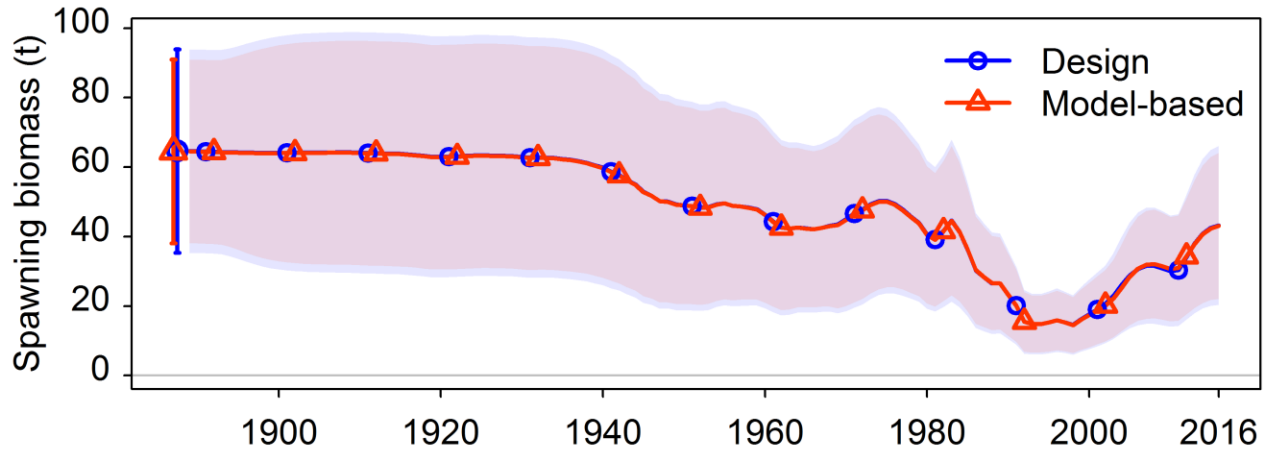- Can track cohorts through OM and both EMs

Spatio-temporal comp-expansion

Lingcod case-study application

- Spatio-temporal has higher input sample size

What if we estimate N$_{effective}$ and time-varying selectivity at the same time?
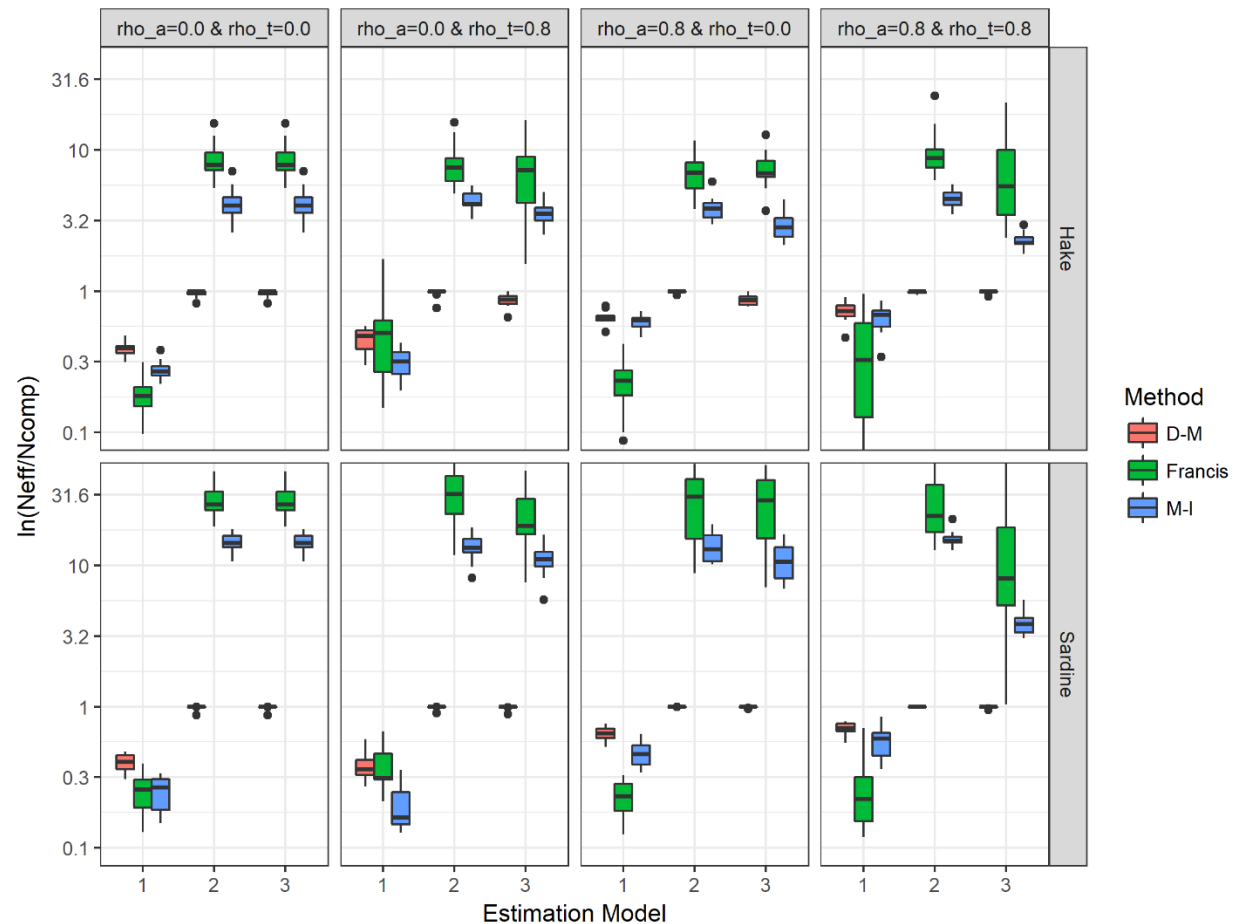
Operating model
Includes variation in selectivity that is correlated over ages ρ_a and years ρ_t

Estimation models
EM1: No estimated deviations
EM2: Estimated deviations are independent for each year and age
EM3: Estimated deviations match the operating model

Suggestion #3:
Reduce cost for doing stuff

Photo: Chris Miller, csmphotos.com

Thorson, Bryan, Hulson, Punt. 2020. Simulation testing a new multi-stage process to measure the effect of increased sampling effort on effective sample size for age and length data. ICES Journal of Marine Science 77:1728–1737.

Reduce cost for doing stuff

Step	Input	Output	Approach to measure effect	Data needed	Responsibility for updating analysis	Processes affecting output
1	$n^*_{nominal}$	n^*_{input}	Non-parametric bootstrap of Hamel-Stewart (2014) sample-size estimator	• Records for sampling data • Age/length subsampling measurements	Survey team	• Variance in age/length composition within and among samples • Variance in biomass/abundance among samples
2	n^*_{input}	$n^*_{effective}$	Analytic formula: $$n^*_{effective} = \frac{\beta n^*_{input}}{\beta + n^*_{input}}$$ where $$\beta = 1 + \theta n_{input}$$	• Stock assessment model fitted to expanded data, and estimating Dirichlet-mulinomial overdispersion	Survey or assessment team	• Degree of overdispersion in fit to expanded composition data (i.e., due to unmodeled variation in growth, selectivity, mortality, etc.)
3	$n^*_{effective}$	$Var^*(X)$	Parametric bootstrap of assessment model	• Stock assessment model with bootstrap simulator	Assessment team	• Which assessment model output is under consideration • Other sources of data available within the model
4	$Var^*(X)$	Management performance	Management strategy evaluation	• Variance and autocorrelation in assessment errors	Research community	• Delays in management (Shertzer and Prager, 2007) • Autocorrelation in errors (Wiedenmann et al., 2015) • Structure of management plan (Wetzel and Punt, 2016)

Reduce cost for doing stuff

Step #2 approach – Theoretical relationship
1. Based on "linear" Dirichlet-multinomial approach to weighting age/length data

$$n_{effective}(t) = \frac{1}{1+\theta} + n_{input}(t)\frac{\theta}{1+\theta}$$

2. Decompose variance into model and sampling error

$$Var_{total} = Var_{sampling} + Var_{model\ mis-specification}$$

… and plug in estimates…

$$\frac{c}{n_{effective}} = \frac{c}{n_{input}} + \sigma^2_{model}$$

… and predict effective sample size $n^*_{effective}$ under a new input sample size n^*_{input}

$$\frac{c}{n^*_{effective}} = \frac{c}{n^*_{input}} + \sigma^2_{model}$$

3. Predict effective sample size $n^*_{effective}$ given new input sample size n^*_{input}

$$n^*_{effective}(t) = \frac{n^*_{input}(t)(1 + \theta n_{input}(t))}{n^*_{input}(t) + \left(1 + \theta n_{input}(t)\right)}$$
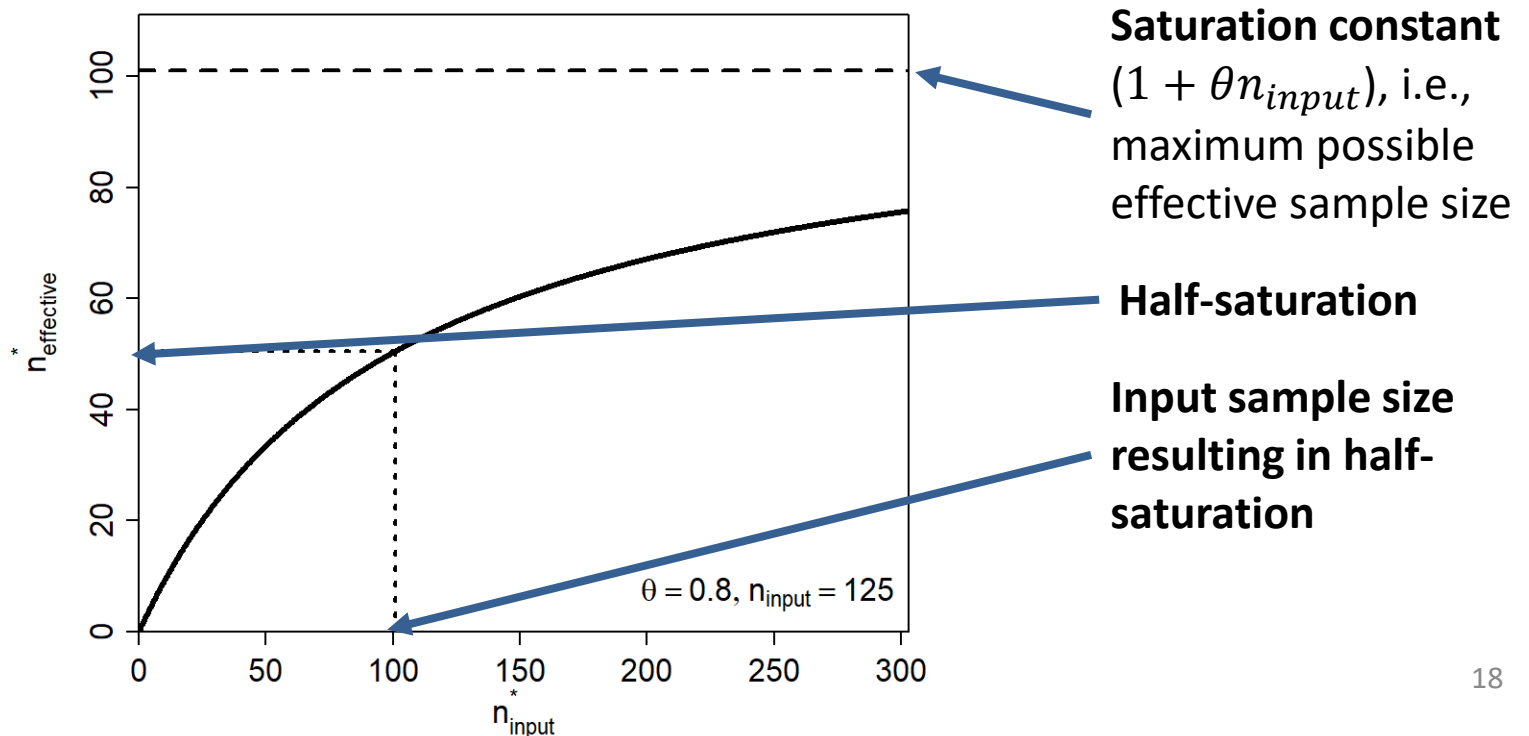
Reduce cost for doing stuff

Step #2 approach – Theoretical relationship

$$n^*_{effective}(t) = \frac{\alpha n^*_{input}(t)}{n^*_{input}(t) + \beta}$$

- a.k.a. Michaelis-Menten relationship with saturation and half-saturation constant $\alpha = \beta = 1 + \theta n_{input}(t)$

Saturation constant
$(1 + \theta n_{input})$, i.e., maximum possible effective sample size

Half-saturation

Input sample size resulting in half-saturation

$\theta = 0.8$, $n_{input} = 125$

Reduce cost for doing stuff

Step #2 approach – Simulation evaluation

- Simulate age-structured dynamics with age-and-time varying selectivity
  - Intentionally mis-specified
- Fit age-structured model with Dirichlet-multinomial and constant selectivity
  - Age data downweighted due to model mis-specification
- Predict change in effective sample size
- Compare with effective sample size given larger input sample size
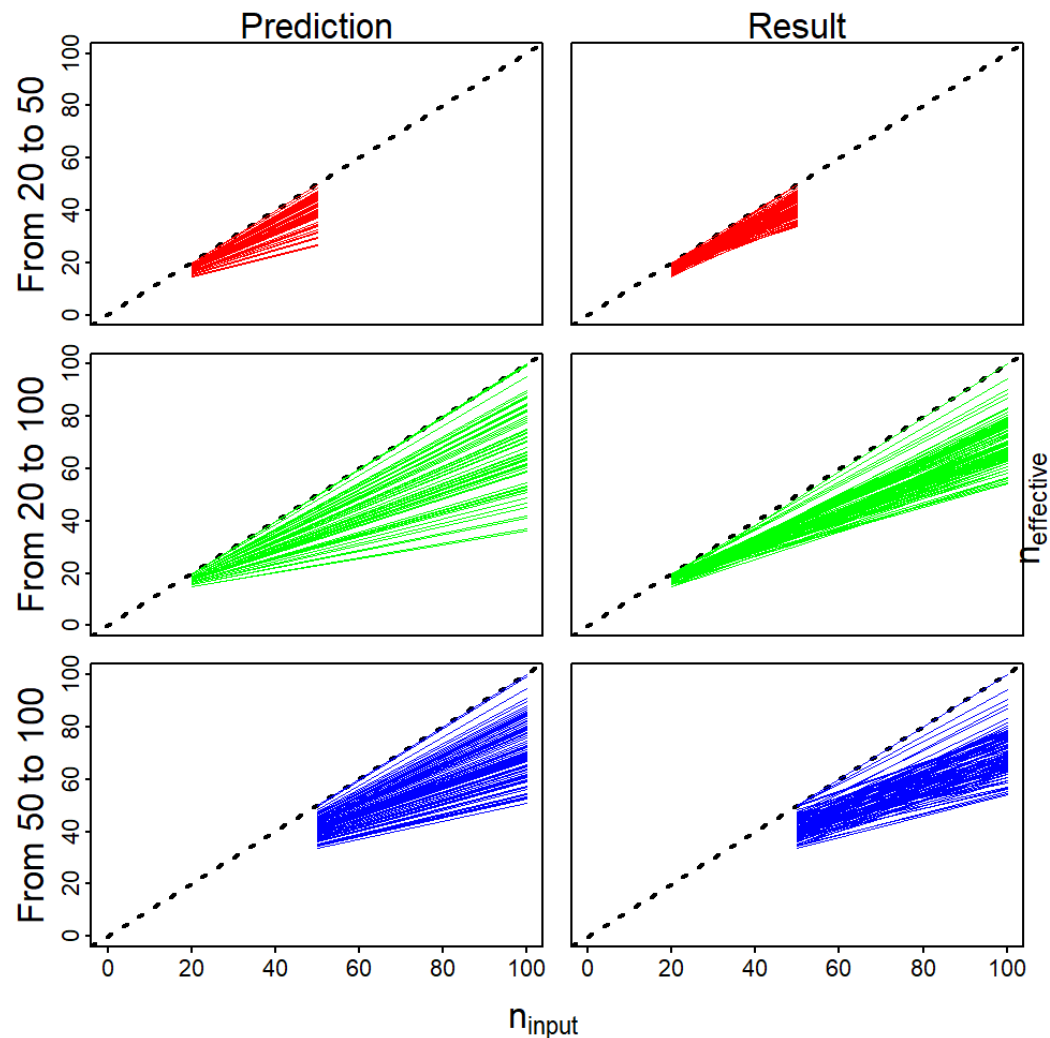
Reduce cost for doing stuff

Future research

- Explore theoretical relationship between $n_{effective}$ and stock-assessment variance

- Pilot-study demonstration
  - Apply to a few selected species
  - Requires operational code for expanding age/length composition data for Step #1

- Explore for generic conclusions
  - Cheaper than routine updates
  - E.g., thornyheads have a higher input sample size for a given number of age-reads than rockfishes

Further reading

- In preparation. Kelli F. Johnson, Cole C. Monnahan, Ian J. Stewart, Ian G. Taylor, James T. Thorson. Simulation residuals provide more meaningful, accurate diagnostics for composition data than the status quo Pearson residuals.

- Thorson, J. T., Bryan, M. D., Hulson, P.-J. F., Xu, H., & Punt, A. E. (2020). Simulation testing a new multi-stage process to measure the effect of increased sampling effort on effective sample size for age and length data. *ICES Journal of Marine Science*, *77*(5), 1728–1737. https://doi.org/10.1093/icesjms/fsaa036

- Thorson, J. T., & Haltuch, M. A. (2018). Spatiotemporal analysis of compositional data: increased precision and improved workflow using model-based inputs to stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, *76*(3), 401–414. https://doi.org/10.1139/cjfas-2018-0015

- Xu, H., Thorson, J. T., & Methot, R. D. (2020). Comparing the performance of three data-weighting methods when allowing for time-varying selectivity. *Canadian Journal of Fisheries and Aquatic Sciences*, *77*(2), 247–263. https://doi.org/10.1139/cjfas-2019-0107