APAIS Calibration Methodology

Jean Opsomer
Westat and Colorado State University

March 20, 2018

Outline

1. Calibration background
2. Background for APAIS adjustment approach
3. 2004-2013
4. 1993-2003
5. 1981-1992

1. Background: Recreational Angler Surveys

- Time series of catch estimates are crucial input in stock assessment models
  - consistency is clearly critical
- Estimates are obtained through two surveys
  - CHTS → FES
  - APAIS (old) → new
- New surveys are significantly improved but have undergone major methodology changes, leading to time series discontinuities

Many official surveys have implemented changes

- Current Population Survey (1994, questionnaire redesign)
- National Household Education Survey (2009, RDD to mail)
- National Crime Victimization Survey (2013, dual frame)
- National Survey of Fishing, Hunting & Wildlife-Associated Recreation (multiple times, mode-questionnaire changes)
- Survey of Graduate Students and Postdoctorates in Science and Engineering (postdoc definition)
- National Resources Inventory (1997, manual to automated photo-interpretation of land cover/use)

How do they deal with survey changes?

- Possible options:
  - do nothing
  - add disclaimers
  - calibrate

- Appropriate in many cases, even for longitudinal surveys
  - small adjustments to methodology, with immaterial effects (common)
  - larger adjustments, but statistical comparison of results reveals no significant effects, e.g. NHES (less common)
- Some repeated surveys make no claims about longitudinal validity of estimates, e.g. FHWAR, Survey of Doctoral Recipients

Add Disclaimers

- When significant changes occur, survey agency alerts data users and provides information about change
- Data users can still perform valid time series analyses, by incorporating changes in models
- This is most commonly implemented approach
  - survey agency does not have to model their data, which is both easier and does not open them up to criticism
  - data users are “free” to choose best way to account for changes

Calibrate

- Develop approaches to preserve integrity of time series, by “matching” estimates before and after change
- Statistical calibration requires overlap sample: side-by-side data of old and new measurements
- CHTS – FES: developed calibration model between modes based on overlap data between both surveys, and incorporating changes in composition of CHTS sample over time (previously reviewed)
- APAIS?

- New APAIS design and estimation procedures implemented in 2013 (wave 2), fully replacing previous methods
- APAIS "pseudo-weights" developed for 2004 – 2013 (wave 1), accounting for
  - selection of site-days as implemented in the field, including alternate sites
  - fraction of days interviewer on site, as fraction of full 24-hour period
- 1981-2003: no weights available, limited/no design information (less for earlier years)

APAIS calibration?

- No overlap period available to fit calibration model
- Very large number of estimates (catch by species by type of trip)

→ Calibrate by adjusting/creating angler-trip weights, preserving micro-data

→ Replace exact calibration by reduction/removal of observed temporal discrepancy in 2013

- Issue: characteristics of trips before 2013 and after 2013 differ more than expected from "typical" angler behavior changes
  - pseudo-weights account for relative frequency of trips by types of sites, waves, modes, and kind-of-day
  - differences still apparent in other characteristics, e.g. area fished, coastal/non-coastal household
- Can we modify weights of pre-2013 trips to correct for trip characteristic discrepancies?

- Consider domain *D* consisting of set of trip characteristics
- Let $U_{D,\,2012}$ = set of trips with those characteristics in 2012, of size $N_{D,\,2012}$
- We know that under valid sampling design settings,

$$\widehat{N}_{D,2012} = \sum_{i \in s} w_i \, I_{\{i \in U_{D,2012}\}} = \sum_{i \in s_D,2012} w_i$$

  is unbiased for $N_{D,\,2012}$, but that is not true here
- Unfeasible calibrated weights

$$w_i^* = \frac{N_{D,2012}}{\widehat{N}_{D,2012}} w_i$$

Ratio calibration (2)

- $N_{D,\,2012}$ unknown, so replace by sample-based quantity obtained from post-2013 design, e.g. $\widehat{N}_{D,\,2014}$
- But: not interested in specific years, so replace by less variable multi-year adjustment

$$w_i^* = \frac{\widehat{N}_{D,\text{new}}}{\widehat{N}_{D,\text{old}}} w_i$$

with $\widehat{N}_{D,\,\text{new}}$ = average of annual estimates for domain D under new design (2013, 2014, 2015, 2016) and $\widehat{N}_{D,\,\text{old}}$ = same under old design (2004, 2005,..., 2013)

Ratio calibration (3)

- Would like to apply ratio corrections to correct for discrepancies in trip distributions by:
  - state and sub-state region
  - year and wave
  - mode
  - area fished
  - coastal/non-coastal household
  - for-hire boat frame membership
- Too many small domains if we consider all possible combinations

- Consider less detailed domains only, and sequentially ratio adjust on each until convergence
- Raking control domains:
  - AF (state, wave, mode, area fished): $\widehat{N}_{D_{,\text{new}_{,\text{AF}}}}$
  - HS (state, wave, mode, coastal/non-coastal household status): $\widehat{N}_{D_{,\text{new}_{,\text{HS}}}}$
  - FH (state, wave, mode, for-hire boat frame status): $\widehat{N}_{D_{,\text{new}_{,\text{FH}}}}$
  - RE (state, wave, mode, substate region): $\widehat{N}_{D_{,\text{new}_{,\text{RE}}}}$

1. Initialize: set t = 0, $w_i^{(t)} = w_i$ for 2004-2013 (wave 1), compute $\widehat{N}_{D\text{, new , AF}}$, $\widehat{N}_{D\text{, new , HS}}$, $\widehat{N}_{D\text{, new , FH}}$, $\widehat{N}_{D\text{, new , RE}}$
2. Let $\widehat{N}_{D,old,AF}^{(t)}$ = averages of estimated AF domain totals for 2004-2012 (include 2013 for wave 1) using weights $w_i^{(t)}$, compute ratios $R_{AF}^{(t)} = \widehat{N}_{D\text{, }new\text{, }AF} / \widehat{N}_{D,old,AF}^{(t)}$, and set $w_{\text{i},AF}^{(t)} = R_{AF}^{(t)} w_i^{(t)}$
3. Starting from $w_{\text{i},AF}^{(t)}$, repeat for HS domains, resulting in ratios $R_{HS}^{(t)}$ and weights $w_{\text{i},HS}^{(t)}$.

4. Starting from $w_{\mathrm{i},HS}^{(t)}$, repeat for FH domains, resulting in ratios $R_{FH}^{(t)}$ and weights $w_{\mathrm{i},FH}^{(t)}$

5. Starting from $w_{\mathrm{i},FH}^{(t)}$, repeat for RE domains, resulting in ratios $R_{RE}^{(t)}$ and weights $w_{\mathrm{i},RE}^{(t)}$

6. Set $w_i^{(t+1)}$ = $w_{\mathrm{i},RE}^{(t)}$

7. Repeat steps 2-6 until convergence (measured by change in $R_{AF}^{(t)}$, $R_{HS}^{(t)}$, $R_{FH}^{(t)}$, $R_{RE}^{(t)}$), and set $w_i^* = w_i^{(t)}$

What happens if fishery changes over time?

- Previous procedure adjusts for design changes if underlying fishery characteristics do not change
- However, what if there are both fishery changes and design changes?

  – lack of data collection overlap under both methods makes confounding unavoidable

  – as long as fishery changes are gradual over time, they can be detected in historical time series

Modification for temporal changes

- Create time series datasets of total trip estimates for each raking control variable for 2004-2013 (wave 1):
  - AF: 145 year-wave estimated totals for each state, mode, area fished
  - (same for HS, FH, RE)
- Fit linear regression and test for significance of slope
  - for categories where slope is not significant: no temporal trend, apply raking as before
  - for categories where slope is significant: temporal trend, replace raking ratio by one computed using 2010-2013 (wave 1) only

Modification for temporal changes (2)

- Computing raking ratios on most recent years avoids removing (most of) time trend in fishery characteristics
- But: increases variability of adjustment

- We would like to apply same procedure, but starting weights not available
- First step: create initial angler-trip weights

– naïve attempt: use CHTS total trip estimates divided by number of intercepted trips

– better attempt: need to account for (unknown) APAIS design changes, so develop proxy for them and include in initial weights

- Count number of site-days with intercepts in state-wave-mode-year domains: $C_{D,1993},\ldots,C_{D,2003}$
- Maximum count across years = $C_{D,\max}$
- Initial weight for angler trip in domain *D* is

$$w_i = \frac{C_{D,\max}}{C_{D,(\text{year})}}$$

→ not calibrated for absolute number of trips in domain, but captures changes in site-day sampling intensity over time

Raking variables and algorithm

- Same as before, plus
  - KOD (state, wave, mode and kind-of-day)
  - MG (state, wave, mode and month groups)
  - AC (state, wave, mode and site activity classes)

→ Account for design effects (already included in initial weights for 2004-2013)

- Raking algorithm as before, with
  - “new” = 2004-2013 (wave 1)
  - “old” = 1993-2003

Modification for temporal changes

- New: create time series datasets of total trip estimates for each raking control variable for 2004-2013 (wave 1)
- Old: create time series datasets of total trip estimates for each raking control variable for 1993-2003
- Fit linear regression and test for significance of slope in old and new time series
    - slope is not significant: apply raking as before
    - slope is significant in old time series: replace raking ratio by one computed using 2001-2003
    - slope is significant in new time series: replace raking ratio by one computed using 2004-2006

- Same procedure as for 1993-2003
  - New period: 1993-2003
  - Old period: 1981-1992
- Temporal trend detection: use 1990-1992 (old) and 1993-1995 (new) if detected