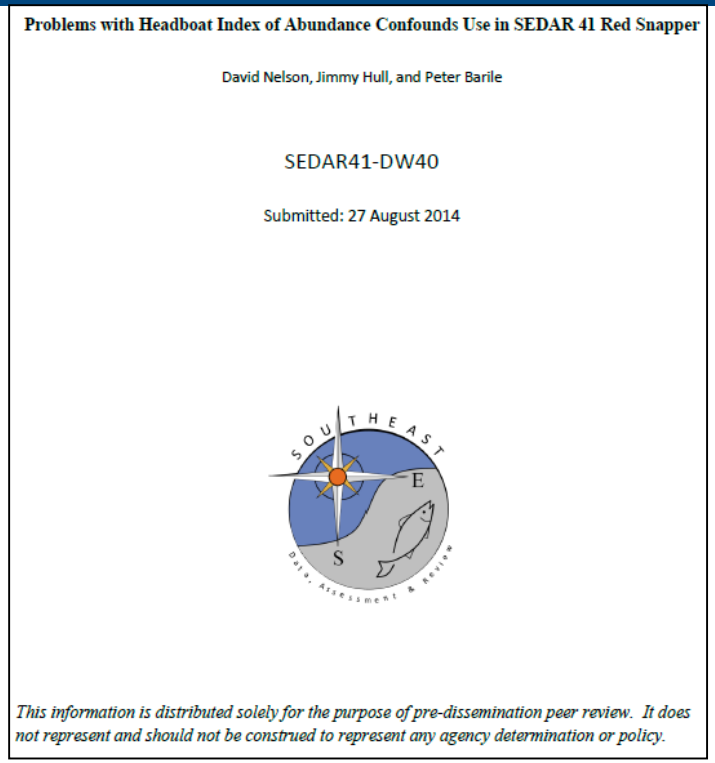
Southeast Region Headboat Survey (SRHS) Data Evaluation

NOAA FISHERIES SEFSC

Impetus

- SEDAR41-DW40 working paper
  - Called into question the data pre-1992
    - Suggested mis-reporting, both by fabrication and by non-reporting
    - Recommended the pre-1992 index be removed
- If data were biased, could affect all the assessments using those data.
- Decided to conduct a full data evaluation.



Problems with Headboat Index of Abundance Confounds Use in SEDAR 41 Red Snapper

David Nelson, Jimmy Hull, and Peter Barile

SEDAR41-DW40

Submitted: 27 August 2014

This information is distributed solely for the purpose of pre-dissemination peer review. It does not represent and should not be construed to represent any agency determination or policy.

Study Design

- Programmatic Component
  - Assessment of current and historic SRHS protocols and policies regarding data quality control and detecting misreported data.
- Analytical Component
  - Set of statistical analyses and metrics used to flag potential misreported data
- Review of the methods

Study Design – Review of the Methods

- Three reviewers provided feedback on the study design: a NOAA senior scientist, a branch chief in the SWFSC, and a fisheries sampling scientist from the AKFSC
- All reviews endorsed the use of the proposed methods to achieve our objectives.
- Some reviewers expressed concern that there may not be a method that would detect average mis-reporting.

Summary of Analytical Results

- 161 (0.04% of the 369,260 trips) extreme outliers were identified in the SRHS data set.

- 74 Vessel/time/area block combinations were flagged for further investigation (11.6% of the 637 vessel-area-time block combinations).

- Only slight underreporting detected, and it was proportional to catch, except in South Florida. Cases of underreporting were found mostly in difficult to identify species (e.g. littlehead porgy).

Discussion - Overview

- In the absence of some independent source of validation, it is generally not possible to determine whether self-reported data that are consistent with others in the dataset are accurate.

- Our approach relied on outlier analysis to identify instances of potential misreporting, followed by detailed investigation of identified records to determine whether a plausible explanation existed or misreporting was likely.

Discussion - Overview

- A primary assumption of this approach is that, if misreporting were prevalent, it was not done in collusion with others to misreport uniformly.

- Similarly, it is unlikely that gross misreporting (collusion by many involved) could have gone undetected by port samplers and SRHS personnel.

- Even though some misreporting could remain undetected by outlier analysis, it is likely to have negligible effects on resulting data products (e.g., abundance indices), because misreported data would be similar to average self-reported data.

Recommendations

Programmatic :

- Continue to evaluate and improve QA/QC procedures for SRHS data.
- Consider re-estimating landings based on extent and magnitude of corrections.
- Link catch records (CRs) to bioprofiles (BPs).
- Digitize Headboat Activity Records and make them available for analysis. (Already underway)
- Consider species identification issues when creating correction factors (k factors) for landings estimation.
- Maintain a living document describing all details of the program procedures and changes in those procedures over time. (Already underway)
- Provide a categorical grouping of the vessels by type (# of anglers, location of fishing, etc.) to facilitate analyses.
- Increase (or continue) efforts to verify data through observer programs and/or whole-haul sampling dockside.

Recommendations

Analytical:

- The vast majority of catch records in the database can be used for analysis and landings estimates.
- Examine the cause of the 161 extreme outliers and correct if possible or remove from the database.
- Consider using a minimum cutoff of number of trips made by a vessel for inclusion in a species-specific index of abundance.
- Identify and filter vessels or trips that fall outside the range of those relevant for analyses of interest.

SRHS Data Evaluation - Review

List of outside (non-Beaufort Lab) reviewers

- 2 from Woods Hole, MA, NEFSC, Population Dynamics Division
- 2 from Silver Spring, MD, Marine Recreational Information Program (MRIP)
- 1 from Miami, FL, SEFSC, Fisheries Statistics Division

General Comments:

+ The analytical methods appear to be appropriate.
+ The results from the analyses described appear to be complete, and they are presented clearly with appropriate measures.
+ The review of the SRHS program and procedures was a sensible approach.
+ The conclusions stated in the report are strongly supported by the results.
+ The report has succeeded in achieving its stated objectives.
- The analyses are better suited for identifying erroneous observation rather than systematic bias.
- Consider a transformation of the catch data.
- The available tools are suitable for identifying outliers, they are less suitable for identifying false reports when the reporter patterns the false report around current average conditions.

Additional Information

- SEDAR 41 working paper (SEDAR41-DW46)

http://sedarweb.org/s41dw46-headboat-data-evaluation

- SEDAR 41 presentation

Contact Julia Byrd (SEDAR), Ken Brennan (SEFSC), or Erik Williams (SEFSC) for a copy of this.

Questions?