SEDAR Stock Assessments:

Transitioning from Benchmarks/Updates to Research/Operational

NOAA FISHERIES

SEFSC

Erik H. Williams

Beaufort, NC Laboratory

The problem: Balancing the three T's

Existing combination of benchmark, standard and update assessments is very transparent, reasonably thorough, but too slow for the demand
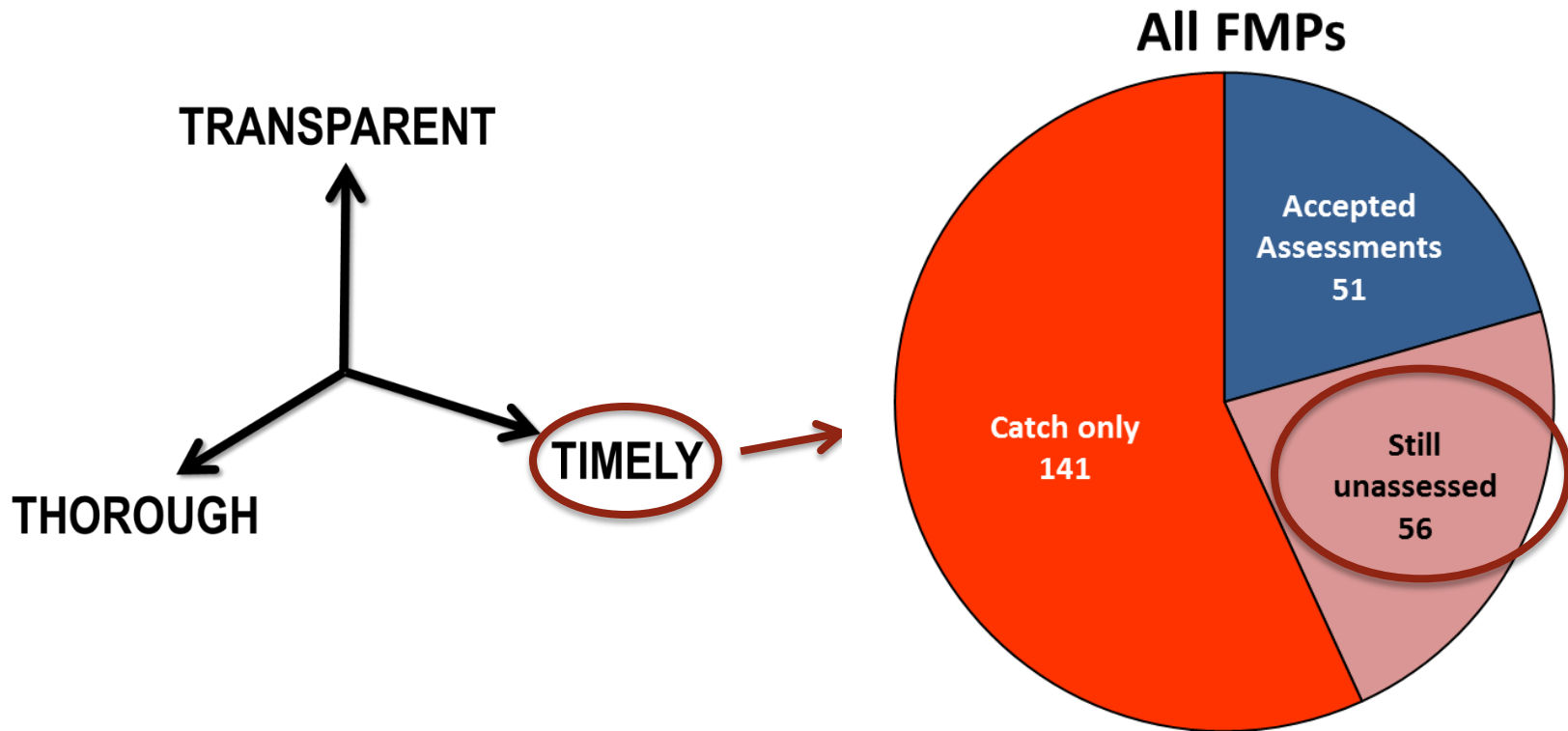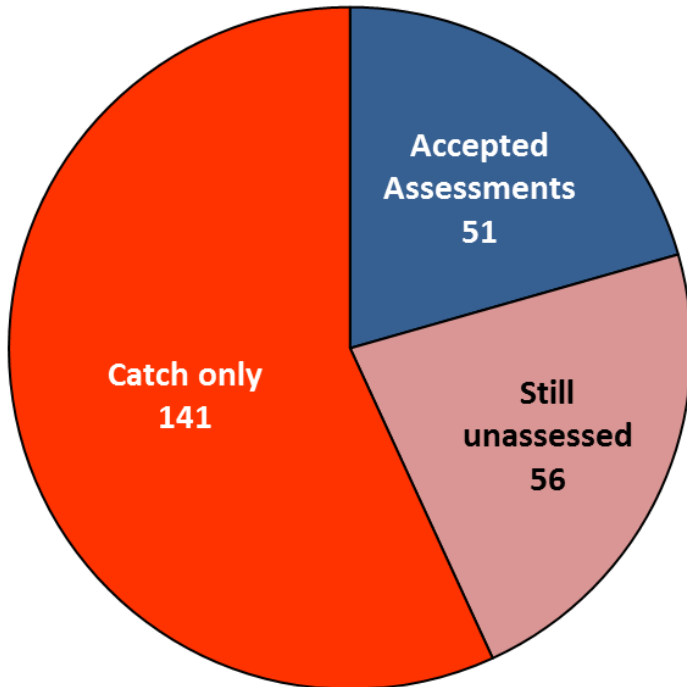

Data from Gulf of Mexico and South Atlantic, last updated 2015

Balancing the three T's

Current assessment rates are too slow ≈ 1 per year per person

TRANSPARENT

THOROUGH

TIMELY

All FMPs

Catch only
141

Accepted
Assessments
51

Still
unassessed
56
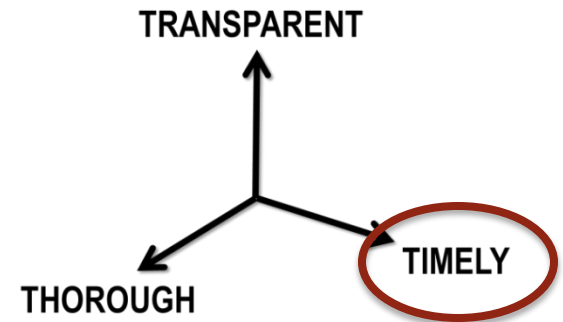
Potential assessment leads: 20 people**

Stocks that can be assessed: 107

Assessment rate in current processes: 1 pyr^{-1}

Average time between assessments: 5.3 years

**Hypothetical and illustrative staff size, actual staff size is smaller and fluctuates*

Data from Gulf of Mexico and South Atlantic, last updated 2015

Existing process

Benchmark

Intended to complete a thorough evaluation that accommodates the input of stakeholders and reviewers while under strict deadlines for providing management advice

Standard

Address specific concerns (expressed in the TORS) without deviating too much from previous benchmark

Update

Deviates as little as possible from previous benchmark

NOAA FISHERIES

Existing process

Benchmark

Intended to complete a thorough evaluation that accommodates the input of stakeholders and reviewers while under strict deadlines for providing management advice

Standard

Address specific concerns (expressed in the TORS) without deviating too much from previous benchmark

Update

Deviates as little as possible from previous benchmark

Issues

Data providers have difficulty meeting deadlines because key decisions made along the way can change what is required

Results often criticized by reviewers, but there is little time to address their concerns

Deadlines are pushed and often missed

Word "benchmark" implies "best" to many when in fact it is the first time some components have been examined and implemented

Existing process

Issues

Benchmark

Intended to complete a thorough evaluation that accommodates the input of stakeholders and reviewers while under strict deadlines for providing management advice

Data providers have difficulty
Can't address suggestions
Deadlines pushed or missed
Loaded language (Benchmark)

Standard

Address specific concerns (expressed in the TORS) without deviating too much from previous benchmark

Reasonably fast, but sometimes criticized by stakeholders who think a "benchmark" is better

Update

Deviates as little as possible from previous benchmark

Fast, but often criticized by stakeholders who think a "benchmark" is better and would like more involvement.

NOAA FISHERIES

Existing process

Benchmark

Intended to complete a thorough evaluation that accommodates the input of stakeholders and reviewers while under strict deadlines for providing management advice

Standard

Address specific concerns (expressed in the TORS) without deviating too much from previous benchmark

Update

Deviates as little as possible from previous benchmark
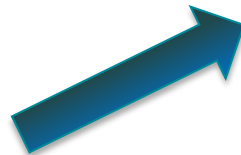
Proposed Changes

Research Cycle

Like a Benchmark, but not intended to produce assessment results for immediate advice to management. The goal is to build a robust tool that will be used to develop timely advice.

Operational Assessment

May follow existing Standard or Update Processes: Designated analysts apply the tool developed by the Research Assessment to the most recent data sets to produce timely management advice.

Research Cycle

- Test, document and review assessment approaches, incorporate new research findings, and evaluate new data streams,

- Conducted similar to current benchmark process with an assessment panel, IPT-style communication and 1-2 workshops

- Review panel meeting with independent external participants (e.g., CIE)

- Findings thoroughly documented as an assessment report, and possibly a NOAA Tech Memo or journal publication commensurate with the degree of novelty of the methods.

- Unresolved issues and ideas for future improvements reported to begin the next cycle of research.

- Not intended to produce assessment results for immediate advice to management, but once vetted, will be operationalized

Operational assessments

- Produce timely advice to management

- Conducted by designated analysts using a suite of previously reviewed procedures and data sets, in consultation with an advisory body comprised of scientists and stakeholders with local expertise

- Minor changes to previous approaches may be considered, if agreed to by the SSC as part of the TORs.

- Findings documented succinctly with an executive summary that makes fishery management advice clearly and quickly accessible

- Anomalies, concerns and research recommendations are documented and made available for future considerations

Advantages of new approach

During research cycles

- Analysts can focus on more thoroughly addressing the major concerns of scientists and stakeholders without the conflicting pressure of finishing the assessment in time for management deadlines

- Suggestions from reviewers can be incorporated and used in the operational phase

- Data providers are not under pressure to provide the most recent data or repeatedly revise inputs

- More opportunities for scientific research that advance the state of the art

During operational assessments

- Standardized, pre-approved approaches will be used such that
  - Implementation errors will be reduced and throughput increased (analysts can focus on updating inputs, implementing only minor changes, and model diagnostics)
  - Assessments will be more reproducible and require less advanced technical skills
  - Data providers will be able to produce inputs more quickly and with minimal effort

- Emphasis will be placed on succinct communication of management advice in plain language (rather than the details of the assessment)

How will it work?

Below is a hypothetical example of two years in the SEDAR cycle with five lead assessment analysts available.

- After two years, 3 research track assessments and 10 operational assessments would be complete

- Long term averages with 5 analysts
  - 1-2 research track assessments per year
  - 4-6 operational assessments per year

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Research Track Assessment					Stock 1													Stock 2						
Research Track Assessment																		Stock 3						
Operational Assessment						Stock 4																		
Operational Assessment				Stock 5							Stock 6							Stock 1						
Operational Assessment	Stock 7							Stock 8							Stock 9									
Operational Assessment					Stock 10						Stock 11									Stock 12				

NOAA FISHERIES

Why make this change now?

We are fast approaching SEDAR 60

- This has provided a tremendous amount of experience and knowledge about the required data, modeling, and communications for our stock assessments
  - Use this experience and knowledge to make the process more efficient.
  - The wheel has been well thought out, designed and built – now lets put it to regular use and not try to re-think it.

Where do we want to be in 20 years?

- Not unreasonable to have annual population estimates for every managed stock
  - This is a step in that direction, shifting us toward more timeliness and efficiency

Questions?

SEDAR stock assessment categories

Operational Stock Assessment

The operational stock assessment category provides management advice quickly and efficiently using previously approved methods and data sources.

- Builds upon approaches developed in previous benchmark and supports incremental improvements.

- Throughput is maximized through a quick and efficient process with few or no public meetings, saving considerable staff time.

- The most recent data available are processed one time based on specifications that are determined in advance (rather than multiple times as is often the case with the current system), saving considerable staff time

- Concise documentation for consistent, standardized public presentation of results.

- Reviews are completed by the Council SSC's (as with current SEDAR update and standard assessments)

- Allows for reasonable flexibility in the model and data to accommodate specific concerns reflected in the Terms of Reference (e.g., previously vetted model approaches and data sets that might be new to the particular stock, or other changes that the SSC feels competent to review).

Steps in the process:

1. Assimilate data necessary for the modeling framework, including the most recently available data. A public meeting (workshop or webinars) should only be required if there is a need to vet the addition of a data stream that is new for the particular stock. (Action: Data Providers)

2. Incorporate data, run the model, and summarize results in a streamlined report. A public meeting (workshop or webinars) should only be required if there is a need to vet changes in the assessment methods previously reviewed and accepted for this particular stock. A change to new software could be considered provided it makes essentially the same calculations and has been reviewed and applied previously to other SEDAR stocks. (Action: Assessment modelers)

3. Review model results. (Action: SSC and Assessment leads)

Expected timeline: 3-6 months

Expected Products: Concise report with an executive summary.

Research Stock Assessment

The research stock assessment category places the emphasis on developing a highly credible stock assessment framework. It should be applied in cases where a new model, hypothesis, or question needs to be answered about a stock/population. It is <u>not</u> intended to provide management advice, but rather set the stage (prototype approach) for operational modeling.

- Serves to answer questions, test hypotheses, or otherwise explore new ideas for assessing a stock or stocks. Establishes scientific credibility of new data types or analysis methods.

- Does not necessarily need to focus on an individual species, such that results might generalize to multiple operational stock assessments.

- Allows for complete flexibility in data and model choice.

- The process should be expected to last up to a year (or more) and involve a series of public meetings. Includes:

  - thorough documentation of new data/methods/performance

  - extensive investigation of model performance

- A hard deadline should be avoided because the necessary steps to achieve a consensus model are too difficult to anticipate. A deadline may hinder options not previously envisioned.

- Reviews should be completed by a panel of independent experts, with the Council SSC's, ultimately providing recommendations for further improvements. Review should be commensurate with the degree of novelty and controversy.

Steps in the process:

1. Schedule the species to be addressed well in advance (2-3 years prior to anticipated completion) so that all relevant data can be processed, analyzed, and finalized for use in the process. Unfortunately much of our data collection involves archiving samples for later analysis. Thus, archived samples for genetics, reproductive measures, and age determination require a fair amount of lead time to complete. Determine stock boundaries as needed. (Action: Data Providers begin data preparations)

2. Hold workshop(s) to assimilate all available data for the species of interest, but not necessarily the most recent data (14 months prior to anticipated completion). Public meetings to be held and input from fishermen will be valuable in understanding the data and its potential uses. Document the proceedings and decisions, particularly where recommendations depart from previously established best practices. (Action: Participants complete assessment report)

3. Data explorations will guide the structure and type of modeling to be built. Build a modeling framework to answer the question/hypothesis. Consider multiple models. Document the final modeling framework being proposed. (Action: Participants complete assessment report)

4. Review modeling framework proposal. Receive recommendations for operational model framework. (Action: CIE and SSC Review and comment on assessment, complete a review report)

Expected timeline: 9-14 months from data workshop completion, but could be longer depending on the hypothesis or question. For example, a question that requires new data collection to answer might require a longer time frame.

Expected Products: Data workshop report, Assessment workshop report, Review report, and an approved/accepted model for use in future operational assessments.

Figure 1. Hypothetical example of two year cycle of the research and operational assessment tracks for five analysts. After two years the results would include 3 research track assessments completed and 10 operational assessments providing management advice. Long term averages for a staff of 5 analysts would work out to 1-2 research track assessments per year and 4-6 operational assessments per year, depending on how many research tracks are chosen in a year.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Research Track Assessment						Stock 1											Stock 2							
Research Track Assessment																	Stock 3							
Operational Assessment					Stock 4																			
Operational Assessment			Stock 5							Stock 6									Stock 1					
Operational Assessment	Stock 7						Stock 8									Stock 9								
Operational Assessment					Stock 10						Stock 11									Stock 12				