ANALYSIS OF ASSESSMENT CAPACITY AND TARGET FREQUENCIES FOR CONDUCTING WEST COAST GROUNDFISH ASSESSMENTS

Introduction

The primary intent of this report is to heighten the awareness of Pacific Fishery Management Council (Council), its advisors, and groundfish process participants of a growing long-term challenge in achieving assessment frequencies that are suitable for each species' biology and importance; referred to in the Prioritization exercise as "target frequency". Over the past decade, there has been a divergence between the growth in the number of species and modeled areas that form the assessment inventory (i.e., the cumulative number of species assessed with Category 1 or 2 methods) and our capacity to develop and review assessments. The calculation of target assessment frequency (i.e., the maximum desired number of years between assessments) draws upon several species' characteristics that bear on how often assessments would ideally be conducted and it plays an important role in the Prioritization scoring. Following a brief overview of that role, this report reviews changes in assessment capacity and the assessment inventory over the past 20 years, and highlights the challenges those developments have presented for refreshing existing assessments with the desired frequency.

Background

The current general framework for evaluating data relevant to prioritizing groundfish species for upcoming assessment cycles was first used as part of the Council's assessment planning process during the spring of 2016, following the publication of a NOAA Technical Memorandum on assessment prioritization the previous summer (Methot, 2015). While the specifics of some calculations and the appearance of the summarized material has evolved over time, including the website developed for the 2024 process, the basic scope and approach remain very similar to that first incarnation.

One of the most important elements of this framework is the identification of a target assessment frequency. Most species tend to retain similar rankings over time, or exhibit slow-moving trends, within many of the individual factors included in the prioritization summary, such as those characterizing fishery importance. The ability of the overall scoring to cycle species among lower and higher rankings, over time, relies to a very large degree on the comparisons of target and realized assessment frequency. In addition to this important role in contributing to dynamic overall rankings, assessment target frequency, when compared to the time since the last assessment, provides a useful metric for evaluating the health of the assessment science enterprise. In other words, how well are we doing at providing sufficiently current scientific information, upon which management decisions can be reliably based?

Drawing from Methot (2015), the derivation of target frequency was, from 2016 through 2022, based largely on each species' mean age in fishery catch, modified by adjustments reflecting recruitment variability, along with fishery and ecosystem importance. This approach did not explicitly take each species' lifespan into account, and also effectively meant that a target frequency could only be calculated for species with existing age-structured assessments. In the 2024 package, mean maximum age is now used instead of mean catch age in the revised derivation of target frequencies. This change yields some increases or decreases in target frequency for some species, but the average remains very similar (6.9 years in 2024 vs 7.0 years in 2022). This change also allows a target frequency to be calculated for all 65 species included in the package. A comparison of the distributions of target-frequency values between the 2024 and 2022 packages, in numbers and percentages, is provided in Table 1. In particular, note that species whose harvest specifications are currently informed by data-limited (i.e., Category 3) methods now have target frequencies. Table 2 lists each of the 65 species included in the 2024 groundfish assessment prioritization package, by its target frequency and grouped by assessment category, with species currently managed using data-limited methods located in the bottom section.

Table 1.--Comparison of the distribution of target assessment frequencies included in the 2022 and 2024 Prioritization packages, grouped in 2024 by whether current management is based on benchmark, update, or data-moderate assessments.

	Target Assessment Frequency (years)				
	4	**6**	**8**	**10**	**Total**
As included in the 2022 Prioritization package					
Species with calculated target frequencies	8	13	7	10	38
Percentage of total	21%	34%	18%	26%	
As included in the 2024 Prioritization package					
Species assessed with benchmark or data-moderate methods (Category 1 or 2)	13	7	9	11	40
Percentage of total	33%	18%	23%	28%	
Species assessed with data-limited methods (Category 3)	4	8	7	6	25
Percentage of total	16%	32%	28%	24%	
All species included in the 2024 package	17	15	16	17	65
Percentage of total	26%	23%	25%	26%	

Table 2.--Target assessment frequencies for the 65 groundfish species, as calculated in the Prioritization package, grouped by whether current management is based on benchmark, update, or data-moderate assessments. (excludes Pacific hake)

4-Year Frequency	6-Year Frequency	8-Year Frequency	10-Year Frequency
40 species whose current management is based on benchmark, update, or data-moderate assessments			
Arrowtooth flounder	Blue/Deacon rockfish	Canary rockfish	Aurora rockfish
Big skate	Brown rockfish	China rockfish	Blackgill rockfish
Black rockfish	Dover sole	Copper rockfish	Cowcod
Bocaccio	Gopher/Black-and-Yellow rockfish	Greenstriped rockfish	Darkblotched rockfish
Cabezon		Longspine thornyhead	Greenspotted rockfish
California scorpionfish	Petrale sole	Pacific ocean perch	Pacific spiny dogfish
Chilipepper	Rex sole	Shortspine thornyhead	Quillback rockfish
English sole	Sablefish	Squarespot rockfish	Rougheye/Blackspotted rockfish
Kelp greenling		Vermilion/Sunset rockfish	
Lingcod			Sharpchin rockfish
Longnose skate			Splitnose rockfish
Widow rockfish			Yelloweye rockfish
Yellowtail rockfish			
25 species whose current management is based on data-limited assessment methods			
Curlfin sole	Flathead sole	Flag rockfish	Bank rockfish
Pacific cod	Grass rockfish	Honeycomb rockfish	Redbanded rockfish
Pacific sanddab	Kelp rockfish	Redstripe rockfish	Rosethorn rockfish
Sand sole	Leopard shark	Rosy rockfish	Shortraker rockfish
	Olive rockfish	Speckled rockfish	Silvergray rockfish
	Rock sole	Starry flounder	Yellowmouth rockfish
	Starry rockfish	Stripetail rockfish	
	Treefish		

Misalignment between Assessment Demand and Capacity

The decade of the 2000s was a period of growth and transformation for U.S. marine assessment science, generally, and for west coast groundfish, in particular. In the wake of the National Marine Fisheries Service's (NMFS) Marine Fisheries Stock Assessment Improvement Plan (NMFS, 2001), increased appropriations supported the building of assessment capacity (and related science and data collection) within NMFS, and within graduate schools across the country. Figure 1 illustrates changes in core groundfish assessment staffing at each west coast Science Center, since 2001, along with a summation of supporting positions and formal student positions for both Science Centers, combined. For purposes of this diagram, core staff includes only those individuals who lead or play a major role in developing Category 1 or 2 assessments. Core numbers do not include supervisory staff who are not actively engaged in developing and leading assessments. Assessment staffing at the Science Centers generally increased from 2001, peaking for both core and support/student positions in 2013. This period of growth coincided with the development of the Council's Stock Assessment Review (STAR) process, as well as new NMFS programs to conduct groundfish surveys and collect data at sea during commercial fishing trips. Benchmark assessments for 10 previously unassessed species were conducted between 2003 and 2009.

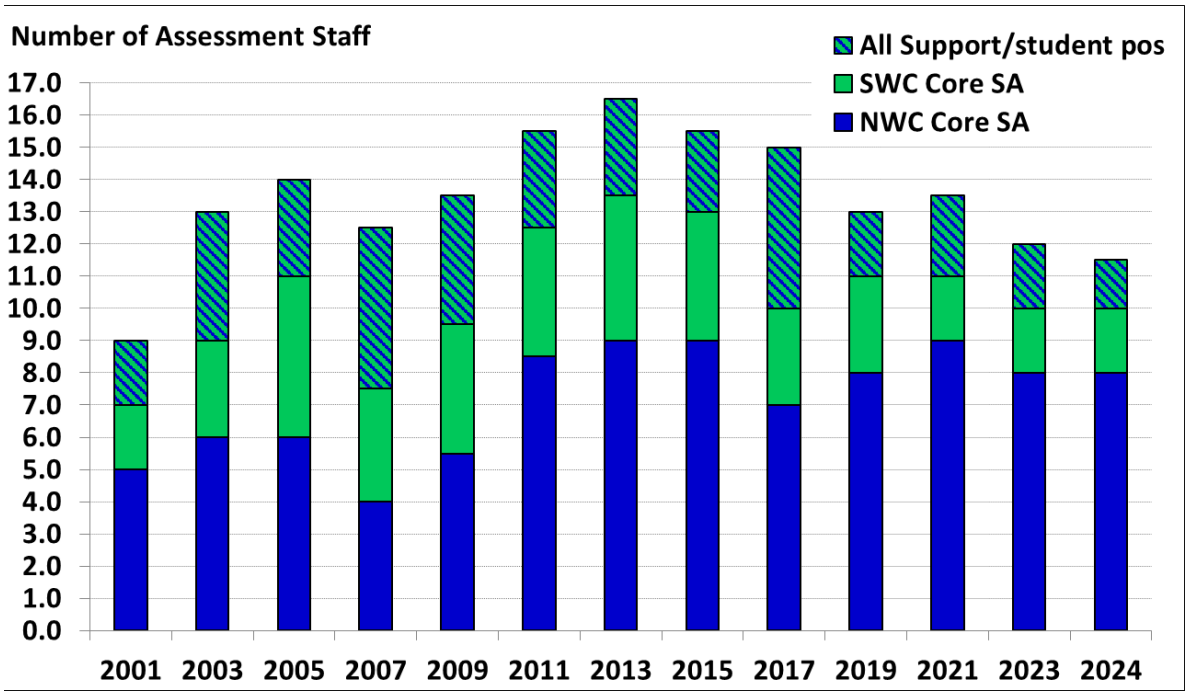

Figure 1.--Number of core groundfish stock assessment (SA) staff, and support or student positions, at the Northwest and Southwest Fishery Science Centers, 2001-2024. (Does not include supervisors who did not conduct assessments.)

Since 2013, however, the combined number of core positions and supporting staff has experienced a gradual 30% decline. For both Science Centers, combined, current core staffing is at its lowest point since 2009 and total core and support staffing is lower than at any point since 2001. Compounding the situation at the Science Centers is the fact that the process has lost numerous (~4-5) state agency positions since the 2000s that were involved in conducting, and frequently leading, groundfish stock assessments.

While assessment capacity since 2013 has ebbed, the 'demand' for assessments, as represented by the number of species (and modeled areas) for which we are trying to maintain benchmark, update, or data-moderate (e.g., Category 1 or 2) assessments, has increased (Figure 2). While the cumulative number of species with Category 1 or 2 assessments has increased by 14% since 2013 (teal line), the number of modeled areas (pink line) used in assessing those species has increased by 50%. The dashed blue line, which replicates the sum of core and support staff from Figure 1, illustrates how staffing and the number of supported species areas modeled increased together through 2013, but have diverged since then. In addition to implications for the ability of assessment staff to 'keep up', these increases in the numbers of species and modeled areas for which we are trying to maintain Category 1 or 2 assessments has also strained age-reading capacity, which has also declined slightly over most of the past decade.

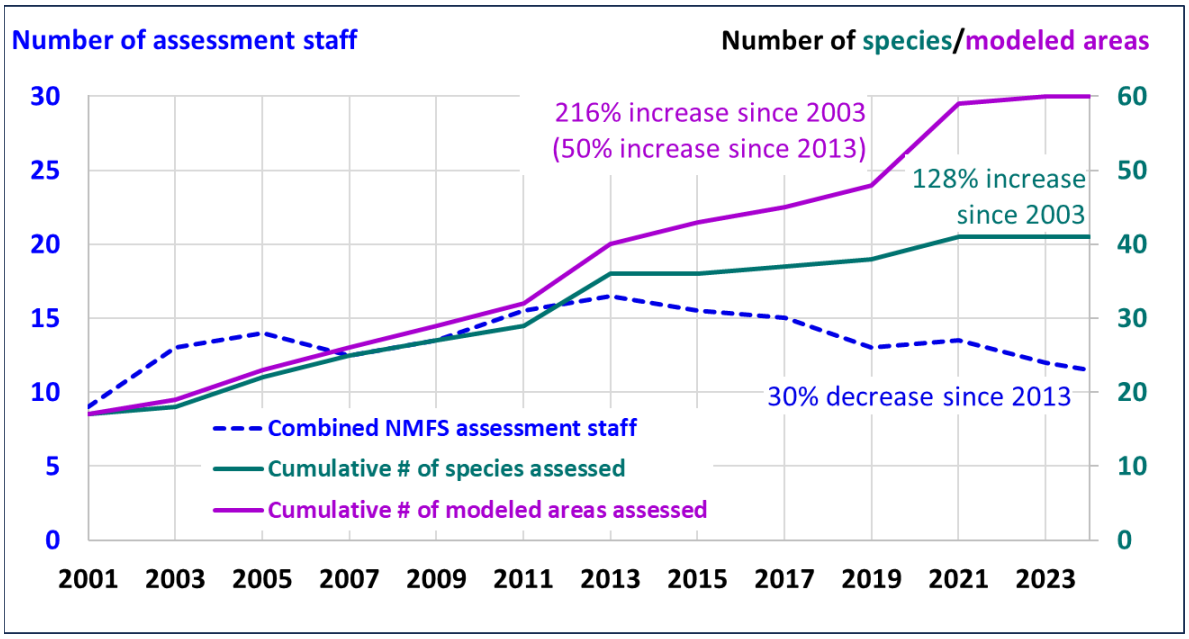

Figure 2.--Cumulative numbers of species assessed using benchmark or data-moderate methods, and the associated number of assessed areas, by 2-year cycle, 2001-2023.

How has the divergence between capacity and demand affected our ability to maintain target frequencies for assessed species? Table 3 summarizes the number of species that would have assessments no older than their target frequencies, as of 2025, based on the time since each was last assessed. For the 40 species with either Category 1 or 2 assessments, only 16 species (40%) would remain within their target frequencies if they are not assessed in 2025. The remaining 24 species (60%) would be beyond their target frequencies if not assessed in 2025, and 21 of those are already beyond theirs, following the 2023 cycle.

Table 3.--Status of the 40 species (excluding Pacific hake) with current management based on benchmark, update, or data-moderate methods (Category 1 and 2), with respect to their target assessment frequencies (calculated from 2025).

	Target assessment frequency (years)				Total Assessed Species	% of Overall Total
	4	6	8	10		
Number of species that:	13	7	9	11	**40**	**100%**
Would remain within target frequency if unassessed in 2025	1	4	5	6	**16**	**40.0%**
Are currently OK, but would exceed target frequency if unassessed in 2025	1	1	1	0	**3**	**7.5%**
Already exceed target frequency	11	2	3	5	**21**	**52.5%**

Table 4 reports the rate of assessment throughput that would be required to keep species assessed at the target frequencies presented in the upper section of Table 2. Using the most-recent number of modeled areas for each species, maintaining target frequencies for all 40 species would require that an average of 13 species (with 20 modeled areas) be assessed each 2-year cycle (Table 4). These amounts are roughly 50% higher than the average of the last four cycles, which are closer to the amount needed to maintain a 10-year frequency for all 40 species. It is worth noting that there are 25 other species included in the prioritization package. Selecting any of them for Category 1 or 2 assessments would increase the average assessment production rate needed to maintain target assessment frequencies.

Table 4.--Comparison of the number of species and area assessments that would be needed to maintain target and 10-year assessment frequencies for the 40 Category 1 and 2 species (excluding Pacific hake) with the average number conducted from 2017 through 2023.

Average number of species or modeled areas:	**Species**	**Modeled Areas**
- that would need to be assessed and reviewed every cycle to maintain **current** target assessment frequencies for all 40 species	13	20
- that were **actually assessed**, per cycle, during 2017-2023	9	13
- that would need to be assessed and reviewed every cycle to assess all 40 species **once every 10 years**	8	12

The numbers of species that have been assessed in each cycle are plotted in Figure 3, by type, along with a dashed line showing the average number that would be needed to achieve the 2024 target assessment frequencies (from Table 4). The most-recent year in which 13 different species were assessed was 2013, and assessments for 8 of those 16 employed index-only data-moderate methods and were reviewed in a single meeting by the Scientific and Statistical Committee's Groundfish Subcommittee.

The numbers of modeled areas included in each cycle's assessments are plotted in Figure 4, by type, along with a dashed line showing the average number that would be needed to achieve the 2024 target assessment frequencies (from Table 4). Developing assessment models for 20 species-areas falls well beyond the output of any past assessment cycle other than 2005 (after which, the phrase, "We will never do *that* again," was commonly heard). That level of output was also only achieved with the help of numerous assessment leads who were either state employees or full-time students, with four species reviewed in every STAR Panel under much simpler Terms of Reference (ToR).

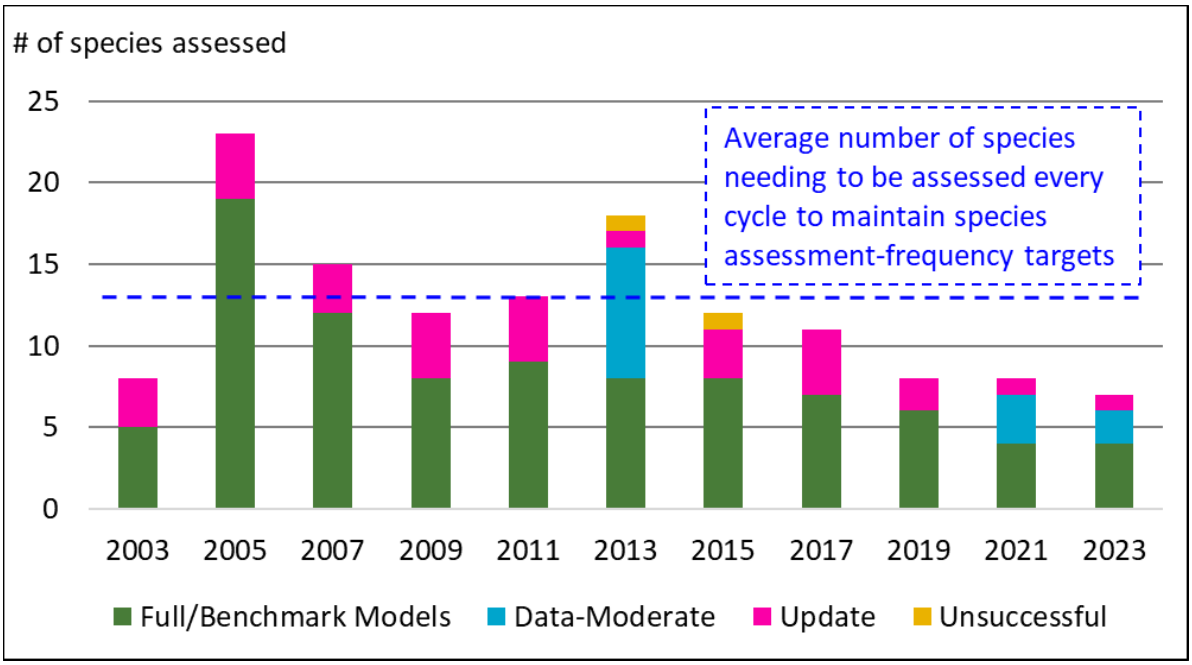

Figure 3.-- Number of species (excluding Pacific hake) assessed biennially, 2003-23, by type, with a reference line showing the average number of species per cycle that would be required to keep 40 species assessed at their target frequencies.

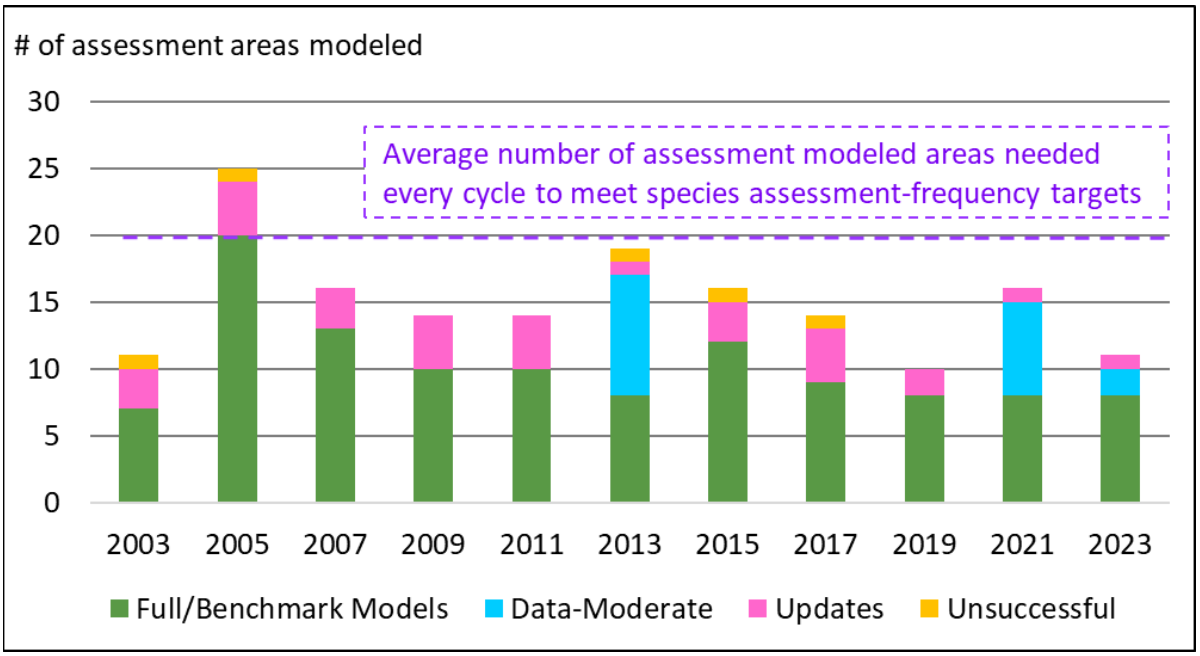

Figure 4.--Number of assessment areas (excluding Pacific hake) modeled biennially, 2003-23, by type; with a reference line showing the average number of modeled areas that would be required to keep 40 species assessed at their target frequencies, given the current number of areas modeled for each species.

Increasing Assessment Complexity

Stock assessments and the documents in which they are presented have increased in complexity and thoroughness over the past 20 years. A major contributing factor to fewer species having been assessed in recent cycles is that the average number of modeled areas per species has increased. This has largely resulted from the Council selecting a higher percentage of nearshore or shallow-shelf species, which are more likely to have population dynamics driven by smaller-scale regional processes (e.g., exploitation, environment, etc.). Around seven nearer-shore species have been assessed for the first time since 2013. Additionally, the number of modeled areas has also increased over time for some of those species and other nearshore ones that had been previously assessed. For example, black rockfish was assessed using two modeled areas in 2003 and 2007, with three areas in 2015, and with four areas in 2023. The assessment cycles from 2005 to 2013 included an average of 1.1 modeled areas per benchmark assessment. From 2015 to 2023, the average was 1.5 modeled areas per benchmark, with an average of 2.0 in each of the last two cycles (2021 and 2023).

As noted above, the current groundfish ToRs place greater demands on those developing assessments than was the case 10-15 years ago. For 2011, the ToR section containing instructions for the Stock Assessment Teams (STATs) contained 4 sentences; for 2023, it was 4 pages long. Appendix B, which provides an outline of all the elements that should be included in full or update assessments was 3 pages in 2011, and 6 pages in 2023. In addition to the increased time required for model development and more-exhaustive evaluation, the impacts of increased requirements and the average number of areas modeled is clearly evident in the page counts of the assessment documents. Table 5 provides a comparison of assessment page counts, by modeled areas and by species, between the 2011 and 2021 cycles. Of particular note is that the average number of pages per benchmark species in 2021 was nearly double the average from 10 years earlier. Although the 2021 data-moderate documents were only about 60% as long as 2011 benchmark documents, per area, on a per-species basis they were 57% longer. In addition to creating larger documents, that better describe models and explore areas of uncertainty, the emergence of requirements to create assessment documents that are compliant with Section 508 accessibility standards has further increased assessor's workload associated with creating assessment documents that commonly have hundreds of tables and figures.

Table 5.--Comparison of assessment document lengths in 2011 and 2021, by assessment type: benchmark, update, or data-moderate (Category 1 or 2) by species and modeled area (including Pacific hake). Page counts include text, tables, figures, and appendices, and exclude documentation of Stock Synthesis input file content.

	2011	2021	% Increase 2011 –> 2021	2023	% Increase 2011 –> 2023
Benchmark					
Number of species	10	5		5	
Number of modeled assessment areas	10	9		9	
Average number of pages, per area	215	230	7%	231	8%
Average number of pages, per species	215	417	94%	416	94%
Data-moderate					
Number of species		3		2	
Number of assessments		8		2	
Average number of pages, per area		126		134	
Average number of pages, per species		337		134	
Update					
Number of species	3	1		1	
Number of assessments	3	1		1	
Average number of pages, per area	100	122	22%	145	45%
Average number of pages, per species	100	122	22%	145	45%
Total					
Number of species	13	9		8	
Number of modeled assessment areas	13	18		12	
Average number of pages, per area	188	178	-5%	208	11%
Average number of pages, per species	188	356	89%	312	66%
Total number of pages	2,447	3,206	31%	2,493	2%

Conclusion

Over the past two decades, this Council has made enviable strides in improving the rigor of its groundfish assessment process, including its reliance upon independent peer review, continuous review and refinement of its groundfish Terms of Reference, extensive vetting of new methods through SSC-led methodology reviews, and use of the Assessment Prioritization package for guiding consideration and selection of species to be assessed. Over time, however, 1) the number of groundfish species for which Category 1 or 2 assessments are desired for management, 2) the assessment development and review workload that is associated with assessing these species (as prescribed by the Fishery Management Plan, the Groundfish Assessment Terms of Reference, and tradition), and 3) the number of stock assessment scientists available to conduct Category 1 or 2 assessments have reduced the Agency's ability to deliver assessments for management use on a schedule that is consistent with identified target assessment frequencies.

Without attention and some manner of intentional action, assessments for some species will inevitably be determined by the SSC to be too old for use by management. Ten years used to be the cutoff for use in management, but that was extended following adoption of a scientific uncertainty buffer that increases with assessment age. However, the Council and SSC have not since articulated a clear process for reviewing and determining when an assessment, even with subsequent catch-only projections, should be considered too old for continued use in setting harvest specifications. The '40-10' rule for setting Annual Catch Limits is an integral part of the harvest policy for stocks with Category 1 or 2 assessments, and its application requires the availability of reliable, contemporary stock status information. Although catch-only projections can help address the impacts of short-term differences between projected an actual catches, they are not intended to provide reliable updating of stock status, particularly when the previous assessment is more than 10 years old. There are currently nine (of 40) species whose assessments are more than 10 years old, and three more will join that group if not assessed in 2025.

The suite of processes that comprise species selection, assessment development and review, and use of assessment results could benefit from a comprehensive joint Council-NMFS review that considers 1) the current throughput capacity for generating new science to inform management, 2) possible ways for increasing the rate of throughput, and 3) possible explicit decisions about which species may need to revert to Category 3 assessments for purposes of informing harvest specifications. In the absence of such a review and responsive correction action, assessments for some species will likely become too old for management use, by default, rather than as a result of deliberate planning that considers the full range of factors driving this predicament. Species that cannot be reassessed with adequate frequency will revert to Category 3 assessments, which will provide less-certain estimates of overfishing levels, along with having larger uncertainty buffers, and provide minimal information about stock status.

References

Methot Jr., Richard D. (editor). 2015. Prioritizing fish stock assessments. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-152, 31 p.

NMFS. 2001. Marine Fisheries Stock Assessment Improvement Plan. Report of the National Marine Fisheries Service National Task Force for Improving Fish Stock Assessments. U.S. Dep. Commerce, NOAA Tech. Memo. NMFS-F/SPO-56, 69 p., 25 appendices.