Center for Independent Experts (CIE) Program
External Independent Peer Review

Remotely Operated Vehicle (ROV) Surveys of Nearshore Stocks California & Oregon

PANEL REVIEW

Alan Williams

February 2020

Executive Summary

The three-day review panel meeting was well-organised and conducted in a highly constructive manner. It enabled key issues to be identified and discussed, provided the opportunity to summarise, package and present the ROV data in a stock assessment context, and was effective in bringing the relevant researchers together – including those from different agencies within individual states. All relevant aspects of the ROV programs were discussed, including revisions based upon feedback from the desk review and additional work done since.

The technical merits and deficiencies of the survey designs and analytical methods were discussed in detail, mostly during the first two days. This report provides summaries and assessments in Section 3.1. It groups the many technical topics in relation to: survey design; assessing the wide range of species for their individual suitability; quantifying the ROV fish counts into density data and exploring the best spatial scales at which to represent these data; characteristics of the seabed mapping data that underpin the expansion of data collected along transects to broader (site and statewide) scales; the conversion of fish counts to biomass; strategies for acquiring mapping data for unmapped areas; and, briefly, future capital investment in ROV infrastructure. Day three concentrated more on the many issues involved in data expansion to broader (site and statewide) scales, including the characteristics and assumptions of models used for the expansion. Both states explored tools and techniques to unpack spatial auto-correlation effects on correlations of fish density with covariates, and on estimates of variance.

Summary of finding

My summary of findings is presented here as the list of suggestions made for extending, standardizing and evolving the work program. Each is detailed in the sections indicated.

<u>Improvements to existing methods (short term)</u>

- Summarise species suitability for ROV surveys [Section 3.2.3] to assist all stakeholders to understand the potential for incorporating ROV data into stock assessments.
- Harmonise the data analysis strategy [Section 3.4.1] using a focussed workshop and data exchange between states to advance specific areas of analysis.
- Evaluate and compare seabed classifications [Section 3.2.5] between states with validation ('cross-walking') of data and methodologies to ensure the mapping data used are reality-checked with ROV observations, comparable and used consistently.

- Review the needs and options for spatial analysis [Section 3.4.1] because this is incomplete. (The Oregon analysis using VAST was an important and encouraging step towards rigorously analyzing how relevant is spatial autocorrelation to variance in expansions of ROV data.)
- Standardise the spatial units of fish count analysis [Section 3.2.6] to identify a single or perhaps complementary pair of scales (segment and transect) to align scale-dependent uncertainties and trade-offs in data processing, and applications, and to enable direct comparison of data.
- Better define the utility of ROV data for inclusion in stock assessments [Section 3.4.1] to assist stakeholders understand the data and facilitate its uptake.
- Develop standard (convergent) operating procedures [Section 3.2.2] for state-specific procedures and data processing protocols, to the extent possible, to enable data integration, now and in the future.
- Refine and standardize methods for converting fish counts to biomass [Section 3.2.7] using stereo imagery and avoiding reliance on co-located proxy (catch) data.

Strategic improvements (long term)

- Implement future surveys based on fishery-specific objectives [Section 3.2.1] so that data have planned replication in space and time, priority gaps in data are filled, and minimise variance in state-wide expansion data through targeted sampling of key areas. New surveys could also consider the utility of spatially-balanced design strategies that emphasise locations of higher importance and incorporate legacy (index) sites. Model validation (below) would also be feasible.
- Validate models by future sampling [Section 3.4.1] of density estimates across methods at site scale using fishery dependent data; across the gradient of predicted fish densities; and to determine fish densities on soft and mixed (non-hard) seabed types.
- Develop a strategy to fill gaps in mapping data [Section 3.2.8] to support more complete future statewide expansions of fish density data.
- Evaluate the needs and cost-benefit for future ROV infrastructure [Section 3.2.4] to consider upgrades (e.g. to position fixing accuracy) and to ensure continuation of this program of work beyond the shelf-life of the vehicles available currently.
- Plan strategically for a centralized ROV data repository [Section 3.2.9] to maximise the likelihood and benefit of re-using data and be compliant with the FAIR principles for data management.

Recommendation

Create a formal project with adequate resources to foster development of this work program. This high-quality science effort has the potential to provide a robust source of data to feed stock assessment of selected commercially and recreationally important inshore rocky reef fish species. No alternative sources of data exist. Draw on the considerable efforts of the science teams involved, and this review process, to establish tactical and strategic workplans.

Specify whether the science reviewed is the best scientific information available

The science presented was of a very high standard. The knowledge, rigor and innovation applied to the data acquisition, data analysis and data products clearly demonstrated the work represents the best information available at this point. The review has identified many areas where the science can be extended in the future and evolved in ways that will improve its ability to meet the demands of the stock assessment process.

Table of contents

1 Background 7

2 Individual Reviewer's Role in Review Activities 9

3 Summary of Findings 10

3.1 Discussion of technical merits and deficiencies of the survey designs and analytical methods (TOR 2) 10

3.1.1 Survey design 10

3.1.2 Field of view, height off bottom, transect width 13

3.1.3 Suitable species for assessment 14

3.1.4 ROV position fixing (and future infrastructure) 15

3.1.5 Seabed data: substratum classification and terrain variables 17

3.1.6 Spatial unit of analysis: transects and segmentation 19

3.1.7 Fish sizing and length-weight conversions 21

3.1.8 Seabed data: missing data and gap-filling 22

3.1.9 Data management 23

3.2 Suggestions for improvements to survey designs and analytical methods (TOR 3) 24

3.2.1 Implement future surveys based on fishery-specific objectives 24

3.2.2 Develop standard (convergent) operating procedures (SOPs) 25

3.2.3 Summarise species suitability for ROV surveys 25

3.2.4 Evaluate needs and cost-benefit for future ROV infrastructure 25

3.2.5 Evaluate and compare seabed classifications 26

3.2.6 Standardise the spatial unit of fish count analysis 26

3.2.7 Refine and standardize methods for converting fish counts to biomass 26

3.2.8 Develop a strategy to fill key gaps in mapping data 27

3.2.9 Plan strategically for a centralized ROV data repository 27

3.3 Evaluate model assumptions, estimates, and major sources of uncertainty (TOR 4) 27

3.3.1 General 27

3.3.2 Model selection, spatial expansions and spatial autocorrelation 28

3.3.2.1 CALIFORNIA 28

3.3.2.2 OREGON 33

3.3.3 Summary 37

3.4 Suggestions for improvements to models and spatial expansion 39

3.4.1 Harmonise the data analysis strategy 39

3.4.2 Validate models by future sampling 40

3.4.3 Review the needs and options for spatial analysis 41

3.4.4 Better define the utility of ROV data in stock assessments 41

3.5 Specific suggestions for short-term and longer-term improvements (TOR 5) 42

3.5.1 Improvements to existing methods (short term) 42
3.5.2 Strategic improvements (long term) 42
3.6 Prepare a Peer Review Report following the Terms of Reference (TOR 6) 42
3.7 Provide a brief description of panel review proceedings (TOR 7) 42
4 Conclusions 44
5 Recommendation 44
6 Appendices 45
6.1 Appendix 1: Bibliography of materials provided for review 45
6.2 Appendix 2: A copy of the CIE Performance Work Statement 46
6.3 Appendix 2: Panel review agenda 54

1 Background

The extensive nearshore marine areas of California and Oregon support a diverse groundfish fauna targeted by commercial and recreational fishers. The associated fishery stock assessment process is challenging because multiple species need attention, catch is taken by multiple fishing methods, and there is high spatial and temporal structure in the fishery regions and patterns of fishing. As a result, catch data acquisition, catch-effort standardisations, and catch history reconstructions are complex, and exploitation status incompletely understood. In addition, many relevant aspects of species' ecology are not known. Against this background, a need for more fishery-independent data sources has been identified over a number of years.

One type of prospectively important fishery-independent data are from visual census using camera platforms that can operate below diver depths (>20 m). Visual surveys have been considered for at least a decade. Camera-based methods have two major advantages over more traditional net-based (trawl) survey sampling techniques: they are non-extractive, and can take data over all seabed types. As such, camera-based methods have the potential to collect fishery-relevant data for rocky-habitat associated species, explicitly account for abundance inside no-take Marine Protected Areas (MPAs), and provide data on overfished species and nearshore species which constrain take of healthy stocks. An added advantage is to avoid the need for research set-asides or other allocative considerations that may arise between fisheries and research sectors.

Remotely operated vehicles (ROVs) provide a potentially effective sampling tool, and both Oregon and California have conducted ROV surveys of rockfish in nearshore areas. Both states have invested in developing seafloor maps which permit estimation of area of habitat types by depth and latitude. This has enabled surveys to focus on fishes associated with rocky reef habitat, and on areas inside and outside of MPAs. Data acquisition and analysis has focused on generating quantitative enumeration of specific species, and developing techniques (including models) to generate regional-scale species/population density estimates. These estimates – as indices of relative abundance or estimates of absolute abundance in the depth and latitudinal areas surveyed – have the potential to be used in stock assessments if there is appropriate accounting for uncertainties in the data, e.g. selectivity and detection probability. There is also the potential to acquire length composition data and habitat ground-truthing data during ROV surveys that may also be relevant to stock assessments and/or management procedures.

The goals and objectives specific to the review of the California and Oregon ROV survey methodologies are to:

1) Evaluate the sampling design used in recent ROV surveys conducted by the states of Oregon and California.
2) Evaluate proposed methods to develop indices or estimates of abundance for these ROV surveys, including using habitat/substrate type and Marine Protected Area designation as covariates.
3) Evaluate proposed methods to estimate size compositions of observed individuals of each species.
4) Identify potential impediments to developing independent indices or estimates of abundance using these ROV surveys and incorporating them into stock assessments.

2 Individual Reviewer's Role in Review Activities

Prior to the panel review I prepared a desktop review in accordance with CIE guidelines. That written review was my first contact with the research programs developing methods to incorporate ROV observations in assessments of stocks of inshore fishes off the coasts of California (CA) and Oregon OR).

My contribution to the review process is based on my working knowledge of visual survey techniques, survey design and analysis, and familiarity with incorporating survey information in stock assessments; these areas are emphasized in my review. I am not a stock assessment scientist (modeler), and therefore my comments on model evaluations and developments are somewhat limited. There are substantially better-informed views contributed to the panel report by other members of the review panel and workshop audience.

I read both the original California and Oregon reports in detail, and the revised reports together with other pertinent information prior to review panel meeting (TOR 1); I participated actively in the full three days of panel review and preparation of the Summary Report.

This report refers to additional reports and papers that provide relevant information (Appendix 1). My review is impartial and independent in accordance with the PWS, OMB guidelines, and the ToRs. As requested I have:

- described in my own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

- discussed my independent view on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

- elaborated on any points raised in the Summary Report that I feel might require further clarification.

3 Summary of Findings

3.1 Discussion of technical merits and deficiencies of the survey designs and analytical methods (TOR 2)

The technical merits and deficiencies of the survey designs and analytical methods were discussed during the open review panel meeting, including revisions based upon feedback from the desk review. The panel meeting considered the generation of ROV video data for assessing abundance of rocky reef fishes – technical aspects, survey design, and much of the analytical methods – during days 1 and 2. Presentations were made on each of the following topics by CA and OR:

- ROV system/ capabilities/ configuration
- Survey design
- Data Processing (transect-level)
- Habitat information
- Transect segmentation – relationship to habitat
- Abundance indices
- Density estimation
- Abundance Estimation
- Length data and conversion of numbers to weight

Each of the presentations included responses to issues raised during the desktop review and reported on additional work done since. In fact, a considerable amount of additional work had been completed by OR since the desktop review. Collectively, the presentations addressed the key needs of the review, and all were of a very high standard.

Below I have recorded what I believe were the key technical aspects of the discussions in the context of assessing the suitability of the ROV for incorporation in stock assessments.

3.1.1 Survey design

In overview, the ROV sampling designs employed in each state are designed to enumerate fishes along strip transects of (mostly) 500 m length targeted at inshore areas of rocky reef (mostly <70 m depth). Many sites have been surveyed in both states, and collectively they appear to well-represent the fishery area and some marine parks. The sampling designs of both programs are appropriate for data expansion because they target reefs delineated by pre-existing mapping data using randomly allocated transects of suitable length within a specified depth range. The designs are generally similar:

- CA survey sites are defined by rectangular boundaries, 500 m wide and extending from shallow to deep, orthogonal to isobaths, for as far as is necessary to reach the seaward limit of the hard ground. Up to 4 km of transect track is covered per site by individual transects that extend across the 500 m width of each site, and which are near depth-parallel. The location of each set of transects within a site is determined by randomly seeding the start point of the shallowest transect and then spacing the remainder transects at regular intervals parallel to the first.

- OR 'survey areas' are located in sub-sections of the coast around rocky reefs, where there is mapping data, and/or other research interests that enable field work opportunities. Surveys randomly identify a sub-set of transects from a larger number of candidates defined within 500 m wide, depth-parallel planning swaths (after eliminating unsuitable transects using decision rules). Bounding boxes define areas using the 20 and 70 m isobaths; in many cases, but not all, 70 m includes the deep limits of the reefs.

I was surprised to learn that the ROV data being considered for stock assessment purposes have been generated from leverage and piggybacking on other projects which have different objectives. These were typically surveys of marine reserves (comparing biological trends inside vs outside), although others had a focus on habitat definition or comparing gears and methods. The obvious consequence is that these designs will not be optimised for the fishery stock assessment application being considered here.

This is neither a criticism of the research efforts to date, or on the veracity of the results presented. It is a comment on the future need for a more strategically planned sampling program. To date, on the one hand it was probably inevitable that a scattershot approach was required to implement surveys over these long coastlines to gather data on rockfishes, develop analysis methods, and establish proof-of-concept for the utility of ROV data for stock assessment. On the other hand, a scattershot of data is more difficult to understand and will have higher variances for predictions due to the lack of planned replication in space and time, and because there are gaps in data.

The summary of ROV sampling effort in Oregon (OR report, Table 2) captures this nicely by showing the distribution of samples over the last decade. There is an impressive tally of samples from four primary sub-regions of the coast, but also large between-year differences in sample effort, and large between-area contrasts in total samples, e.g. low in North Coast, Cape Perpetua and Cape Arago and relatively high at Port Orford, and corresponding contrasts in opportunities for time-series analysis. As a consequence, state-wide expansions across areas (sites) are required to integrate data over years with little scope for temporal effects, including fish catch though time or a 'reserve effect'.

These data are sufficient to estimate local stock relative abundances and suited to explore state-wide expansions, but for more routine and broad-ranging uptake in the stock assessment process it is highly desirable to ensure that future surveys have the ability to meet fishery-related objectives. So, for example, any future survey in the Cape Arago area would sample the extensive deeper (>90 m) reefs to fill this data gap for fishery purposes. But as well, there is a need to invest survey effort into soft habitats because, even though they may support low fish density, they are large and need to be accounted for in fish density expansions. It was noted that data collected to meet other objectives may be inadequate for post-stratification in the fishery context if it generates small samples in some strata, and strata in which samples are no longer random. One imagines that in many instances samples meeting the needs of fishery objectives will also be a value-adding opportunity for marine reserve managers. Whilst 'gap filling' will not be an onerous planning process in many locations, there are locations that are difficult to access for reasons of seasonal sea conditions and/or tide, and that sampling these needs very precise timing. Fishery data acquisition in such locations is more likely to be successful if part of a strategically-planned program.

New surveys could also consider the utility of sampling design strategies to maximise the value of each sample such as those developed by Foster et al. (2017; 2020). Their methods enable transect designs to be generated using randomization where the probability of sampling each cell in a spatial grid is user-defined (the cell inclusion probabilities) to emphasise locations of higher importance, e.g. areas with high biomass or high variability. Part of this process is to assess the trade-offs between repeat sampling (e.g. index sites) vs sampling new areas (e.g. the latter seemed more important to Oregon). Foster et al. (2017) provided a method for considering legacy (index) sites in designs.

My summary recommendations are for standardization, or at least convergence, of designs to the extent possible to minimise uncertainties when data are shared. Data acquisition has to be underpinned by strategic planning that takes account of fishery objectives – and is supported by dedicated funding.

3.1.2 Field of view, height off bottom, transect width

The main purposes of these discussions were (1) to establish the ways in the species-count data were collected and allocated to an area swept on the seabed to provide fish density estimates; (2) to understand any potential biases or inconsistencies in the methods; and (3) to consider the comparability of methods between CA and OR.

The presentations highlighted that a lot of Oregon's set up had been done with close consultation with CA. Thus, the underlying methods were more similar than was evident from comparing the original reports. Most obviously, CA had not previously mentioned their use of artificial light which had appeared to be a fundamental difference in methods. But, in short, there were no technical issues that either compromised the estimation of fish density or that prevented direct comparison of the CA and OR estimates. The methodologies discussed in detail and/or that addressed by feedback in response to the desktop reviews were about estimating field-of-view (FOV), height-off-bottom (HOB), and transect width (TW).

Although slightly different methods were used to estimate FOV – diminishing perspective overlay (CA) and trapezoid overlay (OR) – both were well calibrated and the resulting estimates of the bottom area in which fish were counted were robust. The detail provided by the presentations substantiated that estimates are robust and comparable.

Because the ROV was generally flown higher off bottom by CA than OR, the effect is for CA to observe a wider area of seabed. This has potentially important implications for estimating fish density if there is a bias in fish counts related to either HOV or transect width. This is possible if fewer individuals of flighty species are observed when the ROV is closer to the bottom or fewer individual of cryptic species seen when it is high off the bottom or when water clarity is lower. More fundamentally, because the fish density may be estimated from fish counts within transect segments of fixed length (20 m), then lower density estimates will arise from 20 m segments that are relatively wide. It was also noted that transect area will be affected by smoothing the tracking (USBL) data as this will, typically, reduce the estimate of transect length and therefore area swept. Because fish count is not affected, smoothing may therefore systematically elevate fish density. However, there was strong awareness of all these issues by teams – including having run analytical diagnostics to understand their effects. I believe the methods employed to estimate fish density are not adversely impacted by FOV, HOB or TW effects – and note that potential flow-on, e.g. from using fixed transect length with vary area, is explicitly accounted for in expansion models. I make more notes about this is subsequent sections.

There was quite a bit of discussion about these possible effects – particularly when (1) the ROV is deliberately flown with a great HOB over highly rugose seabed, (2) for 'backsides' when the ROV loses vision of the seabed as it descends after an ascent over high relief, and (3) when the ROV is pitching and creates variable field of view which may be relatively very wide at times. Both programs have protocols, filters and cutoffs to limit any potential biases from these effects by excluding data, but the tradeoffs are to create more gaps and increase fragmentation of transects, and lose data over rugose seabed that rocky reef species are associated with. On balance, I am satisfied that these issues have been thought through carefully and that a pragmatic balance of operating procedures and data processing protocols have been established in both programs that minimise data loss and prevent systematic bias. I think there will be considerable benefit for future work – including for data amalgamation across states – to carefully document operating procedures and standardize these to the extent possible. The data reports for this review process provide a very large step in this direction.

3.1.3 Suitable species for assessment

Many rocky reef species are observed by ROVs, but they have different behavioural characteristics or ecological traits that make them more or less suited to being consistently detected and counted along transects. Unsuitable species are those, for example, that school off the bottom, have cryptic colouration, occur at very low abundance, or that have a marked flight response to disturbance by the ROV. The issue of which species are amenable to consistent detection, within and across programs, was a topic of discussion at many points during the review. There was not a formal list of target species with descriptions of behaviours and traits in the pre-meeting reports or generated during the meeting, but it was clear that there was general agreement on which species were suited to ROV abundance estimation and that there were robust rationales for these. Thus, suitable traits or behaviours included being highly sedentary, exhibiting very little behavioural response to the presence of the ROV, and there being no recognised difference for detection probability between conditions of turbidity and clarity. Suitable species included Vermilion, Copper, Quillback, Yelloweye, and Gopher rockfishes. Less suitable species included those that school off bottom – Blue, Black, Yellowtail, Bocaccio rockfishes – and cryptic or difficult to identify to species – Cabezon, Treefish, various flatfishes. The field scientists with considerable experience of viewing video during ROV operations were confident that behavioural responses to the ROV were not an issue – in part because the ROV can be operated quietly and skillfully by the pilots. These observations are consistent with published literature, including a 2019 review, that were assessed by the CA and OR teams. There is a detailed note about ROV selectivity in the CA response to the desktop review (CA report, pp. 28-29).

CA provided a list of species with potential for generating abundance indices and estimates (their Table 2). I suggest this is developed as a formal list of species suited to quantitative observation by ROV, with rationales, to assist in identifying the opportunities to deploy ROV data to stock assessment and conveying this to third parties. Not all candidate species are suited to analysis across the collective OR, CA and broader coastline because there are latitudinal and local scale patterns in abundance and/or size. There was a feeling in the meeting that it will be worthwhile for exploratory data analysis and to prioritise future work, to select some species that are hard on rock and differentiate those that are also abundant on soft sediments or that make big excursions (e.g. Quillback).

In summary, to assist all stakeholders to understand the potential for incorporating ROV data into stock assessments, I suggest it will be useful to construct a tabulated list of candidate species, integrated across states. This should show behavioral and ecological traits, geographical patterns and variations (e.g. state-specific depth patterns), overall tractability for expansions at particular scales (e.g. local reef areas vs population scale), and rationale for prioritization against (1) the need both to refine or complete analysis of current data, and (2) for the needs of stock assessment processes. The formal process to list/ rank priority species for stock assessments would seem to be an ideal target for a ROV-focused summary.

3.1.4 ROV position fixing (and future infrastructure)

Knowing the position of the ROV is important so that fish count data along transects can be overlaid on maps of seabed habitat or terrain types with known accuracy. The relationship between local-scale fish density and seabed type is the basis for spatially expanding the data to broader areas where seabed type is known or predicted from multibeam sonar (MBS) mapping and/ or geological mapping and interpolation.

Accordingly, there was a lot of discussion about the accuracy of acoustically-derived position fixing estimates (~±4-6 m in both programs), and ways to improve them. Improvements can potentially come from two approaches: first by refining the process to smooth the currently collected USBL tracking data which is often noisy with many spurious position fixes along a transect, or secondly, using an improving tracking sensor system – a Doppler Velocity Logger (DVL) with inertial measurement – instead of a USBL (see Trembanis desktop report, p. 16).

Prospectively, it may be possible to improve the processing of currently collected USBL data using different algorithms. This is a specialist field outside my expertise, but this could be followed up. However, other variables that affect position fixing such as the quality of vessel position data, and varying structure in the water column affecting speed of sound may also introduce errors that are not accounted for. Improvements are undoubtably possible using a DVL system, but as noted by the CA response, this would incur considerable additional cost and processing complexity and could only be justified by a cost-benefit analysis of the effect on the final error estimates in fish abundances. These considerations are part of a larger question related to the replacement and maintenance of vehicles.

However, any potential benefit from improved position fixing is strongly linked to the spatial scale of analysis: the unit of analysis of fish density (e.g. a segment vs whole transect), and the spatial scale and accuracy of seabed data being correlated with fish counts. Thus, improved position fixing may have benefits if data are analysed at a relatively fine (segment) scale (e.g. 20 m length as in much of the exploratory analysis completed to date) and seabed data are terrains defined at relatively fine scale (small neighborhoods of 4 m^2 MBS grid cells). Improved position fixing is less relevant when seabed data exist as relatively coarse scale (e.g. ≥30 x 30 m grid based) substratum polygons, where confidence in substratum classifications is low or uncertain, or when terrain attributes are defined by relatively large neighborhoods (30 x 30 m) of MBS grid cells.

In summary, position-fixing of ROV transects should be considered; greater accuracy may be possible by improving the processing of currently collected USBL data, and would be improved using a DVL system. However, this would need a cost-benefit analysis to determine the effect on the final error estimates in fish abundances – which in turn will depend on the final methods employed for data expansions. Any considerations are part of a larger question related to the maintenance and replacement of vehicles. Any strategic development of this program should evaluate and document the needs for capital funding to maintain/upgrade, replace, or identify alternative suitable rental vehicles, for the ROVs presently being used by CA and OR.

3.1.5 *Seabed data: substratum classification and terrain variables*

Seabed type is a key element of the data required to turn fish counts along transects into estimates of fish abundance over broader areas. Thus, counts (densities) are expanded by determining the relationships between fish density and seabed type at small spatial scales (segment or whole transect) and applying these to broad areas that are either mapped by MBS or classified by geology. Seabed type may be defined in MBS mapping as a 'terrain' by properties including slope, rugosity and aspect, or as an area of a particular substratum type, where the types are pre-defined by geological or geophysical schema.

The methods used to upscale data on seabed type were different between states, and, for me at least, one of the most difficult aspects of the analysis to grasp. So, despite the importance of data expansion for final abundance estimates, within and across states, I found it difficult to track the detail of methods used to initially determine the 'fish-seabed type relationships' and whether or not the methods were comparable. Thus:

- Neither state used a geological classification that exactly matches the standard CMECS (although this may not be important)
- CA used repeat observations of the presence of simple substratum categories to generate a layered habitat characterization of simple soft, mixed and hard habitat types. This was used to generate variables (proportions of hard and mixed bottom types per transect) for the post-stratification of transects and to use in GLMs to identify informative variables for correlating fish density with habitat. GLMs were also used to identify correlations of fish density with seabed terrain type using attributes derived from neighborhoods of 2 x 2 m^2 cells in MBS data and linking these to transect segments. But because MBS data did not exist for all segments, this terrain-based analysis was for a smaller overall area.

- OR used a stepwise-lumping approach to derive two classification options for seabed analysis: (1) rocky reef (all hard substrates sized cobble and larger), and non-reef (small mix and soft) within a buffer zone, and (2) reef only. First, the SGH (Surficial Geologic Habitat) substrate classification was developed from multibeam bathymetry and backscatter data as primary ("Lith1") and secondary ("Lith2") substrates (presumably a % cover estimate), which are defined at different resolutions and with varying data quality, and then mapped as polygons. These are combined into a single mixed classification (“Lith3”) with a multitude of possible combinations. These combinations were reduced and then lumped into habitats known to occur at study sites and a knowledge of fish habitat associations (presumably based on ROV observations). Categories of Lith3 were then combined into six main super-categories termed “UpLith” (rock, rock mix, boulder, large mix, cobble, small mix, soft), which were finally split and mapped as rocky reef and non-reef polygons. OR explored the use of terrain attributes but didn’t persist with them as analysis was eventually completed at a coarser (transect scale).

The geology-based methods provide a generally similar classification. However, the degree of similarity and the ways in which any differences might contribute to variability in estimates of density and abundance between states is difficult to determine. But is this important? The answer lies in several considerations: (1) whether geology is used as the basis for expansion (presently it isn’t for the CA method); (2) the degree of lumping of the classification where there is likely decreasing sensitivity as the degree of lumping increases (e.g. the highly lumped ‘rock vs soft’, or ‘rock only’ classification used by OR will be rather insensitive); (3) the spatial scale of the unit of analysis where there is likely decreasing sensitivity as the spatial unit increases (e.g. from 20 m segment to whole transect); and (4) whether there is an intention or need to merge or directly compare the data between states.

These differences in underlying seabed definitions are important because expansions are sensitive to them. For example, CA expansions will be highly affected by definition of hard bottom which could be using a 2 x 2 m grid scale (cells either hard or not) compared to a 30 x 30 m grid where a neighbourhood which has only to include >10% hard. Because the latter includes more seabed there is then a need to account for the non-hard cells that are missed by the ‘hard only’ approach. These cells probably contain fish at relatively very low density, but they still will add to the estimate – especially where they represent large areas. A cited example is Quillback when juveniles are abundant on isolated cobble – but because habitat is impossible to represent at this scale a combined ROV and trawl survey may be needed. There was discussion about considering a value for the soft bottom in the unfiltered analysis, but this adds additional uncertainty. Another option to deal with this include either including cells within a proximity (although there is now an available reef definition for the CA data, or to apply a more judicious threshold to the cut-off, i.e. higher than >10%.

In summary, I have some uncertainty about whether the same seabed types are represented in the same way across states because the data sourced, mapping methods, classifications and degree of lumping of categories differ. Treating substrate types more consistently as a continuous variable had been considered but it generated data artefacts. There is also the question (raised by another panel member) of "how believable is the mapped habitat data, and do ROV observations match the habitat maps?" Thus, a recommendation would be to consider a validation ('cross-walk') exercise that completes a two-step comparison: first, use ROV observations to do a reality check on geological polygons, and second, implement both ROV data scoring methods on an example dataset from one or other, or both, states.

3.1.6 Spatial unit of analysis: transects and segmentation

Identifying the most appropriate spatial unit(s) of analysis was one of the key discussion points throughout the meeting. Fish and seabed habitat data may be analysed at the scale of entire individual ROV transects; data aggregated at this scale is similar to the integrated samples provided by trawls. The transect scale can also be refined by filtering to remove the sections that are not rocky bottom by overlaying them on the substratum polygons provided by geological classifications. There is opportunity to sub-divide the data taken along transects into segments to capture habitat, substratum and biological variation occurring at finer spatial scales (and to capture variation in operational parameters, mostly notably area swept). Segmentation is possible because data are recorded continuously (at 1-sec intervals) along transects.

There are many strategies for choosing the spatial units of analysis, and the advantages and disadvantages for ROV data were well reviewed previously: in both the California and Oregon reports, my desktop review (from p. 24), and the presentations at the panel review. In brief, each state had explored more than one scale in the work completed before the panel review, but ultimately took different directions:

- CA used a relatively fine-scale unit: a segment of fixed length (20 m) but variable width – which resulted in a degree of variability in the area swept per unit. This approach was strongly driven by the choice of terrain attributes (defined by neighborhoods of adjacent grid cells of MBS bathymetry data) as the basis for correlating fish counts to seabed types.

- OR explored segment-scale data at study site (area) scale, including to a fixed-area strategy (summing transect until segment size was 200 m^2, but used transect scale for the coastwide analysis. This was a contrast to the CA method. Segment-scale (20 m) data were used to explore substrate-specific fish density but this showed strong relationships between segment length and fish density for many species. The transect scale approach eliminated artifacts associated with small segments, reduced the volume of lost data when gaps produced segments below the threshold length, and reduced spatial auto-correlation of observations along transects. OR also used filtering to remove the sections that are not rocky bottom when exploring the expansion methodology.

Whilst a finer-scale (segment-based) approach initially seemed attractive (to me at least) to capture the fine spatial scale of fish-habitat relationships that can exist, there was no compelling evidence for strength or consistency of correlations between fish species and terrain attributes in the CA analysis. This was also the case in the exploratory work with terrains conducted by OR. This may be explained by the fact that even whilst there may be fine-scale relationships between individuals/species and particular geology and geomorphology (terrain), these are hidden by larger scale effects such as currents and food availability and their interactions with terrain at various scales. The OR Cape Perpetua site was given as an example where rocky habitat is limited but fish abundance relatively high; this was interpreted as an island refuge effect. This points to the potential benefit from a landscape ecology approach using species distribution models, however, experience with predicting changes in rockfish distribution and abundance in Puget Sound at hot spots where patterns of currents and recruitment were relatively well-known were unsuccessful. As well, there are other factors that may confound terrain-scale analysis that are not captured, notably historical catch. This potentially influential effect could be captured by variables such as historical fishing effort or distance from port.

Given these considerations, the coarser-scale (transect-based) approach may be more appropriate, at least to large scale expansions, and has the advantage of eliminating artifacts associated with small segments, reducing the volume of lost data when gaps produced segments below the threshold length, and reduced spatial auto-correlation of observations along transects. A comparison of data and methods was suggested – scaling up

CA data to transect level then analyzing it using the OR approach at site scale and then aggregating it. It is possible that coarser scale analysis will pull in more data (those that were rejected at the <12m rule) and may indicate the finer scale (terrain metric) approach is not necessary.

A clear key need for future development of this work is to standardize the spatial unit or units of analysis used. It is too early to be prescriptive about a single standard while the exploration and testing of methods continue – for example the suitability of different scales of analysis to different scale applications (local vs statewide). Ultimately, however, a single or perhaps complementary pair of scales should be considered that have the same set of uncertainties or trade-offs so that direct comparison of data is possible.

3.1.7 Fish sizing and length-weight conversions

Estimates of the sizes of the fish counted are required to convert from counts to biomass for stock assessment purposes. The challenge, without samples of fish 'in the hand' to permit direct measurement, is (1) to measure the fish in imagery with a known level of accuracy and then (2) apply an appropriate length to weight conversion factor, or (3) use an independent source of proxy information such as fish caught in the same areas as ROV samples.

Neither state has yet fully implemented a fish size data acquisition program, although it will be possible because both are deploying calibrated paired cameras and plan to use an industry-standard tool (Event Measure) to take the data. Currently:

- CA identify fish to species or lowest possible taxonomic level and determine size to 1 cm using the scaling lasers; they are flagged in the database for future examination with the stereo sizing method. About 30-50% of fish are sizeable by stereo in CA data. Although the CA group hadn't generated their stereo data, a Moss Landing study did evaluate them relative to visual approximation with paired lasers. This showed that visual approximations were biased upwards (larger size) and the bias was proportional to the size of the fish. A proxy length composition was also considered, and the relatively large sample size in CRFS length composition data was compared to the visual approximations. It was found that the CRFS lengths were biased downwards (smaller size). It was considered of interest to compare the paired camera data to CRFS data given the same direction of biases in relation to the visual estimates.

- OR adopted a similar approach, with length estimated for 12 species and one unidentified rockfish category in relatively broad ranges: < 10 cm, 10 – 30 cm, 30 – 60 cm, > 60 cm. Fish size was recorded only made where fish are broadside near the lasers against a scaleable background. Sparse stereo data were recorded for all 12 species. A primary driver for fish size is the change in size with time in reserves. It was acknowledged that the set of scoring condition is infrequently met, and furthermore, that many factors may have influenced the video reviewer's decision as to whether to make a size estimate for a given fish, and those factors may not have stayed constant over time. Overall, there was poor correspondence between stereo measures and annotator estimates.

In summary, improving fish length estimation using stereo imagery is a necessary but still evolving component of both programs. Laser-based estimate are unreliable, and there was strong discouragement to use proxy length frequency data (e.g. from catches) due to potential but unknown biases from selectivity effects, and possibly ontogenetic patterns of habitat use in some species. Collecting length data on future ROV deployments adds little to the fixed costs of operation (calibrations, and some additional data management overheads). The challenges are to make the cost of data processing affordable (justifiable) and to measure a sufficiently high number of individuals of target species to represent population size structure. Higher numbers of sized fish may result from choosing a different measure less prone to fish orientation (e.g. eye to caudal peduncle) and generating a corresponding length to weight conversion. It was noted that stereo fish lengths with a measure of uncertainty can be applied to uncertainty in biomass. The additional benefit of collecting these data are their relevance to assessing the performance of marine reserves – testing whether they provide refuges enabling fish to grow larger (older). This shared objective should have positive implications for cost-sharing.

3.1.8 Seabed data: missing data and gap-filling

Expanding local scale fish density data requires geological or MBS data for the full state areas of interest. There was much discussion of data deficiencies and gaps in the areas of interest, priority areas, and ways in which these may be addressed or filled. The primary gaps are the 'white zone' (coastal area) and deeper (e.g. >110 m off CA) areas where multibeam mapping data are missing. For example, the white zone in CA stems from the CSMP multibeam data rarely extending inside the state waters boundary which leaves a ~50 to 500 m corridor of unmapped area adjacent to the coastline. This may not be important for some species, but others, e.g. gopher rockfish, has large parts of its population in this zone. There are alternative sources of predictive data, but they are generated in different ways and have varying spatial resolution and are therefore not always consistent in their substratum classifications.

Discussions at the panel meeting revealed prospectively new sources of data for mapping data, and potentially new ways to integrate data. Also, there was good knowledge amongst the science teams of which species the missing shallow and deep data were most relevant to. Given this existing knowledge and a higher level of relevant expertise residing in others at the review, I feel there is no benefit from me providing a lot of further detail here. There are obvious needs to account for areas of species distributions that lack data for future statewide expansions of fish density data. In brief overview, there is a need for (1) short-term review and incorporation of the best available seabed data for the white zones and deeper shelf, and (2) longer term exploration of opportunities to acquire future MBS data. Ideally, the mapping locations of greatest benefit to the fishery assessment process should be identified and ranked, and if possible, used either to initiate fishery-dedicated mapping programs or influence programs being undertaken for other imperatives. An obvious example is the large deep area of unmapped rocky area at Cape Arago which limits statewide expansion in OR. Apparently, there had been a review of priorities for additional MBS mapping on the US west coast, but it was not clear to what extent fishery needs had been considered.

3.1.9 Data management

Both programs have sophisticated means of managing data during field acquisition, post-processing, analysis and visualization. These were evident in the state reports (and exemplified by Oregon's comprehensive library of linked analysis results), and during various discussions in the review. Oregon, at least, has good backend infrastructure, with database linkage enabling anyone else to recreate the project and products from archives; these will support data sharing, partnering, and interoperability. However, for much of the collective ROV operations data and mapping data, a third party would have to know who to talk to.

Detailed suggestions and rationales for integration and transfer of local databased into a central respository were provided in the other CIE report by Arthur Trembanis. He points to the NOAA Fisheries Strategic Initiative on Automated Image Analysis program [https://marineresearchpartners.com/nmfs_aiasi/Home.html] as one example of an established system – and the additional benefit from making data available centrally. In this case, broader access to the derived imagery from the ROV surveys would accelerate the process to develop and test automated machine learning algorithms for fish detection and sizing. His major recommendation was to review and closely align the data collection, archiving, and image analysis to the tools and protocols outlined by the AIASI initiative.

In summary, I think it was acknowledged that there is no centralised repository for ROV data but that this is a sound strategic goal. This need also seemed applicable to other types of stock assessment data. Any such developments move towards compliance with the FAIR principles for data management which are increasingly being linked to funding and publication eligibility.

3.2 Suggestions for improvements to survey designs and analytical methods (TOR 3)

3.2.1 Implement future surveys based on fishery-specific objectives

I was surprised to learn that the ROV data being considered for stock assessment purposes have been generated from leverage and piggybacking on other projects which have different objectives – typically surveys of marine reserves. An obvious consequence is that these designs are not optimal for fishery stock assessment objectives. This is neither a criticism of the research efforts to date, or on the veracity of the results presented. It is a comment on the future need for a more strategically planned sampling program. Data acquired for other purposes is more difficult to understand and will have higher variances for predictions due to the lack of planned replication in space and time, and because there are gaps in data. For example, state-wide expansions across areas (sites) are required to integrate data over years [Section 3.1.1].

These data are sufficient to estimate local stock relative abundances and suited to explore state-wide expansions, but for more routine and broad-ranging uptake in the stock assessment process it is highly desirable to ensure that future surveys have the ability to meet fishery-related objectives. One imagines that in many instances these will be a value-adding opportunity for marine reserve managers. There are locations that are difficult to access for reasons of seasonal sea conditions and/or tide, and that sampling these needs very precise timing and strategic planning [Section 3.1.1]. New surveys could also consider the utility of transect sampling design strategies such as those developed by Foster et al. (2017; 2020) to emphasise locations of higher importance, e.g. areas with high biomass or high variability and consider legacy (index) sites in designs.

In summary I recommend standardization of designs to the extent possible to minimise uncertainties when data are shared. Data acquisition has to be underpinned by strategic planning that takes account of fishery objectives – and is supported by dedicated funding.

3.2.2 Develop standard (convergent) operating procedures (SOPs)

Both programs have protocols, filters and cutoffs to limit any potential biases in ROV data from a variety of effects including variable field-of-view (FOV), height-of-bottom (HOB) and transect width that stem from large HOB over highly rugose seabed, loss of vision on backsides when the ROV descends from high relief, and high FOV due to vehicle pitch. These effects are dealt with by excluding data but the trade-offs are more data gaps, increased fragmentation of transects, and loss of data over rugose seabed that rocky reef species are associated with. Both programs have a pragmatic balance of operating procedures and data processing protocols established by that minimise data loss and prevent systematic bias, but the methods have many differences. I think there will be considerable benefit for future work – including for data amalgamation across states – to carefully document operating procedures and standardize these to the extent possible. The data reports for this review process provide a very large step in this direction [Section 3.1.2].

3.2.3 Summarise species suitability for ROV surveys

A tabulated list of candidate species, integrated across states, will assist all stakeholders to understand the potential for incorporating ROV data into stock assessments. This should show behavioral and ecological traits, geographical patterns and variations (e.g. state-specific depth patterns), overall tractability for expansions at particular scales (e.g. local reef areas vs population scale), and rationale for prioritization against (1) the need both to refine or complete analysis of current data, and (2) for the needs of stock assessment processes [Section 3.1.3]. The formal process to list/ rank priority species for stock assessments would seem to be an ideal target for a ROV-focused summary.

3.2.4 Evaluate needs and cost-benefit for future ROV infrastructure

Improved position-fixing of ROV transects should be considered. Greater accuracy may be possible by improving the processing of currently collected USBL data, and would be improved using a DVL system [Section 3.1.4]. However, each alternative would need a cost-benefit analysis to determine the effect on the final error estimates in fish abundances – which in turn will depend on the final methods employed for data expansions

[Section 3.4.1]. Any considerations are part of a larger question related to the replacement and maintenance of vehicles to enable continuation of this program of work beyond the shelf-life of the vehicle available currently [Section 3.1.5].

3.2.5 Evaluate and compare seabed classifications

I have some uncertainty about whether the same seabed types are represented in the same way across states because the data sourced, mapping methods, classifications and degree of lumping of categories differ. Treating substrate types more consistently as a continuous variable had been considered but it generated data artefacts. There is also the question (raised by another panel member) of "how believable is the mapped habitat data, and do ROV observations match the habitat maps?" [Section 3.1.5]. Thus, a recommendation would be to consider a validation ('cross-walk') exercise that completes a two-step comparison: first, use ROV observations to do a reality check on geological polygons, and second, implement both ROV data scoring methods on an example dataset from one or other, or both, states.

3.2.6 Standardise the spatial unit of fish count analysis

A clear key need for future development of this work is to standardize the spatial unit or units of analysis used. It is too early to be prescriptive about a single standard while the exploration and testing of methods continue – for example the suitability of different scales of analysis to different scale applications (local vs statewide) [Section 3.1.6]. Ultimately, however, a single or perhaps complementary pair of scales should be considered that have the same set of uncertainties or trade-offs so that direct comparison of data is possible.

3.2.7 Refine and standardize methods for converting fish counts to biomass

Improving fish length estimation using stereo imagery is a necessary but still evolving component of both programs. Laser-based estimate are unreliable, and there was strong discouragement to use proxy length frequency data (e.g. from catches) due to potential but unknown biases from selectivity effects, and possibly ontogenetic patterns of habitat use in some species. Collecting length data on future ROV deployments adds little to the fixed costs of operation (calibrations, and some additional data management overheads). The challenges are to make the cost of data processing affordable (justifiable) and to measure a sufficiently high number of individuals of target species to represent population size structure. Higher numbers of sized fish may result from choosing a different measure less prone to fish orientation (e.g. eye to caudal peduncle) and generating a corresponding length to weight conversion. It was noted that stereo fish lengths with a measure of uncertainty can be applied to uncertainty in biomass [Section 3.1.7]. The additional benefit of collecting these data are their relevance to assessing the performance of marine reserves – testing whether they provide refuges enabling fish to grow larger (older). This shared objective should have positive implications for cost-sharing.

3.2.8 Develop a strategy to fill key gaps in mapping data

There are obvious needs to account for areas of species distributions that lack data for future statewide expansions of fish density data. In brief overview, there is a need for (1) short-term review and incorporation of the best available seabed data for the white zones and deeper shelf, and (2) longer term exploration of opportunities to acquire future MBS data in each state. Ideally, the mapping locations of greatest benefit to the fishery assessment process should be identified and ranked, and if possible, used either to initiate fishery-dedicated mapping programs or influence programs being undertaken for other imperatives. An obvious example is the large deep area of unmapped rocky area at Cape Arago which limits statewide expansion in OR [Section 3.1.8]. Apparently, there had been a review of priorities for additional MBS mapping on the US west coast, but it was not clear to what extent fishery needs had been considered.

3.2.9 Plan strategically for a centralized ROV data repository

It was acknowledged that there is no centralised repository for ROV data but that this is a sound strategic goal. This need also seemed applicable to other types of stock assessment data [Section 3.1.9]. Any such developments move towards compliance with the FAIR principles for data management which are increasingly being linked to funding and publication eligibility.

3.3 Evaluate model assumptions, estimates, and major sources of uncertainty (TOR 4)

3.3.1 General

The model assumptions, estimates, and major sources of uncertainty, and ongoing discussion of analytical methods, were discussed during day 3 of the open review panel meeting. Presentations were made on each of the following topics by CA and OR:

- Model selection

- Spatial expansions
- Spatial analysis options
- Utility in stock assessments

3.3.2 Model selection, spatial expansions and spatial autocorrelation

A key requirement of the ROV data for fishery stock assessment purposes is to expand density estimates from transects to broader areas – within study sites, throughout state waters, and potentially across the range of a managed stock. The methods explored here rely on establishing spatial relationships between the abundance of a species and its environment, then using these environmental covariates mapped at large spatial scales to predict (expand) the transect-scale abundances to broader areas. There are many issues to consider in data expansion, including the spatial scales (data resolution and locational uncertainties) of both observations and independently collected covariate data, and the characteristics and assumptions of models used for the expansion. Both programs have used extensive and well-documented sets of methods to evaluate, analyse and expand the data – including exploring relatively simple design-based methods, and more complex model-based methods. These are summarized below using excerpts of text from my desktop review to capture all this detail in a single document. The approaches of each state team are kept separate because the approaches differ substantially.

Spatially auto-correlated data is expected to be a feature of the ROV data because transects are necessarily close in space on reefs that are, in general, relatively small, and that the clusters of samples are grouped in an otherwise very large (statewide) survey area. Spatial non-independence will be stronger when the unit of spatial analysis is small, i.e. transect segments vs whole transects, and when the extent of expansion is limited, e.g. reef scale vs regional scale. Both states explored tools and techniques to unpack spatial auto-correlation effects on correlations of fish density with covariates, and on estimates of variance.
My review of the potential to deal with spatial auto-correlation issues in data is limited due to my lack of relevant experience.

3.3.2.1 CALIFORNIA

California's expansion methods (design- and model-based) utilised multi-beam sonar (MBS) mapping data provided by the California Seabed Mapping Program (CSMP); in the model-based method these data were used to generate raster covariates for presence of hard bottom defined using seafloor roughness as a proxy, and terrain (relief) attributes derived from bathymetry.

- Design-based method: used the derived area of hard bottom habitat in a given depth and latitude; expansion of density data = entire area with CSMP data
- Model-based method: relied on the seafloor area within a given latitude as well as derived relationships between density and depth, latitude and proportion of hard bottom and terrain attributes from GLMs; thus, expansion of density data = a slightly reduced area of CSMP data in which terrain metrics can be generated

3.3.2.1.1 Model selection

GLMs were used to test for significant correlations of gopher, copper, vermillion, China rockfish and kelp greenling density with different suites of environmental variables to identify variables relevant to the expansion of abundance data.

GLMs were applied to two data sets of density data from <150 m depth using:

- variables from all ROV segments ('full')
- variables from only segments with CSMP data (terrain metrics and proportion of hard bottom) ('reduced')

Four suites of variables were tested:

- GLM1 – variables = av. depth, latitude of segments centroid, proportion of hard/mixed bottom, take vs no take (MPA) [area observed included as an offset because it varied between segments]
- GLM2 – included 'effort' variables: distance, width, time (correlated with area)
- GLM3 - + distance, width, time + CSMP hard bottom and terrain
- GLM4 – CSMP hard bottom and terrain only

Five distributions were compared and models tested using the AIC: Regular Poisson, zero-inflated Poisson, quasi-Poisson, zero-inflated quasi-Poisson and negative binomial because fish count data were over-dispersed in segments (including effects of depth or habitat for uncommon species) leading to large numbers of zeros, i.e. a greater skew than regular Poisson or negative binomial distributions. In addition, the binomial was run using counts converted to presence/absence data. Deviance and selection in backward stepwise model selection were evaluated for gopher rockfish to see if important variables were consistently identified between methods (except for zero-inflated models).

3.3.2.1.2 Expansion methods

Expansion used design-based and model-based methods with one species, gopher rockfish, as a test case. Five different methods were considered, and two implemented (#2 for design, and #5 for modelled) for review. The design-based methods could use average density distributions across the important variables identified from the GLMs (depth, latitude and proportion of hard bottom) to post-stratify, with total abundance being the sum of average density x area estimates for each stratum.

- *Method 1. Design-based: statewide densities and hard bottom habitat area.*
- *Method 2. Design-based: Method 1 + depth and latitude.*

The possible model-based methods differ from design-based by making more use of the CSMP data – terrain metrics paired with the centroid of each ROV segment can be estimated to provide variables to be evaluated for correlation to density using GAMs. Coefficients would be derived for variables found to be significantly correlated in GAMs. MGET can then be used to estimate fish density in each CSMP raster grid cell using correlations derived by the GAM. Fish densities in each grid cell are multiplied by the area of each grid cell and summed across the mapped area.

- *Method 3. Model-based: estimate with coefficients from GLM using latitude and depth from the ROV observations.*
- *Method 4. Model-based estimate from GLM using ROV observations paired with CSMP with derived terrain attributes.*
- *Method 5. Model-based estimate from GAM using ROV observations paired with CSMP with derived terrain attributes.*

The California study made a thorough assessment of the suitability of methods for the available data using the results of GLMs that tested for significant correlations of fish density (gopher, copper, vermillion, China rockfish and kelp greenling) using different suites of variables and data distributions.

- GLMs with ROV variables (the variable area of transect segments that have fixed length but varying width was treated as an offset) showed that whilst results varied by species but there were some clear patterns. Quasi-Poisson or negative binomial distributions were the best fits for density data. There was significant over-dispersion for all species – which should be addressed by using the quasi-Poisson or negative binomial models, and zero-inflated models or filtering may be beneficial to eliminate segments outside the distribution of the species in question. Backward stepwise regression on gopher rockfish data clearly showed all four variables should be included in deriving indices of abundance.

- GLMs with ROV and area (effort) variables further showed that the expected correlations of effort (distance, width and time) with density varied across species and model distributions, and the time was consistently significant. However, distance was still considered preferable to time as the basis for defining segments because it will produce a more consistent spatial coverage among segments.

- GLMs evaluating all variables – ROV, area (effort), and CSMP variables (proportion of hard bottom and terrain metrics) – found patterns very similar to above, including appreciably lower AIC scores for the negative binomial across species. There were some differences: copper rockfish was no longer significantly correlated with depth, and take was significant for zero-inflated models, whereas not previously. However, the CSMP data was a mixed bag: the proportion of rocky bottom was consistently highly significant across species and data distribution types, but there were not consistent correlations of terrain metrics for a given species across distribution types – even though they were frequently significant.

- GLMs with a reduced suite of variables – take (no take inside marine reserves), latitude and CSMP – were used to identify the variables suited to model-based expansion of abundance data. These analyses also showed the suitability of distributions other than Poisson: the negative binomial for two species, zero-inflated Poisson for three species, and that over-dispersion was better represented by negative binomial or quasi-Poisson distributions. Latitude, CSMP hard bottom and take were consistently significant variables across species and distributions.

Two of the five methods were implemented (#2 for design, and #5 for modelled) for review.

- *Method 1. Design-based: statewide densities and hard bottom habitat area.*
- *Method 2. Design-based: Method 1 + depth and latitude.*

Both possible methods rely only on ROV survey data, and expand an average fish density (mean and variance from bootstrapping across segments) by multiplying the area of hard (rocky) bottom; Method 2 represented an improvement by also capturing variations in density related to depth and latitude which were significantly correlated with fish distributions, especially gopher rockfish. The strengths of these methods is their relative computational simplicity, and simple assumptions, i.e. they are not prone to spatial error (locational and scale error in data matching) that accompanies the use of CSMP-derived terrain metrics. The significance of depth, but near absence of fish outside the 10-60 m depth range was tackled using a reduced data set (10-60 m only) and stratification across 10 m depth bins. An estimate of variance within stratum assumed – fairly – that habitats across all strata were classified without error from CSMP data.

The obvious weakness in the initial design-based method is that fish on soft-bottom are not accounted for. This arises because expansion uses the CSMP classification of bottom type where raster cells are classified as either hard (rock) or soft bottom based on rugosity as a proxy –and density expanded using only hard bottom cells. This is a negative bias because fish are observed off hard bottom by the ROV. To compensate for this, estimates were made of mixed bottom by calculating proportions of cell types within 30 x 30 m neighborhoods. Because this method generates a full spectrum of hardness – presumably from 0-100% - it provides an opportunity to add a threshold of hardness below which gopher rockfish are assumed to have near-zero abundance. In a second step in Method 2, a threshold of 10% hard (mixed) bottom was chosen – which is rather low. This factor has a very large influence on the final estimate – 10% threshold elevates the total abundance estimate by ~64%. Independently of this, the difference between the depth-stratified and un-stratified total abundance was small (~5%), and the addition of estimates from outside the 10-60 m depth range was negligible.

- *Method 3. Model-based: estimate with coefficients from GLM using latitude and depth from the ROV*
- *Method 4. Model-based estimate from GLM using ROV observations paired with CSMP with derived terrain attributes.*
- *Method 5. Model-based estimate from GAM using ROV observations paired with CSMP with derived terrain attributes.*

Method 3 is a model-based version of Method 2, implemented in R or MGET. It brings in density estimates derived from GLMs, and a fine spatial (raster cell) scale rendition of latitude, depth and hard/soft bottom from CSMP data. It has the same general strengths and weaknesses as Method 2, but, although being more computationally intensive.

Method 4 is a model-based estimate implemented in MGET using density estimates derived from GLMs, and pairing these with terrain metrics derived from CSMP data. This method brings in the terrain attributes – but there are inconsistent relationships with density across species, and there are new and unknown spatial errors introduced. However, it is a potential disadvantage if there are spurious correlations of density with some attributes for some species. Another major advantage is accounting for abundance on soft and hard substrates. However, it could not account for non-linear relationships between density and variables.

Method 5 (implemented) is a model-based estimate implemented in MGET using density estimates derived from GAMs (that allow both flexibility for non-linear variables and implementation of negative binomial distributions) and pairing these with terrain metrics derived from CSMP data. Its strengths and weaknesses are very similar to Method 4, but it has the advantage of being able to capture non-linear relationships, e.g. the increase then decline in density of China rockfish with depth. The GAM provided a good fit and an acceptable pattern of residuals. Predicted densities were mapped for the study area, and an overlay of ROV observations showed reasonable consistency with the predictions. Method 4 using a GLM was also run for comparison. My understanding is that this method uses CSMP data for terrain metrics, but not the computation of mixed bottom (hard bottom at <100%), so it will underestimate fish abundance in the same way the initial design-based method did.

The repeated K-fold cross validations were a useful way to evaluate the models' predictive performance by splitting data into training and validation data sets – proportions of 90-10 and 80-20. Predictions were compared to validations and using different variables: those from the ROV vs. from the CSMP vs. from CSMP without the terrain attribute. There is a problem with the overall comparison because large numbers of zeros affect the analysis, and predictive ability was poor (low R^2). However, the omission of variables showed that leaving out the terrain attribute actually improved the fit, and that predictions were poor without latitude.

3.3.2.1.3 Spatial autocorrelation

The Californian team used two metrics to explore this Moran's I and Ripley's K. I am not able to adequately summarise the findings, but it appeared that neither metric was well suited to this assessment, confirming the obvious close spacing of transects and clustered samples within a very large survey area. There is a paper in an advanced stage of preparation that promises to better explain the results (Perkins et al., in preparation).

3.3.2.2 OREGON

Oregon's expansion methods (design- and model-based) also utilized MBS data to identify rocky habitats, but here, bathymetry and backscatter were used to generate a geological substratum classification (SGH) that was collapsed into simple categories for analysis. The use of GIS generated terrain variables was explored, but not used in analysis. The Oregon program's analysis had been considerably expanded since the desktop review.

Design-based analysis used both the full transect data and the filtered rock-only transect data to generate abundance estimates. Both datasets were applied to the rock area, and for the full transect data only, fish density was applied to the rock buffer area, to compare how the inclusion of the non-rocky portions of the survey data affect abundance estimates. Note, the question of 'where is the reef boundary' is treated here by defining a rock buffer = 211 m, beyond the rock polygon. Full transect data were applied to the rock buffer area to see how inclusion of non-rocky portions affected (depressed) abundance estimates. Applying full transect data to the buffer region aimed to establish a prediction area that includes reef and appropriately limited non-reef areas. Full transects were the primary spatial unit of analysis (mainly to eliminate high fish density artifacts associated with small segments, but also to reduce spatial auto-correlation).
Density estimates were calculated for rock-buffers using full transect estimates; this was important because sometimes the buffer was larger than the area of rock. However, the relatively arbitrary size of the buffer was acknowledged as an obstacle to interpreting its associated abundance estimate, and there was an acknowledged 'substantial problem' expanding substratum-specific density into the rock buffer due to the estimates of variability depending on how data (n= 839 segments) are aggregated in each of the two methods used. There are lower SDs for Method 2 (density x substratum within each site) cf Method 1 (density x substratum pooled across all sites) because it uses a higher number of samples but this represents a source of uncertainty because the assumption of their independence is unsupported.

3.3.2.2.1 Model selection

Filtered (rock-only) transect data were used for expansion and produced slightly higher abundance estimates than the full transect data (as predicted), but the difference was subtle – probably reflecting the predominance of rocky substrates across most transects (noting, conversely, targeted transects at the sandy substratum off Cape Perpetua). The overall similarity of the summed estimates and SD across species is reassuring that the rock-only data were suited to modelling. However, it was noted that Cape Arago had a relatively large influence because of its relatively very large rocky area, e.g. illustrated by kelp greenling – low density x very large reef area. This site swamps other better-sampled sites. More data is needed offshore if this area is to be adequately assessed.

GAMs were selected for the expansion process because of observed non-linear effects of depth, and applied to the density data for:

- filtered (rock-only) transect data from all surveys except Cape Perpetua (where the design differs to deal with sparse rocky reef), and five transects deeper than 70 m at Cape Arago.
- two species with contrasting abundance – kelp greenling (~2600 total observations) and yelloweye rockfish (~180 total observations), to illustrate one potential approach for using the entire dataset to estimate total abundance.
- Response variable = total transect fish count (the nongap, rock-only portion of the transect).
- GAM - variables = Site (each of the 11 sites, "Location" in the model output), a smoothed function of Depth, and Year (as a random-effects variable to address the unexplained variability between different surveys at the same site (Orford, 2010 & 2016; Cascade, 2012 & 2017), not to detect temporal trends). As for California, view area observed included as an offset because it varied between segments.
- A negative binomial distribution was determined to provide the best fit to the data for the focal species.

3.3.2.2.2 Expansion

Unlike California, where several models were considered and two implemented for data expansion, here, expansion of the density data was by individual areas (sites) and regional survey areas:

- Area total abundance = mean transect density (weighted by view area) x area of rock pooled for all sites in each area.
- Region total abundance (entire surveyed area) = sum of 11 sites; SD= square root of the sums of the individual site variances.

The model was used to predict abundance at a set of grid points covering the entire mapped area of rocky reef in the surveyed regions (8.4 million points at 16 m^2 grid resolution). Grid re-sampling was used to provide confidence intervals around abundance estimates, and to reduce computing time the complete grid was sub-sampled to generate a lower resolution prediction grid (1 million points at 135 m^2 grid resolution) using depth at each point in the complete grid). Diagnostics from abundance estimates generated by resampling the prediction grid showed its lower spatial resolution didn't compromise the final abundance estimate.

Unfortunately, I don't understand the underlying assumptions or the mechanics of the redraw method, but this was assessed by experts in the meeting as a robust way to generate a distribution of abundance estimates and confidence intervals – and is documented more thoroughly in the combined panel report. In summary, for the example species discussed (kelp greenling and yelloweye rockfish) the GAM resampled estimate had overall lower variance than either the initial GAM model and design-based estimates. It was noted by the panel that the total number of fish seemed to be on the same scale as estimates available from recent stock assessments.

3.3.2.2.3 Spatial autocorrelation

The Oregon team used the R software package VAST (Vector Autoregressive Spatio-Temporal model) to explicitly incorporate spatial factors. This model appears attractive because it was built to deal with similar effects in trawl-derived estimates of biomass. It uses two linear predictors representing encounter probability (presence/absence) and positive catch rate (metric tons biomass) given an encounter, and these can also be used to implement a zero-inflated analysis on count data. This is ideal because ROV surveys produce many zeros; the first linear predictor represents encounter probability (presence/absence) and the second individual density using a Poisson or other discrete-count distribution. However, VAST is not suited to data expanded into large areas without survey data, and, somewhat ironically, could not be applied to data for kelp greenling because the species was 'too abundant', i.e. its occurrence at every site violated the need for the first model predictor to determine a varying encounter probability. Analysis of data at the Port Orford region, where there are many samples taken in three years, was used to assess the model's potential application, i.e. if spatial or spatio-temporal factors generate structure in ROV dataset leading to bias or lack precision in non-spatial analysis. This model run for lingcod incorporated a spatio-temporal variable and identified 'hotspots' across the three sampling years. However, the resulting total abundance calculated by VAST (3-year mean = 31,680 fish) was consistent with the estimate from the coastwide GLM (Port Orford only): 30,561 fish, with a 95% CI of 22,629 – 41,600.

Again, although I don't understand what goes on 'under the hood' with VAST, it appears to provide another good option for ROV data by enabling all variability associated with variation among years (potentially including natural variation in abundance, and any survey-related differences in detection or swept-area calculation) to be rolled into a single random effect. The result is a single estimate of abundance that accounts for variability among years, while also incorporating any static spatial effects (not spatio-temporal effects). The Oregon team evaluated this by comparing a standard multi-year VAST analysis for yelloweye rockfish at Port Orford followed by a single-year “vessel effect” run with the coastwide yelloweye rockfish GAM model (Port Orford only). The aggregate abundance estimate was 8,520 +/- 1,319 compared to the GAM estimate 8,213, with a resampled 95% CI of 4,687 – 60,379. Thus, the abundance was very similar but the range of estimates much smaller. The ‘single random effect’ option appears to have dealt well with this species characterized by small sample size and patchy distribution.

3.3.3 Summary

There are many issues to consider in data expansion, including the spatial scales (data resolution and locational uncertainties) of both observations and independently collected covariate data, and the characteristics and assumptions of models used for the expansion. Spatially auto-correlated data is expected to be a feature of the ROV data because transects are necessarily close in space on reefs that are, in general, relatively small, and that the clusters of samples are grouped in an otherwise very large (statewide) survey area. Spatial non-independence will be stronger when the unit of spatial analysis is small, i.e. transect segments vs whole transects, and when the extent of expansion is limited, e.g. reef scale vs regional scale. Both states explored tools and techniques to unpack spatial auto-correlation effects on correlations of fish density with covariates, and on estimates of variance. Both programs used extensive and well-documented sets of methods to evaluate, analyse and expand the fish density data – including exploring relatively simple design-based methods, more complex model-based methods, and spatial autocorrelation effects.

The California study made a thorough assessment of the suitability of methods for the available data using the results of GLMs that tested for significant correlations of fish density (gopher, copper, vermillion, China rockfish and kelp greenling) using different suites of variables and data distributions. Data expansion explored five different design-based and model-based methods and implemented one of each. The design-based method used average fish density (mean and variance from bootstrapping across segments) to multiply the area of hard (rocky) bottom, and captured variations in density related to depth and latitude which were significantly correlated. The strength of these methods is their relative computational simplicity, and simple assumptions, but their weakness is a negative bias from not accounted for fish on soft-bottom. A compensatory ‘mixed bottom’ density estimate was added, but this arbitrary value had a very large influence on the final estimate.

The model-based estimate implemented in MGET using density estimates derived from GAMs paired with terrain metrics derived from CSMP data. The GAM provided a good fit and an acceptable pattern of residuals. Predicted densities were mapped for the study area, and an overlay of ROV observations showed reasonable consistency with the predictions. Repeated K-fold cross validations to evaluate the models' predictive performance (training /validation data sets in proportions of 90-10 and 80-20) were generally weakened by the large numbers of zeros in data. However, they showed that predictions were poor without latitude, and that leaving out the terrain attribute improved model fit. The Californian team used Moran's I and Ripley's K to explore spatial autocorrelation in their data, but it appeared that neither metric was well suited to this assessment.

The Oregon program's analysis had been considerably expanded since the desktop review. Oregon's expansion methods (design- and model-based) also utilized MBS mapping data to identify rocky habitats, but here, bathymetry and backscatter were used to generate a geological substratum classification (SGH) that was collapsed into simple categories for analysis. Terrain variables were explored, but not used in analysis. Full transects were the primary spatial unit of analysis (mainly to eliminate high fish density artifacts associated with small segments, but also to reduce spatial auto-correlation). Analysis covered many species but focused on two: kelp greenling and yelloweye rockfish. To account for fish on soft and mixed bottom adjacent to reefs, density estimates were calculated for rock-buffers using full transect estimates. This was important because sometimes the buffer was larger than the area of rock. However, this step was an acknowledged 'substantial problem' because the buffer was arbitrary in size and highly influential on expansion results. As well, estimates of variability are dependent on assumptions about data independence within sites and how data were aggregated (within site or pooled across all sites). Filtered (rock-only) transect data used for expansion produced slightly higher abundance estimates than the full transect data (as predicted), but the difference was subtle – probably reflecting the predominance of rocky substrates across most transects. The overall similarity of the summed estimates and SD across species is reassuring that the rock-only data were suited to modelling (although noting the disproportionate influence of the relatively very large rocky area at Cape Arago). Data expansion was by individual areas (sites) and regional survey areas. Design- and model-based methods were explored, and a redraw method for the GAM model generated a overall lower variance for the two test species (kelp greenling and yelloweye rockfish). It was noted by the panel that the total number of fish seemed to be on the same scale as estimates available from recent stock assessments. The Oregon team used the R software package VAST (Vector Autoregressive Spatio-Temporal model) to explicitly incorporate spatial factors. This model was built to deal with similar effects in trawl-derived estimates of biomass and appears attractive because (1) it is able to deal with large numbers of zero abundances in data (characteristic of ROV surveys); and (2) because it appears to provide a good option for enabling all variability associated with variation among years (potentially including natural variation in abundance, and any survey-related differences in detection or swept-area calculation) to be rolled into a single random effect ("vessel effect" in trawl assessment parlance). However, VAST is not suited to data expanded into large areas without survey data, and, somewhat ironically, could not be applied to data for kelp greenling because the species was 'too abundant', i.e. its occurrence at every site violated the need for the first model predictor to determine a varying encounter probability. VAST analysis of data at Port Orford, where there are many samples taken in three years, to assess whether spatial or spatio-temporal factors generate structure data that lead to bias or lack precision in non-spatial analysis, did identify 'hotspots' across the three sampling years but the resulting 3-year averaged abundance was consistent with the estimate from the coastwide GLM applied to Port Orford. The "vessel effect" analysis for yelloweye rockfish at Port Orford – a single estimate of abundance that accounts for variability among years while also incorporating any static spatial effects (not spatio-temporal effects) – generated an abundance estimate very similar to the coastwide GAM for Port Orford, but the range of estimates much smaller. Thus, the 'single random effect' option appears to have dealt well with this species characterized by small sample size and patchy distribution.

3.4 Suggestions for improvements to models and spatial expansion

3.4.1 Harmonise the data analysis strategy

Discussions at the review identified several opportunities for potentially improving the data analysis strategy [Sections 3.3 and 3.2.5, 3.2.6, 3.2.7]. These provide the basis for the next round of evaluation within and between states and should be a core activity – perhaps the core activity – for future short-term work. The aims would include cross-fertilization (applying the most robust methods to data from both states) and convergence of methods between states to identify, then minimise, sources of variance in the final abundance estimates. Work areas include:

- Explore ways to consistently and rigorously represent 'hard bottom' because this is a key to consistent data expansion [Section 3.1.5]
- Explore ways to consistently and rigorously define and map 'non-hard bottom' (sediments, mixed bottom, unknown areas) adjacent to reefs. A method is needed to

map into non-reef areas that have relatively low fish density but over relatively large areas [Section 3.1.5]

- Evaluate alternative segmentation strategies and linkage to covariates – especially to decide whether it is worth pursuing the use of fine-scale terrain metrics [Section 3.1.6]
- Consider defining fish density-habitat relationships at site scale to improve estimates of variance. Relationships differ between sites but are very tight at some. Reefs are independent, whereas cells in a coastwide analysis best predict what is in adjacent or close cells and this is inconsistent with some tools (e.g. MGET) that assumes independence. Even without statewide estimates of abundance, there is scope for more smaller scale application such as detecting localised depletion
- Determine the optimal spatial scales at which variables derived and applied, e.g. % hard bottom, and terrain metric
- Examine the effects of zeros in analyses where they may be incorrectly interpreted as identifying areas with low variability rather than areas with no values
- Further explore the OR resampling method to account for spatial autocorrelation, especially where segment scale analysis incorrectly assumes independence of density estimates [Section 3.3.2.2.3]

There is a strong general overlap in these needs and tasks with the aims of a Visual Survey Methods Workshop run by a Technical Sub-Committee (TSC) of the Canada-U.S. Groundfish Committee in April 2014. That provides one template for planning future convergence and standardization of analysis methods – but using a focussed workshop and data exchange between states to advance specific areas of analysis.

3.4.2 Validate models by future sampling

There was scattered discussion about the need and opportunities for model validation by future sampling. The three opportunities I recorded were:

- Validation of density estimates across methods at site scale using fishery dependent data (CERF catch data for mapped reefs)
- Sampling across the gradient of predicted fish densities, e.g. in the California data there are few ROV observations from areas predicted to have the highest gopher rockfish densities (see CA Figs 32 & 33).
- Validation of low fish density on non-hard bottom

3.4.3 Review the needs and options for spatial analysis

Overall, a rigorous analysis of spatial autocorrelation has yet to be completed. The Oregon analysis using VAST was an important and encouraging step towards this [Section 3.3.2.2.3]. [Although it identified spatial hotspots at Port Orford in samples taken over three years, a 3-year averaged abundance was consistent with the Port Orford estimate from the coastwide GLM. The "vessel effect" analysis (a single estimate of abundance that accounts for variability among years) generated an abundance estimate very similar to the coastwide GAM, but the range of estimates much smaller. Thus, the 'single random effect' option appears to have dealt well with this species characterized by small sample size and patchy distribution.] It is important to further explore this methodology and apply it across more data, including the CA data. The suggestion from experts at the review was to keep the approach relatively simple, perhaps unless going to segment or terrain metric scales

3.4.4 Better define the utility of ROV data in stock assessments

Packaging the CA and OR information for evaluation in this review was a major step towards evaluating it utility in the stock assessment process, and there was some discussion at the end of the review about specific applications. One is the opportunity to inform the process that prioritises the species for upcoming assessments with a detailed list of suitable species and rationales [Section 3.2.3]. Beyond that, there are a range of data and survey specific items to be considered.

ROV data are scattered across areas and years and cannot provide annual measures to detect changes in abundance, so how can data be treated? The data are sufficiently consistently collected through time to be aggregated into 'super years', e.g. for rolling repeat surveys off CA together. These considerations should inform ideas about multi-year intervals between future surveys.

A good suggestion was to provide future survey reports in a standard form that document the density estimate(s); the standard operating procedure (SOP); and the application of the data to stock assessment. These products would help define the utility of the product and support requests for funding.

There are other several ways in which data and analytical methods should be standardized, or at least converge; these include the underlying classification of seabed types [Section 3.1.5]; the spatial unit of analysis [Section 3.1.6]; the fish size to biomass conversions [Section 3.1.6]; and many aspects of the modelling process [Section 3.3.2].

3.5 Specific suggestions for short-term and longer-term improvements (TOR 5)

My specific suggestions for future improvements to relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term (improvements to existing methods) and longer-term (strategic improvements) time-frames, further summarise those made above in Sections 3.2 and 3.4.

3.5.1 Improvements to existing methods (short term)

- Summarise species suitability for ROV surveys [Section 3.2.3]
- Harmonise the data analysis strategy [Section 3.4.1]
- Evaluate and compare seabed classifications [Section 3.2.5]
- Review the needs and options for spatial analysis [Section 3.4.1]
- Standardise the spatial units of fish count analysis [Section 3.2.6]
- Better define the utility of ROV data for inclusion in stock assessments [Section 3.4.1]
- Develop standard operating procedures [Section 3.2.2]
- Refine and standardize methods for converting fish counts to biomass [Section 3.2.7]

3.5.2 Strategic improvements (long term)

- Implement future surveys based on fishery-specific objectives [Section 3.2.1]
- Validate models by future sampling [Section 3.4.1]
- Develop a strategy to fill gaps in mapping data [Section 3.2.8]
- Evaluate the needs and cost-benefit for future ROV infrastructure [Section 3.2.4]
- Plan strategically for a centralized ROV data repository [Section 3.2.9]

3.6 Prepare a Peer Review Report following the Terms of Reference (TOR 6)

This document is my peer review report; it follows the Terms of Reference.

3.7 Provide a brief description of panel review proceedings (TOR 7)

This was my first involvement in this kind of panel review. On a personal note, I found it was well-organised and conducted in a highly constructive manner. The number of people participating in the meetings varied from session to session but was around thirty on day 2.

This provided opportunity for input from many participants aside from reviewers and was highly beneficial as many insightful views were provided by experts – especially on model selection and performance, and quality and availability of mapping data. The participants included a number of early career researchers, and I know from speaking to some of them, that they benefitted both from exposure to the science, and to the mechanics of the review process.

On a more technical note, each of the science teams delivered excellent presentations that addressed the key needs of the review. The three-person team of John Budrick, Laura Ryley and Michael Prall from California provided complementary talks (data acquisition, survey and mapping, data analysis and models) whilst the one-man team of Scott Marion from Oregon very impressively presented on the entire spectrum of topics. Presentations included responses to issues raised during the desktop review and reported on additional work done since. In fact, a considerable amount of additional work had been completed by OR since the desktop review. The science teams were open about the strengths and weaknesses of the work done to date, and not defensive about the challenges to the work – those acknowledged in reports or raised during the meetings. There were wide-ranging discussions over each of the three days which during which all the relevant topics were discussed (as listed in this report). The key issues were identified and discussed in a constructive and productive manner, the review was highly effective in providing an opportunity to summarise and present the ROV data in a way that will permit its consideration for use in stock assessments, and it enabled the relevant (e.g. 'marine parks' and 'groundfish') people to actually sit down together, in the same room, for an extended period. I think the review panel collectively had a complementary knowledge and skills set that covered the needs of the review and acted as an effective core for the meetings. There are a series of recommendations from the meeting that form part of the CIE reports, and the review panel Summary Report.

4 Conclusions

Review process

The three-day panel review was well-organised and conducted in a highly constructive manner. It enabled key issues to be identified and discussed, provided the opportunity to summarise, package and present the ROV data in a stock assessment context, and was effective in bringing a variety of relevant researchers together, including from different agencies within individual states.

Science

The science presented was of a very high standard. The knowledge, rigor and innovation applied to the data acquisition, data analysis and data products clearly demonstrated the work represents the best information available at this point. The review has identified many areas where the science can be extended in the future and evolved in ways that will improve its ability to meet the demands of the stock assessment process.

Outcomes

Many of the considerable data acquired to date are from surveys designed to meet non-fishery objectives. Dedicated, cross-state funding is required to enable the objectives of future surveys to address the needs of stock assessment in survey design and data acquisition. Targeted funding will also ensure this detailed and encouraging data exploration and interim packaging of results is transformed into a body of work that has direction and purpose. There are a variety of needs to explore, improve and standardize data products that can be addressed by a tactical workplan, and several needs and opportunities to consider more strategically. The review process, including suggestions included in this report, identify and highlight much of what is needed for both the tactical and strategic workplans to be developed and implemented.

5 Recommendation

Create a formal project with adequate resources to foster development of this work program. This high-quality science effort has the potential to provide a robust source of data to feed stock assessment of selected commercially and recreationally important inshore rocky reef fish species. No alternative sources of data exist. Draw on the considerable efforts of the science teams involved, and this review process, to establish tactical and strategic workplans.

6 Appendices

6.1 Appendix 1: Bibliography of materials provided for review

Report 1: Methods for using remotely operated vehicle survey data in assessment of nearshore groundfish stocks along the California coast - update
Authors: Dr. John Budrick, Ms. Laura Ryley and Mr. Mike Prall
Affiliation: California Department of Fish and Wildlife

Report 2: Abundance Estimation for Nearshore Groundfish from ROV Video Surveys of Oregon's Nearshore Rocky Reefs - update
Author: Scott Marion
Affiliation: Oregon Department of Fish and Wildlife

Additional materials consulted

Foster, S.D., Hosack, G.R. Monk, J., Lawrence, E., Barrett, N.S., Williams, A., and Przeslawski, R. (2020). Spatially balanced designs for transect-based surveys. Methods in Ecology and Evolution, 11: 95-105 DOI: 10.1111/2041-210X.13321

Foster, SD, Hosack, GR, Lawrence, E, Przeslawski, Hedge, P, Caley, MJ, Barrett, NS, Williams, A, Li, J, Lynch, T, Dambacher, JM, Sweatman, HPA, and KR Hayes. (2017). Spatially-Balanced Designs that Incorporate Legacy Sites. Methods in Ecology and Evolution, 8: pp 10.1111/2041-210X.12782

6.2 Appendix 2: A copy of the CIE Performance Work Statement

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

Follow-Up Panel Review of Remotely Operated Vehicle (ROV) Surveys Of Nearshore Stocks - California & Oregon

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards.
(http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).
Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

The National Marine Fisheries Service and the Pacific Fishery Management Council is seeking a panel review to follow -up on the FY19 desk review to evaluate and review fishery independent visual survey methodologies, using remotely operate vehicles, for nearshore Groundfish species off the states of Oregon and California.

West coast nearshore groundfish stock assessments have identified the current lack of fishery-independent data sources as a research and data need (PFMC, 2017, Agenda Item E.2, Attachment 1, September 2017). In addition, methods currently utilized in stock assessments do not explicitly account for differential biomass densities inside of no-take Marine Protected Areas (MPAs). Remotely operated vehicles (ROVs) provide a non-lethal sampling method in areas where harvest is prohibited. They also allow collection of data on overfished species and nearshore species that constrain take of healthy stocks. Because ROVs employ only non-lethal data collection methods, they avoid the need for research catch set-asides or other allocative considerations that may arise between fisheries and research sectors.

Both Oregon and California have conducted ROV surveys of rockfish in nearshore areas, focusing on rocky reef habitat, and, in California, on areas inside and outside of MPAs. In both states, resultant information includes density estimates (by transect and habitat) for various species and length data. In addition, the states have developed seafloor maps, allowing estimation of area of habitat types by depth and latitudinal breaks.

Density estimates can be developed in a number of ways, from simple extrapolations to more complex general linear models (GLMs) and generalized additive models (GAMs), including factors that may affect detection probability across sample sites. There is likely to be differential detection by species, gender and size, and by timing of survey as well.

Observed density estimates and indices of relative abundance or estimates of absolute abundance in the depth and latitudinal areas surveyed can be used in stock assessments, given appropriate accounting for selectivity and detection probability, or potentially used in management procedures. Length composition data collected by the surveys may be included in stock assessments or management procedures as well.
The general goals and objectives of Council methodology reviews are to:

1) Ensure that research surveys, data collection, data analyses and other scientific techniques in support of coastal pelagic species (CPS) and groundfish stock assessments are the best available scientific information and facilitate the use of information by the Council;
2) Provide recommendations regarding whether, and if so, how a particular methodology can be applied in future stock assessments;
3) Meet the Magnuson-Stevens Fishery Conservation and Management Reauthorization Act (MSRA) and other legal requirements;
4) Follow a detailed calendar and fulfil explicit responsibilities for all participants to produce required outcomes and reports;
5) Provide an independent external review of survey and analytical methods used to develop data to inform CPS and groundfish stock assessments;
6) Increase understanding and acceptance of CPS and groundfish research methodologies and review by all members of the Council family;
7) Ensure that methodologies not directly related to stock assessments, such as economic analyses or ecosystem-based fishery management approaches, undergo adequate peer review, as appropriate; and
8) Identify research needed to improve assessments, reviews, surveys, analyses, and fishery management in the future.

The goals and objectives specific to the review of the California and Oregon ROV survey methodologies are to:

1) Evaluate the sampling design used in recent ROV surveys conducted by the states of Oregon and California.
2) Evaluate proposed methods to develop indices or estimates of abundance for these ROV surveys, including using habitat/substrate type and Marine Protected Area designation as covariates.
3) Evaluate proposed methods to estimate size compositions of observed individuals of each species.
4) Identify potential impediments to developing independent indices or estimates of abundance using these ROV surveys and incorporating them into stock assessments.

A desk review of the ROV surveys will be held in May, 2019, which will provide initial feedback and recommendations to be considered and explored for the in-person panel review. This methodology review will lead to the development of materials and guidance for future ROV surveys and indices or estimates of abundance for those areas surveyed in Oregon and California, as well as the expansion of such methods to other areas within those states and/or within Washington State.

The specified format and contents of the individual peer review reports are found in **Annex 1**. The Terms of Reference (ToRs) for the review of ROV survey methodologies are listed in **Annex 2**. Lastly, the tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements
NMFS requires two (2) reviewers to conduct an impartial and independent peer review in accordance with the PWS, OMB guidelines, and the ToRs below. The reviewers shall have a working knowledge in visual survey techniques, survey design and analysis, and familiarity with incorporating survey information in stock assessments and have conducted the previous desk review held in FY19.

Tasks for Reviewers

1) <u>Pre-review Background Documents</u>: Two weeks before the peer review, the NMFS Project Contacts will send (by electronic mail or make available at an FTP site) to the CIE reviewer the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contacts will consult with the CIE on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the PWS scheduled deadlines specified herein. The CIE reviewer shall read all documents in preparation for the peer review.

Documents to be provided to the CIE reviewers prior to the methodology review include:

- Reports by the states of California and Oregon describing survey and analysis approaches and preliminary results;
- Reports of the CIE Desk Review.

- The Pacific Fishery Management Council's Scientific and Statistical Committee's Terms of Reference for the Methodology Review Process for Groundfish and Coastal Pelagic Species for 2019-2020;
- Additional supporting documents as available.
- An electronic copy of the data, the parameters, and the software used for developing population indices/estimates and compositional data.

2) Panel Review Meeting: The CIE reviewers shall conduct the independent peer review in accordance with the PWS and ToRs, and shall not serve in any other role unless specified herein. Modifications to the PWS and ToRs cannot be made during the peer review. The CIE reviewers shall actively participate in a professional and respectful manner as members of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

3) Contract Deliverables - Independent CIE Peer Review Report: The CIE reviewers shall complete an independent peer review report in accordance with the PWS. The CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. The CIE reviewer shall complete the independent peer review addressing each ToR as described in **Annex 2**.

4) Other Tasks – Contribution to Summary Report: The CIE reviewers may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. The CIE reviewers are not required to reach a consensus, and should provide a brief summary of their views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

5) Deliver their reports to the Government according to the specified milestones dates.

Foreign National Security Clearance
When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration- system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The place of performance shall be at the contractor's facilities, and in Santa Cruz, California.

Period of Performance
The period of performance shall be from the time of award through March 2020. The CIE reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Within two weeks of award	Contractor selects and confirms reviewers' participation[1]
At least two weeks prior to the panel review meeting	Contractor provides the pre-review documents to the reviewers
February 4 - 6, 2020	Each reviewer participates and conducts an independent peer review during the panel review meeting
Within two weeks after review	Contractor receives draft reports
Within two weeks of receiving draft reports	Contractor submits final reports to the Government

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content; (2) The reports shall address each ToR as specified; and (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $11,000.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

Project Contacts:
Stacey Miller
Fishery Resource, Analysis and Monitoring Division

[1] To ensure consistency of analytical approaches, the CIE reviewers for this panel review are the same personnel who conducted the previous CIE desk review on Remotely Operated Vehicle (ROV) Surveys of Nearshore Stocks in May 2019.

NMFS| Northwest Fisheries Science Center
2032 SE OSU Drive | Newport, Oregon 97365
Phone: 541-867-0535
stacey.miller@noaa.gov

Owen Hamel
Fishery Resource, Analysis and Monitoring Division
NMFS| Northwest Fisheries Science Center
2725 Montlake Boulevard East | Seattle, Washington 98112
Phone: 206-697-3102
owen.hamel@noaa.gov

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent reports shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of each reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Each reviewer should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Each reviewer should discuss their independent view on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Each reviewer should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. The reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent reports shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer reports shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Performance Work Statement
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the proponents of ROV methodologies

Remotely Operated Vehicle (ROV) Surveys of Nearshore Stocks - California & Oregon

1. Become familiar with the reports describing the survey designs, data processing and analysis along with other pertinent information prior to review panel meeting.
2. Discuss the technical merits and deficiencies of the survey designs and analytical methods during the open review panel meeting, including revisions based upon feedback from the desk review.
3. Provide constructive suggestions for improvements.
4. Evaluate model assumptions, estimates, and major sources of uncertainty.
5. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time-frame.
6. Prepare a Peer Review Report that summarizes the Reviewer’s evaluation of the California and Oregon ROV surveys of nearshore stocks following the Terms of Reference.
7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

6.3 Appendix 2: Panel review agenda

Agenda – Panel Review of
Remotely Operated Vehicle (ROV) Surveys of Nearshore Stocks - California & Oregon

Santa Cruz, CA
February 4-6, 2020

2020 Methodology Review of Nearshore ROV Survey Designs and Methodologies

Santa Cruz Laboratory Conference Room (Room 188)
Southwest Fisheries Science Center
110 McAllister Way
Santa Cruz, CA 95060
831-420-3907

This is a public meeting, and time for public comment may be provided at the discretion of the meeting Chair. This is a technical review panel meeting, to review the scientific merits and technical applications of the proposed methodology, and will follow the Pacific Fishery Management Council's (Council) terms of reference for methodology reviews. The Methodology Review Panel will review the reports and produce a report to the full SSC, in advance of the June 2020 Council meeting in San Diego, California. Data collected using this methodology may be used in future groundfish stock assessments.

TUESDAY, FEBRUARY 4

A. 8:30 am — Call to Order, Introductions, Approval of Agenda, Terms of Reference, Assignment of Rapporteur Duties — Owen Hamel, Chair and Council Staff

B. 9:00 am Topic 1 – How are the video data generated?
   a. ROV system/capabilities/configuration
   b. Survey design

9:00 am California Presentation
10:15 am BREAK
10:30 am Oregon Presentation
11:45 am Discussion and Requests

12:30 pm LUNCH

C. 1:30 pm Topic 2- How are the data aggregated and related to habitat / spatial data?
   a. Data Processing (transect-level)
   b. Habitat information
   c. Transect segmentation – relationship to habitat

1:30 pm California Presentation
2:45 pm BREAK
3:00 pm Oregon Presentation
4:15 pm Discussion and Requests

5:00 pm Adjourn for the day

WEDNESDAY, FEBRUARY 5

D. 8:30 am Topic 3 – Analytical Methods
   a. Abundance indicies; density estimation
   b. Abundance Estimation
   c. Length data and conversion of numbers to weight

8:30 am California Presentation
10:30 am BREAK
10:45 am Oregon Presentation

12:30 pm LUNCH

1:30 pm Discussion and Requests

2:45 pm BREAK

E. 3:00 pm Responses to Requests from Tuesday

5:00 pm Adjourn for the day

THURSDAY, FEBRUARY 6

F. 8:30 am Responses to requests from Wednesday

10:15 am BREAK
G. 10:30 am Where do we go from here?
   a. Use in Assessments and Management Advice
   b. Future work

12:30 pm LUNCH

H. 1:30 pm Further discussion and drafting of report

5:00 pm ADJOURN