Center for Independent Experts (CIE) Independent Peer Review Report

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

Alan Williams

August 2019

Executive Summary

The ROV survey methodologies applied by both the California and Oregon programs are of the highest quality – scientifically sound and robust – and are underpinned by several key strengths: the availability of extensive and high resolution MBS mapping to underpin the survey designs and data expansion; operationally sophisticated field equipment; advanced data processing capabilities; and ability to apply leading-edge geospatial modelling tools.

Fishery stock assessments for northeastern Pacific inshore rocky reef fishes have been informed in various ways by ROV-derived fish counts for many years – but methods are still evolving. There are needs to address some uncertainties, and opportunities to review and further develop some operational, technical and analytical aspects.

A current review of the operational, technical and analytical methods employed by California and Oregon, of which this review report forms part, will suggest ways in which a more formal and systematic uptake might be beneficial to stock assessment, and identify the future method developments that are desirable or necessary to support this.

Suggestions for methods development (*also identified by CA and OR report authors)

Scaling the field-of-view

It will be beneficial, for both programs, to review how the field of view of the camera is scaled and how the diminishing perspective viewing guidelines are constructed.

Transect width and ROV height-off-bottom

ROV height-off-bottom (HOB) is related to measurement error, and because it is directly related to transect width, is also influential on the species-specific probability of detection. Programs should review needs for a consistently restricted range of transect width. Evaluate existing data to determine if a bias results from ROV HOB being greater over the most rugose reef and/or reef with high and abrupt relief (including ‘backsides’).

Transect segmentation

It will be beneficial to examine the effect of small (short) segments generated by ‘gaps’ created when bad data are excluded or when transects are segmented by substratum polygons defining geological categories. Small segments create the potential for very high fish densities as the total view area approaches zero.

ROV selectivity

Literature on the topic of ROV selectivity (the reaction of fish to the platforms) should be reviewed, and the sampling design refined and experiments conducted, if necessary, to reflect knowledge and knowledge gaps. Evaluate the Stoner et al. (2008) and Sward et al. (2019) reviews.

Fish size measurements*

Techniques based on calibrated paired-camera imagery are needed to acquire accurate fish size data. These are being implemented but data are not yet evaluated. Evaluation should include the cost of data processing (annotation) against the quantum of measurements required for different stock assessment applications.

Expansion: rugosity and substratum*

Considerable variation in total abundance estimates stem from the different ways MBS-derived covariates (substrata) can be used to expand transect fish density data to total fish abundance over broader areas. The expansions are particularly sensitive to the way in which the proportion of rocky bottom is defined and mapped, and the association of fish density to this metric. Ways to more objectively assess these methods should be explored – noting that the underlying methods of bottom type classification (rugosity proxy vs geology) for rasters differs between states. A stronger link between relief/rugosity and 'hard bottom' may help – possibly by capture of additional relief attributes in ROV imagery.

Expansion: terrain attributes (and spatial error)*

The distributions and abundances of rocky reef fishes are expected to be influenced by seabed terrain; however, terrain metrics had inconsistent correlations with fish density across statistical distributions and across species (California data). This is potentially due to the terrain attributes not capturing the fine scale attributes of reef geomorphic complexity (heterogeneity) that are ecologically important. Adequately capturing this is an important aspect of the methodology to explore further, including by evaluating (as suggested by California authors) deriving terrain attribute variables from CSMP depth data at alternative raster grid resolution. Part of this exploration is to simultaneously consider the role of uncertainty (error) in overlays of spatial data at fine scales that establish the relationships of individual fish observations with habitats in images.

Spatial statistics*

There are several ways in which additional approaches may improve the analyses:

- Explore removal of data outside species' ranges to reduce over-dispersion in data.
- Consider power analysis to identify sample sizes required to achieve lower variances.
- Evaluate improved post-stratification by resampling random selections of segments.
- Analyse spatial autocorrelation and the potential effect on variance estimates.
- Have additional focus on flexible splines in GAMs for area expansions.
- Evaluate utility of k-fold cross-validation, and use of RMSE other tools.
- Evaluate repeated transects in a spatio-temporal model to address temporal variability.

Other models and environmental variables
SDM models using additional environmental variables may provide alternative ways to improve prediction success or cross-validate expansions; state-scale (transboundary) analyses through time could include variables describing oceanographic variability.

Model testing
Model validation should be one objective of future sampling programs – by sampling across the gradient of predicted densities.

Individual reef and seascape perspective
This review has focused on spatial expansion of fish density data to large spatial scales; information packaged at the scale of individual reefs or reef clusters will also have utility in stock assessments.

Transect design
One possibility for future survey designs is to consider spatially-balanced methods to generate transect designs using randomization where the probability of sampling each cell in a spatial grid is user-defined (the cell inclusion probabilities). This is a robust and efficient method to assess ecological patterns.

Gap filling: multi-beam data*
There are gaps in mapping coverage that limit where data can be expanded. Future designs should consider prioritisations for gap-filling. In a similar way, surveys should take account of the needs of the stock assessment process, timing the collection of new data for particular species to the extent possible.

Recommendations

1. Develop a coordinated plan for the ways in which ROV-derived fish abundances (and potentially other habitat and sensor data) can be beneficially incorporated in fishery and conservation management processes – including to address present uncertainties in the estimates of density, relative abundance and total abundance.

2. Review and prioritise the opportunities for method development and enhancement identified here and by the authors of the California and Oregon reports to address these uncertainties.

3. Standardise the ROV-based methods across states to make them interoperable.

4. Fund and implement the research accordingly.

Contents

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 (PFMC, 2017, Agenda Item E.2, Attachment 1, September 2017).

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, e.g. for species including Yelloweye Rockfish (Taylor & Wetzel, 2011). 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 aims of this review are, for the California and Oregon ROV survey methodologies, 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

This written review is 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 and Oregon.

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.

I have read both the California and Oregon reports in detail and referred to a number of additional reports and papers that provide relevant information (Appendix 1). My approach to structuring the review has been to summarise the methods employed by both programs and embed my responses to the ToRs within those summaries. This has been an effective way for me to navigate through the large volume of detailed methodology, is a means for the report authors to check that I have correctly understood their methods, and will hopefully be useful for the panel session.

The structure of the report provides feedback on the California and Oregon programs separately, and to the extent possible, provides an integrated view. The former report is addressed first because it was available first.

My review is impartial and independent in accordance with the PWS, OMB guidelines, and the ToRs.

3 Summary of Findings

3.1 Sampling design used in recent ROV surveys

The ROV sampling design employed in both states is designed to enumerate fishes along strip transects of 500 m length targeted at inshore areas (<70 m depth) of rocky reef. Reef areas are characterized by a predominance of hard seabed substrata – bedrock, large boulders and cobble. Survey design has a heavy reliance on pre-existing maps of seabed types generated variously by geological sampling and high resolution multi-beam sonar (MBS) mapping.

3.1.1 Are sampling designs appropriate to develop estimates or indices of abundance by species in the surveyed areas

3.1.1.1.1 California

The entire program is impressively underpinned by high-resolution MBS mapping (from the California Seafloor Mapping Program – CSMP) which provides extensive coverage of the areas within which fish abundance is being assessed (Point Conception north to the California-Oregon border). The MBS maps are used to identify rocky habitats as study areas and determine the positions of transects. These mapping data represent an essential asset in the context of this program of fishery-related work.

Collection of fish abundance data relies on a suite of cameras carried by a relatively sophisticated observational ROV. The platform is de-coupled from sea surface (vessel) motion, and a suite of sensors aid the vessel skipper and pilot to navigate the ROV along the pre-determined transect lines, to fix the vessel position, and the vessel-relative position of the platform. A large number of spatially explicit samples (10,248 transect segments) were collected.

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 most shallow transect and then spacing the remainder transects at regular intervals parallel to the first.

3.1.1.1.2 Oregon

The program is also impressively underpinned by MBS and geological mapping (from the Oregon State University Active Tectonics and Seafloor Mapping Laboratory – here called the SGH data) which provides extensive, but not complete, coverage of the areas within which fish abundance is being assessed (four major sub-regions: North Coast, Cascade, Arago and Orford). As for California, the MBS maps are used to identify rocky habitats as study areas and determine the positions of transects, and, as such, represent a vital asset enabling the predictive modelling component of the fishery-related program.

Collection of fish abundance data is by a suite of cameras carried by a relatively sophisticated observational ROV. Real time vision and a suite of sensors and enable the vessel skipper and ROV pilot to navigate the ROV along the pre-determined transect lines, and to fix the vessel position and the vessel-relative position of the platform.
Individual transects are typically 500 m long, randomly located, and non-overlapping. As for California, a large number of spatially explicit samples were collected (426 transects).

'Survey areas' are located in sub-sections of the coast around rocky areas, where there is mapping data, and/or other research interests that enable field work opportunities. 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. Analysis of density, including expansion, is restricted to the maximum depth of reefs, and this varies between locations in the range of 45 to 70 m. 'Rock buffers' – regions of ~200 m wide around reef boundaries are treated separately in analyses. Two other types of boundary boxes are added (1) those termed 'deep' which fill non-surveyed gaps within surveyed sub-regions where the seaward reef boundary is <70 m; and (2) sub-regional scale boxes outside surveyed areas to complete a 'coastwide' mapping that provides a state-scale spatial context for the ROV surveyed area. The present analysis is restricted to the survey areas, but would be possible for the entire coastwide area.

3.1.2 Are sampling designs appropriate to estimate size composition by species

Comments relevant to this question are provided in Section 3.4. In short, I don't think there are issues with the survey designs in terms of where data are taken, but there are considerations for the required accuracy of length measurements and required numbers (sample sizes). Calibrated, paired cameras that are adequately separated on the ROV, are required to take data with low and known error. This configuration has been implemented on the California ROV for a while, and under trial on the Oregon ROV, but data from these systems is not presented in the program reports. However, well-established calibration and data processing protocols will enable length data of high accuracy to be extracted from imagery. The answer to how many data are required depends on the application. It seems improbable, given the modest to low numbers of fish per species that are seen, that image-derived length data will satisfy the requirements of models that need length frequency data covering multiple cohorts at a point in time. However, data may have some potential to reveal qualitative ecological patterns, such as locations or concentrations of relatively large fish on reefs.

3.1.3 Are sampling designs appropriate to expand these metrics to areas outside of those surveyed

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. Many sites have been surveyed, and collectively they appear to well-represent the fishery area and some marine parks. There are gaps in mapping coverage that limit where data can be expanded, and future designs should consider prioritisations for gap-filling. In a similar way, surveys should take account of the needs of the stock assessment process, timing the collection of new data for particular species to the extent possible. There are a variety of technical issues that have a bearing on expanding transect-scale data more broadly, but these are more related to data processing and analysis, and discussed in Sections 3.2 and 3.3.

One possibility for future survey designs is to consider work by Foster et al. (in review) who developed methods to generate transect designs using randomization where the probability of sampling each cell in a spatial grid is user-defined (the cell inclusion probabilities). This provides a robust and efficient method to assess ecological patterns.

3.1.4 Are sampling designs appropriate for use in assessment models as indices of abundance or otherwise, or for use in management procedures?

To be suited to stock assessment models and management procedures, survey designs must satisfy a rather long list of requirements. They need to be quantitative, repeatable (past and future), spatially representative of the fishery (and other managed areas), provide metrics suited to stock assessment and for the species of management interest, have known or knowable uncertainties, provide enough samples for the needs of analysis, and be justifiable on a cost-per-sample basis (although in this case there are no scientifically and operationally tractable alternatives).

I believe both surveys being reviewed here meet these requirements – although with a couple of qualifications. First, survey designs vary between states and so repeatability is on a state-specific basis. Additional comments about this are in the next section. Otherwise, and despite the considerable efforts made to date, the sample sizes (numbers of observed individual fish) are low for many species. For example in the case studies presented, the state-scale expansions are based on 230 & 301 individual gopher rockfish (California design & model expansions, respectively), and 180 yelloweye rockfish (Oregon). It would be good to consider some sort of power analysis to determine the numbers of transects required to achieve lower variances in overarching analyses of fish density based on future ROV observations.

3.1.5 Recommendations/suggestions for improvements to sampling designs

My major recommendation is for standardization of designs. Incorporating ROV data more formally into the fishery stock assessment process is necessary – but is also a strategic decision. It makes sense to have comparable data across states, and more ideally across population ranges. Many species are transboundary - across state borders and I guess further afield in the northeastern Pacific. Standardisation of design means that survey methods, survey tools and survey timing should produce sharable data – irrespective of project-specific objectives.

Habitat mapping underlays are critical for expanding ROV-derived data from transects to larger scales. Although the coverage is extensive in both California and Oregon, it does not cover the depth range of interest here, or deeper depths. Future mapping surveys would ideally have a strategic plan for priority gap-filling articulated for the fishery application – and linked to this, a strategic plan for sampling future sites for objectives that contribute to the robustness of the entire data set.

Model validation should be one objective of future sampling programs – by sampling across the gradient of predicted densities. For example, in the California data there are few ROV observations from areas predicted to have the highest gopher rockfish densities (see CA Figs 32 & 33).

3.2 Video and data processing tools/methods, and methods to determine total area surveyed for each transect.

3.2.1 Are the video and data processing methods scientifically sound and robust?

3.2.1.1 TRANSECT AREA SWEPT

The primary aims of these ROV-based methods are to acquire spatially explicit counts of individual fish along strip transects that enable fish densities to be estimated; these data can then be extrapolated across broader areas as indices and/or estimates of abundance for species populations. Fish count and area swept are the two fundamental elements of data underpinning the ROV-based programs. Both programs have used sound and robust methods, but there are some suggested needs and opportunities for fine-tuning.

Two conditions need to be satisfied to quantify the 'area swept'. First, is to incorporate technology and methodology that enable the ROV transect length (distance) x the width (camera field of view) to be measured with known error. Second, is to ensure that field data collection can be successfully and consistently implemented, i.e. that the ROV can be repeatedly flown at a near-constant heading, speed and height off bottom, to maximise collection of data that meet the standards for analysis. A set of relevant parameters that enables comparison of methods between the two states is summarised in Table 1.

3.2.1.1.1 California

Collection of fish abundance data used a suite of cameras carried by a relatively sophisticated observational ROV. The platform is de-coupled from sea surface (vessel) motion, and a suite of sensors aid the vessel skipper and pilot to navigate the ROV along the pre-determined transect lines, to fix the vessel position, and the vessel-relative position of the platform. Data are well managed in a relational database with time-code alignment of data and formal quality control processes.

Very importantly, the methodology employed in this program is founded on an early phase of field experiments; this represents a major strength. Experiments were used to determine the accuracy and precision in estimates of transect distance and width, and to test and validate field protocols (Karpov et al., 2006):

The accuracy of transect length estimates was determined by comparing different methods to measure the distances between several recognizable bottom features spread along the seabed over ~200 m linear distance. Linear (baseline) measurements from

Table 1: Selected characteristics of ROV surveys by California (CA) and Oregon (OR) showing methods in common and if standardized: [similar, Y, (Y) or different, N] together with their possible consequence. Potentially important differences between survey methods highlighted in orange.

Characteristic	CA	OR	Possible consequence	Standard?
HD camera	Y	Y		Y
Consistent camera thru time	?	N	Detection probability	?
Dome port	?	Y	Field of view	N
Parallel lasers	Y	Y		Y
Artificial lighting	N	Y	Detection probability	N
Altitude (ranging sonar/altimeter)	Y	Y	Both used same way?	Y
MBS mapping available	Y	Y	CA - 84% transects	Y
MBS habitat classification methods	Y	Combined UpLith	Methods comparable?; OR - evolved during survey period	N
ROV speed	0.5 to 1 kt	0.5 to 1 kt		Y
ROV height off bottom	<4	0.5 to 1.5	Linked to transect width	N
ROV geolocation (USBL)	Y	Y	Good smoothing protocols for both	Y
Transect length	500m	500m (300m)	Mostly the same	Y
Transect location	Random	Random	OR - non-random excluded	(Y)
Transect direction	Near-parallel to isobaths	Variable	Differ in direction	?
Transect width (m)	1.5 - 3	2-5	Key difference. OR - variable and wider range	N
Segment scale	20 m	20 m	But many smaller	(Y)
Substratum types	%hard	primary & secondary	Some diffs in defining clast size	N
Gap definitions	Fragments	Clear categories	CA applied a 60% rule (12m) as cut off	(Y)
Estimate of spatial error	3 to 6 m	$\pm$4 m (ROV only)	derivation not explained	(Y)
Sites representative	Y	Y		Y
Sites in MPAs	Y	Y		Y
Time series repeats	N?	Some (Redfish rocks)	Generally co-located	?
Effect of depth	?	Orford, Arago	transects cover extent of reef, with exceptions	?
Terrain metrics	Y	N		N

multi-beam maps of known precision were compared with (1) distance of ROV end-to-end (linear) track; (2) distance of ROV tracks plotted from smoothed USBL navigation data; and (3) ROV velocity-based estimates of distance. Overall, the comparisons showed small differences (error) compared to the baseline distance: 2-4 m greater estimated distance across methods, and the linear ROV track measurement lowest at <2 m error. Because errors did not increase with distance, the % error decreased with increasing transect length.

The accuracy of transect width estimates was determined using two different tools. First, parallel lasers projected into the centre of the field-of-view of the forward looking standard video camera were used to scale the width of the video monitor screen. Second, ranging sonar (altimetry) was used to apply a correction (*1.3) to the known width of camera viewing angle. Karpov et al. (2006) tested both methods over relatively smooth and relatively rugose seabed and found there was a strong linear relationship in the width measurements between methods, but that laser widths were smaller than sonar widths. Close to the bottom, this between-tool difference was within the measurement error of the sonar (<0.2 m). However, over rugose seabed and when the ROV was further off bottom (approximately >3 m, equivalent to a viewing width of >4 m), the difference between laser and sonar was greater, and beyond 4 m off bottom the lasers were frequently no longer visible. Thus, field results supported robust decision-making in regard to (1) a calibration of methods, (2) choice of method (it is highly desirable to routinely use sonar-based estimates of width rather than laser-based due to the relative ease and low cost of post-processing data), and (3) a rationale for setting a maximum height off bottom as a cut-off for accepting data. A height off bottom of ~4 m, corresponding to a rise in measurement error and limit of laser penetration, was consistent with a height judged to be the limit for consistently detecting and identifying fish.

One missing element, and a potential weakness, in the explanation of the width (W) estimate is the way in which the camera viewing angle was factored into the sonar-based estimated of distance (D), W_s=1.3D (Karpov et al., 2006 p.81). First, I don’t think there is a description of the camera optical properties in the reference material, but the DOE Vector M4 cameras appear to have flat viewing ports which would mean there will be significant peripheral distortion. Second, there is no mention of an in-water calibration. Thus, it is not clear how the diminishing perspective viewing guidelines for the forward-looking camera are constructed, how the distortion affects the area viewed, or the degree to which fishes were consistently detected and identified across the full transect width. There is reliance on the field-of-view perspective screen overlay to estimate range as well as width. Range is also used to determine whether a fish was within the ROV transect, and precisely where it was in relation to the substratum at the scoring screen mid-point. These details are important because fish data come from the forward-looking camera which is far more susceptible to these uncertainties than the downward-looking camera. (As a post-script to the width considerations, I didn't see any mention of the way in which platform pitch affected the field of view, and how it was dealt with, e.g. via data reduction. Perhaps this was not an issue if heave and pitch was eliminated by the depressor weight, but this is worth documenting in the context of cross-survey compatibility and future surveys).

3.2.1.1.2 Oregon

Technical details provided for the Oregon surveys show the video and data processing tools and methods used are generally similar to those used in California: an observational ROV, sufficiently instrumented in terms of navigation, image and sensor data collection, was used successfully to implement a program of strip transects over rocky reef to generate scientifically sound results (Table 1). Also, similarly to the California program, data are well managed in a relational database with time-code alignment of data and formal quality control processes – in this case, with quality assurance based on extensive reviews of video segments.

In terms of assessing the methods used for survey area calculations, it was not possible to drill down as far into the Oregon methods. Details of the methods used in California came from a peer-reviewed paper that reported results of experiments to quantify ROV transect distance and width (Karpov et al. 2006). However, whilst the Oregon fish observation program appears to have commenced before the California surveys, data in this analysis are all from after the experimental Californian experiments undertaken by Karpov et al. (2006) and both programs used DOE vehicles, so I assume therefore, that the experiences and lessons learned from the experimental work were factored in to the Oregon program. As briefly reported now, positional accuracies were assessed in relation to baselines between known bottom features, and differences (errors) estimated. These appear slightly higher (positional accuracy $\pm$4 m; ROV transect distance $\leq$10% greater than linear) than those reported by Karpov et al. (2006). However, the Oregon data include piloting deviations and other artifacts that are removed from data to eliminate any possibility of bias. (I guess these estimates are also more representative of real field conditions that when the Californian experiments were undertaken).

There is good description of the methods used to determine the accuracy of transect width. These include some details on the process (including camera calibrations) to fit the diminishing perspective viewing overlay for video annotation; the way this process had been rolled out to match the series of cameras used over the duration of the program; and that ROV stability – especially pitch – had been accounted for when calculating and applying transect width estimates to data.

3.2.1.2 FISH SPECIES DETECTION

All sampling gears have selectivity characteristics that may be general and/or specific to individual species and/or environments; to the extent possible, they need to be accounted for in stock assessments. In these ROV programs, the main issue related to gear selectivity is the reaction of fish to camera platforms. There may be behavioural response to ROVs, perhaps species-specific, to the platform because it is moving, generating noise and may carry artificial lights (Oregon surveys). A second issue is whether or not a fish species can be consistently detected with the same probability; this may depend on how visible (or conversely, cryptic) the species is, and whether there are ontogenetic differences in distribution or behaviour.

3.2.1.2.1 California

The methods give a clear rationale for choice of species based on several characteristics: their depth and habitat distributional overlap with survey areas; species believed not to show avoidance behaviour to cameras; species that have sufficient 'visibility' (i.e. are not cryptic); those that have a demersal habit (i.e. not semi-pelagic or schooling off bottom); those that are unambiguously identifiable to species in imagery; and species that are sufficiently abundant to be suited to statistical analysis. Priority was rationalized primarily in relation to near-term stock assessments, or in the case of kelp greenling, that there were no alternative methods to estimate abundance. The questions relating to choice of species are considered further, together with Oregon's, in Table 2 below. The more general issues related to selectivity (the reaction of fish to camera platforms), especially to ROVs because they are moving, generating noise and may carry artificial lights (Oregon surveys), are considered below.

The issues have been addressed by CA, in part, through criteria for adequate sampling conditions, the sampling methodology itself and post-processing methods. Video collected data was only used for density calculations when visibility was sufficient to view the entire video field of view at least 2 m in front of the ROV. During the course of a transect, the angle of the ROV camera relative to the substrata was adjusted by the pilot to maintain an oblique field of view with the horizon slightly below the top of the viewing area thereby

Table 2 List of fishes of interest for abundance assessments by ROV and information on reasons for their selection in the California (CA) and Oregon (OR) programs. First eight species are those of highest apparent interest to one or other program; remainder ordered in taxonomic sequence. Coloured cells indicate notional properties of species lists as follows: from top 8 species, green = priorities for both CA & OR, and/or data provided, yellow = lowest priorities for OR (targeted species, but rare); among other species, those considered poorly suited to analysis of abundance, pink = cryptic or semi-pelagic, orange = co-mingling species

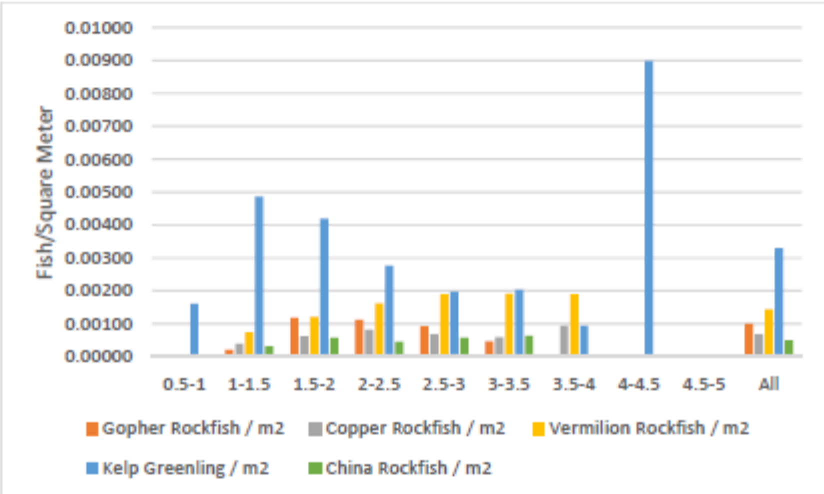
			PRIORITY FOR ROV SURVEY					MINOR/EXCLUDED		
			California			Oregon		OA	CA	Rationale
Seq	Common name	Scientific name	Index	Estimate	Rationale	Target [1]	Key [2]			
1	Kelp Greenling	*Hexagrammos decagrammus*	X	X	Common, no assessment	X	*			
2	Lingcod	*Ophiodon elongatus*	X			X	*			
3	Quillback Rockfish	*Sebastes maliger*	X	X		X	*			
4	Yelloweye Rockfish	*Sebastes ruberrimus*	X			X	*			
5	Gophor Rockfish	*Sebastes carnatus*	X	X	Assess in 2019					
6	Copper Rockfish	*Sebastes caurinus*	X	X	Assess in 2021	X		X		Rare
7	Vermillion Rockfish	*Sebastes miniatus*	X	X	Assess in 2021	X	*	X		Rare
8	China Rockfish	*Sebastes nebulosus*	X	X	Previous assessment	X	*	X		Rare
	Wolf-eel	*Anarrhichthys ocellatus*				X				
	Striped Surfperch	*Embiotoca lateralis*				X				
	Red Irish Lord	*Hemilepidotus hemilepidotus*				X				
	Pacific halibut	*Hippoglossus stenolepis*				X				
	Ratfish	*Hydrolagus colliei*				X				
	Painted Greenling	*Oxylebius pictus*				X				
	Starry Flounder	*Platichthys stellatus*				X				
	Pile Perch	*Rhacochilus vacca*				X				
	Cabezon	*Scorpaenichthys marmoratus*				X	*	X		Cryptic
	Brown Rockfish	*Sebastes auriculatus*	X	X		X				
	Widow Rockfish	*Sebastes entomelas*				X				
	Yellowtail Rockfish	*Sebastes flavidus*				X	*	X		Co-mingle
	Rosethorn Rockfish	*Sebastes helvomaculatus*				X	*			
	Black Rockfish	*Sebastes melanops*	X			X	*	X	X	Semi-pelagic
	Blue or Deacon Rockfish	*Sebastes mystinus or S. diaconus*	X			X	*	X	X	Semi-pelagic
	Tiger Rockfish	*Sebastes nigrocinctus*				X				
	Canary Rockfish	*Sebastes pinniger*	X			X	*	X	X	Co-mingle

1 Suited to visual census + sufficiently abundant

2 *length estimated

ensuring that fish behaving evasively in front of the ROV could be detected. CA found the vast majority of demersal rockfish were relatively unresponsive, but others were relatively cryptic or unavailable to the ROV in mid-water (Table 2).

Many individuals were unidentifiable because they were too far away from the ROV to be adequately illuminated and were most often smaller schooling individuals positioned off the seafloor. California explored the effect of transect width (related to height off bottom by a constant) on fish counts and found that cut-offs were required to eliminate varying detection probability in the narrowest and widest transects (see below). Accordingly, data from transects widths outside the range 1.5 to 3 m were not used in analysis.



Fish density x segment width x species (CA Fig. 7)

3.2.1.2.2 Oregon

A broader range of target species was identified; these included four primary species and two used as case studies for analysis in the Oregon report (see Table 2). All the comments made above about ROV selectivity apply here, although the Oregon ROV also carries artificial lights which is another potential influence on selectivity (see notes above).

I don’t think the cut-offs for transect width were applied in the same way as the California data. Widths are very variable in the Oregon data (Figure 1) with differences appearing to include generally wider transects when the ROV is further off the bottom over rocky bottom vs sandy bottom (e.g. Arago vs Perpetua, respectively). Given the California finding, an approach using cut-off distances at both ends of the range of widths should be considered.

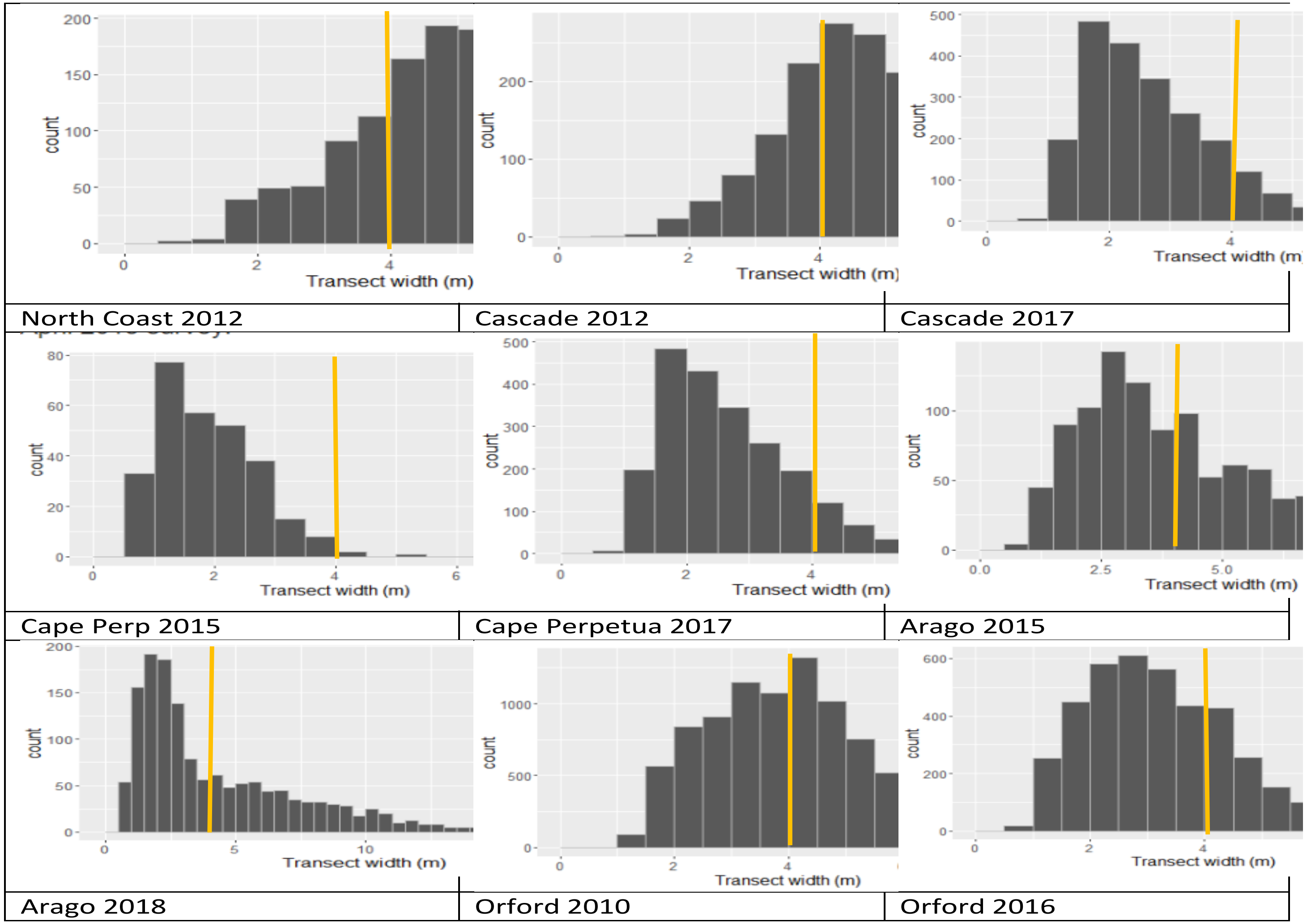


Figure 1: ROV transect width range at Oregon sites (4 m width indicated in each)

3.2.1.3 LOCATIONS AND SEGMENTATION OF TRANSECTS

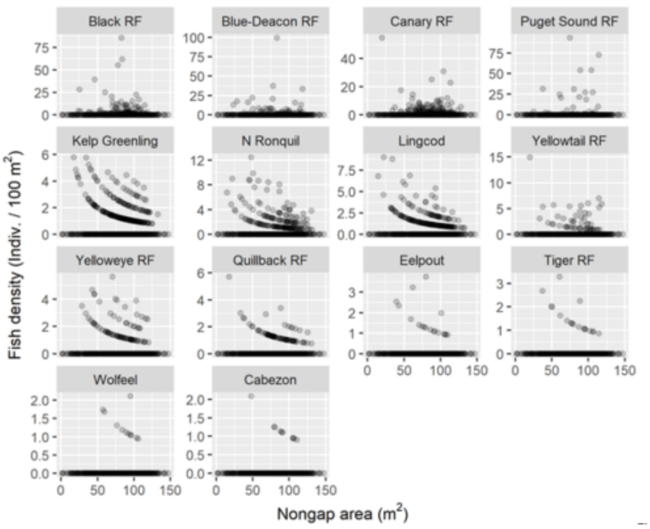
Fish (and habitat) data may be analysed at the scale of entire ROV individual transects; data aggregated at this scale are similar to the integrated samples provided by trawls. However, there is opportunity to sub-divide the data taken along transects to capture habitat, substratum and biological variation occurring at finer spatial scales (and, where necessary, to capture changes in operational parameters). Sub-division, or segmentation, is possible because data are continuous (1-sec intervals) along transects. There are many strategies for defining the lengths of segments (or defining the spatial units of analysis), and the advantages and disadvantages for ROV data are well reviewed in both the California and Oregon reports.

3.2.1.3.1 California

Individual transects were 500 m long, and data were analysed at the 20 m segment scale. Segmentation of transects was considered at several scales ranging from 'microframes' (individual 1-sec observations) to transect, and including two intermediate scales defined by fixed area of 25 m^2 (varying distance and width) and fixed distance of 20 m (varying area). The implications (opportunities and constraints) of each scale were considered in relation to spatial extrapolation (expansion) – including resolution of covariate data, statistical analysis, and potential biases. All alternatives were considered in relation to errors in data geolocation, several of which arise from operational factors (vessel and vehicle locations) and covariate data (especially mapped resolution of substrata and habitats). A particularly important consideration for the Californian data was the alignment of video data with terrain covariates. These are generated from MBS bathymetry using neighborhoods of adjacent cells and may compound uncertainty about spatial alignment.

3.2.1.3.2 Oregon

Transects were mostly 500 m long (300 m from one survey), and data were analysed at transect and 20 m segment scale. Transect scale (summed to produce a single total fish count from a total view area) were used to model and predict total fish abundance and associated variability. Segment-scale data were used to provide substrate-specific density and abundance estimates based on expanding the densities observed in each substratum type. Exploratory analysis using intermediate-sized 20 m long segments (@5 m wide = approximately $\leq$100m^2) showed <u>strong relationships between segment length and fish density</u> for many species (see below). This is an artifact stemming from observations of fish in very small segments. Small (<<20 m) segments are generated by (1) 'gaps' created when bad data are excluded (seven identified data descriptors, e.g. when the ROV is too high off bottom); or (2) when transects are segmented by substratum polygons defining the project-specific 'Lith 3' classification of geological categories that have variable, and sometimes short, extents. Small segments create the potential for very high fish densities as the total view area approaches zero. As a consequence, full transects were chosen as the unit of analysis – mainly to eliminate artifacts associated with small segments, but also to reduce spatial auto-correlation of observations along transects. I am uncertain if/ how this effect was dealt with in California data.



Fish density x segment (nongap) area x species (OR report Fig).

3.2.1.4 SUBSTRATUM, HABITAT AND TERRAIN

3.2.1.4.1 California

I found the description of the substratum determination and scoring difficult to follow in part. The summary categories of hard, mixed and soft habitats are realistic and intuitive, and derived from a familiar (CMECS?) geological grain size classification. However, the way in which video was scored and the metric recorded are not entirely clear. Thus, the aim was to determine % cover, but the overlapping <u>presence</u> of each substratum type is scored, apparently by independent viewings of the video— N=6 views? It is then not clear how combined presences are converted to % cover and how this is used – despite having defined criteria for start and end points of substratum patches (>2 m gaps or no substratum >20% for 3 m). Although I haven't fully grasped the methodology, it is flexible and the ability to classify segments as ecotones is potentially useful.

However, these ROV substratum data were not used in analysis of density distributions because they cannot be consistently mapped to the habitat classifications inherent in the CSMP multi-beam mapping data that is essential to extrapolate the survey site data to broader areas. Instead, these ROV data (relationships between fish density and hard bottom) inform the CSMP habitat classifications generated by terrain (rugosity) algorithms.

There is some uncertainty (error) in all spatial data, and this is relevant to overlays of data at fine scales, e.g. when associating individual fish with habitats in images. In the CA data, the error in ROV observations was estimated to be 3 to 6 m (calculation not explained); MBS mapping data (in depths <85 m) are gridded in 2 x 2 m cells; spatial analysis of terrain (rugosity) metrics uses neighborhoods of at least 3 x 3 cells; and segment level data has to be represented at a point in space (the segment centroid) for analysis. For the CA analysis, 9 base-resolution (2 x 2 m) grid cells were aggregated to form larger (6 x 6 m) cells for analysis to represent habitat at a suitable scale – not too finely to generate unusable and/or meaningless complexity, but not too coarsely to overly smooth rocky relief. Two simple but excellent diagrams (p. 17 and 18) visualize the CA data overlays: a 20 m long transect overlaid on neighborhoods of 3 x 3 and 5 x 5 analysis cells.

Metrics used were those that emphasise variation amongst the cells in a neighborhood because they provide a more meaningful covariate of fish distribution and are less susceptible to spatial location errors (ROV observation + transect location within neighborhood) – depth range, SD depth, slope, SD slope, surface area to planar area. Conversely, BPI, curvature and relative difference to mean (RDMV) were not used because they emphasise the properties of a single (central) cell within the neighborhood. (Note, however, that RDMV also appears to have been used).

3.2.1.4.2 Oregon

The method for recording substratum type from video was different to that used in California. Here, the substratum classification had minor differences relating to clast size and sediments classes, but more importantly the hard bottom categories differed in having two classes of boulders and a slightly different differentiation between boulder and bedrock (4 vs 3 m). Data were recorded at 1-sec intervals, and captured the primary (dominant) and secondary substratum types. Patches were bounded by any >1 sec change in the interpolated data. These substratum data were not used in analysis of density distributions because they cannot be consistently mapped to the habitat classifications inherent in the multi-beam mapping data essential to extrapolate the survey site data to broader areas.

The SGH (Surficial Geologic Habitat) multibeam bathymetry and backscatter data were categorized as primary (“Lith1”) and secondary (“Lith2”) substrates, which are combined into a single mixed classification (“Lith3”) with a multitude of possible combinations. For this analysis, categories of Lith3 were combined into six main super-categories termed “UpLith” (rock, rock mix, boulder, large mix, cobble, small mix, soft), and then into two: rocky reef (all hard substrates sized cobble and larger) and non-reef.

3.2.2 Are the video & data processing methods appropriate for available data?

3.2.2.1.1 California

The methods applied to the video and sensor data are appropriate, and make the most of the available data. Thus, the survey design and its implementation are supported by extensive (near comprehensive) coverage of finely resolved (mostly $4m^2$ grid) multibeam data for the survey areas. The subsequent determination of area swept and fish density are appropriately based on time-stamping, geo-referencing and scaling data provided by a modern suite of instruments and sensors with high accuracy and high logging frequencies – all of which appear well managed in a dedicated database. Collectively, the results showed that operational protocols were sound and that measurements of area are robust. The velocity-based method provides a fall-back if tracking data are unavailable. The subsequent spatial analysis techniques, such as determining the association of fish density with gridded covariate data, are underpinned by suitably post-processed multi-spatial scale data products. Methods were developed to account for (inevitable) weaknesses such as gaps in data. The choice of spatial sampling units took account of data characteristics.

3.2.2.1.2 Oregon

Similar to California, the methods applied to the video and sensor data are appropriate, and make the most of the available data. The survey design and its implementation is supported by extensive coverage of finely resolved (mostly $4m^2$ grids) multibeam data for the survey areas. The subsequent determination of area swept and fish density are appropriately based on time-stamping, geo-referencing and scaling data provided by a modern suite of instruments and sensors with high accuracy and high logging frequencies – all of which appear well managed in a dedicated database. Collectively, the results showed that operational protocols were sound and that measurements of area are robust. The subsequent spatial analysis techniques, such as determining the association of fish density with gridded covariate data, are underpinned by suitably post-processed multi-spatial scale data products. Methods were developed to account for (inevitable) weaknesses such as gaps in data. The choice of spatial sampling units took account of data characteristics.

3.2.3 Recommendations/suggestions for improvement to methods.

3.2.3.1 HEIGHT OFF BOTTOM (TRANSECT WIDTH) IN RELATION TO FISH DETECTION

ROV height off bottom (HOB) is related to measurement error (low or distant heights have higher error measurement); the limit of laser penetration; a maximum height of ~4 m was consistent with a height judged to be the limit for consistently detecting and identifying fish; and because it is directly related to transect width, which is influential on the species-specific probability of detection. California data showed a relationship between transect width (HOB) and probability of fish detection and limited analysis to data from transects of ~<3 m width (~4 m HOB), but I am uncertain how this effect was dealt with in Oregon data. It will be beneficial for both programs to determine potential biases in the HOB effect on probability of detection for the full range of species of interest.

3.2.3.2 SELECTIVITY OF ROV SAMPLING FOR FISHES

The literature on the topic of ROV selectivity for fishes should be reviewed, and the sampling design refined, if necessary, to reflect knowledge and knowledge gaps. The evaluation should include reactions to ROV stimuli that may not have been rigorously quantified since the Stoner et al. (2008) review that concluded, "*associated biases have not been rigorously quantified and remain mostly undocumented or anecdotal*, and that *much remains to be learned about how bias varies among species, age groups, different vehicles, and operating conditions*." There is a recent systematic review of visual assessment of fish assemblages by ROV (Sward et al. 2019), and there are several parameters of potential interest: speed (Trenkel et al. 2004); light (Widder et al. 2005); height off bottom (Laidig et al. 2013; sound (Stoner et al. 2008); behavior Uiblein (2011), Linley et al. (2013), Makwela et al. (2016); Somerton et al. (2017). The potentially most influential bias is *when fish respond at a distance greater than can be detected by observers or video cameras* (Stoner et al., 2008) – as acknowledged in the California report.

3.2.3.3 PERSPECTIVE VIEW FOR FISH COUNTING AND CLASSIFYING SUBSTRATUM

It will be beneficial, for both programs, to review how the diminishing perspective viewing guidelines are constructed, how the distortion affects the area viewed, and the degree to which fishes were consistently detected and identified across the full transect width (for estimates of transect width as well as range).

3.2.3.4 DENSITY ARTEFACTS STEMMING FROM SEGMENT LENGTH

Oregon data showed a strong, often artefactual, relationship between segment length and fish density for many species stemming from observations of fish in very small segments. Small segments create the potential for very high fish densities as the total view area approaches zero. It will be beneficial to examine this relationship further, including because I am uncertain how this effect was dealt with in California data.

3.3 Proposed methods to develop indices or estimates of abundance

3.3.1 Are methods scientifically sound and robust?

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 mostly ideally across the range of a managed stock. Many statistical 'species distribution' or 'habitat suitability' modelling techniques are available to do this. Those contemplated 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 the transect-scale abundances to broader areas. Multi-beam mapping of seabed bathymetry and hardness, sometimes available at high resolution over large areas, is particularly important to a study such as this that examines biological responses to seabed characteristics. 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.

3.3.1.1.1 California

Analysis plan

The California program has used an extensive and well-documented set of methods to evaluate, analyse and expand the data – a relatively simple design-based method, and a complex model-based method. These methods are sound and robust.

Multi-beam sonar (MBS) mapping provided the basis for expanding density estimates for species in both design-based and model-based estimates of abundance; in the model-based method these data were used to generate raster covariates: presence of hard bottom defined using seafloor roughness as a proxy, and terrain (relief) attributes derived from bathymetry.

- <u>Design-based method</u>: used the derived area of hard bottom habitat in a given depth and latitude; <u>expansion of density data = entire area with CSMP data</u>.
- <u>Model-based method</u>: 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, <u>expansion of density data = the area with CSMP data for terrain metrics</u>.

A large number (10,248) of segments contribute to the full data set, but here data were analysed for only five species, and one, gopher rockfish, was used as a case study. The table below summarises the sample sizes – segments and observations, including the reduced data set with terrain metrics:

	Without terrain			With terrain	
Species	**Segments**	**No. fish**		**Segments**	**No. fish**
China rockfish	112	119		91	98
Kelp greenling	1385	1683		1134	1395
Copper rockfish	286	286		208	217
Gopher rockfish	351	429		317	391
Vermillion rockfish	624	1293		501	813

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

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, five distributions were compared and models tested using the AIC: Regular Poisson, zero-inflated Poisson, quasi-Poisson, zero-inflated quasi-Poisson and negative binomial. In addition, the binomial was run using counts converted to presence/absence data.

Deviance and selection in backward stepwise model selection was evaluated for gopher rockfish to see if important variables were consistently identified between methods (except for zero-inflated models).

Expansion methods

There are many options for this stage of the analysis, and what is implemented depends on the data available, trade-offs between strengths and weaknesses of different models, and varying assumptions. Here, expansion used design-based and model-based methods with CSMP data and 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.

In outline, 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 – either by capturing its finer spatial scale to up-scale hard bottom habitat area, or by pairing fish density with terrain attributes. In the latter situation, 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.*

3.3.1.1.2 Oregon

Analysis plan

The Oregon program has used a relatively simple, but nonetheless sound and robust set of methods to evaluate, analyse and expand the data. The methods are well-documented, including with an excellent set of figures that enable drill-down to fine levels of data relationships. As for California, MBS data are used to identify rocky habitats as study areas and determine the positions of transects. Unlike for the California analysis, no GIS generated terrain variables were used in a model-based analysis.

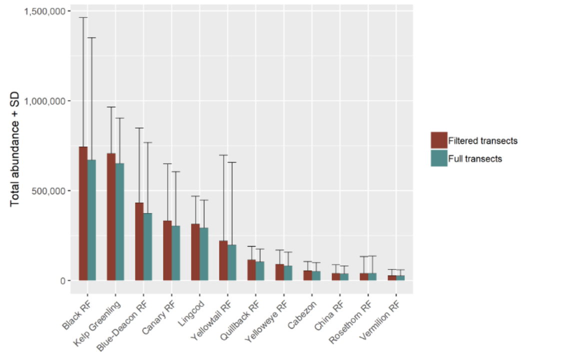
Analysis used both the full transect data and the filtered rock-only transect data to generate abundance estimates; both datasets were applied to the rock area. Note, the question of 'where is the reef boundary' is treated here by defining a rock buffer = 211 m, beyond the rock polygon. Full transect 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. However, the relatively arbitrary size of the buffer was acknowledged as an obstacle to interpreting its associated abundance estimate.

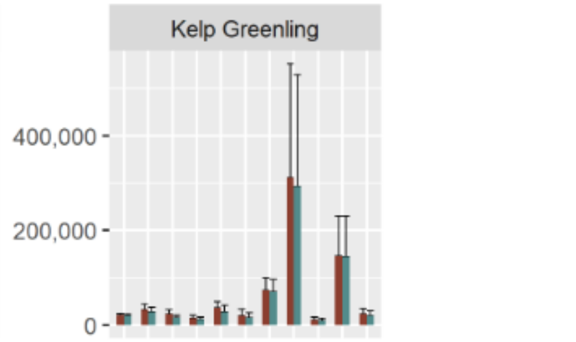
Density estimates were calculated for rock-buffers using full transect estimates. However, I was not certain if this was used, and which method was employed: expanding substratum-specific density into the rock buffer has a problem with estimates of variability depending on how data (n= 839 segments) are aggregated. (Method 1: density x substratum pooled across all sites; Method 2: density x substratum within each site). Method 2 uses a higher number of groupings and the resulting SD is lower; this is a source of uncertainty.
Full transects were the unit of analysis (mainly to eliminate high density artifacts associated with small segments, but also to reduce spatial auto-correlation).

A large number (335 of 426) of transects contributed to the data analysed.

Models and expansion

Filtered (rock-only) transect data were used for expansion.





Rock-only data 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 were suited to modelling. But note, relatively large influence of Cape Arago – e.g. illustrated by kelp greenling (see figure above) – 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 (no supporting information provided – but also apparent in CA data), 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

- GAM - variables = Site (each of the 11 sites, "Location" in the model output), a smoothed function of Depth, and Year [as for California, view area observed included as an offset because it varied between segments].
- The model predicted abundance at a set of grid points covering the entire mapped area of rocky reef in the surveyed regions, with the random Year effect set to zero.

Year was added 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. The North Coast 2012 survey was conducted during the same time frame and on the same vessel as the 2012 Cascade Head survey, so also considered appropriate to include Year to address any oceanographic variability that might have affected both surveys similarly.

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.

3.3.2 Are the methods appropriate for the available data?

Methods applied by both states are appropriate for the available data; they have considered a range of factors to implement the calculation of fish density, and for the identification and treatment of variables used for data spatial expansion (prediction). California has used a more extensive suite of analyses, with comparisons of alternative methods and models backed up with diagnostics. Oregon anticipates this in the future – and clearly there is good scope for collaborative and standardized future development of methods. Nonetheless, there are some areas of method development required to address uncertainties (see below).

3.3.2.1.1 California

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 (area observed 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 <u>proportion of rocky bottom was consistently highly significant across species and data distribution types</u>, <u>but</u> there were no 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, 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. <u>Latitude, CSMP hard bottom and take were consistently significant variables across species and distributions</u>.

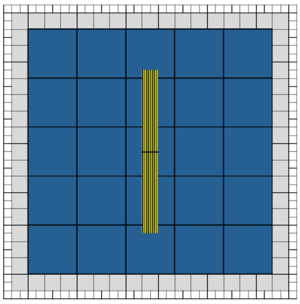
Expansion methods

Five different methods were considered, and two 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; Model 2 represents an improvement by also capturing variations in density related to depth and latitude. The strength 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. Depth and latitude are significantly correlated with fish distributions, especially gopher rockfish, and so it is a no-brainer to use Method 2. 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 in 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. (Here the blue cells made up by 3 x 3 m base level raster grey cells – with a 20 m transect segment superimposed.) 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, may provide a more accurate estimate with less variance. It makes sense to bypass this approach because more sophisticated models are possible – and can tap into the terrain data. However, the terrain metrics have inconsistent relationships with density across species, and there are new and unknown spatial errors introduced, so in some ways this relatively simple model has attraction. It is also provides a more direct comparison with Method 2, which might be valuable.

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 would bring in the terrain attributes – which is a potential strength if it captures relationships between fish density and rugosity. 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 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.

3.3.2.1.2 Oregon

As for California, a negative binomial distribution was determined to provide the best fit to the data for both species – although this appeared considerably better for kelp greenling than yelloweye rockfish. Repeated k-fold cross-validation was used to examine the performance of various potential models. A resampling method was used to generate confidence interval for the predicted total abundance. GAMs were used to establish relationships between fish density and environmental variables. A more formalized model selection procedure is anticipated in the future. Known issues that will be explored in potential subsequent models include the potential for correlations within repeated transects (e.g. the Port Orford 2010 and 2016 surveys) to affect the overall variance. Particularly with the addition of the new 2018 data, there is more opportunity for application of a spatio-temporal model such as VAST, to explicitly address variability over time.

3.3.3 *Are indices or estimates of abundance consistent with input data and population biological characteristics?*

I am not able to comment in relation to population biological characteristics because I am unfamiliar with the biology and ecology of the suite of fishes in focus here, and there is insufficient time available in this review period to gain this knowledge. However, it is possible to make a few observations in terms of the consistency of abundance estimates with input data. Most obviously, there are wide ranges in the estimates for each of the three species analysed that stem from alternative steps in methodologies.

3.3.3.1.1 California

The design and model-based estimates of total abundance for expansion using only hard (rocky) bottom (yellow cells) are similar – but I ended up being uncertain whether the same total habitat areas were compared. Much more importantly, however, estimates with >10% rocky threshold applied (green cells) are considerably (~64%) larger. It is clear the estimate for hard bottom only will underestimate total abundance, whilst a 10% threshold is extremely low. Somewhere between is a more realistic threshold – but where, 20%?? This relationship between fish density and % of hard bottom is a key source of uncertainty in expansions.

Final absolute abundance estimates: California (gopher rockfish)

Method	Characteristics	No. fish	var	Fish (mt)	var
Design #2	Hard only; combined depths	549,995		267	
	Hard only; depth stratified	578,621		281	
	>10% hard; combined depths	915,161		444	
	>10% hard; depth stratified	957,227		464	
Model #4	GLM	498,714		239	
Model #5	GAM	539,759	13,143	259	6.3

3.3.3.1.2 Oregon

Compared to the area-scaled (rocky reef) expansion method to estimate mean density (green cells) and total fish abundance (yellow cells), the GAM model predictions are (1) substantially lower for the abundant kelp greenling and (2) substantially higher for the relatively uncommon yelloweye rockfish. Both are potentially linked to a mismatch between the emphasis of sampling depth and depth distribution of species. Thus, yelloweye increase in abundance with increasing depth in the study area but there was higher intensity of sampling in shallower regions. Conversely, for kelp greenling, a species that is more abundant in shallower depths at some sites, but especially in the large and disproportionately influential Cape Arago region.

Final absolute abundance estimates: Oregon

	Metric	**Kelp greenling**			**Yelloweye rockfish**			**Notes**
1	Overall observed density (indiv. per 100 m^2)	0.751			0.055			Total surveys nongap fish count/ total nongap area
2	Mean transect density (indiv. per 100 m^2)	0.856			0.045			All transects, N=335; weighted by view area
3	Total abundance - expanded (individuals)	1,153,513			60,008			Mean transect density x area of rocky reef
4	GAM predicted total abundance (individuals)	643,336	380,137	1,103,530	137,384	56,475	454,755	Sum of predictions at grid points $\pm$ 2 S.E.
5	GAM prediction of density (indiv. per 100 m2)	0.477	0.282	0.819	0.102	0.042	0.337	Abundance and S.E. converted to density
6	GAM resampled post. 95% CI	510,637	867,024		100,130	290,816		CI s for predicted total abundance

3.3.4 Recommendations/suggestions for improvement to methods.

Collectively, these results and those of Young and Carr (2015), show that considerable variation in total abundance estimates stem from the ways in which environmental covariates from MBS mapping data are used to expand transect data to broader areas. Expansions using only 2 x 2 m grid cells classified as hard (rock) will underestimate population abundance because fish are also found on mixed bottom (proportion of rock <100%) and are unaccounted for. (The binary hard vs soft classification is derived from a vector ruggedness measure and arbitrary cut-off). The design-based methods in the California and Oregon programs explored remedies for this in two ways: (1) (California) by creating grid cells of mixed (hard/soft) bottom at coarser (30 x 30 m neighborhood) scale – but a threshold for the proportion of rocky bottom is not easily defined, and the final abundance estimate is very sensitive to this; or (2) (Oregon) by bounding reef areas with a buffer of intermediate fish density estimated from transects over mixed substrata – but the spatial extent of a buffer is arbitrary, and density variability depends on whether density data are aggregated within site or across sites. Both methods are worthy of further exploration to examine the sensitivity of estimates to the parameters used, and to explore alternative methods for expanding the original derivation of hard vs soft using a rugosity proxy in the CSMP data at raw (2 x 2 m) scale – and see below.

The model-based expansions appear to insufficiently capture reef geomorphic complexity (heterogeneity) that was shown in shallower Californian waters to be strongly and consistently important in explaining spatial variation in rocky reef fish densities – including gopher rockfish (Young and Carr, 2015b). This is attempted in the California methodology, using what appear to be intuitively suitable spatial (neighborhood scales) for reef associated fishes with high habitat fidelity, and using terrain metrics that emphasise rugosity measures internally to neighborhoods, rather than externally (to broader seascape features). However, they had inconsistent correlation with density across statistical distributions and across species.

This was interpreted as being due to the spatial error in pairing terrain attributes with transect centroids, the terrain attributes not capturing the fine scale attributes of rugosity that are ecologically important, or to small sample sizes. All are plausible, and perhaps additive. Whilst the unit scale of analysis is impressively fine (<30 x 30 m) in the context of reef extents and the scale of the study area, it would be worth exploring using other terrain attributes (e.g. one based on multiple thresholds for rugosity, Young and Carr 2015a), especially if they can be represented at finer (18 x 18 m or raw raster scale) with some exploration of sensitivity to the ±3-6 m spatial error in centroid position. Further, the SDM model of Young and Carr (2015b) show other environmental variables may improve prediction success – although theirs (kelp and wave energy) may be less suited to the depth range of interest here. It is worthwhile to note that the numbers of observations supporting analyses are relatively small: 230 & 301 individual gopher rockfish (design & model expansions, respectively), and 180 yelloweye rockfish. Model fits might be improved by removing segments that fall outside the species' depth range – at least to reduce over-dispersion (as suggested by the California authors). As suggested earlier, it would be beneficial to consider some sort of power analysis to determine the numbers of transects required to achieve lower variances in overarching analyses.

Otherwise, there is scope for other methods development (as suggested by the authors).

- Improving post-stratification using methods to better identify suitable strata based on estimates of density and variance from resampling random selections of segments.
- Analysis of spatial autocorrelation and the potential effect on variance estimates.
- Additional focus on flexible splines in GAMs for area expansions; this will be needed, for example, for species with non-linear depth distributions.
- Further examination of inter-annual variability given the unexplained variability between different surveys at the same Oregon sites (Orford, 2010 & 2016; Cascade, 2012 & 2017) and the potential for confounding from oceanographic variability.

- The utility of k-fold cross-validation (and improvements on the use of RMSE therein) vs. other potential tools.
- Exploration of the potential for correlations within repeated transects (e.g. the Port Orford 2010 and 2016 surveys) to affect the overall variance. Particularly with the addition of the new 2018 data, there is more opportunity for application of a spatio-temporal model such as VAST, to explicitly address variability over time.

3.4 Proposed methods to estimate size compositions of individual fish

3.4.1 Are methods scientifically sound, robust, and consistent with accepted practices?

3.4.1.1.1 California

The ROV was equipped with four standard resolution (640 by 480) color cameras: two locally recorded stereo cameras for highly accurate measurements of size. However, to date, the data for fish size (total length) was estimated by the video observer with the use of two parallel lasers placed 10 cm apart aimed to hit the seafloor in the center of the video viewing screen of the forward-facing camera. Fish sizes were estimated to the nearest cm and when possible tagged for future stereo sizing. Criteria for stereo sizing included fish orientation (almost perpendicular) and distance (within two meters) to the cameras.

I didn't see any results presented on fish sizing.

3.4.1.1.2 Oregon

Of the 24 fishes identified to species, length data were estimated for 12 species and one unidentified rockfish category in relatively broad ranges: < 10 cm, 10 – 30 cm, 30 – 60 cm, > 60 cm. Sex was recorded for kelp greenling, which exhibit distinct sexually dimorphic colouration. Fish size was recorded only where fish are broadside near the lasers against a scaleable background.

Sparse data were recorded for all 12 species (total range = 5 to 141 individuals per species); but there was a relatively large tally for the unidentified rockfish category (2727 individuals). With the exception of yelloweye rockfish, which are readily identified as juveniles, it can be inferred that the individuals identified to species (and therefore the only ones included in this analysis) are larger than young of the year. The category "Unidentified rockfish" is largely composed of small (< 10 cm) juvenile rockfish, and included to show the ability of the ROV to detect this size class in large numbers.

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.

3.4.2 Are the methods appropriate for the available data?

There are numerous problems with collecting fish size (length) data in these programs – to date. The background issues are those linked to detection probability discussed in Section 3.3; the technical issues are linked to using lasers that require fish to be broadside and in the field of view close to and at the range of the lasers and/or at a scale-able background. These limitations were acknowledged and explained in the Oregon report.

It is difficult to comment on the appropriateness of methods because little has been done with the data. No results were presented for the California program, whilst the Oregon program provided only a brief summary. There, over the entire program, few data were collected – even for the most abundant species. These data may have some potential to reveal qualitative ecological patterns, such as locations or concentrations of relatively large fish, but the numbers of individuals are so small it would not justify the cost of acquiring length data systematically using this method. There seems little potential to supplement length frequency data taken from extractive survey (trawl) methods for stock assessments.

3.4.3 Are estimated size compositions consistent with input data and population biological characteristics?

There are insufficient data to comment on this.

3.4.4 Recommendations/suggestions for improvement to methods.

The suggestions I have are already mooted in both reports: incorporate techniques based on calibrated paired-camera imagery. These cameras are built into the California program ROV camera suite, but it appears the data have not yet been acquired from imagery. The Oregon program has also recently added stereo cameras to the ROV and calibrated them using SeaGIS software – the industry standard. Post-calibration sizing error assessments indicated < 0.5% error in size estimates using a known size grid in tank trials. This is in line with my expectations for a well-configured system.

Further trials of the Oregon system, together with work on the Californian data, should overcome the technical hurdles for acquiring data with standard operating procedures and known error, and enable an insightful evaluation of the potential scope for collecting length data. The challenges then will be to make the cost of data processing (annotation) affordable (justifiable), and to better understand detection probability. Both aims will be served by collecting length data on future ROV deployments because it adds little to the fixed costs of operation (calibrations, and some additional data management overheads). My recommendation is to continue to deploy the stereo systems as part of all future ROV surveys.

3.5 Identify potential impediments to developing independent indices or estimates of abundance using these ROV surveys and incorporating them into stock assessments.

3.5.1 Are the results informative, robust and could they be incorporated into stock assessments?

ROV surveys results are informative because they are the only means of providing abundance data from reef habitats where a diverse nearshore groundfish fauna is targeted by commercial and recreational fishers. Potentially alternative methods are not well suited to this task. Acoustic methods are non-extractive and may work well for aggregations of single species with known target strength characteristics and a benthopelagic habitat – if there is spatio-temporal ecological information to inform a survey design. However, many of the rocky reef species off California and Oregon are benthic species that mostly occupy the acoustic dead zone, and are not detected by acoustic sampling. Various hook and line methods may provide abundance indices, but uncertainties related to gear selectivity and other factors, and because they are extractive, mean they are unsuited to inshore rockfishes. In contrast, visual census using an ROV provides a suitable and non-extractive method of survey that is quantitative, and can take data from all seabed types. ROV surveys can therefore 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.

Responses to questions about robustness of results and incorporation into stock assessments are covered in Sections 3.1 to 3.4.

3.5.2 Are there limitations to the use and incorporation of the indices or estimates of abundance into stock assessments?

Every sampling methodology has its limitations, but I don't believe those linked to ROV surveys for reef-associated fishes are an impediment to applying the technique successfully in the context of nearshore rocky reef fishes. As noted above, there are no good alternatives. ROV survey methods to derive fish abundance are still evolving, and whilst I'm not aware of the full history of their use in stock assessments, it appears they have been used for some years in other locations, e.g. since 2012 in Alaskan waters (Olson et al. 2016). There are several great strengths for both the California and Oregon programs: the availability of extensive and high resolution MBS mapping to underpin the survey designs and data expansion; operationally sophisticated field equipment; advanced data processing capabilities; and ability to apply leading-edge geospatial modelling tools.

There are a variety of technical methodological issues that should be addressed to make the density and abundance estimates more robust – several have been suggested by the authors of the California and Oregon reports, and I have added some others. These are discussed in more detail elsewhere, and listed in Section 3.6, immediately below.

3.6 Provide guidance on key improvements in survey design or modeling approaches for future considerations.

3.6.1 Consider research recommendations provided and make any additional recommendations or prioritizations warranted.

Below is a summary list of potential improvements; those flagged with an ~~asterix~~ asterisk were also identified by report authors:

3.6.1.1 SCALING THE FIELD-OF-VIEW

It will be beneficial, for both programs, to review how the field of view of the camera used for fish counts is scaled (for counts, to estimate range and width): how the diminishing perspective viewing guidelines are constructed, how the distortion affects the area viewed, and the degree to which fishes were consistently detected and identified across the full transect width.

3.6.1.2 TRANSECT WIDTH AND ROV HEIGHT-OFF-BOTTOM

ROV height-off-bottom (HOB) is related to measurement error, and because it is directly related to transect width, also influential on the species-specific probability of detection. California data showed a relationship between transect width (HOB) and probability of fish detection and limited analysis to data from transects of ~<3 m width (~4 m HOB). Widths are very variable in the Oregon data; an approach using a restricted range of widths should be considered.

Linked to this is the operational aspect of acquiring data over the most rugose reef and/or reef with high and abrupt relief. These areas may be particularly ‘productive’ but may be under-sampled if the ROV is flown further off the bottom, and when off the bottom over ‘backsides’. An evaluation of existing data may determine if a material downward bias results from these effects.

3.6.1.3 TRANSECT SEGMENTATION

Despite aiming for a standard segment length (20 m in both programs), small (short) segments are generated by ‘gaps’ created when bad data are excluded or when transects are segmented by substratum polygons defining geological categories. Small segments create the potential for very high fish densities as the total view area approaches zero. Oregon data showed a strong, often artefactual, relationships between segment length and fish density for many species due to observations of fish in very small segments. It will be beneficial to examine this relationship further, including for California data.

3.6.1.4 ROV SELECTIVITY

Literature on the topic of ROV selectivity (the reaction of fish to the platforms) should be reviewed, and the sampling design refined and experiments conducted, if necessary, to reflect knowledge and knowledge gaps. During survey, ROVs are moving, generating noise and may carry artificial lights (Oregon surveys). The evaluation should include reactions to ROV stimuli that may not have been rigorously quantified since the Stoner et al. (2008) review that found, 'the potentially most influential bias is when fish respond at a distance greater than can be detected by observers or video cameras' – as acknowledged in the California report. There is a recent systematic review of visual assessment of fish assemblages by ROV (Sward et al. 2019).

3.6.1.5 FISH SIZE MEASUREMENTS*

Techniques based on calibrated paired-camera imagery are needed to acquire accurate fish size data. Paired cameras form part of the California ROV camera suite, but it appears the data have not yet been acquired from imagery. The Oregon program has also recently added stereo cameras to its ROV and calibrated them using industry standard software. Further trials of the Oregon system, together with work on the Californian data, are needed – together with methods to make the cost of data processing (annotation) affordable (justifiable), and to better understand detection probability.

3.6.1.6 EXPANSION: RUGOSITY AND SUBSTRATUM*

Considerable variation in total abundance estimates stem from the different ways MBS-derived covariates (substrata, and potentially depth in Oregon data) can be used to expand transect fish density data to total fish abundance over broader areas. The expansions are particularly sensitive to the way in which the proportion of rocky bottom is defined and mapped, and the association of fish density to this metric. A stronger link between relief/ rugosity and 'hard bottom' may help – possibly by capture of additional relief attributes in ROV imagery. Ways to more objectively assess these methods should be explored; noting that the underlying method of bottom type classification (rugosity proxy vs geology) for rasters differs between states. Variation in total abundance estimates potentially were linked to a mismatch between the emphasis of sampling depth and depth distribution of species in Oregon data.

3.6.1.7 EXPANSION: TERRAIN ATTRIBUTES (AND SPATIAL ERROR)*

The distributions and abundances of rocky reef fishes are expected to be influenced by seabed terrain; accordingly, metrics of rugosity are calculated and implemented in the model-based expansions employed by California. However, and disappointingly, they had inconsistent correlation with fish density across statistical distributions and across species. This is potentially due to the terrain attributes not capturing the fine scale attributes of reef geomorphic complexity (heterogeneity) that are ecologically important. These should be evident from the associations seen in ROV footage. Adequately capturing this is an important aspect of the methodology to explore further, including by evaluating (as suggested by California authors) deriving terrain attribute variables from CSMP depth data at alternative raster grid resolution to examine whether the correlations are scale-dependent and whether alternative resolutions result in more consistent correlations. Part of this exploration is to simultaneously consider the role of uncertainty (error) in overlays of spatial data at fine scales that establish the relationships of individual fish observations with habitats in images. Uncertainty stems from several sources: ROV position; point (centroid) representation of transect data; MBS gridding; and terrain neighborhood aggregation (resolution).

3.6.1.8 SPATIAL STATISTICS*

There are several ways in which additional approaches may improve the analyses:

- Exploring whether model fits for over-dispersion are improved by removing segments that fall outside the species' depth ranges.
- Consider power analysis to determine the numbers of new transects required to achieve lower variances in overarching analyses of fish density noting that sample sizes (numbers of observed individual fish) are low for many species.
- Improving post-stratification using methods to better identify suitable strata based on estimates of density and variance from resampling random selections of segments.
- Analysis of spatial autocorrelation and the potential effect on variance estimates.
- Additional focus on flexible splines in GAMs for area expansions; this will be needed, for example, for species with non-linear (e.g. dome shaped) depth distributions.
- Evaluating the utility of k-fold cross-validation (and improvements on the use of RMSE therein) vs. other potential tools.
- Evaluating repeated transects (e.g. Port Orford, 2010, 2016, 2018?) to examine how correlations affect the overall variance and to explore the opportunity for a spatio-temporal model such as VAST to explicitly address variability over time.

3.6.1.9 OTHER MODELS AND ENVIRONMENTAL VARIABLES

SDM models (e.g. Young and Carr, 2015b) show other prediction techniques using additional environmental variables may provide alternative ways to improve prediction success or cross-validate expansions (noting their variables, kelp and wave energy, may be less suited to the depth range of interest here). A consideration for state-scale (transboundary) analyses through time could include variables describing oceanographic variability, e.g. unexplained signals of inter-annual variability in surveys at Oregon sites (Orford, 2010 & 2016; Cascade, 2012 & 2017).

3.6.1.10 MODEL TESTING

Model validation should be one objective of future sampling programs – by sampling across the gradient of predicted densities. For example, in the California data there are few ROV observations from areas predicted to have the highest gopher rockfish densities.

3.6.1.11 INDIVIDUAL REEF AND SEASCAPE PERSPECTIVE

This review has focused on spatial expansion of fish density data to large spatial scales – large sub-regions of individual states; future planning has whole-of-state, and potentially transboundary, scales in mind. It also seems probable that information packaged at different spatial scales will have utility in stock assessments. Thus, relative or absolute abundances at the scale of individual reefs or reef clusters might be of interest to understanding the natural spatial and temporal variation in local scale productivity – particularly in no-take areas.

3.6.1.12 TRANSECT DESIGN

One possibility for future survey designs is to consider spatially-balanced methods to generate transect designs using randomization where the probability of sampling each cell in a spatial grid is user-defined (the cell inclusion probabilities). This is a robust and efficient method to assess ecological patterns.

3.6.1.13 GAP FILLING – MULTI-BEAM DATA*

There are gaps in mapping coverage that limit where data can be expanded. Future designs should consider prioritisations for gap-filling. There is as-yet incomplete mapping coverage, for example, parts of the Oregon south coast regions, specifically the large reef complex at Rogue Reef, and areas south of the Port Orford region. Other areas are of potential interest: the 0-10 m depth range, where data from non-ROV sources will need to be factored in for

species whose distributions extend to the shoreline; deeper than 70 m, for species whose distributions extend deeper – the large Arago reef complex off Oregon is one example.
In a similar way, surveys should take account of the needs of the stock assessment process, timing the collection of new data for particular species to the extent possible.

3.6.2 Provide recommendations on research and monitoring that could improve the reliability of, and information provided by, future visual surveys for assessments of nearshore stocks.

This is covered in the section above.

3.7 Provide a brief description on other aspects of the survey design or model estimation not described above.

The only two things to add in this section are to re-emphasise the need for standardization of methods to the extent possible. Incorporating ROV data more formally into the fishery stock assessment process is a strategic decision. It makes sense to have comparable data across states, and more ideally across population ranges. Many species are transboundary - across state borders and/or more broadly in the northeastern Pacific. Standardisation of design means that survey methods, survey tools and survey timing should be inter-operable, i.e. produce data that can be combined – irrespective of project-specific objectives.

There are several operational, technical and analysis developments required to refine the ROV methodology; these future developments will need project funds.

4 Conclusions

The ROV survey methodologies applied by both the California and Oregon programs are of the highest quality – scientifically sound and robust – and are underpinned by several key strengths: the availability of extensive and high resolution MBS mapping to underpin the survey designs and data expansion; operationally sophisticated field equipment; advanced data processing capabilities; and ability to apply leading-edge geospatial modelling tools.

But methods to estimate fish abundances over large areas from observations along transects are still evolving, and there are needs to address some uncertainties and opportunities to review and further develop some operational, technical and methodological aspects. Perhaps the single biggest opportunity is to standardize the approaches. Because methods differ, there are prospective benefits and efficiencies from alignment, and opportunities to compare and contrast approaches.

Future research work should be directed by priorities that consider the benefits for uptake by fishery and conservation managers, balancing effort and funds against the uncertainties and knowledge gaps identified during the review process.

5 Recommendations

1. Develop a coordinated plan for the ways in which ROV-derived fish abundances (and potentially other habitat and sensor data) can be beneficially incorporated in fishery and conservation management processes – including to address present uncertainties in the estimates of density, relative abundance and total abundance.
2. Review and prioritise the opportunities for method development and enhancement identified here and by the authors of the California and Oregon reports to address these uncertainties.
3. Standardise the ROV-based methods across states to make them interoperable.
4. Fund and implement the research accordingly.

6 Appendices

6.1 Appendix 1: Bibliography of materials consulted for review

Report 1: Methods for using remotely operated vehicle survey data in assessment of nearshore groundfish stocks along the California coast

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

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., Williams, A., and R. Przeslawski (in review). Spatially-Balanced Designs for Transect-Based Surveys. Biometrics.

Laidig, Thomas E. and Krigsman, Lisa M. and Yoklavich, Mary M. (2013) Reactions of fishes to two underwater survey tools, a manned submersible and a remotely operated vehicle. Fishery Bulletin, 111(1), pp. 54-67.

Makwela, M. S., S. E. Kerwath, A. Götz, K. Sink, T. Samaai and C. G. Wilke (2016). "Notes on a remotely operated vehicle survey to describe reef ichthyofauna and habitats - Agulhas Bank, South Africa." Bothalia - African Biodiversity & Conservation 46: 1-7.

Olson, A., Stahl, J., Van Kirk, K., Jeenicke, M. and S. Meyer (2016). Assessment of the demersal shelf rockfish stock complex in the southeast outside district of the Gulf of Alaska. 565-608.

Rooper, Christopher & H. Martin, Michael & L. Butler, John & Jones, Darin & Zimmermann, Mark. (2012). Estimating species and size composition of rockfishes to verify targets in acoustic surveys of untrawlable areas. Fishery Bulletin- National Oceanic and Atmospheric Administration. 110. 317-331.

Somerton, D. A., Williams, K., Campbell, M. 2017. Quantifying the behavior of fish to a towed camera system using stereo optics and target tracking. Fishery Bulletin, U.S. 115:343–354.

Stoner, A. W., Ryer, C. H., Parker, S. J., Auster, P. J., and Wakefield, W. W. (2008). Evaluating the role of fish behavior in surveys conducted with underwater vehicles. Canadian Journal of Fisheries and Aquatic Sciences 65, 1230-1243.

Sward, D., Monk, J. and Barrett, N. (2019). A Systematic Review of Remotely operated Vehicle Surveys for Visually Assessing Fish Assemblages. Frontiers in Marine Science, 6:134. doi: 10.3389/fmars.2019.00134

Taylor, I.G. and W. Wetzel Status of the U.S. yelloweye rockfish resource in 2011 (Update of 2009 assessment model).

Trenkel, V. M., Lorance, P., and Mahévas, S. (2004). Do visual transects provide true population density estimates for deepwater fish? ICES Journal of Marine Science: Journal du Conseil 61, 1050-1056.

Trenkel, VM & Francis, RICC & Lorance, Pascal & Mahévas, S & Rochet, MJ & Tracey, Dianne. (2004b). Availability of deep-water fish to trawling and visual observation from a remotely operated vehicle (ROV). Marine Ecology-progress Series - MAR ECOL-PROGR SER. 284. 293-303. 10.3354/meps284293.

Uiblein, F. (2011). Deep-sea fish behavioral responses to underwater vehicles: differences among vehicles, habitats and species. In ‘Autonomous Underwater Vehicles’. (Ed N. A. Cruz.) pp. 225-238. (InTech: Rijeka, Croatia.)

Widder, E. A., B. H. Robison, K. R. Reisenbichler and S. H. D. Haddock (2005). "Using red light for in situ observations of deep-sea fishes." Deep Sea Research Part I: Oceanographic Research Papers 52(11): 2077-2085.

Williams, K., Rooper, C.N. and Towler, R.D. 2010. Use of stereo camera systems for assessment of rockfish abundance in untrawlable areas and for recording Pollock behavior during midwater trawls. Fishery Bulletin, U.S. 108: 352-362.

Young, M. and Carr, M.H. (2015a). Assessment of Habitat Representation across a Network of Marine Protected Areas with Implications for the Spatial Design of Monitoring. PlosONE, 1-23.

Young, M. and Carr, M.H. (2015b). Application of species distribution models to explain and predict the distribution, abundance and assemblage structure of nearshore temperate reef fishes. Diversity and Distributions, 1-13.

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

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 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.

This methodology review will likely provide feedback on the initial development of materials and guidance for future ROV surveys and the development of 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 desk review of these survey methodologies will be followed-up with an in-person panel review tentatively scheduled for early-December 2019.

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**.

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. Additionally, these CIE reviewers will participate in a follow-on panel review to be held in early December 2019 (See **Attachment A**).

Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Tasks for Reviewers

1) Review the following background materials and reports prior to the review:

Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact(s) in advance of the peer review, including detailed reports of previously conducted and proposed ROV surveys and analysis methods from each of the states of Oregon and California. Two weeks before the desk review, the NMFS Project Contact(s) 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 Contact(s) will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to them in accordance to the PWS scheduled deadlines specified herein. The CIE reviewers 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;
- 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) Desk Review: Each CIE reviewer 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 can not be made during the peer review, and any PWS or ToRs modifications prior to the peer review shall be approved by the NMFS Project Contact.

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

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

Place of Performance
Each CIE reviewer shall conduct an independent peer review as a desk review, therefore no travel is required.

Period of Performance
The period of performance shall be from the time of award through July 2019. Each reviewer's duties shall not exceed 10 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
No later than two weeks prior to the review	Contractor provides the pre-review documents to the reviewers

May 2019	Each reviewer conducts an independent peer review as a desk review
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
Since this is a desk review travel is neither required nor authorized for this contract.

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
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

6.2.1 Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.
2. The main body of the 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.
3. The reviewer report shall include the following appendices:
    a. Appendix 1: Bibliography of materials provided for review
    b. Appendix 2: A copy of the CIE Performance Work Statement

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

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

The specific responsibilities of each of the proponents are to:

1) 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.

2) Evaluate the sampling design used in recent ROV surveys conducted by the states of Oregon and California, addressing the following:
   a. Are sampling designs appropriate
      i. To develop estimates or indices of abundance by species in the surveyed areas
      ii. To estimate size composition by species
      iii. To expand these to areas outside of the those surveyed
      iv. for use in assessment models as indices of abundance or otherwise, or for use in management procedures?
   b. Recommendations/suggestions for improvements to sampling designs

3) Evaluate the video and data processing tools/methods and methods to determine total area surveyed for each transect.
   a. Are methods scientifically sound and robust?
   b. Are the methods appropriate for the available data?
   c. Recommendations/suggestions for improvement to methods.

4) Evaluate proposed methods to develop indices or estimates of abundance from these ROV surveys, including using habitat/substrate type and Marine Protected Area designation as covariates.
   a. Are methods scientifically sound and robust?
   b. Are the methods appropriate for the available data?
   c. Are indices or estimates of abundance consistent with input data and population biological characteristics?
   d. Recommendations/suggestions for improvement to methods.

5) Evaluate proposed methods to estimate size compositions of observed individuals of each species.
   a. Are methods scientifically sound, robust, and consistent with accepted practices?
   b. Are the methods appropriate for the available data?
   c. Are estimated size compositions consistent with input data and population biological characteristics?
   d. Recommendations/suggestions for improvement to methods.

6) Identify potential impediments to developing independent indices or estimates of abundance using these ROV surveys and incorporating them into stock assessments.
    a. Are the results informative, robust and could they be incorporated into stock assessments?
    b. Are there limitations to the use and incorporation of the indices or estimates of abundance into stock assessments?

7) Provide guidance on key improvements in survey design or modeling approaches for future considerations.
    a. Consider research recommendations provided and make any additional recommendations or prioritizations warranted.
    b. Provide recommendations on research and monitoring that could improve the reliability of, and information provided by, future visual surveys for assessments of nearshore stocks.

8) Provide a brief description on other aspects of the survey design or model estimation not described above.