Appendix C. Coastwide Pre-Recruit Indices from SWFSC and NWFSC/PWCC Midwater trawl Surveys (2001-2016)

John Field, E.J. Dick, Nick Grunloh, Xi He, Keith Sakuma and Stephen Ralston
Fisheries Ecology Division, SWFSC, Santa Cruz, CA,

Introduction

This document provides an update of coastwide pre-recruit indices of abundance developed for past stock assessment cycles (Ralston et al. 2015), using data collected during SWFSC, NWFSC and PWCC/NWFSC midwater trawl surveys for young-of-the-year (YOY) pelagic juvenile groundfish. Due to time constraints and complications related to the discovery of a problem in how past indices were developed, this document reports indices for only a handful of those species typically evaluated, with a focus on those being assessed for the 2017 assessment cycle (bocaccio, blue/deacon and yellowtail) and one relatively abundant species from which to evaluate the consequences of the computational issues in past indices (shortbelly rockfish). Some preliminary explorations of an alternative means of developing indices are also included for consideration in review panels of those assessments.

In recent stock assessment cycles, these indices have been developed with guidance from the 2006 Pre-Recruit Survey Workshop (Hastie and Ralston 2007), such that data collected by these different surveys using identical gear and methods could be pooled to develop "coastwide" indices of abundance for YOY *Sebastes* spp. (see Ralston et al. 2013, Ralston and Stewart 2013 and Sakuma et al. 2016 for reviews of data, methods, vessel comparison and select results). This was in recognition that the data collected over a longer time period (1983-present) from the "core" area of the SWFSC survey were likely to present a biased and/or imprecise representation of coastwide YOY abundance due to significant interannual shifts in the spatial distribution of pelagic juvenile YOY (Ralston and Stewart 2013). However, variable ship availability and survey effort make the development of truly "coastwide" indices for some years impossible.

Data Analysis

As in recent assessment cycles, we used only years with the most comprehensive coverage to evaluate the spatial scope appropriate for each individual stock for which an index might be developed. Figure 1 shows haul locations for the different surveys over time, for the SWFSC (1983-2016, fixed stations), NWFSC (2011, 2013-2016, fixed stations) and PWCC/NWFSC (no fixed stations) datasets. Table 1 shows the total number of hauls by 2° latitude bins (the reported latitude in the Table represents the "mean" latitude for that bin, such that latitude 46 includes hauls from 45°- 47° N) for all of the survey data when pooled together. As the years 2004-2009 and 2013-2016 included very comprehensive coastwide coverage (albeit with very little data north of 47°N), these years were used to develop "climatologies" of the spatial distribution of the catch, in order to evaluate where the majority of the catch by species took place, so that "coastwide" indices could be crafted for southerly and northerly distributed species. This time period included years of very high (2009, 2013-2016) as well as very low (but spatially variable, 2005- 2007) abundance, and thus should provide a reasonable characterization of the spatial distributions of most species.

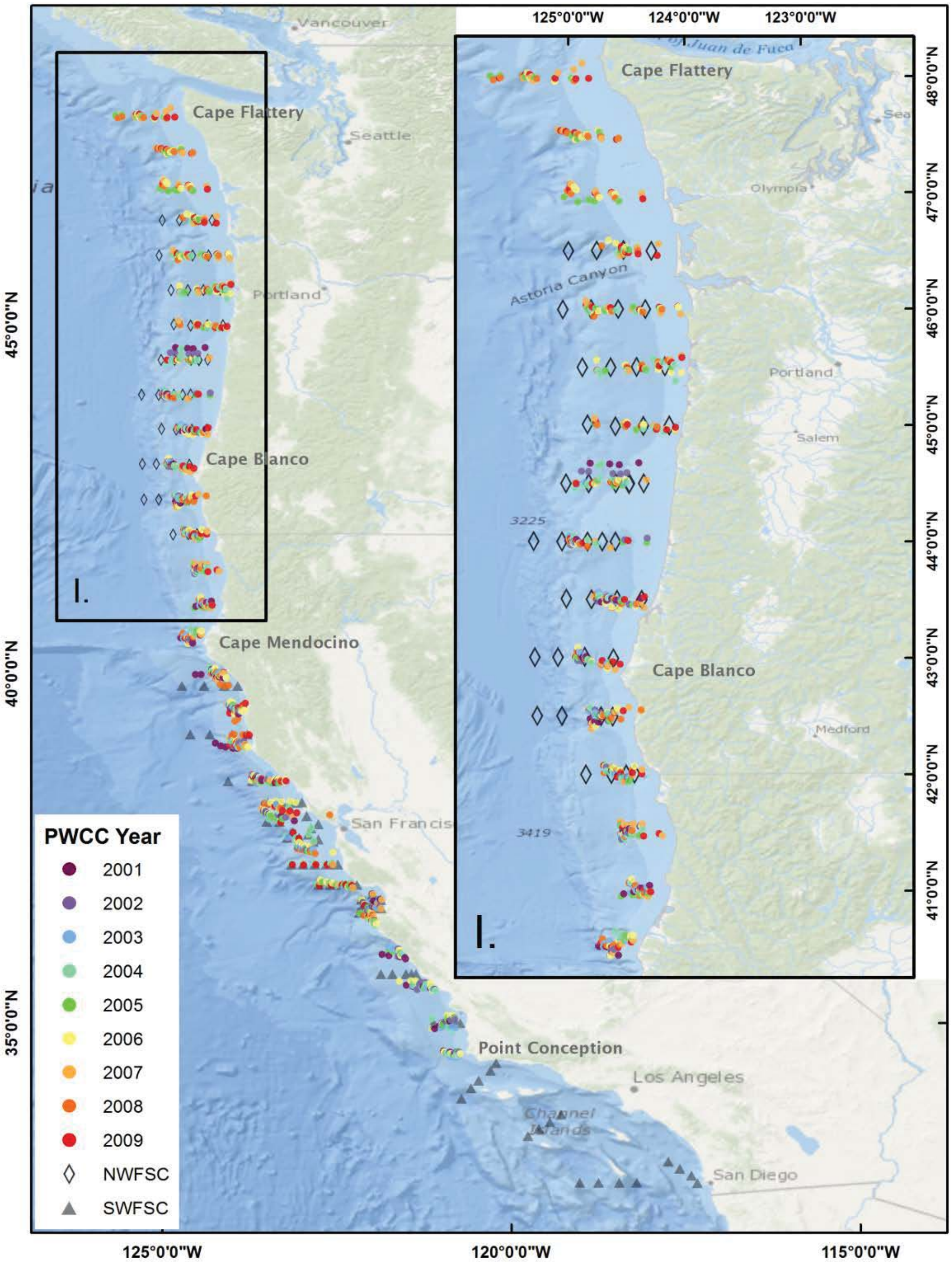

Figure 1: Station and haul locations for SWFSC, NWFSC and PWCC/NWFSC midwater trawl surveys.

Table 1: Number of hauls by year and latitude bin used to develop climatologies of spatial abundance (data prior to 2001 excluded).

only northern species
all species
only southern species

year	latbin 32	34	36	38	40	42	44	46	48	Total
2001		6	68	53	17	17	19			180
2002		6	63	52	19	21	17			178
2003		8	72	71	20	20	19			210
2004	8	27	76	74	28	20	25	20		278
2005	13	27	92	61	35	17	22	21	12	300
2006	14	24	83	86	40	21	20	22	13	323
2007	11	17	78	85	37	25	21	23	16	313
2008	13	20	43	43	37	21	22	18	15	232
2009	7	19	59	79	30	24	23	23	16	280
2010	6	15	44	52	16					133
2011			29	30	19	22	28	24	13	165
2012	3	13	51	27						94
2013	7	21	51	39	17	16	21	13		185
2014	5	13	54	57	16	15	18	9		187
2015	13	25	56	44	18	19	17	13		205
2016	12	26	56	35	6	9	20	12		176

The results of the exploration of catch rate climatologies indicated that some fairly rational generalizations could be made regarding the spatial survey extent that might represent "coastwide" coverage for the different species of rockfish. Specifically, for the "northern" species, widow rockfish (*S. entomelas*), yellowtail rockfish (*S. flavidus*), black rockfish (*S. melanops*), and canary rockfish (*S. pinniger*), the data from the years of the best truly coastwide coverage indicate that 99.7 to 100% of population abundance, as measured by spatial integration of average catch-per-unit-effort ($fish \cdot tow^{-1}$), has occurred within the 36 - 46° N latitudinal bins, representing the area between 35° and 47° N (Table 2). Thus, the best spatial coverage for these species are the years 2004-2009, 2011 and 2013-2016, as reflected by the indices developed for the 2015 assessment cycle (Ralston et al. 2015). By contrast, for blue/deacon rockfish (which have not historically been differentiated to the species level in this survey), catches were very uncommon north of 44 N, and consequently years in which the survey evaluated the region between 36 to 44 could be used for an index.

Similarly, for the "southern" species, chilipepper (*S. goodei*), squarespot rockfish (*S. hopkinsi*), shortbelly rockfish (*S. jordani*), bocaccio (*S. paucispinis*), and stripetail rockfish (*S. saxicola*), between 95 and 100% of the integrated abundance took place within or below the 40° latitude bin (e.g., latitudes 41° and south), although for bocaccio this range extended to the 42° N latitude bin with the addition of 2015-16 data. Thus, the indices developed for the 2017 assessment cycle were limited to those years that included the 32-34 latitude bins up through 42°N for bocaccio; namely 2004-2009, 2013-2016.

Prior to developing the Pre-Recruit index, the raw catch rate data were converted to standard age fish, due to substantial interannual variation in the size distribution of fish collected. To accomplish this, the length of each specimen of a species in a haul was converted to an estimated age using a linear regression of age $N = a + b \times SL$, where N is estimated age in days and SL is standard length (mm). Data used to fit all species-year regressions were generated by sub-sampling fish and counting daily otolith increments (see Woodbury and Ralston 1991). The contribution of each fish in a given haul was then age-adjusted according to:

$$N^{*}_{h,t} = N_{h,t} \exp[-M(100 - t_{hat}]$$

Where N^{*} is the number of fish in 100 day old equivalents, $N_{h,t}$, is the number of fish from haul h of estimated age t and M is the natural mortality rate of pelagic juvenile rockfish (0.04 day^{-1}; see Ralston and Howard 1995, Ralston et al. 2013). Standardized abundances were obtained by summing the number of 100 day old equivalent fishes within a haul. This effectively standardizes the contribution of all fish to a common age of 100 days, i.e., younger fish are downweighted and older fish are up-weighted. The number of age observations for each species is available in the 2015 documentation.

Following discussions during the 2006 Pre-Recruit Survey Workshop related to the strengths and weaknesses of alternative analytical approaches, indices distributed to stock assessment authors in recent assessment cycles (Ralston 2010, Sakuma and Ralston 2012) have been based on an ANOVA index, primarily because of its ability to best account for significant year x latitude interactions, and we continue this practice here. The specific form of the ANOVA mixed-effects model is:

$$\log(C_{i,j,k,l,m,n}+1) = Y_i \times L_j + Z_k + D_l + V_m + \varepsilon_{i,j,k,l,m,n}$$

with all independent variables treated as categorical. Specifically Y_i is a fixed year effect $\{Y_i \in 2001, 2002, \ldots, 2016\}$, L_j is a fixed latitudinal effect $\{L_j \in 32, 34, \ldots, 40\}$, Z_k is a fixed depth effect $\{Z_k \leq 160$ m or $Z > 160$ m$\}$, D_l is a fixed calendar date effect $\{D_l \in 120, 130, \ldots, 170\}$, V_m is a random vessel effect $[V_m \sim N(0,\sigma_v)]$, and $\varepsilon_{i,j,k,l,m,n}$ is normal error term $[\varepsilon \sim N(0,\sigma_\varepsilon)]$ for the n^{th} observation in a stratum. As in the case of the traditional ANOVA model, interactions between latitude and year were explicitly modeled.

Prior to this year, the model was fit to the data using PROC MIXED (SAS Institute Inc. 2004) and the year:latitude parameter estimates were bias-corrected, integrated over latitude, and error estimates summarized in a manner directly analogous to the traditional ANOVA approach. This year the code for developing the indices was migrated from SAS to the R programming language to facilitate future rapid computation of indices. In doing so, a non-trivial issue was discovered related to how the indices were compiled from the year:latitude results. Specifically, the model as previously run summed across latitude parameters in log space, and then backtransformed the sum for each year estimate. However, upon greater consideration it was determined that the appropriate approach is to back-transform the latitude bin results and then sum across latitudes within each year, to produce the annual index in arithmetic space. The use of log(C+1) as a response variable also introduces minor complications with respect to back-transformation to obtain means on the arithmetic scale.

As a consequence of the conflicting time series produced by these two slightly different approaches, we also developed indices based on the well-established delta-GLM model (Lo et al. 1992, Stefánsson, 1996) for these four stocks (as done in earlier assessment cycles as well as Ralston et al. 2013). This model has the greatest potential, in our view, to provide a stopgap approach to developing a YOY index until a deeper modeling exploration can be conducted. The delta-GLM components (binomial and positive models) both contained categorical covariates as described for the ANOVA, above. The delta-GLM was fit using the "rstanarm" package in R to obtain Bayesian posterior distributions of the delta-GLM index. Finally, we also report the resulting indices developed when using the VAST software package (Thorson et al. 2015) on the same data.

Results

We report results of the four modeling approaches (past implementation of the ANOVA approach, "corrected" ANOVA approach, delta-GLM, and VAST) for bocaccio (update assessment) and blue/deacon and yellowtail rockfish (full assessments). We also report results for shortbelly rockfish as this species is the most frequently encountered rockfish in the surveys, has a broad spatial distribution, and thus should provide a better basis for understanding differences in modeling results among these species.

These results are shown in Figure 2, and Table 2 provides the numerical values and the associated CVs. Importantly, upon making the correction to the calculation of the ANOVA indices, the indices for several species appear unusually "flat," particularly for bocaccio but for other species as well, suggesting that even this corrected approach is far less than an ideal means of deriving these indices. Most likely it is the log(catch+1) transformation, which is used to address the issue of large numbers of zeros in the data, that is leading to poor performance of this modeling approach, which was masked by the increased variability in the indices when the summation was done inappropriately.

Relative to the corrected ANOVA, both the delta-GLM and VAST approaches show considerably greater variability in the indices, with high and low values typically ranging from one to several orders of magnitude among different years. Differences in interannual variability between indices derived from the ANOVA and delta-type models (delta-GLM and VAST) also depend on the number of zeros in the data. For example, the corrected ANOVA approach is extremely flat relative to the other two approaches for bocaccio, a species that is fairly rare in these surveys (present in 8.5% of hauls in the nominal range during the 2001-2016 period). However, the ANOVA begins to resemble both the Delta-GLM and the VAST indices for shortbelly rockfish, a species present in a far greater fraction of hauls (34% of hauls in the nominal range during the 2001-2016 period). This lends additional support for the concerns that the log(catch+1) transformation used in the ANOVA method is inappropriate for those species that are rarely encountered in the survey.

Despite these challenges, there are some clear indications in the data, as illustrated in all modeling approaches, of very strong recruitment for some stocks and years, particularly in 2013 for all of these stocks. Such signals were also evident in the 2015 chilipepper assessment update (Field et al. 2015) as well as the 2015 bocaccio assessment (He et al. 2015) and the pending update. Given the consistency of this strong year class with recent observations, the indices should provide some utility for full assessments of blue and yellowtail rockfish this assessment cycle.

Discussion

For bocaccio, the "corrected" ANOVA result is the most consistent with the intent of what had been done in prior assessments, despite the fact that it does not indicate recruitment variability of the magnitude expected from other sources of data (e.g., fishery and survey length frequency data). Consequently, the bocaccio assessment also includes sensitivity analyses that use both the same index (not extended in time) from the 2015 model (the nominally incorrect ANOVA) as well as the indices developed using the delta-GLM and VAST approaches. As none of these approaches suggest unusually strong recruitment since the 2013 year class, which is now largely informed by length composition and other data sources, we think this is a reasonable short-term fix for the purposes of an update.

For the full assessments being conducted in 2017 (blue/deacon, yellowtail rockfish), our current preference would be to use the delta-GLM results. However, the results presented here will need to be refined for the appropriate spatial strata associated with assessment boundaries, and will likely require some additional exploration and documentation. For example, the current VAST outputs include all years regardless of the spatial coverage of the survey, which is inconsistent with previous approaches and should be interpreted with caution (we may have revised in time). The VAST indices also do not include a within-year temporal effect (period effect) to account for the seasonality of sampling, which has varied in surveys throughout the years and has been demonstrated to be an important factor for many species. Consequently, both the delta-GLM and the VAST these results should be considered preliminary, and can be revised and considered in greater detail prior to the full STAR Panels for those two species.

Our intent is to return to alternative means of developing indices, including evaluation delta-GLM models (including the VAST geostatistical approach) as more robust approaches for developing YOY recruitment indices to support West Coast rockfish assessments. Ongoing analyses indicate that in fact there is likely to be considerably more coherence in YOY abundance trends than earlier envisioned, and that the 2005-2006 period was atypical with respect to strong differences in abundance between the historical core survey area and coastwide abundance trends.

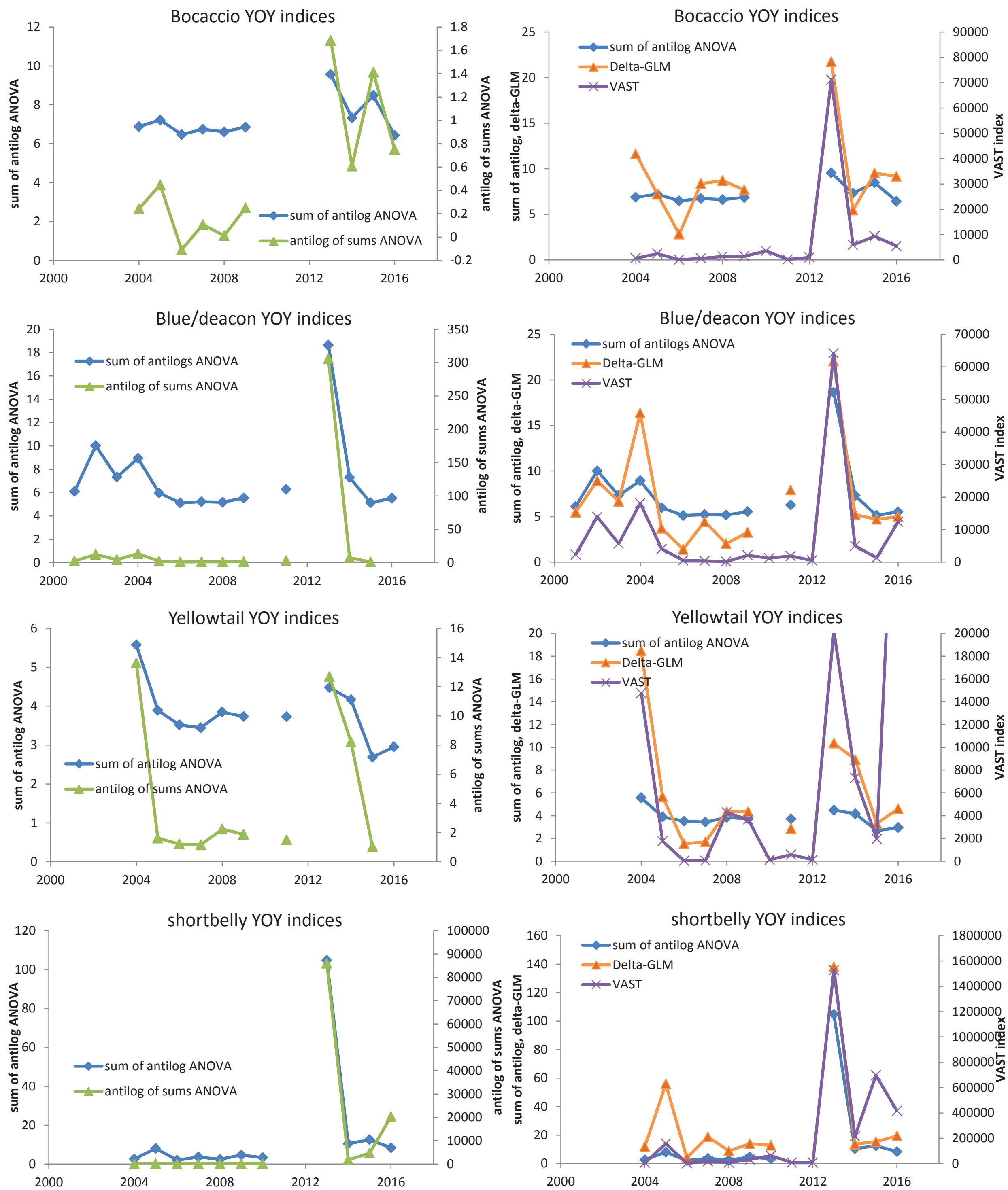

Figure 2: Comparisons of the two ANOVA based indices (using sum of the antilog values or the antilog of the sum of values for the year:latitude interaction model) for YOY rockfish (left panels) and of ANOVA, Delta-GLM and VAST indices for YOY rockfish indices (right panels).

Table 2: Index values and estimated coefficients of variation (CVs) from alternative approaches to developing YOY indices.

Bocaccio	sum antilog ANOVA		antilog sums ANOVA		Delta-GLM		VAST	
	Index	CV	Index	CV	Index	CV	Index	CV
2004	6.878	0.172	1.405	0.256	11.622	0.318	703	0.504
2005	7.216	0.171	1.724	0.237	7.193	0.318	2484	0.364
2006	6.471	0.17	0.987	0.228	2.831	0.46	97	0.75
2007	6.739	0.17	1.227	0.243	8.368	0.349	641	0.499
2008	6.613	0.17	1.115	0.246	8.705	0.355	1377	0.721
2009	6.852	0.171	1.414	0.286	7.692	0.413	1493	0.498
2010							3549	0.54
2011							184	0.791
2012							989	0.887
2013	9.556	0.166	5.94	0.299	21.754	0.378	71157	0.554
2014	7.327	0.169	2.023	0.321	5.458	0.367	5945	0.436
2015	8.481	0.166	4.521	0.251	9.523	0.302	9366	0.33
2016	6.43	0.174	2.333	0.555	9.169	0.438	5430	0.433

Blue/Deacon	sum antilog ANOVA		antilog sums ANOVA		Delta-GLM		VAST	
	Index	CV	Index	CV	Index	CV	Index	CV
2001	6.104	0.279	2.659	0.503	5.482	0.299	2288	0.436
2002	10.024	0.278	12.423	0.495	8.912	0.257	13937	0.289
2003	7.327	0.278	4.685	0.488	6.674	0.244	5729	0.387
2004	8.946	0.278	13.53	0.469	16.367	0.26	18113	0.291
2005	5.97	0.28	2.306	0.473	3.718	0.279	4132	0.311
2006	5.119	0.278	1.16	0.464	1.421	1.553	542	0.855
2007	5.218	0.277	1.274	0.461	4.456	0.375	420	0.52
2008	5.177	0.279	1.225	0.477	2.034	0.526	192	0.629
2009	5.534	0.275	1.683	0.466	3.278	0.314	2129	0.29
2010							1240	0.769
2011	6.283	0.281	3.102	0.5	7.909	0.42	1913	0.557
2012							542	0.855
2013	18.645	0.272	305.436	0.712	22.066	0.328	64142	0.203
2014	7.316	0.271	7.709	0.685	5.221	0.361	5002	0.352
2015	5.129	0.235	1.182	0.637	4.703	0.428	1340	0.54
2016	5.526	0.385	0	0	4.995	0.549	12412	0.475

Yellowtail	sum antilog ANOVA		antilog sums ANOVA		Delta-GLM		VAST	
	Index	CV	Index	CV	Index	CV	Index	CV
2004	5.575	0.314	13.624	0.33	18.472	0.316	14765	0.283
2005	3.892	0.314	1.62	0.333	5.669	0.328	1756	0.357
2006	3.518	0.313	1.214	0.327	1.531	0.72	45	1.078
2007	3.442	0.314	1.159	0.325	1.7	0.69	57	1.057
2008	3.846	0.314	2.239	0.335	4.341	0.324	4280	0.485
2009	3.732	0.31	1.884	0.328	4.354	0.315	3663	0.654
2010							129	0.993
2011	3.726	0.315	1.52	0.35	2.866	0.563	585	0.984
2012							129	0.993
2013	4.477	0.238	12.694	0.487	10.366	0.42	20243	0.474
2014	4.167	0.236	8.213	0.471	8.912	0.444	7323	0.359
2015	2.689	0.21	1.041	0.442	3.315	0.645	1957	0.577
2016	2.954	0.29	0	0	4.603	0.614	42874	0.432

Shortbelly	sum antilog ANOVA		antilog sums ANOVA		Delta-GLM		VAST	
	Index	CV	Index	CV	Index	CV	Index	CV
2004	2.602	0.827	10.099	0.67	11.849	0.666	6091	0.467
2005	8.011	0.854	106.005	0.592	55.807	0.528	157359	0.303
2006	2.04	0.812	3.018	0.578	4.066	0.863	1962	0.576
2007	3.625	0.837	17.624	0.64	18.742	0.62	18509	0.406
2008	2.416	0.81	6.573	0.636	8.838	0.739	7666	0.352
2009	4.676	0.825	79.865	0.826	13.902	0.61	32000	0.402
2010	3.323	0.9	27.044	0.853	12.817	0.931	62008	0.412
2011							7550	1.186
2012							7550	1.186
2013	104.757	1.662	85988.419	0.794	138.074	0.437	1526456	0.287
2014	10.426	1.667	1792.581	0.9	13.662	0.525	214435	0.388
2015	12.477	1.624	4677.989	0.68	15.331	0.45	697206	0.295
2016	8.375	0.468	20330.549	0.838	19.365	0.595	416177	0.366

References

Field, J.C., S. Beyer and X. He. 2015. Status of the Chilipepper Rockfish, *Sebastes goodei*, in the California Current for 2015. Pacific Fishery Management Council Stock Assessment and Fishery Evaluation.

Hastie, J., and S. Ralston. 2007. Pre-recruit survey workshop, September 13-15, 2006, Southwest Fisheries Science Center, Santa Cruz, California, 23 p.

He, X. J.C. Filed, D.E. Pearson, L. Lefebvre, and S. Lindley. 2015. Status of Bocaccio, Sebastes paucispinis, in the Conception, Monterey and Eureka INPFC areas for 2015. Pacific Fisheries Management Council, Portland, OR. http://www.pcouncil.org/wp-content/uploads/2016/05/2015_Bocaccio_Assessment.pdf

Lo, N., L. Jacobson and J. Squire. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Can. J. Fish. Aquat. Sci. 49: 2515-2526.

Ralston, S., and D. F. Howard. 1995. On the development of year-class strength and cohort variability in two northern California rockfishes. Fish. Bull., U. S. 93:710-720.

Ralston, S., K.M. Sakuma and J.C. Field. 2013. Interannual Variation in Pelagic Juvenile Rockfish Abundance– Going With the Flow. Fisheries Oceanography 22: 288–308.

Ralston, S., and I. Stewart. 2013. Anomalous distributions of pelagic juvenile rockfish on the U.S. West Coast in 2005 and 2006. CalCOFI Reports 54: 155-166.

Sakuma, K.M., J.C. Field, B.B. Marinovic, C.N. Carrion, N.J. Mantua and S. Ralston. 2016. Anomalous epipelagic micronekton assemblage patterns in the neritic waters of the California Current in spring 2015 during a period of extreme ocean conditions. Calif. Coop. Oceanic Fish. Invest. Rep. 57: 163-183.

SAS Institute Inc. 2004.

Stefánsson, G. 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES Journal of Marine Science, 53: 577-588.

Thorson, J.T., Shelton, A.O., Ward, E.J. and H.J. Skaug. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES J. Mar. Sci. J. Cons. 72: 1297-1310. doi:10.1093/icesjms/fsu243.

Woodbury, D. P., and S. Ralston. 1991. Interannual variation in growth rates and back-calculated birthdate distributions of pelagic juvenile rockfishes (*Sebastes* spp.) off the central California coast. Fish. Bull., U. S. 89:523-533.