

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 \times latitude interactions, and we continue this practice here. The specific form of the ANOVA mixed-effects model is:

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

with all independent variables treated as categorical. Specifically Y_i is a fixed year effect $\{Y_i \in 2001, 2002, \dots, 2016\}$, L_j is a fixed latitudinal effect $\{L_j \in 32, 34, \dots, 40\}$, Z_k is a fixed depth effect $\{Z_k \leq 160 \text{ m or } Z > 160 \text{ m}\}$, D_l is a fixed calendar date effect $\{D_l \in 120, 130, \dots, 170\}$, V_m is a random vessel effect [$V_m \sim N(0, \sigma_v)$], and $\epsilon_{i,j,k,l,m,n}$ is normal error term [$\epsilon \sim N(0, \sigma_\epsilon)$] 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