# Improving Catch Estimation Methods in Sparsely Sampled Mixed-Stock Fisheries.

Nick Grunloh[a], Edward Dick[b], Don Pearson[b], John Field[b], Marc Mangel[a,c]

March 13, 2018

[a] Center for Stock Assessment Research, University of California, Santa Cruz, Mail Stop SOE-2, Santa Cruz, CA 95064, USA.
[b] Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanographic and Atmospheric Administration, 110 McAllister Way, Santa Cruz, CA 95060, USA.
[c] Department of Applied Mathematics and Statistics, Jack Baskin School of Engineering, University of California, Santa Cruz, Mail Stop SOE-2, Santa Cruz, CA 95064, USA.

**Abstract**

Effective management of exploited fish populations requires accurate estimates of commercial fisheries catches to inform monitoring and assessment efforts. In California, the high degree of heterogeneity in the species composition of many groundfish fisheries, particularly those targeting rockfish (genus *Sebastes*), leads to challenges in sampling all potential strata, or species, adequately. Limited resources and increasingly complex stratification of the sampling system inevitably leads to gaps in sample data. In the presence of sampling gaps, ad-hoc species composition point estimation is currently obtained according to historically derived "data borrowing" (imputation) protocols which introduce unknown bias and do not allow for uncertainty estimation or forecasting. In order to move from the current ad-hoc "data-borrowing" point estimators, we have constructed Bayesian hierarchical models to estimate species compositions, complete with accurate measures of uncertainty, as well as theoretically sound out-of-sample predictions. Furthermore, we introduce a Bayesian model averaging approach for inferring spatial pooling strategies across the over-stratified port sampling system. Our modeling approach, along with a computationally robust system of inference and model exploration, allows us to 1) objectively compare alternative models for estimation of species compositions in landed catch, 2) quantify uncertainty in historical landings, and 3) understand the effect of the highly stratified, and sparse, sampling system on the kinds of inference possible, while simultaneously making the most from the available data.

# 1 Introduction

Estimates of landed catch are a key component of many fishery management systems. Stock assessment models (referred to here as assessments) are often conditioned on time series of annual catch, usually under the assumption that catches are known without error. While some assessment models are able to incorporate uncertainty in catch (e.g. Stock Synthesis; Methot and Wetzel, 2013), reliable estimates of catch uncertainty are often unavailable. Without this information, assessment authors often rely on ad-hoc sensitivity analyses which may or may not be incorporated into management advice and/or fail to propagate catch uncertainty into quantities of interest to managers.

Over the past decade, the estimation of catch and associated uncertainty has become a focus for recreational fisheries in the United States (NAS, 2017). Commercial fisheries, on the other hand, are often assumed to have precise estimates of catch by species. This is due in part to the availability of landing receipts (aka fish tickets) which serve as a record of the weight of fish landed into various market categories (sort groups). As noted by Pearson et al. (2008), it is important to recognize that species and market categories are not synonymous. On the U.S. West Coast, for example, it is common for multiple species to be landed within a single market category (CALCOM 2018, PacFIN 2018). This is expected for categories that are clearly designated as mixed-species categories (e.g. "nearshore rockfish", or species within a particular genus or family). However, in some states, categories that are named after a single species still contain several species, to varying degrees, even after regulations require sorting into a particular category (Pearson et al., 2008). As a result, estimation of landings for a single species based on landing receipt data alone may produce biased estimates of total catch.

Fisherman, dealers, or processors typically decide how to sort species into existing market categories (defined by state agencies) on a landing receipt. Trained port samplers intercept vessels offloading catch or during subsequent processing in order to determine the species composition of catch landed in a given market category (Sen 1984, Crone 1995, Tsou et al. 2015). These species composition data are used to partition, or distribute, the weight of landed catch in a market category across species, a process commonly referred to as catch expansion (Pearson and Almany, 1995). To estimate total landings for a single species, the expanded catch is summed across all relevant market categories. Assuming that landing receipt data are a census of the landed catch, uncertainty in total catch for a given species reflects variability in the species composition among port samples. If information is available to estimate sources of bias in the landing receipt data, e.g. underreporting, then it is possible to also incorporate uncertainty in the bias correction factors as well (Bousquet et al. 2010).

In this study, we quantify sampling uncertainty for estimates of landed catch, although the modeling framework could be extended to include uncertainty or bias in landing receipt data.

Within market categories, particularly those used historically for groupings of highly specious rockfish (*Sebastes* spp), the species composition of landed catch can vary spatially, temporally, by fishing gear, and catch disposition (e.g. fish sold alive or dead). These differences are attributable to many factors, including market preference, fishing behavior, regulatory constraints , and biological/ecological characteristics (e.g. spatial distribution) of the landed species. As a result, estimates of species composition for a given market category are often stratified over time (e.g. quarterly) and across other relevant strata (e.g. ports, gears, catch disposition ). Sampling programs often have limited funds, and attempts to reduce bias in species composition estimates through the introduction of additional strata comes at a cost, namely reduced precision (Cochran, 1977; Tomlinson, 1971).

On the U.S. West Coast, port sampling programs for rockfish and other groundfish allocate effort both spatially and temporally, but many domains of interest (e.g. market category, gear type, catch disposition) remain unsampled or sparsely sampled due to a proliferation of categories over time, logistical constraints, and limited resources (Sen 1986; Crone 1995; Pearson et al. 2008; Tsou et al. 2015). In California, for example, commercial port sampling effort has changed over time and space (Pearson and Almany 1995). For example, regular sampling of California ports north of Point Conception (roughly 34° 27' N. latitude) began as early as 1978, but the more southern ports were rarely sampled prior to 1983. This allocation of effort was largely based on the statewide distribution of landings, diffuse spatial distribution of southern commercial ports, and limitations in funding for port samplers.

When no port samples are collected for landed strata and domains, species composition estimates are borrowed from other strata using deterministic algorithms based on expert opinion. These algorithms have unknown bias and precision. In contrast, model-based estimators are increasingly used to estimate quantities of interest for domains with small sample sizes and/or unsampled strata (sometimes referred to as small area estimation; Rao 2003). As a pilot study, Shelton et al. (2012) developed a Bayesian hierarchical statistical framework for estimating species compositions in rockfish market categories from trawl fisheries from a single port in California in two separate years. Their model has the ability to partially pool information among sparsely sampled strata, predicts species compositions for unsampled strata, and can be combined with landing receipts to estimate total landings by species, across market categories and other strata, along with associated estimates of uncertainty. However, their model considered hierarchical pooling only among quarters within a single year. The authors also underscored the need to better understand performance of alternative models, and to overcome issues with computation time, particularly since commercial port sampling data sets often include hundreds of landed strata spanning decades, multiple ports, gear types, and other domains of interest.

Among the U.S. West Coast states, the challenge of estimating landings for sparsely-sampled mixed stock rockfish fisheries is perhaps greatest for California. Although overall landings have historically been greater for rockfish off of Oregon and Washington, California has a greater number of commercial ports, market categories, and landed species (Pearson and Irwin 1997). California's commercial landings also have greater species diversity among ports due to the geographical range of the coast and the observation that species diversity for this genus is greatest in the Southern California Bight (Love et al. 2002 ). Lastly, California includes two major biogeographic features, Point Conception and Cape Mendocino, that are associated with different physical oceanographic conditions and biological community assemblages (Hickey 1979, Checkley and Barth 2009, Gottscho 2016), and these features are also frequently used as spatial boundaries for stock assessments and management measures. Of particular consequence to the estimation of species compositions is the proliferation of landed market categories over time, particularly during the 1990s, Figure (1)). Sampling effort also leveled off in the mid-1990s, with a reduction in effort in the early 2000s, associated with substantial declines in total catches as well as reductions in sampling resources. The net result of increased stratification and flat (or reduced) sampling effort over time is a decline in mean sample size per stratum, Figure (1). In this situation, it is critical to understand how efforts to reduce bias (e.g. increasing the number of landed market categories) affect precision of the expanded catch estimates.

Of particular consequence to the estimation of species compositions in California is the proliferation of landed market categories over time, particularly during the 1990s, see Figure (1). Sampling effort also leveled off in the mid-1990s, with a reduction in effort in the early 2000s, associated with substantial declines in total catches as well as reductions in sampling resources. The net result of increased stratification and flat (or reduced) sampling effort over time is a decline in mean sample size per stratum, see Figure (1). In this situation, it is critical to understand how efforts to reduce bias (e.g. increasing the number of landed market categories) affect precision of the expanded catch estimates.

Assessments that take catch uncertainty into account are not new (c.f. Doubleday 1976), but most assessments on the U.S. West Coast assume catch is known without error (PFMC 2018). As a result, catch uncertainty is rarely propagated into management reference points. However, the implications of catch uncertainty are not limited to stock assessment efforts. In a management context, catch estimates with large (but unknown) uncertainty may cause managers to react to large, high-frequency deviations in estimated catch, and either impose unnecessary restrictions on a fishery, or mistakenly support excessive harvest. This is particularly an issue for prohibited and/or "choke" species, for which catch information is limited and may be based solely on estimates of discarded catch.

In this study, we evaluate the model-based framework proposed by Shelton et al. (2012) using commercial port sampling data collected in California, U.S.A. We describe species composition data collected by the California Cooperative Groundfish Survey (CCGS, 2017) over the period 1978-1990. We then extend the Shelton et al. framework to address limitations of their approach. Specifically, we evaluate alternative likelihoods to address overdispersion, compare multiple hierarchical structures for pooling information through time, and integrate model predictions across uncertainties in the spatial model structure. Finally, we estimate landed catch by species for both sampled and unsampled strata, and summarize a general framework for quantifying uncertainty including an efficient database design for dissemination of results at any level of aggregation.

# 2 Methods

## 2.1 Data

As outlined in Sen (1984, 1986) the species composition port sampling data are the result of a cluster sampling protocol executed across the many strata of California's commercial fisheries. Each sample is intended to be two fifty- pound clusters selected at random from a stratum. Although port samplers do their best to follow protocol, in reality the port sampling environment does not always allow Sen's original protocol to be followed. Variations in the sampling protocol may result in only a single cluster being taken, or the size of clusters taken to vary from stratum to stratum based on the particular challenges of sampling each stratum.

Samples are recorded as integer pounds for each observed species, across the landed market categories, gear groups, and port complexes in time (quarters within year). Presently there are 71 rockfish market categories, although not all market categories are always used. The number of market categories with recorded landings has gone from less than 25 in 1978 to about 55 in 2014, see Figure (1). Landings are grouped into major fishing gear groups (trawl, hook and line, gillnet, fish pot, or other minor categories) and ten major port complexes spanning the California coast, see Figure (4).

The model based methodology proposed here does not rely strongly upon the cluster sampling structure, but rather views each sample as independent and identically distributed (*i.i.d.*) draws from a data generating model, conditional on a parameterization of the stratification system. So long as the parameterization and data generating model are sufficiently robust for handling the behavior of these data, a conditionally *i.i.d.* model of these data will be practically useful for producing predictions about the data generating system.

That said, for the purpose of modeling these data, it is enough to know which clusters were collected as part of which samples, and how big each cluster actually ended up being. This information is readily available from CALCOM, a database maintained by the California Cooperative Groundfish Survey (CALCOM, 2018). Just as in Shelton et al. (2012), we aggregate all observed clusters within each unique sample so that the total weight sampled is the sum of pounds in each cluster. Similarly the observed weight for a particular species, in each unique sample, is the sum of all of the observed weights across clusters.

Although model based data analysis has the potential to add significant structure to data, a judicious application of these methods must always confront the model with enough empirical information to adequately learn about the system. In this setting some market categories and time periods may not be well enough sampled to learn the parameters of the models presented here. For this reason, we refrain from modeling any period where the minimum possible number of effective parameters exceeds the number of samples for the modeled period. Rather than apply models inappropriately, these landings are speciated as the nominal species for their market category. We later demonstrate that due to prioritization in sampling heavily landed, or otherwise commercially relevant categories, this sample size heuristic leaves relatively few landings to be speciated in a statistically uninformed way (i.e. "nominal" speciation). Thus nominal speciation represents a negligible component of the overall expanded landings for most species.

## 2.2 Likelihood Modeling

For the purposes of accurately modeling not only species composition means, but also higher moments of the data (e.g. variances), it is necessary to recognize model limitations with respect to over-dispersed data. Among the simplest models for count data are the Poisson and binomial models. Both models are typically specified with a single parameter for modeling all of the moments of the data, and thus they rely heavily on their respective data generating processes to accurately represent higher moments in the data. McCullagh and Nelder (1989, pg. 124) commiserate about the prevalence of over- dispersed data in cluster sampling, and explain ways in which cluster sampling itself may result in overdispersion.

Extending the Poisson and binomial models to deal with over-dispersion, typically involves adding additional parameters for the purpose of modeling higher moments of the data. The negative binomial (NB) distribution is often used as an over-dispersed extension of the Poisson model, since it can be expressly written as an infinite mixture of Poisson distributions. The beta-binomial model is used as an over-dispersed extension of the binomial model.

The Poisson and binomial models attempt to model both the mean and residual variance of the data, with a single parameter for each species. By definition these models do not have additional parameters to model the variance, but rather, residual variances in these models are simply transformations of their mean parameters. Only estimating the mean parameters in these cases may not be sufficient to produce models which predict well.

In contrast, the negative binomial and beta-binomial models estimate an additional parameter which can be used to disentangle the mean and residual variance estimates. Thus the negative binomial and beta-binomial models may produce more accurate estimates of the residual variance, while producing more accurate measures of center. We develop an example on a subset of data to evaluate statistical support for overdispersed models, see Appendix (5.2), which we have subsequently used for the purposes of applying at an operational scale.

### 2.2.1 Beta-Binomial Model

For a particular market category, $y_{ijklm\eta}$ is the $i^{th}$ sample of the $j^{th}$ species' weight, in the $k^{th}$ port, caught with the $l^{th}$ gear, in the $\eta^{th}$ quarter, of year $m$. The $y_{ijklm\eta}$ are modeled as $i.i.d.$ observations from a beta-binomial distribution conditional on parameters $\mu_{jklm\eta}$ and $\sigma^2_{jklm\eta}$,

$$y_{ijklm\eta} \overset{i.i.d.}{\sim} BB(\mu_{jklm\eta},\ \sigma^2_{jklm\eta}).$$

Above, $\mu_{jklm\eta}$ is the stratum level mean weight, and $\sigma^2_{jklm\eta}$ is the stratum level residual variance. $\mu_{jklm\eta}$ is related to a linear predictor, $\theta_{jklm\eta}$, via the mean function,

$$\mu_{jklm\eta} = n_{ijklm\eta}\frac{\exp(\theta_{jklm\eta})}{1+\exp(\theta_{jklm\eta})}.$$

Here $n_{ijklm\eta}$ is the observed aggregate cluster size for each sample. Additionally, $\sigma^2_{jklm\eta}$ is related to $\mu_{jklm\eta}$ and the overdispersion parameter, $\rho$, via the following equation,

$$\sigma^2_{jklm\eta} = \mu_{jklm\eta}\Big(1 - \frac{\mu_{jklm\eta}}{n_{ijklm\eta}}\Big)\Big(1 + (n_{ijklm\eta} - 1)\ \rho\Big).$$

$\rho$ is the within-cluster correlation. The situation where $\rho \to 1$ represents identical information content among replicates within a cluster, with maximal overdispersion relative to the binomial distribution. The situation where $\rho \to 0$ represents totally independent information content among replicates within a cluster, and the beta-binomial model approaches the binomial model. $\rho$ explicitly models average overdispersion across all strata within a market category, while $\mu_{jklm\eta}$ gives the model flexibility at the stratum level through the $\theta$ linear predictors,

$$\theta_{jklm\eta} = \beta_0 + \beta^{(s)}_j + \beta^{(p)}_k + \beta^{(g)}_l + \beta^{(t)}_{m\eta}.$$

Firstly, $\theta$ includes a reference level intercept ($\beta_0$). Secondly, $\theta$ is factored among the many strata by additive offsets from $\beta_0$ for each of the species ($\beta^{(s)}_j$), port-complexes ($\beta^{(p)}_k$), and gear-groups ($\beta^{(g)}_l$). Finally year and quarter parameters are indicated generally here inside the $\beta^{(t)}_{m\eta}$ term. Several forms for $\beta^{(t)}_{m\eta}$ are explored each implying a different prior and partial pooling strategies as described in the following section.

## 2.3 Priors

To complete the Bayesian formulation of this model, priors are expressed in a largely diffuse manner.

$$\beta_0 \propto 1$$
$$\left\{\beta^{(s)}_j, \beta^{(p)}_k, \beta^{(g)}_l\right\} \sim N(0, 32^2)$$

Since the $\beta_0$ reference level is chosen arbitrarily, with no conception of which values it may take, no restrictions are placed on the value of the intercept. The species ($\beta^{(s)}_j$), port-complex ($\beta^{(p)}_k$), and gear-group ($\beta^{(g)}_l$) offsets are assigned diffuse normal priors. The large fixed values of the prior variance hyperparameters produce behavior similar to classical fixed effect models for species, port-complex, and gear- group parameters.

In returning to the time parameter model, $\beta^{(t)}_{m\eta}$, it is useful to consider how overparameterized models may cause overfitting and weaken model performance through the bias-variance dilemma (Ramasubramanian, K. and Singh, A., 2016). Simply put, the bias-variance dilemma means that model formulation is not simply a bias reduction task, but rather the goal is to formulate models which reduce bias, while jointly minimizing uncertainty. Janyes (2003, pg. 511) describes how the inclusion of estimation bias via the Bayesian methodology may produce better performing estimates, more quickly, than unbiased counterparts. Among the simplest ways to see the principle is in the structure of the MSE performance metric, and how it can be explicitly written to value both estimator bias and variance, as follows.

$$\text{MSE}(\hat{\theta}) = \mathbb{E}\left[\,(\hat{\theta}-\theta)^2\,\right] = \overbrace{\mathbb{E}\left[\,\left(\hat{\theta}-\mathbb{E}(\hat{\theta})\right)^2\,\right]}^{\text{Var}(\hat{\theta})} + \overbrace{\left(\,\mathbb{E}(\hat{\theta})-\theta\,\right)^2}^{\text{Bias}(\hat{\theta},\theta)^2}$$

Furthermore a model can minimize bias, without regard for estimation uncertainty, by including one model parameter to be fit to each observation. These parameter estimates are totally unbiased, however such a model is also predictively useless since each estimated parameter is specifically bound to a particular observation, and thus such a model does not generalize.

For modeling $\beta^{(t)}_{m\eta}$ we consider a spectrum of models which span a wide range of partially pooled models with several different predictive structures as seen below.

### 2.3.1 (M1)

$$\beta^{(t)}_{m\eta} = \beta^{(y)}_{m} + \beta^{(q)}_{\eta}$$

$$\beta^{(y)}_{m} \sim N(0, 32^2)$$

$$\beta^{(q)}_{\eta} \sim N(0, 32^2)$$

(M1) represents a fixed effects model for additive year and quarter parameters. Here each year and quarter are assigned totally independent and diffuse priors.

### 2.3.2 (M2)

$$\beta^{(t)}_{m\eta} = \beta^{(y)}_{m} + \beta^{(q)}_{\eta}$$

$$\beta^{(y)}_{m} \sim N(0, v^{(y)})$$

$$\beta^{(q)}_{\eta} \sim N(0, v^{(q)})$$

(M2) estimates two hierarchical variance parameters, $v^{(y)}$ and $v^{(q)}$. $v^{(y)}$ has the effect of partially pooling information among year parameters, while $v^{(q)}$ partially pools information among quarter parameters (i.e. treats both year and quarter as "random effects"). The actual degree of pooling among each of the years and quarters is determined by the data.
Depending on the posterior distributions of $v^{(y)}$ and $v^{(q)}$, the $\beta^{(y)}$ and $\beta^{(q)}$ may be shrunk back toward the common mean (for small $v$) or allowed to take largely distinct values (in the case of large estimates of the $v$).

### 2.3.3 (M3)

$$\beta_{m\eta}^{(t)} = \beta_m^{(y)} + \beta_\eta^{(q)} + \beta_{m\eta}^{(y:q)}$$
$$\beta_m^{(y)} \sim N(0, v^{(y)})$$
$$\beta_\eta^{(q)} \sim N(0, v^{(q)})$$
$$\beta_{m\eta}^{(y:q)} \sim N(0, v)$$

(M3) functions similarly as (M2), in that it has hierarchical partial pooling among both the $\beta_m^{(y)}$ and $\beta_\eta^{(q)}$ parameters, except that it introduces a two-way interaction term between year and quarter. This interaction term allows estimates for particular quarters to differ from year to year, as opposed to the previous models in which quarters within a year are assumed to be identical from year to year.

Furthermore the $\beta_{m\eta}^{(y:q)}$ are also modeled hierarchically to introduce a single variance parameter, $v$, shared among all of the $m\eta$ time chunks. Although this interaction term adds many parameters to the model, the shared $v$ parameter functions to shrink extraneous $\beta_{m\eta}^{(y:q)}$ estimates back toward the common stratum mean.

### 2.3.4 (M4)

$$\beta_{m\eta}^{(t)} = \beta_{m\eta}^{(y:q)}$$

$$\beta_{m\eta}^{(y:q)} \sim N(0, v)$$

(M4) simplifies (M3) by excluding year and quarter main effects. This leaves all temporal information in the data to be modeled solely by the quarterly $\beta_{m\eta}^{(y:q)}$ interaction terms. This model represents more opportunity for partial pooling through time than (M3), as fewer time parameters are introduced. Furthermore all of the $\beta_{m\eta}^{(y:q)}$ are hierarchically pooled back toward a single common stratum mean via the single shared variance parameter, $v$.

A model treating the interaction terms as fixed effects (e.g. model (M4) where $v$ is a large fixed value) was not considered. The $\beta_{m\eta}^{(y:q)}$ interaction terms introduce a very large number of parameters, which require regularization in this setting. Furthermore, pooling $\beta_{m\eta}^{(y:q)}$ through time introduces a formalized way to make predictions about unsampled strata.

### 2.3.5 (M5)

$$\beta_{m\eta}^{(t)} = \beta_{m\eta}^{(y:q)}$$

$$\beta_{m\eta}^{(y:q)} \sim N(0, v_\eta)$$

(M5) is largely the same as (M4), but it represents slightly less potential partial pooling through its hierarchical prior variances, $v_\eta$, on $\beta^{(y:q)}_{m\eta}$. Here interaction terms are allowed to partially pool interactions across years, within a common quarter, but since each quarter is assigned a separate variance parameter no pooling is possible between quarters.

#### 2.3.6 (M6)

$$\beta^{(t)}_{m\eta} = \beta^{(y:q)}_{m\eta}$$

$$\beta^{(y:q)}_{m\eta} \sim N(0, v_m)$$

(M6) follows the same idea as (M5), however here interaction terms are allowed to partially pool interactions within a common year, across the quarters of that year, but not between years. (M6) often involves fitting slightly more parameters than (M5) because, at least in this setting, it is typical to model more than four years of data at once.

Historically, regulations have been enacted with the aim of isolating catch in a market category to a single species (sort requirements). This clearly affects the composition of the target market category, but these regulations also affect the species composition of other market categories in which the target species previously occurred. We incorporate this information into the model structure by treating time periods with relatively stable regulatory conditions as independent models. In other words, information is only shared among years in which regulations were similar. For example, a sort requirement for widow rockfish (*S. entomelas*) was initiated in 1983. This not only affected the composition of the widow rockfish market category (269), but also the composition of other categories, such as the unspecified rockfish market category (250). We model the first five years of available data (1978-1982) independently from the years 1983-1990. In 1991, a sort requirement for bocaccio rockfish (*S. paucispinis*) was enacted, which is known to have affected the composition of other categories, including the Chilipepper rockfish market category (254) and the Chilipepper/Bocaccio market category (956). Given these regulation changes, future year groupings might include independently modeled periods in 1991-1999 and 2000-2015, however the analysis of those years has not yet been developed and alternative year groupings may be explored for the most recent landings at a later time.

Hierarchical variance parameters are estimated from the data. As the above models learn the posteriors of the hierarchical variance parameters, it affects the degree of shrinkage as well as the effective number of parameters held within the respective hierarchies (Gelman, 2014). To achieve this, each variance parameter must itself be assigned a prior to be estimated. For all of the hierarchical variance parameters included in the above models $v$ is assigned a diffuse inverse gamma (IG) prior $v \sim IG(1,\ 2 \times 10^3)$.

Finally the overdispersion parameter, $\rho$, is assigned a diffuse normal prior on the logit scale, $\text{logit}(\rho) \sim N(0, 2^2)$. The $N(0, 2^2)$ prior is indeed a symmetric, and far reaching, prior when back transformed to the unit interval. To notice this, it is helpful to realize that the central 95% interval for a $N(0, 2^2)$ (i.e. $0 \pm 3.91$), includes almost the entirety of the back transformed unit interval (i.e. $0.5 \pm 0.48$).

For the purpose of motivating the time parameter structure to be used across all market categories and time periods, each of models M1-M6 are fit to data from market category 250 from 1978-1982. Performing such a model selection exercise across all market categories, in each time period, would not only be impractical but would also risk overfitting models, particually in poorly sampled categories. In the early time periods, market category 250 is the most heavily sampled multi-species market category, and thus it represents an excellent data-set for testing.

## 2.4 Species Composition Prediction

Bayesian inference of the above models gives access to the full posterior distribution of all of the parameters of the model given the data. It is useful to emphasize that in the Bayesian setting, these parameters have full distributions, and they are typically handled as a large number of samples from the joint posterior distribution of the parameters. Once the posterior sampling is complete, this simplifies parameter mean and variance estimation; required moments are simply obtained by computing the desired moments from the posterior samples. Additionally, the fact that the parameters are full distributions means that any functions of those parameters are themselves random variables with the function representing a transformation of those parameters.

To obtain predicted species compositions from the beta-binomial model, first consider the posterior predictive distribution of sampled weight for a particular stratum.

$$p(y^*_{jklm\eta}|y) = \iint \text{BB}\Big(y^*_{jklm\eta}|\mu_{jklm\eta}, \sigma^2_{jklm\eta}\Big) P\Big(\mu_{jklm\eta}, \sigma^2_{jklm\eta}|y\Big) d\mu_{jklm\eta} d\sigma^2_{jklm\eta}.$$

Here BB is the data generating beta-binomial distribution for a predictive observation and $P(\mu_{jklm\eta}, \sigma^2_{jklm\eta}|y)$ is the posterior distribution of the parameters given the observed data. $\mu_{jklm\eta}$ and $\sigma^2_{jklm\eta}$ are integrated numerically via Monte Carlo integration to produce samples from the predictive distribution, $p(y^*_{jklm\eta}|y)$, for sampled weights in the $jklm\eta^{th}$ stratum.

Obtaining predictive species compositions from predictive weights amounts to computing the following transformation,

$$\pi^*_{jklm\eta} = \frac{y^*_{jklm\eta}}{\sum_j y^*_{jklm\eta}} \quad y^*_{klm\eta} \neq 0.$$

For a particular market category, $\pi^*_{jklm\eta}$ is predicted proportion of species $j$ in the $k^{th}$ port, caught with the $l^{th}$ gear, in the $\eta^{th}$ quarter, of year $m$.

## 2.5 Expansion of Landed Catch to Species

For a particular market category, speciated landings simply amounts to the multiplication of the known total landings ($\lambda_{klm\eta}$), reported on landing receipts in the $klm\eta^{th}$ stratum, with the posterior predictive species composition, $\pi^*_{jklm\eta}$, as follows

$$\lambda^*_{jklm\eta} = \lambda_{klm\eta}\pi^*_{jklm\eta}.$$

$\lambda^*_{jklm\eta}$ is then the posterior predictive landings for species $j$ in the $klm\eta^{th}$ stratum of a particular market category. Recall that since $\pi^*_{jklm\eta}$ is a random variable, then so is $\lambda^*_{jklm\eta}$. Computing the variance of $\lambda^*_{jklm\eta}$ simply amounts to computing the variance of random draws from the $\lambda^*_{jklm\eta}$ distribution. Furthermore, any level of aggregation of $\lambda^*_{jklm\eta}$ is easily obtained by summing $\lambda^*_{jklm\eta}$ draws across the desired indices. For example to obtain the distributions of yearly catch of Bocaccio in a particular market category (i.e. aggregated across ports, gears, and quarters) one simply fixes $j$ to Bocaccio, and computes the following transformation of $\lambda^*_{jklm\eta}$,

$$\lambda^*_{j\cdot\cdot m\cdot} = \sum_k \sum_l \sum_\eta \lambda^*_{jklm\eta}$$

.

Distribution summaries such as quantiles, means, or variances may be computed by computing those metrics from the random draws of the resulting $\lambda^*_{j\cdot\cdot m\cdot}$ distributions.

## 2.6 Model Exploration & Averaging

Presently, strata with diminishingly small sample sizes are managed by an ad-hoc "data borrowing" protocol, as outlined in Pearson and Erwin (1997). The protocol for "data borrowing" calls for pooling only when forced to fill holes brought about by unsampled strata. Naturally, such a pooling protocol introduces bias to fill in unsampled strata, however due to the mathematically unstructured way in which this bias is introduced, it is hard to quantitatively justify these "data borrowing" rules.

Model (M4) avoids temporal ad-hoc "borrowing" protocols described in Pearson and

Erwin (1997) by making use of its hierarchical structure to fill temporal holes with a posterior predictive distribution for unseen time periods within the modeled period. This hierarchical structure uses the data to estimate the degree of pooling through time, rather than ad-hoc "data borrowing". In addition, the M4 model allows data collected during periods with similar regulations to inform predictions across both quarters and years, unlike the previous ad-hoc approach which only shared information among quarters within a year.

Despite the benefits of modeling these data as Bayesian hierarchical models, port sampling data still remains sparse. Given the degree of sparsity in these data it is certainly possible that models which consider an additional degree of data pooling between port complexes may offer predictive benefits. In exploring strategies for pooling data across space it is necessary to formalize the port complex pooling scheme in a way which provides a mathematically understandable and scalable structure to build upon.

Given the spatial structure, and complex behavior of, port complex parameters, the typical zero mean hierarchical regularization priors are not appropriate among port complexes. Pooling across spatial categorical parameters in this setting requires the ability to pool port complex parameters back toward an unknown number ($\leq K$) of mean levels. Rather than hierarchically regularize port complex, we frame port complex pooling as a model uncertainty problem, in which we consider some degree of full pooling among port complex, but the exact degree of pooling, and the particular partitioning of the pooled port complexes are not known.

Port complex pooling is achieved by repeatedly fitting model (M4) with different partitionings of the port complex variables within a particular market category and modeling time period. This model exploration exercise explores the possible ways to produce groupings of the existing port complexes so as to discover predictively useful partitionings of the port complexes. Insisting that the port complex groupings be partitions of the available port complexes provides a well-defined mathematical structure for exploring the space of pooled port complexes.

The size of the space of possible pooled models is in the setting is well defined in terms of the size of the set of items to be partitioned, $K$, as given by the Bell numbers ($B_K$),

$$B_K = \sum_{\hat{k}=0}^{K} \frac{1}{\hat{k}!} \left( \sum_{j=0}^{\hat{k}} (-1)^{\hat{k}-j} \binom{\hat{k}}{j} j^K \right).$$

In the case of California the set of items to be partitioned is the set of port complexes in California, of which there are $K = 10$, implying a grand total of $B_{10} = 115975$ ways of partitioning the port complexes in California in each market category and modeled time period. The brute force model selection strategy of computing all 115975 of these partitionings strategies is computationally infeasible. However, not all pooling schemes represent biologically relevant models. For example, it is likely reasonable to pool only among adjacent ports (i.e. no discontinuities between port complex pooling in space) due to species distributions and the presence of biogeographical provinces, and it may be similarly reasonable to assert that similar regions can only extend across a small number of ports.

We limit the set of models to be considered by applying two constraints. 1) Only adjacent port poolings are considered, and 2) the maximum size of a port complex grouping is fixed to be no larger than three port complexes. These two constraints were chosen so as to mirror the currently accepted protocols in Pearson and Erwin (1997), although many other constraints may, in theory, be chosen. When these two simple constraints are applied, the number of models to explore in each modeled period is reduced to a much more manageable 274 models. It this framework, it is not necessary to impose a priori conditions on how data are shared among ports (e.g. never borrowing across Point Conception, as in the current ad hoc approach). However, if the data strongly support a particular port configuration, models with that structure will have greater influence on model predictions.

An exhaustive search of the models in the constrained subspace of $B_{10}$, allows for a concrete comparison of the relative predictive accuracy of each partitioning. Additionally the partitioned models provide a set of candidate models for use in Bayesian Model Averaging (BMA) (Hoeting et al., 1999). BMA, as applied here, allows the model exploration strategy to average models across all potentially relevant partitions of the port complexes, so as to add robustness to final species composition estimates.

For the $\iota^{th}$ model in a set of candidate models $M$, then the BMA weight for $M_\iota$ follows directly from Bayes Theorem as,

$$\omega_\iota = Pr(M_\iota|y) = \frac{p(y|M_\iota)p(M_\iota)}{\sum_\iota p(y|M_\iota)p(M_\iota)}.$$

Where $\omega_\iota$ is the posterior probability that model $\iota$ is the true data generating model of the data, conditional on the subspace of candidate models and the observed data. $\omega_\iota$ is then straightforwardly used to average posteriors across all of the models, as

$$\bar{p}(\theta|y) = \sum_\iota \omega_\iota p(\theta|y, M_\iota).$$

[EJ to insert methods subsection here on landings by species-year-gear, as well as comparison to CALCOM]

# 3 Results

## 3.1 Characteristics of the Landings Data

In the two time periods modeled here, Figures (5 and 6) show how commercial port sampling effort tracks both total landed weight as well as the number of species in market categories accounting for the top 99% of total landings in each time period. Comparable results for the periods 1991-1999 and 2000-2015 are discussed Appendix (5.1), although we have not yet completed modeling for these time periods. Our model is applied to a relatively large proportion of the landings in both time periods, and nominal speciation occurs for a relatively negligible proportion of total landings. This is because 1) port sampling effort is largely opportunistic and implicitly prioritizes heavily landed market categories, and 2) our model is only fit to market categories with more data than model parameters. As a result market categories left with too few samples to fit our model tend to be less landed. Applying the sample size heuristic to determine which market categories are expended by our model results in 96.8% of landed weight being expanded via our model in the 1978-1982 time and 98.3% of landed weight expanded by our model in 1983-1990.

The lower panels of Figures (5 and 6) demonstrate just how many different species are landed into commercially relevant market categories. Although market categories often carry names that label them with a nominal species, Figure (5) makes it abundantly clear that these names can mislead our thinking about the purity, and consistency, of these categories through time. To drive this point, consider the sampled species in market category 267 in 1978-1982. The nominal label for market category 267 is Brown Rockfish, while Brown Rockfish only amounts to a small fraction of that category in 1978-1982. In fact, only 6.3% of the sampled weight in 1978-1982 consisted of Brown Rockfish. In 1978-1982 market category 267 might be better named Widow Rockfish as Widow amounts to 92.6% of sampled weight in this time period, however market category 267 is composed of 99.6% Brown and 0% Widow in recent time periods (see Appendix 5.1).

## 3.2 Predictor and Prior Selection

Recall models (M1)-(M6) differ in the structure of the $\beta^{(t)}_{m\eta}$ time parameters. Table (1) shows the relative support for those model structures. From (M1) to (M4) the models represent a spectrum of models with an increasing potential of shrinkage among time parameters. Models M5 and M6 represent models which build in complexity, from (M4), via the inclusion of multiple hierarchical variance parameters among the interaction terms.

Across all of the time models, model M4 displays consistent support over all other candi-

| | M1 | M2 | M3 | M4 | M5 | M6 |
|---|---|---|---|---|---|---|
| MSE | 0.127245 | 0.127042 | 0.126801 | 0.122373 | 0.127236 | 0.126573 |
| $\Delta$ DIC | 2558.56 | 2259.94 | 2013.21 | 0 | 2175.32 | 2174.71 |
| $\Delta$ WAIC | 2562.65 | 2263.58 | 2009.32 | 0 | 2171.18 | 2170.56 |
| $pr(M\|y)$ | $\approx 0$ | $\approx 0$ | $\approx 0$ | $\approx 1$ | $\approx 0$ | $\approx 0$ |

Table 1: A table showing Mean Squared Error (MSE; computed on the species composition scale), delta deviance information criterion ($\Delta$ DIC), delta widely applicable information criterion ($\Delta$ WAIC), and marginal Bayesian model probabilities ($pr(M|y)$) across the fit of models M1-M6 to data from 1978-1982 in market category 250.

date models considered here. It is worth mentioning that among all of the models considered here, (M4) offers the largest potential for hierarchical partial pooling among the time parameters. Model (M4) represents a model with maximal potential for pooling through time, while still maintain the ability to model differences in seasonality from year to year.

## 3.3 Model Exploration & Averaging

Considering Figure (9), the best partitioned model (first column, $\omega = 0.32$) gives distinct parameters to CRS and ERK, while pooling BRG/BDG, OSF/MNT/MRO, and OSB/OLA/OSD. This model uses five parameters to model the ten ports complexes in California. Given the set of candidate models explained above, the BMA procedure gives this model a probability of 32%. Notice that the only difference among the top four models is in how the port complexes south of Point Conception are handled. In fact, the seven northerly port complexes are identically partitioned in the top four models, which also represent all of the possible partitionings of the southern three port complexes.

In this modeled period it is known that no species composition sampling was done south of Point Conception, thus it is not surprising that these models perform similarly. When no data are present, parameters simply represent place holders for out of sample prediction. Since the port complexes south of Point Conception are not informed by data, the predictions are identical in these categories. Since the first model makes identical predictions to the following three, and does so using the fewest parameters, it is correctly identified as the most parsimonious explanation among these data.

Considering how the top four model configurations share identical structure in the seven northerly port complexes, while exhaustively spanning the candidate partitions south of Point Conception, it is simple to see that BMA assign's approximately 71% marginal probability to the northerly model structure.

The results shown here only represent a single market category across the time period 1978-1982. Similar results for other market categories and time periods are provided in

Appendix (5.3).

### 3.4 Prediction

Repeatedly fitting model (M4) across port complex configurations and applying the BMA procedure, ultimately provides access to posterior predictive distributions of the species compositions ($\pi^*_{jklm\eta}$) within a market category and time interval modeled period. A straight forward way to evaluate the performance of the model in each modeled period is to compare the predictions of the model in each modeled period with the actual observations of species compositions from port samplers.

We evaluate species composition posterior predictive distributions via HDI at three levels containing 68%, 95%, and 99% of posterior predictive probability. Tables (2 and 3) show the proportion of observed species compositions which existed within the HDI across all strata, of each prediction level, in each modeled period. For example, observed species compositions for market category 250 in the 1978-1982 time period fell within the 68% HDI of the posterior predictive distribution 67.1% of the time, Table (2)).

| | 68% | 95% | 99% |
|---|---|---|---|
| 250 | 67.1% | 96.1% | 98.7% |
| 253 | 67.3% | 96.3% | 98.9% |
| 262 | 67.4% | 93.8% | 95.3% |
| 265 | 69.6% | 96.0% | 97.8% |
| 269 | 68.2% | 88.8% | 90.2% |
| 270 | 68.6% | 93.6% | 96.7% |
| 956 | 68.3% | 96.7% | 99.2% |
| 959 | 68.5% | 96.3% | 98.1% |
| 961 | 69.3% | 93.2% | 95.3% |
| AVG | 68.3% | 94.5% | 96.7% |

Table 2: 1978-1982

Table (2 and 3) largely shows that the observed proportion of predicted samples aligns appropriately with the predictions made by the model. Considering the average performance across market categories at each prediction level, it appears that prediction is mostly appropriate with the possible exception of the 99% prediction level. The 99% prediction level appears to slightly under-predict on average, indicating that predictive distributions are slightly lighter in the far tails than the data.

| | 68% | 95% | 99% |
|---|---|---|---|
| 245 | 60.8% | 94.9% | 97.7% |
| 250 | 68.1% | 96.0% | 99.0% |
| 253 | 69.3% | 97.1% | 98.9% |
| 259 | 83.8% | 91.9% | 92.9% |
| 262 | 68.5% | 95.1% | 95.9% |
| 269 | 68.6% | 94.2% | 94.7% |
| 270 | 67.9% | 94.2% | 96.7% |
| 663 | 68.1% | 94.1% | 96.3% |
| 667 | 69.4% | 92.5% | 93.5% |
| 956 | 67.5% | 96.2% | 99.0% |
| 959 | 67.4% | 96.4% | 99.0% |
| 960 | 68.0% | 96.1% | 98.6% |
| 961 | 68.6% | 94.6% | 97.8% |
| AVG | 68.9% | 94.9% | 96.9% |

Table 3: 1983-1990

# 4 Discussion

Described here is a flexible modeling approach, building on the work of Shelton (2012), for speciating sparsely sampled commercial landings. The model described here uses Bayesian statistics to allow, but does not force, the data to inform pooling strategies through both time and space via hierarchical modeling and Bayesian model averaging. Our approach infers the level of pooling in time and space from data, it uses the complete Bayesian statistical framework to produce accurate estimates of uncertainty, and makes predictions for unobserved strata.

Generally speaking the model based statistical framework allows tremendous flexibility in accounting for a dynamic port sampling program. Market forces, regulation changes, and fisherman behavior are a few factors, among the many, which complicate the task of speciating commercial catch. Unlike a purely sample-based statistical framework, model based statistics allows analysts to quickly explore a wide range of hypotheses for estimating species compositions. The models entertained here manage to achieve generally well behaved predictive accuracy (Tables 2 and 3), however these models are by no means perfect. The models presented here simply offer a few fundamental improvements toward estimating species compositions.

Among the largest structural changes improving from the Bayesian methodology in Shelton et. al. (2012) is the recognition of overdispersion in port sampling data. In the absence of highly predictive covariates, modeling overdispersion in port sampling data remains an important modeling consideration. Moving forward, modeling decisions will require a careful consideration of predictive accuracy and bias/variance trade-off, so as to tease better and better performance out of further models. The models presented here offer a great operational starting point and provide a basic framework for further model exploration.

This system provides easy access to estimates of uncertainty in commercial catch, at any aggregation of the stratification system. Making posterior predictive draws from these models widely available, makes it possible for stock assessment scientists to incorporate catch uncertainty into existing assessment models (Methot and Wetzel, 2013) where possible and appropriate.

Due to observed trends in the sampling of more heavily landed, or specious, market categories, the vast majority of commercial landings from 1978-1990 are able to be expanded by our model. Moving into the modern era, regulation changes including the start of the live fish fishery, and the proliferation in the number of market categories, due to mandatory sort requirements, may challenge species composition estimation. However, due to the increased number of strata that these changes introduce, uncertainty estimation for these time periods will prove to be critical. Without a matched increase in sampling effort, alongside increased stratification, the number of samples per stratum falls dramatically and species composition estimation may well become very uncertain.

Given a data sparse setting, model-based strategies of catch estimation provide the best chance of a full statistical treatment of available data. However, a more informed path forward involves either increasing sampling effort, or a simplification of the stratification system. Either of these changes provide models with more data to better infer parameters. Model flexibility and justifiable stratum pooling strategies, will become vital for modeling data-sparse time periods. Although estimates are likely estimable in these sparse time periods, as pooling strategies become more extreme, model fit will suffer as both bias and variance estimates increase.

Future effort in developing models should include an exploration of the effect of landing weight on species compositions, as the current estimation algorithm in CALCOM uses landing weight information in it's calculation. The model-based approach makes testing this hypothesis straight-forward, as the hypothesis may amount to the inclusion of a single slope parameter in the linear predictor, regressing on landing size. Given the current model's agreement with existing data, as well as comlands estimates for well sampled strata, it is unlikely that landing size has an important predictive effect on estimates, however without testing the hypothesis, we can not say whether the effect will prove to be explanatory.

In an attempt to add further flexibility to the models presented here, exploring the possibility of gear-species interactions, as random effects, may prove fruitful. This could improve model performance not only due to differential gear selectivity, but also because some gear groups and port complexes are confounded in the California data due to spatial regulations (e.g. trawl gear is prohibited in Southern California). Furthermore the inclusion of random vessel effects may also find support, perhaps capturing variability in fishing behavior or changes in catch composition related to vessel size, as noted above.

Finally, further large changes to the methods proposed here might include a true multivariate handling of the likelihood. The Beta-Binomial univariate model, used here, suggests that the multivariate Dirichlet-Multinomial extension might be a good model for these data. We have yet to get these models to practically compute (e.g. excessive run times), they may provide appropriate structure across the many species of this system. As affordable computing power increases, predictive distributions and summary statistics from improved models can be easily incorporated into the proposed database structure.

The BMA procedure presented here adds significant robustness and pooling potential to our species composition estimates, however it does so at a substantial computational cost. We have found ways (through parallel computation and constraining the model search) to make the computation tractable, however the solution is a "brute force" approach. Dirichlet Process (DP) Bayesian nonparameteric models (T.S. Ferguson, 1973) are a potential alternative approach to modeling the port parameters.

This approach and analysis is intended to be a template for applying to other highly, or even just moderately comparable examples. For instance, this approach could be applied to historical rockfish catches off of Oregon and Washington, which have a different but comparable history of sampling mixed species market categories that have shifted over time. More importantly, this general approach will be appropriate to additional investigations of the species composition of total landings in other poorly sampled and highly variable mixed stock fisheries that may have widely variable types of available data (e.g. Suter 2010, Benoit 2012, Saldaña-Ruiz et al. 2017, Fields et al. in press).

# 5 Appendix

## 5.1 Appendix A: Modern Market Categories and Sampling

Figures (7 and 8) show structure in port sampling data for two modern unmodled time periods, 1991-1999 and 2000-2015. These modern time periods show similar patterns in port sampling effort as the above describes modeled time periods, although due to the increased number of market categories, port sampling is spread more thinly among the many categories. Namely port sampling effort still seems to track both total landed weight, as well as the number of species in each market category, however this pattern occurs across many more market categories.

In these two modern time periods a key feature of the data is the proliferation of the number of market categories. Figures (7 and 8) show market categories accounting for the top 99% of total landings in each time period. In the modeled periods, 1978-1982 and 1983-1990, the top 99% of total landings are landed into 6 and 12 distinct market categories respectively. In the unmodeled periods shown here, 1991-1999 and 2000-2015, the top 99% of total landings are landed into 20 and 28 distinct market categories respectively.

It was noted in Section (3.1) that market category 267 (nominally Brown Rockfish) was actually composed of a relatively little Brown Rockfish, while instead the market category actually contains mostly Widow Rockfish. Figure (8) shows that in the most modern time period (2000-2015) this pattern reverses as market category 267 is composed almost entirely (99.6%) of Brown Rockfish.

## 5.2 Appendix B: A Motivating Example

To discern between these discrete modeling options we considered Poisson, binomial, negative binomial, and beta-binomial models fit to a subset of the data from market category 250, in the Monterey port complex trawl fishery for the second quarter of 1982. This stratum was selected as a relatively data rich setting, although other stratum produce similar results. This stratum was visited 32 times by port samplers, collecting a total of 59 cluster samples across 55 unique species. For brevity, in this example, we only consider the six most prevalent species (BCAC, CLPR, WDOW, YTRK, BANK, STRK).

Simplified models under each of the discrete likelihoods, mentioned above, are fit to the subset data.

$$y_{ij} \overset{i.i.d.}{\sim} p(y_{ij}|\theta_j, \phi)$$

Here $p$ takes the form of each of the considered Poisson, binomial, negative binomial, and beta-binomial models, $\theta_j$ represents the fixed species parameters, and $\phi$ is included to generally represent the nuisance parameters for modeling overdispersion in the negative binomial, and beta-binomial models.

Table (4) shows model fit measures spanning a wide range of model selection philosophies and yet here they all consistently agree in ranking the likelihood models, with a clear preference for the overdispersed models (NB and BB). The beta-binomial model shows the most overall support and the Poisson model shows the least support. This initial result guides the use of the beta-binomial data generating model for the purposes of building a model to apply at an operational scale.

Figure (10) demonstrates how the beta-binomial model distributes predictive density

| | Poisson | Binomial | NB | BB |
|---|---|---|---|---|
| MSE | 0.06412 | 0.06264 | 0.05171 | 0.04479 |
| $\Delta$ DIC | 1001.41 | 1230.60 | 5.03 | 0 |
| $\Delta$ WAIC | 1079.95 | 1323.75 | 3.43 | 0 |
| $pr(M\|y)$ | $\approx 0$ | $\approx 0$ | $\approx 10^{-7}$ | $\approx 1$ |

Table 4: A table showing Mean Squared Error (MSE; computed on the species composition scale), delta deviance information criterion ($\Delta$ DIC), delta widely applicable information criterion ($\Delta$ WAIC), and marginal Bayesian model probabilities ($pr(M|y)$) across the likelihood models fit.

over the unit interval. Species composition is bounded on [0, 1], thus in the presence of large variability, predictive density may aggregate around the bounds.

Figure (11) visualizes the predictive species composition distributions as 95% Highest Density Intervals (HDI) (colored vertical lines), plotted on top of the predictive means for each model and the observed species compositions (black horizontal lines) from the data in Figure (11).

The large spread of the observed species compositions seen in Figure (11) visually demonstrate the degree of overdispersion present in port sampling data. The Poisson and binomial models disregard this overdispersion to prioritize fitting the data mean. The NB and BB models explicitly model overdispersion in the data, and as such they predict a larger subset of the data. Notably, only the intervals produced by the BB model include the low observed proportions of bocaccio (BCAC) and the high observed proportion of chilipepper rockfish (CLPR) in this example.

The split beta-binomial intervals seen in Figure (11) reflect a large amount of residual variability confined on the unit interval. The beta-binomial is the only model considered here, that estimates such a large degree of variability and thus it is the only model that produces predictive species composition distributions that effectively cover the range of observed species compositions. The predictive intervals in Figure (11) are the smallest possible regions on each of the densities visualizes in Figure (10) so that each intervals contain 95% probability. For the example of STRK, notice that although the predictive HDI in Figure (11) is split, the vast majority of density (seen in Figure (10)) lies directly atop the data.

## 5.3 Appendix C: Spatial Model Averaging

In an effort to provide a flexible spatial pooling strategy among the categorical port complexes in this system, a "brute force" BMA approach is used. In each modeled time period, and market category, 274 variations of model (M4) are fit. Each of the 274 model variations differ only in how they partition port complexes along the California coast. In the extreme cases, port complexes may be modeled as (10) unpooled parameters, or in contrast, port complexes may be pooled into as few as four parameters, see Section (2.6).

In the case of pooling port complexes it is not enough to simply known how many port complexes groups to pool toward, but rather it is imperitive to know which port complexes to pool together. Here we try all port complex pooling schemes, such that only adjacent port complexes are ever pooled, and pooling schemes shall never exceed three port complexes in size. For the case of a 10 port complexes system this results in 274 variations of model (M4), although for a 3 port complex system this only amounts to 4 model variations. In each modeled period and market category all 274 model variation are fit and in an automated fashion these models are averaged togehter so as to approximatly integrate port complex poolings descisions out of the modeling procedure.

Figures (12, 13 and 14) show the spatial model averaging results for all modeled time periods (1978-1982 and 1983-1990) in all modeled market categories. For brevity we only show the five most highly weighted models in each market category.

In 1978-1990, nine market categories are computed by the model. Results from all nine BMA partitioning schemes are shown in Figure (12). Market category 250 is described in more detail in Section (3.3) as it is the most heavily landed and sampled category in the 1978-1982 time period, see Figure (2). Of additional note in the 1978-1982 time period are market categories 270 and 959. The most heavily weighted model in market category 270 ($\omega = 0.08$) includes 7 port complex parameters, with no port complex pooling seen south of San Francisco. Furthermore only two port complex parameters are used from San Francisco north (CRS/ERK and BRG/BDG/OSF). In the north the break between ERK and BRG (i.e. Cape Mendicino) is a very consistent feature, while in the south there is more variability in how to partition port complexes. Additionally it is work noting that the top 5 models shown in market category 270 only amount to about 24% of model probability with the remaining 76% extending into the 269 unshown models for that market category. This represents an example of substantial model uncertainty with respect to the southern port complex model specification. In contrast, consider market category 959, where the first two models amount to 98% of model weight.

In the 1982-1990 time period, see Figures (13 and 14), where 13 market categories were modeled. In contrast with 1978-1982, the 1983-1990 time period shows a lot of variability in the number of parameters used to model port complex. In 1983-1990 market category 262 has its most heavily weighted model only using 4 parameters while market category 663 uses 8 port complex parameters in its most heavily weighted model. Again in market category 663 most of the parameters are used to model the southern port complexes.

In general, across all time periods, notice that each market category has fairly unique port complex partitioning results. The fact that each market category behaves uniquely indicates the complexity of this system. Furthermore the fact that the BMA strategy picks up on these varied market category behaviors indicates the flexibility of this approach. Although the system is very dynamic, key breaks along California biogeographic features, such as Cape Mendocino and/or Point Conception, seem to be recurrent patterns.

Additionally, across all of the modeled periods no market category ever highly weights the fully unpooled model. This suggests the appropriateness of spatial pooling for these sparse data. These data represent the most well sampled time periods in CALCOM an still spatial pooling finds support. Furthermore it is worth reiterating that the fitted model, (M4), is already doing hierarchical temporal partial pooling, as described in Section (2.3), and ever still pooled spatial models outperform the unpooled models.

## 5.4 Appendix D: Nuisance Parameters

Tables (5 - 8) Summarize nuisance parameter posteriors for model M4 fitted in each market category, in each modeled period. Recall high values of $\rho$ indicate overdispersion relative to the binomial model, and small values of $v$ indicate a high degree of temporal pooling in the $\beta_{m\eta}^{(y:q)}$ parameters.

### 5.4.1 Overdispersion Parameter($\rho$)

In general $\rho$ estimates seem to account for a fair degree of overdispersion. Values of $\rho$ never approach the maximal limit ($\rho = 1$), and thus the beta-binomial model seems to be appropriately to modeling the observed residual variance of these data, on average, without structurally underestimating variability.

| MCAT | Mean | Median | SD |
|---|---|---|---|
| 250 | 0.55 | 0.55 | 0.004 |
| 253 | 0.39 | 0.39 | 0.001 |
| 262 | 0.35 | 0.35 | 0.008 |
| 265 | 0.64 | 0.64 | 0.002 |
| 269 | 0.52 | 0.52 | 0.019 |
| 270 | 0.53 | 0.54 | 0.020 |
| 956 | 0.35 | 0.35 | 0.007 |
| 959 | 0.47 | 0.47 | 0.070 |
| 961 | 0.55 | 0.55 | 0.004 |

Table 5: 1978-1982: $\rho$ posterior mean, median, and standard deviation summaries for all modeled market categories.

| MCAT | Mean | Median | SD |
|---|---|---|---|
| 245 | 0.65 | 0.65 | 0.014 |
| 250 | 0.51 | 0.51 | 0.002 |
| 253 | 0.47 | 0.47 | 0.010 |
| 259 | 0.75 | 0.75 | 0.009 |
| 262 | 0.41 | 0.41 | 0.001 |
| 269 | 0.57 | 0.57 | 0.046 |
| 270 | 0.74 | 0.75 | 0.027 |
| 663 | 0.51 | 0.51 | 0.001 |
| 667 | 0.49 | 0.49 | 0.022 |
| 956 | 0.43 | 0.43 | 0.003 |
| 959 | 0.55 | 0.55 | 0.004 |
| 960 | 0.45 | 0.45 | 0.004 |
| 961 | 0.59 | 0.59 | 0.001 |

Table 6: 1983-1990 $\rho$ posterior mean, median, and standard deviation summaries for all modeled market categories.

### 5.4.2 Temporal Pooling($v$)

Each modeled market category and time period require a different degree of pooling. The variance estimates span 2 orders of magnitudes with the smallest estimate (indicating the most temporal shrinkage) occurring in market category 250 in 1983-1990 and the largest estimate (indicating the least temporal shrinkage) occurring in market category 269 in 1983-1990. Modeling each market category separately provides the flexibility to separately characterize these largely distinct temporal pooling behaviors.

| MCAT | Mean | Median | Posterior SD |
|---|---|---|---|
| 250 | 12915.85 | 18523.12 | 8699.87 |
| 253 | 22747.87 | 23063.76 | 1535.53 |
| 262 | 20254.41 | 20506.36 | 2581.87 |
| 265 | 15846.22 | 16694.98 | 7601.15 |
| 269 | 20135.05 | 19975.15 | 4667.11 |
| 270 | 19931.96 | 19955.13 | 6033.35 |
| 956 | 19659.11 | 19795.60 | 1227.99 |
| 959 | 19159.69 | 13375.80 | 19256.94 |
| 961 | 18631.44 | 19498.31 | 7970.44 |

Table 7: 1978-1982 $v$ posterior mean, median, and standard deviation summaries for all modeled market categories.

| MCAT | Mean | Median | Posterior SD |
|---|---|---|---|
| 245 | 20211.82 | 20204.95 | 1276.83 |
| 250 | 236.03 | 192.53 | 134.67 |
| 253 | 20455.18 | 20140.50 | 1521.72 |
| 259 | 20246.14 | 20186.61 | 898.99 |
| 262 | 20445.49 | 20348.56 | 343.70 |
| 269 | 34386.49 | 25951.03 | 24030.32 |
| 270 | 20253.34 | 19908.07 | 9269.02 |
| 663 | 19563.87 | 19624.09 | 331.04 |
| 667 | 20089.55 | 20078.27 | 2723.34 |
| 956 | 20581.67 | 20664.71 | 913.92 |
| 959 | 19242.41 | 18707.09 | 5076.03 |
| 960 | 20059.66 | 20012.80 | 1703.89 |
| 961 | 20127.69 | 20141.04 | 580.80 |

Table 8: 1983-1990 $v$ posterior mean, median, and standard deviation summaries for all modeled market categories.

# 6 Figures

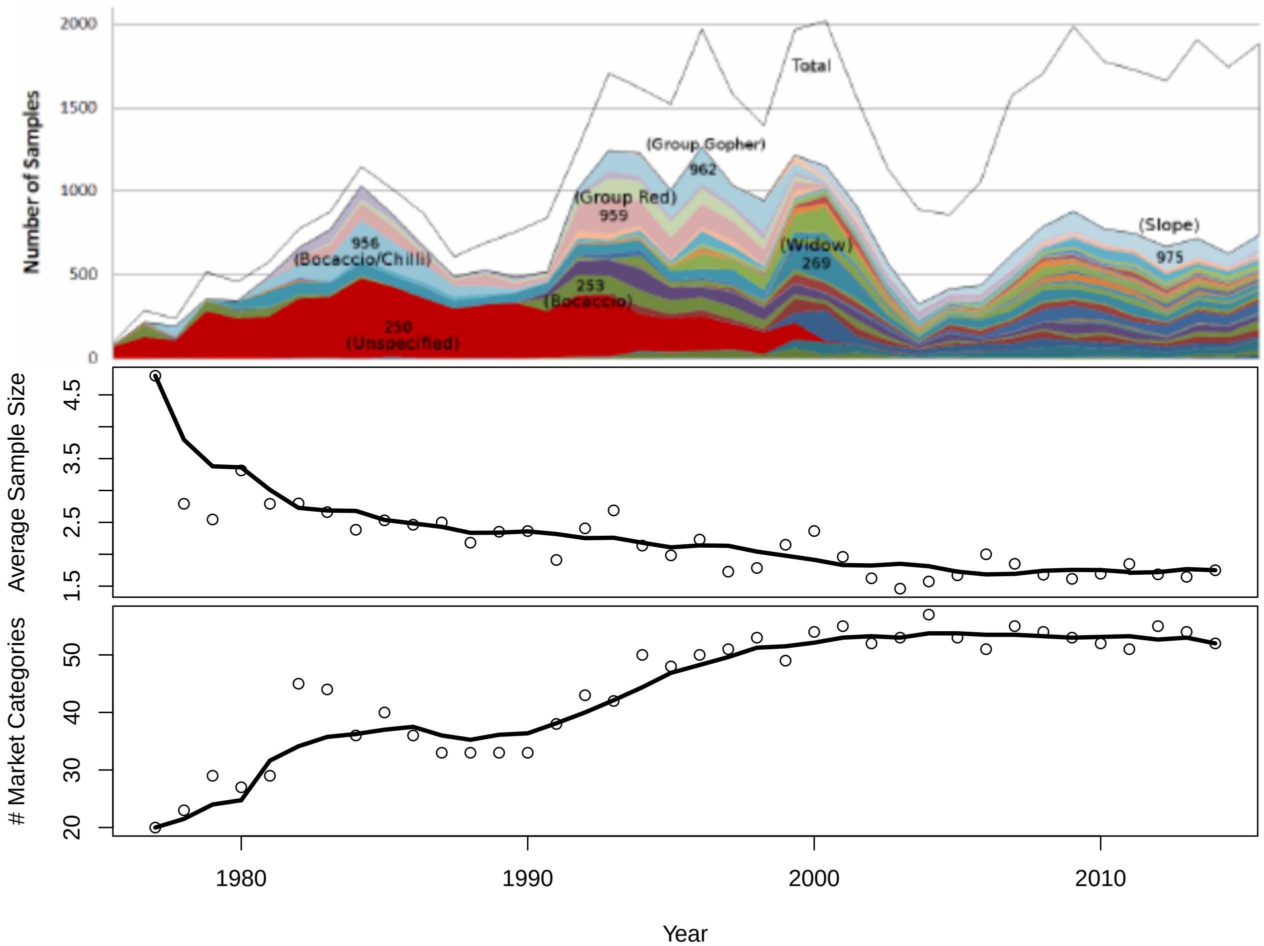


Figure 1: Number of commercial port samples per market category in California, 1978-2014 (upper panel), average sample size per stratum (middle panel), and number of market categories recorded on landing receipts (lower panel). On the lower panels, points indicate observed values, while the black lines represent 8 year moving averages

| | | | | 1978 - 1982 | | | 1983 - 1990 | | |
|---|---|---|---|---|---|---|---|---|---|
| Category | Market Category | Description | Nominal Species or Group | Tons | # Strata | # Samples | Tons | # Strata | # Samples |
| Multi-species | 250 | Rockfish, unspecified | UNSPECIFIED ROCKFISH | 36539.3 | 524 | 1021 | 55332 | 1048 | 2933 |
| | 262 | Thornyheads | THORNYHEADS, UNSPECIFIED | 8512.2 | 202 | 237 | 27929 | 406 | 392 |
| | 956 | Rockfish, group bocaccio/chili | UNSPECIFIED ROCKFISH | 3213.7 | 47 | 127 | 20227 | 655 | 870 |
| | 957 | Rockfish, group bolina | UNSPECIFIED SHELF ROCKFISH | 27.6 | 27 | 0 | 417 | 426 | 1 |
| | 958 | Rockfish, group deepwater reds | ROCKFISH GROUP 3 | 16.3 | 1 | 0 | 19 | 10 | 0 |
| | 959 | Rockfish, group red | UNSPECIFIED ROCKFISH | 225.1 | 41 | 9 | 8883 | 843 | 501 |
| | 960 | Rockfish, group small | UNSPECIFIED ROCKFISH | 1.8 | 6 | 2 | 2223 | 439 | 118 |
| | 961 | Rockfish, group rosefish | ROCKFISH GROUP 6 | 162.1 | 13 | 12 | 4179 | 377 | 327 |
| | 962 | Rockfish, group gopher | UNSPECIFIED ROCKFISH | 0.0 | 0 | 0 | 314 | 225 | 2 |
| | 963 | Rockfish, large red | UNSPECIFIED ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| "Single-species"* | 245 | Rockfish, cowcod | COWCOD | 10.9 | 38 | 1 | 273 | 294 | 31 |
| | 246 | Rockfish, copper (whitebelly)[1] | COPPER ROCKFISH | 6.8 | 18 | 0 | 6 | 93 | 0 |
| | 247 | Rockfish, canary | CANARY ROCKFISH | 0.4 | 1 | 0 | 62 | 51 | 0 |
| | 248 | Rockfish, yelloweye | YELLOWEYE ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 249 | Rockfish, vermilion | VERMILION ROCKFISH | 4.8 | 21 | 0 | 1 | 34 | 0 |
| | 251 | Rockfish, black-and-yellow | BLACK AND YELLOW ROCKFISH | 0.2 | 1 | 0 | 0 | 5 | 0 |
| | 252 | Rockfish, black | BLACK ROCKFISH | 197.4 | 104 | 0 | 403 | 194 | 1 |
| | 253 | Rockfish, bocaccio | BOCACCIO | 14512.9 | 184 | 224 | 1029 | 79 | 44 |
| | 254 | Rockfish, chilipepper | CHILIPEPPER ROCKFISH | 68.7 | 48 | 0 | 90 | 102 | 2 |
| | 255 | Rockfish, greenspotted | GREENSPOTTED ROCKFISH | 10.6 | 6 | 0 | 4 | 7 | 0 |
| | 256 | Rockfish, starry | STARRY ROCKFISH | 2.6 | 10 | 0 | 2 | 7 | 0 |
| | 257 | Rockfish, darkblotched | DARKBLOTCHED ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 258 | Rockfish, China | CHINA ROCKFISH | 78.2 | 68 | 1 | 48 | 147 | 4 |
| | 259 | Rockfish, yellowtail | YELLOWTAIL ROCKFISH | 287.5 | 116 | 0 | 868 | 223 | 11 |
| | 263 | Rockfish, gopher | GOPHER ROCKFISH | 232.4 | 95 | 0 | 35 | 51 | 0 |
| | 264 | Rockfish, pinkrose | PINKROSE ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 265 | Rockfish, yelloweye[2] | YELLOWEYE ROCKFISH | 774.8 | 175 | 27 | 108 | 99 | 0 |
| | 267 | Rockfish, brown | BROWN ROCKFISH | 981.3 | 246 | 9 | 186 | 111 | 3 |
| | 268 | Rockfish, rosy | ROSY ROCKFISH | 0.8 | 3 | 1 | 7 | 14 | 0 |
| | 269 | Rockfish, widow | WIDOW ROCKFISH | 12575.6 | 75 | 132 | 18802 | 374 | 497 |

1. Market category 246 is no longer used since whitebelly rockfish is now considered copper rockfish.
2. Market category 265 was redefined from red rockfish to yelloweye rockfish in 1981 by CDFW.

Figure 2: Landed weight (metric tons), number of landed strata (year, quarter, port complex, and gear group), and number of species composition samples by market category and time period. Market categories created after 1990 are not listed (e.g. 678, 679, 964, and 971-976). *"Single-species" market categories are nominal (in name only); landings in these categories often include a mixture of species.

| | | | | 1978 - 1982 | | | 1983 - 1990 | | |
|---|---|---|---|---|---|---|---|---|---|
| Category | Market Category | Description | Nominal Species or Group | Tons | # Strata | # Samples | Tons | # Strata | # Samples |
| | 270 | Rockfish, splitnose | SPLITNOSE ROCKFISH | 458.7 | 93 | 32 | 3 | 7 | 16 |
| | 271 | Rockfish, Pacific ocean perch | PACIFIC OCEAN PERCH | 175.9 | 65 | 0 | 72 | 60 | 0 |
| | 650 | Rockfish, rougheye | ROUGHEYE ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 651 | Rockfish, olive | OLIVE ROCKFISH | 1.1 | 7 | 0 | 4 | 32 | 0 |
| | 652 | Rockfish, grass | GRASS ROCKFISH | 0.1 | 4 | 0 | 0 | 4 | 0 |
| | 653 | Rockfish, pink | PINK ROCKFISH | 0.1 | 1 | 0 | 0 | 4 | 0 |
| | 654 | Rockfish, greenstriped | GREENSTRIPED ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 655 | Rockfish, copper | COPPER ROCKFISH | 0.4 | 9 | 0 | 43 | 77 | 0 |
| | 656 | Rockfish, blackspotted | BLACKSPOTTED ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 657 | Rockfish, flag | FLAG ROCKFISH | 0.5 | 4 | 0 | 0 | 0 | 0 |
| | 658 | Rockfish, treefish | TREEFISH | 0.0 | 1 | 0 | 0 | 1 | 0 |
| | 659 | Rockfish, kelp | KELP ROCKFISH | 0.0 | 2 | 0 | 0 | 4 | 0 |
| | 660 | Rockfish, honeycomb | HONEYCOMB ROCKFISH | 0.0 | 1 | 0 | 0 | 0 | 0 |
| | 661 | Rockfish, greenblotched | GREENBLTCHED ROCKFISH | 0.1 | 1 | 0 | 0 | 1 | 0 |
| | 662 | Rockfish, bronzespotted | BRONZESPOTTED ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 663 | Rockfish, bank[3] | BANK ROCKFISH | 0.0 | 1 | 0 | 432 | 54 | 15 |
| | 664 | Rockfish, rosethorn | ROSETHORN ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 665 | Rockfish, blue | BLUE ROCKFISH | 176.8 | 117 | 0 | 129 | 194 | 2 |
| | 666 | Rockfish, squarespot | SQUARESPOT ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 667 | Rockfish, blackgill | BLACKGILL ROCKFISH | 9.0 | 3 | 1 | 1213 | 206 | 128 |
| | 668 | Rockfish, stripetail | STRIPETAIL ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 669 | Rockfish, speckled | SPECKLED ROCKFISH | 0.2 | 2 | 0 | 0 | 2 | 0 |
| | 670 | Rockfish, swordspine | SWORDSPINE ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 671 | Rockfish, calico | CALICO ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 672 | Rockfish, shortbelly | SHORTBELLY ROCKFISH | 2.5 | 2 | 0 | 52 | 11 | 1 |
| | 673 | Rockfish, chameleon | CHAMELEON ROCKFISH | 0.0 | 0 | 0 | 0 | 1 | 0 |
| | 674 | Rockfish, aurora | AURORA ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 675 | Rockfish, redbanded | REDBANDED ROCKFISH | 0.0 | 0 | 0 | 1 | 1 | 0 |
| | 676 | Rockfish, Mexican | MEXICAN ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 677 | Rockfish, shortraker | SHORTRAKER ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |
| | 970 | Rockfish, quillback | QUILLBACK ROCKFISH | 0.0 | 0 | 0 | 0 | 0 | 0 |

3. Bank rockfish are sometimes referred to as "red widow" rockfish.

Figure 3: Landed weight (metric tons), number of landed strata (year, quarter, port complex, and gear group), and number of species composition samples by market category and time period. Market categories created after 1990 are not listed (e.g. 678, 679, 964, and 971-976). *"Single-species" market categories are nominal (in name only); landings in these categories often include a mixture of species.

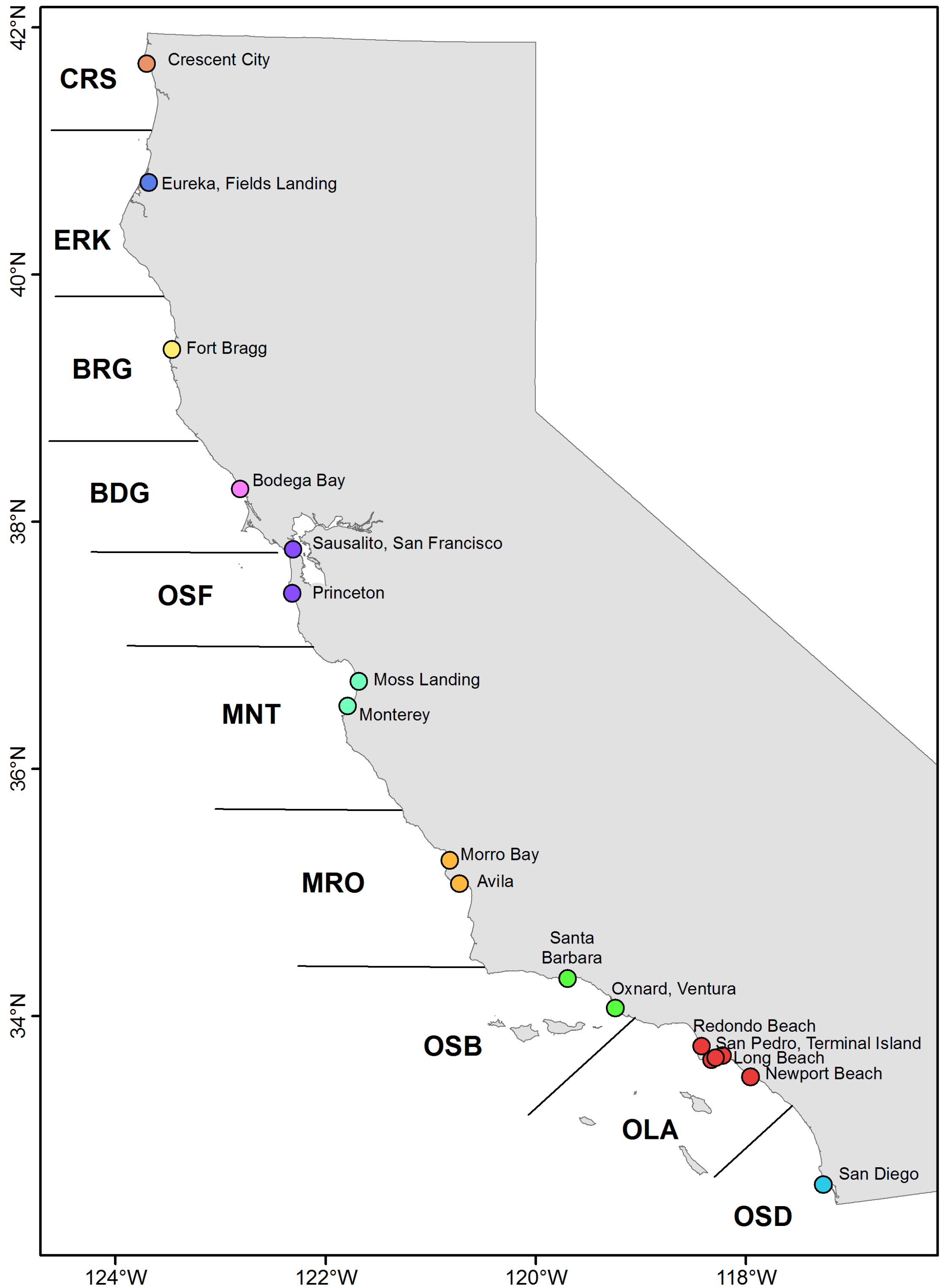


Figure 4: Map showing the ports in California that account for at least 95% of landings. Separating lines show how ports have been aggregated into port complexes.

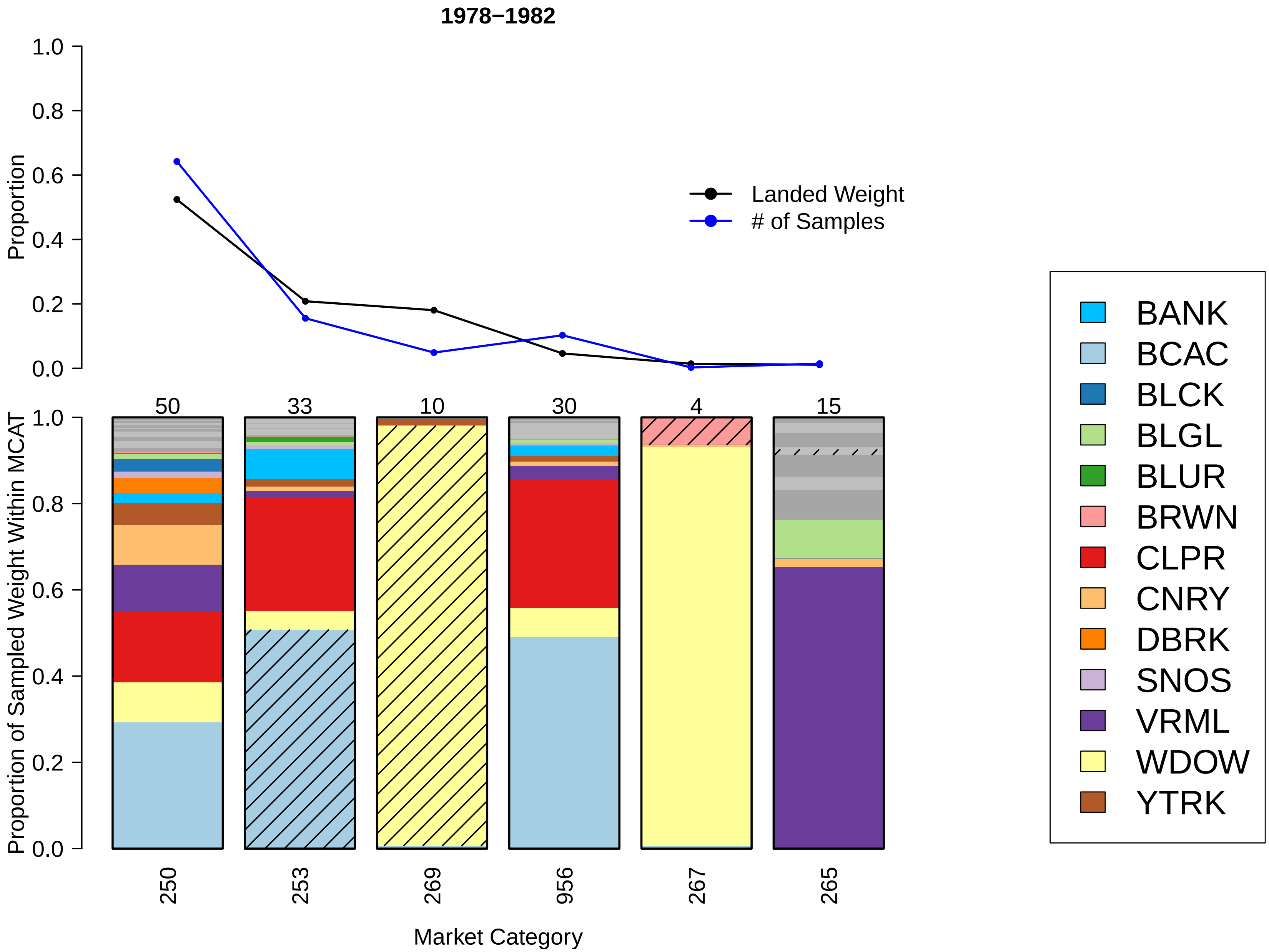


Figure 5: Upper panel shows the proportion of landed weight (black) and number of samples (blue) in each market category for the 1978-1982 time period. Bottom panel shows the proportion of sampled weight for each species in each market category shown. The number above each colored bar indicated the number of species in the market category. Hashing indicates the species that is nominal in relevant the market category.

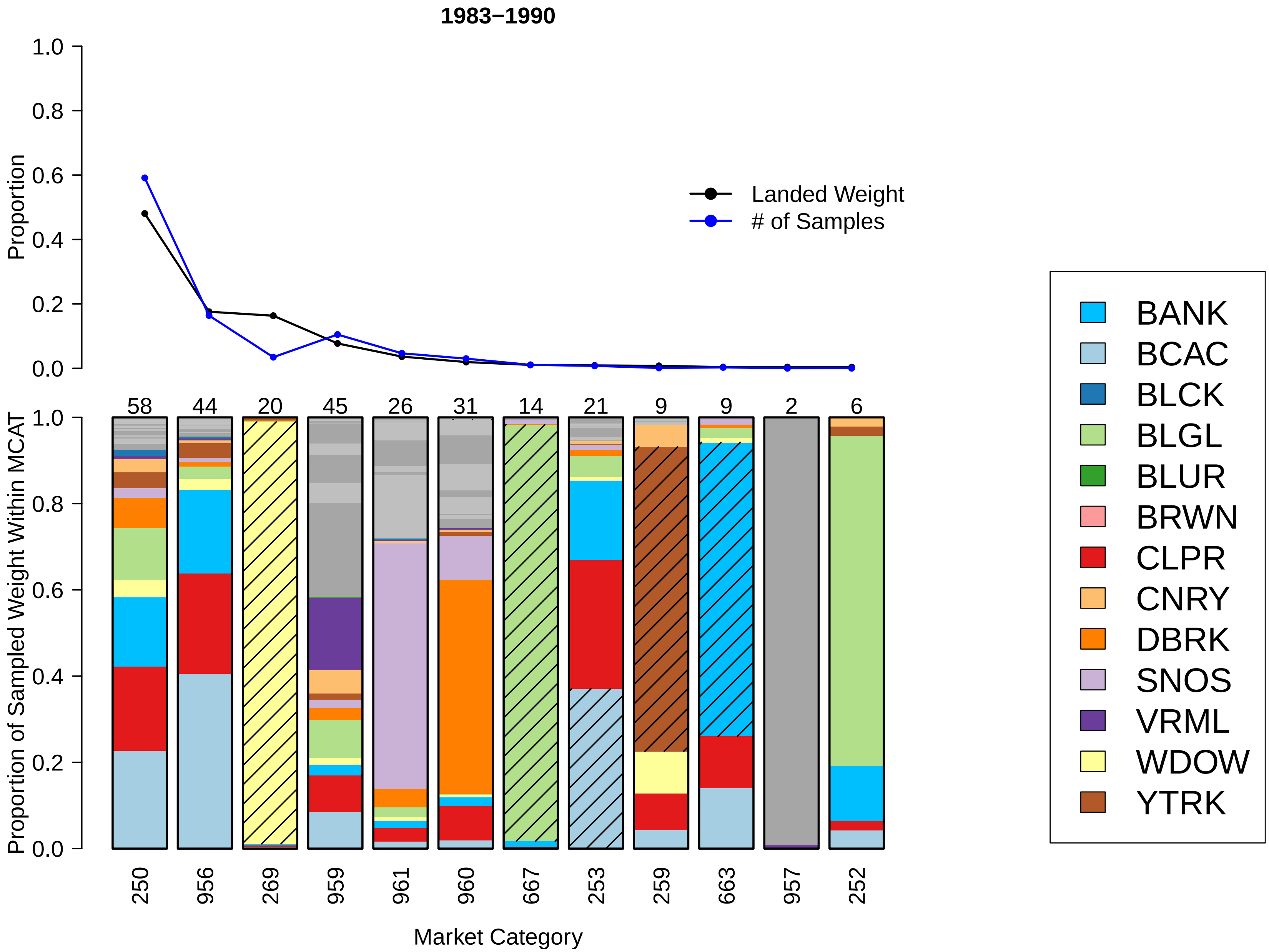


Figure 6: Upper panel shows the proportion of landed weight (black) and number of samples (blue) in each market category for the 1983-1990 time period. Bottom panel shows the proportion of sampled weight for each species in each market category shown. The number above each colored bar indicated the number of species in the market category. Hashing indicates the species that is nominal in relevant the market category.

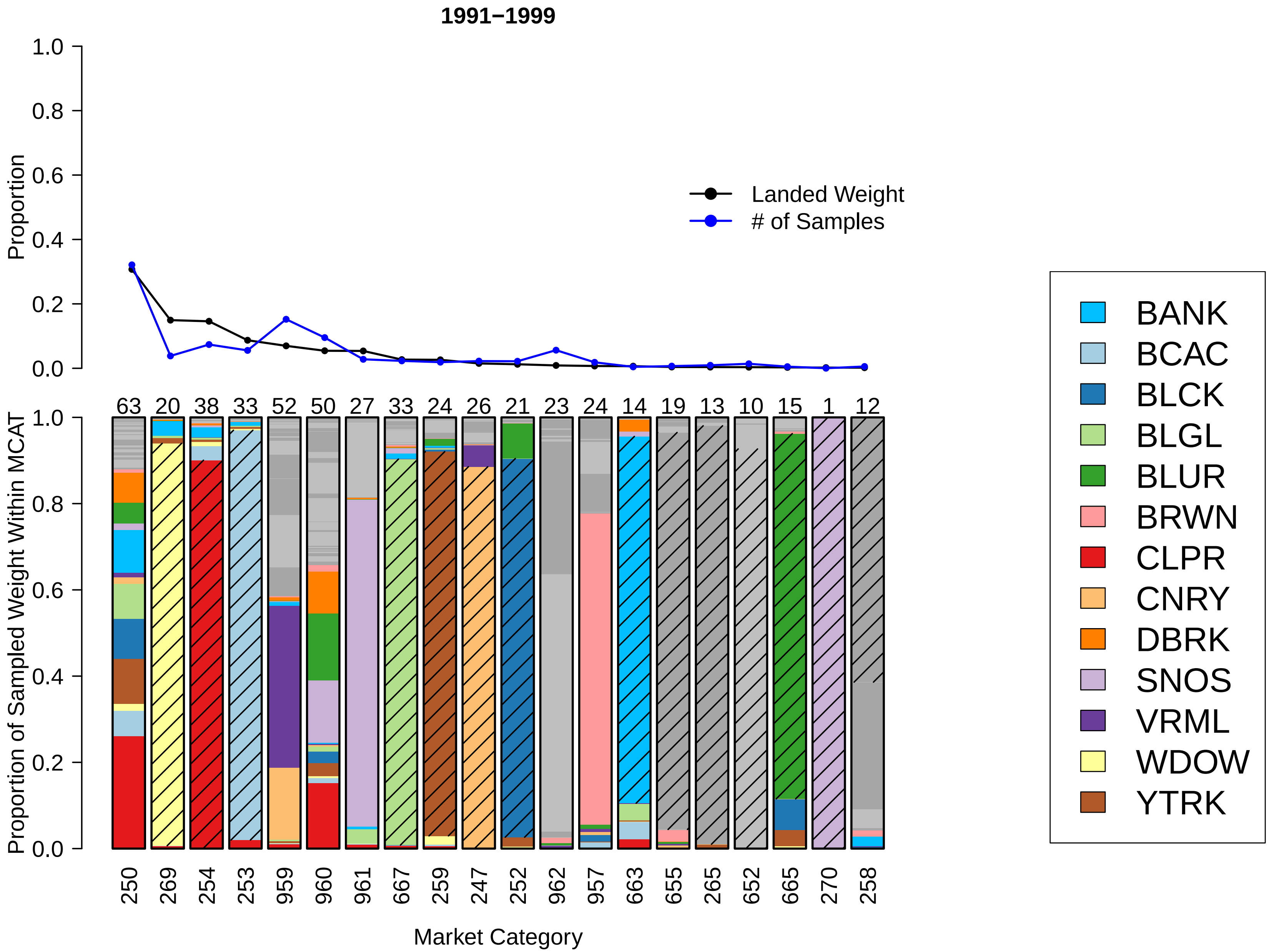


Figure 7: Upper panel shows the proportion of landed weight (black) and number of samples (blue) in each market category for the 1991-1999 time period. Bottom panel shows the proportion of sampled weight for each species in each market category shown. The number above each colored bar indicated the number of species in the market category. Hashing indicates the species that is nominal in relevant the market category.

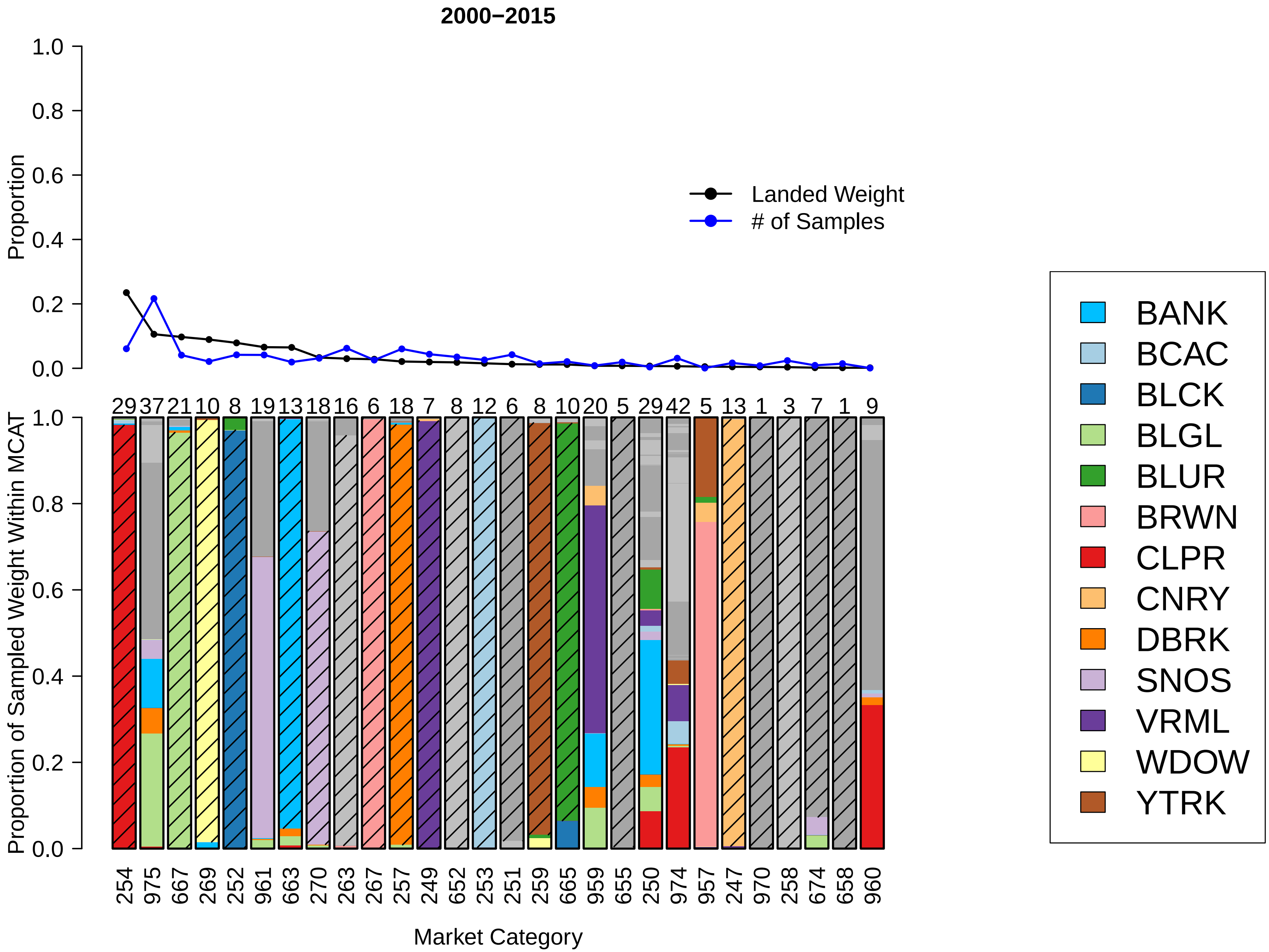


Figure 8: Upper panel shows the proportion of landed weight (black) and number of samples (blue) in each market category for the 2000-2015 time period. Bottom panel shows the proportion of sampled weight for each species in each market category shown. The number above each colored bar indicated the number of species in the market category. Hashing indicates the species that is nominal in relevant the market category.

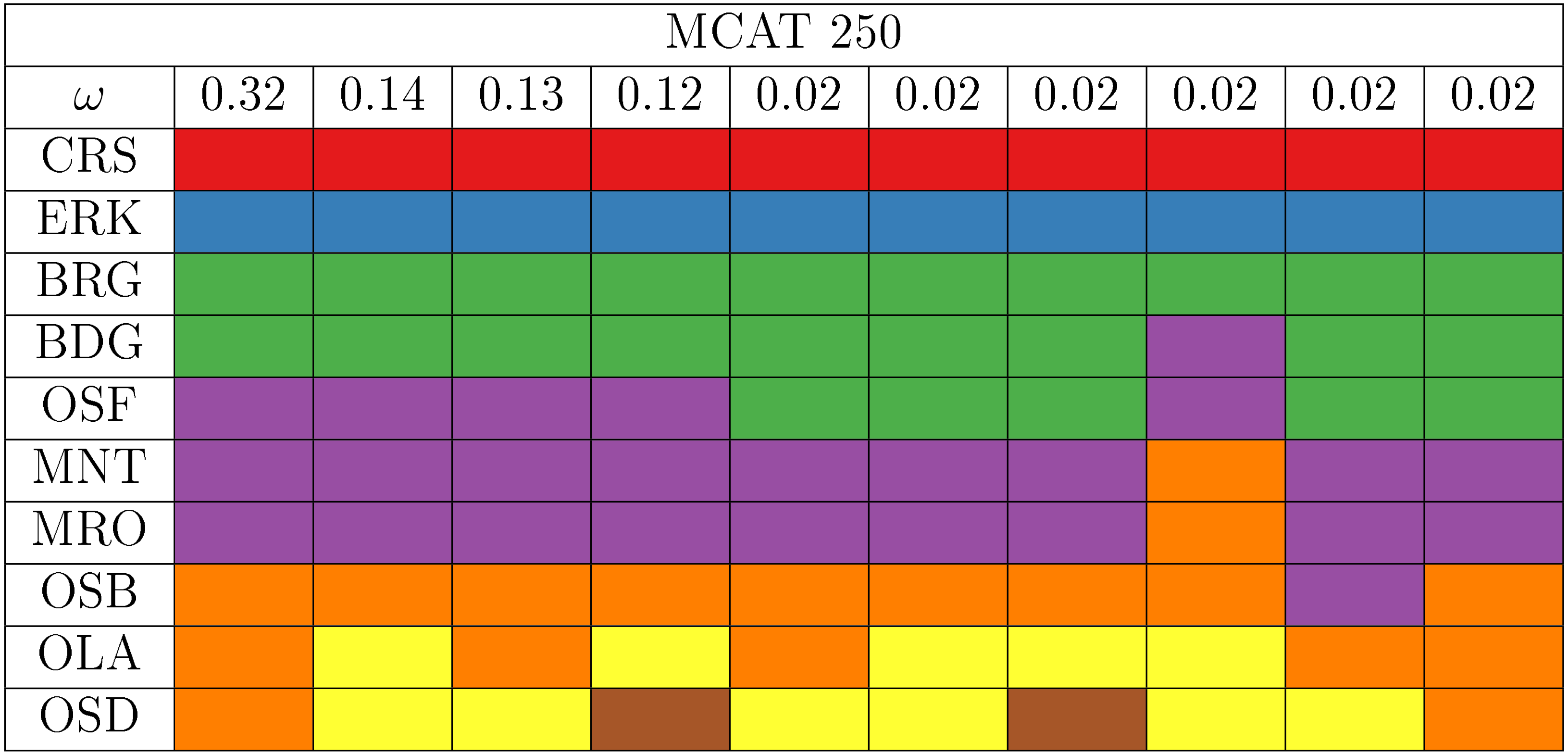


Figure 9: The model averaging results of port complex model selection for market category 250 in 1978 to 1982. BMA weights ($\omega$) for the top 10 models are displayed in the top row (each column is a distinct model). The following ten rows indicate the ten port complexes in California, and the colored cells indicate how port complexes are partitioned in each model.

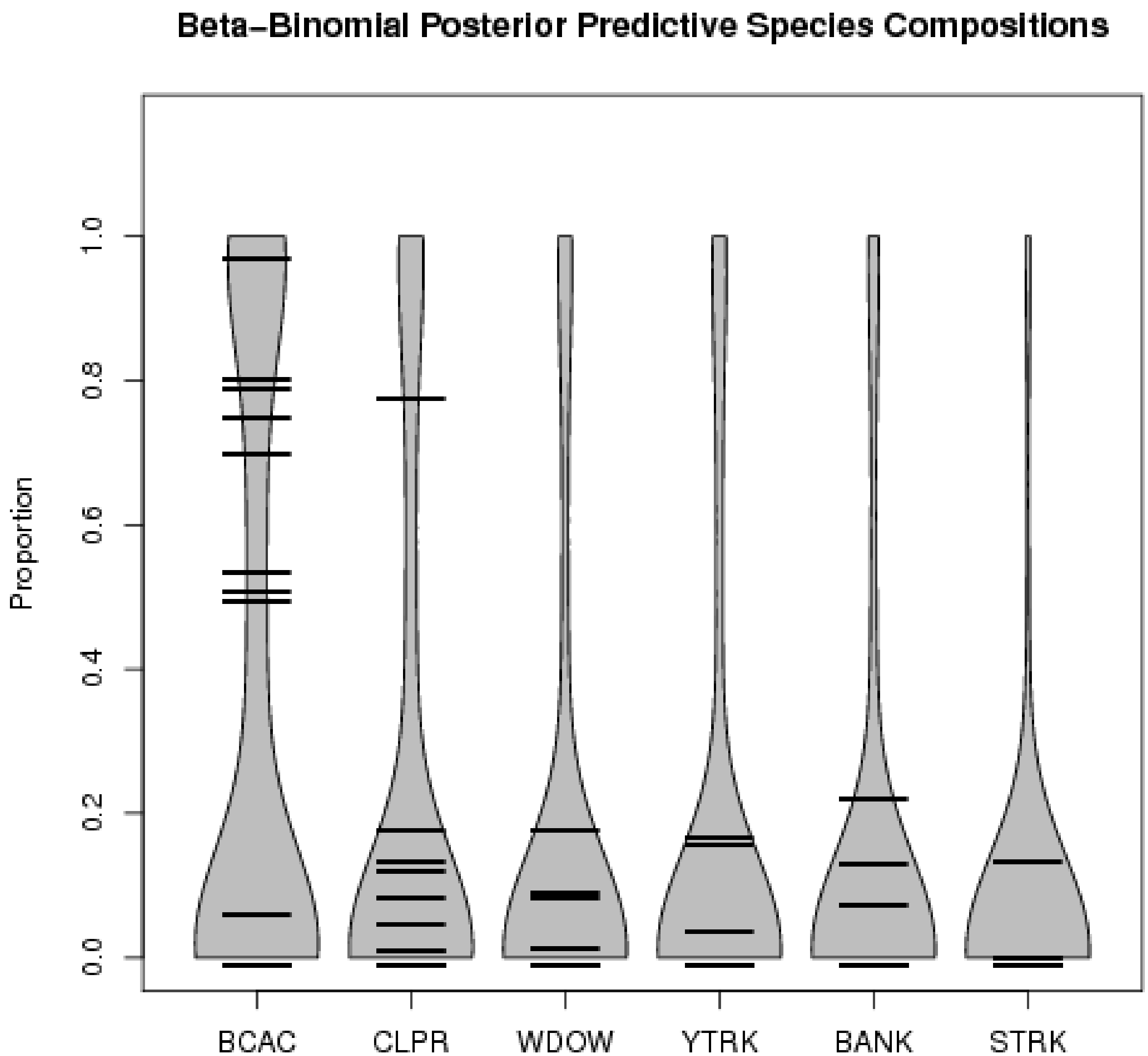


Figure 10: A violin plot of the beta-binomial predictive distributions for the six most prevalent species in the market category 250, Monterey trawl fishery for the second quarter of 1982. Observed species compositions, from port sampling, are plotted atop each density.

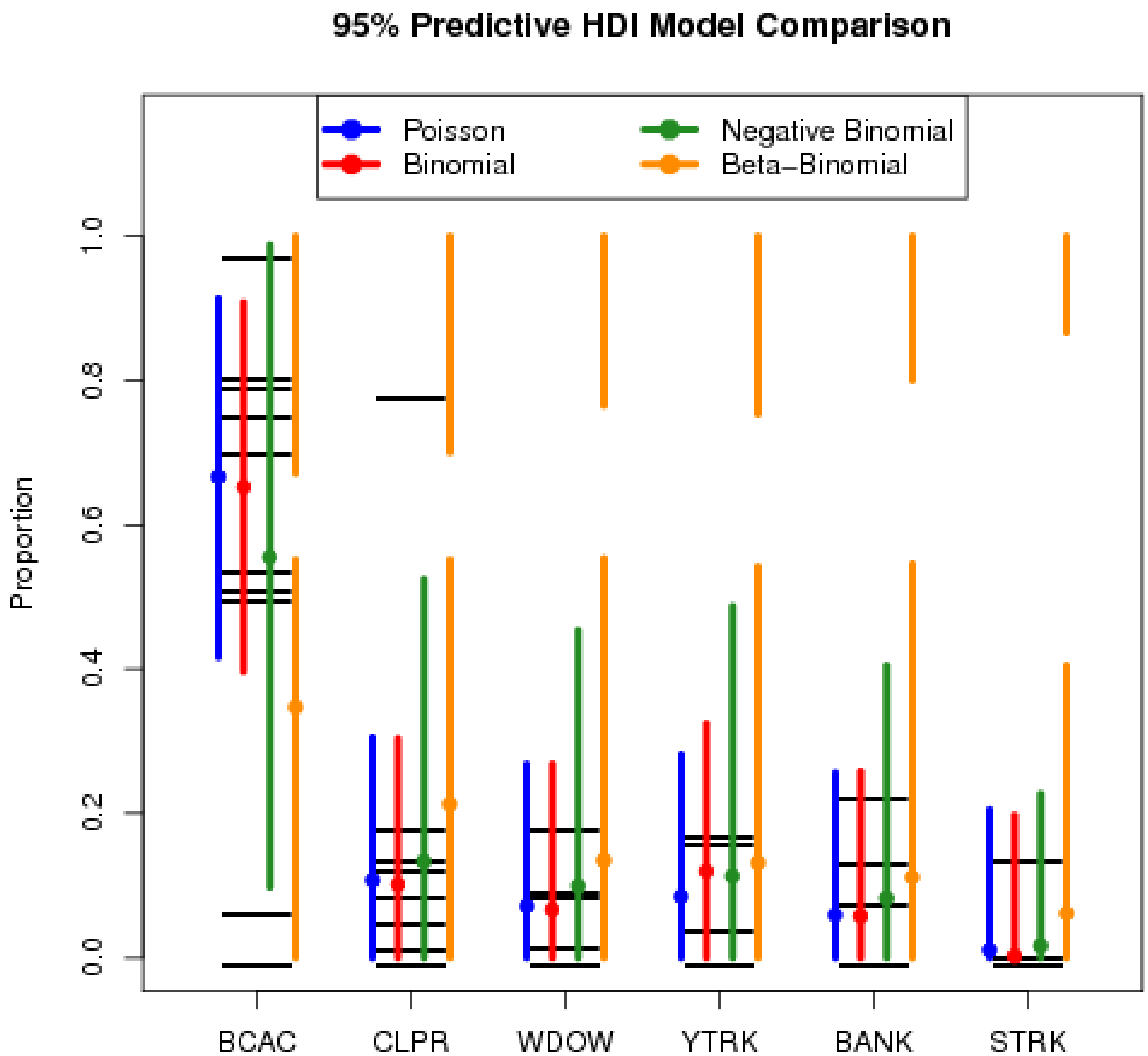


Figure 11: Interval Plot: The predictive species composition distributions as 95% Highest Density Intervals (HDI) (colored vertical lines), plotted on top of the predictive means for each model and the observed species compositions (black horizontal lines) from the data

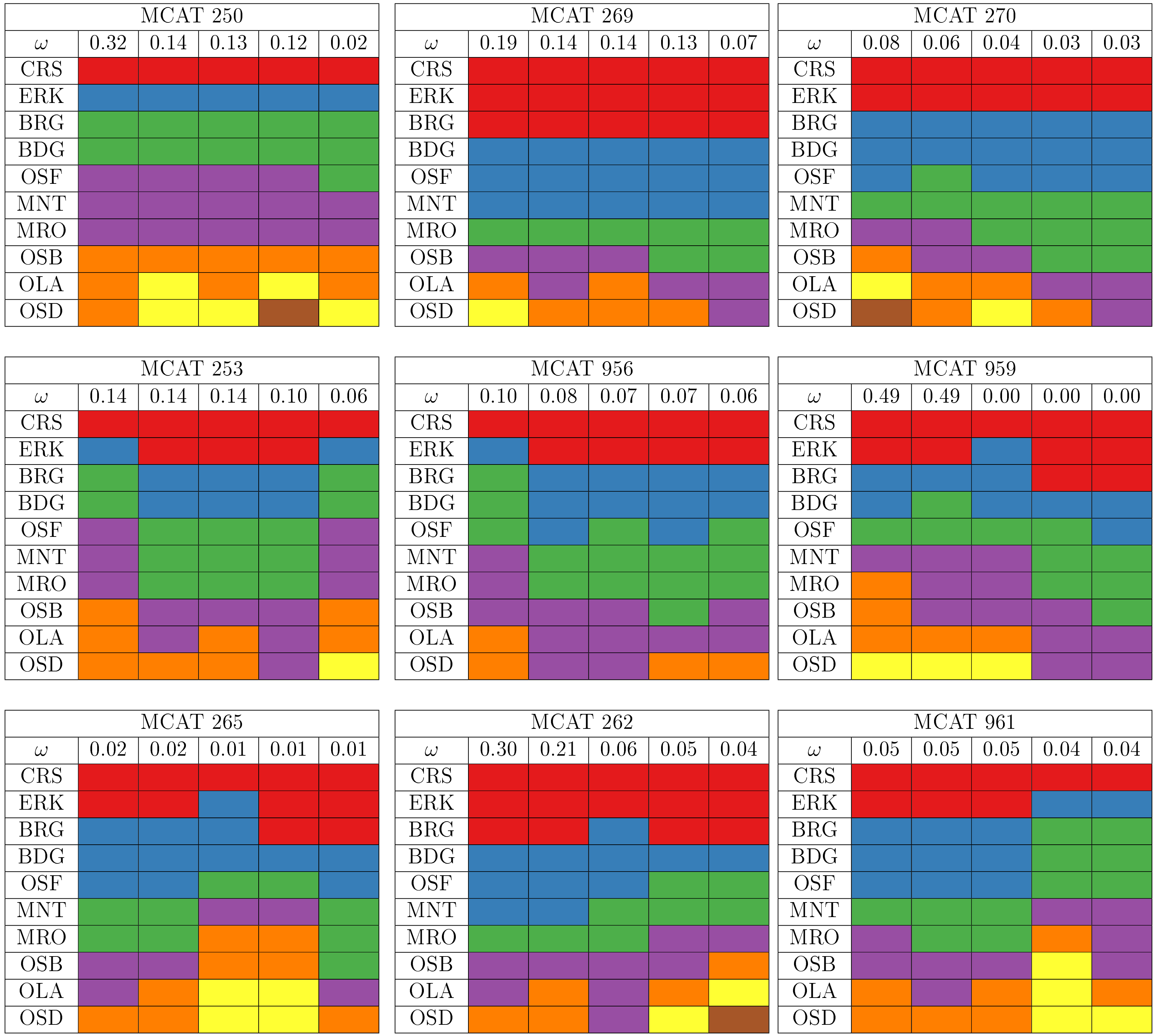


Figure 12: 1978-1982 model averaging results for all modeled market categories.

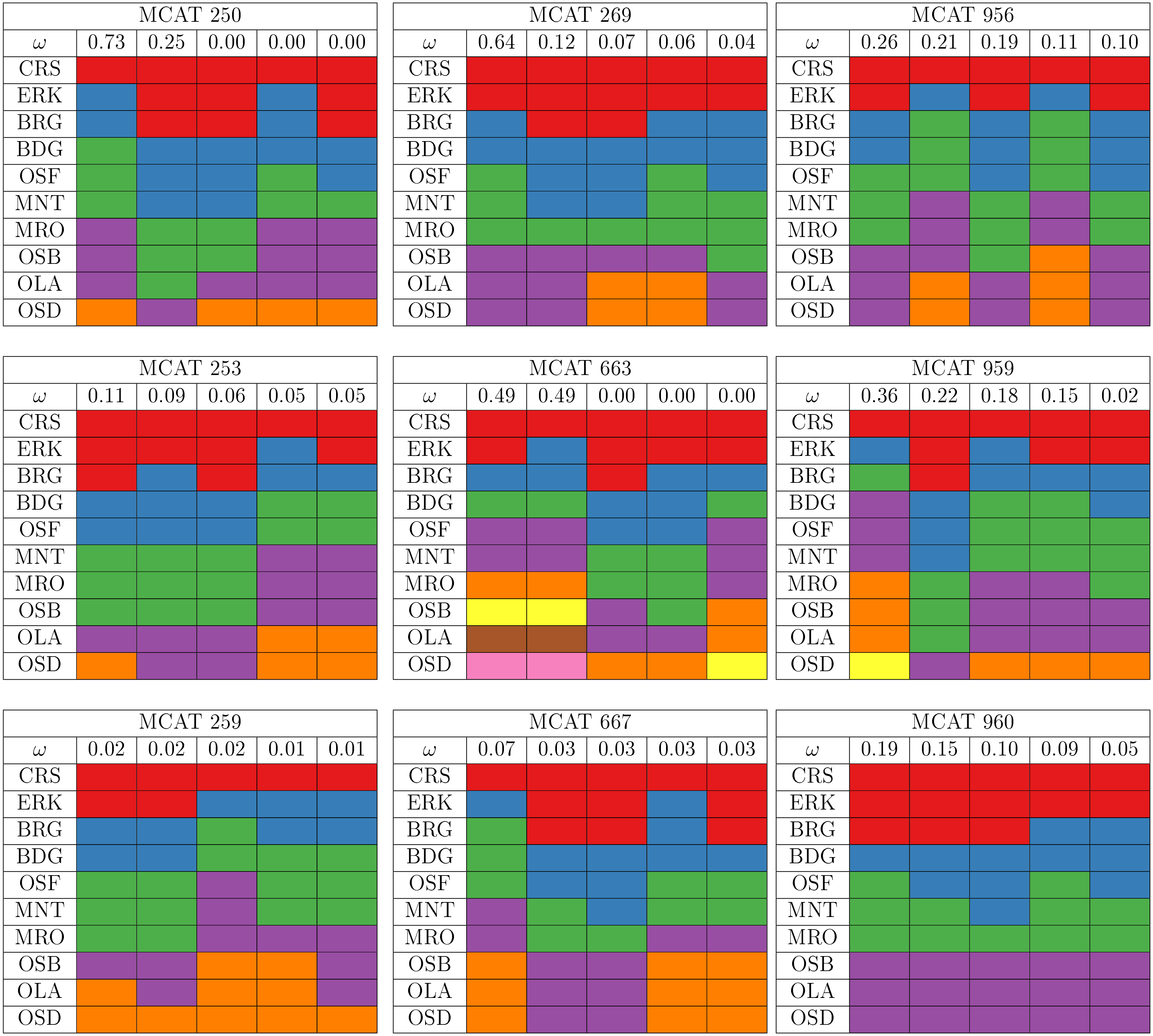


Figure 13: 1983-1990 model averaging results for all modeled market categories.

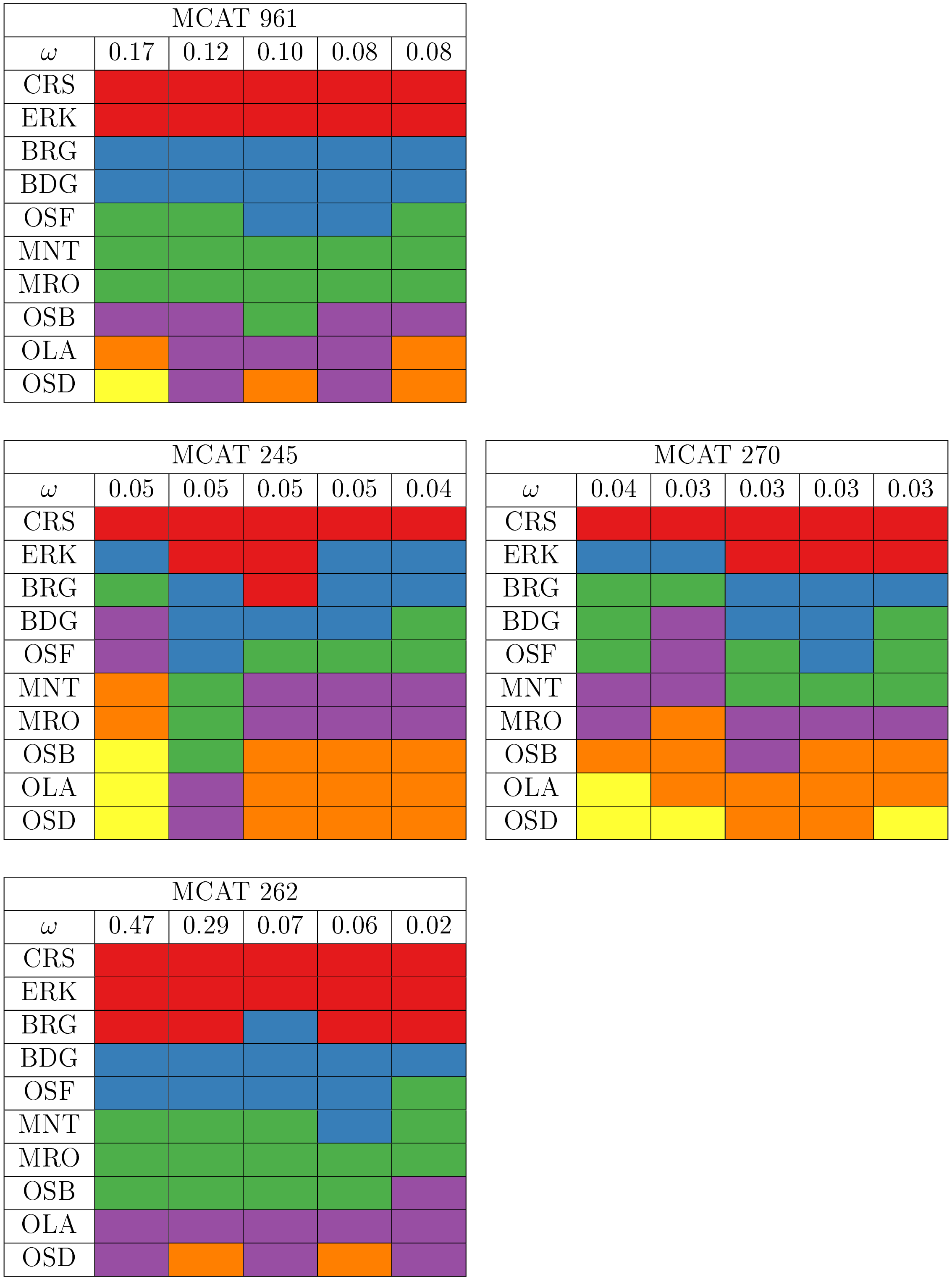


Figure 14: Continued 1983-1990 model averaging results for all modeled market categories.

# References


[1] Benoît, H.P., 2012. An empirical model of seasonal depth-dependent fish assemblage structure to predict the species composition of mixed catches. Canadian journal of fisheries and aquatic sciences, 70(2), pp. 220-232.

[2] Bousquet, N., N. Cadigan, T. Duchesne, and L. Rivest. 2010. Detecting and correcting underreported catches in fish stock assessment: trial of a new method. Canadian Journal of Fisheries and Aquatic Sciences 67:1247-1261.

[3] CALCOM, 2018. California Cooperative Groundfish Survey Database: CDFG, Belmont, CA; PSMFC, Belmont, CA; NMFS, Santa Cruz, CA. http://calcomfish.ucsc.edu.

[4] CCGS (California Cooperative Groundfish Survey). (2017). Pacific States Marine Fisheries Commission, 350 Harbor Boulevard, Belmont, California, 94002. URL: 128.114.3.187.

[5] Checkley Jr, D. M., & Barth, J. A. (2009). Patterns and processes in the California Current System. Progress in Oceanography, 83(1-4), 49-64.

[6] Cochran, W. G. (1977). Sampling Techniques: 3d Ed. Wiley.

[7] Crone, P. R. (1995). Sampling design and statistical considerations for the commercial groundfish fishery of Oregon. Canadian Journal of Fisheries and Aquatic Sciences, 52(4), 716-732.

[8] Doubleday, W. G. (1976). A least squares approach to analyzing catch at age data. Int. Comm. Northwest Atl. Fish. Res. Bull, 12(1), 69-81.

[9] Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. The annals of statistics, 209-230.

[10] Fields, A.T., Fischer, G.A., Shea, S.K., Zhang, H., Abercrombie, D.L., Feldheim, K.A., Babcock, E.A. and Chapman, D.D. in press. Species composition of the international shark fin trade assessed through a retail-market survey in Hong Kong. Conservation Biology. (page proofs available online).

[11] Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014). Bayesian data analysis (Vol. 2). Boca Raton, FL, USA: Chapman & Hall/CRC.

[12] Gottscho, A. D. (2016). Zoogeography of the San Andreas Fault system : Great Pacific Fracture Zones correspond with spatially concordant phylogeographic boundaries in western North America. Biological Reviews, 91(1), 235-254.

[13] Hickey, B. M. (1979). The California Current system—hypotheses and facts. Progress in Oceanography, 8(4), 191-279.

[14] Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. Statistical science, 382-401.

[15] Jaynes, E. T. (2003). Probability theory: The logic of science. Cambridge university press.

[16] Love, M. S., Yoklavich, M., & Thorsteinson, L. K. (2002). The rockfishes of the northeast Pacific. Univ of California Press.

[17] Pearson, D. E., & Almany, G. (1995). The effectiveness of California's commercial rockfish port sampling program.

[18] Ramasubramanian, K., & Singh, A. (2016). Machine Learning Using R. Apress.

[19] Methot Jr, R. D., & Wetzel, C. R. (2013). Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142, 86-99.

[20] McCullagh P. & Nelder, J.A. (1989). Generalized Linear Models, 2nd ed. London: Chapman and Hall.

[21] NAS (National Academies of Sciences, Engineering, and Medicine). 2017. Review of the Marine Recreational Information Program. Washington, DC: The National Academies Press.

[22] PacFIN (Pacific Fisheries Information Network). 2018. https://pacfin.psmfc.org.

[23] Pearson, D. and B. Erwin. 1997. Documentation of California's Commercial Market Sampling Data Entry and Expansion Programs. NOAA Technical Memorandum NMFS, NOAA-TM-NMFS-SWFSC-240. 67 p.

[24] Pearson, D., B. Erwin, and M. Key. 2008. Reliability of California's Groundfish Landing Estimates from 1969-2006. NOAA Technical Memorandum NMFS, NOAA-TM-NMFS-SWFSC-431. 139 p.

[25] PFMC (Pacific Fisheries Management Council). 2018. Groundfish: Stock Assessments, STAR Reports, STAT Reports, Rebuilding Analyses, Terms of Reference. https://www.pcouncil.org/groundfish/stock-assessments.

[26] R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[27] Rao, J. N. K. (2003). Small area estimation. John Wiley & Sons, Ltd.

[28] Rue H., Martino S., Lindgren F., Simpson D., Riebler A. (2013). R-INLA: Approximate Bayesian Inference using Integrated Nested Laplace Approximations. Trondheim, Norway. URL http://www.r-inla.org/.

[29] Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the royal statistical society: Series b (statistical methodology), 71(2), 319-392.

[30] Saldaña-Ruiz, L.E., Sosa-Nishizaki, O. and Cartamil, D., 2017. Historical reconstruction of Gulf of California shark fishery landings and species composition, 1939–2014, in a data-poor fishery context. Fisheries Research, 195, pp.116-129.

[31] Sen, A.R. (1984). Sampling commercial rockfish landings in California. NOAA Tech Memo. NOAA-TM-NMFS-SWFSC-45.

[32] Sen AR. (1986). Methodological problems in sampling commercial rockfish landings. Fish Bull. 84: 409-421.

[33] Shelton, A. O., Dick, E. J., Pearson, D. E., Ralston, S., & Mangel, M. (2012). Estimating species composition and quantifying uncertainty in multispecies fisheries: hierarchical Bayesian models for stratified sampling protocols with missing data. Canadian Journal of Fisheries and Aquatic Sciences, 69(2), 231-246.

[34] Suter, J.M., 2010. An evaluation of the area stratification used for sampling tunas in the eastern Pacific Ocean and implications for estimating total annual catches. Inter-American Tropical Tuna Commission Special Report 18. https://www.iattc.org/SpecialReportsENG.htm

[35] Tomlinson, P. K. (1971). Some sampling problems in fishery work. Biometrics, 631-641.

[36] Tsou, T., L. Wargo, M. Langness, C. Jones, R. LeGoff, D. Downs, V. Okimura, B. Walker, and D. Bacon. 2015. Washington Coastal Commercial Groundfish Port Sampling, 2015 Report. Washington Department of Fish and Wildlife, Fish Program Report Number FPA 15-07. 53 p.