Model-Based Estimation of Average Fish Weights from Recreational Fisheries

E.J. Dick[a,*], Jason Edwards[b], Tien-Shui Tsou[c]

[a] Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 110 McAllister Way, Santa Cruz, CA 95060, USA. Telephone: (831) 420-3947. Email: Edward.Dick@noaa.gov

[b] Pacific States Marine Fisheries Commission, Recreational Fisheries Information Network, 205 SE Spokane St. Suite #100, Portland, OR 97202, USA. Telephone: (503) 595-3100. Email: jedwards@psmfc.org

[c] Washington Department of Fish and Wildlife, 1111 Washington Street SE, Olympia, WA 98501, USA. Telephone: (360) 902-2855. Email: Tien-Shui.Tsou@dfw.wa.gov

[*] Corresponding author

Abstract

Catch estimates from recreational fisheries are an important component of many fishery management plans. Estimates of recreational catch (in weight) on the U.S. West Coast are often derived as the product of catch in numbers of fish and average fish weights. When estimates of average fish weight are imprecise (e.g., due to small sample sizes), the resulting estimates of catch in weight can fluctuate and unnecessarily trigger or delay management actions. This and other challenges associated with average weight estimation are currently addressed through replication of data based on deterministic algorithms ('borrowing rules'). These methods differ among states and do not present a viable method for variance estimation. In this study, we describe a model-based framework for estimation of average fish weights, with an application to the recreational groundfish fishery off Washington, U.S.A. The model-based framework identifies important sources of variability in mean weight, quantifies uncertainty in estimates, pools information to better inform strata with small sample sizes, predicts average weight for unsampled strata, and does not require data replication. We examine the effect of sample size on model-based estimates, and recommend propagation of uncertainty in average catch into estimates of recreational catch in weight.

Keywords

Catch, estimation, recreational, average weight

1. Introduction

Catch monitoring is an important element of fishery management plans around the world (FAO 2020). When stocks are harvested by multiple fisheries, such as recreational and commercial sectors, catch estimates must be standardized across fisheries into common units (e.g., metric tons [mt]). In this way, total catch can be monitored against a target or limit (e.g., Annual Catch Limits, or ACLs, in the United States) (Methot et al. 2014). In waters off the U.S. West Coast (Washington, Oregon, and California), commercial landings of most species are recorded in units of weight. However, catch estimates for marine recreational fisheries are estimated in numbers of fish. To produce estimates of recreational catch in weight, the catch in numbers is multiplied by average fish weight in each stratum. Therefore, fish size data (weights and lengths) are a key component of catch monitoring efforts for these fisheries, but are also important for other aspects of fishery management in the region, including stock assessment.

Estimation of average fish weights from recreational catch is in principle a very straightforward procedure. In practice, however, several issues commonly arise. When estimates of average fish weight are highly variable (e.g., due to small sample sizes), resulting estimates of total recreational catch in weight can fluctuate and unnecessarily trigger (or delay) management action, such as fishery closures. Several mechanisms can contribute to small sample sizes. Sampling of mixed-stock fisheries may provide adequate coverage for primary target species, but only sparsely sample the less common species. Regulations also play a role. When retention of certain species is prohibited, this reduces the number of fish available to land-based samplers. Changing budgets and sampling priorities are a factor. Also, efforts to manage at fine spatial or temporal scales can reduce the amount of data available to generate stable estimates. In some fisheries, many of these factors occur simultaneously, compounding the problem. A recent review of sampling programs for U.S. recreational fisheries recommended that "small area" estimation (SAE) procedures be investigated to reduce variability observed in design-based estimators (NAS 2017). Rao and Molina (2015) describe a "small area" as any domain or grouping (not limited to geographic areas) in which the number of available samples is inadequate to provide estimates with the desired level of precision.

Another challenge is prediction of average weights for strata with no observations. This occurs due to a lack of sampling, or when a species is observed in the catch but a length or weight sample is not collected. In either case, some form of data imputation is needed (Rubin 1987, Gelman and Hill 2006). This process involves selection of an imputation model, as well as a method to characterize uncertainty. Estimation for strata with zero samples requires a model-based approach (Pfeffermann 2013), and can be considered a special case of SAE.

Currently, all three western states use 'borrowing rules' (or simply 'borrowing') to reduce variability in average weight estimates and/or to predict average weights for unobserved strata. Although details of this approach vary among states, it involves a deterministic algorithm that replicates ('borrows') data from observed strata when sample sizes do not meet a pre-specified threshold. These data are assumed to be representative of the sparsely-sampled or unsampled target stratum. Since data are replicated across strata, estimates of uncertainty are not clearly defined and are sometimes ignored (i.e., average weights are treated as a constant). Preferably, estimators of recreational catch in weight should account for uncertainty in both factors, catch in numbers and average weight (Goodman 1960). Estimates of uncertainty are especially important when data standards include precision thresholds, such as those recently established for recreational fishing surveys in the United States[1].

In this paper, we seek to address several of the above-mentioned issues using a model-based approach to estimate average fish weights. Specifically, we estimate average weights using hierarchical regression models (Gelman 2006). These models pool information across strata, 'borrowing strength' to reduce variability in estimates for sparsely-sampled strata, without the need for data replication. For each species, we evaluate covariates used in the borrowing rules (year and month), but also examine spatial (port) effects and vessel characteristics. Posterior predictive distributions for average weight are used for data imputation in unobserved strata. Our model-based approach produces estimates of average weight with uncertainty for both sampled and unsampled strata. These can be combined with estimates of total catch (in numbers) to produce estimates of total catch in weight that reflect uncertainty in both catch and average weight.

We apply our model-based method to marine recreational fishery data from Washington, USA. We compare model-based estimates of average weight to current estimates based on 'borrowing rules' for several groundfish species in the Washington recreational fishery. Species were chosen to include primary targets, as well as uncommon to rare components of the catch. Lastly, we use model selection to evaluate important sources of variability in average weight and to identify a model structure that has good predictive ability, a simple interpretation, and is easily implemented using freely available software.

2. Methods

2.1. Data and current estimation approach

Washington Department of Fish and Wildlife (WDFW) collects fish length and weight data from the coastal groundfish sport fishery in waters off the coast of Washington, USA. These efforts, referred to as "biological sampling," are independent of the sampling design for catch estimation. Biological sampling is conducted by two interrelated groups - the Ocean Sampling Program (OSP) and the Marine Fish Science (MFS) group (Davis and Wargo 2020). The OSP collects length data at all ports, and also collects weight data upon request. For the species examined in this study, OSP samples were available for the months of February through November, as fishing effort is greatly reduced in winter months. The OSP sample unit is a randomly selected boat trip, with four primary sites: Neah Bay, La Push, Westport, and Ilwaco. At the Westport site, the OSP samples fish landed by private boats and the MFS samples carcass lengths from charter boats. From March – September, the MFS also samples at Neah Bay and La Push to enhance data collections. The MFS does not sample sport landings at Ilwaco. Data from both sampling programs are combined to estimate average weight using WDFW's borrowing rules.

[1] NOAA Fisheries. (2020, December 4). *NOAA Fisheries Establishes Recreational Fishing Survey and Data Standards*. Retrieved from https://www.fisheries.noaa.gov/feature-story/noaa-fisheries-establishes-recreational-fishing-survey-and-data-standards

The WDFW borrowing algorithm uses measured fish weights, when available, or weights predicted from measured lengths using the best available weight-length relationship (see supplemental materials for additional details). To remove outliers, fish are excluded if they exceed a maximum length threshold for each species. The algorithm estimates average weights for each species as the arithmetic mean for any year/month stratum having at least 50 observations. If fewer than 50 weights are available in a stratum, data from previous month(s) are added, going back one month at a time, until the 50-fish minimum requirement is either met or exceeded. If fewer than 50 weights are available across all sampled months, an average weight is assigned from a lookup table in the database. Length and weight data are not available for released fish, and released fish are assumed to have the same average weight as retained fish in the same stratum. An analysis of differences in weight between retained and released fish would require a change to WDFW's current sampling design, and is beyond the scope of this analysis. See the supplementary materials for a more detailed description and flowchart of WDFW's borrowing rules for average weight estimation.

We examine data for eight species. These include Black Rockfish (*Sebastes melanops*) and Lingcod (*Ophiodon elongates*), both of which are primary targets of the recreational fishery. We also include species that are less commonly sampled, such as Canary Rockfish (*S. pinniger*) and Quillback Rockfish (*S. maliger*). These were chosen to represent 'data-poor' species for this time period, as they had smaller sample sizes and a greater number of strata with no observations. Other species in the study include China Rockfish (*S. nebulosus*), Copper Rockfish (*S. caurinus*), Cabezon (*Scorpaenichthys marmoratus*), and the Blue/Deacon Rockfish complex, a cryptic species pair (*Sebastes mystinus* and *S. diaconus*) that are not differentiated in the catch, but have generally similar life histories (Bizzarro et al. 2020). We refer to the Blue/Deacon species complex as "Blue Rockfish" for simplicity.

WDFW's biological samples and average weight estimates from the borrowing algorithm were obtained from the Recreational Fisheries Information Network (RecFIN; recfin.org). RecFIN is a repository of marine recreational fishing data for the states of Washington, Oregon, and California. In coordination with WDFW, RecFIN staff linked weight records and estimates of catch (in numbers) from the database to relevant covariates, created a custom view to simplify queries, and facilitated access to data in a format that could be directly imported for our models. All analyses in this study were conducted using the R programming language and environment (R Core Team, 2020).

2.2. Model-based approach

We estimate average fish weight using generalized linear models (GLM) and generalized linear mixed models (GLMM; Gelman 2006). These models offer a well-established framework for identifying important sources of variability (model selection) and evaluating model performance. They also allow us to estimate mean weight for both sampled and unsampled strata. We fit models to the same data used to implement WDFW's borrowing rules. However, we use only the observed data in each stratum, and do not replicate data from earlier months or require a minimum number of samples per stratum. Similar to the borrowing algorithm, we include weights predicted from lengths when direct weight measurements are not available. Hierarchical (mixed) models were fit using the 'rstanarm' and 'lme4' packages and GLMs were fit using the 'glm' function in R (Goodrich et al. 2020, Bates et al. 2015, R Core Team 2020).

We assume that fish weights (the response variables) are gamma distributed since the data are continuous, strictly positive, and have a variance that increases with the mean. We relate mean weight to the linear predictor via a log link function, and evaluate candidate models independently for each species (i.e., species is not a covariate in the linear predictor).

Covariates considered during model selection included year, month, port of landing, and boat mode (private and charter). Port of landing and boat mode are not part of WDFW's borrowing algorithm, but are strata used for catch estimation (in numbers). Therefore, we included them in the set of candidate models. Port of landing could affect average weight if fish near a particular port experienced differences in growth rate and/or exploitation history. Data from small, secondary ports were combined with the primary sites for the model. Specifically, samples from Chinook were combined with Ilwaco, and samples from Ocean Shores (OS) were combined with Westport. Boat modes could also have an effect on the average weight of landed fish. Charter boats, for example, may have access to more distant, less exploited reefs with larger fish. WDFW's catch estimates are also stratified by catch area and trip type (WDFW 2017). These covariates were not included in the candidate set of models because they were not consistently available for weight observations due to differences in the sampling designs for biological data and catch estimation. All covariates were coded as categorical variables ('factors' in R).

As noted earlier, challenges with estimation of average weight include highly variable estimates due to small sample sizes, and prediction for unobserved strata. In addition to models with only 'fixed' effects (single-level models), we evaluate models that estimate group-level parameters using information within the group and across groups (i.e., hierarchical models, 'partial pooling,' or 'random' effects). This is commonly referred to as borrowing strength or shrinkage (Gelman and Hill 2006). In our hierarchical models, the average weight estimate for a group is 'shrunk' toward the mean across groups. The amount of shrinkage depends on the amount of information within the group itself, and predictions for unobserved groups reflect uncertainty among groups.

All linear predictors were coded using the default 'treatment' contrasts for the design matrices. Diffuse priors were selected for regression coefficients in the Bayesian models using the 'rstanarm' package in R. Specifically, normal priors with a zero mean and standard deviation (SD) of 3 kg for the intercept, and normal priors with a zero mean and SD=1 kg for the offsets. An exponential prior with rate parameter equal to 0.25 (mean = 4) was used for the shape parameter of the gamma distribution. Further details about prior specification in 'rstanarm' are described by Goodrich et al. (2020).

We initially fit models for each species using maximum likelihood (for GLMs) or restricted maximum likelihood for GLMMs (using the 'glmer' function in 'lme4'). This approach reduces the time needed to compare models as Bayesian estimation can be slow for larger data sets. It also provides an opportunity to compare parameter estimates between models to confirm that our choice of prior distributions do not significantly influence the results. We use Akaike's Information Criterion (AIC) to narrow the set of candidate models based on the (restricted) maximum likelihood fits, then fit Bayesian hierarchical models using ‘rstanarm’ for final model diagnostics and inference.

Convergence diagnostics for optimization algorithms included monitoring of convergence, defined as gradients less than 0.005. Evidence of convergence for Bayesian models was assessed using Gelman and Rubin’s (1992) diagnostic (“$\hat{R}$”) as reported by ‘rstanarm’ output, for each parameter. Models with $\hat{R}$ values less than 1.1 for all parameters were accepted. Bayesian models were also monitored for divergent transitions in the Hamiltonian Monte Carlo (HMC) algorithm and maximum tree depth of the No-U-Turn-Sampler (Monnahan et al. 2017).

Once a model is selected and passes the above-mentioned convergence diagnostics, we evaluate model performance. For this, we rely on posterior predictive checks (Gabry et al. 2019), whereby the fitted model is used to simulate 10,000 replicate data sets that are then compared to the observed data set. We use a combination of graphical and quantitative checks. Examples of graphical checks include histograms comparing observations or group-level means to predictive distributions. Quantitative checks include measures of predictive coverage, comparing quantiles of the observed weights to quantiles of the posterior predictive distributions.

After selecting a model for average weight estimation, we compare estimates of mean weight derived from the current borrowing algorithm (i.e., with data replication) to estimates from the model-based approach. For each species, we plot group-level posterior means (i.e., every observed combination of year/month/port/mode) from the model, and point estimates from the borrowing algorithm, against arithmetic means of the observed weights.

Finally, we illustrate how uncertainty in estimates of average weight relate to sample size. For each species, we plot percent standard error of the posterior mean as a function of sample size. This analysis can be used to inform target sampling rates and to help optimize allocation of sampling effort.

3. Results

The eight species we examined differed in average weight (grand means), ranging from less than 1 kg (blue and China rockfishes) to over 3 kg for lingcod (Table 1). Two primary targets of the fishery (black rockfish and lingcod) had the largest total sample sizes (N=19,359 and N=10,569, respectively). Canary rockfish had the smallest sample size (N=3,495), primarily due to regulatory limits over the observed time period. Sample sizes for other species were in the range of 4,000 – 5,000 weights. The median number of samples per observed stratum ranged from 10 for quillback rockfish to 49 for black rockfish (Table 1).

Samples were collected over 8 years (2010-2017) and across 4 port groups, 2 boat modes, and 7-10 months depending on the species. Monthly coverage varied across years within each species. During the primary months of March through October, the proportion of *potential* strata (512 possible combinations of year/month/port/mode) having at least one weight sample varied from 13% coverage for canary rockfish to 61% coverage for black rockfish. In practice, however, catch does not occur in all combinations of year, month, port, and mode. So, while the proportion of *landed* strata with weight samples would be larger, the model is still able to generate predictions for all strata (observed or unobserved).

Table 1. Summary of WDFW fish weight data by species, 2010-2017. Grand mean weight is the arithmetic mean across all samples. The proportion of strata sampled is based on 512 year/month/port/mode combinations having at least one sample, excluding winter months (November-February).

Species	Average Weight (grand mean, kg)	Total Sample Size (N)	Median N per sampled stratum	Proportion of strata sampled
Black Rockfish	1.19	19,359	49	0.61
Blue Rockfish	0.96	4,165	16	0.25
Cabezon	2.81	4,186	15	0.40
Canary Rockfish	1.01	3,495	26	0.13
China Rockfish	0.94	4,547	12	0.31
Copper Rockfish	1.61	3,900	12	0.25
Lingcod	3.60	10,569	24	0.60
Quillback Rockfish	1.45	4,467	10	0.31

For this study, we evaluated 14 candidate models for each species (Table 2). We began with GLMs (models 1-8), but due to the sparse nature of the data, these were limited in terms of estimable fixed-effect interaction terms. We also considered a set of GLMMs that all included a 4-way interaction term as a random effect (models 9-14; Table 2). In the GLMMs, all combinations of year/month/port/mode are estimated as deviations sharing a common distribution. This simple random effects structure was chosen as it is highly flexible, easy to interpret, allows for 'partial pooling' of information across strata, and facilitates prediction for unobserved strata.

Based on Delta-AIC (AIC-min(AIC); calculated for each species separately), we found that all the GLMMs showed significant improvements in fit relative to the GLMs (Table 2). This is not surprising given the increase in model complexity associated with the 4-way interaction, and the fact that the penalty term in AIC only counts the random effect as 1 extra parameter (the hierarchical variance parameter). However, the purpose of using AIC in this study was to narrow the set of candidate models prior to evaluating model performance. It seems unlikely that a single model structure could be selected as the 'best' model for all species following any criterion, given differences in biology and fishery characteristics over time and among ports. However, one model may provide adequate performance across a wide range of species.

The GLMMs in the candidate set differ in terms of their fixed effect structures (models 9-14, Table 2). The simplest GLMM (model 9) estimates a single intercept parameter (population mean) as a fixed effect and performed relatively well for many species. Model 10 estimates independent means for each year, along with stratum-specific deviations for each year/month/port/mode that are drawn from a common distribution. Models 11 and 12 are similar, with port group and boat mode replacing the year-specific estimates, respectively. Models 13 and 14 assume additive fixed effects for port/mode and year/port/mode, again with stratum-level deviations estimated via the random effect term.

Table 2. Comparison of GLMs (models 1-8) and GLMMs (models 9-14) based on Delta-AIC (AIC – min(AIC)), by species. Degrees of freedom = d.f. Codes for categorical covariates are yr = year, pt = port group, md = boat mode, and mo = month. Models (GLMMs) with a random intercept (a 4-way interaction term) are indicated by "(1 | yr:mo:pt:md)," following notation used in the 'lme4' and 'rstanarm' packages for R. * indicates failed convergence of the optimizer in glmer.

			Species								
#	Linear Predictor	d.f.	Black Rockfish	Blue Rockfish	Cabezon	Canary Rockfish	China Rockfish	Copper Rockfish	Lingcod	Quillback Rockfish	Average Δ-AIC
1	yr	9	2384	1548	1454	893	2175	633	2106	1431	1578
2	pt	5	2215	1622	1049	1208	658	693	2010	1127	1323
3	md	3	2345	1792	1629	1286	2123	730	2245	1159	1664
4	yr + pt	12	2152	1409	863	877	610	589	1903	1026	1178
5	yr + md	10	2311	1514	1452	893	2001	624	2108	1066	1496
6	pt + md	6	2163	1612	1038	1202	658	692	2012	1086	1308
7	yr + pt + md	13	2118	1389	832	864	610	564	1904	991	1159
8	yr + pt + md + pt:md	16	2001	1389	829	837	493	490	1860	912	1101
9	Intercept + (1 | yr:mo:pt:md)	3	5	0*	24	0	35	5	14*	22	13
10	yr + (1 | yr:mo:pt:md)	10	5	2	29	11*	43	7	16	28	18
11	pt + (1 | yr:mo:pt:md)	6	1	1	1*	3	0	7	0	1	2
12	md + (1 | yr:mo:pt:md)	4	4	2*	25	1	32	5	13	16	12
13	pt + md + (1 | yr:mo:pt:md)	7	1	3	0	5	1	1	0	0	1
14	yr + pt + md + (1 | yr:mo:pt:md)	14*	0*	7*	1*	17*	4*	0*	1*	3*	4

Although the simplest GLMM (#9) was the 'best' among candidate models for canary rockfish, the optimization in 'glmer' failed to converge (gradient > 0.005) when the same model was fit to data for blue rockfish and lingcod (Table 2). In fact, of the 48 models fit using 'glmer,' 13 did not converge. Eight of these cases were for one model (#14), which didn't converge for any species and was not considered further. The remaining 5 cases that did not converge did not show a consistent pattern (Table 2). Convergence issues with 'glmer' are often not an issue when using a Bayesian framework (see discussion for details). Since we used Bayesian models for final inference, and the converged GLMMs consistently outperformed the GLMs, we moved forward with performance testing to determine whether the 'best' model in the set was a good model.

One GLMM (model 13, Table 2) had the lowest average AIC, and also converged for all species. This model specifies additive fixed effects for port group and boat mode, deviations and variance of the random effect term, and a shape (dispersion) parameter for the assumed gamma distribution. We selected model 13 for final inference, performance evaluation, and comparison to the existing borrowing algorithm.

Bayesian fits with model 13 did not show any evidence of lack of convergence, with $\hat{R} < 1.1$ for all parameters, no divergent transitions, and maximum tree depths <15 (the default in 'rstanarm'), for all eight species in our study. Posterior predictive checks suggest that the models were able to adequately reproduce patterns in the observed data (Figures 1-6 illustrate results for black rockfish; see supplementary materials for results from other species). Specifically, marginal means by year, month, port, and mode were calculated for each of 10,000 simulated data sets from the model, and compared to marginal means from the observed data set.

Marginal means by year (Figure 1) were consistent with observed means, even though year was not included as a fixed effect in the model. Since many months may have few or no samples, we also model month as part of the 4-way interaction term. However, monthly marginal means of the posterior predictive distribution show that the model is capable of capturing seasonal variability (Figure 2). The model also predicts greater uncertainty for months with fewer samples (e.g., February and November) compared to peak months during the summer (Figure 2). This demonstrates that deviations in the random effect are able to capture annual and seasonal changes in mean weight, and reproduce the expected relative changes in uncertainty.

Unlike the borrowing algorithm, model 13 accounts for differences in mean weight associated with port group (spatial effects) and boat mode. Black rockfish caught in the Ilwaco/Chinook port group had a mean weight of roughly 1.28 kg, while average weight in Neah Bay was 1.13 kg (Figure 3). Charter boats, on average, caught larger black rockfish than private boats, and this difference was also captured by the model (Figure 4).

Model performance is often best visualized by comparing observed and predicted values. Figure 5 reveals that, relative to the borrowing algorithm, predictions from the model-based approach are less concentrated around the population mean (roughly 1.2 kg, Table 1) and more closely match the observed weights. This is especially true for observed values below 1 kg and above 1.4 kg. This improvement in fit to the observed values when using the model-based approach also holds for the other species we examined (see Supplemental Materials).

Another useful metric of model performance is a comparison of observed data distributions to simulated data sets from the model. If simulated data are consistent with the observed data, then 95% of observations should fall within the 95% central interval, and likewise 50% of observations should fall within a 50% central interval. Using model 13, we found that the 95% highest density intervals for the posterior predictive distributions contained almost exactly 95% of the observations (Table 3). Across species, the 50% highest density interval contained a slightly larger fraction of the data (51-60%; Table 3). This suggests that the data distributions are slightly more concentrated around their central tendencies than the predictive densities from the model.

Table 3. Fraction of observations, by species, that fall within the 50 and 95 percent highest density intervals from the posterior predictive distribution given model 13. PPDs are based on 10,000 simulated data sets.

Species	50%	95%
Black Rockfish	0.533	0.951
Blue Rockfish	0.576	0.947
Cabezon	0.550	0.948
Canary Rockfish	0.599	0.954
China Rockfish	0.556	0.954
Copper Rockfish	0.509	0.961
Lingcod	0.585	0.942
Quillback Rockfish	0.592	0.949

As mentioned earlier, one of the motivations for using borrowing algorithms is to reduce variability in mean weights that results from small sample sizes. Our model-based approach achieves this through ‘partial pooling’ of information among the coefficients of the random effects term. In this framework, stratum-level estimates are ‘shrunken’ toward the population mean by an amount that reflects the amount of information in the available data (Figure 6, black open circles). Estimates from well-informed strata will change very little, while estimates from poorly-informed strata will shrink towards the population mean. The amount of shrinkage depends on the variance of the data as well as the distance from the population mean. For a primary target species like black rockfish, many strata have large sample sizes. As a result, posterior means from the model are largely consistent with the observed, stratum-level means (Figure 6, black open circles). However, when estimated means are very large (e.g., >1.5 kg) or very small (e.g., <0.75 kg), the shrinkage effect becomes much more noticeable in the model-based estimates.

Estimates from the borrowing algorithm are less variable than the stratum-level means (Figure 6, grey solid circles). However, unlike the model-based approach, estimates of mean weight from the borrowing algorithm do not reflect information in the available data. The relationship between the data in a stratum and the estimates from the algorithm is very weak. In fact, some estimates from the borrowing algorithm appear to be ‘shrinking’ in the opposite direction one would expect (away from the population mean). Since model 13 includes fixed effects for port group and boat mode, the random effects are actually estimated as deviations from the 8 possible combinations of port & mode. In that sense, the dashed lines in Figure 5 do not exactly match the population means in the model (observant readers will notice that a single estimate from the model appears to 'shrink' in the wrong direction). However, we show only one horizontal and one vertical reference line representing the population grand mean for the sake of clarity. Shrinkage plots for other species, similar to Figure 5, are available with the supplemental materials.

Given that data are replicated through time, it is not clear how to calculate variance estimates for the current WDFW borrowing rules. As noted earlier, uncertainty in average weight is needed in order to properly estimate uncertainty for total catch in weight. We have shown that estimates of uncertainty from our model-based approach reflect uncertainty in the data for these eight species (Table 3). Another benefit of quantifying uncertainty in mean weight is understanding how data collection affects precision of the estimates. Percent standard errors of the model-based average weight estimates decrease with increasing sample sizes, as one would expect (Figure 7). Sample sizes of 50 or more per stratum result in PSEs less than 7% for all species, and sample sizes of 100 or greater consistently produce PSEs less than 5%.

Small sample sizes (e.g., <25 fish), however, can have PSEs between 10-20%, which may influence estimates of uncertainty for total catch in weight. Predictions for unsampled strata will have even larger PSEs, but in practice most catch estimates used for management are an aggregate across multiple strata (reducing PSEs for the aggregate catch).

4. Discussion

Information about average fish weight from recreational fisheries is needed to meet the requirement for total coast-wide harvest and mortality in federally-managed U.S. fish stocks. Catch estimates from all three West Coast states (Washington, Oregon, and California) are produced in numbers of fish, along with estimates of uncertainty. We propose a model-based approach for estimating mean weight conditional on several other variables (year, month, port, and mode). The same model is used to generate estimates for observed strata and predictive distributions for unsampled strata, both with uncertainty. We show that the model-based estimates are better able to reproduce the observed data than the current borrowing algorithm, which ignores uncertainty in average weight. The model-based estimates can be combined with existing estimates of catch in numbers to produce estimates of total catch in weight that reflect uncertainty in both average weight and catch in numbers (Goodman 1960). This propagation of uncertainty is not possible using the deterministic borrowing algorithm.

Although we illustrate our method with data from Washington State, this study provides a general framework for average weight estimation that warrants consideration by other agencies that employ borrowing rules. Among U.S. West Coast recreational fisheries, deterministic algorithms differ by state, and none of the algorithms quantify uncertainty. This study describes a consistent framework that does not require data replication, identifies important sources of variability in mean weight, quantifies uncertainty, pools information to better inform strata with small sample sizes, imputes average weight for unsampled strata, and helps inform survey design. The model-based approach also appears to work well across a range of species, from 'data-rich' to 'data-poor.'

Variability in mean weight estimates, which can arise due to small sample sizes, is addressed in our method through partial pooling of information (Gelman and Hill 2006). The resulting ‘shrinkage’ of imprecise estimates toward population means reduces the influence of erratic average weights on estimates of catch in biomass. The message in Figure 5 is two-fold. First, the model-based estimates of mean weight are more consistent with observed means, compared to the borrowing algorithm. Second, and equally important, is the fact that the model only shrinks the estimates as far as the population mean, unlike the borrowing algorithm which appears to 'shrink' estimates in the wrong direction. Another advantage of partial pooling is that variance of the random effects can be propagated in predictions for unsampled strata.

SAE using hierarchical (generalized linear mixed-effects) models has a well-established theoretical basis (Ghosh and Rao 1994; Schaible 1996; Pfeffermann 2013; Rao and Molina 2015), but to our knowledge has not been proposed for estimation of average weights for recreational fisheries. As noted by Schaible (1996), small area (“indirect”) estimators are often only considered after a sampling design is developed and implemented. Future research could examine how the use of SAE methods, such as the one we propose, interact with design-based, direct estimation methods and potentially influence the design of fishery surveys in general.

Advantages of the model-based approach also include a well-established framework for model selection. We use AIC to illustrate this approach, but other information criteria such as the Widely-Applicable Information Criterion (WAIC) and Leave-One-Out Information Criterion (LOOIC) are reasonable alternatives (Vehtari et al. 2017). These methods have the advantage of not requiring a strict definition of model dimension for the penalty term. Although AIC treats the random effect term in our models as a single parameter, the use of posterior predictive checks provides a means to ensure that the ‘best’ model in the candidate set is also a good model. While we chose a single model that performs well across a wide range of species, another option would be to tailor models to specific data sets depending on the unique characteristics of a species and/or fishery. For example, canary rockfish was declared overfished in 2000 and rebuilt in 2016 (PFMC 2018). Due to limited landings during that time, a simple model for average weight may be warranted, while a more complex model may be supported in subsequent years with larger sample sizes.

We encountered issues with convergence for some GLMMs that were fit using the glmer package in R. These models were not used for final inference, but the speed of the optimization algorithm in glmer is useful for rapid evaluation of multiple models during the model selection process. With the exception of model 14 (Table 2), failed convergence was uncommon and could potentially be resolved through modification of settings in the optimization algorithm. Advice on resolving these issues is beyond the scope of this paper, but often includes the suggestion to implement a Bayesian model. We did not encounter convergence issues for these data sets when using rstanarm. This is due in part to specification of “weakly informative” priors that have negligible impacts on results but prevent estimation algorithms from ‘getting lost’ in unreasonable regions of the parameter space (Lemoine 2019).

Alternatives to the model-based approach include modification of survey designs to estimate catch in biomass as well as catch in numbers. Essentially, estimation of catch rates in biomass would eliminate the need for separate estimates of catch in numbers and average weight.

Modification to existing survey designs can be expensive, however, and benefits may not warrant the costs. This approach would also only affect future estimates, and average weight estimates would still be required to convert historical estimates of catch from numbers to weight. Hot and cold deck imputation methods are another option for dealing with missing data (Andridge and Little 2017). The model-based approach offers a single solution to both estimation and imputation, with associated variance estimates, but evaluation of performance and careful examination for evidence of model misspecification is key. We find that posterior predictive checks are a powerful tool for evaluating model performance and diagnosing model misspecification (Gelman and Hill 2006).

Our analysis of percent standard errors as a function of sample size can be used to inform allocation of sampling effort across strata. If PSEs below 5% are adequate for catch monitoring, then sampling effort beyond 100 fish per stratum could be re-allocated to address other priorities. Model-based estimates of uncertainty may be less than corresponding design-based estimates due to the partial pooling of information, but this bias-variance tradeoff is by design. Since the borrowing algorithm in this study does not provide an estimate of variance, we find the model-based approach to be a significant improvement, both in terms of point estimates and variance estimation.

Small changes in mean weight can have important implications for catch monitoring and in-season management of primary target species (e.g., black rockfish; H. Hall, WDFW, pers. comm.). Large fluctuations in mean weight due to small sample sizes can produce similar difficulties when they occur for infrequently landed, but constraining species within a mixed-stock fishery (i.e., overfished stocks). The use of borrowing algorithms for average weight estimation is not unique to the U.S. west coast. Similar algorithms have been adopted in the southeast United States (Matter and Turner, 2010; Matter and Rios, 2013). These methods often follow a sequence of increasing aggregation to meet minimum sample sizes, but lack detailed analysis of whether alternative algorithms would better reproduce the observed average weights. Similar to the WDFW example, these algorithms do not provide variance estimates (Southeast Data Assessment and Review, 2016). In these respects, our proposed framework provides a way forward.

Acknowledgements

The authors thank Dr. Michael O'Farrell (NOAA Fisheries) and Dr. John Field (NOAA Fisheries), as well as two anonymous reviewers, for comments and suggestions that have been very helpful in improving the manuscript.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Literature Cited

Andridge, R. and R. Little. 2010. A Review of Hot Deck Imputation for Survey Non-Response. International Statistical Review 78: 40–64.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Bizzarro, J., E. Gilbert-Horvath, E. Dick, A. Berger, K. Schmidt, D. Pearson, C. Petersen, L. Kautzi, R. Miller, J. Field, and J. Garza. 2020. Genetic identification of blue rockfish (*Sebastes mystinus*) and deacon rockfish (*S. diaconus*) to enable life history analyses for stock assessment. Fishery Bulletin 118(1): 37-50.

Davis, R. and L. Wargo. 2020. Enhanced Sampling of the Washington Coast Recreational Groundfish Fishery 2014-2017. Washington Department of Fish and Wildlife. Technical Report No. FPT 20-05.

FAO. 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in action. Rome. https://doi.org/10.4060/ca9229en

Gabry J., Simpson D., Vehtari A., Betancourt M., and Gelman A. 2019. Visualization in Bayesian workflow, J. R. Statist. Soc. A 182, Part 2, 389-402.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7: 457–472.

Gelman, A. and J. Hill. 2006. Missing-data imputation. In Data Analysis Using Regression and Multilevel/Hierarchical Models (Analytical Methods for Social Research, pp. 529-544). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511790942.031

Ghosh, M. and J. Rao. 1994. Small area estimation: an appraisal. Statistical Science 9(1): 55-93.

Goodman, L. 1960. On the Exact Variance of Products. Journal of the American Statistical Association, 55: 708.

Goodrich, B., J. Gabry, I. Ali, and S. Brilleman. 2020. rstanarm: Bayesian applied regression modeling via Stan. R package version 2.21.1 https://mc-stan.org/rstanarm

Lemoine, N. 2019. Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. Oikos 128: 912-928.

Matter, V. and S. Turner. 2010. SEDAR22-DW16. Estimated Recreational Catch in Weight: Method for Filling in Missing Weight Estimates from the Recreational Surveys with Application to Yellowedge Grouper, Tilefish (golden), and Blueline Tilefish. National Marine Fisheries Service, Southeast Fisheries Science Center, Sustainable Fisheries Division. Available online at http://sedarweb.org/docs/wpapers/S22_DW_16_Est_Rec_Catch_in_Weight.pdf.

Matter, V.M. and A. Rios. 2013. MRFSS to MRIP Adjustment Ratios and Weight Estimation Procedures for South Atlantic and Gulf of Mexico Managed Species. SEDAR32-DW02. SEDAR, North Charleston, SC. 6 pp. Available online at http://sedarweb.org/docs/supp/PW7-20_SEDAR32_DW02_Matter%26Rios_2.5.2013_FINAL.pdf.

Methot, R., G. Tromble, D. Lambert, and K. Greene. 2014. Implementing a science-based system for preventing overfishing and guiding sustainable fisheries in the United States. ICES Journal of Marine Science 71(2): 183-194.

Monnahan, C., J. Thorson, and T. Branch. 2017. Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. Methods in Ecology and Evolution 8: 339–348.

NAS. 2017. National Academies of Sciences, Engineering, and Medicine: Review of the Marine Recreational Information Program. Washington, DC: The National Academies Press. https://doi.org/10.17226/24640

Pfeffermann, D. 2013. New important developments in small area estimation. Statistical Science 28(1): 40-68.

PFMC. 2018. Pacific Coast Groundfish Fishery Management Plan for the California, Oregon and Washington Groundfish Fishery: Appendix F, Overfished Species Rebuilding Plans. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, OR 97220. www.pcouncil.org.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rao, J. and I. Molina. 2015. Small Area Estimation, Second Edition. Wiley Series in Survey Methodology. John Wiley & Sons, Inc.

Rubin, D. 1987. Multiple Imputation for Nonresponse in Surveys, New York: John Wiley & Sons, Inc.

Schaible, W. 1996. Indirect Estimators in U.S. Federal Programs. Lecture Notes in Statistics 108. Springer.

Southeast Data Assessment and Review (SEDAR). 2016. Stock Assessment Report: Gulf of Mexico Data-limited Species. SEDAR, North Charleston, SC. 618 pp. Available online at https://www.cio.noaa.gov/services_programs/prplans/pdfs/ID361_SEDAR49_GOMDataLimited_FinalProduct.pdf

Vehtari, A., A. Gelman, and J. Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing, 27, 1413–1432. doi: 10.1007/s11222-016-9696-4.

WDFW. 2017. Washington State Department of Fish and Wildlife, Ocean Sampling Program Overview. Available online: https://www.recfin.org/wp-content/uploads/2017/11/Washington-Ocean-Sampling-Program-OSP-Overview-for-RECFIN-UPDATEDNOV-2....pdf

Figure captions for Dick et al.

Figure 1. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by year (histograms) for black rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

Figure 2. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by month (histograms; February – November) for black rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

Figure 3. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by port group (histograms) for black rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

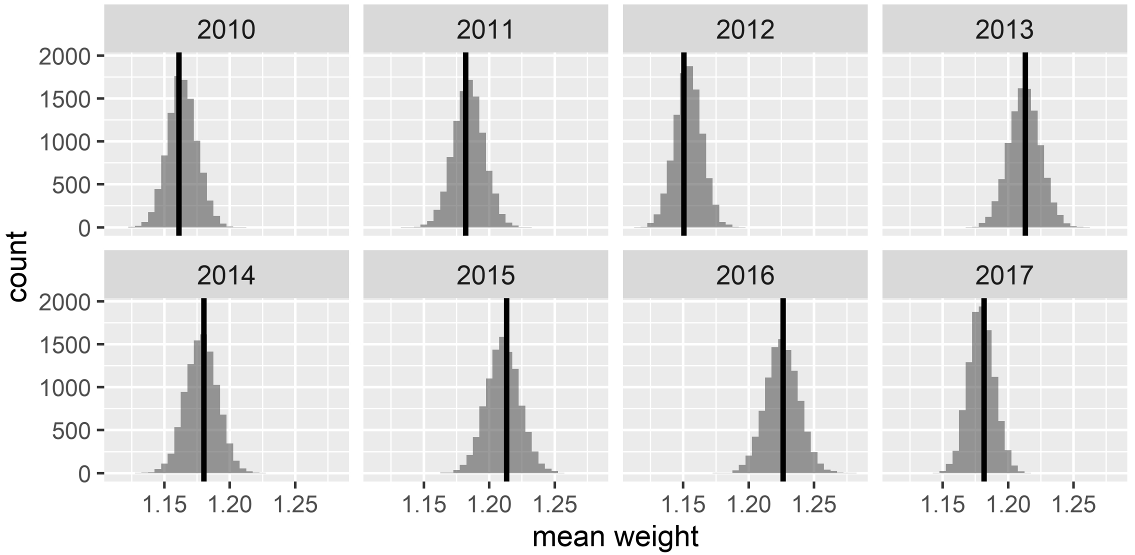
Figure 4. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by boat mode (histograms) for black rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

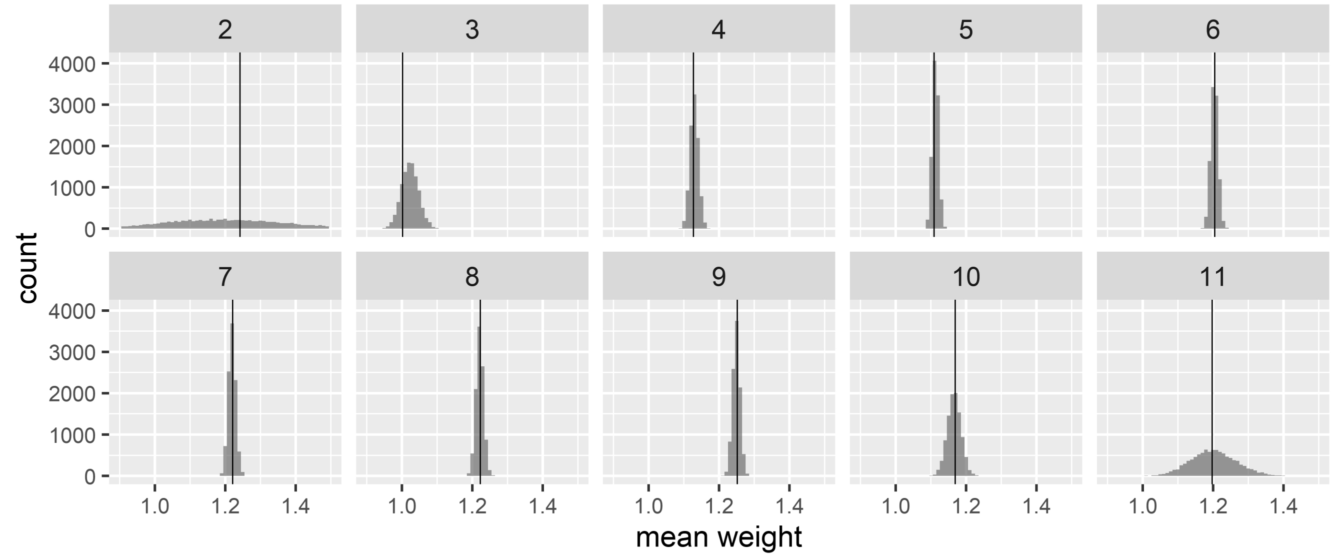
Figure 5. Observed weights [kg] of black rockfish (n=19,359) versus the mean of the posterior predictive distributions (upper panel) and the mean from the borrowing algorithm (lower panel). Solid line is the 1:1 line for reference.

Figure 6. Estimated posterior means [kg] from model 13 (black open circles) and estimates from the borrowing algorithm (grey solid circles) versus observed arithmetic means for black rockfish. Each point represents a single year/month/port/mode combination (316 observed strata). Horizontal and vertical dashed lines are the arithmetic mean weight across all observations (grand mean). Solid line is the line of equality (1:1).

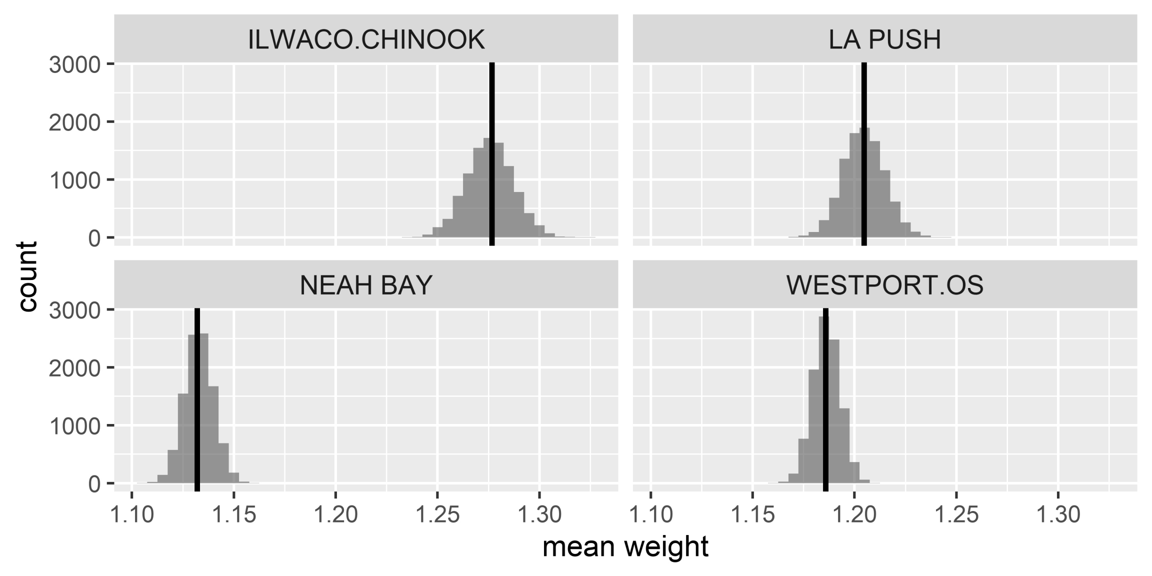
[Figure 7 in color, online only]

Figure 7. Percent standard error (PSE) of the posterior mean by stratum (year/month/port/mode), as a function of sample size and species.

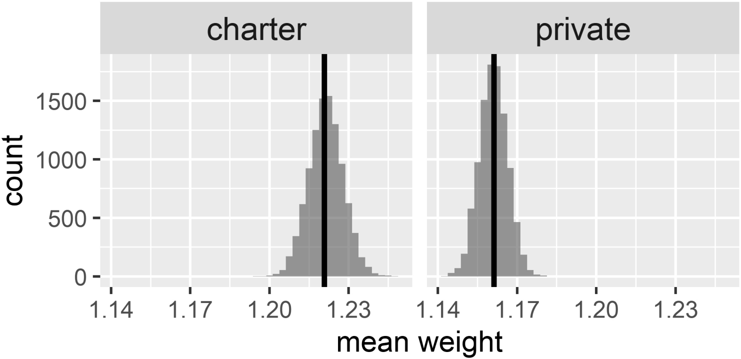

2010
2011
2012
2013
2014
2015
2016
2017
2000
1500
1000
500
0
count
1.15
1.20
1.25
mean weight

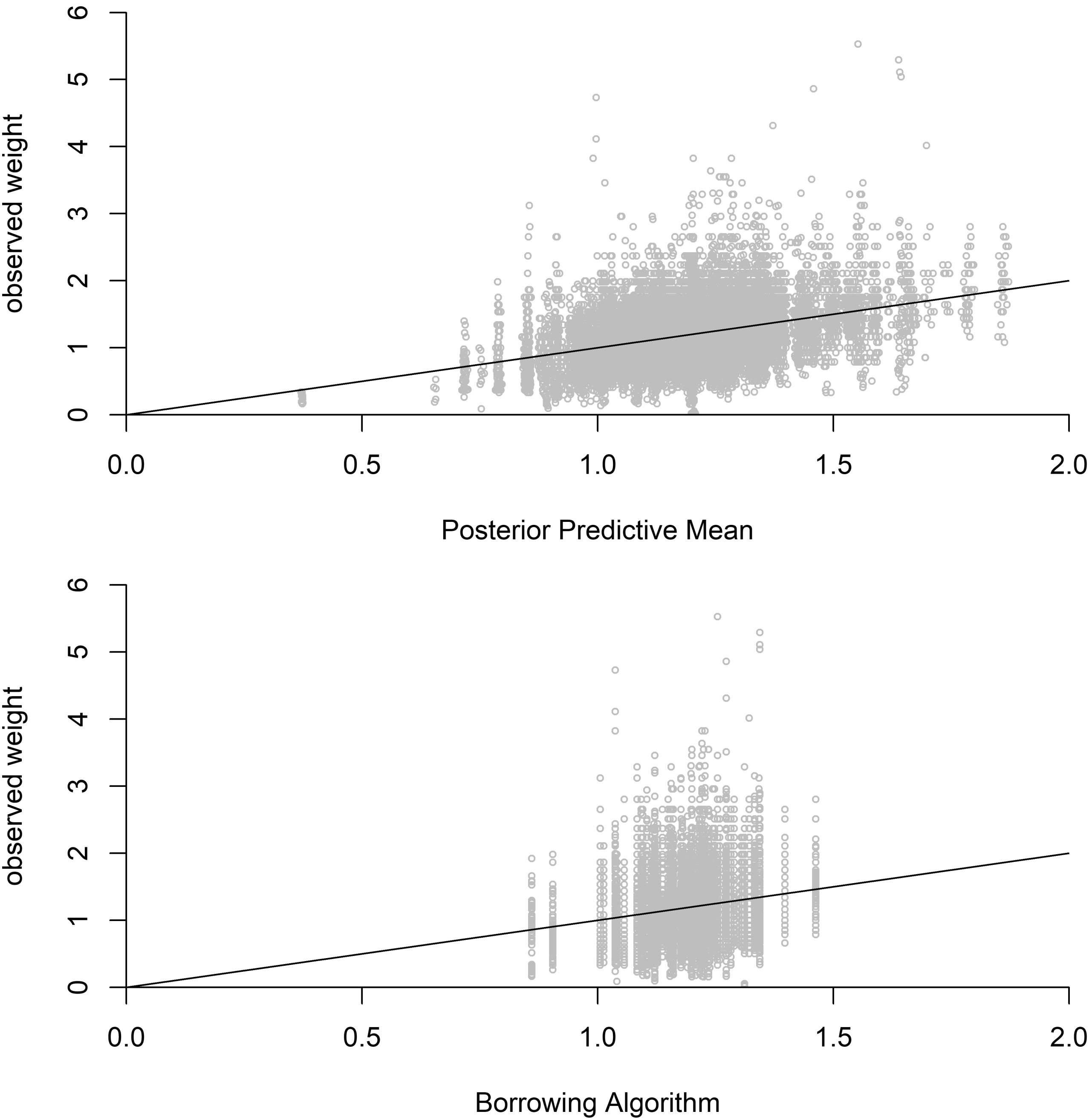


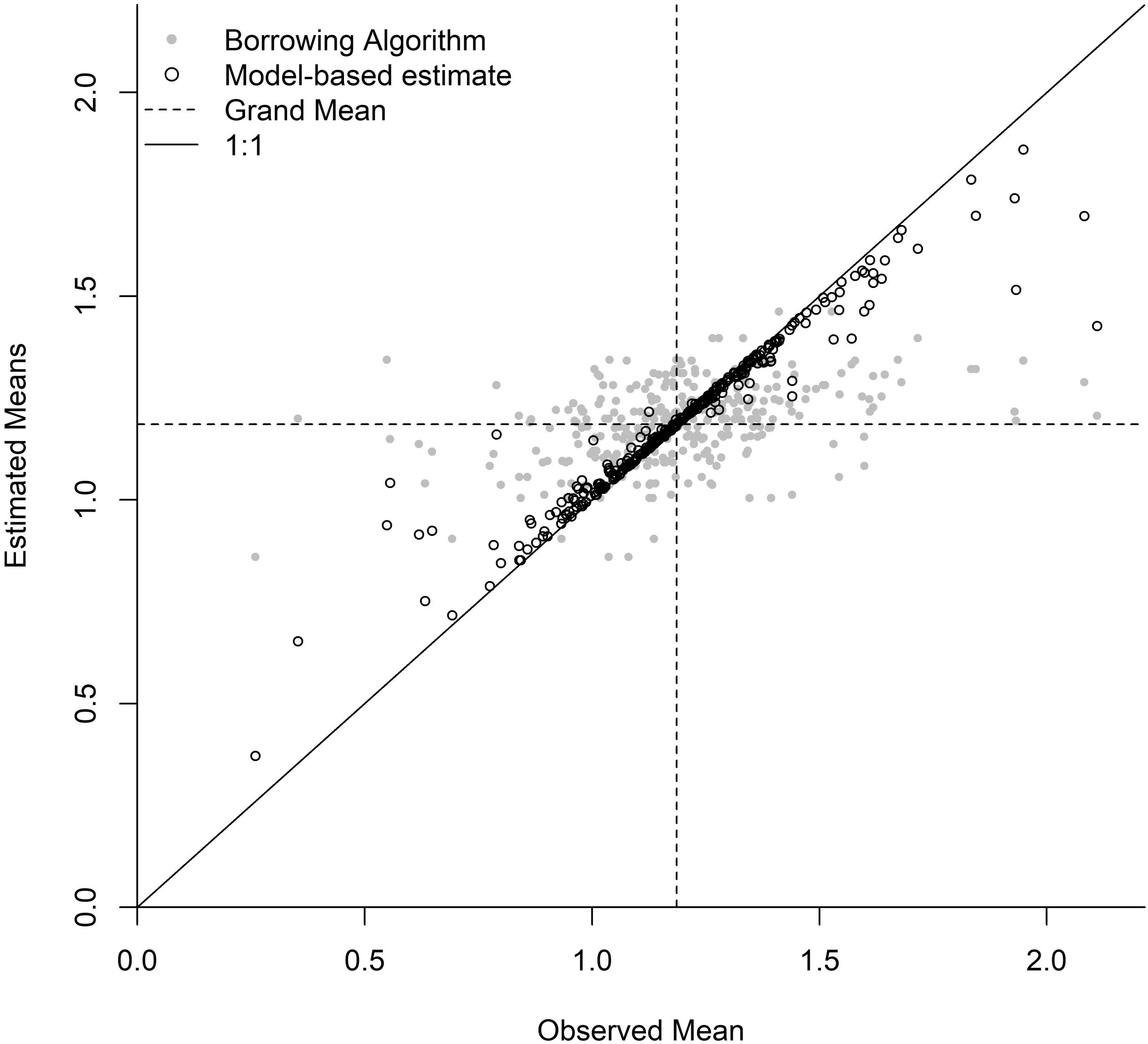
2
3
4
5
6
7
8
9
10
11
count
4000
3000
2000
1000
0
1.0
1.2
1.4
mean weight

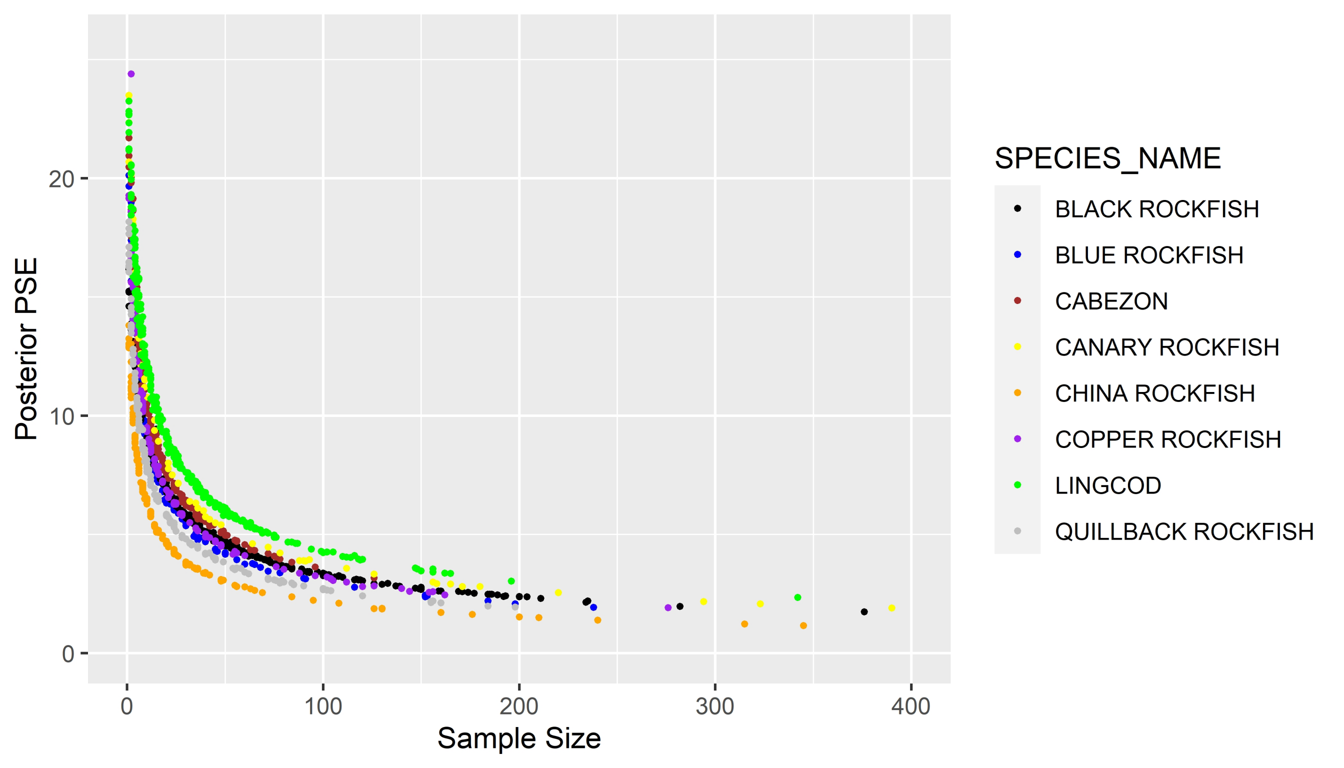


ILWACO.CHINOOK
LA PUSH
NEAH BAY
WESTPORT.OS
count
0
1000
2000
3000
1.10
1.15
1.20
1.25
1.30
mean weight


charter
private
1500
1000
500
0
count
1.14
1.17
1.20
1.23
mean weight


observed weight
0
1
2
3
4
5
6
0.0
0.5
1.0
1.5
2.0
Posterior Predictive Mean
observed weight
0
1
2
3
4
5
6
0.0
0.5
1.0
1.5
2.0
Borrowing Algorithm


Borrowing Algorithm
Model-based estimate
Grand Mean
1:1
Estimated Means
Observed Mean
0.0
0.5
1.0
1.5
2.0


Posterior PSE
0
10
20
Sample Size
0
100
200
300
400
SPECIES_NAME
BLACK ROCKFISH
BLUE ROCKFISH
CABEZON
CANARY ROCKFISH
CHINA ROCKFISH
COPPER ROCKFISH
LINGCOD
QUILLBACK ROCKFISH

Supplementary Materials for Dick et al. ("Model-Based Estimation of Average Weights from Recreational Fisheries")

The Washington average weight procedure does not calculate separate estimates of average weight for retained and released fish, thus for a given stratification, values of retained and released average weight will be equal. The procedure begins by calculating an imputed weight for each individual Washington sample where a length is provided, but no corresponding weight. The equation used to calculate imputed weight is:

$$imputed\ weight = a \times length^b \tag{1}$$

Where a and b are the species-specific length-weight parameters and length is the length of the sampled fish measured in mm. Records where the length of the sampled fish exceed a stored maximum length threshold are excluded by the procedure and thus, not used in the average weight calculation. For each species, the imputed and measured weights are summed along with a count of the number of weight samples for each year-month stratification. These totals are used to calculate an estimated average weight for each processing-month. The degree of stratification employed in the calculation is dependent upon the number of available weight samples at each stratum level (Table S1, Figure S1). For each aggregation level (hereafter "agg level"), the procedure requires ≥ 50 samples to generate an average weight for a given species. The Washington procedure begins at agg level 7 and assigns an average weight to each year, month, agency, and species stratum. The average weight computation begins with the processing-month and works backwards monthly through time, counting samples and stopping on the month where the sample count reaches ≥ 50 for each given species. Using those samples, an average weight is calculated for the current processing-month. If the 50 sample threshold is not reached at agg level 7, and a fixed weight exists for the respective species and current processing-year, the average weight is assigned from the fixed weight table (agg level 7.1). If an average weight from the fixed weight table is not available for the current processing-year, the procedure selects the average weight from the fixed weight table for the most recent available year for the respective species to assign to the processing-month (agg level 7.2). If an average weight does not exist in the fixed weight table for a given species, then an average weight is calculated using Washington sample data from all available years (agg level 14).

Table S1. Levels of aggregation for average weight calculation, as defined for WDFW borrowing rules. Source: RecFIN 2020.

Agg Level	Agency	Year	Month	Species	Agency	Subregion	Port	Trip Type	Mode	Notes
7	WA			X	X					- Requires at least 50 samples to calculate average weight - No date limit on collection of past records to calculate average weight
7.1	WA	X		X	X					- Average weight value from fixed weight table for processing year
7.2	WA			X	X					- Most recent average weight value from fixed weight table
14	WA			X	X					- Average weight calculated using retained samples from WA for all years - No date limit on collection of past records to calculate average weight

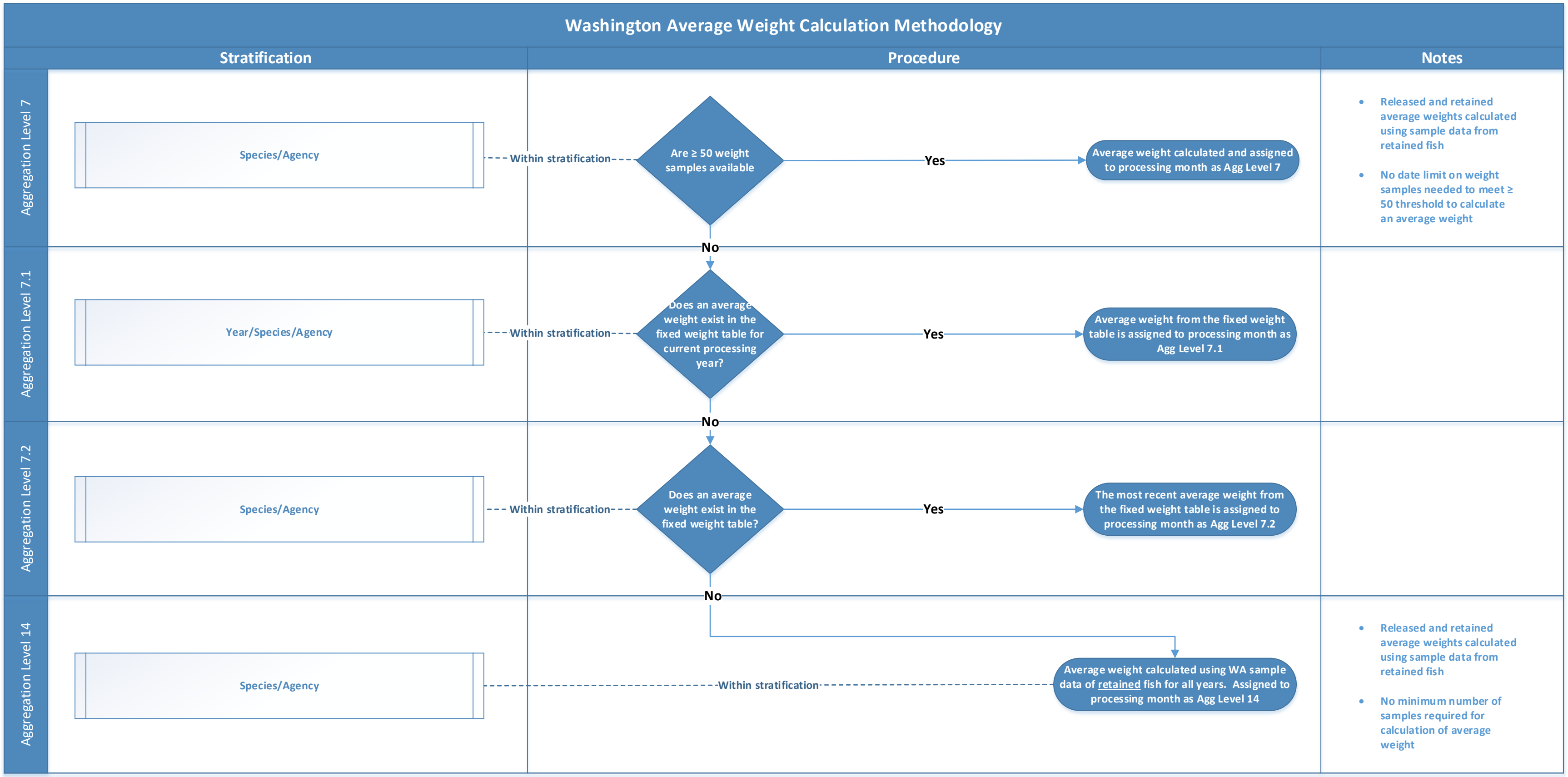


Figure S1. Flowchart of WDFW borrowing rules. Source: RecFIN 2020.

Results for Blue Rockfish

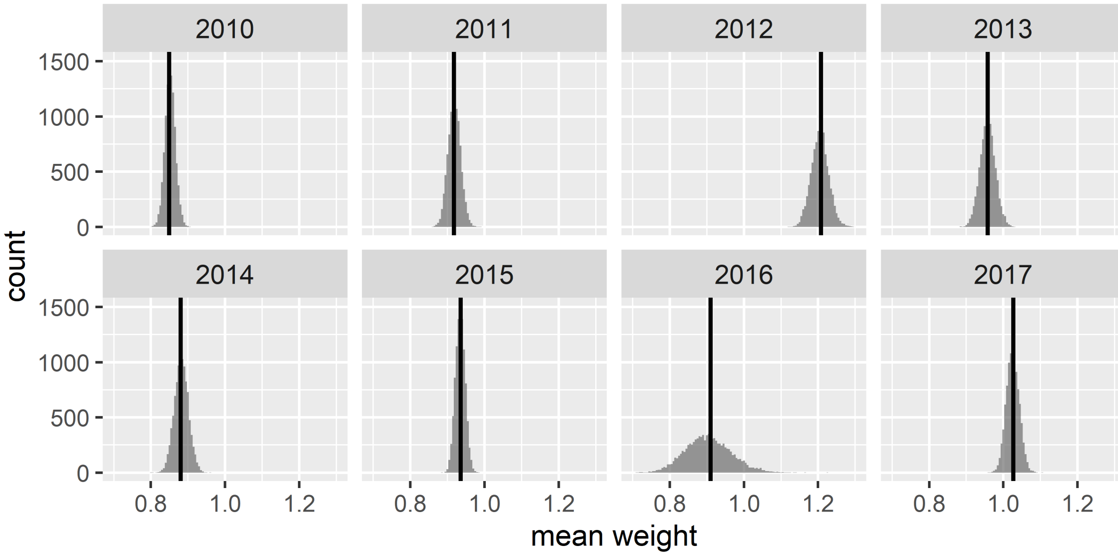


Figure S2. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by year (histograms) for blue rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

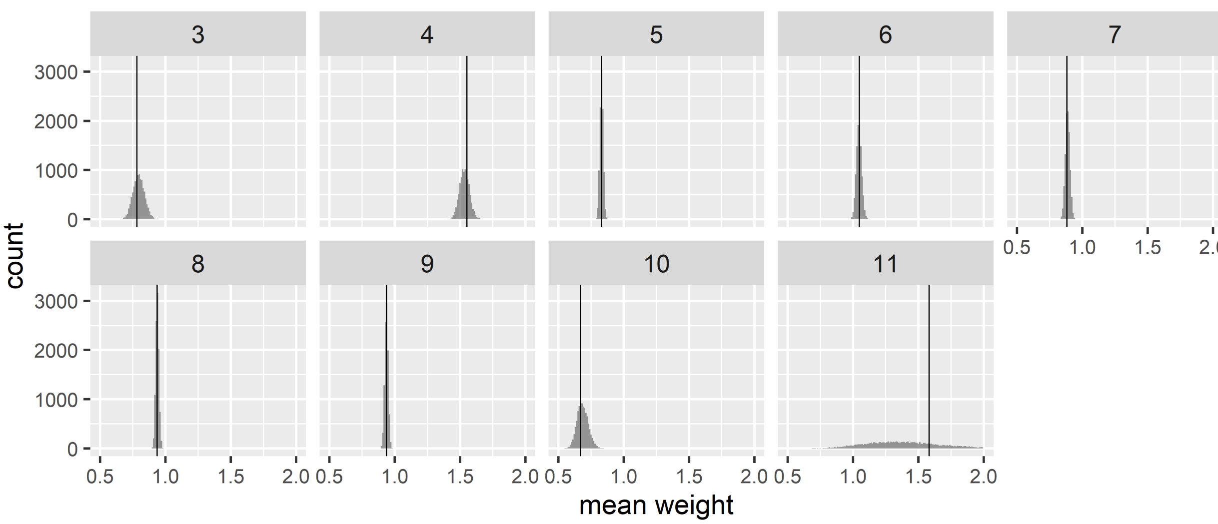


Figure S3. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by month (histograms) for blue rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

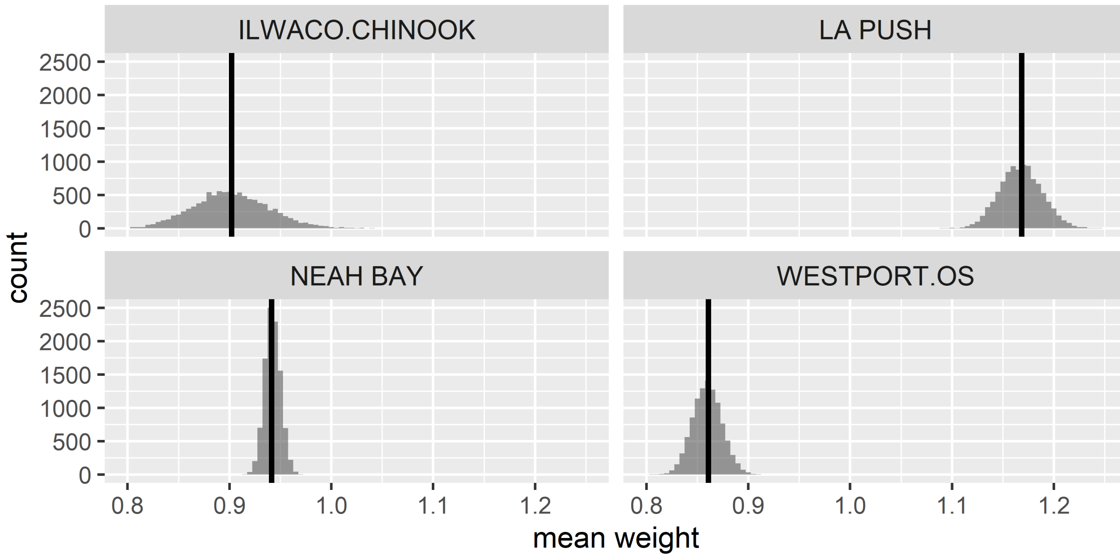


Figure S4. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by port group (histograms) for blue rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

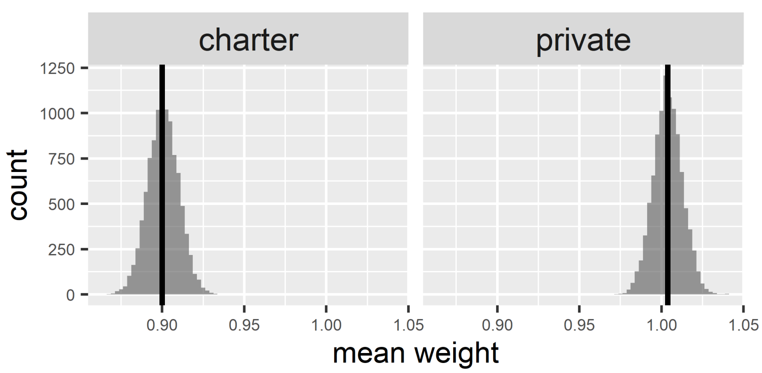


Figure S5. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by boat mode (histograms) for blue rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

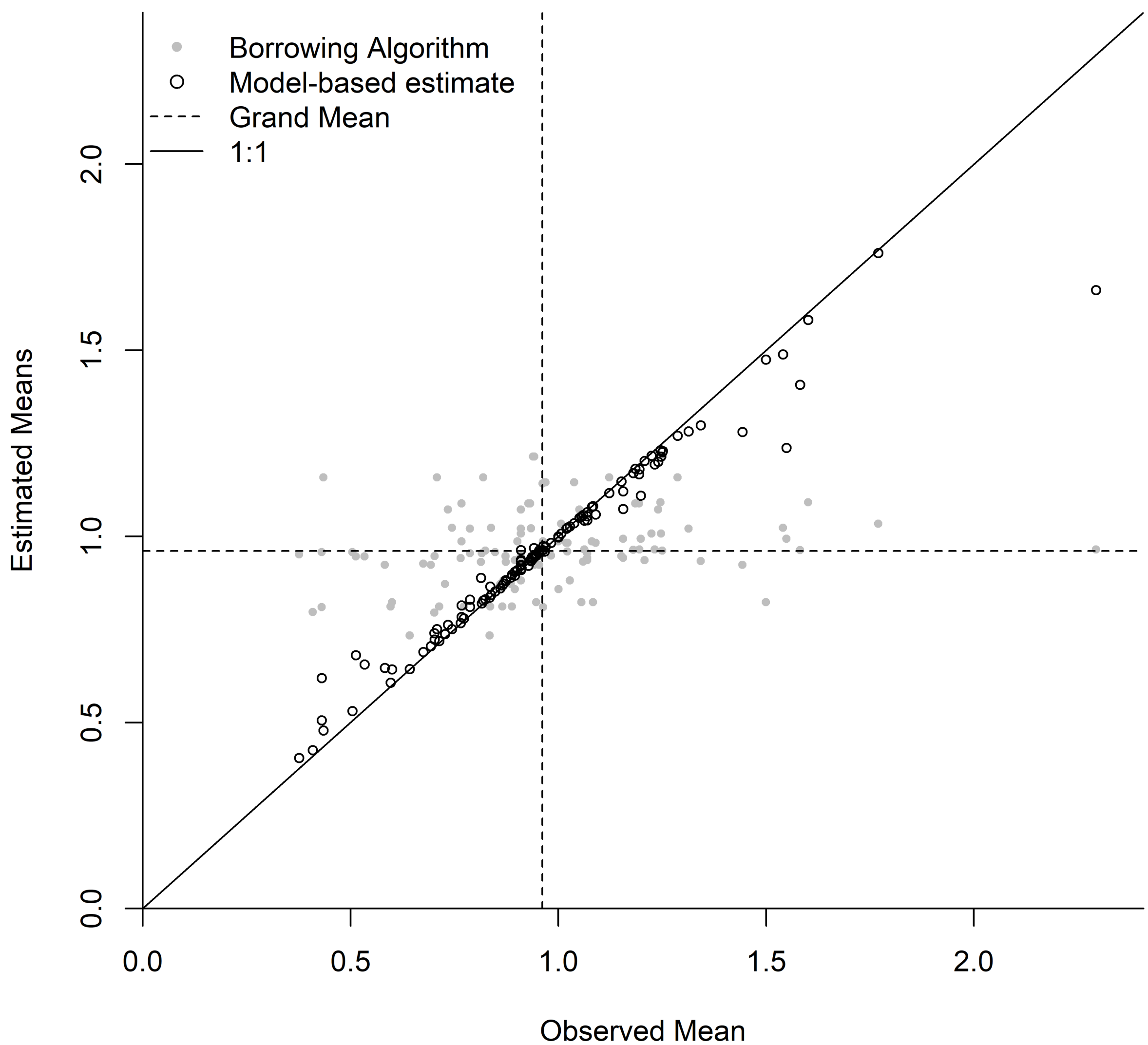


Figure S6. Estimated posterior mean weight [kg] from model 13 (black open circles) and estimates from the borrowing algorithm (grey solid circles) versus observed arithmetic means for blue rockfish. Each point represents a single year/month/port/mode combination (127 observed strata). Horizontal and vertical dashed lines are the arithmetic mean weight across all observations (grand mean). Solid line is the line of equality (1:1).

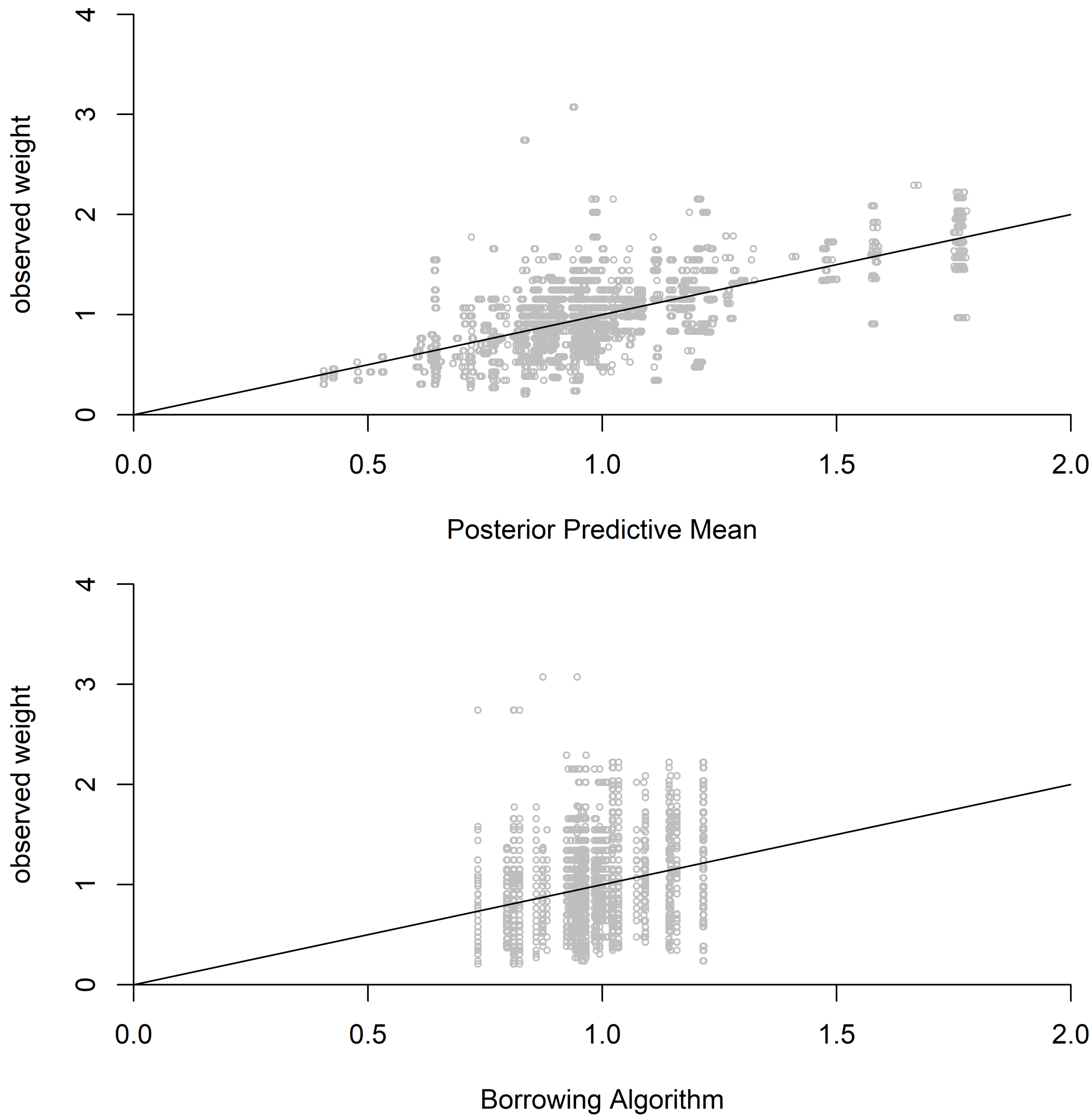


Figure S7. Comparison of observed weights [kg] for blue rockfish to predicted mean weights from model 13 (upper panel) and estimates from the borrowing algorithm (lower panel). Solid black lines are 1:1 for reference.

Results for Cabezon

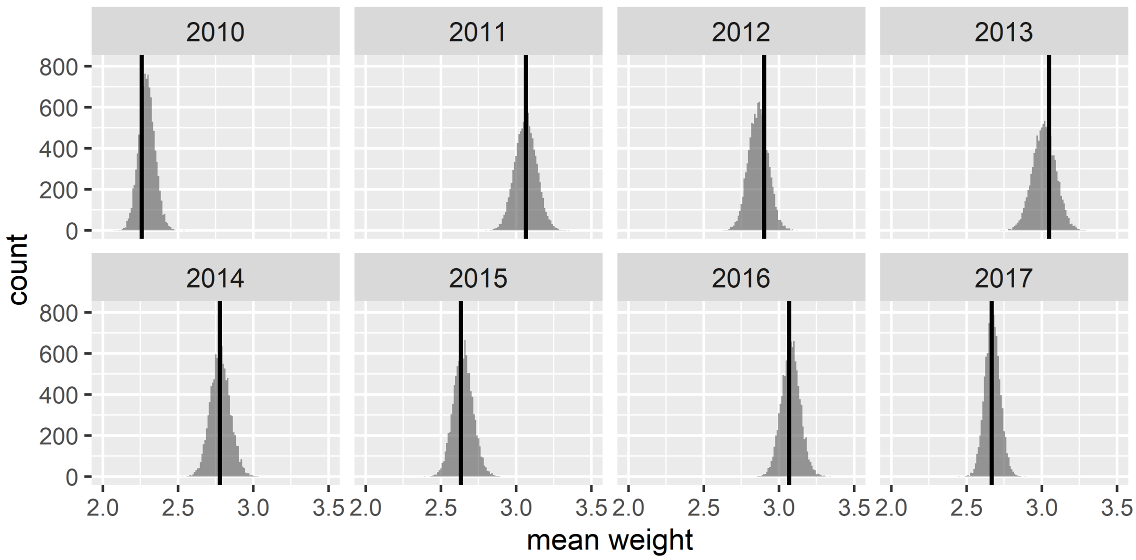


Figure S8. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by year (histograms) for cabezon, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

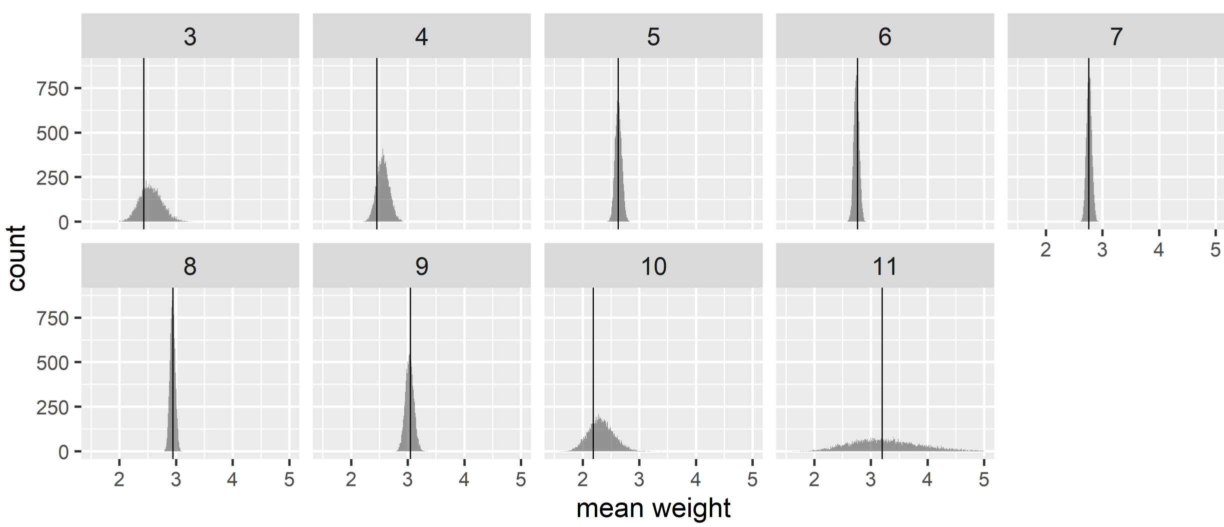


Figure S9. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by month (histograms) for cabezon, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

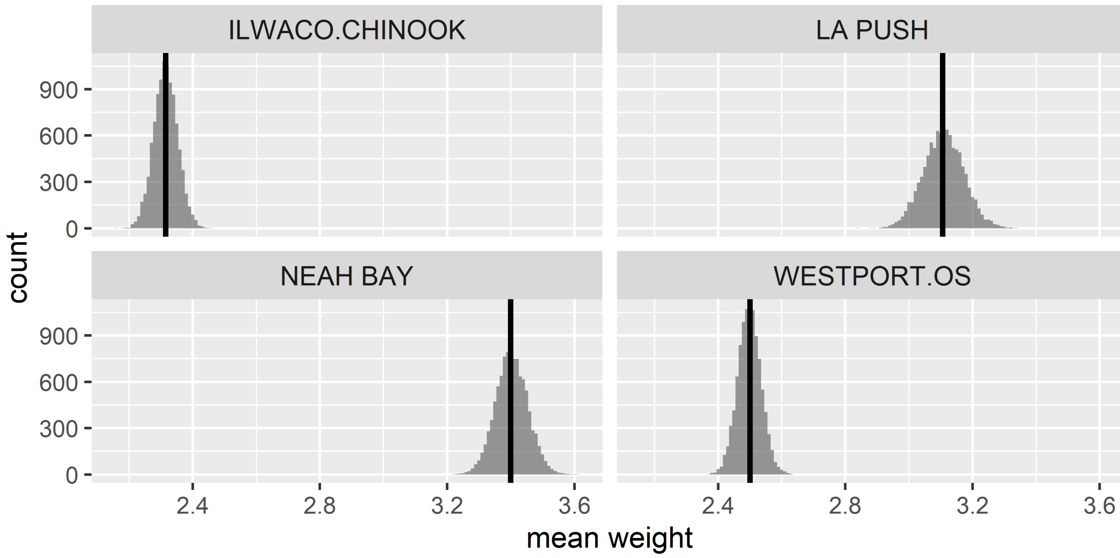


Figure S10. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by port group (histograms) for cabezon, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

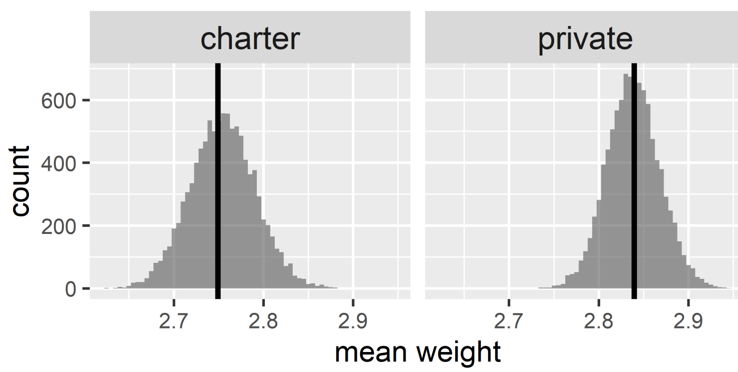


Figure S11. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by boat mode (histograms) for cabezon, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

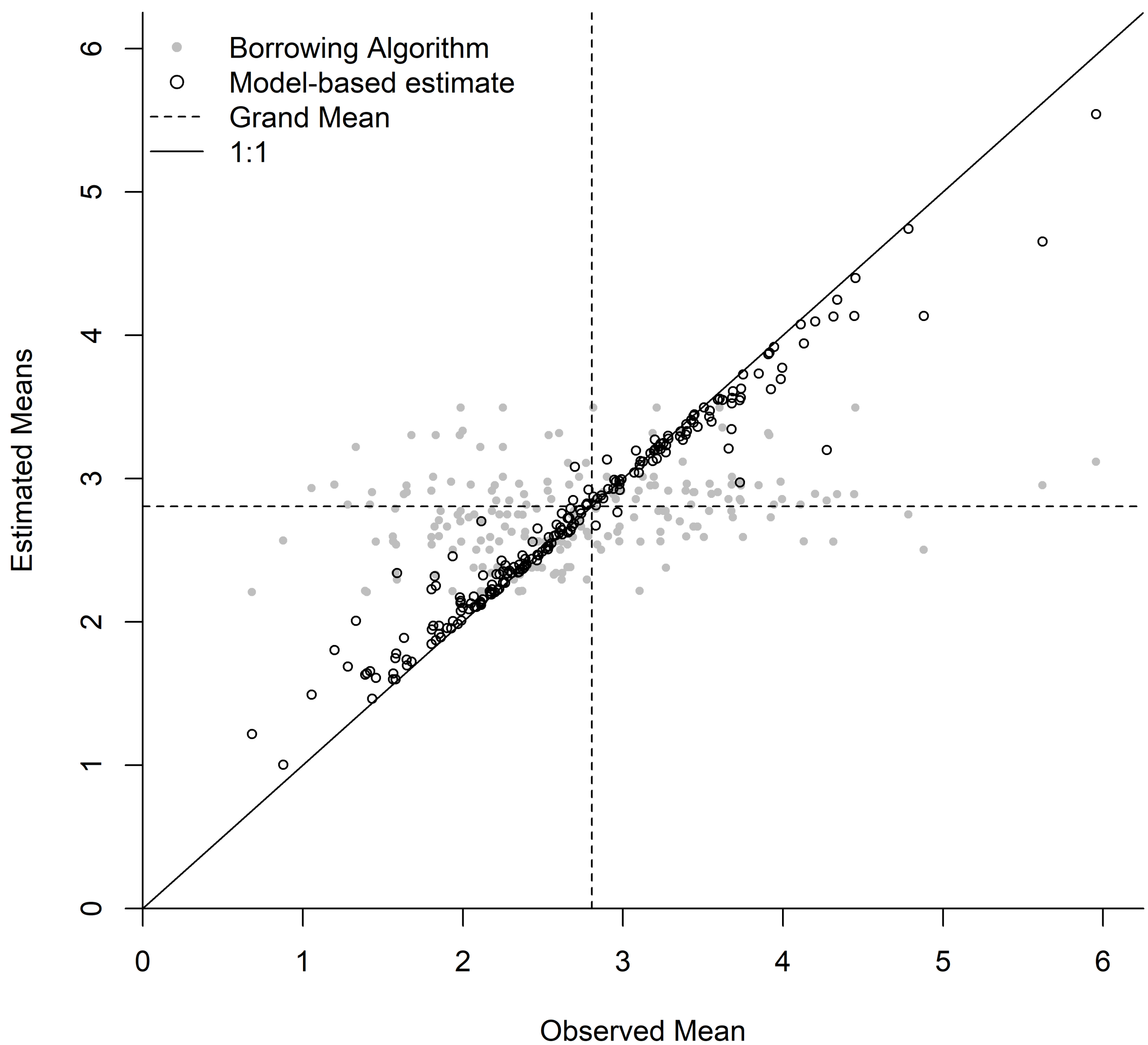


Figure S12. Estimated posterior mean weight [kg] from model 13 (black open circles) and estimates from the borrowing algorithm (grey solid circles) versus observed arithmetic means for cabezon. Each point represents a single year/month/port/mode combination (206 observed strata). Horizontal and vertical dashed lines are the arithmetic mean weight across all observations (grand mean). Solid line is the line of equality (1:1).

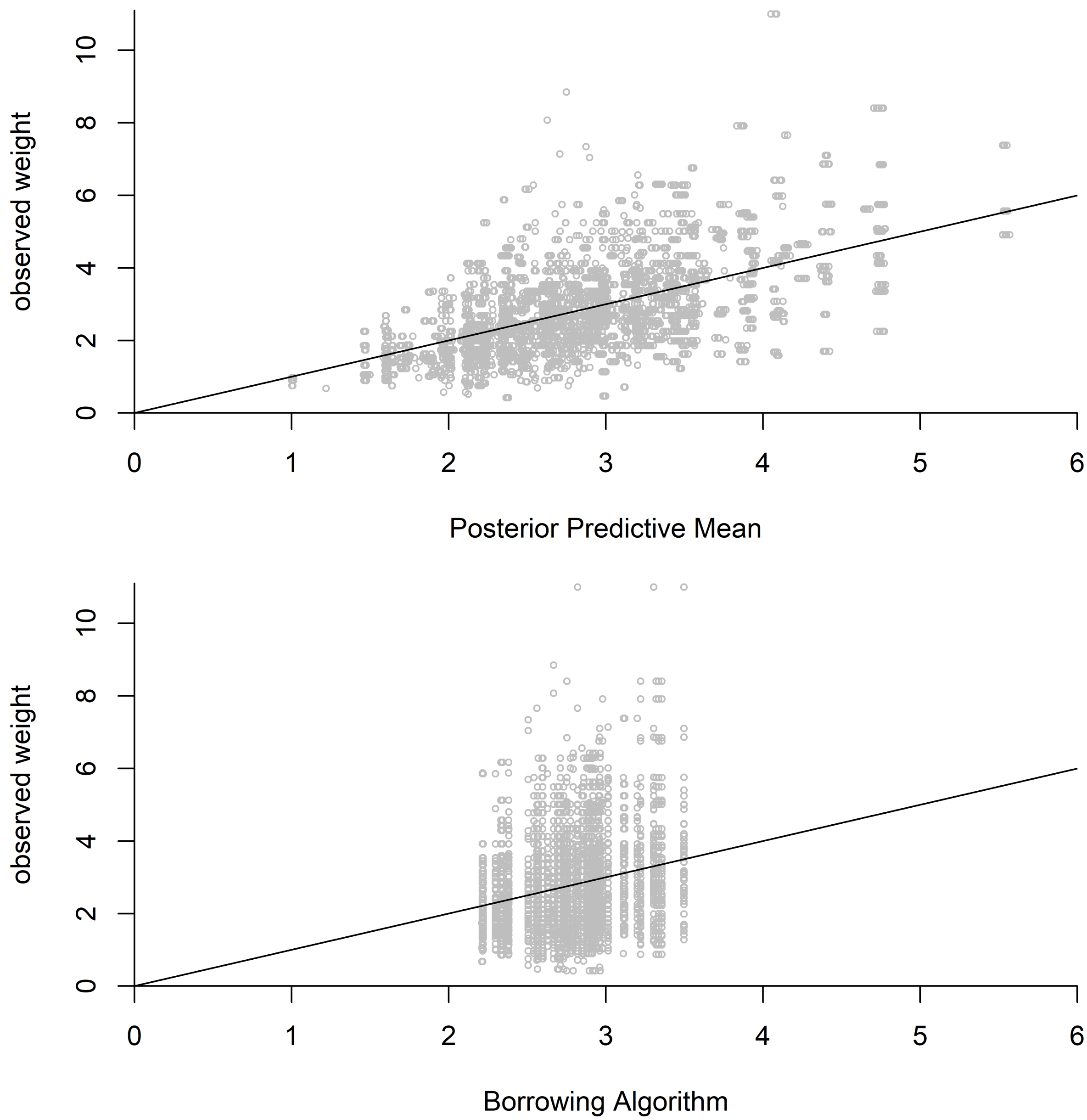


Figure S13. Comparison of observed weights [kg] for cabezon to predicted mean weights from model 13 (upper panel) and estimates from the borrowing algorithm (lower panel). Solid black lines are 1:1 for reference.

Results for Canary Rockfish

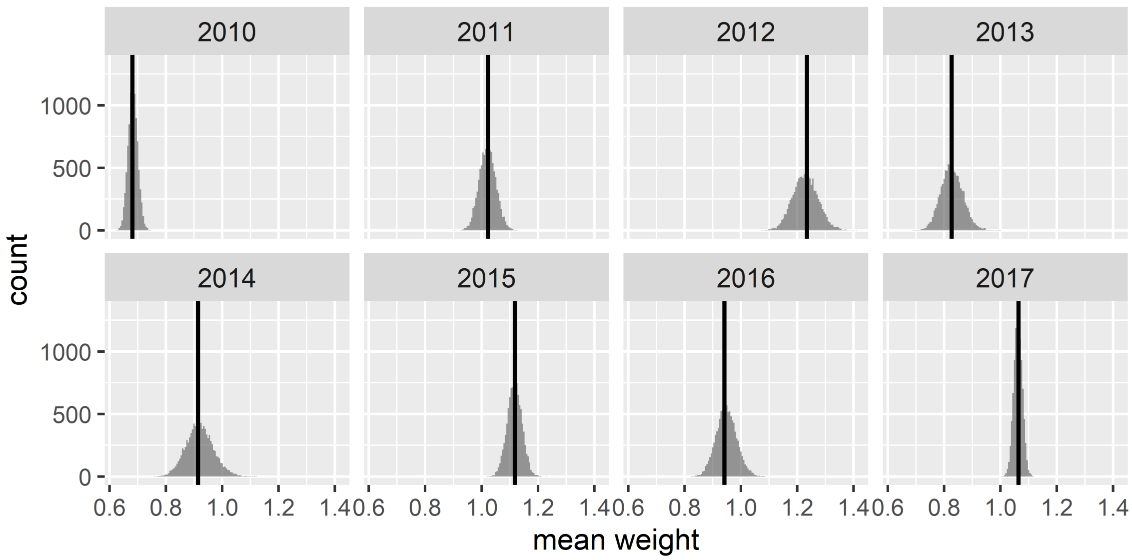


Figure S14. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by year (histograms) for canary rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

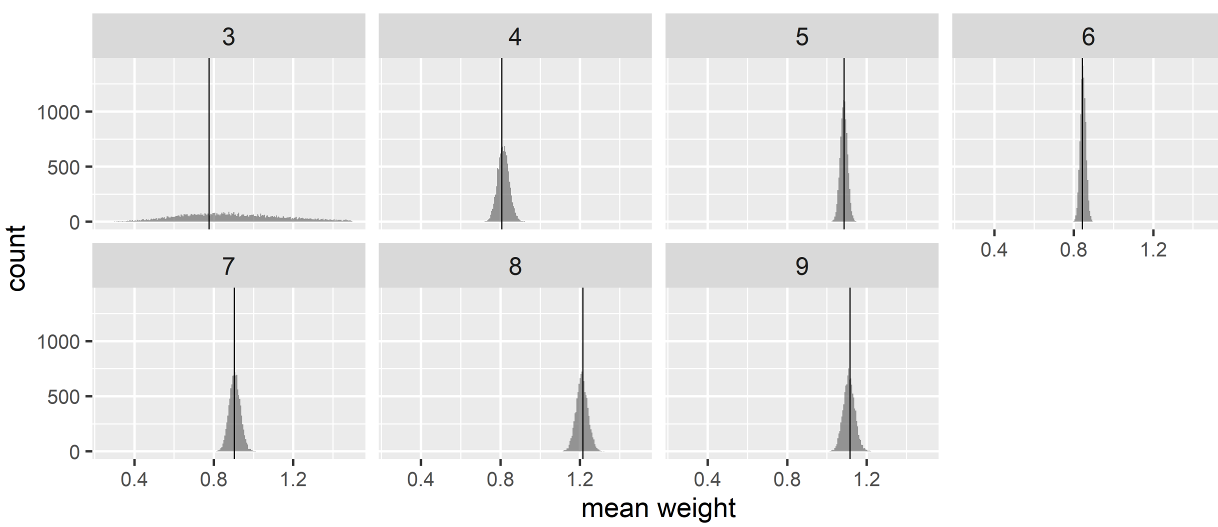


Figure S15. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by month (histograms) for canary rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

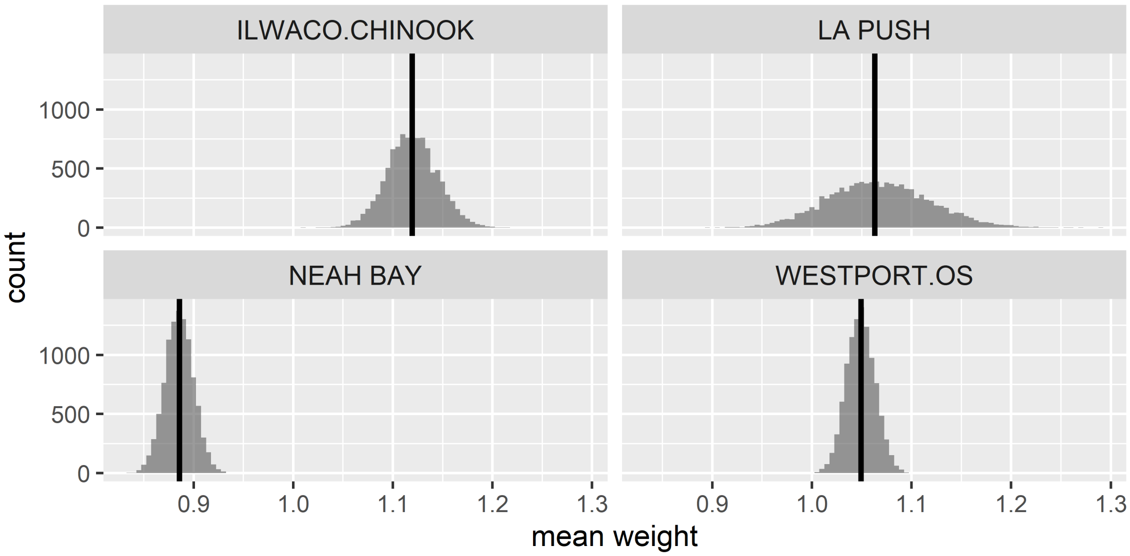


Figure S16. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by port group (histograms) for canary rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

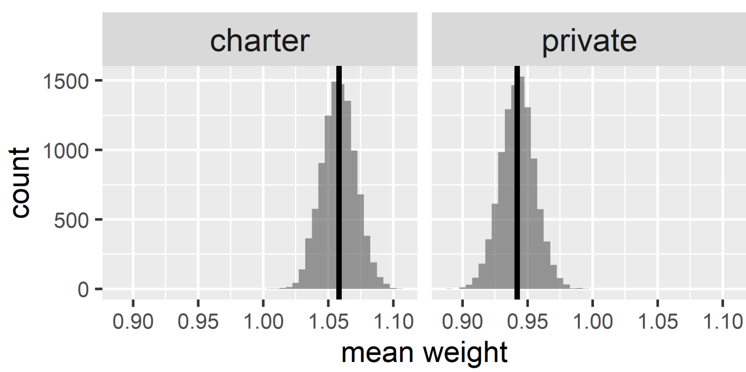


Figure S17. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by boat mode (histograms) for canary rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

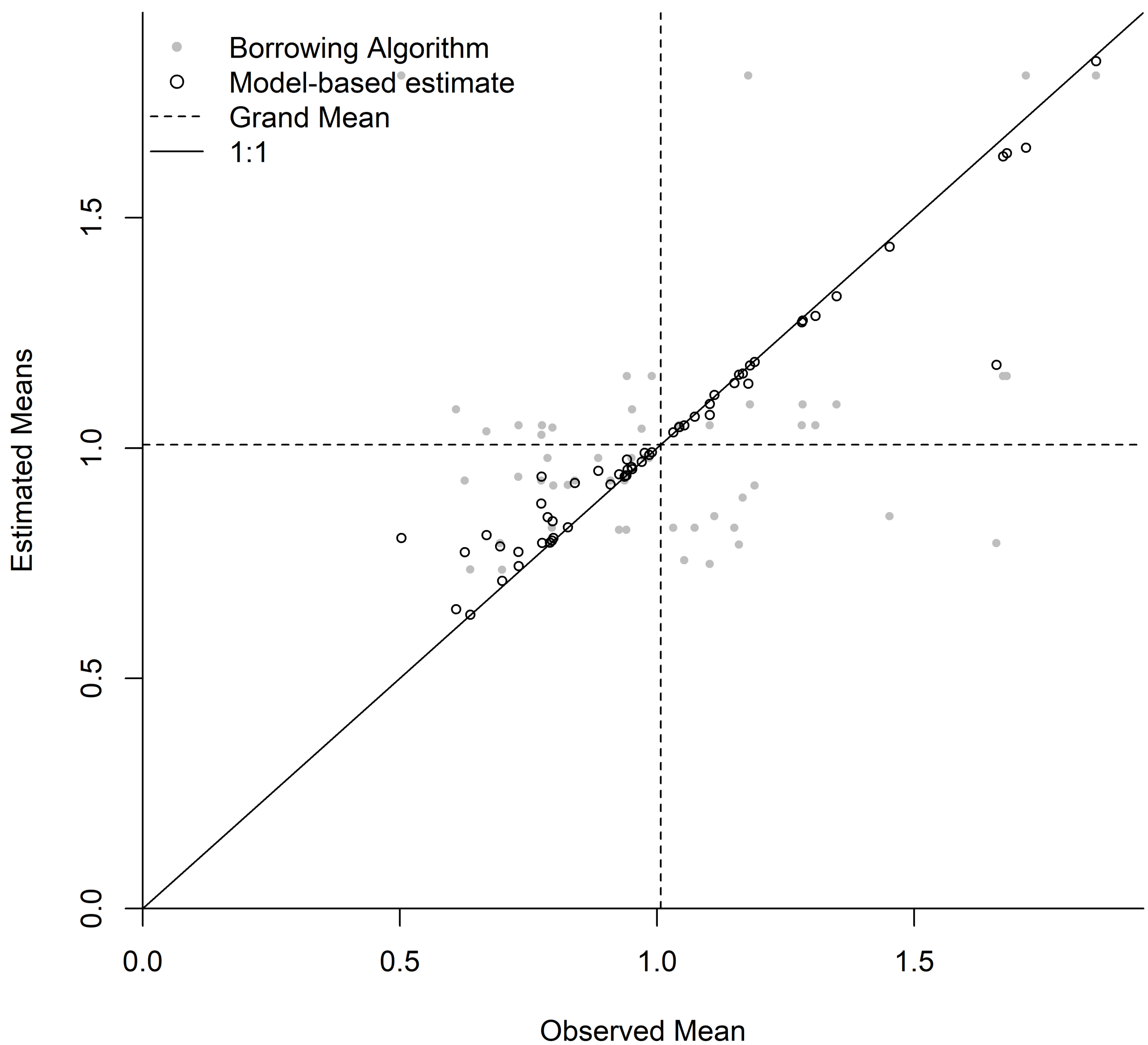


Figure S18. Estimated posterior means from model 13 (black open circles) and estimates from the borrowing algorithm (grey solid circles) versus observed arithmetic means for canary rockfish. Each point represents a single year/month/port/mode combination (56 observed strata). Horizontal and vertical dashed lines are the arithmetic mean weight across all observations (grand mean). Solid line is the line of equality (1:1).

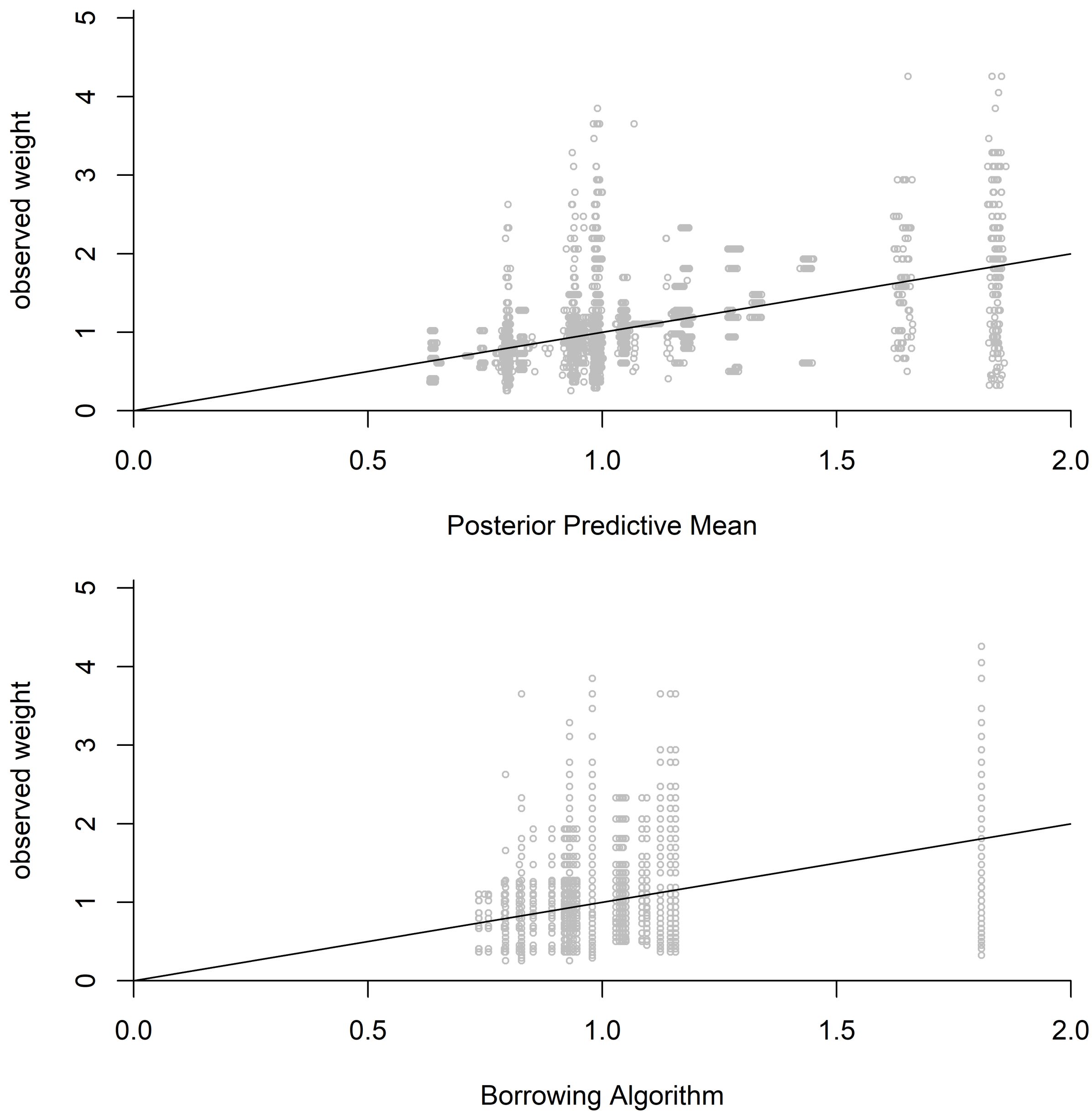


Figure S19. Comparison of observed weights [kg] for canary rockfish to predicted mean weights from model 13 (upper panel) and estimates from the borrowing algorithm (lower panel). Solid black lines are 1:1 for reference.

Results for China Rockfish

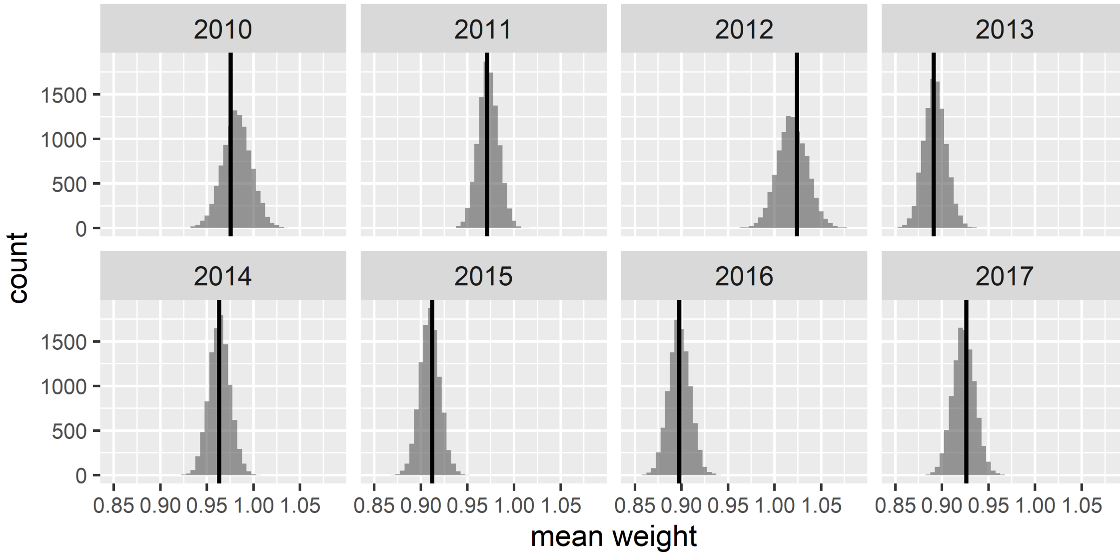


Figure S20. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by year (histograms) for China rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

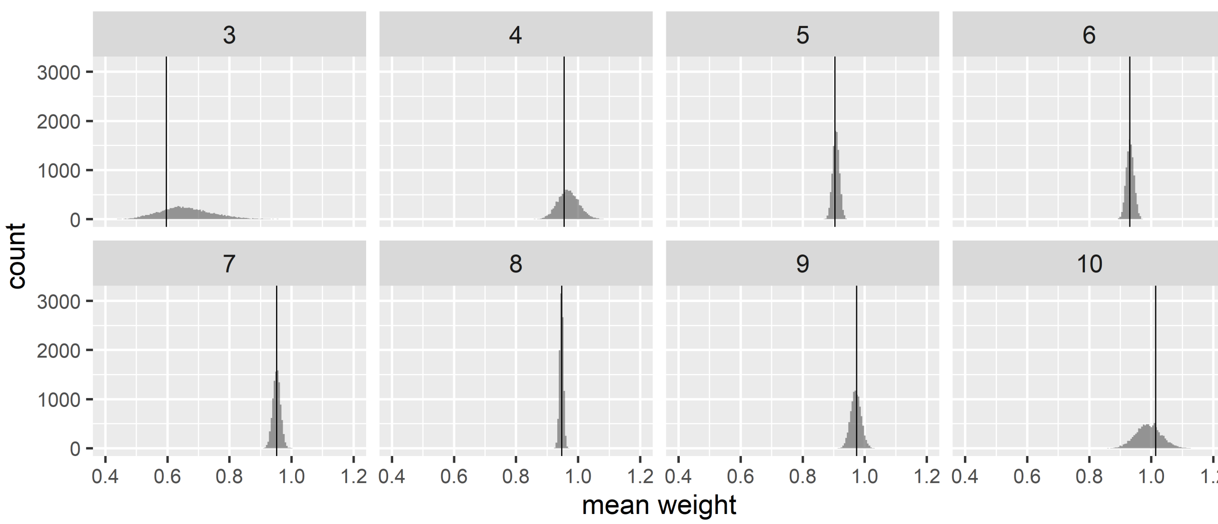


Figure S21. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by month (histograms) for China rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

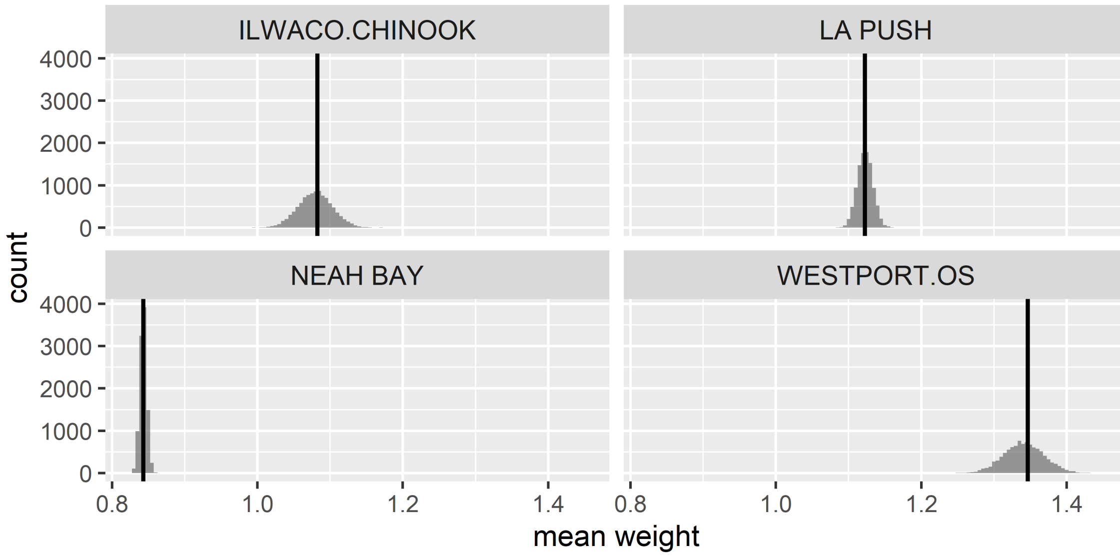


Figure S22. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by port group (histograms) for China rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

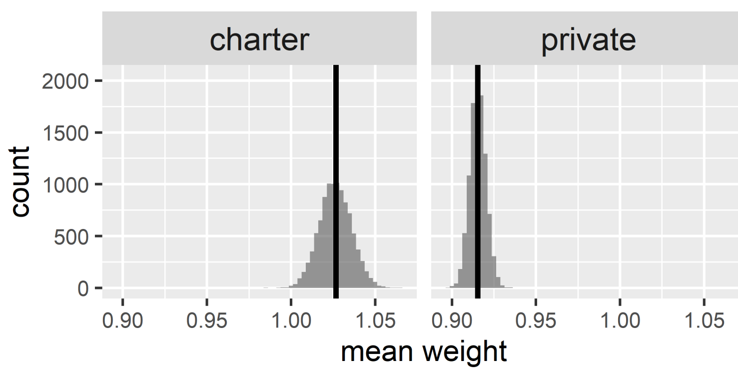


Figure S23. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by boat mode (histograms) for China rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

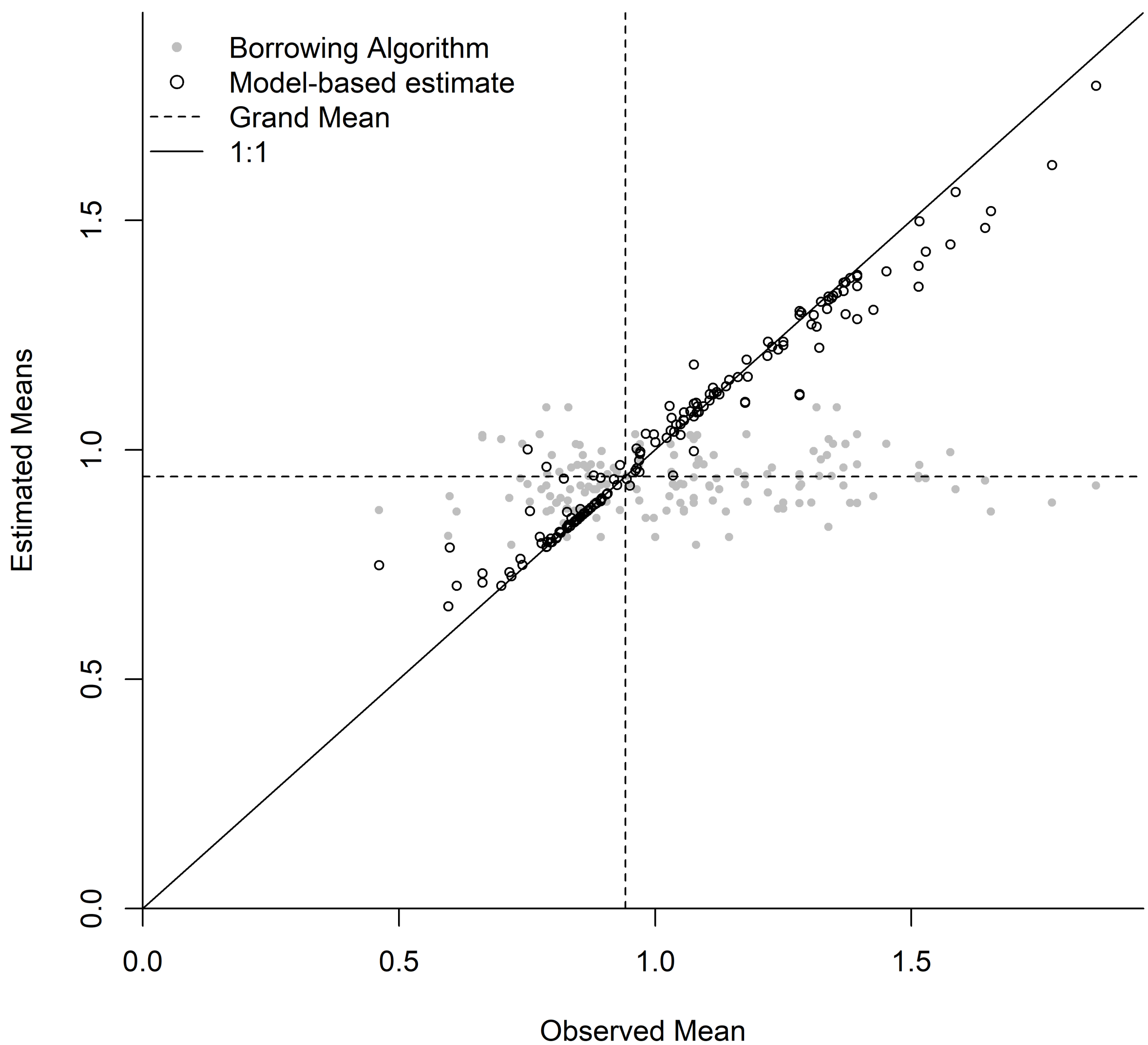


Figure S24. Estimated posterior means [kg] from model 13 (black open circles) and estimates from the borrowing algorithm (grey solid circles) versus observed arithmetic means for China rockfish. Each point represents a single year/month/port/mode combination (157 observed strata). Horizontal and vertical dashed lines are the arithmetic mean weight across all observations (grand mean). Solid line is the line of equality (1:1).

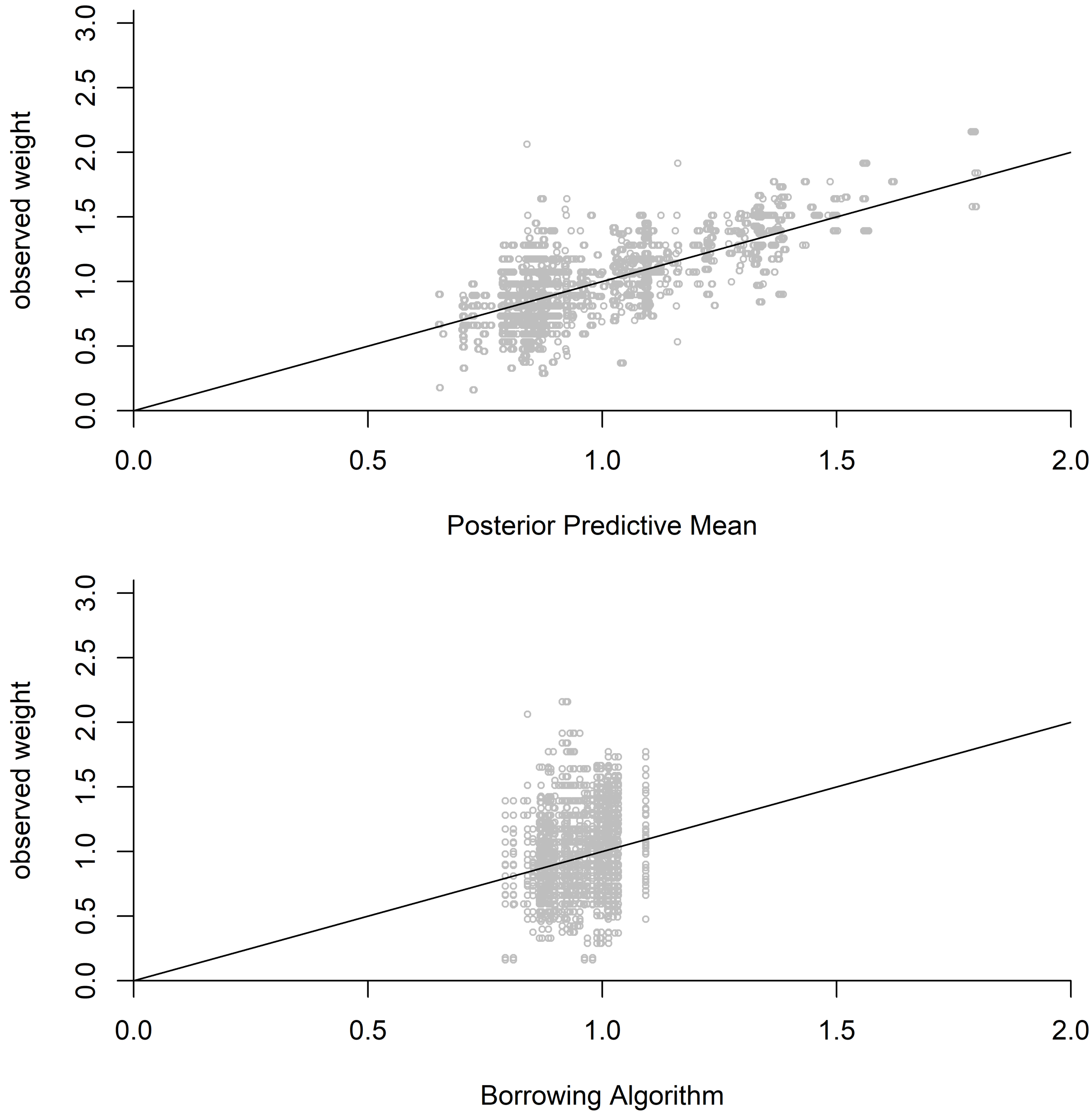


Figure S25. Comparison of observed weights [kg] for China rockfish to predicted mean weights from model 13 (upper panel) and estimates from the borrowing algorithm (lower panel). Solid black lines are 1:1 for reference.

Results for Copper Rockfish

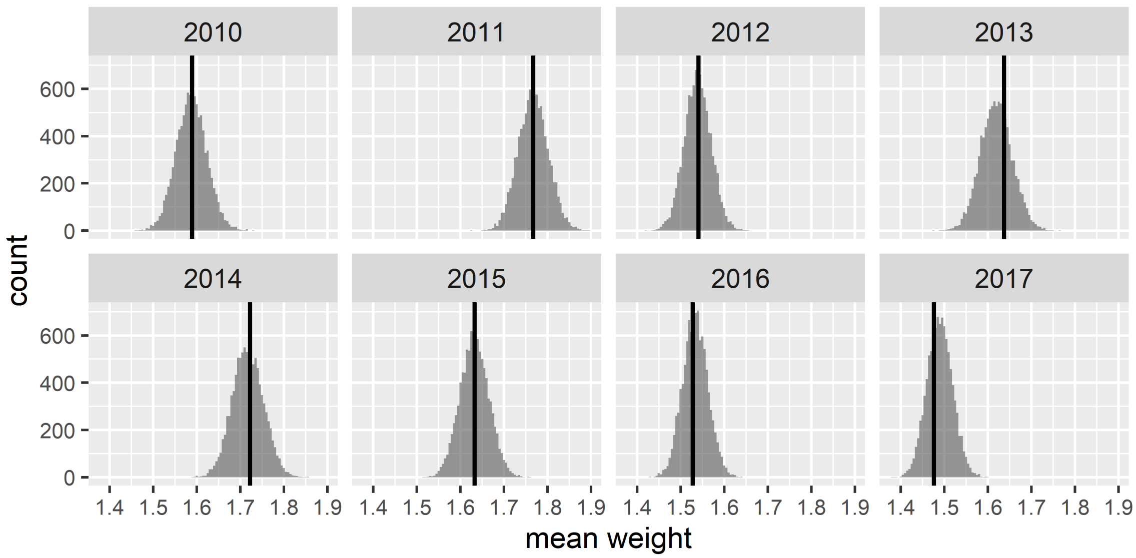


Figure S26. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by year (histograms) for copper rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

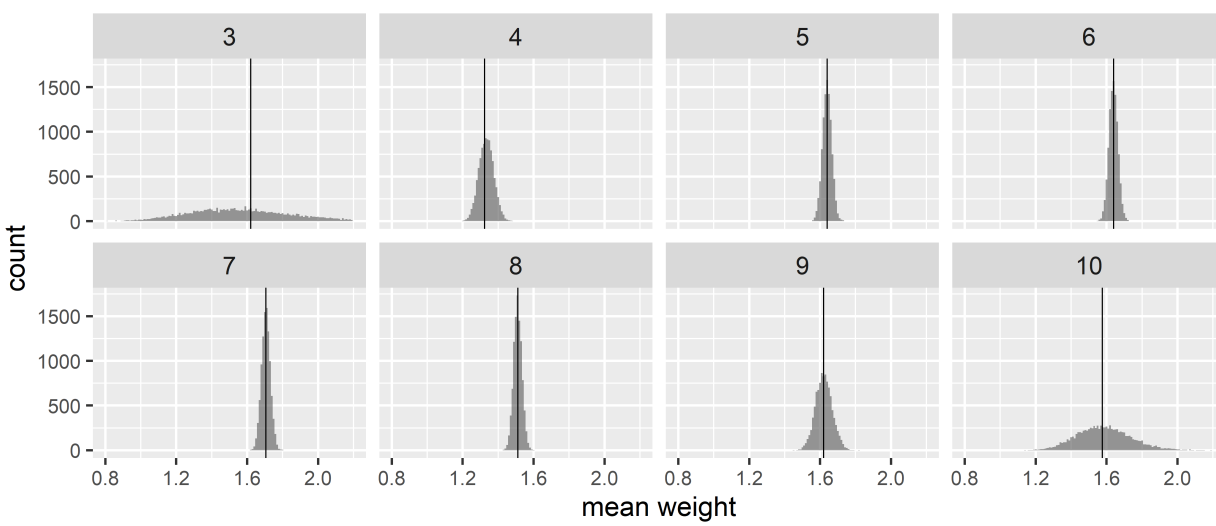


Figure S27. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by month (histograms) for copper rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

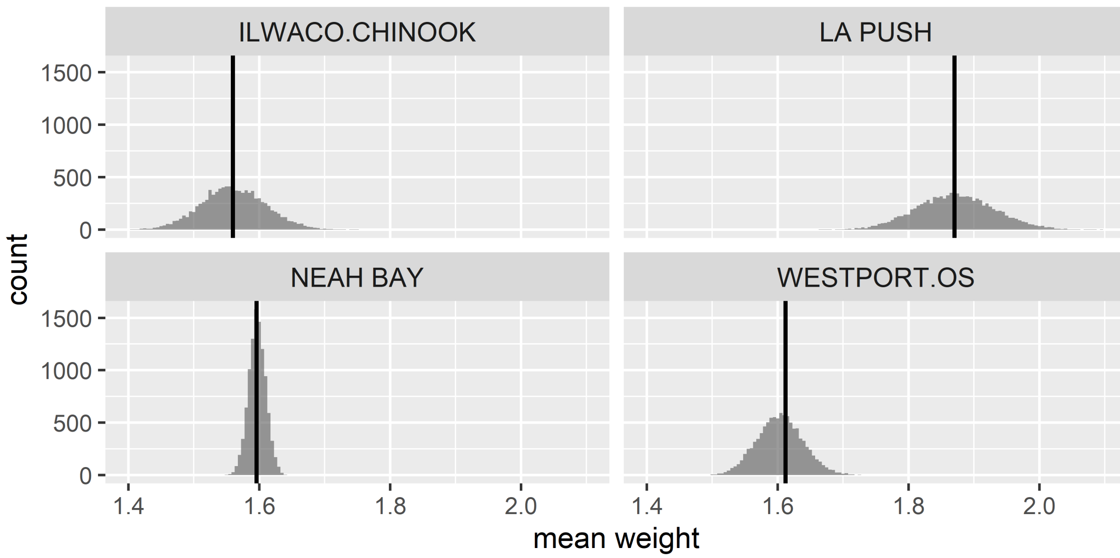


Figure S28. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by port group (histograms) for copper rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

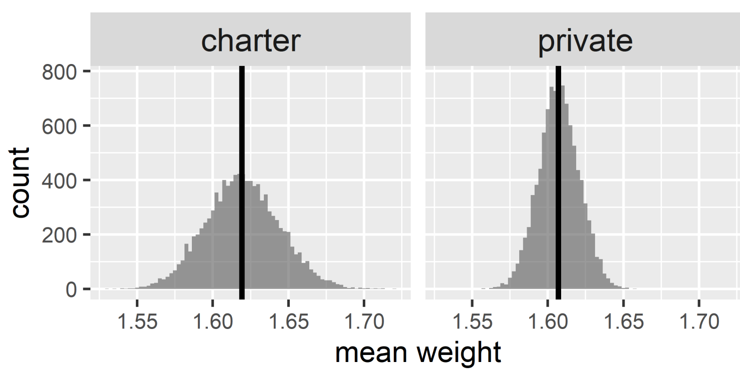


Figure S29. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by boat mode (histograms) for copper rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

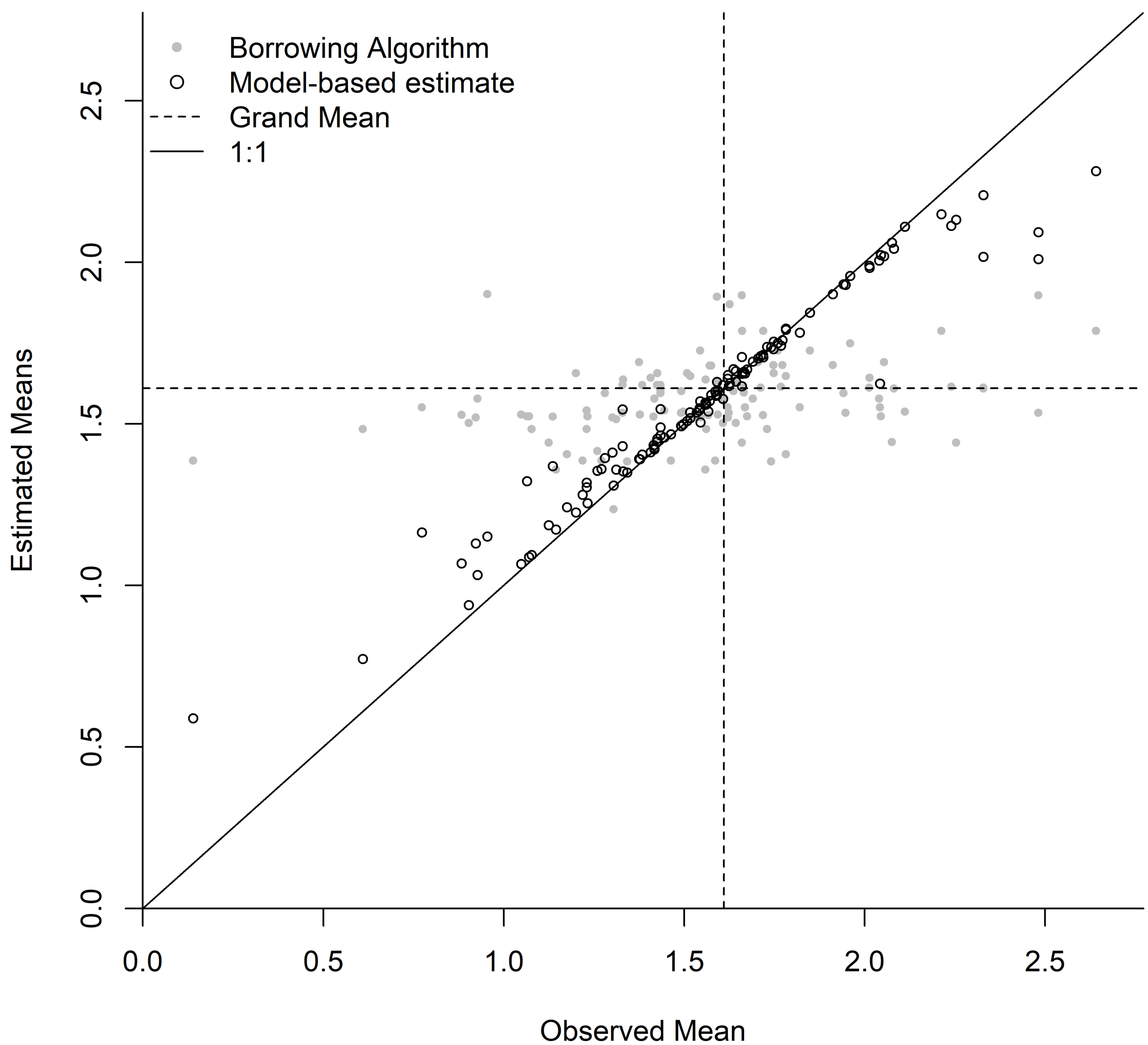


Figure S30. Estimated posterior means [kg] from model 13 (black open circles) and estimates from the borrowing algorithm (grey solid circles) versus observed arithmetic means for copper rockfish. Each point represents a single year/month/port/mode combination (130 observed strata). Horizontal and vertical dashed lines are the arithmetic mean weight across all observations (grand mean). Solid line is the line of equality (1:1).

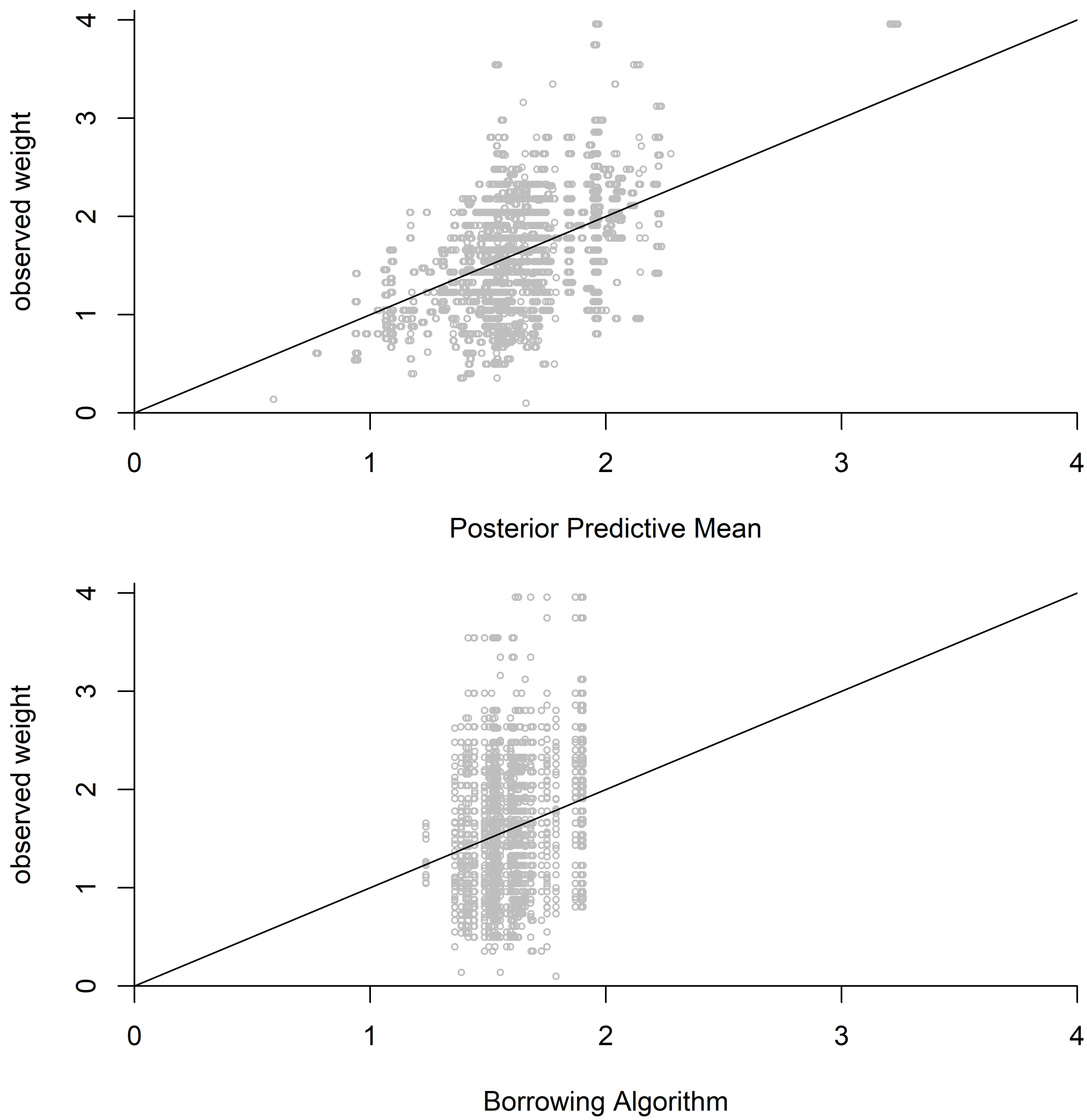


Figure S31. Comparison of observed weights [kg] for copper rockfish to predicted mean weights from model 13 (upper panel) and estimates from the borrowing algorithm (lower panel). Solid black lines are 1:1 for reference.

Results for Lingcod

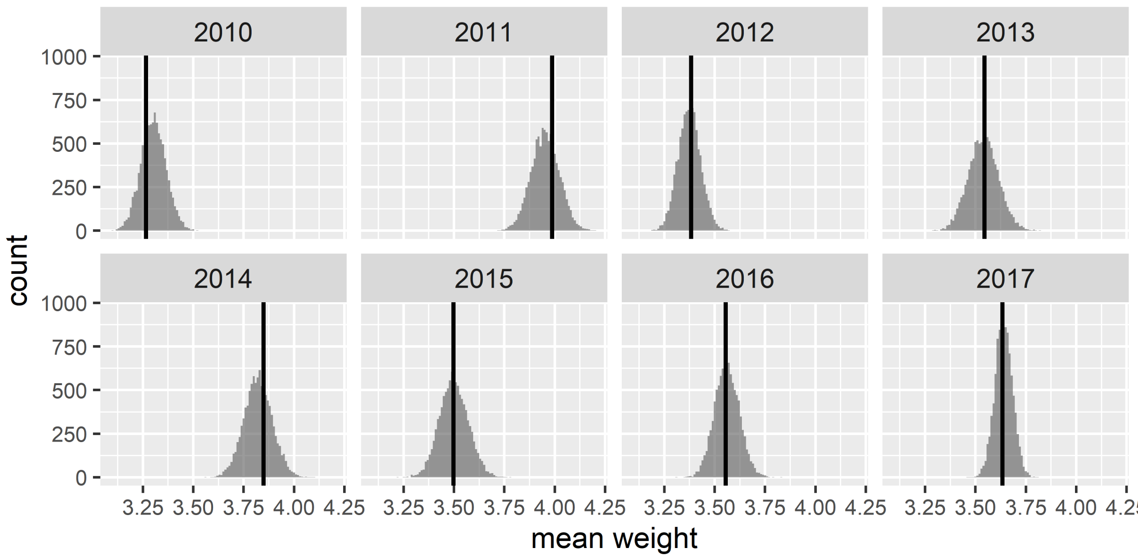


Figure S32. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by year (histograms) for lingcod, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

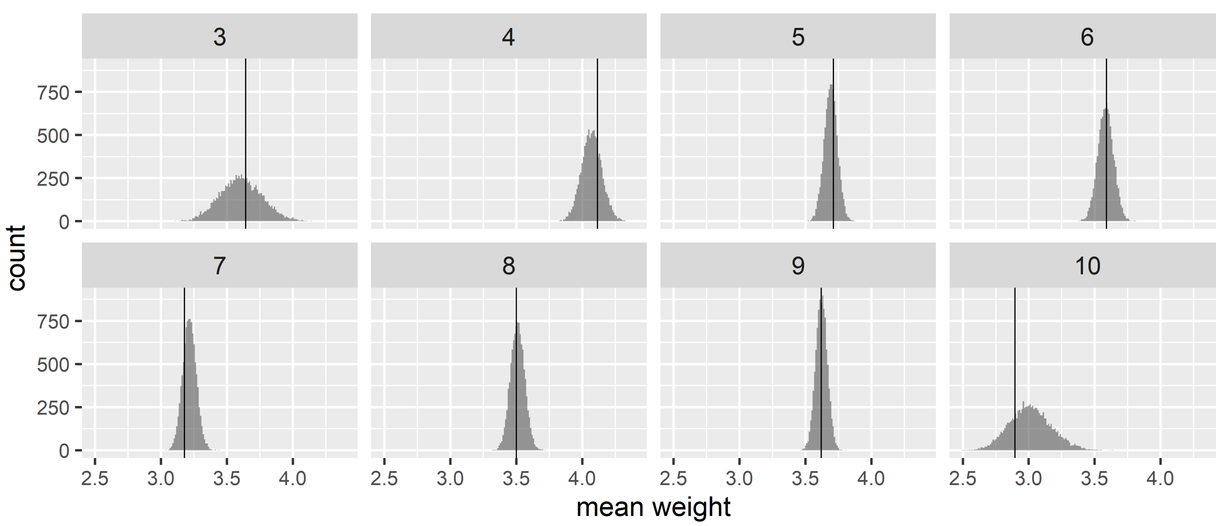


Figure S33. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by month (histograms) for lingcod, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

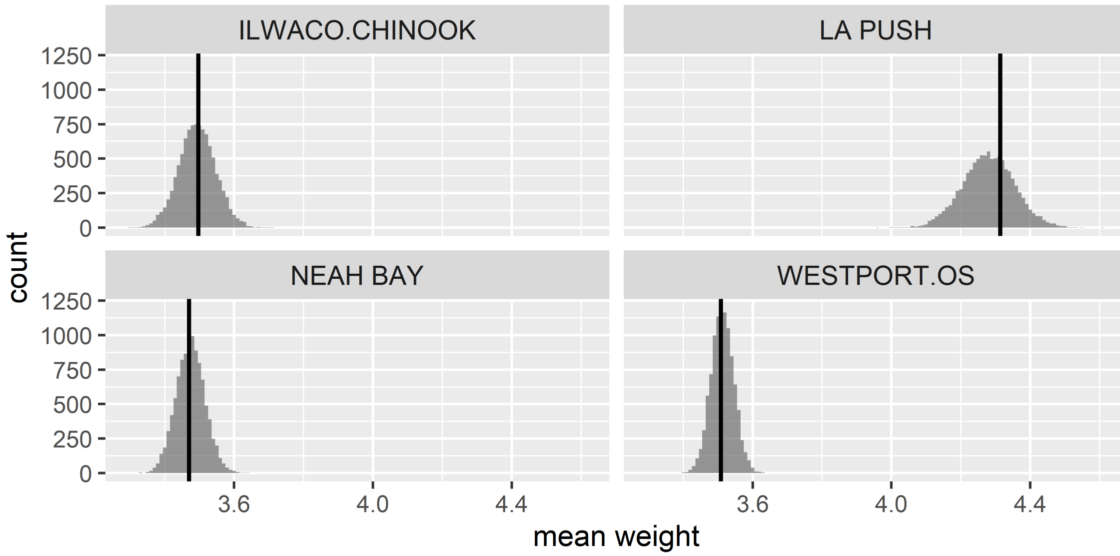


Figure S34. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by port group (histograms) for lingcod, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

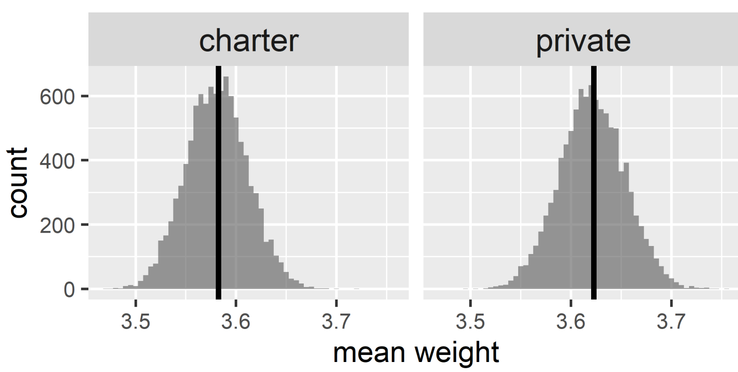


Figure S35. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by boat mode (histograms) for lingcod, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

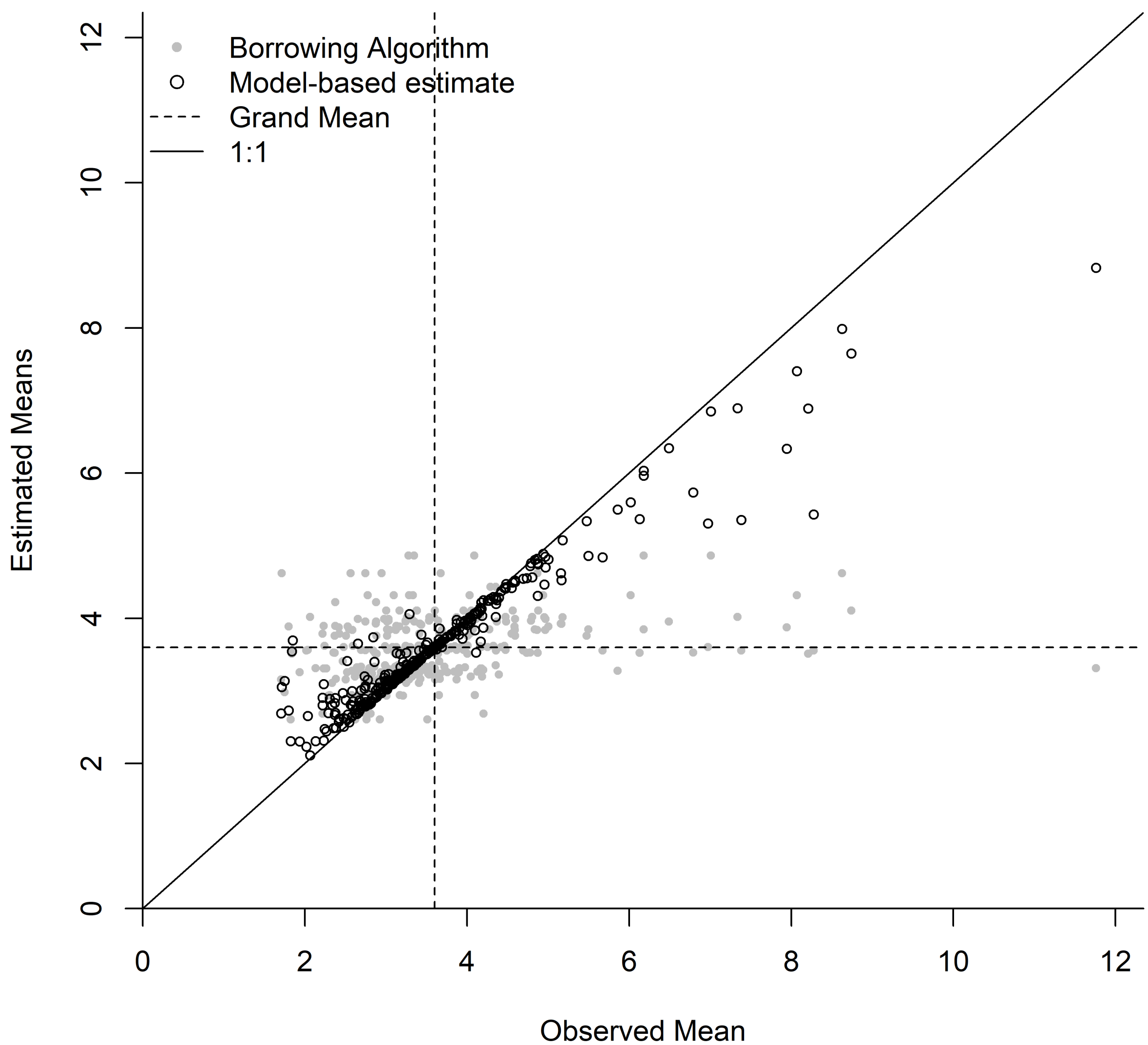


Figure S36. Estimated posterior means [kg] from model 13 (black open circles) and estimates from the borrowing algorithm (grey solid circles) versus observed arithmetic means for lingcod. Each point represents a single year/month/port/mode combination (306 observed strata). Horizontal and vertical dashed lines are the arithmetic mean weight across all observations (grand mean). Solid line is the line of equality (1:1).

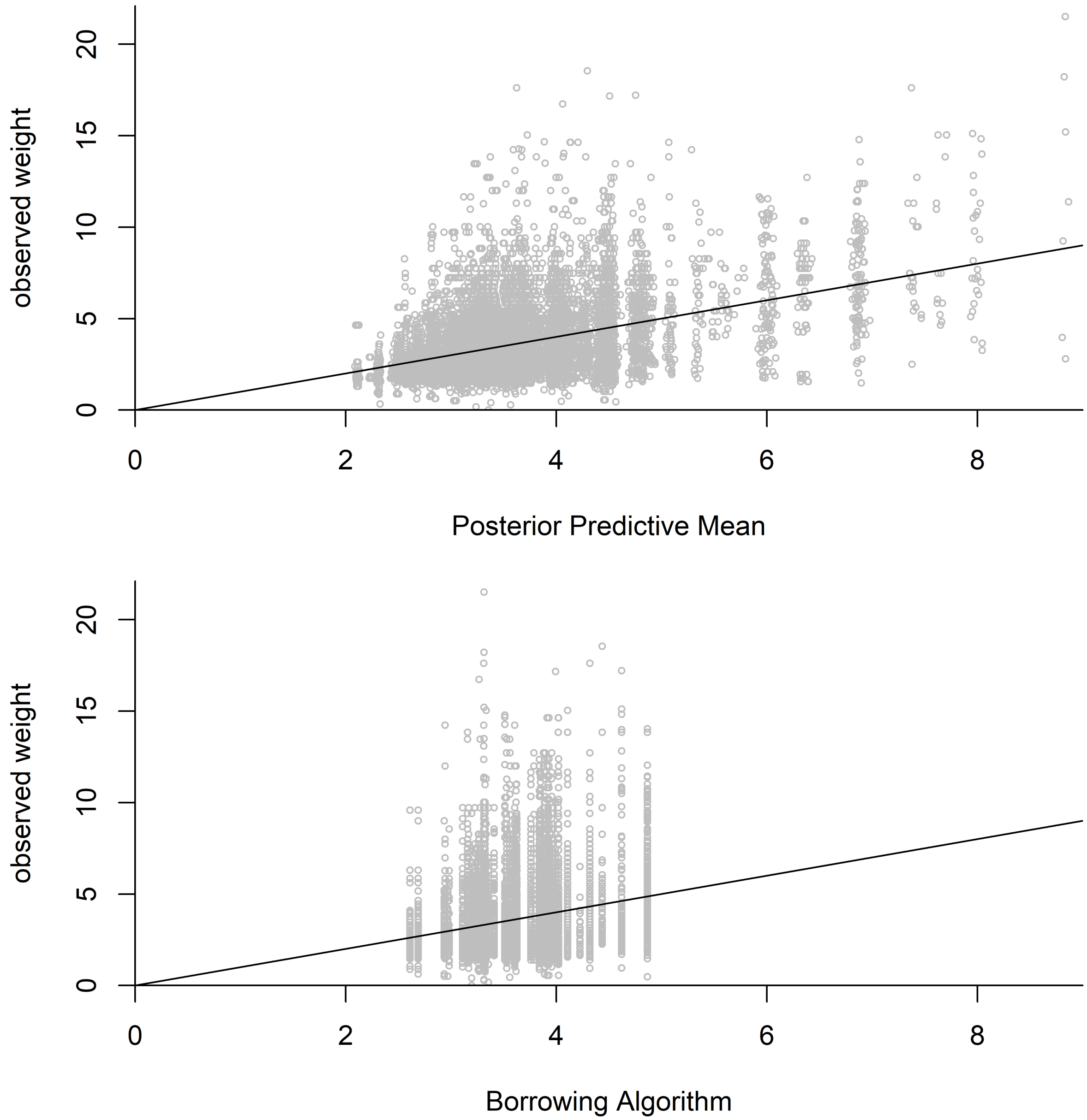


Figure S37. Comparison of observed weights [kg] for lingcod to predicted mean weights from model 13 (upper panel) and estimates from the borrowing algorithm (lower panel). Solid black lines are 1:1 for reference.

Results for Quillback Rockfish

Figure S38. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by year (histograms) for quillback rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

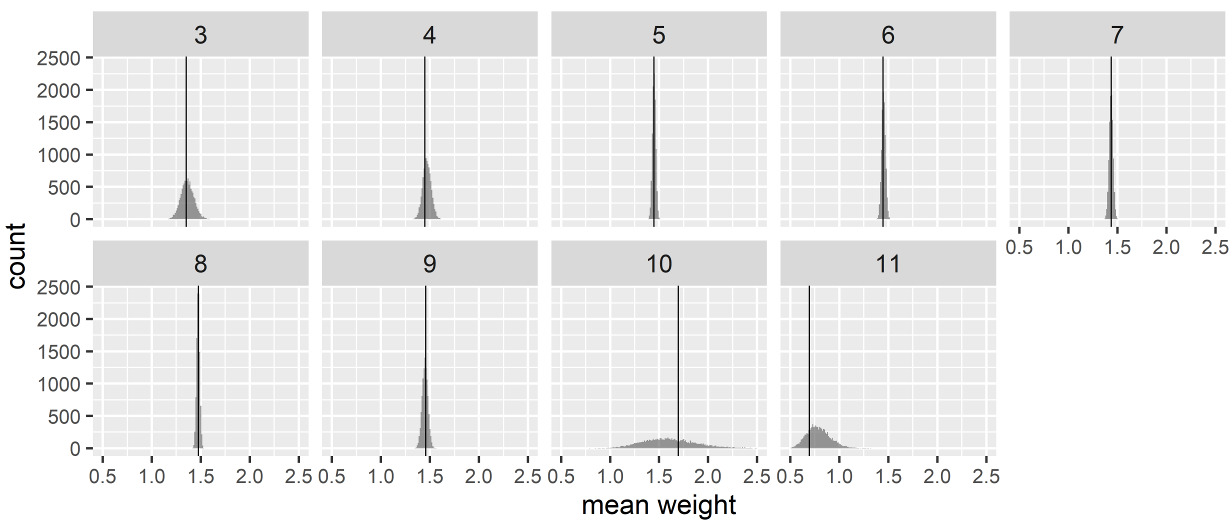


Figure S39. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by month (histograms) for quillback rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

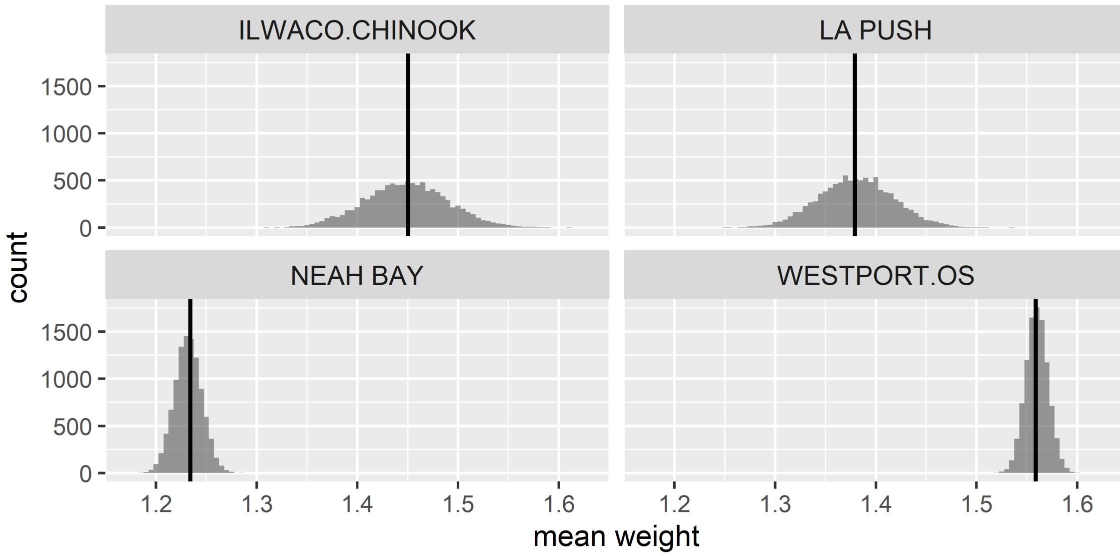


Figure S40. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by port group (histograms) for quillback rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

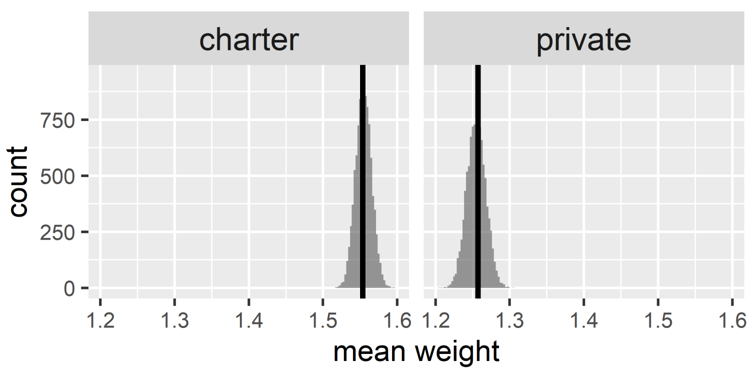


Figure S41. Posterior predictive distributions (PPD) of the marginal mean weight [kg] by boat mode (light) for quillback rockfish, relative to the marginal mean of the data (vertical line). PPDs are based on 10,000 simulated data sets from model 13.

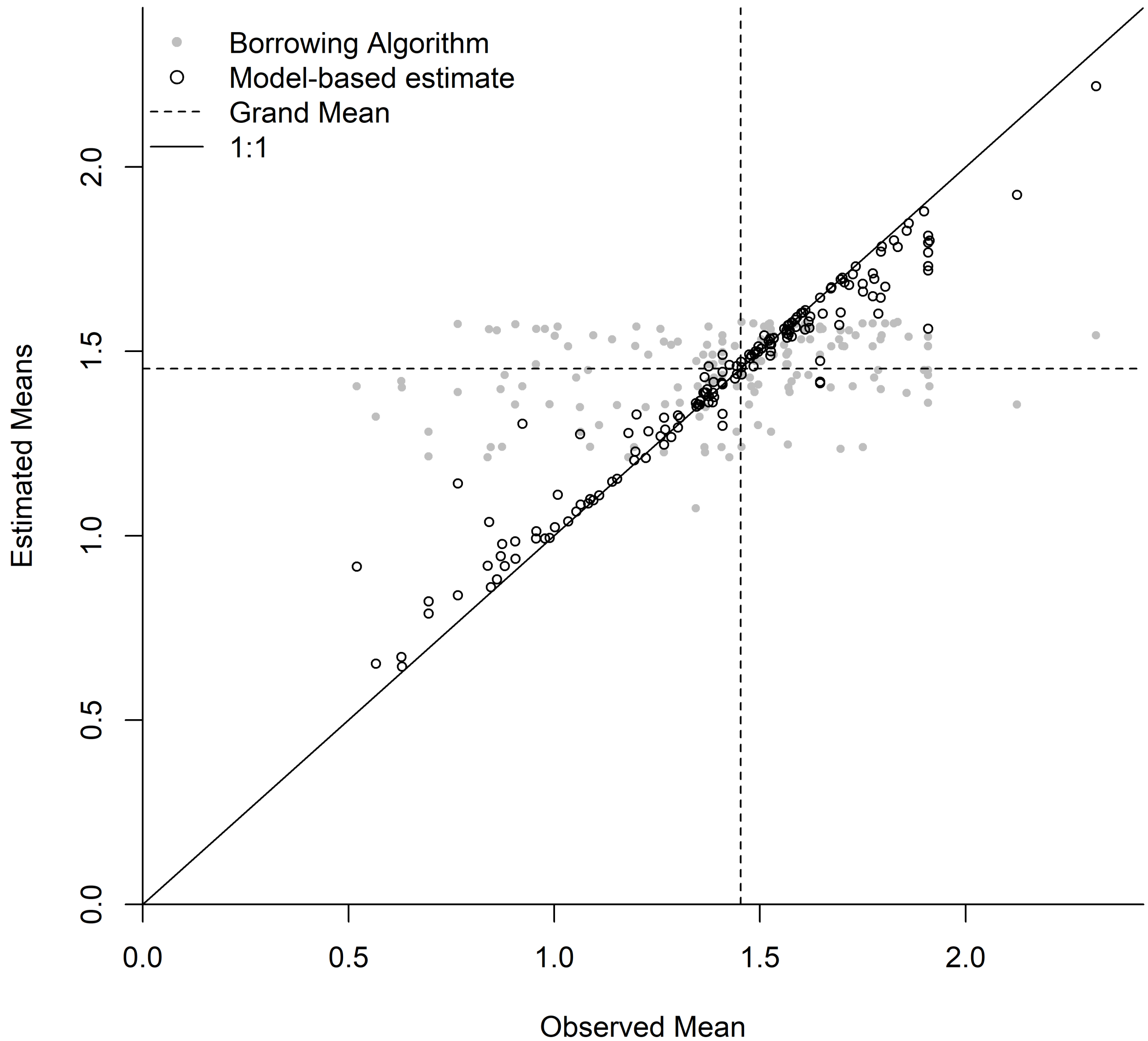


Figure S42. Estimated posterior means [kg] from model 13 (black open circles) and estimates from the borrowing algorithm (grey solid circles) versus observed arithmetic means for quillback rockfish. Each point represents a single year/month/port/mode combination (159 observed strata). Horizontal and vertical dashed lines are the arithmetic mean weight across all observations (grand mean). Solid line is the line of equality (1:1).

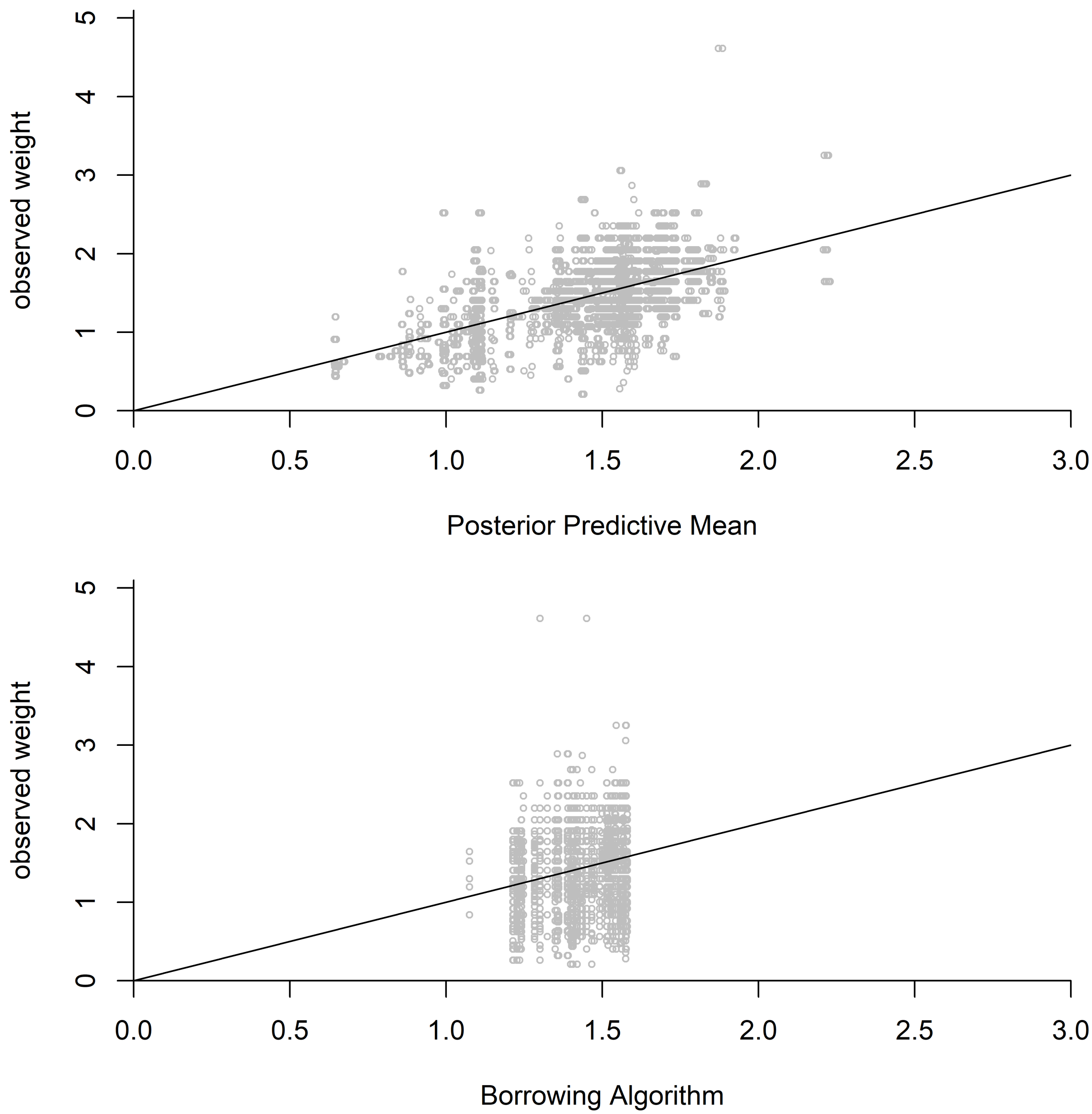


Figure S43. Comparison of observed weights [kg] for quillback rockfish to predicted mean weights from model 13 (upper panel) and estimates from the borrowing algorithm (lower panel). Solid black lines are 1:1 for reference.