Bayesian Inference on Canadian Seafood Mislabelling

Fynn De Vuono-Fraser

STAT4600 Final Report

Introduction

Seafood fraud is the act of seafood producers misinforming consumers and government to increase profit (Oceana). A major practice of seafood fraud is the act of mislabelling products, usually substituting a cheaper product for a more expensive and desirable one. Seafood mislabelling has numerous negative consequences, such as health risks to consumers. Escolar (*Lepidocybium flavobrunneum*) is a common substitute for butterfish and white tuna, is toxic and can cause gastrointestinal symptoms; as a result it is banned in both Korea, and Italy (Oceana, 2018). Seafood mislabelling also allows illegally or unsustainably caught products to be sold on the market, thereby promoting activities such as illegal, unreported, and unregulated (IUU) fishing (Oceana, 2018).

Seafood mislabelling is a widespread global issue. Pardo *et al.* (2016) performed a meta-analysis of 51 studies that covered all continents that used DNA barcoding (Hebert *et al.,* 2003) to identify seafood mislabelling rates. The authors reported an average mislabelling rate of 30% across 4500 samples. This demonstrates the potential prevalence of mislabelling, and how the negative impacts of mislabelling can be widespread. Canada is not exempt from seafood fraud and is well above the global average. A 2018 report from non-profit ocean conservation organization Oceana shows with nearly 400 samples from five Canadian cities that the national mislabelling rate is 44%, with Toronto having the highest mislabelling average at 59% with standard errors 0.025 and 0.05, respectively. Canada has a unique problem where the labelling policies are very relaxed. For example, Canadian guidelines allow over 100 species to be labled rockfish, which includes endangered species (Foster, 2017), such as Bocaccio (*S. paucispinis)* (NOAA).

Knowledge of mislabelling has been present since 1915 when New York diners were selling shark labelled as swordfish (Luque and Donlan, 2019). However, research on the topic has

increased dramatically in the last two decades due to the discovery of a new technology capable of determining the genetic origin of a biological sample. DNA barcoding uses a short section of DNA from a standardized gene region obtained from an unknown sample and compares it to a well curated reference library (The Barcode of Life Data System (BOLD); Ratnasingham and Hebert, 2007). DNA barcoding fundamentally changed species identification from traditional taxonomic methods which were unreliable, inaccurate, and inefficient. Since the inception of DNA barcoding, publications related to seafood mislabelling have increased by 1600% (Google Scholar). Most research has been focused on identifying the presence of seafood mislabelling, and few papers have examined potential explanatory factors. Being able to identify the potential indicators of seafood mislabelling would allow for accurate authenticity testing and traceability, leading to eventual change in government policy and legislation.

Academic Context

In 2016, a controversial paper was published that attempted to look deeper into the global scale of seafood mislabelling. Stawitz *et al.* (2016) used generalized linear regression models to predict probabilities of mislabelling using three covariates: Genus of product, purchase source, and country of purchase. Fixed effects were considered, but potentially important interactions were not. Logistic models were fit using the GAMLSS R package, and the best model was selected using the Akaike Information Criterion (AIC). The paper also examined three summary statistics: mislabelling proportion, percentage change in ex-vessel price (Melnychuk *et al.* 2016), of labelled species and real species, and the difference in conservation status between labelled and real species. To determine conservation status, the authors used the International Union for

Conservation of Nature (IUCN) Red List, a standardized categorization of species' conservation status.

There were some major limitations to this study. Firstly, the use of the IUCN Red List status was very crude. Stawitz *et al.* (2016) failed to account for farmed or wild caught seafood, as the conservation status of a farmed fish has little meaning. Also, for products that could be identified to a genus or species level, which accounted for 32.52% of samples, the authors used the average status of the respective species, which is entirely inaccurate, due to varying conservation status at the Genus taxonomy level. This led the researchers to conclude that a label can often be of more sever conservation status than the true species, which was contradictory to most other research. The authors argued that if less endangered species are being used to substitute more endangered species, then seafood mislabelling might actually be increasing seafood sustainability. These unjust conclusions prompted three response papers (Donlan *et al.* 2017, Mariani *et al*, 2017, and Warner *et al,* 2017) to point out flaws in Stawitz *et al.'s* (2016) study. While their conclusions were unwarranted, Stawitz *et al.* (2016) did show the situation is not as simple as once expected and demonstrated the necessity for further research.

There is limited research in indicators for mislabelling, with most statistical analysis ending at one-way or two-way chi-square tests. The papers that do attempt to predict mislabelling are inconsistent with methodology. Kim *et al.* (2016) constructed a probit regression model with maximum likelihood estimators (MLE) and weighted MLE estimation using data from a 2012 Oceana report in the United States (Oceana, 2012). Vega *et al.* (2022) employed a log-transformed linear regression model to predict mislabelling rates in Mexico. The most interesting method, however, was included in a study by Donlan and Luque (2019), which used a Bayesian logistic meta-analysis. Using two-level hierarchical models, the authors considered seven factors: study

type, seafood type, supply chain location, product form, country, product family, and product. Bayesian inference is synergetic with hierarchical models because it can provide different coefficients for variables on a lower level depending on the values of the upper level's indicators. On the topic of seafood mislabelling, it is intuitive to think it could follow a hierarchical model. For example, different species follow unique farming seasons, which could affect factors such as price, stock, or mislabelling rate. A temporal coefficient might then depend on the value of species.

Methods

Given our response is binary (1 if mislabelled, 0 otherwise), our models would follow logistic regression. The general regression equation for a General Linear Logistic Model is $\log(odds) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i$, where p is the probability of success, which in this context is mislabelling rate. β_0 being the intercept, β_i are the regression coefficients (*i=1…n, with n being the number of predictors),* and x_i are the predictors. We developed models using typical General Linear Model (GLM) approaches, then conducted Bayesian inference. Bayesian inference treats coefficients as random variables, which can optimize them. These random variable distributions, called priors, inform the final model, called the posterior, and are described by Bayes' Rule: Posterior $\propto$ Prior $\times$ Likelihood, where Likelihood is the Maximum Likelihood Estimator (MLE). In order to generate strong posterior coefficients, Bayesian inference makes use of Markov Chain Monte Carlo (MCMC), where posterior estimate is recursively updated. The standard number of iterations per chain is 2000. Usually, a set number of iterations are allocated as a training phase, or burn-in, before averaging the estimates of each iteration to provide a posterior estimate of the model coefficients.

Data

Our goal is to use Bayesian inference to develop a logistic model to predict mislabelling rates of seafood in Ontario. To accomplish this, we obtained access to two datasets. The first comprised of work done over two years by Oceana. Data consisted of 472 samples from 6 cities across Canada, which indicated purchase location, city, year of study, label species, identified species from DNA barcoding, both Latin and common name, and whether or not the sample was mislabelled Being mislabelled classified as if the label was wrong, but generalized labels were deemed correctly labelled. For example, if a label was Atlantic cod, but the real species was pacific cod, it would be considered mislabelled. However, if a label was simply cod, then if the real species was pacific cod, it would be considered correctly labelled. The year of study was 2018 for the cities Victoria, Vancouver, Toronto, Ottawa, and Halifax, and 2019 for Montreal. The second dataset was of sushi samples collected in Ontario from the lab of Dr. Robert Hanner (Department of Integrative Biology, University of Guelph), who has done extensive research on seafood mislabelling rates in North America (Wong and Hanner, 2008). This dataset included 196 samples of sushi across Ontario that noted the follow factors: three process categories (cut/roll, raw/cooked, mixed/single), city, date sampled, price, vendor type (restaurant/grocery), and sushi type (sashimi, nigiri, maki, mixed). All variables were categorical except for price. The three process variables other than sushi type were coded as binary. Each had three categories (for example, process type 2 had categories raw, cooked, or both). This resulted due to the researchers' discretion to decide how to categorize a dish from a restaurant, which could include both raw and cooked components. Given there were an insignificant number of this 'both' category, in each factor it was decided this variable would be removed from the analysis. Given there is theoretical reason to consider process

type (the more processed a sample is, the harder it is to identify to the species level), it should be clear what the process level a dish is in analysis. Source type was also coded as a binary response, 1, from a restaurant, 0, otherwise. There were two observations with no outcome here, and these were omitted from analysis. Expected fish (i.e. the species on the menu, and what the product was sold as) was aggregated to create fewer groups with larger sample sizes. The main example of this was snapper. There were four observations labelled red snapper, and one observation labelled yellowtail snapper. Given there were 12 observations simply labelled snapper, we aggregated the red and yellowtail observations under the snapper label. A similar situation occurred with tuna. 9 observations were labelled red tuna, 14 were white tuna, and 62 were simply tuna. These were all aggregated under the label tuna. Date was converted to two variables: month and season. These give temporal indicators and different levels of generalization Winter was considered as December, January, February, and the rest of the seasons followed with three months each.

There was one observation without a price, which was missing completely at random due to misplacement of the submission form in the lab. This was the only sample from Windsor, and was a mixed roll labelled white and red tuna, sampled August 26th, 2019, from a restaurant. It was also the only sample that was domestic product. We decided to conduct a linear regression model to best predict price and use the response as an estimate for this observation. Using forward stepping variable selection, we identified the optimal model to predict price was $Price = Beta_0 + isCut + isRaw + isRestaurant$

We had decided that the data from the Guelph lab would provide a better model given the number of potential indicators. There was thought to use the data from Oceana to inform prior distributions of our model for city or vendor type, but a test of two proportions showed that the two datasets had significantly different proportions for these categories. Given the Guelph data

only took sushi samples, and the Oceana data took more general samples, it was not a similar enough data to inform priors.

Analysis

We started collating summary statistics of potential indicators, commenting on mislabelling proportion, number of samples, and standard error. We eliminated any variables that included categories with a mislabelling proportion of 0 or 1. These variables overfit the data and included extremely large standard errors when included in a logistic model. We also looked at Fishers' Exact Tests to determine if proportions between categories were significant or not. We used Fishers' Exact Test because there were not at least 5 observations per category, which is an assumption required to conduct Chi-Square test. Given a lot of the contingency tables are large, computing Fishers' P-values using the standard method would be extremely computationally expensive, so we used the simulate.p.values=TRUE setting, which computes p-values through a Monte Carlo simulation. The next step was fitting univariate models of each indicator, using the glm() function with the binomial family, given the response was binary (1 is mislabelled, 0 otherwise). Given the glm() function does not provide a model p-value, we used a drop in deviance test using the anova() function and Chi square test against a null model coded as isMislabelled ~ 1. This model includes an intercept with no variables, and a drop in deviance test against this indicates if a model can predict the response significantly better than an uninformed, constant estimate.

The next step was considering multivariate models. We used Akaike's Information Criterion to compare models, due to the fact it considers goodness of fit and complexity. We

performed forward stepping variable selection starting with univariate models and moving forward. This was finalized with a drop in deviance test against the optimal less complex model. After finding a theoretical best model with forward stepping, we summarized models with the lowest AIC scores to identify any other models of interest with similar significance. We tested for multicollinearity using generalized variance inflation factors between predictors. This was done using the vif() R function under the car package. VIF is calculated as $\frac{1}{1-R_i^2}$, where R_i^2 is the R^2 value of a model of all other parameters, with parameter i as the response. The better the other parameters can predict parameter i, the higher the collinearity is, and the larger R_i^2 will be, which will in turn increase VIF. The closer to 1 the VIF is, the less collinearity is present in the model. The test provided in R under the car package comments on the Generalized Variance Inflation Factor (GVIF) which is equal to $VIF^{\frac{1}{2*df}}$ once a categorical indicator is included in the model. Once categorical variables are considered, the VIF can often give a higher value even when collinearity is not present. The GVIF accounts for this and adjusts for the degrees of freedom. The interpretation of GVIF is similar to VIF, in the closer it is to 1 the less collinearity, however, the value to interpret is $GVIF^2$. If $GVIF^2$ is below 5, there is likely no collinearity, and the closer to 1 the better.

Once we confirmed our models of interest, we considered two-way interaction terms. We did not consider three-way interaction terms or greater because of interpretability and to avoid overfitting the model. After fitting a model with an interaction term, a drop in deviance test was conducted with the original model to test if the added complexity significantly predicted the response better. Any model with interaction of significance was considered.

Once we had finalized a set of models of interest, we implemented Bayesian inference. The R package Rstan was used to run Hamiltonian Monte Carlo (HMC) via the Stan probabilistic programming language. We stated the Rstan standard settings for the chaining process: four chains with 2000 iterations per chain, and 1000 iterations per chain for burn-in/warmup. This provided a standardized process for each model. We used normal distributions for our priors. Prior means for all regression coefficients were centered on their MLEs. The mean was informed as the coefficient estimate, and the standard deviation was the coefficient standard error multiplied by the square root of the number of observations. A normal distribution is intuitive for our priors because we can derive the hyperparameters (mu, sigma) from the MLEs. We also fit our models using uniform priors with a mean of 0, and a standard deviation of 1 to compare accuracy of coefficients. After fitting the models, Rstan outputs the following posterior mean quantities: the coefficient, standard errors, standard deviation, credible intervals, effective number of independent samples in the chain, and Rhat. Credible intervals are similar to frequentist confidence intervals where a range of plausible estimates for the population parameter is given. However, credible intervals allow probabilistic statements about statistical parameters to be made and depend on the priors, which mean a difference in prior will result in a difference in the credible interval. A credible at significance level $(1 - \alpha)\%$ contains $(1 - \alpha)\%$ of the samples drawn in the chaining process. The effective number of independent samples is a measure of collinearity. Collinearity is a measure of the correlation of samples in the chain, and large collinearity means each sample provides little additional information to estimating the posteriors. The larger the effective independent sample size, the less correlation in the chains and the more informed the posterior. Rhat indicates the convergence of a chain and will be very close to 1 if the chain converged effectively to the posterior

distribution. A credible interval follows similar evaluation logic as a confidence interval for coefficients, where if the coefficient is significant, its credible interval would not contain 0.

Results
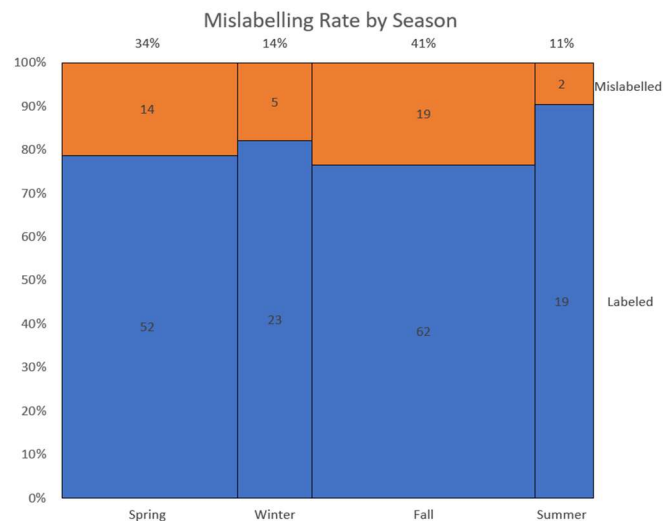
Figure 1: Mosaic plot of mislabelling rates by Season
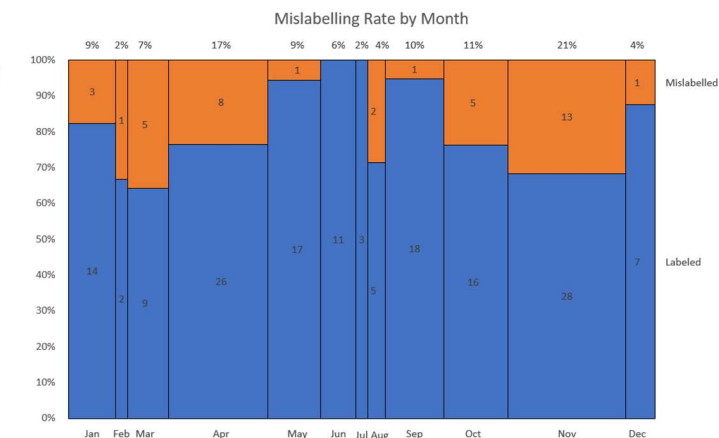
Figure 2: Mosaic plot of mislabelling rates by month

From collating mislabelling proportions, we could find some initial promising patterns. Firstly, we considered month. This mosaic plot shows the mislabelling proportions of each month, and visually there does seem to be a pattern in the months, with mislabelling reaching low averages in the summer. There are some data deficient months, such as July and February. When we plot the seasons instead of months, the significance of this pattern decreases. The width of each bar is proportional to the categories size, and the division of the bar shows the mislabelling proportion.

We can summarize the findings of the Fisher's Exact Test here. As shown one parameter stands out as having significantly different proportions between categories, being the binary

predictos isRestaurant. The only other significant variable at 0.05 significance level is isCut, another binary variable.

It was found that label species overfitted the data tremendously. Every species category except for tuna had a mislabelling proportion of 0 or 1, and when we separated Tuna into its more specific categories (white, red, unknown), each had a mislabelling proportion of 0 or 1. This overfitting meant that labelled species could not be a included in the final model. While this agreed with the Oceana dataset, which found that snapper and butterfish were mislabelled 100% of the time, it is unfortunate that we could not include this in our model, as it was a variable of interest. We ran into a similar issue with IUCN status. If these parameters were included in a model, the coefficients and standard errors would be massive, and would make more useful parameters hard to interpret.

INDICATOR	FISHER'S TEST
City	0.3673
IsRestaurant	0.00001206
Locality	0.5492
Month	0.0559
Season	0.4493
isCut	0.0456
isMixed	0.5875
isRaw	0.1196
sushiType	0.09145

Figure 3: Variables' Fisher's Exact Test

Categories	Labeled	Mislabelled	MislabelledProp	SE	SampleProp
NA	39	9	0.188	0.056	0.245
EN	1	0	0.000	0.000	0.005
LC	85	29	0.254	0.041	0.582
NR	7	0	0.000	0.000	0.036
NT	6	0	0.000	0.000	0.031
VU	18	2	0.100	0.067	0.102

Figure 4: Mislabelling proportions of IUCN status

Category	Labeled	Mislabelled	MislabelledProp	SE	SampleProp
Artificial Crab	1	0	0.000	0.000	0.005
Butterfish	0	9	1.000	0.000	0.046
Clam	5	0	0.000	0.000	0.026
Fish Egg	1	0	0.000	0.000	0.005
Mackerel	2	0	0.000	0.000	0.010
Octopus	2	0	0.000	0.000	0.010
Salmon	11	0	0.000	0.000	0.056
Shrimp	40	0	0.000	0.000	0.204
Snapper	0	17	1.000	0.000	0.087
Tilapia	1	0	0.000	0.000	0.005
Tuna	75	14	0.157	0.039	0.454
Unknown	18	0	0.000	0.000	0.092

Figure 5: Mislabelling proportions of labelled fish

The categories Near Threatened (NT), Not Recognized (NR), and Endangered (EN) have mislabelling proportions of 0. While endangered could be omitted due to the fact it only has one

observation, NT and NR consist of 13 observations, which is not an insignificant portion of the data. We decided to omit IUCN as a potential indicator, not only because of the overfitting, but also because issues with the IUCN as an indicator brought up by response papers of Stawitz *et al.* (2016). Given we cannot verify the exact procedure used to determine a samples IUCN status, we could confirm we would not have the same issues with IUCN status than Stawitz *et al* (2016).

PARAMETER	P-VALUE	AIC
City	0.4006	213.62
isRestaurant	3.01E-06	171.87
Locality	0.6811	194.91
Price	0.0006	181.15
Month	0.0283	192.16
Season	0.3702	194.54
isCut	0.0319	189.08
isMixed	0.5614	193.34
isRaw	0.0218	188.42
sushiType	0.0531	189.34

Figure 6: Univariate model evaluation

The univariate models give us a look into potentially significant predictors. We can see isRestaurant is the most significant by a large margin. There are four other variables with a p-value below 0.05: isRaw (1 if raw, 0 otherwise), isCut (1 if cut, 0 if whole), month and price. We found high correlation between our temporal components and city. We believe the research team that collated the data only visited each city a small number of times and collected their samples in succession. This meant we could not incorporate city and a temporal component in our model. Given the city predictor had a significantly larger AIC than any other predictor, we felt it was reasonable to assume that all cities in Southern Ontario had insignificantly different mislabelling proportions.

Comparator: null, AIC = 192

Variable	AIC	DeltaAIC	Deviance	Significance
isRest	171	-21	3.01E-06	***
locality	195	3	0.6811	
price	181	-11	0.0006	***
month	192	0	0.0283	*
season	194	2	0.3702	
cut	189	-3	0.0319	*
mixed	193	1	0.5614	
raw	188	-4	0.0218	*
sushiType	189	-3	0.0448	*

Comparator: isRest, AIC = 171

Variables	AIC	DeltaAIC	Deviance	Significance
isRest, locality	175	4	0.8007	
isRest, price	168	-3	0.0229	*
isRest, month	177	6	0.1393	
isRest, season	174	3	0.3337	
isRest, cut	171	0	0.1588	
isRest, mixed	173	2	0.9519	
isRest, raw	172	1	0.2869	
isRest, sushiType	176	5	0.49	

Comparator: isRest + Price, AIC = 168

Variables	AIC	DeltaAIC	Deviance	Significance
price, isrest, locality	172	4	0.7507	
price, isRest, month	176	8	0.212	
price, isRest, season	171	3	0.3002	
price, isRest, isCut	170	2	0.5841	
price, isRest, isMixed	170	2	0.8423	
price, isRest, isRaw	169	1	0.3646	
price, isRest, sushiType	173	5	0.603	

Figure 7: Forward stepping model building with deviance values

Our forward stepping variable selection concluded at a bivariate model as the most optimal, being $isRestaurant + Price$, given no more complex models provided a significant p-value at 0.05. However, we looked at other variables with AICs similar to our model. This figure shows the 10 models with the lowest AICs. A model that piqued our interest was $isRest + Price + Season$. This was because a goal in our model was to include a temporal component. This model was also selected along with our optimal model, in order to consider temporality. A temporal component could reveal patterns in the fishing industry that are highly explanatory of mislabelling rates.

Model	AIC
isMislabelled ~ Price + isRest	168.699
isMislabelled ~ Price + isRaw + isRest	169.8772
isMislabelled ~ Price + isCut + isRest	170.3993
isMislabelled ~ Price + isMixed + isRest	170.6594
isMislabelled ~ Price + theSeason + isRest	171.036
isMislabelled ~ Price + isCut + isRaw + isRest	171.6029
isMislabelled ~ Price + isMixed + isRaw + isRest	171.8335
isMislabelled ~ isRest	171.8711
isMislabelled ~ isCut + isRest	171.8851
isMislabelled ~ Price + Locality + isRest	172.1256
isMislabelled ~ Price + theSeason + isRaw + isRest	172.1674

Figure 8: Top 10 models with lowest AIC

At this point collinearity was considered. For our optimal model, we calculated the Variance Inflation Factor (VIF) for price and isRest. For our optimal model, the VIF of our predictors was 1.015, which indicates no evidence of collinearity. For our three-predictor multivariate model, the values of $GVIF^2$ were all less than 1.21, indicating no collinearity.

The different interaction deviance tests for each model are shown. Note that none are significant at 0.05 significance level. However, the interaction of price and isRest was considered, because of theoretical evidnece. The same price of an item at a grocery store and restaurant would result in a different product, due to cost of preparation, ingredients, and scale of the business. We felt this interaction would consider this.

MODEL	INTERACTION	DEVIANCE
price+isRest	price:isRest	0.068
price+isRest+season	price:isRest	0.068
price+isRest+season	price:season	0.4398
price+isRest+season	isRest:season	0.5571

Figure 9: Drop in deviance test of interactions

The last step was to fit the models in Rstan. The four models that were fit were $isRest + Price + Season$ and $isRestaurant + Price,$ each with and without the Price:isRest interaction term. We can look at Rstan output to determine the validity of each model. A rule of thumb is

that the number of effective independent samples (NEFF) should be over 1000 for standard chain settings, and Rhat should be below 1.05. We were immediately able to omit the interaction models from our consideration, as their average NEFF was well below 1000. In fact, the only model without a single NEFF below 1000 was the full model with no interaction. We also compared trace plots for each coefficient for each model. A trace plot traces the estimates of each iteration and give a visual indication of convergence.
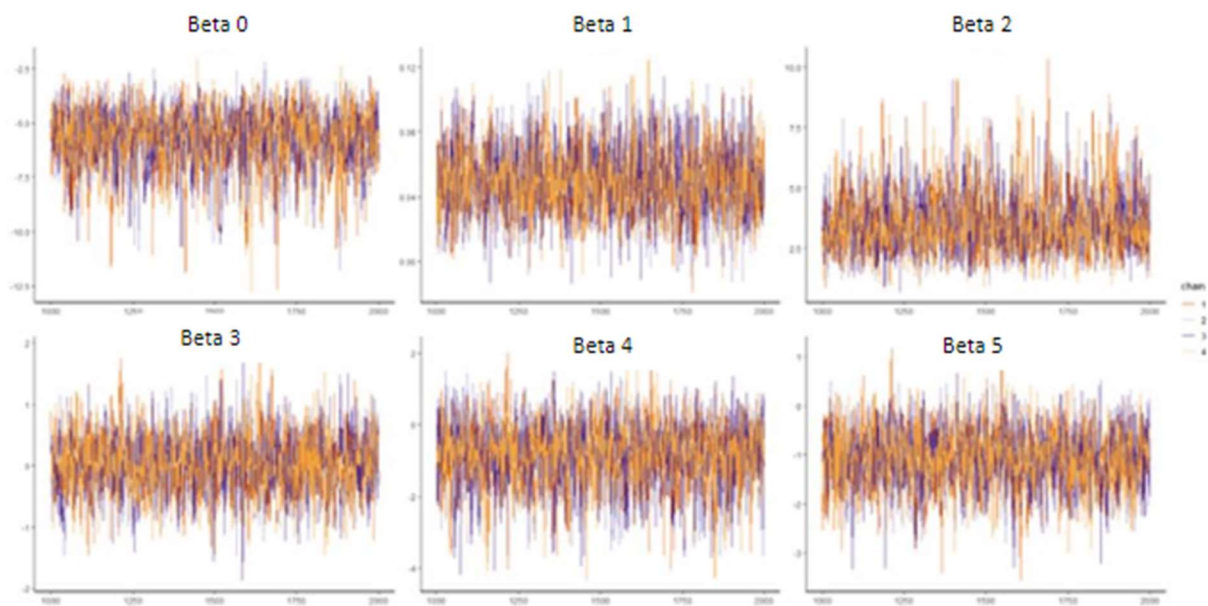
Figure 10: Trace plots for full model without interactions

Nicknamed caterpillar plots due to their appearance, the more full and 'fuzzy' a plot looks shows two things: convergence on an average, and effective sampling, by having large changes in estimates, per iteration. Above are the trace plots for our full model without interactions, and below

are the trace plots for our optimal model with interactions.
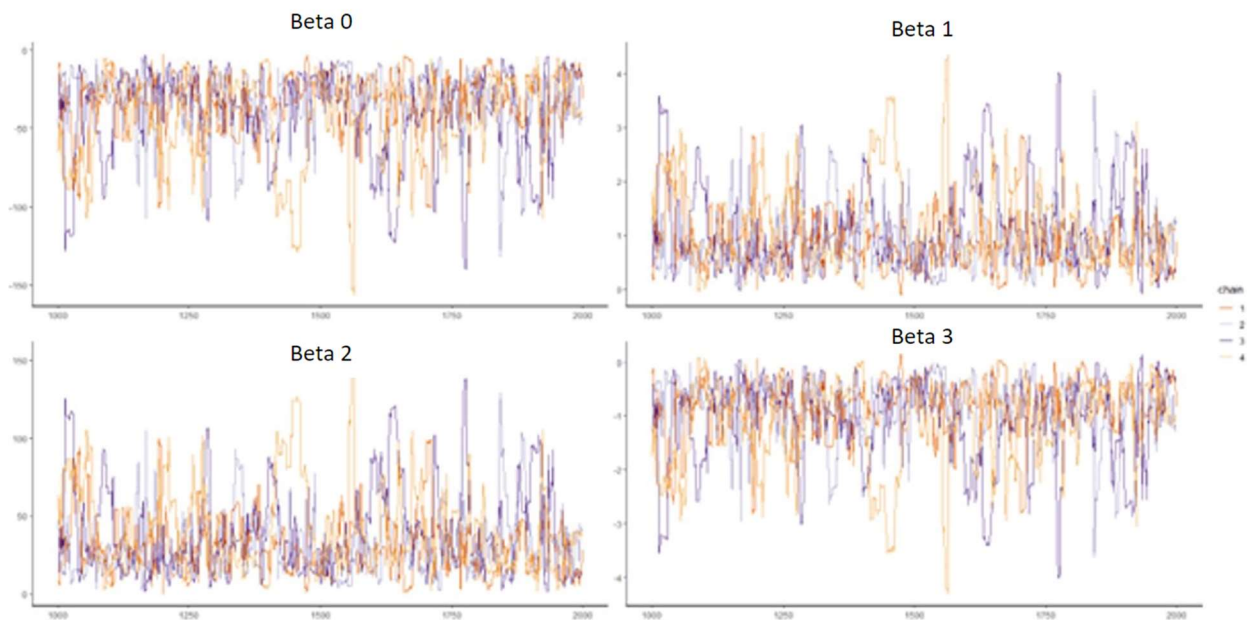
Figure 11: Trace plots for optimal model with interactions

There is an obvious visual difference, and the first plots show better mixing, and better convergence. Another thing to note that the second plots do not seem to follow a normal distribution, with estimates seemingly bounded either above or below 0. This is a clear divergence of the informed priors which were stated as normal.

Coefficient	Mean	SE	CI Lowerbound	CI Upperbound	NEFF
Beta 0	-5.78	0.04	-9.00	-3.75	1561
Beta 1	0.05	0.00	0.01	0.09	2300
Beta 2	3.59	0.03	1.61	6.63	1730
Beta 3	0.03	0.01	-0.92	0.98	2287
Beta 4	-0.84	0.02	-2.87	0.82	2281
Beta 5	-1.06	0.01	-2.29	0.07	2992

Figure 12: Rstan output on favoured model

This figure shows the favored model's ($isRest + Price + Season$) Rstan output. Beta[1] is the intercept, beta[2] is price, beta[3] is isRest, and beta[4-6] represent the factors of season. Each 95% credible interval indicates significance. The credible intervals for beta[4-6] cover 0, but

since they are categorical factors, since the intercept is significant this does not show insignificance.

The comparison made was MLE informed priors against uniform priors. Instead of priors having mean and standard deviation informed from the MLEs, they were set as ~N(0,1). This was to test the validity of the MLEs. What we found was the uninformed models had much larger NEFFs, and smaller standard errors across all four models. This figure shows the trace plots of the optimal model with interaction with uninformed priors, which compared to the informed trace plots is a significant improvement. This is most likely due the small dataset, which created inadequate MLEs.
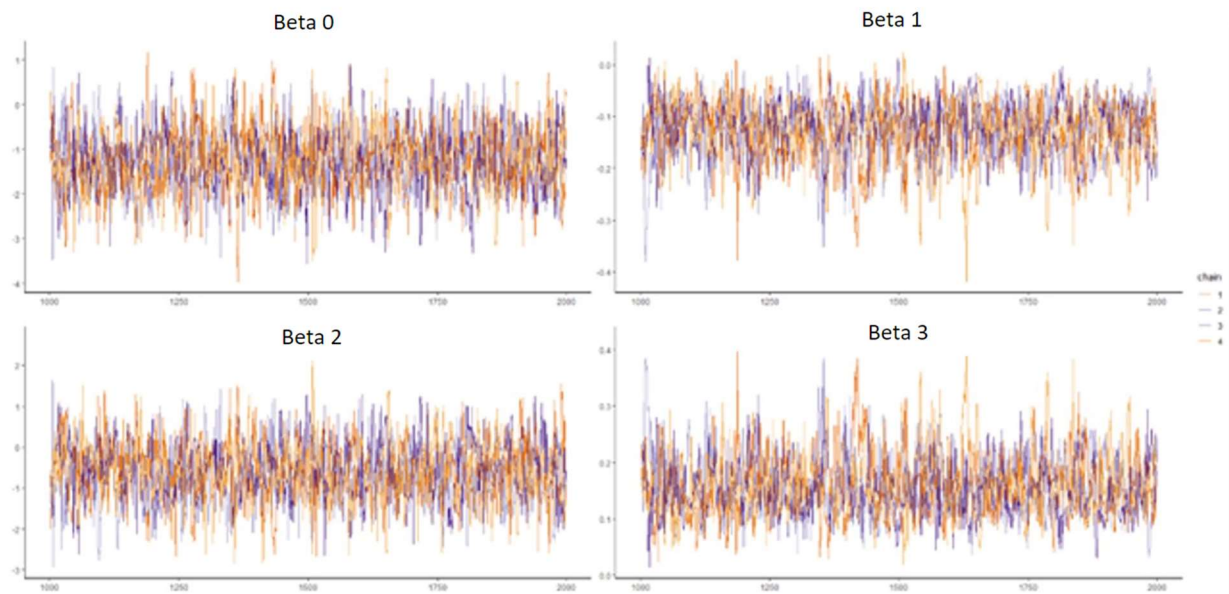
Figure 13: Trace plots for optimal model with interaction with uniform priors

Discussion

We were able to identify the presence of a relationship between mislabelling proportions and multiple indicators, but due to incomplete data we were limited in our findings. A larger dataset would allow for a few major changes in the analysis. First, it would hopefully provide less certainty in labelled fish. This would allow for labelled fish to be an indicator in the model. This would also allow for a hierarchical model with labelled fish to be the first level. We think this is promising especially for a temporal component, due to different fishing seasons for different species. Other than this however, there are many other plausible hierarchical models with different indicators, such as vendor type, or a geographical component. We also believed there was not enough data across all seasons let alone months to provide a robust temporal component.

Another future addition to the analysis is to consider and amend the priors of Bayesian model. Research suggest a Cauchy-distribution for logistic regression: specifically a Cauchy(0, 10) distribution for intercept, and a Cauchy(0, 2.5) on the remaining coefficients (Gelman, 2008). This work only looks at sushi data, which is a very specific subsample of the seafood market and could follow different mislabelling behaviours than the bulk of seafood. Future work will have to be done to consider other sources of seafood.

To conclude, we have shown that an increase in price is associated with an increased odds of a sushi sample being mislabelled, and that restaurants have a higher mislabelling proportion than grocery stores. We have also identified the possibility of a potentially significant temporal component. These are all integral components in a final model to accurately predict seafood mislabelling.

Appendix

Works Cited:

Oceana. (n.d.). *What is Seafood Fraud?*

https://oceana.org/what-seafoodfraud/#:~:text=Take%20Action-,What%20is%20Seafood%20Fraud%3F,conservation%20efforts%20and%20human%20health

Levin, J. (2018). Seafood fraud and mislabelling across Canada. Oceana, Washington, DC.Pardo, M. Á., Jiménez, E., & Pérez-Villarreal, B. (2016). Misdescription incidents in seafood sector. *Food Control*, *62*, 277-283.

Hebert, P. D., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313-321.

Sarah Foster, (2017), *Are you eating seafood in the dark?*. SeaChoice https://www.seachoice.org/are-you-eating-seafood-in-the-dark/

Fisheries, N. O. A. A. (n.d.). *Ecosystem science for endangered and threatened rockfish in Puget sound*. NOAA. Retrieved April 10, 2023, from https://www.fisheries.noaa.gov/west-coast/science-data/ecosystem-science-endangered-and-threatened-rockfish-puget-sound

Luque, G. M., & Donlan, C. J. (2019). The characterization of seafood mislabeling: A global meta-analysis. *Biological Conservation*, *236*, 556-570.

Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). *Molecular ecology notes*, *7*(3), 355-364.

Stawitz, C. C., Siple, M. C., Munsch, S. H., Lee, Q., & Derby, S. R. (2017). Financial and ecological implications of global seafood mislabeling. Conservation Letters, 10(6), 681-689

Donlan, C. J., Luque, G. M., Wilcox, C., Gelcich, S., Koch, G. W., & Hungate, B. A. (2017). Research on seafood fraud deserves better.

Mariani, S., Cawthorn, D. M., & Hanner, R. (2017). Mislabeling seafood does not promote sustainability: A comment on Stawitz et al.(2016). Conservation Letters, 10(6), 781-782.

Warner, K. A., Miller, D. D., & Naaum, A. M. (2017). Unsupported conclusions on net conservation benefits of mislabeling seafood. Conservation Letters, 10(6), 786-787.

Wong, E. H. K., & Hanner, R. H. (2008). DNA barcoding detects market substitution in North American seafood. Food Research International, 41(8), 828-837.

Pardo, Miguel Ángel, Elisa Jiménez, and Begoña Pérez-Villarreal. "Misdescription incidents in seafood sector." Food Control 62 (2016): 277-283.

The IUCN Red List of Threatened Species. IUCN Red List of Threatened Species. (n.d.). Retrieved January 30, 2023, from https://www.iucnredlist.org/

Munguia-Vega, A., Terrazas-Tapia, R., Dominguez-Contreras, J. F., Reyna-Fabian, M., & Zapata-Morales, P. (2022). DNA barcoding reveals global and local influences on patterns of mislabeling and substitution in the trade of fish in Mexico. *Plos one*, *17*(4), e0265960.

Do, T. D., Choi, T. J., Kim, J. I., An, H. E., Park, Y. J., Karagozlu, M. Z., & Kim, C. B. (2019). Assessment of marine fish mislabeling in South Korea's markets by DNA barcoding. *Food Control*, *100*, 53-57.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2013). *Bayesian Data Analysis, third edition, 3rd Edition*. CRC Press.

Comprehensive R Archive Network (CRAN). (2023, January 27). *Generalised additive models for location scale and shape [R package gamlss version 5.4-12]*. The Comprehensive R Archive Network. Retrieved April 10, 2023, from https://cran.r-project.org/web/packages/gamlss/index.html

Comprehensive R Archive Network (CRAN). (2023, March 30). *Companion to applied regression [R package car version 3.1-2]*. The Comprehensive R Archive Network. Retrieved April 10, 2023, from https://cran.r-project.org/web/packages/car/index.html

Team, S. D. (2023, January 16). *RStan: the R interface to Stan*. RStan: The R interface to Stan. Retrieved April 10, 2023, from https://cran.r-project.org/web/packages/rstan/vignettes/rstan.html

Fisher's Exact Test for Count Data. R: Fisher's exact test for count data. (n.d.). Retrieved April 10, 2023, from https://stat.ethz.ch/R-manual/R-devel/library/stats/html/fisher.test.html

MRC CBU Wiki. FAQ/Collinearity - CBU statistics Wiki. (n.d.). Retrieved April 10, 2023, from https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/Collinearity#:~:text=More%20generally%20generalized%20variance%2Dinflation,in%20R%20(%20see%20here)

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models.

Additional Tables and Figures

Mislabelling proportions for variables: Guelph Data

```
SOURCE
              Labeled Mislabelled MislabelledProp         SE SampleProp
D                   1           1       0.5000000 0.35355339 0.01020408
I                  47          12       0.2033898 0.05240365 0.30102041
Missing form        2           0       0.0000000 0.00000000 0.01020408
U                 106          27       0.2030075 0.03487846 0.67857143


LOCATION
              Labeled Mislabelled MislabelledProp         SE SampleProp
G                  63           1        0.015625 0.01550245 0.32653061
Missing form        2           0        0.000000 0.00000000 0.01020408
R                  91          39        0.300000 0.04019185 0.66326531

SINGLE/MIXED
              Labeled Mislabelled MislabelledProp         SE  SampleProp
Missing form        1           0       0.0000000 0.00000000 0.005102041
Mixed              99          22       0.1818182 0.03506315 0.617346939
Single             56          18       0.2432432 0.04987497 0.377551020

CUT/ROLLS
              Labeled Mislabelled MislabelledProp         SE  SampleProp
 cuts              36          17       0.3207547 0.06411533 0.270408163
 rolls            112          23       0.1703704 0.03235730 0.688775510
Missing form        1           0       0.0000000 0.00000000 0.005102041
rolls/cuts          7           0       0.0000000 0.00000000 0.035714286

COOKED/RAW
              Labeled Mislabelled MislabelledProp         SE  SampleProp
 Cooked            10           0       0.0000000 0.00000000 0.051020408
 Raw              129          40       0.2366864 0.03269601 0.862244898
Missing form        1           0       0.0000000 0.00000000 0.005102041
Raw/cooked         16           0       0.0000000 0.00000000 0.081632653
IUCN

         Labeled   Mislabelled MislabelledProp   SE          SampleProp
NA       39        9            0.187500         0.05633674  0.244897959
EN       1         0            0.000000         0.00000000  0.005102041
LC       85        29           0.254386         0.04078976  0.581632653
NR       7         0            0.000000         0.00000000  0.035714286
NT       6         0            0.000000         0.00000000  0.030612245
VU       18        2            0.100000         0.06708204  0.102040816
```

CITY

	Labeled	Mislabelled	MislabelledProp	SE	SampleProp
BRAMPTON	12	1	0.07692308	0.07390530	0.066326531
BRANTFORD	6	0	0.00000000	0.00000000	0.030612245
BURLINGTON	7	3	0.30000000	0.14491377	0.051020408
CAMBRIDGE	6	1	0.14285714	0.13226001	0.035714286
CHATHAM	1	2	0.66666667	0.27216553	0.015306122
FERGUS	1	0	0.00000000	0.00000000	0.005102041
GEORGETOWN	1	0	0.00000000	0.00000000	0.005102041
GUELPH	33	7	0.17500000	0.06007807	0.204081633
HAMILTON	5	3	0.37500000	0.17116330	0.040816327
KITCHENER	5	1	0.16666667	0.15214515	0.030612245
MILTON	2	1	0.33333333	0.27216553	0.015306122
MISSISSAUGA	7	1	0.12500000	0.11692679	0.040816327
NORTH BAY	2	1	0.33333333	0.27216553	0.015306122
NORTH YORK	0	1	1.00000000	0.00000000	0.005102041
OSHAWA	1	0	0.00000000	0.00000000	0.005102041
OTTAWA	18	6	0.25000000	0.08838835	0.122448980
PORT PERRY	3	0	0.00000000	0.00000000	0.015306122
STONEY CREEK	3	1	0.25000000	0.21650635	0.020408163
TORONTO	16	2	0.11111111	0.07407407	0.091836735
TRENTON	1	0	0.00000000	0.00000000	0.005102041
WATERLOO	20	6	0.23076923	0.08262864	0.132653061
WHITBY	1	0	0.00000000	0.00000000	0.005102041
WINDSOR	1	1	0.50000000	0.35355339	0.010204082
WOODBRIDGE	0	1	1.00000000	0.00000000	0.005102041
WOODSTOCK	4	1	0.20000000	0.17888544	0.025510204

EXPECTED GEN

	Labeled	Mislabelled	MislabelledProp	SE	SampleProp
Artificial	1	0	0.0000000	0.0000000	0.005102041
Butterfish	0	9	1.0000000	0.0000000	0.045918367
Clam	5	0	0.0000000	0.0000000	0.025510204
Fish	1	0	0.0000000	0.0000000	0.005102041
Mackerel	2	0	0.0000000	0.0000000	0.010204082
Octopus	2	0	0.0000000	0.0000000	0.010204082
Salmon	11	0	0.0000000	0.0000000	0.056122449
Shrimp	40	0	0.0000000	0.0000000	0.204081633
Snapper	0	17	1.0000000	0.0000000	0.086734694
Tilapia	1	0	0.0000000	0.0000000	0.005102041
Tuna	75	14	0.1573034	0.0385931	0.454081633
Unknown	18	0	0.0000000	0.0000000	0.091836735

```
MONTH
        Labeled    Mislabelled  MislabelledProp        SE SampleProp
Apr     26         8                0.23529412 0.07274670 0.17346939
Aug     5          2                0.28571429 0.17074694 0.03571429
Dec     7          1                0.12500000 0.11692679 0.04081633
Feb     2          1                0.33333333 0.27216553 0.01530612
Jan     14         3                0.17647059 0.09245944 0.08673469
Jul     3          0                0.00000000 0.00000000 0.01530612
Jun     11         0                0.00000000 0.00000000 0.05612245
Mar     9          5                0.35714286 0.12806021 0.07142857
May     17         1                0.05555556 0.05399030 0.09183673
Nov     28         13               0.31707317 0.07267330 0.20918367
Oct     16         5                0.23809524 0.09294286 0.10714286
Sep     18         1                0.05263158 0.05122782 0.09693878
```

Mislabelling proportions for variables: Oceana Data

```
CITY

          Labeled Mislabelled MislabeledProp SampleProp
halifax        55          34      0.3820225 0.18855932
montreal       35          55      0.6111111 0.19067797
ottawa         50          48      0.4897959 0.20762712
toronto        39          57      0.5937500 0.20338983
vancouver      61          23      0.2738095 0.17796610
victoria        5          10      0.6666667 0.03177966

PURCHASE LOCATION

                Labeled Mislabelled MislabeledProp  SampleProp
Fishmarket          2           1       0.3333333 0.006355932
Gift store          0           1       1.0000000 0.002118644
Grocerystore       94          32       0.2539683 0.266949153
Restaurant        148         193       0.5659824 0.722457627
Food Truck          1           0       0.0000000 0.002118644
```

REAL FISH

	Labeled	Mislabelled	MislabeledProp	SampleProp
	0	1	1.00000000	0.002118644
Amberjack	1	28	0.96551724	0.061440678
Atlantic cod	1	0	0.00000000	0.002118644
Bass	4	4	0.50000000	0.016949153
Beeliner	0	3	1.00000000	0.006355932
Catfish	0	8	1.00000000	0.016949153
Char	5	0	0.00000000	0.010593220
Cobia	0	1	1.00000000	0.002118644
Cod	50	11	0.18032787	0.129237288
Crab	2	2	0.50000000	0.008474576
Crayfish	0	1	1.00000000	0.002118644
Crevette	0	1	1.00000000	0.002118644
Cuttlefish	1	0	0.00000000	0.002118644
Dab	1	0	0.00000000	0.002118644
Eel	2	0	0.00000000	0.004237288
elassodon	0	2	1.00000000	0.004237288
Escolar	0	24	1.00000000	0.050847458
flounder	2	2	0.50000000	0.008474576
grouper	0	3	1.00000000	0.006355932
Haddock	3	4	0.57142857	0.014830508
hake	0	3	1.00000000	0.006355932
Hake	1	0	0.00000000	0.002118644
Halibut	32	10	0.23809524	0.088983051
Jaune	0	2	1.00000000	0.004237288
jobfish	0	1	1.00000000	0.002118644
Lingcod	1	1	0.50000000	0.004237288
List	0	3	1.00000000	0.006355932
lobster	4	1	0.20000000	0.010593220
Mackerel	2	0	0.00000000	0.004237288
Monkfish	1	0	0.00000000	0.002118644
Nordique	1	0	0.00000000	0.002118644
perch	1	6	0.85714286	0.014830508
Pickerel	3	0	0.00000000	0.006355932
Pike	0	1	1.00000000	0.002118644
Pollock	0	1	1.00000000	0.002118644
Porgy	0	2	1.00000000	0.004237288
prawn	5	0	0.00000000	0.010593220
redfish	0	1	1.00000000	0.002118644
reedi:	0	1	1.00000000	0.002118644
rockfish	1	5	0.83333333	0.012711864
Sablefish	6	1	0.14285714	0.014830508
Salmon	45	11	0.19642857	0.118644068
Scallop	13	1	0.07142857	0.029661017
seabass	4	0	0.00000000	0.008474576
seabream	0	4	1.00000000	0.008474576
Snapper	0	2	1.00000000	0.004237288
Sole	13	8	0.38095238	0.044491525
Swordfish	1	0	0.00000000	0.002118644
Tilapia	0	34	1.00000000	0.072033898
toothfish	0	3	1.00000000	0.006355932
Trout	8	3	0.27272727	0.023305085
tuna	29	23	0.44230769	0.110169492
Turbot	1	2	0.66666667	0.006355932
Two	0	1	1.00000000	0.002118644
whiting	0	1	1.00000000	0.002118644

```
EXPECTED FISH
            Labeled Mislabelled MislabeledProp  SampleProp
Amberjack        1          0   0.00000000 0.002118644
Bass             8          9   0.52941176 0.036016949
Butterfish       0         12   1.00000000 0.025423729
char             4          0   0.00000000 0.008474576
Char             1          0   0.00000000 0.002118644
Cod             56         20   0.26315789 0.161016949
Crab             2          3   0.60000000 0.010593220
Cuttlefish       1          0   0.00000000 0.002118644
Dora             0          1   1.00000000 0.002118644
Eel              2          0   0.00000000 0.004237288
fish             0          1   1.00000000 0.002118644
Flounder         1          0   0.00000000 0.002118644
Grouper          0          5   1.00000000 0.010593220
Haddock          3          0   0.00000000 0.006355932
Hake             1          3   0.75000000 0.008474576
Halibut         33         12   0.26666667 0.095338983
Lavraki          0          1   1.00000000 0.002118644
Lithrinin        0          1   1.00000000 0.002118644
Lobster          4          1   0.20000000 0.010593220
Loup de mer      0          2   1.00000000 0.004237288
Mackerel         2          0   0.00000000 0.004237288
Matane           0          1   1.00000000 0.002118644
Monkfish         1          0   0.00000000 0.002118644
nordique         1          0   0.00000000 0.002118644
Perch            1          0   0.00000000 0.002118644
Pickerel         3          2   0.40000000 0.010593220
Prawn            5          0   0.00000000 0.010593220
Rockfish         1          0   0.00000000 0.002118644
Sablefish        2          1   0.33333333 0.006355932
Salmon          47         13   0.21666667 0.127118644
Scallop         13          1   0.07142857 0.029661017
Seabream         0          1   1.00000000 0.002118644
Snapper          0         55   1.00000000 0.116525424
Sole            15         16   0.51612903 0.065677966
Steelhead        1          0   0.00000000 0.002118644
Swordfish        1          0   0.00000000 0.002118644
Trout            5          1   0.16666667 0.012711864
Tuna            29         38   0.56716418 0.141949153
Turbot           1          0   0.00000000 0.002118644
White fish       0          1   1.00000000 0.002118644
Yellowtail       0         26   1.00000000 0.055084746
```