AFSC PROCESSED REPORT 2020-04

Alaska Fisheries Science Center
Fisheries Monitoring and Analysis Division

Are Flow Scale and Observer Scale Relationships the Same Among North Pacific At-Sea Processors?

JUNE 2020

U.S. Department of Commerce | National Oceanic and Atmospheric Administration | National Marine Fisheries Service

AFSC Processed Report

This document should be cited as follows:

Faunce, C., A. Molina, and A. Kingham. 2020. Are flow scale and observer scale relationships the same among North Pacific at-sea processors? AFSC Processed Rep. 2020-04, 25 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

.

This document is available online at: https://repository.library.noaa.gov/

Are Flow Scale and Observer Scale Relationships the Same Among North Pacific At-Sea Processors?

C. Faunce[1], A. Molina[2,*] and A. Kingham[1]

[1]Fisheries Monitoring and Analysis Division

Alaska Fisheries Science Center

National Marine Fisheries Service

National Oceanic and Atmospheric Administration

Seattle, WA 98115

[2]Stony Brook University - Science Training and Research to Inform Decisions (STRIDE)

STRIDE Internship with Fisheries Monitoring and Analysis Division

Alaska Fisheries Science Center

National Marine Fisheries Service

National Oceanic and Atmospheric Administration

Seattle, WA 98115

[*]Current Address:

School of Marine and Atmospheric Sciences

Stony Brook University

Stony Brook, NY 11790

June 2020

Abstract

Fishery observers in the North Pacific serve a dual role as scientific samplers and compliance monitors. While this latter role is complex because it can cause conflict with the industry, it serves an important role in maintaining data integrity and helps deter and prevent seafood harvesting fraud. Concerns over the potential for intentional tampering with scales used to weigh catch on some at-sea processing vessels led the Alaska Fisheries Science Center's Fishery Monitoring and Analysis Division (FMA) to collect numerous sample weights from both onboard observer scales and industry scales in the hopes that these would be useful to law enforcement. Comparison of these data were made using a mixed model to account for as many confounding effects as possible and isolate between-vessel effects measured over 5 years. Current data quality control protocols and attention resulted in a data set that was more variable than the effect size to be detected, even after removal of extreme outliers. Consequently, the data are not useful for identifying differences in scale congruence between vessels measured annually without serious data manipulation by analysts -- a process which could potentially negate the utility of the results for the National Oceanic and Atmospheric Administration's Fisheries Office of Law Enforcement (OLE). However, the data may be useful for identifying changes in between-scale agreement within a vessel during a time series for further investigation by OLE.

Contents

Introduction

The effective management of fishery resources under "output controls" requires that catch and its disposition be accurately quantified (Pope 2002). Such information must also be available in real time for each vessel participating in industry cooperatives that can trade catch and bycatch allowances among participants under a common quota.

In the Federal groundfish fisheries off Alaska, vessel- and species-specific catch and disposition information is accurately assessed in real time onboard catcher processors or motherships (CPs) -- vessels that have their own factory on board -- using electronic scales and observers. Flow scales continuously record weight across load cells as catch is moved inside the factory with conveyor belts. Flow scales are extremely useful at providing a timely measurement of total haul weight. Observers are federally-trained biologists who monitor fishing activities to gather unbiased information on catch. Observers in the North Pacific are deployed on every CP trip, and they sample according to a hierarchy where trips > hauls > species composition samples > biological tissue collections (Cahalan et al. 2014, Cahalan and Faunce 2020). Observers obtain multiple random samples of the catch from nearly every haul. They accomplish this by stopping the flow of fish during a random time or weight in the haul, clearing the conveyor belt, collecting all the subsequent fish during a prescribed time or weight, and then taking all the fish and working up a species composition sample that is weighed using a Motion Compensated Platform (MCP) scale (AFSC 2020). Observer samples are used to determine the relative contribution and disposition of each identified species to the total catch, determine size composition, and obtain biological tissues for food habit, maturity, and genetic studies in support of stock assessment scientists and quota managers.

To be useful, scales need to be accurate. Early studies noted that the flow scale in their study performed within ± 3% of the true weight in daily materials tests where a known weight was passed across the scale, but that a small, but consistent positive bias was present in the measurements (Dorn et al. 1995, 1997, 1999). NMFS enacted regulations between 1996 and 2014[1] to help assure the accuracy of scales used for observer data collection and for catch weight records. These requirements include an annual certification of scales through a materials test that allows a technician to calibrate the scale and daily performance of flow scale materials tests in the presence of an observer and/or video so that the weight measured by the flow scale is within ± 3% of the true weight (as defined by the MCP scale). Observers in the North Pacific perform calibration tests on MCP scales each 12-hour shift if working with another observer, and each day if working alone. MCP scales must perform within 0.5% of test weights, and the time and result of

[1] https://www.fisheries.noaa.gov/action/alaska-sea-scales-program-federal-register-rules-and-notices.

MCP scale tests are recorded daily in observer logbooks[2]. The NMFS does not adjust flow scale weights by the error measured in daily tests.

Despite their widespread use by the Alaska groundfish fishing fleet, the performance of flow scales has not been evaluated since the initial studies of Dorn et al. (1995, 1997, 1999), and none have been performed *in situ* under normal fishing operations. There may be considerable bias introduced into catch weights if crew are able to manipulate flow scale readings after tests are performed. As evidence, in January 2012, NOAA's Office of General Counsel issued a Notice of Violation and Assessment (NOVA) for flow scale violations alleged to have occurred on three CPs operating in the North Pacific in 2007, 2008, 2011, and 2012. The parent company later settled for what was at the time the largest monetary penalty of its kind (NOAA OLE 2015).

In light of this case, in 2014 the Fisheries Monitoring and Analysis (FMA) Division of the Alaska Fisheries Science Center began instructing observers to collect flow scale weights corresponding to their sample weights in addition to the MCP weight as part of their regular duties under the premise that they may be useful to the National Oceanic and Atmospheric Administration's Fisheries Office of Law Enforcement (OLE). In theory, the weights from the flow scale and the MCP scale should be the same because they are recordings of the same fish. This study examines whether these comparative data can be used to discriminate between-vessel differences under the assumption that vessels act as individuals and have behaviors that are quantifiable over the period of years.

Methods

In 2013, observers on CPs with flow scales were instructed at the time of their sample to stop the flow scale, clear the scale, record the starting flow scale weight, run the flow scale and collect their sample, record the end flow scale weight, and work up their sample using the MCP scale as normal. Information on the sample, haul, trip and vessel information were recorded as per standard duties. In 2014, after successfully collecting this information for a year, these data began to be entered into databases maintained by the Observer Program.

Our methods were to examine and clean data if needed and then build explanatory statistical models. First, scale data between 2014 and 2018 inclusive were examined for outliers. Our outlier selection process was made difficult because having analysts decide which data to include and which data to exclude may thwart successful prosecution by the OLE if differences were found. Instead, we initially believed that debriefed data which had undergone standard data quality control review provided by FMA

[2] https://www.fisheries.noaa.gov/resource/document/north-pacific-observer-sampling-manual.

would be sufficient for analysis. After initial inspection, however, we realized that we would need to eliminate some of the most egregious data outliers. Towards this end, cumulative distribution curves were examined for each scale as was the raw data relationships between scales. We also used percent difference (D) metrics to identify potential outliers where $D = [(W_f - W_{MCP}) / W_{MCP}] \times 100$ and W_{MCP} is the platform scale weight and W_f is the flow scale weight. Our intent was to identify as few points as possible as outliers. We assumed that where differences occurred, the MCP scale was correct since it is tested to a more strict tolerance and is tared more often than flow scales.

We identified several metrics that might explain between-scale differences. These include how the sample was collected (observers use a hierarchical sampling design where samples are contained within hauls contained in trips), the area fished, the observer, vessel, and fishing gear type. In addition to data collected by the observer, we created four factors that may be useful in describing scale differences: periods before and after an exempted fishing permit (EFP), the frequency of sample collection by the observer, the type of sample taken by the observer, and the fishing sector in which the vessel was participating. Briefly, an EFP that was conducted during the same time period incentivized vessels to discard Pacific halibut before they reached the factory to reduce post-capture mortality[3]. In 2018, the observer scale comparison data collection frequency changed from every sample to only once a haul, so we created a pre- and post-2018 dummy variable. Sample type was also identified as a potential factor associated with scale congruence. Sample type describes whether the observer haul sample was based on a small weight during species diverse hauls, was based on the collection of non-predominant species, or based on the collection of predominant species or species homogenous samples. Finally, fishing sector was created based on the fishing target, vessel, and time of year (Table 1).

We used linear mixed models to investigate the relationship between flow- and MCP scale weight following Bates et al. (2015). Mixed models offer advantages over traditional linear models because they allow the modeler to explicitly account for the non-independence in highly structured, hierarchical data (such as the case of the observer sampling design). Inclusion of sampling hierarchy as random effects can control for pseudoreplication if they are specified correctly, and can help to draw correct inferences about fixed effects, while also reducing the probability of false positives and negatives.

To address whether the relationship between scale weights differ among vessels, we fit a model with flow scale weight as the response variable, MCP scale weight as the continuous predictor, vessel as a categorical fixed effect, and random effects to account for the hierarchical, repeated measures inherent in the data as well as other random effects that could have affected the recorded weights. Models were run

[3] The EFP is now standard practice governed by Federal Regulations. https://www.fisheries.noaa.gov/action/non-pollock-trawl-catcher-processor-halibut-bycatch-handling-and-monitoring-requirements.

with different combinations of factors based on author expertise and increasing complexity, starting with only single order effects (no interaction terms) after examining correlations among factors to avoid collinearity. Models were explored with random effects included as intercepts, slopes, or both (with covariance included). Model performance was compared using analysis of variance and information criterion (AIC). Candidate model selection was based on an author selected subset of evaluated models (Candidate models). The final model was the best-fitting model from the candidate models that could generate bootstrapped confidence intervals for the fixed effect (the effect in question) in a 24-hour period.

Results

Data for flow scale comparisons were numerous and contained several obvious outliers. Which of these reflected actual stochastic errors or faults in either of the two scales is largely unknown, but some of them are likely human error. For example, we identified a series of potential outliers that lie along a line with intercept = +/- 100 kg, indicating a transcription error by the observer. We could not identify a clear cut-off for outlier definition based on cumulative distribution curves of scale weight. However, since the permissible scale weight differences were 3%, we explored -- and were greatly satisfied by the definition of -- an outlier as points that had differences of 300% (Fig. 1). This definition resulted in 0.07% of the data being removed as outliers.

A total of 48 models were explored with 7 candidate models (Table 2). Correlations among variables are depicted in Figure 2. Many of the best-fitting models included the factors of NMFS Area or Observer. However, these factors were likely overfitting the data. For example, a quick investigation of data revealed that models that contained the NMFS Area had many more singularities (where a factor only existed for one sample) than those that did not (Table 3). Similarly, there were a large number of observer:vessel combinations that only occurred once (Fig. 3) and differences between scale weight did not noticeably differ by observer (Fig. 4).

The final model, while not the best-performing model, was selected due to its simplicity and ability to generate confidence bounds around the fixed effects, which was our primary interest (Table 4). The final model contained data from 307,720 samples collected from 125,358 hauls, 3,772 trips, 5 years, 3 FMP Areas, 2 gear types, and 38 vessels by 550 observers. High correlations between gear type, sector, and sample type eliminated the latter two factors from consideration (Fig. 2). Random effects that resulted in differences in scale weights were gear type (with non-pelagic gear having a positive effect and pelagic trawl gear having a negative effect) and FMP Area (with Bering Sea having a positive effect) (Fig. 5). Some hauls and trips had large effect sizes but we cannot rule out that some of these are outliers (Fig. 6). Most importantly, when effect sizes and credible intervals for the fixed effect of vessel were investigated, we could not find any between-vessel differences after accounting for other differences (Fig. 7), and these were validated by percent difference plots (Fig. 8). Model diagnostics revealed satisfactory results for residual and scale-location, however the scale weight data were highly skewed, with large values being more predominant than expected (Fig. 9). Variance explained (%) by intercept terms after fixed effects was as follows: Haul (5.86), Trip (2.84), Year (36.85), FMP (0.10), and residual (54.34) -- Gear was included as a slope and not an intercept.

Discussion

In early work by Dorn et al. (1997), the authors state "*If NMFS requires the use of flow scales to estimate total catch, adequate resources must be devoted to the evaluation of flow scales, monitoring at-sea test results, and auditing records to identify improper uses of the flow scale*". Towards this end, FMA has met the goal of monitoring at-sea test results since flow scales were first used at-sea. In recognition that scale test results may not reflect the accuracy of the scale during the other hours of the day, significant field resources have been expended by FMA towards gaining information to evaluate flow scale accuracy outside of the daily flow scale test. The results of these comparisons warrant an evaluation of their own and are planned in a companion work to this one. However, the question that we address here-- whether or not these between-scale data, numbering in the hundreds of thousands are of use in identifying individual vessel differences as potential evidence of tampering -- leaves us with a sobering result. Despite our efforts, we were unable to identify a significant vessel effect in the data. What makes this sobering is the likely reason for this result is not that vessels did not alter their flow scale, but rather that the data collected are ill-suited to address this question. Although debriefing protocols are in place for these between-scale data, their importance value was set as minimal given the demands on observer and staff time, and no "sanity checks" were in place for data entry. Although permissible flow scale test accuracy must be within 3%, the data that we examined in the final observer database tables were rife with large, impossible between-scale differences. This lack of data quality rigor has two negative consequences. First, it necessitates that the analyst decide what makes good data and not. This mere fact makes any result used for OLE suspect in importance and believability. While we have tried to remove as little data as possible and identified several circumstances where human error was likely, we could not definitively identify errors in the data. Second, since the permissible error in flow scale accuracy was 3% from zero, the accuracy of the between-scale comparison needed to be at least that good to detect behavioral differences. Under current debriefing protocols, these data cannot be used to reliably identify noncompliance among vessels in an annual time period.

This analysis does not negate the importance of between-scale data for examination of potential misuse of flow scales after an observer statement of potential non-compliance has been prepared. For example, it may be possible to review time series data from a particular vessel to determine whether the data on flow scale performance is the same between a period identified as suspicious by an observer and outside this time period (so-called change point or change analysis). Unfortunately, investigation of change points is made difficult by the choice of test, significance value, and other decisions made by the analyst. In addition, change points are subject to the influence of outliers, which was the problem in this modelling exercise. Notwithstanding, the approach does have merit for identifying potential changes in between-scale performance in a time series for forwarding to OLE for further investigation and methods are available to overcome the influence of outliers (Fearnhead and Rigaill 2019).

Some hauls in this study had very large differences that -- if real -- could have consequences to accurate catch estimation. Personal communication with FMA staff and reviews of observer notes identified occurrences when a flow scale reads high due to mud and muck on the load cells. While the frequency of collecting information from the flow scale and MCP scale for comparison has utility, it does not need to be done every sample – the change in the frequency of between-scale data collection by observers to once a haul in 2018 seems prudent, and perhaps that frequency could be reduced even more. However, we do not recommend eliminating the practice completely. Observers represent a unique enforcement resource that can facilitate detection and penalization of violations (Porter 2010), and while this role makes many in the fisheries monitoring community uneasy because it creates the opportunity for conflict between observers and industry, its value cannot be understated. Alaska is unique in the Nation in that only in this region is observer safety and prevention of harassment the top priority of OLE. The mere act of stopping the flow scale belt, clearing it, taking a sample, and weighing it is an act that identifies that the Observer Program is looking at flow scale accuracy in real time, and not just during the daily flow scale tests. This act alone warrants continuation because it generates a positive observer effect, where behavior to tamper with flow scales, if present, is reduced by the highly visible activities by the observer (and assumed by proxy high vigilance by OLE).

The large observer program in the North Pacific no doubt contributes to achieving what has been called by some as the best-managed fisheries in the world (Worm et al., 2009). For its continued success, focus should be placed on collecting data at the accuracy required for its purpose.

Citations

AFSC (Alaska Fisheries Science Center). 2020. 2020 Observer Sampling Manual. Fisheries Monitoring and Analysis Division, North Pacific Groundfish Observer Program. AFSC, 7600 Sand Point Way N.E., Seattle, WA, 98115. 574 p. https://www.fisheries.noaa.gov/resource/document/north-pacific-observer-sampling-manual.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear Mixed-Effects Models using lme4. J. Stat. Software 67: 1-48. doi:10.18637/jss.v067.i01.

Cahalan, J., J. Gasper, and J. Mondragon. 2014. Catch sampling and estimation in the federal groundfish fisheries off Alaska, 2015 edition. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-286, 46 p.

Cahalan, J., and C. Faunce. 2020. Development and implementation of a fully randomized sampling design for a fishery monitoring program. Fishery Bulletin, U.S. 118:87-99.

Dorn, M. W., S. M. Fitzgerald, M. A. Guttormsen, and M. R. Loefflad. 1995. An evaluation of North Pacific groundfish observer program methods of haul weight estimation. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-56, 31 p.

Dorn, M. W., S. K. Gaichas, S. M. Fitzgerald, S. A. and Bibb. 1997. Evaluation of haul weight estimation procedures used by at-sea observers in pollock fisheries off Alaska. AFSC Processed Rep. 97-07, 76 p. Alaska Fish. Sci. Cent., Natl. Mar. Fish. Serv., NOAA, 7600 Sand Point Way NE, Seattle, WA 98115-0070.

Dorn, M. W., S. K. Gaichas, S. M. Fitzgerald, and S. A. Bibb. 1999. Measuring total catch at sea: Use of a motion-compensated flow scale to evaluate observer volumetric methods. N. Am. J. Fish. Manage. 19: 999-1016.

Fearnhead, P., and G. Rigaill 2019. Changepoint detection in the presence of outliers. J. Am. Stat. Assoc. 114:525, 169-183.

NMFS OLE (National Marine Fisheries Service Office of Law Enforcement). 2015. Office of Law Enforcement annual report fiscal year 2015. Available at: https://repository.library.noaa.gov/view/noaa/15635.

Pope, J. 2002. Input and output controls: the practice of fishing effort and catch management in responsible fisheries. In A fishery manager's guidebook: management measures and their application, pp. 75-94. Ed. by K. L. Cochrane. FAO, Rome.

Porter, R. 2010. Fisheries Observers as Enforcement Assets: Lessons from the North Pacific. Mar. Policy 34:583-589.

Worm, B., R. Hilborn, J. K. Baum, T. A. Branch, J. S. Collie, C. Costello, M. J. Fogarty, et al. 2009. Rebuilding global fisheries. Science 325: 578-585.

Acknowledgments

We are indebted to the numerous dedicated observers who made this data possible. Models were performed by AM under guidance of CF and AK using data queries of AK. This work was written by CF edited by AK. We are thankful to the Science Training and Research to Inform DEcisions (STRIDE) program and NOAA for funding for AM.

Table 1. -- Model parameters explored in this study.

Model Parameter (Abbreviation)	Definition
MCP Weight	Motion Compensated Platform Scale Weight. Derived from observer samples.
Vessel (Permit)	Identify of the vessel.
Fishing Trip (Trip)	Defined in observer data.
Haul	Period when fish are brought on board using the trawl net.
Calendar Year (Year)	January – December.
Gear Type (Gear)	Pelagic or non-pelagic trawl nets.
FMP	Fishery Monitoring Plan Area: Bering Sea, Aleutian Islands, and Gulf of Alaska.
Manual Year	Period when observers are trained on a technique for the coming calendar year. Can overlap different calendar years.
Frequency	Prior to 2018 observers collected MCP data on every sample. In 2018 that protocol changed to once a haul. This factor accounts for that sampling difference.
Vessel	Catcher Processor or Mothership.
Observer	Identify of the observer.
NMFS Area (Area)	Statistical area for reporting catch.[1]
Cruise	Combination of observer and vessel.
Halibut EFP (Y or N)	Y or N. Deck sorting practice by vessels to reduce post-capture halibut mortality.
Sector	Fishing activity defined by authors.
Sample Type	Size of sample, highly correlated with sector.

[1]https://www.fisheries.noaa.gov/alaska/sustainable-fisheries/alaska-fisheries-figures-maps-boundaries-regulatory-areas-and-zones

Table 2. -- Suite of models examined, and whether the random effect was included as an intercept (I) or slope (S). Models in bold were run to bootstrap confidence intervals on the fixed effect, but only the highlighted model (#35, fg.s) was successful and was therefore selected as the final model.

		Model Diagnostics					Random Effects												
Model #	**Model Name**	**-logLik**	**AIC**	**Delta LogLik**	**Delta AIC**	**df**	**Trip**	**Haul**	**Calendar Year**	**Manual Year**	**Frequency**	**Observer**	**Area**	**FMP**	**Gear**	**Sector**	**Sample Type**	**Vessel**	**halibut EFP**
1	**area.si**	**1058585.6**	**2117263.3**	**160.1**	**0**	**46**	**I**	**I**	**I**			**I**	**S & I**						
2	area.sic	1058591.5	2117277.1	154.1	13.8	47	I	I	I			I	S & I *						
3	area.cal.si	1058594.5	2117280.9	151.2	17.7	46	I	I	S & I			I	I						I
4	**area.cal.s**	**1058597.3**	**2117284.6**	**148.4**	**21.3**	**45**	**I**	**I**	**S**			**I**	**I**						**I**
5	**obs.area**	**1058598.4**	**2117286.7**	**147.3**	**23.4**	**45**	**I**	**I**	**I**			**I**	**I**						**I**
6	area.cal.sic	1058603.4	2117300.8	142.3	37.5	47	I	I	S & I*			I	I						
7	area.s	1058607.9	2117305.7	137.8	42.5	45	I	I	I			I	S						
8	fmp.cal.si	1058610.8	2117313.6	134.9	50.4	46	I	I	S & I			I		I					I
9	**obs.fmp.gear**	**1058611.0**	**2117313.9**	**134.7**	**50.7**	**46**	**I**	**I**	**I**			**I**		**I**	**I**				
10	fmp.si	1058611.0	2117314.1	134.7	50.8	46	I	I	I			I		S & I					
11	fmp.cal.s	1058612.8	2117315.5	132.9	52.3	45	I	I	S			I		I					
12	fmp.sic	1058611.0	2117316.1	134.7	52.8	47	I	I	I			I		S & I*					
13	obs.fmp	1058613.1	2117316.3	132.5	53	45	I	I	I			I		I					
14	fmp.cal.sic	1058618.5	2117331	127.2	67.7	47	I	I	S & I*			I		I					
15	fmp.s	1058625.8	2117341.5	119.9	78.3	45	I	I	I			I		S					
16	obs	1058657.0	2117402	88.7	138.7	44	I	I	I			I							
17	no.h	1058663.8	2117423.5	81.9	160.3	48	I	I	I				I	I	I	I	I	I	
18	no.g	1058665.2	2117426.5	80.5	163.2	48	I	I	I				I	I		I	I	I	
19	no.f	1058666.3	2117428.5	79.4	165.2	48	I	I	I				I		I	I	I	I	
20	no.t	1058666.4	2117428.8	79.3	165.5	48	I	I	I				I	I	I	I		I	

Table 2. -- Cont.

		Model Diagnostics					Random Effects												
Model #	Model Name	-logLik	AIC	Delta LogLik	Delta AIC	df	Trip	Haul	Calendar Year	Manual Year	Frequency	Observer	Area	FMP	Gear	Sector	Sample Type	Vessel	halibut EFP
21	**thysa**	**1058671.1**	**2117432.3**	**74.6**	**169**	**45**	**I**	**I**	**I**				**I**				**I**		
22	area	1058672.6	2117433.3	73	170	44	I	I	I				I						
23	thyga	1058672.6	2117435.3	73	172	45	I	I	I				I		I				
24	full	1058672.2	2117442.4	73.5	179.1	49	I	I	I				I	I	I	I	I	I	I
25	no.v	1058676.7	2117449.3	69.0	186.1	48	I	I	I				I	I	I	I	I		
26	no.y	1058676.7	2117449.4	69.0	186.1	48	I	I					I	I	I	I	I	I	
27	**fg.si**	**1058678.7**	**2117449.5**	**67.0**	**186.2**	**46**	**I**	**I**	**I**					**I**	**S & I**				
28	fg.sic	1058677.8	2117449.7	67.9	186.4	47	I	I	I					I	S & I*				
29	no.a	1058679.7	2117455.5	65.9	192.2	48	I	I	I					I	I	I	I	I	
30	thygf	1058685.5	2117461	60.2	197.7	45	I	I	I					I	I				I
31	fmp	1058687.6	2117463.2	58.1	200.0	44	I	I	I					I					I
32	thysf	1058686.8	2117463.6	58.9	200.3	45	I	I	I						I		I		I
33	thsa	1058689.6	2117467.2	56.1	203.9	44	I	I					I				I		
34	thga	1058691.4	2117470.8	54.3	207.5	44	I	I					I		I				
35	**fg.s**	**1058691.4**	**2117472.8**	**54.3**	**209.6**	**45**	**I**	**I**	**I**					**I**	**S**				**I**
36	thgf	1058703.2	2117494.3	42.5	231.1	44	I	I						I	I				
37	thsf	1058704.3	2117496.6	41.4	233.3	44	I	I						I			I		
38	sec	1058721.6	2117531.3	24.0	268.0	44	I	I	I							I			
39	ves	1058724.5	2117537.1	21.2	273.8	44	I	I	I									I	
40	gear	1058725.1	2117538.2	20.6	274.9	44	I	I	I						I				
41	man	1058727.0	2117540.1	18.7	276.8	43	I	I		I									
42	cal	1058727.2	2117540.3	18.5	277.1	43	I	I	I										
43	man.reg	1058727.0	2117542.1	18.7	278.8	44	I	I		I	I								
44	cal.reg	1058727.2	2117542.3	18.5	279.1	44	I	I	I		I								
45	hal	1058727.2	2117542.3	18.5	279.1	44	I	I	I										

Table 2. -- Cont.

		Model Diagnostics					Random Effects												
Model #	**Model Name**	**-logLik**	**AIC**	**Delta LogLik**	**Delta AIC**	**df**	**Trip**	**Haul**	**Calendar Year**	**Manual Year**	**Frequency**	**Observer**	**Area**	**FMP**	**Gear**	**Sector**	**Sample Type**	**Vessel**	**halibut EFP**
46	type	1058742.1	2117572.2	3.6	309.0	44	I	I	I								I		
47	trip.haul	1058745.7	2117575.4	0.0	312.1	42	I	I											
48	nest	1058745.7	2117579.4	0.0	316.1	44	I	I											

*With Correlation.

Table 3. -- Effect of different combinations of potential random variables on the amount of singularities in the data.

Grouping Factors (all have permit)	**% Singularities (number of unique hauls over total)**
Year, Sector, Gear, Observer	1.60
Year, FMP, Gear	1.70
Year, FMP, Observer	1.70
Year, Gear, FMP, Observer	1.80
Year, FMP, Area, Observer	7.95
Year, Area, FMP	7.95
Year, Area, Observer	7.95
Year, Gear, FMP, Area	8.00

Table 4. -- Model parameters included or not in the final model.

Model Parameters	Included (Y/N)	Reason
MCP Weight	Y	Dependent variable.
Permit	Y	Variable of interest.
Trip	Y	Needed to capture sampling design.
Haul	Y	Needed to capture sampling design.
Year	Y	Model is better with it included.
Gear	Y	Although not as good as Area or FMP, it is in the best model we could get confidence intervals for.
FMP	Y	Next best suite of models after Area.
Manual Year	N	Same as Calendar Year and less intuitive.
Frequency	N	Model without it is better.
Vessel Type	N	Model without it is better.
Observer	N	Although model is better with it, cannot compute CI when it is included – likely overfitting (See text).
Area	N	Best models have it, but results in singularities and may be overfitting (See Table 3).
Cruise	N	Not needed to capture sampling design because it is used to create trip and haul unique identifiers.
Halibut EFP	N	Not as good as including Area or FMP.
Sector	N	Correlated with FMP and Area, which are both better.
Sample Type	N	Not as good as Area or FMP.

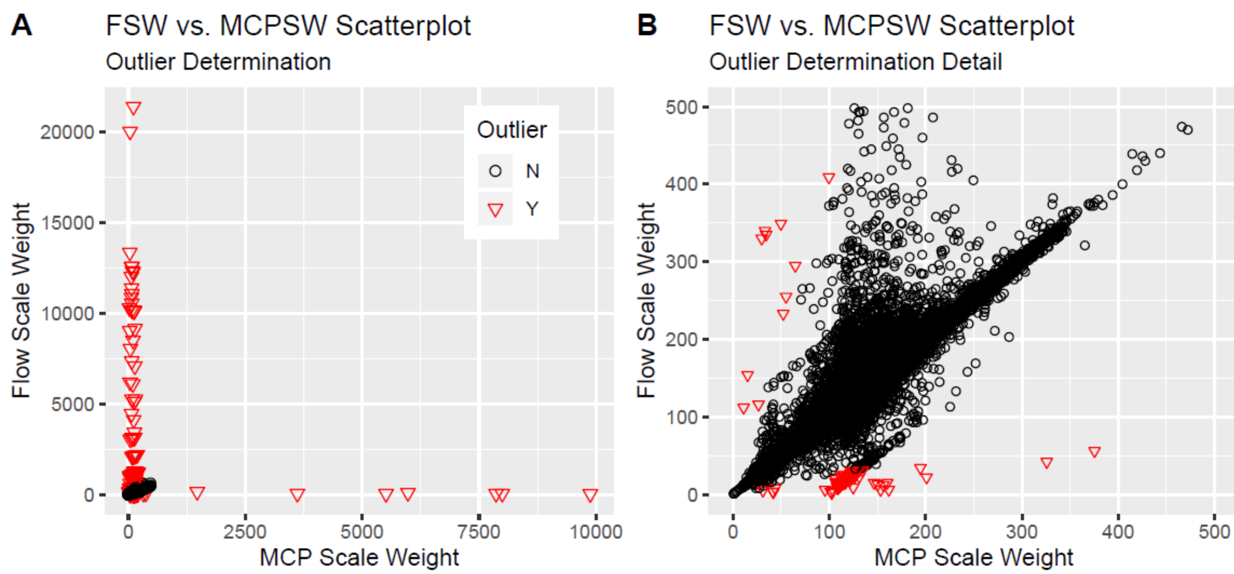

Figure 1. -- Outlier determination plots.

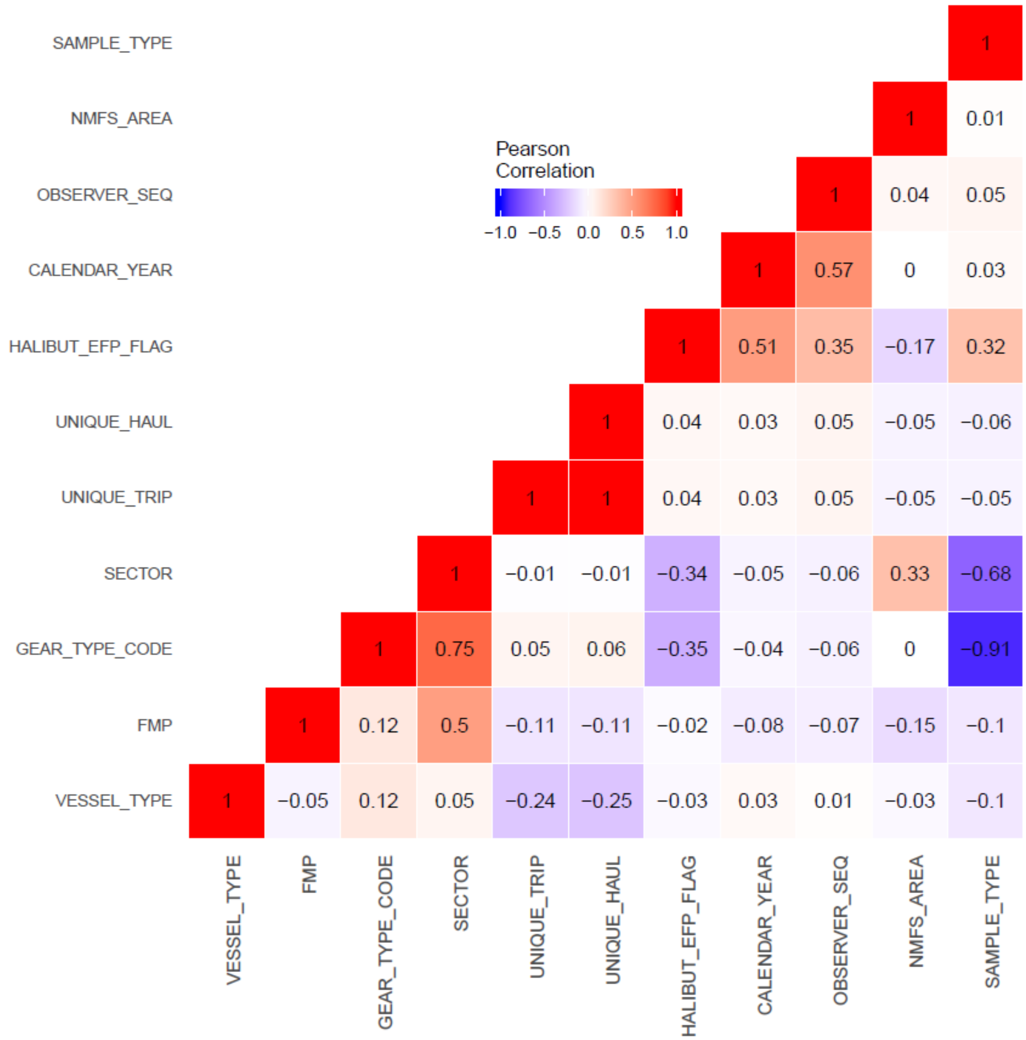

Figure 2. -- Correlation plot of random factors used in this study.

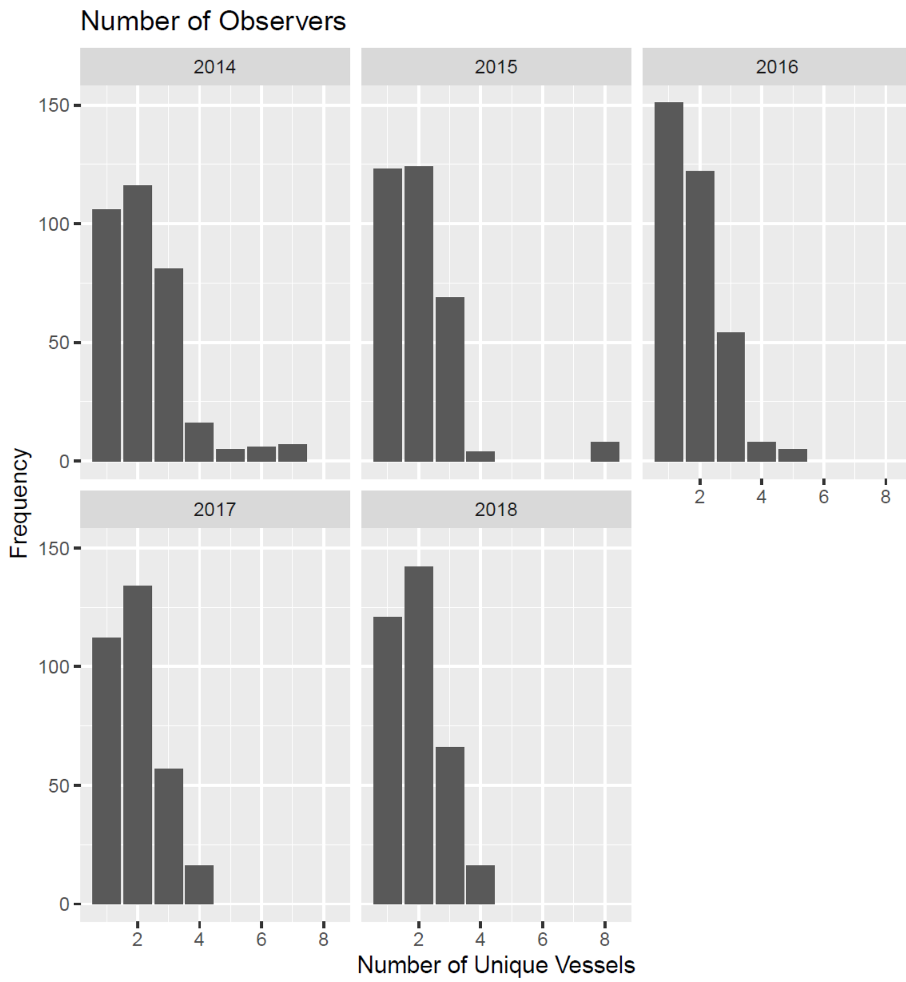

Figure 3. -- Frequency an observer was assigned to different number of unique vessels during a calendar year.

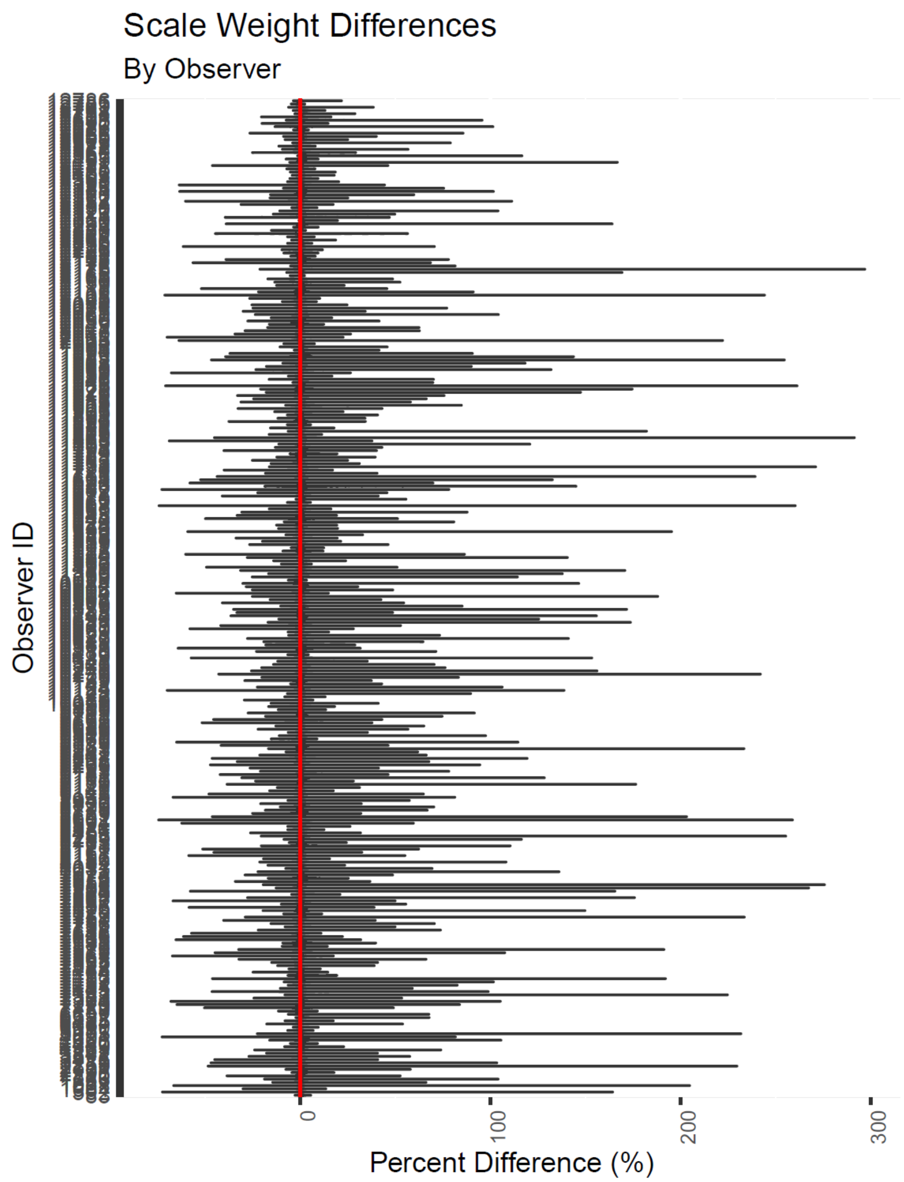

Figure 4. -- Range of differences in scale weight by observer.

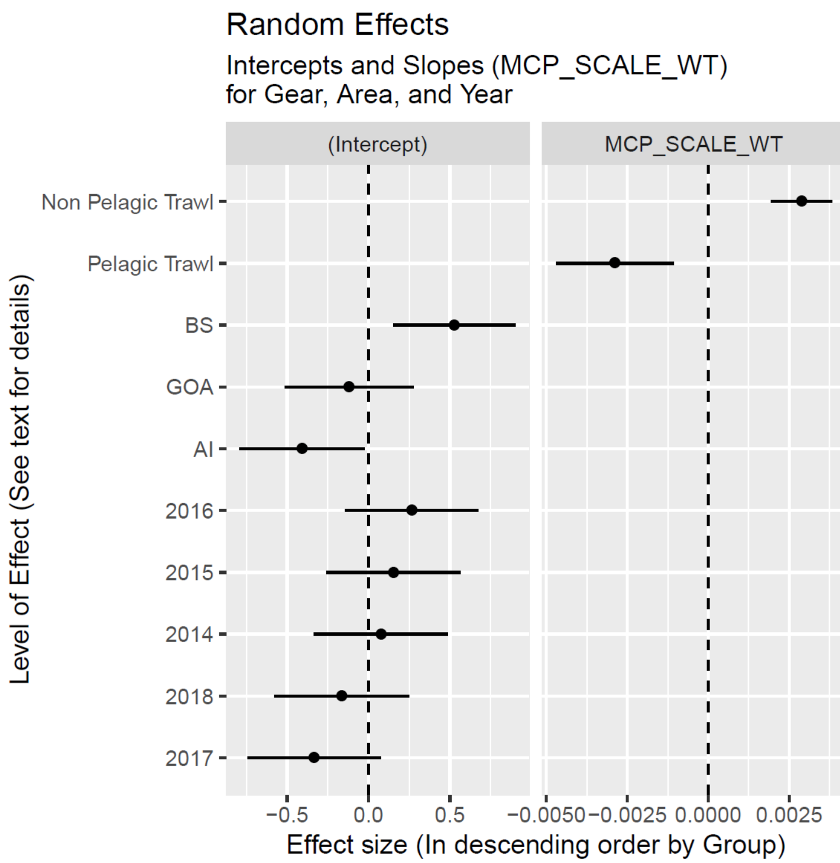

Figure 5. -- Random effect sizes and credible intervals for intercept and slope parameters Gear, FMP, and Year.

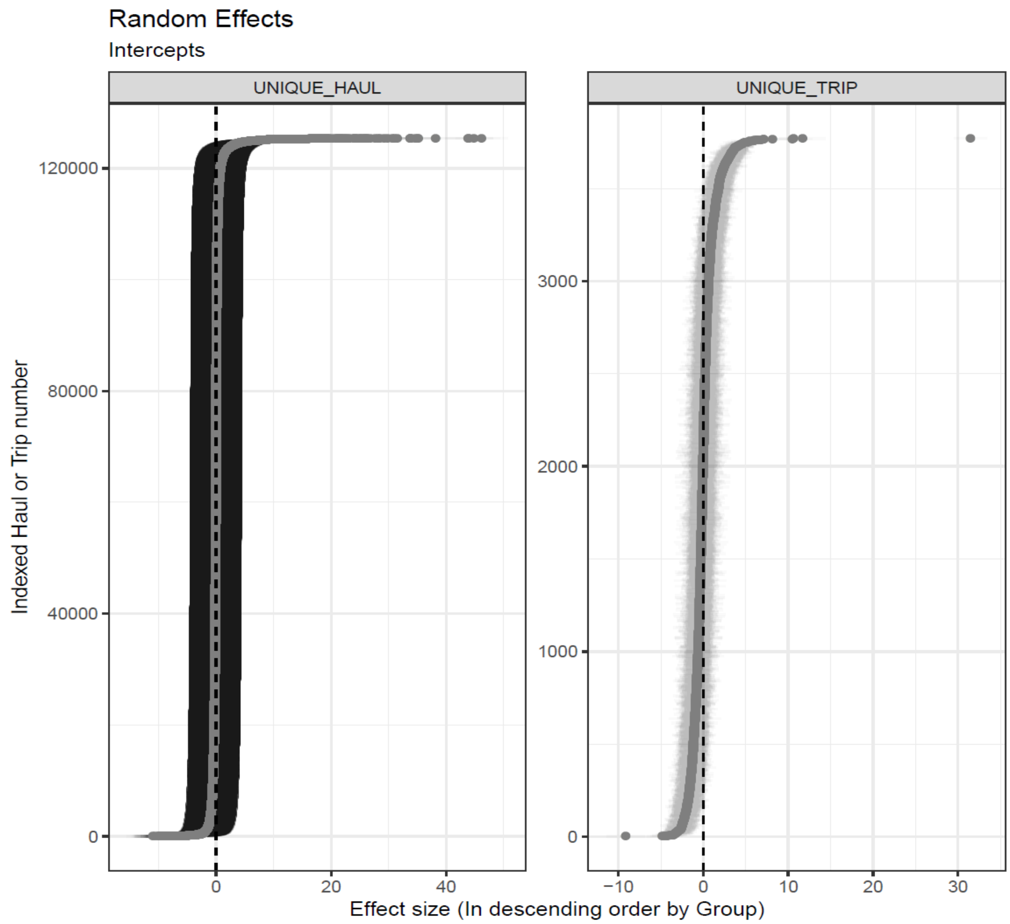

Figure 6. -- Random effect sizes and credible intervals for Haul and Trip.

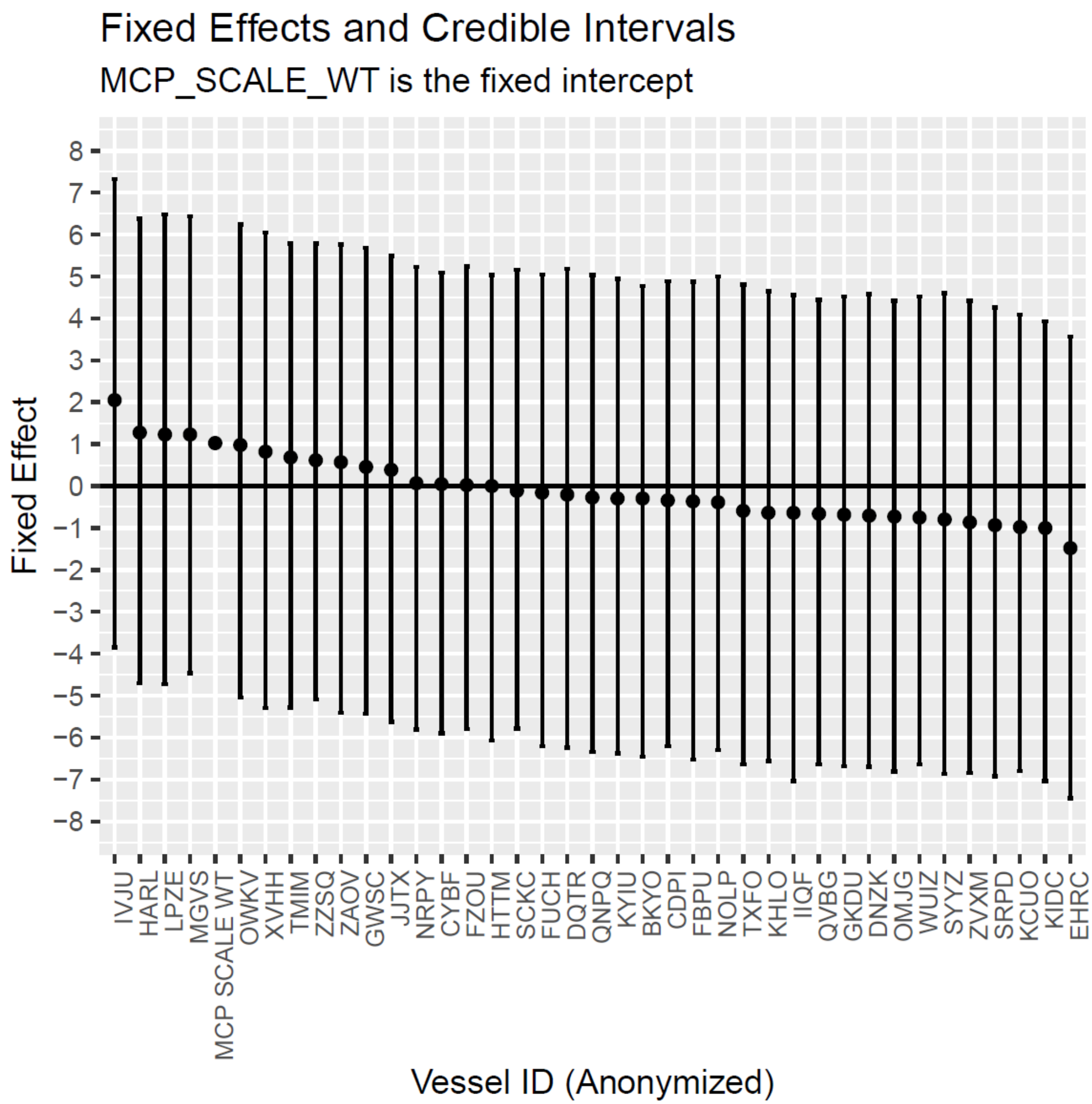

Figure 7. -- Fixed effect sizes and credible intervals.

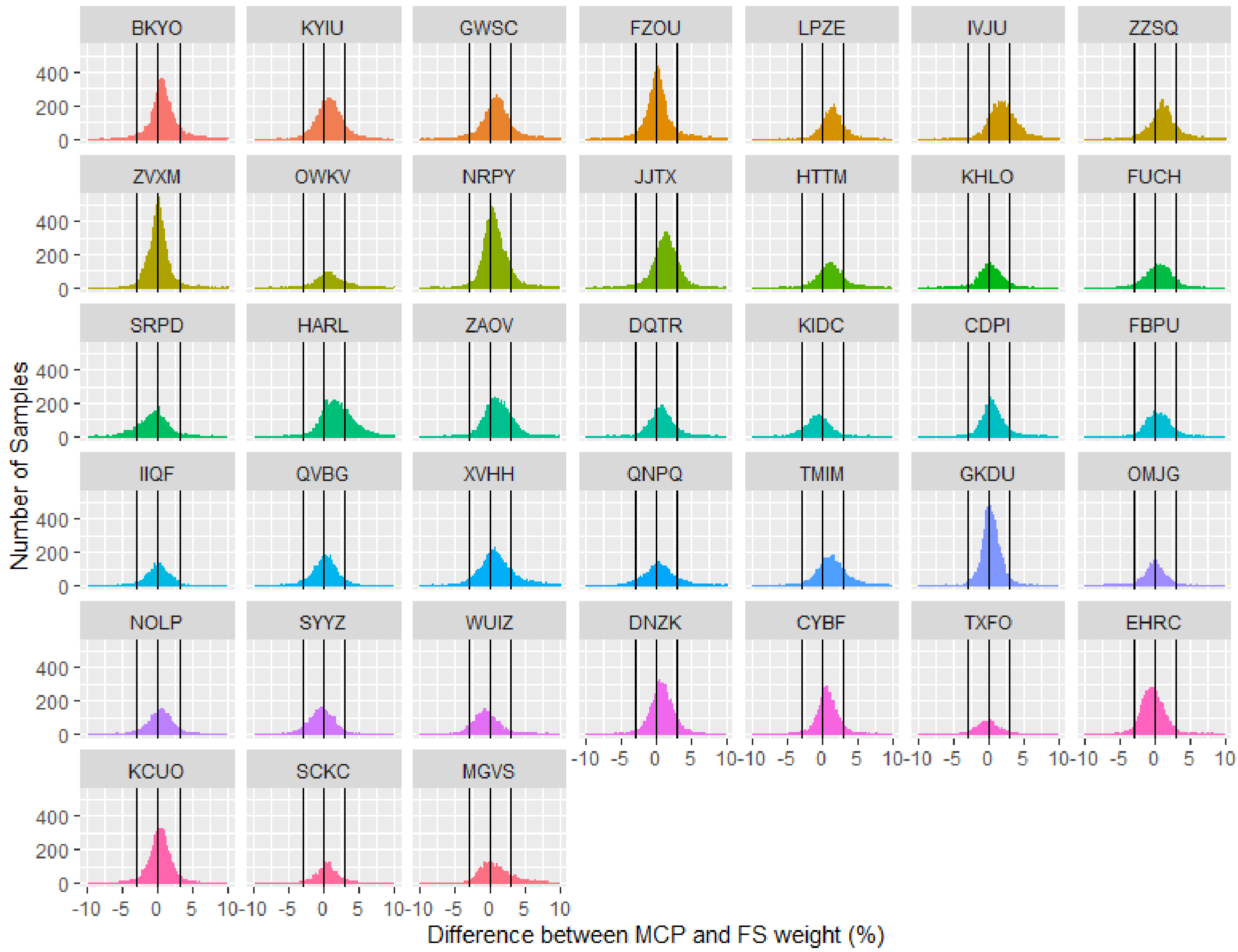

Figure 8. -- Percent differences between scale weights by anonymized vessel name. Vertical solid lines denote -3, 0, and 3% differences, respectively, from left to right.

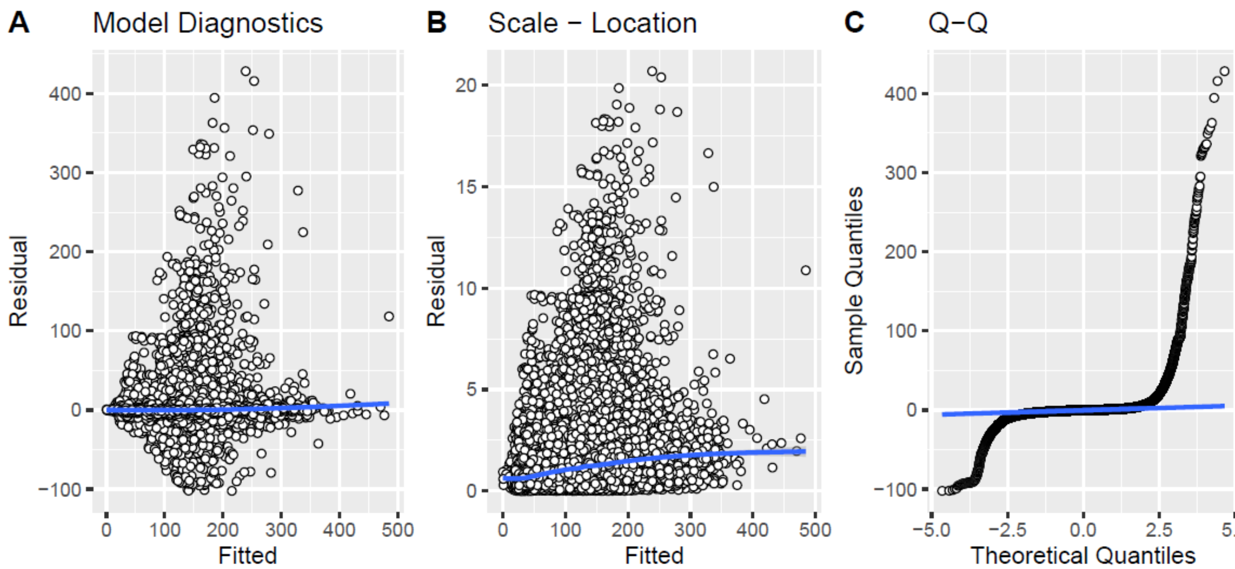

Figure 9. -- Model fit plots.

U.S. Secretary of Commerce
Wilbur L. Ross, Jr.

Acting Under Secretary of Commerce
for Oceans and Atmosphere
Dr. Neil Jacobs

Assistant Administrator for Fisheries
Chris Oliver

June 2020

www.fisheries.noaa.gov

OFFICIAL BUSINESS

National Marine
Fisheries Service
Alaska Fisheries Science Center
7600 Sand Point Way N.E.
Seattle, WA 98115 6349