NOAA Technical Memorandum NMFS-AFSC-406

A Method to Identify Time Periods When Industry and Observer Scale Weights are Incongruent Aboard North Pacific (Alaska) At-sea Processors

C. Faunce and A. Kingham

July 2020

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center

The National Marine Fisheries Service's Alaska Fisheries Science Center uses the NOAA Technical Memorandum series to issue informal scientific and technical publications when complete formal review and editorial processing are not appropriate or feasible. Documents within this series reflect sound professional work and may be referenced in the formal scientific and technical literature.

The NMFS-AFSC Technical Memorandum series of the Alaska Fisheries Science Center continues the NMFS-F/NWC series established in 1970 by the Northwest Fisheries Center. The NMFS-NWFSC series is currently used by the Northwest Fisheries Science Center.

This document should be cited as follows:

Faunce, C., and A. Kingham 2020. A method to identify time periods when industry and observer scale weights are incongruent aboard North Pacific (Alaska) at-sea processors. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-406, 17 p.

This document is available online at:

Document available: https://repository.library.noaa.gov/welcome

This document is available to the public through:
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

www.ntis.gov

NOAA FISHERIES

A Method to Identify Time Periods When Industry and Observer Scale Weights are Incongruent Aboard North Pacific (Alaska) At-sea Processors

Alaska Fisheries Science Center
Fisheries Monitoring and Analysis Division
7600 Sand Point Way
Seattle, WA 98115

U.S. DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Alaska Fisheries Science Center

NOAA Techncial Memorandum NOAA-TM-AFSC-406

July 2020

ABSTRACT

In the Federal groundfish fisheries off Alaska, vessel- and species-specific catch and disposition information is accurately assessed in real time onboard catcher processors and motherships -- vessels that have their own factory on board and process at sea -- using electronic scales and observers who sample nearly every haul on every trip. Flow scales continuously record weight across load cells as catch is moved inside the factory with conveyor belts. To be useful these scales need to be accurate, so a flow scale test must be performed within a 24-hour period and pass within a 3% accuracy. However, flow scale accuracy outside of these tests has not been evaluated for over 20 years, and there may be considerable bias introduced into catch weights if flow scale readings are manipulated after tests are performed. We compared sample weights from flow scales and observer scales under normal fishing operations from over 150,000 hauls from 39 vessels over a 5-year period. Change Point Analyses were used to identify time periods in which the mean difference among scales exceeded permissible amounts for several days in a negative direction indicating weights were biased low. This semi-quantitative approach yielded 12 time periods of concern from five vessels during which the mean differences among scales exceeded permissible amounts, with the vessel scales reading lower. The results of this study were used to focus law enforcement efforts for more detailed investigation and potential action. We conclude with a potential method on how to produce this analysis in a timely fashion to aid fisheries enforcement activities without unduly jeopardizing observer safety.

CONTENTS

INTRODUCTION

In the Federal groundfish fisheries off Alaska, vessel- and species-specific catch and disposition information is accurately assessed in real time onboard catcher-processor and mothership vessels (CPs) -- vessels that have their own factory on board and process catch at sea -- using electronic scales and observers. Flow scales are extremely useful for providing a timely estimate of total haul size in terms of weight since they continuously record all trawl CP catch weight across load cells as catch is moved inside the factory with conveyor belts. 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 obtain multiple random samples of the catch from nearly every haul. All or part of every observer sample on trawl CPs is weighed on a Motion Compensated Platform (MCP) scale suitable for weighing baskets of fish. Observer species composition samples are used to determine the relative contribution and disposition of each identified species to the total catch. Observer biological samples determine size compositions of target and bycatch species and provide biological tissues for food habit, maturity, and genetic studies in support of stock assessment scientists and quota managers (Cahalan and Faunce 2020).

Scales need to be accurate and unbiased. 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 (wherein the flow scale weighted more than the true value) was present in the measurements (Dorn et al. 1995, 1997, 1999). Based on these studies, NMFS enacted regulations between 1996 and 2014 that require flow scales to be certified annually. In addition, the regulations require the scale be calibrated to pass a daily flow scale materials tests that must be performed 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.

Like flow scales, observer MCP scales must also be tested and calibrated daily during normal fishing operations. 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. The MCP scales must perform within 0.5% of test weights, and the time and result of MCP scale tests are recorded in daily observer logbooks. The NMFS does not adjust total haul catch weights derived from flow scale weights by the error measured in daily tests of either scale type.

Despite their widespread use by the Alaska groundfish fishing fleet and the presence of numerous regulations governing the use of flow scales, the performance of flow scales has not been evaluated since the initial studies of Dorn et al. (1995, 1997, 1999), and none of the evaluations were performed *in situ* under normal fishing operations. Therefore, it is unknown how accurate flow scales actually are and if they are prone to systematic errors. Considerable bias could be introduced into catch weights if flow scale readings are manipulated or malfunction after tests are performed.

Observers in the North Pacific make written statements of potential law violations after or during their deployment after conferring with NOAA's Fishery Monitoring and Analysis (FMA) staff of the Alaska Fisheries Science Center. These 'observer statements' are then forwarded to the NOAA's Office of Law Enforcement (OLE). Thus, observers not only represent a data collection resource that is critical for estimating total removals by fishermen, but also represent a unique enforcement resource that can facilitate detection and penalization of violations (Porter, 2010). While these dual roles of the observer makes many in the fisheries monitoring community uneasy because it creates the opportunity for conflict between observers and industry, its value cannot be understated. Based in part on observer statements, in 2012 NOAA's Office of General Counsel issued a Notice of Violation and Assessment (NOVA) in January 2012 for flow scale violations alleged to have occurred on three catcher/processors 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).

The heavy reliance on observer statements towards identifying potential fishery violations prompted the FMA to begin new data collections to support the OLE. In 2013, observers were first instructed to collect flow scale weights that corresponded to the catch weight in their samples via MCP as part of their regular duties. In theory, the weights from the flow scale and the MCP scale should be the same because they are duplicate recordings of the same fish. Despite being collected for numerous years, these data have never been examined. Since the collection of flow scale weights adds to the already numerous tasks an observer must complete, the FMA has interest in maximizing the utility of these data to OLE.

This study examines whether observer data can be used to discriminate within-vessel differences in the flow scale and MCP scale weights throughout time and identify periods when anomalous and suspicious data are present. These "periods of concern" will alert OLE staff of potential problems, and aid decisions by FMA and OLE to continue, modify, or discontinue the practice of collecting scale weight comparison data as part of the regular duties of observers.

METHODS

Data Collection by Observers

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 as a feasibility study for one year, these data were entered into electronic databases maintained by the observer program, known as the 'flow scale/MCP comparison data', or hereafter 'scale data'. Between 2014 and 2017, this protocol was followed for every sample in every haul. In 2018 the protocol changed to ease the burden on observers, and they began collecting these data only on one randomly selected sample per haul.

Data Preparation

Scale data from 2014 through 2019 were pulled from the observer database (NORPAC). Data from each sample were used to calculate a percent difference (D) metric where $D = [(W_f - W_{MCP}) / W_{MCP}] \times 100$ and W_{MCP} is the MCP scale weight and W_f is the flow scale weight. Thus the value of D is negative if the MCP weight is higher than the flow scale weight and positive if the MCP weight is lower than the flow scale weight. It was assumed that when differences occurred, the MCP scale was correct since it is tested and calibrated to more strict tolerance and is tared more often than flow scales.

The frequency of scale data obtained by observers changed from multiple samples prior to 2018 to once a haul for 2018 and 2019. This is problematic for time series analysis since one of the requirements is that data are obtained from equally spaced time periods and are independent, and identically distributed. To help meet these requirements, *D* values from individual samples within a haul were averaged to the haul. Subsequently we assigned empirical cumulative percentiles to each haul *D* value and identified and removed outliers following prior analyses[1].

The Change Point Model

The point at which the statistical properties of a sequence of observations (i.e., a time series) change is called a change point. Change Point Analyses is the detection of change points in a time series. The Change Point Analysis method chosen will be referred to as the model. The same model was applied to each vessel's haul data individually in order to detect the points-in-time when the statistical properties of each vessel's flow scale/MCP comparison data changed. The model applied was a non-parametric multiple Change Point Analysis of mean and variance using the Pruned Exact Linear Time (PELT) algorithm and Change points for a Range of Penalties (CROPS) method of Killick et al. (2012). We performed these models following Killick and Eckley (2014) using the tools provided by Haynes and Killick (2019). Briefly, PELT has been shown to be an efficient and exact method of change point detection, and CROPS allows for the examination of potential number of change points over a range of penalty values to avoid overfitting. CROPS generates a chart where the vertical axis is the change in penalty value,

[1] Kingham, A., and C. H. Faunce. In prep. How final adjusted fishery catch weights are affected by choices of acceptable scale error. Available by request from the authors.

and the horizontal axis is the number of change points. The resulting plot shows a steep decline in the change in penalty value that attenuates as the number of change points increases. Usually there is a point in the plot where the rate of change in the line drops dramatically. This point denotes the point at which maximum parsimony is achieved.

The model has four parameters that need to be specified by the user: a range of penalty values to avoid overfitting the data, the number of quantiles for the test statistic (the test statistic for this model was the empirical distribution (Hayes et al. 2019), and the minimum amount of time to denote as a segment in the time series. Penalty ranges in the model were set at between 1 and 25 × log(the number of haul records), and the value for the number of quantiles for the test statistic was set at 4 × log(the number of haul records). The minimum segment length was set at 29 with the following rationale. Examination of histograms of the frequency of time differences in the scale data revealed that 5 hours was the most common time difference (Fig. 5). Since materials tests must be performed on flow scales every 24 hours, every five successive data points would be related to a flow scale daily test. Under the hypotheses that a vessel that has a Flow Scale out of tolerance for 6 days in a row would indicate serious malfunctioning that would warrant investigation by OLE (a period of concern), a value of 29 was chosen since 24 hours × 6 days = 144 hours, 144 hours / 5 hours per scale data = 28.8 or 29.

In addition to the aforementioned parameters, the model requires that the user select the number of change points. This is accomplished by prompting the user to select a point on the plot provided by CROPS. Given the purpose of change point detection was to identify periods of concern, the decision was made to allow for some overfitting. Therefore, the number of change points selected did not correspond to the most parsimonious solution but to a slightly larger number.

Since this decision results in more change points than would have been selected under more conservative change point selections, a final filtering step was added to the processing of the data. For each time period identified by the model from each vessel, the mean ($\bar{x}$) and a standardized measure of dispersion (Coefficient of Variation, or CV) were calculated where CV was the absolute variance divided by the mean. The results of each time period were labelled as being of low, medium, or high concern following a decision tree using colors to denote level of concern (Fig. 2). Only time periods of high concern ("red") were forwarded to OLE for further investigation. No attempt was made to interpret the factors associated with these periods of high concern.

RESULT AND DISCUSSION

Models from 39 vessels were generated from 153,452 hauls. The number of hauls per vessel ranged between 611 and 8,767 (mean 3,935; interquartile range (IQR): 2,604 – 5,430) and the number of change points ranged between 4 and 57 (mean 26; IQR: 19 – 34). Only five vessels had periods of high concern (Fig. 3); those details are provided in Table 1.

Periods of concern have increased over time. The first period of concern occurred in 2014 on vessel RN during a 20-day time span in June (Table 1). Interestingly, no periods of concern were identified during 2015 or 2016. Only vessel NL had a period of concern in 2017, although this was identified for three time periods lasting between 9 March and 8 August 2017 (Table 1). Despite this long period of time, there are no observer statements of potential violations reported for this vessel except that flow scale tests were not performed on two separate days in June. The years 2018 and 2019 had the most vessels identified as having a period of concern of any year in the time series -- the vessels NL, MB, and RA were identified in 2018 while in 2019 the vessels RA and HU were identified.

The pattern of periods of concern in 2018 and 2019 differed among vessels. For most vessels, a single short period was present in the entire time series. However, the vessels RA, NL, and HU deserve further discussion. Vessel RA appears to have had two summer periods in which large differences in the scales is evident in either direction. The most notable was during May 2018, but this was not identified as a controlled negative, indicating the variance was equal to or larger than the prior time period. However, in 2019, during the same time of the year, vessel RA had a negative scale agreement with lower variance than the period prior and thus it deserves further investigation. Vessel NL had two periods of extended strongly negative difference values June – August 2018 and July – September 2019. From the associated time series it is evident that this vessel can maintain good scale agreement, but during these time periods it did not (Fig. 3). Vessel HU is an interesting time series because this vessels' time series indicates it too can maintain good scale agreement but starts moving in the more negative, more precise direction starting in 2018 and interestingly, continued this trend into 2019 – the period of concern identified was between June and August (Fig. 3). During this time observers on this vessel noted the inconsistencies between scales in written statements of potential violations (Table 1). Across the entire data set, periods of concern predominantly occurred between the months of May and September.

FUTURE DIRECTION

This analysis was conducted on data collected over several years (2014-2019), however this does not need to be the case. Change Point Analyses can examine one or multiple change points and be conducted after the data have been collected (offline) or be done in real time (online). The choice of online or offline Change Point Analysis and rapid or delayed enforcement has implications for the observer program and the observers collecting the data. An advantage to online analysis is that potential flow scale tampering violations could be identified and addressed by OLE through outreach. Online analysis may require online error checking at the point of data entry at the observer program, or an analysis that can handle the presence of potential outliers (e.g., see Fernhead and Regaill 2019). However, rapid enforcement could put the observer and the observer program under direct, immediate, and unwanted scrutiny from captain and crew during the collection of scale comparison data. In contrast, the use of delayed enforcement removes the observer from this immediate cause and effect, but it also increases the magnitude of erroneous catch data.

There may be an intermediate ideal solution for the timing of change point analysis and identification of periods of concern. Catcher processors with flow scales also have cameras on board as part of electronic monitoring requirements for these vessels. Cameras must capture a field of view that includes the scale, where catch enters and leaves the scale, and areas where the crew may stand to adjust the scale. The vessel is required to retain the video for no less than 120 days (unless otherwise specified by NMFS) and be made available upon request by a NMFS employee or any individual authorized by NMFS (679.28(e)(1) (v)). The maximum time between current and an offline Change Point Analysis could be determined from the subtraction of the days necessary to conduct the analysis, identify potential violators, communicate those results to OLE, and obtaining the video footage from 120 days. Periods of 90 days or more would be ideal for observer safety since this is the maximum duration an observer may be continuously deployed onto a catcher processor (679.52).

The process of conducting offline Change Point Analysis need not occur year-round. Since we found a distinct seasonal component to potential violations, it seems prudent to conduct these analyses prior to and after each June – September period.

The results of this study were presented by the authors to the NOAA Office of Law Enforcement Alaska Division in March 2020. The results were determined to be beneficial to their investigations and prosecution, and were then shared with General Council Enforcement Section staff in April 2020. However, it was noted that there was a 5-year statute of limitations against civil (28 U.S.C. § 2462) and criminal prosecution (18 USC § 3282), so future analyses would need to take that into consideration. In addition, it was noted by one officer that although individual data points exceeded 20% difference, the mean values for many periods listed in Table 1 were only slightly below -3, and given the degree of variation in individual measurements, there might be hesitation to prosecute a vessel based on these data. Requests for individual data points was made by the officer to see if there was an obvious pattern over time in the direction and magnitude of between scale differences. These conversations are valuable and mean that change point analyses may be most useful in prompting FMA staff to have observers write statements since these are relied on by OLE more than data for investigating potential violations. Nonetheless, these results warrant the continued data collection by observers under FMA guidance.

ACKNOWLEDGMENTS

We are grateful to all of the dedicated observers whose hard work under challenging conditions made this work possible. Thanks to Adelle Molina for providing the foundation and inspiration for this data analysis. For this work to be successful and meaningful we required many questions to be answered. Special gratitude goes to the Alaska Office of Law Enforcement (especially Jaclyn Smith), and to Rebecca Killick for helping us understand the nuances of the change point package for the R statistical programming language.

CITATIONS

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

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

Dorn, M. W., S. K. Gaichas, S. M. Fitzgerald, and S. A. 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.

Haynes K., and R. Killick. 2019. Change point.np: Methods for Nonparametric Change point Detection. R package version 1.0.1. https://CRAN.R-project.org/package=change point.np.

Killick, R., and I. A. Eckley. 2012. Change point: An R package for change point analysis. J. Stat. Software 58:1-19.

Killick R., P. Fearnhead, and I. A. Eckley. 2012. Optimal detection of change points with a linear computational cost. J. Am. Stat. Assoc. 107:1590-1598.

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.

Porter, R. 2010. Fisheries observers as enforcement assets: Lessons from the North Pacific. Mar. Policy 34:583-589.

Table 1. -- Summary of periods of concern (PERIOD) per vessel (defined by its alias; ALIAS) and observer (OBS). The mean and coefficient of variation (MEAN and CV, respectively) of the percent difference between scale weights is also shown. The field TYPE denotes the type of concern has been identified (see text for details). START and END fields are in the format (Year-Month-Day 24Hour:Minute) and are specific to the haul retrieval time associated with OBS and do not necessarily relate to the start and end of a period. For example, the start of PERIOD 13 for ALIAS NL is 2017-04-08 07:20 but the period ends in 2017-07-03 13:15. Fishing is not continuous within a PERIOD and the reader is referred to Figure 3 for visual depiction of fishing activity during these time periods. The field STATEMENT refers to whether or not an observer statement of potential violations was written during the time period (see text for details).

ALIAS	PERIOD	OBS	START	END	MEAN	CV	TYPE	STATEMENT?
RN	4	19376	2014-06-10 19:45	2014-06-30 02:10	-3.248	0.326	Red: Very low mean and CV	No
NL	12	21635	2017-03-09 14:12	2017-04-08 01:35	-3.247	0.604	Red: Very low mean and CV	No
NL	13	21635	2017-04-08 07:20	2017-04-08 07:20	-4.612	0.137	Red: Very low mean and CV	No
NL	13	21964	2017-06-14 01:40	2017-06-30 23:20	-4.612	0.137	Red: Very low mean and CV	Yes
NL	13	22041	2017-07-03 13:15	2017-07-03 13:15	-4.612	0.137	Red: Very low mean and CV	No
NL	14	22041	2017-07-03 17:05	2017-08-08 12:31	-3.515	1.286	Red: Very low mean and CV	No
NL	18	22788	2018-07-17 03:03	2018-09-03 22:38	-3.033	1.123	Red: Very low mean and CV	No
MB	14	22827	2018-07-23 08:55	2018-08-08 14:40	-4.800	1.610	Red: Very low mean and CV	No
RA	28	22613	2018-05-30 22:18	2018-06-11 08:55	-11.274	11.764	Red: Very low mean	No
RA	37	23403	2019-05-23 06:35	2019-05-30 11:53	-4.132	2.373	Red: Very low mean and CV	No
HU	4	23610	2019-06-30 10:00	2019-08-17 12:50	-4.997	1.555	Red: Very low mean and CV	Yes

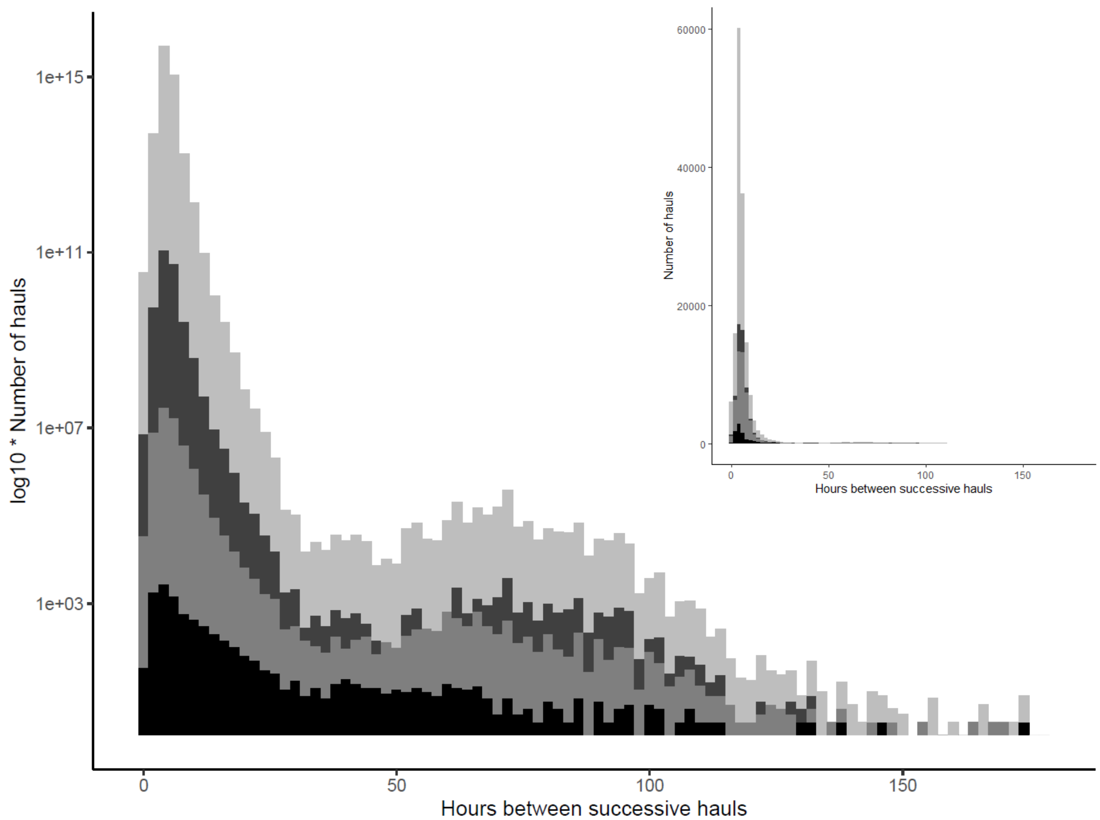

Figure 1. -- Number of hours between hauls when scale data were obtained by observers. Peaks at 5 hours and 72 hours are evident only on log10 scale. The untransformed data are presented in upper right inset.

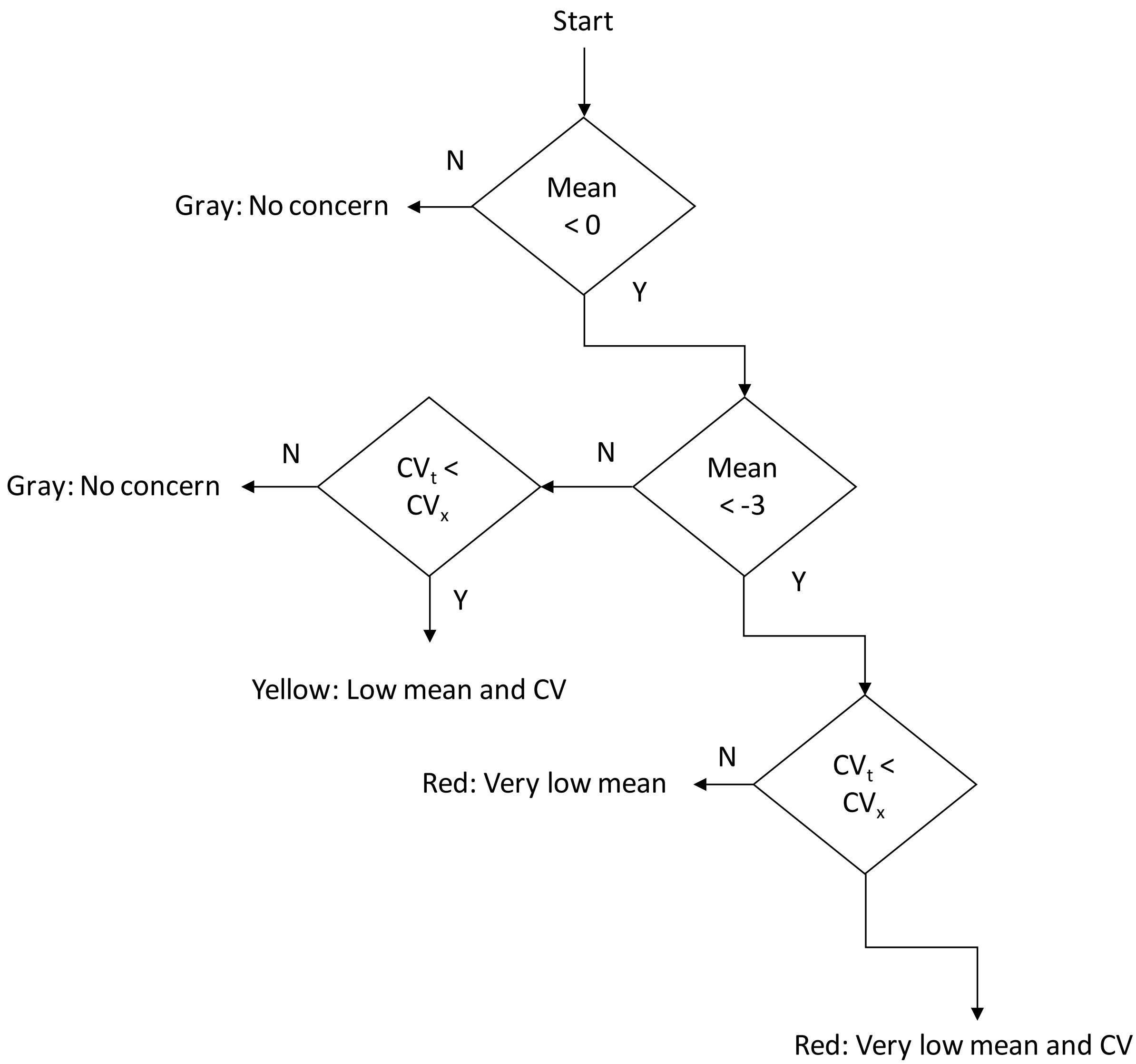

Figure 2. -- Decision tree for labelling periods of concern. Outcomes at the top of the tree are of less concern than those at the bottom (Level of concern: Red > Yellow > Gray). If the time period (t) was 1, the value for CV_x (Coefficient of Variation for period x) was the mean CV for the entire time period. If the time period was greater than 1, CVx was the CV for the prior time period.

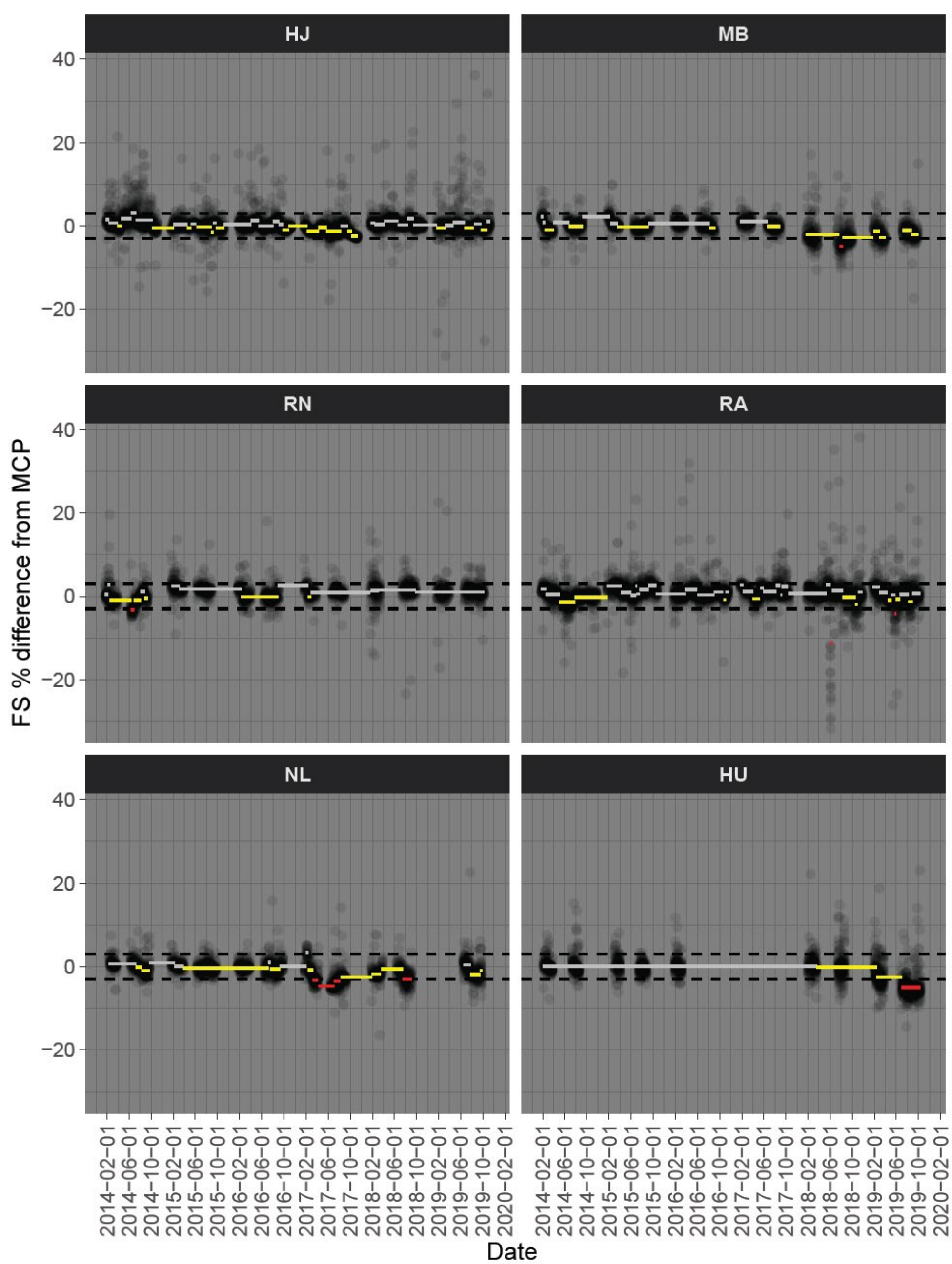

Figure 3. -- Differences in scale weights through time on six vessels denoted by their alias. Dashed horizontal lines denote the maximum permissible error tolerance in daily tests. Dots denote individual hauls while horizontal lines denote time periods identified in Change Point Analysis. Depictions of periods of concern as gray, yellow, or red follow Figure 1. The first upper left panel depicts a vessel with no problem (red) time periods, while the other five panels denote vessels with periods of concern denoted in Table 1.

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

July 2020

www.nmfs.noaa.gov

OFFICIAL BUSINESS

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