Categorising salmon migration behaviour using characteristics of split-beam acoustic data

George Cronkite¹a, Tim Mulligan², John Holmes¹ and Hermann Enzenhofer³

¹ Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC, Canada V9T 6N7
² 3876 Yellow Point Road, Ladysmith, BC, Canada V9G 1E9
³ Fisheries and Oceans Canada, Cultus Lake Salmon Research Laboratory, 4222 Columbia Valley Highway, Cultus Lake, BC, Canada V2R 5B6

Received 15 February 2007; Accepted 22 June 2007

Abstract – Milling behaviour is problematic when using hydroacoustics to estimate the number of migrating fish in rivers. Milling behaviour was observed for adult sockeye salmon migrating upstream in the Wannock River, to their spawning grounds in the tributaries of Owikeno Lake in the central coast area of British Columbia, Canada. We classified the acoustic salmon tracks to separate the milling fish from the actively migrating fish in an attempt to obtain an estimate of sockeye salmon flux as they migrate to their spawning grounds. We used discriminant function analysis and found that three variables measured on each track were sufficient for the classification of the acoustic tracks into milling and non-milling categories with an approximate classification accuracy of 98%. The method we present can also be used to separate tracks of targets of interest from noise or debris tracks that occur in the acoustic data, if discrete track characteristics are ascertained.

Key words: Split-beam / Acoustic noise / Track characteristics / Discriminant function

Résumé – Classification des types de comportement du saumon en migration en utilisant les caractéristiques des données acoustiques d’un sondeur mono-faisceau. Les comportements non-migratoires posent problème quand on veut estimer le nombre de poissons en migration dans les fleuves, en utilisant l’hydroacoustique. Ce comportement non-migratoire est observé chez les saumons rouges (Oncorhynchus nerka) lorsqu’ils remontent le fleuve Wannock pour aller vers leurs aires de ponte situées dans les affluents du lac Owikeno, sur la côte de Colombie Britannique au Canada. Nous classons les suivis (échos-traces) acoustiques des saumons afin de séparer les poissons qui ne migrent pas, des poissons en migration active, dans le but d’obtenir une estimation des flux de saumons rouges lors de leur remontée vers les zones de ponte. Nous utilisons l’analyse de fonction discriminante et nous trouvons que trois variables, mesurées sur chaque déplacement, sont suffisantes pour classer les suivis acoustiques en catégories migrante et non-migrante, avec une précision approximative de 98 %. La méthode que nous présentons peut aussi être utilisée pour séparer les déplacements de cibles présentant un intérêt, des suivis de débris et du bruit présents dans les données acoustiques, si des caractéristiques discrètes sont fixées.

1 Introduction

Split-beam hydroacoustic systems are widely used along the Pacific coast of North America to monitor the abundance of Pacific salmon (Oncorhynchus spp.) returning to their natal streams to spawn (Burwen et al. 1995; Burwen and Bosch 1996; Johnston and Ransom 1994; Johnston et al. 1993; Mulligan and Kieser 1996; Romakkaniemi et al. 1996; Osborne and Daum 1997). Target identification and classification are central to this process because the acoustic data include backscatter from targets of interest (salmon), and background noise related to bubbles and debris in the water.

A further challenge for acoustic escapement estimation is posed by the effects of fish behaviours such as milling, on target tracking and subsequent target classification (Xie et al. 2002). We define milling as non-directed movements within the acoustic beam.

Analysis of riverine acoustic backscatter data relies on target tracking techniques to partition the acoustic data into sets of observations referred to as tracks. Once these tracks are formed and false targets are removed, the number of targets, their direction of travel, and their velocity through the beam can be estimated, and other behavioural variables for classification can be computed (Blackman 1986). The track variables represent the average of several individual echoes, and therefore are more robust for classifying backscatter data.
than a parameter such as target strength, which is measured from a single echo. Target classification to distinguish between valid and invalid targets is currently accomplished manually by scrutinizing electronic echograms in which targets are tracked three-dimensionally over time as they pass through the acoustic beam. We have long recognized that this approach is so labour intensive that it is only tenable when target (fish) densities and passage rates are relatively low, and we have investigated alternatives that would improve the speed and efficiency of this analysis.

Enumerating sockeye salmon (*O. nerka*) escapement into the Owikeno Lake watershed, a major production area on the central coast of British Columbia (Fig. 1), with acceptable levels of accuracy and precision has been difficult and a variety of methods have been tested over the past 40 years. During an acoustic feasibility study on the Wannock River (the only marine outlet) in 2002, Cronkite et al. (2003) observed that the fish were actively migrating through the site (either upstream or downstream) and that a substantial number of fish exhibited near-stationary or milling behaviour in the fast current. The detection of milling or other unusual behaviours in fish at an acoustic site would normally lead to the selection of a different site for run enumeration (Enzenhofer and Cronkite 2000). However, there are a limited number of appropriate acoustic sites on the Wannock River, ruling out the possibility of moving the site to avoid the milling behaviour. Therefore, we sought a statistical method of classifying acoustic fish tracks based on their behaviour in order to remove the non-migrating group from the dataset and improve our ability to provide timely estimates of salmon escapement. In this paper, we use target tracking techniques and discriminant function analysis (DFA) to describe the behaviour and classify salmon into migratory and non-migratory groups. We used data collected on the Wannock River in 2002 to demonstrate the utility of this approach and to learn how these procedures might be used in acoustic site selection.

2 Materials and methods

2.1 Study area

The majority of Rivers Inlet sockeye salmon spawn in the tributaries of Owikeno Lake in the central coast of British Columbia, Canada (Fig. 1). The most practical way to monitor total escapement of these sockeye is by using acoustics in the Wannock River, which drains Owikeno Lake. Field reconnaissance prior to the 2002 migration identified a limited number of acoustic sites based on physical site characteristics outlined by Enzenhofer and Cronkite (2000). The selected site was approximately 300 m from the outlet of the lake at latitude 51° 40.946’ N and longitude 127° 10.652’ W.

2.2 Acoustic data collection

Data from a calibrated split-beam echo sounder provide three-dimensional target location in the acoustic beam as a function of time and allow target strength estimation (Traynor 1986; Traynor and Ehrenberg 1990) and tracking of individual fish in four dimensions (time and 3D location) (Carlson and Jackson 1980). Using the data from tracked fish, movement (vector direction and velocity) and fish flux (number of
Tortuosity

Fig. 3. Depiction of angular distance travelled in the beam, Angle-X and tortuosity, Tor-X.

2.3 Data analysis

Target echoes from the Wannock River were analysed and assigned to tracks using the Alpha Beta algorithm (Blackman 1986) as implemented in ABTrack Fish Tracker software (Pacific Eumetrics 2002). These tracks were then edited using Polaris Echogram Editor software (Pacific Eumetrics 2002) and manually classified by an experienced acoustician using 2-dimensional track projections as either migrating (i.e. upstream and downstream migrating salmon) or milling fish. The locations of the echoes belonging to a track gave the three dimensional trajectory of each target as it moved through the acoustic beam. The data from each aiming angle was treated as a separate data set and was processed independently. The use of tracks as the primary data object has two advantages. First, classification is based on the behaviour of targets as they move through the beam, e.g. target velocity. Second, grouping echoes into tracks smooths measurement errors in the positional data and provides a better description of the average behaviour of a target as it moves through the beam (e.g. as per central limit theorem; Abramowitz and Stegun 1972).

Unambiguous manually classified tracks from the Wannock River data were divided into learning and test (validation) datasets. The learning data contained tracks from 86 migrating and 81 milling fish (0.1% of complete data set). A quadratic discriminant function was derived that best separated tracks in the learning dataset into migratory and non-migratory groups (e.g. Legendre and Legendre 1983). We estimated values of the following behavioural parameters for each track: velocity in the upstream/downstream direction, X-direction (X-Speed), direction of travel in the X-direction, maximum angular distance the track covered along the X-axis (Angle-X, Fig. 3), tortuosity, the ratio of the sum of the echo to echo distances travelled by the fish to the straight line.
distance between the first echo to last echo of a track (Tor-X, Fig. 3), and X-ratio, the ratio of the sum of the echo to echo distances in X divided by the echo to echo distances in XYZ (3 dimensional path). Angle-X is independent of the target range from the transducer, whereas X-distance increases with target range because the width of the acoustic beam increases with range. When defining all of these characters, we smoothed the observed echo locations with a median smoothing algorithm to reduce measurement error (Fig. 4).

Frequency plots (example Fig. 5) of these parameters were visually inspected and Welch Modified Two-Sample t-Tests were run to determine which parameters produced significant separation between the milling and migrating fish groups in the learning dataset. Based on the results of these tests we chose Angle-X, Tor-X and X-Speed for the DFA. S-PLUS was used for all statistical analyses (S-PLUS 1999).

We tested the validity of the discriminant function on an independent subset of tracks from the Wannock data. We used the DFA classification results from the test data as an estimate of the accuracy and precision of the DFA classification procedure. The test data were also manually classified by an experienced acoustician as described above, independent of the DFA testing and these results were used to assess the performance of the DFA model.

Next we chose an additional independent data set from the Wannock data to test the performance of the classification model. This data set included tracks whose classification was not as unambiguous as those used for the learning data (Fig. 6). The test data consisted of 65 tracks from migrating fish and 23 tracks from milling fish (0.05% of complete data set). Once these tracks were classified using the DFA model developed from the learning data, we compared the DFA-classification of the test data tracks with our manual classification of the test data.

Lastly, we manually examined the DFA classification for a 2600 track subset of the Wannock River data (1.5% of complete data set) to compare these results with our manual classification of these tracks. To do this we modified the Polaris Echogram Editor to colour code the tracks by the category assigned by the DFA. We examined the trajectories from DFA-classified tracks and compared them to our manual classification. Having verified that we had achieved adequate accuracy, we proceeded to classify the entire data with the DFA model, and examined the results to determine the fish behaviour and distribution in the Wannock River.

In this application of DFA, we do not have an independent measure of the category for each observed fish track. We are therefore classifying them manually and developing an automatic classification procedure that closely agrees with our manual classification. Since we observed tracks from thousands of fish in the entire data set, the use of an automated classification procedure saved enormous time and effort. In addition, we hoped that the same DFA model might be useful for this acoustic site in future years, if the fish behaviour remained similar over time.

Finally, we examined the entire data to learn if there were other differences between the groups of migrating and milling fish. We were especially interested in differences in the acoustic measure of fish size (target strength) and differences in the spatial and temporal distributions of the fish groups in the river. We also needed to treat the numbers of tracks from the two categories differently when estimating the salmon flux, since an individual milling fish could be tracked several times, whereas a migrating fish would be tracked once.

3 Results

3.1 Migration behaviour apparent in the data

Three categories of salmon behaviour were identified in the Wannock River data from 2002. We believe these represent migrating, both upstream and downstream (Figs. 7a,b), and milling (Figs. 7c,d) fish. Milling fish were typically near-stationary for several minutes until they either slowly moved out of the beam or the acoustic system automatically re-aimed to another position. Milling fish have potential to cause bias in
the estimates if for example, they hold their position or slowly change their position such that they are viewed during a subsequent transducer aim. We have no way of determining if this situation occurred with individual milling fish, and therefore estimates of milling fish were not derived. However, this is not an issue in determining the population of salmon migrating to the spawning grounds as only actively migrating fish are used to derive the flux estimates. The actively migrating fish appeared in the beam for shorter periods of time and displayed characteristic tracks that enter the beam on one side and exit on the other. Migrating fish were seen travelling in both the upstream and downstream directions, but for the data analysed in 2002, less than two percent of the migrating fish were travelling in the downstream direction. In contrast, the milling fish generally showed a near 1:1 ratio of upstream tracks to downstream tracks, which means that estimation of direction of travel of these near-stationary fish was essentially random.

### 3.2 DFA and effectiveness of classifiers

Trials with various combinations of characters lead us to choose the absolute value of AngleX, the log_{10} of TorX and XSpeed as the final set for DFA. T-Test results showed that Angle-X, log_{10} of Tor-X and X-speed displayed significant differences between migrating and milling fish ($t = -22.5$, $df = 137.6$, $p$-value $= 0$; $t = 13.3$, $df = 88.0$, $p$-value $= 0$; $t = 13.9$, $df = 115.3$, $p$-value = 0, respectively). Using the first two variables produced similar classification results as a DF based on all three variables. However, the addition of XSpeed allowed us to differentiate upstream and downstream migrating fish. The results of the DFA demonstrate that reliable classification was achieved (Tables 1 and 2). Of the 167 tracks in the learning data, 163 were correctly classified, representing an overall classification accuracy of 97.6%. The same level of accuracy was achieved for the 2600 track Wannock River data subset when compared with the assessment of the same data set made by a trained acoustician. The 2600 track data subset was more than is required during standard operations, but we wanted to confirm that the method was functioning correctly under many different track scenarios.

The track misclassification rate was low (4 examples in 167) and misclassification occurred when the probability that a track belonged to either group differed substantially from 1. Close inspection of these tracks showed that they were of poor quality, likely due to interference from environmental noise or

![Fig. 6. Plots of the learning data and the test data used in developing and examining the results of the DFA. Plots (a) and (b) are learning data and plots (c) and (d) are test data. The light grey points represent the milling fish and the black points represent the migrating fish. Separation between these two categories is apparent in all four plots.](image-url)
loss of data due to interference between targets (see Cronkite and Enzenhofer 2002). Environmental noise and interference between targets can add or remove accepted echoes from fish tracks or cause inaccurate measurement of target position in the beam. These factors can create track information for which the direction of travel cannot be determined or for which the echoes do not meet criteria used by the echosounder for single target detection and therefore cannot be assigned to a fish target.

3.3 Site specific optimisation of the discriminant function

It is essential that the discriminant function developed for a particular application be tested for effective performance. During testing at an acoustic site at the upper end of Owikeno Lake in 2003 (Secondary Acoustic Site, Fig. 1), we used the DFA routine that was optimised for our acoustic site on the Wannock River in 2002. We knew that the routine was not functioning optimally, but we decided to optimise the routine after the data collection was complete. We obtained a net upstream flux of approximately 32,000 fish in-season but with the optimised routine, which was checked for performance by acousticians, we obtained a net flux of −380 fish post-season. The DFs developed for this site achieved approximately 92% agreement with trained acousticians and the results showed nearly equal upstream and downstream migration. Based on these data, we believe that the fish were moving freely back and forth between the two lakes and some fish may have been dropping back downstream after spawning. Once the DF was optimised for the site, we recognised this behaviour and we rejected this particular acoustic site because accurate estimates of flux were not possible.

3.4 Plots displaying salmon behaviour

Once the fish tracks have been classified into milling and migrating categories, we were then able to make some exploratory plots to learn more about these behaviours. In the Wannock River there was a tendency for spatial separation between two of the three categories of fish with the upstream migrants being closer to the right-bank and the milling fish being closer to the left-bank (Fig. 8). The downstream migrating fish appear to be randomly distributed throughout the river cross-section, although they were few in number.

4 Discussion

We have showed how DFA proved useful in separating tracks into user-defined categories, in the Wannock River case, migrating and milling fish, to allow us to calculate the flux estimates for the migrating portion of the salmon population. We acknowledge that multiple regression analysis could be used for the example we introduce, but we present the DFA method as often researchers will want to define more than two track categories. The Wannock River project was a useful test case for track classification as the fish behaviours were distinct
Fig. 8. Side-view plot displaying the mean detected fish position in the water column relative to the multiple aims of the transducer. The heavy black line indicates the river bottom contour at the acoustic site and the horizontal line at $Y = 0$ indicates the water surface. The small square indicates the location of the transducer and the diverging lines delineate the nominal beam-width of the $4' \times 10'$ transducer. The vertical lines at range from the transducer show the maximum range obtained for each transducer aim. The supplementary data from the DFA are displayed with black points representing the upstream migrating fish and white points representing the milling fish. There are 21 randomly distributed downstream migrating fish on this plot, the positions of which are not easily identified due to the limitations of the small point size and the black and white representation. Note that the scales of the $X$- and $Y$-axes are unequal, causing the vertical beam angles to appear distorted.

and allowed automatic classification of the data into two categories, which saved a great deal of time in “cleaning” the data (i.e., identifying and removing unwanted targets). This automated classification agreed closely with an independent manual classification by an experienced acoustician. Exploratory plots, such as the one shown in Figure 8, are used to check transducer aims and patterns of fish behaviour and these plots are time consuming to produce manually. However, the use of automated classification allows exploratory plots to be prepared easily, and corrections applied if necessary.

DFA can be used to determine the magnitude of problematic fish migration behaviour or debris tracks in the data that may bias flux estimates through a site. The tool requires that some data be collected to act as the learning data set before the best set of DF can be developed and applied to the collected data sets. In the worst case scenario, the most effective set of DF may have to be determined post-season, but we have shown that useful DF can be developed with a relatively small quantity of data, thus allowing the possibility of determining useful DF early in the season and applying them through the remainder of the salmon migration. With this approach, acoustic sites can be quickly assessed and rejected if problematic behaviours are excessive. In the case of the Wannock River we used DFA and achieved flux estimates, but we also determined that milling behaviour among fish in the river was of sufficient magnitude to reject the enumeration site for the long term, since this behaviour creates substantial uncertainty in flux estimates.

Large errors in the estimates of flux can occur if the DFA routines are not carefully optimised and tested for performance on the data. For example, when we tested a second site at the upper end of Owikeno Lake, the overall flux was overestimated when the DFA routine developed for the Wannock River was used. Subsequent optimisation and testing revealed very low net upstream passage along with high counts of fish migrating in both directions, leading us to conclude that the second site was not useful for measuring fish flux, as the fish were migrating upstream and downstream in nearly equal proportions between two lakes. The site was located near several spawning tributaries and it was likely that some of the fish had spawned, they moved back downstream, giving erroneous estimates of the net spawning population above the site.

We stress that each acoustic site may be somewhat unique and the data needs to be examined in detail to determine which variables may be useful in developing DF to categorise the data. We presented some examples that we found useful (Angle-X, Tor-X, X-speed) which may work well in other situations, but certainly other behavioural traits may be useful under differing circumstances, a simple example being target strength. However, target strength alone may not provide good separation between categories as we have measured a significant overlap in target strength distributions between target types, due in part to the large variability of echo to echo target strengths in the riverine environment due to the low signal to noise ratio. If target strength data are combined with other track characteristics then it can be more useful for separating tracks into categories.

Acoustic noise in the data can be identified and removed with the use of DFA as noise data often exhibits characteristics that differ from the fish track characteristics. For example, we were able to remove tracks attributable to drifting log debris simply by using an X-speed DF as the logs displayed a significantly slower average speed than the migrating fish. This is a simple example of categorising tracked data, but often a wider range of discriminant functions needs to be determined to separate the targets of interest from the noise, as we have described for migrating and milling fish. Tracks due to noise from wind, rain, entrained debris and river surface or substrate tend to exhibit characteristics that allow them to be distinguished from the tracks of migrating fish.

We believe that the DFA method is useful in interpreting riverine split-beam acoustic data. The method may also have applications in marine acoustics, allowing the determination of single fish tracks for the species of interest in the presence of acoustic noise or other species, for the determination of single fish target strengths. The DFA method may also be applied to other types of acoustic data such as imaging acoustics, anywhere that distinct traits can be determined for the data, allowing categorisation of the tracks and therefore isolation of the data of interest with minimal effort.

Acknowledgements. We would like to thank the following: Dave Noon for his assistance with the field work and post-season data analyses, Steve Bachen for his advice, assistance and support in the field, Pieter Van Will for his logistical support and advice, Rivers...
and Smiths Enhancement Society for their funding support, and the Owikeno First Nations for their assistance with field work, logistical support and local knowledge.

References

Abramowitz M., Stegun I.A. (Eds.), 1972, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. 9th printing, New York, Dover.


Racca R., 1999, Operational Notes for UWINSTRU systems with NMEA0183 interface. JASCO Research Ltd., Victoria, BC.


