Identifying when tagged fishes have been consumed by piscivorous predators: application of multivariate mixture models to movement parameters of telemetered fishes

Jason Romine[1] , Russ Perry[1], Sam Johnston[2], Chris Fitzer[3], Steve Pagliughi[4], Aaron Blake[5]
[1]USGS Western Fisheries Research Center, [2]Hydroacoustic Technology Inc., [3] Environmental Science Associates, Inc., [4] AECOM, [5]USGS California Water Science Center

I. Introduction

An inherent issue with telemetry of fishes is that they may be preyed upon during the course of telemetry studies potentially leading to incorrect conclusions about movement, behavior or survival. This problem is especially acute in western rivers of the United States where telemetered migrating juvenile salmonids may experience high mortality rates due to predation from piscivorous fishes. More specifically, our concern is with predation of telemetered emigrating juvenile salmonids by non-native Striped Bass (*Morone saxatilis*) and two species of non-native black basses, Smallmouth Bass (*Micropterus dolomieu*) and Spotted Bass (*Micropterus punctulatus*), in the Sacramento-San Joaquin River Delta (Figure 1). Here, telemetry-based survival studies (for example, [1]) assume that tag detections are from live juvenile salmonids, rather than tagged salmonids consumed by predatory fishes (hereafter, consumed smolts). Consumed smolts subsequently detected at downstream locations may lead to inflated survival estimates. Thus, in this example, it is important to differentiate between detections of live tagged smolts and consumed smolts to avoid bias in survival estimates.

To differentiate tracks of live tagged smolts from tagged smolts consumed by predators, we fitted multivariate mixture models to track statistics from a telemetry study conducted in the Sacramento-San Joaquin River Delta. We first estimated the Lévy exponent and tortuosity for each track. We then fitted a bivariate normal mixture model to these statistics to estimate the parameters of the smolt- and predator-specific distributions from the combined bivariate distribution of the track statistics. Given these distributions, we then quantified the probability that any given track exhibited characteristics that were consistent with predator- or smolt-like movement and used this information to classify the track as predator or smolt.

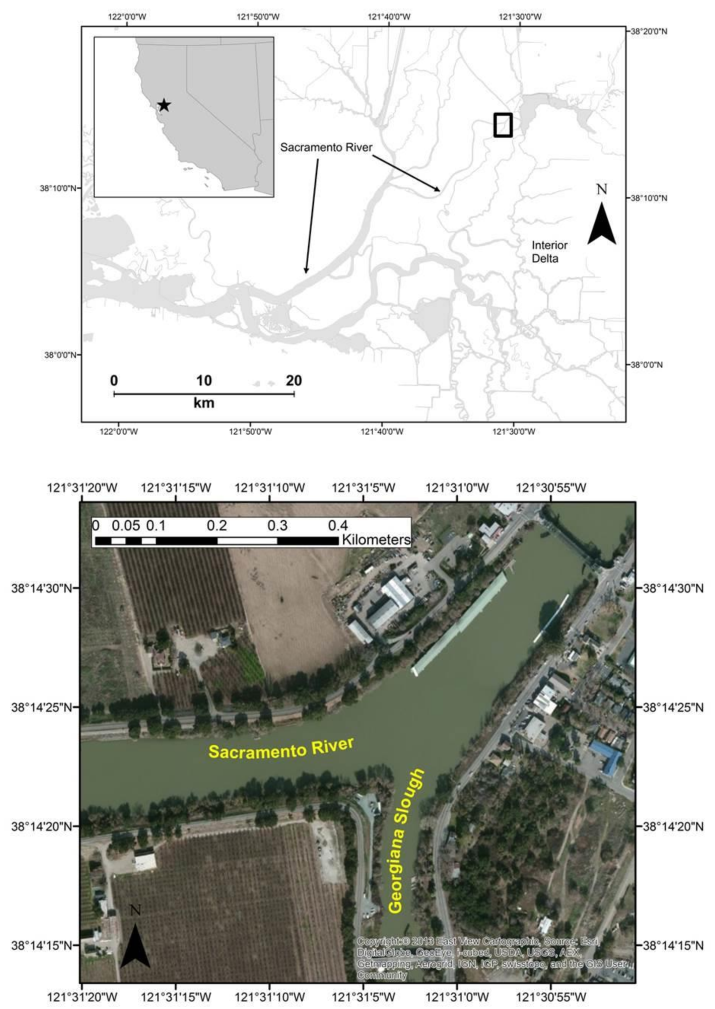


Figure 1. Maps illustrating the study are where the telemetry array was deployed.

II. Methods

Fish tracks encompassing the entire detection history of Chinook Salmon smolts, Steelhead Trout smolts, Striped Bass, Smallmouth Bass and Spotted Bass were used in the analysis. Tracks were broken into discrete track segments if the time between successive detections was greater than 30 min. Each track segment was analyzed separately. In other words, a tag that moved through the array, out of the study area, and then returned after 30 min or more was treated as two separate track segments. This resulted in some tracks consisting of multiple track segments. Tracks with fewer than 60 two-dimensional positions were omitted from the analyses. The ping rates of tags varied from 2 to 4 s. Therefore, we discretized track segments at a time step of 8 s using the adehabitatLT package in R [2] to normalize telemetry data and avoid potential bias in track statistics that might arise due to different ping rates between tags [3]. Two statistics were estimated for each track segment for each fish, tortuosity (τ) and the Lévy exponent (b). Tortuosity (τ) was calculated as a function of the turning angle (θ):

$$\tau = \sqrt{\bar{x}^2 + \bar{y}^2}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} \cos\,(\theta_i)$$

and

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} \sin\,(\theta_i)$$

Here n is the number of relocations and the turning angle (θ) is the change in direction between three successive relocations. A track with tortuosity close to one is considered linear whereas a track with tortuosity near 0.5 is more tortuous or complex.

In Lévy walks, the relation between step length (l) and the frequency of occurrence of a step length follows a power function, $f(l) = al^{-b}$, where a is an intercept parameter and b is the Lévy exponent. Lévy exponents were estimated using the logarithmic binning method following Sims *et al*. [4]. The Lévy exponent was estimated from the slope of the linear regression between log-transformed geometric bin widths and log-transformed bin frequencies of step lengths. A step length is the distance between two successive locations, and the frequency is the number of occurrences of each step length.

After track statistics were estimated for tagged smolts and predators, finite mixture models were fitted to the distributions of track statistics using the mixtools package for R [5]. Finite mixture models are a form of model-based clustering, which uses the expectation maximization algorithm to maximize the likelihood function and estimate parameters of mixed distributions for observations with unknown group membership

Table 1. *A priori* assumptions

Track statistic	Smolt	Predator
Tortuosity (τ)	Higher	Lower
Lévy exponent (b)	Lower	Higher

III. Results

- 1,412 Chinook Salmon: 155 fish with multiple segments
- 259 Steelhead Trout: 41 fish with multiple segments
- 14 Smallmouth Bass: 13 fish with multiple segments
- 6 Spotted Bass: 6 fish with multiple segments
- 29 Striped Bass: 20 fish with multiple segments

Table 2. Mixture model results. Values in parentheses are the standard errors.

Parameter	Mean	Estimate	Standard deviation	Estimate
Lévy exponent, predators	$\mu_{P,b}$	1.84 (0.033)	$\sigma_{P,b}$	1.23 (0.048)
Lévy exponent, smolts	$\mu_{S,b}$	−0.304 (0.008)	$\sigma_{S,b}$	1.46 (0.003)
Tortuosity, predators	$\mu_{P,\tau}$	0.565 (0.037)	$\sigma_{P,\tau}$	0.070 (0.048)
Tortuosity, smolts	$\mu_{S,\tau}$	0.944 (0.001)	$\sigma_{S,\tau}$	0.001 (0.0001)

Figure 2. Distributions of the Lévy exponent (A) and tortuosity (B) for smolt (red line) and predator (green line) populations estimated using a bivariate mixture model of normal distributions. The histogram shows the mixed empirical distribution of track statistics for which the true population assignment is unknown (that is, predator or smolt). The black dashed line shows the distribution of track statistics for known predators.

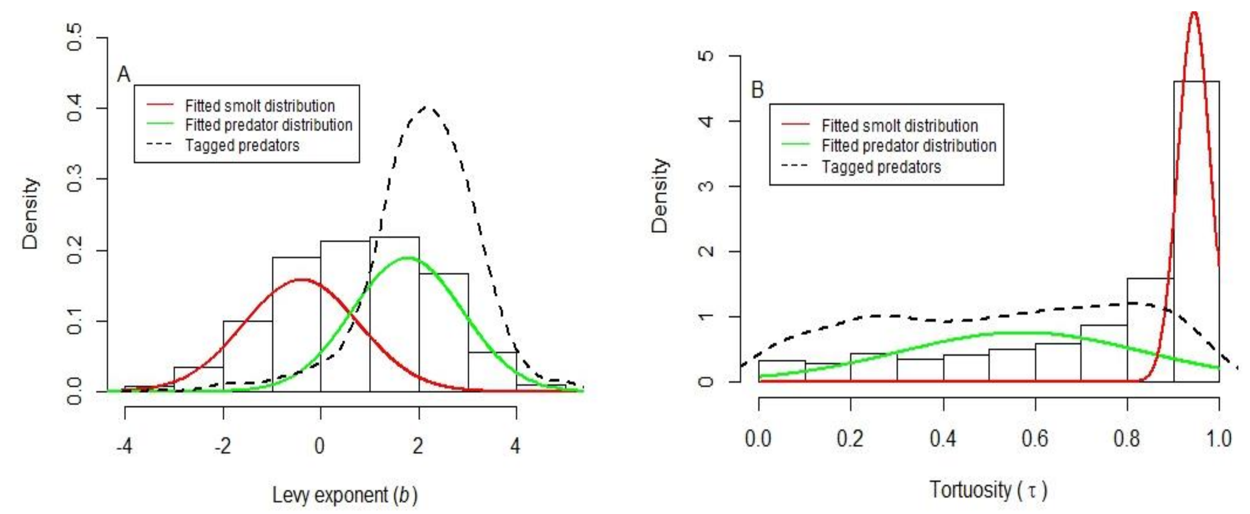


Table 3. Classification of the fish tracks based on mixture model results.

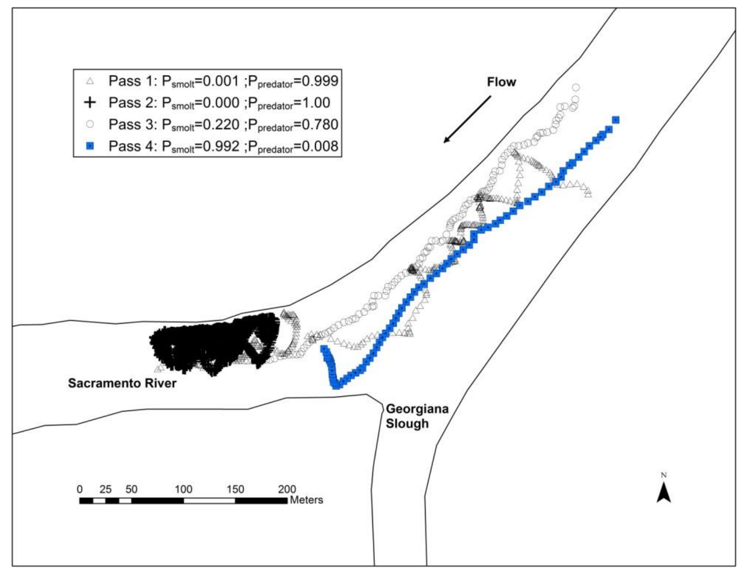
	Model classification	
	Smolt	Predator
Striped Bass	8	21
Smallmouth Bass	2	12
Spotted Bass	0	6
Chinook Salmon	1,131	281
Steelhead Trout	191	68

IV. Conclusions

We believe the mixture model approach is a sound alternative to the manual review of each track, but our approach does not need to eliminate classification schemes that include some level of manual review. Because the mixture model yields a probabilistic estimate of a track's source population, there will be regions of high certainty where a track's characteristics are consistent with those for a smolt or predator, and regions of relative uncertainty where manual review may still provide a useful "second opinion" for a track's classification (Figure 4). For example, one approach would be to divide the probability space into three equal-size regions (that is, 0 to 0.33, 0.33 to 0.66 and 0.66 to 1). Tracks falling in the central region, where the classification is less certain, could be manually reviewed and auxiliary information (for example, movement against the flow) could help inform the classification. Such an approach would provide a more systematic, quantitative method for classifying tracks while still retaining some level of manual review.

It is important to recognize that any classification method, whether statistical or manual, will be unlikely to classify tracks with 100% accuracy because both predators and smolts may exhibit multiple behavioral modes that lead to misclassification. That is, sometimes a predator track may look like a smolt track and sometimes a smolt may act like a predator. This aspect of fish behavior is captured in our mixing model as the overlap in the distributions of track statistics for predator and prey (Figure 2).Specifically, the predator distribution overlaps the smolt distribution, indicating that predator tracks sometimes resemble a smolt track (Figure 3).

Figure 3. Track of tag 2952.15, a tagged striped bass in the study area. The figure illustrates the different behaviors of a striped bass. Segment four had a higher probability of being a smolt (P_{smolt} = 0.971) than a predator ($P_{predator}$ = 0.029), whereas all other segments had higher probabilities of being a predator ($P_{predator}$ > P_{smolt}).



Bibliography

1. Perry RW, Skalski JR, Brandes PL, Sandstrom PT, Klimley AP, Ammann A, MacFarlane B: **Estimating survival and migration route probabilities of juvenile Chinook salmon in the Sacramento-San Joaquin River Delta.** *North Am J Fish Manage* 2010, **30:**142–156.
2. R Development Core Team: *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2011.
3. Laube P, Purves RS: **How fast is a cow? Cross-scale analysis of movement data.** *Trans GIS* 2011, **15:**401–418.
4. Sims DW, Righton D, Pitchford JW: **Minimizing errors in identifying Lévy flight behaviors of organisms.** *J Anim Ecol* 2007, **76:**222–229.
5. Sims DW, Righton D, Pitchford JW: **Minimizing errors in identifying Lévy flight behaviors of organisms.** *J Anim Ecol* 2007, **76:**222–229.

Link to professional page with other publications

Link to full manuscript