

Diagnosing Predated Tags in Telemetry Survival Studies of Migratory Fishes in River Systems

Rebecca A Buchanan (✉ rabuchan@uw.edu)

University of Washington <https://orcid.org/0000-0003-2510-4140>

Steven L. Whitlock

University of Washington

Research Article

Keywords: False Positive, Pattern Recognition, Predator Filter, Salmon, Survival, Tag Predation, Telemetry

Posted Date: October 25th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-991320/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Animal Biotelemetry on April 1st, 2022. See the published version at <https://doi.org/10.1186/s40317-022-00283-1>.

Abstract

Acoustic telemetry is a powerful tool for studying fish behavior and survival that relies on the assumption that tag detection reflects the presence of live study subjects. This assumption is violated in cases where tag signals continue to be recorded after predators have consumed tagged study subjects. When such tag predation is possible, it becomes necessary for researchers to diagnose and remove these nonrepresentative detections. Past studies have employed a variety of data filtering techniques to address the issue, ranging from rule-based algorithms that rely on expert judgements of behavior and movement capabilities of study subjects and their predators to automated pattern-recognition techniques using multivariate analyses. We compare four approaches for flagging suspicious tracks or detection events: two rule-based expert-opinion approaches of differing complexity and two unsupervised pattern recognition approaches with and without data from deliberately tagged predators. We compare and illustrate trade-offs between alternative approaches by applying these filtering techniques to a case study on survival estimation of acoustic-tagged juvenile Chinook salmon (*Oncorhynchus tshawytscha*) in the San Joaquin River, California, United States. Filtering approaches differed in the number and composition of tags suspected of being consumed by predators; the largest differences occurred between the two broad categories, rule-based versus pattern recognition. Despite the appeal of the pattern recognition procedures' automated nature, those approaches nevertheless required investigator judgement in interpreting results. All methods flagged a small subset (5%) of suspicious tags that had exceptionally long residence times and evidence of upstream transitions; 27% of tags showed evidence of predation based on at least one filter. The complex rule-based filter deemed the most tags suspicious (21%) and the simpler pattern recognition method the fewest (10%). Reach-specific survival estimates from the four filters were mostly within 2% of the unfiltered estimates but differences up to 11% were observed. Sensitivity of survival results to tag predation and predator filtering depends on the study setting, spatiotemporal scale of inference, and habitat use of predators. We recommend that survival studies include clear documentation of filtering methods and report on robustness of results to the filtering approach selected.

Background

Aquatic telemetry is a powerful tool for studying the behavior and demographics of fish populations (Drenner et al. 2012; Cooke et al. 2013; Hussey et al. 2015). The value of this technology stems partly from its ability to track uniquely identifiable fish through time and space without requiring either investigator proximity or repeated handling of the study subject after tagging and release (Lucas and Baras 2000). However, spatiotemporal separation raises the possibility that not all tags are tracking a living subject in the manner intended. Tags may be detected after expulsion or displacement, or the tag may be temporarily transferred to a predator after consumption, an event we refer to as "tag predation" (Sandstrom et al. 2013; Muhametsafina et al. 2014; Havn et al. 2017; Jepsen et al. 1998; Evans et al. 2016). The latter possibility results in the predator's behavior being recorded instead of that of the study subject. Both possibilities may produce an inaccurate depiction of the behavior and survival of the study subject and may result in biased inferences if false detections are not accounted for in data analysis (Klinard and Matley 2020).

Tag predation is identified as a potential source of bias in some studies of survival and movement of migratory fishes, and a variety of procedures have been used to identify and remove this type of false detection data; see Klinard and Matley (2020) for a review of methods used to identify tag predation and other forms of mortality in telemetry data sets. The most direct predator diagnoses arise from physical recovery and examination of the tagged individual. In most studies, tag recovery is incidental so indirect diagnosis methods are required. Procedures used to flag purported predator detections have sometimes been referred to as "predator filters" (Buchanan et al. 2018). The bulk of existing predator filters may be classified as either behavior-based or signal-based. Behavior-based filters depend on the assumption of behavioral differences between the focal species and likely predator species (e.g., March et al. 2010; Buchanan et al. 2018; Daniels et al. 2018). Signal-based filters identify predation in the tagging data using changes to the signal generated by the tag; for example, newly developed acoustic tags switch to the "predator signal" when a specially designed coating breaks down in the gastrointestinal tract of a piscivore (Halfyard et al. 2017). This paper addresses behavior-based filters.

Our definition of predator filter is intentionally broad, ranging from simple decision rules to complex algorithmic or statistical approaches. We distinguish between a predator filter and subjective data manipulation by requiring predator filters to be formalized, documented, and applied systematically. In all cases, the effectiveness of a given filter is predicated on the assumption that a distinct behavioral contrast exists between the study subject and their predators and that behavioral differences are detectable on the spatiotemporal scale of the data (Klinard and Matley 2020).

The ability to identify predator detections in a telemetry investigation varies based on the quantity and quality of data available. Typically, the available data are sequences of tag detection times at various fixed-site or mobile acoustic receivers. Movement metrics derived from each sequence, such as migration rate and direction, form the basis of a behavior-based predator filter. Tag detection data may be augmented by environmental data and detections from past or concurrent predator-tagging studies (Thorstad et al. 2012; Gibson et al. 2015). In most cases, some knowledge of the behavior and swimming capabilities of the focal species is required, possibly combined with understanding of predator behavior and gut evacuation rates (Beland et al. 2010; Schultz et al. 2015; Thompson et al. 2015).

The literature addressing the tag predation problem is limited. The types of predator filters previously used include fine-scale assessment of the acoustic signal pattern in the raw telemetry data file (Vogel 2010), 2-D or 3-D tracking of the tag's position in space (Romine et al. 2014), simple or multifaceted decision rules based on expert scientific opinion (Buchanan et al. 2018; Hause 2020), outputs from multivariate clustering procedures (Gibson et al. 2015), random forest classifications (Daniels et al. 2018), multivariate mixture models (Romine et al. 2014), or some combination thereof (Perry et al. 2018). These filtering approaches vary in their spatiotemporal scale (e.g., 2-D tracks, detection events, or the tag's complete detection record), their degree of complexity (e.g., the number and type of metrics used), and whether they incorporate auxiliary information such as environmental data or tag recoveries. A key difference among filters lies in whether tags are seen as suspicious based on exceedance of explicit biological thresholds (rule-based) as opposed to a statistics-based measure of dissimilarity. Rule-based filters use biological knowledge or expert judgement to define criteria that must be met for a detection event to be considered representative of the focal species. A statistics-based filter may generate a set of data-driven rules or may estimate a state (focal species versus

predator) probability or assignment based on patterns in the data. Among statistics-based filters, pattern recognition filters use numerical methods to identify typical behavioral patterns for a population and then flag tags showing aberrant behavior.

The choice of predator filter has the potential to affect study outcome and understanding of the system under investigation, depending on the objectives and scale of the study. However, diagnosing predated tags can be both subjective and time consuming, and little guidance has been provided on when a predator filter is advised, which predation identification procedure to use, and the degree to which study inference may be affected by the predator filter. A comparison among filter types is warranted because of the variable effort required by the different approaches and the potential impact on survival estimation.

In this paper, we discuss considerations in designing and implementing a predator filter. Our context is the common case of studying survival of juvenile migratory fishes in a regulated river using implanted microacoustic tags (McMichael et al. 2010; Klimley et al. 2013; Crossin et al. 2017). In studies of this type, tagged fish released at one end of the study area migrate past a series of interrogation arrays, and their migration survival is characterized using a Cormack-Jolly-Seber (CJS) survival model (Burnham et al. 1987; Skalski et al. 1998). In the remainder of this paper, we assume tagged fish are migrating downstream. Although focused on migration survival studies using acoustic telemetry, the factors we consider are also applicable to other types of telemetry studies. We outline issues pertaining to existing methods of identifying tag predation, limitations to diagnosing predator detections, and potential impacts on survival estimation. We demonstrate two variations on both rule-based and statistics-based predator filters using a case study of subyearling Chinook salmon *Oncorhynchus tshawytscha* in the San Joaquin River of California. The four predator filters we develop are: a simple rule-based filter, a complex rule-based filter, and pattern recognition filters with and without the inclusion of auxiliary movement data from tagged predators. We present novel approaches for assigning detection-level predation events as part of a pattern recognition filter. We compare the behavior of alternative filtering approaches based on the number and pattern of purported predator detections and examine the consequences of each filter using local- and region-scale survival estimates. Finally, we provide guidance on selection of predation identification methods.

Methods

Four alternative predator filters were constructed and implemented for a case study of juvenile Chinook salmon *Oncorhynchus tshawytscha* emigrating from the San Joaquin River in California, United States. The study area, tagging study, data processing, and predator filters are described below.

Study Area

The Sacramento–San Joaquin River delta (“Delta”) of California, United States, is a tidally influenced dendritic inland estuary that connects the productive agricultural region of the Central Valley of California to the San Francisco Bay and the Pacific Ocean (Figure 1). Four runs of Chinook salmon and anadromous rainbow trout *Oncorhynchus mykiss* (steelhead) migrate through the Delta as both juveniles and adults, and all have suffered severe declines in the last century (Moyle 1994). Two runs of salmon and Central Valley steelhead are listed as endangered or threatened under the U.S. Endangered Species Act (1973). The other two salmon runs form the basis of the California–Southern Oregon ocean salmon fishery and are an important component of the prey base for endangered ocean resident populations such as Southern Resident killer whales *Orcinus orca* (NMFS 2009). The Delta also hosts populations of non-native piscivorous fish that feed on native fishes including juvenile salmonids; species suspected to be important predators of juvenile salmonids in the Delta include striped bass *Morone saxatilis*, largemouth bass *Micropterus salmoides*, smallmouth bass *Micropterus dolomieu*, and channel catfish *Ictalurus punctatus* (Cohen and Carlton 1998; Nobriga and Feyrer 2007; Conrad et al. 2016; Grossman 2016; Michel et al. 2018). Additionally, the Delta provides municipal and agricultural water for millions of California citizens, and the challenge of managing water resources as well as protected fishes has focused research attention on the magnitude and sources of mortality of juvenile salmonids in the Delta.

Researchers have used acoustic telemetry and a dense array of fixed-site telemetry receivers to study salmonid survival in the Delta for more than a decade, and the possibility of tag predation within these studies has been recognized as a potential complication (Vogel 2010). The combination of both resident and transient predators, complex hydrodynamic and migratory pathways, and closely spaced telemetry stations increases the likelihood of observing tags after a predation event. Tidal influence in the Delta further complicates the problem, as incoming tides can reverse the flow of the river for extended periods twice a day. These factors make the salmonid telemetry studies in the Delta candidates for predator filters and an appropriate testing ground for novel analytical approaches. We focused our case study on a subset of the data collected in a multiyear salmon study in the southern Delta.

In 2016, an acoustic tagging study of juvenile fall-run Chinook salmon was implemented to estimate survival of these subyearling fish as they emigrated from the San Joaquin basin through the southern Delta (See Tag Predation Supplementary Materials, Supplemental Methods). In April of 2016, 648 hatchery-reared fall-run Chinook salmon were surgically implanted with Vemco V4-180 kHz microacoustic transmitters (“tags”) and released into the San Joaquin River at Durham Ferry, located approximately 20 river km (rkm) upstream of the Delta entrance at Mossdale Bridge and 195 rkm upstream of the Golden Gate Bridge (Figure 1). Surgery, fish handling methods, and fish release methods followed the standard operating procedure outlined in Liedtke et al. (2012). The tags were monitored on fixed-site acoustic receivers located throughout the lower river and Delta (Figure 1, Table 1). An in-tank tag-life study was conducted in April and May in 2016 using 45 randomly selected tags. Additionally, telemetry detection data were available from 37 predator specimens (24 largemouth bass, 5 striped bass, 4 channel catfish, and 4 white catfish *Ameiurus catus*). These known tagged predators had been part of a group of 300 predatory fish tagged with Vemco V9 long-lived acoustic tags in prior telemetry studies in 2014 and 2015 (Hayes et al. 2017).

Table 1

Telemetry stations and regions of lower San Joaquin River and Delta used in predator filter in 2016 Chinook salmon case study. See Figure 1 for location of study areas. Tags were released at Durham Ferry (A0). CRB Filter Region = regions defined for Complex Rule-based predator filter (See Tag Predation Supplementary Materials). Core area = all stations except A1, B2, and C1.

Telemetry station	Name	CRB filter region
A1	Durham Ferry Upstream	I
A2	Durham Ferry Downstream	I
A3	McMullin Road 1	I
A4	McMullin Road 2	I
A5	Banta Carbona	I
A6	Mosssdale	I
A7	River Islands Parkway	I
A8	Lathrop	IIA
A9 ^a	Undine Road	IIA
A10 ^a	Dos Reis Park	IIA
A11 ^a	Frewert Road	IIA
A12 ^a	Brandt Bridge	IIA
A13 ^a	Howard Road	IIA
A14 ^a	Weston Ranch	IIA
A15 ^a	French Camp Slough	IIA
A16	Garwood Bridge	IIA
A17	Navy Bridge	IIA
A18	Calaveras River	IIA
A19	Shipping Channel	IIIA
A20	MacDonald Island	IIIA
B1	Old River East	IIB
B2	Old River South	IIB
C1	Head of Middle River	IIB
R1	Rough and Ready Island	IIA
T1	Turner Cut	IIIA
a = receivers used in associated predator study (Hayes et al. 2017; Michel et al. 2020)		

The southern Delta is a complex habitat of interconnected river channels that includes multiple migration routes (Figure 1). Survival in the San Joaquin River from the region near Mosssdale Bridge (“Mosssdale”) to the Turner Cut Junction (37 rkm from Mosssdale) has been found to be variable and positively related to river discharge entering the Delta from upstream (Buchanan and Skalski 2020). This region also has a sizeable population of non-native predatory fishes (Michel et al. 2018). Misclassification of the predation status of detected tags in this region has the potential to bias survival estimates and the estimated relationships between survival and covariates such as river discharge. For this reason, the spatial focus for this paper was the mainstem San Joaquin River from Durham Ferry downstream to the Turner Cut Junction. This region included the distributary point of Old River (“head of Old River”) where it branches off from the San Joaquin 5 km downstream of Mosssdale.

Tag detections on fixed-site acoustic receivers were used both to assign predation status at the time of the detection events and to estimate survival. The acoustic receivers were grouped into telemetry stations; each telemetry station consisted of one or more receivers to ensure complete coverage of the river channel. Some stations used paired lines of receivers (“dual array”) to enable or enhance detection probability estimation (Figure 1). Telemetry stations were located throughout the San Joaquin River from 1 km upstream of the Durham Ferry release site (station A1) to 3 km downstream of the Turner Cut Junction (A20) and in key channels off the mainstem San Joaquin River: at Rough and Ready Island near Stockton, CA, (R1), in Turner Cut (T1, 1.2 km from the junction), in Old River within 1.0 km of its source (B1), and in Old and Middle rivers within 0.6 km of the source of Middle River (B1, C2). Survival was estimated using detections from all San Joaquin River stations downstream of the release site and from the B1, R1, and T1 stations (“core area”). The

predator filter used detections from all stations but included detections from outside the core area (stations A1, B2, and C1) only if the tag was subsequently detected in the core area

Data Processing

The raw telemetry detection data were first filtered for false positive signal detections (e.g., multipath) and compiled into receiver-specific detection events by researchers at the U.S. Geological Survey (USGS) in Sacramento, California, United States. The USGS Great Lakes Science Center further processed the receiver detection event data into station-specific detection event data in which data from receivers within the same station were pooled. Consecutive detection events at a given telemetry station were separated either by detection at a different station or by a time gap of ≥ 12 hours without detection at the same station. The remaining data processing and analysis was performed by the University of Washington. Water velocity, river stage, and river discharge observations from river gaging stations located throughout the study area were retrieved from two state databases: the California Data Exchange Center (cdec.water.ca.gov/selectQuery.html) and the California Water Data Library (www.water.ca.gov/waterdatalibrary). The river data were recorded at 15-minute intervals; these data were cleaned for obvious errors as described in Buchanan et al. (2021).

The detection event data were passed through one of four distinct predator filter procedures: (1) a simple rule-based filter, (2) a complex rule-based filter, (3) a pattern recognition (cluster analysis) filter restricted to detections of salmon tags (“smolt-only”), and (4) a pattern recognition filter augmented with telemetry data from known predators (“multispecies”). All four filters assigned predator status to individual detection events, but the methods for assigning these events differed between the two primary approaches (rule-based versus pattern recognition). Additionally, tag predation could be assigned at either the beginning or end of a detection event for the two rule-based filters but only at the beginning of a detection event for the two pattern recognition filters. The simple rule-based filter consisted of a small set of binary rules based on several movement metrics, any violation of which signaled tag predation and a predator classification for the tag from that point onward. The complex rule-based filter used additional metrics and a score-based system that classified tags as predated if they showed evidence of unexpected behavior across multiple criteria. It was also distinct from the other filters in that rules were spatially explicit, meaning that thresholds for some metrics depended on the region within the study area. The two pattern recognition filters used common multivariate statistical methods to identify deviant clusters of tags based on patterns of observed movement metrics. The smolt-only pattern recognition filter used statistical procedures to identify groups of suspicious tags based on the observed variation in the metrics data. The multispecies pattern recognition filter used behavioral similarity with known predators as a basis for flagging suspicious smolt tags. Each predator filter was developed and implemented independently, producing four subsets of the detection events that were classified as coming from predators, one for each predator filter. The outputs from the four filters were compared using the total number of tags classified as predated by the end of the detection history and the number of first-time predator classifications that occurred in various regions of the study area: upstream of Mossdale (stations A1–A5); Mossdale through the head of Old River (A6–A8/B1); Undine Road to Howard Road (A9–A13); Weston Ranch (Stockton) to Navy Bridge/Rough and Ready Island (A14–A17/R1); Calaveras River to the Turner Cut Junction (A18–A20/T1); and the head of Middle River (B2, C1) (Figure 1).

The output of each of the filters was further processed by truncating each detection sequence at the last detection event preceding the purported tag predation. The resulting four sets of filtered detection event data were used for survival estimation for the entire study area and for the regions named above. For each predator filter, the filtered data were converted into detection histories that represented the final fate of the tagged fish at each river junction and telemetry station. The detection histories were analyzed in a CJS release-recapture model that included route selection at the head of Old River (Figure 2) to estimate reach-specific and regional survival. Sparse detection data at the R1 and T1 telemetry stations prevented estimation of detection probabilities unique to those sites, and so detections at those sites were pooled with detections at stations A17 and A20, respectively. Tag survival data from the in-tank tag-life study were modeled using the 2013 version of the four-parameter vitality model (Li and Anderson 2013; Skalski and Whitlock 2020), and the resulting tag survival probabilities were used to adjust fish survival estimates for premature tag failure using the methods of Townsend et al. (2006). Survival was also estimated using unfiltered observation data for comparison to the filtered data. The results of the five data sets were assessed by comparing estimates of cumulative survival moving downriver from the release site and regional survival.

Data processing and filtering were conducted in R (R Core Team 2020) using the “cluster” package (v. 2.1.0) for silhouette estimates and the “shipunov” package (v. 1.14) to assess bootstrap cluster stability. The release-recapture model was fit using maximum likelihood in the software Program USER (Lady and Skalski 2009).

Simple Rule-based Filter

The simple rule-based predator filter was constructed using expert opinion to define five rules for classifying detections as coming from either salmon or predators. Metrics and/or rules were derived from methods reported in other salmon survival studies in the Delta. The metrics were: (1) distance traveled upstream on a single upstream trip (Hause 2020), (2) migration rate during each transition between detection events (Buchanan et al. 2018; Johnston et al. 2018), (3) water velocity and river discharge at the start of each detection event (Buchanan et al. 2018; Perry et al. 2018), (4) cumulative average distance traveled per day since release (Perry et al. 2018), and (5) total time spent in the vicinity of a telemetry station (cumulative sum of near field residence time; see Complex Rule-based Filter; Perry et al. 2018). Each metric was computed for each detection event. Water velocity and river discharge at the start of each event were interpolated from the nearest 15-minute velocity and flow observations at the closest gaging station.

Detection events were classified as representing predators if any of the following conditions were met: (1) the distance traveled upstream on a single trip was ≥ 16 km (Hause 2020); (2) an upstream transition had migration rate $> 3 \text{ km}\cdot\text{h}^{-1}$ and was not immediately followed by a downstream transition (Johnston et al. 2018); (3) the tag was observed to move upstream against the direction of river flow (Perry et al. 2018; Hause 2020); (4) the cumulative average distance traveled per day since release was $< 1 \text{ km}\cdot\text{d}^{-1}$; and (5) the total time spent in the vicinity of a station was > 36 h. Rule violations resulted in predator classifications at either the beginning (rules 1–4) or end (rule 5) of the detection event. Upstream movement against the direction of river flow was defined by water velocity $> 0.15 \text{ m}\cdot\text{s}^{-1}$ at the start of an upstream detection event. For telemetry stations A1–A5, water velocity data were unavailable, and river discharge

$> 28.3 \text{ m}^3 \cdot \text{s}^{-1}$ ($1000 \text{ ft}^3 \cdot \text{s}^{-1}$) at the VNS gaging station defined upstream movement against the flow. Neither velocity nor river discharge data were available for the Lathrop telemetry station (A8). Chronological detection data were truncated at the first event classified as predated.

Complex Rule-based Filter

The complex rule-based filter augmented the threshold approach of the simple rule-based filter by using additional metrics, greater specificity and spatial dimension of the rule set, and a more sophisticated score-based system for classifying tags. A similar complex rule-based filter has been used for multiple years of telemetry studies of Chinook salmon in the Delta (Buchanan et al. 2018). The filter used 24 metrics to characterize the behavior of tagged fish and compared observed metrics to criteria determined by expert opinion, literature review, past filters used for studies in this system, and calibration based on examination of the 2016 detection data (See Tag Predation Supplementary Materials, Table S1, Table S2). Physical recovery of five acoustic-tagged salmon in trawl sampling near Mossdale (A6) informed criteria for the 2016 filter (unpublished data). The criteria were spatially defined and fell under nine categories, each of which included one or more metrics: residence time on three spatiotemporal scales (near, mid, and far field), travel time since release, reach-specific migration rate, migration rate scaled by water velocity and fish length (body lengths per second, BLPS), upstream-directed transitions, movements against the direction of water flow, and regional patterns of movement (Table S2). More details on the metrics and filter criteria are provided in the Supplementary Materials.

A score was assigned to each detection event, representing the number of criteria categories violated for that event (range = 0–9). Violations of one or more criteria within a single category increased the event's score by 1. Detection events that earned total scores ≥ 2 were assigned a predator classification, as were all subsequent detections of the tag. The first predator classification could be assigned at either the beginning or the end of a detection event. Assignments to the beginning of an event resulted from violations of the criteria for travel time, migration rate, BLPS, upstream transitions, movements against flow, and unexpected transitions. Assignments to the end of an event resulted from violations of the criteria for residence time and, in some cases, apparent movements against flow or unexpected transitions.

Smolt-Only Pattern Recognition Filter

We applied a pattern recognition filter to detection data using procedures similar to past studies (e.g., Gibson et al. 2015; Perry et al. 2018) and expanded on these methods in a novel way. We used hierarchical cluster analysis to identify individual tags that exhibited suspicious behavior during their detection history. We built upon this approach by following the tag-level cluster analysis with a post hoc analysis to identify the specific detection events when tags first exhibited sufficient abnormal behavior to be classified as predated.

Similar to the rule-based filters, the pattern recognition filter relied on a set of behavioral metrics. We initially computed a list of candidate metrics at the detection-level taken from those used in the complex rule-based filter, and then aggregated these metrics using one or more summary statistics. Each summarized metric reflected a hypothesis about a specific behavioral scenario that might indicate predator-like behavior. For example, migration rate was summarized using the tag-level mean and maximum, based on the hypothesis that deviations in mean migration rate or a stark difference between mean and maximum might reflect a sudden change in movement behavior indicative of tag predation. See Tag Predation Supplementary Materials, Table S3 for a complete description of metrics and hypotheses. After generating an initial list of candidate metrics, we omitted those not suited for a pattern recognition analysis using conventional multivariate statistical methods, such as categorical fields or those with missing data (Legendre and Legendre 2012). We used plots of data distributions and correlograms to determine which variables to remove and applied log transformations to a subset of variables to reduce their skew. The final data set contained 20 explanatory variables. We standardized each metric by subtracting the mean and dividing by the standard deviation.

We then implemented a cluster analysis using the Ward (1963) hierarchical clustering method, based on the squared Euclidean distance matrix (Gibson et al. 2015; Perry et al. 2018). This unsupervised learning algorithm partitioned the tags based on similarities in multivariate patterns of the observed metrics. We used two common statistical techniques to determine the optimal number of clusters: a discontinuity (“elbow”) in the plot of the proportions of variance explained and the full average silhouette statistic (Kaufman and Rousseeuw 2009). After selecting the number of clusters, we assigned predated status to cluster members based on associated behavior patterns and the assumption that a minority of tags were likely to have been consumed before the end of their detection histories. Let T_S indicate the subset of tags classified by the cluster analysis as smolts (non-predated) and T_P the subset of tags classified as predated by the end of the study.

We followed the cluster analysis with a principal components analysis (PCA) in order to examine the behavior of the clustering algorithm and the relative importance of specific metrics for defining clusters. PCA uses the same reference space as the Ward cluster analysis and both methods are based on minimizing (Ward) or decomposing (PCA) the variance of a distance matrix (Murtagh and Legendre 2014). We used the correlation coefficients between the first few principal component axes and the scaled metrics (i.e., loadings) to identify variables most influential in spreading the data in multidimensional space and used resampling procedures to examine clustering stability (Peres-Neto et al. 2003).

After classifying tags into smolt and predator categories, we then used a recursive series of PCA ordinations to identify the detection event for each tag in the T_P set that suggested tag predation. For each increment i in the PCA series, we subset the raw event-level data set to the first i detection events for each tag, where i ranged from 1 through the maximum number of detection events among all tags; all behavioral metrics were summarized and standardized on this incomplete data set and the principal components were computed. We plotted and tracked the relative ordination position of all tags using the value of the first and second principal components. Because all tags were in live smolts at the time of release, the cloud of points begins clustered and then shifts in space as cumulative metrics evolve for each tag with additional detection events, slowing as they settle into their final positions based on the full analysis. We tracked tags that were classified as predators (T_P) and identified the detection event when each tag first moved outside of the bivariate distribution of expected smolt behaviors as defined by the first and second principal axes. For each increment i , we estimated that bivariate distribution using a multivariate

normal distribution defined by the mean ordination $\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix}$ and covariance of the first two principal component values for the T_S tags (i.e., classified as smolts). A significant deviation from that distribution was defined as departure from the ellipse defined by the 95% confidence interval contour about the mean ordination $\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix}$ based on the chi-square distribution with two degrees of freedom (Johnson and Wichern 2007). Because the method is applied recursively, sequentially increasing the number of detection events with each increment, and uses movement beyond a bivariate ellipse as the criterion for estimating the tag predation event, we refer to this event-level assignment method as the “recursive ordination ellipse” (ROE) method.

Because the orientation of the 95% ellipse in the ordination may be unstable for small values of i , we defined a minimum number of events (“event minimum”) required before a tag could be considered abnormal. We defined the event minimum as the increment i_0 ($i_0 \leq 5$) when major rotations of the ROE ellipse ceased and the range of observed values on the principal component axes stabilized. Tags were classified as newly predated upon the first event i ($i \geq i_0$) when their ordination was outside the ROE ellipse boundary. Only tags in the T_P subset were assigned predation events.

Multispecies Pattern Recognition Filter

The multispecies pattern recognition filter used mostly the same methods as the smolt-only pattern recognition filter but augmented the smolt-tag data with detection data from 37 acoustic-tagged known predatory fish. We used the same behavioral metrics as in the smolt-only pattern recognition filter, now defined for both the smolt tags and the known predator tags (i.e., summarizing and standardizing the combined detections of smolt and known predator tags). As in the smolt-only version, we assumed that even if a large number of tagged smolts were consumed during the study, a small (but potentially influential) minority of those tags would be detected after consumption. Thus, we determined the number of clusters by jointly minimizing both the total number of clusters and the heterogeneity of tag identities within clusters. We were particularly interested in defining clusters characterized by normal behavior patterns (closer to the origin of the PCA axes) in which smolt tags appeared as a large majority (i.e., > 90%). All smolt tags assigned to clusters disproportionately composed of known predator tags (>10%) were classified as predated by the end of their detection history (T_P subset of smolt tags); smolt tags assigned to the remaining clusters were classified as smolts (T_S). We used the ROE method to identify the first detection event when the T_P tags were first classified as predated as described above.

Results

A total of 177 tags were classified as predated by one or more predator filtering approaches, ranging from 66 tags flagged by the smolt-only pattern recognition filter to 139 flagged by the complex rule-based filter (Table 2). Detailed detection, travel time, and survival results from the filtering approaches are described below, including the degenerate no-filter approach.

Table 2

Summary of tag detections at key locations and output of rule-based (RB) and pattern recognition (PR) predator filters for 2016 San Joaquin Chinook salmon case study. Only detections in core area or occurring before detections in core area were included. Core area = stations A2–A20, B1, R1, and T1, omits stations A1, B2, and C1 (Figure 1).

Statistic	No filter	Simple RB filter	Complex RB filter	Smolt-only PR filter	Multispecies PR filter
Number of tags released (A0)	648	648	648	648	648
Total number of tags classified in predator	0	93 (14%)	139 (21%)	66 (10%)	89 (14%)
Total number of detection events classified in predator		314 (8%)	396 (10%)	324 (8%)	386 (9%)
Number of first-time predator classifications					
Upstream of Mossdale (A1–A5)		12	17	5	19
Mossdale through head of Old River (A6–A8/B1)		39	67	15	20
Undine Road to Howard Road (A9–A13)		30	37	25	28
Weston Ranch to Navy Bridge/Rough and Ready Island (A14–A17/R1)		5	9	19	20
Calaveras River to Turner Cut Junction (A18–A20/T1)		7	8	2	2
Head of Middle River (B2, C1)		0	1	0	0
Number of total tag detections ^a :					
Banta Carbona (A5)	470	467	467	470	465
Mossdale (A6)	416	415	415	416	415
Lathrop (A8)	250	249	246	249	247
Old River East (B1)	32	32	32	32	32
Brandt Bridge (A12)	108	106	105	104	99
Garwood Bridge (A16)	79	78	77	63	63
Turner Cut Junction (A20/T1)	26	24	26	20	20
a = Tag detection counts at all stations are shown in Tag Predation Supplementary Materials, Table S4.					

No Filter

Of the 648 acoustic tags implanted into juvenile salmon and released at Durham Ferry (A0) in 2016, 618 were detected downstream of the release site, 6 were detected upstream of the release site (A1), and 30 were not detected after release. Three of the tags detected at A1 were also detected downstream; the remaining three tags detected at A1 were excluded from analysis. A total of 250 tags were detected at Lathrop (A8) and 26 tags were detected at the Turner Cut Junction (A20, T1). A total of 32 tags were detected entering Old River (B1), and 1 of these tags was later observed in the San Joaquin River at Lathrop or downstream. A total of 22 tags were detected near the head of Middle River (B2, C1); 1 of these tags was subsequently detected at B1 or in the San Joaquin River, and the remaining 21 tags were excluded from analysis.

Travel time from tag activation to A20 or T1 varied from 4.7 d to 40.4 d and averaged 6.7 d (harmonic mean; 2.4 d to 37.9 d from release; Figure 3). The time to tag failure in the tag-life study ranged from 27.6 d to 46.0 d (Figure 3), and fish survival estimates were adjusted for premature tag failure. Without filtering

for possible tag predation, the total probability of survival from Durham Ferry to the Turner Cut Junction was estimated at 0.044 ($SE=0.009$; Table 3). Cumulative survival decreased markedly from A6 (Mossdale) to A12 (Brandt Bridge) and was largely stable from A12 to A17/R1 (Navy Bridge/Rough and Ready Island). Survival rate per km was lowest between A7 (River Islands Parkway) and A9 (Undine Road); this region included the head of Old River (Figure 4).

Table 3

Regional and total survival estimates for alternative predator filtering approaches for 2016 San Joaquin River Chinook salmon data processed using no filter, simple or complex rule-based (RB) filter, and smolt-only or multispecies pattern recognition (PR) filter. A0 = release site. Standard errors are in parentheses. Telemetry stations are identified in Figure 1 and Table 1. CV = coefficient of variation.

Survival region	No filter	Simple RB	Complex RB	Smolt-only PR	Multispecies PR	CV (%)
Durham Ferry to Turner Cut Junction (A0–A20/T1)	0.044 (0.009)	0.040 (0.008)	0.044 (0.009)	0.036 (0.008)	0.036 (0.008)	10.0
Durham Ferry to Mossdale (A0–A6)	0.643 (0.019)	0.642 (0.019)	0.642 (0.019)	0.643 (0.019)	0.644 (0.019)	0.1
Mossdale to Turner Cut Junction (A6–A20/T1)	0.061 (0.012)	0.055 (0.011)	0.060 (0.012)	0.049 (0.011)	0.049 (0.011)	10.5
Mossdale through head of Old River (A6–A8/B1)	0.665 (0.023)	0.671 (0.023)	0.659 (0.023)	0.673 (0.023)	0.661 (0.023)	0.9
Head of Old River to Weston Ranch (A8–A14)	0.313 (0.030)	0.311 (0.029)	0.318 (0.030)	0.253 (0.028)	0.234 (0.027)	13.7
Weston Ranch to Calaveras River (A14–A18)	0.665 (0.054)	0.637 (0.055)	0.650 (0.054)	0.638 (0.061)	0.687 (0.062)	3.2
Calaveras River to Turner Cut Junction (A18–A20/T1)	0.498 (0.071)	0.470 (0.071)	0.501 (0.071)	0.510 (0.081)	0.523 (0.082)	3.9

Simple Rule-based Filter

The simple rule-based filter classified 93 tags (14% of those released) as predated at some point after release (Figure 5a). The region from Mossdale through the head of Old River (A6 to A8/B1) had the largest number of first-time predator classifications (39), followed by the next region downstream (Undine Road to Howard Road, A9 to A13) with 30 first-time predator classifications (Table 2). The behavior that resulted in the largest number of predator assignments was long residence time near a receiver (filter category 5), which accounted for 61 rule violations (Table 4). Upstream transitions against river flow (category 3) accounted for 27 rule violations, and low average distance per day (category 4) accounted for 6 rule violations. The criteria for length and migration rate of upstream transitions (categories 1 and 2) were not violated for any tags.

Using the detection data passed through the simple rule-based filter, travel time from release to the Turner Cut Junction (A20/T1) ranged from 2.4 d to 18.0 d (harmonic mean = 4.1 d), and survival through this region was estimated at 0.040 ($SE=0.008$). Cumulative survival patterns largely followed those of the unfiltered data (Figure 4).

Table 4

Summary of predation classifications from simple rule-based filter: number of tags first flagged as predated in region because of meeting filter category. See text for more precise definitions of filter categories. Telemetry stations are identified in Figure 1 and Table 1.

Region (station codes)	Filter category					Total ^a
	Long upstream transition	Fast upstream transition	Transition against flow	Low distance per day	Long residence time	
Upstream of Mossdale (A1–A5)	0	0	5	6	1	12
Mossdale through head of Old River (A6–A8/B1)	0	0	7	0	32	39
Undine Road to Howard Road (A9–A13)	0	0	14	0	17	31
Weston Ranch to Navy Bridge/Rough and Ready Island (A14–A17/R1)	0	0	1	0	4	4
Calaveras River to Turner Cut Junction (A18–A20/T1)	0	0	0	0	7	7
Head of Middle River (B2, C1)	0	0	0	0	0	0
Total	0	0	27	6	61	93

a = Total may be greater than row total because a single detection event may violate multiple rules.

Complex Rule-based Filter

The complex rule-based filter classified 139 tags (21% of those released) as predated at some point after release (Figure 5b). Regional patterns agreed with the simple rule-based filter: the region from Mossdale through the head of Old River (A6–A8/B1) had the largest number of predator classifications (67), followed by the region from Undine Road (A9) to Howard Road (A13; Table 2). The near-field and mid-field residence time criteria accounted for the largest number of first-time predator classifications, followed by migration rate and far-field residence time. The time since release and the length- and velocity-adjusted migration rate (body lengths per second) categories accounted for the fewest predator classifications (Table 5).

Using detections passed through the complex rule-based filter, travel time from release to the A20/T1 stations averaged 3.9 d (range = 2.4 d to 10.5 d).

Regional survival estimates ranged from 0.318 ($SE=0.030$) from the head of Old River to Weston Ranch in Stockton (A8–A14) to 0.659 ($SE=0.023$) in the

segment from Mossdale through the head of Old River (A6–A8/B1); overall survival from release to the Turner Cut Junction was estimated at 0.044 ($SE = 0.009$) (Table 3). Cumulative survival patterns largely followed those from the unfiltered data (Figure 4).

Table 5

Summary of predation classifications from complex rule-based filter: number of tags first flagged as predated in region because of rule violation in filter category. See text for more precise definitions of filter categories. Telemetry stations are identified in Figure 1 and Table 1

Category										
Region (Station codes)	Near-field residence time	Mid-field residence time	Far-field residence time	Time since release	Migration rate	BLPS	Upstream	Against flow	Movement patterns	Total ^a
Upstream of Mossdale (A1–A5)	10	10	12	2	12	0	0	0	1	17
Mossdale through head of Old River (A6–A8/B1)	53	52	31	1	25	1	7	0	10	67
Undine Road to Howard Road (A9–A13)	31	26	6	1	16	2	9	1	0	37
Weston Ranch to Navy Bridge/Rough and Ready Island (A14–A17/R1)	6	4	1	0	4	0	0	3	1	9
Calaveras River to Turner Cut Junction (A18–A20/T1)	4	2	2	0	3	0	0	4	3	8
Head of Middle River (B2, C1)	1	1	0	0	0	0	0	0	0	1
Total	105	95	52	4	60	3	16	8	15	139

a = Total may be greater than row total because a single detection event may violate multiple rules

Smolt-only Pattern Recognition Filter

Interpretation of the cluster analysis in the smolt-only pattern recognition filter identified a minority cluster of 66 tags (10%) that had behavior distinct from the remaining 552 tags in the majority cluster (Figure 5c). Although increasing the number of clusters accounted for more variance in the data, there was no obvious optimal cluster number on the basis of variance explained (Figure 7a). The silhouette method indicated two clusters were sufficient (Figure 7b), and further results were based on these two clusters for the smolt-only pattern recognition filter. The first two principal components in the PCA accounted for 56% of the variance (Table 6). The ordination plot for the first two principal component axes depicted a single cloud of data points spread in three directions from the origin. Observations from the minority cluster were dispersed in the direction of positive values for PC1 and PC2 and were associated with metrics involving upstream movement. This cluster was classified as coming from putative predators. The remaining observations were evenly distributed within a cluster extending from the upper left to the lower right quadrants; these observations were classified as belonging to smolts. This grouping contained observations without evidence of upstream movement and spanned a gradient of migration rates. High average migration rate was negatively correlated (180° difference in the angle of loadings) with long residence times and many repeated visits to arrays. We ignored the third principal component in our interpretation of clusters as it mainly reflected differences in downstream migration rate and did not further separate clusters visible in our initial two-dimensional plotting.

The ROE method performed well for the smolt-only analysis, as the ordination points representing the 66 predated tags left the ellipse across a range of detection events and did not return, and only 4 of the points moved outside the ellipse prior to the event minimum ($i_0 = 3$; Figure S3a). The region with the largest number of first-time predation events was from Undine Road (A9) to Howard Road (A13) (25 events; Table 2). The neighboring regions had the next highest predator classifications: 19 tags from Weston Ranch through Navy Bridge/Rough and Ready Island (A14–A17/R1), and 15 tags from Mossdale through the head of Old River (A6–A8/B1). Using the detection events that passed through the smolt-only pattern recognition filter, maximum travel time from release to the Turner Cut Junction (A20/T1) was 37.9 d, and total survival to A20/T1 was estimated at 0.036 ($SE=0.008$; Table 3). Cumulative survival largely followed the patterns from the unfiltered data until station A12 (Brandt Bridge), after which survival began to diverge from the unfiltered estimates. The largest difference in cumulative survival between the smolt-only pattern recognition filter and the no-filter option was observed for A17/R1 (Navy Bridge/Rough and Ready Island; difference = 0.079, $SE=0.034$; Figure 4).

Table 6

Correlations between principal component weights and behavioral metrics as part of two pattern recognition procedures. Principal component analyses were applied to two data sets: (1) behavioral metrics of 618 tagged smolts (“smolt-only”) and (2) the same data set augmented with 37 additional detection records from tagged predator fish species (“multispecies”; largemouth bass, striped bass, channel catfish, and white catfish). The percentage of variance accounted for by each of the first three principal components is shown after the axis label. Correlations are shown between metrics and the weights of the first three principal components; values > 0.6 are in bold and standard errors of correlations are in parentheses.

Behavioral metric	Smolt-only			Multispecies		
	Axis 1 (38%)	Axis 2 (18%)	Axis 3 (12%)	Axis 1 (36%)	Axis 2 (17%)	Axis 3 (14%)
Residence time (mean)	0.70 (0.03)	-0.42 (0.04)	-0.12 (0.04)	0.68 (0.04)	-0.45 (0.06)	0.24 (0.19)
Migration rate (mean)	-0.38 (0.05)	0.50 (0.06)	0.69 (0.05)	-0.37 (0.06)	0.79 (0.03)	0.25 (0.09)
Visits (count)	0.84 (0.01)	-0.20 (0.03)	0.24 (0.03)	0.79 (0.03)	-0.03 (0.07)	0.42 (0.2)
Distance (sum)	0.78 (0.02)	-0.06 (0.05)	0.30 (0.04)	0.71 (0.03)	0.31 (0.06)	0.42 (0.16)
Upstream transitions	0.65 (0.03)	0.65 (0.04)	-0.19 (0.04)	0.66 (0.04)	0.40 (0.08)	-0.46 (0.31)
Detections in high-risk zones	0.64 (0.04)	-0.38 (0.05)	0.05 (0.06)	0.60 (0.05)	-0.23 (0.07)	0.16 (0.1)
Route reversals	0.63 (0.03)	0.50 (0.05)	-0.15 (0.03)	0.72 (0.05)	0.36 (0.08)	-0.39 (0.33)
Visits (max)	0.51 (0.03)	-0.29 (0.08)	-0.19 (0.06)	0.42 (0.02)	-0.31 (0.04)	-0.33 (0.17)
Time since release	0.78 (0.02)	-0.44 (0.04)	-0.06 (0.04)	0.75 (0.04)	-0.48 (0.05)	0.13 (0.14)
Residence time (max)	0.77 (0.02)	-0.40 (0.04)	-0.05 (0.04)	0.74 (0.03)	-0.39 (0.06)	0.31 (0.2)
Upstream distance (max)	0.61 (0.04)	0.61 (0.04)	-0.18 (0.04)	0.68 (0.05)	0.38 (0.09)	-0.45 (0.31)
Flow during residence	0.33 (0.04)	-0.36 (0.06)	-0.01 (0.06)	0.39 (0.04)	-0.39 (0.04)	-0.05 (0.07)
Upstream migration rate (max)	0.66 (0.03)	0.58 (0.04)	-0.12 (0.04)	0.69 (0.06)	0.40 (0.07)	-0.24 (0.23)
Downstream migration rate (max)	0.35 (0.04)	0.16 (0.08)	0.87 (0.02)	0.18 (0.05)	0.68 (0.08)	0.61 (0.21)
Distance over days (max)	-0.05 (0.05)	0.13 (0.08)	0.77 (0.04)	-0.28 (0.04)	0.51 (0.08)	0.56 (0.23)
Migration rate adj. (max[x])	0.80 (0.02)	0.15 (0.04)	0.19 (0.04)	0.64 (0.05)	0.21 (0.04)	0.41 (0.15)
Migration rate (sd)	0.59 (0.03)	-0.35 (0.05)	0.32 (0.05)	0.47 (0.03)	0.08 (0.09)	0.60 (0.29)
Proportion upstream transitions	0.62 (0.04)	0.65 (0.04)	-0.23 (0.04)	0.70 (0.06)	0.35 (0.09)	-0.43 (0.3)
Proportion visits in high-risk zones	0.45 (0.03)	-0.48 (0.05)	0.05 (0.07)	0.38 (0.04)	-0.45 (0.07)	0.26 (0.21)
Proportion of against-flow transitions	0.51 (0.06)	0.44 (0.06)	-0.04 (0.05)	0.63 (0.05)	0.28 (0.07)	-0.31 (0.23)

Multispecies Pattern Recognition Filter

The multispecies pattern recognition filter classified 89 smolt tags (14% of those released) as predated by the end of the study (Table 2). The cluster analysis for the multispecies pattern recognition data set partitioned the study tags into five clusters, the two largest of which were classified as smolts and the remainder as predators. Five clusters were selected in order to jointly minimize the total number of clusters and the number of mixed clusters containing large majorities of smolt tags. The dendrogram in Figure 5 shows the groupings at different levels of clustering; the great majority of known predator tags (35/37) are grouped at the top and bottom. The separation between groups of known predators reflected different patterns of behavior. Dividing the data set into two clusters would have produced a result similar to the smolt-only analysis but overlooked the large number (549) of smolt tags intermixed with the majority of predators (upper cluster). Three clusters were sufficient to isolate a smolt-tag-only grouping (labeled cluster 3 in Figure 5d) from those including known predator tags but left a large mixed cluster that was 93% smolt tags but which contained 62% (23) of the total known predator tags in the data set. Five clusters were necessary to separate the predator-rich grouping at the top of the dendrogram (cluster 1: 20 smolt tags, 21 predator tags) from the two majority-smolt-tag clusters in the middle of the dendrogram (clusters 2 and 3), both of which exhibited behavior patterns more expected from smolts. The fourth cluster included all 66 tags flagged as suspicious in the smolt-only procedure and 3 additional smolt tags. Admixture of 12 known predator tags with smolt tags in this cluster provided additional evidence of tag predation (Figure 5d, cluster 4). The smallest cluster of the five (cluster 5) contained only 2 tags, both of which were known predators (striped bass and white catfish) and deviated significantly from the norm based on their long detection sequences (> 35 events) and high frequency of upstream transitions. Examination of the PCA ordination plot (not shown) and correlations between principal component weights and behavioral metrics for the multispecies procedure (Table 6) indicated that cluster 1 was characterized by abnormally long residence times at telemetry stations. Smolt tags in clusters 1, 4, and 5 were classified as predated. The remaining 529 smolt tags in clusters 2 and 3 were classified as non-predated by the end of the study (Figure 5d).

The ROE method did not perform seamlessly when applied to the multispecies data set. Of the smolt tags that were classified as predators in the cluster analysis (T_p), 21% (19/89) left the boundary of the ellipse before the event minimum ($i_0 = 4$) and 27% (24/89) of predated tags crossed the ellipse boundary more than once (Figure S3b). Leaving the ellipse before the event minimum meant those tags were classified at the fourth detection event ($i = 4$) by default. Many of the points associated with known tagged predators also left the ellipse when only a few detection events were summarized and in similar directions, thus increasing confidence that smolt tags exiting the ellipse early were aberrant in the same way. Tags that returned to the ellipse after initial exit were more concerning, as this pattern suggested that aberrant behavior was only temporary for some purported predated tags.

Similar to the smolt-only pattern recognition filter, the smolt-ellipse departure times for the multispecies filter indicated that the region with the largest number of predation events was between the head of Old River (Undine Road, A9) and south Stockton (Howard Road, A13), followed by the region from Mossdale through the head of Old River (A6–A8/B1) and from Weston Ranch to Navy Bridge/Rough and Ready Island (A14–A17/R1) (Table 2). Maximum travel time from release to the Turner Cut Junction (A20/T1) was 37.9 d, and survival through this region was estimated at 0.036 ($SE=0.008$) (Table 3). Cumulative survival largely followed the patterns from the smolt-only pattern recognition filter, although the estimates from the multispecies filter were consistently slightly lower (non-significant difference) than from the smolt-only pattern recognition filter from Brandt Bridge (A12) through the San Joaquin Shipping Channel (A19) (Figure 4).

Comparison of Filters

Our comparison of the two rule-based filters and the two pattern recognition filters indicated that the filtering approach affected both the number and composition of tags considered suspicious (Table 2, Figure 5). More than twice as many tags were classified as suspect using the complex rule-based filter ($n=139$) compared to the smolt-only pattern recognition approach ($n=66$). A total of 32 tags were classified as predated by all four filtering methods and 54 tags were classified as predated based on only one of the filters. Overall, the four filtering approaches agreed on classifications for 77% of the 618 tags observed. Tags flagged by each of the four filters had significantly more upstream transitions compared to the population average. Although the composition of flagged tags differed across filtering methods, subsets of suspicious tags shared some similar attributes. The population average for the proportion of tags with at least two upstream transitions was 1%, compared with 6% for tags flagged by the complex rule-based filter, 12% for the simple pattern recognition filter, and 9% for the other two filters. The residence time was higher and more variable for flagged tags than for the full population. The harmonic mean residence time of the smolt-only pattern recognition filter was considerably higher at 0.84 hr ($SE=0.20$ hr) than the population mean of 0.06 hr ($SE=0.01$ hr).

There was generally higher agreement in predated tag classifications between related methods. For the rule-based filters, 94% (87 of 93) of the tags flagged by the simple rule-based filter were also flagged by the complex filter (Figure 5a, Figure 5b). The complex rule-based filter flagged an additional 52 tags as suspicious, primarily from Mossdale through the head of Old River (A6–A8/B1; Table 2, Figure 6). Likewise, all tags flagged by the smolt-only pattern recognition filter were also identified by the multispecies pattern recognition filter (Figure 5c, Figure 5d). The additional data provided by the known predators in the multispecies version of the filter resulted in flagging another 23 salmon tags. Overall, the two rule-based filters agreed on the outcome of 560 tags (91% of those observed), and the two pattern recognition filters agreed on the outcome of 595 tags (96%) (Figure 5).

The detection event where a tag's first predator classification was assigned varied spatially among filters (Figure 6). The rule-based filters assigned some predator classifications at the start of detection events (i.e., predation event occurred in the reach between adjacent stations) and some at the end of detection events (i.e., predation event occurred in the vicinity of the station), whereas the pattern recognition filters assigned all predator classifications in the reaches between stations. Despite variability in the location of initial predator classifications, all filters flagged a high proportion of suspicious detections in the region between Banta Carbona and Frewert Road (A5–A11). The predator classifications from the two rule-based filters were concentrated more around the head of Old River (A6–A8), whereas those from the pattern recognition filters tended to concentrate in the San Joaquin River downstream of Old River (A9–A11).

When translated to survival, the differences observed among the four filtering approaches and the degenerate “no filter” approach varied by spatial scale and region (Figure 4). On the reach scale, defined by adjacent telemetry stations (Figure 1), the maximum absolute difference in survival estimates among the filtering approaches ranged from <0.001 for the reach from Banta Carbona to Mossdale (A5–A6) to 0.112 for the reach immediately preceding Garwood Bridge (A15–A16). In the latter reach, survival estimates ranged from 0.875 ($SE=0.039$) for the smolt-only pattern recognition filter to 0.987 ($SE=0.013$) for the complex rule-based filter; the relative variation among survival estimates from the different filtering approaches for that reach was CV (coefficient of variation)=4.6% (See Tag Predation Supplementary Materials, Table S5, Supplemental Results). The reach from Navy Drive Bridge to the Calaveras River station (A17–A18) had the highest relative variation of all reaches (CV=5.5%); survival in this reach was estimated at 0.710–0.739 ($SE\leq 0.061$) for the two pattern recognition filters, compared to 0.654–0.658 ($SE\leq 0.055$) for the two rule-based filters and the no-filter approach (See Tag Predation Supplementary Materials, Table S5, Supplemental Results).

On the regional scale, the reaches from Durham Ferry to Mossdale (A0–A6) were least sensitive to filter choice, with cumulative survival estimates of 0.642–0.644 ($SE=0.019$) for all filtering methods (CV=0.1%; Table 3). The region with the highest sensitivity to filter choice was the head of Old River to the Weston Ranch station near Stockton (A8–A14), where survival estimates ranged from 0.234 ($SE=0.027$) for the multispecies pattern recognition filter to 0.318 ($SE=0.030$) for the complex rule-based filter, and the CV of the survival estimates was 13.7%. The pattern recognition filters tended to produce lower survival estimates until the upstream Stockton reaches (A14), after which the rule-based filters had lower estimates (Table 3, Figure 4). The spatial variation in filter rankings acted to stabilize the cumulative survival estimates on larger spatial scales so that there was little absolute difference in total survival estimates

from release to the Turner Cut Junction: 0.036 ($SE=0.008$) for the pattern recognition filters to 0.044 ($SE=0.009$) for the complex rule-based filter and the “no-filter” approach. Although the variation was small on the absolute scale, the low magnitude of the survival estimates overall resulted in a relatively high CV of 10.0% among survival estimates on this spatial scale. Closer examination of the 2016 data demonstrated that the majority (66%) of tags had detection histories that simply ended before reaching the Turner Cut Junction and did not receive a predated classification from any filter.

The filtering approach had little effect on the distribution of observed travel time through the system for fish that completed travel within approximately one week of release, as demonstrated by the first 75% of the cumulative arrival distributions to the Turner Cut Junction (Figure 3). The primary effect of the predator filters on perceived travel time was observed for the last 10% of the arriving fish. The rule-based filters removed the latest tag detection(s) at the Turner Cut Junction, reducing the maximum observed travel time from approximately 38 d to 18 d using the simple rule-based filter and 10.5 d for the complex rule-based filter. The pattern recognition filters, on the other hand, tended to remove Turner Cut Junction detections between approximately days 8 and 12 but did not remove the latest detections.

Discussion

Predation of study fish is a common occurrence and is likely the proximate cause of mortality in some studies of juvenile fish. Although some studies assume negligible mortality whether from predation or other sources (e.g., behavior studies), survival studies are designed to characterize the magnitude of mortality and the factors that contribute to it. The large majority of tagging studies require that the measures in the data set represent only live study fish rather than a mix of live study fish, predators, and deposited tags. While a variety of methods have been used previously to diagnose tag predation, this is the first study to compare results of different methods and assess the sensitivity of study results to the diagnosis process used. For this data set, we observed considerable variability in predator status classification depending on the predator filter, but minimal effect on survival estimates (e.g., Figure 4) and measures of average travel time. The largest effects were seen in estimates of the upper quantiles of the travel time distribution (e.g., Figure 3).

Despite the variability in predated tag classifications among the filters, the resulting absolute differences in survival estimates tended to be small and within the sampling error. When translated to the relative scale, however, the differences between the various filtered and unfiltered estimates of cumulative survival were larger, ranging up to 19% for survival to the Turner Cut Junction (A20/T1) and 31% for survival to the Navy Bridge/Rough and Ready Island region (A17/R1) (pattern recognition filters, Figure 4). Survival to A17/R1 also had the largest variability among filter estimates (CV = 16.7%). These results are comparable to the large differences in relative survival observed in a 2010 study of subyearling Chinook salmon in the region, in which the estimate of total Delta survival from Mossdale to Delta exit at Chipps Island (55 km downstream of A20/T1) was reduced by 50% when a similar complex rule-based filter was applied to the data set (0.11 vs 0.05; Buchanan et al. 2018).

The four predator filters considered here entailed different tradeoffs in their level of complexity, subjectivity, and interpretability and in the effort required by the researcher. Higher levels of complexity resulted from more data (e.g., multispecies versus smolt-only tag data), increased degrees of freedom represented by more metrics (e.g., complex versus simple rule-based filters), or more decisions that also represented a higher level of effort and subjectivity (e.g., complex rule-based filter compared to the others). On the other hand, the rule-based filters resulted in clearly defined state assignments whereas the cluster analysis in the pattern recognition filters required further assumptions and interpretations to yield state assignments (e.g., the ROE method).

The issue of subjectivity looms large in adapting any predator filter approach and making inferences from its results. In some data sets, there are cases in which particular tag movements are well beyond the swimming capabilities of the focal species and the choice to omit some or all of a tag record is obvious. However, there are likely to be many more cases where subjective judgements can significantly influence findings. The two primary filtering approaches considered here varied in their demands on subject matter knowledge by the researcher and their degree of subjectivity. The metrics and criteria in the rule-based filters were based entirely upon expert knowledge gleaned from literature review, consultation with salmon and bass biologists, and personal experience with similar data sets. Pattern recognition methods that use multivariate or machine learning, on the other hand, prioritize letting signals emerge from the data rather than requiring explicit biological judgements from the researcher. A distinct benefit of the pattern recognition approach lies in its ability to isolate sets of abnormal behavior metrics without requiring the researcher to identify precise numeric criteria associated with smolt-like behavior. Thus, it may appear that pattern recognition filters or similar statistics-based filters remove all subjectivity from the exercise. However, there remains subjectivity in selecting the metrics to include, how to scale or transform the metrics, how many clusters to use, and how to interpret the clusters identified. In our analysis, we were confronted with several options of distance measures, clustering algorithms, statistics for choosing an optimal number of clusters, and ordination approaches. Thus, although pattern recognition approaches remove the researcher's responsibility for setting numeric thresholds, there remain a multitude of decisions, each of which may affect the outcome.

There are limitations to any predator filter's ability to correctly identify all predator detections. Even a well-designed predator filter based on realistic and defensible understanding of the behavior of the focal species and predators may miss some predator detections and/or misclassify some valid detections. A predated tag that has a short detection history may have too few opportunities to demonstrate aberrant behavior to trigger a predator classification. Highly variable behavior in the focal species will also make it difficult to accurately distinguish between predated and non-predated tags. Other factors contributing to variability in filter outcomes include the composition, mobility, and home range of the predator community, the density of the telemetry network and the spatiotemporal scale of inference, and the amount of data available. The availability of predator data in particular may affect the outcome of the multispecies pattern recognition filter. This case study used opportunistic detections of four species of piscivorous fish that had been tagged with long-lived acoustic tags one or two years prior to this study; only 37 individual predators were represented, 24 (65%) of which were largemouth bass and most of which were detected in the segment between the A6 and A12 arrays. A larger and more representative data set of tagged predators captured, released, and detected in a more variable spatial region may have resulted in different predator classifications from the multispecies filter.

Our case study findings suggest that a predator filter may be more important when studying survival over smaller spatial scales, for intermediate levels of survival, or to characterize maximum travel time. Although the need for a predator filter will be apparent only after comparing the filtered and unfiltered data, it may be expected to be more important in some settings than in others. If the home range of the predators lies entirely within the extent of a single reach for which survival is estimated, then tag predation will have no effect on survival estimation because predators are not expected to pass a monitoring station. It is when a predator's range includes one or more monitoring sites that tag predation may bias survival estimation and a predator filter is recommended. Additionally, formal methods may be more necessary in settings in which individuals of the focus species may legitimately move in the direction opposite their eventual target, such as a tidal region with reverse flows; in simpler settings, transitions in an unexpected direction may be easily interpreted as evidence of predation, whereas in more complex settings it is necessary to compare observations with prevailing environmental conditions. Thus, formal predator filters are often omitted from studies of juvenile salmon migratory survival through the 40- to 100-km riverine reaches in the Columbia River; this region has no reverse flows and the predator community is dominated by avian species and piscine species such as Northern pikeminnow *Ptychocheilus oregonensis*, which exhibit mostly small-scale movements near dams or movements upriver during the spring juvenile salmonid emigration (Martinelli and Shively 1997). In this setting, upstream-directed movement may be easily interpreted as evidence of predation. Formal predator filters are more common in studies of salmon migratory survival through the 1- to 10-km tidally influenced reaches in the Sacramento–San Joaquin River delta in the presence of reverse flows and striped bass that may move about the entire south Delta and between the Delta and the Pacific Ocean (Boughton 2020).

Even in settings in which less complicated filtering methods may be suitable, we recommend that researchers clearly document their assumptions and filtering procedures. Unrecognized conceptions of what constitutes realistic behavior can influence results and may have a cascading effect. Furthermore, ad hoc or undocumented procedures bear the risk of investigator drift during implementation, in which a rule for omitting detections may become more or less strict as it is implemented over the complete set of tags and detections. Thus, we recommend identifying and documenting a well-defined procedure to detect predated tags and determine when in the detection history the tag no longer represented the live study fish. The methods should be tailored to the study's focal species, likely predator species, setting, season, and monitoring network. We also recommend that study results be examined for a range of filtering assumptions, at minimum comparing outcomes using the researcher's best understanding of the true state of the detection data (presumably the filtered data) with outcomes from data that are either entirely or partially unfiltered.

More sophisticated modeling techniques and new tag technologies could provide avenues for more robustly addressing the tag predation problem. Modeling techniques designed to account for unobservable or partially observable states, such as hidden Markov and state-space models (Patterson et al. 2009; Alós et al. 2016; Hurme et al. 2019), could be used to propagate the uncertainty in predation status to the final study results. To our knowledge, neither of these approaches has yet been used to account for tag predation. Additionally, there now exist specialized microacoustic transmitters that change their output when they have been predated (predator tags; Seitz et al. 2019; Daniels et al. 2019). For example, manufacturers have designed an acoustic tag that transmits an alternative signal after a polymer coating has broken down in a predator's gut (Weinz et al. 2020). Such predator tags remove the sensitivity of predation detection to behavioral differences. Even with these specialized tags, however, there remain challenges of correctly identifying where and when the predation event occurred, caused in part by variation in digestion rate and the time delay between the predation event and the onset of the predation signal (trigger time) (Halfyard et al. 2017).

Conclusions

Technological advances in telemetry hold great potential for more complete and timely information on individual fish of a widening range of species, life stages, and sizes. With the ability to tag smaller fish and the use of telemetry to study survival over smaller spatiotemporal scales comes the increased risk that tag predation may bias study results. The magnitude of the bias will be apparent only when researchers intentionally investigate the outcomes of one or more predator filters. Rule-based filters have great flexibility but require a high degree of research judgement; pattern-recognition filters are more automated but depend heavily on the individual tags and metrics included. The tag predation problem and the data filtering approaches used to address it warrant serious attention by investigators and further research. We encourage researchers to articulate their assumptions and filtering rules and to report the robustness of their results to the filtering procedure invoked. Such practices should be part of accepted methodology to generate repeatable, defensible science in fisheries research.

Abbreviations

BLPS: body lengths per second; CDWR: California Department of Water Resources; CJS: Cormack-Jolly-Seber; CV: coefficient of variation; rkm: river km; PCA: principal components analysis; ROE: recursive ordination ellipse (ROE); USBR: U.S. Bureau of Reclamation; USFWS: U.S. Fish and Wildlife Service; USGS: U.S. Geological Survey.

Declarations

Ethics Approval and Consent to Participate

Human subjects: not applicable. Fish tagging and handling was performed under authority of the U.S. Fish and Wildlife Service (USFWS), which does not require approval for animal care as it applies to research using hatchery-raised Central Valley fall-run Chinook salmon. We followed a standard operating protocol for fish tagging and handling developed jointly by the U.S. Geological Survey (USGS) and USFWS, based on Adams et al. (1998), Martinelli et al. (1998), and Liedtke et al. (2012).

Consent for Publication

The manuscript does not contain any individual person's data, details, images, or videos.

Availability of Data and Materials

The datasets analyzed in the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Funding

This publication was produced with the financial support of the USFWS. The tagging study was conducted jointly by the USFWS, the U.S. Bureau of Reclamation (USBR), the USGS, and the University of Washington, and funded by the Central Valley Project Improvement Act Restoration Fund, California Department of Water Resources, and the USBR.

Authors' contributions

Both authors contributed equally to the planning, analysis, and writing of the manuscript. Both authors read and approved the final manuscript.

Acknowledgements

Many individuals from several agencies made this project possible. The tagging study was directed by the U.S. Fish and Wildlife Service (USFWS) and the U.S. Bureau of Reclamation (USBR). Individuals from the USFWS, USBR, U.S. Geological Survey (USGS), and California Department of Water Resources (CDWR) implemented the tagging and release components of the project. The USGS provided training for the surgeons, helped design and installed, operated, maintained, and retrieved the acoustic receiver array, and pre-processed the data. Funding for the tagging study and analysis came from the Comprehensive Assessment and monitoring Program for the Central Valley Project Improvement Act, CDWR, and the USBR. Funding for manuscript preparation was provided by the USFWS. The authors are especially grateful to Christopher Holbrook (USGS) for data processing, discussion of filter ideas, and mapping assistance, Pat Brandes and Denise Barnard (USFWS) for information on tagging and release operations, and Craig Scanlan (University of Washington) for manuscript preparation.

References

1. Adams NS, Rondorf DW, Evans SD, Kelly JE. Effects of surgically and gastrically implanted radio transmitters on growth and feeding behavior of juvenile Chinook salmon. *Transactions of the American Fisheries Society*. 1998;127(1):128–1136.
2. Alós J, Palmer M, Balle S, Arlinghaus R. Bayesian state-space modelling of conventional acoustic tracking provides accurate descriptors of home range behavior in a small-bodied coastal fish species. *PLoS ONE*. 2016;11(4):e0154089. <http://doi.org/10.1371/journal.pone.0154089>.
3. Beland KF, Kocik JF, vandeSande J, Sheehan TF. Striped bass predation upon Atlantic salmon smolts in Maine. *Northeast Nat*. 2001;8(3):267–274.
4. Boughton DA. Striped bass on the coast of California: a review. *California Fish and Wildlife*. 2020;106(3):226–257.
5. Buchanan RA, Brandes PL, Skalski JR. Survival of juvenile fall-run Chinook salmon through the San Joaquin River Delta, California, 2010–2015. *North Am J Fish Manage*. 2018;38(3):663–679. <http://doi.org/10.1002/nafm.10063>.
6. Buchanan RA, Skalski JR. Relating survival of fall-run Chinook salmon through the San Joaquin Delta to river flow. *Environ Biol Fish*. 2020;103:389–410. <http://doi.org/10.1007/s10641-019-00918-y>.
7. Burnham KP, Anderson DR, White GC, Brownie C, Pollock KH. Design and analysis methods for fish survival experiments based on release-recapture. *American Fisheries Society. Monograph* 5. 1987.
8. Cohen AN, Carlton JT. Accelerating invasion rate in a highly invaded estuary. *Science*. 1998;279:555–558.
9. Conrad JL, Bibian AJ, Weinersmith KL, De Carion D, Young MJ, Crain P, Hestir EL, Santos MJ, Sih A. Novel species interaction in a highly modified estuary: association of largemouth bass with Brazilian waterweed *Egeria densa*. *Trans Am Fish Soc*. 2016;145:249–263.
10. Cooke SJ, Midwood JD, Thiem JD, Klimley P, Lucas MC, Thorstad EB, Eiler J, Holbrook C, Ebner BC. Tracking animals in freshwater with electronic tags: past, present and future. *Anim Biotelemetry*. 2013;1(1):5. <http://doi.org/10.1186/2050-3385-1-5>.
11. Crossin GT, Heupel MR, Holbrook CM, Hussey NE, Lowerre-Barbieri SK, Nguyen VM, Raby GD, Cooke SJ. Acoustic telemetry and fisheries management. *Ecol Appl*. 2017;27(4):1031–1049. <http://doi.org/10.1002/eap.1533>.
12. Daniels J, Sutton S, Webber D, Carr J. Extent of predation bias present in migration survival and timing of Atlantic salmon smolt (*Salmo salar*) as suggested by a novel acoustic tag. *Anim Biotelemetry*. 2019;7(1): 16. <http://doi.org/10.1186/s40317-019-0178-2>.
13. Drenner SM, Clark TD, Whitney CK, Martins EG, Cooke SJ, Hinch SG. A synthesis of tagging studies examining the behaviour and survival of anadromous salmonids in marine environments. *PLoS One*. 2012;7(3). <http://doi.org/10.1371/journal.pone.0031311>.
14. Evans AF, Payton Q, Turecek A, Cramer B, Collis K, Roby DD, Loschl PJ, Sullivan L, Skalski JR, Weiland M, Dotson C. Avian predation on juvenile salmonids: spatial and temporal analysis based on acoustic and passive integrated transponder tags. *Trans Am Fish Soc*. 2016;145(4):860–877. <http://doi.org/10.1080/00028487.2016.1150881>.
15. Gibson AJF, Halfyard EA, Bradford RG, Stokesbury MJW, and Redden AM. Effects of predation on telemetry-based survival estimates: insights from a study on endangered Atlantic salmon smolts. *Can J Fish Aquat Sci*. 2015;72(5): 728–741. <http://doi.org/10.1139/cjfas-2014-0245>.
16. Grossman, GD. Predation on fishes in the Sacramento–San Joaquin Delta: current knowledge and future directions. *San Franc Estuary Watershed Sci*. 2016;14(2). <http://doi.org/10.15447/sfew.2016v14iss2art8>.

17. Halfyard, EA, Webber D, Del Papa J, Leadley T, Kessel ST, Colborne SF, Fisk AT. Evaluation of an acoustic telemetry transmitter designed to identify predation events. *Methods Ecol Evol.* 2017;8(9):1063–1071. <http://doi.org/10.1111/2041-210X.12726>.
18. Hause C. Outmigration survival of juvenile spring-run Chinook salmon in relation to physicochemical conditions in the San Joaquin River. Master's Thesis: University of California Davis. 2020.
19. Havn TB, Økland F, Teichert MAK, Heermann L, Borchering J, Sæther SA, Tambets M, Diserud OH, and Thorstad EB. Movements of dead fish in rivers. *Anim Biotelemetry.* 2017;5(1):7. <http://doi.org/10.1186/s40317-017-0122-2>.
20. Hayes SA, Huff DD, Demer DA, Michel CJ, Cutter Jr. GR, Demetras NJ, Lehman B, Manugian SC, Lindley ST, Smith JM, Quinn TP. Testing the effects of manipulated predator densities and environmental variables on juvenile salmonid survival in the lower San Joaquin River. Report produced by National Marine Fisheries Service (SWFSC) for the California Department of Water Resources. 2017;266. https://mafiadoc.com/download/testing-the-effects-of-manipulated-predator-densities_5c6adab2097c4723258b4595.html.
21. Hurme E, Gurarie E, Greif S, Herrera LGM, Flores-Martínez JJ, Wilkinson GS, Yovel, Y. Acoustic evaluation of behavioral states predicted from GPS tracking: a case study of a marine fishing bat. *Mov Ecol.* 2019;7(1):21. <http://doi.org/10.1186/s40462-019-0163-7>.
22. Hussey NE, Kessel ST, Aarestrup K, Cooke SJ, Cowley PD, Fisk AT, Harcourt RG, Holland KN, Iverson SJ, Kocik JF, Mills Flemming JE, Whoriskey FG. Aquatic animal telemetry: A panoramic window into the underwater world. *Science.* 2015;348(6240):1255642–1255642. <http://doi.org/10.1126/science.1255642>.
23. Jepsen N, Aarestrup K, Økland F, Rasmussen G. Survival of radiotagged Atlantic salmon (*Salmo salar* L.) – and trout (*Salmo trutta* L.) smolts passing a reservoir during seaward migration. *Hydrobiologia.* 1998;371(0):347. <http://doi.org/10.1023/A:1017047527478>.
24. Johnson RA, Wichern DW. Applied multivariate statistical analysis. 6th ed. Prentice Hall; 2007.
25. Johnston ME, Steel AE, Espe M, Sommer T, Klimley AP, Sandstrom P, Smith D. Survival of juvenile Chinook salmon in the Yolo Bypass and the lower Sacramento River, California. *San Franc Estuary Watershed Sci.* 2018;16(2). <http://doi.org/10.15447/sfew.2018v16iss2art4>.
26. Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. John Wiley & Sons; 2009.
27. Klimley AP, MacFarlane RB, Sandstrom PT, Lindley ST. A summary of the use of electronic tagging to provide insights into salmon migration and survival. *Environ Biol Fish.* 2013;96(2):419–428. <http://doi.org/10.1007/s10641-012-0098-y>.
28. Klinard NV, Matley JK. Living until proven dead: addressing mortality in acoustic telemetry research. *Rev Fish Biol Fisheries.* 2020;30(3):485–499. <http://doi.org/10.1007/s11160-020-09613-z>.
29. Lady, JM, JR Skalski. USER 4: user-specified estimation routine. Columbia Basin Research, School of Aquatic and Fishery Sciences, University of Washington, Seattle. 2009. <http://www.cbr.washington.edu/analysis/apps/user>.
30. Legendre P, Legendre L. Numerical ecology. 3rd ed. Elsevier BV; 2012.
31. Li T, Anderson JJ. Shaping human mortality patterns through intrinsic and extrinsic vitality processes. *Demographic Research.* 2013;28:341–72.
32. Liedtke TL, Beeman JW, Gee LP. A standard operating procedure for the surgical implantation of transmitters in juvenile salmonids. *U.S. Geological Survey.* 2012;2012–1267.
33. Link WA, Yoshizaki J, Bailey LL, Pollock KH. Uncovering a latent multinomial: analysis of mark–recapture data with misidentification. *Biometrics.* 2010;66(1):178–185. <http://doi.org/10.1111/j.1541-0420.2009.01244.x>.
34. Lucas MC, and Baras, E. Methods for studying spatial behaviour of freshwater fishes in the natural environment. *Fish and Fisheries.* 2000;1(4):283–316. <http://doi.org/10.1046/j.1467-2979.2000.00028.x>.
35. March D, Palmer M, Alós J, Grau A, Cardona F. Short-term residence, home range size and diel patterns of the painted comber *Serranus scriba* in a temperate marine reserve. *Mar Ecol Prog Ser.* 2010;400:195–206. <http://doi.org/10.3354/meps08410>.
36. Martinelli TL, Shively RS. Seasonal distribution, movements and habitat associations of northern squawfish in two lower Columbia River reservoirs. *Regulated Rivers: Research & Management.* 1997;13(6):543–556.
37. Martinelli TL, Hansel HC, Shively RS. Growth and physiological responses to surgical and gastric radio transmitter implantation techniques in subyearling Chinook salmon (*Oncorhynchus tshawytscha*). *Hydrobiologia.* 1998;371/372:79–87.
38. McMichael GA, Eppard MB, Carlson TJ, Carter JA, Ebberts BD, Brown RS, Weiland M, Ploskey GR, Harnish RA, Deng ZD. The juvenile salmon acoustic telemetry system: a new tool. *Fisheries.* 2010;35(1):9–22. <https://doi.org/10.1577/1548-8446-35.1.9>.
39. Michel, CJ, Smith JM, Demetras NJ, Huff DD, Hayes SA. Non-native fish predator density and molecular-based diet estimates suggest differing effects of predator species on juvenile salmon in the San Joaquin River, California. *San Franc Estuary and Watershed Sci.* 2018;16(4). <https://doi.org/10.15447/sfew.2018v16iss4art3>
40. Michel CJ, Smith JM, Lehman BM, Demetras NJ, Huff DD, Brandes PL, Israel JA, Quinn TP, Hayes SA. Limitations of active removals to manage predatory fish populations. *North Am J Fish Manage.* 2020;40:3–16. <https://doi.org/10.1002/nafm.10391>.
41. Moyle PB. The decline of anadromous fishes in California. *Conserv Biol.* 1994;8(3):869–870.
42. Muhametsafina A, Midwood JD, Bliss SM, Stamplecoskie KM, Cooke SJ. The fate of dead fish tagged with biotelemetry transmitters in an urban stream. *Aquat Ecol.* 2014;48(1):23–33. <https://doi.org/10.1007/s10452-013-9463-y>.
43. Murtagh, F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J Classif.* 2014;31(3):274–295.
44. National Marine Fisheries Service (NMFS). Endangered species act section 7 consultation: biological opinion and conference opinion on the long-term operations of the Central Valley Project and State Water Project. National Marine Fisheries Service, 2009.

<https://www.fisheries.noaa.gov/resource/document/biological-opinion-and-conference-opinion-long-term-operations-central-valley>. Accessed 15 October 2020.

45. Nobriga ML, Feyrer F. Shallow-water piscivore-prey dynamics in California's Sacramento–San Joaquin delta. *San Franc Estuary and Watershed Sci.* 2007;5(2). <https://doi.org/10.15447/sfew.2007v5iss2art4>.
46. Patterson TA, Basson M, Bravington MV, Gunn JS. Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *J Anim Ecol.* 2009;78(6):1113–1123. <https://doi.org/10.1111/j.1365-2656.2009.01583.x>.
47. Peres-Neto PR, Jackson DA, Somers KM. Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. *Ecology.* 2003;84(9):2347–2363. <https://doi.org/10.1890/00-0634>.
48. Perry RW, Pope AC, Romine JG, Brandes PL, Burau JR, Blake AR, Ammann AJ, Michel CJ. Flow-mediated effects on travel time, routing, and survival of juvenile Chinook salmon in a spatially complex, tidally forced river delta. *Can. J. Fish. Aquat. Sci.* 2018;75(11):1886–1901. <https://doi.org/10.1139/cjfas-2017-0310>.
49. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2020. <http://www.R-project.org>.
50. Romine JG, Perry RW, Johnston SV, Fitzer CW, Pagliughi SW, and Blake AR. Identifying when tagged fishes have been consumed by piscivorous predators: application of multivariate mixture models to movement parameters of telemetered fishes. *Anim Biotelem.* 2014;2(1):3. <https://doi.org/10.1186/2050-3385-2-3>.
51. Sandstrom PT, Ammann AJ, Michel C, Singer G, Chapman ED, Lindley S, MacFarlane RB, Klimley AP. Growth, survival, and tag retention of steelhead trout (*Oncorhynchus mykiss*) and its application to survival estimates. *Environ Biol Fish.* 2013;96(2):145–164. <https://doi.org/10.1007/s10641-012-0051-0>.
52. Schultz AA, Kumagai KK, Bridges BB. Methods to evaluate gut evacuation rates and predation using acoustic telemetry in the Tracy Fish Collection Facility primary channel. *Anim Biotelemetry.* 2015;3(1):13. <https://doi.org/10.1186/s40317-015-0034-y>.
53. Seitz AC, Courtney MB, Evans MD, Manishin K. Pop-up satellite archival tags reveal evidence of intense predation on large immature Chinook salmon (*Oncorhynchus tshawytscha*) in the North Pacific Ocean. *Can J Fish Aquat Sci.* 2019;76(9):1608–1615. <https://doi.org/10.1139/cjfas-2018-0490>.
54. Skalski JR, Smith SG, Iwamoto RN, Williams JG, Hoffmann A. Use of passive integrated transponder tags to estimate survival of migrant juvenile salmonids in the Snake and Columbia rivers. *Can J Fish Aquat Sci.* 1998;55(6):1484–1493.
55. Skalski, JR, Whitlock SL. Vitality models found useful in modeling tag-failure times in acoustic-tag survival studies. *Anim Biotelemetry* 2020;8:26. <https://doi.org/10.1186/s40317-020-00213-z>.
56. Thompson, B.C., Gwinn, D.C., and Allen, M.S. Evacuation times of radio transmitters consumed by largemouth bass. *N Am J Fish Manag.* 2015;35(4):621–625. doi:10.1080/02755947.2015.1017120.
57. Thorstad, E.B., Uglem I, Finstad B, Chittenden CM, Nilsen R, Økland F, Bjørn PA. Stocking location and predation by marine fishes affect survival of hatchery-reared Atlantic salmon smolts: survival and migration of hatchery-reared smolts. *Fish Manag Ecol.* 2012;19(5):400–409. <https://doi.org/10.1111/j.1365-2400.2012.00854.x>.
58. Townsend RL, Skalski JR, Dillingham P, Steig TW. Correcting bias in survival estimation resulting from tag failure in acoustic and radiotelemetry studies. *J. Agric. Biol. Environ. Stat.* 2006;11:183–196. <https://doi.org/10.1198/108571106X111323>.
59. Vogel DA. Evaluation of acoustic-tagged juvenile Chinook salmon movements in the Sacramento–San Joaquin delta during the 2009 Vernalis Adaptive Management Program. Technical Report for San Joaquin River Group Authority. 2010. https://www.waterboards.ca.gov/waterrights/water_issues/programs/bay_delta/bay_delta_plan/water_quality_control_planning/docs/sjrf_spptinfo/vogel. Accessed 5 June 2021.
60. Ward JH. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc.* 1963;58(301):236–244.
61. Weinz AA, Matley JK, Klinard NV, Fisk AT, Colborne S.F. Identification of predation events in wild fish using novel acoustic transmitters. *Anim Biotelemetry.* 2020;8(1):28. <https://doi.org/10.1186/s40317-020-00215-x>.

Figures

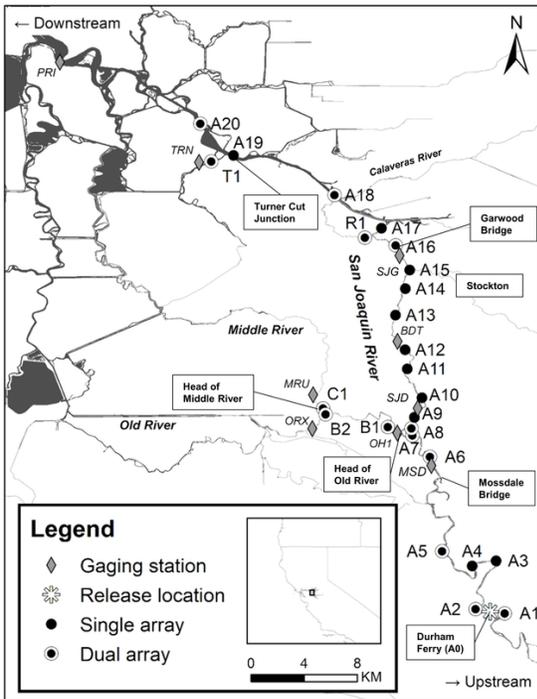


Figure 1

Map of the San Joaquin River and Sacramento–San Joaquin Delta with release site, acoustic telemetry stations (single and dual arrays), and gaging/environmental monitoring stations used in the 2016 acoustic telemetry station.

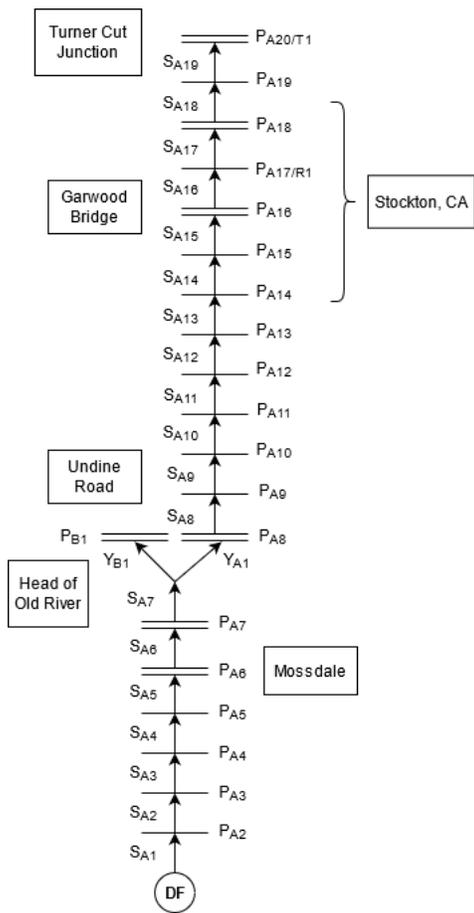


Figure 2
 Schematic of mark-recapture model used to estimate survival in 2016 acoustic telemetry study in the San Joaquin River. DF = release site at Durham Ferry. Horizontal lines indicate telemetry stations, parallel lines indicate dual array stations. Parameters: S = survival probability, P = detection probability, Y = route selection probability.

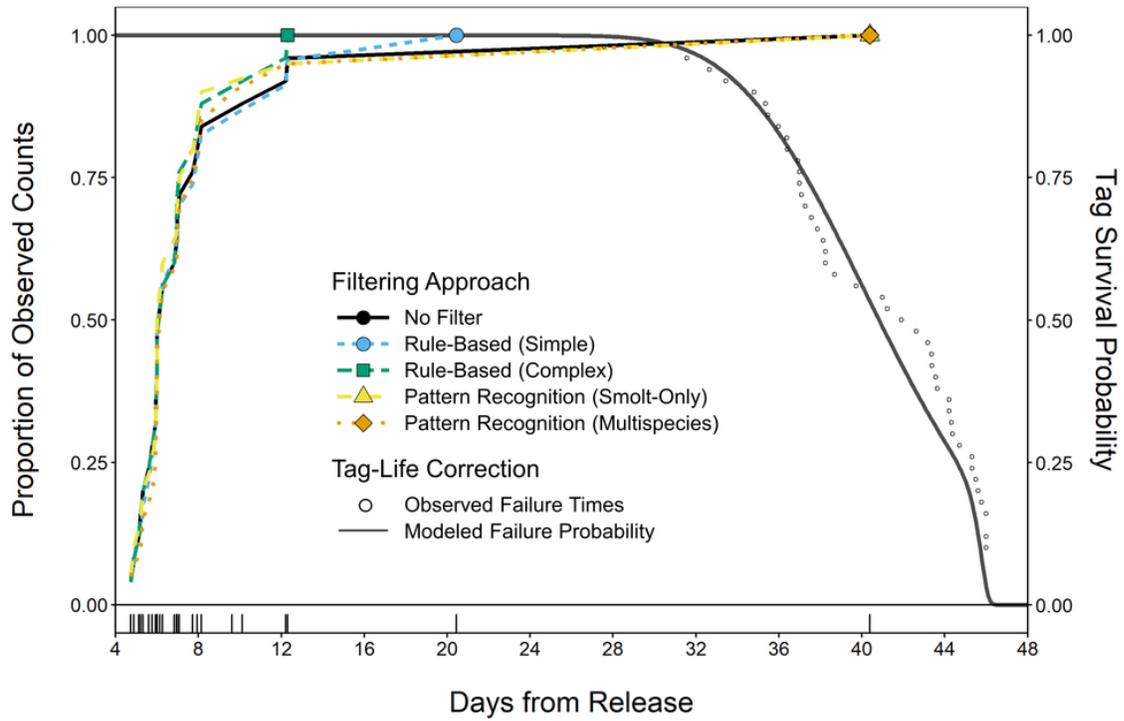


Figure 3
 Arrival timing distribution to the Turner Cut Junction (stations A20, T1) for salmon tags classified by filtering approach (bold and colored lines), observed tag survival from the tag-life study (open dots), and fitted tag survival curve using the four-parameter vitality model. Filled icons indicate last arrival at A20/T1 for the filtering approach; the final detection time observed from the no-filter approach and the pattern recognition filters coincided. Failure times for tag-life data were offset by average delay between tag activation and fish release based on the salmon tags released (2.3 d). The rug plot (horizontal axis) represents the arrival timing of unfiltered salmon tags.

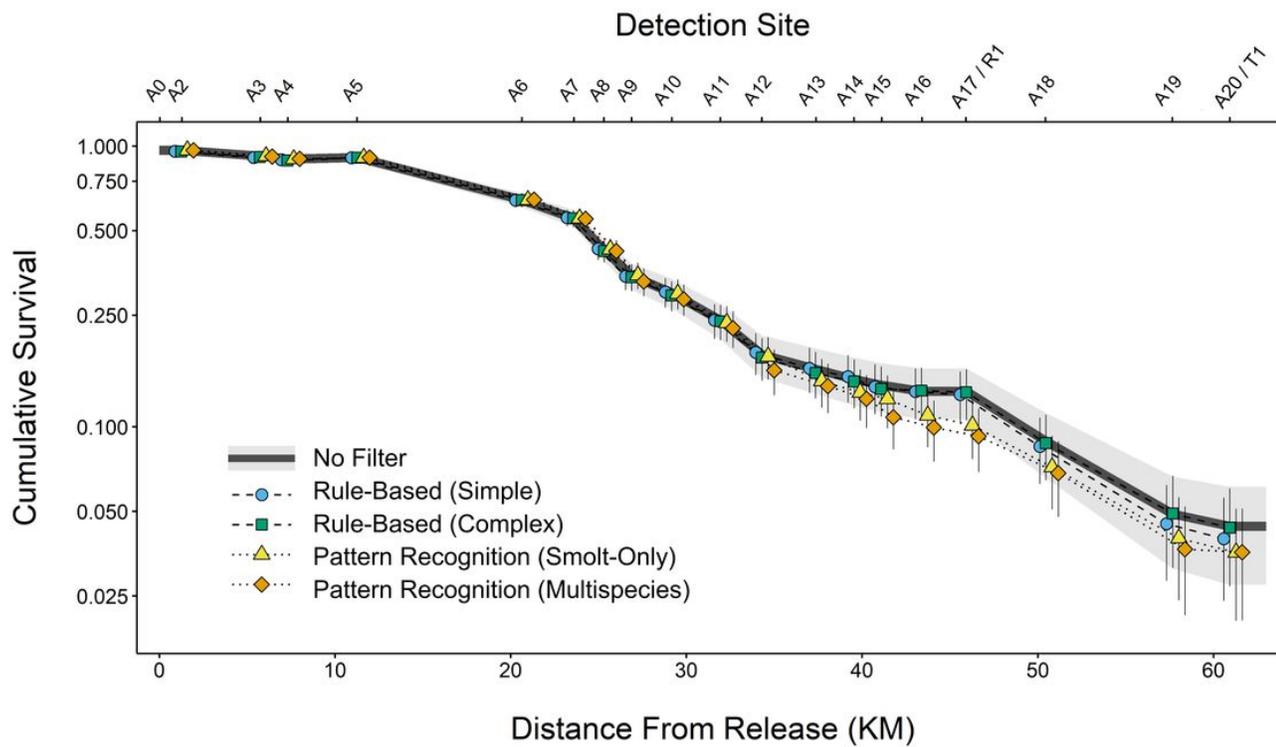


Figure 4
 Cumulative survival probability estimates from release to each telemetry array in the core area (A1 – A20/T1, B1) classified by filtering approach. The horizontal axis is scaled by the distance from release; the vertical axis is shown on the log2 scale to highlight survival < 0.250. A0 = release at Durham Ferry, A20/T1 = Turner Cut Junction. Error bars represent 95% confidence intervals; shading represents 95% confidence interval for no-filter approach.

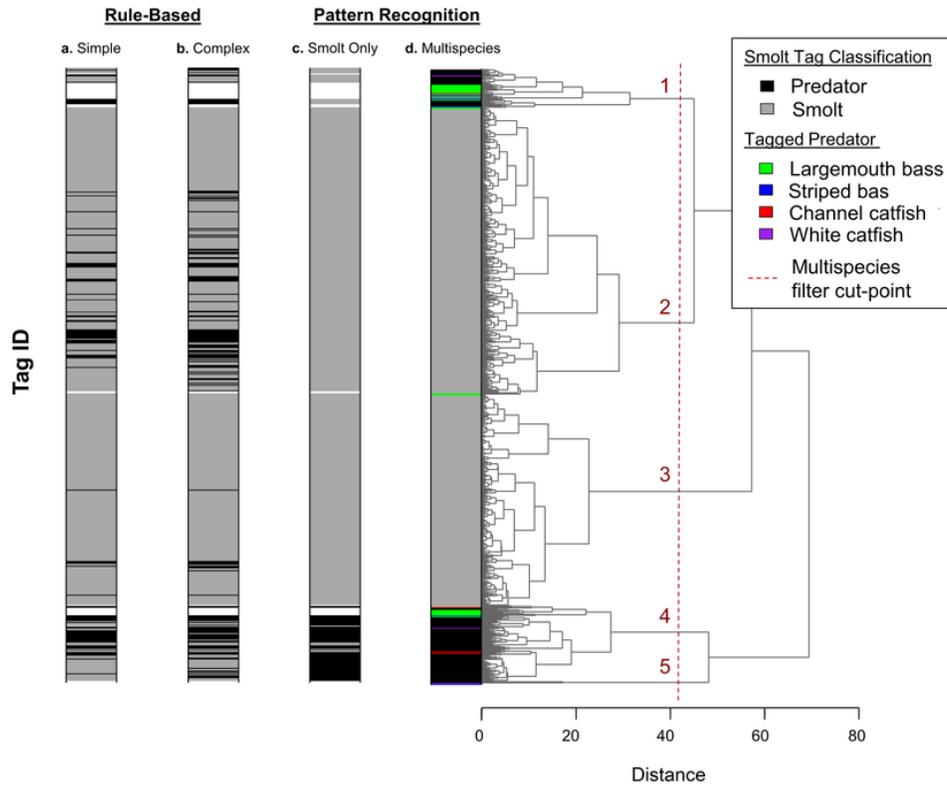
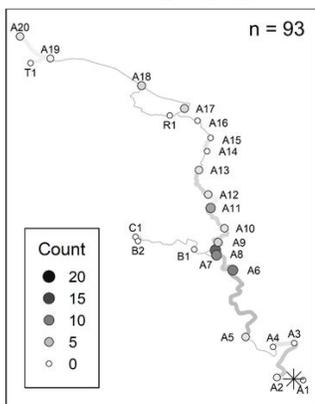


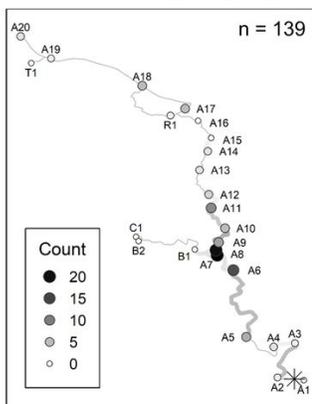
Figure 5

Classifications of smolt tags as predators (black bars) or smolts (gray bars) by the end of the study of acoustic-tagged Chinook salmon smolts according to four filtering approaches: simple rule-based filter (a), complex rule-based filter (b), smolt-only pattern recognition filter (c), and multispecies pattern recognition filter (d). White bars in the smolt-only pattern recognition filter and the rule-based filters are placeholders for the known predator tags used in the multispecies pattern recognition filter. Tags are ordered according to a Ward hierarchical cluster analysis of behavioral metrics from the multispecies pattern recognition filter (dendrogram shown). The red dashed line indicates distance threshold used to divide the multispecies dataset into five clusters (numbered 1–5); clusters 1, 4, and 5 were classified as predated.

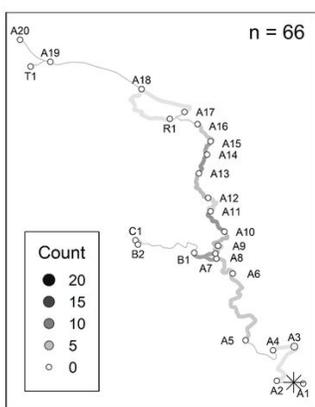
a. Rule-Based (Simple)



b. Rule-Based (Complex)



c. Pattern Recognition (Smolt Only)



d. Pattern Recognition (Multispecies)

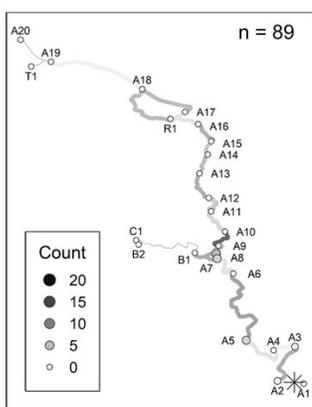


Figure 6
Comparison of locations where predation events were assigned for tags implanted in Chinook salmon smolts in the 2016 case study in the San Joaquin Delta according to four alternative predator filter methods.

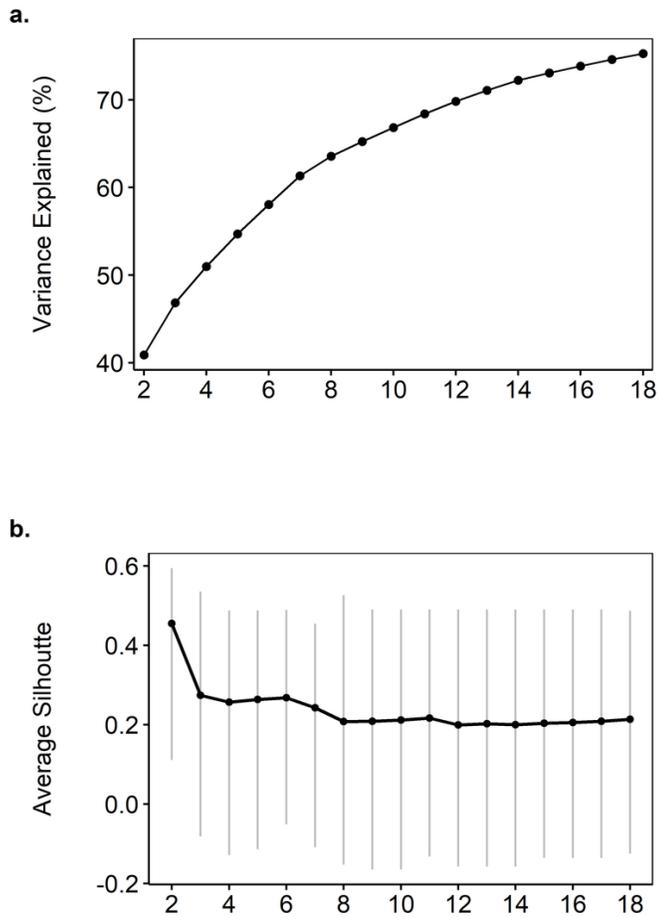


Figure 7
Visualizations used in determining optimal number of clusters in smolt-only pattern recognition analysis for case study of acoustic-tagged Chinook salmon smolts in the San Joaquin River, California, United States.

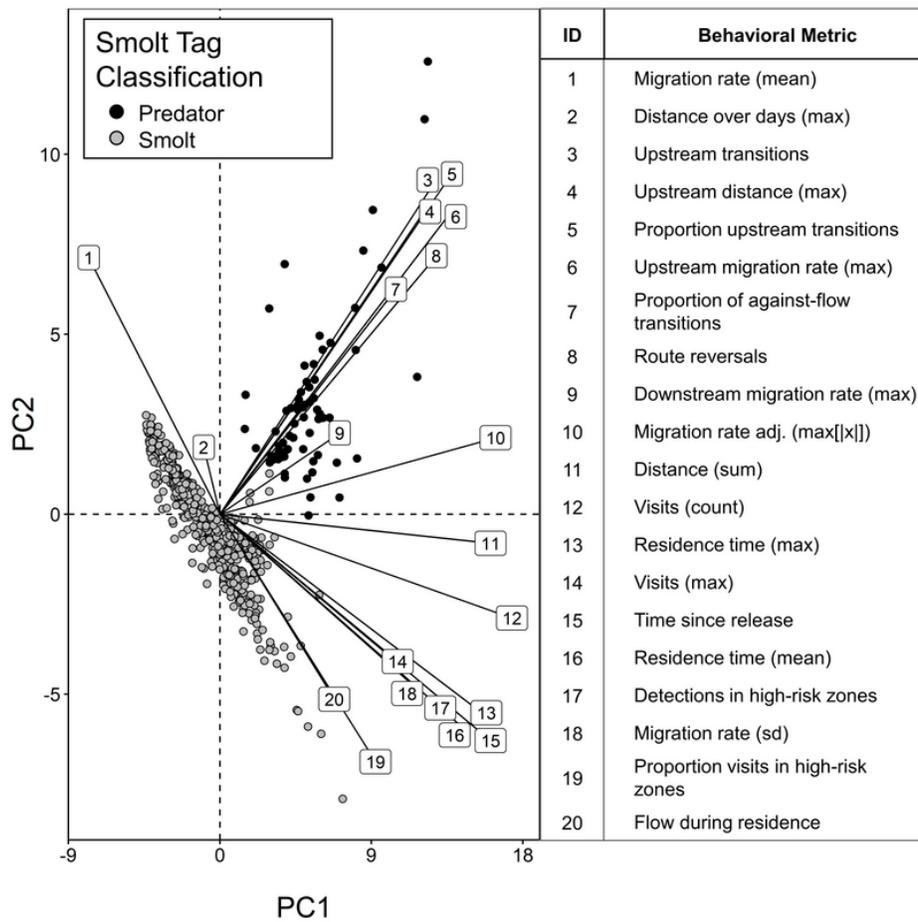


Figure 8
 Ordination plot of smolt tags (points) along the first and second principal component axes based on summarized tag-level metrics for smolt-only pattern recognition filter. Tag classification as smolt or predator was based on partition of data into two clusters in Ward hierarchical cluster analysis. Labeled vectors are relative PCA loadings of observed behavioral metrics from full smolt-tag data set.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TagPredationSupplementaryMaterials.pdf](#)