ADULT AMERICAN SHAD DOWNSTREAM FISH PASSAGE EVALUATION AT HANOVER POND DAM

HANOVER HYDROELECTRIC PROJECT NO. 14550

Prepared for:

New England Hydropower Company, Inc.
Meriden, Connecticut

Prepared by:

Kleinschmidt

Essex, Connecticut
www.KleinschmidtGroup.com

Project No. 4161007.01

September 2019

ADULT AMERICAN SHAD
DOWNSTREAM FISH PASSAGE
EVALUATION AT HANOVER POND DAM

2019

HANOVER POND DAM HYDROELECTRIC PROJECT

Prepared for:

New England HydroPower Company, Inc.
Meriden, Connecticut

Prepared by:

Kleinschmidt

Essex, Connecticut
www.KleinschmidtGroup.com

Project No. 4161007.01

September 2019

Acknowledgment

New England HydroPower Company thanks the
Connecticut Department of Energy and Environmental Protection for their assistance in this study.

ADULT AMERICAN SHAD
DOWNSTREAM FISH PASSAGE
EVALUATION AT HANOVER POND DAM

2019

HANOVER POND DAM HYDROELECTRIC PROJECT

TABLE OF CONTENTS

LIST OF PHOTOS

LIST OF FIGURES

LIST OF TABLES

APPENDICES

J:\4161\007\Docs\Report\Hanover Pond Fish Passage Report_FINAL 10-22-19.docx

ADULT AMERICAN SHAD
DOWNSTREAM FISH PASSAGE
EVALUATION AT HANOVER POND DAM

2019

HANOVER POND DAM HYDROELECTRIC PROJECT

1.0 INTRODUCTION

New England Hydropower Company, LLC (NEHC) designed and conducted a downstream shad passage study in the spring of 2019. The study was developed in consultation with the Connecticut Department of Energy and Environmental Protection (CT DEEP) and the U.S. Fish and Wildlife Service (USFWS). The evaluation of American Shad downstream fish passage at Hanover Dam was conducted to better understand if the downward-closing sluice gate affects downstream passage of Adult American Shad. As specified in the 401 Water Quality Certificate, if the results of the evaluation indicate that operations of the sluice gate inhibits movement down through the Archimedes Screw Generator (ASG) turbine, the Permittee shall implement mitigation measures as directed by the CT DEEP and USFSW.

2.0 PROJECT DESCRIPTION

The Hanover Pond Dam (Dam) is located at river mile 22 on the Quinnipiac River in Meriden, CT. The Dam is 430 feet wide with a spillway that is approximately 242 feet long. On the eastern side of the Dam is a concrete wall, which is bordered by Hanover Road. On the western side of the spillway there is a fish ladder measuring 175-feet long. Also, on the western end of the spillway is the ASG used to generate power. The project is operated run-of-river mode and utilizes the existing head of 14.7 feet (i.e., difference between the upstream and downstream water levels) for the purposes of generation. The ASG is enclosed in a concrete housing structure (Photo 1) consisting of a trough that holds the screw at an optimal angle of 30 degrees. The screw turbine has three blades with a runner diameter of 139.75 inches. The penstock is constructed of precast concrete sections measuring 12 feet in diameter and conveys water through the earthen portion of the Dam to the ASG. The intake is controlled by a hydraulically powered sluice gate to ensure the system maintains a run-of-river flow design. The intake is protected by a trashrack with vertical bars spaced every 10 inches to allow passage of fish and small debris through the ASG.

Flow through the ASG is adjusted by an automated sluice gate to ensure that the reserve flow (30 cfs) is maintained over the Dam and through the low flow notch at all times, remaining flow may be used for power production up to the ASG max flow (192 cfs).

Photograph 1: Western side of Hanover Dam, looking upstream showing the spillway, fish ladder and enclosed ASG

3.0 OBJECTIVES

The goal of this evaluation was to investigate downstream passage effectiveness of adult American Shad. The specific objectives of the evaluation were to determine:

1. If downstream passage occurs through the ASG,
2. The proportion of shad that pass downstream of the Dam through each of the passage routes (i.e., entrainment, through the fishway, and/or passage over the Dam through the notch/spill), and
3. Passage survival through each of the passage routes.

4.0 METHODS

This evaluation was conducted by a study team including NEHC, Kleinschmidt Associates, and the CT DEEP biologists. The methods used in the evaluation were developed in consultation with the study team and the USFWS.

4.1 SHAD PASSAGE MONITORING

Kleinschmidt deployed and tested three radio telemetry monitoring stations in the Hanover Pond Dam Project area, including stations upstream of the Dam (T01), within the intake structure (T02) and downstream of the Dam (T03) to achieve the study objectives. A list of radio telemetry monitoring stations and equipment is provided in Table 1. Photographs 2, 3 and 4 display the location and orientation of each of the three fixed telemetry sites.

In early May 2019, prior to the release of any tagged fish, radio telemetry equipment was installed and calibrated to ensure coverage of the study area as depicted in Figure 1. Complete calibration results for each monitoring stations are in Appendix A - Telemetry Calibration. All fish were tagged and released on May 30, 2019 and the study concluded on July 15, 2019, which coincides with the approximate seasonal end of the shad migration and the time when no more tagged fish were detected in the study area.

Table 1: Shad monitoring locations and equipment for Hanover Pond Dam

Station Location	Station ID	Receiver Station Equipment
Upstream of Hanover Dam	T01	An Orion [1]receiver with two 3-element Yagi antennas
ASG Intake	T02	An Orion receiver with a dipole antenna
Downstream of Hanover Dam	T03	An Orion receiver with a 3-element Yagi antenna

[1] Orion radio telemetry receivers as manufactured and sold by Sigma Eight.

Photograph 2: Upstream of Hanover Pond Dam Yagi Antennas (Site T01)

Photograph 3: Intake Dipole Antenna at Hanover Pond Dam (Site T02)

Photograph 4: Downstream of Hanover Pond Dam Yagi Antenna (Site T03)

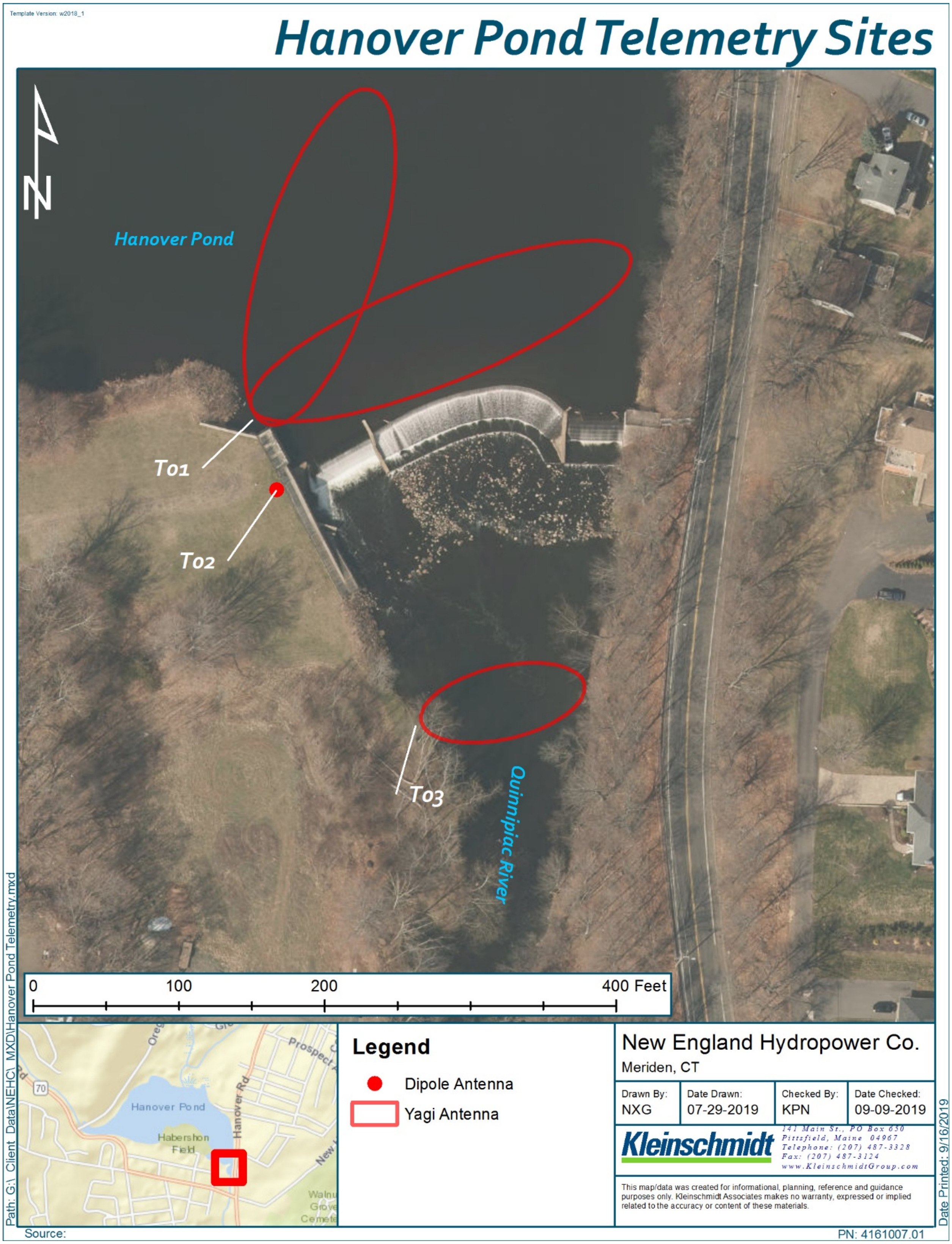

Figure 1: Map of the project area and the coverage of the three fixed telemetry monitoring stations

4.2 Shad Collection, Transport, Tagging and Release

American Shad were collected from the fish lift at the Holyoke Dam and transported by the CT DEEP to Hanover Pond to be radio-tagged and released.

Shad (20) were tagged with TX-PSC-I-80-M Pisces transmitters manufactured by Sigma Eight. The tags measured 9.6 mm by 26 mm and operated on the frequency 149.440 MHz. They were programed with a two-second burst and a motion sensing function to estimate mortality, which defaulted to an eleven-second burst upon activation. Activation of mortality was based on relative motionlessness for a period of 15 minutes. The expected tag life was approximately 90 days.

Tagging priority was given to fish that exhibited vigor, minimal scale loss (less than 10%, evaluated subjectively in the field), and fish that were greater than 400 mm in total length

4.3 Telemetry Analysis

Studies that assess movement of anadromous fish through telemetered river-reaches are complex in nature. Analysis is made difficult because of the presence of false positive signals and receivers with potential for overlapping detection zones. Considerable data cleaning is required before an assessment of movement can occur. Kleinschmidt implemented the following protocol to analyze radio telemetry data:

1. Identify and remove false positive detections with a Naïve Bayes Classifier
2. Reduce any potential overlap between detection zones
3. Assess movement with time-to-event analysis using a competing risks framework

A complete synopsis of the data reduction and statistical methods applied, as well assumptions, is provided in Appendix B Statistical Methods.

Kleinschmidt developed a simple competing risks model to describe movement through the project area. The competing risk model assessed downstream movement through the Hanover Pond Dam, whether through the intake and the ASG or over the Dam via the spillway and the low-flow notch. For the purposes of this report, movement occurs between three telemetered reaches. The model always assume movement from an initial location or spoke. The initial state for this model was the upstream station (T01). The counting process style data were arranged so that the first detection for every fish was always in the initial state (i.e., upstream of the Dam where the tagged shad were released). From there, a detection at any other telemetered site would represent downstream movement. For example, a fish with the detection history beginning upstream (T01), then through the intake (T02), followed by detection downstream of the Dam (T03), passage through the ASG was assumed. If a fish was detected upstream of the Dam (T01), then subsequently downstream of the Dam (T03), the fish either spilled over the Dam or utilized the notch (Figure 2). If a shad were to successfully pass through the ASG, the detection history would be T01, to T02 and finally to T03. The model assumes that both the intake and the downstream are absorbing states, because fish cannot move back upstream once detected at those locations. These secondary movements were queried out of the initial competing risks assessment using methods of Therneau, Crowson, and Atkinson (2016 and 2017) and with data frame filtering in R™.

Operations and flow data used in this evaluation were available in an hourly format. Therefore, flow reported at the time of passage through the ASG is the nearest hourly average flow from the last detection upstream, detections within the intake, and the first detections downstream. If the duration of passage occurred over multiple hours, then the two nearest hourly flows were averaged to calculate passage flows. Similarly, for fish that passed using the notch or via the spillway, passage flow reported represents the average of the flows at the time of last detection upstream and the first detection downstream.

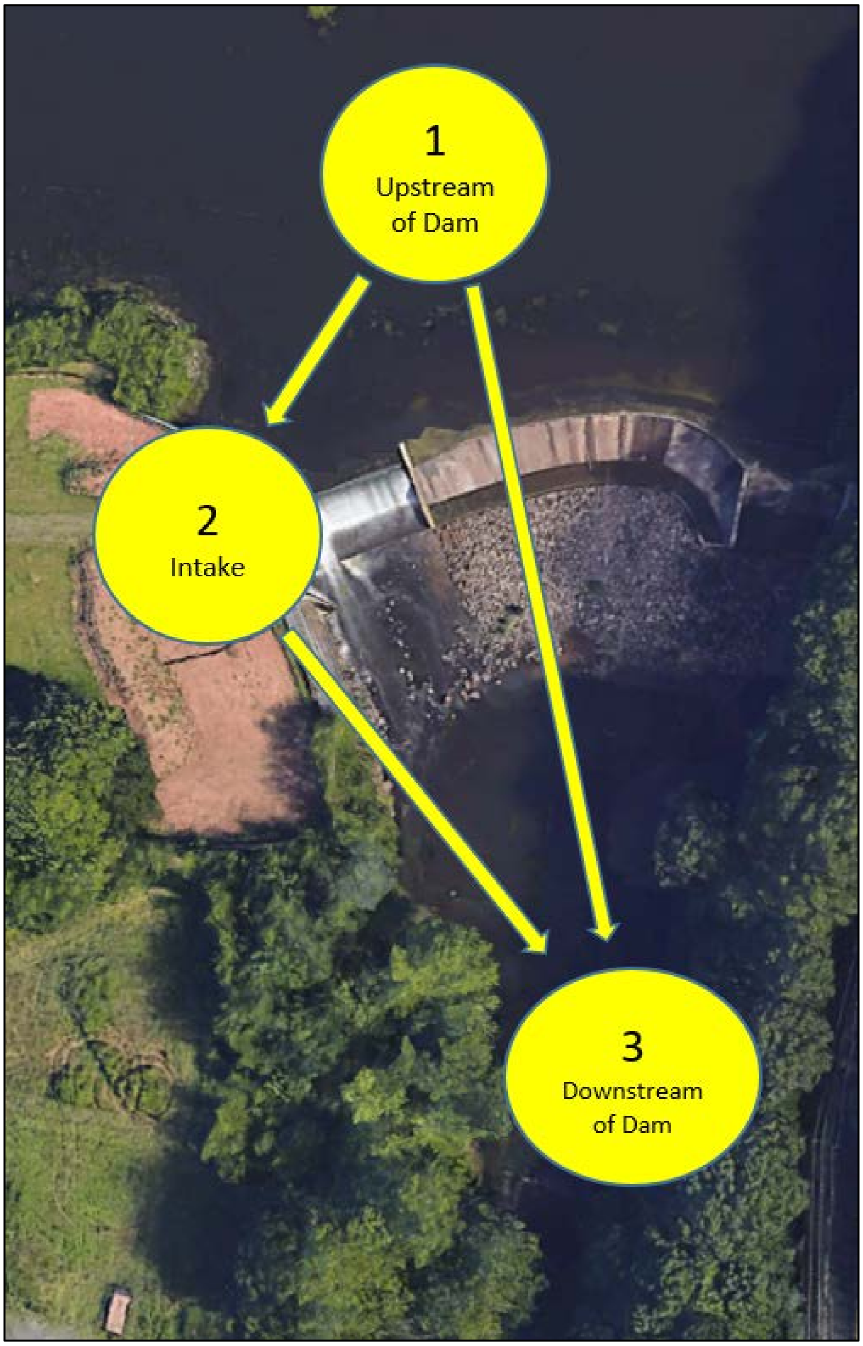

Figure 2: Competing risk model network for Hanover Pond Dam, showing routes of passage

5.0 RESULTS

5.1 SHAD TAGGING

Twenty American Shad were radio tagged on May 30, 2019 and released upstream of the Dam in the northwest corner of the pond from a parking lot accessed from Oregon Road (Figure 3). The average total length of tagged shad was 470 mm, and no shad measuring less than 400 mm were tagged for this study. The CT DEEP transported approximately 50 American Shad from Holyoke Dam to Hanover Pond. The sex ratio of tagged fish (20 fish) was 14 males and 6 females.

Water quality parameters were taken during the tagging process in two locations: in the DEEP holding tank located on the truck and in Hanover Pond at the shad release site (Figure 3). The water quality parameters are located below in Table 2.

Table 2: Water Quality parameters inside the DEEP holding tank and at Hanover Pond release site, May 30, 2019

	Location	
Water Quality parameter	**DEEP Holding Tank**	**Hanover Pond release site**
Temperature (°C)	15.4	17.68
pH	7.66	8.15
DO (mg/l)	21.93	11.72
DO (%)	206.1	123.9

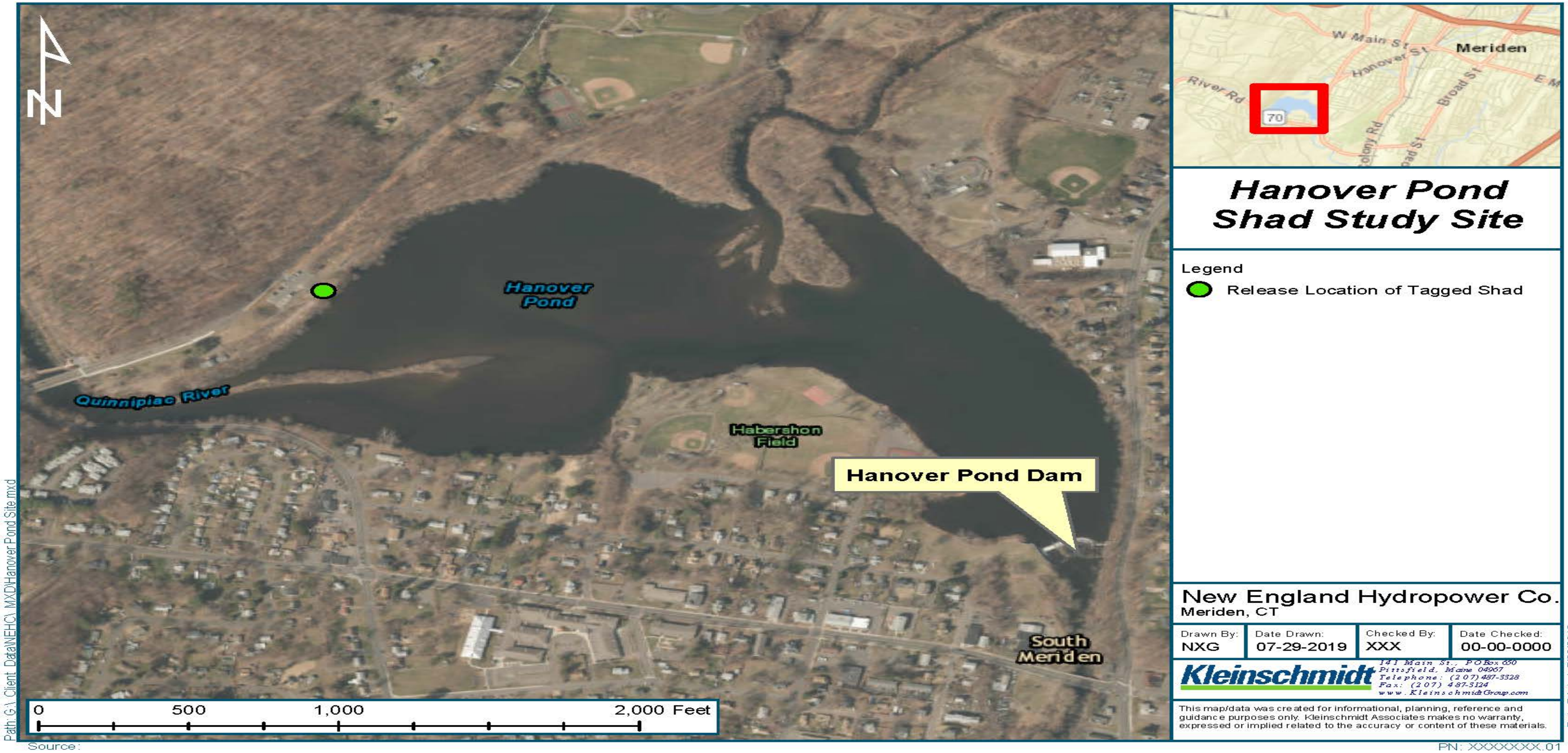

Figure 3: Tagging and release location of adult American Shad

5.2 DOWNSTREAM SHAD PASSAGE

In total, there were 16 tagged shad detected upstream of the Dam at Station T01. Eight of those fish passed downstream and were detected at Station T03. Seven of the eight fish that passed downstream (87.5%) utilized the intake (T02) and passed through the ASG before being detected at the downstream site (T03). One fish passed downstream via the spillway or use of the notch, as evidenced by detections upstream (T01), followed by subsequent detections downstream (T03), with no detections at the intake (T02). The four undetected shad either did not migrate, expelled the tag outside of the range of the telemetry network, or experienced mortality outside of the range of the telemetry network. A fourth but unlikely scenario is that these tagged shad were not detected at any of the monitoring locations as they passed through the system. The fate of the 8 fish that approached the dam but did not pass downstream remains unknown. Many of these fish were detected at Station T01 several times, spanning multiple days, and all of the tags continued to emit a 2-second burst rate throughout the entirety of their upstream detections, suggesting that while detected, they were all still alive. Table 2 displays the frequency and unique ID code of each fish that passed downstream, including the route of passage, the site flow (cfs) at time of passage, the flow through the ASG (cfs) at passage, and the flow over the Dam during time of passage.

The average project flow during the time of passage for the eight fish that passed downstream was 169.9 cfs. The average flow through the ASG and over the notch/Dam via the spillway during passage events was 76.8 cfs and 93.2 cfs, respectively. The fish that passed via the notch/spillway experienced the highest project discharge (195 cfs), with the highest flow through the ASG (108.4 cfs), and the lowest flow over the Dam (86.6 cfs) as compared to the other seven fish that passed (Table 3). All eight fish passed downstream between June 2 and June 13, 2019. Each fish that passed downstream and through T03 appeared to survive as the tags continued to emit a 2-second burst rate and there were no tags that switched to the 11-second burst rate (mortality signal). The telemetry equipment was monitoring continuously until July 15 with no outages and did not detect any more fish moving through the system after June 13, 2019.

Table 3: Description of conditions for fish that passed downstream at Hanover Pond Dam

Frequency & Code	Route of Passage	Date & Hour of Passage	Project Discharge at Passage (cfs)	Flow Through ASG at Passage (cfs)	Spill/Notch Flow at Passage (cfs)
149.440 21	ASG	6/8/2019 2100	142.5	50.8	91.7
149.440 28	ASG	6/13/2019 0700	163.3	71.1	92.2
149.440 31	ASG	6/4/2019 0400	176.5	79.9	96.6
149.440 33	ASG	6/7/2019 2000	150.5	55	95.5
149.440 36	ASG	6/5/2019 0300	165.5	68.7	97.7
149.440 38	ASG	6/3/2019 2100	183.5	91.4	92.1
149.440 39	ASG	6/3/2019 2200	182.5	89.1	93.4
149.440 35	Spill or Notch	6/2/2019 2100	195	108.4	86.6

5.3 ENVIRONMENTAL AND PROJECT OPERATIONS DATA

Total discharge at Hanover Pond Dam was available on an hourly basis and is displayed along with flow through the ASG and flow over the Dam in Figure 4 for the duration of the study (May 30 to July 15, 2019). The maximum total discharge at the site during the study was 828.5 cfs and occurred on June 18, 2019. The minimum discharge at the site during the study was 60.25 cfs and occurred on July 15, 2019. The eight fish that passed downstream did so between June 2 and June 13, experiencing site flows between 142.5 and 195 cfs.

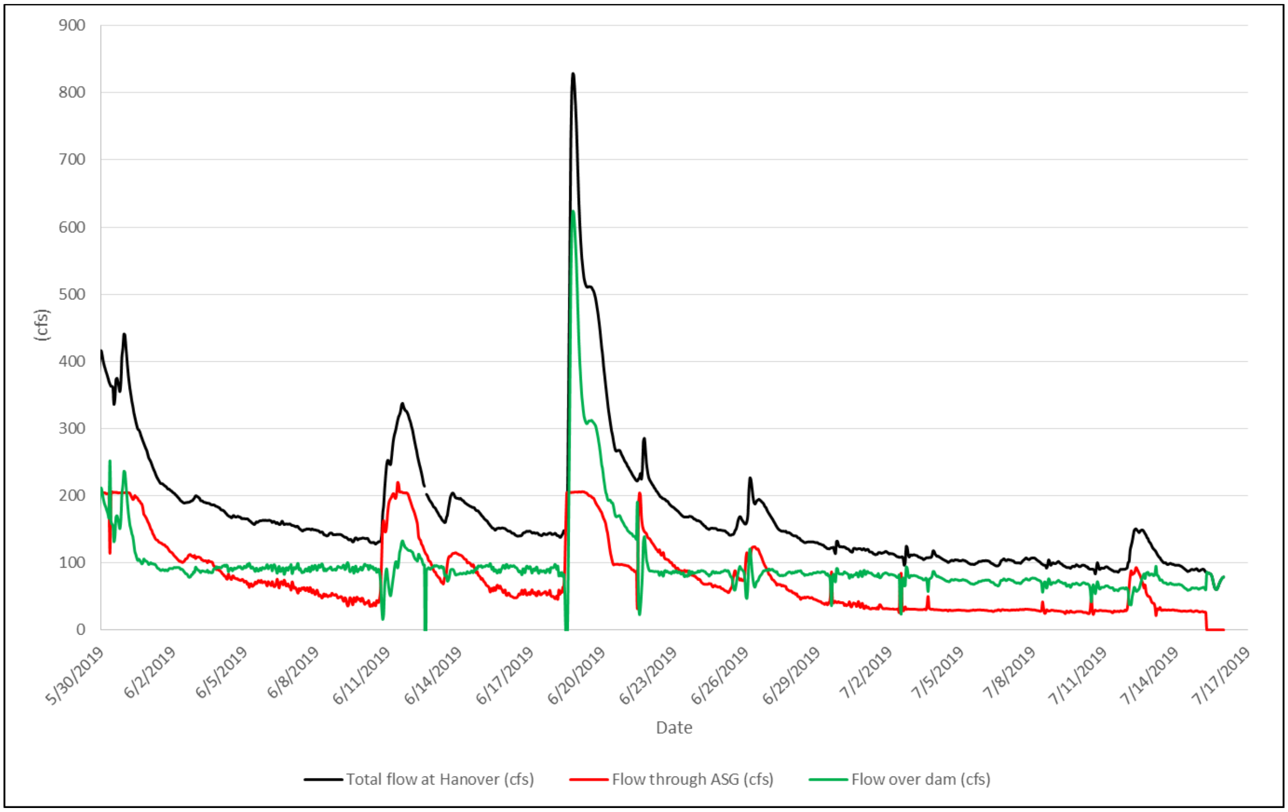

Figure 4: Total flow at Hanover Pond (cfs), flow through the ASG (cfs), and flow over the Dam (cfs) from May 30 to July 15, 2019

6.0 DISCUSSION AND CONCLUSION

This study was successful at collecting, tagging, and monitoring Adult American Shad as they moved through the study area.

A total of eight of the 20 tagged fish passed downstream of Hanover Pond Dam. Downstream passage was most numerous through the ASG turbine as verified by the seven radio tagged shad detected inside of the ASG intake, before being subsequently detected downstream. The dipole antenna in the intake was located immediately upstream of the sluice gate, just before the fish entered the ASG as shown in Photo 3. The range of dipole antenna during testing and calibration did not detect any “test tag” outside of this area; therefore, all fish detected at the intake (T02) are known to have swam under the gate successfully. The seven fish that were detected in the intake were subsequently detected downstream (T03) and remained alive as evidenced by the 2-second burst rate emitted from the radio tags. Downstream passage through the ASG indicates that all seven shad successfully entered the darkened penstock, passed beneath the sluice gate, and utilized the screw safely as indicated by the subsequent 2-second burst detections at T03.

There was one fish that passed downstream via spill or the notch, and was detected downstream with a 2-second burst rate. Therefore, survival for the eight fish that successfully passed downstream through the ASG, and via spillway/notch was 100%. Proportionally, seven of the eight downstream fish (87.5%) passed utilizing the route through the ASG, and one (12.5 %) spilled over the dam or used the notch.

Demonstration of downstream passage through the ASG was successful for this study, and passage survival for all routes of downstream passage was 100%.

There were eight fish that approached the Dam and were detected with a 2-second burst rate at Station T01 but not pass downstream of the Dam (i.e., at T03). The fate of these 8 fish remains unknown, but potentially they may have remained in the pond, shed their tags, or more unlikely passed downstream without being detected at stations T02 or T03.

7.0 LITERATURE CITED

Therneau, T., Crowson, C., & Atkinson, E. (2016, October). Multi-state models and competing risks. From https://cran.r-project.org/web/packages/survival/vignettes/compete.pdf

Therneau, T., Crowson, C., & Atkinson, E. (2017). Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model.

APPENDIX A

TELEMETRY CALIBRATION

APPENDIX A: Telemetry Network Calibration and Equipment Effectiveness

Radio Telemetry Calibration

Each telemetry station was calibrated with a transponder (or radio tag) prior to release of test fish to ensure adequate power readings, range, and proper calibration of equipment. A uniquely coded transponder was used as a '*test tag*' during the calibration period. This code was not repeated for any tags inserted into test fish used for this study. The test tag was attached to fishing line and deployed to a water depth of approximately 3 to 5 feet to mimic the swimming depth of adult American Shad. One member of the field crew remained on land monitoring the receiver output signals and another field staff cast the tag into the targeted detection zone at each telemetry station. Communication via handheld two-way radios allowed transfer of power signals at different locations that were recorded on a map for calibration purposes.

A list of the receivers used for this study is provided in Table 1 of the main report. Orion receivers output an average power number for each detection, which is recorded in decibel levels (db). These numbers are negative, with less negative numbers being higher in signal strength.

All figures of the telemetry stations below show the position of the 'test tag' and the average power levels associated within the detection zones recorded during testing (noted in white). Several test detections were recorded at each location.

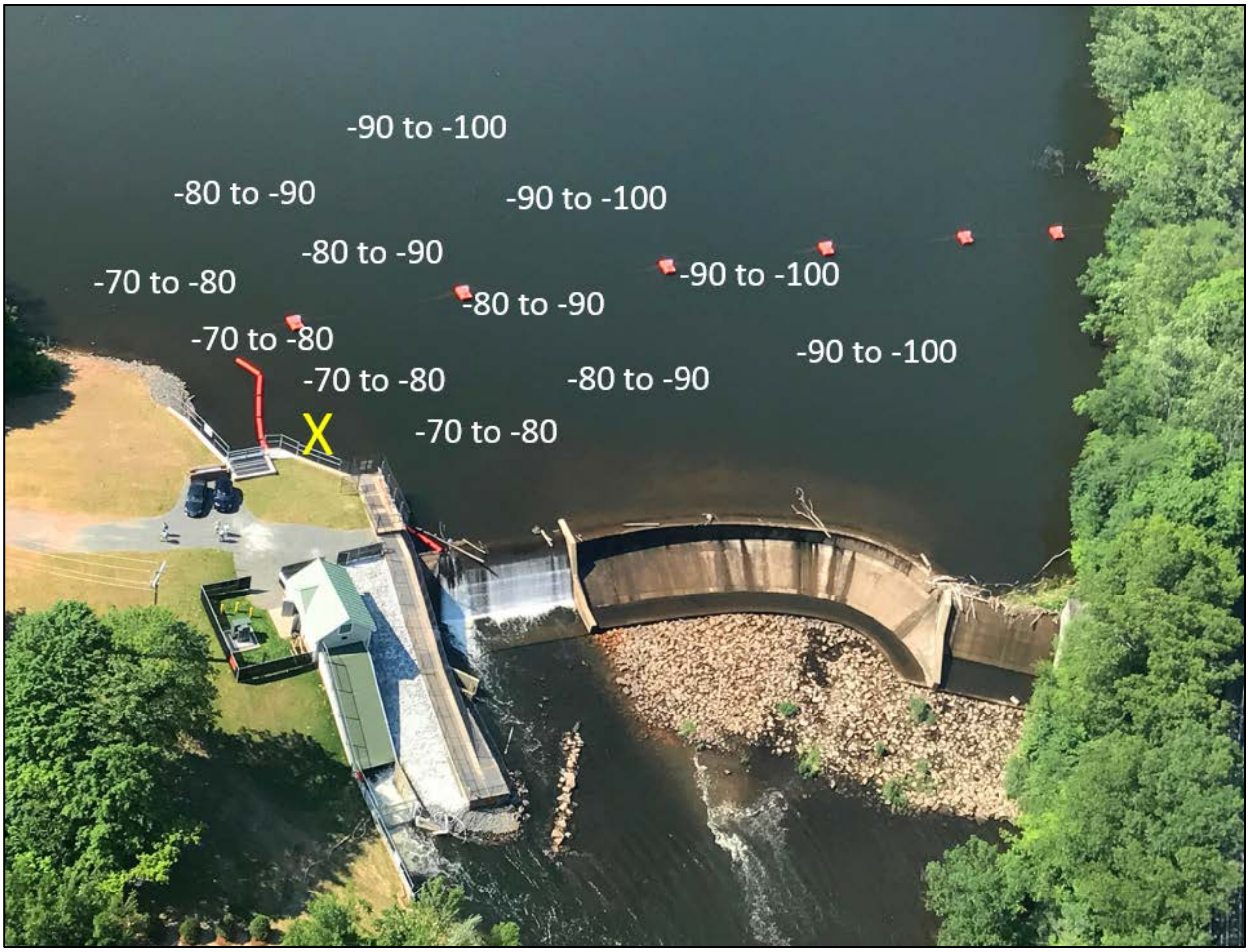

Figure B-1: Telemetry calibration power readings for the orion receiver located upstream of Hanover Dam (T01)

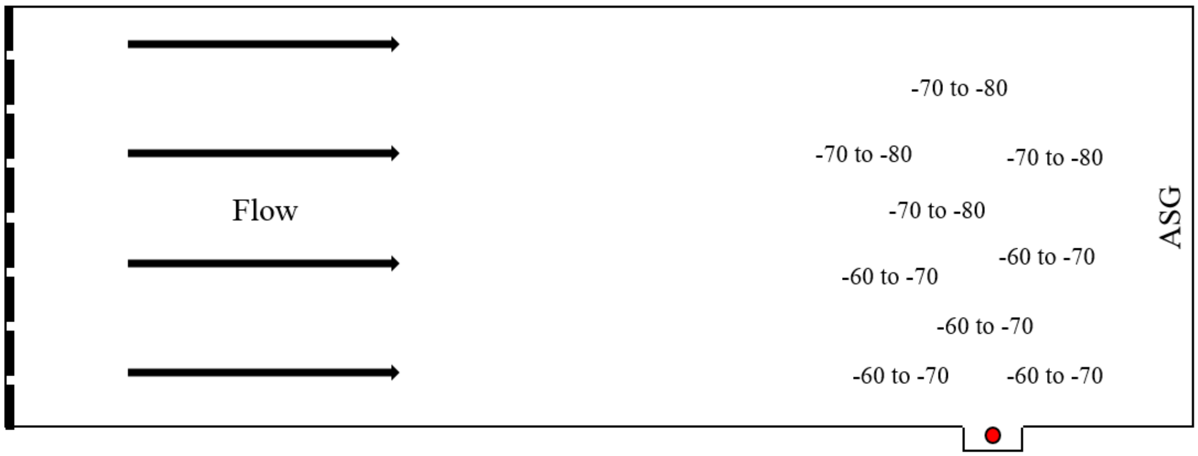

Figure B-2: Telemetry calibration power readings for the Orion receiver located in the intake of Hanover Dam ASG (T02)

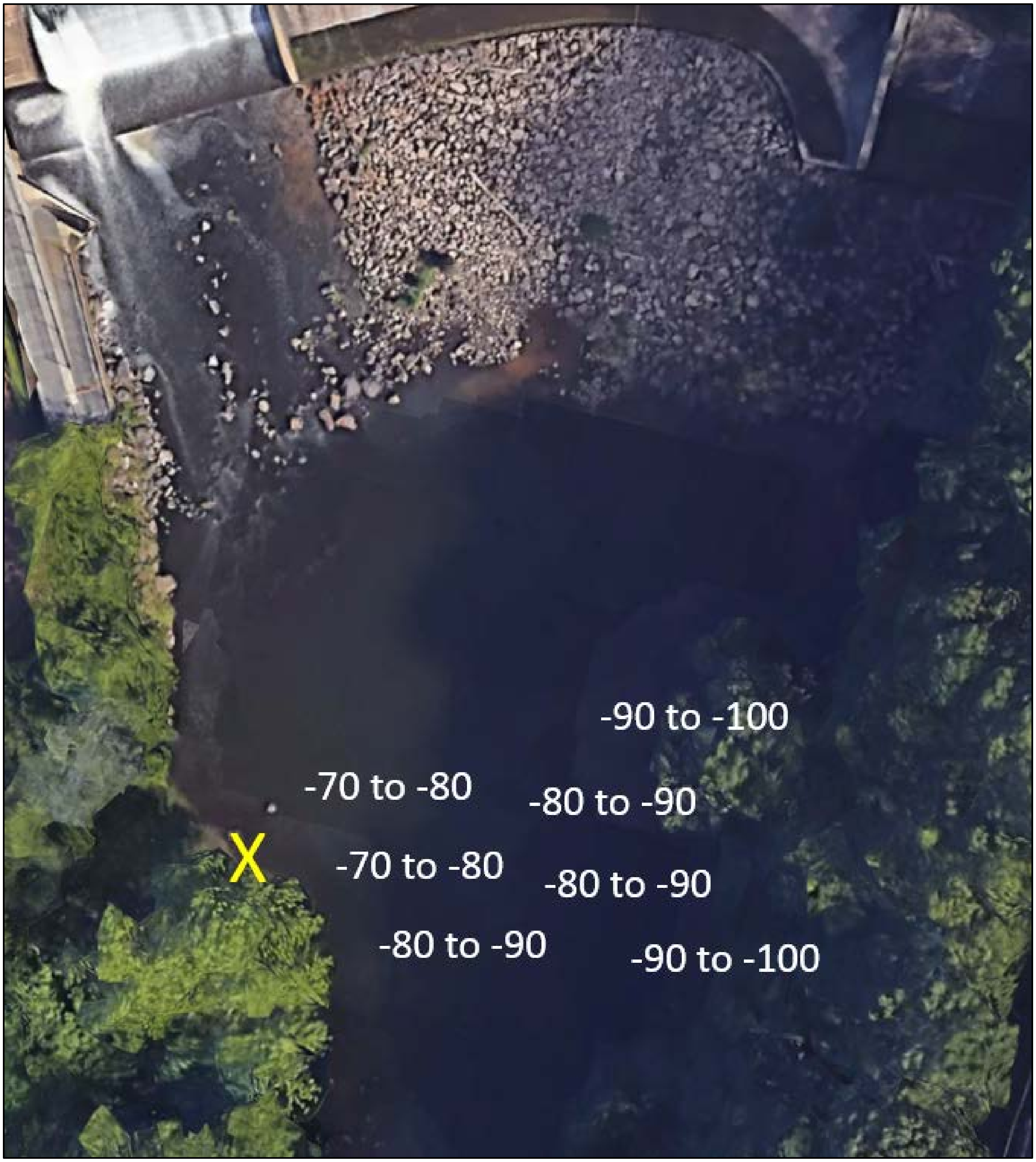

Figure B-3: Telemetry calibration power readings for the Orion receiver located downstream of Hanover Dam (T03)

APPENDIX B

STATISTICAL METHODS

1. METHODS

This appendix contains the unabridged statistical methods used to assess movement of radio-tagged American shad at the Hanover Pond Dam Hydroelectric Project. Statistical assessments of movement with radio telemetry are complex due to the amount of data produced, inclusion of false positives and the possibility of overlap between receiver detection zones. Prior to analyzing movement within a competing risks framework, we implemented an algorithm that identifies and removes false positive detections while not being so strict as to introduce false negatives into the dataset, then we implemented an overlap reduction algorithm that reduced ambiguity in a fish's position. Both of these steps are necessary as they reduce bias and ambiguity that have traditionally plagued assessments of movement with radio telemetry.

a. False Positive Data Reduction

A radio telemetry receiver records four types of detections based upon their binary nature; true positives, true negatives, false positives and false negatives (Beeman and Perry, 2012). True positives and true negatives are valid data points that indicate the presence or absence of a tagged fish. A false positive is a detection of a fish's presence when it is not there, while a false negative is a non-detection of a fish that is there. False negatives arise from a variety of causes including insufficient detection areas, collisions between transmitters, interference from ambient noise, or weak signals (Beeman & Perry, 2012). Inclusion of false negatives may negatively bias statistics as there is no way to know if a fish's absence from a receiver was because it truly wasn't there or if it was not recaptured by the receiver. While the probability of false negatives can be quantified from sample data as the probability of detection, quantifying the rate of false positives (type I error) is more problematic (Beeman & Perry, 2012). Inclusion of false positives in a dataset can bias study results in two ways: they can favor survivability through a project by including fish that weren't there, or increase measures of delay when a fish has already passed. There are no statistical approaches that can reduce bias associated with false positives, therefore they must be identified and removed *a priori*. For the purposes of this study, false positive reduction methods relied upon a Naïve Bayes classifier and an overlap reduction algorithm inspired by nested Russian dolls.

i. Probabilistic Data Reduction – Weight of Evidence

Bayes Rule is a rigorous method for interpreting evidence in the context of previous experience or knowledge (Stone, 2013). Bayes Rule cannot guarantee the correct answer, but rather provides the probability that each alternative answer (either true or false positive) is correct. Bayes theorem updates conditional probabilities (probability of a record being true positive given some data) and is particularly useful when evaluating diagnostic tests (false positives and false negatives).

Specifically, Bayes Rule calculates the posterior probability, or the probability of a hypothesis occurring given some information about its present state, and is written with $P(\theta_i|x_j)$; where θ_i is the hypothesis (true or false positive) and x_j is observed data. Formally, Bayes Rule is expressed as:

$$P(\theta_i|x_j) = \frac{P(x_j|\theta_i)P(\theta_i)}{P(x_j)} \qquad \textit{Equation 1}$$

Where $(x_j|\theta_i)$ is referred to as the likelihood of the j^{th}data occurring given the hypothesis (θ_i); $P(\theta_i)$ is the prior probability of the i^{th}hypothesis (θ); and $P(x_j)$ is the marginal likelihood or evidence. In most applications, including this one, the marginal likelihood is ignored as it has no effect on the relative magnitudes of the posterior probability (Stone, 2013). Therefore, there is no need to waste computational effort by calculating the joint probability. We can state that the posterior probability is approximately equal to the prior probability times the likelihood or:

$$posterior \propto prior * likelihood \qquad \textit{Equation 2}$$

The prior probability is estimated by looking at how often each class (true or false positive) occurs in the training dataset, while the likelihood is estimated from the histogram of the values of each predictor (observed data) in the training dataset given each hypothesis (true or false positive) (Marsland, 2009). A kernel density function was fit for continuous predictors while qualitative predictors replied upon a multinomial probability distribution.

In most circumstances, the data (x) are usually vectors of feature values or predictor variables with n levels (x_n). As the dimensionality of x increases (number of predictor variables increase), the amount of data within each bin of the histogram of related variables shrinks, and it becomes difficult to estimate the posterior probability without more training data (Marsland, 2009). For example, long strings of continuous detections in series may only occur when the power of a detection is fairly high. Therefore, a simplifying assumption, the Naïve Bayes classifier, was employed.

ii. Naïve Bayes Classifier

The Naïve Bayes classifier assumes that the elements (j) of the feature vector x (predictor variables) are conditionally independent of each other given the classification (Marsland, 2009). Therefore, the probability of getting a particular string of feature values of predictor variables is equal to the product of multiplying all of the individual probabilities (Marsland, 2009). The likelihood is given with:

Equation 3

$$P(x_1, \dots, x_n | \theta_i) = \prod_{j=1}^{n} P(x_j | \theta_i)$$

Where n is equal to the number of features or predictor variables in x and θ_i is the hypothesis (either true or false positive). The classifier rule for Naïve Bayes is to select the detection class θ_i for which the following computation is maximized:

Equation 4

$$argmax \left\{ P(\theta_i | x_n) \propto P(\theta_i) * \prod_{j=1}^{n} P\left(x_j \middle| \theta_i\right) \right\}$$

The detection class θ_j with the maximum posterior probability classifies every line of data belonging to a study tag into one of two classes: true or false positive. This is known as the maximum a posteriori or MAP hypothesis (Marsland, 2009).

The Naïve Bayes classifier was nothing more than a database application designed to keep track of which feature gives evidence to which class (Richert & Pedro-Coelho, 2013). However, there were circumstances where a particular feature variable level did not occur for a given detection class in the feature dataset (e.g., false positive detection with very high power and many consecutive hits in series), meaning that the likelihood for that feature given a detection class is zero. When multiplied together, the posterior probability was zero and uninformative. Therefore, the Naïve Bayes classifier used add-one smoothing, which simply adds 1 to all histogram counts (Richert & Pedro-Coelho, 2013). The underlying assumption here is that even if the feature value was not seen in the training dataset for a particular detection class, the resultant likelihood probability would be close to zero allowing for an informative posterior.

The training dataset consists of known true and false positive detections. By placing study tags at strategic locations throughout the study area for the duration of the study, these beacon tags give the algorithm information on what a known true positive detection looks like. On the other hand, known false positive detections are generated by the telemetry receivers themselves, and consist of detections coded toward tags that were not present in the list of tags released for the study.

Following the completion of the study, several predictor features were calculated for each received line of data. Predictor features include a detection history of pulses, the consecutive record hit length, hit ratio, miscode ratio, consecutive detection, detection in series, and power. The pulse detection history consists of a string of 1s and 0s that look forwards and backwards in time from the current detection in series, and identifies whether or not a pulse from that particular tag was detected. For example, if a particular tag had a 2-second burst rate, the algorithm will look forwards and backwards in time 2 seconds, query the entire dataset, and then return 1 if it was detected or 0 if it was not. The algorithm looks forwards and backwards for a user-defined set of detection intervals. Consecutive detection length and hit ratio are derived from this detection history. Consecutive detection length simply counts the number of detections in series, while hit ratio is the ratio of the count of heard detections to the length of the detection history string (Table B-1).

Note from Table B-1 that both detection history events are considerably different, but they have the same hit ratios. The hit ratio counts the number of correctly assigned detections to the total number of detections within a user-defined set of time. The hypothesis behind this predictor stipulates that a detection is more likely to be true when there are less miscoded detections. Consecutive detections and detections in series are binary in nature and quite similar, but the consecutive detection feature was stricter. For consecutive detection to return as true, either the previous or next detection must occur within the next pulse (i.e., 2-second interval). Detections in series allow the previous or next detection to occur at intervals greater than the first pulse; however, recaptures need to be in series. For example, if the pulse rate is 2 seconds and the next consecutive detection was missed, series hit would return true if the next recorded transmission occurred on the 4th or 6th second. In other words, the pulse rate must be a factor of the difference in time between the present detection and next detection for a series hit to return true. The last predictor, power, is hypothesized to be higher for true detections than false positives.

Prior to classification, we assessed the accuracy of the Naïve Bayes false positive detection algorithm with a k-fold cross validation procedure. The cross validation procedure randomly assigned folds (1,…,10) to each row of data. Then, the procedure iterates over each fold. The data assigned to the current fold are classified while the remaining rows served as the training data. Then, the classifications were compared against the known states, compiled into a cross validation table, and assessed with accuracy statistics. we assessed the accuracy of the classifier with the positive predictive value, negative predictive value, sensitivity, and specificity.

Table B-1. Example detection histories with their derived consecutive record length and hit ratio predictor feature levels.

Detections in series originating at the present detection (T_0)							Consecutive Record Length	Hit Ratio
T_{-3}	T_{-2}	T_{-1}	T_0	T_1	T_2	T_3		
0	1	0	1	0	1	0	1	3/7
0	0	1	1	1	0	0	3	3/7

iii. Overlap Reduction

An algorithm inspired by nested-Russian Dolls was developed to reduce overlap and discretize positions in time and space within the telemetry network. If a fish can be placed at a receiver with a limited detection zone (stripped coaxial cables or dipole), then it can be removed from the overlapping detection zone (Yagi) if it is also recaptured there.

Fish will often visit a limited range antenna for a certain amount of time, then leave that detection zone only to return sometime later. This behavior is commonly referred to as a “bout” in the ecological literature (Sibly, Nott, and Fletcher, 1990). Kleinschmidt followed the method of Sibly, Nott and Fletcher (1990) to fit a three-process broken-stick model (piecewise-linear regression with two knots ($k = 2$)). We first calculated the lag between detections for each fish within each discrete detection zone. Then, we binned the lag time into 10-second intervals and counted the number of times a lag interval occurred within each bin. After log-transforming the counts, the three-process broken-stick model was fit using a brute-force procedure that tested every bout-length combination with an ordinary least squares regression. The best three-process model was the one that minimized the total residual error (sum of squares). The first bout process describes a continuous string of detections indicative of a fish being continuously present, the second bout process describes milling behavior at the edge of a detection zone where lags between detections may be 20 – 30 seconds or more, and the third bout process describes the lags between detections where a fish leaves one detection zone completely for another only to come back sometime later.

After deriving the bout criteria for each discrete telemetry location, presences were enumerated. We assumed that a fish left a detection zone at the start of the third process. Therefore, the second knot location in the piecewise linear process model (a.k.a. broken-stick model) described this lag-time. If the lag between detections is equal to or greater than this duration, a fish has left the telemetry location only to return much later. In other words, the fish experiences a new presence. We iterated over every detection, for every fish, at every receiver, applied this logic to each lag time, and then enumerated and described presences at each location with start and end time statistics.

After describing presences at each receiver (time of entrance, time of exit) it is possible to reduce the overlap between receivers that traditionally plague statistical assessments of movement. If we envision overlapping detection zones as a series of nested-Russian Dolls, we can develop a hierarchical data structure that describes these relationships. If a fish is present in a nested antenna while also present in the overlapping antenna, we can remove coincident detections in the overlapping antenna and reduce bias in our statistical models. This hierarchical data structure is known as a directed graph, where nodes are detection zones and the edges describe the hierarchical relationships among them. For this assessment, edges were directed from a larger detection zone towards a smaller. Edges identify the successive neighbors (smaller detection zones) of each parent node (larger detection zone).

2. LITERATURE CITED:

Beeman, J. W., & Perry, R. W. (2012). Telemetry Techniques: A User Guide for Fisheries Research. In N. Adams, J. Beeman, & J. Eiler (Eds.). American Fisheries Society.

Marsland, S. (2009). Machine Learning: An Algorithmic Perspective. Boca Raton, Florida: CRC Press, Taylor and Francis Group.

Richert, W., & Pedro-Coelho, L. (2013). Building Machine-Learning Systems with Python. Birmingham, UK: Packt Publishing.

Sibly, R., H. Nott. and D. Fletcher. (1990). Splitting behavior into bouts. Anim. Behavior 39: 63-69.

Stone, J. V. (2013). Bayes' Rule: A Tutorial Introduction to Bayesian Analysis. Lexington, KY: Sebtel Press.