# Analysis of Factors Affecting Post-release Mortality of Coho Salmon

Richard Hou[1], Andy Lin[2], Xiaotong Liu[3], Weihao Qiu[4]

## Abstract


- This statistical project aims to examine the factors that influence fish survival after a catch-and-release event and make some suggestions on fishing to minimize the mortality rate following the release.
- After statistical analysis, we identify eye injury, hook location, fin damage, and reflex score are the key factors in predicting fish survival after release, employing logistic regressions.


## Introduction

- In recreational fisheries, a significant proportion of fish are released following capture, known as catch-and-release fishing.
- This study aims to identify the factors that influence Coho salmon survival after a catch-and-release event in the ocean.
- Two statistical questions:
  - Assess logistic regression models, as well as alternative models and model selection approaches
  - Determine if condensing some measurements into scores is the optimal strategy

## Methods

- A tagging and tracking approach was employed to investigate the post-release mortality of Coho salmon.
- The cleaned dataset was divided into training data and testing data in a 3:1 ratio.
- Evaluated several logistic regressions with variable selection. All the models used Akaike Information Criterion (AIC).
- Applied alternative classification models and model selection methods such as k-fold cross-validation.
- Models and model selections were repeated to both condensed scores and separate variables, respectively.

## Results

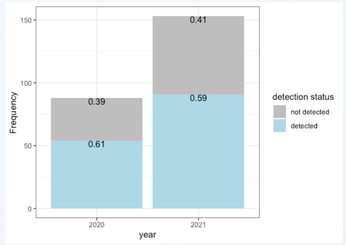


**Figure 1.** Frequency of each response variable level for each year

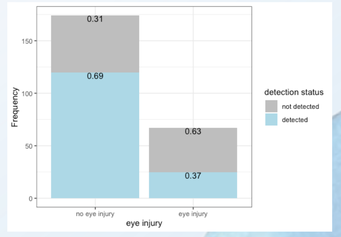


**Figure 2.** Frequency of each response variable level for whether the fish has eye injury

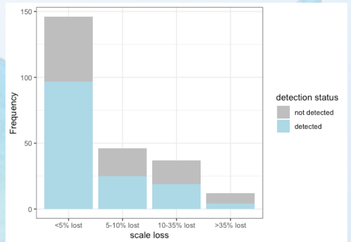


**Figure 3.** Frequency of each response variable level for different scale loss extent

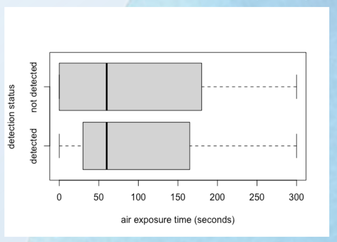


**Figure 4.** Parallel box plots for detection status versus air exposure time

- Figure 1-4: Some exploratory data analysis for categorical and continuous predictor variables

| Models | AIC | Validation Accuracy |
|---|---|---|
| Full Logistic | 227.49 | 0.639 |
| Stepwise Logistic | 213.02 | 0.650 |
| * Logistic with Significant Test | 220.64 | 0.655 |
| ** Logistic removing imbalance | 218.45 | 0.672 |
| *** Best Logistic | 217.99 | 0.707 |

**Table 1.** Summary of Some Representative Models Trained

- The best logistic regression suggests that the fish, which is more likely to survive, would have no eye injury, get hooked at the bottom jaw instead of the top jaw, have a minor fraying fin, and have a lower reflex score.
- Other classification models and model selection methods tested are not better

| Models | Variables Selected |
|---|---|
| * Logistic with Significant Test | Eye injury, hook location, fin damage, air exposure time, scale loss, reflex score |
| ** Logistic removing imbalance | Eye injury, hook location, fin damage, air exposure time, reflex score |
| *** Best Logistic | Eye injury, hook location, fin damage, reflex score |

**Table 2.** Interpretations of Representative Models stated before

- Condensed five categorical predictors into an injury score with test accuracy of 0.653
- Expanded the reflex score to five separate reflex variables with test accuracy of 0.708

## Conclusions

- The results suggest that eye injury, hook location, fin damage, and reflex score (or orientation reflex in detail) are likely to be the key factors.
- Logistic regression is appropriate to identify the significant factors and outperforms other models.
- Using model selection approaches such as AIC alone is unsuitable and probably dangerous. We suggest applying k-fold cross-validation.
- Condensing several measurements into scores (injury score and reflex score) might not be helpful in this study.

## Acknowledgements


- Our client, **Emma Cooke**, for giving us this opportunity to work with real data in a professional environment
- Our supervisors, **Melissa Lee, Nancy Heckman** and **Estella Qi**, for helping and guiding us through this project



[1]ruijieh@student.ubc.ca (UBC Dept. of Computer Science)
[2]andy.lin@ubc.ca (UBC Dept. of Statistics)
[3]xtlau@student.ubc.ca (UBC Dept. of Statistics)
[4]qiu0805@student.ubc.ca (UBC Dept. of Statistics)