Algonquin Turtle Life History Project - Painted Turtle Survival Analysis

Haoluan Chen

2/13/2021

Abstract

Introduction

Algonquin Park is a beautiful place that attracts many tourists to came and enjoyed these wonderful views. At the same time Algonquin park is home to many animal species. Every year starting from 1978 a team of dedicated researchers travel to the Wildlife Research Station in Algonquin Provincial Park, where they live from May to August. The team captured and recaptured the painted turtle then assigned unique notch identification and recorded their size and other statistics. Their goal is to use this long-term dataset to address interesting and timely questions, such as adaptation to northern environments, life-history trade-offs, phenotypic response to climate change, sexual selection, among other topics. However, for our study, we are investigating the relationship between the painted turtles survival rate and its size together with other behaviors.

In Rest of my report contains methods, which I first explain my data cleaning decision and model selection process in the Methods section. Then, I discuss the model result in the Results section, follow by model implication in the Discussion section. Finally in my conclusion, I recap our finding through our statistical analysis and concludes our result regarding to the painted turtles survival rate in the wild.

Methods

Data Cleaning dicision

Based on the information provided from our collaborator, that the sex and size variables are likely to have a impact on the survival rate of the painted turtle. Additionally, midpl was the most reliable measurement for size. Since all the size variable have high correlation, which provide simliar information for our model. I decided to only use midpl in our model to represent the size of painted turtle. After removing the unreliable and Irrelevant variables in the dataset and we left with the following remaining 4 variables.

Variables	Definition
notch	Unique individual ID
sex	Either UNKNOWN (immature individuals can't be sexed) or Male, or Female)
site	site that the turtle was captured
midpl	Plastron length (Measured from between the anterior-most plastral scutes to between the postior-most plastral scutes)

Furthermore, I noticed some input errors with the sex and mass. For Some observation, the sex variable is recorded as "-", so I changed these observations to unknown. For the mass variable, there was one observation recorded as -47. However, looking at other size variables for this observation, we can see that this is a mature turtle. So its mass must be larger than 47. Thus, I replaced the mass of this turtle with unknown. Lastly, there was 6 observation with missing sex variable. By looking at their size, they are all juveniles and I replaced them with UNKNOW indicating juveniles turtle.

Model Selection

Our dataset was obtained using capture and recapture methods. This method cause a high correlation in observations from the same turtle, which made me realized those common statistical models with independent assumption with independent assumption did not apply to our dataset. Additionally, using generalized mixed effect model did not coverage because there is only around 20 dead turtle observation in our dataset. However, we discovered Cormack-Jolly-Seber (CJS) model might be the most appropriate model to fit our dataset. The CSJ model is capable of estimating the survival and the capture probability of individuals and in a population and identifying time-varying factors that affected the survival rate and capture rate using the capture-recapture dataset. Based on those features of the CJS model is the best model in our case.

After comparing various models, I think the model appropriate model is to use midpl as time-varying variable to estimate survival probability and use sex as the static variable to estimate the capture probability. Since we believe that size have an effect on the survival probability and turtles with different sex behave differently in the wild, which have an impact on the capture probability.

Data Manpulation and imputation Process for CJS model

For CJS model, it is required to have observations for each turtle every year, but we only have observations from the turtles that is capture on a given year. For the time-varying covariate midpl, I first remove the turtles does not have any midpl value. Then, I imputed the observations before its first capture year with value 0 and I obtained an estimate that every year a turtle grows 0.15CM in midpl using linear regression to impute the years we did not capture the turtle after its first capture.

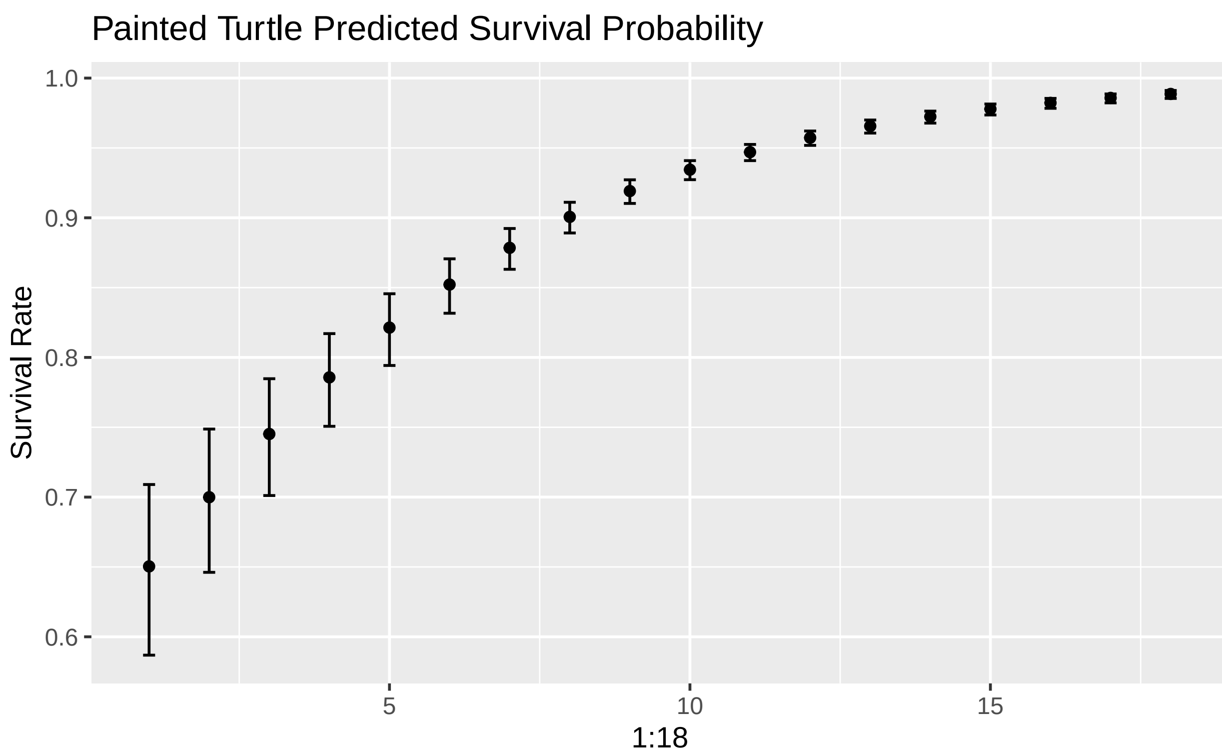
Finally, we generated a capture history for each turtle. Capture history is a string of zeros and ones, where zero indicates the turtle was not capture in that year and 1 indicates the turtle was capture in that year.

3. Result

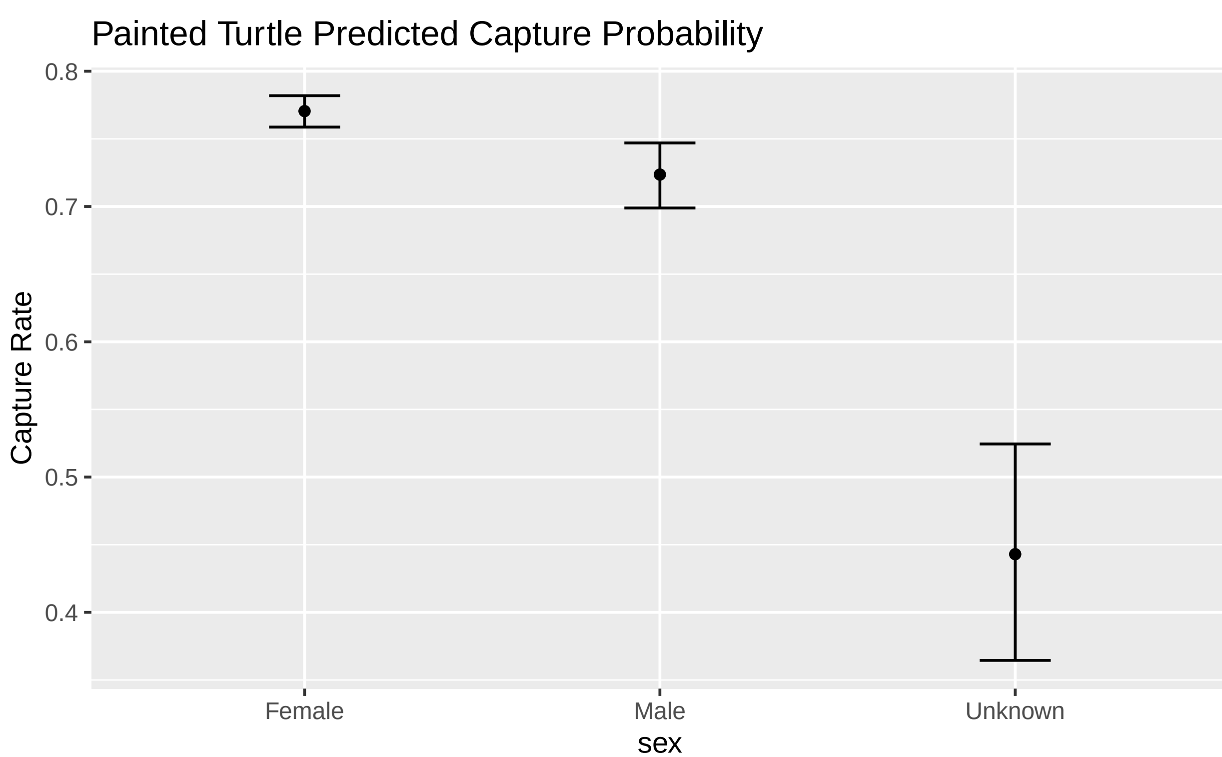
	Estimate	se	lcl	ucl
Phi.(Intercept)	0.395	0.151	0.099	0.690
Phi.midpl	0.226	0.014	0.199	0.254
p.(Intercept)	1.211	0.033	1.146	1.277
p.sexMale	-0.249	0.070	-0.386	-0.112
p.sexUnknown	-1.440	0.170	-1.774	-1.107

From the above table, we see that the odds ratio between the male turtle and female turtle survival rate is 0.499 with a 95% confidence interval of [0.366, 0.681]. This means the odds of survival rate for male turtle are about 50% lower than the odds of survival rate for female. Furthermore, the odds ratio between the Unknown sex turtle (juvenile) and the female turtle survival rate is 0.032 with a 95% confidence interval of [0.025, 0.042]. This means the odds of survival rate for Unknown sex turtles is 96.8% lower than the odds of survival rate for the female turtle.

Additionally, looking at the capture probability p, we see that male turtle has lower odds than female turtle, and unknown sex turtle has even lower odds ratio. The odds ratio of capture probability between male and female is 0.782 with a 95% confidence interval of [0.680, 0.9]. And the odds ratio of capture probability between unknown sex and female is 0.601 with a 95% confidence interval of [0.426, 0.849]



The odds of survival rate for female, male and unknown sex turtle are also reflected on the above plot. We predicted that female turtles have a high survival probability with a small 95% confidence interval. Male turtles have a slightly lower survival rate and slightly wider 95% confidence interval. Lastly, the unknown sex turtle has the lowest survival rate with only around 57%



Our model also predicted that the female turtle has the highest capture probability, around 77%. Male turtles have a slightly lower capture probability, around 72.5%. Lastly, the unknown sex turtle has the lowest estimated capture probability of 67%, with a very wide 95% confidence interval.

4. Disscussion

Result implication

From our CJS model result, we see that the female turtle has the highest survival probability. Male turtles have a slightly lower survival probability. Unsurprisingly, unknown sex turtles have a very low survival probability compared to male and female turtles. This is consistent with our knowledge of painted turtles since the female turtle is generally larger than males, so they are much less vulnerable. And the unknown sex turtle are mostly juvenile, which are very vulnerable.

We also observe the same trend for the capture probability that the female turtle has the highest capture probability, followed by male turtle then, unknown sex turtle. This means that we only capture about 77% of female turtles, 72.5% of male turtles, and only 67% of unknown sex turtle. This means we are likely to overestimate the survival probability because we do not have the data for the whole population. There are estimated 1/3 of juveniles are missing from our data, because they are likely to be dead and been eaten.

Model limitation

Our model has several limitations. First of all, there were many missing values in our dataset, and we have to remove some of the missing values and perform imputation. Removing and imputing the data could have a severe impact on the model result. Secondly, there was some input error in the data, which I chose to remove. This introduces more variation in our model result. Thirdly, some of our assumptions were not satisfied. For example, we cannot ensure that turtle emigrates from the study area is permanent and we are not releasing the turtles right away. Lastly, our model could not take advantage of all the size measurements and other variables in the dataset. Since adding a size measurement as a time-varying covariate leads to a non-significant result, which may because sex and size measurements are highly correlated. And I was not able to add a site random effect to our CJS model.

Conclusion

In conclusion, our model shows female turtles have a very high survival rate and capture rate. Male turtles have a slightly lower survival rate and capture compared to females. And the unknown sex (juveniles) turtle has a much lower survival rate probability and capture probability than female and male. However, there is some limitation of the model due to missing value, input error, imputation methods and the assumption that we may need further investigation.

Acknowledgments

References

Appendixes