Study on the Variation of Lamprey Population with Food Resources

Xianyu Qu*

Institute of Future Technology, Harbin Institute of Technology, Harbin, China, 150001

*Corresponding author: quxianyu2005@163.com

Abstract. Based on the sexually deformable characteristics of the lamprey, this paper combines the logistic model and Bayesian logistic regression model to study the changes in its population numbers with food resources. In this paper, it is assumed that the growth rate is proportional to the food intake before the food reaches saturation. First, based on the traditional logistic model, this paper introduces the sex-related growth index and uses the least square method to fit it. Second, to investigate the relationship between food intake and gender ratio, this paper uses a Bayesian logistic regression model. The methodology of this paper provides a new idea to study the population size of sex-variable organisms, which has some credibility and novelty. At the same time, it also helps a lot in controlling the number of lampreys, which is favorable to the development of fisheries.

Keywords: Logistic Model, Bayesian Logistic Regression Model, Lamprey.

1. Introduction

The lamprey is an invasive fish species that causes huge damage to the ecosystem. Lampreys parasitize other fish, which leads to a significant decline in other fish populations[1]. Therefore, studying changes in the population size of the lamprey is crucial for the development of fisheries. For the lamprey population, food availability is a major external factor affecting its population size[2]. To manage its population, it is important to understand the relationship between food resources and growth rate.

Malthusian growth model and logistic model have been widely used in predicting the growth of the population. Among them, the logistic model is a great improvement over the Malthusian growth model[3]. Logistic equations have a vast range of applications in various areas of natural and social sciences. These equations describe the sales volume of consumer durables, the number of trees in a forest, the number of fish in a pond, and so on. Under specific conditions, their variations usually follow a logistic law[4].

Although the logistic model is widely used, it has some limitations. For example, in the study of organisms, the sexes are immutable. Lampreys are sexually variable and traditional logistic model cannot predict their population. Thus, based on the logistic model, this paper introduced a sex-related growth index. To determine the correlation between nutrient intake and the sex ratio, the Bayesian logistic regression model was used.

2. The Population of Lamprey

Most of the data in this article comes from the Great Lakes Fishery Commission and the Sea Lamprey Research Program.

The key mathematical notations used in this paper are listed in Table 1. Symbols not labeled in the table are introduced in specific models.

Table 1: Notations used in this paper

Symbol	Definitions
$x(t)$	the population of lampreys at time t
E	Environmental capacity
r	Rate of natural increase
r_s	the sex-related growth index

2.1. The Relationship between Food and Population

In Model 1, two assumptions are made. First, it is assumed that food intake (i.e., the nutrients needed for lamprey growth) is positively correlated with the growth rate of the organisms until nutrient saturation occurs. Second, it is assumed that lamprey populations do not experience extreme external changes, such as natural disasters or invasions that result in large changes in population size.

It is assumed that water and food conditions are adequate in the lake or marine habitat, temperatures are suitable, and there is no competition within the population. It is known from the background that the proportion of males is about 56%, which means the proportion of males to females is about 1:1. We assume that the population size $x(t)$ is continuously differentiable and the rate of natural increase r is constant. Also, we assume that the change in population size is closed.

Thus, the population growth from t to $t + \Delta t$ can be $x(t + \Delta t)$ - $x(t)$ = $rx(t)\Delta t$. Then we can derive:

$$\begin{cases} \frac{dx}{dt} = rx \\ x(0) = x_0 \end{cases} \tag{1}$$

The solution is:

$$x(t) = x_0 e^{rt}. \tag{2}$$

However, resources in the watershed are not always sufficient. When the population of lam- prey increases by a certain amount, the rate of natural increase r is no longer a constant but a quantity that decreases as the population increases, so r can be expressed as a decreasing function with population $x(t)$ as the independent variable. The **Logistic** model is used.

Assuming $r(x) = r\text{-}sx$, the maximum population that can be accommodated by environmental resources is E, i.e., the growth rate $r(E) = 0$ when $x = E$. Based on the assumptions, it can be concluded that $r(x) = r\left(1 - \frac{x}{E}\right)$, thus:

$$\begin{cases} \frac{dx}{dt} = r\left(1 - \frac{x}{E}\right)x \\ x(t_0) = x_0 \end{cases} \tag{3}$$

The relationship between population density and time can be obtained by solving the differential equation as above.

$$x(t) = \frac{E}{1+\left(\frac{E}{x_0}\right)e^{-r(t-t_0)}} \tag{4}$$

The logistic model demonstrates that the rate of natural increase of lampreys reaches its maximum at $\frac{E}{2}$, accelerates before this point, and gradually decreases until it eventually reaches zero after this point. This model is an effective representation of the relationship between population size and environmental resources. However, since the availability of food also affects the sex ratio of lampreys, the assumption that the sex ratio is 1:1 is no longer valid. The proportion of males is about 78% when the food supply is low. Additionally, the sex ratio has an impact on the population size of the lamprey. Therefore, the logistic model alone cannot address the population size issue of the lamprey.

Based on the above analysis, the sex-related growth index r_s is introduced and this parameter can replace r in the logistic model.

2.1.1 Nonlinear Regression Fitting of rs

Based on the available data, we can calculate the size of the corresponding r using Equation 4. The relationship between the sex ratio and the total population is known, and we can establish the relationship between r and the sex ratio. The corresponding data are shown in Table 2.

Table 2: Available data about Males% and r_s

Males%	53	54	57	59	60	61	62
r_s	0.1758	0.2058	0.2228	0.3126	0.8841	0.8696	1.0921
Males%	63	65	68	72	73	75	78
r_s	1.404	0.8947	0.6584	0.4532	0.4153	0.1995	0.0921

Through information finding, the scatter plot conforms to GaussAmp:

$$y = y_0 + Ae^{-\frac{(x-x_c)^2}{2w^2}}$$

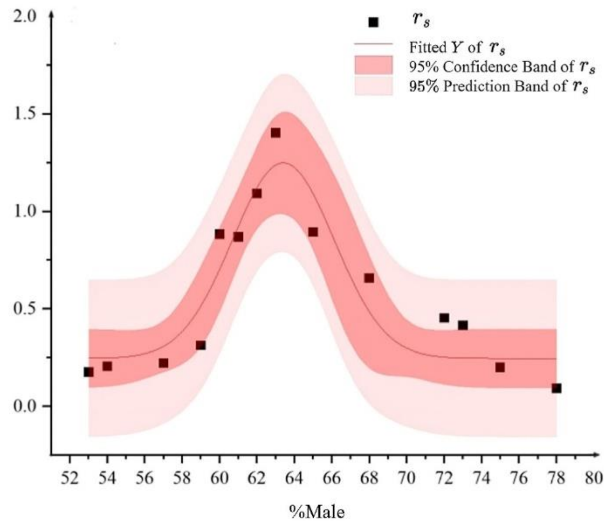

Figure 1: The fitting result of the curve

The values of the parameters were calculated by the least square method using *Python*. The results of fitting the sex ratio and the sex-related growth index r_s are shown in Figure 1. The values of each parameter after fitting are presented in Table 3.

Table 3: The fitting result

parameter	value
y_0	0.24383
X_c	63.38234
w	2.61527
A	1.0185

2.1.2. Check the Fitting Effect

Two methods were applied to analyze the fitting degree[5].

SSE and MSE

First, we calculated the Sum of Squares due to Error (SSE), which represents the sum of squares of the errors of the fitted data and the corresponding points of the original data. The closer the SSE is to 0, the better the model selection and fitting is.

We also calculated the Mean Squared Error (MSE), which represents the sum of squared errors at the corresponding points of the predicted and original data. *n* is the number of observations and *m* is the number of fitted parameters.

There is not much difference between MSE and SSE in the test of fitting effect.

Their corresponding formulas are shown in Eq.(5).

$$\mathrm{SSE} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{5}$$

$$\mathrm{MSE} = \mathrm{SSE}/(n-m) = \frac{1}{n-m}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

We computed the above data in *Python* and got the following results:

SSE = 0.2830517070533513

MSE = 0.023587642254445942

From the above data, it can be seen that the values of SSE and MSE are close to zero, which indicates that the model selection and fitting are good.

R^2

The second method of model testing is to calculate the coefficient of determination (R^2). To calculate the coefficient of determination, we need to first calculate the sum of squares of regression(SSR) and the total sum of squares(SST). The formula for SST is

$$\mathrm{SSE} = \sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{6}$$

which can be decomposed into:

$$\mathrm{SST} = \mathrm{SSE} + \mathrm{SSR}, \mathrm{SSR} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \tag{7}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Then let's go back to calculating R2.

$$\mathrm{R}^2 = \frac{\mathrm{SSR}}{\mathrm{SST}} = \frac{\mathrm{SST}-\mathrm{SSE}}{\mathrm{SST}} = 1 - \frac{\mathrm{SSE}}{\mathrm{SST}} \tag{8}$$

We computed the above data in *Python* and got the following results:

R^2 = 0.995169031304974

Based on the fact that the closer R^2 is to 1, the better the fit of the regression function is, it can be concluded that the fitting is good.

In conclusion, GaussAmp fits the data well with the parameter values.

2.2. The Relationship between the Growing Rate and the Sex Ratio

To explore the relationship between lampreys' growth rate and sex ratio, the second model was developed. In the preceding discussion, how the balance of genders within a population is linked to density has been explored. To understand how external factors like food availability impact this density, it's crucial to determine how food supply affects gender distribution.

We propose that there's a direct, positive link between food consumption (specifically, the nutrients essential for lamprey growth) and the rate at which these creatures grow, up to the point of nutrient sufficiency. Thus, examining how growth rates influence gender ratios can shed light on the food-gender relationship.

Previous research indicates that lampreys finalize their gender differentiation upon reaching a length of 90mm and the growing cycle of lamprey is shown in Figure 2. However, many still possess undifferentiated gonads at this stage, leading to a variable gender ratio and possible sex reversal [6]. This variability persists until the lamprey grows to about 140mm in length.

We draw upon a study by Johnson NS and colleagues [7], which employed radiolabeled tagging to monitor survival and transformation rates among lamprey populations near the Great Lakes' River mouths over time. From their findings, a mathematical formula linking growth rates to gender ratios was derived.

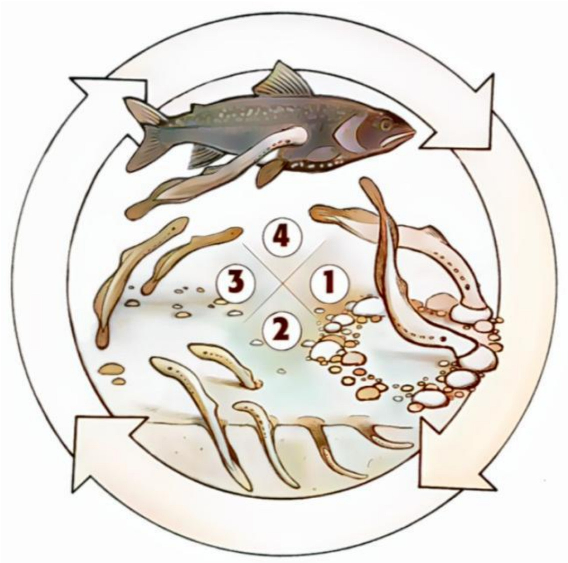

Figure 2: The growing cycle of lamprey

Simple regression models proved inadequate for accurately describing this relationship. Therefore, we applied a Bayesian logistic regression approach to predict the gender composition of adult lampreys in these waters[8]. The method involved treating the gender of each caught lamprey as a Bernoulli random variable, with the male probability modeled accordingly:

$$\text{logit}(p_m) = \alpha + \beta \cdot t_m \tag{9}$$

β denotes the slope and α denotes the intercept, Eq.(10) expresses the relationship between the probability of being male (on a logit scale) and t_m(year). α and β consist of two parts, type-specific population averages and deviations, where δ and μ denote deviations:

$$\alpha = \alpha_0 + \delta \tag{10}$$

$$\beta = \beta_0 + \mu$$

The following vague priors were performed: $\begin{pmatrix}\alpha \\ \beta\end{pmatrix} \sim MVN(0, \Sigma), {}^{\delta \sim N(0,\sigma_\delta)}_{\mu \sim N(0,\sigma_\mu)}$, $\Sigma \sim W_p\left(\begin{pmatrix}1 & 1 \\ 1 & 1\end{pmatrix}, 3\right), {}^{\sigma_\delta \sim Unif(0,100)}_{\sigma_\mu \sim Unif(0,100)}$.

The model was evaluated using R[9]. Point estimates and derived variables were obtained using the median of saved Markov Chain Monte Carlo (MCMC) chains[10]. We end up with the relationship between years to metamorphosis and %Males, as shown in Figure 3.

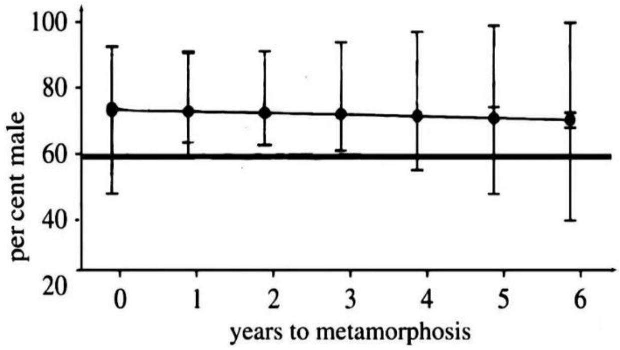

Figure 3: Bayesian logistic regression plot

3. Result

Through the above mathematical analysis, we get the relationship between sex ratio and population density, and the relationship between food and sex ratio. Thus we can conclude that the effect of the external environment, such as food, on population density.

From this, we can conclude that when there is not enough food, the proportion of males is large, the population number grows slowly and the population density is small. When there is sufficient food, the proportion of males is small, the population number grows moderately, and the population density is moderate. When food resources are intermediate, the proportion of males is moderate, the population grows fast and the population density is high.

This conclusion can be represented in Figure 4. Orange lines represent when food resources are moderate, green lines represent when food resources are adequate, and blue lines represent when food resources are inadequate.

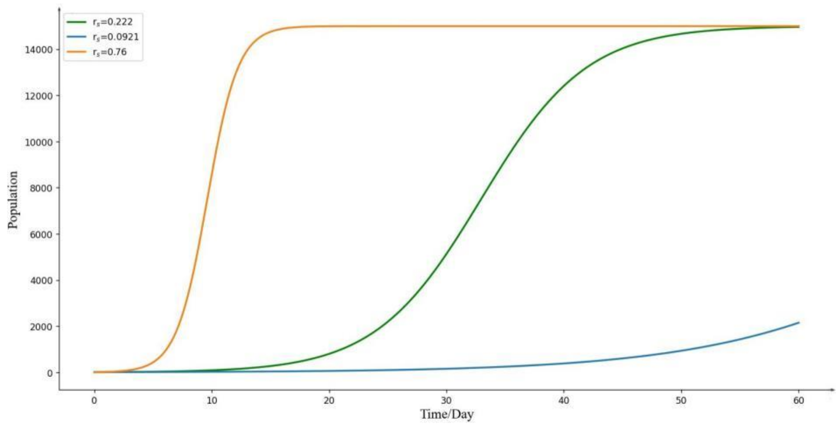

Figure 4: The population curve under different rs

4. Conclusion

In this paper, the logistic regression and the Bayesian logistic regression are used to study the population of lampreys under different food conditions. The methodology of this paper is novel and innovative compared to previous studies on populations of organisms with sex variability.

However, the method has some limitations. Firstly, the results obtained are very dependent on the available data. If the data are inaccurate or insufficient in number, then the credibility of the results is greatly reduced. Secondly, the paper hypothesized that food intake (i.e., nutrients required for lamprey growth) is positively correlated with the growth rate of the organism before nutrient saturation. This mathematical relationship does not always hold in nature.

If the above model is to be improved, the first thing that can be considered is to enhance the processing of the data to increase the credibility of the data. Secondly, the quantitative relationship between food intake and growth rate before food reaches saturation needs to be considered. Finally, external factors other than food that affect the growth of lampreys, such as light and temperature, should be investigated.

References

[1] Lucas M C, Hume J B, Almeida P R, et al. Emerging conservation initiatives for lampreys: Research challenges and opportunities[J]. Journal of Great Lakes Research, 2021, 47: S690- S703.

[2] Almeida P R, Arakawa H, Aronsuu K, et al. Lamprey fisheries: History, trends, and management[J]. Journal of Great Lakes Research, 2021, 47: S159-S185.

[3] ZHAO Tianwei,CHEN Huida. National population projection based on the linear regression model, Malthusian population growth model, and logistic model[J]. Journal of Guangdong Medical University,2023,41(06):623-627.

[4] JIANG Qiyuan, XIE Jinxing. The origin and early development of logistic (logistic) population modeling[J]. Mathematical Modeling and its Applications,2023,12(01):98- 102.

[5] Urang J G, Ebong E D, Akpan A E, et al. A new approach for porosity and permeability prediction from well logs using artificial neural network and curve fitting techniques: A case study of Niger Delta, Nigeria[J]. Journal of Applied Geophysics, 2020, 183: 104207.

[6] Wicks BJ, Barker LA, Morrison BJ, Beamish WH. Gonadal variation in Great Lakes sea lamprey, Petromyzon marinus, larvae. J. Great Lakes Res. 24, 962–968,1998.

[7] Johnson NS, Brenden TO, Swink WD, Lipps, 2016 Survival and metamorphosis of larval sea lamprey (Petromyzon marinus) residing near river mouths in lakes Michigan and Huron, J. Great Lakes Res. 42, 1461– 1469,2016.

[8] Nhu V H, Zandi D, Shahabi H, et al. Comparison of support vector machine, Bayesian logistic regression, and alternating decision tree algorithms for shallow landslide susceptibility mapping along a mountainous road west of Iran[J]. Applied Sciences, 2020, 10(15): 5047.

[9] Giorgi F M, Ceraolo C, Mercatelli D. The R language: an engine for bioinformatics and data science[J]. Life, 2022, 12(5): 648.

[10] Nemeth C, Fearnhead P. Stochastic gradient Markov chain Monte Carlo [J]. Journal of the American Statistical Association, 2021, 116(533): 433-450.