Population Prediction of Finless Porpoise in the Yangtze River

Zhaochen Zan [a], Hang Sun [b], Zhekai Hu [c]
Xi'an Jiaotong University, Xi'an 710049, China
[a] 3089861783@qq.com, [b] 2990227535@qq.com, [c] 798604047@qq.com
Zhaochen Zan and Hang Sun are co-first authors

Abstract. Finless porpoise is the only freshwater mammal in the Yangtze River basin, which has suffered a plummet in population these years. Our team mainly used Improved Leslie Population Model and Improved Differential Equation Model to predict the change of its population. Firstly, we predicted the population of finless porpoises 20 years later in 5 ex-situ reserves and analyse the influence caused by initial sex ratio. We used Leslie Population Model and improved it according to the characteristics of finless porpoises. We divided finless porpoises of different ages into five groups and used Siler's Competition Risk Model and PSO-MLE Algorithm to analyse their mortality. Using the models, we found that only when the initial sex ratio was 1:1 did the reproduction rate reach its maximum, and the population of finless porpoises could reach 1,336 after 20 years in this case. When the sex ratio was imbalanced, the greater the difference from 1:1, the fewer porpoises there were after 20 years. Lastly, we predicted how long it will take for the Yangtze finless porpoise to become functionally extinct without ex-situ conservation. We used Improved Differential Equation Model. According to the prevailing view of the criteria for functional extinction, it could be predicted that without protection, the finless porpoise will become functionally extinct in around 2038, from which we can see the importance of ex-situ conservation in finless porpoises' protection.

Keywords: Ex-situ conservation; Siler's Competition Risk Model; PSO-MLE Algorithm; Improved Differential Equation Model.

1. Introduction

Finless porpoise widely distributes in the middle and lower reaches of the Yangtze River, Poyang Lake and Dongting Lake, which is the only freshwater mammal in the Yangtze River basin. The average life span of the Yangtze finless porpoises is about 20 years. They reach sexual maturity and begin to mate at the age of five. A single female has one baby every two years on average, but due to abortion and other reasons, the chance of successful reproduction under natural conditions is just about 20 percent. In addition, the Yangtze finless porpoise is typically monogamous. Due to their rarity and unusual appearance, the finless porpoise is often called the "smiling angel" of the Yangtze River. Sadly, its population has plummeted in recent years, which decreased from 2,700 in 1991 to 1,012 in 2018. Ex-situ conservation, as one of the most important measures to protect the Yangtze finless porpoise, has been widely adopted in China since the 1980s and achieved outstanding results. At present, China has built five ex-situ conservation bases, namely Hubei Shishou Swan Lake Finless Porpoise Nature Reserve (1990), Anhui Anqing Xijiang Finless Porpoise Nature Reserve (2013), Hubei Jianli Hewangmiao Nature Reserve (1992), Anhui Tongling Freshwater Porpoise Reserve (2015), Institute of Hydrobiology of CAS (1996). The total number of ex-situ groups have reached 155 nowadays. It is worth mentioning that the number of finless porpoises in Swan Lake Reserve has increased from only 5 in 1992 to 101 in 2021.

2. Leslie Population Model Based on PSO-MLE Algorithm

2.1 The Establishment of Improved Leslie Population Model

Since the fertility rate and mortality rate of finless porpoises change with age, we plan to establish a Leslie population model to divide finless porpoises into various age groups to measure the difference of mortality rate in different age groups.

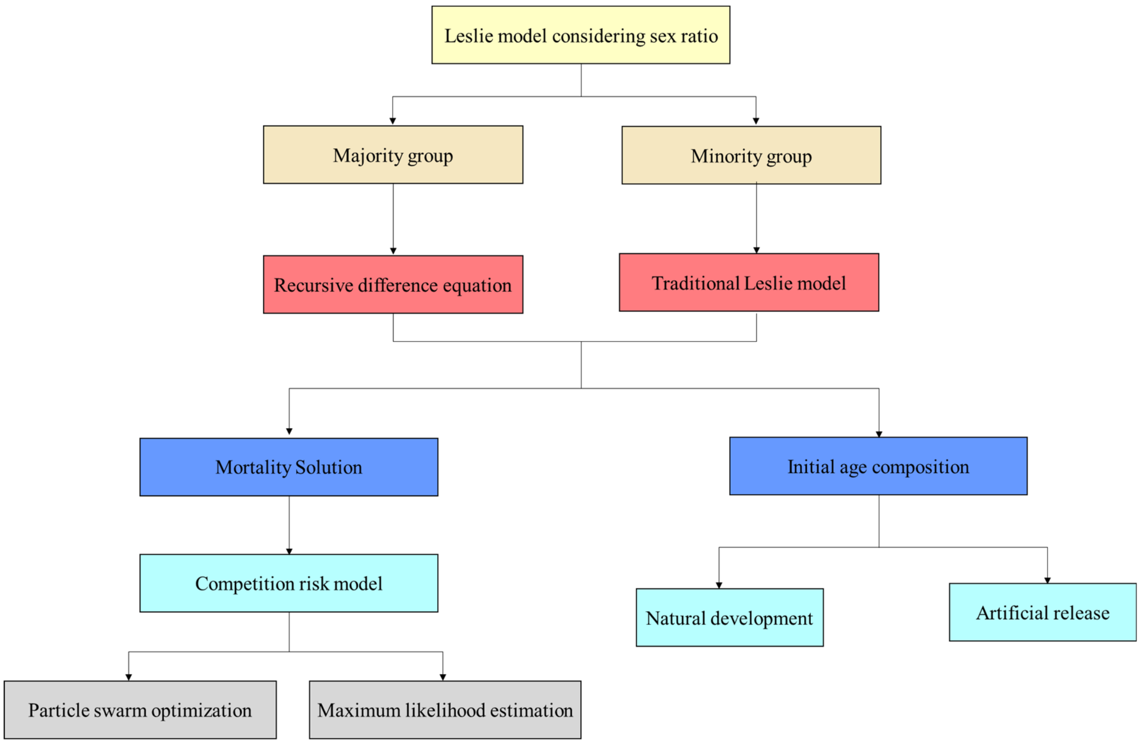

Figure 1. The flow chart

2.1.1 Mortality Solution Based on Competition Risk Model and PSO-MLE Algorithm

The improved Leslie model requires the mortality data of different age groups, while the known mortality information of finless porpoises is hard to find. Since the maximum likelihood estimation method does not need the prior probability information in the statistical inference, which overcomes the fatal weakness of The Bayes method, this paper plans to use this method to solve the problem of the lack of data. The likelihood function is obtained from Siler's competition risk model [1], because it has accurately described the survival or mortality of mammals.

In Siler's model, the survival rate $l(x)$ for a given age is divided into three parts:

$$l(x) = l_j(x) \times l_c(x) \times l_s(x) \tag{1}$$

In this equation, $l_j(x)$ is the exponential growth risk factor caused by youth:

$$l_j(x) = \exp\left(-\frac{a_1}{b_1} \times (1 - \exp(-b_1 x))\right) \tag{2}$$

$l_c(x)$ is the constant risk across all ages:

$$l_c(x) = \exp(-a_2 x) \tag{3}$$

$l_s(x)$ is the exponential growth risk factors due to old age:

$$l_s(x) = \exp\left(\frac{a_3}{b_3} \times (1 - \exp(b_3 x))\right) \tag{4}$$

a_1, a_2, a_3, b_1, b_3 are the five undetermined parameters, and the five parameters will be obtained by maximum likelihood estimation method. According to the number of age groups divided in this paper and related studies of Siler, the maximum likelihood function is given as follows(n is the number of age groups):

$$L(n) = \prod_{i=1}^{n} \frac{l(i)}{\sum_{y=1}^{n} l(y)} \tag{5}$$

In the traditional method of solving likelihood function, it is required to take logarithm of likelihood function and take partial derivative of each parameter respectively to solve the likelihood equation. But this method is very difficult to solve the analytical solution, which requires trial calculation in engineering, and the parameter estimation accuracy is not high, and even there may be no solution problem. From another perspective, the maximum likelihood estimation of probability distribution model is essentially a parameter optimization problem. Considering the good application of PSO algorithm in parameter optimization problems, it is used as a method to solve unknown parameters.

Particle swarm optimization (PSO) is a heuristic algorithm whose basic idea is inspired by the research results of modeling and simulation of bird group behavior. The core content is to make use of the information sharing of individuals in the group to make the movement of the whole group evolve from disorder to order in the problem solving space, so as to obtain the feasible solution of the problem.

For a particular bird, the position of step d is equal to the position of step d-1 plus the product of the speed of step d-1 and the time of movement:

$$x_d = x_{d-1} + v_{d-1} t \tag{6}$$

The speed of the bird at step d is equal to the speed of the bird at the previous step inertia add the self-cognition part and the social cognition part:

$$v_d = \omega v_{d-1} + c_1 r_1 (p_{bd} - x_d) + c_2 r_2 (g_{bd} - x_d) \tag{7}$$

In this equation, c_1 is the individual learning factor, c_2 is the social learning factor, r_1, r_2 are random numbers on [0,1], p_{bd} and g_{bd} are the historical best positions of itself and the group up to step d respectively.

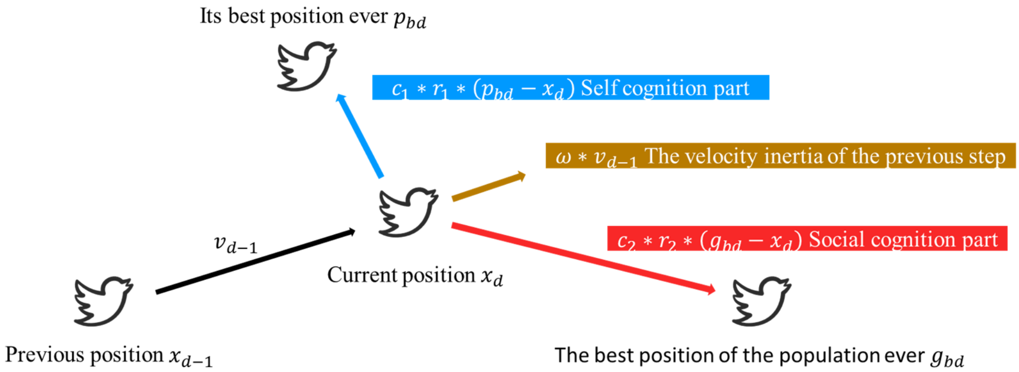

Figure 2. Schematic diagram of particle swarm algorithm

The process of a large number of solutions approaching to the optimal solution is analogous to the process of birds approaching to food. According to five parameters, the search dimension is defined as 5 dimensions, the likelihood function is taken as the fitness function, the initial particle number is 50, the maximum iteration book is 100, and the generation method of the initial solution is random generation. The adaptive dynamic change method [3] proposed by Li Li in 2008 was adopted to determine the change of the inertia weight coefficient according to the convergence degree of population prematurity and individual adaptation value.

Suppose the fitness value of particle p_i is f_i , the fitness value of optimal particle is f_m , the average fitness value of particle swarm $f_{avg} = \frac{1}{n}\sum_{i=1}^{n} f_i$ where the fitness value superior to f_{avg} is averaged to obtain f_{avg}' . Define $\Delta = |f_m - f_{avg}'|$. According to f_i, f_{avg}, f_{avg}' , the population is divided into 3 groups and different adaptive operations are carried out respectively. The adjustment strategies are as follows:

$$w = \begin{cases} w - (w - w_{\min}) \cdot \left| \dfrac{f_i - f_{avg}'}{f_m - f_{avg}'} \right| & f_i > f_{avg}' \\ w & f_{avg} \leq f_i < f_{avg}' \\ 1.5 - \dfrac{1}{1 + k_1 \cdot \exp(-k_2 \cdot \Delta)} & f_i < f_{avg} \end{cases} \tag{8}$$

The first type of particle is the better particle in the population, which is close to the global optimal, so we give it a smaller inertia weight coefficient to enhance its local search ability. The second kind of particle is the qualified particle in the group, both the global search ability and local search ability are qualified, so its inertia weight coefficient does not change. The third kind of particle is the unqualified particle in the population. In this paper, adaptive adjustment of genetic algorithm control parameters is adopted. k_1 and k_2 are the control parameters, k_1 is used to control the upper limit of inertia weight coefficient, and k_2 is used to control the adjustment ability of the above equation. At the end of the algorithm, if the particle distribution is dispersed, Δ will be large, and the above formula can reduce ω and improve the local search ability, so that the population tends to converge. If the particle distribution is concentrated, Δ will be small, and the above equation can increase ω , so as to enhance the particle's probing ability and effectively jump out of the local optimum.

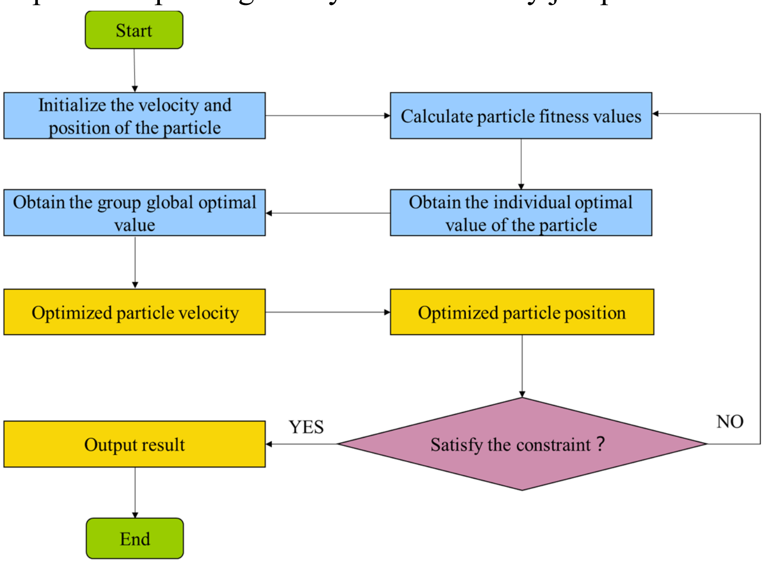

Figure 3. Flow chart of particle swarm optimization

After determining the change method of inertia weight coefficient, a reasonable parameter combination value can be obtained as follows:

$$[a_1, a_2, a_3, b_1, b_3]^T = [-0.0221, 0.0038, 0.0319, 0.2494, -0.5577]^T \tag{9}$$

2.1.2 Determination of Initial Age Composition

Since the initial numbers of finless porpoises in each ex-situ conservation bases are easy to obtain while the numbers of finless porpoises in an intermediate year are difficult, in order to determine the age composition of the finless porpoise population up to now, the Leslie population model and the difference equations established by it will be used.

In the Yangtze finless porpoise population, it is assumed that the maximum survival age is L, and $[0,L]$ is divided into n age groups, the length of each group is L/n. It is assumed that the fertility rate of the i^{th} age group is a_i, and the survival rate is b_i, a_i, b_i are constants which meet:

$$a_i \geq 0(i = 1,2,\dots,n),\ 0 < b_i \leq 1(i = 1,2,\dots,n-1) \tag{10}$$

We select the year 1992 as the starting point of timing, and the number of population at this moment can be obtained by using statistical data, that is, the initial number of finless porpoise in Tian-e-Zhou Conservation Reserve in 1992. The initial release of finless porpoises belongs to the first age group, and the number of the following age group is 0:

$$x^{(0)} = [x_1{}^{(0)}, 0, \dots, 0]^T \tag{11}$$

$$y^{(0)} = [y_1{}^{(0)}, 0, \dots, 0]^T \tag{12}$$

$$x^{(0)} = y^{(0)} \tag{13}$$

In the equation, $x^{(0)}$ and $y^{(0)}$ respectively represent the population quantity distribution vectors of female and male animals when t=0.

We select the interval of the age group L/n as time unit. At moment t_k, the number of females and males in i^{th} age group are respectively expressed as $x_i{}^{(k)}(i = 1, 2, \dots, n)$ and $y_i{}^{(k)}(i = 1, 2, \dots, n)$, the vectors of population quantity distribution in each age group of female population and male population can be respectively expressed as follows:

$$x^{(k)} = [x_1{}^{(k)}, x_2{}^{(k)}, \dots, x_n{}^{(k)}] \tag{14}$$

$$y^{(k)} = [y_1{}^{(k)}, y_2{}^{(k)}, \dots, y_n{}^{(k)}]^T \tag{15}$$

According to the discussion on sex ratio, we assume that the sex ratio is λ (male/female), then:

$$z^{(k)} = \begin{cases} x^{(k)} & \lambda > 1 \\ y^{(k)} & \lambda \leqslant 1 \end{cases} \tag{16}$$

Where $z^{(k)}$ donates the population quantity distribution of each age group of minority groups at moment t_k.

The number of females or males in the first age group of the population at moment t_k should be equal to half of the total number of all individuals born between t_{k-1} and t_k:

$$x_1{}^{(k)} = \frac{1}{2}[a_1 z_1{}^{(k-1)} + a_2 z_2{}^{(k-1)} + \dots + a_n z_n{}^{(k-1)}] \tag{17}$$

$$y_1{}^{(k)} = \frac{1}{2}[a_1 z_1{}^{(k-1)} + a_2 z_2{}^{(k-1)} + \dots + a_n z_n{}^{(k-1)}] \tag{18}$$

Meanwhile, the number of females or males in the $(i+1)^{th}$ age group at t_k should be equal to the number of females or males in the i^{th} age group at t_{k-1} multiplies the corresponding survival rate:

$$x_{i+1}{}^{(k)} = b_i x_i{}^{(k-1)}(i = 1, 2, \dots, n-1) \tag{19}$$

$$y_{i+1}{}^{(k)} = b_i y_i{}^{(k-1)}(i = 1, 2, \dots, n-1) \tag{20}$$

When $\lambda > 1$, the group of females is the minority group according to equation (16):

$$\begin{cases} x_1{}^{(k)} = \dfrac{1}{2}\left[a_1 x_1{}^{(k-1)} + a_2 x_2{}^{(k-1)} + \ldots + a_n x_n{}^{(k-1)}\right] \\ x_2{}^{(k)} = b_1 x_1{}^{(k-1)} \\ x_3{}^{(k)} = \qquad b_2 x_2{}^{(k-1)} \\ \vdots \\ x_n{}^{(k)} = \qquad\qquad b_{n-1} x_{n-1}{}^{(k-1)} \end{cases} \tag{21}$$

Mark the matrix as L:

$$L = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ b_1 & 0 & \cdots & 0 & 0 \\ 0 & b_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & b_{n-1} & 0 \end{bmatrix} \tag{22}$$

So the quation (21) can be expressed as:

$$x^{(k)} = Lx^{(k-1)}, k = 1, 2, 3, \ldots, \tag{23}$$

Where the matrix L is called Leslie Matrix.

According to equation (23) : $x^{(1)} = Lx^{(0)}, x^{(2)} = Lx^{(1)} = L^2 x^{(0)}, \ldots,$ normally:

$$x^{(k)} = Lx^{(k-1)} = L^k x^{(0)}, k = 1, 2, 3, \ldots, \tag{24}$$

According to equation (11)(12) of population quantity distribution at the initial time, the distribution vector of female population quantity at any time can be calculated.

Predictably, the majority group does not satisfy the Leslie matrix, because the number of the first age group of the majority group at moment t_k is determined by the number distribution of the minority group at moment t_{k-1} while other age groups still satisfy the recursive relationship between themselves. Due to the disunity of variables, matrix operation is not enough to solve the number of majority groups. However, when the number distribution of minority groups is known, we can still use the method of gradual iteration of difference equations to get the number distribution of majority groups at every moment:

$$\begin{cases} y_1{}^{(k)} = a_1 x_1{}^{(k-1)} + a_2 x_2{}^{(k-1)} + \ldots + a_n x_n{}^{(k-1)} \\ y_2{}^{(k)} = b_1 y_1{}^{(k-1)} \\ y_3{}^{(k)} = \qquad b_2 y_2{}^{(k-1)} \\ \vdots \\ y_n{}^{(k)} = \qquad\qquad b_{n-1} y_{n-1}{}^{(k-1)} \end{cases} \tag{25}$$

Thus, according to the population distribution at the initial moment, the age distribution of the population can be obtained up to now by calculating step by step according to the interval of age groups. Considering the fact that there is artificial release at some time, the number of female or male animals in the first age group in the population at moment t_k should be equal to half of the total number of all individuals born between t_{k-1} and t_k plus the number of artificial release at that time. So we have modified formula (17) (18) :

$$x_1{}^{(k)} = \frac{1}{2}\left[a_1 z_1{}^{(k-1)} + a_2 z_2{}^{(k-1)} + \ldots + a_n z_n{}^{(k-1)} + qn_0\right] \tag{26}$$

$$y_1{}^{(k)} = \frac{1}{2}\left[a_1 z_1{}^{(k-1)} + a_2 z_2{}^{(k-1)} + \ldots + a_n z_n{}^{(k-1)} + qn_0\right] \tag{27}$$

Where n_0 donates the number of artificial drops at that time and q is the determine factor.According to the information selected, In 1992, 5 finless porpoises were put into Tian-e-Zhou Reserve, 4 were put into Tongling Reserve in 2006, and 15 were put into Anqing Reserve in 2016. In 2021, 6 were put into He Wangmiao Conservation Area. So:

$$q = \begin{cases} 0 & t \neq 3,6,7 \\ 1 & t = 3 \ or \ 6 \ or \ 7 \end{cases} \tag{28}$$

$$n_0 = \begin{cases} 0 & t \neq 3,6,7 \\ 4 & t = 3 \\ 15 & t = 6 \\ 6 & t = 7 \end{cases} \tag{29}$$

According to the relevant life parameters of finless porpoise, $L = 20, n = 5$, and according to the initial data, the age composition vector of the current population is finally solved as follows:

$$S_0 = [0.347, 0.323, 0.199, 0.104, 0.027]^T \tag{30}$$

2.2 The Solution of Improved Leslie Population Model

After obtaining the current age composition structure, it is assumed that the initial population size is N_0, so:

$$X_0 = \frac{1}{1+\lambda} N_0 \tag{31}$$

$$Y_0 = \frac{\lambda}{1+\lambda} N_0 \tag{32}$$

Where X_0 and Y_0 donate the number of females and males at the initial time respectively.

According to the initial age component vector:

$$x_i^{(0)} = S_i^{(0)} X_0 \tag{33}$$

$$y_i^{(0)} = S_i^{(0)} Y_0 \tag{34}$$

Where $S_i^{(0)}$ represents the ratio of initial time in i^{th} age group.

So we know the number of male and female population distribution at the initial moment.Since at this time there is no artificial release, the minority group strictly meet the recursion relationship of Leslie matrix, and we know the initial number of population in 5 ex-situ reservations is $N_0 = 155$, so through equation (33)(34), we can obtain the number of male and female groups respectively at any time.Add them up, we can get the total number of the population.

When $\lambda \leqslant 1$, the analytical method is similar and will not be described here.

When the initial number and age composition are determined, a sex ratio will strictly correspond to a population number in 20 years. When programming, change the sex ratio from 0 to 10 with the step size of 0.1, and draw the curve of the population number in 20 years with the sex ratio as follows:

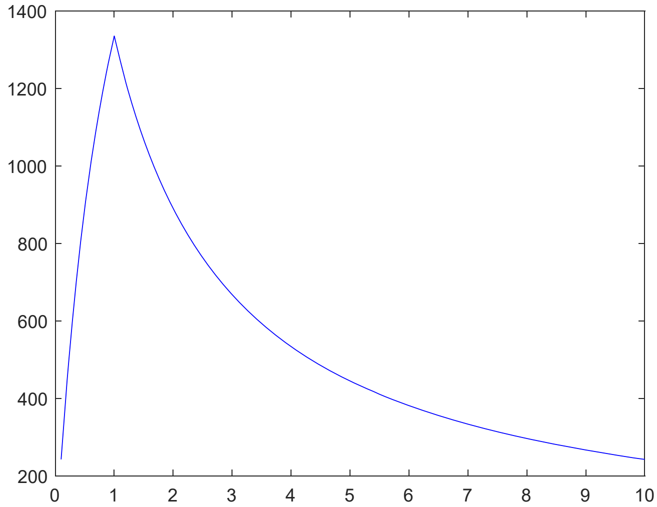

Figure 4. Population size change with sex ratio after 20 years

From Figure 4, we can draw the following conclusions:

1. Populations increase rapidly when the male-to-female sex ratio increases from 0 to 1
2. Populations decline slowly as the male-to-female sex ratio increases in the range of 1 to 10
3. The population size reach its maximum value when the sex ratio is 1, which is 1,336.

For more specific population changes, the initial sex ratio is controlled to remain unchanged, and time is taken as a variable to draw the population change curve under different sex ratios as shown in Figure 5:

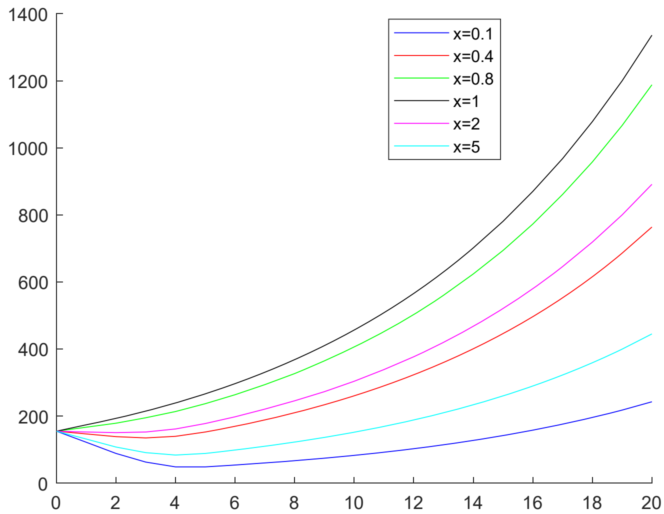

Figure 5. Population changes in 20 years with different sex ratios

3. Improved differential equation model

3.1 The Establishment of Improved Differential Equation Model

3.1.1 Traditional Logistic Model

First, we try to use Logistic model to predict population size. Logistic model is a relatively mature population prediction model, so we do not repeat its basic principles in this paper. We have established the equations as follows:

$$\begin{cases} \dfrac{dX(t)}{dt} = \gamma X(t) - \eta X^2(t) \\ X(t_0) = X(0) \end{cases} \tag{35}$$

Where, γ and η are called the life constant, $X(t_0)$ represents the initial value. By integrating both sides, we can get:

$$X(t) = \frac{\gamma X(t_0)}{\eta X(t_0) - (\gamma - \eta X(t_0) e^{-\gamma(t-t_0)})} \tag{36}$$

Put into three groups of data for fitting, and a reasonable solution can be obtained as γ=-0.1424, η=0.0000755,then we use MATLAB software to draw the curve as shown in the red curve in Figure 6.

Through solving, we find that the sum of squares of residuals is 246,990, indicating that there is a certain deviation from the actual situation, and the fitting effect is not very ideal. So we decided to find new ways to fit.

3.1.2 Improved Differential Equation Model

In any given period of time, the population changes such that the number of finless porpoises is equal to the number of births minus the number of deaths, without taking into account mass deaths caused by human factors and seasonal migration. The number of births is not only related to the breeding ability of the finless porpoise itself, but also related to the population size. The larger the population, the more births.

Suppose the birth rate of finless porpoise per unit time is α and the death rate of finless porpoise per unit time is β, easy to get

$$\frac{dX(t)}{dt} = \alpha X(t) - \beta X(t) \tag{37}$$

Suppose the number of finless porpoises is $X(t_0)$ in t_0. According to this initial condition, the population number of finless porpoise is:

$$X(t) = X(t_0) e^{(\alpha-\beta)(t-t_0)} \tag{38}$$

plug in the initial data t_0=1991, $X(t_0)$=2700, $\alpha - \beta = \xi$, we can get

$$X(t) = 2700 e^{\xi(t-1991)} \tag{39}$$

Put into three groups of data for fitting, we can get $\xi = -0.0363$,use MATLAB software to draw the curve as shown in the green curve in Figure 6. After solving, it is found that the sum of squares of residuals is 97,177, which is significantly lower than prior model, indicating that this model is more reliable. According to the curve, if no ex-situ conservation measures are taken, the population of finless porpoises will drop below 500 around 2038.

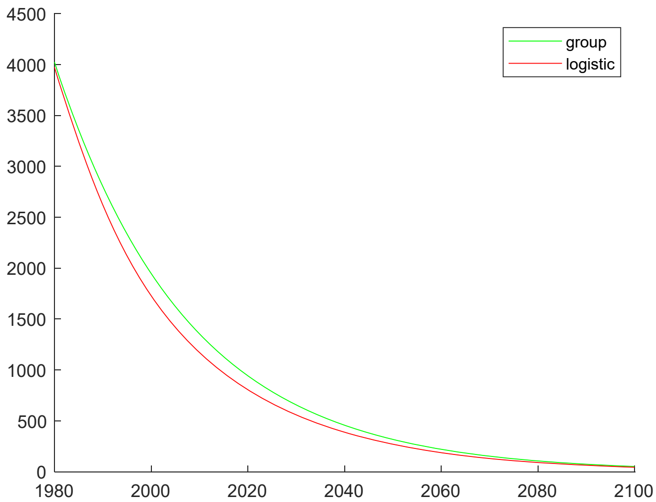

Figure 6. Improved model contrast with traditional logistic model

4. Conclusion

Using the models, we found the best initial sex ratio was 1:1, and the population of finless porpoises could reach 1,336 after 20 years in this case in 5 ex-situ reserves. When the sex ratio was imbalanced, the greater the difference from 1:1, the fewer porpoises there were after 20 years. According to the prevailing view of the criteria for functional extinction, it can be predicted that without protection, the finless porpoise will become functionally extinct around 2038.

In this paper, the mating system of the Yangtze finless porpoise population was considered, and the traditional Leslie model was improved to define the majority population and minority population by sex ratio, which solved the problem of inadequate applicability of the traditional Leslie model. Our model could also be used to predict the population development of other endangered animals in ex-situ reserves.

References

[1] William Siler. A Competing-Risk Model for Animal Mortality[J]. Ecology, 1979, 60(4) : 750-757.

[2] ZHANG X F. Study on age identification, growth and reproduction of finless porpoise [J]. Acta Hydrobiologica Sinica, 1992, 016(004):289-298.

[3] Li Li, Xue Bing, Niu Ben. Inertial Weight Adjustment Strategy of Particle Swarm Optimization algorithm [J]. 2008.

[4] Lei Qingwen, Yan Lei, Lu Dongyang, Bu Jiajun, Luo Yun. Maximum likelihood Estimation of P-III distribution based on Particle Swarm Optimization [J]. China's rural water conservancy and hydropower: 1-9 [2022-01-16].

[5] JIANG W H. Ex situ conservation of the Yangtze finless porpoise [J]. Journal of Anhui University: Natural Science Edition, 2010(4):5.

[6] Li Li, XUE Bing, Niu Ben. Inertial Weight Adjustment Strategy of Particle Swarm Optimization algorithm [J]. 2008.