Our Waterways-An Satisfying Future

参赛队员:田文波(98 机械)

陆峻凌(98环境)

方行玫(98通信)

指导教师:李传东

参赛单位:重庆大学

参赛时间: 2004年2月5~9日

Our Waterways-An Satisfying Future

Wenbo Tian, Junlin Lu, Xingmei Fang Advisor: Chuandong Li

Summary

Zebra mussels are mollusks that have a significant negative influence on some lake's ecosystem and economy because of their colony on various surfaces and high spread. Our aim in this article is to find effective measures that limit the growth of zebra mussels.

In part I, some factors are listed that are possibly related with mussel growth. In part II, we deal with the effect of more than ten environmental factors on mussel growth based on the statistic data taken from lake A. For that purpose, two models are proposed, one of which is based on linear regression, and the other is a grey system model. By computer simulations, we conclude that Cl and K have a desirable effect on zebra mussels, and Ca promotes their growth. As a result, an effective way to decrease the growth of mussels is to increase the amount of Cl and K in a polluted lake and/or, decrease that of Ca.

Part III is focused on the modification of the first one presented in part II by use of the data drawn in lake B. First, a new model is put forward in which the time factor is taken into account. Simulated results show that (1) this model has a satisfactory performance in terms of its prediction capability and its capability of clearly depicting the factors that have significant effect on the spread of zebra mussels, (2) the model is robust (that is, the model is insensitive to the parameters), and (3) lake A is vulnerable to the spread of zebra mussels. Then a few more complicated, but maybe nearer the reality, nonlinear models are listed that are our further work.

In part IV, the data from lakes B and C are compared with those in lake A. It results from the environmental similarity between lake A and B that lake B is much likely to be vulnerable.

In part V, we discuss the relation between the de-icing of roadways near the lake and the growth of mussels. Based on it, we propose a suggestion to the local government that is effective to limit the growth of zebra mussels.

In part VI, a strategy is presented whose idea is to utilize round goby fish, a natural enemy of zebra mussels, to limit or even eliminate zebra mussels. A few valuable models are listed that describe the relation between round gobies and mussels under different environmental conditions. Our further work is to deeply explore the applicability of these models in real environment.

Assumptions and Hypotheses

- The density of Zebra mussels at each site is uniformly distributed;
- Zebra mussel veligers (larvae) are innocuous to the great lakes ecosystem and economy;
- The whole life span of Zebra mussels is six years;
- Other environmental factors that are not mentioned in the problem have negligible effect on the population of Zebra mussels;
- The whole system is close, and the outsiders have negligible effect on the population of Zebra mussels:

Part I: Preliminary Discussions (A Solution to Requirement A)

Since the introduction in the mid 1980s, zebra mussels have spread through all of the Great Lakes and to an increasing number of inland waterways in the United States and Canada. Because Zebra mussels have significantly impacted the Great Lakes ecosystem and economy, many communities are trying to control or eliminate these aquatic pests. Our task is to identify the environmental variables related to the zebra mussel infestation in North American waterways.

The published articles and books on zebra mussels have discussed some factors possibly influencing the spread of zebra mussels. Some of them are listed below:

- □ chemical ingredients□ temperature
- □ their predators
- \Box time and seasons
- □ water quality
- □ current velocity
- particular organic carbon concentration
- □ food quality
- □ surfaces

Different factors may have different effect on zebra mussels. In this article we are focused on identifying how the spread of zebra mussels is influenced by various chemical ingredients by use of the statistic data given in the materials. Our aim is to find out those factors that have significant effect on zebra mussel and, based on it, put forward effective measures of controlling these pests.

Part II: Two Models (A Solution to Requirement B)

In this part, two mathematical models are proposed about the way by which some chemical ingredients are influencing mussel population. The former is based on regression analysis, and the latter utilizes grey system theory and methods. The foundation on which our discussion is based is the statistic data provided. Below are the two models:

Model 1: Linear Regression Model (LRM, for short)

We assume that there is a linear correlation between the mussel population and the chemical ingredients. The model is shown below:

$$\begin{aligned} Y_{n \times 1} &= \sum_{i=1}^{m} a_{i} X_{i} = X_{n \times m} A_{m \times 1} \\ Y_{n \times 1} &= X_{n \times m} A_{m \times 1} & (m = 11, n = 10) \\ \text{Define:} & Y_{n \times 1} &= (y_{1}, y_{2}, \dots, y_{n})^{T}; \\ X_{n \times m} &= (x_{ik})_{n \times m}; \\ A_{m \times 1} &= (a_{1}, a_{2}, \dots, a_{n})^{T} \end{aligned}$$

where Y_i represents the population's average density at the i^{th} site whose QA is S. X_{ij} denotes the mean value of the j^{th} chemical at the i^{th} sampling site, m denotes the number of sampling sites. n denotes the number of factors that are considered. In this place, each value of j denotes Alk, Ca, Chl-a, Cl, Fe, K, Mg, Na, pH, Secchi and Temp, respectively. We take the mean value of each quantity over the sampling data as an estimation of the expectation of the quantity.

Let x denote the value of a specified factor at a site. Then x is a random variable. Let F(x) denote the distribution function of x. For brevity, a few abbreviations are introduced below:

MinV: Minimum Value 25thP: 25th Percencile PMed: Population Median 75thP: 75th Percencile MaxV: Maximum Value.

Figure 1 illustrates the approximate estimation of F(x) based on the statistical data provided.

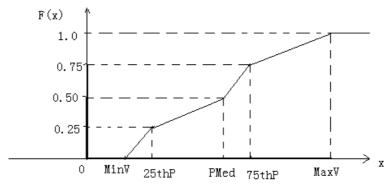


Figure 1: an estimation of the distribution function F(x)

$$E(x) = \int_{MinV}^{MaxV} x dF(x)$$

Below we present an algorithm for estimating the value of X_{ij} :

$$X_{ij1} = (MinV + 25thP)/2$$

$$X_{ij2} = (25thP + PMed)/2$$

$$X_{ij3} = (PMed + 75thP)/2$$

$$X_{ij4} = (75thP + MaxV)/2$$

$$X_{ij} = (X_{ij1} + X_{ij2} + X_{ij3} + X_{ij4})/4$$

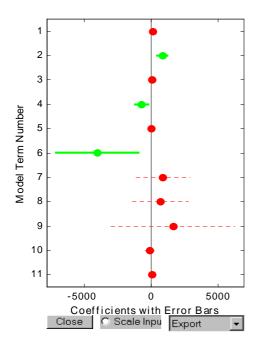
Taking the stepwise linear regression analysis as our basic method and through computer simulations with Matlab, we get the result shown in Figure 2, where the three green nodes (corresponding to the three lines of data whose column numbers are 2, 4, and 6, respectively) denote the factors that have strong pertinence with the mussel's population (that is, they influence the population most greatly).

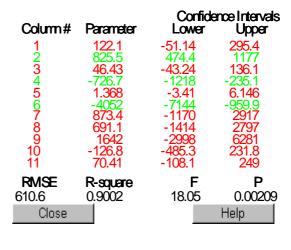
The model with the stepwise linear regression is shown below:

$$Y = 825.5X_2 - 762.7X_4 - 4052X_6 + 593.2$$

Through the above discussions, we conclude that:

- 1) Ca (denoted by column #2) is positively related to the population because its coefficient (825.5) is significantly greater than 0;
- 2) Cl (denoted by column #4) is negatively related to the population because its coefficient (-726.7) is significantly less than 0;
- 3) K (denoted by column #6) is strongly negatively related to the population because its coefficient (-4052) is significantly less than 0.
 - 4) The coefficients of all the remained factors, in its absolute value, are quite small. Consequently, we get the following conclusions:
 - □ Ca makes a positive attribution to the mussel population, and
 - □ Cl and K both limit the mussel population.
 - ☐ The effect of all the remained factors on the population is negligible





Page 4 of 16

Figure 2: the simulation results obtained by use of stepwise regression analysis, where **R** denotes the relational coefficient, **F** denotes F-test, and **P** denotes a probability related to F-test.

Model II: Grey System Model (GSM, for short)

Now we introduce another model known as Grey System Model (GSM). A system with definite (indefinite) information is known as WHITE (BLACK) system. In contrast, a system with definite partially information and indefinite partially information is known as GREY system. We assume that $X_0 = (x_{10}, x_{20}, x_{n0})^T$ is mother serial and

$$X_{1} = (x_{11}, x_{21},, x_{n1})^{T}$$

$$X_{2} = (x_{12}, x_{22},, x_{n2})^{T}$$

$$X_{3} = (x_{13}, x_{23},, x_{n3})^{T}$$
...
$$X_{m} = (x_{1m}, x_{2m},, x_{nm})^{T}$$

are son serial. Define the relational coefficient between Z_i and Z_0 at the K^{th} point as follows:

$$y_{i}(k) = \frac{a + \rho b}{\Delta_{i}(k) + \rho b};$$
where
$$\Delta_{i}(k) = \left| x_{ki} - x_{0i} \right|, i = 1, 2, ..., m; k = 1, 2, ..., n$$

$$a = \min_{1 \le k \le n} \min_{1 \le i \le m} \left\{ \Delta_{i}(k) \right\}$$

$$b = \max_{1 \le k \le n} \max_{1 \le i \le m} \left\{ \Delta_{i}(k) \right\}$$

 ρ denotes the discrimination coefficient $\rho \in (0,1)$,

The relational degree between Z_i and Z_0 is :

$$r_i = \frac{1}{n} \sum_{1}^{n} y_i(k), i = 1, 2, ..., m.$$

Below we propose an algorithm for computing the value of r_i :

Step 1: Get the original data (Select mother index). Here we select the densities of mussels on the plates at the 10 sampling sites, respectively.

Step 2: process the obtained original data in this way: First compute the average value over all original data for each index. Then divide each value of the index by it. The obtained value is denoted by x_{ki} .

Step 3: Compute the relational coefficient

$$y_i(k) = \frac{a + \rho b}{\Delta_i(k) + \rho b} ;$$

where
$$\Delta_i(k) = |x_{ki} - x_{0i}|, i = 1, 2, ..., m; k = 1, 2, ..., n$$

$$a = \min_{1 \le k \le n} \min_{1 \le i \le m} \{\Delta_i(k)\}$$

$$b = \max_{1 \le k \le n} \max_{1 \le i \le m} \{\Delta_i(k)\}$$

$$\rho = 0.5$$

Step 4: calculate relational degree

$$r_j = \frac{1}{n} \sum_{1}^{n} y_j(k), j = 1, 2, ..., m.$$

Step 5: compute the weight r_i of each index r_i :

$$r_{j} = r_{j} / (r_{1} + r_{2} + + r_{m}), j = 1,2,...,m.$$

 r_i denotes the significance the effect of the j^{th} factor on the population.

Step 6: construct the comprehensive evaluation model

$$Z_{k} = \sum_{j=1}^{m} r_{j} \times x_{kj}$$

Step 7: Sort the factors according to the decreasing order of the values of r_j .

By running the above algorithm on the data provided in Requirement B and analyzing the obtained results, we get the following conclusions:

- □ Ca makes a positive attribution to the mussel population, and
- □ Cl and K both limit the mussel population.
- ☐ The effect of all the remained factors on the population is negligible

The conclusion obtained by GSM coincides with that obtained by LRM with a high degree and, consequently, corroborate the reasonableness of LRM and the outcome to some extent.

Part III: Reasonableness of LRM (Requirement C)

1. Corroboration of the Reasonableness

Requirement B's conclusion tells us that Ca, Cl, and K have a much stronger effect on the spread of the mussels than all the other elements sampled. So we can ignore the other factors and only need to test the reasonableness of these three elements if necessary. By use of the additional data on Lake A, we find that among the three elements only Ca is sampled again. Thus we could carry out the following steps to corroborate the reasonableness of LRM:

- 1) Compute the average concentration of Ca at each site basing on the additional data in lake A from another scientist;
 - 2) Replace the average concentration of Ca in LRM by the result of (1);
- 3) Run LRM again on the data obtained in (2) and get a new result (shown in Figure 3).

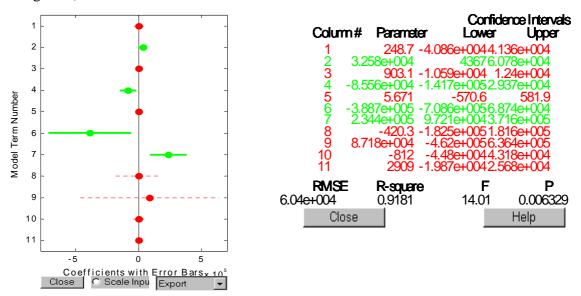


Figure 3: the simulation results obtained by use of stepwise regression analysis, where **R** denotes the relational coefficient, **F** denotes F-test, and **P** denotes a probability related to F-test.

As in part III, we again find that K, Ca, and Cl all have significant effect on Y, which corroborate the reasonableness of LRM. Besides, we also find that the impact of Na(Column#7=Na) is also evident, so Na should be regarded as another factor that has significant impact on Y and, like Ca, the impact is positive.

2. Adjustment

In LRM, we only considered how the chemical ingredients impact Y. In fact, Y depends not only on place but on time (see Figure 4).

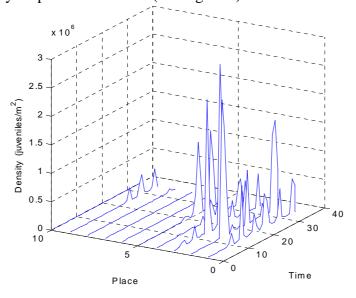


Figure 4: The density depends on both place and time

Thus we adjust our LRM to get a new model by taking time factor into account in the following way:

Model III: Time-Variable Linear Regression Model (TVLRM, for short)

$$Y_{i} = a_{1}T + \sum_{j=2}^{8} a_{j} X_{ij}$$

$$\vec{Y}_{10\times 1} = \vec{X}_{10\times 8} \vec{A}_{8\times 1}$$

where X_{ij} (or T) = Sampling Year-1985

(Sampling Year denotes the year when the sample is taken)

 X_{i2} denotes the average density of TP sampled at location *i* in the same year;

 X_{i3} denotes the average density of DP sampled at location i in the same year;

 X_{i4} denotes the average density of Ca sampled at location i in the same year;

 X_{i5} denotes the average density of Mg sampled at location i in the same year;

 X_{i6} denotes the average density of TN sampled at location i in the same year;

 X_{i7} denotes the average density of Temp sampled at location i in the same year;

 X_{i8} denotes the average density of Chl-a sampled at location i in the same year;

 Y_i denotes the maximum population of one year in each location.

Stepwise linear regression analysis would again lead us to an analogous result as before shown as Figure 5, from which it is easy to deduce the following outcome:

- ☐ Year(Column#=1) has a positive effect on the population;
- □ TP(Column#=2) has a negative effect on the population;
- □ Ca(Column#=4) has a positive effect on the population;
- □ TN(Column#=6) has a positive effect on the population;

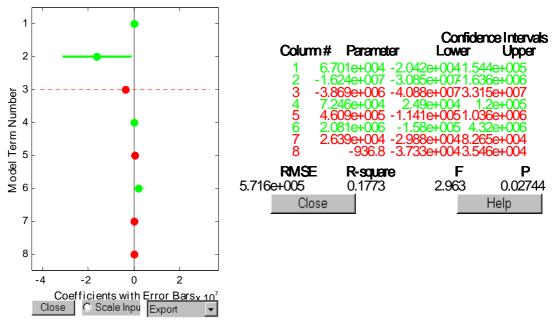


Figure 5: the simulation results obtained by use of stepwise regression analysis, where **R** denotes the relational coefficient, **F** denotes F-test, and **P** denotes a probability related to F-test.

In reality, the effect of some factors on Y is not linear. Hence a few nonlinear models are introduced below that are valuable in some cases:

1) Malthus model: Let N(t) be the number of the single species at time t. Let r be the increasing rate. The model is as follows:

$$\begin{cases} \frac{dN}{dt} = rN(t) \\ N(t) \Big|_{t=t_0} = N_0 \end{cases}$$
 (1)

The solution obtained by the model is: $N(t) = N_0 e^{r(t-t_0)}$.

2) Logistic model: In reality, however, the effect of its density should be considered. Hence we take the following modified model:

$$\frac{1}{N(t)}\frac{dN(t)}{dt} = r - f(N) \tag{2}$$

or

$$\frac{dN(t)}{dt} = NF(N) \tag{3}$$

where f(N) and F(N) are two functions of N(t).

3) Time-Variable Single-Population model: Let K be load volume. K is a function of time that deviates with seasonal variety of time in the following way:

$$K(t) = K_0 + K_1 \cos(\frac{2 \times \pi}{\tau} t) \tag{4}$$

which can be turned to

$$\frac{dN(t)}{dt} = rN(t)\left[1 - \frac{N(t)}{K(t)}\right] \tag{5}$$

The solution to the equation is:

$$N(t) = \left\{r \int_0^t \left[\frac{1}{K(s)} \exp[r(t-s)]ds\right]\right\}^{-1}$$
 (6)

4) Reaction-Spreading model: Logistic model is suitable for the case when the population distributes uniformly in the space. In the opposite case, the population would diffuse from the high-density area to the low-density area. Assuming that the diffusion is isotropic and adding diffusion factors to the Logistic model, we get the following model:

Let U = N/k, s = r, the Logistic model is changed to:

$$\frac{dU}{dt} = SU(1 - U) \tag{7}$$

After adding the diffusion factors to the equation (7), we get a new one:

$$U_t = U_x + SU(1 - U) \tag{8}$$

Then the density is the function of not only time but the place. The equation (8) is called "Fisher equation", in which U represents population density, t as time and x as the space coordinate.

When the distribution density is not even, the equation should be:

$$U_t = U_{xx} + f(U) \tag{9}$$

The equation (9) is called "Reaction-Spreading model". It is a superior model involving not only the time but also the space. The function may be roughly estimated by use of the statistic data provided.

The last model reveals the inner mechanism in theory why zebra mussels are spreading.

3. Sensitivity analysis

For a model incorporating many parameters, it is essential to determine which introduce the greatest error. Given a $\pm 10\%$ deviation in the value of the parameter, we calculate the percentage change in the value that the final system converges to. Table I summarizes the parameters that have less significant effects; This model is fairly insensitive to the values of other parameters.

Factor	%Diff. after (+10%)	%Diff. After (-10%)	
TP	-1.3%	+1.3%	
Ca	+3.6%	-6.2%	
TN	+2.4%	-5.0%	

Table I: Sensitivity of the model to changes in parameters.

Part IV: Comparison and Prediction (Requirement D)

If, in the function obtained by use of TVLRM, the coefficient of the time variable T is positive and its absolute value is significantly big, then the lake is vulnerable to the spread of zebra mussels because, in that case, the amount of mussels would grow rapidly. On the contrary, the lake is invulnerable. Consequently, lake A is vulnerable to the spread of zebra mussels because b, the coefficient of T, is positively big. Now we attempt to determine if lake B and C are vulnerable to the spread of zebra mussels on the basis of environment's comparability.

We listed the average concentrations of common components in lakes A, B and C in table 2, respectively:

Test Lake	PH (PH)	Ca (mg/L)	Mg (mg/L)	Na (mg/L)
Lake A	7.817	17.47	4.21	8.34
Lake B	7.63	11.50	2.36	7.60
Lake C	4.74	1.15	0.23	0.38

Table 2: Average concentrations of common components in lakes A, B and C

The concentrations of common components in lake A are close to those in lake B. But it is not sufficient for us to conclude that lake B is also vulnerable because the average value is not enough to describe the distribution. In contrast, lake A is quite different from lake C in terms of the concentrations of common components. Now we take all the data into account by drawing all the sampling data on the two-dimensional plane of Ca-density and PH-value (shown in Figure 6).

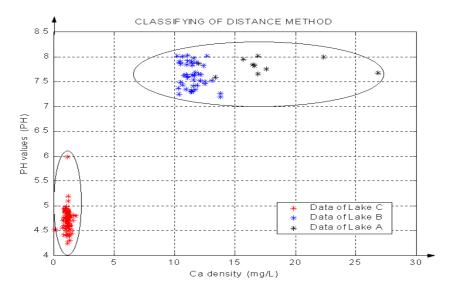


Figure 6: the distribution of the sampling data on the two-dimensional plane of Ca-density and PH-value

On one hand, it is obvious from Figure 6 that the environment of lake B is very similar to that of lake A, while for lake C the case is the opposite. Consequently, we predict that lake B is vulnerable to the spread of zebra mussels. On the other hand, no obvious conclusion can be obtained.

Part V: (Requirement E)

Zebra mussels colonize on various surfaces such as ice. In winter, if the roadways near the lake were covered by ice and no policies are carried out for the de-icing of roadways, zebra mussels would more easily find surfaces they need to colonize. Hence de-icing means destroying the environment in which mussels can easily survive to some extent.

We suggest that the local government plan a timetable of de-icing roadways in order to prevent mussels from quick growth. We predict that, if our suggestion were adopted and carried out for a period of time, the net growth rate of zebra mussels would significantly decrease. Even more optimistically, the absolute amount of mussels would possibly decrease.

Part VI: An Effective Way to Reduce the Growth of the Mussels (Requirement F)(Lotka-Volterra Model)

A model, known as Lotka–Volterra model [William F. Lucas, P116], depicts a way by which a pest is eliminated by one of its natural enemy in a mathematical mode. Round goby fish is one of zebra mussels' predators. Now we use Lotka–Volterra model to simulate the way by which the introduction of enough round gobies reduce the growth of the mussels within and among waterways. Let x(t) and y(t) denote the density of zebra mussels and that of their predators at time t, respectively. Lotka–Volterra model tells us that x(t) and y(t) vary with time in the following way:

$$\begin{cases} \frac{dx}{dt} = x(r_1 - \alpha y) \\ \frac{dy}{dt} = y(-r_2 + \beta x) \end{cases}$$

when the round gobies are not enough to have a significant effect on the spread of zebra mussels, an effective alternative is to raise the gobies artificially and pour them into the lakes where zebra mussels prevail at fixed time periods. In that case, another model, known as modified Lotka–Volterra model is suitable and is listed below:

$$\begin{cases} \frac{dx}{dt} = x(r_1 - \alpha y) \\ \frac{dy}{dt} = y(-r_2 + \beta x) + v \end{cases}$$

where $v \ge 0$ denotes the introducing rate of round goby fish. Since Lotka–Volterra model and modified Lotka–Volterra model has long been used to deal with the problem regarding to population mutual action and proved correct and effective, we believe that it suits for the present case. Consequently, introducing gobies would be a good way to reduce the mussel's population in a reasonable amount.

On the other hand, from the discussions of Part I --- Part IV, we deduce that some chemical components have a negative effect on mussel's population and, hence, can be used as 'pesticides' to achieve the same effect as above.

If the predators are not introduced artificially and some 'pesticide' is introduced

that is assumed not to kill round gobies but mussels, then x(t) and y(t) vary with time in the following way:

$$\begin{cases} \frac{dx}{dt} = x(r_1 - \alpha y) - u(t)x \\ \frac{dy}{dt} = y(-r_2 + \beta x) \end{cases}$$

where x denotes the density of zebra mussel in the lake A, y denotes the density of round goby fish in the lake A, and u(t) denotes the introducing rate of the pesticide, $0 \le u(t) \le b$ (b is a positive constant).

Usually a pesticide kills not only zebra mussels but also the predators. Hence we need a modified model shown below:

$$\begin{cases} \frac{dx}{dt} = x(r_1 - \alpha y) - e_1 u(t) x\\ \frac{dy}{dt} = y(-r_2 + \beta x) - e_2 u(t) y \end{cases}$$

s.t.
$$\begin{cases} x(0) = x_0 \\ y(0) = y_0 \\ 0 \le u(t) \le b \\ x(T) \le \overline{x} \\ y(T) \ge \overline{y} \end{cases}$$

Our aim is to solve the following multi-objective optimization problem:

$$\begin{cases}
\min \int_0^T x(t)dt \\
\min \int_0^T u(t)dt
\end{cases}$$

where u(t) denotes the introducing rate of the pesticide, x denotes the density of zebra mussel in the lake A at time t, y denotes the density of round goby fish in the lake A at time t, e_1 denotes the killing rate of the pesticide to mussels, and e_2 denotes the killing rate of the pesticide to round gobies.

The model above also proved correct in reality. Now we already have a few ways to reduce the number of zebra mussels by introducing different factors such as predators, pesticides, physical methods, or a mixture.

Crafted Report (at the end of the article)

Strengths and Weaknesses of the Models

1) Strengths

The two models proposed I part II that deal with the effect of more than ten environmental factors on mussel growth based on the statistic data taken from lake A are satisfactory because of its success in that it shows that Cl and K have a desirable effect on zebra mussels, and Ca promotes their growth, and, based on it, an effective way is proposed to decrease the growth of mussels is to increase the amount of Cl and K in a polluted lake and/or, decrease that of Ca.

The time-variable model presented in part III is quite good because of its robustness (that is, the model is insensitive to the parameters) as well as its strong prediction capability (the prediction for lake A and lake B are two successful instances).

The models proposed in part VI are very close to our case and, therefore, are worthy doing further study.

2) Weaknesses

Some of our models are nonlinear and, hence, are hard to study. Furthermore, more related factors should be taken into account.

Reference

- 1. William F. Lucas. 1983. 《Modules in Applied Mathem atics. Vol. 4: Life Science Modds》. Springer-Verlag New York, Inc.
- 2. Ren Shanqiang. 1998. *《Mathematical Model》*. Chongqing University Publishing Company, China.
- 3. Ervin Y.Rodin. 1994. 《Mathematical and computer Modelling》. Washington University.
- 4. Jiang QiYuan. 1987. *《Mathematical Model》*. Advanced Educational Publishing Company, China.
- 5. Math Works. 1992. *《Matlab3.5m》*

Crafted Report:

Recommendation to introduce the round goby

With the spread of zebra mussels, the Great Lakes ecosystem and economy are impacted significantly. Zebra mussels colonize on various smooth surfaces, such as docks, boat hulls, commercial fishing nets, water intake pipes and valves. To solve this problem, there are many methods to be adopted, such as chemical treatments, anti-coolant coatings, physical removal, thermal shocking, desiccation, and oxygen deprivation. The method of introducing the round goby that the local community leaders adopted is an advisable choice. The following is our explanation of recommendation to introduce the round goby.

The published articles and books on zebra mussels have pointed out that Zebra mussel is an important component of the gobies' diet in their native range; and in laboratory studies in North America, a single round goby can eat up to 78 zebra mussels a day. Thus, after we introduce the round goby, the population of Zebra mussels will significantly decrease. This is one of the greatest advantages of introducing round gobies.

The round goby can displace native fish, eat their eggs and young, take over optimal habitat, spawn multiple times a season, and survive in poor quality water which give them a competitive advantage. After they reach a new area, they reproduce rapidly, which will threaten the survival of native fish greatly. This is one of the greatest disadvantages of introducing round gobies.

Without exception, everything has two sides. The most important thing that we can do is to minimize its negative effect and utilize its positive effect. Thus, when we adopt the recommendation of introducing round gobies. We must balance advantages and disadvantages and control the number of the round goby introduced.