

Secrets of Altai: An Ecological Model of Birds and Forests

Yurui Wu*, Zhengqing Ou, Qihui Hu, Haixiang Zhang, Chang Shi

Northeastern University at Qinhuangdao, School of Control Engineering, No. 143 Taishan Road, Qinhuangdao Economic & Technological Development Zone, Hebei Province, China

Abstract. With intensified global climate change and human activities, ecosystem stability and biodiversity face significant challenges. Birds, as key ecosystem indicators, are closely linked to vegetation cover, which is vital for assessing ecosystem health and guiding conservation strategies. This study focuses on China's Altai region, exploring the relationship between bird populations and vegetation cover in riparian landscapes through a comprehensive mathematical model and analytical framework encompassing spatial distribution, ecological stability, and dynamic prediction. Utilizing Lotka-Volterra ordinary differential equations (ODEs) and random forest regression, the study elucidates the impact of vegetation changes on bird populations and ecosystems, providing scientific support for regional biodiversity conservation.

1 INTRODUCTION

The Mountain natural ecosystems play a crucial role in maintaining regional biodiversity, regulating microclimates, and protecting ecological balance. The Altai Mountains in China are located in the arid deserts and arid-semiarid zones in the hinterland of Asia, with typical continental climate characteristics: warm springs and autumns, cold winters, no summer throughout the year, and abundant rainfall. They are a core area for studying ecological diversity and ecosystem stability in temperate arid zones. Among them, the riparian zones of the Altai Mountains, with diverse vegetation types and stable water supply, provide an ideal habitat for various bird species, and thus become a typical research area for exploring the relationship between vegetation cover and bird population dynamics.

As sensitive indicators of ecosystem health, birds' population richness, community composition, and temporal-spatial variations are closely related to the vegetation cover and food resources of riparian landscapes. Previous studies have shown that vegetation cover gradient is one of the key factors affecting bird community structure, and the intrinsic connection between the two is of great significance for understanding ecosystem dynamic equilibrium and formulating scientific biodiversity conservation strategies^[1-4]. However, existing research on the riparian zones of the Altai Mountains mostly focuses on single vegetation or bird community surveys, lacking systematic quantitative analysis and mathematical modeling of the temporal-spatial changes of bird populations along the vegetation cover gradient, which limits the in-depth understanding of the ecological mechanism between vegetation and bird populations in this region.

To fill this research gap, this study takes the 630-kilometer-long riparian landscape of China's Altai

Mountains as the research object. Based on the survey data from previous studies, this area is artificially divided into 9 study areas that can reflect the natural vegetation cover gradient. Based on the pre-processed data, this study will conduct in-depth quantitative analysis, construct an original ecosystem stability evaluation system, further clarify the temporal-spatial variation characteristics of bird populations in the study area, and provide a scientific basis for regional biodiversity conservation and riparian ecosystem management.

2 MODEL preparations

2.1 Description of Notation

The meanings of the notations in the model in this paper are shown in the Table 1.

Table 1. Description of Notation.

Notation	Description
T	Total biomass of the ecosystem
n	Total species in the ecosystem of the Altai Mountains in China
l_i	Population density of the i -th species
S	Shannon-Wiener Index
K_{plant}	Carrying capacity of plant resources
p_i	The proportion of biomass of species i in the ecosystem of this region
α	Natural growth rate of prey
β	the predation rate
γ	the number of predators
x	the number of prey
δ	the growth rate of the predator due to predation on the prey
γ	the natural death rate of the predator

* Corresponding author: wyr20041123@163.com

2.2 Data collection and pre-processing

2.2.1 Data collection

To facilitate the validation and evaluation of our model, we collected relevant data on the ecosystem of the Altai Mountains in China, and the data sources are as follows: <https://datadryad.org/dataset/doi:10.5061/dryad.3m66bd1>.

2.2.2 Data pre-processing

In this ecosystem, the majority is composed of bird populations and plants. To further investigate the relationship between ecosystem stability and plant changes, this section categorizes different species of birds into four groups: carnivores, herbivores, insectivores, and omnivores. The quantities of each of the four bird types are counted and organized, and the statistical results are visualized in Fig. 1.

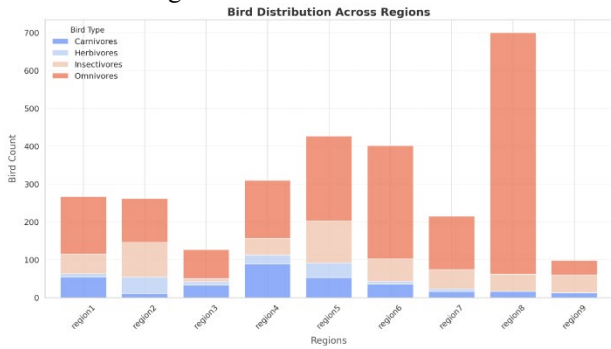


Fig. 1. Bird Distribution Across Regions.

3 MODEL BUILDING

This task presents a dynamic, holistic model with two parts: the first examines the link between plant changes and ecosystem stability, while the second establishes a prediction model for bird population changes influenced by external factors. To address these, two models are developed. The first is a TSL ecological evaluation model that uses total biomass, species diversity, and ecosystem stability as indicators. Second, the random forest algorithm is used to predict the changing trend of bird population under external influences.

3.1 TSL ecological evaluation model

To predict the impact of plant changes on ecosystems, an evaluation model called TSL (Total Biomass, Species Diversity, and Ecosystem Stability) is established. The modeling approach is as follows:

3.1.1 Total Ecosystem Biomass

Otal ecosystem biomass is a key indicator of ecosystem health, reflecting the biotic community's status. Under equilibrium conditions with a fixed number of species, total biomass and ecosystem stability are positively correlated. Higher total biomass indicates greater stability,

while lower biomass suggests fragility and reduced resilience^[5]. For n species in the Altai Mountains ecosystem, with population density l_i and biomass s_i for the i -th species, the total biomass T is:

$$T = \sum_{i=1}^n s_i = \sum_{i=1}^n l_i W_i$$

To address the effects of varying sampling point numbers and area sizes on results, this section calculates total biomass per sampling point to represent overall biological distribution characteristics. A higher average distribution density indicates greater total biomass, suggesting that, under consistent conditions, higher biological density correlates with a more resilient ecosystem capable of withstanding external stimuli. By comparing biological distribution density, we aim to understand how plant changes affect average biological distribution density. The refined total biomass after classification is expressed as a function of average biological distribution density:

$$\eta = \frac{T}{N} = \frac{1}{N} \sum_{i=1}^n s_i = \frac{1}{N} \sum_{i=1}^n l_i W_i$$

Where N is the number of sampling points within the area. This section establishes a connection between average biological distribution density and plant coverage, with the plant coverage also being averaged. The output results are visualized using thermodynamic imagery, as shown in Fig. 2.

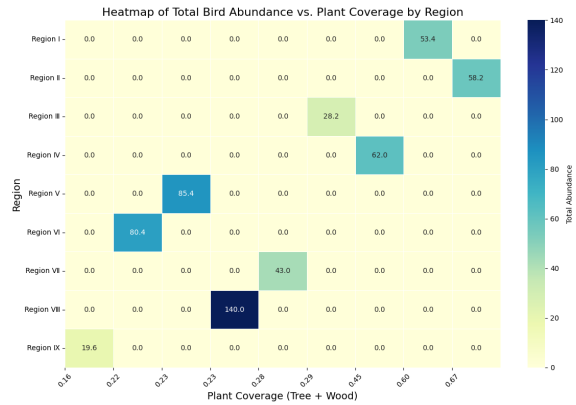


Fig. 2. Heatmap of Total Bird Abundance vs. Plant Coverage by Region.

From the above output results, it can be observed that the average biological distribution density is highest when the plant coverage is around 25%. There is no positive correlation between plant coverage and average biological distribution density; instead, it follows a distribution pattern similar to a normal distribution, characterized by "low on both sides and high in the middle." This means that when plant coverage is either too low or too high, the average distribution density of birds is relatively low.

3.1.2 Shannon-Wiener Diversity Index

Evaluating species diversity is an indispensable step in assessing the condition of an ecosystem. Generally, the higher the species richness, the more balanced the biomass among species, and the better the state of the biological

system. Therefore, this section uses the Shannon-Wiener diversity index to consider both species richness and evenness. Compared to the species richness provided in the dataset, the diversity index set in this chapter considers the physical "entropy" perspective^[6-7]. The higher the species diversity, the greater the disorder, and the higher the "entropy" value, resulting in a larger index. The Shannon-Wiener index S is defined as:

$$S = \sum_{i=1}^n p_i \ln p_i$$

where p_i represents the proportion of the biomass of species i in the ecosystem.

Due to the significant impact of species variety on the diversity index calculation, it is essential to use the original data rather than the pre-processed results. Therefore, the diversity index should be calculated based on the original 81 bird species. After inputting the original data into the diversity index calculation formula, the Shannon-Wiener diversity index of each sampling point in all regions can be calculated.

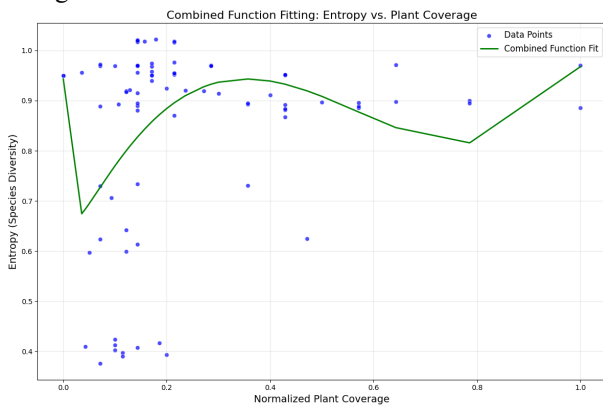


Fig. 3. Combined Function Fitting: Entropy vs. Plant Coverage.

In this section, we will first use function fitting to explore the relationship between the Shannon-Wiener diversity index and plant coverage. We will employ three basic types of functions: polynomial functions, exponential functions, and logarithmic functions, to fit the relationship between the diversity index and plant coverage. After fitting, we will perform nonlinear optimization to ensure that the fitted function adheres to the fundamental logic of function fitting. The diversity index set in this chapter considers the physical "entropy" perspective. This means that we aim to have as many data points as possible fall on the fitted function graph while also ensuring that the function graph reflects the general trend of the changes between the two variables. Finally, we will present the relationship between the diversity index and plant coverage. The regularity curve is showed in Fig. 3.

The output results indicate that the weight of the exponential component is the largest, significantly influencing the fitted function. Graph analysis shows that when plant coverage density is either too high or too low, biodiversity exhibits considerable variability, reflecting complexity and randomness in these areas. Specifically, when plant coverage is at 30%, biodiversity reaches its

highest point, indicating strong ecosystem stability at this level.

3.1.3 Ecosystem Stability Assessment

Stability is one of the important conditions for maintaining the normal function of an ecosystem. When biological stability is good, it means that the ecosystem is in a dynamic equilibrium state, indicating that the distribution of biomass within the ecosystem will change through interactions and regulation among species. In this case, an ecosystem with more species and a complex food web can adjust its structure and function more quickly, making the ecosystem more stable. Based on the above theories and evidence, this chapter expands the Lotka-Volterra model by adding plant coverage as a driving factor to the original model^[8]. The mathematical expressions are as follows.

For the prey:

$$\frac{dx}{dt} = \alpha x \left(1 - \frac{x}{K_{plant}} \right) - \beta xy$$

where x represents the number of prey, specifically herbivorous and insectivorous birds; K_{plant} refers to the carrying capacity of plant resources (determined by trees and wood); α is the natural growth rate of the prey; and β represents the predation rate.

For the predator:

$$\frac{dy}{dt} = \delta xy - \gamma y$$

where y represents the number of predators, specifically carnivorous birds; δ is the growth rate of the predator due to predation on the prey; and γ is the natural death rate of the predator. $K_{plant} = c(\text{tree} + \text{wood})$, c is a proportionality constant used to adjust the magnitude of the carrying capacity.

This section analyzes plant coverage and ecosystem stability, isolating plant coverage by using preprocessed data. Bird populations are grouped into herbivorous, carnivorous, omnivorous, and insectivorous. Under fixed resources, herbivorous and insectivorous birds are treated as prey, carnivorous birds as predators, and omnivorous birds remain neutral. Their numbers are excluded in the simplified model.

To further reduce the impact of the inherent habits and characteristics of different bird species on the experiment, relevant literature has been reviewed. In classical ecological models, it is assumed that the growth rate of the bird populations is 10%, the predation rate is 5%, and the mortality rate of the carnivorous bird population is 8% per cycle. These parameters will be employed to calculate the natural growth rates of the prey (herbivorous and insectivorous birds), as well as the predation rates and mortality rates of the predators (carnivorous birds). This approach will help clarify the dynamics between plant coverage and ecosystem stability while accounting for the complexities introduced by different bird species.

In this context, According to multiple experiments, the parameter biological carrying capacity of plant resources K is taken as $K_{plant} = 1.66 \times 10^{10}$. By incorporating the relevant values into the extended Lotka-Volterra model discussed in this chapter, we can simulate the numerical relationship between predators and prey over the next 200

years based on the predefined assumptions. The changes in their populations will be visualized using function graphs, and the output results are as Figure 4:

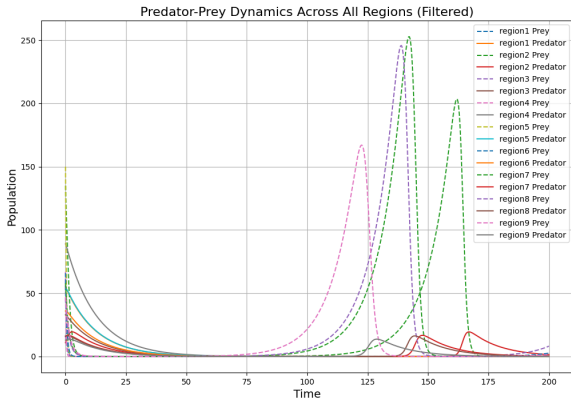


Fig. 4. Predator-Prey Dynamics Across All Regions.

In this simulation, it is assumed initially that prey are abundant and predators are few. As predators exploit prey resources, predator numbers grow slowly. In the mid-stage, predators face intense predation pressure, leading to a sharp decline in their population, while the prey population peaks. However, once the prey population falls rapidly, the predator population also declines. Ultimately, both populations tend to stabilize, though in certain areas extreme predation still causes crashes and cyclical fluctuations. When analyzing ecosystem robustness, the study focuses on resistance to disturbances and recovery ability, examining both prey and predator populations. For prey, despite initially high populations, significant predation pressure causes rapid declines. In Regions 4 and 5, prey populations rebound quickly, indicating strong robustness. In contrast, Regions 6 and 7 show weaker rebound capabilities, with long-term low populations and potential crashes, reflecting poor robustness. For predators, population stability depends on prey resources. Regions 1 and 9 experience rapid predator declines, indicating poor robustness. Conversely, Regions 5 and 6 maintain relatively stable predator populations, demonstrating adaptability to resource fluctuations. Overall, Regions 5 and 6 exhibit higher robustness, supporting long-term dynamic equilibrium, while Regions 1 and 9 show lower robustness, risking population collapse. Region 5 and 6 feature moderate plant coverage (25%), Region 1 has high plant coverage, and Region 9 has low plant coverage.

3.2 Prediction Model

3.2.1 Variables setting

Since this inquiry focuses solely on the population size of bird species, the model is further simplified. It is reduced to a predictive model that uses annual precipitation, elevation, and average annual temperature as independent variables to forecast the abundance of birds in a given area^[9].

Set the input (feature) variables as follows:

$$x_1 = bio_{12}(AP), x_2 = bio_{12}^2, x_3 = elevation$$

$$x_4 = \log(1 + elevation), x_5 = bio_1, x_6 = bio_2^2$$

Let x_1 represent annual precipitation, x_5 represent average annual temperature, and x_2 and x_6 be used to capture nonlinear relationships. The aforementioned four variables reflect climatic characteristics. Additionally, there are two variables to represent geographical features, where x_3 denotes elevation and x_4 is the logarithm of elevation, used to smooth the effects of terrain. These six variables constitute the feature vector $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}^T$.

The output variable (target variable) is defined as y is abundance. The objective function is set as $f(x)$, such that

$$y = f(x_1, x_2, x_3, x_4, x_5, x_6) + \epsilon$$

where $f(x)$ is a nonlinear function learned through random forests, and ϵ is the error term, representing the difference between the predicted and actual values.

3.2.2 Random Forest Regression Model Based on Decision Tree

A single decision tree partitions the data to minimize MSE , with each leaf node $R_{m,t}$ output as the mean target value of samples in that region:

$$f_t(X) = \frac{1}{|R_{m,t}|} \sum_{x_i \in R_{m,t}} y_i$$

A random forest averages T independently trained trees:

$$f_{RF}(X) = \frac{1}{T} \sum_{t=1}^T f_t(X)$$

where the MSE is:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - f_{RF}(x_i))^2$$

where N is the number of training samples, y_i is the actual abundance value of the i -th sample, and $f_{RF}(x_i)$ represents the predicted abundance value of the i -th sample by the random forest.

Finally, integrating the above formula and incorporating the relevant decision variables, we obtain the complete random forest formula:

$$\hat{y}_i = f_{RF}(X_i) = \frac{1}{T} \sum_{t=1}^T \frac{1}{|R_{m,t}|} \sum_{x_i \in R_{m,t}} y_i$$

where \hat{y}_i represents the predicted abundance value of the i -th test sample.

$X = \{x_1, x_2, x_3, x_4, x_5, x_6\}^T$ is the feature vector of the test sample.

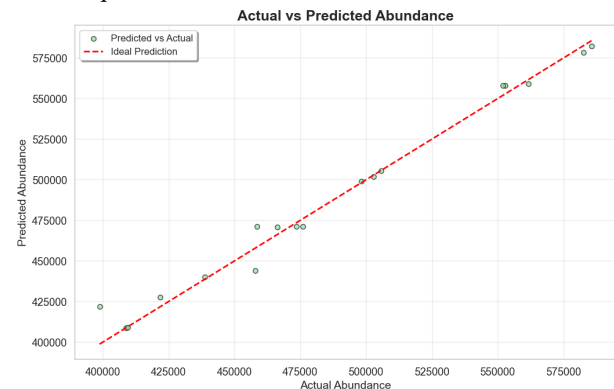


Fig. 5. Python Simulation Results.

Use Python to simulate the established prediction model, output the *MSE* metric and the fitting evaluation metric R^2 , Print the predicted values and actual values generated during the process to a file for backup, and use visualizations to present some of the predicted values alongside the actual values. The output results are shown in the Fig 5.

4 Problem Solving

According to the assessment results of the TSL evaluation model, an analysis from the perspectives of total biomass, diversity, and robustness of the ecosystem indicates that when plant coverage is between 25% and 30%, the ecosystem exhibits strong stability. At this level, the ecosystem demonstrates significant resistance and recovery capacity when facing internal disturbances and external interventions, allowing it to maintain a dynamic equilibrium over an extended period. Conversely, when plant coverage is either too low or too high, the stability of the ecosystem deteriorates, leading to reduced elasticity in response to disturbances, which may result in population collapses and cyclical fluctuations.

The Mean Squared Error (MSE) indicators for abundance vary between hundreds of thousands and tens of millions, with the calculated results falling within a reasonable range. The R^2 value is high, reaching 98%, indicating that the established random forest model effectively meets the objective function set for this chapter. After visualizing the output results, the following image can be obtained, where the red curve represents the predicted curve, effectively illustrating the changes in bird population numbers under the influence of external environmental factors.

References

1. Na Li, Yue-Hua Sun, Hongjun Chu, et al. Bird species diversity in Altai riparian landscapes: Wood cover plays a key role for avian abundance. *Ecology and evolution*. (2019);**9** (17):9634-9643. doi:10.1002/ece3.5493 intr
2. Paul E. Bellamy, Elisabeth C. Charman, Neil Riddle, et al. Impact of woodland agri-environment management on woodland structure and target bird species. *Journal of environmental management*. (2022);**316** (0):115221-115221. doi:10.1016/j.jenvman.2022.115221
3. Golam Kibria. Contamination of coastal and marine bird species with plastics: Global analysis and synthesis. *Marine pollution bulletin*. (2024);**206** (0):116687-116687. doi:10.1016/j.marpolbul.2024.116687
4. Jonathan Lewis-Phillips, Steve Brooks, Carl D. Sayer, et al. Pond management enhances the local abundance and species richness of farmland bird communities. *AGRICULTURE ECOSYSTEMS & ENVIRONMENT*. (2018);**273** (0):130-140. doi:10.1016/j.agee.2018.12.015
5. Ranjan Muthukrishnan. The relationship between native species richness and exotic species richness or occurrence will always be negative when the total number of species is accounted for in statistical models: A response to Beaury et al. *Ecology letters*. (2021);**24** (8):1732-1734. doi:10.1111/ele.13764
6. Song H S, Wang S. Effects of electroacupuncture on behavioral changes, average color change rate per well of intestinal microflora, and Shannon-Wiener diversity index in chronic stress depression rats[J]. *World Journal of Integrated Traditional and Western Medicine*, (2020), **15**(01): 88-91. DOI:10.13935/j.cnki.sjzx.200118..
7. Petra Pártlová, Jan Váchal, Ján Dobrovič, et al. Use of The Shannon Wiener Index To Measure Lag Diversity, A Major Player in Regional Rural Development. *Polish Journal of Management Studies*. (2020);**22** (1):385-400. doi:10.17512/pjms.2020.22.1.25
8. Changwook Yoon. Global dynamics of a Lotka-Volterra type prey-predator model with diffusion and predator-taxis. *APPLICABLE ANALYSIS*. (2021);**101** (16):5557-5570. doi:10.1080/00036811.2021.1898598
9. Abdul-Lateef Balogun, Abdulwaheed Tella. Modelling and investigating the impacts of climatic variables on ozone concentration in Malaysia using correlation analysis with random forest, decision tree regression, linear regression, and support vector regression. *Chemosphere*. (2022);**299** (0):134250-134250. doi:10.1016/j.chemosphere.2022.134250