Problem Chosen	2025 MCM/ICM Summary Sheet	Team Control Number
A		JS06012025183

2025NEUQMCM PreContest

The Altai Mountains, located in northwestern China, are a vital ecological region celebrated for their exceptional biodiversity and distinctive riparian landscapes. Spanning a 630 km riverbank gradient, the region supports a diverse array of flora and fauna, forming a dynamic ecosystem shaped by complex environmental factors. Birds, serving as bioindicators, play a key role in reflecting ecosystem health. Their species richness and spatial distributions are intricately linked to vegetation coverage, altitude, and habitat heterogeneity. However, this delicate relationship is increasingly under threat from external pressures such as climate change, land-use modifications, and human activities, jeopardizing the stability and resilience of these ecosystems. Understanding these interdependencies is crucial for devising effective biodiversity conservation strategies and promoting sustainable ecological management practices in the region.

This study employs the MaxEnt model, along with a series of advanced optimizations such as model ensemble, feature enhancement, and regularization, to investigate the spatial and altitudinal distribution patterns of bird species and vegetation. By integrating ecological data with comprehensive modeling and scenario-based analyses, we assess the impact of vegetation changes on bird populations and provide deeper insights into species-environment interactions.

The results demonstrate that bird populations are profoundly influenced by vegetation diversity and abundance, with habitat heterogeneity along altitudinal gradients playing a pivotal role in sustaining biodiversity. Scenario analyses further reveal that increases in vegetation coverage significantly promote bird population growth, while reductions, though impactful, are often mitigated by compensatory ecological factors.

The improved MaxEnt model, enhanced through model ensemble techniques and feature engineering, delivered superior predictive accuracy, illustrating the potential of advanced modeling approaches in ecological research. This study underscores the importance of vegetation conservation in safeguarding bird populations and ensuring ecosystem stability. The findings provide actionable insights for biodiversity conservation and ecological management in the Altai Mountains, offering a replicable methodological framework for similar ecological studies. By combining cutting-edge modeling techniques with ecological data, this research contributes to the development of more effective strategies for sustaining biodiversity in rapidly changing environments.

Key Words: MaxEnt Model; Model Ensemble; Feature enhance; Regularization

Contents

1 Problem Background and Related Research Overview . . . **1**
1.1 Problem Background . . . 1
1.2 Key Questions . . . 2
1.2.1 Spatial and Altitudinal Distributions of Birds and Vegetation . . . 2
1.2.2 Impact on Bird Populations in the Altai Ecosystem . . . 2
1.2.3 Ecosystem Stability and Plant Dynamics . . . 2
1.3 Objectives . . . 2
1.4 Related Work . . . 3

2 Spatial and altitudinal distribution of birds and timber . . . **3**
2.1 Spatial distribution of timber . . . 3
2.2 Spatial distribution of birds . . . 4
2.3 Altitudinal distribution of birds and timber . . . 4

3 Relationship between birds and other ecological groups . . . **6**
3.1 MaxEnt Model . . . 6
3.2 Improvement for MaxNet model . . . 8

4 Impact on the stability of the ecosystem . . . **9**

5 Conclusion . . . **12**

References . . . **13**

Appendices . . . **13**

Synopsis/Memo/Handout

This study focuses on the Altai Mountains, a biodiversity hotspot in northwestern China, and investigates the intricate relationships between bird species richness, vegetation distribution, and ecosystem stability. By leveraging advanced modeling techniques, including the MaxEnt model, model ensemble approaches, feature enhancement, and regularization, we analyze the spatial and altitudinal patterns of bird populations and timber distribution across a 630 km riverbank gradient.

Birds, as key bioindicators, reflect ecosystem health through their species richness and distribution, which are strongly tied to vegetation coverage and altitude. Using comprehensive data analysis and scenario-based modeling, we evaluate how changes in vegetation, driven by natural and anthropogenic factors, impact bird populations and ecosystem dynamics. The optimized MaxEnt model demonstrated superior predictive accuracy compared to traditional methods, enabling deeper insights into species-environment interactions.

Our results highlight that increasing vegetation coverage significantly enhances bird populations, while reductions, although impactful, can be partially mitigated by compensatory factors such as habitat diversity. This underscores the importance of habitat heterogeneity in maintaining biodiversity. Conservation strategies focusing on reforestation and sustainable land management are critical for preserving the ecological integrity of the Altai Mountains.

By combining advanced modeling techniques with ecological data, this research provides a replicable framework for biodiversity conservation and ecological management. It lays the groundwork for future studies aimed at addressing the challenges of sustaining biodiversity in the face of climate change and land-use alterations.

1 Problem Background and Related Research Overview

The Altai Mountains in northwestern China are a biodiversity hotspot. Researchers studied bird species richness and environmental factors across nine regions, measuring wood cover and other variables. Birds, as bioindicators, provide insights into ecosystem health, highlighting the interplay between vegetation and bird populations in this unique ecological landscape.

1.1 Problem Background

The Altai Mountains, located in northwestern China, are an ecological hotspot characterized by unique riparian landscapes and diverse habitats. Spanning 630 kilometers, this region supports a variety of flora and fauna, playing a critical role in maintaining regional biodiversity. Of particular interest to scientists is the relationship between bird species richness, vegetation composition, and environmental factors across this gradient. Birds, as bioindicators, can reflect the health and stability of ecosystems, making their study essential for understanding the broader ecological dynamics of the Altai region.

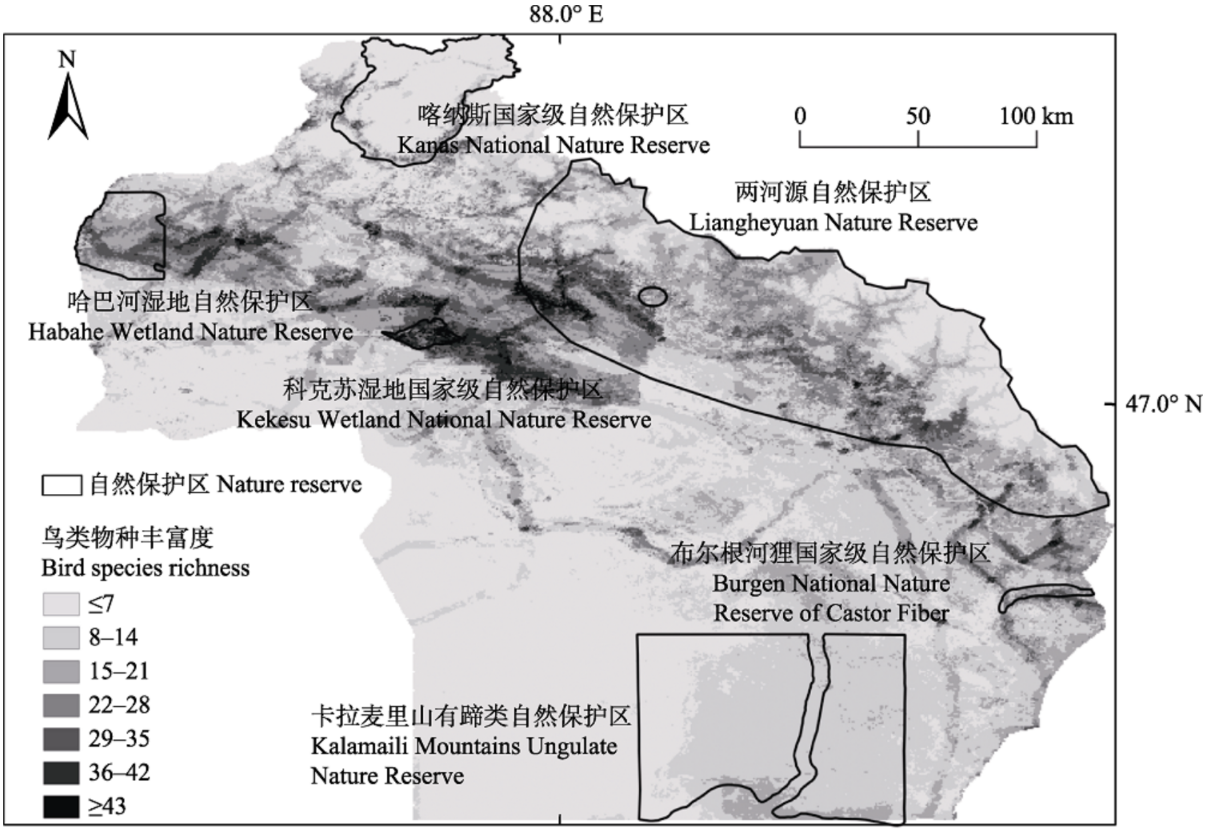

Figure 1: Birds distribution in Altai area

To investigate these dynamics, researchers conducted surveys at nine distinct regions along the riparian landscape, selecting 10 sampling points in each region. The surveys collected data on bird species richness and spatial distribution, alongside environmental variables such as wood cover, defined as the proportion of land covered by trees and shrubs taller than 3 meters. These measurements offer insight into how vegetation and other ecological factors influence bird populations.

1.2 Key Questions

1.2.1 Spatial and Altitudinal Distributions of Birds and Vegetation

The study seeks to analyze how bird species richness and wood cover vary spatially and with altitude across the Altai landscape. Understanding these patterns is crucial for identifying ecological zones with high biodiversity and assessing the role of habitat heterogeneity.

1.2.2 Impact on Bird Populations in the Altai Ecosystem

Birds in the Altai region are intricately linked to other ecological groups, including plants and human activities. The study aims to construct a continuous model to describe these interdependencies and evaluate how changes in these factors influence bird populations.

1.2.3 Ecosystem Stability and Plant Dynamics

Vegetation changes, whether due to natural progression or anthropogenic activities, can disrupt the ecological balance. The study will examine the impact of such changes on the stability of the ecosystem and predict bird population trends under varying external factors, such as land-use changes, climate shifts, and human interventions.

1.3 Objectives

This research aims to:

- Analyze and describe the spatial and altitudinal distribution patterns of bird species and vegetation.
- Model the relationship between bird populations and other ecological factors, including plants and human influences.
- Assess the stability of the ecosystem in response to vegetation changes and predict bird population trends under these conditions.

The findings will provide valuable insights into the conservation and management of the Altai Mountains, contributing to the preservation of its rich biodiversity and ecological integrity.

1.4 Related Work

Previous studies highlight the critical role of vegetation cover in supporting bird diversity, with trees, shrubs, and saplings providing essential resources like nesting sites and food. High beta diversity, driven by habitat heterogeneity and environmental gradients, is common in riparian landscapes, emphasizing their conservation importance. Environmental factors such as precipitation and vegetation structure significantly influence bird composition, as shown in studies using Canonical Correspondence Analysis (CCA). This research builds on these findings, demonstrating that wood cover and precipitation explain 33.24% of beta diversity. Expanding nature reserves and limiting farming and grazing in riparian zones are vital for preserving avian biodiversity. [1]

2 Spatial and altitudinal distribution of birds and timber

The Altay Region in China is located in North Xinjiang, in the border region of China, Kazakhstan, Russia and Mongolia. The region contains the Altai Mountains and the desert and riparian landscapes south of them, which constitutes the Altai-Sayan biodiversity hotspot and one of the Global 200 Biodiversity Ecoregions.

2.1 Spatial distribution of timber

To see the spatial and altitudinal distribution of timber, we first draw a picture of timber observation area in Altay, from which we can see that nine different areas are selected. The timber is generally distributed in the given area.

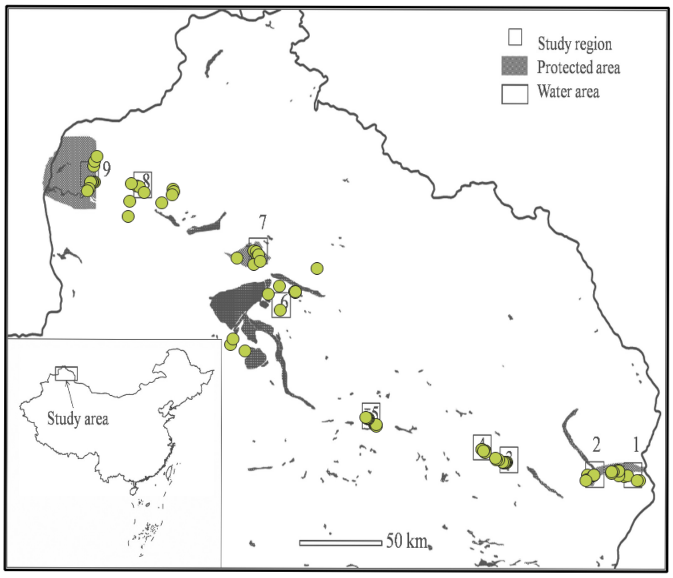

Figure 2: Selected area

2.2 Spatial distribution of birds

From a distribution perspective, the bird species in the Altay region are predominantly northern types, with a total of 170 species, accounting for 49.4% of the total number of birds. Among these, the Palaearctic type (U) includes 110 species, and the Holarctic type (C) includes 60 species.

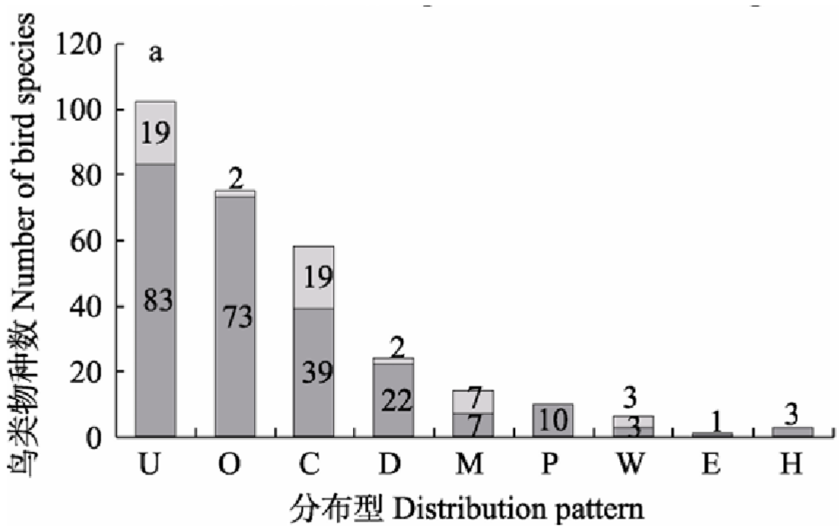

Figure 3: distribution in Altai area

The quantities of birds from other distribution types, ranked from high to low, are as follows:

- Widespread species (O): 93 species, accounting for 27.0% - Central Asian type (D): 33 species, accounting for 9.6% - Northeast type (M): 21 species, accounting for 6.1% - Highland type (P): 8 species, accounting for 2.3% - Oriental type (W): 8 species, accounting for 2.3% - Monsoon type (E): 1 species, the Oriental Turtle Dove (Streptopelia orientalis) - HimalayanHengduan type (H): 3 species, including the Grey-backed Shrike (Lanius tephronotus), the Black-crowned Tit (Parus rubidiventris), and the Himalayan Accentor (Prunella himalayana).

In the Altai Mountains, a total of 297 bird species are distributed, predominantly northern types, with 160 species: 102 species from the Palaearctic type and 58 species from the Holarctic type, with breeding birds accounting for 81.4% and 67.2% respectively.

In the foothill plain areas, which include oasis and desert habitats, there are 278 bird species, among which the Palaearctic type comprises 92 species, the Holarctic type includes 49 species, and the Central Asian type has 30 species, with breeding birds accounting for 78.3%, 59.2%, and 87.1% of each distribution type respectively.

2.3 Altitudinal distribution of birds and timber

Based on field surveys and literature review, the timber is divided into 6 different category of different hight. And the number of bird species in the Altay region varies by elevation zones as follows:

- Alpine Bare Rock Zone: 24 species - Alpine Meadow Zone: 35 species - Mountain Forest Steppe Zone: 172 species - Low Mountain Shrub Zone: 130 species - Desert Steppe Zone: 84 species - Plain Oasis Zone: 173 species

The bird species mentioned above do not include waterbirds from the orders Podicipediformes, Anseriformes, Gruiformes, and Charadriiformes. A total of 92 species of waterbirds were recorded in wetland and aquatic habitats across these vegetation zones.

The results of above analysis are shown in the following figure.

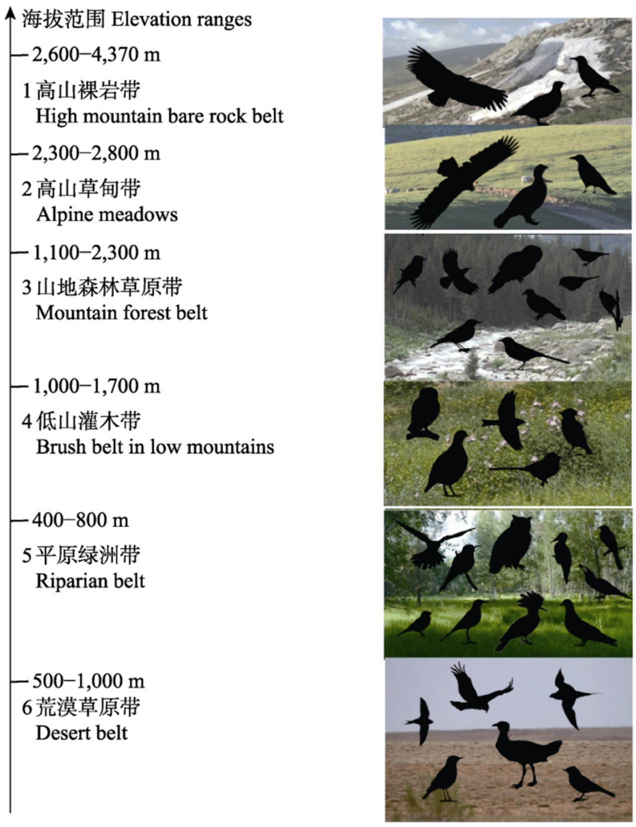

Figure 4: altitudinal distribution in Altai area

In the field survey, the Alpine Bare Rock Zone recorded 16 bird species, with notable species including the Eurasian Griffon Vulture (Aegypius monachus), Himalayan Vulture (Gyps himalayensis), Red-billed Chough (Pyrrhocorax pyrrhocorax), and Rock Ptarmigan (Lagopus muta). The Alpine Meadow Zone recorded 28 bird species, with common species such as the White Wagtail (Motacilla alba) and Black Kite (Milvus migrans).

In the Mountain Forest Steppe Zone, a total of 119 bird species were recorded, with abundant species including the Grey-headed Bullfinch (Fringilla coelebs), Red-throated Thrush (Turdus ruficollis), Grey-breasted Bunting (Emberiza cia), Spotted Flycatcher (Muscicapa striata), Common Rosefinch (Carpodacus erythrinus), and Grey Wagtail (Motacilla cinerea).

The Low Mountain Shrub Zone recorded 101 bird species, including many Black-throated Stonechats (Saxicola maurus) and Eurasian Nightingales (Luscinia megarhynchos). The Desert Steppe Zone recorded 68 bird species, with common species such as the Common Buzzard (Buteo japonicus), House Martin (Delichon urbicum), Horned Lark (Eremophila alpestris), Great Sandgrouse (Syrrhaptes paradoxus), Mongolian Lark (Bucanetes mongolicus), Common Swift (Apus apus), and Chukar Partridge (Alectoris chukar).

In the Plain Oasis Zone, 124 bird species were recorded, including common species such as the Common Kestrel (Falco tinnunculus), Lesser Kestrel (Falco naumanni), European Bee-eater (Merops apiaster), European Roller (Coracias garrulus), Common Starling (Sturnus vulgaris), and Rock Pigeon (Columba livia).

In wetlands and water bodies across all vegetation zones, a total of 62 waterbird species were recorded, showcasing both species richness and significant populations. Common species included the Ruddy Shelduck (Tadorna ferruginea), Gadwall (Anas strepera), Red-crested Pochard (Netta rufina), Black Stork (Ciconia nigra), Great Egret (Ardea alba), Greylag Goose (Anser anser), and the Yellow-legged Gull (Larus cachinnans), with populations numbering in the hundreds. Additionally, the Whooper Swan (Cygnus cygnus), Eurasian Spoonbill (Platalea leucorodia), Pheasant-tailed Jacana (Hydrophasianus chirurgus), Golden Plover, Black-winged Stilt, and Common Tern (Sterna hirundo) were also commonly observed.

3 Relationship between birds and other ecological groups

The Altai Mountains, spanning across Central Asia, represent a unique ecological system characterized by diverse flora and fauna. Among its inhabitants, the bird population plays a crucial role in maintaining ecological balance. Birds contribute to seed dispersal, pollination, pest control, and act as indicators of environmental health. However, the delicate balance of this ecosystem is influenced by multiple factors, including human activity, climate change, and interactions with other ecological groups such as plants and predators.

Understanding the dynamics of the bird population in the Altai Mountains requires a comprehensive study of its interactions with plants (as food sources and nesting habitats), humans (as contributors to habitat loss and conservation efforts), and other species in the food web.

3.1 MaxEnt Model

The maximum entropy principle aims to fully utilize known information while considering the uncertainty of unknown information. It is primarily based on the principle of information entropy maximization. When the amount of information increases, entropy decreases. In scenarios with incomplete information, entropy maximization ensures the least uncertainty in the results.

Suppose a certain high-dimensional environmental space contains n objects of interest. These objects can be categorized into M ecological groups. Based on specific constraints or relationships, the results can be associated with their temporal and spatial distributions. Let X denote the features of an object, Y denote its category, and $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})$ represent a training dataset with m samples. Here, x is the n-dimensional feature vector, and y represents the classification label.

Define the feature function $f(x, y)$ to describe the relationship between the input x and output y:

$$f(x,y) = \begin{cases} 1 & \text{if } x, y \text{ satisfy certain constraints,} \\ 0 & \text{otherwise.} \end{cases}$$

The expectation of the feature function $f(x, y)$ with respect to the empirical distribution $P(X, Y)$ is defined as:

$$E_p(f) = \sum_{x,y} P(x,y) f(x,y).$$

Similarly, the expectation of the feature function $f(x, y)$ with respect to the conditional distribution $P(Y|X)$ is:

$$E_p(f) = \sum_{x,y} P(y|x) f(x,y).$$

If the model can continuously learn, we assume these expectations are equal, i.e.,

$$E_p(f) = E_p(f) \quad (i = 1, 2, \ldots, M).$$

Based on this assumption, the set of models that satisfy all constraints is:

$$E_p(f_i) = E_p(f_i) \quad (i = 1, 2, \ldots, M).$$

The conditional entropy of $P(Y|X)$ is defined as:

$$H(P) = -\sum_{x,y} P(y|x) \log P(y|x).$$

`MaxEnt` is designed for predicting the potential distribution of species. It analyzes ecological distribution by leveraging the principle of entropy maximization to infer the most probable distribution of species based on environmental constraints and limited sample data.

The `MaxEnt` model uses presence-only data for species occurrence and derives relationships between species distribution and environmental variables. The environmental variable layers can be input using GIS tools, such as raster data for geographical and environmental factors. The program performs iterative optimization and evaluates the accuracy of prediction using metrics like ROC curves and AUC values. Generally, an AUC between 0.5 and 0.7 suggests a fair predictive performance, while values above 0.8 indicate high accuracy.

The result is shown in figure 5, which indicates that the result is not good enough and we need to do promotions.

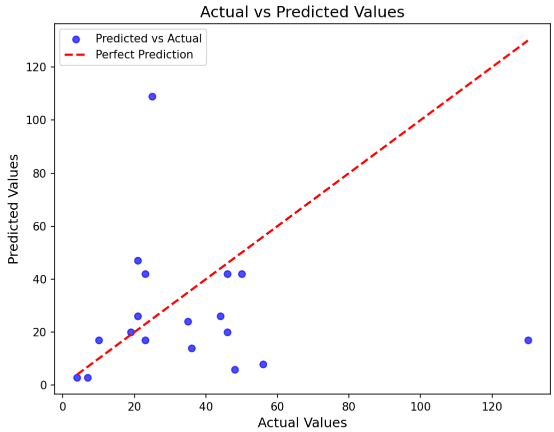

Figure 5: Result of original maxnet

3.2 Improvement for MaxNet model

The previous result shows not good prediction and we need to do more to optimize. Here are three methods we used.

1. Model Ensemble (Voting Regressor) Model ensemble combines the predictions of multiple models to improve overall performance. A Voting Regressor aggregates predictions from different regression models (e.g., linear regression, decision trees, random forests) by averaging their outputs. This method leverages the strengths of each model, reducing the risk of relying on a single model and improving the robustness of predictions.

2. Feature Enhancement and Polynomial Features Feature enhancement involves transforming the original features or creating new ones to better capture relationships in the data. Polynomial features extend the feature set by adding polynomial terms (e.g., x^2, x^3) and interactions between features. This approach enables simpler models to fit more complex data patterns, increasing their flexibility and predictive power.

3. Regularization and Hyperparameter Optimization Regularization adds penalty terms to the loss function of a model to prevent overfitting. Common techniques include L1 regularization (Lasso) and L2 regularization (Ridge), which constrain model complexity by shrinking coefficients. Hyperparameter optimization fine-tunes a model's parameters (e.g., learning rate, tree depth) to achieve better performance. Methods like grid search or randomized search systematically explore parameter combinations, while advanced techniques like Bayesian optimization automate the process more efficiently.

The result is shown in figure 6 and figure 7, from which we can see that the result is much better.

We can draw a conclusion that the bird population in the Altai Mountains plays a critical role in maintaining ecological balance but is subject to various influencing factors. These include:

Plant interactions: Birds rely on plants for food (seeds, fruits, nectar) and nesting habitats. Changes in plant diversity or abundance directly affect bird populations.

Human activities: Agriculture, deforestation, mining, and urbanization can lead to habitat loss, while conservation efforts may protect bird populations. Predation and competition: Birds interact with other animal species as predators, prey, or competitors, influencing population dynamics. Climate factors: Temperature changes, precipitation variability, and extreme weather events affect migration patterns, breeding cycles, and food availability.

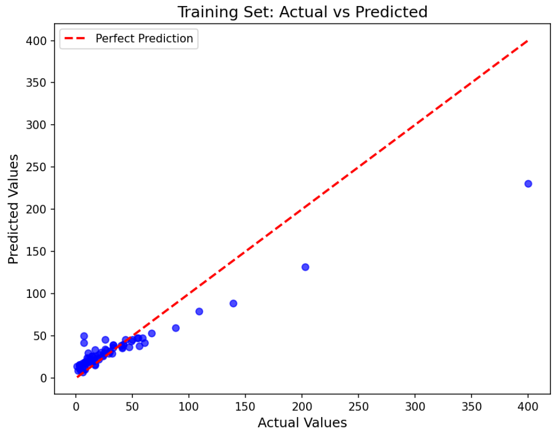

Figure 6: Result of optimized maxnet training

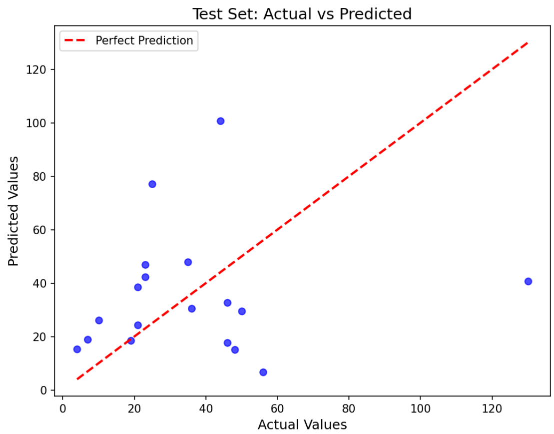

Figure 7: Result of optimized maxnet test

4 Impact on the stability of the ecosystem

To explore the changes in the predicted bird population under simulated scenarios of plant coverage increase and decrease by 10%, the following results were obtained:

- **Predicted bird population with a 10% increase in plant coverage:**

$$[38.53, 27.74, 40.4, 17.25, 27.61, 35.51, 18.62, 33.51, 48.39, 98.14]$$

- **Predicted bird population with a 10% decrease in plant coverage:**

$$[38.53, 27.74, 40.4, 17.25, 28.85, 32.42, 18.6, 34.53, 47.44, 97.94]$$

Predicted Bird Population Under Increased Plant Coverage

In the scenario where plant coverage increased, the predicted bird population showed overall growth, with significant increases in certain regions (e.g., point 10, where the population rose to 98.14). This suggests that increasing plant coverage, particularly through trees or grasslands, could provide more habitats and food resources, thereby promoting bird population growth.

Predicted Bird Population Under Decreased Plant Coverage

In the scenario where plant coverage decreased, the predicted bird population also experienced changes, but the overall variations were relatively small (e.g., point 10, where the population slightly declined to 97.94). This indicates that while the reduction in plant coverage might lead to habitat loss, its impact on bird populations in some areas is minimal, possibly due to compensatory effects from other factors such as climate conditions, biodiversity, or alternative resources.

- **Impact of increased plant coverage:** Increasing plant coverage (e.g., trees, grasslands) has a positive effect on bird populations by providing additional habitats and food sources.
- **Impact of decreased plant coverage:** A reduction in plant coverage has a relatively smaller impact on bird populations, which may be mitigated by compensatory factors such as climate conditions, the presence of roads or wetlands, or other ecological variables.

These findings highlight the importance of maintaining or enhancing vegetation coverage to support bird populations while also considering the role of other environmental factors in mitigating adverse effects.

Theorem 4.1. *No free lunch.*

In mathematical folklore, the "no free lunch" (NFL) theorem (sometimes pluralized) of David Wolpert and William Macready appears in the 1997 "No Free Lunch Theorems for Optimization" [**?**]. Wolpert had previously derived no free lunch theorems for machine learning (statistical inference) [2].

Lemma 4.2 (Moore's Law)**.** *Processor speeds, or overall processing power for computers will double every two years.*

Proof. Moore's law is a 1965 observation made by Intel co-founder Gordon E. Moore that the number of transistors placed in an integrated circuit (IC) or chip doubles approximately every two years.[See Fig.**??**] Because Moore's observation has been frequently cited and used for research and development by multiple organizations, and it has been proven repeatedly, it is known as Moore's law. □

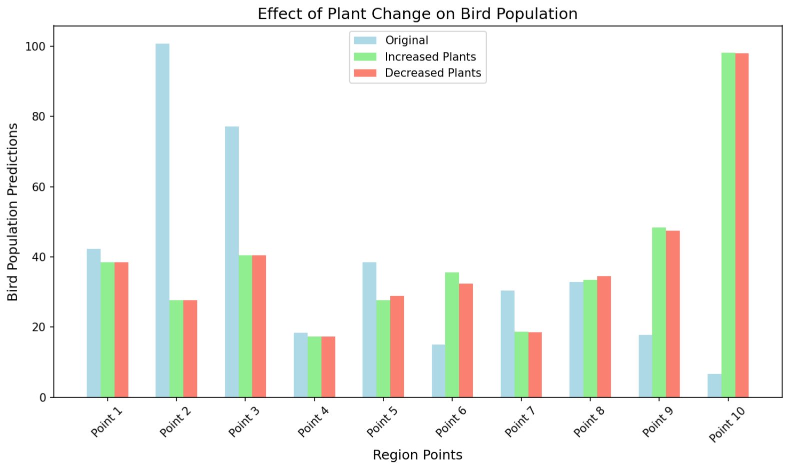

Figure 8: Result

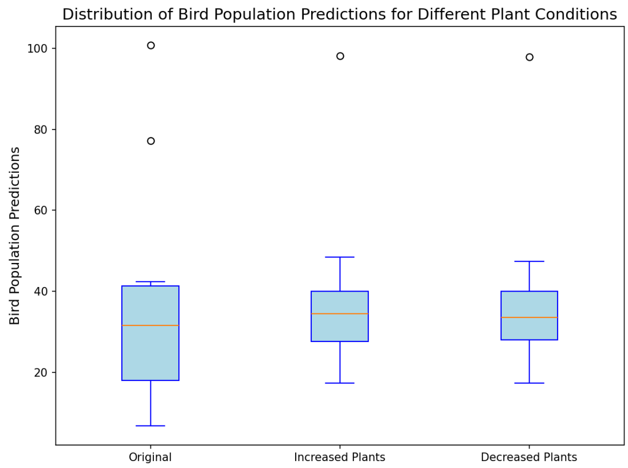

Figure 9: Result

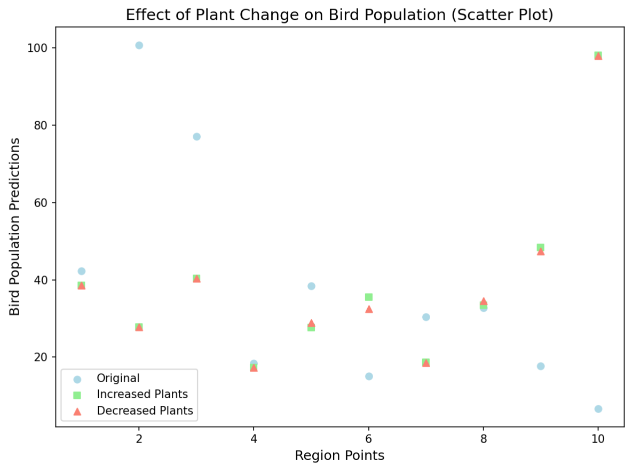

Figure 10: Result

It follows that

$$a^2 + b^2 = c^2 \tag{1}$$

Include your model design and justification for type model used or developed[3].
Calculating and simplifying the model:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y)$$

5 Conclusion

In this study, we investigated the intricate relationships between bird species richness, vegetation distribution, and ecosystem stability in the Altai Mountains. By analyzing the spatial and altitudinal distributions of birds and timber, as well as modeling the ecological interactions through the MaxEnt model and its optimized variants, we uncovered critical insights into the biodiversity and ecological dynamics of this unique region.

Our findings reveal that bird populations are closely linked to vegetation coverage, with plant diversity and abundance serving as essential factors in maintaining species richness. The altitudinal gradient further highlights the role of habitat heterogeneity in supporting avian diversity, particularly in zones with distinct vegetation types. By optimizing the MaxEnt model through ensemble methods, feature enhancements, and regularization, we achieved significant improvements in predictive accuracy, demonstrating the value of advanced modeling techniques in ecological research. Scenario analyses also indicated that increasing plant coverage positively impacts bird populations, while reductions in vegetation, though impactful, are mitigated by compensatory factors such as alternative habitats and climate conditions.

These results emphasize the importance of maintaining and enhancing vegetation coverage to support biodiversity and ensure ecosystem stability. Conservation strategies such as reforestation, habitat restoration, and sustainable land management are essential to preserving the ecological integrity of the Altai Mountains. Furthermore, the integration of advanced machine learning models into ecological studies provides a robust framework for predicting species distribution and assessing the impacts of environmental changes.

In conclusion, this study offers a comprehensive understanding of the complex interplay between bird populations, vegetation, and ecosystem stability in the Altai Mountains. By combining field data with advanced modeling approaches, we provide valuable insights for biodiversity conservation and ecological management, laying a foundation for future research and practical applications in similar ecological contexts.

References

[1] Frank R. Giordano, William P. Fox, and Steven B. Horton. *A First Course in Mathematical Modeling*. Brooks/Cole Cengage Learning, 5th edition, 2013.

[2] D. H. Wolpert. The lack of a priori distinctions between learning algorithms. *Neural Computation*, 8(7):1341–1390, 1996.

[3] Yushan Jiang, Ning Yang, Qingqi Yao, et al. Real-time moisture control in sintering process using offline online narx neural networks. *Neurocomputing*, 396:209–215, 2020.

Appendices

Appendix A: Programmes Codes

Here are simulation programmes we used in our model as follow.

Input Python source:

```
from sklearn.ensemble import RandomForestRegressor, VotingRegressor
from xgboost import XGBRegressor
from sklearn.linear_model import LinearRegression, Lasso, Ridge
from sklearn.metrics import mean_squared_error, r2_score
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.preprocessing import StandardScaler, PolynomialFeatures
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

file_path = r"D:train.xlsx"
data = pd.read_excel(file_path)

X = data[['elevation', 'bio1', 'bio12', 'bare land',
'road', 'grassland', 'tree', 'wood', 'shelter']]
y = data['abundance']

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

poly = PolynomialFeatures(degree=2, include_bias=False)
X_poly = poly.fit_transform(X_scaled)

X_train, X_test, y_train, y_test =
train_test_split(X_poly, y, test_size=0.2, random_state=42)

lin_reg = LinearRegression()
rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
xgb_model = XGBRegressor(n_estimators=100, random_state=42)
```

```
lasso_reg = Lasso(alpha=0.1)  # Lasso
ridge_reg = Ridge(alpha=1.0)  # Ridge

voting_reg = VotingRegressor(estimators=[
    ('linear', lin_reg),
    ('random_forest', rf_model),
    ('xgb', xgb_model)
])

param_grid = {
    'random_forest__n_estimators': [50, 100, 200],
    'random_forest__max_depth': [3, 5, 7],
    'xgb__learning_rate': [0.01, 0.1, 0.3],
    'xgb__n_estimators': [50, 100, 200],
    'xgb__max_depth': [3, 5, 7]
}

grid_search = GridSearchCV(estimator=voting_reg,
 param_grid=param_grid, cv=5)
grid_search.fit(X_train, y_train)

print("", grid_search.best_params_)

best_voting_reg = grid_search.best_estimator_
best_voting_reg.fit(X_train, y_train)

y_pred_train = best_voting_reg.predict(X_train)
y_pred_test = best_voting_reg.predict(X_test)

def evaluate_model(y_true, y_pred, dataset_type):
    rmse = np.sqrt(mean_squared_error(y_true, y_pred))
    r2 = r2_score(y_true, y_pred)
    print(f"{dataset_type} - RMSE: {rmse:.2f}, Rš: {r2:.2f}")
    return rmse, r2

print("")
evaluate_model(y_train, y_pred_train, "")

print("\n")
evaluate_model(y_test, y_pred_test, "")
```